(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 112996912 A
(43)申请公布日 2021.06.18

(21)申请号 201980072798.1

(22)申请日 2019.09.06

(30)优先权数据
62/728007 2018.09.06 US
62/766433 2018.10.17 US
62/773146 2018.11.29 US
62/773150 2018.11.29 US
62/780241 2018.12.15 US

(85)PCT国际申请进入国家阶段日
2021.04.30

(86)PCT国际申请的申请数据
PCT/US2019/050095 2019.09.06

(87)PCT国际申请的公布数据
WO2020/051555 EN 2020.03.12

(71)申请人 加利福尼亚大学董事会
地址 美国加利福尼亚州

(72)发明人 普拉尚特·马利
德拉瓦·卡特埃卡 D·梅鲁兹
陈赓灏 K·M·福特

(74)专利代理机构 青岛联智专利商标事务所有限公司 37101
代理人 阎斌斌 刘丹丹

(51)Int.Cl.
C12N 15/11(2006.01)
C12N 15/113(2006.01)

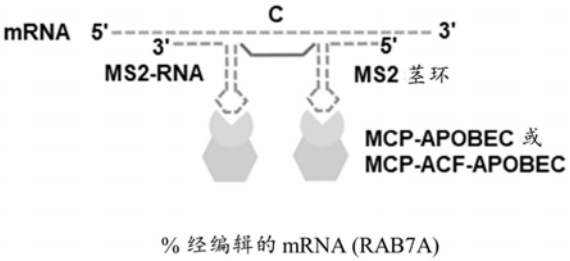
权利要求书7页 说明书81页 附图49页

(54)发明名称

经由改造的ADAR募集的RNA和DNA碱基编辑

(57)摘要

本文公开了经由使用经改造的RNA支架、经改造的DNA支架或DNA-RNA杂交支架，来募集ADAR以催化点突变进行治疗性编辑的系统。该系统包括经改造的ADAR2向导RNA(adRNA)，其携带与靶标RNA的20～100bp的互补性，和在5'端和3'端的任一端或两端的来自GluR2 mRNA的ADAR2募集结构域。



CN 112996912 A

1.一种载体,包括具有编码至少一种RNA编辑实体募集结构域的多核苷酸序列的核酸,其中:

(a)所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列缺少含茎环的二级结构,或

(b)所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括与选自以下项中的至少一个序列的至少约80%的序列同一性:Alu结构域编码序列,载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域编码序列,及其组合。

2.根据权利要求1所述的载体,其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括与所述Alu结构域序列的至少约80%的序列同一性。

3.根据权利要求1所述的载体,其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括与所述APOBEC募集结构域编码序列的至少约80%的序列同一性。

4.根据权利要求1所述的载体,其中所述载体是病毒载体。

5.根据权利要求1所述的载体,其中所述载体是脂质体。

6.根据权利要求1所述的载体,其中所述载体是纳米颗粒。

7.根据权利要求1所述的载体,其中所述至少一种RNA编辑实体募集结构域被配置为募集ADAR蛋白。

8.根据权利要求7所述的载体,其中所述ADAR蛋白是ADAR1、ADAR2或ADAR3蛋白。

9.根据权利要求7或8所述的载体,其中所述ADAR蛋白是人ADAR蛋白。

10.根据权利要求7或8所述的载体,其中所述ADAR蛋白是重组的ADAR蛋白。

11.根据权利要求7或8所述的载体,其中所述ADAR蛋白是经修饰的ADAR蛋白。

12.根据权利要求1所述的载体,其中所述至少一种RNA编辑实体募集结构域被配置为募集APOBEC蛋白。

13.根据权利要求12所述的载体,其中所述APOBEC蛋白是APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H或APOBEC4蛋白。

14.根据权利要求12或13所述的载体,其中所述ADAR蛋白是人ADAR蛋白。

15.根据权利要求12或13所述的载体,其中所述ADAR蛋白是重组的ADAR蛋白。

16.根据权利要求12或13所述的载体,其中所述ADAR蛋白是经修饰的ADAR蛋白。

17.根据权利要求1所述的载体,其中所述至少一种RNA编辑实体募集结构域缺少含茎环的二级结构。

18.根据权利要求1所述的载体,其中所述多核苷酸序列编码至少两种RNA编辑募集结构域。

19.根据权利要求18所述的载体,其中所述至少两种RNA编辑募集结构域中的至少一种是Alu结构域。

20.根据权利要求19所述的载体,其中所述Alu结构域的序列形成含至少一个茎环的二级结构。

21.根据权利要求19所述的载体,其中所述Alu结构域编码序列包括多个Alu重复序列。

22.根据权利要求19所述的载体,其中所述Alu结构域编码序列至少部分是单链的。

23.根据权利要求18所述的载体,其中所述至少两种RNA编辑募集结构域中的至少一种是APOBEC募集结构域。

24.根据权利要求18所述的载体,其中所述至少两种RNA编辑募集结构域编码序列中的至少一种包括与GluR2结构域编码序列的至少约80%的序列同一性。

25.根据权利要求24所述的载体,其中所述至少两种RNA编辑募集结构域中的至少一种是GluR2结构域。

26.根据权利要求18所述的载体,其中所述至少两种RNA编辑募集结构域中的至少一种是Cas13结构域。

27.根据权利要求19或23所述的载体,其中所述至少两种RNA编辑募集结构域是Alu结构域和APOBEC募集结构域。

28.根据权利要求1至27中任一项所述的载体,进一步包括,编码与靶标RNA的至少一部分互补的RNA的核酸。

29.根据权利要求28所述的载体,其中所述编码与靶标RNA的至少一部分互补的RNA的核酸的长度为约10个碱基对(bp)至约1000bp。

30.根据权利要求28所述的载体,其中所述编码至少一种RNA编辑实体募集结构域的核酸,和所述编码与靶标RNA的至少一部分互补的RNA的核酸,包括长度至少约200bp的连续核酸。

31.根据权利要求1至30中任一项所述的载体,其中所述核酸是化学合成的。

32.根据权利要求1至30中任一项所述的载体,其中所述核酸是基因编码的。

33.根据权利要求1至30中任一项所述的载体,其中所述载体包括DNA。

34.根据权利要求33所述的载体,其中所述DNA是双链的。

35.根据权利要求33所述的载体,其中所述DNA是单链的。

36.根据权利要求1至30中任一项所述的载体,其中所述载体包括RNA。

37.根据权利要求1至30中任一项所述的载体,其中所述RNA包括碱基修饰。

38.根据权利要求1至37中任一项所述的载体,其中所述载体是腺相关病毒(AAV)载体。

39.根据权利要求38所述的载体,其中所述AAV是重组的AAV(rAAV)。

40.根据权利要求38所述的载体,其中所述AAV选自由以下项组成的组:AAV1血清型,AAV2血清型,AAV3血清型,AAV4血清型,AAV5血清型,AAV6血清型,AAV7血清型,AAV8血清型,AAV9血清型,任意这些的衍生物,和这些任一种的组合。

41.根据权利要求38所述的载体,其中所述AAV是AAV5血清型或其衍生物。

42.根据权利要求40所述的载体,包括所述AAV的衍生物,其中所述AAV的衍生物包括经修饰的VP1蛋白。

43.根据权利要求3所述的载体,其中所述APOBEC募集结构域选自由以下项组成的组:APOBEC1募集结构域、APOBEC2募集结构域、APOBEC3A募集结构域、APOBEC3B募集结构域、APOBEC3C募集结构域、APOBEC3E募集结构域、APOBEC3F募集结构域、APOBEC3G募集结构域、APOBEC3H募集结构域、APOBEC4募集结构域和其任意组合。

44.根据权利要求1至43中任一项所述的载体,其中所述至少一种RNA编辑实体募集结构域募集至少两种RNA编辑实体,并且其中,编码所述RNA编辑实体的至少两种多核苷酸序列中的至少一种包括与APOBEC蛋白编码序列的至少约80%的同一性。

45.根据权利要求1至43中任一项所述的载体,其中所述至少一种RNA编辑实体募集结构域募集至少两种RNA编辑实体,并且其中,编码所述RNA编辑实体的至少两种多核苷酸序

列中的至少一种包括与ADAR蛋白编码序列的至少约80％的同一性。

46.根据权利要求1至45中任一项所述的载体，其中由所述核酸编码的所述RNA募集结构域包括至少一个茎环。

47.根据权利要求1至46中任一项所述的载体，其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括基本上是十字形的二级结构。

48.根据权利要求1至46中任一项所述的载体，其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括至少两个基本上是十字形的二级结构。

49.根据权利要求48所述的载体，其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列，位于形成所述至少两个基本上是十字形的二级结构的多核苷酸序列之间。

50.根据权利要求47至49中任一项所述的载体，其中所述十字形的二级结构包括，与所述十字形的二级结构的至少一对至少部分互补的链邻接的茎环。

51.根据权利要求1至46中任一项所述的载体，其中所述编码至少一种RNA编辑募集结构域的多核苷酸序列包括基本上是立足点形的二级结构。

52.一种载体，包括编码具有基本上是十字形的二维形状的RNA的核酸，其中所述RNA包括编码RNA编辑实体募集结构域的至少一种序列。

53.根据权利要求52所述的载体，进一步包括编码具有至少一种靶向结构域编码序列的RNA的核酸，该至少一种靶向结构域编码序列与靶标RNA序列的至少一部分互补。

54.根据权利要求53所述的载体，其中所述编码具有至少一种靶向结构域的RNA的核酸进一步包括基本上是线性的二维结构，该至少一种靶向结构域与所述靶标RNA序列的至少一部分互补。

55.一种非天然存在的RNA，其由权利要求1至54中任一项所述的载体编码。

56.一种非天然存在的RNA，包括：第一结构域序列，包括基本上是十字形的二维形状；和，第二结构域序列，具有与所述第一结构域序列连接的基本上是线性的二维结构，

其中所述第一结构域序列编码RNA编辑实体募集结构域，且所述第二结构域序列编码靶向结构域，其中所述第二结构域序列与靶标RNA的至少一部分互补。

57.根据权利要求56所述的非天然存在的RNA，进一步包括附接至所述第二结构域序列的第三结构域序列。

58.根据权利要求57所述的非天然存在的RNA，其中所述第三结构域序列包括RNA编辑实体募集结构域编码序列，所述RNA编辑实体募集结构域编码序列形成具有二维形状基本上是十字形的二级结构。

59.根据权利要求56至59中任一项所述的非天然存在的RNA，其中所述非天然存在的RNA中的至少一个碱基包括化学修饰。

60.根据权利要求56至59中任一项所述的非天然存在的RNA，其中所述非天然存在的RNA中的至少一个糖包括化学修饰。

61.一种包括RNA编辑实体募集结构域和反义结构域序列的核酸，其中，如通过对靶标核酸进行桑格法测序所确定的，当所述核酸与RNA编辑实体和与所述反义结构域的至少一部分互补的靶标核酸接触时，以比与Cas13b蛋白或其活性片段复合的可比较核酸高至少约4倍的效率，修饰所述靶标核酸的至少一个碱基对。

62.一种核酸，包括RNA编辑实体募集结构域和反义结构域，其中，如通过对靶标核酸进

行桑格法测序所确定的，所述核酸在与RNA编辑实体和与所述反义结构域的至少一部分互补的靶标核酸接触时，以比与GluR2结构域和所述反义结构域复合的可比较核酸高至少约4倍的效率，修饰所述靶标核酸的至少一个碱基对。

63.根据权利要求61或62所述的核酸，其中所述核酸包括RNA。

64.根据权利要求51至63中任一项所述的核酸，其中所述靶标核酸包括RNA。

65.根据权利要求64所述的核酸，其中所述RNA是mRNA。

66.根据权利要求65所述的核酸，其中所述mRNA编码蛋白或其部分。

67.根据权利要求66所述的核酸，其中所述蛋白或其部分的功能失调与疾病或病症相关。

68.根据权利要求67所述的核酸，其中所述疾病或病症选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍及其任意组合所组成的组。

69.根据权利要求64所述的核酸，其中所述RNA是小干扰RNA(siRNA)。

70.根据权利要求61至69中任一项所述的核酸，其中所述RNA编辑实体募集结构域包括与GluR2结构域的至少约80％的同一性。

71.根据权利要求61至69中任一项所述的核酸，其中所述RNA编辑实体募集结构域包括与Alu结构域的至少约80％的同一性。

72.根据权利要求61至69中任一项所述的核酸，其中所述RNA编辑实体募集结构域包括与APOBEC募集结构域的至少约80％的同一性。

73.根据权利要求61至72中任一项所述的核酸，其中所述RNA编辑实体募集结构域被配置为募集ADAR蛋白。

74.根据权利要求73所述的核酸，其中所述ADAR蛋白是ADAR1、ADAR2或ADAR3蛋白。

75.根据权利要求73或74所述的核酸，其中所述ADAR蛋白是人ADAR蛋白。

76.根据权利要求73或74所述的核酸，其中所述ADAR蛋白是重组的ADAR蛋白。

77.根据权利要求73或74所述的核酸，其中所述ADAR蛋白是经修饰的ADAR蛋白。

78.根据权利要求61至72中任一项所述的核酸，其中所述RNA编辑实体募集结构域被配置为募集APOBEC蛋白。

79.根据权利要求78所述的核酸，其中所述APOBEC蛋白是APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H或APOBEC4蛋白。

80.根据权利要求78或79所述的核酸，其中所述ADAR蛋白是人ADAR蛋白。

81.根据权利要求78或79所述的核酸，其中所述ADAR蛋白是重组的ADAR蛋白。

82.根据权利要求78或79所述的核酸，其中所述ADAR蛋白是经修饰的ADAR蛋白。

83.根据权利要求61至82中任一项所述的核酸，其中所述核酸是化学合成的。

84.根据权利要求61至82中任一项所述的核酸，其中所述核酸是基因编码的。

85.一种包括序列的核酸，所述序列包括反义结构域、第一茎环形成序列和第二茎环形成序列，其中所述核酸在与以下项接触时，修饰靶标核酸的至少一个碱基对：(a)第一多肽，包括RNA编辑实体的第一部分，和被配置为与所述第一茎环形成序列结合的第一多核苷酸结合结构域；和，(b)第二多肽，包括RNA编辑实体的第二部分，和被配置为与所述第二茎环形成序列结合的第二多核苷酸结合结构域；和，(c)与所述反义结构域的至少一部分互补的靶标核酸。

86.根据权利要求85所述的核酸，其中所述第一茎环或所述第二茎环是MS2茎环。

87.根据权利要求85所述的核酸，其中所述第一茎环或所述第二茎环是BoxB茎环。

88.根据权利要求85所述的核酸，其中所述第一茎环或所述第二茎环是U1A茎环。

89.根据权利要求85至88中任一项所述的核酸，其中所述RNA编辑实体的第一部分或所述RNA编辑实体的第二部分包括ADAR脱氨酶结构域编码序列的N-末端片段。

90.根据权利要求85至88中任一项所述的核酸，其中所述RNA编辑实体的第一部分或所述RNA编辑实体的第二部分包括ADAR脱氨酶结构域编码序列的C-末端片段。

91.根据权利要求85至90中任一项所述的核酸，其中所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域包括MS2外壳蛋白。

92.根据权利要求85至90中任一项所述的核酸，其中所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域包括λN肽。

93.根据权利要求85至90中任一项所述的核酸，其中所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域包括人核酸结合蛋白。

94.根据权利要求93所述的核酸，其中所述人核酸结合蛋白是U1A蛋白、TBP6.7蛋白、人组蛋白茎环结合蛋白或糖皮质激素受体的DNA结合结构域。

95.根据权利要求85至94中任一项所述的核酸，其中所述RNA编辑实体能够在所述靶标核酸上进行腺苷到肌苷的突变。

96.根据权利要求85至94中任一项所述的核酸，其中所述RNA编辑实体能够在所述靶标核酸上进行胞嘧啶到胸腺嘧啶的突变。

97.一种试剂盒，包括在容器中的权利要求1至54中任一项所述的载体，在容器中的权利要求55至60中任一项所述的非天然存在的RNA，或在容器中的权利要求61至96中任一项所述的核酸。

98.根据权利要求97所述的试剂盒，进一步包括注射器。

99.根据权利要求98所述的试剂盒，其中所述容器是所述注射器。

100.一种分离的细胞，包括权利要求1至54中任一项所述的载体，权利要求55至60中任一项所述的非天然存在的RNA，或权利要求61至96中任一项所述的核酸。

101.一种药物组合物，包括单位剂型中的权利要求1至54中任一项所述的载体，单位剂型中的权利要求55至60中任一项所述的非天然存在的RNA，或单位剂型中的权利要求61至96中任一项所述的核酸。

102.根据权利要求101所述的药物组合物，进一步包括药学上可接受的赋形剂、稀释剂或载剂。

103.根据权利要求101所述的药物组合物，其中所述药物组合物包括第二活性成分。

104.一种治疗受试者中的疾病或病症的方法，包括向所述受试者施用权利要求1至54中任一项所述的载体，权利要求55至60中任一项所述的非天然存在的RNA，或权利要求61至96中任一项所述的核酸。

105.根据权利要求104所述的方法，其中所述施用通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。

106.根据权利要求104所述的方法，进一步包括向所述受试者施用第二疗法。

107.根据权利要求104所述的方法，其中所述疾病或病症选自由神经退行性障碍、肌肉

障碍、代谢障碍、眼部障碍及其任意组合组成的组。

108.根据权利要求107所述的方法，其中所述疾病或病症是阿尔茨海默病。

109.根据权利要求107所述的方法，其中所述疾病或病症是肌肉萎缩症。

110.根据权利要求107所述的方法，其中所述疾病或病症是色素性视网膜炎。

111.根据权利要求107所述的方法，其中所述疾病或病症是帕金森病。

112.根据权利要求107所述的方法，其中所述疾病或病症是疼痛。

113.根据权利要求107所述的方法，其中所述疾病或病症是斯塔加特黄斑营养不良。

114.根据权利要求107所述的方法，其中所述疾病或病症是腓骨肌萎缩症。

115.根据权利要求107所述的方法，其中所述疾病或病症是瑞特综合征。

116.根据权利要求104至115中任一项所述的方法，其中，与所述施用之前相比，所述施用足以减少基因的表达。

117.根据权利要求104至115中任一项所述的方法，其中所述施用足以编辑所述受试者中的至少一个点突变。

118.根据权利要求104至115中任一项所述的方法，其中所述施用足以编辑所述受试者中的至少一个终止密码子，从而产生所述终止密码子的通读。

119.根据权利要求104至115中任一项所述的方法，其中所述施用足以在所述受试者中产生外显子跳跃。

120.一种治疗受试者中的肌肉萎缩症的方法，包括向所述受试者施用包括腺相关病毒(AAV)载体的药物组合物，所述腺相关病毒(AAV)载体包括编码第二核酸的第一核酸，其中所述第二核酸包括(a)与肌肉萎缩症有关的RNA序列至少部分互补的反义区域，和(b)至少一种RNA编辑实体募集结构域，其中所述至少一种RNA编辑实体募集结构域不包括茎环，或其中所述至少一种RNA编辑实体募集结构域包括与以下项中的至少一种的至少约80％的序列同一性：Alu结构域、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域及其任意组合。

121.根据权利要求120所述的方法，其中所述药物组合物在单位剂型中。

122.根据权利要求120或121所述的方法，其中所述施用为一周至少一次。

123.根据权利要求120或121所述的方法，其中所述施用为一个月至少一次。

124.根据权利要求120至123中任一项所述的方法，其中所述施用是通过注射进行。

125.根据权利要求124所述的方法，其中所述注射为皮下、静脉内、输注、肌内、鞘内或腹腔注射。

126.根据权利要求120至123中任一项所述的方法，其中所述施用为经皮、经粘膜、经口或经肺部施用。

127.根据权利要求120至126中任一项所述的方法，进一步包括向所述受试者施用第二疗法。

128.一种制造载体的方法，包括：将至少一个拷贝的核酸克隆到所述载体中，其中所述核酸编码至少一种RNA编辑实体募集结构域，且其中编码所述至少一种RNA编辑实体募集结构域的序列不形成包括茎环的二级结构，或其中编码所述至少一种RNA编辑实体募集结构域的核酸包括与选自以下项中的序列的至少约80％的序列同一性：Alu结构域编码序列、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域编码序列及其任意组合。

129.根据权利要求128所述的方法,其中所述载体为病毒载体。

130.根据权利要求129所述的方法,其中所述病毒载体为AAV载体。

131.根据权利要求129或130所述的方法,其中所述病毒载体包括经修饰的VP1蛋白。

132.根据权利要求128所述的方法,其中所述载体为脂质体。

133.根据权利要求128所述的方法,其中所述载体为纳米颗粒。

134.根据权利要求128至133中任一项所述的方法,进一步包括将所述载体转染或转导入分离的人细胞中。

经由改造的ADAR募集的RNA和DNA碱基编辑

[0001] 交叉引用

[0002] 本申请要求享有于2018年9月6日提交的美国临时申请62/728,007、于2018年10月17日提交的美国临时申请62/766,433和于2018年12月15日提交的美国临时申请62/780,241的权益,其全部内容通过引用并入本文中。

[0003] 关于政府支持的声明

[0004] 本公开是在政府的支持下,以美国国立卫生研究院(National Institutes of Health)授予的拨款编号CA222826、GM123313和HG009285进行的。政府拥有本发明的某些权利。

发明内容

[0005] 本公开的一方面提供了一种载体。在某些情况下,所述载体可包括具有编码至少一种RNA编辑实体募集结构域的多核苷酸序列的核酸,其中:(a)所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列不形成含茎环的二级结构,或(b)其中所述编码至少一种RNA编辑实体募集结构域的多核苷酸序列包括与选自以下项中的至少一个序列的至少约80%的序列同一性:Alu结构域编码序列,载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域编码序列,及其任意组合。在某些情况下,所述编码RNA编辑实体募集结构域的多核苷酸序列可以包括与Alu结构域序列的至少约80%的序列同一性。在某些情况下,所述编码RNA编辑实体募集结构域的多核苷酸序列可以包括与APOBEC募集结构域编码序列的至少约80%的序列同一性。在某些情况下,所述载体可以是病毒载体。在某些情况下,所述载体可以是脂质体。在某些情况下,所述载体可以是纳米颗粒。在某些情况下,所述RNA编辑实体募集结构域可以被配置为募集ADAR蛋白。在某些情况下,所述ADAR蛋白可以是ADAR1、ADAR2或ADAR3蛋白。在某些情况下,所述ADAR蛋白可以是人ADAR蛋白。在某些情况下,所述ADAR蛋白可以是重组的ADAR蛋白。在某些情况下,所述ADAR蛋白可以是经修饰的ADAR蛋白。在某些情况下,所述RNA编辑实体募集结构域可以被配置为募集APOBEC蛋白。在某些情况下,所述APOBEC蛋白可以是APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H或APOBEC4蛋白。在某些情况下,所述ADAR蛋白可以是人ADAR蛋白。在某些情况下,所述ADAR蛋白可以是重组的ADAR蛋白。在某些情况下,所述ADAR蛋白可以是经修饰的ADAR蛋白。在某些情况下,所述至少一种RNA编辑实体募集结构域可以不形成含茎环的二级结构。在某些情况下,所述多核苷酸序列可以编码至少两种RNA编辑募集结构域。在某些情况下,所述至少两种RNA编辑募集结构域中的至少一种可以是Alu结构域。在某些情况下,所述Alu结构域序列可以形成包括至少一个茎环的二级结构。在某些情况下,所述Alu结构域编码序列可以包括多个Alu重复序列。在某些情况下,所述Alu结构域编码序列可以是至少部分单链的。在某些情况下,所述至少两种RNA编辑募集结构域中的至少一种可以是APOBEC募集结构域。在某些情况下,所述至少两种RNA编辑募集结构域编码序列中的至少一种可以包括与GluR2结构域编码序列的至少约80%的序列同一性。在某些情况下,所述至少两种RNA编辑募集结构域中的至少一种可以是GluR2结构域。在某些情况下,所述至少两种

RNA编辑募集结构域中的至少一种可以是Cas13结构域。在某些情况下，所述至少两种RNA编辑募集结构域可以是Alu结构域和APOBEC募集结构域。在某些情况下，所述载体可以进一步包括编码可以与靶标RNA的至少一部分互补的RNA的核酸。在某些情况下，所述编码可以与靶标RNA的至少一部分互补的RNA的核酸可以为约10个碱基对(bp)至约1000bp的长度。在某些情况下，所述编码至少一种RNA编辑实体募集结构域的核酸和所述编码可以与靶标RNA的至少一部分互补的RNA的核酸，可以包括长度上至少约200bp的连续核酸。在某些情况下，所述核酸可以是化学合成的。在某些情况下，所述核酸可以是基因编码的。在某些情况下，所述载体可以包括DNA。在某些情况下，所述DNA可以是双链的。在某些情况下，所述DNA可以是单链的。在某些情况下，所述载体可以包括RNA。在某些情况下，所述RNA可以包括碱基修饰。在某些情况下，所述载体可以是腺相关病毒(AAV)载体。在某些情况下，所述AAV可以是重组的AAV(rAAV)。在某些情况下，所述AAV可以选自由AAV1血清型、AAV2血清型、AAV3血清型、AAV4血清型、AAV5血清型、AAV6血清型、AAV7血清型、AAV8血清型、AAV9血清型、这些任何的衍生物，和这些任何的组合组成的组。在某些情况下，所述AAV可以是AAV5血清型或其衍生物。在某些情况下，所述AAV的衍生物可以包括经修饰的VP1蛋白。在某些情况下，所述APOBEC募集结构域可以选自由APOBEC1募集结构域、APOBEC2募集结构域、APOBEC3A募集结构域、APOBEC3B募集结构域、APOBEC3C募集结构域、APOBEC3E募集结构域、APOBEC3F募集结构域、APOBEC3G募集结构域、APOBEC3H募集结构域、APOBEC4募集结构域及其任意组合组成的组。在某些情况下，所述RNA编辑实体募集结构域可以募集至少两种RNA编辑实体，并且其中所述至少两种编码RNA编辑实体的多核苷酸序列中的至少一种包括与APOBEC蛋白编码序列的至少约80％的同一性。在某些情况下，所述RNA编辑实体募集结构域可以募集至少两种RNA编辑实体，并且其中所述至少两种编码RNA编辑实体的多核苷酸序列中的至少一种包括与ADAR蛋白编码序列的至少约80％的同一性。在某些情况下，由所述核酸编码的RNA募集结构域可以包括至少一个茎环。在某些情况下，所述编码RNA编辑募集结构域的多核苷酸序列可以包括可以基本上是十字形的二级结构。在某些情况下，所述编码RNA编辑募集结构域的多核苷酸序列可以包括至少两个基本上是十字形的二级结构。在某些情况下，所述编码RNA编辑实体募集结构域的多核苷酸序列，可以位于形成所述至少两个基本上是十字形的二级结构的多核苷酸序列之间。在某些情况下，所述十字形的二级结构可以包括，与所述十字形的二级结构的至少一对至少部分互补的链邻接的茎环。在某些情况下，所述编码RNA编辑募集结构域的多核苷酸序列可以包括基本上是立足点形的二级结构。在某些情况下，非天然存在的RNA可以由所述载体来编码。在某些情况下，试剂盒可以包括在容器中的所述载体。在某些情况下，所述试剂盒可以进一步包括注射器。在某些情况下，所述容器可以是所述注射器。在某些情况下，分离的细胞可以包括所述载体。在某些情况下，药物组合物可以包括在单位剂型中的所述载体。在某些情况下，所述药物组合物可以进一步包括药学上可接受的赋形剂、稀释剂或载剂。在某些情况下，所述药物组合物可以包括第二活性成分。在某些情况下，治疗受试者中的疾病或病症的方法包括向所述受试者施用所述载体。在某些情况下，所述施用可以是通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。在某些情况下，所述方法可以进一步包括向所述受试者施用第二疗法。在某些情况下，所述疾病或病症可以选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下，所述疾病或病症可以是阿尔茨

海默病。在某些情况下,所述疾病或病症可以是肌肉萎缩症。在某些情况下,所述疾病或病症可以是色素性视网膜炎。在某些情况下,所述疾病或病症可以是帕金森病。在某些情况下,所述疾病或病症可以是疼痛。在某些情况下,所述疾病或病症可以是斯塔加特黄斑营养不良。在某些情况下,所述疾病或病症可以是腓骨肌萎缩症。在某些情况下,所述疾病或病症可以是瑞特综合征。在某些情况下,与所述施用之前相比,所述施用可足以减少基因的表达。在某些情况下,所述施用可足以编辑所述受试者中的至少一个点突变。在某些情况下,所述施用可足以编辑所述受试者中的至少一个终止密码子,从而产生所述终止密码子的通读。在某些情况下,所述施用可足以在所述受试者中产生外显子跳跃。

[0006] 本公开的另一方面提供了一种载体。在某些情况下,所述载体可以包括编码RNA的核酸,该RNA具有基本上可以是十字形的二维形状,其中该RNA包括编码RNA编辑实体募集结构域的至少一种序列。在某些情况下,所述载体可进一步包括编码RNA的核酸,该RNA具有至少一种可与靶标RNA序列的至少一部分互补的靶向结构域编码序列。在某些情况下,所述编码RNA的核酸可进一步包括基本上线性的二维结构,其中该RNA具有至少一种可与靶标RNA序列的至少一部分互补的靶向结构域。在某些情况下,所述载体可以编码非天然存在的RNA。在某些情况下,试剂盒可以包括在容器中的所述载体。在某些情况下,所述试剂盒可以进一步包括注射器。在某些情况下,所述容器可以是所述注射器。在某些情况下,分离的细胞可以包括所述载体。在某些情况下,药物组合物可以包括在单位剂型中的所述载体。在某些情况下,所述药物组合物可以进一步包括药学上可接受的赋形剂、稀释剂或载剂。在某些情况下,所述药物组合物可以包括第二活性成分。在某些情况下,治疗受试者中的疾病或病症的方法包括向所述受试者施用所述载体。在某些情况下,所述施用可以通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。在某些情况下,所述方法可以进一步包括向所述受试者施用第二疗法。在某些情况下,所述疾病或病症可以选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下,所述疾病或病症可以是阿尔茨海默病。在某些情况下,所述疾病或病症可以是肌肉萎缩症。在某些情况下,所述疾病或病症可以是色素性视网膜炎。在某些情况下,所述疾病或病症可以是帕金森病。在某些情况下,所述疾病或病症可以是疼痛。在某些情况下,所述疾病或病症可以是斯塔加特黄斑营养不良。在某些情况下,所述疾病或病症可以是腓骨肌萎缩症。在某些情况下,所述疾病或病症可以是瑞特综合征。在某些情况下,与所述施用之前相比,所述施用可足以减少基因的表达。在某些情况下,所述施用可足以编辑所述受试者中的至少一个点突变。在某些情况下,所述施用可足以编辑所述受试者中的至少一个终止密码子,从而产生所述终止密码子的通读。在某些情况下,所述施用可足以在所述受试者中产生外显子跳跃。

[0007] 本公开的另一方面提供了非天然存在的RNA。在某些情况下,所述非天然存在的RNA可以包括:第一结构域序列,包括基本上是十字形的二维形状;和,第二结构域序列,具有与所述第一结构域序列连接的基本上线性的二维结构,其中所述第一结构域序列编码RNA编辑实体募集结构域且所述第二结构域序列编码靶向结构域,其中所述第二结构域序列可与靶标RNA的至少一部分互补。在某些情况下,所述非天然存在的RNA可进一步包括附接至所述第二结构域序列的第三结构域序列。在某些情况下,所述第三结构域序列可包括RNA编辑实体募集结构域编码序列,所述RNA编辑实体募集结构域编码序列形成具有二维形

状基本上是十字形的二级结构。在某些情况下,所述非天然存在的RNA中的至少一个碱基可包括化学修饰。在某些情况下,所述非天然存在的RNA中的至少一个糖可包括化学修饰。在某些情况下,试剂盒可包括在容器中的所述非天然存在的RNA。在某些情况下,所述试剂盒可进一步包括注射器。在某些情况下,所述容器可以是所述注射器。在某些情况下,分离的细胞可以包括所述非天然存在的RNA。在某些情况下,药物组合物可以包括在单位剂型中的所述非天然存在的RNA。在某些情况下,所述药物组合物可进一步包括药学上可接受的赋形剂、稀释剂或载剂。在某些情况下,所述药物组合物可包括第二活性成分。在某些情况下,治疗受试者中的疾病或病症的方法包括向所述受试者施用所述非天然存在的RNA。在某些情况下,所述施用可通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。在某些情况下,所述方法可进一步包括向所述受试者施用第二疗法。在某些情况下,所述疾病或病症可选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下,所述疾病或病症可以是阿尔茨海默病。在某些情况下,所述疾病或病症可以是肌肉萎缩症。在某些情况下,所述疾病或病症可以是色素性视网膜炎。在某些情况下,所述疾病或病症可以是帕金森病。在某些情况下,所述疾病或病症可以是疼痛。在某些情况下,所述疾病或病症可以是斯塔加特黄斑营养不良。在某些情况下,所述疾病或病症可以是腓骨肌萎缩症。在某些情况下,所述疾病或病症可以是瑞特综合征。在某些情况下,与所述施用之前相比,所述施用可足以减少基因的表达。在某些情况下,所述施用可足以编辑所述受试者中的至少一个点突变。在某些情况下,所述施用可足以编辑所述受试者中的至少一个终止密码子,从而产生所述终止密码子的通读。在某些情况下,所述施用可足以在所述受试者中产生外显子跳跃。

[0008] 本公开的另一方面提供了核酸。在某些情况下,所述核酸可以包括RNA编辑实体募集结构域和反义结构域序列,其中如通过对靶标核酸进行桑格法(Sanger Method)测序所确定的,当所述核酸与RNA编辑实体和与所述反义结构域的至少一部分互补的靶标核酸接触时,以比与Cas13b蛋白或其活性片段复合的可比较核酸高至少约4倍的效率,修饰所述靶标核酸的至少一个碱基对。

[0009] 本公开的另一方面提供了核酸。在某些情况下,所述核酸可包括RNA编辑实体募集结构域和反义结构域,其中,如通过对靶标核酸进行桑格法测序所确定的,所述核酸在与RNA编辑实体和与所述反义结构域的至少一部分互补的靶标核酸接触时,以比与GluR2结构域和所述反义结构域复合的可比较核酸高至少约4倍的效率,修饰所述靶标核酸的至少一个碱基对。在某些情况下,所述核酸可包括RNA。在某些情况下,所述靶标核酸可包括RNA。在某些情况下,所述RNA可以是mRNA。在某些情况下,所述mRNA可以编码蛋白或其部分。在某些情况下,所述蛋白或其部分的功能失调可以与疾病或病症相关。在某些情况下,所述疾病或病症可以选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下,所述RNA可以是小干扰RNA(siRNA)。在某些情况下,所述RNA编辑实体募集结构域可包括与GluR2结构域的至少约80%的同一性。在某些情况下,所述RNA编辑实体募集结构域可包括与Alu结构域的至少约80%的同一性。在某些情况下,所述RNA编辑实体募集结构域可与APOBEC募集结构域的至少约80%的同一性。在某些情况下,所述RNA编辑实体募集结构域可被配置为募集ADAR蛋白。在某些情况下,所述ADAR蛋白可以是ADAR1、ADAR2或ADAR3蛋白。在某些情况下,所述ADAR蛋白可以是人ADAR蛋白。在

某些情况下,所述ADAR蛋白可以是重组的ADAR蛋白。在某些情况下,所述ADAR蛋白可以是经修饰的ADAR蛋白。在某些情况下,所述RNA编辑实体募集结构域可以被配置为募集APOBEC蛋白。在某些情况下,所述APOBEC蛋白可以是APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H或APOBEC4蛋白。在某些情况下,所述ADAR蛋白可以是人ADAR蛋白。在某些情况下,所述ADAR蛋白可以是重组的ADAR蛋白。在某些情况下,所述ADAR蛋白可以是经修饰的ADAR蛋白。在某些情况下,所述核酸可以是化学合成的。在某些情况下,所述核酸可以是基因编码的。在某些情况下,试剂盒可以包括在容器中的所述核酸。在某些情况下,所述试剂盒可进一步包括注射器。在某些情况下,所述容器可以是所述注射器。在某些情况下,分离的细胞可以包括所述核酸。在某些情况下,药物组合物可以包括单位剂型中的所述核酸。在某些情况下,所述药物组合物可进一步包括药学上可接受的赋形剂、稀释剂或载剂。在某些情况下,所述药物组合物可以包括第二活性成分。在某些情况下,治疗受试者中的疾病或病症的方法包括向所述受试者施用所述核酸。在某些情况下,所述施用可以通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。在某些情况下,所述方法可进一步包括向所述受试者施用第二疗法。在某些情况下,所述疾病或病症可以选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下,所述疾病或病症可以是阿尔茨海默病。在某些情况下,所述疾病或病症可以是肌肉萎缩症。在某些情况下,所述疾病或病症可以是色素性视网膜炎。在某些情况下,所述疾病或病症可以是帕金森病。在某些情况下,所述疾病或病症可以是疼痛。在某些情况下,所述疾病或病症可以是斯塔加特黄斑营养不良。在某些情况下,所述疾病或病症可以是腓骨肌萎缩症。在某些情况下,所述疾病或病症可以是瑞特综合征。在某些情况下,与所述施用之前相比,所述施用可足以减少基因的表达。在某些情况下,所述施用可足以编辑所述受试者中的至少一个点突变。在某些情况下,所述施用可足以编辑所述受试者中的至少一个终止密码子,从而产生所述终止密码子的通读。在某些情况下,所述施用可足以在所述受试者中产生外显子跳跃。

[0010] 本公开的另一方面可提供核酸。在某些情况下,所述核酸可包括序列,该序列包括反义结构域、第一茎环形成序列和第二茎环形成序列,其中所述核酸在与以下项接触时修饰靶标核酸的至少一个碱基对:(a)第一多肽,包括RNA编辑实体的第一部分和被配置为与所述第一茎环形成序列结合的第一多核苷酸结合结构域;和,(b)第二多肽,包括RNA编辑实体的第二部分和被配置为与所述第二茎环形成序列结合的第二多核苷酸结合结构域;和,(c)与所述反义结构域的至少一部分互补的靶标核酸。在某些情况下,所述第一茎环或所述第二茎环可以是MS2茎环。在某些情况下,所述第一茎环或所述第二茎环可以是BoxB茎环。在某些情况下,所述第一茎环或所述第二茎环可以是U1A茎环。在某些情况下,所述RNA编辑实体的第一部分或所述RNA编辑实体的第二部分可以包括ADAR脱氨酶结构域编码序列的N-端片段。在某些情况下,所述RNA编辑实体的第一部分或所述RNA编辑实体的第二部分可以包括ADAR脱氨酶结构域编码序列的C-端片段。在某些情况下,所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域可以包括MS2外壳蛋白。在某些情况下,所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域可以包括λN肽。在某些情况下,所述第一多核苷酸结合结构域或所述第二多核苷酸结合结构域可以包括人核酸结合蛋白。在某些情况下,所述人核酸结合蛋白可以是U1A蛋白、TBP6.7蛋白、人组蛋白茎环结合蛋白或糖皮质激

素受体的DNA结合结构域。在某些情况下,所述RNA编辑实体可以能够在所述靶标核酸上进行腺苷到肌苷的突变。在某些情况下,所述RNA编辑实体可以能够在所述靶标核酸上进行胞嘧啶到胸腺嘧啶的突变。在某些情况下,试剂盒可以包括在容器中的所述核酸。在某些情况下,所述试剂盒可以进一步包括注射器。在某些情况下,所述容器可以是所述注射器。在某些情况下,分离的细胞可以包括所述核酸。在某些情况下,药物组合物可以包括在单位剂型中的所述核酸。在某些情况下,所述药物组合物可以进一步包括药学上可接受的赋形剂、稀释剂或载剂。在某些情况下,所述药物组合物可以包括第二活性成分。在某些情况下,治疗受试者中的疾病或病症的方法包括向所述受试者施用所述核酸。在某些情况下,所述施用可以通过静脉注射、肌内注射、鞘内注射、眶内注射、皮下注射或其任意组合来进行。在某些情况下,所述方法可以进一步包括向所述受试者施用第二疗法。在某些情况下,所述疾病或病症可以选自由神经退行性障碍、肌肉障碍、代谢障碍、眼部障碍、细胞增殖障碍(例如赘生物)及其任意组合组成的组。在某些情况下,所述疾病或病症可以是阿尔茨海默病。在某些情况下,所述疾病或病症可以是肌肉萎缩症。在某些情况下,所述疾病或病症可以是色素性视网膜炎。在某些情况下,所述疾病或病症可以是帕金森病。在某些情况下,所述疾病或病症可以是疼痛。在某些情况下,所述疾病或病症可以是斯塔加特黄斑营养不良。在某些情况下,所述疾病或病症可以是腓骨肌萎缩症。在某些情况下,所述疾病或病症可以是瑞特综合征。在某些情况下,与所述施用之前相比,所述施用可足以减少基因的表达。在某些情况下,所述施用可足以编辑所述受试者中的至少一个点突变。在某些情况下,所述施用可足以编辑所述受试者中的至少一个终止密码子,从而产生所述终止密码子的通读。在某些情况下,所述施用可足以在所述受试者中产生外显子跳跃。

[0011] 本公开的另一方面提供了一种治疗受试者中的肌肉萎缩症的方法。在某些情况下,所述方法可以包括:向所述受试者施用包括腺相关病毒(AAV)载体的药物组合物,所述腺相关病毒(AAV)载体包括编码第二核酸的第一核酸,其中所述第二核酸包括(a)可以与肌肉萎缩症有关的RNA序列至少部分互补的反义区域,和(b)至少一种RNA编辑实体募集结构域,其中所述至少一种RNA编辑实体募集结构域不包括茎环,或其中所述至少一种RNA编辑实体募集结构域包括与以下项中的至少一种的至少约80%的序列同一性:Alu结构域、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域及其任意组合。在某些情况下,所述药物组合物可以是单位剂型。在某些情况下,所述施用可以是一周至少一次。在某些情况下,所述施用可以是一个月至少一次。在某些情况下,所述施用可以是通过注射进行。在某些情况下,所述注射可以是皮下、静脉内、输注、肌内、鞘内或腹腔注射。在某些情况下,所述施用可以是经皮、经粘膜、经口或经肺部施用。在某些情况下,所述方法可以进一步包括向所述受试者施用第二疗法。

[0012] 本公开的另一方面可以提供一种制造载体的方法。在某些情况下,所述方法可以包括:将至少一个拷贝的核酸克隆到所述载体中,其中所述核酸编码至少一种RNA编辑实体募集结构域,且其中编码所述至少一种RNA编辑实体募集结构域的序列不形成包括茎环的二级结构,或其中编码所述至少一种RNA编辑实体募集结构域的所述核酸包括与选自以下项中的序列的至少约80%的序列同一性:Alu结构域编码序列、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域编码序列及其任意组合。在某些情况下,所述载体可以是病毒载体。在某些情况下,所述病毒载体可以是AAV载体。在某些情况下,所述病毒载体可以包括经

修饰的VP1蛋白。在某些情况下,所述载体可以是脂质体。在某些情况下,所述载体可以是纳米颗粒。在某些情况下,所述方法进一步包括将所述载体转染或转导入分离的人细胞中。

[0013] 本公开的多个方面涉及经改造的ADAR1或ADAR2向导RNA(“adRNA”),其包括与靶标RNA互补的序列,或可选地基本上由其组成,或进一步基本上由其组成,该序列任选地包括经改造的ADAR2和来源于GluR2 mRNA的ADAR2募集结构域,或可选地基本上由其组成,或进一步基本上由其组成。在一些实施方式中,所述与靶标RNA互补的序列包括约15至200个或可选的20至100个碱基对,或可选地基本上由其组成,或进一步由其组成。在一个方面,所述经改造的adRNA不包括ADAR募集结构域,或不基本上由其组成,或进一步不由其组成。在一些实施方式中,所述ADAR募集结构域包括GluR2 mRNA、Alu重复序列元件或ADAR结合的其他RNA基序,或可选地基本上由其组成,或进一步由其组成。在一个方面,所述经改造的adRNA包括约1至10个ADAR募集结构域,基本上由其组成,或进一步由其组成。在另一方面,所述ADAR2募集结构域可以来源于GluR2 mRNA,并且可以位于所述经改造的adRNA的5’端或3’端。在又一实施方式中,所述经改造的adRNA在5’端和3’端均包括GluR2 mRNA,或可选地基本上由其组成,或进一步由其组成。在进一步的方面,本公开的经改造的adRNA进一步包括两个在与靶标RNA互补的序列两侧的MS2发夹,或可选地基本上由其组成,或进一步由其组成。

[0014] 在一些实施方式中,所述靶标RNA可以是鸟氨酸转氨甲酰酶。本文还提供了一种复合物,其包括在高严格性条件下与互补多核苷酸杂交的本文公开的adRNA,或可选地基本上由其组成,或进一步由其组成。在一个方面,所述多核苷酸可以是DNA。在另一方面,所述多核苷酸可以是RNA。

[0015] 在一个方面,本公开的经改造的adRNA进一步包括编辑诱导元件,或可选地基本上由其组成,或进一步由其组成。

[0016] 进一步的方面涉及经改造的ADAR2向导RNA(“adRNA”),其由选自表1或图2中提供的序列组中的序列编码。所述adRNA可以与载剂,例如药学上可接受的载剂组合,本文提供了这样的实施例。

[0017] 本文还公开了一种经改造的adRNA-snRNA(小核RNA)融合体。在一个方面,所述经改造的adRNA进一步包括N-端线粒体靶向序列(MTS),或可选地基本上由其组成,或还进一步由其组成,以促进所述经改造的adRNA定位至线粒体。在另一方面,本文提供的可以是经改造的进一步包括顺式作用的拉链编码(zipcode),或可选地基本上由其组成,或进一步由其组成,以促进所述经改造的adRNA定位到过氧物酶体、内体和外泌体中。

[0018] 本文进一步提供了小分子可调节的经改造的adRNA。在一个方面,本文公开了经改造的adRNA-适体的融合体。可用于该目的的适体的非限制性实施例包括结合黄素单核苷酸、鸟嘌呤、其他天然的代谢物或糖的适体。本文还公开了完全源自人类的U1A-ADAR融合体。

[0019] 本文还公开了一种复合物,其包括在高严格条件下与互补多核苷酸杂交的本公开的经改造的adRNA,或可选地基本上由其组成,或进一步由其组成。

[0020] 本文还提供了一种载体,其包括一种或多种分离的多核苷酸序列和任选地与所述分离的多核苷酸可操作地连接的调控序列,或可选地基本上由其组成,或进一步由其组成,其中所述分离的多核苷酸序列编码本公开的经改造的adRNA。载体的非限制性实施例包括

质粒或病毒载体，例如逆转录病毒载体、慢病毒载体、腺病毒载体或腺相关病毒载体。

[0021] 本文进一步公开了一种重组细胞，进一步包含或上述载体，或可选地基本上由其组成，或进一步由其组成，其中所述经改造的adRNA可以重组表达。

[0022] 提供了包括一种或多种上述化合物和载剂的组合物。在一个实施方式中，所述组合物可以是药物组合物，因此进一步包括至少一种药学上可接受的载剂或药学上可接受的赋形剂。所述组合物被配制成用于各种递送方式，例如全身性(经口)或局部递送。

[0023] 本文还提供了一种修饰蛋白表达的方法，其包括使编码所述蛋白的多核苷酸与本公开的经改造的adRNA接触，或可选地基本上由其组成，或进一步由其组成，其中所述蛋白的表达是要被修饰的。

[0024] 另外的方面涉及治疗与异常蛋白表达相关的疾病或障碍的方法，其包括以下项，或可选地基本上由其组成，或进一步由其组成：施用有效量的任一种或多种本文公开的经改造的adRNA，和/或将有效量的任一种或多种本文公开的经改造的adRNA用于有此需要的受试者，并用于治疗与异常蛋白表达相关的疾病或障碍。在一个特定的方面，本文提供了一种治疗杜氏肌肉萎缩症的方法，其包括向需要这种治疗的受试者施用有效量的一种或多种本公开的经改造的adRNA，或可选地基本上由其组成，或进一步由其组成。

[0025] 本公开证明了在鸟氨酸转氨甲酰酶缺乏症的spf-ash小鼠模型中，对该方法的体内验证。该模型在外显子4的最后一个核苷酸中带有G->A点突变。经由AAV仅递送adRNA后，在不存在ADAR酶过表达的情况下，观察到了高达1％的点突变校正。

[0026] 其他方面涉及相同或相似的结构，其包括DNA或者DNA与RNA的组合，或可选地基本上由其组成，或进一步由其组成。其他方面涉及试剂盒，其包括上述实施方式中的任何一个或多个，以及用于体外和/或体内使用的说明书。

[0027] 本公开的多个方面可以涉及用于基因编辑的ADAR和APOBEC系统。一些方面涉及用于外显子跳跃的ADAR系统，其包括靶向内含子中的剪接受体和/或分支点的adRNA，和任选地ADAR酶。在一些实施方式中，所述ADAR酶可以是ADAR1、ADAR2，或其各自的突变体或变体。在一些实施方式中，所述突变体或变体可以选自ADAR1(E1008Q)和ADAR2(E488Q)。在一些实施方式中，所述内含子可以被包含在选自抗肌萎缩蛋白、SCN9A或鸟氨酸转氨甲酰酶的基因中。在某些情况下，所述adRNA可以选自序列集1。进一步的方面可以涉及一种治疗以异常基因表达为特征的疾病、障碍或病症的方法，其包括施用所公开的ADAR系统。在一些实施方式中，所述疾病、障碍或病症可以选自杜氏肌肉萎缩症或鸟氨酸转氨甲酰酶缺乏症。在一些实施方式中，所述疾病、障碍或病症可以与疼痛有关。

[0028] 其他方面涉及用于胞嘧啶到胸腺嘧啶编辑的APOBEC系统，其包括创建脂蛋白B mRNA样结构的成对的gRNA，和任选地APOBEC酶。在一些实施方式中，该成对的gRNA可以是序列集2中提供的成对的序列。

附图说明

[0029] 图1是示出ADAR的内源性募集的示意图。

[0030] 图2提供了稳定的支架的列表，其中稳定的支架用于评估(1)改进的效率和(2)募集ADAR1的能力(SEQ ID NO:2～9)。

[0031] 图3示出了示例性adRNA的体外和体内筛选的结果。

[0032] 图4是如何经由携带两个MS2茎环的MS2-RNA来募集MCP-APOBEC融合体或MCP-ACF-APOBEC融合体的示意图。如该图所示，靶标胞嘧啶可以保持单链，以便经由产生凸起可用于APOBEC介导的编辑。通过使用在MS2-RNA中在MS2茎环之间的精确的靶序列(20～30个碱基对)或序列ACATATATGATACAATTTGATCAGTATATT(SEQ ID NO:175)(蓝色)，以及在MS2茎环任一侧上与目标mRNA结合的互补序列(20～30碱基对)(绿色)，来创建凸起。序列ACATATATGATACAATTTGATCAGTATATT(SEQ ID NO:176)获自APOBEC编辑的天然存在的apoB底物。在该设计中使用的MS2茎环序列是aACATGAGGATCACCCATGTc(SEQ ID NO:177)。

[0033] 图5示出了与通过含GluR2结构域的构建体进行编辑相比，通过含Alu结构域的构建体进行的百分比mRNA编辑。

[0034] 图6示出了用于生成图5的数据的、含Alu结构域的构建体(SEQ ID NO:10)的实施例。

[0035] 图7示出了一种构建体的实施例，该构建体包括通过反义结构域连接的两个十字形，且包括Alu结构域。提供了示例性结构(SEQ ID NO:11)和示例性序列(SEQ ID NO:1)。

[0036] 图8示出了多种构建体设计(包括含GluR2结构域的那些构建体)(SEQ ID NO:12～34)和各自的百分比mRNA编辑。

[0037] 图9示出了定位ADAR1和ADAR2对于每一种碱基(例如A、C、G和T)的酶靶向偏好。

[0038] 图10示出了改造具有增强的ADAR1和ADAR2募集潜力的下一代adRNA。第一系列柱子示出了没有利用adRNA的相对活性。第二系列柱子示出了利用含GluR2结构域的构建体的相对活性。第三系列柱子示出了利用含Alu结构域的构建体的相对活性，其中Alu结构域与两个十字形结构相关。

[0039] 图11A～图11D示出了不同的构建体设计。图11A例示了与两个GluR2结构域连接的反义结构域。图11B仅例示了反义结构域。图11C例示了与两个十字形连接的反义结构域。图11D例示了立足点形(toe-hold)。

[0040] 图12示出了不同的构建体设计之间的百分比编辑产量。

[0041] 图13示出了具有错配凸起的短的反义寡核苷酸(AON)与含发夹结构的较长构建体的对比。

[0042] 图14示出了示例性adRNA设计(SEQ ID NO:250和251)。

[0043] 图15示出了具有参数(d、l、m)的示例性adRNA结构(SEQ ID NO:35～37)，其中d＝GluR2结构域的数量，l＝反义结构域的长度，m＝错配的位置。

[0044] 图16示出了在adRNA的存在下，经由内源性ADAR的募集进行RNA编辑的示意图。

[0045] 图17示出了U6启动子转录的反义结构域长度逐渐增加的adRNA，与0、1或2个GluR2结构域结合，评估其在有或没有外源性ADAR2表达的情况下诱导靶向RNA编辑的能力。数值代表平均值+/-SEM(n＝3)。长adRNA可以募集用于RNA编辑的内源性ADAR。

[0046] 图18示出了在有或没有外源性ADAR2表达的情况下，针对一组mRNA测试的化学合成的adRNA型式。鉴定化学修饰以及adRNA的来源。数值代表平均值+/-SEM(n＝3)。

[0047] 图19示出了在经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射)中的正确剪接的OTC mRNA中的体内RNA校正效率。在注射了U6转录的短adRNA的小鼠中看到0.6％的RNA编辑水平。

[0048] 图20示出了Alu adRNA的设计。左：Alu元件的结构。中：如本文所述的设计，其包括

基因座特异性的反义序列,具有与靶标A相对的C错配。右:RNA编辑酶ADAR至靶标的募集。

[0049] 图21示出了示例性Alu向导序列(SEQ ID NOs:38～41)。

[0050] 图22示出了拆分式ADAR2 DD系统的示意图。

[0051] 图23示出了拆分式ADAR2 DD的示例性序列(SEQ ID NO:42),其中拆分式ADAR2 DD具有高亮的用于拆分的潜在位点。

[0052] 图24示出了经由海荧(cypridina)荧光素酶报告因子(Cluc W85X)测定的成对的片段1～16。

[0053] 图25示出了针对Cluc报告因子测定的片段9和片段10。

[0054] 图26示出了示例性序列(SEQ ID NO:43～55)。

[0055] 图27示出了经由U1A(SEQ ID NO:56)的ADAR募集的示意图。

[0056] 图28示出了几种融合构建体的示例性序列(SEQ ID NO:57～68)或一个或多个APOBEC家族成员的示例性序列(SEQ ID NO:57～68)。

[0057] 图29示出了用于募集APOBEC3A的经改造的apRNA的示例性序列(SEQ ID NO:69～73)。

[0058] 图30示出了用于募集MCP-APOBEC3A的经改造的MS2-apRNA的示例性序列(SEQ ID NO:74～78)。

[0059] 图31示出了无ADAR募集和ADAR募集的两种不同的方案,其中ADAR募集允许导致正常荧光素酶表达的核糖体通读。

[0060] 图32A～图32C示出了改造可编程的RNA编辑并表征特异性概况:(图32A)经由利用全长的ADAR2和经改造的adRNA的构建体进行RNA编辑的示意图,其中经改造的adRNA源自于GluR2转录物,或与ADAR1/2脱氨酶结构域和带有adRNA的相应的MS2发夹的MS2外壳蛋白(MCP)融合体。(图32B)通过Sanger测序量化的不同RNA编辑构建体对内源性RAB7A转录物的RNA编辑效率的比较(效率以Sanger峰高G/(A+G)之比来计算)。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n＝3)。(图32C)代表在参考位点上观察到的A->G编辑产量分布的小提琴图,其中发现至少一个处理样品的编辑产量相对于对照样品具有明显变化(费希尔精确检验,FDR＝1%)。蓝色圆圈表示在RAB7A转录物内的目标A位点处的编辑产量。黑色点表示中值脱靶编辑产量。为了使分布的形状更好地可视化,将它们沿y轴的最大范围在所有图上进行了均衡化,并截断在60%的产量处。

[0061] 图33A～图33E示出了在人类疾病的小鼠模型中的体内RNA编辑。(图33A)在杜氏肌肉萎缩症的mdx小鼠模型中恢复抗肌萎缩蛋白表达的DNA和RNA靶向方法的示意图:(i)导致框内切除外显子23的基于双gRNA-CRISPR的方法,和(ii)赭石(ochre)密码子的基于ADAR2和MCP-ADAR1的编辑。(图33B)TA肌肉中的抗肌萎缩蛋白的免疫荧光染色示出了在经处理的样品(肌内注射AAV8-ADAR2、AAV8-ADAR2(E488Q)和AAV8-CRISPR)中表达的部分恢复。在经处理的样品中也看到nNOS定位的部分恢复(比例尺:250μm)。(图33C)在相应的经处理的成年mdx小鼠中,体内TAA->TGG/TAG/TGA的RNA编辑效率。数值代表平均值+/-SEM(分别为n＝4、3、7、3、3、10、3、4个独立的TA肌肉)。(图33D)在鸟氨酸转氨甲酰酶缺乏症的spf^{ash}小鼠模型中的OTC基因座和经由ADAR2介导的RNA编辑校正突变型OTC mRNA的方法的示意图,其中鸟氨酸转氨甲酰酶缺乏症在外显子4的最后一个核苷酸的供体剪接位点处具有G->A点突变。(图33E)在经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射AAV8-ADAR2和AAV8-ADAR2

(E488Q))中,正确剪接的OTC mRNA中的体内RNA校正效率。数值代表平均值+/-SEM(分别为n=4、4、3、3、4、5个独立的动物)。

[0062] 图34A～图34C示出了反义结构域改造。(图34A)使用adRNA支架2来优化adRNA反义区域:长度和与ADAR2募集区的距离在系统上是不同的。数值代表平均值+/-SEM(n=3)(SEQ ID NO:79～102)。(图34B)评估了U6启动子转录的反义结构域长度逐渐增加的adRNA,结合0、1或2个GluR2结构域,其在有或没有外源性ADAR2表达的情况下诱导靶向RNA编辑的能力。数值代表平均值+/-SEM(n=3)。以上所有实验均在HEK 293T细胞中进行。(图34C)实验确认了内源性ADAR1和ADAR2(相对于GAPDH)在HEK 293T和HeLa细胞系中的表达。观察到的水平与在《人类蛋白质图谱(Human Protein Atlas)》(参见proteinatlas.org上的全球性网页(www))中记录的水平相似。

[0063] 图35A～图35B示出了改造MS2 adRNAs。(图35A)系统评估了MS2 adRNA(SEQ ID NO:103～110)的反义RNA靶向结构域。数值代表平均值+/-SEM(n=3)。(图35B)通过MCP-ADAR2 DD-NLS进行的上靶RNA编辑需要MS2 adRNA的共表达。数值代表平均值+/-SEM(n=3)。所有实验均在HEK 293T细胞中进行。

[0064] 图36A～C示出了对一组靶标的RNA编辑产量的分析。(图36A)比较OTC报告因子转录物通过GluR2 adRNA和MS2 adRNA导向RNA编辑构建体,以及基于Cas13b的REPAIR构建体,进行RNA编辑的效率。数值代表平均值+/-SEM(对于报告因子和基于Cas13b的构建体,n=6,对于其他构建体,n=3)。(图36B)针对一组mRNA,在有或没有外源性ADAR2表达的情况下,测试了化学合成的adRNA型式。在图中说明了确切的化学修饰以及adRNA的来源。数值代表平均值+/-SEM(n=3)。(图36C)分析所选择的一连串内源性靶标的RNA编辑产量,以涵盖一系列表达水平。还针对多个内源性mRNA靶标,在有或没有外源性ADAR2表达的情况下,评估了U6转录的没有或具有两个GluR2结构域的长adRNA。甚至在没有外源性ADAR2表达的情况下,也可以在测试的基因座处观察到编辑。数值代表平均值+/-SEM(n=3)。实验在HEK 293T细胞中进行。

[0065] 图37A～D示出了ADAR2变体及其对编辑和特异性的影响。(图37A)比较了RAB7A转录物通过GluR2 adRNA和MS2 adRNA导向RNA编辑构建体,以及基于Cas13b的REPAIR构建体,进行RAB7A转录物的上靶RNA编辑和侧翼腺苷中的编辑。描绘了平均值(n=3)编辑产量(SEQ ID NO:111)。所有实验均在HEK 293T细胞中进行,并且编辑效率计算为Sanger峰高G/(A+G)之比。(图37B)在spf^{ash} OTC报告因子转录物(p=0.037,未配对t检验,双尾;数值代表平均值+/-SEM(n=3))和(图37C)mdx DMD报告因子转录物(分别为p=0.048,p=0.012,未配对t检验,双尾;数值代表平均值+/-SEM(n=3))的体外编辑中,ADAR2(E488Q)显示出比ADAR2更高的效率。(图37D)比较了ADAR2、ADAR2(E488Q)和ADAR2(Δ1-138),对于OTC报告因子转录物(上图)和内源性RAB7A转录物(下图)的编辑效率和特异性概况。热图指示在目标附近的A->G编辑(红色箭头)。数值代表平均值+/-SEM(n=3)。所有实验均在HEK 293T细胞中进行,并将编辑效率计算为Sanger峰高G/(A+G)之比。

[0066] 图38示出了RNA编辑方法(Cas13b-ADAR REPAIR+/-gRNA)的转录组规模的特异性概况。

[0067] 图39示出了RNA编辑方法(ADAR2+/-adRNA)的转录组规模的特异性概况。这些研究使用的型式是GluR2 adRNA(1,20,6)。

[0068] 图40示出了RNA编辑方法(MCP-ADAR1 DD+/-adRNA)的转录组规模的特异性概况。

[0069] 图41示出了RNA编辑方法(MCP-ADAR2 DD+/-adRNA)的转录组规模的特异性概况。

[0070] 图42A～B示出了具有构建体特征的转录组规模编辑的变化。(图42A)箱线图中的每个点都对应于图32中列出的MCP-ADAR构建体之一的经编辑的位点的分数。每个构建体的经编辑的位点的分数,通过将A-至-G编辑产量(见表3)发生显著变化的参考位点数目除以考虑的参考位点总数目8,729,464来计算得出。使用曼-惠特尼U检验比较了水平轴上指示的构建体特征,产生了NLS vs.NES的p值为0.16,ADAR1vs.ADAR2的p值为0.0070,“-adRNA”vs.“+adRNA”的p值为0.72,以及“ADAR WT”vs.“ADAR E>Q”的p值为0.038(在所有条件下n＝8)。(图42B)2D直方图比较了使用每种构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。插图示出了小提琴图,其代表在参考位点观察到的A->G编辑产量的分布,其中相对于对照样品,发现至少一个处理样品的编辑产量具有显著变化(费希尔精确检验,FDR＝1%)。蓝色圆圈表示RAB7A转录物内在目标A-位点处的编辑产量。为了使分布的形状更好地可视化,将它们沿y轴的最大范围在图上进行了均衡化,并截断在60%的产量处。此处的样品对应于转染了携带adRNA的长反义结构域的293T,该adRNA可以经由外源性和/或内源性ADAR募集进行RNA编辑。

[0071] 图43A～E示出了对mdx小鼠中的体外和体内抗肌萎缩蛋白RNA编辑实验的优化和评估。(图43A)利用全长ADAR2以及改造的adRNA或反向adRNA(radRNA)进行RNA编辑的示意图;(ii)琥珀(amber)和赭石(ochre)终止密码子以一步和两步的方式进行的RNA编辑效率。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n＝3)。(图43B)赭石密码子的RNA编辑需要在adRNA或radRNA(SEQ ID NO:112～116)的反义RNA靶向结构域中的两个胞嘧啶错配,来恢复GFP表达。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n＝3)。(图43C)用于在体内递送adRNA和ADAR2的AAV载体的示意图,以及在存在1或2个拷贝的adRNA的情况下,经由AAV载体递送的琥珀和赭石终止密码子的RNA编辑的体外优化(分别为p＝0.0003,p＝0.0001,p＝0.0015,不成对t检验,双尾)。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(对于报告因子,n＝3,对于其他条件,n＝6)。(图43D)代表性的Sanger测序图,显示了在mdx DMD报告因子转录物中赭石终止密码子(TAA->TGG)的编辑(通过NGS进行量化)(SEQ ID NO:117～118)。实验在HEK 293T细胞中进行(n＝3)。(图43E)经处理的mdx小鼠的体内RNA编辑分析的代表性实施例(使用NGS进行量化)(SEQ ID NO:119～130)。

[0072] 图44A～C示出了在mdx小鼠中的体内抗肌萎缩蛋白RNA编辑实验的免疫荧光和免疫印迹的分析。(图44A)在TA肌肉中对抗肌萎缩蛋白的免疫荧光染色示出了经处理的样品(肌肉内注射AAV8-ADAR2、AAV8-ADAR2(E488Q)、AAV8-MCP-ADAR1(E1008Q)NLS)中表达的部分恢复。在经处理的样品中还看到nNOS定位的部分恢复(比例尺:250μm)。(图44B)免疫印迹示出了在mdx小鼠(其注射有编辑工具的两种组分,酶和adRNA)的TA肌肉中抗肌萎缩蛋白表达(1～2.5%)的部分恢复,以及在注射的TA肌肉中注射后长达8周的稳定的ADAR2表达。(图44C)免疫印迹示出了使用AAV8-CRISPR使抗肌萎缩蛋白表达部分恢复(10%)。

[0073] 图45A～E示出了在spf^{ash}小鼠中体外和体内的OTC RNA编辑实验的优化和评估。(图45A)代表性Sanger测序图谱示出了在spf^{ash} OTC报告因子转录物中点突变的校正(使用NGS进行了量化)(SEQ ID NO:131-132)。实验在HEK 293T细胞中进行(n＝3)。(图45B)经处理的spf^{ash}小鼠的体内RNA编辑分析的代表性实施例示出了,在经正确剪接的OTC mRNA中的

点突变的校正(使用NGS进行了量化)(SEQ ID NO:133-139)。(图45C)在经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射A488-ADAR2和AAV8-ADAR2(E488Q))中的OTC pre-mRNA的体内RNA校正效率。数值代表平均值+/-SEM(分别为n＝4、4、3、3、4、5只独立的动物)。(图45D)PCR产物示出了经正确和错误剪接的OTC mRNA。经错误剪接的mRNA被延长了48个碱基对。在用adRNA+ADAR2(E488Q)处理的小鼠中,经错误剪接的mRNA的分数降低了。(图45E)对OTC的免疫印迹示出了,对经处理的成年spf^{ash}小鼠中表达的部分恢复(2.5%～5%),和在注射后三周稳定的ADAR2(E488Q)表达。

[0074] 图46示出了对体内RNA编辑实验的毒性分析。

[0075] 图47是示出经由创建剪接受体和/或分支点突变进行外显子跳跃的示意图。

[0076] 图48是经由APOBEC进行C→T编辑的示意图。

[0077] 图49A～图49D示出了编辑DNA以及DNA /RNA杂交体的两条链的示意图。

[0078] 图50A～图50D示出了鸟氨酸转氨甲酰酶缺乏症模型的研究结果。图50A描绘了在经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射AAV8-ADAR2和AAV8-ADAR2(E488Q))中的体内RNA校正效率。每个数据点代表一只独立的动物。在经剪接的OTC mRNA中测得的编辑效率。误差棒代表+/-SEM。图50B描绘了在经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射A488-ADAR2和AAV8-ADAR2(E488Q))中OTC pre-mRNA的体内RNA校正效率。每个数据点代表一只独立的动物。图50C示出了PCR产物,其显示经正确和错误剪接的OTC mRNA。经错误剪接的mRNA延长了48个碱基对。在用adRNA +ADAR2(E488Q)处理的小鼠中,减少了经错误剪接的mRNA的分数。图50D是OTC的免疫印迹,其示出在经处理的成年spf^{ash}小鼠中表达的部分恢复(2.5%～5%)。

[0079] 图51A～图51B显示了在杜氏肌肉萎缩症模型中的研究结果。图51A描绘了在相应的经处理的成年mdx小鼠中,体内TAA->TGG/TAG/TGA的RNA编辑效率。每个数据点代表一个独立的TA肌肉。误差棒代表+/-SEM。图51B是对抗肌萎缩蛋白的免疫印迹,示出了在相应的经处理的成年mdx小鼠中表达的部分恢复(1～2.5%)。

[0080] 图52提供了关于潜在分支点位置的更多信息。

具体实施方式

[0081] 作为本公开的一个方面,提供了核酸、非天然存在的RNA、包括核酸的载体、组合物和用于RNA编辑的药物组合物。以上或如本文所述的任一个可以被配置用于A(腺苷)至I(肌苷)编辑,C(胞嘧啶)至T(胸腺嘧啶)编辑或其组合。在某些情况下,A至I编辑可解释为或读为C至U突变。在某些情况下,A至I编辑可解释为或读为A至G突变。与天然系统相比,本文所述的核酸、非天然存在的RNA、包括核酸的载体、组合物和药物组合物可提供增强的编辑效率、减少的脱靶编辑、增强的稳定性或体内半衰期,或它们的任意组合。

[0082] 本公开的一个方面提供了一种载体。所述载体可以包括具有多核苷酸序列的核酸,其编码(i)RNA编辑实体募集结构域,(ii)与靶标RNA的至少一部分互补的靶向结构域,(iii)一种以上的任何结构域,或(iv)它们的任意组合。在某些情况下,可以将所述载体施用给受试者,例如有此需要的受试者。在某些情况下,可以将所述载体作为药物组合物的一部分施用给受试者,例如有此需要的受试者。

[0083] 本公开的一个方面提供了非天然存在的RNA。非天然存在的RNA可以包括(i)RNA编

辑实体募集结构域,(ii)与靶标RNA的至少一部分互补的靶向结构域,(iii)一种以上的任何结构域,或(iv)它们的任意组合。在某些情况下,可以将所述非天然存在的RNA施用给受试者,例如有此需要的受试者。在某些情况下,可以将所述非天然存在的RNA作为药物组合物的一部分施用给受试者,例如有此需要的受试者。在某些情况下,可以将所述非天然存在的RNA配制在用于施用的载体中。所述载体可以包括病毒载体、脂质体、纳米颗粒或其任意组合。在某些情况下,所述非天然存在的RNA可以包括具有修饰(例如化学修饰)的至少一种碱基、至少一种糖、一种以上的任一种或它们的组合。

[0084] 本公开的一个方面提供了核酸。所述核酸可以包括(i)RNA编辑实体募集结构域,(ii)与靶标RNA的至少一部分互补的靶向结构域,(iii)一种以上的任何结构域,或(iv)它们的任何组合。在某些情况下,可以将所述核酸施用给受试者,例如有此需要的受试者。在某些情况下,可以将所述核酸作为药物组合物的一部分施用给受试者,例如有此需要的受试者。在某些情况下,可以将核酸配制在用于施用的载体中。所述载体可以包括病毒载体、脂质体、纳米颗粒或其任何组合。所述核酸可以是基因编码的。所述核酸可以是化学合成的。

[0085] 核酸可以包括一种或多种结构域,例如1、2、3、4、5或更多种结构域。在某些情况下,核酸可以包括募集结构域、靶向结构域、一种以上的任何结构域或它们的组合。在某些情况下,核酸可以包括靶向结构域和募集结构域。在某些情况下,核酸可以包括靶向结构域和两个募集结构域。

[0086] 结构域可以形成二维形状或二级结构。例如,靶向结构域、募集结构域或其组合可以形成二级结构,该二级结构可以包括线性区域、十字形或其部分、立足点形,茎环或其任何组合。所述结构域自身可以形成基本上线性的二维结构。所述结构域可以形成可以包括十字形的二级结构。所述结构域可以形成可以包括茎环的二级结构。所述结构域可以形成可以包括立足点形的二级结构。

[0087] 在某些情况下,靶向结构域可以位于募集结构域附近,包括直接相邻,或相邻但被多个核苷酸分隔开。在某些情况下,靶向结构域的两侧可以是两个募集结构域。在某些情况下,两个或更多个募集结构域可以是彼此相邻的。

[0088] 本公开的一个方面包括减少脱靶编辑。本文所述的一种方法包括,通过拆分重组方法,限制ADAR或APOBEC的催化活性。在这样的设计中,第一结构域(例如募集结构域)本身可以是催化失活的,且第二结构域本身可以是催化失活的,但是当重新组装在一起时,这两个结构域一起提供催化活性,以募集ADAR或APOBEC。包含两个结构域的核酸可以在任何数目的位置处分裂,例如在两个结构域之间的位置处分裂。在一些情况下,第一结构域或第二结构域可以包括MS2茎环、BoxB茎环、U1A茎环、任何这些的修饰形式或它们的任何组合。

[0089] 与可以形成不同的二维形状或二级结构的核酸相比,结构域的二维形状或二级结构可影响编辑效率、脱靶效应或其组合。因此,本公开的一个方面包括修饰核酸,使得可以有利地设计二维形状,以增强编辑效率并减少脱靶效应。对包括天然存在的募集结构域的序列进行修饰,还可以增强编辑效率并减少脱靶效应。因此,本公开的一个方面包括修饰核酸,使得可以有利地设计序列(例如合成序列),以增强编辑效率并减少脱靶效应。修饰可以包括改变结构域的长度(例如延长长度),改变天然序列导致二级结构的改变,添加化学修饰,或它们的任何组合。如本文所述的核酸可以提供这些优点。

[0090] 在某些情况下，如本文所述的核酸可以以与可比较核酸相比，高至少约3、4或5倍的效率，修饰靶标核酸的至少一个碱基对，其中可比较核酸与天然募集结构域和反义结构域(与靶标核酸互补)复合。在某些情况下，如本文所述的核酸可以以与可比较核酸相比，高至少约3、4或5倍的效率，修饰靶标核酸的至少一个碱基对，其中可比较核酸与GluR2结构域和反义结构域(与靶标核酸互补)复合。在某些情况下，如本文所述的核酸可以以与可比较核酸相比，高至少约3、4或5倍的效率，修饰靶标核酸的至少一个碱基对，其中可比较核酸与Cas13b蛋白或其活性片段和反义结构域(与靶标核酸互补)复合。效率上的改进可以通过测序方法，例如Sanger法来测量。

[0091] 本公开的一个方面提供了一种载体。所述载体可以包括具有多核苷酸序列的核酸，其编码至少一种RNA编辑实体募集结构域。在某些情况下，所述多核苷酸序列可以不形成含茎环的二级结构。在某些情况下，所述多核苷酸序列可以形成一个或多个茎环。在某些情况下，所述多核苷酸序列可以形成包括十字形的二级结构。在某些情况下，所述多核苷酸序列可以形成可以基本上是线性的二级结构。在某些情况下，所述多核苷酸序列可以包括与包括以下项的一个或多个序列的至少约80％的序列同一性：Alu结构域编码序列、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域编码序列及其任意组合。在某些情况下，所述核酸可以是基因编码的。在某些情况下，所述核酸可以是化学合成的。

[0092] 在某些情况下，多核苷酸序列可以包括与Alu结构域编码序列的至少约80％的序列同一性。在某些情况下，多核苷酸序列可以包括与Alu结构域编码序列的至少约85％的序列同一性。在某些情况下，多核苷酸序列可以包括与Alu结构域编码序列的至少约90％的序列同一性。在某些情况下，多核苷酸序列可以包括与Alu结构域编码序列的至少约95％的序列同一性。在某些情况下，所述Alu结构域编码序列可以是非天然存在的序列。在某些情况下，所述Alu结构域编码序列可以包括经修饰的部分。在某些情况下，所述Alu结构域编码序列可以包括天然存在的Alu结构域序列的一部分。

[0093] 在某些情况下，多核苷酸序列可以包括与APOBEC结构域编码序列的至少约80％的序列同一性。在某些情况下，多核苷酸序列可以包括与APOBEC结构域编码序列的至少约85％的序列同一性。在某些情况下，多核苷酸序列可以包括与APOBEC结构域编码序列的至少约90％的序列同一性。在某些情况下，多核苷酸序列可以包括与APOBEC结构域编码序列的至少约95％的序列同一性。在某些情况下，所述APOBEC结构域编码序列可以是非天然存在的序列。在某些情况下，所述APOBEC结构域编码序列可以包括经修饰的部分。在某些情况下，所述APOBEC结构域编码序列可以包括天然存在的APOBEC结构域序列的一部分。

[0094] 在某些情况下，多核苷酸序列可以包括与GluR2结构域编码序列的至少约80％的序列同一性。在某些情况下，多核苷酸序列可以包括与GluR2结构域编码序列的至少约85％的序列同一性。在某些情况下，多核苷酸序列可以包括与GluR2结构域编码序列的至少约90％的序列同一性。在某些情况下，多核苷酸序列可以包括与GluR2结构域编码序列的至少约95％的序列同一性。在某些情况下，所述GluR2结构域编码序列可以是非天然存在的序列。在某些情况下，所述GluR2结构域编码序列可以包括经修饰的部分。在某些情况下，所述GluR2结构域编码序列可以包括天然存在的GluR2结构域序列的一部分。

[0095] 在某些情况下，多核苷酸序列可以包括与募集ADAR的编码序列的至少约80％的序列同一性。可以分离和纯化，或者可以合成，编码至少一种RNA编辑实体募集结构域的多核

苷酸序列。这样的多核苷酸序列可以被特异性地配置为将ADAR募集到靶位点。所述募集可以包括外源性ADAR募集(其可以被共递送或单独递送)、内源性ADAR募集或其组合。在某些情况下,与天然存在的募集结构域相比,多核苷酸序列可以被特异性地配置为增强ADAR的募集或增强ADAR募集至特定位点的特异性。在某些情况下,所述编码序列可以是非天然存在的序列。在某些情况下,所述编码序列可以包括经修饰的部分。在某些情况下,所述编码序列可以包括天然存在的ADAR募集结构域序列的一部分。可以设想募集ADAR的任何序列(天然的或合成的)均包括在所述多核苷酸序列中。在某些情况下,多核苷酸序列可以包括本文所述的示例性序列。图2、图6、图7、图8、图14、图15、图21、图26、图28、图29、图30和表1包括示例性序列。本文提供的序列包含具有可以编码至少一种RNA编辑实体募集结构域的至少一部分的序列。

[0096] 表1:编码adRNA序列的DNA。由这些序列产生的adRNA序列与所述DNA序列相同,但T被取代成U:

[0097]

双_V2_Rab7a_20_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCC GC**C**AGCTGGATTTCCCAATTCTGAGT**GTGGAATAGTATAACAATATGCT AAATGTTGTTATAGTATCCCAC** （SEQ ID NO:178）

[0098]

双_V2_Rab7a_40_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGC**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:179）
双_V2_Rab7a_60_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACC**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:180）
双_V2_Rab7a_80_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCACCTTACAGGCCTGCA**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:181）
双_V2_Rab7a_100_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCACCTTACAGGCCTGCATTACAGGACTTAAACACATA**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:182）
双_V2_Rab7a_120_6	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCACCTTACAGGCCTGCATTACAGGACTTAAACACATAATCCAAGAATTTCTTACACT**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:183）
双_V2_Rab7a_20_10	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**ATACTGCCGCCAGCTGGATT**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:184）
双_V2_Rab7a_40_20	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**ACTGTACAGAATACTGCCGCCAGCTGGATTTCCCAATTCT**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:185）
双_V2_Rab7a_60_30	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TCTTGTGTCTACTGTACAGAATACTGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACT**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**（SEQ ID NO:186）

[0099]

双_V2_Rab7a_80_40	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**CGTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCA**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC** （SEQ ID NO:187）
双_V2_Rab7a_100_50	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGATAAAAGGCGTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTC**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC** （SEQ ID NO:188）
双_V2_Rab7a_120_60	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**CTTAAGTCTTTGATAAAAGGCGTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCAC**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC** （SEQ ID NO:189）
双_V7_Rab7a_20_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TGCCGC**C**AGCTGGATTTCCCAATTCTGAGT**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**（SEQ ID NO:190）
双_V7_Rab7a_40_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGC**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**（SEQ ID NO:191）
双_V7_Rab7a_60_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACC**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**（SEQ ID NO:192）
双_V7_Rab7a_80_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCACCTTACAGGCCTGCA**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**（SEQ ID NO:193）
双_V7_Rab7a_100_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTCAACCCTCCACCTTACAGGCCTGCATTACAGGACTTAAACACATA**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**（SEQ ID NO:194）

[0100]

双_V7_Rab7a_120_6	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**T GCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAG GGTTCAACCCTCCACCTTACAGGCCTGCATTACAGGACTTAAACACATAA TCCAAGAATTTCTTACACT**GTGGAAGAGGAGAACAATAGGCTAAACGT TGTTCTCGTCTCCCAC** （SEQ ID NO:195）
双_V7_Rab7a_220_10	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**AT ACTGCCGC**C**AGCTGGATT**GTGGAAGAGGAGAACAATAGGCTAAACGT TGTTCTCGTCTCCCAC** （SEQ ID NO:196）
双_V7_Rab7a_420_20	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**A CTGTACAGAATACTGCCGC**C**AGCTGGATTTCCCAATTCT**GTGGAAGAGG AGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC** （SEQ ID NO:197）
双_V7_Rab7a_620_30	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**TC TTGTGTCTACTGTACAGAATACTGCCGC**C**AGCTGGATTTCCCAATTCTGA GTAACACT**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTC TCCCAC** （SEQ ID NO:198）
双_V7_Rab7a_820_40	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**C GTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGC**C**AGCTGGATTTC CCAATTCTGAGTAACACTCTGCAATCCA**GTGGAAGAGGAGAACAATAG GCTAAACGTTGTTCTCGTCTCCCAC** （SEQ ID NO:199）
双_V7_Rab7a_100_50	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**T GATAAAAGGCGTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGC**C** AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTC**G TGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC** （SEQ ID NO:200）
双_V7_Rab7a_120_60	**GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC**CT TAAGTCTTTGATAAAAGGCGTACATAATTCTTGTGTCTACTGTACAGAATA CTGCCGC**C**AGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAAC AGGGTTCAACCCTCCAC**GTGGAAGAGGAGAACAATAGGCTAAACGTT GTTCTCGTCTCCCAC** （SEQ ID NO:201）

[0101] RNA编辑实体募集结构域的多核苷酸序列，可以包括募集任何ADAR蛋白（例如ADAR1、ADAR2、ADAR3或其任意组合）、任何APOBEC蛋白（例如APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H、APOBEC4或其任意组合）或其组合。在某些情况下，募集的ADAR或APOBEC蛋白可以是哺乳动物的。在某些情况下，募集的ADAR或APOBEC蛋白可以是人类的。在某些情况下，募集的ADAR或APOBEC蛋白可以是重组的（例如外源递送的ADAR或APOBEC）、经修饰的（例如外源递送的ADAR或APOBEC）、内源性的（例如内源性ADAR或APOBEC）或其任意组合。

[0102] 在某些情况下，所述至少一种RNA编辑实体募集结构域不形成包括茎环的二级结

构。在某些情况下，所述至少一种RNA编辑实体募集结构域形成包括茎环的二级结构。在某些情况下，所述至少一种RNA编辑实体募集结构域形成不含茎环的二级结构。在某些情况下，所述至少一种RNA编辑实体募集结构域形成包括线性部分的二级结构。在某些情况下，所述至少一种RNA编辑实体募集结构域形成包括十字形或其部分的二级结构。

[0103] 多核苷酸序列可以编码一种以上的RNA编辑募集结构域。在某些情况下，多核苷酸序列可以编码多种募集结构域。在某些情况下，多核苷酸序列可以编码2、3、4、5、6或更多种募集结构域。多种募集结构域可以包含Alu结构域、APOBEC结构域、GluR2结构域、Cas13结构域或其任意组合。在某些情况下，所述Alu结构域、APOBEC结构域、Cas13结构域或GluR2结构域可以是天然存在的募集结构域。在某些情况下，所述Alu结构域、APOBEC结构域、Cas13结构域或GluR2结构域可以是非天然存在的，可以是从天然序列进行修饰的，或可以是重组的。多种募集结构域中的至少一种可以包括单链的序列。多种募集结构域中的至少一种可以包括多个Alu重复序列。多种募集结构域中的至少一种可以形成含茎环的二级结构。多种募集结构域中的至少一种可以形成不含茎环的二级结构。多种募集结构域中的至少一种可以形成包括十字形或其部分的二级结构。多种募集结构域中的至少一种可以形成包括立足点形的二级结构。

[0104] 在某些情况下，核酸可以编码至少一种RNA编辑实体募集结构域。在某些情况下，所述核酸可以编码与靶标RNA的至少一部分互补的RNA。在某些情况下，所述核酸可以编码募集结构域和靶向结构域。在某些情况下，所述核酸可以编码募集结构域，且核酸可以编码靶向结构域。所述靶标RNA的一部分可以包括单个碱基。所述靶标RNA的一部分可以包括多个碱基。所述靶标RNA的一部分可以包括约2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、100、200、300、400、500、600、700、800、900、100个碱基对或更多个。在某些情况下，所述靶标RNA可以包括约1bp至约10bp。在某些情况下，所述靶标RNA可以包括约10bp至约100bp。在某些情况下，所述靶标RNA可以包括约10bp至约500bp。在某些情况下，所述靶标RNA可以包括约10bp至约1000bp。包括靶向结构域和募集结构域的核酸可以包括长度上至少约200bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约150bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约250bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约275bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约300bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约400bp的连续序列。包括靶向结构域和募集结构域的核酸可以包括长度上至少约500bp的连续序列。

[0105] 可以采用载体来递送核酸。载体可以包括DNA，例如双链的DNA或单链的DNA。载体可以包括RNA。在某些情况下，所述RNA可以包括碱基修饰。所述载体可以包括重组的载体。所述载体可以是从天然存在的载体修饰的载体。所述载体可以包括非天然存在的载体的至少一部分。可以利用任何载体。在某些情况下，所述载体可以包括病毒载体、脂质体、纳米颗粒、外泌体、细胞外囊泡或其任意组合。在某些情况下，病毒载体可以包括腺病毒载体、腺相关病毒载体(AAV)、慢病毒载体、逆转录病毒载体、这些载体中的任一种的一部分或其任意组合。在某些情况下，纳米颗粒载体可以包括基于聚合物的纳米颗粒、基于氨基脂质的纳米颗粒、金属纳米颗粒(例如金基纳米颗粒)、这些颗粒中的任一种的一部分或其任意组合。在某些情况下，载体可以包括AAV载体。载体可以经修饰，以包含经修饰的VP1蛋白(例如经修

饰以包含VP1蛋白的AAV载体)。AAV可以包括血清型-例如AAV1血清型、AAV2血清型、AAV3血清型、AAV4血清型、AAV5血清型、AAV6血清型、AAV7血清型、AAV8血清型、AAV9血清型、这些中的任一种的衍生物或其任意组合。

[0106] 除非另有定义,否则本文中使用的所有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含义。除非另外指出,否则本文提供的所有核苷酸序列均以5'至3'的方向来呈现。尽管与本文描述的那些类似或等同的任何方法和材料都可以用于本公开的实践或测试中,但是现在描述优选的方法、装置和材料。本文所引用的所有技术和专利出版物均通过引用整体并入本文中。本文中的任何内容均不得解释为承认本公开无权凭借在先公开而早于该公开。

[0107] 除非另有说明,否则该技术的实践将采用组织培养、免疫学、分子生物学、微生物学、细胞生物学和重组DNA的常规技术,这些技术在本领域技术的范围内。参见,例如Green和Sambrook编辑(2012)Molecular Cloning:A Laboratory Manual,4th edition(分子克隆:实验室手册,第4版);系列Ausubel等人编辑(2015)Current Protocols in Molecular Biology(分子生物学现代实验方案);系列Methods in Enzymology(酶学方法)(学术出版社有限公司,N.Y.);MacPherson et al.(2015)PCR 1:A Practical Approach(PCR 1:实践方法)(牛津大学出版社的IRL出版社);MacPherson等人(1995)PCR 2:A Practical Approach(PCR 2:实践方法);McPherson等人(2006)PCR:The Basics(PCR:基础知识)(嘉兰(Garland)科学);Harlow和Lane eds.(1999)Antibodies,A Laboratory Manual(抗体,实验手册);Greenfield ed.(2014)Antibodies,A Laboratory Manual(抗体,实验手册);Freshney(2010)Culture of Animal Cells:A Manual of Basic Technique,6th edition(动物细胞培养:基础技术手册,第6版);Gait ed.(1984)Oligonucleotide Synthesis(寡核苷酸合成);美国专利号4,683,195;Hames和Higgins eds.(1984)Nucleic Acid Hybridization(核酸杂交);Anderson(1999)Nucleic Acid Hybridization(核酸杂交);Herdewijn编辑(2005)Oligonucleotide Synthesis:Methods and Applications(寡核苷酸合成:方法和应用);Hames和Higgins编辑(1984)Transcription and Translation(转录和翻译);Buzdin和Lukyanov编辑(2007)Nucleic Acids Hybridization:Modern Applications(核酸杂交:现代应用);Immobilized Cells and Enzymes(固定化细胞和酶)(IRL出版社(1986));Grandi ed.(2007)In vitro Transcription and Translation Protocols,2nd edition(体外转录和翻译方案,第2版);Guisan ed.(2006)Immobilization of Enzymes and Cells(酶和细胞的固定);Perbal(1988)A Practical Guide to Molecular Cloning,2nd edition(分子克隆的实用指南,第2版);Miller和Calos编辑,(1987)Gene Transfer Vectors for Mammalian Cells(用于哺乳动物细胞的基因转移载体)(冷泉港实验室);Makrides编辑(2003)Gene Transfer and Expression in Mammalian Cells(哺乳动物细胞中的基因转移和表达);Mayer和Walker编辑(1987)Immunochemical Methods in Cell and Molecular Biology(细胞和分子生物学的免疫化学方法)(学术出版社,伦敦);Lundblad和Macdonald编辑(2010)Handbook of Biochemistry and Molecular Biology,4th edition(生物化学和分子生物学手册;第4版);Herzenberg et al.eds(1996)Weir's Handbook of Experimental Immunology,5th edition(威尔的实验免疫学手册,第5版);和/或,它们的最新版本。

[0108] 在本文说明书中使用的术语仅出于描述特定实施方式的目的，而不旨于限制本公开。

[0109] 所有数值指示，例如pH值、温度、时间、浓度和分子量(包括范围在内)都是近似值，其以(+)或(-)递增1.0或0.1，或者适当地或以+/-15％或10％或5％或2％发生变化。应当理解，尽管并不总是明确指出，但是所有数值指示之前均带有术语“约”。还应理解，尽管并非总是明确指出，但本文所述的试剂仅是示例性的，其等同物是本领域已知的。

[0110] 除非上下文另外指出，否则特别旨于本文所述的本公开的各种特征可以以任何组合来使用。此外，本公开还预期在一些实施方式中，可以排除或省略本文阐述的任何特征或特征的组合。为了说明，如果说明书指出复合物包含组分A、B和C，则特别旨于可以省略和单独或以任何组合放弃A、B或C中的任一种或它们的组合。

[0111] 除非另外明确指出，否则所有指定的实施方式、特征和术语旨于包括所列举的实施方式、特征或术语及其生物学等同物。

[0112] 定义

[0113] 如说明书和权利要求书中所使用的，除非上下文另外明确指出，否则单数形式“一个”、“一种”和“该/所述”包括复数引用。例如，术语“多肽”包括多个多肽，包括其混合物。因此，除非有相反指示，否则在本申请中提出的数值参数是近似值，其可以根据本公开寻求获得的期望特性而变化。

[0114] 如本文所用的术语“约”可意指，如本领域普通技术人员所确定的，特定数值在可接受的误差范围内，其可部分取决于该数值是如何测量或确定的，例如，测量系统的局限性。例如，根据本领域的实践，“约”可以意指±10％。或者，“约”可以表示给定数值的±20％、±10％、±5％或±1％的范围。或者，特别是关于生物系统或工艺，该术语可以意指在数值的一数量级内，5倍以内或2倍以内。在本申请和权利要求书中描述了特定数值的情况下，除非另有说明，否则可以假设术语“约”表示该特定数值在可接受的误差范围内。同样，在提供了数值的范围和/或子范围的情况下，该范围和/或子范围可以包括该范围和/或子范围的端点值。在某些情况下，变化可以包括指定量的20％、10％、5％、1％、0.5％或甚至0.1％的量或浓度。

[0115] 为了本文所列举的数值范围，明确地考虑了它们之间具有相同精确度的每个居间数字。例如，对于6～9的范围，除了6和9外，还可以考虑数字7和8；对于6.0～7.0的范围，明确考虑到数字6.0、6.1、6.2、6.3、6.4、6.5、6.6、6.7、6.8、6.9和7.0。

[0116] 如本文所用的术语“腺嘌呤”、“鸟嘌呤”、“胞嘧啶”、“胸腺嘧啶”、“尿嘧啶”和“次黄嘌呤”(肌苷中的核碱基)指的是像这样的核碱基。

[0117] 术语“腺苷”、“鸟苷”、“胞苷”、“胸苷”、“尿苷”和“肌苷”是指与(脱氧)核糖基糖连接的核碱基。

[0118] 如本文所用的术语“腺相关病毒”或“AAV”是指与该名称相关并属于细小病毒(dependoparvovirus)属细小病毒(Parvoviridae)科的病毒类别的成员。已知该病毒的多种血清型适用于基因递送；所有已知的血清型都可以感染来自各种组织类型的细胞。在现有技术中，公开了至少11种(顺序编号)。用于本文公开的目的的非限制性示例性血清型包括11种血清型中的任一种，例如，AAV2和AAV8。如本文所用的术语“慢病毒”是指与该名称相关并且属于慢病毒属、逆转录病毒(Retroviridae)科的病毒类别的成员。虽然已知某些慢

病毒会引起疾病,但已知其他慢病毒也适用于基因递送。参见,例如,Tomás等人(2013)Biochemistry,Genetics and Molecular Biology:“Gene Therapy-Tools and Potential Applications”(生物化学、遗传学和分子生物学:“基因疗法-工具和潜在应用”)ISBN 978-953-51-1014-9,DOI:10.5772/52534。

[0119] 如本文所用的术语“作用于RNA的腺苷脱氨酶”或“ADAR”可以指可以将RNA序列中的腺苷(A)转化为肌苷(I)的腺苷脱氨酶。ADAR1和ADAR2是参与体内mRNA编辑的两种示例性ADAR。ADAR1的非限制性示例性序列可以在以下参考编号下找到:HGNC:225;Entrez基因:103;Ensembl:ENSG 00000160710;OMIM:146920;UniProtKB:P55265;和,GeneCards:GC01M154554,以及它们的生物等效物。ADAR2的非限制性示例性序列可在以下参考编号下找到:HGNC:226;Entrez基因:104;Ensembl:ENSG00000197381;OMIM:601218;UniProtKB:P78563;和,GeneCards:GC21P045073,以及它们的生物等效物。上面提供了催化结构域的其他非限制性示例性序列。用于指导位点特异性ADAR编辑的正向和反向RNA分别称为“adRNA”和“radRNA”。ADAR1和ADAR2的催化结构域包含在下文提供的序列中。

[0120] ADAR1催化结构域:

[0121] KAERMGFTEVTPVTGASLRRTMLLLSRSPEAQPKTLPLTGSTFHDQIAMLSHRCFNTLTNSFQPSLLGRKILAAIIMKKDSEDMGVVVSLGTGNRCVKGDSLSLKGETVNDCHAEIISRRGFIRFLYSELMKYNSQTAKDSIFEPAKGGEKLQIKKTVSFHLYISTAPCGDGALFDKSCSDRAMESTESRHYPVFENPKQGKLRTKVENGEGTIPVESSDIVPTWDGIRLGERLRTMSCSDKILRWNVLGLQGALLTHFLQPIYLKSVTLGYLFSQGHLTRAICCRVTRDGSAFEDGLRHPFIVNHPKVGRVSIYDSKRQSGKTKETSVNWCLADGYDLEILDGTRGTVDGPRNELSRVSKKNIFLLFKKLCSFRYRRDLLRLSYGEAKKAARDYETAKNYFKKGLKDMGYGNWISKPQEEKNFYLCPV(SEQ ID NO:140)

[0122] ADAR2催化结构域:

[0123] QLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLT(SEQ ID NO:141)

[0124] ADAR的双链RNA结合结构域(dsRBD)包含在下文提供的序列中。

[0125] ADAR dsRBD:

[0126] MDIEDEENMSSSSTDVKENRNLDNVSPKDGSTPGPGEGSQLSNGGGGGPGRKRPLEEGSNGHSKYRLKKRRKTPGPVLPKNALMQLNEIKPGLQYTLLSQTGPVHAPLFVMSVEVNGQVFEGSGPTKKKAKLHAAEKALRSFVQFPNASEAHLAMGRTLSVNTDFTSDQADFPDTLFNGFETPDKAEPPFYVGSNGDDSFSSSGDLSLSASPVPASLAQPPLPVLPPFPPPSGKNPVMILNELRPGLKYDFLSESGESHAKSFVMSVVVDGQFFEGSGRNKKLAKARAAQSALAAIFN(SEQ ID NO:142)

[0127] 应当理解,可以对ADAR和/或其各种结构域的序列进行进一步的突变。例如,本公开提供了ADAR1和ADAR2二者的E488Q和E1008Q突变体,以及ADAR2的“混杂”变体-由C末端缺失导致。这种“混杂”变体被称为“变体”,是因为它在编辑读物中表现出了混杂性,其中几个A接近显示A至G转化的靶序列(在2个不同的基因座上得到了验证)。该变体的序列在下文提供。

[0128] “混杂”ADAR2变体：

[0129] MLRSFVQFPNASEAHLAMGRTLSVNTDFTSDQADFPDTLFNGFETPDKAEPPFYVGSNGDDSFSSSGDLSLSASPVPASLAQPPLPVLPPFPPPSGKNPVMILNELRPGLKYDFLSESGESHAKSFVMSVVVDGQFFEGSGRNKKLAKARAAQSALAAIFNLHLDQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLTP*(SEQ ID NO:143)

[0130] 不受理论约束的情况下，ADAR1中的C-端缺失可以产生相同或相似的作用。

[0131] 术语“Alu结构域”可以指获自Alu转座元件(“Alu元件”)的序列。通常，Alu元件的长度为约300个碱基对。Alu元件通常包括以下结构：十字形-多聚(poly)A5-TAC-polyA6-十字形-polyA尾部，其中两个十字形结构域在核苷酸序列上是相似的。“Alu结构域”可以包括Alu元件的十字形部分。在一些实施方式中，包括十字形结构的两个Alu结构域通过与靶标RNA序列互补的序列连接。

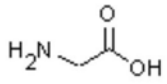
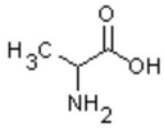
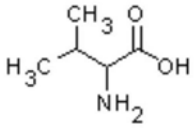
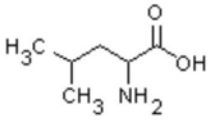
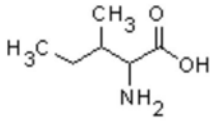
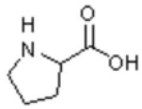
[0132] 如本文所用的术语“APOBEC”可以指，落入参与mRNA编辑的进化上保守的胞苷脱氨酶家族(其催化C至T的编辑，这可以解释为C至U的转化)中的任何蛋白及其等同物。在某些方面，术语APOBEC可以指APOBEC1、APOBEC2、APOBEC3A、APOBEC3B、APOBEC3C、APOBEC3E、APOBEC3F、APOBEC3G、APOBEC3H、APOBEC4或其等同物中的任何一种。本文提供了包括一种或多种APOBEC结构域的融合蛋白的非限制性示例性序列，其融合至ADAR结构域或融合至替代结构域，以使它们适合用于RNA编辑系统。为此，APOBEC可以视为ADAR的等同物，但其可以通过不同的转换来催化编辑。因此，不受理论约束的情况下，本文考虑的用于基于ADAR的编辑系统的所有实施方式，据信可以适于在基于APOBEC的RNA编辑系统中使用。在某些情况下，APOBEC的使用可涉及某些修饰，例如，但不限于，使用特定的向导RNA或“gRNA”来募集酶。

[0133] “适体”可以指，能够以高亲和力和特异性结合各种分子的短的单链寡核苷酸。适体的非限制性实施例描述于Lakhin,A.V.et al.(2013).Acta naturae,5(4),34-43中。

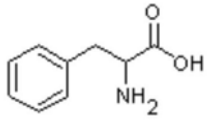
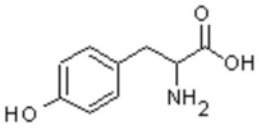
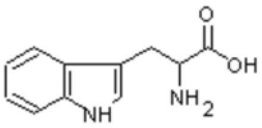
[0134] 如本文所用的术语“包括”旨在表示，组合物和方法包括所列举的要素，但不排除其他要素。除非另外指出，否则开放式术语，例如“包含”、“含有”、“含”、“具有”等表示包括。当使用“基本上由...组成”来定义组合物和方法时，应表示排除对于预期用途而言对该组合具有任何重要意义的其他元素。因此，基本上由本文所定义的元素组成的组合物可能不排除来自分离和纯化方法的痕量污染物，以及药学上可接受的载剂(例如磷酸盐缓冲盐水、防腐剂等)。“由...组成”是指排除其他成分的痕量元素，和用于施用本公开的组合物的实质性方法步骤。由这些过渡术语中的每一个限定的实施方式在本公开的范围内。

[0135] “经典氨基酸”是指在下表中示出的人体中天然存在的20种氨基酸，各自都具有三个字母缩写、一个字母缩写、结构和相应的密码子：

[0136] 非极性脂族残基

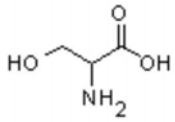
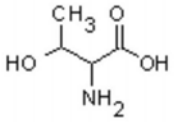
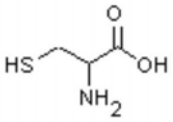
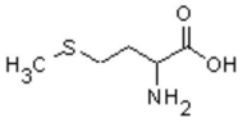
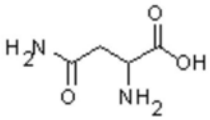
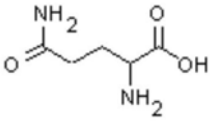
	甘氨酸	Gly	G		GGU GGC GGA GGG
[0137]	丙氨酸	Ala	A		GCU GCC GCA GCG
	缬氨酸	Val	V		GUU GUC GUA GUG
	亮氨酸	Leu	L		UUA UUG CUU CUC CUA CUG
[0138]	异亮氨酸	Ile	I		AUU AUC AUA
	脯氨酸	Pro	P		CCU CCC CCA CCG

[0139] 芳香族残基

	苯丙氨酸	Phe	F		UUU UUC
	酪氨酸	Tyr	Y		UAU UAC
	色氨酸	Trp	W		UGG

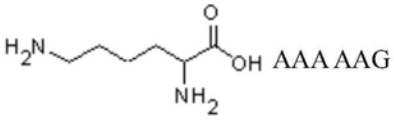
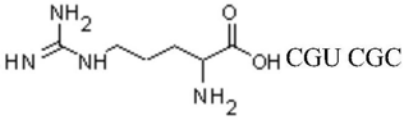
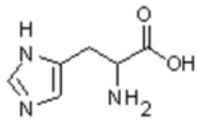
[0140] 极性不带电荷的残基

[0141]

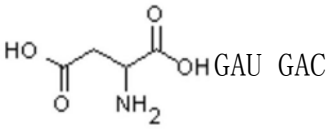
丝氨酸	Ser	S		UCU UCC UCA UCG AGU AGC
苏氨酸	Thr	T		ACU ACC ACA ACG
半胱氨酸	Cys	C		UGU UGC
甲硫氨酸	Met	M		AUG
天冬酰胺	Asn	N		AAU AAC
谷氨酰胺	Gln	Q		CAA CAG

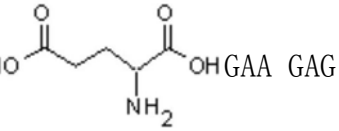
[0142] 带正电荷的残基

[0143]

赖氨酸	Lys	K		AAA AAG
精氨酸	Arg	R		CGU CGC CGA CGG AGA AGG
组氨酸	His	H		CAU CAC

[0144] 带负电荷的残基

[0145] 天冬氨酸Asp D  GAU GAC

[0146] 谷氨酸Glu E  GAA GAG

[0147] 术语“Cas9”可以指该名称所指的CRISPR相关的核酸内切酶。非限制性的示例性Cas9包括金黄色葡萄球菌(Staphylococcus aureus)Cas9、核酸内切酶失活的(nuclease dead)Cas9,以及它们各自的直系同源物和生物学等同物。直系同源物包括,但不限于,化脓链球菌(Streptococcus pyogenes)Cas9(“spCas9”),来自嗜热链球菌(Streptococcus thermophiles)、嗜肺军团菌(Legionella pneumophilia)、乳酸奈瑟球菌(Neisseria lactamica)、脑膜炎奈瑟菌(Neisseria meningitides)、弗朗西丝菌(Francisella novicida)的Cas 9;和,来自各种细菌种类(包括氨基酸球菌属(Acidaminococcus spp.)和新凶手弗朗西丝氏菌(Francisella novicida)U112)的Cpf1(其进行类似于Cas9的切割功能)。例如,可以使用UniProtKB G3ECR1(CAS9_STRTR)以及缺少核酸内切酶活性(例如,在RuvC和HNH结构域中均具有突变)的死亡Cas9或dCas9。术语“Cas9”可进一步指,与其具有至少约60%、65%、70%、75%、80%、85%、90%、95%或99%同一性的参考Cas9的等同物,包括但不限于其他大型Cas9蛋白。在一些实施方式中,Cas9衍生自空肠弯曲杆菌(Campylobacter jejuni),或长度为1000个氨基酸或更少的另一种Cas9直系同源。

[0148] 如本文所用的术语“CRISPR”可以指,依赖于簇状规则间隔的短回文重复通路的序列特异性遗传操作技术。CRISPR可用于进行基因编辑和/或基因调控,以及简单地将蛋白靶向特定的基因组位置。“基因编辑”可以指遗传改造的一种类型,其中通过在多核苷酸序列中引入缺失、插入、单链或双链断裂或碱基取代,来改变靶多核苷酸的核苷酸序列。在某些方面,CRISPR介导的基因编辑利用非同源末端连接(NHEJ)或同源重组的通路来进行编辑。基因调控可以指,增加或减少特定基因产物(例如蛋白或RNA)的产生。

[0149] 如本文所用的术语“缺乏”可以指,低于特定试剂的正常(生理上可接受的)水平。就蛋白而言,缺乏可以指低于全长蛋白的正常水平。

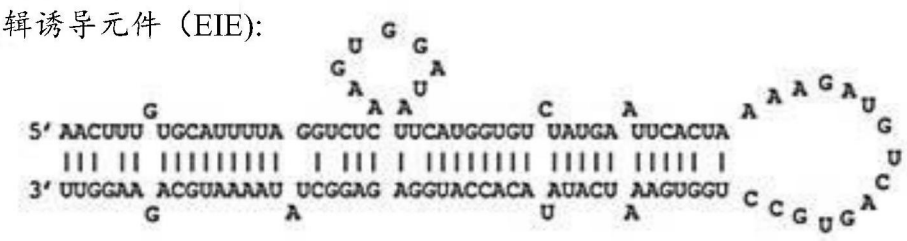
[0150] 如本文所用的术语“可检测的标记物”可以指,至少一种能够直接或间接产生可检测信号的标记物。该标记物的非详尽列表包括(例如通过比色法、荧光、发光)产生可检测信号的酶,例如辣根过氧化物酶、碱性磷酸酶、β-半乳糖苷酶、葡萄糖-6-磷酸脱氢酶;发色团,例如荧光、发光染料;具有通过电子显微镜或通过其电学性质检测到的电子密度的基团,例如电导率、安培法、伏安法、阻抗可检测到的基团,例如其分子的大小足以诱导其物理和/或化学性质的可检测到的修饰,这样的检测可以通过光学方法(例如衍射、表面等离振子共振、表面变化、接触角变化)或物理方法(例如原子力光谱、隧道效应或放射性分子,例如^{32}P、^{35}S或^{125}I)来完成。

[0151] 如本文所用的术语“结构域”可以指蛋白或多肽的特定区域,并且与特定功能相关。例如,“与RNA发夹基序相关的结构域”可以指结合一个或多个RNA发夹的蛋白的结构域。该结合可选地对于特定发夹具有特异性。

[0152] 如本文所用的术语“抗肌萎缩蛋白”是指,与该名称相对应并由基因Dmd编码的蛋白;其非限制性示例在UniProt参考号P11532(对于人类)和P11531(对于小鼠)下找到。

[0153] “编辑诱导元件”可以指主要是双链RNA的结构,这对于有效的RNA编辑是必需的。编辑诱导元件的非限制性实施例描述在Daniel,C.et al.(2017)Genome Biol.18,195中。以下结构(SEQ ID NO:15)提供了编辑诱导元件的另一非限制性实施例:

编辑诱导元件（EIE):

[0154] 

[0155] 进行本公开的模式

[0156] ADAR是天然存在的RNA编辑酶，其催化腺苷水解脱氨为肌苷，肌苷在生物化学上被认为是鸟苷。APOBEC是上文所述的酶，其可以进行类似的功能，但是用于胞嘧啶脱氨至胸腺嘧啶。

[0157] 如应用于多核苷酸的术语“编码”可以指，被称为“编码”多肽的多核苷酸，如果在其天然状态或通过本领域技术人员熟知的方法进行操作时，它可以被转录和/或翻译以产生用于该多肽和/或其片段的mRNA。反义链是这种核酸的互补物，并且可以从其推导编码序列。

[0158] 当涉及特定分子、生物或细胞材料时，术语“等同物”或“生物等同物”可互换使用，并且意指具有极小同源性但仍保持所需结构或功能的那些。

[0159] “真核细胞”包括除原核生物外的所有生命王国。它们可以通过膜围绕核而很容易地识别。动物、植物、真菌和原生生物是真核生物或生物，它们的细胞通过内膜和细胞骨架组织成复杂的结构。最具特征性的膜围绕结构是细胞核。除非特别说明，否则术语“宿主”包括真核宿主，包括例如酵母、高等植物、昆虫和哺乳动物细胞。真核细胞或宿主的非限制性实施例包括猿猴、牛、猪、小鼠、大鼠，禽类、爬行动物和人。

[0160] 如本文所用，“表达”可以指多核苷酸转录成mRNA的过程，和/或转录的mRNA随后被翻译成肽、多肽或蛋白的过程。如果多核苷酸来源于基因组DNA，则表达可包括mRNA在真核细胞中的剪接。

[0161] 如本文所用的术语“功能性”可用于修饰任何分子、生物或细胞材料，以使其实现特定的特定作用。

[0162] 如本文所用的术语“Glur2 mRNA”可以指编码离子型AMPA谷氨酸受体2(“Glur2”)的mRNA，其经历腺苷至肌苷(A->I)的编辑。该mRNA以位点特异性方式募集ADAR。

[0163] 如本文所用的术语“gRNA”或“向导RNA”可以指用于靶向特定多核苷酸序列的向导RNA序列，以利用CRISPR技术进行基因编辑。设计用于靶标特异性的gRNA和供体治疗性多核苷酸的技术是本领域众所周知的。例如，Doench,J.,et al.Nature biotechnology 2014;32(12):1262-7,Mohr,S.et al.(2016)FEBS Journal 283:3232-38,和Graham,D.,et al.Genome Biol.2015;16:260。gRNA包括以下项，或者基本上由以下项组成，或者进一步由以下项组成：包括CRISPR RNA(crRNA)和反式激活CRIPSPR RNA(tracrRNA)的融合多核苷酸；或，包括CRISPR RNA(crRNA)和反式激活CRIPSPR RNA(tracrRNA)的多核苷酸。在某些方面，gRNA是合成的(elley,M.et al.(2016)J of Biotechnology 233(2016)74-83)。

[0164] 在寡核苷酸的背景中，单独使用或与“基序”组合使用的术语“发夹”、“发夹环”、“茎环”和/或“环”指在单链寡核苷酸中形成的结构，当序列在单链内时，其在以相反的方向读取碱基对时是互补的，以形成其构象类似发夹或环的区域。

[0165] “同源性”或“同一性”或“相似性”可以指两个肽之间或两个核酸分子之间的序列相似性。同源性可以通过比较每个序列中的位置来确定，该序列可以为了比较的目的而进行比对。当比较的序列中的一个位置被相同的碱基或氨基酸占据时，则分子在该位置是同源的。序列之间的同源性程度是序列所享有的匹配或同源位置数目的函数。“无关的”或“非同源的”序列与本公开的序列之一享有小于40％的同一性或小于25％的同一性。

[0166] 同源性是指序列与参考序列的％同一性。实际上，任何特定序列是否可以与本文所述的任何序列(其可以对应于本文所述的特定核酸序列)具有至少50％、60％、70％、80％、85％、90％、92％、95％、96％、97％、98％或99％的同一性，这种特定的多肽序列可以常规地使用已知的计算机程序，例如Bestfit程序(威斯康星序列分析包(Wisconsin Sequence Analysis Package)，用于Unix的版本8，遗传学计算机组(Genetics Computer Group)，威斯康星州麦迪逊市科学大道575号，大学研究园，邮政编码53711)。当使用Bestfit或任何其他序列比对程序来确定某特定序列是否与参考序列例如具有95％的同一性时，可以设置参数，以便在整个参考序列的全长上计算同一性百分比，并且允许同源性差距不超过总参考序列的5％。

[0167] 例如，在特定的实施方式中，可以使用FASTDB计算机程序，基于Brutlag等人的算法(Comp.App.Biosci.6:237-245(1990))，来确定参考序列(查询序列，即本公开的序列)和主题序列(也称为全局序列比对)之间的同一性。在某些情况下，其中同一性被狭义解释的特定实施方式中，用于FASTDB氨基酸比对的参数可以包括：评分方案＝PAM(接受的突变百分比)0，k-元组＝2，错配罚分＝1，邻接罚分＝20，随机分组长度＝0，截止分数＝1，窗口大小＝序列长度，空位罚分＝5，空位大小罚分＝0.05，窗口大小＝500或主题序列的长度，以较短者为准。根据该实施方式，如果主题序列由于N-端或C-端缺失(不是由于内部缺失)比查询序列短，则可以对结果进行手动校正，以顾及FASTDB程序在计算全局同一性百分比时，没有考虑到主题序列的N-端和C-端截短这一事实。对于相对于查询序列，在N-端和C-端截短的主题序列，可以通过计算查询序列的在主题序列N-端和C-端两侧的残基数来校正同一性百分比，这些残基不与相应的主题残基匹配/比对，以查询序列总碱基的百分比来表示。确定残基是否匹配/比对，可以通过FASTDB序列比对的结果来确定。然后可以从通过FASTDB程序，使用特定参数计算的同一性百分比中减去该百分比，以得出最终的同一性百分比分数。该最终的同一性百分比分数可以用于该实施方式的目的。在某些情况下，出于手动调整同一性百分比分数的目的，仅考虑主题序列的N-端和C-端与查询序列不匹配/未比对的残基。也就是说，对于该手动校正，仅考虑主题序列的最远N-端和C-端残基外侧的查询残基位置。例如，可以将90个残基的主题序列与100个残基的查询序列进行比对，以确定同一性百分比。缺失发生在主题序列的N-端，因此，FASTDB比对不显在N-端的前10个残基的匹配/比对。10个不配对的残基代表序列的10％(在N-端和C-端不匹配的残基数/查询序列中的残基总数)，因此从FASTDB程序计算的同一性百分比分数中减去10％。如果剩余的90个残基完全匹配，则最终同一性百分比可以是90％。在另一个实施例中，将90个残基的主题序列与100个残基的查询序列进行比较。这次缺失是内部缺失，因此在主题序列的N-端或C-端不存在与查询不匹配/比对的残基。在这种情况下，无法手动校正FASTDB计算出的同一性百分比。再次，仅手动校正在主题序列的N-端和C-端外侧的、与查询序列不匹配/未比对的残基位置(如在FASTDB比对中显示的)。

[0168] “杂交”可以指一种反应，其中一种或多种多核苷酸发生反应形成复合物，该复合物通过核苷酸残基的碱基之间的氢键而稳定。氢键可以通过沃森-克里克碱基配对、霍格斯坦结合或以任何其他的序列特异性方式发生。该复合物可以包括形成双螺旋结构的双链，形成多链复合物的三条或更多条链，自杂交的单链，或它们的任意组合。杂交反应可以构成更广泛过程中的一个步骤，例如PC反应的启动，或通过核酶对多核苷酸的酶促切割。

[0169] 严格杂交条件的实施例包括：约25℃至约37℃的孵育温度；约6x SSC至约10x SSC的杂交缓冲液浓度；约0％至约25％的甲酰胺浓度；和，约4x SSC至约8x SSC的洗涤溶液。温和杂交条件的实施例包括：约40℃至约50℃的孵育温度；约9x SSC至约2x SSC的缓冲液浓度；约30％至约50％的甲酰胺浓度；和，约5x SSC至约2x SSC的洗涤溶液。高严格条件的实施例包括：约55℃至约68℃的孵育温度；约1x SSC至约0.1x SSC的缓冲液浓度；约55％至约75％的甲酰胺浓度；和，约1x SSC、0.1x SSC或去离子水的洗涤溶液。通常，杂交孵育时间为5分钟至24小时，有1、2或更多的洗涤步骤，并且洗涤孵育时间为约1、2或15分钟。SSC为0.15M NaCl和15mM柠檬酸盐缓冲液。应当理解，可以采用使用其他缓冲系统的SSC的等同物。

[0170] 如本文所用的“抑制”是指一种能力，其基本上拮抗、阻止、预防、控制、减慢、破坏、改变、消除、停止或逆转特定因子(例如，传染原)或疾病的活性的发展或严重程度。

[0171] 如本文所用的术语“干扰素”可以指已知与免疫应答相关的一组信号传导蛋白。在本申请的背景中，感兴趣的干扰素是那些导致ADAR表达增强的干扰素。干扰素α与ADAR1之间的相关性是众所周知的，因此，本公开考虑使用干扰素α作为增加内源性ADAR1表达的手段。分离或重组的干扰素α的商业来源包括，但不限于，Sigma-Aldrich(西格玛奥德里奇)、R&D Systems(R&D系统)、Abcam(艾博抗)和Thermo Fisher Scientific(赛默飞世尔科技公司)。或者，可以使用已知的载体和给定的蛋白序列，例如Q6QNB6(人IFNA)，来产生干扰素α。

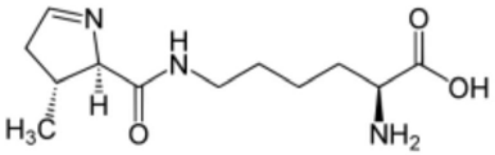
[0172] 如本文所用，术语“分离(的)”可以指基本上不含其他材料的分子或生物或细胞材料。在一方面，术语“分离(的)”可以指分别从存在于天然来源中的其他DNA或RNA，或蛋白或多肽，或细胞或细胞器，或组织或器官中分离的核酸(例如DNA或RNA)，或蛋白或多肽(例如抗体或其衍生物)，或细胞或细胞器，或组织或器官。术语“分离(的)”还可以指，在通过重组DNA技术生产时基本上不含细胞材料、病毒材料或培养基的核酸或肽，或在化学合成时基本不含化学物前体或其他化学物的核酸或肽。此外，“分离的核酸”是指，包括并非天然存在为片段并且可能不会以天然状态发现的核酸片段。术语“分离(的)”在本文中还用于指与其他细胞蛋白中分离的多肽，并且意在涵盖经纯化和重组的多肽。术语“分离(的)”在本文中还用于指与其他细胞或组织分离的细胞或组织，并且意在涵盖经培养和改造的细胞或组织。

[0173] “信使RNA”或“mRNA”是一种核酸分子，其由DNA转录，然后经加工以去除被称为内含子的非编码部分。得到的mRNA从细胞核(或另一存在DNA的位置)输出，并翻译成蛋白。术语“pre-mRNA”可以指在加工去除非编码部分之前的链。

[0174] 如本文所用的术语“突变”可以指，相对于蛋白的共有序列，对编码所述蛋白的核酸序列的改变。“错义”突变导致一个密码子被取代成另一个密码子；“无义”突变将一个密码子从编码特定氨基酸的密码子变为终止密码子。无义突变通常会导致蛋白的翻译被截断。“沉默”突变是对所得的蛋白没有影响的那些突变。如本文所用的术语“点突变”可以指仅影响基因序列中一个核苷酸的突变。“剪接位点突变”是指在pre-mRNA中存在的那些突变

(在加工以去除内含子之前),其导致由于剪接位点的错误划定而使蛋白翻译错误,常常使蛋白截短。突变可以包括单核苷酸变异(SNV)。突变可以包括序列变体、序列变化、序列改变或等位基因变体。可以从参考数据库获得参考DNA序列。突变会影响功能。突变可能不会影响功能。突变可以在DNA水平的一个或多个核苷酸上发生,在核糖核酸(RNA)水平的一个或多个核苷酸上发生,在蛋白水平的一个或多个氨基酸上发生,或以它们的任何组合发生。可以从数据库(例如NCBI参考序列数据库(RefSeq)数据库)中获得参考序列。可以构成突变的具体变化可以包括一个或多个核苷酸,或一个或多个氨基酸的取代、缺失、插入、倒位或转化。突变可以是点突变。突变可以是融合基因。融合对或融合基因可由突变,例如易位、中间缺失、染色体倒位或其任意组合来产生。突变可构成重复序列(例如三倍、四倍或其他)的数量的可变性。例如,突变可以是与给定序列相关的拷贝数的增加或减少(即拷贝数变异或CNV)。突变可以包括在不同等位基因中的两个或更多个序列变化,或在一个等位基因中的两个或更多个序列变化。突变可在一个等位基因的一个位置上包括两个不同的核苷酸,例如嵌合型突变(mosaic)。突变可以在一个等位基因的一个位置上包括两个不同的核苷酸,例如镶嵌型突变(chimeric)。突变可存在于恶性组织中。突变的存在或不存在可以指示发展疾病或病症的风险增加。突变的存在或不存在可以指示疾病或病症的存在。突变可以存在于良性组织中。不存在突变可以指示组织或样品是良性的。或者,不存在突变可能并不表示组织或样品是良性的。本文所述的方法可包括鉴定样品中突变的存在。

[0175] 术语“非经典氨基酸”可以指那些合成的或属于该组之外的其他经修饰的氨基酸,通常是通过化学合成或对经典氨基酸进行修饰(例如氨基酸类似物)而产生的。本公开在本文公开的一些方法和载体中采用了蛋白原性的非经典氨基酸。非限制性的示例性非经典氨基酸是吡咯赖氨酸(Pyl或O),其化学结构提供如下:

[0176] 

[0177] 肌苷(I)是另一示例性的非经典氨基酸,其可以在tRNA中找到,并且对于根据“摇摆碱基配对”的正确翻译是必不可少的。上面提供了肌苷的结构。

[0178] 如本文所用的术语“鸟氨酸转氨甲酰酶”或“OTC”可以指与该名称相对应并由基因Otc编码的蛋白;它的非限制性实施例在UniProt参考号P00480(对于人类)和P11725(对于小鼠)下找到。OTC缺乏症是X连锁遗传病,导致血液中较高的氨浓度。在某些情况下,OTC缺乏症是由外显子4供体剪接位点的G->A剪接位点突变引起的,其导致pre-mRNA的错误剪接。这种突变导致形成延长或携带点突变的蛋白。OTC蛋白水平降低了15～20倍。参见,例如,Hodges,P.E.&Rosenberg,L.E.The spf^{ash} mouse:a missense mutation in the ornithine transcarbamylase gene also causes aberrant mRNA splicing.Proc.Natl.Acad.Sci.U.S.A.86,4142-4146(1989))(显示了所产生的OTC的可选形式)。其序列提供如下:

[0179] OTC pre-mRNA(野生型):

[0180] .....CTCACAGACACCGCTCGGTTTGTAAAACTTTTCTTC.....(SEQ ID NO:144)

[0181] OTC pre-mRNA(突变体):

[0182] .....CTCACAGACACCGCTCAGTTTGTAAAACTTTTCTTC.....(SEQ ID NO:145)

[0183] OTC mRNA(错误剪接的,突变体):

[0184] .....CTCACAGACACCGCTCAGTTTGTAAAACTTTTCTTC.....(SEQ ID NO:146)

[0185] OTC mRNA(正确剪接的,突变体):

[0186] .....CTCACAGACACCGCTCATGTCTTATCTAGCATGACA.....(SEQ ID NO:147)

[0187] OTC mRNA(正确剪接的,野生型):

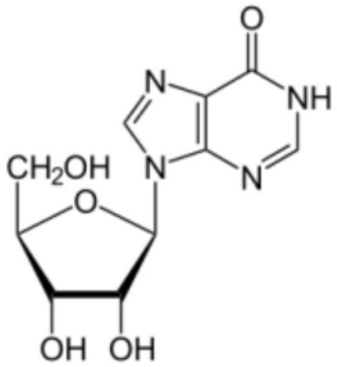
[0188] .....CTCACAGACACCGCTCGTGTCTTATCTAGCATGACA.....(SEQ ID NO:148)

[0189] 如上所示,当存在突变时,可以产生正确的剪接变体;但是,这样的产生导致错义突变,这也可能导致OTC缺乏症。

[0190] 术语“蛋白(质)”、“肽”和“多肽”可互换使用,并且在其最广义上是指两个或更多个亚单位氨基酸、氨基酸类似物或拟肽的化合物。这些亚单位可以通过肽键连接。在另一个实施方式中,亚单位可以通过其他键,例如酯、醚等连接。蛋白或肽可以包含至少两个氨基酸,并且对可以构成蛋白或肽序列的最大氨基酸数没有限制。如本文所用的术语“氨基酸”可以指天然和/或非天然或合成的氨基酸,包括甘氨酸以及D和L光学异构体,氨基酸类似物和模拟肽。如本文所用的术语“融合蛋白”可以指,由来自一种以上的天然存在或重组产生的蛋白的结构域构成的蛋白,其中通常每个结构域发挥不同的功能。在这方面,术语“接头”可以指用于将这些结构域连接在一起的蛋白片段-可选地保留融合的蛋白结构域的构象,和/或防止融合的蛋白结构域之间的不利的相互作用,这些不利的相互作用可能损害其各自的功能。

[0191] 术语“多核苷酸”和“寡核苷酸”可互换使用,是指任何长度的核苷酸的聚合形式,即脱氧核糖核苷酸或核糖核苷酸或其类似物。多核苷酸可以具有任何三维结构,并且可以执行任何已知或未知的功能。以下是多核苷酸的非限制性实施例:基因或基因片段(例如探针、引物、EST或SAGE标签),外显子,内含子,信使RNA(mRNA),转运RNA,核糖体RNA,RNAi,核酶,cDNA,重组多核苷酸,分支多核苷酸,质粒,载体,任何序列的分离DNA,任何序列的分离RNA,核酸探针和引物。多核苷酸可以包括经修饰的核苷酸,例如甲基化的核苷酸和核苷酸类似物。如果存在,则可以在多核苷酸组装之前或之后对核苷酸结构进行修饰。核苷酸的序列可以被非核苷酸组分中断。多核苷酸可在聚合后作进一步修饰,例如通过与标记组分缀合。该术语也可以指双链和单链的分子。除非另有说明或要求,否则本公开的多核苷酸的任何实施方式均涵盖双链形式和两条互补的单链形式中的每一条(已知或预期组成双链形式)。

[0192] 多核苷酸是由以下四种核苷酸碱基的特定序列构成:腺嘌呤(A);胞嘧啶(C);鸟嘌呤(G);胸腺嘧啶(T);和,当多核苷酸为RNA时,胸腺嘧啶为尿嘧啶(U)。在一些实施方式中,多核苷酸可以包括一种或多种其他的核苷酸碱基,例如肌苷(I),一种在次黄嘌呤经由β-N9-糖苷键与核呋喃糖连接时形成的核苷,得到以下的化学结构:

[0193] 

[0194] 肌苷被翻译机器读取为鸟嘌呤(G)。

[0195] 术语“多核苷酸序列”是多核苷酸分子按字母顺序进行的表示。该按字母顺序进行的表示可以输入到具有中央处理单元的计算机的数据库中,并用于生物信息学应用,例如功能基因组学和同源性搜索。

[0196] 如本文所用的术语“纯化标记物”可以指至少一种可用于纯化或鉴定的标记物。这种标记物的非详尽列表包括His、lacZ、GST、麦芽糖结合蛋白、NusA、BCCP、c-myc、CaM、FLAG、GFP、YFP、樱桃(cherry)、硫氧还蛋白、聚(NANP)、V5、Snap、HA、几丁质结合蛋白、软标签(Softag)1,Softag 3、链霉素(Strep)或S-蛋白。合适的直接或间接荧光标记物包括FLAG、GFP、YFP、RFP、dTomato、樱桃(cherry)、Cy3、Cy 5、Cy 5.5、Cy 7、DNP、AMCA、生物素、洋地黄毒苷、塔姆拉(Tamra)、德克萨斯红、罗丹明、Alexa fluors、FITC、TRITC或任何其他荧光染料或半抗原。

[0197] 如本文所用的术语“重组表达系统”是指,用于表达通过重组形成的某些基因材料的一种或多种基因构建体;就此而言,术语“构建体”可与本文所定义的术语“载体”互换。

[0198] 如本文所用的术语“重组蛋白”可以指,通过重组DNA技术产生的多肽,其中通常,将编码多肽的DNA插入到合适的表达载体中,该表达载体继而用于转化宿主细胞以产生异源蛋白。

[0199] 如本文所用的,关于蛋白表达的术语“恢复”可以指在先前蛋白表达由于突变而被截断的情况下,建立全长蛋白表达的能力。在“恢复活性”的背景中,该术语包括在突变导致异常表达(例如,太低或太高)的情况下,使蛋白的表达达到其正常的、未突变的水平。

[0200] 如本文所用的术语“样品”通常是指受试者的任何样品(例如血液样品或组织样品)。样品或其部分可以包括干细胞。样品的一部分可以富集干细胞。可以从样品中分离出干细胞。样品可以包括组织、细胞、血清、血浆、外泌体、体液或其任意组合。体液可以包括尿液、血液、血清、血浆、唾液、粘液、脊髓液、泪液、精液、胆汁、羊水或其任意组合。样品或其部分可以包括获自受试者的细胞外液。样品或其部分可以包括无细胞的核酸、DNA或RNA。可以分析样品或其部分是否存在或不存在一个或多个突变。基因组数据可从样品或其部分获得。样品可以是怀疑或证实具有疾病或病症的样品。样品可以是经由非侵入性技术、微创技术或侵入性技术,从受试者移出的样品。样品或其部分可以通过组织刷、拭子、组织活检、切除的组织、细针抽吸、组织洗涤、细胞学标本、手术切除或其任意组合来获得。样品或其部分可以包括来自一组织类型的组织或细胞。例如,样品可以包括鼻组织、气管组织、肺组织、咽组织、喉组织、支气管组织、胸膜组织、肺泡组织、乳腺组织、膀胱组织、肾脏组织、肝脏组织、结肠组织、甲状腺组织、宫颈组织、前列腺组织、心脏组织、肌肉组织、胰腺组织、肛门组织、

胆管组织、骨组织、脑组织、脊髓组织、肾脏组织、子宫组织、卵巢组织、子宫内膜组织、阴道组织、外阴组织、子宫组织、胃组织、眼组织、窦组织、阴茎组织、唾液腺组织、肠道组织、胆囊组织、胃肠组织、膀胱组织、脑组织、脊髓组织、血液样本或其任意组合。

[0201] 本文使用的术语“测序”可以包括无亚硫酸氢盐测序、亚硫酸氢盐测序、TET辅助亚硫酸氢盐(TAB)测序、ACE测序、高通量测序、马克萨姆·吉利别尔特(Maxam-Gilbert)测序、大规模平行签名测序、波罗尼(Polony)测序、454焦磷酸测序、Sanger测序,依诺米那(Illumina)测序,SOLiD测序,离子激流半导体测序,DNA纳米球测序,赫利克普(Heliscope)单分子测序、单分子实时(SMRT)测序、纳米孔测序、鸟枪法(shot gun)测序、RNA测序、恩尼格玛(Enigma)测序或其任意组合。

[0202] 术语“终止密码子”意指,在信使RNA内的、表明翻译终止的三个核苷酸连续序列。非限制性实施例包括在RNA中的UAG、UAA、UGA,和在DNA中的TAG、TAA或TGA。除非另有说明,否则该术语还包括在DNA或RNA内的引入提前终止密码子的无义突变,这导致所产生的任何蛋白被异常缩短。对应于各终止密码子的tRNA的特定名称为:琥珀(UAG),赭石(ochre)(UAA),和蛋白石(opal)(UGA)。

[0203] “转移核糖核酸”或“tRNA”是帮助将mRNA翻译成蛋白的核酸分子。tRNA具有独特的折叠结构,包括三个发夹环;这些环之一包括编码反密码子的“茎”部分。反密码子识别在mRNA上的相应密码子。每个tRNA都“负荷有”一个与mRNA密码子相对应的氨基酸;这种“负荷”通过酶tRNA合成酶来实现。当tRNA识别出对应于其反密码子的密码子时,tRNA将其负荷的氨基酸转移到正在生长的氨基酸链上,以形成多肽或蛋白。内源性tRNA可以由内源性tRNA合成酶负荷。因此,内源性tRNA通常负荷经典氨基酸。来自外部来源的正交tRNA需要相应的正交tRNA合成酶。此类正交tRNA可以负荷有经典氨基酸和非经典氨基酸。在一些实施方式中,tRNA所负荷的氨基酸可以被可检测地标记,以使得能够在体内进行检测。标记技术是本领域已知的,包括但不限于点击化学,其中含非天然氨基酸的叠氮化物/炔烃通过正交tRNA/合成酶对来添加,因此可以使用包含荧光团或其他此类分子的炔烃/叠氮化物进行检测。

[0204] 如本文所用的术语“治疗”、“处理”等在本文中用来表示,获得期望的药理和/或生理学效果。该效果就完全或部分预防疾病、障碍或病症或其病征或症状而言可以是预防性的,和/或,就部分或完全治愈障碍和/或归因于该障碍的不良作用而言可以是治疗性的。

[0205] 如本文所用的术语“载体”可以指经设计用于在不同宿主之间进行转移的核酸构建体,包括但不限于质粒、病毒、粘粒、噬菌体、BAC、YAC等。“病毒载体”被定义为,重组产生的病毒或病毒颗粒,其包括待在体内、离体或体外递送至宿主细胞中的多核苷酸。在一些实施方式中,质粒载体可以由市售载体来制备。在其他实施方式中,病毒载体可以根据本领域已知的技术,由杆状病毒、逆转录病毒、腺病毒、AAV等来制备。在一个实施方式中,病毒载体是慢病毒载体。病毒载体的实施例包括逆转录病毒载体、腺病毒载体、腺相关病毒载体、α病毒载体等。基于传染性烟草花叶病毒(TMV)的载体可用于制造蛋白,并且据报道已在烟草叶片中表达格瑞弗森(Griffithsin)(O'Keefe et al.(2009)Proc.Nat.Acad.Sci.USA 106(15):6099-6104)。α病毒载体,例如基于塞姆利基森林病毒(Semliki Forest virus)的载体和基于辛德毕斯病毒(Sindbis virus)的载体,也已被开发用于基因治疗和免疫治疗。参见,Schlesinger&Dubensky(1999)Curr.Opin.Biotechnol.5:434-439and Ying et al.

(1999)Nat.Med.5(7):823-827。在基因转移是通过逆转录病毒载体介导的方面,载体构建体可以指一多核苷酸,其包括逆转录病毒基因组或其部分以及感兴趣基因。关于用于基因转移的载体的现代方法的更多细节可以在例如Kotterman et al.(2015)Viral Vectors for Gene Therapy:Translational and Clinical Outlook Annual Review of Biomedical Engineering 17(用于基因治疗的病毒载体:生物医学工程的转化和临床展望年度回顾17)中找到。包括启动子和克隆位点的载体是本领域众所周知的,其中多核苷酸可以可操作地连接进入该克隆位点。这样的载体能够在体外或体内转录RNA,并且可以从诸如安捷伦科技公司(Agilent Technologies)(加利福尼亚州圣克拉拉)和普罗米加生物技术公司(Promega Biotech)(威斯康星州麦迪逊)等来源商购获得。在一方面,启动子是pol III启动子。

[0206] 用于施用AdRNA的药物组合物可以常规地以剂量单位形式存在,并且可以通过药学领域公知的任何方法来制备。药物组合物可以例如通过以下方式来制备:将本文提供的化合物均匀且紧密地与液体载剂、细碎的固体载剂或二者结合,然后,如果需要,将产品成型为所需制剂。在药物组合物中,以足以产生所需治疗效果的量包含本文提供的化合物。例如,该技术的药物组合物可以采取几乎适于任何施用方式,包括例如局部、眼部、经口、经颊、全身、经鼻、注射、输注、经皮、直肠和阴道施用的形式,或适于通过吸入或吹入的形式。

[0207] 对于局部施用,如本领域公知的,化合物可以配制成溶液、凝胶剂、软膏剂、乳膏剂、混悬剂等。

[0208] 全身性制剂包括设计用于通过注射施用的那些(例如,皮下、静脉内、输注、肌内、鞘内或腹腔注射),以及设计用于经皮、经粘膜、经口或经肺部施用的那些。

[0209] 有用的可注射制剂包括在水性或油性媒介物中的、本文提供的化合物的无菌悬浮剂、溶液或乳剂。所述组合物还可以含有配制试剂,例如悬浮剂、稳定剂和/或分散剂。注射用制剂可以单位剂型存在,例如在安瓿或多剂量容器中,并且可以包含添加的防腐剂。

[0210] 或者,可注射制剂可以以粉末形式提供,用于在使用前与合适的媒介物进行重构,该合适的媒介物包括但不限于无菌的、无热原的水、缓冲液和右旋糖溶液。为此,本文提供的化合物可通过任何本领域已知的技术(例如冻干)进行干燥,并在使用前进行重构。

[0211] 对于经粘膜施用,在制剂中使用适合于要渗透的屏障的渗透剂。这样的渗透剂是本领域已知的。

[0212] 对于经口施用,药物组合物可以采用例如锭剂、片剂或胶囊剂的形式,其通过常规方式与药学上可接受的赋形剂例如粘合剂(例如,预糊化的玉米淀粉、聚乙烯吡咯烷酮或羟丙基甲基纤维素);填充剂(例如,乳糖、微晶纤维素或磷酸氢钙);润滑剂(例如,硬脂酸镁、滑石粉或氧化硅);崩解剂(例如,马铃薯淀粉或羟乙酸淀粉钠);或,润湿剂(例如,月桂基硫酸钠)来制备。片剂可以通过本领域众所周知的方法,用例如糖、薄膜或肠溶衣包衣来包被。

[0213] 可以根据本领域已知的用于制备药物组合物的任何方法来制备旨于经口使用的组合物,并且这样的组合物可以包含选自由甜味剂、调味剂、着色剂和防腐剂组成的组中的一种或多种试剂,以提供药学上精美且可口的制剂。片剂含有本文提供的化合物与适合于制造片剂的、无毒的、药学上可接受的赋形剂的混合物。这些赋形剂可以是,例如惰性稀释剂,例如碳酸钙、碳酸钠、乳糖、磷酸钙或磷酸钠;制粒剂和崩解剂(例如,玉米淀粉或海藻酸);粘合剂(例如,淀粉、明胶或阿拉伯胶);和,润滑剂(例如,硬脂酸镁、硬脂酸或滑石粉)。

片剂可以不包被，或它们可以通过已知技术进行包被以延迟在胃肠道中的崩解和吸收，从而在更长的时间内提供持续作用。例如，可以采用延迟材料，例如单硬脂酸甘油酯或二硬脂酸甘油酯。它们也可以通过技术人员熟知的技术进行包被。该技术的药物组合物也可以是水包油乳剂的形式。

[0214] 用于经口施用的液体制剂可以采取例如酏剂、溶液、糖浆剂或混悬剂的形式，或者它们可以以干燥产品的形式存在，用于在使用前用水或其他合适的媒介物进行重构。这样的液体制剂可以通过常规方式，与以下药学上可接受的添加剂一起制备：例如悬浮剂(例如，山梨糖醇糖浆、纤维素衍生物或氢化食用脂肪)；乳化剂(例如，卵磷脂或阿拉伯胶)；非水性媒介物(例如，杏仁油、油性酯，乙醇、cremophore^{TM}或分馏植物油)；和，防腐剂(例如，对羟基苯甲酸甲酯或对羟基苯甲酸丙酯或山梨酸)。该制剂还可以适当地包含缓冲盐、防腐剂、调味剂、着色剂和甜味剂。

[0215] “施用”可以在整个治疗过程中，连续或间歇地以一剂进行。确定最有效的施用手段和剂量的方法是本领域技术人员已知的，并且可以随用于治疗的组合物、治疗的目的、所治疗的靶细胞和所治疗的受试者而变化。可以根据治疗医师选择的剂量水平和模式，进行单次或多次施用。合适的剂量配方和施用试剂的方法是本领域已知的。也可以确定施用途径，并且确定最有效的施用途径的方法是本领域技术人员已知的，并且可以随用于治疗的组合物、治疗目的、所治疗的受试者的健康状况或疾病阶段，以及靶细胞或组织而变化。施用途径的非限制性实施例包括经口施用、经鼻施用、注射和局部应用。

[0216] 施用可以指，可用于使化合物或组合物(例如，DNA构建体、病毒载体或其他)能够递送至所需的生物作用位点的方法。这些方法可以包括对表面(例如皮肤)的外表面进行局部施用(例如，洗剂、乳膏、软膏)。这些方法可以包括肠胃外施用(包括静脉内、皮下、鞘内、腹腔、肌内、血管内或输注)、经口施用、吸入施用、十二指肠内施用、直肠施用。在某些情况下，受试者可以在没有监督的情况下施用该组合物。在某些情况下，受试者可以在医学专业人员(例如，医师、护士、医师助手、护理员、临终关怀工作者等)的监督下施用组合物。在某些情况下，医学专业人员可以施用该组合物。在某些情况下，整容专业人员可以施用该组合物。

[0217] 本文公开的组合物的施用或应用可进行连续或不连续的至少约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、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100天的治疗时间。在某些情况下，治疗时间可以是约1至约30天、约2至约30天、约3至约30天、约4至约30天、约5至约30天、约6至约30天、约7至约30天、约8至约30天、约9至约30天、约10至约30天、约11至约30天、约12至约30天、约13至约30天、约14至约30天、约15至约30天、约16至约30天、约17至约30天、约18至约30天、约19至约30天、约20至约30天、约21至约30天、约22至约30天、约23至约30天、约24至约30天、约25至约30天、约26至约30天、约27至约30天、约28至约30天或约29至约30天。

[0218] 本文公开的组合物的施用或应用一天可进行至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23或24次。在某些情况下，本文公开的组合物的施用或

应用一周可进行至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或21次。在某些情况下，本文公开的组合物的施用或应用一个月可进行至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、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或90次。

[0219] 在某些情况下，组合物可以作为单剂量或分开的剂量来施用/应用。在某些情况下，本文所述的组合物可以在第一时间点和第二时间点施用。在某些情况下，可以施用组合物，使得第一次施用先于另一次施用，其施用时间相差1小时、2小时、4小时、8小时、12小时、16小时、20小时、1天、2天、4天、7天、2周、4周、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月、1年或更长时间。

[0220] 术语“有效量”可以指足以实现所需效果的量。在治疗或预防应用的背景下，有效量将取决于所讨论的病症的类型和严重性，以及个体受试者的特征，例如总体健康、年龄、性别、体重和对药物组合物的耐受性。在免疫原性组合物的背景下，在一些实施方式中，有效量是足以引起针对病原体的保护性应答的量。在其他实施方式中，免疫原性组合物的有效量是足以引起针对抗原的抗体生成的量。在一些实施方式中，有效量是为有需要的受试者赋予被动免疫所需的量。关于免疫原性组合物，在一些实施方式中，除了上述因素之外，有效量还可以取决于预期用途，特定抗原性化合物的免疫原性程度，以及受试者免疫系统的健康/应答性。技术人员可以根据这些和其他因素来确定适当的量。

[0221] 在体外应用的情况下，在一些实施方式中，有效量可以取决于所讨论的应用的大小和性质。它还可取决于体外靶标的性质和敏感性，以及所使用的方法。技术人员可以基于这些和其他考虑因素来确定有效量。根据实施方案，有效量可以包括组合物的一次或多次施用。

[0222] 无需明确引证就可以推断出，除非另有规定，否则当本公开涉及多肽、蛋白、多核苷酸或抗体时，其等同物或生物学等同物旨于在本公开的范围内。如本文所用的术语“其生物等同物”在指代参考蛋白、抗体、多肽或核酸时，旨在与“其等同物”同义，意指具有最小同源性，同时仍保持所需的结构或功能的那些。除非本文具体叙述，否则预期本文提及的任何多核苷酸、多肽或蛋白也包括其等同物。例如，等同物意指至少约70％的同源性或同一性，或至少80％的同源性或同一性，并且可选地，或至少约85％，或可选地至少约90％，或可选地至少约95％，或可选地98％的同源性或同一性，并且显示出与参考蛋白、多肽或核酸的基本等同的生物学活性。或者，当提及多核苷酸时，其等同物是在严格条件下与参考多核苷酸或其互补序列杂交的多核苷酸。

[0223] 本公开提供了用于下文描述的基因和蛋白编辑技术的多肽和/或多核苷酸序列。应该理解，尽管并非总是明确指出，本文提供的序列可用于提供表达产物，以及产生具有相同生物学特性的蛋白的基本相同的序列。这些“生物等同”或“生物活性”多肽由本文所述的等同多核苷酸编码。当使用在默认条件下运行的序列同一性方法进行比较时，它们可以具有与参考多肽至少60％，或可选地至少65％，或可选地至少70％，或可选地至少75％，或可选地至少80％，或可选地至少85％，或可选地至少90％，或可选地至少95％，或可选地至少98％相同的一级氨基酸序列。提供了特定的多肽序列作为特定实施方式的实施例。用具有

相似电荷的备用氨基酸对序列的氨基酸进行修饰。另外，等同多核苷酸是在严格条件下与参考多核苷酸或其互补序列杂交的多核苷酸，或者关于多肽，指由在严格条件下与编码参考物的多核苷酸或其互补链杂交的多核苷酸编码的多肽。或者，等同多肽或蛋白是由等同多核苷酸表达的多肽或蛋白。

[0224] “组合物”通常是指活性剂，例如本公开的adRNA，化合物或组合物，和天然存在的或非天然存在的载剂，惰性(例如，可检测到的试剂或标记物)或活性物质，例如佐剂、稀释剂、粘合剂、稳定剂、缓冲剂、盐、亲脂性溶剂、防腐剂、佐剂等的组合，并包括药学上可接受的载剂。载剂还包括药学赋形剂和添加剂蛋白、肽、氨基酸脂质和碳水化合物(例如，糖，包括单糖、双低聚糖、三低聚糖、四低聚糖，和低聚糖；衍生化糖，如醛糖醇、醛糖酸、酯化糖等；和多糖或糖聚合物)，它们可以单独或以组合形式存在，按重量或体积计，单独或组合占1～99.99％。示例性的蛋白赋形剂包括血清白蛋白，例如人血清白蛋白(HSA)、重组人白蛋白(rHA)、明胶、酪蛋白等。还可起到缓冲作用的代表性氨基酸/抗体组分，包括丙氨酸、精氨酸、甘氨酸、精氨酸、甜菜碱、组氨酸、谷氨酸、天冬氨酸、半胱氨酸、赖氨酸、亮氨酸、异亮氨酸、缬氨酸、蛋氨酸、苯丙氨酸、阿斯巴甜等。碳水化合物赋形剂也旨于涵盖在该技术范围内，其实例包括但不限于单糖，例如果糖、麦芽糖、半乳糖、葡萄糖、D-甘露糖、山梨糖等；双糖，例如乳糖、蔗糖、海藻糖、纤维二糖等；多糖，例如棉子糖、松三糖、麦芽糖糊精、葡聚糖、淀粉等；醛糖醇，例如甘露醇、木糖醇、麦芽糖醇、乳糖醇、木糖醇山梨糖醇(葡萄糖醇)和肌醇。

[0225] 根据本公开使用的组合物，包括细胞、治疗、疗法、药剂、药物和药物制剂，可以包装成剂量单位的形式，以便于施用和剂量均匀。术语“单位剂量”或“剂量”可以指适合在受试者中使用的物理上离散的单位，每个单位包含预定量的组合物，该预定量经计算以与其施用(即适当的途径和方案)相关地产生期望的应答。根据治疗次数和单位剂量所施用的量取决于期望的结果和/或保护。组合物的精确量还取决于执业医师的判断，并且对于每个人而言都是独特的。影响剂量的因素包括受试者的身体和临床状态、施用途径、预期的治疗目标(缓解症状与治愈)，以及特定组合物的效力、稳定性和毒性。在配制后，溶液可以以与剂量制剂相容的方式来施用，所施用的量是治疗或预防上有效的量。制剂可以容易地以多种剂型来施用，例如本文所述的可注射溶液的类型。

[0226] 如本文所用的术语“减少或消除其表达和/或功能”可以指减少或消除所述多核苷酸向mRNA的转录，或可选地减少或消除所述mRNA向肽、多肽或蛋白的翻译，或减少或消除所述肽、多肽或蛋白的功能。在非限制性实施例中，将多核苷酸向mRNA的转录减少至其在野生型细胞中发现的正常水平的至少一半。

[0227] 短语“一线”或“二线”或“三线”可以指患者接受的治疗顺序。一线疗法方案是首先给予的治疗，而二线或三线疗法则是分别在一线疗法后或在二线疗法后进行。美国国家癌症研究所将一线疗法定义为“针对疾病或病症的首次治疗。在癌症患者中，主要治疗可以是手术、化学疗法、放射疗法或这些疗法的组合。本领域技术人员也将一线疗法称为“主要疗法和主要治疗”。参见美国国家癌症研究所的网站cancer.gov，最新访问时间为2017年11月15日。通常，因为患者对一线疗法未显示出阳性的临床或亚临床应答，或一线疗法已经停止了，而对该患者进行后续化疗方案。

[0228] 术语“接触”是指两个或更多个实体之间的直接或间接的结合或相互作用。直接交

互的一个特定实施例是结合。间接相互作用的一个特定实施例是其中一个实体作用于中间分子，而中间分子转而作用于第二个提及的实体。本文所用的接触包括在溶液中、在固相中、在体外、离体、在细胞中和在体内。体内的接触可以称为施用或给药。

[0229] “冷冻保护剂”是本领域已知的，包括但不限于例如蔗糖、海藻糖和甘油。通常使用在生物系统中表现出低毒性的冷冻保护剂。

[0230] 本文公开了，用于在不存在ADAR酶过表达的情况下，进行RNA的位点特异性编辑的adRNA。本文进一步提供了对DNA进行改造A->G编辑。另外，本文提供了，筛选能够进行RNA和DNA的位点特异性C->T编辑的ADAR2突变体。本文还进一步提供了，经由使用与ACF一起表达的APOBEC1，进行RNA的改造C->T编辑。

[0231] 与其他ADAR2系统相比，本公开是独特的，因为它提出了募集内源性ADAR以催化治疗性RNA编辑的新方法。另外，现有技术的系统都没有提供使用ADAR酶来改造C->T编辑的手段。最后，它们没有公开使用APOBEC进行可编程的位点特异性RNA编辑。

[0232] 本文公开了示例性adRNA包括与靶标RNA互补的RNA靶向结构域，和一种或多种ADAR募集结构域。当与其靶标结合时，adRNA能够将ADAR酶募集至靶标RNA。然后，该ADAR酶能够催化靶标腺苷转化至肌苷。不受理论约束的情况下，据信adRNA可以类似地用于募集ADAR2突变体或APOBEC1之一，来影响C->T的RNA的编辑。

[0233] 本文还公开了，使用经改造的adRNA进行了体外和体内实验，以募集内源性ADAR酶。本文还公开了，示出ADAR突变体以及APOBEC1/ACF构建体的C->T编辑效率的实验。

[0234] 如本文所述的病毒载体可以包括，编码至少一种RNA编辑实体募集结构域的核酸序列。在某些情况下，核酸序列可以编码一种以上的RNA编辑实体募集结构域，例如2、3、4或更多种。RNA编辑实体募集结构域可以包括与以下项中的至少一种的至少约80％的序列同一性：Alu结构域、载脂蛋白B mRNA编辑催化多肽样(APOBEC)募集结构域、Cas13结构域、GluR2结构域或其任意组合。所述募集结构域可以包括一种或多种GluR2结构域、一种或多种Alu结构域、一种或多种APOBEC结构域，或它们的任意组合。所述募集结构域可以包括一种以上的GluR2结构域、一种以上的Alu结构域、一种以上的APOBEC结构域、Cas13结构域，或它们的任意组合。所述募集结构域可能不包括Alu结构域。所述募集结构域可能不包括GluR2结构域。所述募集结构域可能不包括APOBEC结构域。所述募集结构域可能不包括Cas13结构域。

[0235] APOBEC募集结构域可以包括APOBEC1募集结构域、APOBEC2募集结构域、APOBEC3A募集结构域、APOBEC3B募集结构域、APOBEC3C募集结构域、APOBEC3D募集结构域、APOBEC3E募集结构域、APOBEC3F募集结构域、APOBEC3G募集结构域、APOBEC3H募集结构域、APOBEC4募集结构域，这些结构域任一种的衍生物，或它们的任意组合。

[0236] 募集结构域可以包括与如本文所述的Alu结构域的任一种的至少约80％的序列同一性。在某些情况下，所述募集结构域可以包括与如本文所述的Alu结构域的任一种的至少约：85％、90％、95％、97％、98％，或99％的序列同一性。

[0237] 募集结构域可以包括与如本文所述的APOBEC结构域的任一种的至少约80％的序列同一性。在某些情况下，所述募集结构域可以包括与如本文所述的APOBEC结构域的任一种的至少约85％、90％、95％、97％、98％或99％的序列同一性。

[0238] 募集结构域可以包括与如本文所述的GluR2结构域的任一种的至少约80％的序列

同一性。在某些情况下，所述募集结构域可以包括与如本文所述的GluR2结构域的任一种的至少约85％、90％、95％、97％、98％或99％的序列同一性。

[0239] 募集结构域可以包括与如本文所述的Cas13结构域的任一种的至少约80％的序列同一性。在某些情况下，所述募集结构域可以包括与如本文所述的Cas13结构域的任一种的至少约85％、90％、95％、97％、98％，或99％的序列同一性。

[0240] 核酸序列可以编码至少1、2、3、4、5个RNA编辑募集结构域。核酸序列可以编码至少2个RNA编辑募集结构域，其中一个是Alu结构域。核酸序列可以编码至少2个RNA编辑募集结构域，其中一个是APOBEC结构域。核酸序列可以编码至少2个RNA编辑募集结构域，其中一个是GluR2结构域。核酸序列可以编码至少2个RNA编辑募集结构域，其中一个是Cas13结构域。

[0241] 募集结构域可以包括一个或多个茎环结构。募集结构域可以包括至少2个茎环结构。募集结构域可以包括至少3个茎环结构。募集结构域可能不包括茎环结构。包括至少一个茎环结构的募集结构域可以是Alu结构域、APOBEC结构域、GluR2结构域、Cas13结构域或其任意组合。

[0242] 募集结构域的至少一部分可以是单链的。在某些情况下，Alu结构域可以是至少部分单链的。在某些情况下，APOBEC结构域可以是至少部分单链的。在某些情况下，GluR2结构域可以是至少部分单链的。在某些情况下，Cas13结构域可以是至少部分单链的。

[0243] 募集结构域可以包括多个重复序列。募集结构域可以包括多个Alu重复序列。

[0244] 在某些情况下，病毒载体可以包括一种或多种RNA编辑募集结构域。在某些情况下，病毒载体可以包括一种以上的RNA编辑募集结构域。在某些情况下，病毒载体可以包括2、3、4、5或更多种RNA编辑募集结构域。核酸序列可以编码一种或多种RNA编辑募集结构域。核酸序列可以编码一种以上的RNA编辑募集结构域。核酸序列可以编码2、3、4、5或更多种RNA编辑募集结构域。核酸序列可以编码至少Alu结构域和GluR2结构域。核酸序列可以编码至少Alu结构域和Cas13结构域。核酸序列可以编码至少Alu结构域和APOBEC结构域。核酸序列可以编码至少GluR2结构域和APOBEC结构域。核酸序列可以编码至少GluR2结构域和Cas13结构域。核酸序列可以编码至少Cas13结构域和APOBEC结构域。

[0245] 核酸序列可以编码可以与靶标RNA的至少一部分互补的靶标RNA。它可以与该靶标RNA的至少一部分互补。可以互补的部分的长度可以是约50个碱基对(bp)至约200bp。可以互补的部分的长度可以是约20bp至约100bp。可以互补的部分的长度可以是约10bp至约50bp。可以互补的部分的长度可以是约50bp至约300bp。修改互补部分的长度，可以提高编辑效率。在某些情况下，与较短的长度相比，该部分的较长的长度可以提高编辑效率。

[0246] 核酸序列可以编码至少一种RNA编辑实体募集结构域，和编码可以与靶标RNA的至少一部分互补的RNA的核酸序列，并且包括长度为至少约200bp的连续核酸序列。该连续核酸序列可以包括长度上约100bp至约300bp的长度。该连续核酸序列可以包括长度上约150bp至约400bp的长度。该连续核酸序列可以包括长度上约200bp至约500bp的长度。该连续核酸序列可以包括长度上约50bp至约300bp的长度。修改该连续核酸序列的长度，可以提高编辑效率。在某些情况下，与较短的长度相比，连续序列的较长的长度可以提高编辑效率。

[0247] 核酸可以包括接头序列，例如位于靶向结构域和募集结构域之间的接头序列。在某些情况下，核酸可以包括序列例如5'-X-(Y-X')n-L-Z-3'，其中X与特定位置下游的靶标

RNA序列互补，X'与特定位置上游的靶标RNA序列互补，Y包括一个或多个可能不与靶标RNA序列互补的核苷酸，n可以是1到10的整数，L可以是包括任意数量的核苷酸(包括0)的接头序列，并且Z可以是被RNA编辑实体识别并与其结合的序列。L也可以由不同的化学键组成，例如(寡)肽连接或PEG连接。

[0248] 核酸可以包括20至数百个核苷酸。在某些情况下，较长的靶向部分对于待编辑的RNA序列的靶位点提供更高的特异性，较小的由于非故意(脱靶)结合而产生的脱靶效应，以及具有更多的空间以在靶向部分自身内创建二级结构(例如茎环结构、十字形、立足点形结构)，错配或摇摆碱基(由于在靶标RNA序列中的、在待编辑位点处或其附近的一个或多个互补碱基不匹配)，等等。在某些情况下，除了与待编辑的核苷酸相对的错配之外，靶向部分可以在靶向部分的整个长度上与靶标RNA序列互补，并且任选地具有1或2摇摆碱基。

[0249] 核酸可以使用各种化学方法和修饰来修饰。在某些情况下，可以通过对磷酸二酯键进行单硫代或二硫代，以分别产生硫代磷酸酯或二硫代磷酸酯，来改变核苷酸之间的规则核苷间连接。对核苷间连接的其他修饰可包括酰胺化或肽接头。可以通过用低级烷基(C1-4，例如2'-O-Me)、烯基(C2-4)、炔基(C2-4)、甲氧基乙基(2'-MOE)或其他取代基来取代2'-O部分，来修饰核糖。在某些情况下，2'OH基团的取代基可以包括甲基、甲氧基乙基或3,3'-二甲基丙烯基。在某些情况下，可以应用锁核酸序列(LNA)，该锁核酸序列(LNA)在核糖环内部包含2'-4'分子内桥(例如，2'氧和4'碳之间的亚甲基桥)键。可以例如通过杂环的氨基化或脱氨基，对嘌呤核碱基和/或嘧啶核碱基进行修饰，以改变它们的性质。

[0250] 病毒载体可以是腺相关病毒(AAV)载体。AAV可以是重组的AAV。AAV可以包括AAV1血清型、AAV2血清型、AAV3血清型、AAV4血清型、AAV5血清型、AAV6血清型、AAV7血清型、AAV8血清型、AAV9血清型，这些任一种的衍生物，或其任意组合。AAV可以选自由以下项组成的组：AAV1血清型、AAV2血清型、AAV3血清型、AAV4血清型、AAV5血清型、AAV6血清型、AAV7血清型、AAV8血清型、AAV9血清型，这些任一种的衍生物，及其任意组合。病毒载体可以是经修饰的病毒载体。病毒载体可以是经修饰的，以包含经修饰的蛋白。在某些情况下，病毒载体可以包括经修饰的VP1蛋白。

[0251] 编码RNA编辑实体募集结构域、靶向结构域或其组合的核酸序列，可以包括基本上是十字形的结构(例如二级结构)。核酸序列可以包括至少两个可以基本上是十字形的结构。募集结构域可以包括可以基本上是十字形的结构。募集结构域可以包括至少两个可以基本上是十字形的结构。可以对核酸序列(例如，编码募集结构域的部分)的二级结构进行修饰，以增强ADAR的募集或结合。对结构进行修饰以增强ADAR的募集或结合，可以包括形成十字形结构。

[0252] RNA编辑实体募集结构域可以位于至少两个可以基本上是十字形的结构之间。靶向结构域可以位于至少两个可以基本上是十字形的结构之间。RNA编辑实体募集结构域的两侧可以有至少一个可以基本上是十字形的结构。靶向结构域的两侧可以有至少一个可以基本上是十字形的结构。

[0253] 十字形结构可以包括与十字形结构的至少一对至少部分互补的链邻接的茎环。十字形结构可以基本上是十字形。十字形结构可以包括基本上少于十字形，例如4个茎环中的3个，或4个茎环中的2个。可以形成十字形的一个或多个茎环可以包括不同的长度。可以形成十字形的一个或多个茎环可以包括相同的长度。可以形成十字形的一个或多个茎环可以

包括一个或多个错配的凸起。

[0254] RNA编辑实体募集结构域可以包括基本上可以是立足点形的结构。RNA编辑实体可以包括一个或多个错配凸起。RNA编辑实体可能不包括错配凸起。RNA编辑实体募集结构域可以包括基本上立足点形的结构、基本上十字形的结构、基本上线性的结构、茎环结构、双茎环结构或其组合。

[0255] 病毒载体可以包括编码具有二维形状的RNA的核酸序列。该二维形状与具有不同二维形状的RNA相比，可以表现处优异的ADAR的募集或结合。可以对编码RNA的核酸序列的序列进行修饰，使得该RNA包括二维形状，该二维形状表现出优异的ADAR的募集或结合。该二维形状可以是基本上十字形、立足点形、茎环或其任意组合。该二维形状可以包括基本上十字形。该二维形状可以包括立足点形。该二维形状可以包括茎环。该二维形状可以是线性的。

[0256] 由核酸序列编码的RNA可以包括第一结构域和第二结构域。第一结构域可以包括十字形，第二结构域可以包括线性结构。第一结构域和第二结构域可以直接或间接连接。第一结构域可以是募集结构域，第二结构域可以是靶向结构域。该RNA可以包括第三结构域。该第三域可以直接或间接地连接到第一结构域或第二结构域。该第三结构域可以是募集结构域。该第三结构域可以包括十字形结构。

[0257] 由核酸序列编码的RNA可以是非天然存在的RNA。由核酸序列编码的RNA可以包括至少一个碱基或至少一个糖包括化学修饰。RNA可以包括两种或更多种化学修饰。化学修饰可以增加RNA的稳定性，例如具有RNA在体内的生物活性半衰期。

[0258] 核酸可以包括一种或多种募集结构域和一种或多种反义结构域。如通过序列方法(如Sanger方法)确定的，当该核酸与RNA编辑实体和与反义结构域的至少一部分互补的靶标核酸接触时，它可以以比与Cas13b蛋白或其活性片段复合的可比较核酸，高至少约2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍的效率，来修饰靶标核酸的至少一个碱基对。该效率可以是高至少约3倍。该效率可以是高至少约4倍。该效率可以高至少约5倍。

[0259] 核酸可以包括一种或多种募集结构域和一种或多种反义结构域。如通过序列方法(例如Sanger方法)确定的，当该核酸与RNA编辑实体和与反义结构域的至少一部分互补的靶标核酸接触时，它可以以比与GluR2蛋白或其活性片段复合的可比较核酸，高至少约2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍的效率，来修饰靶标核酸的至少一个碱基对。该效率可以是高至少约3倍。该效率可以是高至少约4倍。该效率可以是高至少约5倍。

[0260] 与天然募集结构域(例如GluR2结构域)的至少一部分相比，如本文所述的核酸可以提供更高的编辑效率。与经修饰的募集结构域(例如，经修饰的GluR2结构域)的至少一部分相比，核酸可提供更高的编辑效率。

[0261] 靶标核酸可以包括RNA。该RNA可以是mRNA。该RNA可以编码蛋白或其部分。该蛋白或其部分的功能失调可能与疾病或病症有关。如本文所述的组合物、载体、核酸、非天然存在的RNA的施用可以治疗、消除、治愈或减轻疾病或病症的一种或多种症状。

[0262] 疾病或病症可以包括神经退行性疾病、肌肉障碍、代谢障碍、眼部障碍或其任意组合。该疾病或病症可以包括囊性纤维化，白化病，α1-抗胰蛋白酶缺乏症，阿尔茨海默氏病，肌萎缩性侧索硬化症，哮喘，β-地中海贫血，Cadasil综合征，腓骨肌萎缩症(Charcot-Marie-Tooth disease)，慢性阻塞性肺疾病(COPD)，远端脊髓性肌萎缩症(DSMA)，杜氏/贝

克肌肉萎缩症，营养不良性大疱性表皮松解症，大疱性表皮松解症，法布里病(Fabry disease)，因子V莱顿相关疾病，家族性腺瘤，息肉病，半乳糖血症，高雪氏病，6-磷酸葡萄糖脱氢酶，血友病，遗传性血色素沉着病，亨特氏综合征，亨廷顿氏病(Huntington's disease)，贺勒氏(Hurler)综合征，炎性肠病(IBD)，遗传性多凝集综合征，莱伯氏先天性黑蒙病，Lesch-Nyhan综合征，林奇(Lynch)综合征，马凡(Marfan)综合征，粘多糖贮积症，肌肉萎缩症，I型和II型强直性肌营养不良症，神经纤维瘤病，A、B和C型尼曼匹克(Niemann-Pick)病，NY-eso1相关癌症，帕金森氏病，Peutz-Jeghers综合征，苯丙酮尿症，庞贝氏病，原发性睫状病，与凝血酶原突变相关的疾病，例如凝血酶原G20210A突变，肺动脉高压，色素性视网膜炎，桑德霍夫(Sandhoff)病，重度合并免疫缺陷综合症(SCID)，镰状细胞性贫血，脊髓性肌萎缩症，斯塔格氏(Stargardt)病，泰-萨克斯(Tay-Sachs)病，尤塞氏(Usher)综合征，X连锁免疫缺陷，各种形式的癌症(例如，BRCA1和2连锁的乳腺癌和卵巢癌)。该疾病或病症可以包括肌肉萎缩症、鸟氨酸转氨甲酰酶缺乏症、色素性视网膜炎、乳腺癌、卵巢癌、阿尔茨海默氏病、疼痛、Stargardt黄斑肌营养不良、腓骨肌萎缩症、瑞特(Rett)综合征或它们的任意组合。组合物的施用可足以：(a)相对于施用前的基因表达，减少基因表达；(b)在受试者(例如有此需要的受试者)中，编辑至少一个点突变；(c)在受试者中，编辑至少一个终止密码子，以产生终止密码子的通读；(d)在受试者中，产生外显子跳跃；或，(e)它们的任意组合。

[0263] 药物组合物可以包括第一活性成分。该第一活性成分可以包括如本文所述的病毒载体、如本文所述的非天然存在的RNA或如本文所述的核酸。该药物组合物可以被配制成单位剂型。该药物组合物可以包括药学上可接受的赋形剂、稀释剂或载剂。该药物组合物可以包括第二、第三或第四活性成分。

[0264] 本文所述的组合物可以包括赋形剂。赋形剂可以添加到干细胞中，也可以与干细胞一起从它的来源中共分离。赋形剂可以包括冷冻防腐剂，例如DMSO、甘油、聚乙烯吡咯烷酮(PVP)或其任意组合。赋形剂可以包括冷冻防腐剂，例如蔗糖、海藻糖、淀粉，这些的任一种的盐，这些的任一种的衍生物，或它们的任意组合。赋形剂可以包括pH剂(以最大程度地减少组合物组分的氧化或降解)，稳定剂(以防止组合物组分的修饰或降解)，缓冲剂(以增强温度稳定性)，增溶剂(以增加蛋白溶解度)，或它们的任意组合。赋形剂可以包括表面活性剂、糖、氨基酸、抗氧化剂、盐、非离子表面活性剂、增溶剂、甘油三酸酯、醇或其任意组合。赋形剂可以包括碳酸钠、乙酸盐、柠檬酸盐、磷酸盐，聚乙二醇(PEG)、人血清白蛋白(HSA)、山梨糖醇、蔗糖、海藻糖、聚山梨酸酯80、磷酸钠、蔗糖、磷酸二钠、甘露醇、聚山梨酯20、组氨酸、柠檬酸盐、白蛋白、氢氧化钠、甘氨酸、柠檬酸钠、海藻糖、精氨酸、乙酸钠、乙酸盐、HCl、乙二胺四乙酸二钠、卵磷脂、甘油、黄原胶、大豆异黄酮、聚山梨酯80、乙醇、水，替普瑞酮(teprenone)或其任意组合。赋形剂可以是美国药学会(American Pharmaceutical Association)的《药用赋形剂手册》(Handbook of Pharmaceutical Excipients)(1986)中描述的赋形剂。

[0265] 合适的赋形剂的非限制性实施例可以包括缓冲剂、防腐剂、稳定剂、粘合剂、压实剂、润滑剂、螯合剂、分散增强剂、崩解剂、调味剂、甜味剂、着色剂。

[0266] 在某些情况下，赋形剂可以是缓冲剂。合适的缓冲剂的非限制性实施例可以包括柠檬酸钠、碳酸镁、碳酸氢镁、碳酸钙和碳酸氢钙。作为缓冲剂，可以在药物制剂中使用碳酸

氢钠，碳酸氢钾，氢氧化镁，乳酸镁，葡萄糖酸镁，氢氧化铝，柠檬酸钠，酒石酸钠，乙酸钠，碳酸钠，多磷酸钠，多磷酸钾，焦磷酸钠，焦磷酸钾，磷酸氢二钠、磷酸氢二钾，磷酸三钠，磷酸三钾，偏磷酸钾，氧化镁，氢氧化镁，碳酸镁，硅酸镁，乙酸钙，甘油磷酸钙，氯化钙，氢氧化钙，和其它钙盐，或它们的组合。

[0267] 在某些情况下，赋形剂可以包括防腐剂。合适的防腐剂的非限制性实施例可以包括抗氧化剂，例如α-生育酚和抗坏血酸酯，以及抗微生物剂，例如苯甲酸酯类、氯丁醇和苯酚。抗氧化剂可进一步包括，但不限于EDTA，柠檬酸，抗坏血酸，丁基化羟基甲苯(BHT)，丁基化羟基茴香醚(BHA)，亚硫酸钠，对氨基苯甲酸，谷胱甘肽，没食子酸丙酯，半胱氨酸，甲硫氨酸，乙醇和N-乙酰半胱氨酸。在某些情况下，防腐剂可包括有效霉素A，TL-3，原钒酸钠，氟化钠，N-a-甲苯磺酰基-Phe-氯甲基酮，N-a-甲苯磺酰基-Lys-氯甲基酮，抑肽酶(aprotinin)，苯甲基磺酰氟，二异丙基氟磷酸酯，激酶抑制剂，磷酸酶抑制剂，天冬氨酸蛋白水解酶(caspase)抑制剂，颗粒酶抑制剂，细胞粘附抑制剂，细胞分裂抑制剂，细胞周期抑制剂，脂质信号传导抑制剂，蛋白酶抑制剂，还原剂，烷化剂，抗微生物剂，氧化酶抑制剂或其他抑制剂。

[0268] 在某些情况下，药物制剂可以包括粘合剂作为赋形剂。合适的粘合剂的非限制性实施例可以包括淀粉，预糊化淀粉，明胶，聚乙烯吡咯烷酮，纤维素，甲基纤维素，羧甲基纤维素钠，乙基纤维素，聚丙烯酰胺，聚乙烯噁唑烷酮，聚乙烯醇，C12-C18脂肪酸醇，聚乙二醇，多元醇，糖，低聚糖，及其组合。

[0269] 可以用于药物制剂中的粘合剂可以选自：淀粉，例如马铃薯淀粉、玉米淀粉、小麦淀粉；糖，例如蔗糖、葡萄糖、右旋糖、乳糖、麦芽糊精；天然和合成的树胶；明胶；纤维素衍生物，例如微晶纤维素、羟丙基纤维素、羟乙基纤维素、羟丙基甲基纤维素、羧甲基纤维素、甲基纤维素、乙基纤维素；聚乙烯吡咯烷酮(聚维酮)；聚乙二醇(PEG)；蜡；碳酸钙；磷酸钙；醇，例如山梨糖醇、木糖醇、甘露糖醇和水，或它们的组合。

[0270] 在某些情况下，药物制剂可以包括润滑剂作为赋形剂。合适的润滑剂的非限制性实施例可以包括硬脂酸镁硬脂酸钙、硬脂酸锌、氢化植物油、氢化植物油(sterotex)、聚氧乙烯单硬脂酸酯、滑石粉、聚乙二醇、苯甲酸钠、月桂基硫酸钠、月桂基硫酸镁和轻质矿物油。可以在药物制剂中使用的润滑剂可以选自金属硬脂酸盐(例如硬脂酸镁、硬脂酸钙、硬脂酸铝)、脂肪酸酯(例如硬脂富马酸钠)、脂肪酸(例如硬脂酸)、脂族醇、甘油山嵛酸酯、矿物油、石蜡、氢化植物油、亮氨酸、聚乙二醇(PEG)、月桂基金属硫酸盐(例如月桂基硫酸钠、月桂基硫酸镁)、氯化钠、苯甲酸钠、乙酸钠和滑石粉，或它们的组合。

[0271] 在某些情况下，药物制剂可以包括分散增强剂作为赋形剂。合适的分散剂的非限制性实施例可以包括淀粉，海藻酸，聚乙烯吡咯烷酮，瓜尔胶，高岭土，膨润土，纯化的木纤维素，羟乙酸淀粉钠，同晶硅酸盐和微晶纤维素作为高HLB乳化剂表面活性剂。

[0272] 在某些情况下，药物制剂可以包括崩解剂作为赋形剂。在某些情况下，崩解剂可以是非泡腾崩解剂。合适的非泡腾崩解剂的非限制性实施例可包括淀粉例如玉米淀粉、马铃薯淀粉、其预胶化和改性淀粉，甜味剂，粘土例如膨润土，微晶纤维素，藻酸盐，羟乙酸淀粉钠，树胶例如琼脂、瓜尔豆、刺槐豆，梧桐树胶、果胶和黄芪胶。在某些情况下，崩解剂可以是泡腾崩解剂。合适的泡腾崩解剂的非限制性实施例可以包括与柠檬酸组合的碳酸氢钠，和与酒石酸组合的碳酸氢钠。

[0273] 在某些情况下，赋形剂可以包括调味剂。掺入外层中的调味剂可以选自合成调味油和调味芳香剂；天然油；来自植物、叶子、花和果实的提取物；和它们的组合。在某些情况下，调味剂可以选自由以下项组成的组：肉桂油；冬青油；薄荷油；三叶草油；干草油；茴香油；桉树；香草；柑橘油，例如柠檬油、橙油、葡萄和葡萄柚油；和，水果香精，包括苹果、桃、梨、草莓、覆盆子、樱桃、李子、菠萝和杏。

[0274] 在某些情况下，赋形剂可包含甜味剂。合适的甜味剂的非限制性实施例可以包括葡萄糖(玉米糖浆)、右旋糖、转化糖、果糖及其混合物(当不用作载剂时)；糖精及其各种盐，例如钠盐；二肽甜味剂，例如阿斯巴甜；二氢查尔酮类化合物，甘草甜素；甜叶菊(Stevia Rebaudiana)(甜菊糖)；蔗糖的氯代衍生物，例如三氯蔗糖；和，糖醇，例如山梨糖醇、甘露糖醇、几糖醇(sylitol)等。

[0275] 在一个方面，本公开，adRNA，帮助将内源性ADAR酶募集至靶标mRNA并实现位点特异性的A至G编辑。这对于基因治疗具有巨大的潜力，其中递送单一adRNA可以潜在地校正G至A点突变。它还可以实现靶标RNA编辑，而不需要必须过表达RNA编辑酶(如ADAR2)。本公开证明了该技术在体外和体内的适用性。本公开还证明了通过创建长双链RNA，可以在即使没有ADAR募集结构域的情况下也能募集内源性ADAR。此外，使用这些经改造的adRNA，可以在靶标区域的mRNA中创建多个A至G的编辑。在一个方面，该系统使用U6启动子(pol III)转录的adRNA，以及化学合成的adRNA，并且显示出有效的RNA编辑。因此，在一个方面，构建体进一步包括启动子，例如pol II启动子，以转录adRNA。从启动子(例如pol III启动子)的转录可以改善靶标RNA的编辑效率。从pol III启动子的Alu转录物被优先编辑。本文还提供了，来自作为内源性ADAR靶标的Alu重复序列结构的经改造的adRNA。

[0276] 本公开的构建体可用于将adRNA定位于特定的细胞区室。例如，对于核定位，可以使用adRNA-snRNA融合体。类似地，通过添加N-端线粒体靶向序列(MTS)，可以将adRNA定位于线粒体。因此，在一个方面，所述构建体进一步包括N-端线粒体靶向序列(MTS)。通过添加适当的顺式作用的拉链编码，可以将adRNA定位在过氧化物酶体、内体和外泌体中。因此，在另一方面，所述构建体进一步包括适当的顺式作用的拉链编码，用于将adRNA定位于过氧化物酶体、内体和外泌体中。在神经元的情况下，将adRNA定位到内体中可以潜在地使它们能够长距离运输。定位到外泌体中可以潜在地有助于将adRNA传播到邻近细胞。将诸如胆固醇等的部分系链到adRNA可以帮助细胞摄取。因此，在一个方面，所述构建体进一步包括靶向部分，例如胆固醇。

[0277] 在一个方面，本公开证明，为了创建小分子可调控的adRNA，公开了在一方面中使用adRNA-适体，以实现RNA编辑的时间控制，例如结合黄素单核苷酸、鸟嘌呤和其他天然代谢物的适体。结合糖的适体也可以用于这一目的。

[0278] 在一个方面，创建U1A-ADAR融合体完全是人类来源的。剪接体U1A蛋白基序的N-末端RNA识别，以63nM的解离常数结合至其同源的U1发夹II RNA。

[0279] 本公开还提供了进一步包括立足点形的构建体。

[0280] 在一个方面，本公开的构建体可以在没有ADAR酶过表达的情况下使用。

[0281] 因此，在某些方面，本公开的adRNA具有某些组分：长度为约15至约200个碱基对的RNA靶向结构域，其与靶标RNA互补；0～10个ADAR募集结构域，其可以来源于GluR2 mRNA、Alu重复序列元件或ADAR结合的其它RNA基序；和，将ADAR导向靶标腺苷所需的胞嘧啶错配，

该靶标腺苷可能存在于靶向结构域中的任何位置。当此adRNA与其靶标RNA结合时，它将ADAR酶募集至靶标RNA。现在，该ADAR酶可以催化靶标腺苷向肌苷的转化。对于长度超过50个碱基对的adRNA，当在HEK 293T和HeLa中表达时，adRNA细胞在即使没有ADAR募集结构域的情况下也可以募集ADAR，能实现显著水平的靶标RNA编辑。也可以使用单一的adRNA，在靶标mRNA中创建多个A至G编辑。另外，通过利用多种adRNA，可以在同一细胞中编辑多个不同的adRNA。例如，在杜氏肌肉萎缩症的mdxmdx小鼠模型中，不仅可以校正抗肌萎缩蛋白中的突变，而且还可以破坏编码参与无义介导的降解的蛋白的基因的mRNA序列。另一个应用是使用该技术来创建功能丧失、功能获得和显性负突变，并且一方面可以用于癌症筛查、肿瘤进展以及免疫编辑研究。

[0282] 经改造的adRNA

[0283] 本文提供了经改造的ADAR1或ADAR2向导RNA(“adRNA”)，其包括以下项，或可选地基本上由以下项组成，或还进一步由以下项组成：与靶标RNA互补的序列。在一个特定方面，本公开的经改造的adRNA进一步包括与鸟氨酸转氨甲酰酶互补的序列，或可选地基本上由其组成，或还进一步由其组成。

[0284] 在一个方面，本公开的经改造的adRNA进一步包括来源于GluR2 mRNA的ADAR2募集结构域，或可选地基本上由其组成，或还进一步由其组成。在另一方面，本公开的经改造的adRNA进一步包括来源于Alu重复序列的ADAR1募集结构域，或可选地基本上由其组成，或还进一步由其组成。在又一方面，本公开的经改造的adRNA进一步包括在与靶标RNA互补的序列两侧的两个MS2发夹，或可选地基本上由其组成，或还进一步由其组成。在一些实施方式中，与本公开的经改造的adRNA中与靶标RNA互补的序列包括以下项，或可选地基本上由以下项组成，或还进一步由以下项组成：约15至30个碱基对，或约30至45个碱基对，或约45至60个碱基对，或约60至75个碱基对，或约75至90个碱基对，或约90至105个碱基对，或约105至120个碱基对，或约120至135个碱基对，或约135至150个碱基对，或约150至165个碱基对，或约165至180个碱基对，或约180至200个碱基对。在另一方面，它是约40至约200，或约50至约200，或约60至约200，或约70至约200，或约80至约200，或约90至约200，或约100至约200个碱基对。

[0285] 本文公开了一种经改造的adRNA，其包括以下项，或可选地基本上由以下项组成，或还进一步由以下项组成：无ADAR募集结构域，或约1-2个ADAR募集结构域，或约2-3ADAR募集结构域，或约3-4个ADAR募集结构域，或约4-5个ADAR募集结构域，或约5-6个ADAR募集结构域，或约6-7个ADAR募集结构域，或约7-8个ADAR募集结构域，或约8-9个ADAR募集结构域，或约9-10个ADAR募集结构域。在一些实施方案中，ADAR募集结构域包括以下项，或可选地基本上由以下项组成，或还进一步由以下项组成：GluR2 mRNA，Alu重复序列元件，或ADAR结合的其它RNA基序。另外，本文提供了经改造的adRNA，其中来源于GluR2 mRNA的经改造的adRNA的ADAR2募集结构域位于经改造的adRNA的5’端或3’端。在一些实施方案中，GluR2 mRNA位于经改造的adRNA的5’端和3’端。

[0286] 在一个方面，本公开的经改造的adRNA进一步包括编辑诱导元件，或可选地基本上由其组成，或还进一步由其组成。“编辑诱导元件”可以指对于有效的RNA编辑所必需的、主要是双链RNA的结构。编辑诱导元件的非限制性实施例描述于Daniel,C.et al.(2017) Genome Biol.18,195中。

[0287] 在一个特定的方面，本公开的经改造的adRNA由选自表1或图2中提供的一组序列的多核苷酸序列或其各自的等同物编码。

[0288] 本文还公开了一种复合物，其包括在高严格条件下与互补多核苷酸杂交的本公开的经改造的adRNA，或可选地基本上由其组成，或还进一步由其组成。

[0289] 本公开还提供了用于下文描述的基因和蛋白编辑技术的多肽和/或多核苷酸序列。应该理解，尽管并非总是明确指出，本文提供的序列可用于提供表达产物，以及产生具有相同生物学特性的蛋白的基本相同的序列。这些“生物等同”或“生物活性”多肽由本文所述的等同多核苷酸编码。当使用在默认条件下运行的序列同一性方法进行比较时，它们可以具有与参考多肽至少60％，或可选地至少65％，或可选地至少70％，或可选地至少75％，或可选地至少80％，或可选地至少85％，或可选地至少90％，或可选地至少95％，或可选地至少98％相同的一级氨基酸序列。提供了特定的多肽序列作为特定实施方式的实施例。用具有相似电荷的备用氨基酸对序列的氨基酸进行修饰。另外，等同多核苷酸是在严格条件下与参考多核苷酸或其互补序列杂交的多核苷酸，或者关于多肽，指由在严格条件下与编码参考物的多核苷酸或其互补链杂交的多核苷酸编码的多肽。或者，等同多肽或蛋白是由等同多核苷酸表达的多肽或蛋白。

[0290] 本文还公开了一种经改造的adRNA-snRNA(小核RNA)融合体。在一个方面，所述经改造的adRNA进一步包括N-端线粒体靶向序列(MTS)，或可选地基本上由其组成，或还进一步由其组成，以促进所述经改造的adRNA定位至线粒体。在另一方面，本文提供的是经改造的进一步包括顺式作用的拉链编码(zipcode)，或可选地基本上由其组成，或进一步由其组成，以促进所述经改造的adRNA定位到过氧物酶体、内体和外泌体中。在神经元的情况下，将adRNA定位到内体中可以潜在地使它们能够长距离运输。定位到外泌体中可以潜在地有助于将adRNA传播到邻近细胞。将诸如胆固醇等的部分系链到adRNA可以帮助细胞摄取。

[0291] 本文进一步提供了小分子可调控的经改造的adRNA。在一个方面，本文公开了经改造的adRNA-适体的融合体。可用于该目的的适体的非限制性实施例包括结合黄素单核苷酸、鸟嘌呤、其他天然的代谢物或糖的适体。“适体”可以指，能够以高亲和力和特异性结合各种分子的短的单链寡核苷酸。适体的非限制性实施例描述于Lakhin,A.V.et al.(2013).Acta naturae,5(4),34-43中。

[0292] 本文还公开了完全源自人类的U1A-ADAR融合体。剪接体U1A蛋白基序的N-末端RNA识别，以63nM的解离常数结合至其同源的U1发夹II RNA。

[0293] 表达经改造的adRNA的载体和重组细胞

[0294] 本文还提供了一种载体，其包括一种或多种分离的多核苷酸序列和任选地与所述分离的多核苷酸可操作地连接的调控序列，或可选地基本上由其组成，或进一步由其组成，其中所述分离的多核苷酸序列编码本公开的经改造的adRNA。载体的非限制性实施例包括质粒或病毒载体，例如逆转录病毒载体、慢病毒载体、腺病毒载体或腺相关病毒载体。所述载体可以进一步包括本领域已知的靶向序列、拉链编码或立足点形。

[0295] 在一个方面，调控序列包括启动子、增强子元件和/或报告因子，或可选地基本上由其组成，或还进一步由其组成。在一些实施方案中，所述启动子是人U6、小鼠U6启动子、CMV启动子或pol III启动子或pol II启动子。在一个方面，所述载体进一步包括可检测的标记物或纯化标记物，或可选地基本上由其组成，或还进一步由其组成。

[0296] 本文进一步公开了一种重组细胞，其进一步包括上述载体，或可选地基本上由其组成，或还进一步由其组成，其中所述经改造的adRNA被重组表达。

[0297] 经改造的adRNA的组合物

[0298] 本文公开了一种组合物，其包括以下项，或可选地基本上由以下项组成，或还进一步由以下项组成：载剂和本公开的一种或多种经改造的adRNA，编码本公开的经改造的adRNA的分离的多核苷酸，表达本公开的经改造的adRNA的载体，或表达本公开的经改造的adRNA的重组细胞。在一个方面，所述载剂是药学上可接受的载剂或固体支持物。在又一方面，所述组合物进一步包括化学治疗剂或药物，或可选地基本上由其组成，或还进一步由其组成。

[0299] 使用经改造的adRNA的方法

[0300] 本文提供了一种修饰蛋白表达的方法，该方法包括使编码该蛋白的多核苷酸与本公开的经改造的adRNA接触，或可选地基本上由其组成，或还进一步由其组成，其中该蛋白的表达将被修饰。

[0301] 本文还提供了一种治疗与异常蛋白表达有关的疾病或障碍的方法，该方法包括向需要这种治疗的受试者施用有效量的一种或多种本公开的经改造的adRNA，或可选地基本上由其组成，或还进一步由其组成。在一个特定的方面，本文提供了一种治疗杜氏肌肉萎缩症的方法，其包括向需要这种治疗的受试者施用有效量的一种或多种本公开的经改造的adRNA，或可选地基本上由其组成，或还进一步由其组成。

[0302] 在体外应用的情况下，在一些实施方案中，有效量可以取决于所讨论的应用的大小和性质。它还可取决于体外靶标的性质和敏感性，以及所使用的方法。技术人员可以基于这些和其他考虑因素来确定有效量。根据实施方案，有效量可以包括组合物的一次或多次施用。

[0303] 术语“受试者”、“宿主”、“个体”和“患者”在本文中可互换使用，是指动物，通常是哺乳动物。任何合适的哺乳动物可以通过本文所述的方法、细胞或组合物进行治疗。哺乳动物的非限制性实施例包括人，非人灵长类动物(例如猿、长臂猿、黑猩猩，红毛猩猩，猴，猕猴等)，家畜(例如狗和猫)，农场动物(例如马、牛、山羊、绵羊、猪)，和实验动物(例如小鼠、大鼠、兔、豚鼠)。在一些实施方案中，哺乳动物是人。哺乳动物可以是任何年龄或处于任何发育阶段(例如，成年、青少年、儿童、婴儿或子宫内的哺乳动物)。哺乳动物可以是雄性或雌性。哺乳动物可以是怀孕的雌性。在一些实施方案中，受试者是人。在一些实施方案中，受试者患有或怀疑患有癌症或赘生性障碍。在其他实施方案中，受试者患有或怀疑患有与异常蛋白表达相关的疾病或障碍。

[0304] 参考图13，与可以包括短ssRNA(例如长度约35nt)的反义寡核苷酸(AON)设计(图13的左侧)相比，本公开的示例性构建体(图13的右侧)可以是长度约60bp至约100bp、具有优异的靶标特异性的长ssRNA，总序列长度为约150nt至约250nt。本文所述的构建体可以包括经过优化且真正的发夹结构——与可添加以创建发夹样RNA凸起的错配碱基相对，如图13左侧示出的AON设计所示。与AON构建体相比，构建体的真正的发夹设计的优势包括最佳和优异的ADAR募集效率，得到了更高的靶标编辑产量。发夹设计可以完全独立于mRNA靶序列，并且可以轻松地用于任何新的mRNA靶标。脱氨基的靶位点可以是独特、精确的，不需要进行不希望的化学修饰，也没有其它位置发生不希望的脱氨基的风险。相比之下，利用错配

发夹样RNA凸起(如图13左侧所示)的AON设计可以是:(a)募集ADAR的功效有限(不是真正的发夹,结构可能太短);(b)防止重复使用序列和凸起(需要为每个新靶标mRNA进行独特的设计);(c)由于错配而导致对细胞mRNA靶标位点的特异性降低(靶标损伤的风险增加),(d)在凸起中需要2'OMe修饰,以保护其它腺嘌呤免受ADAR脱氨作用(可能在错误的核糖核苷酸处引起不希望的突变)。

[0305] 经由长反义adRNA进行外源性和内源性ADAR的募集

[0306] CRISPR/Cas9系统广泛用于研究中,但仍存在对其体内应用的担忧。两个主要担忧的问题是对基因组的永久编辑,以及可能由于引入细菌来源的系统而引起的免疫应答。结果,RNA编辑作为两种挑战的潜在解决方案已经引起了人们的兴趣。通过应用RNA编辑可以容易地克服第一个挑战;从基因转录的RNA可以是瞬时的,编辑将不会永久改变细胞。但是,因为可能对许多RNA进行编辑以实现表型效果,因此这可能产生降低编辑效率的新的问题。另外,不一定能够克服免疫应答的第二个挑战。编辑RNA的ADAR家族的酶是人类来源的。特别是ADAR1在许多细胞类型中几乎普遍表达;通过利用其对RNA A-至-I(A-至-G)编辑的自然偏好性,可以有许多潜在的治疗应用。这一策略还克服了递送的重大障碍,因为小向导RNA的运输比大体积酶的运输容易得多。因此,本文所述的方法和组合物的目标可以包括改造向导(称为adRNA),以募集内源性ADAR。ADAR可能偏好编辑双链RNA的区域,特别是在具有二级结构的序列附近。因此,对反义结构域长度增加的向导进行改造可以是有利的。adRNA可以具有三个部分:(1)RNA靶向结构域(例如长度约15个碱基对至约200个碱基对),其可以与靶标RNA互补;(2)ADAR结合的约0至约10个ADAR募集结构域(其可以来源于GluR2 mRNA、Alu重复元件或其它RNA基序);和,(3)胞嘧啶错配,其对于将ADAR导向靶标腺苷可以是必需的,其中靶标腺苷可以存在于靶向结构域的任何位置。

[0307] 当该adRNA与其靶标RNA结合时,它可以将ADAR酶募集至靶标RNA。该ADAR酶可以催化靶标腺苷转化至肌苷。有趣的是,当在HEK 293T和HeLa细胞中表达时,长度为50个碱基对以上的长adRNA,甚至可以在没有ADAR募集结构域的情况下募集ADAR,并且可以实现显著水平的靶标RNA编辑。单一的adRNA也可用于在靶标mRNA中创建多个A-至-G编辑。另外,通过利用多种adRNA,可以在同一细胞中编辑多个不同的mRNA。例如,在杜氏肌肉萎缩症的mdx小鼠模型中,不仅可以校正抗肌萎缩蛋白中的突变,而且还可以破坏编码参与无义介导的降解的蛋白的基因的mRNA序列。另一个应用可以是使用该技术来创建功能丧失、功能获得、显性负突变或其任意组合。这在癌症筛查(肿瘤进展以及免疫编辑研究)中至关重要。

[0308] 参考图16,示意图示出了在adRNA存在的情况下,经由募集内源性ADAR来进行RNA编辑。这些adRNA可以作为经化学修饰的RNA或作为U6转录的RNA来提供。参考图17,评估了U6启动子转录的反义结构域长度逐渐增加的adRNA,与0、1或2个GluR2结构域组合,它们在有或没有外源ADAR2表达的情况下诱导靶向RNA编辑的能力。数值代表平均值+/-SEM(n=3)。长adRNA可以募集内源性ADAR进行RNA编辑。参考图18,化学合成的adRNA型式在有或没有外源性ADAR2表达的情况下,针对一组mRNA进行了测试。在图中说明了精确的化学修饰以及adRNA的来源。数值代表平均值+/-SEM(n=3)。参考图19,经处理的成年spf^{ash}小鼠的肝脏(眼眶后注射),正确剪接的OTC mRNA的体内RNA校正效率。在注射了U6转录的短adRNA的小鼠中,RNA编辑水平为0.6%。

[0309] 经由Alu-adRNA募集外源性和内源性ADAR

[0310] Alu基因是基因组中的转座元件,是通过ADAR进行RNA编辑的天然靶标。可以基于Alu元件结构来设计adRNA,以能够通过内源性募集的ADAR进行编辑。可以测试Alu元件结构上的各个位置,其中天然序列可以被替换成与靶标互补的反义序列。可以选择单链接头区域,例如如图20所示的。可以优化反义向导的长度,使长度从约20个碱基改变至约100个碱基,靶向RAB7A基因座。每一向导都可以设计为包括一个或多个错配。错配可以位于反义序列的胞嘧啶和待编辑的靶标腺苷碱基之间。该错配可以位于每一反义序列长度的中间。编辑效率可以与,可募集ADAR1和ADAR2二者的100-50反义向导,和仅可募集ADAR2的GluR2 20-6向导进行比较。

[0311] 参考图20,示出了Alu adRNA的设计。左:Alu元件的结构。中:并入基因座特异性反义序列的设计,其中具有与靶标A相对的C错配。右:RNA编辑酶ADAR至靶标的募集。

[0312] 通过用lipofectamine转染,在293FT细胞中测试了这些向导。每一向导在过表达ADAR1p110、ADAR1p150、ADAR2的细胞中,或无酶过表达的细胞中进行测试,以测试adRNA募集内源性ADAR的能力。在转染后48小时后,收获细胞,提取RNA,转化为cDNA,并扩增RAB7A基因座用于Sanger测序。编辑效率以峰高之比来计算。图5示出了Alu向导长度实验的结果。

[0313] 参考图5,长Alu-v2-100-50向导示出了,在没有过表达ADAR酶的细胞中的得到改进的编辑。ADAR1 p150的过表达导致Alu构建体的编辑率显著较高,而线性100-50向导的编辑率相似。与较短的同种型ADAR1 p110缺乏相比,由于其额外的ds-RNA结合结构域,ADAR1 p150可以优先结合Z-RNA。已知其具有高GC含量的Alu元件会形成Z-DNA和Z-RNA,有助于募集ADAR1 p150。

[0314] 拆分式ADAR2脱氨酶结构域(DD)

[0315] ADAR的过表达可能导致多个转录组范围的脱靶编辑。将ADAR2 DD的催化活性仅限制在靶标mRNA上的能力,可以降低脱靶的数量。创建拆分式ADAR2 DD可能是降低脱靶数量的一种潜在方法。拆分式蛋白重组装或蛋白片段互补可以是研究蛋白-蛋白相互作用的广泛使用的方法。可以设计拆分ADAR2 DD,以使得拆分式ADAR2 DD的每个片段本身都可以在催化上是没有活性的。但是,在存在adRNA的情况下,拆分的两半可以二聚化,以形成对预期的mRNA靶标具有催化活性的酶。

[0316] 可以通过研究ADAR2 DD与其天然存在的底物的复合物的晶体结构,推断溶剂可及性分数,使用预测软件或它们的任何组合,来鉴定用于拆分蛋白的区域。MS2-MCP系统和盒(box)B-λN系统(其可以有效地募集ADAR)可以单独使用或组合使用,来分别募集拆分的ADAR2 DD的N-和C-末端。adRNA可以包括一个MS2茎环和一个boxB发夹,以及与靶标互补的反义结构域。这可以使得在靶标处募集拆分的ADAR2 DD的N-和C-末端,从而可以构成催化活性的DD。

[0317] 参考图22和图23,示出了拆分式ADAR2 DD系统的示意图,并且突出显示了具有用于拆分的位点的ADAR2 DD的示例性序列。

[0318] 参考图24,可以经由海荧(cypridina)荧光素酶报告因子(cluc W85X),来测定成对的片段1～16。片段9和片段10示出了最高的活性。对应于片段9和片段10的拆分位置在图23中用蓝色圈出。

[0319] 参考图25,针对Cluc报道因子测定了片段9和片段10。此外,还可以测试NES-MCP-AD2-C-U变体的A-G编辑。N:N-末端片段,C:C-末端片段,M-M:MS2-MS2adRNA,M-B:MS2-BoxB

adRNA,B-B:BoxB-BoxB adRNA。

[0320] 可以通过利用人RNA结合蛋白,例如(a)U1A,或(b)其进化变体TBP6.7(其没有已知的内源性人发夹靶标),或(c)人组蛋白茎环结合蛋白(SLBP),或(d)糖皮质激素受体的DNA结合结构域,或(e)它们的任意组合,来创建这些构建体的进一步完全人源化的形式。这些蛋白可以融合到ADAR2的N-和C-末端片段,以创建完全人的可编程RNA编辑工具集,其可以以极高的特异性来编辑腺苷。此外,带有两个相应的RNA发夹的嵌合RNA,可用于募集ADAR2片段。本文提供了RNA发夹的序列。

[0321] 可以通过在ADAR2脱氨酶结构域中进行一个或多个突变(例如16个突变),来创建C-U RNA编辑酶。即使具有一个或多个突变,该编辑酶仍可以显示几个转录组范围的A-G和C¬U脱靶。为了改进这种酶的特异性,可以在某些残基处对它进行拆分(例如先前筛选中鉴定出的那些残基),从而开发出用于C-U编辑的拆分式ADAR系统。

[0322] 用于C-U编辑的外源性和内源性APOBEC的募集

[0323] APOBEC(载脂蛋白B mRNA编辑酶,催化性多肽样)是RNA编辑酶,可将胞苷转化为尿嘧啶,并在mRNA水平上创建多样性。可以募集这些实体用于通过向导RNA(例如经改造的募集APOBEC的向导RNA(apRNA))进行RNA的C-U编辑。在某些情况下,可以创建包括一个或多个APOBEC家族成员的融合构建体。例如,融合构建体可以包括针对MCP(MS2外壳蛋白)的ADAT1(腺苷脱氨酶TRNA特异性1)和AID(激活诱导的胞苷脱氨酶)。经改造的MS2-apRNA可用于募集一种或多种MCP融合体。用于该构建体的蛋白序列以及MS2-apRNA序列示于图28～图30中。用于融合构建体的蛋白序列,或在本文所述的任何方法中使用的序列可包括,与图28的任何序列的至少一部分具有至少约:70%,75%,80%,85%,90%,95%,97%,98%、99%或更高的序列同一性。

[0324] 在某些情况下,方法可以包括被配置用于募集APOBEC(例如APOBEC3A)的构建体。募集可以是内源性募集。可以通过一级序列、二级结构或其组合的靶向偏好性,来设计被配置用于募集APOBEC的构建体。一种或多种apRNA的设计可以包括如图29所示的那些序列或其一部分。在某些情况下,可以募集APOBEC(例如APOBEC3A)的序列,或在本文所述的任何方法中使用的序列,可以包括与图29的任何序列的至少一部分具有至少约:70%、75%、80%、85%、90%、95%、97%、98%、99%或更高的序列同一性。

[0325] 为了募集MCP-APOBEC3A,可以设计MS2-apRNA,并且它们的序列可以包括图30的任一种或多种序列。在某些情况下,可以募集MCP-APOBEC3A的序列,或在本文所述的任何方法中使用的序列,可以包括与图30的任何序列的至少一部分具有至少约:70%、75%、80%、85%、90%、95%、97%、98%、99%或更高的序列同一性。

[0326] 荧光素酶测定

[0327] 当ADAR将TAG中的靶标A转化为I(读作G)时,核糖体能够完全翻译蛋白,如图31的示意图所示。可以转染细胞,并且可以在48小时后使用荧光素酶报告基因测定法对其进行可视化。从细胞读出的光可以指示荧光素酶活性的恢复。

[0328] 参考图31,第一个方案证明了一实施例,其中核糖体可以到达荧光素酶基因中的提前TAG终止密码子,并且可以终止翻译,产生截短的无功能的荧光素酶。在第二个方案中,可以通过adRNA将ADAR募集到一位点,在该位点处,它可以将TAG终止密码子编辑成色氨酸的TGG密码子,这可以允许核糖体通读,可以产生正常的荧光素酶表达。这样的系统可以允

许评估通过adRNA进行的适当的ADAR募集。当检测到荧光素酶表达时，可能已经发生了通过adRNA的ADAR募集。当未检测到荧光素酶表达时，可能未发生ADAR募集。

[0329] 试剂盒

[0330] 本文还公开了一种试剂盒，其包括以下项，或可选地基本上以下项组成，或还进一步由以下项组成：本公开的经改造的adRNA，编码本公开的经改造的adRNA的分离的多核苷酸，表达本公开的经改造的adRNA的载体，表达本公开的经改造的adRNA的重组细胞，或本文公开的组合物，和使用说明书。在一个方面，该说明书记载了使用本文公开的经改造的adRNA的方法。

[0331] 试剂盒可包括载体。该载体可以包装在容器中。该试剂盒可以包括非天然存在的RNA。该非天然存在的RNA可以包装在容器中。该试剂盒可以包括注射器。注射器可以是载体、核酸或非天然存在的RNA可以包装在其中的容器。该试剂盒可以包括本文所述的药物组合物。该试剂盒可以包括，用于向有此需要的受试者施用如本文所述的病毒载体、非天然存在的RNA、药物组合物的说明书。

[0332] 实施例

[0333] 以下实施例是非限制性的，并且举例说明可以在使本公开内容生效的各种情况下使用的程序。另外，本文公开的所有参考文献通过引用整体并入本文中。

[0334] 实施例1

[0335] 在不受理论约束的情况下，据信可以募集在哺乳动物中发现的ADAR，以催化点突变的治疗性编辑。经由使用经改造的RNA支架、经改造的DNA支架或DNA-RNA杂交支架，可以将ADAR1或ADAR2募集至靶标RNA或潜在的DNA。可以被这种方法潜在地靶向的组织包括，但不限于，中枢和周围神经系统、肺、肝脏、胃肠道、胰腺、心肌、肾脏和皮肤。

[0336] 本文提出的一个示例性实施方式是经改造的ADAR2向导RNA(adRNA)，其携带与靶标RNA的20～100bp互补性。该经改造的adRNA在5'端或3'端，或在两端，含有来自GluR2 mRNA的ADAR2募集结构域。

[0337] 在鸟氨酸转氨甲酰酶缺乏症的spf^{ash}小鼠模型中对其进行了体内测试。该模型在外显子4的最后一个核苷酸中带有G->A点突变。经由AAV仅递送adRNA后，在不存在ADAR酶过表达的情况下，观察到点突变的高达1％的校正。本公开还示出了，在体外已知表达ADAR的HeLa细胞中的这一功效。仅递送adRNA后，在这些细胞中观察到RNA编辑。

[0338] 此外，该功效可以应用于其他募集结构域。因此，进一步的方面涉及经改造的单链ADAR2向导DNA(adDNA)以及adDNA-RNA杂交体，其具有比adRNA更大的潜在稳定性。还提供了使用本文公开的这些酶的方法。

[0339] 在不受理论约束的情况下，由于ADAR家族的酶催化腺苷水解脱氨为肌苷，因此可以认为这些酶可通过突变ADAR2的三个特定残基，与靶标腺苷相互作用的V351、E396和C451，来用于催化胞嘧啶水解脱氨为胸腺嘧啶。因此，通过这些位点的突变，提供了提供这种催化潜力的方法。

[0340] 创建了在上述三个残基处的所有可能的氨基酸取代，并进行筛选以检验该假设。

[0341] 此外，进行实验以探索其他氨基酸(诸如S486)的作用，这些氨基酸可能能够消除ADAR对UAG编辑位点的固有偏好。

[0342] 为了改造C->T编辑，确定了hAPOBEC1和rAPOBEC1以及Apobec1互补因子(ACF)过表

达的作用。另外，本文进一步提供了MCP-(h/r)APOBEC1-ACF融合蛋白。

[0343] 还提供了目前针对Cas9(或Cpf1)-ADAR脱氨酶结构域融合体(ADAR1、ADAR2及其催化活性突变体E1008Q和E488Q)的腺嘌呤碱基编辑方法，本文进一步提供了通过目前碱基编辑器靶向ssDNA置换链的组合物。为此，将具有A-C凸起(理想地是在接近gRNA的5'端的第一个10bp中)的gRNA结合链作为靶标。

[0344] 附加文件中例示了另外的实施方式，通过引入并入本文中。

[0345] 实施例2

[0346] 参考图6，Alu元件可以成为基于内源性ADAR的RNA编辑的主要靶标。因此，出于可编程的RNA编辑的目的，目标是设计基于Alu元件的ADAR募集RNA(adRNA)，因为它们可以潜在地有效募集内源性ADAR。Alu-adRNA由人U6启动子表达，Alu重复序列之间的接头序列被替换成多种长度的反义结构域，靶向RAB7A转录物。每个反义结构域均在反义区域的中间有一个错配的核苷酸，靶标腺苷对面的胞嘧啶。

[0347] 通过转染293FT细胞，连同ADAR1p110、ADAR1p150、ADAR2一起，或在没有过表达的酶的情况下，在体外测试Alu-adRNA，以证明内源性ADAR的募集。转染后48小时收获细胞，提取RNA，并使用随机六聚体或寡核苷酸(oligo)-dT引物，将其转化为cDNA。然后扩增RAB7A基因座，并发送至桑格(Sanger)测序。编辑效率计算为Sanger峰高G/(A+G)的比值。

[0348] 实施例3

[0349] 载体设计和构建

[0350] 将0、1或2个拷贝的GluR2 adRNA克隆到AAV载体中，其中该AAV载体包含人U6和小鼠U6启动子，以及CMV启动子，驱动GFP或全长人ADAR2酶或其超高活性突变体ADAR2(E488Q)的表达。类似地，将1或2个拷贝的MS2 adRNA克隆到AAV载体中，其中该AAV载体带有MCP-ADAR1或MCP-ADAR2脱氨酶结构域融合体及其高活性突变体。为了构建GFP报告因子，GFP-Amber、GFP-Ochre和GFP-Opal，合成了三个基因块，分别用'TAG'、'TAA'和'TGA'取代了野生型GFP的Y39残基，并克隆到CAG启动子的下游。为了构建OTC和DMD报告因子，将带有要编辑的靶标腺苷的spf^{ash} OTC和mdx DMD转录物的200bp片段，克隆到CAG启动子的下游。

[0351] 哺乳动物细胞的培养和转染

[0352] 所有的HEK 293T细胞均在37℃和5％CO2气氛的培养箱中，在补充了10％FBS和1％抗生素-抗真菌药(Thermo Fisher)的杜氏改良伊格尔培养基(Dulbecco's Modified Eagle Medium)中生长。所有体外转染实验均使用商业转染试剂Lipofectamine 2000(Thermo Fisher)，在HEK 293T细胞中进行。所有涉及报告因子的体外RNA编辑实验均使用400ng报告因子质粒和800ng adRNA+酶质粒，在24孔板中进行。所有靶向内源性转录物的体外RNA编辑实验均使用800ng adRNA/酶质粒，在24孔板中进行。基于dCas13b-ADAR2DDE488Q的RNA编辑实验使用800ng的酶质粒(Addgene#103864)和800ng的gRNA质粒进行。细胞在25～30％汇合度时进行转染，并在转染后60小时收获，用于对编辑进行定量。化学合成的adRNA(通过IDT或Synthego合成)使用Lipofectamine 3000(Thermo Fisher)，以20pmol/孔的量进行转染。

[0353] AAV载体的产生

[0354] 使用HEK 293T细胞，通过三重转染方法生产AAV8颗粒，并通过碘克沙醇梯度进行纯化。转染时的汇合度约为80％。转染前两个小时，将补充有10％FBS的DMEM添加到HEK

293T细胞中。每种病毒均在5x 15cm的平板中产生，其中每一平板均用7.5μg pXR-8、7.5μg重组转移载体、7.5μg pHelper载体，使用PEI(1μg/μL线性PEI在1xDPBS中，pH 4.5，使用HCl)，以4:1的PEI:DNA质量比，进行转染。将混合物在RT下孵育10分钟，然后滴加到细胞培养基上。72小时后收获病毒，并使用碘克沙醇密度梯度超速离心法进行纯化。然后，使用50kDA过滤器(Millipore)，用补充有50mM NaCl和0.0001％的Pluronic F68(Thermo Fisher)的1 x PBS(pH 7.2)透析病毒，至终体积～1mL，并根据标准(ATCC VR-1616)，使用针对ITR区域的特异性引物通过qPCR进行定量。

[0355] AAV-ITR-F:5'-CGGCCTCAGTGAGCGA-3'(SEQ ID NO:149)和

[0356] AAV-ITR-R:5'-GGAACCCCTAGTGATGGAGTT-3'(SEQ ID NO:150)。

[0357] 实施例4-通过RNA引导的腺苷脱氨酶在体内进行点突变的RNA编辑

[0358] 设计了一种系统，该系统通过作用于RNA的腺苷脱氨酶(ADAR)酶，与相关的ADAR向导RNA(adRNA)，来进行序列特异性的RNA碱基编辑。该系统经过系统性地改造以利用ADAR，并通过两种人类疾病的小鼠模型，全面评估了其在体外和体内的特异性和活性。在某些情况下，该平台可针对不同应用实现RNA的可调和可逆的改造。

[0359] 腺苷到肌苷的RNA编辑，转录后的RNA修饰，是由作用于RNA的腺苷脱氨酶(ADAR)酶催化的。肌苷是腺苷的一种脱氨基形式，其在生物化学上被识别为鸟嘌呤。最近，多项研究证明了ADAR介导的靶向RNA编辑。在这些基础上，对两个正交工具集进行了改造，用于在体外和体内进行序列特异性的可编程的RNA碱基编辑。具体而言，利用了一种系统，该系统经由ADAR1/2与相关的ADAR向导RNA(adRNA)，用于靶向RNA编辑(图32A)。adRNA部分地包括可编程的反义区域，该可编程的反义区域与靶标RNA序列互补，带有与靶标腺苷相对的错配胞苷。此外，它们在一种型式中，携带0、1或2个从天然存在的ADAR底物GluR2 pre-mRNA改造而成的ADAR-募集结构域(以下称为GluR2 adRNA)；在第二种型式中，携带2个位于反义区域两侧的MS2发夹(以下称为MS2 adRNA)。通过评估多种支架变体，包括基于G-C与A-U配对，添加编辑诱导元件，和反义结构域长度和失配位置修饰的诱变支架，对GluR2 adRNA进行了系统性优化，以增强外源性和/或内源性ADAR的募集(图8，图34A～C)。经由系统性反义结构域长度和错配位置修饰，结合使用超高活性型式的脱氨酶结构域，和携带核定位(NLS)与输出(NES)信号的型式，反过来优化了后者MS2 adRNA型式，以利用包含与MS2外壳蛋白(MCP)融合的ADAR1或ADAR2的脱氨酶结构域(DD)的合成蛋白(图32B，图35A，41B)。

[0360] 使用最近开发的基于Cas13b的RNA编辑系统，对上述两种系统的活性进行了全面评估和基准测试。这些体外实验表明：(1)经改造的构建体能够有效地进行靶向RNA编辑，产量可与基于Cas13b的系统相当(图32B，图36A，表2，表3)，U6转录的adRNA和化学合成的adRNA都是有效的形式(图36B)；(2)带有长反义结构域的adRNA(带有或不带有GluR2结构域)足以募集外源表达的ADAR，在一定程度上也能募集内源性ADAR，以实现有效的RNA编辑(图32B，图34B，40C，42C)；(3)基于MS2 adRNA和相应的MCP-ADAR1/2融合体的构建体，示出了最高且最稳健的活性，包括一大组跨越一系列不同表达水平选择的内源性基因(图32B，图36C)；(4)使用在MCP-ADAR1/2融合体中的NES和/或超高活性脱氨酶结构域，一致地产生对靶标腺苷更高的RNA编辑产量，但也导致编辑侧翼序列中的非靶标腺苷的更高的倾向性(图32B，图37A)。为了进一步验证这一点，通过缺失ADAR2中的天然NLS结构域(Δ1-138)，产生了类似的杂泛性(promiscuity)(图37B～图37D)；和，5)这两个工具集可操作地正交：具

体地，评估MCP-ADAR2脱氨酶结构域融合体与共表达的MS2 adRNA或GluR2 adRNA的编辑效率，并显示出仅通过前者的上靶编辑。相反，观察到全长ADAR2被GluR2 adRNA而非MS2 adRNA募集(图35B)。

[0361] 在证明了该工具集的稳健活性之后，经由分析通过该系统实现的整个转录组范围的脱靶A->G编辑，来研究它的特异性概况(图32C)。为此，用每种构建体来转染HEK 293T细胞，并通过RNA-seq进行分析。使用未转染的细胞作为对照。从每个样品中，收集到约4000万条独特比对的测序读数。使用费希尔精确检验，来量化在每个具有足够读取覆盖率的参考腺苷位点处，相对于未转染细胞，A->G编辑产量的显著变化。计算在任何样品中检测到至少一次A->G编辑事件的位点的数量。在这些中，发现具有统计上显著的A->G编辑、1％的错误发现率(FDR)和至少1.1的倍数变化的位点的数量在很大范围内变化，从最低的MCP-ADAR2 DD-NLS构建体，到最高的MCP-ADAR1 DD(E1008Q)-NES(图38～图41，表4、5)。为了研究编辑产量的分布，考虑了其编辑产量在至少一个样品中发生显著变化的A-位点生成了小提琴图(图32)。综上，RNA-seq实验表明，整个转录组范围的脱靶编辑是：1)在具有NLS的MCP-ADAR构建体中的普遍性低于具有NES的构建体；2)在MCP-ADAR2构建体中的普遍性低于MCP-ADAR1构建体；3)在野生型MCP-ADAR构建体中的普遍性低于E>Q超高活性突变体(图42A，表5)；和，4)脱靶主要归因于ADAR过表达，而仅使用adRNA导致最少数量的脱靶(图42B)。

[0362] 表2：adRNA和gRNA反义序列的列表：

[0363]

名称	adRNA/gRNA反义序列（5'至3'）
mOTC†	ACAAACCGAGCGGTGTCTGT (SEQ ID NO:202)
mDMD†	GCCATTCCATTGCTCTTTCA (SEQ ID NO:203)
RAB7A (20, 6)	TGCCGCCAGCTGGATTTCCC (SEQ ID NO:204)
CCNB1 (20, 6)	CTGTACCAGCCAGTCAATTA (SEQ ID NO:205)
DAXX (20, 6)	CTTCTCCACAGCCCGAAGCA (SEQ ID NO:206)
CKDN2 (20, 6)	CTCCTCCACCCGACCCCGGG (SEQ ID NO:207)
GAPDH (20, 6)	GGGTGCCAAGCAGTTGGTGG (SEQ ID NO:208)
ALDOA (20, 6)	CTTGTCCACCTTGATGCCCA (SEQ ID NO:209)
ARHGAP8 (20, 6)	TTCATCCAATGGCTGGTTAT (SEQ ID NO:210)
CKB (20, 6)	CAAGGCCAAGGGCTCGCCAG (SEQ ID NO:211)
KRAS (20, 6)	TCCAACCACCACAAGTTTAT (SEQ ID NO:212)
Cas13b_RAB7A	TACAGAATACTGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGC (SEQ ID NO:213)
Cas13b_mOTC	GAAAAGTTTTACAAACCGAGCGGTGTCTGTGAGACTTTCATTCACACCCA (SEQ ID NO:214)
Cas9_mDMD_1	ATAATTTCTATTATATTACA (SEQ ID NO:215)
Cas9_mDMD_2	ATTTCAGGTAAGCCGAGGTT (SEQ ID NO:216)
RAB7A (20, 10)	ATACTGCCGCCAGCTGGATT (SEQ ID NO:217)
RAB7A (60, 6)	TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGG TTCAACC (SEQ ID NO:218)
RAB7A (60, 30)	TCTTGTGTCTACTGTACAGAATACTGCCGCCAGCTGGATTTCCCAATTCTGAGT AACACT (SEQ ID NO:219)
RAB7A (100, 6) (SEQ ID NO:220)	TGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGG TTCAACCCTCCACCTTACAGGCCTGCATTACAGGACTTAAACACATA
RAB7A (100, 50) (SEQ ID NO:221)	TGATAAAAGGCGTACATAATTCTTGTGTCTACTGTACAGAATACTGCCGCCAGC TGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACAGGGTTC
KRAS (100, 50) (SEQ ID NO:222)	TGAATTAGCTGTATCGTCAAGGCACTCTTGCCTACGCCACCAGCTCCAACCAC CACAAGTTTATATTCAGTCATTTTCAGCAGGCCTCTCTCCCGCACCT
CKB (100, 50) (SEQ ID NO:223)	ATCAAAAAAATAAACTCTACCAAGGGTGACGGAAGTCTCTACAGCAAGGCCAA GGGCTCGCCAGACGGCGAACATCAGGGGTGCATGGTGGGCACTGCCC

[0364] 表3:用于下一代测序(NGS)分析的引物的列表。

[0365]

名称	序列（5'至3'）
mDMD_NGS_F†	CTCTCTGTACCTTATCTTAGTGTTACTGA (SEQ ID NO:224)
mDMD_NGS_R†	ATTTCTGGCATATTTCTGAAGGTG (SEQ ID NO:225)
mOTC_NGS_F†	ACCCTTCCTTTCTTACCACACA (SEQ ID NO:226)
mOTC_剪接的_NGS_R†	CAGGGTGTCCAGATCTGATTGTT (SEQ ID NO:227)
mOTC_未剪接的_NGS_R†	CTTCTCTTTTAAACTAACCCATCAGAGTT (SEQ ID NO:228)
CCNB1_NGS_F	CAAGCAGTCAGACCAAAATACCTACTG (SEQ ID NO:229)
CCNB1_NGS_R	TCTTAGCATGCTTCGATGTGGCATA (SEQ ID NO:230)
DAXX_NGS_F	CATCAACAAGCCAGGGCCTG (SEQ ID NO:231)
DAXX_NGS_R	GAAGAGGAAATGTCCGTCTCCAC (SEQ ID NO:232)
RAB7A_NGS_F	AGGCCTGTAAGGTGGAGGG (SEQ ID NO:233)
RAB7A_NGS_R	(SEQ ID NO:234) TGAAATAACGGCAATTTATCCATTGCACATAC
CDKN2A_NGS_F	GGGAGCAGCATGGAGCCTT (SEQ ID NO:234)
CDKN2A_NGS_R	TCCGACCGTAACTATTCGGTGC (SEQ ID NO:236)
GAPDH_NGS_F	TGGGTGTGAACCATGAGAAGTAT (SEQ ID NO:237)
GAPDH_NGS_R	TGGCATGGACTGTGGTCATG (SEQ ID NO:238)
CKB_NGS_F	CCTAACTTATTGCCTGGGCAGTG (SEQ ID NO:239)
CKB_NGS_R	GCATCAGCAGTATCTTAGCCATCAA (SEQ ID NO:240)
NGS_KRAS_F	CAGAGGCTCAGCGGCTCC (SEQ ID NO:241)
NGS_KRAS_R	TAGCTGTATCGTCAAGGCACTC (SEQ ID NO:242)
ARHGAP8_NGS_F	CACACCTGTCTGTGCACTTGTA (SEQ ID NO:243)
ARHGAP8_NGS_R	CGGTCCACAGCTCAGGAACC (SEQ ID NO:244)
ALDOA_NGS_F	ACCAGAAGGCGGATGATGGG (SEQ ID NO:245)
ALDOA_NGS_R	CTCAGACAGCCCATCCAACC (SEQ ID NO:246)
KRAS_NGS_R2	TACTACTTGCTTCCTGTAGGAATCCTC (SEQ ID NO:247)
CKB_NGS_F2	AGCCCTGCTGCTTCCTAACTT (SEQ ID NO:248)
CKB_NGS_R2	ACCCTAGTTTATTTCAGCATCAGCAG (SEQ ID NO:249)

[0366] 表4:来自高通量测序实验的RNA-seq读数的记录。来自高通量测序实验的RNA-seq读数的记录。给定的计数表示配对末端测序的读数配偶(read mate),而不是读数对(mate pair)。列是:sn,样品名称;nt,解复用后的原始读数的总数目;nu,与参考基因组唯一对齐的成对读数的数目;nd,重复对中的读数的数目;nr,剩余读数的数目;df,下采样分数。名为“293T”、“293T L2”、“293T L8”和“293T L4”的样品均取自相同的对照库,但在Illumina仪器的不同泳道上进行了测序:

[0367]

sn	nt	nu	nd	nr	df
293TL1(0)	108434072	86792552	21376362	65416190	0.474401

293T(1)	107737550	86652062	21362310	65289752	0.47532
MCP-ADAR1 DD-NES-adRNA(2)	75469304	57715184	11907804	45807380	0.677479
MCP-ADAR1 DD-NES+adRNA(3)	76113978	55714058	11591434	44122624	0.703347
MCP-ADAR1 DD-NLS-adRNA(4)	96485146	79023222	17911404	61111818	0.507815
MCP-ADAR1 DD-NLS+adRNA(5)	70684382	56425658	8656076	47769582	0.64965
MCP-ADAR1 DD(E1008Q)-NES-adRNA(6)	73073334	54369950	10484274	43885676	0.707145
MCP-ADAR1 DD(E1008Q)-NES+adRNA(7)	95946852	71154174	18408668	52745506	0.588363
MCP-ADAR1 DD(E1008Q)-NLS-adRNA(8)	54654264	43441234	6314734	37126500	0.835886
MCP-ADAR1 DD(E1008Q)-NLS+adRNA(9)	78346400	59725272	11074654	48650618	0.637886
MCP-ADAR2 DD-NES-adRNA(10)	89534306	74166552	14629650	59536902	0.521249
MCP-ADAR2 DD-NES+adRNA(11)	80859886	66904932	12911706	53993226	0.574767
MCP-ADAR2 DD(E488Q)-NES-adRNA(12)	79769278	65570598	11792520	53778078	0.577066
MCP-ADAR2 DD(E488Q)-NES+adRNA(13)	98084994	80214200	20639602	59574598	0.520919
MCP-ADAR2 DD-NLS-adRNA(14)	75862216	60748320	14040780	46707540	0.664422
MCP-ADAR2 DD-NLS+adRNA(15)	80473694	66106146	12359830	53746316	0.577407
MCP-ADAR2 DD(E488Q)-NLS-adRNA(16)	48576248	40372732	6032488	34340244	0.903707
MCP-ADAR2 DD(E488Q)-NLS+adRNA(17)	72617598	58732422	12388120	46344302	0.66963
293T L8(18)	68191034	54705760	10254952	44450808	0.698154
293T+GFP(19)	88146982	64109834	15072908	49036926	0.63286
ADAR2-adRNA(20)	86641852	69462198	18251916	51210282	0.606002
ADAR2+adRNA(21)	74048950	59010378	16071512	42938866	0.722737
ADAR2(E488Q)-adRNA(22)	81927154	65842936	16273572	49569364	0.626063
ADAR2(E488Q)+adRNA(23)	74616248	56997066	17856714	39140352	0.792878
Cas13b-ADAR2 DD(E488Q)-gRNA(24)	72072754	54678392	12500074	42178318	0.73577
Cas13b-ADAR2 DD(E488Q)+gRNA(25)	116274658	91188024	29488244	61699780	0.502976
293T L4(26)	70234868	56894104	11595590	45298514	0.685089
MS2 adRNA(27)	78457766	51176442	19161354	32015088	0.969341
GluR2 adRNA(28)	65855024	41791326	10757802	31033524	1
gRNA(29)	89132098	59288978	19226242	40062736	0.774623
最小	48576248	40372732	6032488	31033524	0.474401
最大	116274658	91188024	29488244	65416190	1
总计	2588377968	2005961258	481864884	1524096374	21.68535

[0368] 表5:来自比对的RNA-seq读数的A->G编辑产量量化的结果。来自比对的RNA-seq读数的A->G编辑产量量化的结果。列是:样品名称;总位点,在处理样品与对照样品之间的至少一个比较中,A->G编辑产量有显著变化的参考位点的总数目;改变的位点,将处理样品与对照样品(其是表4中的第一个样品)进行比较时,发现A->G编辑产量具有显著变化的参考A-位点的数目;上靶编辑产量,在RAB7A mRNA内预期的目标A-位点上观察到的编辑产量;中值编辑产量,除目标位点外,在考虑到的所有其他位点上的中值产量:

[0369]

样品名称	总位点	变化的位点	上靶编辑产量	中值编辑产量
293T L2	382978	0	0	
293T+GFP	382978	6	0	0.295913155
gRNA	382978	32	0.013333333	0.142102619
Cas13b-ADAR2 DD(E488Q)-gRNA	382978	112853	0.025316456	0.105263158
Cas13b-ADAR2 DD(E488Q)+gRNA	382978	49432	0.11637931	0.092205807

GluR2 adRNA	382978	23	0	0.144542773
ADAR2-adRNA	382978	5769	0	0.157894737
ADAR2+adRNA	382978	18573	0.27638191	0.14
ADAR2(E488Q)-adRNA	382978	25732	0	0.131578947
ADAR2(E488Q)+adRNA	382978	125409	0.268398268	0.150943396
MS2 adRNA	382978	19	0.006666667	0.169230769
MCP-ADAR1 DD-NLS-adRNA	382978	20481	0.006849315	0.079069767
MCP-ADAR1 DD-NLS+adRNA	382978	28537	0.159763314	0.084745763
MCP-ADAR1 DD(E1008Q)-NLS-adRNA	382978	90182	0	0.112
MCP-ADAR1 DD(E1008Q)-NLS+adRNA	382978	110565	0.261627907	0.118081181
MCP-ADAR1 DD-NES-adRNA	382978	116165	0.017142857	0.097222222
MCP-ADAR1 DD-NES+adRNA	382978	101183	0.366459627	0.096618357
MCP-ADAR1 DD(E1008Q)-NES-adRNA	382978	226634	0.010416667	0.123076923
MCP-ADAR1 DD(E1008Q)-NES+adRNA	382978	195533	0.418604651	0.12244898
MCP-ADAR2 DD-NLS-adRNA	382978	3760	0.006756757	0.056173674
MCP-ADAR2 DD-NLS+adRNA	382978	4740	0.07	0.076555024
MCP-ADAR2 DD(E488Q)-NLS-adRNA	382978	28028	0.014778325	0.095238095
MCP-ADAR2 DD(E488Q)-NLS+adRNA	382978	38087	0.113122172	0.098591549
MCP-ADAR2 DD-NES-adRNA	382978	9489	0.021276596	0.09375
MCP-ADAR2 DD-NES+adRNA	382978	20249	0.416216216	0.102564103
MCP-ADAR2 DD(E488Q)-NES-adRNA	382978	35287	0.004672897	0.09929078
MCP-ADAR2 DD(E488Q)-NES+adRNA	382978	42715	0.278350515	0.101351351

[0370] 在进行了这些体外研究之后,利用了adRNA和外源性ADAR表达构建体形式,评估了该系统在基因治疗应用中的体内RNA靶向性,因为这些始终实现了最高的体外RNA编辑产量。首先评估了杜氏肌肉萎缩症(Duchenne Muscular Dystrophy,DMD)的mdx小鼠模型,该模型在抗肌萎缩蛋白基因的外显子23上带有ochre终止位点。这种选择还受到以下事实的启发:无义突变通常可能导致近11%的上述引起可遗传人类疾病的基因损伤,以及接近20%的影响基因编码区的与疾病相关的单碱基取代。因此,此处对RNA编辑策略的验证可以具有广泛的治疗应用。为此,首先在体外优化了终止密码子的RNA编辑(图43)。值得注意的是,观察到添加第二拷贝的adRNA显著提高了靶向效率(图43C),因此在体内研究中采用了双adRNA递送方式。然后将构建体包装到AAV8中,并将2E+12载体基因组(vg)/肌肉注射到mdx小鼠的胫前肌(TA)或腓肠肌中。为了进一步验证该方法,经由对外显子23进行基于CRISPR-Cas9的切除(图33A),同时打靶了mdx小鼠。注射后四周或八周,从mdx小鼠、野生型小鼠、经adRNA靶向和非靶向对照以及CRISPR-Cas9处理的小鼠中,收集了TA和腓肠肌的肌肉。免疫荧光染色显示,经由靶向RNA编辑,可以明显地恢复抗肌萎缩蛋白的表达(图33B,图44A)。另外,在肌膜处也恢复了nNOS活性(图33B,图44A)。在经处理的小鼠中,观察到了RNA编辑产率(TAA->TGG/TAG/TGA)高达3.6%,TAA->TGG高达2.4%(图33C,图43E)。对经处理的肌肉进行的免疫印迹证实了免疫荧光的观察结果,表明了1～2.5%的蛋白恢复(图44B)。作为基准,注射带有CRISPR-Cas9的载体的肌肉也预期地导致了在一部分肌肉细胞中恢复抗肌萎缩蛋白的表达(图33B),对经处理的肌肉进行的免疫印迹证实了高达10%的蛋白恢复

(图44C)。

[0371] 为了进一步证实该方式的有效性,接下来在一种独立的人类疾病小鼠模型,鸟氨酸转氨甲酰酶(OTC)缺陷的雄性稀疏毛皮灰(sparse fur ash,spf^{ash})小鼠模型中,评估了ADAR介导的RNA编辑。spf^{ash}小鼠在OTC基因第四个外显子的最后一个核苷酸中携带有G->A点突变,这导致OTC mRNA缺乏并产生突变蛋白。最近的研究已经表明,使用CRISPR-Cas9和基于同源重组的策略,可在新生小鼠中对这种突变进行强有力的校正。为了测试该系统在spf^{ash} OTC的mRNA中编辑点突变(图33D)的有效性,在体外对构建体进行了评估(图45A)。将构建体包装到具有高肝脏嗜性的AAV8中,并以2.5E+12vg/小鼠的量注射到10～12周龄的spf^{ash}小鼠中。注射后三至四周,从spf^{ash}小鼠、野生型同窝仔畜,和经ADAR2靶向和非靶向载体处理的spf^{ash}小鼠,收集肝脏样品,并经由NGS评估相应的编辑效率。值得注意的是,在递送adRNA和ADAR2后,在正确剪接的OTC mRNA中观察到了0.8～4.7%的经编辑mRNA,有趣的是,仅adRNA导致了低但显著的RNA编辑产量(图33E)。此外,在递送超高活性的ADAR2突变体(E488Q)后,在正确剪接的OTC mRNA中观察到高编辑分数(4.6～33.8%)(图33E,图45B),在OTC pre-mRNA中观察到4.6～8.2%的编辑分数(图45C),并证实了错误剪接产物减少(图45D)。对经处理的肝脏样品进行的免疫印迹证实了OTC蛋白的部分恢复(2.5～5%)(图45E)。

[0372] 总起来讲,结果建立了RNA导向的ADAR在点突变的体内RNA编辑中的效用。在某些情况下,ADAR酶的序列偏好、RNA折叠、固有半衰期、定位、翻译机制和驻留的RNA结合蛋白可能会影响RNA中的靶标位点的可接近性和可编辑性,并且可能是重要的设计参数,是为实现有效的靶向而要考虑的因素。例如,在mdx模型中,基于ADAR的RNA编辑方式可能必须与无义介导的突变抗肌萎缩蛋白mRNA衰变进行竞争,并且还需要在非理想侧翼核苷酸的情况下实现两个A->I取代,以消除提前终止密码子以及对RNA稳定性和功能的潜在影响。此外,在spf^{ash}模型中,靶向瞬时OTC pre-mRNA的需要可能需要快速的靶标衔接和编辑。还可能需要进一步解决该系统的重要局限性,例如由固有的酶-RNA结合诱导的脱靶,进行性,杂泛性(promiscuity),由递送形式本身(诸如脂质、纳米颗粒或病毒)对干扰素应答的刺激,进而导致增加的内源性ADAR表达、adRNA诱导RNAi的潜力以及adRNA的反义结构域的脱靶杂交(其可能会产生有害作用)。在这方面,研究表明了在全身注射高活性ADAR突变体的小鼠中具有毒性(图46)。这些研究对于帮助系统性改善这种方式的特异性和安全性是至关重要的。在考虑用于基因治疗的RNA靶向时,尤其是经由使用非整合载体时,另一个重要的考虑因素可以是,由于经编辑的mRNA和效应因子的有限的半衰期,需要定期重新施用效应构建体。在这方面,与基于CRISPR的RNA编辑方式相比,RNA导向的ADAR策略可以直接与人类治疗相关,因为相同的形式仅利用效应RNA和人类蛋白。此外,随着ADAR的广泛表达,例如,在大多数人体组织中的ADAR1和尤其在肺和脑中的ADAR2,这些ADAR经由带长反义结构域的adRNA进行的内源性募集(如图32,图33E和图34,图36中所示),为有效的RNA编辑提出了非常有吸引力的策略。随着不断改进,该工具集可对各种基础科学和治疗应用产生广泛的影响。

[0373] 图32:改造可编程的RNA编辑并表征特异性概况:(A)经由利用全长ADAR2和源自于GluR2转录物的经改造的adRNA的构建体,或MS2外壳蛋白(MCP)融合体与ADAR1/2脱氨酶结构域和相应的带有adRNA的MS2发夹的构建体,进行RNA编辑的示意图。(B)通过Sanger测序

量化的不同RNA编辑构建体对内源性RAB7A转录物的RNA编辑效率的比较(效率以Sanger峰高G/(A+G)之比来计算)。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n＝3)。(C)代表在参考位点上观察到的A->G编辑产量分布的小提琴图,其中发现至少一个处理样品的编辑产量相对于对照样品具有明显变化(费希尔精确检验,FDR＝1％)。蓝色圆圈表示在RAB7A转录物内的目标A位点处的编辑产量。黑色点表示中值脱靶编辑产量。为了使分布的形状更好地可视化,将它们沿y轴的最大范围在图上进行了均衡化,并截断在60％的产量处。

[0374] 图33:在人类疾病的小鼠模型中的体内RNA编辑:(A)在杜氏肌肉萎缩症的mdx小鼠模型中恢复抗肌萎缩蛋白表达的DNA和RNA靶向方法的示意图:(i)导致框内切除外显子23的基于双gRNA-CRISPR的方法,和(ii)赭石密码子的基于ADAR2和MCP-ADAR1的编辑。(B)TA肌肉中的抗肌萎缩蛋白的免疫荧光染色示出了,在经处理的样品(肌内注射AAV8-ADAR2、AAV8-ADAR2(E488Q)和AAV8-CRISPR)中表达的部分恢复。在经处理的样品中也看到nNOS定位的部分恢复(比例尺:250μm)。(C)在相应的经处理的成年mdx小鼠中,体内TAA->TGG/TAG/TGA的RNA编辑效率。数值代表平均值+/-SEM(分别为n＝4、3、7、3、3、10、3、4个独立的TA肌肉)。(D)在鸟氨酸转氨甲酰酶缺乏症的spf^{ash}小鼠模型中,OTC基因座和经由ADAR2介导的RNA编辑校正突变型OTC mRNA的方法的示意图,其中鸟氨酸转氨甲酰酶缺乏症在外显子4的最后一个核苷酸的供体剪接位点处具有G->A点突变。(E)在经处理的成年Spf^{ash}小鼠的肝脏(眼眶后注射AAV8-ADAR2和AAV8-ADAR2(E488Q))中,正确剪接的OTC mRNA中的体内RNA校正效率。数值代表平均值+/-SEM(分别为n＝4、4、3、3、4、5个独立的动物)。

[0375] 载体的设计和构建

[0376] 将1或2个拷贝的adRNA克隆到AAV载体中,该AAV载体包含人U6和小鼠U6启动子,以及驱动酶表达的CMV启动子。为了构建GFP报告因子——GFP-琥珀(Amber)、GFP-Ochre和GFP-欧泊(Opal),分别用'TAG'、'TAA'和'TGA'取代野生型GFP的Y39残基而合成了三种基因块,并将它们克隆到CAG启动子的下游。为了构建OTC和DMD报告因子,将带有要编辑的目标腺苷的spf^{ash} OTC和mdx DMD转录物的200bp片段克隆到CAG启动子的下游。

[0377] 哺乳动物细胞的培养和转染

[0378] 所有的HEK 293T细胞均在37℃和5％CO_2气氛的培养箱中,在补充有10％FBS和1％抗生素-抗真菌药的杜氏改良伊格尔培养基(Thermo Fisher)中生长。所有的体外转染实验均使用商业转染试剂Lipofectamine 2000(Thermo Fisher),在HEK 293T细胞中进行。涉及报告因子的所有体外RNA编辑实验,均使用400ng报告因子质粒和800ng adRNA+酶质粒,在24孔板中进行。靶向内源性转录物的所有体外RNA编辑实验,均使用800ng adRNA/酶质粒,在24孔板中进行。基于dCas13b-ADAR2DDE488Q的RNA编辑实验,使用800ng酶质粒(Addgene#103864)和800ng gRNA质粒,在细胞为25～30％汇合度时进行转染,并在转染后60小时收获,以量化编辑。化学合成的adRNA(由IDT或Synthego合成)使用Lipofectamine 3000(Thermo Fisher),以20pmol/孔的量进行转染。

[0379] AAV载体的产生

[0380] 使用HEK 293T细胞,经由三重转染方法来产生AAV8颗粒,并经由碘克沙醇梯度进行纯化。转染时的汇合度约为80％。在转染前两个小时,将补充有10％FBS的DMEM添加到HEK 293T细胞中。每种病毒均在5x 15cm的平板中产生,其中各平板使用PEI(1μg/uL线性PEI在

1xDPBS pH 4.5中,使用HCl),以4:1的PEI:DNA质量比,使用7.5μg pXR-8、7.5μg重组转移载体,7.5μg pHelper载体进行转染。将混合物在RT下孵育10分钟,然后滴加到细胞培养基上。72小时后收获病毒,并使用碘克沙醇密度梯度超速离心法进行纯化。然后,使用50kDA滤膜(Millipore),用补充有50mM NaCl和0.0001%的Pluronic F68(Thermo Fisher)的1x PBS(pH 7.2)透析病毒,至终体积～1mL,并根据标准(ATCC VR-1616),使用对ITR区域的特异性引物通过qPCR进行量化。

[0381] AAV-ITR-F:5'-CGGCCTCAGTGAGCGA-3'(SEQ ID NO:149),和

[0382] AAV-ITR-R:5'-GGAACCCCTAGTGATGGAGTT-3'(SEQ ID NO:150)。

[0383] RNA的分离和下一代测序文库的制备

[0384] 根据制造商的方案,使用RNeasy+通用迷你试剂盒(RNeasy Plus Universal Mini Kit,Qiagen),从动物组织中提取RNA。使用RNeasy迷你试剂盒(Qiagen)从细胞中提取RNA。使用Protoscript II第一链cDNA合成试剂盒(NEB),从500ng RNA合成cDNA。下一代测序文库的制备如下。简而言之,使用KAPA Hifi HotStart PCR混合物(Kapa Biosystems),通过PCR扩增1μl上面制备的cDNA,其中PCR使用扩增目标位点周围约150bp。纯化PCR产物(Qiagen PCR纯化试剂盒/凝胶提取试剂盒),以消除副产物。使用NEBNext Multiplex Oligos for Illumina试剂盒(NEB),来构建库。用索引引物,扩增10ng投入的DNA。然后将样品合并,并加载到Illumina Miseq(150bp单端运行)或Hiseq(100bp双端运行)上。使用CRISPResso进行数据分析(Pinello,L.et al.2016)。对于所有体内实验,至少分析100,000个读数。使用NEBNext Poly(A)mRNA磁性分离模块和用于Illumina的NEBNext超RNA文库制备试剂盒(NEBNext Ultra RNA Library Prep Kit for Illumina),从300ng RNA制备RNA-seq文库。将样品合并,并加载到Illumina Hiseq(100bp双端运行)上。

[0385] Spf^{ash}小鼠中OTC mRNA编辑产量的量化

[0386] spf^{ash}小鼠带有三种形式的OTC RNA:pre-mRNA,正确剪接的mRNA,和由于在内含子4中使用了隐蔽剪接位点48个碱基对而形成的错误剪接的延长的mRNA。使正确剪接的mRNA的总数目为X,错误剪接的变异体的总数目为Y,pre-mRNA为Z。Xe、Ye和Ze表示三种形式中的A->G编辑的mRNA。理想情况下,mRNA的编辑产量可计算为(Xe+Ye+Ze)/(X+Y+Z)。但是,因为不可能使用相同的引物来扩增剪接的变体和pre-mRNA变体,因此图34E示出了在正确剪接的mRNA中被编辑的转录物的分数(Xe/X),该正确剪接的mRNA转而将被翻译以产生OTC蛋白。另外,图48C示出了在pre-mRNA中被编辑的转录物的分数(Ze/Z)。该分数在正确剪接后将有助于OTC蛋白的形成。最后,错误剪接的mRNA导致产生延长16个氨基酸的蛋白,其被选择性降解。在图48D中,示出了与X和Y相对应的条带。

[0387] 动物实验

[0388] 所有动物程序均按照加利福尼亚大学圣地亚哥分校的机构动物护理和使用委员会(IACUC)批准的方案进行。小鼠获自杰克逊实验室。使用2E+12vg/肌肉,将AAV注射到mdx小鼠(C57BL/10ScSn-Dmd^{mdx}/J)的腓肠肌或TA肌肉中。使用2.5E+12vg/小鼠,经由眼眶后注射,将AAV注射到spf^{ash}小鼠(B6EiC3Sn a/A-$Otc^{spf-ash}$/J)中。将呈现出毛发粗糙、移动缓慢并且略微弓背的小鼠称为患病小鼠,并被安乐死。

[0389] 免疫荧光

[0390] 将收获的腓肠肌或TA肌肉放入含OCT化合物(VWR)的模具中,并在液氮中快速冷

冻。将10μm的切片切在经预处理的组织学载玻片上。载玻片用4％多聚甲醛进行固定。分别用针对抗肌萎缩蛋白的C端结构域(1:200,Abcam 15277)和nNOS的N端结构域(1:100,Immunostar 24431)的兔多克隆抗体,然后用驴抗兔的Alexa546二抗(1:400,Thermo Fisher),来检测抗肌萎缩蛋白和nNOS。

[0391] 免疫印迹

[0392] 将来自mdx小鼠的肌肉活检和来自spf^{ash}小鼠的肝脏活检,在含蛋白酶抑制剂鸡尾酒(Roche)的RIPA缓冲液(Sigma)中进行片段化,并在冰上孵育1小时,同时进行间歇涡旋。将样品在4℃,以15500xg离心30分钟,分离上清液,并用Pierce考马斯+(Pierce Coomassie Plus,Bradford)分析试剂盒(Thermo Fisher)进行量化。将蛋白分离物与4x利姆里(Laemmli)上样缓冲液(Biorad)和2-巯基乙醇(Biorad)混合,并在100℃煮沸10分钟。将来自肌肉活检的100μg总蛋白或来自肝脏活检的60μg总蛋白,上样至有Tris-甘氨酸-SDS缓冲液(Biorad)的4～15％Mini Protean TGX凝胶(Biorad)的每个孔中,并在100V电泳60分钟。将来自肌肉活检的蛋白,在34V过夜,转移到硝酸纤维素膜上,而将来自肝脏活检的蛋白在1X tris-甘氨酸转移缓冲液中,在4℃,65V转移1小时30分钟。将印迹在5％奶-TBST中封闭1小时。印迹用兔抗-抗肌萎缩蛋白(1:200,Abcam 15277)、兔抗-GAPDH(1:4000,Cell Signaling 2118S)、兔抗-OTC(1:800,Abcam 203859)和小鼠抗-ADAR2(1:150,Santa Cruz Biotechnology 73409),在5％奶-TBST中,于4℃探测过夜。印迹用TBST洗涤,然后与抗兔或抗小鼠的辣根过氧化物酶偶联的二抗(Cell Signaling),在5％奶-TBST中孵育1小时。在用TBST洗涤后,使用SuperSignal免疫毫微微化学发光底物(SuperSignal West Femto Chemiluminescent Substyeild,Thermo Fisher)和X射线胶片,使印迹可视化。

[0393] 统计学和可重复性

[0394] 体外实验:体外实验一次至少有3次独立的重复。体内实验:对于mdx小鼠模型,进行了两次基于ADAR2和MCP-ADAR1(E1008Q)NLS的实验。两轮实验都产生了一致的RNA编辑效率,抗肌萎缩蛋白免疫荧光和抗肌萎缩蛋白的恢复(如免疫印迹所示)。进行了一次基于ADAR2(E488Q)和CRISPR-Cas9的实验。对于spf^{ash}小鼠模型,基于小鼠的可用性,进行了两次实验。在两轮实验中,OTC转录物、剪接的mRNA和pre-mRNA的RNA编辑效率均一致。在实验组1中,对动物进行了RT-PCR和免疫印迹。

[0395] 1.RNA A->G编辑的量化

[0396] (a)RNA-seq读数比对

[0397] 使用STAR aligner版本2.6.0c(Dobin A等人2013),将每个读数配偶具有100个碱基的RNA-seq读数对,与GRCh38参考基因组进行比对。使用来自GENCODE发行28(GRCh38.p12)的主要组装注释,来构建基因组索引。除以下相关设置外,使用默认参数来运行STAR:readMapNumber＝-1,alignSJoverhangMin＝5,alignSJDBoverhangMin＝1,alignEndsType＝EndToEnd,outFilterMismatchNmax＝10,outFilterMultimapNmax＝1,outSAMunmapped＝None(无),outSAMmultNmax＝1。使用Samtools排序(sort),按基因组坐标对所得的唯一比对对的读数进行排序(Li H.等,2009)。重复的读数对使用Samtools标记(markup)进行标记,并从后续分析中去除。表4中报告了每个样品的总的、比对的、重复和剩余的读数(非成对)的记录。

[0398] (b)用于量化编辑产量的参考位点的选择

[0399] 对具有A-至-G编辑产量显著变化的位点(见下文)的评估,对各样品可获得的唯一比对读数的数目敏感。为了在比较不同样品时,在显著编辑位点方面最大程度地减少潜在误差,使用了Samtools视图(选项-s),对每个HEK293T样品的唯一比对读数进行了下采样,并在表4中报告了下采样分数。用样品中最小的唯一比对读数的数目除以下采样的样品可获得的唯一比对读数,来计算这些分数。未对对照样品(表4中的第一个)的读数进行下采样。然后,使用Samtools mpileup对下采样的读数进行处理。对该工具的输出进行解析,以提取GRCh38参考基因组序列中,在每个A-位点和T-位点处的比对读数中发现的每个碱基的计数。忽略插入和缺失。下游分析中省略了读数覆盖率小于10的参考位点。在表4中列出的样品中,剩余的读数覆盖率至少为10的参考A-位点和T-位点的数目变化了～15%。在不进行下采样的情况下,则发现该数目的变化为～50%。在至少一个样品的比对读数中,通过选择所有样品所共有的,和分别在参考A-位点或T-位点处观察到至少一次G或C的那些位点,从读数覆盖率至少为10的参考A-位点和T-位点,选择总位点(A-位点和T-位点)的最终列表。丢弃不是所有样品所共有的,或者在所有样品中观察到的编辑事件为零的那些站点。

[0400] (c)A-至-G编辑产量的显著变化的评估

[0401] 为了揭示A-至-G编辑量的显著变化,考虑了几对对照和处理样品。对于每一对,对照样品均是表4中列出的第一个样品,而处理样品是图32中显示的样品之一。对于每对比较的样品,以及对于如上所述选择的每个参考A-位点,使用2x2权变矩阵C进行费舍尔(Fisher)精确检验,其中权变矩阵C具有如下定义的项目:C1,1＝对照样品中观察到的除G以外的碱基的计数,C2,1＝对照样品中观察到的G碱基的计数,C1,2＝测试样品中观察到的除G以外的碱基的计数,C2,2＝测试样品中观察到的G碱基的计数。对于每个选定的参考T-位点,都使用类似的权变矩阵,不同的是,在上述定义中G被C取代。使用本杰米尼-霍奇伯格(Benjamini-Hochberg)方法,调整所有选定参考位点和给定样品比较所计算出的p值,以进行多次测试。具有小于1%的错误发现率(FDR)的调整后的p值并且具有编辑产量至少1.1的倍数变化的A-位点和T-位点,被视为分别在正向和反向转录物上具有A-至-G编辑产量的显著变化。每次样品比较时这些位点的计数在图38～图42中示为N_{sig},并报告在表5的“变化的位点”列中。计算了具有A-至-G编辑产量显著变化的参考位点的总数目。这些位点处的编辑产量用于构建图32中所示的分布。使用在RAB7A转录物中的预期目标A-位点处观察到的计数,将每一样品的上靶A-至-G编辑产量(如图32和图38～图42中的蓝色圆圈所示)估计为$C_{2,2}/(C_{1,2}+C_{2,2})$。这些数值报告在表5中的“编辑产量”列中。在选择了参考组装hg38后,通过将以下序列提交给BLAT,发现预期的目标A-位点的基于1(1-based)的基因组坐标为chr3:128814202:

[0402] AGCGGCAGTATTCTGTACAGTAGACACAAGAATTATGTACGCCTTTTATCA(SEQ ID NO:151)。

[0403] 图8-改造GluR2 adRNA:支架结构域改造。adRNA支架的序列信息:突出显示了ADAR募集结构域、反义RNA靶向结构域和胞嘧啶错配。将经突变以创建稳定支架的碱基对进行编号并以红色突出显示,而编辑诱导元件基序则以绿色显示。显示了通过Sanger测序定量的,由此生成的用于OTC报告因子转录物的支架的编辑效率的量化。数值代表平均值+/-SEM(n＝3)。实验在HEK 293T细胞中进行。

[0404] 图34-改造GluR2 adRNAs:反义结构域改造。(a)使用adRNA支架2来优化adRNA反义区域:长度和与ADAR2募集区的距离在系统上是不同的。数值代表平均值+/-SEM(n＝3)。(b)

评估了U6启动子转录的反义结构域长度逐渐增加的adRNA,结合0、1或2个GluR2结构域,其在有或没有外源性ADAR2表达的情况下诱导靶向RNA编辑的能力。数值代表平均值+/-SEM(n＝3)。这些数据的一部分在图1b中重复使用。以上所有实验均在HEK 293T细胞中进行。(c)实验确认了内源性ADAR1和ADAR2(相对于GAPDH)在HEK 293T和HeLa细胞系中的表达。观察到的水平与《人类蛋白质图谱(Human Protein Atlas)》(参见proteinatlas.org上的全球性网页)中记录的水平相似。

[0405] 图35-改造MS2 adRNA。(a)系统评估了MS2 adRNA的反义RNA靶向结构域。数值代表平均值+/-SEM(n＝3)。(b)通过MCP-ADAR2 DD-NLS进行的上靶RNA编辑需要MS2 adRNA的共表达。数值代表平均值+/-SEM(n＝3)。实验在HEK293T细胞中进行。

[0406] 图36-对一组靶标的RNA编辑产量的分析。(A)比较OTC报告因子转录物通过GluR2 adRNA和MS2 adRNA导向RNA编辑构建体,以及基于Cas13b的REPAIR构建体,进行RNA编辑的效率。数值代表平均值+/-SEM(对于报告因子和基于Cas13b的构建体,n＝6,对于其他构建体,n＝3)。(B)针对一组mRNA,在有或没有外源性ADAR2表达的情况下,测试了化学合成的adRNA型式。在图中说明了确切的化学修饰以及adRNA的来源。数值代表平均值+/-SEM(n＝3)。(C)分析所选择的一连串内源性靶标的RNA编辑产量,以涵盖一系列表达水平。还针对多个内源性mRNA靶标,在有或没有外源性ADAR2表达的情况下,评估了U6转录的没有或具有两个GluR2结构域的长adRNA。甚至在没有外源性ADAR2表达的情况下,也可以在测试的基因座处观察到编辑。数值代表平均值+/-SEM(n＝3)。实验在HEK 293T细胞中进行。

[0407] 图37-ADAR2变体及其对编辑和特异性的影响。(A)比较RAB7A转录物通过GluR2 adRNA和MS2 adRNA导向RNA编辑构建体,以及基于Cas13b的REPAIR构建体,进行RAB7A转录物的上靶RNA编辑和在侧翼腺苷中的编辑。描绘了平均值(n＝3)编辑产量。实验在HEK 293T细胞中进行,并且编辑效率计算为Sanger峰高G/(A+G)之比。(B)在spf^{ash} OTC报告因子转录物(p＝0.037,未配对t检验,双尾;数值代表平均值+/-SEM(n＝3))和(C)mdx DMD报告因子转录物(分别为p＝0.048,p＝0.012,未配对t检验,双尾;数值代表平均值+/-SEM(n＝3))的体外编辑中,ADAR2(E488Q)显示出比ADAR2更高的效率。(d)比较了ADAR2、ADAR2(E488Q)和ADAR2(Δ1-138),对于OTC报告因子转录物(上图)和内源性RAB7A转录物(下图)的编辑效率和特异性概况。热图指示在目标附近的A->G编辑(红色箭头)。数值代表平均值+/-SEM(n＝3)。实验在HEK 293T细胞中进行,并将编辑效率计算为Sanger峰高G/(A+G)之比。

[0408] 图38-RNA编辑方法(Cas13b-ADAR REPAIR+/-gRNA)的转录组规模(scale)的特异性概况。2D直方图比较了使用每种Cas13b-ADAR2构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。每个直方图表示同一组8,729,464个参考位点,其中读数覆盖率至少为10,并且在至少一个样本中检测到至少一个推定编辑事件。当处理样品与对照样品进行比较时,以红色突出显示的统计箱子(bin)包含A->G编辑产量有显著变化的位点。每个图中的红色交叉表示具有调整最小的p值的100个位点。蓝色圆圈表示RAB7A转录物内预期的目标A-位点。接近左下角的箱子中的大量计数可能不仅对应于测试和对照样品中的较低编辑量,而且还对应于测序误差和比对误差。接近每个图右上角的箱子中的大量计数可能对应于纯合的单核苷酸多态性(SNP),以及参考基因组和实验中使用的HEK293T细胞系的基因组之间的其他差异。

[0409] 图39-RNA编辑方法(ADAR2+/-adRNA)的转录组规模的特异性概况。这些研究使用

的型式是GluR2 adRNA(1,20,6)。2D直方图比较了使用每种ADAR构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。图38中提供了更多详细信息。

[0410] 图40-RNA编辑方法(MCP-ADAR1 DD+/-adRNA)的转录组规模的特异性概况。2D直方图比较了使用每种ADAR构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。图38中提供了更多详细信息。

[0411] 图41-RNA编辑方法(MCP-ADAR2 DD+/-adRNA)的转录组规模的特异性概况。2D直方图比较了使用每种ADAR构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。图38中提供了更多详细信息。

[0412] 图42-具有构建体特征的转录组规模编辑特异性的变化。(A)箱线图中的每个点都对应于图32中列出的MCP-ADAR构建体之一的经编辑的位点的分数。每个构建体的经编辑的位点的分数,通过将A-至-G编辑产量(见表3)发生显著变化的参考位点数目除以考虑的参考位点总数目8,729,464来计算得出。使用曼-惠特尼U检验比较了水平轴上指示的构建体特征,产生了NLS vs.NES的p值为0.16,ADAR1 vs.ADAR2的p值为0.0070,“-adRNA”vs.“+adRNA”的p值为0.72,以及“ADAR WT”vs.“ADAR E>Q”的p值为0.038(在所有条件下n=8)。(B)2D直方图比较了使用每种构建体观察到的转录组范围的A->G编辑产量(y轴)与使用对照样品观察到的产量(x轴)。图38中提供了更多的详细信息。插图示出了小提琴图,其代表在参考位点观察到的A->G编辑产量的分布,其中相对于对照样品,发现至少一个处理样品的编辑产量具有显著变化(费希尔精确检验,FDR=1%)。蓝色圆圈表示RAB7A转录物内在目标A-位点处的编辑产量。为了使分布的形状更好地可视化,将它们沿y轴的最大范围在图上进行了均衡化,并截断在60%的产量处。此处的样品对应于转染了携带adRNA的长反义结构域的293T,该adRNA可以经由外源性和/或内源性ADAR募集进行RNA编辑。

[0413] 图43-对mdx小鼠中的体外和体内抗肌萎缩蛋白RNA编辑实验的优化和评估。(A)利用全长ADAR2以及改造的adRNA或反向adRNA(radRNA)进行RNA编辑的示意图;(ii)琥珀(amber)和赭石(ochre)终止密码子以一步和两步的方式进行的RNA编辑效率。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n=3)。(B)赭石密码子的RNA编辑需要在adRNA或radRNA的反义RNA靶向结构域中的两个胞嘧啶错配,来恢复GFP表达。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(n=3)。(C)用于在体内递送adRNA和ADAR2的AAV载体的示意图,以及在存在1或2个拷贝的adRNA的情况下,经由AAV载体递送的琥珀和赭石终止密码子的RNA编辑的体外优化(分别为p=0.0003,p=0.0001,p=0.0015,不成对t检验,双尾)。实验在HEK 293T细胞中进行。数值代表平均值+/-SEM(对于报告因子,n=3,对于其他条件,n=6)。(D)代表性的Sanger测序图,显示了在mdx DMD报告因子转录物中赭石终止密码子(TAA->TGG)的编辑(通过NGS进行量化)。实验在HEK 293T细胞中进行(n=3)。(E)经处理的mdx小鼠的体内RNA编辑分析的代表性实施例(使用NGS进行量化)。

[0414] 图44-在mdx小鼠中的体内抗肌萎缩蛋白RNA编辑实验的免疫荧光和免疫印迹的分析。(A)在TA肌肉中对抗肌萎缩蛋白的免疫荧光染色示出了经处理的样品(肌肉内注射AAV8-ADAR2、AAV8-ADAR2(E488Q)、AAV8-MCP-ADAR1(E1008Q)NLS)中表达的部分恢复。在经处理的样品中还看到nNOS定位的部分恢复(比例尺:250μm)。(B)免疫印迹示出了在mdx小鼠(其注射有编辑工具的两种组分,酶和adRNA)的TA肌肉中抗肌萎缩蛋白表达(1～2.5%)的

部分恢复，以及在注射的TA肌肉中注射后长达8周的稳定的ADAR2表达。(C)免疫印迹示出了，使用AAV8-CRISPR使抗肌萎缩蛋白表达部分恢复(10％)。

[0415] 图45-在spf^{ash}小鼠中体外和体内的OTC RNA编辑实验的优化和评估。(A)代表性Sanger测序图谱示出了在spf^{ash} OTC报告因子转录物中点突变的校正(使用NGS进行了量化)。实验在HEK 293T细胞中进行(n＝3)。(B)经处理的spf^{ash}小鼠的体内RNA编辑分析的代表性实施例示出了，在经正确剪接的OTC mRNA中的点突变的校正(使用NGS进行了量化)。(C)在经处理的成年Spf^{ash}小鼠的肝脏(眼眶后注射A488-ADAR2和AAV8-ADAR2(E488Q))中的OTC pre-mRNA的体内RNA校正效率。数值代表平均值+/-SEM(分别为n＝4、4、3、3、4、5只独立的动物)。(D)PCR产物示出了经正确和错误剪接的OTC mRNA。经错误剪接的mRNA被延长了48个碱基对。在用adRNA+ADAR2(E488Q)处理的小鼠中，经错误剪接的mRNA的分数降低了。(E)对OTC的免疫印迹示出了，对经处理的成年spf^{ash}小鼠中表达的部分恢复(2.5％～5％)，和在注射后三周稳定的ADAR2(E488Q)表达。

[0416] 图46-对体内RNA编辑实验的毒性分析。对动物实验的总结，记录了AAV的施用途径、递送的构建体以及注射后3周的注射小鼠的健康状况。

[0417] 实施例5

[0418] 本文公开了使用ADAR1的E100Q突变体(包含在MCP-ADAR1(E100Q)中)，在杜氏肌肉萎缩症的mdx小鼠模型中的实验结果。在某些情况下，与ADAR2和ADAR2的E488Q突变体相比，这种突变体可以改善体内的编辑产量。本文还公开了以通过编辑内含子中的剪接受体位点或分支点，以改变剪接模式的方式来应用ADAR系统，从而导致外显子跳跃。另外，设想并举例说明了使用APOBEC的方法。本文展示的实施例中，在感兴趣的mRNA上创建载脂蛋白B mRNA的局部结构，可以使得C→T RNA编辑成为可能，其中ACF-APOBEC复合物结合该载脂蛋白B mRNA。本文还公开了利用ADAR酶对RNA和DNA进行可编程的编辑的方法。

[0419] 为了利用ADAR编辑系统，无论是否有外源性ADAR1或ADAR2，都可以生成adRNA。实施例为给定靶标提供了感兴趣的特异性adRNA。此外，本文提供了示例性的化学合成的adRNA，其在整个反义区域中的第一个和最后3、6、9或12个核苷酸中具有2'-O-甲基3'硫代磷酸酯修饰，和/或除了以错配位点为中心的3个核苷酸(以下有下划线)之外具有2'-O-甲基修饰，再加上靶向部分(如GalNc或胆固醇或穿透细胞的肽)可用于在有或没有外源性ADAR1/2过表达的情况下，在细胞或体内(尤其是在肝脏、肺和脑中)改造靶向RNA编辑。

[0420] 总体设计：

[0421] adRNA22：

[0422] GTGGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCCACNNNNNNCNNNNNNNNNNNNNN(SEQ ID NO:152)

[0423] adRNA32：

[0424] GGTGTCGAGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCTCGACACCNNNNNNCNNNNNNNNNNNNNN(SEQ ID NO:153)

[0425] 具体实施例：

[0426] adRNA22_RAB7A：

[0427] GTGGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCCACTGCCGCCAGCTGGATTTCCC(SEQ ID NO:154)

[0428] adRNA32_RAB7A：

[0429] GGTGTCGAGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCTCGACACCTGCCGCCAGCTGGATTTCCC(SEQ ID NO:155)

[0430] adRNA22_CKDN2A：

[0431] GTGGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCCACCTCCTCCACCCGACCCCGGG(SEQ ID NO:156)

[0432] adRNA32_CKDN2A：

[0433] GGTGTCGAGAAgAGgAgAACAATATGCTAAATGTTGTTcTcGTcTCCTCGACACCCTCCTCCACCCGACCCCGGG(SEQ ID NO:157)

[0434] 下面提供的实施例描述了为进一步阐明本文公开的实施方式而采取的多个步骤，诸如：

[0435] (1)递送具有MS2-adRNA的MCP-ADAR构建体，该MS2-adRNA靶向在mdx小鼠的抗肌萎缩蛋白转录物中的提前终止密码子；

[0436] (2)通过在杜氏肌肉萎缩症的mdx小鼠模型中创建点突变来测试外显子跳跃，并将adRNA与ADAR2/ADAR2(E488Q)一起递送，或MS-adRNA与MCP-ADAR1/ADAR(E1008Q)一起递送，或ms-adRNA与MCP-ADAR2/ADAR2(E488Q)一起递送-用于在剪接位点产生A→G取代；

[0437] (3)在spc-ash小鼠的OTC转录物中，测试剪接位点突变的编辑；

[0438] (4)测试鸟氨酸转氨甲酰酶缺乏症的spf-ash小鼠模型中，测试对突变体剪接位点的编辑以改变的剪接模式；

[0439] (5)通过在剪接位点产生A→G取代以敲低SCN9A转录物，打靶小鼠的SCN9A转录物以对改造疼痛的不敏感性；

[0440] (6)改造APOBEC gRNA，以在感兴趣的目标mRNA上创建脂蛋白B mRNA的局部结构；和

[0441] (7)通过利用RNA、DNA和RNA-DNA杂交体来募集ADAR，测试编辑RNA和DNA中的ADAR。

[0442] 在实施例中，观察到在编辑mdx小鼠的抗肌萎缩蛋白转录物中的提前终止密码子时，使用MS2-adRNA和MCP-ADAR1(E1008Q)，获得了增强的RNA产量和蛋白表达。在使用adRNA递送ADAR2(E488Q)后，还证实了在spf-ash小鼠模型中错误剪接的OTC mRNA的减少。

[0443] 因此，本公开描述了一种能够实现RNA的位点特异性A->G编辑的系统。这种方法可用于编辑剪接受体位点和分支点，以改变剪接。这也可以实现外显子跳跃。进一步的实施方式考虑了使用DNA的A->G编辑。另外，本公开通过使用与ACF1一起表达的APOBEC1，提供了在RNA中进行C->T编辑的潜力。

[0444] 本公开描述了第一个体内位点特异性的RNA编辑。实际上，在肌肉萎缩症的mdx小鼠模型中，MCP-ADAR1(E1008Q)的使用显示出比以前的构建体更高的编辑效率。进一步的体外和体内测试，检查了(1)RNA编辑用于改变剪接，以及(2)通过APOBEC与过表达的ACF一起，进行C->T编辑的功效，和(3)用于编辑DNA的ADAR。

[0445] 与其他ADAR2系统(例如Stafforst，Zhang和Rosenthal实验室)和基于Cas13d的剪接抑制相比，当前的ADAR系统是独一无二的。

[0446] 如本文所述，上述ADAR系统包括以下项，或基本上由以下项组成，或进一步由以下项组成：与靶标RNA互补的RNA靶向结构域，和一种或多种能够募集ADAR的ADAR募集结构域。

引入这些组成部分后，ADAR酶可以催化靶标腺苷向肌苷的转化，从而修复点突变。本公开描述了使用ADAR2或MCP-ADAR1/2-NLS或其超高活性突变体，在内含子的剪接受体位点和/或分支点处创建A->G编辑，以改造外显子跳跃。跳过的对称外显子导致形成截短的蛋白。这种策略可用于跳过DMD的mdx小鼠模型的抗肌萎缩蛋白转录物中的外显子23，其在该外显子中带有提前终止密码子。跳过外显子23导致功能性截短的抗肌萎缩蛋白的翻译。另外，跳过不对称的外显子会导致移码突变。因此，通过跳过必需的外显子或不对称的外显子，可以改造基因敲低。此外，编辑起始密码子ATG和科扎克(Kozak)/夏因-达尔加诺(Shine-Dalgarno)序列，也可以帮助改变翻译效率并导致基因敲低。外显子跳跃策略可用于打靶SCN9A转录物，以改造对疼痛的不敏感性。可以在感兴趣的mRNA靶点处创建载脂蛋白B mRNA的局部结构，用于小心地将要编辑的C定位在与载脂蛋白B mRNA中天然存在的位点类似的位置上，其中ACF-APOBEC复合物结合该载脂蛋白B mRNA。这可以通过过表达成对的adRNA来创建载脂蛋白B mRNA样结构来实现，该结构可以通过过表达MCP-ACF1和APOBEC1来进行编辑。

[0447] 实施例6-经由创建剪接受体和/或分支点突变来进行外显子跳跃

[0448] 本公开展示了使用ADAR2或MCP-ADAR1/2-NLS或它们的超高活性突变体，在内含子的剪接受体位点和/或分支点创建A->G编辑，以改造外显子跳跃。跳过对称外显子导致形成截短的蛋白。在本实施例中，该策略可用于跳过DMD的mdx小鼠模型的抗肌萎缩蛋白转录物中的外显子23，其在该外显子中带有提前终止密码子。跳过外显子23导致功能性截短的抗肌萎缩蛋白的翻译。另外，跳过不对称外显子会导致移码突变。因此，通过跳过必需的外显子或不对称的外显子，可以改造基因敲低。此外，编辑起始密码子ATG和Kozak/Shine-Dalgarno序列，有助于改变翻译效率并导致基因的敲低。该策略可用于打靶SCN9A转录物，以改造对疼痛的不敏感性。

[0449] 在下文中提供了用于实现该目的的示例性序列。

[0450] 序列集1：示例性的adRNA序列

[0451] 用于在mdx小鼠的抗肌萎缩蛋白转录物中进行外显子23跳跃的adRNA序列(SEQ ID NO:158-159)：

[0452]

adDMD-I22-E23	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**GAGCC**C**CAAAATTAAATAGA
adDMD-E23-I23	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TTACC**C**GAAATTTTCGAAGT

[0453] 用于在mdx小鼠的抗肌萎缩蛋白转录物中进行外显子23跳跃的MS2-adRNA序列(SEQ ID NO:160-161)：

[0454]

MS2-DMD-I22-E23	**aACATGAGGATCACCCATGTc**GAGCC**C**CAAAATTAAATAGA**aACATGAGGATCACCCATGTc**
MS2-DMD-E23-I23	**aACATGAGGATCACCCATGTc**TTACC**C**GAAATTTTCGAAGT**aACATGAGGATCACCCATGTc**

[0455] 用于编辑mdx小鼠中的提前终止密码子的MS2-adRNA序列(SEQ ID 162)：

[0456]

MS2_DMD	**aACATGAGGATCACCCATGTc**CCATT**CC**ATTGCTCTTTCAA**aACATGAGGATCACCCATGTc**

[0457] 用于在小鼠的SCN9A转录物中进行外显子跳跃的adRNA序列(SEQ ID NO:163-167)

[0458]

adRNA_SCN9A_I8_E9_F	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**ACTGCCCACAGATGAACAAG
adRNA_SCN9A_I4_E5_F	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**TGTACCCGAAGGAGAGAATA
adRNA_SCN9A_I2_E3_F	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**AAAGTCCGAGGAGGAAAAAG

[0459]

adRNA_SCN9A_I5_E6_F	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**AATACCCGTAGGATTAAATC
adRNA_SCN9A_I13_E14_F	**GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC**AAGTTCCGGAAAACAAACAA

[0460] 用于在小鼠的SCN9A转录物中进行外显子跳跃的MS2-adRNA序列(SEQ ID NO:168-172)

[0461]

MS2_adRNA_SCN9A_I8_E9_F	**aACATGAGGATCACCCATGTc**ACTGCCCACAGATGAACAAG**aACATGAGGATCACCCATGTc**
MS2_adRNA_SCN9A_I4_E5_F	**aACATGAGGATCACCCATGTc**TGTACCCGAAGGAGAGAATA**aACATGAGGATCACCCATGTc**
MS2_adRNA_SCN9A_I2_E3_F	**aACATGAGGATCACCCATGTc**AAAGTCCGAGGAGGAAAAAG**aACATGAGGATCACCCATGTc**
MS2_adRNA_SCN9A_I5_E6_F	**aACATGAGGATCACCCATGTc**AATACCCGTAGGATTAAATC**aACATGAGGATCACCCATGTc**
MS2_adRNA_SCN9A_I13_E14_F	**aACATGAGGATCACCCATGTc**AAGTTCCGGAAAACAAACAA**aACATGAGGATCACCCATGTc**

[0462] 实施例7-经由APOBEC进行的C→T编辑

[0463] 可以创建载脂蛋白B mRNA的局部结构,以将要编辑的C小心地定位在与载脂蛋白B mRNA中天然存在的位点类似的位置,其中ACF-APOBEC复合物在感兴趣的mRNA靶标处结合载脂蛋白B mRNA。这是通过过表达成对的gRNA,来创建载脂蛋白B mRNA样结构来实现的,该结构可以通过过表达以下一种或多种组合来编辑:

[0464] 1.MCP-ACF1和APOBEC1

[0465] 2.MCP-ACF1和MCP-APOBEC1

[0466] 3.MCP-接头-ACF-接头-APOBEC1

[0467] 4.MCP-APOBEC1和ACF1

[0468] 序列集2:示例性C→T编辑序列

[0469] 用于C-T编辑的MS2-adRNA序列(SEQ ID NO:173-174)

[0470]

MS2-gRNA-1	acatatatgataNNNNNNNNNNaACATGAGGATCACCCATGTcNNNNNNNNNNNNNNNNNNNN
MS2-gRNA-2	NNNNNNNNNNNNNNNNNNNNaACATGAGGATCACCCATGTcNNNNNNNNNNttgatcagtatatta

[0471] 实施例8-与癌症相关的点突变的创建

[0472] 癌症通路中涉及的几个基因携带单个氨基酸取代。通过在这些基因的mRNA序列中创建A->G取代,可以创建显性负突变体、组成活性突变体和催化失活突变体。这些基因中的一部分包括KRAS、HRAS、JAK2、GSK3β、β-连环蛋白(catenin)、SmoM2,含半胱氨酸的天冬氨酸蛋白水解酶(Caspase)3、Caspase 8、TGF-β、p53。

[0473] 实施例9-编辑DNA以及DNA/RNA杂交体的两条链

[0474] 由于已显示出ADAR可编辑双链RNA以及DNA-RNA杂合体的两条链,因此可以经由单链DNA或DNA-RNA杂合体来募集ADAR,以编辑DNA和RNA。这可用于修改目前针对Cas9(或Cpf1)-ADAR脱氨酶结构域融合体(ADAR1、ADAR2及其催化活性突变体E1008Q和E488Q)的腺嘌呤碱基编辑方法,而不是通过目前碱基编辑器靶向ssDNA置换链,可以靶向具有A-C凸起的gRNA结合链,理想地是在接近gRNA的5'端的第一个10bp中。

[0475] 实施例10-鸟氨酸转氨甲酰酶缺乏症

[0476] 在spf^{ash}小鼠中递送了ADAR2(E488Q)和adRNA。除了正确剪接的mRNA之外,spf^{ash}小鼠还携带经错误剪接的延长的mRNA变体,该变体是由于在外显子4中使用隐含剪接位点(48个碱基对)而形成。在递送了ADAR2(E488Q)与adRNA后,证实了错误剪接的产物减少。还观察到剪接的mRNA的高达33.9%的高效RNA编辑产量。在pre-mRNA中,观察到高达8%的产量。还观察到注射后3周内,蛋白恢复了2.5～5%。这证明了RNA编辑在剪接位点突变的校正和改变剪接模式中的效用。(图50A～图50D)

[0477] 实施例11-杜氏肌肉萎缩症

[0478] 在mdx小鼠中,与MS2-adRNA一起测试了MCP-ADAR1(E1008Q)。观察到,与ADAR2和ADAR2(E488Q)酶相比,增强的RNA编辑产量和抗肌萎缩蛋白表达的恢复。(图51A～图51B)

[0479] 等同物

[0480] 除非另有定义,否则本文所用的所有技术和科学术语均具有与该技术所属领域的普通技术人员通常所理解的相同的含义。

[0481] 本文说明性描述的技术可以在不存在本文未具体公开的任一个元件或多个元件、一种限制或多种限制的情况下适当地实践。因此,例如,术语"包括"、"包含"、"含有"等应被广义地且不受限制地理解。另外,本文所采用的术语和表达已被用作描述性术语,而非限制性术语,并且无意在使用这些术语和表达时排除所示出和描述的特征或其部分的任何等同形式,但是应当认识到,在所要求保护的技术范围内可以进行各种修改。

[0482] 因此,应该理解,这里提供的材料、方法和实施例是代表优选的方面,是示例性的,并且不意在限制技术范围。

[0483] 本文中已经广泛且概括地描述了该技术。落入一般公开范围内的每一个较窄的种类和次类分组也构成该技术的一部分。这包括对技术的一般性描述,但有条件限制或否定限制,以从种类物中删除任何主题,而不管是否在本文中特别叙述了被删除的材料。

[0484] 另外,在以马库什组来描述技术的多个特征或方面的情况下,本领域技术人员将认识到,该技术也由此以马库什组的任何单独的成员或成员的子群来描述。

[0485] 本文提及的所有出版物、专利申请、专利和其他参考文献都明确地通过引用整体并入,其程度就好像各自都单独通过引用并入。如有冲突,以说明书(包括定义)为准。

[0486] 具体实施方式

[0487] 本文公开了许多组合物、方法和系统。下面公开了这些组合物、方法和系统的具体的示例性实施方式。

[0488] 部分1

[0489] 实施方式1.一种经改造的ADAR1或ADAR2向导RNA("adRNA"),包括:与靶标RNA互补的序列。

[0490] 实施方式2.实施方式1所述的经改造的adRNA,进一步包括来源于GluR2mRNA的ADAR2募集结构域。

[0491] 实施方式3.实施方式1所述的经改造的adRNA,进一步包括两个在与靶标RNA互补的序列两侧的MS2发夹。

[0492] 实施方式4.实施方式1～3中任一项所述的经改造的adRNA,其中所述与靶标RNA互补的序列包括约20至100个碱基对。

[0493] 实施方式5.实施方式2或4所述的经改造的adRNA,其中来源于GluR2 mRNA的所述ADAR2募集结构域位于所述经改造的adRNA的5'末端或3'末端。

[0494] 实施方式6.实施方式5所述的经改造的adRNA,在所述经改造的adRNA的5'末端和3'末端均包含GluR2 mRNA。

[0495] 实施方式7.实施方式5或6所述的经改造的adRNA,进一步包括编辑诱导元件。

[0496] 实施方式8.上述实施方式中任一项所述的经改造的adRNA,其中所述靶标RNA是鸟氨酸转氨甲酰酶。

[0497] 实施方式9.一种经改造的ADAR2向导RNA("adRNA"),由选自表1或图2提供的序列组或其各自的等同物中的多核苷酸序列来编码。

[0498] 实施方式10.一种分离的多核苷酸,编码实施方式1～9中任一项所述的经改造的adRNA,或其各自的等同物。

[0499] 实施方式11.一种载体,包括一种或多种实施方式10所述的分离的多核苷酸,或编码实施方式9所述的经改造的adRNA的多核苷酸序列,和任选地可操作地与所述分离的多核苷酸连接的调控序列。

[0500] 实施方式12.实施方式11所述的载体,其中所述调控序列包括启动子、增强子元件和/或报告因子。

[0501] 实施方式13.实施方式12所述的载体,其中所述启动子为人U6、小鼠U6启动子或CMV启动子。

[0502] 实施方式14.实施方式11～13中任一项所述的载体,进一步包括可检测的标记物或纯化标记物。

[0503] 实施方式15.实施方式14所述的载体,其中所述载体是质粒或病毒载体。

[0504] 实施方式16.实施方式15所述的载体,其中所述载体选自由逆转录病毒载体、慢病毒载体、腺病毒载体和腺相关病毒载体组成的组。

[0505] 实施方式17.一种重组细胞,进一步包括实施方式11～16中任一项所述的载体,其中所述改造的adRNA被重组表达。

[0506] 实施方式18.一种组合物,包括载剂,和实施方式1～9中任一项所述的一种或多种经改造的adRNA,实施方式10所述的分离的多核苷酸,实施方式11～16中任一项所述的载体,或实施方式17所述的重组细胞。

[0507] 实施方式19.实施方式18所述的组合物,进一步包括化学治疗剂或药物。

[0508] 实施方式20.实施方式18或19所述的组合物,其中所述载体是药学上可接受的载剂或固体支持物。

[0509] 实施方式21.一种修饰蛋白表达的方法,包括使编码所述蛋白的多核苷酸与实施方式1～9中任一项的经改造的adRNA接触。

[0510] 实施方式22.实施方式21所述的方法,其中所述接触是在体外或在体内。

[0511] 实施方式23.一种治疗与异常蛋白表达相关的疾病或障碍的方法,包括向需要这种治疗的受试者,施用有效量的一种或多种实施方式1～9中任一项所述的经改造的adRNA。

[0512] 实施方式24.实施方式23所述的方法,其中所述疾病或障碍是杜氏肌肉萎缩症。

[0513] 实施方式25.实施方式23或24所述的方法,其中所述受试者是动物。

[0514] 实施方式26.实施方式23或24所述的方法,其中所述动物是哺乳动物。

[0515] 实施方式27.有效量的一种或多种实施方式1～9中任一项所述的经改造的adRNA在治疗与异常蛋白表达相关的疾病或障碍中的用途。

[0516] 实施方式28.实施方式的27所述的用途,其中所述疾病或障碍是杜氏肌肉萎缩症。

[0517] 实施方式29.一种试剂盒,包括:实施方式1～9中任一项所述的经改造的adRNA、实施方式10所述的分离的多核苷酸、实施方式11～16中任一项所述的载体、实施方式17所述的重组细胞或实施方式18～20中任一项所述的组合物,和使用说明书。

[0518] 实施方式30.实施方式19所述的试剂盒,其中所述说明书记载了实施方式21～26中任一项所述的方法。

[0519] 实施方式31.一种复合物,包括实施方式1～9中任一项所述的adRNA,所述adRNA在高严格性条件下与互补多核苷酸杂交。

[0520] 部分2

[0521] 实施方式1.一种经改造的ADAR2向导RNA(“adRNA”),包括:与靶标RNA互补的序列,和来源于GluR2 mRNA的ADAR2募集结构域。

[0522] 实施方式2.实施方式1所述的经改造的adRNA,其中所述与靶标RNA互补的序列包括约20至100个碱基对。

[0523] 实施方式3.实施方式1或实施方式2所述的经改造的adRNA,其中所述来源于GluR2 mRNA的ADAR2募集结构域位于所述经改造的adRNA的5’末端或3’末端。

[0524] 实施方式4.实施方式3所述的经改造的adRNA,在所述经改造的adRNA的5’末端和3’末端均包括GluR2 mRNA。

[0525] 实施方式5.前述实施方式中任一项所述的经改造的adRNA,其中所述靶标RNA是鸟氨酸转氨甲酰酶。

[0526] 实施方式6.一种经改造的ADAR2向导RNA(“adRNA”),由选自表1或图2中提供的序列组中的序列编码。

[0527] 实施方式7.一种修饰蛋白表达的方法,包括使编码所述蛋白的多核苷酸与实施方式1～6中任一项所述的经改造的adRNA接触。

[0528] 实施方式8.实施方式7所述的方法,其中所述接触是在体外或在体内。

[0529] 实施方式9.一种治疗与异常蛋白表达相关的疾病或障碍的方法,包括对需要这种治疗的受试者施用有效量的一种或多种实施方式1至6中任一项所述的经改造的adRNA。

[0530] 实施方式10.实施方式9所述的方法,其中所述受试者是动物。

[0531] 实施方式11.实施方式9所述的方法,其中所述动物是哺乳动物。

[0532] 实施方式12.有效量的一种或多种实施方式1至6中任一项所述的经改造的adRNA在治疗与异常蛋白表达相关的疾病或障碍中的用途。

[0533] 实施方式13.一种试剂盒,包括实施方式1至6中任一项所述的经改造的adRNA和使

用说明书。

[0534] 实施方式14.实施方式13所述的试剂盒,其中所述说明书记载了实施方式7或实施方式9所述的方法。

[0535] 实施方式15.一种组合物,包括实施方式1～6中任一项所述的AdRNA和载剂。

[0536] 实施方式16.实施方式15所述的组合物,其中所述载体是药学上可接受的载剂。

[0537] 实施方式17.一种复合物,包括实施方式1～6中任一项所述的AdRNA,所述AdRNA在高严格性条件下与互补多核苷酸杂交。

[0538] 部分3

[0539] 实施方式1.一种用于外显子跳跃的ADAR系统,包括靶向内含子中的剪接受体和/或分支点的adRNA,以及任选地ADAR酶。

[0540] 实施方式2.权利要求1所述的ADAR系统,其中所述ADAR酶是ADAR1、ADAR2或其各自的突变体或变体。

[0541] 实施方式3.权利要求2所述的ADAR系统,其中所述突变体或变体选自ADAR1(E1008Q)和ADAR2(E488Q)。

[0542] 实施方式4.权利要求1至3中任一项所述的ADAR系统,其中所述内含子包含在选自抗肌萎缩蛋白、SCN9A或鸟氨酸转氨甲酰酶的基因中。

[0543] 实施方式5.权利要求1至4中任一项所述的ADAR系统,其中所述adRNA选自序列集1。

[0544] 实施方式6.一种治疗以异常基因表达为特征的疾病、障碍或病症的方法,包括施用权利要求1至5中任一项所述的ADAR系统。

[0545] 实施方式7.权利要求6所述的方法,其中所述疾病、障碍或病症选自杜氏肌肉萎缩症或鸟氨酸转氨甲酰酶缺乏症。

[0546] 实施方式8.权利要求6或权利要求7所述的方法,其中所述疾病、障碍或病症与疼痛有关。

[0547] 实施方式9.一种用于胞嘧啶至胸腺嘧啶编辑的APOBEC系统,包括创建脂蛋白B mRNA样结构成对的gRNA,和任选地APOBEC酶。

[0548] 实施方式10.根据权利要求9所述的APOBEC系统,其中所述成对的gRNA是在序列集2中提供的成对的序列。

[0549] 参考文献

[0550] Montiel-Gonzalez,M.F.,Vallecillo-Viejo,I.,Yudowski,G.A.&Rosenthal,J.J.C.Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing.Proc.Natl.Acad.Sci.U.S.A.110,18285-90(2013).

[0551] Wettengel,J.,Reautschnig,P.,Geisler,S.,Kahle,P.J.&Stafforst,T.Harnessing human ADAR2 for RNA repair-Recoding a PINK1 mutation rescues mitophagy.Nucleic Acids Res.45,2797-2808(2017).

[0552] Fukuda,M.et al.Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing.Sci.Rep.7,41478(2017).

[0553] Heep,M.,Mach,P.,Reautschnig,P.,Wettengel,J.&Stafforst,T.Applying

Human ADAR1p110 and ADAR1p150 for Site-Directed RNA Editing—G/C Substitution Stabilizes GuideRNAs against Editing.Genes(Basel).8,34(2017).

[0554] Vallecillo-Viejo,I.C.,Liscovitch-Brauer,N.,Montiel-Gonzalez,M.F.,Eisenberg,E.&Rosenthal,J.J.C.Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme.RNA Biol.15,104-114(2018).

[0555] Sinnamon,J.R.et al.Site-directed RNA repair of endogenous Mecp2 RNA in neurons.Proc.Natl.Acad.Sci.U.S.A.114,E9395-E9402(2017).

[0556] Cox,D.B.T.et al.RNA editing with CRISPR-Cas13.Science eaaq0180(2017).doi:10.1126/science.aaq0180

[0557] Montiel-González,M.F.,Vallecillo-Viejo,I.C.&Rosenthal,J.J.C.An efficient system for selectively altering genetic information within mRNAs.Nucleic Acids Res.44,gkw738(2016).

[0558] Hanswillemenke,A.,Kuzdere,T.,Vogel,P.,Jeely,G.&Stafforst,T.Site-Directed RNA Editing in vivo Can Be Triggered by the Light-Driven Assembly of an Artificial Riboprotein.doi:10.1021/jacs.5b10216

[0559] Woolf,T.M.,Chase,J.M.&Stinchcomb,D.T.Toward the therapeutic editing of mutated RNA sequences.Biochemistry 92,8298-8302(1995).

[0560] Stafforst,T.&Schneider,M.F.An RNA-Deaminase Conjugate Selectively Repairs Point Mutations.Angew.Chemie Int.Ed.51,11166-11169(2012).

[0561] Schneider,M.F.,Wettengel,J.,Hoffmann,P.C.&Stafforst,T.Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans.Nucleic Acids Res.42,e87(2014).

[0562] Azad,M.T.A.,Bhakta,S.&Tsukahara,T.Site-directed RNA editing by adenosine deaminase acting on RNA for correction of the genetic code in gene therapy.Gene Ther.24,779-786(2017).

[0563] Vogel,P.et al.Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs.Nat.Methods 15,535-538(2018).

[0564] Smalley,E.First AAV gene therapy poised for landmark approval.Nat.Biotechnol.2017 3511(2017).

[0565] Calcedo,R.&Wilson,J.M.AAV Natural Infection Induces Broad Cross-Neutralizing Antibody Responses to Multiple AAV Serotypes in Chimpanzees.Hum.Gene Ther.Clin.Dev.27,79-82(2016).

[0566] Harbison,C.E.et al.Examining the cross-reactivity and neutralization mechanisms of a panel of mabs against adeno-associated virus serotypes 1 and 5.J.Gen.Virol.93,(2012).

[0567] Riviere,C.,Danos,O.&Douar,A.M.Long-term expression and repeated administration of AAV type 1,2 and 5 vectors in skeletal muscle of immunocompetent adult mice.Gene Ther.13,1300-1308(2006).

[0568] Mays,L.E.&Wilson,J.M.The Complex and Evolving Story of T cell Activation to AAV Vector-encoded Transgene Products.Mol.Ther.19,16-27(2011).

[0569] Chen,J.,Wu,Q.,Yang,P.,Hsu,H.C.&Mountz,J.D.Determination of specific CD4 and CD8 T cell epitopes after AAV2-and AAV8-hF.IX gene therapy.Mol Ther 13,260-269(2006).

[0570] Ertl,H.C.J.&High,K.A.Impact of AAV Capsid-Specific T-Cell Responses on Design and Outcome of Clinical Gene Transfer Trials with Recombinant Adeno-Associated Viral Vectors:An Evolving Controversy.Hum.Gene Ther.28,328-337(2017).

[0571] Hinderer,C.et al.Severe toxicity in nonhuman primates and piglets following high-dose intravenous administration of an AAV vector expressing human SMN.Hum.Gene Ther.hum.2018.015(2018).doi:10.1089/hum.2018.015

[0572] Mingozzi,F.et al.Prevalence and pharmacological modulation of humoral immunity to AAV vectors in gene transfer to synovial tissue.Gene Ther 20,417-424(2013).

[0573] Basner-Tschakarjan,E.,Bijjiga,E.&Martino,A.T.Pre-clinical assessment of immune responses to adeno-associated virus(AAV)vectors.Frontiers in Immunology 5,(2014).

[0574] Harding,F.A.,Stickler,M.M.,Razo,J.&DuBridge,R.B.The immunogenicity of humanized and fully human antibodies:Residual immunogenicity resides in the CDR regions.MAbs 2,256-265(2010).

[0575] De Groot,a S.,Knopp,P.M.&Martin,W.De-immunization of therapeutic proteins by T-cell epitope modification.Dev.Biol.(Basel).122,171-194(2005).

[0576] Ferdosi,S.R.et al.Multifunctional CRISPR/Cas9 with engineered immunosilenced human T cell epitopes.bioRxiv(2018).

[0577] Veronese,F.M.&Mero,A.The impact of PEGylation on biological therapies.BioDrugs 22,315-329(2008).

[0578] Armstrong,J.K.et al.Antibody against poly(ethylene glycol)adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients.Cancer 110,103-111(2007).

[0579] Kurosaki,T.,Kometani,K.&Ise,W.Memory B cells.Nat.Rev.Immunol.15,149-159(2015).

[0580] Zabel,F.et al.Distinct T helper cell dependence of memory B-cell proliferation versus plasma cell differentiation.Immunology 150,329-342(2017).

[0581] Chylinski,K.,Makarova,K.S.,Charpentier,E.&Koonin,E.V.Classification and evolution of type II CRISPR-Cas systems.Nucleic Acids Research 42,6091-6105(2014).

[0582] Burstein,D.et al.New CRISPR-Cas systems from uncultivated

microbes.Nat.Publ.Gr.542,(2017).

[0583] Makarova,K.S.,Zhang,F.&Koonin,E.V.SnapShot:Class 2 CRISPR-Cas Systems.Cell 168,328-328.e1(2017).

[0584] Ellis,B.L.et al.A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus(AAV1-9)and one engineered adeno-associated virus serotype.Virol.J.10,74 (2013).

[0585] Andreatta,M.&Nielsen,M.Gapped sequence alignment using artificial neural networks:application to the MHC class I system.Bioinformatics 32,511-517(2015).

[0586] Andreatta,M.et al.Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification.Immunogenetics 67,641-650(2015).

[0587] Truong,D.-J.J.et al.Development of an intein-mediated split-Cas9 system for gene therapy.Nucleic Acids Res.43,6450-6458(2015).

[0588] Moreno,A.M.et al.In Situ Gene Therapy via AAV-CRISPR-Cas9-Mediated Targeted Gene Regulation.Mol.Ther.0,(2018).

[0589] Benveniste,O.et al.Prevalence of Serum IgG and Neutralizing Factors Against Adeno-Associated Virus(AAV)Types 1,2,5,6,8,and 9 in the Healthy Population:Implications for Gene Therapy Using AAV Vectors.Hum.Gene Ther.21, 704-712(2010).

[0590] Hsu,P.D.,Lander,E.S.&Zhang,F.Development and applications of CRISPR-Cas9 for genome engineering.Cell 157,1262-1278(2014).

[0591] Ran,F.A.et al.In vivo genome editing using Staphylococcus aureus Cas9.Nature 520,186-190(2015).

[0592] Clemente,T.,Dominguez,M.R.,Vieira,N.J.,Rodrigues,M.M.&Amarante-Mendes,G.P.In vivo assessment of specific cytotoxic T lymphocyte killing.Methods 61,105-109(2013).

[0593] Weinmann,J.&Grimm,D.Next-generation AAV vectors for clinical use:an ever-accelerating race.Virus Genes 53,707-713(2017).

[0594] Gabriel,N.et al.Bioengineering of AAV2 Capsid at Specific Serine, Threonine,or Lysine Residues Improves Its Transduction Efficiency in vitro and in vivo.Hum.Gene Ther.Methods 24,80-93(2013).

[0595] Zhong,L.et al.Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses.Proc.Natl.Acad.Sci.105,7827-7832(2008).

[0596] Shen,S.et al.Engraftment of a Galactose Receptor Footprint onto Adeno-associated Viral Capsids Improves Transduction Efficiency.J.Biol.Chem.288,28814-28823(2013).

[0597] Zinn,E.et al.In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector.Cell Rep.12,1056-1068(2017).

[0598] Deverman,B.E.et al.Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain.Nat.Biotechnol.34,204-209(2016).

[0599] Smith,J.K.&Agbandje-McKenna,M.Creating an arsenal of Adeno-associated virus(AAV)gene delivery stealth vehicles.PLOS Pathog.14,e1006929(2018).

[0600] Bartel,M.,Schaffer,D.&Büning,H.Enhancing the clinical potential of AAV vectors by capsid engineering to evade pre-existing immunity.Front.Microbiol.2,204(2011).

[0601] Hui,D.J.et al.AAV capsid CD8+T-cell epitopes are highly conserved across AAV serotypes.Mol.Ther.-Methods Clin.Dev.2,15029(2015).

[0602] Gao,G.,Vandenberghe,L.&Wilson,J.New Recombinant Serotypes of AAV Vectors.Curr.Gene Ther.5,285-297(2005).

[0603] Grieger,J.C.,Choi,V.W.&Samulski,R.J.Production and characterization of adeno-associated viral vectors.Nat.Protoc.1,1412-1428(2006).

[0604] Mingozzi,F.&Büning,H.Adeno-Associated Viral Vectors at the Frontier between Tolerance and Immunity.Front.Immunol.6,120(2015).

[0605] Mingozzi,F.&High,K.A.Therapeutic in vivo gene transfer for genetic disease using AAV:progress and challenges.Nat.Rev.Genet.12,341-355(2011).

[0606] Wagner,D.L.et al.High prevalence of S.pyogenes Cas9-specific T cell sensitization within the adult human population-A balanced effector/regulatory T cell response.bioRxiv 295139(2018).doi:10.1101/295139

[0607] Gernoux,G.,Wilson,J.M.&Mueller,C.Regulatory and Exhausted T Cell Responses to AAV Capsid.Hum.Gene Ther.28,338-349(2017).

[0608] Zhu,J.,Huang,X.&Yang,Y.The TLR9-MyD88 pathway is critical for adaptive immune responses to adeno-associated virus gene therapy vectors in mice.J.Clin.Invest.119,2388-2398(2009).

[0609] Rogers,G.L.&Herzog,R.W.TLR9 and Dendritic Cells Are Required for CD8+ T Cell Responses to the AAV Capsid.Blood 124,(2014).

[0610] Faust,S.M.et al.CpG-depleted adeno-associated virus vectors evade immune detection.J.Clin.Invest.123,2994-3001(2013).

[0611] Hirsch,M.&Wu,J.319.Rationally Designed AAV Inverted Terminal Repeats Enhance Gene Targeting.Mol.Ther.24,S129(2016).

[0612] Léger,A.et al.Adeno-Associated Viral Vector-Mediated Transgene Expression Is Independent of DNA Methylation in Primate Liver and Skeletal Muscle.PLoS One 6,e20881(2011).

[0613] Deaton,A.M.&Bird,A.CpG islands and the regulation of transcription.Genes Dev.25,1010-22(2011).

[0614] Shao,W.et al.Double-stranded RNA innate immune response activation

from long-term adeno-associated virus vector transduction.JCI Insight 3,(2018).

[0615] Peisley,A.et al.Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition.Proc.Natl.Acad.Sci.U.S.A.108,21010-5(2011).

[0616] Borel,F.,Kay,M.A.&Mueller,C.Recombinant AAV as a platform for translating the therapeutic potential of RNA interference.Mol.Ther.22,692-701(2014).

[0617] Xie,J.,Burt,D.R.&Gao,G.Adeno-associated virus-mediated microRNA delivery and therapeutics.Semin.Liver Dis.35,81-8(2015).

[0618] Li,J.,Kim,S.G.&Blenis,J.Rapamycin:one drug,many effects.Cell Metab.19,373-9(2014).

[0619] Meliani,A.et al.Antigen-selective modulation of AAV immunogenicity with tolerogenic rapamycin nanoparticles enables successful vector re-administration.Nat.Commun.9,4098(2018).

[0620] Battaglia,M.et al.Rapamycin promotes expansion of functional CD4+CD25+FOXP3+regulatory T cells of both healthy subjects and type 1 diabetic patients.J.Immunol.177,8338-47(2006).

[0621] Lee,E.J.,Guenther,C.M.&Suh,J.Adeno-associated virus(AAV)vectors:Rational design strategies for capsid engineering.Curr.Opin.Biomed.Eng.7,58-63(2018).

[0622] Xie,Q.et al.The atomic structure of adeno-associated virus(AAV-2),a vector for human gene therapy.Proc.Natl.Acad.Sci.U.S.A.99,10405-10(2002).

[0623] Walters,R.W.et al.Structure of Adeno-Associated Virus Serotype 5.J.Virol.78,3361-3371(2004).

[0624] Kotterman,M.A.,Chalberg,T.W.&Schaffer,D.V.Viral Vectors for Gene Therapy:Translational and Clinical Outlook.Annu.Rev.Biomed.Eng.17,63-89(2015).

[0625] Grimm,D.et al.In vitro and In vivo Gene Therapy Vector Evolution via Multispecies Interbreeding and Retargeting of Adeno-Associated Viruses.J.Virol.82,5887-5911(2008).

[0626] Chew,W.L.et al.A multifunctional AAV-CRISPR-Cas9 and its host response.Nat.Methods 13,868-874(2016).

[0627] Katrekar,D.et al.In vivo RNAediting of point mutations via RNA-guided adenosine deaminases.Nat.Methods(2019).doi:10.1038/s41592-019-0323-0

[0628] Kim,D.D.Y.et al.Widespread RNA editing of embedded alu elements in the human transcriptome.Genome Res.14,1719-25(2004).

[0629] Herbert,A.Z-DNA and Z-RNA in human disease.Commun.Biol.2,7(2019).

[0630] Koeris,M.,Funke,L.,Shrestha,J.,Rich,A.&Maas,S.Modulation of ADAR1

editing activity by Z-RNA in vitro.Nucleic Acids Res.33,5362-70(2005).

[0631] Yang,J.-H.et al.Intracellular localization of differentially regulated RNA-specific adenosine deaminase isoforms in inflammation.J.Biol.Chem.278,45833-42(2003).

[0632] Abudayyeh,O.O.et al.A cytosine deaminase for programmable single-base RNAediting.Science 365,382-386(2019).

[0633] Point Mutation-Definition,Types and Examples|Biology Dictionary.

[0634] Lueck,J.D.et al.Engineered transfer RNAs for suppression of premature termination codons.Nat.Commun.10,822(2019).

[0635] Jonsson,T.et al.A mutation in APP protects against Alzheimer's disease and age-related cognitive decline.Nature 488,96-99(2012).

[0636] Sun,J.&Roy,S.The physical approximation of APP and BACE-1:A key event in alzheimer's disease pathogenesis.Dev.Neurobiol.78,340-347(2018).

[0637] Sun,J.et al.CRISPR/Cas9 editing of APP C-terminus attenuatesβ-cleavage and promotesα-cleavage.Nat.Commun.10,53(2019).

[0638] Melcher,T.et al.A mammalian RNA editing enzyme.Nature 379,460-464(1996).

[0639] Montiel-Gonzalez,M.F.,Vallecillo-Viejo,I.,Yudowski,G.A.&Rosenthal,J.J.C.Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing.Proc.Natl.Acad.Sci.U.S.A.110,18285-90(2013).

[0640] Wettengel,J.,Reautschnig,P.,Geisler,S.,Kahle,P.J.&Stafforst,T.Harnessing human ADAR2 for RNA repair-Recoding a PINK1 mutation rescues mitophagy.Nucleic Acids Res.45,2797-2808(2017).

[0641] Fukuda,M.et al.Construction of a guide-RNA for site-directed RNA mutagenesis utilising intracellular A-to-I RNA editing.Sci.Rep.7,41478(2017).

[0642] Vallecillo-Viejo,I.C.,Liscovitch-Brauer,N.,Montiel-Gonzalez,M.F.,Eisenberg,E.&Rosenthal,J.J.C.Abundant off-target edits from site-directed RNA editing can be reduced by nuclear localization of the editing enzyme.RNA Biol.15,104-114(2018).

[0643] Sinnamon,J.R.et al.Site-directed RNA repair of endogenous Mecp2 RNA in neurons.Proc.Natl.Acad.Sci.U.S.A.114,E9395-E9402(2017).

[0644] Cox,D.B.T.et al.RNA editing with CRISPR-Cas13.Science 358,1019-1027(2017).

[0645] Azad,M.T.A.,Bhakta,S.&Tsukahara,T.Site-directed RNA editing by adenosine deaminase acting on RNA for correction of the genetic code in gene therapy.Gene Ther.24,779-786(2017).

[0646] Vogel,P.et al.Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs.Nat.Methods 15,535-538(2018).

[0647] Daniel,C.,Widmark,A.,Rigardt,D.&Öhman,M.Editing inducer elements increases A-to-I editing efficiency in the mammalian transcriptome.Genome Biol.18,195(2017).

[0648] Woolf,T.M.,Chase J.M.,&Stinchcomn D.T.Toward the therapeutic editing of mutated RNA sequences.Proc.Natl.Acad.Sci.USA 92,8298-8302(1995).

[0649] Chung,H.et al.Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown.Cell 172,811-824.e14(2018).

[0650] Desterro,J.M.P.et al.Dynamic association of RNA-editing enzymes with the nucleolus.J.Cell Sci.116,1805-1818(2003).

[0651] Mort,M.,Ivanov,D.,Cooper,D.N.&Chuzhanova,N.A.A meta-analysis of nonsense mutations causing human genetic disease.Hum.Mutat.29,1037-1047(2008).

[0652] Nelson,C.E.et al.In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy.Science 351,403-7(2016).

[0653] Tabebordbar,M.et al.In vivo gene editing in dystrophic mouse muscle and muscle stem cells.Science(80-.).351,407-411(2016).

[0654] Long,C.et al.Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy.Science(80-.).351,400-403(2016).

[0655] Hodges,P.E.&Rosenberg,L.E.The spf^{ash} mouse:a missense mutation in the ornithine transcarbamylase gene also causes aberrant mRNA splicing.Proc.Natl.Acad.Sci.U.S.A.86,4142-4146(1989).

[0656] Yang,Y.et al.A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice.Nat.Biotechnol.34,334-338(2016).

[0657] Chen,L.et al.Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma.Nat.Med.19,209-216(2013).

[0658] Mann CJ,Honeyman K,Cheng AJ,et al.Antisense-induced exon skipping and synthesis of dystrophin in the mdx mouse.Proceedings of the National Academy of Sciences of the United States of America.2001;98(1):42-47.

[0659] S.Konermann,P.Lotfy,N.J.Brideau,J.Oki,M.N.Shokhirev,P.D.Hsu,Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors.Cell 173,665-676.e14(2018).

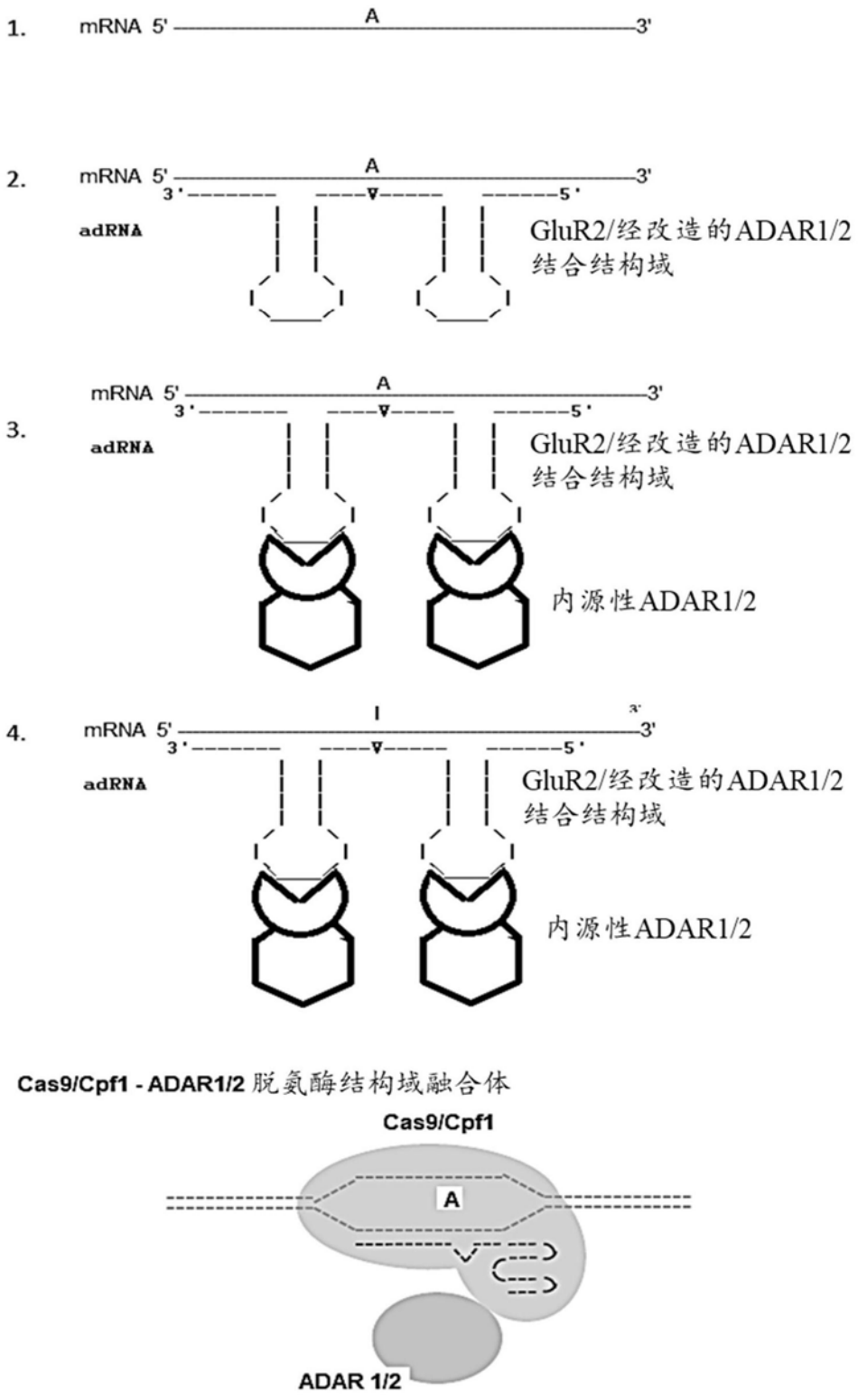
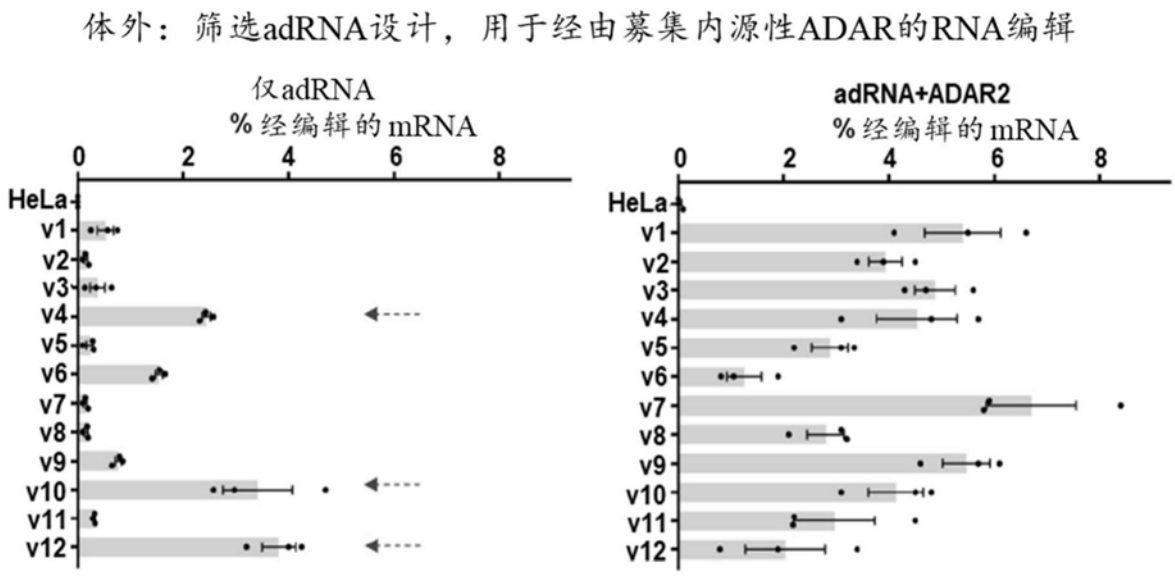

ADAR的内源性募集
步骤：
1.
mRNA 5'
A
3'
2.
mRNA 5'
A
3'
3'
5'
adRNA
GluR2/经改造的ADAR1/2
结合结构域
3.
mRNA 5'
A
3'
3'
5'
adRNA
GluR2/经改造的ADAR1/2
结合结构域
内源性ADAR1/2
4.
mRNA 5'
3'
3'
5'
adRNA
GluR2/经改造的ADAR1/2
结合结构域
内源性ADAR1/2
Cas9/Cpf1 - ADAR1/2 脱氨酶结构域融合体
Cas9/Cpf1
A
ADAR 1/2

图1

```
v3 :   GTGGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCCAC NNNNNNCNNNNNNNNNNNNNN
v4 : GGGTGGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCCACCT NNNCNNNNNNNNNNNNNNN
v5 :   GGTGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCACC NNNNNNCNNNNNNNNNNNNNN
v6 :   GGTGAAGAGGAGAACAATATGCTAAATGTTGTTCTCGTCTCCACC NNNNNNNCNNNNNNNNNNNNN
v7 :   GTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCAC NNNNNNCNNNNNNNNNNNNNN
v8 : GGGTGGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCCACCT NNNCNNNNNNNNNNNNNNN
v9 :   GGTGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCACC NNNNNNCNNNNNNNNNNNNNN
v10:   GGTGAAGAGGAGAACAATAGGCTAAACGTTGTTCTCGTCTCCACC NNNNNNNCNNNNNNNNNNNNN
```

图2



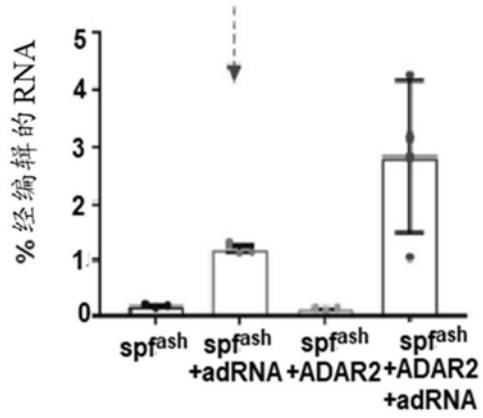


图3

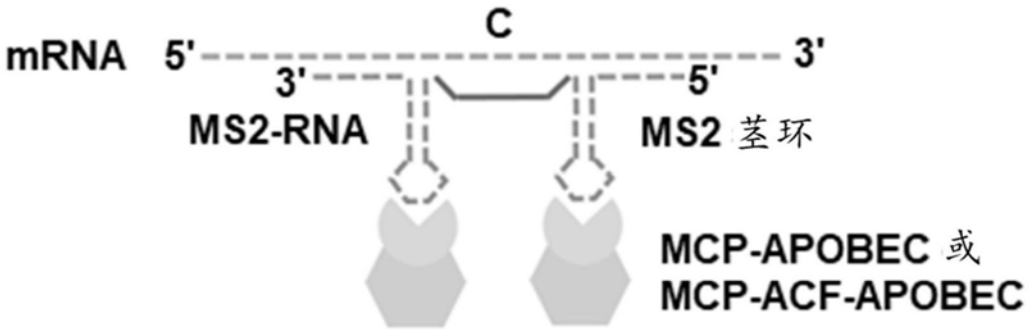

C
mRNA 5'
3'
3'
5'
MS2-RNA
MS2 茎环
MCP-APOBEC 或
MCP-ACF-APOBEC

图4

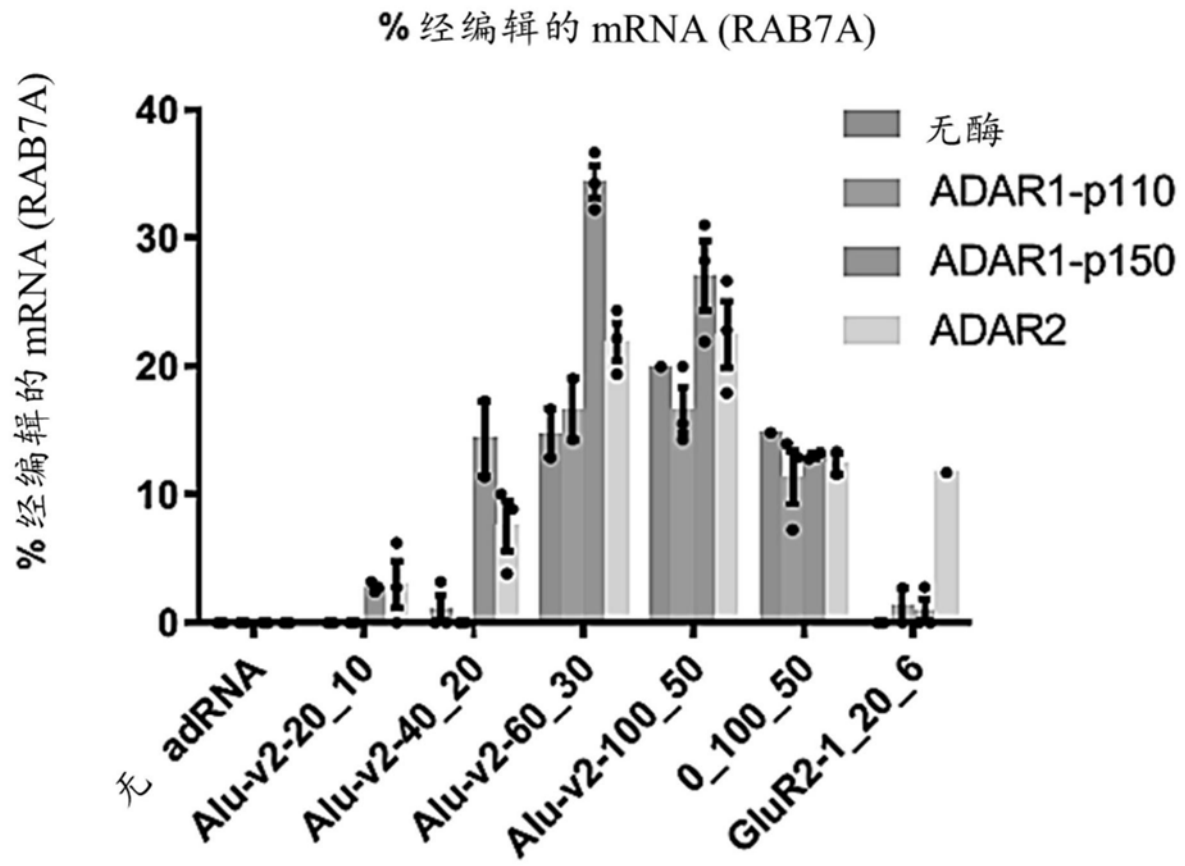

% 经编辑的 mRNA (RAB7A)
% 经编辑的 mRNA (RAB7A)
40
30
20
10
0
无酶
ADAR1-p110
ADAR1-p150
ADAR2
无 adRNA
Alu-v2-20_10
Alu-v2-40_20
Alu-v2-60_30
Alu-v2-100_50
0_100_50
GluR2-1_20_6

图5

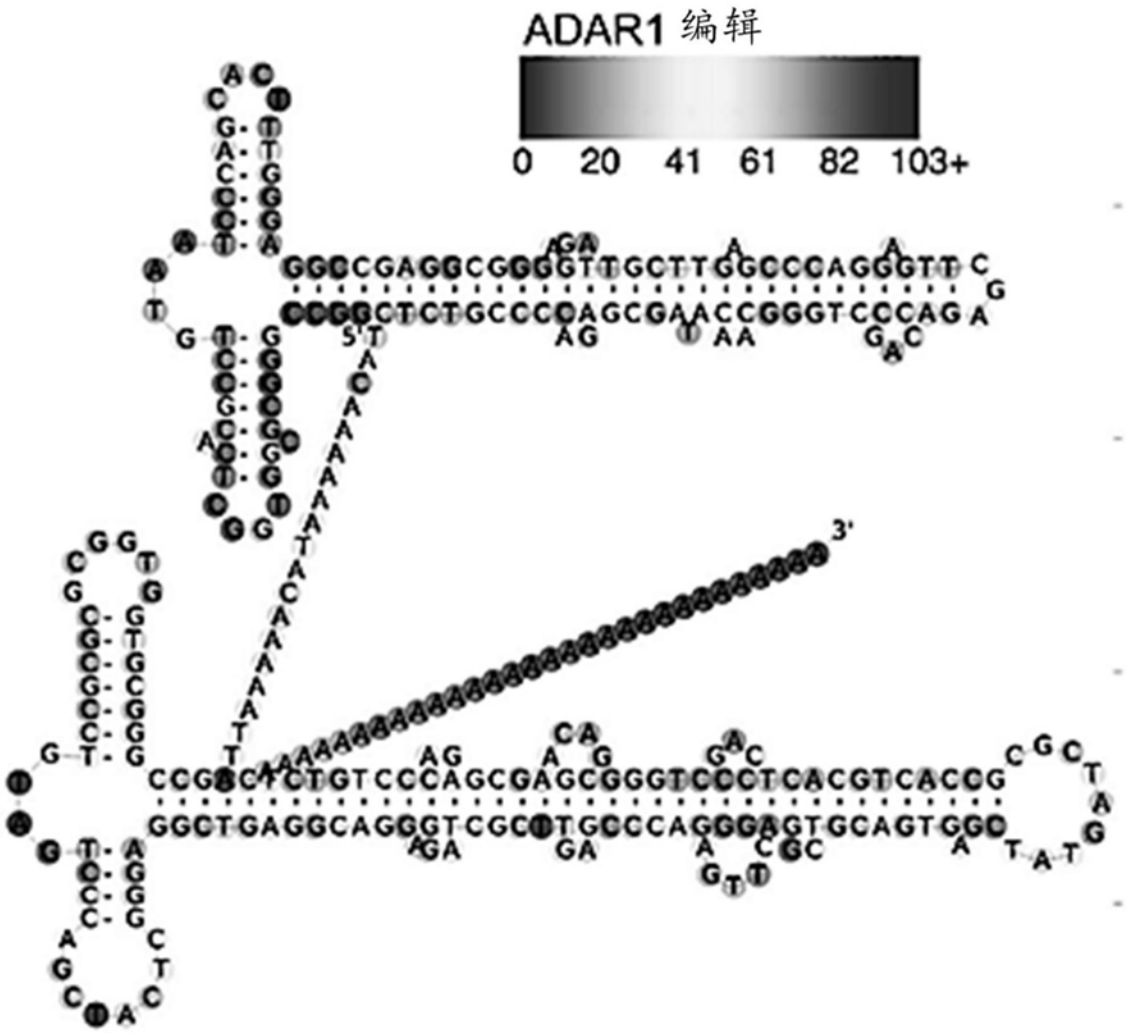
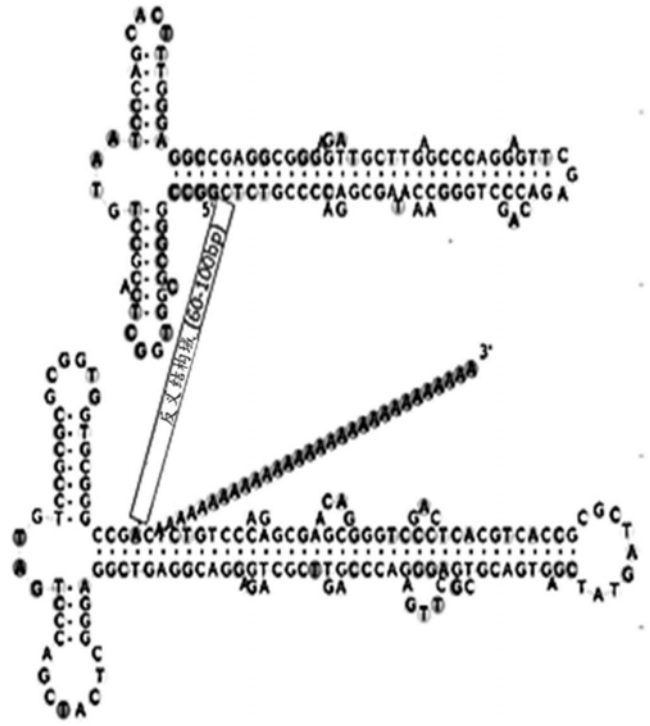


图6

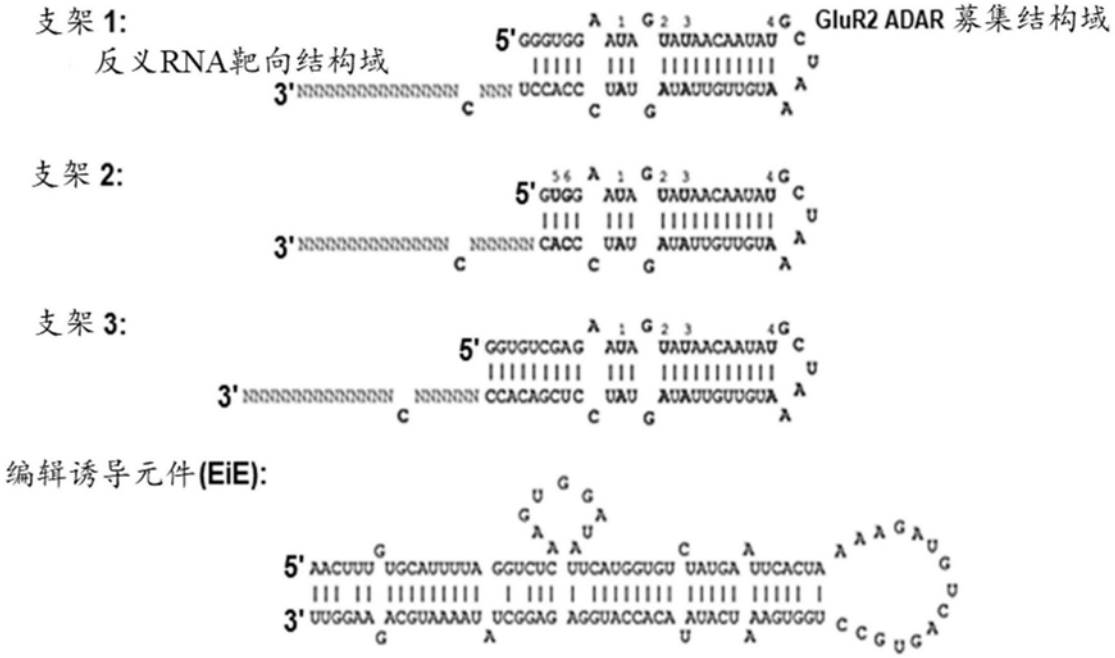


SEQ ID NO: 1

实施例:5'--

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTG
GGAGGCCGAGGCGGGGAGATTGCTTGAGCCCAGGAGTTC
GAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnAGCCGGGCGTGGTGGCGCGCGCCTGTAGT
CCCAGCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAG
CCCAGGAGTTCGAGGCTGCAGTGAGCTATGATCGCGCCACT
GCACTCCAGCCTGGGCGACAGAGCGAGACCCTGTCTC -- 3'

图7



根据靶标:3′-TTCTTTTCAAAATGTTTGACTCGCCACAGACACTCTGAAAGTAAGT-5′

1.1	1:U-A	2:U-A	3:U-A	4:U-A		AACCGAGCGGTGTCTGTGA
1.2	1:G-C	2:G-C	3:G-C	4:U-A		AACCGAGCGGTGTCTGTGA
1.3	1:G-C	2:G-C	3:G-C	4:G-C		AACCGAGCGGTGTCTGTGA
2.1	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG
2.2	1:G-C	2:G-C	3:G-C	4:U-A	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG
2.3	1:G-C	2:G-C	3:G-C	4:G-C	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG
2.4	1:G-C	2:G-C	3:G-C	4:U-A	5:G-C 6:U-A	ACAAACCGAGCGGTGTCTGTG
2.5	1:G-C	2:G-C	3:G-C	4:U-A	5:G-C 6:U-A	TACAAACCGAGCGGTGTCTGT
2.6	1:G-C	2:G-C	3:G-C	4:G-C	5:G-C 6:U-A	ACAAACCGAGCGGTGTCTGTG
2.7	1:G-C	2:G-C	3:G-C	4:G-C	5:G-C 6:U-A	TACAAACCGAGCGGTGTCTGT
3.1		1:U-A	2:U-A	3:U-A	4:U-A	TACAAACCGAGCGGTGTC
3.2		1:G-C	2:G-C	3:G-C	4:U-A	TACAAACCGAGCGGTGTC
3.2		1:G-C	2:G-C	3:G-C	4:G-C	TACAAACCGAGCGGTGTC
4.1	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	CAAACCGAGCGGTGTCTGTG
4.2	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	TACAAACCGAGCGGTGTCTG
4.3	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	TTTACAAACCGAGCGGTGTC
4.4	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG
4.5	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG
4.6	1:U-A	2:U-A	3:U-A	4:U-A	5:U-A 6:G-C	ACAAACCGAGCGGTGTCTGTG

图8

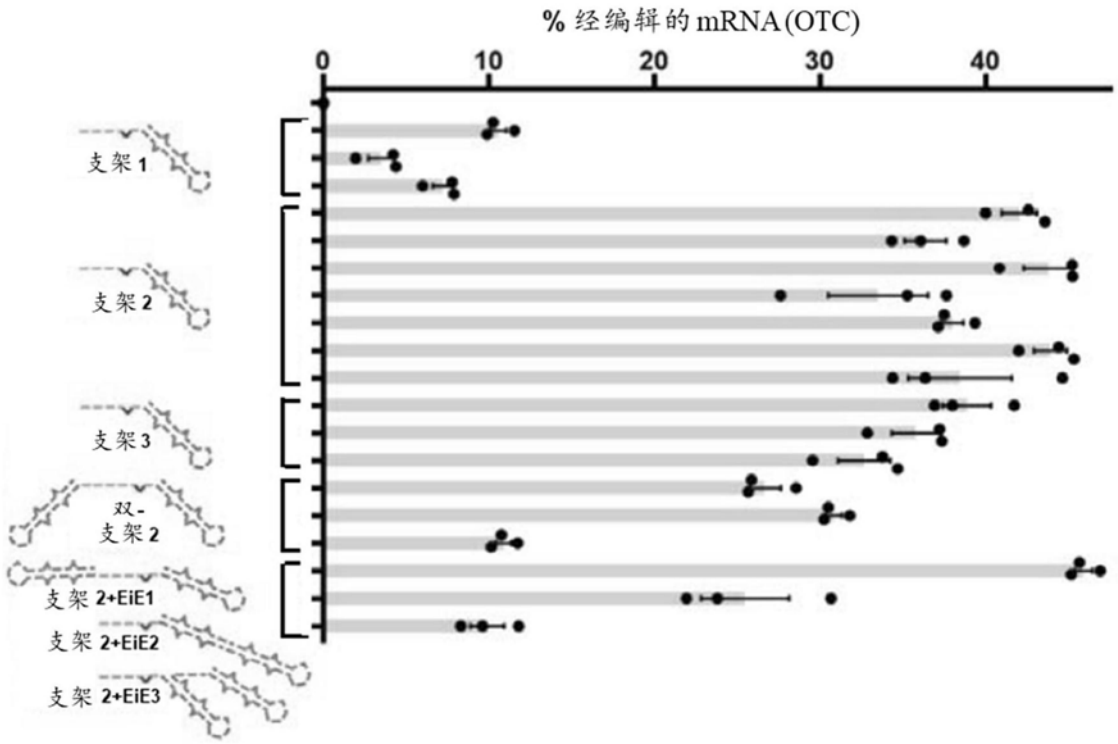

% 经编辑的 mRNA (OTC)
0
10
20
30
40
支架 1
支架 2
支架 3
双-
支架 2
支架 2+EiE1
支架 2+EiE2
支架 2+EiE3

图8(续)

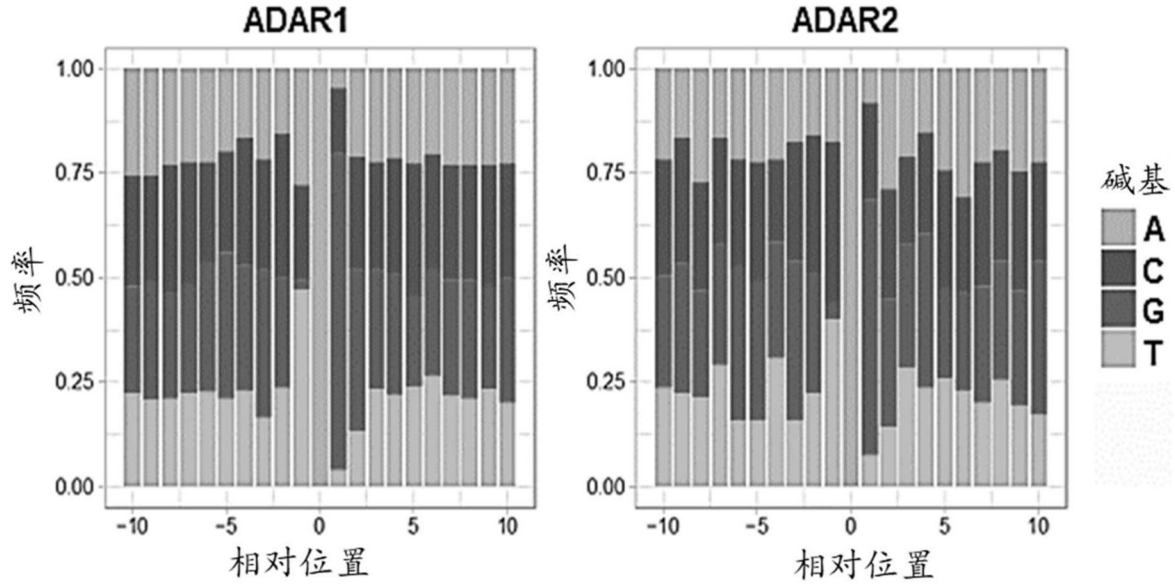

ADAR1
ADAR2
1.00
0.75
0.50
0.25
0.00
频率
-10
-5
0
5
10
相对位置
碱基
A
C
G
T

图9

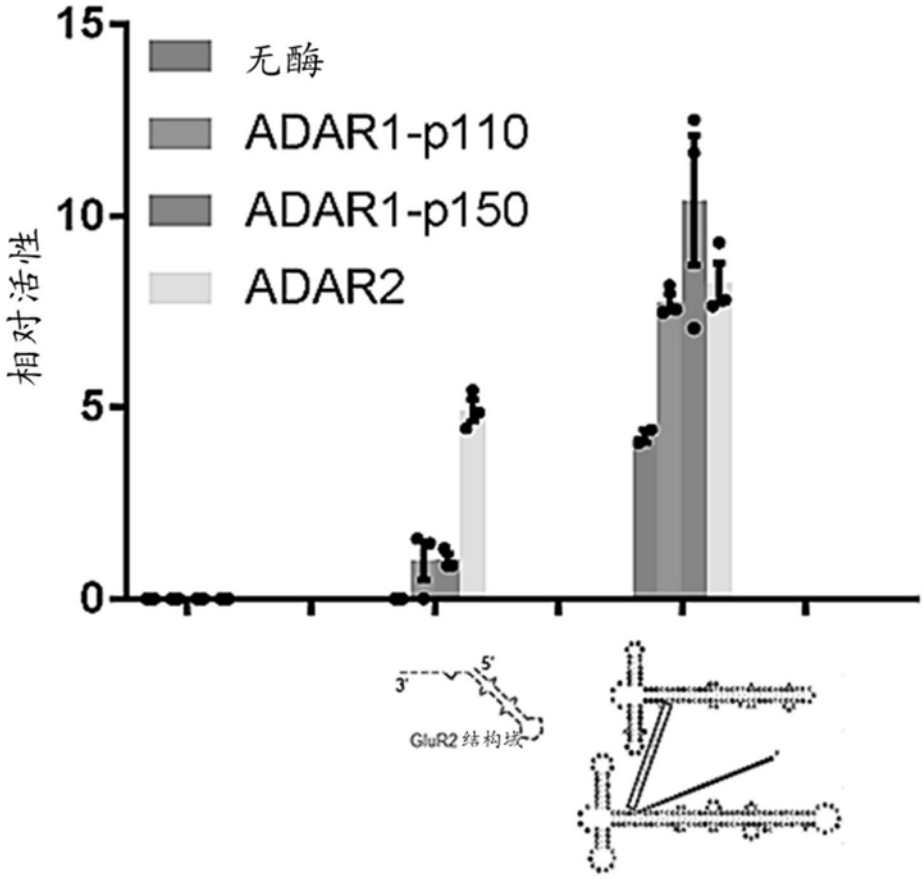


图10

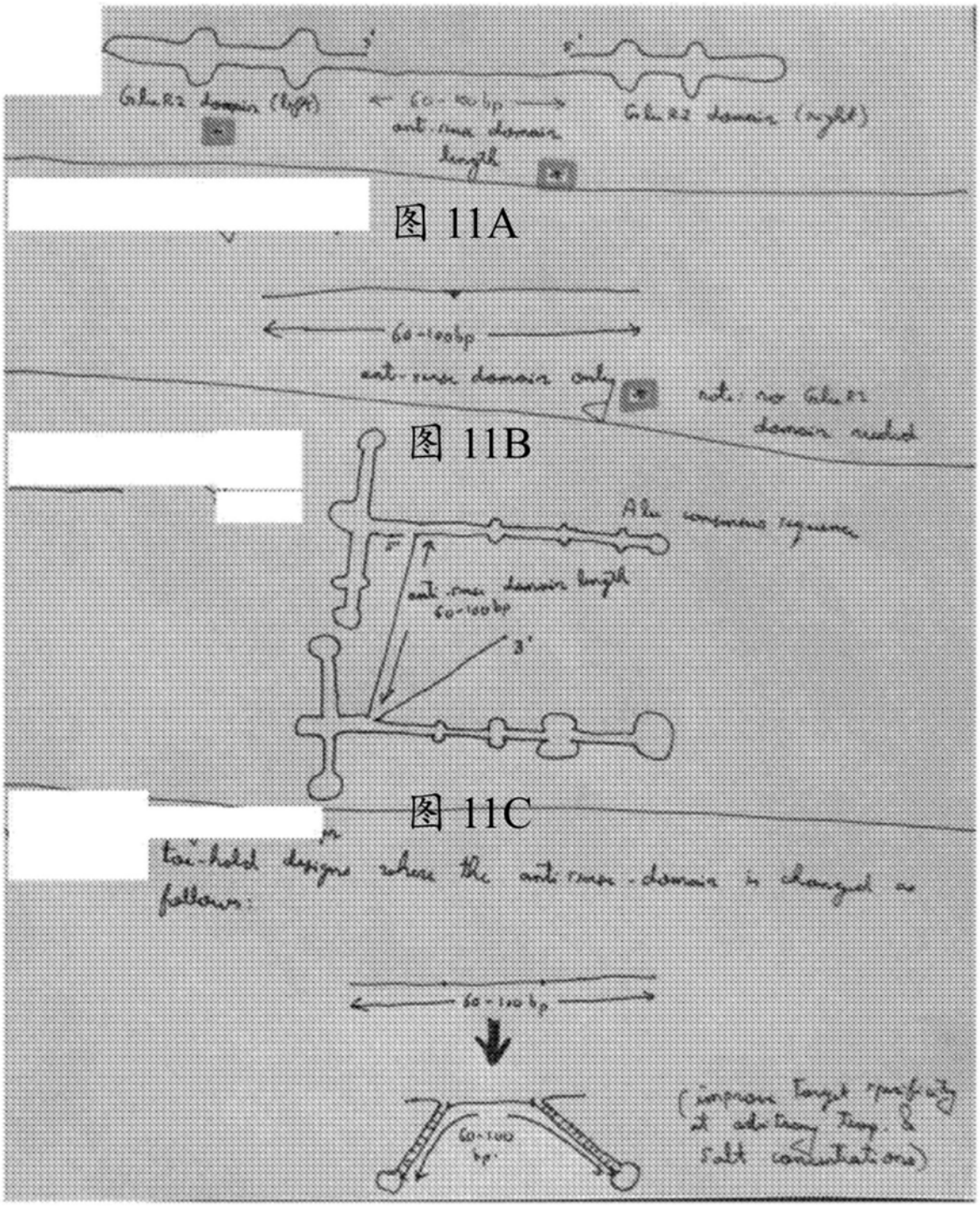


图 11A

图 11B

图 11C

图 11D

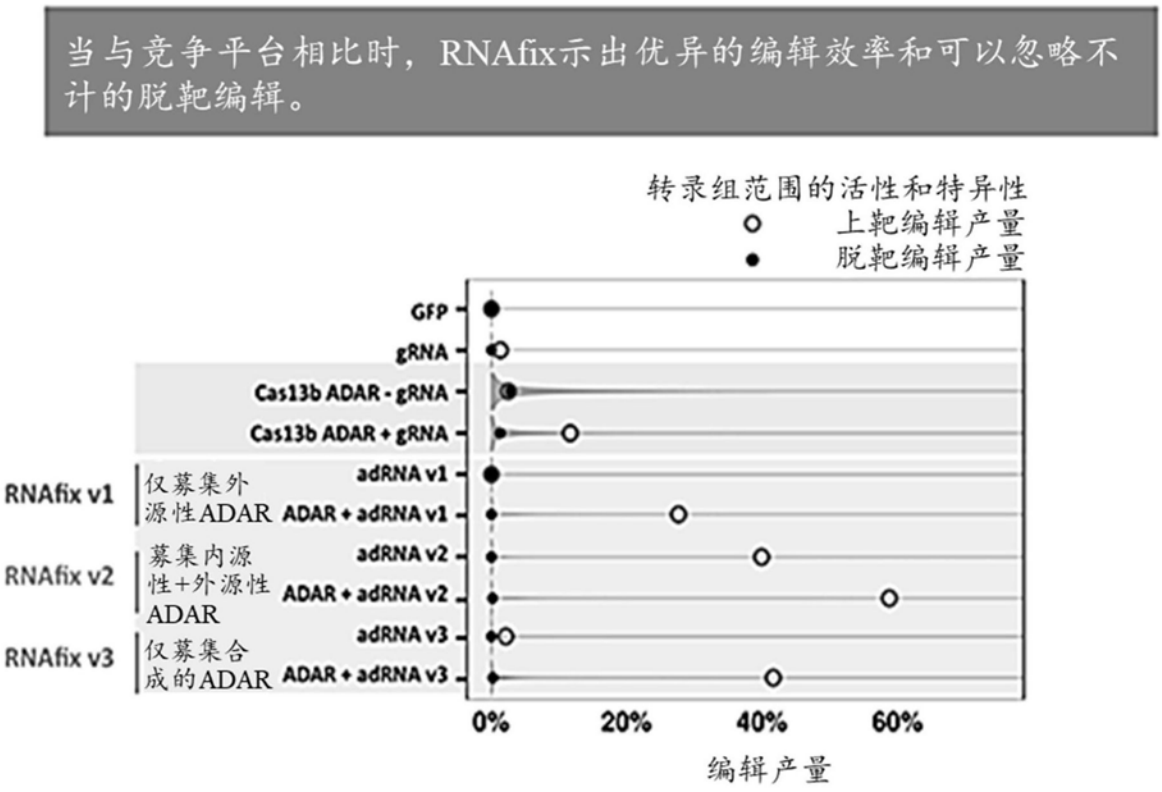

当与竞争平台相比时，RNAfix示出优异的编辑效率和可以忽略不计的脱靶编辑。
转录组范围的活性和特异性
上靶编辑产量
脱靶编辑产量
GFP
gRNA
Cas13b ADAR - gRNA
Cas13b ADAR + gRNA
RNAfix v1
仅募集外源性ADAR
adRNA v1
ADAR + adRNA v1
RNAfix v2
募集内源性+外源性ADAR
adRNA v2
ADAR + adRNA v2
RNAfix v3
仅募集合成的ADAR
adRNA v3
ADAR + adRNA v3
0%
20%
40%
60%
编辑产量

图12

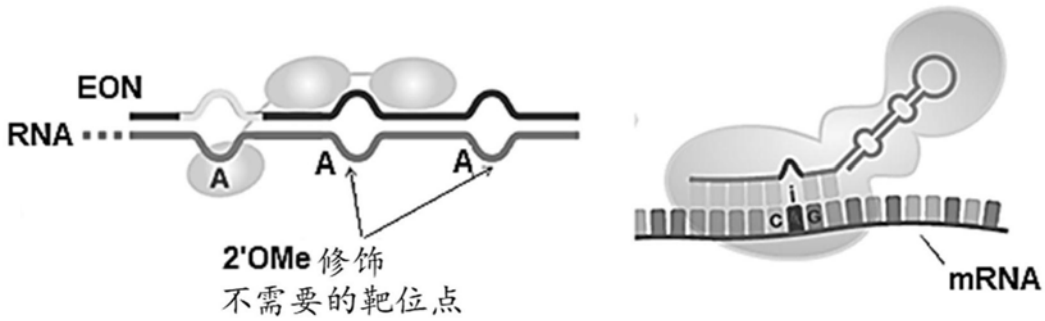

EON
RNA
A
A
A
2'OMe 修饰
不需要的靶位点
C
G
mRNA

图13

adRNA_0_100_50的通用设计：

NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

adRNA_2_100_50的通用设计：

GTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCACNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNGTGGAATAGTATAACAATATGCTAAATGTTGTTATAGTATCCCAC

图14

代表性的adRNA结构

参数

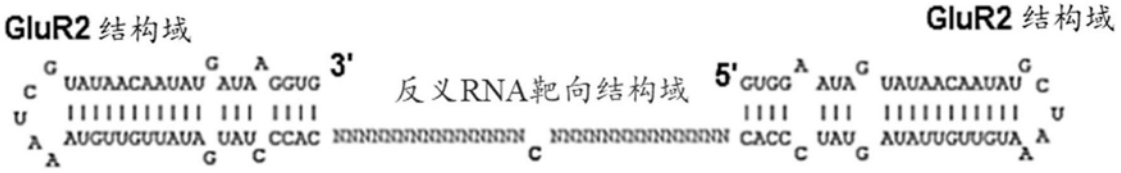
(d,l,m)

d: GluR2结构域的数量

l:反义结构域的长度

m: 错配的位置

adRNA (2,30,15)



adRNA (0,30,15)



adRNA (1,20,6)

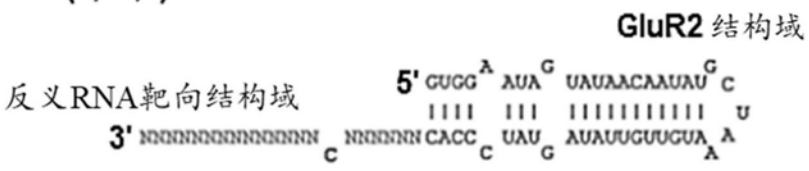


图15

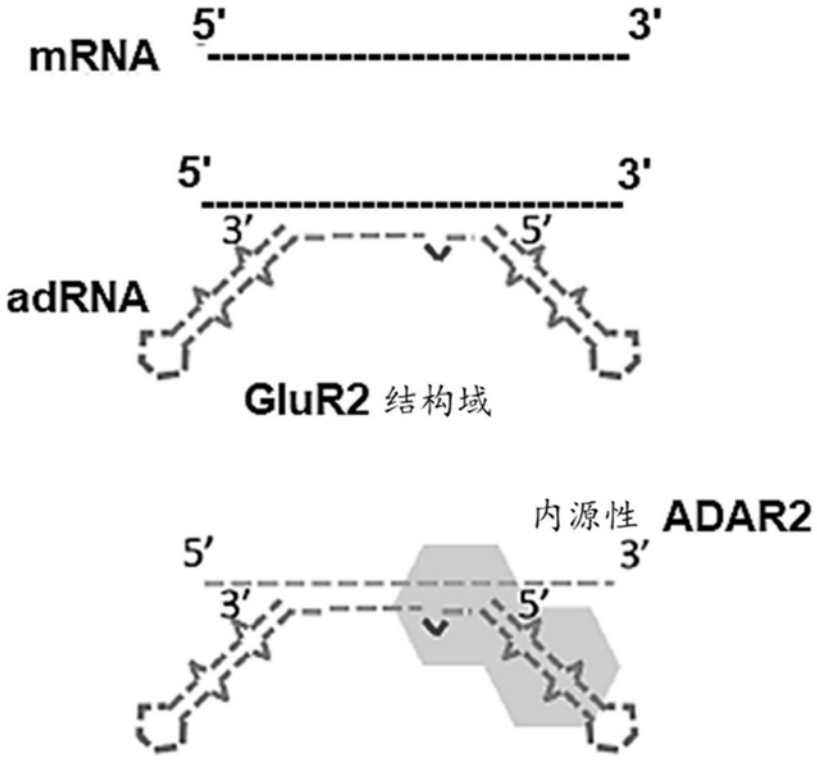

mRNA
5'
3'
5'
3'
3'
5'
adRNA
GluR2 结构域
内源性 ADAR2
5'
3'
3'
5'

图16

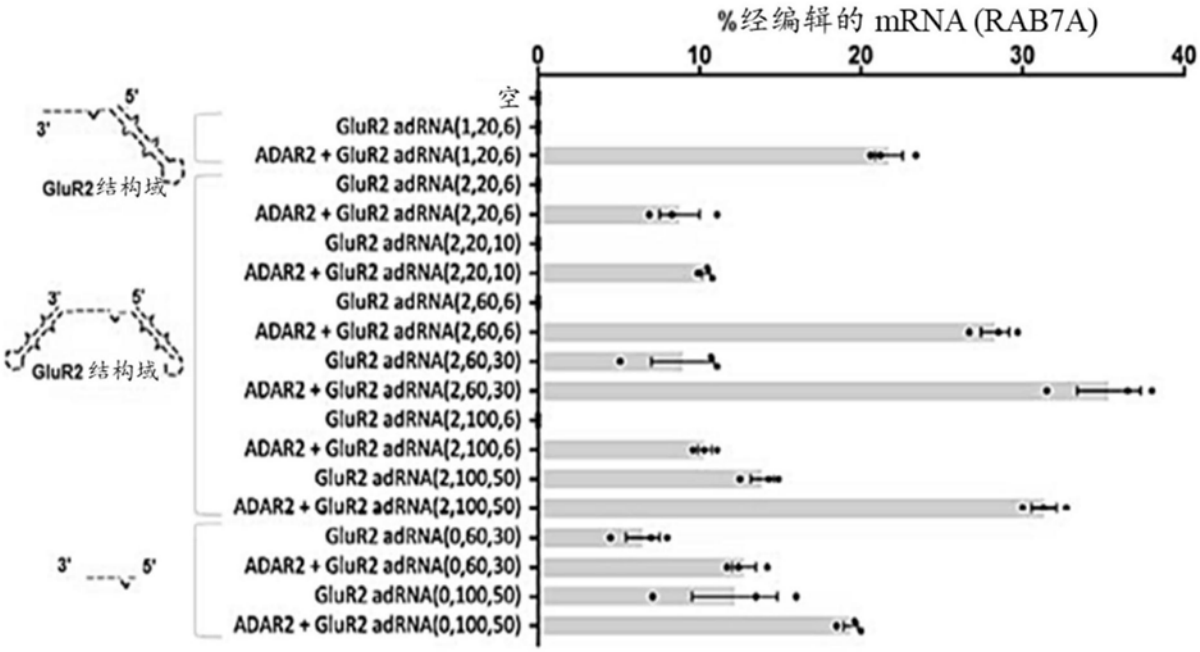

%经编辑的 mRNA (RAB7A)
0
10
20
30
40
空
GluR2 adRNA(1,20,6)
ADAR2 + GluR2 adRNA(1,20,6)
GluR2 adRNA(2,20,6)
ADAR2 + GluR2 adRNA(2,20,6)
GluR2 adRNA(2,20,10)
ADAR2 + GluR2 adRNA(2,20,10)
GluR2 adRNA(2,60,6)
ADAR2 + GluR2 adRNA(2,60,6)
GluR2 adRNA(2,60,30)
ADAR2 + GluR2 adRNA(2,60,30)
GluR2 adRNA(2,100,6)
ADAR2 + GluR2 adRNA(2,100,6)
GluR2 adRNA(2,100,50)
ADAR2 + GluR2 adRNA(2,100,50)
GluR2 adRNA(0,60,30)
ADAR2 + GluR2 adRNA(0,60,30)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)
GluR2结构域
GluR2 结构域

图17

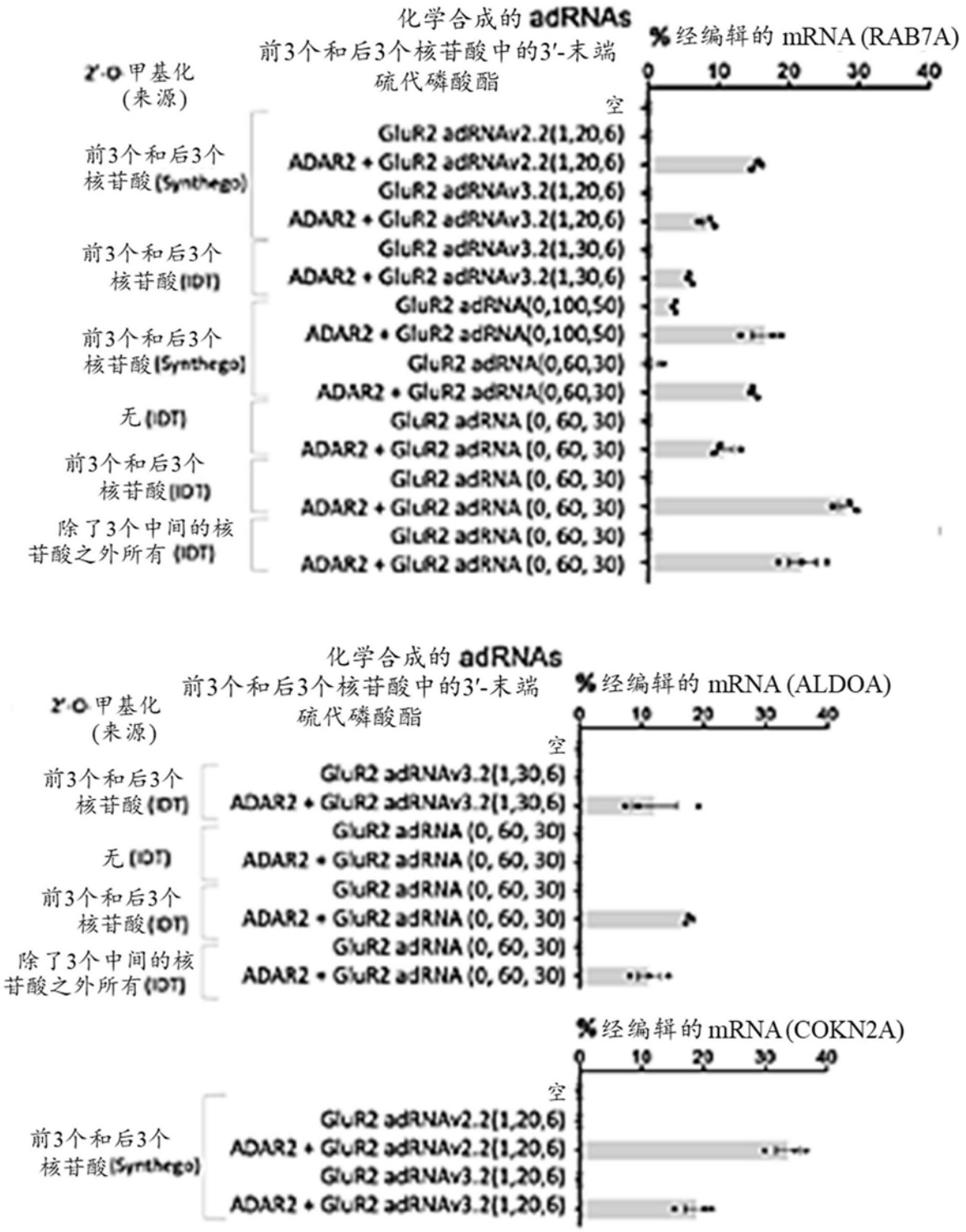

化学合成的 adRNAs
2'-O 甲基化 (来源)
前3个和后3个核苷酸中的3'-末端硫代磷酸酯
% 经编辑的 mRNA (RAB7A)
0 10 20 30 40
空
GluR2 adRNAv2.2(1,20,6)
ADAR2 + GluR2 adRNAv2.2(1,20,6)
GluR2 adRNAv3.2(1,20,6)
ADAR2 + GluR2 adRNAv3.2(1,20,6)
GluR2 adRNAv3.2(1,30,6)
ADAR2 + GluR2 adRNAv3.2(1,30,6)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)
GluR2 adRNA(0,60,30)
ADAR2 + GluR2 adRNA(0,60,30)
GluR2 adRNA (0, 60, 30)
ADAR2 + GluR2 adRNA (0, 60, 30)
前3个和后3个核苷酸(Synthego)
前3个和后3个核苷酸(IDT)
前3个和后3个核苷酸(Synthego)
无(IDT)
前3个和后3个核苷酸(IDT)
除了3个中间的核苷酸之外所有(IDT)
化学合成的 adRNAs
2'-O 甲基化 (来源)
前3个和后3个核苷酸中的3'-末端硫代磷酸酯
% 经编辑的 mRNA (ALDOA)
0 10 20 30 40
空
GluR2 adRNAv3.2(1,30,6)
ADAR2 + GluR2 adRNAv3.2(1,30,6)
GluR2 adRNA (0, 60, 30)
ADAR2 + GluR2 adRNA (0, 60, 30)
GluR2 adRNA (0, 60, 30)
ADAR2 + GluR2 adRNA (0, 60, 30)
GluR2 adRNA (0, 60, 30)
ADAR2 + GluR2 adRNA (0, 60, 30)
前3个和后3个核苷酸(IDT)
无(IDT)
前3个和后3个核苷酸(IDT)
除了3个中间的核苷酸之外所有(IDT)
% 经编辑的 mRNA (COKN2A)
0 10 20 30 40
空
GluR2 adRNAv2.2(1,20,6)
ADAR2 + GluR2 adRNAv2.2(1,20,6)
GluR2 adRNAv3.2(1,20,6)
ADAR2 + GluR2 adRNAv3.2(1,20,6)
前3个和后3个核苷酸(Synthego)

图18

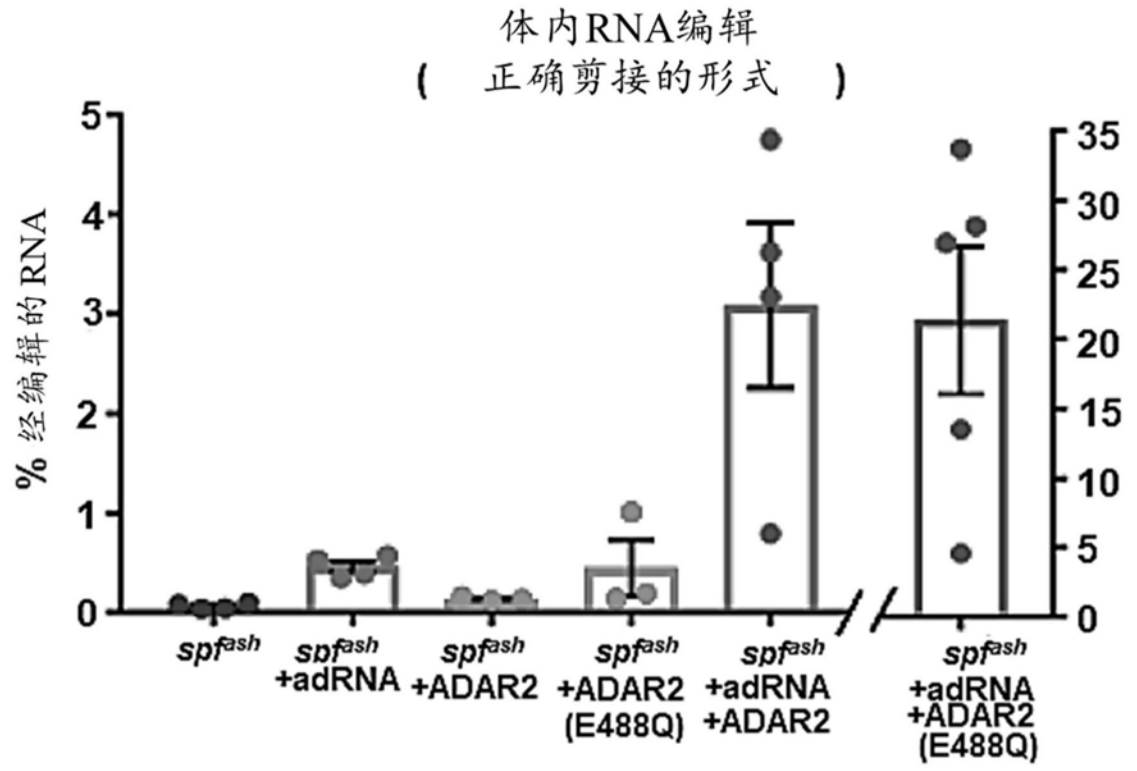

体内RNA编辑
(正确剪接的形式)
% 经编辑的 RNA
0
1
2
3
4
5
0
5
10
15
20
25
30
35
spfash
spfash +adRNA
spfash +ADAR2
spfash +ADAR2 (E488Q)
spfash +adRNA +ADAR2
spfash +adRNA +ADAR2 (E488Q)

图19

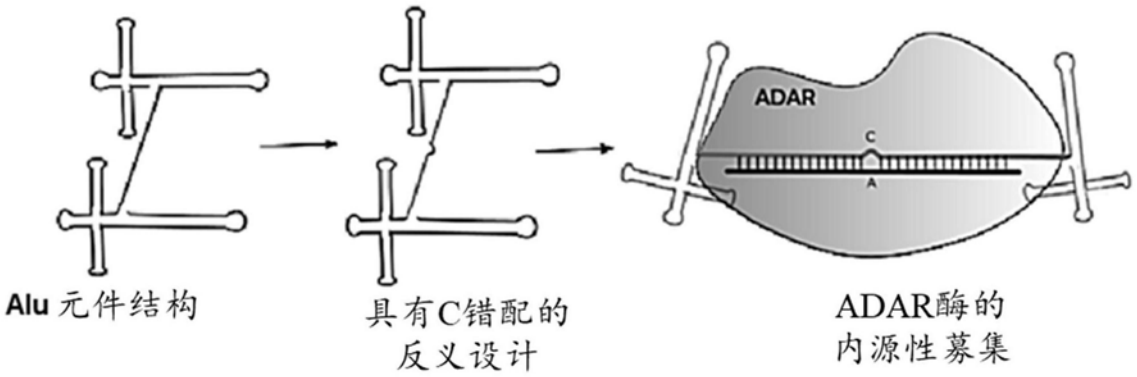

ADAR
C
A
Alu 元件结构
具有C错配的
反义设计
ADAR酶的
内源性募集

图20

Alu v2 导向序列:

Alu 元件

反义向导

靶标C错配

Alu-v2-20-10:

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGGAGATTGC
TTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCATACTGCCGCCA
GCTGGATTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGG
AGGATCGCTTGAGCCCAGGAGTTCGAGGCTGCAGTGAGCTATGATCGCGCCACTGCACTCCAGC
CTGGGCGACAGAGCGAGACCCTGTCTC

Alu-v2-40-20:

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGGAGATTGC
TTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCACTGTACAGAAT
ACTGCCGCCAGCTGGATTTCCCAATTCTAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCT
ACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTCGAGGCTGCAGTGAGCTATGA
TCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACCCTGTCTC

Alu-v2-60-30:

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGGAGATTGC
TTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTCTTGTGTCTAC
TGTACAGAATACTGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACAGCCGGGCGTGGTGGCG
CGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGGATCGCTTGAGCCCAGGAGTTCG
AGGCTGCAGTGAGCTATGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACCCTGTC
TC

Alu-v2-100-50:

GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGGAGATTGCTTGAGC
CCAGGAGTTCGAGACCAGCCTGGGCAACATAGCGAGACCCCGTCTCTGATAAAAGGCGTACATAATTCTT
GTGTCTACTGTACAGAATACTGCCGCCAGCTGGATTTCCCAATTCTGAGTAACACTCTGCAATCCAAACA
GGGTTCAGCCGGGCGTGGTGGCGCGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGGATCGC
TTGAGCCCAGGAGTTCGAGGCTGCAGTGAGCTATGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGC
GAGACCCTGTCTC

图21

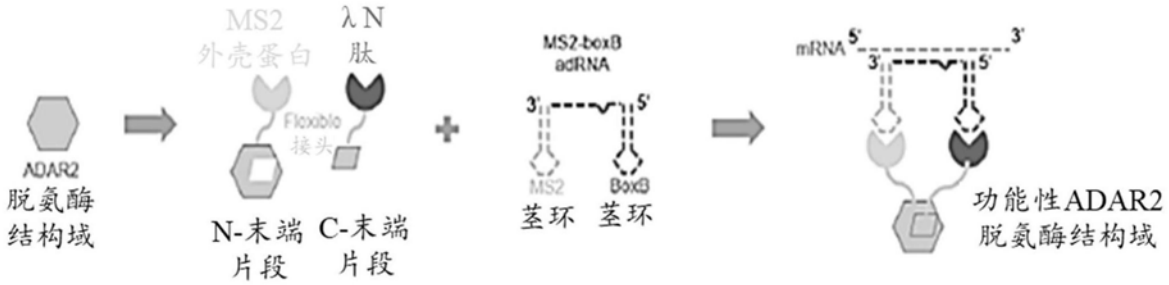


图22

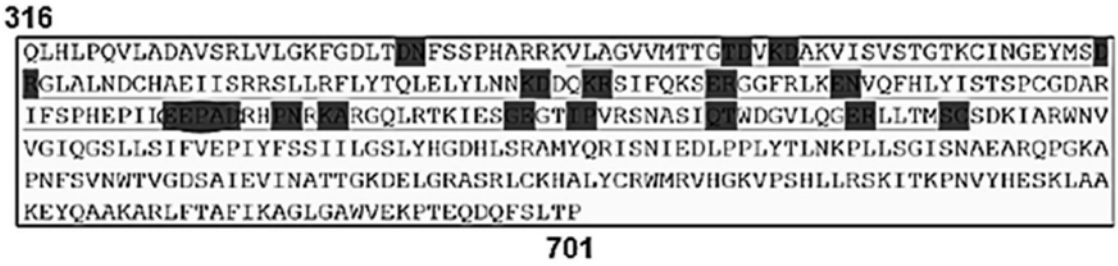

316
701

图23

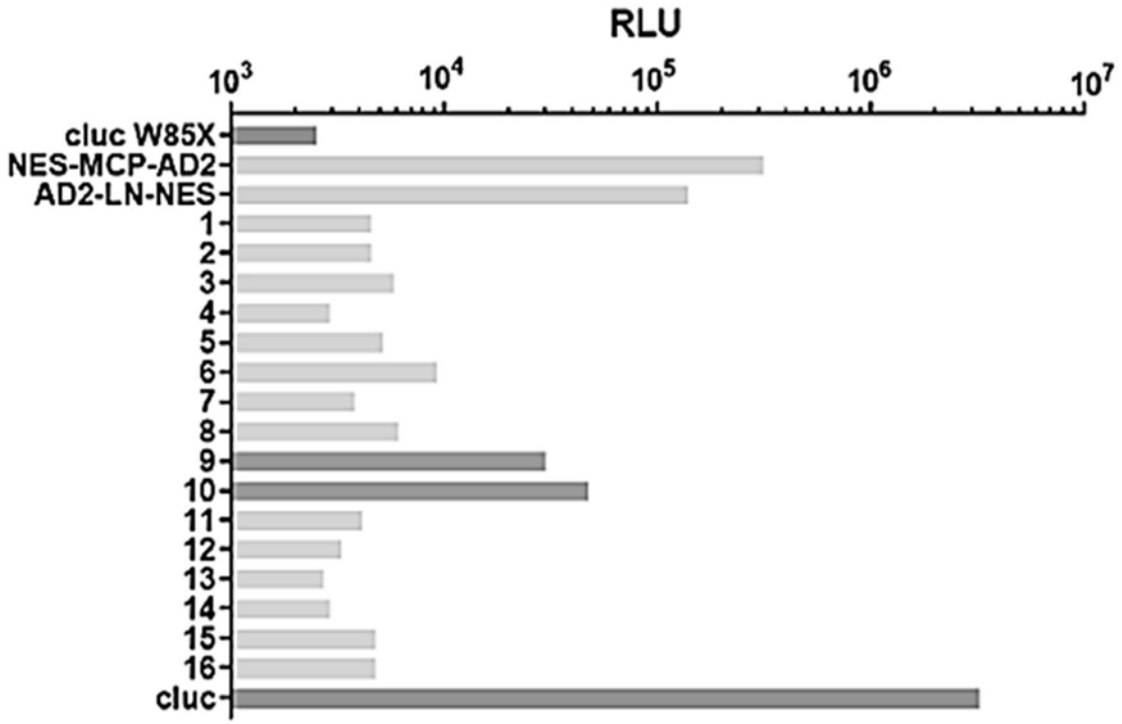

RLU
10^3
10^4
10^5
10^6
10^7
cluc W85X
NES-MCP-AD2
AD2-LN-NES
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
cluc

图24

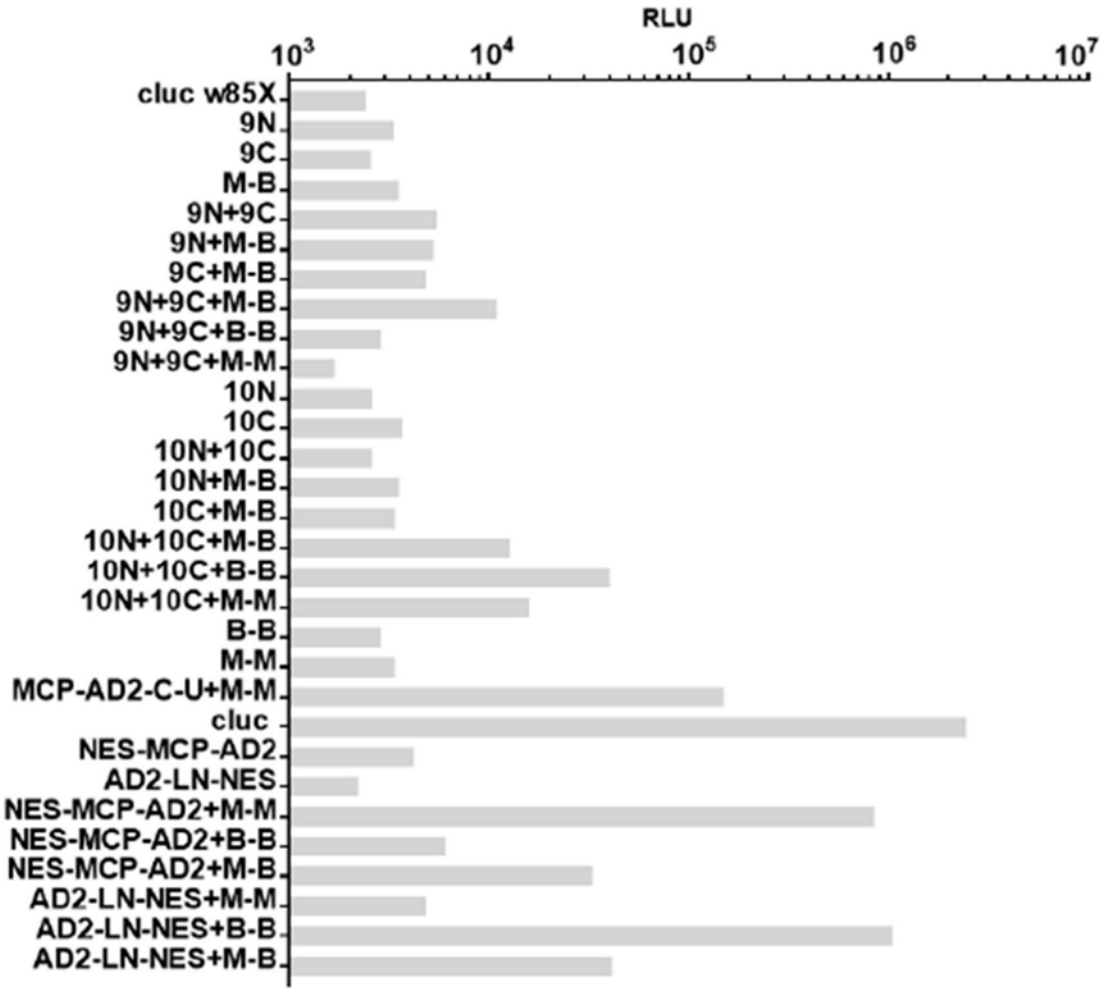

RLU
10^3
10^4
10^5
10^6
10^7
cluc w85X
9N
9C
M-B
9N+9C
9N+M-B
9C+M-B
9N+9C+M-B
9N+9C+B-B
9N+9C+M-M
10N
10C
10N+10C
10N+M-B
10C+M-B
10N+10C+M-B
10N+10C+B-B
10N+10C+M-M
B-B
M-M
MCP-AD2-C-U+M-M
cluc
NES-MCP-AD2
AD2-LN-NES
NES-MCP-AD2+M-M
NES-MCP-AD2+B-B
NES-MCP-AD2+M-B
AD2-LN-NES+M-M
AD2-LN-NES+B-B
AD2-LN-NES+M-B

图25

Cluc W85X 报告因子:

ATGAAGACCTTAATTCTTGCCGTTGCATTAGTCTACTGCGCCACTGTTCATTGCCAGGACTGTCCTTACG
AACCTGATCCACCAAACACAGTTCCAACTTCCTGTGAAGCTAAAGAAGGAGAATGTATTGATAGCAGCTG
TGGCACCTGCACGAGAGACATACTATCAGATGGACTGTGTGAAAATAAACCAGGAAAAACATGTTGCCGA
ATGTGTCAGTATGTAATTGAATGCAGAGTAGAGGCCGCAGGA**TAG**TTTAGAACATTCTATGGAAAGAGAT
TCCAGTTCCAGGAACCTGGTACATACGTGTTGGGTCAAGGAACCAAGGGCGGCGACTGGAAGGTGTCCAT
CACCCTGGAGAACCTGGATGGAACCAAGGGGGCTGTGCTGACCAAGACAAGACTGGAAGTGGCTGGAGAC
ATCATTGACATCGCTCAAGCTACTGAGAATCCCATCACTGTAAACGGTGGAGCTGACCCTATCATCGCCA
ACCCGTACACCATCGGCGAGGTCACCATCGCTGTTGTTGAGATGCCAGGCTTCAACATCACCGTCATTGA
GTTCTTCAAACTGATCGTGATCGACATCCTCGGAGGAAGATCTGTAAGAATCGCCCCAGACACAGCAAAC
AAAGGAATGATCTCTGGCCTCTGTGGAGATCTTAAAATGATGGAAGATACAGACTTCACTTCAGATCCAG
AACAACTCGCTATTCAGCCTAAGATCAACCAGGAGTTTGACGGTTGTCCACTCTATGGAAATCCTGATGA
CGTTGCATACTGCAAAGGTCTTCTGGAGCCGTACAAGGACAGCTGCCGCAACCCCATCAACTTCTACTAC
TACACCATCTCCTGCGCCTTCGCCCGCTGTATGGGTGGAGACGAGCGAGCCTCACACGTGCTGCTTGACT
ACAGGGAGACGTGCGCTGCTCCCGAAACTAGAGGAACCTGCGTTTTGTCTGGACATACTTTCTACGATAC
ATTTGACAAAGCAAGATACCAATTCCAGGGTCCCTGCAAGGAGATTCTTATGGCCGCCGACTGTTTCTGG
AACACTTGGGATGTGAAGGTTTCACACAGGAATGTTGACTCTTACACTGAAGTAGAGAAAGTACGAATCA
GGAAACAATCGACTGTAGTAGAACTCATTGTTGATGGAAAACAGATTCTGGTTGGAGGAGAAGCCGTGTC
CGTCCCGTACAGCTCTCAGAACACTTCCATCTACTGGCAAGATGGTGACATACTGACTACAGCCATCCTA
CCTGAAGCTCTGGTGGTCAAGTTCAACTTCAAGCAACTGCTCGTCGTACATATTAGAGATCCATTCGATG
GTAAGACTTGCGGTATTTGCGGTAACTACAACCAGGATTTCAGTGATGATTCTTTTGATGCTGAAGGAGC
CTGTGATCTGACCCCCAACCCACCGGGATGCACCGAAGAACAGAAACCTGAAGCTGAACGACTCTGCAAT
AGTCTCTTCGCCGGTCAAAGTGATCTTGATCAGAAATGTAACGTGTGCCACAAGCCTGACCGTGTCGAAC
GATGCATGTACGAGTATTGCCTGAGGGGACAACAGGGTTTCTGTGACCACGCATGGGAGTTCAAGAAAGA
ATGCTACATAAAGCATGGAGACACCCTAGAAGTACCAGATGAATGCAAATAA

MS2/BoxB-adRNA 序列:

BoxB_Luc-TAG_MS2_F	**ggccctgaaaaagggcc**TGTTCTAAA**CCA**TCCTGCGGCCTC**aACATGAGGATCACCCATGc**
MS2_Luc-TAG_BoxB_F	**aACATGAGGATCACCCATGTc**CTAAA**CCA**TCCTGCGGCCTCTACTCT**ggccctgaaaaagggcc**
MS2_Luc-TAG_MS2_F	**aACATGAGGATCACCCATGTcCTAAACCATCCTGCGGCCTCaACATGAGGATCACCCATGTc**
BoxB_Luc-TAG_BoxB_F	**ggccctgaaaaagggccTGTTCTAAACCATCCTGCGGCCTCTACTCTggccctgaaaaagggcc**

NES-FLAG-MCP-N- 末端 -AD2-9:

MLPPLERLTLGSDYKDDDDKGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYK
VTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS
GIYGGSGSGAGSGSPAGGGAPGSGGGSQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMT
TGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQK
SERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILE

图26

C- 末端 -AD2-9-LN-HA-NES:

MEPADRHPNRKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLS
IFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTV
GDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKAR
LFTAFIKAGLGAWVEKPTEQDQFSLTPGGSGSGAGSGSPAGGGAPGSGGGSNARTRRRERRAEKQAQWKA
ANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQW
KAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGSYPYDVPDYAGSLPPLERLTL

NES-FLAG-MCP-N- 末端 -AD2-10:

MLPPLERLTLGSDYKDDDDKGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYK
VTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS
GIYGGSGSGAGSGSPAGGGAPGSGGGSQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMT
TGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNKDDQKRSIFQK
SERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPA

C- 末端 -AD2-10-LN-HA-NES:

MDRHPNRKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFV
EPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSVNWTVGDS
AIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFT
AFIKAGLGAWVEKPTEQDQFSLTPGGSGSGAGSGSPAGGGAPGSGGGSNARTRRRERRAEKQAQWKAANG
GGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAA
NGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGSYPYDVPDYAGSLPPLERLTL

NES-FLAG-MCP-N- 末端 -AD2-C-U- 变体 -9:

MLPPLERLTLGSDYKDDDDKGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYK
VTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS
GIYGGSGSGAGSGSPAGGGAPGSGGGSQLHLPQVLADAVSRLVIGKFGDLTDNFSSPHARRIGLAGVVMT
TGTDVKDAKVICVSTGSKCINGEYLSDRGLALNDCHAEIVSRRSLLRFLYTQLELYLNNEDDQKRSIFQK
SERGGFRLKENIQFHLYISTSPCGDARIFSPHEAILE

C- 末端 -AD2-C-U- 变体 -9-LN-HA-NES:

MEPADRHPNRKARGQLRTKIEAGQGTIPVRNNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLS
IFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLTGISNAEARQPGKAPIFSVNWTV
GDSAIEVINATTGKGELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHETKLAAKEYQAAKAR
LFTAFIKAGLGAWVEKPTEQDQFSLTPGGSGSGAGSGSPAGGGAPGSGGGSNARTRRRERRAEKQAQWKA
ANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQW
KAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGSYPYDVPDYAGSLPPLERLTL

NES-FLAG-MCP-N- 末端 -AD2-C-U- 变体 -10:

MLPPLERLTLGSDYKDDDDKGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYK
VTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS
GIYGGSGSGAGSGSPAGGGAPGSGGGSQLHLPQVLADAVSRLVIGKFGDLTDNFSSPHARRIGLAGVVMT
TGTDVKDAKVICVSTGSKCINGEYLSDRGLALNDCHAEIVSRRSLLRFLYTQLELYLNNEDDQKRSIFQK
SERGGFRLKENIQFHLYISTSPCGDARIFSPHEAILEEPA

图26续

C- 末端 -AD2-C-U- 变体 -10-LN-HA-NES:

MDRHPNRKARGQLRTKIEAGQGTIPVRNNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFV
EPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLTGISNAEARQPGKAPIFSVNWTVGDS
AIEVINATTGKGELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHETKLAAKEYQAAKARLFT
AFIKAGLGAWVEKPTEQDQFSLTPGGSGSGAGSGSPAGGGAPGSGGGSNARTRRRERRAEKQAQWKAANG
GGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAA
NGGGGSGGGGSGGGGSNARTRRRERRAEKQAQWKAANGSYPYDVPDYAGSLPPLERLTL

图26续

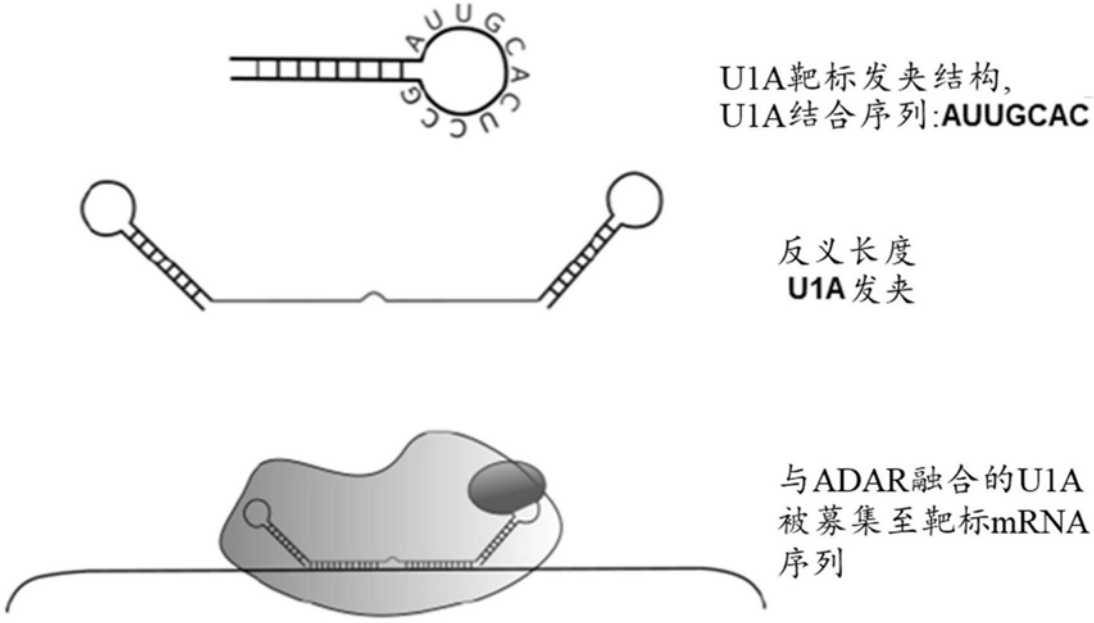


图27

MCP-hAPOBEC1

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SSGSETPGTSESATPESMTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRS
SGKNTTNHVEVNFIKKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLFWHM
DQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMMLYALELHCIILSLPPCL
KISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR

MCP-hAPOBEC2

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMAQKEEAAVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVEYSSGRNKTF
LCYVVEAQGKGGQVQASRGYLEDEHAAAHAEEAFFNTILPAFDPALRYNVTWYVSSSPCAACADRIIKTL
SKTKNLRLLILVGRLFMWEEPEIQAALKKLKEAGCKLRIMKPQDFEYVWQNFVEQEEGESKAFQPWEDIQ
ENFLYYEEKLADILK

MCP-hAPOBEC3A

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQAKNLLCGFYGR
HAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEA
LQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLDEHSQALSGRLRAILQNQGN

MCP-hAPOBEC3B

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQVYFKPQYHAEM
CFLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFLSEHPNVTLTISAARLYYYWERDYRRALCRLS
QAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKFDENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVL
RRRQTYLCYEVERLDNGTWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFI
SWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFV
YRQGCPFQPWDGLEEHSQALSGRLRAILQNQGN

MCP-hAPOBEC3C

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRNQVDSETHCHAER
CFLSWFCDDILSPNTKYQVTWYTSWSPCPDCAGEVAEFLARHSNVNLTIFTARLYYFQYPCYQEGLRSLS
QEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKGLKTNFRLLKRRLRESLQ

图28

MCP-hAPOBEC3D

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGPVLPKRQSNHRQ
EVYFRFENHAEMCFLSWFCGNRLPANRRFQITWFVSWNPCLPCVVKVTKFLAEHPNVTLTISAARLYYYR
DRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVCNEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMY
PHIFYFHFKNLLKACGRNESWLCFTMEVTKHHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPN
TNYEVTWYTSWSPCPECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDF
VSCWKNFVYSDDEPFKPWKGLQTNFRLLKRRLREILQ

MCP-hAPOBEC3F

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRLDAKIFRGQVYSQPEHHAEMC
FLSWFCGNQLPAYKCFQITWFVSWTPCPDCVAKLAEFLAEHPNVTLTISAARLYYYWERDYRRALCRLSQ
AGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFDDNYAFLHRTLKEILRNPMEAMYPHIFYFHFKNLRK
AYGRNESWLCFTMEVVKHHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSP
CPECAGEVAEFLARHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYNDDE
PFKPWKGLKYNFLFLDSKLQEILE

MCP-hAPOBEC3G

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQVYSELKYHPEMR
FFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDMATFLAEDPKVTLTIFVARLYYFWDPDYQEALRSLC
QKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWNNLPKYYILLHIMLGEILRHSMDPPTFTFNFNNE
PWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRV
TCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDT
FVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQEN

MCP-hAPOBEC3H

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFINEIKSMGLD
ETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLRLLCGSQVPVEVMGFP
EFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLERIKIPGVRAQGRYMDILCDAEV

图28续

MCP-hAPOBEC4

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMEPIYEEYLANHGTIVKPYYWLSFSLDCSNCPYHIRTGEEARVSLTEFCQIFGFPYGTTFPQTKHLTFY

ELKTSSGSLVQKGHASSCTGNYIHPESMLFEMNGYLDSAIYNNDSIRHIILYSNNSPCNEANHCCISKMY
NFLITYPGITLSIYFSQLYHTEMDFPASAWNREALRSLASLWPRVVLSPISGGIWHSVLHSFISGVSGSH
VFQPILTGRALADRHNAYEINAITGVKPYFTDVLLQTKRNPNTKAQEALESYPLNNAFPGQFFQMPSGQL
QPNLPPDLRAPVVFVLVPLRDLPPMHMGQNPNKPRNIVRHLNMPQMSFQETKDLGRLPTGRSVEIVEITE
QFASSKEADEKKKKKGKK

MCP-AID

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHVELLFLRYISDWD
LDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFCEDRKAEPEGLRRLHRAGVQIAIM
TFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSRQLRRILLPLYEVDDLRDAFRTLGL

MCP-hADAT1

MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGAGSGSPAGGGAPGSGGG
SMWTADEIAQLCYEHYGIRLPKKGKPEPNHEWTLLAAVVKIQSPADKACDTPDKPVQVTKEVVSMGTGTK
CIGQSKMRKNGDILNDSHAEVIARRSFQRYLLHQLQLAATLKEDSIFVPGTQKGVWKLRRDLIFVFFSSH
TPCGDASIIPMLEFEDQPCCPVFRNWAHNSSVEASSNLEAPGNERKCEDPDSPVTKKMRLEPGTAAREVT
NGAAHHQSFGKQKSGPISPGIHSCDLTVEGLATVTRIAPGSAKVIDVYRTGAKCVPGEAGDSGKPGAAFH
QVGLLRVKPGRGDRTRSMSCSDKMARWNVLGCQGALLMHLLEEPIYLSAVVIGKCPYSQEAMQRALIGRC
QNVSALPKGFGVQELKILQSDLLFEQSRSAVQAKRADSPGRLVPCGAAISWSAVPEQPLDVTANGFPQGT
TKKTIGSLQARSQISKVELFRSFQKLLSRIARDKWPHSLRVQKLDTYQEYKEAASSYQEAWSTLRKQVFG
SWIRNPPDYHQFK

图28续

靶标基因座：GFP L202S

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAA
ACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTT
CATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAG
TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG
TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGG
CGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC
AAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGG
TGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACAC
CCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAATCATTAT**TCG**AGCACTCAGTCCGCCCTGAGCAAA
GACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA
TGGACGAGCTGTACAAGTACTCAGATCTCGAGCTCAAGTAG

apRNA_v1:

gcgcctccatcgaagtgtaaagccatctatcgatgggaaaGTCTTTGCTCTCCCCGGACACTCTGCTCCT
TAAATGAAACACGGGCTCGTGCACGGGGCCaaagcgcctccatcgaagtgtaaagccatctatcgatggg

apRNA_v2:

gccatctatcgatgggaaaGTCTTTGCTCTCCCCGGACACTCTGCTCCTTAAATGAAACACGGGCTCGTG
CACGGGGCCaaagccatctatcgatggg

apRNA_v3:

gccatctatcgatgggaaaCGGACTGAGTGCTCaaAATGATTGTCGGGCaaagccatctatcgatggg

apRNA_v4:

gccatctatcgatgggaaaTGCTCAGGGCGGACTGAGaaaaTTGTCGGGCAGCAGCACGaaagccatcta
tcgatggg

图29

MS2-apRNA_v1:

aACATGAGGATCACCCATGTcGTCTTTGCTCTCCCCGGACACTCTGCTCCTTAAATGAAACACGGGCTCG
TGCACGGGGCCaACATGAGGATCACCCATGTc

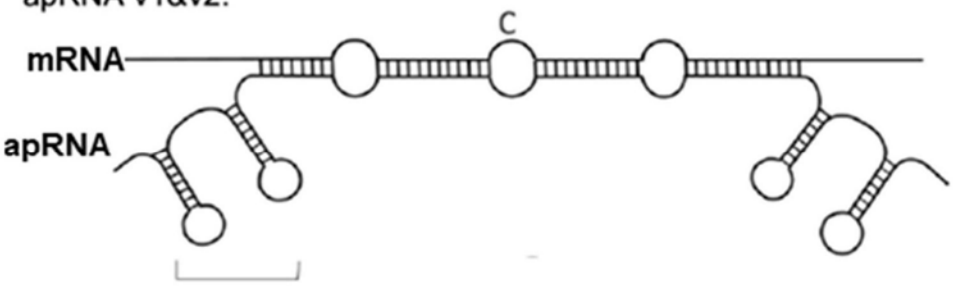
MS2-apRNA_v2:

aACATGAGGATCACCCATGTcCGGACTGAGTGCTCaaAATGATTGTCGGGCaACATGAGGATCACCCATG
Tc

MS2-apRNA_v3:

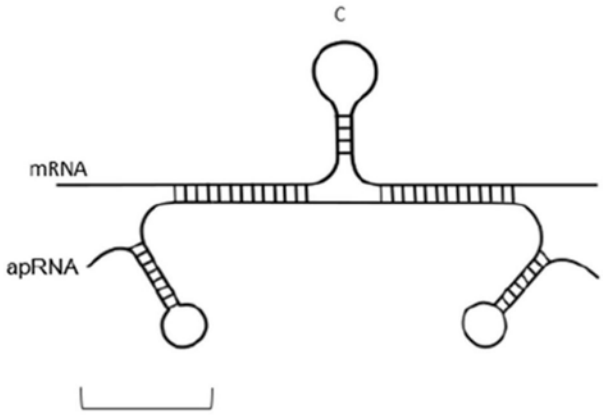
aACATGAGGATCACCCATGTcTGCTCAGGGCGGACTGAGaaaaTTGTCGGGCAGCAGCACGaACATGAGG
ATCACCCATGTc

apRNA v1&v2:



基于内源性靶标的
APO3A募集
结构域

apRNA v3&v4:



基于内源性靶标的
APO3A募集
结构域

APO3A募集结构域序列:

1. ASCC2: gcgcctccatcgaagtgt
2. SDHB: gccatctatcgatggg

图30

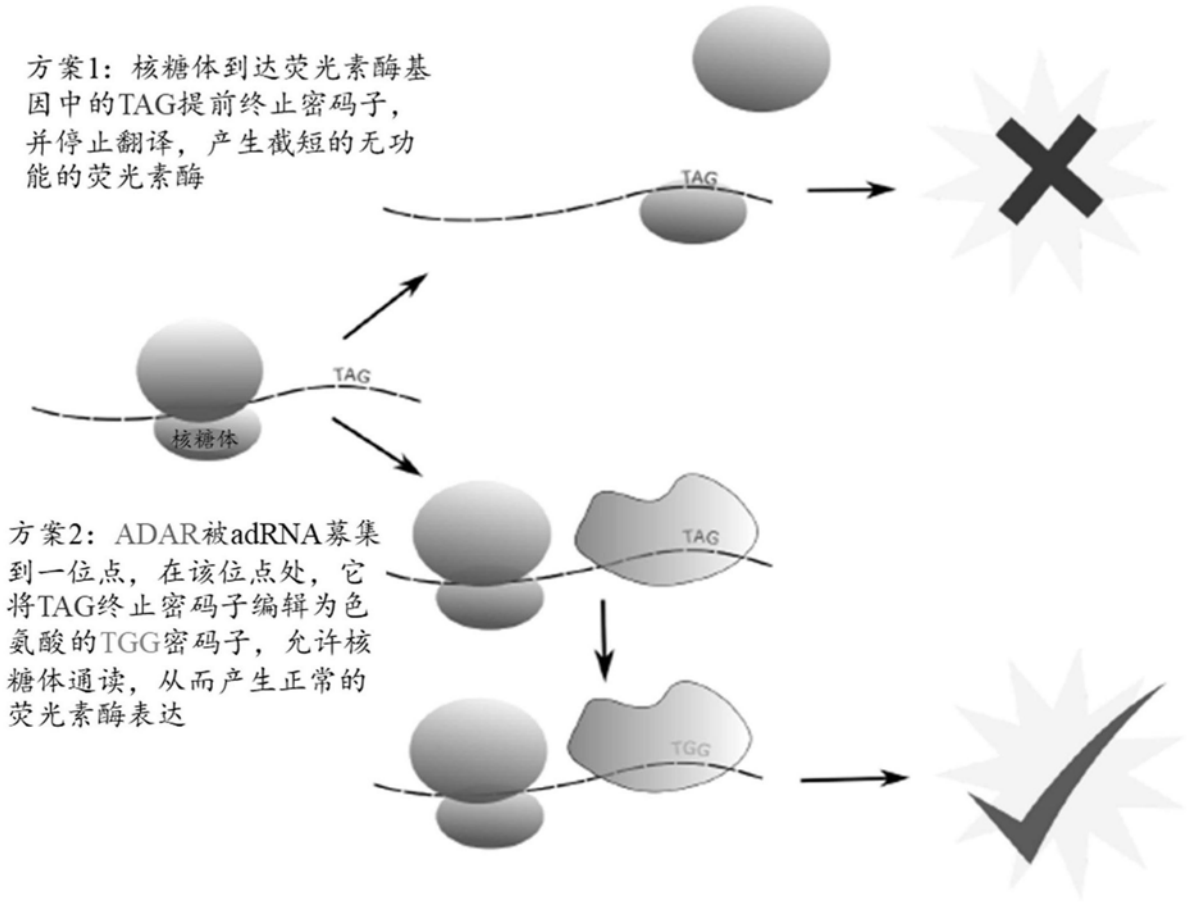

方案1：核糖体到达荧光素酶基因中的TAG提前终止密码子，并停止翻译，产生截短的无功能的荧光素酶
TAG
核糖体
方案2：ADAR被adRNA募集到一位点，在该位点处，它将TAG终止密码子编辑为色氨酸的TGG密码子，允许核糖体通读，从而产生正常的荧光素酶表达
TGG

图31

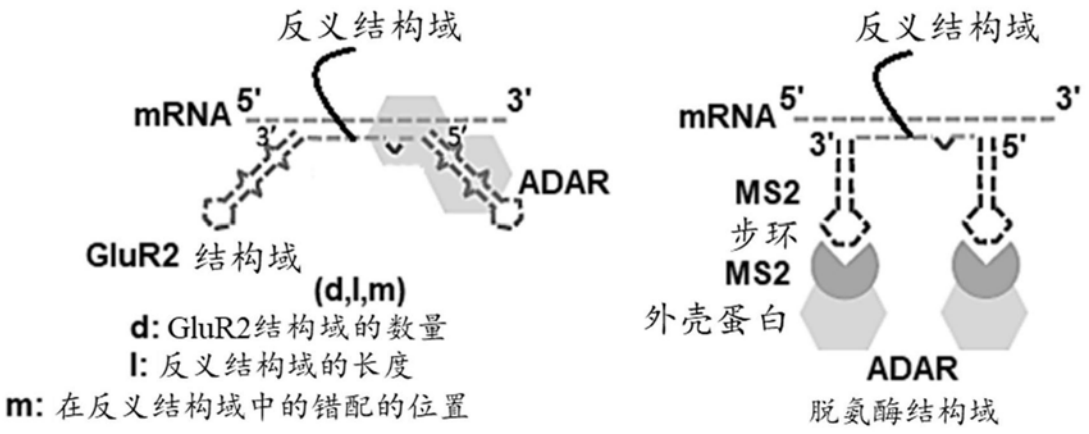

反义结构域
mRNA 5'
3'
3'
5'
ADAR
GluR2 结构域
(d,l,m)
d: GluR2结构域的数量
l: 反义结构域的长度
m: 在反义结构域中的错配的位置
反义结构域
mRNA 5'
3'
3'
5'
MS2
步环
MS2
外壳蛋白
ADAR
脱氨酶结构域

图32A

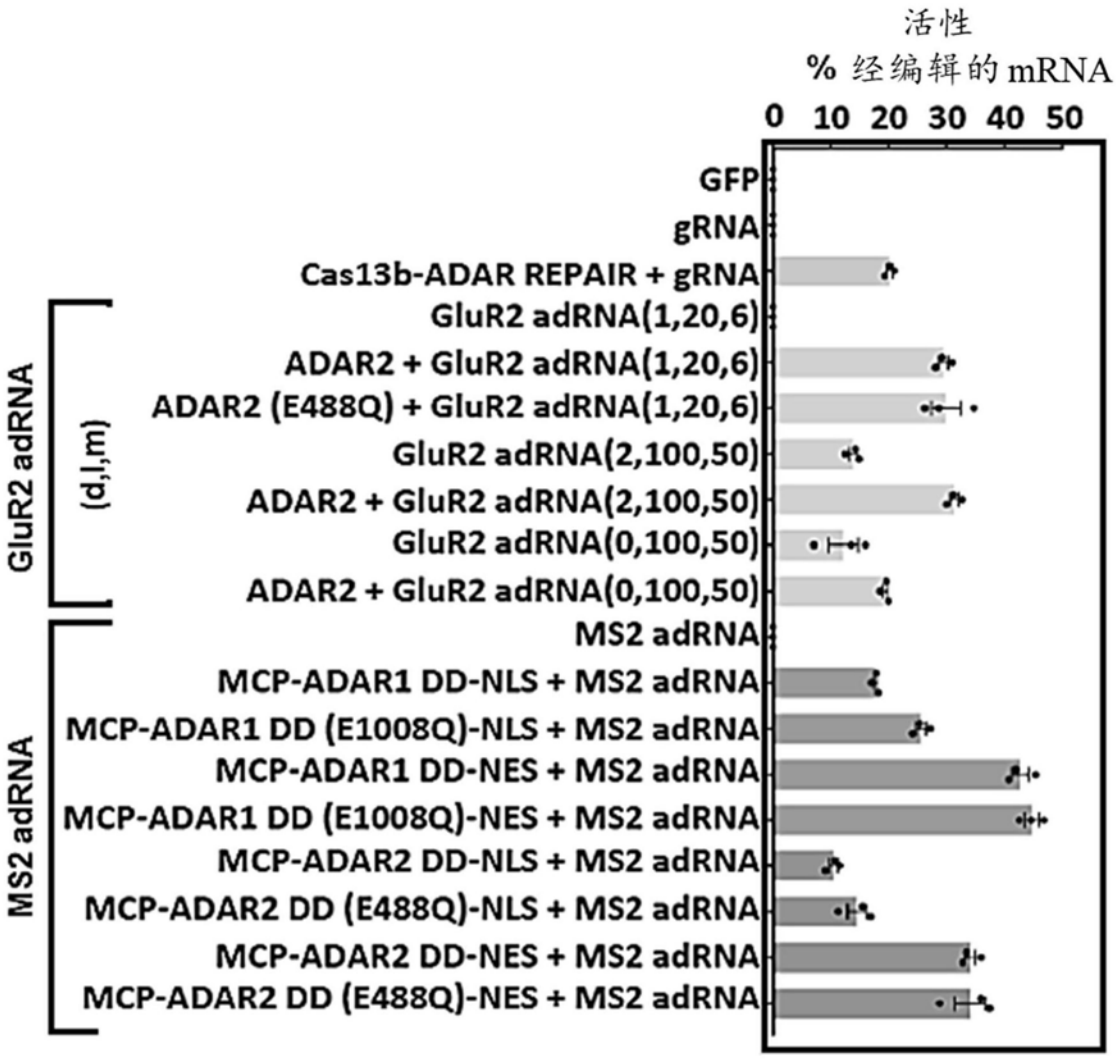

活性
% 经编辑的 mRNA
0 10 20 30 40 50
GFP
gRNA
Cas13b-ADAR REPAIR + gRNA
GluR2 adRNA(1,20,6)
ADAR2 + GluR2 adRNA(1,20,6)
ADAR2 (E488Q) + GluR2 adRNA(1,20,6)
GluR2 adRNA(2,100,50)
ADAR2 + GluR2 adRNA(2,100,50)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)
MS2 adRNA
MCP-ADAR1 DD-NLS + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NLS + MS2 adRNA
MCP-ADAR1 DD-NES + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NES + MS2 adRNA
MCP-ADAR2 DD-NLS + MS2 adRNA
MCP-ADAR2 DD (E488Q)-NLS + MS2 adRNA
MCP-ADAR2 DD-NES + MS2 adRNA
MCP-ADAR2 DD (E488Q)-NES + MS2 adRNA
GluR2 adRNA
(d,l,m)
MS2 adRNA

图32B

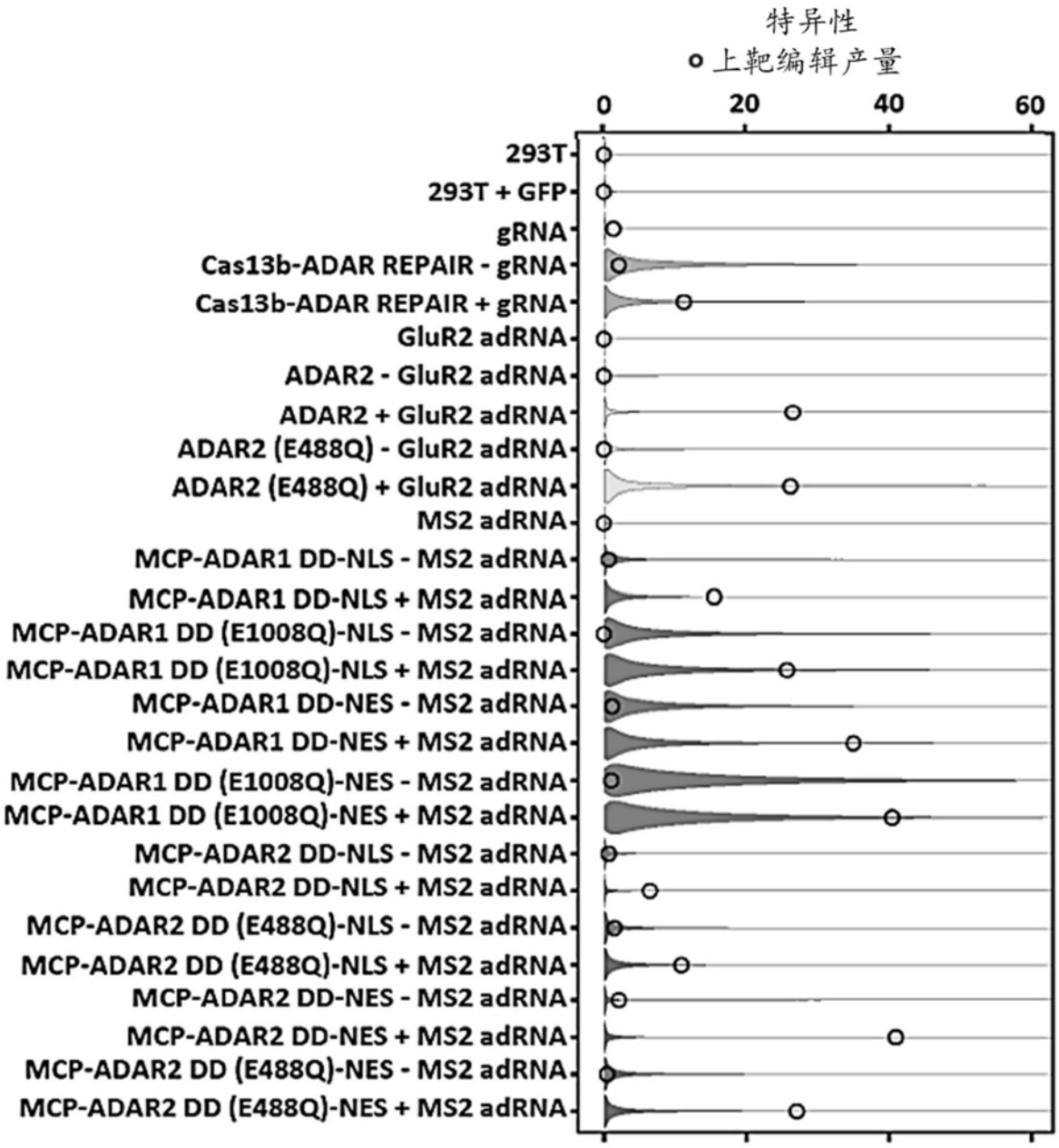

特异性
上靶编辑产量
0
20
40
60
293T
293T + GFP
gRNA
Cas13b-ADAR REPAIR - gRNA
Cas13b-ADAR REPAIR + gRNA
GluR2 adRNA
ADAR2 - GluR2 adRNA
ADAR2 + GluR2 adRNA
ADAR2 (E488Q) - GluR2 adRNA
ADAR2 (E488Q) + GluR2 adRNA
MS2 adRNA
MCP-ADAR1 DD-NLS - MS2 adRNA
MCP-ADAR1 DD-NLS + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NLS - MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NLS + MS2 adRNA
MCP-ADAR1 DD-NES - MS2 adRNA
MCP-ADAR1 DD-NES + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NES - MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NES + MS2 adRNA
MCP-ADAR2 DD-NLS - MS2 adRNA
MCP-ADAR2 DD-NLS + MS2 adRNA
MCP-ADAR2 DD (E488Q)-NLS - MS2 adRNA
MCP-ADAR2 DD (E488Q)-NLS + MS2 adRNA
MCP-ADAR2 DD-NES - MS2 adRNA
MCP-ADAR2 DD-NES + MS2 adRNA
MCP-ADAR2 DD (E488Q)-NES - MS2 adRNA
MCP-ADAR2 DD (E488Q)-NES + MS2 adRNA

图32C

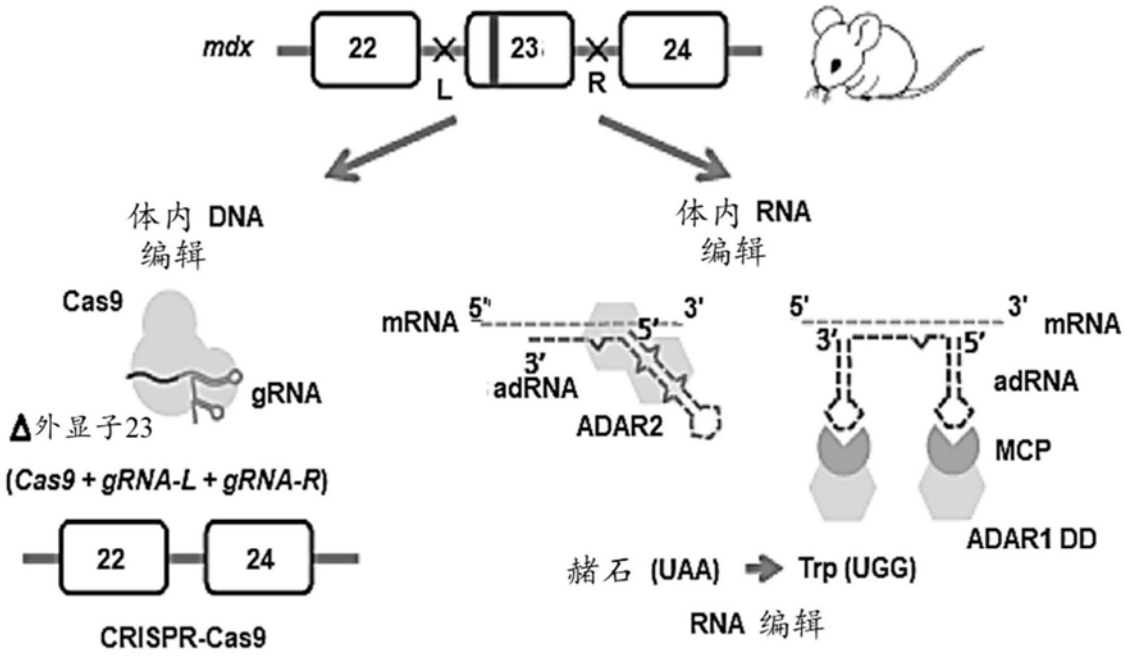

mdx
22
23
24
L
R
体内 DNA
编辑
体内 RNA
编辑
Cas9
gRNA
Δ外显子23
(Cas9 + gRNA-L + gRNA-R)
22
24
CRISPR-Cas9
mRNA
5'
3'
adRNA
ADAR2
MCP
ADAR1 DD
赭石 (UAA)
Trp (UGG)
RNA 编辑

图33A

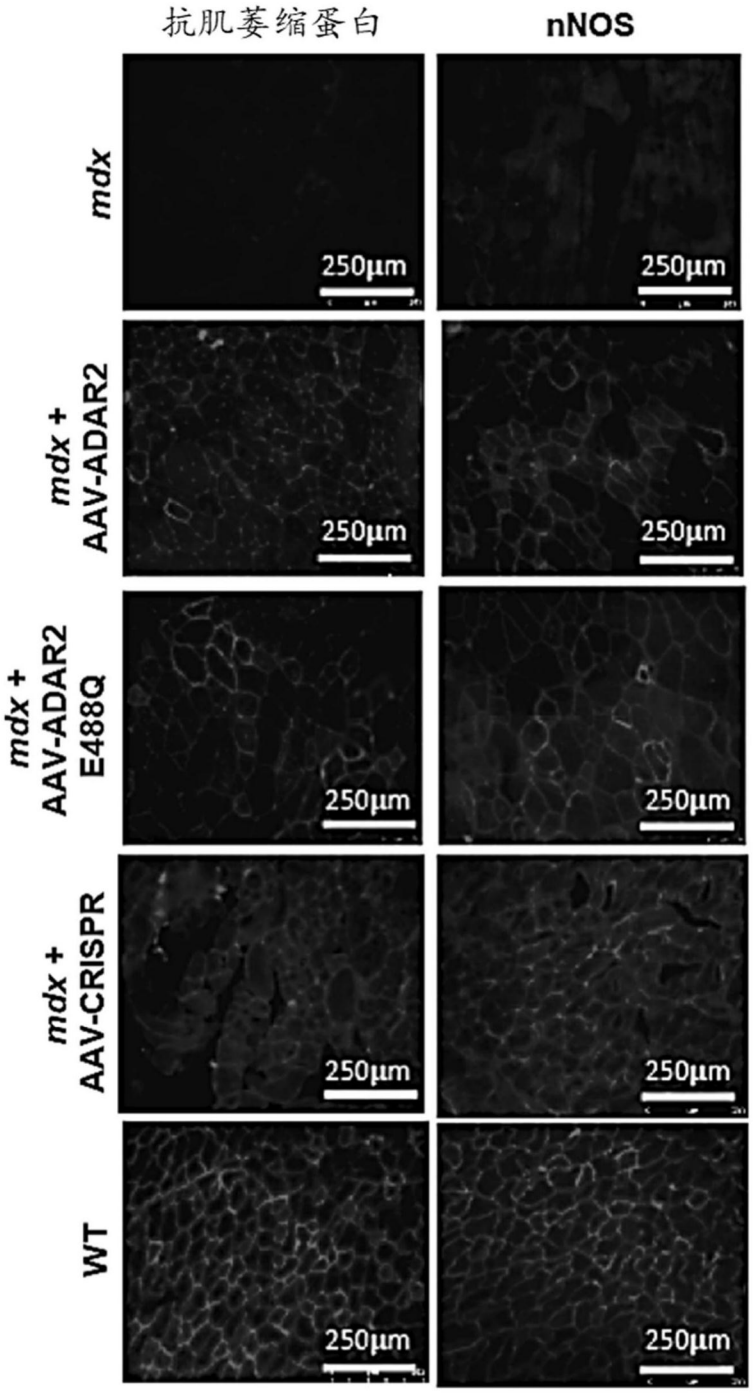

抗肌萎缩蛋白
nNOS
mdx
mdx + AAV-ADAR2
mdx + AAV-ADAR2 E488Q
mdx + AAV-CRISPR
WT
250μm

图33B

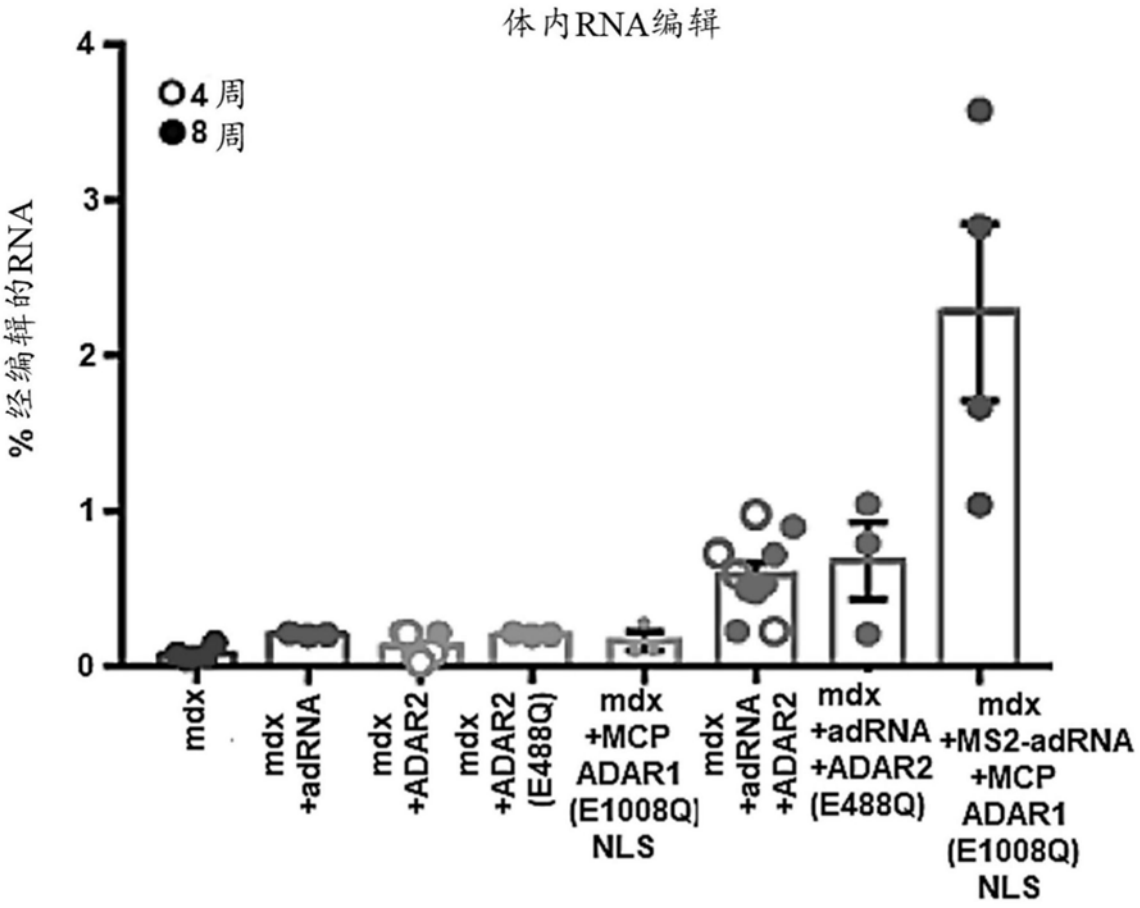

体内RNA编辑
4周
8周
% 经编辑的RNA
4
3
2
1
0
mdx
mdx +adRNA
mdx +ADAR2
mdx +ADAR2 (E488Q)
mdx +MCP ADAR1 (E1008Q) NLS
mdx +adRNA +ADAR2
mdx +adRNA +ADAR2 (E488Q)
mdx +MS2-adRNA +MCP ADAR1 (E1008Q) NLS

图33C

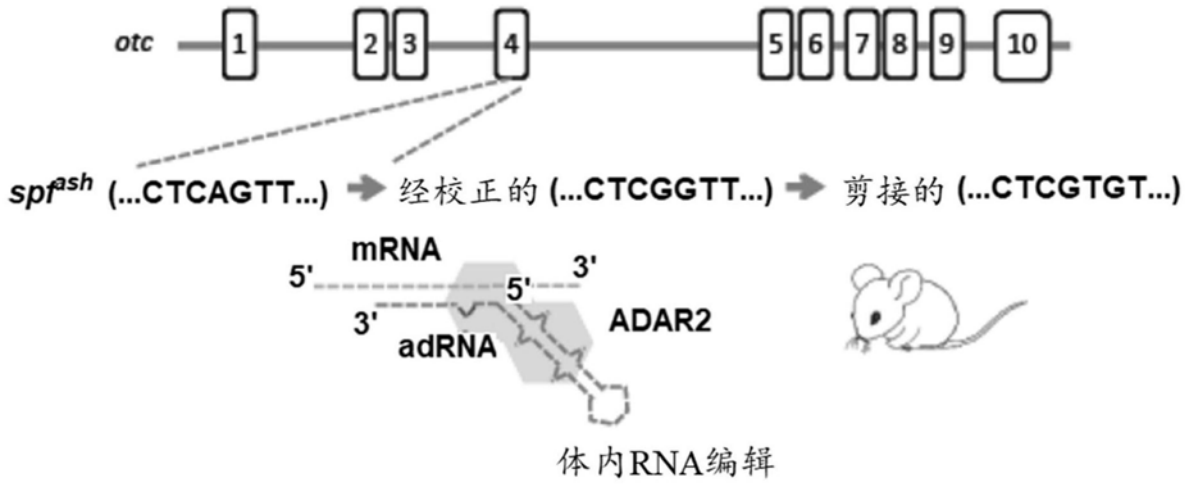

otc
1
2
3
4
5
6
7
8
9
10
spfash (...CTCAGTT...) 经校正的 (...CTCGGTT...) 剪接的 (...CTCGTGT...)
mRNA
5'
3'
5'
3'
adRNA
ADAR2
体内RNA编辑

图33D

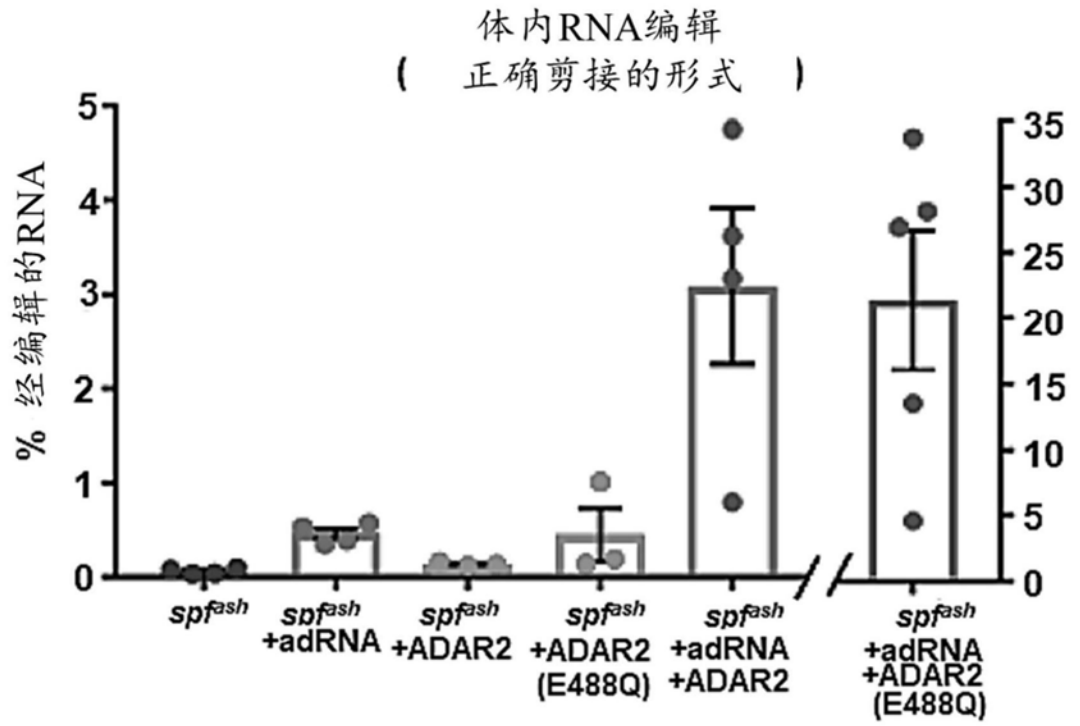

体内RNA编辑
(正确剪接的形式)
% 经编辑的RNA
0
1
2
3
4
5
0
5
10
15
20
25
30
35
spfash
spfash +adRNA
spfash +ADAR2
spfash +ADAR2 (E488Q)
spfash +adRNA +ADAR2
spfash +adRNA +ADAR2 (E488Q)

图33E

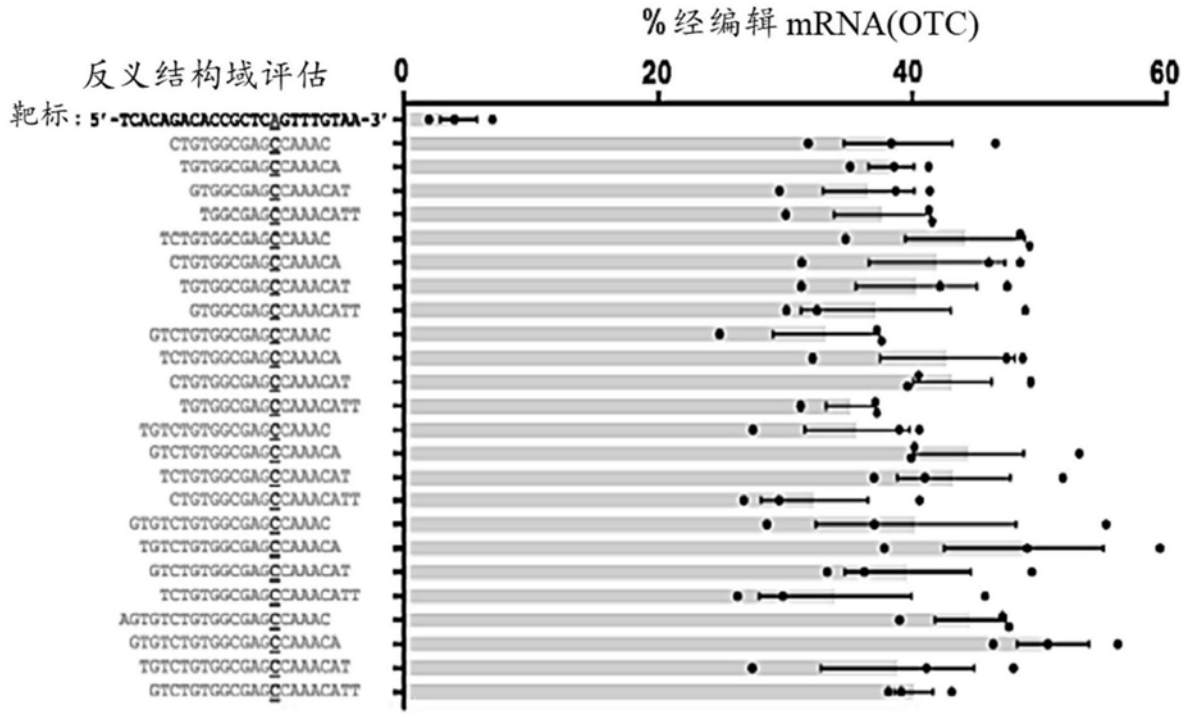

% 经编辑 mRNA(OTC)
反义结构域评估
0
20
40
60
靶标：5'-TCACAGACACCGCTCAGTTTGTAA-3'
CTGTGGCGAGCCAAAC
TGTGGCGAGCCAAACA
GTGGCGAGCCAAACAT
TGGCGAGCCAAACATT
TCTGTGGCGAGCCAAAC
CTGTGGCGAGCCAAACA
TGTGGCGAGCCAAACAT
GTGGCGAGCCAAACATT
GTCTGTGGCGAGCCAAAC
TCTGTGGCGAGCCAAACA
CTGTGGCGAGCCAAACAT
TGTGGCGAGCCAAACATT
TGTCTGTGGCGAGCCAAAC
GTCTGTGGCGAGCCAAACA
TCTGTGGCGAGCCAAACAT
CTGTGGCGAGCCAAACATT
GTGTCTGTGGCGAGCCAAAC
TGTCTGTGGCGAGCCAAACA
GTCTGTGGCGAGCCAAACAT
TCTGTGGCGAGCCAAACATT
AGTGTCTGTGGCGAGCCAAAC
GTGTCTGTGGCGAGCCAAACA
TGTCTGTGGCGAGCCAAACAT
GTCTGTGGCGAGCCAAACATT

图34A

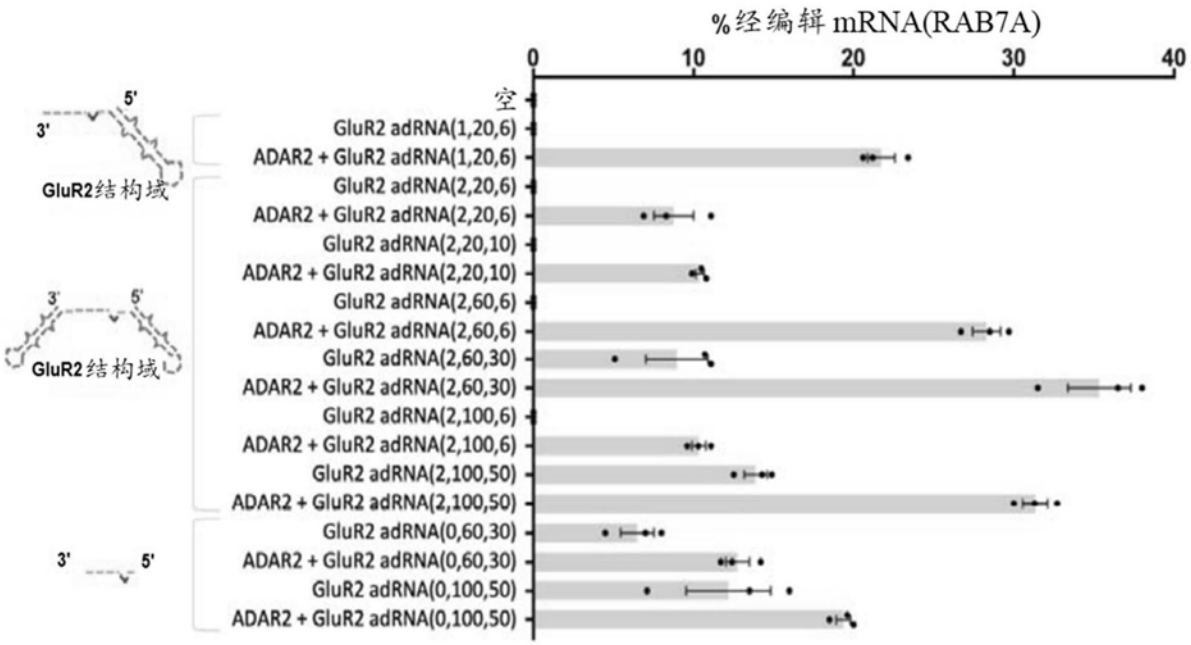

%经编辑 mRNA(RAB7A)
0
10
20
30
40
5'
3'
GluR2结构域
3'
5'
GluR2结构域
3'
5'
空
GluR2 adRNA(1,20,6)
ADAR2 + GluR2 adRNA(1,20,6)
GluR2 adRNA(2,20,6)
ADAR2 + GluR2 adRNA(2,20,6)
GluR2 adRNA(2,20,10)
ADAR2 + GluR2 adRNA(2,20,10)
GluR2 adRNA(2,60,6)
ADAR2 + GluR2 adRNA(2,60,6)
GluR2 adRNA(2,60,30)
ADAR2 + GluR2 adRNA(2,60,30)
GluR2 adRNA(2,100,6)
ADAR2 + GluR2 adRNA(2,100,6)
GluR2 adRNA(2,100,50)
ADAR2 + GluR2 adRNA(2,100,50)
GluR2 adRNA(0,60,30)
ADAR2 + GluR2 adRNA(0,60,30)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)

图34B

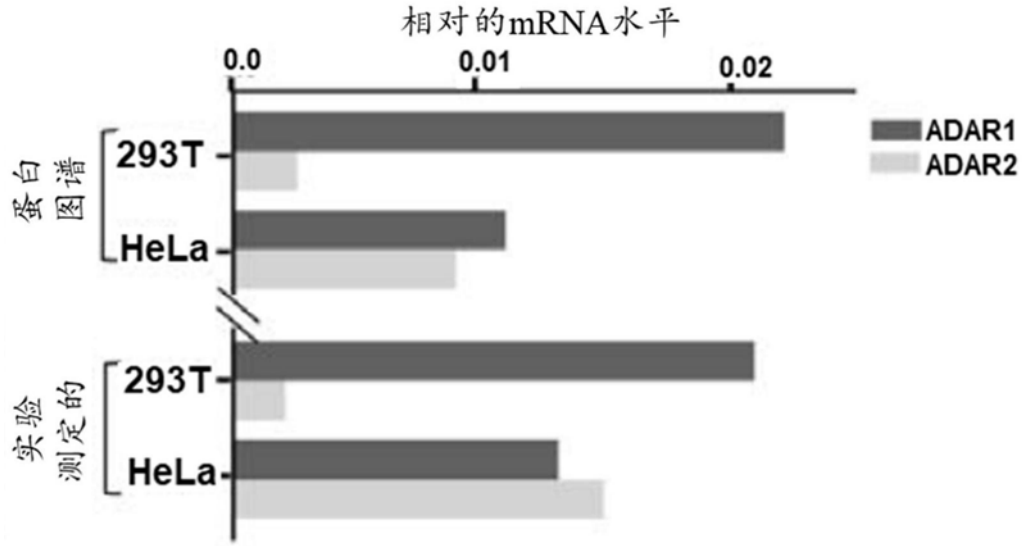

相对的mRNA水平
0.0
0.01
0.02
ADAR1
ADAR2
蛋白图谱
293T
HeLa
实验测定的
293T
HeLa

图34C

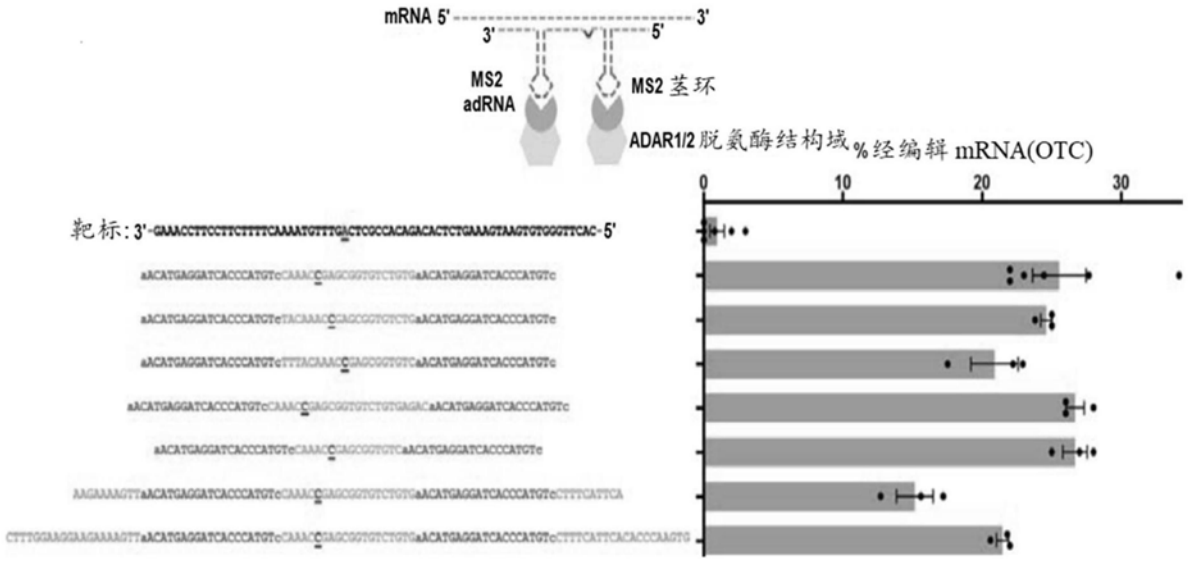

mRNA 5'
3'
3'
5'
MS2 adRNA
MS2 茎环
ADAR1/2 脱氨酶结构域
% 经编辑 mRNA(OTC)
0
10
20
30
靶标: 3'-GAAACCTTCCTTCTTTTCAAAATGTTTGACTCGCCACAGACACTCTGAAAGTAAGTGTGGGTTCAC-5'

图35A

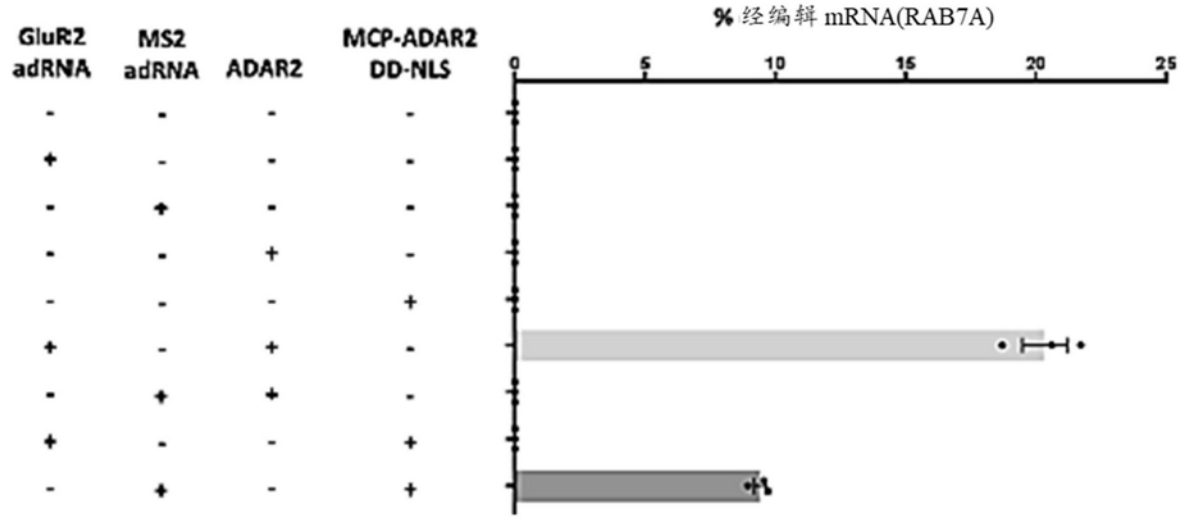
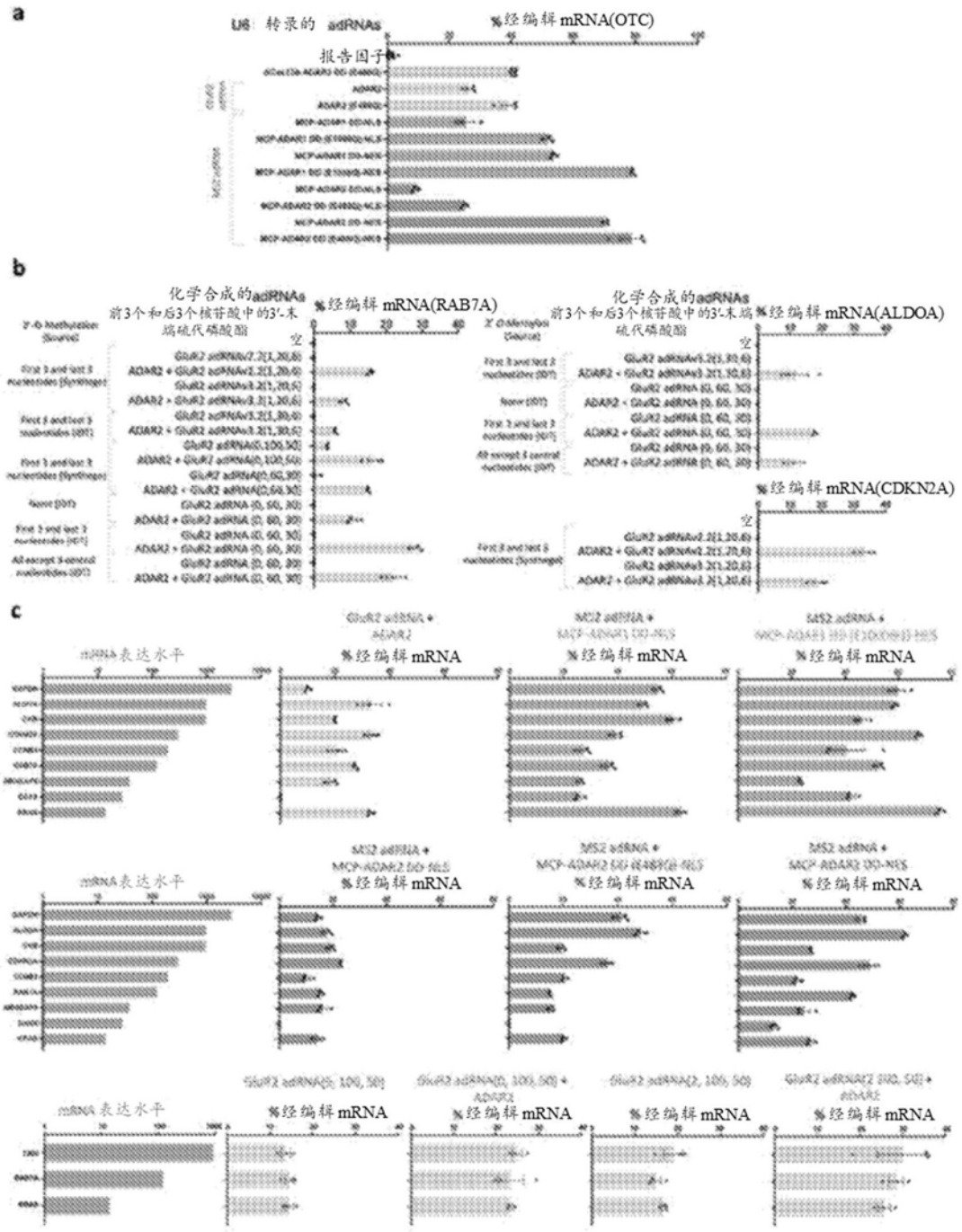

GluR2 adRNA
MS2 adRNA
ADAR2
MCP-ADAR2 DD-NLS
% 经编辑 mRNA(RAB7A)
0
5
10
15
20
25

图35B



补充图4：各种表达水平的一系列基因的编辑率的表征

图36A-C

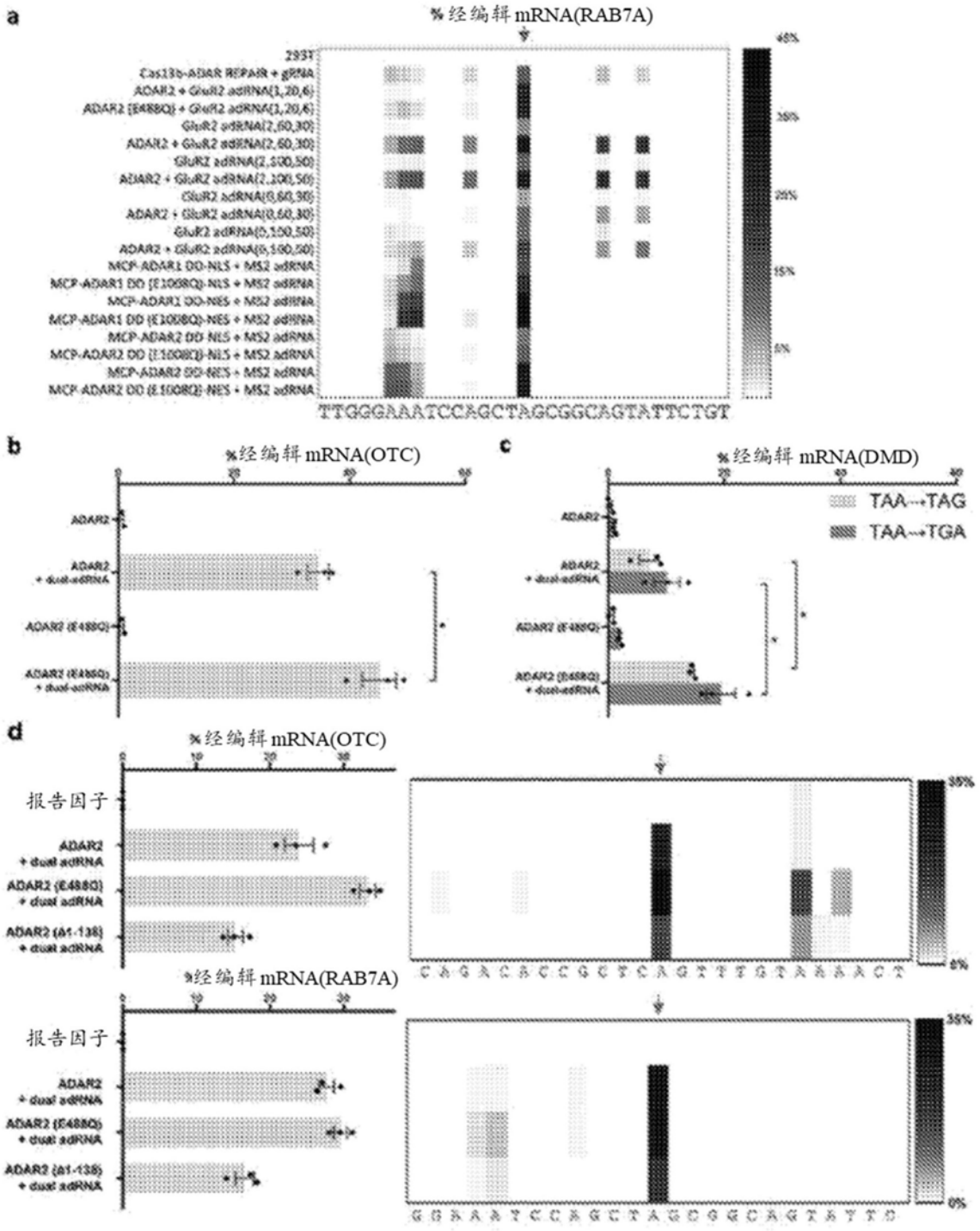

a
% 经编辑 mRNA(RAB7A)
293T
Cas13b-ADAR REPAIR + gRNA
ADAR2 + GluR2 adRNA(1,20,6)
ADAR2 (E488Q) + GluR2 adRNA(1,20,6)
GluR2 adRNA(2,60,30)
ADAR2 + GluR2 adRNA(2,60,30)
GluR2 adRNA(2,100,50)
ADAR2 + GluR2 adRNA(2,100,50)
GluR2 adRNA(0,60,30)
ADAR2 + GluR2 adRNA(0,60,30)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)
MCP-ADAR1 DD-NLS + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NLS + MS2 adRNA
MCP-ADAR1 DD-NES + MS2 adRNA
MCP-ADAR1 DD (E1008Q)-NES + MS2 adRNA
MCP-ADAR2 DD-NLS + MS2 adRNA
MCP-ADAR2 DD (E1008Q)-NLS + MS2 adRNA
MCP-ADAR2 DD-NES + MS2 adRNA
MCP-ADAR2 DD (E1008Q)-NES + MS2 adRNA
TTGGGAAATCCAGCTAGCGGCAGTATTCTGT
45%
30%
25%
15%
5%
b
% 经编辑 mRNA(OTC)
ADAR2
ADAR2 + dual-adRNA
ADAR2 (E488Q)
ADAR2 (E488Q) + dual-adRNA
c
% 经编辑 mRNA(DMD)
TAA→TAG
TAA→TGA
ADAR2
ADAR2 + dual-adRNA
ADAR2 (E488Q)
ADAR2 (E488Q) + dual-adRNA
d
% 经编辑 mRNA(OTC)
报告因子
ADAR2 + dual adRNA
ADAR2 (E488Q) + dual adRNA
ADAR2 (Δ1-138) + dual adRNA
% 经编辑 mRNA(RAB7A)
报告因子
ADAR2 + dual adRNA
ADAR2 (E488Q) + dual adRNA
ADAR2 (Δ1-138) + dual adRNA
35%
0%
C A G A C A C C G C T C A G T T T G T A A A A C T
35%
0%
G G A A A T C C A G C T A G C G G C A G T A T T C

图37A-D

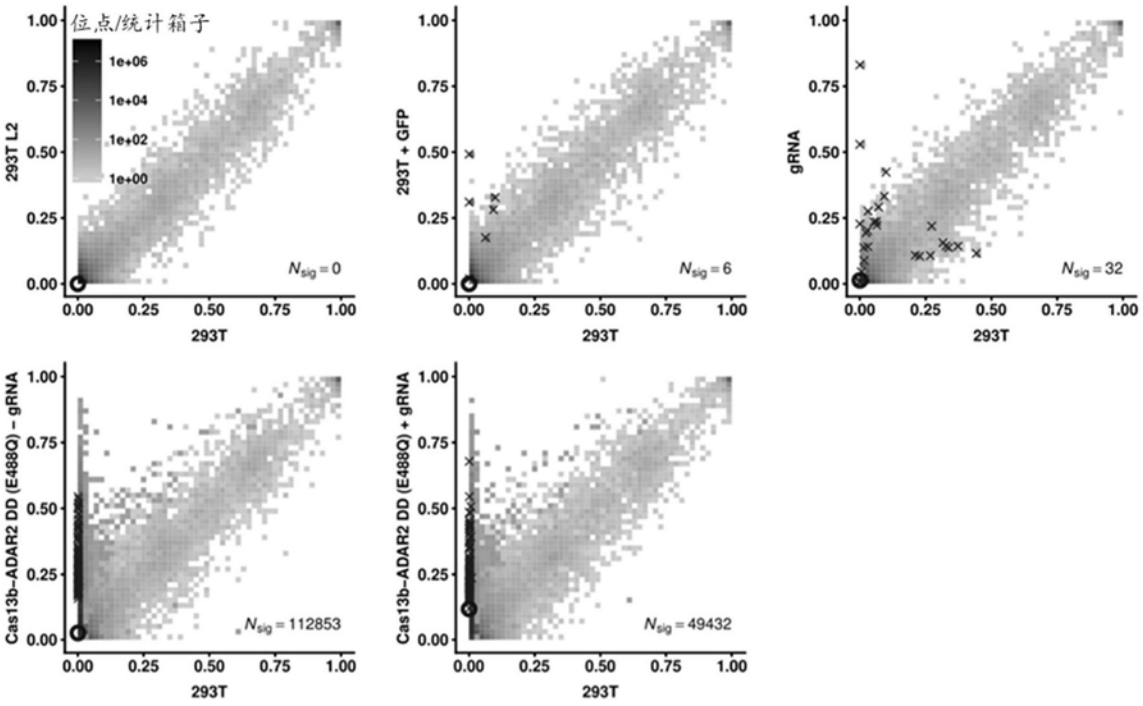

位点/统计箱子
1e+06
1e+04
1e+02
1e+00
293T L2
293T + GFP
gRNA
Cas13b–ADAR2 DD (E488Q) – gRNA
Cas13b–ADAR2 DD (E488Q) + gRNA
293T
$N_{sig} = 0$
$N_{sig} = 6$
$N_{sig} = 32$
$N_{sig} = 112853$
$N_{sig} = 49432$
0.00
0.25
0.50
0.75
1.00

图38

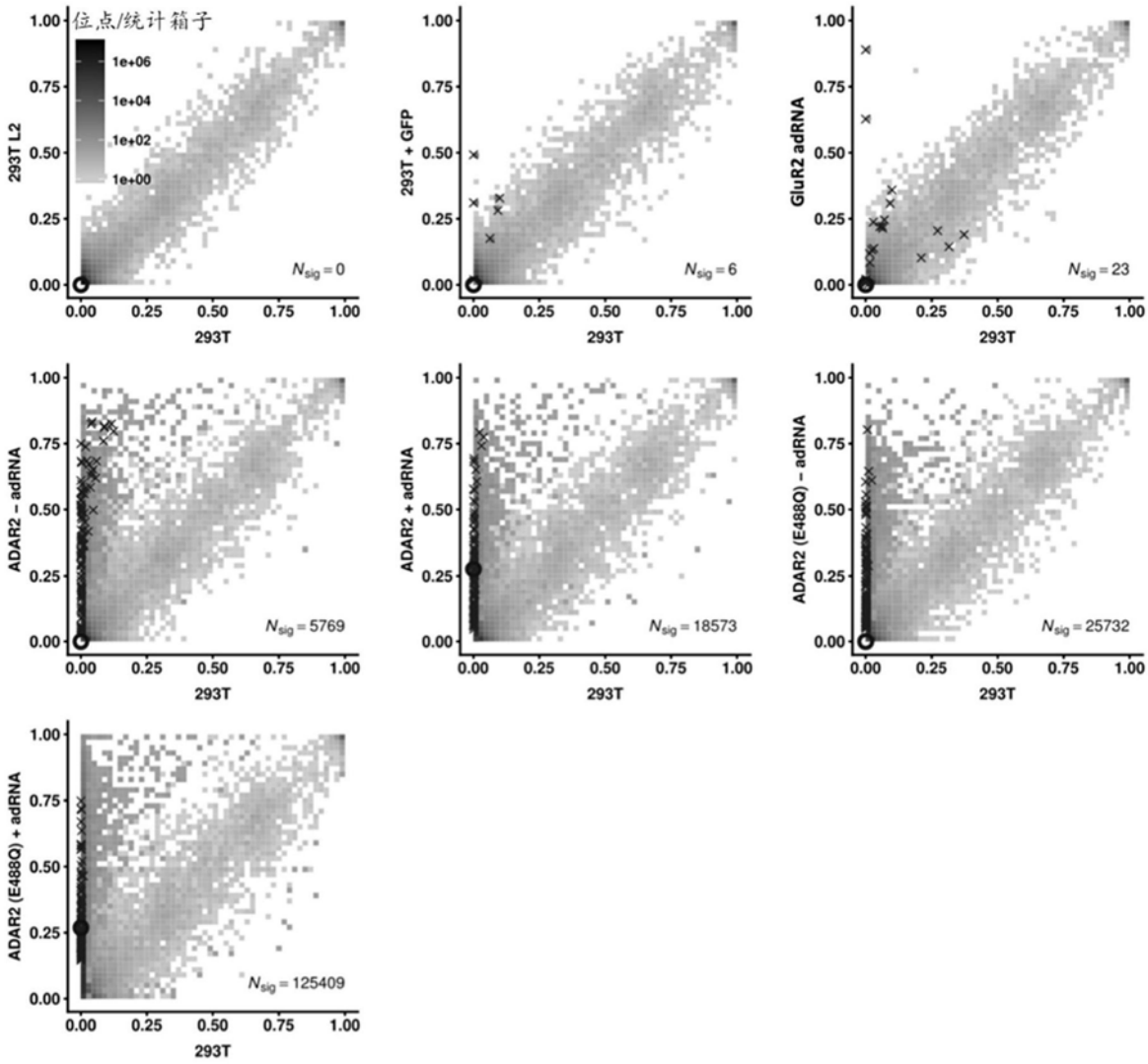

位点/统计箱子
1e+06
1e+04
1e+02
1e+00
293T L2
293T + GFP
GluR2 adRNA
ADAR2 – adRNA
ADAR2 + adRNA
ADAR2 (E488Q) – adRNA
ADAR2 (E488Q) + adRNA
293T
N_{sig} = 0
N_{sig} = 6
N_{sig} = 23
N_{sig} = 5769
N_{sig} = 18573
N_{sig} = 25732
N_{sig} = 125409
0.00
0.25
0.50
0.75
1.00

图39

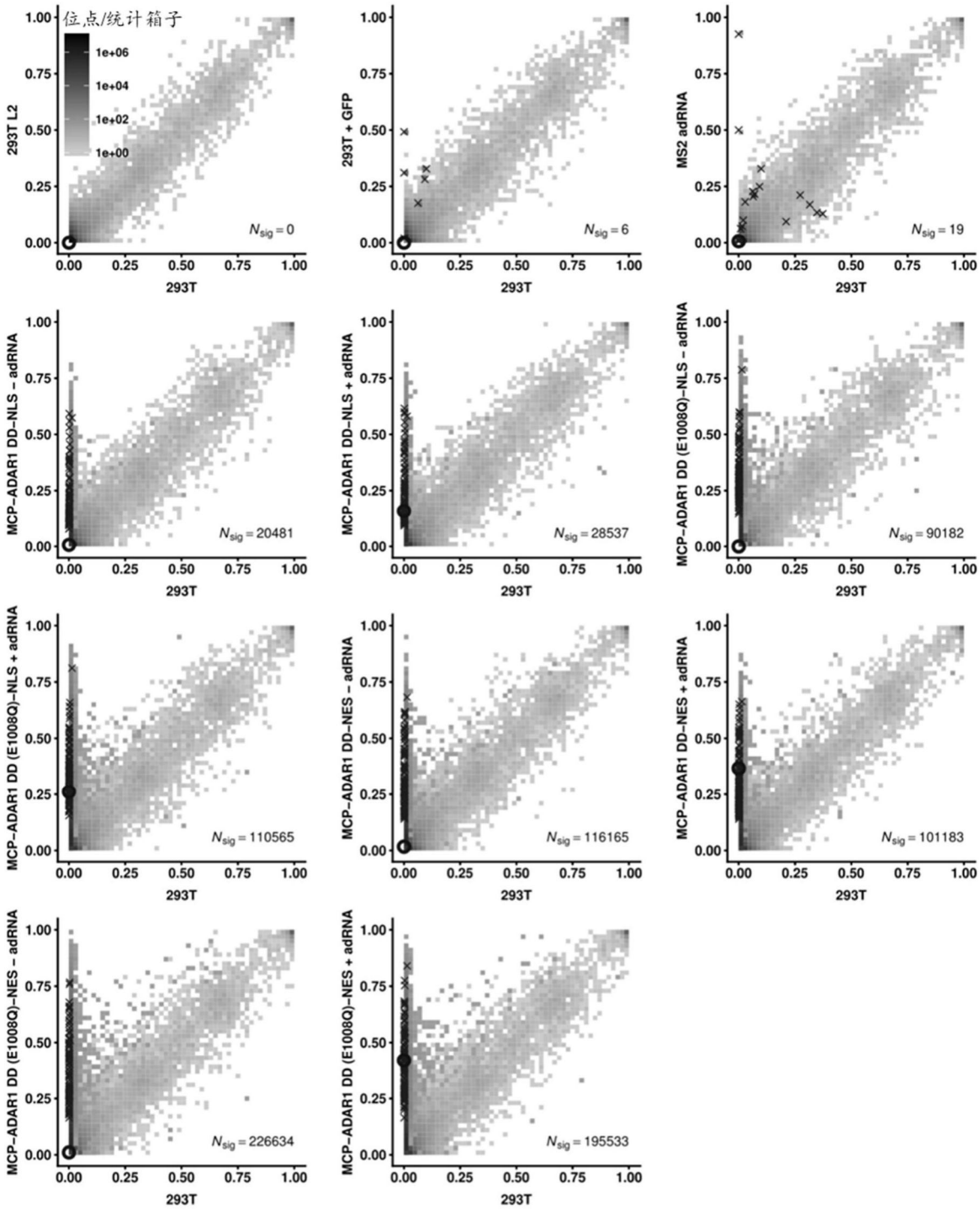

位点/统计箱子
1e+06
1e+04
1e+02
1e+00
293T L2
293T + GFP
MS2 adRNA
MCP–ADAR1 DD–NLS – adRNA
MCP–ADAR1 DD–NLS + adRNA
MCP–ADAR1 DD (E1008Q)–NLS – adRNA
MCP–ADAR1 DD (E1008Q)–NLS + adRNA
MCP–ADAR1 DD–NES – adRNA
MCP–ADAR1 DD–NES + adRNA
MCP–ADAR1 DD (E1008Q)–NES – adRNA
MCP–ADAR1 DD (E1008Q)–NES + adRNA
293T
$N_{sig} = 0$
$N_{sig} = 6$
$N_{sig} = 19$
$N_{sig} = 20481$
$N_{sig} = 28537$
$N_{sig} = 90182$
$N_{sig} = 110565$
$N_{sig} = 116165$
$N_{sig} = 101183$
$N_{sig} = 226634$
$N_{sig} = 195533$
0.00
0.25
0.50
0.75
1.00

图40

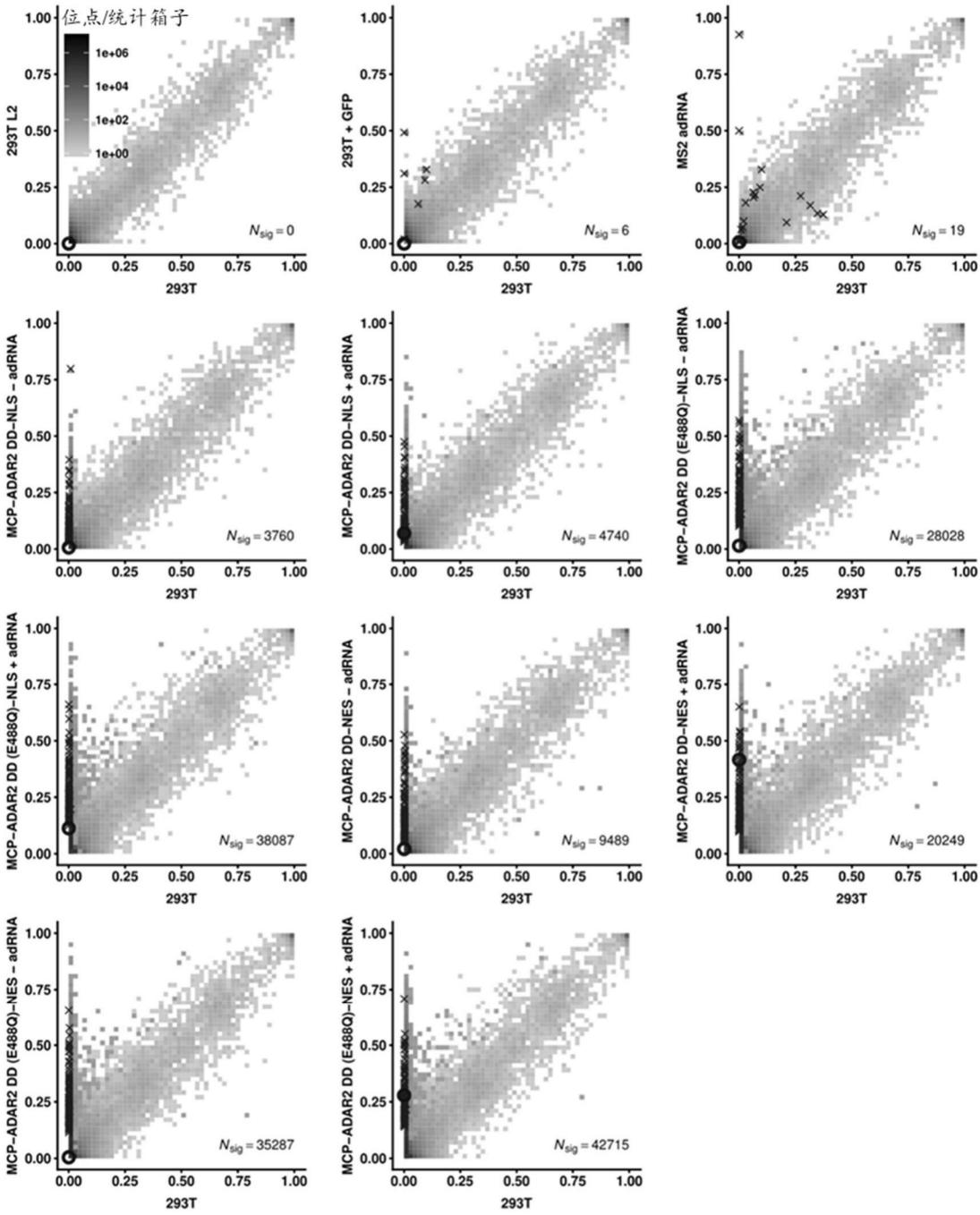

位点/统计箱子
1e+06
1e+04
1e+02
1e+00
293T L2
293T + GFP
MS2 adRNA
MCP–ADAR2 DD–NLS – adRNA
MCP–ADAR2 DD–NLS + adRNA
MCP–ADAR2 DD (E488Q)–NLS – adRNA
MCP–ADAR2 DD (E488Q)–NLS + adRNA
MCP–ADAR2 DD–NES – adRNA
MCP–ADAR2 DD–NES + adRNA
MCP–ADAR2 DD (E488Q)–NES – adRNA
MCP–ADAR2 DD (E488Q)–NES + adRNA
293T
$N_{sig} = 0$
$N_{sig} = 6$
$N_{sig} = 19$
$N_{sig} = 3760$
$N_{sig} = 4740$
$N_{sig} = 28028$
$N_{sig} = 38087$
$N_{sig} = 9489$
$N_{sig} = 20249$
$N_{sig} = 35287$
$N_{sig} = 42715$

图41

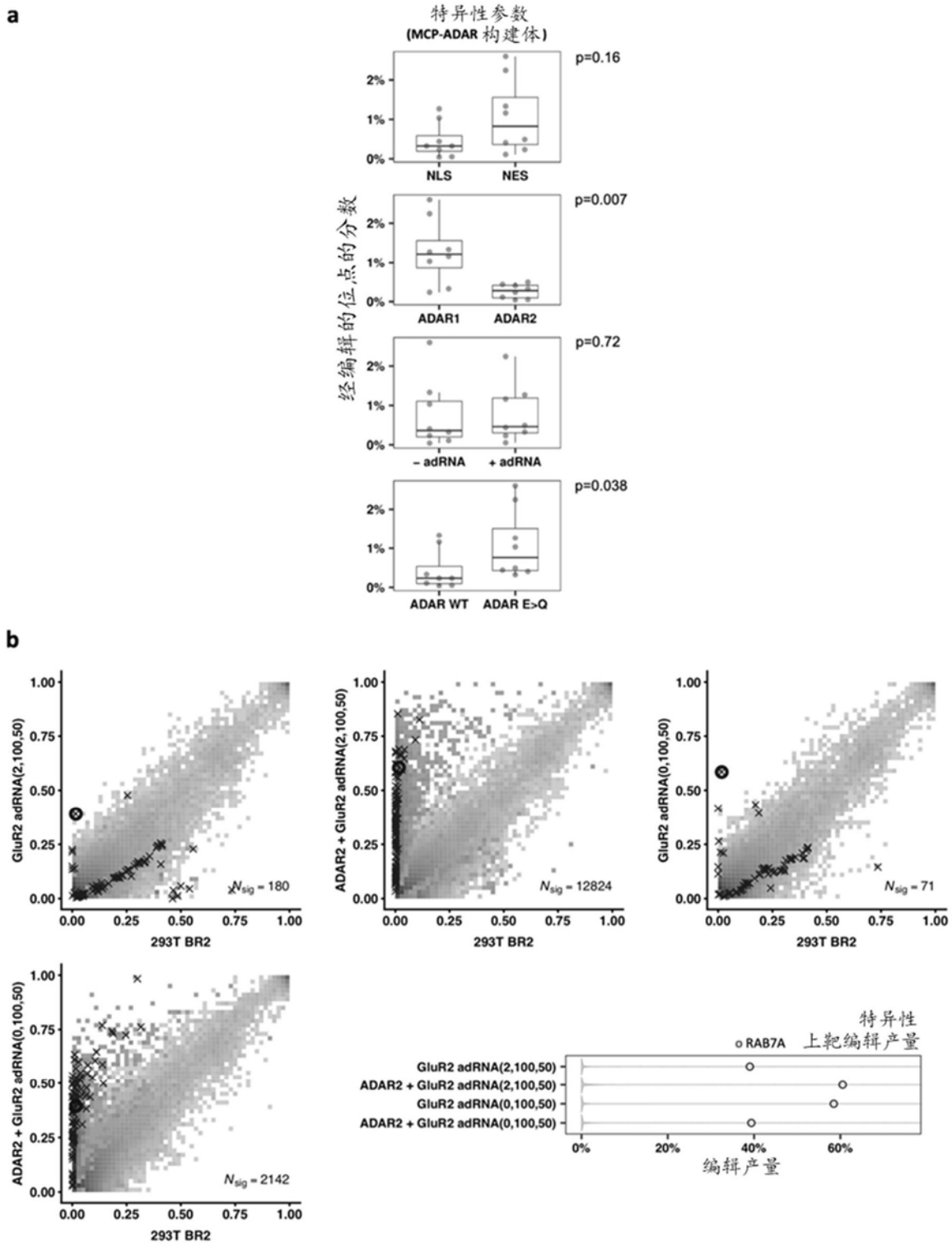

a
特异性参数
(MCP-ADAR 构建体)
经编辑的位点的分数
p=0.16
NLS
NES
p=0.007
ADAR1
ADAR2
p=0.72
– adRNA
+ adRNA
p=0.038
ADAR WT
ADAR E>Q
b
GluR2 adRNA(2,100,50)
293T BR2
N_{sig} = 180
ADAR2 + GluR2 adRNA(2,100,50)
N_{sig} = 12824
GluR2 adRNA(0,100,50)
N_{sig} = 71
ADAR2 + GluR2 adRNA(0,100,50)
N_{sig} = 2142
特异性
上靶编辑产量
RAB7A
GluR2 adRNA(2,100,50)
ADAR2 + GluR2 adRNA(2,100,50)
GluR2 adRNA(0,100,50)
ADAR2 + GluR2 adRNA(0,100,50)
编辑产量

图42A-B

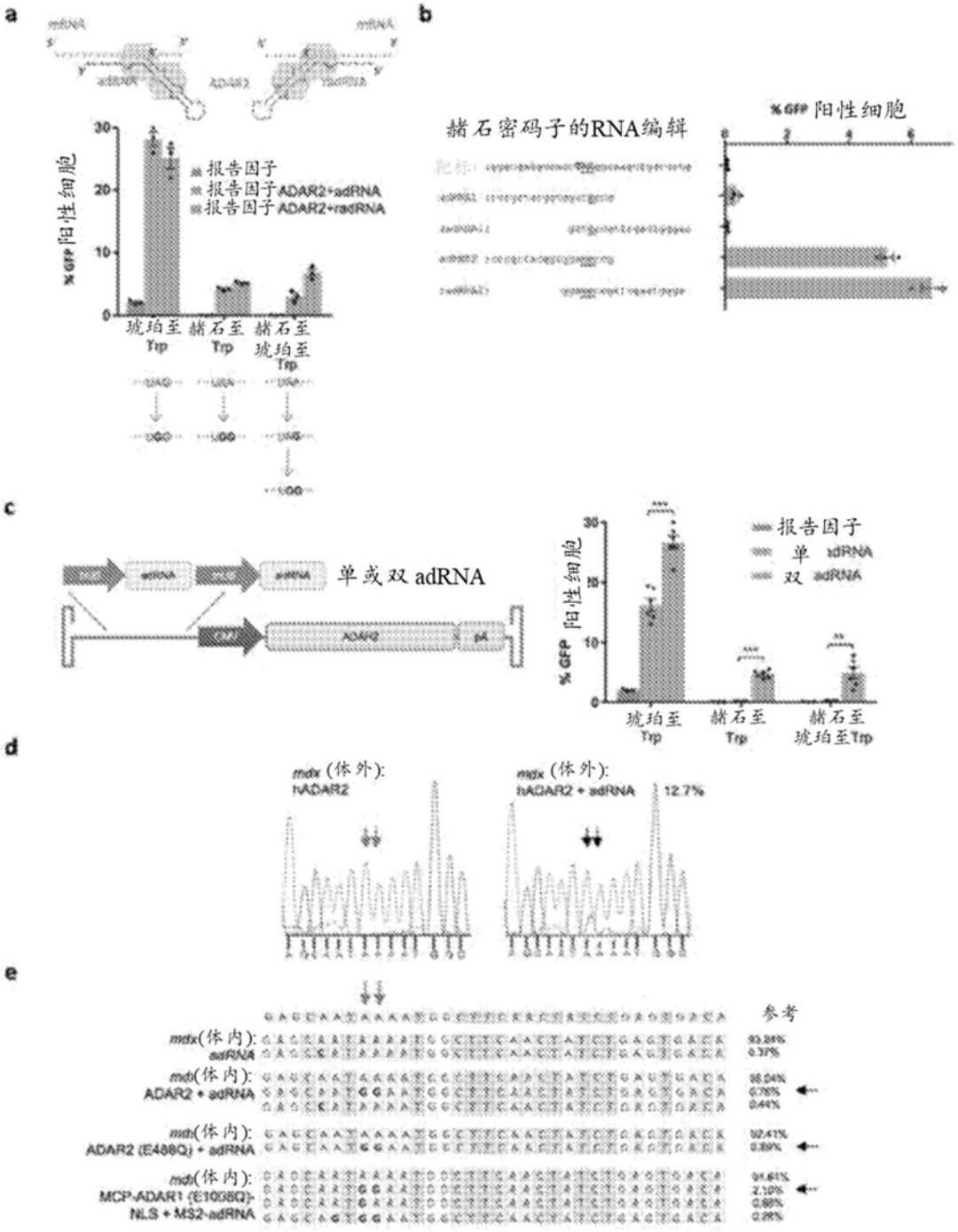

a
mRNA
adRNA
ADAR2
% GFP 阳性细胞
报告因子
报告因子+ADAR2+adRNA
报告因子+ADAR2+radRNA
琥珀至 Trp
赭石至 Trp
赭石至 琥珀至 Trp
b
赭石密码子的RNA编辑
靶标:
% GFP 阳性细胞
c
单或双 adRNA
ADAR2
pA
% GFP 阳性细胞
报告因子
单 adRNA
双 adRNA
琥珀至 Trp
赭石至 Trp
赭石至 琥珀至Trp
d
mdx (体外): hADAR2
mdx (体外): hADAR2 + adRNA
12.7%
e
参考
mdx(体内): adRNA
mdx(体内): ADAR2 + adRNA
mdx(体内): ADAR2 (E488Q) + adRNA
mdx(体内): MCP-ADAR1 (E1008Q)-NLS + MS2-adRNA

图43A-E

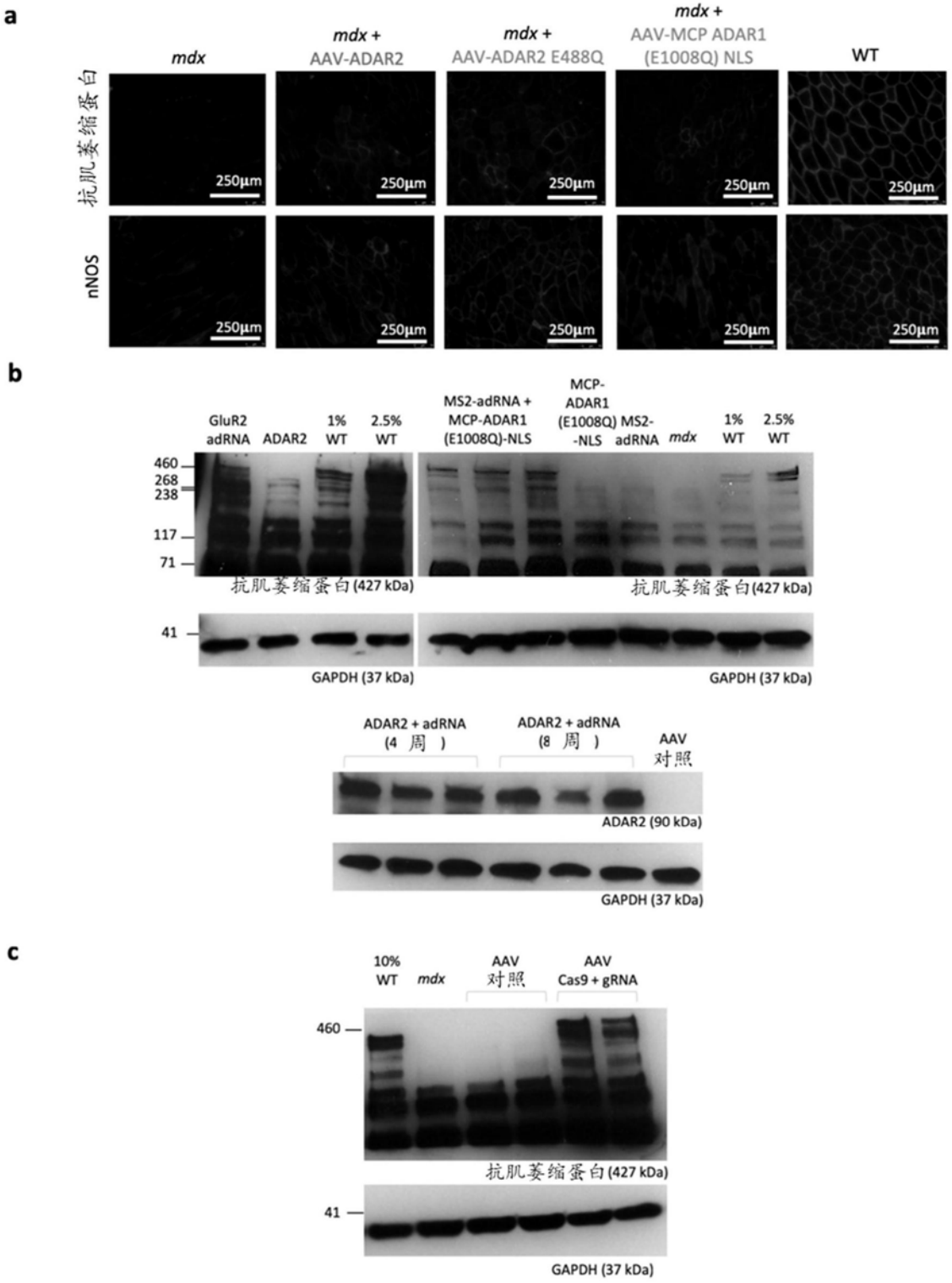

a
mdx
mdx + AAV-ADAR2
mdx + AAV-ADAR2 E488Q
mdx + AAV-MCP ADAR1 (E1008Q) NLS
WT
抗肌萎缩蛋白
nNOS
250μm
b
GluR2 adRNA
ADAR2
1% WT
2.5% WT
MS2-adRNA + MCP-ADAR1 (E1008Q)-NLS
MCP-ADAR1 (E1008Q) -NLS
MS2-adRNA
mdx
460
268
238
117
71
抗肌萎缩蛋白(427 kDa)
41
GAPDH (37 kDa)
ADAR2 + adRNA (4 周)
ADAR2 + adRNA (8 周)
AAV 对照
ADAR2 (90 kDa)
c
10% WT
mdx
AAV 对照
AAV Cas9 + gRNA
460
抗肌萎缩蛋白(427 kDa)
41
GAPDH (37 kDa)

图44A-C

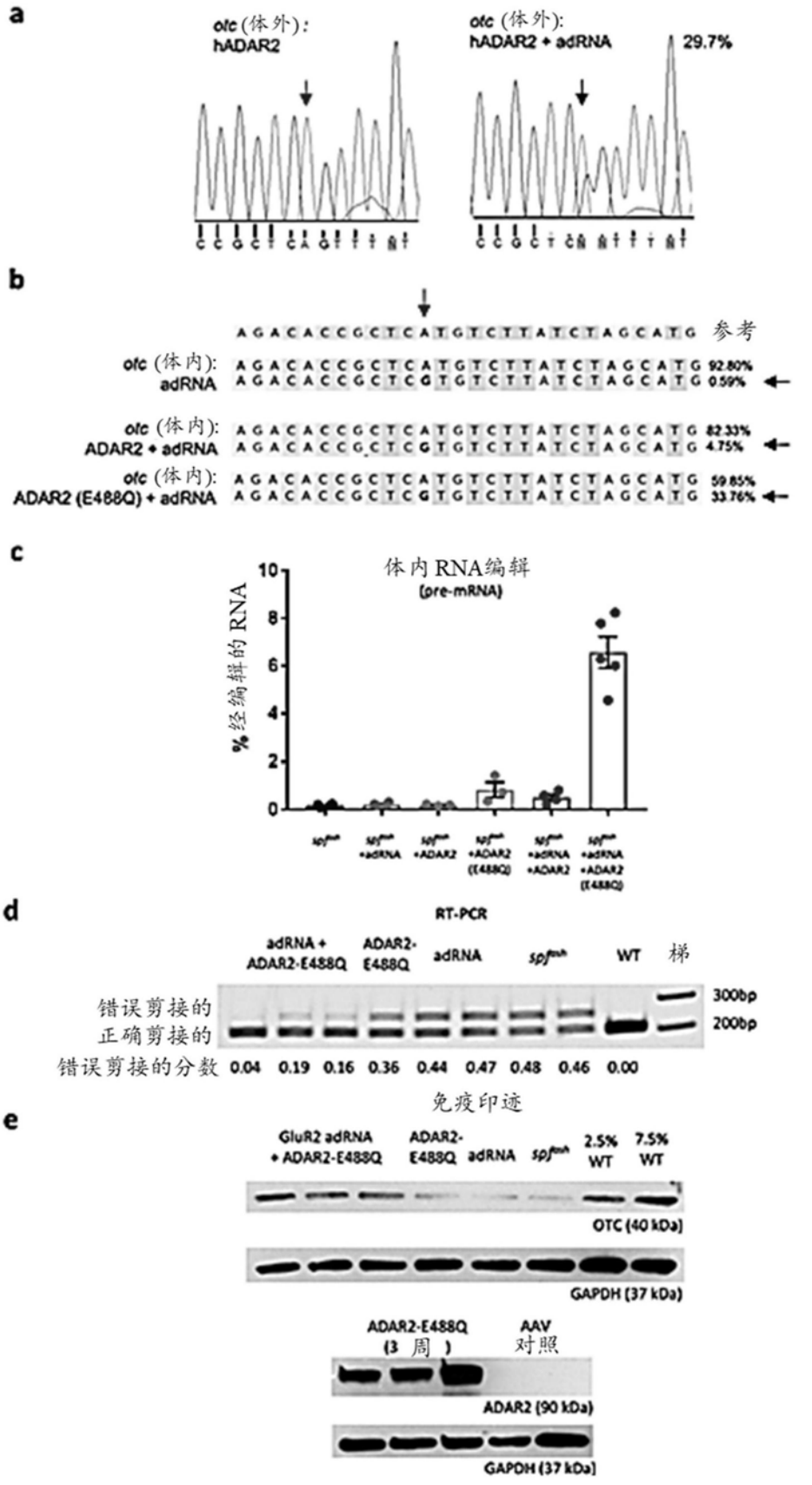
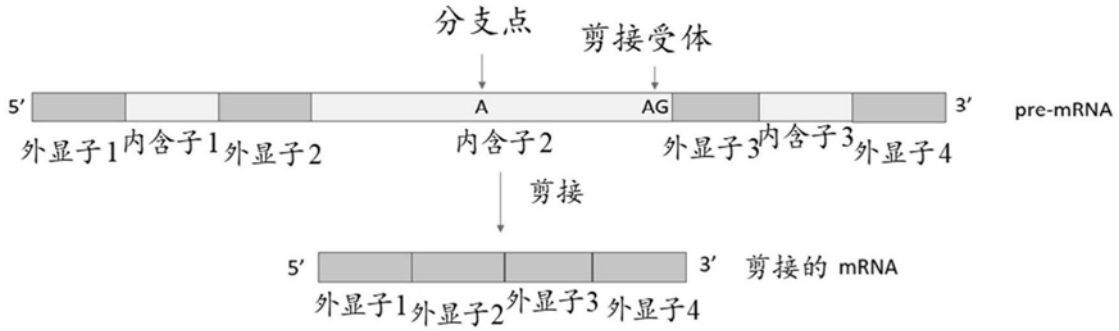

a
otc (体外):
hADAR2
otc (体外):
hADAR2 + adRNA
29.7%
b
参考
otc (体内):
adRNA
92.80%
0.59%
otc (体内):
ADAR2 + adRNA
82.33%
4.75%
otc (体内):
ADAR2 (E488Q) + adRNA
59.85%
33.76%
c
体内 RNA编辑
(pre-mRNA)
% 经编辑的 RNA
d
RT-PCR
adRNA + ADAR2-E488Q
ADAR2-E488Q
adRNA
WT
梯
300bp
200bp
错误剪接的
正确剪接的
错误剪接的分数 0.04 0.19 0.16 0.36 0.44 0.47 0.48 0.46 0.00
e
免疫印迹
GluR2 adRNA + ADAR2-E488Q
ADAR2-E488Q
adRNA
2.5% WT
7.5% WT
OTC (40 kDa)
GAPDH (37 kDa)
ADAR2-E488Q (3 周)
AAV 对照
ADAR2 (90 kDa)
GAPDH (37 kDa)

图45A-E

微扰	递送途径	% 患病小鼠 (n=总小鼠 ;测定 = 3 周)
盐水	眼眶后	0% (n=8)
盐水	眼眶后	0% (n=13)
盐水	眼眶后	50% (n=14)
盐水	肌肉内	0% (n=9)
盐水	肌肉内	6% (n=18)
盐水	肌肉内	17% (n=6)
盐水	肌肉内	9% (n=11)
盐水	肌肉内	0% (n=3)

图46

正确剪接:



经由在剪接受体位点和/或分支点处创建A->G突变来进行外显子跳跃:

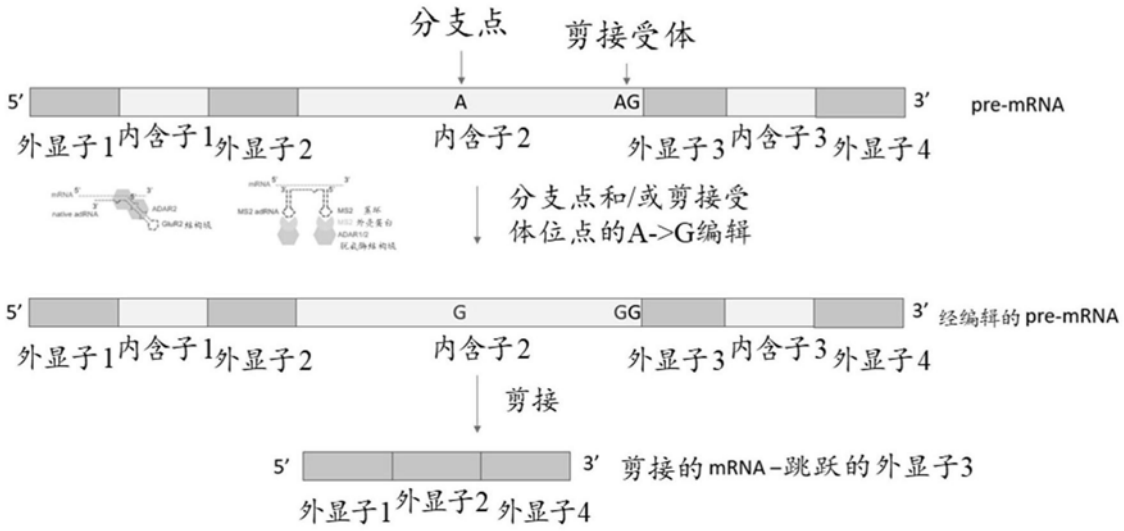


图47

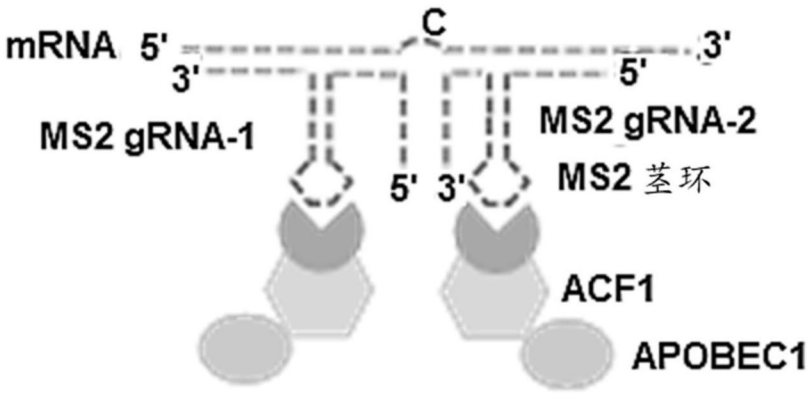


图48



图49A



图49B



图49C

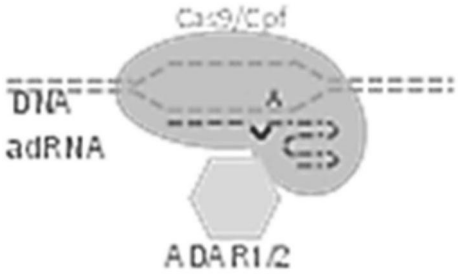


图49D

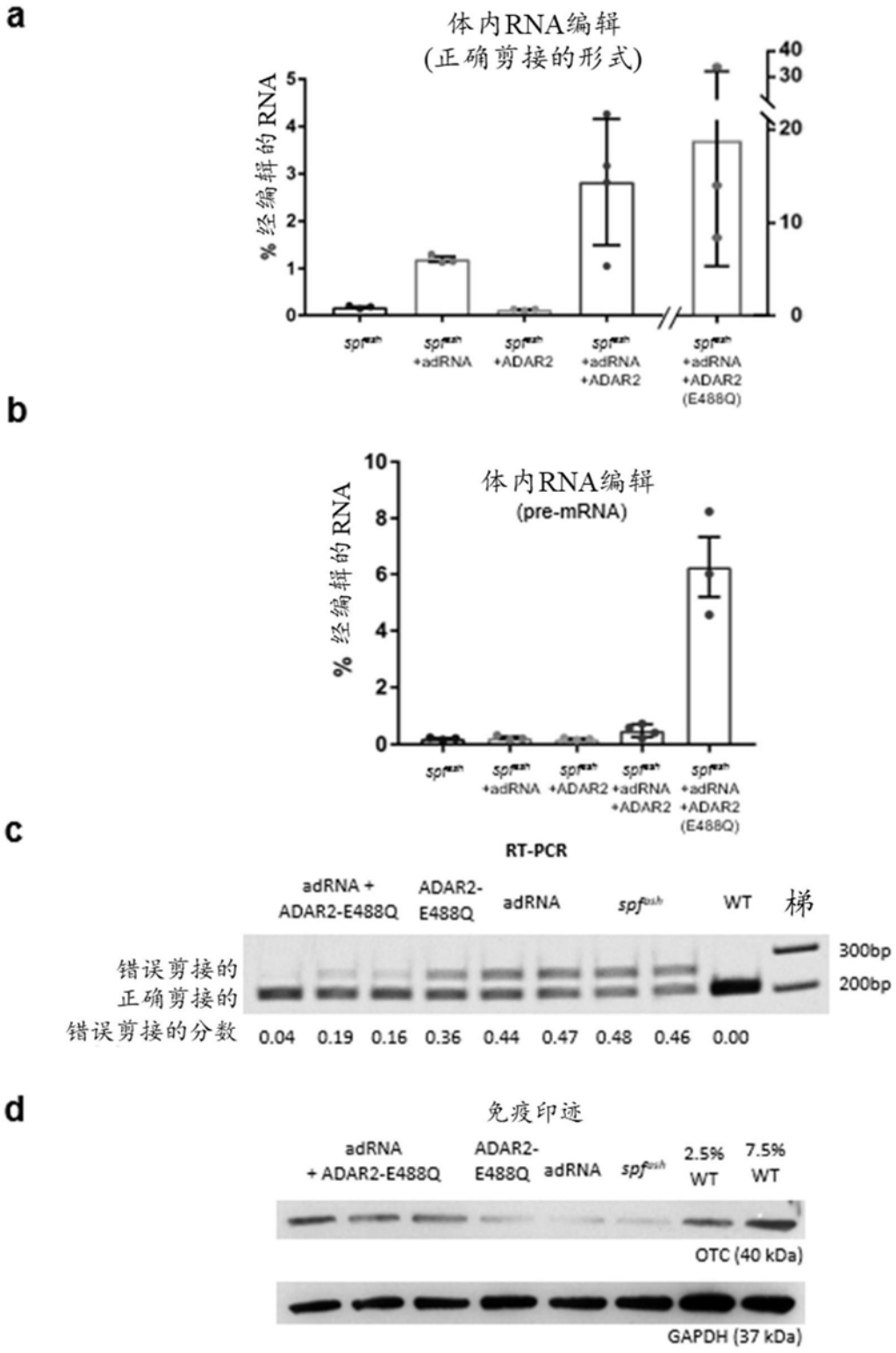

a
体内RNA编辑
(正确剪接的形式)
% 经编辑的 RNA
+adRNA
+ADAR2
+adRNA
+ADAR2
+adRNA
+ADAR2
(E488Q)
b
体内RNA编辑
(pre-mRNA)
% 经编辑的 RNA
+adRNA
+ADAR2
+adRNA
+ADAR2
+adRNA
+ADAR2
(E488Q)
c
RT-PCR
adRNA + ADAR2-E488Q
ADAR2-E488Q
adRNA
WT
梯
300bp
200bp
错误剪接的
正确剪接的
错误剪接的分数
0.04 0.19 0.16 0.36 0.44 0.47 0.48 0.46 0.00
d
免疫印迹
adRNA + ADAR2-E488Q
ADAR2-E488Q
adRNA
2.5% WT
7.5% WT
OTC (40 kDa)
GAPDH (37 kDa)

图50A-D

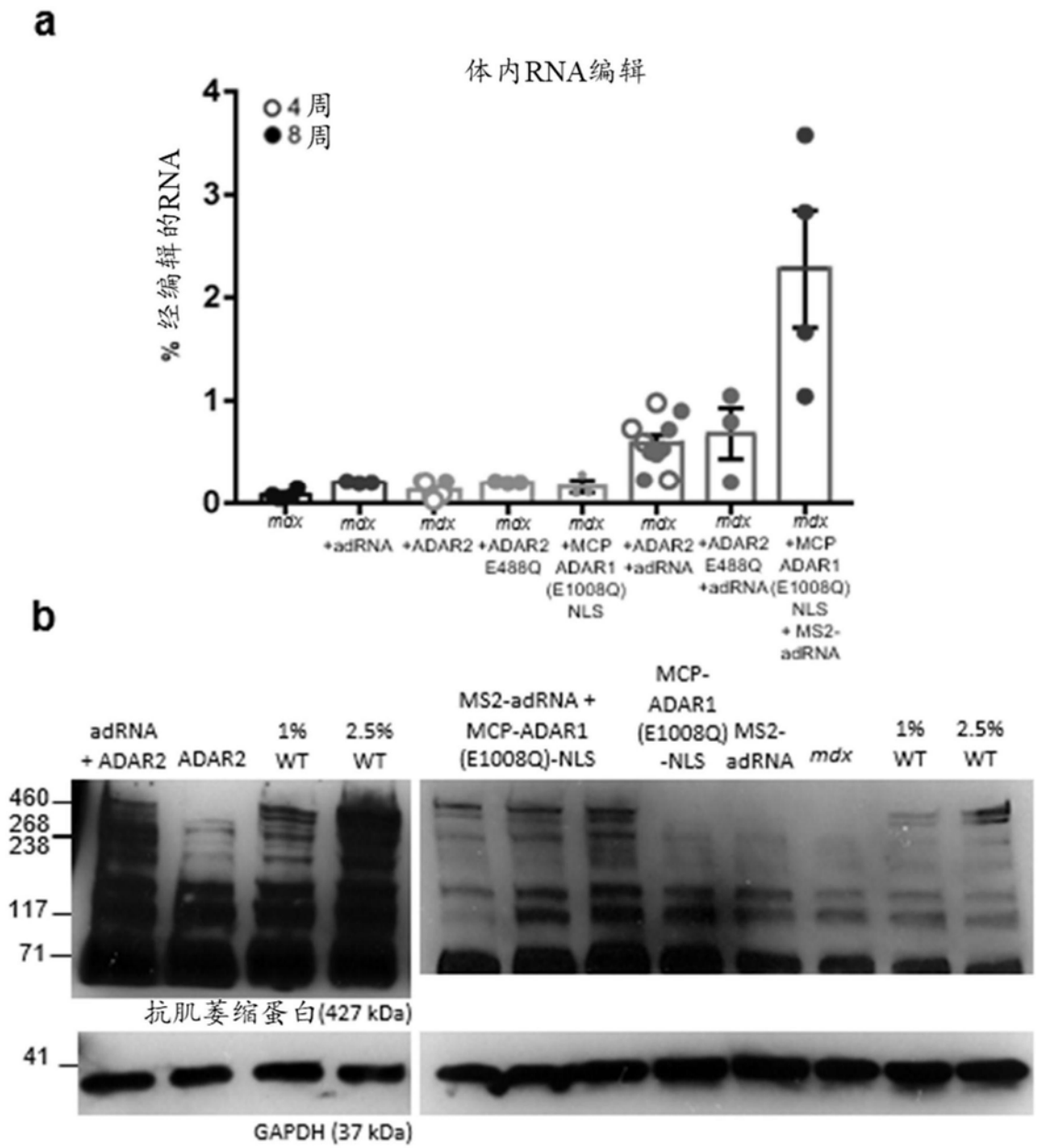

a
体内RNA编辑
4周
8周
% 经编辑的RNA
0
1
2
3
4
mdx
mdx +adRNA
mdx +ADAR2
mdx +ADAR2 E488Q
mdx +MCP ADAR1 (E1008Q) NLS
mdx +ADAR2 +adRNA
mdx +ADAR2 E488Q +adRNA
mdx +MCP ADAR1 (E1008Q) NLS + MS2-adRNA
b
adRNA + ADAR2
ADAR2
1% WT
2.5% WT
MS2-adRNA + MCP-ADAR1 (E1008Q)-NLS
MCP-ADAR1 (E1008Q) -NLS
MS2-adRNA
mdx
1% WT
2.5% WT
460
268
238
117
71
抗肌萎缩蛋白(427 kDa)
41
GAPDH (37 kDa)

图51A-B

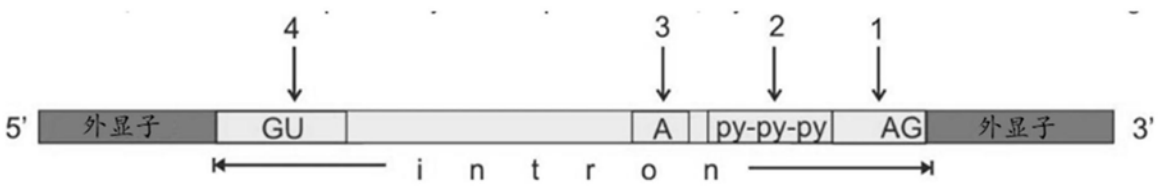

4
3
2
1
5'
外显子
GU
A
py-py-py
AG
外显子
3'
intron

图52