(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 111511388 A
(43)申请公布日 2020.08.07

(21)申请号 201880074422.X
(22)申请日 2018.09.21
(30)优先权数据
62/561,669 2017.09.21 US
(85)PCT国际申请进入国家阶段日
2020.05.15
(86)PCT国际申请的申请数据
PCT/US2018/052247 2018.09.21
(87)PCT国际申请的公布数据
WO2019/060746 EN 2019.03.28
(71)申请人 博德研究所
地址 美国马萨诸塞州
申请人 麻省理工学院 哈佛学院院长等
(72)发明人 F·张 J·戈滕贝格
D·B·T·科克斯 O·阿布达耶
S·坎南
(74)专利代理机构 北京市金杜律师事务所 11256
代理人 陈文平 袁元
(51)Int.Cl.
A61K 38/43(2006.01)
C12N 15/09(2006.01)
C12N 15/10(2006.01)
C12N 9/16(2006.01)

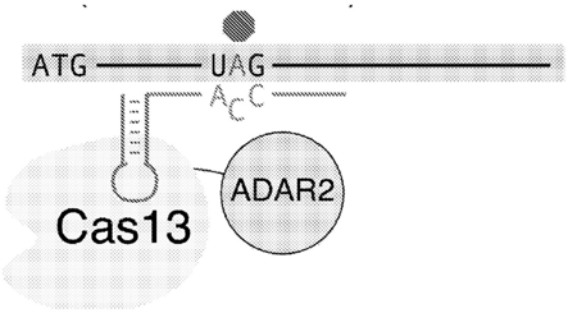
权利要求书2页 说明书283页
序列表（电子公布） 附图90页

(54)发明名称

用于靶向核酸编辑的系统、方法和组合物

(57)摘要

本公开提供了用于靶向和编辑核酸的系统、方法和组合物。特别地，本发明提供了非天然存在的或工程化的RNA靶向系统，所述系统包含靶向RNA的Cas13蛋白、至少一种指导分子和至少一种腺苷脱氨酶蛋白或其催化结构域。



CN 111511388 A

1.一种工程化的非天然存在的系统，所述系统包含无催化活性的Cas13效应蛋白(dCas13)或编码所述无催化活性的Cas13效应蛋白的核苷酸序列。

2.如权利要求1所述的系统，其中所述dCas13蛋白在C端、N端或这两端处被截短。

3.如权利要求2所述的系统，其中所述dCas13在所述C端上被截短至少20个、至少40个、至少60个、至少80个、至少100个、至少120个、至少140个、至少160个、至少180个、至少200个、至少220个、至少240个、至少260个或至少300个氨基酸。

4.如权利要求2所述的系统，其中所述dCas13在所述N端上被截短至少20个、至少40个、至少60个、至少80个、至少100个、至少120个、至少140个、至少160个、至少180个、至少200个、至少220个、至少240个、至少260个或至少300个氨基酸。

5.如权利要求2所述的系统，其中所述Cas13效应蛋白的所述截短形式已在C端Δ984-1090、C端Δ1026-1090、C端Δ1053-1090、C端Δ934-1090、C端Δ884-1090、C端Δ834-1090、C端Δ784-1090或C端Δ734-1090处被截短，其中所述截短的氨基酸位置对应于普雷沃氏菌属种P5-125 Cas13b蛋白的氨基酸位置。

6.如权利要求2所述的系统，其中所述Cas13效应蛋白的所述截短形式已在C端Δ795-1095处被截短，其中所述截短的氨基酸位置对应于鸭疫里默氏杆菌Cas13b蛋白的氨基酸位置。

7.如权利要求2所述的系统，其中所述Cas13效应蛋白的所述截短形式已在C端Δ875-1175、C端Δ895-1175、C端Δ915-1175、C端Δ935-1175、C端Δ955-1175、C端Δ975-1175、C端Δ995-1175、C端Δ1015-1175、C端Δ1035-1175、C端Δ1055-1175、C端Δ1075-1175、C端Δ1095-1175、C-端Δ1115-1175、C-端Δ1135-1175、C-端Δ1155-1175处被截短，其中氨基酸位置对应于咽喉卟啉单胞菌Cas13b蛋白的氨基酸位置。

8.如权利要求2所述的系统，其中所述Cas13效应蛋白的所述截短形式已在N端Δ1-125、N端Δ1-88或N端Δ1-72处被截短，其中所述截短的氨基酸位置对应于普雷沃氏菌属种P5-125 Cas13b蛋白的氨基酸位置。

9.如权利要求1所述的系统，其中所述dCas13在所述Cas13效应蛋白的HEPN结构域处包含Cas13效应蛋白的截短形式。

10.如权利要求1所述的系统，其中所述Cas13效应蛋白是Cas13a。

11.如权利要求1所述的系统，其中所述Cas13效应蛋白是Cas13b。

12.如权利要求1所述的系统，其中所述Cas13效应蛋白是Cas13c或Cas13d。

13.如权利要求1所述的系统，所述系统还包含功能组分，或者其中所述核苷酸序列还编码所述功能组分。

14.如权利要求13所述的系统，其中所述功能组分是碱基编辑组分。

15.如权利要求14所述的系统，其中所述碱基编辑组分包括腺苷脱氨酶、胞苷脱氨酶或它们的催化结构域。

16.如权利要求15所述的系统，其中所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域融合至所述dCas13。

17.如权利要求16所述的系统，其中所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域通过接头融合至所述dCas13。

18.如权利要求15所述的系统，其中所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催

化结构域插入到所述dCas13的内环中。

19.如权利要求15所述的系统，其中所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域连接至衔接蛋白。

20.如权利要求19所述的系统，其中所述衔接蛋白选自MS2、PP7、Qβ、F2、GA、fr、JP501、M12、R17、BZ13、JP34、JP500、KU1、M11、MX1、TW18、VK、SP、FI、ID2、NL95、TW19、AP205、ΦCb5、ΦCb8r、ΦCb12r、ΦCb23r、7s和PRR1。

21.如权利要求15所述的系统，其中所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域是人类、头足类动物或果蝇蛋白。

22.如权利要求1所述的系统，其中所述dCas13是裂解Cas13效应蛋白。

23.如权利要求22所述的系统，其中所述裂解Cas13效应蛋白是第一裂解Cas13效应蛋白并且能够融合至第二裂解Cas13效应蛋白以形成催化活性Cas13效应蛋白。

24.如权利要求23所述的系统，其中所述第一Cas13效应蛋白和所述第二Cas13效应蛋白的融合是可诱导的。

25.如权利要求13所述的系统，其中所述功能组分是转录因子或其活性结构域。

26.如权利要求25所述的系统，其中所述dCas13与所述转录因子或其活性结构域融合。

27.如权利要求1所述的系统，所述系统还包含指导序列。

28.一种载体系统，所述载体系统包含一种或多种编码如前述权利要求中任一项所述的dCas13的载体。

29.一种包含如权利要求1至27中任一项所述的系统的体外或离体宿主细胞或其子代或者细胞系或其子代。

30.一种用于对靶序列进行可编程和靶向式的碱基编辑的方法，所述方法包括将如权利要求1-21中任一项所述的系统递送至细胞。

31.如权利要求30所述的方法，所述方法还包括确定所述目标靶序列以及选择最有效地使存在于所述靶序列中的腺嘌呤或胞苷脱氨基的腺苷脱氨酶、胞苷脱氨酶或它们的催化结构域。

32.如权利要求31所述的方法，其中所述脱氨补救了由含有病原性T(U)→C、A→G、G→A或C→T点突变的转录物引起的疾病。

33.一种用于修饰靶基因座处的核酸的方法，所述方法包括递送：(1)如权利要求1所述的系统，其中所述dCas13是第一裂解Cas13效应蛋白；和(2)第二裂解Cas13效应蛋白，其中所述第一裂解Cas13效应蛋白和所述第二裂解Cas13效应蛋白能够形成催化活性Cas13效应蛋白。

34.一种用于调控细胞中的靶基因的方法，所述方法包括递送如权利要求1所述的系统，其中所述dCas13融合至转录因子或其活性结构域。

用于靶向核酸编辑的系统、方法和组合物

[0001] 相关申请的交叉引用

[0002] 本申请要求2017年9月21日提交的美国临时申请号62/561,669的权益。以上确认的申请的全部内容特此以引用方式完全并入本文。

[0003] 关于联邦资助研究的声明

[0004] 本发明是根据由美国国立卫生研究院(National Institutes of Health)授予的授权号MH110049在政府支持下完成的。政府享有本发明的某些权利。

[0005] 电子序列表引用

[0006] 电子序列表(“BROD-2305WP_ST25.txt”;大小为1,409,943字节,创建日期为2018年9月14日)的内容以引用方式整体并入本文。

技术领域

[0007] 本发明总体上涉及用于靶向和编辑核酸,特别地用于目标靶基因座处腺嘌呤的可编程脱氨的系统、方法和组合物。

背景技术

[0008] 在基因组测序技术和分析方法中的最新进展显著加速了对与范围广泛的生物功能和疾病相关联的遗传因子进行编目和映射的能力。精确的基因组靶向技术是通过允许个体遗传元件的选择性干扰而使得因果性遗传变异的系统性逆向工程化成为可能、以及推进合成生物学、生物技术应用、和医学应用所需要的。虽然基因组编辑技术,如设计师锌指、转录激活因子样效应子(TALE)、或归巢大范围核酸酶(homing meganuclease)对于产生靶向的基因组干扰是可得的,但是仍然需要采用新颖的策略和分子机制且是负担得起的、易于建立的、可扩展的、并且便于靶向真核基因组内的多个位置的新的基因组工程化技术。这将为基因组工程化和生物技术的新应用提供主要资源。

[0009] 先前已经报道了胞嘧啶的可编程脱氨,其可用于校正A→G和T→C点突变。例如,Komor等人,Nature(2016)533:420-424报道了在非靶向DNA链中通过APOBEC1胞嘧啶脱氨酶进行的胞嘧啶脱氨,Cas9-指导RNA复合物与靶向DNA链的结合导致该非靶向DNA移位,从而使得胞嘧啶转化为尿嘧啶。另参见Kim等人,Nature Biotechnology(2017)35:371-376;Shimatani等人,Nature Biotechnology(2017)doi:10.1038/nbt.3833;Zong等人,Nature Biotechnology(2017)doi:10.1038/nbt.3811;Yang Nature Communication(2016)doi:10.1038/ncomms13330。

发明内容

[0010] 在一方面,本公开包括一种工程化的非天然存在的系统,所述系统包含无催化活性的Cas13效应蛋白(dCas13)或编码所述无催化活性的Cas13效应蛋白的核苷酸序列。

[0011] 在一些实施方案中,所述dCas13蛋白在C端、N端或这两端处被截短。在一些实施方案中,所述dCas13在所述C端上被截短至少20个、至少40个、至少60个、至少80个、至少100

个、至少120个、至少140个、至少160个、至少180个、至少200个、至少220个、至少240个、至少260个或至少300个氨基酸。在一些实施方案中，所述dCas13在所述N端上被截短至少20个、至少40个、至少60个、至少80个、至少100个、至少120个、至少140个、至少160个、至少180个、至少200个、至少220个、至少240个、至少260个或至少300个氨基酸。在一些实施方案中，所述Cas13效应蛋白的所述截短形式已在C端Δ984-1090、C端Δ1026-1090、C端Δ1053-1090、C端Δ934-1090、C端Δ884-1090、C端Δ834-1090、C端Δ784-1090或C端Δ734-1090处被截短，其中所述截短的氨基酸位置对应于普雷沃氏菌属种P5-125 Cas13b蛋白的氨基酸位置。在一些实施方案中，所述Cas13效应蛋白的所述截短形式已在C端Δ795-1095处被截短，其中所述截短的氨基酸位置对应于鸭疫里默氏杆菌Cas13b蛋白的氨基酸位置。在一些实施方案中，所述Cas13效应蛋白的所述截短形式已在C端Δ875-1175、C端Δ895-1175、C端Δ915-1175、C端Δ935-1175、C端Δ955-1175、C端Δ975-1175、C端Δ995-1175、C端Δ1015-1175、C端Δ1035-1175、C端Δ1055-1175、C端Δ1075-1175、C端Δ1095-1175、C-端Δ1115-1175、C-端Δ1135-1175、C-端Δ1155-1175处被截短，其中氨基酸位置对应于咽喉卟啉单胞菌Cas13b蛋白的氨基酸位置。在一些实施方案中，所述Cas13效应蛋白的所述截短形式已在N端Δ1-125、N端Δ1-88或N端Δ1-72处被截短，其中所述截短的氨基酸位置对应于普雷沃氏菌属种P5-125 Cas13b蛋白的氨基酸位置。在一些实施方案中，所述dCas13在所述Cas13效应蛋白的HEPN结构域处包含Cas13效应蛋白的截短形式。

[0012] 在一些实施方案中，所述Cas13效应蛋白是Cas13a。在一些实施方案中，所述Cas13效应蛋白是Cas13b。在一些实施方案中，所述Cas13效应蛋白是Cas13c或Cas13d。

[0013] 在一些实施方案中，所述系统还包含功能组分，或者其中所述核苷酸序列还编码所述功能组分。在一些实施方案中，所述功能组分是碱基编辑组分。在一些实施方案中，所述碱基编辑组分包括腺苷脱氨酶、胞苷脱氨酶或它们的催化结构域。在一些实施方案中，所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域融合至所述dCas13。在一些实施方案中，所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域通过接头融合至所述dCas13。在一些实施方案中，所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域插入到所述dCas13的内环中。在一些实施方案中，所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域连接至衔接蛋白。在一些实施方案中，所述衔接蛋白选自MS2、PP7、Qβ、F2、GA、fr、JP501、M12、R17、BZ13、JP34、JP500、KU1、M11、MX1、TW18、VK、SP、FI、ID2、NL95、TW19、AP205、ΦCb5、ΦCb8r、ΦCb12r、ΦCb23r、7s和PRR1。在一些实施方案中，所述腺苷脱氨酶、所述胞苷脱氨酶或所述它们的催化结构域是人类、头足类动物或果蝇蛋白。

[0014] 在一些实施方案中，所述dCas13是裂解Cas13效应蛋白。在一些实施方案中，所述裂解Cas13效应蛋白是第一裂解Cas13效应蛋白，并且能够融合至第二裂解Cas13效应蛋白以形成催化活性Cas13效应蛋白。在一些实施方案中，所述第一Cas13效应蛋白和所述第二Cas13效应蛋白的融合是可诱导的。

[0015] 在一些实施方案中，所述功能组分是转录因子或其活性结构域。在一些实施方案中，所述dCas13与所述转录因子或其活性结构域融合。在一些实施方案中，所述系统还包含指导序列。

[0016] 在另一方面，本公开包括一种载体系统，所述载体系统包含一种或多种编码本文所公开的dCas13的载体。

[0017] 在另一方面，本公开包括包含本文所公开的系统的体外或离体宿主细胞或其子代或者细胞系或其子代。

[0018] 在另一方面，本公开包括一种用于对靶序列进行可编程和靶向式的碱基编辑的方法，所述方法包括将本文所公开的系统递送至细胞。在一些实施方案中，所述方法还包括确定所述目标靶序列以及选择最有效地使存在于所述靶序列中的腺嘌呤或胞苷脱氨基的腺苷脱氨酶、胞苷脱氨酶或它们的催化结构域。在一些实施方案中，所述脱氨补救了由含有病原性T(U)→C、A→G、G→A或C→T点突变的转录物引起的疾病。

[0019] 在另一方面，本公开包括一种用于修饰靶基因座处的核酸的方法，所述方法包括递送(1)如权利要求1所述的系统，其中所述dCas13是第一裂解Cas13效应蛋白；和(2)第二裂解Cas13效应蛋白，其中所述第一裂解Cas13效应蛋白和所述第二裂解Cas13效应蛋白能够形成催化活性Cas13效应蛋白。

[0020] 在另一方面，本公开包括一种用于调控细胞中的靶基因的方法，所述方法包括递送如本文所公开的系统，其中所述dCas13融合至转录因子或其活性结构域。

[0021] 结合以下对所示示例性实施方案的详细说明，示例性实施方案的这些和其他方面、目的、特征和优点对于本领域普通技术人员而言将变得显而易见。

附图说明

[0022] 通过参考阐述了其中可能利用了本发明原理的说明性实施方案的以下详细说明及其附图，将获得对本发明的特征和优点的理解：

[0023] 图1示出了用于目标靶RNA序列处腺嘌呤的靶向脱氨的本发明的示例性实施方案，在本文中以Cas13b蛋白为例。

[0024] 图2示出了RNA编辑作为在转录水平上(如当使用Cas13b时)治疗人类疾病的治疗策略的发展。通过靶向荧光素酶转录物中的工程化的早期终止密码子的Cas13-ADAR2融合进行RNA碱基编辑的示意图。

[0025] 图3用以恢复荧光素酶表达的指导物位置和长度优化。

[0026] 图4腺嘌呤脱氨酶蛋白的示例性序列(SEQ ID NO:650-656)。

[0027] 图5示例性实施方案中使用的指导物(SEQ ID NO:657-659和762)。

[0028] 图6:编辑效率与编辑碱基有关，所述编辑碱基离DR更远且具有长RNA双链体，这可以通过延长指导物长度来实现。

[0029] 图7编辑位点离DR/蛋白质结合区越远，编辑效率越高。

[0030] 图8编辑位点离DR的距离。

[0031] 图9A和图9B:将ADAR1或ADAR2融合至Cas13b12(双R HEPN突变体)的N端或C端。指导物与荧光素酶中的终止密码子完美匹配。信号似乎与编辑碱基与指导物的5'末端之间的距离有关，较短的距离可以提供更好的编辑。

[0032] 图10:用于Cas13a/Cas13b干扰的Cluc/Gluc平铺。

[0033] 图11A-图11C:通过NGS进行的ADAR编辑定量(荧光素酶报告基因)。

[0034] 图12:通过NGS进行的ADAR编辑定量(KRAS和PPIB)。

[0035] 图13A-图13C:来自RNA Seq的Cas13a/b+shRNA特异性。

[0036] 图14:用以减少脱靶的错配特异性(A:A或A:G)(SEQ ID No:661-668)。

[0037] 图15:用于中靶活性的错配。

[0038] 图16:ADAR基序优选性。

[0039] 图17:用以增强RNA编辑效率的大气泡。

[0040] 图18:编辑转录物中的多个A(SEQ ID No:669-672)。

[0041] 图19:用于RNA编辑的指导物长度滴定。

[0042] 图20A-图20F:哺乳动物密码子优化的Cas13b直系同源物介导高效RNA敲低。(图20A)代表性Cas13a、Cas13b和Cas13c基因座及相关crRNA的示意图。(图20B)用以测量HEK293FT细胞中的Cas13a切割活性的荧光素酶测定的示意图。(图20C)使用两种不同的靶向Cluc的指导物与19种Cas13a、15种Cas13b和5种Cas13c直系同源物进行的RNA敲低的效率。针对在非靶向指导物控制条件下的表达将荧光素酶表达归一化。(图20D)用两种不同的靶向Cluc的指导RNA,用三种不同的NLS和NES标签测定在C部分进行的前7个直系同源物的活性。(图20E)比较Cas13b12和Cas13a2(LwCas13a)针对Gluc和Cluc的敲低活性。将指导物沿转录物平铺,并且对Cas13b12与Cas13a2之间的指导物进行位置匹配。(图20F)Cas13a2、Cas13b6、Cas13b11和Cas13b12针对内源性KRAS转录物的指导物敲低,与相应的shRNA进行比较。

[0043] 图21A-图21G:Cas13酶介导哺乳动物细胞中的特异性RNA敲低。(图21A)用于Cas13a/b错配特异性测试的半简并靶序列的示意图(SEQ ID NO:673-294)。(图21B)Cas13a/b的单错配敲低数据的热图。对于每种酶针对非靶向(NT)指导物将敲低归一化。(图21C)Cas13a的双错配敲低数据。每个错配的位置在X轴和Y轴上指示。敲低数据是给定位置集合的所有双错配的总和。对于每种酶针对(NT)指导物将数据归一化。(图21D)Cas13b的双错配敲低数据。有关说明,请参见C。(图21E)RNA-seq数据,比较了对于位置匹配的指导物Cas13 a/b和shRNA的全转录组特异性。Y轴表示靶向条件的读取计数,X轴表示非靶向条件的计数。(图21F)从Cas13 a/b和shRNA的RNA-seq数据计算的RNA表达。(图21G)来自RNA-seq数据的Cas13 a/b和shRNA的显著脱靶。使用FDR<0.05计算显著脱靶。

[0044] 图22A-图22F:无催化活性的Cas13b-ADAR融合物使得可以在哺乳动物细胞中进行靶向RNA编辑。(图22A)用Cas13b-ADAR融合蛋白进行RNA编辑以去除海萤(Cypridina)荧光素酶转录物上的终止密码子的示意图。(图22B)针对多个指导物位置,与野生型ADAR2融合的Cas13b和与超活性ADAR2 E488Q突变体融合的Cas13b之间的RNA编辑比较。针对海肾(Gaussia)荧光素酶对照值将荧光素酶表达归一化。(图22C)被设计来靶向编辑位点周围的不同位置的30、50、70和84nt指导物之间的RNA编辑比较。(图22D)显示用于测序定量的指导物相对于海萤荧光素酶转录物上的终止密码子的位置和长度的示意图(SEQ ID NO:695)。(图22E)如通过测序定量的,每种指导物设计在海萤荧光素酶转录物上的相应腺嘌呤碱基处的中靶和脱靶编辑效率。(图22F)具有与靶向腺嘌呤相对的不同碱基的指导物的荧光素酶读出。

[0045] 图23A:Cas13b-ADAR融合物的内源性RNA编辑。对内源性KRAS和PPIB基因座的内源性Cas13b12-ADAR编辑的下一代测序。靶向每个转录物的两个不同区域,并且定量靶向腺嘌呤附近的所有腺嘌呤处的A->G编辑。

[0046] 图24:确定最佳指导物位置的策略。

[0047] 图25A-图25C:(图25A)Cas13b-huADAR2促进突变荧光素酶转录物的修复。(图25B)

Cas13b-huADAR1促进突变荧光素酶转录物的修复。(图25C)人类ADAR1和人类ADAR2的比较。

[0048] 图26:E488Q与wt dADAR2编辑的比较。E488Q是dADAR2的超活性突变体。

[0049] 图27A-图27B:由Cas13b-huADAR2-E488Q靶向的转录物含有预期的A-G编辑。(图27A)热图。(图27B)模板中的位置。热图中仅示出A位点的至G编辑率。

[0050] 图28A-图28B:指导物的内源性平铺。(图28A)KRAS:热图(顶部)。模板中的位置(底部)。热图中仅示出A位点的至G编辑率。(图28B)PPIB:热图(顶部)。模板中的位置(底部)。热图中仅示出A位点的至G编辑率。

[0051] 图29:非靶向编辑。

[0052] 图30:接头优化。

[0053] 图31:Cas13b ADAR可以用于校正患者的在表达的cDNA中的病原性A>G突变。

[0054] 图32:Cas13b-ADAR对5'G基序有轻微限制。

[0055] 图33:筛选退化的PFS位置以影响编辑效率。所有PFS(4-N)种类的编辑度都高于非靶向种类(SEQ ID NO:696-699)。

[0056] 图34:减少靶转录物中的脱靶编辑。

[0057] 图35:减少靶转录物中的脱靶编辑。

[0058] 图36A-图36B:Cas13b-ADAR转录组特异性。中靶编辑是71%。(图36A)靶向指导物;482个显著位点。(图36B)非靶向指导物;949个显著位点。注意染色体0是Gluc并且染色体1是Cluc;之后按次序是人类染色体。

[0059] 图37A-图37B:Cas13b-ADAR转录组特异性。(图37A)靶向指导物。(图37B)非靶向指导物。

[0060] 图38:与竞争性ADAR编辑策略相比,Cas13b具有最高的效率。

[0061] 图39A-图39D:竞争性RNA编辑系统。(图39A-图39B)BoxB;中靶编辑是63%;(图39A)靶向指导物-2020个显著位点;(图39B)非靶向指导物-1805个显著位点。(图39C-图39D)Stafforst;中靶编辑是36%;(图39C)靶向指导物-176个显著位点;(图39D)非靶向指导物-186个显著位点。

[0062] 图40:ADAR的剂量滴定。crRNA的量是恒定的。

[0063] 图41A-图41D:剂量反应对特异性的影响。(图41A-图41B)150ng Cas13-ADAR;中靶编辑是83%;(图41A)靶向指导物-1231个显著位点;(图41B)非靶向指导物-520个显著位点。(图41C-41D)10ng Cas13-ADAR;中靶编辑是80%;(图41C)靶向指导物-347个显著位点;(图41D)非靶向指导物-223个显著位点。

[0064] 图42A-图42B:ADAR1似乎比ADAR2更具特异性。中靶编辑是29%。(图42A)靶向指导物;11个显著位点。(图42B)非靶向指导物;6个显著位点。注意染色体0是Gluc并且染色体1是Cluc;之后按次序是人类染色体。

[0065] 图43A-图43D:ADAR特异性突变体具有增强的特异性。(图43A)靶向指导物。(图43B)非靶向指导物。(图43C)靶向指导物与靶向指导物之比。(图43D)靶向指导物和非靶向指导物。

[0066] 图44:沿每个残基与RNA靶标的接触点绘制的ADAR突变体荧光素酶结果。

[0067] 图45:ADAR特异性突变体具有增强的特异性。紫色点是选择用于全转录组脱靶NGS分析的突变体。红色点是起点(即E488Q突变体)。需注意,所有其他突变体也都具有E488Q突

变。

[0068] 图46A-图46B:根据NGS,ADAR突变体更具特异性。(图46A)中靶。(图46B)脱靶。

[0069] 图47A-图47B:有关ADAR特异性突变体的荧光素酶数据与NGS匹配。(图47A)选择用于NGS的靶向指导物。(图47B)选择用于NGS的非靶向指导物。荧光素酶数据与图46中的NGS数据匹配。在非靶向指导物下具有较少活性的直系同源物在整个转录组中具有较少的脱靶,并且它们的中靶编辑效率可以通过靶向指导物荧光素酶条件来预测。

[0070] 图48:Cas13b 12的C端截短在ADAR编辑中仍然高度活跃。

[0071] 图49:A-49G用于RNA敲低的高活性Cas13b直系同源物的表征结果图49A)常规Cas13基因座和相应crRNA结构的示意图。图49B)使用两种不同指导物评估19种Cas13a、15种Cas13b和7种Cas13c直系同源物的荧光素酶敲低。将使用两种指导物具有有效敲低的直系同源物标记上其宿主生物体名称。图49C)通过针对Gluc平铺指导物并测量荧光素酶表达来比较PspCas13b和LwaCas13a敲低活性。图49D)通过针对Cluc平铺指导物并测量荧光素酶表达来比较PspCas13b和LwaCas13a敲低活性。图49E)针对LwaCas13a(红色)和shRNA(黑色)在非靶向对照(x轴)对比Gluc靶向条件(y轴)的RNA-seq文库中检测到的所有基因的以log2(每百万的转录物(TPM))值计的表达水平。显示的是三个生物学重复的平均值。标记了Gluc转录物数据点。图49F)针对PspCas13b(蓝色)和shRNA(黑色)在非靶向对照(x轴)对比Gluc靶向条件(y轴)的RNA-seq文库中检测到的所有基因的以log2(每百万的转录物(TPM))值计的表达水平。显示的是三个生物学重复的平均值。标记了Gluc转录物数据点。图49G)从图49E和图49F中的转录组范围分析获得的LwaCas13a、PspCas13b和shRNA的Gluc敲低的显著脱靶数量。

[0072] 图50A-图50E:用于RNA编辑的工程化dCas13b-ADAR融合物图50A)由dCas13b-ADAR融合蛋白进行的RNA编辑的示意图。图50B)海萤荧光素酶W85X靶标和靶向指导物设计的示意图(SEQ ID NO:700和701)。图50C)使用长度为30、50、70或84nt的平铺指导物的Cas13b-dADAR1(左)和Cas13b-ADAR2-cd(右)的荧光素酶活性恢复的定量。图50D)用于靶向海萤荧光素酶W85X的靶标位点的示意图。图50E)靶向海萤荧光素酶W85X的50nt指导物的A->I编辑的测序定量(SEQ ID NO:702)。

[0073] 图51A-图51D:图51A测量通过REPAIRv1进行RNA编辑的序列灵活性,筛选确定通过REPAIRv1进行RNA编辑的原间隔区侧接位点(PFS)优选性的示意图。(SEQ ID NO:703)。图51B)在两个不同的编辑位点,所有4-N PFS组合的RNA编辑效率分布。图51C)在所有可能的3碱基基序上,在Cluc W85处REPAIRv1的编辑百分比的定量结果(SEQ ID NO:704)。图51D)针对所有可能的3碱基基序在Clue W85处进行RNA编辑的5'和3'碱基优选性的热图。

[0074] 图52A-图52F:使用REPAIRv1校正疾病相关突变图52A)用于靶向AVPR2878G>A的靶标和指导物设计的示意图(SEQ ID NO:705-708)。图52B)使用具有三种不同指导物设计的REPAIRv1以不同的百分比校正了AVPR2中的878G>A突变。图52C)用于靶向FANCC 1517G>A的靶标和指导物设计的示意图。(SEQ ID No:709-712)。图52D)使用具有三种不同指导物设计的REPAIRv1以不同的百分比校正了FANCC中的1517G>A突变。图52E)对使用REPAIRv1对34种不同的疾病相关G>A突变的编辑百分比的定量结果。图52F)对如通过ClinVar数据库所注释可以被校正的所有可能的G>A突变的分析结果。相对于如在Gluc转录物上定量的每基序而言的REPAIRv1的编辑效率,显示了ClinVar中所有G>A突变的编辑基序的分布。

[0075] 图53A-图53D:表征REPAIRv1的特异性图53A)KRAS靶位点和指导物设计的示意图。(SEQ ID NO:713-720)。图53B)平铺KRAS靶向指导物的编辑百分比的定量结果。显示了在中靶和相邻腺苷酸位点处的编辑百分比。对于每种指导物,双链体RNA的区域由红色矩形表示。图53C)使用Cluc靶向指导物,REPAIRv1的全转录组显著RNA编辑位点。中靶位点Cluc位点(254A>G)以橙色突出显示。图53D)使用非靶向指导物,REPAIRv1的全转录组显著RNA编辑位点。

[0076] 图54A-图54F:合理诱变ADAR2以提高REPAIRv1的特异性图54A)沿着关键ADAR2脱氨酶残基与dsRNA靶标之间的接触示意图绘制的各种dCas13-ADAR2突变体的荧光素酶信号恢复的定量结果及其特异性得分。特异性得分被定义为在靶向指导物与非靶向指导物条件之间的荧光素酶信号的比率。图54B)各种dCas13-ADAR2突变体的荧光素酶信号恢复的定量结果及其特异性评分。图54C)通过mRNA的全转录组测序得出的每个dCas13-ADAR2突变体的中靶编辑得分以及显著脱靶数量的测量结果。图54D)使用靶向Cluc中的早期终止位点的指导物,REPAIRv1和REPAIRv2的全转录组显著RNA编辑位点。中靶Cluc位点(254A>G)以橙色突出显示。图54E)围绕中靶Cluc编辑位点(254A>G)的RNA测序读取突出显示了REPAIRv1与REPAIRv2之间的脱靶编辑的差异。所有A>G编辑均以红色突出显示,而测序误差均以蓝色突出显示(SEQ ID NO:721)。图54F)使用靶向在内源性KRAS和PPIB转录物中的框外UAG位点的指导物通过REPAIRv1和REPAIRv2进行的RNA编辑。在每个条件行的右侧,将中靶编辑分数显示为横向条形图。由指导RNA形成的双链体区域由红色轮廓框显示。

[0077] 图55A-图55C:用于Cas13b直系同源物的体内效率和PFS确定的细菌筛选。图55A)用于确定Cas13b直系同源物的PFS的细菌测定的示意图。将具有β-内酰胺酶靶向间隔区的Cas13b直系同源物与β-内酰胺酶表达质粒共转化并进行双选择。(SEQ ID NO:722)。图55B)如通过菌落形成单位(cfu)所测量的靶向β-内酰胺酶的Cas13b直系同源物的干扰活性的定量结果。图55C)如通过耗尽序列从细菌测定中所确定的Cas13b直系同源物的PFS logo。

[0078] 图56A-图56E:Cas13b敲低的优化和错配特异性的进一步表征。图56A)使用与多种核定位标签和核输出标签融合的前2个Cas13a和前4个Cas13b直系同源物来测量使用两种不同指导物的Gluc敲低。图56B)测量用四种不同指导物LwaCas13a、RanCas13b、PguCas13b和PspCas13b的KRAS敲低,并将其与四个位置匹配的shRNA对照物进行比较。图56C)用于评估LwaCas13a和PspCas13b敲低的特异性的单错配和双错配质粒文库的示意图。每个可能的单错配和双错配都存在于靶序列中以及直接侧接靶位点5'端和3'端的3个位置中。(SEQ ID NO:723-735)。图56D)将具有指示的单错配的转录物的耗尽水平绘制为LwaCas13a和PspCas13b条件的热图。图56E)将具有指示的双错配的转录物的耗尽水平绘制为LwaCas13a和PspCas13b条件的热图(SEQ ID No:736)。

[0079] 图57A-图57F:dCas13-ADAR2 RNA编辑的设计参数的表征结果图57A)野生型Cas13b和无催化活性的H133A/H1058A Cas13b(dCas13b)的Gluc靶向敲低效率。图57B)使用平铺Cluc靶向指导物测试的,融合至野生型ADAR2催化结构域或高活性E488Q突变型ADAR2催化催化结构域的dCas13b的荧光素酶活性恢复的定量结果。图57C)靶向海萤荧光素酶W85X的30nt指导物的指导物设计和A->I编辑的测序定量结果(SEQ ID NO:737-745)。图57D)靶向PPIB的50nt指导物的指导物设计和A->I编辑的测序定量结果(SEQ ID NO:746-753)。图57E)接头选择对REPAIRv1的荧光素酶活性恢复的影响。图57F)靶向腺苷酸对面的

碱基身份对REPAIRv1的荧光素酶活性恢复的影响。(SEQ ID NO:754-755)。

[0080] 图58:G>A突变的ClinVar基序分布。对于所有G>A突变在ClinVar数据库中观察到的每个可能的三联体基序的数量。

[0081] 图59:dCas13b的截短仍然具有功能性RNA编辑。如通过荧光素酶信号的恢复所测量的,dCas13b的各种N端和C端截短允许RNA编辑。

[0082] 图60A-图60F:其他可编程ADAR系统与dCas13-ADAR2编辑器的比较。图60A)两种可编程ADAR方案的示意图:基于BoxB的靶向和全长ADAR2靶向。在BoxB方案中(顶图),将ADAR2脱氨酶结构域(ADAR2$_{DD}$(E488Q))融合至称为兰布达N(λN)的小细菌病毒蛋白,所述小细菌病毒蛋白特异性地结合称为BoxB-λ的小RNA序列)。然后,含有两个BoxB-λ发夹的指导RNA可以指导ADAR2$_{DD}$(E488Q)-λN进行位点特异性编辑。在全长ADAR2方案中(底图),ADAR2的dsRNA结合结构域结合指导RNA中的发夹,从而允许进行可编程的ADAR2编辑。(SEQ ID NO:756-760)。图60B)使用靶向Cluc的指导物和非靶向指导物,BoxB-ADAR2$_{DD}$(E488Q)的全转录组显著RNA编辑位点。中靶Cluc位点(254A>G)以橙色突出显示。图60C)使用靶向Cluc的指导物和非靶向指导物,ADAR2的全转录组显著RNA编辑位点。中靶Cluc位点(254A>G)以橙色突出显示。图60D)使用靶向Cluc的指导物和非靶向指导物,REPAIRv1的全转录组显著RNA编辑位点。中靶Cluc位点(254A>G)以橙色突出显示。图60E)对于针对Cluc的靶向指导物,BoxB-ADAR2$_{DD}$(E488Q)、ADAR2、和REPAIRv1的中靶编辑率百分比的定量结果。图60F)可编程ADAR系统在不同靶向和非靶向条件之间的脱靶位点的重叠。

[0083] 图61A-图61C:dCas13b-ADAR2突变体的效率和特异性图61A)对于Cluc靶向指导物和非靶向指导物,dCas13b-ADAR2$_{DD}$(E488Q)突变体的荧光素酶活性恢复的定量结果。图61B)靶向指导物和非靶向指导物比率与如通过全转录组测序定量的RNA编辑脱靶数量之间的关系图61C)dCas13b-ADAR2$_{DD}$(E488Q)突变体的全转录组脱靶RNA编辑位点数量与中靶Cluc编辑效率的定量结果。

[0084] 图62A-图62B:dCas13b-ADAR2$_{DD}$(E488Q)突变体的全转录组RNA编辑特异性图62A)使用靶向Cluc的指导物,dCas13b-ADAR2$_{DD}$(E488Q)突变体的全转录组显著RNA编辑位点。中靶Cluc位点(254A>G)以橙色突出显示。图62B)使用非靶向指导物,dCas13b-ADAR2$_{DD}$(E488Q)突变体的全转录组显著RNA编辑位点。

[0085] 图63A-图63C:dCas13b-ADAR2$_{DD}$(E488Q)编辑的脱靶中的基序偏倚的表征结果。图63A)对于每个dCas13b-ADAR2$_{DD}$(E488Q)突变体,显示了转录组中所有A>G脱靶编辑中存在的基序。图63B)针对使用靶向指导物和非靶向指导物的REPAIRv1,显示了每基序身份的脱靶A>G编辑的分布。图63C)针对使用靶向指导物和非靶向指导物的REPAIRv2,显示了每基序身份的脱靶A>G编辑的分布。

[0086] 图64A-图63F:对REPAIRv1和REPAIRv2脱靶的进一步表征结果。图64A)REPAIRv1的每转录物脱靶数量的直方图。图64B)REPAIRv2的每转录物脱靶数量的直方图。图64C)REPAIRv1脱靶的变异效应预测。图64D)REPAIRv1脱靶的潜在致癌作用分布。图64E)REPAIRv2脱靶的变异效应预测。图64F)REPAIRv2脱靶的潜在致癌作用分布。

[0087] 图65A-图65C:REPAIRv1和REPAIRv2的RNA编辑效率和特异性。图65A)对于REPAIRv1和REPAIRv2,在靶向腺苷及相邻位点处使用KRAS靶向指导物1的情况下KRAS的编辑百分比的定量结果。图65B)对于REPAIRv1和REPAIRv2,在靶向腺苷及相邻位点处使用

KRAS靶向指导物3的情况下KRAS的编辑百分比的定量结果。图65C)对于REPAIRv1和REPAIRv2,在靶向腺苷及相邻位点处使用PPIB靶向指导物2的情况下PPIB的编辑百分比的定量结果。

[0088] 图66:使用A>G RNA编辑器,所有潜在密码子变化的展示。图66A)通过A>I编辑启用的所有潜在密码子转换的表。图66B)展示了通过A>I编辑启用的所有潜在密码子转换的密码子表。

[0089] 图67示出了Cas13b6上RNA编辑活性的测试结果。

[0090] 图68示出了Cas13b11上RNA编辑活性的测试结果。

[0091] 图69示出了Cas13b12上RNA编辑活性的测试结果。

[0092] 图70示出了dCas13b6(Δ795-1095)-REPAIR的示意图以及该构建体的编辑率。

[0093] 图71示出了通过递送dCas13b6(Δ795-1095)-REPAIR来调节钾离子通道Kcna1中的位点(已经是天然的ADAR2底物)处的编辑。

[0094] 图72示出了dCas13b6(Δ795-1095)-GS-HIVNES-GS-huADAR2dd的示意图。

[0095] 图73对pgCas13b C端截短进行REPAIR测定的结果。

[0096] 图74示出了可以与本文的系统和方法一起使用的ADAR直系同源物的比对树。

[0097] 本文中的附图仅用于说明目的,而不一定按比例绘制。

具体实施方式

[0098] 一般定义

[0099] 除非另有规定,否则本文所用的技术和科学术语具有与本公开所属领域的普通技术人员通常所理解的相同的含义。在分子生物学中常用的术语和技术的定义可以在以下文献中找到:Molecular Cloning:A Laboratory Manual,第2版(1989)(Sambrook、Fritsch和Maniatis);Molecular Cloning:A Laboratory Manual,第4版(2012)(Green和Sambrook);Current Protocols in Molecular Biology(1987)(F.M.Ausubel等人编辑);the series Methods in Enzymology(Academic Press,Inc.):PCR 2:A Practical Approach(1995)(M.J.MacPherson、B.D.Hames和G.R.Taylor编辑):Antibodies,A Laboraotry Manual(1988)(Harlow和Lane编辑):Antibodies A Laboraotry Manual,第2版2013(E.A.Greenfield编辑);Animal Cell Culture(1987)(R.I.Freshney编辑);Benjamin Lewin,Genes IX,Jones and Bartlet出版,2008(ISBN 0763752223);Kendrew等人(编辑),The Encyclopedia of Molecular Biology,Blackwell Science Ltd.出版,1994(ISBN0632021829);Robert A.Meyers(编辑),Molecular Biology and Biotechnology:a Comprehensive Desk Reference,VCH Publishers,Inc.出版,1995(ISBN9780471185710);Singleton等人,Dictionary of Microbiology and Molecular Biology第2版,J.Wiley&Sons(New York,N.Y.1994),三月,Advanced Organic Chemistry Reactions,Mechanisms and Structure第4版,John Wiley&Sons(New York,N.Y.1992);和Marten H.Hofker和Jan van Deursen,Transgenic Mouse Methods and Protocols,第2版(2011)。

[0100] 参考2016年6月17日提交的美国临时申请62/351,662和62/351,803、2016年8月17日提交的美国临时申请62/376,377、2016年10月19日提交的美国临时申请62/410,366、2016年12月9日提交的美国临时申请62/432,240、2017年3月15日提交的美国临时申请62/

471,792和2017年4月12日提交的美国临时申请62/484,786。参考2017年6月19日提交的国际PCT申请PCT/US2017/038154。参考2017年3月15日提交的美国临时申请62/471,710(标题为“新颖Cas13B直系同源物CRISPR酶和系统(Novel Cas13B Orthologues CRISPR Enzymes and Systems)”,代理人编号:BI-10157VP 47627.04.2149)。进一步参考2016年12月9日提交的美国临时申请62/432,553、2017年2月8日提交的美国临时申请62/456,645和2017年3月15日提交的美国临时62/471,930(标题为“基于CRISPR效应子系统的诊断(CRISPR Effector System Based Diagnostics)”,代理人编号BI-10121BROD 0842P),以及2017年4月12日提交的有待转让的美国临时申请(标题为“基于CRISPR效应子系统的诊断(CRISPR Effector System Based Diagnostics)”,代理人BI-10121BROD 0843P)。

[0101] 除非上下文另有明确指示,否则如本文所用的单数形式“一个”、“一种”和“所述”包括单数和复数指示物。

[0102] 术语“任选的”或“任选地”意指后续描述的事件、情形或替代物可能发生或可能不发生,并且该描述包括事件或情形发生的情况和不发生的情况。

[0103] 通过端点表述的数值范围包括在对应范围内的所有数值和分数,以及所表述的端点。

[0104] 如本文所用的术语“约”或“近似”当涉及诸如参数、量、时距等可测量的值时,有意涵盖指定值的变化和从指定值的变化,诸如指定值和从指定值+/-10%或更小、+/-5%或更小、+/-1%或更小和+/-0.1%或更小的变化,只要此类变化适于在所公开的发明中执行即可。应当理解,修饰语“约”或“近似”所涉及的值本身也是具体地且优选地公开的。

[0105] 如本文所用,“生物样品”可以含有全细胞和/或活细胞和/或细胞碎片。生物样品可以含有(或来源于)“体液”。本发明涵盖以下实施方案,其中体液选自羊水、房水、玻璃体液、胆汁、血清、母乳、脑脊液、耳屎(耳垢)、乳糜、食糜、内淋巴液、外淋巴液、渗出液、粪便、女性射液、胃酸、胃液、淋巴液、粘液(包括鼻引流和痰)、心包液、腹膜液、胸膜液、脓液、发炎性分泌物、唾液、皮脂(皮油)、精液、痰涎、滑液、汗液、眼泪、尿液、阴道分泌物、呕吐物以及这些物质中一者或多者的混合物。生物样品包括细胞培养物、体液、来自体液的细胞培养物。体液可以例如通过穿刺或其他收集或采样程序从哺乳生物体获得。

[0106] 术语“受试者”、“个体”和“患者”在本文中可互换使用,是指脊椎动物,优选地是哺乳动物,更优选地是人类。哺乳动物包括但不限于鼠、猴、人类、农场动物、运动型动物和宠物。还包括在体内获得或在体外培养的生物实体的组织、细胞和他们的子代。

[0107] 在本说明书通篇提及“一个实施方案”、“实施方案”、“示例性实施方案”意指结合实施方案描述的特定特征、结构或特性包括于本发明的至少一个实施方案中。因此,在本说明书通篇各处出现短语“在一个实施方案中”、“在实施方案中”或“示例性实施方案”不一定全部指代同一实施方案,但也有可能如此。此外,如本领域技术人员从本公开将明显了解的,在一个或多个实施方案中,特定特征、结构或特性可以按任何适合的方式组合。此外,虽然本文所述的一些实施方案包括其他实施方案中所包括的一些特征而非其他特征,但不同实施方案的特征的组合有意处于本发明的范围内。例如,在所附权利要求中,任何所要求保护的实施方案可以按任何组合使用。

[0108] 先在将C2c2称为Cas13a。应当理解,本文中的术语“C2c2”与“Cas13a”可互换使用。如本文所用,Cas13可以指代Cas13a或Cas13b或Cas13c或Cas13d或Cas13家族中的其他成

员。在一些实例中,Cas13可以是Cas13a。在一些实例中,Cas13可以是Cas13b。在一些实例中,Cas13可以是Cas13c。在一些实例中,Cas13可以是Cas13d。

[0109] 本文所引用的所有公布、公布的专利文献和专利申请特此以引用方式并入,如同每个单独公布、公布的专利文献或专利申请被确切地且单独地指明为以引用方式整体并入。

[0110] 在下文中描述各种实施方案。应当指出的是,具体实施方案不旨在作为详尽的描述或作为对本文所论述的更广泛方面的限制。结合特定实施方案描述的一个方面不必限于该实施方案,而是可以与任何其他一个或多个实施方案一起实践。

[0111] 综述

[0112] 本文所公开的实施方案提供了将无催化活性的Cas效应蛋白(例如Cas13)用于各种应用的系统、构建体和方法。一般来讲,本文所公开的系统可以包含无催化活性的Cas效应蛋白(例如dCas13)。所述无催化活性的Cas13(dCas13)可以包括Cas13效应蛋白例如在C端、在N端或在这两端上的截短。在一些实施方案中,所述系统包含dCas13。所述dCas13可以是任何Cas13亚型蛋白的无催化活性形式。例如,dCas13可以是dCas13a、dCas13b、dCas13c或dCas13d。在一些实例中,所述dCas13可以是来自普雷沃氏菌属种(Prevotella sp.)P5-125、鸭疫里默氏杆菌(Riemerella anatipestifer)或咽喉卟啉单胞菌(Porphyromonas gulae)的修饰的Cas13效应蛋白。

[0113] 所述系统还可以包括用于期望应用的一种或多种功能组分。在一些实例中,所述系统可以包含碱基编辑组分,例如腺苷脱氨酶或胞苷脱氨酶,或其催化结构域。在一些实例中,所述系统可以包含转录因子或其活性结构域。在这些情况下,所述无催化活性的Cas效应蛋白可用于将所述转录因子递送至靶序列。在一些实例中,所述无催化活性的Cas效应蛋白可以是裂解Cas效应蛋白。所述裂解的Cas可以与另一种裂解Cas一起使用,以用于各种应用。例如,两个裂解Cas效应蛋白可以融合以形成催化活性的Cas效应蛋白。例如使用诱导剂,融合可以是可控的。

[0114] V型CRISPR-Cas蛋白

[0115] 本申请描述了使用V型CRISPR-Cas蛋白的方法。本文以Cas13为例,由此鉴定了多种直系同源物或同系物。对于本领域技术人员将显而易见的是,可以鉴定其他直系同源物或同系物,并且本文所述的任何功能物可以被工程化为其他直系同源物,包括包含来自多种直系同源物的片段的嵌合酶。

[0116] 鉴定新颖的CRISPR-Cas基因座的计算方法描述于EP3009511或US2016208243中,并且可包括以下步骤:检测编码Cas1蛋白的所有重叠群;鉴定cas1基因20kB内的所有预测蛋白编码基因;将已鉴定的基因与Cas蛋白特异性谱进行比较,并预测CRISPR阵列;选择含有大于500个氨基酸(>500aa)的蛋白质的未分类候选CRISPR-Cas基因座;使用诸如PSI-BLAST和HHPred的方法分析选定的候选物,以筛选已知的蛋白质结构域,从而鉴定新颖的2类CRISPR-Cas基因座(另参见Schmakov等人2015,Mol Cell.60(3):385-97)。除以上提及的步骤之外,还可以通过在宏基因组学数据库中搜索另外的同系物来另外对候选物进行分析。另外地或可替代地,为了将搜索扩展到非自主的CRISPR-Cas系统,可以使用CRISPR阵列作为种子执行相同的过程。

[0117] 在一方面,检测编码Cas1蛋白的所有重叠群通过GenemarkS执行,GenemarkS是一

种基因预测程序,其如以下进一步描述:“GeneMarkS:a self-training method for prediction of gene starts in microbial genomes.Implications for finding sequence motifs in regulatory regions.”John Besemer,Alexandre Lomsadze and Mark Borodovsky,Nucleic Acids Research(2001)29,第2607-2618页,以引用方式并入本文。

[0118] 在一方面,可以通过将已鉴定的基因与Cas蛋白特异性谱进行比较并根据NCBI保守结构域数据库(CDD)对其进行注释来鉴定所有预测的蛋白质编码基因,CDD是由针对古代结构域和全长蛋白质的完好注释的多序列比对模型集合组成的蛋白质注释资源。这些可作为位置特异性得分矩阵(PSSM)用于经由RPS-BLAST快速鉴定蛋白质序列中的保守结构域。CDD内容物包括NCBI编策的结构域,这些结构域使用3D结构信息来明确地限定结构域边界并为序列/结构/功能关系提供洞见;以及从多个外部来源数据库(Pfam、SMART、COG、PRK、TIGRFAM)导入的结构域模型。在另一方面,使用PILER-CR程序预测CRISPR阵列,PILER-CR程序是用于发现CRISPR重复序列的公共领域软件,如“PILER-CR:fast and accurate identification of CRISPR repeats”,Edgar,R.C.,BMC Bioinformatics,Jan 20;8:18(2007)所述,该文献以引用方式并入本文。

[0119] 在另一方面,使用PSI-BLAST(位置特异性迭代基本局部比对搜索工具)来进行逐例分析。PSI-BLAST使用蛋白质-蛋白质BLAST从经检测高于给定得分阈值的序列的多序列比对得出位置特异性得分矩阵(PSSM)或序型。该PSSM用于进一步在数据库中搜索新匹配项,并使用这些新检测到的序列进行更新以进行后续迭代。因此,PSI-BLAST提供了一种检测蛋白质之间远距离关系的手段。

[0120] 在另一方面,使用HHpred进行逐例分析,HHpred是一种用于序列数据库搜索和结构预测的方法,该方法与BLAST或PSI-BLAST一样容易使用,同时在发现远同系物方面更加灵敏。实际上,HHpred的灵敏度可与目前可用的功能最强大的结构预测服务器相竞争。HHpred是第一台基于序型隐蔽马尔科夫模型(HMM)的成对比较的服务器。大多数传统的序列搜索方法搜索序列数据库(例如UniProt或NR),而HHpred搜索比对数据库(如Pfam或SMART)。这大大简化了多个序列簇而非混乱的单一序列的命中物列表。所有主要的公开可得的序型和比对数据库均可通过HHpred获得。HHpred接受单一查询序列或多重比对作为输入。仅在几分钟内,它就以类似于PSI-BLAST的易于阅读的格式返回搜索结果。搜索选项包括局部或全局比对以及对二级结构相似性评分。HHpred可以生成成对的查询模板序列比对、合并的查询模板多重比对(例如用于传递搜索),以及由MODELLER软件根据HHpred比对计算出的3D结构模型。

[0121] Cas13的直系同源物

[0122] 术语“直系同源物(orthologue)”(在本文中也称为“直系同源物(ortholog)”)和“同系物(homologue)”(本文中也称为“同系物(homolog)”)在本领域中是众所周知的。作为进一步指导,如本文所用的蛋白质的“同系物”是与作为其同系物的蛋白质发挥相同或类似功能的相同种类的蛋白质。同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。如本文所用的蛋白质的“直系同源物”是与作为其直系同源物的蛋白质发挥相同或类似功能的不同种类的蛋白质。直系同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。同系物和直系同源物可以通过同源建模(参见例如Greer,Science第

228卷(1985)1055和Blundell等人Eur J Biochem vol 172(1988),513)或“结构BLAST”(Dey F,Cliff Zhang Q,Petrey D,Honig B.Toward a"structural BLAST":using structural relationships to infer function.Protein Sci.2013 Apr;22(4):359-66.doi:10.1002/pro.2225.)来鉴定。另参见Shmakov等人(2015)了解在CRISPR-Cas基因座领域中的申请。同源蛋白质可以是但不需要是结构上相关的,或仅是部分结构上相关的。

[0123] Cas13基因存在于若干种不同的细菌基因组中,典型地与cas1、cas2和cas4基因以及CRISPR盒(例如新凶手弗朗西丝菌(Francisella cf.novicida)Fx1的FNFX1_1431-FNFX1_1428)在同一基因座中。因此,此推定的新颖CRISPR-Cas系统的布局似乎与II-B型的布局类似。此外,与Cas9类似,Cas13蛋白含有与转座子ORF-B同源的易于鉴定的C端区,并且包含活性的RuvC样核酸酶、富含精氨酸的区和Zn指(不存在于Cas9中)。然而,与Cas9不同,Cas13还存在于没有CRISPR-Cas环境的若干种基因组中,并且其与ORF-B的相对较高相似性表明其可能是转座子组分。表明如果此是真正的CRISPR-Cas系统并且Cas13是Cas9的功能类似物,则其将是新颖CRISPR-Cas类型,即V型(参见Annotation and Classification of CRISPR-Cas Systems.Makarova KS,Koonin EV.Methods Mol Biol.2015;1311:47-75)。然而,如本文所述,将Cas13指代为亚型V-A以将其与C2c1p区分,该C2c1p不具有相同的结构域结构并且因此被指代为亚型V-B。

[0124] 本发明涵盖来源于被指代为亚型V-A的Cas13基因座的Cas13效应蛋白的用途。在本文中,此类效应蛋白也称为“Cas13p”,例如Cas13蛋白(并且这种效应蛋白或Cas13蛋白或来源于Cas13基因座的蛋白也称为“CRISPR-Cas蛋白”)。

[0125] 在特定实施方案中,效应蛋白是来自包括以下的属的生物体的Cas13效应蛋白:链球菌属(Streptococcus)、弯曲杆菌属(Campylobacter)、Nitratifractor、葡萄球菌属(Staphylococcus)、细小棒菌属(Parvibaculum)、罗氏菌属(Roseburia)、奈瑟氏菌属(Neisseria)、葡糖醋杆菌属(Gluconacetobacter)、固氮螺菌属(Azospirillum)、Sphaerochaeta、乳酸杆菌属(Lactobacillus)、真细菌属(Eubacterium)、棒状杆菌属(Corynebacter)、肉杆菌属(Carnobacterium)、红细菌属(Rhodobacter)、李斯特菌属(Listeria)、帕鲁迪菌属(Paludibacter)、梭菌属(Clostridium)、毛螺旋菌科(Lachnospiraceae)、Clostridiaridium、纤毛菌属(Leptotrichia)、弗朗西丝菌属、军团杆菌属(Legionella)、脂环酸芽孢杆菌属(Alicyclobacillus)、甲烷嗜甲基菌属(Methanomethyophilus)、卟啉单胞菌属(Porphyromonas)、普雷沃氏菌属、拟杆菌门(Bacteroidetes)、创伤球菌属(Helcococcus)、钩端螺旋体属(Leptospira)、脱硫弧菌属(Desulfovibrio)、脱硫盐碱杆菌属(Desulfonatronum)、丰佑菌科(Opitutaceae)、肿块芽孢杆菌属(Tuberibacillus)、芽孢杆菌属(Bacillus)、短芽孢杆菌属(Brevibacilus)、甲基杆菌属(Methylobacterium)、丁酸弧菌属(Butyvibrio)、异域菌门菌属(Perigrinibacterium)、副真杆菌属(Pareubacterium)、莫拉氏菌属(Moraxella)、硫微螺菌属(Thiomicrospira)或氨基酸球菌属。在特定实施方案中,Cas13效应蛋白选自选自以下的属的生物体:真细菌属、钩端螺旋体科、纤毛菌属、弗朗西丝菌属、甲烷嗜甲基菌属、卟啉单胞菌属、普雷沃氏菌属、钩端螺旋体属、丁酸弧菌属、异域菌门菌属、副真杆菌属、莫拉氏菌属、硫微螺菌属或氨基酸球菌属。

[0126] 在另外的特定实施方案中,Cas13效应蛋白来自选自以下的生物体:变异链球菌

(S.mutans)、无乳链球菌(S.agalactiae)、似马链球菌(S.equisimilis)、血链球菌(S.sanguinis)、肺炎链球菌;空肠弯曲杆菌(C.jejuni)、大肠弯曲杆菌(C.coli);N.salsuginis、N.tergarcus;耳葡萄球菌(S.auricularis)、肉葡萄球菌(S.carnosus);脑膜炎奈瑟氏菌(N.meningitides)、淋病奈瑟氏菌(N.gonorrhoeae);单核增生李斯特菌(L.monocytogenes)、伊氏李斯特菌(L.ivanovii);肉毒梭菌(C.botulinum)、艰难梭菌(C.difficile)、破伤风梭菌(C.tetani)、索氏梭菌(C.sordellii)、稻田氏钩端螺旋体(L.inadai)、土拉弗朗西斯菌(F.tularensis)1、易北普雷沃氏菌(P.albensis)、毛螺科菌(L.bacterium)、解蛋白丁酸弧菌(B.proteoclasticus)、异域菌门菌(P.bacterium)、狗口腔卟啉单胞菌(P.crevioricanis)、解糖胨普雷沃氏菌(P.disiens)和猕猴卟啉单胞菌(P.macacae)。

[0127] 效应蛋白可包含嵌合效应蛋白,所述嵌合效应蛋白包含来自第一效应蛋白(例如Cas13)直系同源物的第一片段和来自第二效应蛋白(例如Cas13)直系同源物的第二片段,并且其中第一和第二效应蛋白直系同源物是不同的。第一和第二效应蛋白(例如Cas13)直系同源物中的至少一者可以包含来自包括以下的生物体的效应蛋白(例如Cas13):链球菌属、弯曲杆菌属、Nitratifractor、葡萄球菌属、细小棒菌属、罗氏菌属、奈瑟氏菌属、葡糖醋杆菌属、固氮螺菌属、Sphaerochaeta、乳酸杆菌属、真细菌属、棒状杆菌属、肉杆菌属、红细菌属、李斯特菌属、帕鲁迪菌属、梭菌属、毛螺旋菌科、Clostridiaridium、纤毛菌属、弗朗西丝菌属、军团杆菌属、脂环酸芽孢杆菌属、甲烷嗜甲基菌属、卟啉单胞菌属、普雷沃氏菌属、拟杆菌门、创伤球菌属、钩端螺旋体属、脱硫弧菌属、脱硫盐碱杆菌属、丰佑菌科、肿块芽孢杆菌属、芽孢杆菌属、短芽孢杆菌属、甲基杆菌属、丁酸弧菌属、异域菌门菌属、副真杆菌属、莫拉氏菌属、硫微螺菌属或氨基酸球菌属;例如包含第一片段和第二片段的嵌合效应蛋白,其中第一片段和第二片段各自选自包括以下的生物体的Cas13:链球菌属、弯曲杆菌属、Nitratifractor、葡萄球菌属、细小棒菌属、罗氏菌属、奈瑟氏菌属、葡糖醋杆菌属、固氮螺菌属、Sphaerochaeta、乳酸杆菌属、真细菌属、棒状杆菌属、肉杆菌属、红细菌属、李斯特菌属、帕鲁迪菌属、梭菌属、毛螺旋菌科、Clostridiaridium、纤毛菌属、弗朗西丝菌属、军团杆菌属、脂环酸芽孢杆菌属、甲烷嗜甲基菌属、卟啉单胞菌属、普雷沃氏菌属、拟杆菌门、创伤球菌属、钩端螺旋体属、脱硫弧菌属、脱硫盐碱杆菌属、丰佑菌科、肿块芽孢杆菌属、芽孢杆菌属、短芽孢杆菌属、甲基杆菌属、丁酸弧菌属、异域菌门菌属、副真杆菌属、莫拉氏菌属、硫微螺菌属或氨基酸球菌属,其中第一片段和第二片段并非来自相同细菌;例如,包含第一片段和第二片段的嵌合效应蛋白,其中第一片段和第二片段各自选自包括以下的生物体的Cas13:变异链球菌、无乳链球菌、似马链球菌、血链球菌、肺炎链球菌;空肠弯曲杆菌、大肠弯曲杆菌;N.salsuginis、N.tergarcus;耳葡萄球菌、肉葡萄球菌;脑膜炎奈瑟氏菌、淋病奈瑟氏菌;单核增生李斯特菌、伊氏李斯特菌;肉毒梭菌、艰难梭菌、破伤风梭菌、索氏梭菌;土拉弗朗西斯菌1、易北普雷沃氏菌、毛螺科菌MC2017 1、解蛋白丁酸弧菌、异域菌门菌GW2011_GWA2_33_10、帕库氏菌GW2011_GWC2_44_17、史密斯氏菌属种SCADC、氨基酸球菌属种BV3L6、毛螺科菌MA2020、候选白蚁甲烷枝原体、挑剔真细菌、牛眼莫拉氏菌237、稻田钩端螺旋体、毛螺科菌ND2006、狗口腔卟啉单胞菌3、解糖胨普雷沃氏菌和猕猴卟啉单胞菌,其中第一片段和第二片段并非来自相同细菌。

[0128] 在更优选的实施方案中,Cas13p来源于选自以下的细菌种类:土拉弗朗西斯菌1、

易北普雷沃氏菌、毛螺科菌MC2017 1、解蛋白丁酸弧菌、异域菌门菌GW2011_GWA2_33_10、帕库氏菌GW2011_GWC2_44_17、史密斯氏菌属种SCADC、氨基酸球菌属种BV3L6、毛螺科菌MA2020、候选白蚁甲烷枝原体、挑剔真细菌、牛眼莫拉氏菌237、牛眼莫拉氏菌AAX08_00205、牛眼莫拉氏菌AAX11_00205、丁酸弧菌属种NC3005、硫微螺菌属种XS5、稻田钩端螺旋体、毛螺科菌ND2006、狗口腔卟啉单胞菌3、解糖胨普雷沃氏菌和猕猴卟啉单胞菌。在某些实施方案中，Cas13p来源于选自氨基酸球菌属种BV3L6、毛螺科菌MA2020的细菌种类。在某些实施方案中，效应蛋白来源于土拉弗朗西斯菌1的亚种，包括但不限于土拉弗朗西斯菌新凶手(Novicida)亚种。在某些优选的实施方案中，Cas13p来源于选自以下的细菌种类：氨基酸球菌属种BV3L6、毛螺科菌ND2006、毛螺科菌MA2020、牛眼莫拉氏菌AAX08_00205、牛眼莫拉氏菌AAX11_00205、丁酸弧菌属种NC3005、硫微螺菌属种XS5。

[0129] 在特定实施方案中，如本文所提及的Cas13的同系物或直系同源物与本文所公开的示例性Cas13蛋白具有至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同源性或同一性。在另外的实施方案中，如本文所提及的Cas13的同系物或直系同源物与野生型Cas13具有至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同一性。在Cas13具有一个或多个突变(是突变的)的情况下，如本文所提及的所述Cas13的同系物或直系同源物与突变的Cas13具有至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同一性。

[0130] 在一个实施方案中，Cas13蛋白可以是包括但不限于以下的属的生物体的直系同源物：氨基酸球菌属种、毛螺科菌或牛眼莫拉氏菌；在特定实施方案中，V型Cas蛋白可以是包括但不限于以下的属的生物体的直系同源物：氨基酸球菌属种BV3L6、毛螺科菌ND2006(LbCas13)或牛眼莫拉氏菌237。在特定实施方案中，如本文所提及的Cas13的同系物或直系同源物与本文所公开的Cas13序列中的一者或多者具有至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同源性或同一性。在另外的实施方案中，如本文所提及的Cas13的同系物或直系同源物与野生型FnCas13、AsCas13或LbCas13具有至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同一性。

[0131] 在特定实施方案中，本发明的Cas13蛋白与FnCas13、AsCas13或LbCas13具有至少60％，更特别地至少70％，诸如至少80％，更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同源性或同一性。在另外的实施方案中，如本文所提及的Cas13蛋白与野生型AsCas13或LbCas13具有至少60％，诸如至少70％，更特别地至少80％、更优选地至少85％、甚至更优选地至少90％，诸如例如至少95％的序列同一性。在特定实施方案中，本发明的Cas13蛋白与FnCas13具有少于60％的序列同一性。技术人员将理解，这包括Cas13蛋白的截短形式，由此在截短形式的长度上确定序列同一性。在特定实施方案中，Cas13酶不是FnCas13。

[0132] 修饰的Cas13酶

[0133] 在特定实施方案中，令人感兴趣的是利用本文所定义的工程化的Cas13蛋白(诸如Cas13)，其中所述蛋白与包含RNA的核酸分子复合以形成CRISPR复合物，其中当在CRISPR复合物中时，所述核酸分子靶向一个或多个靶多核苷酸基因座，所述蛋白与未修饰的Cas13蛋白相比包含至少一种修饰，并且其中包含修饰的蛋白的CRISPR复合物与包含未修饰的Cas13蛋白的复合物相比具有改变的活性。应当理解，当在本文中提及CRISPR“蛋白”时，

Cas13蛋白质优选地是修饰的CRISPR-Cas蛋白(例如,具有增加的或减少的(或没有)酶活性),如非限制地包括Cas13。术语“CRISPR蛋白”可以与“CRISPR-Cas蛋白”互换使用,不论所述CRISPR蛋白与野生型CRISPR蛋白相比是否具有改变的,如增加的或减少的(或没有)酶活性。

[0134] Cas13核酸酶一级结构的计算分析揭示了三个不同的区。第一是C端RuvC样结构域,其是仅功能表征的结构域。第二是N端α-螺旋区并且第三是位于RuvC样结构域与α-螺旋区之间的混合的α区和β区。

[0135] 预测非结构化区的若干小段在Cas13一级结构之内。对于小的蛋白质序列的裂解和插入而言,不同的Cas13直系同源物内的暴露于溶剂且不保守的非结构化区是优选的侧面。另外,这些侧面可以用于在Cas13直系同源物之间产生嵌合蛋白。

[0136] 基于以上信息,可以产生突变体,这些突变体使得酶失活或将双链核酸酶修饰为具有切口酶活性。在替代实施方案中,此信息用于开发具有减少的脱靶效应的酶(在本文其他地方描述)。

[0137] 在以上所述的某些Cas13酶中,酶是通过一个或多个残基(在RuvC结构域中)的突变来修饰的,这些残基包括但不限于参照AsCas13(氨基酸球菌属种BV3L6)的氨基酸位置编号的位置R909、R912、R930、R947、K949、R951、R955、K965、K968、K1000、K1002、R1003、K1009、K1017、K1022、K1029、K1035、K1054、K1072、K1086、R1094、K1095、K1109、K1118、K1142、K1150、K1158、K1159、R1220、R1226、R1242和/或R1252。在某些实施方案中,包含所述一个或多个突变的Cas13酶具有修改的,更优选地增加的对靶标的特异性。

[0138] 在以上所述的某些非天然存在的CRISPR-Cas蛋白中,酶是通过一个或多个残基(在RAD50中)的突变来修饰的,这些残基包括但不限于参照AsCas13(氨基酸球菌属种BV3L6)的氨基酸位置编号的位置K324、K335、K337、R331、K369、K370、R386、R392、R393、K400、K404、K406、K408、K414、K429、K436、K438、K459、K460、K464、R670、K675、R681、K686、K689、R699、K705、R725、K729、K739、K748和/或K752。在某些实施方案中,包含所述一个或多个突变的Cas13酶具有修改的,更优选地增加的对靶标的特异性。

[0139] 在某些Cas13酶中,酶是通过一个或多个残基的突变来修饰的,这些残基包括但不限于参照AsCas13(氨基酸球菌属种BV3L6)的氨基酸位置编号的位置R912、T923、R947、K949、R951、R955、K965、K968、K1000、R1003、K1009、K1017、K1022、K1029、K1072、K1086、F1103、R1226和/或R1252。在某些实施方案中,包含所述一个或多个突变的Cas13酶具有修改的,更优选地增加的对靶标的特异性。

[0140] 在某些实施方案中,Cas13酶是通过一个或多个残基的突变来修饰的,这些残基包括但不限于参照LbCas13(毛螺科菌ND2006)的氨基酸位置编号的位置R833、R836、K847、K879、K881、R883、R887、K897、K900、K932、R935、K940、K948、K953、K960、K984、K1003、K1017、R1033、R1138、R1165和/或R1252。在某些实施方案中,包含所述一个或多个突变的Cas13酶具有修改的,更优选地增加的对靶标的特异性。

[0141] 在某些实施方案中,Cas13酶是通过一个或多个残基的突变来修饰的,这些残基包括但不限于参照AsCas13(氨基酸球菌属种BV3L6)的氨基酸位置编号的位置K15、R18、K26、Q34、R43、K48、K51、R56、R84、K85、K87、N93、R103、N104、T118、K123、K134、R176、K177、R192、K200、K226、K273、K275、T291、R301、K307、K369、S404、V409、K414、K436、K438、K468、D482、

K516、R518、K524、K530、K532、K548、K559、K570、R574、K592、D596、K603、K607、K613、C647、R681、K686、H720、K739、K748、K757、T766、K780、R790、P791、K796、K809、K815、T816、K860、R862、R863、K868、K897、R909、R912、T923、R947、K949、R951、R955、K965、K968、K1000、R1003、K1009、K1017、K1022、K1029、A1053、K1072、K1086、F1103、S1209、R1226、R1252、K1273、K1282和/或K1288。在某些实施方案中，包含所述一个或多个突变的Cas13酶具有修改的，更优选地增加的对靶标的特异性。

[0142] 在某些实施方案中，酶是通过一个或多个残基的突变来修饰的，这些残基包括但不限于参照FnCas13(新凶手弗朗西斯菌U112)的氨基酸位置编号的位置K15、R18、K26、R34、R43、K48、K51、K56、K87、K88、D90、K96、K106、K107、K120、Q125、K143、R186、K187、R202、K210、K235、K296、K298、K314、K320、K326、K397、K444、K449、E454、A483、E491、K527、K541、K581、R583、K589、K595、K597、K613、K624、K635、K639、K656、K660、K667、K671、K677、K719、K725、K730、K763、K782、K791、R800、K809、K823、R833、K834、K839、K852、K858、K859、K869、K871、R872、K877、K905、R918、R921、K932、I960、K962、R964、R968、K978、K981、K1013、R1016、K1021、K1029、K1034、K1041、K1065、K1084和/或K1098。在某些实施方案中，包含所述一个或多个突变的Cas13酶具有修改的，更优选地增加的对靶标的特异性。

[0143] 在某些实施方案中，酶是通过一个或多个残基的突变来修饰的，这些残基包括但不限于参照LbCas13(毛螺科菌ND2006)的氨基酸位置编号的位置K15、R18、K26、K34、R43、K48、K51、R56、K83、K84、R86、K92、R102、K103、K116、K121、R158、E159、R174、R182、K206、K251、K253、K269、K271、K278、P342、K380、R385、K390、K415、K421、K457、K471、A506、R508、K514、K520、K522、K538、Y548、K560、K564、K580、K584、K591、K595、K601、K634、K640、R645、K679、K689、K707、T716、K725、R737、R747、R748、K753、K768、K774、K775、K785、K787、R788、Q793、K821、R833、R836、K847、K879、K881、R883、R887、K897、K900、K932、R935、K940、K948、K953、K960、K984、K1003、K1017、R1033、K1121、R1138、R1165、K1190、K1199和/或K1208。在某些实施方案中，包含所述一个或多个突变的Cas13酶具有修改的，更优选地增加的对靶标的特异性。

[0144] 在某些实施方案中，酶是通过一个或多个残基的突变来修饰的，这些残基包括但不限于参照MbCas13(牛眼莫拉氏菌237)的氨基酸位置编号的位置K14、R17、R25、K33、M42、Q47、K50、D55、K85、N86、K88、K94、R104、K105、K118、K123、K131、R174、K175、R190、R198、I221、K267、Q269、K285、K291、K297、K357、K403、K409、K414、K448、K460、K501、K515、K550、R552、K558、K564、K566、K582、K593、K604、K608、K623、K627、K633、K637、E643、K780、Y787、K792、K830、Q846、K858、K867、K876、K890、R900、K901、M906、K921、K927、K928、K937、K939、R940、K945、Q975、R987、R990、K1001、R1034、I1036、R1038、R1042、K1052、K1055、K1087、R1090、K1095、N1103、K1108、K1115、K1139、K1158、R1172、K1188、K1276、R1293、A1319、K1340、K1349和/或K1356。在某些实施方案中，包含所述一个或多个突变的Cas13酶具有修改的，更优选地增加的对靶标的特异性。

[0145] 在一个实施方案中，Cas13蛋白是通过在参照AsCas13的氨基酸位置编号的S1228处的突变(例如S1228A)来修饰的。参见Yamano等人，Cell 165:949-962(2016)，该文献以引用方式整体并入本文。

[0146] 在某些实施方案中，Cas13蛋白已被修饰以识别非天然PAM，如识别具有以下序列

或包含以下序列的PAM:YCN、YCV、AYV、TYV、RYN、RCN、TGYV、NTTN、TTN、TRTN、TYTV、TYCT、TYCN、TRTN、NTTN、TACT、TYCC、TRTC、TATV、NTTV、TTV、TSTG、TVTS、TYYS、TCYS、TBYS、TCYS、TNYS、TYYS、TNTN、TSTG、TTCC、TCCC、TATC、TGTG、TCTG、TYCV或TCTC。在特定实施方案中,所述突变的Cas13包含在AsCas13的位置11、12、13、14、15、16、17、34、36、39、40、43、46、47、50、54、57、58、111、126、127、128、129、130、131、132、133、134、135、136、157、158、159、160、161、162、163、164、165、166、167、168、169、170、171、172、173、174、175、176、177、178、532、533、534、535、536、537、538、539、540、541、542、543、544、545、546、547、548、549、550、551、552、553、554、555、556、565、566、567、568、569、570、571、572、573、574、575、592、593、594、595、596、597、598、599、600、601、602、603、604、605、606、607、608、609、610、611、612、613、614、615、616、617、618、619、620、626、627、628、629、630、631、632、633、634、635、636、637、638、642、643、644、645、646、647、648、649、651、652、653、654、655、656、676、679、680、682、683、684、685、686、687、688、689、690、691、692、693、707、711、714、715、716、717、718、719、720、721、722、739、765、768、769、773、777、778、779、780、781、782、783、784、785、786、871、872、873、874、875、876、877、878、879、880、881、882、883、884或1048或在Cas13直系同源物中与其对应的位置处的一个或多个突变的氨基酸残基;优选地包含在位置130、131、132、133、134、135、136、162、163、164、165、166、167、168、169、170、171、172、173、174、175、176、177、536、537、538、539、540、541、542、543、544、545、546、547、548、549、550、551、552、570、571、572、573、595、596、597、598、599、600、601、602、603、604、605、606、607、608、609、610、611、612、613、614、615、630、631、632、646、647、648、649、650、651、652、653、683、684、685、686、687、688、689或690处的一个或多个突变的氨基酸残基;

[0147] 在某些实施方案中,Cas13蛋白被修饰成具有增加的活性,即更宽的PAM特异性。在特定实施方案中,Cas13是通过一个或多个残基的突变来修饰的,这些残基包括但不限于AsCas13的位置539、542、547、548、550、551、552、167、604和/或607或AsCas13直系同源物、同系物或变体的相应位置,优选地在542或542和607处的突变的氨基酸残基,其中所述突变优选地为542R和607R,诸如S542R和K607R;或优选地在位置542和548(且任选地552)处的突变的氨基酸残基,其中所述突变优选地为542R和548V(且任选地552R),诸如S542R和K548V(且任选地N552R);或LbCas13的位置532、538、542和/或595,或AsCas13直系同源物、同系物或变体的相应位置,优选地在532或532和595处的突变的氨基酸残基,其中所述突变优选地为532R和595R,诸如G532R和K595R;或优选地在位置532和538(且任选地542)处的突变的氨基酸残基,其中所述突变优选地为532R和538V(且任选地542R),诸如G532R和K538V(且任选地Y542R),最优选地其中所述突变为AsCas13的S542R和K607R、S542R和K548V,或S542R、K548V和N552R。

[0148] 灭活/失活的Cas13蛋白

[0149] 在Cas13蛋白具有核酸酶活性的情况下,Cas13蛋白可以被修饰成具有减弱的核酸酶活性,例如,与野生型酶相比,具有至少70%、至少80%、至少90%、至少95%、至少97%或100%的核酸酶失活;或者换句话说,Cas13酶有利地具有非突变型或野生型Cas13酶或CRISPR-Cas蛋白的核酸酶活性的约0%,或不超过非突变型或野生型Cas13酶的核酸酶活性的约3%或约5%或约10%,这些酶例如是属于非突变型或野生型新凶手弗朗西斯菌U112(FnCas13)、氨基酸球菌属种BV3L6(AsCas13)、毛螺科菌ND2006(LbCas13)或牛眼莫拉氏菌

237(MbCas13 Cas13酶或CRISPR-Cas蛋白)。有可能通过将突变引入到Cas13及其直系同源物的核酸酶结构域中实现此举。

[0150] 在本发明的优选实施方案中,使用至少一种Cas13蛋白,其为Cas13切口酶。更特别地,使用Cas13切口酶,其不使靶链产生切口而是能够仅使与靶链互补的链,即在本文中也称为与指导序列不互补的链的非靶DNA链产生切口。更特别地,Cas13切口酶是Cas13蛋白,所述Cas13蛋白包含来自氨基酸球菌属种的Cas13的Nuc结构域中的位置1226A或Cas13直系同源物中的相应位置处的精氨酸突变。在另外的特定实施方案中,该酶包含精氨酸至丙氨酸取代或R1226A突变。本领域技术人员将理解,在酶不是AsCas13的情况下,可以在相应位置的残基处形成突变。在特定实施方案中,Cas13是FnCas13,并且突变是在位置R1218处的精氨酸上。在特定实施方案中,Cas13是LbCas13,并且突变是在位置R1138处的精氨酸上。在特定实施方案中,Cas13是MbCas13,并且突变是在位置R1293处的精氨酸上。

[0151] 在某些实施方案中,另外地或可替代地使用经工程化的并且可以包含降低或消除核酸酶活性的一个或多个突变的CRISPR-Cas蛋白。FnCas13p RuvC结构域中的氨基酸位置包括但不限于D917A、E1006A、E1028A、D1227A、D1255A、N1257A、D917A、E1006A、E1028A、D1227A、D1255A和N1257A。申请人还鉴定了与PD-(D/E)XK核酸酶超家族和HincII样内切核酸酶最类似的推定的第二核酸酶结构域。在此推定的核酸酶结构域中产生的大幅度降低核酸酶活性的点突变包括但不限于N580A、N584A、T587A、W609A、D610A、K613A、E614A、D616A、K624A、D625A、K627A和Y629A。在优选的实施方案中,FnCas13p RuvC结构域中的突变是D917A或E1006A,其中D917A或E1006A突变使FnCas13效应蛋白的DNA切割活性完全失活。在另一个实施方案中,FnCas13p RuvC结构域中的突变是D1255A,其中突变的FnCas13效应蛋白具有明显降低的核溶解活性。

[0152] 更特别地,失活的Cas13酶包括在AsCas13的氨基酸位置As908、As993、As1263或Cas13直系同源物中的相应位置突变的酶。另外,失活的Cas13酶包括在LbCas13的氨基酸位置Lb832、925、947或1180或Cas13直系同源物中的相应位置突变的酶。更特别地,失活的Cas13酶包括包含AsCas13的突变AsD908A、AsE993A、AsD1263A或Cas13直系同源物中的相应突变中的一个或多个的酶。另外,失活的Cas13酶包括包含LbCas13的突变LbD832A、E925A、D947A或D1180A或Cas13直系同源物中的相应突变中的一个或多个的酶。

[0153] 突变还可以在邻近残基处,例如在靠近以上指出的参与核酸酶活性的那些的氨基酸处形成。在一些实施方案中,仅RuvC结构域是失活的,而在其他实施方案中,另一推定的核酸酶结构域是失活的,其中效应蛋白复合物充当切口酶并且仅切割一条DNA链。在一个优选的实施方案中,其他推定的核酸酶结构域是HincII样内切核酸酶结构域。

[0154] 失活的Cas13或Cas13切口酶可以具有缔合的(例如经由融合蛋白)一个或多个功能结构域,包括例如腺苷脱氨酶或其催化结构域。在一些情况下,有利的是另外提供至少一个异源NLS。在一些情况下,将NLS定位在N端处是有利的。一般来讲,一个或多个功能结构域在失活的Cas13或Cas13切口酶上的定位是允许功能结构域的正确空间取向,从而以属性化的功能效应影响靶标的定位。例如,当功能结构域是腺苷脱氨酶其催化结构域时,腺苷脱氨酶催化结构域被定位成允许其接触靶腺嘌呤并使靶腺嘌呤脱氨基的空间取向。此可以包括除Cas13的N端/C端之外的位置。在一些实施方案中,腺苷脱氨酶蛋白或其催化结构域被插入到Cas13的内环中。

[0155] 在某些实施方案中，如本文所述的根据本发明的效应蛋白(CRISPR酶;Cas13;效应蛋白)是无催化活性的或死亡的Cas13效应蛋白(dCas13)。在一些实施方案中，dCas13效应子包含在核酸酶结构域中的突变。在一些实施方案中，dCas13效应蛋白已被截短。在一些实施方案中，dCas13在以下一个或多个位置处包含Cas13效应蛋白的截短形式:Cas13效应蛋白的C端、Cas13效应蛋白的N端或Cas13效应蛋白的一个或多个结构域(例如HEPN结构域)。例如，dCas13可以在Cas13效应蛋白的C端包含Cas13效应蛋白的截短形式。例如，dCas13可以在Cas13效应蛋白的N端包含Cas13效应蛋白的截短形式。例如，dCas13可以在Cas13效应蛋白的HEPN结构域处包含Cas13效应蛋白的截短形式。在一些实施方案中，此类截短可以减小所述Cas13效应子和所述一种或多种功能组分的融合蛋白的大小。在一些情况下，截短的Cas13可能仍保持其RNA结合功能。

[0156] 在一些实施方案中，Cas13效应子的C端可以被截短。例如，Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少50个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸、至少120个氨基酸、至少140个氨基酸、至少150个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少250个氨基酸、至少260个氨基酸、或至少300个氨基酸、或至少350个氨基酸、或最多120个氨基酸、或最多140个氨基酸、或最多160个氨基酸、或最多180个氨基酸、或最多200个氨基酸、或最多250个氨基酸、或最多300个氨基酸、或最多350个氨基酸、或最多400个氨基酸可以被截短。Cas13截短的实例包括C端Δ984-1090、C端Δ1026-1090和C端Δ1053-1090、C端Δ934-1090、C端Δ884-1090、C端Δ834-1090、C端Δ784-1090和C端Δ734-1090，其中氨基酸位置对应于普雷沃氏菌属种P5-125 Cas13b蛋白的氨基酸位置。Cas13截短的实例还包括C端Δ795-1095，其中氨基酸位置对应于鸭疫里默氏杆菌Cas13b蛋白的氨基酸位置。Cas13截短的实例还包括C端Δ875-1175、C端Δ895-1175、C端Δ915-1175、C端Δ935-1175、C端Δ955-1175、C端Δ975-1175、C端Δ995-1175、C端Δ1015-1175、C端Δ1035-1175、C端Δ1055-1175、C端Δ1075-1175、C端Δ1095-1175、C-端Δ1115-1175、C-端Δ1135-1175、C-端Δ1155-1175，其中氨基酸位置对应于咽喉卟啉单胞菌Cas13b蛋白的氨基酸位置。技术人员将理解，可以为其他Cas13b直系同源物或其他Cas13类型或亚型(诸如Cas13a、Cas13c或Cas13d)设计类似的截短。

[0157] 在一些实施方案中，Cas13效应蛋白的N端可以被截短。例如，Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少50个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸、至少120个氨基酸、至少140个氨基酸、至少150个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少250个氨基酸、至少260个氨基酸、或至少300个氨基酸、或至少350个氨基酸、或最多120个氨基酸、或最多140个氨基酸、或最多160个氨基酸、或最多180个氨基酸、或最多200个氨基酸、或最多250个氨基酸、或最多300个氨基酸、或最多350个氨基酸、或最多400个氨基酸可以被截短。Cas13截短的实例包括N端Δ1-125、N端Δ1-88或N端Δ1-72，其中截短的氨基酸位置对应于普雷沃氏菌种P5-125 Cas13b蛋白的氨基酸位置。

[0158] 在一些实施方案中，Cas13效应蛋白的N端和C端都可以被截短。例如，Cas13效应子的C端的至少20个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少

140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少40个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少60个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少80个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少100个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少120个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少140个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少160个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少180个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少200个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个

氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少220个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少240个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少260个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少280个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少300个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的C端的至少350个氨基酸可以被截短，并且Cas13效应子的N端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少20个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少40个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少60个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端

的至少80个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少100个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少120个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少140个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少160个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少180个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少200个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少220个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少240个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少260个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60

个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少280个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少300个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。例如，Cas13效应子的N端的至少350个氨基酸可以被截短，并且Cas13效应子的C端的至少20个氨基酸、至少40个氨基酸、至少60个氨基酸、至少80个氨基酸、至少100个氨基酸酸、至少120个氨基酸、至少140个氨基酸、至少160个氨基酸、至少180个氨基酸、至少200个氨基酸、至少220个氨基酸、至少240个氨基酸、至少260个氨基酸、至少300个氨基酸、或至少350个氨基酸可以被截短。

[0159] CRISPR-Cas蛋白和指导物

[0160] 在本发明的方法和系统中，使用了CRISPR-Cas蛋白和相应的指导分子。更特别地，CRISPR-Cas蛋白是2类CRISPR-Cas蛋白。在某些实施方案中，所述CRISPR-Cas蛋白是Cas13。CRISPR-Cas系统不需要生成靶向特定序列的定制蛋白，而是单个Cas蛋白可以通过指导分子进行编程以识别特定核酸靶标，换句话说，可以使用所述指导分子将Cas酶蛋白募集至目标靶基因座的特定核酸。

[0161] 指导分子

[0162] 2类V型CRISPR-Cas蛋白的指导分子或指导RNA包含tracr配对序列(在内源性CRISPR系统的情形中涵盖“正向重复序列”)和指导序列(在内源性CRISPR系统的情形中也称为“间隔区”)。实际上，与II型CRISPR-Cas蛋白成对比，Cas13蛋白不依赖于tracr序列的存在。在一些实施方案中，本文所述的CRISPR-Cas系统或复合物不包含和/或不依赖于tracr序列的存在(例如，若Cas蛋白是Cas13)。在某些实施方案中，指导分子可以包含融合或连接至指导序列或间隔区序列的正向重复序列，基本上由其组成，或由其组成。

[0163] 一般来讲，CRISPR系统由促进在靶序列的位点处CRISPR复合物形成的元件表征。在形成CRISPR复合物的情形中，“靶序列”是指指导序列被设计成与其具有互补性的序列，其中靶DNA序列与指导序列之间的杂交促进CRISPR复合物的形成。

[0164] 术语“指导分子”和“指导RNA”在本文可互换使用，是指能够与CRISPR-Cas蛋白形成复合物并且包含与靶核酸序列具有足够的互补性以与靶核酸序列杂交并引导复合物序列特异性地结合至靶核酸序列的指导序列的基于RNA的分子。如本文所述，指导分子或指导RNA特别涵盖具有一种或多种化学修饰(例如，通过化学连接两个核糖核苷酸或通过用一种或多种脱氧核糖核苷酸替代一种或多种核糖核苷酸)的基于RNA的分子。

[0165] 如本文所用，在CRISPR-Cas系统的情形中，术语“指导序列”包括与靶核酸序列具有足够互补性以与靶核酸序列杂交并且引导核酸靶向复合物序列特异性地结合至靶核酸

序列的任何多核苷酸序列。在本发明的情形中，靶核酸序列或靶序列是包含待脱氨基的靶腺苷(在本文中也称为“靶腺苷”)的序列。在一些实施方案中，除预期dA-C错配外，当使用合适的比对算法最佳比对时，互补程度为约或大于约50％、60％、75％、80％、85％、90％、95％、97.5％、99％或更大。最佳比对可以借助于用于比对序列的任何合适算法来确定，其非限制性实例包括史密斯-沃特曼算法(Smith-Waterman algorithm)、尼德曼-翁施算法(Needleman-Wunsch algorithm)、基于巴罗斯-维勒变换(Burrows-Wheeler Transform)的算法(例如巴罗斯-维勒比对仪(Burrows Wheeler Aligner))、ClustalW、Clustal X、BLAT、Novoalign(Novocraft Technologies;在www.novocraft.com上可得)、ELAND(Illumina,San Diego,CA)、SOAP(在soap.genomics.org.cn上可得)和Maq(在maq.sourceforge.net上可得)。指导序列(在核酸靶向指导RNA内)引导核酸靶向复合物序列特异性地结合至靶核酸序列的能力可以通过任何合适的测定来评定。举例来说，足以形成核酸靶向复合物的核酸靶向CRISPR系统的组分，包括有待测试的指导序列，可以提供给具有相应靶核酸序列的宿主细胞，诸如通过用编码核酸靶向复合物的组分的载体转染，继而诸如通过如本文所述的Surveyor测定评定靶核酸序列内的优先靶向(例如切割)。类似地，可以在试管中通过提供靶核酸序列、核酸靶向复合物的组分(包括有待测试的指导序列)和不同于测试指导序列的对照指导序列，以及在测试指导序列与对照指导序列反应之间比较靶序列处或附近的结合或切割速率来评估靶核酸序列(或其附近的序列)的切割。其他测定可能存在，并且将为本领域技术人员所想到。可以选择指导序列并且因此选择核酸靶向指导RNA以靶向任何靶核酸序列。

[0166] 在一些实施方案中，指导分子包含被设计成与靶序列具有至少一个错配的指导序列，使得在指导序列与靶序列之间所形成的RNA双链体包含在指导序列中的与靶A相对的非配对C，以便于靶序列上的脱氨。在一些实施方案中，除该A-C错配外，当使用合适的比对算法最佳比对时，互补程度为约或大于约50％、60％、75％、80％、85％、90％、95％、97.5％、99％或更大。

[0167] 在某些实施方案中，指导分子的指导序列或间隔区长度为15至50nt。在某些实施方案中，指导RNA的间隔区长度为至少15个核苷酸。在某些实施方案中，间隔区长度为15至17nt，例如15、16或17nt；17至20nt，例如17、18、19或20nt；20至24nt，例如20、21、22、23或24nt；23至25nt，例如23、24或25nt；24至27nt，例如24、25、26或27nt；27-30nt，例如27、28、29或30nt；30-35nt，例如30、31、32、33、34或35nt；或35nt或更长。在某些示例性实施方案中，指导序列为15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39 40、41、42、43、44、45、46、47 48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99或100nt。

[0168] 在一些实施方案中，指导序列是长度为10至50nt之间，但更特别地为约20-30nt，有利地为约20nt、23-25nt或24nt的RNA序列。选择指导序列以确保其与包含待脱氨基的腺苷的靶序列杂交。对此将在下文更详细地描述。选择可涵盖进一步的步骤，这些步骤可以提高脱氨的功效和特异性。

[0169] 在一些实施方案中，指导序列长约20nt至约30nt，并且与靶DNA链杂交以形成几乎完美匹配的双链体，除了在靶腺苷位点处具有dA-C错配。特别地，在一些实施方案中，dA-C

错配位于靠近靶序列的中心(并且因此在指导序列与靶序列杂交后的双链体的中心),从而将腺苷脱氨酶限制于狭窄的编辑窗口(例如,约4bp宽)。在一些实施方案中,靶序列可以包含多于一个待脱氨基的靶腺苷。在另外的实施方案中,靶序列可还包含在靶腺苷位点3'端的一个或多个dA-C错配。在一些实施方案中,为了避免在靶序列中非预期的腺嘌呤位点处的脱靶编辑,可以将指导序列设计成在对应于所述非预期的腺嘌呤的位置上包含非配对鸟嘌呤以引入dA-G错配,对于某些腺苷脱氨酶(诸如ADAR1和ADAR2)而言这不具有催化益处。参见Wong等人,RNA7:846-858(2001),该文献以引用方式整体并入本文。

[0170] 在一些实施方案中,使用具有规范长度(例如,对于AsCas13为约24nt)的Cas13指导序列与靶DNA形成RNA双链体。在一些实施方案中,使用比典型长度(例如,对于AsCas13为>24nt)长的Cas13指导分子与靶DNA形成RNA双链体,包括在Cas13指导RNA-靶DNA复合物之外形成双链体。当有意使给定的一段核苷酸内的多于一个腺嘌呤脱氨基时,这可能是令人感兴趣的。在替代实施方案中,维持典型指导序列长度的限制是令人感兴趣的。在一些实施方案中,指导序列被设计来在Cas13指导物的典型长度之外引入dA-C错配,这可以减少Cas13的空间位阻并且增加腺苷脱氨酶与dA-C错配之间的接触频率。

[0171] 在一些实施方案中,选择指导分子的序列(正向重复序列和/或间隔区)以减少指导分子内的二级结构的程度。在一些实施方案中,当最佳折叠时,核酸靶向指导RNA的约或小于约75%、50%、40%、30%、25%、20%、15%、10%、5%、1%或更少的核苷酸参与自身互补碱基配对。最佳折叠可以通过任何合适的多核苷酸折叠算法来确定。一些程序是基于计算最小吉布斯自由能(Gibbs free energy)。一种此类算法的实例是如Zuker和Stiegler(Nucleic Acids Res.9(1981),133-148)所描述的mFold。另一个示例性折叠算法是维也纳大学(University of Vienna)的理论化学研究所(Institute for Theoretical Chemistry)开发的使用质心结构预测算法的在线网络服务器RNAfold(参见例如A.R.Gruber等人,2008,Cell 106(1):23-24;和PA Carr和GM Church,2009,Nature Biotechnology 27(12):1151-62)。

[0172] 在一些实施方案中,降低指导分子对RNA切割例如对通过Cas13的切割的易感性是令人感兴趣的。因此,在特定实施方案中,对指导分子进行调整以避免被Cas13或其他RNA切割酶切割。

[0173] 在某些实施方案中,指导分子包含非天然存在的核酸和/或非天然存在的核苷酸和/或核苷酸类似物和/或化学修饰。优选地,这些非天然存在的核酸和非天然存在的核苷酸位于指导序列之外。非天然存在的核酸可以包括例如天然和非天然存在的核苷酸的混合物。非天然存在的核苷酸和/或核苷酸类似物可在核糖、磷酸和/或碱基部分被修饰。在本发明的实施方案中,指导核酸包含核糖核苷酸和非核糖核苷酸。在一个这样的实施方案中,指导物包含一种或多种核糖核苷酸和一种或多种脱氧核糖核苷酸。在本发明的实施方案中,指导物包含一种或多种非天然存在的核苷酸或核苷酸类似物,诸如具有硫代磷酸酯键联的核苷酸、包含在核糖环的2'和4'碳原子之间的亚甲基桥的锁定核酸(LNA)或桥接核酸(BNA)。修饰的核苷酸的其他实例包括2'-O-甲基类似物、2'-脱氧类似物或2'-氟类似物。修饰的碱基的其他实例包括但不限于2-氨基嘌呤、5-溴-尿苷、假尿苷、肌苷、7-甲基鸟苷。指导RNA化学修饰的实例包括但不限于在一个或多个末端核苷酸处并入2'-O-甲基(M)、2'-O-甲基3'硫代磷酸酯(MS)、S-约束乙基(cEt),或2'-O-甲基3'硫代PACE(MSP)。此类化学修饰

的指导物与未修饰的指导物相比可以包含增加的稳定性和增加的活性，不过中靶对脱靶特异性不可预测。(参见Hendel,2015,Nat Biotechnol.33(9):985-9,doi:10.1038/nbt.3290,2015年6月29日在线发布,Ragdarm等人,0215,PNAS,E7110-E7111;Allerson等人,J.Med.Chem.2005,48:901-904;Bramsen等人,Front.Genet.,2012,3:154;Deng等人,PNAS,2015,112:11870-11875;Sharma等人,MedChemComm.,2014,5:1454-1471;Hendel等人,Nat.Biotechnol.(2015)33(9):985-989;Li等人,Nature Biomedical Engineering,2017,1,0066DOI:10.1038/s41551-017-0066)。在一些实施方案中，指导RNA的5'和/或3'端被包括荧光染料、聚乙二醇、胆固醇、蛋白质或检测标签在内的多种功能性部分修饰。(参见Kelly等人,2016,J.Biotech.233:74-83)。在某些实施方案中，指导物在结合至靶DNA的区域中包含核糖核苷酸，并在结合至Cas13的区域中包含一个或多个脱氧核糖核苷酸和/或核苷酸类似物。在本发明的实施方案中，将脱氧核糖核苷酸和/或核苷酸类似物并入工程化的指导结构(诸如但不限于茎环区和种子区)中。对于Cas13指导物，在某些实施方案中，修饰不在茎环区的5'柄(5'-handle)中。指导物的茎环区的5'柄中的化学修饰可能会废除其功能(参见Li等人,Nature Biomedical Engineering,2017,1:0066)。在某些实施方案中，指导物的至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、35个、40个、45个、50个或75个核苷酸经化学修饰。在一些实施方案中，指导物的3'或5'端的3-5个核苷酸经化学修饰。在一些实施方案中，在种子区中仅引入较小的修饰，诸如2'-F修饰。在某些实施方案中，在指导物的3'端引入2'-F修饰。在某些实施方案中，指导物的5'端和/或3'端的3至5个核苷酸用2'-O-甲基(M)、2'-O-甲基3'硫代磷酸酯(MS)、S-约束乙基(cEt)或2'-O-甲基3'硫代PACE(MSP)化学修饰。这样的修饰可以提高基因组编辑效率(参见Hendel等人,Nat.Biotechnol.(2015)33(9):985-989)。在某些实施方案中，指导物的所有磷酸二酯键被硫代磷酸酯(PS)取代以增强基因破坏的水平。在某些实施方案中，指导物的5'端和/或3'端的多于5个核苷酸用2'-O-Me、2'-F或S-约束乙基(cEt)进行化学修饰。这种化学修饰的指导物可以介导增强的基因破坏水平(参见Ragdarm等人,0215,PNAS,E7110-E7111)。在本发明的一个实施方案中，指导物被修饰成在其3'和/或5'端包含化学部分。这样的部分包括但不限于胺、叠氮化物、炔、硫代基、二苯并环辛炔(DBCO)或若丹明。在某些实施方案中，化学部分通过接头诸如烷基链缀合至指导物。在某些实施方案中，修饰的指导物的化学部分可用于将指导物附接至另一分子，诸如DNA、RNA、蛋白质或纳米粒子。这种化学修饰的指导物可用于识别或富集一般由CRISPR系统编辑的细胞(参见Lee等人,eLife,2017,6:e25312,DOI:10.7554)。

[0174] 在一些实施方案中，指导物包括修饰的Cas13 crRNA，其具有5'柄和指导区段，所述指导区段还包括种子区和3'端。在一些实施方案中，修饰的指导物可以与以下任一种Cas13结合使用:氨基酸球菌属种BV3L6 Cas13(AsCas13);土拉弗朗西斯菌新凶手亚种U112 Cas13(FnCas13);李斯特氏菌(L.bacterium)MC2017 Cas13(Lb3Cas13);解蛋白丁酸弧菌Cas13(BpCas13);帕库氏菌GWC2011_GWC2_44_17Cas13(PbCas13);异域菌门菌GW2011_GWA_33_10Cas13(PeCas13);稻田氏钩端螺旋体Cas13(LiCas13);史密斯氏菌属种SC_K08D17 Cas13(SsCas13);李斯特氏菌MA2020 Cas13(Lb2Cas13);狗口腔卟啉单胞菌Cas13(PcCas13);猕猴卟啉单胞菌Cas13(PmCas13);候选白蚁甲烷支原体Cas13(CMtCas13);挑剔

真杆菌Cas13(EeCas13)；牛眼莫拉氏菌237Cas13(MbCas13)；解糖胨普雷沃氏菌Cas13(PdCas13)；或李斯特氏菌ND2006 Cas13(LbCas13)。

[0175] 在一些实施方案中，对指导物的修饰是化学修饰、插入、缺失或裂解。在一些实施方案中，化学修饰包括但不限于并入2'-O-甲基(M)类似物、2'-脱氧类似物、2-硫代尿苷类似物，N6-甲基腺苷类似物、2'-氟类似物、2-氨基嘌呤、5-溴-尿苷、假尿苷(Ψ)、N1-甲基假尿苷(me1Ψ)、5-甲氧基尿苷(5moU)、肌苷、7-甲基鸟苷、2'-O-甲基3'硫代磷酸酯(MS)、S-约束乙基(cEt)、硫代磷酸酯(PS)或2'-O-甲基3'硫代PACE(MSP)。在一些实施方案中，指导物包含一种或多种硫代磷酸酯修饰。在某些实施方案中，指导物的至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或25个核苷酸经化学修饰。在某些实施方案中，种子区中的一个或多个核苷酸经化学修饰。在某些实施方案中，在3'端的一个或多个核苷酸经化学修饰。在某些实施方案中，5'柄中的核苷酸均未经化学修饰。在一些实施方案中，种子区中的化学修饰是次要修饰，诸如并入2'-氟类似物。在具体实施方案中，种子区的一个核苷酸被2'-氟类似物替代。在一些实施方案中，在3'端中的5至10个核苷酸经化学修饰。在Cas13 CrRNA的3'端处的此类化学修饰可改善Cas13活性(参见Li等人，Nature Biomedical Engineering，2017，1:0066)。在具体实施方案中，3'端中的1个、2个、3个、4个、5个、6个、7个、8个、9个或10个核苷酸被2'-氟类似物替代。在具体实施方案中，3'端中的1个、2个、3个、4个、5个、6个、7个、8个、9个或10个核苷酸被2'-O-甲基(M)类似物替代。

[0176] 在一些实施方案中，指导物的5'柄的环经修饰。在一些实施方案中，指导物的5'柄的环被修饰为具有缺失、插入、裂解或化学修饰。在某些实施方案中，修饰的环包含3个、4个或5个核苷酸。在某些实施方案中，环包含序列UCUU、UUUU、UAUU或UGUU。

[0177] 在一些实施方案中，指导分子与单独的非共价连接的序列(其可以是DNA或RNA)形成茎环。在特定实施方案中，首先使用标准亚磷酰胺合成方案合成形成指导物的序列(Herdewijn，P.，编辑，Methods in Molecular Biology Col 288，Oligonucleotide Synthesis:Methods and Applications，Humana Press，New Jersey(2012))。在一些实施方案中，可以使用本领域已知的标准方案将这些序列官能化成含有适于连接的官能团(Hermanson，G.T.，Bioconjugate Techniques，Academic Press(2013))。官能团的实例包括但不限于羟基、胺、羧酸、羧酸卤化物、羧酸活性酯、醛、羰基、氯代羰基、咪唑基羰基、肼基、氨基脲、硫代氨基脲、硫醇、马来酰亚胺、卤代烷基、磺酰基、烯丙基、炔丙基、二烯、炔和叠氮化物。一旦该序列被官能化，就可以在该序列与正向重复序列之间形成共价化学键或键联。化学键的实例包括但不限于基于以下的那些：氨基甲酸酯、醚、酯、酰胺、亚胺、脒、氨基三嗪、脲、二硫键、硫醚、硫酯、硫代磷酸酯、二硫代磷酸酯、磺酰胺、磺酸酯、砜(fulfone)、亚砜、脲、硫脲、酰肼、肟、三唑、光不稳定键联、C-C键形成基团(诸如狄尔斯-阿尔德环加成对(Diels-Alder cyclo-addition pair)或闭环复分解对(ring-closing metathesis pair))以及迈克尔反应对(Michael reaction pair)。

[0178] 在一些实施方案中，这些茎环形成序列可以经化学合成。在一些实施方案中，化学合成使用自动固相寡核苷酸合成机并利用2'-乙酰氧基乙基原酸酯(2'-ACE)(Scaringe等人，J.Am.Chem.Soc.(1998)120:11820-11821；Scaringe，Methods Enzymol.(2000)317:3-18)或2'-硫代氨基甲酸酯(2'-TC)化学品(Dellinger等人，J.Am.Chem.Soc.(2011)133:

11540-11546;Hendel等人,Nat.Biotechnol.(2015)33:985-989)。

[0179] 在某些实施方案中,指导分子(能够将Cas13导向至靶基因座)包含(1)能够与靶基因座杂交的指导序列和(2)tracr配对或正向重复序列,其中正向重复序列位于指导序列的上游(即5')。在特定实施方案中,Cas13指导序列的种子序列(即对于识别靶基因座处的序列和/或与其杂交而言至关重要的序列)大约在指导序列的前10个核苷酸之内。在特定实施方案中,Cas13为FnCas13,并且种子序列大约在指导序列的5'端的前5nt之内。

[0180] 在特定实施方案中,指导分子包含连接至正向重复序列的指导序列,其中正向重复序列包含一个或多个茎环或优化的二级结构。在特定实施方案中,正向重复序列的最小长度为16nt,并且具有单个茎环。在另外的实施方案中,正向重复序列的长度大于16nt,优选大于17nt,并且具有多于一个茎环或优化的二级结构。在特定实施方案中,指导分子包含连接至全部或部分天然正向重复序列的指导序列或由其组成。典型的V型Cas13指导分子包含(以3'至5'方向):指导序列、第一互补段("重复序列")、环(长度通常为4或5个核苷酸)、第二互补段(与重复序列互补的"抗重复序列"),以及poly A(通常为RNA中的poly U)尾巴(终止子)。在某些实施方案中,正向重复序列保留其天然架构并形成单个茎环。在特定实施方案中,指导物架构的某些方面可以例如通过特征的添加、减去或取代进行修饰,而指导物架构的某些其他方面得以保留。工程化指导分子修饰(包括但不限于插入、缺失和取代)的优选位置包括指导物末端和指导分子的在与Cas13蛋白和/或靶标复合时暴露的区域,例如正向重复序列的茎环。

[0181] 在特定实施方案中,茎含有互补X和Y序列,包含至少约4bp,但也涵盖具有更多(例如5个、6个、7个、8个、9个、10个、11个或12个)或更少(例如3个、2个)的碱基对的茎。因此,可以涵盖例如X2-10和Y2-10(其中X和Y代表任何互补的核苷酸组)。在一方面,由X和Y核苷酸组成的茎与环一起将在总体二级结构中形成完整的发夹;并且,这可能是有利的,并且碱基对的数量可以是形成完整发夹的任何数量。在一方面,只要保留整个指导分子的二级结构,就可以容忍任何互补的X:Y碱基配对序列(例如,就长度而言)。在一方面,连接由X:Y碱基对组成的茎的环,可以是不干扰指导分子的总体二级结构的长度相同(例如4或5个核苷酸)或更长的任何序列。在一方面,茎环可还包含例如MS2适体。在一方面,茎含有互补X和Y序列,包含约5-7bp,但也涵盖具有更多或更少的碱基对的茎。在一方面,涵盖非沃森-克里克(Watson-Crick)碱基配对,这种配对通常以其他方式保留该位置处茎环的架构。

[0182] 在特定实施方案中,将指导分子的天然发夹或茎环结构延伸或由延伸的茎环替代。已证实茎的延伸可以增强指导分子与CRISPR-Cas蛋白的组装(Chen等人Cell.(2013);155(7):1479-1491)。在特定实施方案中,茎环的茎延伸了至少1个、2个、3个、4个、5个或更多个互补碱基对(即对应于在指导分子中添加2个、4个、6个、8个、10个或更多个核苷酸)。在特定实施方案中,它们位于茎的端部,邻近茎环的环。

[0183] 在特定实施方案中,可以通过对指导分子的序列进行不影响其功能的轻微修饰来降低指导分子对RNA酶或表达降低的易感性。例如,在特定实施方案中,可以通过修饰指导分子序列中的推定Pol-III终止子(4个连续的U)来消除转录的过早终止,诸如U6 Pol-III的过早转录。当在指导分子的茎环中需要这种序列修饰时,优选通过碱基对翻转来确保。

[0184] 在优选的实施方案中,可以修饰正向重复序列使其包含一个或多个蛋白结合RNA适体。在特定实施方案中,可以包括一个或多个适体,诸如优化的二级结构的一部分。此类

适体可能能够结合如本文进一步详述的噬菌体外壳蛋白。

[0185] 在一些实施方案中,指导分子与包含至少一个待编辑的靶腺苷残基的靶DNA链形成双链体。指导RNA分子与靶DNA链杂交后,腺苷脱氨酶结合至双链体并催化DNA-RNA双链体内包含的一个或多个靶腺苷残基脱氨基。

[0186] 可以选择指导序列并且因此选择核酸靶向指导RNA以靶向任何靶核酸序列。靶序列可以是DNA。靶序列可以是基因组DNA。靶序列可以是线粒体DNA。

[0187] 在某些实施方案中,靶序列应缔合至PAM(原间隔区相邻基序)或PFS(原间隔区侧接序列或位点);即,由CRISPR复合物识别的短序列。取决于CRISPR-Cas蛋白的性质,应选择靶序列,使得其在DNA双链体中的互补序列(在本文中也称为非靶序列)在PAM的上游或下游。在CRISPR-Cas蛋白是Cas13蛋白的本发明的实施方案中,靶序列的互补序列在PAM的下游或3'。PAM的精确序列和长度要求因所使用的Cas13蛋白而异,但PAM通常是与原间隔区(即靶序列)相邻的2-5个碱基对序列。在下文中提供了用于不同Cas13直系同源物的天然PAM序列的实例,并且技术人员将能够鉴定与给定Cas13蛋白一起使用的另外的PAM序列。

[0188] 此外,例如,如Kleinstiver BP等人Engineered CRISPR-Cas9 nucleases with altered PAM specificities.Nature.2015年7月23日;523(7561):481-5.doi:10.1038/nature14592中针对Cas9所述,对PAM相互作用(PI)结构域的工程化可以允许对PAM特异性进行编程,提高靶位点识别的保真度,并提高CRISPR-Cas蛋白的多功能性。如本文进一步详述的,技术人员将理解可以采用类似方式修饰Cas13蛋白。

[0189] 在特定实施方案中,选择指导序列以确保脱氨酶在待脱氨基的腺嘌呤上的最佳效率。可以考虑靶链中腺嘌呤相对于Cas13切口酶的切割位点的位置。在特定实施方案中,令人感兴趣的是确保切口酶将作用于非靶链上邻近待脱氨基的腺嘌呤的位置处。例如,在特定实施方案中,Cas13切口酶切除非靶向链中在PAM下游的17个核苷酸(例如AsCas13、LbCas13)或PAM下游的18个核苷酸(例如FnCas13),并且令人感兴趣的是这样设计指导物:对应于待脱氨基的腺嘌呤的胞嘧啶位于指导序列中,在相应非靶链序列中切口酶切割位点上游或下游10bp内。

[0190] 在特定实施方案中,指导物是受护航的指导物。所谓"受护航的"是指将Cas13CRISPR-Cas系统或复合物或指导物递送至细胞内的选定时间或位置,从而在空间或时间上控制Cas13 CRISPR-Cas系统或复合物或指导物的活性。例如,可以通过对适体配体(诸如细胞表面蛋白或其他局部细胞组分)具有结合亲和力的护航性RNA适体序列控制Cas13 CRISPR-Cas系统或复合物或指导物的活性和目的地。或者,护航性适体可以例如对细胞上或细胞中的适体效应子作出反应,所述适体效应子诸如有瞬时效应子,诸如在特定时间施加至细胞的外部能源。

[0191] 受护航的Cas13 CRISPR-Cas系统或复合物具有指导分子,其具有被设计来改善指导分子的结构、架构、稳定性、基因表达或它们的任何组合的功能结构。这样的结构可包括适体。

[0192] 适体是可以例如使用一种称为指数富集的配体系统进化(SELEX;Tuerk C,Gold L:"Systematic evolution of ligands by exponential enrichment:RNA ligands to bacteriophage T4 DNA polymerase."Science 1990,249:505-510)的技术进行设计或选择以与其他配体紧密结合的生物分子。核酸适体可以例如选自随机序列寡核苷库,它们对

大范围的生物医学相关靶具有高结合亲和力和特异性,这揭示了适体的广泛治疗实用性(Keefe、Anthony D.、Supriya Pai和Andrew Ellington."Aptamers as therapeutics."Nature Reviews Drug Discovery 9.7(2010):537-550)。这些特征还揭示了适体作为药物递送媒介物的广泛用途(Levy-Nissenbaum,Etgar等人"Nanotechnology and aptamers:applications in drug delivery."Trends in biotechnology 26.8(2008):442-449;和Hicke BJ,Stephens AW."Escort aptamers:a delivery service for diagnosis and therapy."J Clin Invest 2000,106:923-928.)。还可以构建充当分子开关、通过改变特性来响应问询(que)的适体,诸如结合荧光团以模拟绿色荧光蛋白活性的RNA适体(Paige、Jeremy S.、Karen Y.Wu和Samie R.Jaffrey."RNA mimics of green fluorescent protein."Science 333.6042(2011):642-646)。先前还已提出适体可以用作靶向的siRNA治疗性递送系统的组分,例如靶向细胞表面蛋白(Zhou,Jiehua和John J.Rossi."Aptamer-targeted cell-specific RNA interference."Silence 1.1(2010):4)。

[0193] 因此,在特定实施方案中,例如通过一种或多种适体对指导分子进行修饰,所述一种或多种适体被设计来改善指导分子的递送,包括递送穿过细胞膜、到达细胞内隔室或进入细胞核。加上所述一种或多种适体或无所述一种或多种适体,这样的结构可包括一个或多个部分,以便使指导分子可递送、可诱导或响应于选定的效应子。因此,本发明包括响应于正常或病理生理条件的指导分子,所述生理条件包括但不限于pH、低氧、O2浓度、温度、蛋白质浓度、酶浓度、脂质结构、曝光、机械破坏(例如超声波)、磁场、电场或电磁辐射。

[0194] 诱导型系统的光响应性可以经由隐花色素-2和CIB1的激活和结合来实现。蓝光刺激诱导隐花色素-2中的激活的构象变化,导致其结合配偶体CIB1募集。这种结合是快速且可逆的,在脉冲刺激后的<15秒内达到饱和,并在刺激结束后的<15分钟内恢复至基线。这些快速的结合动力学使得系统暂时仅受转录/翻译和转录物/蛋白质降解的速度限制,而不受诱导剂的吸收和清除的限制。隐花色素-2的激活还是高度敏感的,使得可以使用低光强度刺激并减轻了光毒性的风险。此外,在诸如完整的哺乳动物脑的情形下,可变的光强度可用于控制受激区域的大小,从而获得比单独的载体递送所能提供的精度更高的精度。

[0195] 本发明考虑了诸如电磁辐射、声能或热能的能源来诱导指导物。有利地,电磁辐射是可见光的组分。在优选的实施方案中,光是波长为约450至约495nm的蓝光。在特别优选的实施方案中,波长为约488nm。在另一个优选的实施方案中,光刺激是经由脉冲实现的。光功率可以在大约0-9mW/cm2的范围内。在优选的实施方案中,每15秒低至0.25秒的刺激范式应该会导致最大的激活。

[0196] 化学或能量敏感型指导物在由于化学源的结合或能量而被诱导时可能会发生构象变化,使其成为指导物并具有Cas13 CRISPR-Cas系统或复合物功能。本发明可涉及施加化学源或能量以具有指导物功能和Cas13 CRISPR-Cas系统或复合物功能;并且任选地进一步确定基因组基因座的表达已改变。

[0197] 此化学诱导型系统有几种不同的设计:1.由脱落酸(ABA)可诱导的基于ABI-PYL的系统(参见例如http://stke.sciencemag.org/cgi/content/abstract/sigtrans;4/164/rs2);2.由雷帕霉素可诱导的基于FKBP-FRB系统(或基于雷帕霉素的相关化学品)(参见例如http://www.nature.com/nmeth/journal/v2/n6/full/nmeth763.html);3.由赤霉素(GA)可诱导的基于GID1-GAI的系统(参见例如http://www.nature.com/nchembio/

journal/v8/n5/full/nchembio.922.html)。

[0198] 化学诱导型系统可以是由4-羟基他莫昔芬(4OHT)可诱导的基于雌激素受体(ER)的系统(参见例如http://www.pnas.org/content/104/3/1027.abstract)。雌激素受体的一种称为ERT2的突变配体结合结构域在与4-羟基他莫昔芬结合后易位到细胞的细胞核中。在本发明的另外的实施方案中,任何核受体、甲状腺激素受体、视黄酸受体、雌激素受体、雌激素相关受体、糖皮质激素受体、孕激素受体、雄激素受体的任何天然存在或工程化的衍生物都可以用于与基于ER的诱导型系统类似的诱导型系统。

[0199] 另一种诱导型系统是基于使用由能量、热或无线电波可诱导的基于瞬时受体电位(TRP)离子通道的系统进行的设计(参见例如http://www.sciencemag.org/content/336/6081/604)。这些TRP家族蛋白响应于不同的刺激,包括光和热。当这种蛋白质被光或热激活时,离子通道将打开并允许诸如钙的离子进入质膜。这种离子涌流将与连接至多肽(包括指导物和Cas13 CRISPR-Cas复合物或系统的其他组分)的细胞内离子相互作用配偶体结合,并且所述结合将诱导所述多肽的亚细胞定位的变化,从而使得整个多肽进入细胞的细胞核。一旦进入细胞核,所述指导蛋白和Cas13 CRISPR-Cas复合物的其他组分就将呈活性状态并调节细胞中的靶基因表达。

[0200] 尽管光激活可以是有利的实施方案,但是有时对于光可能不穿透皮肤或其他器官的体内应用而言可能是尤为不利的。在这种情况下,可以考虑其他具有类似效果的能量激活方法,特别是电场能和/或超声。

[0201] 优选地在体内条件下,使用约1V/cm至约10k V/cm的一个或多个电脉冲,基本上如本领域中所述施加电场能。代替脉冲或加上脉冲,可以采用连续方式递送电场。可以施加电脉冲,持续1微秒与500毫秒之间,优选地1微秒与100毫秒之间。可以连续地或以脉冲方式施加电场,持续约5分钟。

[0202] 如本文所用,“电场能”是细胞暴露于其中的电能。在体内条件下,电场的强度优选为约1V/cm至约10kV/cm或更大(参见WO97/49450)。

[0203] 如本文所用,术语“电场”包括在可变电容和电压下的一个或多个脉冲,并且包括指数波和/或方形波和/或调制波和/或调制方形波形式。对电场和电的提及应视为包括对细胞环境中电位差的存在的提及。如本领域中已知的,可以通过静电、交流电(AC)、直流电(DC)等来建立这样的环境。电场可以是均匀的、不均匀的或其他方式的,并且可以以时间依赖性方式改变强度和/或方向。

[0204] 还可以按任何顺序和任何组合单次或多次施加电场,以及单次或多次施加超声。超声和/或电场可以作为单次或多次连续施加或作为脉冲来递送(脉冲式递送)。

[0205] 电穿孔已用于体外和体内程序中,以将异物引入活细胞中。在体外应用中,首先将活细胞样品与目标剂混合,接着将它们放置在电极(诸如平行板)之间。接着,电极向细胞/植入物混合物施加电场。执行体外电穿孔的系统的实例包括Electro Cell Manipulator ECM600产品和Electro Square Porator T820,这两者均由Genetronics,Inc的BTX分部制造(参见美国专利号5,869,326)。

[0206] 已知的电穿孔技术(体外和体内)都通过向位于治疗区域周围的电极施加短暂的高压脉冲来发挥作用。电极之间产生的电场使细胞膜暂时变为多孔的,此时目标剂进入细胞。在已知的电穿孔应用中,此电场包括持续约100微秒的大约1000V/cm的单个方形波脉

冲。这样的脉冲可以例如在Electro Square Porator T820的已知应用中产生。

[0207] 在体外条件下,电场的强度优选为约1V/cm至约10kV/cm。因此,电场的强度可以为1V/cm、2V/cm、3V/cm、4V/cm、5V/cm、6V/cm、7V/cm、8V/cm、9V/cm、10V/cm、20V/cm、50V/cm、100V/cm、200V/cm、300V/cm、400V/cm、500V/cm、600V/cm、700V/cm、800V/cm、900V/cm、1kV/cm、2kV/cm、5kV/cm、10kV/cm、20kV/cm、50kV/cm或更大。在体外条件下,更优选为约0.5kV/cm至约4.0kV/cm。在体内条件下,电场的强度优选为约1V/cm至约10kV/cm。然而,当递送至靶位点的脉冲数量增加时,电场强度可能降低。因此,设想以较低的场强脉冲式递送电场。

[0208] 优选地,采用多个脉冲的形式,诸如具有相同强度和电容的双脉冲或具有变化强度和/或电容的顺序脉冲来施加电场。如本文所用,术语“脉冲”包括在可变电容和电压下的一个或多个电脉冲,并且包括指数波和/或方形波和/或调制波/方形波形式。

[0209] 优选地,将电脉冲作为选自指数波形式、方形波形式、调制波形式和调制方形波形式的波形递送。

[0210] 优选的实施方案采用低压直流电。因此,申请人公开了以1V/cm与20V/cm之间的场强向细胞、组织或组织块施加电场,持续时间为100毫秒或更长,优选为15分钟或更长。

[0211] 有利地,以约0.05W/cm2至约100W/cm2的功率水平施用超声。可以使用诊断性超声或治疗性超声,或它们的组合。

[0212] 如本文所用,术语“超声”是指一种由机械振动组成的能量形式,所述机械振动的频率特别高以至于超出人类的听觉范围。超声频谱的频率下限通常可以取为约20kHz。大多数诊断性超声应用采用1至15MHz’的频率(Ultrasonics in Clinical Diagnosis,P.N.T.Wells,编辑,第二版,出版社Churchill Livingstone[Edinburgh,London&NY,1977])。

[0213] 在诊断性和治疗性应用中皆已使用超声。当用作诊断性工具(诊断性超声)时,通常在高达约100mW/cm2的能量密度下使用超声(FDA推荐),但也使用过高达750mW/cm2的能量密度。在物理疗法中,通常使用高达约3至4W/cm2范围内的超声作为能源(WHO推荐)。在其他治疗性应用中,可以在短时间内采用更高强度的超声,例如100W/cm至1kW/cm2(或甚至更高)的HIFU。在本说明书中使用的术语“超声”旨在涵盖诊断性超声、治疗性超声和聚焦超声。

[0214] 聚焦超声(FUS)允许在不使用侵入式探头的情况下递送热能(参见Morocz等人1998,Journal of Magnetic Resonance Imaging,第8卷,第1期,第136-142页。聚焦超声的另一种形式是高强度聚焦超声(HIFU),Moussatov等人,Ultrasonics(1998),第36卷,第8期,第893-900页以及TranHuuHue等人,Acustica(1997),第83卷,第6期,第1103-1106页中对此进行了综述。

[0215] 优选地,采用诊断性超声和治疗性超声的组合。但是,该组合并非旨在进行限制,而是本领域技术人员将理解可以使用超声的任何多种组合。另外,能量密度、超声频率和暴露时间是可以改变的。

[0216] 优选地,超声能源暴露的功率密度为约0.05至约100Wcm-2。甚至更优选地,超声能源暴露的功率密度为约1至约15Wcm-2。

[0217] 优选地,超声能源暴露的频率为约0.015至约10.0MHz。更优选地,超声能源暴露的频率为约0.02至约5.0MHz或约6.0MHz。最优选地,以3MHz的频率施加超声。

[0218] 优选地，暴露持续约10毫秒至约60分钟的时段。优选地，暴露持续约1秒至约5分钟的时段。更优选地，施加超声持续约2分钟。然而，取决于有待破坏的特定靶细胞，暴露可以持续更长的持续时间，例如持续15分钟。

[0219] 有利地，将靶组织暴露于超声能源，超声能源的声功率密度为约0.05Wcm-2至约10Wcm-2，频率在约0.015至约10MHz的范围内(参见WO 98/52609)。但是替代方案也是可能的，例如超声能源暴露的声功率密度高于100Wcm-2，但持续缩短的时间段，例如1000Wcm-2持续毫秒范围或更小的时段。

[0220] 优选地，超声施加呈多个脉冲的形式;因此，可以采用任何组合的连续波和脉冲波(脉冲式超声递送)。例如，可以施加连续波超声，之后施加脉冲波超声，反之亦然。可以采用任何顺序和组合将其重复任意次数。可以在连续波超声的背景下施加脉冲波超声，并且可以使用任何组数的任何数量的脉冲。

[0221] 优选地，超声可包括脉冲波超声。在高度优选的实施方案中，以0.7Wcm-2或1.25Wcm-2的功率密度以连续波形式施加超声。如果使用了脉冲超声波，则可以采用更高的功率密度。

[0222] 超声的使用是有利的，因为像光一样，超声可以精确地聚焦在靶标上。此外，超声是有利的，因为与光不同，超声可以更深地聚焦到组织中。因此，它更适于完整组织穿透(诸如但不限于肝叶)或完整器官(诸如但不限于整个肝脏或整个肌肉，诸如心脏)疗法。另一个重要的优点是超声是非侵入式刺激，可用于多种多样的诊断性和治疗性应用。举例来说，超声在医学成像技术以及骨科疗法中是众所周知的。此外，适于向受试者脊椎动物施加超声的仪器是广泛可得的，并且它们使用在本领域中是众所周知的。

[0223] 在特定实施方案中，通过二级结构修饰指导分子以增加CRISPR-Cas系统的特异性，并且所述二级结构可以防御核酸外切酶活性并允许对指导序列进行5'添加(在本文中也称为受保护的指导分子)。

[0224] 在一方面，本发明提供了将“保护性RNA”与指导分子的序列杂交的方法，其中“保护性RNA”是与指导分子的3'端互补的RNA链，从而产生部分双链的指导RNA。在本发明的实施方案中，用完全互补的保护性序列保护错配的碱基(即，指导分子中不形成指导序列的一部分的碱基)降低了靶DNA结合至3'端处的错配碱基对的可能性。在本发明的特定实施方案中，在指导分子内还可以存在包含延伸长度的其他序列，使得指导物在指导分子内包含保护性序列。该“保护性序列”确保指导分子除“暴露的序列”(包含与靶序列杂交的指导序列的一部分)之外还包括“受保护的序列”。在特定实施方案中，通过保护性指导物的存在将指导分子修饰为包含二级结构(诸如发夹)。有利地，存在三个或四个至三十个或更多个(例如约10个或更多个)具有与受保护序列、指导序列或两者互补的连续碱基对。有利的是，受保护部分不妨碍CRISPR-Cas系统与其靶相互作用的热力学。通过提供包括部分双链的指导分子在内的这种延伸，指导分子被认为是受保护的并且改善了CRISPR-Cas复合物的特异性结合，同时维持特定活性。

[0225] 在特定实施方案中，使用了截短的指导物(tru-指导物)，即包含相对于典型的指导序列长度，长度被截短的指导序列的指导分子。如Nowak等人(Nucleic Acids Res(2016)44(20):9555-9564)所述，此类指导物可允许催化活性CRISPR-Cas酶结合其靶标，而不切割靶DNA。在特定实施方案中，使用了截短的指导物，所述截短的指导物允许靶标的结合，但仅

保留了CRISPR-Cas酶的切口酶活性。

[0226] Crispr-Cas酶

[0227] 未经修饰形式的CRISPR-Cas蛋白是一种催化活性蛋白。这意味着在形成核酸靶向复合物(包含与靶序列中或附近(例如,离靶序列1、2、3、4、5、6、7、8、9、10、20、50、或更多个碱基对之内)的一条或两条DNA链杂交的指导RNA)后,所述靶序列被修饰(例如,被切割)。如本文所用,术语“与目标靶基因座相缔合的一个或多个序列”是指在靶序列附近的序列(例如,离靶序列1、2、3、4、5、6、7、8、9、10、20、50、或更多个碱基对之内,其中所述靶序列被包含在目标靶基因座中)。未修饰的催化活性Cas13蛋白产生交错切口,由此切割位点通常位于靶序列内。更特别地,交错切口通常是PAM远端的13-23个核苷酸。在特定实施方案中,非靶链上的切口是PAM下游的17个核苷酸(即,在PAM下游的核苷酸17与18之间),而靶链上的切口(即,与指导序列杂交的链)又出现了离PAM互补序列较远的4个核苷酸(这是3’链上PAM互补序列上游的21个核苷酸,或在PAM互补序列上游的核苷酸21与22之间)。

[0228] 在根据本发明的方法中,优选地相对于相应的野生型酶使CRISPR-Cas蛋白突变,使得所述突变的CRISPR-Cas蛋白缺乏切割含有靶序列的靶基因座的一条或两条DNA链的能力。在特定实施方案中,使Cas13蛋白的一个或多个催化结构域突变以产生仅切割靶序列的一条DNA链的突变的Cas蛋白。

[0229] 在特定实施方案中,可以相对于相应的野生型酶使CRISPR-Cas蛋白突变,使得所述突变的CRISPR-Cas蛋白基本上缺乏所有的DNA切割活性。在一些实施方案中,当突变酶的切割活性为所述酶的非突变形式的核酸切割活性的约不超过25%、10%、5%、1%、0.1%、0.01%或更低时,认为CRISPR-Cas蛋白基本上缺乏所有的DNA和/或RNA切割活性;一个实例可以是当突变形式的核酸切割活性与非突变形式相比是零或可忽略不计时。

[0230] 在本文提供的方法的某些实施方案中,CRISPR-Cas蛋白是仅切割一条DNA链的突变CRISPR-Cas蛋白,即切口酶。更特别地,在本发明的情形中,所述切口酶确保在非靶序列(即,在靶序列的相反DNA链上并且是PAM序列的3’的序列)内的切割。作为进一步指导且非限制性地,来自氨基酸球菌属种的Cas13的Nuc结构域中的精氨酸至丙氨酸取代(R1226A)将Cas13从切割两条链的核酸酶转化为切口酶(切割单条链)。本领域技术人员将理解,在酶不是AsCas13的情况下,可以在相应位置的残基处形成突变。在特定实施方案中,Cas13是FnCas13,并且突变是在位置R1218处的精氨酸上。在特定实施方案中,Cas13是LbCas13,并且突变是在位置R1138处的精氨酸上。在特定实施方案中,Cas13是MbCas13,并且突变是在位置R1293处的精氨酸上。

[0231] 在本文所提供的方法的某些实施方案中,CRISPR-Cas蛋白具有降低的催化活性或没有催化活性。在CRISPR-Cas蛋白是Cas13蛋白的情况下,突变可包括但不限于催化性RuvC样结构域中的一个或多个突变,诸如D908A或E993A(依据AsCas13中的位置)。

[0232] 在一些实施方案中,当突变酶的DNA切割活性为所述酶的非突变形式的DNA切割活性的约不超过25%、10%、5%、1%、0.1%、0.01%或更低时,认为CRISPR-Cas蛋白基本上缺乏所有的DNA切割活性;一个实例可以是当突变形式的DNA切割活性与非突变形式相比为零或可忽略不计时。在这些实施方案中,将CRISPR-Cas蛋白用作通用DNA结合蛋白。突变可以是人工引入的突变或者功能获得或功能丧失突变。

[0233] 除以上所述的突变外,还可以另外对CRISPR-Cas蛋白进行修饰。如本文所用,关于

CRISPR-Cas蛋白的术语“修饰的”通常是指与衍生其的野生型Cas蛋白相比，CRISPR-Cas蛋白具有一个或多个修饰或突变(包括点突变、截短、插入、缺失、嵌合体、融合蛋白等)。所谓衍生的，是指在与野生型酶具有高度序列同源性的意义上，衍生酶主要基于野生型酶，但是已经以本领域已知或如本文所述的某种方式对衍生酶进行了突变(修饰)。

[0234] 对CRISPR-Cas蛋白的另外的修饰可能会或可能不会导致功能改变。举例来说，且特别地就CRISPR-Cas蛋白而言，不导致功能改变的修饰包括例如针对表达到特定宿主中进行密码子优化，或向核酸酶提供特定标记物(例如用于可视化)。可能导致功能改变的修饰还可能包括突变，包括点突变、插入、缺失、截短(包括裂解核酸酶)等。融合蛋白可包括但不限于例如与异源结构域或功能结构域(例如定位信号、催化结构域等)形成的融合物。在某些实施方案中，可以组合各种不同的修饰(例如，具有催化活性的突变核酸酶进一步融合至功能结构域(例如)以诱导DNA甲基化；或另一种核酸修饰，如包括但不限于断裂(例如通过不同的核酸酶(结构域))、突变、缺失、插入、替代、连接、消化、断裂或重组)。如本文所用，“改变的功能性”包括但不限于改变的特异性(例如改变的靶标识别、增加的(例如“增强的”Cas蛋白)或降低的特异性，或改变的PAM识别)、改变的活性(例如增加的或降低的催化活性，包括无催化活性的核酸酶或切口酶)和/或改变的稳定性(例如与去稳定结构域融合)。合适的异源结构域包括但不限于核酸酶、连接酶、修复蛋白、甲基转移酶、(病毒)整合酶、重组酶、转座酶、argonaute、胞苷脱氨酶、反转录子、II族内含子、磷酸酶、磷酸化酶、磺酰化酶(sulpfurylase)、激酶、聚合酶、核酸外切酶等。所有这些修饰的实例在本领域中都是已知的。将理解的是，如本文所提及的“修饰的”核酸酶，且特别地“修饰的”Cas或“修饰的”CRISPR-Cas系统或复合物优选地仍具有与多核酸相互作用或结合的能力(例如与指导分子复合)。这种修饰的Cas蛋白可与如本文所述的脱氨酶蛋白或其活性结构域组合。

[0235] 在某些实施方案中，CRISPR-Cas蛋白可包含一种或多种使活性和/或特异性增强的修饰，例如包括使靶向或非靶向链稳定的突变残基(例如eCas9；“Rationally engineered Cas9 nucleases with improved specificity”，Slaymaker等人(2016)，Science，351(6268):84-88，以引用方式整体并入本文)。在某些实施方案中，工程化CRISPR蛋白的改变或修饰的活性包括增加的靶向效率或减少的脱靶结合。在某些实施方案中，工程化CRISPR蛋白的改变的活性包括修改的切割活性。在某些实施方案中，改变的活性包括对靶多核苷酸基因座的增加的切割活性。在某些实施方案中，改变的活性包括对靶多核苷酸基因座的降低的切割活性。在某些实施方案中，改变的活性包括对脱靶多核苷酸基因座的降低的切割活性。在某些实施方案中，修饰的核酸酶的改变的或修改的活性包括改变的解旋酶动力学。在某些实施方案中，修饰的核酸酶包含改变蛋白质与包含RNA的核酸分子(在Cas蛋白的情况下)、或靶多核苷酸基因座的链、或脱靶多核苷酸的链的缔合的修饰。在本发明的一方面，工程化CRISPR蛋白包含改变CRISPR复合物的形成的修饰。在某些实施方案中，改变的活性包括对脱靶多核苷酸基因座的增加的切割活性。因此，在某些实施方案中，相较于脱靶多核苷酸基因座，对靶多核苷酸基因座的特异性增加。在其他实施方案中，相较于脱靶多核苷酸基因座，对靶多核苷酸基因座的特异性降低。在某些实施方案中，突变导致脱靶效应(例如切割或结合特性、活性或动力学)降低，诸如在Cas蛋白的情况下，例如导致对靶标与指导RNA之间的错配的耐受性降低。其他突变可能导致脱靶效应(例如切割或结合特性、活性或动力学)增加。其他突变可能导致中靶效应(例如切割或结合特性、活性或

动力学)增加或降低。在某些实施方案中,突变引起改变的(例如增加或降低的)解旋酶活性、功能性核酸酶复合物(例如CRISPR-Cas复合物)的缔合或形成。在某些实施方案中,如上所述,突变导致PAM识别改变,即相较于未修饰的Cas蛋白,可能(另外地或替代地)识别不同的PAM。为了增强特异性,特别优选的突变包括带正电的残基和/或(进化的)保守的残基,诸如保守的带正电残基。在某些实施方案中,此类残基可被突变为不带电荷的残基,诸如丙氨酸。

[0236] 碱基编辑

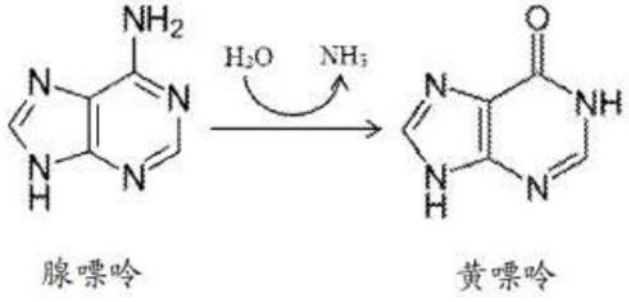
[0237] 在一些实施方案中,所述系统可以用于碱基编辑。例如,本文所公开的系统包含靶向组分和碱基编辑组分。所述靶向组分的功能是将碱基编辑组分特异性地靶向其中一个或多个核苷酸有待被编辑的靶核苷酸序列。在一些实例中,靶成分可以是无催化活性的Cas13效应蛋白(例如dCas13)。接着,碱基编辑组分可以催化化学反应,以将靶序列中的第一核苷酸转化为第二核苷酸。例如,碱基编辑器可以催化腺嘌呤的转化,使得其被细胞的转录或翻译机械装置读取为鸟嘌呤,反之亦然。同样地,碱基编辑组分可以催化胞苷转化为尿嘧啶,反之亦然。在某些示例性实施方案中,可以通过以诸如腺嘌呤脱氨酶或胞苷脱氨酶的已知碱基编辑器开始衍生碱基编辑器,并且使用诸如定向进化的方法对其进行修饰以衍生新的功能。定向进化技术在本领域中是已知的,并且可以包括在WO 2015/184016“遗传排列的高通量组装(High-Throughput Assembly of Genetic Permuatations)”中描述的那些。

[0238] 在某些示例性实施方案中,结合组分可以是DNA结合蛋白或其功能结构域,或者是RNA结合结构域或其功能结构域。结合组分可以结合在靶基因座处或附近的序列、基序或结构特征。结构特征可以包括发夹、四环或核酸的其他二级结构特征。如本文所用,“邻近”意指腺苷脱氨酶在其中可以完成其碱基编辑功能的靶基因座的距离和/或取向内。在某些示例性实施方案中,其结合组分可以结合指导分子的序列、基序或结构序列。指导分子包含与目标靶基因座杂交并由此将腺苷脱氨酶引导至所述靶基因座的序列。

[0239] 腺苷脱氨酶

[0240] 在某些示例性实施方案中,碱基编辑组分是腺苷脱氨酶。腺苷脱氨酶被靶向组分募集至靶基因座。在某些示例性实施方案中,腺苷脱氨酶连接至RNA结合蛋白或其功能片段,或DNA结合蛋白或其功能片段连接。

[0241] 如本文所用,术语“腺苷脱氨酶”或“腺苷脱氨酶蛋白”是指能够催化如下所示的将腺嘌呤(或分子的腺嘌呤部分)转化为次黄嘌呤(或分子的次黄嘌呤部分)的水解脱氨反应的蛋白质、多肽或蛋白质或多肽的一个或多个功能结构域。在一些实施方案中,含腺嘌呤的分子是腺苷(A),而含次黄嘌呤的分子是肌苷(I)。含腺嘌呤的分子可以是脱氧核糖核酸(DNA)或核糖核酸(RNA)。

[0242] 

[0243] 根据本公开，可以与本公开结合使用的腺苷脱氨酶包括但不限于称为作用于RNA的腺苷脱氨酶(ADAR)的酶家族成员、称为作用于tRNA的腺苷脱氨酶(ADAT)的酶家族成员，以及其他含腺苷脱氨酶结构域(ADAD)的家族成员。根据本公开，腺苷脱氨酶能够靶向RNA/DNA和RNA双链体中的腺嘌呤。实际上，Zheng等人(Nucleic Acids Res.2017,45(6):3369-3377)证明ADAR可以在RNA/DNA和RNA/RNA双链体上实现腺苷到肌苷的编辑反应。在特定实施方案中，如下文所详述，已对腺苷脱氨酶进行了修饰以增强其编辑RNA/DNAn RNA双链体中的DNA的能力。

[0244] 在一些实施方案中，腺苷脱氨酶来源于一种或多种后生动物种类，包括但不限于哺乳动物、鸟、青蛙、鱿鱼、鱼、蝇和蠕虫。在一些实施方案中，腺苷脱氨酶是人类、鱿鱼或果蝇腺苷脱氨酶。

[0245] 在一些实施方案中，腺苷脱氨酶是人类ADAR，包括hADAR1、hADAR2、hADAR3。在一些实施方案中，腺苷脱氨酶是秀丽隐杆线虫(Caenorhabditis elegans)ADAR蛋白，包括ADR-1和ADR-2。在一些实施方案中，腺苷脱氨酶是果蝇ADAR蛋白，包括dAdar。在一些实施方案中，腺苷脱氨酶是鱿鱼近海长鳍鱿鱼(Loligo pealeii)ADAR蛋白，包括sqADAR2a和sqADAR2b。在一些实施方案中，腺苷脱氨酶是人类ADAT蛋白。在一些实施方案中，腺苷脱氨酶是果蝇ADAT蛋白。在一些实施方案中，腺苷脱氨酶是人类ADAD蛋白，包括TENR(hADAD1)和TENRL(hADAD2)。

[0246] 在一些实施方案中，腺苷脱氨酶蛋白识别双链核酸底物中的一个或多个靶腺苷残基并将其转化为一个或多个肌苷残基。在一些实施方案中，双链核酸底物是RNA-DNA杂合双链体。在一些实施方案中，腺苷脱氨酶蛋白识别双链底物上的结合窗口。在一些实施方案中，结合窗口含有至少一个靶腺苷残基。在一些实施方案中，结合窗口在约3bp至约100bp的范围内。在一些实施方案中，结合窗口在约5bp至约50bp的范围内。在一些实施方案中，结合窗口在约10bp至约30bp的范围内。在一些实施方案中，结合窗口为约1bp、2bp、3bp、5bp、7bp、10bp、15bp、20bp、25bp、30bp、40bp、45bp、50bp、55bp、60bp、65bp、70bp、75bp、80bp、85bp、90bp、95bp或100bp。

[0247] 在一些实施方案中，腺苷脱氨酶蛋白包含一个或多个脱氨酶结构域。不意欲受理论的束缚，预期脱氨酶结构域用于识别双链核酸底物中所含的一个或多个靶腺苷(A)残基并将其转化为一个或多个肌苷(I)残基。在一些实施方案中，脱氨酶结构域包含活性中心。在一些实施方案中，活性中心包含锌离子。在一些实施方案中，在A至I编辑过程中，靶腺苷残基处的碱基配对被破坏，并且靶腺苷残基从双螺旋中“翻转”出来以变得可被腺苷脱氨酶接近。在一些实施方案中，在活性中心内或附近的氨基酸残基与靶腺苷残基5’端的一个或多个核苷酸相互作用。在一些实施方案中，在活性中心内或附近的氨基酸残基与靶腺苷残基3’端的一个或多个核苷酸相互作用。在一些实施方案中，在活性中心内或附近的氨基酸残基进一步与相反链上与靶腺苷残基互补的核苷酸相互作用。在一些实施方案中，氨基酸残基与核苷酸的2’羟基形成氢键。

[0248] 在一些实施方案中，腺苷脱氨酶包含人类ADAR2全蛋白(hADAR2)或其脱氨酶结构域(hADAR2-D)。在一些实施方案中，腺苷脱氨酶是与hADAR2或hADAR2-D同源的ADAR家族成员。

[0249] 特别地，在一些实施方案中，同源ADAR蛋白是人类ADAR1(hADAR1)或其脱氨酶结构

域(hADAR1-D)。在一些实施方式中,hADAR1-D的甘氨酸1007对应于甘氨酸487hADAR2-D,而hADAR1-D的谷氨酸1008对应于hADAR2-D的谷氨酸488。

[0250] 在一些实施方案中,腺苷脱氨酶包含hADAR2-D的野生型氨基酸序列。在一些实施方案中,腺苷脱氨酶包含在hADAR2-D序列中的一个或多个突变,这样hADAR2-D的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0251] hADAR1和hADAR2蛋白的某些突变已描述于Kuttan等人,Proc Natl Acad Sci U S A.(2012)109(48):E3295-304;Want等人ACS Chem Biol.(2015)10(11):2512-9;和Zheng等人Nucleic Acids Res.(2017)45(6):3369-337,所述文献各自以引用方式整体并入本文。

[0252] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的甘氨酸336处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置336处的甘氨酸残基被天冬氨酸残基替代(G336D)。

[0253] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的甘氨酸487处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置487处的甘氨酸残基被具有相对小的侧链的非极性氨基酸残基替代。例如,在一些实施方案中,位置487处的甘氨酸残基被丙氨酸残基替代(G487A)。在一些实施方案中,位置487处的甘氨酸残基被缬氨酸残基替代(G487V)。在一些实施方案中,位置487处的甘氨酸残基被具有相对大的侧链的氨基酸残基取代。在一些实施方案中,位置487处的甘氨酸残基被精氨酸残基替代(G487R)。在一些实施方案中,位置487处的甘氨酸残基被赖氨酸残基替代(G487K)。在一些实施方案中,位置487处的甘氨酸残基被色氨酸残基替代(G487W)。在一些实施方案中,位置487处的甘氨酸残基被酪氨酸残基替代(G487Y)。

[0254] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的谷氨酸488处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置488处的谷氨酸残基被谷氨酰胺残基替代(E488Q)。在一些实施方案中,位置488处的谷氨酸残基被组氨酸残基替代(E488H)。在一些实施方案中,位置488处的谷氨酸残基被精氨酸残基替代(E488R)。在一些实施方案中,位置488处的谷氨酸残基被赖氨酸残基替代(E488K)。在一些实施方案中,位置488处的谷氨酸残基被天冬酰胺残基替代(E488N)。在一些实施方案中,位置488处的谷氨酸残基被丙氨酸残基替代(E488A)。在一些实施方案中,位置488处的谷氨酸残基被甲硫氨酸残基替代(E488M)。在一些实施方案中,位置488处的谷氨酸残基被丝氨酸残基替代(E488S)。在一些实施方案中,位置488处的谷氨酸残基被苯丙氨酸残基替代(E488F)。在一些实施方案中,位置488处的谷氨酸残基被赖氨酸残基替代(E488L)。在一些实施方案中,位置488处的谷氨酸残基被色氨酸残基替代(E488W)。

[0255] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的苏氨酸490处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置490处的苏氨酸残基被半胱氨酸残基替代(T490C)。。在一些实施方案中,位置490处的苏氨酸残基被丝氨酸残基替代(T490S)。在一些实施方案中,位置490处的苏氨酸残基被丙氨酸残基替代(T490A)。在一些实施方案中,位置490处的苏氨酸残基被苯丙氨酸残基替代(T490F)。在一些实施方案中,位置490处的苏氨酸残基被酪氨酸残基替代(T490Y)。在一些实施方案中,位置490处的苏氨酸残基被丝氨酸残基替代(T490R)。在一些实施方案中,位置490处的苏氨酸残基被丙氨酸残基替代(T490K)。在一些实施方案中,位置490处的苏氨酸残基被苯丙氨酸残基替代

(T490P)。在一些实施方案中,位置490处的苏氨酸残基被酪氨酸残基替代(T490E)。

[0256] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的缬氨酸493处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置493处的缬氨酸残基被丙氨酸残基替代(V493A)。在一些实施方案中,位置493处的缬氨酸残基被丝氨酸残基替代(V493S)。在一些实施方案中,位置493处的缬氨酸残基被苏氨酸残基替代(V493T)。在一些实施方案中,位置493处的缬氨酸残基被精氨酸残基替代(V493R)。在一些实施方案中,位置493处的缬氨酸残基被天冬氨酸残基替代(V493D)。在一些实施方案中,位置493处的缬氨酸残基被脯氨酸残基替代(V493P)。在一些实施方案中,位置493处的缬氨酸残基被甘氨酸残基替代(V493G)。

[0257] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的丙氨酸589处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置589处的丙氨酸残基被缬氨酸残基替代(A589V)。

[0258] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的天冬酰胺597处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置597处的天冬酰胺残基被赖氨酸残基替代(N597K)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被精氨酸残基替代(N597R)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被丙氨酸残基替代(N597A)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被谷氨酸残基替代(N597E)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被组氨酸残基替代(N597H)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被甘氨酸残基替代(N597G)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置597处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置597处的天冬酰胺残基被酪氨酸残基替代(N597Y)。在一些实施方案中,位置597处的天冬酰胺残基被苯丙氨酸残基替代(N597F)。

[0259] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的丝氨酸599处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置599处的丝氨酸残基被苏氨酸残基替代(S599T)。

[0260] 在一些实施方案中,腺苷脱氨酶包含在hADAR2-D氨基酸序列的天冬酰胺613处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,位置613处的天冬酰胺残基被赖氨酸残基替代(N613K)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置613处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置613处的天冬酰胺残基被精氨酸残基替代(N613R)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置613处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置613处的天冬酰胺残基被丙氨酸残基替代(N613A)。在一些实施方案中,腺苷脱氨酶包含在氨基酸序列的位置613处的突变,在野生型序列中该位置有天冬酰胺残基。在一些实施方案中,位置613

处的天冬酰胺残基被谷氨酸残基替代(N613E)。

[0261] 在一些实施方案中,为了提高编辑效率,腺苷脱氨酶可包含以下突变中的一个或多个:G336D、G487A、G487V、E488Q、E488H、E488R、E488N、E488A、E488S、E488M、T490C、T490S、V493T、V493S、V493A、V493R、V493D、V493P、V493G、N597K、N597R、N597A、N597E、N597H、N597G、N597Y、A589V、S599T、N613K、N613R、N613A、N613E(基于hADAR2-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。

[0262] 在一些实施方案中,为了降低编辑效率,腺苷脱氨酶可包含以下突变中的一个或多个:E488F、E488L、E488W、T490A、T490F、T490Y、T490R、T490K、T490P、T490E、N597F(基于hADAR2-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。在特定实施方案中,可能令人感兴趣的是使用功效降低的腺苷脱氨酶来减少脱靶效应。

[0263] 在一些实施方案中,为了减少脱靶效应,腺苷脱氨酶可包含在以下位置处的突变中的一个或多个:R348、V351、T375、K376、E396、C451、R455、N473、R474、K475、R477、R481、S486、E488、T490、S495、R510(基于hADAR2-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。在一些实施方案中,腺苷脱氨酶包含在E488和选自R348、V351、T375、K376、E396、C451、R455、N473、R474、K475、R477、R481、S486、T490、S495、R510的一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在T375处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在N473处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在V351处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在E488和T375处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在E488和N473处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在E488和V351处,以及任选地在一个或多个其他位置处的突变。在一些实施方案中,腺苷脱氨酶包含在E488处以及T375、N473和V351中的一个或多个位置处的突变。

[0264] 在一些实施方案中,为了减少脱靶效应,腺苷脱氨酶可包含选自以下的突变中的一个或多个:R348E、V351L、T375G、T375S、R455G、R455S、R455E、N473D、R474E、K475Q、R477E、R481E、S486T、E488Q、T490A、T490S、S495T和R510E(基于hADAR2-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。在一些实施方案中,腺苷脱氨酶包含突变E488Q和选自R348E、V351L、T375G、T375S、R455G、R455S、R455E、N473D、R474E、K475Q、R477E、R481E、S486T、T490A、T490S、S495T和R510E的一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变T375G或T375S,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变N473D,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变V351L,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变E488Q和T375G或T375G,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变E488Q和N473D,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变E488Q和V351L,以及任选地一个或多个另外的突变。在一些实施方案中,腺苷脱氨酶包含突变E488Q以及T375G/S、N473D和V351L中的一个或多个。

[0265] 在某些实施方案中,通过gRNA的化学修饰实现编辑的改善和脱靶修饰的减少。按照Vogel等人(2014),Angew Chem Int Ed,53:6267-6271,doi:10.1002/anie.201402634(以引用方式整体并入本文)中所例示进行化学修饰的gRNA降低了脱靶活性并提高了中靶

效率。2'-O-甲基和硫代磷酸酯修饰的指导RNA通常可提高细胞中的编辑效率。

[0266] 已证明ADAR显示出对编辑过的A的任一侧的相邻核苷酸的优选性(www.nature.com/nsmb/journal/v23/n5/full/nsmb.3203.html,Matthews等人(2017),Nature Structural Mol Biol,23(5):426-433,以引用方式整体并入本文)。因此,在某些实施方案中,选择gRNA、靶标和/或ADAR针对基序优选性进行优化。

[0267] 在体外已证实有意错配允许编辑非优选的基序(https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gku272;Schneider等人(2014),Nucleic Acid Res,42(10):e87);Fukuda等人(2017),Scientific Reports,7,doi:10.1038/srep41478,以引用方式整体并入本文)。因此,在某些实施方案中,为了提高在非优选的5'或3'相邻碱基上的RNA编辑效率,引入了相邻碱基中的有意错配。

[0268] 结果表明,ADAR脱氨酶结构域的靶向窗口中的与C相对的A可先于其他碱基优先被编辑。此外,在靶向碱基的几个碱基内与U配对的A碱基显示出低Cas13b-ADAR融合物编辑水平,这表明该酶可以灵活地编辑多个A。参见例如图18。这两个观察结果表明,可以通过使待编辑的所有A与C错配来指定Cas13b-ADAR融合物的活性窗口中的多个A以进行编辑。因此,在某些实施方案中,活性窗口中的多个A:C错配被设计来创建多个A:I编辑。在某些实施方案中,为了阻遏活性窗口中潜在的脱靶编辑,将非靶A与A或G配对。

[0269] 术语"编辑特异性"和"编辑优选性"在本文可互换使用,是指在双链底物中特定腺苷位点处A至I编辑的程度。在一些实施方案中,底物编辑优选性是由靶腺苷残基的5'最近邻和/或3'最近邻决定的。在一些实施方案中,腺苷脱氨酶对底物的5'最近邻的优选性排序为U>A>C>G(">"表示更大的优选性)。在一些实施方案中,腺苷脱氨酶对底物的3'最近邻的优选性排序为G>C～A>U(">"表示更大的优选性;"～"表示相似的优选性)。在一些实施方案中,腺苷脱氨酶对底物的3'最近邻的优选性排序为G>C>U～A(">"表示更大的优选性;"～"表示相似的优选性)。在一些实施方案中,腺苷脱氨酶对底物的3'最近邻的优选性排序为G>C>A>U(">"表示更大的优选性)。在一些实施方案中,腺苷脱氨酶对底物的3'最近邻的优选性排序为C～G～A>U(">"表示更大的优选性;"～"表示相似的优选性)。在一些实施方案中,腺苷脱氨酶对含有靶腺苷残基的三联体序列的优选性排序为TAG>AAG>CAC>AAT>GAA>GAC(">"表示更大的优选性),中心A为靶腺苷残基。

[0270] 在一些实施方案中,腺苷脱氨酶的底物编辑优选性受腺苷脱氨酶蛋白中是否存在核酸结合结构域影响。在一些实施方案中,为了修改底物编辑优选性,将脱氨酶结构域与双链RNA结合结构域(dsRBD)或双链RNA结合基序(dsRBM)连接。在一些实施方案中,dsRBD或dsRBM可以来源于ADAR蛋白,诸如hADAR1或hADAR2。在一些实施方案中,使用包含至少一个dsRBD和脱氨酶结构域的全长ADAR蛋白。在一些实施方案中,所述一个或多个dsRBM或dsRBD处于脱氨酶结构域的N端。在其他实施方案中,所述一个或多个dsRBM或dsRBD处于脱氨酶结构域的C端。

[0271] 在一些实施方案中,腺苷脱氨酶的底物编辑优选性受酶活性中心内或附近的氨基酸残基影响。在一些实施方案中,为了修改底物编辑优选性,腺苷脱氨酶可包含以下突变中的一个或多个:G336D、G487R、G487K、G487W、G487Y、E488Q、E488N、T490A、V493A、V493T、V493S、N597K、N597R、A589V、S599T、N613K、N613R(基于hADAR2-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。

[0272] 特别地，在一些实施方案中，为了降低编辑特异性，腺苷脱氨酶可包含以下突变中的一个或多个：E488Q、V493A、N597K、N613K（基于hADAR2-D中的氨基酸序列位置），以及同源ADAR蛋白中对应于以上突变的突变。在一些实施方案中，为了增加编辑特异性，腺苷脱氨酶可包含突变T490A。

[0273] 在一些实施方案中，为了增加对具有最接近5'G的靶腺苷（A），诸如包含三联体序列GAC（中心A为靶腺苷残基）的底物的编辑优选性，腺苷脱氨酶可包含以下突变中的一个或多个：G336D、E488Q、E488N、V493T、V493S、V493A、A589V、N597K、N597R、S599T、N613K、N613R（基于hADAR2-D中的氨基酸序列位置），以及同源ADAR蛋白中对应于以上突变的突变。

[0274] 特别地，在一些实施方案中，腺苷脱氨酶包含突变E488Q或同源ADAR蛋白中的相应突变，以便编辑包含以下三联体序列的底物：GAC、GAA、GAU、GAG、CAU、AAU、UAC，中心A为靶腺苷残基。

[0275] 在一些实施方案中，腺苷脱氨酶包含hADAR1-D的野生型氨基酸序列。在一些实施方案中，腺苷脱氨酶包含在hADAR1-D序列中的一个或多个突变，这样hADAR1-D的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0276] 在一些实施方案中，腺苷脱氨酶包含在hADAR1-D氨基酸序列的甘氨酸1007处，或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中，位置1007处的甘氨酸残基被具有相对小的侧链的非极性氨基酸残基替代。例如，在一些实施方案中，位置1007处的甘氨酸残基被丙氨酸残基替代（G1007A）。在一些实施方案中，位置1007处的甘氨酸残基被缬氨酸残基替代（G1007V）。在一些实施方案中，位置1007处的甘氨酸残基被具有相对大的侧链的氨基酸残基取代。在一些实施方案中，位置1007处的甘氨酸残基被精氨酸残基替代（G1007R）。在一些实施方案中，位置1007处的甘氨酸残基被赖氨酸残基替代（G1007K）。在一些实施方案中，位置1007处的甘氨酸残基被色氨酸残基替代（G1007W）。在一些实施方案中，位置1007处的甘氨酸残基被酪氨酸残基替代（G1007Y）。另外，在其他实施方案中，位置1007处的甘氨酸残基被亮氨酸残基替代（G1007L）。在其他实施方案中，位置1007处的甘氨酸残基被苏氨酸残基替代（G1007T）。在其他实施方案中，位置1007处的甘氨酸残基被丝氨酸残基替代（G1007S）。

[0277] 在一些实施方案中，腺苷脱氨酶包含在hADAR1-D氨基酸序列的谷氨酸1008处，或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中，位置1008处的谷氨酸残基被具有相对大的侧链的极性氨基酸残基替代。在一些实施方案中，位置1008处的谷氨酸残基被谷氨酰胺残基替代（E1008Q）。在一些实施方案中，位置1008处的谷氨酸残基被组氨酸残基替代（E1008H）。在一些实施方案中，位置1008处的谷氨酸残基被精氨酸残基替代（E1008R）。在一些实施方案中，位置1008处的谷氨酸残基被赖氨酸残基替代（E1008K）。在一些实施方案中，位置1008处的谷氨酸残基被非极性或小极性氨基酸残基替代。在一些实施方案中，位置1008处的谷氨酸残基被苯丙氨酸残基替代（E1008F）。在一些实施方案中，位置1008处的谷氨酸残基被色氨酸残基替代（E1008W）。在一些实施方案中，位置1008处的谷氨酸残基被甘氨酸残基替代（E1008G）。在一些实施方案中，位置1008处的谷氨酸残基被异亮氨酸残基替代（E1008I）。在一些实施方案中，位置1008处的谷氨酸残基被缬氨酸残基替代（E1008V）。在一些实施方案中，位置1008处的谷氨酸残基被脯氨酸残基替代（E1008P）。在一些实施方案中，位置1008处的谷氨酸残基被丝氨酸残基替代（E1008S）。在其他实施方案中，位置1008

处的谷氨酸残基被天冬酰胺残基替代(E1008N)。在其他实施方案中,位置1008处的谷氨酸残基被丙氨酸残基替代(E1008A)。在其他实施方案中,位置1008处的谷氨酸残基被甲硫氨酸残基替代(E1008M)。在一些实施方案中,位置1008处的谷氨酸残基被亮氨酸残基替代(E1008L)。

[0278] 在一些实施方案中,为了提高编辑效率,腺苷脱氨酶可包含以下突变中的一个或多个:E1007S、E1007A、E1007V、E1008Q、E1008R、E1008H、E1008M、E1008N、E1008K(基于hADAR1-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。

[0279] 在一些实施方案中,为了降低编辑效率,腺苷脱氨酶可包含以下突变中的一个或多个:E1007R、E1007K、E1007Y、E1007L、E1007T、E1008G、E1008I、E1008P、E1008V、E1008F、E1008W、E1008S、E1008N、E1008K(基于hADAR1-D中的氨基酸序列位置),以及同源ADAR蛋白中对应于以上突变的突变。

[0280] 在一些实施方案中,腺苷脱氨酶的底物编辑优选性、效率和/或选择性受酶活性中心内或附近的氨基酸残基影响。在一些实施方案中,腺苷脱氨酶包含在hADAR1-D序列的谷氨酸1008位置处,或同源ADAR蛋白中的相应位置处的突变。在一些实施方案中,所述突变是E1008R,或同源ADAR蛋白中的相应突变。在一些实施方案中,E1008R突变体对在相反链上具有错配的G残基的靶腺苷残基具有提高的编辑效率。

[0281] 在一些实施方案中,腺苷脱氨酶蛋白还包含或连接至一个或多个双链RNA(dsRNA)结合基序(dsRBM)或结构域(dsRBD),以识别并结合至双链核酸底物。在一些实施方案中,腺苷脱氨酶与双链底物之间的相互作用由一种或多种另外的蛋白质因子(包括CRISPR/CAS蛋白质因子)介导。在一些实施方案中,腺苷脱氨酶与双链底物之间的相互作用进一步由一种或多种核酸组分(包括指导RNA)介导。

[0282] 在某些示例性实施方案中,定向进化可以用于设计修饰的ADAR蛋白,除了将腺嘌呤脱氨为次黄嘌呤之外,所述ADAR蛋白还能够催化另外的反应。例如

[0283] 根据本发明,腺苷脱氨酶的底物是在指导分子与其DNA靶标结合后形成的RNA/DNAn RNA双链体,该双链体之后与CRISPR-Cas酶形成CRISPR-Cas复合物。RNA/DNA或DNA/RNAn RNA在本文中也称为“RNA/DNA杂合体”、“DNA/RNA杂合体”或“双链底物”。指导分子和CRISPR-Cas酶的特殊特征在下文详述。

[0284] 如本文所用,术语“编辑选择性”是指由腺苷脱氨酶编辑的双链底物上所有位点的分数。不受理论的束缚,预期腺苷脱氨酶的编辑选择性受双链底物的长度和二级结构(诸如错配碱基、凸环和/或内环的存在)影响。

[0285] 在一些实施方案中,当底物是长于50bp的完全碱基配对的双链体时,腺苷脱氨酶可能能够使双链体内的多个腺苷残基(例如,所有腺苷残基的50%)脱氨基。在一些实施方案中,当底物短于50bp时,腺苷脱氨酶的编辑选择性受靶腺苷位点处错配存在的影响。特别地,在一些实施方案中,在相反链上具有错配的胞苷(C)残基的腺苷(A)残基被高效脱氨基。在一些实施方案中,在相反链上具有错配的鸟苷(G)残基的腺苷(A)残基被跳过而不进行编辑。

[0286] 用ADAR1/ADAR2进行RNA编辑

[0287] 具有G至A改变(在转录/编码区/剪接序列中)的遗传性疾病可以通过本文所公开的系统和方法来治疗。

[0288] ADAR使腺苷(A)突变为肌苷(I),细胞翻译机制将肌酐(I)视为鸟苷(G)进行处理。在一些实施方案中,可以使用dCas13-ADAR融合物进行RNA编辑以便逆转含有病原性G至A改变的转录物中的遗传突变的效应。可以在双链RNA中进行ADAR RNA编辑,其中编辑的碱基与U或C相对。可以通过使用指导RNA创建双链RNA结构以及将ADAR遗传融合至Cas13b的死亡型式来实现Cas13b-ADAR(例如dCas13b-ADAR)RNA编辑。

[0289] 在ClinVar的所有早期终止病原性突变中,9135中的985个(10.7%)是G至A突变,并且是由Cas13-ADAR融合编辑器进行校正的潜在候选者。在ClinVar的所有病原性突变中,97941中的15531个(15.8%)是G至A突变。

[0290] 可以通过本文的系统和方法来治疗的具有G>A突变的疾病的实例包括:梅尔-戈林综合征(Meier-Gorlin syndrome)、塞克尔综合征(Seckel syndrome)4、乔伯特综合征(Joubert syndrome)5、莱伯氏先天性黑蒙症(Leber congenital amaurosis)10、夏科-马里-图思病(Charcot-Marie-Tooth disease)2型、脑白质病、乌谢尔综合征(Usher syndrome)2C型、脊髓小脑性共济失调28、糖原贮积病III型、原发性高草酸尿症I型、长QT综合征2、西奥格林-拉尔逊氏综合征(Sjögren-Larsson syndrome)、遗传性果糖尿病、神经母细胞瘤、肌萎缩性侧索硬化症9型、卡尔曼综合征(Kallmann syndrome)1、肢带型肌营养不良2L型、家族性腺瘤性息肉病1、家族性3型高脂蛋白血症、阿尔兹海默病1型、异染性脑白质营养不良。

[0291] 还可以使用本文的方法和系统治疗以下疾病。

[0292] 提前终止疾病

[0293] 早期终止疾病的特征在于早期终止密码子中的突变,这些突变是通过引入终止的单核苷酸多态性、改变蛋白质的翻译框架并产生新的终止密码子的插入缺失,或优先引入具有早期终止的外显子的替代剪接而引起的。通过经由A至I编辑去除以这些方式产生的终止密码子,利用本文的系统和方法使用ADAR进行RNA编辑可以挽救涉及提前终止的疾病。如果SNP不是G至A,而是产生无义突变,则将无义突变改变为错义突变可能会带来一些临床益处。

[0294] 在不进行种系编辑的情况下改变能育性突变

[0295] RNA编辑相对于DNA编辑的一个优势是在SNP影响能育性的情况下,通过基因组编辑进行校正必然会导致种系编辑,由此可能产生道德或安全隐患。用本文的系统和方法进行RNA编辑可以校正这些突变而不会对基因组造成永久性影响,从而规避了这些问题。

[0296] 剪接改变

[0297] 前mRNA需要特定的剪接供体和受体序列,才能通过剪接体进行加工。剪接受体位点含有不变的AG序列,所述序列是接受剪接供体序列的攻击和内含子去除所必需的。通过将Cas13b-ADAR融合物靶向前mRNA并将AG剪接受体位点编辑为IG,可以使剪接受体位点失活,从而跳过下游外显子。与目前使用化学修饰的反义寡核苷酸实现外显子跳读的方法相比,这种剪接改变方法具有优势。可以对Cas13b-ADAR进行基因编码,以允许长期跳过外显子。此外,Cas13b-ADAR可以产生突变来促进跳读,这可能比如同使用反义寡核苷酸那样通过双链RNA来掩蔽剪接供体/受体位点更为稳健。

[0298] 新抗原

[0299] 癌症中的新抗原是由于缺陷错配修复引起的突变而在肿瘤细胞中表达的新颖抗

原。针对新抗原对T细胞进行工程化是有利的，原因是由于抗原仅在肿瘤细胞中表达，T细胞将没有脱靶活性并且因此没有毒性。利用本文的系统和方法进行RNA编辑，可以将Cas13-ADAR融合物靶向癌细胞，从而在转录物中引入突变，这样可以引入氨基酸改变和新抗原，如此可以使用嵌合抗原受体T细胞靶向所述新抗原。这种方法比DNA碱基编辑器更好，因为它是瞬时的，因此由于脱靶递送而导致永久编辑非肿瘤细胞的风险很小。

[0300] 改变微小RNA靶标(用于肿瘤阻遏因子)

[0301] ADAR天然地编辑mRNA来生成或去除微小RNA靶标，从而调节表达。可编程RNA编辑可用于经由改变靶向区域来上调或下调微小RNA靶标。此外，微小RNA本身是ADAR的天然底物，对微小RNA进行的可编程RNA编辑可以降低或增强其相应靶标上的功能。

[0302] 沿区域进行多次编辑(指导物中存在多个错配)

[0303] 通过在所需腺苷靶标对面的指导区引入错配并产生有利于A至I编辑的气泡，可以精确地靶向Cas13-ADAR融合物从而编辑特定的腺苷。通过在指导物/靶标双链体中的不同腺苷位点中引入多个这些错配，可以一次引入多个突变。

[0304] 对PTC的TAA的逆转(两个A至G)

[0305] 涉及早期终止密码子改变的许多疾病都涉及TAA终止密码子，TAA终止密码子需要A至I改变才能进行校正，而TAG或TGA终止密码子仅需要一个A至I编辑。可以使用两种方法来逆转TAA终止密码子。(1)如本文所述，可以针对TAA密码子中的两个腺苷在指导物中引入两个错配。(2)可以使用两个指导阵列依序将每个腺苷转化为肌苷。所述阵列中的第一指导物可以含有针对第一腺苷的突变，而第二指导物则可以具有针对此改变的互补性，并且具有针对该终止密码子中的第二腺苷的错配。

[0306] 癌症(GOF、LOF突变逆转)

[0307] 癌症中的许多致癌改变涉及G至A突变，这些突变将功能获得或功能丧失表型引入突变的蛋白质中。本文的RNA编辑系统和方法可以被适当定位以校正这些改变并减少肿瘤发生。

[0308] 新碱基优选性设计

[0309] 已发现对于催化活性，当前的ADAR1/2蛋白具有周围碱基优选性(,pubs.acs.org/doi/10.1021/acschembio.5b00711)，这可能对某些应用构成限制。ADAR变体的合理诱变或定向进化以及碱基优选性改变或放宽可以提高可编程RNA编辑的多功能性。

[0310] 可以使用的ADAR直系同源物

[0311] 各种ADAR直系同源物可用于本文的方法和系统。可以使用的ADAR直系同源物包括：

[0312]

ADAR 名称	蛋白质序列
智人_ADAR2_E_Q_突变体	QLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRK VLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSD RGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQ KRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFS PHEPILEEPADRHPNRKARGQLRTKIESGQGTIPVRS NASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQG SLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIED LPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVG DSAIEVINATTGKDELGRASRLCKHALYCRWMRVH GKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARL FTAFIKAGLGAWVEKPTEQDQFSLT*
智人_ADAR1_E_Q_突变体	SLGTGNRCVKGDSLSLKGETVNDCHAEIISRRGFIRF LYSELMKYNSQTAKDSIFEPAKGGEKLQIKKTVSFH LYISTAPCGDGALFDKSCSDRAMESTESRHYPVFEN PKQGKLRTKVENGQGTIPVESSDIVPTWDGIRLGER LRTMSCSDKILRWNVLGLQGALLTHFLQPIYLKSVT LGYLFSQGHLTRAICCRVTRDGSAFEDGLRHPFIVN HPKVGRVSIYDSKRQSGKTKETSVNWCLADGYDLE ILDGTRGTVDGPRNELSRVSKKNIFLLFKKLCSFRY RRDLLRLSYGEAKKAARDYETAKNYFKKGLKDMG YGNWISKPQEEKNF
普通章鱼(Octopus vulgaris)_ADAR1_E_Q_突变体	SVGTGNRCLTGDHLSLEGNSVNDSHAEMITRRGFL RYLYRHLLEYDAEVPNDLFEKGERSICRIKTNITFHL YISTAPCGDGALFSPRDTDSSNAKMEEENKHIHNPT FSSSVQGLLRTKVEGGQGTIPIDADFTEQTWDGIQR GERLRTMSCSDKICRWNVVGLQGALLSHFIEPIYLD SLTLGYLYDHGHLARAVCCRIERGEASVNQLLPEG YRLNHPWLGRVTACDPPRETQKTKSLSINWCYDDE KSEVLDGTAGICYTAIEKNLFSRLTKHNLYEEFKRV CRKFDRNDLLTAPSYNKAKMMATPFQTAKNVMLK KLKENNCGTWVSKPIEEEMF
乌贼(Sepia)_ADAR1_E_Q_突变体	SVGTGNRCLTGDRLSLEGNSVNDSHAEMVTRRGFL RYLYKHLLEYDPEKPHDLFEKGERSLCRIKTNITFHL YISTAPCGDGALFSPRDTDSSNVKVDEENKHVHNP TFSSSVQGLLRTKVEGGQGTIPIDADFTEQTWDGIQ RGERLRTMSCSDKICRWNVVGLQGALLSHFVEPIYL ESLTLGYLYDHGHLARAVCCRIERGEASVNQLLPE GYRLNHPWLGRVTACDPPRETQKTKSLSINWCYDD EKSEVLDGTAGICYTAIEKNLFSRLTKHSLYEEFKK VCQKFEREDLLNVTSYNKAKMMAIPFQTAKNVML KKLKENNCGTWVSKPIEEEMF

[0313]

普通章鱼_ADAR2_E_Q_突变体	GIGTGTKCINGEHMSDRGFGVNDCHAEIIARRCFLR YIYDQLELHLSDNSDVRNSSIFELRDKGGYQLKENI QFHLYISTAPCGDARIFSPHGQDVETGDRHPNRKAR GQLRTKIESGQGTIPVRTSGVIQTWDGVLEGERLLT MSCSDKIARWNVLGIQGSLLSHFMNPIYLESIILGSL YHSDHLSRAMYSRISIIENLPEPFHLNRPFLSGISSPE SRQPGKAPNFGINWRKEDETFEVINAMTGRVEGGS VSRICKQALFGRFMSLYGKLSSLTGQSVTTRPTHYS DAKAAVMEYQLAKQCVFQAFQKAGLGNWVQKPIE QDQF
乌贼_ADAR2_E_Q_突变体	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLR FIYDQLEMHLSEDPEVRGQSVFELRDGGGYKLKPNI HFHLYISTAPCGDARIFSPHGQDVETGDRHPNRKAR GQLRTKIESGQGTIPVRSSGFIQTWDGVLEGERLLT MSCSDKIARWNVLGIQGALLCHFMHPIYLESIILGSL YHSDHLSRAVYCRIASIENLPDLFQLNRPFLSGISSPE SRQPGKAPNFGINWRRNDDTFEVINAMTGRVEGGN MSRICKQALFDRFMNLYGRLSSLTGQSVTTRPTLYS EAKAAVMEYQLAKQCVFQAFQKAGLGNWVQKPIE QDQF
乳光枪乌贼(Doryteusthis opalescens)_ADAR2_E_Q_突变体	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLR FIYDQLELHLSDNAEVRGQSIFELRDAGGYKLKPNI QFHLYISTAPCGDARIFSPHGQDVETGDRHPNRKAR GQLRTKIESGQGTIPVRSSGFIQTWDGVLEGERLLT MSCSDKIARWNVLGVQGALLCHFMHPIYLESIILGS LYHSDHLSRAVYCRIAAIENLPDLFRLNRPFLSGISS PESRQPGKAPNFGINWRRNDDSFEVINAMTGRVEG GSMSRICKQALFDRFMNLYGKLSSLTGQSVTTRPAL YSEAKATVMEYQLAKQCVFQAFQKAGLGNWVQK PIEQDQF

[0314] 需注意,以上列出的直系同源物具有高活性E至Q突变。可以使用头足类动物直系同源物。已知头足类动物的RNA编辑率极高,因此这些直系同源物可能具有高于人类变体的活性或改变的碱基优选性。

[0315] 这些直系同源物的比对树在图74中示出。

[0316] 在人体细胞中具有增强活性的ADAR突变体.先前已经报告了在体外或在酵母中具有改变的活性的ADAR突变体。本公开还提供了筛选或合理设计在人类细胞的情形中具有增强活性的突变体的方法,这些突变体可以提高基于ADAR的可编程RNA编辑系统或构建体的效率或特异性。

[0317] 肌苷产生的生物应用.用ADAR进行RNA编辑会产生肌苷,当在转录物中多次出现时,肌苷可以与内源性生物途径相互作用,进而加剧细胞和组织中的炎症。多个肌苷碱基的产生会加剧炎症,尤其是在炎症会导致清除的细胞中。另外的肌苷产生也可以用于使转录物不稳定。

[0318] 去除上游起始密码子以促进下游ORF的蛋白质表达(ATG突变).反义寡核苷酸可用于阻断上游起始密码子位点,以促进下游起始密码子处的蛋白质表达。这可以允许出于治

疗目的提高内源蛋白质水平。Cas13-ADAR融合物可以通过将ATG位点转化为ITG(GTG)位点并且由此去除内源转录物中的上游密码子且由此增强蛋白质翻译而达到类似的效果。这可能是允许增强基因表达的应用。一个实例包括在镰状细胞病和地中海贫血中提高胎儿血红蛋白水平。

[0319] 用于C至U或任何转换的ADAR诱变.通过合理诱变或定向进化,可以将ADAR(例如,在直系同源物部分中列出的那些)制成为C至U编辑器或任何碱基转换的编辑器。

[0320] APOBEC RNA编辑

[0321] C至U转化.所述系统可以包括Cas13b-APOBEC融合物(例如dCas13b-APOBEC融合物)。这些融合物使得可以对RNA进行C至U编辑。APOBEC底物是ssRNA,并且这些系统可以靶向在指导物/靶标双链体周围的RNA区域。

[0322] 经由终止密码子引入进行C至U敲除.除了校正在细胞生命周期中出现的病原性U至C突变外,Cas13-APOBEC融合物还可以通过将CAA、CGA或CAG分别转化为TAA、TGA或TAG来引入终止密码子。

[0323] APOBEC直系同源物.可以使用所述系统筛选APOBEC直系同源物。筛选其他APOBEC直系同源物与Cas13(例如dCas13)的融合物可以提高C至U编辑的效率,或者可以允许其他类型的碱基转化。

[0324] 使APOBEC突变以实现dsRNA/碱基翻转/活性增加.利用编辑体进行RNA插入/缺失。所述系统还可以包含RNA编辑体(例如,融合至Cas13诸如dCas13的RNA编辑体)。此类系统可以向转录物中导入靶向序列缺失或插入。

[0325] 参考文献

[0326] Human SNPs resulting in premature stop codons and protein truncation.Hum Genomics,2006;2(5);274-286.

[0327] The genetics of human infertility by functional interrogation of SNPs in mice.PNAS August 18,2015 112(33)10431-10436

[0328] microRNA editing in seed region aligns with cellular changes in hypoxic conditionns.Nucleic Acids Research,Volume 44,Issue 13,27July 2016,Pages 6298-6308

[0329] RNA Editing Modulates Human Hepatic Aryl Hydrocarbon Receptor Expressionby Creating MicroRNA Recognition Sequence.January 8,/2016,The Journal of BiologicalChemistry,291,894-903.

[0330] Mechanistic insights inro editing-site specificity of ADARs.PNAS November 27,2012 109 (48)E3295-E3304.

[0331] A Phenotypic Screen for Functional Mutants of Human Adenosine DeaminaseActing on RNA 1.ACS Chem.Biol.,2015,10(11),pp 2512-2519.

[0332] Inosine-ContainingRNA Is a Novel Innate Immune Recognirion Element and Reduces RSV Infection.PLoS One,2011,6(10):e26463.

[0333] Inosine-Mediated Modulation of RNA Sensing by Toll-Like Receptor 7 (TLR7)and TLR8.J Virol.,2014Jan;88(2);799-810.

[0334] Specific cleavage of hyper-eited dsRNAs.The EMBO Journal(2001)20,

4243-4252.

[0335] Translation efficiency of mRNAs is increased by antisense oligonucletidestargeting upstream open reading frames.Nature Biotechnology volume 34,pages 875-880(2016).

[0336] The mechanism of Uinsertion/deletion RNA editing in kinetoplastid mitochondriaNucleic Acids Res.1997 Oct 1,25(19):3751-3759.

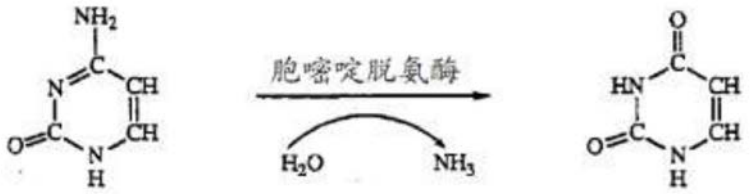
[0337] Functions and Regulation of RNA Editing by ADAR Deaminases.Annual Review of Biochemistry,Vol.79:321-349(Volume publication date 7/7/2010).

[0338] Trade-off betwcen Transcriptome Plasticity and Genome Evolution inCephalopods.Cell.Volume 169,lssue,2,6 April 2017,Pages191-202.ell.

[0339] An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-Udeamination of DNA.PNAS May 8,2007 104(19)7821-7826.

[0340] 胞苷脱氨酶

[0341] 如本文所用,术语“胞苷脱氨酶”或“胞苷脱氨酶蛋白”是指能够催化如下所示的将胞嘧啶(或分子的胞嘧啶部分)转化为尿嘧啶(或分子的尿嘧啶部分)的水解脱氨反应的蛋白质、多肽或蛋白质或多肽的一个或多个功能结构域。在一些实施方案中,含胞嘧啶的分子是胞苷(C),而含尿嘧啶的分子是尿苷(U)。含胞嘧啶的分子可以是脱氧核糖核酸(DNA)或核糖核酸(RNA)。

[0342] 

胞嘧啶
(4-氨基-2-氧代嘧啶)

尿嘧啶
(2,4-二氧代嘧啶)

[0343] 根据本公开,可以与本公开结合使用的胞苷脱氨酶包括但不限于称为载脂蛋白B mRNA编辑复合物(APOBEC)家族脱氨酶的酶家族成员、激活诱导的脱氨酶(AID)或胞苷脱氨酶1(CDA1)。在特定实施方案中,所述脱氨酶是APOBEC1脱氨酶、APOBEC2脱氨酶、APOBEC3A脱氨酶、APOBEC3B脱氨酶、APOBEC3C脱氨酶和APOBEC3D脱氨酶、APOBEC3E脱氨酶、APOBEC3F脱氨酶、APOBEC3G脱氨酶、APOBEC3H脱氨酶、或APOBEC4脱氨酶。

[0344] 在本发明的方法和系统中,胞苷脱氨酶能够靶向DNA单链中的胞嘧啶。在某些示例性实施方案中,胞苷脱氨酶可以在存在于结合组分例如结合的Cas13外部的单链上进行编辑。在其他示例性实施方案中,胞苷脱氨酶可以在局部气泡,诸如由靶编辑位点处指导序列的错配所形成的局部气泡处进行编辑。在某些示例性实施方案中,胞苷脱氨酶可以包含有助于集中活性区域的突变,诸如在Kim等人,Nature Biotechnology(2017)35(4):371-377(doi:10.1038/nbt.3803中公开的那些突变。

[0345] 在一些实施方案中,胞苷脱氨酶来源于一种或多种后生动物种类,包括但不限于哺乳动物、鸟、青蛙、鱿鱼、鱼、蝇和蠕虫。在一些实施方案中,胞苷脱氨酶是人类、灵长类动物、牛、狗大鼠或小鼠胞苷脱氨酶。

[0346] 在一些实施方案中，胞苷脱氨酶是人类APOBEC，包括hAPOBEC1或hAPOBEC3。在一些实施方案中，胞苷脱氨酶是人类AID。

[0347] 在一些实施方案中，胞苷脱氨酶蛋白识别RNA双链体的单链气泡中的一个或多个靶胞嘧啶残基并将它们转化为一个或多个尿嘧啶残基。在一些实施方案中，胞苷脱氨酶蛋白识别RNA双链体的单链气泡上的结合窗口。在一些实施方案中，结合窗口含有至少一个靶胞嘧啶残基。在一些实施方案中，结合窗口在约3bp至约100bp的范围内。在一些实施方案中，结合窗口在约5bp至约50bp的范围内。在一些实施方案中，结合窗口在约10bp至约30bp的范围内。在一些实施方案中，结合窗口为约1bp、2bp、3bp、5bp、7bp、10bp、15bp、20bp、25bp、30bp、40bp、45bp、50bp、55bp、60bp、65bp、70bp、75bp、80bp、85bp、90bp、95bp或100bp。

[0348] 在一些实施方案中，胞苷脱氨酶蛋白包含一个或多个脱氨酶结构域。不意欲受理论的束缚，预期脱氨酶结构域用于识别RNA双链体的单链气泡中所含的一个或多个靶胞嘧啶(C)残基并将它们转化为一个或多个尿嘧啶(U)残基。在一些实施方案中，脱氨酶结构域包含活性中心。在一些实施方案中，活性中心包含锌离子。在一些实施方案中，在活性中心内或附近的氨基酸残基与靶胞嘧啶残基5'端的一个或多个核苷酸相互作用。在一些实施方案中，在活性中心内或附近的氨基酸残基与靶胞嘧啶残基3'端的一个或多个核苷酸相互作用。

[0349] 在一些实施方案中，胞苷脱氨酶包含人类APOBEC1全蛋白(hAPOBEC1)或其脱氨酶结构域(hAPOBEC1-D)或其C端截短型式(hAPOBEC-T)。在一些实施方案中，胞苷脱氨酶是与hAPOBEC1、hAPOBEC-D或hAPOBEC-T同源的APOBEC家族成员。在一些实施方案中，胞苷脱氨酶包含人类AID1全蛋白(hAID)或其脱氨酶结构域(hAID-D)或其C端截短型式(hAID-T)。在一些实施方案中，胞苷脱氨酶是与hAID、hAID-D或hAID-T同源的AID家族成员。在一些实施方案中，hAID-T是C端被截短约20个氨基酸的hAID。

[0350] 在一些实施方案中，胞苷脱氨酶包含胞嘧啶脱氨酶的野生型氨基酸序列。在一些实施方案中，胞苷脱氨酶在胞嘧啶脱氨酶序列中包含一个或多个突变，这样胞嘧啶脱氨酶的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0351] APOBEC1和APOBEC3蛋白的某些突变已描述于Kim等人，Nature Biotechnology(2017)35(4):371-377(doi:10.1038/nbt.3803)；和Harris等人Mol.Cell(2002)10:1247-1253，所述文献各自以引用方式整体并入本文。

[0352] 在一些实施方案中，胞苷脱氨酶是在对应于大鼠APOBEC1中的W90、R118、H121、H122、R126或R132的氨基酸位置处包含一个或多个突变的APOBEC1脱氨酶，或者是在对应于人类APOBEC3G中的W285、R313、D316、D317X、R320或R326的氨基酸位置处包含一个或多个突变的APOBEC3G脱氨酶。

[0353] 在一些实施方案中，胞苷脱氨酶包含在大鼠APOBEC1氨基酸序列的色氨酸^{90}处，或同源APOBEC蛋白中的相应位置，诸如APOBEC3G的色氨酸^{285}处的突变。在一些实施方案中，位置90处的色氨酸残基被酪氨酸或苯丙氨酸残基替代(W90Y或W90F)。

[0354] 在一些实施方案中，胞苷脱氨酶包含在大鼠APOBEC1氨基酸序列的精氨酸^{118}处，或同源APOBEC蛋白中的相应位置处的突变。在一些实施方案中，位置118处的精氨酸残基被丙氨酸残基替代(R118A)。

[0355] 在一些实施方案中，胞苷脱氨酶包含在大鼠APOBEC1氨基酸序列的组氨酸^{121}处，或

同源APOBEC蛋白中的相应位置处的突变。在一些实施方案中，位置121处的组氨酸残基被精氨酸残基替代(H121R)。

[0356] 在一些实施方案中，胞苷脱氨酶包含在大鼠APOBEC1氨基酸序列的组氨酸^{122}处，或同源APOBEC蛋白中的相应位置处的突变。在一些实施方案中，位置122处的组氨酸残基被精氨酸残基替代(H122R)。

[0357] 在一些实施方案中，胞苷脱氨酶包含在大鼠APOBEC1氨基酸序列的精氨酸^{126}处，或同源APOBEC蛋白中的相应位置，诸如APOBEC3G的精氨酸^{320}处的突变。在一些实施方案中，位置126处的精氨酸残基被丙氨酸残基替代(R126A)或被谷氨酸替代(R126E)。

[0358] 在一些实施方案中，胞苷脱氨酶包含在APOBEC1氨基酸序列的精氨酸^{132}处，或同源APOBEC蛋白中的相应位置处的突变。在一些实施方案中，位置132处的精氨酸残基被谷氨酸残基替代(R132E)。

[0359] 在一些实施方案中，为了使编辑窗口的宽度变窄，胞苷脱氨酶可以包含以下突变中的一个或多个：W90Y、W90F、R126E和R132E(基于大鼠APOBEC1的氨基酸序列位置)，以及同源APOBEC蛋白中对应于以上突变的突变。

[0360] 在一些实施方案中，为了降低编辑效率，胞苷脱氨酶可包含以下突变中的一个或多个：W90A、R118A、R132E(基于大鼠APOBEC1的氨基酸序列位置)，以及同源APOBEC蛋白中对应于以上突变的突变。在特定实施方案中，可能令人感兴趣的是使用功效降低的胞苷脱氨酶来减少脱靶效应。

[0361] 在一些实施方案中，胞苷脱氨酶是野生型大鼠APOBEC1(rAPOBEC1)或其催化结构域。在一些实施方案中，胞苷脱氨酶在rAPOBEC1序列中包含一个或多个突变，这样rAPOBEC1的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0362] rAPOBEC1：

[0363] MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK

[0364] 在一些实施方案中，胞苷脱氨酶是野生型人类APOBEC1(hAPOBEC1)或其催化结构域。在一些实施方案中，胞苷脱氨酶在hAPOBEC1序列中包含一个或多个突变，这样hAPOBEC1的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0365] APOBEC1：

[0366] MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTTNHVEVNFIKKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLFWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMMLYALELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR

[0367] 在一些实施方案中，胞苷脱氨酶是野生型人类APOBEC3G(hAPOBEC3G)或其催化结构域。在一些实施方案中，腺苷脱氨酶在hAPOBEC3G序列中包含一个或多个突变，这样hAPOBEC3G的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0368] hAPOBEC3G：

[0369] MELKYHPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWD

PDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYYILLHIMLGEILRHSMDPPTFTFN FNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTCF TSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQ PWDGLDEHSQDLSGRLRAILQNQEN

[0370] 在一些实施方案中,胞苷脱氨酶是野生型海七鳃鳗(Petromyzon marinus)CDA1(pmCDA1)或其催化结构域。在一些实施方案中,腺苷脱氨酶在pmCDA1序列中包含一个或多个突变,这样pmCDA1的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0371] pmCDA1:

[0372] MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQSGTERGIHAEI FSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGHTLKIWACKLYYEKNARNQIGLWNLRDNGV GLNVMVSEHYQCCRKIFIQSSHNQLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAV

[0373] 在一些实施方案中,胞苷脱氨酶是野生型人类AID(hAID)或其催化结构域。在一些实施方案中,腺苷脱氨酶在pmCDA1序列中包含一个或多个突变,这样pmCDA1的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0374] hAID:

[0375] MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHVELLFLRYISDWD LDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPYLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIMTFKDYFY CWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLPLYEVDDLRDAFRTLGLLD

[0376] 在一些实施方案中,胞苷脱氨酶是hAID(hAID-DC)或其催化结构域的截短型式。在一些实施方案中,胞苷脱氨酶在hAID-DC序列中包含一个或多个突变,这样hAID-DC的编辑效率和/或底物编辑优选性可根据特定需要而改变。

[0377] hAID-DC:

[0378] MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHVELLFLRYISDWD LDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIMTFKDYFY CWNTFVENHERTFKAWEGLHENSVRLSRQLRRILL

[0379] 胞苷脱氨酶的其他实施方案公开于标题为“核酸酶编辑器及其用途(Nucleobase Editor and Uses Thereof)”的WO WO2017/070632中,所述文献以引用方式整体并入本文。

[0380] 在一些实施方案中,胞苷脱氨酶具有包围易于进行脱氨编辑的核苷酸的有效脱氨窗口。因此,在一些实施方案中,“编辑窗口宽度”是指在给定靶位点上胞苷脱氨酶的编辑效率超过该靶位点的一半最大值的核苷酸位置的数目。在一些实施方案中,胞苷脱氨酶的编辑窗口宽度在约1至约6个核苷酸范围内。在一些实施方案中,胞苷脱氨酶的编辑窗口宽度是1、2、3、4、5或6个核苷酸。

[0381] 不意欲受理论的束缚,预期在一些实施方案中,接头序列的长度影响编辑窗口宽度。在一些实施方案中,编辑窗口宽度随着接头长度延长(例如,约3至约21个氨基酸)而增加(例如,约3至约6个核苷酸)。在非限制性实例中,16个残基的接头提供了约5个核苷酸的有效脱氨窗口。在一些实施方案中,指导RNA的长度影响编辑窗口宽度。在一些实施方案中,缩短指导RNA导致胞苷脱氨酶的有效脱氨窗口变窄。

[0382] 在一些实施方案中,胞苷脱氨酶的突变影响编辑窗口宽度。在一些实施方案中,CD官能化的CRISPR系统的胞苷脱氨酶组分包含一个或多个突变,所述一个或多个突变降低胞

苷脱氨酶的催化效率，如此防止该脱氨酶在每个DNA结合事件中使多个胞苷脱氨基。在一些实施方案中，使APOBEC1的残基90处的色氨酸(W90)或同源序列中的相应色氨酸残基突变。在一些实施方案中，将无催化活性的Cas13融合至或连接至包含W90Y或W90F突变的APOBEC1突变体。在一些实施方案中，使APOBEC3G的残基285处的色氨酸(W285)或同源序列中的相应色氨酸残基突变。在一些实施方案中，将无催化活性的Cas13融合至或连接至包含W285Y或W285F突变的APOBEC3G突变体。

[0383] 在一些实施方案中，CD官能化的CRISPR系统的胞苷脱氨酶组分包含一个或多个突变，所述一个或多个突变降低胞苷对脱氨酶活性位点的非最佳呈递的耐受性。在一些实施方案中，胞苷脱氨酶包含一个或多个突变，所述一个或多个突变改变脱氨酶活性位点的底物结合活性。在一些实施方案中，胞苷脱氨酶包含一个或多个突变，所述一个或多个突变改变由脱氨酶活性位点识别并结合的DNA的构象。在一些实施方案中，胞苷脱氨酶包含一个或多个突变，所述一个或多个突变改变底物对脱氨酶活性位点的可及性。在一些实施方案中，使APOBEC1的残基126处的精氨酸(R126)或同源序列中的相应精氨酸残基突变。在一些实施方案中，将无催化活性的Cas13融合至或连接至包含R126A或R126E突变的APOBEC1。在一些实施方案中，使APOBEC3G的残基320处的色氨酸(R320)或同源序列中的相应精氨酸残基突变。在一些实施方案中，将无催化活性的Cas13融合至或连接至包含R320A或R320E突变的APOBEC3G突变体。在一些实施方案中，使APOBEC1的残基132处的精氨酸(R132)或同源序列中的相应精氨酸残基突变。在一些实施方案中，将无催化活性的Cas13融合至或连接至包含R132E突变的APOBEC1突变体。

[0384] 在一些实施方案中，CD官能化的CRISPR系统的APOBEC1结构域包含选自W90Y、W90F、R126A、R126E和R132E的一个、两个或三个突变。在一些实施方案中，APOBEC1结构域包含双突变，即W90Y和R126E。在一些实施方案中，APOBEC1结构域包含双突变，即W90Y和R132E。在一些实施方案中，APOBEC1结构域包含双突变，即R126E和R132E。在一些实施方案中，APOBEC1结构域包含三个突变，即W90Y、R126E和R132E。

[0385] 在一些实施方案中，如本文所公开的胞苷脱氨酶中的一个或多个突变将编辑窗口宽度减小至约2个核苷酸。在一些实施方案中，如本文所公开的胞苷脱氨酶中的一个或多个突变将编辑窗口宽度减小至约1个核苷酸。在一些实施方案中，如本文所公开的胞苷脱氨酶中的一个或多个突变减小编辑窗口宽度，同时仅最小程度地或适度地影响酶的编辑效率。在一些实施方案中，如本文所公开的胞苷脱氨酶中的一个或多个突变减小编辑窗口宽度，而没有降低酶的编辑效率。在一些实施方案中，如本文所公开的胞苷脱氨酶中的一个或多个突变能够区分相邻的胞苷核苷酸，否则这些胞苷核苷酸将被胞苷脱氨酶以相似的效率编辑。

[0386] 在一些实施方案中，胞苷脱氨酶蛋白还包含或连接至一个或多个双链RNA(dsRNA)结合基序(dsRBM)或结构域(dsRBD)，以识别并结合至双链核酸底物。在一些实施方案中，胞苷脱氨酶与底物之间的相互作用由一种或多种另外的蛋白质因子(包括CRISPR/CAS蛋白质因子)介导。在一些实施方案中，胞苷脱氨酶与底物之间的相互作用进一步由一种或多种核酸组分(包括指导RNA)介导。

[0387] 根据本发明，胞苷脱氨酶的底物是包含目标胞嘧啶的RNA双链体的DNA单链气泡，其在指导分子与其DNA靶标结合接着与CRISPR-Cas酶形成CRISPR-Cas复合物后可接近胞苷

脱氨酶，由此胞嘧啶脱氨酶融合至或能够结合至CRISPR-Cas复合物的一种或多种组分(即CRISPR-Cas酶和/或指导分子)。指导分子和CRISPR-Cas酶的特殊特征在下文详述。

[0388] RNA结合蛋白

[0389] 在某些示例性实施方案中，RNA结合蛋白或其功能结构域包含RNA识别基序。在一些实例中，RNA结合蛋白可以是Cas效应蛋白，诸如Cas13效应蛋白。在一些情况下，RNA结合蛋白可以是无催化活性的Cas效应蛋白，诸如dCas13。

[0390] 在某些示例性实施方案中，RNA结合蛋白包含锌指基序。RNA结合蛋白或其功能结构域可以包含Cys2-His2、Gag-knuckle、Treble-clet、锌带、Zn2/Cys6类基序。

[0391] 在某些示例性实施方案中，RNA结合蛋白可以包含Pumilio同源结构域。

[0392] 在一方面，本发明提供了用于对RNA中，更特别地是目标RNA序列中的腺嘌呤进行靶向脱氨的方法。根据本发明的方法，通过可以特异性地结合至靶序列的CRISPR-Cas复合物将腺苷脱氨酶(AD)蛋白特异性地募集至目标RNA序列中的相关腺嘌呤。为了实现这一点，可以将腺苷脱氨酶蛋白共价连接至CRISPR-Cas酶或作为单独的蛋白提供，但是进行调整以确保将其募集至CRISPR-Cas复合物。

[0393] 在本发明方法的特定实施方案中，通过将腺苷脱氨酶或其催化结构域融合至CRISPR-Cas蛋白(其为Cas13蛋白)来确保将腺苷脱氨酶募集至靶基因座。由两种分离的蛋白质产生融合蛋白的方法在本领域中是已知的，并且通常涉及间隔区或接头的使用。Cas13蛋白可以在其N端或C端融合至腺苷脱氨酶蛋白或其催化结构域。在特定实施方案中，CRISPR-Cas蛋白是无活性的或死亡的Cas13蛋白并且连接至脱氨酶蛋白或其催化结构域的N端。

[0394] 如关于融合蛋白使用的术语“接头”是指将所述蛋白连接以形成融合蛋白的分子。通常，除了连接蛋白质或保持蛋白质之间的一些最小距离或其他空间关系之外，此类分子不具有特殊的生物活性。然而，在某些实施方案中，可以选择接头以影响接头和/或融合蛋白的某些特性，诸如接头的折叠、净电荷或疏水性。接头(例如非核苷酸环)可以具有任何长度。在一些实施方案中，接头具有等于约0-16个核苷酸的长度。在一些实施方案中，接头具有等于约0-8个核苷酸的长度。在一些实施方案中，接头具有等于约0-4个核苷酸的长度。在一些实施方案中，接头具有等于约2个核苷酸的长度。示例性接头设计另外描述于WO2011/008730中。

[0395] 适用于本发明方法的接头是本领域技术人员所熟知的，并且包括但不限于直链或支链碳接头、杂环碳接头或肽接头。但是，如本文所用，接头还可以是共价键(碳-碳键或碳-杂原子键)。在特定实施方案中，接头用于将CRISPR-Cas蛋白和腺苷脱氨酶分开足以确保每种蛋白保留其所需功能特性的距离。优选的肽接头序列采用柔性的延伸构象，并且不表现出发展有序二级结构的倾向。在某些实施方案中，接头可以是化学部分，其可以是单体、二聚体、多聚体或聚合物。优选地，接头包含氨基酸。柔性接头中的典型氨基酸包括Gly、Asn和Ser。因此，在特定实施方案中，接头包含Gly、Asn和Ser氨基酸中的一种或多种的组合。其他接近中性的氨基酸，诸如Thr和Ala，也可以用于接头序列中。示例性接头公开于Maratea等人(1985)，Gene 40:39-46；Murphy等人(1986)Proc.Nat'l.Acad.Sci.USA 83:8258-62；美国专利号4,935,233；和美国专利号4,751,180中。例如，可以使用GlySer接头GGS、GGGS或GSG。GGS、GSG、GGGS或GGGGS接头可以按3个(诸如$(GGS)_3$(SEQ ID NO:12)、$(GGGGS)_3$)(SEQ ID

NO:1)或5、6、7、9甚至12个或更多个的重复使用,以提供合适的长度。在特定实施方案中,在本文中优选使用诸如(GGGGS)$_3$(SEQ ID NO:1)的接头。可以优选地使用(GGGGS)$_6$(SEQ ID NO:4)、(GGGGS)$_9$(SEQ ID NO:7)或(GGGGS)$_{12}$)(SEQ ID NO:13)作为替代品。其他优选的替代品是(GGGGS)1)(SEQ ID NO:14)、(GGGGS)$_2$)(SEQ ID NO:15)、(GGGGS)$_4$)(SEQ ID NO:2)、(GGGGS)$_5$)(SEQ ID NO:3)、(GGGGS)$_7$)(SEQ ID NO:5)、(GGGGS)$_8$)(SEQ ID NO:6)、(GGGGS)$_{10}$)(SEQ ID NO:8)或(GGGGS)$_{11}$)(SEQ ID NO:9)。在又一个实施方案中,将LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR(SEQ ID NO:11)用作接头。在又另一个实施方案中,接头是XTEN接头。在特定实施方案中,CRISPR-cas蛋白是Cas13蛋白并且通过LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR(SEQ ID NO:11)接头连接至脱氨酶蛋白或其催化结构域。在另外的特定实施方案中,Cas13蛋白在C端通过LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR(SEQ ID NO:11)接头连接至脱氨酶蛋白或其催化结构域的N端。此外,N端和C端的NLS也可以充当接头(例如PKKKRKVEASSPKKRKVEAS(SEQ ID NO:16))。

[0396] 在本发明方法的特定实施方案中,将腺苷脱氨酶蛋白或其催化结构域作为单独的蛋白递送至细胞或在细胞内表达,但对其进行了修饰以使其能够连接至Cas13蛋白或指导分子。在特定实施方案中,这是通过使用存在于多种多样的噬菌体外壳蛋白中的正交RNA结合蛋白或衔接蛋白/适体组合来确保的。此类外壳蛋白的实例包括但不限于:MS2、Qβ、F2、GA、fr、JP501、M12、R17、BZ13、JP34、JP500、KU1、M11、MX1、TW18、VK、SP、FI、ID2、NL95、TW19、AP205、ΦCb5、ΦCb8r、ΦCb12r、ΦCb23r、7s和PRR1。适体可以是通过反复轮次的体外选择或SELEX(指数富集的配体系统进化)进行工程化以便与特定靶标结合的天然存在的或合成的寡核苷酸。

[0397] 在本发明的方法和系统的特定实施方案中,指导分子提供有一个或多个可以募集衔接蛋白的不同的RNA环或不同的序列。可以通过插入一个或多个不同的RNA环或不同的序列来延长指导分子,而不会与Cas13蛋白发生冲突,这些不同的RNA环或不同的序列可以募集可以结合至所述不同的RNA环或不同的序列的衔接蛋白。Konermann(Nature 2015,517(7536):583-588)中提供了修饰指导物的示例及其在将效应结构域募集至CRISPR-Cas复合物中的用途。在特定实施方案中,适体是选择性地结合哺乳动物细胞中的二聚化MS2噬菌体外壳蛋白的最小发夹适体,并且被引入到指导分子中,诸如茎环和/或四环中。在这些实施方案中,将腺苷脱氨酶蛋白融合至MS2。然后将腺苷脱氨酶蛋白与CRISPR-Cas蛋白和相应的指导RNA共同递送。

[0398] 如本文所用,术语“AD官能化的CRISPR系统”是指核酸靶向和编辑系统,所述系统包含(a)CRISPR-Cas蛋白,更特别地是无催化活性的Cas13蛋白;(b)包含指导序列的指导分子;(c)腺苷脱氨酶蛋白或其催化结构域;其中所述腺苷脱氨酶蛋白或其催化结构域共价或非共价地连接至所述CRISPR-Cas蛋白或所述指导分子,或者适于在递送之后连接至所述CRISPR-Cas蛋白或所述指导分子;其中所述指导序列与靶序列基本上互补,但包含对应于针对脱氨靶向的A的非配对C,导致由所述指导序列和所述靶序列形成的RNA双链体中出现A-C错配。对于在真核细胞中应用,CRISPR-Cas蛋白和/或腺苷脱氨酶优选地带有NLS标签。

[0399] 在一些实施方案中,将组分(a)、组分(b)和组分(c)作为核糖核蛋白复合物递送至细胞。可以经由一种或多种脂质纳米粒子递送核糖核蛋白复合物。

[0400] 在一些实施方案中,将组分(a)、组分(b)和组分(c)作为一种或多种RNA分子诸如

编码CRISPR-Cas蛋白、腺苷脱氨酶蛋白和任选地衔接蛋白的一种或多种指导RNA和一种或多种mRNA分子递送至细胞。可以经由一种或多种脂质纳米粒子递送RNA分子。

[0401] 在一些实施方案中，将组分(a)、组分(b)和组分(c)作为一种或多种DNA分子递送至细胞。在一些实施方案中，一种或多种DNA分子包含在一种或多种载体诸如病毒载体(例如AAV)中。在一些实施方案中，所述一种或多种DNA分子包含可操作地配置成表达CRISPR-Cas蛋白、指导分子和腺苷脱氨酶蛋白或其催化结构域的一个或多个调控元件，任选地其中所述一个或多个调控元件包括诱导型启动子。

[0402] 在某些实施方案中，CRISPR-Cas蛋白是死亡Cas13。在一些实施方案中，死亡Cas13是死亡Cas13a蛋白，其包含在HEPN结构域中的一个或多个突变。在一些实施方案中，死亡Cas13a包含对应于韦德纤毛菌(LwaCas13a)中的R474A和R1046A的突变。在一些实施方案中，死亡Cas13是死亡Cas13b蛋白，所述死亡Cas13b蛋白包含源自动物溃疡伯格菌ATCC 43767的Cas13b蛋白的R116A、H121A、R1177A、H1182A中的一者或多者，或Cas13b直系同源物的其相应氨基酸位置中的突变。

[0403] 在一些实施方案中，指导分子能够与RNA序列内包含待脱氨基的腺嘌呤的靶序列杂交，以形成包含与所述腺嘌呤相对的非配对胞嘧啶的RNA双链体。在形成RNA双链体后，指导分子与Cas13蛋白形成复合物，并引导所述复合物结合目标靶RNA序列处的RNA多核苷酸。下文提供了有关AD官能化的CRISPR-Cas系统中的指导物方面的细节。

[0404] 在一些实施方案中，使用具有典型长度的Cas13指导RNA例如LawCas13与靶DNA形成RNA双链体。在一些实施方案中，使用比典型长度长的Cas13指导分子例如LawCas13a与靶DNA形成RNA双链体，包括在Cas13指导RNA-靶DNA复合物之外形成双链体。

[0405] 在至少第一设计中，AD官能化的CRISPR系统包含(a)融合或连接至CRISPR-Cas蛋白的腺苷脱氨酶，其中所述CRISPR-Cas蛋白是无催化活性的；和(b)指导分子，其包含被设计来在指导序列与靶序列之间所形成的RNA双链体中引入A-C错配的指导序列。在一些实施方案中，CRISPR-Cas蛋白和/或腺苷脱氨酶在N端或C端或这两者上带有NLS标签。

[0406] 在至少第二设计中，AD官能化的CRISPR系统包含(a)无催化活性的CRISPR-Cas蛋白；(b)指导分子，其包含被设计来在指导序列与靶序列之间所形成的RNA双链体中引入A-C错配的指导序列，和能够结合衔接蛋白(例如MS2外壳蛋白或PP7外壳蛋白)的适体序列(例如MS2 RNA基序或PP7 RNA基序)；和(c)融合或连接至衔接蛋白的腺苷脱氨酶，其中适体和衔接蛋白的结合将腺苷脱氨酶募集至在指导序列与靶序列之间所形成的RNA双链体，以在A-C错配的A处进行靶向脱氨。在一些实施方案中，衔接蛋白和/或腺苷脱氨酶在N端或C端或这两者上带有NLS标签。CRISPR-Cas蛋白也可以带有NLS标签。

[0407] 在一些实施方案中，使用具有典型长度(例如，约15-30nt)的Cas13指导RNA与靶RNA形成RNA双链体。在一些实施方案中，使用比典型长度(例如，>30nt)长的Cas13指导分子与靶RNA形成RNA双链体，包括在Cas13指导RNA-靶RNA复合物之外形成双链体。

[0408] 在至少第一设计中，CD官能化的CRISPR系统包含(a)融合或连接至CRISPR-Cas蛋白的胞苷脱氨酶，其中所述CRISPR-Cas蛋白是无催化活性的Cas13；和(b)指导分子，所述指导分子包含任选地被设计来(A)处于目标胞嘧啶的上游或下游，或(B)在指导序列与靶序列之间所形成的RNA双链体中引入C-A/U错配的指导序列。在一些实施方案中，CRISPR-Cas蛋白和/或胞苷脱氨酶在N端或C端或这两者上带有NES标签。

[0409] 在至少第二设计中，CD官能化的CRISPR系统包含(a)CRISPR-Cas蛋白，所述CRISPR-Cas蛋白是无催化活性的Cas13；和(b)指导分子，所述指导分子包含任选地被设计来(A)处于目标胞嘧啶的上游或下游，或(B)在指导序列与靶序列之间所形成的RNA双链体中引入C-A/U错配的指导序列，和能够结合至衔接蛋白(例如MS2外壳蛋白或PP7外壳蛋白)的适体序列(例如MS2 RNA基序或PP7 RNA基序)；和(c)融合或连接至衔接蛋白的胞苷脱氨酶，其中适体和衔接蛋白的结合将胞苷脱氨酶募集至在指导序列与靶序列之间所形成的RNA双链体，以在靶序列外部的C处或在任选的C-A/U错配的C处进行靶向脱氨。在一些实施方案中，衔接蛋白和/或胞苷脱氨酶在N端或C端或这两者上带有NES标签。CRISPR-Cas蛋白也可以带有NES标签。

[0410] 使用不同的适体和相应的衔接蛋白还可以实现正交基因编辑。在结合使用腺苷脱氨酶与胞苷脱氨酶用于正交基因编辑/脱氨的实例中，用不同的RNA环修饰靶向不同基因座的sgRNA，以分别募集MS2-腺苷脱氨酶和PP7-胞苷脱氨酶(或PP7-腺苷脱氨酶和MS2-胞苷脱氨酶)，导致目标靶基因座处的A或C分别发生正交脱氨。PP7是噬菌体假单胞菌的RNA结合外壳蛋白。同MS2一样，它结合特定的RNA序列和二级结构。PP7 RNA识别基序与MS2的不同。因此，PP7和MS2可以多路复用，以同时在不同的基因组基因座介导不同的效应。例如，靶向基因座A的sgRNA可以用MS2环修饰，从而募集MS2-腺苷脱氨酶；而靶向基因座B的另一sgRNA可以用PP7环修饰，从而募集PP7-胞苷脱氨酶。因此，在同一细胞中实现了正交的基因座特异性修饰。可以扩展该原理以并入其他正交RNA结合蛋白。

[0411] 在至少第三设计中，AD官能化的CRISPR系统包含(a)插入到CRISPR-Cas蛋白的内环或非结构化区中的腺苷脱氨酶，其中所述CRISPR-Cas蛋白无催化活性或者是切口酶；和(b)指导分子，其包含被设计来在指导序列与靶序列之间所形成的RNA双链体中引入A-C错配的指导序列。

[0412] 可以借助晶体结构鉴定适用于腺苷脱氨酶插入的CRISPR-Cas蛋白裂解位点。如果直系同源物与预期的CRISPR-Cas蛋白之间存在相对高程度的同源性，则可以使用直系同源物的晶体结构。

[0413] 裂解位置可以位于某一区域或环内。优选地，裂解位置在氨基酸序列的中断不会导致结构特征(例如α-螺旋或β-折叠)部分或全部破坏的地方出现。非结构化区(由于结构化程度不足以使其在晶体中“冻结”而未出现在晶体结构中的区域)通常是首选选项。非结构化区或外环内的位置可能不必完全是上面提供的数字，而是可能会变化，例如以上给出的位置任一侧的1个、2个、3个、4个、5个、6个、7个、8个、9个甚至10个氨基酸(取决于环的大小)，只要裂解位置仍落在外环的非结构化区内即可。

[0414] 本文所述的系统可用于靶向RNA多核苷酸序列内的特定腺嘌呤以进行脱氨。例如，指导分子可以与CRISPR-Cas蛋白形成复合物，并引导所述复合物结合目标RNA多核苷酸中的靶RNA序列。由于指导序列被设计成具有非配对C，因此在指导序列与靶序列之间所形成的RNA双链体包含A-C错配，该错配引导腺苷脱氨酶与同非配对C相对的A接触并使其脱氨基，从而将其转化为肌苷(I)。由于肌苷(I)碱基与C配对并且在细胞过程中具发挥类似G的功能，因此本文所述的A的靶向脱氨可用于校正不合需要的G-A和C-T突变，以及获得期望的A-G和T-C突变。

[0415] 在一些实施方案中，所述系统可用于体外RNA多核苷酸分子中的靶向脱氨。在一些

实施方案中，AD官能化的CRISPR系统是用于细胞内DNA分子中的靶向脱氨。细胞可以是真核细胞，诸如动物细胞、哺乳动物细胞、人类细胞或植物细胞。

[0416] 本发明还涉及一种通过使用AD官能化的CRISPR系统进行靶向脱氨来治疗或预防疾病的方法，其中A的脱氨补救了由含有病原性G→A或C→T点突变的转录物引起的疾病。可以利用本发明来治疗或预防的疾病的实例包括癌症、梅尔-戈林综合征、塞克尔综合征4、乔伯特综合征5、莱伯氏先天性黑蒙症10；夏科-马里-图思病2型；夏科-马里-图思病2型；乌谢尔综合征2C型；脊髓小脑性共济失调28；脊髓小脑性共济失调28；脊髓小脑性共济失调28；长QT综合征2；西奥格林-拉尔逊氏综合征；遗传性果糖尿病；遗传性果糖尿病；神经母细胞瘤；神经母细胞瘤；卡尔曼综合征1；卡尔曼综合征1；卡尔曼综合征1；异染性脑白质营养不良。

[0417] 本发明还涉及一种用于敲除或敲低基因的不合需要的活性的方法，其中在所述基因的转录物处A的脱氨导致功能丧失。例如，在一个实施方案中，通过AD官能化的CRISPR系统进行靶向脱氨可引起无义突变，使得内源性基因中出现提前终止密码子。这可能改变内源性基因的表达，并使得编辑后的细胞具有期望的性状。在另一个实施方案中，通过AD官能化的CRISPR系统进行靶向脱氨可引起非保守的错义突变，使得内源性基因中出现不同氨基酸残基的密码。这可能改变内源性基因的功能，并使得编辑后的细胞具有期望的性状。

[0418] 本发明还涉及一种通过使用AD官能化的CRISPR系统进行靶向脱氨而获得的修饰的细胞或其子代，其中与靶向脱氨之前的相应细胞相比，所述修饰的细胞在目标靶RNA序列中包含I或G而非A。修饰的细胞可以是真核细胞，诸如动物细胞、植物细胞、哺乳动物细胞或人类细胞。

[0419] 在一些实施方案中，修饰的细胞是治疗性T细胞，诸如适合于CAR-T疗法的T细胞。修饰可以使得在治疗性T细胞中出现一种或多种期望的性状，包括但不限于免疫检查点受体(例如PDA、CTLA4)的表达降低、HLA蛋白(例如B2M、HLA-A)的表达降低以及内源性TCR的表达降低。

[0420] 在一些实施方案中，修饰的细胞是产生抗体的B细胞。修饰可以使得在B细胞中出现一种或多种期望的性状，包括但不限于抗体产生增强。

[0421] 本发明还涉及一种修饰的非人类动物或修饰的植物。修饰的非人类动物可以是农场动物。修饰的植物可以是农作物。

[0422] 本发明还涉及一种用于细胞疗法的方法，所述方法包括向有需要的患者施用本文所述的修饰的细胞，其中所述修饰的细胞的存在补救了所述患者的疾病。在一个实施方案中，用于细胞疗法的修饰的细胞是能够识别和/或攻击肿瘤细胞的CAR-T细胞。在另一个实施方案中，用于细胞疗法的修饰的细胞是干细胞，诸如神经干细胞、间充质干细胞、造血干细胞或iPSC细胞。

[0423] 本发明另外涉及一种适用于修饰目标靶基因座中的腺嘌呤的工程化的非天然存在的系统，所述系统包含：包含指导序列的指导分子，或编码所述指导分子的核苷酸序列；CRISPR-Cas蛋白，或编码所述CRISPR-Cas蛋白的一个或多个核苷酸序列；腺苷脱氨酶蛋白或其催化结构域，或编码其的一个或多个核苷酸序列；其中所述腺苷脱氨酶蛋白或其催化结构域共价或非共价地连接至所述CRISPR-Cas蛋白或所述指导分子，或者适于在递送之后连接至所述CRISPR-Cas蛋白或所述指导分子；其中所述指导序列能够与目标RNA多核苷酸

内包含腺嘌呤的靶序列杂交,但是在对应于所述腺嘌呤的位置处包含胞嘧啶。

[0424] 本发明另外涉及一种适用于修饰目标靶基因座中的腺嘌呤的工程化的非天然存在的载体系统,所述载体系统包含一种或多种载体,所述一种或多种载体包含:第一调控元件,所述第一调控元件可操作地连接至一个或多个编码包含指导序列的指导分子的核苷酸序列;第二调控元件,所述第二调控元件可操作地连接至编码CRISPR-Cas蛋白的核苷酸序列;和任选地编码腺苷脱氨酶蛋白或其催化结构域的核苷酸序列,所述核苷酸序列受第一调控元件或第二调控元件的控制或者可操作地连接至第三调控元件;其中如果所述编码腺苷脱氨酶蛋白或其催化结构域的核苷酸序列可操作地连接至第三调控元件,则所述腺苷脱氨酶蛋白或其催化结构域适于在表达之后连接至所述指导分子或所述Crispr-Cas蛋白;其中所述指导序列能够与所述靶基因座内包含腺嘌呤的靶序列杂交,但在对应于所述腺嘌呤的位置上包含胞嘧啶;其中组分(a)、组分(b)和组分(c)位于所述系统的相同或不同载体上。

[0425] 本发明另外涉及包含本文所述的工程化的非天然存在的系统或载体系统的体外、离体或体内宿主细胞或细胞系或其子代。宿主细胞可以是真核细胞,诸如动物细胞、植物细胞、哺乳动物细胞或人类细胞。

[0426] 碱基切除修复抑制剂

[0427] 在一些实施方案中,所述系统还包含碱基切除修复(BER)抑制剂。不希望受任何特定理论的束缚,对I:T配对存在的细胞DNA修复反应可能导致细胞中核碱基编辑效率的降低。烷基腺嘌呤DNA糖基化酶(也称为DNA-3-甲基腺嘌呤糖基化酶、3-烷基腺嘌呤DNA糖基化酶或N-甲基嘌呤DNA糖基化酶)催化细胞DNA中次黄嘌呤的去除,这可能启动碱基切除修复,结果I:T配对逆转成A:T配对。

[0428] 在一些实施方案中,BER抑制剂是烷基腺嘌呤DNA糖基化酶的抑制剂。在一些实施方案中,BER抑制剂是人类烷基腺嘌呤DNA糖基化酶的抑制剂。在一些实施方案中,BER抑制剂是多肽抑制剂。在一些实施方案中,BER抑制剂是结合次黄嘌呤的蛋白质。在一些实施方案中,BER抑制剂是结合DNA中的次黄嘌呤的蛋白质。在一些实施方案中,BER抑制剂是无催化活性的烷基腺嘌呤DNA糖基化酶蛋白或其结合结构域。在一些实施方案中,BER抑制剂是不从DNA中切除次黄嘌呤的无催化活性的烷基腺嘌呤DNA糖基化酶蛋白或其结合结构域。能够抑制(例如,在空间上阻断)烷基腺嘌呤DNA糖基化酶碱基切除修复酶的其他蛋白质在本公开的范围之内。另外,阻断或抑制碱基切除修复的任何蛋白质也在本公开的范围之内。

[0429] 不希望受任何特定理论的束缚,可通过结合编辑链、阻断编辑碱基、抑制烷基腺嘌呤DNA糖基化酶、抑制碱基切除修复、保护编辑碱基和/或促进未编辑链固定的分子抑制碱基切除修复。据信,使用本文所述的BER抑制剂可以提高能够催化A至I变化的腺苷脱氨酶的编辑效率。

[0430] 因此,在以上论述的AD官能化的CRISPR系统的第一设计中,可以将CRISPR-Cas蛋白或腺苷脱氨酶融合或连接至BER抑制剂(例如烷基腺嘌呤DNA糖基化酶的抑制剂)。在一些实施方案中,BER抑制剂可以包含在以下结构之一中(nCas13=Cas13切口酶;dCas13=死亡Cas13):[AD]-[任选的接头]-[nCas13/dCas13]-[任选的接头]-[BER抑制剂];[AD]-[任选的接头]-[BER抑制剂]-[任选的接头]-[nCas13/dCas13];[BER抑制剂]-[任选的接头]-[AD]-[任选的接头]-[nCas13/dCas13];[BER抑制剂]-[任选的接头]-[nCas13/dCas13]-

[任选的接头]-[AD];[nCas13/dCas13]-[任选的接头]-[AD]-[任选的接头]-[BER抑制剂];[nCas13/dCas13]-[任选的接头]-[BER抑制剂]-[任选的接头]-[AD]。

[0431] 类似地,在以上论述的AD官能化的CRISPR系统的第二设计中,可以将CRISPR-Cas蛋白、腺苷脱氨酶或衔接蛋白融合或连接至BER抑制剂(例如烷基腺嘌呤DNA糖基化酶的抑制剂)。在一些实施方案中,BER抑制剂可以包含在以下结构之一中(nCas13=Cas13切口酶;dCas13=死亡Cas13):[nCas13/dCas13]-[任选的接头]-[BER抑制剂];[BER抑制剂]-[任选的接头]-[nCas13/dCas13];[AD]-[任选的接头]-[衔接子]-[任选的接头]-[BER抑制剂];[AD]-[任选的接头]-[BER抑制剂]-[任选的接头]-[衔接子];[BER抑制剂]-[任选的接头]-[AD]-[任选的接头]-[衔接子];[BER抑制剂]-[任选的接头]-[衔接子]-[任选的接头]-[AD];[衔接子]-[任选的接头]-[AD]-[任选的接头]-[BER抑制剂];[衔接子]-[任选的接头]-[BER抑制剂]-[任选的接头]-[AD]。

[0432] 在以上论述的AD官能化的CRISPR系统的第三设计中,可以将BER抑制剂插入CRISPR-Cas蛋白的内环或非结构化区中。

[0433] 对细胞核的靶向

[0434] 在一些实施方案中,本发明的方法涉及修饰目标靶基因座中的腺嘌呤,其中所述靶基因座处于细胞内。为了改善在本发明的方法中使用的CRISPR-Cas蛋白和/或腺苷脱氨酶蛋白或其催化结构域对细胞核的靶向,可能有利的是为这些组分之一或两者提供一个或多个核定位序列(NLS)。

[0435] 在优选的实施方案中,在本发明的情形中使用的NLS与蛋白质是异源的。NLS的非限制性实例包括来源于以下的NLS序列:SV40病毒大T抗原的NLS,其具有氨基酸序列PKKKRKV(SEQ ID NO:17)或PKKKRKVEAS(SEQ ID NO:18);来自核质蛋白的NLS(例如核质蛋白两分NLS,其具有序列KRPAATKKAGQAKKKK(SEQ ID NO:19));c-myc NLS,其具有氨基酸序列PAAKRVKLD(SEQ ID NO:20)或RQRRNELKRSP(SEQ ID NO:21);hRNPA1 M9 NLS,其具有序列NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY(SEQ ID NO:22);来自输入蛋白-α的IBB结构域的序列RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV(SEQ ID NO:23);肌瘤T蛋白的序列VSRKRPRP(SEQ ID NO:24)和PPKKARED(SEQ ID NO:25);人类p53的序列PQPKKKPL(SEQ ID NO:26);小鼠c-abl IV的序列SALIKKKKKMAP(SEQ ID NO:27);流感病毒NS1的序列DRLRR(SEQ ID NO:28)和PKQKKRK(SEQ ID NO:29);肝炎病毒δ抗原的序列RKLKKKIKKL(SEQ ID NO:30);小鼠Mx1蛋白的序列REKKKFLKRR(SEQ ID NO:31);人类聚(ADP-核糖)聚合酶的序列KRKGDEVDGVDEVAKKKSKK(SEQ ID NO:32);以及类固醇激素受体(人类)糖皮质激素的序列RKCLQAGMNLEARKTKK(SEQ ID NO:33)。通常,一个或多个NLS具有足够的强度来驱动可检测量的DNA靶向Cas蛋白在真核细胞的细胞核中的积聚。通常,核定位活性的强度可以源自CRISPR-Cas蛋白中NLS的数量、所使用的特定NLS或这些因素的组合。可以通过任何合适的技术来检测细胞核中的积聚。例如,(如)结合用于检测细胞核的位置的手段(例如对细胞核具有特异性的染色剂诸如DAPI),可以将可检测的标记物融合至核酸靶向蛋白,从而可以可视化细胞内的位置。也可以从细胞中分离细胞核,接着可以通过用于检测蛋白质的任何合适的方法(诸如免疫组织化学、蛋白质印迹或酶活性测定)来分析细胞核的内容。也可以通过针对靶序列处的核酸靶向复合物形成的作用的测定(例如针对脱氨酶活性的测定),或针对受DNA靶向复合物形成/或DNA靶向影响而改变的基因表达的测定),如与未暴露于

CRISPR-Cas蛋白和脱氨酶蛋白的对照物，或暴露于缺少一个或多个NLS的CRISPR-Cas和/或脱氨酶蛋白的对照物相比，间接确定细胞核中的积聚。

[0436] 可以为CRISPR-Cas和/或腺苷脱氨酶蛋白提供1个或多个，诸如2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个异源NLS。在一些实施方案中，蛋白质在氨基末端处或附近包含约或多于约1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个NLS，在羧基末端处或附近包含约或多于约1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或多个NLS，或这些的组合(例如在氨基末端处为零或至少一个或多个NLS，或在羧基末端处为零或一个或多个NLS)。当存在多于一个NLS时，可以彼此独立地选择每个NLS，使得单个NLS可以多于一个拷贝存在和/或与一个或多个其他NLS组合以一个或多个拷贝存在。在一些实施方案中，当NLS的最近的氨基酸沿着多肽链距N端或C端约1个、2个、3个、4个、5个、10个、15个、20个、25个、30个、40个、50个或更多个氨基酸内时，认为NLS接近N端或C端。在CRISPR-Cas蛋白的优选实施方案中，NLS附接至蛋白质的C端。

[0437] 在本文提供的方法的某些实施方案中，CRISPR-Cas蛋白和脱氨酶蛋白作为单独的蛋白质被递送至细胞或在细胞内表达。在这些实施方案中，CRISPR-Cas和脱氨酶蛋白各自可以被提供有一个或多个如本文所述的NLS。在某些实施方案中，CRISPR-Cas和脱氨酶蛋白作为融合蛋白被递送至细胞或在细胞内表达。在这些实施方案中，CRISPR-Cas和脱氨酶蛋白之一或两者都被提供有一个或多个NLS。如上所述，当腺苷脱氨酶融合至衔接蛋白(诸如MS2)时，可以在衔接蛋白上提供一个或多个NLS，条件是这不干扰适体结合。在特定实施方案中，一个或多个NLS序列还可以充当腺苷脱氨酶与CRISPR-Cas蛋白之间的接头序列。

[0438] 在某些实施方案中，本发明的指导物包含针对衔接蛋白的特异性结合位点(例如适体)，所述衔接蛋白可以连接或融合至腺苷脱氨酶或其催化结构域。当这样的指导物形成CRISPR复合物(即结合至指导物和靶标的CRISPR-Cas蛋白)时，衔接蛋白结合腺苷脱氨酶或其催化结构域，并且与所述衔接蛋白缔合的腺苷脱氨酶或其催化结构域被定位成有利于属性化的功能生效的空间取向。

[0439] 技术人员将理解，对允许衔接子+腺苷脱氨酶结合但未正确定位衔接子+腺苷脱氨酶(例如由于CRISPR复合物的三维结构内的空间位阻)的指导物的修饰是未预期的修饰。如本文所述，一种或多种修饰的指导物可在四环、茎环1、茎环2或茎环3处修饰，优选地在四环或茎环2中，且最优选地在四环处环和茎环2两者中修饰。

[0440] 正交无催化活性的CRISPR-Cas蛋白的用途

[0441] 在特定实施方案中，将Cas13切口酶与正交无催化活性的CRISPR-Cas蛋白结合使用以提高所述Cas13切口酶的效率(如Chen等人2017,Nature Communications8:14958;doi:10.1038/ncomms14958中所述)。更特别地，通过与所述系统中所用的Cas13切口酶不同的PAM识别位点来表征正交无催化活性的CRISPR-Cas蛋白，并且选择相应的指导序列以结合至邻近所述系统的Cas13切口酶的靶序列的靶序列。如在本发明的情形中使用的正交无催化活性的CRISPR-Cas蛋白不形成所述系统的一部分，而是仅用于增加所述Cas13切口酶的效率，并且与如在本领域中针对所述CRISPR-Cas蛋白描述的标准指导分子结合使用。在特定实施方案中，所述正交无催化活性的CRISPR-Cas蛋白是死亡CRISPR-Cas蛋白，即包含一种或多种消除所述CRISPR-Cas蛋白的核酸酶活性的突变。在特定实施方案中，无催化活性的正交CRISPR-Cas蛋白被提供有两种或更多种能够与邻近Cas13切口酶的靶序列的靶序

列杂交的指导分子。在特定实施方案中，至少两种指导分子用于靶向所述无催化活性的CRISPR-Cas蛋白，其中至少一种指导分子能够与所述系统的Cas13切口酶的靶序列的5”靶序列杂交并且至少一种指导分子能够与Cas13切口酶的靶序列的3’靶序列杂交，由此所述一个或多个靶序列可以处于与Cas13切口酶的靶序列相同或相反的DNA链上。在特定实施方案中，选择正交无催化活性的CRISPR-Cas蛋白的一种或多种指导分子的指导序列，以使靶序列邻近用于AD官能化的CRISPR的靶向(即Cas13切口酶的靶向)的指导分子的靶序列。在特定实施方案中，正交无催化活性的CRISPR-Cas酶的一个或多个靶序列各自与Cas13切口酶的靶序列分开大于5个但小于450个碱基对。与正交无催化活性的CRISPR-Cas蛋白一起使用的指导物的靶序列与所述系统的靶序列之间的最佳距离可以由技术人员确定。在特定实施方案中，正交CRISPR-Cas蛋白是II类II型CRISPR蛋白。在特定实施方案中，正交CRISPR-Cas蛋白是II类V型CRISPR蛋白。在特定实施方案中，无催化活性的正交CRISPR-Cas蛋白在特定实施方案中，无催化活性的正交CRISPR-Cas蛋白已经如本文别处所述进行了修饰，以改变其PAM特异性。在特定实施方案中，Cas13蛋白切口酶是这样的一种切口酶：其本身在人类细胞中活性有限，但与无活性的正交CRISPR-Cas蛋白和一种或多种相应的邻近指导物组合可确保所需的切口酶活性。

[0442] Cas13效应蛋白复合物可以递送功能效应子

[0443] CRISPR-Cas13介导的敲除能通过RNA水平上的突变消除表达，不同与此的是，CRISPR-Cas13敲除允许暂时减少基因表达，方式是使用人工转录因子，例如经由使Cas13蛋白的一个或多个切割结构域中的残基突变从而产生无催化活性的Cas13蛋白。无催化活性的Cas13与指导RNA或crRNA复合，并定位至该指导RNA或crRNA的靶向结构域所指定的RNA序列，但是它不能切割靶标。无活性的Cas13蛋白与效应子结构域例如转录阻抑结构域融合，使得能够将该效应子募集至由指导RNA指定的任何位点。

[0444] 在一些实施方案中，dCas13可用于递送转录因子或其活性结构域。dCas13可以与转录因子或其活性结构域融合。所得复合物可以在核酸，例如指导序列的指导下递送。

[0445] 裂解蛋白质

[0446] 注意，在这种情形下，并且更一般地对于本文所述的各种应用，可以设想使用靶向RNA的效应蛋白的裂解型式。实际上，这不仅可以允许增加的特异性，而且对于递送也可以是有利的。可以在Cas13酶的两个部分基本上包含功能性Cas13的意义上将Cas13(例如Cas13b)裂解。理想的是，裂解应始终使所述一个或多个催化结构域不受影响。Cas13c可以充当核酸酶，或者它可以是死亡的Cas13c，所述死亡的Cas13c由于其催化结构域中通常发生一个或多个突变，它实质上是一种具有很少催化活性或没有催化活性的RNA结合蛋白。

[0447] 裂解Cas13的每一半可以融合至二聚化配偶体。举例来说但不限于，采用雷帕霉素敏感的二聚化结构域允许产生化学诱导型裂解Cas13，用于暂时控制Cas13的活性。因此，可以通过将Cas13裂解成两个片段来使其可在化学上诱导，并且雷帕霉素敏感的二聚化结构域可用于Cas13的受控重组。可以将裂解Cas13c的两个部分视为裂解Cas13c的N’端部分和C’端部分。融合通常发生在Cas13的裂解点。换句话说，裂解Cas13c的N’端部分的C’端融合至一个二聚体一半部分，而C’端部分的N’端融合至另一个二聚体一半部分。

[0448] 从断裂是新产生的含义上讲，不必裂解Cas13。裂解点通常是在计算机上设计的，并且克隆到构建体中。裂解Cas13c的两个部分，即N’端和C’端部分一起形成全Cas13，所述

全Cas13优选地包含野生型氨基酸(或编码它们的核苷酸)的至少70％或更多,优选野生型氨基酸(或编码它们的核苷酸)的至少80％或更多,优选至少90％或更多,优选至少95％或更多,且最优选至少99％或更多。一些修整是可能的,并且设想了突变体。非功能结构域可能会被完全去除。重要的是可以将这两个部分放在一起,并且恢复或重构所需的Cas13c功能。二聚体可以是同源二聚体或异源二聚体。

[0449] 在一些情况下,第一裂解Cas13效应蛋白和第二裂解Cas13效应蛋白能够彼此融合形成催化活性Cas13效应蛋白。融合可能是可诱导的。

[0450] 优化的功能RNA靶向系统

[0451] 因此,在一方面本发明提供了一种用于将功能组分特异性地递送至RNA环境的系统。可以使用包含允许不同组分特异性地靶向RNA的本发明的RNA靶向效应蛋白的CRISPR系统确保这一点。更特别地,此类组分包括激活因子或阻抑因子,诸如RNA翻译、降解等的激活因子或阻抑因子。

[0452] 根据一方面,本发明提供了非天然存在的或工程化的组合物,所述组合物包含含有能够与细胞中的目标靶序列杂交的指导序列的指导RNA或crRNA,其中所述指导RNA或crRNA通过插入结合衔接蛋白的一个或多个不同的RNA序列而加以修饰。在特定实施方案中,RNA序列可以结合至两个或更多个衔接蛋白(例如适体),并且其中每个衔接蛋白与一个或多个功能结构域缔合。当存在超过一个功能结构域时,这些功能结构域可以相同或不同,例如两个是相同的,或者两个是不同的激活因子或阻抑因子。在一方面,本发明提供了本文讨论的组合物,其中所述一个或多个功能结构域附接至所述RNA靶向酶,使得在结合至所述靶RNA时,所述功能结构域处于允许所述功能结构域以其属性化的功能发挥作用的空间取向;在一方面,本发明提供了本文讨论的组合物,其中所述组合物包含具有至少三个功能结构域的CRISPR-Cas13复合物,其中至少一个功能结构域与所述RNA靶向酶缔合,并且至少两个功能结构域与gRNA或crRNA缔合。

[0453] CRISPR开发和使用

[0454] 可以基于以下文章中所陈述的CRISPR-Cas开发和使用的方面,特别是涉及CRISPR蛋白复合物的递送以及RNA指导的内切核酸酶在细胞和生物体中的使用的方面进一步说明和扩展本发明:

[0455] ➢Multiplex genome engineering using CRISPR-Cas systems.Cong,L.,Ran,F.A.,Cox,D.,Lin,S.,Barretto,R.,Habib,N.,Hsu,P.D.,Wu,X.,Jiang,W.,Marraffini,L.A.,&Zhang,F.Science Feb 15;339(6121):819-23(2013);

[0456] ➢RNA-guided editing of bacterial genomes using CRISPR-Cas systems.Jiang W.,Bikard D.,Cox D.,Zhang F,Marraffini LA.Nat Biotechnol Mar;31(3):233-9(2013);

[0457] ➢One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR-Cas-Mediated Genome Engineering.Wang H.,Yang H.,Shivalila CS.,Dawlaty MM.,Cheng AW.,Zhang F.,Jaenisch R.Cell May 9;153(4):910-8(2013);

[0458] ➢Optical control of mammalian endogenous transcription and epigenetic states.Konermann S,Brigham MD,Trevino AE,Hsu PD,Heidenreich M,Cong L,Platt RJ,Scott DA,Church GM,Zhang F.Nature.Aug 22;500(7463):472-6.doi:

10.1038/Nature12466.Epub 2013Aug 23(2013)；

[0459] ➢Double Nicking by RNA-Guided CRISPR Cas9for Enhanced Genome Editing Specificity.Ran,FA.,Hsu,PD.,Lin,CY.,Gootenberg,JS.,Konermann,S.,Trevino,AE.,Scott,DA.,Inoue,A.,Matoba,S.,Zhang,Y.,&Zhang,F.Cell Aug 28.pii:S0092-8674(13)01015-5(2013-A)；

[0460] ➢DNA targeting specificity of RNA-guided Cas9 nucleases.Hsu,P.,Scott,D.,Weinstein,J.,Ran,FA.,Konermann,S.,Agarwala,V.,Li,Y.,Fine,E.,Wu,X.,Shalem,O.,Cradick,TJ.,Marraffini,LA.,Bao,G.,&Zhang,F.Nat Biotechnoldoi:10.1038/nbt.2647(2013)；

[0461] ➢Genome engineering using the CRISPR-Cas9 system.Ran,FA.,Hsu,PD.,Wright,J.,Agarwala,V.,Scott,DA.,Zhang,F.Nature Protocols Nov;8(11):2281-308(2013-B)；

[0462] ➢Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells.Shalem,O.,Sanjana,NE.,Hartenian,E.,Shi,X.,Scott,DA.,Mikkelson,T.,Heckl,D.,Ebert,BL.,Root,DE.,Doench,JG.,Zhang,F.Science Dec 12.(2013)；

[0463] ➢Crystal structure of cas9 in complex with guide RNA and target DNA.Nishimasu,H.,Ran,FA.,Hsu,PD.,Konermann,S.,Shehata,SI.,Dohmae,N.,Ishitani,R.,Zhang,F.,Nureki,O.Cell Feb 27,156(5):935-49(2014)；

[0464] ➢Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells.WuX.,Scott DA.,Kriz AJ.,Chiu AC.,Hsu PD.,Dadon DB.,Cheng AW.,Trevino AE.,Konermann S.,Chen S.,Jaenisch R.,Zhang F.,Sharp PA.Nat Biotechnol.Apr 20.doi:10.1038/nbt.2889(2014)；

[0465] ➢CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling.Platt RJ,Chen S,Zhou Y,Yim MJ,Swiech L,Kempton HR,Dahlman JE,Parnas O,EisenhaureTM,Jovanovic M,Graham DB,Jhunjhunwala S,Heidenreich M,Xavier RJ,Langer R,Anderson DG,Hacohen N,Regev A,Feng G,Sharp PA,Zhang F.Cell 159(2):440-455DOI:10.1016/j.cell.2014.09.014(2014)；

[0466] ➢Development and Applications of CRISPR-Cas9 for Genome Engineering,Hsu PD,Lander ES,Zhang F.,Cell.Jun 5;157(6):1262-78(2014).

[0467] ➢Genetic screens in human cells using the CRISPR-Cas9 system,Wang T,Wei JJ,Sabatini DM,Lander ES.,Science.January 3;343(6166):80-84.doi:10.1126/science.1246981(2014)；

[0468] ➢Rational design of highly active sgRNAs for CRISPR-Cas9-mediated geneinactivation,Doench JG,Hartenian E,Graham DB,Tothova Z,Hegde M,Smith I,Sullender M,Ebert BL,Xavier RJ,Root DE.,(published online 3 September 2014)NatBiotechnol.Dec;32(12):1262-7(2014)；

[0469] ➢In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9,Swiech L,Heidenreich M,Banerjee A,Habib N,Li Y,Trombetta J,Sur M,Zhang F.,(published online 19 October 2014)Nat Biotechnol.Jan;33(1):102-6

(2015)；

[0470] ➢Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex,Konermann S,Brigham MD,Trevino AE,Joung J,Abudayyeh OO,Barcena C,Hsu PD,Habib N,Gootenberg JS,Nishimasu H,Nureki O,Zhang F.,Nature.Jan29;517(7536):583-8(2015).

[0471] ➢A split-Cas9 architecture for inducible genome editing and transcriptionmodulation,Zetsche B,Volz SE,Zhang F.,(published online 02 February 2015)NatBiotechnol.Feb;33(2):139-42(2015)；

[0472] ➢Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis,Chen S,Sanjana NE,Zheng K,Shalem O,Lee K,Shi X,Scott DA,Song J,Pan JQ,Weissleder R,Lee H,Zhang F,Sharp PA.Cell 160,1246-1260,March 12,2015 (multiplex screen in mouse),and

[0473] ➢In vivo genome editing using Staphylococcus aureus Cas9,Ran FA,Cong L,YanWX,Scott DA,Gootenberg JS,Kriz AJ,Zetsche B,Shalem O,Wu X,Makarova KS,Koonin EV,Sharp PA,Zhang F.,(published online 01 April 2015),Nature.Apr9;520(7546):186-91(2015).

[0474] ➢Shalem et al.,“High-throughput functional genomics using CRISPR-Cas9,”Nature Reviews Genetics 16,299-311(May 2015).

[0475] ➢Xu et al.,“Sequence determinants of improved CRISPR sgRNA design,”Genome Research 25,1147-1157(August 2015).

[0476] ➢Parnas et al.,“A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks,”Cell 162,675-686(July 30,2015).

[0477] ➢Ramanan et al.,CRISPR-Cas9cleavage of viral DNA efficiently suppresses hepatitis B virus,”Scientific Reports 5:10833.doi:10.1038/srep10833(June 2,2015)

[0478] ➢Nishimasu et al.,Crystal Structure of Staphylococcus aureus Cas9,”Cell 162,1113-1126(Aug.27,2015)

[0479] ➢BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis,Canver et al.,Nature 527(7577):192-7(Nov.12,2015)doi:10.1038/nature15521.Epub2015Sep 16.

[0480] ➢Cas13Is a Single RNA-Guided Endonuclease of a Class 2CRISPR-Cas System,Zetsche et al.,Cell 163,759-71(Sep 25,2015).

[0481] ➢Discovery and Functional Characterization of Diverse Class 2CRISPR-Cas Systems,Shmakov et al.,Molecular Cell,60(3),385-397doi:10.1016/j.molcel.2015.10.008Epub October 22,2015.

[0482] ➢Rationally engineered Cas9nucleases with improved specificity,Slaymaker et al.,Science 2016Jan 1 351(6268):84-88doi:10.1126/science.aad5227.Epub 2015Dec 1.

[0483] ➢Gao et al,“Engineered Cas13Enzymes with Altered PAM Specificities,”

bioRxiv091611;doi:http://dx.doi.org/10.1101/091611(Dec.4,2016).

[0484] 将它们各自以引用方式并入本文,可以在本发明的实践中考虑,并在下文简要论述:

[0485] ➢Cong等人基于嗜热链球菌(Streptococcus thermophilus)Cas9还有化脓链球菌(Streptococcus pyogenes)Cas9两者工程化了II型CRISPR-Cas系统以用于在真核细胞中使用,并且证实了Cas9核酸酶可以通过短RNA引导以诱导在人类和小鼠细胞中DNA的精确切割。他们的研究进一步显示Cas9在转化成一种切口酶时可以用来以最低诱变活性促进在真核细胞中的同源定向修复。另外,他们的研究证实多个指导序列可以被编码进单一CRISPR阵列中以使得能够在哺乳动物基因组内的内源性基因组基因座位点处同时编辑若干,证实了RNA指导的核酸酶技术的容易可编程性和广泛可应用性。这种使用RNA以编程细胞内序列特异性DNA切割的能力定义了新一类的基因组工程化工具。这些研究进一步显示,其他CRISPR基因座可能是可移植入哺乳动物细胞中的,并且还可以介导哺乳动物基因组切割。重要地,可以设想的是CRISPR-Cas系统的若干方面可以进一步改善以增加其效率和多功能性。

[0486] ➢Jiang等人使用成簇的规律间隔的短回文重复序列(CRISPR)-关联的Cas9内切核酸酶,与双-RNA复合,以在肺炎链球菌(Streptococcus pneumoniae)和大肠杆菌(Escherichia coli)的基因组中引入精确的突变。该方法依赖于在靶向基因组位点处的双-RNA:Cas9-引导的切割,以杀死未突变的细胞,并且回避对选择性标记物或反选择系统的需要。该研究报道通过改变短CRISPR RNA(crRNA)的序列以使单一和多个核苷酸变化被携带在编辑模板上而重编程双-RNA:Cas9特异性。该研究显示,同时使用两种crRNA使得多重诱变成为可能。另外,当该方法与重组工程组合使用时,在肺炎链球菌中使用描述的方法回收的接近100%的细胞包含所需的突变,并且在大肠杆菌中回收的65%包含突变。

[0487] ➢Wang等人(2013)使用CRISPR-Cas系统用于一步生成携带多基因中突变的小鼠,所述小鼠传统上是以多步通过在胚胎干细胞中的连续重组和/或小鼠的与单一突变的耗时性杂交生成的。CRISPR-Cas系统将大大加速功能上丰富的基因和上位基因相互作用的体内研究。

[0488] ➢Konermann等人(2013)解决了在本领域中对通用和稳固技术的需要,其使得能够基于CRISPR Cas9酶以及还有转录激活因子样效应子对DNA-结合结构域进行光调节和化学调节

[0489] ➢Ran等人(2013-A)描述了将Cas9切口酶突变体与配对的指导RNA相组合以引入靶向的双链断裂的方法。这解决了以下问题:来自微生物CRISPR-Cas系统的Cas9核酸酶通过指导序列而靶向特异性基因组基因座,所述指导序列可以耐受与该DNA靶标的某些错配并由此促进不希望的脱靶诱变。因为基因组中的单独切口以高保真性被修复,所以同时经由适当偏移指导RNA而形成切口对于双链断裂是必需的,并且所述切口形成扩大了特异性识别的碱基的数目以用于靶标切割。作者证实了使用配对的切口形成可以降低在细胞系中的脱靶活性50至1,500倍,并且从而促进在小鼠受精卵中的基因敲除而不牺牲中靶切割效率。这个通用策略使得多种多样的要求高特异性的基因组编辑应用成为可能。

[0490] ➢Hsu等人(2013)表征了在人类细胞中SpCas9靶向特异性以告知靶位点的选择并避免脱靶效应。该研究评估了在293T和293FT细胞中>100个预测的基因组脱基因座处>700

个指导RNA变体和SpCas9诱导的插入缺失突变水平。这些作者示出SpCas9以序列依赖性方式耐受指导RNA与靶DNA之间在不同位置处的错配，对错配的数目、位置和分布敏感。作者进一步表明，SpCas9介导的切割不受DNA甲基化的影响，并且SpCas9和指导RNA的剂量可被滴定为使得脱靶修饰最小化。另外，为了促进哺乳动物基因组工程化应用，这些作者报道提供基于网络的软件工具以指导靶序列的选择和验证连同脱靶分析。

[0491] ➢Ran等人(2013-B)描述了用于在哺乳动物细胞中经由非同源末端连接(NHEJ)或同源定向修复(HDR)的Cas9-介导的基因组编辑、连同产生修饰的细胞系(以用于下游功能研究)的一组工具。为了最小化脱靶切割，这些作者进一步描述了一种双-切口策略，使用的是Cas9切口酶突变体与配对的指导RNA。由这些作者提供的方案经实验得出用于选择靶位点、评估切割效率和分析脱靶活性的指南。这些研究显示，以靶设计开始，基因修饰可以在少至1-2周内实现，并且修饰的克隆细胞系可以在2-3周内得以衍生。

[0492] ➢Shalem等人描述了一种新的在全基因组范围上探询基因功能的方式。他们的研究显示，递送基因组范围的CRISPR-Cas9敲除(GeCKO)文库利用64,751个独特的指导序列靶向18,080个基因，这使得在人类细胞中阴性和阳性选择筛选两者成为可能。首先，这些作者显示，使用该GeCKO文库来鉴定癌症和多能干细胞中对于细胞活力至关重要的基因。接着，在黑素瘤模型中，这些作者针对基因进行筛选，这些基因的丧失涉及对维罗非尼(一种抑制突变体蛋白激酶BRAF的治疗剂)的抗性。他们的研究显示，最高级候选物包括先前验证的基因NF1和MED12连同新颖的命中物NF2、CUL3、TADA2B和TADA1。这些作者观察到在靶向相同基因的独立指导RNA之间的高水平的一致性以及高比率的命中确认，并且因此证实了采用Cas9进行基因组范围筛选的前景。

[0493] ➢Nishimasu等人以2.5A°分辨率报道了与sgRNA及其靶DNA复合的化脓链球菌Cas9的晶体结构。该结构揭示了一种由靶识别和核酸酶叶片组成的两叶片架构，其将sgRNA:DNAn RNA双链体容纳在它们的界面处的带正电的凹槽中。然而识别叶片对于结合sgRNA和DNA是至关重要的，核酸酶叶片含有HNH和RuvC核酸酶结构域，这些结构域适合地被定位为分别用于靶DNA的互补和非互补链的切割。核酸酶叶还含有负责与原间隔区邻近基序(PAM)相互作用的羧基末端结构域。这种高分辨率结构和伴随的功能分析已经揭示了RNA指导的由Cas9进行的DNA靶向的分子机制，由此为合理设计新的通用基因组编辑技术做好准备。

[0494] ➢Wu等人标定了在小鼠胚胎干细胞(mESC)中，来自化脓链球菌的无催化活性的Cas9(dCas9)(加载有单一指导RNA(sgRNA))的全基因组结合位点。这些作者显示，测试的四种sgRNA中的每一种将dCas9靶向至数十与数千个之间的基因组位点，这些基因组位点频繁地通过sgRNA中的5-核苷酸种子区和NGG原间隔区邻近基序(PAM)表征。染色质不可接近性降低了dCas9与具有匹配种子序列的其他位点的结合；因此70％的脱靶位点是与基因相关联的。这些作者显示，在用催化活性的Cas9转染的mESC中对295个dCas9结合位点的靶向测序鉴定出超过背景水平的仅一个突变位点。这些作者提出了一种针对Cas9结合和切割的两态模型，其中种子匹配触发了结合但是需要与靶DNA的广泛配对用于切割。

[0495] ➢Platt等人建立了Cre依赖性Cas9敲入式小鼠。这些作者证实了在神经元、免疫细胞和内皮细胞中，使用腺相关病毒(AAV)-、慢病毒-、或粒子介导的指导RNA递送进行体内以及离体基因组编辑。

[0496] ➢Hsu等人(2014)是综述文章,其总体论述了CRISPR-Cas9从酸奶到基因组编辑的历史,包括细胞的遗传筛选。

[0497] ➢Wang等人(2014)涉及使用基因组范围的慢病毒单一指导RNA(sgRNA)文库进行的适于阳性和阴性选择两者的合并的功能丧失遗传筛选方法。

[0498] ➢Doench等人创建了sgRNA池,覆盖六个内源性小鼠基因和三个内源性人类基因的一组的全部可能靶位点,并且通过抗体染色和流式细胞术定量地评定了这些sgRNA产生其靶基因的无效等位基因的能力。这些作者显示PAM的优化改善了活性并且还提供了用于设计sgRNA的在线工具。

[0499] ➢Swiech等人证实了AAV介导的SpCas9基因组编辑可以使能进行脑中的基因功能的反向遗传学研究。

[0500] ➢Konermann等人(2015)论述了在有和没有接头的情况下,在指导物(诸如茎或四环)上的适当位置处,附接多种效应结构域的能力,这些效应结构域例如转录激活因子、功能和表观基因组调节子。

[0501] ➢Zetsche等人证明了Cas9酶可以分为两部分,因此Cas9的激活组件可以得到控制。

[0502] ➢Chen等人涉及通过证实以下进行多重筛选:在小鼠中基因组光度的体内CRISPR-Cas9筛选揭示了调节肺转移的基因。

[0503] ➢Ran等人(2015)涉及SaCas9以及其编辑基因组的能力,并且证实不能从生物化学测定外推。

[0504] ➢Shalem等人(2015)描述了无催化活性的Cas9(dCas9)融合用于综合地阻抑(CRISPRi)或激活(CRISPRa)表达的方式,示出使用Cas9用于基因组范围的筛选(包括阵列和合并筛选)的进展、使基因组基因座失活的敲除方法、以及调节转录活性的策略。

[0505] ➢Xu等人等人(2015)评定了在基于CRISPR的筛选中促成单一指导RNA(sgRNA)效率的DNA序列特征。这些作者探索了CRISPR-Cas9敲除的效率以及在切割位点处的核苷酸优选性。这些作者还发现对于CRISPRi/a的序列优选性基本上不同于对于CRISPR-Cas9敲除的序列优选性。

[0506] ➢Parnas等人(2015)将合并的全基因组CRISPR-Cas9文库引入树突细胞(DC)中,以鉴定控制由细菌脂多糖(LPS)对肿瘤坏死因子(Tnf)的诱导的基因。对Tlr4信号传导的已知调节子和先前未知的候选物进行鉴定,并根据对于对LPS的典型反应的不同效果分成三个功能模块。

[0507] ➢Ramanan等人(2015)证明了在受感染细胞中对病毒附加体DNA(cccDNA)的切割。HBV基因组作为3.2kb双链附加体DNA种类存在于受感染的肝细胞的细胞核中,该种类称为共价闭合环状DNA(cccDNA),其是HBV生命周期中的关键组分,其复制不受目前疗法的抑制。这些作者显示特异性靶向HBV的高度保守区的sgRNA稳固地阻遏病毒复制并耗尽cccDNA。

[0508] ➢Nishimasu等人(2015)报道了与单一指导RNA(sgRNA)及其双链DNA靶标复合的SaCas9的晶体结构,其含有5'-TTGAAT-3'PAM和5'-TTGGGT-3'PAM。SaCas9与SpCas9的结构比较突出显示出结构保存和差别,解释了它们不同的PAM特异性和直系同源性sgRNA识别。

[0509] ➢Canver等人(2015)证明了基于CRISPR-Cas9的非编码基因组元件的功能探索。这些作者开发了合并的CRISPR-Cas9指导RNA文库以进行人类和小鼠BCL11A增强子的原位

饱和诱变,这揭示了增强子的关键特征。

[0510] ➢Zetsche等人(2015)报道了Cas13的表征,Cas13是来自新凶手弗朗西丝菌(Francisella novicida)U112的2类CRISPR核酸酶,其具有与Cas9不同的特征。Cas13是一种缺乏tracrRNA的单一RNA指导的内切核酸酶,利用富含T的原间隔区邻近基序,并且经由交错的DNA双链断裂切割DNA。

[0511] ➢Shmakov等人(2015)报道了三种不同的2类CRISPR-Cas系统。两种系统CRISPR酶(C2c1和C2c3)含有与Cas13远缘相关的RuvC样内切核酸酶结构域。与Cas13不同,C2c1依赖crRNA和tracrRNA两者进行DNA切割。第三种酶(C2c2)含有两个预测的HEPN RNA酶结构域,并且不依赖tracrRNA。

[0512] ➢Slaymaker等人(2016)报道了使用结构指导的蛋白质工程化来改善化脓链球菌Cas9(SpCas9)的特异性。这些作者开发了“增强的特异性”SpCas9(eSpCas9)变异体,该变异体维持了稳固的中靶切割,并降低了脱靶效应。

[0513] 本文提供的方法和工具是针对Cas13示例,Cas13是一种不使用tracrRNA的II型核酸酶。如本文所述,已经在不同的细菌种类中鉴定了Cas13的直系同源物。可以使用本领域中描述的方法鉴定具有类似特性的另外的II型核酸酶(Shmakov等人2015,60:385-397;Abudayeh等人2016,Science,5;353(6299))。在特定实施方案中,此类用于鉴定新颖CRISPR效应蛋白的方法可包括以下步骤:从数据库中选择编码种子的序列,所述种子鉴定CRISPR Cas基因座的存在;鉴定位于种子的10kb内、在选定序列中包含开放阅读框(ORF)的基因座;从中选择包含多个ORF的基因座,其中仅单个ORF编码新颖的CRISPR效应子,所述新颖的CRISPR效应子具有大于700个氨基酸并且与已知的CRISPR效应子具有不超过90%的同源性。在特定实施方案中,种子是与CRISPR-Cas系统共用的蛋白质,诸如Cas1。在另外的实施方案中,使用CRISPR阵列作为种子以鉴定新的效应蛋白。

[0514] 已经证实了本发明的有效性。可以例如通过电穿孔转染包含Cas13和crRNA的预组装重组CRISPR-Cas13复合物,从而产生高突变率且不存在可检测的脱靶突变。Hur,J.K.等人,Targeted mutagenesis in mice by electroporation of Cas13ribonucleoproteins,Nat Biotechnol.2016年6月6日.doi:10.1038/nbt.3596。全基因组分析显示,Cas13具有高度特异性。根据一种量度,在人类HEK293T细胞中确定的Cas13的体外切割位点明显少于SpCas9的体外切割位点。Kim,D.等人,Genome-wide analysis reveals specificities of Cas13 endonucleases in human cells,Nat Biotechnol.2016年6月6日.doi:10.1038/nbt.3609。在果蝇中已经证明了采用Cas13的有效多重系统,该系统使用了从包含本发明tRNA的阵列中加工得到的gRNA。Port,F.等人,Expansion of the CRISPR toolbox in an animal with tRNA-flanked Cas9 and Cas13gRNAs.doi:http://dx.doi.org/10.1101/046417。

[0515] 另外,“Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing”,Shengdar Q.Tsai、Nicolas Wyvekens、Cyd Khayter、Jennifer A.Foden、Vishal Thapar、Deepak Reyon、Mathew J.Goodwin、Martin J.Aryee,J.Keith Joung Nature Biotechnology 32(6):569-77(2014)涉及二聚体RNA指导的FokI核酸酶,该酶识别扩展序列并且可以在人类细胞中高效编辑内源性基因。

[0516] 关于CRISPR/Cas系统、其组分、及此类组分的递送(包括方法、材料、递送媒介物、

载体、粒子、以及其制造和使用(包括关于量和配制品)),连同表达CRISPR-Cas的真核细胞、表达CRISPR-Cas的真核生物(诸如小鼠)的一般信息,参考以下:美国专利号8,999,641、8,993,233、8,697,359、8,771,945、8,795,965、8,865,406、8,871,445、8,889,356、8,889,418、8,895,308、8,906,616、8,932,814和8,945,839;美国专利公布US 2014-0310830(美国申请序列号14/105,031)、US 2014-0287938 A1(美国申请序列号14/213,991)、US 2014-0273234 A1(美国申请序列号14/293,674)、US2014-0273232 A1(美国申请序列号14/290,575)、US 2014-0273231(美国申请序列号14/259,420)、US 2014-0256046 A1(美国申请序列号14/226,274)、US 2014-0248702 A1(美国申请序列号14/258,458)、US 2014-0242700 A1(美国申请序列号14/222,930)、US 2014-0242699 A1(美国申请序列号14/183,512)、US 2014-0242664 A1(美国申请序列号14/104,990)、US 2014-0234972 A1(美国申请序列号14/183,471)、US 2014-0227787 A1(美国申请序列号14/256,912)、US 2014-0189896 A1(美国申请序列号14/105,035)、US 2014-0186958(美国申请序列号14/105,017)、US 2014-0186919 A1(美国申请序列号14/104,977)、US 2014-0186843 A1(美国申请序列号14/104,900)、US 2014-0179770 A1(美国申请序列号14/104,837)和US 2014-0179006 A1(美国申请序列号14/183,486)、US 2014-0170753(美国申请序列号14/183,429);US 2015-0184139(美国申请序列号14/324,960);14/054,414欧洲专利申请EP 2 771 468(EP13818570.7)、EP 2 764 103(EP13824232.6)和EP 2 784 162(EP14170383.5);以及PCT专利公布WO2014/093661(PCT/US2013/074743)、WO2014/093694(PCT/US2013/074790)、WO2014/093595(PCT/US2013/074611)、WO2014/093718(PCT/US2013/074825)、WO2014/093709(PCT/US2013/074812)、WO2014/093622(PCT/US2013/074667)、WO2014/093635(PCT/US2013/074691)、WO2014/093655(PCT/US2013/074736)、WO2014/093712(PCT/US2013/074819)、WO2014/093701(PCT/US2013/074800)、WO2014/018423(PCT/US2013/051418)、WO2014/204723(PCT/US2014/041790)、WO2014/204724(PCT/US2014/041800)、WO2014/204725(PCT/US2014/041803)、WO2014/204726(PCT/US2014/041804)、WO2014/204727(PCT/US2014/041806)、WO2014/204728(PCT/US2014/041808)、WO2014/204729(PCT/US2014/041809)、WO2015/089351(PCT/US2014/069897)、WO2015/089354(PCT/US2014/069902)、WO2015/089364(PCT/US2014/069925)、WO2015/089427(PCT/US2014/070068)、WO2015/089462(PCT/US2014/070127)、WO2015/089419(PCT/US2014/070057)、WO2015/089465(PCT/US2014/070135)、WO2015/089486(PCT/US2014/070175)、WO2015/058052(PCT/US2014/061077)、WO2015/070083(PCT/US2014/064663)、WO2015/089354(PCT/US2014/069902)、WO2015/089351(PCT/US2014/069897)、WO2015/089364(PCT/US2014/069925)、WO2015/089427(PCT/US2014/070068)、WO2015/089473(PCT/US2014/070152)、WO2015/089486(PCT/US2014/070175)、WO2016/049258(PCT/US2015/051830)、WO2016/094867(PCT/US2015/065385)、WO2016/094872(PCT/US2015/065393)、WO2016/094874(PCT/US2015/065396)、WO2016/106244(PCT/US2015/067177)。

[0517] 还提到了2015年6月17日提交的美国申请62/180,709,受保护的指导RNA(PROTECTED GUIDE RNAS(PGRNAS));2014年12月12日提交的美国申请62/091,455,受保护的指导RNA(PROTECTED GUIDE RNAS(PGRNAS));2014年12月24日提交的美国申请62/096,708,受保护的指导RNA(PROTECTED GUIDE RNAS(PGRNAS));2014年12月12日提交的美国申

请62/091,462、2014年12月23日提交的美国申请62/096,324、2015年6月17日提交的美国申请62/180,681和2015年10月5日提交的美国申请62/237,496,CRISPR转录因子的死亡指导物(DEAD GUIDES FOR CRISPR TRANSCRIPTION FACTORS);2014年12月12日提交的美国申请62/091,456和2015年6月17日提交的美国申请62/180,692,用于CRISPR-CAS系统的受护航的和官能化的指导物(ESCORTED AND FUNCTIONALIZED GUIDES FOR CRISPR-CAS SYSTEMS);2014年12月12日提交的美国申请62/091,461,用于关于造血干细胞(HSC)的基因组编辑的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR GENOME EDITING AS TO HEMATOPOETIC STEM CELLS(HSCs));2014年12月19日提交的美国申请62/094,903,通过基因组范围的插入捕获测序对双链断裂和基因组重排的无偏鉴定(UNBIASED IDENTIFICATION OF DOUBLE-STRAND BREAKS AND GENOMIC REARRANGEMENT BY GENOME-WISE INSERT CAPTURE SEQUENCING);2014年12月24日提交的美国申请62/096,761,用于序列操纵的系统、方法及优化的酶以及指导物支架的工程化(ENGINEERING OF SYSTEMS,METHODS AND OPTIMIZED ENZYME AND GUIDE SCAFFOLDS FOR SEQUENCE MANIPULATION);2014年12月30日提交的美国申请62/098,059、2015年6月18日提交的美国申请62/181,641和2015年6月18日提交的美国申请62/181,667,RNA靶向系统(RNA-TARGETING SYSTEM);2014年12月24日提交的美国申请62/096,656和2015年6月17日提交的62/181,151,具有稳定化结构域或与稳定化结构域相关联的CRISPR(CRISPR HAVING OR ASSOCIATED WITH DESTABILIZATION DOMAINS);2014年12月24日提交的美国申请62/096,697,具有AAV或与AAV相关联的CRISPR(CRISPR HAVING OR ASSOCIATED WITH AAV);2014年12月30日提交的美国申请62/098,158,工程化CRISPR复合物插入靶向系统(ENGINEERED CRISPR COMPLEX INSERTIONAL TARGETING SYSTEMS);2015年4月22日提交的美国申请62/151,052,用于胞外核外报告的细胞靶向(CELLULAR TARGETING FOR EXTRACELLULAR EXOSOMAL REPORTING);2014年9月24日提交的美国申请62/054,490,使用粒子递送组分靶向病症和疾病的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING PARTICLE DELIVERY COMPONENTS);2014年2月12日提交的美国申请61/939,154,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物(SYSTEMS,METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年9月25日提交的美国申请62/055,484,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物(SYSTEMS,METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年12月4日提交的美国申请62/087,537,用于用优化的功能性CRISPR-CAS系统进行序列操纵的系统、方法和组合物(SYSTEMS,METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年9月24日提交的美国申请62/054,651,用于对体内多种癌症突变的竞争建模的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO);2014年10月23日提交

的美国申请62/067,886,用于对体内多种癌症突变的竞争建模的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR MODELING COMPETITION OF MULTIPLE CANCER MUTATIONS IN VIVO);2014年9月24日提交的美国申请62/054,675和2015年6月17日提交的美国申请62/181,002,在神经元细胞/组织中CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN NEURONAL CELLS/TISSUES);2014年9月24日提交的美国申请62/054,528,在免疫疾病或病症中CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS IN IMMUNE DISEASES OR DISORDERS);2014年9月25日提交的美国申请62/055,454,用于使用细胞穿透肽(CPP)靶向病症和疾病的CRISPR-CAS系统和组合物的递送、用途以及治疗应用(DELIVERY,USE AND THERAPEUTIC APPLICATIONS OF THE CRISPR-CAS SYSTEMS AND COMPOSITIONS FOR TARGETING DISORDERS AND DISEASES USING CELL PENETRATION PEPTIDES(CPP));2014年9月25日提交的美国申请62/055,460,多功能性CRISPR复合物和/或优化的酶连接的功能性CRISPR复合物(MULTIFUNCTIONAL-CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES);2014年12月4日提交的美国申请62/087,475和2015年6月18日提交的62/181,690,用优化的功能性CRISPR-CAS系统进行功能性筛选(FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年9月25日提交的美国申请62/055,487,用优化的功能性CRISPR-CAS系统进行功能性筛选(FUNCTIONAL SCREENING WITH OPTIMIZED FUNCTIONAL CRISPR-CAS SYSTEMS);2014年12月4日提交的美国申请62/087,546和2015年6月18日提交的62/181,687,多功能性CRISPR复合物和/或优化的酶连接的功能性CRISPR复合物(MULTIFUNCTIONAL CRISPR COMPLEXES AND/OR OPTIMIZED ENZYME LINKED FUNCTIONAL-CRISPR COMPLEXES);2014年12月30日提交的美国申请62/098,285,对肿瘤生长和转移的CRISPR介导的体内建模以及遗传筛选(CRISPR MEDIATED IN VIVO MODELING AND GENETIC SCREENING OF TUMOR GROWTH AND METASTASIS)。

[0518] 提到了2015年6月18日提交的美国申请62/181,659和2015年8月19日提交的美国申请62/207,318,用于序列操纵的CAS9直系同源物和变体的系统、方法、酶和指导物支架的工程化和优化(ENGINEERING AND OPTIMIZATION OF SYSTEMS,METHODS,ENZYME AND GUIDE SCAFFOLDS OF CAS9 ORTHOLOGS AND VARIANTS FOR SEQUENCE MANIPULATION)。提到了2015年6月18日提交的美国申请62/181,663和2015年10月22日提交的美国申请62/245,264,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS);2015年6月18日提交的美国申请62/181,675、2015年10月22日提交的美国申请62/285,349、2016年2月17日提交的62/296,522和2016年4月8日提交的美国申请62/320,231,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS);2015年9月24日提交的美国申请62/232,067、2015年12月18日提交的美国申请14/975,085、欧洲申请号16150428.7、2015年8月16日提交的美国申请62/205,733、2015年8月5日提交的美国申请62/201,542、2015年7月16日提交的美国申请62/193,507和2015年6月18日提交的美国申请62/181,739,其各自的标题为新颖的CRISPR酶和系统(NOVELCRISPR ENZYMES AND SYSTEMS);以及2015年10月22日提交的美国申请62/

245,270,新颖的CRISPR酶和系统(NOVEL CRISPR ENZYMES AND SYSTEMS)。还提到了2014年2月12日提交的美国申请61/939,256和2014年12月12日提交的WO 2015/089473(PCT/US2014/070152),其各自的标题为用于序列操纵的具有新架构的系统、方法和优化的指导物组合物的工程化(ENGINEERING OF SYSTEMS,METHODS AND OPTIMIZED GUIDE COMPOSITIONS WITH NEW ARCHITECTURES FOR SEQUENCE MANIPULATION)。还提到了2015年8月15日提交的PCT/US2015/045504、2015年6月17日提交的美国申请62/180,699和2014年8月17日提交的美国申请62/038,358,其各自的标题为使用CAS9切口酶进行基因组编辑(GENOME EDITING USING CAS9 NICKASES)。

[0519] 这些专利、专利公布和申请的每一者,以及在其中或在它们的审查程序期间引用的所有文献(“申请引用文献”)以及在这些申请引用文献中引用或参考的所有文献,连同其中提到的或在其中任何文献中提到并以引用方式并入本文的针对任何产品的任何说明书、说明、产品规格和产品表,特此以引用方式并入本文,并且可以在本发明的实践中采用。所有文献(例如这些专利、专利公布和申请以及申请引用文献)在如同每个单独文献被确切地且单独地指明为以引用方式并入的相同程度上以引用方式并入本文。

[0520] RNA靶向效应蛋白在调节细胞状态中的用途

[0521] 在某些实施方案中,与crRNA复合的Cas13在结合至靶RNA后被激活,并随后切割任何附近的ssRNA靶标(即“附带”或“旁观者”效应)。一旦被同源靶标引发,Cas13即可切割其他(非互补)RNA分子。这种混杂的RNA切割可能潜在地引起细胞毒性,或者以其他方式影响细胞生理或细胞状态。

[0522] 因此,在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞休眠或在诱导细胞休眠中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞周期停滞或在诱导细胞周期停滞中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于减少细胞生长和/或细胞增殖或在减少细胞生长和/或细胞增殖中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞无能或在诱导细胞无能中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞凋亡或在诱导细胞凋亡中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞坏死或在诱导细胞坏死中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导细胞死亡或在诱导细胞死亡中使用。在某些实施方案中,如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统用于诱导程序性细胞死亡或在诱导程序性细胞死亡中使用。

[0523] 在某些实施方案中,本发明涉及一种用于诱导细胞休眠的方法,所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中,本发明涉及一种用于诱导细胞周期停滞的方法,所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中,本发明涉及一种用于减少细胞生长和/或细胞增殖的方法,所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中,本发明涉及一种用于诱导细胞无能的方法,所述方法包括引入或诱导如本文所述的非天然存在的或

工程化的组合物、载体系统或递送系统。在某些实施方案中，本发明涉及一种用于诱导细胞凋亡的方法，所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中，本发明涉及一种用于诱导细胞坏死的方法，所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中，本发明涉及一种用于诱导细胞死亡的方法，所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。在某些实施方案中，本发明涉及一种用于诱导程序性细胞死亡的方法，所述方法包括引入或诱导如本文所述的非天然存在的或工程化的组合物、载体系统或递送系统。

[0524] PAM的测定

[0525] 如下可以确保PAM的测定。此实验与大肠杆菌中的StCas9异源表达的类似工作(Sapranauskas,R.等人Nucleic Acids Res 39,9275-9282(2011))极为相似。申请人将含有PAM和抗性基因两者的质粒引入到异源大肠杆菌中，接着铺板在相应抗生素上。如果存在质粒的DNA切割，则申请人观察不到有活力的菌落。

[0526] 在进一步细节中，如下针对DNA靶标进行测定。在此测定中使用两种大肠杆菌菌株。一种携带编码来自细菌菌株的内源性效应蛋白基因座的质粒。另一种菌株携带空质粒(例如pACYC184，对照菌株)。将所有可能的7或8bp PAM序列呈递在抗生素抗性质粒(具有氨苄青霉素抗性基因的pUC19)上。将PAM定位成靠近原间隔区1的序列(内源性效应蛋白基因座中的第一间隔区的DNA靶标)。克隆了两个PAM文库。一个具有原间隔区的8个随机bp 5’(例如总的65536个不同PAM序列＝复杂度)。另一个文库具有原间隔区的7个随机bp 3’(例如总复杂度是16384个不同的PAM)。将两个文库克隆成具有平均500个质粒/可能的PAM。用5’PAM和3’PAM文库在单独的转化中转化测试菌株和对照菌株并且将转化的细胞分别铺在氨苄青霉素板上。使用质粒的识别和随后的切割/干扰使得细胞对氨苄青霉素易感并且阻止了生长。转化后大约12小时，收获由测试菌株和对照菌株形成的所有菌落并且分离出质粒DNA。使用质粒DNA作为用于PCR扩增和随后的深度测序的模板。未转化的文库中的所有PAM的表现度显示转化细胞中的PAM的预期表现度。对照菌株中发现的所有PAM的表现度显示真实的表现度。测试菌株中的所有PAM的表现度显示哪个PAM未被酶识别并且与对照菌株的比较允许提取出耗尽PAM的序列。

[0527] 对于某些野生型Cas13直系同源物，已鉴定出以下PAM：氨基酸球菌属种BV3L6Cas13(AsCas13)、毛螺科菌ND2006 Cas13(LbCas13)和易北普雷沃氏菌(PaCas13)可以切割以TTTV PAM开头的靶位点，其中V是A/C或G，FnCas13p可以切割以TTN开头的位点，其中N是A/C/G或T。牛眼莫拉氏菌AAX08_00205、牛眼莫拉氏菌AAX11_00205、丁酸弧菌属种NC3005、硫微螺菌属种XS5或毛螺科菌MA2020 PAM是5’TTN，其中N是A/C/G或T。天然PAM序列是TTTV或BTTV，其中B是T/C或G，V是A/C或G，并且效应蛋白是腔隙莫拉氏菌(Moraxella lacunata)Cas13。

[0528] 密码子优化的核酸序列

[0529] 在效应蛋白有待以核酸形式施用的情况下，本申请设想了使用密码子优化的CRISPR-Cas V型蛋白，并且更特别地是Cas13编码核酸序列(和任选地蛋白序列)。在这种情况下，密码子优化的序列的实例是对于在真核生物中表达而优化，例如人类(即，对于在人类中表达而优化)，或对于在如本文所论述的另一种真核生物、动物或哺乳动物中表达而优

化的序列;参见WO 2014/093622(PCT/US2013/074667)中的作为密码子优化的序列的实例的SaCas9人类密码子优化的序列(根据本领域的知识和本公开,密码子优化的编码核酸分子,尤其是关于效应蛋白(例如Cas13)在技术人员的能力范围之内)。虽然这是优选的,但将了解,其他实例可能存在,并且对于除人类以外的宿主种类的密码子优化或对于特定器官的密码子优化是已知的。在一些实施方案中,编码DNA/RNA-靶向Cas蛋白的酶编码序列对于在特定细胞,诸如真核细胞中表达进行密码子优化。真核细胞可以是特定生物体的那些或来源于特定生物体的那些,所述生物体诸如植物或哺乳动物,包括但不限于人类,或如本文所论述的非人类真核生物或动物或哺乳动物,例如小鼠、大鼠、兔、犬、家畜或非人类哺乳动物或灵长类动物。在一些实施方案中,可以排除有可能不会对人或动物带来任何实质性医学益处的修改人类的种系遗传身份的过程和/或修改动物的遗传身份的过程,以及由此类过程产生的动物。一般来讲,密码子优化是指修饰核酸序列用于增强在目标宿主细胞中的表达的过程,这个过程是通过用在宿主细胞的基因中较频繁或最频繁使用的密码子替代原生序列的至少一个密码子(例如约或大于约1个、2个、3个、4个、5个、10个、15个、20个、25个、50个或更多个密码子),同时维持原生氨基酸序列。不同种类对特定氨基酸的某些密码子展现特定偏性。密码子偏性(生物体之间密码子使用的差异)常常与信使RNA(mRNA)的翻译效率相关,据信所述效率继而尤其取决于所翻译的密码子的特性和特定转移RNA(tRNA)分子的可用性。细胞中所选tRNA的主导性一般反映了肽合成中最频繁使用的密码子。因此,可以基于密码子优化来调整基因用于给定的生物体中最佳基因表达。密码子使用表可容易得到,例如,在www.kazusa.orjp/codon/上可得的“密码子使用数据库(Codon Usage Database)”中,并且这些表可以按许多方式进行改编。参见Nakamura,Y.,等人“Codon usage tabulated from the international DNA sequence databases:status for the year 2000”Nucl.Acids Res.28:292(2000)。对于在特定宿主细胞中表达对特定序列进行密码子优化的计算机算法也可得到,诸如Gene Forge(Aptagen;Jacobus,PA)也可得到。在一些实施方案中,编码DNA/RNA-靶向Cas蛋白的序列中的一个或多个密码子(例如1个、2个、3个、4个、5个、10个、15个、20个、25个、50个或更多个或所有密码子)对应于对于特定氨基酸最频繁使用的密码子。关于酵母中的密码子使用,参考在http://www.yeastgenome.org/community/codon_usage.shtml上可得的在线酵母基因组数据库(online Yeast Genome database);或关于酵母中的密码子选择,参考Bennetzen和Hall,J Biol Chem.1982年3月25日;257(6):3026-31。关于包括藻类在内的植物中的密码子使用,参考Codon usage in higher plants,green algae,and cyanobacteria,Campbell and Gowri,Plant Physiol.1990年1月;92(1):1-11.;以及关于植物基因中的密码子使用,参考Murray等人,Nucleic Acids Res.1989年1月25日17(2):477-98;或Selection on the codon bias of chloroplast and cyanelle genes in different plant and algal lineages,Morton BR,J Mol Evol.1998年4月;46(4):449-59。

[0530] 在某些示例性实施方案中,CRISPR Cas蛋白选自表1。

[0531] 表1

[0532]

C2c2 直系同源物	代码	多字母
沙氏纤毛菌	C2-2	Lsh
韦德纤毛菌 F0279 (Lw2)	C2-3	Lw2
斯氏李斯特菌(Listeria seeligeri)	C2-4	Lse
毛螺科菌 MA2020	C2-5	LbM
毛螺科菌 NK4A179	C2-6	LbNK179
嗜氨基[梭菌]([Clostridium] aminophilum) DSM 10710	C2-7	Ca
鸡肉杆菌(Carnobacterium gallinarum) DSM 4847	C2-8	Cg
鸡肉杆菌 DSM 4847	C2-9	Cg2
普根沼杆菌(Paludibacter propionicigenes) WB4	C2-10	Pp

[0533]

维森李斯特菌(Listeria weihenstephanensis) FSL R9-0317	C2-11	Lwei
李斯特菌科菌 FSL M6-0635	C2-12	LbFSL
韦德纤毛菌 F0279	C2-13	Lw
荚膜红细菌(Rhodobacter capsulatus) SB 1003	C2-14	Rc
荚膜红细菌 R121	C2-15	Rc
荚膜红细菌 DE442	C2-16	Rc

[0534] 在某些示例性实施方案中,CRISPR效应蛋白是选自表2的Cas13a蛋白。

[0535] 表2

[0536]

c2c2-5 (SEQ ID NO: 34)	1	毛螺科菌 MA2020
c2c2-6 (SEQ ID NO: 35)	2	毛螺科菌 NK4A179
c2c2-7 (SEQ ID NO: 36)	3	嗜氨基[梭菌]DSM 10710
c2c2-8 (SEQ ID NO: 37)	5	鸡肉杆菌 DSM 4847
c2c2-9 (SEQ ID NO: 38)	6	鸡肉杆菌 DSM 4847
c2c2-10 (SEQ ID NO: 39)	7	普根沼杆菌 WB4
c2c2-11 (SEQ ID NO: 40)	9	维森李斯特菌 FSL R9-0317
c2c2-12 (SEQ ID NO: 41)	10	李斯特菌科菌 FSL M6-0635 = 纽约李斯特菌(Listeria newyorkensis) FSL M6-0635
c2c2-13(SEQ ID NO: 42)	12	韦德纤毛菌 F0279
c2c2-14(SEQ ID NO: 43)	15	荚膜红细菌 SB 1003
c2c2-15(SEQ ID NO: 44)	16	荚膜红细菌 R121
c2c2-16 (SEQ ID NO: 45)	17	荚膜红细菌 DE442
c2c2-2 (SEQ ID NO: 46)		
c2c2-3(SEQ ID NO:		韦德纤毛菌(Lw2)

[0537]

47)		
c2c2-4 (SEQ ID NO: 48)		斯氏李斯特菌

[0538]

C2-17(SEQ ID NO: 49)		颊纤毛菌 C-1013-b
C2-18 (SEQ ID NO: 50)		解半纤维素赫氏菌
C2-19 (SEQ ID NO: 51)		直肠[真杆菌]
C2-20 (SEQ ID NO: 52)		真杆菌科菌 CHKCI004
C2-21 (SEQ ID NO: 53)		布劳特氏菌属种马赛-P2398
C2-22 (SEQ ID NO: 54)		纤毛菌属种口腔泰肯菌(Leptotrichia sp. oral taxon) 879 str.F0557
C2-23 (SEQ ID NO: 55)		毛螺科菌 NK4A144
C2-24 (SEQ ID NO: 56)		凝集绿屈挠菌(Chloroflexus aggregans)
C2-25 (SEQ ID NO: 57)		桔红色脱甲基菌(Demequina aurantiaca)
C2-26 (SEQ ID NO: 58)		海旋菌属种(Thalassospira sp.) TSL5-1
C2-27 (SEQ ID NO: 59)		SAMN04487830_13920 [假丁酸弧菌属种 OR37]
C2-28 (SEQ ID NO: 60)		SAMN02910398_00008 [丁酸弧菌属种 YAB3001]
C2-29 (SEQ ID NO: 61)		布劳特氏菌属种马赛-P2398
C2-30 (SEQ ID NO: 62)		纤毛菌属种马赛(Leptotrichia sp. Marseille)-P3007
C2-31 (SEQ ID NO: 63)		伊华拟杆菌(Bacteroides ihuae)
C2-32		SAMN05216357_1045 [紫单孢菌科菌

[0539]

(SEQ ID NO: 64)		(Porphyromonadaceae bacterium) KH3CP3RA]
C2-33 (SEQ ID NO: 65)		河岸李斯特菌(Listeria riparia)
C2-34 (SEQ ID NO: 66)		陌生非适应螺菌(Insolitispirillum peregrinum)

[0540] 在某些示例性实施方案中,CRISPR效应蛋白是选自表3的Cas13b蛋白。

[0541] 表3

动物溃疡伯格菌 (SEQ ID NO: 67)	1
中间普雷沃氏菌 (SEQ ID NO: 68)	2
颊普雷沃氏菌 (SEQ ID NO: 69)	3
牙龈卟啉单胞菌(Porphyromonas gingivalis) (SEQ ID NO: 70)	4
化脓拟杆菌(Bacteroides pyogenes) (SEQ ID NO: 71)	5
另枝菌属种 ZOR0009 (SEQ ID NO: 72)	6
普雷沃氏菌属种 MA2016 (SEQ ID NO: 73)	7a
普雷沃氏菌属种 MA2016 (SEQ ID NO: 74)	7b
鸭疫里默氏杆菌 (SEQ ID NO: 75)	8
桔红色普雷沃氏菌 (SEQ ID NO: 76)	9
解糖普雷沃氏菌 (SEQ ID NO: 77)	10
HMPREF9712_03108 [拟香味类香味菌(Myroides odoratimimus) CCUG 10230] (SEQ ID NO: 78)	11
中间普雷沃氏菌 (SEQ ID NO: 79)	12
犬咬二氧化碳嗜纤维菌 (SEQ ID NO: 80)	13
咽喉卟啉单胞菌 (SEQ ID NO: 81)	14
普雷沃氏菌属种 P5-125 (SEQ ID NO: 82)	15
嗜鳃黄杆菌 (SEQ ID NO: 83)	16
香味类香味菌 (SEQ ID NO: 84)	17

[0542]

[0543]

柱状黄杆菌(Flavobacterium columnare) (SEQ ID NO: 85)	18
牙龈卟啉单胞菌 (SEQ ID NO: 86)	19
卟啉单胞菌属种 COT-052 OH4946 (SEQ ID NO: 87)	20
中间普雷沃氏菌 (SEQ ID NO: 88)	21
PIN17_0200 [中间普雷沃氏菌 17] (SEQ ID NO: 89)	AFJ07523
中间普雷沃氏菌 (SEQ ID NO: 90)	BAU18623
HMPREF6485_0083 [颊普雷沃氏菌 ATCC 33574] (SEQ ID NO: 91)	EFU31981
HMPREF9144_1146 [苍白普雷沃氏菌(Prevotella pallens) ATCC 700821] (SEQ ID NO: 92)	EGQ18444
HMPREF9714_02132 [拟香味类香味菌 CCUG 12901] (SEQ ID NO: 93)	EHO08761
HMPREF9711_00870 [拟香味类香味菌 CCUG 3837] (SEQ ID NO: 94)	EKB06014
HMPREF9699_02005 [动物溃疡伯格菌 ATCC 43767] (SEQ ID NO: 95)	EKB54193
HMPREF9151_01387 [解糖普雷沃氏菌 F0055] (SEQ ID NO: 96)	EKY00089
A343_1752 [牙龈卟啉单胞菌 JCVI SC001] (SEQ ID NO: 97)	EOA10535
HMPREF1981_03090 [化脓拟杆菌 F0041] (SEQ ID NO: 98)	ERI81700
HMPREF1553_02065 [牙龈卟啉单胞菌 F0568] (SEQ ID NO: 99)	ERJ65637
HMPREF1988_01768 [牙龈卟啉单胞菌 F0185] (SEQ ID NO: 100)	ERJ81987
HMPREF1990_01800 [牙龈卟啉单胞菌 W4087] (SEQ ID NO: 101)	ERJ87335
M573_117042 [中间普雷沃氏菌 ZT] (SEQ ID NO: 102)	KJJ86756
A2033_10205 [拟杆菌门细菌(Bacteroidetes bacterium) GWA2_31_9]] (SEQ ID NO: 103)	OFX18020.1
SAMN05421542_0666 [济州岛金黄杆菌(Chryseobacterium jejuense)]] (SEQ ID NO: 104)	SDI27289.1
SAMN05444360_11366 [卡耐克金黄杆菌(Chryseobacterium carnipullorum)]] (SEQ ID NO: 105)	SHM52812.1
SAMN05421786_1011119 [解脲酸金黄杆菌	SIS70481.1

[0544]

(Chryseobacterium ureilyticum)]] (SEQ ID NO: 106)	
颊普雷沃氏菌] (SEQ ID NO: 107)	WP_004343581
牙龈卟啉单胞菌] (SEQ ID NO: 108)	WP_005873511
牙龈卟啉单胞菌] (SEQ ID NO: 109)	WP_005874195
苍白普雷沃氏菌] (SEQ ID NO: 110)	WP_006044833
香味类香味菌] (SEQ ID NO: 111)	WP_006261414
香味类香味菌] (SEQ ID NO: 112)	WP_006265509
普雷沃氏菌属种 MSX73] (SEQ ID NO: 113)	WP_007412163
牙龈卟啉单胞菌] (SEQ ID NO: 114)	WP_012458414
普根沼杆菌] (SEQ ID NO: 115)	WP_013446107
牙龈卟啉单胞菌] (SEQ ID NO: 116)	WP_013816155
柱状黄杆菌] (SEQ ID NO: 117)	WP_014165541
扭曲冷弯曲菌(Psychroflexus torquis)] (SEQ ID NO: 118)	WP_015024765
鸭疫里默氏杆菌] (SEQ ID NO: 119)	WP_015345620
胸膜炎普雷沃氏菌(Prevotella pleuritidis)] (SEQ ID NO: 120)	WP_021584635
牙龈卟啉单胞菌 (SEQ ID NO: 121)	WP_021663197
牙龈卟啉单胞菌 (SEQ ID NO: 122)	WP_021665475
牙龈卟啉单胞菌 (SEQ ID NO: 123)	WP_021677657
牙龈卟啉单胞菌 (SEQ ID NO: 124)	WP_021680012
牙龈卟啉单胞菌 (SEQ ID NO: 125)	WP_023846767
法氏普雷沃氏菌(Prevotella falsenii) (SEQ ID NO: 126)	WP_036884929
胸膜炎普雷沃氏菌 (SEQ ID NO: 127)	WP_036931485
[牙龈卟啉单胞菌 (SEQ ID NO: 128)	WP_039417390

[0545]

咽喉卟啉单胞菌 (SEQ ID NO: 129)	WP_039418912
咽喉卟啉单胞菌 (SEQ ID NO: 130)	WP_039419792
咽喉卟啉单胞菌 (SEQ ID NO: 131)	WP_039426176
咽喉卟啉单胞菌 (SEQ ID NO: 132)	WP_039431778
咽喉卟啉单胞菌 (SEQ ID NO: 133)	WP_039437199
咽喉卟啉单胞菌 (SEQ ID NO: 134)	WP_039442171
咽喉卟啉单胞菌 (SEQ ID NO: 135)	WP_039445055
猫咬二氧化碳嗜纤维菌(Capnocytophaga cynodegmi) (SEQ ID NO: 136)	WP_041989581
普雷沃氏菌属种 P5-119 (SEQ ID NO: 137)	WP_042518169
普雷沃氏菌属种 P4-76 (SEQ ID NO: 138)	WP_044072147
普雷沃氏菌属种 P5-60 (SEQ ID NO: 139)	WP_044074780
厦门褐指藻杆菌(Phaeodactylibacter xiamenensis) (SEQ ID NO: 140)	WP_044218239
黄杆菌属种 316 (SEQ ID NO: 141)	WP_045968377
咽喉卟啉单胞菌 (SEQ ID NO: 142)	WP_046201018
WP_047431796 (SEQ ID NO: 143)	金黄杆菌属种 YR477
鸭疫里默氏杆菌 (SEQ ID NO: 144)	WP_049354263
牙龈卟啉单胞菌 (SEQ ID NO: 145)	WP_052912312
牙龈卟啉单胞菌 (SEQ ID NO: 146)	WP_058019250
柱状黄杆菌 (SEQ ID NO: 147)	WP_060381855
牙龈卟啉单胞菌 (SEQ ID NO: 148)	WP_061156470
牙龈卟啉单胞菌 (SEQ ID NO: 149)	WP_061156637
鸭疫里默氏杆菌 (SEQ ID NO: 150)	WP_061710138
柱状黄杆菌	WP_063744070

[0546]

(SEQ ID NO: 151)	
鸭疫里默氏杆菌 (SEQ ID NO: 152)	WP_064970887
海洋中华微杆菌(Sinomicrobium oceani) (SEQ ID NO: 153)	WP_072319476.1
噬琼胶赖兴巴赫氏菌(Reichenbachiella agariperforans) (SEQ ID NO: 154)	WP_073124441.1

[0547] 在某些示例性实施方案中,CRISPR效应蛋白是来自表4的Cas13c蛋白。

[0548] 表4

[0549]

坏死梭杆菌肠形亚种(Fusobacterium necrophorum subsp. funduliforme) ATCC 51357 contig00003 (SEQ ID NO: 155)
坏死梭杆菌 DJ-2 contig0065, 全基因组鸟枪序列 (SEQ ID NO: 156)
坏死梭杆菌 BFTR-1 contig0068 (SEQ ID NO: 157)
坏死梭杆菌肠形亚种 1_1_36S cont1.14 (SEQ ID NO: 158)
极臭梭杆菌(Fusobacterium perfoetens) ATCC 29250 T364DRAFT_scaffold00009.9_C (SEQ ID NO: 159)
溃疡梭杆菌(Fusobacterium ulcerans) ATCC 49185 cont2.38 (SEQ ID NO: 160)
厌氧盐杆菌属种(Anaerosalibacter sp.) ND1 基因组组件马赛厌氧盐杆菌 ND1 (SEQ ID NO: 161)

[0550] 其他DNA结合蛋白

[0551] 在某些示例性实施方案中,DNA结合蛋白可以是转录激活因子样效应子(TALE)。适用于本发明情形的示例性TALE包括但不限于美国专利申请号2012/0270273和WO/2013/082519中公开的那些。

[0552] 递送

[0553] 在一些实施方案中,可以将AD官能化的系统的组分以各种形式递送,这些形式诸如有DNA/RNA或RNA/RNA或蛋白质RNA的组合。例如,可以将Cas13蛋白作为DNA编码多核苷酸或RNA编码多核苷酸或作为蛋白质递送。可以将指导物作为DNA编码多核苷酸或RNA递送。设想了所有可能的组合,包括混合的递送形式。

[0554] 在一些方面,本发明提供了包括以下步骤的方法:向宿主细胞递送一种或多种多核苷酸(诸如本文所述的一种或多种载体)、其一种或多种转录物和/或由其转录的一种或多种蛋白质。

[0555] 载体

[0556] 一般来讲,术语“载体”是指能够转运它所连接的另一个核酸的核酸分子。载体是复制子,诸如质粒、噬菌体或粘粒,可以向该复制子中插入另一个DNA区段以便使得该插入

的区段复制。通常，载体在与适当控制元件缔合时能够复制。载体包括但不限于单链、双链或部分双链的核酸分子；包含一个或多个游离端、不包含游离端(例如环状)的核酸分子；包含DNA、RNA或两者的核酸分子；以及本领域中已知的多核苷酸的其他种类。一种类型的载体是“质粒”，其是指环状双链DNA环，可以诸如通过标准分子克隆技术向该环中插入另外的DNA区段。另一种类型的载体是病毒载体，其中病毒来源的DNA或RNA序列存在于包装到病毒(例如逆转录病毒、复制缺陷型逆转录病毒、腺病毒、复制缺陷型腺病毒和腺相关病毒)中的载体中。病毒载体还包括由转染到宿主细胞中的病毒携带的多核苷酸。某些载体能够在引入它们的宿主细胞中自主复制(例如具有细菌复制起点的细菌载体和附加型哺乳动物载体)。其他载体(例如非附加型哺乳动物载体)在引入到宿主细胞中之后被整合到宿主细胞的基因组中，并且因此随着宿主基因组一起复制。此外，某些载体能够引导它们可操作地连接的基因的表达。此类载体在本文中称为“表达载体”。用于真核细胞并且在真核细胞中产生表达的载体在本文中可以称为“真核表达载体”。在重组DNA技术中有效用的常用表达载体常常呈质粒的形式。

[0557] 重组表达载体可以包含处于适于在宿主细胞中表达核酸的形式的本发明的核酸，这意味着重组表达载体包含一个或多个调控元件，这些调控元件可以基于用于表达的宿主细胞来选择，可操作地连接至有待表达的核酸序列。在重组表达载体内，“可操作地连接”旨在意指目标核苷酸序列以允许核苷酸序列表达(例如，在体外转录/翻译系统中或当该载体被引入到宿主细胞时在宿主细胞中)的方式连接至一个或多个调控元件。有利的载体包括慢病毒和腺伴随病毒并且此类载体类型还可以被选择用于靶向特定细胞类型。

[0558] 关于重组和克隆方法，提及2004年9月2日以US 2004-0171156A1公布的美国专利申请10/815,730，该专利的内容以引用方式整体并入本文。

[0559] 术语“调控元件”旨在包括启动子、增强子、内部核糖体进入位点(IRES)以及其他表达控制元件(例如转录终止信号，诸如多聚腺苷酸化信号和聚U序列)。此类调控元件描述于例如Goeddel,GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY 185,Academic Press,San Diego,Calif.(1990)。调控元件包括引导核苷酸序列在许多类型的宿主细胞中连续表达的那些元件和引导核苷酸序列仅在某些宿主细胞中表达的那些元件(例如组织特异性调控序列)。组织特异性启动子可以引导主要在希望的目标组织诸如肌肉、神经元、骨骼、皮肤、血液、特定器官(例如肝脏、胰脏)、或特定细胞类型(例如淋巴细胞)中的表达。调控元件还可以时间依赖性方式诸如细胞周期依赖性或发育阶段依赖性方式引导表达，这可以是或也可以不是组织特异性或细胞类型特异性的。在一些实施方案中，载体包含一个或多个pol III启动子(例如1个、2个、3个、4个、5个或更多个pol III启动子)、一个或多个pol II启动子(例如1个、2个、3个、4个、5个或更多个pol II启动子)、一个或多个pol I启动子(例如1个、2个、3个、4个、5个或更多个pol I启动子)或它们的组合。pol III启动子的实例包括但不限于U6和H1启动子。pol II启动子的实例包括但不限于逆转录病毒劳斯氏肉瘤病毒(RSV)LTR启动子(任选地具有RSV增强子)、巨细胞病毒(CMV)启动子(任选地具有CMV增强子)[参见例如Boshart等人,Cell,41:521-530(1985)]、SV40启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EF1α启动子。术语“调控元件”还涵盖增强子元件，诸如WPRE;CMV增强子;HTLV-I的LTR中的R-U5’区段(Mol.Cell.Biol.,第8(1)卷，第466-472页，1988)；SV40增强子；以及兔β-球蛋白的外显子2与3之间的内含子序列

(Proc.Natl.Acad.Sci.USA.,第78(3)卷,第1527-31页,1981)。本领域技术人员将了解的是,表达载体的设计可以取决于诸如有待转化的宿主细胞的选择、所希望的表达水平等因素。载体可以引入到宿主细胞中从而产生由本文所述的核酸编码的转录物、蛋白质或肽,包括融合蛋白或肽(例如,成簇的规律间隔的短回文重复序列(CRISPR)转录物、蛋白质、酶、其突变体形式、其融合蛋白等)。关于调控序列,提及美国专利申请10/491,026,该专利的内容以引用方式整体并入本文。关于启动子,提及PCT公布WO 2011/028929和美国申请12/511,940,这些专利的内容以引用方式整体并入本文。

[0560] 有利的载体包括慢病毒和腺伴随病毒并且此类载体类型还可以被选择用于靶向特定细胞类型。

[0561] 在特定实施方案中,使用融合至腺苷脱氨酶的指导RNA和(任选地修饰或突变的)CRISPR-Cas蛋白的双顺反子载体。融合至腺苷脱氨酶的指导RNA和(任选地修饰或突变的)CRISPR-Cas蛋白的双顺反子表达载体是优选的。一般来讲且具体地讲,在此实施方案中,融合至腺苷脱氨酶的(任选地修饰或突变的)CRISPR-Cas蛋白优选地通过CBh启动子驱动。RNA可以优选地通过Pol III启动子诸如U6启动子驱动。理想的是,将两者结合。

[0562] 载体可以被设计为在原核细胞或真核细胞中表达CRISPR转录物(例如核酸转录物、蛋白质或酶)。例如,CRISPR转录物可以在细菌细胞(诸如大肠杆菌)、昆虫细胞(使用杆状病毒表达载体)、酵母细胞、或哺乳动物细胞中表达。合适的宿主细胞在Goeddel,GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY 185,Academic Press,San Diego,Calif.(1990)中进一步论述。或者,重组表达载体可以例如使用T7启动子调控序列和T7聚合酶来进行体外转录和翻译。

[0563] 载体可以在原核生物或原核细胞中引入并增殖。在一些实施方案中,使用原核生物扩增有待引入到真核细胞中的载体拷贝或者作为产生有待引入到真核细胞中的载体的中间载体(例如,扩增作为病毒载体包装系统的一部分的质粒)。在一些实施方案中,使用原核生物扩增载体拷贝并表达一种或多种核酸,如以便提供用于递送至宿主细胞或宿主生物体的一种或多种蛋白质来源。原核生物中的蛋白质表达最常在大肠杆菌中利用含有引导融合蛋白或非融合蛋白表达的组成型启动子或诱导型启动子的载体进行。融合载体将许多氨基酸添加到其中编码的蛋白质,诸如添加到重组蛋白的氨基末端。此类融合载体可以用于一种或多种目的,诸如:(i)增加重组蛋白的表达;(ii)增加重组蛋白的溶解度;以及(iii)通过充当亲和纯化中的配体来帮助纯化重组蛋白。通常,在融合表达载体中,蛋白水解切割位点被引入在融合部分与重组蛋白的接点处,以使得重组蛋白能够与融合部分分离,从而随后纯化该融合蛋白。此类酶及其同源识别序列包括因子Xa、凝血酶以及肠激酶。示例性融合表达载体包括pGEX(Pharmacia Biotech Inc;Smith and Johnson,1988.Gene 67:31-40)、pMAL(New England Biolabs,Beverly,Mass.)以及pRIT5(Pharmacia,Piscataway,N.J.),它们分别将谷胱甘肽S-转移酶(GST)、麦芽糖E结合蛋白或蛋白A融合至靶重组蛋白。合适的诱导型非融合大肠杆菌表达载体的实例包括pTrc(Amrann等人,(1988)Gene 69:301-315)和pET 11d(Studier等人,GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY185,Academic Press,San Diego,Calif.(1990)60-89)。在一些实施方案中,载体是酵母表达载体。用于在酵母酿酒酵母中表达的载体的实例包括pYepSec1(Baldari等人,1987.EMBO J.6:229-234)、pMFa(Kuijan和Herskowitz,1982.Cell 30:933-943)、

pJRY88(Schultz等人,1987.Gene 54:113-123)、pYES2(Invitrogen Corporation,San Diego,Calif.)以及picZ(InVitrogen Corp,San Diego,Calif.)。在一些实施方案中,载体使用杆状病毒表达载体驱动昆虫细胞中的蛋白质表达。可用于在培养的昆虫细胞(例如SF9细胞)表达蛋白质的杆状病毒载体包括pAc系列(Smith等人,1983.Mol.Cell.Biol.3:2156-2165)和pVL系列(Lucklow和Summers,1989.Virology 170:31-39)。

[0564] 在一些实施方案中,载体能够使用哺乳动物表达载体驱动一种或多种序列在哺乳动物细胞中的表达。哺乳动物表达载体的实例包括pCDM8(Seed,1987.Nature 329:840)和pMT2PC(Kaufman等人,1987.EMBO J.6:187-195)。当用于哺乳动物细胞时,表达载体的控制功能典型地是由一个或多个调控元件提供的。例如,常用的启动子是来源于多瘤、腺病毒2、巨细胞病毒、猿猴病毒40,以及在此披露和本领域已知的其他来源。对于原核细胞和真核细胞二者的其他合适表达系统,参见例如Sambrook等人,MOLECULAR CLONING:A LABORATORY MANUAL.第2版,Cold Spring Harbor Laboratory,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,1989中的第16章和第17章。

[0565] 在一些实施方案中,重组哺乳动物表达载体能够引导核酸优先在特定细胞类型中表达(例如,组织特异性调控元件用于表达核酸)。组织特异性调控元件在本领域中是已知的。合适的组织特异性启动子的非限制性实例包括白蛋白启动子(肝脏特异性;Pinkert等人,1987.Genes Dev.1:268-277)、淋巴特异性启动子(Calame和Eaton,1988.Adv.Immunol.43:235-275)(具体地说T细胞受体)(Winoto和Baltimore,1989.EMBO J.8:729-733)和免疫球蛋白(Baneiji等人,1983.Cell 33:729-740;Queen和Baltimore,1983.Cell 33:741-748)的启动子)、神经元特异性启动子(例如神经丝启动子;Byrne和Ruddle,1989.Proc.Natl.Acad.Sci.USA 86:5473-5477)、胰腺特异性启动子(Edlund等人,1985.Science 230:912-916),以及乳腺特异性启动子(例如乳清启动子;美国专利号4,873,316和欧洲申请公布号264,166)。还涵盖发育调节的启动子,例如鼠hox启动子(Kessel和Gruss,1990.Science 249:374-379)和α-胎蛋白启动子(Campes和Tilghman,1989.Genes Dev.3:537-546)。关于这些原核载体和真核载体,提及美国专利6,750,059,该专利的内容以引用方式整体并入本文。本发明的其他实施方案可以涉及病毒载体的使用,关于此使用提及美国专利申请13/092,085,该专利的内容以引用方式整体并入本文。组织特异性调控元件在本领域中是已知的,并且就这一点而言,提及美国专利7,776,321,该专利的内容以引用方式整体并入本文。在一些实施方案中,调控元件可操作地连接至CRISPR系统的一个或多个元件,以便驱动该CRISPR系统的一个或多个元件表达。

[0566] 在一些实施方案中,驱动核酸靶向系统的一个或多个元件表达的一种或多种载体被引入到宿主细胞中,以使得该核酸靶向系统的这些元件的表达能引导核酸靶向复合物在一个或多个靶位点处形成。例如,核酸靶向效应蛋白和核酸靶向指导RNA可以各自可操作地连接至单独载体上的单独调控元件。核酸靶向系统的一种或多种RNA可以被递送至转基因核酸靶向效应蛋白动物或哺乳动物,例如组成型地或诱导型地或条件型地表达核酸靶向效应蛋白的动物或哺乳动物;或以其他方式表达核酸靶向效应蛋白或具有含有核酸靶向效应蛋白的细胞的动物或哺乳动物,诸如通过在先向这些动物或哺乳动物施用编码或体内表达核酸靶向效应蛋白的一种或多种载体的方式。或者,从相同或不同调控元件表达的这些元件的两种或更多种可以组合在单一载体中,其中一种或多种另外的载体提供核酸靶向系统

在第一载体中不包含的任何组分。组合于单一载体中的核酸靶向系统元件可以布置为任何适合的取向，诸如一个元件位于相对于第二元件的5'(“上游”)或相对于该第二元件的3'(“下游”)。一个元件的编码序列可以位于第二元件的编码序列的相同链或相反链上，并且取向为相同或相反方向。在一些实施方案中，单一启动子驱动编码核酸靶向效应蛋白的转录物和嵌入一个或多个内含子序列之内(例如，各自在不同内含子中、两个或更多个在至少一个内含子中，或所有在单一内含子中)的核酸靶向指导RNA的表达。在一些实施方案中，核酸靶向效应蛋白和核酸靶向指导RNA可以可操作地连接至同一启动子并且从该同一启动子表达。用于表达核酸靶向系统的一个或多个元件的递送媒介物、载体、粒子、纳米粒子、配制品以及其组分如在前述文献诸如WO 2014/093622(PCT/US2013/074667)中所使用的。在一些实施方案中，载体包含一个或多个插入位点，诸如限制性内切核酸酶识别序列(也称之为“克隆位点”)。在一些实施方案中，一个或多个插入位点(例如，约或超过约1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个插入位点)位于一种或多种载体的一个或多个序列元件的上游和/或下游。当使用多个不同的指导序列时，可以使用单一表达构建体来使核酸靶向活性靶向细胞内的多个不同的相应靶序列。例如，单一载体可以包含约或超过约1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、15个、20个或更多个指导序列。在一些实施方案中，可以提供约或超过约1个、2个、3个、4个、5个、6个、7个、8个、9个、10个或更多个含有此指导序列的载体，并且任选地将其递送至细胞中。在一些实施方案中，载体包含可操作地连接至编码核酸靶向效应蛋白的酶编码序列的调控元件。可以单独地递送核酸靶向效应蛋白或一种或多种核酸靶向指导RNA;并且有利的是经由粒子复合物递送这些中的至少一者。核酸靶向效应蛋白mRNA可以在核酸靶向指导RNA之前递送，以给出时间以待核酸靶向效应蛋白表达。核酸靶向效应蛋白mRNA可以在施用核酸靶向指导RNA之前1-12小时(优选约2-6小时)施用。或者，核酸靶向效应蛋白mRNA和核酸靶向指导RNA可以一起施用。有利地，指导RNA的第二加强剂量可以在初始施用核酸靶向效应蛋白mRNA+指导RNA之后1-12小时(优选约2-6小时)施用。为了实现最有效的基因组修饰水平，核酸靶向效应蛋白mRNA和/或指导RNA的附加施用可能是有用的。

[0567] 常规的基于病毒和非病毒的基因转移方法可以用于在哺乳动物细胞或靶组织中引入核酸。此类方法可以用于向培养基或宿主生物体中的细胞施用编码核酸靶向系统的组分的核酸。非病毒载体递送系统包括DNA质粒、RNA(例如本文所述的载体的转录物)、裸核酸，以及与递送媒介物诸如脂质体复合的核酸。病毒载体递送系统包括DNA和RNA病毒，这些病毒在递送至细胞后具有附加型基因组或整合型基因组。对于基因疗法程序的综述，参见Anderson,Science 256:808-813(1992);Nabel和Felgner,TIBTECH 11:211-217(1993);Mitani和Caskey,TIBTECH 11:162-166(1993);Dillon,TIBTECH 11:167-175(1993);Miller,Nature 357:455-460(1992);Van Brunt,Biotechnology 6(10):1149-1154(1988);Vigne,Restorative Neurology and Neuroscience8:35-36(1995);Kremer和Perricaudet,British Medical Bulletin 51(1):31-44(1995);Haddada等人,Current Topics in Microbiology and Immunology,Doerfler和Böhm(编辑)(1995);以及Yu等人,Gene Therapy 1:13-26(1994)。

[0568] 非病毒递送核酸的方法包括包括脂质转染、核转染、微注射、基因枪、病毒体、脂质体、免疫脂质体、聚阳离子或脂质:核酸轭缀合物、裸DNA、人工病毒体以及剂增强DNA摄取。

脂质转染描述于例如美国专利号5,049,386、4,946,787;以及4,897,355)并且脂质转染试剂是商业上销售的(例如Transfectam™和Lipofectin™)。适于多核苷酸的有效受体识别脂质转染的阳离子脂质和中性脂质包括Felgner,WO 91/17424;WO91/16024的那些。递送可以是递送至细胞(例如体外或离体施用)或靶组织(例如体内施用)。

[0569] 质粒递送涉及将指导RNA克隆到CRISPR-Cas蛋白表达质粒中,并在细胞培养物中转染DNA。质粒主链是商业上可得的,并且不需要特定的设备。它们的优势为是模块化的,能够携带不同尺寸的CRISPR-Cas编码序列(包括编码更大尺寸蛋白质的序列)以及选择标记物。同时质粒的优点是它们可以确保瞬时但持续的表达。然而,质粒的递送不是直接的,使得体内效率通常很低。持续表达也可能是不利的,因为这可以增加脱靶编辑。另外,CRISPR-Cas蛋白的过度积累可能对细胞有毒。最后,质粒总是具有在宿主基因组中随机整合dsDNA的风险,更特别地是考虑到产生双链断裂(中靶和脱靶)。

[0570] 脂质:核酸复合物(包括靶向脂质体,诸如免疫脂质复合物)的制备是本领域技术人员所熟知的(参见例如Crystal,Science 270:404-410(1995);Blaese等人,Cancer Gene Ther.2:291-297(1995);Behr等人,Bioconjugate Chem.5:382-389(1994);Remy等人,Bioconjugate Chem.5:647-654(1994);Gao等人,Gene Therapy 2:710-722(1995);Ahmad等人,Cancer Res.52:4817-4820(1992);美国专利4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028和4,946,787)。对此将在下文更详细地论述。

[0571] 使用用于递送核酸的基于RNA或DNA病毒的系统利用了用于将病毒靶向身体内的特定细胞并且将病毒有效负载运输到核内的高度进化的方法。病毒载体可以直接施用至患者(体内),或者它们可以用于体外处理细胞,并且修饰的细胞可以任选地施用至患者(离体)。常规的基于病毒的系统可以包括用于基因转移的逆转录病毒、慢病毒、腺病毒、腺相关病毒以及单纯疱疹病毒载体。通过逆转录病毒、慢病毒以及腺相关病毒基因转移方法整合在宿主基因组中是可能的,这常常导致插入的转基因长期表达。另外,已经在许多不同的细胞类型和靶组织中观察到高转导效率。

[0572] 可以通过并入外源包膜蛋白,扩展靶细胞的潜在靶群体而改变逆转录病毒的向性。慢病毒载体是能够转导或感染非分裂细胞并典型地产生较高病毒效价的逆转录病毒载体。因此,逆转录病毒基因转移系统的选择将取决于靶组织。逆转录病毒载体由顺式作用长末端重复序列组成,这些长末端重复序列具有包装多达6-10kb的外源序列的能力。最低量的顺式作用LTR对于载体的复制和包装而言是足够的,然后使用这些载体将治疗基因整合到靶细胞中,以提供永久的转基因表达。广泛使用的逆转录病毒载体包括基于鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猴免疫缺陷病毒(SIV)、人类免疫缺陷病毒(HIV)以及它们的组合的那些(参见例如Buchscher等人,J.Virol.66:2731-2739(1992);Johann等人,J.Virol.66:1635-1640(1992);Sommnerfelt等人,Virol.176:58-59(1990);Wilson等人,J.Virol.63:2374-2378(1989);Miller等人,J.Virol.65:2220-2224(1991);PCT/US94/05700)。

[0573] 在瞬时表达是优选的应用中,可以使用基于腺病毒的系统。基于腺病毒的载体能够在许多细胞类型中具有极高转导效率并且无需细胞分裂。用此类载体,已经获得了较高的效价和表达水平。可以在相对简单的系统中大量地产生此载体。还可以使用腺相关病毒

(“AAV”)载体转导具有靶核酸的细胞,例如,在体外产生核酸和肽,以及用于体内和离体基因疗法程序(参见例如West等人,Virology 160:38-47(1987);美国专利号4,797,368;WO 93/24641;Kotin,Human Gene Therapy 5:793-801(1994);Muzyczka,J.Clin.Invest.94:1351(1994)。重组AAV载体的构建描述于许多出版物中,包括美国专利号5,173,414;Tratschin等人,Mol.Cell.Biol.5:3251-3260(1985);Tratschin等人,Mol.Cell.Biol.4:2072-2081(1984);Hermonat和Muzyczka,PNAS 81:6466-6470(1984);以及Samulski等人,J.Virol.63:03822-3828(1989)。

[0574] 本发明提供了含有以下各项或基本上由以下各项组成的AAV:编码CRISPR系统的外源性核酸分子,例如包含第一盒或由该第一盒组成的多个盒,该第一盒包含启动子、编码CRISPR-相关(Cas)蛋白(推定的核酸酶或解旋酶蛋白)(例如Cas13)的核酸分子和终止子或基本上由其组成,以及两个或更多个、有利地多至载体的包装尺寸限度,例如总计(包括该第一盒在内)五个包含启动子、编码指导RNA(gRNA)的核酸分子和终止子或基本上由其组成的盒(例如,每个盒示意性地表示为启动子-gRNA1-终止子、启动子-gRNA2-终止子...启动子-gRNA(N)-终止子,其中N是可以插入的数值,该数值是载体包装尺寸限度的上限);或者两个或更多个单独的rAAV,其各自含有CRISPR系统的一个或多于一个盒,例如第一rAAV含有包含启动子、编码Cas例如Cas(Cas13)的核酸分子和终止子或基本上由其组成的第一盒,并且第二rAAV含有一个或多个各自包含启动子、编码指导RNA(gRNA)的核酸分子和终止子或基本上由其组成的盒(例如,每个盒示意性地表示为启动子-gRNA1-终止子、启动子-gRNA2-终止子...启动子-gRNA(N)-终止子,其中N是可以插入的数值,该数值是载体包装尺寸限度的上限)。或者,由于Cas13可以处理自己的crRNA/gRNA,因此可以使用单一crRNA/gRNA阵列用于多重基因编辑。因此,rAAV可以含有单个包含启动子、多个crRNA/gRNA和终止子或基本上由其组成的表达盒(例如,示意性地表示为启动子-gRNA1-gRNA2…gRNA(N)终止子,其中N是可以插入的数值,该数值是载体包装尺寸限度的上限),而不是包括多个盒来递送gRNA。参见Zetsche等人Nature Biotechnology 35,31-34(2017),该文献以引用方式整体并入本文。由于rAAV是DNA病毒,因此本文关于AAV或rAAV的论述中的核酸分子有利地是DNA。在一些实施方案中,启动子有利地是人类突触蛋白I启动子(hSyn)。用于将核酸递送至细胞的另外的方法是本领域技术人员已知的。参见例如US20030087817,该专利以引用方式并入本文。

[0575] 在另一个实施方案中,考虑了科卡尔(Cocal)水泡病毒包膜假型化逆转录病毒载体粒子(参见例如转让给Fred Hutchinson Cancer Research Center的美国专利公布号20120164118)。科卡尔病毒属于水泡病毒属,并且是哺乳动物水疱性口炎的致病物。科卡尔病毒最初从特立尼达拉岛(Trinidad)的螨虫中分离(Jonkers等人,Am.J.Vet.Res.25:236-242(1964)),并且在特立尼达拉岛、巴西和阿根廷已经确定了由昆虫、牛和马引起的感染。感染哺乳动物的许多水泡病毒已从天然感染的节肢动物中分离,这表明它们是载体传播(vector-borne)的。水泡病毒抗体在生活于农村地区的人中是常见的,而这些病毒是地方性的并且是实验室采集的;人类的感染通常导致流感样症状。科卡尔病毒包膜糖蛋白与印第安纳VSV-G享有71.5%氨基酸水平的一致性,并且对水泡病毒的包膜基因的系统发育比较显示科卡尔病毒在血清学上不同于水泡病毒内的印第安纳VSV-G株,但与其紧密相关。Jonkers等人,Am.J.Vet.Res.25:236-242(1964)以及Travassos da Rosa等人,

Am.J.Tropical Med.&Hygiene 33:999-1006(1984)。科卡尔水泡病毒包膜蛋白假型化逆转录病毒载体粒子可以包括慢病毒、α逆转录病毒、β逆转录病毒、γ逆转录病毒、δ逆转录病毒以及ε逆转录病毒载体粒子，这些载体粒子可以包含逆转录病毒Gag、Pol，和/或一种或多种辅助蛋白以及科卡尔水泡病毒包膜蛋白。在这些实施方案的某些方面，Gag、Pol和辅助蛋白是慢病毒和/或γ逆转录病毒。

[0576] 在一些实施方案中，用本文所述的一种或多种载体瞬时或非瞬时转染宿主细胞。在一些实施方案中，当细胞天然地出现在受试者体内(任选地有待重新引入其中)时将其转染。在一些实施方案中，转染的细胞是从受试者获得。在一些实施方案中，细胞来源于从受试者获得的细胞，诸如细胞系。用于组织培养的多种多样的细胞系在本领域中是已知的。细胞系的实例包括但不限于C8161、CCRF-CEM、MOLT、mIMCD-3、NHDF、HeLa-S3、Huh1、Huh4、Huh7、HUVEC、HASMC、HEKn、HEKa、MiaPaCell、Panc1、PC-3、TF1、CTLL-2、C1R、Rat6、CV1、RPTE、A10、T24、J82、A375、ARH-77、Calu1、SW480、SW620、SKOV3、SK-UT、CaCo2、P388D1、SEM-K2、WEHI-231、HB56、TIB55、Jurkat、J45.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、NRK、NRK-52E、MRC5、MEF、Hep G2、HeLa B、HeLa T4、COS、COS-1、COS-6、COS-M6A、BS-C-1猴肾上皮细胞、BALB/3T3小鼠胚成纤维细胞、3T3 Swiss、3T3-L1、132-d5人类胎儿成纤维细胞;10.1小鼠成纤维细胞、293-T、3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、A253、A431、A-549、ALC、B16、B35、BCP-1cells、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C3H-10T1/2、C6/36、Cal-27、CHO、CHO-7、CHO-IR、CHO-K1、CHO-K2、CHO-T、CHO Dhfr-/-、COR-L23、COR-L23/CPR、COR-L23/5010、COR-L23/R23、COS-7、COV-434、CML T1、CMT、CT26、D17、DH82、DU145、DuCaP、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、HEK-293、HeLa、Hepa1c1c7、HL-60、HMEC、HT-29、Jurkat、JY细胞、K562细胞、Ku812、KCL22、KG1、KYO1、LNCap、Ma-Mel 1-48、MC-38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、MDCK II、MDCK II、MOR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NALM-1、NW-145、OPCN/OPCT细胞系、Peer、PNT-1A/PNT 2、RenCa、RIN-5F、RMA/RMAS、Saos-2细胞、Sf-9、SkBr3、T2、T-47D、T84、THP1细胞系、U373、U87、U937、VCaP、Vero cells、WM39、WT-49、X63、YAC-1、YAR，以及它们的转基因品种。细胞系可从本领域技术人员已知的多种来源获得(参见例如美国典型培养物保藏中心(ATCC)(Manassus,Va.))。

[0577] 在特定实施方案中，所述系统的一种或多种组分的瞬时表达和/或存在可能是令人感兴趣的，诸如用以减少脱靶效应。在一些实施方案中，用本文所述的一种或多种载体转染的细胞用于建立包含一种或多种载体来源的序列的新细胞系。在一些实施方案中，使用用如本文所述的系统的组分瞬时转染(诸如通过一种或多种载体进行瞬时转染，或用RNA进行转染)并且通过CRISPR复合物的活性修饰的细胞建立细胞系，所述细胞系包含含有修饰但缺少任何其他外源性序列的细胞。在一些实施方案中，用本文所述的一种或多种载体瞬时或非瞬时转染的细胞，或来源于此类细胞的细胞系用于评定一种或多种测试化合物。

[0578] 在一些实施方案中，设想将RNA和/或蛋白质直接引入宿主细胞。例如，可以将CRISPR-Cas蛋白作为编码mRNA与体外转录的指导RNA一起递送。此类方法可以减少确保CRISPR-Cas蛋白起效的时间，并且进一步防止CRISPR系统组分的长期表达。

[0579] 在一些实施方案中，本发明的RNA分子以脂质体或转化脂(lipofectin)脂质转染

蛋白配制品递送，并且可以通过本领域技术人员所熟知的方法来制备。此类方法描述于例如美国专利5,593,972、5,589,466和5,580,859，这些专利以引用方式并入本文。已开发了特别旨在增强并改善siRNA到哺乳动物细胞的递送的递送系统(参见例如Shen等人FEBS Let.2003,539:111-114;Xia等人,Nat.Biotech.2002,20:1006-1010;Reich等人,Mol.Vision.2003,9:210-216;Sorensen等人,J.Mol.Biol.2003,327:761-766;Lewis等人,Nat.Gen.2002,32:107-108以及Simeoni等人,NAR 2003,31,11:2717-2724)，并且这些递送系统可以适用于本发明。siRNA最近已成功用于抑制灵长类动物中的基因表达(参见，例如Tolentino等人,Retina 24(4):660，该文献也可以适用于本发明)。

[0580] 实际上，RNA递送是可用的体内递送方法。有可能使用脂质体或纳米粒子将Cas13、腺苷脱氨酶和指导RNA递送至细胞中。因此，CRISPR-Cas蛋白(诸如Cas13)的递送、腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接蛋白)的递送和/或本发明的RNA的递送可以可以呈RNA形式并且经由微泡、脂质体或一个粒子或多个粒子来进行。例如，可以将Cas13 mRNA、腺苷脱氨酶mRNA和指导RNA包装到脂质体粒子中以进行体内递送。脂质体转染试剂诸如来自Life Technologies的lipofectamine和市场上的其他试剂可以有效地将RNA分子递送至肝脏中。

[0581] RNA递送手段还优选包括经由粒子的RNA递送(Cho,S.、Goldberg,M.、Son,S.、Xu,Q.、Yang,F.、Mei,Y.、Bogatyrev,S.、Langer,R.和Anderson,D.,Lipid-like nanoparticles for small interfering RNA delivery to endothelial cells,AdvancedFunctional Materials,19:3112-3118,2010)或经由外泌体的递送(Schroeder,A.、Levins,C.、Cortez,C.、Langer,R.和Anderson,D.,Lipid-based nanotherapeutics for siRNA delivery,Journal of Internal Medicine,267:9-21,2010,PMID:20059641)。实际上，已显示外泌体在递送siRNA中特别有用，其为与CRISPR系统有一些相似之处的系统。例如，El-Andaloussi S等人(“Exosome-mediated delivery of siRNA in vitro and in vivo.”Nat Protoc.2012年12月;7(12):2112-26.doi:10.1038/nprot.2012.131.电子版2012年11月15日)描述了外泌体如何成为用于跨不同生物屏障的药物递送的有希望的工具并且可以用于体外和体内递送siRNA。他们的方法是通过包含与肽配体融合的外泌体蛋白的表达载体的转染来产生靶向的外泌体。然后将这些外泌体纯化并且由转染的细胞上清液进行表征，然后将RNA加载到外泌体中。根据本发明的递送或施用可以使用外泌体进行，特别地但不限于脑。维生素E(α-生育酚)可以与CRISPR Cas缀合并且与高密度脂蛋白(HDL)一起递送至脑，方式为例如与(Uno)等人(HUMAN GENE THERAPY 22:711-719(2011年6月))完成的用于将短干扰RNA(siRNA)递送至脑的方式类似。经由填充有磷酸盐缓冲盐水(PBS)或游离TocsiBACE或Toc-siBACE/HDL并且与脑灌注试剂盒3(Brain Infusion Kit 3)(Alzet)连接的微型渗透泵(型号1007D;Alzet,Cupertino,CA)灌注小鼠。将脑灌注插管置于在正中线的前囟的后方约0.5mm，以便灌注到背侧第三脑室中。Uno等人发现，通过相同的ICV灌注方法，少至3nmol的Toc-siRNA与HDL可以诱导相当程度的靶标减少。在本发明中对于人类可以考虑类似剂量的缀合至α-生育酚并且与HDL共同施用靶向脑的CRISPR Cas，例如，可以考虑靶向脑的约3nmol至约3μmol的CRISPR Cas。Zou等人((HUMAN GENE THERAPY 22:465-475(2011年4月))描述了靶向PKCγ的短发夹RNA的慢病毒介导的递送方法，其用于在大鼠脊髓中的体内基因沉默。Zou等人通过鞘内导管施用约10μl的具有1x109个转导单位

(TU)/ml的滴度的重组慢病毒。在本发明中对于人类可以考虑类似剂量的靶向脑的在慢病毒载体中表达的CRISPR Cas,例如,可以考虑靶向脑的在具有1x10^{9}个转导单位(TU)/ml的滴度的慢病毒中的约10-50ml的CRISPR Cas。

[0582] 载体的剂型

[0583] 在一些实施方案中,载体(例如质粒或病毒载体)是例如通过肌内注射递送至目标组织,而有时经由静脉内、经皮、鼻内、经口、粘膜、或其他递送方法进行递送。这种递送可以经由单剂量或多剂量来进行。本领域技术人员理解的是,本文有待递送的实际剂量可以在很大程度上取决于多种因素而变化,诸如载体选择、靶细胞、生物体或组织、有待治疗的受试者的一般状况、所寻求的转化/修饰的程度、施用途径、施用方式、所寻求的转化/修饰的类型等。

[0584] 这样的剂量还可以含有例如载剂(水、盐水、乙醇、甘油、乳糖、蔗糖、磷酸钙、明胶、葡聚糖、琼脂、果胶、花生油、芝麻油等)、稀释剂、药学上可接受的载剂(例如磷酸盐缓冲盐水)、药学上可接受的赋形剂,和/或本领域已知的其他化合物。该剂型还可以含有一种或多种药学上可接受的盐,诸如例如,矿物酸盐诸如盐酸盐、氢溴酸盐、磷酸盐、硫酸盐等;以及有机酸盐,诸如乙酸盐、丙酸盐、丙二酸盐、苯甲酸盐等。另外,本文也可以存在辅助物质,诸如润湿剂或乳化剂、pH缓冲物质、凝胶或胶凝材料、调味剂、着色剂、微球体、聚合物、悬浮剂等。此外,还可以存在一种或多种其他常规药用成分,诸如防腐剂、保湿剂、悬浮剂、表面活性剂、抗氧化剂、抗结剂、填充剂、螯合剂、包衣剂、化学稳定剂等,尤其是在该剂型呈可重构形式时。合适的示例性成分包括微晶纤维素、羧甲基纤维素钠、聚山梨酯80、苯乙醇、三氯叔丁醇、山梨酸钾、抗坏血酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸酯、乙基香兰素、甘油、苯酚、对氯酚、明胶、白蛋白以及它们的组合。对药学上可接受的赋形剂的彻底论述可获自REMINGTON'S PHARMACEUTICAL SCIENCES(Mack Pub.Co.,N.J.1991),该文献以引用方式并入本文。

[0585] 在本文的一个实施方案中,递送是经由腺病毒进行的,其可以是含有至少1x105个腺病毒载体粒子(也称为粒子单位,pu)的单次加强剂量。在本文的一个实施方案中,该剂量优选为腺病毒载体的至少约1x106个粒子(例如约1x106-1x1012个粒子)、更优选至少约1x107个粒子、更优选至少约1x108个粒子(例如约1x108-1x1011个粒子或约1x108-1x1012个粒子)、且最优选至少约1x100个粒子(例如约1x109-1x1010个粒子或约1x109-1x1012个粒子),或甚至至少约1x1010个粒子(例如约1x1010-1x1012个粒子)。或者,该剂量包含不超过约1x1014个粒子、优选不超过约1x1013个粒子、甚至更优选不超过约1x1012个粒子、甚至更优选不超过约1x1011个粒子、且最优选不超过约1x1010个粒子(例如不超过约1x109个粒子)。因此,该剂量可以含有单剂量的腺病毒载体,其具有例如约1x106粒子单位(pu)、约2x106pu、约4x106pu、约1x107pu、约2x107pu、约4x107pu、约1x108pu、约2x108pu、约4x108pu、约1x109pu、约2x109pu、约4x109pu、约1x1010pu、约2x1010pu、约4x1010pu、约1x1011pu、约2x1011pu、约4x1011pu、约1x1012pu、约2x1012pu或约4x1012pu的腺病毒载体。参见,例如,在2013年6月4日授权的授予Nabel等人的美国专利号8,454,972B2中的腺病毒载体;该专利通过引用结合在此,以及在其第29栏第36-58行的剂量。在本文的一个实施方案中,腺病毒是经由多剂量递送的。

[0586] 在本文的一个实施方案中,递送是经由AAV进行的。用于针对人类的AAV的体内递

送的治疗有效剂量被认为处于含有从约1x1010至约1x1010个功能AAV/ml溶液的从约20至约50ml的盐水溶液的范围内。可以调整该剂量以便使治疗益处相对于任何副作用平衡。在本文的一个实施方案中，AAV剂量大致处于约1x105至1x1050个基因组AAV、约1x108至1x1020个基因组AAV、约1x1010至约1x1016个基因组，或约1x1011至约1x1016个基因组AAV的浓度范围内。人类剂量可以是约1x1013个基因组AAV。这样的浓度能以约0.001ml至约100ml、约0.05至约50ml，或约10至约25ml的载剂溶液进行递送。通过建立剂量反应曲线的常规试验，本领域普通技术人员可以容易地确立其他有效剂量。参见，例如，2013年3月26日授权的授予Hajjar等人的美国专利号8,404,658B2，在第27栏，第45-60行。

[0587] 在本文的一个实施方案中，递送是经由质粒进行的。在此类质粒组合物中，该剂量应当是足以引发反应的质粒的量。例如，在质粒组合物中的质粒DNA的适当量可以是约0.1至约2mg，或约1μg至约10μg/70kg个体。本发明的质粒将大体上包含(i)启动子；(ii)可操作地连接至所述启动子的编码CRISPR-Cas蛋白的序列；(iii)选择性标记物；(iv)复制起点；以及(v)在(ii)的下游并可操作地连接至(ii)的转录终止子。质粒还可以编码CRISPR复合物的RNA组分，但是这些组分中的一者或多者还可以被编码在不同载体上。

[0588] 本文的剂量是基于平均70kg的个体。施用频率在医学或兽医学从业者(例如医师、兽医师)或本领域熟练的科学家的范围之内。还应注意的是，实验中使用的小鼠典型地是约20g，根据小鼠实验可以扩展到70kg的个体。

[0589] 用于本文所提供的组合物的剂量包括用于重复施用或重复给药的剂量。在特定实施方案中，在数周、数月或数年的时段内重复施用。可以进行合适的测定以获得最佳剂量方案。重复施用可以允许使用较低的剂量，这可以积极地影响脱靶修饰。

[0590] RNA递送

[0591] 在特定实施方案中，使用基于RNA的递送。在这些实施方案中，CRISPR-Cas蛋白的mRNA、腺苷脱氨酶的mRNA(其可以融合至CRISPR-Cas蛋白或衔接子)与体外转录的指导RNA一起递送。Liang等人描述了使用基于RNA的递送的有效基因组编辑(Protein Cell.2015年5月;6(5):363-372)。在一些实施方案中，可以通过化学方式修饰编码Cas13和/或腺苷脱氨酶的mRNA，与质粒编码的Cas13和/或腺苷脱氨酶相比，这可以使活性提高。例如，一个或多个mRNA中的尿苷可以被假尿苷(Ψ)、N1-甲基假尿苷(me1Ψ)、5-甲氧基尿苷(5moU)部分或完全取代。参见Li等人，NatureBiomedical Engineering 1,0066DOI:10.1038/s41551-017-0066(2017)，该文献以引用方式整体并入本文。

[0592] RNP递送

[0593] 在特定实施方案中，预复合的指导RNA、CRISPR-Cas蛋白和腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接子)作为核糖核蛋白(RNP)递送。RNP的优势在于，与RNA方法相比，它们带来的快速编辑效应更大，因为该过程避免了对转录的需要。一个重要的优点是RNP递送都是瞬时的，从而减少了脱靶效应和毒性问题。以下作者已观察到不同细胞类型中的有效基因组编辑:Kim等人(2014,Genome Res.24(6):1012-9);Paix等人(2015,Genetics 204(1):47-54);Chu等人(2016,BMC Biotechnol.16:4)，以及Wang等人(2013,Cell.9;153(4):910-8)。

[0594] 在特定实施方案中，以如WO2016161516中所述的基于多肽的穿梭剂的方式递送核糖核蛋白。WO2016161516描述了使用包含内体泄漏结构域(ELD)的合成肽对多肽货物的有

效转导，所述内体泄漏结构域(ELD)可操作地连接至细胞穿透结构域(CPD)，连接至富含组氨酸的结构域和CPD。类似地，这些多肽可用于在真核细胞中递送基于CRISPR效应子的RNP。

[0595] 粒子

[0596] 在一些方面或实施方案中，可以使用包含递送粒子配制品的组合物。在一些方面或实施方案中，配制品包含CRISPR复合物，该复合物包含CRISPR蛋白和指导物，该指导物引导CRISPR复合物特异性地结合至靶序列。在一些实施方案中，递送粒子包含基于脂质的粒子，任选地脂质纳米粒子，或阳离子脂质和任选地生物可降解的聚合物。在一些实施方案中，阳离子脂质包括1,2-二油酰基-3-三甲基铵-丙烷(DOTAP)。在一些实施方案中，亲水性聚合物包括乙二醇或聚乙二醇。在一些实施方案中，递送粒子还包含脂蛋白，优选地胆固醇。在一些实施方案中，递送粒子的直径小于500nm、任选地直径小于250nm、任选地直径小于100nm、任选地直径为约35nm至约60nm。

[0597] 示例性粒子递送复合物在2017年4月14日提交的标题为“大有效负载的新型递送(Nove Delivery of Large Payloads)”的美国临时申请62/485,625中进一步公开。

[0598] 已知若干类型的粒子递送系统和/或配制品可用于各种各样的生物医学应用。一般来讲，粒子被定义为，就其传输和特性而言表现为整体的小物体。粒子根据直径被进一步分类。粗粒子的覆盖范围在2,500与10,000纳米之间。细粒子的尺寸在100与2,500纳米之间。超细粒子或纳米粒子的尺寸通常在1与100纳米之间。100nm限制是基于以下事实，即将粒子与块状材料区分开的新颖特性典型地在100nm以下的临界长度尺度上形成。

[0599] 如本文所用，粒子递送系统/配制品被定义为包含根据本发明的粒子的任何生物递送系统/配制品。根据本发明的粒子是具有小于100微米(μm)的最大尺寸(例如直径)的任何实体。在一些实施方案中，本发明的粒子具有小于10μm的最大尺寸。在一些实施方案中，本发明的粒子具有小于2000纳米(nm)的最大尺寸。在一些实施方案中，本发明的粒子具有小于1000纳米(nm)的最大尺寸。在一些实施方案中，本发明的粒子具有小于900nm、800nm、700nm、600nm、500nm、400nm、300nm、200nm或100nm的最大尺寸。典型地，本发明的粒子具有500nm或更小的最大尺寸(例如直径)。在一些实施方案中，本发明的粒子具有250nm或更小的最大尺寸(例如直径)。在一些实施方案中，本发明的粒子具有200nm或更小的最大尺寸(例如直径)。在一些实施方案中，本发明的粒子具有150nm或更小的最大尺寸(例如直径)。在一些实施方案中，本发明的粒子具有100nm或更小的最大尺寸(例如直径)。在本发明的一些实施方案中使用较小的粒子，例如具有50nm或更小的最大尺寸的粒子。在一些实施方案中，本发明的粒子具有在25nm与200nm之间的最大尺寸。

[0600] 就本发明而言，优选使用纳米粒子或脂质包膜递送CRISPR复合物的一种或多种组分，例如CRISPR-Cas蛋白或mRNA、或腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接子)或mRNA、或指导RNA。其他递送系统或载体可以与本发明的纳米粒子方面结合使用。

[0601] 一般来讲，“纳米粒子”是指直径小于1000nm的任何粒子。在某些优选的实施方案中，本发明的纳米粒子具有500nm或更小的最大尺寸(例如直径)。在其他优选的实施方案中，本发明的纳米粒子具有在25nm与200nm之间的最大尺寸。在其他优选的实施方案中，本发明的纳米粒子具有100nm或更小的最大尺寸。在其他优选的实施方案中，本发明的纳米粒子具有在35nm与60nm之间的最大尺寸。应当理解，在适当的情况下，本文中对粒子或纳米粒子的提及可以互换。

[0602] 应当理解，粒子的尺寸将取决于是在加载之前还是加载之后测量而有所不同。因此，在特定实施方案中，术语“纳米粒子”可能仅适用于加载前的粒子。

[0603] 本发明中涵盖的纳米粒子可以提供为不同的形式，例如为固体纳米粒子(例如金属(诸如银、金、铁、钛)、非金属、基于脂质的固体、聚合物)、纳米粒子的悬浮液，或它们的组合。可以制备金属、绝缘体和半导体纳米粒子，以及杂合结构(例如核-壳纳米粒子)。如果由半导体材料制备的纳米粒子足够小(典型地低于10nm)以至于出现电子能级的量子化，则这些纳米粒子还可以是标记量子点。此类纳米级粒子作为药物载剂或成像剂用于生物医学应用中并且可以适于本发明中的类似目的。

[0604] 半固体和软纳米粒子以被制造出并且处于本发明的范围之内。半固体性质的原型纳米粒子是脂质体。目前，临床上将各种类型的脂质体纳米粒子用作抗癌药物和疫苗的递送系统。一半亲水并且另一半疏水的粒子称为杰那斯(Janus)粒子，并且对于稳定乳液是特别有效的。它们可以在水/油界面处自组装，并且充当固体表面活性剂。

[0605] 粒子表征(包括例如表征形态、尺寸等)是使用多种不同的技术进行的。常用技术是电子显微术(TEM、SEM)、原子力显微镜(AFM)、动态光散射(DLS)、X-射线光电子光谱法(XPS)、粉末X-射线衍射(XRD)、傅里叶变换红外光谱法(FTIR)、基质辅助激光解吸/电离飞行时间质谱法(MALDI-TOF)、紫外-可见光谱法、双偏振干涉法以及核磁共振(NMR)。可以针对针对天然粒子(即加载前)或在加载货物(在本文中货物是指例如CRISPR-Cas系统的一种或多种组分，例如CRISPR-Cas蛋白或mRNA、或腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接子)或mRNA、或指导RNA或它们的任何组合，并且可以包括另外的载剂和/或赋形剂)之后进行表征(尺寸测量)以便为本发明的任何体外、离体和/或体内应用的递送提供具有最佳尺寸的粒子。在某些优选实施方案中，粒子尺寸(例如直径)表征是基于使用动态激光散射(DLS)的测量。关于粒子、其制备和使用方法以及其测量，提及美国专利号8,709,843；美国专利号6,007,845；美国专利号5,855,913；美国专利号5,985,309；美国专利号5,543,158；以及James E.Dahlman和Carmen Barnes等人Nature Nanotechnology(2014)的出版物，2014年5月11日在线公布，doi:10.1038/nnano.2014.84。

[0606] 本发明范围内的粒子递送系统可以采用任何形式提供，包括但不限于固体、半固体、乳液或胶体粒子。这样，本文所述的任何递送系统，包括但不限于例如基于脂质的系统、脂质体、胶束、微泡、外泌体或基因枪，可以被提供作为本发明范围内的粒子递送系统。

[0607] 可以使用粒子或脂质包膜同时递送CRISPR-Cas蛋白mRNA、腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接子)或mRNA，以及指导RNA；例如，本发明的CRISPR-Cas蛋白和RNA，例如作为复合物，可以经由如Dahlman等人、WO2015089419A2及其中引用的文献中的粒子诸如7C1来递送(参见例如James E.Dahlman和Carmen Barnes等人Nature Nanotechnology(2014)，2014年5月11日在线公布，doi:10.1038/nnano.2014.84)，例如递送粒子包含脂质或类脂质(lipidoid)和亲水性聚合物，例如阳离子脂质和亲水性聚合物，例如其中阳离子脂质包括1,2-二油酰基-3-三甲基铵-丙烷(DOTAP)或1,2-二十四酰基-sn-甘油基-3-磷酸胆碱(DMPC)并且/或者其中亲水性聚合物包括乙二醇或聚乙二醇(PEG)；并且/或者其中粒子还包含胆固醇(例如，来自以下项的粒子：配制品1＝DOTAP 100、DMPC 0、PEG 0、胆固醇0；配制品编号2＝DOTAP 90、DMPC 0、PEG 10、胆固醇0；配制品编号3＝DOTAP 90、DMPC 0、PEG 5、胆固醇5)，其中使用有效的多步方法形成粒子，其中第一步，将效应蛋白

和RNA例如在无菌、不含核酸酶的1X PBS中例如在室温下以例如1:1的摩尔比在一起混合例如30分钟;并且分开地,将如适用于该配制品的DOTAP、DMPC、PEG和胆固醇溶解于醇(例如100%乙醇)中;并且,将这两种溶液混合在一起以形成含有这些复合物的粒子)。

[0608] 可以使用粒子或脂质包膜同时递送核酸靶向效应蛋白(例如,V型蛋白诸如Cas13)mRNA和指导RNA。合适的粒子的实例包括但不限于US 9,301,923中描述的那些。

[0609] 例如,Su X、Fricke J、Kavanagh DG、Irvine DJ("In vitro and in vivo mRNA delivery using lipid-enveloped pH-responsive polymer nanoparticles"Mol Pharm.2011年6月6日;8(3):774-87.doi:10.1021/mp100390w.电子版2011年4月1日)描述了生物可降解的核-壳结构粒子,其具有由磷脂双层壳包封的聚(β-氨基酯)(PBAE)核。这些被开发用于体内mRNA递送。该pH响应性PBAE组分被选择为促进内体破坏,而该脂质表面层被选择为将聚阳离子核的毒性最小化。因此,这些对于递送本发明的RNA是优选的。

[0610] 在一个实施方案中,考虑了基于自组装生物粘附聚合物的粒子/纳米粒子,这些粒子可以适用于肽的经口递送、肽的静脉内递送以及肽的鼻递送,均递送至脑。还考虑了其他实施方案,诸如疏水性药物的经口吸收和眼部递送。分子包膜技术涉及被保护并递送至疾病位点的工程化聚合物包膜(参见例如Mazza,M.等人ACSNano,2013.7(2):1016-1026;Siew,A.等人Mol Pharm,2012.9(1):14-28;Lalatsa,A.等人JContr Rel,2012.161(2):523-36;Lalatsa,A.等人,Mol Pharm,2012.9(6):1665-80;Lalatsa,A.等人Mol Pharm,2012.9(6):1764-74;Garrett,N.L.等人J Biophotonics,2012.5(5-6):458-68;Garrett,N.L.等人J Raman Spect,2012.43(5):681-688;Ahmad,S.等人J Royal Soc Interface 2010.7:S423-33;Uchegbu,I.F.Expert Opin Drug Deliv,2006.3(5):629-40;Qu,X.等人Biomacromolecules,2006.7(12):3452-9以及Uchegbu,I.F.等人Int J Pharm,2001.224:185-199)。考虑了约5mg/kg的剂量,取决于靶组织,采用单剂量或多剂量。

[0611] 可以使用由丹·安德森实验室(Dan Anderson's lab)在MIT开发的可以将RNA递送至癌细胞以便使肿瘤生长停止的粒子/纳米粒子并且/或者使这些粒子/纳米粒子适于本发明的AD官能化的CRISPR-Cas系统。具体地说,安德森实验室开发了用于新生物材料和纳米配制品的合成、纯化、表征和配制的全自动化组合系统。参见例如Alabi等人,Proc Natl Acad Sci U S A.2013年8月6号;110(32):12881-6;Zhang等人,Adv Mater.2013年9月6日;25(33):4641-5;Jiang等人,Nano Lett.2013年3月13日;13(3):1059-64;Karagiannis等人,ACS Nano.2012年10月23日;6(10):8484-7;Whitehead等人,ACS Nano.2012年8月28日;6(8):6922-9和Lee等人,Nat Nanotechnol.2012年6月3日;7(6):389-93。

[0612] 美国专利申请20110293703涉及类脂质化合物,这些化合物在多核苷酸的施用中也是特别有用的,它们可以适用于递送本发明的AD官能化的CRISPR-Cas系统。在一方面,氨基醇类脂质化合物与有待递送至细胞或受试者的剂组合而形成微粒子、纳米粒子、脂质体或胶束。有待通过粒子、脂质体或胶束递送的剂可以呈气体、液体或固体的形式,并且该剂可以是多核苷酸、蛋白质、肽或小分子。氨基醇类脂质化合物可以与其他氨基醇类脂质化合物、聚合物(合成的或天然的)、表面活性剂、胆固醇、碳水化合物、蛋白质、脂质等组合而形成粒子。然后这些粒子可以任选地与药物赋形剂组合而形成药物组合物。

[0613] 美国专利公布号20110293703也提供了制备氨基醇类脂质化合物的方法。使胺的一种或多种等效物与环氧化物封端化合物的一种或多种等效物在合适条件下反应而形成

本发明的氨基醇类脂质化合物。在某些实施方案中，胺的所有氨基基团与环氧化物封端化合物充分反应而形成叔胺。在其他实施方案中，胺的所有氨基基团未与环氧化物封端化合物完全反应形成叔胺，由此生成在氨基醇类脂质化合物中的伯胺或仲胺。将这些伯胺或仲胺照原样留下或者可以使其与另一种亲电体诸如不同的环氧化物封端化合物反应。如本领域技术人员将理解的，使胺与未过量的环氧化物封端化合物反应将产生多种不同的具有不同数目的尾部的氨基醇类脂质化合物。某些胺可以被两个环氧化物衍生的化合物尾部完全官能化，而其他分子不会被环氧化物衍生的化合物尾部完全官能化。例如，二胺或多胺可以包括离开该分子的不同氨基部分的一个、二个、三个、或四个环氧化物衍生的化合物尾部，从而产生伯胺、仲胺和叔胺。在某些实施方案中，并不是所有氨基基团都被完全官能化。在某些实施方案中，使用了两种相同类型的环氧化物封端化合物。在其他实施方案中，使用了两种或更多种不同的环氧化物封端化合物。氨基醇类脂质化合物的合成是用或不用溶剂进行的，并且该合成可以在30℃-100℃的范围内，优选在大约50℃-90℃的较高温度下进行。任选地，可以将制备的氨基醇类脂质化合物纯化。例如，可以将氨基醇类脂质化合物的混合物纯化以产生具有特定数目的环氧化物衍生的化合物尾部的氨基醇类脂质化合物。或者可以将混合物纯化以产生特定的立体异构体或区域异构体。可以使用卤代烷(例如碘甲烷)或其他烷化剂将这些氨基醇类脂质化合物烷化，并且/或者可以将它们酰化。

[0614] 美国专利公布号20110293703也提供了通过本发明方法制备的氨基醇类脂质化合物的文库。可以使用涉及液体处理器、机器人、微量滴定板、计算机等的高通量技术制备并且/或者筛选这些氨基醇类脂质化合物。在某些实施方案中，筛选了这些氨基醇类脂质化合物的将多核苷酸或其他剂(例如蛋白质、肽、小分子)转染到细胞中的能力。

[0615] 美国专利公布号20130302401涉及已经使用组合聚合制备的一类聚(β-氨基醇)(PBAA)。本发明的PBAA可以在生物技术和生物医学应用中用作涂层(诸如用于医疗装置或植入物的膜或多层膜的涂层)、添加剂、材料、赋形剂、生物防污剂(non-biofouling agent)、微图案化剂以及细胞封装剂。当用作表面涂层时，这些PBAA在体外和体内均引发不同水平的炎症，这取决于它们的化学结构。这类材料的巨大化学多样性允许我们鉴定出在体外抑制巨噬细胞激活的聚合物涂层。此外，在羧化聚苯乙烯微粒的皮下移植之后，这些涂层减少了炎症细胞的募集，并且减轻了纤维化。这些聚合物可以用于形成用于细胞封装的聚电解质复合物胶囊。本发明还可以具有许多其他的生物应用，诸如抗微生物涂层、DNA或siRNA递送，以及干细胞组织工程化。美国专利公布号20130302401的教导内容可以适用于本发明的AD官能化的CRISPR-Cas系统。

[0616] 可以例如通过电穿孔转染包含Cas13、腺苷脱氨酶(其可以融合至Cas13或衔接蛋白)和指导RNA的预组装重组CRISPR-Cas复合物，从而产生高突变率且不存在可检测的脱靶突变。Hur,J.K.等人,Targeted mutagenesis in mice by electroporation of Cas13 ribonucleoproteins,Nat Biotechnol.2016年6月6日.doi:10.1038/nbt.3596。

[0617] 就局部递送至脑而言，这可以通过多种方式来实现。例如，可以例如通过注射经纹状体内递送材料。注射可以经由颅骨切开术立体定位地进行。

[0618] 在一些实施方案中，可以使用基于糖的粒子，例如GalNAc，如本文所述且参考WO2014118272(以引用方式并入本文)以及Nair,JK等人,2014,Journal of the American Chemical Society 136(49),16958-16961)以及在此的教导内容，除非另外表明，否则特别

涉及适用于所有粒子的递送。这可以被认为是基于糖的粒子，并且本文提供了关于其他粒子递送系统和/或配制品的更多细节。GalNAc因此可以被认为是在本文所述的其他粒子的意义上的粒子，使得一般用途和其他考虑因素(例如所述粒子的递送)也适用于GalNAc粒子。溶液相缀合策略可以例如用于将作为PFP(五氟苯酚)酯激活的三触角GalNAc簇(分子量约2000)附接至5'-己基氨基修饰的寡核苷酸上(5′-HA ASO，分子量约8000Da；Østergaard等人，Bioconjugate Chem.，2015，26(8)，第1451-1455页)。类似地，已描述了用于体内核酸递送的聚(丙烯酸酯)聚合物(参见WO2013158141，该专利以引用方式并入本文)。在另外的替代实施方案中，为了改善递送，可以使用预先混合的CRISPR纳米粒子(或蛋白复合物)与天然存在的血清蛋白(Akinc A等人，2010，Molecular Therapy第18卷第7期，1357-1364)。

[0619] 纳米线团

[0620] 此外，可以使用纳米线团来递送所述系统，例如，如在以下文献中所述的：Sun W等人，Cocoon-like self-degradable DNA nanoclew for anticancer drug delivery.，J Am Chem Soc.2014年10月22日；136(42)：14722-5.doi:10.1021/ja5088024.电子版2014年10月13日；或Sun W等人，Self-Assembled DNA Nanoclews for the Efficient Delivery of CRISPR-Cas9 for Genome Editing.，Angew Chem Int Ed Engl.2015年10月5日；54(41)：12029-33.doi:10.1002/anie.201506030.电子版2015年8月27日。

[0621] LNP

[0622] 在一些实施方案中，递送是通过将Cas13蛋白或mRNA形式封装在脂质粒子诸如LNP中而进行的。因此，在一些实施方案中，考虑了脂质纳米粒子(LNP)。抗转甲状腺素蛋白小干扰RNA已被封装在脂质纳米粒子中并且被递送至人类(参见例如Coelho等人，N Engl J Med 2013；369：819-29)，并且此系统可以适于并应用于本发明的CRISPR Cas系统。考虑了静脉内施用约0.01至约1mg/kg体重的剂量。考虑了降低输注相关反应的风险的药物，诸如考虑了地塞米松、对乙酰氨基酚、苯海拉明或西替利嗪，以及雷尼替丁。还考虑了约0.3mg/kg的多剂量，每4周一次，五个剂量。

[0623] LNP已经显示在将siRNA递送至肝脏中是高度有效的(参见例如Tabernero等人，Cancer Discovery，2013年4月，第3卷，第4期，第363-470页)，并且因此被考虑用于将编码CRISPR Cas的RNA递送至肝脏。可以考虑6mg/kg的LNP的约四个剂量的用量，每两周一次。Tabernero等人证明，在以0.7mg/kg给予LNP前2个周期之后，观察到肿瘤消退，并且在6个周期结束之后，患者已经实现了部分反应，具有淋巴结转移完全消退以及肝脏肿瘤的显著萎缩。在40个剂量之后在此患者中获得完全反应，此患者在接受经过26个月的剂量之后保持缓解和完全治疗。具有RCC和在用VEGF途径抑制剂进行的在先疗法之后进展的包括肾脏、肺以及淋巴结的肝外位点疾病的两位患者在所有位点的疾病都在大约8至12个月内保持稳定，并且一位具有PNET和肝转移的患者继续在18个月(36个剂量)的延伸研究中保持疾病稳定。

[0624] 然而，必须将LNP的电荷考虑在内。当阳离子脂质与带负电的脂质结合时，诱导促进细胞内递送的非双层结构。因为带电荷的LNP在静脉内注射之后迅速从循环中清除，所以开发了具有低于7的pKa值的可电离阳离子脂质(参见例如Rosin等人，Molecular Therapy，第19卷，第12期，第1286-2200页，2011年12月)。带负电的聚合物诸如RNA可以低pH值(例如pH 4)加载到LNP中，在此pH时可电离脂质展示出正电荷。然而，在生理学pH值下，LNP表现出

与更长的循环时间相容的低表面电荷。已经关注了四种可电离阳离子脂质，即1,2-二亚油酰基-3-二甲基铵-丙烷(DLinDAP)、1,2-二亚油基氧基-3-N,N-二甲基氨基丙烷(DLinDMA)、1,2-二亚油基氧基-酮基-N,N-二甲基-3-氨基丙烷(DLinKDMA)，以及1,2-二亚油基-4-(2-二甲基氨基乙基)-[1,3]-二氧戊环(DLinKC2-DMA)。已经显示，含有这些脂质的LNP siRNA系统在体内肝细胞中表现出显著不同的基因沉默特性，具有根据采用因子VII基因沉默模型的DLinKC2-DMA>DLinKDMA>DLinDMA>>DLinDAP系列而变化的潜能(参见例如Rosin等人，Molecular Therapy，第19卷，第12期，第1286-2200页，2011年12月)。可以考虑在LNP中或与LNP相关联的1μg/ml的LNP或CRISPR-Cas RNA的剂量，尤其是对于含有DLinKC2-DMA的配制品而言。

[0625] LNP的制备和CRISPR Cas封装可以使用并且/或者调适自Rosin等人，Molecular Therapy，第19卷，第12期，第1286-2200页，2011年12月)。阳离子脂质1,2-二亚油酰基-3-二甲基铵-丙烷(DLinDAP)、1,2-二亚油基氧基-3-N,N-二甲基氨基丙烷(DLinDMA)、1,2-二亚油基氧基酮基-N,N-二甲基-3-氨基丙烷(DLinK-DMA)、1,2-二亚油基-4-(2-二甲基氨基乙基)-[1,3]-二氧戊环(DLinKC2-DMA)、(3-o-[2″-(甲氧基聚乙二醇2000)琥珀酰基]-1,2-二肉豆蔻酰基-sn-乙二醇(PEG-S-DMG)，以及R-3-[(ω-甲氧基-聚(乙二醇)2000)氨甲酰基]-1,2-二肉豆蔻酰氧基丙基-3-胺(PEG-C-DOMG)可以由Tekmira Pharmaceuticals(Vancouver,Canada)提供或合成。胆固醇可购自Sigma(St Louis,MO)。特定的CRISPR Cas RNA可以封装在含有DLinDAP、DLinDMA、DLinK-DMA和DLinKC2-DMA的LNP中(阳离子脂质:DSPC:CHOL:PEGS-DMG或PEG-C-DOMG，摩尔比为40:10:40:10)。在必要时，可并入0.2％SP-DiOC18(Invitrogen,Burlington,Canada)来评定细胞摄取、细胞内递送和生物分布。可以通过以下方式来进行封装：将由阳离子脂质:DSPC:胆固醇:PEG-c-DOMG(40:10:40:10摩尔比)组成的脂质混合物溶解在乙醇中，直至最终脂质浓度为10mmol/l。可以将脂质的此乙醇溶液逐滴添加到pH 4.0的50mmol/l柠檬酸盐中以形成多层囊泡，从而产生30％乙醇(体积/体积)的终浓度。在使用挤出机(Northern Lipids,Vancouver,Canada)通过两个重叠的80nm Nuclepore聚碳酸酯过滤器挤出多层囊泡之后，可以形成大的单层囊泡。可以通过以下方式来实现封装：将溶解在含有30％乙醇(体积/体积)的pH 4.0的50mmol/l柠檬酸盐中的2mg/ml的RNA逐滴添加到挤出的预成形的大单层囊泡中，并且在31℃下孵育30分钟，伴随持续混合直至最终的RNA/脂质重量比为0.06/1(重量/重量)。通过使用Spectra/Por 2再生纤维素透析膜在pH 7.4的磷酸盐缓冲盐水(PBS)中透析16小时进行乙醇的去除以及配制缓冲液的中和。可以使用NICOMP 370型粒径分析仪、囊泡/强度模式以及高斯拟合通过动态光散射测定纳米粒子粒径分布(Nicomp Particle Sizing,Santa Barbara,CA)。所有三个LNP系统的粒径可能为约70nm。可以通过使用VivaPureD MiniH柱(Sartorius Stedim Biotech)从在透析前后收集的样品中去除游离RNA来确定RNA封装效率。可以从洗脱的纳米粒子中提取封装的RNA并且将其在260nm下量化。通过使用来自Wako Chemicals USA(Richmond,VA)的胆固醇E酶测定测量囊泡中的胆固醇含量来确定RNA与脂质的比率。结合本文对LNP和PEG脂质的论述，PEG化的脂质体或LNP同样适用于CRISPR-Cas系统或其组分的递送。

[0626] 可以在含有50:10:38.5摩尔比的DLinKC2-DMA、DSPC和胆固醇的乙醇中制备脂质预混物溶液(20.4mg/ml总脂质浓度)。可以0.75:1的摩尔比(乙酸钠:DLinKC2-DMA)将乙酸

钠添加到脂质预混物中。随后可以通过将该混合物与1.85倍体积的柠檬酸盐缓冲液(10mmol/l,pH 3.0)在剧烈搅拌下合并来使脂质水合,从而使得在含有35%乙醇的水性缓冲液中自发形成脂质体。可以在37℃下孵育该脂质体溶液以允许粒径的时间依赖性增加。可以通过动态光散射(Zetasizer Nano ZS,Malvern Instruments,Worcestershire,UK)在孵育过程中的不同时间处去除等分试样来研究脂质体尺寸的变化。一旦实现所希望的粒径,就可以将水性PEG脂质溶液(储备溶液=在35%(体积/体积)乙醇中的10mg/ml PEG-DMG)添加到该脂质体混合物中,以产生3.5%总脂质的最终PEG摩尔浓度。在添加PEG-脂质之后,这些脂质体应该其大小,有效抑制进一步生长。然后可以大约1:10(重量:重量)的RNA与总脂质比率将RNA添加到空脂质体中,然后在37℃下孵育30分钟以形成加载的LNP。随后可以将该混合物在PBS中透析过夜,并且用0.45-μm的注射器式过滤器进行过滤。

[0627] 球形核酸(SNA™)构建体和其他纳米粒子(特别是金纳米粒子)也被考虑作为将CRISPR-Cas系统递送至预期靶标的手段。大量数据表明,基于核酸官能化的金纳米粒子的AuraSense Therapeutics'Spherical Nucleic Acid(SNA™)构建体是可用的。

[0628] 可以与本文的教导内容结合使用的文献包括:Cutler等人,J.Am.Chem.Soc.2011133:9254-9257;Hao等人,Small.2011 7:3158-3162;Zhang等人,ACS Nano.20115:6962-6970;Cutler等人,J.Am.Chem.Soc.2012 134:1376-1391;Young等人,Nano Lett.2012 12:3867-71;Zheng等人,Proc.Natl.Acad.Sci.USA.2012 109:11975-80;Mirkin,Nanomedicine 2012 7:635-638;Zhang等人,J.Am.Chem.Soc.2012134:16488-1691;Weintraub,Nature 2013 495:S14-S16,Choi等人,Proc.Natl.Acad.Sci.USA.2013 110(19):7625-7630;Jensen等人,Sci.Transl.Med.5,209ra152(2013)以及Mirkin等人,Small,10:186-192。

[0629] 具有RNA的自组装纳米粒子可以用PEG化的聚乙烯亚胺(PEI)构建,其中Arg-Gly-Asp(RGD)肽配体附接在聚乙二醇(PEG)的远端。例如,这一系统已经被用作靶向表达整合素的肿瘤新血管系统和递送抑制血管内皮生长因子受体2(VEGF R2)表达以及由此实现抑制肿瘤血管新生的siRNA的手段(参见例如Schiffelers等人,Nucleic Acids Research,2004,第32卷,第19期)。纳米束(Nanoplex)可以通过以下方式来制备:将等体积的阳离子聚合物水溶液和核酸水溶液混合,以产生在2至6范围上的可电离氮(聚合物)相比磷酸盐(核酸)的净摩尔过量。在阳离子聚合物与核酸之间的静电相互作用导致聚复合物的形成,该聚合物具有约100nm的平均粒径分布,此后称为纳米束。设想了CRISPR Cas的约100至200mg的剂量,用于Schiffelers等人的自组装纳米粒子中的递送。

[0630] Bartlett等人(PNAS,2007年9月25日,第104卷,第39期)的纳米复合物也可以适用于本发明。Bartlett等人的纳米复合物通过以下方式来制备:将等体积的阳离子聚合物水溶液和核酸水溶液混合,以产生在2至6范围上的可电离氮(聚合物)相比磷酸盐(核酸)的净摩尔过量。在阳离子聚合物与核酸之间的静电相互作用导致聚复合物的形成,该聚合物具有约100nm的平均粒径分布,此后称为纳米束。Bartlett等人的DOTA-siRNA的合成如下:1,4,7,10-四氮杂环十二烷-1,4,7,10-四乙酸单(N-羟基琥珀酰亚胺酯)(DOTA-NHS酯)订购自Macrocyclics(Dallas,TX)。将于碳酸盐缓冲液(pH 9)中的具有100倍摩尔过量的DOTA-NHS-酯的胺修饰的RNA有义链添加到微量离心管中。通过在室温下搅拌4小时使这些内容物反应。将DOTA-RNA有义缀合物用乙醇沉淀,重新悬浮在水中,并且退火到未修饰的反义链上

以产生DOTA-siRNA。所有液体均用Chelex-100(Bio-Rad,Hercules,CA)预处理,以去除痕量金属污染物。可以通过使用含有环糊精的聚阳离子形成Tf靶向和非靶向的siRNA纳米粒子。典型地,以3(+/-)的进料比和0.5克/升的siRNA浓度在水中形成纳米粒子。用Tf(金刚烷-PEG-Tf)修饰在靶向纳米粒子表面上的百分之一的金刚烷-PEG分子。将纳米粒子悬浮在用于注射的5%(重量/体积)葡萄糖载剂溶液中。

[0631] Davis等人(Nature,第464卷,2010年4月15日)进行了使用靶向的纳米粒子递送系统的RNA临床试验(临床试验登记号NCT00689065)。在21天周期的第1、3、8和10天通过30min的静脉内输注向患有标准护理疗法难治的实体癌的患者施用靶向的纳米粒子剂量。纳米粒子由合成递送系统组成,该系统含有:(1)线性的基于环糊精的聚合物(CDP);(2)展示在纳米粒子外部上的用于接合癌细胞表面上的TF受体(TFR)的人类转铁蛋白(TF)靶向配体;(3)亲水性聚合物(用于促进纳米粒子在生物流体中的稳定性的聚乙二醇(PEG));以及(4)被设计来降低RRM2(先前在临床中使用的序列指代为siR2B+5)表达的siRNA。长久以来已知TFR在恶性细胞中被下调,并且RRM2是一种确立的抗癌靶标。已经显示这些纳米粒子(临床版本指代为CALAA-01)在非人类灵长类动物中的多剂量研究中耐受性良好。虽然已经通过脂质体递送向患有慢性粒细胞白血病的单一患者施用了siRNA,但是Davis等人的临床试验是初期人类试验,该试验用靶向递送系统全身性地递送siRNA并且治疗患有实体癌的患者。为了确定该靶向递送系统是否能够将功能性siRNA有效递送至人类肿瘤,Davis等人研究了来自三个不同的剂量组群的三位患者的活组织检查;患者A、B和C,他们均患有转移性黑素瘤并且分别接受了18、24和30mg m-2siRNA的CALAA-01剂量。对于本发明的CRISPR Cas系统,还可以考虑类似的剂量。用含有线性的基于环糊精的聚合物(CDP)、展示在纳米粒子外部上的用于接合癌细胞表面上的TF受体(TFR)的人类转铁蛋白(TF)靶向配体和/或亲水性聚合物(例如,用于促进纳米粒子在生物流体中的稳定性的聚乙二醇(PEG))的纳米粒子,可以实现本发明的递送。

[0632] 以引用方式并入本文的美国专利号8,709,843提供了用于将含有治疗剂的粒子靶向递送至组织、细胞和细胞内区室的药物递送系统。本发明提供了包含缀合至表面活性剂、亲水性聚合物或脂质的聚合物的靶向粒子。以引用方式并入本文的美国专利号6,007,845提供了下述粒子:所述粒子具有通过将多官能化合物与一种或多种疏水性聚合物和一种或多种亲水性聚合物共价地连接而形成的多嵌段共聚合物的核,并且含有生物活性材料。以引用方式并入本文的美国专利No.5,855,913提供了下述颗粒组合物:所述颗粒组合物含有具有小于0.4g/cm3振实密度以及介于5μm与30μm之间的平均直径的空气动力学光粒子,在其表面上并入了表面活性剂,以用于向肺部系统的药物递送。以引用方式并入本文的美国专利号5,985,309提供了下述粒子:所述粒子并入了表面活性剂和/或带正电荷或带负电荷的治疗剂或诊断剂和带相反电荷的带电荷分子的亲水性或疏水性复合物,以用于向肺部系统的递送。以引用方式并入本文的美国专利号5,543,158提供了生物可降解的可注射粒子,所述粒子具有生物可降解实心核,该核在其表面上含有生物活性材料和聚(烷撑二醇)部分。以引用方式并入本文的WO2012135025(也以US20120251560公布)描述了缀合的聚乙烯亚胺(PEI)聚合物和缀合的氮杂大环(统称为“缀合微质体(lipomer)”或“微质体”)。在某些实施方案中,可以设想的是,此类缀合微质体可以用于CRISPR-Cas系统的情形中以在体外、离体和在体内实现基因组干扰,从而修饰基因表达,包括调节蛋白质表达。

[0633] 在一个实施方案中，纳米粒子可以是环氧化物修饰的脂质-聚合物，有利地是7C1(参见例如James E.Dahlman和Carmen Barnes等人Nature Nanotechnology(2014)，2014年5月11日在线发布，doi:10.1038/nnano.2014.84)。C71是通过使C15环氧化物封端脂质与PEI600以14:1摩尔比反应合成的，并且与C14PEG2000一起进行配制以产生纳米粒子(直径介于35与60nm之间)，这些纳米粒子在PBS溶液中在至少40天内保持稳定。

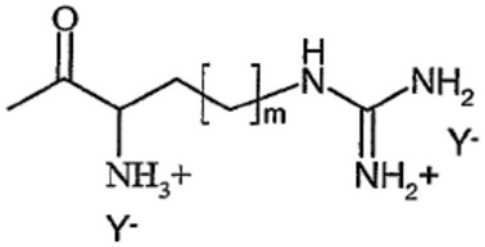
[0634] 环氧化物修饰的脂质-聚合物可以用于将本发明的CRISPR-Cas系统递送至肺部细胞、心血管细胞或肾细胞，然而本领域的技术人员可以调适该系统以递送至其他靶器官。设想了约0.05至约0.6mg/kg的剂量范围。还设想了经数天或数周的剂量，其中总剂量为约2mg/kg。

[0635] 在一些实施方案中，用于递送RNA分子的LNP通过本领域已知的方法来制备，所述方法诸如在例如WO 2005/105152(PCT/EP2005/004920)、WO 2006/069782(PCT/EP2005/014074)、WO 2007/121947(PCT/EP2007/003496)和WO 2015/082080(PCT/EP2014/003274)中所述的那些方法，这些文献以引用方式并入本文。明确地旨在增强和改善siRNA到哺乳动物细胞中的递送的LNP描述于例如Aleku等人，Cancer Res.，68(23):9788-98(2008年12月1日)；Strumberg等人，Int.J.Clin.Pharmacol.Ther.，50(1):76-8(2012年1月)；Schultheis等人，J.Clin.Oncol.，32(36):4141-48(2014年12月20日)以及Fehring等人，Mol.Ther.，22(4):811-20(2014年4月22日)(这些文献以引用方式并入本文)，并且可以应用于本发明的技术。

[0636] 在一些实施方案中，LNP包括在WO 2005/105152(PCT/EP2005/004920)、WO 2006/069782(PCT/EP2005/014074)、WO 2007/121947(PCT/EP2007/003496)和WO 2015/082080(PCT/EP2014/003274)中所公开的任何LNP。

[0637] 在一些实施方案中，LNP包含至少一种具有式I的脂质：

[0638] (Formula I)，其中R1和R2各自且独立地选自包括烷基的组，n是1与4之间的任何整数，并且R3是选自包括赖氨酰基、鸟氨酰基、2,4-二氨基丁酰基、组氨酰基和根据式II的酰基部分的组的酰基：

[0639] 

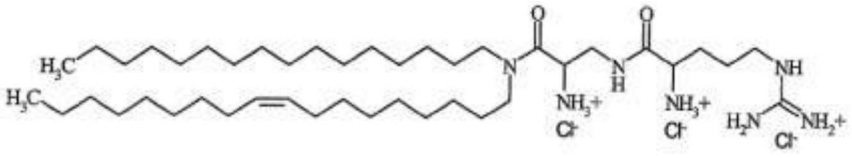
(式 II)，

[0640] 其中m是1至3的任何整数，并且Y^-是药学上可接受的阴离子。在一些实施方案中，根据式I的脂质包括至少两个不对称的C原子。在一些实施方案中，式I的对映异构体包括但不限于R-R;S-S;R-S和S-R对映异构体。

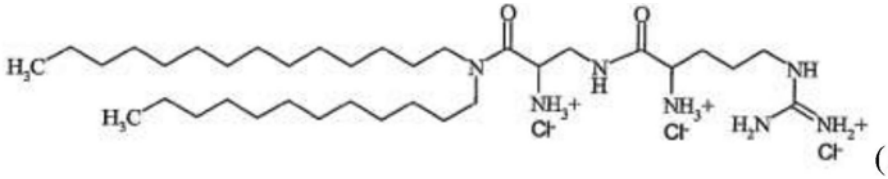
[0641] 在一些实施方案中，R1是月桂基并且R2是肉豆蔻基。在另一个实施方案中，R1是棕榈基并且R2是油基。在一些实施方案中，m是1或2。在一些实施方案中，Y-选自卤化物、乙酸盐或三氟乙酸盐。

[0642] 在一些实施方案中，LNP包含一种或多种选自以下的脂质：

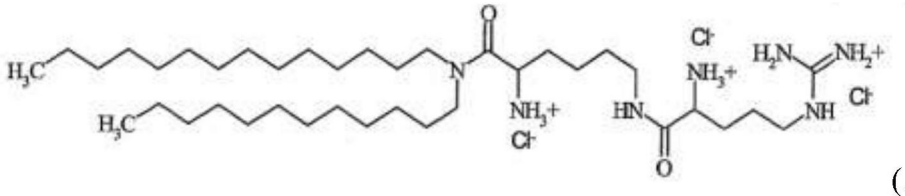
[0643] -精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐(式III)：

[0644] 
(式 III);

[0645] -精氨酰基-2,3-二氨基丙酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐(式IV)：

[0646] 
以及
(式 IV);

[0647] -精氨酰基-赖氨酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐(式V)：

[0648] 
(式 V)。

[0649] 在一些实施方案中,LNP还包含某种成分。举例来说,但不作为限制,在一些实施方案中,所述成分选自肽、蛋白质、寡核苷酸、多核苷酸、核酸或它们的组合。在一些实施方案中,所述成分是抗体,例如单克隆抗体。在一些实施方案中,所述成分是选自例如核酶、适体、镜像异构适配体(spiegelmer)、DNA、RNA、PNA、LNA或它们的组合的核酸。在一些实施方案中,核酸是指导RNA和/或mRNA。

[0650] 在一些实施方案中,LNP的成分包括编码CRIPSR-Cas蛋白的mRNA。在一些实施方案中,LNP的成分包括编码II型或V型CRIPSR-Cas蛋白的mRNA。在一些实施方案中,LNP的成分包括编码腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或衔接蛋白)的mRNA。

[0651] 在一些实施方案中,LNP的成分还包括一种或多种指导RNA。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至血管内皮。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至肺内皮。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至肝脏。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至肺。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至心脏。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至脾。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至肾脏。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至胰腺。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至脑。在一些实施方案中,LNP被配置成递送前述mRNA并将RNA导向至巨噬细胞。

[0652] 在一些实施方案中,LNP还包含至少一种辅助脂质。在一些实施方案中,辅助脂质选自磷脂和类固醇。在一些实施方案中,磷脂是磷酸的二酯和/或单酯。在一些实施方案中,磷脂是磷酸甘油酯和/或鞘脂。在一些实施方案中,类固醇是天然存在的和/或基于部分氢化的环戊[a]菲的合成化合物。在一些实施方案中,类固醇含有21至30个C原子。在一些实施方案中,类固醇是胆固醇。在一些实施方案中,辅助脂质选自1,2-二植烷酰基-sn-甘油基-3-磷酸乙醇胺(DPhyPE)、神经酰胺和1,2-二油酰基sn-甘油基-3-磷酸乙醇胺(DOPE)。

[0653] 在一些实施方案中,所述至少一种辅助脂质包含选自包括PEG部分、HEG部分、聚羟乙基淀粉(聚HES)部分和聚丙烯部分的组的部分。在一些实施方案中,所述部分具有在约

500至10,000Da之间或在约2,000至5,000Da之间的分子量。在一些实施方案中,PEG部分选自1,2-二硬脂酰基-sn-甘油基-3磷酸乙醇胺、1,2-二烷基-sn-甘油基-3-磷酸乙醇胺和神经酰胺-PEG。在一些实施方案中,PEG部分具有在约500至10,000Da之间或在约2,000至5,000Da之间的分子量。在一些实施方案中,PEG部分具有2,000Da的分子量。

[0654] 在一些实施方案中,辅助脂质为组合物的总脂质含量的约20摩尔%至80摩尔%。在一些实施方案中,辅助脂质组分为LNP的总脂质含量的约35摩尔%至65摩尔%。在一些实施方案中,LNP包含占LNP的总脂质含量的50摩尔%的脂质和50摩尔%的辅助脂质。

[0655] 在一些实施方案中,LNP包含-3-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐、-精氨酰基-2,3-二氨基丙酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐或-精氨酰基-赖氨酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐中的任一者与DPhyPE的组合,其中DPhyPE的含量为LNP的总脂质含量的约80摩尔%、65摩尔%、50摩尔%和35摩尔%。在一些实施方案中,LNP包含-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐(脂质)和1,2-二植烷酰基-sn-甘油基-3-磷酸乙醇胺(辅助脂质)。在一些实施方案中,LNP包含-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐(脂质)、1,2-二植烷酰基-sn-甘油基-3-磷酸乙醇胺(第一辅助脂质)和1,2-二硬脂酰基-sn-甘油基-3-磷酸乙醇胺-PEG2000(第二辅助脂质)。

[0656] 在一些实施方案中,第二辅助脂质为总脂质含量的约0.05摩尔%至4.9摩尔%之间或约1摩尔%至3摩尔%之间。在一些实施方案中,LNP包含为总脂质含量的约45摩尔%至50摩尔%之间的脂质、为总脂质含量的约45摩尔%至50摩尔%之间的第一辅助脂质,条件是存在为总脂质含量的约0.1摩尔%至5摩尔%之间、约1摩尔%至4摩尔%之间或约2摩尔%的PEG化的第二辅助脂质,其中脂质、第一辅助脂质和第二辅助脂质的含量之和为总脂质含量的100摩尔%,并且其中第一辅助脂质和第二辅助脂质之和为总脂质含量的50摩尔%。在一些实施方案中,LNP包含:(a)50摩尔%的-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐、48摩尔%的1,2-二植烷酰基-sn-甘油基-3-磷酸乙醇胺;以及2摩尔%的1,2-二硬脂酰基-sn-甘油基-3-磷酸乙醇胺-PEG2000;或者(b)50摩尔%的-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐、49摩尔%的1,2-二植烷酰基-sn-甘油基-3-磷酸乙醇胺;以及1摩尔%的N(羰基-甲氧基聚乙二醇-2000)-1,2-二硬脂酰基-sn-甘油基3-磷酸乙醇胺,或其钠盐。

[0657] 在一些实施方案中,LNP含有核酸,其中核酸骨架磷酸盐与阳离子脂质氮原子的进料比为约1:1.5-7或约1:4。

[0658] 在一些实施方案中,LNP还包含屏蔽化合物,所述屏蔽化合物可以在体内条件下从脂质组合物中去除。在一些实施方案中,屏蔽化合物是生物惰性化合物。在一些实施方案中,屏蔽化合物在其表面或分子本身上不携带任何电荷。在一些实施方案中,屏蔽化合物是聚乙二醇(PEG)、基于羟乙基葡萄糖(HEG)的聚合物、聚羟乙基淀粉(聚HES)和聚丙烯。在一些实施方案中,PEG、HEG、聚HES和聚丙烯的重量为约500至10,000Da之间或约2000至5000Da之间。在一些实施方案中,屏蔽化合物是PEG2000或PEG5000。

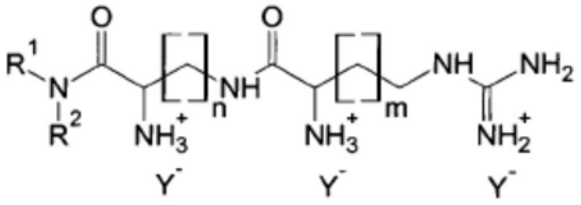
[0659] 在一些实施方案中,LNP包含至少一种脂质、第一辅助脂质和在体内条件下可以从脂质组合物中去除的屏蔽化合物。在一些实施方案中,LNP还包含第二辅助脂质。在一些实施方案中,第一辅助脂质是神经酰胺。在一些实施方案中,第二辅助脂质是神经酰胺。在一

些实施方案中，神经酰胺包含至少一个6至10个碳原子的短碳链取代基。在一些实施方案中，神经酰胺包含8个碳原子。在一些实施方案中，屏蔽化合物附接至神经酰胺。在一些实施方案中，屏蔽化合物附接至神经酰胺。在一些实施方案中，屏蔽化合物共价附接至神经酰胺。在一些实施方案中，屏蔽化合物附接至LNP中的核酸。在一些实施方案中，屏蔽化合物共价附接至核酸。在一些实施方案中，屏蔽化合物通过接头附接至核酸。在一些实施方案中，接头在生理条件下被切割。在一些实施方案中，接头选自ssRNA、ssDNA、dsRNA、dsDNA、肽、S-S接头和pH敏感接头。在一些实施方案中，接头部分附接至核酸有义链的3'端。在一些实施方案中，屏蔽化合物包含pH敏感接头或pH敏感部分。在一些实施方案中，pH敏感接头或pH敏感部分是阴离子接头或阴离子部分。在一些实施方案中，在酸性环境中，阴离子接头或阴离子部分带较少的阴离子或为中性的。在一些实施方案中，pH敏感接头或pH敏感部分选自低聚(谷氨酸)、低聚酚盐和二亚乙基三胺五乙酸。

[0660] 在先前段落中的任何LNP实施方案中，LNP可以具有约50至600mosmole/kg之间、约250至350mosmole/kg之间或约280至320mosmole/kg之间的重量克分子渗透压浓度，并且/或者其中由脂质和/或一种或两种辅助脂质和屏蔽化合物形成的LNP具有约20至200nm之间、约30至100nm之间或约40至80nm之间的粒度。

[0661] 在一些实施方案中，屏蔽化合物提供了更长的体内循环时间，并且允许含核酸的LNP的更好的生物分布。在一些实施方案中，屏蔽化合物防止LNP与血清化合物或其他体液或细胞质膜(例如，向其施用LNP的脉管系统的内皮层的细胞质膜)的化合物相互作用。另外地或可替代地，在一些实施方案中，屏蔽化合物还防止免疫系统的元件立即与LNP相互作用。另外地或可替代地，在一些实施方案中，屏蔽化合物充当抗调理化合物。不希望受到任何机制或理论的束缚，在一些实施方案中，屏蔽化合物形成覆盖物或外壳，所述覆盖物或外壳减少了LNP可用于与其环境相互作用的表面区域。另外地或可替代地，在一些实施方案中，屏蔽化合物屏蔽LNP的总电荷。

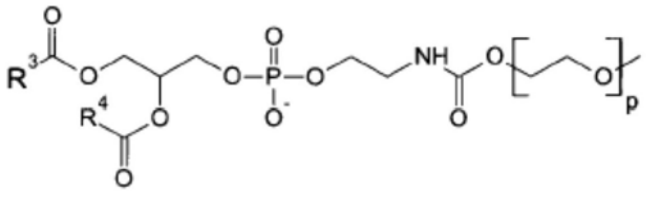
[0662] 在另一个实施方案中，LNP包含至少一种具有式VI的脂质：

[0663] 

(式 VI)，

[0664] 其中n是1、2、3或4，其中m是1、2或3，其中Y^-是阴离子，其中R^1和R^2各自单独地且独立地选自由以下组成的组：直链C12-C18烷基和直链C12-C18烯基；固醇化合物，其中所述固醇化合物是选自由胆固醇和豆甾醇组成的组；以及PEG化脂质，其中所述PEG化脂质包含PEG部分，其中所述PEG化脂质选自由以下组成的组：

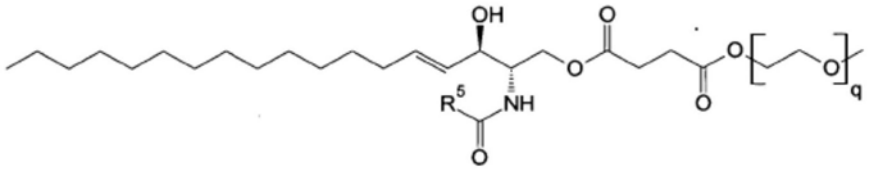
[0665] 式VII的PEG化磷酸乙醇胺：

[0666] 

(式 VII)，

[0667] 其中R^3和R^4单独地且独立地是直链C13-C17烷基，并且p是15至130之间的任何整数；

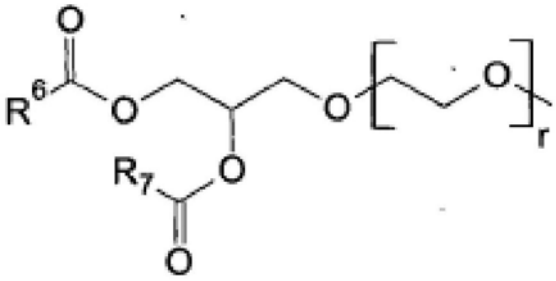
[0668] 式VIII的PEG化神经酰胺：

[0669] 

(式 VIII)；

[0670] 其中R^5是直链C7-C15烷基，并且q是15至130之间的任何数字；以及

[0671] 式IX的PEG化二酰基甘油：

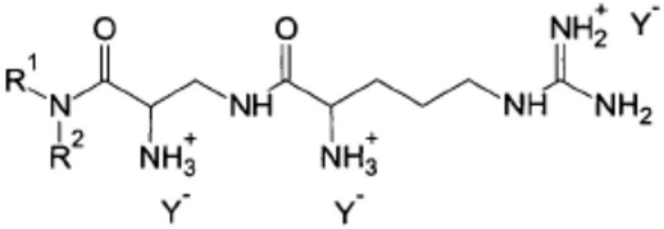
[0672] 

(式 IX)，

[0673] 其中R^6和R^7各自单独地且独立地是直链C11-C17烷基，并且r是15至130的任何整数。

[0674] 在一些实施方案中，R^1和R^2彼此不同。在一些实施方案中，R^1是棕榈基并且R^2是油基。在一些实施方案中，R^1是月桂基并且R^2是肉豆蔻基。在一些实施方案中，R^1和R^2是相同的。在一些实施方案中，R^1和R^2各自单独地且独立地选自由以下组成的组：C12烷基、C14烷基、C16烷基、C18烷基、C12烯基、C14烯基、C16烯基和C18烯基。在一些实施方案中，C12烯基、C14烯基、C16烯基和C18烯基各自包含一个或两个双键。在一些实施方案中，C18烯基是在C9与C10之间具有一个双键的C18烯基。在一些实施方案中，C18烯基是顺式-9-十八烷基。

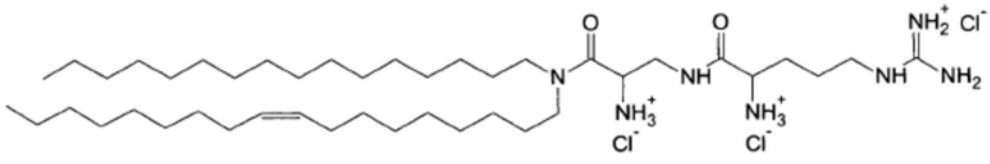
[0675] 在一些实施方案中，阳离子脂质是式X的化合物：

[0676] 

在一些实施方案中，

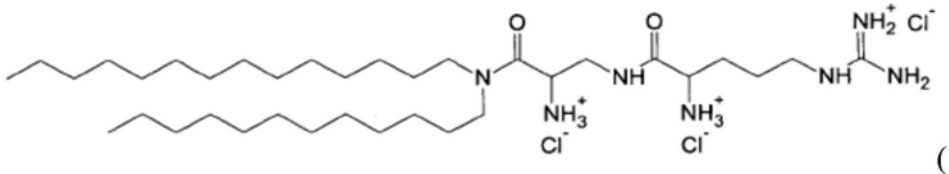
(式 X)。

Y^-选自卤化物、乙酸盐和三氟乙酸盐。在一些实施方案中，阳离子脂质是式III的-精氨酰基-2,3-二氨基丙酸-N-棕榈基-N-油基-酰胺三盐酸盐：

[0677] 

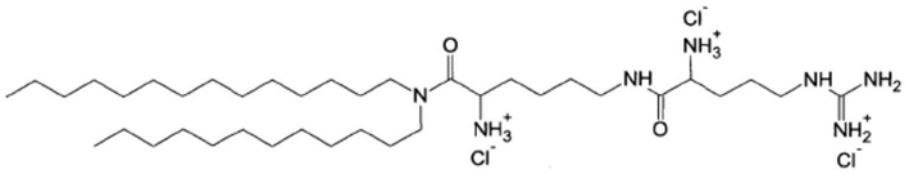
(式 III)。

[0678] 在一些实施方案中，阳离子脂质是式IV的-精氨酰基-2,3-二氨基丙酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐：

[0679] 

(式 IV)。

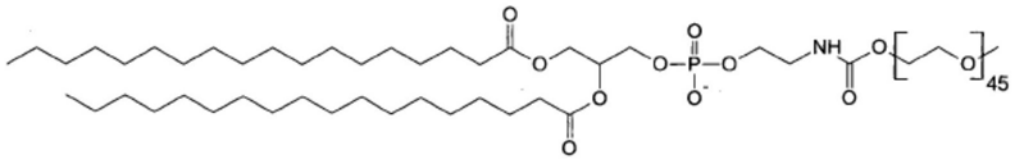
在一些实施方案中,阳离子脂质是式V的-精氨酰基-赖氨酸-N-月桂基-N-肉豆蔻基-酰胺三盐酸盐:

[0680] 

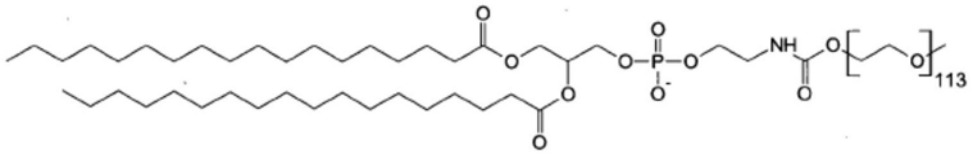
(式 V)。

[0681] 在一些实施方案中,固醇化合物是胆固醇。在一些实施方案中,固醇化合物是豆固醇(stigmasterin)。

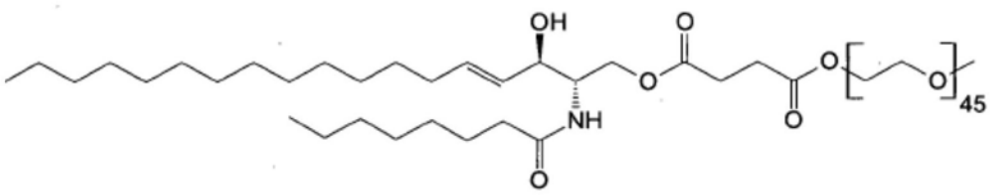
[0682] 在一些实施方案中,PEG化脂质的PEG部分具有约800至5,000Da的分子量。在一些实施方案中,PEG化脂质的PEG部分的分子量为约800Da。在一些实施方案中,PEG化脂质的PEG部分的分子量为约2,000Da。在一些实施方案中,PEG化脂质的PEG部分的分子量为约5,000Da。在一些实施方案中,PEG化脂质是式VII的PEG化磷酸乙醇胺,其中R^3和R^4各自单独地且独立地是直链C13-C17烷基,并且p是18、19或20,或44、45或46,或113、114或115中的任何整数。在一些实施方案中,R^3和R^4是相同的。在一些实施方案中,R^3和R^4是不同的。在一些实施方案中,R^3和R^4各自单独地且独立地选自由以下组成的组:C13烷基、C15烷基和C17烷基。在一些实施方案中,式VII的PEG化磷酸乙醇胺是1,2-二硬脂酰基-sn-甘油基-3-磷酸乙醇胺-N-[甲氧基(聚乙二醇)-2000](铵盐):

[0683] 

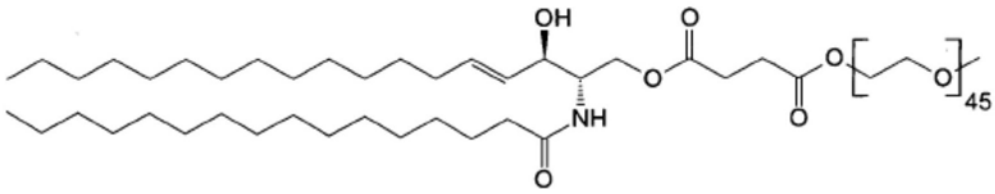
[0684] (式XI)。在一些实施方案中,式VII的PEG化磷酸乙醇胺是1,2-二硬脂酰基-sn-甘油基-3-磷酸乙醇胺-N-[甲氧基(聚乙二醇)-5000](铵盐):

[0685] 

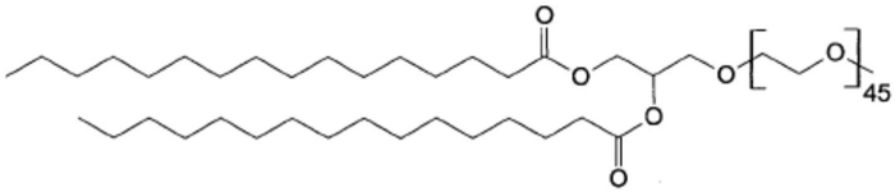
[0686] (式XII)。在一些实施方案中,PEG化脂质是式VIII的PEG化神经酰胺,其中R^5是直链C7-C15烷基,并且q是18、19或20,或44、45或46,或113、114或115中的任何整数。在一些实施方案中,R^5是直链C7烷基。在一些实施方案中,R^5是直链C15烷基。在一些实施方案中,式VIII的PEG化神经酰胺是N-辛酰基-鞘氨醇-1-{琥珀酰基[甲氧基(聚乙二醇)2000]}:

[0687] 

[0688] (式XIII)。在一些实施方案中,式VIII的PEG化神经酰胺是N-棕榈酰基-鞘氨醇-1-{琥珀酰基[甲氧基(聚乙二醇)2000]}:

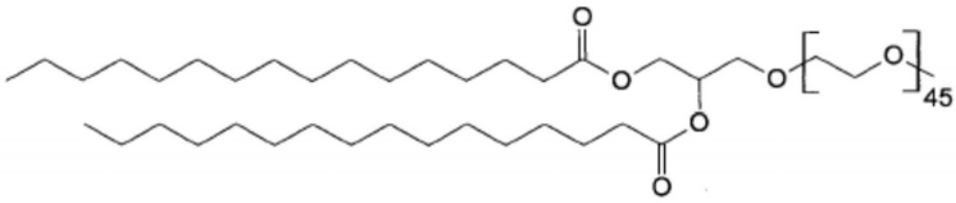
[0689] 

[0690] (式XIV)。在一些实施方案中,PEG化脂质是式IX的PEG化二酰基甘油,其中R^6和R^7各自单独地且独立地是直链C11-C17烷基,并且r是18、19或20,或44、45或46,或113、114或115中的任何整数。在一些实施方案中,R^6和R^7是相同的。在一些实施方案中,R^6和R^7是不同的。在一些实施方案中,R^6和R^7各自单独地且独立地选自由以下组成的组:C17烷基、直链C15烷基和直链C13烷基。在一些实施方案中,式IX的PEG化二酰基甘油是1,2-二硬脂酰基-sn-甘油[甲氧基(聚乙二醇)2000]:

[0691] 

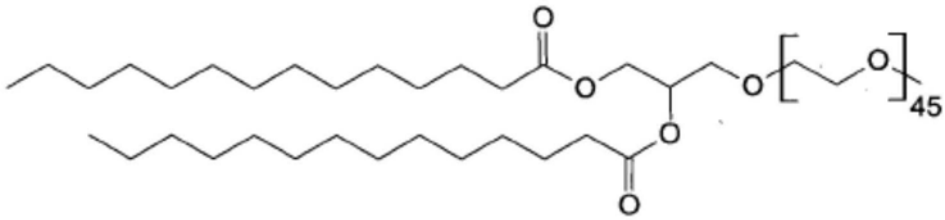
(式 XV)。

在一些实施方案中,式IX的PEG化二酰基甘油是1,2-二棕榈酰基-sn-甘油[甲氧基(聚乙二醇)2000]:

[0692] 

(式 XVI)。

在一些实施方案中,式IX的PEG化二酰基甘油是:

[0693] 

(式 XVII)。

在一些实施方案中,LNP包含至少一种选自式III、IV和V的阳离子脂质、至少一种选自胆固醇和豆固醇的固醇,并且其中PEG化脂质是选自式XI和XII中的至少一者。在一些实施方案中,LNP包含至少一种选自式III、IV和V的阳离子脂质、至少一种选自胆固醇和豆固醇的固醇,并且其中PEG化脂质是选自式XIII和XIV中的至少一者。在一些实施方案中,LNP包含至少一种选自式III、IV和V的阳离子脂质、至少一种选自胆固醇和豆固醇的固醇,并且其中PEG化脂质是选自式XV和XVI中的至少一者。在一些实施方案中,LNP包含式III的阳离子脂质、作为固醇化合物的胆固醇,并且其中PEG化脂质是式XI。

[0694] 在先前段落中的任何LNP实施方案中,其中阳离子脂质组合物的含量在约65摩尔%至75摩尔%之间,固醇化合物的含量在约24摩尔%至34摩尔%之间,并且PEG化脂质的

含量在约0.5摩尔％至1.5摩尔％之间，其中脂质组合物中阳离子脂质、固醇化合物和PEG化脂质的含量之和为100摩尔％。在一些实施方案中，阳离子脂质为约70摩尔％，固醇化合物的含量为约29摩尔％，并且PEG化脂质的含量为约1摩尔％。在一些实施方案中，LNP为70摩尔％的式III、29摩尔％的胆固醇和1摩尔％的式XI。

[0695] 外泌体

[0696] 外泌体是转运RNA和蛋白质的内源性纳米囊泡，并且可以将RNA递送至脑和其他靶器官。为了降低免疫原性，Alvarez-Erviti等人(2011,Nat Biotechnol 29:341)使用了用于外泌体产生的自我衍生的树突细胞。通过将树突细胞工程化为表达Lamp2b(一种外泌体膜蛋白，融合至神经元特异性RVG肽)实现对脑的靶向。通过电穿孔使纯化的外泌体加载外源性RNA。静脉内注射的RVG靶向的外泌体将GAPDH siRNA特异性地递送至脑中的神经元、小胶质细胞、少突神经胶质细胞，导致特异性的基因敲低。预暴露于RVG外泌体未减弱敲低，并且在其他组织中未观察到非特异性摄取。通过BACE1的强的mRNA(60％)和蛋白质(62％)敲低证明了外泌体介导的siRNA递送的治疗潜能，BACE1是阿尔茨海默病中的治疗靶标。

[0697] 为了获得免疫惰性的外泌体库，Alvarez-Erviti等人收获了来自具有均质主要组织相容性复合体(MHC)单倍型的近交C57BL/6小鼠的骨髓。由于未成熟树突细胞产生大量的缺乏T细胞激活物诸如MHC-II和CD86的外泌体，Alvarez-Erviti等人选择了具有粒细胞/巨噬细胞集落刺激因子(GM-CSF)的树突细胞，持续7天。次日，使用良好建立的超速离心方案从培养上清液中纯化外泌体。产生的外泌体在物理上是均质的，具有直径为80nm的粒径分布峰，正如通过纳米粒子跟踪分析(NTA)和电子显微术所测定。Alvarez-Erviti等人获得了6-12μg的外泌体(基于蛋白质浓度测量的)/106个细胞。

[0698] 接着，Alvarez-Erviti等人研究了使用适于纳米级应用的电穿孔方案给修饰的外泌体加载外源性货物的可能性。由于电穿孔对于纳米级的膜粒子尚未良好表征，使用非特异性Cy5标记的RNA用于电穿孔方案的经验优化。在外泌体超速离心和溶解之后测定了封装的RNA量。在400V和125μF下的电穿孔产生RNA的最大保留并且用于所有的后续实验。

[0699] Alvarez-Erviti等人向正常C57BL/6小鼠施用被封装在150μg的RVG外泌体中的150μg的每种BACE1 siRNA并且将敲低效率与四只对照小鼠进行比较：未处理的小鼠、仅用RVG外泌体注射的小鼠、用与体内阳离子脂质体试剂复合的BACE1 siRNA注射的小鼠、以及用与RVG-9R复合的BACE1 siRNA注射的小鼠，该RVG肽与静电结合至siRNA的9个D-精氨酸缀合。在施用之后3天，分析皮层组织样品，并且在siRNA-RVG-9R处理的小鼠和siRNARVG外泌体处理的小鼠中均观察到显著的蛋白质敲低(45％，$P<0.05$，相对于62％，$P<0.01$)，这是由于BACE1 mRNA水平的显著降低(分别为66％[+或-]15％，$P<0.001$和61％[+或-]13％，$P<0.01$)。此外，申请人证明了在RVG外泌体处理的动物中总[β]-淀粉样蛋白1-42水平上的显著降低(55％，$P<0.05$)，该β淀粉样蛋白是在阿尔茨海默病理学中的淀粉样白斑的主要组分。所观察到的降低大于在心室内注射BACE1抑制剂之后的正常小鼠中展示的β淀粉样蛋白1-40降低。Alvarez-Erviti等人在BACE1切割产物上进行了5'-cDNA末端快速扩增(RACE)，这提供了经由siRNA的RNAi-介导的敲低的证据。

[0700] 最后，Alvarez-Erviti等人通过评定IL-6、IP-10、TNFα和IFN-α血清浓度研究了RNA-RVG外泌体是否诱导了体内免疫反应。在外泌体处理之后，类似于与强有力地刺激IL-6分泌的siRNA-RVG-9R成对比的siRNA转染试剂处理，登记了在所有细胞因子上的非显著性

变化，证实了外泌体处理的免疫惰性属性(profile)。假定外泌体仅封装20％的siRNA，用RVG外泌体的递送比RVG-9R递送显得更有效，因为用少五倍的siRNA实现了相当的mRNA敲低和更好的蛋白质敲低，而没有相应水平的免疫刺激。这个实验证明了RVG外泌体技术的治疗潜力，这种技术潜在地适用于与神经变性疾病相关的基因的长期沉默。Alvarez-Erviti等人的外泌体递送系统可以适用于将本发明的AD官能化的CRISPR-Cas系统递送至治疗靶标，尤其是神经变性疾病。对于本发明可以考虑封装在约100至1000mg的RVG外泌体中的约100至1000mg的CRISPR Cas的剂量。

[0701] El-Andaloussi等人(Nature Protocols 7,2112-2126(2012))公开了可以如何利用来源于培养细胞的外泌体用于体外和体内递送RNA。这个方案首先描述了通过转染包含与肽配体融合的外泌体蛋白的表达载体产生靶向的外泌体。接着，El-Andaloussi等人解释了如何纯化和表征来自转染的细胞上清液的外泌体。接着，El-Andaloussi等人详述了将RNA加载到外泌体中的关键步骤。最后，El-Andaloussi等人概述了如何使用外泌体有效地在体外递送RNA以及体内递送至小鼠脑中。还提供了预期结果的实例，其中外泌体介导的RNA递送通过功能测定和成像来评估。整个方案进行约3周。根据本发明的递送或施用可以使用由自我衍生的树突细胞产生的外泌体进行。根据本文的教导内容，这可以用于本发明的实践中。

[0702] 在另一个实施方案中，考虑了Wahlgren等人(Nucleic Acids Research,2012,第40卷,第17期e130)的血浆外泌体。外泌体是由包括树突细胞(DC)、B细胞、T细胞、肥大细胞、上皮细胞和肿瘤细胞的许多细胞类型产生的纳米尺寸的囊泡(30-90nm大小)。这些囊泡通过晚期内体的向内出芽而形成，然后在与质膜融合后释放到细胞外环境。因为外泌体天然地在细胞之间运送RNA，所以这种特性在基因疗法中可能有用，并且根据本公开可以用于本发明的实践中。

[0703] 来自血浆的外泌体可以通过以下方式制备：在900g离心血沉棕黄层持续20分钟以便分离血浆，之后收获细胞上清液，在300g离心10分钟以便消除细胞，并且在16 500g离心30分钟，之后通过0.22mm过滤器进行过滤。通过在120 000g超速离心70分钟使外泌体沉淀。根据在RNAi人类/小鼠启动试剂盒(RNAi Human/Mouse Starter Kit,Quiagen,Hilden,Germany)中的制造商的说明进行siRNA到外泌体中的化学转染。siRNA以终浓度2mmol/ml添加到100ml PBS中。在加入HiPerFect转染试剂之后，将混合物在室温下孵育10分钟。为了去除过量的胶束，使用醛/硫酸盐乳胶珠再分离外泌体。可以类似于siRNA进行CRISPR Cas到外泌体中的化学转染。外泌体可以与从健康供体的外周血中分离的单核细胞和淋巴细胞共培养。因此，可以考虑的是，可以将含有CRISPR Cas的外泌体引入到人类的单核细胞和淋巴细胞中并且以自体方式再引入到人类中。因此，可以使用血浆外泌体进行根据本发明的递送或施用。

[0704] 脂质体

[0705] 可以用脂质体进行根据本发明的递送或施用。脂质体是球形囊泡结构，其由围绕内部水性区室的单层或多层脂质双层以及相对不可渗透的外部亲脂性磷脂双层构成。脂质体作为药物递送载剂受到了相当的重视，因为它们是生物相容、无毒的，可以递送亲水性和亲脂性药物分子，保护它们的货物免于被血浆酶降解，并且转运它们的负载跨过生物膜和血脑屏障(BBB)(对于评述，参见例如，Spuch和Navarro,Journal of Drug Delivery,第

2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0706] 可以由几种不同类型的脂质制造脂质体;然而,磷脂最常用来产生作为药物载剂的脂质体。虽然当脂质膜与水性溶液混合时脂质体形成是自发的,但是也可以通过使用均质机、超声破碎器或挤出设备通过以振荡的形式施加力使其加速(对于评述,参见例如,Spuch和Navarro,Journal of Drug Delivery,第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0707] 可以将几种其他添加剂添加到脂质体中以修改其结构和特性。例如,可以将胆固醇或鞘磷脂添加到脂质体混合物中,以便帮助稳定化脂质体结构并且防止脂质体内部货物的泄漏。此外,脂质体由氢化卵磷脂酰胆碱或卵磷脂酰胆碱、胆固醇和磷酸二鲸蜡脂制备,并且脂质体的平均囊泡尺寸被调整到约50nm和100nm。(对于评述,参见例如,Spuch和Navarro,Journal of Drug Delivery,第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0708] 脂质体配制品可以主要由天然磷脂和脂质诸如1,2-二硬脂酰基-sn-甘油基-3-磷脂酰胆碱(DSPC)、鞘磷脂、卵磷脂酰胆碱和单唾酸酰神经节苷酯构成。由于这种配制品仅仅由磷脂组成,脂质体配制品已经遇到了许多挑战,其中之一是在血浆中的不稳定性。已经作出战胜这些挑战的若干尝试,特别是在脂质膜的处理方面。这些尝试之一集中于胆固醇的处理。将胆固醇添加到常规配制品中减缓了封装的生物活性化合物到血浆中的迅速释放,或者添加1,2-二油酰基-sn-甘油基-3-磷酸乙醇胺(DOPE)增加了稳定性(对于评述,参见例如,Spuch和Navarro,Journal of Drug Delivery,第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0709] 在一个特别有利的实施方案中,特洛伊木马(Trojan Horse)脂质体(也称为分子特洛伊木马)是令人希望的并且方案可见于http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long。这些粒子允许转基因在血管内注射之后递送至整个脑。在不受限制的情况下,据信表面缀合有特异性抗体的中性脂质粒子允许经由胞吞作用跨过血脑屏障。特洛伊木马脂质体可以用于将核酸酶的CRISPR家族经由血管内注射递送至脑,这将允许全脑转基因动物,而不需要胚胎操纵。对于在脂质体中的体内施用,可以考虑约1-5g的DNA或RNA。

[0710] 在另一个实施方案中,AD官能化的CRISPR Cas系统或其组分可以在脂质体中施用,所述脂质体诸如稳定的核酸-脂质粒子(SNALP)(参见例如Morrissey等人,Nature Biotechnology,第23卷,第8期,2005年8月)。考虑每日静脉内注射约1、3或5mg/kg/天的SNALP中的被靶向的特异性CRISPR Cas。日治疗可以经过约三天,并且然后每周治疗持续约五周。在另一个实施方案中,还考虑了通过以约1或2.5mg/kg的剂量静脉内注射施用封装有特异性CRISPR Cas的SNALP)(参见例如Zimmerman等人,Nature Letters,第441卷,2006年5月4日)。SNALP配制品可以含有为2:40:10:48的摩尔百分比的脂质3-N-[(w甲氧基聚(乙二醇)2000)氨甲酰基]-1,2-二肉豆蔻氧基-丙胺(PEG-C-DMA)、1,2-二亚油基氧基-N,N-二甲基-3-氨基丙烷(DLinDMA)、1,2-二硬脂酰基-sn-甘油基-3-磷酸胆碱(DSPC)和胆固醇(参见例如Zimmerman等人,Nature Letters,第441卷,2006年5月4日)。

[0711] 在另一个实施方案中,已经证明稳定的核酸-脂质粒子(SNALP)将分子有效地递送至高度血管化的HepG2-衍生的肝脏肿瘤,但是不递送至血管化不良的HCT-116衍生的肝脏

肿瘤(参见例如Li,Gene Therapy(2012)19,775-780)。可以通过以下方式制备SNALP脂质体:使用25:1的脂质/siRNA比率和48/40/10/2的胆固醇/D-Lin-DMA/DSPC/PEG-C-DMA的摩尔比,用二硬脂酰磷脂酰胆碱(DSPC)、胆固醇和siRNA配制D-Lin-DMA和PEG-C-DMA。所得的SNALP脂质体的尺寸为约80-100nm。

[0712] 在又一个实施方案中,SNALP可以包含合成胆固醇(Sigma-Aldrich,St Louis,MO,USA)、二棕榈酰磷脂酰胆碱(Avanti Polar Lipids,Alabaster,AL,USA)、3-N-[(w-甲氧基聚(乙二醇)2000)氨甲酰基]-1,2-二肉豆蔻氧基丙胺,以及阳离子的1,2-二亚油基氧基-3-N,N二甲基氨基丙烷(参见例如Geisbert等人,Lancet 2010;375:1896-905)。可以考虑例如静脉内推注施用约2mg/kg总CRISPR Cas/剂的剂量。

[0713] 在又一个实施方案中,SNALP可以包含合成胆固醇(Sigma-Aldrich)、1,2-二硬脂酰基-sn-甘油基-3-磷酸胆碱(DSPC;Avanti Polar Lipids Inc.)、PEG-cDMA,以及1,2-二亚油基氧基-3-(N;N-二甲基)氨基丙烷(DLinDMA)(参见例如Judge,J.Clin.Invest.119:661-673(2009))。用于体内研究的配制品可以包含约9:1的最终脂质/RNA质量比。

[0714] Alnylam Pharmaceuticals的Barros和Gollob已对RNAi纳米药物的安全性进行了评论(参见例如Advanced Drug Delivery Reviews 64(2012)1730-1737)。稳定的核酸脂质粒子(SNALP)由四种不同的脂质构成-在低pH下为阳离子的可电离脂质(DLinDMA)、中性辅助脂质、胆固醇以及可扩散的聚乙二醇(PEG)-脂质。该粒子的直径为大约80nm并且在生理pH下是电中性的。在配制期间,可电离脂质用于在粒子形成期间使脂质与阴离子RNA缩合。当在渐增的酸性内体条件下带正电荷时,可电离脂质还介导SNALP与内体膜的融合,从而能够将RNA释放到细胞质中。PEG-脂质在配制期间稳定化粒子并且减少聚集,并且随后提供改善药代动力学特性的中性的亲水性外部。

[0715] 到目前为止,已经使用具有RNA的SNALP配制品开始了两个临床项目。Tekmira Pharmaceuticals最近在具有升高的LDL胆固醇的成年志愿者中完成了SNALP-ApoB I期单剂量研究。ApoB主要是在肝脏和空肠中表达,并且是为VLDL和LDL的组装和分泌所必需的。十七位受试者接受了SNALP-ApoB的单剂量(跨7个剂量水平的剂量递增)。没有肝脏毒性(预期为基于临床前研究的潜在剂量限制性毒性)的证据。处于最高剂量的(两位中的)一位受试者经历了与免疫系统刺激一致的流感样症状,于是做出结束该试验的决定。

[0716] Alnylam Pharmaceuticals已经类似地推出了ALN-TTR01,其采用以上所述的SNALP技术并且靶向突变体和野生型TTR的肝细胞产生,从而治疗TTR淀粉样变性(ATTR)。已经描述了三种ATTR综合征:家族性淀粉样变性多神经病(FAP)和家族性淀粉样心肌病(FAC)-两者均由TTR中的常染色体显性突变引起;以及由野生型TTR引起的老年全身性淀粉样变性(SSA)。最近在患有ATTR的患者中完成了ALN-TTR01的安慰剂对照单剂量递增I期试验。向31位患者(23位用研究药物,8位用安慰剂)在0.01至1.0mg/kg(基于siRNA)的剂量范围内以15分钟静脉内输注施用ALN-TTR01。治疗耐受性良好,其中在肝功能试验中没有显著增加。在≥0.4mg/kg时在23位患者的3位中注意到输注相关反应;所有患者均对减慢输注速率做出了反应并且所有患者继续参与研究。在处于1mg/kg的最高剂量(如根据临床前和NHP研究预期的)的两位患者中注意到血清细胞因子IL-6、IP-10和IL-1ra的最小与瞬时升高。在1mg/kg时观察到ALN-TTR01的预期药效动力学效应,即血清TTR的降低。

[0717] 在又一个实施方案中,可以通过将阳离子脂质、DSPC、胆固醇以及PEG-脂质例如以

40:10:40:10的摩尔比分别溶解在例如乙醇中来制备SNALP(参见Semple等人,Nature Niotechnology,第28卷,第2期,2010年2月,第172-177页)。将脂质混合物添加到水性缓冲液(50mM柠檬酸盐,pH 4)中,混合至最终的乙醇和脂质浓度分别为30%(体积/体积)和6.1mg/ml,并且使得其在22℃下平衡2分钟,然后挤出。使用Lipex挤出仪(Northern Lipids),在22℃下将水合脂质挤出通过两个重叠的80nm孔径大小的过滤器(Nuclepore),直到获得如通过动态光散射分析测定的70-90nm直径的囊泡为止。这大致需要1-3次通过。将siRNA(溶解在50mM柠檬酸盐中,pH为4的含有30%乙醇的水性溶液)以约5ml/min的速率在混合下添加到预平衡的(35℃)囊泡中。在达到0.06(重量/重量)的最终靶siRNA/脂质比率之后,将混合物在35℃下另外孵育30分钟,以允许囊泡重组和siRNA的封装。然后去除乙醇并且通过透析或切向流渗滤用PBS(155mM NaCl,3mM Na2HPO4,1mM KH2PO4,pH 7.5)替换外部缓冲液。使用受控的逐步稀释法工艺将siRNA封装在SNALP中。KC2-SNALP的脂质构分为以57.1:7.1:34.3:1.4的摩尔比使用的DLin-KC2-DMA(阳离子脂质)、二棕榈酰磷脂酰胆碱(DPPC;Avanti Polar Lipids)、合成胆固醇(Sigma)和PEG-C-DMA。在形成加载的粒子后,将SNALP在PBS中透析并且在使用之前通过0.2μm的过滤器灭菌过滤。平均粒径为75-85nm,并且90%-95%的siRNA被封装在脂质粒子之内。用于体内测试的在配制品中的最终siRNA/脂质比率为约0.15(重量/重量)。在临使用之前将含有因子VII siRNA的LNP-siRNA系统在无菌PBS中稀释到适当浓度,并且通过侧尾静脉以10ml/kg的总体积静脉内施用配制品。这种方法和这些递送系统可以外推到本发明的AD官能化的CRISPR Cas系统。

[0718] 其他脂质

[0719] 其他阳离子脂质,诸如氨基脂质2,2-二亚油基-4-二甲基氨基乙基-[1,3]-二氧戊环(DLin-KC2-DMA)可以类似于SiRNA地用来封装CRISPR Cas或其组分或对其编码的一个或多个核酸分子(参见例如Jayaraman,Angew.Chem.Int.Ed.2012,51,8529-8533),因此可以用于本发明的实践中。可以考虑具有下列脂质组成的预成型囊泡:分别处于摩尔比40/10/40/10的氨基脂质、二硬脂酰磷脂酰胆碱(DSPC)、胆固醇和(R)-2,3-双(十八烷氧基)丙基-1-(甲氧基聚(乙二醇)2000)丙基碳酸酯(PEG-脂质),以及大约0.05(w/w)的FVII siRNA/总脂质比率。为了确保在70-90nm范围内的窄粒径分布以及0.11+0.04(n=56)的低多分散性指数,可以在添加指导RNA之前将粒子通过80nm的膜挤出达三次。可以使用含有高度有效的氨基脂质16的粒子,其中四种脂质组分16、DSPC、胆固醇和PEG-脂质的摩尔比(50/10/38.5/1.5)可以被进一步优化,以增强体内活性。

[0720] Michael S D Kormann等人("Expression of therapeutic proteins after delivery of chemically modified mRNA in mice:Nature Biotechnology,第29卷,第154-157页(2011))描述了脂质包膜用于递送RNA的用途。在本发明中,脂质包膜的使用也是优选的。

[0721] 在另一个实施方案中,脂质可以用本发明的AD官能化的CRISPR Cas系统或其一种或多种组分或对其编码的一个或多个核酸分子一起配制而形成脂质纳米粒子(LNP)。脂质包括但不限于,DLin-KC2-DMA4、C12-200和辅助脂质二硬脂酰磷脂酰胆碱、胆固醇和PEG-DMG,可以使用自发囊泡形成程序将其与CRISPR Cas而不是siRNA一起配制(参见例如Novobrantseva,Molecular Therapy-Nucleic Acids(2012)1,e4;doi:10.1038/mtna.2011.3)。组分摩尔比可以是约50/10/38.5/1.5(DLin-KC2-DMA或C12-200/二硬脂酰

磷脂酰胆碱/胆固醇/PEG-DMG)。在DLin-KC2-DMA和C12-200脂质纳米粒子(LNP)的情况下，最终脂质:siRNA的重量比分别为约12:1和9:1。配制品可以具有约80nm的平均粒子直径，具有>90％的包封效率。可以考虑3mg/kg的剂量。

[0722] Tekmira在美国和国外具有一组针对LNP和LNP配制品的不同方面的大约95个同族专利(参见例如美国专利号7,982,027;7,799,565;8,058,069;8,283,333;7,901,708;7,745,651;7,803,397;8,101,741;8,188,263;7,915,399;8,236,943和7,838,658，以及欧洲专利号1766035;1519714;1781593和1664316)，所有这些专利均可用于并且/或者适于本发明。

[0723] AD官能化的CRISPR Cas系统或其组分或对其编码的一个或多个核酸分子可以封装在PLGA微球中进行递送，诸如进一步描述于美国公布申请20130252281和20130245107以及20130244279(转让给Moderna Therapeutics)中，这些申请涉及包含修饰的核酸分子的组合物的配制品的多个方面，这些核酸分子可以编码蛋白质、蛋白质前体、或该蛋白质或该蛋白质前体的部分或完全加工形式。该配制品可以具有50:10:38.5:1.5-3.0(阳离子脂质:融合脂质:胆固醇:PEG脂质)的摩尔比。PEG脂质可以选自但不限于PEG-c-DOMG、PEG-DMG。融合脂质可以是DSPC。另参见Schrum等人，Delivery and Formulation of Engineered Nucleic Acids，美国公布申请20120251618。

[0724] Nanomerics的技术着手解决针对广泛治疗学的生物利用度挑战，包括基于低分子量疏水性药物、肽以及核酸的治疗剂(质粒、siRNA、miRNA)。该技术已经证明了明显优势的特异性的施用途径包括经口途径、跨血脑屏障的转运、向实体瘤的递送，以及向眼部的递送。参见例如Mazza等人，2013，ACS Nano.2013年2月26日;7(2):1016-26;Uchegbu和Siew，2013，J Pharm Sci.102(2):305-10以及Lalatsa等人，2012，J Control Release.2012年7月20日;161(2):523-36。

[0725] 美国专利公布号20050019923描述了用于向哺乳动物身体递送生物活性分子诸如多核苷酸分子、肽和多肽和/或药剂的阳离子树状聚合物。树状聚合物适用于将生物活性分子的递送靶向至例如肝脏、脾、肺、肾脏或心脏(或甚至脑)。树状聚合物是由简单的支化单体单元以逐步方式制备的合成性3维大分子，其性质和功能性可以容易地进行控制和改变。树状聚合物经由向多功能核(发散式合成法)或朝向多功能核(收敛式合成法)重复加成结构单元来合成，并且结构单元的3维壳的每次加成使得更高级别的树状聚合物形成。聚丙烯亚胺树状聚合物从二氨基丁烷核开始，通过对伯胺的丙烯腈的双迈克尔加成反应向其上添加两倍数目的氨基基团，然后进行腈的氢化。这导致氨基基团的加倍。聚丙烯亚胺树状聚合物含有100％的可质子化氮以及高达64个末端氨基基团(5级，DAB 64)。可质子化基团通常是能够在中性pH下接受质子的胺基。树状聚合物作为基因递送剂的用途在很大程度上集中于聚酰胺-胺和含磷化合物的用途，其中胺/酰胺的混合物或N--P(O2)S分别作为缀合单元，没有报道关于更低级别的聚丙烯亚胺树状聚合物用于基因递送的用途的著作。还研究了作为pH敏感的控制释放系统的聚丙烯亚胺树状聚合物，其用于药物递送以及当被外周氨基酸基团化学修饰时用于它们的客体分子的封装。还研究了聚丙烯亚胺树状聚合物的细胞毒性和其与DNA的相互作用以及DAB 64的转染效力。

[0726] 美国专利公布号20050019923是基于与早期报道相反的观察:阳离子树状聚合物诸如聚丙烯亚胺树状聚合物展示出合适的特性，诸如特异性靶向和低毒性，其用于靶向递

送生物活性分子诸如遗传物质。此外，阳离子树状聚合物的衍生物也展示出适用于生物活性分子的靶向递送的特性。另参见，生物活性聚合物(Bioactive Polymers)，美国公布申请20080267903，其公开了“不同的聚合物，包括阳离子聚胺聚合物和树枝状聚合物显示出具有抗增殖活性，并且因此可用于治疗特征为不希望的细胞增殖的病症，诸如新生物和肿瘤、炎性病症(包括自身免疫性病症)、银屑病和动脉粥样硬化。这些聚合物可以作为活性剂单独使用，或者作为其他治疗剂(诸如药物分子或用于基因疗法的核酸)的递送媒介物。在此类情况下，聚合物的自身固有的抗肿瘤活性可以补足有待递送的剂的活性。”这些专利公布的公开内容可以与本文的教导内容结合使用，以用于递送一种或多种AD官能化的CRISPR Cas系统或其一种或多种组分或对其编码的一个或多个核酸分子。

[0727] 超电荷蛋白

[0728] 超电荷蛋白是一类具有非常高的正或负的理论净电荷的工程化或天然存在的蛋白质并且可以用于递送一种或多种AD官能化的CRISPR Cas系统或其一种或多种组分或对其编码的一个或多个核酸分子。超负电荷蛋白和超正电荷蛋白两者都表现出显著的抵抗热诱导或化学诱导的聚集的能力。超正电荷蛋白还能够穿透哺乳动物细胞。使货物诸如质粒DNA、RNA或其他蛋白质与这些蛋白质缔合可以使得这些大分子到体外和体内的哺乳动物细胞中的功能递送成为可能。在2007年报道了超电荷蛋白的产生和表征(Lawrence等人2007，Journal of the American Chemical Society 129,10110-10112)。

[0729] RNA和质粒DNA到哺乳动物细胞中的非病毒递送对于研究和治疗应用都是有价值的(Akinc等人，2010，Nat.Biotech.26,561-569)。纯化的+36GFP蛋白(或其他超正电荷蛋白)与RNA在适当的无血清培养基中混合并且使得其在添加到细胞中之前复合。在这个阶段包含血清抑制超电荷蛋白-RNA复合物的形成并且降低治疗效果。已经发现以下方案对于多种细胞系是有效的(McNaughton等人，2009，Proc.Natl.Acad.Sci.USA 106,6111-6116)(然而，应当进行改变蛋白质和RNA剂量的预试验来优化用于特定细胞系的程序)：(1)在治疗前一天，以1x105个细胞/孔铺在48孔板中。(2)在治疗当天，将纯化的+36GFP蛋白在无血清的培养基中稀释至终浓度200nM。添加RNA至50nM的终浓度。涡旋混合并且在室温下孵育10分钟。(3)在孵育期间，从细胞抽出培养基并且用PBS洗涤一次。(4)在孵育+36GFP和RNA之后，向细胞添加蛋白质-RNA复合物。(5)将细胞与复合物在37℃下孵育4小时。(6)在孵育之后，抽出培养基并且用20U/mL的肝素PBS洗涤三次。用含血清的培养基另外孵育细胞48小时或更长，这取决于用于活性的测定。(7)通过免疫印迹、qPCR、表型测定或其他适当的方法分析细胞。

[0730] 进一步发现，+36GFP在一系列细胞中是有效的质粒递送试剂。由于质粒DNA是一种比siRNA大的货物，有效复合质粒需要成比例地更大的+36GFP蛋白。为了有效质粒递送，申请人已经开发了一种带有C端HA2肽标签的+36GFP变体，这种肽是一种已知的来源于流感病毒血凝素蛋白的内体破坏肽。以下方案在多种细胞中是有效的，但是如上所述，建议针对特定细胞系和递送应用优化质粒DNA和超电荷蛋白的剂量：(1)在治疗前一天，以1x105/孔铺在48孔板中。(2)在治疗当天，将纯化的þ36GFP蛋白在无血清的培养基中稀释至终浓度2mM。添加1mg质粒DNA。涡旋混合并且在室温下孵育10分钟。(3)在孵育期间，从细胞抽出培养基并且用PBS洗涤一次。(4)在孵育þ36GFP和质粒DNA之后，向细胞轻轻添加蛋白质-DNA复合物。(5)将细胞与复合物在37C下孵育4小时。(6)在孵育之后，抽出培养基并且用PBS洗

涤。在含血清培养基中孵育细胞，并且另外孵育24-48小时。(7)在适当时分析质粒递送(例如，通过质粒驱动的基因表达)。

[0731] 另参见，例如，McNaughton等人，Proc.Natl.Acad.Sci.USA 106,6111-6116(2009)；Cronican等人，ACS Chemical Biology 5,747-752(2010)；Cronican等人，Chemistry&Biology 18,833-838(2011)；Thompson等人，Methods in Enzymology 503,293-319(2012)；Thompson,D.B.，等人，Chemistry&Biology 19(7),831-843(2012)。超电荷蛋白的这些方法可以用于并且/或者适于本发明的AD官能化的CRISPR Cas系统的递送。这些系统结合本文的教导内容可以用于递送一种或多种AD官能化的CRISPR Cas系统或其一种或多种组分或对其编码的一个或多个核酸分子。

[0732] 细胞穿透肽(CPP)

[0733] 在又一个实施方案中，考虑了细胞穿透肽(CPP)用于AD官能化的CRISPR Cas系统的递送。CPP是促进各种分子货物(从纳米级粒子至小化学分子和大的DNA片段)的细胞摄取的短肽。如本文所用，术语“货物”包括但不限于由以下组成的组：治疗剂、诊断性探针、肽、核酸、反义寡核苷酸、质粒、蛋白质、粒子(包括纳米粒子)、脂质体、发色团、小分子以及放射性物质。在本发明的方面中，货物还可以包括AD官能化的CRISPR Cas系统的任何组分或整个AD官能化的CRISPR Cas系统。本发明的各方面还提供了用于将所需货物递送至受试者中的方法，所述方法包括：(a)制备包含本发明的细胞穿透肽和所需货物的复合物，以及(b)向受试者经口地、关节内地、腹膜内地、鞘内地、动脉内地(intrarterially)、鼻内地、实质内地(intraparenchymally)、皮下地、肌内地、静脉内地、真皮地、直肠内地或局部地施用复合物。货物通过经由共价键的化学键联或通过非共价的相互作用与肽缔合。

[0734] CPP的功能是将货物递送至细胞中，这是一种通常通过胞吞作用发生的过程，其中货物被递送至活哺乳动物细胞的内体。细胞穿透肽具有不同的尺寸、氨基酸序列并且带电荷，但是所有CPP具有一种独特的特征，该特征是易位质膜并且将各种分子货物递送至细胞质或细胞器的能力。CPP易位可以被分类成三种主要的进入机制：直接穿透膜中、胞吞作用介导的进入，以及通过瞬时结构的形成的易位。CPP在医学中发现了许多应用，在治疗不同疾病包括癌症中作为药物递送剂，和病毒抑制剂以及用于细胞标记的造影剂。后者的实例包括充当用于GFP、MRI造影剂或量子点的载剂。CPP作为用于研究和医学的体外及体内递送载体具有极大潜力。CPP典型地具有下述的氨基酸组成，该氨基酸组成含有高相对丰度的带正电荷的氨基酸诸如赖氨酸或精氨酸或具有含有极性/带电荷氨基酸和非极性、疏水性氨基酸的交替图案的序列。这两种类型的结构分别称为聚阳离子的或两亲性的。CPP的第三种类别是仅含有非极性残基的疏水性肽，具有低净电荷或具有对于细胞摄取关键的疏水性氨基酸基团。所发现的初始CPP中之一是来自人类免疫缺陷病毒1(HIV-1)的反激活转录激活因子(Tat)，发现其高效地从周围介质被许多培养中的细胞类型摄取。从此以后，多种已知的CPP得到了相当地扩展并且产生了具有更有效的效应蛋白转导特性的小分子合成类似物。CPP包括但不限于穿透素、Tat(48-60)、转运素和(R-AhX-R4)(Ahx＝氨基己酰基)。

[0735] 美国专利8,372,951提供了来源于嗜酸性粒细胞阳离子蛋白(ECP)的CPP，该CPP表现出非常高的细胞穿透效率和低毒性。还提供了将带有其货物的CPP递送至脊椎动物受试者中的方面。CPP及其递送的另外方面描述于美国专利8,575,305；8;614,194和8,044,019中。CPP可用于递送AD官能化的CRISPR-Cas系统或其组分。可以用于递送AD官能化的

CRISPR-Cas系统或其组分的CPP还被提供于手稿“通过细胞穿透肽介导的Cas9蛋白和指导RNA的递送进行的基因破坏(Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA)”,Suresh Ramakrishna、Abu-Bonsrah Kwaku Dad、Jagadish Beloor等人Genome Res.2014年4月2日,该文献以引用方式并入,其中证实了用CPP缀合的重组Cas9蛋白和CPP复合的指导RNA的处理导致人类细胞系中的内源性基因破坏。在论文中,Cas9蛋白经由硫醚键缀合至CPP,而指导RNA与CPP复合,形成了稠合的带正电荷的粒子。已经显示,用修饰的Cas9和指导RNA同时和顺序地治疗人类细胞,包括胚胎干细胞、真皮成纤维细胞、HEK293T细胞、HeLa细胞和胚胎癌细胞,导致有效的基因破坏,伴随相对于质粒转染而言降低的脱靶突变。

[0736] 气雾剂递送

[0737] 对肺病进行治疗的受试者可以在自主呼吸时例如每侧肺部接受药物有效量的支气管递送的雾化的AAV载体系统。因此,一般而言,对于AAV递送,雾化的递送是优选的。腺病毒或AAV粒子可以用于递送。可以将合适的基因构建体克隆到递送载体中,这些基因构建体各自可操作地连接至一个或多个调控序列。

[0738] 包装和启动子

[0739] 用于驱动CRISPR-Cas蛋白和腺苷脱氨酶编码核酸分子表达的启动子可以包括AAV ITR,其可以用作启动子。这对于消除对另外的启动子元件(可能在载体中占用空间)的需要是有利的。空出来的另外的空间可以用于驱动另外的元件(gRNA等)的表达。另外,ITR活性是相对较弱的,因此可以用于降低由于Cas13的过表达所致的潜在毒性。

[0740] 对于普遍表达,可以使用的启动子包括:CMV、CAG、CBh、PGK、SV40、铁蛋白重链或轻链等。对于脑或其他CNS表达,可以使用启动子:用于所有神经元的突触蛋白I(SynapsinI)、用于兴奋性神经元的CaMKIIα、用于GABA能神经元的GAD67或GAD65或VGAT。对于肝脏表达,可以使用白蛋白启动子。对于肺表达,可以使用SP-B。对于内皮细胞,可以使用ICAM。对于造血细胞,可以使用IFNβ或CD45。对于成骨细胞,可以使用OG-2。

[0741] 用于驱动指导RNA的启动子可以包括Pol III启动子,诸如U6或H1,以及使用Pol II启动子和内含子盒来表达指导RNA。

[0742] 腺相关病毒(AAV)

[0743] CRISPR-Cas蛋白、腺苷脱氨酶和一种或多种指导RNA可以使用腺相关病毒(AAV)、慢病毒、腺病毒或其他质粒或病毒载体类型进行递送,特别地,使用来自以下文献的配方和剂量:例如,美国专利号8,454,972(针对腺病毒的配方、剂量)、8,404,658(针对AAV的配方、剂量)和5,846,946(针对DNA质粒的配方、剂量)以及来自临床试验和关于涉及慢病毒、AAV和腺病毒的临床试验的出版物的配方和剂量。例如,对于AAV,施用途径、配方和剂量可以如美国专利号8,454,972并且如涉及AAV的临床试验。对于腺病毒,施用途径、配方和剂量可以如美国专利号8,404,658并且如涉及腺病毒的临床试验。对于质粒递送,施用途径、配方和剂量可以如美国专利号5,846,946并且如涉及质粒的临床试验。剂量可以基于或外推为平均70kg的个体(例如,男性成人),并且可以针对不同重量和种类的患者、受试者、哺乳动物进行调整。施用频率在医学或兽医学从业者(例如医师、兽医师)的范围之内,其取决于常规因素,包括患者或受试者的年龄、性别、一般健康状况、其他状况以及着手解决的特定病状或症状。可以将病毒载体注射到目标组织中。对于细胞类型特异性基因组修饰,Cas13和腺

苷脱氨酶的表达可以由细胞类型特异性启动子驱动。例如，肝脏特异性表达可以使用白蛋白启动子，而神经元特异性表达(例如靶向CNS病症)可以使用突触蛋白I启动子。

[0744] 就体内递送而言，AAV相比于其他病毒载体是有利的，这是由于两个原因：低毒性(这可能是由于纯化方法不需要细胞粒子的超速离心所致，而超速离心可能激活免疫反应)；以及引起插入诱变的低概率，原因在于它未整合到宿主基因组中。

[0745] AAV的包装限度为4.5或4.75Kb。这意味着Cas13以及启动子和转录终止子必须都适于相同的病毒载体。大于4.5或4.75Kb的构建体将导致病毒产生显著减少。SpCas9是相当大的，该基因自身超过4.1Kb，使其难于包装到AAV中。因此本发明的实施方案包括利用更短的Cas13同系物。

[0746] 关于AAV，AAV可以是AAV1、AAV2、AAV5或它们的任何组合。可以相对于有待被靶向的细胞从这些AAV中选择AAV；例如，可以选择用于靶向脑或神经元细胞的AAV血清型1、2、5或杂合衣壳AAV1、AAV2、AAV5或它们的任何组合；并且可以选择用于靶向心脏组织的AAV4。AAV8可用于递送至肝脏。本文的启动子和载体是单独优选的。就这些细胞而论(参见Grimm, D.等人，J.Virol.82:5887-5911(2008))，某些AAV血清型的列表如下：

[0747]

细胞系	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
HeLa	3	100	2.0	0.1	6.7	1	0.2	0.1
HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
单核细胞	1111	100	ND	ND	125	1429	ND	ND
未成熟 DC	2500	100	ND	ND	222	2857	ND	ND
成熟 DC	2222	100	ND	ND	333	3333	ND	ND

[0748] 慢病毒

[0749] 病毒是复杂的反转录病毒，其具有在有丝分裂细胞和有丝分裂后细胞两者中感染并表达其基因的能力。最为人熟知的慢病毒是人类免疫缺陷病毒(HIV)，其使用其他病毒的包膜糖蛋白来靶向广泛范围的细胞类型。

[0750] 慢病毒可以如下制备。在克隆pCasES10(含有慢病毒转移质粒骨架)之后，将处于低传代数(p＝5)的HEK293FT接种在T-75烧瓶中，以在转染之前的一天在具有10％胎牛血清而没有抗生素的DMEM中达到50％汇合。在20小时之后，将培养基更换为OptiMEM(无血清)培养基，并且在4小时后进行转染。将细胞用10μg的慢病毒转移质粒(pCasES10)和下列包装质粒转染：5μg的pMD2.G(VSV-g假型)和7.5ug的psPAX2(gag/pol/rev/tat)。在具有阳离子脂质递送剂(50uL的Lipofectamine 2000和100ul的Plus试剂)的4mL OptiMEM中进行转染。在

6小时之后,将培养基更换为具有10％胎牛血清的无抗生素的DMEM。这些方法在细胞培养期间使用血清,但是优选无血清的方法。

[0751] 慢病毒可以如下纯化。在48小时后收获病毒上清液。首先清除上清液中的碎片,然后通过0.45um低蛋白结合(PVDF)过滤器进行过滤。然后将它们在超速离心机中以24,000rpm旋转2小时。将病毒沉淀重新悬浮在50ul的DMEM中,在4C下过夜。然后将它们等分,并且立即在-80℃下冷冻。

[0752] 在另一个实施方案中,还考虑了基于马感染性贫血病毒(EIAV)的最小非灵长类动物慢病毒载体,特别是对于眼部基因疗法而言(参见例如Balagaan,J Gene Med 2006;8:275-285)。在另一个实施方案中,还考虑了一种经由视网膜下注射递送用于治疗湿型年龄相关性黄斑变性的、表达血管生成抑制蛋白(内皮抑素和血管抑素)的基于马感染性贫血病毒的慢病毒基因疗法载体RetinoStat®(参见例如Binley等人,HUMAN GENE THERAPY 23:980-991(2012年9月)),并且此载体可以被修改用于本发明的AD官能化的CRISPR-Cas系统。

[0753] 在另一个实施方案中,自我失活性慢病毒载体可以用于并且/或者适于本发明的AD官能化的CRISPR-Cas系统核,该自我失活性慢病毒载体具有靶向由HIV tat/rev共享的共有外显子的siRNA、核仁定位TAR诱饵和抗CCR5特异性锤头状核酶(参见例如DiGiusto等人(2010)Sci Transl Med 2:36ra43)。可以收集最少2.5×106个CD34+细胞/每千克患者体重并且以2×106个细胞/ml的密度在X-VIVO 15培养基(Lonza)中预刺激16至20小时,该培养基含有2μmol/L-谷氨酰胺、干细胞因子(100ng/ml)、Flt-3配体(Flt-3L)(100ng/ml)和促血小板生成素(10ng/ml)(CellGenix)。可以用慢病毒以感染复数5在75-cm2的包覆有纤连蛋白(25mg/cm2)(RetroNectin,Takara Bio Inc.)的组织培养瓶中转导预刺激的细胞,持续16至24小时。

[0754] 慢病毒载体已公开于帕金森病的治疗中,参见例如美国专利公布号20120295960以及美国专利号7303910和7351585。慢病毒载体还已公开于眼部疾病的治疗中,参见例如美国专利公布号20060281180、20090007284、US20110117189;US20090017543;US20070054961、US20100317109。还已公开了将慢病毒载体递送至脑,参见例如美国专利公布号US20110293571;US20110293571、US20040013648、US20070025970、US20090111106和美国专利号US7259015。

[0755] 在非动物生物体中的应用

[0756] 所述一种或多种系统(例如,单一或多重)可以与农作物基因组的研究进展结合来使用。本文所述的系统可以用于进行有效且性价比高的植物基因或基因组探询或编辑或操纵—例如,用于快速研究并且/或者选择并且/或者探询并且/或者比较并且/或者操纵并且/或者转化植物基因或基因组;例如,以便为一种或多种植物产生、鉴定、开发、优化或赋予一种或多种性状或一种或多种特征或者以转化植物基因组。因此,可以存在植物、具有新性状或特征组合的新植物或具有增强的性状的新植物的改善的产生方法。关于定点整合(SDI)或基因编辑(GE)或任何近反向育种(Near Reverse Breeding)(NRB)或反向育种(RB)技术中的植物,可以使用所述系统。利用本文所述的Cas13效应蛋白系统的方面可能类似于CRISPR-Cas(例如CRISPR-Cas9)系统在植物中的使用,并且提及亚利桑那大学(University of Arizona)网站“CRISPR-PLANT”(http://www.genome.arizona.edu/crispr/)(得到宾州州立大学(Penn State)和AGI的支持)。本发明的实施方案可以用于在植物中或在先前已使

用RNAi或类似基因组编辑技术的情况中进行基因组编辑;参见例如,Nekrasov,“Plant genome editing made easy:targeted mutagenesis in model and crop plants using the CRISPR-Cas system,”Plant Methods 2013,9:39(doi:10.1186/1746-4811-9-39);Brooks,“Efficient gene editing in tomato in the first generation using the CRISPR-Cas9 system,”Plant Physiology,2014年9月,第114.247577页;Shan,“Targeted genome modification of crop plants using a CRISPR-Cas system,”Nature Biotechnology 31,686-688(2013);Feng,“Efficient genome editing in plants using a CRISPR-Cas system,”Cell Research(2013)23:1229-1232.doi:10.1038/cr.2013.114;2013年8月20日在线公布;Xie,“RNA-guided genome editing in plants using a CRISPR-Cas system,”Mol Plant.2013年11月;6(6):1975-83.doi:10.1093/mp/sst119.电子版2013年8月17日;Xu,“Gene targeting using the Agrobacterium tumefaciens-mediated CRISPR-Cas system in rice,”Rice 2014,7:5(2014);Zhou等人,“Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate:CoA ligase specificity and Redundancy,”New Phytologist(2015)(论坛)1-4(仅在www.newphytologist.com处在线可得);Caliando等人,“Targeted DNA degradation using a CRISPR device stably carried in the host genome,NATURE COMMUNICATIONS 6:6989,DOI:10.1038/ncomms7989,www.nature.com/naturecommunications DOI:10.1038/ncomms7989;美国专利号6,603,061-土壤杆菌属介导的植物转化方法(Agrobacterium-Mediated Plant Transformation Method);美国专利号7,868,149-植物基因组序列及其用途(Plant Genome Sequences and Uses Thereof)以及US 2009/0100536-具有增强的农艺性状的转基因植物(Transgenic Plants with Enhanced Agronomic Traits),每份文献的所有内容和公开内容均以引用方式整体并入本文。在本发明的实践中,Morrell等人“农作物基因组:进展与应用(Crop genomics:advances and applications)”,Nat Rev Genet.2011年12月29日;13(2):85-96的内容和公开内容;每份文献以引用方式并入本文,包括关于本文的实施方案如何可以就植物而使用。因此,除非另外表明,否则本文提及动物细胞也可以加上必要的变更应用于植物细胞;并且,本文的具有降低的脱靶效应的酶和采用此类酶的系统可以用于植物应用,包括本文提及的那些。

[0757] 系统在植物和酵母中的应用

[0758] 一般来讲,术语“植物”涉及植物界中通过细胞分裂特征性生长、含有叶绿体并具有包含纤维素的细胞壁的任何不同光合作用生物体、真核生物体、单细胞生物体或多细胞生物体。术语植物涵盖单子叶植物和双子叶植物。确切地说,这些植物旨在包括但不限于被子植物和裸子植物,诸如刺槐、苜蓿、苋菜、苹果、杏、朝鲜蓟、白蜡树、芦笋、鳄梨、香蕉、大麦、豆类、甜菜、桦树、山毛榉、黑莓、蓝莓、西兰花、抱子甘蓝、卷心菜、油菜、哈密瓜、胡萝卜、木薯、花椰菜、雪松、谷类、芹菜、栗子、樱桃、大白菜、柑橘、克莱门氏小柑橘、三叶草、咖啡、玉米、棉花、豇豆、黄瓜、柏树、茄子、榆树、菊苣、桉树、茴香、无花果、冷杉、天竺葵、葡萄、葡萄柚、落花生、地樱桃、树胶铁杉、山核桃木、羽衣甘蓝、奇异果、甘蓝、落叶松、莴苣、韭、柠檬、青柠、洋槐、松树、孔雀草、玉米、芒果、枫树、甜瓜、粟、蘑菇、芥菜、坚果、橡树、燕麦、油棕、秋葵、洋葱、橙、观赏植物或花或树木、木瓜、棕榈、荷兰芹、欧洲防风草、豌豆、桃、花生、

梨、泥炭(peat)、胡椒、柿子、木豆、松树、菠萝、大蕉、李子、石榴、马铃薯、南瓜、菊苣、萝卜、油菜籽、覆盆子、稻、黑麦、高粱、红花、黄华柳、大豆、菠菜、云杉、南瓜属植物果实、草莓、糖甜菜、甘蔗、向日葵、甘薯、甜玉米、橘子、茶、烟草、番茄、树类、黑小麦、草坪草、芜菁、藤本植物、胡桃、豆瓣菜、西瓜、小麦、山药、紫杉以及西葫芦。术语植物还涵盖藻类,藻类主要是光能自养生物,它们主要一致缺少根、叶以及其他表征高等植物的器官。

[0759] 用于使用如本文所述的系统进行基因组编辑的方法可以用于对基本上任何植物赋予所需的性状。针对本文所述的所需生理以及农学特征,可以使用本公开的核酸构建体和以上提及的各种转化方法对多种多样的植物和植物细胞系统进行工程化。在优选实施方案中,用于工程化的靶植物和植物细胞包括但不限于那些单子叶植物和双子叶植物,诸如农作物(包括谷类作物(例如小麦、玉米、稻、粟、大麦)、果实作物(例如番茄、苹果、梨、草莓、橙)、饲料作物(例如苜蓿)、根用蔬菜作物(例如胡萝卜、马铃薯、甜菜、山药)、叶类蔬菜作物(例如莴苣、菠菜);开花植物(例如矮牵牛、玫瑰、菊花)、松柏植物以及松树(例如松杉、云杉);植物修复中所使用的植物(例如重金属累积植物);油料作物(例如向日葵、油菜种子)和用于实验目的植物(例如拟南芥属)。因此,这些方法和系统可以用于遍及广泛范围的植物,像例如与属于以下目的双子叶植物:木兰目(Magniolales)、八角目(Illiciales)、樟目(Laurales)、胡椒目(Piperales)、马兜铃目(Aristochiales)、睡莲目(Nymphaeales)、毛茛目(Ranunculales)、罂粟目(Papeverales)、瓶子草科(Sarraceniaceae)、昆栏树目(Trochodendrales)、金缕梅目(Hamamelidales)、杜仲目(Eucomiales)、塞子木目(Leitneriales)、杨梅目(Myricales)、壳斗目(Fagales)、木麻黄目(Casuarinales)、石竹目(Caryophyllales)、肉穗果目(Batales)、蓼目(Polygonales)、蓝雪目(Plumbaginales)、五桠果目(Dilleniales)、山茶目(Theales)、锦葵目(Malvales)、荨麻目(Urticales)、玉蕊目(Lecythidales)、堇菜目(Violales)、杨柳目(Salicales)、白花菜目(Capparales)、杜鹃花目(Ericales)、岩梅目(Diapensales)、柿树目(Ebenales)、报春花目(Primulales)、蔷薇目(Rosales)、豆目(Fabales)、川草目(Podostemales)、小二仙草目(Haloragales)、桃金娘目(Myrtales)、山茱萸目(Cornales)、山龙眼目(Proteales)、檀香目(San tales)、大花草目(Rafflesiales)、卫矛目(Celastrales)、大戟目(Euphorbiales)、鼠李目(Rhamnales)、无患子目(Sapindales)、胡桃目(Juglandales)、牻牛儿苗目(Geraniales)、远志目(Polygalales)、伞形目(Umbellales)、龙胆目(Gentianales)、花葱目(Polemoniales)、唇形目(Lamiales)、车前草目(Plantaginales)、玄参目(Scrophulariales)、桔梗目(Campanulales)、茜草目(Rubiales)、川续断目(Dipsacales)以及菊目(Asterales);这些方法和CRISPR-Cas系统可以用于单子叶植物,诸如属于以下目的单子叶植物:泽泻目(Alismatales)、水鳖目(Hydrocharitales)、茨藻目(Najadales)、霉草目(Triuridales)、鸭跖草目(Commelinales)、谷精草目(Eriocaulales)、帚灯草目(Restionales)、禾本目(Poales)、灯芯草目(Juncales)、莎草科(Cyperales)、香蒲目(Typhales)、凤梨目(Bromeliales)、姜目(Zingiberales)、槟榔目(Arecales)、环花目(Cyclanthales)、露兜树目(Pandanales)、天南星目(Arales)、百合目(Lilliales)以及兰目(Orchid ales),或者用于属于裸子植物(Gymnospermae)的植物,例如属于松杉目(Pinales)、银杏目(Ginkgoales)、苏铁目(Cycadales)、南洋杉目(Araucariales)、柏目(Cupressales)以及麻黄目(Gnetales)的植物。

[0760] 本文所述的系统和使用方法可以用于广泛范围的植物种类,这些植物种类包括在下面的双子叶植物、单子叶植物或裸子植物属的非限制性列表中:颠茄属(Atropa)、油丹属(Alseodaphne)、腰果属(Anacardium)、落花生属(Arachis)、琼楠属(Beilschmiedia)、芸苔属(Brassica)、红花属(Carthamus)、木防己属(Cocculus)、巴豆属(Croton)、甜瓜属(Cucumis)、柑橘属(Citrus)、西瓜属(Citrullus)、辣椒属(Capsicum)、长春花属(Catharanthus)、椰子属(Cocos)、咖啡属(Coffea)、南瓜属(Cucurbita)、胡萝卜属(Daucus)、杜氏木属(Duguetia)、花菱草属(Eschscholzia)、榕属(Ficus)、草莓属(Fragaria)、海罂粟属(Glaucium)、大豆属(Glycine)、棉属(Gossypium)、向日葵属(Helianthus)、橡胶树属(Hevea)、天仙子属(Hyoscyamus)、莴苣属(Lactuca)、卷枝藤属(Landolphia)、亚麻属(Linum)、木姜子属(Litsea)、番茄属(Lycopersicon)、羽扇豆属(Lupinus)、木薯属(Manihot)、马郁兰属(Majorana)、苹果属(Malus)、苜蓿属(Medicago)、烟草属(Nicotiana)、木犀榄属(Olea)、银胶菊属(Parthenium)、罂粟属(Papaver)、鳄梨属(Persea)、菜豆属(Phaseolus)、黄连木属(Pistacia)、豌豆属(Pisum)、梨属(Pyrus)、李属(Prunus)、萝卜属(Raphanus)、蓖麻属(Ricinus)、千里光属(Senecio)、防己属(Sinomenium)、千金藤属(Stephania)、欧白芥属(Sinapis)、茄属(Solanum)、可可属(Theobroma)、三叶草属(Trifolium)、胡芦巴属(Trigonella)、蚕豆属(Vicia)、蔓长春花属(Vinca)、葡萄属(Vilis)和豇豆属(Vigna);以及葱属(Allium)、须芒草属(Andropogon)、画眉草属(Aragrostis)、天门冬属(Asparagus)、燕麦属(Avena)、狗牙根属(Cynodon)、油棕属(Elaeis)、羊茅属(Festuca)、羊茅黑麦草属(Festulolium)、萱草属(Heterocallis)、大麦属(Hordeum)、浮萍属(Lemna)、毒麦属(Lolium)、芭蕉属(Musa)、稻属(Oryza)、黍属(Panicum)、狼尾草属(Pannesetum)、梯牧草属(Phleum)、早熟禾属(Poa)、黑麦属(Secale)、高粱属(Sorghum)、小麦属(Triticum)、玉蜀黍属(Zea)、冷杉属(Abies)、杉木属(Cunninghamia)、麻黄属(Ephedra)、云杉属(Picea)、松属(Pinus),以及黄杉属(Pseudotsuga)。

[0761] 所述系统以及使用方法还可以用于遍及广泛范围的“藻类”或“藻类细胞”;包括例如选自若干真核生物门的藻类,其包括红藻植物门(Rhodophyta)(红藻)、绿藻门(Chlorophyta)(绿藻)、褐藻门(Phaeophyta)(褐藻)、硅藻门(Bacillariophyta)(硅藻)、真眼点藻纲(Eustigmatophyta)和沟鞭藻类以及原核门蓝藻门(Cyanobacteria)(蓝绿藻)。术语“藻类”包括例如选自以下的藻类:双眉藻属(Amphora)、鱼腥藻属(Anabaena)、纤维藻属(Anikstrodesmis)、丛粒藻属(Botryococcus)、角毛藻属(Chaetoceros)、衣藻属(Chlamydomonas)、绿藻属(Chlorella)、绿球藻属(Chlorococcum)、小环藻属(Cyclotella)、筒柱藻属(Cylindrotheca)、杜氏藻属(Dunaliella)、球石藻属(Emiliana)、眼虫属(Euglena)、红球藻属(Hematococcus)、等鞭金藻属(Isochrysis)、单鞭金藻属(Monochrysis)、单针藻属(Monoraphidium)、微绿球藻属(Nannochloris)、拟微绿球藻属(Nannnochloropsis)、舟形藻属(Navicula)、肾鞭藻属(Nephrochloris)、肾爿藻属(Nephroselmis)、菱形藻属(Nitzschia)、节球藻属(Nodularia)、念珠藻属(Nostoc)、髓球藻属(Oochromonas)、卵囊藻(Oocystis)、颤藻属(Oscillartoria)、巴夫藻属(Pavlova)、褐指藻属(Phaeodactylum)、扁藻属(Playtmonas)、颗石藻属(Pleurochrysis)、紫菜属(Porhyra)、假鱼腥藻属(Pseudoanabaena)、塔胞藻属(Pyramimonas)、裂丝藻属

(Stichococcus)、聚球藻属(Synechococcus)、集胞藻属(Synechocystis)、四片藻属(Tetraselmis)、海链藻属(Thalassiosira)以及束毛藻属(Trichodesmium)。

[0762] 可以根据本发明的方法处理植物的一部分,即“植物组织”,以产生改良的植物。植物组织还涵盖植物细胞。如本文所用的术语“植物细胞”是指活体植物的个体单元,或者在完整全株中或者为在体外组织培养基中、在培养基或琼脂上、在生长培养基或缓冲液的悬浮液中或作为高等组织单元一部分生长的分离形式,例如像植物组织、植物器官或全株。

[0763] “原生质体”是指这样的植物细胞,其保护性细胞壁使用例如机械或酶手段被完全去除或部分去除从而形成活体植物的完整生物化学活性单元,该活性单元在适当生长条件下可以重新形成细胞壁、增殖并再生成全株。

[0764] 术语“转化”广泛地是指通过借助于土壤杆菌或多种化学或物理方法之一来引入DNA从而对植物宿主进行遗传修饰的过程。如本文所用,术语“植物宿主”是指植物,包括植物的任何细胞、组织、器官、或子代。许多合适的植物组织或植物细胞可以被转化,并且包括但不限于原生质体、体细胞胚胎、花粉、叶、幼苗、茎、愈伤组织、匍伏茎、试管块茎、以及胚芽。植物组织还是指这种植物、种子、子代、繁殖体(不管有性还是无性生殖的)、以及这些中任一项的后代(诸如插条或种子)的任何克隆。

[0765] 如本文所用,术语“转化的”是指已引入外源DNA分子诸如构建体的细胞、组织、器官或生物体。所引入的DNA分子可以整合到受体细胞、组织、器官或生物体的基因组DNA中,以使得引入的DNA分子被传递至后续子代。在这些实施方案中,“转化的”或“转基因的”细胞或植物还可以包括细胞或植物的子代以及通过育种程序采用这种转化的植物作为杂交的母体并表现出因引入的DNA分子的存在而产生的改变的表型的子代。优选地,转基因植物是能育的并且能够将引入的DNA通过有性繁殖传递给子代。

[0766] 术语“子代”诸如转基因植物的子代是由植物或转基因植物生出的、由其产生的或从其来源的子代。所引入的DNA分子还可以被瞬时引入到受体细胞中,使得所引入的DNA分子不被后续子代遗传并且因此不被认为是“转基因的”。因此,如本文所用,“非转基因”植物或植物细胞是不含有稳定整合到其基因组中的外源DNA的植物。

[0767] 如本文所用,术语“植物启动子”是能够启动植物细胞转录的启动子,无论其是否来源自植物细胞。示例性的合适植物启动子包括但不限于,那些获得自植物、植物病毒以及细菌(诸如包含在植物细胞中表达的基因的土壤杆菌属或根瘤菌属)的启动子。

[0768] 如本文所用,“真菌细胞”是指真菌界内的任何类型的真核细胞。真菌界内的门包括子囊菌门(Ascomycota)、担子菌门(Basidiomycota)、芽枝霉门(Blastocladiomycota)、壶菌门(Chytridiomycota)、球囊菌门(Glomeromycota)、微孢子虫目(Microsporidia)以及新美鞭菌门(Neocallimastigomycota)。真菌细胞可包括酵母、霉菌和丝状真菌。在一些实施方案中,真菌细胞是酵母细胞。

[0769] 如本文所用,术语“酵母细胞”是指子囊菌门和担子菌门内的任何真菌细胞。酵母细胞可以包括芽殖酵母细胞、裂殖酵母细胞和霉菌细胞。不限于这些生物体,在实验室和工业环境中使用的许多类型的酵母菌是子囊菌门的部分。在一些实施方案中,酵母细胞是酿酒酵母(S.cerervisiae)、马克思克鲁维酵母(Kluyveromyces marxianus)或东方伊萨酵母(Issatchenkia orientalis)细胞。其他酵母细胞可以包括但不限于假丝酵母属种(Candida spp.)(例如白色念珠菌(Candida albicans))、亚罗酵母属种(Yarrowia spp.)

(例如亚罗解脂酵母(Yarrowia lipolytica))、毕赤酵母属种(Pichia spp.)(例如巴斯德毕赤酵母(Pichia pastoris))、克鲁维酵母菌属种(Kluyveromyces spp.)(例如产乳糖酶酵母(Kluyveromyces lactis)和马克思克鲁维酵母)、脉孢菌属种(Neurospora spp.)(例如粗糙脉孢菌(Neurospora crassa))、镰刀菌属种(Fusarium spp.)(例如尖孢镰刀菌(Fusarium oxysporum)),以及伊萨酵母属种(Issatchenkia spp.)(例如东方伊萨酵母,又称为库德里阿兹威毕赤酵母(Pichia kudriavzevii)和Candida acidothermophilum)。在一些实施方案中,真菌细胞是丝状真菌细胞。如本文所用,术语"丝状真菌细胞"是指以丝状体(即菌丝或菌丝体)生长的任何类型的真菌细胞。丝状真菌细胞的实例可以包括但不限于曲霉属种(Aspergillus spp.)(例如黑曲霉(Aspergillus niger))、木霉属种(Trichoderma spp.)(例如里氏木霉(Trichoderma reesei))、根霉属种(Rhizopus spp.)(例如米根霉(Rhizopus oryzae))以及被孢霉属种(Mortierella spp.)(例如深黄被孢霉(Mortierella isabellina))。

[0770] 在一些实施方案中,真菌细胞是工业菌株。如本文所用,"工业菌株"是指工业工艺(例如,商业或工业规模的产品生产)中使用的或从中分离的任何真菌细胞菌株。工业菌株可以指代典型地在工业工艺中使用的真菌种类,或者它可以指代还可用于非工业目的(例如,实验室研究)的真菌种类的分离株。工业工艺的实例可以包括发酵(例如,在食品或饮料产品的生产中)、蒸馏、生物燃料生产、化合物生产以及多肽生产。工业菌株的实例可以包括但不限于JAY270和ATCC4124。

[0771] 在一些实施方案中,真菌细胞是多倍体细胞。如本文所用,"多倍体"细胞可以指代其基因组以多于一个拷贝存在的任何细胞。多倍体细胞可以指代天然地发现处于多倍体状态的细胞类型,或者它可以指代已被诱导以多倍体状态存在的细胞(例如,通过特异性调节、改变、失活、激活、或者减数分裂、胞质分裂或DNA复制的修饰)。多倍体细胞可以指代其整个基因组是多倍体的细胞,或者它可以指代在特定目标基因组基因座中为多倍体的细胞。不希望受理论的束缚,认为与在单倍体细胞中相比指导RNA的丰度在多倍体细胞的基因组工程化中可能更经常地是速率限制性组分,并且因此使用本文所述的系统的方法可以利用使用某一真菌细胞类型的优点。

[0772] 在一些实施方案中,真菌细胞是二倍体细胞。如本文所用,"二倍体"细胞可以指代其基因组以两个拷贝存在的任何细胞。二倍体细胞可以指代天然地发现处于二倍体状态的细胞类型,或者它可以指代已被诱导以二倍体状态存在的细胞(例如,通过特异性调节、改变、失活、激活、或者减数分裂、胞质分裂或DNA复制的修饰)。例如,酿酒酵母菌株S228C可以维持处于单倍体或二倍体状态。二倍体细胞可以指代其整个基因组是二倍体的细胞,或者它可以指代在特定目标基因组基因座中为二倍体的细胞。在一些实施方案中,真菌细胞是单倍体细胞。如本文所用,"单倍体"细胞可以指代其基因组以一个拷贝存在的任何细胞。单倍体细胞可以指代天然地发现处于单倍体状态的细胞类型,或者它可以指代已被诱导以单倍体状态存在的细胞(例如,通过特异性调节、改变、失活、激活、或者减数分裂、胞质分裂或DNA复制的修饰)。例如,酿酒酵母菌株S228C可以维持处于单倍体或二倍体状态。单倍体细胞可以指代其整个基因组是单倍体的细胞,或者它可以指代在特定目标基因组基因座中为单倍体的细胞。

[0773] 如本文所用,"酵母表达载体"是指含有编码RNA和/或多肽的一个或多个序列的核

酸并且还可以含有控制一个或多个核酸表达的任何所需元件,以及使得能够在酵母细胞内复制并维持表达载体的任何元件。许多合适的酵母表达载体及其特征在本领域中是已知的;例如,各种载体和技术示出于Yeast Protocols,第2版,Xiao,W.,编辑(Humana Press,New York,2007)以及Buckholz,R.G.和Gleeson,M.A.(1991)Biotechnology(NY)9(11):1067-72。酵母载体可以包含但不限于着丝粒(CEN)序列、自主复制序列(ARS)、可操作地连接至目标序列或基因的启动子(诸如RNA聚合酶III启动子)、终止子(诸如RNA聚合酶III终止子)、复制起点和标记基因(例如,营养缺陷型、抗生素型、或其他选择性标记物)。在酵母菌中使用的表达载体的实例可以包括质粒、酵母人工染色体、2μ质粒、酵母整合型质粒、酵母复制型质粒、穿梭载体以及附加型质粒。

[0774] 系统组分在植物和植物细胞基因组中的稳定整合

[0775] 在特定实施方案中,设想的是引入编码所述系统的组分的多核苷酸,以稳定整合到植物细胞的基因组中。在这些实施方案中,可以依据指导RNA和/或腺苷脱氨酶和Cas13的融合蛋白在何时、何处及在何条件下表达,对转化载体或表达系统的设计进行调整。

[0776] 在特定实施方案中,设想的是将所述系统的组分稳定地引入植物细胞的基因组DNA中。另外地或可替代地,设想的是引入所述系统的组分,以将其稳定整合到植物细胞器的DNA中,诸如但不限于质体、线粒体或叶绿体。

[0777] 用于稳定整合到植物细胞基因组中的表达系统可以含有以下元件中的一者或多者:可用于在植物细胞中表达指导RNA和/或腺苷脱氨酶和Cas13的融合蛋白的启动子元件;增强表达的5'非翻译区;进一步增强在某些细胞(诸如单子叶植物细胞)中表达的内含子元件;为插入编码指导RNA和/或腺苷脱氨酶和Cas13的融合蛋白的序列以及其他所需元件提供方便的限制性位点的多克隆位点;以及为所表达的转录物提供有效终止的3'非翻译区。

[0778] 表达系统的元件可以处于一个或多个表达构建体上,所述一个或多个表达构建体是环状的(诸如质粒或转化载体),或者是非环状的(诸如线性双链DNA)。

[0779] 在特定实施方案中,AD官能化的CRISPR表达系统至少包含:编码与植物中的靶序列杂交的指导RNA(gRNA)的核苷酸序列,并且其中所述指导RNA包含指导序列和同向重复序列;以及编码腺苷脱氨酶和Cas13的融合蛋白的核苷酸序列,其中组分(a)或(b)位于相同或不同构建体上,并且由此不同核苷酸序列可以处于植物细胞内可操作的相同或不同调控元件的控制下。

[0780] 可以通过多种常规技术将含有所述系统的组分的一个或多个DNA构建体和(在适用情况下)模板序列引入到植物、植物部分或植物细胞的基因组中。所述方法大体上包括以下步骤:选择合适的宿主细胞或宿主组织、将所述一个或多个构建体引入到宿主细胞或宿主组织中,以及由其再生植物细胞或植物。

[0781] 在特定实施方案中,可以使用诸如但不限于电穿孔、微注射、植物细胞原生质体的雾化束注入的技术将DNA构建体引入到植物细胞中,或者可以使用诸如DNA粒子轰击的基因枪法直接将这些DNA构建体引入到植物组织中(另参见Fu等人,Transgenic Res.2000年2月;9(1):11-9)。粒子轰击的基础是使包覆有目标一种或多种基因的粒子朝向细胞加速,从而导致粒子穿透原生质并且典型地稳定整合到基因组中。(参见例如Klein等人,Nature(1987),Klein et ah,Bio/Technology(1992),Casas et ah,Proc.Natl.Acad.Sci.USA(1993).)。

[0782] 在特定实施方案中,可以通过土壤杆菌介导的转化将含有所述系统的组分的DNA构建体引入到植物中。可以将DNA构建体与合适的T-DNA侧接区组合,并且将它们引入到常规根瘤土壤杆菌(Agrobacterium tumefaciens)宿主载体中。通过感染植物或通过用含有一种或多种Ti(肿瘤诱导)质粒的土壤杆菌属细菌培育植物原生质体,可以将外源DNA并入植物基因组中。(参见例如Fraley等人(1985)、Rogers等人(1987)和美国专利号5,563,055)。

[0783] 植物启动子

[0784] 为了确保在植物细胞中适当表达,典型地将本文所述的系统的组分置于植物启动子(即在植物细胞中可操作的启动子)的控制下。设想了使用不同类型的启动子。

[0785] 组成型植物启动子是能够在所有或几乎所有植物组织的所有或几乎所有植物发育阶段表达可读框(ORF)的启动子(称为"组成型表达")。组成型启动子的一个非限制性实例是花椰菜花叶病毒35S启动子。"调节启动子"是指非组成型地、但是以时间和/或空间调节方式指导基因表达的启动子,并且包括组织特异性、组织优选型及诱导型启动子。不同启动子可以指导在不同组织或细胞类型中、或在不同发育阶段、或响应于不同环境条件的基因表达。在特定实施方案中,AD官能化的CRISPR组分中的一者或多者在组成型启动子(诸如花椰菜花叶病毒35S启动子)的控制下表达,组织优选型启动子可以用于靶向特定植物组织中某些细胞类型,例如叶或根的维管细胞或种子的特定细胞内的增强的表达。用于所述系统的特定启动子的实例可见于Kawamata等人,(1997)Plant Cell Physiol 38:792-803;Yamamoto等人,(1997)Plant J 12:255-65;Hire等人,(1992)Plant Mol Biol 20:207-18;Kuster等人,(1995)Plant Mol Biol 29:759-72;以及Capana等人,(1994)Plant Mol Biol 25:681-91。

[0786] 在有限的情况下,为了避免脱氨酶的非特异性活性,诱导型启动子可以有利地表达所述系统的组分中的一者或多者。在特定实施方案中,所述系统的一个或多个元件在诱导型启动子的控制下表达。允许空间时间控制基因编辑或基因表达的诱导型启动子的实例可以使用能量形式。能量形式可以包括但不限于声能、电磁辐射、化学能和/或热能。诱导型系统的实例包括四环素诱导型启动子(Tet-On或Tet-Off)、小分子双杂交转录激活系统(FKBP、ABA等)、或光诱导型系统(光敏色素、LOV结构域或隐花色素),诸如以序列特异性方式引导转录活性改变的光诱导型转录效应子(LITE)。光诱导型系统的组分可以包括腺苷脱氨酶和Cas13的融合蛋白、光反应性细胞色素异源二聚体(例如,来自阿拉伯芥)。诱导型DNA结合蛋白及其使用方法的另外的实例提供于US 61/736465和US 61/721,283中,这些专利以引用方式整体并入本文。

[0787] 在特定实施方案中,可以通过使用例如化学调节型启动子来实现瞬时表达或诱导型表达,即由此应用外源性化学品诱导基因表达。也可以通过化学阻抑型启动子来获得对基因表达的调节,其中应用化学品阻抑基因表达。化学诱导型启动子包括但不限于:由苯磺酰胺除草剂安全剂激活的玉米ln2-2启动子(De Veylder等人,(1997)Plant Cell Physiol 38:568-77)、由用作萌前除草剂的疏水性亲电子化合物激活的玉米GST启动子(GST-11-27,WO 93/01294),以及由水杨酸激活的烟草PR-1a启动子(Ono等人,(2004)Biosci Biotechnol Biochem 68:803-7)。还可以在本文中使用由抗生素调节的启动子,诸如四环素诱导型启动子和四环素阻抑型启动子(Gatz等人,(1991)Mol Gen Genet 227:229-37;美

国专利号5,814,618和5,789,156)。

[0788] 特定植物细胞器中的易位和/表达

[0789] 表达系统可以包含在特定植物细胞器中易位并且/或者表达的元件。

[0790] 叶绿体靶向

[0791] 在特定实施方案中,设想的是将所述系统用于特异性地修饰叶绿体基因或确保在叶绿体中的表达。出于此目的,使用叶绿体转化方法或者将AD官能化的CRISPR组分区室化至叶绿体的方法。例如,在质粒基因组中引入遗传修饰可以减少生物安全性问题,诸如通过花粉的基因流。

[0792] 叶绿体转化的方法在本领域中是已知的,并且包括粒子轰击、PEG处理和微注射。另外,涉及转化盒从核基因组易位至质粒的方法可以如WO2010061186所述地使用。

[0793] 或者,设想的是将AD官能化的CRISPR组分中的一者或多者靶向植物叶绿体。这是通过在表达构建体中结合编码叶绿体转运肽(CTP)或质体转运肽的序列来实现的,该序列可操作地连接至编码腺苷脱氨酶和Cas13的融合蛋白的序列的5'区。在易位到叶绿体中的过程中,在处理步骤中去除CTP。所表达蛋白质的叶绿体靶向是本领域技术人员所熟知的(参见例如Protein Transport into Chloroplasts,2010,Annual Review of Plant Biology,第61卷:157-180)。在此类实施方案中,还期望将指导RNA靶向植物叶绿体。例如在US 20040142476中描述了可以用于借助于叶绿体定位序列来将指导RNA易位到叶绿体中的方法和构建体,该专利以引用方式并入本文。可以将构建体的此类变型并入本发明的表达系统中以有效地使所述系统组分易位。

[0794] 在藻类细胞中引入编码所述系统的多核苷酸.

[0795] 转基因藻类(或其他植物诸如油菜)可能在植物油或生物燃料诸如醇类(尤其是甲醇和乙醇)或其他产品的生产中特别有用。这些可以被工程化以表达或过表达高水平的油或醇类,以供在油或生物燃料行业中使用。

[0796] US 8945839描述了一种用于使用Cas9工程化微藻(莱茵衣藻(Chlamydomonas reinhardtii)细胞)的方法。使用类似的工具,本文所述系统的方法可以应用于衣藻属种和其他藻类。在特定实施方案中,在藻类中引入CRISPR-Cas蛋白(例如Cas13)、腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或适体结合衔接蛋白)和指导RNA,其使用在组成型启动子诸如Hsp70A-Rbc S2或β2-微管蛋白的控制下表达腺苷脱氨酶和Cas13的融合蛋白的载体进行表达。指导RNA任选地使用含有T7启动子的载体递送。或者,可以将Cas13 mRNA和体外转录的指导RNA递送至藻类细胞。电穿孔方案可由本领域技术人员获得,诸如来自GeneArt衣藻属工程化试剂盒(GeneArt Chlamydomonas Engineering kit)的标准推荐方案。

[0797] 在酵母细胞中引入系统组分

[0798] 在特定实施方案中,本发明涉及所述系统用于酵母细胞的基因组编辑的用途。可以用于引入编码所述系统组分的多核苷酸的转化酵母细胞的方法描述于Kawai等人,2010,Bioeng Bugs.2010年11月-12月;1(6):395-403)。非限制性实例包括通过醋酸锂处理(其还可以包括运载DNA和PEG处理)、轰击或通过电穿孔来转化酵母细胞。

[0799] 在植物和植物细胞中瞬时表达系统组分

[0800] 在特定实施方案中,设想的是在植物细胞中瞬时表达指导RNA和/或CRISPR-Cas基因。在这些实施方案中,仅当细胞中存在指导RNA、CRISPR-Cas蛋白(例如Cas13)和腺苷脱氨

酶(其可以融合至CRISPR-Cas蛋白或适体结合衔接蛋白)时,所述系统可以确保靶基因的修饰,这样使得可以进一步控制基因组修饰。因为CRISPR-Cas蛋白的表达是瞬时的,自此类植物细胞再生的植物典型地不含外源DNA。在特定实施方案中,CRISPR-Cas蛋白由植物细胞稳定表达,并且指导序列被瞬时表达。

[0801] 在特定实施方案中,可以使用植物病毒载体将所述系统组分引入植物细胞中(Scholthof等人1996,Annu Rev Phytopathol.1996;34:299-323)。在另外的特定实施方案中,所述病毒载体是来自DNA病毒的载体。例如,双粒病毒组(例如卷心菜曲叶病毒、豆黄矮病毒、小麦矮化病毒、番茄曲叶病毒、玉米条纹病毒、烟草曲叶病毒或番茄金色花叶病毒)或矮缩病毒组(例如蚕豆坏死黄脉病毒)。在其他特定实施方案中,所述病毒载体是来自RNA病毒的载体。例如,烟草脆裂病毒组(例如烟草扰乱病毒、烟草花叶病毒)、马铃薯X病毒组(例如马铃薯X病毒),或大麦病毒组(例如,大麦条纹花叶病毒)。植物病毒复制基因组是非整合型载体。

[0802] 在特定实施方案中,用于瞬时表达系统的载体是例如pEAQ载体,该载体被专门定制用于在原生质体中进行土壤杆菌介导的瞬时表达(Sainsbury F.等人,Plant Biotechnol J.2009Sep;7(7):682-93)。使用修饰的卷心菜叶曲病毒(CaLCuV)载体以在表达CRISPR酶的稳定转基因植物中表达指导RNA证明了基因组位置的精确靶向(Scientific Reports 5,文章编号:14926(2015),doi:10.1038/srep14926)。

[0803] 在特定实施方案中,可以将编码指导RNA和/或CRISPR-Cas基因的双链DNA片段瞬时导入到植物细胞中。在此类实施方案中,以足够的量提供引入的双链DNA片段以修饰细胞,但在预期时间段过去之后或者在一次或多次细胞分裂之后不再持续。用于在植物中进行直接DNA转移的方法是技术人员已知的(参见例如Davey等人Plant Mol Biol.1989年9月;13(3):273-85。)

[0804] 在其他实施方案中,将编码CRISPR-Cas蛋白(例如Cas13)和/或腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或适体结合衔接蛋白)的RNA多核苷酸引入到植物细胞中,然后植物细胞被生成足够量的蛋白质的宿主细胞翻译并加工以修饰该细胞(在至少一种指导RNA存在下),该引入在预期时间段过去之后或者在一次或多次细胞分裂之后不再持续。用于将mRNA引入植物原生质体以进行瞬时表达的方法是技术人员已知的(参见例如Gallie,Plant Cell Reports(1993),13;119-122)。

[0805] 还设想了以上所述的不同方法的组合。

[0806] 将系统组分递送至植物细胞

[0807] 在特定实施方案中,令人感兴趣的是将所述系统的一种或多种组分直接递送至植物细胞。对于产生非转基因植物,这是尤其令人感兴趣的(参见下文)。在特定实施方案中,在植物或植物细胞外部制备所述系统组分中的一者或多者并将其传递至细胞。例如,在特定实施方案中,在体外制备CRISPR-Cas蛋白,之后将其引入植物细胞中。可以通过本领域技术人员已知的各种方法来制备CRISPR-Cas蛋白,包括重组生产。在表达之后,将CRISPR-Cas蛋白分离,如果需要的话进行再折叠,进行纯化和任选地进行处理以去除任何纯化标签(诸如His标签)。一旦获得粗的、部分纯化的、或更完全纯化的CRISPR-Cas蛋白,就可以将该蛋白引入植物细胞中。

[0808] 在特定实施方案中,将CRISPR-Cas蛋白与靶向目标基因的指导RNA混合,以形成预

组装的核糖核蛋白。

[0809] 可以经由电穿孔、通过用CRISPR-Cas相关基因产品包覆的粒子的轰击、通过化学转染或通过跨细胞膜转运的一些其他手段，将这些单独的组分或预组装的核糖核蛋白引入到植物细胞中。例如，已证明用预组装CRISPR核糖核蛋白转染植物原生质体确保了植物基因组的靶向修饰(如Woo等人Nature Biotechnology,2015;DOI:10.1038/nbt.3389)。

[0810] 在特定实施方案中，使用纳米粒子将所述系统组分引入到植物细胞中。这些组分，无论是蛋白质或核酸或它们的组合，都可以上载到纳米粒子上或包装在纳米粒子中并且适用于这些植物(例如像WO 2008042156和US 20130185823所述的)。特别地，本发明的实施方案包括上载有或包装有以下各项的纳米粒子：编码CRISPR-Cas蛋白(例如Cas13)的一个或多个DNA分子、编码腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或适体结合衔接蛋白)的一个或多个DNA分子，以及编码指导RNA的DNA分子和/或如WO2015089419中所述的分离的指导RNA。

[0811] 将所述系统的一种或多种组分引入植物细胞中的其他手段是使用细胞穿透肽(CPP)。因此，在特定实施方案中，本发明包括含有连接至CRISPR-Cas蛋白的细胞穿透肽的组合物。在本发明的特定实施方案中，将CRISPR-Cas蛋白和/或指导RNA与一种或多种CPP偶联以有效地将它们转运到植物原生质体内。Ramakrishna(Genome Res.2014年6月;24(6):1020-7，针对人类细胞中的Cas9)。在其他实施方案中，CRISPR-Cas基因和/或指导RNA由一个或多个环状或非环状DNA分子编码，所述一个或多个环状或非环状DNA分子与一个或多个CPP偶联以便进行植物原生质体递送。然后将植物原生质体再生为植物细胞并进一步再生为植物。CPP通常被描述为来源于蛋白质或来源于能够以非受体依赖性方式跨细胞膜转运生物分子的嵌合序列的小于35个氨基酸的短肽。CPP可以是阳离子肽、具有疏水性序列的肽、两亲性肽、具有富含脯氨酸的序列和抗微生物序列的肽，以及嵌合肽或二分肽(Pooga和Langel2005)。CPP能够穿透生物膜，并且如此触发不同生物分子跨细胞膜移动到细胞质中，并能改善它们的细胞内通路，并且因此促进生物分子与靶标的相互作用。CPP的实例包括：Tat(通过HIV 1型进行病毒复制所需的核转录激活蛋白)、穿透素、卡波济(Kaposi)成纤维细胞增长因子(FGF)信号肽序列、整联蛋白β3信号肽序列;聚精氨酸肽Arg序列、富含鸟嘌呤的分子转运体、甜箭头肽(sweet arrow peptide)等。

[0812] 使用系统制备遗传修饰的非转基因植物

[0813] 在特定实施方案中，本文所述的方法用于修饰内源性基因或修饰其表达而不永久性引入到任何外源基因的植物的基因组中，包括编码CRISPR组分的外源基因，以便避免在植物基因组中存在外源DNA。这可能是令人感兴趣的，因为对非转基因植物的规则要求较不严格。

[0814] 在特定实施方案中，这是通过所述系统组分的瞬时表达来确保的。在特定实施方案中，一种或多种组分是在一种或多种病毒载体上表达的，所述一种或多种表达载体产生足够的CRISPR-Cas蛋白、腺苷脱氨酶和指导RNA，以一致地稳定地确保根据本文所述的方法修饰目标基因。

[0815] 在特定实施方案中，在植物原生质体中确保系统构建体的瞬时表达并且因此该构建体并未整合到基因组中。有限的表达窗口足以允许所述系统确保如本文所述的靶基因的修饰。

[0816] 在特定实施方案中，所述系统的不同组分借助于如上文所述的粒子递送分子诸如纳米粒子或CPP分子单独地或混合地引入在植物细胞、原生质体或植物组织中。

[0817] 所述系统组分的表达可以通过腺苷脱氨酶的脱氨酶活性诱导基因组的靶向修饰。上文所述的不同策略允许CRISPR介导的靶向基因组编辑，而无需将所述系统组分引入到植物基因组中。瞬时导入到植物细胞中的组分典型地在杂交后去除。

[0818] 植物培养和再生

[0819] 在特定实施方案中，具有修饰基因组并且通过本文所述的任何方法产生或获得的植物细胞可以被培养至再生成具有转化或修饰表型并因此具有所需表型的全株。常规的再生技术是本领域技术人员所熟知的。此类再生技术的特定实例依赖于组织培养生长培养基中某些植物激素的操纵，并且典型地依赖于已与所需核苷酸序列一起引入的杀生物剂和/或除草剂标记物。在另外的特定实施方案中，植物再生是从培养的原生质体、植物愈伤组织、外植体、器官、花粉、胚胎或其部分获得的(参见例如Evans等人(1983)，Handbook of Plant Cell Culture;Klee等人(1987)Ann.Rev.of Plant Phys.)。

[0820] 在特定实施方案中，如本文所述的转化或改良的植物可以自体受精以提供本发明的纯合改良植物(对于DNA修饰是纯合的)的种子，或者可以与非转基因植物或不同的改良植物杂交以提供纯合植物的种子。当将重组DNA引入到植物细胞中时，这种杂交所产生的植物是对于重组DNA分子为杂合的植物。通过与改良植物杂交并包含遗传修饰(可以是重组DNA)而获得的此类纯合植物和杂合植物在此被称为“子代”。子代植物是从原始转基因植物传代并且含有通过本文提供的方法引入的基因组修饰或重组DNA分子的植物。或者，遗传修饰植物可以是通过以上所述方法之一使用所述系统来获得的，因此无外源DNA并入该基因组中。通过进一步育种获得的此类植物的子代也可以含有遗传修饰。育种是通过常用于不同农作物的任何育种方法来进行(例如Allard,Principles of Plant Breeding,John Wiley&Sons,NY,U.of CA,Davis,CA,50-98(1960)。

[0821] 生成具有增强的农艺性状的植物

[0822] 本文素所提供的系统可用于引入靶向的A-G和T-C突变。通过在单一细胞中共表达旨在实现多种修饰的多个靶向RNA，可以确保多重基因组修饰。此技术可以用于高度精确工程化植物以使其具有改良的特征，包括增强的营养品质、增加的对疾病的抗性和对生物和非生物胁迫的抗性，以及增加的有商业价值的植物产品或异源化合物的产生。

[0823] 在特定实施方案中，如本文所述的系统用于引入靶向的A-G和T-C突变。这种突变可以是无义突变(例如，提前终止密码子)或错义突变(例如，编码不同的氨基酸残基)。当某些内源性基因中的A-G和T-C突变可以赋予或促成所需性状时，这是令人感兴趣的。

[0824] 本文所述的方法通常导致生成“改良植物”，在这点上它们与野生型植物相比具有一种或多种期望的性状。在特定实施方案中，获得的植物、植物细胞或植物部分是包含并入植物的所有或部分细胞的基因组中的外源性DNA序列的转基因植物。在特定实施方案中，获得非转基因遗传修饰植物、植物部分或细胞，在这点上没有外源性DNA序列并入植物的任何植物细胞的基因组中。在此类实施方案中，改良植物是非转基因的。当仅确保内源性基因的修饰并且在植物基因组中未引入或维持外源基因时，所得遗传修饰农作物不含有外源基因并且因此可以基本上认为是非转基因的。

[0825] 在特定实施方案中，多核苷酸通过DNA病毒(例如双粒病毒组)或RNA病毒(例如烟

草脆裂病毒组)递送至细胞。在特定实施方案中,引入步骤包括将含有编码CRISPR-Cas蛋白、腺苷脱氨酶和指导RNA的一个或多个多核苷酸序列的T-DNA递送至植物细胞,其中所述递送是经由土壤杆菌进行的。可以将编码所述系统的组分的多核苷酸序列可操作地连接至启动子,诸如组成型启动子(例如花椰菜花叶病毒35S启动子)或细胞特异性启动子或诱导型启动子。在特定实施方案中,通过微粒轰击引入多核苷酸。在特定实施方案中,所述方法还包括在引入步骤之后筛选植物细胞,以确定目标基因的表达是否已被修饰。在特定实施方案中,所述方法包括从植物细胞再生植物的步骤。在另外的实施方案中,所述方法包括使植物杂交育种以获得遗传上所需的植物谱系。

[0826] 在以上所述的方法的特定实施方案中,抗病性农作物是通过靶向突变疾病易感性基因或植物防卫基因中的编码负调节子的基因(例如,Mlo基因)来获得的。在特定实施方案中,耐除草剂农作物是通过靶向取代植物基因诸如编码乙酰乳酸合酶(ALS)和原卟啉原氧化酶(PPO)的基因的特定核苷酸来生成的。在特定实施方案中,通过靶向突变编码负调节子的基因而产生的具有非生物胁迫耐受性的耐干旱和盐农作物、通过靶向突变Waxy基因而产生的低直链淀粉谷物、通过靶向突变糊粉层中的主要脂肪酶基因而产生的具有降低的酸败性的稻谷或其他谷物等。在特定实施方案中。编码目标性状的内源性基因的更广泛列表列出在下文中。

[0827] 使用系统修饰多倍体植物

[0828] 许多植物都是多倍体的,这意味着它们携带其基因组的复制拷贝—像在小麦中,有时多至六个。根据本发明利用所述系统的方法可以被“多重化”以影响基因的所有拷贝,或者一次靶向许多基因。例如,在特定实施方案中,本发明的方法用于同时确保不同基因中负责阻遏针对疾病的防卫的功能丧失突变。在特定实施方案中,本发明的方法用于同时阻遏小麦植物细胞内TaMLO-A1、TaMLO-B1和TaMLO-D1核酸序列的表达并且由该细胞再生小麦植物,以便确保该小麦植物抵抗白粉病(另参见WO2015109752)。

[0829] 赋予农艺性状的示例性基因

[0830] 在特定实施方案中,本发明涵盖在内源性基因及其调控元件中包括靶向A-G和T-C突变的方法,所述内源性基因诸如以下列出的基因:

[0831] 1.赋予对害虫或疾病的抗性的基因:

[0832] 植物疾病抗性基因.可以用克隆的抗性基因转化植物以工程化对特定病原体菌株具有抗性的植物。参见例如Jones等人,Science 266:789(1994)(cloning of the tomato Cf-9gene for resistance to Cladosporium fulvum);Martin等人,Science 262:1432(1993)(tomato Pto gene for resistance to Pseudomonas syringae pv.tomato encodes a protein kinase);Mindrinos等人,Cell 78:1089(1994)(Arabidopsmay be RSP2 gene for resistance to Pseudomonas syringae)。可以对在病原体感染期间被上调或下调的植物基因进行工程化以获得病原体抗性。参见例如Thomazella等人,bioRxiv 064824;doi:https://doi.org/10.1101/064824Epub.2016年7月23日(tomato plants with deletions in the SlDMR6-1 which is normally upregulated during pathogen infection)。

[0833] 赋予对害虫诸如大豆囊胞线虫的抗性的基因。参见例如,PCT申请WO 96/30517;PCT申请WO 93/19181。

[0834] 苏云金芽孢杆菌蛋白,参见例如Geiser等人,Gene 48:109(1986)。

[0835] 凝集素,参见例如Van Damme等人,Plant Molec.Biol.24:25(1994。

[0836] 维生素结合蛋白,诸如抗生物素蛋白,参见PCT申请US93/06487,该申请教导了抗生物素蛋白和抗生物素蛋白同系物作为针对害虫的杀幼虫剂的用途。

[0837] 酶抑制剂诸如蛋白酶或朊酶抑制剂或淀粉酶抑制剂。参见例如Abe等人,J.Biol.Chem.262:16793(1987);Huub等人,Plant Molec.Biol.21:985(1993));Sumitani等人,Biosci.Biotech.Biochem.57:1243(1993)和美国专利号5,494,813。

[0838] 昆虫特异性激素或信息素,诸如蜕皮类固醇或保幼激素、或其变体、基于它的模拟物、或其拮抗剂或激动剂。参见例如Hammock等人,Nature 344:458(1990)。

[0839] 昆虫特异性肽或神经肽,这些肽在表达时破坏受影响害虫的生理。例如Regan,J.Biol.Chem.269:9(1994)和Pratt等人,Biochem.Biophys.Res.Comm.163:1243(1989)。另参见美国专利号5,266,317。

[0840] 在自然界中由蛇、黄蜂或任何其他生物体产生的昆虫特异性毒液。例如,参见Pang等人,Gene 116:165(1992)。

[0841] 引起单萜、倍半萜、甾体、异羟肟酸、苯丙素衍生物或具有杀昆虫活性的另一种非蛋白质分子超积累的酶。

[0842] 参与生物活性分子修饰(包括翻译后修饰)的酶;例如,糖解酶、蛋白水解酶、脂解酶、核酸酶、环化酶、转氨酶、酯酶、水解酶、磷酸酶、激酶、磷酸化酶、聚合酶、弹性蛋白酶、壳多糖酶以及葡聚糖酶,无论是天然还是合成的。参见PCT申请WO93/02197;Kramer等人,Insect Biochem.Molec.Biol.23:691(1993)和Kawalleck等人,Plant Molec.Biol.21:673(1993)。

[0843] 刺激信号转导的分子.例如,参见Botella等人,Plant Molec.Biol.24:757(1994)和Griess等人,Plant Physiol.104:1467(1994)。

[0844] 病毒侵入蛋白或源于此的复合物毒素。参见Beachy等人,Ann.rev.Phytopathol.28:451(1990)。

[0845] 在自然界中由病原体或寄生虫产生的发育阻滞蛋白。参见Lamb等人,Bio/Technology 10:1436(1992)和Toubart等人,Plant J.2:367(1992)。

[0846] 在自然界中由植物产生的发育阻滞蛋白。例如,Logemann等人,Bio/Technology10:305(1992)。

[0847] 在植物中,病原体通常是宿主特异性的。例如,一些镰刀菌种类将引起番茄枯萎病但仅攻击番茄,而其他镰刀菌种类仅攻击小麦。植物具有现有和诱导的防卫以抵抗大部分病原体。跨植物各代的突变和重组事件导致引起易感性的遗传变异性,特别是当病原体以比植物更大频率繁殖时。在植物中可以存在非宿主抗性,例如宿主和病原体是不相容的或者可以存在针对所有病原体种族的部分抗性,这些抗性典型地是通过许多基因来控制的,并且/或者也存在对一些病原体种族而不是其他种族的完全抗性。此抗性典型地是通过几种基因控制的。使用多种方法和系统组分,现在存在一种在此预先诱导特异性突变的新工具。因此,可以分析抗性基因来源基因组,并且在具有所需特征或形状的植物中,使用所述多种方法和所述系统组分来诱导抗性基因增加。本发明系统可以比先前的诱变剂更精确地完成此分析并且因此加速并改善植物育种程序。

[0848] 2.涉及植物疾病的基因,诸如WO 2013046247中列出的基因:

[0849] 稻谷病害:稻瘟病菌(Magnaporthe grise)、宫部旋孢腔菌(Cochliobolus miyabeanus)、立枯丝核菌(Rhizoctonia solani)、藤仓赤霉(Gibberella fujikuroi);小麦病害:白粉病菌(Erysiphe graminis)、禾谷镰刀菌(Fusarium graminearum)、燕麦镰刀菌(F.avenaceum)、黄色镰刀菌(F.culmorum)、雪霉叶枯菌(Microdochium nivale)、条形柄锈菌(Puccinia striiformis)、禾柄锈菌(P.graminis)、隐匿柄锈菌(P.recondita)、粉红雪腐病菌(Micronectriella nivale)、核瑚菌属种(Typhula sp.)、小麦黑粉菌(Ustilago tritici)、小麦网腥黑穗病菌(Tilletia caries)、小麦基腐病菌(Pseudocercosporella herpotrichoides)、禾生球腔菌(Mycosphaerella graminicola)、小麦壳多孢(Stagonospora nodorum)、偃麦草核腔菌(Pyrenophora tritici-repentis);大麦病害:白粉病菌、禾谷镰刀菌、燕麦镰刀菌、黄色镰刀菌、雪霉叶枯菌、条形柄锈菌、禾柄锈菌、大麦柄锈菌(P.hordei)、裸黑粉菌(Ustilago nuda)、大麦云纹斑病菌(Rhynchosporium secalis)、圆核腔菌(Pyrenophora teres)、禾旋孢腔菌(Cochliobolus sativus)、麦类核腔菌(Pyrenophora graminea)、立枯丝核菌:玉米病害:玉米黑粉菌(Ustilago maydis)、异旋孢腔菌(Cochliobolus heterostrophus)、高粱胶尾孢(Gloeocercospora sorghi)、多堆柄锈菌(Puccinia polysora)、玉米灰斑病菌(Cercospora zeae-maydis)、立枯丝核菌;

[0850] 柑橘病害:柑橘间座壳菌(Diaporthe citri)、柑橘痂囊腔菌(Elsinoe fawcetti)、指状青霉菌(Penicillium digitatum)、桔青霉菌(P.italicum)、寄生疫霉(Phytophthora parasitica)、柑橘褐腐疫霉(Phytophthora citrophthora);苹果病害:苹果链核盘菌(Monilinia mali)、苹果树腐烂病菌(Valsa ceratosperma)、苹果白粉病菌(Podosphaera leucotricha)、互隔交链孢菌苹果致病型(Alternaria alternata apple pathotype)、苹果黑星病菌(Venturia inaequalis)、尖孢炭疽(Colletotrichum acutatum)、恶疫霉(Phytophtora cactorum);

[0851] 梨病害:梨黑星病菌(Venturia nashicola)、梨黑星菌(V.pirina)、互隔交链孢霉日本梨致病型(Alternaria alternata Japanese pear pathotype)、梨胶锈菌(Gymnosporangium haraeanum)、恶疫霉;

[0852] 桃病害:桃褐腐病菌(Monilinia fructicola)、嗜果枝孢菌(Cladosporiumcarpophilum)、拟茎点霉属种(Phomopsis sp.);

[0853] 葡萄病害:痂囊腔菌(Elsinoe ampelina)、檬果炭疽病菌(Glomerella cingulata)、葡萄白粉菌(Uninula necator)、葡萄锈病菌(Phakopsora ampelopsidis)、葡萄球座菌(Guignardia bidwellii)、葡萄霜霉菌(Plasmopara viticola);

[0854] 柿子病害:柿盘长孢菌(Gloesporium kaki)、柿角斑病菌(Cercospora kaki)、柿叶球腔菌(Mycosphaerela nawae);

[0855] 瓠果病害:瓜类炭疽菌(Colletotrichum lagenarium)、黄瓜白粉病菌(Sphaerotheca fuliginea)、甜瓜球腔菌(Mycosphaerella melonis)、尖孢镰刀菌、黄瓜霜霉病菌(Pseudoperonospora cubensis)、疫霉属种(Phytophthora sp.)、腐霉属种(Pythium sp.);

[0856] 番茄病害:茄链格孢菌(Alternaria solani)、番茄叶霉病菌(Cladosporium fulvum)、致病疫霉菌(Phytophthora infestans);番茄丁香假单胞菌(Pseudomonas

syringae pv.Tomato);辣椒疫霉菌(Phytophthora capsici);黄单胞菌(Xanthomonas)

[0857] 茄子病害:茄褐纹病菌(Phomopsis vexans)、二孢白粉菌(Erysiphe cichoracearum);十字花科蔬菜病害:萝卜链格孢菌(Alternaria japonica)、白菜白斑病菌(Cercosporella brassicae)、根肿病菌(Plasmodiophora brassicae)、寄生霜霉菌(Peronospora parasitica);

[0858] 大葱病害:葱柄锈菌(Puccinia allii)、大葱霜霉(Peronospora destructor);

[0859] 大豆病害:大豆紫斑病菌(Cercospora kikuchii)、大豆痂囊腔菌(Elsinoe glycines)、菜豆间座壳大豆变种(Diaporthe phaseolorum var.sojae)、大豆壳针孢(Septoria glycines)、大豆尾孢(Cercospora sojina)、豆薯层锈菌(Phakopsora pachyrhizi)、大豆疫霉病菌(Phytophthora sojae)、立枯丝核菌、棒抱叶斑病菌(Corynespora casiicola)、核盘菌(Sclerotinia sclerotiorum);

[0860] 芸豆病害:菜豆炭疽病菌(Colletrichum lindemthianum);

[0861] 花生病害:花生黑斑病菌(Cercospora personata)、花生褐斑病菌(Cercospora arachidicola)、齐整小核菌(Sclerotium rolfsii);

[0862] 豌豆病害豌豆:豌豆白粉菌(Erysiphe pisi);

[0863] 马铃薯病害:茄链格孢菌、致病疫霉菌、马铃薯疫霉绯腐病菌(Phytophthora erythroseptica)、马铃薯粉状疮痂病菌(Spongospora subterranean,f.sp.Subterranean);

[0864] 草莓病害:薄草单丝壳菌(Sphaerotheca humuli)、檬果炭疽病菌;

[0865] 茶病害:茶网饼病菌(Exobasidium reticulatum)、茶白星病菌(Elsinoe leucospila)、拟盘多毛孢属种(Pestalotiopsis sp.)、茶炭疽菌(Colletotrichum theae-sinensis);

[0866] 烟草病害:烟草赤星病菌(Alternaria longipes)、二孢白粉菌、烟草炭疽病菌(Colletotrichum tabacum)、烟草霜霉菌(Peronospora tabacina)、烟草疫霉菌(Phytophthora nicotianae);

[0867] 油菜籽病害:核盘菌、立枯丝核菌;

[0868] 棉花病害:立枯丝核菌;

[0869] 甜菜病害:甜菜尾孢菌(Cercospora beticola)、水稻纹枯病菌(Thanatephorus cucumeris)、水稻纹枯病菌、螺壳状丝囊霉(Aphanomyces cochlioides);

[0870] 玫瑰病害:蔷薇双壳菌(Diplocarpon rosae)、蔷薇单丝壳茵(Sphaerotheca pannosa)、蔷薇霜霉(Peronospora sparsa);

[0871] 菊花和菊科病害:莴苣盘枝霉(Bremia lactuca)、野菊壳针抱(Septoria chrysanthemi-indici)、堀氏菊柄锈菌(Puccinia horiana);

[0872] 各种植物的病害:瓜果腐霉菌(Pythium aphanidermatum)、德巴利氏腐霉菌(Pythium debarianum)、禾草腐霉菌(Pythium graminicola)、畸雌腐霉菌(Pythium irregulare)、终极腐霉菌(Pythium ultimum)、灰葡萄孢菌(Botrytis cinerea)、核盘菌;

[0873] 萝卜病害:甘蓝链格孢菌(Alternaria brassicicola);

[0874] 结缕草病害:同果核盘菌(Sclerotinia homeocarpa)、立枯丝核菌;

[0875] 香蕉病害:香蕉黑条叶斑病菌(Mycosphaerella fijiensis)、香蕉黄条叶斑病菌

(Mycosphaerella musicola)；

[0876] 向日葵病害:向日葵霜霉病菌(Plasmopara halstedii)；

[0877] 在不同植物生长初期阶段由以下各项引起的种子疾病或病害:曲霉属种、青霉属种(Penicillium spp.)、镰刀菌属种(Fusarium spp.)、赤霉菌属种(Gibberella spp.)、木霉属种、根串珠霉属种(Thielaviopsis spp.)、根霉属种、毛霉菌属种(Mucor spp.)、伏革菌属种(Corticium spp.)、茎点霉属种(Rhoma spp.)、丝核菌属种(Rhizoctonia spp.)、色二孢属种(Diplodia spp.)等；

[0878] 由杆菌属种(Polymixa spp.)、油壶菌属种(Olpidium spp.)等介导的各种植物病毒病。

[0879] 3.赋予对除草剂的抗性的基因的实例：

[0880] 对抑制生长点或分生组织的除草剂的抗性,该除草剂诸如咪唑啉酮或硫酰脲,例如分别在Lee等人,EMBO J.7:1241(1988)和Miki等人,Theor.Appl.Genet.80:449(1990)中所述的。

[0881] 草甘膦耐受性(分别由例如突变体5-烯醇丙酮莽草酸-3-磷酸合酶(EPSP)基因、aroA基因和草甘膦乙酰转移酶(GAT)基因赋予的抗性)、或者对于其他膦羧基化合物诸如草铵膦的抗性(由来自链霉菌种类(包括吸水链霉菌(Streptomyces hygroscopicus)和绿色产色链霉菌(Streptomyces viridichromogenes))的草丁膦乙酰基转移酶(PAT)基因赋予)、以及对吡啶氧基或苯氧基丙酸和环异己酮的抗性(由ACC酶抑制剂编码基因赋予)。参见例如美国专利号4,940,835和美国专利号6,248,876、美国专利号4,769,061、欧洲专利号0 333 033和美国专利号4,975,374。另参见欧洲专利号0242246;DeGreef等人,Bio/Technology 7:61(1989);Marshall等人,Theor.Appl.Genet.83:435(1992);Castle等人的WO 2005012515,以及WO 2005107437。

[0882] 对抑制光合成的除草剂的抗性,该除草剂诸如三嗪(psbA和gs+基因)或苯基氰(腈水解酶基因),以及谷胱甘肽S-转移酶,如在Przibila等人,Plant Cell 3:169(1991);美国专利号4,810,648和Hayes等人,Biochem.J.285:173(1992)中所述的。

[0883] 编码使除草剂去毒的酶或对抑制具有抗性的突变体谷氨酰胺合酶的基因,例如在美国专利申请序列号11/760,602中所述的。或者去毒酶是编码草丁膦乙酰转移酶的酶(诸如来自链霉菌种类的bar或pat蛋白)。草丁膦乙酰转移酶例如描述于美国专利号5,561,236;5,648,477;5,646,024;5,273,894;5,637,489;5,276,268;5,739,082;5,908,810和7,112,665。

[0884] 羟基苯丙酮酸双氧化酶(HPPD)抑制剂,即天然存在的HPPD抗性酶,或者编码突变或嵌合HPPD酶的基因,如在WO 96/38567、WO 99/24585和WO 99/24586、WO 2009/144079、WO 2002/046387,或美国专利号6,768,044中所述的。

[0885] 4.涉及非生物胁迫耐受性的基因的实例：

[0886] 能够在植物细胞或植物中减少聚(ADP-核糖)聚合酶(PARP)基因的表达/或活性的转基因,如在WO 00/04173或WO/2006/045633中所述的。

[0887] 能够减少植物或植物细胞的PARG编码基因的表达和/或活性的转基因,如例如在WO 2004/090140中所述的。

[0888] 编码烟酰胺腺嘌呤二核苷酸补救合成途径的植物功能酶的转基因,这些酶包括烟

酰胺酶、烟酰酸磷酸核糖基转移酶、烟酸单核苷酸腺嘌呤转移酶、烟酰胺腺嘌呤二核苷酸合成酶或烟酰胺磷酸核糖基转移酶，如在例如EP 04077624.7、WO 2006/133827、PCT/EP07/002,433、EP 1999263或WO 2007/107326中所述的。

[0889] 涉及碳水化合物生物合成的酶包括在例如EP 0571427、WO 95/04826、EP 0719338、WO 96/15248、WO 96/19581、WO 96/27674、WO 97/11188、WO 97/26362、WO 97/32985、WO 97/42328、WO 97/44472、WO 97/45545、WO 98/27212、WO 98/40503、WO99/58688、WO 99/58690、WO 99/58654、WO 00/08184、WO 00/08185、WO 00/08175、WO 00/28052、WO 00/77229、WO 01/12782、WO 01/12826、WO 02/101059、WO 03/071860、WO 2004/056999、WO 2005/030942、WO 2005/030941、WO 2005/095632、WO 2005/095617、WO 2005/095619、WO 2005/095618、WO 2005/123927、WO 2006/018319、WO 2006/103107、WO 2006/108702、WO 2007/009823、WO 00/22140、WO 2006/063862、WO 2006/072603、WO 02/034923、EP 06090134.5、EP 06090228.5、EP 06090227.7、EP 07090007.1、EP 07090009.7、WO 01/14569、WO 02/79410、WO 03/33540、WO 2004/078983、WO 01/19975、WO 95/26407、WO 96/34968、WO 98/20145、WO 99/12950、WO 99/66050、WO 99/53072、美国专利号6,734,341、WO 00/11192、WO 98/22604、WO 98/32326、WO 01/98509、WO 01/98509、WO 2005/002359、美国专利号5,824,790、美国专利号6,013,861、WO 94/04693、WO 94/09144、WO 94/11520、WO 95/35026或WO 97/20936中所述的酶；或者如EP0663956、WO 96/01904、WO 96/21023、WO 98/39460和WO 99/24593所公开的涉及多聚果糖(尤其是菊粉和果聚糖类型)的产生的酶；如WO 95/31553、US 2002031826、美国专利号6,284,479、美国专利号5,712,107、WO 97/47806、WO 97/47807、WO 97/47808和WO 00/14249所公开的涉及α-1,4-葡聚糖的产生的酶；如WO 00/73422所公开的涉及α-1,6分支α-1,4-葡聚糖的产生的酶；如例如WO 00/47727、WO 00/73422、EP 06077301.7、美国专利号5,908,975和EP 0728213所公开的涉及交替糖的产生的酶；如例如WO 2006/032538、WO 2007/039314、WO 2007/039315、WO 2007/039316、JP 2006304779和WO 2005/012529所公开的涉及透明质酸的产生的酶。

[0890] 提高抗旱性的基因。例如，WO 2013122472公开了功能性泛素蛋白连接酶蛋白(UPL)、更具体地说是UPL3的缺乏或水平降低导致所述植物对水的需求减少或者对干旱的抗性提高。具有增加的耐旱性的转基因植物的其他实例公开于例如US 2009/0144850、US 2007/0266453和WO 2002/083911。US2009/0144850描述了一种由于DR02核酸表达的改变而展示耐旱性表型的植物。US 2007/0266453描述了一种由于DR03核酸表达的改变而展示耐旱性表型的植物，并且WO 2002/08391 1描述了一种由于在保卫细胞中表达的ABC转运体的活性降低而具有增加的对干旱胁迫的耐受性的植物。另一个实例是Kasuga和合著者(1999)的著作，他们描述了在正常生长条件下编码DREB1 A的cDNA在转基因植物中的过表达激活了许多胁迫耐受性基因的表达并且导致对干旱、盐负荷以及寒冷的耐受性提高。然而，在正常生长条件下DREB1A的表达也导致严重的生长迟缓(Kasuga(1999)Nat Biotechnol 17(3)287-291)。

[0891] 在另外的特定实施方案中，可以通过影响特定植物性状来改良农作物植物。例如，通过开发耐杀虫剂植物、提高植物的抗病性、提高植物的昆虫和线虫抗性、提高植物针对寄生杂草的抗性、提高植物的耐旱性、提高植物的营养价值、提高植物的胁迫耐受性、避免自花授粉、提高植物饲料消化率生物量、提高谷物产量等。在下文中提供了若干特定的非限制

性实例。

[0892] 除单一基因的靶向突变之外，系统可以被设计成允许在植物中靶向突变多基因、缺失染色体片段、位点特异性整合转基因、体内定点诱变、以及精确基因替代或等位基因交换。因此，本文描述的方法在基因发现和验证、突变和顺基因育种，以及杂交育种中具有广泛的应用。这些应用有助于产生新一代的具有各种改良的农艺性状的遗传修饰农作物，这些农艺性状诸如除草剂耐受性、抗病性、非生物胁迫耐受性、高产率以及优等品质。

[0893] 使用系统产生雄性不育植物

[0894] 杂交植物与自交植物相比典型地具有有利的农艺性状。然而，对于自花授粉植物，杂交体传代可能是有挑战的。在不同植物类型中，已经修饰了对植物能育性，更特别地雄性能育性至关重要的基因。例如，在玉米中，至少鉴定了两种对能育性至关重要的基因(Amitabh Mohanty International Conference on New Plant Breeding Molecular Technologies Technology Development And Regulation,2014年10月9-10日,Jaipur,India;Svitashev等人Plant Physiol.2015年10月;169(2):931-45;Djukanovic等人Plant J.2013Dec;76(5):888-99)。本文提供的方法和系统可以用于靶向雄性能育性所需要的基因，以便生成雄性不育植物，这些植物可以易于杂交以生成杂交体。在特定实施方案中，本文所提供的系统用于靶向诱变细胞色素P450-样基因(MS26)或大范围核酸酶基因(MS45)，从而向玉米植物赋予雄性不育性。如此遗传改变的玉米植物可以用于杂交育种程序。

[0895] 增加植物的能育性阶段

[0896] 特定实施方案，本文提供的方法和系统用于延长植物诸如稻谷植物的能育性阶段。例如，可以靶向稻谷能育性阶段基因诸如Ehd3，以便生成基因中的突变并且可以选择小植物以延长再生植物能育性阶段(如CN 104004782中所述的)。

[0897] 使用系统生成目标农作物的遗传变异

[0898] 农作物植物中野生种质和遗传变异的可用性是农作物改良程序的关键，但是来自农作物植物的种质的可用多样性是有限的。本发明设想了用于生成目标种质的遗传变异的多样性的方法。在所述系统的此应用中，提供了靶向植物基因组中的不同位置的指导RNA文库并且将该文库与CRISPR-Cas蛋白和腺苷脱氨酶一起引入到植物细胞中。以这种方式，可以生成基因组范围的点突变和基因敲除的集合。在特定实施方案中，所述方法包括由如此获得的细胞生成植物部分或植物以及针对目标性状筛选细胞。靶基因可以包含编码区和非编码区两者。在特定实施方案中，性状是胁迫耐受性，并且所述方法是用于生成胁迫耐受性农作物品种的方法。

[0899] 使用AD官能化的CRISPR影响果实催熟

[0900] 催熟是果实和蔬菜成熟过程中的正常阶段。仅在催熟开始后的几天，催熟致使果实或蔬菜不可食用。这个过程给农民和消费者都造成了重大损失。在特定实施方案中，本发明的方法用于减少乙烯的产生。这是通过确保以下各项中的一者或多者来确保的:a.阻遏ACC合酶基因表达。ACC(1-氨基环丙烷-1-羧酸)合酶是负责将S-腺苷甲硫氨酸(SAM)转化成ACC的酶;这是乙烯生物合成中的第二个步骤至最后一个步骤。当合酶基因的反义(“镜像”)或截短的拷贝插入到植物基因组中时会阻碍酶表达;b.插入ACC脱氨酶基因。从一种常见的非病原性土壤细菌绿针假单胞菌(Pseudomonas chlororaphis)获得编码该酶的基因。它将ACC转化为一种不同的化合物，从而减少可用于产生乙烯的ACC量;c.插入SAM水解酶基因。

这种方法类似于ACC脱氨酶,其中当乙烯前体代谢物的量减少时乙烯产生受到阻碍;在此情况下,SAM被转化为高丝氨酸。从大肠杆菌T3噬菌体获得编码该酶的基因;以及d.阻遏ACC氧化酶基因表达。ACC氧化酶是催化ACC氧化成乙烯的酶,这是乙烯生物合成途径中的最后一个步骤。使用本文所述的方法,下调ACC氧化酶基因,导致乙烯产生受到阻遏,从而延迟果实催熟。在特定实施方案中,对于以上所述修饰另外地或可替代地,本文所述的方法用于修饰乙烯受体,以便干扰由果实获得的乙烯信号。在特定实施方案中,修饰,更特别地说阻遏了编码乙烯结合蛋白的ETR1基因的表达。在特定实施方案中,对于以上所述修饰另外地或可替代地,本文所述的方法用于修饰编码多聚半乳糖醛酸酶(PG)的基因的表达,所述多聚半乳糖醛酸酶是负责分解果胶(维持植物细胞壁完整性的物质)的酶。果胶分解发生在催熟过程开始时,导致水果软化。因此,在特定实施方案中,本文描述的方法用于在PG基因中引入突变或阻遏PG基因的激活,以便减少所产生的PG酶的量,从而延迟果胶降解。

[0901] 因此,在特定实施方案中,所述方法包括使用所述系统来确保如上所述的植物细胞基因组的一种或多种修饰,并且由所述细胞再生植物。在特定实施方案中,植物是番茄植物。

[0902] 增加植物的保存期限

[0903] 在特定实施方案中,本发明的方法用于修饰涉及产生影响植物或植物部分的保存期限的化合物的基因。更特别地,修饰是在防止马铃薯块茎中的还原糖积累的基因中。在高温处理时,这些还原糖与游离氨基酸反应,产生棕色苦味产物和高水平的丙烯酰胺,丙烯酰胺是一种潜在的致癌物。在特定实施方案中,本文提供的方法用于减少或抑制液泡转化酶基因(VInv)的表达,所述液泡转化酶基因编码将蔗糖分解为葡萄糖和果糖的蛋白质(Clasen等人DOI:10.1111/pbi.12370)。

[0904] 使用系统确保增值的性状

[0905] 在特定实施方案中,所述系统用于产生营养提高的农作物。在特定实施方案中,本文提供的方法适于生成“功能性食品”,即可以提供超过它所含有的传统营养物的健康益处的修饰的食品或食品成分,并且/或者适于生成“营养食品”,即可以被视为食品或食品的一部分并且提供健康益处(包括预防和治疗疾病)的物质。在特定实施方案中,营养食品可用于预防并且/或者治疗癌症、糖尿病、心血管疾病以及高血压中的一者或多者。

[0906] 营养提高的农作物的实例包括(Newell-McGloughlin,Plant Physiology,2008年7月,第147卷,第939-953页):

[0907] 改良的蛋白质品质、含量和/或氨基酸组成,诸如对于以下各项所描述的:百喜草(Luciani等人2005,Florida Genetics Conference Poster)、油菜(Roesler等人,1997,Plant Physiol 113 75-81)、玉米(Cromwell等人,1967,1969J Anim Sci 26 1325-1331;O'Quin等人2000J Anim Sci 78 2144-2149;Yang等人2002,Transgenic Res 11 11-20;Young等人2004,Plant J 38 910-922)、马铃薯(Yu J和Ao,1997Acta Bot Sin 39329-334;Chakraborty等人2000,Proc Natl Acad Sci USA 97 3724-3729;Li等人2001)Chin Sci Bull 46 482-484、稻(Katsube等人1999,Plant Physiol 120 1063-1074)、大豆(Dinkins等人2001,Rapp 2002,In Vitro Cell Dev Biol Plant 37 742-747)、甘薯(Egnin和Prakash 1997,In Vitro Cell Dev Biol 33 52A)。

[0908] 必需氨基酸含量,诸如对于以下各项所描述的:油菜(Falco等人1995,Bio/

Technology 13 577-582)、羽扇豆(White等人2001,J Sci Food Agric 81 147-154)、玉米(Lai和Messing,2002,Agbios 2008GM crop database(March 11,2008))、马铃薯(Zeh等人2001,Plant Physiol 127 792-802)、高粱(Zhao等人2003,Kluwer Academic Publishers,Dordrecht,The Netherlands,第413-416页)、大豆(Falco等人1995Bio/Technology 13 577-582;Galili等人2002Crit Rev Plant Sci 21 167-204)。

[0909] 油类和脂肪酸,诸如对于油菜(Dehesh等人(1996)Plant J 9 167-172[PubMed];Del Vecchio(1996)INFORM International News on Fats,Oils and Related Materials 7230-243;Roesler等人(1997)Plant Physiol 113 75-81[PMC free article][PubMed];Froman和Ursin(2002,2003)Abstracts of Papers of the American Chemical Society 223U35;James等人(2003)Am J Clin Nutr 77 1140-1145[PubMed];Agbios(2008,above);棉花(Chapman等人(2001).J Am Oil Chem Soc 78 941-947;Liu等人(2002)J Am Coll Nutr 21 205S-211S[PubMed];O'Neill(2007)Australian Life Scientist.http://www.biotechnews.com.au/index.php/id;866694817;fp;4;fpid;2(June 17,2008)、亚麻籽(Abbadi等人,2004,Plant Cell 16:2734-2748)、玉米(Young等人,2004,Plant J 38 910-922)、油棕(Jalani等人1997,J Am Oil Chem Soc 74 1451-1455;Parveez,2003,AgBiotechNet 113 1-8)、稻(Anai等人,2003,Plant Cell Rep 21 988-992)、大豆(Reddy和Thomas,1996,Nat Biotechnol 14 639-642;Kinney和Kwolton,1998,Blackie Academic and Professional,London,pp 193-213)、向日葵(Arcadia,Biosciences 2008)

[0910] 碳水化合物,诸如对于以下各项所描述的果聚糖:菊苣(Smeekens(1997)Trends Plant Sci 2 286-287;Sprenger等人(1997)FEBS Lett 400 355-358;Sévenier等人(1998)Nat Biotechnol 16 843-846)、玉米(Caimi等人(1996)Plant Physiol 110 355-363)、马铃薯(Hellwege等人,1997Plant J 12 1057-1065)、甜菜(Smeekens等人1997,同上);菊粉,诸如对于马铃薯所述(Hellewege等人2000,Proc Natl Acad Sci USA 97 8699-8704);淀粉,诸如对于稻所述的(Schwall等人(2000)Nat Biotechnol 18 551-554;Chiang等人(2005)Mol Breed 15 125-143),

[0911] 维生素和类胡萝卜素,诸如对于以下各项所述的:油菜(Shintani和DellaPenna(1998)Science 282 2098-2100)、玉米(Rocheford等人(2002).J Am Coll Nutr 21191S-198S;Cahoon等人(2003)Nat Biotechnol 21 1082-1087;Chen等人(2003)Proc Natl Acad Sci USA 100 3525-3530)、芥菜籽(Shewmaker等人(1999)Plant J 20401-412)、马铃薯(Ducreux等人,2005,J Exp Bot 56 81-89)、稻(Ye等人(2000)Science287 303-305)、草莓(Agius等人(2003),Nat Biotechnol 21 177-181)、番茄(Rosati等人(2000)Plant J 24 413-419;Fraser等人(2001)J Sci Food Agric 81 822-827;Mehta等人(2002)Nat Biotechnol 20 613-618;Díaz de la Garza等人(2004)Proc Natl Acad Sci USA101 13720-13725;Enfissi等人(2005)Plant Biotechnol J 3 17-27;DellaPenna(2007)Proc Natl Acad Sci USA 104 3675-3676。

[0912] 功能性次级代谢产物,诸如对于以下各项所述的:苹果(芪类,Szankowski等人(2003)Plant Cell Rep 22:141-149)、苜蓿(白藜芦醇,Hipskind和Paiva(2000)Mol Plant Microbe Interact 13 551-562)、猕猴桃(白藜芦醇,Kobayashi等人(2000)Plant Cell Rep19 904-910)、玉米和大豆(类黄酮,Yu等人(2000)Plant Physiol 124 781-794)、马铃

薯(花青素和生物碱糖苷,Lukaszewicz等人(2004)J Agric Food Chem 52 1526-1533)、稻(类黄酮和白藜芦醇,Stark-Lorenzen等人(1997)Plant Cell Rep 16 668-673;Shin等人(2006)Plant Biotechnol J 4 303-315)、番茄(+白藜芦醇、绿原酸、类黄酮、芪;Rosati等人(2000)同上;Muir等人(2001)Nature 19 470-474;Niggeweg等人(2004)Nat Biotechnol 22 746-754;Giovinazzo等人(2005)Plant Biotechnol J 3 57-69)、小麦(咖啡酸和阿魏酸、白藜芦醇;United Press International(2002));以及

[0913] 矿物质可用性,诸如对于以下各项所述的:苜蓿(植酸酶,Austin-Phillips等人(1999)http://www.molecularfarming.com/nonmedical.html)、生菜(Lettuse)(铁;Goto等人(2000)Theor Appl Genet 100 658-664)、稻(铁,Lucca等人(2002)J Am Coll Nutr 21184S-190S)、玉米、大豆和小麦(植酸酶,Drakakaki等人(2005)Plant Mol Biol 59869-880;Denbow等人(1998)Poult Sci 77 878-881;Brinch-Pedersen等人(2000)Mol Breed 6 195-206)。

[0914] 在特定实施方案中,增值的性状与存在于植物中的化合物的设想的健康益处相关。例如,在特定实施方案中,通过应用本发明的方法来确保以下化合物中的一者或多者的合成的修改或者诱导/增加它们的合成,以获得增值的农作物:

[0915] • 类葫萝卜素,诸如存在于胡萝卜中的α-胡萝卜素,其中和可引起对细胞的损害的自由基;或者存在于各种果实和蔬菜中的β-胡萝卜素,其中和自由基

[0916] • 存在于绿色蔬菜中的叶黄素,其有助于维持健康视力

[0917] • 存在于番茄和番茄产品中的番茄红素,认为其降低前列腺癌风险;

[0918] • 存在于柑橘和玉米中的玉米黄素,其有助于维持健康视力

[0919] • 膳食纤维,诸如存在于麦麸中的不溶性纤维,其可以降低乳腺癌和/或结肠癌风险;以及存在于燕麦中的β葡聚糖;存在于车前子(Psylium)和全谷粒中的可溶性纤维,其可以降低心血管疾病(CVD)风险;

[0920] • 脂肪酸,诸如ω-3脂肪酸,其可以降低CVD风险并提高心理功能和视功能;缀合亚油酸,其可以改善身体组成,可以减小某些癌症风险;以及GLA,其可以降低癌症和CVD的炎症风险,可以改善身体组成;

[0921] • 黄酮类,诸如存在于小麦中的羟基苯乙烯,其具有抗氧化剂样活性,可以降低退行性疾病风险;存在于果实和蔬菜中的黄酮醇、儿茶酚和鞣酸,它们中和自由基并且可以降低癌症风险;

[0922] • 葡萄糖异硫氰酸酯、吲哚、异硫氰酸酯,诸如存在于十字花科蔬菜(花椰菜、羽衣甘蓝)、辣根中的萝卜硫素,其中和自由基,可以降低癌症风险;

[0923] • 酚类,诸如存在于葡萄中的芪类,其可以降低退行性疾病、心脏病和癌症的风险,可以有延年益寿功效;以及存在于蔬菜和柑橘中的咖啡酸和阿魏酸,它们具有抗氧化剂样活性,可以降低退行性疾病、心脏病和眼病的风险;以及存在于可可中的表儿茶素,其具有抗氧化剂样活性,可以降低退行性疾病和心脏病的风险;

[0924] • 存在于玉米、大豆、小麦和木制油中的植物甾烷醇/固醇,它们可以通过减低血胆固醇水平来降低冠心病风险;

[0925] • 存在于洋姜、胡葱、洋葱粉中的果聚糖、菊糖、低聚果糖,它们可以提高胃肠健康;

[0926] • 存在于大豆中的皂苷,其可以减低LDL胆固醇;

[0927] • 存在于大豆中的大豆蛋白质,其可以降低心脏病风险;

[0928] • 植物雌激素,诸如存在于大豆中的异黄酮,其可以减少绝经期症状(诸如热潮红),可以减少骨质疏松症和CVD;以及存在于亚麻、黑麦和蔬菜中的木脂素,其可以防止心脏病和一些癌症,可以降低LDL胆固醇、总胆固醇。;

[0929] • 硫化物和硫醇,诸如存在于洋葱、大蒜、橄榄、韭葱和青葱(scallon)中的二烯丙基硫;以及存在于十字花科蔬菜中的烯丙基甲基三硫、二硫醇硫酮,它们可以降低LDL胆固醇,帮助维持健康免疫系统;以及

[0930] • 鞣酸,诸如存在于蔓越橘、可可中的原花色素,其可以提高泌尿道健康,可以降低CVD和高血压风险。

[0931] 此外,本发明的方法还设想了修改蛋白质/淀粉功能性、保质期、味道/美学、纤维品质、以及减少过敏原、抗营养素以及毒素的形状。

[0932] 因此,本发明涵盖了用于产生具有营养增加价值的植物的方法,所述方法包括使用如本文所述的系统将编码涉及产生具有增加的营养价值的组分的酶的基因引入到植物细胞中并且由所述植物细胞再生植物,所述植物的特征在于所述具有增加的营养价值的组分的表达增加。在特定实施方案中,所述系统用于例如通过修饰控制此化合物代谢的一种或多种转录因子来间接修改这些化合物的内源性合成。上文描述了用于使用所述系统将目标基因引入到植物细胞中并且/或者修饰内源性基因的方法。

[0933] 在已修饰为赋予增值性状的植物中的一些具体修饰实例是:具有修饰的脂肪酸代谢的植物,方式为例如用硬脂酰-ACP去饱和酶的反义基因转化植物以增加植物硬脂酸含量的。参见Knultzon等人,Proc.Natl.Acad.Sci.U.S.A.89:2624(1992)。另一个实例涉及例如通过克隆接着再引入与可以负责特征为低水平植酸的玉米突变体的单一等位基因相关联的DNA来降低植酸酯含量。参见Raboy等人,Maydica 35:383(1990)。

[0934] 类似地,在强启动子控制下调节玉米糊粉层中黄酮类的产生的玉米(玉蜀黍)Tf s C1和R的表达导致拟南芥属(阿拉伯芥)中的花青素的高积累速率,推测是通过激活整个途径(Bruce等人,2000,Plant Cell 12:65-80)。DellaPenna(Welsch等人,2007Annu Rev Plant Biol 57:711-738)发现Tf RAP2.2及其相互作用配偶体SINAT2增加拟南芥叶中的胡萝卜素形成作用。在转基因拟南芥中表达Tf Dof1诱导了编码用于产生碳架、标记性增加氨基酸含量以及减少Glc水平的酶的基因的上调(Yanagisawa,2004Plant Cell Physiol 45:386-391),并且DOF Tf AtDof1.1(OBP2)上调了拟南芥的葡萄糖异硫氰酸酯生物合成途径中的所有步骤(Skirycz等人,2006Plant J 47:10-24)。

[0935] 减少植物中的过敏原

[0936] 在特定实施方案中,本文提供的方法用于生成具有减少的水平的过敏原的植物,从而使得它们对于消费者而言更安全。在特定实施方案中,所述方法包括修饰负责产生植物过敏原的一种或多种基因的表达。例如,在特定实施方案中,所述方法包括下调植物细胞诸如黑麦草植物细胞中的Lol p5基因的表达并且由所述植物细胞再生植物以便减少所述植物的花粉的过敏原性(Bhalla等人1999,Proc.Natl.Acad.Sci.USA第96卷:11676-11680)。

[0937] 花生过敏和对豆类过敏总体上是真实而严重的健康问题。本发明的系统可以用于

鉴定然后突变编码此类豆类的过敏原性蛋白的基因。不限于此类基因和蛋白质,Nicolaou等人鉴定了花生、大豆、扁豆、豌豆、羽扇豆、青豆以及绿豆中的过敏原性蛋白。参见Nicolaou等人,Current Opinion in Allergy and Clinical Immunology 2011;11(3):222)。

[0938] 用于目标内源性基因的筛选方法

[0939] 本文提供的方法进一步允许鉴定编码涉及产生具有增加的营养价值的组分的酶的有价值基因或者通常是影响跨种类、门和植物界的目标农艺性状的基因。通过使用如本文所述的CRISPR系统选择性地靶向例如编码植物代谢途径的酶的基因,可以鉴定负责植物某些营养方面的基因。类似地,通过选择性地靶向可以影响所期望的农艺性状的基因,可以鉴定相关基因。因此,本发明涵盖了用于编码涉及产生具有特定营养价值和/或农艺性状的化合物的酶的基因的筛选方法。

[0940] 系统在植物和酵母中的进一步应用

[0941] 系统在生物燃料生产中的用途

[0942] 如本文所用,术语“生物燃料”是由植物和植物来源的资源制成的替代燃料。可以从有机物质中提取可再生生物燃料,已通过碳固定方法获得或者通过使用或转化生物质制成有机物质的能量。此生物质可以直接用于生物燃料或者可以通过热转化、化学转化和生物化学转化来转化成含有能量的便利物质。这种生物质转化可以产生固体、液体或气体形式的燃料。存在两种类型的生物燃料:生物乙醇和生物柴油。生物乙醇主要是通过纤维素(淀粉)的糖发酵过程产生的,纤维素大部分来源于玉米和甘蔗。在另一方面,生物柴油主要是由油料作物诸如油菜籽、棕榈和大豆产生的。生物燃料主要用于运输。

[0943] 增强用于生物燃料生产的植物特性

[0944] 在特定实施方案中,使用如本文所述的系统的方法用于改变细胞壁的特性,以便促进关键性水解剂进入,从而更有效地释放用于发酵的糖。在特定实施方案中,修改纤维素和/或木质素的生物合成。纤维素是细胞壁的主要组分。纤维素和木质素的生物合成是共调节的。通过减少植物中的木质素比例,可以增加纤维素的比例。在特定实施方案中,本文所述的方法用于下调植物中的木质素生物合成,以便增加可发酵的碳水化合物。更特别地,本文所述的方法用于下调如WO 2008064289A2所公开的选自由以下组成的组的至少第一木质素生物合成基因:4-香豆酸酯3-羟化酶(C3H)、苯丙氨酸氨-裂解酶(PAL)、肉桂酸酯4-羟化酶(C4H)、羟基肉桂酰转移酶(HCT)、咖啡酸O-甲基转移酶(COMT)、咖啡酰辅酶A 3-O-甲基转移酶(CCoAOMT)、阿魏酸酯5-羟化酶(F5H)、肉桂醇脱氢酶(CAD)、肉桂酰辅酶A-还原酶(CCR)、4-香豆酸酯辅酶A连接酶(4CL)、单木质醇-木质素-特异性糖基转移酶,以及醛脱氢酶(ALDH)。

[0945] 在特定实施方案中,本文所述的方法用于产生在发酵过程中生成较低水平乙酸的植物生物质(另参见WO 2010096488)。更特别地,本文所公开的方法用于生成与Cas1L同源的突变,以减少多糖乙酰化。

[0946] 修饰用于生物燃料生产的酵母

[0947] 在特定实施方案中,本文所提供的系统用于通过重组微生物进行生物乙醇生产。例如,所述系统可以用于工程化微生物,诸如酵母,以由可发酵糖生成生物燃料或生物聚合物并且任选地能够降解来源于作为可发酵糖来源的农业废弃物的植物来源的木质纤维素。

在一些实施方案中,所述系统用于修饰与生物燃料生产途径竞争的内源性代谢途径。

[0948] 因此,在更具体的实施方案中,本文所述的方法用于如下修饰微生物:修饰编码所述宿主细胞的代谢途径中的酶的至少一种核酸,其中所述途径产生除来自丙酮酸酯的乙醛或来自乙醛的乙醇之外的代谢物,并且其中所述修饰导致所述代谢物的产生减少,或者引入编码所述酶的抑制剂的至少一种核酸。

[0949] 修饰用于植物油或生物燃料生产的藻类和植物

[0950] 例如,转基因藻类或其他植物诸如油菜可能在植物油或生物燃料诸如醇类(尤其是甲醇和乙醇)的生产中特别有用。这些可以被工程化以表达或过表达高水平的油或醇类,以供在油或生物燃料行业中使用。

[0951] 根据本发明的特定实施方案,所述系统用于生成可用于生物燃料生产的富含脂质的硅藻。

[0952] 在特定实施方案中,设想的是特异性地修饰涉及改变由藻类细胞产生的脂质的量和/或脂质的品质的基因。编码涉及脂肪酸合成途径的酶的基因的实例可以编码具有例如以下酶活性的蛋白质:乙酰辅酶A羧化酶、脂肪酸合酶、3-酮乙基_酰基-运载蛋白合酶III、甘油-3-磷酸脱氢酶(G3PDH)、烯酰-酰基运载蛋白还原酶(烯酰-ACP-还原酶)、甘油-3-磷酸酰基转移酶、溶血磷脂酰基转移酶或二酰甘油酰基转移酶、磷脂:二酰甘油酰基转移酶、磷脂酸磷酸酶、脂肪酸硫酯酶诸如软脂酰蛋白硫酯酶,或苹果酸酶活性。在另外的实施方案中,设想的是生成具有增加的脂质积累的硅藻。这可以通过靶向减少脂质分解代谢的基因来实现。对于用于本发明的方法中特别令人感兴趣的是涉及激活三酰基甘油和游离脂肪酸的基因,以及直接涉及脂肪酸的β氧化的基因,诸如脂酰辅酶A合成酶、3-酮脂酰辅酶A硫解酶、脂酰辅酶A氧化酶活性以及磷酸葡萄糖变位酶。本文所述的系统和方法可以用于特异性地激活硅藻中的此类基因以增加其脂质含量。

[0953] 诸如微藻的生物被广泛用于合成生物学。Stovicek等人(Metab.Eng.Comm.,2015;2:13描述了工业用酵母例如酿酒酵母的基因组编辑,以有效产生用于工业生产的有力菌株。Stovicek使用了对于酵母进行密码子优化的CRISPR-Cas9系统来同时破坏内源性基因的两个等位基因并且敲除异源基因。Cas9和指导RNA由基因组或附加型2μ基载体位置表达。这些作者们还证实基因破坏效率可以通过优化Cas9和指导RNA表达水平来提高。Hlavová等人(Biotechnol.Adv.2015)论述了使用诸如CRISPR的技术靶向核基因和叶绿体基因进行插入诱变和筛选来开发微藻种类或菌株。

[0954] US 8945839描述了一种用于使用Cas9工程化微藻(莱茵衣藻细胞)的方法。使用类似的工具,本文所述系统的方法可以应用于衣藻属种和其他藻类。在特定实施方案中,在藻类中引入CRISPR-Cas蛋白(例如Cas13)、腺苷脱氨酶(其可以融合至CRISPR-Cas蛋白或适体结合衔接蛋白)和指导RNA,其使用在组成型启动子诸如Hsp70A-Rbc S2或β2-微管蛋白的控制下表达CRISPR-Cas蛋白和任选地腺苷脱氨酶的载体进行表达。指导RNA将使用含有T7启动子的载体递送。或者,可以将mRNA和体外转录的指导RNA递送至藻类细胞。电穿孔方法遵循来自GeneArt衣藻属工程化试剂盒的标准推荐方法。

[0955] 系统在产生能够进行脂肪酸生产的微生物中的用途

[0956] 在特定实施方案中,本发明的方法用于生成能够产生脂肪酸酯诸如脂肪酸甲酯(“FAME”)和脂肪酸乙酯(“FAEE”)的遗传工程化微生物。

[0957] 典型地，宿主细胞可以被工程化以通过表达或过表达编码硫酯酶的基因、编码脂酰辅酶A合酶的基因以及编码酯合酶的基因来由存在于培养基中的碳源诸如醇产生脂肪酸酯。因此，本文提供的方法用于修饰微生物，以便过表达或引入硫酯酶基因、编码脂酰辅酶A合酶的基因以及编码酯合酶的基因。在特定实施方案中，硫酯酶基因选自tesA、'tesA、tesB、fatB、fatB2、fatB3、fatA1或fatA。在特定实施方案中，编码脂酰辅酶A合酶的基因选自fadDJadK、BH3103、pfl-4354、EAV15023、fadD1、fadD2、RPC_4074、fadDD35、fadDD22、faa39，或编码具有相同特性的酶的鉴定的基因。在特定实施方案中，编码酯合酶的基因是编码来自以下各项的合酶/脂酰辅酶A:二酰基甘油酰基转移酶的基因:霍霍巴(Simmondsia chinensis)、不动杆菌属种ADP、泊库岛食烷菌(Alcanivorax borkumensis)、铜绿假单胞菌(Pseudomonas aeruginosa)、亚德海床杆菌(Fundibacter jadensis)、阿拉伯芥或真养产碱杆菌(Alkaligenes eutrophus)，或其变体。

[0958] 另外地或可替代地，本文提供的方法用于减少以下各项中的至少一者在所述微生物中的表达:编码脂酰辅酶A脱氢酶的基因、编码外膜蛋白受体的基因，以及编码脂肪酸生物合成转录调节子的基因。在特定实施方案中，诸如通过引入突变使这些基因中的一者或多者失活。在特定实施方案中，编码脂酰辅酶A脱氢酶的基因是fadE。在特定实施方案中，编码脂肪酸生物合成的转录调节子的基因编码DNA转录阻抑因子，例如fabR。

[0959] 另外地或可替代地，所述微生物被修饰为减少以下各项中的至少一者的表达:编码丙酮酸甲酸裂解酶的基因、编码乳酸脱氢酶的基因或二者。在特定实施方案中，编码丙酮酸甲酸裂解酶的基因是pflB。在特定实施方案中，编码乳酸脱氢酶的基因是IdhA。在特定实施方案中，诸如通过在其中引入突变使这些基因中的一者或多者失活。

[0960] 在特定实施方案中，微生物埃希菌属(Escherichia)、芽孢杆菌属(Bacillus)、乳酸杆菌属(Lactobacillus)、红球菌属(Rhodococcus)、聚球蓝细菌属(Synechococcus)、集胞藻属(Synechoystis)、假单胞菌属、曲霉属、木霉属、脉孢菌属、镰刀菌属、腐质霉属(Humicola)、根毛霉属(Rhizomucor)、克鲁维酵母属、毕赤酵母属、毛霉菌属、蚀丝霉属(Myceliophtora)、青霉属、平革菌属(Phanerochaete)、侧耳属(Pleurotus)、栓菌属(Trametes)、金孢子菌属(Chrysosporium)、酵母菌属(Saccharomyces)、寡养单胞菌属(Stenotrophamonas)、裂殖酵母属(Saccharomyces)、亚罗酵母属或链霉菌属。

[0961] 系统在产生能够进行有机酸生产的微生物中的用途

[0962] 本文提供的方法还用于工程化能够更具体地说由戊糖或己糖生产有机酸的微生物。在特定实施方案中，所述方法包括将外源性LDH基因引入到微生物中。在特定实施方案中，所述微生物中的有机酸生产另外地或可替代地通过使编码涉及内源性代谢途径的蛋白质的内源性基因失活来增加，所述代谢途径产生除目标有机酸之外的代谢物，并且/或者其中所述内源性代谢途径消耗有机酸。在特定实施方案中，所述修饰确保减少除目标有机酸之外的代谢物的产生。根据特定实施方案，所述方法用于引入其中消耗有机酸的内源性途径或编码涉及产生除目标有机酸之外的代谢物的内源性途径的产物的基因的至少一种工程化基因缺失和/或失活。在特定实施方案中，所述至少一种工程化基因缺失或失活是在编码选自由以下组成的组的酶的一种或多种基因中:丙酮酸脱羧酶(pdc)、延胡索酸还原酶、醇脱氢酶(adh)、乙醛脱氢酶、磷酸烯醇丙酮酸羧化酶(ppc)、D-乳酸脱氢酶(d-ldh)、L-乳酸脱氢酶(l-ldh)、乳酸2-单加氧酶。在另外的实施方案中，所述至少一种工程化基因缺失和/

或失活是在编码丙酮酸脱羧酶(pdc)的内源性基因中。

[0963] 在另外的实施方案中,微生物被工程化以产生乳酸,并且所述至少一种工程化基因缺失和/或失活是在编码乳酸脱氢酶的内源性基因中。另外地或可替代地,微生物包含编码细胞色素依赖性乳酸脱氢酶诸如细胞色素B2依赖性L-乳酸脱氢酶的内源性基因的至少一种工程化基因缺失或失活。

[0964] 系统在产生改良的利用木糖或纤维二糖的酵母菌株中的用途

[0965] 在特定实施方案中,所述系统可以应用于选择改良的利用木糖或纤维二糖的酵母菌株。易错PCR可以用于扩增涉及木糖利用或纤维二糖利用途径的一种(或多种)基因。涉及木糖利用途径和纤维二糖利用途径的基因的实例可以包括但不限于描述于Ha,S.J.等人(2011)Proc.Natl.Acad.Sci.USA 108(2):504-9以及Galazka,J.M.等人(2010)Science 330(6000):84-6中的那些基因。如WO2015138855所述,各自在这种选择的基因中包含随机突变的双链DNA分子的所得文库可以与所述系统的组分共转化到酵母菌株(例如S288C)中,并且可以选择具有增加的木糖或纤维二糖利用能力的菌株。

[0966] 系统在产生用于类异戊二烯生物合成的改良酵母菌株中的用途

[0967] Tadas Jakočiūnas等人描述了多重CRISPR/Cas9系统在面包酵母酿酒酵母的一个转化步骤中用于基因组工程化多至5个不同基因组基因座的成功应用(Metabolic Engineering第28卷,2015年3月,第213-222页),得到具有高甲羟戊酸酯(它是工业上重要的异戊二烯生物合成途径的关键中间体)产量的菌株。在特定实施方案中,所述系统可以应用于如本文所述的用于鉴定在异戊二烯合成中使用的另外高产的酵母菌株的多重基因组工程化方法中。

[0968] 改良的植物和酵母细胞

[0969] 本发明还提供了通过本文提供的方法可获得并且通过本文提供的方法获得的植物和酵母细胞。通过本文所述的方法获得的改良的植物可以适用于通过表达确保例如对植物害虫、除草剂、干旱、低温或高温、过量水等耐受的基因来进行食品或饲料生产。

[0970] 通过本文所述的方法获得的改良的植物,尤其是农作物和藻类可以适用于通过表达例如比野生型中通常所见更高的蛋白质、碳水化合物、营养素或维生素水平来进行食品或饲料生产。就这一点而言,改良的植物,尤其是豆类和块茎类是优选的。

[0971] 改良的藻类或其他植物诸如油菜可能在植物油或生物燃料诸如醇类(尤其是甲醇和乙醇)的生产中特别有用。这些可以被工程化以表达或过表达高水平的油或醇类,以供在油或生物燃料行业中使用。

[0972] 本发明还提供了改良的植物部分。植物部分包括但不限于叶、茎、根、块茎、种子、胚乳、胚珠以及花粉。如本文所设想的植物部分可以是有活力的、无活力的、可再生的和/或不可再生的。

[0973] 本文还涵盖的是提供根据本发明的方法生成的植物细胞和植物。在本发明的范围内还包括通过传统育种方法产生的包含遗传修饰的植物的配子、种子、胚胎(合子胚或体细胞胚)、子代或杂交体。此类植物可以含有插入在靶序列处或代替靶序列的异源或外源DNA序列。或者,此类植物可以仅含有在一个或多个核苷酸中的改变(突变、缺失、插入、取代)。这样,此类植物与祖代植物的不同之处仅在于特定修饰的存在。

[0974] 因此,本发明提供了通过本方法产生的植物、动物或细胞或其子代。子代可以是产

生的植物或动物的克隆，或者可以由通过与相同种类的其他个体杂交以使另外的期望的性状渗入其后代来进行的有性繁殖产生。在多细胞生物体(特别是动物或植物)的情况下，细胞可以是体内或离体的。

[0975] 用于使用如本文所述的系统进行基因组编辑的方法可以用于对基本上任何植物、藻类、真菌、酵母等赋予所需的性状。针对本文所述的所需生理以及农艺特征，可以使用本公开的核酸构建体和以上提及的各种转化方法对多种多样的植物、藻类、真菌、酵母等以及植物藻类、真菌、酵母细胞或组织系统系统进行工程化。

[0976] 在特定实施方案中，本文所述的方法用于修饰内源性基因或修饰其表达而不永久性引入到任何外源基因的植物、藻类、真菌、酵母等的基因组中，包括编码CRISPR组分的外源基因，以便避免在植物基因组中存在外源DNA。这可能是令人感兴趣的，因为对非转基因植物的规则要求较不严格。

[0977] 本文所述的方法通常导致生成“改良的植物、藻类、真菌、酵母等”，在这点上它们与野生型植物相比具有一种或多种期望的性状。在特定实施方案中，获得非转基因遗传修饰植物、藻类、真菌、酵母等、部分或细胞，在这点上没有外源性DNA序列并入植物的任何细胞的基因组中。在此类实施方案中，改良的植物、藻类、真菌、酵母等是非转基因的。当仅确保内源性基因的修饰并且在植物、藻类、真菌、酵母等基因组中未引入或维持外源基因时，所得遗传修饰农作物不含有外源基因并且因此可以基本上认为是非转基因的。所述系统用于植物、藻类、真菌、酵母等基因组编辑的不同应用包括但不限于：编辑内源性基因以赋予目标农业性状。赋予农艺性状的示例性基因包括但不限于赋予对害虫或疾病的抗性的基因；涉及植物疾病的基因，诸如WO2013046247中列出的基因；赋予对除草剂、杀真菌剂等的抗性的基因；涉及(非生物)胁迫耐受性的基因。CRISPR-Cas系统的用途的其他方面包括但不限于：产生(雄性)不育植物；增加植物/藻类等的能育性阶段；生成目标农作物的遗传变异；影响果实催熟；增加植物/藻类等的保存期限；减少植物/藻类等中的过敏原；确保增值的性状(例如营养提高)；用于目标内源性基因的筛选方法；生物燃料、脂肪酸、有机酸等的生产。

[0978] 系统可以用于非人类生物体

[0979] 在一方面，本发明提供了一种非人类的真核生物体；优选地是多细胞真核生物体，这些生物体包含根据任何所述实施方案的真核宿主细胞。在其他方面，本发明提供了一种真核生物体；优选地是多细胞真核生物体，这些生物体包含根据任何所述实施方案的真核宿主细胞。在这些方面的一些实施方案中，生物体可以是动物；例如哺乳动物。而且，生物体可以是节肢动物，诸如昆虫。本发明还可以扩展到其他农业应用，例如像农场和生产动物。例如，猪具有许多特征，这些特征使得它们作为生物医学模型是有吸引力的，尤其是在再生医学中。特别地，具有重症联合免疫缺陷(SCID)的猪可以提供用于再生医学、异种移植(还在本文别处论述)以及肿瘤发展的有用模型，并且将有助于开发用于人类SCID患者的疗法。Lee等人(Proc Natl Acad Sci U S A.2014年5月20日；111(20):7260-5)利用一种报告基因指导的转录激活因子样效应核酸酶(TALEN)系统，以高效率生成对体细胞中的重组激活基因(RAG)2的靶向修饰，包括影响两种等位基因的一些修饰。可以将所述系统应用于类似的系统。

[0980] Lee等人的方法(Proc Natl Acad Sci U S A.2014年5月20日；111(20):7260-5)

可以如下类似地应用于本发明。突变的猪是通过靶向修饰胎儿成纤维细胞中的RAG2，随后进行SCNT和胚胎转移而产生的。将编码CRISPR Cas和报告基因的构建体电穿孔到胎儿来源的成纤维细胞中。在48小时后，表达绿色荧光蛋白的转染细胞以估计每孔单个细胞的稀释度分到96孔板的单个孔中。通过扩增侧接任何CRISPR Cas切割位点的基因组DNA片段随后对PCR产物进行测序来筛选RAG2的靶向修饰。在筛选并确保不存在位点外突变之后，将携带RAG2的靶向修饰的细胞用于SCNT。去除极体连同卵母细胞的一部分相邻细胞质(推测含有中期II板)，并且将供体细胞置于卵黄周隙中。然后对重构的胚胎进行电穿孔，以使供体细胞与卵母细胞融合，接着进行化学激活。将激活的胚胎在具有0.5μM Scriptaid(S7817；Sigma-Aldrich)的猪受精卵培养基(Porcine Zygote Medium)3(PZM3)中孵育14-16小时。接着洗涤胚胎以去除Scriptaid并且在PZM3中培养，直到它们转移到代孕猪的输卵管为止。

[0981] 本发明还适用于修饰其他动物诸如牛的SNP。Tan等人(Proc Natl Acad Sci U S A.2013年10月8日；110(41)：16526-16531)使用质粒、rAAV和寡核苷酸模板扩增家畜基因编辑工具包，以包括转录激活因子样(TAL)效应核酸酶(TALEN)-和成簇的规律间隔的短回文重复序列(CRISPR)/Cas9-刺激性同源定向修复(HDR)。根据他们的方法将基因特异性指导RNA序列克隆到Church实验室指导RNA载体(Addgene ID:41824)中(Mali P等人(2013)RNA-Guided Human Genome Engineering via Cas9.Science 339(6121)：823-826)。Cas9核酸酶通过共转染hCas9质粒(Addgene ID:41815)或由RCIScript-hCas9合成的mRNA来提供。此RCIScript-hCas9通过将来自hCas9质粒(涵盖hCas9 cDNA)的XbaI-AgeI片段亚克隆到RCIScript质粒中来构建。

[0982] Heo等人(Stem Cells Dev.2015年2月1日；24(3)：393-402.doi:10.1089/scd.2014.0278.电子版2014年11月3日)报道了在牛基因组中使用牛多能细胞和成簇的规律间隔的短回文重复序列(CRISPR)/Cas9核酸酶的高效基因靶向。首先，Heo等人通过异位表达山中因子(yamanaka factor)并且进行GSK3β和MEK抑制剂(2i)处理来由牛体成纤维细胞生成诱导的多能干细胞(iPSC)。Heo等人观察到，这些牛iPSC在畸胎瘤的基因表达和发育潜力方面高度类似于天然多能干细胞。此外，对于牛NANOG基因座特异的CRISPR-Cas9核酸酶在牛iPSC和胚胎的牛基因组中显示高度有效的编辑。

[0983] Igenity®提供了一种对诸如牛的动物执行并传播经济上重要的经济性状的性状的谱图分析，这些性状诸如胴体组成、胴体质量、母体和繁殖性状以及平均日增重。综合性Igenity®谱图分析以DNA标记物(最常是单核苷酸多态性或SNP)的发现开始。在Igenity®谱图之后的所有标记物是由研究机构的独立科学家发现的，这些研究机构包括大学、研究团体以及政府机构诸如USDA。然后在Igenity®验证群体中分析标记物。Igenity®使用代表各种生产环境和生物类型的多种资源种群，通常与来自牛肉产业的种畜、母犊牛、饲育场和/或包装部门的行业伙伴一起工作，以收集不能普遍获得的表型。牛基因组数据库是广泛可得的，参见例如NAGRP牛基因组协调程序(http://www.animalgenome.org/cattle/maps/db.html)。因此，本发明可以应用于靶向牛SNP。本领域技术人员可以利用用于靶向SNP的以上方案并且将它们应用于牛SNP，例如，如Tan等人或Heo等人所述的。

[0984] Qingjian Zou等人(Journal of Molecular Cell Biology，在2015年10月12日在线先行公布)证明通过靶向狗肌生成抑制蛋白(MSTN)基因(骨骼肌质量的负调节子)的第一外显子可增加狗的肌肉质量。首选，通过将sgRNA靶向MSTN与Cas9载体共转染到犬胚胎成纤

维细胞(CEF)中来验证sgRNA的效率。之后，通过微注射具有正常形态学的胚胎以及Cas9 mRNA和MSTN sgRNA的混合物并且将受精卵自身移植到同一母狗的输卵管来生成MSTN KO狗。与其野生型同窝出生母狗相比，敲除小狗在大腿上显示明显的肌肉表型。这也可以使用本文所提供的系统来进行。

[0985] 家畜-猪

[0986] 在一些实施方案中，家畜中的病毒靶标可以包括例如猪巨噬细胞上的猪CD163。CD163与PRRSv(猪繁殖与呼吸综合征病毒，它是一种动脉炎病毒)的感染(认为是通过病毒细胞侵入)相关联。PRRSv的感染，尤其是对猪肺泡巨噬细胞(可见于肺中)的感染导致先前不能治愈的猪综合征(“神秘猪病”或“蓝耳病”)，使得家猪遭受(包括)繁殖障碍、体重减轻和高死亡率。常常可见机会性感染诸如流行性肺炎、脑膜炎和耳肿胀，这是因为通过巨噬细胞活性丧失会引起免疫缺陷。由于抗生素使用的增加和经济损失(估计每年660百万美元)，这也具有重大的经济和环境影响。

[0987] 如密苏里大学(University of Missouri)的Kristin M Whitworth和Randall Prather博士等人(Nature Biotech 3434，2015年12月07日在线公布)与Genus Plc合作报道的，使用CRISPR-Cas9靶向CD163，当编辑的猪的后代暴露于PRRSv时它们是有抗性的。使在CD163的外显子7中均具有突变的一个雄性起始者和一种雌性起始者繁殖产生后代。雄性起始者具有一个等位基因的外显子7中的11-bp的缺失，这导致结构域5中氨基酸45处的移码突变和错义翻译以及氨基酸64处的后一个提前终止密码子。另一个等位基因具有外显子7中的2-bp添加和在前内含子中的377-bp缺失，预测这引起结构域5的前49个氨基酸的表达，随后是在氨基酸85处的提前终止密码子。母猪在一个等位基因中具有7bp添加，预测该添加在翻译时表达结构域5的前48个氨基酸，随后是在氨基酸70处的提前终止密码子。母猪的另一个等位基因是不可扩增的。预测选定的后代是无效动物(CD163-/-)，即CD163敲除。

[0988] 因此，在一些实施方案中，猪肺泡巨噬细胞可以被CRISPR蛋白靶向。在一些实施方案中，猪CD163可以被CRISPR蛋白靶向。在一些实施方案中，猪CD163可以通过诱导DSB或通过插入或缺失来敲除，例如外显子7的靶向缺失或修饰，包括以上所述的那些缺失或修饰中的一种或多种，或者在该基因的其他区域中，例如外显子5的缺失或修饰。

[0989] 还设想了编辑的猪及其子代，例如CD163敲除猪。这可以是出于家畜、育种或建模目的(即，猪模型)。还提供了包含基因敲除的精液。

[0990] CD163是富含半胱氨酸清道夫受体(SRCR)超家族的成员。基于体外研究，蛋白质的SRCR结构域5是负责启封和释放病毒基因组的结构域。这样，也可以靶向SRCR超家族的其他成员，以便评定对其他病毒的抗性。PRRSV还是哺乳动物动脉炎病毒组的成员，所述病毒组还包括鼠乳酸脱氢酶病毒、猴出血热病毒和马动脉炎病毒。动脉炎病毒享有重要的发病机理特性，包括巨噬细胞向性和引起严重疾病和持续感染二者的能力。因此，可以例如通过猪CD163或其在其他种类中的同系物来提供动脉炎病毒以及特别地鼠乳酸脱氢酶病毒、猴出血热病毒和马动脉炎病毒，并且还提供鼠、猴和马的模型以及敲除。

[0991] 实际上，此方法可以扩展到引起其他家畜疾病且可以传播到人类的病毒或细菌，诸如猪流感病毒(SIV)株，包括丙型流感和称为H1N1、H1N2、H2N1、H3N1、H3N2以及H2N3的甲型流感亚型，以及以上提及的肺炎、脑膜炎和水肿。

[0992] 在一些实施方案中，本文所述的系统可以用于遗传性修饰猪基因组以使一个或多

个猪内源性逆转录病毒(PERV)基因座失活,从而促进猪向人类异种移植的临床应用。参见Yang等人,Science 350(6264):1101-1104(2015),该文献以引用方式整体并入本文。在一些实施方案中,本文所述的系统可以用于产生不包含任何活性猪内源性逆转录病毒(PERV)基因座的遗传修饰的猪。

[0993] 利用系统进行治疗性靶向

[0994] 如将显而易见的,设想的是所述系统可以用于靶向任何目标多核苷酸序列。本发明提供了一种非天然存在的或工程化的组合物、或编码所述组合物的组分的一种或多种多核苷酸、或包含编码所述组合物的组分的一种或多种多核苷酸的载体或递送系统,其用于体内、离体或体外修饰靶细胞,并且所述修饰可以这样一种方式实施:改变细胞,使得一旦被修饰,CRISPR修饰的细胞的子代或细胞系保留改变的表型。修饰的细胞和子代可以是多细胞生物体的一部分,诸如在将CRISPR系统离体或体内应用于所需细胞类型的情况下的植物或动物。CRISPR发明可以是一种治疗性治疗方法。治疗性治疗方法可以包括基因或基因组编辑,或基因疗法。

[0995] 过继细胞疗法

[0996] 本发明还考虑使用本文所述的系统来修饰用于过继疗法的细胞。因此,本发明的方面涉及过继性转移对于选定抗原诸如肿瘤相关抗原特异的免疫系统细胞诸如T细胞(参见Maus等人,2014,Adoptive Immunotherapy for Cancer or Viruses,Annual Review of Immunology,第32卷:189-225;Rosenberg和Restifo,2015,Adoptive cell transfer as personalized immunotherapy for human cancer,Science第348卷,第6230期,第62-68页;以及Restifo等人,2015,Adoptive immunotherapy for cancer:harnessing the T cell response.Nat.Rev.Immunol.12(4):269-281;以及Jenson和Riddell,2014,Design and implementation of adoptive therapy with chimeric antigen receptor-modified T cells.Immunol Rev.257(1):127-144)。可以例如采用各种策略来通过改变T细胞受体(TCR)的特异性,例如通过引入具有选定的肽特异性的新TCRα和β链来遗传性修饰T细胞(参见美国专利号8,697,854;PCT专利公布:WO2003020763、WO2004033685、WO2004044004、WO2005114215、WO2006000830、WO2008038002、WO2008039818、WO2004074322、WO2005113595、WO2006125962、WO2013166321、WO2013039889、WO2014018863、WO2014083173;美国专利号8,088,379)。

[0997] 作为TCR修饰的替代方案或者除TCR修饰之外,可以使用嵌合抗原受体(CAR),以便生成对于选定靶标诸如恶性肿瘤细胞特异的免疫反应细胞诸如T细胞,其中已经描述了多种多样的受体嵌合构建体(参见美国专利号5,843,728;5,851,828;5,912,170;6,004,811;6,284,240;6,392,013;6,410,014;6,753,162;8,211,422;以及PCT公布WO9215322)。替代CAR构建体可以被表征为属于连续世代。第一代CAR典型地由对于抗原特异的抗体单链可变片段组成,例如包括连接至特异性抗体的VH的VL,通过柔性接头连接,例如通过CD8α铰链结构域和CD8α跨膜结构域,连接至CD3ζ或FcRγ的跨膜和细胞内信号传导结构域(scFv-CD3ζ或scFv-FcRγ;参见美国专利号7,741,465;美国专利号5,912,172;美国专利号5,906,936)。第二代CAR并入一种或多种共刺激分子的细胞内结构域,诸如内结构域内的CD28、OX40(CD134)或4-1BB(CD137)(例如scFv-CD28/OX40/4-1BB-CD3ζ;参见美国专利号8,911,993;8,916,381;8,975,071;9,101,584;9,102,760;9,102,761)。第三代CAR包括共刺激内

结构域的组合，诸如CD3ζ-链、CD97、GDI la-CD18、CD2、ICOS、CD27、CD154、CDS、OX40、4-1BB或CD28信号传导结构域(例如scFv-CD28-4-1BB-CD3ζ或scFv-CD28-OX40-CD3ζ；参见美国专利号8,906,682；美国专利号8,399,645；美国专利号5,686,281；PCT公布号WO2014134165；PCT公布号WO2012079000)。或者，可以通过在抗原特异性T细胞中表达CAR来协调共刺激，所述抗原特异性T细胞被选择以便在伴随共刺激的情况下例如由专职抗原呈递细胞上的抗原接合其天然αβTCR之后得以激活并扩增。此外，另外的工程化受体可以被提供在免疫反应细胞上，例如以提高T细胞攻击的靶向并且/或者最小化副作用。

[0998] 可以使用替代技术转化免疫反应靶细胞，所述替代技术诸如原生质体融合、脂转染、转染或电穿孔。可以使用多种多样的载体，诸如逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体、质粒或转座子(诸如睡美人易位子)(参见美国专利号6,489,458；7,148,203；7,160,682；7,985,739；8,227,432)，这些载体可以用于例如使用通过CD3ζ和CD28或CD137传导信号的第2代抗原特异性CAR来引入CAR。病毒载体可以例如包括基于HIV、SV40、EBV、HSV或BPV的载体。

[0999] 被靶向转化的细胞可以例如包括T细胞、自然杀伤细胞(NK)、细胞毒性T淋巴细胞(CTL)、调节T细胞、人类胚胎干细胞、肿瘤浸润淋巴细胞(TIL)或淋巴样细胞可以由其分化的多能干细胞。表达所需CAR的T细胞可以例如通过与γ-辐射激活和增殖细胞(AaPC)共培养来选择，这些激活和增殖细胞共表达癌症抗原和共刺激分子。工程化CAR T细胞可以例如通过在可溶性因子诸如IL-2和IL-21存在下在AaPC上共培养来扩增。可以例如执行此扩增以提供记忆CAR+T细胞(这些细胞可以例如通过非酶数字阵列和/或多板块(multi-panel)流式细胞术来测定)。以这种方式，可以提供具有针对携带抗原的肿瘤的特异性细胞毒性活性的CAR T细胞(任选地与所需趋化因子诸如干扰素-γ的产生相结合)。这种类型的CAR T细胞例如可以用于动物模型中，例如以威慑肿瘤异种移植物。

[1000] 诸如上述方法的方法可以适于提供例如通过施用有效量的免疫反应细胞来治疗患有诸如瘤形成的疾病的患者并且/或者增加所述患者的存活的方法，所述免疫反应细胞包含结合选定抗原的抗原识别受体，其中所述结合激活免疫反应细胞，从而治疗或预防所述疾病(诸如瘤形成、病原体感染、自身免疫病症或同种异体移植反应)。在具有或不具有淋巴细胞耗尽过程的情况下，例如在使用环磷酰胺的情况下，CAR T细胞疗法的给药可以例如涉及施用106至109个细胞/kg。

[1001] 在一个实施方案中，可以向经受免疫阻遏治疗的患者施用所述治疗。细胞或细胞群体可以被制成由于编码至少一种免疫阻遏剂的受体的基因的失活而抵抗此免疫阻遏剂。不受理论的束缚，免疫阻遏治疗应有助于在患者内选择和扩增根据本发明的免疫反应细胞或T细胞。

[1002] 可以采用任何常规方式执行根据本发明的细胞或细胞群体的施用，这些方式包括通过雾化吸入、注射、摄取、输血、植入或移植。可以向患者皮下、真皮内、瘤内、节点内、髓内、肌内、通过静脉内注射或淋巴管内注射、或者腹膜内施用所述细胞或细胞群体。在一个实施方案中，本发明的细胞组合物优选地通过静脉内注射来施用。

[1003] 所述细胞或细胞群体的施用可以由104-109个细胞/kg体重、优选地105至106个细胞/kg体重(包括这些范围内细胞数目的所有整数值)的施用组成。在具有或不具有淋巴细胞耗尽过程的情况下，例如在使用环磷酰胺的情况下，CAR T细胞疗法的给药可以例如涉及

施用106至109个细胞/kg。细胞或细胞群体可以按一个或多个剂量施用。在另一个实施方案中,有效量的细胞作为单一剂量施用。在另一个实施方案中,有效量的细胞作为多于一个剂量在一段时间内施用。施用的时间在管理医师的判断内并且取决于患者的临床状况。细胞或细胞群体可以从任何来源诸如血库或供体获得。虽然个体需要不同,但是针对特定疾病或疾患的给定细胞类型的有效量的最佳范围的确定在本领域技术人员技能之内。有效量是指提供治疗或预防益处的量。施用的剂量将取决于接受者的年龄、健康和体重、同期治疗(如果有的话)的类型、治疗的频率以及所需的作用性质。

[1004] 在另一个实施方案中,肠胃外施用有效量的细胞或包含那些细胞的组合物。施用可以是静脉内施用。可以通过在肿瘤内注射直接进行施用。

[1005] 为了防止可能的不良反应,可以用转基因安全开关装备工程化的免疫反应细胞,所述转基因安全开关呈致使这些细胞易于暴露于特定信号的转基因形式。例如,单纯疱疹病毒胸苷激酶(TK)可以这种方式使用,例如通过在干细胞移植之后引入到用作供体淋巴细胞输注的同种异体T淋巴细胞中(Greco等人,Improving the safety of cell therapy with the TK-suicide gene.Front.Pharmacol.2015;6:95)。在此类细胞中,施用核苷前药诸如更昔洛韦或阿昔洛韦引起细胞死亡。替代安全开关构建体包括诱导型半胱天冬酶9,其例如通过施用使两个非功能性icasp9分子连接在一起形成活性酶的小分子二聚体来触发。已经描述了用于实施细胞增殖控制的多种多样的替代方法(参见美国专利公布号20130071414;PCT专利公布WO2011146862;PCT专利公布WO2014011987;PCT专利公布WO2013040371;Zhou等人BLOOD,2014,123/25:3895-3905;Di Stasi等人,The New England Journal of Medicine 2011;365:1673-1683;Sadelain M,The New England Journal of Medicine 2011;365:1735-173;Ramos等人,Stem Cells 28(6):1107-15(2010))。

[1006] 在过继疗法的另一个改进方案中,如本文所述的AD官能化的CRISPR-Cas系统进行的基因组编辑可以用于使免疫反应细胞适于替代实施方式,例如提供编辑的CAR T细胞(参见Poirot等人,2015,Multiplex genome edited T-cell manufacturing platform for"off-the-shelf"adoptive T-cell immunotherapies,Cancer Res 75(18):3853)。例如,免疫反应细胞可以被编辑为缺失一些或所有类别的HLA II型和/或I型分子的表达,或者敲除可以抑制所需免疫反应的选定基因诸如PD1基因。

[1007] 可以使用如本文所述的系统来编辑细胞。可以通过本文所述的任何方法将系统递送至免疫细胞。在优选的实施方案中,将细胞离体编辑并转移至有需要的受试者。可以编辑免疫反应细胞、CAR-T细胞或用于过继细胞转移的任何细胞。可以进行编辑以消除潜在的同种异体反应性T-细胞受体(TCR)、破坏化学治疗剂的靶标、阻断免疫校验点、激活T细胞并且/或者增加功能耗竭或功能障碍的CD8+T-细胞的分化和/或增殖(参见PCT专利公布:WO2013176915、WO2014059173、WO2014172606、WO2014184744和WO2014191128)。编辑可以导致基因失活。

[1008] T细胞受体(TCR)是响应于抗原呈递而参与激活T细胞的细胞表面受体。TCR通常是由组装形成异源二聚体的两条链α和β形成,并且与CD3-转导亚基缔合形成存在于细胞表面的T细胞受体复合物。TCR的每条α和β链由免疫球蛋白样N端可变区(V)和恒定(C)区、疏水性跨膜结构域以及短胞质区组成。如对于免疫球蛋白分子,α和β链的可变区是通过V(D)J重组而生成,从而在T细胞群体内形成多种抗原特异性。然而,与识别完整抗原的免疫球蛋白相

比,T细胞通过与MHC分子缔合的加工肽片段来激活,从而将额外维度引入到由T细胞进行的抗原识别中,这被称为MHC限制。通过T细胞受体识别供体与受体之间的MHC差异导致T细胞增殖和移植物抗宿主疾病(GVHD)的潜在发展。TCRα或TCRβ的失活可以导致TCR从T细胞表面消除,从而阻止了同种抗原的识别并因此产生GVHD。然而,TCR破坏通常导致CD3信号传导组分的消除并且改变另外的T细胞扩增的方式。

[1009] 同种异体细胞被宿主免疫系统迅速排斥。已证明存在于非辐射血液产品中的同种异体淋巴细胞将持续不超过5至6天(Boni,Muranski等人2008Blood1;112(12):4746-54)。因此,为了防止同种异体细胞排斥,通常必须在一定程度上阻遏宿主的免疫系统。然而,在过继细胞转移的情况下,使用免疫阻遏药物也对引入的治疗性T细胞具有有害作用。因此,为了在这些情况下有效使用过继免疫治疗方法,引入的细胞将需要抵抗免疫阻遏治疗。因此,在特定实施方案中,本发明还包括优选地通过使编码免疫阻遏剂的靶标的至少一种基因失活来修饰T细胞以使其抵抗免疫阻遏剂的步骤。免疫阻遏剂是一种通过若干作用机制之一阻遏免疫功能的物质。免疫阻遏剂可以是但不限于钙调磷酸酶抑制剂、雷帕霉素的靶标、白介素-2受体α-链阻断剂、肌苷单磷酸脱氢酶的抑制剂、二氢叶酸还原酶的抑制剂、皮质类固醇或免疫阻遏抗代谢物。本发明允许通过使T细胞中的免疫阻遏剂的靶标失活来对用于免疫疗法的T细胞赋予免疫阻遏抗性。作为非限制性实例,免疫阻遏剂的靶标可以是免疫阻遏剂的受体,诸如:CD52、糖皮质激素受体(GR)、FKBP家族基因成员以及亲环蛋白家族基因成员。

[1010] 免疫检验点是减慢或停止免疫反应并且防止因免疫细胞的不受控活性所致的过度组织损害的抑制性途径。在某些实施方案中,所靶向的免疫检查点是程序性死亡-1(PD-1或CD279)基因(PDCD1)。在其他实施方案中,所靶向的免疫检查点是细胞毒性T-淋巴细胞相关抗原(CTLA-4)。在其他实施方案中,所靶向的免疫检查点是CD28和CTLA4 Ig超家族的另一个成员,诸如BTLA、LAG3、ICOS、PDL1或KIR。在另外的其他实施方案中,所靶向的免疫检查点是TNFR超家族的成员,诸如CD40、OX40、CD137、GITR、CD27或TIM-3。

[1011] 其他免疫检验点包括含有Src同源2结构域的蛋白酪氨酸磷酸酶1(SHP-1)(Watson HA等人,SHP-1:the next checkpoint target for cancer immunotherapy?Biochem Soc Trans.2016年4月15日;44(2):356-62)。SHP-1是一种广泛表达的抑制性蛋白酪氨酸磷酸酶(PTP)。在T细胞中,它是抗原依赖性激活和增殖的负调节子。它是一种胞质蛋白,并且因此不适于抗体介导的疗法,但是它在激活和增殖中的作用使得它成为过继转移策略中用于遗传操纵的有吸引力的靶标,诸如嵌合抗原受体(CAR)T细胞。免疫检查点还可以包括具有Ig和ITIM结构域的T细胞免疫受体(TIGIT/Vstm3/WUCAM/VSIG9)和VISTA(Le Mercier I等人,(2015)Beyond CTLA-4and PD-1,the generation Z of negative checkpoint regulators.Front.Immunol.6:418)。

[1012] WO2014172606涉及使用MT1和/或MT1抑制剂来增加耗竭的CD8+T细胞的增殖和/或活性并且减少CD8+T细胞耗竭(例如,减少功能耗竭或不反应的CD8+免疫细胞)。在某些实施方案中,在过继性转移的T细胞中通过基因编辑靶向金属硫蛋白。

[1013] 在某些实施方案中,基因编辑的靶标可以是涉及免疫检验点蛋白的表达的至少一个靶向基因座。此类靶标可以包括但不限于CTLA4、PPP2CA、PPP2CB、PTPN6、PTPN22、PDCD1、ICOS(CD278)、PDL1、KIR、LAG3、HAVCR2、BTLA、CD160、TIGIT、CD96、CRTAM、LAIR1、SIGLEC7、

SIGLEC9、CD244(2B4)、TNFRSF10B、TNFRSF10A、CASP8、CASP10、CASP3、CASP6、CASP7、FADD、FAS、TGFBRII、TGFRBRI、SMAD2、SMAD3、SMAD4、SMAD10、SKI、SKIL、TGIF1、IL10RA、IL10RB、HMOX2、IL6R、IL6ST、EIF2AK4、CSK、PAG1、SIT1、FOXP3、PRDM1、BATF、VISTA、GUCY1A2、GUCY1A3、GUCY1B2、GUCY1B3、MT1、MT2、CD40、OX40、CD137、GITR、CD27、SHP-1或TIM-3。在优选的实施方案中,靶向涉及PD-1或CTLA-4基因的表达的基因座。在其他优选的实施方案中,靶向基因的组合,诸如但不限于PD-1和TIGIT。

[1014] 在其他实施方案中,编辑至少两种基因。基因对可以包括但不限于PD1和TCRα、PD1和TCRβ、CTLA-4和TCRα、CTLA-4和TCRβ、LAG3和TCRα、LAG3和TCRβ、Tim3和TCRα、Tim3和TCRβ、BTLA和TCRα、BTLA和TCRβ、BY55和TCRα、BY55和TCRβ、TIGIT和TCRα、TIGIT和TCRβ、B7H5和TCRα、B7H5和TCRβ、LAIR1和TCRα、LAIR1和TCRβ、SIGLEC10和TCRα、SIGLEC10和TCRβ、2B4和TCRα、2B4和TCRβ。

[1015] 无论是在T细胞的遗传修饰之前还是之后,T细胞都可以通常使用如例如以下文献所述的方法来激活并扩增:美国专利6,352,694;6,534,055;6,905,680;5,858,358;6,887,466;6,905,681;7,144,575;7,232,566;7,175,843;5,883,223;6,905,874;6,797,514;6,867,041;以及7,572,631。T细胞可以在体外或体内扩增。

[1016] 本发明的实践采用免疫学、生物化学、化学、分子生物学、微生物学、细胞生物学、基因组学以及重组DNA领域中已知的技术,这些技术在本领域技能范围之内。参见MOLECULAR CLONING:A LABORATORY MANUAL,第2版(1989)(Sambrook、Fritsch和Maniatis);MOLECULAR CLONING:A LABORATORY MANUAL,第4版(2012)(Green和Sambrook);CURRENT PROTOCOLS IN MOLECULAR BIOLOGY(1987)(F.M.Ausubel等人编辑);METHODS IN ENZYMOLOGY系列(Academic Press,Inc.);PCR 2:A PRACTICAL APPROACH(1995)(M.J.MacPherson、B.D.Hames和G.R.Taylor编辑);ANTIBODIES,A LABORATORY MANUAL(1988)(Harlow和Lane编辑);ANTIBODIES A LABORATORY MANUAL,第2版(2013)(E.A.Greenfield编辑);以及ANIMAL CELL CULTURE(1987)(R.I.Freshney编辑)。

[1017] 疾病相关突变和病原性SNP的校正

[1018] 在一方面,本文所述的发明提供了用于修饰靶基因座处的腺苷残基的方法,目的是补救并且/或者预防由或可能由G至A或C至T点突变或病原性单核苷酸多态性(SNP)引起的疾病状况。

[1019] 影响脑和中枢神经系统的疾病

[1020] 与影响脑和中枢神经系统的各种疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开,这些疾病包括但不限于阿尔茨海默病、帕金森病、自闭症、肌萎缩性侧索硬化症(ALS)、精神分裂症、肾上腺脑白质营养不良、艾卡尔迪-Goutieres综合征(Aicardi Goutieres syndrome)、法布里病(Fabry disease)、莱施-奈恩综合征(Lesch-Nyhan syndrome)和门克斯病(Menkes Disease)。因此,本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1021] 阿尔茨海默病

[1022] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与阿尔茨海默病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在

于选自PSEN1、PSEN2和APP的至少一种基因中，至少包括以下各项：

[1023] NM_000021.3(PSEN1):c.796G>A(p.Gly266Ser)

[1024] NM_000484.3(APP):c.2017G>A(p.Ala673Thr)

[1025] NM_000484.3(APP):c.2149G>A(p.Val717Ile)

[1026] NM_000484.3(APP):c.2137G>A(p.Ala713Thr)

[1027] NM_000484.3(APP):c.2143G>A(p.Val715Met)

[1028] NM_000484.3(APP):c.2141C>T(p.Thr714Ile)

[1029] NM_000021.3(PSEN1):c.438G>A(p.Met146Ile)

[1030] NM_000021.3(PSEN1):c.1229G>A(p.Cys410Tyr)

[1031] NM_000021.3(PSEN1):c.487C>T(p.His163Tyr)

[1032] NM_000021.3(PSEN1):c.799C>T(p.Pro267Ser)

[1033] NM_000021.3(PSEN1):c.236C>T(p.Ala79Val)

[1034] NM_000021.3(PSEN1):c.509C>T(p.Ser170Phe)

[1035] NM_000447.2(PSEN2):c.1289C>T(p.Thr430Met)

[1036] NM_000447.2(PSEN2):c.717G>A(p.Met239Ile)

[1037] NM_000447.2(PSEN2):c.254C>T(p.Ala85Val)

[1038] NM_000021.3(PSEN1):c.806G>A(p.Arg269His)

[1039] NM_000484.3(APP):c.2018C>T(p.Ala673Val)。

[1040] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自PSEN1、PSEN2和APP的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防阿尔茨海默病的方法。

[1041] 帕金森病

[1042] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与帕金森病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，在一些实施方案中，病原性突变/SNP存在于选自SNCA、PLA2G6、FBXO7、VPS35、EIF4G1、DNAJC6、PRKN、SYNJ1、CHCHD2、PINK1、PARK7、LRRK2、ATP13A2和GBA的至少一种基因中，至少包括以下各项：

[1043] NM_000345.3(SNCA):c.157G>A(p.Ala53Thr)

[1044] NM_000345.3(SNCA):c.152G>A(p.Gly51Asp)

[1045] NM_003560.3(PLA2G6):c.2222G>A(p.Arg741Gln)

[1046] NM_003560.3(PLA2G6):c.2239C>T(p.Arg747Trp)

[1047] NM_003560.3(PLA2G6):c.1904G>A(p.Arg635Gln)

[1048] NM_003560.3(PLA2G6):c.1354C>T(p.Gln452Ter)

[1049] NM_012179.3(FBXO7):c.1492C>T(p.Arg498Ter)

[1050] NM_012179.3(FBXO7):c.65C>T(p.Thr22Met)

[1051] NM_018206.5(VPS35):c.1858G>A(p.Asp620Asn)

[1052] NM_198241.2(EIF4G1):c.3614G>A(p.Arg1205His)

[1053] NM_198241.2(EIF4G1):c.1505C>T(p.Ala502Val)

[1054] NM_001256865.1(DNAJC6):c.2200C>T(p.Gln734Ter)

[1055] NM_001256865.1(DNAJC6):c.2326C>T(p.Gln776Ter)

[1056] NM_004562.2(PRKN):c.931C>T(p.Gln311Ter)

[1057] NM_004562.2(PRKN):c.1358G>A(p.Trp453Ter)

[1058] NM_004562.2(PRKN):c.635G>A(p.Cys212Tyr)

[1059] NM_203446.2(SYNJ1):c.773G>A(p.Arg258Gln)

[1060] NM_001320327.1(CHCHD2):c.182C>T(p.Thr61Ile)

[1061] NM_001320327.1(CHCHD2):c.434G>A(p.Arg145Gln)

[1062] NM_001320327.1(CHCHD2):c.300+5G>A

[1063] NM_032409.2(PINK1):c.926G>A(p.Gly309Asp)

[1064] NM_032409.2(PINK1):c.1311G>A(p.Trp437Ter)

[1065] NM_032409.2(PINK1):c.736C>T(p.Arg246Ter)

[1066] NM_032409.2(PINK1):c.836G>A(p.Arg279His)

[1067] NM_032409.2(PINK1):c.938C>T(p.Thr313Met)

[1068] NM_032409.2(PINK1):c.1366C>T(p.Gln456Ter)

[1069] NM_007262.4(PARK7):c.78G>A(p.Met26Ile)

[1070] NM_198578.3(LRRK2):c.4321C>T(p.Arg1441Cys)

[1071] NM_198578.3(LRRK2):c.4322G>A(p.Arg1441His)

[1072] NM_198578.3(LRRK2):c.1256C>T(p.Ala419Val)

[1073] NM_198578.3(LRRK2):c.6055G>A(p.Gly2019Ser)

[1074] NM_022089.3(ATP13A2):c.1306+5G>A

[1075] NM_022089.3(ATP13A2):c.2629G>A(p.Gly877Arg)

[1076] NM_022089.3(ATP13A2):c.490C>T(p.Arg164Trp)

[1077] NM_001005741.2(GBA):c.1444G>A(p.Asp482Asn)

[1078] m.15950G>A。

[1079] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自SNCA、PLA2G6、FBXO7、VPS35、EIF4G1、DNAJC6、PRKN、SYNJ1、CHCHD2、PINK1、PARK7、LRRK2、ATP13A2和GBA的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防帕金森病的方法。

[1080] 自闭症

[1081] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与自闭症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自MECP2、NLGN3、SLC9A9、EHMT1、CHD8、NLGN4X、GSPT2和PTEN的至少一种基因中，至少包括以下各项：

[1082] NM_001110792.1(MECP2):c.916C>T(p.Arg306Ter)

[1083] NM_004992.3(MECP2):c.473C>T(p.Thr158Met)

[1084] NM_018977.3(NLGN3):c.1351C>T(p.Arg451Cys)

[1085] NM_173653.3(SLC9A9):c.1267C>T(p.Arg423Ter)

[1086] NM_024757.4(EHMT1):c.3413G>A(p.Trp1138Ter)

[1087] NM_020920.3(CHD8):c.2875C>T(p.Gln959Ter)

[1088] NM_020920.3(CHD8):c.3172C>T(p.Arg1058Ter)

[1089] NM_181332.2(NLGN4X):c.301C>T(p.Arg101Ter)

[1090] NM_018094.4(GSPT2):c.1021G>A(p.Val341Ile)

[1091] NM_000314.6(PTEN):c.392C>T(p.Thr131Ile)

[1092] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自MECP2、NLGN3、SLC9A9、EHMT1、CHD8、NLGN4X、GSPT2和PTEN的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防自闭症的方法。

[1093] 肌萎缩性侧索硬化症(ALS)

[1094] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与ALS相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于选自SOD1、VCP、UBQLN2、ERBB4、HNRNPA1、TUBA4A、SOD1、TARDBP、FIG4、OPTN、SETX、SPG11、FUS、VAPB、ANG、CHCHD10、SQSTM1和TBK1的至少一种基因中,至少包括以下各项:

[1095] NM_000454.4(SOD1):c.289G>A(p.Asp97Asn)

[1096] NM_007126.3(VCP):c.1774G>A(p.Asp592Asn)

[1097] NM_007126.3(VCP):c.464G>A(p.Arg155His)

[1098] NM_007126.3(VCP):c.572G>A(p.Arg191Gln)

[1099] NM_013444.3(UBQLN2):c.1489C>T(p.Pro497Ser)

[1100] NM_013444.3(UBQLN2):c.1525C>T(p.Pro509Ser)

[1101] NM_013444.3(UBQLN2):c.1573C>T(p.Pro525Ser)

[1102] NM_013444.3(UBQLN2):c.1490C>T(p.Pro497Leu)

[1103] NM_005235.2(ERBB4):c.2780G>A(p.Arg927Gln)

[1104] NM_005235.2(ERBB4):c.3823C>T(p.Arg1275Trp)

[1105] NM_031157.3(HNRNPA1):c.940G>A(p.Asp314Asn)

[1106] NM_006000.2(TUBA4A):c.643C>T(p.Arg215Cys)

[1107] NM_006000.2(TUBA4A):c.958C>T(p.Arg320Cys)

[1108] NM_006000.2(TUBA4A):c.959G>A(p.Arg320His)

[1109] NM_006000.2(TUBA4A):c.1220G>A(p.Trp407Ter)

[1110] NM_006000.2(TUBA4A):c.1147G>A(p.Ala383Thr)

[1111] NM_000454.4(SOD1):c.112G>A(p.Gly38Arg)

[1112] NM_000454.4(SOD1):c.124G>A(p.Gly42Ser)

[1113] NM_000454.4(SOD1):c.125G>A(p.Gly42Asp)

[1114] NM_000454.4(SOD1):c.14C>T(p.Ala5Val)

[1115] NM_000454.4(SOD1):c.13G>A(p.Ala5Thr)

[1116] NM_000454.4(SOD1):c.436G>A(p.Ala146Thr)

[1117] NM_000454.4(SOD1):c.64G>A(p.Glu22Lys)

[1118] NM_000454.4(SOD1):c.404G>A(p.Ser135Asn)

[1119] NM_000454.4(SOD1):c.49G>A(p.Gly17Ser)

[1120] NM_000454.4(SOD1):c.217G>A(p.Gly73Ser)

[1121] NM_007375.3(TARDBP):c.892G>A(p.Gly298Ser)

[1122] NM_007375.3(TARDBP):c.943G>A(p.Ala315Thr)

[1123] NM_007375.3(TARDBP):c.883G>A(p.Gly295Ser)

[1124] NM_007375.3(TARDBP):c.*697G>A

[1125] NM_007375.3(TARDBP):c.1144G>A(p.Ala382Thr)

[1126] NM_007375.3(TARDBP):c.859G>A(p.Gly287Ser)

[1127] NM_014845.5(FIG4):c.547C>T(p.Arg183Ter)

[1128] NM_001008211.1(OPTN):c.1192C>T(p.Gln398Ter)

[1129] NM_015046.5(SETX):c.6407G>A(p.Arg2136His)

[1130] NM_015046.5(SETX):c.8C>T(p.Thr3Ile)

[1131] NM_025137.3(SPG11):c.118C>T(p.Gln40Ter)

[1132] NM_025137.3(SPG11):c.267G>A(p.Trp89Ter)

[1133] NM_025137.3(SPG11):c.5974C>T(p.Arg1992Ter)

[1134] NM_004960.3(FUS):c.1553G>A(p.Arg518Lys)

[1135] NM_004960.3(FUS):c.1561C>T(p.Arg521Cys)

[1136] NM_004960.3(FUS):c.1562G>A(p.Arg521His)

[1137] NM_004960.3(FUS):c.1520G>A(p.Gly507Asp)

[1138] NM_004960.3(FUS):c.1483C>T(p.Arg495Ter)

[1139] NM_004960.3(FUS):c.616G>A(p.Gly206Ser)

[1140] NM_004960.3(FUS):c.646C>T(p.Arg216Cys)

[1141] NM_004738.4(VAPB):c.166C>T(p.Pro56Ser)

[1142] NM_004738.4(VAPB):c.137C>T(p.Thr46Ile)

[1143] NM_001145.4(ANG):c.164G>A(p.Arg55Lys)

[1144] NM_001145.4(ANG):c.155G>A(p.Ser52Asn)

[1145] NM_001145.4(ANG):c.407C>T(p.Pro136Leu)

[1146] NM_001145.4(ANG):c.409G>A(p.Val137Ile)

[1147] NM_001301339.1(CHCHD10):c.239C>T(p.Pro80Leu)

[1148] NM_001301339.1(CHCHD10):c.176C>T(p.Ser59Leu)

[1149] NM_001142298.1(SQSTM1):c.-47-1924C>T

[1150] NM_003900.4(SQSTM1):c.1160C>T(p.Pro387Leu)

[1151] NM_003900.4(SQSTM1):c.1175C>T(p.Pro392Leu)

[1152] NM_013254.3(TBK1):c.1340+1G>A

[1153] NM_013254.3(TBK1):c.2086G>A(p.Glu696Lys)

[1154] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自SOD1、VCP、UBQLN2、ERBB4、HNRNPA1、TUBA4A、SOD1、TARDBP、FIG4、OPTN、SETX、SPG11、FUS、VAPB、ANG、CHCHD10、SQSTM1和TBK1的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防ALS的方法。

[1155] 精神分裂症

[1156] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与精神分裂症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自PRODH、SETD1A和SHANK3的至少一种基因中，至少包括以下各项：

[1157] NM_016335.4(PRODH):c.1292G>A(p.Arg431His)

[1158] NM_016335.4(PRODH):c.1397C>T(p.Thr466Met)

[1159] NM_014712.2(SETD1A):c.2209C>T(p.Gln737Ter)

[1160] NM_033517.1(SHANK3):c.3349C>T(p.Arg1117Ter)

[1161] NM_033517.1(SHANK3):c.1606C>T(p.Arg536Trp)

[1162] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自PRODH、SETD1A和SHANK3的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防精神分裂症的方法。

[1163] 肾上腺脑白质营养不良

[1164] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与肾上腺脑白质营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于ABCD1基因中，至少包括以下各项：

[1165] NM_000033.3(ABCD1):c.421G>A(p.Ala141Thr)

[1166] NM_000033.3(ABCD1):c.796G>A(p.Gly266Arg)

[1167] NM_000033.3(ABCD1):c.1252C>T(p.Arg418Trp)

[1168] NM_000033.3(ABCD1):c.1552C>T(p.Arg518Trp)

[1169] NM_000033.3(ABCD1):c.1850G>A(p.Arg617His)

[1170] NM_000033.3(ABCD1):c.1396C>T(p.Gln466Ter)

[1171] NM_000033.3(ABCD1):c.1553G>A(p.Arg518Gln)

[1172] NM_000033.3(ABCD1):c.1679C>T(p.Pro560Leu)

[1173] NM_000033.3(ABCD1):c.1771C>T(p.Arg591Trp)

[1174] NM_000033.3(ABCD1):c.1802G>A(p.Trp601Ter)

[1175] NM_000033.3(ABCD1):c.346G>A(p.Gly116Arg)

[1176] NM_000033.3(ABCD1):c.406C>T(p.Gln136Ter)

[1177] NM_000033.3(ABCD1):c.1661G>A(p.Arg554His)

[1178] NM_000033.3(ABCD1):c.1825G>A(p.Glu609Lys)

[1179] NM_000033.3(ABCD1):c.1288C>T(p.Gln430Ter)

[1180] NM_000033.3(ABCD1):c.1781-1G>A

[1181] NM_000033.3(ABCD1):c.529C>T(p.Gln177Ter)

[1182] NM_000033.3(ABCD1):c.1866-10G>A

[1183] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地至少存在于ABCD1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防肾上腺脑白质营养不良的方法。

[1184] 艾卡尔迪-Goutieres综合征

[1185] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与艾卡尔迪-Goutieres综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自TREX1、RNASEH2C、ADAR和IFIH1的至少一种基因中，至少包括以下各项：

[1186] NM_016381.5(TREX1):c.794G>A(p.Trp265Ter)

[1187] NM_033629.4(TREX1):c.52G>A(p.Asp18Asn)

[1188] NM_033629.4(TREX1):c.490C>T(p.Arg164Ter)

[1189] NM_032193.3(RNASEH2C):c.205C>T(p.Arg69Trp)

[1190] NM_001111.4(ADAR):c.3019G>A(p.Gly1007Arg)

[1191] NM_022168.3(IFIH1):c.2336G>A(p.Arg779His)

[1192] NM_022168.3(IFIH1):c.2335C>T(p.Arg779Cys)

[1193] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自TREX1、RNASEH2C、ADAR和IFIH1的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防艾卡尔迪-Goutieres综合征的方法。

[1194] 法布里病

[1195] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与法布里病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于GLA基因中，至少包括以下各项：

[1196] NM_000169.2(GLA):c.1024C>T(p.Arg342Ter)

[1197] NM_000169.2(GLA):c.1066C>T(p.Arg356Trp)

[1198] NM_000169.2(GLA):c.1025G>A(p.Arg342Gln)

[1199] NM_000169.2(GLA):c.281G>A(p.Cys94Tyr)

[1200] NM_000169.2(GLA):c.677G>A(p.Trp226Ter)

[1201] NM_000169.2(GLA):c.734G>A(p.Trp245Ter)

[1202] NM_000169.2(GLA):c.748C>T(p.Gln250Ter)

[1203] NM_000169.2(GLA):c.658C>T(p.Arg220Ter)

[1204] NM_000169.2(GLA):c.730G>A(p.Asp244Asn)

[1205] NM_000169.2(GLA):c.369+1G>A

[1206] NM_000169.2(GLA):c.335G>A(p.Arg112His)

[1207] NM_000169.2(GLA):c.485G>A(p.Trp162Ter)

[1208] NM_000169.2(GLA):c.661C>T(p.Gln221Ter)

[1209] NM_000169.2(GLA):c.916C>T(p.Gln306Ter)

[1210] NM_000169.2(GLA):c.1072G>A(p.Glu358Lys)

[1211] NM_000169.2(GLA):c.1087C>T(p.Arg363Cys)

[1212] NM_000169.2(GLA):c.1088G>A(p.Arg363His)

[1213] NM_000169.2(GLA):c.605G>A(p.Cys202Tyr)

[1214] NM_000169.2(GLA):c.830G>A(p.Trp277Ter)

[1215] NM_000169.2(GLA):c.979C>T(p.Gln327Ter)

[1216] NM_000169.2(GLA):c.422C>T(p.Thr141Ile)

[1217] NM_000169.2(GLA):c.285G>A(p.Trp95Ter)

[1218] NM_000169.2(GLA):c.735G>A(p.Trp245Ter)

[1219] NM_000169.2(GLA):c.639+919G>A

[1220] NM_000169.2(GLA):c.680G>A(p.Arg227Gln)

[1221] NM_000169.2(GLA):c.679C>T(p.Arg227Ter)

[1222] NM_000169.2(GLA):c.242G>A(p.Trp81Ter)

[1223] NM_000169.2(GLA):c.901C>T(p.Arg301Ter)

[1224] NM_000169.2(GLA):c.974G>A(p.Gly325Asp)

[1225] NM_000169.2(GLA):c.847C>T(p.Gln283Ter)

[1226] NM_000169.2(GLA):c.469C>T(p.Gln157Ter)

[1227] NM_000169.2(GLA):c.1118G>A(p.Gly373Asp)

[1228] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地至少存在于GLA基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防法布里病的方法。

[1229] 莱施-奈恩综合征

[1230] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与莱施-奈恩综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于HPRT1基因中,至少包括以下各项:

[1231] NM_000194.2(HPRT1):c.151C>T(p.Arg51Ter)

[1232] NM_000194.2(HPRT1):c.384+1G>A

[1233] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地至少存在于HPRT1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防莱施-奈恩综合征的方法。

[1234] 门克斯病

[1235] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与门克斯病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于ATP7A基因中,至少包括以下各项:

[1236] NM_000052.6(ATP7A):c.601C>T(p.Arg201Ter)

[1237] NM_000052.6(ATP7A):c.2938C>T(p.Arg980Ter)

[1238] NM_000052.6(ATP7A):c.3056G>A(p.Gly1019Asp)

[1239] NM_000052.6(ATP7A):c.598C>T(p.Gln200Ter)

[1240] NM_000052.6(ATP7A):c.1225C>T(p.Arg409Ter)

[1241] NM_000052.6(ATP7A):c.1544-1G>A

[1242] NM_000052.6(ATP7A):c.1639C>T(p.Arg547Ter)

[1243] NM_000052.6(ATP7A):c.1933C>T(p.Arg645Ter)

[1244] NM_000052.6(ATP7A):c.1946+5G>A

[1245] NM_000052.6(ATP7A):c.1950G>A(p.Trp650Ter)

[1246] NM_000052.6(ATP7A):c.2179G>A(p.Gly727Arg)

[1247] NM_000052.6(ATP7A):c.2187G>A(p.Trp729Ter)

[1248] NM_000052.6(ATP7A):c.2383C>T(p.Arg795Ter)

[1249] NM_000052.6(ATP7A):c.2499-1G>A

[1250] NM_000052.6(ATP7A):c.2555C>T(p.Pro852Leu)

[1251] NM_000052.6(ATP7A):c.2956C>T(p.Arg986Ter)

[1252] NM_000052.6(ATP7A):c.3112-1G>A

[1253] NM_000052.6(ATP7A):c.3466C>T(p.Gln1156Ter)

[1254] NM_000052.6(ATP7A):c.3502C>T(p.Gln1168Ter)

[1255] NM_000052.6(ATP7A):c.3764G>A(p.Gly1255Glu)

[1256] NM_000052.6(ATP7A):c.3943G>A(p.Gly1315Arg)

[1257] NM_000052.6(ATP7A):c.4123+1G>A

[1258] NM_000052.6(ATP7A):c.4226+5G>A

[1259] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地至少存在于ATP7A基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防门克斯病的方法。

[1260] 眼病

[1261] 与各种眼部疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开,这些疾病包括但不限于斯特格病(Stargardt Disease)、巴比二氏综合征(Bardet-Biedl Syndrome)、锥杆营养不良、先天性静止性夜盲症、乌谢尔综合征(Usher Syndrome)、莱伯氏先天性黑蒙症(Leber Congenital Amaurosis)、色素性视网膜炎和色盲症。因此,本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1262] 斯特格病

[1263] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与斯特格病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于ABCA4基因中,至少包括以下各项:

[1264] NM_000350.2(ABCA4):c.4429C>T(p.Gln1477Ter)

[1265] NM_000350.2(ABCA4):c.6647C>T(p.Ala2216Val)

[1266] NM_000350.2(ABCA4):c.5312+1G>A

[1267] NM_000350.2(ABCA4):c.5189G>A(p.Trp1730Ter)

[1268] NM_000350.2(ABCA4):c.4352+1G>A

[1269] NM_000350.2(ABCA4):c.4253+5G>A

[1270] NM_000350.2(ABCA4):c.3871C>T(p.Gln1291Ter)

[1271] NM_000350.2(ABCA4):c.3813G>A(p.Glu1271=)

[1272] NM_000350.2(ABCA4):c.1293G>A(p.Trp431Ter)

[1273] NM_000350.2(ABCA4):c.206G>A(p.Trp69Ter)

[1274] NM_000350.2(ABCA4):c.3322C>T(p.Arg1108Cys)

[1275] NM_000350.2(ABCA4):c.1804C>T(p.Arg602Trp)

[1276] NM_000350.2(ABCA4):c.1937+1G>A

[1277] NM_000350.2(ABCA4):c.2564G>A(p.Trp855Ter)

[1278] NM_000350.2(ABCA4):c.4234C>T(p.Gln1412Ter)

[1279] NM_000350.2(ABCA4):c.4457C>T(p.Pro1486Leu)

[1280] NM_000350.2(ABCA4):c.4594G>A(p.Asp1532Asn)

[1281] NM_000350.2(ABCA4):c.4919G>A(p.Arg1640Gln)

[1282] NM_000350.2(ABCA4):c.5196+1G>A

[1283] NM_000350.2(ABCA4):c.6316C>T(p.Arg2106Cys)

[1284] NM_000350.2(ABCA4):c.3056C>T(p.Thr1019Met)

[1285] NM_000350.2(ABCA4):c.52C>T(p.Arg18Trp)

[1286] NM_000350.2(ABCA4):c.122G>A(p.Trp41Ter)

[1287] NM_000350.2(ABCA4):c.1903C>T(p.Gln635Ter)

[1288] NM_000350.2(ABCA4):c.194G>A(p.Gly65Glu)

[1289] NM_000350.2(ABCA4):c.3085C>T(p.Gln1029Ter)

[1290] NM_000350.2(ABCA4):c.4195G>A(p.Glu1399Lys)

[1291] NM_000350.2(ABCA4):c.454C>T(p.Arg152Ter)

[1292] NM_000350.2(ABCA4):c.45G>A(p.Trp15Ter)

[1293] NM_000350.2(ABCA4):c.4610C>T(p.Thr1537Met)

[1294] NM_000350.2(ABCA4):c.6112C>T(p.Arg2038Trp)

[1295] NM_000350.2(ABCA4):c.6118C>T(p.Arg2040Ter)

[1296] NM_000350.2(ABCA4):c.6342G>A(p.Val2114=)

[1297] NM_000350.2(ABCA4):c.6658C>T(p.Gln2220Ter)

[1298] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于ABCA4基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防斯特格病的方法。

[1299] 巴比二氏综合征

[1300] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与巴比二氏综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于选自BBS1、BBS2、BBS7、BBS9、BBS10、BBS12、LZTFL1和TRIM32的至少一种基因中,至少包括以下各项:

[1301] NM_024649.4(BBS1):c.416G>A(p.Trp139Ter)

[1302] NM_024649.4(BBS1):c.871C>T(p.Gln291Ter)

[1303] NM_198428.2(BBS9):c.263+1G>A

[1304] NM_001178007.1(BBS12):c.1704G>A(p.Trp568Ter)

[1305] NM_001276378.1(LZTFL1):c.271C>T(p.Arg91Ter)

[1306] NM_031885.3(BBS2):c.1864C>T(p.Arg622Ter)

[1307] NM_198428.2(BBS9):c.1759C>T(p.Arg587Ter)

[1308] NM_198428.2(BBS9):c.1789+1G>A

[1309] NM_024649.4(BBS1):c.432+1G>A

[1310] NM_176824.2(BBS7):c.632C>T(p.Thr211Ile)

[1311] NM_012210.3(TRIM32):c.388C>T(p.Pro130Ser)

[1312] NM_031885.3(BBS2):c.823C>T(p.Arg275Ter)

[1313] NM_024685.3(BBS10):c.145C>T(p.Arg49Trp)

[1314] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自BBS1、BBS2、BBS7、BBS9、BBS10、BBS12、LZTFL1和TRIM32的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防巴比二氏综合征的方法。

[1315] 锥杆营养不良

[1316] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与锥杆营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于选自RPGRIP1、DRAM2、ABCA4、ADAM9和CACNA1F的至少一种基因中,至少包括以下各项:

[1317] NM_020366.3(RPGRIP1):c.154C>T(p.Arg52Ter)

[1318] NM_178454.5(DRAM2):c.494G>A(p.Trp165Ter)

[1319] NM_178454.5(DRAM2):c.131G>A(p.Ser44Asn)

[1320] NM_000350.2(ABCA4):c.161G>A(p.Cys54Tyr)

[1321] NM_000350.2(ABCA4):c.5714+5G>A

[1322] NM_000350.2(ABCA4):c.880C>T(p.Gln294Ter)

[1323] NM_000350.2(ABCA4):c.6079C>T(p.Leu2027Phe)

[1324] NM_000350.2(ABCA4):c.3113C>T(p.Ala1038Val)

[1325] NM_000350.2(ABCA4):c.634C>T(p.Arg212Cys)

[1326] NM_003816.2(ADAM9):c.490C>T(p.Arg164Ter)

[1327] NM_005183.3(CACNA1F):c.244C>T(p.Arg82Ter)

[1328] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自RPGRIP1、DRAM2、ABCA4、ADAM9和CACNA1F的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防锥杆营养不良的方法。

[1329] 先天性静止性夜盲症

[1330] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与先天性静止性夜盲症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于选自GRM6、TRPM1、GPR179和CACNA1F的至少一种基因中,至少包括以下各项:

[1331] NM_000843.3(GRM6):c.1462C>T(p.Gln488Ter)

[1332] NM_002420.5(TRPM1):c.2998C>T(p.Arg1000Ter)

[1333] NM_001004334.3(GPR179):c.673C>T(p.Gln225Ter)

[1334] NM_005183.3(CACNA1F):c.2576+1G>A

[1335] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自GRM6、TRPM1、GPR179和CACNA1F的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防先天性静止性夜盲症的方法。

[1336] 乌谢尔综合征

[1337] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与乌谢尔综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自MYO7A、USH1C、CDH23、PCDH15、USH2A、ADGRV1、WHRN和CLRN1的至少一种基因中，至少包括以下各项：

[1338] NM_000260.3(MYO7A):c.640G>A(p.Gly214Arg)

[1339] NM_000260.3(MYO7A):c.1200+1G>A

[1340] NM_000260.3(MYO7A):c.141G>A(p.Trp47Ter)

[1341] NM_000260.3(MYO7A):c.1556G>A(p.Gly519Asp)

[1342] NM_000260.3(MYO7A):c.1900C>T(p.Arg634Ter)

[1343] NM_000260.3(MYO7A):c.1963C>T(p.Gln655Ter)

[1344] NM_000260.3(MYO7A):c.2094+1G>A

[1345] NM_000260.3(MYO7A):c.4293G>A(p.Trp1431Ter)

[1346] NM_000260.3(MYO7A):c.5101C>T(p.Arg1701Ter)

[1347] NM_000260.3(MYO7A):c.5617C>T(p.Arg1873Trp)

[1348] NM_000260.3(MYO7A):c.5660C>T(p.Pro1887Leu)

[1349] NM_000260.3(MYO7A):c.6070C>T(p.Arg2024Ter)

[1350] NM_000260.3(MYO7A):c.470+1G>A

[1351] NM_000260.3(MYO7A):c.5968C>T(p.Gln1990Ter)

[1352] NM_000260.3(MYO7A):c.3719G>A(p.Arg1240Gln)

[1353] NM_000260.3(MYO7A):c.494C>T(p.Thr165Met)

[1354] NM_000260.3(MYO7A):c.5392C>T(p.Gln1798Ter)

[1355] NM_000260.3(MYO7A):c.5648G>A(p.Arg1883Gln)

[1356] NM_000260.3(MYO7A):c.448C>T(p.Arg150Ter)

[1357] NM_000260.3(MYO7A):c.700C>T(p.Gln234Ter)

[1358] NM_000260.3(MYO7A):c.635G>A(p.Arg212His)

[1359] NM_000260.3(MYO7A):c.1996C>T(p.Arg666Ter)

[1360] NM_005709.3(USH1C):c.216G>A(p.Val72=)

[1361] NM_022124.5(CDH23):c.7362+5G>A

[1362] NM_022124.5(CDH23):c.3481C>T(p.Arg1161Ter)

[1363] NM_022124.5(CDH23):c.3628C>T(p.Gln1210Ter)

[1364] NM_022124.5(CDH23):c.5272C>T(p.Gln1758Ter)

[1365] NM_022124.5(CDH23):c.5712+1G>A

[1366] NM_022124.5(CDH23):c.5712G>A(p.Thr1904=)

[1367] NM_022124.5(CDH23):c.5923+1G>A

[1368] NM_022124.5(CDH23):c.6049+1G>A

[1369] NM_022124.5(CDH23):c.7776G>A(p.Trp2592Ter)

[1370] NM_022124.5(CDH23):c.9556C>T(p.Arg3186Ter)

[1371] NM_022124.5(CDH23):c.3706C>T(p.Arg1236Ter)

[1372] NM_022124.5(CDH23):c.4309C>T(p.Arg1437Ter)

[1373] NM_022124.5(CDH23):c.6050-9G>A

[1374] NM_033056.3(PCDH15):c.3316C>T(p.Arg1106Ter)

[1375] NM_033056.3(PCDH15):c.7C>T(p.Arg3Ter)

[1376] NM_033056.3(PCDH15):c.1927C>T(p.Arg643Ter)

[1377] NM_001142772.1(PCDH15):c.400C>T(p.Arg134Ter)

[1378] NM_033056.3(PCDH15):c.3358C>T(p.Arg1120Ter)

[1379] NM_206933.2(USH2A):c.11048-1G>A

[1380] NM_206933.2(USH2A):c.1143+1G>A

[1381] NM_206933.2(USH2A):c.11954G>A(p.Trp3985Ter)

[1382] NM_206933.2(USH2A):c.12868C>T(p.Gln4290Ter)

[1383] NM_206933.2(USH2A):c.14180G>A(p.Trp4727Ter)

[1384] NM_206933.2(USH2A):c.14911C>T(p.Arg4971Ter)

[1385] NM_206933.2(USH2A):c.5788C>T(p.Arg1930Ter)

[1386] NM_206933.2(USH2A):c.5858-1G>A

[1387] NM_206933.2(USH2A):c.6224G>A(p.Trp2075Ter)

[1388] NM_206933.2(USH2A):c.820C>T(p.Arg274Ter)

[1389] NM_206933.2(USH2A):c.8981G>A(p.Trp2994Ter)

[1390] NM_206933.2(USH2A):c.9304C>T(p.Gln3102Ter)

[1391] NM_206933.2(USH2A):c.13010C>T(p.Thr4337Met)

[1392] NM_206933.2(USH2A):c.14248C>T(p.Gln4750Ter)

[1393] NM_206933.2(USH2A):c.6398G>A(p.Trp2133Ter)

[1394] NM_206933.2(USH2A):c.632G>A(p.Trp211Ter)

[1395] NM_206933.2(USH2A):c.6601C>T(p.Gln2201Ter)

[1396] NM_206933.2(USH2A):c.13316C>T(p.Thr4439Ile)

[1397] NM_206933.2(USH2A):c.4405C>T(p.Gln1469Ter)

[1398] NM_206933.2(USH2A):c.9570+1G>A

[1399] NM_206933.2(USH2A):c.8740C>T(p.Arg2914Ter)

[1400] NM_206933.2(USH2A):c.8681+1G>A

[1401] NM_206933.2(USH2A):c.1000C>T(p.Arg334Trp)

[1402] NM_206933.2(USH2A):c.14175G>A(p.Trp4725Ter)

[1403] NM_206933.2(USH2A):c.9390G>A(p.Trp3130Ter)

[1404] NM_206933.2(USH2A):c.908G>A(p.Arg303His)

[1405] NM_206933.2(USH2A):c.5776+1G>A

[1406] NM_206933.2(USH2A):c.11156G>A(p.Arg3719His)

[1407] NM_032119.3(ADGRV1):c.2398C>T(p.Arg800Ter)

[1408] NM_032119.3(ADGRV1):c.7406G>A(p.Trp2469Ter)

[1409] NM_032119.3(ADGRV1):c.12631C>T(p.Arg4211Ter)

[1410] NM_032119.3(ADGRV1):c.7129C>T(p.Arg2377Ter)

[1411] NM_032119.3(ADGRV1):c.14885G>A(p.Trp4962Ter)

[1412] NM_015404.3(WHRN):c.1267C>T(p.Arg423Ter)

[1413] NM_174878.2(CLRN1):c.619C>T(p.Arg207Ter)

[1414] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自MYO7A、USH1C、CDH23、PCDH15、USH2A、ADGRV1、WHRN和CLRN1的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防增强型乌谢尔综合征的方法。

[1415] 莱伯氏先天性黑蒙症

[1416] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与莱伯氏先天性黑蒙症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自TULP1、RPE65、SPATA7、AIPL1、CRB1、NMNAT1和PEX1的至少一种基因中，至少包括以下各项：

[1417] NM_003322.5(TULP1):c.1495+1G>A

[1418] NM_000329.2(RPE65):c.11+5G>A

[1419] NM_018418.4(SPATA7):c.322C>T(p.Arg108Ter)

[1420] NM_014336.4(AIPL1):c.784G>A(p.Gly262Ser)

[1421] NM_201253.2(CRB1):c.1576C>T(p.Arg526Ter)

[1422] NM_201253.2(CRB1):c.3307G>A(p.Gly1103Arg)

[1423] NM_201253.2(CRB1):c.2843G>A(p.Cys948Tyr)

[1424] NM_022787.3(NMNAT1):c.769G>A(p.Glu257Lys)

[1425] NM_000466.2(PEX1):c.2528G>A(p.Gly843Asp)

[1426] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自TULP1、RPE65、SPATA7、AIPL1、CRB1、NMNAT1和PEX1的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防莱伯氏先天性黑蒙症的方法。

[1427] 色素性视网膜炎

[1428] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与色素性视网膜炎相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自CRB1、IFT140、RP1、IMPDH1、PRPF31、RPGR、ABCA4、RPE65、EYS、NRL、FAM161A、NR2E3、USH2A、RHO、PDE6B、KLHL7、PDE6A、CNGB1、BEST1、C2orf71、PRPH2、CA4、CERKL、RPE65、PDE6B和ADGRV1的至少一种基因中，至少包括以下各项：

[1429] NM_001257965.1(CRB1):c.2711G>A(p.Cys904Tyr)

[1430] NM_014714.3(IFT140):c.3827G>A(p.Gly1276Glu)

[1431] NM_006269.1(RP1):c.2029C>T(p.Arg677Ter)

[1432] NM_000883.3(IMPDH1):c.931G>A(p.Asp311Asn)

[1433] NM_015629.3(PRPF31):c.1273C>T(p.Gln425Ter)

[1434] NM_015629.3(PRPF31):c.1073+1G>A

[1435] NM_000328.2(RPGR):c.1387C>T(p.Gln463Ter)

[1436] NM_000350.2(ABCA4):c.4577C>T(p.Thr1526Met)

[1437] NM_000350.2(ABCA4):c.6229C>T(p.Arg2077Trp)

[1438] NM_000329.2(RPE65):c.271C>T(p.Arg91Trp)

[1439] NM_001142800.1(EYS):c.2194C>T(p.Gln732Ter)

[1440] NM_001142800.1(EYS):c.490C>T(p.Arg164Ter)

[1441] NM_006177.3(NRL):c.151C>T(p.Pro51Ser)

[1442] NM_001201543.1(FAM161A):c.1567C>T(p.Arg523Ter)

[1443] NM_014249.3(NR2E3):c.166G>A(p.Gly56Arg)

[1444] NM_206933.2(USH2A):c.2209C>T(p.Arg737Ter)

[1445] NM_206933.2(USH2A):c.14803C>T(p.Arg4935Ter)

[1446] NM_206933.2(USH2A):c.10073G>A(p.Cys3358Tyr)

[1447] NM_000539.3(RHO):c.541G>A(p.Glu181Lys)

[1448] NM_000283.3(PDE6B):c.892C>T(p.Gln298Ter)

[1449] NM_001031710.2(KLHL7):c.458C>T(p.Ala153Val)

[1450] NM_000440.2(PDE6A):c.1926+1G>A

[1451] NM_001297.4(CNGB1):c.2128C>T(p.Gln710Ter)

[1452] NM_001297.4(CNGB1):c.952C>T(p.Gln318Ter)

[1453] NM_004183.3(BEST1):c.682G>A(p.Asp228Asn)

[1454] NM_001029883.2(C2orf71):c.1828C>T(p.Gln610Ter)

[1455] NM_000322.4(PRPH2):c.647C>T(p.Pro216Leu)

[1456] NM_000717.4(CA4):c.40C>T(p.Arg14Trp)

[1457] NM_201548.4(CERKL):c.769C>T(p.Arg257Ter)

[1458] NM_000329.2(RPE65):c.118G>A(p.Gly40Ser)

[1459] NM_000322.4(PRPH2):c.499G>A(p.Gly167Ser)

[1460] NM_000539.3(RHO):c.403C>T(p.Arg135Trp)

[1461] NM_000283.3(PDE6B):c.2193+1G>A

[1462] NM_032119.3(ADGRV1):c.6901C>T(p.Gln2301Ter)

[1463] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自CRB1、IFT140、RP1、IMPDH1、PRPF31、RPGR、ABCA4、RPE65、EYS、NRL、FAM161A、NR2E3、USH2A、RHO、PDE6B、KLHL7、PDE6A、CNGB1、BEST1、C2orf71、PRPH2、CA4、CERKL、RPE65、PDE6B和ADGRV1的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防色素性视网膜炎的方法。

[1464] 色盲症

[1465] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与色盲症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自

CNGA3、CNGB3和ATF6的至少一种基因中，至少包括以下各项：

[1466] NM_001298.2(CNGA3):c.847C>T(p.Arg283Trp)

[1467] NM_001298.2(CNGA3):c.101+1G>A

[1468] NM_001298.2(CNGA3):c.1585G>A(p.Val529Met)

[1469] NM_019098.4(CNGB3):c.1578+1G>A

[1470] NM_019098.4(CNGB3):c.607C>T(p.Arg203Ter)

[1471] NM_019098.4(CNGB3):c.1119G>A(p.Trp373Ter)

[1472] NM_007348.3(ATF6):c.970C>T(p.Arg324Cys)

[1473] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自CNGA3、CNGB3和ATF6的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防色盲症的方法。

[1474] 影响听力的疾病

[1475] 与影响听力的各种疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于耳聋和非综合征性听力丧失。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1476] 耳聋

[1477] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与耳聋相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自FGF3、MYO7A、STRC、ACTG1、SLC17A8、TMC1、GJB2、MYH14、COCH、CDH23、USH1C、GJB2、MYO7A、PCDH15、MYO15A、MYO3A、WHRN、DFNB59、TMC1、LOXHD1、TMPRSS3、OTOGL、OTOF、JAG1和MARVELD2的至少一种基因中，至少包括以下各项：

[1478] NM_005247.2(FGF3):c.283C>T(p.Arg95Trp)

[1479] NM_000260.3(MYO7A):c.652G>A(p.Asp218Asn)

[1480] NM_000260.3(MYO7A):c.689C>T(p.Ala230Val)

[1481] NM_153700.2(STRC):c.4057C>T(p.Gln1353Ter)

[1482] NM_001614.3(ACTG1):c.721G>A(p.Glu241Lys)

[1483] NM_139319.2(SLC17A8):c.632C>T(p.Ala211Val)

[1484] NM_138691.2(TMC1):c.1714G>A(p.Asp572Asn)

[1485] NM_004004.5(GJB2):c.598G>A(p.Gly200Arg)

[1486] NM_004004.5(GJB2):c.71G>A(p.Trp24Ter)

[1487] NM_004004.5(GJB2):c.416G>A(p.Ser139Asn)

[1488] NM_004004.5(GJB2):c.224G>A(p.Arg75Gln)

[1489] NM_004004.5(GJB2):c.95G>A(p.Arg32His)

[1490] NM_004004.5(GJB2):c.250G>A(p.Val84Met)

[1491] NM_004004.5(GJB2):c.428G>A(p.Arg143Gln)

[1492] NM_004004.5(GJB2):c.551G>A(p.Arg184Gln)

[1493] NM_004004.5(GJB2):c.223C>T(p.Arg75Trp)

[1494] NM_024729.3(MYH14):c.359C>T(p.Ser120Leu)

[1495] NM_004086.2(COCH):c.151C>T(p.Pro51Ser)

[1496] NM_022124.5(CDH23):c.4021G>A(p.Asp1341Asn)

[1497] NM_153700.2(STRC):c.4701+1G>A

[1498] NM_153676.3(USH1C):c.496+1G>A

[1499] NM_004004.5(GJB2):c.131G>A(p.Trp44Ter)

[1500] NM_004004.5(GJB2):c.283G>A(p.Val95Met)

[1501] NM_004004.5(GJB2):c.298C>T(p.His100Tyr)

[1502] NM_004004.5(GJB2):c.427C>T(p.Arg143Trp)

[1503] NM_004004.5(GJB2):c.109G>A(p.Val37Ile)

[1504] NM_004004.5(GJB2):c.-23+1G>A

[1505] NM_004004.5(GJB2):c.148G>A(p.Asp50Asn)

[1506] NM_004004.5(GJB2):c.134G>A(p.Gly45Glu)

[1507] NM_004004.5(GJB2):c.370C>T(p.Gln124Ter)

[1508] NM_004004.5(GJB2):c.230G>A(p.Trp77Ter)

[1509] NM_004004.5(GJB2):c.231G>A(p.Trp77Ter)

[1510] NM_000260.3(MYO7A):c.5899C>T(p.Arg1967Ter)

[1511] NM_000260.3(MYO7A):c.2005C>T(p.Arg669Ter)

[1512] NM_033056.3(PCDH15):c.733C>T(p.Arg245Ter)

[1513] NM_016239.3(MYO15A):c.3866+1G>A

[1514] NM_016239.3(MYO15A):c.6178-1G>A

[1515] NM_016239.3(MYO15A):c.8714-1G>A

[1516] NM_017433.4(MYO3A):c.2506-1G>A

[1517] NM_015404.3(WHRN):c.1417-1G>A

[1518] NM_001042702.3(DFNB59):c.499C>T(p.Arg167Ter)

[1519] NM_138691.2(TMC1):c.100C>T(p.Arg34Ter)

[1520] NM_138691.2(TMC1):c.1165C>T(p.Arg389Ter)

[1521] NM_144612.6(LOXHD1):c.2008C>T(p.Arg670Ter)

[1522] NM_144612.6(LOXHD1):c.4714C>T(p.Arg1572Ter)

[1523] NM_144612.6(LOXHD1):c.4480C>T(p.Arg1494Ter)

[1524] NM_024022.2(TMPRSS3):c.325C>T(p.Arg109Trp)

[1525] NM_173591.3(OTOGL):c.3076C>T(p.Gln1026Ter)

[1526] NM_194248.2(OTOF):c.4483C>T(p.Arg1495Ter)

[1527] NM_194248.2(OTOF):c.2122C>T(p.Arg708Ter)

[1528] NM_194248.2(OTOF):c.2485C>T(p.Gln829Ter)

[1529] NM_001038603.2(MARVELD2):c.1498C>T(p.Arg500Ter)

[1530] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自FGF3、MYO7A、STRC、ACTG1、SLC17A8、TMC1、GJB2、MYH14、COCH、CDH23、USH1C、GJB2、MYO7A、PCDH15、MYO15A、MYO3A、WHRN、DFNB59、TMC1、LOXHD1、

TMPRSS3、OTOGL、OTOF、JAG1和MARVELD2的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防耳聋的方法。

[1531] 非综合征性听力丧失

[1532] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与非综合征性听力丧失相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自GJB2、POU3F4、MYO15A、TMPRSS3、LOXHD1、OTOF、MYO6、OTOA、STRC、TRIOBP、MARVELD2、TMC1、TECTA、OTOGL和GIPC3的至少一种基因中，至少包括以下各项：

[1533] NM_004004.5(GJB2):c.169C>T(p.Gln57Ter)

[1534] NM_000307.4(POU3F4):c.499C>T(p.Arg167Ter)

[1535] NM_016239.3(MYO15A):c.8767C>T(p.Arg2923Ter)

[1536] NM_024022.2(TMPRSS3):c.323-6G>A

[1537] NM_024022.2(TMPRSS3):c.916G>A(p.Ala306Thr)

[1538] NM_144612.6(LOXHD1):c.2497C>T(p.Arg833Ter)

[1539] NM_194248.2(OTOF):c.2153G>A(p.Trp718Ter)

[1540] NM_194248.2(OTOF):c.2818C>T(p.Gln940Ter)

[1541] NM_194248.2(OTOF):c.4799+1G>A

[1542] NM_004999.3(MYO6):c.826C>T(p.Arg276Ter)

[1543] NM_144672.3(OTOA):c.1880+1G>A

[1544] NM_153700.2(STRC):c.5188C>T(p.Arg1730Ter)

[1545] NM_153700.2(STRC):c.3670C>T(p.Arg1224Ter)

[1546] NM_153700.2(STRC):c.4402C>T(p.Arg1468Ter)

[1547] NM_024022.2(TMPRSS3):c.1192C>T(p.Gln398Ter)

[1548] NM_001039141.2(TRIOBP):c.6598C>T(p.Arg2200Ter)

[1549] NM_016239.3(MYO15A):c.7893+1G>A

[1550] NM_016239.3(MYO15A):c.5531+1G>A

[1551] NM_016239.3(MYO15A):c.6046+1G>A

[1552] NM_144612.6(LOXHD1):c.3169C>T(p.Arg1057Ter)

[1553] NM_001038603.2(MARVELD2):c.1331+1G>A

[1554] NM_138691.2(TMC1):c.1676G>A(p.Trp559Ter)

[1555] NM_138691.2(TMC1):c.1677G>A(p.Trp559Ter)

[1556] NM_005422.2(TECTA):c.5977C>T(p.Arg1993Ter)

[1557] NM_173591.3(OTOGL):c.4987C>T(p.Arg1663Ter)

[1558] NM_153700.2(STRC):c.3493C>T(p.Gln1165Ter)

[1559] NM_153700.2(STRC):c.3217C>T(p.Arg1073Ter)

[1560] NM_016239.3(MYO15A):c.5896C>T(p.Arg1966Ter)

[1561] NM_133261.2(GIPC3):c.411+1G>A

[1562] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自GJB2、POU3F4、MYO15A、TMPRSS3、LOXHD1、OTOF、MYO6、

OTOA、STRC、TRIOBP、MARVELD2、TMC1、TECTA、OTOGL和GIPC3的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防非综合征性听力丧失的方法。

[1563] 血液病症

[1564] 与各种血液病症相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于β地中海贫血、甲型血友病、乙型血友病、丙型血友病和威斯科特-奥尔德里奇综合征。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1565] β地中海贫血

[1566] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与β地中海贫血相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于HBB基因中，至少包括以下各项：

[1567] NM_000518.4(HBB):c.-137C>T

[1568] NM_000518.4(HBB):c.-50-88C>T

[1569] NM_000518.4(HBB):c.-140C>T

[1570] NM_000518.4(HBB):c.316-197C>T

[1571] NM_000518.4(HBB):c.93-21G>A

[1572] NM_000518.4(HBB):c.114G>A(p.Trp38Ter)

[1573] NM_000518.4(HBB):c.118C>T(p.Gln40Ter)

[1574] NM_000518.4(HBB):c.92+1G>A

[1575] NM_000518.4(HBB):c.315+1G>A

[1576] NM_000518.4(HBB):c.92+5G>A

[1577] NM_000518.4(HBB):c.-50-101C>T

[1578] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于HBB基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防β地中海贫血的方法。

[1579] 甲型血友病

[1580] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与甲型血友病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于F8基因中，至少包括以下各项：

[1581] NM_000132.3(F8):c.3169G>A(p.Glu1057Lys)

[1582] NM_000132.3(F8):c.902G>A(p.Arg301His)

[1583] NM_000132.3(F8):c.1834C>T(p.Arg612Cys)

[1584] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于F8基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防甲型血友病的方法。

[1585] 莱登因子V(Factor V Leiden)

[1586] 在一些实施方案中，本文所述的方法、统和组合物用于校正莱登因子V突变。这种致病单点突变(G1746->A)代表了高加索人群的遗传性多因素易栓症中最丰富的遗传风险因素。由于点突变，在凝血因子F5的蛋白C依赖性蛋白水解切割位点(R533R534)上出现单个氨基酸取代(R534[→]Q)。杂合缺陷仅伴随血栓形成风险的微小增加(约8倍)，而纯合缺陷具有更为明显的作用(风险增加了>80倍)。19定向RNA编辑有潜力通过以RNA水平进行修复弥补这种遗传缺陷。

[1587] 乙型血友病

[1588] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与乙型血友病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于F9基因中，至少包括以下各项：

[1589] NM_000133.3(F9):c.835G>A(p.Ala279Thr)

[1590] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于F9基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防乙型血友病的方法。

[1591] 丙型血友病

[1592] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与丙型血友病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于F11基因中，至少包括以下各项：

[1593] NM_000128.3(F11):c.400C>T(p.Gln134Ter)

[1594] NM_000128.3(F11):c.1432G>A(p.Gly478Arg)

[1595] NM_000128.3(F11):c.1288G>A(p.Ala430Thr)

[1596] NM_000128.3(F11):c.326-1G>A

[1597] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于F11基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防丙型血友病的方法。

[1598] 威斯科特-奥尔德里奇综合征

[1599] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与威斯科特-奥尔德里奇综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于WAS基因中，至少包括以下各项：

[1600] NM_000377.2(WAS):c.37C>T(p.Arg13Ter)

[1601] NM_000377.2(WAS):c.257G>A(p.Arg86His)

[1602] NM_000377.2(WAS):c.777+1G>A

[1603] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于WAS基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防威斯科特-奥尔德里奇综合征的方法。

[1604] 肝脏疾病

[1605] 与各种肝脏疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于运甲状腺素蛋白淀粉样变性、α-1抗胰蛋白酶缺乏症、威尔逊氏病(Wilson's disease)和苯丙酮尿症。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1606] 运甲状腺素蛋白淀粉样变性

[1607] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与运甲状腺素蛋白淀粉样变性相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于TTR基因中，至少包括以下各项：

[1608] NM_000371.3(TTR):c.424G>A(p.Val142Ile)

[1609] NM_000371.3(TTR):c.148G>A(p.Val50Met)

[1610] NM_000371.3(TTR):c.118G>A(p.Val40Ile)

[1611] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于TTR基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防运甲状腺素蛋白淀粉样变性的方法。

[1612] α-1-抗胰蛋白酶缺乏症

[1613] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与α-1-抗胰蛋白酶缺乏症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于SERPINA1基因中，至少包括以下各项：

[1614] NM_000295.4(SERPINA1):c.538C>T(p.Gln180Ter)

[1615] NM_001127701.1(SERPINA1):c.1178C>T(p.Pro393Leu)

[1616] NM_001127701.1(SERPINA1):c.230C>T(p.Ser77Phe)

[1617] NM_001127701.1(SERPINA1):c.1096G>A(p.Glu366Lys)

[1618] NM_000295.4(SERPINA1):c.1177C>T(p.Pro393Ser)

[1619] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于SERPINA1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防α1-抗胰蛋白酶缺乏症的方法。

[1620] 威尔逊氏病

[1621] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与威尔逊氏病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于ATP7B基因中，至少包括以下各项：

[1622] NM_000053.3(ATP7B):c.2293G>A(p.Asp765Asn)

[1623] NM_000053.3(ATP7B):c.3955C>T(p.Arg1319Ter)

[1624] NM_000053.3(ATP7B):c.2865+1G>A

[1625] NM_000053.3(ATP7B):c.3796G>A(p.Gly1266Arg)

[1626] NM_000053.3(ATP7B):c.2621C>T(p.Ala874Val)

[1627] NM_000053.3(ATP7B):c.2071G>A(p.Gly691Arg)

[1628] NM_000053.3(ATP7B):c.2128G>A(p.Gly710Ser)

[1629] NM_000053.3(ATP7B):c.2336G>A(p.Trp779Ter)

[1630] NM_000053.3(ATP7B):c.4021G>A(p.Gly1341Ser)

[1631] NM_000053.3(ATP7B):c.3182G>A(p.Gly1061Glu)

[1632] NM_000053.3(ATP7B):c.4114C>T(p.Gln1372Ter)

[1633] NM_000053.3(ATP7B):c.1708-1G>A

[1634] NM_000053.3(ATP7B):c.865C>T(p.Gln289Ter)

[1635] NM_000053.3(ATP7B):c.2930C>T(p.Thr977Met)

[1636] NM_000053.3(ATP7B):c.3659C>T(p.Thr1220Met)

[1637] NM_000053.3(ATP7B):c.2605G>A(p.Gly869Arg)

[1638] NM_000053.3(ATP7B):c.2975C>T(p.Pro992Leu)

[1639] NM_000053.3(ATP7B):c.2519C>T(p.Pro840Leu)

[1640] NM_000053.3(ATP7B):c.2906G>A(p.Arg969Gln)

[1641] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于ATP7B基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防威尔逊氏病的方法。

[1642] 苯丙酮尿症

[1643] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与苯丙酮尿症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于PAH基因中，至少包括以下各项：

[1644] NM_000277.1(PAH):c.1315+1G>A

[1645] NM_000277.1(PAH):c.1222C>T(p.Arg408Trp)

[1646] NM_000277.1(PAH):c.838G>A(p.Glu280Lys)

[1647] NM_000277.1(PAH):c.331C>T(p.Arg111Ter)

[1648] NM_000277.1(PAH):c.782G>A(p.Arg261Gln)

[1649] NM_000277.1(PAH):c.754C>T(p.Arg252Trp)

[1650] NM_000277.1(PAH):c.473G>A(p.Arg158Gln)

[1651] NM_000277.1(PAH):c.727C>T(p.Arg243Ter)

[1652] NM_000277.1(PAH):c.842C>T(p.Pro281Leu)

[1653] NM_000277.1(PAH):c.728G>A(p.Arg243Gln)

[1654] NM_000277.1(PAH):c.1066-11G>A

[1655] NM_000277.1(PAH):c.781C>T(p.Arg261Ter)

[1656] NM_000277.1(PAH):c.1223G>A(p.Arg408Gln)

[1657] NM_000277.1(PAH):c.1162G>A(p.Val388Met)

[1658] NM_000277.1(PAH):c.1066-3C>T

[1659] NM_000277.1(PAH):c.1208C>T(p.Ala403Val)

[1660] NM_000277.1(PAH):c.890G>A(p.Arg297His)

[1661] NM_000277.1(PAH):c.926C>T(p.Ala309Val)

[1662] NM_000277.1(PAH):c.441+1G>A

[1663] NM_000277.1(PAH):c.526C>T(p.Arg176Ter)

[1664] NM_000277.1(PAH):c.688G>A(p.Val230Ile)

[1665] NM_000277.1(PAH):c.721C>T(p.Arg241Cys)

[1666] NM_000277.1(PAH):c.745C>T(p.Leu249Phe)

[1667] NM_000277.1(PAH):c.442-1G>A

[1668] NM_000277.1(PAH):c.842+1G>A

[1669] NM_000277.1(PAH):c.776C>T(p.Ala259Val)

[1670] NM_000277.1(PAH):c.1200-1G>A

[1671] NM_000277.1(PAH):c.912+1G>A

[1672] NM_000277.1(PAH):c.1065+1G>A

[1673] NM_000277.1(PAH):c.472C>T(p.Arg158Trp)

[1674] NM_000277.1(PAH):c.755G>A(p.Arg252Gln)

[1675] NM_000277.1(PAH):c.809G>A(p.Arg270Lys)

[1676] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于PAH基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防苯丙酮尿症的方法。

[1677] 肾脏疾病

[1678] 与各种肾脏疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开,这些疾病包括但不限于常染色体隐性遗传性多囊肾病和肾性肉毒碱转运缺陷。因此,本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1679] 常染色体隐性遗传性多囊肾病

[1680] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与常染色体隐性遗传性多囊肾病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于PKHD1基因中,至少包括以下各项:

[1681] NM_138694.3(PKHD1):c.10444C>T(p.Arg3482Cys)

[1682] NM_138694.3(PKHD1):c.9319C>T(p.Arg3107Ter)

[1683] NM_138694.3(PKHD1):c.1480C>T(p.Arg494Ter)

[1684] NM_138694.3(PKHD1):c.707+1G>A

[1685] NM_138694.3(PKHD1):c.1486C>T(p.Arg496Ter)

[1686] NM_138694.3(PKHD1):c.8303-1G>A

[1687] NM_138694.3(PKHD1):c.2854G>A(p.Gly952Arg)

[1688] NM_138694.3(PKHD1):c.7194G>A(p.Trp2398Ter)

[1689] NM_138694.3(PKHD1):c.10219C>T(p.Gln3407Ter)

[1690] NM_138694.3(PKHD1):c.107C>T(p.Thr36Met)

[1691] NM_138694.3(PKHD1):c.8824C>T(p.Arg2942Ter)

[1692] NM_138694.3(PKHD1):c.982C>T(p.Arg328Ter)

[1693] NM_138694.3(PKHD1):c.4870C>T(p.Arg1624Trp)

[1694] NM_138694.3(PKHD1):c.1602+1G>A

[1695] NM_138694.3(PKHD1):c.1694-1G>A

[1696] NM_138694.3(PKHD1):c.2341C>T(p.Arg781Ter)

[1697] NM_138694.3(PKHD1):c.2407+1G>A

[1698] NM_138694.3(PKHD1):c.2452C>T(p.Gln818Ter)

[1699] NM_138694.3(PKHD1):c.5236+1G>A

[1700] NM_138694.3(PKHD1):c.6499C>T(p.Gln2167Ter)

[1701] NM_138694.3(PKHD1):c.2725C>T(p.Arg909Ter)

[1702] NM_138694.3(PKHD1):c.370C>T(p.Arg124Ter)

[1703] NM_138694.3(PKHD1):c.2810G>A(p.Trp937Ter)

[1704] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于PKHD1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防常染色体隐性遗传性多囊肾病的方法。

[1705] 肾性肉毒碱转运缺陷

[1706] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与肾性肉毒碱转运缺陷相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于SLC22A5基因中，至少包括以下各项：

[1707] NM_003060.3(SLC22A5):c.760C>T(p.Arg254Ter)

[1708] NM_003060.3(SLC22A5):c.396G>A(p.Trp132Ter)

[1709] NM_003060.3(SLC22A5):c.844C>T(p.Arg282Ter)

[1710] NM_003060.3(SLC22A5):c.505C>T(p.Arg169Trp)

[1711] NM_003060.3(SLC22A5):c.1319C>T(p.Thr440Met)

[1712] NM_003060.3(SLC22A5):c.1195C>T(p.Arg399Trp)

[1713] NM_003060.3(SLC22A5):c.695C>T(p.Thr232Met)

[1714] NM_003060.3(SLC22A5):c.845G>A(p.Arg282Gln)

[1715] NM_003060.3(SLC22A5):c.1193C>T(p.Pro398Leu)

[1716] NM_003060.3(SLC22A5):c.1463G>A(p.Arg488His)

[1717] NM_003060.3(SLC22A5):c.338G>A(p.Cys113Tyr)

[1718] NM_003060.3(SLC22A5):c.136C>T(p.Pro46Ser)

[1719] NM_003060.3(SLC22A5):c.506G>A(p.Arg169Gln)

[1720] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于SLC22A5基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防肾性肉毒碱转运缺陷的方法。

[1721] 肌肉疾病

[1722] 与各种肌肉疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于杜兴氏肌营养不良(Duchenne muscular

dystrophy)、贝克肌营养不良(Becker muscular dystrophy)、肢带型肌营养不良、埃-德二氏肌营养不良(Emery-Dreifuss muscular dystrophy)和面肩肱型肌营养不良。因此,本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1723] 杜兴氏肌营养不良

[1724] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与杜兴氏肌营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于DMD基因中,至少包括以下各项:

[1725] NM_004006.2(DMD):c.2797C>T(p.Gln933Ter)

[1726] NM_004006.2(DMD):c.4870C>T(p.Gln1624Ter)

[1727] NM_004006.2(DMD):c.5551C>T(p.Gln1851Ter)

[1728] NM_004006.2(DMD):c.3188G>A(p.Trp1063Ter)

[1729] NM_004006.2(DMD):c.8357G>A(p.Trp2786Ter)

[1730] NM_004006.2(DMD):c.7817G>A(p.Trp2606Ter)

[1731] NM_004006.2(DMD):c.7755G>A(p.Trp2585Ter)

[1732] NM_004006.2(DMD):c.5917C>T(p.Gln1973Ter)

[1733] NM_004006.2(DMD):c.5641C>T(p.Gln1881Ter)

[1734] NM_004006.2(DMD):c.5131C>T(p.Gln1711Ter)

[1735] NM_004006.2(DMD):c.4240C>T(p.Gln1414Ter)

[1736] NM_004006.2(DMD):c.3427C>T(p.Gln1143Ter)

[1737] NM_004006.2(DMD):c.2407C>T(p.Gln803Ter)

[1738] NM_004006.2(DMD):c.2368C>T(p.Gln790Ter)

[1739] NM_004006.2(DMD):c.1683G>A(p.Trp561Ter)

[1740] NM_004006.2(DMD):c.1663C>T(p.Gln555Ter)

[1741] NM_004006.2(DMD):c.1388G>A(p.Trp463Ter)

[1742] NM_004006.2(DMD):c.1331+1G>A

[1743] NM_004006.2(DMD):c.1324C>T(p.Gln442Ter)

[1744] NM_004006.2(DMD):c.355C>T(p.Gln119Ter)

[1745] NM_004006.2(DMD):c.94-1G>A

[1746] NM_004006.2(DMD):c.5506C>T(p.Gln1836Ter)

[1747] NM_004006.2(DMD):c.1504C>T(p.Gln502Ter)

[1748] NM_004006.2(DMD):c.5032C>T(p.Gln1678Ter)

[1749] NM_004006.2(DMD):c.457C>T(p.Gln153Ter)

[1750] NM_004006.2(DMD):c.1594C>T(p.Gln532Ter)

[1751] NM_004006.2(DMD):c.1150-1G>A

[1752] NM_004006.2(DMD):c.6223C>T(p.Gln2075Ter)

[1753] NM_004006.2(DMD):c.3747G>A(p.Trp1249Ter)

[1754] NM_004006.2(DMD):c.2861G>A(p.Trp954Ter)

[1755] NM_004006.2(DMD):c.9563+1G>A

[1756] NM_004006.2(DMD):c.4483C>T(p.Gln1495Ter)

[1757] NM_004006.2(DMD):c.4312C>T(p.Gln1438Ter)

[1758] NM_004006.2(DMD):c.8209C>T(p.Gln2737Ter)

[1759] NM_004006.2(DMD):c.4071+1G>A

[1760] NM_004006.2(DMD):c.2665C>T(p.Arg889Ter)

[1761] NM_004006.2(DMD):c.2202G>A(p.Trp734Ter)

[1762] NM_004006.2(DMD):c.2077C>T(p.Gln693Ter)

[1763] NM_004006.2(DMD):c.1653G>A(p.Trp551Ter)

[1764] NM_004006.2(DMD):c.1061G>A(p.Trp354Ter)

[1765] NM_004006.2(DMD):c.8914C>T(p.Gln2972Ter)

[1766] NM_004006.2(DMD):c.6118-1G>A

[1767] NM_004006.2(DMD):c.4729C>T(p.Arg1577Ter)

[1768] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于DMD基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防杜兴氏肌营养不良的方法。

[1769] 贝克肌营养不良

[1770] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与贝克肌营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于DMD基因中,至少包括以下各项:

[1771] NM_004006.2(DMD):c.3413G>A(p.Trp1138Ter)

[1772] NM_004006.2(DMD):c.358-1G>A

[1773] NM_004006.2(DMD):c.10108C>T(p.Arg3370Ter)

[1774] NM_004006.2(DMD):c.6373C>T(p.Gln2125Ter)

[1775] NM_004006.2(DMD):c.9568C>T(p.Arg3190Ter)

[1776] NM_004006.2(DMD):c.8713C>T(p.Arg2905Ter)

[1777] NM_004006.2(DMD):c.1615C>T(p.Arg539Ter)

[1778] NM_004006.2(DMD):c.3151C>T(p.Arg1051Ter)

[1779] NM_004006.2(DMD):c.3432+1G>A

[1780] NM_004006.2(DMD):c.5287C>T(p.Arg1763Ter)

[1781] NM_004006.2(DMD):c.5530C>T(p.Arg1844Ter)

[1782] NM_004006.2(DMD):c.8608C>T(p.Arg2870Ter)

[1783] NM_004006.2(DMD):c.8656C>T(p.Gln2886Ter)

[1784] NM_004006.2(DMD):c.8944C>T(p.Arg2982Ter)

[1785] NM_004006.2(DMD):c.5899C>T(p.Arg1967Ter)

[1786] NM_004006.2(DMD):c.10033C>T(p.Arg3345Ter)

[1787] NM_004006.2(DMD):c.10086+1G>A

[1788] NM_004019.2(DMD):c.1020G>A(p.Thr340=)

[1789] NM_004006.2(DMD):c.1261C>T(p.Gln421Ter)

[1790] NM_004006.2(DMD):c.1465C>T(p.Gln489Ter)

[1791] NM_004006.2(DMD):c.1990C>T(p.Gln664Ter)

[1792] NM_004006.2(DMD):c.2032C>T(p.Gln678Ter)

[1793] NM_004006.2(DMD):c.2332C>T(p.Gln778Ter)

[1794] NM_004006.2(DMD):c.2419C>T(p.Gln807Ter)

[1795] NM_004006.2(DMD):c.2650C>T(p.Gln884Ter)

[1796] NM_004006.2(DMD):c.2804-1G>A

[1797] NM_004006.2(DMD):c.3276+1G>A

[1798] NM_004006.2(DMD):c.3295C>T(p.Gln1099Ter)

[1799] NM_004006.2(DMD):c.336G>A(p.Trp112Ter)

[1800] NM_004006.2(DMD):c.3580C>T(p.Gln1194Ter)

[1801] NM_004006.2(DMD):c.4117C>T(p.Gln1373Ter)

[1802] NM_004006.2(DMD):c.649+1G>A

[1803] NM_004006.2(DMD):c.6906G>A(p.Trp2302Ter)

[1804] NM_004006.2(DMD):c.7189C>T(p.Gln2397Ter)

[1805] NM_004006.2(DMD):c.7309+1G>A

[1806] NM_004006.2(DMD):c.7657C>T(p.Arg2553Ter)

[1807] NM_004006.2(DMD):c.7682G>A(p.Trp2561Ter)

[1808] NM_004006.2(DMD):c.7683G>A(p.Trp2561Ter)

[1809] NM_004006.2(DMD):c.7894C>T(p.Gln2632Ter)

[1810] NM_004006.2(DMD):c.9361+1G>A

[1811] NM_004006.2(DMD):c.9564-1G>A

[1812] NM_004006.2(DMD):c.2956C>T(p.Gln986Ter)

[1813] NM_004006.2(DMD):c.883C>T(p.Arg295Ter)

[1814] NM_004006.2(DMD):c.31+36947G>A

[1815] NM_004006.2(DMD):c.10279C>T(p.Gln3427Ter)

[1816] NM_004006.2(DMD):c.433C>T(p.Arg145Ter)

[1817] NM_004006.2(DMD):c.9G>A(p.Trp3Ter)

[1818] NM_004006.2(DMD):c.10171C>T(p.Arg3391Ter)

[1819] NM_004006.2(DMD):c.583C>T(p.Arg195Ter)

[1820] NM_004006.2(DMD):c.9337C>T(p.Arg3113Ter)

[1821] NM_004006.2(DMD):c.8038C>T(p.Arg2680Ter)

[1822] NM_004006.2(DMD):c.1812+1G>A

[1823] NM_004006.2(DMD):c.1093C>T(p.Gln365Ter)

[1824] NM_004006.2(DMD):c.1704+1G>A

[1825] NM_004006.2(DMD):c.1912C>T(p.Gln638Ter)

[1826] NM_004006.2(DMD):c.133C>T(p.Gln45Ter)

[1827] NM_004006.2(DMD):c.5868G>A(p.Trp1956Ter)

[1828] NM_004006.2(DMD):c.565C>T(p.Gln189Ter)

[1829] NM_004006.2(DMD):c.5089C>T(p.Gln1697Ter)

[1830] NM_004006.2(DMD):c.2512C>T(p.Gln838Ter)

[1831] NM_004006.2(DMD):c.10477C>T(p.Gln3493Ter)

[1832] NM_004006.2(DMD):c.93+1G>A

[1833] NM_004006.2(DMD):c.4174C>T(p.Gln1392Ter)

[1834] NM_004006.2(DMD):c.3940C>T(p.Arg1314Ter)参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于DMD基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防贝克肌营养不良的方法。

[1835] 肢带型肌营养不良

[1836] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与肢带型肌营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP存在于选自SGCB、MYOT、LMNA、CAPN3、DYSF、SGCA、TTN、ANO5、TRAPPC11、LMNA、POMT1和FKRP的至少一种基因中,至少包括以下各项:

[1837] NM_000232.4(SGCB):c.31C>T(p.Gln11Ter)

[1838] NM_006790.2(MYOT):c.164C>T(p.Ser55Phe)

[1839] NM_006790.2(MYOT):c.170C>T(p.Thr57Ile)

[1840] NM_170707.3(LMNA):c.1488+1G>A

[1841] NM_170707.3(LMNA):c.1609-1G>A

[1842] NM_000070.2(CAPN3):c.1715G>A(p.Arg572Gln)

[1843] NM_000070.2(CAPN3):c.2243G>A(p.Arg748Gln)

[1844] NM_000070.2(CAPN3):c.145C>T(p.Arg49Cys)

[1845] NM_000070.2(CAPN3):c.1319G>A(p.Arg440Gln)

[1846] NM_000070.2(CAPN3):c.1343G>A(p.Arg448His)

[1847] NM_000070.2(CAPN3):c.1465C>T(p.Arg489Trp)

[1848] NM_000070.2(CAPN3):c.1714C>T(p.Arg572Trp)

[1849] NM_000070.2(CAPN3):c.2306G>A(p.Arg769Gln)

[1850] NM_000070.2(CAPN3):c.133G>A(p.Ala45Thr)

[1851] NM_000070.2(CAPN3):c.499-1G>A

[1852] NM_000070.2(CAPN3):c.439C>T(p.Arg147Ter)

[1853] NM_000070.2(CAPN3):c.1063C>T(p.Arg355Trp)

[1854] NM_000070.2(CAPN3):c.1250C>T(p.Thr417Met)

[1855] NM_000070.2(CAPN3):c.245C>T(p.Pro82Leu)

[1856] NM_000070.2(CAPN3):c.2242C>T(p.Arg748Ter)

[1857] NM_000070.2(CAPN3):c.1318C>T(p.Arg440Trp)

[1858] NM_000070.2(CAPN3):c.1333G>A(p.Gly445Arg)

[1859] NM_000070.2(CAPN3):c.1957C>T(p.Gln653Ter)

[1860] NM_000070.2(CAPN3):c.1801-1G>A

[1861] NM_000070.2(CAPN3):c.2263+1G>A

[1862] NM_000070.2(CAPN3):c.956C>T(p.Pro319Leu)

[1863] NM_000070.2(CAPN3):c.1468C>T(p.Arg490Trp)

[1864] NM_000070.2(CAPN3):c.802-9G>A

[1865] NM_000070.2(CAPN3):c.1342C>T(p.Arg448Cys)

[1866] NM_000070.2(CAPN3):c.1303G>A(p.Glu435Lys)

[1867] NM_000070.2(CAPN3):c.1993-1G>A

[1868] NM_003494.3(DYSF):c.3113G>A(p.Arg1038Gln)

[1869] NM_001130987.1(DYSF):c.5174+1G>A

[1870] NM_001130987.1(DYSF):c.159G>A(p.Trp53Ter)

[1871] NM_001130987.1(DYSF):c.2929C>T(p.Arg977Trp)

[1872] NM_001130987.1(DYSF):c.4282C>T(p.Gln1428Ter)

[1873] NM_001130987.1(DYSF):c.1577-1G>A

[1874] NM_003494.3(DYSF):c.5529G>A(p.Trp1843Ter)

[1875] NM_001130987.1(DYSF):c.1576+1G>A

[1876] NM_001130987.1(DYSF):c.4462C>T(p.Gln1488Ter)

[1877] NM_003494.3(DYSF):c.5429G>A(p.Arg1810Lys)

[1878] NM_003494.3(DYSF):c.5077C>T(p.Arg1693Trp)

[1879] NM_001130978.1(DYSF):c.1813C>T(p.Gln605Ter)

[1880] NM_003494.3(DYSF):c.3230G>A(p.Trp1077Ter)

[1881] NM_003494.3(DYSF):c.265C>T(p.Arg89Ter)

[1882] NM_003494.3(DYSF):c.4434G>A(p.Trp1478Ter)

[1883] NM_003494.3(DYSF):c.3478C>T(p.Gln1160Ter)

[1884] NM_001130987.1(DYSF):c.1372G>A(p.Gly458Arg)

[1885] NM_003494.3(DYSF):c.4090C>T(p.Gln1364Ter)

[1886] NM_001130987.1(DYSF):c.2409+1G>A

[1887] NM_003494.3(DYSF):c.1708C>T(p.Gln570Ter)

[1888] NM_003494.3(DYSF):c.1956G>A(p.Trp652Ter)

[1889] NM_001130987.1(DYSF):c.5004-1G>A

[1890] NM_003494.3(DYSF):c.331C>T(p.Gln111Ter)

[1891] NM_001130978.1(DYSF):c.5776C>T(p.Arg1926Ter)

[1892] NM_003494.3(DYSF):c.6124C>T(p.Arg2042Cys)

[1893] NM_003494.3(DYSF):c.2643+1G>A

[1894] NM_003494.3(DYSF):c.4253G>A(p.Gly1418Asp)

[1895] NM_003494.3(DYSF):c.610C>T(p.Arg204Ter)

[1896] NM_003494.3(DYSF):c.1834C>T(p.Gln612Ter)

[1897] NM_003494.3(DYSF):c.5668-7G>A

[1898] NM_001130978.1(DYSF):c.3137G>A(p.Arg1046His)

[1899] NM_003494.3(DYSF):c.1053+1G>A

[1900] NM_003494.3(DYSF):c.1398-1G>A

[1901] NM_003494.3(DYSF):c.1481-1G>A

[1902] NM_003494.3(DYSF):c.2311C>T(p.Gln771Ter)

[1903] NM_003494.3(DYSF):c.2869C>T(p.Gln957Ter)

[1904] NM_003494.3(DYSF):c.4756C>T(p.Arg1586Ter)

[1905] NM_003494.3(DYSF):c.5509G>A(p.Asp1837Asn)

[1906] NM_003494.3(DYSF):c.5644C>T(p.Gln1882Ter)

[1907] NM_003494.3(DYSF):c.5946+1G>A

[1908] NM_003494.3(DYSF):c.937+1G>A

[1909] NM_003494.3(DYSF):c.5266C>T(p.Gln1756Ter)

[1910] NM_003494.3(DYSF):c.3832C>T(p.Gln1278Ter)

[1911] NM_003494.3(DYSF):c.5525+1G>A

[1912] NM_003494.3(DYSF):c.3112C>T(p.Arg1038Ter)

[1913] NM_000023.3(SGCA):c.293G>A(p.Arg98His)

[1914] NM_000023.3(SGCA):c.850C>T(p.Arg284Cys)

[1915] NM_000023.3(SGCA):c.403C>T(p.Gln135Ter)

[1916] NM_000023.3(SGCA):c.409G>A(p.Glu137Lys)

[1917] NM_000023.3(SGCA):c.747+1G>A

[1918] NM_000023.3(SGCA):c.229C>T(p.Arg77Cys)

[1919] NM_000023.3(SGCA):c.101G>A(p.Arg34His)

[1920] NM_000023.3(SGCA):c.739G>A(p.Val247Met)

[1921] NM_001256850.1(TTN):c.87394C>T(p.Arg29132Ter)

[1922] NM_213599.2(ANO5):c.762+1G>A

[1923] NM_213599.2(ANO5):c.1213C>T(p.Gln405Ter)

[1924] NM_213599.2(ANO5):c.1639C>T(p.Arg547Ter)

[1925] NM_213599.2(ANO5):c.1406G>A(p.Trp469Ter)

[1926] NM_213599.2(ANO5):c.1210C>T(p.Arg404Ter)

[1927] NM_213599.2(ANO5):c.2272C>T(p.Arg758Cys)

[1928] NM_213599.2(ANO5):c.41-1G>A

[1929] NM_213599.2(ANO5):c.172C>T(p.Arg58Trp)

[1930] NM_213599.2(ANO5):c.1898+1G>A

[1931] NM_021942.5(TRAPPC11):c.1287+5G>A

[1932] NM_170707.3(LMNA):c.1608+1G>A

[1933] NM_007171.3(POMT1):c.1864C>T(p.Arg622Ter)

[1934] NM_024301.4(FKRP):c.313C>T(p.Gln105Ter)

[1935] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自SGCB、MYOT、LMNA、CAPN3、DYSF、SGCA、TTN、ANO5、TRAPPC11、LMNA、POMT1和FKRP的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防肢带型肌营养不良的方法。

[1936] 埃-德二氏肌营养不良

[1937] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与埃-德二氏肌营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于EMD或SYNE1基因中，至少包括以下各项：

[1938] NM_000117.2(EMD):c.3G>A(p.Met1Ile)

[1939] NM_033071.3(SYNE1):c.11908C>T(p.Arg3970Ter)

[1940] NM_033071.3(SYNE1):c.21721C>T(p.Gln7241Ter)

[1941] NM_000117.2(EMD):c.130C>T(p.Gln44Ter)

[1942] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于EMD或SYNE1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防埃-德二氏肌营养不良的方法。

[1943] 面肩肱型肌营养不良

[1944] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与面肩肱型肌营养不良相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于SMCHD1基因中，至少包括以下各项：

[1945] NM_015295.2(SMCHD1):c.3801+1G>A

[1946] NM_015295.2(SMCHD1):c.1843-1G>A

[1947] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于SMCHD1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防面肩肱型肌营养不良的方法。

[1948] 先天性代谢缺损(IEM)

[1949] 与各种IEM相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于原发性高草酸尿症1型、精氨酸琥珀酸裂解酶缺乏症、鸟氨酸氨甲酰基转移酶缺乏症和枫糖尿病。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[1950] 原发性高草酸尿症1型

[1951] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与原发性高草酸尿症1型相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于AGXT基因中，至少包括以下各项：

[1952] NM_000030.2(AGXT):c.245G>A(p.Gly82Glu)

[1953] NM_000030.2(AGXT):c.698G>A(p.Arg233His)

[1954] NM_000030.2(AGXT):c.466G>A(p.Gly156Arg)

[1955] NM_000030.2(AGXT):c.106C>T(p.Arg36Cys)

[1956] NM_000030.2(AGXT):c.346G>A(p.Gly116Arg)

[1957] NM_000030.2(AGXT):c.568G>A(p.Gly190Arg)

[1958] NM_000030.2(AGXT):c.653C>T(p.Ser218Leu)

[1959] NM_000030.2(AGXT):c.737G>A(p.Trp246Ter)

[1960] NM_000030.2(AGXT):c.1049G>A(p.Gly350Asp)

[1961] NM_000030.2(AGXT):c.473C>T(p.Ser158Leu)

[1962] NM_000030.2(AGXT):c.907C>T(p.Gln303Ter)

[1963] NM_000030.2(AGXT):c.996G>A(p.Trp332Ter)

[1964] NM_000030.2(AGXT):c.508G>A(p.Gly170Arg)

[1965] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于AGXT基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防原发性高草酸尿症1型的方法。

[1966] 精氨酸琥珀酸裂解酶缺乏症

[1967] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与精氨酸琥珀酸裂解酶缺乏症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于ASL基因中,至少包括以下各项:

[1968] NM_001024943.1(ASL):c.1153C>T(p.Arg385Cys)

[1969] NM_000048.3(ASL):c.532G>A(p.Val178Met)

[1970] NM_000048.3(ASL):c.545G>A(p.Arg182Gln)

[1971] NM_000048.3(ASL):c.175G>A(p.Glu59Lys)

[1972] NM_000048.3(ASL):c.718+5G>A

[1973] NM_000048.3(ASL):c.889C>T(p.Arg297Trp)

[1974] NM_000048.3(ASL):c.1360C>T(p.Gln454Ter)

[1975] NM_000048.3(ASL):c.1060C>T(p.Gln354Ter)

[1976] NM_000048.3(ASL):c.35G>A(p.Arg12Gln)

[1977] NM_000048.3(ASL):c.446+1G>A

[1978] NM_000048.3(ASL):c.544C>T(p.Arg182Ter)

[1979] NM_000048.3(ASL):c.1135C>T(p.Arg379Cys)

[1980] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于ASL基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防面肩肱型肌营养不良的方法。

[1981] 鸟氨酸氨甲酰基转移酶缺乏症

[1982] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与鸟氨酸氨甲酰基转移酶缺乏症相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于OTC基因中,至少包括以下各项:

[1983] NM_000531.5(OTC):c.119G>A(p.Arg40His)

[1984] NM_000531.5(OTC):c.422G>A(p.Arg141Gln)

[1985] NM_000531.5(OTC):c.829C>T(p.Arg277Trp)

[1986] NM_000531.5(OTC):c.674C>T(p.Pro225Leu)

[1987] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A

或C至T突变/SNP、特别地存在于OTC基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防鸟氨酸氨甲酰基转移酶缺乏症的方法。

[1988] 枫糖尿病

[1989] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与枫糖尿病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自BCKDHA、BCKDHB、DBT和DLD的至少一种基因中，至少包括以下各项：

[1990] NM_000709.3(BCKDHA):c.476G>A(p.Arg159Gln)

[1991] NM_183050.3(BCKDHB):c.3G>A(p.Met1Ile)

[1992] NM_183050.3(BCKDHB):c.554C>T(p.Pro185Leu)

[1993] NM_001918.3(DBT):c.1033G>A(p.Gly345Arg)

[1994] NM_000709.3(BCKDHA):c.940C>T(p.Arg314Ter)

[1995] NM_000709.3(BCKDHA):c.793C>T(p.Arg265Trp)

[1996] NM_000709.3(BCKDHA):c.868G>A(p.Gly290Arg)

[1997] NM_000108.4(DLD):c.1123G>A(p.Glu375Lys)

[1998] NM_000709.3(BCKDHA):c.1234G>A(p.Val412Met)

[1999] NM_000709.3(BCKDHA):c.288+1G>A

[2000] NM_000709.3(BCKDHA):c.979G>A(p.Glu327Lys)

[2001] NM_001918.3(DBT):c.901C>T(p.Arg301Cys)

[2002] NM_183050.3(BCKDHB):c.509G>A(p.Arg170His)

[2003] NM_183050.3(BCKDHB):c.799C>T(p.Gln267Ter)

[2004] NM_183050.3(BCKDHB):c.853C>T(p.Arg285Ter)

[2005] NM_183050.3(BCKDHB):c.970C>T(p.Arg324Ter)

[2006] NM_183050.3(BCKDHB):c.832G>A(p.Gly278Ser)

[2007] NM_000709.3(BCKDHA):c.1036C>T(p.Arg346Cys)

[2008] NM_000709.3(BCKDHA):c.288+9C>T

[2009] NM_000709.3(BCKDHA):c.632C>T(p.Thr211Met)

[2010] NM_000709.3(BCKDHA):c.659C>T(p.Ala220Val)

[2011] NM_000709.3(BCKDHA):c.964C>T(p.Gln322Ter)

[2012] NM_001918.3(DBT):c.1291C>T(p.Arg431Ter)

[2013] NM_001918.3(DBT):c.251G>A(p.Trp84Ter)

[2014] NM_001918.3(DBT):c.871C>T(p.Arg291Ter)

[2015] NM_000056.4(BCKDHB):c.1016C>T(p.Ser339Leu)

[2016] NM_000056.4(BCKDHB):c.344-1G>A

[2017] NM_000056.4(BCKDHB):c.633+1G>A

[2018] NM_000056.4(BCKDHB):c.952-1G>A

[2019] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自BCKDHA、BCKDHB、DBT和DLD的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突

变/SNP来治疗或预防枫糖尿病的方法。

[2020] 癌症相关疾病

[2021] 与各种癌症和癌症相关疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于乳腺癌-卵巢癌和林奇综合征(Lynch syndrome)。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[2022] 乳腺癌-卵巢癌

[2023] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与乳腺癌-卵巢癌相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于BRCA1或BRCA2基因中，至少包括以下各项：

[2024] NM_007294.3(BRCA1):c.5095C>T(p.Arg1699Trp)

[2025] NM_000059.3(BRCA2):c.7558C>T(p.Arg2520Ter)

[2026] NM_007294.3(BRCA1):c.2572C>T(p.Gln858Ter)

[2027] NM_007294.3(BRCA1):c.3607C>T(p.Arg1203Ter)

[2028] NM_007294.3(BRCA1):c.5503C>T(p.Arg1835Ter)

[2029] NM_007294.3(BRCA1):c.2059C>T(p.Gln687Ter)

[2030] NM_007294.3(BRCA1):c.4675+1G>A

[2031] NM_007294.3(BRCA1):c.5251C>T(p.Arg1751Ter)

[2032] NM_007294.3(BRCA1):c.5444G>A(p.Trp1815Ter)

[2033] NM_000059.3(BRCA2):c.9318G>A(p.Trp3106Ter)

[2034] NM_000059.3(BRCA2):c.9382C>T(p.Arg3128Ter)

[2035] NM_000059.3(BRCA2):c.274C>T(p.Gln92Ter)

[2036] NM_000059.3(BRCA2):c.6952C>T(p.Arg2318Ter)

[2037] NM_007294.3(BRCA1):c.1687C>T(p.Gln563Ter)

[2038] NM_007294.3(BRCA1):c.2599C>T(p.Gln867Ter)

[2039] NM_007294.3(BRCA1):c.784C>T(p.Gln262Ter)

[2040] NM_007294.3(BRCA1):c.280C>T(p.Gln94Ter)

[2041] NM_007294.3(BRCA1):c.5542C>T(p.Gln1848Ter)

[2042] NM_007294.3(BRCA1):c.5161C>T(p.Gln1721Ter)

[2043] NM_007294.3(BRCA1):c.4573C>T(p.Gln1525Ter)

[2044] NM_007294.3(BRCA1):c.4270C>T(p.Gln1424Ter)

[2045] NM_007294.3(BRCA1):c.4225C>T(p.Gln1409Ter)

[2046] NM_007294.3(BRCA1):c.4066C>T(p.Gln1356Ter)

[2047] NM_007294.3(BRCA1):c.3679C>T(p.Gln1227Ter)

[2048] NM_007294.3(BRCA1):c.1918C>T(p.Gln640Ter)

[2049] NM_007294.3(BRCA1):c.963G>A(p.Trp321Ter)

[2050] NM_007294.3(BRCA1):c.718C>T(p.Gln240Ter)

[2051] NM_000059.3(BRCA2):c.9196C>T(p.Gln3066Ter)

[2052] NM_000059.3(BRCA2):c.9154C>T(p.Arg3052Trp)

[2053] NM_007294.3(BRCA1):c.3991C>T(p.Gln1331Ter)

[2054] NM_007294.3(BRCA1):c.4097-1G>A

[2055] NM_007294.3(BRCA1):c.1059G>A(p.Trp353Ter)

[2056] NM_007294.3(BRCA1):c.1115G>A(p.Trp372Ter)

[2057] NM_007294.3(BRCA1):c.1138C>T(p.Gln380Ter)

[2058] NM_007294.3(BRCA1):c.1612C>T(p.Gln538Ter)

[2059] NM_007294.3(BRCA1):c.1621C>T(p.Gln541Ter)

[2060] NM_007294.3(BRCA1):c.1630C>T(p.Gln544Ter)

[2061] NM_007294.3(BRCA1):c.178C>T(p.Gln60Ter)

[2062] NM_007294.3(BRCA1):c.1969C>T(p.Gln657Ter)

[2063] NM_007294.3(BRCA1):c.2275C>T(p.Gln759Ter)

[2064] NM_007294.3(BRCA1):c.2410C>T(p.Gln804Ter)

[2065] NM_007294.3(BRCA1):c.2869C>T(p.Gln957Ter)

[2066] NM_007294.3(BRCA1):c.2923C>T(p.Gln975Ter)

[2067] NM_007294.3(BRCA1):c.3268C>T(p.Gln1090Ter)

[2068] NM_007294.3(BRCA1):c.3430C>T(p.Gln1144Ter)

[2069] NM_007294.3(BRCA1):c.3544C>T(p.Gln1182Ter)

[2070] NM_007294.3(BRCA1):c.4075C>T(p.Gln1359Ter)

[2071] NM_007294.3(BRCA1):c.4201C>T(p.Gln1401Ter)

[2072] NM_007294.3(BRCA1):c.4399C>T(p.Gln1467Ter)

[2073] NM_007294.3(BRCA1):c.4552C>T(p.Gln1518Ter)

[2074] NM_007294.3(BRCA1):c.5054C>T(p.Thr1685Ile)

[2075] NM_007294.3(BRCA1):c.514C>T(p.Gln172Ter)

[2076] NM_007294.3(BRCA1):c.5239C>T(p.Gln1747Ter)

[2077] NM_007294.3(BRCA1):c.5266C>T(p.Gln1756Ter)

[2078] NM_007294.3(BRCA1):c.5335C>T(p.Gln1779Ter)

[2079] NM_007294.3(BRCA1):c.5345G>A(p.Trp1782Ter)

[2080] NM_007294.3(BRCA1):c.5511G>A(p.Trp1837Ter)

[2081] NM_007294.3(BRCA1):c.5536C>T(p.Gln1846Ter)

[2082] NM_007294.3(BRCA1):c.55C>T(p.Gln19Ter)

[2083] NM_007294.3(BRCA1):c.949C>T(p.Gln317Ter)

[2084] NM_007294.3(BRCA1):c.928C>T(p.Gln310Ter)

[2085] NM_007294.3(BRCA1):c.5117G>A(p.Gly1706Glu)

[2086] NM_007294.3(BRCA1):c.5136G>A(p.Trp1712Ter)

[2087] NM_007294.3(BRCA1):c.4327C>T(p.Arg1443Ter)

[2088] NM_007294.3(BRCA1):c.1471C>T(p.Gln491Ter)

[2089] NM_007294.3(BRCA1):c.1576C>T(p.Gln526Ter)

[2090] NM_007294.3(BRCA1):c.160C>T(p.Gln54Ter)

[2091] NM_007294.3(BRCA1):c.2683C>T(p.Gln895Ter)

[2092] NM_007294.3(BRCA1):c.2761C>T(p.Gln921Ter)

[2093] NM_007294.3(BRCA1):c.3895C>T(p.Gln1299Ter)

[2094] NM_007294.3(BRCA1):c.4339C>T(p.Gln1447Ter)

[2095] NM_007294.3(BRCA1):c.4372C>T(p.Gln1458Ter)

[2096] NM_007294.3(BRCA1):c.5153G>A(p.Trp1718Ter)

[2097] NM_007294.3(BRCA1):c.5445G>A(p.Trp1815Ter)

[2098] NM_007294.3(BRCA1):c.5510G>A(p.Trp1837Ter)

[2099] NM_007294.3(BRCA1):c.5346G>A(p.Trp1782Ter)

[2100] NM_007294.3(BRCA1):c.1116G>A(p.Trp372Ter)

[2101] NM_007294.3(BRCA1):c.1999C>T(p.Gln667Ter)

[2102] NM_007294.3(BRCA1):c.4183C>T(p.Gln1395Ter)

[2103] NM_007294.3(BRCA1):c.4810C>T(p.Gln1604Ter)

[2104] NM_007294.3(BRCA1):c.850C>T(p.Gln284Ter)

[2105] NM_007294.3(BRCA1):c.1058G>A(p.Trp353Ter)

[2106] NM_007294.3(BRCA1):c.131G>A(p.Cys44Tyr)

[2107] NM_007294.3(BRCA1):c.1600C>T(p.Gln534Ter)

[2108] NM_007294.3(BRCA1):c.3286C>T(p.Gln1096Ter)

[2109] NM_007294.3(BRCA1):c.3403C>T(p.Gln1135Ter)

[2110] NM_007294.3(BRCA1):c.34C>T(p.Gln12Ter)

[2111] NM_007294.3(BRCA1):c.4258C>T(p.Gln1420Ter)

[2112] NM_007294.3(BRCA1):c.4609C>T(p.Gln1537Ter)

[2113] NM_007294.3(BRCA1):c.5154G>A(p.Trp1718Ter)

[2114] NM_007294.3(BRCA1):c.5431C>T(p.Gln1811Ter)

[2115] NM_007294.3(BRCA1):c.241C>T(p.Gln81Ter)

[2116] NM_007294.3(BRCA1):c.3331C>T(p.Gln1111Ter)

[2117] NM_007294.3(BRCA1):c.3967C>T(p.Gln1323Ter)

[2118] NM_007294.3(BRCA1):c.415C>T(p.Gln139Ter)

[2119] NM_007294.3(BRCA1):c.505C>T(p.Gln169Ter)

[2120] NM_007294.3(BRCA1):c.5194-12G>A

[2121] NM_007294.3(BRCA1):c.5212G>A(p.Gly1738Arg)

[2122] NM_007294.3(BRCA1):c.5332+1G>A

[2123] NM_007294.3(BRCA1):c.1480C>T(p.Gln494Ter)

[2124] NM_007294.3(BRCA1):c.2563C>T(p.Gln855Ter)

[2125] NM_007294.3(BRCA1):c.1066C>T(p.Gln356Ter)

[2126] NM_007294.3(BRCA1):c.3718C>T(p.Gln1240Ter)

[2127] NM_007294.3(BRCA1):c.3817C>T(p.Gln1273Ter)

[2128] NM_007294.3(BRCA1):c.3937C>T(p.Gln1313Ter)

[2129] NM_007294.3(BRCA1):c.4357+1G>A

[2130] NM_007294.3(BRCA1):c.5074+1G>A

[2131] NM_007294.3(BRCA1):c.5277+1G>A

[2132] NM_007294.3(BRCA1):c.2338C>T(p.Gln780Ter)

[2133] NM_007294.3(BRCA1):c.3598C>T(p.Gln1200Ter)

[2134] NM_007294.3(BRCA1):c.3841C>T(p.Gln1281Ter)

[2135] NM_007294.3(BRCA1):c.4222C>T(p.Gln1408Ter)

[2136] NM_007294.3(BRCA1):c.4524G>A(p.Trp1508Ter)

[2137] NM_007294.3(BRCA1):c.5353C>T(p.Gln1785Ter)

[2138] NM_007294.3(BRCA1):c.962G>A(p.Trp321Ter)

[2139] NM_007294.3(BRCA1):c.220C>T(p.Gln74Ter)

[2140] NM_007294.3(BRCA1):c.2713C>T(p.Gln905Ter)

[2141] NM_007294.3(BRCA1):c.2800C>T(p.Gln934Ter)

[2142] NM_007294.3(BRCA1):c.4612C>T(p.Gln1538Ter)

[2143] NM_007294.3(BRCA1):c.3352C>T(p.Gln1118Ter)

[2144] NM_007294.3(BRCA1):c.4834C>T(p.Gln1612Ter)

[2145] NM_007294.3(BRCA1):c.4523G>A(p.Trp1508Ter)

[2146] NM_007294.3(BRCA1):c.5135G>A(p.Trp1712Ter)

[2147] NM_007294.3(BRCA1):c.1155G>A(p.Trp385Ter)

[2148] NM_007294.3(BRCA1):c.4987-1G>A

[2149] NM_000059.3(BRCA2):c.9573G>A(p.Trp3191Ter)

[2150] NM_000059.3(BRCA2):c.1945C>T(p.Gln649Ter)

[2151] NM_000059.3(BRCA2):c.217C>T(p.Gln73Ter)

[2152] NM_000059.3(BRCA2):c.523C>T(p.Gln175Ter)

[2153] NM_000059.3(BRCA2):c.2548C>T(p.Gln850Ter)

[2154] NM_000059.3(BRCA2):c.2905C>T(p.Gln969Ter)

[2155] NM_000059.3(BRCA2):c.4689G>A(p.Trp1563Ter)

[2156] NM_000059.3(BRCA2):c.4972C>T(p.Gln1658Ter)

[2157] NM_000059.3(BRCA2):c.1184G>A(p.Trp395Ter)

[2158] NM_000059.3(BRCA2):c.2137C>T(p.Gln713Ter)

[2159] NM_000059.3(BRCA2):c.3217C>T(p.Gln1073Ter)

[2160] NM_000059.3(BRCA2):c.3523C>T(p.Gln1175Ter)

[2161] NM_000059.3(BRCA2):c.4783C>T(p.Gln1595Ter)

[2162] NM_000059.3(BRCA2):c.5800C>T(p.Gln1934Ter)

[2163] NM_000059.3(BRCA2):c.6478C>T(p.Gln2160Ter)

[2164] NM_000059.3(BRCA2):c.7033C>T(p.Gln2345Ter)

[2165] NM_000059.3(BRCA2):c.7495C>T(p.Gln2499Ter)

[2166] NM_000059.3(BRCA2):c.7501C>T(p.Gln2501Ter)

[2167] NM_000059.3(BRCA2):c.7887G>A(p.Trp2629Ter)

[2168] NM_000059.3(BRCA2):c.8910G>A(p.Trp2970Ter)

[2169] NM_000059.3(BRCA2):c.9139C>T(p.Gln3047Ter)

[2170] NM_000059.3(BRCA2):c.9739C>T(p.Gln3247Ter)

[2171] NM_000059.3(BRCA2):c.582G>A(p.Trp194Ter)

[2172] NM_000059.3(BRCA2):c.7963C>T(p.Gln2655Ter)

[2173] NM_000059.3(BRCA2):c.8695C>T(p.Gln2899Ter)

[2174] NM_000059.3(BRCA2):c.8869C>T(p.Gln2957Ter)

[2175] NM_000059.3(BRCA2):c.1117C>T(p.Gln373Ter)

[2176] NM_000059.3(BRCA2):c.1825C>T(p.Gln609Ter)

[2177] NM_000059.3(BRCA2):c.2455C>T(p.Gln819Ter)

[2178] NM_000059.3(BRCA2):c.2881C>T(p.Gln961Ter)

[2179] NM_000059.3(BRCA2):c.3265C>T(p.Gln1089Ter)

[2180] NM_000059.3(BRCA2):c.3283C>T(p.Gln1095Ter)

[2181] NM_000059.3(BRCA2):c.3442C>T(p.Gln1148Ter)

[2182] NM_000059.3(BRCA2):c.3871C>T(p.Gln1291Ter)

[2183] NM_000059.3(BRCA2):c.439C>T(p.Gln147Ter)

[2184] NM_000059.3(BRCA2):c.4525C>T(p.Gln1509Ter)

[2185] NM_000059.3(BRCA2):c.475+1G>A

[2186] NM_000059.3(BRCA2):c.5344C>T(p.Gln1782Ter)

[2187] NM_000059.3(BRCA2):c.5404C>T(p.Gln1802Ter)

[2188] NM_000059.3(BRCA2):c.5773C>T(p.Gln1925Ter)

[2189] NM_000059.3(BRCA2):c.5992C>T(p.Gln1998Ter)

[2190] NM_000059.3(BRCA2):c.6469C>T(p.Gln2157Ter)

[2191] NM_000059.3(BRCA2):c.7261C>T(p.Gln2421Ter)

[2192] NM_000059.3(BRCA2):c.7303C>T(p.Gln2435Ter)

[2193] NM_000059.3(BRCA2):c.7471C>T(p.Gln2491Ter)

[2194] NM_000059.3(BRCA2):c.7681C>T(p.Gln2561Ter)

[2195] NM_000059.3(BRCA2):c.7738C>T(p.Gln2580Ter)

[2196] NM_000059.3(BRCA2):c.7886G>A(p.Trp2629Ter)

[2197] NM_000059.3(BRCA2):c.8140C>T(p.Gln2714Ter)

[2198] NM_000059.3(BRCA2):c.8363G>A(p.Trp2788Ter)

[2199] NM_000059.3(BRCA2):c.8572C>T(p.Gln2858Ter)

[2200] NM_000059.3(BRCA2):c.8773C>T(p.Gln2925Ter)

[2201] NM_000059.3(BRCA2):c.8821C>T(p.Gln2941Ter)

[2202] NM_000059.3(BRCA2):c.9109C>T(p.Gln3037Ter)

[2203] NM_000059.3(BRCA2):c.9317G>A(p.Trp3106Ter)

[2204] NM_000059.3(BRCA2):c.9466C>T(p.Gln3156Ter)

[2205] NM_000059.3(BRCA2):c.9572G>A(p.Trp3191Ter)

[2206] NM_000059.3(BRCA2):c.8490G>A(p.Trp2830Ter)

[2207] NM_000059.3(BRCA2):c.5980C>T(p.Gln1994Ter)

[2208] NM_000059.3(BRCA2):c.7721G>A(p.Trp2574Ter)

[2209] NM_000059.3(BRCA2):c.196C>T(p.Gln66Ter)

[2210] NM_000059.3(BRCA2):c.7618-1G>A

[2211] NM_000059.3(BRCA2):c.8489G>A(p.Trp2830Ter)

[2212] NM_000059.3(BRCA2):c.7857G>A(p.Trp2619Ter)

[2213] NM_000059.3(BRCA2):c.1261C>T(p.Gln421Ter)

[2214] NM_000059.3(BRCA2):c.1456C>T(p.Gln486Ter)

[2215] NM_000059.3(BRCA2):c.3319C>T(p.Gln1107Ter)

[2216] NM_000059.3(BRCA2):c.5791C>T(p.Gln1931Ter)

[2217] NM_000059.3(BRCA2):c.6070C>T(p.Gln2024Ter)

[2218] NM_000059.3(BRCA2):c.7024C>T(p.Gln2342Ter)

[2219] NM_000059.3(BRCA2):c.961C>T(p.Gln321Ter)

[2220] NM_000059.3(BRCA2):c.9380G>A(p.Trp3127Ter)

[2221] NM_000059.3(BRCA2):c.8364G>A(p.Trp2788Ter)

[2222] NM_000059.3(BRCA2):c.7758G>A(p.Trp2586Ter)

[2223] NM_000059.3(BRCA2):c.2224C>T(p.Gln742Ter)

[2224] NM_000059.3(BRCA2):c.5101C>T(p.Gln1701Ter)

[2225] NM_000059.3(BRCA2):c.5959C>T(p.Gln1987Ter)

[2226] NM_000059.3(BRCA2):c.7060C>T(p.Gln2354Ter)

[2227] NM_000059.3(BRCA2):c.9100C>T(p.Gln3034Ter)

[2228] NM_000059.3(BRCA2):c.9148C>T(p.Gln3050Ter)

[2229] NM_000059.3(BRCA2):c.9883C>T(p.Gln3295Ter)

[2230] NM_000059.3(BRCA2):c.1414C>T(p.Gln472Ter)

[2231] NM_000059.3(BRCA2):c.1689G>A(p.Trp563Ter)

[2232] NM_000059.3(BRCA2):c.581G>A(p.Trp194Ter)

[2233] NM_000059.3(BRCA2):c.6490C>T(p.Gln2164Ter)

[2234] NM_000059.3(BRCA2):c.7856G>A(p.Trp2619Ter)

[2235] NM_000059.3(BRCA2):c.8970G>A(p.Trp2990Ter)

[2236] NM_000059.3(BRCA2):c.92G>A(p.Trp31Ter)

[2237] NM_000059.3(BRCA2):c.9376C>T(p.Gln3126Ter)

[2238] NM_000059.3(BRCA2):c.93G>A(p.Trp31Ter)

[2239] NM_000059.3(BRCA2):c.1189C>T(p.Gln397Ter)

[2240] NM_000059.3(BRCA2):c.2818C>T(p.Gln940Ter)

[2241] NM_000059.3(BRCA2):c.2979G>A(p.Trp993Ter)

[2242] NM_000059.3(BRCA2):c.3166C>T(p.Gln1056Ter)

[2243] NM_000059.3(BRCA2):c.4285C>T(p.Gln1429Ter)

[2244] NM_000059.3(BRCA2):c.6025C>T(p.Gln2009Ter)

[2245] NM_000059.3(BRCA2):c.772C>T(p.Gln258Ter)

[2246] NM_000059.3(BRCA2):c.7877G>A(p.Trp2626Ter)

[2247] NM_000059.3(BRCA2):c.3109C>T(p.Gln1037Ter)

[2248] NM_000059.3(BRCA2):c.4222C>T(p.Gln1408Ter)

[2249] NM_000059.3(BRCA2):c.7480C>T(p.Arg2494Ter)

[2250] NM_000059.3(BRCA2):c.7878G>A(p.Trp2626Ter)

[2251] NM_000059.3(BRCA2):c.9076C>T(p.Gln3026Ter)

[2252] NM_000059.3(BRCA2):c.1855C>T(p.Gln619Ter)

[2253] NM_000059.3(BRCA2):c.4111C>T(p.Gln1371Ter)

[2254] NM_000059.3(BRCA2):c.5656C>T(p.Gln1886Ter)

[2255] NM_000059.3(BRCA2):c.7757G>A(p.Trp2586Ter)

[2256] NM_000059.3(BRCA2):c.8243G>A(p.Gly2748Asp)

[2257] NM_000059.3(BRCA2):c.8878C>T(p.Gln2960Ter)

[2258] NM_000059.3(BRCA2):c.8487+1G>A

[2259] NM_000059.3(BRCA2):c.8677C>T(p.Gln2893Ter)

[2260] NM_000059.3(BRCA2):c.250C>T(p.Gln84Ter)

[2261] NM_000059.3(BRCA2):c.6124C>T(p.Gln2042Ter)

[2262] NM_000059.3(BRCA2):c.7617+1G>A

[2263] NM_000059.3(BRCA2):c.8575C>T(p.Gln2859Ter)

[2264] NM_000059.3(BRCA2):c.8174G>A(p.Trp2725Ter)

[2265] NM_000059.3(BRCA2):c.3187C>T(p.Gln1063Ter)

[2266] NM_000059.3(BRCA2):c.9381G>A(p.Trp3127Ter)

[2267] NM_000059.3(BRCA2):c.2095C>T(p.Gln699Ter)

[2268] NM_000059.3(BRCA2):c.1642C>T(p.Gln548Ter)

[2269] NM_000059.3(BRCA2):c.8608C>T(p.Gln2870Ter)

[2270] NM_000059.3(BRCA2):c.3412C>T(p.Gln1138Ter)

[2271] NM_000059.3(BRCA2):c.4246C>T(p.Gln1416Ter)

[2272] NM_000059.3(BRCA2):c.6475C>T(p.Gln2159Ter)

[2273] NM_000059.3(BRCA2):c.7366C>T(p.Gln2456Ter)

[2274] NM_000059.3(BRCA2):c.7516C>T(p.Gln2506Ter)

[2275] NM_000059.3(BRCA2):c.8969G>A(p.Trp2990Ter)

[2276] NM_000059.3(BRCA2):c.6487C>T(p.Gln2163Ter)

[2277] NM_000059.3(BRCA2):c.2978G>A(p.Trp993Ter)

[2278] NM_000059.3(BRCA2):c.7615C>T(p.Gln2539Ter)

[2279] NM_000059.3(BRCA2):c.9106C>T(p.Gln3036Ter)

[2280] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于BRCA1或BRCA2基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防乳腺癌-卵巢癌的方法。

[2281] 林奇综合征

[2282] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与林奇综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于

选自MSH6、MSH2、EPCAM、PMS2和MLH1的至少一种基因中，至少包括以下各项：

[2283] NM_000179.2(MSH6):c.1045C>T(p.Gln349Ter)

[2284] NM_000251.2(MSH2):c.1384C>T(p.Gln462Ter)

[2285] NM_002354.2(EPCAM):c.133C>T(p.Gln45Ter)

[2286] NM_002354.2(EPCAM):c.429G>A(p.Trp143Ter)

[2287] NM_002354.2(EPCAM):c.523C>T(p.Gln175Ter)

[2288] NM_000179.2(MSH6):c.2680C>T(p.Gln894Ter)

[2289] NM_000251.2(MSH2):c.350G>A(p.Trp117Ter)

[2290] NM_000179.2(MSH6):c.2735G>A(p.Trp912Ter)

[2291] NM_000179.2(MSH6):c.3556+1G>A

[2292] NM_000251.2(MSH2):c.388C>T(p.Gln130Ter)

[2293] NM_000535.6(PMS2):c.1912C>T(p.Gln638Ter)

[2294] NM_000535.6(PMS2):c.1891C>T(p.Gln631Ter)

[2295] NM_000249.3(MLH1):c.454-1G>A

[2296] NM_000251.2(MSH2):c.1030C>T(p.Gln344Ter)

[2297] NM_000179.2(MSH6):c.2330G>A(p.Trp777Ter)

[2298] NM_000179.2(MSH6):c.2191C>T(p.Gln731Ter)

[2299] NM_000179.2(MSH6):c.2764C>T(p.Arg922Ter)

[2300] NM_000179.2(MSH6):c.2815C>T(p.Gln939Ter)

[2301] NM_000179.2(MSH6):c.3020G>A(p.Trp1007Ter)

[2302] NM_000179.2(MSH6):c.3436C>T(p.Gln1146Ter)

[2303] NM_000179.2(MSH6):c.3647-1G>A

[2304] NM_000179.2(MSH6):c.3772C>T(p.Gln1258Ter)

[2305] NM_000179.2(MSH6):c.3838C>T(p.Gln1280Ter)

[2306] NM_000179.2(MSH6):c.706C>T(p.Gln236Ter)

[2307] NM_000179.2(MSH6):c.730C>T(p.Gln244Ter)

[2308] NM_000249.3(MLH1):c.1171C>T(p.Gln391Ter)

[2309] NM_000249.3(MLH1):c.1192C>T(p.Gln398Ter)

[2310] NM_000249.3(MLH1):c.1225C>T(p.Gln409Ter)

[2311] NM_000249.3(MLH1):c.1276C>T(p.Gln426Ter)

[2312] NM_000249.3(MLH1):c.1528C>T(p.Gln510Ter)

[2313] NM_000249.3(MLH1):c.1609C>T(p.Gln537Ter)

[2314] NM_000249.3(MLH1):c.1613G>A(p.Trp538Ter)

[2315] NM_000249.3(MLH1):c.1614G>A(p.Trp538Ter)

[2316] NM_000249.3(MLH1):c.1624C>T(p.Gln542Ter)

[2317] NM_000249.3(MLH1):c.1684C>T(p.Gln562Ter)

[2318] NM_000249.3(MLH1):c.1731+1G>A

[2319] NM_000249.3(MLH1):c.1731+5G>A

[2320] NM_000249.3(MLH1):c.1732-1G>A

[2321] NM_000249.3(MLH1):c.1896G>A(p.Glu632=)

[2322] NM_000249.3(MLH1):c.1989+1G>A

[2323] NM_000249.3(MLH1):c.1990-1G>A

[2324] NM_000249.3(MLH1):c.1998G>A(p.Trp666Ter)

[2325] NM_000249.3(MLH1):c.208-1G>A

[2326] NM_000249.3(MLH1):c.2101C>T(p.Gln701Ter)

[2327] NM_000249.3(MLH1):c.2136G>A(p.Trp712Ter)

[2328] NM_000249.3(MLH1):c.2224C>T(p.Gln742Ter)

[2329] NM_000249.3(MLH1):c.230G>A(p.Cys77Tyr)

[2330] NM_000249.3(MLH1):c.256C>T(p.Gln86Ter)

[2331] NM_000249.3(MLH1):c.436C>T(p.Gln146Ter)

[2332] NM_000249.3(MLH1):c.445C>T(p.Gln149Ter)

[2333] NM_000249.3(MLH1):c.545G>A(p.Arg182Lys)

[2334] NM_000249.3(MLH1):c.731G>A(p.Gly244Asp)

[2335] NM_000249.3(MLH1):c.76C>T(p.Gln26Ter)

[2336] NM_000249.3(MLH1):c.842C>T(p.Ala281Val)

[2337] NM_000249.3(MLH1):c.882C>T(p.Leu294=)

[2338] NM_000249.3(MLH1):c.901C>T(p.Gln301Ter)

[2339] NM_000251.2(MSH2):c.1013G>A(p.Gly338Glu)

[2340] NM_000251.2(MSH2):c.1034G>A(p.Trp345Ter)

[2341] NM_000251.2(MSH2):c.1129C>T(p.Gln377Ter)

[2342] NM_000251.2(MSH2):c.1183C>T(p.Gln395Ter)

[2343] NM_000251.2(MSH2):c.1189C>T(p.Gln397Ter)

[2344] NM_000251.2(MSH2):c.1204C>T(p.Gln402Ter)

[2345] NM_000251.2(MSH2):c.1276+1G>A

[2346] NM_000251.2(MSH2):c.1528C>T(p.Gln510Ter)

[2347] NM_000251.2(MSH2):c.1552C>T(p.Gln518Ter)

[2348] NM_000251.2(MSH2):c.1720C>T(p.Gln574Ter)

[2349] NM_000251.2(MSH2):c.1777C>T(p.Gln593Ter)

[2350] NM_000251.2(MSH2):c.1885C>T(p.Gln629Ter)

[2351] NM_000251.2(MSH2):c.2087C>T(p.Pro696Leu)

[2352] NM_000251.2(MSH2):c.2251G>A(p.Gly751Arg)

[2353] NM_000251.2(MSH2):c.2291G>A(p.Trp764Ter)

[2354] NM_000251.2(MSH2):c.2292G>A(p.Trp764Ter)

[2355] NM_000251.2(MSH2):c.2446C>T(p.Gln816Ter)

[2356] NM_000251.2(MSH2):c.2470C>T(p.Gln824Ter)

[2357] NM_000251.2(MSH2):c.2536C>T(p.Gln846Ter)

[2358] NM_000251.2(MSH2):c.2581C>T(p.Gln861Ter)

[2359] NM_000251.2(MSH2):c.2634G>A(p.Glu878=)

[2360] NM_000251.2(MSH2):c.2635C>T(p.Gln879Ter)

[2361] NM_000251.2(MSH2):c.28C>T(p.Gln10Ter)

[2362] NM_000251.2(MSH2):c.472C>T(p.Gln158Ter)

[2363] NM_000251.2(MSH2):c.478C>T(p.Gln160Ter)

[2364] NM_000251.2(MSH2):c.484G>A(p.Gly162Arg)

[2365] NM_000251.2(MSH2):c.490G>A(p.Gly164Arg)

[2366] NM_000251.2(MSH2):c.547C>T(p.Gln183Ter)

[2367] NM_000251.2(MSH2):c.577C>T(p.Gln193Ter)

[2368] NM_000251.2(MSH2):c.643C>T(p.Gln215Ter)

[2369] NM_000251.2(MSH2):c.645+1G>A

[2370] NM_000251.2(MSH2):c.652C>T(p.Gln218Ter)

[2371] NM_000251.2(MSH2):c.754C>T(p.Gln252Ter)

[2372] NM_000251.2(MSH2):c.792+1G>A

[2373] NM_000251.2(MSH2):c.942G>A(p.Gln314＝)

[2374] NM_000535.6(PMS2):c.949C>T(p.Gln317Ter)

[2375] NM_000249.3(MLH1):c.306+1G>A

[2376] NM_000249.3(MLH1):c.62C>T(p.Ala21Val)

[2377] NM_000251.2(MSH2):c.1865C>T(p.Pro622Leu)

[2378] NM_000179.2(MSH6):c.426G>A(p.Trp142Ter)

[2379] NM_000251.2(MSH2):c.715C>T(p.Gln239Ter)

[2380] NM_000249.3(MLH1):c.350C>T(p.Thr117Met)

[2381] NM_000251.2(MSH2):c.1915C>T(p.His639Tyr)

[2382] NM_000251.2(MSH2):c.289C>T(p.Gln97Ter)

[2383] NM_000251.2(MSH2):c.2785C>T(p.Arg929Ter)

[2384] NM_000249.3(MLH1):c.131C>T(p.Ser44Phe)

[2385] NM_000249.3(MLH1):c.1219C>T(p.Gln407Ter)

[2386] NM_000249.3(MLH1):c.306+5G>A

[2387] NM_000251.2(MSH2):c.1801C>T(p.Gln601Ter)

[2388] NM_000535.6(PMS2):c.1144+1G>A

[2389] NM_000251.2(MSH2):c.1984C>T(p.Gln662Ter)

[2390] NM_000249.3(MLH1):c.381-1G>A

[2391] NM_000535.6(PMS2):c.631C>T(p.Arg211Ter)

[2392] NM_000251.2(MSH2):c.790C>T(p.Gln264Ter)

[2393] NM_000251.2(MSH2):c.366+1G>A

[2394] NM_000249.3(MLH1):c.298C>T(p.Arg100Ter)

[2395] NM_000179.2(MSH6):c.3013C>T(p.Arg1005Ter)

[2396] NM_000179.2(MSH6):c.694C>T(p.Gln232Ter)

[2397] NM_000179.2(MSH6):c.742C>T(p.Arg248Ter)

[2398] NM_000249.3(MLH1):c.1039-1G>A

[2399] NM_000249.3(MLH1):c.142C>T(p.Gln48Ter)

[2400] NM_000249.3(MLH1):c.1790G>A(p.Trp597Ter)

[2401] NM_000249.3(MLH1):c.1961C>T(p.Pro654Leu)

[2402] NM_000249.3(MLH1):c.2103+1G>A

[2403] NM_000249.3(MLH1):c.2135G>A(p.Trp712Ter)

[2404] NM_000249.3(MLH1):c.588+5G>A

[2405] NM_000249.3(MLH1):c.790+1G>A

[2406] NM_000251.2(MSH2):c.1035G>A(p.Trp345Ter)

[2407] NM_000251.2(MSH2):c.1255C>T(p.Gln419Ter)

[2408] NM_000251.2(MSH2):c.1861C>T(p.Arg621Ter)

[2409] NM_000251.2(MSH2):c.226C>T(p.Gln76Ter)

[2410] NM_000251.2(MSH2):c.2653C>T(p.Gln885Ter)

[2411] NM_000251.2(MSH2):c.508C>T(p.Gln170Ter)

[2412] NM_000251.2(MSH2):c.862C>T(p.Gln288Ter)

[2413] NM_000251.2(MSH2):c.892C>T(p.Gln298Ter)

[2414] NM_000251.2(MSH2):c.970C>T(p.Gln324Ter)

[2415] NM_000179.2(MSH6):c.4001G>A(p.Arg1334Gln)

[2416] NM_000251.2(MSH2):c.1662-1G>A

[2417] NM_000535.6(PMS2):c.1882C>T(p.Arg628Ter)

[2418] NM_000535.6(PMS2):c.2174+1G>A

[2419] NM_000535.6(PMS2):c.2404C>T(p.Arg802Ter)

[2420] NM_000179.2(MSH6):c.3991C>T(p.Arg1331Ter)

[2421] NM_000179.2(MSH6):c.2503C>T(p.Gln835Ter)

[2422] NM_000179.2(MSH6):c.718C>T(p.Arg240Ter)

[2423] NM_000249.3(MLH1):c.1038G>A(p.Gln346=)

[2424] NM_000249.3(MLH1):c.245C>T(p.Thr82Ile)

[2425] NM_000249.3(MLH1):c.83C>T(p.Pro28Leu)

[2426] NM_000249.3(MLH1):c.884G>A(p.Ser295Asn)

[2427] NM_000249.3(MLH1):c.982C>T(p.Gln328Ter)

[2428] NM_000251.2(MSH2):c.1046C>T(p.Pro349Leu)

[2429] NM_000251.2(MSH2):c.1120C>T(p.Gln374Ter)

[2430] NM_000251.2(MSH2):c.1285C>T(p.Gln429Ter)

[2431] NM_000251.2(MSH2):c.1477C>T(p.Gln493Ter)

[2432] NM_000251.2(MSH2):c.2152C>T(p.Gln718Ter)

[2433] NM_000535.6(PMS2):c.703C>T(p.Gln235Ter)

[2434] NM_000249.3(MLH1):c.2141G>A(p.Trp714Ter)

[2435] NM_000251.2(MSH2):c.1009C>T(p.Gln337Ter)

[2436] NM_000251.2(MSH2):c.1216C>T(p.Arg406Ter)

[2437] NM_000179.2(MSH6):c.3202C>T(p.Arg1068Ter)

[2438] NM_000251.2(MSH2):c.1165C>T(p.Arg389Ter)

[2439] NM_000249.3(MLH1):c.1943C>T(p.Pro648Leu)

[2440] NM_000249.3(MLH1):c.200G>A(p.Gly67Glu)

[2441] NM_000249.3(MLH1):c.793C>T(p.Arg265Cys)

[2442] NM_000249.3(MLH1):c.2059C>T(p.Arg687Trp)

[2443] NM_000249.3(MLH1):c.677G>A(p.Arg226Gln)

[2444] NM_000249.3(MLH1):c.2041G>A(p.Ala681Thr)

[2445] NM_000249.3(MLH1):c.1942C>T(p.Pro648Ser)

[2446] NM_000249.3(MLH1):c.676C>T(p.Arg226Ter)

[2447] NM_000251.2(MSH2):c.2038C>T(p.Arg680Ter)

[2448] NM_000179.2(MSH6):c.1483C>T(p.Arg495Ter)

[2449] NM_000179.2(MSH6):c.2194C>T(p.Arg732Ter)

[2450] NM_000179.2(MSH6):c.3103C>T(p.Arg1035Ter)

[2451] NM_000179.2(MSH6):c.892C>T(p.Arg298Ter)

[2452] NM_000249.3(MLH1):c.1459C>T(p.Arg487Ter)

[2453] NM_000249.3(MLH1):c.1731G>A(p.Ser577＝)

[2454] NM_000249.3(MLH1):c.184C>T(p.Gln62Ter)

[2455] NM_000249.3(MLH1):c.1975C>T(p.Arg659Ter)

[2456] NM_000249.3(MLH1):c.199G>A(p.Gly67Arg)

[2457] NM_000251.2(MSH2):c.1076+1G>A

[2458] NM_000251.2(MSH2):c.1147C>T(p.Arg383Ter)

[2459] NM_000251.2(MSH2):c.181C>T(p.Gln61Ter)

[2460] NM_000251.2(MSH2):c.212-1G>A

[2461] NM_000251.2(MSH2):c.2131C>T(p.Arg711Ter)

[2462] NM_000535.6(PMS2):c.697C>T(p.Gln233Ter)

[2463] NM_000535.6(PMS2):c.1261C>T(p.Arg421Ter)

[2464] NM_000251.2(MSH2):c.2047G>A(p.Gly683Arg)

[2465] NM_000535.6(PMS2):c.400C>T(p.Arg134Ter)

[2466] NM_000535.6(PMS2):c.1927C>T(p.Gln643Ter)

[2467] NM_000179.2(MSH6):c.1444C>T(p.Arg482Ter)

[2468] NM_000179.2(MSH6):c.2731C>T(p.Arg911Ter)

[2469] NM_000535.6(PMS2):c.943C>T(p.Arg315Ter)

[2470] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自BCKDHA、BCKDHB、DBT和DLD的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防林奇综合征的方法。

[2471] 其他遗传疾病

[2472] 与其他遗传疾病相关联的病原性G至A或C至T突变/SNP已在ClinVar数据库中报告并在表A中公开，这些疾病包括但不限于马凡综合征、赫尔勒综合征(Hurler syndrome)、糖

原贮积病和囊性纤维化。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性G至A或C至T突变/SNP的方法。

[2473] 马凡综合征

[2474] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与马凡综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP至少存在于FBN1基因中，至少包括以下各项：

[2475] NM_000138.4(FBN1):c.1879C>T(p.Arg627Cys)

[2476] NM_000138.4(FBN1):c.1051C>T(p.Gln351Ter)

[2477] NM_000138.4(FBN1):c.184C>T(p.Arg62Cys)

[2478] NM_000138.4(FBN1):c.2855-1G>A

[2479] NM_000138.4(FBN1):c.3164G>A(p.Cys1055Tyr)

[2480] NM_000138.4(FBN1):c.368G>A(p.Cys123Tyr)

[2481] NM_000138.4(FBN1):c.4955G>A(p.Cys1652Tyr)

[2482] NM_000138.4(FBN1):c.7180C>T(p.Arg2394Ter)

[2483] NM_000138.4(FBN1):c.8267G>A(p.Trp2756Ter)

[2484] NM_000138.4(FBN1):c.1496G>A(p.Cys499Tyr)

[2485] NM_000138.4(FBN1):c.6886C>T(p.Gln2296Ter)

[2486] NM_000138.4(FBN1):c.3373C>T(p.Arg1125Ter)

[2487] NM_000138.4(FBN1):c.640G>A(p.Gly214Ser)

[2488] NM_000138.4(FBN1):c.5038C>T(p.Gln1680Ter)

[2489] NM_000138.4(FBN1):c.434G>A(p.Cys145Tyr)

[2490] NM_000138.4(FBN1):c.2563C>T(p.Gln855Ter)

[2491] NM_000138.4(FBN1):c.7466G>A(p.Cys2489Tyr)

[2492] NM_000138.4(FBN1):c.2089C>T(p.Gln697Ter)

[2493] NM_000138.4(FBN1):c.592C>T(p.Gln198Ter)

[2494] NM_000138.4(FBN1):c.6695G>A(p.Cys2232Tyr)

[2495] NM_000138.4(FBN1):c.6164-1G>A

[2496] NM_000138.4(FBN1):c.5627G>A(p.Cys1876Tyr)

[2497] NM_000138.4(FBN1):c.4061G>A(p.Trp1354Ter)

[2498] NM_000138.4(FBN1):c.1982G>A(p.Cys661Tyr)

[2499] NM_000138.4(FBN1):c.6784C>T(p.Gln2262Ter)

[2500] NM_000138.4(FBN1):c.409C>T(p.Gln137Ter)

[2501] NM_000138.4(FBN1):c.364C>T(p.Arg122Cys)

[2502] NM_000138.4(FBN1):c.3217G>A(p.Glu1073Lys)

[2503] NM_000138.4(FBN1):c.4460-8G>A

[2504] NM_000138.4(FBN1):c.4786C>T(p.Arg1596Ter)

[2505] NM_000138.4(FBN1):c.7806G>A(p.Trp2602Ter)

[2506] NM_000138.4(FBN1):c.247+1G>A

[2507] NM_000138.4(FBN1):c.2495G>A(p.Cys832Tyr)

[2508] NM_000138.4(FBN1):c.493C>T(p.Arg165Ter)

[2509] NM_000138.4(FBN1):c.5504G>A(p.Cys1835Tyr)

[2510] NM_000138.4(FBN1):c.5863C>T(p.Gln1955Ter)

[2511] NM_000138.4(FBN1):c.6658C>T(p.Arg2220Ter)

[2512] NM_000138.4(FBN1):c.7606G>A(p.Gly2536Arg)

[2513] NM_000138.4(FBN1):c.7955G>A(p.Cys2652Tyr)

[2514] NM_000138.4(FBN1):c.3037G>A(p.Gly1013Arg)

[2515] NM_000138.4(FBN1):c.8080C>T(p.Arg2694Ter)

[2516] NM_000138.4(FBN1):c.1633C>T(p.Arg545Cys)

[2517] NM_000138.4(FBN1):c.7205-1G>A

[2518] NM_000138.4(FBN1):c.4621C>T(p.Arg1541Ter)

[2519] NM_000138.4(FBN1):c.1090C>T(p.Arg364Ter)

[2520] NM_000138.4(FBN1):c.1585C>T(p.Arg529Ter)

[2521] NM_000138.4(FBN1):c.4781G>A(p.Gly1594Asp)

[2522] NM_000138.4(FBN1):c.643C>T(p.Arg215Ter)

[2523] NM_000138.4(FBN1):c.3668G>A(p.Cys1223Tyr)

[2524] NM_000138.4(FBN1):c.8326C>T(p.Arg2776Ter)

[2525] NM_000138.4(FBN1):c.6354C>T(p.Ile2118=)

[2526] NM_000138.4(FBN1):c.1468+5G>A

[2527] NM_000138.4(FBN1):c.1546C>T(p.Arg516Ter)

[2528] NM_000138.4(FBN1):c.4615C>T(p.Arg1539Ter)

[2529] NM_000138.4(FBN1):c.5368C>T(p.Arg1790Ter)

[2530] NM_000138.4(FBN1):c.1285C>T(p.Arg429Ter)

[2531] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于FBN1基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防马凡综合征的方法。

[2532] 赫尔勒综合征

[2533] 在一些实施方案中,本文所述的方法、系统和组合物用于校正与赫尔勒综合征相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中,病原性突变/SNP至少存在于IDUA基因中,至少包括以下各项:

[2534] NM_000203.4(IDUA):c.972+1G>A

[2535] NM_000203.4(IDUA):c.1855C>T(p.Arg619Ter)

[2536] NM_000203.4(IDUA):c.152G>A(p.Gly51Asp)

[2537] NM_000203.4(IDUA):c.1205G>A(p.Trp402Ter)

[2538] NM_000203.4(IDUA):c.208C>T(p.Gln70Ter)

[2539] NM_000203.4(IDUA):c.1045G>A(p.Asp349Asn)

[2540] NM_000203.4(IDUA):c.1650+5G>A

[2541] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A

或C至T突变/SNP、特别地存在于IDUA基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防赫尔勒综合征的方法。

[2542] 糖原贮积病

[2543] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与糖原贮积病相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于选自GAA、AGL、PHKB、PRKAG2、G6PC、PGAM2、GBE1、PYGM和PFKM的至少一种基因中，至少包括以下各项：

[2544] NM_000152.4(GAA):c.1927G>A(p.Gly643Arg)

[2545] NM_000152.4(GAA):c.2173C>T(p.Arg725Trp)

[2546] NM_000642.2(AGL):c.3980G>A(p.Trp1327Ter)

[2547] NM_000642.2(AGL):c.16C>T(p.Gln6Ter)

[2548] NM_000642.2(AGL):c.2039G>A(p.Trp680Ter)

[2549] NM_000293.2(PHKB):c.1546C>T(p.Gln516Ter)

[2550] NM_016203.3(PRKAG2):c.1592G>A(p.Arg531Gln)

[2551] NM_000151.3(G6PC):c.248G>A(p.Arg83His)

[2552] NM_000151.3(G6PC):c.724C>T(p.Gln242Ter)

[2553] NM_000151.3(G6PC):c.883C>T(p.Arg295Cys)

[2554] NM_000151.3(G6PC):c.247C>T(p.Arg83Cys)

[2555] NM_000151.3(G6PC):c.1039C>T(p.Gln347Ter)

[2556] NM_000152.4(GAA):c.1561G>A(p.Glu521Lys)

[2557] NM_000642.2(AGL):c.2590C>T(p.Arg864Ter)

[2558] NM_000642.2(AGL):c.3682C>T(p.Arg1228Ter)

[2559] NM_000642.2(AGL):c.118C>T(p.Gln40Ter)

[2560] NM_000642.2(AGL):c.256C>T(p.Gln86Ter)

[2561] NM_000642.2(AGL):c.2681+1G>A

[2562] NM_000642.2(AGL):c.2158-1G>A

[2563] NM_000290.3(PGAM2):c.233G>A(p.Trp78Ter)

[2564] NM_000152.4(GAA):c.1548G>A(p.Trp516Ter)

[2565] NM_000152.4(GAA):c.2014C>T(p.Arg672Trp)

[2566] NM_000152.4(GAA):c.546G>A(p.Thr182＝)

[2567] NM_000152.4(GAA):c.1802C>T(p.Ser601Leu)

[2568] NM_000152.4(GAA):c.1754+1G>A

[2569] NM_000152.4(GAA):c.1082C>T(p.Pro361Leu)

[2570] NM_000152.4(GAA):c.2560C>T(p.Arg854Ter)

[2571] NM_000152.4(GAA):c.655G>A(p.Gly219Arg)

[2572] NM_000152.4(GAA):c.1933G>A(p.Asp645Asn)

[2573] NM_000152.4(GAA):c.1979G>A(p.Arg660His)

[2574] NM_000152.4(GAA):c.1465G>A(p.Asp489Asn)

[2575] NM_000152.4(GAA):c.2512C>T(p.Gln838Ter)

[2576] NM_000158.3(GBE1):c.1543C>T(p.Arg515Cys)

[2577] NM_005609.3(PYGM):c.1726C>T(p.Arg576Ter)

[2578] NM_005609.3(PYGM):c.1827G>A(p.Lys609=)

[2579] NM_005609.3(PYGM):c.148C>T(p.Arg50Ter)

[2580] NM_005609.3(PYGM):c.613G>A(p.Gly205Ser)

[2581] NM_005609.3(PYGM):c.1366G>A(p.Val456Met)

[2582] NM_005609.3(PYGM):c.1768+1G>A

[2583] NM_001166686.1(PFKM):c.450+1G>A

[2584] 参见表A。因此，本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于选自GAA、AGL、PHKB、PRKAG2、G6PC、PGAM2、GBE1、PYGM和PFKM的至少一种基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防糖原贮积病的方法。

[2585] 囊性纤维化

[2586] 在一些实施方案中，本文所述的方法、系统和组合物用于校正与囊性纤维化相关联的一种或多种病原性G至A或C至T突变/SNP。在一些实施方案中，病原性突变/SNP存在于CFTR基因中，至少包括以下各项：

[2587] NM_000492.3(CFTR):c.3712C>T(p.Gln1238Ter)

[2588] NM_000492.3(CFTR):c.3484C>T(p.Arg1162Ter)

[2589] NM_000492.3(CFTR):c.1766+1G>A

[2590] NM_000492.3(CFTR):c.1477C>T(p.Gln493Ter)

[2591] NM_000492.3(CFTR):c.2538G>A(p.Trp846Ter)

[2592] NM_000492.3(CFTR):c.2551C>T(p.Arg851 Ter)

[2593] NM_000492.3(CFTR):c.3472C>T(p.Arg1158Ter)

[2594] NM_000492.3(CFTR):c.1475C>T(p.Ser492Phe)

[2595] NM_000492.3(CFTR):c.1679G>A(p.Arg560Lys)

[2596] NM_000492.3(CFTR):c.3197G>A(p.Arg1066His)

[2597] NM_000492.3(CFTR):c.3873+1G>A

[2598] NM_000492.3(CFTR):c.3196C>T(p.Arg1066Cys)

[2599] NM_000492.3(CFTR):c.2490+1G>A

[2600] NM_000492.3(CFTR):c.3718-1G>A

[2601] NM_000492.3(CFTR):c.171G>A(p.Trp57Ter)

[2602] NM_000492.3(CFTR):c.3937C>T(p.Gln1313Ter)

[2603] NM_000492.3(CFTR):c.274G>A(p.Glu92Lys)

[2604] NM_000492.3(CFTR):c.1013C>T(p.Thr338Ile)

[2605] NM_000492.3(CFTR):c.3266G>A(p.Trp1089Ter)

[2606] NM_000492.3(CFTR):c.1055G>A(p.Arg352Gln)

[2607] NM_000492.3(CFTR):c.1654C>T(p.Gln552Ter)

[2608] NM_000492.3(CFTR):c.2668C>T(p.Gln890Ter)

[2609] NM_000492.3(CFTR):c.3611G>A(p.Trp1204Ter)

[2610] NM_000492.3(CFTR):c.1585-8G>A

[2611] NM_000492.3(CFTR):c.223C>T(p.Arg75Ter)

[2612] NM_000492.3(CFTR):c.1680-1G>A

[2613] NM_000492.3(CFTR):c.349C>T(p.Arg117Cys)

[2614] NM_000492.3(CFTR):c.1203G>A(p.Trp401Ter)

[2615] NM_000492.3(CFTR):c.1240C>T(p.Gln414Ter)

[2616] NM_000492.3(CFTR):c.1202G>A(p.Trp401Ter)

[2617] NM_000492.3(CFTR):c.1209+1G>A

[2618] NM_000492.3(CFTR):c.115C>T(p.Gln39Ter)

[2619] NM_000492.3(CFTR):c.1116+1G>A

[2620] NM_000492.3(CFTR):c.1393-1G>A

[2621] NM_000492.3(CFTR):c.1573C>T(p.Gln525Ter)

[2622] NM_000492.3(CFTR):c.164+1G>A

[2623] NM_000492.3(CFTR):c.166G>A(p.Glu56Lys)

[2624] NM_000492.3(CFTR):c.170G>A(p.Trp57Ter)

[2625] NM_000492.3(CFTR):c.2053C>T(p.Gln685Ter)

[2626] NM_000492.3(CFTR):c.2125C>T(p.Arg709Ter)

[2627] NM_000492.3(CFTR):c.2290C>T(p.Arg764Ter)

[2628] NM_000492.3(CFTR):c.2353C>T(p.Arg785Ter)

[2629] NM_000492.3(CFTR):c.2374C>T(p.Arg792Ter)

[2630] NM_000492.3(CFTR):c.2537G>A(p.Trp846Ter)

[2631] NM_000492.3(CFTR):c.292C>T(p.Gln98Ter)

[2632] NM_000492.3(CFTR):c.2989-1G>A

[2633] NM_000492.3(CFTR):c.3293G>A(p.Trp1098Ter)

[2634] NM_000492.3(CFTR):c.4144C>T(p.Gln1382Ter)

[2635] NM_000492.3(CFTR):c.4231C>T(p.Gln1411Ter)

[2636] NM_000492.3(CFTR):c.4234C>T(p.Gln1412Ter)

[2637] NM_000492.3(CFTR):c.579+5G>A

[2638] NM_000492.3(CFTR):c.595C>T(p.His199Tyr)

[2639] NM_000492.3(CFTR):c.613C>T(p.Pro205Ser)

[2640] NM_000492.3(CFTR):c.658C>T(p.Gln220Ter)

[2641] NM_000492.3(CFTR):c.1117-1G>A

[2642] NM_000492.3(CFTR):c.3294G>A(p.Trp1098Ter)

[2643] NM_000492.3(CFTR):c.1865G>A(p.Gly622Asp)

[2644] NM_000492.3(CFTR):c.743+1G>A

[2645] NM_000492.3(CFTR):c.1679+1G>A

[2646] NM_000492.3(CFTR):c.1657C>T(p.Arg553Ter)

[2647] NM_000492.3(CFTR):c.1675G>A(p.Ala559Thr)

[2648] NM_000492.3(CFTR):c.165-1G>A

[2649] NM_000492.3(CFTR):c.200C>T(p.Pro67Leu)

[2650] NM_000492.3(CFTR):c.2834C>T(p.Ser945Leu)

[2651] NM_000492.3(CFTR):c.3846G>A(p.Trp1282Ter)

[2652] NM_000492.3(CFTR):c.1652G>A(p.Gly551Asp)

[2653] NM_000492.3(CFTR):c.4426C>T(p.Gln1476Ter)

[2654] NM_000492.3:c.3718-2477C>T

[2655] NM_000492.3(CFTR):c.2988+1G>A

[2656] NM_000492.3(CFTR):c.2657+5G>A

[2657] NM_000492.3(CFTR):c.2988G>A(p.Gln996=)

[2658] NM_000492.3(CFTR):c.274-1G>A

[2659] NM_000492.3(CFTR):c.3612G>A(p.Trp1204Ter)

[2660] NM_000492.3(CFTR):c.1646G>A(p.Ser549Asn)

[2661] NM_000492.3(CFTR):c.3752G>A(p.Ser1251Asn)

[2662] NM_000492.3(CFTR):c.4046G>A(p.Gly1349Asp)

[2663] NM_000492.3(CFTR):c.532G>A(p.Gly178Arg)

[2664] NM_000492.3(CFTR):c.3731G>A(p.Gly1244Glu)

[2665] NM_000492.3(CFTR):c.1651G>A(p.Gly551Ser)

[2666] NM_000492.3(CFTR):c.1585-1G>A

[2667] NM_000492.3(CFTR):c.1000C>T(p.Arg334Trp)

[2668] NM_000492.3(CFTR):c.254G>A(p.Gly85Glu)

[2669] NM_000492.3(CFTR):c.1040G>A(p.Arg347His)

[2670] NM_000492.3(CFTR):c.273+1G>A

[2671] 参见表A。因此,本发明的一个方面涉及一种用于通过校正一种或多种病原性G至A或C至T突变/SNP、特别地存在于CFTR基因中的一种或多种病原性G至A或C至T突变/SNP、且更特别地以上所述的一种或多种病原性G至A或C至T突变/SNP来治疗或预防囊性纤维化的方法。

[2672] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性2乳腺癌-卵巢癌相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA2基因中(HGVS:U43746.1:n.7829+1G>A)。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性2乳腺癌-卵巢癌的方法。

[2673] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与遗传因子IX缺乏症相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于F9基因中的GRCh38:ChrX:139537145处,导致Arg至Gln取代。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防遗传因子IX缺乏症的方法。

[2674] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与β+地中海贫血、β地中海贫血和重型β地中海贫血相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于HBB基因中的GRCh38:Chr11:5226820处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防β+地中海贫血、β地中海贫血和

重型β地中海贫血的方法。

[2675] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与马凡综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于FBN1基因中(IVS2DS,G-A,+1)，如Yamamoto等人J Hum Genet.2000;45(2):115-8所报道的。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防马凡综合征的方法。

[2676] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与威斯科特-奥尔德里奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于WAS基因的内含子6的位置-1处(IVS6AS,G-A,-1)，如Kwan等人(1995)所报道的。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防威斯科特-奥尔德里奇综合征的方法。

[2677] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117590440处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化的方法。

[2678] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化和遗传性胰腺炎相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117606754处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化和遗传性胰腺炎的方法。

[2679] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117587738处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化的方法。

[2680] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与透克综合征(Turcot syndrome)和林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MSH2基因中的GRCh38:Chr2:47470964处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防透克综合征和林奇综合征的方法。

[2681] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117642437处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化的方法。

[2682] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征II和林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37001058处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征II和林奇综合征的方法。

[2683] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117642594处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化的方法。

[2684] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于CFTR基因中的GRCh38:Chr7:117592658处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防囊性纤维化的方法。

[2685] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43057051处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2686] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与二氢嘧啶脱氢酶缺乏症、希尔施普龙病1(Hirschsprung disease 1)、氟尿嘧啶反应、嘧啶类似物反应-毒性/ADR、卡培他滨反应-毒性/ADR、氟尿嘧啶反应-毒性/ADR、喃氟啶(tegafur)反应-毒性/ADR相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于DPYD基因中的GRCh38:Chr1:97450058处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防二氢嘧啶脱氢酶缺乏症、希尔施普龙病1、氟尿嘧啶反应、嘧啶类似物反应-毒性/ADR、卡培他滨反应-毒性/ADR、氟尿嘧啶反应-毒性/ADR、喃氟啶反应-毒性/ADR的方法。

[2687] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MSH2基因中的GRCh38:Chr2:47478520处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2688] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37011819处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2689] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37014545处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2690] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37011867处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2691] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征相关联的病原性A至G(A>G)突变或SNP，其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37025636处。因此，本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2692] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与林奇综合征

相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:37004475处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征的方法。

[2693] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与林奇综合征和遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MSH2基因中的GRCh38:Chr2:47416430处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征和遗传性癌症易感综合征的方法。

[2694] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与林奇综合征和遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MSH2基因中的GRCh38:Chr2:47408400处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征和遗传性癌症易感综合征的方法。

[2695] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与林奇综合征和遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MLH1基因中的GRCh38:Chr3:36996710处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防林奇综合征和遗传性癌症易感综合征的方法。

[2696] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43067696处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌的方法。

[2697] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性2乳腺癌-卵巢癌和遗传性乳腺癌及卵巢癌综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA2基因中的GRCh38:Chr13:32356610处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性2乳腺癌-卵巢癌和遗传性乳腺癌及卵巢癌综合征的方法。

[2698] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与原发性扩张型心肌病和原发性家族性肥厚型心肌病相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MYH7基因中的GRCh38:Chr14:23419993处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防原发性扩张型心肌病和原发性家族性肥厚型心肌病的方法。

[2699] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与原发性家族性肥厚型心肌病、躯干前曲症(camptocormism)和肥厚型心肌病相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MYH7基因中的GRCh38:Chr14:23415225处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防原发性家族性肥厚型心肌病、躯干前曲症和肥厚型心肌病的方法。

[2700] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性乳腺癌、家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相

关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA2基因中的GRCh38:Chr13:32357741处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性乳腺癌、家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2701] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与原发性扩张型心肌病、肥厚型心肌病、心肌病和左心室肌致密化不全相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于MYH7基因中的GRCh38:Chr14:23431584处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防原发性扩张型心肌病、肥厚型心肌病、心肌病和左心室肌致密化不全的方法。

[2702] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43067607处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2703] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征、遗传性癌症易感综合征和乳腺癌相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43047666处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征、遗传性癌症易感综合征和乳腺癌的方法。

[2704] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA2基因中的GRCh38:Chr13:32370558处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2705] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征、遗传性癌症易感综合征和乳腺癌相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43074330处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征、遗传性癌症易感综合征和乳腺癌的方法。

[2706] 在一些实施方案中,本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相关联的病原性A至G(A>G)突变或SNP,其中所述病原性A>G突变或SNP位于BRCA1基因中的GRCh38:Chr17:43082403处。因此,本发明的另一方面涉及一种用于通过校正前述病原性A>G突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2707] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与囊性纤维化和遗传性胰腺炎相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于CFTR基因中的GRCh38:Chr7:117639961处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防囊性纤维化和遗传性胰腺炎的方法。

[2708] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性2乳腺癌-卵巢癌相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于BRCA2基因中的GRCh38:Chr13:32336492处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性2乳腺癌-卵巢癌的方法。

[2709] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于BRCA1基因中的GRCh38:Chr17:43063365处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌的方法。

[2710] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性1乳腺癌-卵巢癌相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于BRCA1基因中的GRCh38:Chr17:43093613处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性1乳腺癌-卵巢癌的方法。

[2711] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性乳腺癌和家族性1乳腺癌-卵巢癌相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于BRCA1基因中的GRCh38:Chr17:43093931处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性乳腺癌和家族性1乳腺癌-卵巢癌的方法。

[2712] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性肥厚型心肌病1、原发性家族性肥厚型心肌病和肥厚型心肌病相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于MYH7基因中的GRCh38:Chr14:23429279处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性肥厚型心肌病1、原发性家族性肥厚型心肌病和肥厚型心肌病的方法。

[2713] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于BRCA2基因中的GRCh38:Chr13:32356472处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性2乳腺癌-卵巢癌、遗传性乳腺癌及卵巢癌综合征以及遗传性癌症易感综合征的方法。

[2714] 在一些实施方案中，本文所述的方法、系统和组合物用于校正据信与家族性肥厚型心肌病1、原发性家族性肥厚型心肌病、家族性限制型心肌病和肥厚型心肌病相关联的病原性C至T(C>T)突变或SNP，其中所述病原性C>T突变或SNP位于MYH7基因中的GRCh38:Chr14:23429005处。因此，本发明的另一方面涉及一种用于通过校正前述病原性C>T突变或SNP来治疗或预防家族性肥厚型心肌病1、原发性家族性肥厚型心肌病、家族性限制型心肌病和肥厚型心肌病的方法。

[2715] 另外的病原性A>G突变和SNP可见于ClinVar数据库中。因此，本公开的另一方面涉

及使用本文所述的方法、系统和组合物对ClinVar中所列的病原性A>G突变或SNP进行校正以治疗或预防与其相关联的疾病或疾患。

[2716] 另外的病原性C>T突变和SNP可见于ClinVar数据库中。因此，本公开的另一方面涉及使用本文所述的方法、系统和组合物对ClinVar中所列的病原性C>T突变或SNP进行校正以治疗或预防与其相关联的疾病或疾患。

[2717] 在一方面，本文所述的发明提供了用于修饰靶基因座处的胞苷残基的方法，目的是补救并且/或者预防由或可能由T(U)至C或A至G点突变或病原性单核苷酸多态性(SNP)引起的疾病状况。

[2718] 与各种疾病相关的病原性T(U)至C或A至G突变/SNP已报告于ClinVar数据库中，包括但不限于遗传疾病、癌症、代谢性疾病或溶酶体贮积病。因此，本发明的一个方面涉及一种用于校正与如以下所论述的这些疾病中的任一种相关联的一种或多种病原性T(U)至C或A至G突变/SNP的方法。

[2719] 在一些实施方案中，本文所述的方法、系统和组合物用于校正ClinVar数据库中报告的一种或多种病原性T(U)至C或A至G突变/SNP。在一些实施方案中，本文所述的方法、系统和组合物用于校正与标题为“核酸酶编辑器及其用途(NucleobaseEditor and Uses Thereof)”的WO2017/070632中公开的任何疾病或病症相关的一种或多种病原性T(U)至C或A至G的突变/SNP，所述文献以引用方式整体并入本文。可以治疗的示例性疾病或病症包括但不限于3-甲基戊二酸尿症2型、46,XY性腺发育不全、4-α-羟苯丙酮酸羟化酶缺乏症、6-丙酮酰-四氢蝶呤合酶缺乏症、色盲症、酸不稳定亚单位缺乏症、肢端发育不全、肢端红角化病、ACTH抵抗、ACTH非依赖性大结节肾上腺增生、激活PBK-δ综合征、急性间歇性卟啉症、急性骨髓性白血病、亚当斯-奥利弗综合征(Adams-Oliver syndrome)1/5/6、腺苷酸琥珀酸裂解酶缺乏症、肾上腺脑白质营养不良、成人神经元蜡样脂褐质沉积症、伴有眼球运动失调的成人发病型共济失调、睡眠相位前移综合征、年龄相关性黄斑变性、阿拉吉勒综合征(Alagille syndrome)、亚历山大病(Alexander's disease)、艾伦-赫恩登-杜德利综合征(Allan-Herndon-Dudley syndrome)、X连锁隐性遗传性阿尔波特综合征、儿童交替性偏瘫症、伴有肺静脉错位的肺泡毛细血管发育不良、釉质发育不全、淀粉样变性运甲状腺素蛋白淀粉样变性、肌萎缩性侧索硬化症、贫血(非球形红细胞溶血性，由G6PD缺乏症引起)、贫血(常染色体隐性遗传性吡哆醇难治性铁粒幼红细胞性)、无甲畸形、抗凝血酶III缺乏症、主动脉瘤、再生障碍性贫血、载脂蛋白C2缺乏症、表观盐皮质激素过量、芳香酶缺乏症、致心律失常性右心室心肌病、家族性肥厚型心肌病、肥厚型心肌病、先天性多发性关节挛缩症、天冬氨葡糖氨尿症、窒息性胸营养不良、伴有维生素E缺乏症的共济失调、共济失调(痉挛性)、心房颤动、房间隔缺损、非典型性溶血性尿毒综合征、常染色体显性遗传性CD11C+/CD1C+树突细胞缺乏症、伴有线粒体DNA缺失的常染色体显性遗传进行性外眼肌麻痹、Baraitser-Winter综合征、巴特综合征(Bartter syndrome)、基底神经节钙化、贝克威斯-韦德曼氏综合征(Beckwith-Wiedemann syndrome)、良性家族性新生儿癫痫、良性肩腓肌营养不良、巨大血小板综合征、β地中海贫血、β-D-甘露糖苷贮积症、Bietti晶状角膜视网膜营养不良、胆汁酸吸收不良、生物素酶缺乏症、Borjeson-Forssman-Lehmann综合征、Boucher Neuhauser综合征、Bowen-Conradi综合征、指过短、Brown-Vialetto-Van laere综合征、Brugada综合征、心律失常、心面皮肤综合征、心肌病、Carnevale综合征、肉碱棕榈酰转移酶II缺乏症、卡

朋特综合征(Carpenter syndrome)、白内障、儿茶酚胺多态性室性心动过速、中央轴空病、染色体1、9和16的着丝粒不稳定性和免疫缺陷、常染色体显性遗传性脑动脉病、脑眼面骨骼综合征、蜡样脂褐质沉积症、夏科-马里-图思病、胆固烷醇贮积病、软骨钙质沉着病、软骨发育不良、慢性进行性多发性硬化症、辅酶Q10缺乏症、科恩综合征(Cohen syndrome)、因子V和因子VIII联合缺乏症、联合免疫缺陷、联合氧化磷酸化缺乏症、1 7-α-羟化酶/17,20-裂解酶部分性联合缺乏症、补体因子d缺乏症、17-α-羟化酶/17,20-裂解酶完全联合缺乏症、锥杆营养不良、先天性挛缩性细长指、先天性去糖基化障碍、先天性脂肪瘤过度生长、卵巢赘生物、PIK3CA相关的过度生长疾病谱、先天性长QT综合征、先天性肌营养不良、先天性肌肥大-脑综合征、先天性肌无力综合征、伴有纤维类型不均衡的先天性肌病、Eichsfeld型先天性肌营养不良、先天性静止性夜盲症、角膜营养不良、狄兰吉氏综合征(Cornelia de Lange syndrome)、颅骨干骺端发育不良、克里格勒纳贾尔综合征(Crigler Najjar syndrome)、克鲁宗综合征、伴有骨营养不良的皮肤松弛症、发绀、囊性纤维化、胱氨酸病、细胞色素c氧化酶缺乏症、线粒体复合物I缺乏症、D-2-羟基戊二酸尿症、Danon病、伴有迷路发育不全小耳畸形和小牙畸形(LAMM)的耳聋、耳聋、乙酰辅酶A乙酰转移酶缺乏症、铁氧化酶缺乏症、UDP葡萄糖-己糖-1-磷酸尿苷酰转移酶缺乏症、索二氏病(Dejerine-Sottas disease)、德斯比夸综合征(Desbuquois syndrome)、DFNA、糖尿病2型、糖尿病耳聋综合征、Diamond-Blackfan贫血、弯曲变形性发育不良、二氢蝶啶还原酶缺乏症、二氢嘧啶酶缺乏症、扩张型心肌病、播散性非典型性分枝杆菌感染、远端关节挛缩症、远端遗传性运动神经元病、Donnai Barrow综合征、杜兴氏肌营养不良、贝克肌营养不良、遗传性泛发性色素异常症、先天性角化不良、肌张力障碍、早期婴儿癫痫性脑病、Ehlers-Danlos综合征、Eichsfeld型先天性肌营养不良、埃-德二氏肌营养不良、釉质-肾综合征、逆向性营养不良性大疱性表皮松解症、大疱性表皮松解症、癫痫、发作性共济失调、可变性红斑皮肤角化症、红细胞生成性原卟啉症、锻炼不耐受、渗出性玻璃体视网膜病变、法布里病、因子V缺乏症、因子VII缺乏症、因子xiii缺乏症、家族性腺瘤性息肉病、乳腺癌、卵巢癌、寒冷性荨麻疹!、慢性婴儿神经,皮肤和关节综合征、偏瘫型偏头痛、高胆固醇血症、肥厚型心肌病、初级α低脂蛋白血症、低钾血症-低镁血症、青少年痛风、高脂蛋白血症、内脏淀粉样变性、低磷酸盐血症性维生素D难治性佝偻病、FG综合征、眼外肌纤维化、芬兰先天性肾病综合征、局灶性癫痫、局灶节段性肾小球硬化症、额颞骨发育不良、额颞叶痴呆、果糖-二磷酸酶缺乏症、Gamstorp-Wohlfart综合征、神经节糖苷唾液酸酶缺乏症、GATA-I相关的血小板减少症、高雪氏病(Gaucher disease)、巨轴索神经病变、血小板无力症、肾小球囊性肾病、肾小球病、糖皮质激素抵抗、葡萄糖-6-磷酸转运缺陷、戊二酸尿症、糖原贮积病、Gorlin综合征、前脑无裂畸形、GRACILE综合征、出血性毛细血管扩张症、血色素沉着症、血红蛋白H病、溶血性贫血、噬血细胞性淋巴组织细胞增生症、结肠癌、Myhre综合征、脑白质病、遗传性因子IX缺乏症、遗传性因子VIII缺乏症、遗传因子XI缺乏症、遗传性果糖尿病、遗传性非息肉性结直肠赘生物、遗传性胰腺炎、遗传性嗜铬细胞增多症、椭圆细胞增多症、内脏异位、异位、组织细胞性髓性网状细胞增生症、组织细胞增多症-淋巴结病附加症综合征、由醛缩酶A缺乏症引起的HNSHA、全羧化酶合成酶缺乏症、高半胱氨酸血症、Rowel-Evans综合征、葡萄胎、高尿钙高钙血症、超免疫球蛋白D、甲羟戊酸尿症、高胰岛素血症低血糖、高钾血症周期性麻痹、von Eulenburg先天性副肌强直、高脂蛋白血症、高锰血症、高甲硫胺酸血症、高磷酸酶血症、高

血压、低镁血症、低β脂蛋白血症、低钙血症、低促性腺激素性性腺功能减退症、低促性腺激素性性腺功能减退症、少汗性外胚层发育不良、高IgM免疫缺陷、少汗性X连锁性外胚层发育不良、低镁血症、甲状旁腺功能减退、特发性纤维化肺泡炎、免疫缺陷、免疫球蛋白A缺乏症、婴儿低磷酸酯酶症、婴儿帕金森症-肌张力障碍、胰岛素依赖性糖尿病、中间型枫糖尿病、髋-髌骨发育不良、胰岛细胞增生、孤立性生长激素缺乏症、孤立性促黄体素缺乏症、异戊酸血症、乔伯特综合征、青少年息肉病综合征、青少年视网膜劈裂、卡尔曼综合征(Kallmann syndrome)、卡塔格内综合征、Kugelberg-Welander病、晶状体角膜营养不良、莱伯氏先天性黑蒙症、莱伯氏视神经萎缩、左心室肌致密化不全、利氏病(Leigh disease)、线粒体复合物I缺乏症、矮怪病综合征、关节挛缩症、前角细胞病、白细胞粘附缺陷、脑白质营养不良、脑白质病、卵巢脑白质营养不良、L-铁蛋白缺乏症、李-佛美尼综合征、肢带型肌营养不良-抗肌萎缩相关糖蛋白病、Loeys-Dietz综合征、长QT综合征、大头畸型/自闭症综合征、黄斑角膜营养不良、黄斑营养不良、恶性高热易感性、前列腺恶性肿瘤、枫糖尿病、马登-沃克样综合征、马凡综合征、玛丽恩纳(Marie Unna)遗传性少毛症、肥大细胞病、胎粪性肠梗阻、中链酰基辅酶A脱氢酶缺乏症、梅尔尼克-弗雷泽综合征(Melnick-Fraser syndrome)、智力迟钝、Merosin缺陷型先天性肌营养不良、间皮瘤、异染性脑白质营养不良、干骺端发育不良、高铁血红蛋白血症、甲基丙二酸尿症、高半胱氨酸尿症、小头畸形、脉络膜视网膜病变、淋巴水肿、小眼畸形、轻度非PKU型高苯丙氨酸血症、米切尔-莱利综合征、线粒体3-羟基-3-甲基戊二酰辅酶A合酶缺乏症、线粒体复合物I缺乏症、线粒体复合物III缺乏症、线粒体肌病、粘多糖症III、粘多糖贮积症、多发性硫酸酯酶缺乏症、肌无力综合征、结核分枝杆菌、绿过氧物酶缺乏症、Myhre综合征、肌阵挛性癫痫、肌原纤维肌病、肌红蛋白尿、肌病、近视、先天性肌强直、Navajo脑肝病、杆状体肌病、胃赘生物、肾源性尿崩症、肾消耗病、肾病综合征、神经纤维瘤病、中性脂质贮积病、尼曼-皮克病(Niemann-Pick disease)、非酮症性高甘氨酸血症、努南综合征(Noonan syndrome)、努南综合征样病症、诺卢姆病(Norum disease)、黄斑变性、N端乙酰转移酶缺乏症、眼皮肤白化病、眼齿指发育不良、Ohdo综合征、视神经发育不全、鸟氨酸氨甲酰基转移酶缺乏症、口面指综合征、成骨不全症、骨硬化病、卵巢发育不全、厚甲症、非表皮松解性掌跖角化病、Papillon-Lef\xc3\xa8vre综合征、Haim-Munk综合征、牙周炎、皮肤剥脱综合征、Pendred综合征、过氧化物酶体脂肪酰辅酶A还原酶I病症、过氧化物酶体生物发生障碍、发否氏综合征、苯丙酮尿症、苯丙酮尿症、非PKU型高苯丙氨酸血症、垂体激素缺乏症、毛发红糠疹、结节性多动脉炎、多囊肾病、多囊脂肪膜骨发育不良、多小脑回畸形、脑桥小脑发育不全、汗孔角化病、后柱性共济失调、原发性红斑性肢痛症、高草酸尿症、进行性家族性肝内胆汁淤积、进行性假性类风湿发育不良、丙酸血症、假两性畸形、假性醛固酮减少症、弹性纤维假黄瘤样病症、嘌呤核苷磷酸化酶缺乏症、吡哆醛5-磷酸依赖性癫痫、肾发育不良、视网膜色素营养不良、小脑性共济失调、骨骼发育不良、网状细胞发育不全、色素性视网膜炎、乌谢尔综合征、视网膜母细胞瘤、视网膜病、RRM2B相关的线粒体疾病、鲁宾斯坦-泰比综合征(Rubinstein-Taybi syndrome)、Schnyder晶状角膜营养不良、皮脂腺瘤、重度先天性中性粒细胞减少症、重度婴儿肌阵挛性癫痫、重度X连锁性肌管肌病、关节骨指甲发育不良、面部畸形、少毛症、短肋胸椎发育异常、唾液酸贮积病、唾液酸贮积症、铁粒幼细胞性贫血、小纤维神经病变、史密斯-马吉尼斯综合征(Smith-Magenis syndrome)、索斯比眼底营养不良、痉挛性共济失调、痉挛性截瘫、生精障碍、球形红细胞增多症、鞘磷

脂/胆固醇脂质沉着症、脊髓小脑性共济失调、裂手/足畸形、脊椎干骺端发育不良、扁平椎体致死性骨骼发育不良、头颈部鳞状细胞癌、斯特格病(Stargardt disease)、蔗糖酶异麦芽糖酶缺乏症、婴儿猝死综合征、瓣上主动脉狭窄、表面活性物质代谢功能障碍、丹吉尔病(Tangier disease)、塔顿-布朗-拉赫曼综合征(Tatton-Brown-rahman syndrome)、胸主动脉瘤和主动脉夹层形成、易栓症、甲状腺激素抵抗、TNF受体相关性周期性发热综合征(TRAPS)、牙齿发育不全、尖端扭转(Torsades de pointes)、大动脉转位、特-柯二氏综合征、结节性硬化症综合征、酪氨酸酶阴性眼皮肤白化病、酪氨酸酶阳性眼皮肤白化病、酪氨酸血症、UDP葡萄糖-4-表异构酶缺乏症、乌尔里希先天性肌营养不良、贝特莱姆肌病乌谢尔综合征、紫外线敏感综合征、范德伍德综合征、窝翼状赘蹼综合征、超长链酰基辅酶A脱氢酶缺乏症、膀胱输尿管反流、玻璃体视网膜脉络膜病变、Von Rippel-Lindau综合征、血管性血友病(von Willebrand disease)、瓦登伯革氏综合征、Warsaw破损综合征、WFSI相关病症、威尔逊氏病、着色性干皮病、X连锁性无丙种球蛋白血症、X连锁性遗传性运动和感觉神经病变、X连锁性重度联合免疫缺陷和齐薇格综合征(Zellweger syndrome)。

[2720] 在某些实施方案中，本文所述的方法、系统和组合物用于校正如下表中提供的一种或多种病原性T(U)至C或A至G突变/SNP。

[2721]

候选物	基因	疾病
NM_007262.4(PARK7):c.497T>C (p.Leu166Pro)	PARK7	帕金森病 7
NM_174936.3(PCSK9):c.646T>C (p.Phe216Leu)	PCSK9	常染色体显性遗传性高胆固醇血症 3
NM_000642.2(AGL):c.3083+2T>C	AGL	糖原贮积病 III 型
NM_213653.3(HFE2):c.842T>C (p.Ile281Thr)	HFE2	血色素沉着症 2A 型

[2722]

NM_170707.3(LMNA):c.799T>C (p.Tyr267His)	LMNA	原发性扩张型心肌病|未提供
NM_000488.3(SERPINC1):c.1141T>C (p.Ser381Pro)	SERPINC1	抗凝血酶 III 缺乏症
NM_000465.3(BARD1):c.1159T>C (p.Phe387Leu)	BARD1	家族性乳腺癌|未指明|遗传性癌症易感综合征
NM_000030.2(AGXT):c.613T>C (p.Ser205Pro)	AGXT	原发性高草酸尿症，I 型|未提供
NM_001302946.1(TRNT1):c.668T>C (p.Ile223Thr)	TRNT1	伴有 B 细胞免疫缺陷、周期性发热和发育迟缓的铁粒幼细胞性贫血
NM_138694.3(PKHD1):c.8068T>C (p.Trp2690Arg)	PKHD1	常染色体隐性遗传性多囊肾病
NM_000162.3(GCK):c.1169T>C (p.Ile390Thr)	GCK	青年成人发病型糖尿病，2 型
NM_017890.4(VPS13B):c.7504+2T>C	VPS13B	科恩综合征 (Cohen syndrome)
NM_000155.3(GALT):c.512T>C (p.Phe171Ser)	GALT	UDP 葡萄糖-己糖-1-磷酸尿苷酰转移酶缺乏症
NM_000277.1(PAH):c.691T>C (p.Ser231Pro)	PAH	苯丙酮尿症|未提供
NM_000138.4(FBN1):c.4531T>C (p.Cys1511Arg)	FBN1	马凡综合征
NM_000527.4(LDLR):c.1745T>C (p.Leu582Pro)	LDLR	家族性高胆固醇血症

[2723] 在以下编号的段落中描述了一些示例性实施方案。

[2724] 1.一种适用于修饰靶基因座处的核酸的工程化的非天然存在的系统，所述系统包含a)靶向组分或编码所述靶向组分的核苷酸序列；和b)碱基编辑组分或编码所述碱基编辑组分的核苷酸序列。

[2725] 2.如段落1所述的系统，其中所述靶向结构域是DNA结合蛋白或其功能部分，或RNA结合蛋白或其功能部分。

[2726] 3.如段落2所述的系统，其中所述DNA结合蛋白是Cas9或Cpf1，并且所述系统还包含含有指导序列的指导分子或编码所述指导分子的核苷酸。

[2727] 4.如段落2所述的系统，其中所述RNA结合蛋白是Cas13蛋白，并且所述系统还包含含有指导序列的指导分子或编码所述指导分子的核苷酸序列。

[2728] 5.如段落3或4所述的系统，其中所述Cas9、Cpf1或Cas13是死亡Cas9、死亡Cpf1或死亡Cas13。

[2729] 6.如前述段落中任一项所述的系统，其中所述碱基编辑组分是腺苷脱氨酶或其催化结构域。

[2730] 7.如段落1所述的系统，其中所述腺苷脱氨酶或其催化结构域融合至所述靶向组分。

[2731] 8.如段落6所述的系统,其中所述腺苷脱氨酶或其催化结构域通过接头融合至所述靶向组分。

[2732] 9.如前述段落中任一项所述的系统,其中所述腺苷脱氨酶蛋白或其催化结构域是人类、头足类动物或果蝇腺苷脱氨酶蛋白或其催化结构域。

[2733] 10.如前述段落中任一项所述的系统,其中所述腺苷脱氨酶蛋白或其催化结构域已被修饰以包含在hADAR2-D氨基酸序列的谷氨酸488处、或同源ADAR蛋白中的相应位置处的突变。

[2734] 11.如段落10所述的系统,其中在位置488或同源ADAR蛋白中的相应位置处的所述谷氨酸残基被谷氨酰胺残基替代(E488Q)。

[2735] 12.如前述段落中任一项所述的系统,其中所述腺苷脱氨酶蛋白或其催化结构域是包含突变E488Q的突变hADAR2d或包含突变E1008Q的突变hADAR1d。

[2736] 13.一种载体系统,所述载体系统包含一种或多种编码如前述段落中任一项所述的系统的载体。

[2737] 14.一种包含如段落1至12中任一项所述的系统的体外或离体宿主细胞或其子代或者细胞系或其子代。

[2738] 15.一种用于对靶基因座进行可编程和靶向式的碱基编辑的方法,所述方法包括递送如段落1-12中任一项所述的系统。

[2739] 16.如段落15所述的反腐,其中所述方法包括确定所述目标靶序列以及选择最有效地使存在于所述靶序列中的所述腺嘌呤脱氨基的所述腺苷脱氨酶蛋白或其催化结构域。

[2740] 17.如段落16所述的方法,其中在所述目标靶RNA中所述腺嘌呤的所述脱氨补救了由含有病原性G→A或C→T点突变的转录物引起的疾病。

[2741] 18.如段落17所述的方法,其中所述疾病选自梅尔-戈林综合征、塞克尔综合征4、乔伯特综合征5、莱伯氏先天性黑蒙症10;夏科-马里-图思病2型;夏科-马里-图思病2型;乌谢尔综合征2C型;脊髓小脑性共济失调28;脊髓小脑性共济失调28;脊髓小脑性共济失调28;长QT综合征2;西奥格林-拉尔逊氏综合征;遗传性果糖尿病;遗传性果糖尿病;神经母细胞瘤;神经母细胞瘤;卡尔曼综合征1;卡尔曼综合征1;卡尔曼综合征1;异染性脑白质营养不良、雷特综合征、肌萎缩性侧索硬化症10型、李-佛美尼综合征,或表5中所列出的疾病。

[2742] 在以下实施例中进一步描述了本发明,这些实施例不限制权利要求中所描述的本发明的范围。

[2743] 实施例

[2744] 实施例1

[2745] 腺嘌呤脱氨酶(AD)使双链RNA中特定位点处的腺嘌呤脱氨基。

[2746] 一些AD可以在DNA-RNAn RNA双链体上实现腺嘌呤脱氨(例如Zheng等人,Nucleic Acids Research 2017),这一事实提供了利用在非活性Cas13的RNA指导的DNA结合期间所形成的R-环中在指导RNA与其互补DNA靶标之间所形成的RNA体来开发RNA指导的AD的独特机会。通过使用无活性的Cas13来募集AD,AD酶随后将作用于RNA-DNAn RNA双链体中的腺嘌呤。

[2747] 在一个实施方案中,使用以下突变来获得无活性的Cas13诸如Cas13b:R116A、H121A、R1177A和H1182A。为了提高通过AD进行编辑的效率,使用突变ADAR,诸如包含突变

E488Q的突变hADAR2d。

[2748] 用于将AD募集至特定基因座的设计：

[2749] 1.带有NLS标签的无活性Cas13在N端或C端融合至AD。使用多种接头，包括诸如GSG5(SEQ ID NO:10)的柔性接头或诸如LEPGEKPYKCPECGKSFSQSGALTRHQRTHTR的较不灵活接头。

[2750] 2.用诸如MS2结合位点的适体修饰指导RNA支架(例如Konermann等人，Nature2015)。将带有NLS标签的AD-MS2结合蛋白融合物与(带有NLS标签的无活性或Cas13b)和相应的指导RNA共同引入到靶细胞中。

[2751] 3.将AD插入带有NLS标签的无活性或切口酶Cas13的内环中。

[2752] RNA指导物的设计：

[2753] 1.设计长度与Cas13蛋白的天然指导序列相对应的指导序列以靶向目标RNA。

[2754] 2.使用比典型长度长的RNA指导物以在蛋白质指导RNA-靶DNA复合物之外形成RNA双链体。

[2755] 对于这些RNA指导物设计中的每一个，与靶RNA链上腺嘌呤相对的RNA碱基将被指定为C而不是U。

[2756] AD的选择和设计：

[2757] 使用了多种AD，每种AD具有不同的活性水平。这些AD

[2758] 1.人类ADAR(hADAR1、hADAR2、hADAR3)

[2759] 2.鱿鱼普通章鱼ADAR

[2760] 3.鱿鱼乌贼(Sepia)ADAR；乳光枪乌贼ADAR

[2761] ADAT(人类ADAT、果蝇ADAT)

[2762] 还可以使用突变以增加ADAR针对DNA-RNAn RNA双链体反应的活性。例如，对于人类ADAR基因，使用hADAR1d(E1008Q)或hADAR2d(E488Q)突变以增加其针对DNA-RNA靶标的活性。

[2763] 每种ADAR具有不同级别的序列情形要求(sequence context requirement)。例如，对于hADAR1d(E1008Q)，tAg和aAg位点被有效脱氨基，而aAt和cAc的编辑效率较低，而gAa和gAc的编辑效率甚至更低。然而，情形要求因不同的ADAR而异。

[2764] 图1中提供了该系统一种型式的示意图。图4中提供了示例性AD蛋白的氨基酸序列。

[2765] 在本文说明性地描述的实施方案可以在缺少本文未具体公开的任何一种或多种要素、一种或多种限制的情况下实践。因此，例如术语“包含”、“包括”、“含有”等应当被理解为开放性的，而非限制性的。另外，本文所采用的术语和表达被用作描述性术语而非用于限制，并且无意在使用这些术语和表达时排除所示和所描述的特征的任何等效物或其部分，但是应当认识到在所要求保护的技术的范围内可以进行各种修改。另外，短语“基本上由……组成”将被理解为包括具体叙述的那些要素以及不会实质性影响所要求保护的技术的基本和新颖特征的那些附加要素。短语“由……组成”排除未指定的任何要素。

[2766] 本公开不限于本申请中所描述的特定实施方案。如对于本领域技术人员将显而易见的，在不脱离本公开的精神和范围的情况下可以进行许多修改和变型。除本文列举的那些方法和组合物之外，根据前述描述在本公开范围内的功能上等效的方法和组合物对于本

领域技术人员将是显而易见的。这些修改和变型也旨在落入所附权利要求的范围内。本公开仅由所附权利要求的术语以及这些权利要求所享有的等效物的全部范围来限制。应当理解，本公开不限于特定的方法、试剂、化合物组合物或生物系统，它们当然可以变化。还应理解，本文使用的术语仅用于描述特定实施方案的目的，而并非旨在进行限制。

[2767] 此外，在以马库什组来描述本公开的特征或方面的情况下，本领域技术人员将认识到，由此也以马库什组的任何单独成员或成员的子组来描述本公开。

[2768] 如本领域技术人员将理解的，出于任何目的和所有目的，特别是就提供书面描述而言，本文公开的所有范围还涵盖其任何和所有可能的子范围及子范围的组合。所列出的任何范围可以容易地被认为充分描述且使得同一范围能够细分为至少相等的二分之一、三分之一、四分之一、五分之一、十分之一等。作为非限制性实例，本文所论述的每个范围可以容易地细分为下三分之一、中三分之一和上三分之一等。如本领域技术人员还将理解的，所有语言诸如“多至”、“至少”、“大于”、“小于”等包括所叙述的数字并且指代可随后细分为以上论述的子范围的范围。最后，如本领域技术人员将理解的，范围包括每个单独成员。

[2769] 本说明书中引用的所有公布、专利申请、授权专利和其他文献以引用方式并入本文，如同每个单独公布、专利申请、授权专利或其他文献被确切地且单独地指明为以引用方式整体并入。若以引用方式并入的文本中所含的定义与本公开中的定义相抵触，则将其排除。

[2770] 实施例2

[2771] 用于Cas13a/Cas13b干扰的Cluc/Gluc平铺

[2772] 为了比较Cas13a和Cas13b之间的敲低效率，分别用24个或96个指导物对海萤荧光素酶和海肾荧光素酶基因进行平铺(图10)。将指导物与Cas13a和Cas13b匹配，结果显示Cas13b的敲低效率增加，针对每个基因，所有指导物(一个指导物除外)都显示出Cas13b的效率更高。

[2773] 通过NGS进行的ADAR编辑定量

[2774] 通过设计具有提前终止密码子UAG的荧光素酶报告基因，测试Cas13b-ADAR2RNA的编辑效率，所述提前终止密码子UAG阻止荧光素酶的表达(图11A)。设计7个不同长度的指导物并相对于UAG终止密码子定位，这些指导物都含有针对UAG中的A的C错配。已知C错配会在编辑位点处产生气泡，该气泡是ADAR催化结构域所偏爱的。Cas13b-ADAR2进行的RNA编辑会将UAG转换为UIG(UGG)，这引入了色氨酸而不是终止密码子，并且允许进行翻译。在HEK293FT细胞中指导物和Cas13b12-ADAR2融合物的表达使荧光素酶表达恢复到不同水平，其中指导物5的恢复程度最大(图11B)。一般来讲，随着编辑位点离crRNA的3’端(正向重复序列所处的位置并且因此是蛋白质结合的位置)更远，从指导物1-指导物5，编辑水平不断提高。这可能指示crRNA:靶标双链体中被蛋白质结合的部分是ADAR催化结构域不可接近的。指导物5、指导物6和指导物7显示出最大的活性，因为编辑位点位于远离DR/蛋白质结合区域的指导物的远端，而且因为它们的指导物更长，产生受ADAR偏好的更长的RNA双链体。当编辑位点在RNA双链体中间时，ADAR活性最佳。针对非靶向指导物条件，将荧光素酶活性的相对表达归一化。

[2775] 对这些样品进行测序以精确定量RNA编辑效率(图11C)。编辑效率列在指导物标签旁边的括号中。总体而言，通过测序确定的编辑百分比与荧光素酶表达恢复的相对水平相

符，如图11B所示。指导物5显示出最多的RNA编辑，中靶A处的G转化率为45％。在一些情况下，该区域中存在少量的脱靶A-G编辑。这些编辑可以通过在指导序列中引入G错配来减少，所述G错配不受ADAR催化结构域偏好。

[2776] 除了编辑荧光素酶报告转录物外，还设计了指导物来编辑KRAS和PPIB转录物中的框外UAG位点，其中两个指导物靶向每个转录物(图12)。这些指导物的设计原理与上述指导物5相同(45nt的间隔区，编辑位点距3'DR的距离为27nt，具有针对编辑位点的腺苷的C错配)。KRAS指导物能够在中靶腺苷处实现6.5％和13.7％的编辑，而PPIB指导物则能够实现7.7％和9.2％的编辑。这些指导物中的一些也存在一些脱靶，可以通过在间隔区中针对附近可能存在的脱靶腺苷设计G错配来减少脱靶。看起来脱靶似乎发生在靶腺苷的双链体区3'上。

[2777] 来自RNA Seq的Cas13a/b+shRNA特异性

[2778] 为了确定Cas13b12敲低的特异性，对整个转录组中的所有mRNA进行RNA测序(图13A)。针对非靶向指导物比较了指导物靶向Gluc和KRAS的敲低，发现Cas13a2和Cas13b12对靶转录物具有特异性敲低(图13A中的红色点)，而shRNA具有许多脱靶，如分布中更大的变异所证实的。对于这些条件中的每一个，显著脱靶数量显示在图13B中。对于变化超过2倍或低于0.8倍的任何脱靶转录物，均需通过FDR校正的t检验($p<0.01$)测量显著脱靶。与针对shRNA条件发现的数百个脱靶相比，Cas13a和Cas13b条件的脱靶极少。每个条件的敲低效率显示在图13C中。

[2779] 用以减少脱靶的错配特异性(A:A或A:G)

[2780] 为了减少靶腺苷编辑位点附近的腺苷的脱靶，设计了与潜在脱靶腺苷具有G或A错配的指导物(图14和下表)。G或A错配在活性上不为ADAR催化结构域所偏好。

[2781]

名称	指导物
具有 C 错配的荧光素酶指导物 WT (SEQ ID NO: 162)	catagaatgttctaaaCCAtcctgcggcctctactctg cattcaa
具有 C 错配和 1 个 G MM 的荧光素酶指导物 WT (SEQ ID NO: 163)	catagaatgttcGaaaCCAtcctgcggcctctactct gcattcaa
具有 C 错配和 2 个 G MM 的荧光素酶指导物 WT (SEQ ID NO: 164)	catagaatgGtcGaaaCCAtcctgcggcctctactct gcattcaa
具有 C 错配和 1 个 G MM 的荧光素酶指导物 WT (SEQ ID NO: 165)	catagaatgttcAaaaCCAtcctgcggcctctactct gcattcaa
具有 C 错配和 2 个 G MM 的荧光素酶指导物 WT (SEQ ID NO: 166)	catagaatgAtcAaaaCCAtcctgcggcctctactct gcattcaa
具有 C 错配的 KRAS 指导物 WT (SEQ ID NO: 167)	ggtttctccatcaattacCacttgcttcctgtaggaatcct ctatt
具有 C 错配和 1 个 G MM 的 KRAS 指导物(SEQ ID No: 168)	ggtttctccatcaatGacCacttgcttcctgtaggaatc ctctatt
具有 C 错配和 2 个 G MM 的 KRAS 指导物(SEQ ID No: 169)	ggtttctccatcaaGGacCacttgcttcctgtaggaatc ctctatt
具有 C 错配和 1 个 A MM 的 KRAS 指导物(SEQ ID No: 170)	ggtttctccatcaatAacCacttgcttcctgtaggaatc ctctatt
具有 C 错配和 2 个 A MM 的 KRAS 指导物(SEQ ID No: 171)	ggtttctccatcaaAAacCacttgcttcctgtaggaatc ctctatt
具有 C 错配的 PPIB 指导物 WT (SEQ ID NO: 172)	gcctttctctcctgtagcCaaggccacaaaattatccac tgttttt
具有 C 错配和 1 个 G MM 的 PPIB 指导物 WT (SEQ ID No: 173)	gcctttctctcctgGagcCaaggccacaaaattatcca ctgttttt

[2782]

具有 C 错配和 1 个 A MM 的 PPIB 指导物 WT (SEQ ID No: 174)	gcctttctctcctgAagcCaaggccacaaaattatcca ctgttttt

[2783] 上表中的指导物被设计为针对要编辑的中靶腺苷具有C错配，并且针对已知脱靶位点具有G或A错配(基于上述RNA测序结果)。间隔序列中的错配是大写的。

[2784] 用于中靶活性的错配

[2785] 先前对ADAR2催化结构域的研究表明，与靶A相对的不同碱基可以影响肌苷编辑的数量(Zheng等人(2017)，Nucleic Acid Research，45(6)：3369-3377)。具体来说，发现U和C与ADAR编辑的自然A相对，而G和A则不然。为了测试A和G与编辑后的A的错配是否可以用于阻遏ADAR活性，在荧光素酶报告基因测定中，使用所有其他3种可能的碱基对已知具有C错配的活性指导物进行了测试(图16)。通过测定荧光素酶活性来定量相对活性。下表中提供了指导序列。

[2786]

错配	指导序列
错配 C (SEQ ID NO: 175)	GcatagaatgttctaaaCCAtcctgcggcctcta ctctgcattcaa
错配 G (SEQ ID NO: 176)	GcatagaatgttctaaaCGAtcctgcggcctcta ctctgcattcaa
错配 T (SEQ ID NO: 177)	GcatagaatgttctaaaCTAtcctgcggcctctac tctgcattcaa
错配 A (SEQ ID NO: 178)	GcatagaatgttctaaaCAAtcctgcggcctcta ctctgcattcaa

[2787]

[2788] 通过化学修饰gRNA改善编辑并减少脱靶修饰

[2789] 按照Vogel等人(2014)，Angew Chem Int Ed，53：6267-6271，doi：10.1002/anie.201402634)中所例示进行化学修饰的gRNA降低了脱靶活性并提高了中靶效率。2'-O-甲基和硫代磷酸酯修饰的指导RNA通常可提高细胞中的编辑效率。

[2790] 基序优选性

[2791] 已证明ADAR显示出对编辑过的A的任一侧的相邻核苷酸的优选性(www.nature.com/nsmb/journal/v23/n5/full/nsmb.3203.html，Matthews等人(2017)，Nature Structural Mol Biol，23(5)：426-433)。通过用围绕靶向A的可变碱基靶向海萤荧光素酶转录物来系统地测试优选性(图17)。

[2792] 用以增强RNA编辑效率的大气泡

[2793] 为了提高对非优选5'或3'相邻碱基的RNA编辑效率，在相邻碱基中引入了有意错配，已在体外证明此举可以允许编辑非优选基序(https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gku272；Schneider等人(2014)，Nucleic Acid Res，42(10)：e87)；Fukuda等人(2017)，Scienticic Reports，7，doi：10.1038/srep41478)。测试了另外的错配(诸如鸟苷取代)，以查看它们是否降低自然优选性(图18)。

[2794] 编辑转录物中的多个A

[2795] 结果表明，ADAR脱氨酶结构域的靶向窗口中的与C相对的A可先于其他碱基优先被编辑。此外，在靶向碱基的几个碱基内与U配对的A碱基显示出低Cas13b-ADAR融合物编辑水平，这表明该酶可以灵活地编辑多个A(图19)。这两个观察结果表明，可以通过使待编辑的所有A与C错配来指定Cas13b-ADAR融合物的活性窗口中的多个A以进行编辑。为了对此进行测试，采用了优化实验中最有前景的指导物并且在活性窗口中设计了多个A:C错配，以测试

创建多个A:I编辑的可能性。使用NGS定量此实验的编辑率。为了阻遏活性窗口中的潜在脱靶编辑，将非靶A与A或G配对(取决于碱基优选性实验的结果)。

[2796] 用于RNA编辑的指导物长度滴定

[2797] ADAR天然地作用于长度>20bp的分子间或分子内RNA双链体(另参见Nishikura等人(2010)，Annu Rev Biochem，79:321-349)。结果展示，更长的crRNA产生更长的双链体，它具有更高的活性水平。为了针对RNA编辑活性系统地比较不同长度的指导物的活性，在荧光素酶报告基因测定中我们设计了30、50、70和84个碱基的指导物以校正终止密码子(图20A-图20F和下表)。我们设计了这些指导物，使得编辑的A的位置存在于在mRNA:crRNA双链体内相对于crRNA特异性决定区的3'端所有可能的偶数距离(即+2、+4等)处。

[2798]

序列名称	序列
指导物_Cas13bC-luc_30_ADAR0顶部	GCATCCTGCGGCCTCTACTCTGCATTCAATT (SEQ ID NO: 179)
指导物_Cas13bC-luc_30_ADAR1顶部	GACCATCCTGCGGCCTCTACTCTGCATTCAA (SEQ ID NO: 180)
指导物_Cas13bC-luc_30_ADAR2顶部	GAAACCATCCTGCGGCCTCTACTCTGCATTC (SEQ ID NO: 181)
指导物_Cas13bC-luc_30_ADAR3顶部	GCTAAACCATCCTGCGGCCTCTACTCTGCAT (SEQ ID NO: 182)
指导物_Cas13bC-luc_30_ADAR4顶部	GTTCTAAACCATCCTGCGGCCTCTACTCTGC (SEQ ID NO: 183)
指导物_Cas13bC-luc_30_ADAR5顶部	GTGTTCTAAACCATCCTGCGGCCTCTACTCT (SEQ ID NO: 184)
指导物_Cas13bC-luc_30_ADAR6顶部	GAATGTTCTAAACCATCCTGCGGCCTCTACT (SEQ ID NO: 185)
指导物_Cas13bC-luc_30_ADAR7顶部	GAGAATGTTCTAAACCATCCTGCGGCCTCTA (SEQ ID NO: 186)
指导物_Cas13bC-luc_30_ADAR8顶部	GATAGAATGTTCTAAACCATCCTGCGGCCTC (SEQ ID NO: 187)
指导物_Cas13bC-luc_30_ADAR9顶部	GCCATAGAATGTTCTAAACCATCCTGCGGCC (SEQ ID NO: 188)
指导物_Cas13bC-luc_3	GTTCCATAGAATGTTCTAAACCATCCTGCGG (SEQ ID NO: 189)

[2799]

0_ADAR10 顶部	
指导物_Cas13bC-luc_30_ADAR11 顶部	GCTTTCCATAGAATGTTCTAAACCATCCTGC (SEQ ID NO: 190)
指导物_Cas13bC-luc_30_ADAR12 顶部	GCTCTTTCCATAGAATGTTCTAAACCATCCT (SEQ ID NO: 191)
指导物_Cas13bC-luc_30_ADAR13 顶部	GATCTCTTTCCATAGAATGTTCTAAACCATC (SEQ ID NO: 192)
指导物_Cas13bC-luc_30_ADAR14 顶部	GGAATCTCTTTCCATAGAATGTTCTAAACCA (SEQ ID NO: 193)
指导物_Cas13bC-luc_50_ADAR0 顶部	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA (SEQ ID NO: 194)
指导物_Cas13bC-luc_50_ADAR1 顶部	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG (SEQ ID NO: 195)
指导物_Cas13bC-luc_50_ADAR2 顶部	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC (SEQ ID NO: 196)
指导物_Cas13bC-luc_50_ADAR3 顶部	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACAT (SEQ ID NO: 197)
指导物_Cas13bC-luc_50_ADAR4 顶部	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC (SEQ ID NO: 198)
指导物_Cas13bC-luc_50_ADAR5 顶部	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC (SEQ ID NO: 199)
指导物_Cas13bC-luc_50_ADAR6 顶部	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG (SEQ ID NO: 200)
指导物_Cas13bC-luc_50_ADAR7 顶部	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC (SEQ ID NO: 201)
指导物_Cas13bC-luc_50_ADAR8 顶部	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT (SEQ ID NO: 202)
指导物	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA

[2800]

_Cas13bC-luc_50_ADAR9 顶部	CTCTGCATTCAATTAC (SEQ ID NO: 203)
指导物_Cas13bC-luc_50_ADAR10 顶部	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT (SEQ ID NO: 204)
指导物_Cas13bC-luc_50_ADAR11 顶部	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA (SEQ ID NO: 205)
指导物_Cas13bC-luc_50_ADAR12 顶部	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC (SEQ ID NO: 206)
指导物_Cas13bC-luc_50_ADAR13 顶部	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT (SEQ ID NO: 207)
指导物_Cas13bC-luc_50_ADAR14 顶部	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC (SEQ ID NO: 208)
指导物_Cas13bC-luc_50_ADAR15 顶部	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT (SEQ ID NO: 209)
指导物_Cas13bC-luc_50_ADAR16 顶部	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT (SEQ ID NO: 210)
指导物_Cas13bC-luc_50_ADAR17 顶部	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA (SEQ ID NO: 211)
指导物_Cas13bC-luc_50_ADAR18 顶部	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC (SEQ ID NO: 212)
指导物_Cas13bC-luc_50_ADAR19 顶部	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC (SEQ ID NO: 213)
指导物_Cas13bC-luc_50_ADAR20 顶部	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG (SEQ ID NO: 214)

[2801]

指导物_Cas13bC-luc_50_ADAR21 顶部	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC (SEQ ID NO: 215)
指导物_Cas13bC-luc_50_ADAR22 顶部	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT (SEQ ID NO: 216)
指导物_Cas13bC-luc_50_ADAR23 顶部	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC (SEQ ID NO: 217)
指导物_Cas13bC-luc_50_ADAR24 顶部	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA (SEQ ID NO: 218)
指导物_Cas13bC-luc_70_ADAR0 顶部	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTAT (SEQ ID NO: 219)
指导物_Cas13bC-luc_70_ADAR1 顶部	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTT (SEQ ID NO: 220)
指导物_Cas13bC-luc_70_ADAR2 顶部	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGT (SEQ ID NO: 221)
指导物_Cas13bC-luc_70_ADAR3 顶部	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTG (SEQ ID NO: 222)
指导物_Cas13bC-luc_70_ADAR4 顶部	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCC (SEQ ID NO: 223)
指导物_Cas13bC-luc_70_ADAR5 顶部	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTT (SEQ ID NO: 224)
指导物_Cas13bC-luc_70_ADAR6 顶部	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTT (SEQ ID NO: 225)
指导物_Cas13bC-luc_70_ADAR7 顶部	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGT (SEQ ID NO: 226)
指导物_Cas13bC-luc_70_ADAR8 顶部	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACAT (SEQ ID NO: 227)

[2802]

指导物_Cas13bC-luc_70_ADAR9 顶部	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAAC (SEQ ID NO: 228)
指导物_Cas13bC-luc_70_ADAR10 顶部	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA (SEQ ID NO: 229)
指导物_Cas13bC-luc_70_ADAR11 顶部	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG (SEQ ID NO: 230)
指导物_Cas13bC-luc_70_ADAR12 顶部	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC (SEQ ID NO: 231)
指导物_Cas13bC-luc_70_ADAR13 顶部	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACAT (SEQ ID NO: 232)
指导物_Cas13bC-luc_70_ADAR14 顶部	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC (SEQ ID NO: 233)
指导物_Cas13bC-luc_70_ADAR15 顶部	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC (SEQ ID NO: 234)
指导物_Cas13bC-luc_70_ADAR16 顶部	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG (SEQ ID NO: 235)
指导物_Cas13bC-luc_70_ADAR17 顶部	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC (SEQ ID NO: 236)
指导物_Cas13bC-luc_70_ADAR18 顶部	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT (SEQ ID NO: 237)
指导物_Cas13bC-luc_70_ADAR19 顶部	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTAC (SEQ ID NO: 238)
指导物_Cas13bC-luc_7	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT

[2803]

0_ADAR20 顶部	(SEQ ID NO: 239)
指导物_Cas13bC-luc_70_ADAR21 顶部	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA (SEQ ID NO: 240)
指导物_Cas13bC-luc_70_ADAR22 顶部	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC (SEQ ID NO: 241)
指导物_Cas13bC-luc_70_ADAR23 顶部	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT (SEQ ID NO: 242)
指导物_Cas13bC-luc_70_ADAR24 顶部	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC (SEQ ID NO: 243)
指导物_Cas13bC-luc_70_ADAR25 顶部	GATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT (SEQ ID NO: 244)
指导物_Cas13bC-luc_70_ADAR26 顶部	GGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT (SEQ ID NO: 245)
指导物_Cas13bC-luc_70_ADAR27 顶部	GACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA (SEQ ID NO: 246)
指导物_Cas13bC-luc_70_ADAR28 顶部	GACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC (SEQ ID NO: 247)
指导物_Cas13bC-luc_70_ADAR29 顶部	GCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC (SEQ ID NO: 248)
指导物_Cas13bC-luc_70_ADAR30 顶部	GCCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG (SEQ ID NO: 249)
指导物_Cas13bC-luc_70_ADAR31 顶部	GGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC (SEQ ID NO: 250)

[2804]

指导物_Cas13bC-luc_70_ADAR32 顶部	GTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT (SEQ ID NO: 251)
指导物_Cas13bC-luc_70_ADAR33 顶部	GCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC (SEQ ID NO: 252)
指导物_Cas13bC-luc_70_ADAR34 顶部	GTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA (SEQ ID NO: 253)
指导物_Cas13bC-luc_84_ADAR0 顶部	GCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAGTCCA (SEQ ID NO: 254)
指导物_Cas13bC-luc_84_ADAR1 顶部	GACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAGTC (SEQ ID NO: 255)
指导物_Cas13bC-luc_84_ADAR2 顶部	GAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACACAG (SEQ ID NO: 256)
指导物_Cas13bC-luc_84_ADAR3 顶部	GCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCACAC (SEQ ID NO: 257)
指导物_Cas13bC-luc_84_ADAR4 顶部	GTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTCAC (SEQ ID NO: 258)
指导物_Cas13bC-luc_84_ADAR5 顶部	GTGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTTTC (SEQ ID NO: 259)
指导物_Cas13bC-luc_84_ADAR6 顶部	GAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTATTT (SEQ ID NO: 260)
指导物_Cas13bC-luc_84_ADAR7 顶部	GAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTTAT (SEQ ID NO: 261)
指导物_Cas13bC-luc_84_ADAR8 顶部	GATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGTTT (SEQ ID NO: 262)

[2805]

指导物_Cas13bC-luc_84_ADAR9 顶部	GCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTGGT (SEQ ID NO: 263)
指导物_Cas13bC-luc_84_ADAR10 顶部	GTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCCTG (SEQ ID NO: 264)
指导物_Cas13bC-luc_84_ADAR11 顶部	GCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTTCC (SEQ ID NO: 265)
指导物_Cas13bC-luc_84_ADAR12 顶部	GCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTTTT (SEQ ID NO: 266)
指导物_Cas13bC-luc_84_ADAR13 顶部	GATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGTTT (SEQ ID NO: 267)
指导物_Cas13bC-luc_84_ADAR14 顶部	GGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACATGT (SEQ ID NO: 268)
指导物_Cas13bC-luc_84_ADAR15 顶部	GTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAACAT (SEQ ID NO: 269)
指导物_Cas13bC-luc_84_ADAR16 顶部	GACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCAAC (SEQ ID NO: 270)
指导物_Cas13bC-luc_84_ADAR17 顶部	GGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGGCA (SEQ ID NO: 271)
指导物_Cas13bC-luc_84_ADAR18 顶部	GTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTCGG (SEQ ID NO: 272)
指导物_Cas13bC-luc_84_ADAR19 顶部	GCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACACATTC (SEQ ID NO: 273)
指导物_Cas13bC-luc_8	GTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTA

[2806]

4_ADAR20 顶部	CATACTGACACAT (SEQ ID NO: 274)
指导物_Cas13bC-luc_84_ADAR21 顶部	GGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGACAC (SEQ ID NO: 275)
指导物_Cas13bC-luc_84_ADAR22 顶部	GCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTGAC (SEQ ID NO: 276)
指导物_Cas13bC-luc_84_ADAR23 顶部	GACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATACTG (SEQ ID NO: 277)
指导物_Cas13bC-luc_84_ADAR24 顶部	GGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACATAC (SEQ ID NO: 278)
指导物_Cas13bC-luc_84_ADAR25 顶部	GATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTACAT (SEQ ID NO: 279)
指导物_Cas13bC-luc_84_ADAR26 顶部	GGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATTAC (SEQ ID NO: 280)
指导物_Cas13bC-luc_84_ADAR27 顶部	GACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAATT (SEQ ID NO: 281)
指导物_Cas13bC-luc_84_ADAR28 顶部	GACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTCAA (SEQ ID NO: 282)
指导物_Cas13bC-luc_84_ADAR29 顶部	GCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCATTC (SEQ ID NO: 283)
指导物_Cas13bC-luc_84_ADAR30 顶部	GCCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGCAT (SEQ ID NO: 284)
指导物_Cas13bC-luc_84_ADAR31 顶部	GGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCTGC (SEQ ID NO: 285)

[2807]

指导物_Cas13bC-luc_84_ADAR32 顶部	GTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACTCT (SEQ ID NO: 286)
指导物_Cas13bC-luc_84_ADAR33 顶部	GCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTACT (SEQ ID NO: 287)
指导物_Cas13bC-luc_84_ADAR34 顶部	GTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTCTA (SEQ ID NO: 288)
指导物_Cas13bC-luc_84_ADAR35 顶部	GGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCCTC (SEQ ID NO: 289)
指导物_Cas13bC-luc_84_ADAR36 顶部	GTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGGCC (SEQ ID NO: 290)
指导物_Cas13bC-luc_84_ADAR37 顶部	GCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGCGG (SEQ ID NO: 291_
指导物_Cas13bC-luc_84_ADAR38 顶部	GGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCTGC (SEQ ID NO: 292)
指导物_Cas13bC-luc_84_ADAR39 顶部	GCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATCCT (SEQ ID NO: 293)
指导物_Cas13bC-luc_84_ADAR40 顶部	GCGCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCATC (SEQ ID NO: 294)
指导物_Cas13bC-luc_84_ADAR41 顶部	GGTCGCCGCCCTTGGTTCCTTGACCCAACACGTATGTACCAGGTTCCTGGAACTGGAATCTCTTTCCATAGAATGTTCTAAACCA (SEQ ID NO: 295)

[2808] 逆转因果疾病突变

[2809] 合成了具有病原性G>A突变(其将早期终止位点引入基因中)的下表中的三个基因并且将它们整合到非人类细胞系中。测试了Cas13b12-ADAR2通过将终止密码子UAG改变为UIG(UGG)并由此恢复蛋白质翻译来校正转录物的能力。

[2810]

蛋白质长	全长候选物	基因	疾病

[2811]

	度(aa)			
	498	NM_004992.3(MECP2):c.311G>A (p.Trp104Ter)	MECP2	雷特综合征 (Rett syndrome)
	414	NM_007375.3(TARDBP):c.943G>A (p.Ala315Thr)	TARDBP	肌萎缩性侧索硬化症 10 型
	393	NM_000546.5(TP53):c.273G>A (p.Trp91Ter)	TP53	Li-Fraumeni 综合征 (Li-Fraumeni syndrome)/遗传性癌症易感综合征

[2812] 下表5中示出了另外48个病原性G>A突变连同伴随的疾病。围绕这些突变合成了200bp的片段而非整个基因，并将这些片段克隆在GFP前面。当早期停止位点恢复时，将允许GFP的翻译，并且可以通过高通量荧光以及RNA测序测量校正。

[2813] 表5

[2814]

	候选物	基因	疾病
1	NM_004006.2(DMD):c.3747G>A (p.Trp1249Ter)	DMD	杜兴氏肌营养不良
2	M_000344.3(SMN1):c.305G>A (p.Trp102Ter)	SMN1	脊髓型肌萎缩，II 型|Kugelberg-Welander 病
3	NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)	CFTR	囊性纤维化|遗传性胰腺炎|未提供|ataluren 反应-功效;
4	NM_004562.2(PRKN):c.1358G>A (p.Trp453Ter)	PRKN	帕金森病 2
5	NM_017651.4(AHI1):c.2174G>A (p.Trp725Ter)	AHI1	乔伯特综合征 3
6	NM_000238.3(KCNH2):c.3002G>A (p.Trp1001Ter)	KCNH2	长 QT 综合征|未提供
7	NM_000136.2(FANCC):c.1517G>A (p.Trp506Ter)	FANCC|C9orf3	范可尼贫血，互补群 C
8	NM_001009944.2(PKD1):c.12420G>A (p.Trp4140Ter)	PKD1	成人型多囊肾病
9	NM_177965.3(C8orf37):c.555G>A (p.Trp185Ter)	C8orf37	色素性视网膜炎 64
10	NM_000833.4(GRIN2A):c.3813G>A (p.Trp1271Ter)	GRIN2A	局灶性癫痫，伴有言语障碍，且伴有或不伴有智力迟钝
11	NM_000548.4(TSC2):c.2108G>A (p.Trp703Ter)	TSC2	结节性硬化症 2|结节性硬化症综合征
12	NM_000267.3(NF1):c.7044G>A (p.Trp2348Ter)	NF1	神经纤维瘤病，1 型

[2815]

13	NM_000520.5(HEXA):c.1454G>A (p.Trp485Ter)	HEXA	泰-萨二氏病(Tay-Sachs disease)
14	NM_130838.1(UBE3A):c.2304G>A (p.Trp768Ter)	UBE3A	安格曼综合征(Angelman syndrome)
15	NM_000543.4(SMPD1):c.168G>A (p.Trp56Ter)	SMPD1	尼曼-皮克病(Niemann-Pick disease)，A型
16	NM_000218.2(KCNQ1):c.1175G>A (p.Trp392Ter)	KCNQ1	长QT综合征
17	NM_000256.3(MYBPC3):c.3293G>A (p.Trp1098Ter)	MYBPC3	原发性家族性肥厚型心肌病
18	NM_000038.5(APC):c.1262G>A (p.Trp421Ter)	APC	家族性腺瘤性息肉病1
19	NM_000249.3(MLH1):c.1998G>A (p.Trp666Ter)	MLH1	林奇综合征
20	NM_000054.4(AVPR2):c.878G>A (p.Trp293Ter)	AVPR2	X连锁性肾原性尿崩症
21	NM_001204.6(BMPR2):c.893G>A (p.W298*)	BMPR2	原发性肺动脉高压
22	NM_004560.3(ROR2):c.2247G>A (p.Trp749Ter)	ROR2	指过短B1型
23	NM_000518.4(HBB):c.114G>A (p.Trp38Ter)	HBB	β^0^地中海贫血/β地中海贫血
24	NM_024577.3(SH3TC2):c.920G>A (p.Trp307Ter)	SH3TC2	夏科-马里-图思病，4C型
25	NM_206933.2(USH2A):c.9390G>A (p.Trp3130Ter)	USH2A	乌谢尔综合征，2A型
26	NM_000179.2(MSH6):c.3020G>A (p.Trp1007Ter)	MSH6	林奇综合征
27	NM_002977.3(SCN9A):c.2691G>A (p.Trp897Ter)	SCN9A|LOC101929680	常染色体隐性遗传性先天性疼痛不敏感
28	NM_000090.3(COL3A1):c.30G>A (p.Trp10Ter)	COL3A1	Ehlers-Danlos综合征，4型
29	NM_000551.3(VHL):c.263G>A (p.Trp88Ter)	VHL	Von Hippel-Lindau综合征|未提供
30	NM_015627.2(LDLRAP1):c.65G>A (p.Trp22Ter)	LDLRAP1	常染色体隐性遗传性高胆固醇血症
31	NM_000132.3(F8):c.3144G>A (p.Trp1048Ter)	F8	遗传性因子VIII缺乏症

[2816]

32	NM_002185.4(IL7R):c.651G>A (p.Trp217Ter)	IL7R	常染色体隐性遗传性 T 细胞阴性 B 细胞阳性 NK 细胞阳性重度联合免疫缺陷
33	NM_000527.4(LDLR):c.1449G >A (p.Trp483Ter)	LDLR	家族性高胆固醇血症
34	NM_002294.2(LAMP2):c.962G >A (p.Trp321Ter)	LAMP2	Danon 病
35	NM_000271.4(NPC1):c.1142G >A (p.Trp381Ter)	NPC1	尼曼-皮克病，C1 型
36	NM_000267.3(NF1):c.1713G> A (p.Trp571Ter)	NF1	神经纤维瘤病，1 型
37	NM_000035.3(ALDOB):c.888 G>A (p.Trp296Ter)	ALDOB	遗传性果糖尿病
38	NM_000090.3(COL3A1):c.383 3G>A (p.Trp1278Ter)	COL3A1	Ehlers-Danlos 综合征，4 型
39	NM_001369.2(DNAH5):c.8465 G>A (p.Trp2822Ter)	DNAH5	原发性纤毛运动障碍症
40	NM_178443.2(FERMT3):c.48G >A (p.Trp16Ter)	FERMT3	白细胞粘附缺陷，III 型
41	NM_005359.5(SMAD4):c.906 G>A (p.Trp302Ter)	SMAD4	青少年息肉病综合征
42	NM_032119.3(ADGRV1):c.740 6G>A (p.Trp2469Ter)	ADGRV1	乌谢尔综合征，2C 型
43	NM_000206.2(IL2RG):c.710G> A (p.Trp237Ter)	IL2RG	X 连锁性重度联合免疫缺陷
44	NM_007294.3(BRCA1):c.5511 G>A (p.Trp1837Ter)	BRCA1	家族性乳腺癌|家族性乳腺癌-卵巢癌 1
45	NM_130799.2(MEN1):c.1269G >A (p.Trp423Ter)	MEN1	遗传性癌症易感综合征
46	NM_000071.2(CBS):c.162G>A (p.Trp54Ter)	CBS	由 CBS 缺乏症引起的高胱氨酸尿症
47	NM_000059.3(BRCA2):c.582G >A (p.Trp194Ter)	BRCA2	家族性乳腺癌|家族性乳腺癌-卵巢癌 2
48	NM_000053.3(ATP7B):c.2336 G>A (p.Trp779Ter)	ATP7B	威尔逊氏病

[2817] 实施例3

[2818] 有效且精确的核酸编辑就治疗遗传性疾病而言具有广阔的前景，尤其是在RNA水平上，在RNA水平上可以挽救疾病相关转录物，以产生功能性蛋白质产品。VI型CRISPR-Cas系统含有可编程的单效应RNA指导的RNA酶Cas13。在此，我们概述了VI型系统的多样性，以工程化能够进行稳健敲低的Cas13直系同源物，并且通过使用无催化活性的Cas13(dCas13)

来将腺苷脱氨酶活性引导至哺乳动物细胞中的转录物来证实RNA编辑。通过将ADAR2脱氨酶结构域与dCas13融合并工程化指导RNA以产生最佳的RNA双链体底物，我们实现了从特定的单个腺苷到肌苷(在翻译过程中被读出为鸟苷)的靶向编辑，其效率常规地在20％-40％的范围内，并且最高达89％。该系统(称之为用于可编程A至I替换的RNA编辑(REPAIR))可以进一步工程化以实现高特异性。工程化变体REPAIRv2显示出的特异性增加大于170倍，同时保持了稳健的中靶A至I编辑。我们使用REPAIRv2编辑含有已知病原性突变的全长转录物，并且产生适合于包装在腺相关病毒(AAV)载体中的功能性截短形式。REPAIR提出了一个有前途的RNA编辑平台，所述平台在研究、治疗学和生物技术方面具有广泛的适用性。精确的核酸编辑技术对于研究细胞功能和作为新颖疗法非常有价值。尽管当前的编辑工具(诸如Cas9核酸酶)可以实现对基因组基因座的可编程修饰，但是由于插入或缺失或需要供体模板进行精确编辑，编辑常常是异质的。碱基编辑器(诸如dCas9-APOBEC融合物)允许进行编辑而不会产生双链断裂，但由于胞苷脱氨酶活性的性质(编辑靶标窗口中的任何胞苷)，因此可能缺乏精确性。此外，对原间隔区相邻基序(PAM)的要求限制了可能的编辑位点的数量。在此，我们描述了精确且灵活的RNA碱基编辑工具的开发，所述工具使用了来自VI型原核生物成簇规律间隔短回文重复序列(CRISPR)自适应免疫系统的RNA指导的RNA靶向Cas13酶。

[2819] 精确的核酸编辑技术对于研究细胞功能和作为新颖疗法非常有价值。当前基于可编程核酸酶(诸如原核生物成簇规律间隔短回文重复序列(CRISPR)相关的核酸酶Cas9(1-4)或Cpf1(5))的编辑工具已广泛用于介导靶向DNA切割，进而通过非同源末端连接(NHEJ)驱动靶向的基因破坏或通过模板依赖性同源直接修复(HDR)驱动精确的基因编辑(6)。NHEJ利用在分裂细胞和有丝分裂后细胞中均活跃的宿主机器，并且通过产生可导致蛋白质编码基因移码的插入或插入缺失突变的混合物，提供有效的基因破坏。相比之下，HDR由表达主要限于复制细胞的宿主机器介导。因此，有丝分裂后细胞中基因编辑能力的开发仍然是主要挑战。最近，DNA碱基编辑器(诸如使用无催化活性的Cas9(dCas9)将胞苷脱氨酶活性靶向至特定的基因组靶标，以在靶标窗口内实现胞嘧啶向胸腺嘧啶的转化)允许在不产生DNA双链断裂的情况下进行编辑并显著降低插入缺失的形成(7,8)。然而，由于Cas9在编辑位点处需要原间隔区相邻基序(PAM)，因此DNA碱基编辑器的靶向范围受到限制(9)。在此，我们描述了使用VI型CRISPR相关RNA指导的RNA酶Cas13开发精确且灵活的RNA碱基编辑技术(10-13)。

[2820] Cas13酶具有两个介导精确RNA切割的结合高级真核生物和原核生物核苷酸(HEPN)的内切RNA酶结构域(10,11)。迄今为止，已鉴定出三个Cas13蛋白家族:Cas13a(以前称为C2c2)、Cas13b和Cas13c(12,13)。我们最近报道，Cas13a酶可以适合用作核酸检测的工具(14)以及哺乳动物和植物细胞的RNA敲低和转录物追踪的工具(15)。Cas13酶的RNA指导性质使其成为RNA结合和干扰应用的有吸引力工具。

[2821] 作用于酶的RNA(ADAR)家族的腺苷脱氨酶通过使腺苷水解脱氨为肌苷来介导转录物的内源性编辑，肌苷是在翻译和剪接中与鸟苷功能上等同的核碱基(16)。有两种功能性人ADAR直系同源物ADAR1和ADAR2，它们由N端双链RNA结合结构域和C端催化脱氨结构域组成。ADAR1和ADAR2的内源性靶位点含有实质性的双链同一性，并且催化结构域需要双链体区域才能在体外和体内进行有效编辑(17,18)。尽管ADAR蛋白具有优选的可能限制靶向的

潜在灵活性的编辑基序,但超活性突变体(诸如ADAR(E488Q)(19))放松了序列限制,并且提高腺苷到肌苷的编辑率。ADAR优先使RNA双链体中与胞苷碱基相对的腺苷脱氨基(20),从而为精确的碱基编辑提供了有希望的机会。尽管以前的方法已经经由RNA指导物工程化靶向ADAR融合物(21-24),但是尚未报道这些方法的特异性,并且它们各自的靶向机制依赖于RNA-RNA杂交,而不借助于可能增强靶标识别和严格性的蛋白质配偶体。

[2822] 在此,我们测定了整个Cas13酶家族在哺乳动物细胞中的RNA敲低活性,并且鉴定出来自普雷沃氏菌属种P5-125的Cas13b直系同源物(PspCas13b)对于哺乳动物细胞应用而言是最有效且最具特异性的。然后,我们将ADAR2脱氨酶结构域(ADARDD)与无催化活性的PspCas13b融合,并且证实了用于对报告基因和内源性转录物以及疾病相关突变的可编程A到I(G)替换(REPAIR)的RNA编辑。最后,我们采用合理的诱变方案来改善dCas13b-ADAR2DD融合物的特异性,以产生特异性增加超过170倍的REPAIRv2。

[2823] 方法

[2824] 细菌构建体的设计和克隆

[2825] 在Lac启动子的控制下,将经哺乳动物密码子优化的Cas13b构建体克隆到氯霉素抗性pACYC184载体中。然后,在pJ23119启动子的控制下,将由BsaI限制性位点隔开的两个相应的正向重复(DR)序列插入Cas13b的下游。最后,使用T4 PNK(New England Biolabs)将用于靶向间隔区的寡核苷酸磷酸化、退火、并且使用T7连接酶(Enzymatics)连接到BsaI消化的载体中,以生成靶向Cas13b载体。

[2826] 细菌PFS筛选

[2827] 通过使用NEB Gibson Assembly(New England Biolabs)插入含有带有2个5'随机化核苷酸和4个3'随机化核苷酸的Cas13b靶标的PCR产物来克隆氨苄青霉素抗性质粒用于PFS筛选,所述5'随机化核苷酸和所述3'随机化核苷酸由在紧邻氨苄青霉素抗性基因bla起始密码子下游的靶位点隔开。然后将100ng氨苄青霉素抗性靶质粒与65-100ng氯霉素抗性Cas13b细菌靶向质粒一起电穿孔到Endura电感受态细胞中。将质粒添加到细胞中,在冰上孵育15分钟,使用制造商的方案进行电穿孔,然后将950uL的回收培养基添加到细胞中,之后在37C下生长一小时。将生长物铺板到氯霉素和氨苄青霉素双重选择板上。将生长物的连续稀释液用于估计cfu/ng DNA。铺板后16小时,将细胞从板上刮下,并且使用Qiagen Plasmid Plus Maxi试剂盒(Qiagen)收获存活的质粒DNA。通过PCR扩增存活的Cas13b靶序列及其侧翼区域,并且使用Illumina NextSeq进行测序。为了评估PFS优选性,提取原始文库中含有随机化核苷酸的位置,并使用自定义python脚本提取存在于两个生物复制品中的仅相对于载体条件耗尽的序列。然后,将Cas13b条件下的PFS丰度与仅载体对照之比的-log2用于计算优选的基序。具体地,将所有-log2(样品/载体)耗尽比高于特定阈值的序列都用于生成序列基序的weblogo(weblogo.berkeley.edu)。表9中列出了用于针对每个Cas13b直系同源物生成weblogo的特定耗尽比值。

[2828] 用于RNA干扰的哺乳动物构建体的设计和克隆

[2829] 为了产生用于在哺乳动物细胞中测试Cas13直系同源物的载体,对经哺乳动物密码子优化的Cas13a、Cas13b和Cas13c基因进行PCR扩增,并且在EF1α启动子的控制下,将其金门(golden-gate)克隆到含有双重NLS序列和C端msfGFP的哺乳动物表达载体中。为了进一步优化,在EF1α启动子的控制下,将Cas13直系同源物金门克隆到含有不同C端定位标签

的目的载体中。

[2830] 通过扩增海肾荧光素酶和海萤荧光素酶编码DNA、EF1α和CMV启动子的PCR和使用NEB Gibson Assembly(New England Biolabs)的组件来克隆双重荧光素酶报告基因。

[2831] 为了表达Cas13a、Cas13b或Cas13c直系同源物的哺乳动物指导RNA,合成了具有金门受体位点的相应正向重复序列,并且通过限制性消化克隆将其在U6表达下克隆。然后通过金门克隆将各个指导序列克隆到每个直系同源物的相应表达主链中。

[2832] 针对Cas13干扰特异性的合并错配文库的克隆

[2833] 通过PCR产生合并错配文库靶位点。将含有G-荧光素酶中的在靶标内每个位置处含有94%正确碱基和2%每个错误碱基的混合物的半简并靶序列的寡核苷酸用作一个引物,并且将对应于G-荧光素酶的非靶向区的寡核苷酸用作PCR反应中的第二引物。然后使用NEB Gibson assembly(New England Biolabs)将错配文库靶标克隆到双重荧光素酶报告基因中以代替野生型G-荧光素酶。

[2834] 用于RNA编辑的哺乳动物构建体的设计和克隆

[2835] 经由在HEPN结构域的催化位点处的两个组氨酸到丙氨酸突变(H133A/H1058A)使PspCas13b失去催化活性(dPspCas13b)。合成人类ADAR1和ADAR2的脱氨酶结构域,并且对其进行PCR扩增以gibson克隆到pcDNA-CMV载体主链中,并且经由GS或GSGGGGS

[2836] (SEQ ID NO:296)接头在C端处与dPspCas13b融合。对于我们测试不同接头的实验,我们在dPspCas13b与ADAR2dd之间克隆了以下其他接头:GGGGSGGGGSGGGGS、EAAAK(SEQ ID NO:297)、GGSGGSGGSGGSGGSGGS(SEQ ID NO:298)和SGSETPGTSESATPES(XTEN)(SEQ ID NO:299)。通过将适当的突变体gibson克隆到dPspCas13b-GSGGGGS(SEQ ID NO:761)主链中来产生特异性突变体。

[2837] 通过在用于敲低实验的荧光素酶报告载体中产生W85X突变(TGG>TAG),产生了用于测量RNA编辑活性的荧光素酶报告载体。该报告载体表达功能性Gluc以作为归一化对照,但是由于添加了早期终止位点,所以所述功能性Gluc是缺陷型Cluc。为了测试ADAR编辑基序优选性,我们克隆了在Cluc的第85位密码子(XAX)处腺苷周围的每个可能基序。

[2838] 为了测试REPAIR的PFS优选性,我们克隆了合并的质粒文库,所述文库在靶区域和腺苷编辑位点的上游含有6个碱基对的简并PFS序列。所述文库是由Integrated DNA Technologies(IDT)以超聚物形式合成,并且通过退火引物并将DNA聚合酶I的Klenow片段(New England Biolabs)填充在序列中而成为双链的。然后将该含有简并序列的dsDNA片段gibson克隆到消化的报告载体中,然后将该载体异丙醇沉淀并纯化。然后将克隆的文库电穿孔到Endura感受态大肠杆菌细胞(Lucigen)中并且铺板在245mmx245mm方形生物测定板(Nunc)上。16小时后,收获菌落并且使用无内毒素的MACHEREY-NAGEL midiprep试剂盒进行中量制备。通过下一代测序验证克隆的文库。

[2839] 为了克隆疾病相关突变以测试REPAIR活性,选择了与ClinVar中定义的疾病发病机理相关的34个G>A突变,并且在CMV启动子下将这些突变周围的200bp区域金门克隆在mScarlett与EGFP之间。选择AVPR2和FANCC中的两个另外的G>A突变用于在EF1α表达下Gibson克隆整个基因序列。

[2840] 为了表达用于REPAIR的哺乳动物指导RNA,合成了具有金门受体位点的PspCas13b正向重复序列,并且通过限制性消化克隆将其在U6表达下克隆。然后通过金门克隆将各个

指导物克隆到该表达主链中。

[2841] 哺乳动物细胞培养

[2842] 在HEK293FT品系(美国典型培养物保藏中心(ATCC))中进行了哺乳动物细胞培养实验,将所述品系在具有高葡萄糖、丙酮酸钠和GlutaMAX(Thermo Fisher Scientific)、额外补充有1×青霉素-链霉素(Thermo Fisher Scientific)和10%胎牛血清(VWR Seradigm)的杜氏改良型伊格尔培养基(Dulbecco's Modified Eagle Medium)中生长。将细胞保持在低于80%的汇合度。

[2843] 除非另有说明,否则所有转染均用Lipofectamine 2000(Thermo Fisher Scientific)在涂覆有聚-D-赖氨酸(BD Biocoat)的96孔板中进行。在转染前十六个小时将细胞以大约20,000个细胞/孔铺板,以确保在转染时达到90%汇合度。对于板上的每个孔,将转染质粒与Opti-MEM I还原血清培养基(Thermo Fisher)合并至总共25μl。分别地,将24.5ul的Opti-MEM与0.5ul的Lipofectamine 2000合并。然后将质粒和Lipofectamine溶液合并,并且孵育5分钟,然后将其吸移到细胞上。

[2844] RNA敲低哺乳动物细胞测定

[2845] 为了评定在具有报告构建体的哺乳动物细胞中的RNA靶向,将150ng的Cas13构建体与300ng的指导表达质粒和12.5ng的敲低报告构建体共转染。转染后48小时,从细胞中移取出含有分泌型荧光素酶的培养基,将其在PBS中以1:5稀释,并且使用注射方案在读板仪(Biotek Synergy Neo2)上使用BioLux海萤和Biolux海肾荧光素酶测定试剂盒(New England Biolabs)测量活性。进行的所有重复都是生物学重复。

[2846] 为了靶向内源性基因,将150ng的Cas13构建体与300ng的指导表达质粒共转染。转染后48小时,将细胞裂解并且收获RNA并使用先前描述的Cells-to-Ct试剂盒(Thermo Fisher Scientific)的[引用协议(CITE PROTOCOLS)]修改版进行逆转录。使用针对用于KRAS转录物的TaqMan qPCR探针(Thermo Fisher Scientific)、GAPDH对照探针(Thermo Fisher Scientific)和快速高级主混合物(Thermo Fisher Scientific),通过qPCR测量cDNA表达。在LightCycler 480仪器II(Roche)上读取qPCR反应,以384孔格式进行四次5ul技术重复。

[2847] 使用错配靶标的合并文库评估RNA特异性

[2848] 使用接种在6孔板中的HEK293FT细胞测试了Cas13干扰错配靶标文库的能力。使用2400ng Cas13载体、4800ng指导物和240ng错配的靶标文库转染约70%汇合细胞。转染后48小时,收获细胞并且使用QIAshredder(Qiagen)和Qiagen RNeasy微型试剂盒提取RNA。按照制造商的基因特异性引发方案使用qScript Flex cDNA合成试剂盒(Quantabio)和Gluc特异性RT引物,对1ug提取RNA进行逆转录。然后扩增cDNA并在Illumina NextSeq上测序。

[2849] 通过对每个序列的读数进行计数来分析测序,并且通过确定log2(-读取计数比)值来计算耗尽得分,其中读取计数比是靶向指导物条件与非靶向指导物条件的读取计数之比。该得分值表示序列上Cas13的活性水平,较高的值表示更强的耗尽,并且因此表示较高的Cas13切割活性。确定了单错配和双错配序列的单独分布,并将其绘制为热图,所述热图具有针对每个错配身份的耗尽得分。对于双错配序列,绘制了给定位置处所有可能的双错配的平均值。

[2850] 通过RNA测序对哺乳动物细胞中的Cas13进行全转录组分析

[2851] 为了测量全转录组特异性,共转染了150ng Cas13构建体、300ng指导表达质粒和15ng敲低报告构建体;对于shRNA条件,共转染了300ng shRNA靶向质粒、15ng敲低报告构建体和150ng EF1α驱动mCherry(以平衡报告基因负载)。转染后48小时,用RNeasy Plus微型试剂盒(Qiagen)纯化RNA,使用NEBNext Poly(A)mRNA磁性分离模块(New England Biolabs)选择mRNA,并且将其制备用于用Illumina的NEBNext Ultra RNA文库制备试剂盒(New England Biolabs)进行测序。然后在NextSeq(Illumina)上对RNA测序文库进行测序。

[2852] 为了分析全转录组测序数据,使用Bowtie和RSEM版本1.2.31利用默认参数参考序列GRCh38组件比对读数[引用RSEM:在用或不用参考基因组的情况下从RNA测序数据进行准确的转录定量]。将转录物表达定量为log2(TPM+1),过滤基因以使log2(TPM+1)>2.5。为了选择差异表达的基因,仅考虑差异变化>2或<.75的基因。在三个靶向重复品与非靶向重复品上学生T检验评估差异表达的统计显著性,并且通过Benjamini-Hochberg程序针对<0.01%的错误发现率进行过滤。

[2853] 哺乳动物细胞转染中的ADAR RNA编辑

[2854] 为了评估哺乳动物细胞中的REPAIR活性,我们转染了150ng REPAIR载体、300ng指导表达质粒和40ng RNA编辑报告基因。48小时后,从细胞中收获RNA并且使用先前描述的方法[引用JJ]和基因特异性逆转录引物进行逆转录。然后使用NEBNext高保真度2X PCR主混合物对提取的cDNA进行两轮PCR,以添加Illumina衔接子和样品条形码。然后在Illumina NextSeq或MiSeq上对文库进行下一代测序。然后在测序窗口内评估所有腺苷处的RNA编辑率。

[2855] 在荧光素酶报告基因被靶向用于RNA编辑的实验中,我们在RNA收获之前还收获了具有分泌型荧光素酶的培养基。在这种情况下,因为校正后的Cluc可能处于较低水平,因此我们没有稀释培养基。我们使用注射方案在读板仪(Biotek Synergy Neo2)上使用BioLux海萤和Biolux海肾荧光素酶测定试剂盒(New England Biolabs)测量了荧光素酶活性。进行的所有重复都是生物学重复。

[2856] PFS结合哺乳动物筛选

[2857] 为了确定PFS对编辑效率的贡献,将625ng PFS靶标文库、4.7ug指导物和2.35ug REPAIR共转染在铺板在225cm2烧瓶中的HEK293FT细胞上。将质粒与33ul PLUS试剂(Thermo Fisher Scientific)混合,用Opti-MEM调节至533ul,孵育5分钟,再与30ul Lipofectamine 2000和500ul Opti-MEM合并,再孵育5分钟,然后将其吸移到细胞上。转染后48小时,使用RNeasy Plus微型试剂盒(Qiagen)收获RNA,使用基因特异性引物用qScript Flex(Quantabio)逆转录,并使用NEBNext高保真度2X PCR主混合物(New England Biolabs)以两轮PCR进行扩增以添加Illumina衔接子和样品条形码。将所述文库在Illumina NextSeq上测序,并且将靶腺苷处的RNA编辑率与PFS身份进行映射。为了增加覆盖率,PFS在计算上被折叠为4个核苷酸。计算每个PFS的REPAIR编辑率,在减去相应PFS的非靶向率的情况下将其在生物学复制品上求平均值。

[2858] 全转录组测序以评价ADAR编辑特异性

[2859] 为了分析整个转录组的脱靶RNA编辑位点,我们使用RNeasy Plus微型制备试剂盒(Qiagen)在转染后48小时从细胞中收获总RNA。然后使用NEBNext Poly(A)mRNA磁性分离模块(NEB)富集mRNA级分,然后将此RNA制备用于使用Illumina的NEBNext Ultra RNA文库制

备试剂盒(NEB)进行测序。然后将文库在Illumina NextSeq上进行测序并加载成使得每个样品读取至少500万次。

[2860] 靶向和全转录组实验的RNA编辑分析

[2861] 为了分析全转录组RNA编辑RNA测序数据,将序列文件随机降采样至500万个读取。使用添加了Gluc和Cluc序列的参考序列GRCh38组件生成了索引,并且使用Bowtie/RSEM版本1.3.0对读数进行比对和定量。然后对比对BAM进行分类并且使用REDitool[引用]用以下参数分析RNA编辑位点:-t 8-e-d-l-U[AG或TC]-p-u-m20-T6-0-W-v 1-n 0.0。在未转染或EGFP转染的条件下发现的任何显著的编辑均视为SNP或转染伪像,并将其从脱靶分析中过滤掉。如果Fisher的精确检验得出的p值小于0.5,并且3个生物学重复品中至少有2个鉴定出编辑位点,则认为脱靶是显著的。

[2862] 为了分析每个脱靶的预测变异效应,使用作为UCSC基因组浏览器工具套件的一部分的变异注释整合器(https://genome.ucsc.edu/cgi-bin/hgVai)使用SIFT和PolyPhen-2注释对脱靶编辑位点列表进行了分析。要声明脱靶基因是否是致癌基因,请参阅COSMIC癌症体细胞突变目录(cancer.sanger.ac.uk)中的致癌注释数据库。

[2863] 为了分析REPAIR构建体是否干扰RNA水平,将从RSEM分析输出的每百万转录物(TPM)值用于表达计数,并且通过取log2(TPM+1)转化为对数空间。为了找到差异调控的基因,在三个靶向指导物重复品与三个非靶向指导物重复品上进行了学生t检验。仅对log2(TPM+1)值大于2.5的基因进行统计分析,并且只有倍数变化大于2或小于0.8的基因才被认为是差异调控的。如果基因具有小于0.01的错误发现率,则报告所述基因。

[2864] 结果

[2865] 对哺乳动物细胞中Cas13家族成员的综合表征

[2866] 我们先前开发了LwaCas13a用于哺乳动物敲低应用,但是它需要msfGFP稳定结构域才能有效敲低,尽管特异性很高,但敲低效率并没有始终低于50%(15)。我们试图通过表征一组遗传多样的Cas13家族成员以评定其在哺乳动物细胞中的RNA敲低活性来鉴定出一种更稳健的RNA靶向CRISPR系统(图49A)。我们将21个Cas13a、15个Cas13b和7个Cas13c哺乳动物密码子优化的直系同源物(表6)克隆到具有N端和C端核输出信号(NES)序列和C端msfGFP的表达载体中,以增强蛋白质稳定性。为了测定在哺乳动物细胞中的干预,我们设计了双重报告构建体,所述构建体在单独的启动子下表达直系同源的海肾(Gluc)和海萤(Cluc)荧光素酶,从而使一种荧光素酶充当Cas13干预活性的量度,而另一种则可以作为内部对照。对于每个直系同源物,我们使用来源于氨苄青霉素干预测定的Cas13b PFS基序(图55;表7)和来自先前Cas13a活性报道的3'H PFS设计了与PFS相容的指导RNA(10)。

[2867] 我们用Cas13表达、指导RNA和报告质粒转染了HEK293FT细胞,并在48小时后定量了靶向Gluc的水平。对每个Cas13直系同源物测试两个指导RNA揭示了一系列的活性水平,包括在两个指导RNA上相对于最近表征的LwaCas13a具有相似或增加干预的五个Cas13b直系同源物(图49B)。我们选择了这五个Cas13b直系同源物,以及前两个Cas13a直系同源物,以进行进一步的工程化。

[2868] 接下来,我们测试了在没有msfGFP的情况下,Cas13介导的Gluc的敲低,以选择不需要稳定结构域来实现稳健活性的直系同源物。我们假设,除了msfGFP外,Cas13活性还可能受到亚细胞定位的影响,如先前针对LwaCas13a优化所报道的(15)。因此,我们测试了在C

端与六个不带msfGFP的不同定位标签之一融合的七个选择的Cas13直系同源物的干扰活性。使用荧光素酶报告基因测定，我们发现在C端与HIV Rev基因NES融合的PspCas13b和PguCas13b和在C端与MAPK NES融合的RanCas13b具有最高水平的干扰活性(图56A)。为了进一步区分热点直系同源物的活性水平，我们将三个优化的Cas13b构建体与最佳的LwaCas13a-msfGFP融合物和shRNA针对其使用位置匹配的指导物敲低KRAS转录物的能力进行比较(图56B)。我们观察到PspCas13b的干扰水平最高(平均敲低率为62.9％)，由此选择了该PspCas13b来与LwaCas13a进行进一步比较。

[2869] 为了更严格地定义PspCas13b和LwaCas13a的活性水平，我们设计了沿Gluc和Clue平铺的位置匹配指导物，并且使用我们的荧光素酶报告基因测定测定了它们的活性。我们分别测试了靶向Gluc和Cluc的93个和20个位置匹配指导物，发现相对于LwaCas13a，PspCas13b具有始终增加的敲低水平(PspCas13b的平均值为92.3％，而LwaCas13a的敲低率为40.1％)(图49C、图49D)。

[2870] Cas13哺乳动物干扰活性的特异性

[2871] 为了表征PspCas13b和LwaCas13a的干扰特异性，我们设计了在整个靶序列以及三个侧翼5’和3’碱基对中含有单错配和双错配的荧光素酶靶标的质粒文库(图56C)。我们用LwaCas13a或PspCas13b、靶向未修饰靶序列的固定指导RNA以及对应于适当系统的错配靶标文库转染了HEK293FT细胞。然后，我们对未切割的转录物进行了靶向RNA测序，以定量错配靶序列的耗尽。我们发现LwaCas13a和PspCas13b具有相对不耐受单错配的中心区域，所述中心区域对于PspCas13b靶标从碱基对12-26扩展并且对于LwaCas13a靶标从碱基对13-24扩展(图56D)。与单突变相比，双错配的耐受性甚至更低，在较大的窗口中观察到的敲低活性很小，在各自的靶标中，所述窗口对于PspCas13b从碱基对12-29扩展并且对于LwaCas13a从碱基对8-27扩展(图56E)。另外，由于在靶序列的5’和3’端侧翼的三个核苷酸中含有错配，我们可以评定PFS对Cas13敲低活性的限制。测序表明，几乎所有PFS组合都可实现稳健的敲低，这表明对于测试的任一种酶，可能都不存在PFS对哺乳动物细胞中的干扰的限制。这些结果表明Cas13a和Cas13b显示出相似的序列约束和对错配的敏感性。

[2872] 接下来，我们表征了PspCas13b和LwaCas13a在转录组的mRNA级分上的干扰特异性。我们进行了转录组范围的mRNA测序，以检测显著差异表达的基因。LwaCas13a和PspCas13b展示了稳健的Gluc敲低作用(图49E、图49F)，并且与位置匹配的显示数百个脱靶的shRNA相比具有高度特异性(图49G)。

[2873] Cas13-ADAR融合物实现靶向RNA编辑

[2874] 鉴于PspCas13b在哺乳动物细胞中实现了一致、稳健且特异性的mRNA敲低，我们设想其可以适合作为RNA结合平台来募集ADAR的脱氨酶结构域($ADAR_{DD}$)以进行可编程RNA编辑。为了工程化缺乏核酸酶活性的PspCas13b(dPspCas13b，据此称之为dCas13b)，我们使HEPN结构域中的保守催化残基突变，并且观察到荧光素酶RNA敲低活性的丧失(图57A)。我们假设dCas13b-$ADAR_{DD}$融合物可以通过指导RNA募集以靶向腺苷，其中杂交的RNA产生ADAR活性所需的双链体底物(图50A)。为了提高靶腺苷脱氨率，我们在最初的RNA编辑设计中引入了另外两个修饰：我们引入了与靶腺苷相对的错配胞苷(先前据报道此举增加了脱氨频率)，并将dCas13b与含有超活化突变以增强催化活性的人类ADAR1或ADAR2的脱氨酶结构域($ADAR1_{DD}$(E1008Q)(25)或$ADAR2_{DD}$(E488Q)(19))融合。

[2875] 为了测试dCas13b-ADAR$_{DD}$的活性，我们通过引入无义突变(W85X(UGG->UAG))在Cluc上产生RNA编辑报告基因，所述突变可通过A->I编辑(图50B)而功能上修复为野生型密码子并且然后检测为Cluc荧光的恢复。我们在靶腺苷上均匀平铺其中间隔区长度为30、50、70或84个核苷酸的指导物，以确定最佳的指导物布置和设计(图50C)。我们发现dCas13b-ADAR1$_{DD}$需要更长的指导序列来修复Cluc报告基因，而dCas13b-ADAR2$_{DD}$在测试的所有指导序列长度下均是功能性的(图50C)。我们还发现，超活性E488Q突变提高了编辑效率，因为具有野生型ADAR2$_{DD}$的荧光素酶的恢复减少(图57B)。根据此活性展示，我们选择dCas13b-ADAR2$_{DD}$(E488Q)用于进一步表征，并且将此方法指定为可编程A至I替换的RNA编辑版本1(REPAIRv1)。

[2876] 为了验证荧光素酶活性的恢复是由于真正的编辑事件引起的，我们直接通过逆转录和靶向的下一代测序测量了经历REPAIRv1的Cluc转录物的编辑。我们测试了在靶位点周围的30nt和50nt的间隔区，并且发现两种指导物长度均导致预期的A至I编辑，而50nt的间隔区实现了更高的编辑百分比(图50D、图50E、图57C)。我们还观察到，50nt间隔区增加了在非靶向腺苷处的编辑倾向，这可能是由于双链体RNA区域的增加(图50E、图57C)。

[2877] 接下来我们靶向内源性基因PPIB。我们设计了平铺PPIB的50nt的间隔区，并且发现我们可以高达28％的编辑效率编辑PPIB转录物(图57D)。为了测试REPAIR是否可以被进一步优化，我们修改了dCas13b与ADAR2$_{DD}$(E488Q)之间的接头(图57E，表8)，并且发现接头选择适度影响了荧光素酶活性的恢复。

[2878] 定义用于RNA编辑的序列参数

[2879] 鉴于我们可以在测试位点实现精确的RNA编辑，我们想要表征针对转录组中任何RNA靶标进行系统编程的序列限制。序列限制可能源自dCas13b靶向限制(诸如PFS)或ADAR序列优选性(26)。为了研究对REPAIRv1的PFS限制，我们设计了质粒文库，该文库在Cluc转录物上的靶位点的5'端带有一系列四个随机化核苷酸(图51A)。我们靶向UAG或AAC基序内的中心腺苷并且发现对于这两个基序，所有PFS均展示出可检测水平的RNA编辑，其中大多数PFS在靶位点具有大于50％的编辑(图51B)。接下来，我们试图确定REPAIRv1中的ADAR2$_{DD}$是否具有紧邻靶碱基侧翼的任何序列限制，如先前已关于ADAR2$_{DD}$报道的(26)。我们测试了直接在靶腺苷周围的5'和3'侧翼核苷酸的每种可能组合(图51C)，并且发现REPAIRv1能够编辑所有基序(图51D)。最后，我们分析了与间隔区序列中的靶A相对的碱基的身份是否影响编辑效率，并且发现A-C错配具有最高的荧光素酶恢复率，其中A-G、A-U和A-A具有大大降低的REPAIRv1活性(图57F)。

[2880] 使用REPAIRv1对疾病相关人类突变的校正

[2881] 为了证明REPAIRv1系统对于哺乳动物细胞中的RNA编辑的广泛适用性，我们针对两种与疾病相关突变设计了REPAIRv1指导物：X连锁性肾原性糖尿病中的878G>A(AVPR2 W293X)和范科尼贫血中的1517G>A(FANCC W506X)。我们将针对携带这些突变的基因的cDNA的表达构建体转染到HEK293FT细胞中，并且测试REPAIRv1是否可以校正突变。使用含有50nt间隔区的指导RNA，我们能够实现AVPR2的35％校正和FANCC的23％校正(图52A-图52D)。然后，我们测试了REPAIRv1校正34种疾病相关G>A突变(表9)的能力，并且发现我们能够在33个位点处实现显著编辑，编辑效率最高达28％(图52E)。我们选择的突变只是ClinVar数据库中病原性G至A突变的一部分(5,739个)，所述数据库还包括另外11,943个G

至A变异(图52F和图58)。因为没有序列限制,所以REPAIRv1能够潜在地编辑所有这些疾病相关突变,尤其是考虑到无论靶基序如何,我们都观察到了显著编辑(图51C和图52G)。

[2882] 将REPAIRv1系统递送至患病细胞为用于治疗性用途的先决条件,并且因此我们试图设计可包装到治疗相关病毒载体诸如腺相关病毒(AAV)载体的REPAIRv1构建体。AAV载体具有4.7kb的包装限度,它不能容纳大尺寸的dCas13b-ADAR$_{DD}$(4473bp)连同启动子和表达调控元件。为了减小尺寸,对于RNA编辑活性我们测试了许多与ADAR2$_{DD}$(E488Q)融合的dCas13的N端和C端截短。我们发现所有测试的C端截短仍有功能并且能够恢复荧光酶信号(图59),并且最大的截短C端Δ984-1090(融合蛋白的总尺寸,4,152bp)足够小以适应AAV载体的包装限度。

[2883] REPAIRv1的全转录物组特异性

[2884] 尽管PspCas13b的RNA敲低具有高特异性,在荧光素酶平铺实验中,在指导物:靶标双链体内观察到脱靶腺苷编辑(图50E)。为了查看这是否是普遍存在的现象,我们评估内源性转录物KRAS并且测量靶腺苷附近的脱靶编辑程度(图53A)。我们发现对于KRAS而言,虽然中靶编辑率为23%,但在靶位点周围存在许多也具有可检测的A至G编辑的位点(图53B)。

[2885] 由于在指导物:靶标双链体内观察到的脱靶编辑,所以我们通过对所有mRNA进行RNA测序来评估所有可能的转录物组脱靶。RNA测序揭示存在大量A至G脱靶事件,其中在靶向条件下具有1,732个脱靶并且在非靶向条件下具有925个脱靶,其中828个脱靶重叠(图53C、图53D)。在整个转录物组的所有编辑位点中,中靶编辑位点具有最高的编辑率,A至G转化率为89%。

[2886] 鉴于Cas13靶向的高特异性,我们有理由认为脱靶可由ADAR引起。先前已描述了两种RNA指导的ADAR系统(图60A)。第一种系统利用ADAR2$_{DD}$与结合至BoxB-λ RNA发夹的小病毒蛋白兰布达N(λN)的融合物(22)。具有双链BoxB-λ发夹的指导物RNA指导ADAR2$_{DD}$编辑指导RNA中编码的位点(23)。第二种系统设计利用全长ADAR2(ADAR2)和具有ADAR2的双链RNA结合结构域(dsRBD)识别的发夹的指导RNA(21,24)。我们分析了这两种系统与REPAIRv1相比的编辑效率并且发现BoxB-ADAR2和ADAR2系统分别证实了63%和36%编辑率,相比之下REPAIRv1实现了89%编辑率(图60B-图60E)。另外,BoxB和ADAR2系统在靶向指导物条件下分别产生2018和174个观察到的脱靶,相比之下在REPAIRv1靶向指导物条件下产生1,229个脱靶。值得注意的是,使用两种基于ADAR2$_{DD}$的系统(REPAIRv1和BoxB)的所有条件均在其脱靶中显示高重叠百分比,而ADAR2系统具有一组显著不同的脱靶(图60F)。靶向条件与非靶向条件之间以及REPAIRv1与BoxB条件之间的脱靶重叠表明,ADAR2$_{DD}$独立于dCas13靶向驱动脱靶(图60F)。

[2887] 通过合理蛋白质工程化改进REPAIRv1的特异性

[2888] 为了改进REPAIR的特异性,我们采用ADAR2$_{DD}$(E488Q)的结构指导的蛋白质工程化。由于脱靶的独立于指导物的性质,我们假设去稳定ADAR2$_{DD}$(E488Q)-RNA结合将选择性地减少脱靶编辑,但由于dCas13b将ADAR2$_{DD}$(E488Q)黏连到靶位点所致的局部浓度增加而保持中靶编辑。我们在ADAR2DD(E488Q)背景下诱变处理先前确定接触靶RNA的双链体区域的ADAR2$_{DD}$(E488Q)残基(图54A)(18)。为了评定效率和特异性,我们在假设非靶向条件下检测到的背景荧光素酶恢复将指示更广泛脱靶活性的情况下使用靶向和非靶向指导物测试了17种单一突变体。我们发现在所选残基处的突变对靶向和非靶向指导物的荧光素酶活性具

有显著影响(图54A、图54B、图61A)。大部分突变体显著改善靶向指导物的荧光素酶活性或者增加靶向指导物与非靶向指导物的活性比率,我们将该比率称为特异性得分(图54A、图54B)。我们选择这些突变体中的一个子集(图54B)用于通过下一代测序进行全转录物组特异性分析。正如预期,由全转录物组测序测量的脱靶与突变体的特异性得分相关(图61B)。我们发现除了ADAR2$_{DD}$(E488Q/R455E),所有测序的REPAIRv1突变体都可以有效地编辑报告转录物(图54C),其中许多突变体显示脱靶数量减少(图61C、图62)。我们进一步探索特异性突变体的脱靶的周围基序,并且发现REPAIRv1和大部分工程化突变体表现出对齐编辑的强3'G优选性,这与表征的ADAR2基序一致(图63A)(26)。我们选择突变体ADAR2$_{DD}$(E488Q/T375G)进行未来的实验,因为它在四种突变体中具有最高的编辑百分比和最低数量的全转录物组脱靶,我们将其称为REPAIRv2。与REPAIRv1相比,REPAIRv2表现出更高的特异性,转录组脱靶从1732个减少至10个(图54D)。在Cluc的靶向腺苷周围的区域中,REPAIRv2具有减少的脱靶编辑,在测序痕迹图中可见(图54E)。在来源于下一代测序的基序中,REPAIRv1呈现针对3'G的强优选性,但对所有基序均显示脱靶编辑(图63B);相比之下,REPAIRv2仅编辑最强脱靶基序(图63C)。编辑在转录物上的分布严重错位,其中高度编辑的基因具有超过60个编辑(图64A、图64B),而REPAIRv2仅多次编辑一个转录物(EEF1A1)(图64D-64F)。据预测REPAIRv1脱靶编辑产生许多变体,包括1000个错义突变(图64C),其中有93个致癌事件(图64D)。相比之下,REPAIRv2仅具有6个错义突变(图64E),这些错义突变不具有致癌结果(图64F)。预测的脱靶效应的此减小将REPAIRv2与其他RNA编辑方法区分开。

[2889] 我们将REPAIRv2靶向内源性基因以测试特异性增强突变是否减少靶转录物中的附近编辑同时保持高效率中靶编辑。对于靶向KRAS或PPIB的指导物,我们发现与REPAIRv1不同,REPAIRv2不具有可检测的脱靶编辑,并且可分别以27.1%和13%有效编辑中靶腺苷(图54F、图54G)。此特异性扩展至另外的靶位点,包括在REPAIRv1的非靶向条件下展示高水平背景的区域,诸如其他KRAS或PPIB靶位点(图65A-图65C)。总之,REPAIRv2消除了编辑的腺苷周围的双链体区域中的脱靶并且显示显著增强的全转录物组特异性。

[2890] 结论

[2891] 我们已在此表明RNA指导的RNA靶向的VI-B型效应子Cas13b能够进行高度有效且特异性的RNA敲低,为用于询问必需基因和非编码RNA以及将细胞过程控制在转录物水平下的改进工具提供基础。无催化活性的Cas13b(dCas13b)保留可编程的RNA结合能力,我们在此通过将dCas13b融合至腺苷脱氨酶ADAR2以实现精确A至I编辑来支持此能力,我们将此系统称为REPAIRv1(可编程A至I置换的RNA编辑版本1)。对系统进行进一步工程化可产生REPAIRv2,所述方法相对于当前的编辑平台具有相当的或增加的活性以及大幅提高的特异性。

[2892] 尽管Cas13b表现出高保真度,但是使用dCas13b-ADAR2$_{DD}$融合物的初始结果显示了数以千计的脱靶。为了解决这个问题,我们采用合理的诱变策略来改变接触RNA双链体的ADAR2$_{DD}$残基,从而鉴定当融合至dCas13b时能够精确、有效且高特异性编辑的变体ADAR2$_{DD}$(E488Q/T375G)。使用此变体的编辑效率相当于或好于使用两种当前可用系统BoxB-ADAR$_{DD}$或ADAR2编辑实现的效率。此外,REPAIRv2系统在整个转录物组中仅产生10个可观察的脱靶,至少比两种替代性编辑技术好一个数量级。

[2893] REPAIR系统与其他核酸编辑工具相比提供了许多优点。首先,可以通过将胞苷置

于所需腺苷对面的指导物延伸部内以形成对于ADAR编辑活性而言理想的有利A-C错配来在指导物中编辑精确靶位点。其次,Cas13不具有靶向序列限制,诸如PFS或PAM,并且也不具有靶腺苷周围的基序优选性,从而允许转录组中的任何腺苷被REPAIR系统潜在靶向。然而,我们注意到,DNA碱基编辑器可以靶向有义链或反义链,而REPAIR系统局限于转录的序列,由此限制我们可以靶向的可能编辑位点的总数目。然而,由于使用REPAIR靶向的更灵活性质,此系统可以影响到的ClinVar内的编辑(图52C)比Cas9-DNA碱基编辑器更多。第三,REPAIR系统直接将靶腺苷脱氨基成肌苷并且不依赖于内源性修复途径(诸如碱基切除或错配修复),以生成所需编辑结果。因此,REPAIR应在不能支持其他形式的编辑的非分裂细胞中是可能的。第四,RNA编辑可以是瞬时的,提供了对编辑结果进行时序控制的可能性。此特性将可能适用于治疗由细胞状态的临时变化造成的疾病,诸如局部炎症。

[2894] REPAIR系统为额外工程化提供了多个机会。Cas13b具有前crRNA加工活性(13),从而允许多重编辑多个变体,这些变体单独地不能改变疾病风险,但一起可具有加成作用和疾病改善可能。合理设计方法的延伸部分(诸如将有前景的突变组合)可以进一步增加该系统的特异性和效率,而公正的筛选方法可以识别用于改进REPAIR活性和特异性的额外残基。

[2895] 目前,通过REPAIR可实现的碱基转化局限于由腺苷生成肌苷;dCas13与其他催化性RNA编辑结构域的额外融合物诸如APOBEC可实现胞苷至尿苷的编辑。另外,ADAR的诱变可放松底物优选性以靶向胞苷,从而允许C->U编辑器探索由双链RNA底物要求赋予的特异性增强。使用无催化活性的DNA靶向CRISPR效应子(诸如dCas9或dCpf1),通过形成DNA-RNA异源双链体靶标(27)或诱变处理ADAR结构域也可以实现DNA底物上腺苷至肌苷的编辑。

[2896] REPAIR可应用于一系列其中A至I(A至G)编辑可逆转或减慢疾病进展的治疗适应症(图66)。首先,可以使用在疾病相关组织中用于靶向有原因的孟德尔遗传学G至A突变的REPAIR的表达来恢复有害突变并治疗疾病。例如,在脑组织中经由AAV进行稳定REPAIR表达可以用于矫正引起局灶性癫痫的GRIN2A错义突变c.2191G>A(Asp731Asn)(28)或引起早发性阿尔兹海默病的APP错义突变c.2149G>A(Val717Ile)(29)。其次,REPAIR可以用于通过改变参与疾病相关信号转导的蛋白质的功能来治疗疾病。例如,REPAIR编辑将允许重新编码激酶的靶标,即一些丝氨酸、苏氨酸和酪氨酸残基(图66)。在疾病相关蛋白中的这些残基的磷酸化影响了许多病症的疾病进展,这些病症包括阿尔兹海默病和多发性神经变性病状(30)。第三,REPAIR可以用于改变表达的改善风险的G至A变体的序列以优先减小患者进入疾病状态的机会。最引人注目的情况是“保护性”风险改善等位基因,其显著减小进入疾病状态的机会并且在一些情况下赋予额外健康益处。例如,REPAIR可以用于在功能上模拟分别预防心血管疾病和银屑病性关节炎的PCSK9和IFIH1的A至G等位基因(31)。最后,REPAIR可以用于在治疗上修饰用于外显子调节疗法的剪接受体和供体位点。REPAIR可以将AU改变为IU或将AA改变为AI,它们分别是共有5’剪接供体或3’剪接受体位点的功能等效物,从而形成新的剪接接合部。另外,REPAIR编辑可以使共有3’剪接受体位点发生AG->IG突变以促进相邻下游外显子的跳读,这是已得到显著关注的用于治疗DMD的治疗性策略。调节剪接位点在反义寡核苷酸具有一些成效的疾病中具有广泛应用,诸如用于调节SMN2剪接以用于治疗脊髓型肌萎缩(32)。

[2897] 我们证实了PspCas13b酶作为RNA敲低和RNA编辑工具的用途。用于可编程RNA结合

的dCas13b平台具有许多应用，包括实况转录物成像、剪接修饰、靶向定位转录物、拉下RNA结合蛋白以及表观转录组修饰。在此，我们使用dCas13形成REPAIR，将其添加到现有成套的核酸编辑技术。REPAIR提供了用于治疗遗传疾病或模拟保护性等位基因的新方法并且建立了作为用于修改遗传功能的有用工具的RNA编辑。

[2898] 表6本研究中所使用的Cas13直系同源物

[2899]

Cas13 ID	Cas13 缩写	寄主生物体	蛋白质登录号
Cas13a1	LshCas13a	沙氏纤毛菌	WP_018451595.1
Cas13a2	LwaCas13a	韦德纤毛菌(Lw2)	WP_021746774.1
Cas13a3	LseCas13a	斯氏李斯特菌	WP_012985477.1
Cas13a4	LbmCas13a	毛螺科菌 MA2020	WP_044921188.1
Cas13a5	LbnCas13a	毛螺科菌 NK4A179	WP_022785443.1
Cas13a6	CamCas13a	嗜氨基[梭菌]DSM 10710	WP_031473346.1
Cas13a7	CgaCas13a	鸡肉杆菌 DSM 4847	WP_034560163.1
Cas13a8	Cga2Cas13a	鸡肉杆菌 DSM 4847	WP_034563842.1
Cas13a9	Pprcas13a	普根沼杆菌 WB4	WP_013443710.1
Cas13a10	LweCas13a	维森李斯特菌 FSL R9-0317	WP_036059185.1
Cas13a11	LbfCas13a	李斯特菌科菌 FSL M6-0635	WP_036091002.1
Cas13a12	Lwa2cas13a	韦德纤毛菌 F0279	WP_021746774.1
Cas13a13	RcsCas13a	荚膜红细菌 SB 1003	WP_013067728.1
Cas13a14	RcrCas13a	荚膜红细菌 R121	WP_023911507.1
Cas13a15	RcdCas13a	荚膜红细菌 DE442	WP_023911507.1
Cas13a16	LbuCas13a	颊纤毛菌 C-1013-b	WP_015770004.1
Cas13a17	HheCas13a	解半纤维素赫氏菌	CRZ35554.1
Cas13a18	EreCas13a	直肠[真杆菌]	WP_055061018.1
Cas13a19	EbaCas13a	真杆菌科菌 CHKCI004	WP_090127496.1
Cas13a20	BmaCas13a	布劳特氏菌属种马赛-P2398	WP_062808098.1
Cas13a21	LspCas13a	纤毛菌属种口腔泰肯菌 879 str.F0557	WP_021744063.1

	Cas13b1	BzoCas13b	动物溃疡伯格菌	WP_002664492
	Cas13b2	PinCas13b	中间普雷沃氏菌	WP_036860899
	Cas13b3	PbuCas13b	颊普雷沃氏菌	WP_004343973
	Cas13b4	AspCas13b	另枝菌属种 ZOR0009	WP_047447901
	Cas13b5	PsmCas13b	普雷沃氏菌属种 MA2016	WP_036929175
	Cas13b6	RanCas13b	鸭疫里默氏杆菌	WP_004919755
	Cas13b7	PauCas13b	桔红色普雷沃氏菌	WP_025000926
	Cas13b8	PsaCas13b	解糖普雷沃氏菌	WP_051522484
	Cas13b9	Pin2Cas13b	中间普雷沃氏菌	WP_061868553
	Cas13b10	CcaCas13b	犬咬二氧化碳嗜纤维菌	WP_013997271
	Cas13b11	PguCas13b	咽喉卟啉单胞菌	WP_039434803
	Cas13b12	PspCas13b	普雷沃氏菌属种 P5-125	WP_044065294
	Cas13b13	FbrCas13b	嗜鳃黄杆菌	WP_014084666
	Cas13b14	PgiCas13b	牙龈卟啉单胞菌	WP_053444417
	Cas13b15	Pin3Cas13b	中间普雷沃氏菌	WP_050955369
[2900]	Cas13c1	FnsCas13c	坏死梭杆菌肠形亚种 ATCC 51357 contig00003	WP_005959231.1
	Cas13c2	FndCas13c	坏死梭杆菌 DJ-2 contig0065，全基因组鸟枪序列	WP_035906563.1
	Cas13c3	FnbCas13c	坏死梭杆菌 BFTR-1 contig0068	WP_035935671.1
	Cas13c4	FnfCas13c	坏死梭杆菌肠形亚种 1_1_36S cont1.14	EHO19081.1
	Cas13c5	FpeCas13c	极臭梭杆菌 ATCC 29250 T364DRAFT_scaffold00009.9_C	WP_027128616.1
	Cas13c6	FulCas13c	溃疡梭杆菌 ATCC 49185 cont2.38	WP_040490876.1
	Cas13c7	AspCas13c	厌氧盐杆菌属种 ND1 基因组组件马赛厌氧盐杆菌 ND1	WP_042678931.1

[2901] 表7细菌筛选中的PFS截断

[2902]

Cas13b 直系同源物	关键字	-用于生成 PFS 基序的 Log_2 耗尽得分
动物溃疡伯格菌	1	2
中间普雷沃氏菌基因座 1	2	1
颊普雷沃氏菌	3	3
另枝菌属种 ZOR0009	4	1
普雷沃氏菌属种 MA2016	5	2
鸭疫里默氏杆菌	6	4
桔红色普雷沃氏菌	7	1
解糖普雷沃氏菌	8	0

[2903]

中间普雷沃氏菌基因座 2	9	0
犬咬二氧化碳嗜纤维菌	10	3
咽喉卟啉单胞菌	11	4
普雷沃氏菌属种 P5-125	12	2.1
嗜鳃黄杆菌	13	1
牙龈卟啉单胞菌	14	3
中间普雷沃氏菌基因座 2	15	4

[2904] 表8在本研究中用于在哺乳动物细胞中进行RNA编辑的dCas13b-ADAR接头序列.

[2905]

图	接头
50C	GSGGGGS (SEQ ID NO: 774)
50E	GS
57B	GSGGGGS
57C	GS
57D	GS
57E: GS	GS
57E: GSGGGGS	GSGGGGS
57E: (GGGS)3	GGGGSGGGGSGGGGS(SEQ ID NO: 1)
57E: Rigid	EAAAK(SEQ ID NO: 297)
57E: (GGS)6	GGSGGSGGSGGSGGSGGS(SEQ ID NO: 298))
57E: XTEN	SGSETPGTSESATPES(SEQ ID NO: 299)
51B	GS
57F	GS
51C	GS
52B	GS
52D	GS
52E	GS
51A : Δ984-1090 、Δ1026-1090、Δ1053-1090	GS
51A: Δ1-125、Δ1-88、Δ1-72	GSGGGGS (SEQ ID NO: 774)
53B	GS
53C	GS
53D	GS
60A	GS
60C	GS
60D	GS
61D	GS
54A	GS
62A	GS
54B	GS
62B	GS
62C	GS
63A	GS
63B	GS
54C	GS
54D	GS

[2906]

54E	GS
54F	GS
66A	GS
66A	GS

[2907] 表9疾病相关突变的疾病信息

[2908]

全长候选物	基因	疾病
NM_000054.4(AVPR2):c.878G>A (p.Trp293Ter)	AVPR2	X 连锁性肾原性尿崩症
NM_000136.2(FANCC):c.1517G>A (p.Trp506Ter)	FANCC	范可尼贫血，互补群 C
额外模拟候选物		
候选物	**基因**	**疾病**
NM_000206.2(IL2RG):c.710G>A (p.Trp237Ter)	IL2RG	X 连锁性重度联合免疫缺陷
NM_000132.3(F8):c.3144G>A (p.Trp1048Ter)	F8	遗传性因子 VIII 缺乏症
NM_000527.4(LDLR):c.1449G>A (p.Trp483Ter)	LDLR	家族性高胆固醇血症
NM_000071.2(CBS):c.162G>A (p.Trp54Ter)	CBS	由 CBS 缺乏症引起的高胱氨酸尿症
NM_000518.4(HBB):c.114G>A (p.Trp38Ter)	HBB	β^0^地中海贫血/β 地中海贫血
NM_000035.3(ALDOB):c.888G>A (p.Trp296Ter)	ALDOB	遗传性果糖尿病
NM_004006.2(DMD):c.3747G>A (p.Trp1249Ter)	DMD	杜兴氏肌营养不良
NM_005359.5(SMAD4):c.906G>A (p.Trp302Ter)	SMAD4	青少年息肉病综合征
NM_000059.3(BRCA2):c.582G>A (p.Trp194Ter)	BRCA2	家族性乳腺癌|家族性乳腺癌-卵巢癌 2
NM_000833.4(GRIN2A):c.3813G>A (p.Trp1271Ter)	GRIN2A	局灶性癫痫，伴有言语障碍，且伴有或不伴有智力迟钝
NM_002977.3(SCN9A):c.2691G>A (p.Trp897Ter)	SCN9A	常染色体隐性遗传性先天性疼痛不敏感
NM_007375.3(TARDBP):c.943G>A (p.Ala315Thr)	TARDBP	肌萎缩性侧索硬化症 10 型
NM_000492.3(CFTR):c.3846G>A (p.Trp1282Ter)	CFTR	囊性纤维化|遗传性胰腺炎|未提供|ataluren 反应-功效;
NM_130838.1(UBE3A):c.2304G>A (p.Trp768Ter)	UBE3A	安格曼综合征 (Angelman syndrome)
NM_000543.4(SMPD1):c.168G>A (p.Trp56Ter)	SMPD1	尼曼-皮克病 (Niemann-Pick disease)，A 型
NM_206933.2(USH2A):c.9390G>A	USH2A	乌谢尔综合征，2A 型

[2909]

(p.Trp3130Ter)		
NM_130799.2(MEN1):c.1269G>A (p.Trp423Ter)	MEN1	遗传性癌症易感综合征
NM_177965.3(C8orf37):c.555G>A (p.Trp185Ter)	C8orf37	色素性视网膜炎 64
NM_000249.3(MLH1):c.1998G>A (p.Trp666Ter)	MLH1	林奇综合征
NM_000548.4(TSC2):c.2108G>A (p.Trp703Ter)	TSC2	结节性硬化症 2|结节性硬化症综合征
NM_000267.3(NF1):c.7044G>A (p.Trp2348Ter)	NF1	神经纤维瘤病，1 型
NM_000179.2(MSH6):c.3020G>A (p.Trp1007Ter)	MSH6	林奇综合征
NM_000344.3(SMN1):c.305G>A (p.Trp102Ter)	SMN1	脊髓型肌萎缩，II 型|Kugelberg-Welander 病
NM_024577.3(SH3TC2):c.920G>A (p.Trp307Ter)	SH3TC2	夏科-马里-图思病，4C 型
NM_001369.2(DNAH5):c.8465G>A (p.Trp2822Ter)	DNAH5	原发性纤毛运动障碍症
NM_004992.3(MECP2):c.311G>A (p.Trp104Ter)	MECP2	雷特综合征
NM_032119.3(ADGRV1):c.7406G>A (p.Trp2469Ter)	ADGRV1	乌谢尔综合征，2C 型
NM_017651.4(AHI1):c.2174G>A (p.Trp725Ter)	AHI1	乔伯特综合征 3
NM_004562.2(PRKN):c.1358G>A (p.Trp453Ter)	PRKN	帕金森病 2
NM_000090.3(COL3A1):c.3833G>A (p.Trp1278Ter)	COL3A1	Ehlers-Danlos 综合征，4 型
NM_007294.3(BRCA1):c.5511G>A (p.Trp1837Ter)	BRCA1	家族性乳腺癌|家族性乳腺癌-卵巢癌 1
NM_000256.3(MYBPC3):c.3293G>A (p.Trp1098Ter)	MYBPC3	原发性家族性肥厚型心肌病
NM_000038.5(APC):c.1262G>A (p.Trp421Ter)	APC	家族性腺瘤性息肉病 1
NM_001204.6(BMPR2):c.893G>A (p.W298*)	BMPR2	原发性肺动脉高压

[2910] 表10：本研究中所使用的关键性质粒

[2911]

质粒名称	说明
pAB0006	CMV-荧光素酶-polyA EF1a-G-荧光素酶-polyA
pAB0040	CMV-荧光素酶(STOP85)-polyA EF1a-G-荧光素酶-polyA
pAB0048	pCDNA-ADAR2-N 端 B12-HIV NES
pAB0050	pAB0001-A02 (crRNA 主链)
pAB0051	pAB0001-B06 (crRNA 主链)

[2912]

pAB0052	pAB0001-B11 (crRNA 主链)
pAB0053	pAB0001-B12 (crRNA 主链)
pAB0014.B6	EF1a-BsiWI-Cas13b6-NES-mapk
pAB0013.B11	EF1a-BsiWI-Cas13b11-NES-HIV
pAB0013.B12	EF1a-BsiWI-Cas13b12-NES-HIV
pAB0051	pAB0001-B06 (crRNA 主链)
pAB0052	pAB0001-B11 (crRNA 主链)
pAB0053	pAB0001-B12 (crRNA 主链)
pAB0079	pCDNA-ADAR1hu-EQ 突变体-N 端目的载体
pAB0085	pCDNA-ADAR2 (E488Q)hu-EQ 突变体-N 端目的载体
pAB0095	EF1a-BsiWI-Cas13-B12-NES-HIV，具有双 H HEPN 突变体
pAB0114	pCDNA-wtADAR2hu-EQ 突变体-N 端目的载体
pAB0120	最佳荧光素酶 ADAR 指导物(来自 wC0054 的指导物 24)
pAB0122	用于 ADAR 实验的 pAB0001-B12 NT 指导物
pAB0151	pCDNA-ADAR2hu-EQ 突变体-N 端目的载体 C 端 Δ984-1090
pAB0180	T375G 特异性突变体
pAB0181	T375G Cas13b C 端 Δ984-1090

[2913] 表11:在本研究中用于在哺乳动物细胞中敲低的指导物/shRNA序列

[2914]

名称	间隔区序列	干扰机制	注意	第一图
细菌 PFS 指导物	GCCAGCUUUCCGGGCAUUGGCUUCCAUC(SEQ ID NO: 300)	Cas13b	用于所有直系同源物	
Cas13a-Gluc 指导物 1	GCCAGCTTTCCGGGCATTGGCTTCCATC (SEQ ID NO: 301)	Cas13a	用于所有 Cas13a 直系同源物	图 49B
Cas13a-Gluc 指导物 2	ACCCAGGAATCTCAGGAATGTCGACGAT (SEQ ID NO: 302)	Cas13a	用于所有 Cas13a 直系同源物	图 49B
Cas13a-非靶向指导物(LacZ)	AGGGTTTTCCCAGTCACGACGTTGTAAA (SEQ ID NO: 303)	Cas13a	用于所有 Cas13a 直系同源物	图 49B
Cas13b-Gluc 指导物 1.1	GGGCATTGGCTTCCATCTCTTTGAGCACCT (SEQ ID NO: 304)	Cas13b	用于直系同源物 1-3、6、7、10、11、12、14、15	图 49B
Cas13b-Gluc 指导物 1.2	GUGCAGCCAGCUUUCCGGGCAUUGGCUUCC (SEQ ID NO: 305)	Cas13b	用于直系同源物 4	图 49B
Cas13b-Gluc 指导物 1.3	GCAGCCAGCUUUCCGGGCAUUGGCUUCCAU (SEQ ID NO: 306)	Cas13b	用于直系同源物 5	图 49B
Cas13b-Gluc	GGCUUCCAUCUCUU	Cas13b	用于直系同源物	图 49B

[2915]

指导物 1.4	UGAGCACCUCCAGCGG (SEQ ID NO: 307)		8、9	
Cas13b-Gluc 指导物 1.5	GGAAUGUCGACGAUCGCCUCGCCUAUGCCG (SEQ ID NO: 308)	Cas13b	用于直系同源物 13	图 49B
Cas13b-Gluc 指导物 2.1	GAAUGUCGACGAUCGCCUCGCCUAUGCCGC (SEQ ID NO: 309)	Cas13b	用于直系同源物 1-3、6、7、10、11、14、15	图 49B
Cas13b-Gluc 指导物 2.2	GACCUGUGCGAUGAACUGCUCCAUGGGCUC (SEQ ID NO: 310)	Cas13b	用于直系同源物 12	图 49B
Cas13b-Gluc 指导物 2.2	GUGUGGCAGCGUCCUGGGAUGAACUUCUUC (SEQ ID NO: 311)	Cas13b	用于直系同源物 4	图 49B
Cas13b-Gluc 指导物 2.3	GUGGCAGCGUCCUGGGAUGAACUUCUUCAU (SEQ ID NO: 312)	Cas13b	用于直系同源物 5	图 49B
Cas13b-Gluc 指导物 2.4	GCUUCUUGCCGGGCAACUUCCCGCGGUCAG (SEQ ID NO: 313)	Cas13b	用于直系同源物 8、9	图 49B
Cas13b-Gluc 指导物 2.6	GCAGGGUUUUCCCAGUCACGACGUUGUAAAA (SEQ ID NO: 314)	Cas13b	用于直系同源物 13	图 49B
Cas13b- 非靶向指导物	GCAGGGUUUUCCCAGUCACGACGUUGUAAAA (SEQ ID NO: 315)	Cas13b	用于所有直系同源物	图 49B
Cas13a-Gluc 指导物 -RNASeq	ACCCAGGAAUCUCAGGAAUGUCGACGAU (SEQ ID NO: 316)	Cas13a		图 49E
shRNA-Gluc 指导物	CAGCTTTCCGGGCATTGGCTT (SEQ ID NO: 317)	shRNA		图 49F
Cas13b-Gluc 指导物 -RNASeq	CCGCUGGAGGUGCUCAAAGAGAUGGAAGCC (SEQ ID NO: 318)	Cas13b		图 49F
Cas13a-Gluc-指导物-1	GCCAGCTTTCCGGGCATTGGCTTCCATC (SEQ ID NO: 319)	Cas13a		图 56A
Cas13a-Gluc-	ACCCAGGAATCTCAG	Cas13a		图 56A

	指导物-2	GAATGTCGACGAT (SEQ ID NO: 320)			
	Cas13b-Gluc-最佳指导物-1	GGGCATTGGCTTCCATCTCTTTGAGCACCT (SEQ ID NO: 321)	Cas13b		图 56A
	Cas13b-Gluc-最佳指导物-2	GAAUGUCGACGAUCGCCUCGCCUAUGCCGC (SEQ ID NO: 322)	Cas13b		图 56A
	Cas13a KRAS 指导物 1	CAAGGCACTCTTGCCTACGCCACCAGCT (SEQ ID NO: 323)	Cas13a		图 56B
	Cas13a KRAS 指导物 2	TCATATTCGTCCACAAAATGATTCTGAA (SEQ ID NO: 324)	Cas13a		图 56B
	Cas13a KRAS 指导物 3	ATTATTTATGGCAAATACACAAAGAAAG (SEQ ID NO: 325)	Cas13a		图 56B
	Cas13a KRAS 指导物 4	GAATATCTTCAAATGATTTAGTATTATT (SEQ ID NO: 326)	Cas13a		图 56B
	Cas13a KRAS 指导物 5	ACCATAGGTACATCTTCAGAGTCCTTAA (SEQ ID NO: 327)	Cas13a		图 56B
[2916]	Cas13b KRAS 指导物 1	GTCAAGGCACTCTTGCCTACGCCACCAGCT (SEQ ID NO: 328)	Cas13b		图 56B
	Cas13b KRAS 指导物 2	GATCATATTCGTCCACAAAATGATTCTGAA (SEQ ID NO: 329)	Cas13b		图 56B
	Cas13b KRAS 指导物 3	GTATTATTTATGGCAAATACACAAAGAAAG (SEQ ID NO: 330)	Cas13b		图 56B
	Cas13b KRAS 指导物 4	GTGAATATCTTCAAATGATTTAGTATTATT (SEQ ID NO: 331)	Cas13b		图 56B
	Cas13b KRAS 指导物 5	GGACCATAGGTACATCTTCAGAGTCCTTAA (SEQ ID NO: 332)	Cas13b		图 56B
	shRNA KRAS 指导物 1	aagagtgccttgacgatacagcCTCGAGgctgtatcgtcaaggcactctt (SEQ ID NO: 333)	shRNA		图 56B
	shRNA KRAS 指导物 2	aatcattttgtggacgaatatCTCGAGatattcgtccacaaaatgatt (SEQ ID NO: 334)	shRNA		图 56B
	shRNA KRAS	aaataatactaaatcatttgaCTC	shRNA		图 56B

[2917]

指导物 3	GAGtcaaatgatttagtattattt (SEQ ID NO: 335)			
shRNA KRAS 指导物 4	aataatactaaatcatttgaaCTC GAGttcaaatgatttagtattatt (SEQ ID NO: 336)	shRNA		图 56B
shRNA KRAS 指导物 5	aaggactctgaagatgtacctCT CGAGaggtacatcttcagagtc ctt (SEQ ID NO: 337)	shRNA		图 56B

[2918] 表12:用于Gluc敲低的指导序列

名称	间隔区序列	位置	注意	第一图
Gluc 平铺指导物 1	GAGATCAGGGCAAACAGAA CTTTGACTCCC (SEQ ID NO: 338)	2	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 2	GGATGCAGATCAGGGCAAA CAGAACTTTGA (SEQ ID NO: 339)	7	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 3	GCACAGCGATGCAGATCAG GGCAAACAGAA (SEQ ID NO: 340)	13	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 4	GCTCGGCCACAGCGATGCAG ATCAGGGCAA (SEQ ID NO: 341)	19	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 5	GGGGCTTGGCCTCGGCCACA GCGATGCAGA (SEQ ID NO: 342)	28	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 6	GTGGGCTTGGCCTCGGCCAC AGCGATGCAG (SEQ ID NO: 343)	29	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 7	GTCTCGGTGGGCTTGGCCTC GGCCACAGCG (SEQ ID NO: 344)	35	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 8	GTTCGTTGTTCTCGGTGGGC TTGGCCTCGG (SEQ ID NO: 345)	43	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 9	GGAAGTCTTCGTTGTTCTCG GTGGGCTTGG (SEQ ID NO: 346)	49	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 10	GATGTTGAAGTCTTCGTTGT TCTCGGTGGG (SEQ ID NO: 347)	54	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 11	GCGGCCACGATGTTGAAGTC TTCGTTGTTC (SEQ ID NO: 348)	62	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GTGGCCACGGCCACGATGTT GAAGTCTTCG	68	应注意 Cas13a 间隔	49C

[2919]

[2920]

导物 12	(SEQ ID NO: 349)		区在 5'端处被截短两个核苷酸	
Gluc 平铺指导物 13	GGTTGCTGGCCACGGCCACGATGTTGAAGT (SEQ ID NO: 350)	73	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 14	GTCGCGAAGTTGCTGGCCACGGCCACGATG (SEQ ID NO: 351)	80	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 15	GCCGTGGTCGCGAAGTTGCTGGCCACGGCC (SEQ ID NO: 352)	86	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 16	GCGAGATCCGTGGTCGCGAAGTTGCTGGCC (SEQ ID NO: 353)	92	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 17	GCAGCATCGAGATCCGTGGTCGCGAAGTTG (SEQ ID NO: 354)	98	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 18	GGGTCAGCATCGAGATCCGTGGTCGCGAAG (SEQ ID NO: 355)	101	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 19	GCTTCCCGCGGTCAGCATCGAGATCCGTGG (SEQ ID NO: 356)	109	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 20	GGGGCAACTTCCCGCGGTCAGCATCGAGAT (SEQ ID NO: 357)	115	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 21	GTCTTGCCGGGCAACTTCCCGCGGTCAGCA (SEQ ID NO: 358)	122	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 22	GGCAGCTTCTTGCCGGGCAACTTCCCGCGG (SEQ ID NO: 359)	128	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 23	GCCAGCGGCAGCTTCTTGCCGGGCAACTTC (SEQ ID NO: 360)	134	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 24	GCACCTCCAGCGGCAGCTTCTTGCCGGGCA (SEQ ID NO: 361)	139	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 25	GCTTTGAGCACCTCCAGCGGCAGCTTCTTG (SEQ ID NO: 362)	146	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 26	GCATCTCTTTGAGCACCTCCAGCGGCAGCT (SEQ ID NO: 363)	151	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GTCCATCTCTTTGAGCACCT	153	应注意 Cas13a 间隔	49C

[2921]

导物 27	CCAGCGGCAG (SEQ ID NO: 364)		区在 5'端处被截短两个核苷酸	
Gluc 平铺指导物 28	GGGCATTGGCTTCCATCTCTTTGAGCACCT (SEQ ID NO: 365)	163	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 29	GTCCGGGCATTGGCTTCCATCTCTTTGAGC (SEQ ID NO: 366)	167	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 30	GGCCAGCTTTCCGGGCATTGGCTTCCATCT (SEQ ID NO: 367)	175	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 31	GGGTGCAGCCAGCTTTCCGGGCATTGGCTT (SEQ ID NO: 368)	181	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 32	GAGCCCCTGGTGCAGCCAGCTTTCCGGGCA (SEQ ID NO: 369)	188	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 33	GATCAGACAGCCCCTGGTGCAGCCAGCTTT (SEQ ID NO: 370)	195	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 34	GGCAGATCAGACAGCCCCTGGTGCAGCCAG (SEQ ID NO: 371)	199	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 35	GACAGGCAGATCAGACAGCCCCTGGTGCAG (SEQ ID NO: 372)	203	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 36	GTGATGTGGGACAGGCAGATCAGACAGCCC (SEQ ID NO: 373)	212	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 37	GACTTGATGTGGGACAGGCAGATCAGACAG (SEQ ID NO: 374)	215	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 38	GGGGCGTGCACTTGATGTGGGACAGGCAGA (SEQ ID NO: 375)	223	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 39	GCTTCATCTTGGGCGTGCACTTGATGTGGG (SEQ ID NO: 376)	232	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 40	GTGAACTTCTTCATCTTGGGCGTGCACTTG (SEQ ID NO: 377)	239	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 41	GGGATGAACTTCTTCATCTTGGGCGTGCAC (SEQ ID NO: 378)	242	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GTGGGATGAACTTCTTCATC	244	应注意 Cas13a 间隔	49C

[2922]

导物 42	TTGGGCGTGC (SEQ ID NO: 379)		区在 5'端处被截短两个核苷酸	
Gluc 平铺指导物 43	GGGCAGCGTCCTGGGATGAACTTCTTCATC (SEQ ID NO: 380)	254	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 44	GGGTGTGGCAGCGTCCTGGGATGAACTTCT (SEQ ID NO: 381)	259	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 45	GTTCGTAGGTGTGGCAGCGTCCTGGGATGA (SEQ ID NO: 382)	265	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 46	GCGCCTTCGTAGGTGTGGCAGCGTCCTGGG (SEQ ID NO: 383)	269	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 47	GTCTTTGTCGCCTTCGTAGGTGTGGCAGCG (SEQ ID NO: 384)	276	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 48	GCTTTGTCGCCTTCGTAGGTGTGGCAGCGT (SEQ ID NO: 385)	275	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 49	GTGCCGCCCTGTGCGGACTCTTTGTCGCCT (SEQ ID NO: 386)	293	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 50	GTATGCCGCCCTGTGCGGACTCTTTGTCGC (SEQ ID NO: 387)	295	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 51	GCCTCGCCTATGCCGCCCTGTGCGGACTCT (SEQ ID NO: 388)	302	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 52	GGATCGCCTCGCCTATGCCGCCCTGTGCGG (SEQ ID NO: 389)	307	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 53	GATGTCGACGATCGCCTCGCCTATGCCGCC (SEQ ID NO: 390)	315	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 54	GCAGGAATGTCGACGATCGCCTCGCCTATG (SEQ ID NO: 391)	320	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 55	GAATCTCAGGAATGTCGACGATCGCCTCGC (SEQ ID NO: 392)	325	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 56	GCCCAGGAATCTCAGGAATGTCGACGATCG (SEQ ID NO: 393)	331	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GCCTTGAACCCAGGAATCTC	338	应注意 Cas13a 间隔	49C

[2923]

导物 57	AGGAATGTCG (SEQ ID NO: 394)		区在 5'端处被截短两个核苷酸	
Gluc 平铺指导物 58	GCCAAGTCCTTGAACCCAGGAATCTCAGGA (SEQ ID NO: 395)	344	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 59	GTGGGCTCCAAGTCCTTGAACCCAGGAATC (SEQ ID NO: 396)	350	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 60	GCCATGGGCTCCAAGTCCTTGAACCCAGGA (SEQ ID NO: 397)	353	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 61	GGAACTGCTCCATGGGCTCCAAGTCCTTGA (SEQ ID NO: 398)	361	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 62	GTGCGATGAACTGCTCCATGGGCTCCAAGT (SEQ ID NO: 399)	367	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 63	GGACCTGTGCGATGAACTGCTCCATGGGCT (SEQ ID NO: 400)	373	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 64	GACAGATCGACCTGTGCGATGAACTGCTCC (SEQ ID NO: 401)	380	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 65	GACACACAGATCGACCTGTGCGATGAACTG (SEQ ID NO: 402)	384	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 66	GTGCAGTCCACACACAGATCGACCTGTGCG (SEQ ID NO: 403)	392	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 67	GCCAGTTGTGCAGTCCACACACAGATCGAC (SEQ ID NO: 404)	399	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 68	GGGCAGCCAGTTGTGCAGTCCACACACAGA (SEQ ID NO: 405)	404	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 69	GTTTGAGGCAGCCAGTTGTGCAGTCCACAC (SEQ ID NO: 406)	409	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 70	GAAGCCCTTTGAGGCAGCCAGTTGTGCAGT (SEQ ID NO: 407)	415	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 71	GCACGTTGGCAAGCCCTTTGAGGCAGCCAG (SEQ ID NO: 408)	424	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GACTGCACGTTGGCAAGCCC	428	应注意 Cas13a 间隔	49C

[2924]

导物 72	TTTGAGGCAG (SEQ ID NO: 409)		区在 5'端处被截短两个核苷酸	
Gluc 平铺指导物 73	GGGTCAGAACACTGCACGTTGGCAAGCCCT (SEQ ID NO: 410)	437	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 74	GCAGGTCAGAACACTGCACGTTGGCAAGCC (SEQ ID NO: 411)	439	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 75	GAGCAGGTCAGAACACTGCACGTTGGCAAG (SEQ ID NO: 412)	441	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 76	GGCCACTTCTTGAGCAGGTCAGAACACTGC (SEQ ID NO: 413)	452	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 77	GCGGCAGCCACTTCTTGAGCAGGTCAGAAC (SEQ ID NO: 414)	457	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 78	GTGCGGCAGCCACTTCTTGAGCAGGTCAGA (SEQ ID NO: 415)	459	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 79	GAGCGTTGCGGCAGCCACTTCTTGAGCAGG (SEQ ID NO: 416)	464	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 80	GAAAGGTCGCACAGCGTTGCGGCAGCCACT (SEQ ID NO: 417)	475	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 81	GCTGGCAAAGGTCGCACAGCGTTGCGGCAG (SEQ ID NO: 418)	480	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 82	GGGCAAAGGTCGCACAGCGTTGCGGCAGCC (SEQ ID NO: 419)	478	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 83	GTGGATCTTGCTGGCAAAGGTCGCACAGCG (SEQ ID NO: 420)	489	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 84	GCACCTGGCCCTGGATCTTGCTGGCAAAGG (SEQ ID NO: 421)	499	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 85	GTGGCCCTGGATCTTGCTGGCAAAGGTCGC (SEQ ID NO: 422)	495	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指导物 86	GTGATCTTGTCCACCTGGCCCTGGATCTTG (SEQ ID NO: 423)	509	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
Gluc 平铺指	GCCCCTTGATCTTGTCCACC	514	应注意 Cas13a 间隔	49C

	导物 87	TGGCCCTGGA (SEQ ID NO: 424)		区在 5'端处被截短两个核苷酸	
	Gluc 平铺指导物 88	GCCCTTGATCTTGTCCACCTGGCCCTGGAT (SEQ ID NO: 425)	513	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	Gluc 平铺指导物 89	GCCTTGATCTTGTCCACCTGGCCCTGGATC (SEQ ID NO: 426)	512	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	Gluc 平铺指导物 90	GGCAAAGGTCGCACAGCGTTGCGGCAGCCA (SEQ ID NO: 427)	477	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	Gluc 平铺指导物 91	GCAAAGGTCGCACAGCGTTGCGGCAGCCAC (SEQ ID NO: 428)	476	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
[2925]	Gluc 平铺指导物 92	GAAGGTCGCACAGCGTTGCGGCAGCCACTT (SEQ ID NO: 429)	474	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	Gluc 平铺指导物 93	GAGGTCGCACAGCGTTGCGGCAGCCACTTC (SEQ ID NO: 430)	473	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	非靶向指导物 1	GGTAATGCCTGGCTTGTCGACGCATAGTCTG (SEQ ID NO: 431)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	非靶向指导物 2	GGGAACCTTGGCCGTTATAAAGTCTGACCAG (SEQ ID NO: 432)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C
	非靶向指导物 3	GGAGGGTGAGAATTTAGAACCAAGATTGTTG (SEQ ID NO: 433)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49C

[2926] 表13:用于Cluc敲低的指导序列

名称	间隔区序列	位置	注意	第一图
Cluc 平铺指导物 1	GAGTCCTGGCAATGAACAGTGGCGCAGTAG (SEQ ID NO: 434)	32	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 2	GGGTGCCACAGCTGCTATCAATACATTCTC (SEQ ID NO: 435)	118	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
[2927] Cluc 平铺指导物 3	GTTACATACTGACACATTCGGCAACATGTT (SEQ ID NO: 436)	197	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 4	GTATGTACCAGGTTCCTGGAACTGGAATCT (SEQ ID NO: 437)	276	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 5	GCCTTGGTTCCATCCAGGTTCTCCAGGGTG	350	应注意 Cas13a 间隔区在 5'端处被截短两个	49D

	(SEQ ID NO: 438)		核苷酸	
Cluc 平铺指导物 6	GCAGTGATGGGATTCTCAG TAGCTTGAGCG (SEQ ID NO: 439)	431	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 7	GAGCCTGGCATCTCAACAA CAGCGATGGTG (SEQ ID NO: 440)	512	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 8	GTGTCTGGGGCGATTCTTA CAGATCTTCCT (SEQ ID NO: 441)	593	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 9	GCTGGATCTGAAGTGAAGT CTGTATCTTCC (SEQ ID NO: 442)	671	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 10	GGCAACGTCATCAGGATTT CCATAGAGTGG (SEQ ID NO: 443)	747	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 11	GAGGCGCAGGAGATGGTG TAGTAGTAGAAG (SEQ ID NO: 444)	830	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 13	GAGGGACCCTGGAATTGGT ATCTTGCTTTG (SEQ ID NO: 445)	986	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
[2928] Cluc 平铺指导物 14	GGTAAGAGTCAACATTCCT GTGTGAAACCT (SEQ ID NO: 446)	1066	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 15	GACCAGAATCTGTTTTCCA TCAACAATGAG (SEQ ID NO: 447)	1143	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 16	GATGGCTGTAGTCAGTATG TCACCATCTTG (SEQ ID NO: 448)	1227	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 17	GTACCATCGAATGGATCTC TAATATGTACG (SEQ ID NO: 449)	1304	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 18	GAGATCACAGGCTCCTTCA GCATCAAAAGA (SEQ ID NO: 450)	1380	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 19	GCTTTGACCGGCGAAGAGA CTATTGCAGAG (SEQ ID NO: 451)	1461	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 20	GCCCCTCAGGCAATACTCG TACATGCATCG (SEQ ID NO: 452)	1539	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
Cluc 平铺指导物 21	GCTGGTACTTCTAGGGTGT CTCCATGCTTT	1619	应注意 Cas13a 间隔区在 5'端处被截短两个	49D

[2929]

	(SEQ ID NO: 453)		核苷酸	
非靶向指导物 1	GGTAATGCCTGGCTTGTCG ACGCATAGTCTG (SEQ ID NO: 454)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
非靶向指导物 2	GGGAACCTTGGCCGTTATA AAGTCTGACCAG (SEQ ID NO: 455)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D
非靶向指导物 3	GGAGGGTGAGAATTTAGA ACCAAGATTGTTG (SEQ ID NO: 456)	不适用	应注意 Cas13a 间隔区在 5'端处被截短两个核苷酸	49D

[2930] 表14:在本研究中用于在哺乳动物细胞中进行RNA编辑的指导序列.将错配碱基翻转大写

[2931]

名称	间隔区序列	注意	第一图
平铺 30 nt 30 错配距离	gCatcctgcggcctctactctgcattcaat t (SEQ ID NO: 457)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 28 错配距离	gacCatcctgcggcctctactctgcattc aa (SEQ ID NO: 458)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 26 错配距离	gaaacCatcctgcggcctctactctgcat tc (SEQ ID NO: 459)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 24 错配距离	gctaaacCatcctgcggcctctactctgc at (SEQ ID NO: 460)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 22 错配距离	gttctaaacCatcctgcggcctctactctg c (SEQ ID NO: 461)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 20 错配距离	gtgttctaaacCatcctgcggcctctactc t (SEQ ID NO: 462)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 18 错配距离	gaatgttctaaacCatcctgcggcctcta ct (SEQ ID NO: 463)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 16 错配距离	gagaatgttctaaacCatcctgcggcctc ta (SEQ ID NO: 464)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 14 错配距离	gatagaatgttctaaacCatcctgcggcc tc (SEQ ID NO: 465)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 12 错配距离	gccatagaatgttctaaacCatcctgcgg cc (SEQ ID NO: 466)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 10 错配距离	gttccatagaatgttctaaacCatcctgcg g (SEQ ID NO: 467)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 8 错配距离	gctttccatagaatgttctaaacCatcctg c	具有用于 U6 表达的 5' G	50C

[2932]

	(SEQ ID NO: 468)		
平铺 30 nt 6 错配距离	gctctttccatagaatgttctaaacCatcct (SEQ ID NO: 469)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 4 错配距离	gatctctttccatagaatgttctaaacCatc (SEQ ID NO: 470)	具有用于 U6 表达的 5' G	50C
平铺 30 nt 2 错配距离	ggaatctctttccatagaatgttctaaacC a (SEQ ID NO: 471)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 50 错配距离	gCatcctgcggcctctactctgcattcaat tacatactgacacattcggca (SEQ ID NO: 472)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 48 错配距离	gacCatcctgcggcctctactctgcattc aattacatactgacacattcgg (SEQ ID NO: 473)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 46 错配距离	gaaacCatcctgcggcctctactctgcat tcaattacatactgacacattc (SEQ ID NO: 474)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 44 错配距离	gctaaacCatcctgcggcctctactctgc attcaattacatactgacacat (SEQ ID NO: 475)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 42 错配距离	gttctaaacCatcctgcggcctctactctg cattcaattacatactgacac (SEQ ID NO: 476)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 40 错配距离	gtgttctaaacCatcctgcggcctctactc tgcattcaattacatactgac (SEQ ID NO: 477)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 38 错配距离	gaatgttctaaacCatcctgcggcctcta ctctgcattcaattacatactg (SEQ ID NO: 478)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 36 错配距离	gagaatgttctaaacCatcctgcggcctc tactctgcattcaattacatac (SEQ ID NO: 479)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 34 错配距离	gatagaatgttctaaacCatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 480)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 32 错配距离	gccatagaatgttctaaacCatcctgcgg cctctactctgcattcaattac (SEQ ID NO: 481)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 30 错配距离	gttccatagaatgttctaaacCatcctgcg gcctctactctgcattcaatt (SEQ ID NO: 482)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 28 错配距离	gctttccatagaatgttctaaacCatcctg cggcctctactctgcattcaa (SEQ ID NO: 483)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 26 错配距离	gctctttccatagaatgttctaaacCatcct gcggcctctactctgcattc (SEQ ID NO: 484)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 24 错	gatctctttccatagaatgttctaaacCatc ctgcggcctctactctgcat	具有用于 U6 表达	50C

[2933]

配距离	(SEQ ID NO: 485)	的 5' G	
平铺 50 nt 22 错配距离	ggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgc (SEQ ID NO: 486)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 20 错配距离	gtggaatctctttccatagaatgttctaaacCatcctgcggcctctactct (SEQ ID NO: 487)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 18 错配距离	gactggaatctctttccatagaatgttctaaacCatcctgcggcctctact (SEQ ID NO: 488)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 16 错配距离	ggaactggaatctctttccatagaatgttctaaacCatcctgcggcctcta (SEQ ID NO: 489)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 14 错配距离	gtggaactggaatctctttccatagaatgttctaaacCatcctgcggcctc (SEQ ID NO: 490)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 12 错配距离	gcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcc (SEQ ID NO: 491)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 10 错配距离	gttcctggaactggaatctctttccatagaatgttctaaacCatcctgcgg (SEQ ID NO: 492)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 8 错配距离	gggttcctggaactggaatctctttccatagaatgttctaaacCatcctgc (SEQ ID NO: 493)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 6 错配距离	gcaggttcctggaactggaatctctttccatagaatgttctaaacCatcct (SEQ ID NO: 494)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 4 错配距离	gaccaggttcctggaactggaatctctttccatagaatgttctaaacCatc (SEQ ID NO: 495)	具有用于 U6 表达的 5' G	50C
平铺 50 nt 2 错配距离	ggtaccaggttcctggaactggaatctctttccatagaatgttctaaacCa (SEQ ID NO: 496)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 70 错配距离	gCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttat (SEQ ID NO: 497)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 68 错配距离	gacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggttt (SEQ ID NO: 498)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 66 错配距离	gaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggt (SEQ ID NO: 499)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 64 错配距离	gctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctg	具有用于 U6 表达的 5' G	50C

[2934]

	(SEQ ID NO: 500)		
平铺 70 nt 62 错配距离	gttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcc (SEQ ID NO: 501)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 60 错配距离	gtgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttt (SEQ ID NO: 502)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 58 错配距离	gaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttt (SEQ ID NO: 503)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 56 错配距离	gagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgt (SEQ ID NO: 504)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 54 错配距离	gatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacat (SEQ ID NO: 505)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 52 错配距离	gccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaac (SEQ ID NO: 506)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 50 错配距离	gttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggca (SEQ ID NO: 507)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 48 错配距离	gctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcgg (SEQ ID NO: 508)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 46 错配距离	gctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattc (SEQ ID NO: 509)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 44 错配距离	gatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacat (SEQ ID NO: 510)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 42 错配距离	ggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacac (SEQ ID NO: 511)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 40 错配距离	gtggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgac (SEQ ID NO: 512)	具有用于 U6 表达的 5' G	50C

	平铺 70 nt 38 错配距离	gactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactg (SEQ ID NO: 513)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 36 错配距离	ggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatac (SEQ ID NO: 514)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 34 错配距离	gtggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacat (SEQ ID NO: 515)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 32 错配距离	gcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattac (SEQ ID NO: 516)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 30 错配距离	gttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaatt (SEQ ID NO: 517)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 28 错配距离	gggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaa (SEQ ID NO: 518)	具有用于 U6 表达的 5' G	50C
[2935]	平铺 70 nt 26 错配距离	gcaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattc (SEQ ID NO: 519)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 24 错配距离	gaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcat (SEQ ID NO: 520)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 22 错配距离	ggtaccaggttcctggaactggaatctcttccatagaatgttctaaacCatcctgcggcctctactctgc (SEQ ID NO: 521)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 20 错配距离	gatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactct (SEQ ID NO: 522)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 18 错配距离	ggtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctact (SEQ ID NO: 523)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 16 错配距离	gacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctcta (SEQ ID NO: 524)	具有用于 U6 表达的 5' G	50C
	平铺 70 nt 14 错	gacacgtatgtaccaggttcctggaactg	具有用于 U6 表达	50C

[2936]

配距离	gaatctctttccatagaatgttctaaacCatcctgcggcctc (SEQ ID NO: 525)	的 5' G	
平铺 70 nt 12 错配距离	gcaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcc (SEQ ID NO: 526)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 10 错配距离	gcccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcgg (SEQ ID NO: 527)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 8 错配距离	ggacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgc (SEQ ID NO: 528)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 6 错配距离	gttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcct (SEQ ID NO: 529)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 4 错配距离	gccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatc (SEQ ID NO: 530)	具有用于 U6 表达的 5' G	50C
平铺 70 nt 2 错配距离	gttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCa (SEQ ID NO: 531)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 84 错配距离	gCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttattttcacacagtcca (SEQ ID NO: 532)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 82 错配距离	gacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttattttcacacagtc (SEQ ID NO: 533)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 80 错配距离	gaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttattttcacacag (SEQ ID NO: 534)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 78 错配距离	gctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttattttcacac (SEQ ID NO: 535)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 76 错配距离	gttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttcctggtttattttcac (SEQ ID NO: 536)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 74 错配距离	gtgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggc	具有用于 U6 表达的 5' G	50C

[2937]

	aacatgtttttcctggtttattttc (SEQ ID NO: 537)		
平铺 84 nt 72 错配距离	gaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggtttattt (SEQ ID NO: 538)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 70 错配距离	gagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggtttat (SEQ ID NO: 539)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 68 错配距离	gatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggttt (SEQ ID NO: 540)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 66 错配距离	gccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctggt (SEQ ID NO: 541)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 64 错配距离	gttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcctg (SEQ ID NO: 542)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 62 错配距离	gctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgtttttcc (SEQ ID NO: 543)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 60 错配距离	gctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttttt (SEQ ID NO: 544)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 58 错配距离	gatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgttt (SEQ ID NO: 545)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 56 错配距离	ggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacatgt (SEQ ID NO: 546)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 54 错配距离	gtggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaacat (SEQ ID NO: 547)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 52 错配距离	gactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggcaac (SEQ ID NO: 548)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 50 错配距离	ggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcggca	具有用于 U6 表达的 5' G	50C

[2938]

	(SEQ ID NO: 549)		
平铺 84 nt 48 错配距离	gtggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattcgg (SEQ ID NO: 550)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 46 错配距离	gcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacattc (SEQ ID NO: 551)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 44 错配距离	gttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacacat (SEQ ID NO: 552)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 42 错配距离	gggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgacac (SEQ ID NO: 553)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 40 错配距离	gcaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactgac (SEQ ID NO: 554)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 38 错配距离	gaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatactg (SEQ ID NO: 555)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 36 错配距离	ggtaccaggttcctggaactggaatctcttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacatac (SEQ ID NO: 556)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 34 错配距离	gatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattacat (SEQ ID NO: 557)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 32 错配距离	ggtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaattac (SEQ ID NO: 558)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 30 错配距离	gacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaatt (SEQ ID NO: 559)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 28 错配距离	gacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattcaa (SEQ ID NO: 560)	具有用于 U6 表达的 5' G	50C
平铺 84 nt 26 错配距离	gcaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcattc (SEQ ID NO: 561)	具有用于 U6 表达的 5' G	50C

	平铺 84 nt 24 错配距离	gcccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgcat (SEQ ID NO: 562)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 22 错配距离	ggacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactctgc (SEQ ID NO: 563)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 20 错配距离	gttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctactct (SEQ ID NO: 564)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 18 错配距离	gccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctctact (SEQ ID NO: 565)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 16 错配距离	gttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctcta (SEQ ID NO: 566)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 14 错配距离	gggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcctc (SEQ ID NO: 567)	具有用于 U6 表达的 5' G	50C
[2939]	平铺 84 nt 12 错配距离	gttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcggcc (SEQ ID NO: 568)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 10 错配距离	gccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgcgg (SEQ ID NO: 569)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 8 错配距离	ggcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcctgc (SEQ ID NO: 570)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 6 错配距离	gccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatcct (SEQ ID NO: 571)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 4 错配距离	gcgccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCatc (SEQ ID NO: 572)	具有用于 U6 表达的 5' G	50C
	平铺 84 nt 2 错配距离	ggtcgccgcccttggttccttgacccaacacgtatgtaccaggttcctggaactggaatctctttccatagaatgttctaaacCa (SEQ ID NO: 573)	具有用于 U6 表达的 5' G	50C
	ADAR 非靶向指	GTAATGCCTGGCTTGTCG	具有用于 U6 表达	50C

[2940]

导物	ACGCATAGTCTG (SEQ ID NO: 574)	的 5' G	
用于 TAG 基序的 PFS 结合筛选指导物	gaaaacgcaggttcctcCagtttcggga gcagcgcacgtctccctgtagtc (SEQ ID NO: 575)	具有用于 U6 表达的 5' G	51B
用于 AAC 基序的 PFS 结合筛选指导物	gacgcaggttcctctagCttcgggagca gcgcacgtctccctgtagtcaag (SEQ ID NO: 576)	具有用于 U6 表达的 5' G	51B
PFS 结合筛选非靶向指导物	GTAATGCCTGGCTTGTCG ACGCATAGTCTG (SEQ ID NO: 577)	具有用于 U6 表达的 5' G	51B
基序优选性靶向指导物	gatagaatgttctaaacCatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 578)	具有用于 U6 表达的 5' G	51C
基序优选性非靶向指导物	GTAATGCCTGGCTTGTCG ACGCATAGTCTG (SEQ ID NO: 579)	具有用于 U6 表达的 5' G	51C
PPIB 平铺指导物 50 错配距离	gCaaggccacaaaattatccactgtttttg gaacagtctttccgaagagac (SEQ ID NO: 580)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 42 错配距离	gcctgtagcCaaggccacaaaattatcc actgtttttggaacagtctttcc (SEQ ID NO: 581)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 34 错配距离	gctttctctcctgtagcCaaggccacaaa attatccactgtttttggaaca (SEQ ID NO: 582)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 26 错配距离	ggccaaatcctttctctcctgtagcCaag gccacaaaattatccactgttt (SEQ ID NO: 583)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 18 错配距离	gtttttgtagccaaatcctttctctcctgtag cCaaggccacaaaattatc (SEQ ID NO: 584)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 10 错配距离	gatttgctgtttttgtagccaaatcctttctct cctgtagcCaaggccaca (SEQ ID NO: 585)	具有用于 U6 表达的 5' G	57D
PPIB 平铺指导物 2 错配距离	gacgatggaatttgctgtttttgtagccaaa tcctttctctcctgtagcCa (SEQ ID NO: 586)	具有用于 U6 表达的 5' G	57D
靶向指导物，相对碱基 G	gatagaatgttctaaacGatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 587)	具有用于 U6 表达的 5' G	57D
靶向指导物，相对碱基 A	gatagaatgttctaaacAatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 588)	具有用于 U6 表达的 5' G	57D
靶向指导物，相对碱基 C	gatagaatgttctaaacTatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 589)	具有用于 U6 表达的 5' G	57D
AVPR2 指导物 37	ggtcccacgcggccCacagctgcacca ggaagaagggtgcccagcacagca	具有用于 U6 表达	52A

[2941]

错配距离	(SEQ ID NO: 590)	的 5' G	
AVPR2 指导物 35 错配距离	ggggtcccacgcggccCacagctgcac caggaagaagggtgcccagcacag (SEQ ID NO: 591)	具有用于 U6 表达的 5' G	52A
AVPR2 指导物 33 错配距离	gccgggtcccacgcggccCacagctgc accaggaagaagggtgcccagcac (SEQ ID NO: 592)	具有用于 U6 表达的 5' G	52A
FANCC 指导物 37 错配距离	gggtgatgacatccCaggcgatcgtgtg gcctccaggagcccagagcagga (SEQ ID NO: 593)	具有用于 U6 表达的 5' G	52B
FANCC 指导物 35 错配距离	gagggtgatgacatccCaggcgatcgtg tggcctccaggagcccagagcag (SEQ ID NO: 594)	具有用于 U6 表达的 5' G	52B
FANCC 指导物 32 错配距离	gatcagggtgatgacatccCaggcgatc gtgtggcctccaggagcccagag (SEQ ID NO: 595)	具有用于 U6 表达的 5' G	52B
合成疾病基因靶标 IL2RG	ggtggctccattcactcCaatgctgagca cttccacagagtgggttaaagc (SEQ ID NO: 596)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 F8	gtttctaatatattttgCcagactgatggac tattctcaattaataatgat (SEQ ID NO: 597)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 LDLR	gagatgttgctgtggatCcagtccacagc cagcccgtcgggggcctggatg (SEQ ID NO: 598)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 CBS	gcaggccggcccagctgCcaggtgcac ctgctcggagcatcgggccggatc (SEQ ID NO: 599)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 HBB	gcaaagaacctctgggtCcaagggtaga ccaccagcagcctgcccagggcc (SEQ ID NO: 600)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 ALDOB	gaagagaaacttagtttCcagggctttgg tagagggcaaaggttgatagca (SEQ ID NO: 601)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 DMD	gtcagcctagtgcagagCcactggtagtt ggtggttagagtttcaagttcc (SEQ ID NO: 602)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 SMAD4	ggctcattgtgaacaggCcagtaatgtcc gggatggggcggcataggcggg (SEQ ID NO: 603)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 BRCA2	gtagctaaagaacttgaCcaagacatatc aggatccacctcagctcctaga (SEQ ID NO: 604)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 GRIN2A	ggggcattgttctgtgcCcagtcctgctg gtagacctgctccccggtggct (SEQ ID NO: 605)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 SCN9A	gagaagtcgttcatgtgCcaccgtggga gcgtacagtcatcattgatcttg (SEQ ID NO: 606)	具有用于 U6 表达的 5' G	52E

	合成疾病基因靶标 TARDBP	gggattaatgctgaacgCaccaaagttcatcccaccacccatattactacc (SEQ ID NO: 607)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 CFTR	gctccaaaggctttcctCcactgttgcaaagttattgaatcccaagacaca (SEQ ID NO: 608)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 UBE3A	gatgaatgaacgatttcCcagaactccctaatcagaacagagtccctggta (SEQ ID NO: 609)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 SMPD1	ggagcctctgccggagcCcagagaaccccgagagtcagacagagccagcgcc (SEQ ID NO: 610)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 USH2A	ggcttccgtggagacacCcaatcaatttgaagagatcttgaagtgatgcca (SEQ ID NO: 611)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 MEN1	gtgggactgccctcctcCcatttgcagatgccgtcgtagaatcgcagcagg (SEQ ID NO: 612)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 C8orf37	gcttcttcaatagttctCcagctacactggcaggcatatgcccgtgttcct (SEQ ID NO: 613)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 MLH1	gattccttttcttcgtcCcaattcacctcagtggctagtcgaagaatgaag (SEQ ID NO: 614)	具有用于U6表达的5' G	52E
[2942]	合成疾病基因靶标 TSC2	gcagcttcagcaccttcCagtcagactcctgcttcaagcactgcagcagga (SEQ ID NO: 615)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 NF1	gccatttgcttgcagtgCcactccagaggattccggattgccataaatact (SEQ ID NO: 616)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 MSH6	gttcaatagttttggtcCagtatcgtttacagcccttcttggtagatttca (SEQ ID NO: 617)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 SMN1	ggcaaccgtcttctgacCaaatggcagaacatttgtccccaactttccact (SEQ ID NO: 618)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 SH3TC2	gcgactttccaatgaacCactgaagcccaggtatgacaaagccgatgatct (SEQ ID NO: 619)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 DNAH5	gtttacactcatgcttcCacagctttaacagatcatttggttccttgatga (SEQ ID NO: 620)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 MECP2	gcttaagcttccgtgtcCagccttcaggcagggtggggtcatcatacatgg (SEQ ID NO: 621)	具有用于U6表达的5' G	52E
	合成疾病基因靶标 ADGRV1	ggacagctgggctgatcCatgatgtcatccagaaacactggggaccctcag (SEQ ID NO: 622)	具有用于U6表达的5' G	52E
	合成疾病基因靶	gtctcatctcaactttcCatatccgtatcat	具有用于U6表达	52E

标 AHI1	ggaatcatagcatcctgtaa (SEQ ID NO: 623)	的 5' G	
合成疾病基因靶标 PRKN	gcatgcagacgcggttcCactcgcagcc acagttccagcaccactcgagcc (SEQ ID NO: 624)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 COL3A1	gttggttagggtcaaccCagtattctccac tcttgagttcaggatggcaga (SEQ ID NO: 625)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 BRCA1	gctacactgtccaacacCcactctcgggt caccacaggtgcctcacacatc (SEQ ID NO: 626)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 MYBPC3	gctgcactgtgtaccccCagagctccgt gttgccgacatcctggggtggct (SEQ ID NO: 627)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 APC	gagcttcctgccactccCaacaggtttca cagtaagcgcgtatctgttcca (SEQ ID NO: 628)	具有用于 U6 表达的 5' G	52E
合成疾病基因靶标 BMPR2	gacggcaagagcttaccCagtcacttgt gtggagacttaaatacttgcata (SEQ ID NO: 629)	具有用于 U6 表达的 5' G	52E
KRAS 平铺指导物 50 错配距离	gCaaggccacaaaattatccactgtttttg gaacagtctttccgaagagac (SEQ ID NO: 630)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 42 错配距离	gcctgtagcCaaggccacaaaattatcc actgtttttggaacagtctttcc (SEQ ID NO: 631)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 34 错配距离	gctttctctcctgtagcCaaggccacaaa attatccactgtttttggaaca (SEQ ID NO: 632)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 26 错配距离	ggccaaatcctttctctcctgtagcCaag gccacaaaattatccactgttt (SEQ ID NO: 633)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 18 错配距离	gtttttgtagccaaatcctttctctcctgtag cCaaggccacaaaattatc (SEQ ID NO: 634)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 10 错配距离	gatttgctgtttttgtagccaaatcctttctct cctgtagcCaaggccaca (SEQ ID NO: 635)	具有用于 U6 表达的 5' G	53A
KRAS 平铺指导物 2 错配距离	gacgatggaatttgctgtttttgtagccaaa tcctttctctcctgtagcCa (SEQ ID NO: 636)	具有用于 U6 表达的 5' G	53A
KRAS 平铺非靶向指导物	GTAATGCCTGGCTTGTCG ACGCATAGTCTG (SEQ ID NO: 637)	具有用于 U6 表达的 5' G	53A
用于转录组特异性的荧光素酶 W85X 靶向指导物	gatagaatgttctaaacCatcctgcggcc tctactctgcattcaattacat (SEQ ID NO: 638)	具有用于 U6 表达的 5' G	53B
用于转录组特异	GCAGGGTTTTCCCAGTCA	具有用于 U6 表达	53C

[2943]

性的非靶向指导物	CGACGTTGTAAAGTTG (SEQ ID NO: 639)	的 5' G	
内源性 KRAS 指导物 2	gtcaaggcactcttgccCacgccaccag ctccaactaccacaagtttatat (SEQ ID NO: 640)	具有用于 U6 表达的 5' G	54F
内源性 PPIB 指导物 1	gcaaagatcacccggccCacatcttcatc tccaattcgtaggtcaaaatac (SEQ ID NO: 641)	具有用于 U6 表达的 5' G	54G
内源性 KRAS 指导物 1	GcgccaccagctccaacCaccacaagt ttatattcagtcattttcagcagg (SEQ ID NO: 642)	具有用于 U6 表达的 5' G	54F
内源性 KRAS 指导物 3	GtttctccatcaattacCacttgcttcctgt aggaatcctctattGTtgga (SEQ ID NO: 643)	具有用于 U6 表达的 5' G	54F
内源性 PPIB 指导物 2	GctttctctcctgtagcCaaggccacaaa attatccactgtttttggaaca (SEQ ID NO: 644)	具有用于 U6 表达的 5' G	54G
内源性非靶向指导物	GTAATGCCTGGCTTGTCG ACGCATAGTCTG (SEQ ID NO: 645)	具有用于 U6 表达的 5' G	54F
BoxB Cluc 指导物	tctttccataGGCCCTGAAAAA GGGCCtgttctaaacCatcctgcggc ctctactcGGCCCTGAAAAAG GGCCattcaattac (SEQ ID NO: 646)	具有用于 U6 表达的 5' G	62B
BoxB 非靶向指导物	cagctggcgaGGCCCTGAAAA AGGGCCggggatgtgcCgcaaggc gattaagttggGGCCCTGAAAA AGGGCCacgccagggt (SEQ ID NO: 647)	具有用于 U6 表达的 5' G	62B
Stafforst 全长 ADAR2 指导物 1	GTGGAATAGTATAACAAT ATGCTAAATGTTGTTATA GTATCCCACtctaaaCCAtcctg cgGGGCCCTCTTCAGGGCC C (SEQ ID NO: 648)	具有用于 U6 表达的 5' G	62C
Stafforst 全长 ADAR2 非靶向指导物	GTGGAATAGTATAACAAT ATGCTAAATGTTGTTATA GTATCCCACaccctggcgttaccc aGGGCCCTCTTCAGGGCCC (SEQ ID NO: 649)	具有用于 U6 表达的 5' G	62C

[2944]

[2945] 参考文献

[2946] 1.P.D.Hsu,E.S.Lander,F.Zhang,Devlopment and applications of CRISPR-Cas9 for genome engineering.Cell 157,1262-1278(2014).

[2947] 2.A.C.Komor,A.H.Badran,D.R.Liu,CRISPR-Based Tcchnologies for the Manipulation of Eukaryotic Genomes,Cell 168,20-36(2017).

[2948] 3.L.Cong et al.,Multiplex genome engineering using CRISPR/Cas sy stems.Science 339,819-823(2013).

[2949] 4.P.Mali et al.,RNA-guided human genome engineering via Cas9.Science

339,823-826(2013).

[2950] 5.B.Zetsche et al.,Cpfl is a single RNA-guided endonuclease of a class 2CRISPR-Cas system.Cell 163,759-771(2015).

[2951] 6.H.Kim,J.S.Kim,A guide to genome engineering with programmable nucleases.Nat Rev Genet 15,321-334(2014).

[2952] 7.A.C.Komor,Y.B.Kim,M.S.Packer,J.A.Zuris,D.R.Liu,Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage.Nature 533,420-424(2016).

[2953] 8.K.Nishida et al.,Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems.Science 353,(2016).

[2954] 9.Y.B.Kim et al.,Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions.NatBiotechnol 35,371-376(2017).

[2955] 10.O.O.Abudayyeh et al.,C2c2 is a single-component programmable RNA-guidedRNA-targeting CRISPR effector.Science 353,aaf5573(2016).

[2956] 11.S.Shmakov et al.,Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems.Mol Cell 60,385-397(2015).

[2957] 12.S.Shmakov et al.,Diversity and evolution of class 2 CRISPR-Cas systems.Nat Rev Microbiol 15,169-182(2017).

[2958] 13.A.A.Smargon et al.,Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28.Mol Cell 65,618-630 e6 17(2017).

[2959] 14.J.S.Gootenberg et al.,Nucleic acid detection with CRISPR-Cas13a/C2c2.Science 356,438-442(2017).

[2960] 15.O.O.Abuddyyeh et al.,RNA targeting with CRISPR-Cas 13a.Nature in press,(2017).

[2961] 16.K.Nishikura,Functions and regulation of RNA editing by ADAR deaminases.Annu ReV Biochem 79,321-349(2010).

[2962] 17.B.L_Bass,HL Weintraub,An unwinding activity that covalently modifies its double-stranded RNA sub strate.Cell 55,1089-1098(1988).

[2963] 18.M.M.Matthews et al.,Structures of human ADAR2bound to dsRNA reveal base-flipping mechanism and basis for site selectivity.Nat Struct Mol Biol 23,426-433(2016).

[2964] 19.A.Kuttan,B.L.Bass,Mechanistic insights into editing-site specificity of ADARs.Proc Natl Acad Sci U S A 109,E3295-3304(2012).

[2965] 20.S.K.Wong,S.Sato.D.W.Lazinski,Substrate recognition by ADAR1 and ADAR2.RNA 7,846-858(2001).

[2966] 21.M.Fukuda et al.,Construction of a guide-RNA for site-directed RNA mutagemesisutilising intracellular A-to-I RNA editing.Sci Rep 7,41478(2017).

[2967] 22.M.F.Montiel-Gonzalez,I.Vallecillo-Viejo,G.AYudowski,J.J.Rosenthal,Correction of mutations within the cystic fibrosis transmembrane conductace regulator hy site-directed RNA editing.Proc Natl Acad Sci U S A110,18285-18290(2013).

[2968] 23.M.F.Montiel-Gonzalez,I.C.Vallecillo-Viejo.J.J.Rosenthal,An efficientsy sy stemfor selectively altering genetic information with in mRNAs.Nucleuc Acids Res 44,e157(2016).

[2969] 24.J.Wettegel,P.Reautschnig,S.Geisler,P.J.Kahle,T.Stafforst,Harnessing humanADAR2 for RNA repair-Recoding a PINkI mutatiob resces mitophagy.Nucleic Acids Res 45,2797-2808(2017).

[2970] 25.Y.Wang,J.Havel,P.A.Beal,A Phenotypic Screen for Functional Mutants of Human Adenosine Deaminase Acting on RNA 1.ACS Chem Biol 10,2512-2519(2015).

[2971] 26.K.A.Lehmann,B.L.Bass,Double-stranded RNA adenosine deaminases ADARIand ADAR2 have overlapping specificitie.Biochemistry 39,12875-12884(2000).

[2972] 27.Y.Zhenng,C.Lorenzo,P.A.Beal,DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA.Nucleic Acids Res 45,3369-3377(2017).

[2973] 28.K.Gao et al.,A de novo loss-of-function GRIN2A mutation associated with childhood focal epilepsy and acquired epileptic aphasia PLoS One 12,e0170818(2017).

[2974] 29.H.M.Lanoiselee et al.,App,PSEN1,and PSEN2mutations in early-onset Alzheimer disease:A genetic screening study of familial and sporadic cases.pLoS Med 14,e1002270(2017).

[2975] 30.C.Ballatore,V.M.Lee,J.Q.Trojanowski,Tau-mediated neurodegeneration in Alzheimer's disease and related disorders.Nat Rev Neurosci 8,663-672(2007).

[2976] 31.Y.Li et al.,Carriers of rare missense variants in IFIHI are protected from psoriasis.J Invest Dermatol 130,2768-2772(2010).

[2977] 32.R.S.Finkel et al.,Treatment of infantile-onset spinal muscular atrophy with nusinersen:a phase 2,open-label,dose-escalation study.Lancet 388,3017-3026(2016).

[2978] 实施例4-

[2979] 产生了用于RNA编辑的dCas13b的截短。

[2980] 自N端和C端两端产生20个氨基酸间隔的截短,并在A至I RNA编辑的情形下加以测试。选择dCas13b6(Δ795-1095)-GS-HIVNES-GS-huADAR2dd作为用于包装到AAV中的最终构建体。

[2981] 测试了Cas13b直系同源物,包括Cas13b6(RanCas13b)、Cas13b11(PguCas13b)和Cas13b12(PspCas13b)。对于每种直系同源物,我们从Cas13b直系同源物的N端(红色)和C端

(蓝色)末端逐一截下20个氨基酸。接着我们将该截短的蛋白融合至ADAR2dd(E488Q),并利用海肾荧光素酶作为表达对照,评估该融合蛋白对海萤荧光素酶中的提前终止密码子的RNA编辑活性。我们发现从Cas13b6(RanCas13b)的C端截下300个氨基酸可保留约2/3的RNA编辑活性,同时将蛋白质缩短到可接受地接近AAV的包装限度。AAV大体上能够包装约4.7kb的基因组,并且对于较大的基因组,包装效率会降低。在Cas13b6、Cas13b11和Cas13b12上的测试结果分别显示在图67-图69中。

[2982] 使用AAV将dCas13b6(Δ795-1095)-REPAIR构建体递送至原代神经元培养物中。使用原代大鼠皮层神经元培养物来测试对MAP2转录物的编辑。对于递送,使用AAV1/AAV2杂合衣壳。测量错配距离,即从正向重复序列到指导RNA间隔区序列中的错配碱基的距离。

[2983] 如图70中的示意图所示,我们使用从C端去除了300个氨基酸的Cas13b6与ADAR2dd(E488Q)融合,以便对原代大鼠皮层神经元的MAP2转录物进行沉默突变。MAP2是神经元特异性标记物,此外,REPAIR蛋白在人类突触蛋白启动子下表达,以增加系统的神经元特异性表达。我们使用AAV1/AAV2杂合衣壳将其递送至原代大鼠皮层神经元培养物中,并且针对靶位点处指定的指导RNA设计实现了高达20%的编辑(图70)。

[2984] 我们还使用该系统调节钾离子通道Kcna1中的位点(已经是天然的ADAR2底物)处的编辑,方式是使用AAV1/2杂合衣壳将图71上详述的系统递送至原代大鼠皮层神经元培养物中。我们发现,尽管我们能够通过将REPAIR系统传递至这些细胞来增加编辑的转录物的百分比,但不一定需要指导RNA来引导这种活性。这可能是因为该位点已经是天然的ADAR2底物,因此无需使用指导RNA诱导此位点处的编辑。结果显示,过表达含有ADAR2催化结构域的构建体可能足以在这些天然底物处开始饱和编辑。

[2985] 图72示出了含有dCas13b6(Δ795-1095)-GS-HIVNES-GS-huADAR2dd的质粒的示意图。质粒的序列显示在下表中。

[2986]

图 72 中所示的质粒的序列

[2987]

gacggatcgggagatctcccgatcccctatggtgcactctcagtacaatctgctctgatgccgcatagttaagc
cagtatctgctccctgcttgtgtgttggaggtcgctgagtagtgcgcgagcaaaatttaagctacaacaaggca
aggcttgaccgacaattgcatgaagaatctgcttagggttaggcgttttgcgctgcttcgcgatgtacgggcca
gatatacgcgttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatat
ggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccattgacg
tcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggt
aaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaat
ggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcat
cgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttcc
aagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaa
caactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctgg
ctaactagagaacccactgcttactggcttatcgaaattaatacgactcactatagggagacccaagctggcta
gcgtttaaacttaagcttgccaccATGgagaagcccctgctgcctaacgtgtataccctgaagcacaagttct
tttggggcgccttcctgaatatcgcccggcacaacgcctttatcaccatctgccacatcaatgagcagctgggc
ctgaaaaccccagcaacgacgataagatcgtggacgtggtgtgcgagacatggaacaatatcctgaacaat
gaccacgatctgctgaagaagtcccagctgaccgagctgatcctgaagcacttccctttctgacagccatgtg
ctaccacccccctaagaaggagggcaagaagaagggccaccagaaggagcagcagaaggagaaggaga
gcgaggcacagtcccaggcagaggccctgaacccatccaagctgatcgaggccctggagatcctggtgaat
cagctgcactctctgGCAaactactatagcGCCtataagcacaagaagcccgacgccgagaaggatatct
tcaagcacctgtataaggccttcgacgcctctctgagaatggtgaaggaggattataaggcccacttcaccgtg
aatctgacaagggactttgcccacctgaaccgcaagggcaagaacaagcaggacaatcccgacttcaacaga
tacagattcgagaaggatggcttctttaccgagtctggcctgctgttctttacaaatctgtttctggacaagaggg
atgcctattggatgctgaagaaggtgtccggcttcaaggcctctcacaagcagcgcgagaagatgaccacag
aggtgttttgcaggagccgcatcctgctgcccaagctgaggctggagtcccgctacgaccacaaccagatgct
gctggatatgctgtctgagctgagcagatgtcctaagctgctgtatgagaagctgagcgaggagaataagaag
cacttccaggtggaggccgacggctttctggatgagatcgaggaggagcagaacccattcaaggacaccctg
atccggcaccaggatagattcccctactttgccctgaggtatctggacctgaatgagtccttcaagtctatccgct
ttcaggtggatctgggcacataccactattgtatctacgacaagaagatcggcgatgagcaggagaagaggca
cctgacccgcacactgctgtccttcggccggctgcaggactttaccgagatcaacagaccccaggagtggaa
ggccctgaccaaggacctggattacaaggagacatccaatcagcctttcatctctaagaccacaccacactatc
acatcaccgacaacaagatcggctttcggctgggcacaagcaaggagctgtacccctccctggagatcaagg
atggcgccaatagaatcgccaagtacccttataactccggcttcgtggcccacgcctttatctctgtgcacgagc
tgctgcctctgatgttctaccagcacctgaccggcaagagcgaggacctgctgaaggagacagtgcggcaca
tccagagaatctataaggacttcgaggaggagcggatcaataccatcgaggatctggagaaggcaaaccagg
gcagactgccactgggagcctttcccaagcagatgctgggcctgctgcagaataagcagcctgatctgtccga
gaaggccaagatcaagatcgagaagctgatcgccgagacaaagctgctgtctcacaggctgaacacaaagct
gaagagctccccaaagctgggcaagcggagagagaagctgatcaagacaggcgtgctggccgactggctg
gtgaaggatttcatgcgctttcagcccgtggcctacgacgcccagaatcagcctatcaagtctagcaaggcca
actccaccgagttctggtttatcaggcgcgccctggccctgtacggaggagagaagaataggctggagggct
atttcaagcagaccaacctgatcggcaacacaaatccacacccttcctgaacaagtttaattggaaggcctgc
aggaatctggtggatttctaccagcagtatctggagcagcgcgagaagtttctggaggccatcaagaaccagc
catgggagccctaccagtattgcctgctgctgaagatccctaaggagaacaggaagaatctggtgaagggatg
ggagcagggaggaatctctctgccacggggcctgtttaccgaggccatcagagagacactgtctgaggacct
gatgctgagcaagccaatccgcaaggagatcaagaagcacggccgggtgggcttcatcagcagagccatca
ccctgtactttaaggagaagtatcaggataagcaccagagcttctacaatctgtcctataagctggaggcaaag
gcaccactgggatcTCTTCAACTGCCTCCACTTGAAAGACTGACACTGGGA
TCCcagctgcatttaccgcaggttttagctgacgctgtctcacgcctggtcctgggtaagtttggtgacctgac
cgacaacttctcctcccctcacgctcgcagaaaagtgctggctggagtcgtcatgacaacaggcacagatgtt
aaagatgccaaggtgataagtgtttctacaggaacaaaatgtattaatggtgaatacatgagtgatcgtggcctt
gcattaaatgactgccatgcagaaataatatctcggagatccttgctcagatttctttatacacaacttgagcttta

[2988]

cttaaataacaaagatgatcaaaaaagatccatctttcagaaatcagagcgaggggggtttaggctgaaggag
aatgtccagtttcatctgtacatcagcacctctccctgtggagatgccagaatcttctcaccacatgagccaatcc
tggaagaaccagcagatagacacccaaatcgtaaagcaagaggacagctacggaccaaaatagagtctggt
Caggggacgattccagtgcgctccaatgcgagcatccaaacgtgggacggggtgctgcaaggggagcggc
tgctcaccatgtcctgcagtgacaagattgcacgctggaacgtggtgggcatccagggatcActgctcagcat
tttcgtggagcccatttacttctcgagcatcatcctgggcagcctttaccacggggaccacctttccagggccat
gtaccagcggatctccaacatagaggacctgccacctctctacaccctcaacaagcctttgctcagtggcatca
gcaatgcagaagcacggcagccagggaaggcccccaacttcagtgtcaactggacggtaggcgactccgct
attgaggtcatcaacgccacgactgggaaggatgagctgggccgcgcgtcccgcctgtgtaagcacgcgttg
tactgtcgctggatgcgtgtgcacggcaaggttccctcccacttactacgctccaagattaccaagcccaacgt
gtaccatgagtccaagctggcggcaaaggagtaccaggccgccaaggcgcgtctgttcacagccttcatcaa
ggcggggctgggggcctgggtggagaagcccaccgagcaggaccagttctcactcacgTAAgcggccg
ctcgagtctagagggcccgtttaaacccgctgatcagcctcgactgtgccttctagttgccagccatctgttgttt
gcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg
catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattg
ggaagacaatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagctggg
gctctaggggtatccccacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcg
tgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccgg
ctttccccgtcaagctctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcgaccccaaa
aaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagt
ccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttata
agggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattaattctgtg
gaatgtgtgtcagttagggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctc
aattagtcagcaaccaggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctc
aattagtcagcaaccatagtcccgccctaactccgcccatcccgcccctaactccgcccagttccgcccattc
tccgccccatggctgactaatttttttttatttatgcagaggccgaggccgcctctgcctctgagctattccagaagt
agtgaggaggcttttttggaggcctaggcttttgcaaaaagctcccgggagcttgtatatccattttcggatctga
tcaagagacaggatgaggatcgtttcgcatgattgaacaagatggattgcacgcaggttctccggccgcttgg
gtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgt
cagcgcaggggcgcccggttctttttgtcaagaccgacctgtccggtgccctgaatgaactgcaggacgagg
cagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcgg
gaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaa
agtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccaccaag
cgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaag
agcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctc
gtcgtgacccatggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgt
ggccggctgggtgtggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcg
gcgaatgggctgaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcatcgccttctatcgcc
ttcttgacgagttcttctgagcgggactctggggttcgaaatgaccgaccaagcgacgcccaacctgccatcac
gagatttcgattccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggacgccggctggatga
tcctccagcgcggggatctcatgctggagttcttcgcccaccccaacttgtttattgcagcttataatggttacaa
ataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcat
caatgtatcttatcatgtctgtataccgtcgacctctagctagagcttggcgtaatcatggtcatagctgtttcctgt
gtgaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcct
aatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagc
tgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcact
gactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcca
cagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaa
ggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcag
aggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctg

[2989]

ttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacg
ctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccg
accgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagca
gccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaact
acggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggta
gctcttgatccggcaaacaaaccaccgctggtagcggtttttttgtttgcaagcagcagattacgcgcagaaaa
aaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaaggg
attttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaa
gtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctattt
cgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagt
gctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagg
gccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagta
agtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttg
gtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcg
gttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcac
tgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgag
aatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaac
tttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagt
tcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaa
caggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttt
caatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaa
atagggttccgcgcacatttccccgaaaagtgccacctgacgtc

[2990] 实施例5-

[2991] 用于RNA编辑的哺乳动物构建体的设计和克隆

[2992] 为了产生Cas12b的截短型式，设计了引物来对从C端逐渐截下60bp(20个氨基酸)直至900bp(总共15个截短)的dCas13b进行PCR扩增，并且将这些截短的Cas13b基因Gibson克隆到上述pcDNA-CMV-ADAR2主链中。在U6启动子下，使用金门克隆法将靶向Cluc的指导RNA克隆到在间隔区序列目的位点的3'端处包含该直系同源物的正向重复序列的哺乳动物表达载体中。

[2993] 所使用的荧光素酶报告基因是Cox等人(2017)用来测量RNA编辑的CMV-Cluc(W85X)EF1α-Gluc双重荧光素酶报告基因。该报告载体表达功能性Gluc以作为归一化对照，但是由于添加了W85X早期终止位点，所以所述功能性Gluc是缺陷型Cluc。

[2994] 哺乳动物细胞中的REPAIR编辑

[2995] 为了评估哺乳动物细胞中的REPAIR活性，我们转染了150ng REPAIR载体、300ng指导表达质粒和45ng RNA编辑报告基因。接着在48小时后，我们用分泌的荧光素酶收获培养基，并在Dulbecco的磷酸盐缓冲盐水(PBS)中将培养基以1:10的比例稀释(将10μl培养基放入90μl PBS中)。我们使用注射方案在读板仪(Biotek Synergy Neo2)上使用BioLux海萤和Biolux海肾荧光素酶测定试剂盒(New England Biolabs)测量了荧光素酶活性。进行的所有重复都是生物学重复。

[2996] 将与ADAR脱氨酶结构域融合的核酸酶-死亡Cas13b用于REPAIR，以实现靶向RNA碱基编辑。AAV介导的递送通常用于基因疗法，但是REPAIR超出了AAV载货能力的大小限度。先前我们显示，普雷沃氏菌属种P5-125(PspCas13b)的C端截短不会降低REPAIR活性。我们还使用了来自咽喉卟啉单胞菌的另一种Cas13b直系同源物(PguCas13b)，与PbuCas13b相比，PguCas13b在哺乳动物细胞中稳定表达并显示出高活性。基于PbuCas13b和PguCas13b之间

的比对，我们对PguCas13b进行截短以去除HEPN2结构域，将其融合至ADAR，并且测试了其使用REPAIR系统执行碱基编辑的能力。令人惊讶地是，这些截短的突变体不仅保留了RNA靶向，而且一些突变体在RNA编辑方面明显更有效(图73)。

[2997] ***

[2998] 在不脱离本发明的范围和精神的情况下，本发明所述的方法、药物组合物和药盒的各种修改和变型对于本领域技术人员而言是显而易见的。尽管已结合具体实施方案描述了本发明，但是应当理解，本发明能够进行进一步修改，并且所要求保护的本发明不应该不当地受到这些具体实施方案的限制。实际上，对于本领域技术人员而言显而易见的用于实施本发明的所述模式的各种变型旨在落入本发明的范围内。本申请旨在覆盖本发明的任何变型、用途或改型，这些变型、用途或改型总体遵循本发明的原理且包括在本发明所属领域内已知习惯性实践内的与本公开内容有所偏离的内容，并且可以在陈述之前应用于本文的基本特征。

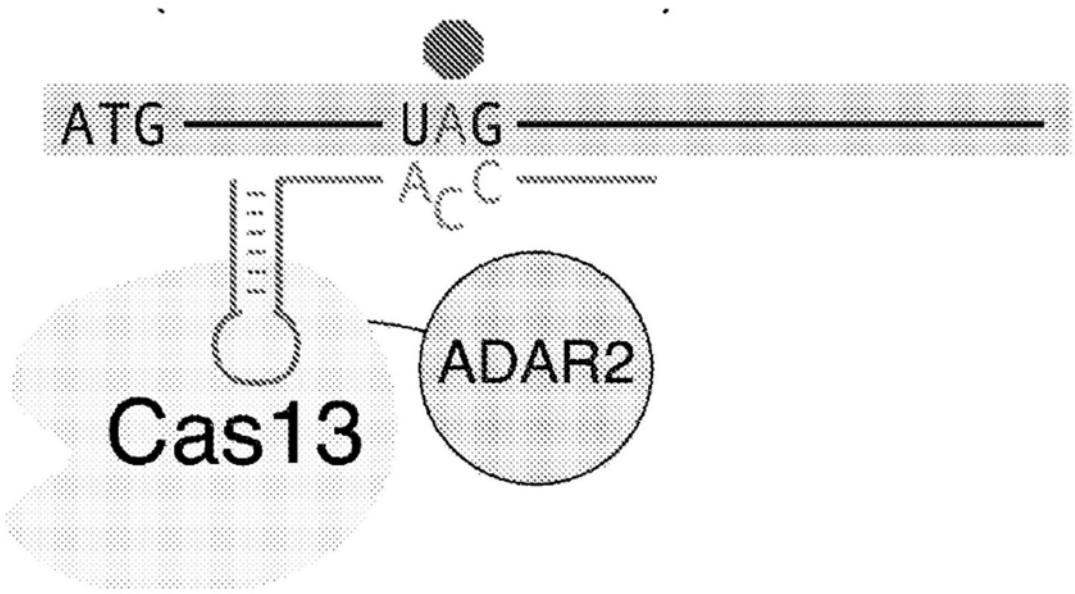

ATG
UAG
AC C
ADAR2
Cas13

图1

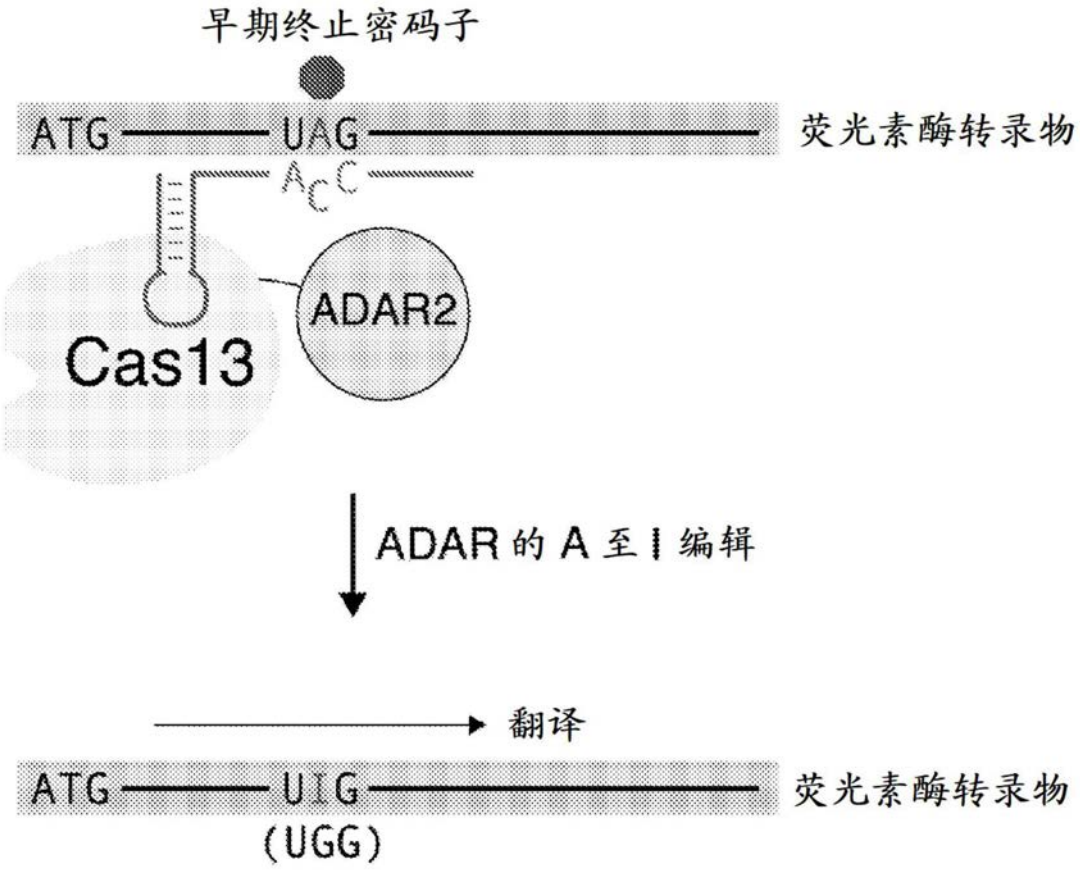

早期终止密码子
ATG
UAG
荧光素酶转录物
AC C
ADAR2
Cas13
ADAR 的 A 至 I 编辑
翻译
ATG
UIG
(UGG)
荧光素酶转录物

图2

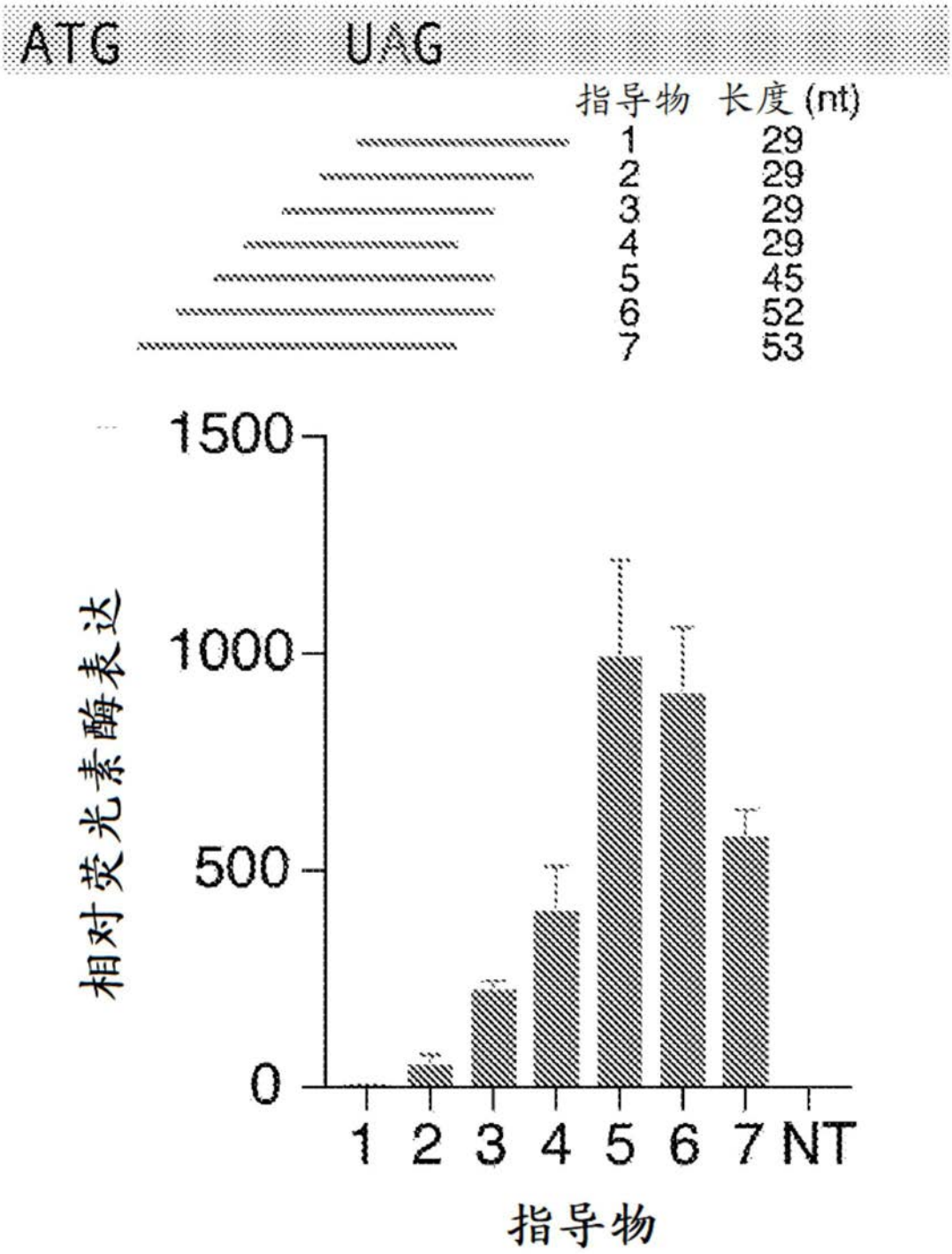

ATG
UAG
指导物 长度 (nt)
1 29
2 29
3 29
4 29
5 45
6 52
7 53
1500
1000
500
0
相对荧光素酶表达
1 2 3 4 5 6 7 NT
指导物

图3

ADAR 名称	蛋白质序列
智人_ADAR2_E_Q_突变体 (SEQ ID NO: 650)	QLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMT TGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSL LRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYIST SPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGQGTIPVR SNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEP IYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISN AEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKH ALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKA RLFTAFIKAGLGAWVEKPTEQDQFSLT*
智人_ADAR1_E_Q_突变体 (SEQ ID NO: 651)	SLGTGNRCVKGDSLSLKGETVNDCHAEIISRRGFIRFLYSELMKY NSQTAKDSIFEPAKGGEKLQIKKTVSFHLYISTAPCGDGALFDKS CSDRAMESTESRHYPVFENPKQGKLRTKVENGQGTIPVESSDIVP TWDGIRLGERLRTMSCSDKILRWNVLGLQGALLTHFLQPIYLKS VTLGYLFSQGHLTRAICCRVTRDGSAFEDGLRHPFIVNHPKVGR VSIYDSKRQSGKTKETSVNWCLADGYDLEILDGTRGTVDGPRNE LSRVSKKNIFLLFKKLCSFRYRRDLLRLSYGEAKKAARDYETAK NYFKKGLKDMGYGNWISKPQEEKNF
普通章鱼_ADAR1_E_Q_突变体 (SEQ ID NO: 652)	SVGTGNRCLTGDHLSLEGNSVNDSHAEMITRRGFLRYLYRHLLE YDAEVPNDLFEKGERSICRIKTNITFHLYISTAPCGDGALFSPRDT DSSNAKMEEENKHIHNPTFSSSVQGLLRTKVEGGQGTIPIDADFT EQTWDGIQRGERLRTMSCSDKICRWNVVGLQGALLSHFIEPIYL DSLTLGYLYDHGHLARAVCCRIERGEASVNQLLPEGYRLNHPW LGRVTACDPPRETQKTKSLSINWCYDDEKSEVLDGTAGICYTAIE KNLFSRLTKHNLYEEFKRVCRKFDRNDLLTAPSYNKAKMMATP FQTAKNVMLKKLKENNCGTWVSKPIEEEMF
乌贼_ADAR1_E_Q_突变体 (SEQ ID NO: 653)	SVGTGNRCLTGDRLSLEGNSVNDSHAEMVTRRGFLRYLYKHLL EYDPEKPHDLFEKGERSLCRIKTNITFHLYISTAPCGDGALFSPRD TDSSNVKVDEENKHVHNPTFSSSVQGLLRTKVEGGQGTIPIDADF TEQTWDGIQRGERLRTMSCSDKICRWNVVGLQGALLSHFVEPIY LESLTLGYLYDHGHLARAVCCRIERGEASVNQLLPEGYRLNHPW LGRVTACDPPRETQKTKSLSINWCYDDEKSEVLDGTAGICYTAIE KNLFSRLTKHSLYEEFKKVCQKFEREDLLNVTSYNKAKMMAIPF QTAKNVMLKKLKENNCGTWVSKPIEEEMF
普通章鱼_ADAR2_E_Q_突变体 (SEQ ID NO: 654)	GIGTGTKCINGEHMSDRGFGVNDCHAEIIARRCFLRYIYDQLELH LSDNSDVRNSSIFELRDKGGYQLKENIQFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRTSGVIQTWDG VLEGERLLTMSCSDKIARWNVLGIQGSLLSHFMNPIYLESIILGSL YHSDHLSRAMYSRISIIENLPEPFHLNRPFLSGISSPESRQPGKAPN FGINWRKEDETFEVINAMTGRVEGGSVSRICKQALFGRFMSLYG KLSSLTGQSVTTRPTHYSDAKAAVMEYQLAKQCVFQAFQKAGL GNWVQKPIEQDQF

图4

乌贼_ADAR2_E_Q_突变体 (SEQ ID NO: 655)	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLRFIYDQLEMH LSEDPEVRGQSVFELRDGGGYKLKPNIHFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRSSGFIQTWDG VLEGERLLTMSCSDKIARWNVLGIQGALLCHFMHPIYLESIILGSL YHSDHLSRAVYCRIASIENLPDLFQLNRPFLSGISSPESRQPGKAP NFGINWRRNDDTFEVINAMTGRVEGGNMSRICKQALFDRFMNL YGRLSSLTGQSVTTRPTLYSEAKAAVMEYQLAKQCVFQAFQKA GLGNWVQKPIEQDQF
乳光枪乌贼_ADAR2_E_Q_突变体 (SEQ ID NO: 656)	GIGTGTKCINGEYMNDRGFAVNDCHAEIIARRCFLRFIYDQLELH LSDNAEVRGQSIFELRDAGGYKLKPNIQFHLYISTAPCGDARIFSP HGQDVETGDRHPNRKARGQLRTKIESGQGTIPVRSSGFIQTWDG VLEGERLLTMSCSDKIARWNVLGVQGALLCHFMHPIYLESIILGS LYHSDHLSRAVYCRIAAIENLPDLFRLNRPFLSGISSPESRQPGKA PNFGINWRRNDDSFEVINAMTGRVEGGSMSRICKQALFDRFMNL YGKLSSLTGQSVTTRPALYSEAKATVMEYQLAKQCVFQAFQKA GLGNWVQKPIEQDQF

图4(续)

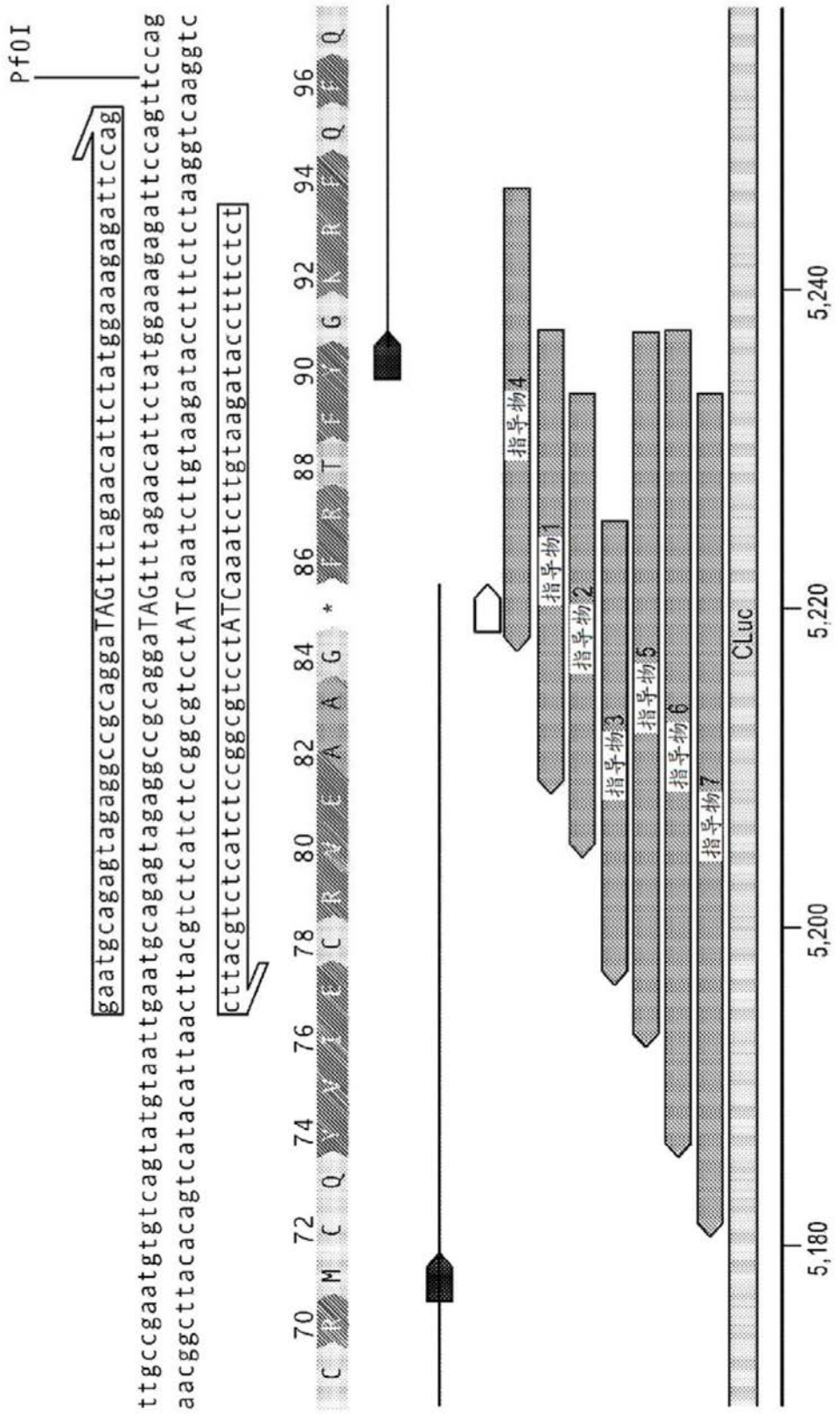

PfoI
gaatgcagagtagaggccgcaggaTAGtttagaacattctatggaaagagattccag
ttgccgaatgtgtcagtatgtaattgaatgcagagtagaggccgcaggaTAGtttagaacattctatggaaagagattccagtccag
aacggcttacacagtcatacattaacttacgtctcatctccggcgtcctATCaaatcttgtaagataccttttctctaaggtcaaggtc
cttacgtctcatctccggcgtcctATCaaatcttgtaagataccttttctct
70 72 74 76 78 80 82 84 86 88 90 92 94 96
C R M C Q Y X E C R X E A A G * F R T F Y G K R F Q
指导物4
指导物1
指导物2
指导物3
指导物5
指导物6
指导物7
CLuc
5,180
5,200
5,220
5,240

图5

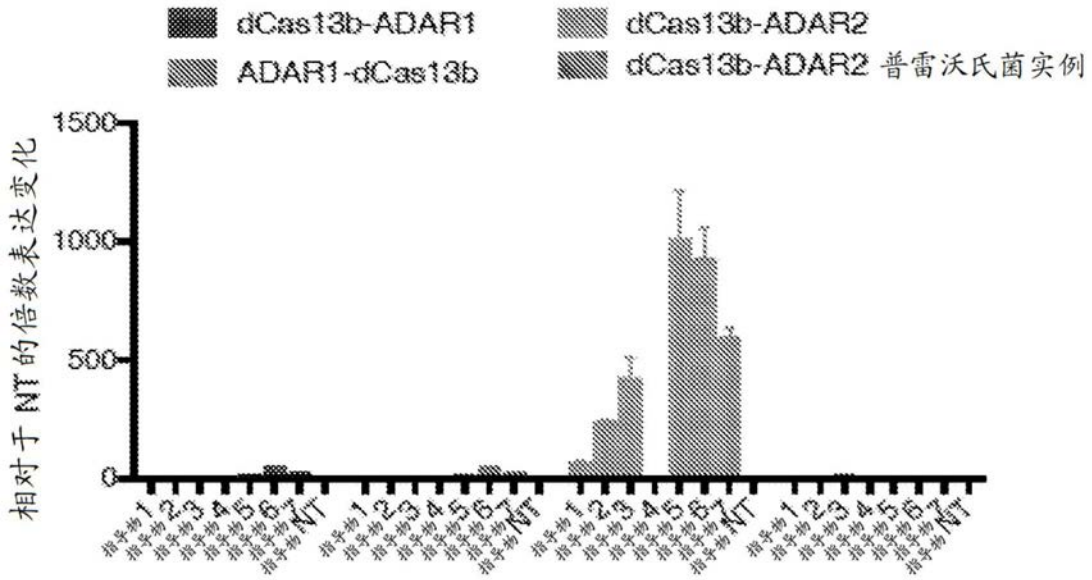

dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
dCas13b-ADAR2 普雷沃氏菌实例
相对于 NT 的倍数表达变化
1500
1000
500
0
指导物 1
指导物 2
指导物 3
指导物 4
指导物 5
指导物 6
指导物 7
指导物 NT

图6

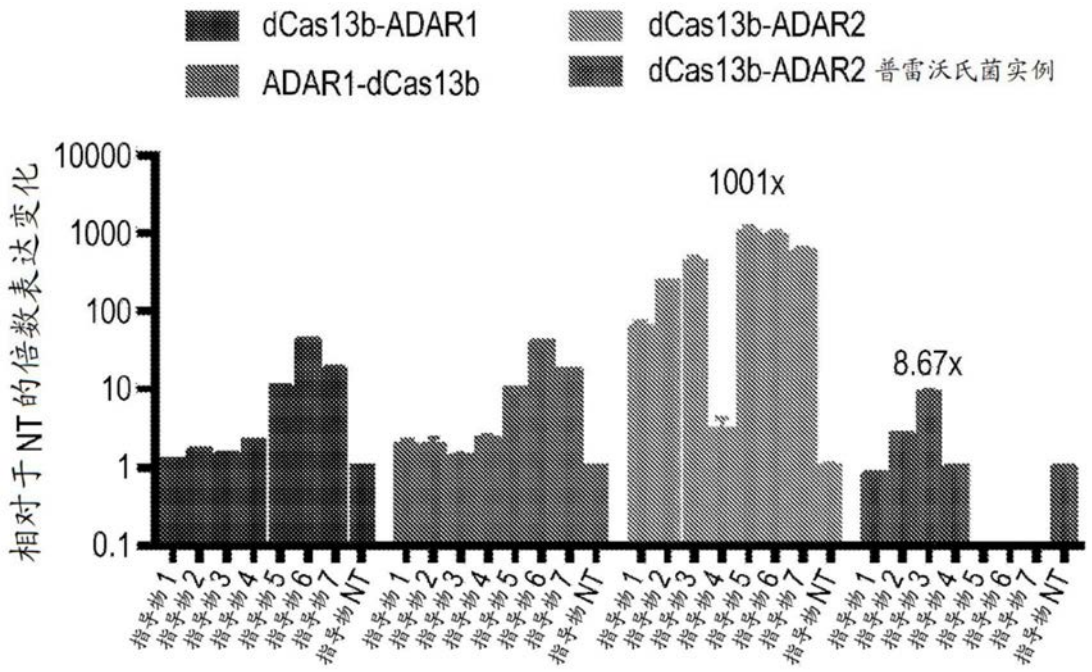

dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
dCas13b-ADAR2 普雷沃氏菌实例
相对于 NT 的倍数表达变化
10000
1000
100
10
1
0.1
1001x
8.67x
指导物 1
指导物 2
指导物 3
指导物 4
指导物 5
指导物 6
指导物 7
指导物 NT

图7

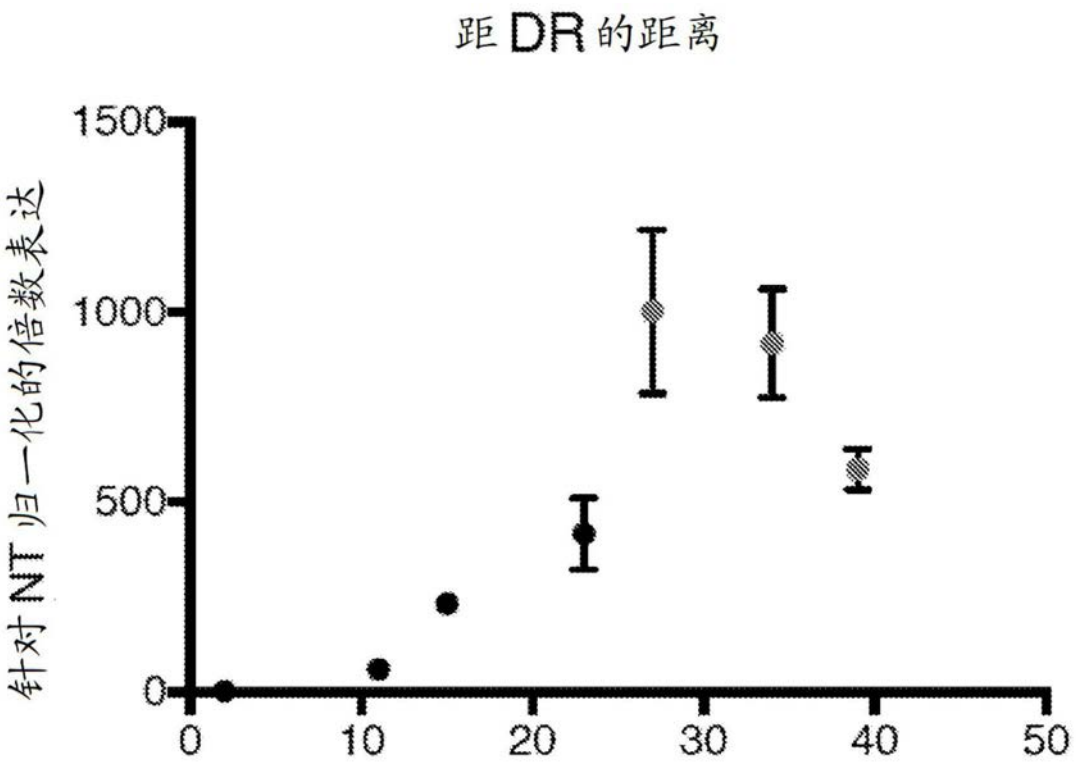

距DR的距离
针对NT归一化的倍数表达
1500
1000
500
0
0
10
20
30
40
50

图8

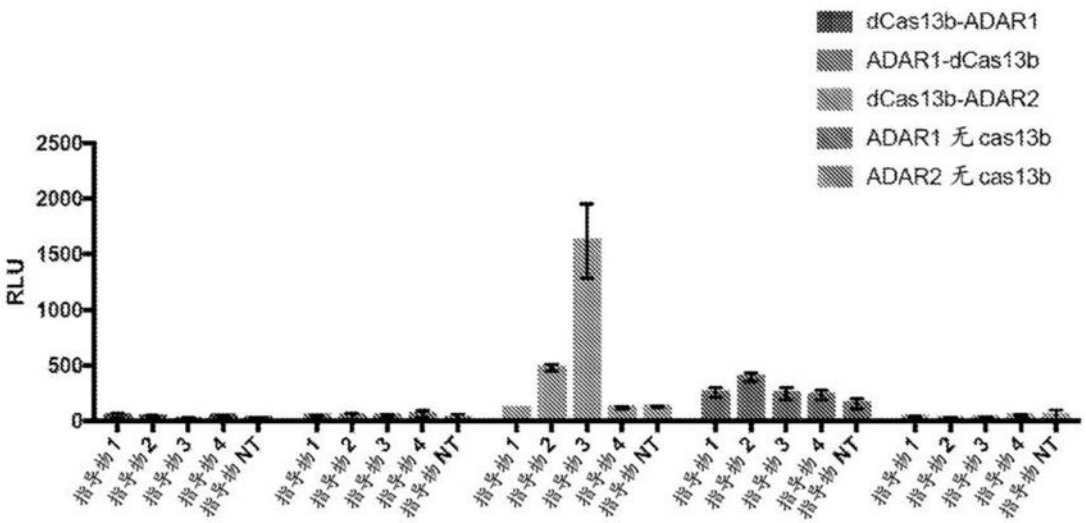

dCas13b-ADAR1
ADAR1-dCas13b
dCas13b-ADAR2
ADAR1 无 cas13b
ADAR2 无 cas13b
RLU
2500
2000
1500
1000
500
0
指导物 1
指导物 2
指导物 3
指导物 4
指导物 NT

图9A

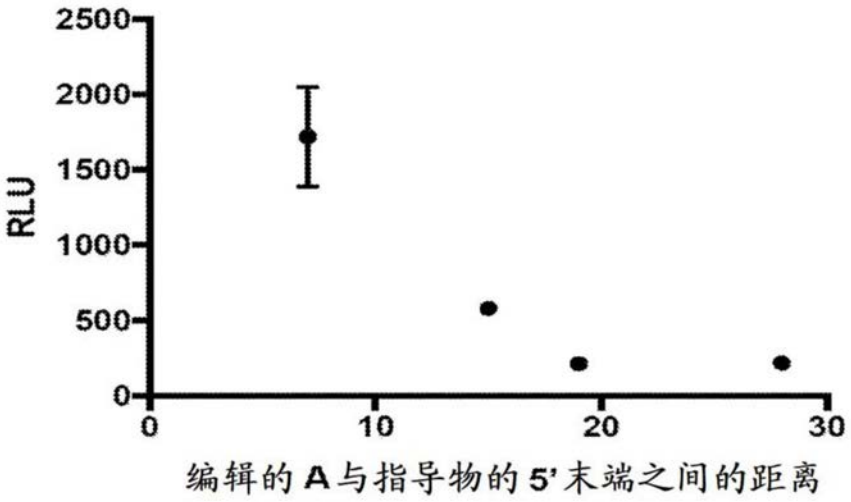


图9B

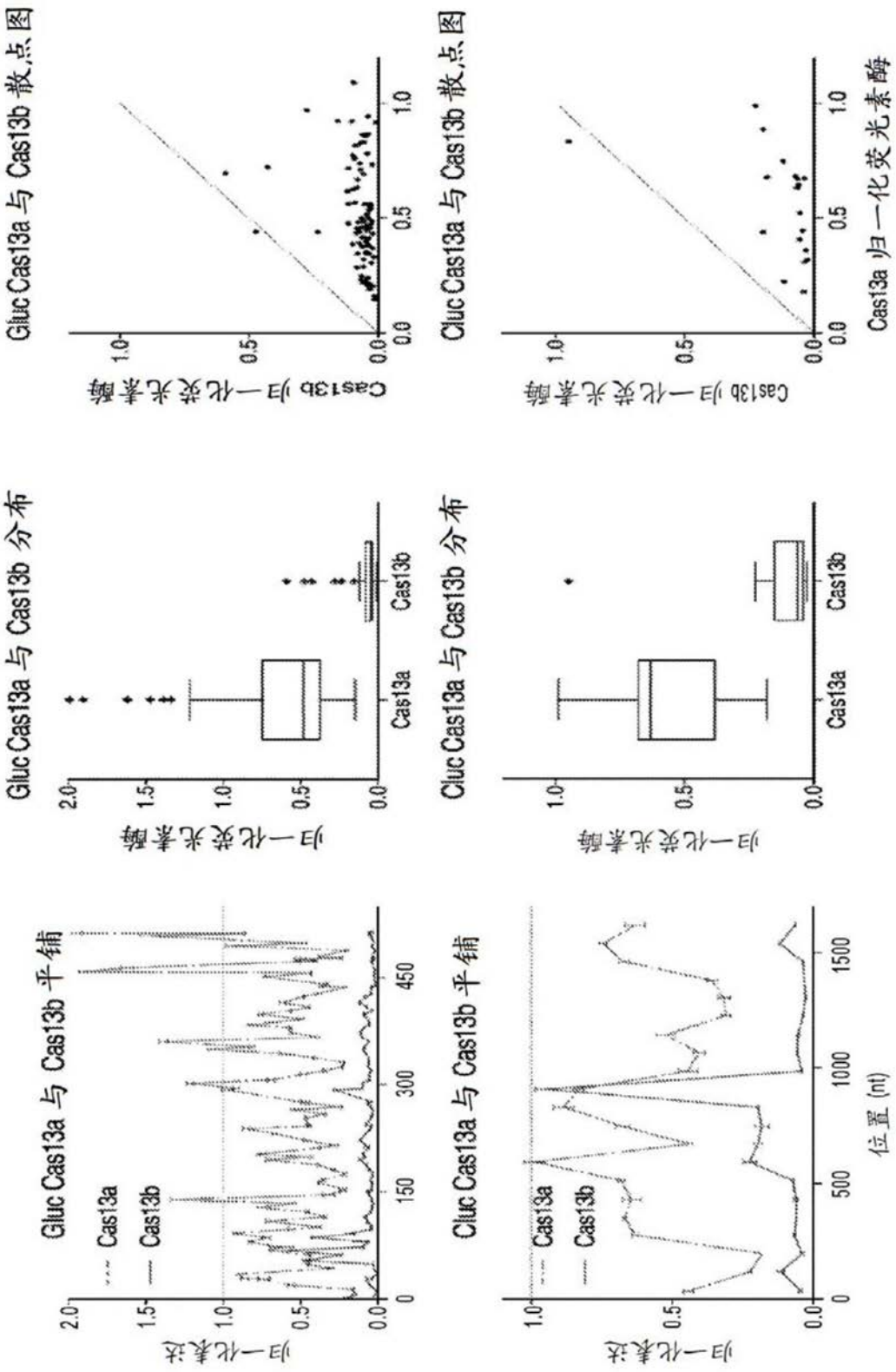

Gluc Cas13a 与 Cas13b 平铺
Cas13a
Cas13b
归一化表达
位置 (nt)
Gluc Cas13a 与 Cas13b 分布
归一化荧光素酶
Cas13a
Cas13b
Gluc Cas13a 与 Cas13b 散点图
Cas13b 归一化荧光素酶
Cas13a 归一化荧光素酶
Cluc Cas13a 与 Cas13b 平铺
Cluc Cas13a 与 Cas13b 分布
Cluc Cas13a 与 Cas13b 散点图

图10



图11A

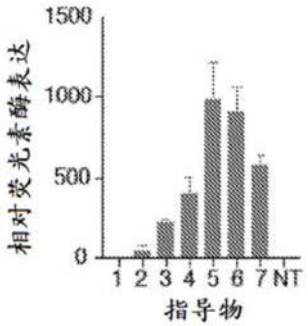


图11B

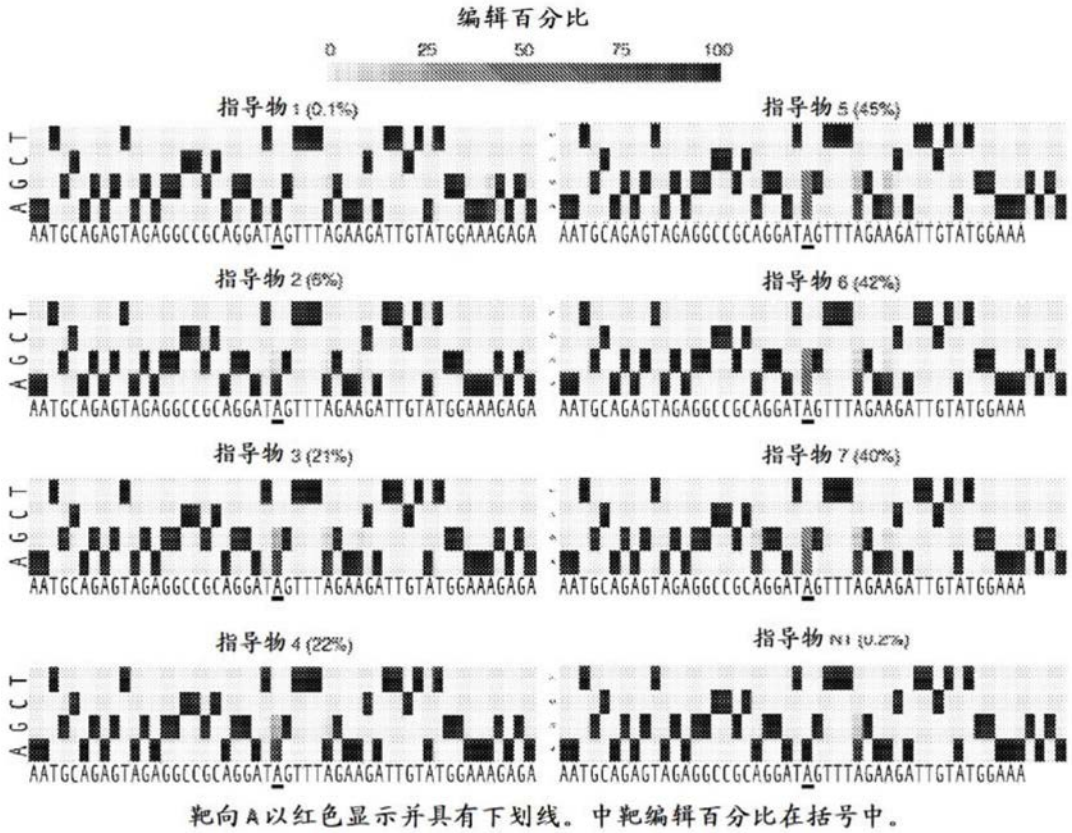
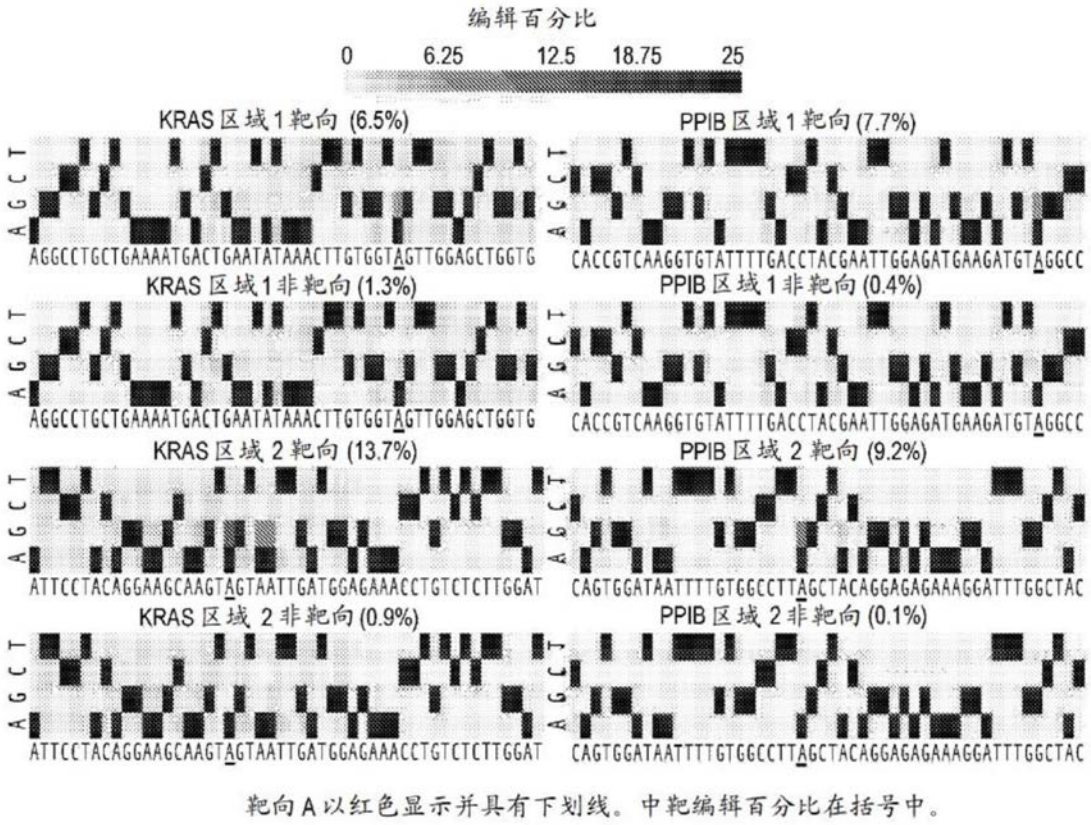


图11C



靶向 A 以红色显示并具有下划线。中靶编辑百分比在括号中。

图12

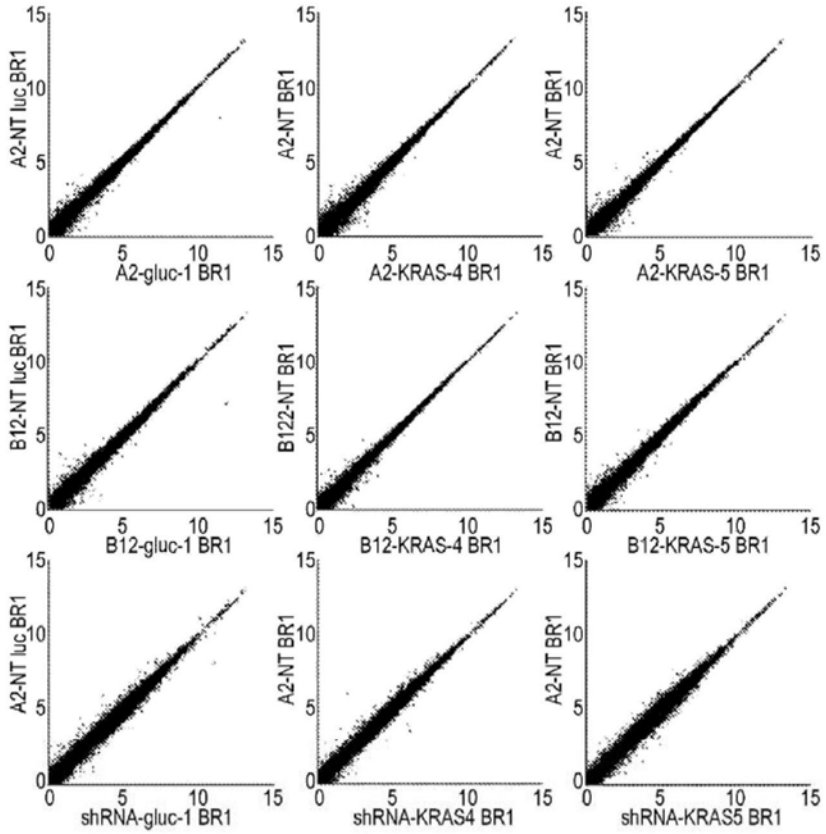

A2-NT luc BR1
A2-gluc-1 BR1
A2-NT BR1
A2-KRAS-4 BR1
A2-NT BR1
A2-KRAS-5 BR1
B12-NT luc BR1
B12-gluc-1 BR1
B122-NT BR1
B12-KRAS-4 BR1
B12-NT BR1
B12-KRAS-5 BR1
A2-NT luc BR1
shRNA-gluc-1 BR1
A2-NT BR1
shRNA-KRAS4 BR1
A2-NT BR1
shRNA-KRAS5 BR1

图13A

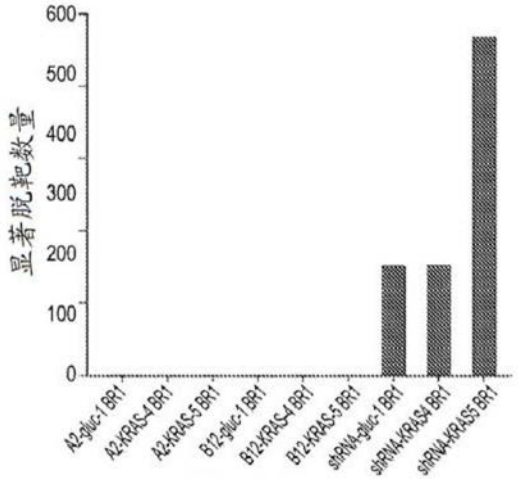

显著脱靶数量
A2-gluc-1 BR1
A2-KRAS-4 BR1
A2-KRAS-5 BR1
B12-gluc-1 BR1
B12-KRAS-4 BR1
B12-KRAS-5 BR1
shRNA-gluc-1 BR1
shRNA-KRAS4 BR1
shRNA-KRAS5 BR1

图13B

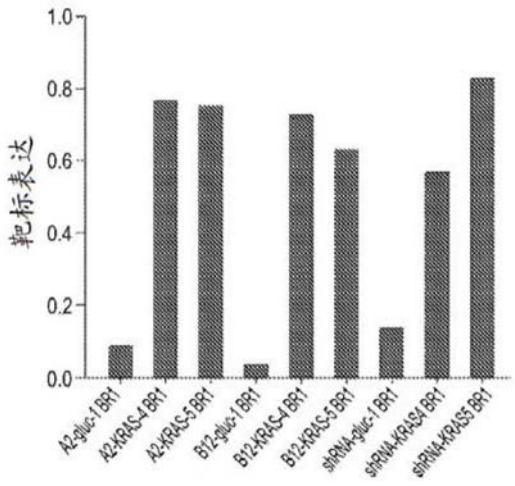

1.0
0.8
0.6
0.4
0.2
0.0
靶标表达
A2-gluc-1 BR1
A2-KRAS-4 BR1
A2-KRAS-5 BR1
B12-gluc-1 BR1
B12-KRAS-4 BR1
B12-KRAS-5 BR1
shRNA-gluc-1 BR1
shRNA-KRAS4 BR1
shRNA-KRAS5 BR1

图13C

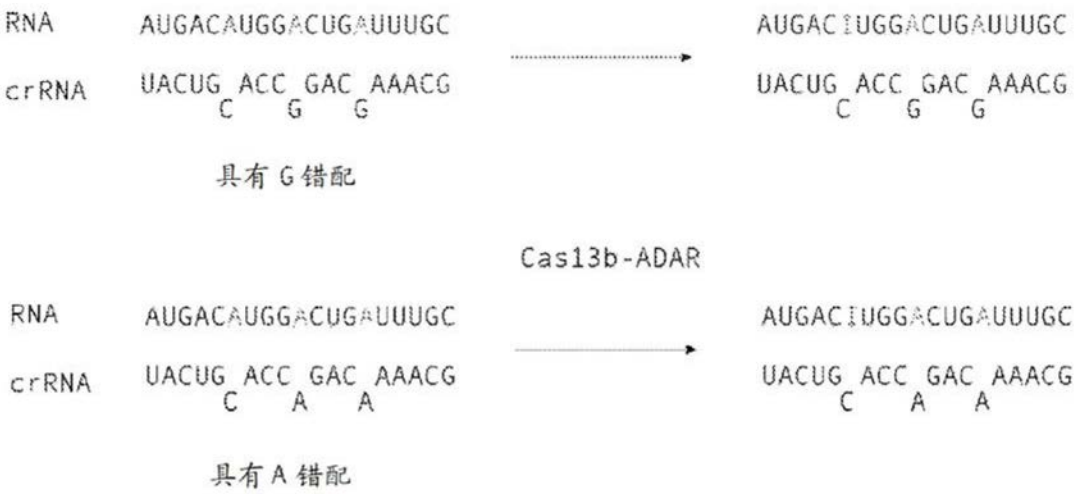

RNA
AUGACAUGGACUGAUUUGC
crRNA
UACUG ACC GAC AAACG
C G G
具有 G 错配
AUGACIUGGACUGAUUUGC
UACUG ACC GAC AAACG
C G G
Cas13b-ADAR
RNA
AUGACAUGGACUGAUUUGC
crRNA
UACUG ACC GAC AAACG
C A A
具有 A 错配
AUGACIUGGACUGAUUUGC
UACUG ACC GAC AAACG
C A A

图14

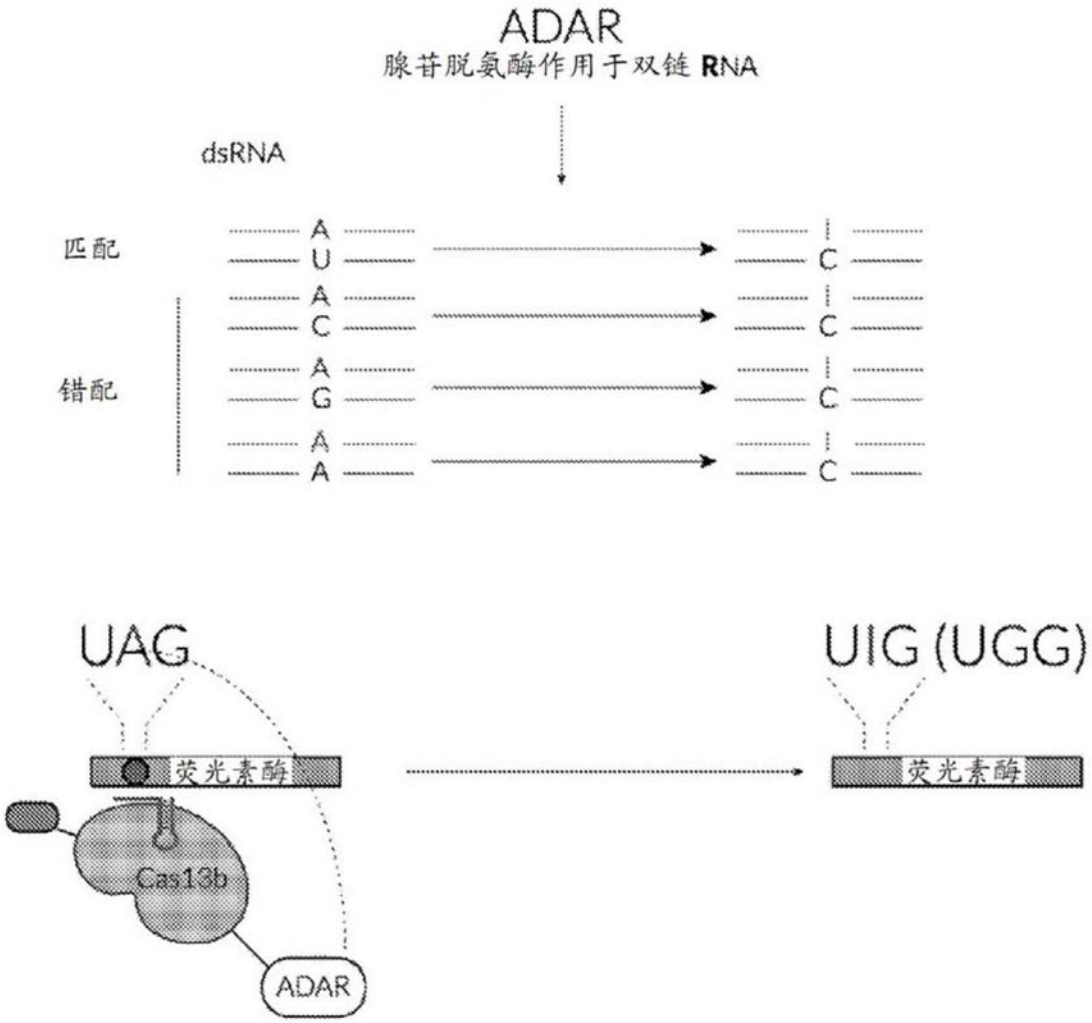

ADAR
腺苷脱氨酶作用于双链 RNA
dsRNA
匹配
错配
A
U
C
A
C
C
A
G
C
A
A
C
UAG
UIG (UGG)
荧光素酶
荧光素酶
Cas13b
ADAR

图15

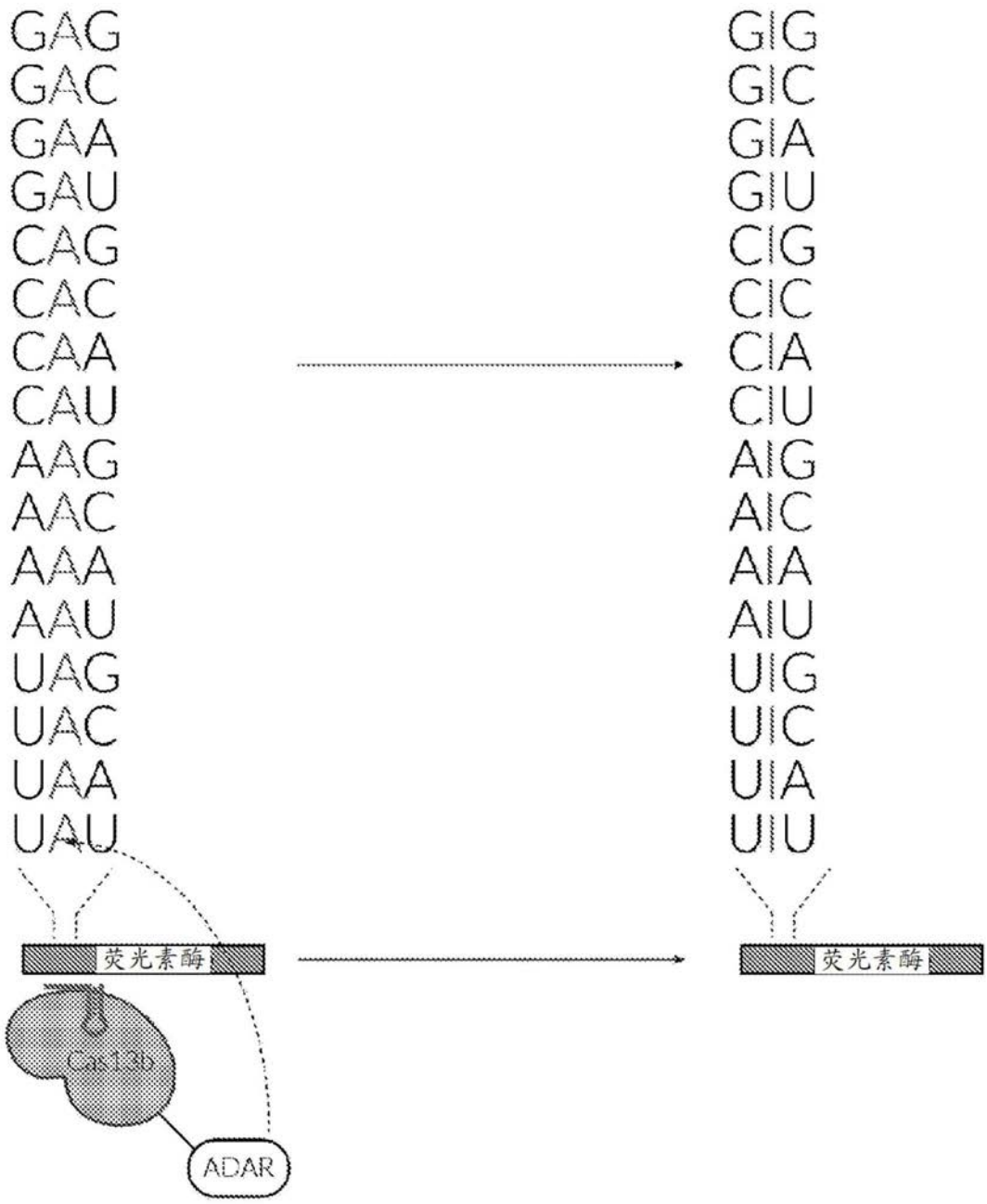

GAG
GAC
GAA
GAU
CAG
CAC
CAA
CAU
AAG
AAC
AAA
AAU
UAG
UAC
UAA
UAU
G|G
G|C
G|A
G|U
C|G
C|C
C|A
C|U
A|G
A|C
A|A
A|U
U|G
U|C
U|A
U|U
荧光素酶
荧光素酶
Cas13b
ADAR

图16

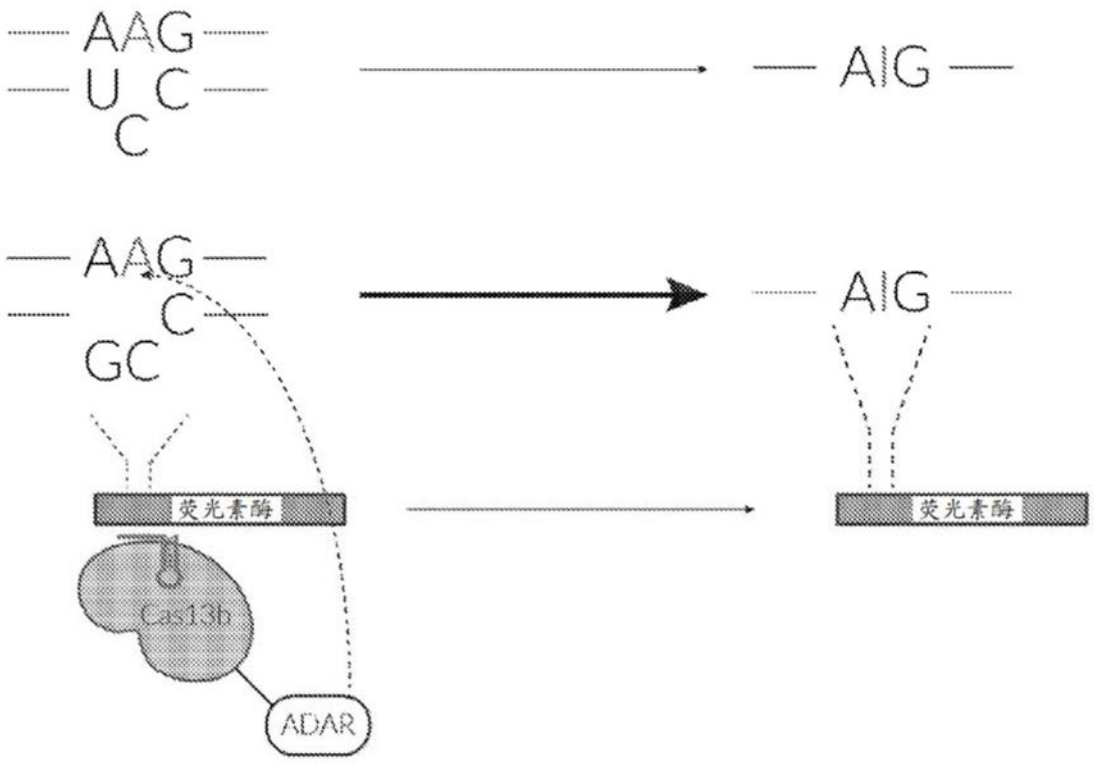

AAG
U C
C
AIG
AAG
C
GC
AIG
荧光素酶
荧光素酶
Cas13b
ADAR

图17

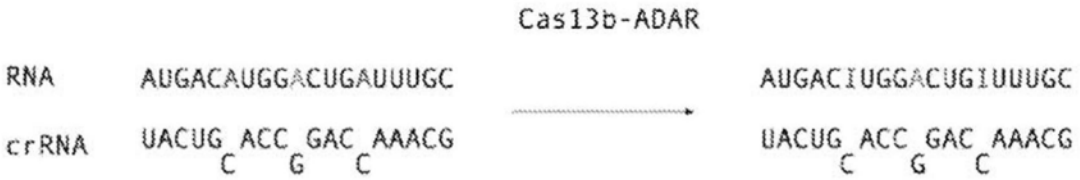

Cas13b-ADAR
RNA
AUGACAUGGACUGAUUUGC
AUGACIUGGACUGIUUUGC
crRNA
UACUG ACC GAC AAACG
C G C
UACUG ACC GAC AAACG
C G C

图18

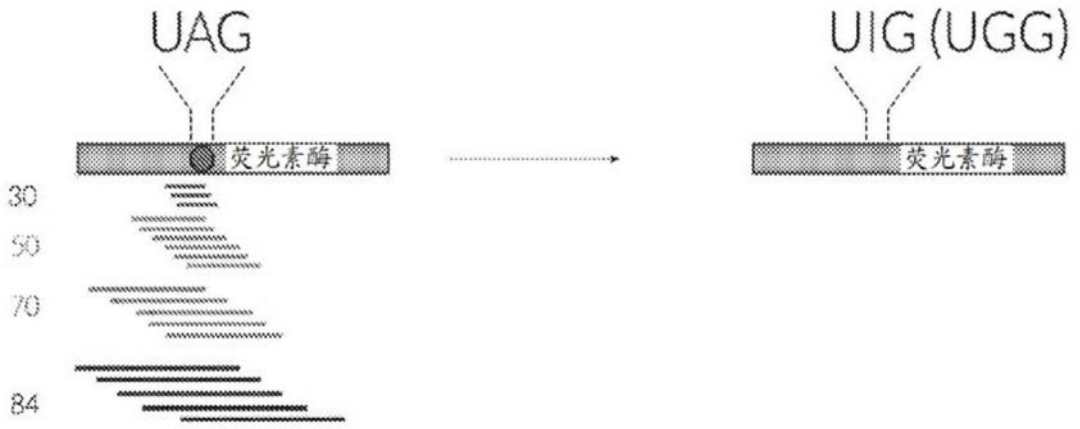

UAG
荧光素酶
UIG (UGG)
荧光素酶
30
50
70
84

图19

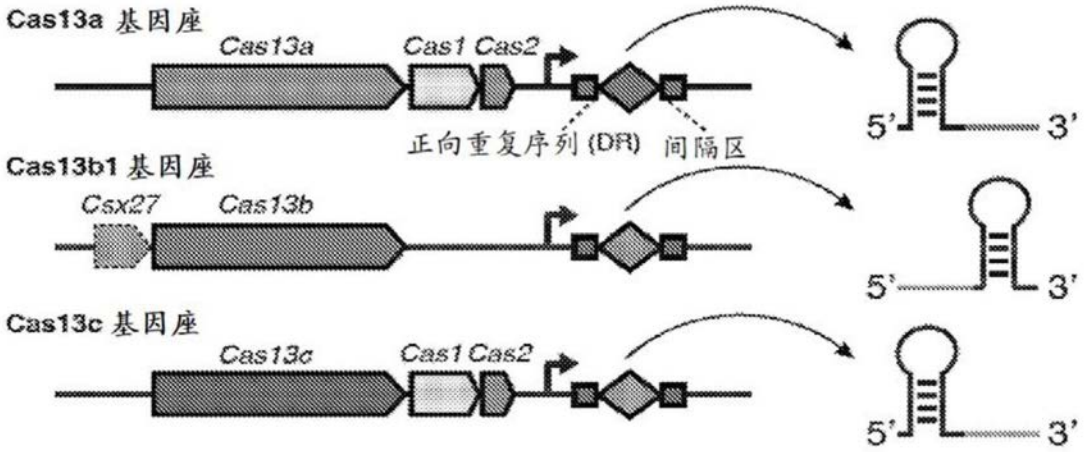

Cas13a 基因座
Cas13a
Cas1 Cas2
正向重复序列(DR)
间隔区
5'
3'
Cas13b1 基因座
Csx27
Cas13b
5'
3'
Cas13c 基因座
Cas13c
Cas1 Cas2
5'
3'

图20A

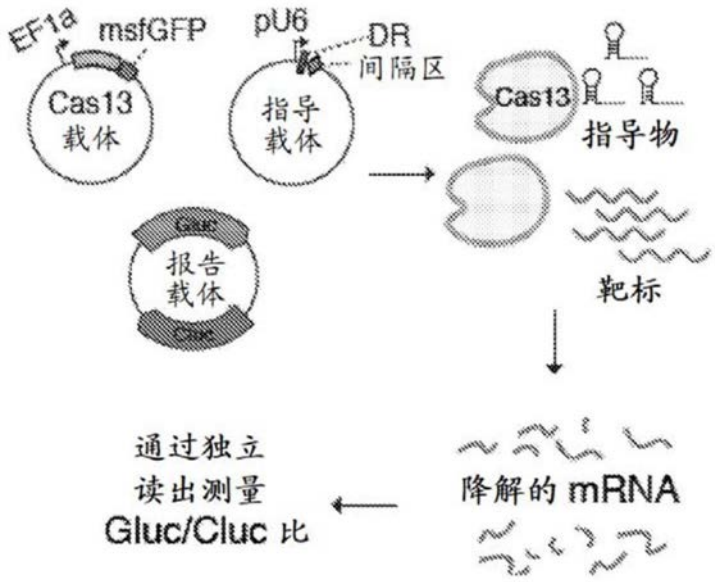

EF1a
msfGFP
Cas13
载体
pU6
DR
间隔区
指导
载体
Cas13
指导物
Gluc
报告
载体
Cluc
靶标
降解的 mRNA
通过独立
读出测量
Gluc/Cluc 比

图20B

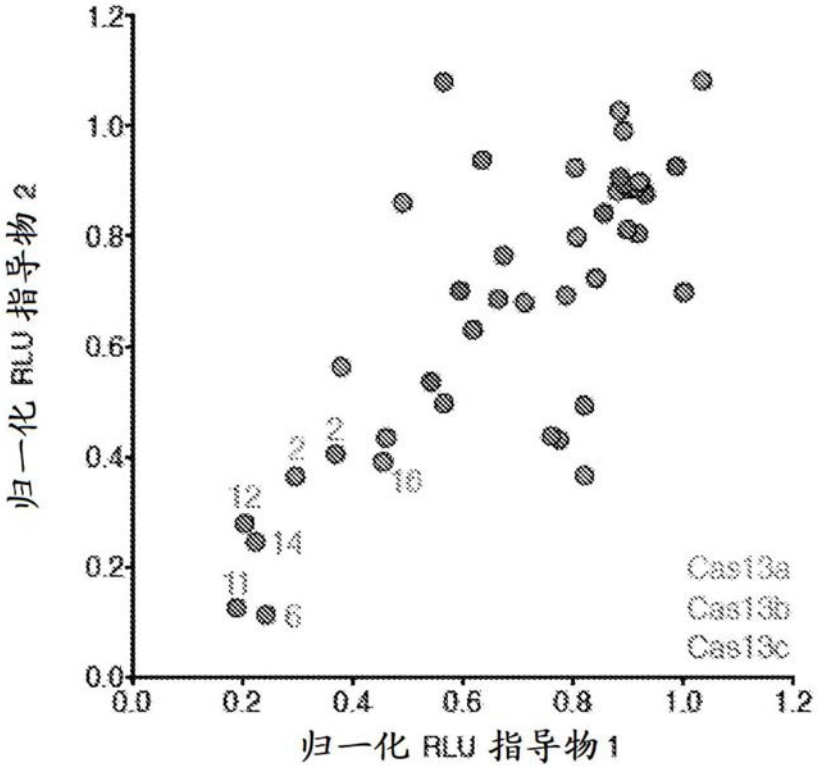

归一化 RLU 指导物 2
归一化 RLU 指导物 1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
12
14
11
6
2
2
16
Cas13a
Cas13b
Cas13c

图20C

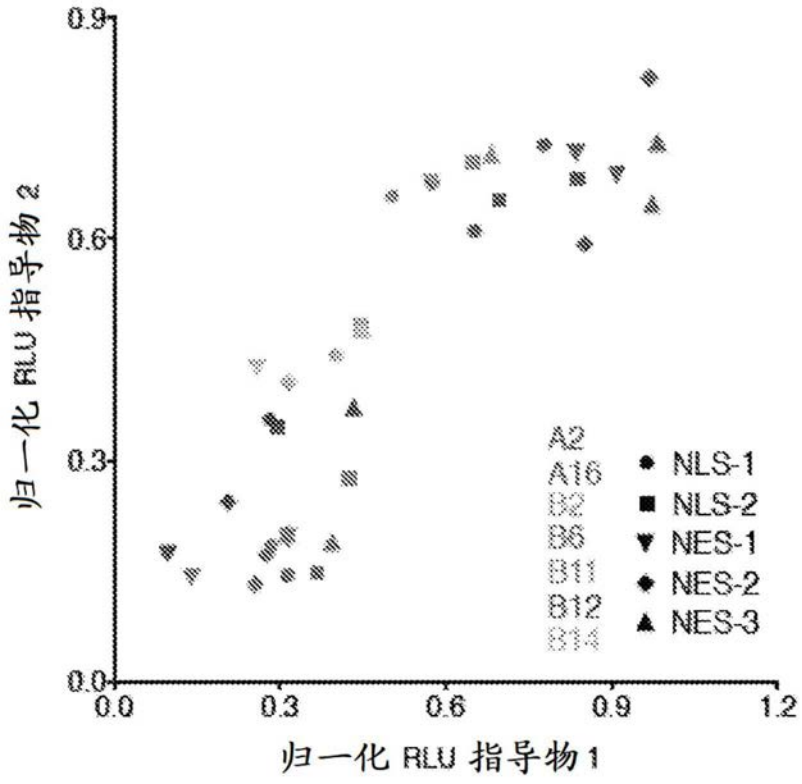

归一化 RLU 指导物 2
归一化 RLU 指导物 1
0.0
0.3
0.6
0.9
1.2
A2
A16
B2
B6
B11
B12
B14
NLS-1
NLS-2
NES-1
NES-2
NES-3

图20D

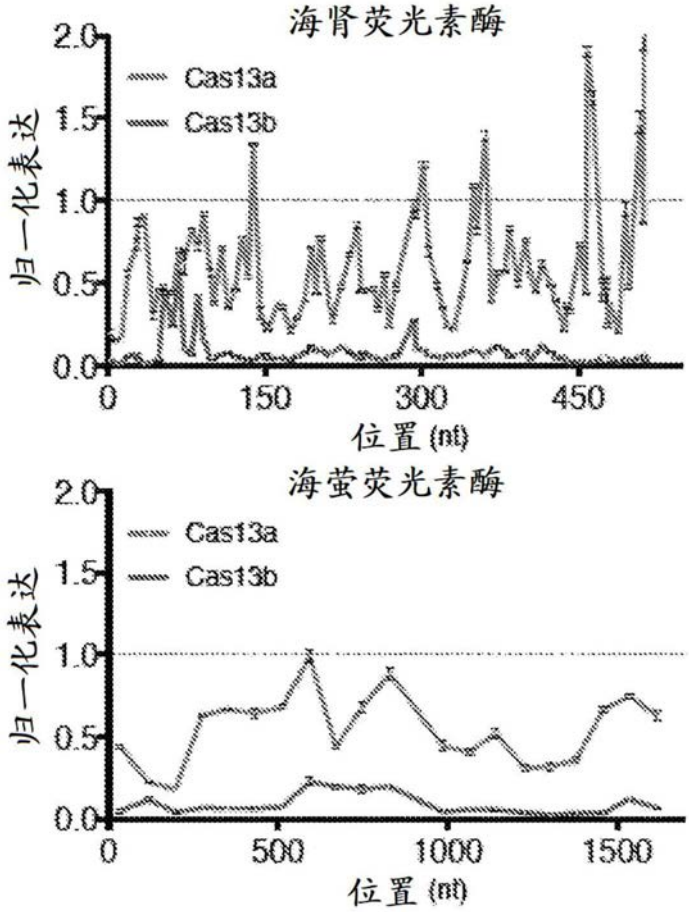

海肾荧光素酶
Cas13a
Cas13b
归一化表达
位置 (nt)
海萤荧光素酶
Cas13a
Cas13b
归一化表达
位置 (nt)

图20E

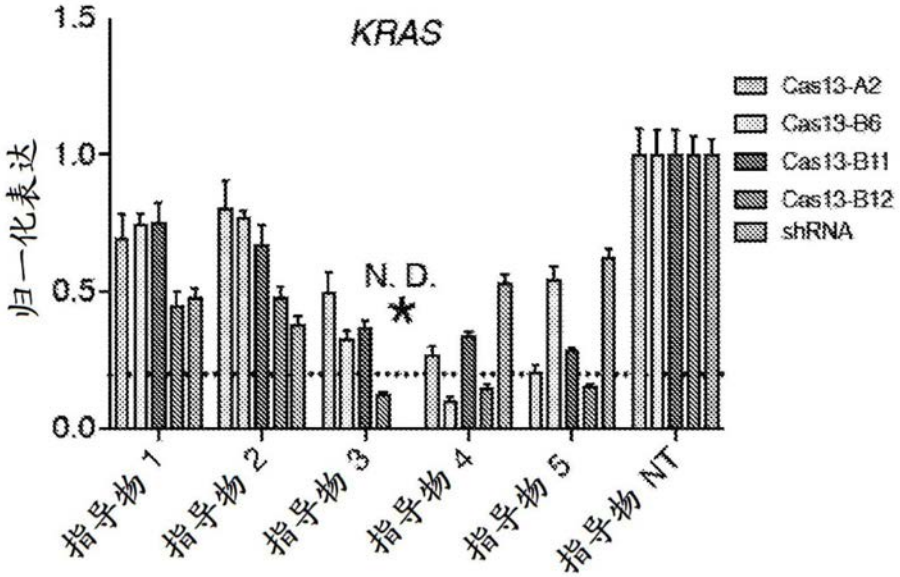

KRAS
Cas13-A2
Cas13-B6
Cas13-B11
Cas13-B12
shRNA
归一化表达
N. D.
指导物 1
指导物 2
指导物 3
指导物 4
指导物 5
指导物 NT

图20F

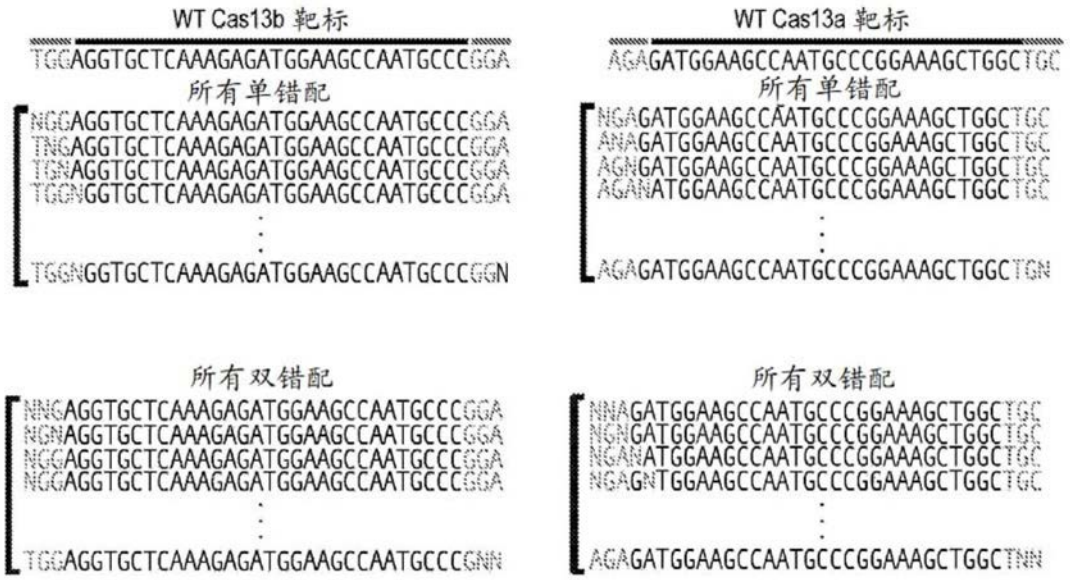


图21A

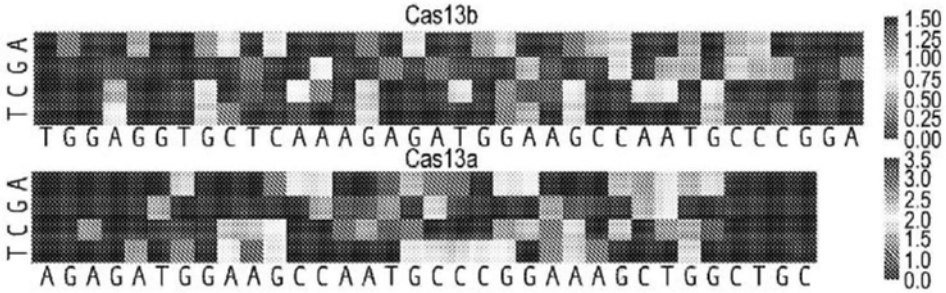


图21B

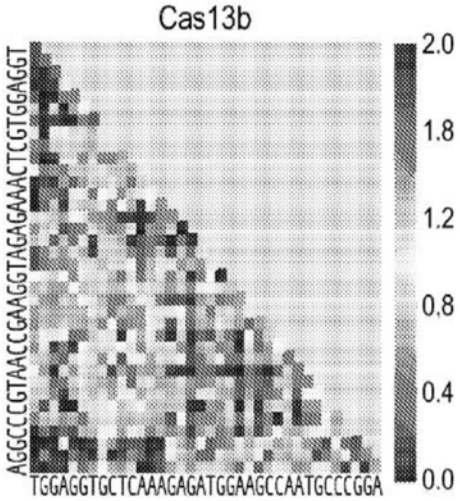


图21C

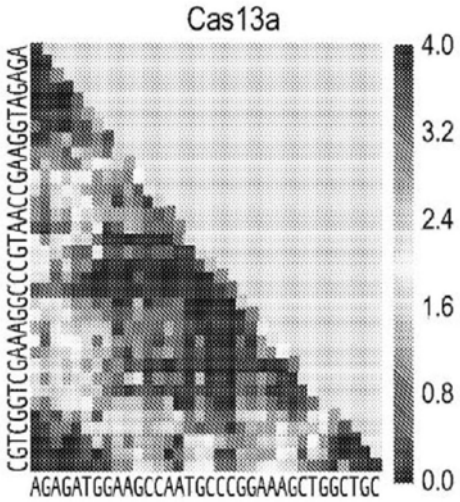

Cas13a
4.0
3.2
2.4
1.6
0.8
0.0
CGTCGGTCGAAAGGCCCGTAACCGAAGGTAGAGA
AGAGATGGAAGCCAATGCCCGGAAAGCTGGCTGC

图21D

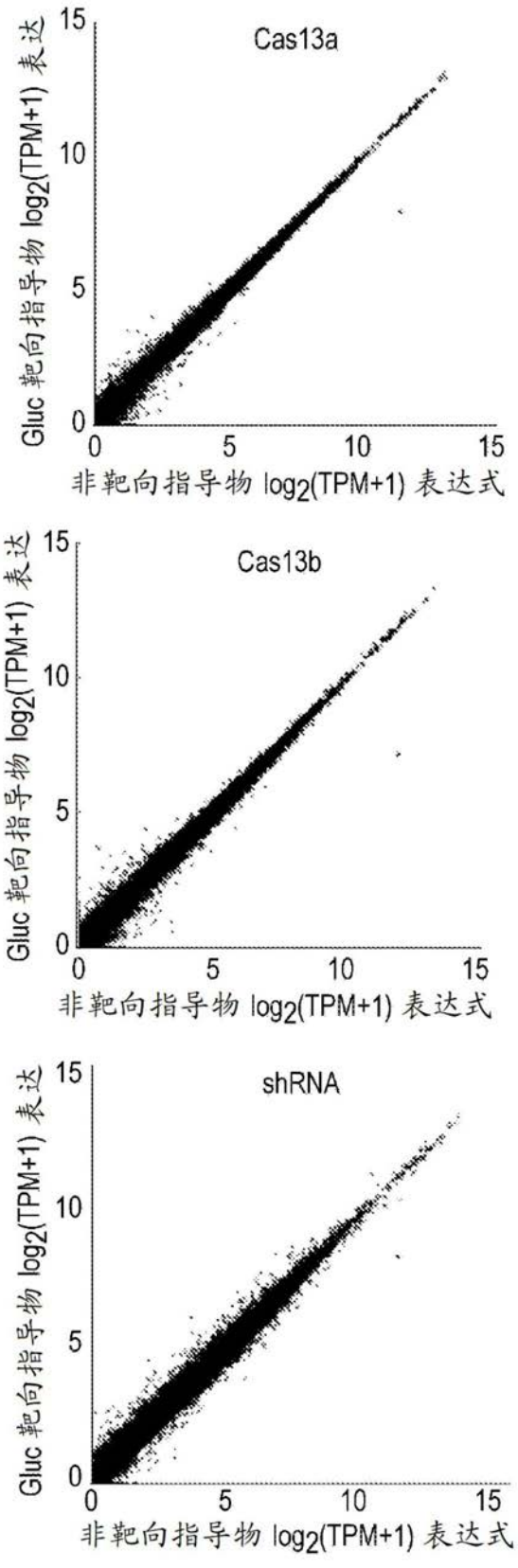

Cas13a
Gluc 靶向指导物 log2(TPM+1) 表达
非靶向指导物 log2(TPM+1) 表达式
Cas13b
Gluc 靶向指导物 log2(TPM+1) 表达
非靶向指导物 log2(TPM+1) 表达式
shRNA
Gluc 靶向指导物 log2(TPM+1) 表达
非靶向指导物 log2(TPM+1) 表达式
0
5
10
15

图21E

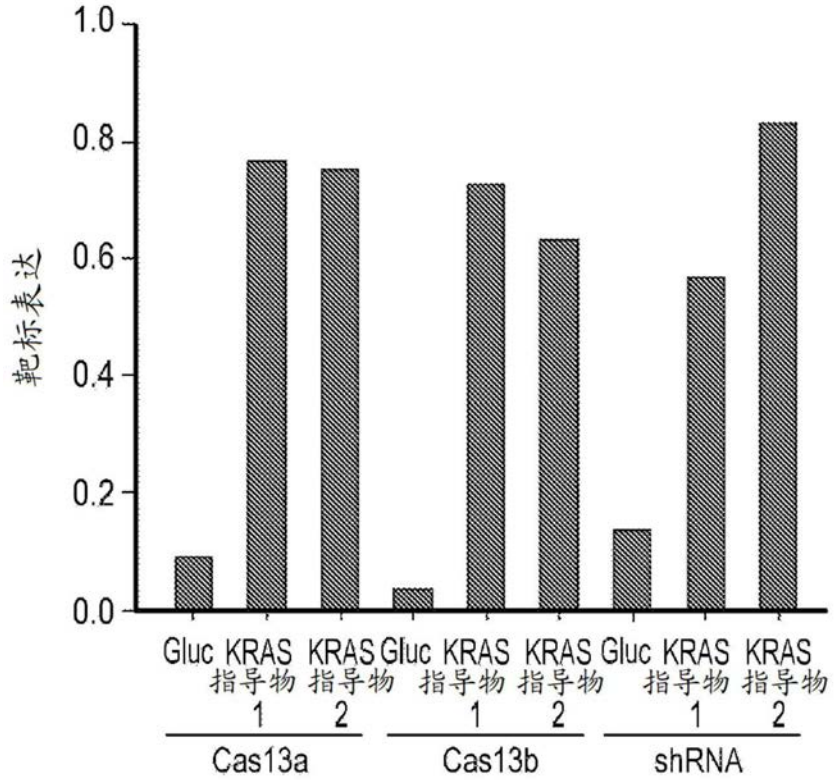

1.0
0.8
0.6
0.4
0.2
0.0
靶标表达
Gluc
KRAS 指导物 1
KRAS 指导物 2
Cas13a
Cas13b
shRNA

图21F

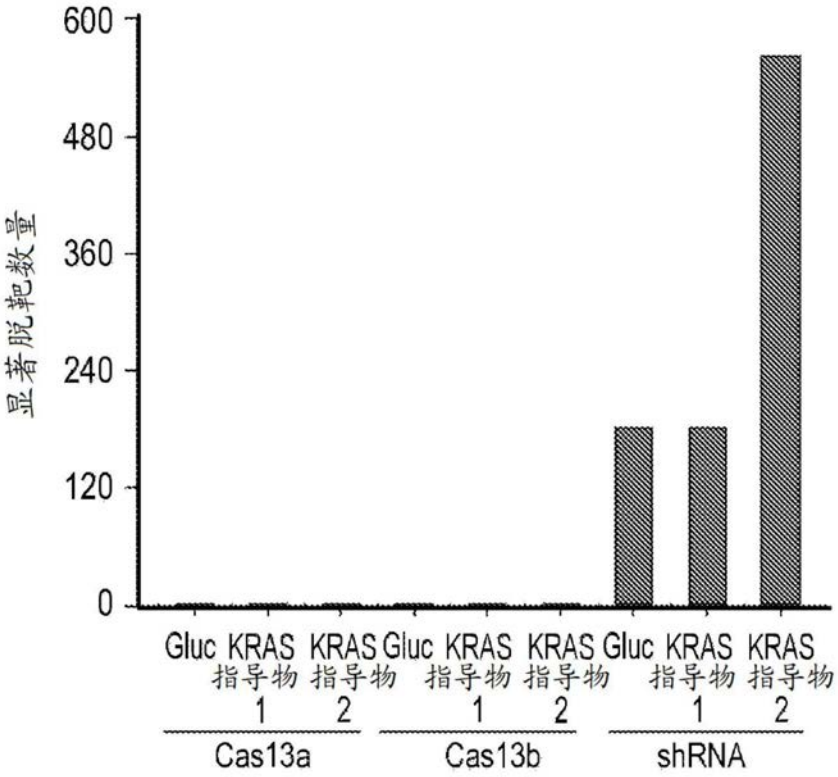

600
480
360
240
120
0
显著脱靶数量
Gluc
KRAS 指导物 1
KRAS 指导物 2
Cas13a
Cas13b
shRNA

图21G

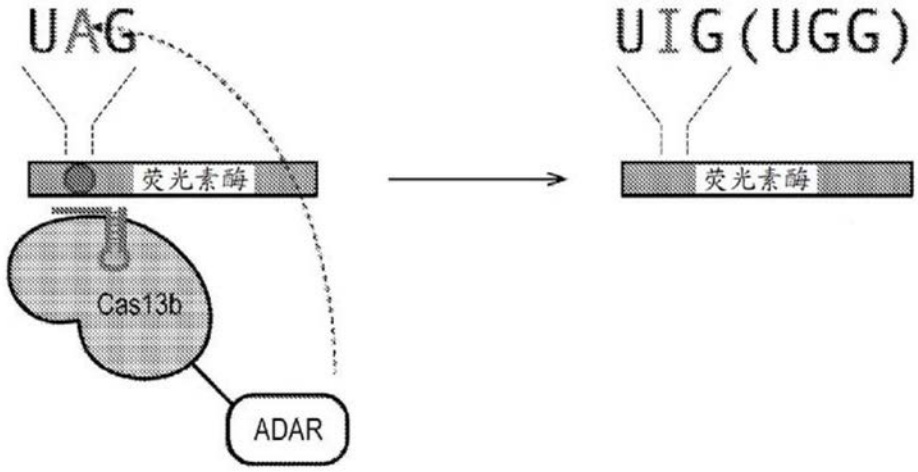

UAG
UIG(UGG)
荧光素酶
荧光素酶
Cas13b
ADAR

图22A

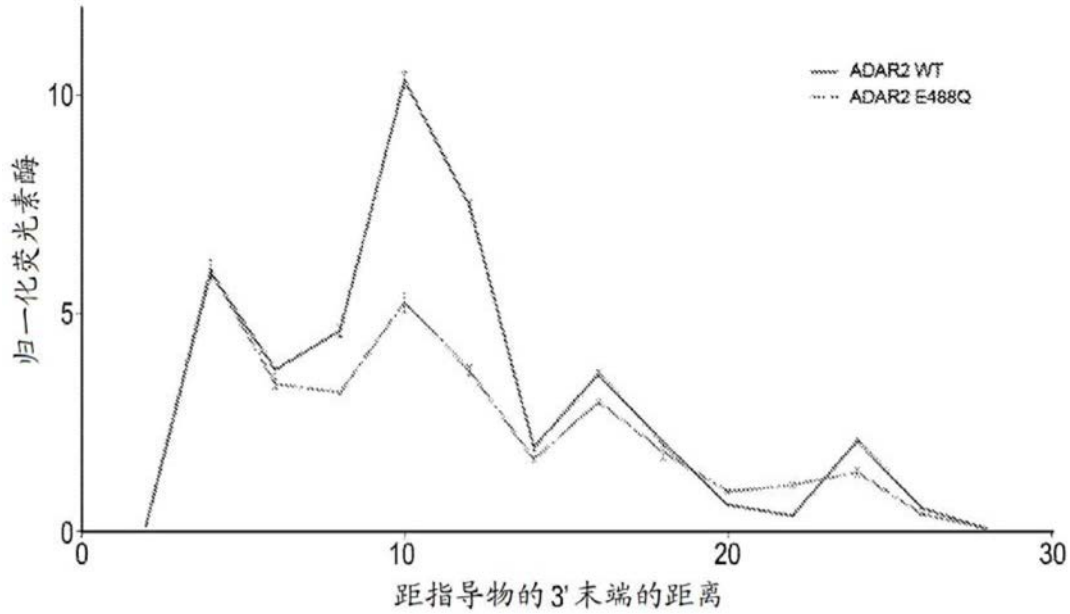

ADAR2 WT
ADAR2 E488Q
归一化荧光素酶
0
5
10
0
10
20
30
距指导物的3'末端的距离

图22B

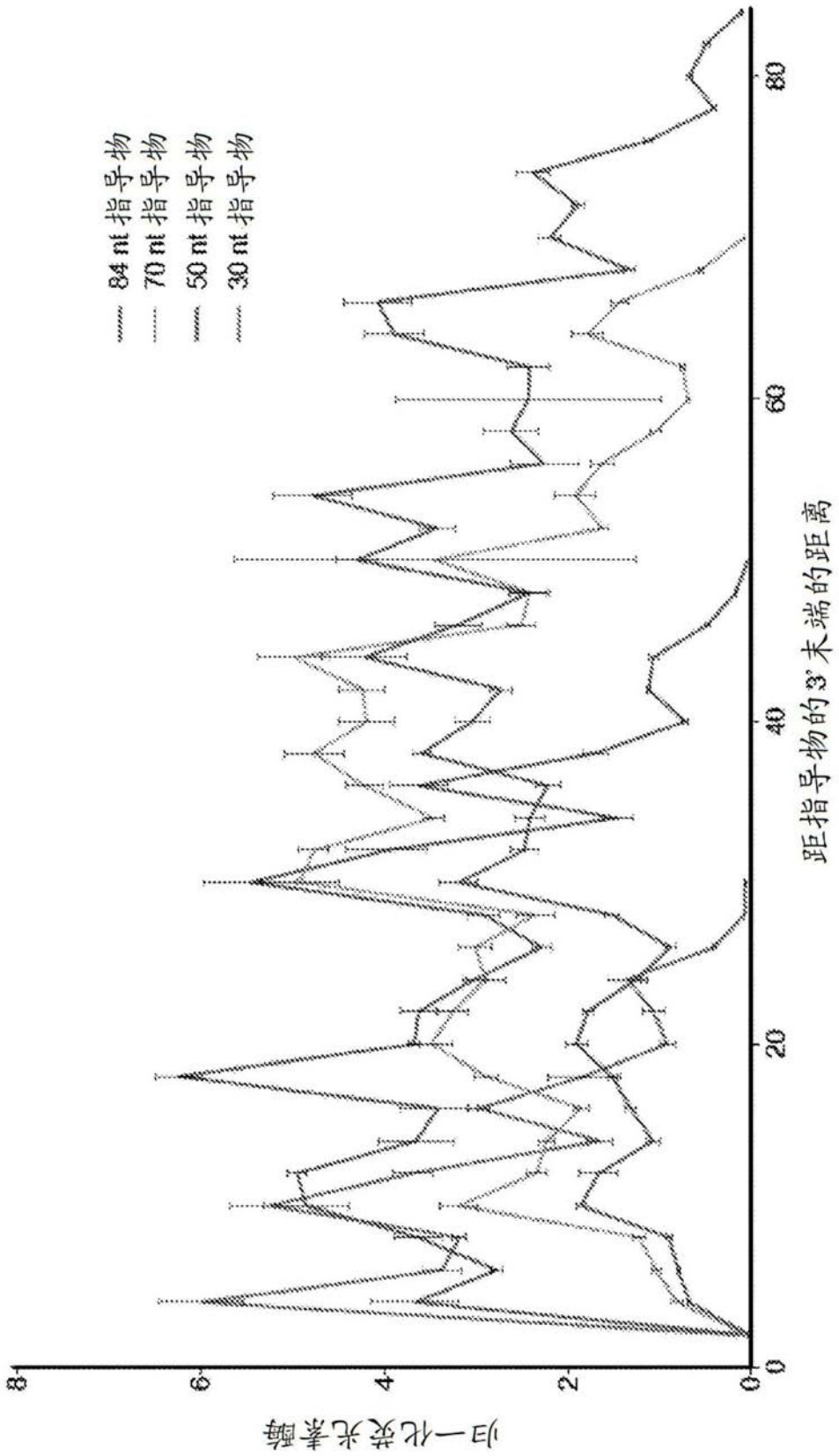

84 nt 指导物
70 nt 指导物
50 nt 指导物
30 nt 指导物
归一化荧光素酶
0
2
4
6
8
0
20
40
60
80
距指导物的3'末端的距离

图22C

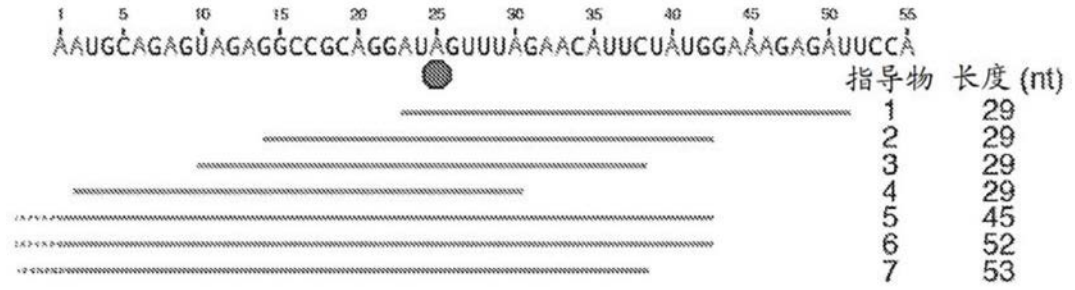


图22D

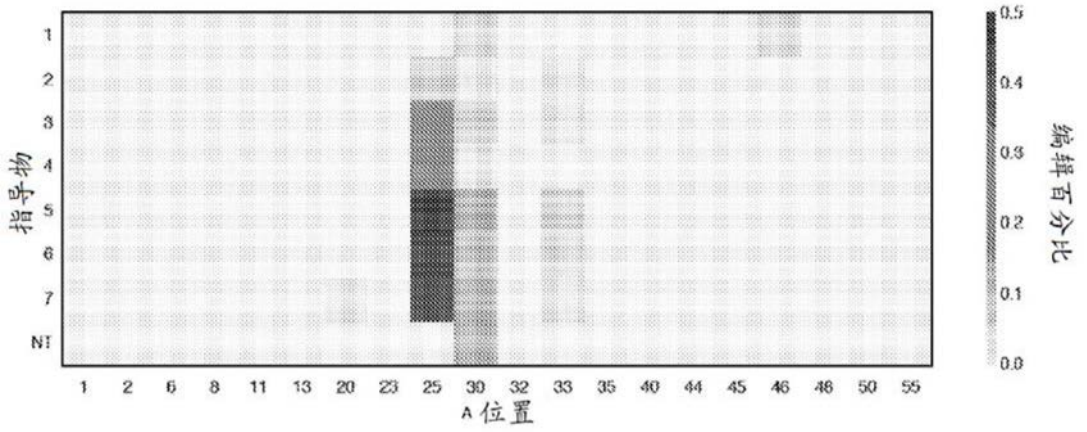


图22E

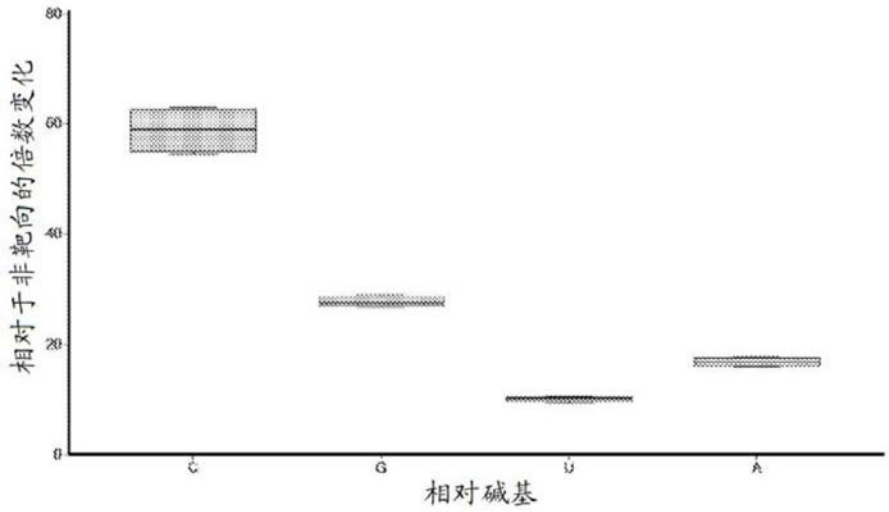


图22F

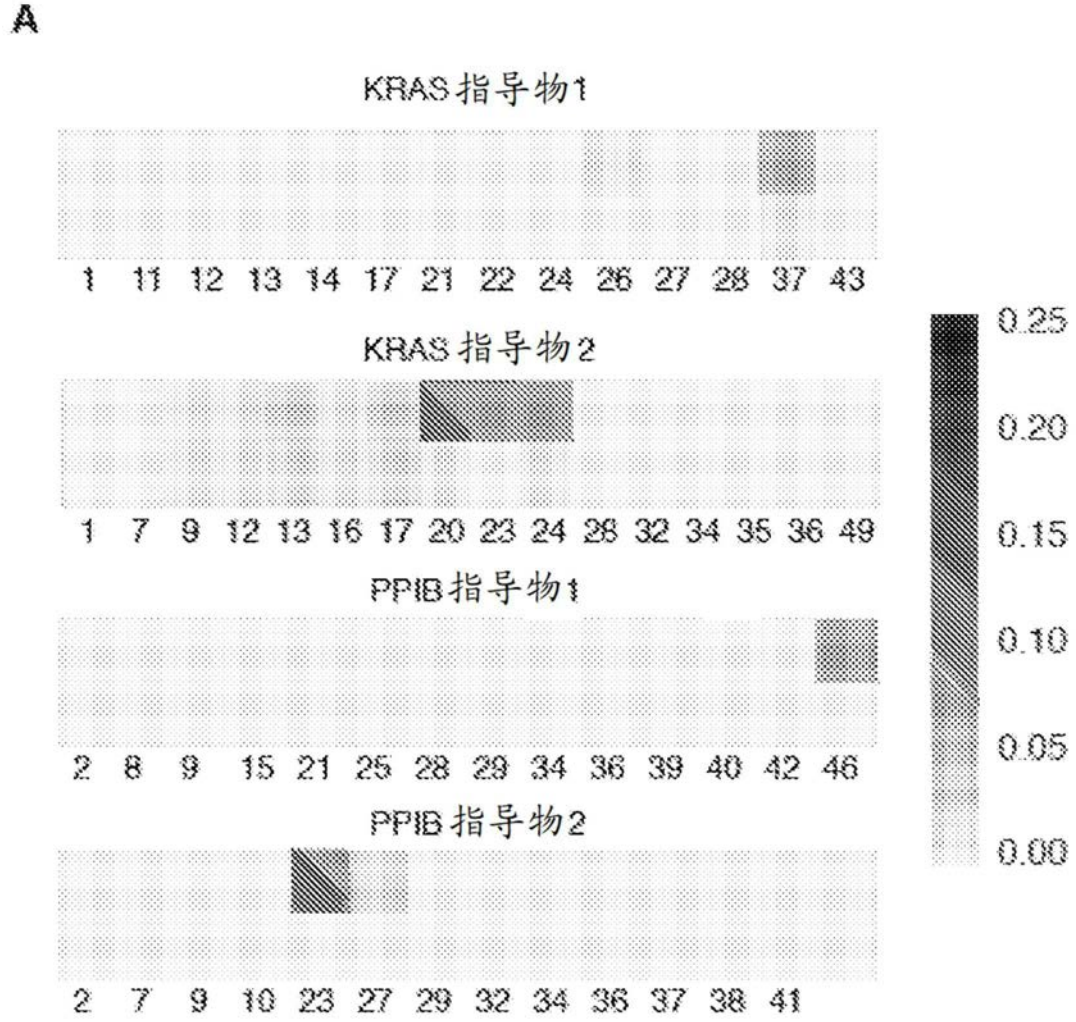

A
KRAS指导物1
1 11 12 13 14 17 21 22 24 26 27 28 37 43
KRAS指导物2
1 7 9 12 13 16 17 20 23 24 28 32 34 35 36 49
PPIB指导物1
2 8 9 15 21 25 28 29 34 36 39 40 42 46
PPIB指导物2
2 7 9 10 23 27 29 32 34 36 37 38 41
0.25
0.20
0.15
0.10
0.05
0.00

图23

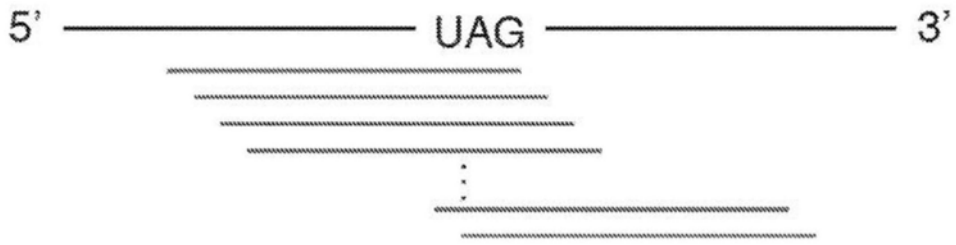

5'
UAG
3'

图24

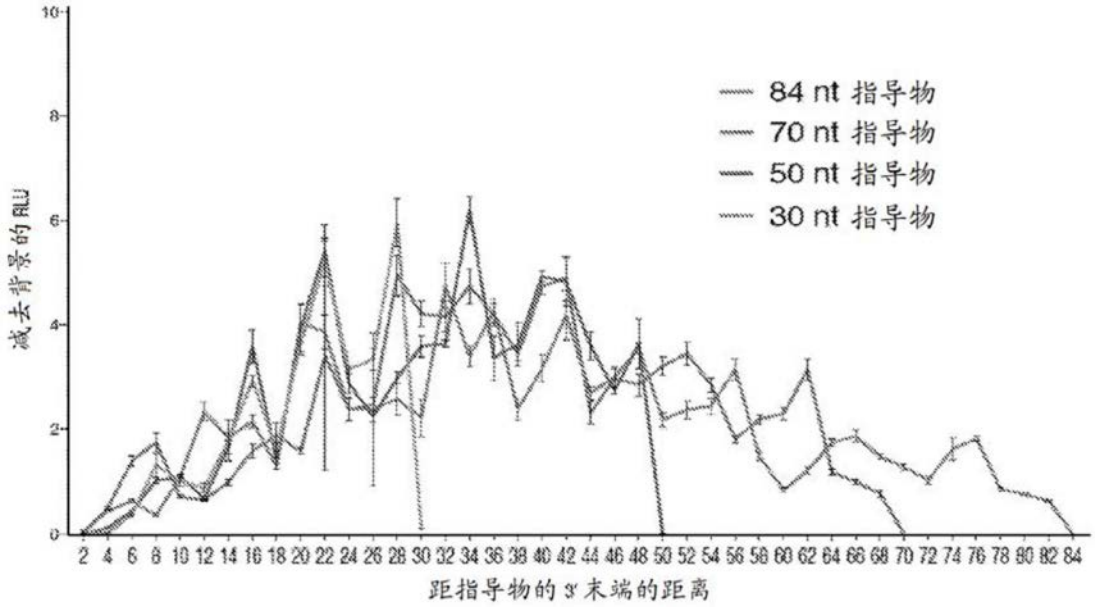

84 nt 指导物
70 nt 指导物
50 nt 指导物
30 nt 指导物
减去背景的RLU
距指导物的3'末端的距离

图25A

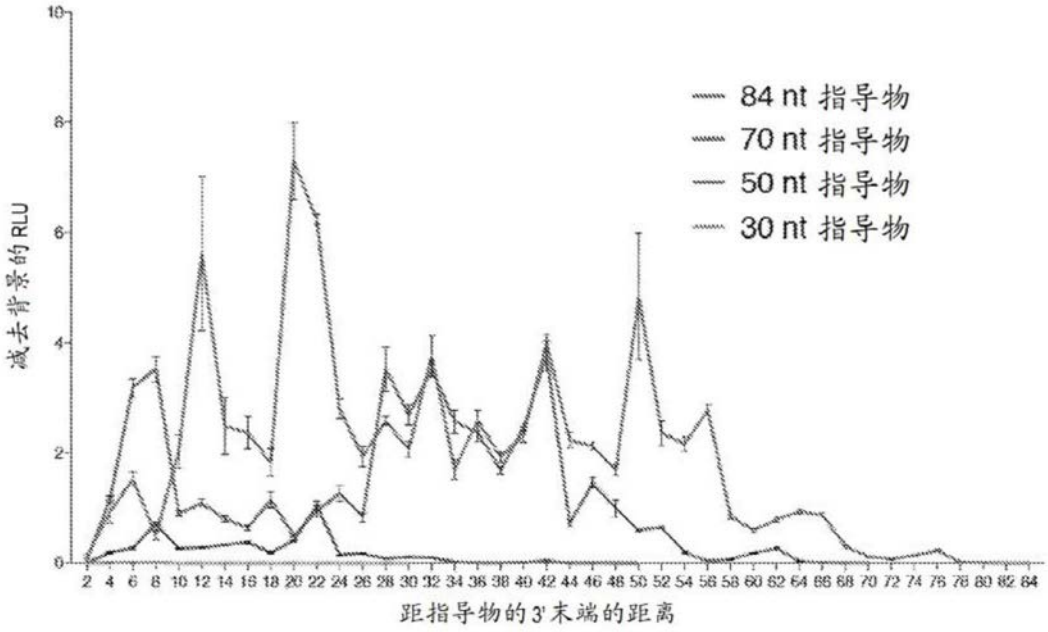

84 nt 指导物
70 nt 指导物
50 nt 指导物
30 nt 指导物
减去背景的RLU
距指导物的3'末端的距离

图25B

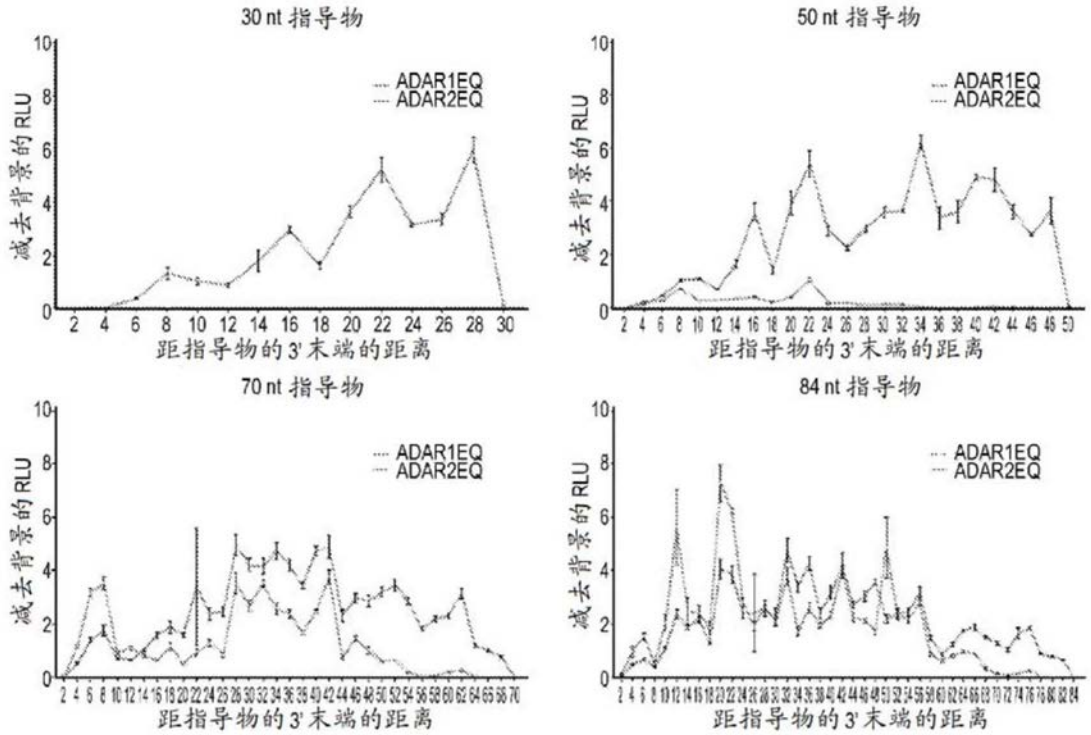

30 nt 指导物
50 nt 指导物
70 nt 指导物
84 nt 指导物
ADAR1EQ
ADAR2EQ
减去背景的 RLU
距指导物的 3' 末端的距离

图25C

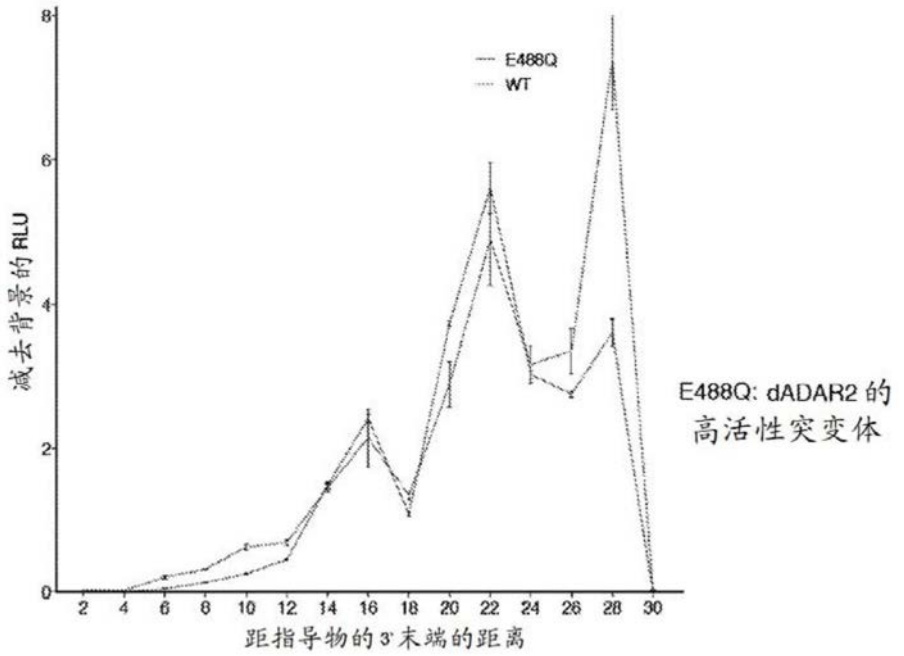

E488Q
WT
减去背景的 RLU
距指导物的 3' 末端的距离
E488Q: dADAR2 的
高活性突变体

图26

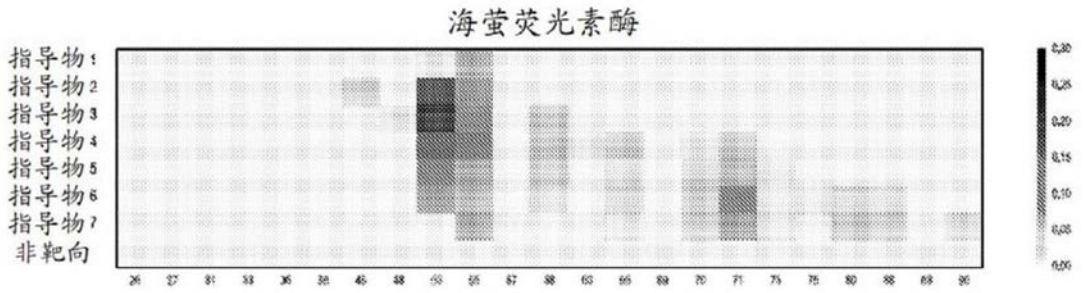
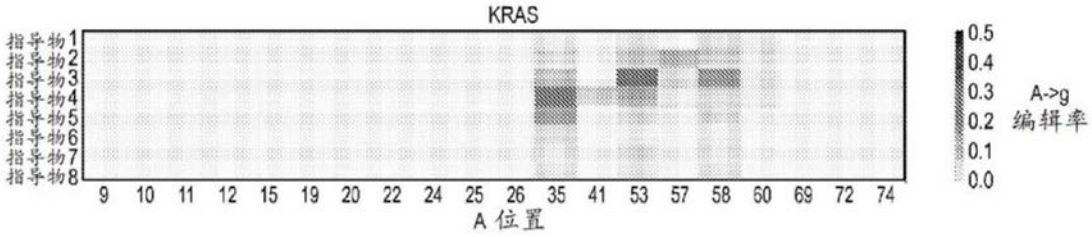


图27A

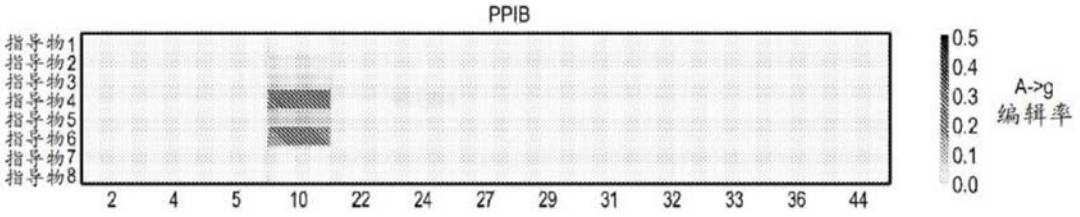
指导物	3' 起始位置	5' 起始位置	A-C 位置
1	1	50	50
2	9	59	50
3	17	67	50
4	25	75	50
5	33	83	50
6	41	91	50
7	49	99	50
NT			

图27B



指导物	指导物中的 A-C 错配位置
1	2
2	10
3	18
4	26
5	34
6	42
7	50
8	非靶向

图28A



指导物	指导物中的A:C错配位置
1	2
2	10
3	18
4	26
5	34
6	42
7	50
8	非靶向

图28B

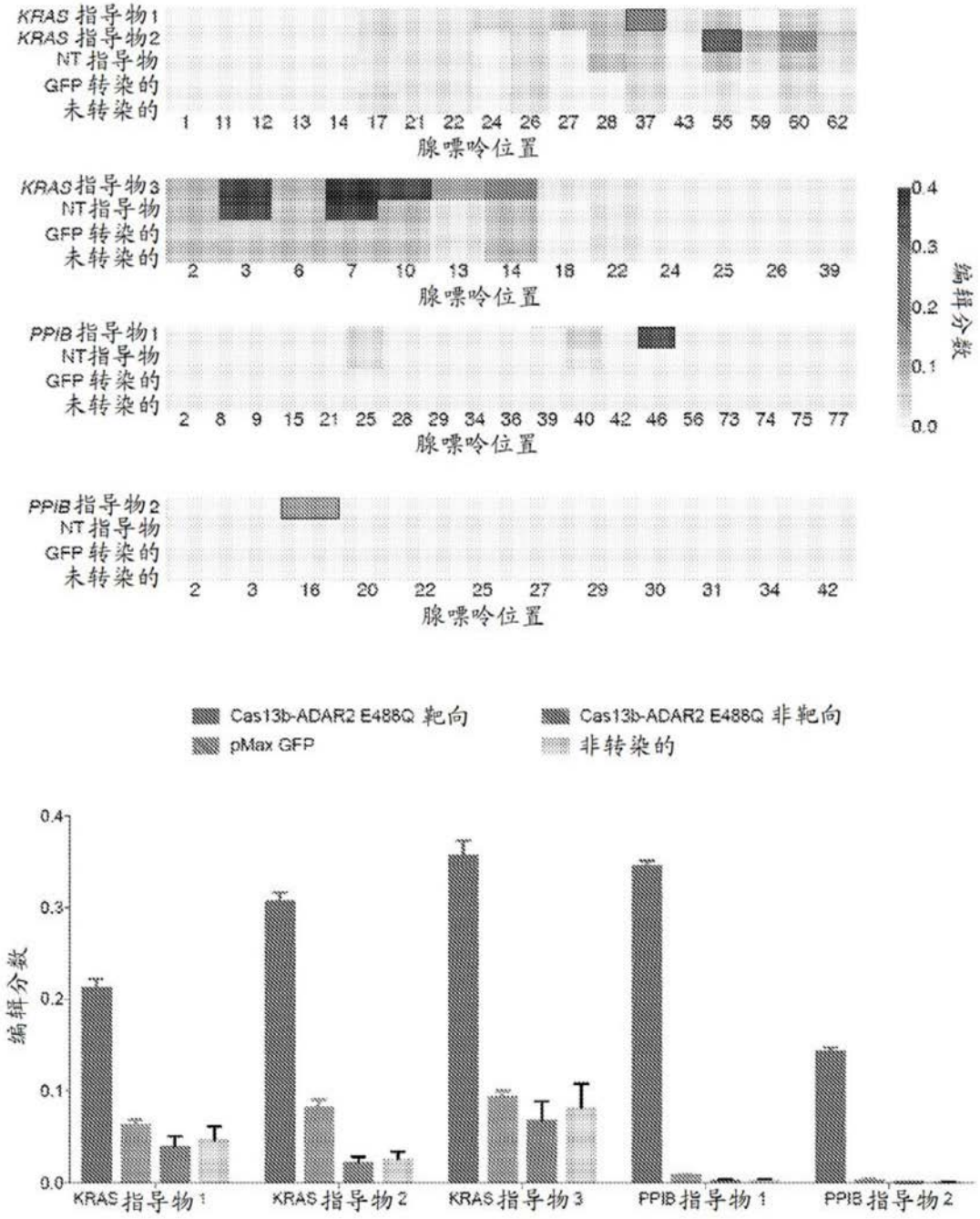

KRAS 指导物1
KRAS 指导物2
NT 指导物
GFP 转染的
未转染的
1 11 12 13 14 17 21 22 24 26 27 28 37 43 55 59 60 62
腺嘌呤位置
KRAS指导物3
NT指导物
GFP 转染的
未转染的
2 3 6 7 10 13 14 18 22 24 25 26 39
腺嘌呤位置
PPIB指导物1
NT指导物
GFP 转染的
未转染的
2 8 9 15 21 25 28 29 34 36 39 40 42 46 56 73 74 75 77
腺嘌呤位置
PPIB指导物2
NT指导物
GFP 转染的
未转染的
2 3 16 20 22 25 27 29 30 31 34 42
腺嘌呤位置
0.4
0.3
0.2
0.1
0.0
编辑分数
Cas13b-ADAR2 E488Q 靶向
Cas13b-ADAR2 E488Q 非靶向
pMax GFP
非转染的
编辑分数
KRAS 指导物 1
KRAS 指导物 2
KRAS 指导物 3
PPIB 指导物 1
PPIB 指导物 2

图29

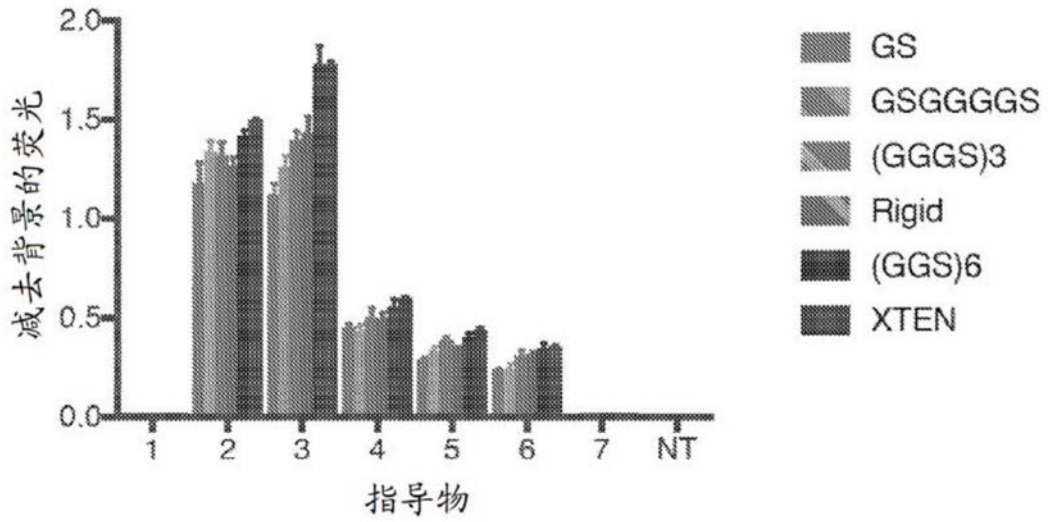

减去背景的荧光
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
NT
指导物
GS
GSGGGGS
(GGGS)3
Rigid
(GGS)6
XTEN

图30

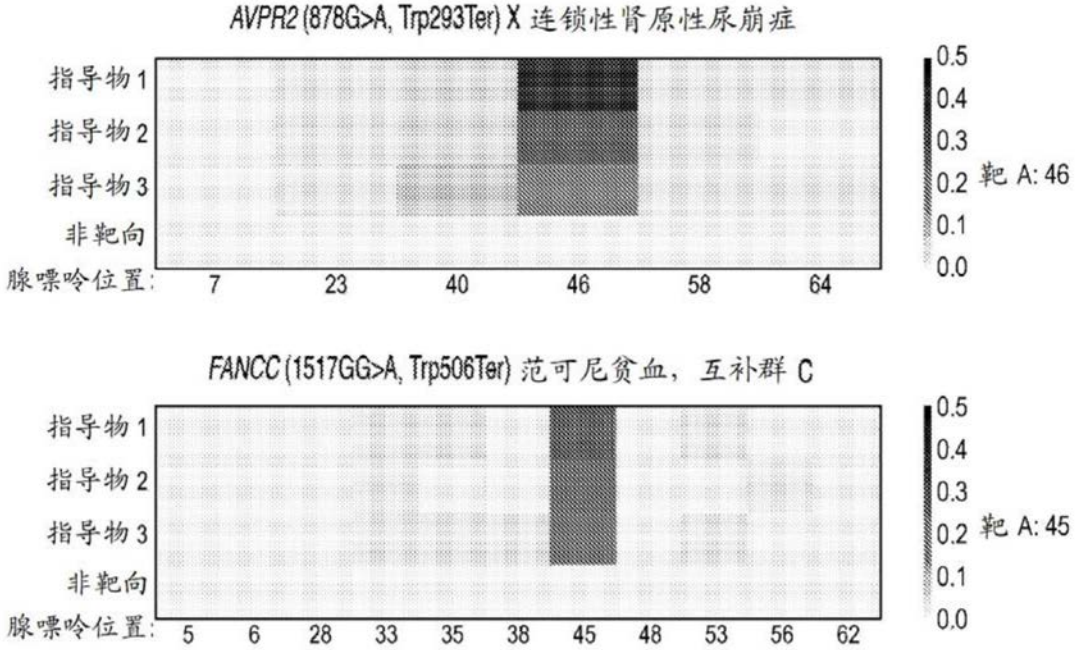

AVPR2 (878G>A, Trp293Ter) X 连锁性肾原性尿崩症
指导物 1
指导物 2
指导物 3
非靶向
腺嘌呤位置:
7
23
40
46
58
64
0.5
0.4
0.3
0.2
0.1
0.0
靶 A: 46
FANCC (1517GG>A, Trp506Ter) 范可尼贫血，互补群 C
指导物 1
指导物 2
指导物 3
非靶向
腺嘌呤位置:
5
6
28
33
35
38
45
48
53
56
62
0.5
0.4
0.3
0.2
0.1
0.0
靶 A: 45

图31

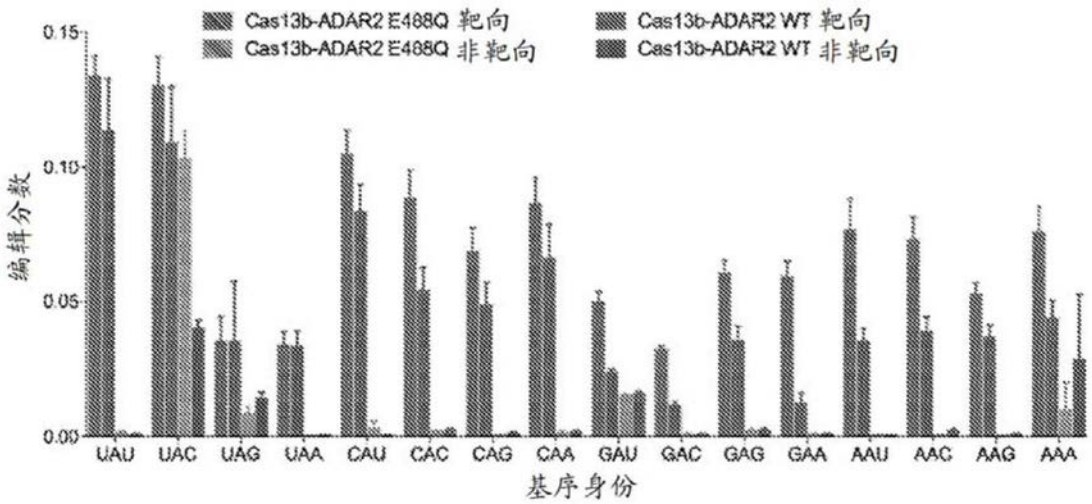
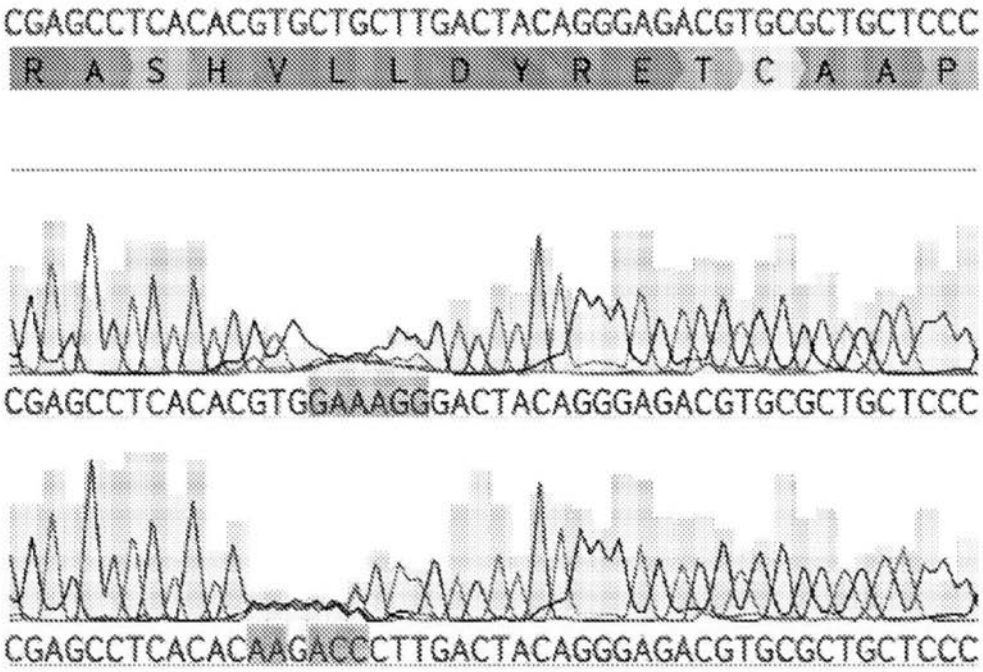

Cas13b-ADAR2 E488Q 靶向
Cas13b-ADAR2 E488Q 非靶向
Cas13b-ADAR2 WT 靶向
Cas13b-ADAR2 WT 非靶向
编辑分数
0.15
0.10
0.05
0.00
UAU UAC UAG UAA CAU CAC CAG CAA GAU GAC GAG GAA AAU AAC AAG AAA
基序身份

图32


CGAGCCTCACACGTGCTGCTTGACTACAGGGAGACGTGCGCTGCTCCC
R A S H V L L D Y R E T C A A P
CGAGCCTCACACGTGGAAAGGGACTACAGGGAGACGTGCGCTGCTCCC
CGAGCCTCACACAAGACCCTTGACTACAGGGAGACGTGCGCTGCTCCC

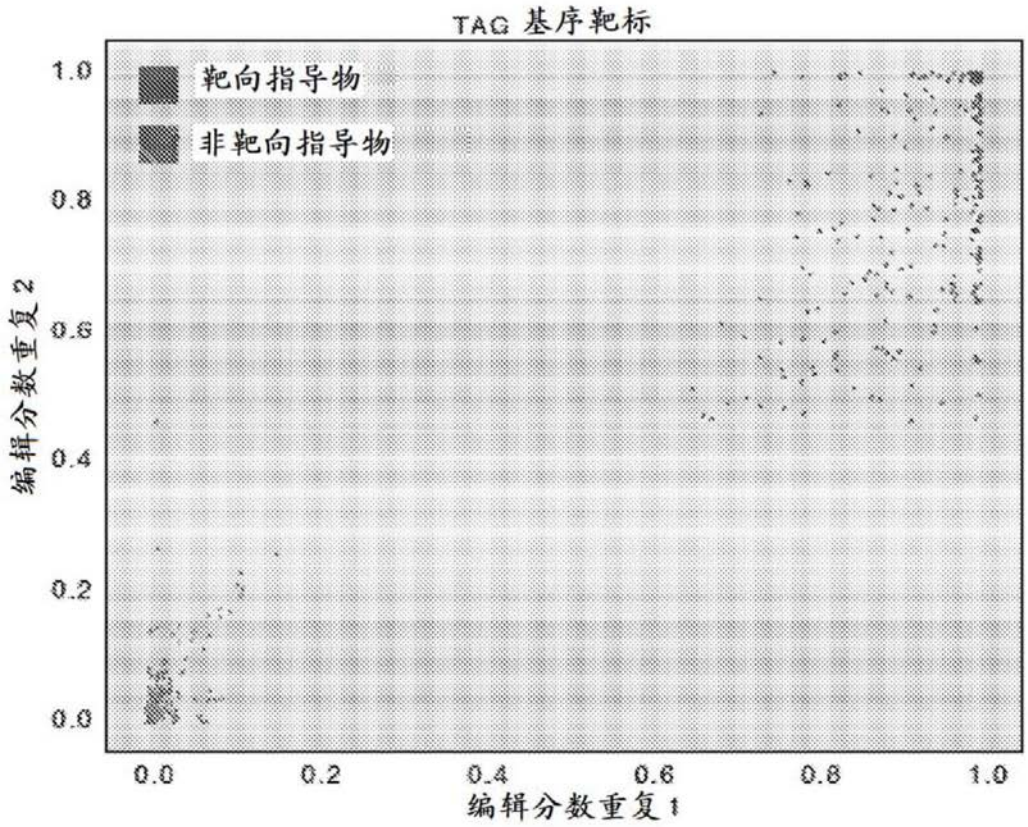

TAG 基序靶标
靶向指导物
非靶向指导物
编辑分数重复 2
编辑分数重复 1
0.0
0.2
0.4
0.6
0.8
1.0

图33

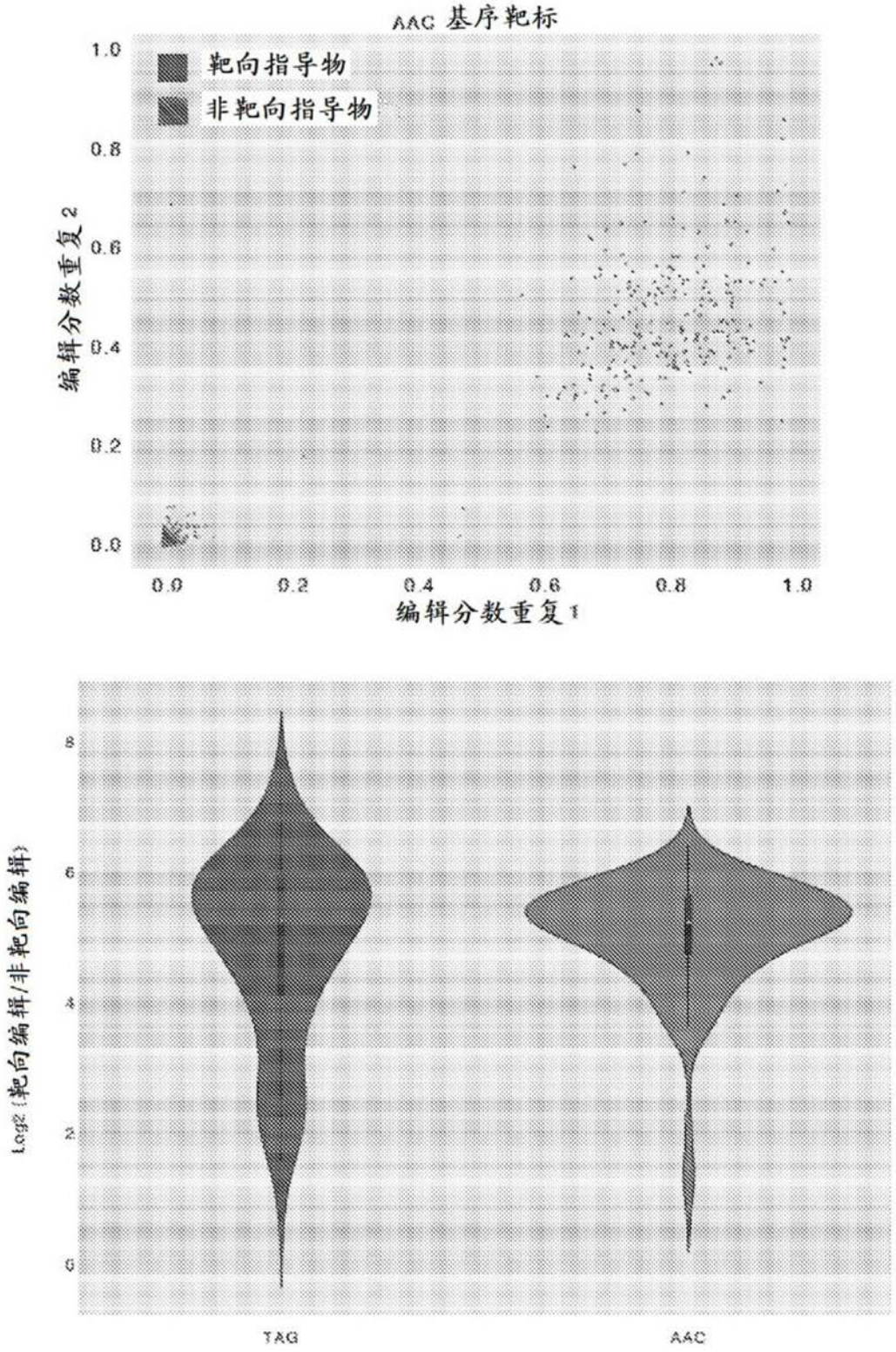

AAC 基序靶标
靶向指导物
非靶向指导物
编辑分数重复 2
编辑分数重复 1
0.0
0.2
0.4
0.6
0.8
1.0
Log2(靶向编辑/非靶向编辑)
TAG
AAC

图33(续)

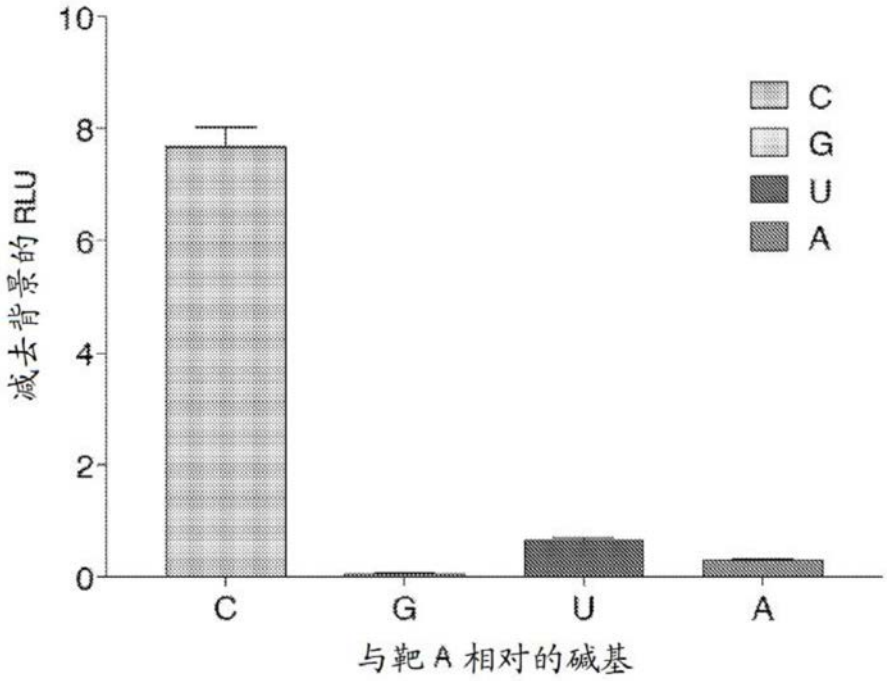

10
8
6
4
2
0
减去背景的 RLU
C
G
U
A
C
G
U
A
与靶 A 相对的碱基

图34

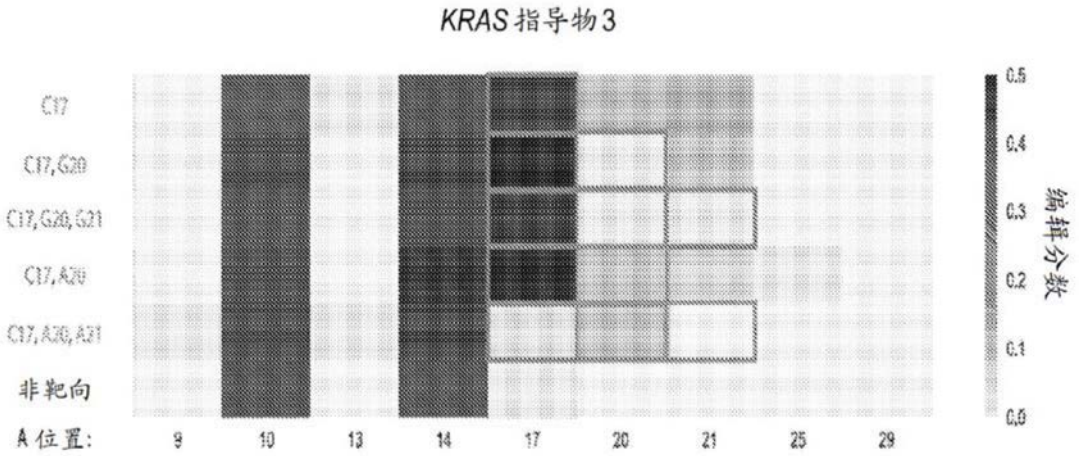

KRAS 指导物 3
C17
C17,G20
C17,G20,G21
C17,A20
C17,A20,A21
非靶向
A 位置:
9
10
13
14
17
20
21
25
29
0.5
0.4
0.3
0.2
0.1
0.0
编辑分数

图35

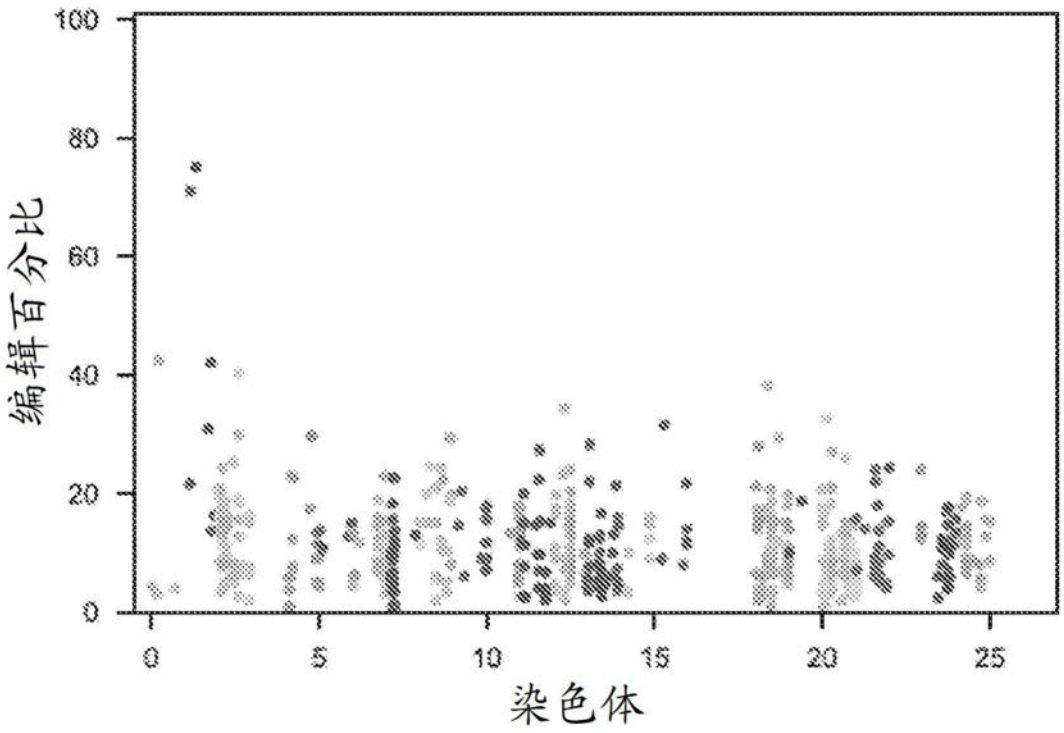

编辑百分比
染色体
0
20
40
60
80
100
0
5
10
15
20
25

图36A

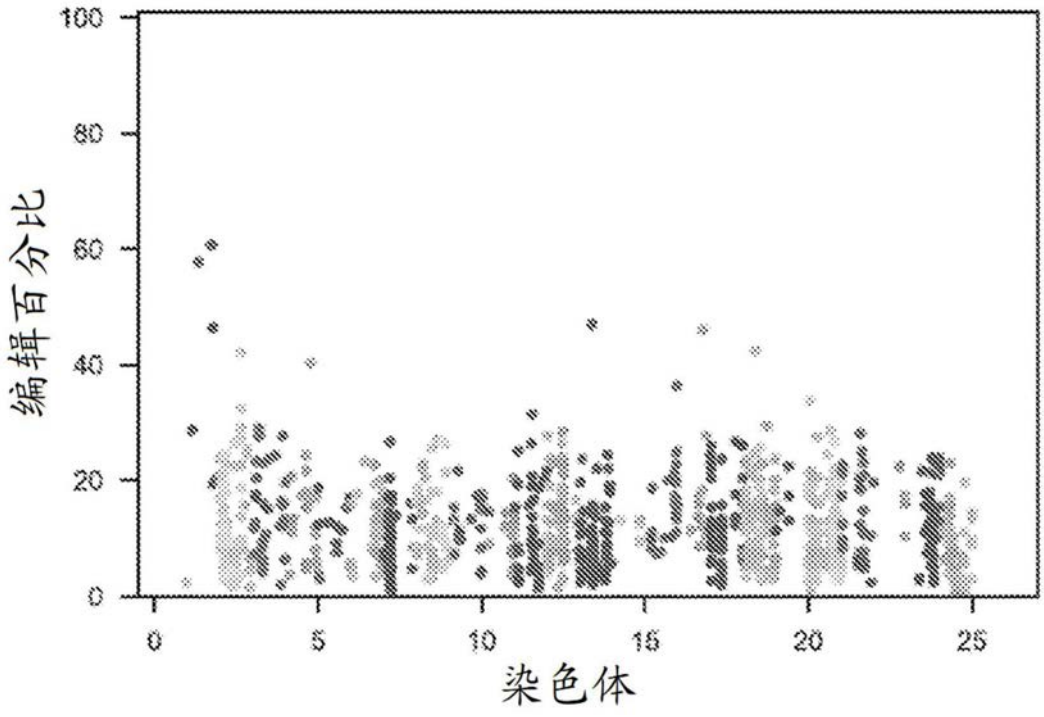

编辑百分比
染色体
0
20
40
60
80
100
0
5
10
15
20
25

图36B

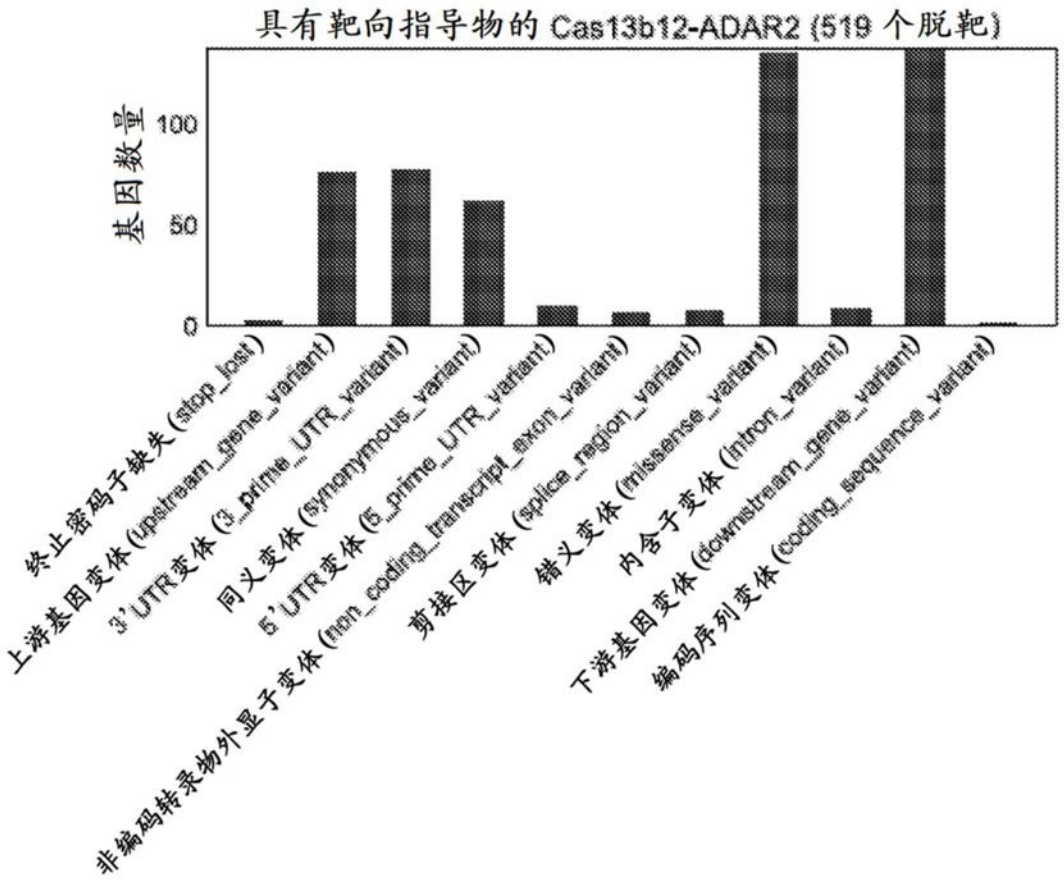

具有靶向指导物的 Cas13b12-ADAR2 (519 个脱靶)
基因数量
0
50
100
终止密码子缺失 (stop_lost)
上游基因变体 (upstream_gene_variant)
3'UTR变体 (3_prime_UTR_variant)
同义变体 (synonymous_variant)
5'UTR变体 (5_prime_UTR_variant)
非编码转录物外显子变体 (non_coding_transcript_exon_variant)
剪接区变体 (splice_region_variant)
错义变体 (missense_variant)
内含子变体 (intron_variant)
下游基因变体 (downstream_gene_variant)
编码序列变体 (coding_sequence_variant)

图37A

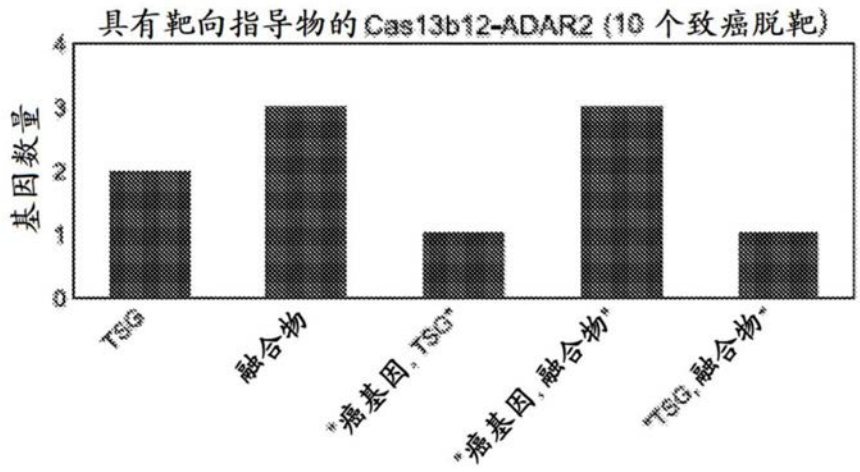

具有靶向指导物的 Cas13b12-ADAR2 (10 个致癌脱靶)
基因数量
0
1
2
3
4
TSG
融合物
癌基因, TSG
癌基因, 融合物
TSG, 融合物

图37B

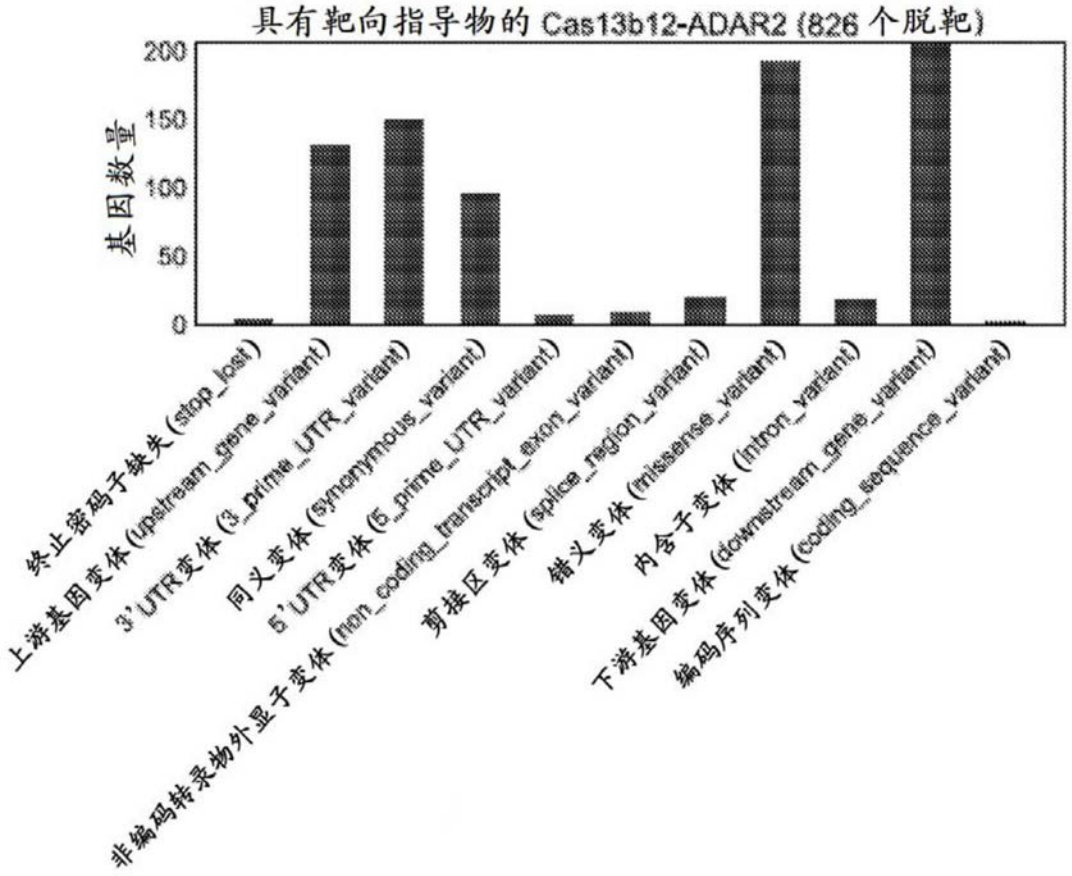

具有靶向指导物的 Cas13b12-ADAR2 (826 个脱靶)
基因数量
0
50
100
150
200
终止密码子缺失 (stop_lost)
上游基因变体 (upstream_gene_variant)
3'UTR变体 (3_prime_UTR_variant)
同义变体 (synonymous_variant)
5'UTR变体 (5_prime_UTR_variant)
非编码转录物外显子变体 (non_coding_transcript_exon_variant)
剪接区变体 (splice_region_variant)
错义变体 (missense_variant)
内含子变体 (intron_variant)
下游基因变体 (downstream_gene_variant)
编码序列变体 (coding_sequence_variant)

图37C

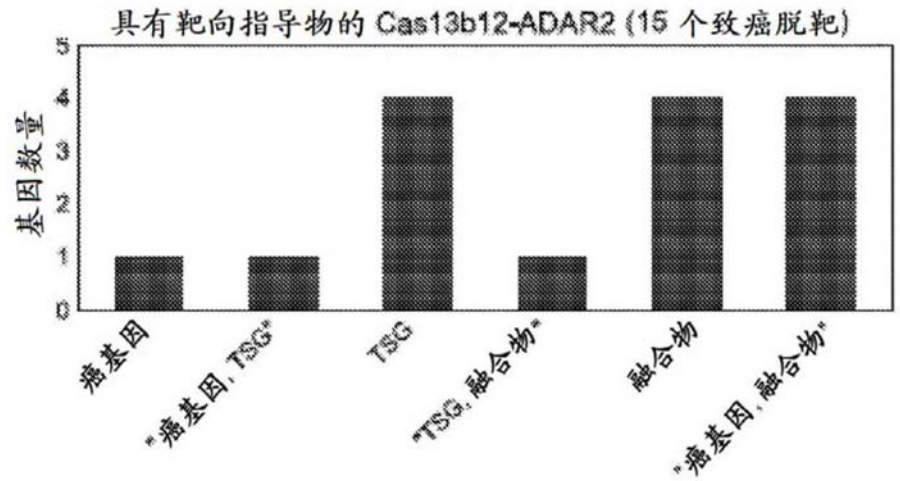

具有靶向指导物的 Cas13b12-ADAR2 (15 个致癌脱靶)
基因数量
0
1
2
3
4
5
癌基因
"癌基因, TSG"
TSG
"TSG, 融合物"
融合物
"癌基因, 融合物"

图37D

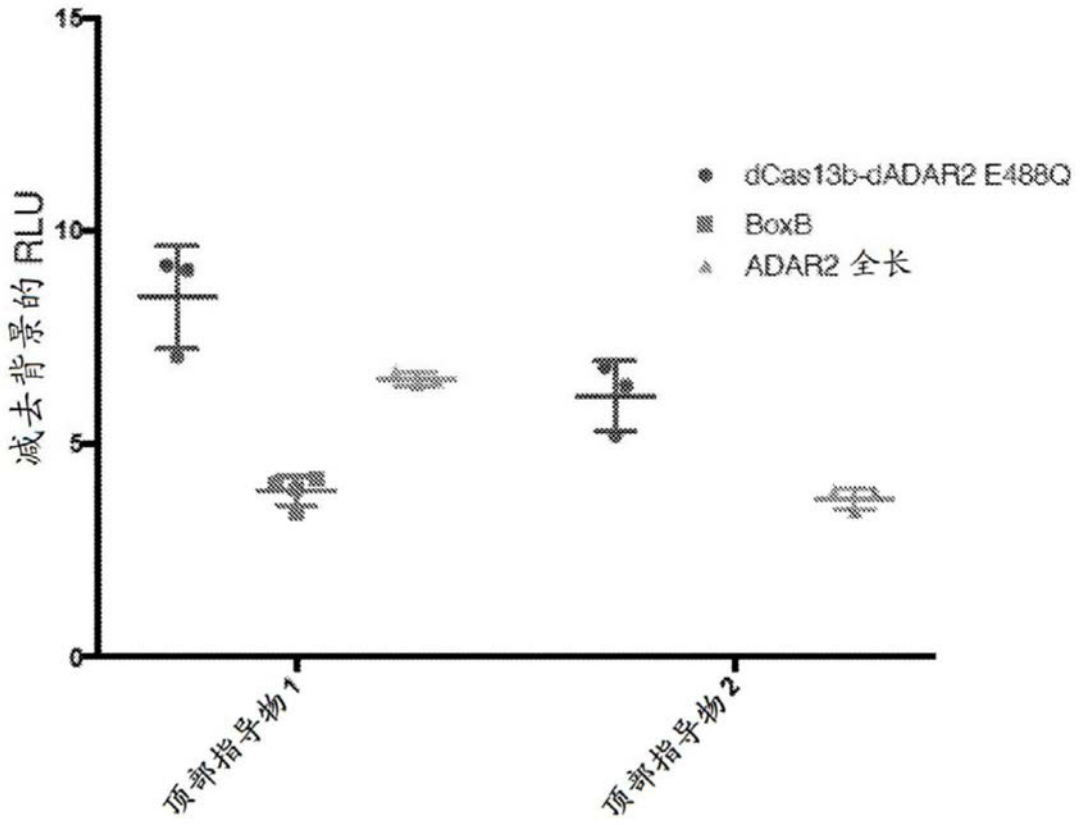

减去背景的RLU
15
10
5
0
dCas13b-dADAR2 E488Q
BoxB
ADAR2 全长
顶部指导物1
顶部指导物2

图38

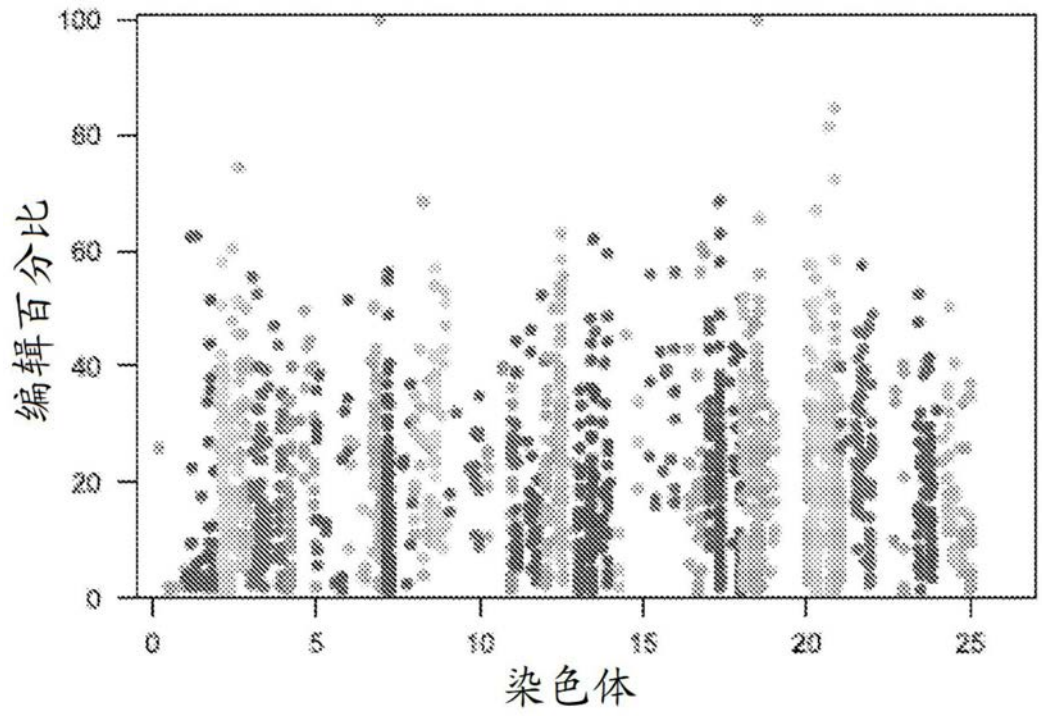

编辑百分比
100
80
60
40
20
0
0
5
10
15
20
25
染色体

图39A

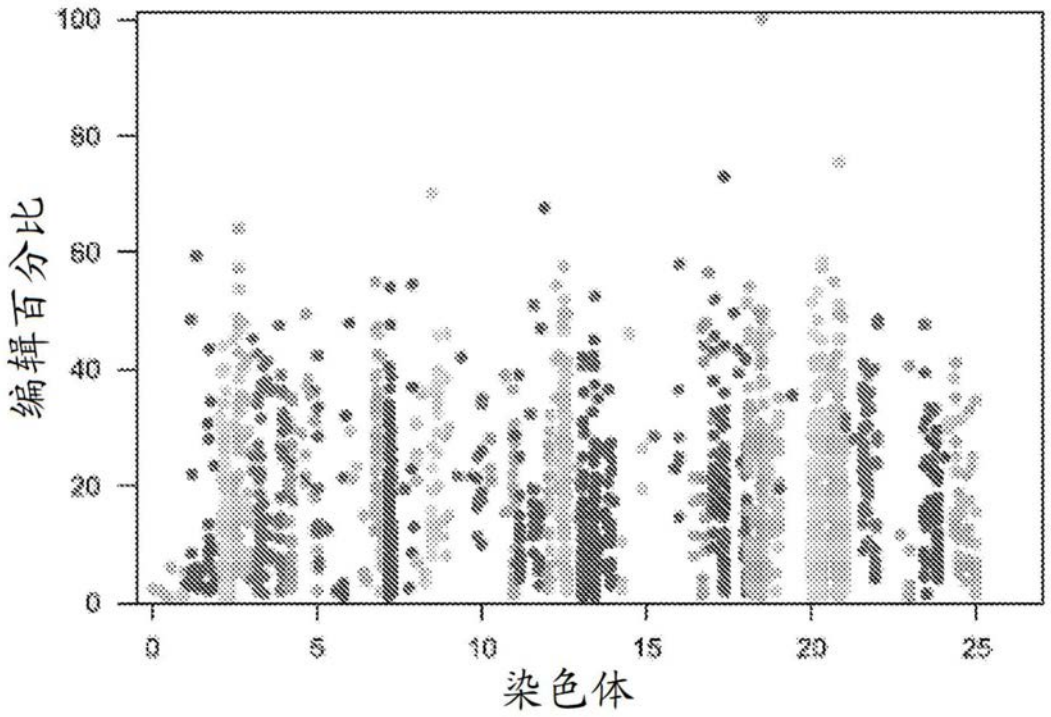

100
80
60
40
20
0
编辑百分比
0
5
10
15
20
25
染色体

图39B

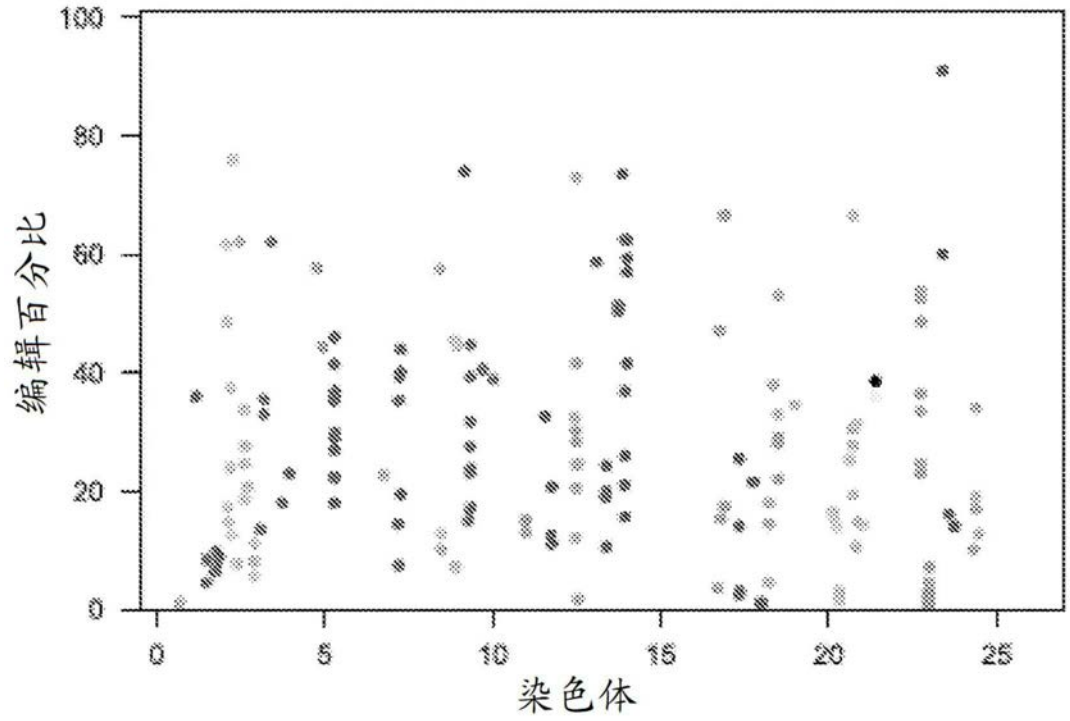

100
80
60
40
20
0
编辑百分比
0
5
10
15
20
25
染色体

图39C

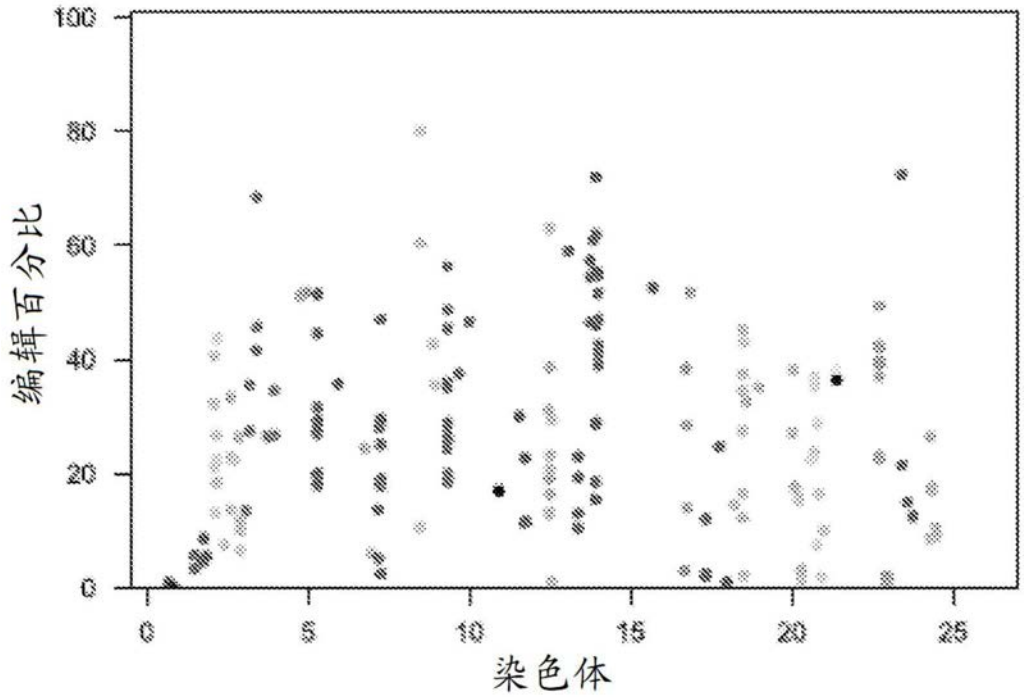


图39D

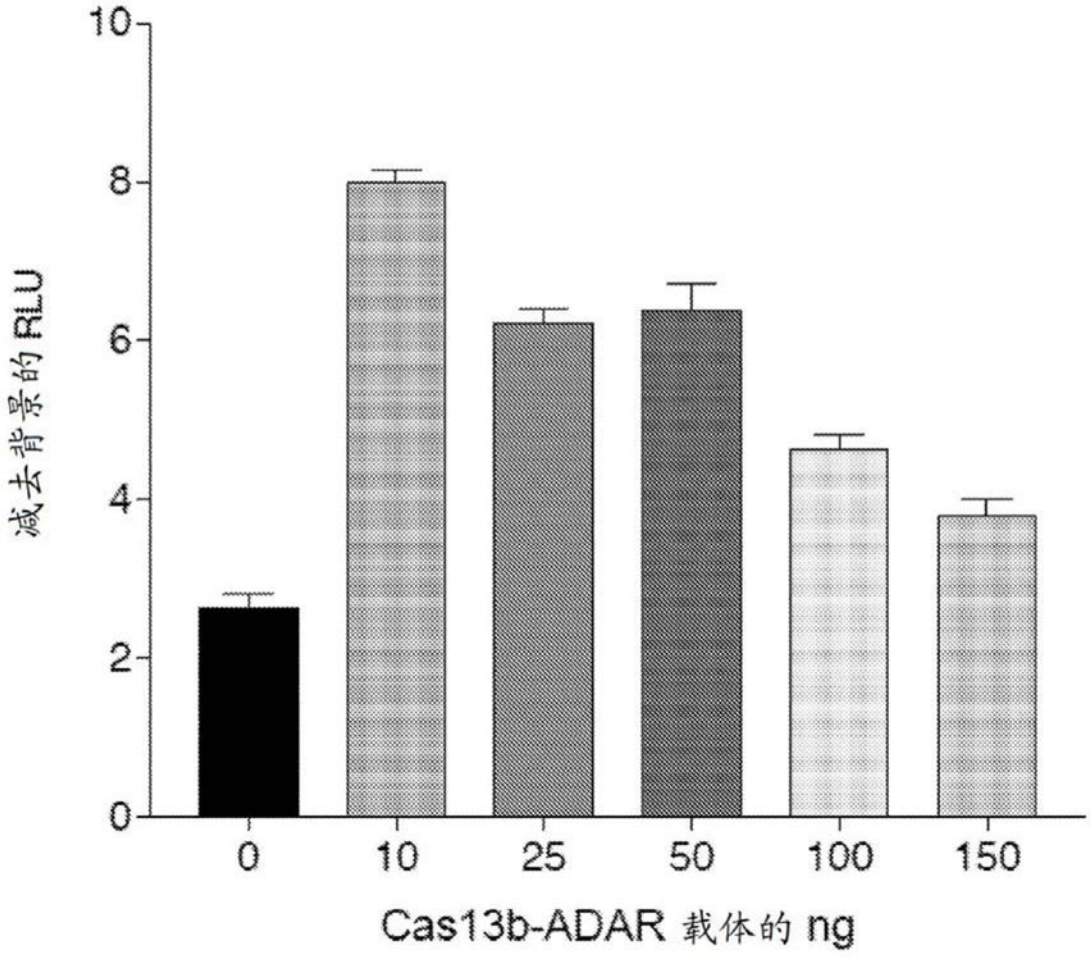


图40

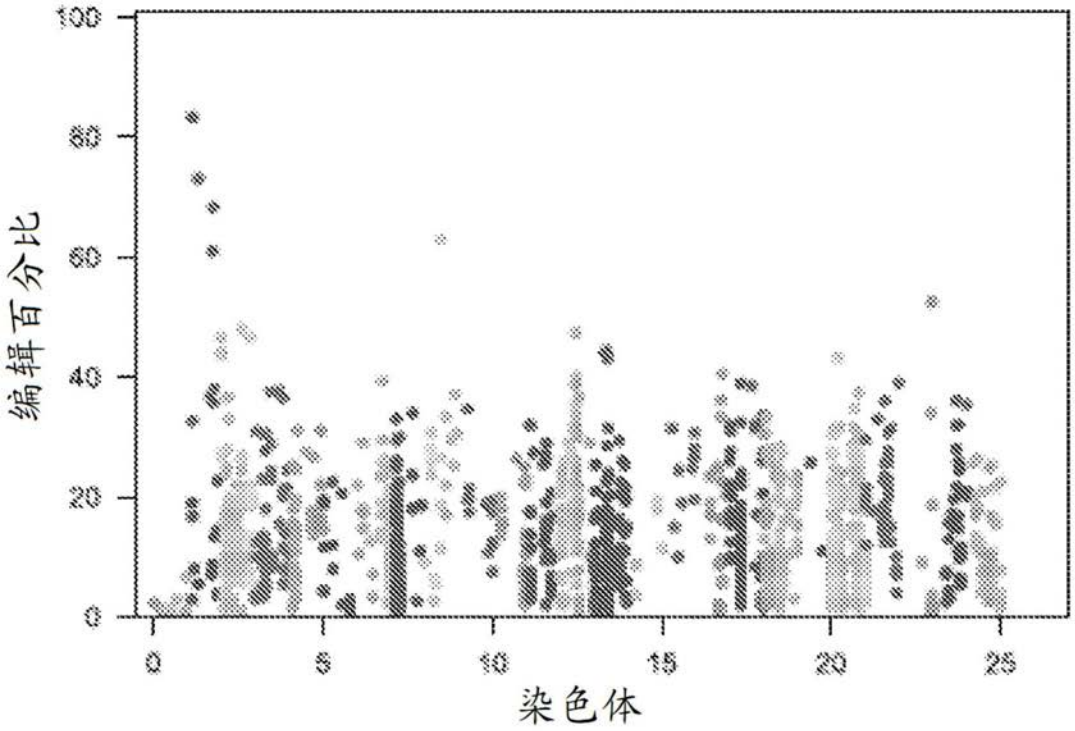

编辑百分比
染色体
0
20
40
60
80
100
0
5
10
15
20
25

图 41A

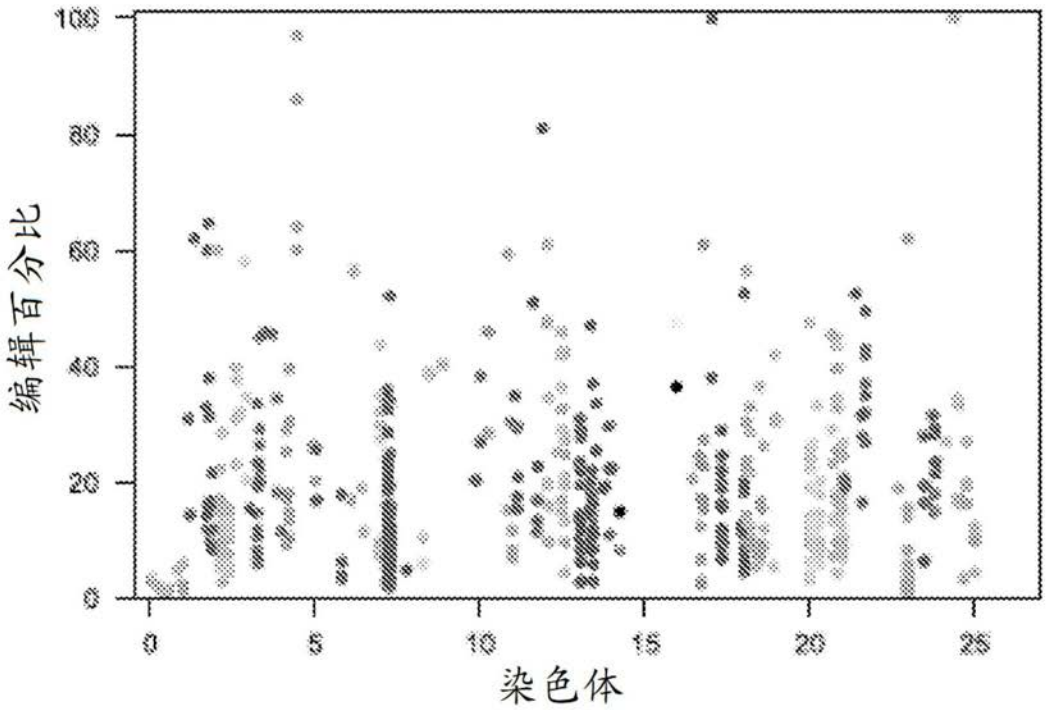

编辑百分比
染色体
0
20
40
60
80
100
0
5
10
15
20
25

图 41B

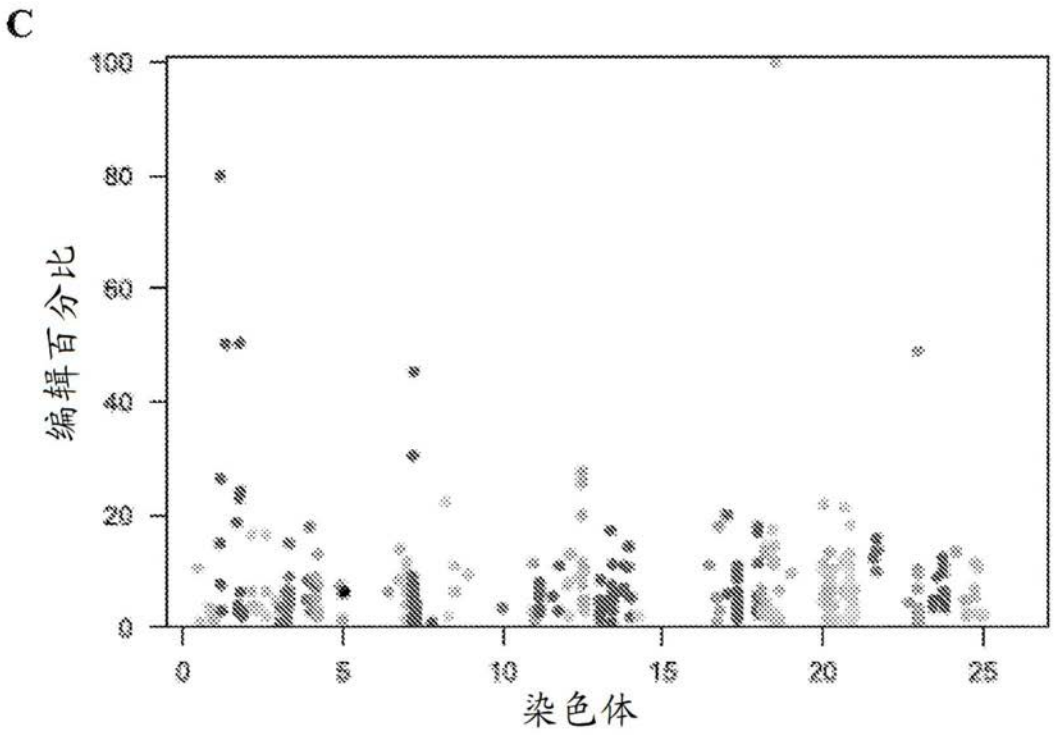


图 41C

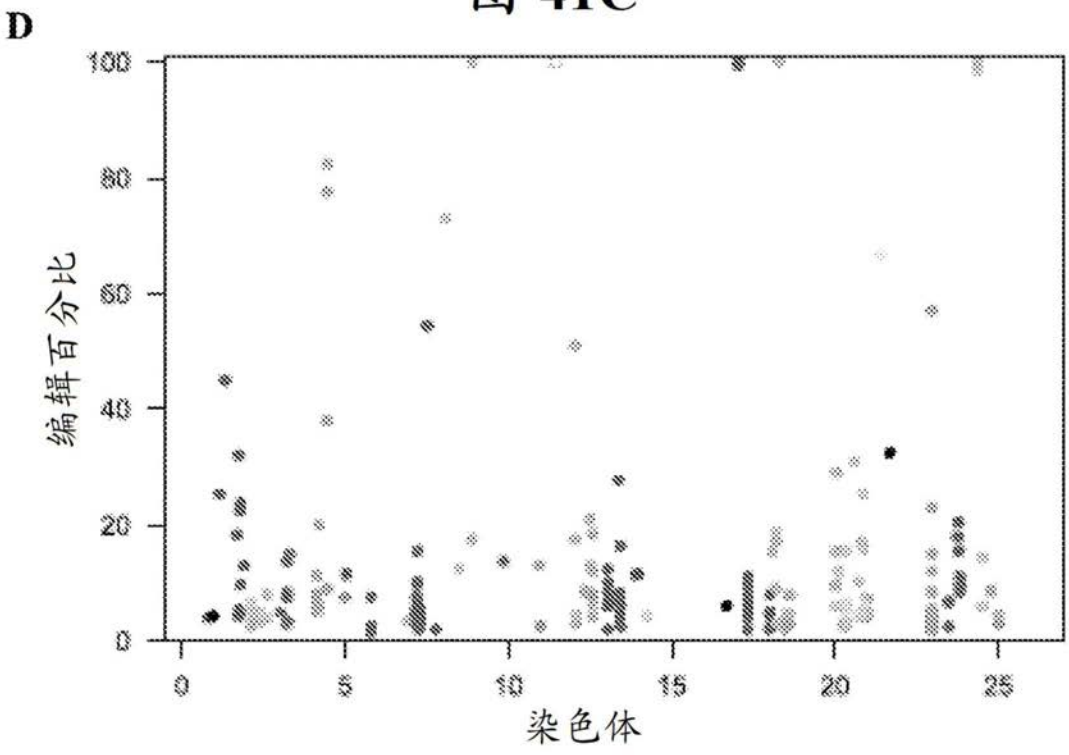


图 41D

图41(续)

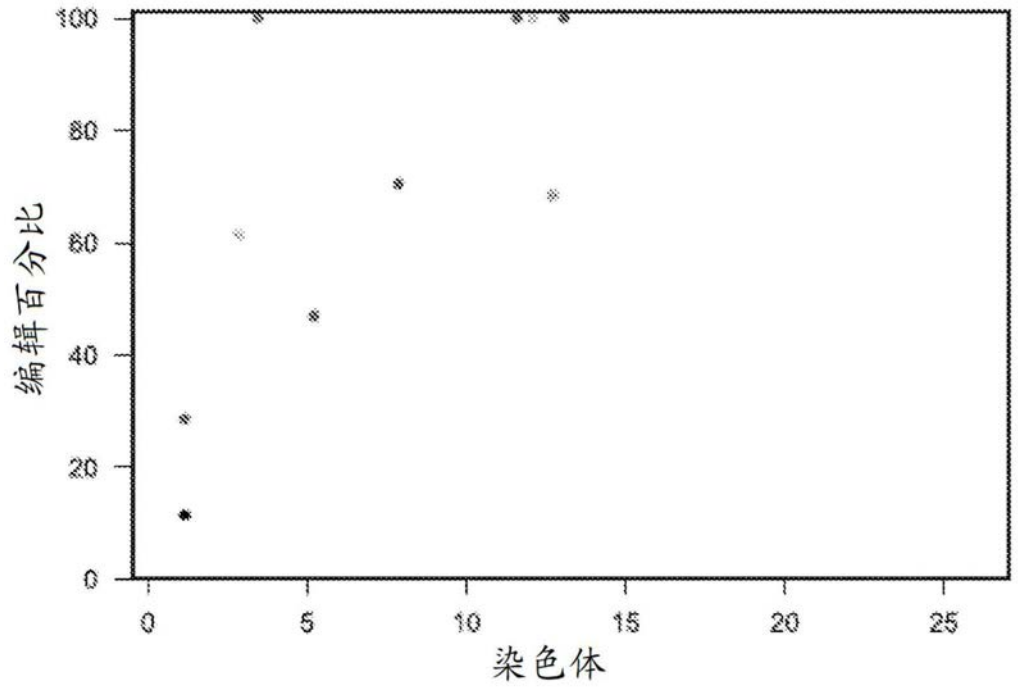

100
80
60
40
20
0
编辑百分比
0
5
10
15
20
25
染色体

图42A

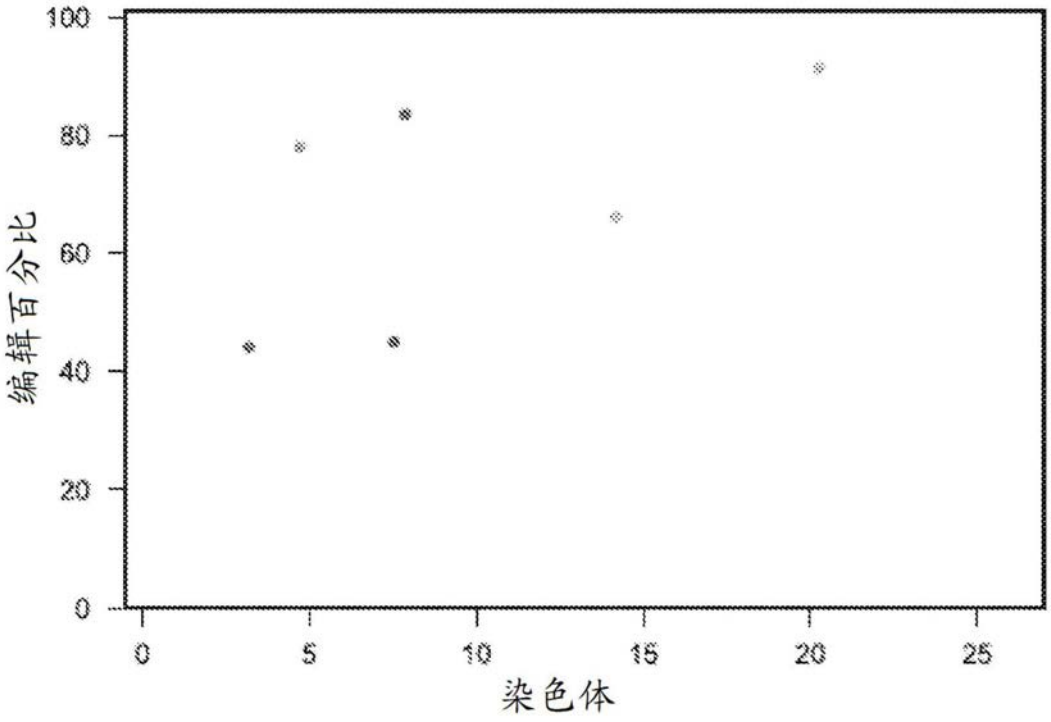

100
80
60
40
20
0
编辑百分比
0
5
10
15
20
25
染色体

图42B

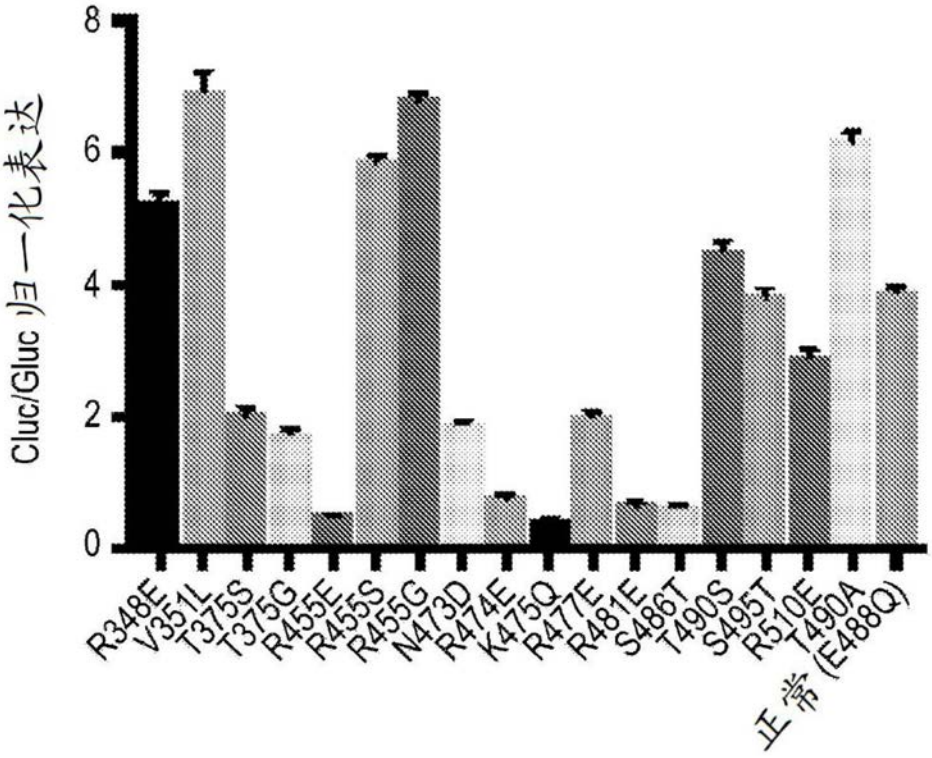

Cluc/Gluc 归一化表达
8
6
4
2
0
R348E
V351L
T375S
T375G
R455E
R455S
R455G
N473D
R474E
K475Q
R477E
R481E
S486T
T490S
S495T
R510E
T490A
正常 (E488Q)

图43A

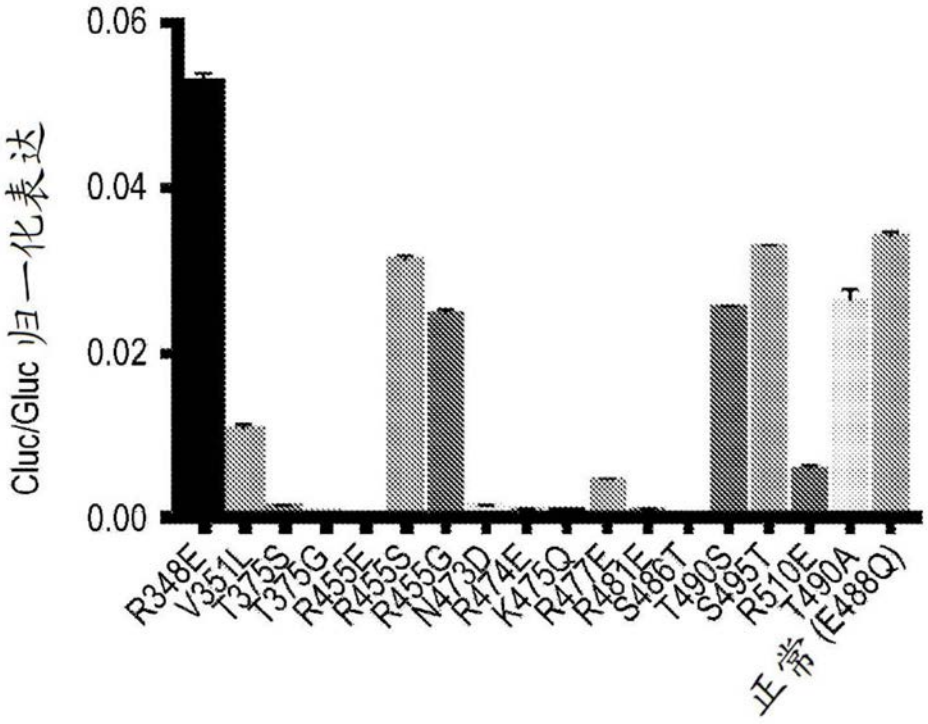

Cluc/Gluc 归一化表达
0.06
0.04
0.02
0.00
R348E
V351L
T375S
T375G
R455E
R455S
R455G
N473D
R474E
K475Q
R477E
R481E
S486T
T490S
S495T
R510E
T490A
正常 (E488Q)

图43B

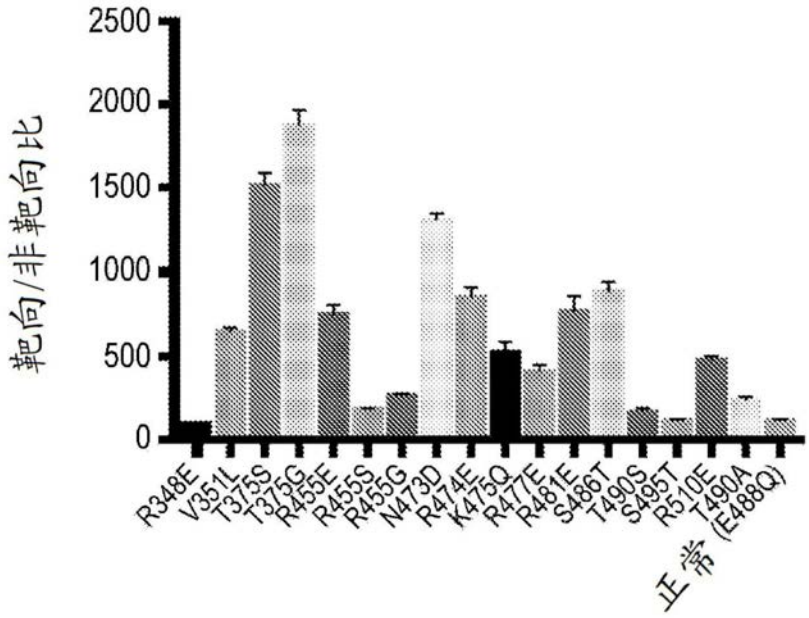

2500
2000
1500
1000
500
0
靶向/非靶向比
R348E
V351L
T375S
T375G
R455E
R455S
R455G
N473D
R474E
K475Q
R477E
R481E
S486T
T490S
S495T
R510E
T490A
正常 (E488Q)

图43C

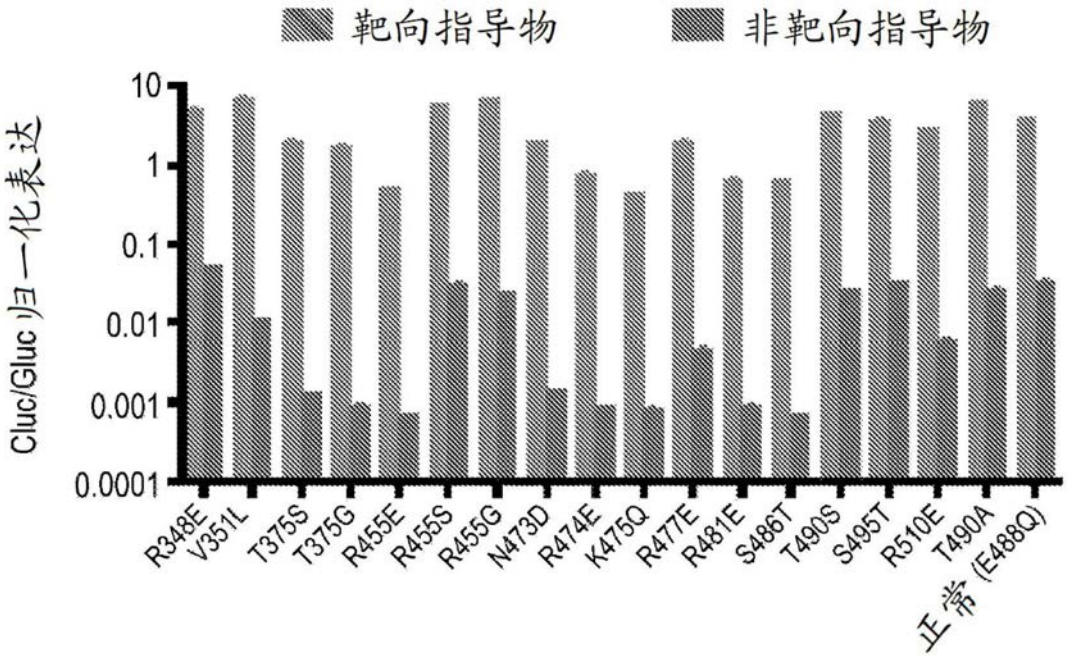

靶向指导物
非靶向指导物
10
1
0.1
0.01
0.001
0.0001
Cluc/Gluc 归一化表达
R348E
V351L
T375S
T375G
R455E
R455S
R455G
N473D
R474E
K475Q
R477E
R481E
S486T
T490S
S495T
R510E
T490A
正常 (E488Q)

图43D

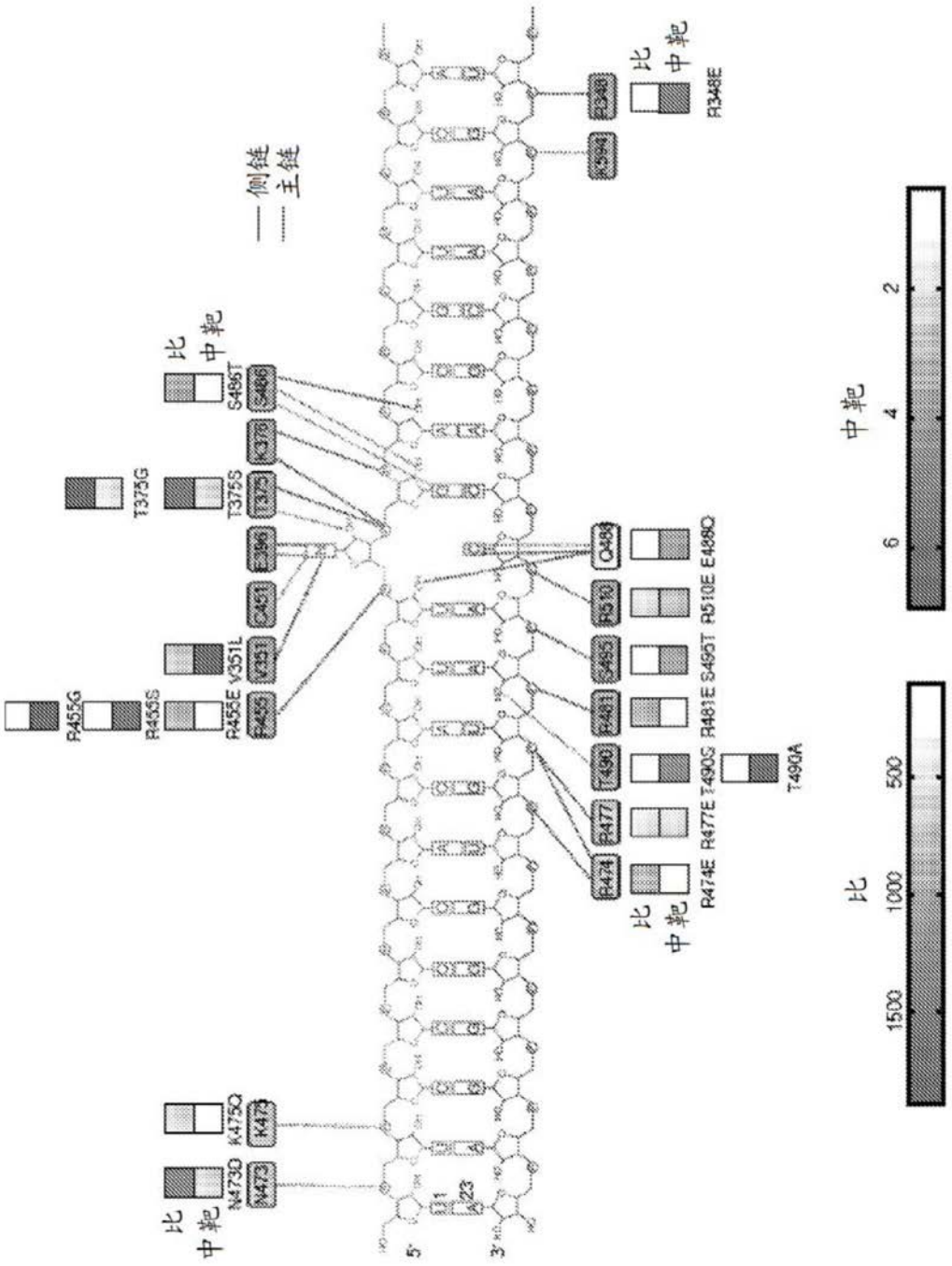

侧链
主链
比
中靶
R348E
S495T
T375S
T375G
R455G
R455S
R455E V351L
K475Q
N473D
R474E R477E T490S R481E S495T R510E E488Q
T490A
5'
3'
1
23
比
中靶
2
4
6
500
1000
1500

图44

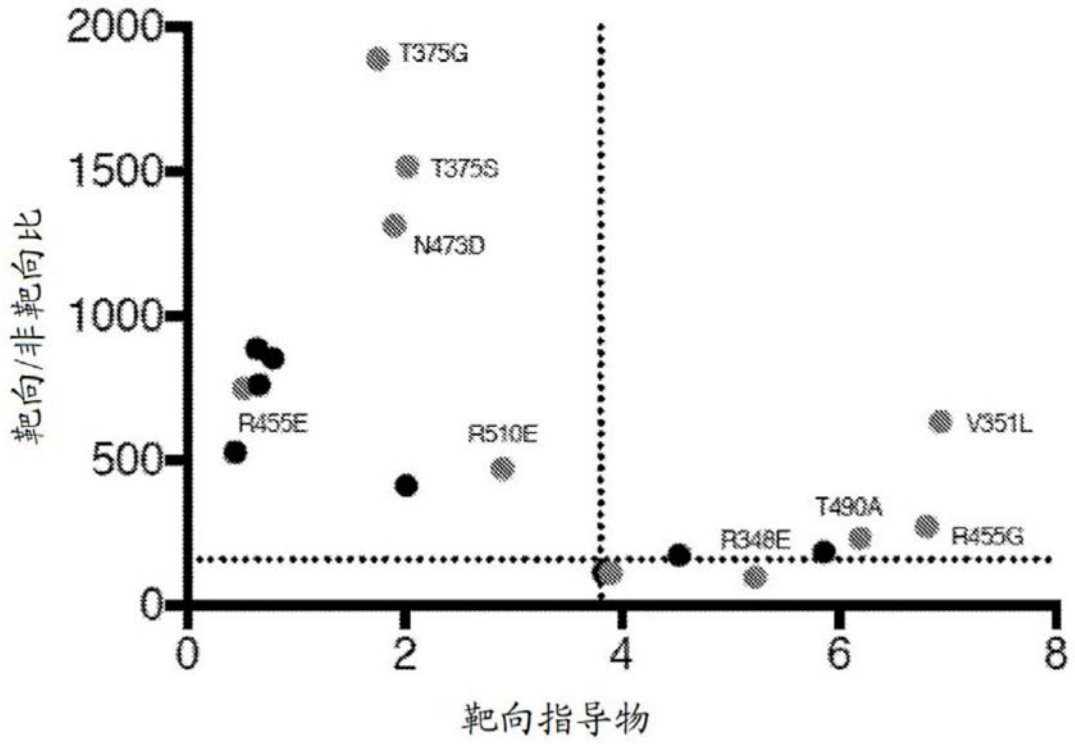

2000
1500
1000
500
0
0
2
4
6
8
靶向/非靶向比
靶向指导物
T375G
T375S
N473D
R455E
R510E
V351L
T490A
R348E
R455G

图45

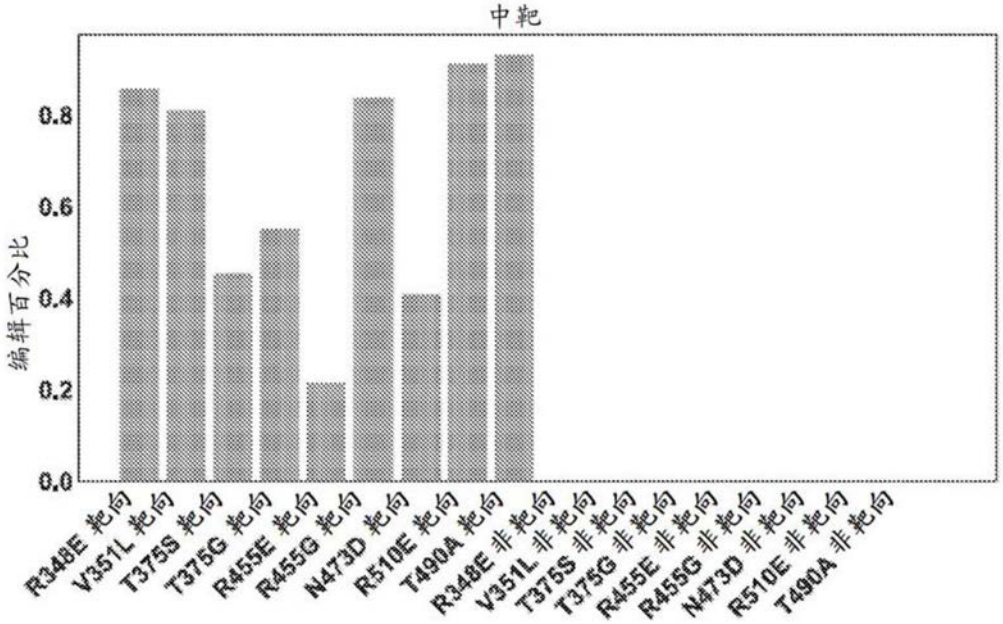

中靶
0.8
0.6
0.4
0.2
0.0
编辑百分比
R348E 靶向
V351L 靶向
T375S 靶向
T375G 靶向
R455E 靶向
R455G 靶向
N473D 靶向
R510E 靶向
T490A 靶向
R348E 非靶向
V351L 非靶向
T375S 非靶向
T375G 非靶向
R455E 非靶向
R455G 非靶向
N473D 非靶向
R510E 非靶向
T490A 非靶向

图46A

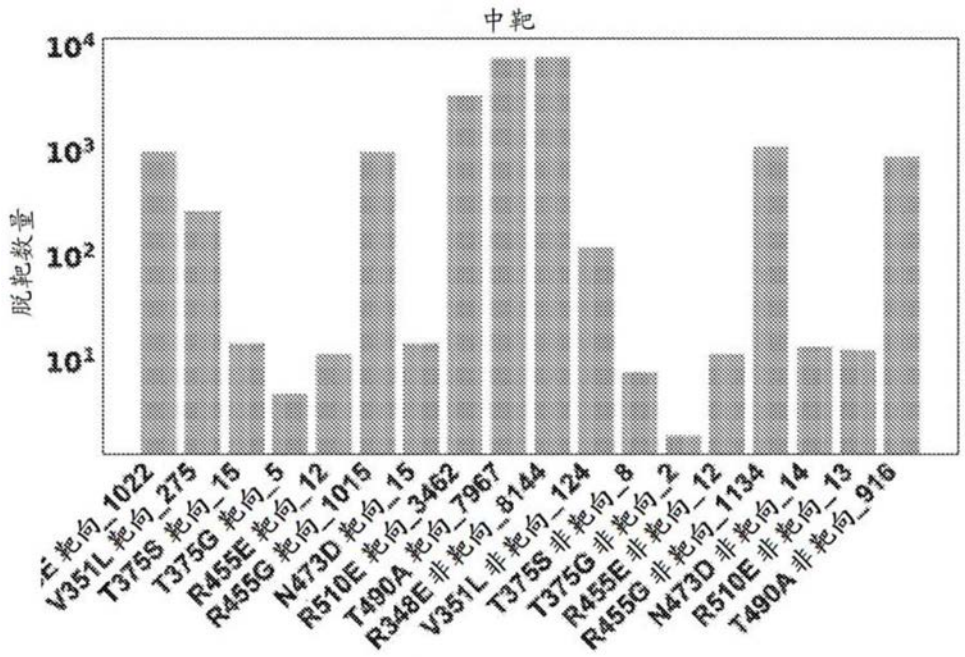

中靶
脱靶数量
10^4
10^3
10^2
10^1
V351L 脱向_1022
T375S 脱向_275
T375G 脱向_15
R455E 脱向_5
R455G 脱向_12
N473D 脱向_1015
R510E 脱向_15
T490A 脱向_3462
R348E 非脱向_7967
V351L 非脱向_8144
T375S 非脱向_124
T375G 非脱向_8
R455E 非脱向_2
R455G 非脱向_12
N473D 非脱向_1134
R510E 非脱向_14
T490A 非脱向_13
非脱向_916

图46B

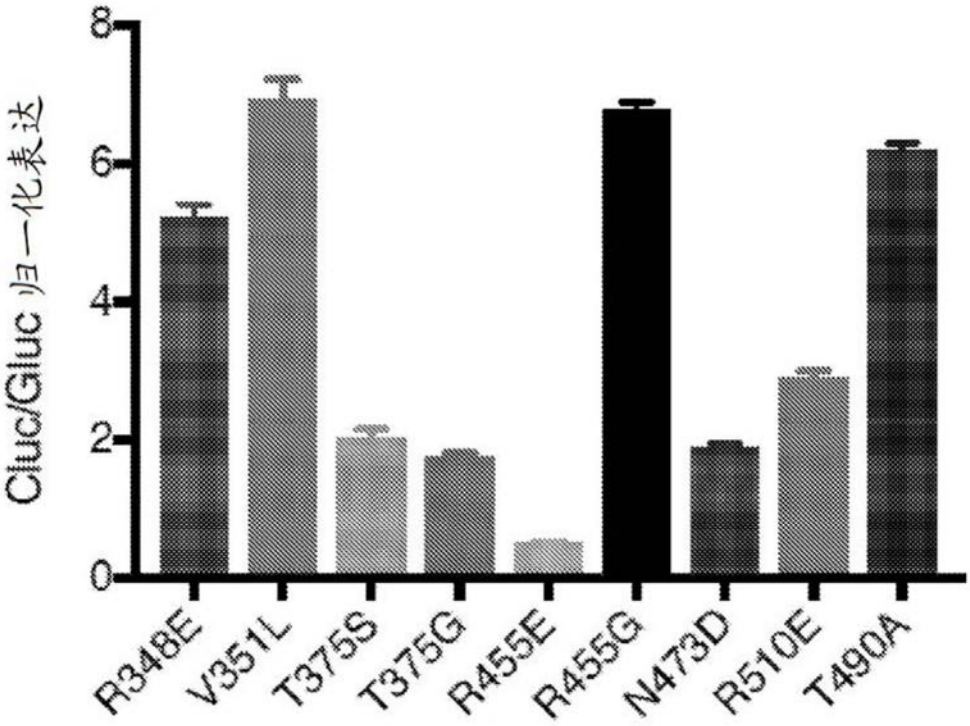

Cluc/Gluc 归一化表达
8
6
4
2
0
R348E
V351L
T375S
T375G
R455E
R455G
N473D
R510E
T490A

图47A

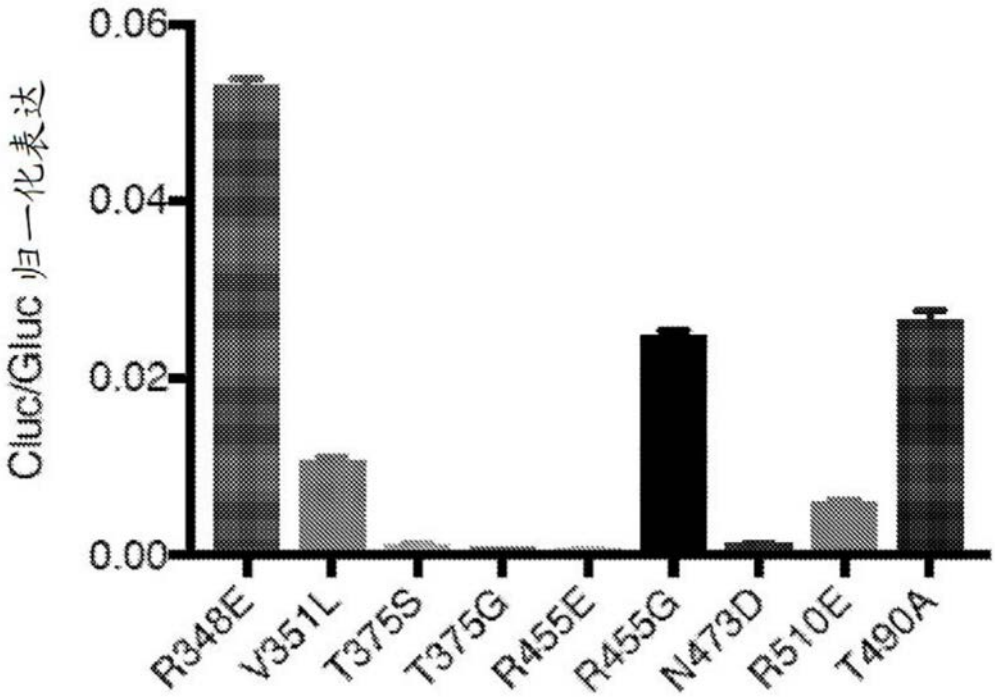

Cluc/Gluc 归一化表达
0.06
0.04
0.02
0.00
R348E
V351L
T375S
T375G
R455E
R455G
N473D
R510E
T490A

图47B

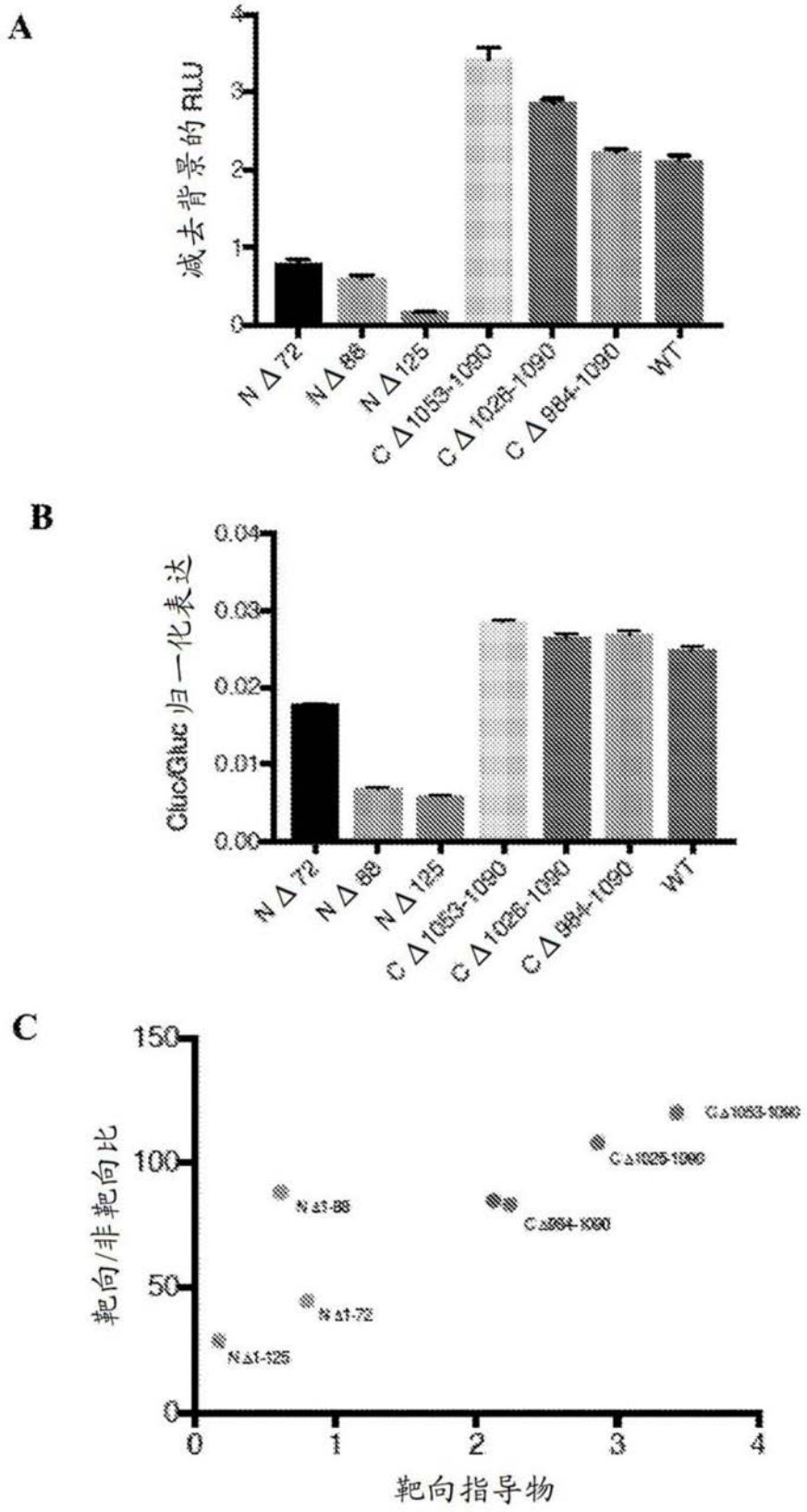

A
减去背景的RLU
N Δ 72
N Δ 88
N Δ 125
C Δ 1053-1090
C Δ 1026-1090
C Δ 984-1090
WT
B
Gluc/Gluc 归一化表达
0.00
0.01
0.02
0.03
0.04
C
靶向/非靶向比
靶向指导物
0
50
100
150
0
1
2
3
4

图48

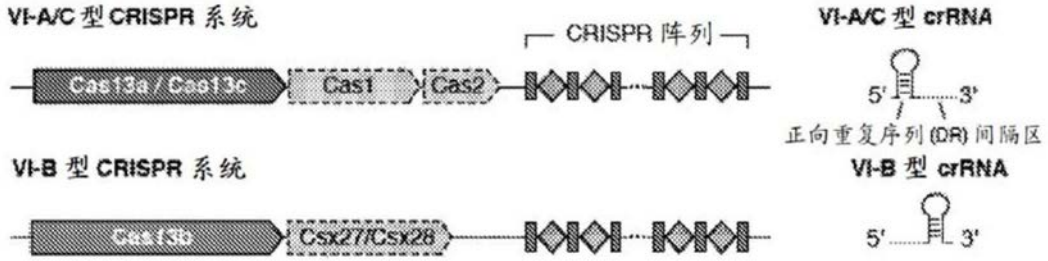


图49A

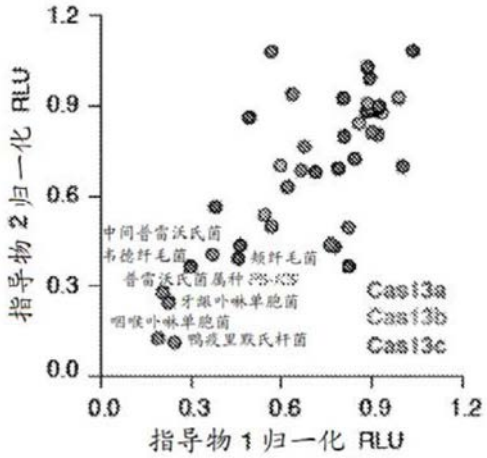


图49B

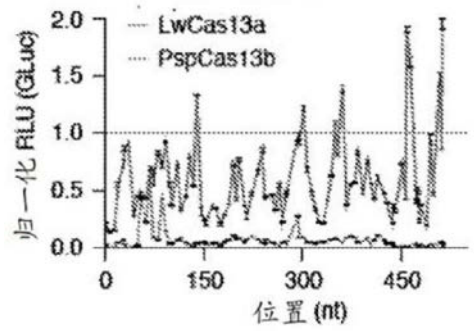


图49C

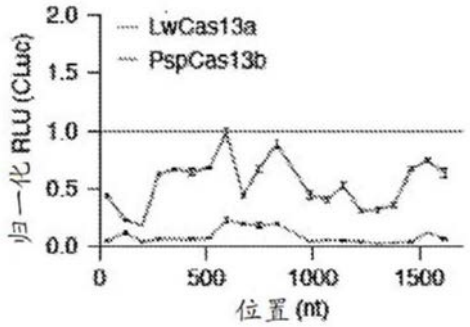


图49D

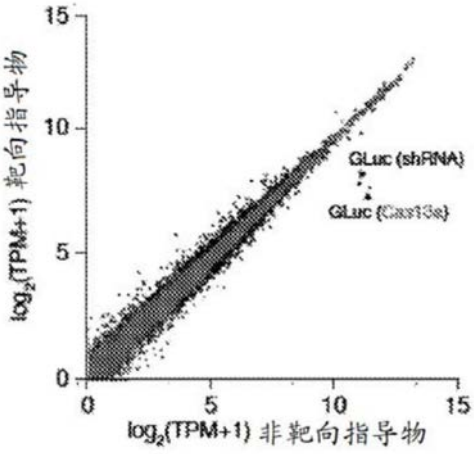


图49E

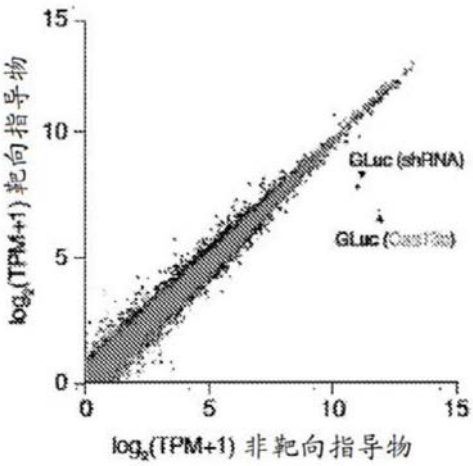


图49F

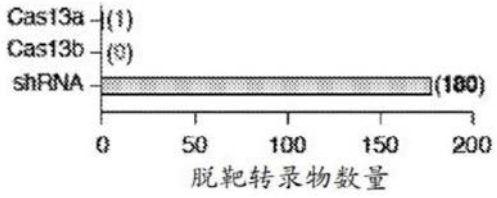


图49G

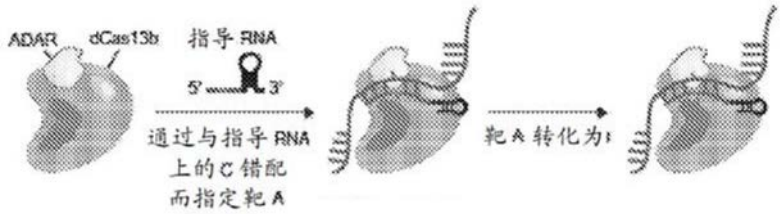


图50A

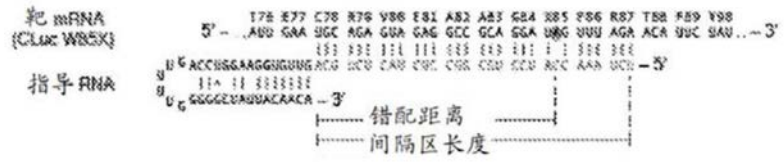


图50B

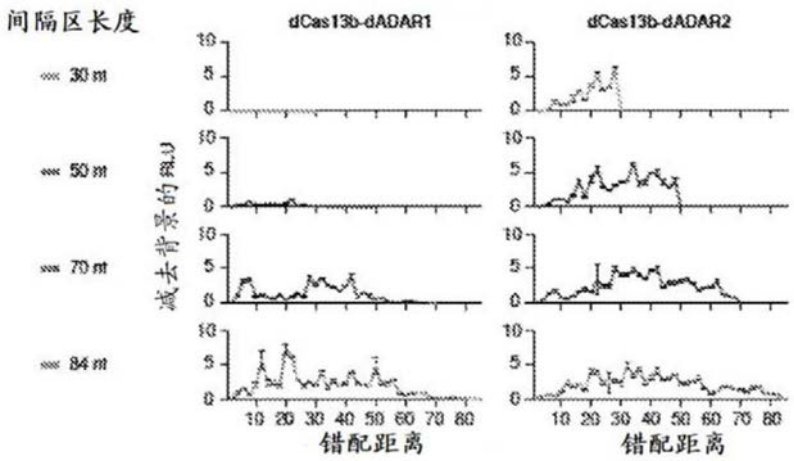


图50C

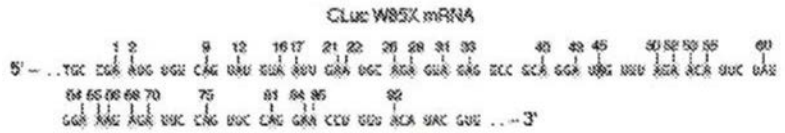


图50D

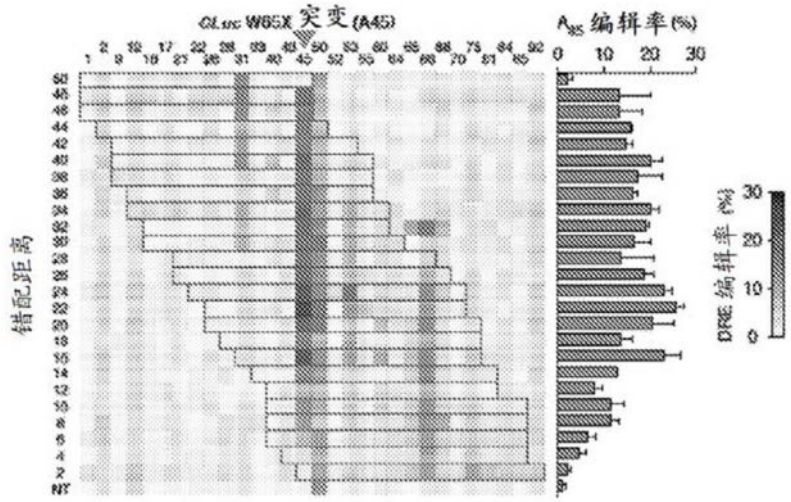


图50E



图51A

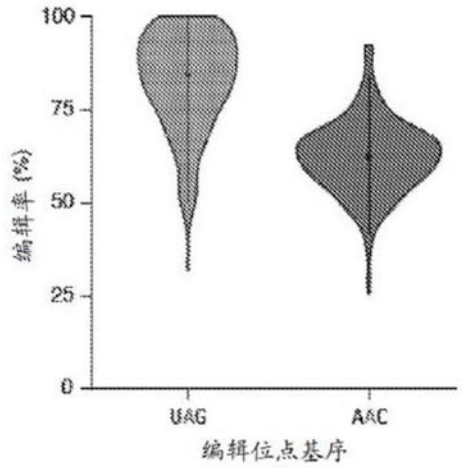


图51B

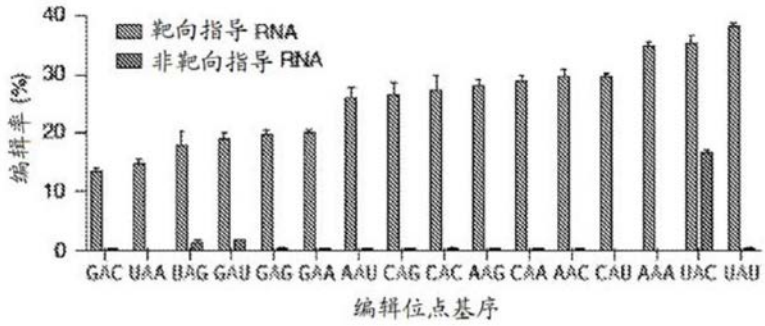


图51C

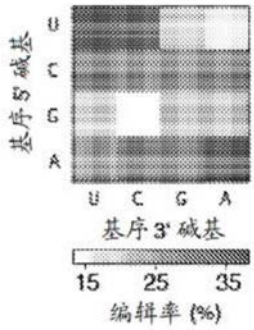


图51D


AVPR2 mRNA
(878G>A, Trp293Ter) X 连锁性肾原性尿崩症
指导 RNA

图52A

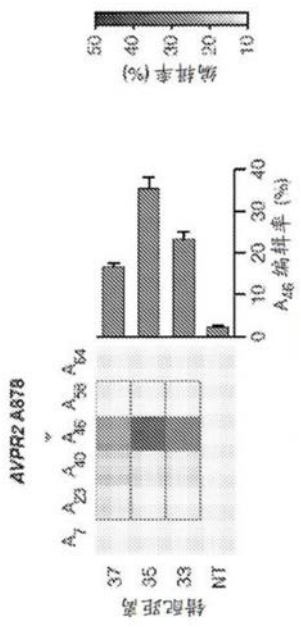

编辑率 (%)
AVPR2 A878
错配距离
NT
A56 编辑率 (%)

图52B


FANCC mRNA
(1517G>A, Trp506Ter) 范可尼贫血，互补群 C
靶标 RNA
5'
3'

图52C

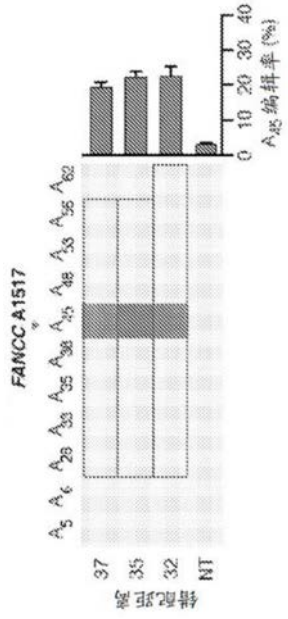

FANCC A1517
错配距离
37
35
32
NT
0 10 20 30 40
A45 编辑率 (%)

图52D

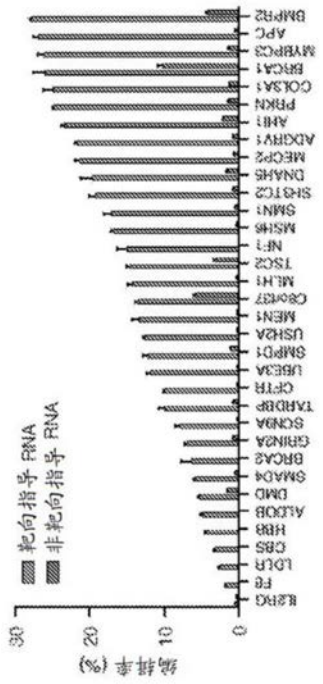


图52E

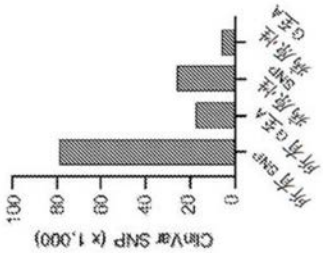


图52F

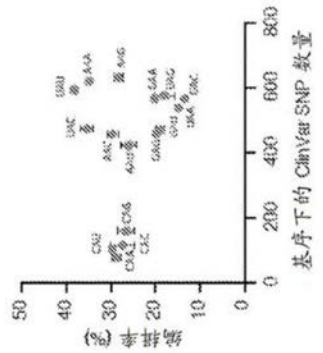


图52G

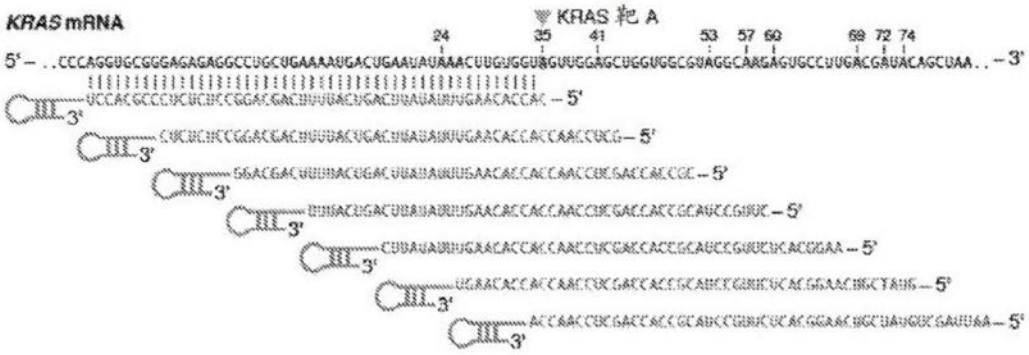


图53A

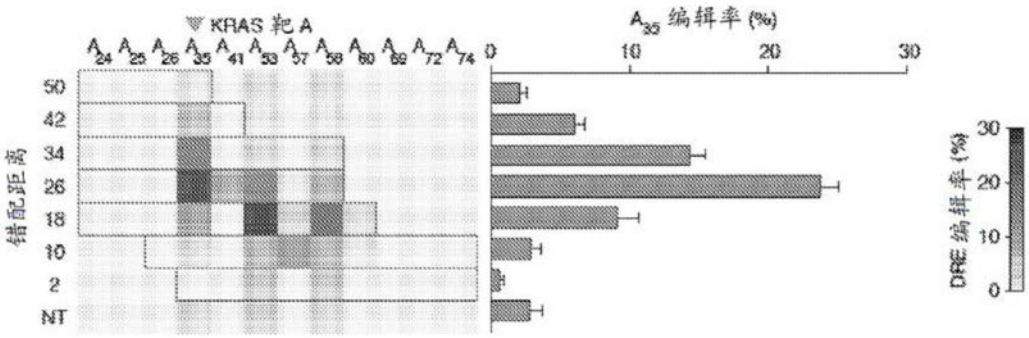


图53B

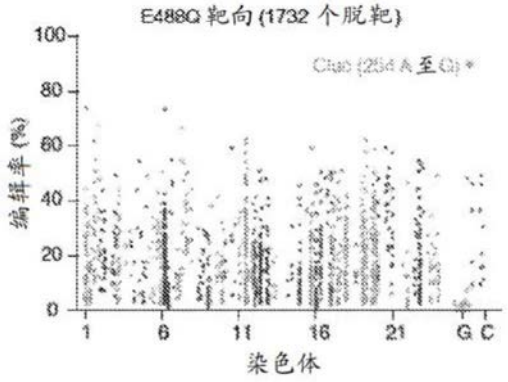


图53C

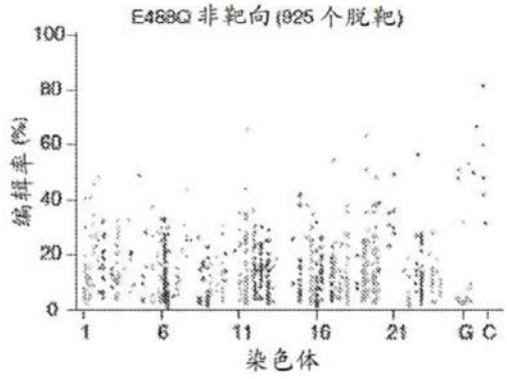


图53D

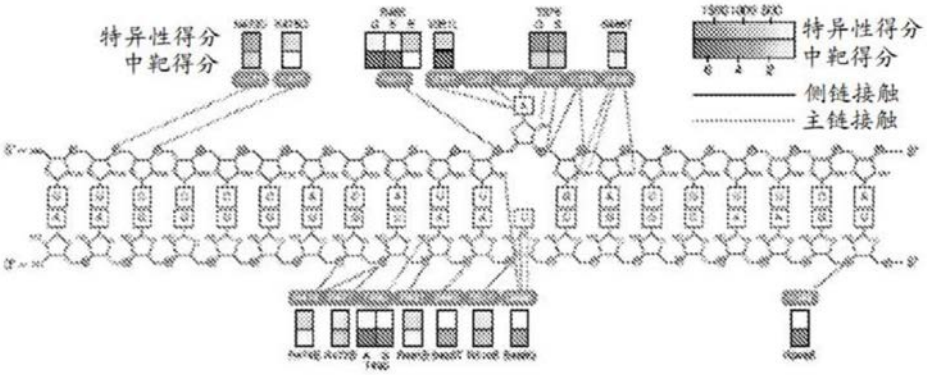


图54A

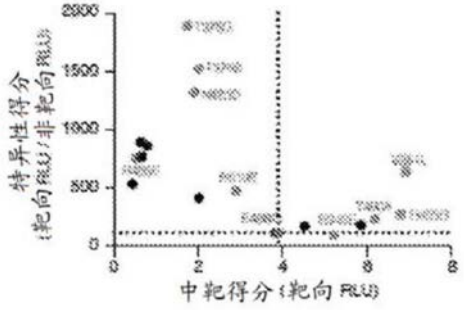


图54B

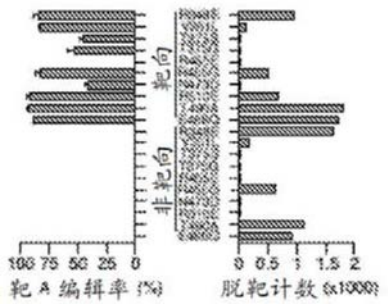


图54C

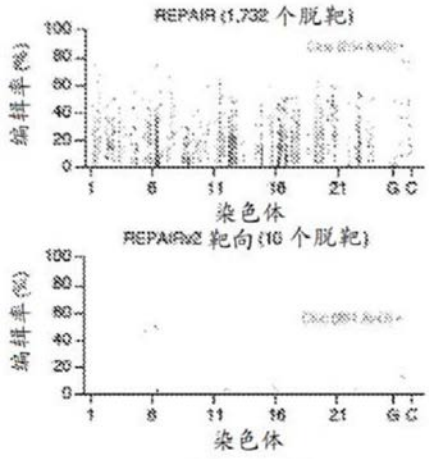


图54D



图54E

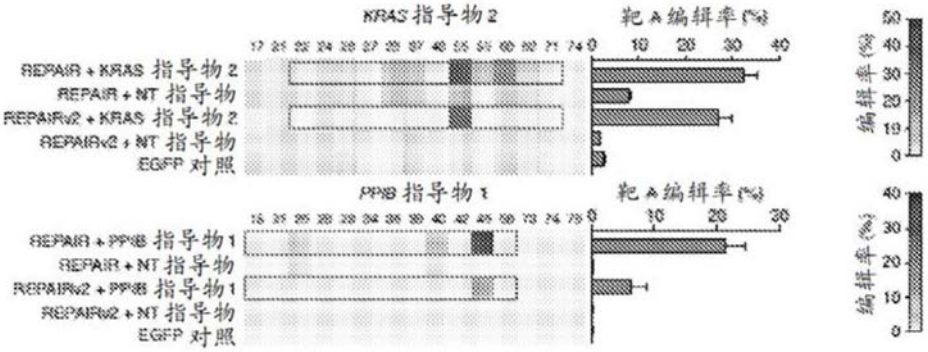


图54G

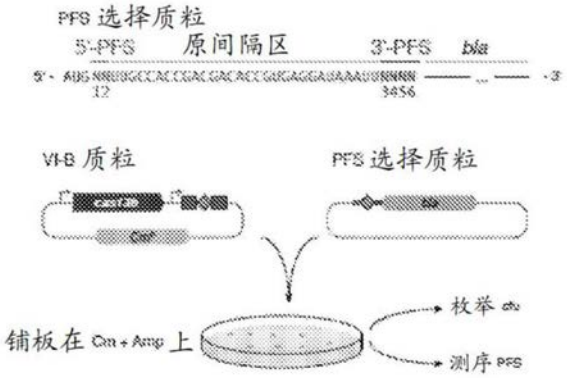


图55A

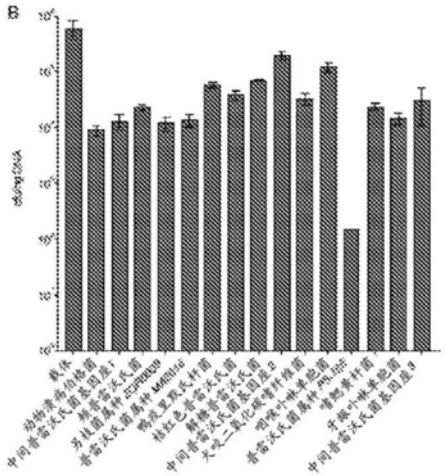


图55A

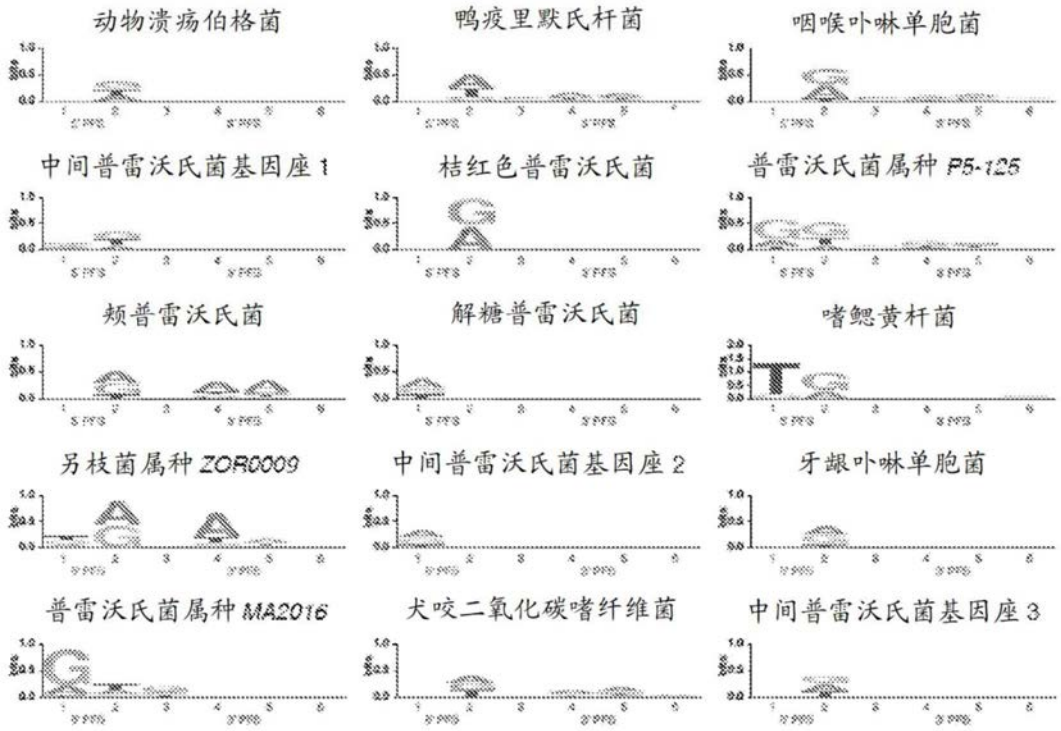

动物溃疡伯格菌
鸭疫里默氏杆菌
咽喉卟啉单胞菌
中间普雷沃氏菌基因座 1
桔红色普雷沃氏菌
普雷沃氏菌属种 P5-125
颊普雷沃氏菌
解糖普雷沃氏菌
嗜鳃黄杆菌
另枝菌属种 ZOR0009
中间普雷沃氏菌基因座 2
牙龈卟啉单胞菌
普雷沃氏菌属种 MA2016
犬咬二氧化碳嗜纤维菌
中间普雷沃氏菌基因座 3

图55C

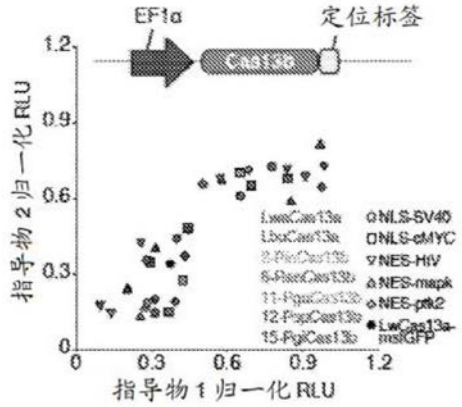

EF1α
定位标签
Cas13a
指导物 2 归一化 RLU
指导物 1 归一化 RLU
0
0.3
0.6
0.9
1.2
LwaCas13a
LbuCas13a
8-PauCas13b
11-PguCas13b
12-PspCas13b
15-PgiCas13b
NLS-SV40
NLS-cMYC
NES-HIV
NES-mapk
NES-pik2
LwCas13a-msfGFP

图56A

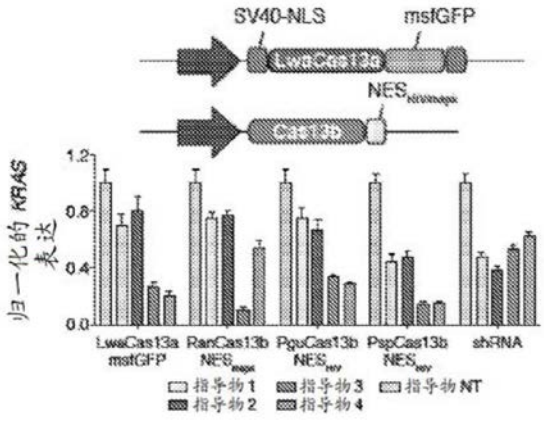


图56B



图56C

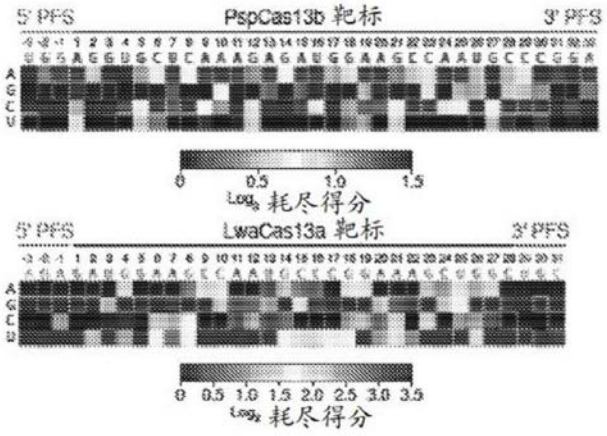


图56D

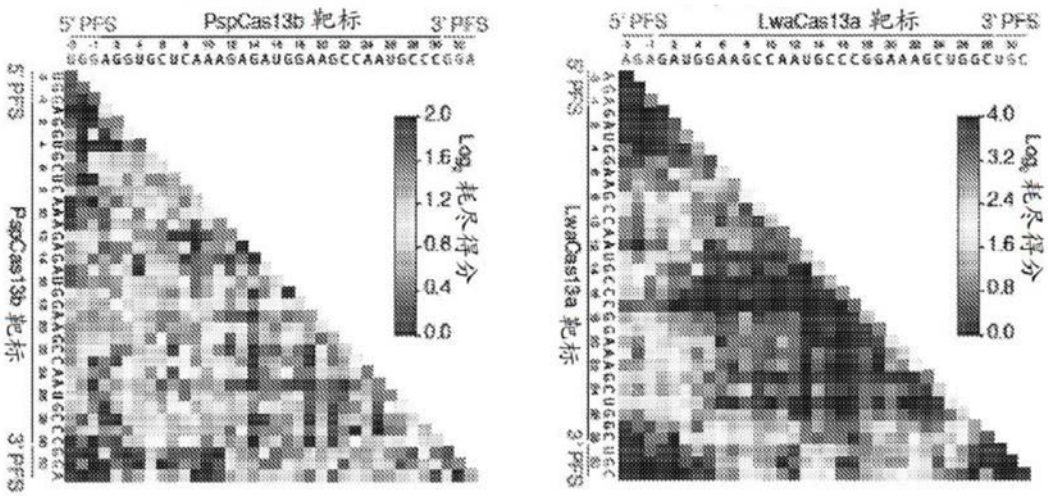


图56E

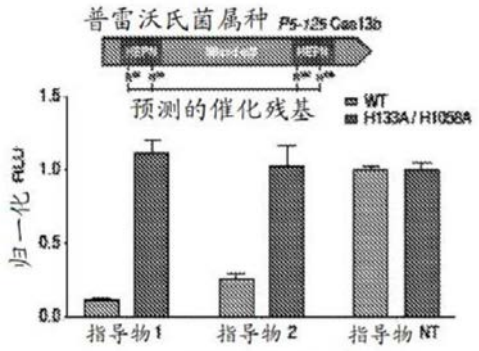


图57A

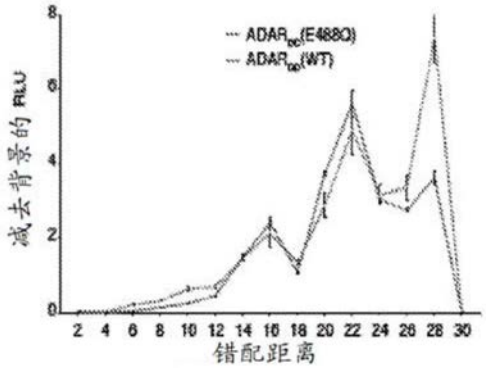


图57B

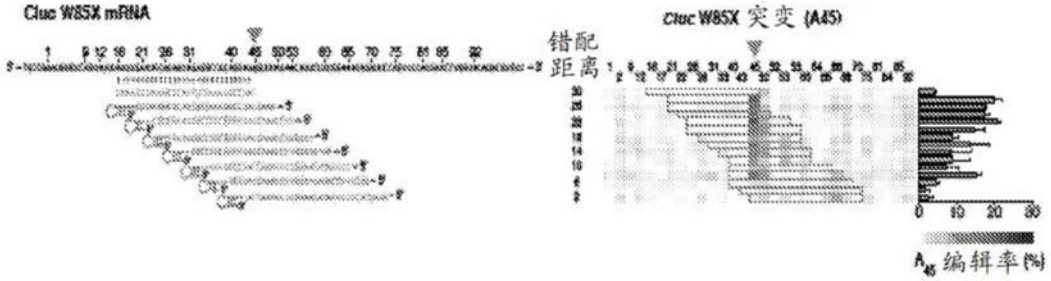


图57C



图57D

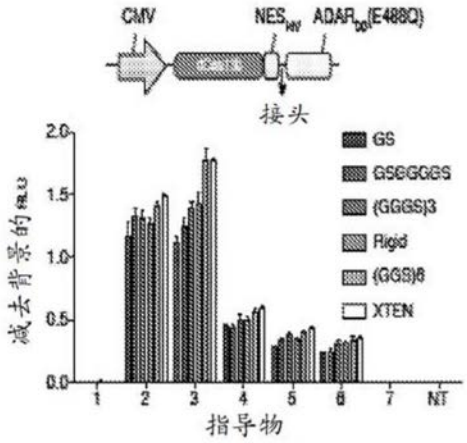


图57E

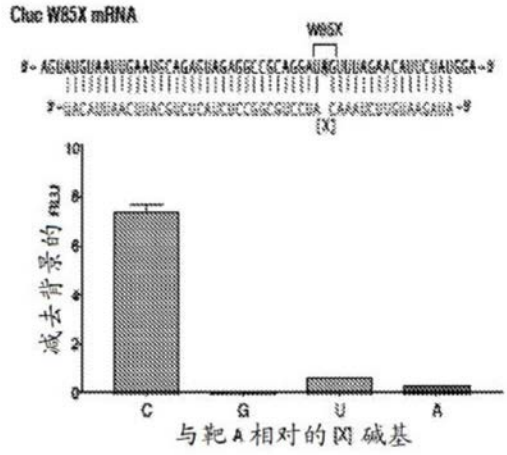

Cluc W85X mRNA
W85X
减去背景的
C
G
U
A
与靶A相对的[X]碱基

图57F

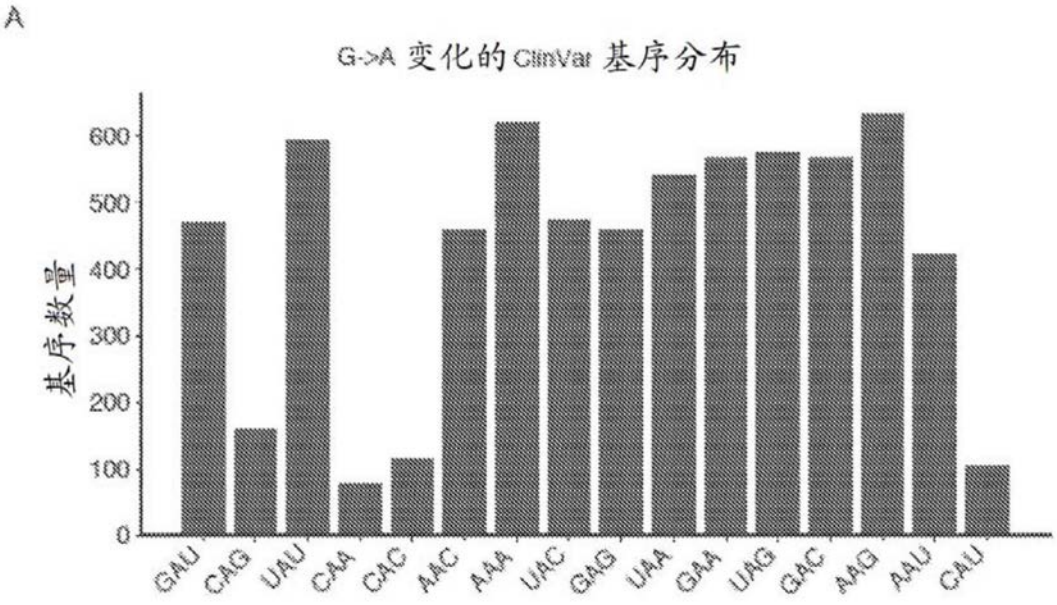

A
G->A 变化的ClinVar 基序分布
基序数量
0
100
200
300
400
500
600
GAU
CAG
UAU
CAA
CAC
AAC
AAA
UAC
GAG
UAA
GAA
UAG
GAC
AAG
AAU
CAU

图58

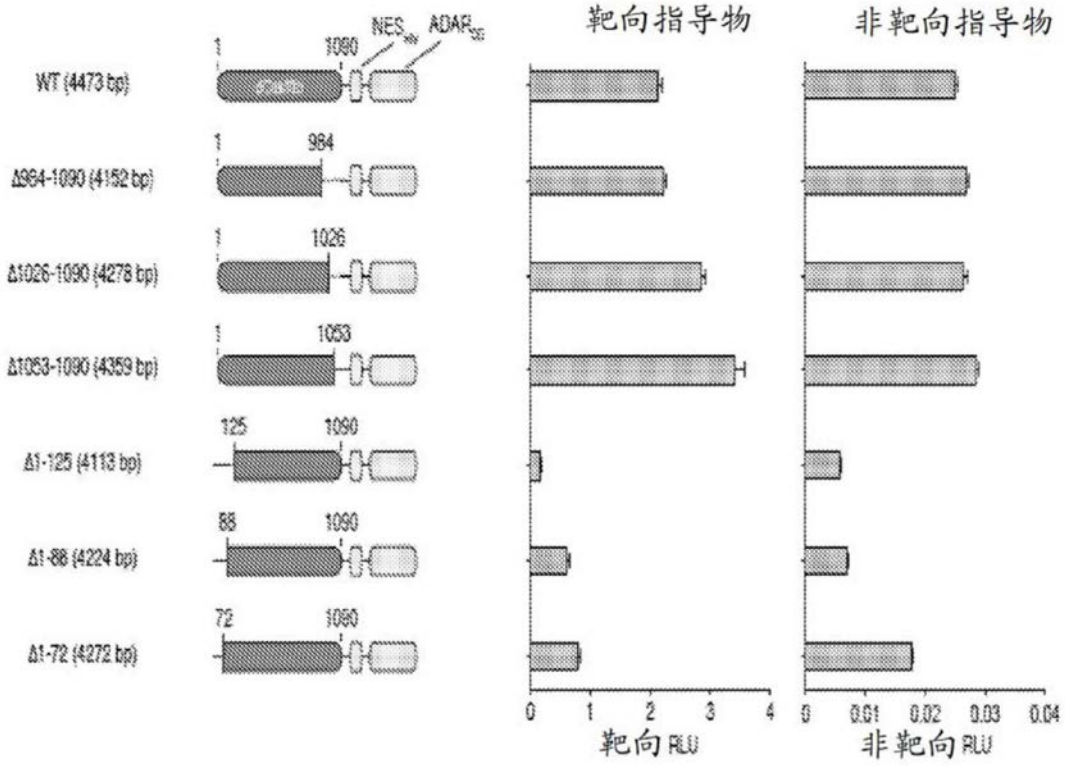

靶向指导物
非靶向指导物
NES
ADAR
WT (4473 bp)
Δ964-1090 (4152 bp)
Δ1026-1090 (4278 bp)
Δ1053-1090 (4359 bp)
Δ1-125 (4113 bp)
Δ1-88 (4224 bp)
Δ1-72 (4272 bp)
靶向 RLU
非靶向 RLU

图59

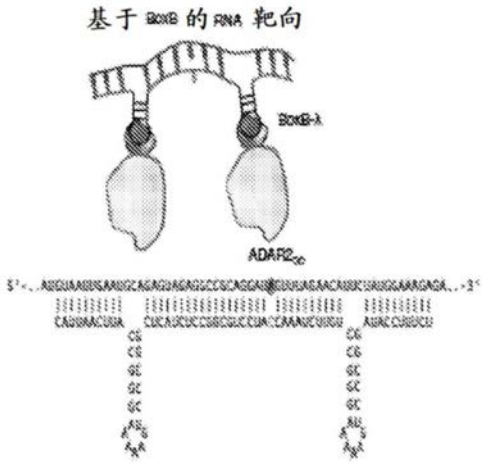

基于 BOXB 的 RNA 靶向
BOXB-λ
ADAR2

图60A

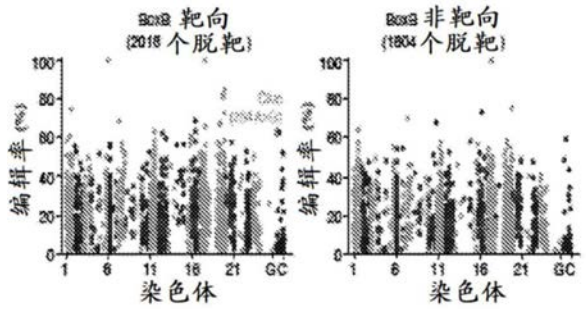


图60B

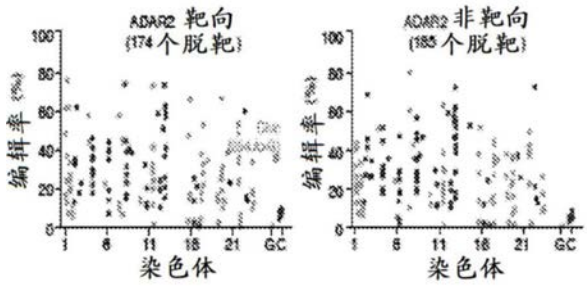


图60C

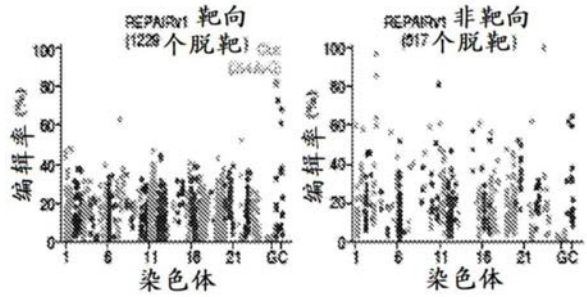


图60D

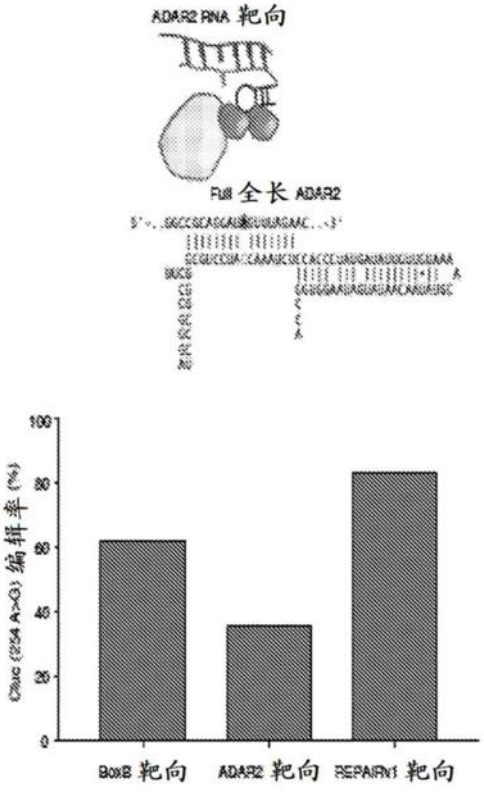

ADAR2 RNA 靶向
Full 全长 ADAR2
编辑率 (%)
100
80
60
40
20
0
BoxB 靶向
ADAR2 靶向
REPAIRV1 靶向

图60E

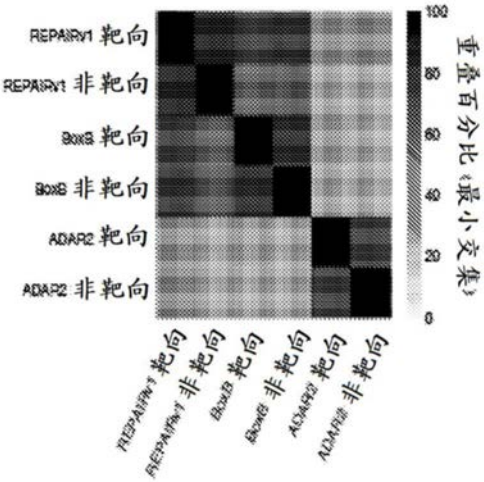

REPAIRV1 靶向
REPAIRV1 非靶向
BoxB 靶向
BoxB 非靶向
ADAR2 靶向
ADAR2 非靶向
重叠百分比（最小交集）
100
80
60
40
20
0

图60F

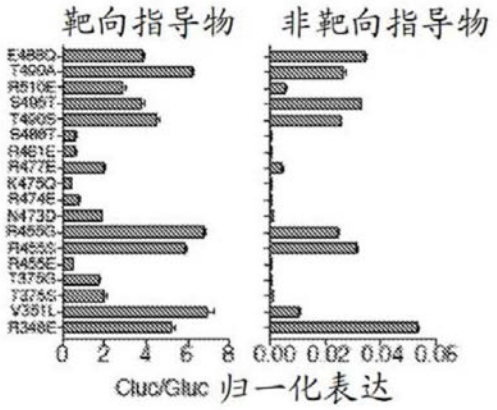


图61A

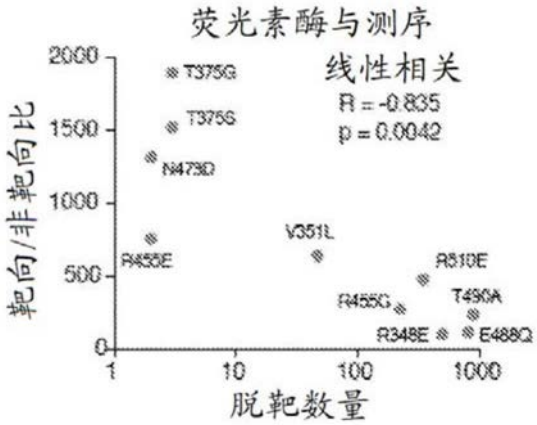


图61B

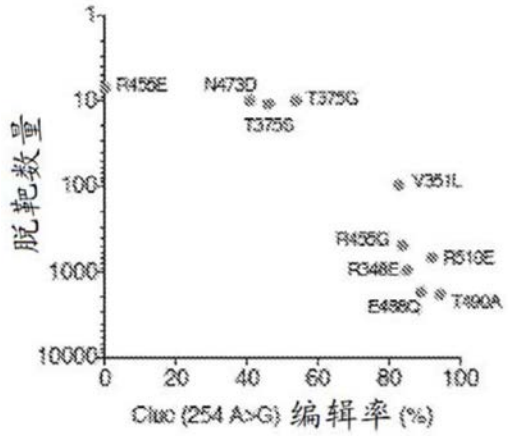


图61C

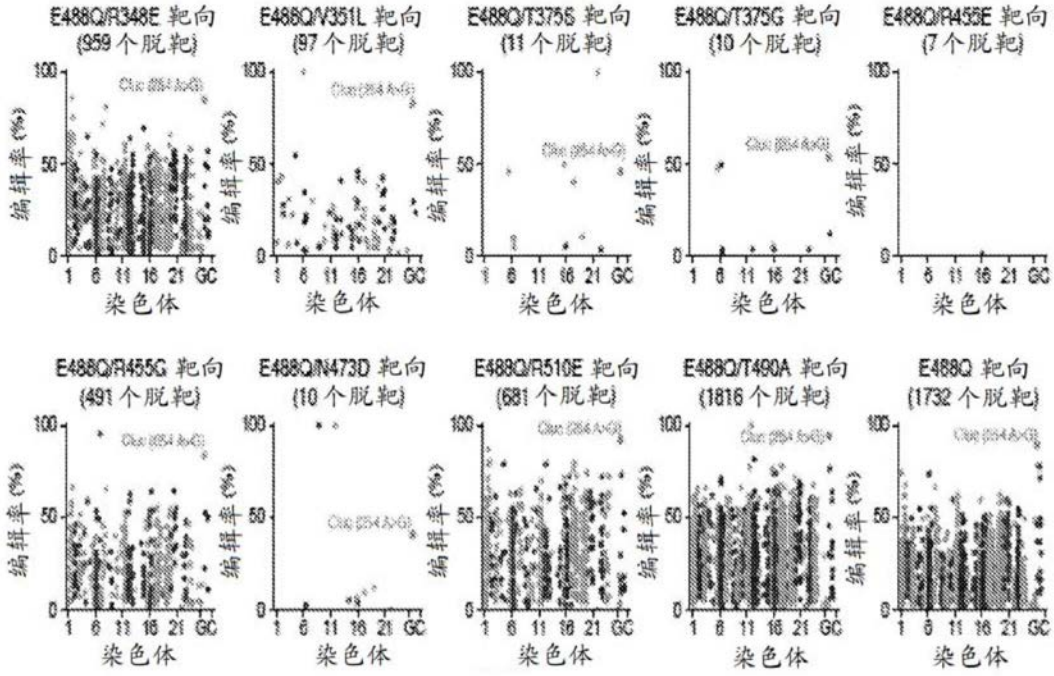

E488Q/R348E 靶向
(959 个脱靶)
E488Q/V351L 靶向
(97 个脱靶)
E488Q/T375S 靶向
(11 个脱靶)
E488Q/T375G 靶向
(10 个脱靶)
E488Q/R455E 靶向
(7 个脱靶)
E488Q/R455G 靶向
(491 个脱靶)
E488Q/N473D 靶向
(10 个脱靶)
E488Q/R510E 靶向
(681 个脱靶)
E488Q/T490A 靶向
(1816 个脱靶)
E488Q 靶向
(1732 个脱靶)
编辑率 (%)
染色体
1 6 11 16 21 GC

图62A

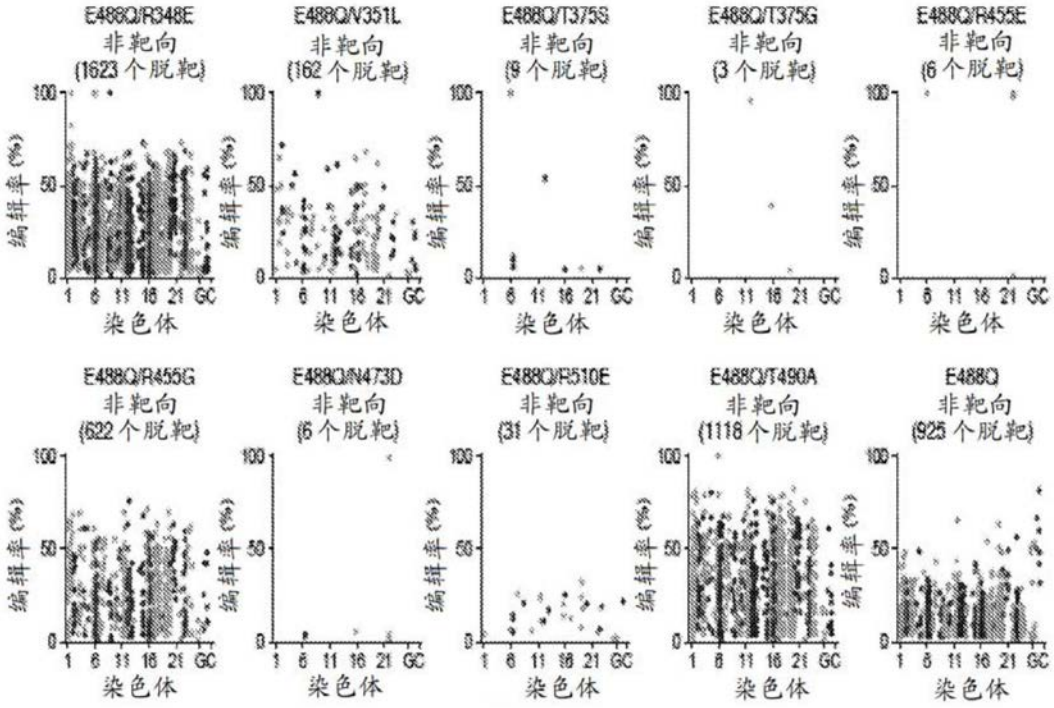

E488Q/R348E
非靶向
(1623 个脱靶)
E488Q/V351L
非靶向
(162 个脱靶)
E488Q/T375S
非靶向
(9 个脱靶)
E488Q/T375G
非靶向
(3 个脱靶)
E488Q/R455E
非靶向
(6 个脱靶)
E488Q/R455G
非靶向
(622 个脱靶)
E488Q/N473D
非靶向
(6 个脱靶)
E488Q/R510E
非靶向
(31 个脱靶)
E488Q/T490A
非靶向
(1118 个脱靶)
E488Q
非靶向
(925 个脱靶)
编辑率 (%)
染色体
1 6 11 16 21 GC

图62B

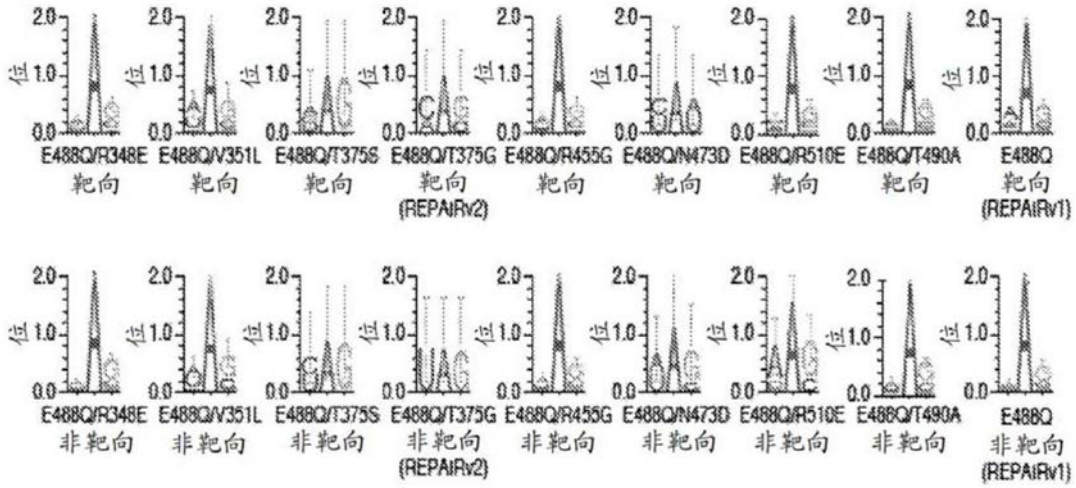

E488Q/R348E 靶向
E488Q/V351L 靶向
E488Q/T375S 靶向
E488Q/T375G 靶向 (REPAIRv2)
E488Q/R455G 靶向
E488Q/N473D 靶向
E488Q/R510E 靶向
E488Q/T490A 靶向
E488Q 靶向 (REPAIRv1)
E488Q/R348E 非靶向
E488Q/V351L 非靶向
E488Q/T375S 非靶向
E488Q/T375G 非靶向 (REPAIRv2)
E488Q/R455G 非靶向
E488Q/N473D 非靶向
E488Q/R510E 非靶向
E488Q/T490A 非靶向
E488Q 非靶向 (REPAIRv1)

图63A

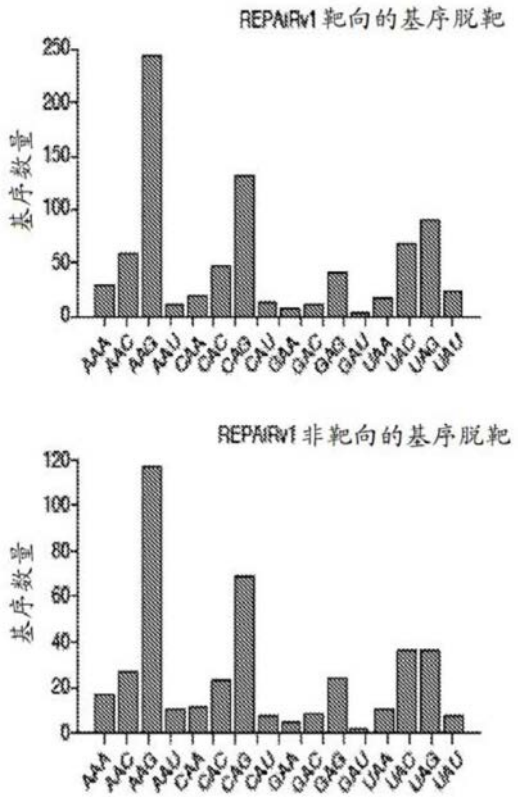

REPAIRv1 靶向的基序脱靶
基序数量
REPAIRv1 非靶向的基序脱靶
基序数量
AAA AAC AAG AAU CAA CAC CAG CAU GAA GAC GAG GAU UAA UAC UAG UAU

图63B

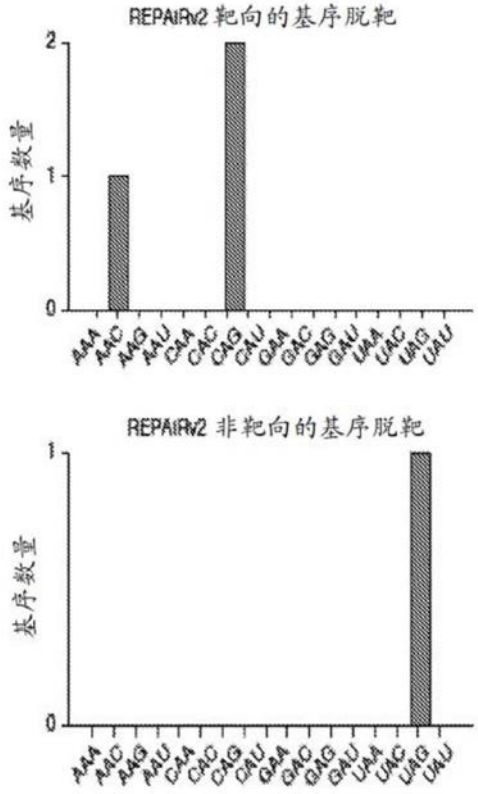

REPAIRv2 靶向的基序脱靶
基序数量
2
1
0
AAA AAC AAG AAU CAA CAC CAG CAU GAA GAC GAG GAU UAA UAC UAG UAU
REPAIRv2 非靶向的基序脱靶
基序数量
1
0
AAA AAC AAG AAU CAA CAC CAG CAU GAA GAC GAG GAU UAA UAC UAG UAU

图63C

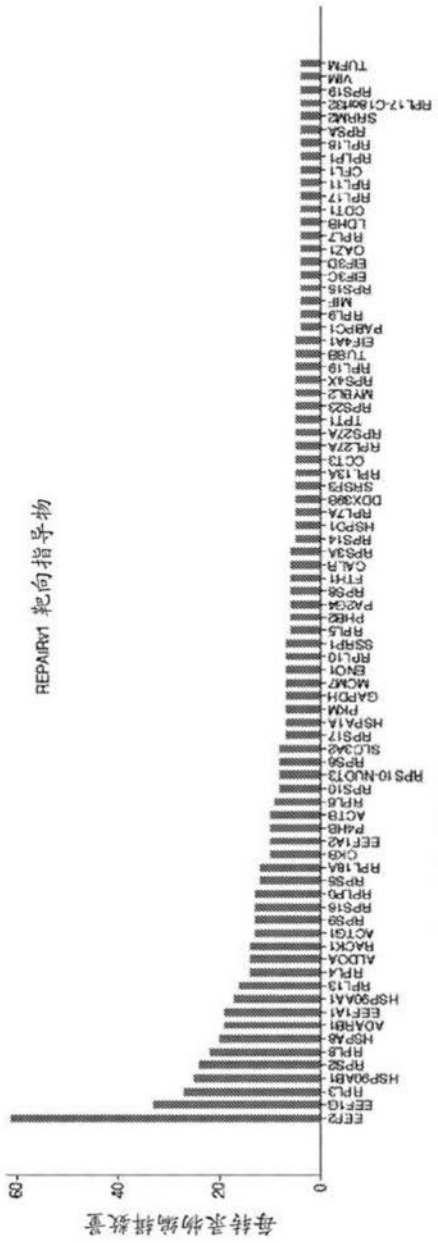

REPAIRv1 靶向指导物
EEF2
EEF1G
RPL3
HSP90AB1
RPS2
RPL8
HSPA8
ADARB1
EEF1A1
HSP90AA1
RPL13
RPLA
ALDOA
PAICK1
ACTG1
RPS9
RPS16
RPLP0
RPS5
RPL18A
CYB
EEF1A2
PABPC1
ACTB
RPL6
RPS10
RPS10-NUDT3
RPS6
SLC3A2
RPS17
HSPA1A
PKM
GAPDH
MCM7
ENO1
RPL10
SSRP1
RPL5
PHB2
P4G4
RPS8
FTH1
CALR
RPS3A
HSPD1
RPS14
RPS3
RPL7A
DDX39B
SHSP1
RPL13A
CCT3
RPL27A
RPS27A
TPT1
RPS23
MYBL2
RPS4X
RPL19
TUBB
EIF4A1
PABPC1
RPL9
MIF
RPS15
EIF3C
EIF3D
CAZ1
RPL7
LDHB
CCT1
RPL17
RPL11
CFL1
RPLP1
RPL18
RPSA
SNRNP2
RPL17-C18orf32
RPS19
VIM
TUFM
60
40
20
0
每转录物缺辑数量

图64A

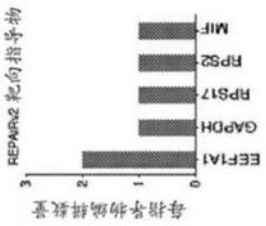

REPAIRv2 靶向指导物
EEF1A1
GAPDH
RPS17
RPS2
MIF
3
2
1
0
每指导物缺辑数量

图64B

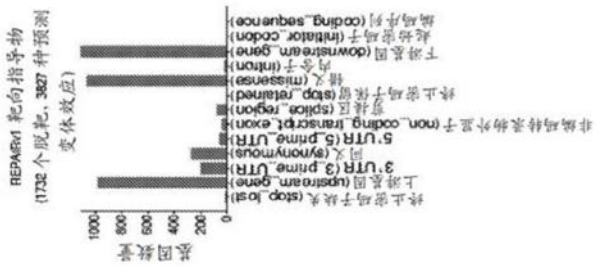


图64C

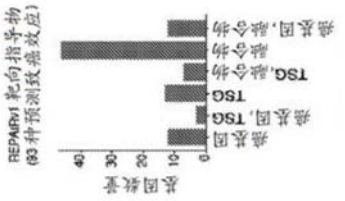


图64D



图64E

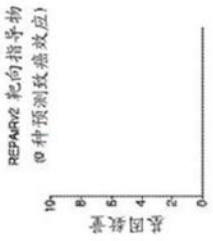


图64F

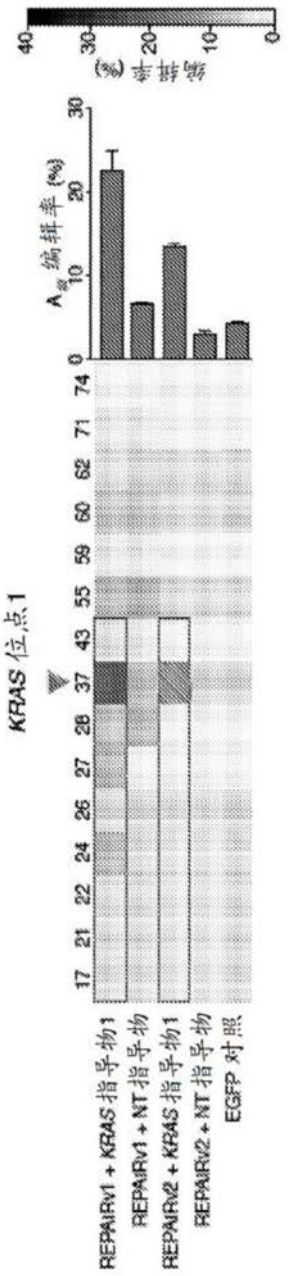

编辑率 (%)
40
30
20
10
0
A碱基编辑率 (%)
30
20
10
0
KRAS 位点1
17
21
22
24
26
27
28
37
43
55
59
60
62
71
74
REPAIRv1 + KRAS 指导物1
REPAIRv1 + NT 指导物
REPAIRv2 + KRAS 指导物1
REPAIRv2 + NT 指导物
EGFP 对照

图65A

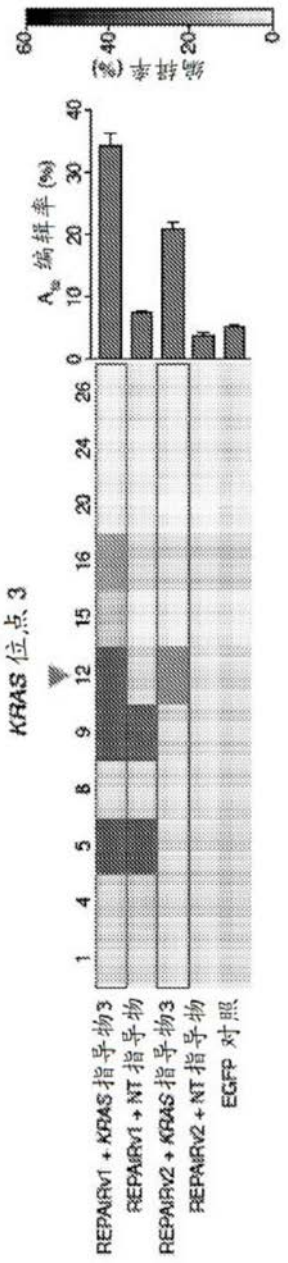

编辑率 (%)
60
40
20
0
A₅₀ 编辑率 (%)
40
30
20
10
0
KRAS 位点 3
1
4
5
8
9
12
15
16
20
24
26
REPAIRv1 + KRAS 指导物3
REPAIRv1 + NT 指导物
REPAIRv2 + KRAS 指导物3
REPAIRv2 + NT 指导物
EGFP 对照

图65B

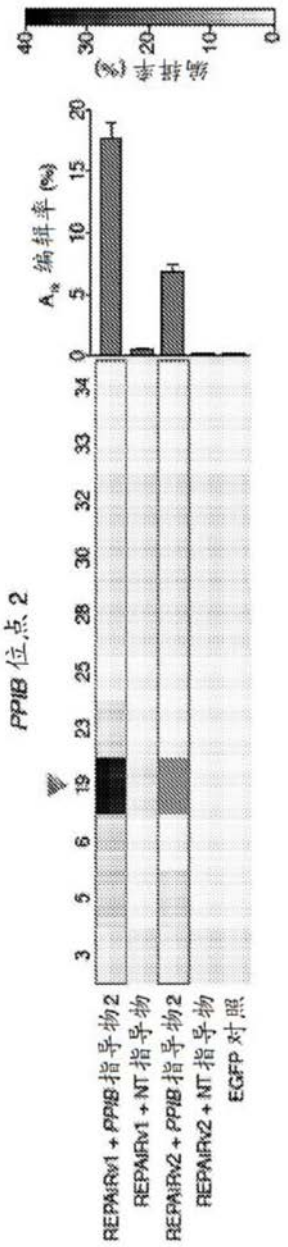

40
30
20
10
0
编辑率 (%)
A* 编辑率 (%)
0
5
10
15
20
PPIB 位点 2
3
5
6
19
23
25
28
30
32
33
34
REPAIRv1 + PPIB 指导物2
REPAIRv1 + NT 指导物
REPAIRv2 + PPIB 指导物2
REPAIRv2 + NT 指导物
EGFP 对照

图65C

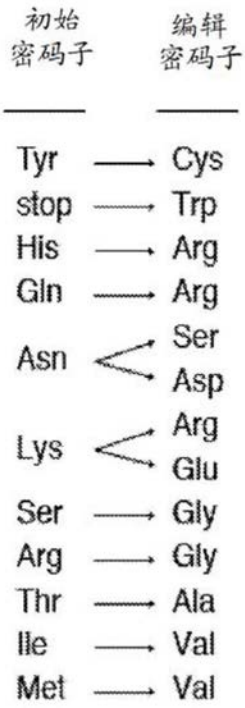

初始
密码子
编辑
密码子
Tyr → Cys
stop → Trp
His → Arg
Gln → Arg
Asn → Ser
Asn → Asp
Lys → Arg
Lys → Glu
Ser → Gly
Arg → Gly
Thr → Ala
Ile → Val
Met → Val

图66A

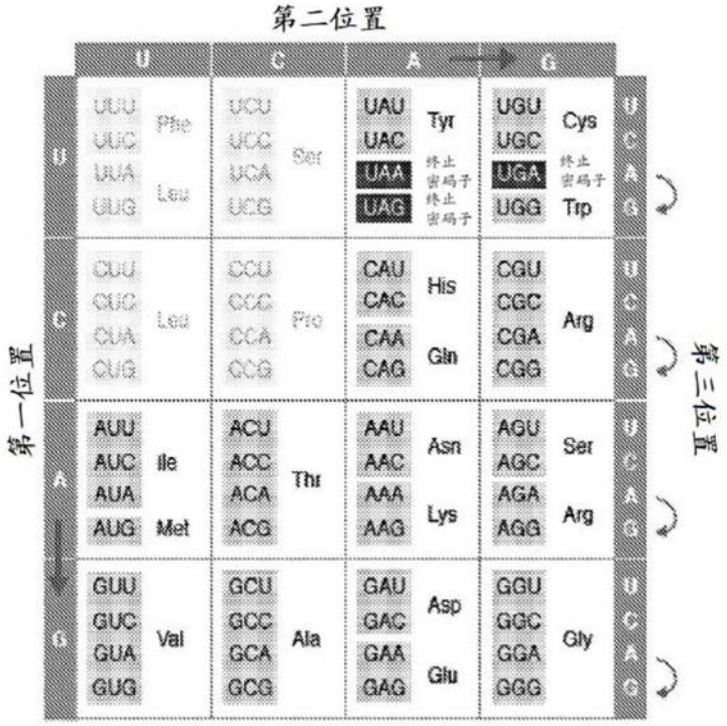

第二位置
U
C
A
G
第一位置
第三位置
UUU Phe
UUC
UUA Leu
UUG
UCU
UCC
UCA Ser
UCG
UAU Tyr
UAC
UAA 终止密码子
UAG 终止密码子
UGU Cys
UGC
UGA 终止密码子
UGG Trp
CUU
CUC
CUA Leu
CUG
CCU
CCC
CCA Pro
CCG
CAU His
CAC
CAA Gln
CAG
CGU
CGC
CGA Arg
CGG
AUU
AUC Ile
AUA
AUG Met
ACU
ACC
ACA Thr
ACG
AAU Asn
AAC
AAA Lys
AAG
AGU Ser
AGC
AGA Arg
AGG
GUU
GUC Val
GUA
GUG
GCU
GCC
GCA Ala
GCG
GAU Asp
GAC
GAA Glu
GAG
GGU
GGC Gly
GGA
GGG

图66B

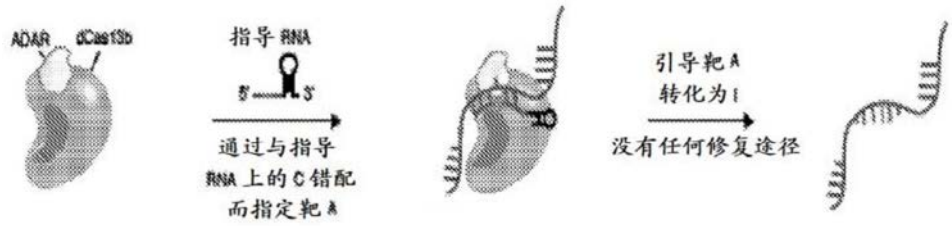


图66C



图66D



图66E

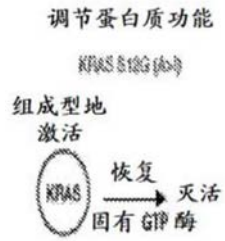


图66F

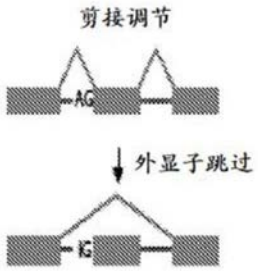

剪接调节
AG
外显子跳过
AG

图66G

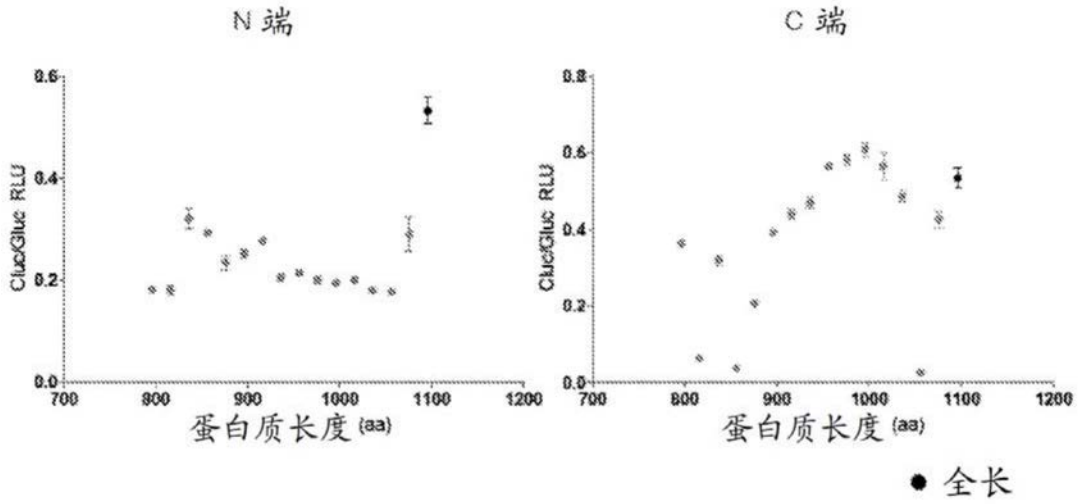

N 端
C 端
Cluc/Gluc RLU
蛋白质长度 (aa)
全长

图67

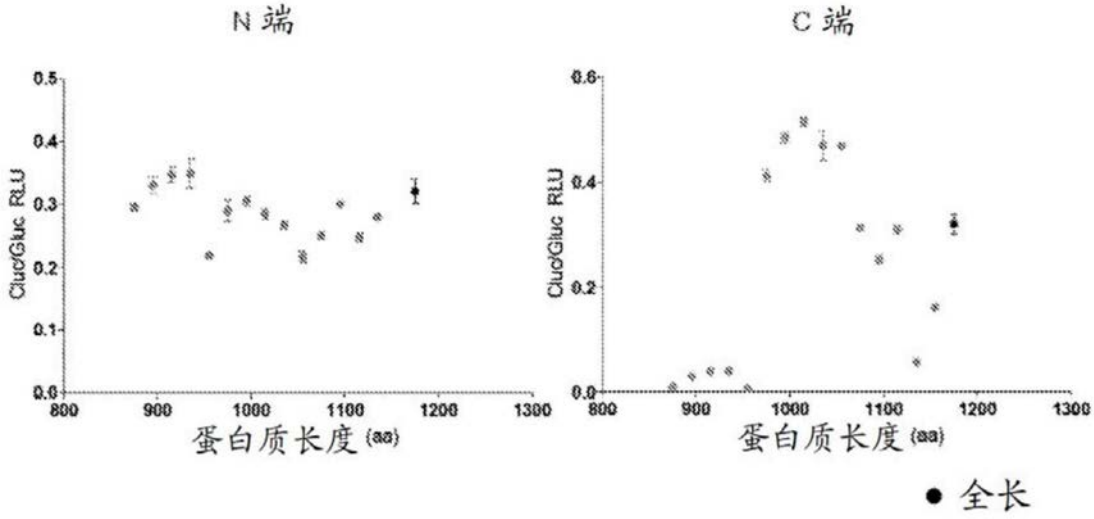

N 端
C 端
Cluc/Gluc RLU
蛋白质长度 (aa)
全长

图68

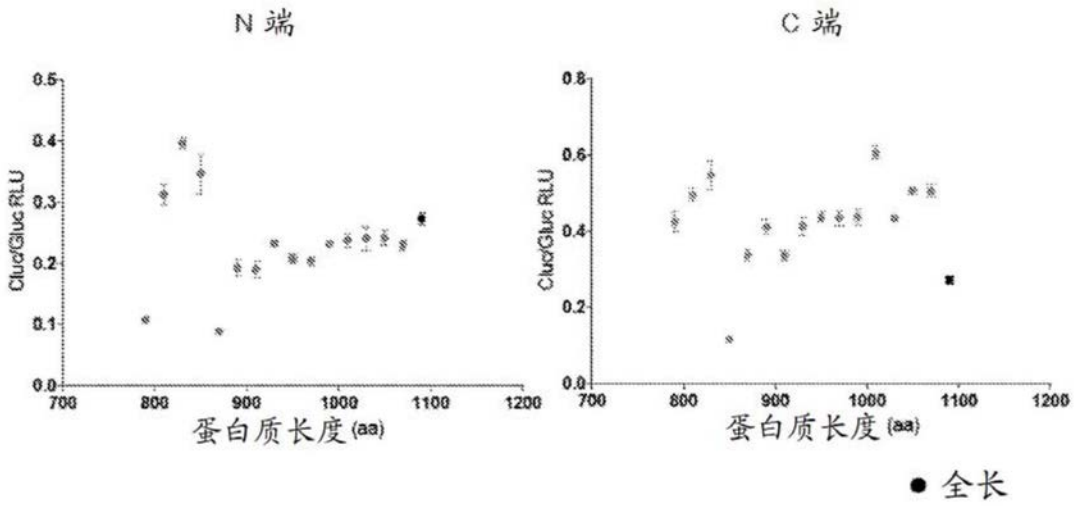

N 端
C 端
Cluc/Gluc RLU
蛋白质长度(aa)
全长

图69

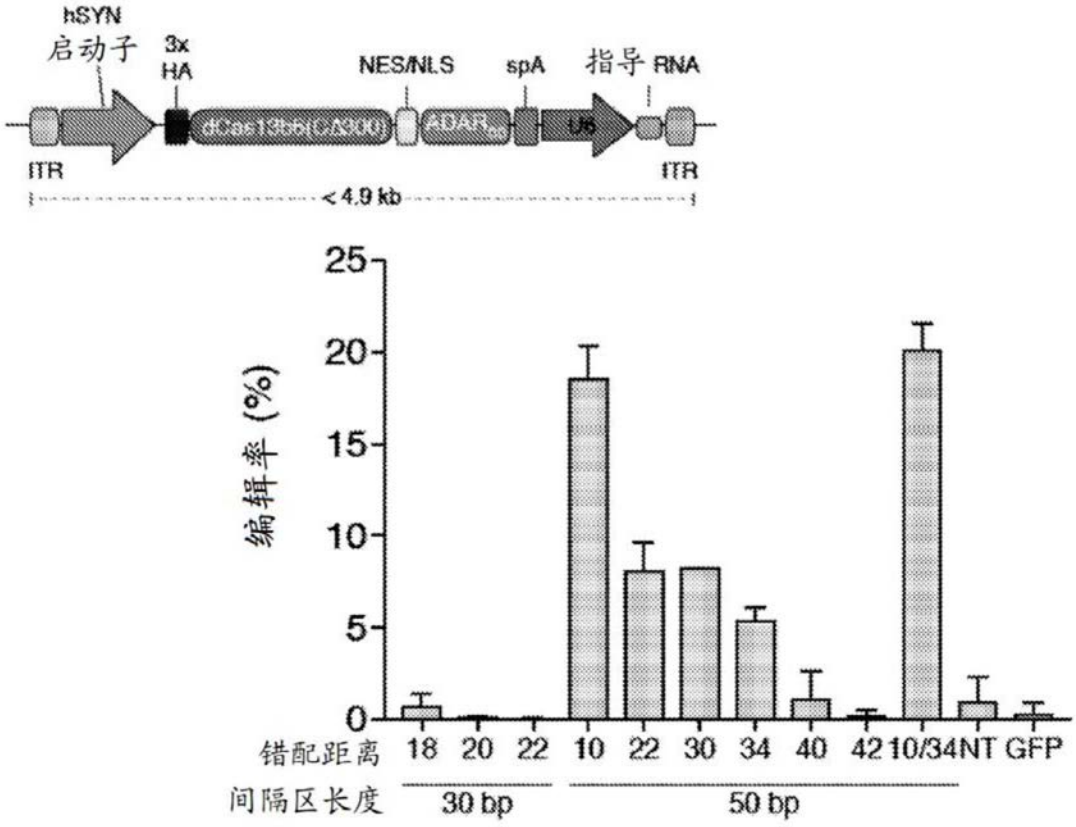

hSYN
启动子
3x
HA
NES/NLS
spA
指导 RNA
ITR
ITR
< 4.9 kb
编辑率 (%)
错配距离 18 20 22 10 22 30 34 40 42 10/34NT GFP
间隔区长度 30 bp 50 bp

图70

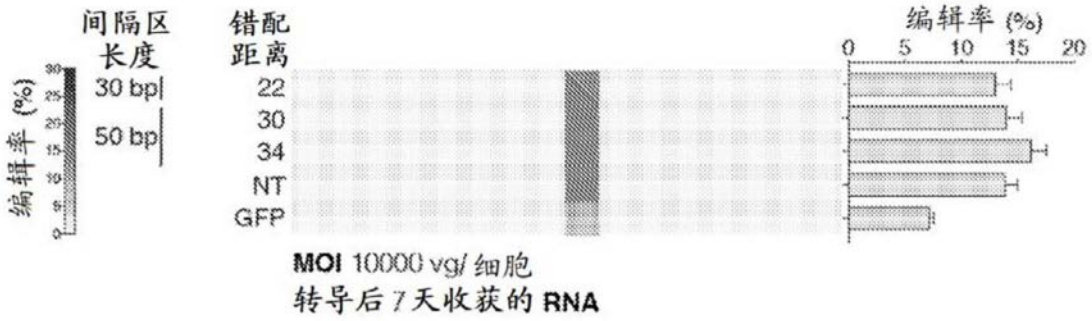

间隔区
长度
30 bp
50 bp
错配
距离
22
30
34
NT
GFP
编辑率 (%)
0 5 10 15 20
编辑率 (%)
MOI 10000 vg/细胞
转导后7天收获的RNA

图71

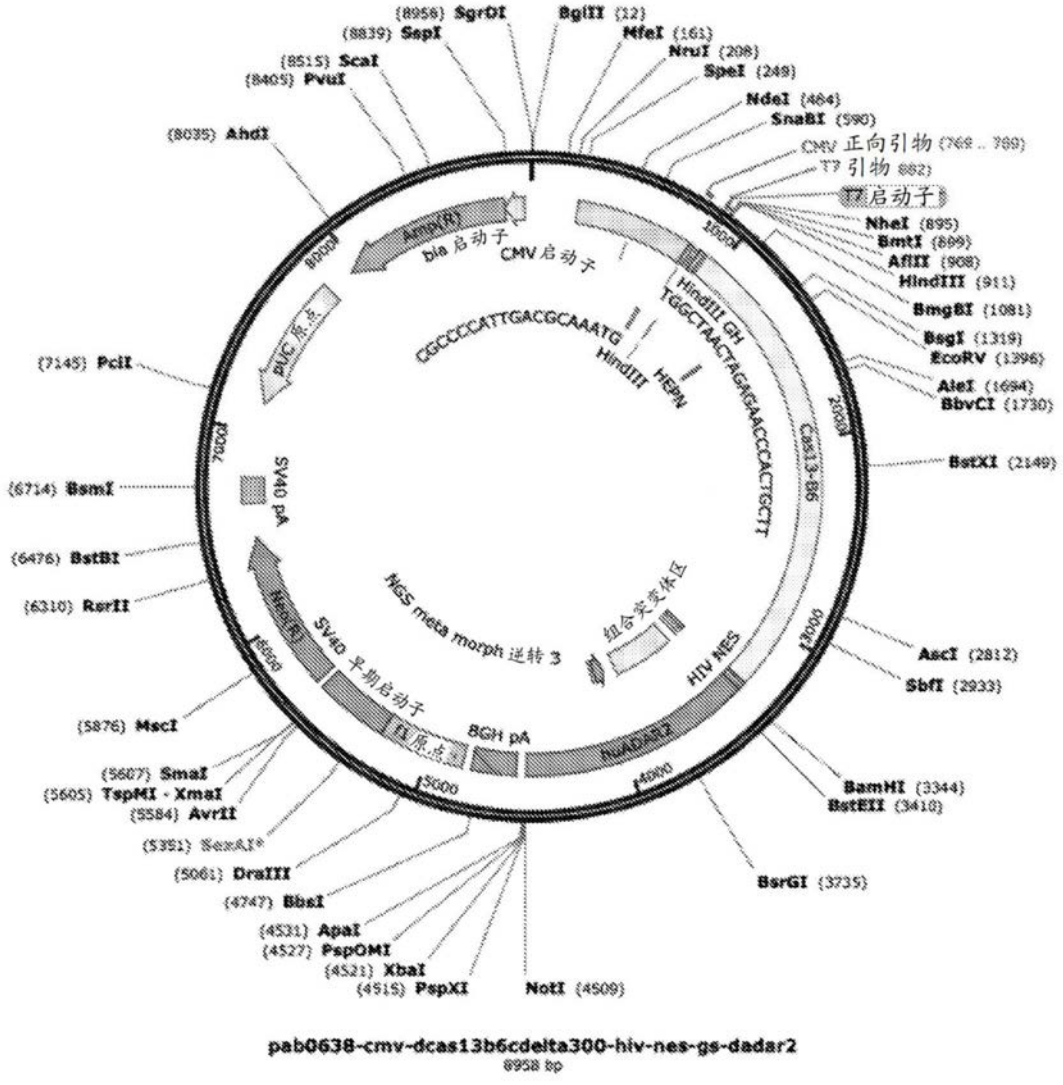

(8958) SgrDI
BglII (12)
(8839) SspI
MfeI (161)
(8515) ScaI
NruI (208)
(8405) PvuI
SpeI (249)
NdeI (484)
(8035) AhdI
SnaBI (590)
CMV 正向引物 (769 .. 789)
T7 引物 (882)
T7 启动子
NheI (895)
BmtI (899)
AflII (908)
HindIII (911)
BmgBI (1081)
BsgI (1319)
EcoRV (1396)
(7148) PciI
AleI (1694)
BbvCI (1730)
BstXI (2149)
(6714) BsmI
(6476) BstBI
(6310) RsrII
AscI (2812)
SbfI (2933)
(5876) MscI
(5607) SmaI
(5605) TspMI - XmaI
(5584) AvrII
(5351) KasAI*
(5061) DraIII
(4747) BbsI
(4531) ApaI
(4527) PspOMI
(4521) XbaI
(4515) PspXI
NotI (4509)
BamHI (3344)
BstEII (3410)
BsrGI (3735)
Amp(R) bla 启动子
CMV 启动子
HindIII GH
TGGCTAACTAGAGAACCCACTGCTT
CGCCCCATTGACGCAAATG
HindIII
HEPN
Cas13-B6
pUC 原点
SV40 pA
NGS meta morph 逆转 3
组合突变体区
HIV NES
Kan(R)
SV40 早期启动子
f1 原点
BGH pA
NUADAR2
pab0638-cmv-dcas13b6cdelta300-hiv-nes-gs-dadar2
8958 bp

图72

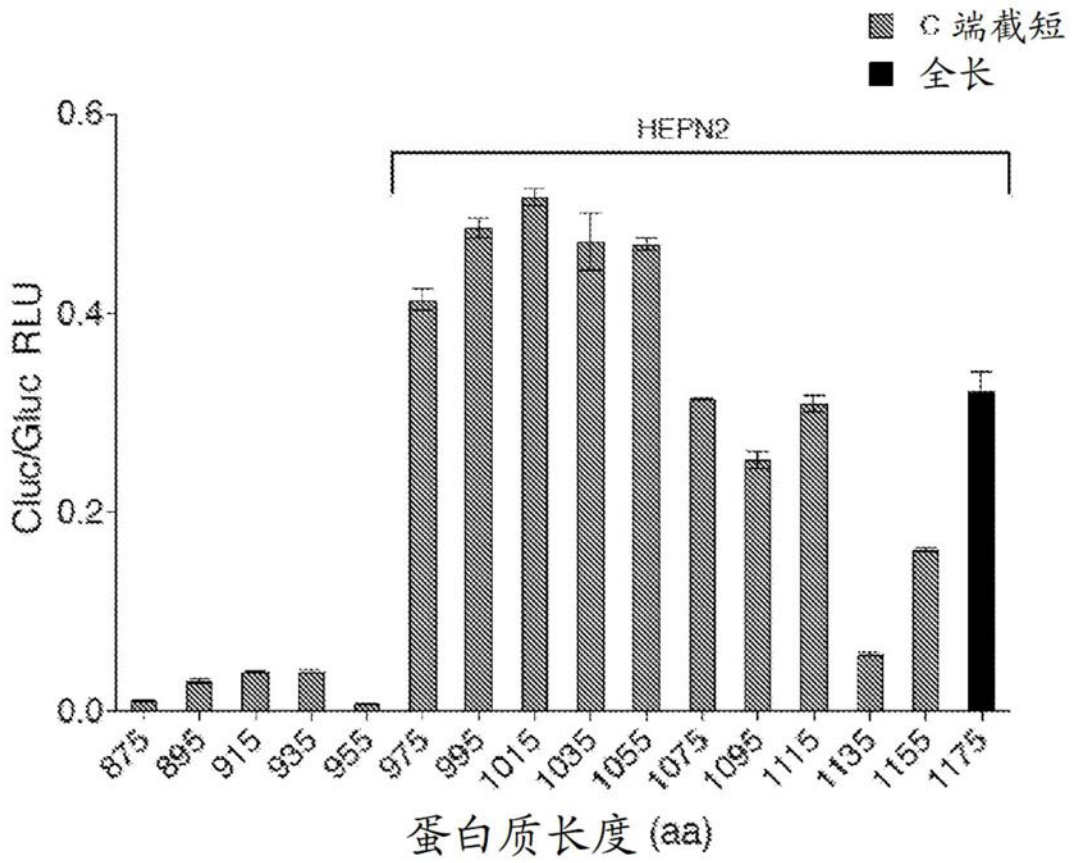

C端截短
全长
HEPN2
0.6
0.4
0.2
0.0
Cluc/Gluc RLU
875
895
915
935
955
975
995
1015
1035
1055
1075
1095
1115
1135
1155
1175
蛋白质长度(aa)

图73

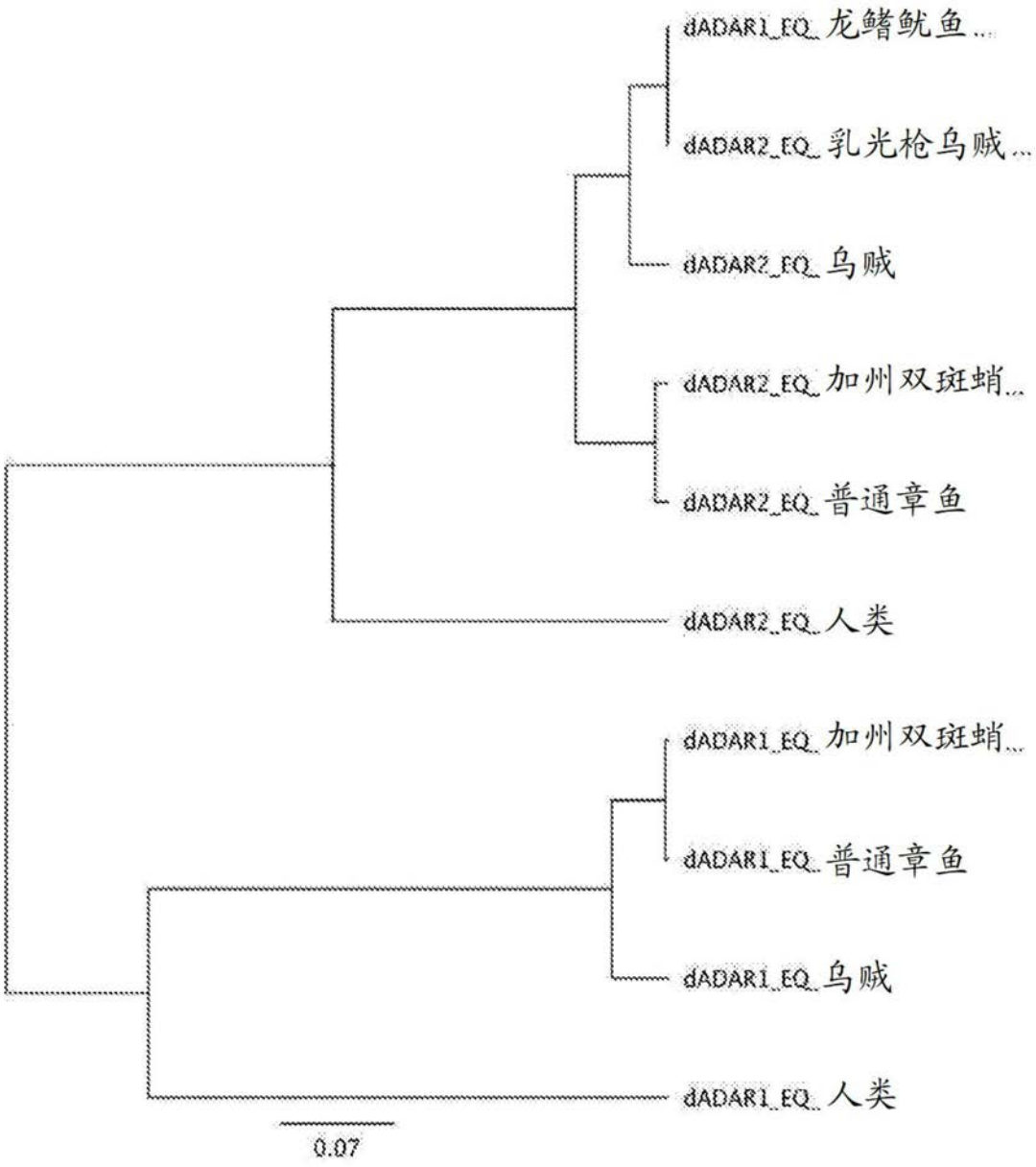

dADAR1_EQ_龙鳍鱿鱼...
dADAR2_EQ_乳光枪乌贼...
dADAR2_EQ_乌贼
dADAR2_EQ_加州双斑蛸...
dADAR2_EQ_普通章鱼
dADAR2_EQ_人类
dADAR1_EQ_加州双斑蛸...
dADAR1_EQ_普通章鱼
dADAR1_EQ_乌贼
dADAR1_EQ_人类
0.07

图74