



(12)发明专利

(10)授权公告号 CN 105765070 B

(45)授权公告日 2020.07.10

(21)申请号 201480058811.5

(72)发明人 D.R.刘 J.H.胡

(22)申请日 2014.09.05

(74)专利代理机构 北京市柳沈律师事务所

(65)同一申请的已公布的文献号

申请公布号 CN 105765070 A

11105

(43)申请公布日 2016.07.13

代理人 张文辉

(30)优先权数据

(51)Int.Cl.

C12N 15/113(2006.01)

61/874,682 2013.09.06 US

(56)对比文件

US 6057153 A,2000.05.02,

14/326,329 2014.07.08 US

US 2009/0234109 A1,2009.09.17,

14/326,340 2014.07.08 US

WO 2006/042112 A2,2006.04.20,

14/326,361 2014.07.08 US

Mali等.CAS9 transcriptional

(85)PCT国际申请进入国家阶段日

activators for target specificity

2016.04.26

screening and paired nickases for

(86)PCT国际申请的申请数据

cooperative genome engineering.《nature

PCT/US2014/054252 2014.09.05

biotechnology》.2013,第31卷(第9期),图1.

(87)PCT国际申请的公布数据

审查员 张锦广

W02015/035139 EN 2015.03.12

(73)专利权人 哈佛大学的校长及成员们

权利要求书4页 说明书25页

地址 美国马萨诸塞州

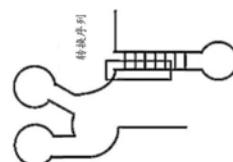
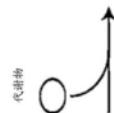
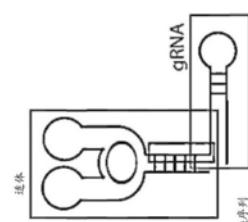
序列表10页 附图12页

(54)发明名称

可变换CAS9核酸酶及其用途

(57)摘要

本公开内容的一些方面提供了用于控制RNA可编程内切核酸酶(诸如Cas9)的活性和/或改善RNA可编程内切核酸酶(诸如Cas9)的特异性的组合物、方法、系统和试剂盒。例如,提供了工程化改造为以“开启”或“关闭”状态存在的引导RNA(gRNA),其控制RNA可编程内切核酸酶的结合以及因此切割活性。本公开内容的一些方面提供了mRNA感测性gRNA,其基于靶mRNA的存在或缺乏调控RNA-可编程内切核酸酶的活性。本公开内容的一些方面提供了gRNA,其基于延长的DNA(xDNA)的存在或缺乏调控RNA-可编程内切核酸酶的活性。



1. 一种复合物, 其包含:

(i) 单一引导RNA (sgRNA), 所述单一引导RNA包含

(1) sgRNA主链序列;

(2) RNA适体; 和

(3) 与DNA靶位点互补的引导序列,

其中(3)的引导序列在一侧上侧翼有(2)的RNA适体和在另一侧上侧翼有(1)的sgRNA主链序列,

其中所述sgRNA在缺乏与所述适体结合的特定配体的情况下不与所述DNA靶位点杂交; 和

(ii) Cas9蛋白。

2. 权利要求1的复合物, 其中所述RNA适体包含结合所述引导序列的序列。

3. 权利要求1的复合物, 其中所述RNA适体包含与所述引导序列互补的至少4个连续核苷酸的序列。

4. 权利要求1的复合物, 其中所述RNA适体由所述特定配体结合。

5. 权利要求4的复合物, 其中所述特定配体是小分子、代谢物、肽或核酸。

6. 权利要求5的复合物, 其中所述sgRNA:配体:Cas9复合物结合所述DNA靶位点。

7. 权利要求1-6中任一项的复合物, 其中所述RNA适体自核开关 (riboswitch) 衍生。

8. 权利要求7的复合物, 其中衍生所述RNA适体的所述核开关选自茶碱核开关、焦磷酸硫胺素 (TPP) 核开关、腺昔钴胺素 (AdoCbl) 核开关、S-腺昔甲硫氨酸 (SAM) 核开关、SAH核开关、黄素单核苷酸 (FMN) 核开关、四氢叶酸核开关、赖氨酸核开关、甘氨酸核开关、嘌呤核开关、G1mS核开关或前Q核苷 (pre-queosine) 1 (PreQ1) 核开关。

9. 权利要求8的复合物, 其中所述RNA适体自茶碱核开关衍生, 并且包含SEQ ID NO:3。

10. 权利要求1-6中任一项的复合物, 其中所述RNA适体是非天然存在的。

11. 权利要求10的复合物, 其中使用指数富集配体系统进化 (systematic evolution of ligands by exponential enrichment, SELEX) 平台将所述RNA适体工程化改造为结合所述特定配体。

12. 权利要求1-6中任一项的复合物, 其中所述sgRNA的非RNA适体部分包含至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140或至少150个核苷酸, 并且所述适体包含至少20、至少30、至少40、至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140、至少150、至少175、至少200、至少250或至少300个核苷酸。

13. 单一引导RNA (sgRNA), 其包含:

(1) sgRNA主链序列;

(2) RNA适体; 和

(3) 与DNA靶位点互补的引导序列,

其中(3)的引导序列在一侧上侧翼有(2)的RNA适体和在另一侧上侧翼有(1)的sgRNA主链序列, 和

其中所述sgRNA在缺乏与所述适体结合的特定配体的情况下不与DNA靶位点杂交。

14. 权利要求13的sgRNA, 其中所述RNA适体包含结合所述引导序列的序列。

15. 权利要求14的sgRNA,其中所述RNA适体包含与所述引导序列互补的至少4个连续核苷酸的序列。

16. 权利要求15的sgRNA,其中所述sgRNA在缺乏与所述RNA适体结合的所述特定配体的情况下不结合Cas9。

17. 权利要求15的sgRNA,其中在所述RNA适体被所述特定配体结合时,所述sgRNA结合Cas9。

18. 权利要求15的sgRNA,其中所述sgRNA在缺乏或存在与所述RNA适体结合的特定配体的情况下结合Cas9。

19. 权利要求13-18中任一项的sgRNA,其中所述特定配体是小分子、代谢物、肽或核酸。

20. 权利要求19的sgRNA,其中所述RNA适体自核开关(riboswitch)衍生。

21. 权利要求20的sgRNA,其中衍生所述RNA适体的所述核开关选自茶碱核开关、焦磷酸硫胺素(TPP)核开关、腺昔钴胺素(AdoCbl)核开关、S-腺昔甲硫氨酸(SAM)核开关、SAH核开关、黄素单核苷酸(FMN)核开关、四氢叶酸核开关、赖氨酸核开关、甘氨酸核开关、嘌呤核开关、G1mS核开关或前Q核昔1(PreQ1)核开关。

22. 权利要求21的sgRNA,其中所述RNA适体自茶碱核开关衍生,并且包含SEQ ID NO:3。

23. 权利要求22的sgRNA,其中所述RNA适体是非天然存在的。

24. 权利要求23的sgRNA,其中使用指数富集配体系统进化(systematic evolution of ligands by exponential enrichment, SELEX)平台将所述RNA适体工程化改造为结合所述特定配体。

25. 权利要求13-18中任一项的sgRNA,其中所述sgRNA的非RNA适体部分包含至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140或至少150个核苷酸,并且所述适体包含至少20、至少30、至少40、至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140、至少150、至少175、至少200、至少250或至少300个核苷酸。

26. 一种多核苷酸,其编码权利要求13-25中任一项的sgRNA。

27. 一种载体,其包含权利要求26的多核苷酸。

28. 一种用于重组表达的载体,其包含编码权利要求13-25中任一项的sgRNA的多核苷酸。

29. 一种细胞,其包含用于表达权利要求13-25中任一项的sgRNA的遗传构建体。

30. 权利要求29的细胞,其中所述细胞表达Cas9蛋白。

31. 一种用于位点特异性DNA切割的体外方法,其包括在Cas9蛋白切割DNA的条件下使DNA与复合物接触,所述复合物包含(i)权利要求13-25中任一项的sgRNA,其中所述DNA包含DNA靶位点,(ii)与所述sgRNA的RNA适体结合的特定配体,和(iii)Cas9蛋白。

32. 权利要求31的体外方法,其中所述DNA在细胞中。

33. 权利要求32的体外方法,其中所述细胞是真核细胞。

34. 一种用于诱导细胞中的位点特异性DNA切割的体外方法,其包括:

(a)使细胞接触或在细胞内表达权利要求13-25中任一项的sgRNA,其中所述细胞包含含有DNA靶位点的DNA;

(b)使细胞接触或在细胞内表达Cas9蛋白;并且

(c) 使所述细胞与所述特定配体接触,导致切割所述DNA的sgRNA:配体:Cas9复合物的形成。

35. 一种用于诱导细胞中的位点特异性DNA切割的体外方法,其包括:

(a) 使细胞与包含Cas9蛋白和权利要求13-25中任一项的sgRNA的复合物接触,其中所述细胞包含含有DNA靶位点的DNA,并且

(b) 使细胞与所述特定配体接触,导致切割所述DNA的sgRNA:配体:Cas9复合物的形成。

36. 权利要求35的体外方法,其中同时或以任何次序序贯实施步骤(a)和(b)。

37. 权利要求34-36中任一项权利要求的体外方法,其中所述细胞是真核细胞。

38. 一种试剂盒,其包含权利要求13-25中任一项的sgRNA。

39. 一种试剂盒,其包含编码权利要求13-25中任一项的sgRNA的多核苷酸。

40. 一种试剂盒,其包含用于重组表达的载体,其中所述载体包含编码权利要求13-25中任一项的sgRNA的多核苷酸。

41. 一种试剂盒,其包含细胞,该细胞包含用于表达权利要求13-25中任一项的sgRNA的遗传构建体。

42. 权利要求41的试剂盒,其包含Cas9蛋白。

43. 权利要求38-42中任一项的试剂盒,其进一步包含一种或多种Cas9蛋白。

44. 一种用于位点特异性DNA切割的体外方法,其包括使DNA与权利要求1-12中任一项的复合物接触。

45. 一种用于诱导细胞中的位点特异性DNA切割的体外方法,其包括

(a) 使细胞与权利要求1-12中任一项的复合物接触,其中所述细胞包含含有DNA靶位点的DNA,并且

(b) 使细胞与结合所述特定配体接触,导致切割所述DNA的sgRNA:配体:Cas9复合物的形成。

46. 一种用于诱导细胞中的位点特异性DNA切割的离体方法,其包括

(a) 使细胞接触或在细胞内表达权利要求1-12中任一项的复合物或权利要求13-25中任一项的sgRNA,其中所述细胞包含含有DNA靶位点的DNA,并且

(b) 使细胞与所述特定配体接触,导致切割所述DNA的sgRNA:配体:Cas9复合物的形成,其中所述细胞是从受试者获得的。

47. 权利要求46的离体方法,其进一步包括使所述细胞接触Cas9蛋白或在所述细胞内表达Cas9蛋白。

48. 权利要求46或47所述的离体方法,其中所述细胞是真核细胞。

49. 权利要求48所述的离体方法,其中所述受试者是人。

50. 权利要求46所述的离体方法,其中同时或以任何次序序贯实施步骤(a)和(b)。

51. 权利要求1-12中任一项所述的复合物在制备用于个体中的基因组修饰的药物中的用途。

52. 权利要求13-25中任一项的sgRNA在制备用于个体中的基因组修饰的药物中的用途。

53. 权利要求51或52的用途,其中所述基因组修饰包含位点特异性DNA切割。

54. 一种药物组合物,其包含权利要求1-12中任一项所述的复合物和药学可接受赋形

剂。

55. 一种药物组合物,其包含权利要求13-25中任一项的sgRNA和药学可接受赋形剂。

可变换CAS9核酸酶及其用途

[0001] 相关申请

[0002] 本申请要求2014年7月8日提交的美国申请U.S.S.N.14/326,329,2014年7月8日提交的美国申请U.S.S.N.14/326,340,和2014年7月8日提交的美国申请U.S.S.N.14/326,361的在35U.S.C. §365 (c) 下的优先权,而且还要求2013年9月6日提交的美国临时申请U.S.S.N.61/874,682的在35U.S.C. §119 (e) 下的优先权,每篇通过提及并入本文。

[0003] 发明背景

[0004] 位点特异性内切核酸酶在理论上容许靶向操作基因组内的单一位点,并且在基因靶向的背景中以及对于治疗应用是有用的。在多种生物体(包括哺乳动物)中,位点特异性内切核酸酶已经用于通过刺激非同源末端连接或同源重组来进行基因组工程。在提供有力的研究工具外,位点特异性核酸酶还具有作为基因治疗剂的潜力,并且两种位点特异性内切核酸酶最近已经进入临床试验:一种(即CCR5-2246)靶向人CCR-5等位基因作为抗HIV治疗办法的一部分(NCT00842634,NCT01044654,NCT01252641),和另一种(VF24684)靶向人VEGF-A启动子作为抗癌治疗办法的一部分(NCT01082926)。

[0005] 在没有脱靶活性或仅具有最小脱靶活性的情况下对意图核酸酶靶位点的特异性切割是位点特异性内切核酸酶的临床应用,和还是基础研究应用中的高效率基因组操作的前提。例如,工程化位点特异性结合域的不完全特异性已经与除了意图靶物外的基因组基因座的不想要变化和细胞毒性联系起来。然而,目前可用的大多数核酸酶展现出相当大的脱靶活性,并且如此可能不适合于临床应用。一种在临床和研究背景中使用的出现的核酸酶平台是RNA引导性核酸酶,诸如Cas9。虽然这些核酸酶能够结合引导指导特异性靶位点切割的RNA(gRNA),但是仍对某些Cas9:gRNA复合物观察到脱靶活性(Pattanayak et al., "High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity." Nat Biotechnol. 2013; doi:10.1038/nbt.2673)。因此,需要将核酸酶工程化改造为具有改善的特异性的技术。

[0006] 发明概述

[0007] 本公开内容的一些方面基于下述认识,即一些工程化位点特异性内切核酸酶的报告的毒性基于脱靶DNA切割。此外,一般不能在分子水平上控制存在的RNA引导性核酸酶的活性,例如,以将核酸酶从“关闭”转换成“开启”状态。控制核酸酶的活性可以降低招致脱靶效应的可能性。本公开内容的一些方面提供了控制RNA-可编程内切核酸酶(诸如Cas9内切核酸酶)的结合和/或切割活性的策略、组合物、系统和方法。

[0008] 因而,本公开内容的一个实施方案提供了RNA引导性核酸酶复合物,其包含“可变换”引导RNA(gRNA)。例如,在一些实施方案中,本发明提供了复合物,其包含:(i)包含适体的gRNA,其中gRNA在缺乏与适体结合的特定配体的情况下不与靶核酸杂交;和(ii)Cas9蛋白。在一些实施方案中,适体由配体结合。在一些方面中,配体是任何分子。在一些方面中,配体是小分子、代谢物、碳水化合物、肽、蛋白质或核酸。在一些实施方案中,gRNA:配体:Cas9复合物结合靶核酸并且介导靶核酸的切割。参见例如图1。

[0009] 根据另一个实施方案,提供了包含适体的gRNA。在一些实施方案中,gRNA在缺乏与

适体结合的配体的情况下不与靶核酸杂交。此类gRNA可以称为“可变换gRNA”。例如,在一些方面中,gRNA在缺乏与适体的配体的情况下不结合Cas9。参见例如图1A-B。在一些实施方案中,在适体被对适体特异性的配体结合时,gRNA结合Cas9。在一些实施方案中,gRNA在缺乏或存在与适体结合的特定配体的情况下结合Cas9,但是仅在存在与适体结合的配体的情况下结合靶核酸。在一些方面中,配体是任何分子。在一些方面中,配体是小分子、代谢物、碳水化合物、肽、蛋白质或核酸。在一些实施方案中,适体是RNA适体,例如,自核开关衍生的RNA适体。在一些实施方案中,衍生适体的核开关选自茶碱核开关、焦磷酸硫胺素 (TPP) 核开关、腺苷钴胺素 (AdoCbl) 核开关、S-腺苷甲硫氨酸 (SAM) 核开关、SAH核开关、黄素单核苷酸 (FMN) 核开关、四氢叶酸核开关、赖氨酸核开关、甘氨酸核开关、嘌呤核开关、G1mS核开关或前Q核苷₁ (PreQ1) 核开关。在一些实施方案中,适体自茶碱核开关衍生,并且包含SEQ ID NO:3。在其它实施方案中,适体是非天然存在的,并且在一些方面中,使用指数富集配体系统进化(systematic evolution of ligands by exponential enrichment, SELEX) 平台被工程化改造为结合特定配体。在一些实施方案中,gRNA的非适体部分包含至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140或至少150个核苷酸,并且适体包含至少20、至少30、至少40、至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140、至少150、至少175、至少200、至少250或至少300个核苷酸。

[0010] 根据另一个实施方案,提供了使用发明性Cas9变体进行位点特异性DNA切割的方法。例如,在一些方面中,方法包括在Cas9蛋白切割DNA的条件下使DNA与复合物接触,所述复合物包含 (i) 包含适体的gRNA,其中gRNA包含结合DNA的部分的序列, (ii) 与gRNA的适体结合的特定配体,和 (iii) Cas9蛋白。

[0011] 根据另一个实施方案,提供了用于诱导细胞中的位点特异性DNA切割的方法。例如,在一些实施方案中,方法包括: (a) 使细胞接触或在细胞内表达包含适体的gRNA,其中所述gRNA包含能够结合DNA靶序列的序列; (b) 使细胞接触或在细胞内表达Cas9蛋白;并且 (c) 使细胞与结合gRNA的适体的配体接触,导致切割所述DNA靶物的gRNA:配体:Cas9复合物的形成。在一些实施方案中,细胞在细胞内生成配体,例如作为生理或病理生理过程的一部分。在一些实施方案中,方法包括 (a) 使细胞与包含Cas9蛋白和包含适体的gRNA的复合物接触,其中所述gRNA包含能够结合DNA靶序列的序列,并且 (b) 使细胞与结合gRNA的适体的配体接触,导致切割所述DNA靶物的gRNA:配体:Cas9复合物的形成。在一些方面中,同时或以任何次序序贯实施步骤 (a) 和 (b)。在一些实施方案中,在体外进行方法,而在其它实施方案中,在体内进行方法。

[0012] 根据另一个实施方案,提供了包含mRNA感测性gRNA的RNA引导性核酸酶复合物。例如,在一些实施方案中,复合物包含Cas9蛋白和gRNA,其中gRNA包含: (i) 与靶核酸的区域杂交的区域; (ii) 与区域 (i) 的序列部分或完全杂交的另一个区域; 和 (iii) 与转录物 (mRNA) 的区域杂交的区域。

[0013] 根据另一个实施方案,提供了mRNA感测性gRNA,例如,其包含: (i) 与靶核酸的区域杂交的区域; (ii) 与区域 (i) 的序列部分或完全杂交的另一个区域; 和 (iii) 与转录物 (mRNA) 的区域杂交的区域。参见例如图2。在一些实施方案中,区域 (i)、(ii) 和 (iii) 的每个序列包含至少5、至少10、至少15、至少20或至少25个核苷酸。在一些方面中,gRNA形成茎-环

结构,其中所述茎包含与区域(ii)的部分或整个序列杂交的区域(i)的序列,并且所述环由区域(iii)的部分或整个序列形成。在一些实施方案中,区域(ii)和(iii)都在区域(i)的5'或3'。参见例如图2A对图2C。在一些实施方案中,茎-环结构在缺乏与区域(iii)的序列杂交的转录物的情况下形成。在此例子中,gRNA被说成为“关闭”状态。参见例如图2A,2C。在一些实施方案中,转录物对区域(iii)的序列的结合导致茎-环结构的解折叠,或者阻止茎-环结构的形成,使得区域(ii)的序列不与区域(i)的序列杂交。在此例子中,gRNA被说成为“开启”状态。参见例如图2B,2D。在一些实施方案中,gRNA结合Cas9蛋白,并且区域(i)的序列在区域(iii)的序列结合(例如“感测”)转录物时与靶核酸杂交。

[0014] 根据另一个实施方案,提供了用于位点特异性DNA切割的方法,例如其包括使DNA与包含与mRNA感测性gRNA结合的Cas9蛋白的复合物接触,其中mRNA感测性gRNA被mRNA结合,从而容许复合物结合并切割DNA。

[0015] 根据另一个实施方案,提供了延长的DNA识别(xDNA感测性)gRNA。参见例如图3。在一些实施方案中,xDNA感测性gRNA包含:(i)与靶核酸的区域杂交的区域;(ii)与区域(i)的序列部分或完全杂交的另一个区域;和(iii)与靶核酸的另一个区域杂交的区域。在一些实施方案中,区域(i)和(ii)的每种序列包含至少5、至少10、至少15、至少20或至少25个核苷酸;并且区域(iii)的序列包含至少5、至少10、至少15、至少20、至少25、至少30、至少40、至少50、至少75或至少100个核苷酸。在一些实施方案中,gRNA形成茎-环结构,其中所述茎包含与区域(ii)的部分或整个序列杂交的区域(i)的序列,并且环由区域(iii)的部分或整个序列形成。在一些实施方案中,区域(ii)和(iii)都在区域(i)的5'或3'。参见例如图3A对图3C。在一些实施方案中,茎-环结构在缺乏与(iii)中的序列互补并结合的靶核酸的区域的情况下形成。参见例如图3A,C。在一些实施方案中,靶核酸的区域对区域(iii)的序列的结合导致茎-环结构的解折叠,或者阻止茎-环结构的形成,使得区域(ii)的序列不与区域(i)的序列杂交。参见例如图3B,D。在一些实施方案中,gRNA结合Cas9蛋白,并且(i)中的序列在(iii)中的序列结合靶核酸时结合靶核酸。

[0016] 根据另一个实施方案,提供了复合物,其包含xDNA感测性gRNA和Cas9蛋白,任选地其中复合物包含靶核酸。在一些实施方案中,复合物的形成导致靶核酸的切割。

[0017] 根据另一个实施方案,提供了用于位点特异性DNA切割的方法,其包括使DNA与包含xDNA感测性gRNA和Cas9的复合物接触。

[0018] 可以对细胞(例如体外或体内细胞)中的DNA进行本文中提供的任何方法。在一些实施方案中,对真核细胞中的DNA进行本文中提供的任何方法。在一些实施方案中,真核细胞在个体,例如人中。

[0019] 根据另一个实施方案,提供了多核苷酸,例如,其编码本文中描述的gRNA、复合物或蛋白质(例如Cas9蛋白)。在一些实施方案中,提供了包含本文中描述的多核苷酸的载体。在一些实施方案中,提供了用于重组表达本文中描述的任何gRNA、复合物或蛋白质(例如Cas9蛋白)的载体。在一些实施方案中,细胞包含用于重组表达本文中描述的任何gRNA、复合物或蛋白质(例如Cas9蛋白)的遗传构建体。

[0020] 在一些实施方案中,提供了试剂盒。例如,提供了包含本文中描述的任何gRNA、复合物或蛋白质(例如Cas9蛋白)的试剂盒。在一些实施方案中,提供了包含本文中描述的多核苷酸的试剂盒。在一些实施方案中,提供了包含用于重组表达的载体的试剂盒,其中载体

包含编码本文中描述的任何gRNA、复合物或蛋白质(例如Cas9蛋白)的多核苷酸。在一些实施方案中,提供了包含细胞的试剂盒,所述细胞包含用于表达本文中描述的任何gRNA、复合物或蛋白质(例如Cas9蛋白)的遗传构建体。

[0021] 本发明的其它优点、特征和用途从本发明的某些实施方案的详细描述;附图(其是示意性的而并不意图按比例绘制);和权利要求书看会是明显的。

[0022] 附图简述

[0023] 图1A-D显示了涉及与适体连接的gRNA的本发明的某些实施方案。(A)在此图中,示意性描述包含适体的可变换gRNA。在缺乏结合适体的特定配体(在本文为代谢物)的情况下,负责结合靶核酸的序列与适体(参见远左侧,描绘为“转换序列”的区域)方面杂交。在结合代谢物后,适体经历构象变化,使得负责结合靶核酸的序列不再与适体序列杂交,容许它与靶物杂交。(B)在转换成“开启”状态后,gRNA在与Cas9结合时将核酸酶引导至靶位点,在那里它与靶位点杂交,容许Cas9切割靶核酸的每条链。(C-D)在此图中,与gRNA连接的适体自茶碱核开关衍生。在缺乏茶碱的情况下(C),适体(描绘为“茶碱核开关”的方面结合负责结合靶核酸(通过小图顶部的双链序列描绘)的序列(描绘为“切割靶物的引导”)方面,从而阻碍gRNA与靶核酸杂交。在图1C中,从上至下,序列对应于SEQ ID NO:4-6。当适体被茶碱(描述为结合适体序列的固体小分子)结合时(D),它经历构象变化,其导致没有与靶核酸杂交的“引导”序列。在图1D中,从上至下,序列对应于SEQ ID NO:4、5和11。

[0024] 图2A-D显示了涉及mRNA感测性gRNA的本发明的某些实施方案。(A-B)在此图中,描绘了包含5'转录物传感器(transcript sensor)/引导块基序(guide block motif)的gRNA。在缺乏某种mRNA的情况下(A),转录物传感器的方面保持未结合,导致形成茎-环结构,其阻断负责结合靶核酸(通过小图顶部的双链序列描绘)的序列(描绘为“切割靶物的引导”)的某些方面,从而阻止gRNA与靶核酸杂交。在图2A中,从上至下,序列对应于SEQ ID NO:4、5和7。在存在与转录物传感器杂交的mRNA的情况下(B),gRNA经历构象变化,其导致没有与靶核酸杂交的“引导”序列,序列(从上至下)对应于SEQ ID NO:4、5、7和12。(C-D)类似地,可以对包含3'转录物传感器/引导块的gRNA应用策略,使得在缺乏mRNA的情况下(C),gRNA为“关闭”状态(从上至下,序列对应于SEQ ID NO:4、5和8),并且在存在mRNA的情况下(D),gRNA为“开启”状态(从上至下,序列对应于SEQ ID NO:4、5和12)。

[0025] 图3A-D显示了涉及延长的DNA(xDNA)识别策略的本发明的某些实施方案。(A-B)在此实施方案中,描绘了包含5'xDNA传感器/引导块基序的gRNA。xDNA传感器基序与靶核酸的其它方面(例如在“切割靶物的引导”序列外)互补并杂交。(A)在缺乏正确的靶序列(例如包含“引导”序列的靶物及xDNA传感器序列的靶物两者)的情况下,xDNA传感器的方面保持未结合,导致形成茎-环结构,其阻断负责结合靶核酸(通过小图顶部的双链序列描绘)的序列(描绘为“切割靶物的引导”)的某些方面,从而阻止gRNA与靶核酸杂交,序列(从上至下)对应于SEQ ID NO:4、5和9。在存在与xDNA传感器的部分杂交的正确靶核酸的情况下(B),gRNA经历构象变化,其导致没有与靶核酸杂交的“引导”序列,序列(从上至下)对应于13、14和9。如此,仅在存在正确的靶核酸的情况下,确实发生gRNA(和关联Cas9蛋白)的结合。这有效增加(即延长)被例如Cas9:gRNA复合物识别的靶核苷酸的数目,这增加特异性。(C-D)类似地,可以对包含3'xDNA传感器/引导块的gRNA应用策略,从而在缺乏靶核酸的情况下(C),gRNA为“关闭”状态(从上至下,序列对应于SEQ ID NO:4、5和10),并且在存在靶核酸的情况下

(D),gRNA为“开启”状态,从上至下,序列对应于SEQ ID NO:4、15、5、16和10。

[0026] 定义

[0027] 如本文中和权利要求书中使用的,单数形式“一个”、“一种”和“该/所述”包括单数和复数提及物,除非上下文另有清楚指示。如此,例如,提及“药剂”包括单一药剂和多种此类药剂。

[0028] 术语“适体”指结合特定靶分子(例如特定配体)的核酸或肽分子。在一些实施方案中,配体对适体的结合诱导适体和例如与适体缀合或连接的其它分子的构象变化。在一些实施方案中,经由重复轮次的体外选择或者等同地SELEX(指数富集配体系统进化)工程化改造核酸(例如DNA或RNA)适体以结合各种分子靶物,例如,小分子、大分子、代谢物、蛋白质、蛋白质、碳水化合物、金属、核酸、细胞、组织和生物体。用于将适体工程化改造为结合小分子的方法是本领域中已知的,并且包括那些记载于美国专利No.5,580,737和8,492,082; Ellington and Szostak, "In vitro selection of RNA molecules that bind specific ligands." Nature. 1990;346:818-822; Tuerk and Gold, "Systematic evolution of ligands by exponential enrichment:RNA ligands to bacteriophage T4 DNA polymerase." Science. 1990;249:505-510; Burke and Gold, "RNA aptamers to the adenosine moiety of S-adenosyl methionine:structural inferences from variations on a theme and the reproducibility of SELEX." Nucleic Acids Res. 1997;25 (10) :2020-4; Ulrich et al., "DNA and RNA aptamers:from tools for basic research towards therapeutic applications." Comb Chem High Throughput Screen. 2006;9 (8) :619-32; Svobodová et al., "Comparison of different methods for generation of single-stranded DNA for SELEX processes." Anal Bioanal Chem. 2012; 404:835-842;在此通过提及并入每篇的全部内容。在自然界中也找到核酸适体,例如,那些形成核开关的部分的。“核开关”是mRNA分子的调节区段,其结合小分子,例如代谢物,从而导致由mRNA编码的蛋白质(例如参与结合核开关的代谢物生成的蛋白质)的生成的变化。核开关在概念上经常分成两个部分:适体和表达平台(例如mRNA)。适体直接结合小分子(例如代谢物),并且mRNA相应适体变化而经历结构变化。通常,mRNA的结构变化导致蛋白质表达的降低或抑制。可以使用本领域中的常规方法从核开关中克隆(例如分离)适体,并且使用适体控制与其连接的其它分子(例如RNA,DNA)的活性。另外,可以使用已知的方法再工程化改造自然界中找到的适体以结合合成的、非天然的小分子配体以控制与其连接的其它分子的活性。参见例如Dixon et al., "Reengineering orthogonally selective riboswitches." PNAS2010;107 (7) :2830-2835,其全部内容在此通过提及并入。下文是包括非限制性的一批核开关,包括适体:

[0029] 钴胺素核开关(又是B₁₂-元件),其结合腺苷酰钴胺素(维生素B₁₂的辅酶形式)以调节钴胺素生物合成和转运钴胺素和类似代谢物,和其它基因。参见例如Nahvi et al., "Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes." Nucleic Acids Res. 2004;32:143-150; Vitreschak et al., "Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element." RNA. 2003;9:1084-1097;在此通过提及并入每篇的全部内容。

[0030] 环状二-GMP核开关结合信号传导分子环状二-GMP以调节由此第二信使控制的多

种基因。至少两类环状二-GMP核开关是已知的：环状二-GMP-I核开关和环状二-GMP-II核开关。参见例如Sudarsan et al., “Riboswitches in eubacteria sense the second messenger cyclic di-GMP.” *Science*. 2008; 321 (5887) :411-3; Lee et al., “An allosteric self-splicing ribozyme triggered by a bacterial second messenger.” *Science*. 2010; 329 (5993) :845-8;在此通过提及并入每篇的全部内容。

[0031] FMN核开关(也是RFN-元件)结合黄素单核苷酸(FMN)以调节核黄素生物合成和转运。参见例如Winkler et al., “An mRNA structure that controls gene expression by binding FMN.” *Proc Natl Acad Sci USA*. 2002; 99 (25) :15908-15913; Serganov et al., “Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch.” *Nature*. 2009; 458 (7235) :233-7;在此通过提及并入每篇的全部内容。

[0032] G1mS核开关是一种在由葡萄糖胺-6-磷酸结合时切割自身的核酶。参见例如Winkler et al., “Control of gene expression by a natural metabolite-responsive ribozyme.” *Nature*. 2004; 428:281-286; Jansen et al., “Backbone and nucleobase contacts to glucosamine-6-phosphate in the glmS ribozyme.” *Nat Struct Mol Biol*. 2006; 13:517-523; Hampel and Tinsley, “Evidence for preorganization of the glmS ribozyme ligand binding pocket.” *Biochemistry*. 2006; 45:7861-7871;在此通过提及并入每篇的全部内容。

[0033] 甘氨酸核开关结合甘氨酸以调节甘氨酸代谢基因,包括使用甘氨酸作为能源。参见例如Mandalet al., “A glycine-dependent riboswitch that uses cooperative binding to control gene expression.” *Science*. 2004; 306 (5694) :275-279; Kwon and Strobel, “Chemical basis of glycine riboswitch cooperativity.” *RNA*. 2008; 14 (1) :25-34;在此通过提及并入每篇的全部内容。

[0034] 赖氨酸核开关(又是L-box)结合赖氨酸以调节赖氨酸生物合成、分解代谢和转运。参见例如Sudarsan et al., “An mRNA structure in bacteria that controls gene expression by binding lysine.” *Genes Dev*. 2003; 17:2688-2697; Grundy et al., “The L box regulon:Lysine sensing by leader RNAs of bacterial lysine biosynthesis genes.” *Proc.Natl.Acad.Sci.USA*. 2003; 100:12057-12062;在此通过提及并入每篇的全部内容。

[0035] PreQ1核开关结合前-Q核苷(pre-queuosine)₁,以调节参与此前体至Q核苷的合成或转运的基因。至少两种独特类别的PreQ1核开关是已知的:PreQ1-I核开关和PreQ1-II核开关。参见例如Roth et al., “A riboswitch selective for the queuosine precursor preQ1contains an unusually small aptamer domain,” *Nat Struct Mol Biol*. 2007; 14 (4) :308-317; Klein et al., “Cocrystal structure of a class I preQ1riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase,” *Nat.Struct.Mol.Biol*. 2009; 16 (3) :343-344; Kang et al., “Structural Insights into riboswitch control of the biosynthesis of queuosine,a modified nucleotide found in the anticodon of tRNA.” *Mol.Cell* 33 2009; (6) :784-90; Meyer et al., “Confirmation of a second natural preQ1aptamer class in Streptococcaceae bacteria.” *RNA* 2008; 14 (4) :685;在此通过提及并入每篇的全部内容。

[0036] 嘌呤核开关结合嘌呤以调节嘌呤代谢和转运。嘌呤核开关的不同形式结合鸟嘌呤(最初称为G-框的形式)或腺嘌呤。对鸟嘌呤或腺嘌呤的特异性完全取决于与特定位置(例如Y74)处的核开关中的单一嘧啶的Watson-Crick相互作用。在鸟嘌呤核开关中,此残基通常是胞嘧啶(例如C74),在腺嘌呤核开关中,它通常是尿嘧啶(例如U74)。嘌呤核开关的同源类型结合脱氧鸟苷,但是比单一核苷酸突变具有更重大的差异。参见例如Serganov et al., "Structural basis for discriminative regulation of gene expression by adenine-and guanine-sensing mRNAs." *Chem Biol.* 2004;11(12):1729-41; Batey et al., "Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine." *Nature.* 2004;432(7015):411-415; Mandal and Breaker, "Adenine riboswitches and gene activation by disruption of a transcription terminator." *Nat Struct Mol Biol.* 2004;11(1):29-35;在此通过提及并入每篇的全部内容。

[0037] SAH核开关结合S-腺苷高半胱氨酸以调节参与再循环此代谢物的基因,所述代谢物在甲基化反应中使用S-腺苷甲硫氨酸时生成。参见例如Wang et al., "Riboswitches that Sense S-adenosylhomocysteine and Activate Genes Involved in Coenzyme Recycling." *Mol. Cell.* 2008;29(6):691-702; Edwards et al., "Structural basis for recognition of S-adenosylhomocysteine by riboswitches." *RNA* 2010;16(11):2144-2155;在此通过提及并入每篇的全部内容。

[0038] SAM核开关结合S-腺苷甲硫氨酸(SAM)以调节甲硫氨酸和SAM生物合成和转运。至少4种SAM核开关是已知的: SAM-I(最初称作S-框)、SAM-II、S_{MK}框核开关和Sam-IV。SAM-I在细菌中是普遍的,但是SAM-II仅存在于alpha-、beta-和几种gamma-蛋白菌中。认为S_{MK}框核开关仅存在于乳杆菌目(Lactobacillales)中。SAM-IV核开关与SAM-I核开关具有相似的配体结合核心,但是在独特支架的背景中。参见例如Montange et al., "Structure of the S-adenosyl methionine riboswitch regulatory mRNA element." *Nature.* 2006;441:1172-1175; Winkler et al., "An mRNA structure that controls gene expression by binding Sadenosylmethionine." *Nat Struct Biol.* 2003;10:701-707; Zasha et al., "The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches." *RNA.* 2008;14(5):822-828;在此通过提及并入每篇的全部内容。

[0039] 四氢叶酸核开关结合四氢叶酸以调节合成和转运基因。参见例如Ames et al., "A eubacterial riboswitch class that senses the coenzyme tetrahydrofolate." *Chem. Biol.* 2010;17(7):681-5; Huang et al., "Long-range pseudoknot interactions dictate the regulatory response in the tetrahydrofolate riboswitch." *Proc. Natl. Acad. Sci. U.S.A.* 2011;108(36):14801-6; Trausch et al., "The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer." *Structure.* 2011;19(10):1413-23;在此通过提及并入每篇的全部内容。

[0040] 茶碱核开关通过SELEX鉴定,并且选择性结合小分子茶碱。该适体包含茶碱结合需要的15个核苷酸的核心基序。参见例如Jenison et al., "High-resolution molecular discrimination by RNA." *Science.* 1994;263:1425-1429; Zimmerman et al., "Molecular

interactions and metal binding in the theophylline-binding core of an RNA aptamer.”RNA.2000;6 (5) :659-67;Suess et al.,“A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo.”Nucleic Acids Res.2004;32 (4) :1610-1614;在此通过提及并入每篇的全部内容。还可参见例如图1C-D。

[0041] TPP核开关(也是THI-框)结合硫胺素焦磷酸(TPP)以调节硫胺素生物合成和转运,以及类似代谢物的转运。认为它是至今为止在真核生物中找到的唯一核开关。参见例如Edwards et al.,“Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition.”Structure2006;14 (9) :1459-68;Winkler et al.,“Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression.”Nature.2002;419 (6910) :952-956;Serganov et al.,“Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch.”Nature.2006;441 (7097) :1167-1171;在此通过提及并入每篇的全部内容。

[0042] 术语“Cas9”或“Cas9核酸酶”指包含Cas9蛋白或其片段的RNA引导性核酸酶。Cas9核酸酶有时又称为cas9核酸酶或CRISPR(成簇的规律间隔性短回文重复序列(clustered regularly interspaced short palindromic repeat))关联核酸酶。CRISPR是一种适应性免疫系统,其提供针对活动的遗传元件(例如病毒、可转座元件和接合质粒)的保护。CRISPR簇含有间隔物、与前述活动元件互补的序列和靶侵入性核酸。CRISPR簇被转录并加工成CRISPR RNA(crRNA)。在II型CRISPR系统中,pre-crRNA的正确加工需要反式编码小RNA(trans-encoded small RNA,tracrRNA)、内源核糖核酸酶3(rnc)和Cas9蛋白。tracrRNA充当核糖核酸酶3辅助性pre-crRNA加工的引导。随后,Cas9/crRNA/tracrRNA以内切核水解方式(endonucleolytically)切割与间隔物互补的线性或环状dsDNA靶物。首先,以内切核水解方式切割不与crRNA互补的靶链,然后以外切核水解方式3’-5’修剪。实际上,DNA结合和切割通常需要蛋白质和两种RNA种类。然而,可以工程化改造单一引导RNA(“sgRNA”,或仅“gRNA”),从而掺入crRNA和tracrRNA两者的方面掺入单一RNA分子中。参见例如Jinek M.,Chylinski K.,Fonfara I.,Hauer M.,Doudna J.A.,Charpentier E.Science 337:816-821 (2012),在此通过提及并入每篇的全部内容。Cas9识别CRISPR重复序列中的短基序(PAM或原间隔物(protospacer)相邻基序)以帮助区分自身或非自身。Cas9核酸酶序列和结构是本领域技术人员公知的(参见例如“Complete genome sequence of an M1 strain of Streptococcus pyogenes.”Ferretti J.J.,McShan W.M.,Ajdic D.J.,Savic D.J.,Savic G.,Lyon K.,Primeaux C.,Sezate S.,Suvorov A.N.,Kenton S.,Lai H.S.,Lin S.P.,Qian Y.,Jia H.G.,Najar F.Z.,Ren Q.,Zhu H.,Song L.expand/collapse author list McLaughlin R.E.,Proc.Natl.Acad.Sci.U.S.A.98:4658-4663 (2001);“CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.”Deltcheva E.,Chylinski K.,Sharma C.M.,Gonzales K.,Chao Y.,Pirzada Z.A.,Eckert M.R.,Vogel J.,Charpentier E.,Nature 471:602-607 (2011);及“A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.”Jinek M.,Chylinski K.,Fonfara I.,Hauer M.,Doudna J.A.,Charpentier E.Science 337:816-821 (2012),通

过提及将每篇的全部内容收入本文)。已经在多种物种中描述了Cas9直向同系物,包括但不限于酿脓链球菌 (*S. pyogenes*) 和嗜热链球菌 (*S. thermophilus*)。其它合适的Cas9核酸酶和序列基于本公开内容对于本领域技术人员会是明显的,并且此类Cas9核酸酶和序列包含来自生物体的Cas9序列和披露于Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) RNA Biology 10:5, 726-737 (其全部内容通过提及并入本文) 的基因座。在一些实施方案中,包含Cas9的蛋白质或其片段的蛋白质称为“Cas9变体”。Cas9变体与Cas9或其片段共享同源性。例如,Cas9变体与野生型Cas9是至少约70%相同的,至少约80%相同的,至少约90%相同的,至少约95%相同的,至少约98%相同的,至少约99%相同的,至少约99.5%相同的,或至少约99.9%。在一些实施方案中,Cas9变体包含Cas9的片段(例如gRNA结合域或DNA切割域),使得片段与野生型Cas9的相应片段是至少约70%相同的,至少约80%相同的,至少约90%相同的,至少约95%相同的,至少约98%相同的,至少约99%相同的,至少约99.5%相同的,或至少约99.9%。在一些实施方案中,野生型Cas9对应于来自酿脓链球菌的Cas9 (NCBI参照序列:NC_017053.1, SEQ ID NO:1 (核苷酸);SEQ ID NO:2 (氨基酸))。

[0043] ATGGATAAGAAACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCGGTGATCACTGA
TGATTATAAGGTTCCGTCTAAAAGTTCAAGGTTCTGGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATA
GGGGCTCTTTATTGCGAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACAC
GTCGGAAGAACGTATTGTTATCTACAGGAGATTTTCAAATGAGATGGCGAAAGTAGATGATAGTTCTTCA
TCGACTTGAAGAGTCTTTGGTGAAGAACAGACATGAACGTACCTATTGGAAATATAGTAGAT
GAAGTTGCTTATCATGAGAAATATCCAACATCTATCATCTGCAGAAAAATTGGCAGATTCTACTGATAAACCGG
ATTGCGCTTAATCTATTGGCCTTAGCGCATATGATTAAGTTCGTGGCATTGGATTGAGGGAGATTAAA
TCCTGATAATAGTGTGGACAAACTATTATCCAGTTGGTACAAATCTACAATCAATTATTGAAGAAAACCT
ATTAACGCAAGTAGAGTAGATGCTAAAGCGATTCTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCA
TTGCTCAGCTCCCCGGTGAGAAGAGAAATGGCTTGGGAATCTCATTGCTTGTCAATTGGGATTGACCCCTAA
TTTAAATCAAATTGATTGGCAGAAGATGCTAAATTACAGCTTCAAAGATACTACGATGATGATTAGAT
AATTATTGGCGCAAATTGGAGATCAATATGCTGATTGTTGGCAGCTAAGAATTATCAGATGCTATTTC
TTTCAGATATCCTAAGAGTAAATAGTGAATAACTAAGGCTCCCTATCAGCTTCAATGATTAAGCGCTACGATGA
ACATCATCAAGACTTGAACCTTTAAAAGCTTACTTCGACAACAACCTCCAGAAAAGTATAAGAAATCTTTT
GATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGAGCTAGCCAAGAAGAATTTATAAATTATCAAAC
CAATTAGAAAAATGGATGGTACTGAGGAATTATTGGTGAAGAAACTAACTCGTAAGAGATTGCTGCGCAAGCAACG
GACCTTGACAACGGCTCTATTCCCCATCAAATTCACTGGGTGAGCTGCATGCTATTGAGAAGACAAGAAC
TTTATCCATTAAAGACAATCGTGAGAAGATTGAAAAATCTTGACTTTCAATTCCATTATGTTGGC
CATGGCGCTGGCAATAGTCGTTGCATGGATGACTCGGAAGTCTGAAGAACAAATTACCCATGGAATTG
AGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTGAACGCATGACAAACTTGATAAAAATCTTCAAAT
GAAAAAGTACTACCAAAACATAGTTGCTTATGAGTATTGAGGTTATAACGAATTGACAAAGGTCAAATATG
TTACTGAGGAATGCGAAAACCAGCATTCTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTACTCTTCAAAC
AAATCGAAAAGTAACCGTTAAGCAATTAAAGAAGATTATTCAAAAAAATAGAATGTTGATAGTGTGAAATT
TCAGGAGTTGAAGATAGATTAAATGCTCATTAGGCGCCTACCATGATTGCTAAAATTAAAGATAAAGATT
TTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTAACATTGACCTTATTGAAGATAGGGGAT

GATTGAGGAAAGACTAAACATATGCTCACCTCTTGATGATAAGGTGATGAAACAGCTAACGTCGCCGTAT
 ACTGGTTGGGACGTTGTCTGAAAATTGATTAATGGTATTAGGGATAAGCAATCTGGCAAACAAATATTAGATT
 TTTGAAATCAGATGGTTGCCAATCGCAATTATGCAGCTGATCCATGATGATAGTTGACATTAAAGAAGA
 TATTCAAAAAGCACAGGTGTCTGACAAGGCCATAGTTACATGAACAGATGCTAACCTAGCTGGCAGTCCTGCT
 ATTAAGGTTAGTACAGACTGTAAAATTGTTGATGAAGTGGTCAAAGTAATGGGCATAAGCCAGAAAATA
 TCGTTATTGAAATGGCACGTGAAAATCAGACAACCTAAAAGGGCCAGAAAATTGCGAGAGCGTATGAAACGAAT
 CGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTCTAAAGAGCATCCTGAAATACTCAATTGCAAATGAA
 AAGCTCTATCTTACACAGTAACTACAGTAACTGAGAGACATGTATGTGGACCAAGAATTAGATATTACGTTAAGTG
 ATTATGATGTCGATCACATTGTCACAAAGTTCATTAAGACGATTCAATAGACAATAAGGTACTAACCGTTC
 TGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAGATGAAAAGTATTGGAGACAA
 CTTCTAAACGCCAAGTTAACACTAACGTAAGTTGATAATTAAACGAAAGCTGAACGTGGAGGTTGAGTGAAC
 TTGATAAAAGCTGGTTTATCAAACGCCAATTGTTGAAACTCGCCAAATCACTAACGATGTGGCACAAATTGGA
 TAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTATTGAGAGGTTAAAGTGAATTACCTTAAACGTTAA
 TTAGTTCTGACTTCCGAAAAGATTCCAATTCTATAAGTACGTGAGGTTAACATTACCATCATGCCATGATG
 CGTATCTAAATGCCGTCGGAACTGCTTGATTAAGAAATATCCAAAACCTGAATCGGAGTTGTCTATGGTGA
 TTATAAAGTTATGATGTTGTTAAATGATTGCTAACGTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAATATTG
 TTTACTCTAATATCATGAACTTCTCAAACAGAAATTACACTGCAAATGGAGAGATTGCAAACGCCCTCTAA
 TCGAAACTAATGGGGAAACTGGAGAAATTGCTGGATAAAGGGCGAGATTGCAAACGTCAGTGAAGTATTGTC
 CATGCCCAAGTCAATTGCAAGAAAACAGAAGTACAGACAGCGGATTCTCCAAGGAGTCATTACCAAAA
 AGAAAATCGGACAAGCTTATTGCTCGAAAAAGACTGGGATCCAAAAAATGGTTGATAGTCAAACCG
 TAGCTTATTCACTGCTAGTGGTCTAAGGTGGAAAAGGGAAATCGAAGAAGTAAACCGTTAAAGAGTTACT
 AGGGATCACAATTGGAAGAAGTTCTTGAAAAAATCGATTGACTTTAGAAGCTAAAGGATATAAGGAA
 GTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTGAGTTAGAAAACGGTGTAAACGGATGCTGG
 CTAGTGCCGAGAATTACAAAAGGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTAGTTATTTAGCTAG
 TCATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAACAATTGTTGTGGAGCAGCATAAGCATTAT
 TTAGATGAGATTATTGAGCAAATCAGTGAATTCTAAGCGTTATTAGCAGATGCCATTAGATAAAAGTTC
 TTAGTGCATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTACATTACGTTGAC
 GAATCTGGAGCTCCGCTGTTAAATATTGATACAACAATTGATCGAAACGATACGTCTACAAAAGAA
 GTTTAGATGCCACTCTTATCCATCAATCCATCACTGGCTTTATGAAACACGCATTGAGTCAGCTAGGAG
 GTGACTGA (SEQ ID NO:1)

[0044] MDKKYSIGLDIGTNSVGWAVITDDYKVPSSKFVLGNTDRHSIKKNLIGALLFGSGETAETRLKRTA
 RRRYTRRKNRICYLQEIFSNEAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLAD
 STDKADLRLIYLALAHMIKFRGHFIEGDLNPDNSDVDKLFQIQLVQIYNQLFEENPINASRVDAKAILSRSKSR
 RLENLIAQLPGEKRNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDLDNLLAQIGDQYADLFLAAKNL
 SDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKYKEIFFDQSKNGYACYIDGGASQEEF
 YKFIKPILEKMDGTEELLVQLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFKNLPNEKVLPHSLLYEYFTVYNEL
 TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAYHDLLKI
 IKDKDFLDNEENEDILEDIVLTLLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRTGWGRSLRKLINGIRDKQSG

KTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGHSLHEQIANLAGSPAIIKKGILQTVKIVDELVKVMG
HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELD
INRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER
GGLSELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY
HHAHDAYLNAVVGTLIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI
RKRPLIETNGETGEIVWDKGKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENG
RKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADA
NLDKVLSAYNKHRDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRID
LSQLGGD (SEQ ID NO:2)

[0045] 术语“缀合”指两个实体的联合,例如,两个分子,诸如两个蛋白质、两个域(例如结合域和切割域)、或蛋白质和药剂(例如蛋白质结合域和小分子)的联合。在一些方面中,连接在蛋白质(例如RNA可编程核酸酶)和核酸(例如引导RNA)之间。联合可以是例如经由直接或间接(例如经由接头)共价连接。在一些实施方案中,连接是共价的。在一些实施方案中,经由连接两个分子的接头缀合两个分子。例如,在RNA的两个部分彼此缀合的一些实施方案中,例如适体(或核酸感测域)和gRNA,可以经由多核苷酸接头(例如连接一个RNA的3'端至另一个RNA的5'端的核苷酸序列)缀合两个RNA。在一些实施方案中,接头包含至少1、至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少15、至少20、至少25或至少30个核苷酸。

[0046] 如本文中在核酸序列的背景中使用的,术语“共有序列”指代表多个相似序列中在每个位置处找到的最频繁的核苷酸残基的计算的序列。通常,通过序列比对确定共有序列,其中彼此比较相似序列,并且计算类似的序列基序。

[0047] 如本文中使用的,术语“有效量”指足以引发期望的生物应答的生物活性剂的量。例如,在一些实施方案中,核酸酶的有效量可以指足以诱导切割由核酸酶特异性结合并切割的期望靶位点(优选以最小脱靶切割或无脱靶切割)的核酸酶量。如熟练技术人员会领会, the effective amount of 药剂(例如核酸酶、杂合蛋白、融合蛋白、蛋白质二聚体、蛋白质(或蛋白质二聚体)和多核苷酸的复合物、或多核苷酸可以随各种因素,如例如期望的生物应答、特定的等位基因、基因组、靶位点、细胞、或靶定的组织、和使用的药剂而变化。

[0048] 如本文中使用的,术语“工程化/工程化改造”指已经人为设计、生成、制备、合成和/.或制造的核酸分子、蛋白质复合物、复合物、物质、或实体。因而,工程化产物是自然界中不存在的产物。

[0049] 如本文中使用的,术语“接头”指连接两个相邻分子或模块(例如适体(或核酸感测域)和gRNA的化学基团或分子。通常,接头位于两个基团、分子或其它模块之间,或者侧翼有两个基团、分子或其它模块,并且经由共价键彼此连接,如此连接两者。在一些实施方案中,接头是核苷酸接头。在一些实施方案中,核苷酸接头包含至少1、至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少15、至少20、至少25或至少30个核苷酸。在一些实施方案中,接头是氨基酸或多个氨基酸(例如肽或蛋白质)。在一些实施方案中,接头是有机分子、基团、多聚体或化学模块。

[0050] 如本文中使用的,术语“突变”指序列(例如核酸或氨基酸序列)内的残基取代为另

一种残基,或者序列内的一个或多个残基的取代或插入。突变在本文中通常通过鉴定后面有序列内的残基位置的初始残基及通过新取代残基的身份描述。用于进行本文中提供的氨基酸取代(突变)的方法是本领域中已知的,并且由例如,Green and Sambrook,Molecular Cloning:A Laboratory Manual(4th ed.,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.(2012))提供。

[0051] 如本文中使用的,术语“核酸酶”指能够切割连接核酸分子中的核苷酸残基的磷酸二酯键的药剂,例如蛋白质或小分子。在一些实施方案中,核酸酶是蛋白质,例如可以结合核酸分子并且切割连接核酸分子内的核苷酸残基的磷酸二酯键的酶。核酸酶可以是内切核酸酶(其切割多核苷酸链内的磷酸二酯键)或者外切核酸酶(其切割多核苷酸链末端的磷酸二酯键)。在一些实施方案中,核酸酶是位点特异性核酸酶,其结合和/或切割特定核苷酸序列内的特定磷酸二酯键,所述特定核苷酸序列在本文中又称为“识别序列”、“核酸酶靶位点”或“靶位点”。在一些实施方案中,核酸酶是RNA引导性(即RNA-可编程的)核酸酶,其与具有与靶位点互补的序列的RNA(例如引导RNA,“gRNA”)复合(例如结合),从而提供核酸酶的序列同一性。在一些实施方案中,核酸酶识别单链靶位点。在其它实施方案中,核酸酶识别双链靶位点,例如,双链DNA靶位点。许多天然存在的核酸酶(例如许多天然存在的DNA限制性核酸酶)的靶位点是本领域技术人员公知的。在许多情况下,DNA核酸酶(诸如EcoRI、HindIII或BamHI)识别长度4至10个碱基对的回文双链DNA靶位点,并且在靶位点内的特定位置处切割两个DNA链之每个。一些内切核酸酶对称切割双链核酸靶位点,即在相同位置处切割两条链,从而末端包含碱基成对核苷酸,在本文中又称为平端。其它内切核酸酶不对称切割双链核酸靶位点,即在不同位置处切割每条链,从而末端包含不成对核苷酸。在双链DNA分子末端的不成对核苷酸又称为“突出端”,例如作为“5’-突出端”或作为“3’-突出端”,这取决于不成对核苷酸形成相应DNA链的5’还是5’端。以不成对核苷酸为末端的双链DNA分子末端又称为粘性末端,因为它们可以“粘附于”包含互补不成对核苷酸的其它双链DNA分子末端。核酸酶蛋白质通常包含“结合域”(其介导蛋白质与核酸底物的相互作用,并且在一些情况中还特异性结合靶位点)和“切割域”(其催化核酸主链内的磷酸二酯键的切割)。在一些实施方案中,核酸酶蛋白质可以以单体形式结合并且切割核酸分子,而在其它实施方案中,核酸酶蛋白质必须二聚化或多聚化以切割靶核酸分子。天然存在的核酸酶的结合域和切割域,及可以融合以创建核酸酶结合特异性靶位点的模块结合域和切割域是本领域技术人员公知的。例如,RNA可编程核酸酶(例如Cas9)的结合域,或具有无活性DNA切割域的Cas9蛋白可以用作结合域(例如其结合gRNA以指导对靶位点的结合),从而特异性结合期望的靶位点,并且与切割域(例如FokI的切割域)融合或缀合以创建切割靶位点的工程化核酸酶。

[0052] 如本文中使用的,术语“核酸”和“核酸分子”指包含核碱基和酸性模块的化合物,例如核苷、核苷酸、或核苷酸的多聚体。通常,多聚体核酸,例如包含3个或更多个核苷酸的核酸分子是线性分子,其中相邻核苷酸经由磷酸二酯连接彼此连接。在一些实施方案中,“核酸”指个别的核酸残基(例如核苷酸和/或核苷)。在一些实施方案中,“核酸”指包含3个或更多个个别的核苷酸残基的寡核苷酸链。如本文中使用的,术语“寡核苷酸”和“多核苷酸”可以互换使用,指核苷酸的多聚体(例如至少3个核苷酸的串)。在一些实施方案中,“核酸”涵盖RNA以及单链和/或双链DNA。核酸可以是天然存在的,例如,在基因组、转录物、

mRNA、tRNA、rRNA、siRNA、snRNA、质粒、粘粒、染色体、染色单体、或其它天然存在的核酸分子的背景中。另一方面，核酸分子可以是非天然存在的分子，例如重组DNA或RNA、人工染色体、工程化基因组、或其片段、或合成DNA、RNA、DNA/RNA杂合物，或者包括非天然存在的核苷酸或核苷。此外，术语“核酸”、“DNA”、“RNA”和/或相似术语包括核酸类似物，即具有磷酸二酯主链外的类似物。核酸可以从天然来源中纯化，使用重组表达系统生成，并且任选纯化，化学合成，等等。在适当的情况下，例如在化学合成分子的情况下，核酸可以包含核苷类似物，诸如具有经化学修饰的碱基或糖的类似物，和主链修饰。核酸序列以5'至3'方向呈现，除非另有指示。在一些实施方案中，核酸是或者包含天然核苷（例如腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷、和脱氧胞苷）；核苷类似物（例如2-氨基腺苷、2-硫代胸苷（2-thiothymidine）、肌苷、吡咯并-嘧啶、3-甲基腺苷、5-甲基胞苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮腺苷（7-deazaadenosine）、7-脱氮鸟苷、8-氧化腺苷（8-oxoadenosine）、8-氧化鸟苷（8-oxoguanosine）、0(6)-甲基鸟嘌呤、和2-硫代胞苷）；经化学修饰的碱基；经生物修饰的碱基（例如甲基化碱基）；插入碱基；经修饰的糖类（例如2'-氟核糖、核糖、2'-脱氧核糖、阿拉伯糖和己糖）；和/或经修饰的磷酸酯基团（例如硫代磷酸酯和5'-N-氨基磷酸酯连接）。

[0053] 如本文中使用的，术语“药物组合物”指可以在治疗疾病或病症的背景中对受试者施用的组合物。在一些实施方案中，药物组合物包含活性成分，例如核酸酶或编码核酸酶的核酸，和药学可接受赋形剂。

[0054] 如本文中使用的，术语“增殖性疾病”指细胞或组织稳态由于细胞或细胞群体展现出异常升高的增殖速率而受到扰乱的任何疾病。增殖性疾病包括超增殖性疾病，诸如新生前（pre-neoplastic）增生性状况和新生性疾病。新生性疾病以异常的细胞增殖为特征，并且包括良性和恶性新生物两者。恶性新生物又称为癌症。

[0055] 术语“蛋白质”、“肽”和“多肽”在本文中可互换使用，并且指通过肽（酰胺）键连接在一起的氨基酸残基的多聚体。该术语指任何大小、结构或功能的蛋白质、肽或多肽。通常，蛋白质、肽或多肽会长至少3个氨基酸。蛋白质、肽或多肽可以指个别的蛋白质或蛋白质集合。例如，可以通过添加化学实体，诸如碳水化合物基团、羟基基团、磷酸基团、法呢基基团、异法呢基基团、脂肪酸基团、用于缀合、功能化或其它修饰的接头等修饰蛋白质、肽或多肽中的一个或多个氨基酸。蛋白质、肽或多肽也可以是单一分子或者可以是多分子复合物。蛋白质、肽或多肽可以仅是天然存在的蛋白质或肽的片段。蛋白质、肽或多肽可以是天然存在的、重组的、或合成的、或者其任何组合。如本文中使用的，术语“融合蛋白”指包含来自至少两种不同蛋白质的蛋白质域的杂合多肽。一种蛋白质可以位于融合蛋白的氨基端（N端）部分或者位于羧基端（C端）蛋白质，如此分别形成“氨基端融合蛋白”或“羧基端融合蛋白”。蛋白质可以包含不同域，例如，核酸结合域（例如Cas9的gRNA结合域，其指导蛋白质对靶位点的结合）和核酸切割位点。在一些实施方案中，蛋白质包含蛋白质性部分，例如构成核酸结合域的氨基酸序列，和有机化合物，例如可以起核酸切割剂的作用的化合物。在一些实施方案中，蛋白质与核酸（例如RNA）在复合物中，或者与核酸（例如RNA）联合。可以通过本领域中已知的任何方法生成本文中提供的任何蛋白质。例如，可以经由重组蛋白表达和纯化（其尤其适合于包含肽接头的融合蛋白）生成本文中提供的蛋白质。用于重组蛋白表达和纯化的方法是公知的，并且包括那些由Green and Sambrook, Molecular Cloning: A Laboratory

Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) 描述的,其全部内容通过提及并入本文。

[0056] 术语“RNA可编程核酸酶”和“RNA引导性核酸酶”在本文中可互换使用,并且指与一种或多种不作为切割靶物的RNA形成复合物(例如结合或联合)的核酸酶。在一些实施方案中,RNA可编程核酸酶在与RNA的复合物中时可以称为核酸酶:RNA复合物。通常,结合的RNA称为引导RNA (gRNA)。gRNA可以以两个或更多个RNA的复合物或以单一RNA分子存在。以单一RNA分子存在的gRNA可以称为单一引导RNA (sgRNA),尽管“gRNA”可互换使用,指以单一分子或以两个或更多个分子的复合物存在的引导RNA。通常,以单一RNA种类存在的gRNA包含至少两个域:(1)与靶核酸共享同源性的域(例如,并且指导Cas9复合物对靶物的结合);和(2)结合Cas9蛋白的域。在一些实施方案中,域(2)是“sgRNA”主链,如图1-4中任一描述的。在一些实施方案中,域(2)对应于称为tracrRNA的序列,并且包含茎-环结构。例如,在一些实施方案中,域(2)与tracrRNA同源,如Jinek et al., Science 337:816-821 (2012) 的图1E中描绘的,其全部内容通过提及并入本文。在一些实施方案中,域2与图1-4中任一的“sgRNA主链”或如Jinek et al., Science 337:816-821 (2012) 描述的tracrRNA是至少90%、至少95%、至少98%或至少99%相同的。在一些实施方案中,gRNA包含两个或更多个域(1)和(2),并且可以称为“延长的gRNA”。例如,延长的gRNA会结合两个或更多个Cas9蛋白,并且在两个或更多个独特区域处结合靶核酸。gRNA包含与靶位点互补的核苷酸序列,其介导核酸酶/RNA复合物对所述靶位点的结合,并且提供核酸酶:RNA复合物的序列特异性。结合靶核酸的gRNA的序列可以包含与靶物的区域互补并且适合于核酸酶:RNA复合物结合的任何序列。在一些实施方案中,RNA-可编程核酸酶是(CRISPR-associated system)Cas9内切核酸酶,例如,来自酿脓链球菌的Cas9 (Csn1) (参见例如“Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.” Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L. expand/collapse author list McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001); “CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.” Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607 (2011); 及“A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. Science 337:816-821 (2012),每篇的全部内容通过提及并入本文。

[0057] 由于RNA可编程核酸酶(例如Cas9)使用RNA:DNA杂交来测定靶DNA切割位点,这些蛋白质能够在原则上切割由指导RNA规定的任何序列。使用RNA-可编程核酸酶,诸如Cas9进行位点特异性切割(例如以修饰基因组)的方法是本领域中已知的(参见例如Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823 (2013); Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823-826 (2013); Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology 31, 227-229 (2013); Jinek, M. et al. RNA-programmed genome editing in human cells. eLife 2, e00471 (2013); Dicarlo, J.E. et

a1.Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems.核酸s research (2013) ;Jiang, W. et al. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology 31, 233-239 (2013) ;每篇的全部内容通过提及并入本文)。

[0058] 术语“小分子”和“有机化合物”在本文中可互换使用,并且指具有相对较低的分子量的分子,无论是天然存在的或人工创建的(例如经由化学合成)。通常,有机化合物含有碳。有机化合物可以含有多个碳-碳键、立构中心、和其它官能团(例如胺、羟基、羧基或杂环状环)。在一些实施方案中,有机化合物是单体的,并且具有小于约1500g/mol的分子量。在某些实施方案中,小分子的分子量小于约1000g/mol或小于约500g/mol。在某些实施方案中,小分子是药物,例如,已经被合适的政府机构或管理机关视为在人或动物中使用方面安全且有效的药物。在某些实施方案中,已知小分子结合适体。在一些实施方案中,有机化合物是抗生素药物,例如,抗癌抗生素,诸如dynemicin、新制癌菌素(neocarzinostatin)、加利车霉素(calicheamicin)、埃斯波霉素(esperamicin)、博来霉素(bleomycin)或其衍生物。

[0059] 如本文中使用的,术语“受试者”指个别的生物体,例如,个别的哺乳动物。在一些实施方案中,受试者是人。在一些实施方案中,受试者是非人哺乳动物。在一些实施方案中,受试者是非人灵长类。在一些实施方案中,受试者是啮齿类。在一些实施方案中,受试者是绵羊、山羊、牛、猫或犬。在一些实施方案中,受试者是脊椎动物、两栖动物、爬虫动物、鱼、昆虫、苍蝇或线虫类。在一些实施方案中,受试者是研究动物。在一些实施方案中,受试者是经遗传工程化的,例如经遗传工程化的非人受试者。受试者可以是任何性别的并且在发育的任何阶段。

[0060] 如本文中在核酸酶的背景中使用的,术语“靶核酸”和“靶基因组”分别指包含给定核酸酶的至少一个靶位点的核酸分子或基因组。

[0061] 术语“靶位点”在本文中与术语“核酸酶靶位点”可互换使用,指被核酸酶结合并且切割的核酸分子内的序列。靶位点可以是单链的或双链的。在RNA引导性(例如RNA可编程的)核酸酶(例如包含Cas9 gRNA结合域和活性Cas9 DNA切割域的蛋白质二聚体)的背景中,靶位点通常包含与RNA可编程核酸酶的gRNA互补的核苷酸序列,和gRNA互补序列相邻的3'端的原间隔物相邻基序(PAM)。对于RNA引导性核酸酶Cas9,在一些实施方案中,靶位点可以是20个碱基对加3个碱基对PAM(例如NNN,其中N代表任何核苷酸)。通常,PAM的第一个核苷酸可以是任何核苷酸,而两个下游核苷酸是根据特定的RNA引导性核酸酶规定的。用于RNA引导性核酸酶(诸如Cas9)的示例性靶位点是本领域技术人员已知的,并且包括但不限于NNG、NGN、NAG和NGG,其中N代表任何核苷酸。另外,来自不同物种(例如嗜热链球菌代替酿脓链球菌)的Cas9核酸酶识别包含序列NGGNG的PAM。其它PAM序列是已知的,包括但不限于NNAGAAW和NAAR(参见例如Esvelt and Wang, Molecular Systems Biology, 9:641 (2013),其全部内容通过提及并入本文)。例如,RNA引导性核酸酶(诸如Cas9)的靶位点可以包含结构[N_z]-[PAM],其中独立地,每个N是任何核苷酸,并且z是1-50的整数。在一些实施方案中,z是至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少25、至少30、至少35、至少40、至少45或至少50。在一些实施方案中,z是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、或50。在一些实施方案中，Z是20。在一些实施方案中，“靶位点”也可以指核酸分子内被核酸酶结合但不切割的序列。

[0062] 术语“处理/治疗”指目的为逆转、减轻疾病或病症或其一种或多种症状、延迟疾病或病症或其一种或多种症状的发作、或者抑制疾病或病症或其一种或多种症状的过程的临床干预，如本文中描述的。如本文中使用的，“处理/治疗”指目的为逆转、减轻疾病或病症或其一种或多种症状、延迟疾病或病症或其一种或多种症状的发作、或者抑制疾病或病症或其一种或多种症状的过程的临床干预，如本文中描述的。在一些实施方案中，可以在已经形成一种或多种症状后和/或在已经诊断出疾病后施用治疗/处理。在其它实施方案中，可以在缺乏症状的情况下施用治疗，例如以预防疾病或延迟症状的发作或抑制疾病的发作或进展。例如，可以在症状发作前对易感性个体施用治疗/处理(例如鉴于症状史和/或鉴于遗传或其它易感性因素)。也可以在症状已经消退后继续治疗/处理，例如，以预防或延迟其复发。

[0063] 术语“载体”指包含本发明的一种或多种重组多核苷酸(例如那些编码本文中提供的gRNA和/或Cas9蛋白的重组多核苷酸)的多核苷酸。载体包括但不限于质粒、病毒载体、粘粒、人工染色体和噬菌粒。载体是能够在宿主细胞中复制，并且进一步以可以切割载体并且可以接受期望的核酸序列插入的一个或多个内切核酸酶限制性位点为特征的载体。载体可以含有一种或多种适合于用于鉴定和/或选择细胞的标志物序列，所述细胞已经或尚未用载体转化或遗传修饰。标志物包括例如，编码提高或降低对抗生素(例如卡那霉素、氨苄青霉素)或其它化合物的抗性或敏感性的蛋白质的基因、编码其活性通过本领域中已知的标准测定法可检测的酶(例如β-半乳糖苷酶、碱性磷酸酶或萤光素酶)的基因、和明显影响经转化的或经转染的细胞、宿主、集落或噬斑的表型的基因。本发明涵盖任何适合于转化宿主细胞(例如大肠杆菌、哺乳动物细胞，诸如CHO细胞、昆虫细胞等)的载体，例如属于pUC系列、pGEM系列、pET系列、pBAD系列、pTET系列、或pGEX系列的载体。在一些实施方案中，载体适合于转化用于重组蛋白生成的宿主细胞。用于选择和工程化改造载体和宿主细胞以表达gRNA和/或蛋白质(例如本文中提供的那些)、转化细胞、和表达/纯化重组蛋白的方法是本领域中公知的，并且由例如Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) 提供。

[0064] 发明详述

[0065] 位点特异性核酸酶是用于体外和体内靶向基因组修饰的有力工具。一些位点特异性核酸酶在理论上可以实现一定水平的对靶切割位点的特异性，其会容许在不影响任何其它基因组位点的情况下靶向用于切割的基因组中的单一独特位点。已经报告了活细胞中的核酸酶切割触发DNA修复机制，其经常例如经由同源重组或非同源末端连接而导致经切割并且修复的基因组序列的修饰。因而，基因组内的特定独特序列的靶向切割打开了用于体细胞中的基因靶向和基因修饰的新通道，包括难以用常规的基因靶向方法操作的细胞，诸如许多人体细胞或胚胎干细胞。可以在临床背景中使用核酸酶介导的对疾病相关序列，例如HIV/AIDS患者中的CCR-5等位基因或肿瘤新血管化必需的基因的修饰，并且两种位点特异性核酸酶目前在临床试验中(Perez, E.E. et al., “Establishment of HIV-1

resistance in CD4+T cells by genome editing using zinc-finger nucleases.” Nature Biotechnology.26, 808-816 (2008); ClinicalTrials.gov identifiers: NCT00842634, NCT01044654, NCT01252641, NCT01082926)。可以使用位点特异性核酸酶治疗的其它疾病包括例如与三重扩充相关的疾病(例如亨廷顿(Huntington)氏疾病、营养不良性肌强直(myotonic dystrophy)、脊髓小脑共济失调(spino-cerebellar ataxias)等)、囊性纤维化(cystic fibrosis) (通过靶向CFTR基因)、癌症、自身免疫性疾病、和病毒性感染。

[0066] 用位点特异性核酸酶介导的修饰的一个重要问题是脱靶核酸酶效应,例如与意图的靶序列相差一个或多个核苷酸的基因组序列的切割。脱靶切割的不期望副作用范围为基因靶向事件期间对不想要基因座的插入到临床情况中的重度并发症。由对受试者施用的内切核酸酶对编码必需基因功能或肿瘤阻抑物基因的序列的脱靶切割可以导致受试者的疾病或甚至死亡。因而,期望在设计具有使脱靶效应最小化的最大机会的核酸酶中采用新策略。

[0067] 在一些方面中,本公开内容的方法和组合物通过提供控制RNA引导性核酸酶的短暂活性和/或提高特异性的手段而代表相对于先前方法和组合物的改善。例如,在与靶物互补的RNA(例如gRNA)形成复合物后,本领域中已知的RNA引导性核酸酶(天然存在的和那些工程化的两者)通常结合并且切割DNA。本发明的方面涉及下述认识,即对RNA引导性核酸酶:RNA复合物对其靶物的结合时机具有短暂控制会通过最小化或控制复合物能够结合并且切割靶物的时间量来降低脱靶效应的可能性。另外,仅结合要切割的靶位点的工程化gRNA(例如使用具有在缺乏靶物的情况下阻断结合的延长的靶物识别域的gRNA)改善RNA引导性核酸酶的特异性,并且降低脱靶效应的机会。

[0068] 可以使用本文中提供的策略、方法、组合物、试剂盒和系统来控制任何RNA引导性核酸酶(例如Cas9)的活性和/或改善任何RNA引导性核酸酶(例如Cas9)的特异性。基于本公开内容,适合于与如本文中描述的经修饰的gRNA一起使用的核酸酶对于本领域技术人员会是明显的。

[0069] 在某些实施方案中,利用本文中提供的策略、方法、组合物、试剂盒和系统来控制RNA-引导性(例如RNA-可编程)核酸酶活性的时机。鉴于典型RNA引导性核酸酶在形成核酸酶:RNA复合物后识别并切割靶序列,本文中提供的经修饰的gRNA容许控制靶物结合和切割。其它方面提供了gRNA,其工程化改造为仅当存在意图的靶位点时结合靶位点,从而改善RNA-引导性核酸酶的特异性。虽然Cas9:gRNA复合物已经成功用于修饰细胞(Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 339, 819-823 (2013); Mali, P. et al. RNA-guided human genome engineering via Cas9. Science. 339, 823-826 (2013); Jinek, M. et al. RNA-programmed genome editing in human cells. eLife 2, e00471 (2013))和生物体(Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature Biotechnology. 31, 227-229 (2013))两者,使用Cas9:引导RNA复合物修饰斑马鱼胚胎的研究以与ZFN和TALEN相似的比率观察到毒性(例如脱靶效应)(Hwang, W.Y. et al. Nature Biotechnology. 31, 227-229 (2013))。因而,本公开内容的方面目的在于使用新的gRNA平台来降低Cas9脱靶效应的机会,所述新的gRNA平台控制靶物结合和切割的时机和/或改善RNA引导性核酸酶的特异性。

[0070] 虽然与DNA和DNA切割性核酸酶诸如Cas9特别相关,本文中提供的发明构思、方法、

组合物、策略、试剂盒和系统在此方面不受限制,而是可以应用于利用核酸模板诸如RNA来指导对靶核酸的结合的任何核酸:核酸酶系统。

[0071] 经修饰的指导RNA (gRNA)

[0072] 本公开内容的一些方面提供了工程化改造为具有“开始”和“关闭”状态两者的gRNA。然后,在一些方面中,gRNA可以统称为“可变换gRNA”。例如,当gRNA在阻止gRNA对靶核酸的结合的结构状态中时,可变换gRNA被说成为“关闭”状态。在一些方面中,在“关闭”状态中的gRNA可以结合其关联RNA引导性核酸酶(例如Cas9),然而,核酸酶:gRNA复合物(当gRNA为“关闭”状态时)不能结合靶核酸来介导切割。在其它方面中,在“关闭”状态中的gRNA不能结合其靶序列或RNA引导性核酸酶,诸如Cas9。相反,当gRNA在容许gRNA对靶核酸结合的结构状态(例如作为与RNA引导性核酸酶诸如Cas9的复合物)中时,可变换gRNA被说成为“开启”状态。本公开内容的一些实施方案提供了包含与RNA引导性核酸酶(诸如Cas9)有关的发明性gRNA的复合物及其使用方法。本公开内容的一些实施方案提供了编码此类gRNA和/或RNA引导性核酸酶(例如Cas9)的核酸。本公开内容的一些实施方案提供了包含此类编码核酸的表达构建体。

[0073] 基于适体的gRNA

[0074] 在一个实施方案中,提供了包含适体的gRNA。参见例如图1。例如,在一些实施方案中,gRNA经由核苷酸接头与适体连接,如本文中描述的。适体通常是基于RNA或肽的分子,其以亲和力(例如,竞争抗体:抗原相互作用)结合特定配体。在一些实施方案中,适体以约1nM-10μM、约1nM-1μM、约1nM-500nM或约1nM-100nM的K_d结合其配体。凭借基于RNA的适体,例如,那些在mRNA的核开关中找到的,配体对适体域的结合导致控制mRNA表达(例如翻译)的构象变化。RNA适体已经成功克隆,并且适合于其它分子,例如以控制基因表达,或者已经使用SELEX针对特定配体工程化改造/选择(参见例如Dixon et al., “Reengineering orthogonally selective riboswitches.” PNAS2010;107 (7):2830-2835;Suess et al., “A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo.” Nucleic Acids Res.2004;32 (4):1610-1614;Ellington and Szostak, “In vitro selection of RNA molecules that bind specific ligands.” Nature.1990;346:818-822;Tuerk and Gold, “Systematic evolution of ligands by exponential enrichment:RNA ligands to bacteriophage T4DNA polymerase.” Science.1990;249:505-510;Burke and Gold, “RNA aptamers to the adenosine moiety of S-adenosyl methionine:structural inferences from variations on a theme and the reproducibility of SELEX.” Nucleic Acids Res.1997;25 (10):2020-4;Ulrich et al., “DNA and RNA aptamers:from tools for basic research towards therapeutic applications.” Comb Chem High Throughput Screen.2006;9 (8):619-32;Svobodová et al., “Comparison of different methods for generation of single-stranded DNA for SELEX processes.” Anal Bioanal Chem.2012;404:835-842;每篇的全部内容在此通过提及并入)。结合适体的配体包括但不限于小分子、代谢物、碳水化合物、蛋白质、肽或核酸。如图1中显示的,与适体连接的gRNA在缺乏结合适体的特定配体的情况下以“关闭”状态存在。通常,“关闭”状态由结构特征介导,所述结构特征阻止与靶核酸杂交的整个或部分gRNA序列自由杂交靶核酸。例如,在一些方面中,设计包含适体的gRNA,使得适体序列的一部分

与整个或部分与靶物杂交的gRNA序列杂交。可以使用本领域中已知的方法工程化改造结合靶物的gRNA的序列(例如在图1C,D中描绘为“切割靶物的引导”,在本文中称为“引导”序列),以包含靶向任何期望核酸靶物的任何序列,并且因此不限于例示性的图中描绘的序列。类似地,任何合适的适体可以在gRNA序列的5’或3’连接,并且使用本领域中的常规方法修饰以包含会与gRNA中的特定引导序列杂交的核苷酸。在一些实施方案中,与本文中提供的任何gRNA连接的适体是RNA适体,如本文中描述的。在一些实施方案中,RNA适体衍生自(例如克隆自)核开关。任何核开关可以在RNA适体中使用。例示性的核开关包括但不限于茶碱核开关、焦磷酸硫胺素(TPP)核开关、腺昔钴胺素(AdoCbl)核开关、S-腺昔甲硫氨酸(SAM)核开关、SAH核开关、黄素单核苷酸(FMN)核开关、四氢叶酸核开关、赖氨酸核开关、甘氨酸核开关、嘌呤核开关、G1mS核开关和前Q核苷₁(PreQ1)核开关。在一些实施方案中,RNA适体衍生自茶碱核开关。在一些实施方案中,自茶碱核开关衍生的适体包含SEQ ID N0:3。在一些实施方案中,可以修饰SEQ ID N0:3的加下划线粗体部分,使得本文中的任何核苷酸用任何其它核苷酸替换,和/或可以通过添加或缺失1或更多个核苷酸修饰。例如,可以修饰加下划线粗体部分,使得包含与部分或整个与靶核酸杂交的gRNA序列的序列。参见例如图1C。在一些实施方案中,RNA适体与SEQ ID N0:3是至少80%、至少85%、至少90%、至少95%、至少98%或至少99%相同的。

[0075] 5’-GGUGAUACCAG**CAUCGUCUUGAUG**CCCUUGGCAGCACC-3’ (SEQ ID N0:3)

[0076] 在一些实施方案中,适体是非天然存在的(例如不存在于自然界中)。例如,在一些实施方案中,使用SELEX工程化改造或者从文库选择适体。在一些实施方案中,适体包含至少20、至少30、至少40、至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140、至少150、至少175、至少200、至少250、或至少300个核苷酸。在一些实施方案中,适体包含20-200、20-150、20-100、或20-80个核苷酸。在一些实施方案中,提供的RNA(例如包含与适体连接的gRNA的RNA)的gRNA部分包含至少50、至少60、至少70、至少80、至少90、至少100、至少110、至少120、至少130、至少140、至少150、至少175、或至少200个核苷酸。在一些实施方案中,gRNA部分包含60-150、60-100、或60-80个核苷酸。

[0077] mRNA感测性gRNA

[0078] 根据另一个实施方案,提供了在某些条件下(例如在存在代谢物、小分子、核酸等的情况下)结合靶核酸的gRNA。在一些实施方案中,使gRNA在结构上阻止(例如为“关闭”状态)结合靶物(例如与靶物杂交),除非另一个分子结合gRNA(例如与gRNA杂交),导致与“开启”状态对应的结构重排。在一些实施方案中,特定转录物(例如mRNA)对gRNA的结合使gRNA从“关闭”状态转向“开启”状态。参见例如图2。此类gRNA称为“mRNA感测性”gRNA。例如,在一些方面中,提供了gRNA,其包含:(i)与靶核酸区域杂交的区域(例如“引导”序列);(ii)与区域(i)的序列部分或完全杂交的另一个区域(例如“引导块”序列);和(iii)与转录物(mRNA)的区域杂交的区域(例如“转录物传感器”)。在一些实施方案中,每个区域(例如i-iii)包含至少5、至少10、至少15、至少20或至少25个核苷酸。在一些实施方案中,gRNA形成茎-环结构。在一些实施方案中,茎包含与区域(ii)的部分或整个序列杂交的区域(i)的序列,并且环由区域(iii)的部分或整个序列形成。在一些实施方案中,区域(ii)和(iii)在区域(i)的5’或3’。参见例如图2A对图2C。可以使用本领域中已知的方法工程化改造结合靶物的gRNA的序列(例如“引导”序列),以包括靶向任何期望核酸靶物的任何序列,并且因此不限于例

示性的图中描绘的序列。类似地,可以使用本领域中常规的方法工程化改造区域(iii) (例如转录物传感器)以包含与感兴趣mRNA杂交的任何序列。类似地,可以使用本领域中常规的方法工程化改造区域(ii)以包含与部分或整个“引导”序列杂交的序列。例如,在一些方面中,mRNA是在细胞中表达时,期望靶核酸(例如基因)的基因组修饰的。如此,在缺乏mRNA的情况下,gRNA当投递到细胞(或在细胞中表达)时仍然为“关闭”状态。当存在(表达)mRNA时,它结合gRNA的转录物传感器,导致阻止“引导”序列对靶核酸杂交的茎-环结构的解折叠,从而“开启”gRNA。参见例如图2B和2D。以“开启”状态提供的gRNA能够联合并引导RNA引导性核酸酶(例如Cas9蛋白)以结合靶核酸。

[0079] 延长的-xDNA感测性(xDNA感测性)gRNA

[0080] 根据另一个实施方案,提供了保持为“关闭”状态的经修饰的gRNA,除非gRNA与靶核酸的至少两个独特区域杂交。参见例如图3。此类gRNA提供了改善的对RNA引导性核酸酶(例如Cas9)的特异性,因为它们有效延长特定gRNA/靶物相互作用的识别序列。此类gRNA称为“xDNA感测性”gRNA(“x”是“延长的”DNA识别的缩写)。例如,提供了gRNA,其包含:(i)与靶核酸区域杂交的区域(例如“引导”序列);(ii)与区域(i)的序列部分或完全杂交的另一个区域(例如“引导块”序列);和(iii)与靶核酸的另一个区域杂交的区域(例如“xDNA传感器”)。在一些实施方案中,xDNA传感器必须首先结合靶核酸,之后引导序列能够结合靶物。在一些实施方案中,xDNA传感器结合引导序列结合的靶物的相同链。在一些实施方案中,xDNA传感器和引导序列结合靶核酸的不同链。在一些实施方案中,区域(i)和(ii)的序列包含至少5、至少10、至少15、至少20或至少25个核苷酸。在一些实施方案中,区域(iii)的序列包含至少5、至少10、至少15、至少20、至少25、至少30、至少40、至少50、至少75或至少100个核苷酸。在一些实施方案中,gRNA形成茎-环结构。例如,在一些实施方案中,茎包含与区域(ii)的部分或整个序列杂交的区域(i)的序列,并且环由区域(iii)的部分或整个序列形成。在一些实施方案中,区域(i)和(iii)包含gRNA中相邻的序列。在一些实施方案中,区域(ii)和(iii)在区域(i)的5'或3'。参见例如图3A对图3C。在一些实施方案中,区域(ii)位于区域(i)和(iii)之间。可以使用本领域中已知的方法工程化改造结合靶物的gRNA的序列(例如“引导”序列)以包含靶向任何期望的核酸靶物的任何序列,并且因此不限于示性的图中描绘的序列。类似地,可以使用本领域中常规的方法工程化改造区域(iii)(例如xDNA传感器)以包含与靶核酸的另一个区域(例如与被“引导”序列靶向的区域不同的区域)杂交的任何序列。同样地,可以使用本领域中常规的方法工程化改造区域(ii)以包含与部分或整个“引导”序列杂交的序列。如此,在缺乏正确的靶核酸(例如包含两个区域的靶物,gRNA设计为与所述两个区域杂交)的情况下,gRNA在投递至细胞(或者在细胞中表达)时仍然为“关闭”状态。不希望限于任何具体的理论,预期在gRNA(例如与Cas9联合时)与靶核酸开始接触时,xDNA传感器与靶物杂交,这继而解开阻断“引导”序列的茎-环结构,将gRNA“开启”。若它是正确的靶核酸,则引导序列然后会与靶物杂交,并且任选地复合物会切割靶核酸。参见例如图3B和3D。

[0081] 复合物

[0082] 在一些实施方案中,提供了复合物,其包含本文中提供的任何RNA/gRNA(例如包含与适体连接的gRNA的RNA,感测mRNA的gRNA,或者包含xDNA传感器的gRNA)。在一些方面中,提供了复合物,其包含与RNA引导性核酸酶联合的提供的RNA/gRNA。在一些实施方案中,RNA

引导性核酸酶是Cas9、Cas9的变体、或Cas9的片段,例如如本文中描述的。在一些实施方案中,RNA引导性核酸酶是Cas9蛋白的任何形式,如2013年9月6日提交的美国临时专利申请U.S.S.N.61/874,609,题目为“Cas9 Variants And Uses Thereof”和2013年9月6日提交的美国临时专利申请U.S.S.N.61/874,746,题目为“Delivery System For Functional Nucleases”中提供的,在此通过提及完整收录每篇的全部内容。

[0083] 在一些实施方案中,复合物进一步包含配体,例如结合与RNA引导性核酸酶联合的RNA适体的配体,如本文中描述的。在一些实施方案中,复合物(例如包含提供的RNA(gRNA):配体:Cas9蛋白)结合并且任选切割靶核酸。在一些方面中,包含“感测性”gRNA(例如mRNA或xDNA)和Cas9的复合物结合并且任选切割靶核酸。

[0084] 药物组合物

[0085] 在一些实施方案中,提供本文中描述的任何gRNA作为药物组合物的一部分。在一些实施方案中,药物组合物进一步包含与发明性gRNA形成复合物的RNA引导性核酸酶(例如Cas9)。例如,一些实施方案提供了药物组合物,其包含如本文中提供的gRNA和RNA引导性核酸酶或编码此类gRNA和/或核酸酶的核酸,和药学可接受赋形剂。任选地,药物组合物可以包含一种或多种别的治疗活性物质。

[0086] 在一些实施方案中,对受试者(例如对人受试者)施用本文中提供的组合物以实现受试者中的靶向基因组修饰。在一些实施方案中,自受试者获得细胞,并且使该细胞与联合RNA引导性核酸酶或其编码核酸的提供的gRNA离体接触。在一些实施方案中,将自受试者取出并且与发明性gRNA:核酸酶复合物离体接触的细胞再导入受试者中,任选地在已经在细胞中实现或检出期望的基因组修饰后。投递包含核酸酶的药物组合物的方法是已知的,并且记载于例如,美国专利No.6,453,242;6,503,717;6,534,261;6,599,692;6,607,882;6,689,558;6,824,978;6,933,113;6,979,539;7,013,219;和7,163,824,全部的公开内容通过提及完整并入本文。虽然本文中提供的药物组合物的描述主要涉及适合于对人施用的药物组合物,但是熟练技术人员应当理解,此类组合物一般适合于对所有类别的动物或生物体施用。修饰适合于对人施用的药物组合物以使组合物适合于对各种动物施用是完全了解的,并且普通技术的兽医药理学家可以仅通过常规(若有的话)实验设计和/或进行此类修饰。涵盖药物组合物施用的受试者包括但不限于人和/或其它灵长类;哺乳动物、驯养动物、宠物和商业相关哺乳动物,诸如牛、猪、马、山羊、猫、犬、小鼠和/或大鼠;和/或禽类,包括商业相关禽类,诸如鸡、鸭、鹅和/或火鸡。

[0087] 可以通过药理学领域中已知或后来开发的任何方法制备本文中描述的药物组合物的制剂。一般地,此类制备方法包括下述步骤,即使活性成分与赋形剂和/或一种或多种其它辅助成分联合,然后若必要和/或想要的话,运输和/或包装产品成期望的单剂量或多剂量形式。

[0088] 另外,药物组合物可以包含药学可接受赋形剂,如本文中使用的,其包括任何和所有溶剂、分散介质、稀释剂或其它液体媒介物、分散或悬浮助剂、表面活性剂、等张剂、增稠剂或乳化剂、防腐剂、固体粘合剂、润滑剂等,如适合于期望的特定剂量形式的。Remington's The Science and Practice of Pharmacy, 21st Edition, A.R. Gennaro (Lippincott, Williams & Wilkins, Baltimore, MD, 2006; 其通过提及完整并入本文)公开了配制药物组合物中使用的各种赋形剂及其已知的制备技术。关于其它适合于生成包含核酸酶的药物组合

物的方法、试剂、赋形剂和溶剂,还可参见PCT申请PCT/US2010/055131(公开文本号W02011053982 A8,2010年11月2日),其通过提及完整并入本文。除非任何常规赋形剂介质与物质或其衍生物不相容,诸如通过产生任何不想要的生物效应或以其它方式与药物组合物的任何其它组分以有害的方式相互作用,涵盖其用途在本公开内容的范围内。

[0089] 在一些实施方案中,可以使用根据本发明的组合物来制备多种疾病、病症和/或状况之任一,包括但不限于下列一项或多项:自身免疫性病症(例如糖尿病、狼疮、多发性硬化、银屑病、类风湿性关节炎);炎性病症(例如关节炎、盆腔炎性疾病);传染病(例如病毒性感染(例如HIV、HCV、RSV)、细菌性感染、真菌性感染、败血症);神经学病症(例如阿尔茨海默氏病、亨廷顿氏病;孤独症;迪谢内肌营养不良(Duchenne muscular dystrophy));心血管病症(例如动脉粥样硬化、高胆固醇血症(hypercholesterolemia)、血栓形成(thrombosis)、凝固病症、血管生成性病症,诸如黄斑变性);增殖性病症(例如癌症,良性新生物);呼吸病症(例如慢性阻塞性肺疾病);消化病症(例如炎性肠病、溃疡);肌肉骨骼病症(例如纤维肌痛(fibromyalgia),关节炎(arthritis));内分泌、代谢和营养病症(例如糖尿病,骨质疏松);泌尿学病症(例如肾疾病);心理学病症(例如抑郁,精神分裂症);皮肤病症(例如创伤、湿疹);血液和淋巴病症(例如贫血,血友病)等。

[0090] 用于位点特异性核酸切割的方法

[0091] 在本公开内容的另一个实施方案中,提供了用于位点特异性核酸(例如DNA)切割的方法。在一些实施方案中,方法包括使DNA与本文中描述的任何Cas9:RNA复合物接触。例如,在一些实施方案中,方法包括在适合于Cas9核酸酶切割DNA的条件下使DNA与复合物接触,所述复合物包含:(i)与如本文中描述的适体连接的gRNA,其中gRNA包含与DNA的部分结合的序列;(ii)与gRNA的适体结合的配体;和(iii)RNA引导性核酸酶(例如Cas9蛋白)。

[0092] 在一些实施方案中,提供了用于诱导细胞中的位点特异性DNA切割的方法。在一些实施方案中,方法包括:(a)使细胞接触或在细胞内表达包含如本文中描述的适体的gRNA,其中gRNA包含能够结合DNA靶序列的序列;(b)使细胞接触或在细胞内表达RNA引导性核酸酶(例如Cas9蛋白);以及(c)使细胞与结合gRNA的适体的特定配体接触,导致切割DNA靶物的gRNA:配体:Cas9复合物的形成。在一些实施方案中,方法包括:(a)使细胞与复合物接触,所述复合物包含Cas9蛋白和包含如本文中描述的适体的gRNA,其中gRNA包含能够结合DNA靶序列的序列;并且(b)使细胞与结合gRNA适体的特定配体接触,导致切割DNA靶物的gRNA:配体:Cas9复合物的形成。在一些实施方案中,同时进行步骤(a)和(b)。在一些实施方案中,序贯进行步骤(a)和(b)。如此,在一些实施方案中,其中在细胞与复合物接触后使细胞与配体接触,实现切割的控制,因为切割仅在已经将配体投递到细胞后发生。在这些方法的一些实施方案中,不将配体投递到细胞,而是由细胞内部生成,例如作为生理学或病理生理学过程的一部分。

[0093] 在一些实施方案中,提供了用于位点特异性DNA切割的方法,其利用如本文中描述的mRNA感测性gRNA。例如,在一些实施方案中,方法包括使DNA与复合物接触,所述复合物包含RNA引导性核酸酶(例如Cas9蛋白)和mRNA感测性gRNA,其中gRNA包含:(i)与靶核酸的区域杂交的区域;(ii)与区域(i)的序列部分或完全杂交的另一个区域;和(iii)与转录物(mRNA)的区域杂交的区域。在一些实施方案中,在区域(iii)中的序列与mRNA杂交后发生切割。

[0094] 在其它实施方案中,提供了用于位点特异性DNA切割的方法,其利用如本文中描述的xDNA感测性gRNA。例如,在一些实施方案中,方法包括使DNA与复合物接触,所述复合物包含RNA引导性核酸酶(例如Cas9蛋白)和xDNA感测性gRNA,其中gRNA包含:(i)与靶核酸的区域杂交的区域;(ii)与区域(i)的序列部分或完全杂交的另一个区域;和(iii)与靶核酸的另一个区域杂交的区域。在一些实施方案中,在区域(iii)中的序列与不被“引导”序列靶向的靶核酸区域杂交后发生切割。

[0095] 在一些实施方案中,可以对细胞中的DNA进行本文中提供的任何方法。例如,在一些实施方案中,由本文中提供的任何包含RNA/gRNA的复合物接触的DNA在真核细胞中。在一些实施方案中,真核细胞在个体中。在一些实施方案中,个体是人。在一些实施方案中,在体外进行本文中提供的任何方法。在一些实施方案中,在体内进行本文中提供的任何方法。

[0096] 多核苷酸、载体、细胞、试剂盒

[0097] 在本公开内容的另一个实施方案中,提供了多核苷酸,其编码本文中描述的任何gRNA(和任选地任何Cas9蛋白)。例如,提供了编码本文中描述的任何gRNA和/或Cas9蛋白的多核苷酸,例如用于重组表达和纯化发明性gRNA,或包含其的复合物,例如包含发明性gRNA和RNA引导性核酸酶(例如Cas9蛋白)的复合物。在一些实施方案中,提供的多核苷酸包含单独或与编码本文中描述的任何Cas9蛋白的序列组合的一种或多种编码gRNA的序列。

[0098] 在一些实施方案中,提供了编码本文中描述的任何gRNA(和任选地任何Cas9蛋白)的载体,例如用于重组表达和纯化发明性gRNA,或包含发明性gRNA和RNA引导性核酸酶(例如Cas9蛋白)的复合物。在一些实施方案中,载体包含或工程化改造为包含多核苷酸,例如那些在本文中描述的。在一些实施方案中,载体包含一种或多种编码gRNA和/或任何Cas9蛋白(例如如本文中描述)的序列。通常,载体包含与启动子可操作连接的编码发明性gRNA的序列,使得在宿主细胞中表达gRNA。

[0099] 在一些实施方案中,提供了用于重组表达和纯化本文中描述的任何gRNA(和任选地任何Cas9蛋白)的细胞。细胞包括适合于重组RNA表达和任选地蛋白质表达的任何细胞,例如,包含表达或能够表达发明性gRNA的遗传构建体的细胞(例如已经用本文中描述的一种或多种载体转化的细胞,或者具有基因组修饰的细胞,所述细胞从已经掺入细胞基因组中的等位基因表达发明性gRNA和任选地本文中提供的任何Cas9蛋白)。用于转化细胞,遗传修饰细胞以及在此类细胞中表达基因和蛋白质的方法是本领域中公知的,并且包括那些由例如Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)) 和Friedman and Rossi, Gene Transfer: Delivery and Expression of DNA and RNA, A Laboratory Manual (1st ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2006)) 提供的。

[0100] 本公开内容的一些方面提供了试剂盒,其包含本文中提供的任何发明性gRNA或复合物和任选地本文中描述的任何Cas9蛋白。在一些实施方案中,试剂盒包含任何多核苷酸,其编码提供的gRNA和任选地任何Cas9蛋白。在一些实施方案中,试剂盒包含用于重组表达任何发明性gRNA和任选地任何Cas9蛋白的载体。在一些实施方案中,试剂盒包含细胞,该细胞包含用于表达本文中提供的任何发明性gRNA、复合物和任选地任何Cas9蛋白的遗传构建体。在一些实施方案中,试剂盒包含赋形剂和用法说明,其关于使任何发明性组合物与赋形

剂接触以产生组合物,该组合物适合于使核酸与例如发明性gRNA和RNA引导性核酸酶诸如Cas9的复合物接触。在一些实施方案中,组合物适合于接触基因组内的核酸。在一些实施方案中,组合物适合于将发明性组合物(例如gRNA,其与Cas9的复合物)投递到细胞。在一些实施方案中,组合物适合于将发明性组合物(例如gRNA,其与Cas9的复合物)投递至受试者。在一些实施方案中,赋形剂是药学可接受赋形剂。

[0101] 等同方案和范围

[0102] 本领域技术人员会认可或者仅使用常规实验便确认本文中描述的发明的具体实施方案的许多等同方案。本发明的范围不限于上述描述,而是如所附权利要求书中列出的。

[0103] 在权利要求书中,冠词诸如“一个/种”和“该/所述”可以意指一个/种或超过一个/种,除非相反指出或者另有从上下文下看明显。如果一个、超过一个、或全部的组成员在给定的产物或方法中存在、采用、或与给定的产物或方法以其它方式相关,那么认为在组的一个或多个成员间包含“或”的权利要求或描述得到满足,除非相反指出或者另有从上下文下看明显。本发明包括组的一个成员在给定产物或方法中确切存在、采用、或与给定的产物或方法以其它方式相关的实施方案。本发明还包括超过一个或全部组成员在给定产物或方法中确切存在、采用、或与给定的产物或方法以其它方式相关的实施方案。

[0104] 此外,应当理解,本发明涵盖所有变型、组合、和排列,其中来自一个或多个权利要求或说明书的相关部分的一个或多个限定、要素、字句、描述性术语等引入另一项权利要求。例如,从属于另一项权利要求的任何权利要求可以修饰为包含从属于同一基础权利要求的任何其它权利要求中找到的一个或多个限定。此外,在权利要求叙述组合物的情况下,应当理解,包括出于本文中公开的任何目的使用组合物的方法,和根据本文中公开的任何生成方法或本领域中已知的任何生成组合物的方法,除非另有指示或者除非可引起冲突或不一致性对于本领域普通技术人员会是明显的。

[0105] 在要素以列表呈现的情况下,例如在马库斯组形式中,应当理解还公开了要素的每个亚组,并且可以从该组除去任何要素。还注意到,术语“包括”意图是开放的,并且容许包括其他要素或步骤。应当理解,一般地,在发明或发明的方面称为包含特定要素、特征、步骤等的情况下,本发明的某些实施方案或本发明的范围由此类元素和/或特征组成或基本上由此类元素、特征、步骤等组成。为了简洁的目的,那些实施方案在本文中尚未用同样的文字明确列出。如此,对于包含一个或多个要素、特征、步骤等的本发明的每个实施方案,本发明还提供了由或基本上由那些要素、特征、步骤等组成的实施方案。

[0106] 在给出范围的情况下,包括端点。此外,应当理解,除非另外指出或者另外从上下文和/或本领域普通技术人员的理解看明显,表示为范围的数值在本发明的不同实施方案中可以假设所述范围内的任何具体的数值(到范围下限单位的十分之一),除非尚显文另有明确叙述。还应当理解,除非另外指出或者另外从上下文和/或本领域普通技术人员的理解看明显,表示为范围的数值可以假设给定范围内的任何子范围,其中子范围的端点以与范围下限单位的十分之一的精确性相同的程度表示。

[0107] 另外,应当理解本发明的任何具体的实施方案可以明确从任何一个或多个权利要求排除。在给出范围的情况下,可以从一个或多个权利要求明确排除范围内的任何数值。本发明的组合物和/或方法的任何实施方案、要素、特征、应用或方面可以从任何一个或多个权利要求排除。为了简短的目的,本文中没有明确列出排除一个或多个要素、特征、目的或

方面的所有实施方案。

[0108] 本文中提及的所有出版物、专利和序列数据库条目(包括上文列出的那些项)在此通过提及完整收入,就像明确并个别指定通过提及并入每篇个别的出版物或专利一样。在冲突的情况下,以本申请(包括本文中的定义)为准。

序列表

<110> 哈佛大学的校长及成员们
<120> 可变换 CAS9 核酸酶及其用途
<130> H0824.70148W000
<140> PCT/US2014/054252
<141> 2014-09-05
<150> US 14/326, 329
<151> 2014-07-08
<150> US 14/326, 340
<151> 2014-07-08
<150> US 14/326, 361
<151> 2014-07-08
<150> US 61/874, 682
<151> 2013-09-06
<160> 16
<170> PatentIn version 3.5
<210> 1
<211> 4104
<212> DNA
<213> 酿脓链球菌

[0001]	<400> 1 atggataaga aatactcaat aggcttagat atcggcacaa atagcgtcg atggcggtg atcaactgatg attataaggt tccgtctaaa aagtcaagg ttctggaaa tacagaccgc cacagiatca aaaaaaatct tatagggct cttttatttg gcagtggaga gacagcggaa gcgactcgctc tcaaaccgac agctcgtaga aggtatacac gtcggaaagaa tcgtatttg tatctacagg agatitttc aaatgagatg gcgaaagtagt atgatagttt ctticatcgaa cttgaagagt cttttttgtt ggaagaagac aagaagcatg aacgtcatcc tattttggaa aatatactatgat atgaagtgc ttatcatgag aaatatccaa ctatctatca tctgcggaaa aaattggcag attctactgta taaagcggat ttgcgtttaa tctatttgc cttagcgcatt atgattaagt ttcgtggtca tttttgatt gagggagatt taaatcctga taatagtgt gtggacaaac tatttaccca gttggtacaa atctacaatc aattatttga agaaaaccct attaaacccaa gtagagttaga tgctaaagcg attctttctg cacgatttagt taaatcaaga cgatttagaaa atctcattgc tcagctcccc ggtgagaaga gaaatggctt gtttggaaat ctcattgcctt tgcattggg attgaccctt aattttaaat caaattttga tttggcagaa gatgttaat tacagtttc aaaagataact tacgtatgtt atttagataaa tttatttgc caaattggag atcaaatatgc tgatttgc ttggcagctt agaattttatc agatgttatt ttactttcag atatcctaaatg agttaatagt gaaataacta aggctccccctt atcagcttca atgattaagc gctacgtatgaa acatcatcaa gacttgcattc ttttaaaagc tttagttcgaa caacaacttc cagaaaaagta taaagaaatc ttttttgcattc aatcaaaaaa cggatatgc ggttataatttgc atgggggagc tagccaagaa gaattttata aattttatcaa accaatttt gaaaaaatgg atggactgaa ggaatttttg gtgaaactaa atcgtgaaga tttgtcgcc gaaaaaatgg atggactgaa ggaatttttg gtgaaactaa atcgtgaaga tttgtcgcc	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200
--------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------

cgtaaaaaaag actgggatcc aaaaaaatat ggtggtttg atagtccaac ggtagcttat 3420
 tcagtccatg tggttgctaa ggtggaaaaa gggaaatcga agaagttaaa atccgttaaa 3480
 gagttactag ggatcacaat tatggaaaga agttcccttg aaaaaaatcc gattgacttt 3540
 tttagaagcta aaggatataa ggaagttaaa aaagacttaa tcattaaact acctaaatat 3600
 agtcttttg agttagaaaaa cggtcgtaaa cggtcgctgg ctgtgcgg agaattacaa 3660
 aaaggaaatg agctggctct gccaagcaaa tatgtgaatt ttttatattt agctagtcat 3720
 tatgaaaagt tgaaggtag tccagaagat aacgaacaaa aacaattgtt tgtggagcag 3780
 cataaggcatt atttagatga gattatttag caaatcagtg aatttctaa gcgtgttatt 3840
 ttagcagatg ccaattttaga taaagttctt agtgcata acaaacatag agacaaacca 3900
 atacgtgaac aagcagaaaaa tattattcat ttatttacgt tgacgaatct tggagctccc 3960
 gctgctttta aatatttta tacaacaatt gatcgtaaac gatatacgtc tacaaaagaa 4020
 gtttagatg ccactcttat ccatcaatcc atcactggtc tttatgaaac acgcattgat 4080
 ttgagtcagc taggaggtga ctga 4104

<210> 2
 <211> 1367
 <212> PRT
 <213> 酿脓链球菌

<400> 2

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
[0003] 1 5 10 15

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

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

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

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

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

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

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

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

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

145	150	155	160
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro			
165	170	175	
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Ile Tyr			
180	185	190	
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Arg Val Asp Ala			
195	200	205	
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn			
210	215	220	
Leu Ile Ala Gln Leu Pro Gly Glu Lys Arg Asn Gly Leu Phe Gly Asn			
225	230	235	240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe			
245	250	255	
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp			
260	265	270	
Asp Asp Leu Asp Asn Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp			
275	280	285	
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp			
[0004] 290	295	300	
Ile Leu Arg Val Asn Ser Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser			
305	310	315	320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Lys			
325	330	335	
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe			
340	345	350	
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser			
355	360	365	
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp			
370	375	380	
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg			
385	390	395	400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu			
405	410	415	
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe			
420	425	430	
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile			

435

440

445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

[0005] Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590

Ala Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620

Leu Phe Glu Asp Arg Gly Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Tyr
 645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly His Ser Leu
 705 710 715 720

His Glu Gln Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

	725	730	735
	Ile Leu Gln Thr Val Lys Ile Val Asp Glu Leu Val Lys Val Met Gly		
	740	745	750
	His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr		
	755	760	765
	Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu		
	770	775	780
	Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val		
	785	790	795
	Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln		
	805	810	815
	Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu		
	820	825	830
	Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Ile Lys Asp		
	835	840	845
	Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly		
	850	855	860
[0006]	Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn		
	865	870	875
	Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe		
	885	890	895
	Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys		
	900	905	910
	Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys		
	915	920	925
	His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu		
	930	935	940
	Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys		
	945	950	955
	Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu		
	965	970	975
	Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val		
	980	985	990
	Gly Thr Ala Leu Ile Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val		
	995	1000	1005
	Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys		

1010	1015	1020
Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr		
1025	1030	1035
Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn		
1040	1045	1050
Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr		
1055	1060	1065
Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg		
1070	1075	1080
Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu		
1085	1090	1095
Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg		
1100	1105	1110
Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys		
1115	1120	1125
Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu		
1130	1135	1140
[0007] Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser		
1145	1150	1155
Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe		
1160	1165	1170
Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu		
1175	1180	1185
Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe		
1190	1195	1200
Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu		
1205	1210	1215
Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn		
1220	1225	1230
Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro		
1235	1240	1245
Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His		
1250	1255	1260
Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg		
1265	1270	1275
Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr		

1280	1285	1290
Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile		
1295	1300	1305
Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe		
1310	1315	1320
Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr		
1325	1330	1335
Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly		
1340	1345	1350
Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
1355	1360	1365
<210> 3		
<211> 38		
<212> RNA		
<213> 人工序列		
<220>		
<223> 合成多核苷酸		
<400> 3		38
ggugauacca gcaucgucuu gaugccuug gcagcacc		
[0008] <210> 4		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成多核苷酸		
<400> 4		24
ggcagatgt a gtgttccac aggg		
<210> 5		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成多核苷酸		
<400> 5		24
ccctgtggaa acactacatc tgcc		
<210> 6		
<211> 133		
<212> RNA		
<213> 人工序列		
<220>		
<223> 合成多核苷酸		
<400> 6		60
ggugauacca gcaucgucuu gaugccuug gcagcaccgc cugcgcagg gguaucaggc		
agauguagug uuuccacagu uuuagagcua ugcugaaaag cauagcaagu uaaaaauaagg		120

cuaguccguu auc	133
<210> 7	
<211> 107	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 7	
uacaucugcc uugugagagu ugaaguugua uggcagaugu aguguuucca caguuuuaga	60
gcuaugcuga aaagcauagc aaguuaaaaau aaggcuaguc cguuauc	107
<210> 8	
<211> 90	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 8	
ggcagauhua guguuuccac aguuuuagag cuaucugaa aagcauagca aguuaaaaauu	60
auguugaagu ugagagugu uacaucugcc	90
<210> 9	
<211> 107	
<212> RNA	
<213> 人工序列	
[0009]	
<220>	
<223> 合成多核苷酸	
<400> 9	
uacaucugcc uugugagagu ugaaguugua uggcagaugu aguguuucca caguuuuaga	60
gcuaugcuga aaagcauagc aaguuaaaaau aaggcuaguc cguuauc	107
<210> 10	
<211> 117	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 10	
ggcagauhua guguuuccac aguuuuagag cuaucugaa aagcauagca aguuaaaaaua	60
aggcuagucc guuaacaacu ugaaaaagug gugaaguuga gaguguuuac aucugcc	117
<210> 11	
<211> 134	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 11	
ggugauacca gcaucguuug augccuugg cagcacgcu ggcgcagggg uaucaacagg	60
cagauguagu guuuccacag uuuuagagcu augcugaaaa gcauagcaag uaaaaauaag	120

gcuaguccgu uauc	134
<210> 12	
<211> 21	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 12	
auacaacuuc aacucucaca a	21
<210> 13	
<211> 45	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 13	
ttgtgagagt tgaagttgta tggcagatgt agtgtttcca caggg	45
<210> 14	
<211> 45	
<212> DNA	
[0010] <213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 14	
ccctgtggaa acactacatc tgccatacaa cttcaactct caca	45
<210> 15	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 15	
tatgttgaag ttgagagtgt t	21
<210> 16	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成多核苷酸	
<400> 16	
aacactctca acttcaacat a	21

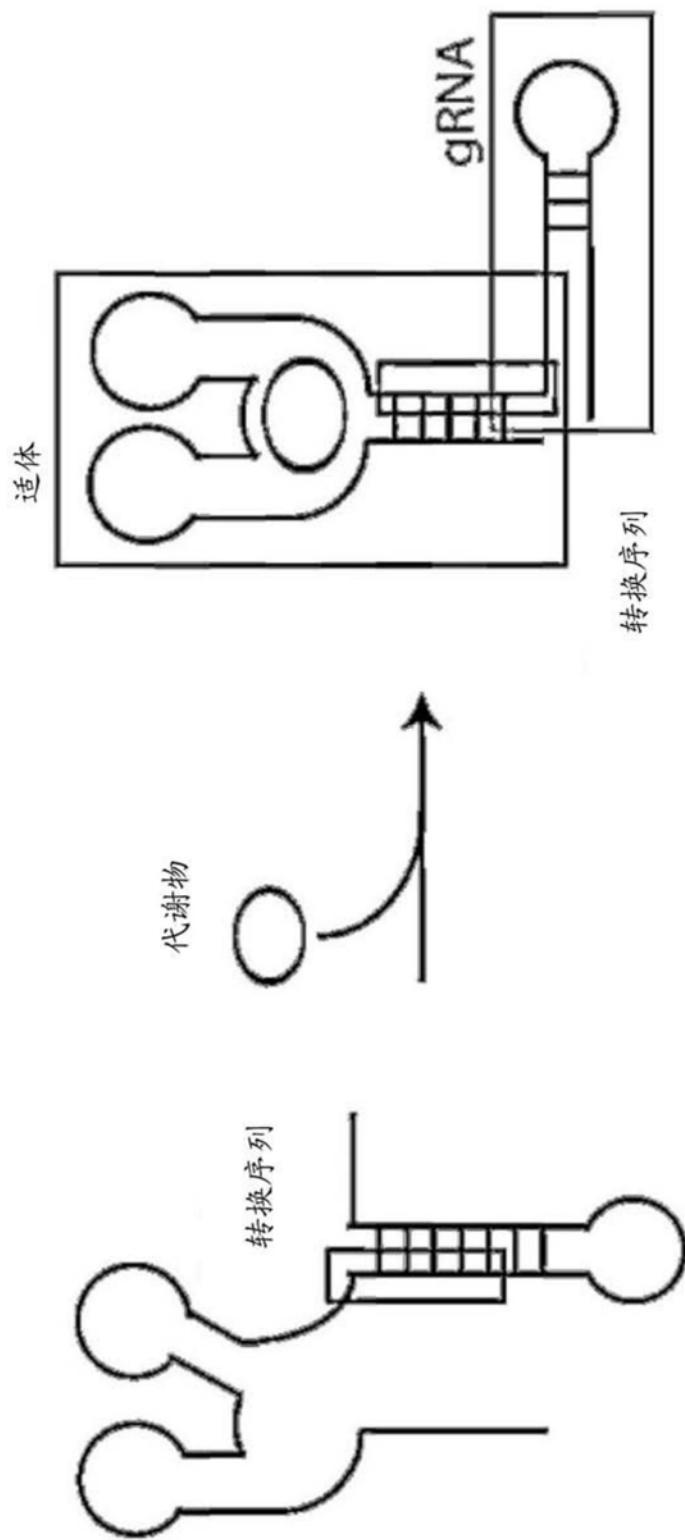


图1A

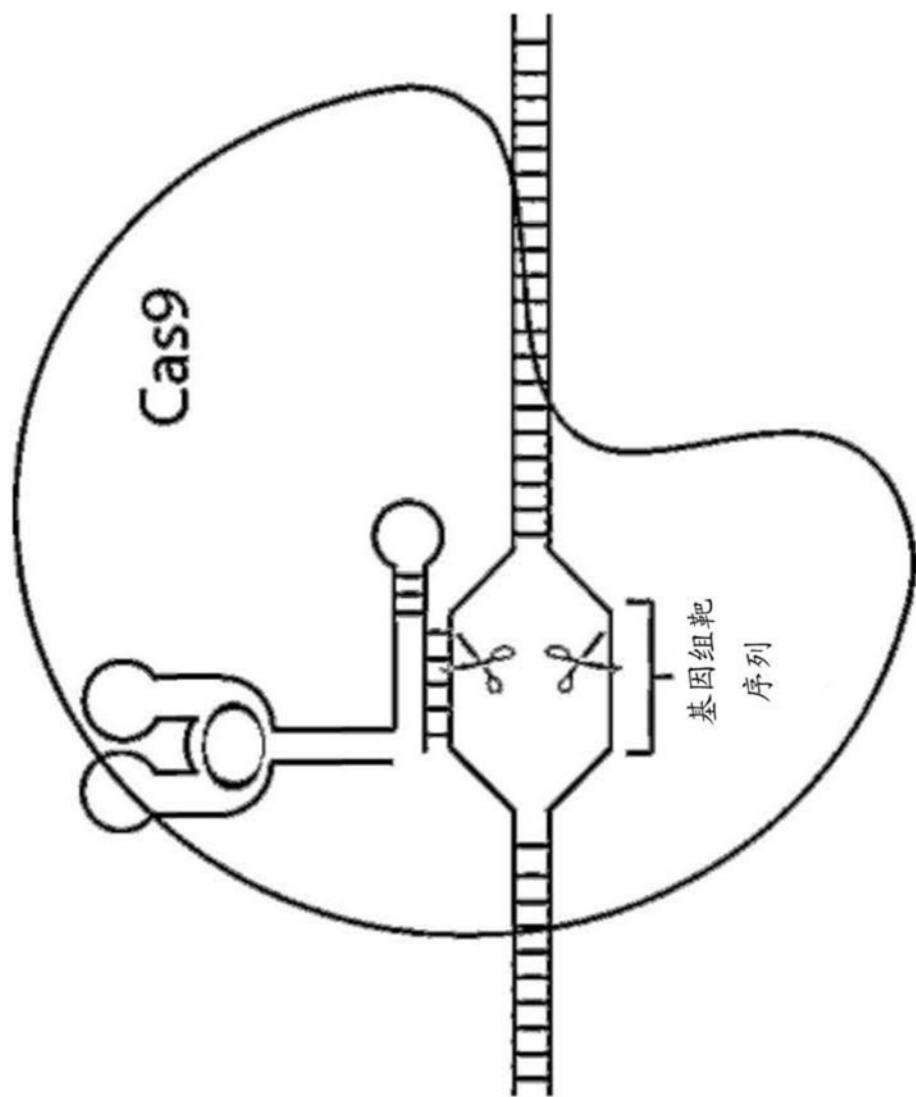


图1B

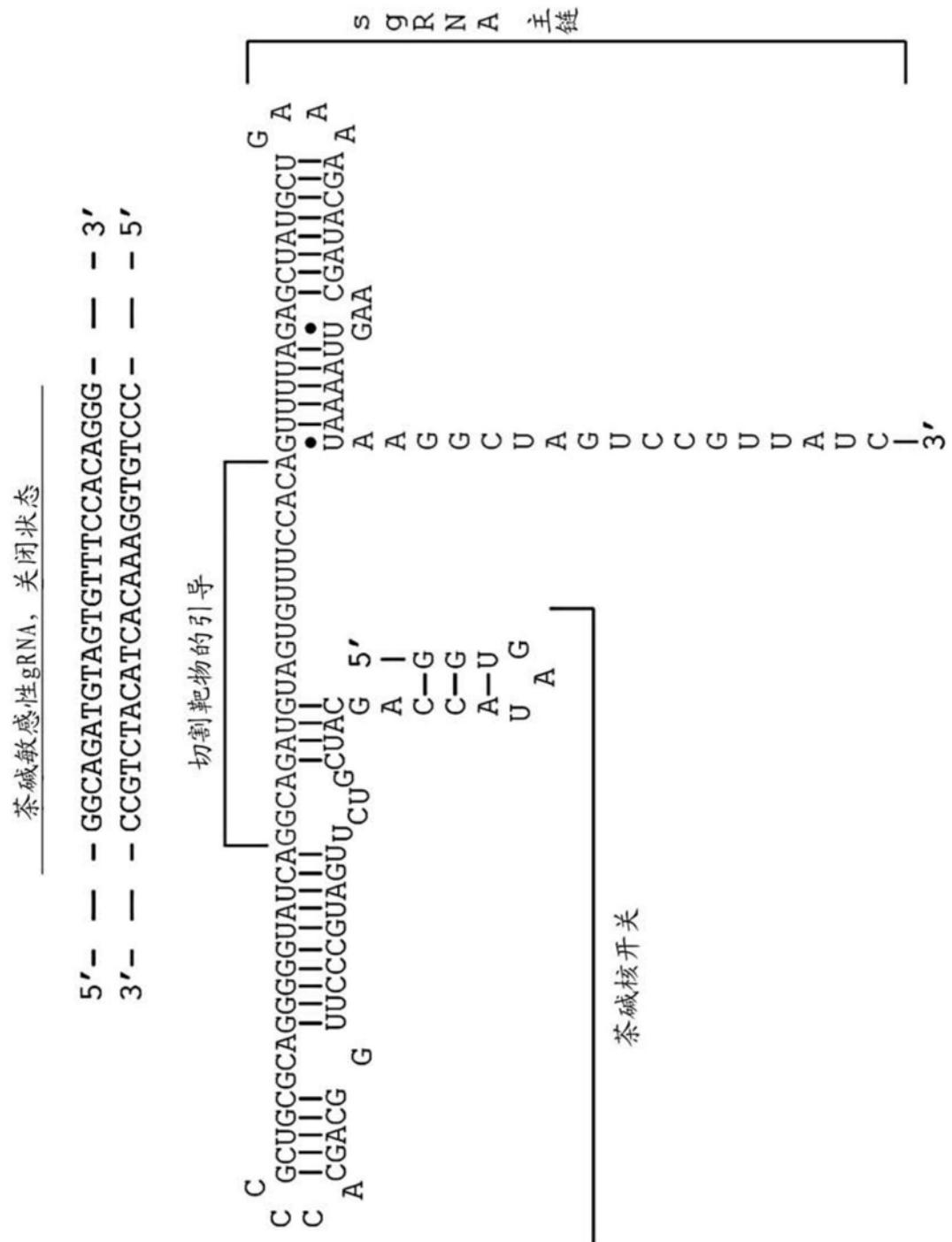


图1C

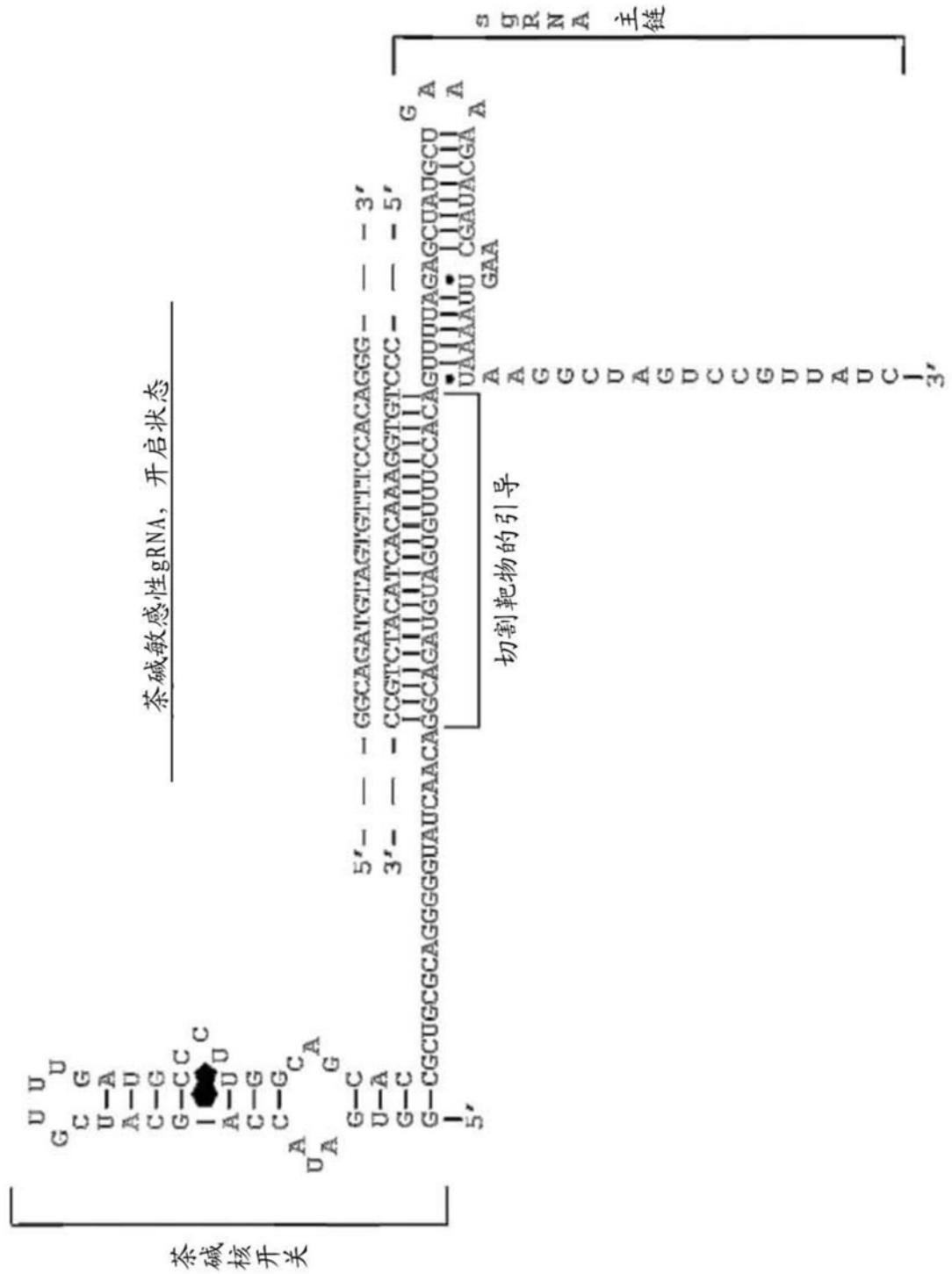


图1D

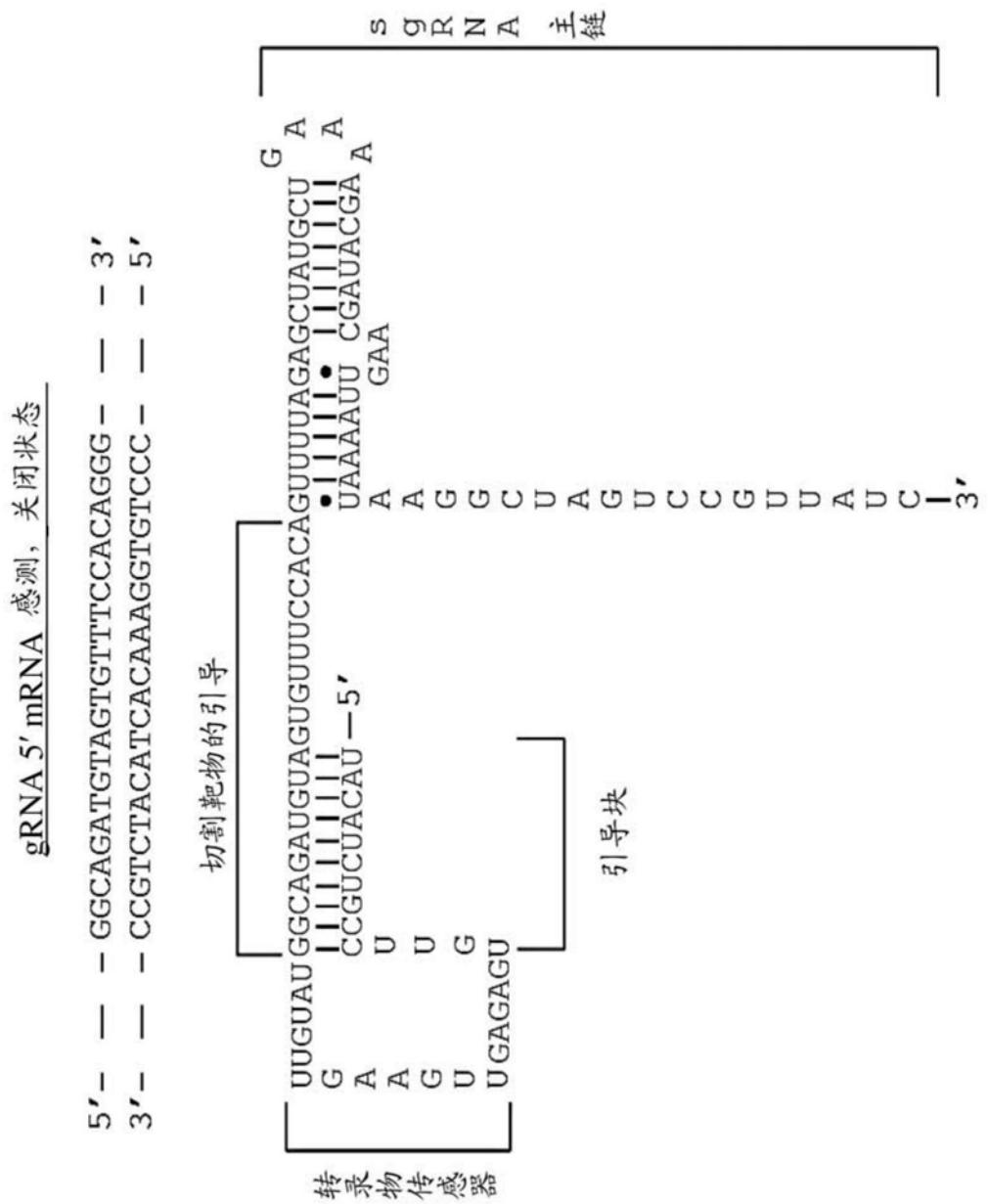


图2A

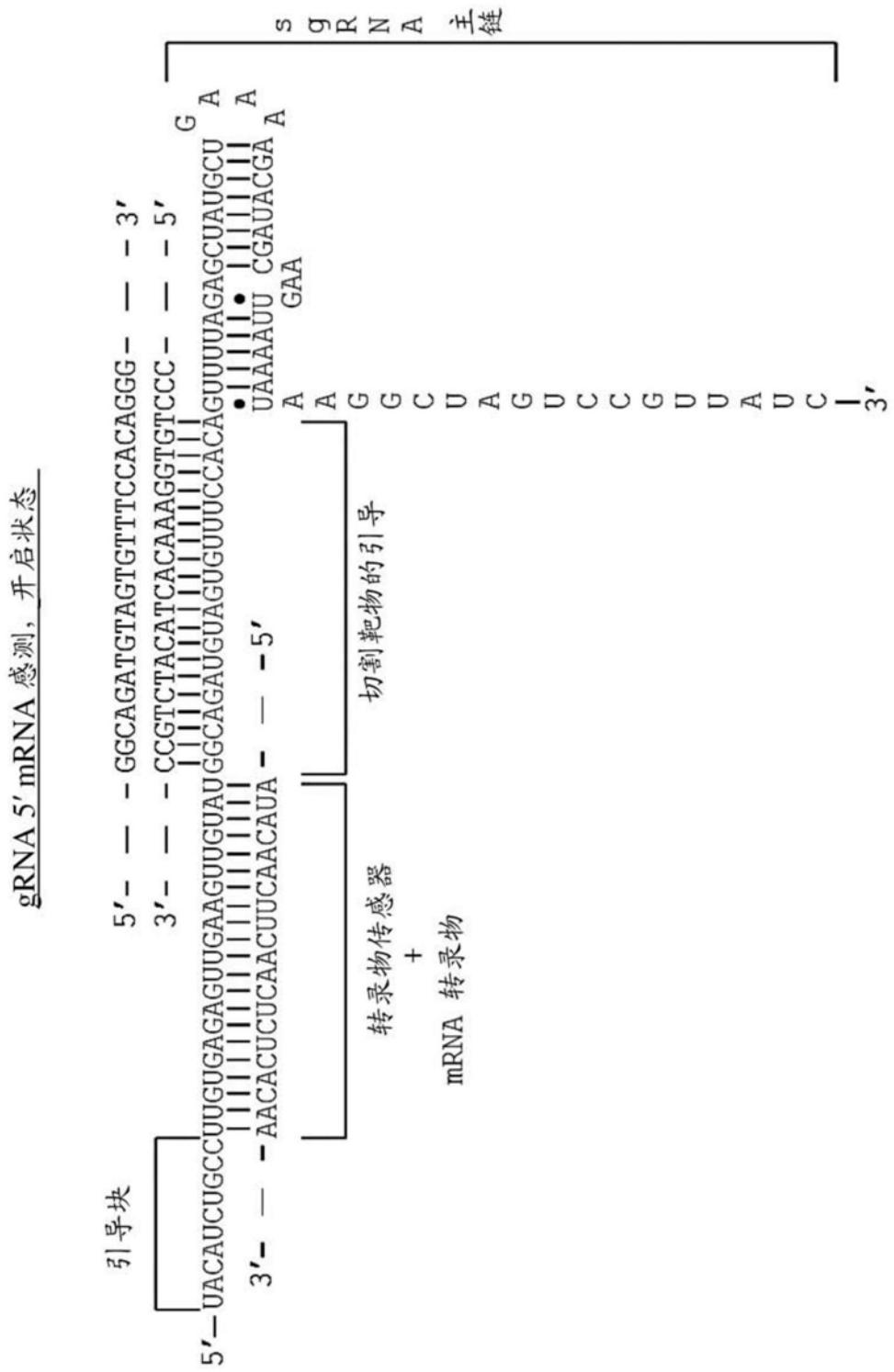
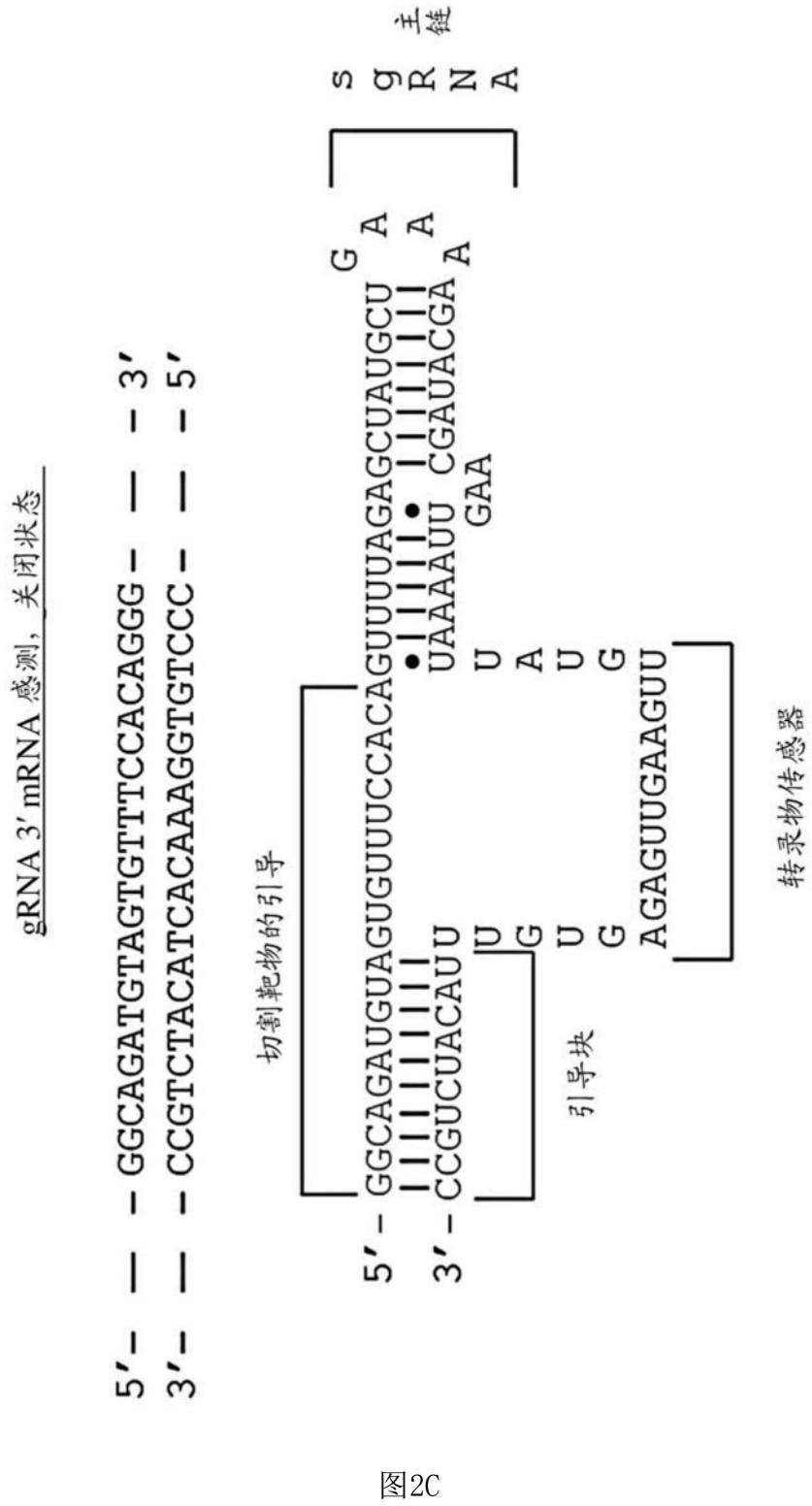


图 2B



gRNA 3' mRNA 感测，开启状态

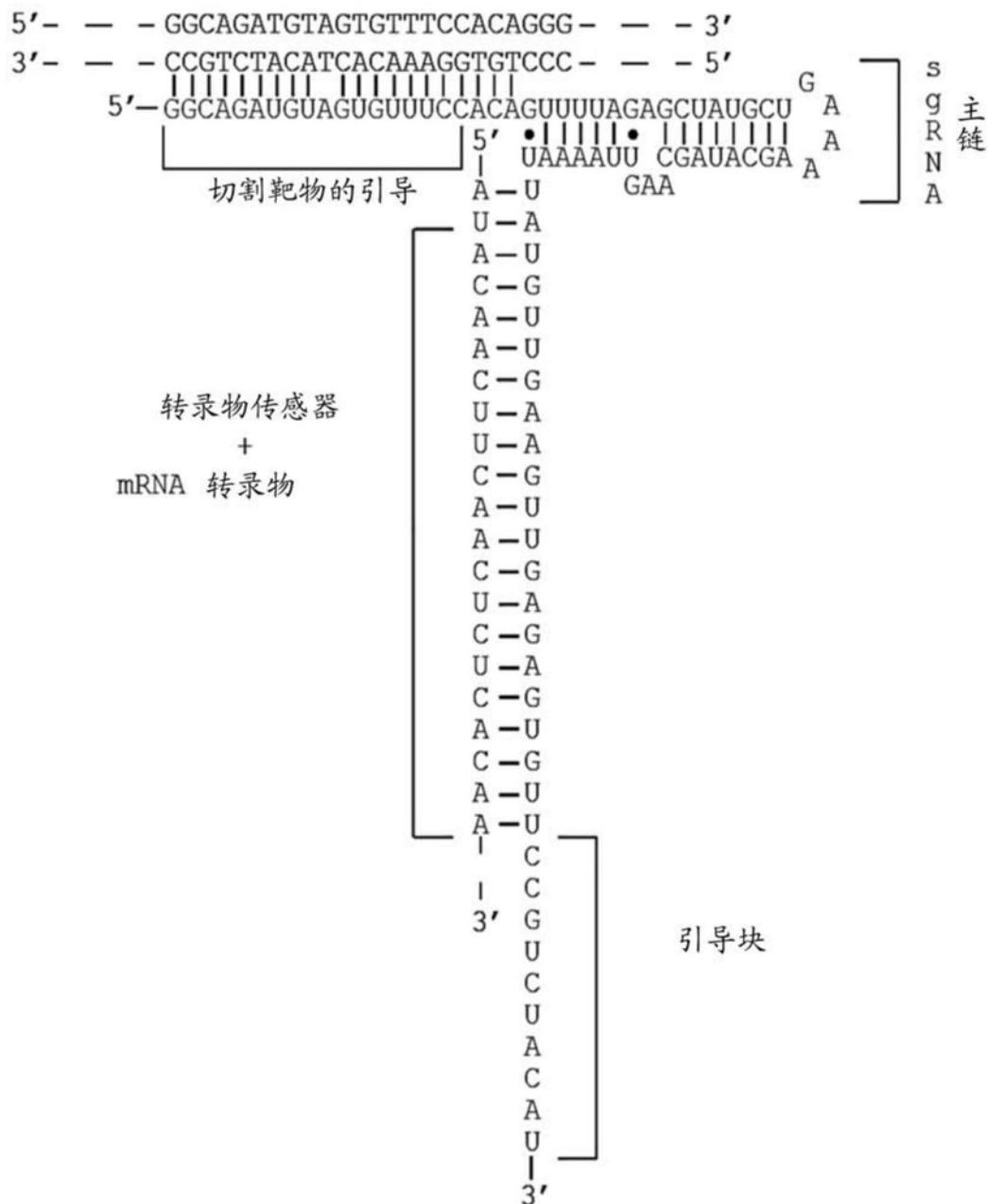


图 2D

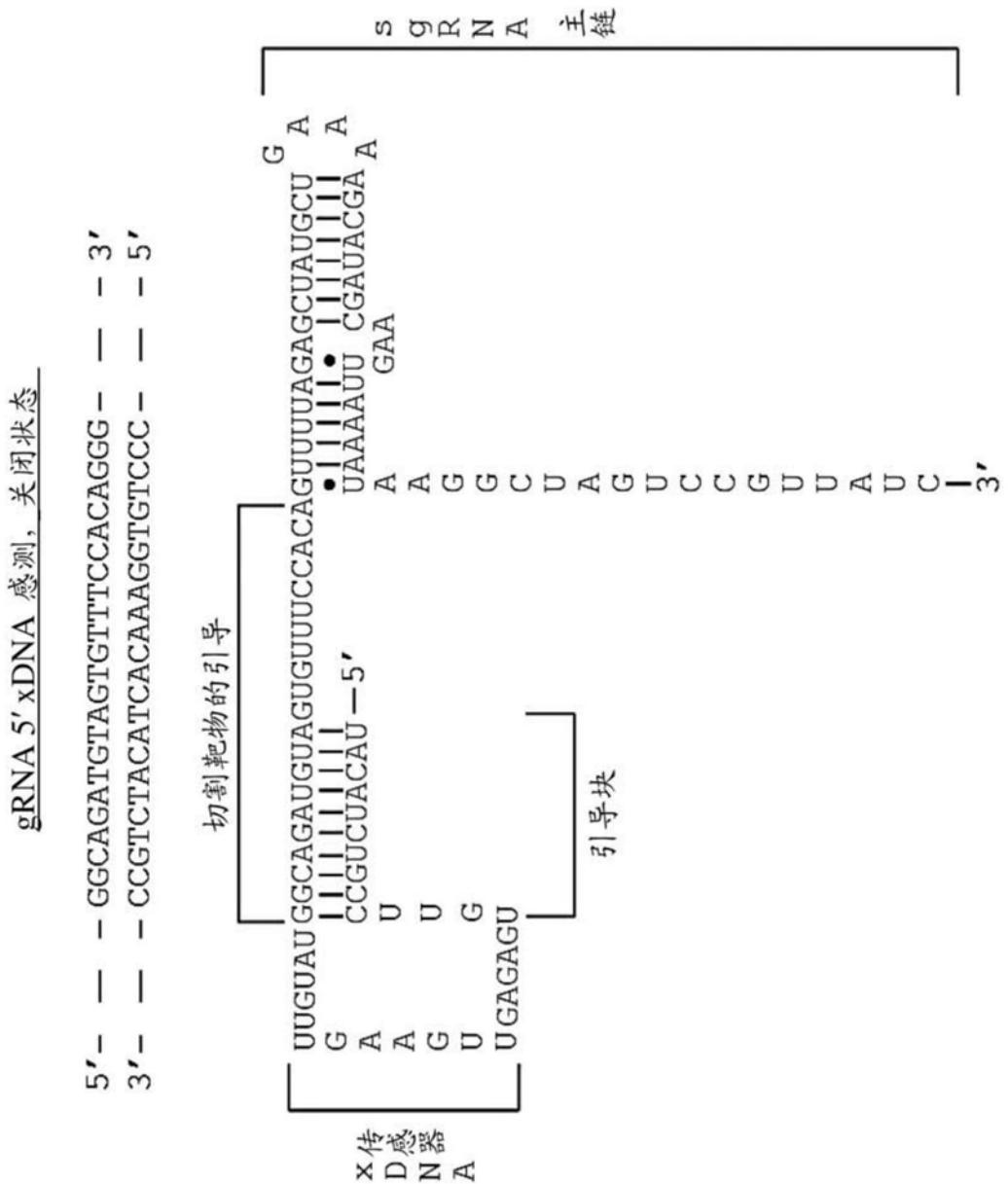


图3A

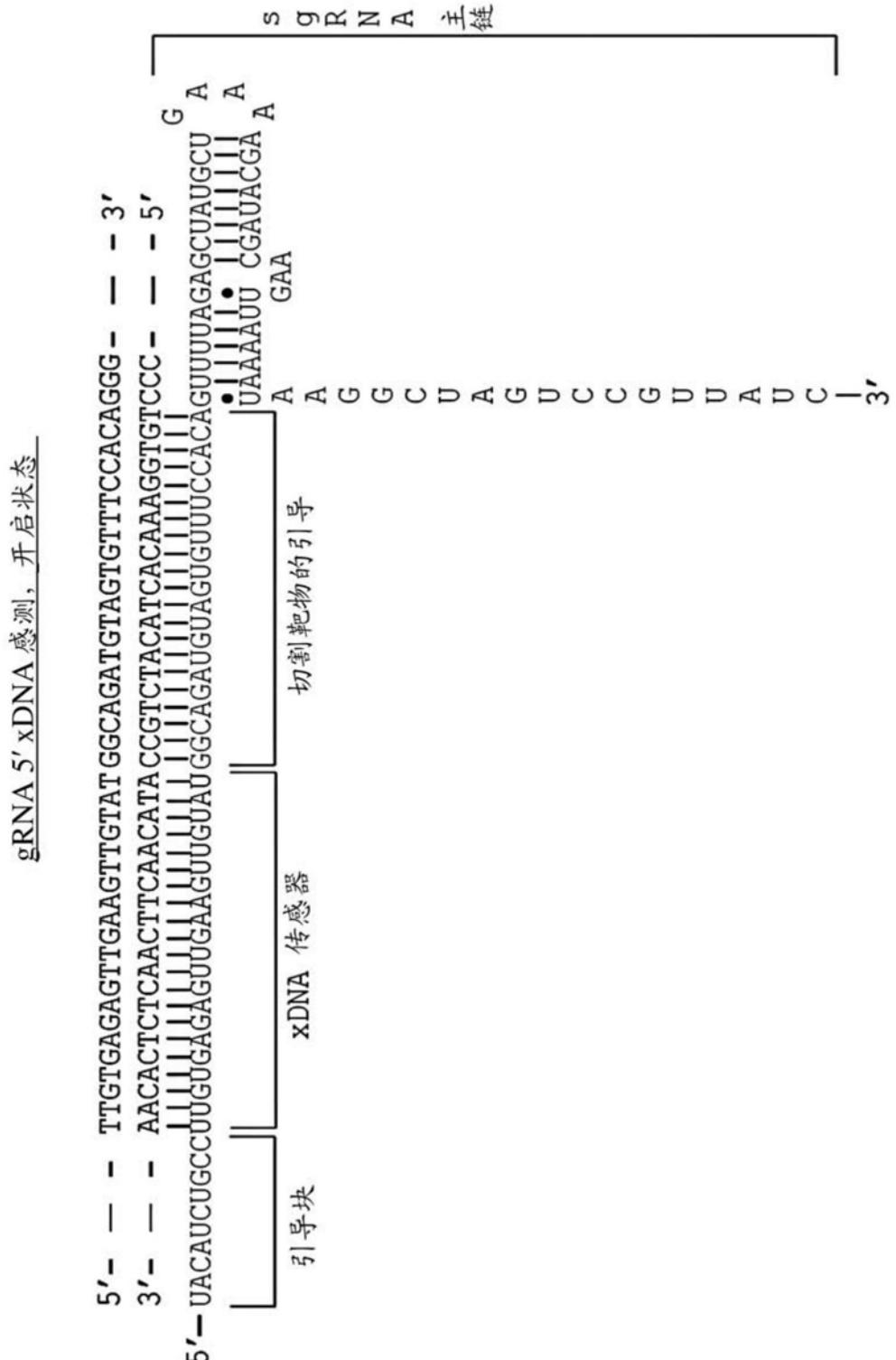


图3B

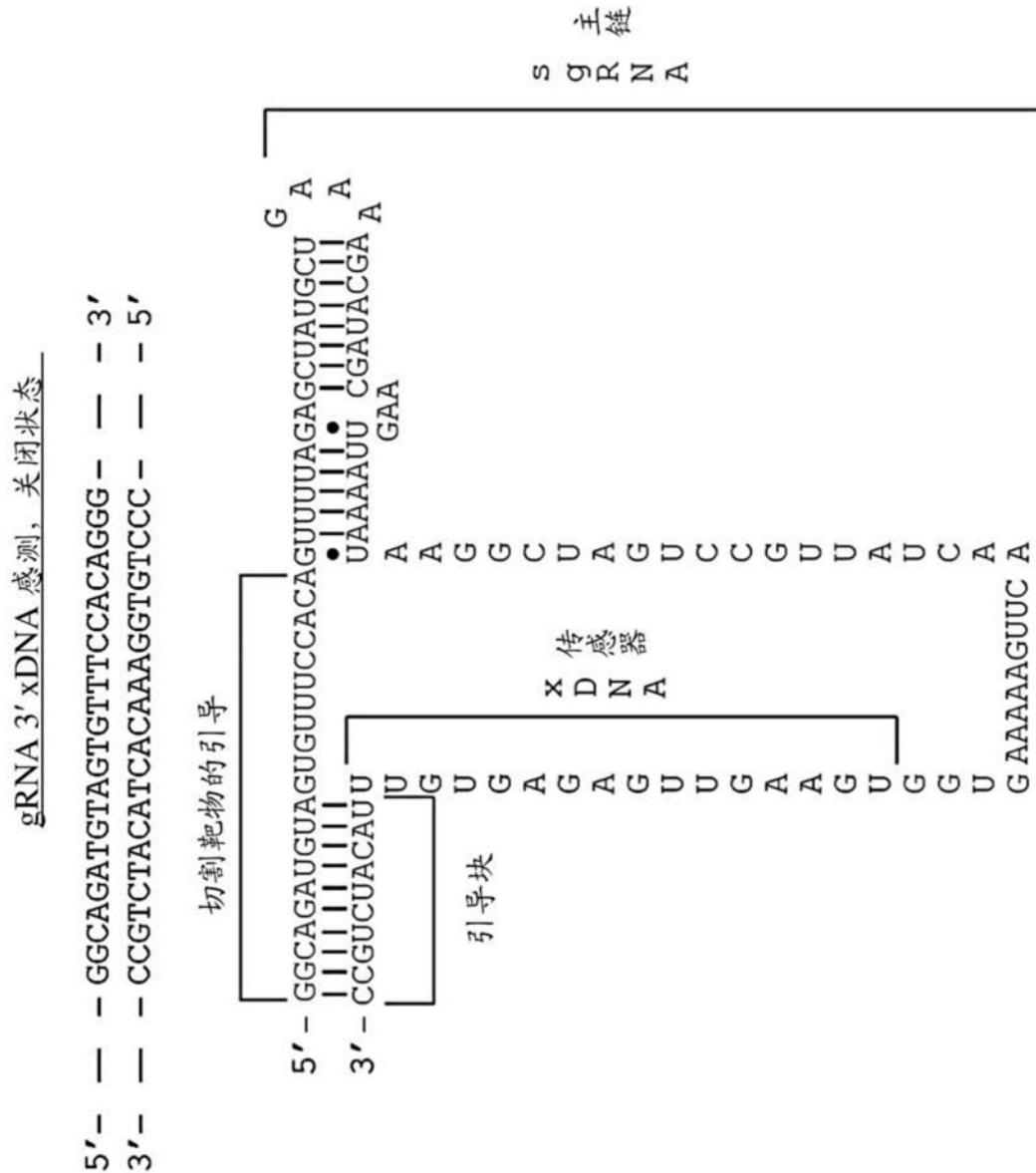


图3C

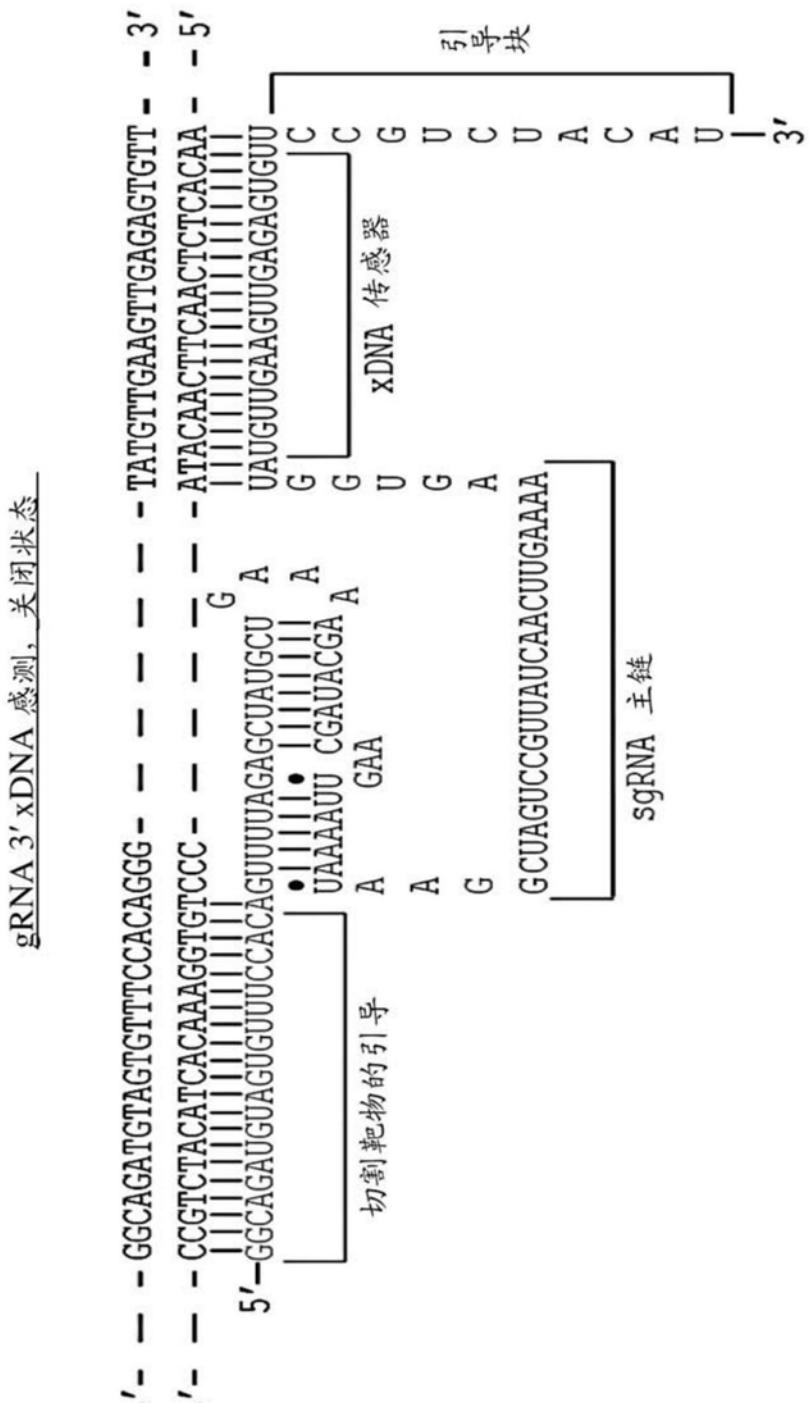


图3D