



(12) 发明专利申请

(10) 申请公布号 CN 113286619 A

(43) 申请公布日 2021.08.20

(21) 申请号 201980051039.7	(74) 专利代理机构 北京坤瑞律师事务所 11494
(22) 申请日 2019.06.07	代理人 封新琴
(30) 优先权数据 62/682,276 2018.06.08 US	(51) Int.Cl. A61K 48/00 (2006.01) C12N 9/22 (2006.01) C12N 15/62 (2006.01)
(85) PCT国际申请进入国家阶段日 2021.01.29	
(86) PCT国际申请的申请数据 PCT/US2019/036050 2019.06.07	
(87) PCT国际申请的公布数据 W02019/236998 EN 2019.12.12	
(71) 申请人 洛卡纳生物股份有限公司 地址 美国加利福尼亚州	
(72) 发明人 D·A·内尔斯 R·巴特拉 E·杨	

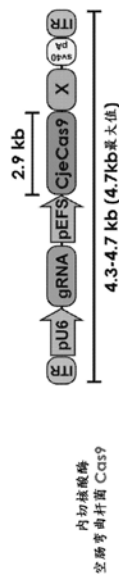
权利要求书3页 说明书156页 附图6页

(54) 发明名称

用于调节适应性免疫的组合物和方法

(57) 摘要

公开了用于同时提供基因疗法并且防止受试者的免疫系统对通过所述基因疗法修饰的细胞的适应性免疫应答的组合物和方法。在一些实施方案中,本公开文本的组合物修饰与疾病或障碍相关的RNA分子的表达水平,以及抑制适应性免疫应答的组分的表达或活性,以从受试者的免疫系统中掩蔽所述修饰的细胞。



1. 一种组合物,其包含含有特异性结合靶RNA序列的指导RNA (gRNA) 序列的核酸序列,其中所述靶RNA序列编码适应性免疫应答的蛋白质组分,并且其中所述gRNA序列包含含有编码所述蛋白质组分的核酸序列的一部分的间隔子序列,并且其中所述蛋白质组分选自 β -2-微球蛋白 (β 2M)、人白细胞抗原A (HLA-A)、人白细胞抗原B (HLA-B)、人白细胞抗原C (HLA-C)、分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS)、ICOS配体 (ICOSLG)、OX40L、白细胞介素12 (IL12) 和CC趋化因子受体7 (CCR7)。

2. 根据权利要求1所述的组合物,其中所述适应性免疫应答选自I型主要组织相容性复合物 (MHC I)、II型主要组织相容性复合物 (MHC II)、T细胞受体 (TCR)、共刺激分子及其组合。

3. 根据权利要求1所述的组合物,其中所述间隔子序列的长度是约20或21个核苷酸。

4. 根据权利要求1所述的组合物,其中所述间隔子序列和所述靶RNA序列是彼此的反向互补体。

5. 根据权利要求1所述的组合物,其中所述gRNA序列包含与CRISPR/Cas多肽或其部分特异性结合的支架序列。

6. 根据权利要求5所述的组合物,其中所述CRISPR/Cas多肽或其部分选自Cas9、Cpf1、Cas13a、Cas13b、Cas13c和CasRX/Cas13d,其中所述CRISPR/Cas多肽具有天然的、降低的或无效的活性。

7. 根据权利要求1所述的组合物,其中所述核酸序列包含驱动所述gRNA序列的表达的启动子。

8. 根据权利要求7所述的组合物,其中所述启动子选自聚合酶III启动子和tRNA启动子。

9. 根据权利要求8所述的组合物,其中所述聚合酶III启动子是U6启动子。

10. 根据权利要求1所述的组合物,其中所述间隔子序列是特异性结合第一靶RNA序列的第一间隔子序列,并且其中所述组合物还包含特异性结合第二靶RNA序列的第二间隔子序列,其中所述第一间隔子序列与第二间隔子序列结合不同的靶RNA序列。

11. 根据权利要求10所述的组合物,其中所述gRNA序列是第一gRNA序列,并且其中所述第二间隔子序列包含在第二gRNA序列内。

12. 根据权利要求10所述的组合物,其中所述第二靶RNA序列编码适应性免疫应答的蛋白质组分。

13. 根据权利要求10所述的组合物,其中所述第二间隔子序列包含编码选自以下的蛋白质组分的核酸序列的一部分: β -2-微球蛋白 (β 2M)、人白细胞抗原A (HLA-A)、人白细胞抗原B (HLA-B)、人白细胞抗原C (HLA-C)、分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS)、ICOS配体 (ICOSLG)、OX40L、白细胞介素12 (IL12) 和CC趋化因子受体7 (CCR7)。

14. 根据权利要求10所述的组合物,其中所述第二间隔子序列包含选自以下的核酸序列的至少1、2、3、4、5、6或7个重复:CUG (SEQ ID NO:18)、CCUG (SEQ ID NO:19)、CAG (SEQ ID NO:80)、GGGGCC (SEQ ID NO:81) 及其组合。

15. 一种组合物,其包含含有以下的核酸序列:(a) 特异性结合第一靶RNA序列的第一指导RNA (gRNA) 序列;以及(b) 特异性结合第二靶RNA序列的第二gRNA,其中所述第一靶RNA序

列编码适应性免疫应答的蛋白质组分,并且其中所述第一gRNA序列包含含有编码所述蛋白质组分的核酸序列的一部分的间隔子序列,并且其中所述蛋白质组分选自 β -2-微球蛋白(β 2M)、人白细胞抗原A(HLA-A)、人白细胞抗原B(HLA-B)、人白细胞抗原C(HLA-C)、分化群28(CD28)、分化群80(CD80)、分化群86(CD86)、可诱导T细胞共刺激分子(ICOS)、ICOS配体(ICOSLG)、OX40L、白细胞介素12(IL12)和CC趋化因子受体7(CCR7)。

16. 一种组合物,其包含含有以下的核酸序列:(a) 根据权利要求1所述的指导RNA(gRNA)序列;以及(b) 编码融合蛋白的核酸序列,其中所述融合蛋白包含第一RNA结合多肽和第二RNA结合多肽,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

17. 一种组合物,其包含含有以下的核酸序列:(a) 根据权利要求11所述的第一和第二指导RNA(gRNA)序列;以及(b) 编码融合蛋白的核酸序列,其中所述融合蛋白包含第一RNA结合多肽和第二RNA结合多肽,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

18. 一种组合物,其包含含有以下的核酸序列:(a) 特异性结合第一RNA分子内的第一靶RNA序列的第一指导RNA(gRNA),其中所述第一靶RNA序列编码适应性免疫应答的蛋白质组分;(b) 特异性结合第二RNA分子内的第二靶RNA序列的第二指导RNA(gRNA);以及(c) 编码融合蛋白的核酸序列,其中所述融合蛋白包含第一RNA结合多肽、第二RNA结合多肽,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

19. 根据权利要求18所述的组合物,其中所述第一gRNA序列包含含有编码选自以下的蛋白质的核酸序列的一部分的间隔子序列: β -2-微球蛋白(β 2M)、HLA-A、HLA-B、HLA-C、CD28、CD80、CD86、ICOSLG、OX40L、IL12和CCR7。

20. 根据权利要求18所述的组合物,其中所述第一RNA结合多肽或其部分是CRISPR/Cas多肽或其部分。

21. 根据权利要求20所述的组合物,其中所述CRISPR/Cas多肽或其部分选自Cas9、Cpf1、Cas13a、Cas13b、Cas13c和CasRX/Cas13d,其中所述CRISPR/Cas多肽具有天然的、降低的或无效的活性。

22. 根据权利要求18所述的组合物,其中所述第二RNA结合多肽以与RNA缔合的方式结合RNA。

23. 根据权利要求22所述的组合物,其中所述第二RNA结合多肽以切割RNA的方式与RNA缔合。

24. 根据权利要求18所述的组合物,其中所述核酸序列包含启动子。

25. 根据权利要求18所述的组合物,其中所述第二gRNA包含含有选自以下的序列的至少1、2、3、4、5、6或7个重复的间隔子序列:CUG(SEQ ID NO:18)、CCUG(SEQ ID NO:19)、CAG(SEQ ID NO:80)、GGGGCC(SEQ ID NO:81)及其组合。

26. 根据权利要求18所述的组合物,其中所述融合蛋白包含NLS、NES或标签。

27. 一种载体,其包含根据权利要求18所述的组合物。

28. 根据权利要求27所述的载体,其中所述载体选自:腺相关病毒、逆转录病毒、慢病毒、腺病毒、纳米颗粒、胶束、脂质体、阳离子脂质体/DNA复合物、聚合物囊泡、聚合物/DNA复合物和树枝状聚合物。

29. 一种细胞,其包含根据权利要求28所述的载体。

30. 根据权利要求18所述的组合物,其中所述第二RNA结合多肽选自:RNA酶1、RNA酶4、RNA酶6、RNA酶7、RNA酶8、RNA酶2、RNA酶6PL、RNA酶L、RNA酶T2、RNA酶11、RNA酶T2样蛋白、NOB1、ENDO V、ENDO G、ENDOD1、hFEN1、hSLFN14、hLACTB2、APEX2、ANG、HRSP12、ZC3H12A、RIDA、PDL6、NTHL、KIAA0391、APEX1、AGO2、EXO G、ZC3H12D、ERN2、PELO、YBEY、CPSF4L、hCG_2002731、ERCC1、RAC1、RAA1、RAB1、DNA2、FLJ35220、FLJ13173、ERCC4、RNA酶1 (K41R)、RNA酶1 (K41R、D121E)、RNA酶1 (K41R、D121E、H119N)、RNA酶1 (H119N)、RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)、RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E)、RNA酶1 (R39D、N67D、N88A、G89D、R91D)、TENM1、TENM2、RNA酶K、TALEN、ZNF638和hSMG6 PIN。

用于调节适应性免疫的组合物和方法

技术领域

[0001] 本公开文本涉及分子生物学,并且更具体而言涉及用于修饰涉及适应性免疫应答的RNA分子的表达和活性的组合物和方法。

相关申请的交叉引用

[0002] 本申请要求2018年6月8日提交的美国专利申请号62/682,276的优先权,将其内容通过引用以其整体并入本文。将2019年6月7日提交的国际申请号PCT/US2019/036021、2019年6月7日提交的美国专利申请号16/434,689和2018年6月8日提交的美国专利申请号62/682,271的内容通过引用以其整体并入本文。

序列表的并入

[0003] 将2019年6月6日创建的大小为2.93MB的名为“LOCN_003_001W0_SeqList_ST25”的文本文件的内容通过引用以其整体特此并入。

背景技术

[0004] 对于同时提供基因疗法并且抑制当例如通过病毒载体递送所述基因疗法时可能出现的适应性免疫应答,在本领域中存在长期但未得到满足的需要。本公开文本提供了用于以序列特异性方式特异性靶向RNA分子的组合物和方法,其在体内提供基因疗法,同时从受试者的免疫系统中掩蔽修饰的细胞,从而防止对所述修饰的细胞的适应性免疫应答。

发明内容

[0005] 本公开文本提供了一种组合物,其包含含有特异性结合靶RNA序列的指导RNA (gRNA) 序列的核酸序列,其中所述靶RNA序列编码适应性免疫应答的蛋白质组分,并且其中所述gRNA序列包含含有编码所述蛋白质组分的核酸序列的一部分的间隔子序列,并且其中所述蛋白质组分选自 β -2-微球蛋白(β 2M)、人白细胞抗原A (HLA-A)、人白细胞抗原B (HLA-B)、人白细胞抗原C (HLA-C)、分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子(ICOS)、ICOS配体(ICOSLG)、OX40L、白细胞介素12 (IL12) 和CC趋化因子受体7 (CCR7)。

[0006] 本公开文本还提供了一种组合物,其包含(a)包含特异性结合RNA分子内的靶序列的指导RNA (gRNA) 的第一序列,其中所述靶序列包含编码适应性免疫应答的组分的序列;以及(b)编码融合蛋白的序列,所述序列包含编码第一RNA结合多肽的序列和编码第二RNA结合多肽的序列,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0007] 本公开文本提供了一种组合物,其包含:(a)包含特异性结合第一RNA分子内的第一靶序列的指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;和(b)包含特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列;以及(c)编码融合蛋白的序列,所述序列包含编码第一RNA结合多肽的序列和编

码第二RNA结合多肽的序列,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0008] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一靶序列或所述第二靶序列包含至少一个重复的序列。

[0009] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一序列包含能够在真核细胞中表达所述gRNA的第一启动子和/或所述第二序列包含能够在真核细胞中表达所述gRNA的第二启动子。在一些实施方案中,所述第一启动子和所述第二启动子是相同的。在一些实施方案中,所述第一启动子和所述第二启动子是不同的。

[0010] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一序列和所述第二序列包含能够在真核细胞中表达所述第一gRNA和所述第二gRNA的启动子。

[0011] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中gRNA序列包含能够在真核细胞中表达所述gRNA的启动子,所述真核细胞是动物细胞。在一些实施方案中,所述动物细胞是哺乳动物细胞。在一些实施方案中,所述动物细胞是人细胞。

[0012] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中gRNA序列包含能够在真核细胞中表达所述gRNA的启动子,所述启动子是组成型活性启动子。

[0013] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中gRNA序列包含能够在真核细胞中表达所述gRNA的启动子,所述gRNA序列包含从能够驱动RNA聚合酶的表达的启动子分离或衍生的序列。在一些实施方案中,所述启动子序列是从U6启动子分离或衍生的。

[0014] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中gRNA序列包含能够在真核细胞中表达所述gRNA的启动子,所述启动子包含从能够驱动转移RNA (tRNA) 的表达的启动子分离或衍生的序列。在一些实施方案中,所述启动子序列是从以下启动子分离或衍生的:丙氨酸tRNA启动子、精氨酸tRNA启动子、天冬酰胺tRNA启动子、天冬氨酸tRNA启动子、半胱氨酸tRNA启动子、谷氨酰胺tRNA启动子、谷氨酸tRNA启动子、甘氨酸tRNA启动子、组氨酸tRNA启动子、异亮氨酸tRNA启动子、亮氨酸tRNA启动子、赖氨酸tRNA启动子、甲硫氨酸tRNA启动子、苯丙氨酸tRNA启动子、脯氨酸tRNA启动子、丝氨酸tRNA启动子、苏氨酸tRNA启动子、色氨酸tRNA启动子、酪氨酸tRNA启动子或缬氨酸tRNA启动子。在一些实施方案中,所述启动子序列是从缬氨酸tRNA启动子分离或衍生的。

[0015] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组

合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 包含所述第一gRNA的序列还包含与所述第一靶RNA序列特异性结合的第一间隔子序列。在一些实施方案中, 所述第一间隔子序列与所述靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。在一些实施方案中, 所述第一间隔子序列与所述靶RNA序列具有100%互补性。在一些实施方案中, 所述第一间隔子序列包含20个核苷酸或由其组成。在一些实施方案中, 所述第一间隔子序列包含21个核苷酸或由其组成。在一些实施方案中, 所述第一间隔子序列包含编码 β -2-微球蛋白 (β 2M) 蛋白的氨基酸序列的20个核苷酸或由其组成。在一些实施方案中, 所述第一间隔子序列包含以下的氨基酸序列的20个核苷酸或由其组成:

MSRSVALAVL ALLSLSGLEA IQRTPKIQVY SRHPADIEVD LLKNGERIEK VEHSDFSFSK
DWSFYLLYYT EFTPTEKDEY ACRVNHVTL S QPKIVKWDRD M (SEQ ID NO:88)。

[0016] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 包含所述第一gRNA的序列还包含与所述第一RNA结合蛋白特异性结合的第一支架序列。在一些实施方案中, 所述第一支架序列包含茎环结构。在一些实施方案中, 所述支架序列包含90个核苷酸或由其组成。在一些实施方案中, 所述支架序列包含93个核苷酸或由其组成。在一些实施方案中, 所述支架序列包含以下序列: GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:12) 或GUUUUAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:13)。

[0017] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 包含所述第二gRNA的序列还包含与所述第二靶RNA序列特异性结合的第二间隔子序列。在一些实施方案中, 所述第二间隔子序列与所述第一靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。在一些实施方案中, 所述第二间隔子序列与所述靶RNA序列具有100%互补性。在一些实施方案中, 所述第二间隔子序列包含20个核苷酸或由其组成。在一些实施方案中, 所述第二间隔子序列包含21个核苷酸或由其组成。在一些实施方案中, 所述第二间隔子序列包含或还包含含有以下序列的至少1、2、3、4、5、6或7个重复的序列: CUG (SEQ ID NO:18)、CCUG (SEQ ID NO:19)、CAG (SEQ ID NO:80)、GGGGCC (SEQ ID NO:81) 或其任何组合。

[0018] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 包含所述第二gRNA的序列还包

含与所述第一RNA结合蛋白特异性结合的第二支架序列。在一些实施方案中,所述第二支架序列包含茎环结构。在一些实施方案中,所述支架序列包含85个核苷酸或由其组成。在一些实施方案中,所述支架序列包含以下序列:GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAU AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:12) 或GUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:13)。

[0019] 在本公开文本的组合物的一些实施方案中,所述gRNA不结合或不选择性结合所述RNA分子内的第二序列。

[0020] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一gRNA不结合或不选择性结合所述第一RNA分子内的第二序列。

[0021] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二gRNA不结合或不选择性结合所述第二RNA分子内的第二序列。

[0022] 在本公开文本的组合物的一些实施方案中,RNA基因组或RNA转录组包含所述RNA分子。

[0023] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,RNA基因组或RNA转录组包含所述第一RNA分子或所述第二RNA分子。

[0024] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一RNA结合蛋白包含CRISPR-Cas蛋白。在一些实施方案中,所述CRISPR-Cas蛋白是II型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cas9多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白是V型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cpf1多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白是VI型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cas13多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白包含天然RNA核酸酶活性。在一些实施方案中,所述天然RNA核酸酶活性被降低或抑制。在一些实施方案中,所述天然RNA核酸酶活性被增加或诱导。在一些实施方案中,所述CRISPR-Cas蛋白包含天然DNA核酸酶活性并且其中所述天然DNA核酸酶活性被抑制。在一些实施方案中,所述CRISPR-Cas蛋白包含突变。在一些实施方案中,所述CRISPR-Cas蛋白的核酸酶结构域包含所述突变。在一些实施方案中,所述突变发生在编

码所述CRISPR-Cas蛋白的核酸中。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的氨基酸中。在一些实施方案中,所述突变包含取代、插入、缺失、移码、倒位或转座。在一些实施方案中,所述突变包含核酸酶结构域、所述核酸酶结构域内的结合位点、所述核酸酶结构域内的活性位点或所述核酸酶结构域内的至少一个必需氨基酸残基的缺失。

[0025] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一RNA结合蛋白包含Pumilio和FBF (PUF) 蛋白或其RNA结合部分。在一些实施方案中,所述第一RNA结合蛋白包含基于Pumilio的联合体(Pumilio-based assembly, PUMBY) 蛋白或其RNA结合部分。

[0026] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一RNA结合蛋白不需要多聚化以用于RNA结合活性。在一些实施方案中,所述第一RNA结合蛋白不是多聚体复合物的单体。在一些实施方案中,多聚体蛋白复合物不包含所述第一RNA结合蛋白。

[0027] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一RNA结合蛋白与所述RNA分子内的靶序列选择性结合。在一些实施方案中,所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的亲和力。在一些实施方案中,所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的高亲和力或不选择性结合所述第二序列。在一些实施方案中,RNA基因组或RNA转录组包含所述RNA分子。

[0028] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第一RNA结合蛋白包含在2个与1300个之间的氨基酸,包括端点。

[0029] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,编码所述第一RNA结合蛋白的序列还包含编码核定位信号(NLS) 的序列。在一些实施方案中,编码核定位信号(NLS) 的所述序列定位于编码所述第一RNA结合蛋白的序列的3'。在一些实施方案中,所述第一RNA结合蛋白包含在所述蛋白质的C末端的NLS。

[0030] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二

RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 编码所述第一RNA结合蛋白的序列还包含编码第一NLS的第一序列和编码第二NLS的第二序列。在一些实施方案中, 编码所述第一NLS或所述第二NLS的序列定位于编码所述第一RNA结合蛋白的序列的3'。在一些实施方案中, 所述第一RNA结合蛋白包含在所述蛋白质的C末端的所述第一NLS或所述第二NLS。

[0031] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含核酸酶结构域或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶1或由其组成。在一些实施方案中, 所述RNA酶1蛋白包含SEQ ID NO:20或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶4或由其组成。在一些实施方案中, 所述RNA酶4蛋白包含SEQ ID NO:21或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶6或由其组成。在一些实施方案中, 所述RNA酶6蛋白包含SEQ ID NO:22或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶7或由其组成。在一些实施方案中, 所述RNA酶7蛋白包含SEQ ID NO:23或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶8或由其组成。在一些实施方案中, 所述RNA酶8蛋白包含SEQ ID NO:24或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶2或由其组成。在一些实施方案中, 所述RNA酶2包含SEQ ID NO:25或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶6PL或由其组成。在一些实施方案中, 所述RNA酶6PL蛋白包含SEQ ID NO:26或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶L或由其组成。在一些实施方案中, 所述RNA酶L蛋白包含SEQ ID NO:27或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶T2或由其组成。在一些实施方案中, 所述RNA酶T2蛋白包含SEQ ID NO:28或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶11或由其组成。在一些实施方案中, 所述RNA酶11蛋白包含SEQ ID NO:29或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含RNA酶T2样蛋白或由其组成。在一些实施方案中, 所述RNA酶T2样蛋白包含SEQ ID NO:30或由其组成。

[0032] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含NOB1多肽或由其组成。在一些实施方案中, 所述NOB1多肽包含SEQ ID NO:31或由其组成。

[0033] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含内切核酸酶或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含内切核酸酶V (ENDO V) 或由其组成。在一些实施方案中, 所述ENDO V包含SEQ ID NO:32或由其组成。在一些实施方案中, 所述第二RNA结合蛋白包含内切核酸酶G (ENDO G) 或由其组成。在一些实施方案中, 所

述ENDOG包含SEQ ID NO:33或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含内切核酸酶D1 (ENDOD1) 或由其组成。在一些实施方案中,所述ENDOD1包含SEQ ID NO:34或由其组成。

[0034] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含人瓣状内切核酸酶-1 (hFEN1) 或由其组成。在一些实施方案中,所述hFEN1包含SEQ ID NO:35或由其组成。

[0035] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含人斯库拉芬蛋白 (Schlafen) 14 (hSLFN14) 多肽或由其组成。在一些实施方案中,所述hSLFN14包含SEQ ID NO:36或由其组成。

[0036] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含人 β -内酰胺酶样蛋白2 (hLACTB2) 多肽或由其组成。在一些实施方案中,所述hLACTB2包含SEQ ID NO:37或由其组成。

[0037] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含无嘌呤/无嘧啶 (AP) 内切脱氧核糖核酸酶 (APEX2) 多肽或由其组成。在一些实施方案中,所述APEX2包含SEQ ID NO:38或由其组成。在一些实施方案中,所述APEX2包含SEQ ID NO:39或由其组成。

[0038] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含血管生成素 (ANG) 多肽或由其组成。在一些实施方案中,所述ANG包含SEQ ID NO:40或由其组成。

[0039] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含热反应蛋白12 (HRSP12) 多肽或由其组成。在一些实施方案中,所述HRSP12包含SEQ ID NO:41

或由其组成。

[0040] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含含锌指CCCH型12A (Zinc Finger CCCH-Type Containing 12A, ZC3H12A) 或由其组成。在一些实施方案中,所述ZC3H12A包含SEQ ID NO:42或由其组成。在一些实施方案中,所述ZC3H12A包含SEQ ID NO:43或由其组成。

[0041] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含反应性中间亚胺脱氨酶A (Reactive Intermediate Imine Deaminase A, RIDA) 多肽或由其组成。在一些实施方案中,所述RIDA多肽包含SEQ ID NO:44或由其组成。

[0042] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含磷脂酶D家族成员6 (PDL6) 多肽或由其组成。在一些实施方案中,所述PDL6多肽包含SEQ ID NO:126或由其组成。

[0043] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含内切核酸酶III样蛋白1 (NTHL) 多肽或由其组成。在一些实施方案中,所述NTHL多肽包含SEQ ID NO:123或由其组成。

[0044] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含线粒体核糖核酸酶P催化亚基 (KIAA0391) 多肽或由其组成。在一些实施方案中,所述KIAA0391多肽包含SEQ ID NO:127或由其组成。

[0045] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含无嘌呤或无嘧啶位点裂解酶 (APEX1) 多肽或由其组成。在一些实施方案中,所述APEX1多肽包含SEQ ID NO:125或由其组成。

[0046] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组

合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含阿尔古蛋白 (argonaute) 2 (AGO2) 多肽或由其组成。在一些实施方案中, 编码所述AGO2多肽包含SEQ ID NO:128或由其组成。

[0047] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含线粒体核酸酶EXOG (EXOG) 多肽或由其组成。在一些实施方案中, 所述EXOG多肽包含SEQ ID NO:129或由其组成。

[0048] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含含锌指CCCH型12D (ZC3H12D) 多肽或由其组成。在一些实施方案中, 所述ZC3H12D多肽包含SEQ ID NO:130或由其组成。

[0049] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含内质网核信号转导蛋白2 (ERN2) 多肽或由其组成。在一些实施方案中, 所述ERN2多肽包含SEQ ID NO:131或由其组成。

[0050] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含回力球mRNA监督和核糖体挽救因子 (pelota mRNA surveillance and ribosome rescue factor, PELO) 多肽或由其组成。在一些实施方案中, 所述PELO多肽包含SEQ ID NO:132或由其组成。

[0051] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列, 所述第二RNA结合蛋白包含YBEY金属肽酶 (YBEY) 多肽或由其组成。在一些实施方案中, 所述YBEY多肽包含SEQ ID NO:133或由其组成。

[0052] 在本公开文本的组合物的一些实施方案中, 包括以下那些实施方案: 其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列, 其中所述第一靶序列包含编码适应性免疫应答的组分的序列; 以及含有特异性结合第二

RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含剪切和多聚腺苷酸化特异性因子4样蛋白 (CPSF4L) 多肽或由其组成。在一些实施方案中,所述CPSF4L多肽包含SEQ ID NO:134或由其组成。

[0053] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含hCG_2002731多肽或由其组成。在一些实施方案中,所述hCG_2002731多肽包含SEQ ID NO:135或由其组成。在一些实施方案中,编码所述hCG_2002731多肽的序列包含SEQ ID NO:136或由其组成。

[0054] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含切除修复交叉互补组1 (ERCC1) 多肽或由其组成。在一些实施方案中,所述ERCC1多肽包含SEQ ID NO:137或由其组成。

[0055] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含ras相关C3肉毒杆菌毒素底物1亚型 (RAC1) 多肽或由其组成。在一些实施方案中,所述RAC1多肽包含SEQ ID NO:138或由其组成。

[0056] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含核糖核酸酶A A1 (RAA1) 多肽或由其组成。在一些实施方案中,所述RAA1多肽包含SEQ ID NO:139或由其组成。

[0057] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含Ras相关蛋白 (RAB1) 多肽或由其组成。在一些实施方案中,所述RAB1多肽包含SEQ ID NO:140或由其组成。

[0058] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含DNA复制解旋酶/核酸酶2 (DNA2) 多肽或由其组成。在一些实施方案中,所述DNA2多肽包含SEQ

ID NO:141或由其组成。

[0059] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含FLJ35220多肽或由其组成。在一些实施方案中,所述FLJ35220多肽包含SEQ ID NO:142或由其组成。

[0060] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含FLJ13173多肽或由其组成。在一些实施方案中,所述FLJ13173多肽包含SEQ ID NO:143或由其组成。

[0061] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含DNA修复内切核酸酶XPF (ERCC4) 多肽或由其组成。在一些实施方案中,所述ERCC4多肽包含SEQ ID NO:124或由其组成。

[0062] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R)) 多肽或由其组成。在一些实施方案中,所述RNA酶1 (K41R) 多肽包含SEQ ID NO:116或由其组成。

[0063] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R、D121E)) 多肽或由其组成。在一些实施方案中,所述RNA酶1 (RNA酶1 (K41R、D121E)) 多肽包含SEQ ID NO:117或由其组成。

[0064] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R、D121E、H119N)) 多肽或由其组成。在一些实施方案中,所述RNA酶1 (RNA酶1 (K41R、D121E、H119N)) 多肽包含SEQ ID NO:118或由其组成。

[0065] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,

其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(H119N))多肽或由其组成。在一些实施方案中, RNA酶1(RNA酶1(H119N))多肽包含SEQ ID NO:119或由其组成。

[0066] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,所述RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽包含SEQ ID NO:120或由其组成。

[0067] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,所述RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E))多肽包含SEQ ID NO:121或由其组成。

[0068] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,所述RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D))多肽包含SEQ ID NO:122或由其组成。

[0069] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含特诺伊林(Teneurin)跨膜蛋白1(TENM1)多肽或由其组成。在一些实施方案中,所述TENM1多肽包含SEQ ID NO:144或由其组成。

[0070] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含特诺伊林跨膜蛋白1(TENM2)多肽或由其组成。在一些实施方案中,所述TENM2多肽包含SEQ ID NO:145或由其组成。

[0071] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,

其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含核糖核酸酶 κ (RNA酶K)多肽或由其组成。在一些实施方案中,所述RNA酶K蛋白包含SEQ ID NO: 204或由其组成。

[0072] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含转录激活因子样效应物核酸酶(TALEN)多肽或所述多肽的核酸酶结构域或由其组成。在一些实施方案中,所述TALEN多肽包含以下序列或由其组成:

```
1 MRIGKSSGWL NESVSLEYEH VSPPTRPRDT RRRPRAAGDG GLAHLHRRRLA VGYAEDTPRT
61 EARSPAPRRP LPVAPASAPP APSLVPEPPM PVSLPAVSSP RFSAGSSAAI TDPFPSLPPT
121 PVLYAMAREL EALSDATWQP AVPLPAEPPT DARRGNTVFD EASASSPVIA SACPQAFASP
181 PRAPRSARAR RARTGGDAWP APTFLSRPSS SRIGRDVFGK LVALGYSREQ IRKLKQESLS
241 EIAKYHTTLT GQGFTHADIC RISRRRQSLR VVARNYPELA AALPELTRAH IVDIARQMSG
301 DLALQALLPV ATALTAAPLR LSASQIATVA QYGERPAIQA LYRLRRKLTR APLHLTPQQV
361 VAIASNTGGK RALEAVCVQL PVLRAAPYRL STEQVVAIAS NKGKQALEA VKAHLDDLGL
421 APYVLDTEQV VAIASHNGGK QALEAVKADL LDLRGAPYAL STEQVVAIAS HNGGKQALEA
481 VKADLLELRG APYALSTEQV VAIASHNGGK QALEAVKAHL LDLRGVPYAL STEQVVAIAS
541 HNGGKQALEA VKAQLDLRG APYALSTAQV VAIASNGGGK QALEGIGEQL LKLRTAPYGL
601 STEQVVAIAS HDGGKQALEA VGAQLVALRA APYALSTEQV VAIASNKGK QALEAVKAQL
661 LELRGAPYAL STAQVVAIAS HDGGNQALEA VGTQLVALRA APYALSTEQV VAIASHDGGK
721 QALEAVGAQL VALRAAPYAL NTEQVVAIAS SHGGKQALEA VRALFPDLRA APYALSTAQL
781 VAIASNPGGK QALEAVRALF RELRAAPYAL STEQVVAIAS NHGGKQALEA VRALFRGLRA
841 APYGLSTAQV VAIASSNGGK QALEAVWALL PVLRAPTYDL NTAQIVAIAS HDGGKPALEA
901 VWAKLPVLRG APYALSTAQV VAIACISGQQ ALEAIEAHMP TLRQASHSL PERVAIACI
961 GGRSAVEAVR QGLPVKAIRR IRREKAPVAG PPPASLGPTP QELVAVLHFF RAHQQPRQAF
1021 VDAAAFQAT RPALLRLLSS VGVTEIEALG GTIPDATERW QRLGRLGFR PATGAAAPSP
1081 DSLQGFAQSL ERTLGSFGMA GQSACSPHRK RPAETAIAPR SIRRSPNNAG QPSEPWPQQL
1141 AWLQRRKRRTA RSHIRADSAA SVPANLHLGT RAQFTPDRLR AEPGPIMQAH TSPASVSFGS
1201 HVAFEPGLPD PGTPTSADLA SFEAEPFGVG PLDFHLDWLL QILET (SEQ ID NO: 205)。
```

在一些实施方案中,所述TALEN多肽包含以下序列或由其组成:

1 mdpirsrtps parellpgpq pdrvqptadr ggappaggpl dglparrtms rtrlpsppap
61 spafsagsfs dllrqfdpsl ldtsllldsmp avgtphataa paecdevqsg lraaddpppt
121 vrvavtaarp prakpaprrr aaqpsdaspa aqvdlrtlgy sqqqqekikp kvgstvaqhh
181 ealvghgfth ahivalsrhp aalgtvavky qdmiaalpea thedivgvvk qwsagaralea
241 lltvagelrg pplqltdgql vkiakrggvt aveavhasrn altgaplnlt paqvaiasn
301 nggkqaletv qrllpvlcqa hgltpaqvva iashdggkqa letmqrllpv lcqahglppd
361 qvvaiaasnig gkqaletvqr llpvlcqhahg ltpdqvvaia shgggkqale tvqrllpvlc
421 qahgltpdqv vaiashdggk qaletvqrll pvlcqhahgt pdqvvaiaasn gggkqaletv
481 qrllpvlcqa hgltpdqvva iasnggkqal etvqrllpvl cqahgltpdq vvaiahdgg
541 kqaletvqrl lpvlcqhthgl tpaqvvaiahdggkqalet vqqlpvlcqhahgltpdqvv
601 aiasniggkq alatvqrllp vlcqhahgltp dqvvaiaasn gggkqaletvq rllpvlcqhah
661 gltpdqvvai asngggkqal etvqrllpvl cqahgltpdqv vvaiaasnigg kqaletvqrl
721 lpvlcqhahgl tpaqvvaiahdggkqalet vqrllpvlcqhahgltpdqvv aiasniggkq
781 aletvqrllp vlcqhahgltp eqvvaiaasn gggkqaletvq rllpvlcqhah gltpdqvvai
841 asngggkqal etvqrllpvl cqahgltpaq vvaiaasnigg kqaletvqrl lpvlcqhahgl
901 tlaqvvaiaasniggkqalet vqrllpvlcqhahgltpdqvv aiasniggkq aletvqrllp
961 vlcqhahgltp dqvvaiaasn gggkqaletvq rllpvlcqhah gltpdqvvai asngggkqale
1021 tvqrllpvlc qdhgltpdqv vaiaasniggk qaletvqrll pvlcqhahgltpdqvvai
1081 gggkqaletvq aqlsrpdal aaltndhlva laclggrpam davkkglpha pelirrvnrr
1141 igertshrva dyaqvrvle ffqchshpay afdeamtqfg msrnlvqlf rrvgvtelea
1201 rggtlppasq rwdrlqasg mkrakpspts aqtpdqaslh afadslerdl dapspmhgd
1261 qtgassrkrs rsdravtgps aqhsfevrp eqrdalhlpl swrvkrprtr igggldpdpgt
1321 piaadlaass tvmweqdaap fagaaddfpa fneeelawlm ellpqsgsvg gti (SEQ ID
NO: 206)。

[0073] 在本公开文本的组合物的一些实施方案中,包括以下那些实施方案:其中所述组合物包含含有特异性结合第一RNA分子内的第一靶序列的第一指导RNA(gRNA)的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;以及含有特异性结合第二RNA分子内的第二靶序列的第二指导RNA(gRNA)的第二序列,所述第二RNA结合蛋白包含锌指核酸酶多肽或所述多肽的核酸酶结构域或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含ZNF638多肽或所述多肽的核酸酶结构域或由其组成。在一些实施方案中,所述ZNF638多肽包含以下序列或由其组成:

1 MSRPREFNPRG DFPLQRPRAP NPSGMRPPGP FMRPGSMGLP RFYPAGRARG IPHRFAGHES
61 YQNMGPQRMN VQVTQHRTDP RLTKEKLDHFH EAQQKKKGKPH GSRWDDEPHI SASVAVKQSS
121 VTQVTEQSPK VQSRYTKESA SSILASFGLS NEDLEELSR YDEQLTPENM PLILRDIRMR
181 KMGRRLPNLP SQSRNKETLG SEAVSSNVID YGHASKYGYT EDPLEVRIYD PEIPTDEVEN
241 EFQSQQNISA SVPNPNVICN SMFPVEDVFR QMDFPGESSN NRSFFSVESG TKMSGHLHISG
301 GQSVLEPIKS VNQSINQTVS QTMSQSLIPP SMNQQPFSSE LISSVSQQER IPHEPVINSS
361 NVHVGSRGSK KNYQSQADIP IRSPFGIVKA SWLPKFSHAD AQKMKRLPTP SMMNDYYAAS
421 PRIFPHLCSL CNVECSHLKD WIQHQNTSTH IESCRQLRQQ YPDWNPEILP SRRNEGNRKE
481 NETPRRRSHS PSPRRSRRSS SSHRFRRSRS PMHYMYRPRS RSPRICHRFI SRYRSRSRSR
541 SPYRIRNPFR GSPKCFRSVS PERMSRRSVR SSDRKKALED VVQRSRGHGTE FNKQKHLEAA
601 DKGHSPAQKP KTSSGTKPSV KPTSATKSDS NLGGHSIRCK SKNLEDDTLS ECKQVSDKAV
661 SLQRKLREKQ SLHYGSVLLI TELPEDGCTE EDVRKLFQPF GKVNDVLIVP YRKEAYLEME
721 FKEAITAIMK YIETTPLTIK GKSVKICVPG KKKAQNKEVK KKTLESKKVS ASTLKRDATA
781 SKAVEIVTST SAAKTGQAKA SVAKVNKSTG KSASSVKSVV TVAVKGNKAS IKTAKSGGKK
841 SLEAKKTGNV KNKDSNKPVT IPENSEIKTS IEVKATENCA KEAISDAALE ATENEPLNKE
901 TEEMCVMLVS NLPNKGYSVE EVYDLAKPFG GLKDILILSS HKKAYIEINR KAAESMVKFY
961 TCFPVLM DGN QLSISMAPEN MNIKDEEAIF ITLVKENDPE ANIDTIYDRF VHLDNLPEDG
1021 LQCVLCVGLQ FGKVDHVFH SNRNKAILQL DSPESAQSMY SFLKQNPQNI GDHMLTCSLS
1081 PKIDLPEVQI EHDPELEKES PGLKNSPIDE SEVQTATDSP SVKPNELEEE STPSIQTETL

1141 VQQEEPCEEE AEKATCDSDF AVETLELETQ GEEVKKEIPL VASASVSIEQ FTENAEECAL
1201 NQQMFNSDLE KKGAEIINPK TALLPSDSVF AEERNLKGIL EESPSEAEDF ISGITQTMVE
1261 AVAEVEKNET VSEILPSTCI VTLVPGIPTG DEKTVDKKNI SEKKGNMDEK EEKEFNTKET
1321 RMDLQIGTEK AEKNEGRMDA EKVEKMAAMK EKPAENTLFK AYPNKGVGQA NKPDETSKTS
1381 ILAVSDVSSS KPSIKAVIVS SPKAKATVSK TENQKSFPKS VPRDQINA EK KLSAKEFGLL
1441 KPTSARSGLA ESSSKFKPTQ SSLTRGGSGR ISALQGKLSK LDYRDITKQS QETEARPSIM
1501 KRDDSNKTL AEQNTKNPKS TTGRSSKSKE EPLFPFNLDE FVTVDEVIEE VNPSQAKQNP
1561 LKGKRKETLK NVPFSELNLK KKKGKTSTPR GVEGELSFVT LDEIGEEEDA AAHLAQALVT
1621 VDEVIDEEEL NMEEMVKNSN SLFTLDELID QDDCISHSEP KDVTVLSVAE EQDLLKQERL
1681 VTVDEIGEVE ELPLNESADI TFATLNTKGN EGDTVRDSIG FISSQVPEDP STLVTVDEIQ
1741 DDSSDLHLVT LDEVTEEDED SLADFNNLKE ELNFVTVDEV GEEEDGDNDL KVELAQSKND
1801 HPTDKKGNRK KRAVDTKKTK LESLSQVGPV NENVMEEDLK TMIERHLTAK TPTKRVRIGK
1861 TLPSEKAVVT EPAKGEEAFQ MSEVDEESGL KDSEPERKRK KTEDSSSGKS VASDVPEELD
1921 FLVPKAGFFC PICSLFYSGE KAMTNHCKST RHKQNTKFM AKQRKEKEQN EAEERSSR

(SEQ ID NO: 207)。

[0074] 在本公开文本的组合物的一些实施方案中,所述组合物还包含(a)包含特异性结合于RNA分子内的gRNA的序列;以及(b)编码核酸酶的序列。在一些实施方案中,编码核酸酶的所述序列包含从CRISPR/Cas蛋白分离或衍生的序列。在一些实施方案中,所述CRISPR/Cas蛋白是从以下中的任一种分离或衍生的:I型、IA型、IB型、IC型、ID型、IE型、IF型、IU型、III型、IIIA型、IIIB型、IIIC型、IIID型、IV型、IVA型、IVB型、II型、IIA型、IIB型、IIC型、V型或VI型CRISPR/Cas蛋白。在一些实施方案中,编码核酸酶的所述序列包含从TALEN或其核酸酶结构域分离或衍生的序列。在一些实施方案中,编码核酸酶的所述序列包含从锌指核酸酶或其核酸酶结构域分离或衍生的序列。在一些实施方案中,所述靶序列包含编码适应性免疫应答的组分的序列。

[0075] 本公开文本提供了一种包含本公开文本的组合物载体。在一些实施方案中,所述载体是病毒载体。在一些实施方案中,所述载体包含从慢病毒、腺病毒、腺相关病毒(AAV)载体或逆转录病毒分离或衍生的序列。在一些实施方案中,所述载体无复制能力。

[0076] 本公开文本提供了一种包含本公开文本的组合物载体。在一些实施方案中,所述载体是病毒载体。在一些实施方案中,所述载体包含从腺相关载体(AAV)分离或衍生的序列。在一些实施方案中,所述腺相关病毒(AAV)是分离的AAV。在一些实施方案中,所述腺相关病毒(AAV)是自身互补的腺相关病毒(scAAV)。在一些实施方案中,所述腺相关病毒(AAV)是重组腺相关病毒(rAAV)。在一些实施方案中,所述腺相关病毒(AAV)包含从血清型AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11或AAV12的AAV分离或衍生的序列。在一些实施方案中,所述腺相关病毒(AAV)包含从血清型AAV9的AAV分离或衍生的序列。在一些实施方案中,所述腺相关病毒(AAV)包含从Anc80分离或衍生的序列。

[0077] 本公开文本提供了一种包含本公开文本的组合物的载体。在一些实施方案中,所述载体是病毒载体。在一些实施方案中,所述载体是逆转录病毒。

[0078] 本公开文本提供了一种包含本公开文本的组合物的载体。在一些实施方案中,所述载体是病毒载体。在一些实施方案中,所述载体是慢病毒。

[0079] 本公开文本提供了一种包含本公开文本的组合物的载体。在一些实施方案中,所述载体是非病毒载体。在一些实施方案中,所述非病毒载体包含纳米颗粒、胶束、脂质体或阳离子脂质体/DNA复合物(lipoplex)、聚合物囊泡(polymersome)、聚合物/DNA复合物(polyplex)或树枝状聚合物。

[0080] 本公开文本提供了一种包含本公开文本的载体的组合物。

[0081] 本公开文本提供了一种包含本公开文本的载体的细胞。

[0082] 本公开文本提供了一种包含本公开文本的细胞的细胞。

[0083] 在本公开文本的细胞的一些实施方案中,所述细胞是哺乳动物细胞。在一些实施方案中,所述细胞是人细胞。

[0084] 在本公开文本的细胞的一些实施方案中,所述细胞是免疫细胞。在一些实施方案中,所述免疫细胞是T淋巴细胞(T细胞)。在一些实施方案中,所述T细胞是效应T细胞、辅助T细胞、记忆T细胞、调节性T细胞、自然杀伤T细胞、粘膜相关恒定T细胞或 $\gamma\delta$ T细胞。

[0085] 在本公开文本的细胞的一些实施方案中,所述细胞是免疫细胞。在一些实施方案中,所述免疫细胞是抗原呈递细胞。在一些实施方案中,所述抗原呈递细胞是树突细胞、巨噬细胞或B细胞。在一些实施方案中,所述抗原呈递细胞是体细胞。

[0086] 在本公开文本的细胞的一些实施方案中,所述细胞是免疫细胞。在一些实施方案中,所述细胞是健康细胞。在一些实施方案中,所述细胞不是健康细胞。在一些实施方案中,所述细胞是从患有疾病或障碍的受试者分离或衍生的。

[0087] 本公开文本提供了一种包含本公开文本的细胞的组合物。

[0088] 本公开文本提供了一种包含本公开文本的多个细胞的组合物。

[0089] 本公开文本提供了一种从适应性免疫应答中掩蔽细胞的方法,其包括使本公开文本的组合物与所述细胞接触以产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述RNA分子编码适应性免疫应答的组分。在一些实施方案中,所述细胞是体内的、体外的、离体的或原位的。在一些实施方案中,所述细胞是体外的或离体的。在一些实施方案中,多个细胞包含所述细胞。在一些实施方案中,使所述多个细胞中的每个细胞与所述组合物接触,从而产生多个修饰的细胞。在一些实施方案中,所述方法还包括向受试者施用所述修饰的细胞。在一些实施方案中,所述方法还包括向受试者施用所述多个修饰的细胞。在一些实施方案中,所述细胞是自体的。在一些实施方案中,所述细胞是同种异体的。在一些实施方案中,所述多个修饰的细胞是自体的。在一些实施方案中,所述多个修饰的细胞是同种异体的。在一些实施方案中,适应性免疫应答的所述组分包含以下的组分或由其组成:I型主要组织相容性复合物(MHC I)、II型主要组织相容性复合物(MHC II)、T细胞受体(TCR)、共刺激分子或其组合。在一些实施方案中,所述MHC I组分包含 $\alpha 1$ 链、 $\alpha 2$ 链、 $\alpha 3$ 链或 $\beta 2M$ 蛋白。在一些实施方案中,适应性免疫应答的所述组分包含MHC I $\beta 2M$ 蛋白或由其组成。在一些实施方案中,所述MHC II组分包含 $\alpha 1$ 链、 $\alpha 2$ 链、 $\beta 1$ 链或 $\beta 2$ 链。在一些实施方案中,所述TCR组分包含 α 链和 β 链。在一些实施方案中,所述共刺激分子包含分化群

28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS) 或ICOS配体 (ICOSLG) 蛋白。在一些实施方案中,适应性免疫应答的蛋白质组分是但不限于 β -2-微球蛋白 (β 2M)、人白细胞抗原A (HLA-A)、人白细胞抗原B (HLA-B)、人白细胞抗原C (HLA-C)、分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS)、ICOS配体 (ICOSLG)、OX40L、白细胞介素12 (IL12) 或CC趋化因子受体7 (CCR7)。

[0090] 本公开文本提供了一种防止或减少受试者的适应性免疫应答的方法,其包括向所述受试者施用治疗有效量的本公开文本的组合物,其中使所述组合物与所述受试者的至少一个细胞接触,产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述RNA分子编码适应性免疫应答的组分。

[0091] 本公开文本提供了一种治疗受试者的疾病或障碍的方法,其包括向所述受试者施用治疗有效量的本公开文本的组合物,其中使所述组合物与所述受试者的至少一个细胞接触,产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述组合物防止或减少对所述修饰的细胞的适应性免疫应答。

[0092] 在本公开文本的方法的一些实施方案中,适应性免疫应答的所述组分包含以下的组分或由其组成:I型主要组织相容性复合物 (MHC I)、II型主要组织相容性复合物 (MHC II)、T细胞受体 (TCR)、共刺激分子或其组合。在一些实施方案中,所述MHC I组分包含 α 1链、 α 2链、 α 3链或 β 2M蛋白。在一些实施方案中,适应性免疫应答的所述组分包含MHC I β 2M蛋白或由其组成。在一些实施方案中,所述MHC II组分包含 α 1链、 α 2链、 β 1链或 β 2链。在一些实施方案中,所述TCR组分包含 α 链和 β 链。在一些实施方案中,所述共刺激分子包含分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS) 或ICOS配体 (ICOSLG) 蛋白。

[0093] 在本公开文本的治疗疾病或障碍的方法的一些实施方案中,所述疾病或障碍是遗传疾病或障碍。在一些实施方案中,所述疾病或障碍是单基因遗传疾病或障碍。在一些实施方案中,所述疾病或障碍是由微卫星不稳定性引起的。在一些实施方案中,所述微卫星不稳定性发生在DNA序列的至少1、2、3、4、5或6个重复的基序处。在一些实施方案中,RNA分子包含所述DNA序列的转录物,并且其中所述组合物与所述RNA分子的靶序列结合,所述靶序列包含至少1、2、3、4、5或6个重复的基序。

[0094] 在本公开文本的方法的一些实施方案中,全身性施用所述组合物。在一些实施方案中,静脉内施用所述组合物。在一些实施方案中,通过注射或输注施用所述组合物。

[0095] 在本公开文本的方法的一些实施方案中,局部施用所述组合物。在一些实施方案中,通过骨内、眼内、脑内或脊柱内途径施用所述组合物。在一些实施方案中,通过注射或输注施用所述组合物。

[0096] 在本公开文本的方法的一些实施方案中,所述组合物的治疗有效量是单剂量。

[0097] 在本公开文本的方法的一些实施方案中,所述组合物是非基因组整合的。

附图说明

[0098] 所述专利或申请文件含有至少一张彩色附图。在请求并支付必要的费用后,官方将会提供带有一张或多张彩色附图的本专利或专利申请公开案的副本。

[0099] 图1A是描绘了示例性RNA内切核酸酶-空肠弯曲杆菌 (*C. jejuni*) Cas9融合蛋白的

示意图。

[0100] 图1B是描绘了在E43和E67 CjeCas9-内切核酸酶融合物二者的存在下Zika NS5的表达水平的变化的图,所述融合物具有含有如表8中指示的各种靶向NS5的间隔子序列的sgRNA。将Zika NS5表达展示为相对于加载有含有对照(λ)间隔子序列的sgRNA的内切核酸酶的变化倍数。

[0101] 图2A是用加载有含有靶向Zika NS5的间隔子序列的sgRNA的CjeCas9-内切核酸酶融合物转染的细胞的荧光显微镜检查图像。

[0102] 图2B是描绘了如与加载有不靶向Zika NS5的sgRNA的CjeCas9-内切核酸酶融合物相比在加载有靶向Zika NS5的适当sgRNA的CjeCas9-内切核酸酶融合物的存在下Zika NS5的表达的变化的图。

[0103] 图3是用于本公开文本的组合物中的示例性内切核酸酶的列表。

[0104] 图4是描绘了编码示例性RNA内切核酸酶-空肠弯曲杆菌Cas9融合蛋白和用于在基因疗法的背景下调节免疫应答的两种gRNA分子的构建体的示意图。本发明描述了使用基于CRISPR的基因疗法或在AAV中编码的其他非自身蛋白同时改变宿主基因表达以防止对非自身蛋白的适应性免疫应答来解决人类疾病的方式。在一个实施方案中,所述AAV颗粒(左)携带一对指导RNA和CRISPR相关(Cas)蛋白。所述指导物分别靶向与适应性免疫应答相关的基因和促进治疗效果的基因(或基因产物)。递送至靶组织后,所述靶向免疫应答的指导物减少与抗原呈递(β -2-微球蛋白,B2M)或T细胞的共刺激(ICOSLG、CD80、CD86、OX40L、IL12、CCR7)相关的基因的表达。抗原呈递抑制防止对治疗性转基因(如Cas蛋白)具有特异性的T辅助(Th)细胞的形成,同时共刺激抑制防止对转基因具有特异性的Th细胞的激活。

具体实施方式

[0105] 本公开文本提供了用于通过靶向修饰的细胞的RNA分子同时治疗疾病并且同时从适应性免疫应答中掩蔽所述修饰的细胞的组合物和方法。通过抑制或减少在所述修饰的细胞中适应性免疫应答的组分的表达,所述修饰的细胞对于宿主免疫系统不可见。例如,本公开文本的组合物可以同时靶向与遗传疾病或障碍相关的RNA分子和编码MHC I的 β 2M亚基的RNA分子。通过选择性靶向编码MHC I的 β 2M亚基的RNA分子,所述组合物防止所述修饰的细胞在所述修饰的细胞表面上展示一种或多种衍生自RNA靶向构建体、载体或其组合的抗原肽。因此,受试者的免疫系统不会将所述修饰的细胞鉴定为含有外源序列,并且不会尝试针对所述修饰的细胞产生免疫应答。此方法提高了治疗所述遗传疾病或障碍的治疗功效,同时避免了基因疗法的常见副作用。

靶向RNA的融合蛋白组合物

[0106] 本公开文本提供了一种组合物,其包含(a)包含特异性结合RNA分子内的靶序列的指导RNA(gRNA)的序列;以及(b)编码融合蛋白的序列,所述序列包含编码第一RNA结合多肽的序列和编码第二RNA结合多肽的序列,其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性,其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0107] 在本公开文本的组合物的一些实施方案中,所述靶序列包含至少一个重复的序

列。

[0108] 在本公开文本的组合物的一些实施方案中,所述gRNA序列包含能够在真核细胞中表达所述gRNA的启动子。

[0109] 在本公开文本的组合物的一些实施方案中,所述真核细胞是动物细胞。在一些实施方案中,所述动物细胞是哺乳动物细胞。在一些实施方案中,所述动物细胞是人细胞。

[0110] 在本公开文本的组合物的一些实施方案中,所述启动子是组成型活性启动子。在一些实施方案中,所述启动子序列是从能够驱动RNA聚合酶的表达的启动子分离或衍生的。在一些实施方案中,所述启动子序列是从U6启动子分离或衍生的。在一些实施方案中,所述启动子序列是从能够驱动转移RNA (tRNA) 的表达的启动子分离或衍生的。在一些实施方案中,所述启动子序列是从以下启动子分离或衍生的:丙氨酸tRNA启动子、精氨酸tRNA启动子、天冬酰胺tRNA启动子、天冬氨酸tRNA启动子、半胱氨酸tRNA启动子、谷氨酰胺tRNA启动子、谷氨酸tRNA启动子、甘氨酸tRNA启动子、组氨酸tRNA启动子、异亮氨酸tRNA启动子、亮氨酸tRNA启动子、赖氨酸tRNA启动子、甲硫氨酸tRNA启动子、苯丙氨酸tRNA启动子、脯氨酸tRNA启动子、丝氨酸tRNA启动子、苏氨酸tRNA启动子、色氨酸tRNA启动子、酪氨酸tRNA启动子或缬氨酸tRNA启动子。在一些实施方案中,所述启动子序列是从缬氨酸tRNA启动子分离或衍生的。

[0111] 在本公开文本的组合物的一些实施方案中,包含所述gRNA的序列还包含与所述靶RNA序列特异性结合的间隔子序列。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有100%互补性。在一些实施方案中,所述间隔子序列包含20个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含21个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含以下序列或由其组成:

UGGAGCGAGCAUCCCCAAA (SEQ ID NO:1)、GUUUGGGGAUGCUCGUCCA (SEQ ID NO:2)、CCCUCACUGCUGGGGAGUCC (SEQ ID NO:3)、GGACUCCCCAGCAGUGAGGG (SEQ ID NO:4)、GCAACUGGAUCAUUUGCUG (SEQ ID NO:5)、GCAGCAAAUUGAUCCAGUUGC (SEQ ID NO:6)、GCAUUCUUAUCUGGUCAGUGC (SEQ ID NO:7)、GCACUGACCAGAUAAAGAAUG (SEQ ID NO:8)、GAGCAGCAGCAGCAGCAGCAG (SEQ ID NO:9)、GCAGGCAGGCAGGCAGGCAGG (SEQ ID NO:10)、GCCCCGGCCCCGGCCCCGGC (SEQ ID NO:11) 或GCTGCTGCTGCTGCTGCTGC (SEQ ID NO:84)、GGGGCCGGGGCCGGGGCCGG (SEQ ID NO:74)、GGGCCGGGGCCGGGGCCGGG (SEQ ID NO:75)、GGCCGGGGCCGGGGCCGGGG (SEQ ID NO:76)、GCCGGGGCCGGGGCCGGGGC (SEQ ID NO:77)、CCGGGGCCGGGGCCGGGGCC (SEQ ID NO:78) 或CGGGGCCGGGGCCGGGGCCG (SEQ ID NO:79)。

[0112] 在本公开文本的组合物的一些实施方案中,包含所述gRNA的序列还包含与所述靶RNA序列特异性结合的间隔子序列。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有100%互补性。在一些实施方案中,所述间隔子序列包含20个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含21个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含以下序列或由其组成:

GUGAUAAGUGGAAUGCCAUG (SEQ ID NO:14)、CUGGUGAACUCCGAUAGUG (SEQ ID NO:15) 或GAGATATAGCCTGGTGGTTC (SEQ ID NO:16)。

[0113] 在本公开文本的组合物的一些实施方案中,包含所述gRNA的序列还包含与所述靶RNA序列特异性结合的间隔子序列。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。在一些实施方案中,所述间隔子序列与所述靶RNA序列具有100%互补性。在一些实施方案中,所述间隔子序列包含20个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含21个核苷酸或由其组成。在一些实施方案中,所述间隔子序列包含以下序列或由以下序列组成,所述序列包含序列CUG (SEQ ID NO:18)、CCUG (SEQ ID NO:19)、CAG (SEQ ID NO:80)、GGGGCC (SEQ ID NO:81) 或其任何组合的至少1、2、3、4、5、6或7个重复。

[0114] 在本公开文本的组合物的一些实施方案中,包含所述gRNA的序列还包含与所述第一RNA结合蛋白特异性结合的支架序列。在一些实施方案中,所述支架序列包含茎环结构。在一些实施方案中,所述支架序列包含90个核苷酸或由其组成。在一些实施方案中,所述支架序列包含93个核苷酸或由其组成。在一些实施方案中,所述支架序列包含以下序列或由其组成:GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:83)。在一些实施方案中,所述支架序列包含以下序列或由其组成:

GGACAGCAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUU (SEQ ID NO:17)。在一些实施方案中,所述支架序列包含以下序列或由其组成:

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:82) 或GUUUUAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:13)。

[0115] 在本公开文本的组合物的一些实施方案中,所述gRNA不结合或不选择性结合所述RNA分子内的第二序列。

[0116] 在本公开文本的组合物的一些实施方案中,RNA基因组或RNA转录组包含所述RNA分子。

[0117] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白包含CRISPR-Cas蛋白。在一些实施方案中,所述CRISPR-Cas蛋白是II型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cas9多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白包含天然RNA核酸酶活性。在一些实施方案中,所述天然RNA核酸酶活性被降低或抑制。在一些实施方案中,所述天然RNA核酸酶活性被增加或诱导。在一些实施方案中,所述CRISPR-Cas蛋白包含天然DNA核酸酶活性并且所述天然DNA核酸酶活性被抑制。在一些实施方案中,所述CRISPR-Cas蛋白包含突变。在一些实施方案中,所述CRISPR-Cas蛋白的核酸酶结构域包含所述突变。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的核酸中。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的氨基酸中。在一些实施方案中,所述突变包含取代、插入、缺失、移码、倒位或转座。在一些实施方案中,所述突变包含核酸酶结构域、所述核酸酶结构域内的结合位点、所述核酸酶结构域内的活性位点或所述核酸酶结构域内的至少一个必需氨基酸残基的缺失。

[0118] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白包含CRISPR-

Cas蛋白。在一些实施方案中,所述CRISPR-Cas蛋白是V型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cpf1多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白包含天然RNA核酸酶活性。在一些实施方案中,所述天然RNA核酸酶活性被降低或抑制。在一些实施方案中,所述天然RNA核酸酶活性被增加或诱导。在一些实施方案中,所述CRISPR-Cas蛋白包含天然DNA核酸酶活性并且所述天然DNA核酸酶活性被抑制。在一些实施方案中,所述CRISPR-Cas蛋白包含突变。在一些实施方案中,所述CRISPR-Cas蛋白的核酸酶结构域包含所述突变。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的核酸中。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的氨基酸中。在一些实施方案中,所述突变包含取代、插入、缺失、移码、倒位或转座。在一些实施方案中,所述突变包含核酸酶结构域、所述核酸酶结构域内的结合位点、所述核酸酶结构域内的活性位点或所述核酸酶结构域内的至少一个必需氨基酸残基的缺失。

[0119] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白包含CRISPR-Cas蛋白。在一些实施方案中,所述CRISPR-Cas蛋白是VI型CRISPR-Cas蛋白。在一些实施方案中,所述第一RNA结合蛋白包含Cas13多肽或其RNA结合部分。在一些实施方案中,所述第一RNA结合蛋白包含Cas13d多肽或其RNA结合部分。在一些实施方案中,所述CRISPR-Cas蛋白包含天然RNA核酸酶活性。在一些实施方案中,所述天然RNA核酸酶活性被降低或抑制。在一些实施方案中,所述天然RNA核酸酶活性被增加或诱导。在一些实施方案中,所述CRISPR-Cas蛋白包含天然DNA核酸酶活性并且所述天然DNA核酸酶活性被抑制。在一些实施方案中,所述CRISPR-Cas蛋白包含突变。在一些实施方案中,所述CRISPR-Cas蛋白的核酸酶结构域包含所述突变。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的核酸中。在一些实施方案中,所述突变发生在编码所述CRISPR-Cas蛋白的氨基酸中。在一些实施方案中,所述突变包含取代、插入、缺失、移码、倒位或转座。在一些实施方案中,所述突变包含核酸酶结构域、所述核酸酶结构域内的结合位点、所述核酸酶结构域内的活性位点或所述核酸酶结构域内的至少一个必需氨基酸残基的缺失。

[0120] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白包含Pumilio和FBF (PUF) 蛋白。在一些实施方案中,所述第一RNA结合蛋白包含基于Pumilio的联合体(PUMBY) 蛋白。在一些实施方案中,本公开文本的PUF1蛋白包含以下的氨基酸序列或由其组成:

MDKSKQMNIN NLSNIPEVID PGITIPYEE EYENNGESNS QLQQQPQKLG SYRSRAGKFS	60
NTLSNLLPSI SAKLHHSKKN SHGKNGAEFS SSNNSSQSTV ASKTPRASPS RSKMMESSID	120
GVTMDRPGSL TPPQDMEKLV HFPDSSNNFL IPAPRGSSDS FNLPHQISRT RNNTMSSQIT	180
SISSIAPKPR TSSGIWSSNA SANDPMQQHL LQQLOPTTSN NTTNSNTLND YSTKTAYFDN	240
MVSTSGSQMA DNKMNTNNLA IPNSVWSNTR QRSQSNASSI YTDAPLYEQP ARASISSHYT	300
IPTQESPLIA DEIDPQSINW VTMDPTVPSI NQISNLLPTN TISISNVFPL QHQQPQLNNA	360
INLTSTSLAT LCSKYGEVIS ARTLRNLNMA LVEFSSVES A VKALDSLQ GK EVSMIGAPSK	420
ISFAKILPMH QQPPQFLLNS QGLPLGLENN NLQPQPLLQE QLFNGAVTFQ QQGNVSIPVF	480
NQSQSQSQHQ NHSSGSAGFS NVLHGYNNNN SMHGNNNSA NEKEQCPFPL PPPNVNEKED	540
LLREIIELEF ANSDEYQINS LIKKSINHKG TSDTQNF GPL PEPLSGREFD PPKLRELKRS	600
IDSNAFSDLE IEQLAIAMLD ELPELSSDYL GNTIVQKLFE HSSDIKDIM LRKTSKYLTS	660
MGVHKNGTWA CQKMITMAHT PRQIMQVTQG VKDYCTPLIN DQFGNYVIQC VLKFGFPWNQ	720
FIFESIIANF WVIVQNRYGA RAVRACLEAH DIVTPEQSIV LSAMIVTYAE YLSTNSNGAL	780
LVTWFLDTSV LPNRHSILAP RLTKRIVELC GHRLASLTIL KVLNYRGDDN ARKIILDSLF	840
GNVNAHDSSP PKELTKLLCE TNYGPTFVHK VLAMPLLEDD LRAHIKQVR KVLTDSTQIQ	900
PSRRLLEEVG LASPSSTHNK TKQQQQQHHN SSISHMFATP DTSGQHMRGL SVSSVKSGGS	960
KHTTMNTTTT NGSSASTLSP GQPLNANSNS SMGYFSYPGV FVSGFSGNA SNGYAMNDD	1020
LSSQFDMLNF NNGTRLSLPQ LSLTNHNNTT MELVNNVGSS QPHTNNNNNN NNTNYDDNT	1080
VFETLTLHSA N	1091

(SEQ ID NO: 208)。

在一些实施方案中,本公开文本的PUF3蛋白包含以下的氨基酸序列或由其组成:

1 MEMNMDMD MELASIVSSL SALSHSNNNG GQAAAAGIVN GGAAGSQQIG GFRRSSFTTA
61 NEVDSEILLL HGSSESSPIF KKTALSVGTA PPFSTNSKKF FGNGGNYYQY RSTDATSLSS

121 ASYNNYHTHH TAANLGKNNK VNHLLGQYSA SIAGPVYYNG NDNNNSGGEG FFEKFGKSLI
181 DGTRELESQD RPDVANTQSQ FISKSVSNAS LDTQNTFEQN VESDKNFNKL NRNTTNSGSL
241 YHSSSNSGSS ASLESENAHY PKRNIWNVAN TPVFRPSNNP AAVGATNVAL PNQQDGPANN
301 NFPPYMNGFP PNQFHQGPY QNFPNYLIGS PSNFISQMIS VQIPANEDTE DSNKKKKKA
361 NRPSSVSSPS SPPNNSPFPF AYPNPMFMF PPPLSAPQQQ QQQQQQQQQE DQQQQQQQEN
421 PYIYYPTNP IPVKMPKDEK TFKKRNNKNH PANNSNNANK QANPYLENSI PTKNTSKKNA
481 SSKSNESTAN NHKSHSHSH HSQSLQQQQQ TYHRSPLEQ LRNSSSDKNS NSNMSLKDIF
541 GHSLEFCKDQ HGSRFIQREL ATSPASEKEV IFNEIRDDAI ELSNDVFGNY VIOKFFFEFGS
601 KIQKNTLVDQ FKGNMQLSL QMYACRVIQK ALEYIDSNQR IELVLELSDS VLQMIKDQNG
661 NHVIQKAJET IPIEKLPFIL SSLTGHYHL STHSYGCRVI QRLLEFGSSE DQESILNELK
721 DFIPYLIQDQ YGNYVIQYVL QQDQFTNKEM VDIKQEIIET VANNVVEYSK HKFASNVVEK
781 SILEGSKNQK DLIISKILPR DKNHALNLED DSPMILMIK QFANYVIQKL VNVSEGEKK
841 LIVIAIRAYL DKLKNSNSLG NRHLASVEKL AALVENAEV (SEQ ID NO: 209)。

在一些实施方案中,本公开文本的PUF4蛋白包含以下的氨基酸序列或由其组成:

1 MSTKGLKEEI DDVPSVDPV SETVNSALEQ LQLDDPEENA TSNAFANKVS QDSQFANGPP
61 SQMFPHQMM GGMGFMPYSQ MMQVPHNPCP FFPPPDFNDP TAPLSSSPLN AGGPPMLFKN
121 DSLPFQMLSS GAAVATQGGQ NLNPLINDNS MKVLPIASAD PLWTHSNVPG SASVAIEETT
181 ATLQESLPSK GRESNNKASS FRRQTFHALS PTDLINAANN VTLSKDFQSD MQNFSKAKKP
241 SVGANNAKT RTQSFIDNT PSSTSFIPPT NSVSEKLSDF KIETSKEDLI NKTAPAKKES
301 PTTYGAAYPY GGPLLQNPFI MPGHPHNISS PIYGIRSPFP NSYEMGAQFQ PFSPILNPTS
361 HSLNANSPIP LTQSPIHLAP VLNPSNSVA FSDMKNDGGK PTTDNDKAGP NVRMDLINPN
421 LGPSMQPFHI LPPQQNTPPP PWLYSTPPPF NAMVPPHLLA QNHMPLMNSA NNKHHGRNNN
481 SMSSHNDNDN IGNSNYNNKD TGRSNVGKMK NMKNSYHGY NNNNNNNNNN NNNNSNATN
541 SNSAEKQRKI EESSRFADAV LDQYIGSIHS LCKDQHGRF LQQLDILGS KAADAIFEET
601 KDYTVELMTD SFGNYLIQKL LEEVTTEQRI VLTKISSPHF VEISLNPHGT RALQKLIETI
661 KTDEEAQIVV DSLRPYTVQL SKDLNGNHVI QKCLQRLKPE NFQFIFDAIS DSCIDIATHR
721 HGCCVLQRCL DHGTTEQCDN LCDKLLALVD KLTLDPFNGY VVQYIITKEA EKNKYDYTHK
781 IVHLLKPRAI ELSIHKFGSN VIEKILKTAI VSEPMILEIL NNGGETGIQS LLNDSYGNVY
841 LQTALDISHK QNDYLYKRLS EIVAPLLVGP IRNTPHGKRI IGMHLHDS (SEQ ID NO:
210)。

在一些实施方案中,本公开文本的PUF5蛋白包含以下的氨基酸序列或由其组成:

1 MSDSTGRINS KASDSSSISD HQTADLSIFN GSFDGGAFSS SNIPLFNFMG TGNQRFQYSP
61 HPFAKSSDPC RLAALTPSTP KGPLNLTPAD FGLADFSVGN ESFADFTANN TSFVGNVQSN
121 VRSTRLLPAW AVDNNGNIRD DLTLDQDVSN GSLIDFAMDR TGVKFLERHF PEDHDNEMHF
181 VLFDKLTEQG AVFTSLCRSA AGNFIIQKFV EHATLDEQER LVRKMCNGL IEMCLDKFAC
241 RVVQMSIQKF DVSIAMKLVE KISSLDLPL CTDQCAIHVL QKVVKLLPIS AWSFFVKFLC
301 RDDNLMTVCQ DKYGCRLVQQ TIDKLSDNPK LHCFTNRLQL LHGLMTSVAR NCFRLSSNEF
361 ANYVVQYVIK SSGVMEMYRD TIEKCLLRN ILSMSQDKYA SHVVEGAFLF APPLLLSEMM
421 DEIFDGYVKD QETNRDALDI LLFHQYGNV VQQMISICIS ALLGKEERKM VASEMRLYAK
481 WFDRIKNRVN RHSGRLERFS SGKKIIESLQ KLNVPMTMTN EMPYWAMPT PLMDISAHFM
541 NKLNFQKNSV FDE (SEQ ID NO: 211)。

在一些实施方案中,本公开文本的PUF6蛋白包含以下的氨基酸序列或由其组成:

1 MTPNRRSTDY YNMLGASFDF DPDFSLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDLNDY
61 IFNSGSGSSD DETPPPAPI FISLEEVLN GLLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKLTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMGDGL LVMCKDKFAC
181 RVVQLALQKF DHSNVFQLIQ ELSTFDLAAM CTDQISIHVI QRVVKQLPVD MWTFFVHFLS
241 SGDSLMAVCQ DKYGCRLVQQ VIDRLAENPK LPCFKFRIQL LHSLMTCIVR NCYRLSSNEF
301 ANYVIQYVIK SSGIMEMYRD TIIDKCLLRN LLSMSQDKYA SHVIEGAFLF APPALLHEMM
361 EEIFSGYVKD VELNRDALDI LLFHQYGNV VQQMISICTA ALIGKEERQL PPAILLLYSG
421 WYEKMKQKRVL QHASRLERFS SGKKIIDSVM RHGVPTAAAI NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 212)。

在一些实施方案中,本公开文本的PUF7蛋白包含以下的氨基酸序列或由其组成:

1 MTPNRRSTDY YNMLGASFDF DPDFSLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDSDSY
61 IFNSGSGSSD AETPAPVAPI FISLEDVLN GLLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKFTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMGDGL LAMCKDKFAC
181 RVVQLALQKF DHSNVFQLIQ ELSTFDLAAM CTDQISIHVI QRVVKQLPVD MWTFFVHFLS
241 SGDSLMAVCQ DKYGCRLVQQ VIDRLAENPK LPCFKFRIQL LHSLMTCIVR NCYRLSSNEF
301 ANYVIQYVIK SSGIMEMYRD TIIDKCLLRN LLSMSQDKYA SHVIEGAFLF APPALLHEMM
361 EEIFSGYVKD VESNRDALDI LLFHQYGNV VQQMISICTA ALIGKEEREL PPAILLLYSG
421 WYEKMKQKRVL QHASRLERFS SGKKIIDSVM RHGVPTAAAV NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 213)。

在一些实施方案中,本公开文本的PUF8蛋白包含以下的氨基酸序列或由其组成:

1 MSRPISIGNT CTFDPSASPI ESLGRSIGAQ KIVDSVCGSP IRSYGRHIST NPKNERLPDT
 61 PEFQFATYMH QGGKVIGQNT LHMFGTPPSC YCAQENIPIS SNVGHVLSTI NNNYMNHQYN
 121 GSNMFSNQMT QMLQAQAYND LQMHQAHSQS IRVPVQPSAT GIFSNPYREP TTTDDLTRY
 181 RANPAMMKNL KLSDIRGALL KFAKDQVGSF FIQQELASSK DRFEKDSIFD EVVSNADELV
 241 DDIFGNYVVQ KFFEYGEERH WARLVDIAID RVPEYAFQMY ACRVLQKALE KINEPLQIKI
 301 LSQIRHVIHR CMKDQNGNHV VQKAIEKVSP QYVQFIVDTL LESSNTIYEM SVDPYGCRVV
 361 QRCLEHCSPS QTKPVIGQIH KRFDEIANNQ YGNYVVQHVI EHGSEEDRMV IVTRVSNNLF
 421 EFATHKYSSN VIEKCLEQGA VYHKSMIVGA ACHHQEGSVP IVVQMMKDQY ANYVVQKMFD
 481 QVTSEQRREL ILTVRPHIPV LRQFPHGKHI LAKLEKYFQK PAVMSYPYQD MQGSH (SEQ
 ID NO: 214)。

在一些实施方案中,本公开文本的PUF9蛋白包含以下的氨基酸序列或由其组成:

1 MADPNWAYAP PTNYYADHSI AKPIMISGGH PSQDQGHSPK SESFGQSVTT AFNGMVDNLV
 61 GSPSSSVQQR NYFTTTPFPI SRSPNDRNDD KIMGNGSYGV PIPIPQDGVV QGTPDFQMTF
 121 FLQGGHLLIG GSPNGPVQVS GNWYSGGAGI FSTMQQADPS NGMPGMAAEF VNNENGMPGP
 181 NGMHQQAMIS GSPPFPYQNM MNLTTSFGAM GLGPQQIQQR DPQMFQQPIL HEPIQGMAQN
 241 GFGQQVFFFTQ MQNQQHPOGQ AQQQLQQLAQ QHQQQQNSQQ FFGQGPNGMG NGGVMNDWSQ
 301 RSFGMPQQQA QQNGLPNFS QNPPRRRGPE DPNGQTPKTL QDIKNNVIEF AKDQHGSRFI
 361 QQKLERASLR DKAAIFTPVL ENAEELMTDV FGNYVIQKFF EFGNNEQRNQ LVGTIRGNVM
 421 KLALQMYGCR VIQKALEYVE EKYQHEILGE MEGQVLKCVK DQNGNHVIQK VIERVEPERL
 481 QFIIDAFTKN NSDNVYTLVS HPYGCRVIQR VLEYCNEEQK QPVLDAQIH LKQLVLDQYG
 541 NYVIQHVIEH GSPSDKEQIV QDVISDDLK FAQHKFASNV IEKCLTFGGH AERNLIIDKV
 601 CGDPNDPSPP LLQMMKDPFA NYVVQKMLDV ADPQHRKKIT LTIKPHIATL RKNYFGKHIL
 661 LKLEKYFAKQ APANSSNSSS NDQIYEHSPF DIPLGADFSN HPF (SEQ ID NO: 215)。

[0121] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白不需要多聚化以用于RNA结合活性。在一些实施方案中,所述第一RNA结合蛋白不是多聚体复合物的单体。在一些实施方案中,多聚体蛋白复合物不包含所述第一RNA结合蛋白。

[0122] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白与所述RNA分子内的靶序列选择性结合。在一些实施方案中,所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的亲合力。在一些实施方案中,所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的高亲合力或不选择性结合所述第二序列。

[0123] 在本公开文本的组合物的一些实施方案中,RNA基因组或RNA转录组包含所述RNA分子。

[0124] 在本公开文本的组合物的一些实施方案中,所述第一RNA结合蛋白包含在2个与

1300个之间的氨基酸,包括端点。

[0125] 在本公开文本的组合物的一些实施方案中,编码所述第一RNA结合蛋白的序列还包含核定位信号(NLS)。在一些实施方案中,编码核定位信号(NLS)的所述序列定位于编码所述第一RNA结合蛋白的序列的3'。在一些实施方案中,所述第一RNA结合蛋白包含在所述蛋白质的C末端的NLS。

[0126] 在本公开文本的组合物的一些实施方案中,编码所述第一RNA结合蛋白的序列还包含编码第一NLS的第一序列和编码第二NLS的第二序列。在一些实施方案中,编码所述第一NLS或所述第二NLS的序列定位于编码所述第一RNA结合蛋白的序列的3'。在一些实施方案中,所述第一RNA结合蛋白包含在所述蛋白质的C末端的所述第一NLS或所述第二NLS。

[0127] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含核酸酶结构域或由其组成。在一些实施方案中,所述第二RNA结合蛋白以与RNA缔合的方式结合RNA。在一些实施方案中,所述第二RNA结合蛋白以切割RNA的方式与RNA缔合。

[0128] 在本公开文本的组合物的一些实施方案中,编码所述第二RNA结合蛋白的序列包含RNA酶或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶1多肽或由其组成。在一些实施方案中,RNA酶1多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGLCKPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNCKYKSNSSMHIITDCRLTNGSRYPNCAYRTSPKERHIIIVACEGSPYVPVHFDASVEDST (SEQ ID NO:20)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶4多肽或由其组成。在一些实施方案中,RNA酶4多肽包含以下序列或由其组成:

QDGMVQRFLRQHVHPEETGGSDRYCDLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICSTTNIQCKNGKMNCHGVVVKVTDCTDGTSSRAPNCRYRAIASTRRVVIACEGNPQVPVHFDG (SEQ ID NO:21)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶6多肽或由其组成。在一些实施方案中,RNA酶6多肽包含以下序列或由其组成:

WPKRLTKAHWFEIQHIQPSPLQCNRAMSGINNYTQHCKHQNTFLHDSFQ

NVAAVCDLLSIVCKNRRHNCHQSSKPVNMTDCRLTSGKYPQCRYSAQAQYKFFIVACDPPQKSDPPYKLVPVHLDL (SEQ ID NO:22)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶7多肽或由其组成。在一些实施方案中,RNA酶7多肽包含以下序列或由其组成:

APARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINKHTKRCKDLNTFLHEPFSSVAATCQTPKIAACKNGDKNCHQSHGPVSLTMCKLTSGKYPNCRYKEKRQNKSYVACKPPQKKDSQQFHLVPVHLDRVL (SEQ ID NO:23)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶8多肽或由其组成。在一些实施方案中,RNA酶8多肽包含以下序列或由其组成:

TSSQWFKTQHVQPSQACNSAMSIINKYTERCKDLNTFLHEPFSSVAITCQTPNIAACKNSCKNCHQSHGPMSLTMGELTSGKYPNCRYKEKHLNTPYIVACDPPQQGDPGYPLVPVHLDKVV (SEQ ID NO:24)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶2多肽或由其组成。在一些实施方案中,RNA酶2多肽包含以下序列或由其组成:

KPPQFTWAQWFETQHINMTSQQCTNAMQVINNYQRRCKNQNTFLLTTFANVVNVCNPNMTCPNKTCKNCHHSGSQVPLIHCNLTTPSPQNISNCRYAQTANMFYIVACDNRDQRRDPPQYPVVPVHLDRII (SEQ ID NO:25)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶6多肽或由其组成。在一些实施方案中,RNA酶6多肽包含以下序列或由其组成:

DKRLRDNHEWKKLIMVQHPETVCEKIQNDCRDPPDYWTIHGLWPKSEGCNRSWPFNLEEIKKNWME
ITDSSLPSMGPAPPRWMRSTPRRSTLAEAWNSTGSWTSTGGCALPPAALPSGDLCCRPSLTAGSRGVGVDLTAL
HQLLHVHYSATGIIPEECSEPTKPFQIILHHDTEWVQSIGMPIWGTISSSESAIGKNEESQPACAVLSHDS (SEQ
ID NO:26)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶L多肽或由其组成。在一些
实施方案中,RNA酶L多肽包含以下序列或由其组成:

AAVEDNHLLIKAVQNEDVDLVQQLLEGGANVNFQEEEGWTPHNAVQMSREDIVELLRLRHGADPVLK
KKNATPFIILAAIAGSVKdLLKLFLSKGADVNECDFYGFATFMEAAYGKVKALKFLYKRGANVNLRRKTKEDQER
LRKGGATALMDAAEKGHVEVLKILLDEMGADVACDNMGRNALIHALLSSDDSDVEAITHLLLDHGADVNRGERG
KTPLILAVEKKHLGLVQRLLQEHEIEINDTSDGKTALLLAVELKLKKIAELLCKRGASTDCGDLVMTARRNYDHS
LVKVLSSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLKFFIDEKYKIADTSEGGIYLGfyEKQEVAV
KTFCEGSPRAQREVSCQLSSRENSHLVTFYGSSEHRGHLFVCVTLCQTEACLDVHRGEDVENEDEFARNVLSS
IFKAVQELHLSCGYTHQDLQPQNILIDSKKAAHLADFDKSIKWAGDPQEVKRDLEDLGRVLVYVVKGSISFEDLK
AQSNEEVVQLSPDEETKDLIHRLFHPGEHVRDCLSDLLGHPFFWTWESRYRTLNRVGNESDIKTRKSESEILRLQ
PGPSEHSKSFDKWTTKINCEVMKKMNKFYEKRGNFYQNTVGDLLKFIRNLGEHIDEKHKMKLKIIGDPSLYFQKT
FPDLVIYVYTKLQNTYRKHFPQTHSPNKPQCDGAGGASGLASPGC (SEQ ID NO:27)。在一些实施方案
中,所述第二RNA结合蛋白包含RNA酶T2多肽或由其组成。在一些实施方案中,RNA酶T2多肽
包含以下序列或由其组成:

VQHPETVCEKIQNDCRDPPDYWTIHGLWPKSEGCNRSWPFNLEEIKDLLPEMRAYWPDVHSFPNR
SRFWKHEWEKHGTCAAQVDALNSQKKYFGRSLELYRELDLNSVLLKLGIKPSINYYQVADFKDALARYGVIPKIQ
CLPPSQDEEVQTIGQIELCLTKQDQQLQNCTEPGEQPSPKQEVWLANGAAESRGLRVCEGDPVFYPPPKTKH (SEQ
ID NO:28)。在一些实施方案中,所述第二RNA结合蛋白包含RNA酶11多肽或由其组成。在一
些实施方案中,RNA酶11多肽包含以下序列或由其组成:

EASESTMKIKEEFTDEEMQYDMAKSGQEKQTIEILMNPILLVKNTSLSMSKDDMSSTLLTFRSLHYN
DPKGNSSGNDKECCNDMTVWRKVSEANGSCKWSNNFIRSSTEVMRRVHRAPSCKFVQNPGISCCESLELENTVCQF
TTGKQFPRCQYHSVTSLEKILTVLTGHSLMSWLVCGLSKL (SEQ ID NO:29)。在一些实施方案中,所述第
二RNA结合蛋白包含RNA酶T2样多肽或由其组成。在一些实施方案中,RNA酶T2样多肽包含以
下序列或由其组成:

XLGGADKRLRDNHEWKKLIMVQHPETVCEKIQNDCRDPPDYWTIHGLWPKSEGCNRSWPFNLEEIK
DLLPEMRAYWPDVHSFPNRSRFBKHEWEKHGTCAAQVDALNSQKKYFGRSLELYRELDLNSVLLKLGIKPSINYY
QTTEEDLNLDEPTTEDTAEVTHVLLHSALFGEIGPRRW (SEQ ID NO:30)。

[0129] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含突变的RNA
酶或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1
(K41R)) 多肽或由其组成。在一些实施方案中,RNA酶1 (K41R) 多肽包含以下序列或由其组
成:KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNCY
KSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYVPVHFDASVEDST (SEQ ID NO:116)。在一
些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R、D121E)) 多肽或由其
组成。在一些实施方案中,RNA酶1 (RNA酶1 (K41R、D121E)) 包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHIIVACEGSPYVPVHFEASVEDST (SEQ ID NO:117)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(K41R、D121E、H119N))多肽或由其组成。在一些实施方案中,RNA酶1(RNA酶1(K41R、D121E、H119N))多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCRPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHIIIVACEGSPYVPVNFEASVEDST (SEQ ID NO:118)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(H119N))多肽或由其组成。在一些实施方案中,RNA酶1(RNA酶1(H119N))多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHIIIVACEGSPYVPVNFASVEDST (SEQ ID NO:119)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCKPVNTFVHEPLVDVQNVCFQEKVTCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHIIIVACEGSPYVPVNFASVEDST (SEQ ID NO:120)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E))多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDGCRPVNTFVHEPLVDVQNVCFQEKVTCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHIIIVACEGSPYVPVNFEASVEDST (SEQ ID NO:121)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N))多肽或由其组成。在一些实施方案中,RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D))多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDGCRPVNTFVHEPLVDVQNVCFQEKVTCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHIIIVACEGSPYVPVHFDASVEDST (SEQ ID NO:122)。

在一些实施方案中,所述第二RNA结合蛋白包含突变的RNA酶1(RNA酶1(R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E))多肽或由其组成,所述多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDGCRPVNTFVHEPLVDVQNVCFQEKVTCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHIIIVACEGSPYVPVNFEASVEDST (SEQ ID NO:225)。

[0130] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含NOB1多肽或由其组成。在一些实施方案中,NOB1多肽包含以下序列或由其组成:

APVEHVVDAGAFRLRHAALQDIGKNIYTIREVVTEIRDKATRRRLAVLPYELRFKEPLPEYVRLVTEF
SKKTGDYPSLSATDIQVLALTYQLEAEFVGVSHLKQEPQKVKVSSSIQHPETPLHISGFHLPYKPKPPQETEKGHS
ACEPENLEFSSFMFWRNPLPNIDHELQELLIDRGEDVPSEEEEEENGFEEDRKDDSDDDGGGWITPSNIKQIQQEL
EQCDVPEDVRVGCLTTDFAMQNVLLQMGLHVLAVNGMLIREARSYILRCHGCFKTTSDMSRVFCSHCNKTLLKVS
VTV (SEQ ID NO:31)。

[0131] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含内切核酸酶或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含内切核酸酶V(ENDOV)或由其

组成。在一些实施方案中,ENDOV多肽包含以下序列或由其组成:

AFSGLQRVGGVDVSVFKGDSVRACASLVLSFPELEVVEESRMVSLTAPYVSGFLAFREVPFLLELV
QQLREKEPGLMPQVLLVDGNGVLHHRGFGVACHLGLVLTDLPCVGVAKKLLQVDGLENNALHKEKIRLLQTRGDSFP
LLGDSGTVLGMALRSHDRSTRPLYISVGHMSLEAAVRLTCCCCRFRIPEPVRQADICSREHIRKS (SEQ ID NO:
32)。在一些实施方案中,所述第二RNA结合蛋白包含内切核酸酶G (ENDO G) 多肽或由其组成。
在一些实施方案中,ENDO G多肽包含以下序列或由其组成:

AELPPVPGGPRGPGLAKYGLPGLAQLKSRESYVLCYDPRTRGALWVVEQLRPERLRGDGDRRECDFR
EDDSVHAYHRATNADYRGSGFDRGHLAAAANHRWSQKAMDDTFYLSNVAPQVPHLNQNAWNNLEKYSRSLTRSYQN
VYVCTGPLFLPRTEADGKSYVKYQVIGKNHVAVPTHFFKVLILEAAGGQIELRTYVMPNAPVDEAIPLERFLVPIE
SIERASGLLFVPNILARAGSLKAITAGSK (SEQ ID NO:33)。在一些实施方案中,所述第二RNA结合
蛋白包含内切核酸酶D1 (ENDOD1) 多肽或由其组成。在一些实施方案中,ENDOD1多肽包含以
下序列或由其组成:

RLVGEEEAGFGCEDKFFYAGTPPAGLAADSHVKICQRAEGAERFATLYSTRDRIPVYSAFRAPRPAPG
GAEQRWLVEPQIDDPNSNLEEAINAEAITSVNSLGSKQALNTDYLDSDYQRGQLYPFSLSSDVQVATFTLTNSAP
MTQSFQERWYVNLHSLMDRALTPQCGSGEDLYILTGTVPDSDYRVKDKVAVPEFVWLAACCAVPGGGWAMGFVKHTR
DSDI IEDVMVKDLQKLLPFNPQLFQNNCGETE QDTEKMKKILEVVNQIQDEERMVQSQKSSSPLSSTRSKRSTLLP
PEASEGSSSFLGKLMGFIATPFIKLFQLIYYLVVAILKNIVYFLWCVTQKQVINGIESCLYRLGSATISYFMAIGEE
LVSIPWKVLKVVAKVIRALLRILCCLLKAICRVLSIPVRVLVDVATFPVYTMGAIPVCKDIALGLGGTVSLLFDT
AFGTLGGLFQVVFVSVCKRIGYKVTFDNSGEL (SEQ ID NO:34)。在一些实施方案中,所述第二RNA结
合蛋白包含人瓣状内切核酸酶-1 (hFEN1) 多肽或由其组成。在一些实施方案中,hFEN1多肽
包含以下序列或由其组成:

MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLI AVRQGGDVLQNEEGETTSHLMGMF
YRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDCKHL
LSLMGIPYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQELGLNQ
EQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLD PESVEL
KWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGA
AGKFGRGK (SEQ ID NO:35)。在一些实施方案中,所述第二RNA结合蛋白包含DNA修复内切核
酸酶XPF (ERCC4) 多肽或由其组成。在一些实施方案中,ERCC4多肽包含以下序列或由其组
成:

MESGQPARRIAMAPLLEYERQLVLELLD TDGLVVCARGLGADRLLYHFLQLHCHPACLVLVLNTQPAE
EEYFINQLKIEGVEHLPRRV TNEITSNSRYEVYTQGGVIFATSRILVVDFLTDRIPSDLITGILVYRAHRIIESCQ
EAFILRLFRQKNKRGFIKAFTDNAFAFDTG FCHVERVMRNLFVRKLYLWPRFHVAVNSFLEQHKPEVVEIHVSMT P
TMLAIQTAILDILNACLKELKCHNPSLEVEDLSLENAIGKPFDKTIRHYLDPLWHQLGAKTKSLVQDLKILRTL LQ
YLSQYDCVTFLNLLLESLRATEKAFGQNSGWFLDSSTSMFINARARVYHLPDAKMSKKEKISEKMEIKEGEGILWG
(SEQ ID NO:124)。

[0132] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含内切核酸
酶III样蛋白1 (NTHL) 多肽或由其组成。在一些实施方案中,NTHL多肽包含以下序列或由其
组成:

CSPQESGMTALSARMLTRSRSLGPGAGPRGCREEPGLRRREAAAEARKSHSPVKRPRKAQRLRVAYE

GSDSEKGEAEPLKVPVWEPQDWQQQLVNIRAMRNKKDAPVDHLGTEHCYDSSAPPKVRRYQVLLSLMLSSQTKDQ
VTAGAMQRLRARGLTVDLSILQTDDATLGKLIYPVGFWRSKVKYIKQTSAILQQHYGGDIPASVAELVALPGVGP
KMAHLAMAVAWGTVSGIAVDTHVHRIANRLRWTKKATKSPEETRAALEEWLPRELWHEINGLLVGFGQQTCLPVHPRC
HACLNQALCPAAQGL (SEQ ID NO:123)。

[0133] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含人斯库拉
芬蛋白14 (hSLFN14) 多肽或由其组成。在一些实施方案中,hSLFN14多肽包含以下序列或由
其组成:

ESTHVEFKRFTTKKVIPRIKEMLPHYVSFAFANTQGGYVLIGVDDKSKEVVGCKWEKVNPDLLKKEIEN
CIEKLPTFHFCCEKPKVNFTTKILNVYQKDVLGDYVCVIQVEPFCCVVFAPDSWIMKDNSVTRLTAEQWVVMML
DTQSAPPSLVTDYNSCLISSASSARKSPGYPIKVHKFKEALQ (SEQ ID NO:36)。

[0134] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含人 β -内酰
胺酶样蛋白2 (hLACTB2) 多肽或由其组成。在一些实施方案中,hLACTB2多肽包含以下序列或
由其组成:

TLQGTNTYLVTGTPRRILIDTGEPAPIEYISCLKQALTEFNATIQEIVVTHWHRDHSGGIGDICKSIN
NDTTYCIKKLPRNPQREEIIGNGEQQYVYLKGDVVIKTEGATLRVLYTPGHTDDHMLLLEENAI FSGDCILGEG
TTVFEDLYDYMNSLKELLKIKADIIYPGHGPVIHNAEAKIQQYISHRNIREQQILTLFRENFEKSFTVMELVKIIY
KNTPENLHEMAKHNNLLHLKKLEKEGKIFSNTDPDKKWK AHL (SEQ ID NO:37)。

[0135] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含无嘌呤/无
嘧啶 (AP) 内切脱氧核糖核酸酶 (APEX) 多肽或由其组成。在一些实施方案中,所述第二RNA结
合蛋白包含无嘌呤/无嘧啶 (AP) 内切脱氧核糖核酸酶 (APEX2) 多肽或由其组成。在一些实施
方案中,APEX2多肽包含以下序列或由其组成:

MLRVVSWNINGIRRPLQGVANQEPSNCAAVAVGRILDELADIVCLQETKVTRDALTEPLAIVEGYNS
YFSFSRNRSGYSGVATFCKDNATPVAAEEGLSGLFATQNGDVGCYGNMDEFTQEELRALDSEGRALLTQHKIRTWE
GKEKTLTLINVYCPHADPGRPERLVFKMRFYRLLQIRAEALLAAGSHV IILGDLNTAHRPIDHWDVNLECFEEDP
GRKWMDSLLSNLGCQSASHVGPFI DSYRCFQPKQEGAF TCWSAVTGARHLNYGSRLDYVLGDRTLVIDTFQASFL
PEVMGSDHCPVGAVLSVSSVPAKQCPPLCTRFLPEFAGTQLKILRFLVPLEQSPVLEQSTLQHNNQTRVQTCQNK
QVRSTRPQPSQVGSSRGQKNLKSYPQSPSCPQASPDIELPSLPLMSALMTPKTPEEKAVAKVVKGAKTSEAKDE
KELRTSFWKSVLGAPLRTPLCGHREPCVMRTVKKPGPNLGRRFYMCARPRGPPTDPSSRCNFFLWSRPS (SEQ ID
NO:38)。在一些实施方案中,APEX2多肽包含以下序列或由其组成:

MLRVVSWNINGIRRPLQGVANQEPSNCAAVAVGRILDELADIVCLQETKVTRDALTEPLAIVEGYNS
YFSFSRNRSGYSGVATFCKDNATPVAAEEGLSGLFATQNGDVGCYGNMDEFTQEELRALDSEGRALLTQHKIRTWE
GKEKTLTLINVYCPHADPGRPERLVFKMRFYRLLQIRAEALLAAGSHV IILGDLNTAHRPIDHWDVNLECFEEDP
GRKWMDSLLSNLGCQSASHVGPFI DSYRCFQPKQEGAF TCWSAVTGARHLNYGSRLDYVLGDRTLVIDTFQASFL
PEVMGSDHCPVGAVLSVSSVPAKQCPPLCTRFLPEFAGTQLKILRFLVPLEQSP (SEQ ID NO:39)。在一些实
施方案中,所述第二RNA结合蛋白包含无嘌呤或无嘧啶位点裂解酶 (APEX1) 多肽或由其组
成。在一些实施方案中,APEX1多肽包含以下序列或由其组成:

PKRGKKGAVAEDGDEL RTEPEAKKSKTAAKNDKEAAGEGPALYEDPPDQKTSPSGKPATLKICSWNV
DGLRAWIKKKGLDWVKEEAPDILCLQETKCSENKLPAELQELPGLSHQYWSAPSDKEGYSGVGLLSRQCPLKVSYG
IGDEEHDQEGRVIVAEFDSFVLVTAYVPNAGRGLVRLEYRQRWDEAFRKFLKGLASRKPLVLCGDLNVAHEEIDLR

NPKGNNKKNAGFTPQERQGFCELLQAVPLADSRHLYPNTPTYAYTFWTYMMNARSKNVGWRLDYFLLS (SEQ ID NO:125)。

[0136] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含血管生成素(ANG)多肽或由其组成。在一些实施方案中,ANG多肽包含以下序列或由其组成:

QDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHGKRSIKAICENKNGNPHREN LRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVACENGLPVHLDQSIFRRP (SEQ ID NO:40)。

[0137] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含热反应蛋白12(HRSP12)多肽或由其组成。在一些实施方案中,HRSP12多肽包含以下序列或由其组成:

SSLIRRVISTAKAPGAIGPYSQAVLVDRTIYISGQIGMDPSSGQLVSGGVAEEAKQALKNMGEILKAA GCDFTNVVKTTVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGSRIEIEAVAIQGPLTTASL (SEQ ID NO:41)。

[0138] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含含锌指CCCH型12A(ZC3H12A)多肽或由其组成。在一些实施方案中,ZC3H12A多肽包含以下序列或由其组成:

GGGTPKAPNLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFLERGHDTITVFV PSWRKEQPRPDVPI TDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIVKLAYESDGIVVSN DTYRDLQGERQE WKRFIEERLLMYSFVNDKFMPPDDPLGRHGPSLDNFLRKKPLTLE (SEQ ID NO:42)。

在一些实施方案中,ZC3H12A多肽包含以下序列或由其组成:

SGPCGEKPVLEASPTMSLWFE DSHSRQGT PRPGQELAAEEASALELQMKVDFFRKL GYSSTEIHSVL QKLG VQADTNTVLGELVKHGTATERERQTSPDPCQLPLVPRGGGTPKAPNLEPPLPEEEKEGSDLRPVVIDGSNV AMSHGNKEVFSCRGILLAVNWFLERGHDTITVFVPSWRKEQPRPDVPI TDQHILRELEKKKILVFTPSRRVGGKRV VCYDDRFIVKLAYESDGIVVSN DTYRDLQGERQE WKRFIEERLLMYSFVNDKFMPPDDPLGRHGPSLDNFLRKKPL TLEHRKQPCPYGRKCTYG I KCRFFHPERPSCPQRSVADEL RANALLSPPRAPSKDKNGRRPSPSSQSSSLLTESEQ CSLDGKKLGAQASPGSRQEGLTQTYAPSGRSLAPSGGSGSSFGPTDWLPQTLD SLPYVSQDCLDSGIGSLESQMSE LWGVRGGGPGEPPPRAPYTGYS PYGSELPATAAFSAFGRAMGAGHFSVPADYPPAPPAFPFPREYWSEPYPLPPPT SVLQEPPVQSPGAGRSPWGRAGSLAKEQASVYTKLCGVFP PHLVEAVMGRFPQLLDPQQLAAEILSYKSQHPSE (SEQ ID NO:43)。

[0139] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含反应性中间亚胺脱氨酶A(RIDA)多肽或由其组成。在一些实施方案中,RIDA多肽包含以下序列或由其组成:

SSLIRRVISTAKAPGAIGPYSQAVLVDRTIYISGQIGMDPSSGQLVSGGVAEEAKQALKNMGEILKAA GCDFTNVVKTTVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGSRIEIEAVAIQGPLTTASL (SEQ ID NO:44)。

[0140] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含磷脂酶D家族成员6(PDL6)多肽或由其组成。在一些实施方案中,PDL6多肽包含以下序列或由其组成:

EALFFPSQVTCTEALLRAPGAELAE LPEGCPGLPHGESALSRLLRALLAARASLDLCLFAFSSPQLG RAVQLLHQRGVRVRVVTDCDYMALNGSQIGLLRKAGIQVRHDQDPGYMHKFAIVDKRVLITGSLNWTQAIQNNR ENVLITEDDEYVRLFLEEFERIWEQFNPTKYTFFPKKSHGSCAPPVSRAGGRLLSWHRTC GTSSSESQT (SEQ ID NO:126)。

[0141] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含线粒体核糖核酸酶P催化亚基(KIAA0391)多肽或由其组成。在一些实施方案中,KIAA0391多肽包含以下序列或由其组成:

KARYKTLEPRGYSLIRGLIHSDRWREALLLLEDIKKVITPSKKNYNCIQGALLHQDVNTAWNLYQE
LLGHDIVPMLETLKAFFDFGKDIKDDNYSNKLDDILSYLRNNQLYPGESFAHSIKTWFEVPGKQWKQGF₂TVRKS
GQCSGCGKTIESIQLSPEEYECLKGKIMRDVIDGGDQYRKTTPELKRFEF₂IKSRPPFDVVIDGLNVAKMFPKVR
ESQLLLNVVSQ₂LAKRNLRLVLGRKHMLRRSSQWSRDEMEEVQKQASCF₂FADDISEDDPFLLYATLHSGNHCRFIT
RDLMRDHKACLPDAKTQRLFFKWQQGHQLAIVNRFPGSKLTFQRILSYDTVVQTTGDSWHIPYDEDLVERCSCEVP
TKWLCLHQKT (SEQ ID NO:127)。

[0142] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含阿尔古蛋白2(AG02)多肽或由其组成。

在本公开文本的组合物的一些实施方案中,AG02多肽包含以下序列或由其组成:

SVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTVLMATQCVQMKNVQRTTPQTLSNLCLKI
NVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDAHPNRYCATVRVQQHRQEIIQDLA
AMVRELLIQFYKSTRFKPTRIIFYRDGVSEGGFQQVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTD
KNERVGKSGNIPAGTTVDTKITHPTEFDYLCSHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTR
SVSIPAPAYYAHLVAFRARYHLVDKEHDSAEGSHTSGQSNRQDHQALAKAVQVHQDTLRTMYFA (SEQ ID NO:
128)。

[0143] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含线粒体核糖核酸酶EXOG(EXOG)多肽或由其组成。在一些实施方案中,EXOG多肽包含以下序列或由其组成:

QGAEGALTGKQPDGSAEKAVLEQFGFPLTGTEARCYTNHALSYDQAKRVPRWVLEHISKSKIMGDADR
KHCKFKPDPNIPPTFSAFNEDYVGSWSRGHMAPAGNNKFSSKAMAETFYLSNIVPQDFDNNSGYWNRIEMYCREL
TERFEDVWVVSGLPLTPQTRGDGKKIVSYQVIGEDNVAVPSHLYKVILARRSSVSTEPLALGAFVVPNEAIGFQ₂Q
LTEFQVSLQDLEKLSGLVFFPHLDRTSDIRNICSVDTC₂LLDFQEFTLYLSTRKIEGARSVLRLEKIMENLKNAEI
EPDDYFMSRYEKKLEELKAKEQSGTQIRKPS (SEQ ID NO:129)。

[0144] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含含锌指CCCH型12D(ZC3H12D)多肽或由其组成。在一些实施方案中,ZC3H12D多肽包含以下序列或由其组成:

EHPSKMEFFQKLGYDREDVLRVLGKLGE₂ALVNDVLQELIRTGSRPGALEHPAAPRLVPRGSCGVPDS
AQRGPGTALEEDFRTLASSLRPIVIDGSNVAMSHGNKETFSCRGIKLAVDWFRDRGHTYIKVFVPSWRKDP₂PRADT
PIREQHVLAELERQAVLVYTPSRKVHGKRLVCYDDRYIVKVAYEQDGVIVSNDNYRDLQSENPEWKWFIEQRLLMF
SFVNDRFMPDDPLGRHG₂PSLSNFLSRKPKPEPSWQHCPYGGKCTYGIKCKFYHPERPHHAQLAVADELRAKTGA
RPGAGAEERPPRAPGGSAGARAAPREPF₂AHSLPPARGSPDLAALRGFSRLAFSDDLGLPPLPVPACSLTPRL
GGPDWVSAGGRVPGPLSLPSPESQFSPGDLPPPPGLQLQPRGEHRPRDLHGDLLSPRRPPDDPWARPPRSDFPGR
SVWAEPAWGDGATGGLSVYATEDDEGDARARARIALYSVFPRDQVDRVMAAFPELSDLARLILLVQRCQSAGAPLG
KP (SEQ ID NO:130)。

[0145] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含内质网核信号转导蛋白2(ERN2)多肽或由其组成。在一些实施方案中,ERN2多肽包含以下序列或由其组成:

RQQQPQVVEKQQETPLAPADFAHISQDAQSLHSGASRRSQKRLQSPSKQAQPLDDPEAEQLTVVGKIS
FNPKDVLGRGAGGTFVFRGQFEGRAVAVKRLRECFGLVRREVQLLQESDRHPNVLRYFCTERGPQFHYLELCR
ASLQEYVENPDLDRGGLEPEVVLQQLMSGLAHLHSLHIVHRDLKPGNILITGPDSSQLGRVVLSDFGCLKLPAGR
CSFSLHSGIPGTEGWMAPELLQLLPPDSPTS AVDIFSAGCVFYVLSGGSHPFGDSLYRQANILTGAPCLAHLEEE
VHDKVVARDLVGAMLSPLPQPRPSAPQVLAHPFFWSRAKQLQFFQDVS DWLEKESEQEPLVRALEAGGCAVVRDNW
HEHISMP LQTDLRKFRSYKGT SVRDLLRAVRNKKHHYREL PVEVRQALGQVPDGFVQYFTNRFPRLLLHTHRAMRS
CASESFLPYPPDSEARRPCPGATGR (SEQ ID NO:131)。

[0146] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含回力球 mRNA 监督和核糖体挽救因子 (PELO) 多肽或由其组成。在一些实施方案中,PELO 多肽包含以下序列或由其组成:

KLVRKNIEKDNAGQVTLVPEEPEDMWHTYNLVQVGDSLRASTIRKVQTESSTGSGVSNRVRTTLTLCV
EAIDFDSQACQLRVKGTNIQENEYVKMGAYHTIELEPNRQFTLAKKQWDSVVLERIEQACDPAWSADVAAVVMQEG
LAHICLVTPSMTLTRAKEVNIPRKRKGNC SQH DRALERFYEQVVQAIQRHIHF DVVKC ILVASPGFVREQFCDYL
FQQAVKTDNKLLENRSKFLQVHASSGHKYS LKEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRAFYGLK
QVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSSLHVS GEQLS QLTGVAAILRFPVPEL
SDQEGDSSSEED (SEQ ID NO:132)。

[0147] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含YBEY金属肽酶 (YBEY) 多肽或由其组成。在一些实施方案中,YBEY 多肽包含以下序列或由其组成:

SLVIRNLQRVIPIRRAPLRSKIEIVRRILGVQKFDLGIICVDNKNIQHINRIYDRNVPTDVLSPFPH
EHLKAGEFPQPDFDDYNLGDIFLGVEYIFHQCKENEDYNDVLTVTATHGLCHLLGFTHGTEAEWQQMFQKEKAVL
DELGRRTGTRLQPLTRGLFGGS (SEQ ID NO:133)。

[0148] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含剪切和多聚腺苷酸化特异性因子4样蛋白 (CPSF4L) 多肽或由其组成。在一些实施方案中,CPSF4L 包含以下序列或由其组成:

QEVIAGLERFTFAFEKDVEMQKGTGLLPFQGMDSASAVCNFFTKGLCEKGKLCPRHRDRGEKMMVCK
HWLRGLCKKGDHCKFLHQYDLTRMPECFYFSKFGDCSNKECSFLHVKPAFKSQDCPWYDQGFCKDGPLCKYRHVPR
IMCLNYLVGFCPEGPKCQFAQKIREFKLLPGSKI (SEQ ID NO:134)。

[0149] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含hCG_2002731多肽或由其组成。在一些实施方案中,hCG_2002731多肽包含以下序列或由其组成:

KLVRKNIEKDNAGQVTLVPEEPEDMWHTYNLVQVGDSLRASTIRKVQTESSTGSGVSNRVRTTLTLCV
EAIDFDSQACQLRVKGTNIQENEYVKMGAYHTIELEPNRQFTLAKKQWDSVVLERIEQACDPAWSADVAAVVMQEG
LAHICLVTPSMTLTRAKEVNIPRKRKGNC SQH DRALERFYEQVVQAIQRHIHF DVVKC ILVASPGFVREQFCDYM
FQQAVKTDNKLLENRSKFLQVHASSGHKYS LKEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRAFYGLK
QVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSSLHVS GEQLS QLTGVAAILRFPVPEL
SDQEGDSSSEED (SEQ ID NO:135)。在一些实施方案中,hCG_2002731多肽包含以下序列或由其组成:

DPAWSADVAAVVMQEGLAHICLVTPSMTLTRAKEVNIPRKRKGNC SQH DRALERFYEQVVQAIQRHI
HFDVVKC ILVASPGFVREQFCDYMFQQAVKTDNKLLENRSKFLQVHASSGHKYS LKEALCDPTVASRLSDTKAAG
EVKALDDFYKMLQHEPDRAFYGLKQVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSL

HVSGEQLSQLTGVAAILRFPVPELSDQEGDSSSEED (SEQ ID NO:136)。

[0150] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含切除修复交叉互补组1 (ERCC1) 多肽或由其组成。在一些实施方案中,ERCC1多肽包含以下序列或由其组成:

MDPGKDKEGVPQPSGPPARKKFVIPLDEDEVPPGVRGNPVLKFVRNPWEFGDVIPDYVLGQSTCALF
LSLRYHNLHPDYIHGRLQSLGKNFALRVLLVQVDVKDPQQALKELAKMCILADCTLILAWSPEEAGRYLETYKAYE
QKPADLLMEKLEQDFVSRVTECLTTVKSVNKTDSTLLTTFGSLEQLIAASREDLALCPGLGPQK (SEQ ID NO:
137)。

[0151] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含ras相关C3肉毒杆菌毒素底物1亚型 (RAC1) 多肽或由其组成。在一些实施方案中,RAC1多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNVCFQEKVCTKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHIIIVACEGSPYVPVHFDASVEDST (SEQ ID NO:138)。

[0152] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含核糖核酸酶A A1 (RAA1) 多肽或由其组成。在一些实施方案中,RAA1多肽包含以下序列或由其组成:

QDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHGNKRSIKAICENKNGNPHREN
LRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVVACENGLPVHLDQSIFRRP (SEQ ID NO:139)。

[0153] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含Ras相关蛋白 (RAB1) 多肽或由其组成。在一些实施方案中,RAB1多肽包含以下序列或由其组成:

GLGLVQPSYGQDGMVQRFLRQHVPHEETGGSDRYCNLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICS
TTNIQCKNGKMNCHEGVVKVTDRCRTGSSRAPNCRYRAIASTRRVIIACEGNPQVPVHFDG (SEQ ID NO:140)。

[0154] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含DNA复制解旋酶/核酸酶2 (DNA2) 多肽或由其组成。在一些实施方案中,DNA2多肽包含以下序列或由其组成:

XSAVDNILLKLAKFKIGFLRLGQIQKVHPAIQQFTEQEICRSKSIKSLALLEELYNSQLIVATTCMGI
NHPIFSRKIFDFCIVDEASQISQPICLGPLFFSRRFVLVGDHQQLPPLVLNREARALGMSESLFKRLEQNKSAVVQ
LTVQYRMNSKIMSLSNKLTYEGKLECGSDKVANAVINLRHFKDVKLELEFYADYSDNPWLMGVFEPNNPVCFLNTD
KVPAPQVEKGGVSNVTEAKLIVFLTSIFVKAGCSPSDIGIIAPYRQQLKIINDLLARSIGMVEVNTVDKYQGRDK
SIVLVSFVRSNKDGTVGELLKDWRRRLNVAITRAKHKLILLGCVPSLNCYPPLEKLLNHLNSEKLISFFFCIWSHLI
ALL (SEQ ID NO:141)。

[0155] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含FLJ35220多肽或由其组成。在一些实施方案中,FLJ35220多肽包含以下序列或由其组成:

MALRSHDRSTRPLYISVGHMSLEAAVRLTCCCFRIFPEPVRQADICSREHIRKSLGLPGPPTPRSP
KAQRPVACPKGDSGESSALC (SEQ ID NO:142)。

[0156] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含FLJ13173多肽或由其组成。在一些实施方案中,FLJ13173多肽包含以下序列或由其组成:

CYTNHALSYDQAKRVPRWVLEHISKSKIMGDADRKHKCFKPDNPPTFSAFNEDYVGSWSRGHMAP
AGNNKFSSKAMAETFYLSNIVPQDFDNNSGYWNRIEMYCRELTERFEDVWVVSGLPLTLPQTRGDGKKIVSYQVIGE
DNVAVPSHLYKVLARRSSVSTEPLALGAFVVPNEAIGFQPQLTEFQVSLQDLEKLSGLVFFPHLDRT (SEQ ID

NO:143)。

[0157] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含特诺伊林跨膜蛋白(TENM)多肽或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含特诺伊林跨膜蛋白1(TENM1)多肽或由其组成。在一些实施方案中,TENM1多肽包含以下序列或由其组成:

VTVSQMTSVLNGKTRRFADIQLQHGCALCFNIRYGTTVVEEKNHVLEIARQRAVAQAWTKEQRRLQEGE
EGIRAWTEGEKQQLLSTGRVQGYDGYFVLSVEQYLELSDSANNIHFMRQSEIGRR (SEQ ID NO:144)。在一些
实施方案中,所述第二RNA结合蛋白包含特诺伊林跨膜蛋白2(TENM2)多肽或由其组成。

在一些实施方案中,TENM2多肽包含以下序列或由其组成:

TVSQPTLLVNGKTRRFTNIEFQYSTLLLSIRYGLTPDTLDEEKARVLDQARQRALGTAWAKEQQKARD
GREGSRLWTEGEKQQLLSTGRVQGYEGYYVLPVEQYPELADSSSNIQFLRQNMKGK (SEQ ID NO:145)。

在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含核糖核酸酶_K(RNA酶K)多肽或由其组成。在一些实施方案中,RNA酶K多肽包含以下序列或由其组成:

MGWLRPGPRPLCPPARASWAFSHRFPSPAPRRSPTPFFMASLLCCGPKLAACGIVLSAWGVIMLIML
GIFNVHSAVLIEDVPFTEKDFENGPNQNIYNYEVSYNCFIAAGLYLLLGGFSFCQVRLNKRKEYMVR (SEQ ID
NO:204)。

在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含转录激活因子样效应物核酸酶(TALEN)多肽或所述多肽的核酸酶结构域或由其组成。

在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含锌指核酸酶多肽或所述多肽的核酸酶结构域或由其组成。在一些实施方案中,所述第二RNA结合蛋白包含ZNF638多肽或所述多肽的核酸酶结构域或由其组成。

[0158] 在本公开文本的组合物的一些实施方案中,所述第二RNA结合蛋白包含从人SMG6蛋白衍生的PIN结构域或由其组成,所述人SMG6蛋白通常也称为端粒酶结合蛋白EST1A亚型3,NCBI参考序列:NP_001243756.1。在一些实施方案中,来自hSMG6的PIN以Cas融合蛋白的形式在本文中使用并且作为内部对照。

指导RNA

[0159] 术语指导RNA(gRNA)与单一指导RNA(sgRNA)在整个本公开文本中可互换使用。

[0160] 本公开文本的指导RNA(gRNA)可以包含间隔子序列和支架序列。在一些实施方案中,指导RNA是单一指导RNA(sgRNA),其包含连续间隔子序列和支架序列。在一些实施方案中,间隔子序列和支架序列是连续的。在一些实施方案中,支架序列包含“同向重复”(DR)序列。DR序列是指CRISPR基因座(天然存在于细菌基因组或质粒中)中的重复性序列,其中散布有间隔子序列。众所周知,如果相关CRISPR基因座的序列是已知的,则将能够推断相应Cas蛋白的DR序列。在一些实施方案中,间隔子序列和支架序列是不连续的。在一些实施方案中,编码本公开文本的指导RNA的序列包含由接头序列隔开的间隔子序列和支架序列或由其组成。在一些实施方案中,接头序列可以包含1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50个或之间的任何数量的核苷酸或由其组成。在一些实施方案中,接头序列可以包含至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50个或之间的任何数量的核苷酸。

[0161] 本公开文本的指导RNA(gRNA)可以包含非天然存在的核苷酸。在一些实施方案中,

本公开文本的指导RNA或编码所述指导RNA的序列包含修饰的或合成的RNA核苷酸或由其组成。示例性的修饰的RNA核苷酸包括但不限于假尿苷(Ψ)、二氢尿苷(D)、肌苷(I)、和7-甲基鸟苷(m7G)、次黄嘌呤、黄嘌呤、黄苷、7-甲基鸟嘌呤、5,6-二氢尿嘧啶、5-甲基胞嘧啶、5-甲基胞苷、5-羟甲基胞嘧啶、异鸟嘌呤和异胞嘧啶。

[0162] 本公开文本的指导RNA(gRNA)可以结合靶序列内的修饰的RNA。在靶序列内,本公开文本的指导RNA(gRNA)可以结合修饰的RNA。示例性的表观遗传或转录后修饰的RNA包括但不限于2'-O-甲基化(2'-OMe)(2'-O-甲基化发生在核糖部分的游离2'-OH的氧上)、N6-甲基腺苷(m6A)和5-甲基胞嘧啶(m5C)。

[0163] 在本公开文本的组合物的一些实施方案中,本公开文本的指导RNA包含编码非编码C/D盒核仁小RNA(snoRNA)序列的至少一个序列。在一些实施方案中,snoRNA序列包含与靶RNA互补的至少一个序列,其中所述RNA分子的靶序列包含至少一个2'-OMe。在一些实施方案中,snoRNA序列包含与靶RNA互补的至少一个序列,其中与靶RNA互补的所述至少一个序列包含盒C基序(RUGAUGA)和盒D基序(CUGA)。

[0164] 本公开文本的间隔子序列与RNA分子的靶序列结合。本公开文本的间隔子序列可以包含CRISPR RNA(crRNA)。本公开文本的间隔子序列包含与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列或由其组成。在与RNA分子的靶序列结合后,间隔子序列可以将支架序列和融合蛋白中的一种或多种指导至所述RNA分子。在一些实施方案中,与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列与所述靶序列具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或之间的任何百分比的同一性。在一些实施方案中,与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列与所述靶序列具有100%同一性。

[0165] 本公开文本的支架序列结合本公开文本的第一RNA结合多肽。本公开文本的支架序列可以包含反式作用RNA(tracrRNA)。本公开文本的支架序列包含与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列或由其组成。在与RNA分子的靶序列结合后,支架序列可以将融合蛋白指导至所述RNA分子。在一些实施方案中,与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列与所述靶序列具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或之间的任何百分比的同一性。在一些实施方案中,与RNA分子的靶序列具有足够互补性以与所述靶序列选择性结合的序列与所述靶序列具有100%同一性。可替代地,或另外地,在一些实施方案中,本公开文本的支架序列包含与本公开文本的融合蛋白的第一RNA结合蛋白或第二RNA结合蛋白结合的序列或由其组成。在一些实施方案中,本公开文本的支架序列包含二级结构或三级结构。示例性二级结构包括但不限于螺旋、茎环、凸起、四环和假结。示例性三级结构包括但不限于螺旋的A形式、螺旋的B形式和螺旋的Z形式。示例性三级结构包括但不限于扭曲的或螺旋化的茎环。示例性三级结构包括但不限于扭曲的或螺旋化的假结。在一些实施方案中,本公开文本的支架序列包含至少一种二级结构或至少一种三级结构。在一些实施方案中,本公开文本的支架序列包含一种或多种二级结构或者一种或多种三级结构。

[0166] 在本公开文本的组合物的一些实施方案中,指导RNA或其部分与本公开文本的RNA分子中的四环基序选择性结合。在一些实施方案中,RNA分子的靶序列包含四环基序。在一些实施方案中,四环基序是“GRNA”基序,所述基序包含GAAA、GUGA、GCAA或GAGA的序列中的

一种或多种或由其组成。

[0167] 在本公开文本的组合物的一些实施方案中,与RNA分子的靶序列结合的指导RNA或其部分与所述RNA分子的靶序列杂交。在一些实施方案中,与第一RNA结合蛋白或与第二RNA结合蛋白结合的指导RNA或其部分与第一RNA结合蛋白或与第二RNA结合蛋白共价结合。在一些实施方案中,与第一RNA结合蛋白或与第二RNA结合蛋白结合的指导RNA或其部分与第一RNA结合蛋白或与第二RNA结合蛋白非共价结合。

[0168] 在本公开文本的组合物的一些实施方案中,指导RNA或其部分包含在10个与100个之间的核苷酸(包括端点)或由所述核苷酸组成。在一些实施方案中,本公开文本的间隔子序列包含在10个与30个之间的核苷酸(包括端点)或由其组成。在一些实施方案中,本公开文本的支架序列包含15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个核苷酸或由其组成。在一些实施方案中,本公开文本的间隔子序列包含20个核苷酸或由其组成。在一些实施方案中,本公开文本的间隔子序列包含21个核苷酸或由其组成。在一些实施方案中,本公开文本的支架序列包含在10个与100个之间的核苷酸(包括端点)或由其组成。在一些实施方案中,本公开文本的支架序列包含30、35、40、45、50、55、60、65、70、76、80、87、90、95、100个或之间的任何数量的核苷酸或由其组成。在一些实施方案中,本公开文本的支架序列包含在85个与95个之间的核苷酸(包括端点)或由其组成。在一些实施方案中,本公开文本的支架序列包含85个核苷酸或由其组成。在一些实施方案中,本公开文本的支架序列包含90个核苷酸或由其组成。在一些实施方案中,本公开文本的支架序列包含93个核苷酸或由其组成。

[0169] 在本公开文本的组合物的一些实施方案中,指导RNA或其部分不包含核定位序列(NLS)。

[0170] 在本公开文本的组合物的一些实施方案中,指导RNA或其部分不包含与原间隔子相邻基序(PAM)互补的序列。

[0171] 本公开文本的治疗或药物组合物不包含PAMmer寡核苷酸。在其他实施方案中,任选地,非治疗或非药物组合物可以包含PAMmer寡核苷酸。术语“PAMmer”是指以下寡核苷酸,其包含能够与指导核苷酸序列可编程的RNA结合蛋白相互作用的PAM序列。PAMmer的非限制性例子描述于通过引用并入本文的O'Connell等人Nature 516,第263-266页(2014)中。PAM序列是指包含约2至约10个核苷酸的原间隔子相邻基序。PAM序列对与其相互作用的指导核苷酸序列可编程的RNA结合蛋白具有特异性,并且是本领域中已知的。例如,酿脓链球菌(*Streptococcus pyogenes*)PAM具有序列5'-NGG-3',其中“N”是任何核碱基,之后有两个鸟嘌呤(“G”)核碱基。新凶手弗朗西丝菌(*Francisella novicida*)的Cas9识别规范PAM序列5'-NGG-3',但是已经被工程化以识别PAM 5'-YG-3'(其中“Y”是嘧啶),从而添加至可能的Cas9靶标的范围内。新凶手弗朗西丝菌的Cpf1核酸酶识别PAM 5'-TTTN-3'或5'-YTN-3'。

[0172] 在本公开文本的组合物的一些实施方案中,指导RNA或其部分包含与原间隔子侧翼序列(PFS)互补的序列。在一些实施方案中,包括以下那些实施方案:其中指导RNA或其部分包含与PFS互补的序列,第一RNA结合蛋白可以包含从Cas13蛋白分离或衍生的序列。在一些实施方案中,包括以下那些实施方案:其中指导RNA或其部分包含与PFS互补的序列,第一RNA结合蛋白可以包含编码Cas13蛋白或其RNA结合部分的序列。在一些实施方案中,所述指导RNA或其部分不包含与PFS互补的序列。

[0173] 在本公开文本的组合物的一些实施方案中,本公开文本的指导RNA序列包含用于驱动指导RNA的表达的启动子。在一些实施方案中,包含本公开文本的指导RNA序列的载体包含用于驱动指导RNA的表达的启动子。在一些实施方案中,所述启动子是组成型启动子。在一些实施方案中,启动子是组织特异性和/或细胞类型特异性启动子。在一些实施方案中,启动子诱导型启动子。在一些实施方案中,启动子是杂合启动子或重组启动子。在一些实施方案中,启动子是能够在哺乳动物中驱动表达的启动子。在一些实施方案中,启动子是能够在人细胞中表达的启动子。在一些实施方案中,启动子是能够表达指导RNA序列并且将表达限制于细胞核的启动子。在一些实施方案中,启动子是人RNA聚合酶启动子或从人RNA聚合酶启动子分离或衍生的启动子序列。在一些实施方案中,启动子是U6启动子或从编码U6启动子的序列分离或衍生的序列。在一些实施方案中,启动子是人tRNA启动子或从人tRNA启动子的序列分离或衍生的启动子序列。在一些实施方案中,启动子是人缬氨酸tRNA启动子或从人缬氨酸tRNA启动子分离或衍生的启动子序列。

[0174] 在本公开文本的组合物的一些实施方案中,启动子还包含调节元件。在一些实施方案中,包含启动子的载体还包含调节元件。在一些实施方案中,调节元件增强指导RNA的表达。示例性调节元件包括但不限于增强子元件、内含子、外显子或其组合。

[0175] 在本公开文本的组合物的一些实施方案中,本公开文本的载体包含指导RNA序列、用于驱动指导RNA的表达的启动子和用于增强指导RNA的表达的调节元件中的一种或多种。在本公开文本的组合物的一些实施方案中,所述载体还包含编码本公开文本的融合蛋白的核酸序列。

融合蛋白

[0176] 本公开文本的融合蛋白包含第一RNA结合蛋白和第二RNA结合蛋白。在一些实施方案中,沿着编码融合蛋白的序列,编码第一RNA结合蛋白的序列定位于编码第二RNA结合蛋白的序列的5'。在一些实施方案中,沿着编码融合蛋白的序列,编码第一RNA结合蛋白的序列定位于编码第二RNA结合蛋白的序列的3'。

[0177] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从蛋白质分离或衍生的序列,所述蛋白质能够结合RNA分子。在一些实施方案中,编码第一RNA结合蛋白的序列包含从蛋白质分离或衍生的序列,所述蛋白质能够选择性结合RNA分子,并且不结合DNA分子、哺乳动物DNA分子或任何DNA分子。在一些实施方案中,编码第一RNA结合蛋白的序列包含从蛋白质分离或衍生的序列,所述蛋白质能够结合RNA分子,并且诱导所述RNA分子中的断裂。在一些实施方案中,编码第一RNA结合蛋白的序列包含从蛋白质分离或衍生的序列,所述蛋白质能够结合RNA分子,诱导所述RNA分子中的断裂,并且不结合DNA分子、哺乳动物DNA分子或任何DNA分子。在一些实施方案中,编码第一RNA结合蛋白的序列包含从蛋白质分离或衍生的序列,所述蛋白质能够结合RNA分子,诱导所述RNA分子中的断裂,并且既不结合DNA分子、哺乳动物DNA分子或任何DNA分子,也不诱导所述DNA分子中的断裂。

[0178] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从没有DNA核酸酶活性的蛋白质分离或衍生的序列。

[0179] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从具有DNA核酸酶活性的蛋白质分离或衍生的序列,其中在将本公开文本的组合物与RNA分子接触或引入本公开文本的细胞中或受试者体内时,所述DNA核酸酶活性不诱导DNA分子、哺

乳动物DNA分子或任何DNA分子中的断裂。

[0180] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从具有DNA核酸酶活性的蛋白质分离或衍生的序列,其中所述DNA核酸酶活性是失活的,并且其中在将本公开文本的组合物与RNA分子接触或引入本公开文本的细胞中或受试者体内时,所述DNA核酸酶活性不诱导DNA分子、哺乳动物DNA分子或任何DNA分子中的断裂。在一些实施方案中,编码第一RNA结合蛋白的序列包含使DNA核酸酶活性失活或降低至以下水平的突变,在所述水平下,在将本公开文本的组合物与RNA分子接触或引入本公开文本的细胞中或受试者体内时,所述DNA核酸酶活性不诱导DNA分子、哺乳动物DNA分子或任何DNA分子中的断裂。在一些实施方案中,编码第一RNA结合蛋白的序列包含使DNA核酸酶活性失活或降低的突变,并且所述突变包含对编码第一RNA结合蛋白或其核酸酶结构域的核酸序列或氨基酸序列的取代、倒位、转座、插入、缺失或其任何组合中的一种或多种。

[0181] 在本公开文本的组合物的一些实施方案中,编码本文公开的RNA指导的融合蛋白的第一RNA结合蛋白的序列包含从CRISPR Cas蛋白分离或衍生的序列。在一些实施方案中,所述CRISPR Cas蛋白包含II型CRISPR Cas蛋白。在一些实施方案中,所述II型CRISPR Cas蛋白包含Cas9蛋白。本公开文本的示例性Cas9蛋白可以从任何物种分离或衍生,所述物种包括但不限于细菌或古菌。本公开文本的示例性Cas9蛋白可以从任何物种分离或衍生,所述物种包括但不限于酿脓链球菌、地中海富盐菌 (*Haloferax mediteranii*)、结核分枝杆菌 (*Mycobacterium tuberculosis*)、土拉热弗朗西丝菌新凶手亚种 (*Francisella tularensis* subsp. *novicida*)、多杀巴斯德菌 (*Pasteurella multocida*)、脑膜炎奈瑟球菌 (*Neisseria meningitidis*)、空肠弯曲杆菌 (*Campylobacter jejune*)、嗜热链球菌 (*Streptococcus thermophilus*)、红嘴鸥弯曲杆菌CF89-12 (*Campylobacter lari* CF89-12)、鸡毒支原体F株 (*Mycoplasma gallisepticum* str. F)、卤水硝酸盐裂解菌DSM 16511株 (*Nitratifractor salsuginis* str. DSM 16511)、食清洁剂细小棒菌 (*Parvibaculum lavamentivorans*)、肠道罗斯拜瑞氏菌 (*Roseburia intestinalis*)、灰色奈瑟球菌 (*Neisseria cinerea*)、重氮营养葡萄糖酸醋杆菌 (*Gluconacetobacter diazotrophicus*)、固氮螺菌 (*Azospirillum*) B510、球状螺旋菌巴迪株 (*Sphaerochaeta globus* str. Buddy)、柱状黄杆菌 (*Flavobacterium columnare*)、塔夫河栖河菌 (*Fluviicola taffensis*)、嗜粪拟杆菌 (*Bacteroides coprophilus*)、运动支原体 (*Mycoplasma mobile*)、香肠乳杆菌 (*Lactobacillus farciminis*)、巴氏链球菌 (*Streptococcus pasteurianus*)、约氏乳杆菌 (*Lactobacillus johnsonii*)、伪中间型葡萄球菌 (*Staphylococcus pseudintermedius*)、龈沟产线菌 (*Filifactor alocis*)、齿垢密螺旋体 (*Treponema denticola*)、嗜肺军团菌巴黎株 (*Legionella pneumophila* str. Paris)、华德萨特氏菌 (*Sutterella wadsworthensis*)、白喉棒状杆菌 (*Corynebacter diphtherias*)、金黄色葡萄球菌 (*Streptococcus aureus*) 和新凶手弗朗西丝菌。

[0182] 本公开文本的示例性的野生型酿脓链球菌Cas9蛋白可以包含以下氨基酸序列或由其组成:

1 MDKKYSIGLD IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE
61 ATRLKRTARR RYTRRKNRIC YLQEIFSNEK AKVDDSFHHR LEESFLVEED KKHERHPIFG
121 NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLPNDNSD
181 VDKLFIQLVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN
241 LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLNLLA QIGDQYADLF LAAKNLSDAI
301 LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI FFDQSKNGYA
361 GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH
421 AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE
481 VVDKGASQAS FIERMTNFDK NLPNEKVLPK HSLLEYFTV YNELTKVKYV TEGMRKPAFL
541 SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI
601 IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKYA HLFDDKVMKQ LKRRRYTGWG
661 RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFQMQLHDD SLTFKEDIQK AQVSGQGDSL
721 HEHIANLAGS PAIKKGILQT VKVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER
781 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH
841 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKMK NYWRQLLNAK LITQRKFDNL
901 TKAERGGLSE LDKAGFIKQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
961 KLVSDFRKDF QFYKVEINN YHHAHDAYLN AVVGITALIKK YPKLESEFVY GDYKVYDVRK
1021 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
1081 ATVRKVLSP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPPK YGGFDSPTVA
1141 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK
1201 YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
1261 QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHARDK PIREQAENII HLFTLTNLGA
1321 PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO: 147)。

[0183] 核酸酶失活的酿脓链球菌Cas9蛋白可以包含丙氨酸(A)取代位置10的天冬氨酸(D)以及丙氨酸(A)取代位置840的组氨酸(H)。本公开文本的示例性的核酸酶失活的酿脓链球菌Cas9蛋白可以包含以下氨基酸序列或由其组成(D10A和H840A加粗并加下划线)：

1 MDKKYSIGLA IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE
61 ATRLKRTARR RYTRRKNRIC YLQEIFSNE AKVDDSFHR LEESFLVEED KKHERHPIFG
121 NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLPDNDSD
181 VDKLFIQLVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN
241 LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLNLLA QIGDQYADLF LAAKNLSDAI
301 LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI FFDQSKNGYA
361 GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH
421 AILRRQEDFY PFLKDNREKI EKILTFRIPI YVGPLARGNS RFAWMTRKSE ETITPWNFEE
481 VVDKGASQAS FIERMTNFDK NLPNEKVLKP HSLLEYFTV YNELTKVKYV TEGMRKPAFL
541 SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI
601 IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKYA HLFDDKVMKQ LKRRRYTGWG
661 RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL
721 HEHIANLAGS PAIKKGILQT VKVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER
781 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVAA
841 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL
901 TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
961 KLVSDFRKDF QFYKVEINN YHHAHDAYLN AVVGTAIIK YPKLESEFVY GDYKVYDVRK
1021 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
1081 ATVRKVLSP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA
1141 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLKP
1201 YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
1261 QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHHRDK PIREQAENII HLFTLTNLGA
1321 PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO: 148)。

[0184] 核酸酶失活的酿脓链球菌Cas9蛋白可以包含RuvC核酸酶结构域或其部分、HNH结构域、DNA酶活性位点、包含DNA酶活性位点的 $\beta\beta\alpha$ -金属折叠或其部分或者其任何组合的缺失。

[0185] 其他示例性Cas9蛋白或其部分可以包含以下氨基酸序列或由所述氨基酸序列组成。

[0186] 在一些实施方案中,所述Cas9蛋白可以是酿脓链球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLPDNDSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR

RLENLIAQLPGEKKNGLFGNLI ALSLGLTPNFKS NFDAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNL
SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PYVVGPLARGNSRFAWMTRKSEETITPWNFEFVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNEL
TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI
IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKS DGFANRNFMLIHDDSLTFKEDIQKAQVSGQGS LHEHIANLAGSPA I KKGILQTVKVVDELVKVMG
RHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINN
YHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG
GFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS SF EKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELN
GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIEQISEFSKRVI LAD
ANLDKVL SAYNKH RDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRI
DLSQLGGD (SEQ ID NO:149)

[0187] 在一些实施方案中,所述Cas9蛋白可以是金黄色葡萄球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MKRN YILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQRVK
KLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEDTGNELSTKEQISRNSKA
LEEKYVAELQLERLKKDG EVRGSINRFKTS DYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGE GSP
FGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYEFQI IENVFKQKKKPTL
KQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEI IENAELLDQIAKILTIYQSSEDIQEELTNLN
SELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSP
VVKRSFIQSIKVINAI IKKYGLPNDII IELAREKNSKDAQKMINEMQKRNRQTNERIEE IIRTTGKENAKYLIEKI
KLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQYLS SSDSKIS
YETFKKHILNLA KGKGRISKTKEYLLEERDINRF SVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKV KSI
NGGFTSFLRRKWKFKKERNKG YKHH AEDAL I IANADFIK EWKKLDKAKKVMENQMFE EKQAESMPEIETE QEYKE
IFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLL
MYHHDPQTYQKLKLIMEQY GDEKNPLYKYEEETGN YLTKYSKKDNGPV IKKIKYYGNKLN AHLDITDDYPNSRNKV
VKLSLKP YRFDVYLDNGVYKFVTVKNLDV IKKENYYEVNSKCYEEAKK LKKSINQAEFIASFYNNDLIKINGEL YR
VIGVNNDLLNRIE VNMIDITYREYLENMNDKRPPRI IKTIASKTQS I KKYSTDILGNLYEVKSKKHPQIIKKG (SEQ
ID NO:150)

[0188] 在一些实施方案中,所述Cas9蛋白可以是嗜热链球菌CRISPR1Cas9,并且可以包含以下氨基酸序列或由其组成:

MSDLVLGLDIGIGSVGVGILNKVTGEI IHKNSRIFPAAQAENNLVRR TNRQGRRLARRKKHRRVRLNR
LFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFI ALKNMVKHRGISYLD DASDDGNSSVG DYAQIVKENS K
QLETKTPGQIQ LERYQTYGQLRGDF TVEKD GK KHR LINVFP TSAYRSEALRILQTQQEFNPQITDEFINRYLEILT
GKRKYHGP GNEKSRTDYGRYRTSGETLDNIFGILIGKCTFY PDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKL

SKEQKNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL
DKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMT
ILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQ
KIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGKQLATKIRLWHQQGERCLYTGTISIHDLINNSNQFEVDHI
LPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTL SNKKKEYLLTEEDISKFDVR
KKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLN
LWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFDVTLKSKEFEDSILFSYQVDSKFNKISDATIY
ATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKE
VPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYDYSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEI
LGLKYADLQFDKGTGTYKISQEKYNDIKKKEGVSDSEFKFTLYKNDLLL VKDTETKEQQLFRLSRTMPKQKHVY
ELKPYDKQKFEGGEALIKVLGNVANSQGCKGLGKSNISYKVRTDVLGNQHI IKNEGDKPKLDF (SEQ ID NO:
151)。

[0189] 在一些实施方案中,所述Cas9蛋白可以是脑膜炎奈瑟球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRFERAEVPKTGDSLAMARRLARSV
RRLTRRRARHLLRARRLLKREGVLQAADFDEGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQK
EFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNYTAERFIWLTKLNNLRILEQGSERP
LTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSP
NLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTTEEKIYLPPIPADEIRNPVVLRLSQAARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKAAAKFREYFPNFVGEPSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLVLGSENQKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDEGFKERNLNDTRYVNR
LCQFVADRMRLTGKGGKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTP
FVSRAPNRKMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLDLEKMNREREPPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHGIADNATMVRVDVFEKGDKYLVPIYSWQVAKGILPDRVV
QKGDEEDWQLIDDSFNFKFSLHPNDLVEVITTKARMFYGFASCHRG TGNINIRIHDL DHKIGKNGILEGIGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:152)。

[0190] 在一些实施方案中,所述Cas9蛋白可以是食清洁剂细小棒菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIPEARDPDGTPLNQRRQKRMRRQLRRRRI
RRKALNETLHEAGFLPAYGSADWPVMADEPYELRRRGLEGLSAYEFGRAIYHLAQHRHFKGRELESDTPDPDV
DDEKEAANERAATLKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVAEEFERLWEVQSKFHPALKSEEMRARI
SDTIFAQRPVFWRKNTLGECRFMPGEPLCPKGSWLSQRRMLEKLNNLAIAGGNARPLDAEERDAILSKLQQQASMSW
PGVRSALKALYKQGEPGAESLKFNLGEGESKLLGNALEAKLADMFGPDWPAHPRKQEI RHAVHERLWAADYGE
TPDKKRVIILSEKDRKAHREAAANSFVADFGITGEQAAQLQALKLPTGWEPYSIPALNLF LAELEKGERFGALVNG
PDWEGWRRNTNPHRNQPTGEILDKLPSASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYGKPDRI RIEVGRD
VGKSKREREI IQSGIRRNEKQRKKATEDLIKNGIANPSRDDVEKWILWKEGQERCPYTGDI GFNALFREGRYEVE

HIWPRSRSFNDSRNTLCRKDVNIEKGNRMPFEAFGHDEDRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMP
EDFAARQLNDTRYAAKQILAQKLRLWPDGMPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDA
LTVACTHPGMTNKLSTRYWLRRDDPRAEKPALTPPWDTIRADA EKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDI
KTKSGTYRQFVTRKKIESLSKGELDEIRDPRIKEIVAAHVAGRGDPKKAFFPPYPCVSPGGPEIRKVRLTSKQQLN
LMAQTGNGYADLGSNHIIAIYRLPDGKADFEIVSLFDASRRLAQRNPVQRTRADGASFVMSLAAGEAIMIPEGSK
KGIWIVQGVWASGQVLERDTHSTTRPMPNPILKDDAKKVSIDPIGRVRPSND (SEQ ID NO:153)。

[0191] 在一些实施方案中,所述Cas9蛋白可以是白喉棒状杆菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MKYHVGIDVGTFSVGLAAIEVDDAGMPIKTLSLVSHIHDSGLDPDEIKSAVTRLASSGIARRTRRLYR
RKRRRLQQLDKFIQRQGWPIELEDYSDPLYPWKVRAELAASYIADEKERGEKLSVALRHIAHRGWNPYAKVSS
LYLPDGPSTDAFKAIREEIKRASGQVPETATVGQMVTLCELGTLLRGEGLVLSARLQQSDYAREIQEICRMQEIG
QELYRKIIDVVFAAESPKGSASSRVGKDPLQPGKNRALKASDAFQRYRIAALIGNLRVRVDGEKRILSVEEKNLVF
DHLVNLTPKKEPEWVTIAEILGIDRGQLIGTATMTDDGERAGARPPTHDTNRSIVNSRIAPLVDWWKTASALEQHA
MVKALSNAEVDDFDSPEGAKVQAFFADLDDDVHAKLDSLHLPVGRAAYSEDTLVRLTRRMLSDGVDLYTARLQEFQ
IEPSWTPPTPRIGEPVGNPAVDRVLKTVSRWLESATKTWGAPERVIIIEHVREGFVTEKRAREMDGDMRRRAARNAK
LFQEMQEKLNVQGKPSRADLWRYQSVQRQNCQAYCGSPITFSNSEMDHIVPRAGQGSTNTRENLVAVCHRCNQSK
GNTPFIAIWAKNTSIEGVSVEKEAVERTRHWVTDGMRSTDFKKFTKAVVERFQRATMDEEIDARSMESVAMANELR
SRVAQHFASHGTTVRVYRGSALTAEARRASGISGKLKFFDGVGKSRLDRRHHAIDAAVIAFTSDYVAETLAVRSNLK
QSQAHRQEAPQWREFTGKDAEHRAAWRVWCQKMEKLSALLTEDLRDDRVRVMSNVRLRLNGSAHKETIGKLSKVK
LSSQLSVSDIDKASSEALWCALTREPGFDPKEGLPANPERHIRVNGTHVYAGDNIGLFPVSAGSIALRGGYAELGS
SFHHARVYKITSGKKPAFAMLRVYTIDLLPYRNQDLFSVELKPQTMSMRQAEKKLRDALATGNAEYLGWLVDDEL
VVDTSKIATDQVKAWEAELGTIRRWVRDGGFFSPSKLRLRLQMSKEGKESAPELSKIIDRPGWLPAVNKLFSQG
NVTVVRDLSLGRVRLESTAHLPVTWKVQ (SEQ ID NO:154)。

[0192] 在一些实施方案中,所述Cas9蛋白可以是巴氏链球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MTNGKILGLDIGIASVGVGIIIEAKTGKVVHANSRLFSAANAENNAERRGFRGSRRLNRRKKHRVKRVR
DLFEKYGIVTDFRNLNLPYELRVKGLTEQLKNEELFAALRTISKRRGISYLDAAEDDSTGSTDYAKSIDENRRL
KNKTPGQIQLERLKYGQLRGNFTVYDENGEAHLINVFSTSDYEKEARKILETQADYNKKITAEIFIDYVEILTQ
KRKYHGPNGEKSRTDYGRFRDGTTLNIFGILIGKCNFYPPDEYRASKASYTAQEYNFLNDLNNLKVSTETGKLS
TEQKESLVEFAKNTATLGPALLKEIAKILDCVDEIKGYREDDKGKPDLHTFEPYRKLKFNLESINIDDLREVI
DKLADILTLNTEREGIEDAIKRNLPNQFTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNELIPELYATSDEQMTI
LTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAAVKKYGDFDKIVIEPRDKNADDEKKFID
KRKNENKKEKDDALKRAAYLYNSSDKLPDEVFHGNKQLETKIRLWYQQGERCLYSGKPISIQELVHNSNNFEIDHI
LPLSLSFDDSLANKVLVYAWTNQEGKQKTPYQVIDSMDAAWSFREMCDYVLKQKGLGKKRDRYLLTTENIDKIEVK
KKFIERNLVDTRYASRVVLSLQSLRELKDKTKSVVRGQFTSQLRRKWKIDKSRETYHHHAVDALIIAASSQLK
LWEKQDNPMFVDYGNQVVDKQTEILSVSDDEYKELVFQPPYQGFVNTISSKGFEDILFSYQVDSKYNRKVSDA
TIYSTRKAKIGKDKKEETYVLGKIKDIYSQNGFDTFIKKYNKDKTQFLMYQKDSLWENVIEVILRDYPTTKKSED
GKNDVKCNPFEEYRRENGLICKYSKKGKGTPIKSLKYYDKKLGNCIDITPEESRNKVILQSINPWRADVFPNPETL
KYELMGLKYSDSLFEKGTGNYHISQEKYDAIKEKEGIGKKSEFKFTLYRNDLILIKDIASGEQEIYRFLSRTMPNV

NHYVELKPYDKEKFDNVQELVEALGEADKVGRCIKGLNKPNI SIYKVRTDVLGNKYFVKKKGDKPKLDFKNNKK (SEQ ID NO:155)。

[0193] 在一些实施方案中,所述Cas9蛋白可以是灰色奈瑟球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MAAFKPNPMNYILGLDIGIASVGWAIVEIDEEENPIRLIDLGVRFERAEVPKTGDSLAAARRLARSV
RRLTRRAHRLLRARLLKREGVLQAADFDEGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVADNTHALQTGDFRTPAELALNKFEEKSGHIRNQRGDYSHTFNRKDLQAEINLLFEKQK
EFGNPHVSDGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPTPKAAKN TYTAERFVWLTCLNNLRILEQGSERP
LTDTERATLMDEPYRKSCLTYAQARKLLDLDDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPL
NLSPELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKQVQISLKLRRIVPLMEQGNRYDEACTEI
YGDHYGKKNTTEEKIYLPPIPADEIRNPVVLRLSQAARKVINGVRRYRGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKSAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLALGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRI LLQKFDEDFGKERNLNDTRYINRF
LCQFVADHMLLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTIAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKAHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRLLAEKLSSRPEAVHKYVTP
FISRPNRKMSGQGHMETVKSARKLDEGISVLRVPLTQLKLKLEKMNRREREKPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKAVRVEQVQKTGVVHNGIADNATIVRVDVFEKGGKYLYLPIYSWQVAKGILPDRVV
QKGDEEDWTVMDDSFEFKFLYANDLIKLTAKKNEFLGYFVSLNRATGAIDIRTHDSTDSTKGKNGIFQSVGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:156)。

[0194] 在一些实施方案中,所述Cas9蛋白可以是红嘴鸥弯曲杆菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MRILGFDIGINSIGWAFVENDELKDCGVRIFTKAENPKNKESLALPRRNARSSRRRLKRRKARLIAIK
RILAKELKLNKYDYAADGELPKAYEGSLASVYELRYKALTQNLTKDLARVILHIAKHRGYMNKNEKKSNDKKG
KILSALKNNALKLENYQSVGEYFYKEFFQKYKKN TKNF IKIRNTKDNYN NCVLSSDLEKELKLI LEKQKEFGYNYS
EDFINEILKVAFFQRPLKDFSHLVGACTFFEEKRACKNSYS AWEFVALTKIINEIKSLEKISGEIVPTQTINEVL
NLILDKGSITYKKFRSCINLHESISFKSLKYDKENAENAKLIDFRKLVEFKKALGVHSLSRQELDQISTHITLIK
NVKLKTVLEKYNLSNEQINNLEIEFN DYINLSFKALGMILPLMREGKRYDEACEIANLKP KTVDEKKDFLP
SIFAHELSNPVVNRAISEYRKVLNALLKKYGVH KIHLELARDVGLSKKAREKIEKEQKENQAVNAWALKECENIG
LKASAKNILKLKLWKEQKEIC IYSGNKISIEHLKDEKALEVDHIYPYSRSFDDSFINKVLVFTKENQEKL NKT
PFEAFGKNIEKWSKIQT LAQNLPYKKKNKILDENFKDKQ QEDFISRNLNDTRYIATLIAKYTKEYLNFLLLSE
NENANLKS GEKGSKIHVQTISGMLTSVLRHTWGFDDKDRNNHLHHALDAIIVAYSTNSIIKAFSDFRKNQELL
KARFYAKELTSDNYKHQVKFFEPFKSFREKILSKIDEIFVSKPPRKRARRALHKDTFHSENKIIDKCSYNSKEGL
QIALSCGRVRKIGTKYVENDTIVRVDIFKKQNKFYAIP IYAMDFALGILPNKIVITGDKNNNPKQWQTIDESY
EFCFSLYKNDLILLQKKNMQEPEFAYYNDFSISTSSICVEKHDNKFENLTSNQKLLFSNAKEGSVKVESLGIQNL
KVFEKYIITPLGDKIKADFQPRENISLKTSKKYGLR (SEQ ID NO:157)。

[0195] 在一些实施方案中,所述Cas9蛋白可以是齿垢密螺旋体Cas9,并且可以包含以下氨基酸序列或由其组成:

MKKEIKDYFLGLDVGTGSGWAVTDTDYKLLKANRKDLWGMRCFETAETA EVRRLHRGARRRIERRKK
RIKLLQELFSQEIAKTDEGFFQRMKESPFYAEDKTI LQENTLFNDKDFADKTYHKAYPTINHLIKAWIENKV
KPD

RLLYLACHNIIKKRGHFLFEGDFDSENQFDTSIQALFEYLREDMEVDIDADSQKVKEILKDSSLKNSEKQSRNLKI
LGLKPSDKQKKAITNLSGNKINFADLYDNPDLDKAEKNSISFSKDDFDALSDDLASILGDSFELLLKAKAVYNCS
VLSKVIGDEQYLSFAKVKIYEKHKTDLTCLKNVKKHFPKDYKKVFGYNKNEKNNNNYSGYVGCKTKSKKLIINN
SVNQEDFYKFLKTILSAKSEIKEVNDILTEIETGTFLPKQISKSNAEIPYQLRKMELEKILSNAEKHFSFLKQKDE
KGLSHSEKIIMLLTFKIPYYIGPINDNHKKFFPDRCVWVKKKEKSPSGKTPWNFFDHIDKEKTAEAFITSRNFTCT
YLVGESVLPKSSLLYSEYTVLNEINNLIIDGKNICDIKQKIYEDLFKKYKKITQKQISTFIKHEGICNKTDE
VIILGIDKECTSSLSYIELKNIFGKQVDEISTKNMLEEIIIRWATIYDEGEGKTILKTKIKAEGKYCSDEQIKKI
LNLKFSGWGRLSRKFLFETVTSEMPGFSEPVNIITAMRETQNNLMELLSSEFTFTENIKKINSGFEDAEKQFSYDGL
VKPLFLSPSVKMLWQTLKLVEISHITQAPPKIFIEMAKGALEPARTKTRKILQDLYNNCKNDADAFSSEIK
DLSGKIENEDNLRRLSDKLYLYTQLGKCMYCGKPIEIGHVFDTSNYDIDHIYPQSKIKDDISNRVLVCCSNKN
KEDKYPLKSEIQSKQRGFWNFLQRNFIISLEKLNRLTRATPISDDETAQFIARQLVETRQATKVAKVLEKMPFET
KIVYSKAETVSMFRNKFDIVKCREINDFHHAHDAYLNIVVGNVYNTKFTNNPWNFIKEKRDNPKIADTYNYYKVF
YDVKRNNITAWKGGTIITVKDMLKRNTPIYTRQAACKKGELFNQTIMKKGLGQHPLKKEGPFNSISKYGGYNKVS
AAYYTLIEYEEKGNKIRSLETIPLYLVKDIQKDQDVLSYLTDLGKKEFKILVPKIKINSLLKINGFPCHITGKT
NDSFLLRPVAVQFCCSNNEVLYFKKIIIRFSEIRSQREKIGKTIISPYEDLSFRSYIKENLWKKTKNDEIGEKEFYDLL
QKKNLEIYDMLLTKHKDTIYKKRPNSATIDILVKGKEKFKSLIENQFEVILEILKLFSATRNVSDLQHIGGSKYS
GVAKIGNKISSLDNCILYQSITGIFEKRIDLLKV (SEQ ID NO:158)。

[0196] 在一些实施方案中,所述Cas9蛋白可以是变形链球菌 (*S. mutans*) Cas9,并且可以包含以下氨基酸序列或由其组成:

MKKPYSIGLDIGTNSVGWAVVTDDYKVPAAKMKVLGNTDKSHIEKNLLGALLFDSGNTAEDRRLKRTA
RRRYTRRRNRILYLQEIFSEEMGKVDDSFHRLSDSFLVTEDKRGERHPIFGNLEEEVKYHENFPTIYHLRQYLAD
NPEKVDLRLVYLALAHIIKFRGHFLIEGKFDRNNDVQRLFQEFLAVYDNTFENSSSLQEQNQVEEILTDKISKSA
KKDRVLKLPNEKSNRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDYEEEELEVLLAQIGDNYAELFLSAKKL
YDSILLSGILTVTDVGTAKPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSVSKDGYAGYIDGKTNQEAF
YKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEFYFPFLADNQDRIEKLTFRI
PPYVGPLARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPHSLLYEKFTVYNEL
TKVKYKTEQGKTAFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKVFNASYGTYHDL
KILDKDFLDNSKNEKILEDIVLTLTFEDREMIRKRLNYSDDLTKQVKKLERRHYTGWGRLSAELIHGIRNKES
RKTILDYLIIDGNSNRNFMQLINDDALSFKEEIAKAQVIGETDNLNQVVSADIAGSPAIKKILQSLKIVDELVKIM
GHQPENIVVEMARENQFTNQGRNSQQRLKGLTDSIKEFGSQILKEHPVENSQQLQNDRLFLYYLQNGRDMYTGEEL
DIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYWSKLLSAKLITQRKFDNLTAE
RGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTETDENKKIRQVKIVTLKSNLVSNFRKEFELYKVIN
YHHAHDAYLNAVIGKALLGVYPQLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKD
EHISNIKKVLSYPQVNIIVKKVEEQTGGSKEKILPKGNSDKLIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIE
KGKSKKLKTVKALVGVTIMEKMTFERDPVAFLEKGYRNVQEENI IKLPKYSLFKLENGRKRLLASARELQKGNEI
VLPHLGLTLLYHAKNIHKVDEPKHLDYVDKHKDEFKELLVVSNSFSKKYTLAEGNLEKIKELYAQNGEDLKELAS
SFINLLTFTAIGAPATFKFFDKNIDRKRYTSTTEILNATLIHQSIITGLYETRIDLNKLGGD (SEQ ID NO:159)

[0197] 在一些实施方案中,所述Cas9蛋白可以是嗜热链球菌CRISPR 3Cas9,并且可以包含以下氨基酸序列或由其组成:

MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKAYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDELTLLGYIGDDYSDVFLKAKKLYDAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKYTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPPFLAKNKERIEKILTFRIPIYYVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLNIIINDKEFLDDSSNEAIIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGLSAKLINGIRDEKSGNTILDYLIIDDGISNRNFMQLIHDDALSFKKKIQKAQIIIGDEDKGNIKEVVKSLPGSPAIIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGSNSQQLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYLQNGKDMYTGDDLIDRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKRKTFWYQLLSKLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAVRTVKIITLKTLSVSQRKDFELYKVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAGISNSFTVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLEKGYKDIELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEEFLYYYILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO:160)

[0198] 在一些实施方案中,所述Cas9蛋白可以是空肠弯曲杆菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLARRKARLNHLKHLIANEFKLNIEDYQSFDESLAKAYKGSLSIPYELRFALNELLSKQDFARVILHIAKRRGYDDIKNSDDKEGAILKAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCIASFLKDELKLIFKKQREFGFSFSKKFEEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGILYTKDDLNALLEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGYTIFIEFKKYKEFIKALGEHNSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTPNVVLRAIKEYRKVLNALLKKYGVHKKINIELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMKNVLVFTKQKQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPKKQKRILDKNYKDEKQNFKDRNLNDTRYIARLVLNKYTDYLDLPLSDDENTKLNQTKGSKVHVEAKSGMLTSALRHTWGFSKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFEPFSGFRQKVLKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVKALELGKIRKVNKGIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCSLYKDSLILIQTKDMQEPFVYYNAFTSSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK (SEQ ID NO:161)

[0199] 在一些实施方案中,所述Cas9蛋白可以是多杀巴斯德菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MQTTNLSYILGLDLGIASVGWAVVEINENEDPIGLIDVGVRIFERAEPKGTGESLALSRRLARSTRRLIRRRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEWGAVLLHLIKHRGYLSKRKNESQTNN

KELGALLSGVAQNHQLQSDDYRTPAELALKKFAKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCK
EHIQQYMTTELLMWQKPALSGEAILKMLGKCTHEKNEFKAACKHTYSAERFVWLTKLNNLRILEDGAERALNEEERQL
LINHPYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIRKALENQGLKDTWQDLAKKPDLLD
EIGTAFSLYKTDEDIQQYLTKNVPNSVINALLVSLNFDKFIELSLKSLRKILPLMEQGKRYDQACREIYGHYGEA
NQKTSQLLPAIPAQEIIRNPVVLRTLSQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQKQEDNRTKRES
AVQKFKELFSDFSSEPKSKDILKFRLYEQQHGKCLYSGKEINIHRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLA
SENQNKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAAKKQRLLTQVIDDNKFIDRNLDTRYIARFLSNYIQEN
LLLVGKNKNVFTPNGQITALLRSRWGLIKARENNNRHHALDAIVVACATPSMQQKITRFRFKEVHPYKIENRYE
MVDQESGEIISPFPPEWAYFRQEVNIRVFDNHPDVTVLKEMPLDRPQANHQFVQPLFVSRAPTRKMSGQGHMETIK
SAKRLAEGISVLRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPAKAFATPFYKQGGQVKAIRVEQVQ
KSGVLVRENNGVADNASIVRTDVFIKNNKFFLVPIYTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLPND
LVELKTKKEYFFGYIIGLDRATGNISLKEHDGEISKGKGVYRVGVKLALSFEKYQVDELGKNRQICRPQQRQPV
(SEQ ID NO:162)

[0200] 在一些实施方案中,所述Cas9蛋白可以是新凶手弗朗西丝菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MNFKILPIAIDLGVKNTGVFSAFYQKGTSLERLDNKNKGKVELSKDSYTLLMNNRTARRHQRRGIDRK
QLVKRLFKLIWTEQLNLEWDKDTQQAISFLFNRRGFSFITDGYSPEYLNIVPEQVKAILMDIFDDYNGEDDLDSYL
KLATEQESKISEIYNKLMQKILEFKLMKLCTDIKDDKVSTKTLKEITSYEFELLADYLANYESLKTQKFSYTDKQ
GNLKELSYHHDKYNIQEFILKRHATINDRILDTLLTDDLDIWNFNFEKFDKNEEKLQNQEDKDHIQAHLHHFVF
AVNLIKSEMASGGRHRSQYFQEITNVLDENNHQEGYLKNFCENLHNKKYSNLSVKNLVNLIGNLSNLELKPLRKYF
NDKIHAKADHWDEQKFTETTYCHWILGEWRVGVKDQDKDGAKYSYKDLCELKQKVTAGLVDFLLELDPCRTIPP
YLDNNNRKPPKCQSLILNPKFLDNQYPNWQYLQELKKLQSIQNYLDSFETDLKVLKSSKDQPYFVEYKSSNQQA
SGQRDYKDLARILQFIFDRVKASDELLNEIYFQAKKLKQKASSELEKLESSKKLDEVIANSQLSQILKSQHTNG
IFEQGTFLHLVCKYYKQRQRARDSRLYIMPEYRYDKKLHKYNNTGRFDDDNQLLTYCNHKPRQKRYQLLNDLAGVL
QVSPNFLKDKIGSDDDLFISKWLVEHIRGFKKACEDSLKIQKDNRGLLNHKINIARNTKGKCEKEIFNLICKIEGS
EDKKGNKYKHGLAYELGVLLFGEPNEASKPEFDRKIKKFNSIYSFAQIQQAFAERKGNANTCAVCSADNAHRMQQI
KITEPVEDNKDKIILSAKAQRLPAIPTRIVDGAVKKMATILAKNIVDDNWQNIKQVLSAKHQLHIPIITESNAFEF
EPALADVKGKSLKDRRKALERISPENIFKDKNNRIKEFAKGISAYSGANLTDGDFDGAKEELDHIIPRSHKKYGT
LNDEANLICVTRGDNKNKGNRIFCLRDADNYKLKQFETDDLEIEKKIADTIWDANKKDFKFGNYRSFINLTPQE
QKAFRHALFLADENPIKQAVIRAINNRNRTFVNGTQRYFAEVLANNIYLAKKENLNTDKISFDYFGIPTIGNRG
IAEIRQLYEKVDSDIQAYAKGDKPQASYSHLIDAMLAFCIAADEHRNDGSIGLEIDKNYSLYPLDKNTGEVFTKDI
FSQIKITDNEFSDDKLVRKKAIEGFNTHRQMTDGIYAENYLPILIHKELNEVRKGYTWKNSEEIKIFKGKKYDIQ
QLNNLVYCLKFVDKPISIDIQISTLEELRNILTTNNIAATAEYYYINLKTQKLHEYIENYNTALGYKKYSKEMEF
LRSLAYRSERVKIKSIDDVKQVLDKDSNFIIGKITLPFKKEWQRLYREWQNTTIKDDYEFLKSFFNVKSITKLHKK
VRKDFSLPISTNEGKFLVKKRTWDNNFIYQILNDSRADGTPFIPAFDISKNEIVEAIDSFTSKNIFWLPKNI
ELQKVDNKNIFAIDTSKWFEVETPSDLRDIGIATIQYKIDNNSRPKVRVKLDYVIDDDSKINYFMNHSLLKSRYPD
KVLEILKQSTIIEFESSGFNKTIKEMLMKLAGIYNETSNN (SEQ ID NO:163)

[0201] 在一些实施方案中,所述Cas9蛋白可以是布氏乳杆菌 (*Lactobacillus buchneri*) Cas9,并且可以包含以下氨基酸序列或由其组成:MKVNYYHIGLDIGTSSIGWVAIGKDGKPLRVKKG

TAIGARLFQEGNPAADRRMFRTTRRRLSRRKWRLKLLIEIFDPYITPVDSTFFARLKQSNLSPKDSRKEFKGSMLF
PDLTDMQYHKNYPTIYHLRHALMTQDKKFDIRMVYLAIHIVKYRGNFLNSTPVDSFKASKVDFVDQFKKLNELYA
AINPEESFKINLANSEDIGHQFLDPSIRKFDKKKQIPKIVPVMNDKVTDRNLNGKIASEIIHAILGYKAKLDVVLQ
CTPVDSKPPWALKFDDDEDIDAKLEKILPEMDENQQSIVAILQNLYSQVTLNQIVPNGMSLSESMIEKYNDHHDHLKL
YKKLIDQLADPKKKAVLKAYSQYVGDDGKVEQAEFWSSVKKNLDDSELSKQIMDLIDAEKFMKPQRTSQNGVIP
HQLHQRELDEIIHQSKYYPWLVEINPNKHDHLAKYKIEQLVAFRVPYYVGPMTPKDQAESAETVFSWMERKGT
ETGQITPWNFDEKVDKASANRFIKRMTTKDTYLIGEDVLPDESLLYEKFKVLNELNMVRVNGKLLKVADKQAIQFQ
DLFENYKHVSVKKLQNYIKAKTGLPSDPEISGLSDPEHFNNSLGTYNDFKKLFGSKVDEPDLQDDFEKIVEWSTVF
EDKKILREKLNEITWLSQQKQDVLESSRYQGWGRLSKKLLTGIVNDQGERIIDKLWNTNKNFMQIQSDDDFAKRIH
EANADQMCAVDVEDVLADAYTSPQNKKAIRQVVKVDDIQKAMGGVAPKYISIEFTRSEDRNPRRTISRQRQLENT
LKDTAKSLAKSINPELLSELDNAKSKKGLTDRLYLYFTQLGKDIYTGEPINIDELNKYDIDHILPQAFIKDNSLD
NRVLVLTAVNNGKSDNVPLRMFGAKMGHFWQLAEAGLISKRKLKNLQTDPTISKYAMHGFIRRQLVETSQVIKL
VANILGDKYRNDDTKIIETARMNHQMRDEFGFIKNREINDYHHAFFDAYLTAFLGRYLHYRIKLRPYFVYGDFKK
FREDKVTMRNFNHLHDLTDDTQEKIADAETGEVIWDRENSIQQLKDVYHYKFMLISHEVYTLRGAMFNQTVYPASD
AGKRKLIPVKADRPVNVYGGYSGSADAYMAIVRIHNKKGDYRVVGVPMRALDRLDAAKNVSDADFDRALKDVLAP
QLTKTKKSRKTGEITQVIEDFEIVLGKVMYRQLMIDGDKKFMLGSSTYQYNAKQLVLSQSVKTLASKGRDLPLQE
SMDYNVYTEILDKNVQYFSLYDMNKFRLHKLNLGFSKFI SFPNHNVLGDNTKVSSGKREILQEILNGLHANPTFGN
LKDVGITTPFGQLQPPNGILLSDETKIRYQSPTGLFERTVSLKDL (SEQ ID NO:164)

[0202] 在一些实施方案中,所述Cas9蛋白可以是无害李斯特菌 (*Listeria innocua*) Cas9,并且可以包含以下氨基酸序列或由其组成:

MKKPYTIGLDIGTNSVGWAVLTDQYDLVKKRMKIAGDSEKKQIKKNFWGVRLFDEGQTAADRRMARTA
RRRIERRRNRI SYLQGIFAEEMSKTDANFFCRLSDSFYVDNEKRNSRHPFFATIEEEVEYHKNYPTIYHLREELVN
SSEKADLRLVYLALAHIIKYRGNFLIEGALDTQNTSVDGIYKQFIQTYNQVFASGIEDGSLKKLEDNKDVAKILVE
KVTRKEKLERILKLYPGEKSAGMFAQFISLIVSGKGNFQKPFDLIEKSDIECAKDSYEEDLESLLALIGDEYAEFL
VAAKNAYSAVVLSSIIITVAETETNAKLSASMIERFDTHEEDLGELKAFIKLHLPKHYYEIFSNTTEKHGYAGYIDGK
TKQADFYKYMKMTLENIEGADYFIAKIEKENFLRKQRTFDNGAIPHQLHLEELEAILHQQAKYYPFLKENYDKIKS
LVTFRIPYFVGPLANGQSEFAWLTRKADGEIRPWNIEEKVDFGKSAVDFIEKMTNKDTYLPKENVLPKHSLCYQKY
LVYNELTKVRYINDQGKTSYFSGQEKEQIFNDLFKQKRKVKKKDLELFLRNMSHVESPTIEGLEDSFNSSYSTYHD
LLKVGIKQEILDNPVNTEMLNIVKILTVFEDKRMIEQLQQFSDVLDGVVLKKLERRHYTGWGRLSAKLLMGIRD
KQSHLTILDYLMNDDGLNRNLMQLINDSNLSFKSIIIEKEQVTTADKDIQSIVADLAGSPAIIKKGILQSLKIVDELV
SVMGYPPQTIVVEMARENQTTGKGKNSRPRYKSLEKAIKEFGSQILKEHPTDNQELRNRLYLYLQNGKDMYTG
QDLDIHNLSNYDIDHIVPQSFITDNSIDNLVLTSSAGNREKGDVPPLEIVRKRKVFWEKLYQGNLMSKRKFDYLT
KAERGGLTEADKARFIHRQLVETRQITKNVANILHQRFNIEKDDHGNTMKQVRIVTLKSALVSQFRKQFQLYKVRD
VNDYHHAHDAYLNGVVANTLLKVYPQLEPEFVYGDYHQFDWFKANKATAKKQFYTNIMLFFAQKDRIIDENGEILW
DKKYLDTVKKVMSYRQMNIKKTEIQKGEFSKATIKPKGNSSKLIPRKTNWDPMKYGGLDSPNMAYAVVIEYAKGK
NKLVEFKKIIIRVTIMERKAFKDEKAFLEEQGYRQPKVLAKLPKYTLYECEEGRRRMLASANEAAQGNQQLPNHL
VTLLHHAANCEVSDGKSLDYIESNREMFAELLAHVSEFAKRYTLAEANLNKINQLFEQNKEGDIKAIASFSVDLMA
FNAMGAPASFKFETTIERKRYNNLKELLNSTIIYQSITGLYESRKRLDD (SEQ ID NO:165)

[0203] 在一些实施方案中,所述Cas9蛋白可以是嗜肺军团菌Cas9,并且可以包含以下氨

基酸序列或由其组成:

MESSQILSPIGIDLGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRVRNKK
RNQFVKRVALQLFQHILSRDLNAKEETALCHYLNNGYTYVDTDLDEYIKDETTINLLKELLPSESEHNFIDWFLQ
KMQSSSEFRKILVSKVEEKDDKELKNAVKNIKNFITGFEKNSVEGHRHRKVYFENIKSDITKDNQLDSIKKKIPSV
CLSNLLGHLNQLWKNLHRYLAKNPKQFDEQTFGNEFLRMLKNFRHLKGSQESLAVRNLIQQLEQSQDYISILEKT
PPEITIPPYEARTNTGMEKDQSLLLNPEKLNLYPNWRNLIPGIIDAHPFLEKDLEHTKLDRKRRIISPSKQDEKR
DSYILQRYLDLNNKIDKFKIKKQLSFLGQKQLPANLIETQKEMETHFNSSLSVSVLIQIASAYNKEREDAAQGIWF
DNAFSLCELSNINPPRKQKILPLLVGAILSEDFINNKKDWAKFKIFWNTHKIGRTSLKSKCKEIEEARKNSGNAFK
IDYEEALNHPEHSNNKALIKIIQTIPDIIQAIQSHLGHNDSQALIYHNPFSLSQLYTILETKRDGFHKNCVAVTCE
NYWRSQKTEIDPEISYASRLPADSVRPFQGLARMMQRLAYEIAMAKWEQIKHIPDNSSLLIPIYLEQNRFEFEES
FKKIKGSSSDKTLEQAIEKQNIQWEEKFQRIINASMNICPYKGASIGGQGEIDHIYPRSLSKKHFGVIFNSEVNLI
YCSSLQGNREKKEEHYLLHLSPLYLKHQFGTDNVSDIKNFISQNVANIKKYISFHLLTPEQQAARHALFLDYDDE
AFKTITKFLMSQQKARVNGTQKFLGKQIMEFLSTLADSKQLQLEFSIKQITAEVVDHRELLSKQEPKLVKSQQS
FPSHAIDATLTMSIGLKEFPQFSQELDNSWFINHLMPDEVHLNPVRSKEKYNKPNISSTPLFKDSLVAERFIPVWV
KGETFAIGFSEKDLFEIKPSNKEKLFLLKTYSTKNPGESLQELQAKSKAKWLYFPINKTLALEFLHHYFHKEIVT
PDDTTVCHFINSRLRYTCKESITVKILKEPMPVLSVKFESSKKNVLGSFKHTIALPATKDWERLFNHPNFLALKAN
PAPNPKEFNEFIRKYFLSDNNPNSDIPNNGHNIPKQKHAKVRKVSPLVPGNAGTMMRIRKDNKGQPLYQLQTI
DDTPSMGIQINEDRLVKQEVLMDAYKTRNLSTIDGINNSEGQAYATFDNWLTLPVSTFKPEIIKLEMKPHSKTRRY
IRITQSLADFIKTIDEALMIKPSDSIDDPLNMPNEIVCKNKLFGNELKPRDGMKIVSTGKIVTYEFESDSTPQWI
QTLVYTQLKKQP (SEQ ID NO:166)

[0204] 在一些实施方案中,所述Cas9蛋白可以是嗜乳糖奈瑟球菌(*N.lactamica*) Cas9,并且可以包含以下氨基酸序列或由其组成:

MAAFKPNPMNYILGLDIGIASVGWAMVEVDEEENPIRLIDLGVRFERAEVPKTGDSLAMARRLARSV
RRLTRRRARHLLRARRLLKREGVLQDADFENGLVKSLPNTPWQLRAAALDRKLTCLWSAVLLHLVKHRYLSQR
KNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEINLLFEKQK
EFGNPHVSDGLKEDIETLLMAQRPALSGDAVQKMLGHCTFEPAPKAAKNTYTAERFIWLTCLNNLRILEQGSERP
LTDTERATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPL
NLSTELQDEIGTAFSLFKTDKIDTGRLLKDRVQPEILEALLKHISFDKVFQISLKLARRIVPLMEQGKRYDEACAEI
YGDHYCKKNAEEKIYLPPIPADEIRNPVVLRLSQARKVINCVVRRYGGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKAAAKFREYFPNFVGEPSKSDILKLRLYEQQHGKCLYSGKEINLVRLNEKGYVEIDHALPFSRTWDDSF
NNKVLVLGSENQNGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRIILLQKFDEEGFKERNLNDTRYVNR
LCQFVADHILLTGKGRVVFASNGQITNLLRGFWGLRKVRTENDRHHALDAVVVACSTVAMQKQITRFVRYKEMNA
FDGKTIDKETGEVLHQKAHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPL
FVSRAPNRKMSGQGHMETVKSARKLDEGISVLRVPLTQLKLKGLEKMNREREPPLYDALKALETHKDDPAKAF
EPFYKYDKAGSRTQQVKAVRIEQVQKTGVWRNHNGIADNATMVRVDVFEKGKYYLVPIYSWQVAKGILPDRVV
AFKDEEDWTVMDDSFVFLYANDLIKLTAKKNEFLGYFVSLNRATGAIDIRTHDSTDSTKGKNGIFQSVGVKTAL
SFQKNQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:167)

[0205] 在一些实施方案中,所述Cas9蛋白可以是脑膜炎奈瑟球菌Cas9,并且可以包含以下氨基酸序列或由其组成:

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRLARSV
RRLTRRAHRLLRARLLKREGVLQAADFENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFSRKDLQAEILLFEKQK
EFGNPHVSGGLKEGIIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERP
LTDTERTATLMDEPYRKSCLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPL
NLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTTEKIYLPPIPADEIRNPVVLRLASQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLVLGSENQKGNQTPYEYFNGKDNSREWQEFKARVETSFRPRSKKQRIILLQKFDEDFGKERNLNDTRYVNR
LCQFVADRMRLTGKGGKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTP
FVSRAPNRKMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLDLEKMNREREPPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKA VRVEQVQKTGVVVRNHNGIADNATMVRVDVFEKGDYLYLPIYSWQVAKGILPDR
AVVQKDEEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFYGFASCHRG TGNINIRIHDLDHKIGKNGILEGIGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:168)

[0206] 在一些实施方案中,所述Cas9蛋白可以是长双歧杆菌 (*B. longum*) Cas9,并且可以包含以下氨基酸序列或由其组成:

MLSRQLLGASHLARPVSYSYNVQDNDVHCSYGERCFMRGKRYRIGIDVGLNSVGLAAVEVSDENSPVR
LLNAQSVIHDGGVDPQKNKEAITRKNMSGVARRTRMRRRKRERLHKLDMLLGKFGYPVIEPESLDKPFEEWHVRA
ELATRYIEDDELRRESISIALRHMARHRGWRNPYRQVDSLISDNPYSKQYGELKEKAKAYNDATAAEEESTPAQL
VVAMLDAGYAEAPRLRWRTGSKKPAEGYLPVRLMQEDNANELKQIFRVQRVPADEWKPLFRSVFYAVSPKGS
AEQ RVGQDPLAPEQARALKASLAFQEYRIANVITNLRIKDASAELRKLTVDEKQSIYDQLVSPSSE
DITWSDLCDFLGF KRSQ LKGVGSLTEDGEERISSRPPRLTSVQRIYESDNKIRKPLVAWWSASDNEHEAMIRLLSNTVDIDK
VREDVA YASAI EFIDGLDDDALTKLDSVDLPSGRAAYSVELQKLTRQMLTTDDDLHEARKTLFNVTDSWRPPADPIGEPLG
NPSVDRVLKNVNRYL MNCQQRWGNPVSVNIEHVRSSFSVAFARKDKREYEKNEKRSIFRSSLSEQLRADEQMEK
VRESDLRRLEAIQRQNGQCLYCGRTITFRTCEMDHIVPRKGVGSTNTRTNFAAVCAECNRMKSNTPF
AIWARSEDA QTRGVSLAEAKKRVMTFTNPKSYAPREVKAFKQAVIARLQQTEDDAAIDNRSIESV
AWMADELHRRIDWYFNAKQ YVNSASIDDAEATMKTTSVVFQGRVTASARRAAGIEGKIHFIGQQSKTRLD
RRHHA VDASVIAMMNTAAAQTLME RESLRESQRLIGLMPGERSWKEYPYEGTSRYESFHLWLDNMDV
LLELLNDALDNDRIAVMQSQRYVLGNSIAHDAT IHPLEKVPLGSAMSADLIRRASTPALWCALTRLPD
YDEKEGLPEDSHREIRVHDTRY SADDEMGGFFASQAAQIAVQ EGSADIGSAIHHARVYRCWKTNAK
GVRKYFYGMIRVFQTDLLRACHDDLFTVPLPPQSISMRYGEPRVVQALQSGN AQYLGSLVVGDEIEMDF
SSLDVDGQIGEYLQFFSQFSGGNLAWKHVVVDGFFNQTLRIRPRYLAAEGLAKAFSDD VVPDGVQKIVTK
QGWLPPVNTASKTAVRIVRRNAFGEPRLSSAHMPCSWQWRHE (SEQ ID NO:169)

[0207] 在一些实施方案中,所述Cas9蛋白可以是嗜粘蛋白艾克曼菌 (*A. muciniphila*) Cas9,并且可以包含以下氨基酸序列或由其组成:

MSRSLTFSFDIGYASIGWAVIASASHDDADPSVCGCGTVLFPKDDCQAFKRREYRRLRRNIRSRRVRI
ERIGRLLVQAQIIITPEMKETSGHPAPFYLAASEALKGHRTLAPIELWHVLRWYAHNRGYDNNASWSNSLSE
DGGNGE DTERVKHAQDLMDKHGTATMAETICRELKEEGKADAPMEVSTPAYKNLNTAFPRLIVEKEVRR
ILELSAPLIPGL TAEIIELIAQHHLTTEQRGVLLQHGIKLARRYRGSLLFGQLIPRFDNRIISRCPVTWAQV
YEAE LKKGNSEQSAR

ERAEKLSKVPTANCPEFYEYRMARILCNIRADGEPLSAEIRRELMNQARQEGKLTAKASLEKAISSRLGKETETNVS
NYFTLHPDSEEALYLNPAVEVLQRSGIGQILSPSVYRIAANRLRRGKSVTPNYLLNLLKSRGESGEALEKKIEKES
KKKEADYADTPLPKPYATGRAPYARTVLKKVVEEILDGEDPTRPARGEAHPDGELKAHDGCLYCLDSTDSSVNQHQ
KERRLDTMTNHLVRHRMLILDRLLKDLIQDFADGQKDRISRVCEVVGKELTTFSAMDSKKIQRELTLRQKSHTDA
VNRLKRKLPGKALSANLIRKCRIAMDMNWTCPFTGATYGDHELENLELEHIVPHSFRQSNALSSLVLTWPGVNRMK
GQRTGYDFVEQEENPVDPKPNLHICSLNNYRELVEKLDDKKGHEDRRRKKKRRKALLMVRGLSHKHQSQNHEAMK
EIGMTEGMMTQSSHLMKLACKSIKTSPLDAHIDMIPGAVTAEVKAWDVFGVFKELCPEAADPDSGKILKENLRSL
THLHHALDACVLGLIPYIIPAHHNGLLRVLAMRRIPEKLIPQVRPVANQRHYVLNDDGRMMLRDLASLKENIRE
QLMEQRVIQHVPADMGGALLKETMQRVLSVDGSGEDAMVLSKKKDGKKEKNQVKASKLVGVFPEGPSKLKALKAA
IEIDGNYGVALDPKPVVIRHIKVFKRIMALKEQNGGKPVIRILKKGMLIHLTSSKDPKHAGVWRIESIQDSKGGVKL
DLQRAHCAVPKNKTHECNWREVDLISLLKKYQMKRYPTSYTGTPR (SEQ ID NO:170)

[0208] 在一些实施方案中,所述Cas9蛋白可以是兰氏臭杆菌 (*O. laneus*) Cas9,并且可以包含以下氨基酸序列或由其组成:

METTLGIDLGTNSIGLALVDQEEHQILYSGVRIFPEGINKDTIGLGEKEESRNATRRAKRQMRRQYFR
KKLRKAKLLELLIAYDMCPLKPEDVRRWKNWDKQKSTVRQFPDTPAFREWLKQNPYELRKQAVTEDVTRPELGRI
LYQMIQRRGFLSSRKGEKGIKFTGKDRMVGIDETRKNLQKQTLGAYLYDIAPKNGEKYRFRTERVRARYTLRDMY
IREFEIIWQRQAGHLGLAHEQATRKNIFLEGSATNVRNSKLITHLQAKYGRGHVLIEDTRITVTFLPLKEVLGG
KIEIEEEQLKFKSNESVLFWQRPLRSQKSLSKCVFEGRNFYDPVHQWIIAGPTPAPLSHPEFEEFRAYQFINNI
IYGKNEHLTAIQREAVFELMCTESKDFNFEKIPKHLKLFKFNFDTTKVPACTTISQLRKLFPHPVWEEKREEIW
HCFYFYDDNTLLFEKLQKDIALQTNDEKIKKIRLSESYGNVSLKAIRRNIPYLKGYAYSTAVLLGGIRNSFGKR
FEYFKEYEPEIEKAVCRILKEKNAEGEVIRKIKDYLHNRFGFAKNDRAFQKLYHHSQAITTQAQKERLPETGNLR
NPVQQLNELRRTVNKLLATCREKYGPSFKFDHIHVEMGRELRSSKTEREKQSRQIRENEKKNEAAKVLAEYGL
KAYRDNQKYLKYKEIEEKGGTVCCPYTGKTLNISHTLGSNSVQIEHIIPYSISLDDSLANKTLCDATFNREKGE
LTPYDFYQKDPSPKKGASSWEEIEDRAFRLLPYAKAQRFIRRKQESNEFISRQLNDTRYISKKAVEYLSAICSD
VKAFFPGLTAELRHLWGLNNILQSAPDITFPLPVSATENHREYVITNEQNEVIRLFPKQGETPRTEKGELLTGE
VERKVFRCCKGMQEFQTDVSDGKYWRRIKLSSSVTWSPLFAPKPIADGQIVLKGRIEKGVFVCNQLKQKLKTGLPD
GSYWISLPVISQTFKEGESVNNSKLTSQQVQLFGRVREGIFRCHNYQCPASGADGNFWCTLDTDTAQPAFTPIKNA
PPGVGGGQIILTGDVDDKGIFHADDLHYELPASLPKGKYYGIFTVESCDPTLIPIELSAKPTSKGENLIEGNIWV
DEHTGEVRFDPKKNREDQRHHAIDAIVIALSSQSLFQRLSTYNARRENKKRGLDSTEHFSPWPFGAQDVRQSVVP
LLVSYKQNPCTLCKISKTLKYDGKKIHSCGNAVGRQLHKETVYGQRTAPGATEKSYHIRKDIRELKTSHKIGKVVD
ITIRQMLLKHLQENYHIDITQEFNIPSNAFFKEGVYRIFLPNKHGEPVPIKKIRMKEELGNAERLKDNIQYVNPR
NNHHVMIYQDADGNLKEEIVSFWSVIERQNGGQPIYQLPREGRNIVSILQINDTFLIGLKEEPEVYRNDLSTLSK
HLYRVQKLSGMYTFRHHLASTLNNEREEFRIQSLEAWKRANPVKVQIDEIGRITFLNGPLC (SEQ ID NO:
171)。

[0209] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从CRISPR Cas蛋白分离或衍生的序列。在一些实施方案中,所述CRISPR Cas蛋白包含V型CRISPR Cas蛋白。在一些实施方案中,所述V型CRISPR Cas蛋白包含Cpf1蛋白。本公开文本的示例性Cpf1蛋白可以从任何物种分离或衍生,所述物种包括但不限于细菌或古菌。本公开文本的示例性Cpf1蛋白可以从任何物种分离或衍生,所述物种包括但不限于土拉热弗朗

西丝菌新凶手亚种、氨基酸球菌属物种 (*Acidaminococcus* sp.) BV3L6 和毛螺科细菌物种 (*Lachnospiraceae* bacterium sp.) ND2006。本公开文本的示例性 Cpf1 蛋白可以是核酸酶失活的。

[0210] 本公开文本的示例性野生型土拉热弗朗西丝菌新凶手亚种 Cpf1 (FnCpf1) 蛋白可以包含以下氨基酸序列或由其组成：

```
1 MSIQEFVNK YSLSKTLRFE LIPQGKTLEN IKARGLILDD EKRAKDYKKA KQIIDKYHQF
61 FIEEILSSVC ISEDLLQNYS DVYFKLKKSD DDNLQKDFKS AKDTIKKQIS EYIKDSEKFK
121 NLFNQNLIDA KKGQESDLIL WLKQSKDNGI ELFKANSDIT DIDEALEIIK SFKGWTTFYK
181 GFHENRKNVY SSNDIPTSII YRIVDDNLPK FLENKAKYES LKDKAPEAIN YEQIKKDLAE
241 ELTFDIDYKT SEVNQRVFSL DEVFEIANFN NYLNQSGITK FNTIIGGKFV NGENTKRKGI
301 NEYINLYSQQ INDKTLKKYK MSVLFKQILS DTESKSFVID KLEDDSDVVT TMQSFYEQIA
361 AFKTVEEKSI KETLSLLFDD LKAQKLDLSK IYFKNDKSLT DLSQQVFDDY SVIGTAVLEY
421 ITQQIAPKNL DNPSKKEQEL IAKKTEKAKY LSLETIKLAL EEFNKHRDID KQCRFEEILA
481 NFAAIPMIFD EIAQNKDNLQ QISIKYQNQG KKDLLQASAE DDVKAIKDLL DQTNNLLHKL
541 KIFHISQSED KANILDKDEH FYLVFEECYF ELANIVPLYN KIRNYITQKP YSDEKFKLNF
601 ENSTLANGWD KNKEPDNTAI LFIKDDKYLY GVMNKNNKI FDDKAIKENK GEGYKKIVYK
661 LLPGANKMLP KVFFSAKSIK FYNPSEDILR IRNHSTHTKN GSPQKGYEKF EFNIEDCRKF
721 IDFYKQSISK HPEWKDFGFR FSDTQRYNSI DEFYREVENQ GYKLTFENIS ESYIDSVVNQ
781 GKLYLFQIYN KDFSAYSKGR PNLHTLYWKA LFDERNLQDV VYKLNGEAEF FYRKQSIPKK
841 ITHPAKEAIA NKNKDNPKKE SVFEYDLIKD KRFTEDKFFF HCPITINFKS SGANKFNDEI
901 NLLLKEKAND VHILSIDRGE RHLAYYTLVD GKGNIKQDT FNIIGNDRMK TNYHDKLAAI
961 EKDRDSARKD WKKINNIKEM KEGYLSQVVH EIAKLIVIEYN AIVVFEDLNF GFKRGRFKVE
1021 KQVYQKLEKM LIEKLNLYLVF KDNEFDKTGG VLRAYQLTAP FETFKKMGKQ TGIIYYVPAG
1081 FTSKICPVTG FVNQLYPKYE SVSKSQEFFS KFDKICYNLD KGYFEFSFDY KNFGDKAAKG
1141 KWTIASFGSR LINFRNSDKN HNWDTREVPY TKELEKLLKD YSIEYGHGEC IKAACGESD
1201 KKFFAKLTSV LNTILQMRNS KTGTELDLYI SPVADVNGNF FDSRQAPKNM PQDADANGAY
1261 HIGLKGLMLL GRIKNNQEGK KLNLVIKNEE YFEFVQNRNN (SEQ ID NO: 172)。
```

[0211] 本公开文本的示例性野生型毛螺科细菌物种 ND2006 Cpf1 (LbCpf1) 蛋白可以包含以下氨基酸序列或由其组成：

1 AASKLEKFTN CYLSKTLRF KAIPVGKTQE NIDNKRLIVE DEKRAEDYKG VKKLLDRYYL
61 SFINDVLHSI KLKNLNYYIS LFRKKTRTEK ENKELENLEI NLRKEIAKAF KGAAGYKSLF
121 KKDIETILP EAADDKDEIA LVNSFNGFTT AFTGFFDNRE NMFSEEAKST SIAFRCINEN
181 LTRYISNMDI FEKVDAIFDK HEVQEIKEKI LNSDYDVEDF FEGEFFNFVL TQEGIDVYNA
241 IIGGFVTESG EKI KGLNEYI NLYNAKTKQA LPKFKPLYKQ VLSDRESLSF YGEGYTSDEE
301 VLEVFRNTLN KNSEIFSSIK KLEKLFKNFD EYSSAGIFVK NGPAISTISK DIFGEWNLIR
361 DKWNAEYDDI HLKKKAVVTE KYEDDRRSF KKIGSFSLEQ LQEYADADLS VVEKLKEIII
421 QKVDEIYKVY GSSEKLFAD FVLEKSLKKN DAVVAIMKDL LDSVKSFENY IKAFFGEGKE
481 TNRDESFYGD FVLAYDILLK VDHIYDAIRN YVTQKPYSKD KFKLYFQNPQ FMGGWDKDKE
541 TDYRATILRY GSKYYLAIMD KKYAKCLQKI DKDDVNGNYE KINYKLLPGP NKMLPKVFFS
601 KKWMAYYNPS EDIQKIYKNG TFKKGDMFNL NDCHKLIDFF KDSISRYPKW SNAYDFNFSE
661 TEKYKDIAGF YREVEEQGYK VSFESASKKE VDKLVEEGKL YMFQIYNKDF SDKSHGTPNL
721 HTMYFKLLFD ENNHGQIRLS GGAELFMRR SLKKEELVVH PANSPIANKN PDNPKKTTTL
781 SYDVYKDKRF SEDQYELHIP IAINKCPKNI FKINTEVRVL LKHDDNPYVI GIDRGERNLL
841 YIVVVDGKGN IVEQYSLNEI INNFGIRIK TDYHSLDDKK EKERFEARQN WTSIENIKEL
901 KAGYISQVVH KICELVEKYD AVIALEDLNS GFKNSRVKVE KQVYQKFEKM LIDKLNMYVD
961 KKSNPCATGG ALKGYQITNK FESFKSMSTQ NGFIFYIPAW LTSKIDPSTG FVNLLKTKYT
1021 SIADSKKFIS SFDRIMYVPE EDLFEFALDY KNFSRTDADY IKKWKLYSYG NRIRIFAAAK
1081 KNNVFAWEEV CLTSAYKELF NKYGINYQQG DIRALLCEQS DKAFYSSFMA LMSLMLQMRN
1141 SITGRTDVDF LISPVKNSDG IFYDSRNYEA QENAILPKNA DANGAYNIAR KVLWAIGQFK
1201 KAEDEKLDKV KIAISNKEWL EYAQTSVK (SEQ ID NO: 173)。

[0212] 本公开文本的示例性野生型氨基酸球菌属物种BV3L6 Cpf1 (AsCpf1) 蛋白可以包含以下氨基酸序列或由其组成：

```

1 MTQFEGFTNL YQVSKTLRFE LIPQGKTLKH IQEQGFIEED KARNDHYKEL KPIIDRIYKT
61 YADQCLQLVQ LDWENLSAAI DSYRKEKTEE TRNALIEEQA TYRNAIHDFY IGRTDNLTD
121 INKRHAIEYK GLFKAELFNG KVLKQLGTVT TTEHENALLR SFDKFTTYFS GFYENRKNVF
181 SAEDISTAIP HRIVQDNFPA FKENCHIFTR LITAVPSLRE HFENVKKAIG IFVSTSIEEV
241 FSFPFYNQLL TQTQIDLYNQ LLGGISREAG TEKIKGLNEV LNLAIQKNDE TAHIIASLPH
301 RFIPLFKQIL SDRNTLSFIL EEFKSDEEVI QSFCKYKTLL RNENVLETAE ALFNEINLSID
361 LTHIFISHKK LETISSALCD HWDTLRNALY ERRISELTGK ITKSAKEKVQ RSLKHEDINL
421 QEIIISAAGKE LSEAFKQKTS EILSHAHAAAL DQPLPTTLKK QEEKEILKSQ LDSLLGLYHL
481 LDWFAVDESN EVDPEFSARL TGIKLEMEPS LSFYNKARNY ATKKPYSVEK FKLNFQMPTL
541 ASGWDVNKEK NNGAILFVKN GLYYLGIMPQ QKGRYKALSF EPTEKTSEGF DKMYDYDFPD
601 AAKMIPKCST QLKAVTAHFQ THTTPILLSN NFIEPLEITK EIYDLNNPEK EPKKFQTAYA
661 KKTGDQKGYR EALCKWIDFT RDFLSKYTKT TSIDLSSLRP SSQYKDLGEY YAEINPLLYH
721 ISFQRIAEKE IMDAVETGKL YLFQIYNKDF AKGHHGKPNL HTLYWTGLFS PENLAKTSIK
781 LNGQAELFYR PKSRMKRMAH RLGEKMLNKK LKDQKTPIPD TLYQELYDYV NHRLSHDLSD
841 EARALLPNVI TKEVSHEIIK DRRFTSDKFF FHVPITLNYQ AANSPSKFNQ RVNAYLKEHP
901 ETPIIGIDRG ERNLIYITVI DSTGKILEQR SLNTIQQFDY QKKLDNREKE RVAARQAWSV
961 VGTIKDLKQG YLSQVIHEIV DLMIHQAVV VLENLNFQFK SKRTGIAEKA VYQQFEKMLI
1021 DKLNCLVLKD YPAEKVGGVL NPYQLTDQFT SFAKMGTSQG FLFYVPAPYT SKIDPLTGFV
1081 DPFVWKTIKN HESRKHFLG FDFLHYDVKT GDFILHFKMN RNLSFQRGLP GFMPAWDIVF
1141 EKNETQFQAK GTPFIAGKRI VPVIENHRFT GRYRDLYPAN ELIALLEEKI IVFRDGSNIL
1201 PKLLENDSSH AIDTMVALIR SVLQMRNSNA ATGEDYINSP VRDLNGVCDF SRFQNPWPM
1261 DADANGAYHI ALKGQLLLLH LKESKDLKLQ NGISNQDWLA YIQELRN (SEQ ID NO: 174)。

```

[0213] 在本公开文本的组合物的一些实施方案中,编码第一RNA结合蛋白的序列包含从CRISPR Cas蛋白分离或衍生的序列。在一些实施方案中,所述CRISPR Cas蛋白包含VI型CRISPR Cas蛋白或其部分。在一些实施方案中,所述VI型CRISPR Cas蛋白包含Cas13蛋白或其部分。本公开文本的示例性Cas13蛋白可以从任何物种分离或衍生,所述物种包括但不限于细菌或古菌。本公开文本的示例性Cas13蛋白可以从任何物种分离或衍生,所述物种包括但不限于韦德纤毛菌(*Leptotrichia wadei*)、西尔李斯特菌血清变型(*Listeria seeligeri* serovar)1/2b(菌株ATCC 35967/DSM 20751/CIP 100100/SLCC 3954)、毛螺科细菌、嗜氨基梭菌(*Clostridium aminophilum*) DSM10710、鸡肉杆菌(*Carnobacterium gallinarum*) DSM 4847、产丙酸沼杆菌(*Paludibacter propionigenes*) WB4、韦氏李斯特菌(*Listeria weihenstephanensis*) FSL R9-0317、韦氏李斯特菌FSL R9-0317、细菌FSL M6-0635(纽约李斯特菌(*Listeria newyorkensis*))、韦德纤毛菌F0279、荚膜红细菌

(*Rhodobacter capsulatus*) SB 1003、荚膜红细菌R121、荚膜红细菌DE442和溃疡棒状杆菌(*Corynebacterium ulcerans*)。本公开文本的示例性Cas13蛋白可以是DNA核酸酶失活的。本公开文本的示例性Cas13蛋白包括但不限于Cas13a、Cas13b、Cas13c、Cas13d及其直系同源物。本公开文本的示例性Cas13b蛋白包括但不限于亚型1和2,在本文中分别称为Csx27和Csx28。

[0214] 示例性Cas13a蛋白包括但不限于:

Cas13a 编号	Cas13a 缩写	生物名称	登录号	同向重复序列
Cas13a1	LshCas13a	沙氏纤毛菌 (<i>Leptotrichia shahii</i>)	WP_018451595.1	CCACCCCAATATCGAAGGGGACTAA AAC (SEQ ID NO: 175)
Cas13a2	LwaCas13a	韦德纤毛菌	WP_021746774.1	GATTTAGACTACCCCAAAAACGAAG GGGACTAAAAC (SEQ ID NO: 176)
Cas13a3	LseCas13a	西尔李斯特菌	WP_012985477.1	GTAAGAGACTACCTCTATATGAAAG AGGACTAAAAC (SEQ ID NO: 177)
Cas13a4	LbmCas13a	毛螺科细菌 MA2020	WP_044921188.1	GTATTGAGAAAAGCCAGATATAGTT GGCAATAGAC (SEQ ID NO: 178)
Cas13a5	LbnCas13a	毛螺科细菌 NK4A179	WP_022785443.1	GTTGATGAGAAGAGCCCAAGATAGA GGGCAATAAC (SEQ ID NO: 179)
Cas13a6	CamCas13a	嗜氨基[梭菌] DSM 10710	WP_031473346.1	GTCTATTGCCCTCTATATCGGGCTGT TCTCCAAAC (SEQ ID NO: 180)
Cas13a7	CgaCas13a	鸡肉杆菌 DSM 4847	WP_034560163.1	ATTAAAGACTACCTCTAAATGTAAG AGGACTATAAC (SEQ ID NO: 181)

Cas13a8	Cga2Cas13a	鸡肉杆菌 DSM 4847	WP_034563842.1	AATATAAACTACCTCTAAATGTAAG AGGACTATAAC (SEQ ID NO: 182)
Cas13a9	Pprcas13a	产丙酸沼杆菌 WB4	WP_013443710.1	CTTGTGGATTATCCCAAAATTGAAG GGAAGTACAAC (SEQ ID NO: 183)
Cas13a10	LweCas13a	韦氏李斯特菌 FSL R9-0317	WP_036059185.1	GATTTAGAGTACCTCAAAATAGAAG AGGTCTAAAAC (SEQ ID NO: 184)
Cas13a11	LbfCas13a	李斯特菌科 (Listeriaceae) 细菌 FSL M6-0635 (纽 约李斯特菌)	WP_036091002.1	GATTTAGAGTACCTCAAAACAAAAG AGGACTAAAAC (SEQ ID NO: 185)
Cas13a12	Lwa2cas13a	韦德纤毛菌 F0279	WP_021746774.1	GATATAGATAACCCCAAAAACGAAG GGATCTAAAAC (SEQ ID NO: 186)
Cas13a13	RcsCas13a	荚膜红细菌 SB 1003	WP_013067728.1	GCCTCACATCACCGCCAAGACGACG GCGGACTGAAC (SEQ ID NO: 187)
Cas13a14	RcrCas13a	荚膜红细菌 R121	WP_023911507.1	GCCTCACATCACCGCCAAGACGACG GCGGACTGAAC (SEQ ID NO: 188)
Cas13a15	RcdCas13a	荚膜红细菌 DE442	WP_023911507.1	GCCTCACATCACCGCCAAGACGACG GCGGACTGAAC (SEQ ID NO: 189)

[0215] 本公开文本的示例性野生型Cas13a蛋白可以包含以下氨基酸序列或由其组成:

```

1  MGNLFQHKRW YEVRDKKDFK IKRKVKVKN YDGNKYILNI NENNNKEKID NKNFIRKYIN
61  YKKNNDILKE FTRKFHAGNI LFKLKGKEGI IRIENDDDFL ETEEVVLYIE AYGKSEKLKA
121 LGITKKKIID EAIRQGITKD DKKIEIKRQE NEEIEIDIR DEYTNKTLND CSIILRIEN
181 DELETKSIY EIFKNINMSL YKIEEKIEN ETEKVFENRY YEEHLREKLL KDDKIDVILT
241 NFMEIREKIK SNLEILGFVK FYLNVGDDKK KSKNKKMLVE KILNINVDLT VEDIADFKV
301 ELEFWNITKR IEKVKKVNE FLEKRRNRTY IKSIVLLDKH EKFKIERENK KDKIVKFFVE
361 NIKNNSIKEK IEKILAEFKI DELIKKLEKE LKKGNCDEI FGIFKKHYKV NFDSKKFSKK
421 SDEEKELYKI IYRYLKGRIE KILVNEQKVR LKKMEKIEIE KILNESILSE KILKRVKQYT
481 LEHIMYLGKL RHNDIDMTTV NTDDFSRLHA KEELDLELIT FFASTNMELN KIFSRENINN

```

541 DENIDFFGGD REKNYVLDKK ILNSIKIIR DLDIFDNKNN ITNNFIRKFT KIGTNERNRI
 601 LHAISKERDL QGTQDDYNKV INIIQNLKIS DEEVSKALNL DVVFKDKKNI ITKINDIKIS
 661 EENNNDIKYL PSFSKVLPEI LNLYRNNPKN EPFDTIETEK IVLNALIYVN KELYKKLILE
 721 DDLEENESKN IFLQELKCTL GNIDEIDENI IENYYKNAQI SASKGNNAI KKYQKKVIEC
 781 YIGYLRKNYE ELFDSDFKM NIQEIKKQIK DINDNKTYER ITVKTSDKTI VINDDFEYII
 841 SIFALLNSNA VINKIRNRFF ATSVWLNTSE YQNIIDILDE IMQLNTRLNE CITENWNLNL
 901 EEFIQMKKEI EKDFDDFKIQ TKKEIFNNYY EDIKNNILTE FKDDINGCDV LEKKLEKIVI
 961 FDEETKFEID KKSNIQDEQ RKLSNINKKD LKKKVDQYIK DKDQEIISKI LCRIIFNSDF
 1021 LKKYKKEIDN LIEDMESENE NKFQEIIYPK ERKNELYIYK KNLFNLNIGNP NFDKIYGLIS
 1081 NDIKMADAKF LFNIDGKNIR KNKISEIDAI LKNLNDKLNQ YSKEYKEYI KKLKENDDFE
 1141 AKNIQNKNYK SFEKDYNRVS EYKKIRDLVE FNYLNKIESY LIDINWKLAI QMARFERDMH
 1201 YIVNGLRELQ IIKLSGYNTG ISRAYPKRNG SDGFYTTTAY YKFFDEESYK KFEKICYGFG
 1261 IDLSENSEIN KPENESIRNY ISHFYIVRNP FADYSIAEQI DRVSNLLSYS TRYNNSTYAS
 1321 VFEVFKKDVN LDYDELKKKF KLIGNNDILE RLMKPKKVSF LELESYNSDY IKNLIIEELLT
 1381 KIENTNDTL (SEQ ID NO: 190)

[0216] 示例性Cas13b蛋白包括但不限于：

物种	Cas13b登录号	Cas13b大小(aa)
产丙酸沼杆菌WB4	WP_013446107.1	1155
普雷沃菌属物种P5-60	WP_044074780.1	1091
普雷沃菌属物种P4-76	WP_044072147.1	1091
普雷沃菌属物种P5-125	WP_044065294.1	1091
普雷沃菌属物种P5-119	WP_042518169.1	1091
犬咬二氧化碳嗜纤维菌 (Capnocytophaga canimorsus) Cc5	WP_013997271.1	1200
厦门褐指藻杆菌 (Phaeodactylibacter xiamenensis)	WP_044218239.1	1132
牙龈卟啉单胞菌W83	WP_005873511.1	1136
牙龈卟啉单胞菌F0570	WP_021665475.1	1136
牙龈卟啉单胞菌ATCC 33277	WP_012458151.1	1136
牙龈卟啉单胞菌F0185	ERJ81987.1	1136
牙龈卟啉单胞菌F0185	WP_021677657.1	1136

牙龈卟啉单胞菌SJD2	WP_023846767.1	1136
牙龈卟啉单胞菌F0568	ERJ65637.1	1136
牙龈卟啉单胞菌W4087	ERJ87335.1	1136
牙龈卟啉单胞菌W4087	WP_021680012.1	1136
牙龈卟啉单胞菌F0568	WP_021663197.1	1136
牙龈卟啉单胞菌	WP_061156637.1	1136
咽喉卟啉单胞菌	WP_039445055.1	1136
酿脓拟杆菌 (Bacteroides pyogenes) F0041	ERI81700.1	1116
酿脓拟杆菌JCM 10003	WP_034542281.1	1116
别样杆菌属物种 (Alistipes sp.) ZOR0009	WP_047447901.1	954
嗜鳃黄杆菌 (Flavobacterium branchiophilum) FL-15	WP_014084666.1	1151
普雷沃菌属物种MA2016	WP_036929175.1	1323
拟香味类香味菌CCUG 10230	EHO06562.1	1160
拟香味类香味菌CCUG 3837	EKB06014.1	1158
拟香味类香味菌CCUG 3837	WP_006265509.1	1158
拟香味类香味菌CCUG 12901	WP_006261414.1	1158
拟香味类香味菌CCUG 12901	EHO08761.1	1158
拟香味类香味菌 (NZ_CP013690.1)	WP_058700060.1	1160
动物溃疡伯格菌ATCC 43767	EKB54193.1	1225
狗咬二氧化碳嗜纤维菌 (Capnocytophaga cynodegmi)	WP_041989581.1	1219
动物溃疡伯格菌ATCC 43767	WP_002664492.1	1225
黄杆菌属物种 (Flavobacterium sp.) 316	WP_045968377.1	1156
扭曲冷弯曲菌 (Psychroflexus torquis) ATCC 700755	WP_015024765.1	1146
柱状黄杆菌ATCC 49512	WP_014165541.1	1180
柱状黄杆菌	WP_060381855.1	1214
柱状黄杆菌	WP_063744070.1	1214
柱状黄杆菌	WP_065213424.1	1215
金黄杆菌属物种 (Chryseobacterium sp.) YR477	WP_047431796.1	1146

鸭疫里默氏杆菌 (<i>Riemerella anatipestifer</i>) ATCC 11845 = DSM 15868	WP_004919755.1	1096
鸭疫里默氏杆菌 RA-CH-2	WP_015345620.1	949
鸭疫里默氏杆菌	WP_049354263.1	949
鸭疫里默氏杆菌	WP_061710138.1	951
鸭疫里默氏杆菌	WP_064970887.1	1096
解糖普雷沃菌 (<i>Prevotella saccharolytica</i>) F0055	EKY00089.1	1151
解糖普雷沃菌 JCM 17484	WP_051522484.1	1152
颊普雷沃菌 ATCC 33574	EFU31981.1	1128
颊普雷沃菌 ATCC 33574	WP_004343973.1	1128
颊普雷沃菌 D17	WP_004343581.1	1128
普雷沃菌属物种 MSX73	WP_007412163.1	1128
苍白普雷沃菌 ATCC 700821	EGQ18444.1	1126
苍白普雷沃菌 ATCC 700821	WP_006044833.1	1126
中间普雷沃菌 (<i>Prevotella intermedia</i>) ATCC 25611 = DSM 20706	WP_036860899.1	1127
中间普雷沃菌	WP_061868553.1	1121
中间普雷沃菌 17	AFJ07523.1	1135
中间普雷沃菌	WP_050955369.1	1133
中间普雷沃菌	BAU18623.1	1134
中间普雷沃菌 ZT	KJJ86756.1	1126
桔红色普雷沃菌 (<i>Prevotella aurantiaca</i>) JCM 15754	WP_025000926.1	1125
胸膜炎普雷沃菌 (<i>Prevotella pleuritidis</i>) F0068	WP_021584635.1	1140
胸膜炎普雷沃菌 JCM 14110	WP_036931485.1	1117
弗氏普雷沃菌 (<i>Prevotella falsenii</i>) DSM 22864 = JCM 15124	WP_036884929.1	1134
咽喉卟啉单胞菌	WP_039418912.1	1176
卟啉单胞菌属物种 (<i>Porphyromonas</i> sp.) COT-052 OH4946	WP_039428968.1	1176
咽喉卟啉单胞菌	WP_039442171.1	1175

咽喉卟啉单胞菌	WP_039431778.1	1176
咽喉卟啉单胞菌	WP_046201018.1	1176
咽喉卟啉单胞菌	WP_039434803.1	1176
咽喉卟啉单胞菌	WP_039419792.1	1120
咽喉卟啉单胞菌	WP_039426176.1	1120
咽喉卟啉单胞菌	WP_039437199.1	1120
牙龈卟啉单胞菌TDC60	WP_013816155.1	1120
牙龈卟啉单胞菌ATCC 33277	WP_012458414.1	1120
牙龈卟啉单胞菌A7A1-28	WP_058019250.1	1176
牙龈卟啉单胞菌JCVI SC001	EOA10535.1	1176
牙龈卟啉单胞菌W50	WP_005874195.1	1176
牙龈卟啉单胞菌	WP_052912312.1	1176
牙龈卟啉单胞菌AJW4	WP_053444417.1	1120
牙龈卟啉单胞菌	WP_039417390.1	1120
牙龈卟啉单胞菌	WP_061156470.1	1120

[0217] 本公开文本的示例性野生型动物溃疡伯格菌ATCC 43767Cas13b (BzCas13b) 蛋白可以包含以下氨基酸序列或由其组成:

```

1 menktslgnn iyympfkpqq ksyfagyfna amentdsvfr elgkrlkgke ytsenffdai
61 fkenislvey eryvklldsy fpmarlldkk evpikerken fkknfkgiik avrdlnrfyt
121 hkehgeveit deifgvldem lkstvlvtvkk kkvkttdktke ilkksiekql dilcqkkley
181 lrdtarkiee krrnqrerge kelvapfkys dkrddliaai yndafdvaid kkkdskess
241 kakyntksdp qqeegdlkip iskngvvfl1 slfltkqeih afkskiagfk atvideatvs
301 eatvshgkns icfmatheif shlaykklkr kvrtaeinyg eaenaeqlsv yaketlmmqm
361 ldelskvpdv vyqnlsedvq ktfieldwney lkenngdvgt meeeqvihpv irkryedkfn
421 yfairfldef aqfptlrfqv hlgnylhdsr pkenlisdr r ikekitvfgr lselehkkal
481 fikntetned rehyweifpn pnydfpkeni svndkdfpia gsildrekqp vagkigikvk
541 llngqyvsev dkavkahqlk qrkaskpsiq niieivpin esnpkeaivf ggqptaylsm
601 ndihsilyef fdkwekkkek lekkgelr keigkelekk ivgkiqaiq qiidkdnak
661 ilkpyqdgns taidkeklik dlkqeqnilq klkdeqtvre keyndfiayq dknreinkvr
721 drnhkqylkd nlrkypeap arkevlyyre kgkvavwlan dikrfmptdf knewkgeqhs

```

781 llqkslayye qckeelknll pekvfqhlpf klgyfqqky lyqfytcyld krleyisglv
 841 qqaenfksen kvfkkvenec fkflkkqnyt hkeldarvqs ilgypifler gfmdekptii
 901 kgktfkgnea lfadwfryyk eyqnfqtfyd tenyplvele kkqadrkrkt kiyqqkndv
 961 ftllmakhif ksvfkqdsid qfsledlyqs reerlgnqr arqtgerntn yiwnktvdlk
 1021 lcdgkitven vklknvgdfi kyeydqrva flkyeeniew qafilkeske eenypyvver
 1081 eieqyekvrr eellkevhli eeyilekvkd keilkkgdnq nfkyiilngl lkqlknedve
 1141 sykvfnlnte pedvninqlk qeatdleqka fvltyirnkf ahnqlpkkef wdycqekygk
 1201 iekektayaey faevfkkeke alik (SEQ ID NO: 191)。

[0218] 在本公开文本的组合物的一些实施方案中,编码所述第一RNA结合蛋白或RNA指导的靶RNA结合蛋白的序列包含从CasRX/Cas13d蛋白分离或衍生的序列。CasRX/Cas13d是VI-D型CRISPR-Cas系统的效应子。在一些实施方案中,所述CasRX/Cas13d蛋白是可以切割或结合RNA的RNA指导的RNA内切核酸酶。在一些实施方案中,所述CasRX/Cas13d蛋白可以包括一个或多个高等真核生物和原核生物核苷酸结合 (HEPN) 结构域。在一些实施方案中,所述CasRX/Cas13d蛋白可以包括野生型或突变的HEPN结构域。在一些实施方案中,所述CasRX/Cas13d蛋白包括无法切割RNA但可以加工指导RNA的突变的HEPN结构域。在一些实施方案中,所述CasRX/Cas13d蛋白不需要原间隔子侧翼序列。

[0219] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群6049000251 (CasRX/Cas13d Gut_metagenome_contig6049000251):

LYLTSFGKGN AAVIEQKIEP ENGYRVTGMQ ITPSITVNKA TDESVRFRVK RKIAQKDEFI	60
ADNPMHEGRH RIEPSAGSDM LGLKTKLEKY YFGKEFDDNL HIQIIYNILD IEKILAVYST	120
NITA	124

(SEQ ID NO: 54)。

[0220] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群546000275:

MDSYRPKLYK LIDFCIFKHY HEYTEISEKN VDTLRAAVSE EQKESFYADE AKRLWGIFDK	60
QFLGFCKKIN VWVNGSHEKE ILGYIDKDAY RKKSDVSYFS KFLYAMSFFL DGKEINDLLT	120
TLINKFDNIA SFISTAKELD AEIDRILEKK LDPVTGKPLK GKNSFRNFIA NNVIENKRFI	180
YVIKFCNPKN VLKLVKNTKV TEFVLKRMPE SQIDRYSSC IDTEKNPSVD KKISDLAEMI	240
KKIAFDDFRN VRQKTRTREE SLEKERFKAV IGLYLTVVYL LIKNLVNVNS RYVMAFHCL	300
RDAKLYGINI GKNYIELTED LCRENENSRS AYLARNKRLR DCVKQNI DNA KNMKSKEK	358

(SEQ ID NO: 57)。

[0221] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群4114000374:

DTKINPQTWL YQLENTPLDL NEYRDTLDHF FDERFNEINE HFVTQATNL CIMKEVFPDE	60
DFKSIADLYY DFIVVKSYPN IGFSIKKLR KMLELPEAKR VTSTEMDSVR SKLYKLIDFC	120
IFKHHEKPE TVEMIVSMLR AYTSEDME	149

(SEQ ID NO: 61)。

[0222] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群721000619：

KEGSTMAKNE KKKSTAKALG LKSSFVNND IYMTSFGKGN KAVLEKKITE NTIENKSDTT	60
YFDVINRDPK GFTLEGRRIA DMTAFSNDPK YHVNVDNGKF LEDQLGARSE LEKKVFGRTF	120
DDNVHIQLIH NILDIEKIMA QYVSDIVYLL HNTIKRDMND DIMGYISIRN SFDDFCHPER	180
IPDRKAKDNL QKQHDIFDE ILKCGRLAYF GNAFFEDGSD NKEIAKLKRY KEIYHIIALM	240
GSLRQSYFHG ENSDKNFQGP TWAYTLESNL TGKYKEFKDT LDKTFDERYE MISKDFGSTN	300
MVNLQILEEL LKMLYGNVSP	320

(SEQ ID NO: 67)。

[0223] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群2002000411：

EKQNKAKYQA IISLYLMVMY QIVKNMIYVN SRYVIAFHCL ERDSNQLLGR FNSRDASMYN	60
KLTKQFITDK YLNDGAQGCS KKVGNLYLSHN ITCCSDELK EYRNQVDHFA VVRMIGKYAA	120
DIGKFSTWFE LYHYVMQRRI FDKRNPLSET ERTYKQLIAK HHTYCKDLVK ALNTPFGYNL	180
ARYKNLSIGE LFDRNNYNK TKET	204

(SEQ ID NO: 69)。

[0224] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群1352000311：

LIDFLIYDLY YNRKPARIEE IVDKLRESVN DEEKESIYSA ETKYVYEALG KVLVRSKKY	60
LNGATIRDLK NRYDAKTANR IWDISEHSKS GHVNCFCCLI YMTLMLDGK EINDLLTTLV	120
NKFDNIASFI DVMDELGLEH SFTDNYKMFA DSKAICLDLQ FINSFARMSK IDDEKSKRQL	180
FRDALVVLDI GDKNEDWIEK YLTSDIFKRD ENGNKIDGK RDFRNFIANN VIKSARFKYL	240
VKYSSADGMI KKKNEKLIS FVLEQLPETQ IDRYYESCGL DCAVADRKVR IEKLTGLIRD	300
MRFDNFRGVN YSNDACKKDK QAKAKYQAI SLYLMVLYQI VKNMIYVNSR YVIAFHCLER	360
DLLFFNIELD NSYQYSNCNE LTKFIKDKY MKEGALGFNM KAGRYLTKNI GNCSNELRKI	420
YRNQVDHFAV VRKIGNYAAD IASVGSWE	449

(SEQ ID NO: 71)。

[0225] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群10037000527：

YMDQNFANS DAWAIHVYRNK IQHLDVVRHA DMYIGDIREF HSWFELYHYI IQRRIIDQYA 60
YESTPGSSRD GSAIIDEERL NPATRRYFRL ITTYKT 96

(SEQ ID NO: 72)。

[0226] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群238000329：

RYDKDRSKIY TMMDFVIYRY YIDNNNSID FINKLRSSID EKSKEKLYNE EANRLWNKLG 60
EYMLYIKEFN GKLASRTPDR DGNISEFVES LPKIHRLLP GQKISNFSKL MYLLTMFLDG 120
KEINDLLTTL INKFENIQGF LDIMPEINVN AKFEPEYVFF NKSHEIAGEL KLIKGFQMG 180
EPAATLKLEM TADAIKILGT EKEDAELIKL AESLFKDENG KLLGNKQHGM RNFIGNNVIK 240
SKRFHYLIRY GDPAHLHKIA TNKNVVRFLV GRIADMQKKQ GQKGKNQIDR YYEVCVGND 300
IKKTIEEKID ALTDIIVNMN YDQFEKKKAV IENQNRGKTF EEKNKYKRDN AEREKFKKII 360
SLYLTVIYHI LKNIVNVNSR YILGFHCLEK DKQLYIEKYN KDKLDGFVAL TKFCLGDEER 420
YEDLKAKAQA SIQALETANP KLYAKYMNYS DEEKKEEFKK QLNRRERVKNA RNAYLKNIKN 480
YIMIRLQLRD QTDSSGYLCG EFRDKVAHLE VARHAHEYI 519

(SEQ ID NO: 73)。

[0227] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群2643000492：

NGEIVSLAEK EAFSAKIADK NIGCKIENKQ FRHPKGYDVI ADNPIYKGSP RQDMLGLKET 60
LEKRYFSPSD SIDNVRVQVA HNILDIEKIL AEYITNAVYS FDNIAGFGKD IIGDDFSPVY 120
TYDKFEKSDR YEYFKNLLNN SRLGYYGQAF FECDDSKENK KKKDAIKCYN IIALLSGLRH 180
W 181

(SEQ ID NO: 84)。

[0228] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群874000057：

MSKNKESYAK GMGLKSALVS GSKVYMTSFE GGNDKLEKV VENSEIVSLA EKESFSAEIF 60
KKNIGCKIEN KKFHKPKRYD VIADNPLYKG SVRQDMLGLK ETLKRYFNS ADGTDNVCIQ 120
VIHNILDIEK ILAEYITNAV YSFDNIAGFG EDIIGMGGFK PIYTYKQFKE PDKYNKKFDD 180
ILNNSRLGYY GKAFTEKNDL KHNPNNKKRD KNPYILKYDN ECYIIALLS GLRHWNIHSH 240
AKDDLVSRYW LYNLDSILNR EYISTLNYLY DDIADDELTS FSKNSSANVN YIAETLNIDP 300
SEFAQQYFRF SIMKEQKNMG FNVSKLREIM LDRKELSDIR DNHRVFDSIR SKLYTMMDFV 360
IYRYYIEEAA KTEAENRNLP ENEKKISEKD FFVINLRGSF DENQKEKLYI EEAKRLWEKL 420
KDIMLKIKEF RGEKVKEYKK 440

(SEQ ID NO: 85)。

[0229] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群4781000489:

LDKQLDYEYI RTLNMFNDI ADELTRTF SK NSAANVNYIA ETLNIDPNKF AEQYFRFSIM	60
KEQKNLGFNL TKLRESMLDR RELSDIRDNH NVFDSIRPKL YTMMDFVIYK HYIDEAKKTE	120
AENKSLPDDR KNLSEKD	137

(SEQ ID NO: 86)。

[0230] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群12144000352:

RMGEPVANTK RVMIDAVKI LGTDLSDEL KEMADSF FSKD SDGNLLKKKG HGMRNFITNN	60
VIKNKRFHYL IRYGDP AHLH EIAKNEA	87

(SEQ ID NO: 87)。

[0231] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群5590000448:

VHNNEEKDLI KYTWLYNL DK YLDAEYITL NYMYNDIGDE LTDSFSKNSA ANINYIAETL	60
GIDPKTF AEQ YFRFSIMKEQ KNLGFNLTKL REVMLDRKDM SEIRENHND F DSIRAKVYTM	120
MDFVIYRYII EEA AKVNAAN KSLPDNEKSL SEKDIFVISL RGSFNEDQKD RLYYDEA QRL	180
WSKVGKLMKL IKKFRGKDTR KYKNMGTPRI RRLIPEGRDI STFSKL MYAL TMFLDGKEIN	240
DLTTLINKF DNIQSFLKVM PLIGVNAKFA EEYSFFNNSE KIADELRLIK SFARMGEPVA	300
DARRAMYIDA IRILGTDLS DELKALADSF SLDENGKLG KGKHGMRNFI INNVITNKRF	360
HYLIRYGNPV HLHEIAKNEA VVKFVLGRIA DIQKKQGQNG KNQIDRYYET CIGK	414

(SEQ ID NO: 88)。

[0232] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群525000349:

MSKKENRKS Y VKGLGLKSTL VSDSKVYLT FADGSNAKLE KCVENNKIIC ISNDKEAFAA	60
SIANKNVGYK IKNDEKFRHP KGYDIISNP LLHNNSVQQD MLGLKNVLEK RYFGKSSGGD	120
NNLCIQIIHN IIDIEKILSE YIPNVVYAFN NIAGFKDEHN NIIDIIGTQT YNSSYTYADF	180
SKDKSDKKYI EFQKLLKNKR LGYWGKAFFT GQGNNAKVRQ ENQCFHIIAL LISLRNWATH	240
SNELDKHTKR TWLYKLDDTN ILNAEYVKTL NYLYDTIADE LTKSFSKNGA VNVNYLAKKY	300
NIKDDLPGFS EQYFRFSIMK EQKNLGFNIS KLRENMLDFK DMSVI	345

(SEQ ID NO: 89)。

[0233] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d肠道_宏基因组_重叠群7229000302:

KKISSLTKFC LGESDEKKLK ALAKKSLEEL KTTNSKLYEN YIKYSDERKA EEAKRQINRE	60
RAKTAMNAHL RNTKWNDIMY GQLKDLADSK SRICSEFRNK AAHLEVARYA HMYINDISEV	120
KSYFRLYHYI MQRRIIDVIE NNPKAKYEGK VKVYFEDVKK NKKYNKNLLK LMCVPFGYCI	180
PRFKNLSIEQ MFDNETDNS DKKKEK	206

(SEQ ID NO: 90)。

[0234] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群3227000343：

IGDISEVNSY FQLYHYIMQR ILIDKIGSKT TGKAKEYFDS VIVNKKYDDR LLKLLCSPLG	60
YCLTRYKDLS IEALFDMNEA AKYDKLNKER KNKKK	95

(SEQ ID NO: 91)。

[0235] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_重叠群7030000469：

SIRSKLYTMM DFVIYRYIE ESAKAAENK PSESDFVIR LRGSFNENQK EELYIEEAER	60
LWKKFGEIML KIKEFRGEKV KEYKKEVPRI ERILPHGKDI SAFSKLMYML SMFLD	115

(SEQ ID NO: 92)。

[0236] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d肠道_宏基因组_P17E0k2120140920,_c87000043：

MYFSKMIYML TYFLDGKEIN DLLTTLISKF DNIKEFLKIM KSSAVDVECE LTAGYKLFND	60
SQRITNELFI VKNIASMRKP AASAKLTMFR DALITLGIDD KITDDRISSEI LKLKEKGKGI	120
HGLRNFITNN VIESSRFVYL IKYANAQKIR EVAKNEKVVM FVLGGIPDTQ IERYKSCVE	180
FPDMNSSLEA KRSELARMIK NISFDDFKNV KQQAKGRENK AKERAKAVIG LYLT	234

(SEQ ID NO: 93)。

[0237] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OBVH01003037.1,人肠宏基因组序列(也发现于WGS重叠群emb|OBXZ01000094.1|和emb|OBJF01000033.1|)：

MAKKKRITAK ERKQNHRELL MKKADSNAEK EKAKKPVVEN KPDTAISKDN TPKPNKEIKK	60
SKAKLAGVKW VIKANDDVAY ISSFGKGNNNS VLEKRIMGDV SSVNKNDSHM YVNPKYTKKN	120
YEIKNGFSSG SSLVTYPNKP DKNSGMDALC LKPYFEKDFE GHIFTDNMHI QAIYNIFDIE	180
KILAKHITNI IYTVNSFDRN YNQSGNDTIG FGLNYRVPYS EYGGGKDSNG EPKNQSKWEK	240
RDNFIKFYNE SKPHLGYYEN IFYDHGEPIS EEKFYNYLNI LNFIRNNTFH YKDDDIELYS	300
ENYSEEFVFI NCLNKFVKNK FKNVNKNFIS NEKNNLYIIL NAYGKDTENV EVVKKYSKEL	360
YKLSVLKTNK NLGVNVKKLR ESAIEYGYCP LPYDKEKEVA KLSSVKKHLY KTYDFVITHY	420
LNSNDKLLLE IVETLRLSKN DDEKENVYKK YAEKLFKADD VINPIKAISK LFARKGNKLF	480
KEKIIIKKEY IEDVSIKNI YDFTKVIFFM TCFLDGKEIN DLLTNIISKL QVIEDHNNVI	540
KFISNNKDAV YKDYSKYAI FRNAGKATE LEAIIKSIARM ENKIENAPQE PLLKDALLSL	600
GVSDDTKVL NTKYFYFDSK EKTDKQSQKV STFLMNNVIN NNREKYVIKY INPADINGLA	660
KNRYLVKFVL SKIPEEQIDS YYKLFSEEEE PGCEEKIKLL TKKISKLNQ TLFENNKIPN	720
VEKEKKKAI TLYFTIVYIL VKNLVNINGL YTLALYFVER DGYFYKDICG KKDKKKSYND	780
VDYLLLP EIF SGSKYREETK NLKLPKEKDR DIMKKYLPND KDREKYNKFF TAYRNNIVHL	840
NIIAKLSELT KNIDKDINSY FDIYHYCTQR VMFNCKEKN DVVLAKMKDL AHKSDCNEF	900
SSKHTYPFSS AVLRFMNLPF AYNVPRFKNL SYKKFFDKQ	939

(SEQ ID NO: 94)。

[0238] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群tpg|DJXD01000002.1| (未培育的瘤胃球菌属(Ruminococcus)联合体,UBA7013,来自绵羊肠道宏基因组)：

MKKQKSKKTV SKTSGLKEAL SVQGTVIMTS FGKGNMANLS YKIPSSQKPQ NLNSSAGLKN	60
VEVSGKKIKF QGRHPKIATT DNPLFKPQPG MDLLCLKDKL EMHYFGKTFD DNIHIQLIYQ	120
ILDIEKILAV HVNNIVFTLD NVLHPQKEEL TEDFIGAGGW RINLDYQTLR GQTNKYDRFK	180
NYIKRKELLY FGEAFYHENE RRYEEDIFAI LTLLSALRQF CFHSDLSSDE SDHVNSFWLY	240
QLEDQLSDEF KETLSILWEE VTERIDSEFL KTNTVNLHIL CHVFPKESKE TIVRAYYEFL	300

IKKSFKNMGF SIKKLREIML EQSDLKSFKE DKYNSVRAKL YKLFDIFIITY YYDHHAFFEKE	360
ALVSSLRSSL TEENKEEIIYI KTARTLASAL GADFKKAAAD VNAKNIRDYQ KKANDYRISF	420
EDIKIGNTGI GYFSELIYML TLLLDGKEIN DLLTTLINKF DNIISFIDIL KKLNLFEKFK	480
PEYADFFNMT NCRYTLEELR VINSIARMQK PSADARKIMY RDALRILGMD NRPDEEIDRE	540
LERTMPVGAD GKFIKKGQGF RNFIASNVIE SSRFHLYLVRY NNPHKTRTLV KNPNVVKFVL	600
EGIPETQIKR YFDVCKGQEI PPTSDKSAQI DVLARIISSV DYKIFEDVPQ SAKINKDDPS	660
RNFSDALKKQ RYQAIVSLYL TVMYLITKNL VYVNSRYVIA FHCLERDAFL HGVTLPKMKNK	720
KIVYSQLTTH LLTDKNYTTY GHLKNQKGHR KQYVLVKNL QNSDITAVSS FRNIVAHISV	780
VRNSNEYISG IGELHSYFEL YHYLVQSMIA KNNWYDTS HQ PKTAEYLNLL KKHHTYCKDF	840
VKAYCIPFGY VVPYKNTLI NELFDRNNPN PEPKEEV	877

(SEQ ID NO: 95)。

[0239] CasRX/Cas13d宏基因组命中(无蛋白质登录号):重叠群tpg|D JXD01000002.1|(未培育的瘤胃球菌属联合体,UBA7013,来自绵羊肠道宏基因组)(SEQ ID NO:95)的示例性同向重复序列包含以下核酸序列或由其组成:

CasRX/Cas13d DR:

caactacaac cccgtaaaaa tacgggggttc tgaaac 36

[0240] (SEQ ID NO:96)。

[0241] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成:

CasRX/Cas13d宏基因组命中(无蛋白质登录号):重叠群0GZC01000639.1(人肠道宏基因组联合体):

MKKKNIRATR EALKAQKIKK SQENEALKKQ KLAEEAAQKR REELEKKNLA QWEETSAEGR	60
RSRVKAVGVK SVFVVGDDLY LATFGNGNET VLEKKITPDG KITTFPEEET FTAKLKFAQT	120
EPTVATSIGI SNGRIVLPEI SVDNPLHTM QKNTIKRSAG EDILQLKDVL ENRYFDRSFN	180
DDLHIRLIYN ILDIEKILAE YTTNAVFAID NVSGCSDDFL SNFSTRNQWD EFQNPEQHRE	240
HFGNKDNVIC SVKKQQDLFF NFFKNNRIGY FGKAFFHAES ERKIVKKTEK EVYHILTIG	300
SLRQWITHST EGGISRLWLY QLEDALSREY QETMNNCYNs TIYGLQKDFE KTNAPNLNFL	360
AEILGKNASE LAEPYFRFII TKEYKNLGFS IKTLREMLLD QPDLQEIREN HNVYDSIRSK	420
LYKMIDFVLV YAYSNERKSK ADALASNLRS AITEDAKKRI YQNEADQLWT SYQELFKRIR	480
GFKGAQVKEY SSKNMPPIQ KQIQNILKPA EQVTYFTKLM YLLTMFLDGK EINDLLTTLI	540
NKFDNISSLL KTMEQLELQT TFKEDYTFFQ QSSRLCKEIT QLKSFARMGN PISNLKEVMM	600

```

VDAIQILGTE KSEQELQSMA CFFFRDKNGK KLNTGEHGMR NFIGNNVISN TRFQYLIRYG      660
NPQKLHTLSQ NETVVRFVLS RIAKNQRVQG MNGKNQIDRY YETCGGTNSW SVSEEEKINF      720
LCKILTNSMY DQFQDVQSG AEITAEKRRK KERYKAIISL YLTVLYQLIK NLVNINARYI      780
IAFHCLERDA ILYSSKFNTS INLKKRYTAL TEMILGYETD EKARRKDTRT VYEKAEAAKN      840
RHLKNVKWNC KTRENLENAD KNAIVAFRNI VAHLWIIRDA DRFITGMGAM KRYFDCYHYL      900
LQRELGYILE KSNQGSEYTK KSLEKVQQYH SYCKDFLHML CLPFAYCIPR YKNLSIAELF      960
DRHEPEAEPK EEASSVNNSQ FITT                                             984

```

(SEQ ID NO: 97)。

[0242] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OHBM01000764.1 (人肠道宏基因组联合体)：

```

XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX      60
XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX      120
XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX      180
XXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXHPLQKRYR YLTSTNLKSF      240
ETYKNNLVNK KKFDLDRVKK IPQLAYFGSA FYNTPEDTSA KITKTKIKSN EEIYYTFMLL      300
STARNFSAHY LDRNRAKSSD AEDFDGTSVI MYNLDNEELY KKLYNKKVHM ALTGMKKVLD      360
ANFNKKVEHL NNSFIKNSAK DFVILCEVLG IKSREKTKF VKDYYDFVVR KNYKHLGFSV      420
KELRELLFAN HDSNKYIKEF DKISNKKFDS VRSRLNRLAD YIIYDYNNKN NAKVSDLVKY      480
LRAAADDEQK KKIYLNESIN LVKSGILERI KKILPKLNGK IIGNMQPDST ITASMLHNTG      540
KDWHPISENA HYFTKWIYTL TLFMDGKEIN DLVTTLINKF DNIAFIEVL KSQSVCTHFS      600
EERKMFIISA EICSELSAMN SFARMEAPGA SSKRAMFVEA ARILGDNRSK EELEEFDTL      660
FDKSASKKEK GFRNFIRNNV VDSNRFKYLT RYTDTSVKA FSNKALVKF AIKDIPQEQI      720
LRYNSCFGAS SERYNDGMS DKLVEAIGKI NLMQFNGVIQ QADRNMLPEE KKKANAQKEK      780
YKSIIRLYLT VCYLFFKNLV YVNSRYSAF YNLEKDRSLF EINGELKPTG KFDEGHYTGL      840
VKLFIDNGWI NPRASAYLTV NLANSDETAI RTRNTAEHL EALRNADKYL NDLKQFDSYF      900
EIIHYITQRN IKEKCEMLKE QTVKYNNDDL KYHGYSKDFV KALCVFPGYN LPRFKNLSID      960
ALFDKNDKRE KLKKGFE

```

(SEQ ID NO: 98)。

[0243] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OHCP01000044.1 (人肠道宏基因组联合体)：

MAKKITAKQK REEKERLNKQ KWAKNDSVII VPETKEEIKT GEIQDNNRKR SRQKSQAKAM	60
GLKAVLSFDN KIAIASFVSS KNAKSSHIER ITDKEGTTIS VNSKMFESSV NKRDIKIEKR	120
ITIEEPQQDG TIKKEEGVK STTCNPYFKV GGKDYIGIKE IAEEHFFGRA FPNENLRVQI	180
AYNIFDVQKI LGTFVNNIIY SFYNLSRDEV QSDNDVIGML YSISDYDRQK ETETFLQAKS	240
LLKQTEAYYA YFDDVFKKNK KPDKNKEGDN SKQYQENLRH NFNILRVLSF LRQICMHAEV	300
HVSDDEGCTR TQNYTDSLEA LFNISKAFGK KMPCLKTLID NIYSKGINAI NDEFVKNKGN	360
NLYILSKVYP NEKREVLRE YYNFVCKEG SNIGISTRKL KETMIAQNMP SLKEENTYRN	420
KLYTVMNFIL VRELKNCATI REQMIKELRA NMDEEEGRDR IYSKYAKEIY LYVKDKLKL	480
LNVFKEEAEG IIIPGKEDPV KFSHGKLDKK EIESFCLTTK NTEDITKVIY FLCKFLDGKE	540
INELCCAMMN KLDGISDLIE TAKQCGEDVE FVDQFKCLSK CATMSNQIRI VKNISRMKKE	600
MTIDNDTIFL DALELLGRKI EKYQKDKNGD YVKDEKGKKV YTKDYNFQD MFFEGKNHRV	660
RNFVSNNVIK SKWFSYVVRV NKPAECQALM RNSKLVKFAL DELPDSQIEK YYISVFGEKS	720
SSSNEEMRRE LLKKLCDFSV RGFLDEIVLL SEDEMKQKDK FSEKEKKKSL IRLYLTVIYL	780
ITKSMVKINT RFSIACATYE RDYILLCQSE KAERAWEKGA TAFALTRKFL NHDKPTFEQY	840
YTREREISAM PQEKRKELRK ENDQLKKTH YSKHAYCYIV DNVNLTGAV ANDNGRGLPC	900
LSEKNDNANL FLEMRNKIVH LNVVHDMVKY INEIKNITSY YAFFCYVLQR MIIGNNSNEQ	960
NKFKAKYSKT LQEFGTYSKD LMWVLNLPFA YNLPRYKNLS NEQLFYDEEE RMEKIVGRKN	1020
DSR	1023

(SEQ ID NO: 99)。

[0244] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OGDF01008514.1| (人肠道宏基因组联合体)：

MTETKPKRED IAKTPAAKSR SKAAGLKSTF AVNGSVLLTS FGRGNDVPE KLITEKAVSE	60
INTVKPRFSV EKPATSYSSS FGIKSHISAT ADNPLAGRAP VGEDAIHAK VLEQRVFGKT	120
FSDDNIHIQL IYNILDIRKI LSTYANNVVF TINSMRRLDE YDREQDYLG YLTGNSYERL	180
LDIADKYAVD GEDWRNTAAG ISNDFEKKQF QTINGFWDLL DMIEPYMCYF SEAFFCETTV	240
KDPDSGRIVP CLEQRSDGDI YNILRILSIV RQTCMHDNAS MRTVMFTLGQ NSVRDRKNGF	300
DELAELLDYL YDEKIDIVNR DFLRNQKNNI ELLSRIYGSS ADSPERDRLV QNFYDFRVLS	360
QDKNLGFSIK KLREKLLDSP ALSVVRSKKY DTMRSKIYSL IDFMIYRKFS ENHVAVDDEV	420
EELRSLLTED EKESAYSRTWA ETLINDGFAQ EILVKLLPQT DPAVIGKIKG KKLLNDSIAG	480

IKLKKDASFF TKIINVLCMF QDGKEINELV SSLVNKFANI QSFVDVMRSQ GIDSGFTADY	540
AMFAESGRIS RELHILKGIA RMQHSIAGLG DVKIYGSDDK FHGVSRRVYT DAAYILGFGE	600
RSEDNDGYVD DYVSSKLLGG ADKNLRNFIT NNVIKNRRFL YTVRYMNPKR AKKLVQNDAL	660
VVLALSGIPE TQIDRYKSC IEKRSFNPDL NEKIAALSEM ITTLKIDDFE DVKQNPEKNA	720
NYEAKKNQRI SKERYKACIG LYLTVLYLIC KNLVKINARY SIAIGCLERD TQLHGVDFKG	780
AAYMTRDVFI AKGWINPKKP TVKSIKEQYA FLTPYIFTTY RNMIHLAAV TNAYKYIPQM	840
DRFKSWFHLI HTVIQHSLIQ QYEYDRDYGR KGAPVVSERV LQLEQCREH SNYSRDLHI	900
LNLPFGYNLP RYLNLSSEKY FDANAI	926

[0245] (SEQ ID NO:100)。

[0246] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OGPN01002610.1(人肠道宏基因组联合体)：

MAKITAKQK REEKERLNKQ KWAKQDTPVV PKSKTEEKPV AASDDKLLKT TQVKKVQTKS	60
KAKAMGLKTV LSFDDKIAIA SFVNDKKTCL PHIERITDKS GTTIHENARM FDSSVDEQNV	120
NIEKRMTIEE KQNDGTFKKD EKDVKATICN PYFKTCGKDY IGIKDVAEKY FFGKTFPNEN	180
LRVQIAYNVF DIQKILGTYV NNIIYSFYNL RRDGKSDVDI IGSLYAFADF DNQLKDKPAF	240
REAKDLLKNT EAYFSYFGDV FKKSCKGKKD ENNEDYEKNL RHNFNVLRLV SFLRQICHTA	300
YVKCTGGAKN NGDSTKVEAE SLDALFNITE YFAKTAPELS KTINEIYKEG IDRINNDFTV	360
NGKNNLYILS KVYPDMQRNE LVKKYYQFVV CKEGNNVGIN TRKLKESIIS QHPWITTPQD	420
NNKANDYESC RHKLYTIMCF ILVAELDAHE SIRDNMVAEL RANMDGDDGR DAIYEKYAKD	480
IYHIVKDKLL AMQKVFDEEL VPKVEGKND PQQFTHGKLG KKEIESFCLS DKNTSDIAKV	540
VYFLCNFLDG KEINELCCAM MNKFDGIGDL IDTAKQCGEE VKFIEEFACL SNCRKITNDI	600
RVAKSISKMK NKVNIDNDII YLDAIELLGR KIEKYQKDEN GKILLGTDGK RLYTQEYKYF	660
NDMFFNAGNH KVRNFIANNV MQSKWFFYV RYNKPAECQI IMRNKTLVKF TLDDLPLDQI	720
QRYYSVFGD NNMPAVDEMR KRLDKINQF SVRGFLDELD EIVLMSDEES KRNSSEKEQ	780
KKSLIRLYLT IAYLITKSMV KINTRFSIAC AMYERDYALL CQSEMKGGPW DGGAQALAVT	840
RKFLNHDREV FDRYCAREAE IARLPSEERK PLRKANDKLL KQTHYTNHSY TYIVNNLNSF	900
TDIDYCAKDV GLPAPNDKND NASILGEMRN DIAHLNIVHD MVKYIEELKD ISSYYAFYCY	960
VLQRRLVGKD PNCQNKFKAK YAKELNDYGT YNKNLMWMLN LPFAYNLPRY KNLSSEFLFY	1020
DMEYNKKDDE	1030

(SEQ ID NO: 101)。

[0247] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：来自重叠群emb|OBLI01020244和emb|OBLI01038679(来自猪肠道宏基因组)：

MAKKITAKQR REERERQNKQ KWAKKQADAT AVFECEADIK PADSKDEDCT NIYIKREKKK	60
TQAKAMGLKT VLGFDNKIAI ASFMSSKDSK SSHIERITDP NGKTIREDRV MFDSNVDECS	120
INLEKRMTVE ERQKDGTIKK DEKDVKSTIC NPYSNECGKD YIGIKSVAEE LFFGRTPND	180
NLRVQIAYNI FDIQKILGTY INNIIYSFY NLSRDESQSDN DVIGTLYMLK DFDGQKETDT	240
FRQARALLER TEAYYSYFDN VFKKIDKNKK KSDDCKRERN EILRYNFENVL RVL SFLRQIC	300
AHAQVKISNE HDREKGGGLV DSLDALFNIS RFFDAVAPEL NEVINSVYSK GIDDINDNFV	360
KNGKNNFYIL SKIYPEVARE DLLREYYYFV VSKEGNNIGI STKKLKEAII VQDMSYKSE	420
DYD TYRNKLY TVLCFILVKE LNERTTIREQ MVADLRANMN GDIGREDIYS KYAKIIYAQV	480
KPRFDTMKSA FEEEEAKDVIV PDKKKPVKFS HGKLDKNEIE RFCITSANTD SVAKIIYFLC	540
KFLDGKEINE LCCAMMNKLD GINDLIETAE QCGAKVEFVD KFSVLSNCET ISDQIRIVKS	600
ISKMKKEIAI DNDTIFLDAL ELLGRKIDKY KKDATGKYLK DENGKYLYSK EYDDFQYMFF	660
KDSHRVRNFI SNSVIKSKWF SYIVRYNQPS ECRAIMKNKT LVKFALDELP DLQIQRYFVA	720
LYGDEDLPSY GEMRKILLKK LHDFSIGFL DEIVLLSDLD MESQDKYCEK EQKKSFLRLY	780
LTIAYLITKS MVKINTRFSI ACATYERDYA LLCASNKQER AWSSGATALA LTRRFLNQDK	840
LIFEKHYARE GEISKLPKEE RKAMRKVNDQ LLKRTHFSKH SYCYIVDNVN RLTGGE CRTD	900
KRVLPVLNEK NDNAGILLDF RKTIAHLNVV HKMVDYVDEI KGITSYYAFF CYVLQRM LVG	960
NNLNEKNAIK EKYSATVKSF GTYSKDFMWL INLPFAYNLP RYKNLSNEQL FYDEEERNET	1020
EEQIDRL	1027

(SEQ ID NO: 102)。

[0248] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：
CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群OIZX01000427.1：

MAKKKK TARQ LREEMQQQRK QAIQKQQEQR QEKAAAARET AAPEQPAAAP VPKRQRKSLA	60
KAAGLKS NFI LDPQRRTTVM TAFGQGSTAI LEKQIVDRAI SDLQPVQQFQ VEPASAAKYR	120
LKNSRVRFPN VTADDPLYRR KDGGFVPGMD ALRRKNVLEQ RFFGKSFADN IHIQMIYSIL	180
DIHKILAAAS GHIVHLLNIV NGSKDRDFIG M LAAHVLYNE LNEEAKRSIA DFCKSPRLIY	240
YSAAFYETLD NGKSERRSNE DIFNILALMT CLRNFS SHHS IAIKVKDYSA AGLYNLRR LG	300
PDMKKMLDTF YTEAFIQLNQ SFQDHNTTNL TCLFDILNIS DSARQKQLAE EFYRYVVFKE	360

QKNLGFSVRK LREEMLLLPD AAVIADKRYD TCRSKLYNLM DFLILRVYRT GRADRCDKLP	420
EALRAALTDE EKAVVYHKEA LSLWNEMRTL ILDGLLPQMT PENLSRLSGQ KRKGELSDD	480
AMLKECLYEP GPVPEDAAPE EANAIFYCRM IYLATLFMDG KEINTLLTTL ISKFENIAAF	540
LQTMEQLNIE AELGPEYAMF TRSRVAEQL RVINSFALMK KPQVNAKQQL YRAAVTLLGT	600
EDPDGVTDEM LCIDPVTGKM LPPNQRHHGD TGLRNFIANN VVESRRFQYL IRYSDPAQLH	660
QLASNKKLVR FVLSSIPDTQ INRYETCGQ TRLAGRAAKV EFLTDMIAAI RFDQFRDYNQ	720
KERGANTQKE RYKAMLGlyQ TVLYLAVKNL VNINARVMA FHCVERDMFL YDGELTDPKG	780
ESVSAFLAVN GKKGVQPQYL LLTQLFIRRD YLKRSACEQI QHNMENISDR LLREYRNAVA	840
HLNVIAHLAD YSADMREITS YYGLYHYLMQ RHLFKRHAWQ IRQPERPTEE EQKLIEQEQK	900
QLAWEKALFD KTLQYHSYNK DLVKALNAPF GYNLARYKNL SIEPLFSKEA APAAEIKATH	960
A	961

(SEQ ID NO: 103)。

[0249] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群OCTW011587266.1：

MKQNDRENNN KIKKSAKAV GVKSLARLSD GSTVVSFGK GAAAELES LI TGGEIRKLS D	60
KAILEITDDT QKNAYNVKS SRIPNLTART DKLSKSGMD DLGFKRELEL EVFGQCFDD S	120
IHIQIAHAVF DIQKSLAAVI PNVLYTLNNL DRSYSTDN TS DKKDIIGN TL NYQHSYESFN	180
VEKRGFEFTEY YNAAKDRFSY FPDILCVLEK VNGKDRYQPK SEKDAFNVLS SVNMLRNSLF	240
HFAPKSN DGK ARIAVFKNQF DSDFSHITST VNKIYSAKIA GVNENFLNNE GNNLYIILKA	300
TNWDIKKIVP QLYRFSVLKS DKNMGFNMRK LREFAVESKN IDLSRLNDKF LTNNRKKLYK	360
VIDFIIYYHL NKVLKDSFVD DFVAALRASQ SEEEKEKLYA QYSERLFADE GLKSAIKKAV	420
DMISDTKSNI FKMKTPLDKA LIENIKVNSD ASDFCKLIYV FTRFLDGKEI NILLNSLIKK	480
FQDIHSFN TT VKKLSEN LI INADYVDDYS LFEQSGTVAR ELMLIKSISK MDFGLDNINL	540
SFMYDDALRT LGVSDENLPE VKREYFGGTK NLSAYIRNNV LENRRFKYVI KYIHPSDVQK	600
IACNKAIAGF VLNRMPTQI KRYYDSL INK GATDIQAQAK ALLDCITGIS FDAIKDDKHL	660
HKSKEKSPQR SADREKKAM LTLYYTIVYI FVKQMLHINS LYTIGFFYLE RDQRFIYSRA	720
KKENKNPSKN SYLNDFRSVT AYFIPSEIMK RIEKNENKGF LEDFEALWNS CGKTSRLRKE	780
DVLLYARYIS PDHALKNYKM ILNSYRNKIA HINVIMSAGK YTGGIKRMDS YFSVFQHLVQ	840
CDILSNPN NK GKCFESES LK PLLLD MKFDG TDEKLYSKRL TRALNIPFGY NVPRYKNLTF	900
EKIYLKSSIN E	911

(SEQ ID NO: 104)。

[0250] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OGNF01009141.1：

MADIDKKKSS AKAAGLKSTF VLENNKLLMT SFGNGNKAVI EKIIDEKVDS INEPEVFSVT	60
PCDKKFELQP AKRGLAADSL VDNPLKSKKT AGDDAIHSRK FLERQFFDGN TFNDNIHIQL	120
IYNILDIEKI LSVHVNDIVY SVNNILSRGE GMEYNDYIGT LNLKSFETYK NNLVNKKKFD	180
LDRVKKIPQL AYFGSAFYNT PEDTSAKITK TKIKSNEEIY YTFMLLSTAR NFSAHYLDRN	240
RAKSSDAEDF DGTSVIMYNL DNEELYKKLY NKKVHMALTG MKKVLDANFN KKVEHLNNSF	300
IKNSAKDFVI LCEVLGIKSR DEKTKFVKDY YDFVVRKNYK HLGFSVKELR ELLFANHDSN	360
KYIKEFDKIS NKKFDSVRSR LNRLADYIIY DYYNKNNAKV SDLVKYLRAA ADDEQKKKIY	420
LNESINLVKS GILERIKKIL PKLNGKIIGN MQPDSTITAS MLHNTGKDWL PISENAHYFT	480
KWIYTLTLFM DGKEINDLVT TLINKFDNIA SFIEVLKSQS VCTHFSEERK MFIDSAEICS	540
ELSAMNSFAR MEAPGASSKR AMFVEAARIL GDNRSKEELE EYFDTLFDKS ASKKEKGFRN	600
FIRNNVVDN RFKYLTRYTD TSSVKAFSNN KALVKFAIKD IPQEQILRYN NSCFGASERY	660
YNDGMSDKLV EAIGKINLMQ FNGVIQQADR NMLPEEKKA NAQKEKYKSI IRLYLTVCYL	720
FFKNLVYVNS RYSAFYNLE KDRSLFEING ELKPTGKFDE GHYTGLVKLF IDNGWINPRA	780
SAYLTVNLAN SDETAIRTFR NTAEHLEALR NADKYLNDLK QFDSYFEIYH YITQRNIKEK	840
CEMLKEQTVK YNNDLLKYHG YSKDFVKALC VPFGYNLPRF KNLSIDALFD KNDKREKLKK	900
GFED	904

(SEQ ID NO: 105)。

[0251] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d宏基因组命中(无蛋白质登录号)：重叠群emb|OIEN01002196.1：

MERQKRKMKS KSKMAGVKS FVIGDELLMT SFGDGDDAVL EKIDIDENGVV NDCRNPAAYD	60
AVYGTDSIRV KKTNNNIRAK VNNPLAKSNI RSEESALFRT RVNEYKREQK DKYETLFFGK	120
TFDDNIHIQL ISKILDIEKT FSVVIGNIVY AINNLSLEQS IDRPIDIFGD KNTQGSLRE	180
DNDYLKTMPL RCEYLFHNIL NSDSDNNSKM NYNKVNKGKE EKDNRRNNENI EKLKKALEVI	240
KIIRVDSFHG VDGIGKDQKF PRSKYNLAVN YNEEIQKTIS EPFNRKVEEV QQDFYRNSCV	300
NIDFLKEIMY GSNYTDGSD SLECSYFNFA ILKQNKMGF SITSIRECLL DLYELNFESM	360
QNLPRANSF CDFLIYDYC KNESERANLV DCLRSAASEE EKKNIFYQTA ERVKEKFRNA	420

FNRISRFDAS YIKNSREKNL SGGSSSLPKYS FIEGFTKRSK KINDNDEKNA DLFCNMLYYL	480
AQFLDGKEIN IFLTSIHNI QNIDSFLKVM KEGMECKFQ KDFKMFHAG HVAKKIEIVI	540
SLAKMKKTLD FYNAQALKDA VTILGVSKKH QYLDMNSYLD FYMFDNRSGA TGKNAGKDH	600
LRNFLVSNVI RSRKFNYLSR YSNLAEVKKL AQNPSLVQFV LSRIEPLIC RYYESSQGIS	660
SEGITIDEQI KKLTGIIIVDM NIDSFENINN GEIGMRYSKA TPQSIERRNQ MRVCVGLYLN	720
VLYQIEKNLM NVNARYVLAF AFAERDALML NFTLEECKKN KKRSSGGFSF IEMTQFFIDK	780
KLFKVATEAI KKNVLKYNGN PESLNHIPGE YICKNMEGYH ENTVRNFRNM VAHLTAVARV	840
PLYISEVTQI DSYVALYHYC MQMNILQGIE QSGKILDNIK LKNALENARV HRTYSKDAVK	900
YLCLPFAYNI SRYKALTIKD LFDWTEYSCK KDE	933

(SEQ ID NO: 106)。

[0252] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：
CasRX/Cas13d宏基因组命中(无蛋白质登录号):重叠群e-k87_11092736:

MKRQKTFAKR IGIKSTVAYG QGKYAITTFG KGSKAETIAVR SADPPEETLP TESDATLSIH	60
AKFAKAGRDG REFKCGDVDE TRIHTSRSEY ESLISNPAES PREDYLGGLG TLERKFFGDE	120
YPKDNLRIQI IYSILDIQKI LGLYVEDILH FVDGLQDEPE DLVGLGLGDE KMQKLLSKAL	180
PYMGFFGSTD VFKVTKKREE RAAADEHNAK VFALGAIQ KLAHFKWKES LAIFGANANM	240
PIRFFQGATG GRQLWNDVIA PLWKKRIERV RKSFLSNSAK NLWVLYQVFK DDTDEKKKAR	300
ARQYYHFSVL KEGKNLGFNL TKTREYFLDK FFPIFHSSAP DVKRKVDTFR SKFYAILDFI	360
IYEASVSVAN SGQMKGKVPW KGAI DNALVK LREAPDEEAK EKIYNVLAAS IRNDSLFLRL	420
KSACDKFGAE QNRPVFPNEL RNNRDIRNVR SEWLEATQDV DAAAFVQLIA FLCNFLEGKE	480
INELVTALIK KFEIGIQAID LLRNLEGVDS IRFENEFALE NDDKGNMAGR IARQLRLLAS	540
VGKMKPDMTD AKRVLYKSAL EILGAPPDEV SDEWLAENIL LDKSNNDYQK AKKTVNPFRN	600
YIAKNVITSR SFYYLVRYAK PTAVRKLSMN PKIVRYVLKR LPEKQVASY SAIWTQSES	660
SNEMVKLIEM IDRLTTEIAG FSFAVLKDKK DSIVSASRES RAVNLEVERL KKLTTLMSI	720
AYIAVKSLVK VNARYFIAYS ALERDLYFFN EKYGEEFRLH FIPYELNGKT CQFEYLAILK	780
YYLARDEETL KRKCEICEEI KVGCEKHKK ANPPYEYDQE WIDKKKALNS ERKACERRLH	840
FSTHWAQYAT KRDNMAKHP QKWYDILASH YDELLALQAT GWLATQARND AEHLNPVNEF	900
DVYIEDLRRY PEGTPKNKDY HIGSYFEIYH YIRQRAYLEE VLAKRKEYRD SGSFTDEQLD	960
KLQKILDDIR ARGSDKNLL KLEYLPFAYN LPRYKNLTTE ALFDDDSVSG KKRVAEWRE	1020
EKTREAEREQ RRQR	1034

(SEQ ID NO: 107)。

[0253] CasRX/Cas13d宏基因组命中(无蛋白质登录号):重叠群e-k87_11092736 (SEQ ID NO:107)的示例性同向重复序列包含以下核酸序列或由其组成:

CasRX/Cas13d同向重复1:gtgagaagtc tccttatggg gagatgctac 30
(SEQ ID NO:108)。

[0254] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d Ga0129306_1000735：

MQKQREQQTV TDESERKKKP LKSGAKAAGL KSVFVLSEGG ELLTSFGRGN EAVPEKRVTG	60
GTIANARTDN KEAFSAALQN KRFEVFGRTA GSSDDPLAVS RAPGQDLIGA KTALEERYFG	120
RAFADNIHQ VIYAIQDINK ILAVHANNIV YTLNNLDREA DPETDDFIGS GYLTLKNTFE	180
TYCDPAALNE REREKVTVSK QHFDAFMQNP RLAYYGNAFF RKLSKAERLA RGREIFDKES	240
PERRQEILGS RGKNKSVDDDE IRALAPEWVK REERDVYSEL VLMSELRQSC FHGQQKNSAR	300
IFRLDNDLGP GVDGARELLD RLYAEKINDL RSFDKTSASS NFRLLFNAYH ADNEKKKELA	360
QEFYRFSVLK VSKNTGFSIR TLREKIIEDH AAQYRDKIYD SMRKKLFSTF DFFLWRFYEE	420
REDEAEELRA CLRAARSDEE KEQIYAEAAA SCWPSVKPFV ESVAATLCDV VKGRTKLNKL	480
KLSADESTLV RNAIDGVRIS PRASYFTKLI YLMTLFLDGK EINDLLTTLI HAFENIDSFL	540
SVLGSERLER TFDANYRIFA DSGVIAQELR AVNSFARMTT EPFNSKLVMF EDAAQLFGMS	600
GGLVEHAEEL REYLDNKMLD KTKLRLLPDG KVDTGFRNFI ISNVTESRRF RYLVRycePR	660
AVRDYMSCRP LIRLTLRDMP DTILRRYYEQ SVGAATVDRE RILDTLADKL LSLRFTDFEN	720
VNQORANAERN REKQKMMGII SLYLNVAYQI VKNLVYVNAR YTMAYHCAER DTELLLNAAAG	780
EGNLLRRDRS WPARLHLPRR ALARRRDRVE VMERDVARGP EAYNRDEWLK LVRTLRRREKR	840
VCDNLHNNYA YLCGADAEPG DASLSLLFVY RNKAAHLSVL NKGGRLSGDL KEAKSWFYVY	900
HFLMQRVLEE EFRNTQALPE RLRELLMAE RYRGCSKDLI KVLNLTFAYN LPRYKNLSID	960
GRFDKNHPDP SDE	973

(SEQ ID NO: 109)。

[0255] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d Ga0129317_1008067：

MKKQKKSLVK AAGLKSAFVV GDSVYLTSGF KGNAARLDTK INPDNSTERY VSDSEKHTLK	60
INSITDTELR LSGPFPKQAE AKNPETHKKN EQKNTRQDML GLKSTLEKfy FGSTFDDNIH	120

IQIIHNIQDI AKILAAHSNN AGYALDNMLA YQGVFSDMI GYMGTSRTFD NYDPNHKNNK	180
DDFRFLKLPR LGYFGSAFYS QKGKDFEKRS DEEVYNICAL MGQIRQCCFH GKQEKYQLKW	240
LYNFHNFKSN KPFLDTLDKH FDEMIDRINK NFIKNNTPD L IILSGLYPDM AKKELVRLFY	300
DFTTVKEYKN MGFSVKKLRE KMLESEEASD FRDKDYDSVR RKLYKLMDFC IYYLYYSDSE	360
RNENLVSRRLR ESLTDENKDI IYSKEAKIVW NELRKKFSTI LDNVKGSNIK KLENVKEKFI	420
SEDEFDDIKL DIDISYFSKL MYVMCYFLDG KEINDLLTTL VSKFDNIGSI IEAATQIGIN	480
IEFIDDFKFF DRSKDISVEL NIIRNFARMQ APVPNAKRAM QEDAIRILGG SEEDIFSILD	540
DMTGYDKSGK KLAQSKKGFR NFIINNIVES SRFKYIVRYS NPQKIRKLAN NSVVVG FVLG	600
KLPDAQIESY FNSCLPNRVY STPDKARESL RDMLHNISFN DFADVKQDDR RATPEEKVEK	660
ERYKAIIGLY LTVMYHLVKN LVYVNSRYVM AFHCLERDAM HYDVSLDNYR DLIRHLISEG	720
DSSCNHFISH NRRMRDCIEE NVKNSEQLIF GKEDAVIRFR NNVAHLSAIR NANEYIGDIR	780
EITSYFALYH YLMQRKLIDD CKVNDTAHKY FEQLTKYKTY VMDMVKALCS PFGYNLPREFK	840
NLSIEGKFDM HESK	854

(SEQ ID NO: 110)。

[0256] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d Ga0224415_10048792：

MSKKENRKS Y VKGLGLKSTL VSDSKVYLT FADGSNAKLE KCVENNKIIC ISNDKEAFAA	60
SIANKNVGYK IKNDEKFRHP KGYDIISNNP LLHNSVQQD MLGLKNVLEK RYFGKSSGGD	120
NNLCIQIIHN IIDIEKILSE YIPNVVYAFN NIAGFKDEHN NIIDIIGTQT YNSSYTYADF	180
SKDKSDKKYI EFQKLLKNKR LGYWGKAFFT GQGNNAKVRQ ENQCFHIIAL LISLRNWATH	240
SNELDKHTKR TWLYKLDDTN ILNAEYVKTL NYLYDTIADE LTKSFSKNGA VNVNYLAKKY	300
NIKDDLPGFS EQYFRFSIMK EQKNLGFNIS KLRENMLDFK DMSVIRDDHN RYDKDRSKIY	360
TMMDFVIYRY YIDNNNSID FINKLRSSID EKSKEKLYNE EANRLWNK LK EYMLYIKEFN	420
GKLASRTPDR DGNISEFVES LPKIHRL LPR GQKISNFSKL MYLLTMFLDG KEINDLLTTL	480
INKFENIQGF LDIMPEINVN AKFEPEYVFF NKSHEIAGEL KLIKGFAQMG EPAATLKLEM	540
TADAIKILGT EKEDAELIKL AESLFKDENG KLLGNKQHGM RNFIGNNVIK SKRFHYLIRY	600
GDPALHLKIA TNKNVVR FVL GRIADMQKKQ GQKGKNQIDR YEV CVGNKD IKKTIEEKID	660
ALTDIIVNMN YDQFEKKKAV IENQNRGKTF EEKNKYKRDN AEREKFKKII SLYLTVIYHI	720
LKNIVNVNSR YILGFHCLER DKQLYIEKYN KDKLDGFVAL TKFCLGDEER FEDLKAKAQA	780
SIQALETANP KLYAKYMNYS DEEKKEEFKK QLNRRERVKNA RNAYLKNIKN YIMIRLQLRD	840
QTDSSGYLCG EFRDKVAHLE VARHAHEYIG NIKEVNSYFQ LYHYIMQCRL YDVLKNNTKA	900

EAMVKGKAKE YFEALEKEGT YNDKLLKIAC VPFGYCIPRY KNLSMEELFD MNEEKKFKKK 960
 APENT 965

(SEQ ID NO: 111)。

[0257] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d 160582958_基因49834：

MKNSVTFKLI QAQENKEAAR KKAKDIAEQA RIAKRNGVVK KEENRINRIQ IEIQTQKKS 60
 TQNAYHLKSL AKAAGVKSUF AIGNDLLMTG FGPNDATIE KRVFQNR AIE TLSSPEQYSA 120
 EFQNKQFKIK GNIKVLNHST QKMEEIQTEL QDYNRPHFD LLGCKNVLEQ KYFGRTFSDN 180
 IHVQIAYNIM DIEKLLTPYI NNIIYTLNEL MRDNSKDDFF GCDSHFSVAY LYDELKAGYS 240
 DRLKTKPNLS KNIDRIWNNF CNYMNSDSGN TEARLAYFGE LFYKPKETGD AKSDYKTHLS 300
 NNQKEEWELK SDKEVYNIFA ILCDLRHFCF HGESITPSGK PFPYNLEKNL FPEAKQVLNS 360
 LFEEKAESLG AEAFGKTAGK TDVSILLKVF EKEQASQKEQ QALLKEYYDF KVQKTYKNMG 420
 FSIKKLREAI MEIPDAAKFK DDLYSSLRHK LYGLDFDILV KHFLDTS DSE NLQNN DIFRQ 480
 LRACRCEEEK DQVYRSIAVK VWEKVKKKEL NMFKQVVVIP SLSKDELKQM EMTKNTELLS 540
 SIETISTQAS LFSEMIFMMT YLLDGKEINL LCTSLIEKFE NIASFNEVLK SPQIGYETKY 600
 TEGYAFFKNA DKTAKELRQV NNMARMTKPL GGVNTKCVMY NEAAKILGAK PMSKAELESV 660
 FNLDNHDTY SPSGKKIPNK NFRNFIINNV ITSRRFLYLI RYGNPEKIRK IAINPSIISF 720
 VLKQIPDEQI KRYYPFCIGK RTDDVTLMRD ELGKMLQSVN FEQFSRVNNK QNAKQNPNGE 780
 KARLQACVRL YLTVPYLFIK NMVNINARV LAFHCLERDH ALCFNSRKLN DDSYNEMANK 840
 FQMRKAKKE QYEKEYKCKK QETGTAHTKK IEKLNQQIAY IDKDIKNMHS YTCRNYRNLV 900
 AHLNVVSKLQ NYVSELPNDY QITSYFSFYH YCMLGLMEK VSSKNIPLVE SLKNEANDAQ 960
 SYSAKKTLEY FDLIEKNRTY CKDFLKALNA PFSYNLPRFK NLSIEALFDK NIVYEQADLK 1020
 KE 1022

(SEQ ID NO: 112)。

[0258] CasRX/Cas13d蛋白的示例性同向重复序列可以包含以下序列或由其组成：

[0259] CasRX/Cas13d 160582958_基因49834 (SEQ ID NO:112) 包含以下核酸序列或由其组成：

CasRX/Cas13d DR：

gaactacacc cctctgttct ttaggggtc taacac 36

(SEQ ID NO:113)。

[0260] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d 250twins_35838_GL0110300：

MGNKQRVSAQ KRRENAKLCN QQKARQAESQ RDKIKNMNVE KMKNINTNDI KHTKTTAKKL	60
GLKSTIIADK KIILTSFINE QSSKTANIEK VAGFKGDTID TISYTPRMFR SEINPGEIVI	120
SKGDDLSEFA NPANFPPIGRD YVKIRSALEK QYFGKEFPED NLHVQIAYNV ADIKKILSVY	180
INNIIYMFYN LARSEEDIF YNSQSENSGR DCDVIGSLYY QASYRNQDAN RFEKDGKKKA	240
IDSLLDDTRA YYTYFDGLFS VPKREDDGKI KESEKEKAKD QNFDVLRLLS VGRQLTFHSD	300
KSNNEAYLFD LSKLTRAAQD ENRRQDIQSL LNILNSTCRS NLEGVNGDFV KHAKNNLYVL	360
NQLYPSLKAN DLIGEYYNFI VKKENRNIGI RLITVRELII EHNYTNLKDS KYD TYRNKIY	420
TVLNFILFRE IQENSIAIKN FREKLRSTEK AEQPALYQAF ANKIYPMVQA KFAKAIDLFE	480
EQYKTKFKSE FKGGISIENM QQQNILLQTE NIDYFSKYVL FLTKFLDGKE INELLCALIN	540
KFDNIADLLD ISKQIGTPVV FCADYESLND AAKIAENIRL IKNIAHLRPA IQEAQSSKDN	600
ADAAGTPATL LIDAYNMLNT DIQLVYGEAA YEELRKDLFE RKNGTKYNKK GKKVDVYDHK	660
FRNFLINNV I KSKWFFYIAK YVKPADCAKM MSNKKMIEFA LRDLPETQIK RYYTITIGNE	720
ALGDAESLKG VIIEQLHAFS IKNTLLSIKN MGEGEYKIQQ IGSSKEKLKA IVNLYLT VAY	780
LLTKSLVKVN IRFSIAFGCL ERDLVLQKKS EKKFDAINE ILLEDDKIRK ECDKERAQAK	840
TLPRELAQER FAQIKRRESG CYFKSYHVD YLSKNSNEFK QNHIDFAVTS YRNNVEHLNV	900
VHCMTKYFSE VKDVKSYYGV YCYIMQRLC DELIKNQDK PDVRQTFEEY NRLLKDHGTY	960
SKNLMWLLNF PFAYNLARYK NLSNEDLFNA KNNDQKSK	998

(SEQ ID NO: 114)。

[0261] 示例性CasRX/Cas13d蛋白可以包含以下序列或由其组成：

CasRX/Cas13d 250twins_36050_GL0158985：

MKKKHQSAAE KRQVKKLKNQ EKAQKYASEP SPLQSDTAGV ECSQKKT VVS HIASSKTLAK	60
AMGLKSTLVM GDKLVITSFA ASKAVGGAGY KSANIEKITD LQGRVIEEHE RMFSADVGEK	120
NIELSKNDCH TNVNNPVVTN IGKDYIGLKS RLEQEFGKT FENDNLHVQL AYNILDIKKI	180
LGTYVNNIIY IFYNLNRAGT GRDERMYDDL IGTLYAYKPM EAQQT YLLKG DKDMRRFEEV	240
KQLLQNTSAY YVYYGTLFEK VKAKSKKEQR AKEAEIDACT AHNYDVLRLS SLMRQLCMHS	300
VAGTAFKLAE SALFNIEDVL SADLKEILDE AFSGAVNKLN DGFVQHSGNN LYVLQQLYPN	360
ETIERIAEKY YRLTVRKEDL NMGVNIKKLR ELIVGQYFPE VLDKEYDLSK NGDSVVTYRS	420
KIYTVMN YIL LYYLEDHDSS RESMVEALRQ NREGDEGKEE IYRQFAKKVW NGVSGLFGVC	480
LNLFKTEKRN KFRSKVALPD VSGAAYMLSS ENIDYFVKML FVCKFLDGK EINELLCALI	540

```
NKFDNIADIL DAAQCGSSV WFVDSYRFFE RSRRISAQIR IVKNIASKDF KSKKDSDES      600
YPEQLYLDAL ALLGDVISKY KQNRDGSVVI DDQGNVLTQ QYKFRYEFF EEIKRDESGG      660
IKYKSGKPE YNHQRRNFIL NNVLKSKWFF YVVKYNRPSS CRELMKNKEI LRFVLRDIPD      720
SQVRRYFKAV QGEEAYASAE AMRTRLVDAL SQFSVTACLD EVGGMTDKEF ASQRAVDSKE      780
KLRAIIRLYL TVAYLITKSM VKVNTREFSIA FSVLERDYLL LIDGKKKSSD YTGEDMLALT      840
RKFGVEDAGL YREWKEKNAE AKDKYFDKAE RKKVLRQNDK MIRKMHFTPH SLNYVQKNLE      900
SVQSNGLAAV IKEYRNAVAH LNIINRLDEY IGSARADSY SLYCYCLQMY LSKNFSVGYL      960
INVQKQLEEH HTYMKDLMWL LNIPFAYNLA RYKNLSNEKL FYDEEAAAEEK ADKAENERGE    1020
```

(SEQ ID NO: 115)。

[0262] Yan 等人 (2018) Mol Cell. 70 (2) : 327-339 (doi:10.1016/j.molcel.2018.02.2018) 和 Konermann 等人 (2018) Cell 173 (3) : 665-676 (doi:10.1016/j.cell/2018.02.033) 已经描述了 CasRX/Cas13d 蛋白, 将所述两篇参考文献都通过引用以其整体并入本文。还参见 W0 公开号 W0 2018/183703 (CasM) 和 W02019/006471 (Cas13d), 将其通过引用以其整体并入本文。

[0263] 本公开文本的示例性野生型 Cas13d 蛋白可以包含以下氨基酸序列或由其组成: Cas13d (生黄瘤胃球菌 (*Ruminococcus flavefaciens*) XPD3002) 序列:

```
1 IEKKKSFAKG MGVKSTLVSG SKVYMTTFAE GSDARLEKIV EGDSIRSVNE GEAFSAEMAD
61 KNAGYKIGNA KFSHPKGYAV VANNPLYTGP VQQDMLGLKE TLEKRYFGES ADGNDNICIQ
121 VIHNLIDIEK ILAEYITNAA YAVNNISGLD KDIIGFGKFS TVYTYDEFKD PEHHRAAFNN
181 NDKLINAICA QYDEFDNFLD NPRLGYFGQA FFSKEGRNYI INYGNECYDI LALLSGLAHW
241 VVANNEEESR ISRTWLYNLD KNLDNEYIST LNYLYDRITN ELTNSFSKNS AANVNYIAET
301 LGINPAEFAE QYFRFSIMKE QKNLGFNITK LREVMLDRKD MSEIRKNHKV FDSIRTKVYT
361 MMDFVIYRYI IEEDAKVAAA NKSLPDNEKS LSEKDIFVIN LRGSFNDDQK DALYYDEANR
421 IWRKLENIMH NIKEFRGNKT REYKKKDAPR LPRILPAGRD VSAFSKLMYA LTMFLDGKEI
481 NDLLTTLINK FDNIQSFLKV MPLIGVNAKF VEEYAFFKDS AKIADELRLI KSFARMGEPI
541 ADARRAMYID AIRILGTNLS YDELKALADT FSLDENGKSL KKGKHGMRNF IINNVISNKR
601 FHYLIRYGBP AHLHEIAKNE AVVKFVLGRI ADIQKKQGGN GKNQIDRYYE TCIGKDKGKS
661 VSEKVDALTK IITGMNYDQF DKKRSVIEDT GRENAEREKF KKIISLYLTV IYHILKNIVN
721 INARYVIGFH CVERDAQLYK EKGYNLKK LEEKGFSSVT KLCAGIDETA PDKRKDVEKE
781 MAERAKESID SLESANPKLY ANYIKYSDEK KAEFTTRQIN REKAKTALNA YLRNTKWNVI
841 IREDLLRIDN KTCTLFANKA VALEVARYVH AYINDIAEVN SYFQLYHYIM QRIIMNERYE
901 KSSGKVSEYF DAVNDEKKYN DRLLKLLCVP FGYCIPRFKN LSIEALFDRN EAAKFDKEKK
961 KVSGNS (SEQ ID NO: 45) 。
```

[0264] 本公开文本的示例性野生型Cas13d蛋白可以包含以下氨基酸序列或由其组成：

[0265] Cas13d (重叠群e-k87_11092736)：

MRKQKTFKRIGIKSTVAYGQGYAITTFGKGSKAETIAVRSADPPEETLPTESDATLSIHAKFAKAGR
DGREFKCGDVDETRIHTSRSEYESLISNPAESPREDYLGKGTLEKFFGDEYPKDNLRIQIIYSILDIQKILGLY
VEDILHFVDGLQDEPEDLVGLGLGDEKMQKLLSKALPYMGFFGSTDFKVTKKREERAADEHNAKVFRALGAIRQ
KLAHFKWKESLAIFGANANMPIRFFQGATGGRQLWNDVIAPLWKKRIERVRSFLSNSAKNLWVLYQVFKDDTDEK
KKARARQYYHFSVLKEGKNLGFNLTKTREYFLDKFFPIFHSSAPDVKRKVDTFRSKFYAILDFIIYEASVSVANS
QMGKVAPWKGAI DNALVKLREAPDEEAKEKIYNVLAASIRNDSLFLRLKSACDKFGAEQNRPVFPNELRNNRDIRN
VRSEWLEATQDVEDAAAFVQLIAFLCNFLEGKEINELVTALIKKFEGIQALIDLLRNLEGVDSIRFENEFALFNDDK
GNMAGRIARQLRLASVGMKPDMDAKRVLYKSALEILGAPPDEVSEWLAENILLDKSNNDYQKAKKTVNPFNRN
YIAKNVITSRSFYLLVRYAKPTAVRKLSNPKIVRYVLKRLPEKQVASYYSAIWTQSESNSNEMVKLIEMIDRLTT
EIAGFSFAVLKDKKDSIVSASRESRAVNLEVERLKKLTTLYMSIAYIAVKS LVKNARYFIAYSALERDLYFFNEK
YGEEFRLHFIPYELNGKTCQFEYLAILKYYLARDEETLKRKCEICEEIKVGCEKHKKNANPPYEYDQEWIDKKKAL
NSERKACERRLHFSTHWAQYATKRDENMAKHPQKWYDILASHYDELLALQATGWLATQARND AEHLNPVNEFDVYI
EDLRRYPEGTPKNKDYHIGSYFEIYHYIRQRAYLEEVLA KRKEYRDSGSFTDEQLDKLQKILDDIRAGSYDKNLL
KLEYLPFAYNLPYKNTTTEALFDDDSVSGKKRVAEWREKERTREAREQRRQR (SEQ ID NO:46)。

[0266] Cas13d (重叠群e-k87_11092736) (SEQ ID NO:46) 的示例性同向重复序列包含以下核酸序列或由其组成：

[0267] Cas13d (重叠群e-k87_11092736) 同向重复序列：

GTGAGAAGTCTCCTTATGGGGAGATGCTAC (SEQ ID NO:47)。

[0268] 本公开文本的示例性野生型Cas13d蛋白可以包含以下氨基酸序列或由其组成：

[0269] Cas13d (160582958_基因49834)：

MKNSVTFKLIQAQENKEAARKKAKDIAEQARIAKRNGVVKKEENRINRIQIEIQTQKKSNTQNAYHLK
SLAKAAGVKS VFAIGNDLLMTGFGPGNDATIEKRVFQNRAIETLSSPEQYSAEFQNKQFKIKGNIKVLNHSTQKME
EIQTELQDNYNRPHFDLLGCKNVLEQKYFGRTFSDNIHVQIAYNIMDIEKLLTPYINNIIYTLNELMRD NSKDDFF
GCDSHFSVAYLYDELKAGYSRDLKTKPNLSKNIDRIWNNFCNYMNSDSGNTEARLAYFGELFYKPKETGDAKSDYK
THLSNNQKEEWELKSDKEVYNIFAILCDLRHFCTHGESITPSGKPFYPNLEKNLFPEAKQVLNSLFEEKAESLGAE
AFGKTAGKTDVSILLKVFEKEQASQKEQQALLKEYYDFKVQKTYKNMGFSIKKLREAIM EIPDAAKFDDLYSSLR
HKLYGLFDFILVKHFLDTSSEN LQNNIDFRQLRACRCEEEKDQVYRSIAVKVWEKVKKELNMFQV VVIPSLSK
DELKQMEMTKNTELLSSIETISTQASLFSEMIFMMTYLLDGKEINLLCTSLIEKFENIASFNEVLKSPQIGYETKY
TEGYAFFKNADKTAKELRQVNNMARM TKPLGGVNTKCVMYNEAAKILGAKPMSKAELESVFNLDNHDTYSPSGKK
IPNKNFRNFIINN VITSRRFLYLIRYGNPEKIRKIAINPSIISFVLKQIPDEQIKRYYP PCIGKRTDDVTLMRDEL
GKMLQSVNFEQFSRVNNKQNAKQNPNGEKARLQACVRLYLTVPYLF IKNMVNINARYVLA FHCLERDHALCFNSRK
LNDDSYNEMANKFQMVRKAKKEQEYKEYCKKQETGTAHTKKIEKLNQQIAYIDKDIKNMHSYTCRNYRN LVAHLN
VVSKLQNYVSELPNDYQITSYFSFYHYCMQLGLMEKVSSKNIPLVESLKNEANDAQSYS AKKTLEYFDLIEKNRTY
CKDFLKALNAPFSYNLPRFKNLSIEALFDKNIVYEQADLKKE (SEQ ID NO:48)。

[0270] Cas13d (160582958_基因49834) (SEQ ID NO:48) 的示例性同向重复序列包含以下核酸序列或由其组成：

[0271] Cas13d (160582958_基因49834) 同向重复序列：

GAACTACACCCCTCTGTTCTTGTAGGGGTCTAACAC (SEQ ID NO:49)。

[0272] 本公开文本的示例性野生型Cas13d蛋白可以包含以下氨基酸序列或由其组成：

[0273] Cas13d(重叠群tpg|DJXD01000002.1|；未培育的瘤胃球菌属联合体,UBA7013,来自绵羊肠道宏基因组)：

MKKQKSKKTVSKTSGLKEALSVQGTIVMTSFGKGNMANLSYKIPSSQKPQNLNSSAGLKNVEVSGKKI
KFQGRHPKIATTDNPLFKPQPGMDLLCLKDKLEMHYFGKTFDDNIHIQLIYQILDIEKILAVHVNNIVFTLDNVLH
PQKEELTEDFIGAGGWRINLDYQTLRGQTNKYDRFKNYIKRKELLYFGEAFYHENERRYEEDIFAILTLLSALRQF
CFHSDLSSDES DHVNSFWLYQLEDQLSDEFKETLSILWEEVTERIDSEFLKTNTVNLHILCHVFPKESKETIVRAY
YEFLIKKSFKNMGFSIKKLREIMLEQSDLKSFKEDKYNVRAKLYKLDFDIITYYYDHHAFEKEALVSSLRSSLTE
ENKEEIIYIKTARTLASALGADFKKAAADVNAKNIRDYQKKANDYRISFEDIKIGNTGIGYFSELIYMLTLLLDGKE
INDLLTTLINKFDNIISFIDILKKLNLEFKFKPEYADFFNMTNCRYTLEELRVINSIARMQKPSADARKIMYRDAL
RILGMDNRPDEEIDRELERTMPVGADGKFIKKGQGRNFIASNVISSRFHYLVRYNNPHKTRTLVKNPNVVKFVL
EGIPETQIKRYFDVCKGQEIPPTSDKSAQIDVLARIISSVDYKIFEDVPQSAKINKDDPSRNFSDALKKQRYQAIV
SLYLTVMYLITKNLVYVNSRYVIAFHCLERDAFLHGVTLPKMNKKIVYSQLTTHLLTDKNYTTYGHLKNQKGHRKW
YVLVKNLQNSDITAVSSFRNIVAHISVVRNSNEYISGIGELHSYFELYHYLVQSMIAKNNWYDTSHPKTAEYLN
NLKKHHTYCKDFVKAYCIPFGYVVPYKNTINELFDRNNPNPEPKKEEV (SEQ ID NO:50)。

[0274] Cas13d(重叠群tpg|DJXD01000002.1|；未培育的瘤胃球菌属联合体,UBA7013,来自绵羊肠道宏基因组) (SEQ ID NO:50) 的示例性同向重复序列包含以下核酸序列或由其组成：

[0275] Cas13d(重叠群tpg|DJXD01000002.1|；未培育的瘤胃球菌属联合体,UBA7013,来自绵羊肠道宏基因组) :CAACTACAACCCCGTAAAAATACGGGGTCTGAAAC (SEQ ID NO:51)。

[0276]

[0277] 在本公开文本的一些实施方案中,CjeCas9-内切核酸酶融合物和gRNA分子可以包含以下的核酸序列或由其组成：

E43-CjeCas9和sgRNA质粒 (U6:N=sgRNA间隔子,E43,CjeCas9) gtttattacaggac
agcagagatccagtttggttaattaaggtaccgagggcctatttcccatgattccttcataatttgcataatacgata
caaggctgtagagagataattagaattaatttgactgtaaacacaaagatatattagtacaaaatacgtgacgtaga
aagtaataatttcttggttagtttgagttttaaataattatgttttaaaatggactatcatatgcttacgtaactt
gaaagtatttctgatttcttggtttatatatcttGTGGAAAGGACGAAACACCNNNNNNNNNNNNNNNNNNGTTT
TAGTCCCTGAAGGGACTAAAAATAAGAGTTTGCAGGACTCTGCGGGTTACAATCCCCTAAAACCGCTTTTTTTCC
TGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAG
GGTTAATTGCGGAATTCGCTAGCTAGGTCTTGAAGGAGTGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGC
GCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGG
GGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCA
GTAGTCGCGGTGAACGTTCTTTTCGCAACGGGTTTGCCGCCAGAACACAGGACCGGTTCTAGAGCGCTATTTAGA
ACCatgTGTTCCTCCCAAGAATCTGGCATGACCGCTCTTTCAGCGAGGATGTTGACGCGAAGCAGATCCCTGGGAC
CTGGGGCCGGGCCACGAGGGTGTGCGGAAGAACCAGGACCGTTGCGACGGAGGGAAGCAGCAGCGGAAGCTCGGAA
ATCCCATTCTCCGGTTAAACGACCCCGCAAGGCACAACGGCTCAGGGTTGCTTACGAGGGGAGCGATTCCGAAAAG
GGTGAAGGAGCAGAGCCCTTGAAGGTTCCAGTATGGGAACCCAGGATTGGCAGCAGCAGCTTGTAACATCCGAG

CAATGAGGAACAAAAAGATGCACCTGTTGATCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAA
AGTCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGGACCAGGTTACGGCCGGAGCAATG
CAACGGCTTCGGGCACGGGACTCACGGTCGATAGCATTTCGAGACCGATGACGCAACATTGGGTAAACTCATAT
ATCCAGTTGGCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAACAACATTACGGAGG
TGATATACCCGCAAGCGTAGCTGAACTGGTAGCACTGCCGGGCGTCGGTCCCAAAATGGCACATCTGGCTATGGCG
GTTGCTTGGGGAACGGTGTCTGGTATCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGGTGGACTA
AAAAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGAGTGGCTTCCACGAGAGCTGTGGCATGAAAT
CAATGGATTGCTGGTTGGTTTCGGGCAGCAGACATGCTTGCCCGTGCACCCCCGGTGTCTATGCTTGCTTGAACCAG
GCTTTGTGCCAGCTGCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATCTGAGTCTGCGACCCCGGAGAGCa
caaacGCGGAATCCTGGCCTTCGcgATTGGCATTAGCAGCATCGGCTGGGCATTCTCTGAAAACGACGAACTGAA
GGATTGCGGCGTGCGAATTTTCACTAAGGTCGAAAATCCCAAACTGGTGAATCACTCGCTCTCCCTAGACGACTG
GCACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCATCTTAAACACCTTATTGCAAATGAGT
TTAAACTGAATTATGAGGACTACCAATCCTTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCCC
GTATGAGCTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTTGCTCGCGTGATTCTCCACATAGCG
AAGAGGCGAGGATACGATGACATTA AAAACAGTGATGATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAGCAAA
ATGAAGAGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCCAGAAGTTCAAAGAAAA
TAGCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGAAAGATGTATTGCGCAATCTTTCCTCAAGGAC
GAGCTCAAATTGATTTTCAAGAAACAAAGGGAATTTGGGTTTCAGCTTCTCAAAAAAATTTGAGGAAGAGGTTCTGA
GCGTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAATTGTAGTTTCTTCACCGATGAAAA
ACGGGCGCCAAAAAATAGCCCTTTGGCTTTTATGTTTGTGCTCTGACTCGCATCATTAACTGCTCAACAACCTT
AAAAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAACGAAGTTTGAAGAACGGGACTT
TGACCTACAAACAAACCAAAAAGCTTCTTGGTCTCAGTGATGACTACGAATTCAAGGGAGAAAAAGGGACATATTT
CATCGAATTCAAGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACCTGTCTCAAGATGATCTCAACGAA
ATTGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGCGTTGGCGAAGTATGACCTTAACCAAA
ACCAAATAGATAGCCTCAGCAAGTTGGAATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGTCAC
CCCCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATCTCAAGGTTGCTATTAACGAAGAC
AAAAAAGATTTCTCCAGCTTTCAATGAGACTTACTATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCA
TCAAAGAGTATCGAAAGGTCTGAATGCTTTGCTCAAAAAATACGGTAAGGTACACAAAATAAATATTGAGCTCGC
AAGGGAGGTCGGTAAGAACCACTCCCAGCGCGCCAAAATAGAAAAGGAACAGAATGAAAATTACAAAGCGAAAAAG
GACGCCGAGCTCGAGTGCGAAAAGCTGGGCCTGAAAATAAACAGCAAGAACATTCTCAAACCTCCGCTCTTCAAAG
AACAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAATCTCCGATCTTCAAGACGAGAAGATGCTCGAAAT
AGACgcgATATATCCATATAGCAGGTCTTTTGACGATTCTTACATGAATAAAGTGCTTGTTTTCTACTAAGCAGAAT
CAGGAAAAGTTGAATCAGACCCCTTTGAGGCCTTTGGCAACGACTCAGCAAAGTGGCAGAAGATCGAGGTCTTGG
CTAAGAATCTTCTACTAAGAAACAGAAAAGGATATTGGATAAGAACTATAAAGACAAAGAACAAAAGAACTTTAA
AGACCGCAACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACACAAAAGATTATTTGGACTTCTTG
CCGCTGTCTGATGATGAGAACACGAACTCAACGACACGCAAAAGGGGTCTAAAGTCCACGTCTGAAGCTAAATCTG
GGATGCTCACCTCAGCATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACCATGCCAT
TGACGCAGTTATCATAGCGTATGCCAATAATTCATAGTAAAAGCGTTTAGCGACTTCAAGAAGGAACAAGAGTCC
AACAGCGCCGAGCTCTACGCAAAAAAGATTAGTGAACCTGACTACAAAAACAAAAGAAAATCTTTGAGCCGTTCA

GCGGATTTTCGACAGAAGGTATTGGATAAAATAGATGAAATTTTCGTGAGCAAACCCGAAAGGAAAAAGCCCTCAGG
CGCCTTGCACGAAGAGACTTTCAGGAAGGAAGAGGAATTCTACCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAG
GCTCTCGAACTTGAAAAGATTAGGAAGGTGAACGGCAAGATAGTGAAAAACGGCGATATGTTCCGGGTGATATCT
TCAAACATAAAAAACGAATAAATTTTATGCTGTGCCTATATACACTATGGACTTCGCACCTAAGGTCTGCCGAA
TAAGGCGGTAGCCCGATCTAAAAAAGGCGAAATTAAGGACTGGATTTTGTATGGATGAAAATTACGAGTTCTGCTTT
TCTCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACCGGAATTCGTGTATTACAACGCTT
TTACTTCCTCTACGGTATCTTTGATTGTCTCCAAACATGACAACAAATTCGAAACACTCAGTAAAAACCAAAGAT
TCTCTTTAAAAATGCGAACGAGAAAAGAAGTAATTGCAAAATCAATTGGCATCCAAAATTTGAAAGTTTTTGAAAAA
TATATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACAGCGAGAGGACTTCAAAAAATCAGGTCCAC
CCAAGAAAAACGCAAGGTGGAAGATCCGAAGAAAAAGCGAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCGGCT
GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCAACTGTTTATTGCAGCTTATAATGG
TTACAAATAAAGCAATAGCATCACAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCC
AAACTCATCAATGTATCTTATCATGTCTGTATACCG (SEQ ID NO:202)。

[0278] 在本公开文本的一些实施方案中,CjeCas9-内切核酸酶融合物和gRNA分子可以包含以下的核酸序列或由其组成:

E67-CjeCas9和sgRNA质粒 (U6:N=sgRNA间隔子,E67,CjeCas9) gtttattacagggac
agcagagatccagtttggttaattaaggtaccgagggcctatttcccatgattccttcataatttgcatatacgata
caaggctgttagagagataattagaattaatttgactgtaaacacaaagatatagtagacaaaatacgtgacgtaga
aagtaataatttcttggttagtttgagttttaaattatgttttaaaatggactatcatatgcttaccgtaactt
gaaagtatttctgatttcttggtttatatatcttGTGGAAAGGACGAAACACCNNNNNNNNNNNNNNNNNNGTTT
TAGTCCCTGAAGGGACTAAAATAAAGAGTTGCGGGACTCTGCGGGTTACAATCCCCTAAAACCGCTTTTTTTC
TGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCTTTAGTGAG
GGTTAATTGCGGAATTCGCTAGCTAGGTCTTGAAAGGAGTGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGC
GCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGATCCGGTGCCTAGAGAAGGTGGCGCGG
GGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCA
GTAGTCGCGGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGACCGGTTCTAGAGCGCTATTTAGA
ACCAatgCAGGAGGTAATAGCGGGGCTTGAGCGATTTACCTTTGCCTTCGAAAAAGACGTAGAGATGCAGAAGGGAA
CCGGCCTGCTCCCATTTCAAGGTATGGACAAATCAGCATCTGCCGTGTGCAATTTTTTCACCAAGGGTCTGTGTGA
AAAGGGGAAGCTCTGTCCATTTGCCATGATCGCGGAGAGAAGATGGTGGTGTGTAAGCACTGGCTGAGAGGGCTT
TGCAAAAAAGGCGACCACTGCAATTTCTTACCAATATGACCTGACTCGAATGCCTGAGTGTTATTTTTACAGTA
AGTTCCGTGACTGTAGCAACAAAGAATGCAGCTTCTTGCATGTCAAACCAGCATTCAAGTCACAGGATTGCCCGTG
GTACGATCAGGGTTTTTGCAAGGACGGTCCCCTCTGCAATATCGACACGTACCCAGAATTATGTGCCTTAATTAC
CTGGTCGGCTTCTGTCTGAAGGGCCAAAATGTCAGTTTGCTCAAAAAATTCGCGAGTTCAAATTGCTCCCTGGGT
CTAAAATTTGGGAACCCAGGATTGGCAGCAGCAGCTTGTAACATCCGAGCAATGAGGAACAAAAAAGATGCACC
TGTTGATCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAAAAGTCCGCCGGTATCAGGTTCTGTTG
AGTTTGATGCTGAGTAGTACAGTAAGGACCAGGTTACGGCCGGAGCAATGCAACGGCTTCGGGCACGGGGACTCA
CGGTCGATAGCATTTTGCAGACCGATGACGCAACATTGGGTAAACTCATATATCCAGTTGGCTTCTGGCGGAGCAA
AGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAACAACATTACGGAGGTGATATACCCGCAAGCGTAGCTGAA
CTGGTAGCACTGCCGGGCGTCGGTCCCAAAATGGCACATCTGGCTATGGCGGTTGCTTGGGGAACGGTGTCTGGTA

TCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGGTGGACTAAAAAGCCACTAAGTCTCCTGAAGA
AACACGGGCTGCTCTGGAAGAGTGGCTTCCACGAGAGCTGTGGCATGAAATCAATGGATTGCTGGTTGGTTTCGGG
CAGCAGACATGCTTGCCCGTGCACCCCGGTGTCATGCTTGCTTGAACCAGGCTTTGTGCCAGCTGCCAGGGCC
TGAGTGGAAGTGAGACACCGGGAACATCTGAGTCTGCGACCCCGAGAGCacaaacGCGCGAATCCTGGCCTTCGc
gATTGGCATTAGCAGCATCGGCTGGGCATTCTCTGAAAACGACGAACTGAAGGATTGCGGCGTGCGAATTTTCACT
AAGGTCGAAAAATCCCAAACTGGTGAATCACTCGCTCTCCCTAGACGACTGGCACGCTCCGCACGAAAGAGGCTTG
CCCGCCGCAAGGCACGCTTGAACCATCTTAAACACCTTATTGCAATGAGTTTAACTGAATTATGAGGACTACCA
ATCCTTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCCCGTATGAGCTCCGGTTCAGAGCACTC
AACGAACTGCTGTCCAAACAGGATTTTGCTCGCGTGATTCTCCACATAGCGAAGAGGCGAGGATACGATGACATTA
AAAACAGTGATGATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAGCAAAATGAAGAGAAGCTCGCTAACTATCA
ATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCCAGAAGTTCAAAGAAAATAGCAAGGAATTTACTAATGTCCGG
AATAAAAAGGAGTCTTACGAAAGATGTATTGCGCAATCTTTCCTCAAGGACGAGCTCAAATTGATTTTCAAGAAAC
AAAGGGAATTTGGGTTCACTTCTCAAAAAATTTGAGGAAGAGGTTCTGAGCGTTGCCTTTTACAAACGCGCCCT
TAAGGACTTCTCACATCTCGTAGGGAATTGTAGTTTCTTACCGATGAAAAACGGGCGCCAAAAAATAGCCCTTTG
GCTTTTATGTTTGTGCTCTGACTCGCATTAATCTGCTCAACAACCTTAAAAACACGGAAGGGATTCTGTACA
CAAAGGATGATCTGAACGCTCTGCTTAACGAAGTTTTGAAGAACGGGACTTTGACCTACAAACAAACCAAAAAGCT
TCTTGGTCTCAGTGATGACTACGAATTCAAGGGAGAAAAAGGGACATATTTTCATCGAATTCAAGAAGTATAAGGAG
TTCATCAAAGCCTTGGGCGAGCACAACCTGTCTCAAGATGATCTCAACGAAATTGCTAAGGATATCACTCTGATTA
AAGACGAGATCAAGCTCAAAAAGGCGTTGGCGAAGTATGACCTTAACCAAAACCAAAATAGATAGCCTCAGCAAGTT
GGAATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGTCACCCCTTGATGCTTGAAGGAAAGAAA
TATGATGAGGCATGTAATGAGCTGAATCTCAAGGTTGCTATTAACGAAGACAAAAAAGATTTCTCCAGCTTTCA
ATGAGACTTACTATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCATCAAAGAGTATCGAAAGGTCCTGAA
TGCTTTGCTCAAAAAATACGGTAAGGTACACAAAATAAATATTGAGCTCGCAAGGGAGGTGCGTAAGAACCACTCC
CAGCGCGCCAAAAATAGAAAAGGAACAGAATGAAAATTACAAAGCGAAAAAGGACGCCGAGCTCGAGTGCGAAAAGC
TGGGCCTGAAAAATAAACAGCAAGAACATTCTCAAACCTCCGCTCTTCAAAGAACAAAAAGAATTTTGTGCTTATAG
TGGTGAGAAAAATAAAATCTCCGATCTTCAAGACGAGAAGATGCTCGAAATAGACgcgATATATCCATATAGCAGG
TCTTTTGACGATTCTTACATGAATAAAGTGCTTGTTTTCACTAAGCAGAATCAGGAAAAGTTGAATCAGACCCCT
TTGAGGCCTTTGGCAACGACTCAGCAAAGTGGCAGAAGATCGAGGTCTTGGCTAAGAATCTTCCTACTAAGAAACA
GAAAAGGATATTGGATAAGAACTATAAAGACAAAGAACAAAGAAGCTTTAAAGACCGCAACCTCAATGACACCAGA
TACATAGCAAGATTGGTTCTGAACCTACACAAAAGATTATTTGGACTTCTTGCCGCTGTCTGATGATGAGAACACGA
AACTCAACGACACGCAAAAAGGGTCTAAAGTCCACGTCAAGCTAAATCTGGGATGCTCACCTCAGCATTGAGGCA
TACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACCATGCCATTGACGCAGTTATCATAGCGTATGCC
AATAATTCAATAGTAAAAGCGTTTAGCGACTTCAAGAAGGAACAAGAGTCCAACAGCGCCGAGCTCTACGCAAAAA
AGATTAGTGAACCTGACTACAAAAACAAAGAAAATTCTTTGAGCCGTTTCAAGGATTTCGACAGAAGGTATTGGA
TAAAATAGATGAAATTTTCGTGAGCAAACCCGAAAGGAAAAAGCCCTCAGGCGCCTTGACGAAGAGACTTTTCAAG
AAGGAAGAGGAATTCTACCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAGGCTCTCGAACTTGGAAGATTAGGA
AGGTGAACGGCAAGATAGTAAAAACGGCGATATGTTCCGGGTTGATATCTTCAAACATAAAAAAACGAATAAATT
TTATGCTGTGCCTATATACACTATGGACTTCGCACTTAAGGTCTGCCGAATAAGGCGGTAGCCCGATCTAAAAAA
GGCGAAATTAAGGACTGGATTTTGATGGATGAAAATTACGAGTTCTGCTTTTCTCTCTACAAGGATTCCCTTATAT

TGATACAGACGAAAGATATGCAGGAACCGGAATTCGTGTATTACAACGCTTTTACTTCCTCTACGGTATCTTTGAT
TGTCTCCAAACATGACAACAAATTCGAAACACTCAGTAAAAACCAAAAGATTCTCTTTAAAAATGCGAACGAGAAA
GAAGTAATTGCAAAATCAATTGGCATCCAAATTTGAAAGTTTTTGAAAAATATATAGTATCTGCCCTCGGAGAGG
TTACTAAAGCGGAATTTAGACAGCGAGAGGACTTCAAAAAATCAGGTCCACCCAAGAAAAAACGCAAGGTGGAAGA
TCCGAAGAAAAAGCGAAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCT
CATGCTGGAGTTCTTCGCCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACA
AATTTTACAAATAAAGCATTTTTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATG
TCTGTATACCG (SEQ ID NO:203)。

gRNA靶序列

[0279] 在本公开文本的组合物的一些实施方案中，RNA分子的靶序列包含对应于第一RNA结合蛋白和/或第二RNA结合蛋白的序列基序。

[0280] 在本公开文本的组合物和方法的一些实施方案中，所述序列基序是疾病或障碍的标志。

[0281] 本公开文本的序列基序可以从基因组序列中发现的外来或外源序列的序列分离或衍生，并且因此翻译为本公开文本的mRNA分子或在本公开文本的RNA序列中发现的外来或外源序列的序列。

[0282] 本公开文本的序列基序可以包含内源序列中引起疾病或障碍的突变或由其组成。所述突变可以包含序列取代、倒位、缺失、插入、转座或其任何组合，或者由其组成。

[0283] 本公开文本的序列基序可以包含重复的序列或由其组成。在一些实施方案中，所述重复的序列可能与微卫星不稳定性 (MSI) 相关。一个或多个基因座处的MSI是由于本公开文本的细胞的DNA错配修复机制受损所致。可以将DNA的超变序列转录为本公开文本的包含靶序列的mRNA，所述靶序列包含所述超变序列或由其组成。

[0284] 本公开文本的序列基序可以包含生物标记或由其组成。所述生物标记可以指示患上疾病或障碍的风险。所述生物标记可以指示健康基因 (低或无可确定的患上疾病或障碍的风险)。所述生物标记可以指示编辑的基因。示例性生物标记包括但不限于单核苷酸多态性 (SNP)、序列变异或突变、表观遗传标记、剪接受体位点、外源序列、异源序列及其任何组合。

[0285] 本公开文本的序列基序可以包含二级、三级或四级结构或由其组成。所述二级、三级或四级结构可以是内源的或天然存在的。所述二级、三级或四级结构可以是诱导的或非天然存在的。所述二级、三级或四级结构可以由内源、外源或异源序列编码。

[0286] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列包含在2个与100个之间的核苷酸或核酸碱基 (包括端点) 或由其组成。在一些实施方案中，RNA分子的所述靶序列包含在2个与50个之间的核苷酸或核酸碱基 (包括端点) 或由其组成。在一些实施方案中，RNA分子的所述靶序列包含在2个与20个之间的核苷酸或核酸碱基 (包括端点) 或由其组成。

[0287] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列是连续的。在一些实施方案中，RNA分子的所述靶序列是不连续的。例如，RNA分子的所述靶序列可以包含不连续的一个或多个核苷酸或核酸碱基或由其组成，因为一个或多个间断的核苷酸定位于所述靶序列的核苷酸之间。

[0288] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列是天然存在的。在一些实施方案中，RNA分子的所述靶序列是非天然存在的。示例性的非天然存在的靶序列可以包含序列变异或突变、嵌合序列、外源序列、异源序列、嵌合序列、重组序列、包含修饰的或合成的核苷酸的序列或其任何组合，或者由其组成。

[0289] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列与本公开文本的指导RNA结合。

[0290] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列与本公开文本的第一RNA结合蛋白结合。

[0291] 在本公开文本的组合物和方法的一些实施方案中，RNA分子的靶序列与本公开文本的第二RNA结合蛋白结合。

RNA分子

[0292] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子包含靶序列。在一些实施方案中，本公开文本的RNA分子包含至少一个靶序列。在一些实施方案中，本公开文本的RNA分子包含一个或多个靶序列。在一些实施方案中，本公开文本的RNA分子包含两个或更多个靶序列。

[0293] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子是天然存在的RNA分子。在一些实施方案中，本公开文本的RNA分子是非天然存在的分子。示例性的非天然存在的RNA分子可以包含序列变异或突变、嵌合序列、外源序列、异源序列、嵌合序列、重组序列、包含修饰的或合成的核苷酸的序列或其任何组合，或者由其组成。

[0294] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子包含从病毒分离或衍生的序列或由其组成。

[0295] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子包含从原核生物分离或衍生的序列或由其组成。在一些实施方案中，本公开文本的RNA分子包含从古菌的物种或菌株或者细菌的物种或菌株分离或衍生的序列或由其组成。

[0296] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子包含从真核生物分离或衍生的序列或由其组成。在一些实施方案中，本公开文本的RNA分子包含从以下的物种分离或衍生的序列或由其组成：原生动物、寄生虫、原生生物、藻类、真菌、酵母、变形虫、蠕虫、微生物、无脊椎动物、脊椎动物、昆虫、啮齿类动物、小鼠、大鼠、哺乳动物或灵长类动物。在一些实施方案中，本公开文本的RNA分子包含从人分离或衍生的序列或由其组成。

[0297] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的RNA分子包含从生物或病毒的基因组的编码序列衍生的序列或由其组成。在一些实施方案中，本公开文本的RNA分子包含初级RNA转录物、前体信使RNA(前体mRNA)或信使RNA(mRNA)或由其组成。在一些实施方案中，本公开文本的RNA分子包含尚未加工的基因产物(例如转录物)或由其组成。在一些实施方案中，本公开文本的RNA分子包含已经进行转录后加工的基因产物(例如包含5'帽和3'多聚腺苷酸化信号的转录物)或由其组成。在一些实施方案中，本公开文本的RNA分子包含已经进行选择剪接的基因产物(例如剪接变体)或由其组成。在一些实施方案中，本公开文本的RNA分子包含已经进行非编码序列和/或内含子序列的去除的基因产物(例如信使RNA(mRNA))或由其组成。

[0298] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的RNA分子包含从非编码序列衍生的序列(例如非编码RNA(ncRNA))或由其组成。在一些实施方案中,本公开文本的RNA分子包含核糖体RNA或由其组成。在一些实施方案中,本公开文本的RNA分子包含小ncRNA分子或由其组成。本公开文本的示例性小RNA分子包括但不限于微小RNA(miRNA)、小干扰(siRNA)、piwi相互作用RNA(piRNA)、核仁小RNA(snoRNA)、小核RNA(snRNA)、细胞外或外泌体RNA(exRNA)和小卡哈尔体特异性RNA(scaRNA)。在一些实施方案中,本公开文本的RNA分子包含长ncRNA分子或由其组成。本公开文本的示例性长RNA分子包括但不限于X染色体失活特异性转录物(Xist)和HOX转录物反义RNA(HOTAIR)。

[0299] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在细胞内间隙中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在胞质溶胶面中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在核中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在囊泡、细胞的膜结合区室或细胞器中接触。

[0300] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在细胞外间隙中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在外泌体中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在脂质体、聚合物囊泡、胶束或纳米颗粒中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在细胞外基质中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在液滴中接触。在一些实施方案中,本公开文本的RNA分子与本公开文本的组合物在微流体液滴中接触。

[0301] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的RNA分子包含单链序列或由其组成。在一些实施方案中,本公开文本的RNA分子包含双链序列或由其组成。在一些实施方案中,所述双链序列包含两个RNA分子。在一些实施方案中,所述双链序列包含一个RNA分子和一个DNA分子。在一些实施方案中,包括以下那些实施方案:其中所述双链序列包含一个RNA分子和一个DNA分子,本公开文本的组合物选择性结合并任选地选择性切割所述RNA分子。

融合蛋白

[0302] 在本公开文本的组合物和方法的一些实施方案中,所述组合物包含编码靶RNA结合融合蛋白的序列,所述序列包含(a)编码第一RNA结合多肽或其部分的序列;以及(b)编码第二RNA结合多肽的序列,其中所述第一RNA结合多肽结合靶RNA,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0303] 在一些实施方案中,靶RNA结合融合蛋白是RNA指导的靶RNA结合融合蛋白。RNA指导的靶RNA结合融合蛋白包含至少一种RNA结合多肽,其对应于将所述RNA结合多肽指导至靶RNA的gRNA。RNA指导的靶RNA结合融合蛋白包括但不限于RNA结合多肽,其是基于CRISPR/Cas的RNA结合多肽或其部分。

[0304] 在一些实施方案中,靶RNA结合融合蛋白不是RNA指导的靶RNA结合融合蛋白,并且因此包含至少一种RNA结合多肽,其能够在没有相应gRNA序列的情况下结合靶RNA。此类非指导的RNA结合多肽包括但不限于作为PUF(Pumilio和FBF同源家族)的至少一种RNA结合蛋白或其RNA结合部分。这种类型的RNA结合多肽可以用于代替gRNA指导的RNA结合蛋白如

CRISPR/Cas。参与介导mRNA稳定性和翻译的PUF蛋白(以果蝇(*Drosophila*) Pumilio和秀丽隐杆线虫(*C.elegans*) fem-3结合因子命名)的独特RNA识别模式是本领域中熟知的。也是本领域中已知的人Pumilio1的PUF结构域与同源RNA序列紧密结合,并且可以修饰其特异性。其含有八个PUF重复,它们识别八个保守RNA碱基,且每个重复识别单个碱基。由于每个重复中的两条氨基酸侧链识别相应碱基的Watson-Crick边缘并决定该重复的特异性,PUF结构域可以被设计为特异性结合大多数8-nt RNA。Wang等人,Nat Methods.2009;6(11):825-830。还参见WO 2012/068627,将其通过引用以其整体并入本文。

[0305] 在本公开文本的非指导的RNA结合融合蛋白的一些实施方案中,所述融合蛋白包含作为PUMBY(基于Pumilio的联合体)蛋白的至少一种RNA结合蛋白或其RNA结合部分。已经以天然和修饰形式广泛用于靶向RNA的RNA结合蛋白PumHD(Pumilio同源结构域,PUF家族的成员)已经被工程化以产生一组四个规范蛋白质模块,其中的每个靶向一个RNA碱基。这些模块(即,Pumby,代表基于Pumilio的联合体)可以以不同组成和长度的链进行链状结合,以结合所需靶RNA。此类Pumby-RNA相互作用的特异性高,且Pumby链与携带相对于靶序列的三个或更多个错配的RNA序列的结合不可检测。Katarzyna等人,PNAS,2016;113(19):E2579-E2588。还参见US2016/0238593,将其通过引用以其整体并入本文。

[0306] 在本公开文本的组合物的一些实施方案中,至少一种RNA结合蛋白或其RNA结合部分是PPR蛋白。PPR蛋白(从植物衍生的具有三角状五肽重复(PPR)基序的蛋白质)是核编码的并且仅在RNA水平上受控制细胞器(叶绿体和线粒体),切割,翻译,剪接,RNA编辑,特异性作用于RNA稳定性的基因。PPR蛋白通常是35个氨基酸的基序,并且具有以下结构,其中PPR基序是约10个连续氨基酸。PPR基序的组合可以用于与RNA的序列选择性结合。PPR蛋白通常由约10个重复结构域的PPR基序构成。PPR结构域或RNA结合结构域可以被配置为无催化活性的。将WO 2013/058404通过引用以其整体并入本文。

[0307] 在一些实施方案中,本文公开的融合蛋白在所述至少两种RNA结合多肽之间包含接头。在一些实施方案中,所述接头是肽接头。在一些实施方案中,所述肽接头包含三肽GGG的一个或多个重复。在其他实施方案中,所述接头是非肽接头。在一些实施方案中,所述非肽接头包含聚乙二醇(PEG)、聚丙二醇(PPG)、共-聚(乙二醇/丙二醇)、聚氧乙烯(POE)、聚氨基甲酸酯、聚磷腈、多糖、葡聚糖、聚乙烯醇、聚乙烯吡咯烷酮、聚乙烯基乙醚、聚丙烯酰胺、聚丙烯酸酯、聚氰基丙烯酸酯、脂质聚合物、甲壳素、透明质酸、肝素或烷基接头。

[0308] 在一些实施方案中,所述至少一种RNA结合蛋白不需要多聚化以用于RNA结合活性。在一些实施方案中,所述至少一种RNA结合蛋白不是多聚体复合物的单体。在一些实施方案中,多聚体蛋白复合物不包含所述RNA结合蛋白。在一些实施方案中,所述至少一种RNA结合蛋白与所述RNA分子内的靶序列选择性结合。在一些实施方案中,所述至少一种RNA结合蛋白不包含对所述RNA分子内的第二序列的亲和力。在一些实施方案中,所述至少一种RNA结合蛋白不包含对所述RNA分子内的第二序列的高亲和力或不选择性结合所述第二序列。在一些实施方案中,所述至少一种RNA结合蛋白包含在2个与1300个之间的氨基酸,包括端点。

[0309] 在一些实施方案中,本文公开的编码融合蛋白的所述至少一种RNA结合蛋白的序列还包含编码核定位信号(NLS)的序列。在一些实施方案中,编码核定位信号(NLS)的所述序列定位于编码所述RNA结合蛋白的序列的3'。在一些实施方案中,所述至少一种RNA结合

蛋白包含在所述蛋白质的C末端的NLS。在一些实施方案中,编码所述至少一种RNA结合蛋白的序列还包含编码第一NLS的第一序列和编码第二NLS的第二序列。在一些实施方案中,编码所述第一NLS或所述第二NLS的序列定位于编码所述RNA结合蛋白的序列的3'。在一些实施方案中,所述至少一种RNA结合蛋白包含在所述蛋白质的C末端的第一NLS或第二NLS。在一些实施方案中,所述至少一种RNA结合蛋白还包含NES(核输出信号)或其他肽标签或分泌信号。

[0310] 在一些实施方案中,本文公开的融合蛋白包含所述至少一种RNA结合蛋白作为第一RNA结合蛋白以及包含核酸酶结构域或由其组成的第二RNA结合蛋白。在一些实施方案中,所述第二RNA结合蛋白以与RNA缔合的方式结合RNA。在一些实施方案中,所述第二RNA结合蛋白以切割RNA的方式与RNA缔合。

[0311] 在一些实施方案中,所述第二RNA结合多肽被可操作地配置到在所述第一RNA结合多肽的C末端的第一RNA结合多肽。在一些实施方案中,所述第二RNA结合多肽被可操作地配置到在所述第一RNA结合多肽的N末端的第一RNA结合多肽。

载体

[0312] 在本公开文本的组合物和方法的一些实施方案中,载体包含本公开文本的指导RNA。在一些实施方案中,所述载体包含本公开文本的至少一种指导RNA。在一些实施方案中,所述载体包含本公开文本的一种或多种指导RNA。在一些实施方案中,所述载体包含本公开文本的两种或更多种指导RNA。在一些实施方案中,所述载体还包含本公开文本的融合蛋白。在一些实施方案中,所述融合蛋白包含第一RNA结合蛋白和第二RNA结合蛋白。

[0313] 在本公开文本的组合物和方法的一些实施方案中,第一载体包含本公开文本的指导RNA,并且第二载体包含本公开文本的融合蛋白。在一些实施方案中,所述第一载体包含本公开文本的至少一种指导RNA。在一些实施方案中,所述第一载体包含本公开文本的一种或多种指导RNA。在一些实施方案中,所述第一载体包含本公开文本的两种或更多种指导RNA。在一些实施方案中,所述融合蛋白包含第一RNA结合蛋白和第二RNA结合蛋白。在一些实施方案中,所述第一载体和所述第二载体是相同的。在一些实施方案中,所述第一载体和所述第二载体是不同的。

[0314] 在本公开文本的组合物和方法的一些实施方案中,所述载体是或包含“双组分RNA靶向系统”的组分,所述双组分RNA靶向系统包含(a)编码本公开文本的靶向RNA的融合蛋白的核酸序列;以及(b)单一指导RNA(sgRNA)序列,其包含:在其5'端,与靶RNA序列杂交或特异性结合的RNA序列(例如,间隔子序列);和在其3'端,能够与所述融合蛋白的CRISPR/Cas蛋白特异性结合或缔合的RNA序列(例如,支架序列);并且其中所述双组分RNA靶向系统在PAMmer不存在的情况下识别并改变细胞中的所述靶RNA。在一些实施方案中,所述双组分系统的序列包含在单一(例如,一元)载体内。在一些实施方案中,所述双组分系统的间隔子序列靶向选自以下的重复序列:CUG、CCUG、CAG和GGGGCC。在一些实施方案中,所述双组分系统的间隔子序列靶向涉及适应性免疫应答的RNA序列。在一些实施方案中,所述双组分系统的间隔子序列包含编码适应性免疫应答的蛋白质组分的核酸序列的一部分,并且其中所述蛋白质组分选自 β -2-微球蛋白(β 2M)、人白细胞抗原A(HLA-A)、人白细胞抗原B(HLA-B)、人白细胞抗原C(HLA-C)、分化群28(CD28)、分化群80(CD80)、分化群86(CD86)、可诱导T细胞共刺激分子(ICOS)、ICOS配体(ICOSLG)、OX40L、白细胞介素12(IL12)和CC趋化因子受体7

(CCR7)。在一些实施方案中,所述双组分系统包含间隔子,其是编码适应性免疫应答的蛋白质组分的核酸序列的一部分,并且其长度是约20或21个核苷酸。在一些实施方案中,所述双组分系统包含第一和第二间隔子,其包含在单一的(singular) gRNA内。在一些实施方案中,所述双组分系统包含第一和第二间隔子序列,其包含在第一和第二gRNA序列内。在一些实施方案中,所述第一间隔子靶向重复序列,并且所述第二间隔子靶向涉及适应性免疫应答的RNA。

[0315] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的载体是病毒载体。在一些实施方案中,所述病毒载体包含从逆转录病毒分离或衍生的序列。在一些实施方案中,所述病毒载体包含从慢病毒分离或衍生的序列。在一些实施方案中,所述病毒载体包含从腺病毒分离或衍生的序列。在一些实施方案中,所述病毒载体包含从腺相关病毒(AAV)分离或衍生的序列。在一些实施方案中,所述病毒载体无复制能力。在一些实施方案中,所述病毒载体是分离的或重组的。在一些实施方案中,所述病毒载体是自身互补的。

[0316] 在本公开文本的组合物和方法的一些实施方案中,所述病毒载体包含从腺相关病毒(AAV)分离或衍生的序列。在一些实施方案中,所述病毒载体包含从血清型AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11或AAV12的AAV分离或衍生的反向末端重复序列或衣壳序列。在一些实施方案中,所述病毒载体无复制能力。在一些实施方案中,所述病毒载体是分离的或重组的(rAAV)。在一些实施方案中,所述病毒载体是自身互补的(scAAV)。

[0317] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的载体是非病毒载体。在一些实施方案中,所述载体包含以下项或由其组成:纳米颗粒、胶束、脂质体或阳离子脂质体/DNA复合物、聚合物囊泡、聚合物/DNA复合物或树枝状聚合物。在一些实施方案中,所述载体是表达载体或重组表达系统。如本文所用,术语“重组表达系统”是指用于表达通过重组形成的某些遗传物质的遗传构建体。

[0318] 在本公开文本的组合物和方法的一些实施方案中,本文提供的表达载体、病毒载体或非病毒载体包括但不限于表达控制元件。如本文所用的“表达控制元件”是指调节编码序列如基因的表的任何序列。示例性表达控制元件包括但不限于启动子、增强子、微小RNA、转录后调节元件、多聚腺苷酸化信号序列和内含子。例如,表达控制元件可以是组成型的、诱导型的、阻抑型的或组织特异性的。“启动子”是以下控制序列,其是多核苷酸序列中控制转录起始和速率的区域。它可以含有调节蛋白和分子可以结合的遗传元件,如RNA聚合酶和其他转录因子。在一些实施方案中,启动子对表的控制是组织特异性的。非限制性的示例性启动子包括CMV、CBA、CAG、Cbh、EF-1a、PGK、UBC、GUSB、UCOE、hAAT、TBG、结蛋白(Desmin)、MCK、C5-12、NSE、突触蛋白(Synapsin)、PDGF、MecP2、CaMKII、mGluR2、NFL、NFH、nβ2、PPE、ENK、EAAT2、GFAP、MBP和U6启动子。“增强子”是DNA中可以由激活蛋白结合以增加转录的可能性或频率的区域。非限制性的示例性增强子和转录后调节元件包括CMV增强子和WPRE。

[0319] 在本公开文本的组合物和方法的一些实施方案中,本文提供的表达载体、病毒载体或非病毒载体包括但不限于用于建构“多顺反子(multicistronic)”或“多顺反子(polycistronic)”或“双顺反子”或“三顺反子”构建体(即,具有双重或三重或多重编码区或外显子)的载体元件,如IRES或2A肽位点,并且因此将具有从mRNA表来自单一构建体的

两种或更多种蛋白质的能力。多顺反子载体从同一mRNA同时表达两种或更多种单独蛋白质。最广泛用于构建多顺反子构型的两种策略是通过使用IRES或2A自切割位点。“IRES”是指用于多顺反子载体构建体内的病毒、原核或真核来源的内部核糖体进入位点或其部分。在一些实施方案中，IRES是允许以非帽依赖性方式进行翻译起始的RNA元件。术语“自切割肽”或“编码自切割肽的序列”或“2A自切割位点”是指在载体构建体内用于并入位点以促进核糖体跳跃且因此从单一启动子产生两种多肽的连接序列，此类自切割肽包括但不限于T2A和P2A肽或编码自切割肽的序列。

[0320] 在一些实施方案中，所述载体是病毒载体。在一些实施方案中，所述载体是腺病毒载体、腺相关病毒 (AAV) 载体或慢病毒载体。在一些实施方案中，所述载体是逆转录病毒载体、腺病毒/逆转录病毒嵌合体载体、单纯疱疹病毒I或II载体、细小病毒载体、网状内皮组织增殖病病毒载体、脊髓灰质炎病毒载体、乳头状瘤病毒载体、痘苗病毒载体或者并入两种或更多种病毒载体的有利方面的任何杂合或嵌合载体。在一些实施方案中，所述载体还包含与多核苷酸可操作地连接的一种或多种表达控制元件。在一些实施方案中，所述载体还包含一种或多种选择标记。在一些实施方案中，所述AAV载体具有低毒性。在一些实施方案中，所述AAV载体不并入宿主基因组中，从而具有低的引起插入诱变的概率。在一些实施方案中，所述AAV载体可以编码4.5kb至4.75kb的一系列总多核苷酸。在一些实施方案中，可以用于任何本文所述的组合物、系统、方法和试剂盒中的示例性AAV载体可以包括AAV1载体、修饰的AAV1载体、AAV2载体、修饰的AAV2载体、AAV3载体、修饰的AAV3载体、AAV4载体、修饰的AAV4载体、AAV5载体、修饰的AAV5载体、AAV6载体、修饰的AAV6载体、AAV7载体、修饰的AAV7载体、AAV8载体、AAV9载体、AAV.rh10载体、修饰的AAV.rh10载体、AAV.rh32/33载体、修饰的AAV.rh32/33载体、AAV.rh43载体、修饰的AAV.rh43载体、AAV.rh64R1载体和修饰的AAV.rh64R1载体及其任何组合或等效物。在一些实施方案中，所述慢病毒载体是有整合酶能力的慢病毒载体 (integrase-competent lentiviral vector, ICLV)。在一些实施方案中，所述慢病毒载体可以是指转基因质粒载体，以及与相关质粒 (例如，包装质粒、rev表达质粒、包膜质粒) 结合的转基因质粒载体，以及能够通过病毒或病毒样进入机制将外源核酸引入细胞中的基于慢病毒的颗粒。慢病毒载体是本领域中熟知的 (参见例如，Trono D. (2002) *Lentiviral vectors*, New York:Spring-Verlag Berlin Heidelberg和Durand等人 (2011) *Viruses* 3 (2):132-159doi:10.3390/v3020132)。在一些实施方案中，可以用于任何本文所述的组合物、系统、方法和试剂盒中的示例性慢病毒载体可以包括人免疫缺陷病毒 (HIV) 1载体、修饰的人免疫缺陷病毒 (HIV) 1载体、人免疫缺陷病毒 (HIV) 2载体、修饰的人免疫缺陷病毒 (HIV) 2载体、白脸白眉猴 (sooty mangabey) 猴免疫缺陷病毒 (SIV_{SM}) 载体、修饰的白脸白眉猴猴免疫缺陷病毒 (SIV_{SM}) 载体、非洲绿猴猴免疫缺陷病毒 (SIV_{AGM}) 载体、修饰的非洲绿猴猴免疫缺陷病毒 (SIV_{AGM}) 载体、马传染性贫血病毒 (EIAV) 载体、修饰的马传染性贫血病毒 (EIAV) 载体、猫免疫缺陷病毒 (FIV) 载体、修饰的猫免疫缺陷病毒 (FIV) 载体、维斯纳/梅迪病毒 (Visna/maedi virus) (VNV/VMV) 载体、修饰的维斯纳/梅迪病毒 (VNV/VMV) 载体、羊关节炎-脑炎病毒 (CAEV) 载体、修饰的羊关节炎-脑炎病毒 (CAEV) 载体、牛免疫缺陷病毒 (BIV) 或修饰的牛免疫缺陷病毒 (BIV)。

核酸

[0321] 本文提供了编码用于本文所述的基因转移和表达技术中的本文公开的融合蛋白

的核酸序列。虽然没有总是明确陈述,但是应当理解,本文提供的序列可以用于提供表达产物以及产生具有相同生物特性的蛋白质的基本上相同的序列。这些“生物等效的”或“生物活性的”或“等效的”多肽是由如本文所述的等效多核苷酸编码的。当使用在默认条件下运行的序列同一性方法比较时,它们可以具有与参考多肽至少60%、或可替代地至少65%、或可替代地至少70%、或可替代地至少75%、或可替代地至少80%、或可替代地至少85%、或可替代地至少90%、或可替代地至少95%、或可替代地至少98%相同的初级氨基酸序列。提供了特定多肽序列作为具体实施方案的例子。用具有类似电荷的可替代氨基酸对序列中的氨基酸进行修饰。另外,等效多核苷酸是在严格条件下与参考多核苷酸或其互补体杂交的多核苷酸,或者就多肽而言,是由在严格条件下与参考编码多核苷酸或其互补链杂交的多核苷酸编码的多肽。可替代地,等效多肽或蛋白质是从等效多核苷酸表达的多肽或蛋白质。

[0322] 本文公开的核酸序列(例如,多核苷酸序列)可以是密码子优化的,密码子优化是本领域熟知的技术。在本文公开的一些实施方案中,示例性Cas序列(如例如,SEQ ID NO:46 (Cas13d))被密码子优化以用于在人细胞中表达。密码子优化涉及以下事实,即不同细胞在对特定密码子的使用方面有所不同。该密码子偏好对应于特定tRNA在细胞类型中的相对丰度的偏好。通过改变序列中的密码子以与相应tRNA的相对丰度匹配,可能增加表达。还可能通过故意选择已知相应tRNA在特定细胞类型中罕见的密码子来减少表达。哺乳动物细胞以及多种其他生物的密码子使用表是本领域中已知的。基于遗传密码,可以产生编码例如Cas蛋白的核酸序列。在一些实施方案中,这样的序列被优化以用于在宿主细胞或靶细胞中表达,所述宿主细胞或靶细胞是如用于表达Cas蛋白的宿主细胞或在其中实践所公开方法的细胞(如在哺乳动物细胞例如人细胞中)。特定物种的密码子偏好和密码子使用表可以用于工程化编码Cas蛋白的分离的核酸分子(如编码与其相应野生型蛋白具有至少80%、至少85%、至少90%、至少92%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列同一性的蛋白质的核酸分子),其利用该特定物种的密码子使用偏好。例如,本文公开的Cas蛋白可以被设计为具有特定目标生物优先使用的密码子。在一个例子中,Cas核酸序列被优化以用于在人细胞中表达,如与其相应野生型或起源核酸序列具有至少70%、至少80%、至少85%、至少90%、至少92%、至少95%、至少98%或至少99%序列同一性的Cas核酸序列。在一些实施方案中,编码至少一种Cas蛋白(其可能是载体的一部分)的分离的核酸分子包括被密码子优化以用于在真核细胞中表达的至少一个Cas蛋白编码序列或被密码子优化以用于在人细胞中表达的至少一个Cas蛋白编码序列。在一个实施方案中,这样的密码子优化的Cas编码序列与其相应野生型或起源序列具有至少80%、至少85%、至少90%、至少92%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列同一性。在另一个实施方案中,真核细胞密码子优化的核酸序列编码与其相应野生型或起源蛋白具有至少85%、至少90%、至少92%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列同一性的Cas蛋白。在另一个实施方案中,可以常规地产生含有功能等效核酸的多个克隆,所述功能等效核酸是如序列不同但编码相同Cas蛋白序列的核酸。编码序列中的沉默突变是由于遗传密码的简并性(即,冗余)所致,借此多于一种密码子可以编码相同氨基酸残基。因此,例如,亮氨酸可以由CTT、CTC、CTA、CTG、TTA或TTG编码;丝氨酸可以由TCT、TCC、TCA、TCG、AGT或AGC编码;天冬酰胺可以由AAT或AAC编码;天冬氨酸可以由GAT或GAC编码;半胱氨酸可以由TGT或TGC编码;丙氨酸可以由GCT、GCC、GCA或GCG编码;谷氨酰胺可以由CAA或

CAG编码;酪氨酸可以由TAT或TAC编码;并且异亮氨酸可以由ATT、ATC或ATA编码。显示标准遗传密码的表格可以在多个来源发现(参见例如,Stryer,1988,Biochemistry,第3版,W.H.5Freeman and Co.,NY)。

[0323] “杂交”是指一种或多种多核苷酸反应形成通过核苷酸残基的碱基之间的氢键合稳定化的复合物的反应。氢键合可以通过Watson-Crick碱基配对、Hoogsteen结合或以任何其他序列特异性方式来进行。所述复合物可以包含形成双链体结构的两条链、形成多链复合物的三条或更多条链、单条自杂交链或这些的任何组合。杂交反应可以构成更广泛过程(如PC反应的起始或核酶对多核苷酸的酶促切割)中的步骤。

[0324] 严格杂交条件的例子包括: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;以及约4x SSC至约8x SSC的洗涤溶液。中等杂交条件的例子包括: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;以及约5x SSC至约2x SSC的洗涤溶液。高严格性条件的例子包括: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;以及约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个、2个或更多个洗涤步骤,并且洗涤孵育时间为约1、2或15分钟。SSC是0.15M NaCl和15mM柠檬酸盐缓冲液。应理解,可以采用使用其他缓冲液系统的SSC的等效物。

“同源性”或“同一性”或“相似度”是指两个肽之间或两个核酸分子之间的序列相似度。同源性可以通过比较每个序列中的位置来确定,所述序列可以出于比较目的进行比对。在所比较序列中的位置由相同的碱基或氨基酸占据时,则所述分子在该位置是同源的。序列之间的同源性程度随着序列共有的匹配或同源位置的数量而变。“无关的”或“非同源的”序列与本发明的一个序列共有小于40%同一性、或可替代地小于25%同一性。

细胞

[0325] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的细胞是原核细胞。

[0326] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的细胞是真核细胞。在一些实施方案中,所述细胞是哺乳动物细胞。在一些实施方案中,所述细胞是牛、鼠、猫、马、猪、犬、猿猴或人细胞。在一些实施方案中,所述细胞是非人哺乳动物细胞,如非人灵长类动物细胞。

[0327] 在一些实施方案中,本公开文本的细胞是体细胞。在一些实施方案中,本公开文本的细胞是性细胞。在一些实施方案中,本公开文本的性细胞不是人细胞。

[0328] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的细胞是干细胞。在一些实施方案中,本公开文本的细胞是胚胎干细胞。在一些实施方案中,本公开文本的胚胎干细胞不是人细胞。在一些实施方案中,本公开文本的细胞是多潜能干细胞或多能干细胞。在一些实施方案中,本公开文本的细胞是成体干细胞。在一些实施方案中,本公开文本的细胞是诱导多能干细胞(iPSC)。在一些实施方案中,本公开文本的细胞是造血干细胞(HSC)。

[0329] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是免疫细胞。在一些实施方案中,本公开文本的免疫细胞是淋巴细胞。在一些实施方案中,本公开文

本的免疫细胞是T淋巴细胞(本文也称为T细胞)。本公开文本的示例性T细胞包括但不限于幼稚T细胞、效应T细胞、辅助T细胞、记忆T细胞、调节T细胞(Treg)和 $\gamma\delta$ T细胞。在一些实施方案中,本公开文本的免疫细胞是B淋巴细胞。在一些实施方案中,本公开文本的免疫细胞是自然杀伤细胞。在一些实施方案中,本公开文本的免疫细胞是抗原呈递细胞。

[0330] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是肌肉细胞。在一些实施方案中,本公开文本的肌肉细胞是成肌细胞或肌细胞。在一些实施方案中,本公开文本的肌肉细胞是心肌细胞、骨骼肌细胞或平滑肌细胞。在一些实施方案中,本公开文本的肌肉细胞是横纹肌细胞。

[0331] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是上皮细胞。在一些实施方案中,本公开文本的上皮细胞形成鳞状细胞上皮、立方细胞上皮、柱状细胞上皮、层状细胞上皮、假复层柱状细胞上皮或移行细胞上皮。在一些实施方案中,本公开文本的上皮细胞形成腺体,包括但不限于松果腺、胸腺、垂体、甲状腺、肾上腺、顶质分泌腺、全质分泌腺、局质分泌腺、浆液腺、粘液腺和皮脂腺。在一些实施方案中,本公开文本的上皮细胞接触器官的外表面,所述器官包括但不限于肺、脾、胃、胰腺、膀胱、肠、肾、胆囊、肝、喉或咽。在一些实施方案中,本公开文本的上皮细胞接触血管或静脉的外表面。

[0332] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是神经元细胞。在一些实施方案中,本公开文本的神经元细胞是中枢神经系统的神经元。在一些实施方案中,本公开文本的神经元细胞是脑或脊髓的神经元。在一些实施方案中,本公开文本的神经元细胞是视网膜的神经元。在一些实施方案中,本公开文本的神经元细胞是脑神经或视神经的神经元。在一些实施方案中,本公开文本的神经元细胞是周围神经系统的神经元。在一些实施方案中,本公开文本的神经元细胞是神经胶质细胞或胶质细胞。在一些实施方案中,本公开文本的胶质细胞是中枢神经系统的胶质细胞,包括但不限于少突胶质细胞、星形胶质细胞、室管膜细胞和小胶质细胞。在一些实施方案中,本公开文本的胶质细胞是周围神经系统的胶质细胞,包括但不限于施万细胞和卫星细胞。

[0333] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是原代细胞。

[0334] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是培养的细胞。

[0335] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是体内的、体外的、离体的或原位的。

[0336] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的体细胞是自体的或同种异体的。

掩蔽本公开文本的修饰的细胞

[0337] 本公开文本的组合物同时递送基因疗法并且防止从基因疗法构建体或相关递送载体衍生的抗原的表达,防止所述抗原在本公开文本的修饰的细胞的表面上展示。

[0338] 通过抑制或减少在所述修饰的细胞中适应性免疫应答的组分的表达,所述修饰的细胞对于宿主免疫系统不可见。例如,本公开文本的组合物可以同时靶向与遗传疾病或障碍相关的RNA分子和编码MHC I的 $\beta 2M$ 亚基的RNA分子。通过选择性靶向编码MHC I的 $\beta 2M$ 亚基的RNA分子,所述组合物防止所述修饰的细胞在所述修饰的细胞表面上展示一种或多种衍

生自RNA靶向构建体、载体或其组合的抗原肽。因此,受试者的免疫系统不会将所述修饰的细胞鉴定为含有外源序列,并且不会尝试针对所述修饰的细胞产生免疫应答。此方法提高了治疗所述遗传疾病或障碍的治疗功效,同时避免了基因疗法的常见副作用。

[0339] 在本公开文本的组合物和方法的一些实施方案中,适应性免疫应答的所述组分包含以下的组分或由其组成:I型主要组织相容性复合物(MHC I)、II型主要组织相容性复合物(MHC II)、T细胞受体(TCR)、共刺激分子或其组合。在一些实施方案中,所述MHC I组分包含 α 1链、 α 2链、 α 3链或 β 2M蛋白。在一些实施方案中,适应性免疫应答的所述组分包含MHC I β 2M蛋白或由其组成。在一些实施方案中,所述MHC II组分包含 α 1链、 α 2链、 β 1链或 β 2链。在一些实施方案中,所述TCR组分包含 α 链和 β 链。在一些实施方案中,所述共刺激分子包含分化群28(CD28)、分化群80(CD80)、分化群86(CD86)、可诱导T细胞共刺激分子(ICOS)或ICOS配体(ICOSLG)蛋白。

[0340] MHC I的 α 链可以由HLA基因(包括但不限于HLA-A、HLA-B和HLA-C)编码。

[0341] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码从HLA-A基因衍生的 α 链,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

```
1 atggccgtca tggcgccccg aaccctcgtc ctgctactct cgggggctct ggccctgacc
61 cagacctggg cgggctctca ctccatgagg tattttcttca catccgtgtc ccggccccggc
121 cgcggggagc cccgcttcat cgcagtgggc tacgtggacg acacgcagtt cgtgcggttc
181 gacagcgacg ccgcgagcca gaggatggag ccgcgggcgc cgtggataga gcaggaggggt
241 ccggagtatt gggacgggga gacacggaaa gtgaaggccc actcacagac tcaccgagtg
301 gacctgggga ccctgcgcgg ctactacaac cagagcgagg ccggttctca caccgtccag
361 aggatgtgtg gctgcgacgt ggggtcggac tggcgcttcc tccgcgggta ccaccagtac
421 gcctacgacg gcaaggatta catcgccctg aaagaggacc tgcgctcttg gaccgcggcg
481 gacatggcag ctacagaccac caagcacaag tgggaggcgg cccatgtggc ggagcagttg
541 agagcctacc tggagggcac gtgcgtggag tggctccgca gatacctgga gaacgggaag
601 gagacgctgc agcgcacgga cgccccaaa acgcatatga ctcaccacgc tgtctctgac
661 catgaagcca ccctgaggtg ctgggccctg agcttctacc ctgcggagat cacactgacc
721 tggcagcggg atggggagga ccagaccag gacacggagc tcgtggagac caggcctgca
781 ggggatggaa ccttcagaa gtgggcggct gtggtggtgc cttctggaca ggagcagaga
841 taaacctgcc atgtgcagca tgagggtttg cccaagcccc tcaccctgag atgggagccg
901 tcttccagc ccaccatccc catcgtgggc atcattgctg gcctggttct ctttgagct
961 gtgatcactg gagctgtggt cgctgctgtg atgtggagga ggaagagctc agatagaaaa
1021 ggagggagct actctcaggc tgcaagcagt gacagtgcc agggctctga tgtgtctctc
1081 acagcttgta aagtgtga (SEQ ID NO: 216)。
```

[0342] 本公开文本的组合可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码从HLA-B基因衍生的 α 链,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

```
1  tggtagtagga gaagaggat caggacgaag tcccaggccc cgggcggggc tctcagggtc
61  tcaggctccg agggccgcgt ctgcaatggg gaggcgcagc gttggggatt cccactccc
121 acgagtttca cttcttctcc caacctatgt cgggtccttc ttccaggata ctcgtagcgc
181 gtccccatth cccactccca ttgggtgtcg ggtgtctaga gaagccaatc agcgtagcgc
```

241 tgggtcccagt tctaaagtcc ccacgcaccc acccggaactc agaattctcct cagacgccga
301 gatgcggggtc acggcacccc gaaccgtcct cctgctgctc tggcgggccc tggccctgac
361 cgagacctgg gccggtgagt gcgggtcggc agggaaatgg cctctgtggg gaggagcgag
421 gggaccgcag gcgggggggc aggacccggg gagccgcgcc gggaggaggg tcgggggggt
481 ctacagccct cctcgcccc aggtccac tccatgaggt atttcacac cgccatgtcc
541 cggcccggcc gcgggggagc ccgcttcac accgtgggct acgtggacga cacgctgttc
601 gtgaggttcg acagcgacgc cagagtcgc aggaaggagc cgcggggccc atggatagag
661 caggaggggc cggagtattg ggaccgggag acacagatct ccaagaccaa cacacagact
721 taccgagaga gcctgcggaa cctgcgcggc tactacaacc agagcgaggc cggtagtgatga
781 ccccgggccc gggcgaggt cagactccc catccccac gtacggcccg ggtcgccccg
841 agtctccggg tccgagatcc gccccctga ggccgcggga cccgccaga cctcgaccg
901 gcgagagccc caggcgcgtt taccggttt cattttcagt tgaggccaaa atccccgcg
961 gttggtcggg gcggggcggg gcggggctcg ggggacggg ctgaccgcg ggccctgggc
1021 aggggtctcac acttggcaga ggatgtatgg ctgcgacctg gggcccgacg ggcgcctcct
1081 ccgcgggtat aaccagttag cctacgacgg caaggattac atcgccctga acgaggacct
1141 gagctcctgg accgcggcg acaccgcggc tcagatcacc cagcgcaagt gggaggcggc
1201 ccgtgtggcg gagcaggaca gagcctacct ggagggcctg tgcgtggagt cgctccgag
1261 atacctggag aacgggaagg agacgtgca gcgcgcgggt accaggggca gtggggagcc
1321 ttccccatct cctataggtc gccggggatg gcctcccacg agaagaggag gaaaatggga
1381 tcagcgctag aatgtcgccc tccctgaat ggagaatggc atgagtttcc ctgagtttcc
1441 tctgagggcc cctcttctc tctaggaaa taaggaatga cgtctctgag gaaatggagg
1501 ggaagacagt cctagaata ctgatcagg gtcccccttg acccctgcag cagccttggg
1561 aaccgtgact ttctctcag gccttgtct ctgcctcaca ctcaagtgtg ttggggctct
1621 gattccagca cttctgagtc actttacct cactcagatc gggagcagaa gtccctgttc
1681 cccgctcaga gactcgaact ttccaatgaa taggagatta tccaggtgc ctgcgtccag
1741 gctggtgtct gggttctgtg ccccttccc acccaggtg tctgtccat tctcaggctg
1801 gtcacatggg tggctcctagg gtgtcccatg agagatgcaa agcgctgaa tttctgact
1861 ctccccatca gacccccaa agacacatgt gaccaccac cccatctctg accatgaggc
1921 caccctgagg tgctgggccc tgggcttcta cctgcggag atcacactga cctggcagcg
1981 ggatggcgag gaccaaactc aggacaccga gcttgtggag accagaccag caggagatag
2041 aaccttcag aagtgggcag ctgtgtgtgt gccttctgga gaagagcaga gataacatg

2101 ccatgtacag catgaggggc tgccgaagcc cctcaccctg agatggggta aggaggggga
2161 tgaggggtca tatctgttct cagggaaagc aggagccctt ctggagccct tcagcagggg
2221 cagggccctt catcttcccc tcctttccca gagccatctt cccagtccac catccccatc
2281 gtgggcattg ttgctggcct ggctgtccta gcagttgtgg tcatcggagc tgtggctcgt
2341 actgtgatgt gtaggaggaa gagctcaggt agggaagggg tgaggggtgg ggtctgggtt
2401 ttcttgtccc actgggggtt tcaagcccca ggtagaagtg ttccctgcct cattactggg
2461 aagcagcatc cacacagggg ctaacgcagc ctgggaccct gtgtgccagc acttactctt
2521 ttgtgcagca catgtgacaa tgaaggacgg atgtatcgcc ttgatgggtg tgggtgtggg
2581 gtccatgattc cagcattcat gagtcagggg aaggccctg ctaaggacag accttaggag
2641 ggcagttggg ccaggaccca cacttgcttt cctcgtgttt cctgatcctg ccttgggtct
2701 gtagtcatac ttctggaaat tccttttggg tccaagacga ggaggttctt ctaagatctc
2761 atggccctgc ttccctccag tcccctcaca ggacattttc ttccacagg tggaaaagga
2821 gggagctact ctcaggctgc gtgtaagtgg tgggggtggg agtgtggagg agctcaccca
2881 cccataatt cctcctgtcc cacgtctcct gagggctctg accaggctct gttttgttc
2941 tactccagcc agcgacagtg cccagggtc tgatgtgtct ctcacagctt gaaaagggtga
3001 gattcttggg gtctagagtg ggtgggggtg cgggtctggg ggtgggtggg gcagtgggga
3061 aaggcctggg taatggagat tctttgattg ggatgtttcg cgtgtgtggg gggctgttca
3121 gagtgtcatc acttaccatg actaaccaga atttgttcat gactgttgtt ttctgtagcc
3181 tgagacagct gtcttgtgag ggactgagat gcaggatttc ttcacgcctc cccttgtga
3241 cttcaagagc ctctggcatc tctttctgca aaggcacctg aatgtgtctg cgtccctgtt
3301 agcataatgt gaggaggtgg agagacagcc cacccttgtg tccactgtga cccctgttcg
3361 catgctgacc tgtgtttcct cccca (SEQ ID NO: 217)。

[0343] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码从HLA-C基因衍生的 α 链,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 tccgcagtcc cggttctaaa gtccccagtc acccaccgg actcacattc tccccagagg
61 ccgagatgcg ggtcatggcg ccccgagccc tctcctgct gctctcgga ggctggccc
121 tgaccgagac ctgggcctgc tcccactcca tgaggatatt cgacaccgcc gtgtcccggc
181 ccggccgcgg agagccccgc ttcattctcag tgggctacgt ggacgacacg cagttcgtgc
241 ggttcgacag cgacgccgcg agtccgagag gggagccgcg ggcgccgtgg gtggagcagg
301 aggggcccga gtattgggac cgggagacac agaagtacaa gcgccaggca caggctgacc

361 gagtgagcct gcggaacctg cgcggctact acaaccagag cgaggacggg tctcacaccc
421 tccagaggat gtctggctgc gacctggggc ccgacgggcg cctcctccgc gggatgacc
481 agtccgccta cgacggcaag gattacatcg cctgaacga ggacctgcgc tcttgaccg
541 ccgcggaacac cgcggctcag atcaccacgc gcaagttgga ggcgcccggt gcggcggagc
601 agctgagagc ctacctggag ggcacgtgcg tggagtggct ccgcagatac ctggagaacg
661 ggaaggagac gctgcagcgc gcagaacccc caaagacaca cgtgaccac caccctct
721 ctgaccatga ggccaccctg aggtgctggg cctgggctt ctaccctgcg gagatcacac
781 tgacctggca gcgggatggg gaggaccaga ccaggacac cgagcttggt gagaccaggc
841 cagcaggaga tggaaccttc cagaagtggg cagctgtggt ggtgccttct ggacaagagc
901 agagatacac gtgccatatg cagcacgagg ggctgcaaga gccctcacc ctgagctggg
961 agccatcttc ccagcccacc atcccatca tgggcatcgt tgctggcctg gctgtcctgg
1021 ttgtcctagc tgtccttgga gctgtggtca ccgctatgat gtgtaggagg aagagctcag
1081 gtggaaaagg agggagctgc tctcaggctg cgtgcagcaa cagtgccag ggctctgatg
1141 agtctctcat cacttgtaaa gcctgagaca gctgcctgtg tgggactgag atgcaggatt
1201 tcttcacacc tctcctttgt gacttcaaga gcctctggca tctctttctg caaaggcacc
1261 tgaatgtgtc tgcgttctg ttagcataat gtgaggaggt ggagagacag cccacccccg
1321 tgtccaccgt gaccctgtc ccacactga cctgtgttcc ctcccgatc atctttctg
1381 ttccagagag gtggggctgg atgtctccat ctctgtctca aattcatggt gcactgagct
1441 gcaacttctt acttcctaa tgaagttaag aacctgaata taaatttggt ttctcaaata
1501 tttgctatga agcgttgatg gattaattaa ataagtcaat tctagaagt tgagagagca
1561 aataaagacc tgagaacctt ccagaa (SEQ ID NO: 218)。

[0344] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码从HLA-C基因衍生的 α 链,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 tccgcagtcc cggttctaaa gtccccagtc acccaccgg actcacattc tcccagagg
61 ccgagatgcg ggtcatggcg ccccgagccc tctcctgct gctctcgga ggctggccc
121 tgaccgagac ctgggcctgc tccactcca tgaggatatt cgacaccgcc gtgtcccgcc
181 ccggccgcgg agagccccgc ttcattctcag tgggctacgt ggacgacacg cagttcgtgc
241 ggttcgacag cgacgcgcg agtccgagag gggagccgcg ggcccggtgg gtggagcagg
301 aggggcccga gtattgggac cgggagacac agaactacaa gcgccaggca caggctgacc
361 gagtgagcct gcggaacctg cgcggctact acaaccagag cgaggacggg tctcacaccc

421 tccagaggat gtatggctgc gacctggggc ccgacgggcg cctcctccgc gggatatgacc
481 agtccgccta cgacggcaag gattacatcg ccctgaacga ggacctgcmc tcttgaccg
541 ccgcgacac cgcggtcag atcaccacgc gcaagttgga ggcgcccggt gcggcgagc
601 agctgagagc ctacctggag ggcacgtgcg tggagtggct ccgcagatac ctggagaacg
661 ggaaggagac gctgcagcgc gcagaacccc caaagacaca cgtgaccac caccctct
721 ctgaccatga ggccaccctg aggtgctggg ccctgggctt ctaccctgcg gagatcacac
781 tgacctggca gcgggatggg gaggaccaga ccaggacac cgagcttggt gagaccaggc
841 cagcaggaga tggaaccttc cagaagtggg cagctgtggt ggtgccttct ggacaagagc
901 agagatacac gtgccatatg cagcacgagg ggctgcaaga gcccctcacc ctgagctggg
961 agccatcttc ccagcccacc atcccatca tgggcatcgt tgctggcctg gctgtcctgg
1021 ttgtcctagc tgtccttgga gctgtggtca ccgctatgat gtgtaggagg aagagctcag
1081 gtggaaaagg agggagctgc tctcaggctg cgtgcagcaa cagtgccag ggctctgatg
1141 agtctctcat cacttgtaaa gcctgagaca gctgcctgtg tgggactgag atgcaggatt
1201 tcttcacacc tctcctttgt gacttcaaga gcctctggca tctctttctg caaaggcgtc
1261 tgaatgtgtc tgcgttctg ttagcataat gtgaggaggt ggagagacag cccaccccg
1321 tgtccaccgt gaccctgtc ccacactga cctgtgttcc ctcccgatc atctttctg
1381 ttccagagag gtggggctgg atgtctccat ctctgtctca aattcatggt gactgagct
1441 gcaacttctt acttcctaa tgaagttaag aacctgaata taaatttggt ttctcaaata
1501 tttgctatga agcgttgatg gattaattaa ataagtcaat tctagaagt tgagagagca
1561 aataaagacc tgagaacctt ccagaa (SEQ ID NO: 219)。

[0345] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码B2M蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 attcctgaag ctgacagcat tcgggcgag atgtctcgt ccgtggcctt agctgtgtc
61 gcgctactct ctctttctg cctggaggct atccagcgt ctccaaagat tcaggtttac
121 tcacgtcatc cagcagagaa tggaaagtca aatttcctga attgctatgt gtctgggtt
181 catccatccg acattgaagt tgacttactg aagaatggag agagaattga aaaagtggag
241 cattcagact tgtctttcag caaggactgg tctttctatc tcttgacta cactgaattc
301 acccccactg aaaaagatga gtatgcctgc cgtgtgaacc atgtgacttt gtcacagccc
361 aagatagtta agtgggatcg agacatgtaa gcagcatcat ggaggtttga agatgccga
421 tttggattgg atgaattcca aattctgctt gcttgctttt taatattgat atgcttatac

481 acttacactt tatgcacaaa atgtagggtt ataataatgt taacatggac atgatcttct
541 ttataattct actttgagtg ctgtctccat gtttgatgta tctgagcagg ttgtccaca
601 ggtagctcta ggagggctgg caacttagag gtggggagca gagaattctc ttatccaaca
661 tcaacatctt ggtcagattt gaactcttca atctcttgca ctcaaagctt gttaagatag
721 ttaagcgtgc ataagttaac ttccaattta catactctgc ttagaatttg ggggaaaatt
781 tagaaatata attgacagga ttattggaaa tttgttataa tgaatgaaac attttgtcat
841 ataagattca tatttacttc ttatacattt gataaagtaa ggcatgggtg tggttaatct
901 ggtttatttt tgttccacaa gttaaataaa tcataaaaact tgatgtgtta tctcttatat
961 ctactccca ctattacccc tttattttca aacagggaaa cagtcttcaa gttccacttg
1021 gtaaaaaatg tgaaccctt gtatatagag tttggctcac agtgtaaagg gcctcagtga
1081 ttcacatttt ccagattagg aatctgatgc tcaaagaagt taaatggcat agttggggtg
1141 acacagctgt ctagtgggag gccagccttc tatattttag ccagcgttct ttctgcggg
1201 ccaggtcatg aggagtatgc agactctaag agggagcaaa agtatctgaa ggatttaata
1261 ttttagcaag gaatagatat acaatcatcc ctgggtctcc ctgggggatt ggtttcagga
1321 ccccttcttg gacaccaa atctatggat ttaagtcct tctataaaat ggtatagtat
1381 ttgcatataa cctatccaca tcctcctgta tactttaaat catttctaga ttacttgtaa
1441 tacctaatac aatgtaaatg ctatgcaa atgtgttatt gttaaggaa taatgacaag
1501 aaaaaaaagt ctgtacatgc tcagtaaaga cacaaccatc ccttttttcc ccagtgttt
1561 ttgatccatg gtttgctgaa tccacagatg tggagcccct ggatacggaa ggcccgtgt
1621 actttgaatg acaaataaca gatttaaaat tttcaaggca tagttttata cctga (SEQ
ID NO: 220)。

[0346] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD28蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 taaagtcac aaaacaacgt tatatcctgt gtgaaatgct gcagtcagga tgccttggtg
61 tttgagtgcc ttgatcatgt gccctaagg gatggtggcg gtggtggtg ccgtggatga
121 cggagaactc caggccttgg caggtgcgtc tttcagttcc cctcacactt cgggttctc
181 ggggaggagg ggctggaacc ctagcccatc gtcaggacaa agatgctcag gctgctcttg
241 gctctcaact tattcccttc aattcaagta acaggaaaca agattttggt gaagcagtcg
301 cccatgcttg tagcgtacga caatgcggtc aaccttagct gcaagtattc ctacaatctc
361 ttctcaaggg agttccgggc atcccttcac aaaggactgg atagtgtgtt ggaagtctgt

421 gttgtatatg ggaattactc ccagcagctt cagggtttact caaaaacggg gttcaactgt
481 gatgggaaat tgggcaatga atcagtgaca ttctacctcc agaatttgta tgttaaccaa
541 acagatattt acttctgcaa aattgaagtt atgtatcctc ctctttacct agacaatgag
601 aagagcaatg gaaccattat ccatgtgaaa gggaaacacc tttgtccaag tccctatatt
661 cccggacctt ctaagccctt ttgggtgctg gtgggtggtg gtggagtcct ggcttgctat
721 agcttgctag taacagtggc ctttattatt ttctgggtga ggagtaagag gagcaggctc
781 ctgcacagtg actacatgaa catgactccc cgccgccccg ggcccacccg caagcattac
841 cagccctatg ccccaccacg cgacttcgca gcctatcgct cctgacacgg acgcctatcc
901 agaagccagc cggctggcag ccccatctg ctcaatatca ctgctctgga taggaaatga
961 ccgccatctc cagccggcca cctcaggccc ctggtgggcc accaatgcca atttttctcg
1021 agtgactaga ccaaatatca agatcatttt gagactctga aatgaagtaa aagagatttc
1081 ctgtgacagg ccaagtctta cagtgccatg gccacattc caactacca tgtacttagt
1141 gacttgactg agaagttagg gtagaaaaca aaaagggagt ggattctggg agcctcttcc
1201 ctttctcact cacctgcaca tctcagtcaa gcaaagtgtg gtatccacag acattttagt
1261 tgcagaagaa aggctaggaa atcattcctt ttggttaa at ggggtgttaa tcttttggtt
1321 agtgggttaa acggggtaag ttagagtagg gggagggata ggaagacata tttaaaaacc
1381 attaaaaacac tgtctccacac tcatgaaatg agccacgtag ttctatttta atgctgtttt
1441 cttttagttt agaaatacat agacattgtc ttttatgaat totgatcata tttagtcatt
1501 ttgaccaa at gagggatttg gtcaaatgag ggattccctc aaagcaatat caggtaaacc
1561 aagttgcttt cctcaactccc tgtcatgaga cttcagtgtt aatgttcaca atatactttc
1621 gaaagaataa aatagttctc ctacatgaag aaagaatatg tcaggaaata aggtcacttt
1681 atgtcaaa at tatttgagta ctatgggacc tggcgcagtg gotcatgctt gtaatccag
1741 cactttggga ggccgaggtg ggcagatcac ttgagatcag gaccagcctg gtcaagatgg
1801 tgaaactccg tctgtactaa aaatacaaaa tttagcttgg cctggtggca ggcacctgta
1861 atcccagctg cccaagaggc tgaggcatga gaatcgcttg aacctggcag gcgagggttg
1921 cagtgagccg agatagtgc acagctctcc agcctgggag acagagttag actocatctc
1981 aaacaacaac aacaacaaca acaacaaca caaaccacaa aattatttga gtactgtgaa
2041 ggattatttg tctaacagtt cattccaatc agaccaggta ggagctttcc tgtttcatat
2101 gtttcagggg tgacacagtt gtctctttta tgtcggtgtg gagatccaaa gtgggtgtg
2161 gaaagagcgt ccataggaga agtgagaata ctgtgaaaaa gggatgttag cattcattag
2221 agtatgagga tgagtcccaa gaaggttctt tggaaggagg acgaatagaa tggagtaatg

2281 aaattcttgc catgtgctga ggagatagcc agcattaggt gacaatcttc cagaagtggg
2341 caggcagaag gtgccctggg gagagctcct ttacagggac tttatgtggg ttagggctca
2401 gagctccaaa actctgggct cagctgctcc tgtaccttgg aggtccattc acatgggaaa
2461 gtattttggg atgtgtcttt tgaagagagc atcagagttc ttaagggact gggtaaggcc
2521 tgaccctgaa atgaccatgg atatTTTTct acctacagtt tgagtcaact agaatatgcc
2581 tggggacctt gaagaatggc ccttcagtgg ccctcaccat ttgttcatgc ttcagttaat
2641 tcagggtgttg aaggagctta ggttttagag gcacgtagac ttgggttcaag tctcggttagt
2701 agttgaatag cctcaggcaa gtcaactgcc acctaatg atggttcttc aactataaaa
2761 tggagataat ggttaciaat gtctcttctt atagtataat ctccataagg gcatggccca
2821 agtctgtctt tgactctgcc tatccctgac atttagtagc atgcccgaca tacaatgtta
2881 gctattggta ttattgccat atagataaat tatgtataaa aattaaactg ggcaatagcc
2941 taagaagggg ggaatattgt aacacaaatt taaaccact acgcagggat gaggtgctat
3001 aatatgagga ccttttaact tccatcatTT tctgttttct tgaaatagtt tatcttgtaa
3061 tgaaatataa ggcacctccc acttttatgt atagaaagag gtcttttaat ttttttttaa
3121 tgtgagaagg aaggaggag taggaatctt gagattccag atcgaaaata ctgtactttg
3181 gttgattttt aagtgggctt ccattccatg gatttaatca gtcccaagaa gatcaaactc
3241 agcagtactt gggtgctgaa gaactgttgg atttaccctg gcacgtgtgc cacttgccag
3301 cttcttgggc acacagagtt cttcaatcca agttatcaga ttgtatttga aaatgacaga
3361 gctggagagt tttttgaaat ggcagtggca aataaataaa tacttttttt taaatggaaa
3421 gaactgatct atggtaataa atgattttgt tttctgactg gaaaaatagg cctactaaag
3481 atgaatcaca cttgagatgt ttcttactca ctctgcacag aaacaaagaa gaaatgttat
3541 acagggaagt ccgtttttcac tattagtatg aaccaagaaa tggttcaaaa acagtggtag
3601 gagcaatgct ttcatagttt cagatatggg agttatgaag aaaacaatgt catttgctgc
3661 tattattgta agagtcttat aattaatggg actcctataa tttttgattg tgagctcacc
3721 tatttggggt aagcatgcca atttaaagag accaagtgtg tgtacattat gttctacata
3781 ttcagtgata aaattactaa actactatat gtctgcttta aatttgact ttaatattgt
3841 cttttgggtat taagaaagat atgctttcag aatagatatg cttcgctttg gcaaggaatt
3901 tggatagaac ttgctattta aaagagggtg ggggtaaatc cttgtataaa tctocagttt
3961 agcctttttt gaaaaagcta gactttcaaa tactaatttc acttcaagca gggtaagttt
4021 ctgggttgggt tgcttgactt cagtcacaat ttcttatcag accaatggct gacctctttg
4081 agatgtcagg ctaggcttac ctatgtgttc tgtgtcatgt gaatgctgag aagtttgaca

4141 gagatccaac ttcagccttg accccatcag tccctcgggt taactaactg agccacogggt
4201 cctcatgggt attttaatga ggggtattgat ggttaaagtc atgtctgata cttatccca
4261 gccatttgca ctgccagctg ggaactatac cagacctgga tactgatccc aaagtgttaa
4321 attcaactac atgctggaga ttagagatgg tgccaataaa ggaccagaa ccaggatctt
4381 gattgctata gacttattaa taatccagggt caaagagagt gacacacact ctctcaagac
4441 ctgggggtgag ggagtctgtg ttatctgcaa ggccatttga ggctcagaaa gtctctcttt
4501 cctatagata tatgcatact ttctgacata taggaatgta tcaggaatac tcaaccatca
4561 caggcatggt cctacctcag ggccctttaca tgtcctgttt actctgtcta gaatgtcctt
4621 ctgtagatga cctggcttgc ctgctcacc ctcaggctct tgcctcaagt tcatctcttc
4681 ccctagttaa actaccccac accctgtctg ctttccttgc ttatttttct ccatagcatt
4741 ttaccatctc ttacattaga ctttttctt atttatttgt agtttataag cttcatgagg
4801 caagtaactt tgctttgttt cttgctgtat ctccagtgc cagagcagt cctggtatat
4861 aataaatatt tattgactga gtgaaaaaaaa aaaaaaaaaa (SEQ ID NO: 221)。

[0347] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD28蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 taaagtcatac aaaacaacgt tatatcctgt gtgaaatgct gcagtcagga tgccttgtgg
61 tttgagtgcc ttgatcatgt gccctaaggg gatggtggcg gtggtggtgg ccgtggatga
121 cggagactct caggccttgg cagggtgcgtc ttccagttcc cctcacactt cgggttcctc
181 ggggaggagg ggctggaacc ctagcccatc gtcaggacaa agatgctcag gctgctcttg
241 gctctcaact tattcccttc aattcaagta acaggaaaca agatttttgt gaagcagtcg
301 cccatgcttg tagcgtacga caatgcggtc aaccttagct ggaaacacct ttgtccaagt
361 ccctatttc ccggaccttc taagcccttt tgggtgctgg tgggtggttg tggagtcttg
421 gcttgctata gcttgctagt aacagtggcc ttattattt tctgggtgag gagtaagagg
481 agcaggctcc tgcacagtga ctacatgaac atgactcccc gccgccccgg gccacccgc
541 aagcattacc agccctatgc cccaccagc gacttcgcag cctatcgctc ctgacacgga
601 cgcctatcca gaagccagcc ggctggcagc ccccatctgc tcaatatcac tgctctggat
661 aggaaatgac cgccatctcc agccggccac ctcaggcccc tgttgggcca ccaatgccaa
721 tttttctcga gtgactagac caaatatcaa gatcattttg agactctgaa atgaagtaaa
781 agagatttcc tgtgacaggc caagtcttac agtgccatgg ccacattcc aacttaccat
841 gtacttagtg acttgactga gaagttaggg tagaaaacaa aaagggagtg gattctggga

901 gcctcttccc ttctcactc acctgcacat ctcagtcaag caaagtgtgg tatccacaga
961 catttttagtt gcagaagaaa ggctaggaaa tcattccttt tggttaaatg ggtgtttaat
1021 ctttttggtta gtgggttaaa cggggtaagt tagagtaggg ggagggatag gaagacatat
1081 ttaaaaacca ttaaaacact gtctccact catgaaatga gccacgtagt tcctatttaa
1141 tgctgttttc ctttagttta gaaatacata gacattgtct tttatgaatt ctgatcatat
1201 ttagtcattt tgaccaaag agggatttgg tcaaagagg gattccctca aagcaatato
1261 aggtaaacca agttgcttcc ctactccct gtcagtggac ttcagtgtta atgttcacaa
1321 tatactttcg aaagaataaa atagttctcc tacatgaaga aagaatatgt caggaaataa
1381 ggtcacttta tgtcaaaatt atttgagtac tatgggacct ggcgagtggt ctcatgcttg
1441 taatcccagc actttgggag gccgaggtgg gcagatcact tgagatcagg accagcctgg
1501 tcaagatggt gaaactccgt ctgtactaaa aatacaaaat ttagcttggc ctggtggcag
1561 gcacctgtaa tcccagctgc ccaagaggct gaggcagtag aatcgcttga acctggcagg
1621 cggaggttgc agtgagccga gatagtcca cagctctcca gcctgggcca cagagtga
1681 ctccatctca aacaacaaca acaacaaca caacaacaac aaaccacaaa attatttgag
1741 tactgtgaag gattatttgg ctaacagttc attccaatca gaccaggtag gagctttcct
1801 gtttcatatg ttccagggtt gcacagttgg tctctttaat gtcggtgtgg agatccaaag
1861 tgggttgtgg aaagagcgtc cataggagaa gtgagaatac tgtgaaaaag ggatgttagc
1921 attcattaga gtatgaggat gagtcccaag aaggttcttt ggaaggagga cgaatagaat
1981 ggagtaatga aattcttgcc atgtgctgag gagatagcca gcattaggtg acaatcttcc
2041 agaagtggtc aggcagaagg tgccctggtg agagctcctt tacagggact ttatgtgggt
2101 tagggctcag agctccaaaa ctctgggctc agctgctcct gtaccttga ggtccattca
2161 catgggaaag tattttggaa tgtgtctttt gaagagagca tcagagttct taagggactg
2221 ggtaaggcct gacctgaaa tgaccatgga ttttttcta cctacagttt gagtcaacta
2281 gaatatgcct ggggaccttg aagaatggcc cttcagtggt cctcaccatt tgttcatgct
2341 tcagttaatt caggtgttga aggagcttag gtttttagagg cacgtagact tggttcaagt
2401 ctcgtttagta gttgaatagc ctcaggcaag tcaactgcca cctaagatga tggttcttca
2461 actataaaat ggagataatg gttacaaatg tctcttccta tagtataatc tccataaggg
2521 catggcccaa gtctgtcttt gactctgcct atccctgaca tttagtagca tgcccgacat
2581 acaatgttag ctattggtat tattgccata tagataaatt atgtataaaa attaaactgg
2641 gcaatagcct aagaaggggg gaattattgta acacaaattt aaaccacta cgcagggatg
2701 aggtgctata atatgaggac cttttaactt ccatcatttt cctgtttctt gaaatagttt

2761 atcttgtaat gaaatataag gcacctccca cttttatgta tagaaagagg tcttttaatt
2821 tttttttaat gtgagaagga agggaggagt aggaatcttg agattccaga tcgaaaatac
2881 tgtacttttg ttgattttta agtgggcttc cattccatgg atttaatcag tcccaagaag
2941 atcaaaactca gcagtacttg ggtgctgaag aactgttga tttaccctgg cacgtgtgcc
3001 acttgccagc ttcttgggca cacagagttc ttcaatccaa gttatcagat tgtatttgaa
3061 aatgacagag ctggagagtt ttttgaaatg gcagtggcaa ataaataaat actttttttt
3121 aaatggaaag acttgatcta tggtaataaa tgattttgtt ttctgactgg aaaaataggc
3181 ctactaaaga tgaatcacac ttgagatggt tcttactcac tctgcacaga aacaaagaag
3241 aaatgttata cagggaggtc cgttttcaact attagtatga accaagaaat ggttcaaaaa
3301 cagtggtagg agcaatgctt tcatagtctc agatatggta gttatgaaga aaacaatgtc
3361 atttgctgct attattgtaa gagtcttata attaatggta ctctataat ttttgattgt
3421 gagctcacct atttgggtta agcatgccaa tttaaagaga ccaagtgtat gtacattatg
3481 ttctacatat tcagtataa aattactaaa ctactatatg totgctttaa atttgtactt
3541 taatattgtc ttttggtatt aagaaagata tgctttcaga atagatatgc ttcgctttgg
3601 caaggaattt ggatagaact tgctatttaa aagaggtgtg gggtaaactc ttgtataaat
3661 ctccagttta gccttttttg aaaaagctag actttcaaact actaatttca cttcaagcag
3721 ggtacgtttc tggtttggtt gcttgacttc agtcacaatt tottatcaga ccaatggctg
3781 acctctttga gatgtcaggc taggcttacc tatgtgttct gtgtcatgtg aatgtgaga
3841 agtttgacag agatccaact tcagccttga ccccatcagt cctcggggtt aactaactga
3901 gccaccggtc ctcatggcta ttttaatgag ggtattgatg gttaaatgca tgtctgatcc
3961 cttatcccag ccatttgcac tgccagctgg gaactatacc agacctggat actgatccca
4021 aagtgttaaa ttcaactaca tgctggagat tagagatggt gccataaag gaccagaac
4081 caggatcttg attgctatag acttattaat aatccaggtc aaagagagtg acacacactc
4141 totcaagacc tggggtgagg gagtctgtgt tatctgcaag gccatttgag gctcagaaag
4201 tctctctttc ctatagatat atgcatactt tctgacatat aggaatgtat caggaatact
4261 caaccatcac aggcattgtc ctacctcagg gcctttacat gtctgttta ctctgtctag
4321 aatgtccttc ttagatgac ctggcttgcc tcgtcacctc tcaggctcct gctcaagtgt
4381 catcttctcc cctagttaaa ctacccaca cctgtctgc tttccttgct tatttttctc
4441 catagcattt taccatctct tacattagac atttttctta tttatttgta gtttataagc
4501 ttcattgaggc aagtaacttt gctttgttct ttgctgtatc tccagtgcc agagcagtgc
4561 ctggtatata ataaatattt attgactgag tgaaaaaaaa aaaaaaaaa (SEQ ID NO:
222)。

[0348] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列

的gRNA,所述RNA分子编码CD28蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

```
1 taaagtcatc aaaacaacgt tatatcctgt gtgaaatgct gcagtcagga tgccttgtgg
61 tttgagtgcc ttgatcatgt gccctaaggg gatggtggcg gtggtggtgg ccgaggatga
121 cggagactct caggccttgg cagggtcgtc tttcagttcc cctcacactt cgggttccctc
181 ggggaggagg ggctggaacc ctagcccatc gtcaggacaa agatgctcag gctgctcttg
241 gctctcaact tattcccttc aattcaagta acagggaaac acctttgtcc aagtccccta
301 tttcccgga cttctaagcc cttttgggtg ctggtggtgg ttggtggagt cctggcttgc
361 tatagcttgc tagtaacagt ggcctttatt attttctggg tgaggagtaa gaggagcagg
421 ctctgcaca gtgactacat gaacatgact ccccgccgcc ccgggcccac ccgcaagcat
481 taccagccct atgccccacc acgcgacttc gcagcctatc gctcctgaca cggacgccta
541 tccagaagcc agccggctgg cagcccccat ctgctcaata tcaactgctc ggataggaaa
601 tgaccgccat ctccagccgg ccacctcagg cccctgttgg gccaccaatg ccaatttttc
661 tcgagtgact agaccaaata tcaagatcat tttgagactc tgaaatgaag taaaagagat
721 ttctgtgac aggccaagtc ttacagtgcc atggcccaca ttccaactta ccatgtactt
781 agtgacttga ctgagaagtt agggtagaaa acaaaaaggg agtggattct gggagcctct
841 tccctttctc actcacctgc acatctcagt caagcaaagt gtggtatcca cagacathtt
901 agttgcagaa gaaaggctag gaaatcattc cttttggtta aatgggtgtt taatcttttg
961 gttagtgggt taaacggggt aagttagagt agggggaggg ataggaagac atatttaaaa
1021 accattaaaa cactgtctcc cactcatgaa atgagccacg tagttcctat ttaatgctgt
1081 tttccttttag tttagaaata catagacatt gtcttttatg aattctgac atatttagtc
1141 attttgacca aatgagggat ttggtcaa atgaggattcc ctcaaagcaa tatcaggtaa
1201 accaagttgc tttcctcact cctgtcatg agacttcagt gttaatgttc acaatatact
1261 ttcgaaagaa taaaatagtt ctctacatg aagaaagaat atgtcaggaa ataaggtcac
1321 tttatgtcaa aattatttga gtactatggg acctggcgca gtggctcatg cttgtaatcc
1381 cagcactttg ggaggccgag gtgggcagat cacttgagat caggaccagc ctggtcaaga
1441 tggtgaaact ccgtctgtac taaaaataca aaatttagct tggcctggtg gcaggcacct
1501 gtaatccag ctgcccaaga ggctgaggca tgagaatcgc ttgaacctgg caggcggagg
1561 ttgcagtgag ccgagatagt gccacagctc tccagcctgg gcgacagagt gagactccat
```

1621 ctcaaacaac aacaacaaca acaacaacaa caacaaacca caaaattatt tgagtactgt
1681 gaaggattat ttgtotaaca gttcattcca atcagaccag gtaggagctt tctgttttca
1741 tatgttttcag ggttgacacag ttggtctctt taatgtcggg gtggagatcc aaagtggggt
1801 gtggaaagag cgtccatagg agaagtgaga atactgtgaa aaagggatgt tagcattcat
1861 tagagtatga ggatgagtcc caagaagggt ctttggaagg aggacgaata gaatggagta
1921 atgaaattct tgccatgtgc tgaggagata gccagcatta ggtgacaatc ttccagaagt
1981 ggtcagggcag aagggtgccct ggtgagagct cctttacagg gactttatgt ggtttagggc
2041 tcagagctcc aaaactctgg gctcagctgc tctgtacct tggaggcca ttccatggg
2101 aaagtatttt ggaatgtgtc ttttgaagag agcatcagag ttcttaaggg actgggtaag
2161 gcctgaccct gaaatgacca tggatatttt tctacctaca gtttgagtca actagaatat
2221 gcctggggac cttgaagaat ggcccttcag tggccctcac catttgttca tgcttcagtt
2281 aattcagggtg ttgaaggagc ttaggtttta gaggcacgta gacttggttc aagtctcggt
2341 agtagttgaa tagcctcagg caagtcactg cccacctaag atgatgggtc ttcaactata
2401 aaatggagat aatggttaca aatgtctctt cctatagtat aatctccata agggcatggc
2461 ccaagtctgt ctttgactct gcctatccct gacatttagt agcatgcccg acatacaatg
2521 ttagctattg gtattattgc catatagata aattatgtat aaaaattaaa ctgggcaata
2581 gcctaagaag gggggaatat tgtaacacaa atttaaacc actacgcagg gatgagggtc
2641 tataatatga ggacotttta acttccatca ttttctgtt tottgaaata gtttatcttg
2701 taatgaaata taaggcacct cccactttta tgtatagaaa gaggtctttt aattttttt
2761 taatgtgaga aggaaggag gagtaggaat cttgagattc cagatcgaaa atactgtact
2821 ttggttgatt ttaagtggg cttccattcc atggatttaa tcagtcccaa gaagatcaaa
2881 ctcagcagta cttgggtgct gaagaactgt tggatttacc ctggcacgtg tgccacttgc
2941 cagcttcttg ggacacaga gttcttcaat ccaagttatc agattgtatt tgaaaatgac
3001 agagctggag agttttttga aatggcagtg gcaaataaat aaatactttt ttttaaatgg
3061 aaagacttga tctatggtaa taaatgattt tgttttctga ctggaaaaat aggctacta
3121 aagatgaatc acacttgaga tgtttcttac tcaactctgca cagaaacaaa gaagaaatgt
3181 tatacaggga agtccgtttt cactattagt atgaaccaag aaatgggttca aaaacagtgg
3241 taggagcaat gctttcatag tttcagatat ggtagttatg aagaaaacaa tgctatttgc
3301 tgctattatt gtaagagtct tataattaat ggtactccta taatttttga ttgtgagctc
3361 acctatttgg gttaagcatg ccaatttaa gagaccaagt gtatgtacat tatgttctac
3421 atattcagtg ataaaattac taaactacta tatgtctgct ttaaatttgt actttaatat

3481 tgtcttttgg tattaagaaa gatatgcttt cagaatagat atgcttcgct ttggcaagga
3541 atttgatag aacttgctat ttaaaagagg tgtggggtaa atccttgat aaatctccag
3601 tttagccttt ttgaaaaag ctgactttc aaataactaat ttcacttcaa gcagggtacg
3661 tttctgggtt gtttgcttga cttcagtcac aatttcttat cagaccaatg gctgacctct
3721 ttgagatgct aggctaggct tacctatgtg ttctgtgtca tgtgaatgct gagaagtgtg
3781 acagagatcc aacttcagcc ttgaccccat cagtccctcg ggttaactaa ctgagccacc
3841 ggtcctcatg gctattttta tgagggtatt gatgggttaa tgcattgtct atcccttacc
3901 ccagccattt gcactgccag ctgggaacta taccagacct ggatactgat cccaaagtgt
3961 taaattcaac tacatgctgg agattagaga tgggtccaat aaaggacca gaaccaggat
4021 cttgattgct atagacttat taataatcca ggtcaaagag agtgacacac actctctcaa
4081 gacctggggt gagggagtct gtgttatctg caaggccatt tgaggctcag aaagtctctc
4141 tttctatag atatatgcat actttctgac atataggaat gtatcaggaa tactcaacca
4201 tcacaggcat gttcctacct cagggccttt acatgtcctg tttactctgt ctagaatgct
4261 cttctgtaga tgacctggct tgctctgtca cccttcaggt ccttgctcaa gtgtcatctt
4321 ctcccctagt taaactaccc cacacctgt ctgctttcct tgcttatttt tctccatagc
4381 attttaccat ctcttacatt agacattttt cttattttatt tgtagtttat aagcttcatg
4441 aggcaagtaa ctttgctttg tttcttgctg tatctccagt gccagagca gtgcctggta
4501 tataataaat atttattgac tgagtgaaaa aaaaaaaaaa aaa (SEQ ID NO: 223)。

[0349] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD80蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 gacaagtact gagtgaactc aaaccctctg taaagtaaca gaagttagaa ggggaaatgt
61 cgctctctg aagattaccc aaagaaaaag tgatttgta ttgctttata gactgtaaga
121 agagaacatc tcagaagtgg agtcttacct tgaatcaaaa ggatttaaag aaaaagtggg
181 atttttcttc agcaagctgt gaaactaaat ccacaacctt tggagacca ggaacacctt
241 ccaatctctg tgtgttttgt aaacatcact ggagggtctt ctacgtgagc aattggattg
301 tcacagccc tgctgtttt gcacctggga agtgccctgg tcttacttgg gtccaaattg
361 ttggctttca cttttgacct taagcatctg aagccatggg ccacacacgg aggcaggga
421 catcaccatc caagtgtcca tacctcaatt tctttcagct cttggtgctg gctggctctt
481 ctacttctg ttcagggtgt atccacgtga ccaaggaagt gaaagaagt gcaacgctgt
541 cctgtggtca caatgtttct gttgaagagc tggcacaac tcgcatctac tggcaaaagg

601 agaagaaaat ggtgctgact atgatgtctg gggacatgaa tatatggccc gagtacaaga
661 accggaccat ctttgatata actaataacc tctccattgt gatcctggct ctgcgccat
721 ctgacgaggg cacatacgag tgtgttggtc tgaagtatga aaaagacgct ttcaagcggg
781 aacacctggc tgaagtgaag ttatcagtca aagctgactt cctacacct agtatatctg
841 actttgaaat tccaacttct aatattagaa ggataatttg ctcaacctct ggagggtttc
901 cagagcctca cctctcctgg ttggaaaatg gagaagaatt aaatgccatc aacacaacag
961 tttccaaga tctgaaact gagctctatg ctgttagcag caaactggat ttcaatatga
1021 caaccaacca cagcttcatg tgtctcatca agtatggaca tttagagtg aatcagacct
1081 tcaactggaa tacaaccaag caagagcatt ttctgataa cctgctcca tctgggcca
1141 ttaccttaat ctcatgaaat ggaatttttg tgatatgctg cctgacctac tgctttgcc
1201 caagatgcag agagagaagg aggaatgaga gattgagaag ggaaagtgtg cgccctgtat
1261 aacagtgtcc gcagaagcaa ggggctgaaa agatctgaag gtccacctc catttgcaat
1321 tgacctcttc tgggaacttc ctcatgga caagattacc ccaccttgc ctttacgtat
1381 ctgctcttag gtgcttcttc acttcagttg ctttgaggga agtgtctaga ggaatatggt
1441 gggcacagaa gtgctctgg tgacctgat caagggtgtt tgaaatgcag aattcttgag
1501 ttctggaagg gacttttagg aataccagtg ttattaatga caaaggcact gaggccag
1561 gaggtgacct gaattataaa ggccagcgcc agaaccaga tttcctaact ctggtgctct
1621 ttccctttat cagtttgact gtggcctgtt aactggtata tacatatata tgtcaggcaa
1681 agtctgctg gaagtagaat ttgtccaata acagggtcaac ttcagagact atctgatttc
1741 ctaatgtcag agtagaagat tttatgctgc tgtttacaaa agccaatgt aatgcatagg
1801 aagtatggca tgaacatctt taggagacta atggaaatat tattggtgtt taccagtat
1861 tccatTTTTT tcattgtgtt ctctattgt gctctctcac tccccatga ggtacagcag
1921 aaaggagAAC tatccaaac taatttctc tgacatgtaa gacgaatgat ttaggtacgt
1981 caaagcagta gtcaaggagg aaagggatag tccaaagact taactggttc atattggact
2041 gataatctct ttaaattggc ttatgctagt ttgacctcat ttgtaaaata tttatgagaa
2101 agttctcatt taaaatgaga tcgttggtta cagtgtatgt actaagcagt aagctatctt
2161 caaatgtcta aggtagtaac tttccatagg gctccttag atccctaaga tggtttttc
2221 tcttggtat ttctgggtct ttctgacatc agcagagAAC tggaaagaca tagccaactg
2281 ctgttcatgt tactcatgac tctttctct aaaactgcct tccacaattc actagaccag
2341 aagtggacgc aacttaagct gggataatca cattatcatc tgaaaatctg gagttgaaca
2401 gcaaaagaag acaacatttc tcaaatgcac atctcatggc agctaagcca catggctggg

2461 atttaaagcc tttagagcca gcccatggct ttagctacct cactatgctg cttcacaaac
2521 cttgctcctg tgtaaaacta tattctcagt gtagggcaga gaggtctaac accaacataa
2581 ggtactagca gtgtttcccg tattgacagg aatacttaac tcaataattc ttttcttttc
2641 catttagtaa cagttgtgat gactatgttt ctattctaag taattcctgt attctacagc
2701 agatactttg tcagcaatac taagggaaga aacaaagttg aaccgtttct ttaataa

(SEQ ID NO: 224)

[0350] 本公开文本的与本公开文本的编码CD80蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自包含SEQ ID NO:330至SEQ ID NO:3067中任一个的序列的核酸或由其组成。

[0351] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD86蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 agtcattgcc gaggaaggct tgcacagggt gaaagctttg cttctctgct gctgtaacag
61 ggactagcac agacacacgg atgagtgggg tcatttccag atattaggtc acagcagaag
121 cagccaaaat ggatccccag tgcactatgg gactgagtaa cattctcttt gtgatggcct
181 tcctgctctc tgggtgctgct cctctgaaga ttcaagctta tttcaatgag actgcagacc
241 tgccatgcca atttgcaaac tctcaaaacc aaagcctgag tgagctagta gtattttggc
301 aggaccagga aaacttggtt ctgaatgagg tatacttagg caaagagaaa tttgacagtg
361 ttcattccaa gtatatgggc cgcacaagtt ttgattcgga cagttggacc ctgagacttc
421 acaatcttca gatcaaggac aagggttgt atcaatgtat catccatcac aaaaagccca
481 caggaatgat tcgcatccac cagatgaatt ctgaactgtc agtgcttgct aacttcagtc
541 aacctgaaat agtaccaatt tctaataata cagaaaatgt gtacataaat ttgacctgct
601 catctataca cggttacca gaacctaaga agatgagtgt tttgctaaga accaagaatt
661 caactatcga gtatgatggt attatgcaga aatctcaaga taatgtcaca gaactgtacg
721 acgtttccat cagcttgtct gtttcattcc ctgatgttac gagcaatatg accatcttct
781 gtattctgga aactgacaag acgcggcttt tatcttcacc tttctctata gagcttgagg
841 accctcagcc tccccagac cacattcctt ggattacagc tgtacttcca acagttatta
901 tatgtgtgat ggttttctgt ctaattctat ggaaatggaa gaagaagaag cggcctcgca
961 actcttataa atgtggaacc aacacaatgg agagggaaga gagtgaacag accaagaaaa
1021 gagaaaaaat ccatatacct gaaagatctg atgaagccca gcgtgttttt aaaagttcga
1081 agacatcttc atgcgacaaa agtgatacat gtttttaatt aaagagtaaa gcccatacaa

1141 gtattcattt tttctaccct ttcctttgta agttcctggg caaccttttt gattttcttcc
1201 agaaggcaaa aagacattac catgagtaat aagggggctc caggactccc tctaagtgga
1261 atagcctccc tgtaactcca gctctgctcc gtatgccaag aggagacttt aattctctta
1321 ctgcttcttt tcacttcaga gcacacttat gggccaagcc cagcttaatg gctcatgacc
1381 tggaaataaa atttaggacc aatacctcct ccagatcaga ttcttctctt aatttcatag
1441 attgtgtttt ttttttaaat agacctctca atttctggaa aactgccttt tatctgcccc
1501 gaattctaag ctggtgcccc actgaatttt gtgtacctgt gactaaacaa ctacctctc
1561 agtctgggtg ggacttatgt atttatgacc ttatagtgtt aatatcttga aacatagaga
1621 tctatgtact gtaatagtgt gattactatg ctctagagaa aagtctaccc ctgctaagga
1681 gttctcatcc ctctgtcagg gtcagtaagg aaaacggtgg cctagggtag aggcaacaat
1741 gagcagacca acctaaattt ggggaaatta ggagaggcag agatagaacc tggagccact
1801 tctatctggg ctgttgctaa tattgaggag gcttgcccc cccaacaagc catagtggag
1861 agaactgaat aaacaggaaa atgccagagc ttgtgaacct tgtttctctt gaagaactga
1921 ctagtgagat ggctgggga agctgtgaaa gaaccaaag agatcacaat actcaaaaga
1981 gagagagaga gaaaaagag agatcttgat ccacagaaat acatgaaatg tctgggtctgt
2041 ccaccccatc aacaagtctt gaaacaagca acagatggat agtctgtcca aatggacata
2101 agacagacag cagtttccct ggtggtcagg gaggggtttt ggtgataccc aagttattgg
2161 gatgtcatct tcttggaagc agagctgggg agggagagcc atcaccttga taatgggatg
2221 aatggaagga ggcttaggac tttccactcc tggctgagag aggaagagct gcaacggaat
2281 taggaagacc aagacacaga tcaccgggg cttacttagc ctacagatgt cctacgggaa
2341 cgtgggctgg ccagcatag ggctagcaaa tttgagttgg atgattgttt ttgctcaagg
2401 caaccagagg aaacttgcac acagagacag atatactggg agaaatgact ttgaaaacct
2461 ggctctaagg tgggatcact aagggatggg gcagtctctg cccaaacata aagagaactc
2521 tggggagcct gagccacaaa aatgttcctt tattttatgt aaacctcaa gggttataga
2581 ctgcatgct agacaagctt gtccatgtaa tattcccatg tttttaccct gccctgcct
2641 tgattagact cctagcacct ggctagtttc taacatgttt tgtgcagcac agtttttaat
2701 aaatgcttgt tacattcatt taaaaaaaaaaaaa. (SEQ ID NO: 226)

[0352] 本公开文本的与本公开文本的编码CD86蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自SEQ ID NO:3068至SEQ ID NO:5783中任一个的序列的核酸或由其组成。

[0353] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列

的gRNA,所述RNA分子编码CD86蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

```
1  ccotttctgt atttgagttc taccgtcagt cctggcatta tttctctctc tacaaggagc
61  cttaggaggt acggggagct cgcaaatact ccttttggtt tattcttacc accttgcttc
121 tgtgttcctt gggaatgctg ctgtgcttat gcatctggtc tctttttgga gctacagtgg
181 acaggcattt gtgacagcac tatgggactg agtaacattc tctttgtgat ggcttctctg
241 ctctctggtg ctgctcctct gaagattcaa gcttatttca atgagactgc agacctgcca
301 tgccaatttg caaactctca aaaccaaagc ctgagtgagc tagtagtatt ttggcaggac
361 caggaaaact tggttctgaa tgaggtatac ttaggcaaag agaaatttga cagtgttcat
421 tccaagtata tgggccgcac aagttttgat tctggacagt ggacctgag acttcacaat
481 cttcagatca aggacaaggg cttgtatcaa tgtatcatcc atcacaaaaa gcccacagga
541 atgattcgca tccaccagat gaattctgaa ctgtcagtgc ttgctaactt cagtcaacct
601 gaaatagtac caatttctaa tataacagaa aatgtgtaca taaatttgac ctgctcatct
661 atacacgggt acccagaacc taagaagatg agtgttttgc taagaaccaa gaattcaact
721 atcgagtatg atggtattat gcagaaatct caagataatg tcacagaact gtacgacgtt
781 tccatcagct tgtctgtttc attccctgat gttacgagca atatgaccat cttctgtatt
841 ctggaaactg acaagacgcg gcttttatct tcacctttct ctatagagct tgaggacctt
901 cagcctcccc cagaccacat tccttggatt acagctgtac ttccaacagt tattatatgt
961 gtgatggttt tctgtctaata tctatggaaa tggaagaaga agaagcggcc tcgcaactct
1021 tataaatgtg gaaccaacac aatggagagg gaagagagtg aacagaccaa gaaaagagaa
1081 aaaatccata tacctgaaag atctgatgaa gccacgctg tttttaaaag ttcgaaagaca
1141 tcttcatgcg acaaaagtga tacatgtttt taattaaaga gtaaagccca tacaagtatt
1201 cattttttct accctttcct ttgtaagttc ctgggcaacc tttttgattt cttccagaag
1261 gcaaaaagac attaccatga gtaataaggg ggctccagga ctccctctaa gtggaatagc
1321 ctccctgtaa ctccagctct gctccgtatg ccaagaggag actttaattc tcttactgct
1381 tcttttctact tcagagcaca cttatgggcc aagcccagct taatggctca tgacctggaa
1441 ataaaattta ggaccaatac ctctccaga tcagattctt ctcttaattt catagattgt
1501 gttttttttt taaatagacc tctcaatttc tggaaaactg ccttttatct gccagaatt
1561 ctaagctggg gcccactga attttggtga cctgtgacta aacaactacc tctcagttct
1621 ggggtgggact tatgtattta tgaccttata gtgttaatat cttgaaacat agagatctat
1681 gtactgtaat agtgtgatta ctatgctcta gagaaaagtc taccctgct aaggagttct
```

1741 catccctctg tcagggctcag taaggaaaac ggtggcctag ggtacaggca acaatgagca
1801 gaccaaccta aatttgggga aattaggaga ggcagagata gaacctggag ccacttctat
1861 ctgggctggt gctaattattg aggaggcttg cccacccaa caagccatag tggagagaac
1921 tgaataaaca ggaaaatgcc agagcttggtg aacctgttt ctcttgaaga actgactagt
1981 gagatggcct ggggaagctg tgaaagaacc aaaagagatc acaatactca aaagagagag
2041 agagagaaaa aagagagatc ttgatccaca gaaatacatg aaatgtctgg tctgtccacc
2101 ccatcaacaa gtcttgaaac aagcaacaga tggatagtct gtccaaatgg acataagaca
2161 gacagcagtt tccctgggtg tcagggaggg gttttggtga tacccaagtt attgggatgt
2221 catcttctctg gaagcagagc tggggaggga gagccatcac cttgataatg ggatgaatgg
2281 aaggaggctt aggactttcc actcctggct gagagaggaa gagctgcaac ggaattagga
2341 agaccaagac acagatcacc cggggcttac ttagcctaca gatgtcctac gggaacgtgg
2401 gctggcccag catagggcta gcaaatttga gttggatgat tgtttttgct caaggcaacc
2461 agaggaaact tgcatacaga gacagatata ctgggagaaa tgactttgaa aacctggctc
2521 taagggtgga tactaagggt atggggcagt ctctgccaa acataaagag aactctgggg
2581 agcctgagcc acaaaaatgt tcctttatct tatgtaaacc ctcaagggtt atagactgcc
2641 atgctagaca agcttgtcca tgtaatatcc ccatgttttt accctgcccc tgctttgatt
2701 agactcctag cacctggcta gtttctaaca tgttttgtgc agcacagttt ttaataaatg
2761 cttgtttacat tcattttaaaaaaaaaaaaaa. (SEQ ID NO: 227)

[0354] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD86蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 ccctttctgt atttgagttc taccgtcagt cctggcatta tttctctctc tacaaggagc
61 cttaggaggt acggggagct cgcaaatact ccttttggtt tattcttacc accttgcttc
121 tgtgttcctt gggaatgctg ctgtgcttat gcatctggtc tctttttgga gctacagtgg
181 acaggcatctt gtgacagcac tatgggactg agtaacatcc tctttgtgat ggcttctctg
241 ctctctgggtg ctgctcctct gaagattcaa gcttatttca atgagactgc agacctgcca
301 tgccaatttg caaactctca aaaccaaagc ctgagtgagc tagtagtatt ttggcaggac
361 caggaaaaact tggttctgaa tgaggtatac ttaggcaaag agaaatttga cagtgttcat
421 tccaagtata tgggccgcac aagttttgat tcggacagtt ggacctgag acttcacaat
481 cttcagatca aggacaagggt cttgtatcaa tgtatcatcc atcacaaaaa gccacagga
541 atgattcgca tccaccagat gaattctgaa ctgtcagtgc ttgctaactt cagtcaacct

601 gaaatagtagtac caattttctaa tataacagaa aatgtgtaca taaatttgac ctgctcatct
661 atacacgggtt acccagaacc taagaagatg agtggtttgc taagaaccaa gaattcaact
721 atcgagtatg atggtattat gcagaaatct caagataatg tcacagaact gtacgacggt
781 tccatcagct tgtctgttct attcctgat gttacgagca atatgacat cttctgtatt
841 ctggaaactg acaagacgcg gcttttatct tcacctttct ctataggaac caacacaatg
901 gagagggaag agagtgaaca gaccaagaaa agagaaaaaa tccatatacc tgaaagatct
961 gatgaagccc agcgtgtttt taaaagtctg aagacatctt catgagacaa aagtataca
1021 tggttttaaat taaagagtaa agcccatata agtattcatt tttctaccc tttccttgt
1081 aagttcctgg gcaacctttt tgatttcttc cagaaggcaa aaagacatta ccatgagtaa
1141 taagggggct ccaggactcc ctctaagtgg aatagcctcc ctgtaactcc agctctgctc
1201 cgtatgcaa gaggagactt taattctctt actgcttctt ttcacttcag agcacactta
1261 tgggccaagc ccagcttaat ggctcatgac ctggaaataa aatttaggac caatacctcc
1321 tccagatcag attctctctt taatttcata gattgtgttt tttttttaa tagacctctc
1381 aatttctgga aaactgcctt ttatctgccc agaattctaa gctgggtgcc cactgaattt
1441 tgtgtacctg tgactaaaca actacctct cagtctgggt gggacttatg tatttatgac
1501 cttatagtgt taatatcttg aaacatagag atctatgtac tgtaatagtg tgattactat
1561 gctctagaga aaagtctacc cctgctaagg agttctcatc cctctgtcag ggtcagtaag
1621 gaaaacggtg gcctagggtg caggcaacaa tgagcagacc aacctaaatt tggggaaatt
1681 aggagaggca gagatagaac ctggagccac ttctatctgg gctgttgcta atattgagga
1741 ggcttgcccc acccaacaag ccatagtgga gagaactgaa taaacaggaa aatgccagag
1801 cttgtgaacc ctgtttctct tgaagaactg actagtgaga tggcctgggg aagctgtgaa
1861 agaacaaaa gagatcacia tactcaaaag agagagagag agaaaaaaga gagatcttga
1921 tccacagaaa tacatgaaat gtctggtctg tccacccat caacaagtct tgaaacaagc
1981 aacagatgga tagtctgtcc aatggacat aagacagaca gcagtttccc tgggtggcag
2041 ggaggggttt tggatgatacc caagttattg ggatgtcatc ttctggaag cagagctggg
2101 gagggagagc catcaccttg ataattggat gaatggaag aggcttagga ctttccactc
2161 ctggctgaga gaggaagagc tgcaacggaa ttaggaagac caagacacag atcaccggg
2221 gcttacttag cctacagatg tcctacggga acgtgggctg gccagcata gggctagcaa
2281 atttgagttg gatgattgtt tttgctcaag gcaaccagag gaaacttgca tacagagaca
2341 gatatactgg gagaaatgac tttgaaaacc tggctctaag gtgggatcac taagggatgg
2401 ggcagtctct gccaaacat aaagagaact ctggggagcc tgagccacaa aatgttctc

2461 ttattttatg taaacctca agggttatag actgccatgc tagacaagct tgtccatgta
2521 atattcccat gtttttacc tgccctgcc ttgattagac tccatgcacc tggetagttt
2581 ctaacatggt ttgtgcagca cagtttttaa taaatgcttg ttacattcat taaaaaaaa
2641aaaaaa. (SEQ ID NO: 228)

[0355] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD86蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 agtcattgcc gaggaaggct tgcacagggt gaaagctttg cttctctgct gctgtaacag
61 ggactagcac agacacacgg atgagtgggg tcatttccag atattaggtc acagcagaag
121 cagccaaaat ggatccccag tgcactatgg gactgagtaa cattctcttt gtgatggcct
181 tcctgctctc tgctaacttc agtcaacctg aaatagtacc aatttctaataaacagaaa
241 atgtgtacat aaatttgacc tgctcatcta tacacgggta ccagaacct aagaagatga
301 gtgttttgct aagaaccaag aattcaacta tcgagtatga tgggtattatg cagaaatctc
361 aagataatgt cacagaactg tacgacgttt ccatcagctt gtctgtttca ttccctgatg
421 ttacgagcaa tatgaccatc ttctgtattc tggaaactga caagacgcgg cttttatctt
481 cacctttctc tatagagctt gaggacctc agcctcccc agaccacatt ccttgatta
541 cagctgtact tccaacagtt attatatgtg tgatggtttt ctgtctaatt ctatggaaat
601 ggaagaagaa gaagcggcct cgcaactctt ataaatgtgg aaccaacaca atggagaggg
661 aagagagtga acagaccaag aaaagagaaa aaatccatat acctgaaaga tctgatgaag
721 ccagcgtgt ttttaaagt tcgaagacat cttcatgca caaagtgat acatgttttt
781 aattaaagag taaagcccat acaagtattc attttttcta ccttttctt tgtaagttcc
841 tgggcaacct ttttgatttc ttccagaagg caaaaagaca ttaccatgag taataagggg
901 gctccaggac tccctctaag tggaatagcc tccctgtaac tccagctctg ctccgtatgc
961 caagaggaga ctttaattct cttactgctt cttttcactt cagagcacac ttatgggcca
1021 agcccagctt aatggctcat gacctgaaa taaaatttag gaccaatacc tccctcagat
1081 cagattcttc tcttaatttc atagattgtg tttttttttt aaatagacct ctcaatttct
1141 ggaaaactgc cttttatctg ccagaattc taagctgggtg cccactgaa tttgtgtac
1201 ctgtgactaa acaactacct cctcagtctg ggtgggactt atgtatttat gaccttatag
1261 tgttaatatc ttgaaacata gagatctatg tactgtaata gtgtgattac tatgctctag
1321 agaaaagtct acccctgcta aggagttctc atccctctgt cagggtcagt aaggaaaacg
1381 gtggcctagg gtacaggcaa caatgagcag accaacctaa atttggggaa attaggagag

1441 gcagagatag aacctggagc cacttctatc tgggctgttg ctaatattga ggaggcttgc
1501 cccacccaac aagccatagt ggagagaact gaataaacag gaaaatgcc a gagcttgtga
1561 accctgtttc tcttgaagaa ctgactagt agatggcctg gggaagctgt gaaagaacca
1621 aaagagatca caatactcaa aagagagaga gagagaaaaa agagagatct tgatccacag
1681 aaatacatga aatgtctggg ctgtccaccc catcaacaag tcttgaaca agcaacagat
1741 ggatagtctg tccaaatgga cataagacag acagcagttt ccttgggtgg cagggagggg
1801 ttttgggtgat acccaagtta ttgggatgtc atcttctctg aagcagagct ggggagggag
1861 agccatcacc ttgataatgg gatgaatgga aggaggctta ggactttcca ctctggctg
1921 agagaggaag agctgcaacg gaattaggaa gaccaagaca cagatcacc ggggcttact
1981 tagcctacag atgtcctacg ggaacgtggg ctggcccagc atagggctag caaatttgag
2041 ttggatgatt gtttttgctc aaggcaacca gaggaaactt gcatacagag acagatatac
2101 tgggagaaat gactttgaaa acctggctct aagggtggat cactaaggga tggggcagtc
2161 tctgccccaa cataaagaga actctgggga gctgagcca caaaatgtt cttttatttt
2221 atgtaaaccc tcaagggta tagactgcca tgctagacaa gcttgtccat gtaatatcc
2281 catgttttta cctgcccct gccttgatta gactcctagc acctggctag tttctaacat
2341 gttttgtgca gcacagtttt taataaatgc ttgttacatt catttaaaaa aaaaaaaaa.

(SEQ ID NO: 229)

[0356] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CD86蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 agtcattgcc gaggaaggct tgcacaggt gaaagctttg cttctctgct gctgtaacag
61 ggactagcac agacacacgg atgagtgggg tcatttccag atattaggtc acagcagaag
121 cagccaaaat ggatccccag tgggtgctgct cctctgaaga ttcaagctta tttcaatgag
181 actgcagacc tgccatgcc aattgcaaac tctcaaaacc aaagcctgag tgagctagta
241 gtattttggc aggaccagga aaacttggtt ctgaatgagg tatacttagg caaagagaaa
301 ttgacagtg ttcatccaa gtatatgggc cgcacaagtt ttgattcgga cagttggacc
361 ctgagacttc acaatcttca gatcaaggac aagggttgt atcaatgtat catccatcac
421 aaaaagccca caggaatgat tcgcatccac cagatgaatt ctgaactgtc agtgcttgc
481 aacttcagtc aacctgaaat agtaccaatt tctaataata cagaaaatgt gtacataaat
541 ttgacctgct catctataca cggttacca gaacctaga agatgagtgt tttgctaaga
601 accaagaatt caactatcga gtatgatggg attatgcaga aatctcaaga taatgtcaca

661 gaactgtacg acgtttccat cagcttgtct gtttcattcc ctgatgttac gagcaatatg
721 accatcttct gtattctgga aactgacaag acgcggcttt tatcttcacc tttctctata
781 gagcttgagg accctcagcc tccccagac cacattcctt ggattacagc tgtacttcca
841 acagttatta tatgtgtgat ggttttctgt ctaattctat ggaaatggaa gaagaagaag
901 cggcctcgca actcttataa atgtggaacc aacacaatgg agagggaaga gagtgaacag
961 accaagaaaa gagaaaaaat ccatatacct gaaagatctg atgaagccca gcggtgtttt
1021 aaaagttcga agacatcttc atgcgacaaa agtgatacat gtttttaatt aaagagtaaa
1081 gccatacaa gtattcattt tttctaccct ttcctttgta agttcctggg caacctttt
1141 gatttcttcc agaaggcaaa aagacattac catgagtaat aagggggctc caggactccc
1201 tctaagtgga atagcctccc tgtaactcca gctctgctcc gtatgccaag aggagacttt
1261 aattctctta ctgcttcttt tcaactcaga gcacacttat gggccaagcc cagcttaatg
1321 gctcatgacc tggaaataaa atttaggacc aatacctcct ccagatcaga ttcttctctt
1381 aatttcatag attgtgtttt ttttttaa atagacctctca atttctggaa aactgccttt
1441 tatctgcccga gaattctaag ctggtgcccc actgaatttt gtgtacctgt gactaaacaa
1501 ctacctctc agtctgggtg ggacttatgt atttatgacc ttatagtgtt aatatcttga
1561 aacatagaga tctatgtact gtaatagtgt gattactatg ctctagagaa aagtctaccc
1621 ctgctaagga gttctcatcc ctctgtcagg gtcagtaagg aaaacgggtg cctagggtag
1681 aggcaacaat gagcagacca acctaaattt ggggaaatta ggagaggcag agatagaacc
1741 tggagccact tctatctggg ctgttgctaa tattgaggag gcttgcccca cccaacaagc
1801 catagtggag agaactgaat aaacaggaaa atgccagagc ttgtgaaccc tgtttctctt
1861 gaagaactga ctagtgagat ggctgggga agctgtgaaa gaaccaaag agatcacaat
1921 actcaaaaga gagagagaga gaaaaagag agatcttgat ccacagaaat acatgaaatg
1981 tctggtctgt ccacccatc aacaagtctt gaaacaagca acagatggat agtctgtcca
2041 aatggacata agacagacag cagtttccct ggtggtcagg gaggggtttt ggtgataccc
2101 aagttattgg gatgtcatct tcttggaagc agagctgggg agggagagcc atcaccttga
2161 taatgggatg aatggaagga ggcttaggac tttccactcc tggctgagag aggaagagct
2221 gcaacggaat taggaagacc aagacacaga tcaccggggg ctacttagc ctacagatgt
2281 cctacgggaa cgtgggctgg ccagcatag ggctagcaaa tttgagttgg atgattgttt
2341 ttgctcaagg caaccagagg aaacttgc atagagacag atatactggg agaaatgact
2401 ttgaaaacct ggctctaagg tgggatcact aagggatggg gcagtctctg cccaacata
2461 aagagaactc tggggagcct gagccacaaa aatgttccct tattttatgt aaacctcaa

2521 gggttataga ctgccatgct agacaagctt gtccatgtaa tattcccatg tttttaccct
2581 gccctgcct tgattagaact cctagcacct ggctagtttc taacatgttt tgtgcagcac
2641 agtttttaaat aaatgcttgt tacattcatt taaaaaaaaaaaaa.

(SEQ ID NO: 230)

[0357] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码ICOSLG蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 AGTTAGAGCC GATCTCCCGC GCCCCGAGGT TGCTCCTCTC CGAGGTCTCC CGCGGCCCAA
61 GTTCTCCGCG CCCCAGAGTC TCCGCGCCCC GAGGTCTCCG CGGCCCGAGG TCTCCGCCCCG
121 CACCATGCGG CTGGGCAGTC CTGGACTGCT CTTCCCTGCTC TTCAGCAGCC TTCGAGCTGA
181 TACTCAGGAG AAGGAAGTCA GAGCGATGGT AGGCAGCGAC GTGGAGCTCA GCTGCGCTTG
241 CCCTGAAGGA AGCCGTTTTG ATTTAAATGA TGTTTACGTA TATTGGCAAA CCAGTGAGTC
301 GAAAACCGTG GTGACCTACC ACATCCCACA GAACAGCTCC TTGGAAAACG TGGACAGCCG
361 CTACCGGAAC CGAGCCCTGA TGTCAACGGC CGGCATGCTG CGGGGCGACT TCTCCCTGCG
421 CTTGTTCAAC GTCACCCCCC AGGACGAGCA GAAGTTTCAC TGCCTGGTGT TGAGCCAATC
481 CCTGGGATTC CAGGAGGTTT TGAGCGTTGA GGTACACTG CATGTGGCAG CAACTTCAG
541 CGTGCCCGTC GTCAGCGCCC CCCACAGCCC CTCCCAGGAT GAGCTCACCT TCACGTGTAC
601 ATCCATAAAC GGCTACCCCA GGCCCAACGT GTACTGGATC AATAAGACGG ACAACAGCCT
661 GCTGGACCAG GCTCTGCAGA ATGACACCGT CTTCTTGAAC ATGCGGGGCT TGTATGACGT
721 GGTCAGCGTG CTGAGGATCG CACGGACCCC CAGCGTGAAC ATTGGCTGCT GCATAGAGAA
781 CGTGCTTCTG CAGCAGAACC TGACTGTCGG CAGCCAGACA GGAAATGACA TCGGAGAGAG
841 AGACAAGATC ACAGAGAATC CAGTCAGTAC CGGCGAGAAA AACGCGGCCA CGTGGAGCAT
901 CCTGGCTGTC CTGTGCCTGC TTGTGGTCGT GGCGGTGGCC ATAGGCTGGG TGTGCAGGGA
961 CCGATGCCTC CAACACAGCT ATGCAGGTGC CTGGGCTGTG AGTCCGGAGA CAGAGCTCAC
1021 TGGTGAGTTT GCCGTGGGAA GCAGCAGGTT CTGGGGGGCC CAGGGGAGGC TTGGCTGCCA
1081 GCTGTCTTTC AGAGTTTCAA AAAACTTTCA AAAGGCAAAA GTCCCTTGCC TTGAACAACT
1141 GTTGTTCTTG GAGACGCAGC GAAGCCCTCG ATGGTGCGCA TGGCATTTCC TGCAGCCTCC
1201 CCTTGGCATG GGATGGCATC CTGGTGTGCA CTTTGTGACA CTGCGATGGG ATTTTCCCAA
1261 CATGCACAGA AGCAGAGAGA CGAGTGCTAG ACCCCGCGC TCCCCAGTGC CCAGCCCCGA
1321 CCAGGGTGTC CAGGGCGGGT CCAGGCACCG GCGCCCAGCC CCCATGGGGT GTCCGGAGTG
1381 GGTCCAGGCA CCGGCGCCCA GCCCCGTGG GGTGTCCAGG GCGGGTCCAG GCACCGGCGC

1441 CCAGCCCCTG TGGGGTGTCC GGAGTGGGTC CGGGCACCGC CAGCTTCTCT CTGTGGCAGC
1501 CACTCCTGCA GCTCTCGTTT GCCCTCAGT TCCAGGAGCA ACATAGATGT GGATTCCTGT
1561 CCAATTTGGG AAAAATGTCC ACACACGGTC ACCCACCTGG CAGGTGCCTC TGGCTGCAAG
1621 GGGCGCTGGG CTTCGCAGGC AGGCCAGCCG GGCTCCCCGC CATGGGCCAG GATCCCCTCC
1681 GAGCCCTGTT TGCCGCCCAG GAGAAGGGGT TCCCCGGGA CAGTGGGCTC AGGGTGTGCG
1741 CAGCCACCAT GCTGTGGTGT CACCTGTGGA CCCAGGCGAG CTGATGGCCG ACCGCAGAAA
1801 CGCACTTCCA AGGCCAGGTC GGCCCATCCA GATGATGCAG GAACACAGCT TGCTAAAAAC
1861 ACGGCCGGCC TGTTCCCGTC GGAGCCAGTC GAAGTCCCT GAACAGGCCG CTGTTTCCGA
1921 AGCTTTAAAC CCTGTGTTTC CACCAAGCTG AGTCCTGAGA AAACCGACGT CTGCCTGCAG
1981 AAGGGAAAGG GGTGCTTCAT GTTCCTCTCT CTCCTTCATC TCCCT

(SEQ ID NO: 231)。

[0358] 本公开文本的与本公开文本的编码IOSLG蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自SEQ ID NO:5784至SEQ ID NO:7789中任一个的序列的核酸或由其组成。

[0359] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码OX40L蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 GGCCCTGGGA CCTTGCCTA TTTTCTGATT GATAGGCTTT GTTTTGTCTT TACCTCCTTC
61 TTTCTGGGGA AAACCTCAGT TTTATCGCAC GTTCCCCTTT TCCATATCTT CATCTTCCCT
121 CTACCCAGAT TGTGAAGATG GAAAGGGTCC AACCCCTGGA AGAGAATGTG GGAAATGCAG
181 CCAGGCCAAG ATTCGAGAGG AACAAGCTAT TGCTGGTGGC CTCTGTAATT CAGGGACTGG
241 GGCTGCTCCT GTGCTTCACC TACATCTGCC TGCACTTCTC TGCTCTTCAG GTATCACATC
301 GGTATCCTCG AATTCAAAGT ATCAAAGTAC AATTTACCGA ATATAAGAAG GAGAAAGGTT
361 TCATCCTCAC TTCCCAAAG GAGGATGAAA TCATGAAGGT GCAGAACAAC TCAGTCATCA
421 TCAACTGTGA TGGGTTTTAT CTCATCTCCC TGAAGGGCTA CTTCTCCCAG GAAGTCAACA
481 TTAGCCTTCA TTACCAGAAG GATGAGGAGC CCCTCTTCCA ACTGAAGAAG GTCAGGTCTG
541 TCAACTCCTT GATGGTGGCC TCTCTGACTT ACAAAGACAA AGTCTACTTG AATGTGACCA
601 CTGACAATAC CTCCCTGGAT GACTTCCATG TGAATGGCGG AGAACTGATT CTTATCCATC
661 AAAATCCTGG TGAATTCTGT GTCCTTTGAG GGGCTGATGG CAATATCTAA AACCAGGCAC
721 CAGCATGAAC ACCAAGCTGG GGGTGGACAG GGCATGGATT CTTCAATTGCA AGTGAAGGAG
781 CCTCCCAGCT CAGCCACGTG GGATGTGACA AGAAGCAGAT CCTGGCCCTC CCGCCCCAC

841 CCCTCAGGGA TATTTAAAC TTATTTTATA TACCAGTTAA TCTTATTTAT CCTTATATTT
901 TCTAAATTGC CTAGCCGTC CACCCCAAGA TTGCCTTGAG CCTACTAGGC ACCTTTGTGA
961 GAAAGAAAAA ATAGATGCCT CTTCTTCAAG ATGCATTGTT TCTATTGGTC AGGCAATTGT
1021 CATAATAAAC TTATGTCATT GAAAACGGTA CCTGACTACC ATTTGCTGGA AATTGACAT
1081 GTGTGTGGCA TTATCAAAAT GAAGAGGAGC AAGGAGTGAA GGAGTGGGGT TATGAATCTG
1141 CCAAAGGTGG TATGAACCAA CCCCTGGAAG CCAAAGCGGC CTCTCCAAGG TTAAATTGAT
1201 TGCAGTTTGC ATATTGCCTA AATTTAACT TTCTCATTTG GTGGGGGTTT AAAAGAAGAA
1261 TCAGCTTGTG AAAATCAGG ACTTGAAGAG AGCCGTCTAA GAAATACCAC GTGCTTTTTT
1321 TCTTTACCAT TTTGCTTTC CAGCCTCCAA ACATAGTTAA TAGAAATTTT CCTTCAAAGA
1381 ACTGTCTGGG GATGTGATGC TTTGAAAAAT CTAATCAGTG ACTTAAGAGA GATTTTCTTG
1441 TATACAGGGA GAGTGAGATA ACTTATTGTG AAGGGTTAGC TTTACTGTAC AGGATAGCAG
1501 GGAAGTGGAC ATCTCAGGGT AAAAGTCAGT ACGGATTTTA ATAGCCTGGG GAGGAAAACA
1561 CATTCCTTGC CACAGACAGG CAAAGCAACA CATGCTCATC CTCCTGCCTA TGCTGAGATA
1621 CGCACTCAGC TCCATGTCTT GTACACACAG AAACATTGCT GGTTTCAAGA AATGAGGTGA
1681 TCCTATTATC AAATTCAATC TGATGTCAA TAGCACTAAG AAGTTATTGT GCCTTATGAA
1741 AAATAATGAT CTCTGTCTAG AAATACCATA GACCATATAT AGTCTCACAT TGATAATTGA
1801 AACTAGAAGG GTCTATAATC AGCCTATGCC AGGGCTTCAA TGGAATAGTA TCCCCTTATG
1861 TTTAGTTGAA ATGTCCCCTT AACTTGATAT AATGTGTTAT GCTTATGGCG CTGTGGACAA
1921 TCTGATTTTT CATGTCAACT TTCCAGATGA TTTGTAACCT CTCTGTGCCA AACCTTTTAT
1981 AAACATAAAT TTTTGAGATA TGTATTTTAA AATTGTAGCA CATGTTTCCC TGACATTTTC
2041 AATAGAGGAT ACAACATCAC AGAATCTTTC TGGATGATTC TGTGTTATCA AGGAATTGTA
2101 CTGTGCTACA ATTATCTCTA GAATCTCCAG AAAGGTGGAG GGCTGTTTCG CCTTACACTA
2161 AATGGTCTCA GTTGGATTTT TTTTTCCTGT TTTCTATTTT CTCTTAAGTA CACCTTCAAC
2221 TATATTCCCA TCCCTCTATT TTAATCTGTT ATGAAGGAAG GTAAATAAAA ATGCTAAATA
2281 GAAGAAATTG TAGGTAAGGT AAGAGGAATC AAGTTCTGAG TGGCTGCCAA GGCCTCACA
2341 GAATCATAAT CATGGCTAAA TATTTATGGA GGGCCTACTG TGGACCAGGC ACTGGGCTAA
2401 ATACTTACAT TTACAAGAAT CATTCTGAGA CAGATATTCA ATGATATCTG GCTTCACTAC
2461 TCAGAAGATT GTGTGTGTGT TTGTGTGTGT GTGTGTGTGT GTATTTCACT TTTTGTATT
2521 GACCATGTTT TGCAAAATTG CAGTTACTCA GTGAGTGATA TCCGAAAAG TAAACGTTTA
2581 TGACTATAGG TAATATTTAA GAAAATGCAT GGTTCATTTT TAAGTTTGA ATTTTATCT
2641 ATATTTCTCA CAGATGTGCA GTGCACATGC AGGCCTAAGT ATATGTTGTG TGTGTTGTTT

2701 GTCTTTGATG TCATGGTCCC CTCTCTTAGG TGCTCACTCG CTTTGGGTGC ACCTGGCCTG
2761 CTCTTCCCAT GTTGGCCTCT GCAACCACAC AGGGATATTT CTGCTATGCA CCAGCCTCAC
2821 TCCACCTTCC TTCCATCAAA AATATGTGTG TGTGTCTCAG TCCCTGTAAG TCATGTCCTT
2881 CACAGGGAGA ATTAACCCTT CGATATACAT GGCAGAGTTT TGTGGGAAAA GAATTGAATG
2941 AAAAGTCAGG AGATCAGAAT TTAAATTTG ACTTAGCCAC TAACTAGCCA TGTAACCTTG
3001 GGAAAGTCAT TTCCCATTTT TGGGTCTTGC TTTTCTTTCT GTTAAATGAG AGGAATGTTA
3061 AATATCTAAC AGTTTAGAAT CTTATGCTTA CAGTGTTATC TGTGAATGCA CATATTAAAT
3121 GTCTATGTTT TTGTTGCTAT GAGTCAAGGA GTGTAACCTT CTCCTTTACT ATGTTGAATG
3181 TATTTTTTTC TGGACAAGCT TACATCTTCC TCAGCCATCT TTGTGAGTCC TTCAAGAGCA
3241 GTTATCAATT GTTAGTTAGA TATTTTCTAT TTAGAGAATG CTTAAGGGAT TCCAATCCCG
3301 ATCCAAATCA TAATTTGTTC TTAAGTATAC TGGGCAGGTC CCCTATTTTA AGTCATAATT
3361 TTGTATTTAG TGCTTTCCTG GCTCTCAGAG AGTATTAATA TTGATATTAA TAATATAGTT
3421 AATAGTAATA TTGCTATTTA CATGGAAACA AATAAAAGAT CTCAGAATTC ACTAAAAAAA
3481 AAAA (SEQ ID NO: 232)。

[0360] 本公开文本的与本公开文本的编码OX40L蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自SEQ ID NO:7790至SEQ ID NO:11254中任一个的序列的核酸或由其组成。

[0361] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码IL12蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 TTTCGCTTTC ATTTTGGGCC GAGCTGGAGG CGGCGGGGCC GTCCCGGAAC GGCTGCGGCC
61 GGGCACCCCG GGAGTTAATC CGAAAGCGCC GCAAGCCCCG CGGGCCGGCC GCACCGCACG
121 TGTCACCGAG AAGCTGATGT AGAGAGAGAC ACAGAAGGAG ACAGAAAGCA AGAGACCAGA
181 GTCCCGGGAA AGTCCTGCCG CGCCTCGGGA CAATTATAAA AATGTGGCCC CCTGGGTCAG
241 CCTCCAGCC ACCGCCCTCA CCTGCCGCGG CCACAGGTCT GCATCCAGCG GCTCGCCCTG
301 TGTCCCTGCA GTGCCGGCTC AGCATGTGTC CAGCGCGCAG CCTCCTCCTT GTGGCTACCC
361 TGGTCCTCCT GGACCACCTC AGTTTGGCCA GAAACCTCCC CGTGGCCACT CCAGACCCAG
421 GAATGTTCCC ATGCCTTCAC CACTCCCAA ACCTGCTGAG GGCCGTCAGC AACATGCTCC
481 AGAAGGCCAG ACAAACTCTA GAATTTTACC CTTGCACTTC TGAAGAGATT GATCATGAAG
541 ATATCACAAA AGATAAAACC AGCACAGTGG AGGCCTGTTT ACCATTGGAA TTAACCAAGA
601 ATGAGAGTTG CCTAAATTCC AGAGAGACCT CTTTCATAAC TAATGGGAGT TGCCTGGCCT

661 CCAGAAAGAC CTCTTTTATG ATGGCCCTGT GCCTTAGTAG TATTTATGAA GACTTGAAGA
721 TGTACCAGGT GGAGTTCAAG ACCATGAATG CAAAGCTTCT GATGGATCCT AAGAGGCAGA
781 TCTTTCTAGA TCAAAACATG CTGGCAGTTA TTGATGAGCT GATGCAGGCC CTGAATTTCA
841 ACAGTGAGAC TGTGCCACAA AAATCCTCCC TTGAAGAACC GGATTTTTAT AAAACTAAAA
901 TCAAGCTCTG CATACTTCTT CATGCTTTCA GAATTCGGGC AGTGACTATT GATAGAGTGA
961 TGAGCTATCT GAATGCTTCC TAAAAAGCGA GGTCCCTCCA AACCGTTGTC ATTTTTATAA
1021 AACTTTGAAA TGAGGAAACT TTGATAGGAT GTGGATTAAG AACTAGGGAG GGGGAAAGAA
1081 GGATGGGACT ATTACATCCA CATGATACCT CTGATCAAGT ATTTTGTACA TTTACTGTGG
1141 ATAAATTGTT TTTAAGTTTT CATGAATGAA TTGCTAAGAA GGGAAAATAT CCATCCTGAA
1201 GGTGTTTTTC ATTCACTTTA ATAGAAGGGC AAATATTTAT AAGCTATTTT TGTACCAAAG
1261 TGTTTGTGGA AACAAACATG TAAGCATAAC TTATTTTAAA ATATTTATTT ATATAACTTG
1321 GTAATCATGA AAGCATCTGA GCTAACTTAT ATTTATTTAT GTTATATTTA TTAAATTATT
1381 TATCAAGTGT ATTTGAAAAA TATTTTTAAG TGTTCTAAAA ATAAAAGTAT TGAATTAAAG
1441 TGAAAAAAA (SEQ ID NO: 233)。

[0362] 本公开文本的与本公开文本的编码IL12蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自SEQ ID NO:11255至SEQ ID NO:12685中任一个的序列的核酸或由其组成。

[0363] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,所述RNA分子编码CCR7蛋白,所述间隔子序列包含以下的序列的20个核苷酸或由其组成:

1 CACTTCCTCC CCAGACAGGG GTAGTGCGAG GCCGGGCACA GCCTTCCTGT GTGGTTTTAC
61 CGCCCAGAGA GCGTCATGGA CCTGGGGAAA CCAATGAAAA GCGTGCTGGT GGTGGCTCTC
121 CTTGTCATTT TCCAGGTATG CCTGTGTCAA GATGAGGTCA CGGACGATTA CATCGGAGAC
181 AACACCACAG TGGACTACAC TTTGTTTCGAG TCTTTGTGCT CCAAGAAGGA CGTGCGGAAC
241 TTAAAGCCT GGTTCCTCCC TATCATGTAC TCCATCATTT GTTTCGTGGG CCTACTGGGC
301 AATGGGCTGG TCGTGTTGAC CTATATCTAT TTCAAGAGGC TCAAGACCAT GACCGATACC
361 TACCTGCTCA ACCTGGCGGT GGCAGACATC CTCTTCCTCC TGACCCTTCC CTTCTGGGCC
421 TACAGCGCGG CCAAGTCCTG GGTCTTCGGT GTCCACTTTT GCAAGCTCAT CTTTGCCATC
481 TACAAGATGA GCTTCTTCAG TGGCATGCTC CTAATTCTTT GCATCAGCAT TGACCGCTAC
541 GTGGCCATCG TCCAGGCTGT CTCAGCTCAC CGCCACCGTG CCCGCGTCCT TCTCATCAGC
601 AAGCTGTCCT GTGTGGGCAT CTGGATACTA GCCACAGTGC TCTCCATCCC AGAGCTCCTG

661 TACAGTGACC TCCAGAGGAG CAGCAGTGAG CAAGCGATGC GATGCTCTCT CATCACAGAG
721 CATGTGGAGG CCTTTATCAC CATCCAGGTG GCCCAGATGG TGATCGGCTT TCTGGTCCCC
781 CTGCTGGCCA TGAGCTTCTG TTACCTTGTC ATCATCCGCA CCCTGCTCCA GGCACGCAAC
841 TTTGAGCGCA ACAAGGCCAT CAAGGTGATC ATCGCTGTGG TCGTGGTCTT CATAGTCTTC
901 CAGCTGCCCT ACAATGGGGT GGTCCCTGGCC CAGACGGTGG CCAACTTCAA CATCACCAGT
961 AGCACCTGTG AGCTCAGTAA GCAACTCAAC ATCGCCTACG ACGTCACCTA CAGCCTGGCC
1021 TGCGTCCGCT GCTGCGTCAA CCCTTTCTTG TACGCCTTCA TCGGCGTCAA GTTCCGCAAC
1081 GATCTCTTCA AGCTCTTCAA GGACCTGGGC TGCCTCAGCC AGGAGCAGCT CCGGCAGTGG
1141 TCTTCCTGTC GGCACATCCG GCGCTCCTCC ATGAGTGTGG AGGCCGAGAC CACCACCACC
1201 TTCTCCCCAT AGGCGACTCT TCTGCCTGGA CTAGAGGGAC CTCTCCCAGG GTCCCTGGGG
1261 TGGGGATAGG GAGCAGATGC AATGACTCAG GACATCCCCC CGCCAAAAGC TGCTCAGGGA
1321 AAAGCAGCTC TCCCCTCAGA GTGCAAGCCC CTGCTCCAGA AGATAGCTTC ACCCCAATCC
1381 CAGCTACCTC AACCAATGCC AAAAAAAGAC AGGGCTGATA AGCTAACACC AGACAGACAA
1441 CACTGGGAAA CAGAGGCTAT TGTCCCCTAA ACCAAAAACT GAAAGTGAAA GTCCAGAAAC
1501 TGTTCCACC TGCTGGAGTG AAGGGGCCAA GGAGGGTGAG TGCAAGGGGC GTGGGAGTGG
1561 CCTGAAGAGT CCTCTGAATG AACCTTCTGG CCTCCCACAG ACTCAAATGC TCAGACCAGC
1621 TCTTCCGAAA ACCAGGCCTT ATCTCCAAGA CCAGAGATAG TGGGGAGACT TCTTGGCTTG
1681 GTGAGGAAAA GCGGACATCA GCTGGTCAAA CAAACTCTCT GAACCCCTCC CTCCATCGTT
1741 TTCTTCACTG TCCTCCAAGC CAGCGGGAAT GGCAGCTGCC ACGCCGCCCT AAAAGCACAC
1801 TCATCCCCTC ACTTGCCGCG TCGCCCTCCC AGGCTCTCAA CAGGGGAGAG TGTGGTGTTC
1861 CCTGCAGGCC AGGCCAGCTG CCTCCGCGTG ATCAAAGCCA CACTCTGGGC TCCAGAGTGG
1921 GGATGACATG CACTCAGCTC TTGGCTCCAC TGGGATGGGA GGAGAGGACA AGGGAAATGT
1981 CAGGGGCGGG GAGGGTGACA GTGGCCGCCC AAGGCCACG AGCTTGTCTT TTGTTCTTTG
2041 TCACAGGGAC TGAAAACCTC TCCTCATGTT CTGCTTTCGA TTCGTTAAGA GAGCAACATT
2101 TTACCCACAC ACAGATAAAG TTTTCCCTTG AGGAAACAAC AGCTTTAAAA GAAAAAGAAA
2161 AAAAAAGTCT TTGGTAAATG GCAAAAAAAA AAAAAA AAAA

(SEQ ID NO: 234)。

[0364] 本公开文本的与本公开文本的编码CCR7蛋白的RNA分子的靶序列特异性结合的示例性gRNA间隔子序列可以包含具有选自SEQ ID NO:12686至SEQ ID NO:14872中任一个的序列的核酸或由其组成。

[0365] 本公开文本的组合物可以包含含有与RNA分子的靶序列特异性结合的间隔子序列的gRNA,其中所述间隔子序列和所述靶序列是彼此的反向互补体。在一些实施方案中,本公

开文本的组合物可以包含单一(即,单一的)gRNA,其包含a)与第一靶RNA序列特异性结合的第一间隔子序列以及b)与第二靶RNA序列特异性结合的第二间隔子序列,其中所述第一和第二间隔子序列各自结合不同的靶RNA序列。在一些实施方案中,结合不同的靶RNA序列的第一和第二间隔子序列不包含在单一(即,单一的)gRNA内,而是第一间隔子序列包含在第一gRNA内,并且第二间隔子序列包含在第二gRNA序列内。在一些实施方案中,本文公开的间隔子序列包含编码适应性免疫应答的蛋白质组分的核酸序列的一部分,其中所述蛋白质组分选自 β -2-微球蛋白(β 2M)、人白细胞抗原A(HLA-A)、人白细胞抗原B(HLA-B)、人白细胞抗原C(HLA-C)、分化群28(CD28)、分化群80(CD80)、分化群86(CD86)、可诱导T细胞共刺激分子(ICOS)、ICOS配体(ICOSLG)、OX40L、白细胞介素12(IL12)和CC趋化因子受体7(CCR7)。在一些实施方案中,作为编码适应性免疫应答的蛋白质组分的核酸序列的一部分的间隔子的长度是约20或21个核苷酸。

[0366] 本公开文本的所有核苷酸序列可以互换地包括尿嘧啶(U)或胸腺嘧啶(T)。

[0367] sgRNA的示例性的非限制性的靶向Zika NS5的间隔子序列包括但不限于:gcaatgatcttcatgttgggagc(SEQ ID NO:196)、gaaccttggtgatgaactcttc(SEQ ID NO:197)、gttggtgattagagcttcattc(SEQ ID NO:198)和gagtgatcctcgttcaagaatcc(SEQ ID NO:199)。

[0368] sgRNA的示例性的非限制性的靶向 λ NS5的间隔子序列包括但不限于:GTGATAAGTGAATGCCATG(SEQ ID NO:200)和GNNNNNNNNNNNNNNNNNNNGUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCUUUUUUU(SEQ ID NO:201)。

同时治疗疾病并且防止免疫应答的方法

[0369] 本公开文本提供了用于通过将基因疗法递送至细胞同时治疗受试者的疾病或障碍并且防止对接受所述基因疗法的细胞的免疫应答的组合物和方法。例如,可以将图4中所示的组合物施用至受试者,其中gRNA 1与编码适应性免疫应答的组分的RNA分子内的靶序列结合,并且gRNA2与和疾病或障碍相关的RNA分子内的靶序列结合。通过靶向编码适应性免疫应答的组分的RNA分子,gRNA1防止与所述组合物或包含所述组合物的载体相关的抗原在所述细胞的表面上展示,从而从所述受试者的免疫系统中掩蔽所述细胞。gRNA2同时靶向第二RNA分子以治疗本公开文本的疾病或障碍。

[0370] 在可替代的实施方案中,图4中所示的组合物的gRNA1和gRNA2例如可以各自靶向编码适应性免疫应答的组分的不同的RNA分子。例如,在gRNA1靶向编码 β 2M多肽的RNA分子的时候,gRNA2靶向共刺激分子(ICOSLG、CD80、CD86、OX40L、IL12或CCR7)。

[0371] 在一些实施方案中,本公开文本的组合物可以包含至少1、2、3、4、5、6、7、8、9或10种gRNA或由其组成。

[0372] 在一些实施方案中,本公开文本的组合物可以包含以下或由其组成:至少1、2、3、4、5、6、7、8、9或10种gRNA,所述gRNA的表达是在组成型启动子(例如U6)的控制下;以及包含第一RNA结合蛋白和第二RNA结合蛋白的融合蛋白,所述融合物的表达是在病毒启动子(其可以任选地是组成型的(例如EFS))的控制下。

[0373] 在一些实施方案中,本公开文本的组合物可以包含以下或由其组成:至少1、2、3、4、5、6、7、8、9或10种gRNA,所述gRNA的表达是在第一启动子的控制下;以及包含第一RNA结

合蛋白和第二RNA结合蛋白的融合蛋白,所述融合物的表达是在第二启动子的控制下,其中所述第一启动子驱动至少1、2、3、4、5、6、7、8、9或10种gRNA的更强的表达,并且所述第二启动子驱动所述融合蛋白的表达。在一些实施方案中,本公开文本的组合物可以包含以下或由其组成:至少1、2、3、4、5、6、7、8、9或10种gRNA,所述gRNA的表达是在第一启动子的控制下;以及包含第一RNA结合蛋白和第二RNA结合蛋白的融合蛋白,所述融合物的表达是在第二启动子的控制下,其中所述第一启动子驱动至少1、2、3、4、5、6、7、8、9或10种gRNA的更弱的表达,并且所述第二启动子驱动所述融合蛋白的表达。通过改变驱动本公开文本的组合物的gRNA相比于融合蛋白组分的表达的启动子的相关强度,所述组合物可以以按比率计的剂量提供,同时从同一载体表达所述gRNA和所述融合蛋白。因此,本公开文本的组合物可以包含以下gRNA,其结合与两种或更多种疾病以及适应性免疫应答的两种或更多种组分相关的RNA分子。在一些实施方案中,所述本公开文本的组合物可以包含本文公开的融合蛋白,其中至少一种融合配偶体蛋白是内切核酸酶,如但不限于RNA酶1、RNA酶4、RNA酶6、RNA酶7、RNA酶8、RNA酶2、RNA酶6PL、RNA酶L、RNA酶T2、RNA酶11、RNA酶T2样蛋白、NOB1、ENDOV、ENDOG、ENDOD1、hFEN1、hSLFN14、hLACTB2、APEX2、ANG、HRSP12、ZC3H12A、RIDA、PDL6、NTHL、KIAA0391、APEX1、AGO2、EXO2、ZC3H12D、ERN2、PELO、YBEY、CPSF4L、hCG_2002731、ERCC1、RAC1、RAA1、RAB1、DNA2、FLJ35220、FLJ13173、ERCC4、RNA酶1 (K41R)、RNA酶1 (K41R、D121E)、RNA酶1 (K41R、D121E、H119N)、RNA酶1 (H119N)、RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)、RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E)、RNA酶1 (R39D、N67D、N88A、G89D、R91D)、TENM1、TENM2、RNA酶K、TALEN、ZNF638、或hSMG6的PIN。

使用方法

[0374] 本公开文本提供了修饰本公开文本的RNA分子或由所述RNA分子编码的蛋白质的表达水平的方法,所述方法包括在适合于所述指导RNA或所述融合蛋白(或其部分)中的一种或多种与所述RNA分子结合的条件下使所述组合物与所述RNA分子接触。

[0375] 本公开文本提供了修饰由RNA分子编码的蛋白质的活性的方法,所述方法包括在适合于所述指导RNA或所述融合蛋白(或其部分)中的一种或多种与所述RNA分子结合的条件下使所述组合物与所述RNA分子接触。

[0376] 本公开文本提供了修饰本公开文本的RNA分子或由所述RNA分子编码的蛋白质的表达水平的方法,所述方法包括在适合于所述指导RNA或所述融合蛋白(或其部分)中的一种或多种与所述RNA分子结合的条件下使所述组合物与包含所述RNA分子的细胞接触。在一些实施方案中,所述细胞是体内的、体外的、离体的或原位的。在一些实施方案中,所述组合物包含载体,其包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物。在一些实施方案中,所述载体是AAV。

[0377] 本公开文本提供了修饰由RNA分子编码的蛋白质的活性的方法,所述方法包括在适合于所述指导RNA或所述融合蛋白(或其部分)中的一种或多种与所述RNA分子结合的条件下使所述组合物与包含所述RNA分子的细胞接触。在一些实施方案中,所述细胞是体内的、体外的、离体的或原位的。在一些实施方案中,所述组合物包含载体,其包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物。在一些实施方案中,所述载体是AAV。

[0378] 本公开文本提供了修饰本公开文本的RNA分子或由所述RNA分子编码的蛋白质的表达水平的方法,所述方法包括在适合于RNA核酸酶活性的条件下使所述组合物与所述RNA

分子接触,其中所述融合蛋白诱导所述RNA分子中的断裂。

[0379] 本公开文本提供了修饰由RNA分子编码的蛋白质的活性的方法,所述方法包括在适合于RNA核酸酶活性的条件下使所述组合物与所述RNA分子接触,其中所述融合蛋白诱导所述RNA分子中的断裂。

[0380] 本公开文本提供了修饰本公开文本的RNA分子或由所述RNA分子编码的蛋白质的表达水平的方法,所述方法包括在适合于RNA核酸酶活性的条件下使所述组合物与包含所述RNA分子的细胞接触,其中所述融合蛋白诱导所述RNA分子中的断裂。在一些实施方案中,所述细胞是体内的、体外的、离体的或原位的。在一些实施方案中,所述组合物包含载体,其包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物。在一些实施方案中,所述载体是AAV。

[0381] 本公开文本提供了修饰由RNA分子编码的蛋白质的活性的方法,所述方法包括在适合于RNA核酸酶活性的条件下使所述组合物与包含所述RNA分子的细胞接触,其中所述融合蛋白诱导所述RNA分子中的断裂。在一些实施方案中,所述细胞是体内的、体外的、离体的或原位的。在一些实施方案中,所述组合物包含载体,其包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物。在一些实施方案中,所述载体是AAV。

[0382] 本公开文本提供了治疗疾病或障碍的方法,所述方法包括向受试者施用治疗有效量的本公开文本的组合物。

[0383] 本公开文本提供了治疗疾病或障碍的方法,所述方法包括向受试者施用治疗有效量的本公开文本的组合物,其中所述组合物包含载体,所述载体包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物,并且其中所述组合物修饰本公开文本的RNA分子或由所述RNA分子编码的蛋白质的表达水平。

[0384] 本公开文本提供了治疗疾病或障碍的方法,所述方法包括向受试者施用治疗有效量的本公开文本的组合物,其中所述组合物包含载体,所述载体包含含有本公开文本的指导RNA和本公开文本的融合蛋白的组合物,并且其中所述组合物修饰由RNA分子编码的蛋白质的活性。

[0385] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于遗传疾病或障碍。在一些实施方案中,所述遗传疾病或障碍是单基因疾病或障碍。在一些实施方案中,所述单基因疾病或障碍是常染色体显性疾病或障碍、常染色体隐性疾病或障碍、X染色体连锁(X连锁)疾病或障碍、X连锁显性疾病或障碍、X连锁隐性疾病或障碍、Y连锁疾病或障碍或线粒体疾病或障碍。在一些实施方案中,所述遗传疾病或障碍是多基因疾病或障碍。在一些实施方案中,所述遗传疾病或障碍是多基因疾病或障碍。在一些实施方案中,所述单基因疾病或障碍是常染色体显性疾病或障碍,包括但不限于亨廷顿病(Huntington's disease)、神经纤维瘤病1型、神经纤维瘤病2型、马凡综合征(Marfan syndrome)、遗传性非息肉病性结直肠癌、遗传性多发性外生骨疣、血管性血友病(Von Willebrand disease)和急性间歇性卟啉病。在一些实施方案中,所述单基因疾病或障碍是常染色体隐性疾病或障碍,包括但不限于白化病、中链酰基辅酶A脱氢酶缺乏症、囊性纤维化、镰状细胞病、泰-萨克斯病(Tay-Sachs disease)、尼曼-皮克病(Niemann-Pick disease)、脊髓性肌萎缩和罗伯茨综合征(Roberts syndrome)。在一些实施方案中,所述单基因疾病或障碍是X连锁疾病或障碍,包括但不限于肌营养不良、杜氏肌营养不良

(Duchenne muscular dystrophy)、血友病、肾上腺脑白质营养不良 (ALD)、雷特综合征 (Rett syndrome) 和血友病A。在一些实施方案中,所述单基因疾病或障碍是线粒体障碍,包括但不限于利伯氏遗传性视神经病变 (Leber's hereditary optic neuropathy)。

[0386] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于免疫疾病或障碍。在一些实施方案中,所述免疫疾病或障碍是免疫缺陷疾病或障碍,包括但不限于B细胞缺乏症、T细胞缺乏症、嗜中性粒细胞减少症、无脾、补体缺乏症、获得性免疫缺陷综合征 (AIDS) 和由于医学干预所致的免疫缺陷 (免疫抑制是医学疗法的预期或不利影响)。在一些实施方案中,所述免疫疾病或障碍是自身免疫性疾病或障碍,包括但不限于失弛缓症、艾迪生病 (Addison's disease)、成人斯蒂尔病 (Adult Still's disease)、无丙种球蛋白血症、斑秃、淀粉样变性、抗GBM/抗TBM肾炎、抗磷脂综合征、自身免疫性血管性水肿、自身免疫性自主神经机能异常、自身免疫性脑脊髓炎、自身免疫性肝炎、自身免疫性内耳病 (AIED)、自身免疫性心肌炎、自身免疫性卵巢炎、自身免疫性睾丸炎、自身免疫性胰腺炎、自身免疫性视网膜病变、自身免疫性荨麻疹、轴突和神经元神经病 (AMAN)、巴洛病 (Baló disease)、白塞病 (Behcet's disease)、良性粘膜类天疱疮、大疱性类天疱疮、卡斯尔曼病 (Castleman disease, CD)、乳糜泻、美洲锥虫病 (Chagas disease)、慢性炎症性脱髓鞘性多发性神经病 (CIDP)、慢性复发性多病灶性骨髓炎 (CRMO)、变应性肉芽肿性血管炎 (Churg-Strauss Syndrome, CSS) 或嗜伊红细胞性肉芽肿病 (EGPA)、瘢痕性类天疱疮、寇甘综合征 (Cogan's syndrome)、冷凝集素病、先天性心脏传导阻滞、柯萨奇病毒性心肌炎、CREST综合征、克罗恩病 (Crohn's disease)、疱疹样皮炎、皮炎、德维克病 (Devic's disease) (视神经脊髓炎)、盘状狼疮、德雷斯勒综合征 (Dressler's syndrome)、子宫内膜异位症、嗜酸性食道炎 (EoE)、嗜酸性筋膜炎、结节性红斑、特发性混合性冷球蛋白血症 (Essential mixed cryoglobulinemia)、伊文思综合征 (Evans syndrome)、纤维肌痛、纤维化肺泡炎、巨细胞动脉炎 (颞动脉炎)、巨细胞性心肌炎、肾小球肾炎、肺出血-肾炎综合征 (Goodpasture's syndrome)、肉芽肿性多血管炎、格雷夫斯病 (Graves' disease)、格林-巴利综合征 (Guillain-Barre syndrome)、桥本甲状腺炎 (Hashimoto's thyroiditis)、溶血性贫血、过敏性紫癜 (Henoch-Schonlein purpura, HSP)、妊娠疱疹或妊娠性类天疱疮 (PG)、化脓性汗腺炎 (HS) (反常性痤疮)、低丙球蛋白血症、IgA肾病、IgG4相关性硬化性疾病、免疫性血小板减少性紫癜 (ITP)、包涵体肌炎 (IBM)、间质性膀胱炎 (IC)、幼年型关节炎、幼年型糖尿病 (1型糖尿病)、幼年型肌炎 (JM)、川崎病 (Kawasaki disease)、兰伯特-伊顿综合征 (Lambert-Eaton syndrome)、白细胞破碎性血管炎、扁平苔癣、硬化性苔癣、木样结膜炎、线性IgA疾病 (LAD)、狼疮、慢性莱姆病 (Lyme disease chronic)、梅尼埃病 (Meniere's disease)、显微镜下多血管炎 (MPA)、混合性结缔组织病 (MCTD)、蚕蚀性角膜溃疡 (Mooren's ulcer)、穆-哈二氏病 (Mucha-Habermann disease)、多灶性运动神经病 (MMN) 或MMNCB、多发性硬化症、重症肌无力、肌炎、发作性睡病、新生儿狼疮、视神经脊髓炎、嗜中性粒细胞减少症、眼部瘢痕性类天疱疮、视神经炎、复发性风湿病 (PR)、PANDAS、副肿瘤性小脑变性 (PCD)、阵发性睡眠性血红蛋白尿 (PNH)、帕里-龙贝格综合征 (Parry Romberg syndrome)、睫状体扁平部炎 (周边葡萄膜炎)、帕-特二氏综合征 (Parsonnage-Turner syndrome)、天疱疮、周围神经病变、静脉周围性脑脊髓炎 (Perivenous encephalomyelitis)、恶性贫血 (PA)、POEMS综合征、结节性多动脉炎、多腺体综合征I型、II型、III型、风湿性多肌痛、多发性肌

炎、心肌梗死后综合征、心包切开后综合征、原发性胆汁性肝硬化、原发性硬化性胆管炎、孕酮性皮炎、银屑病、银屑病关节炎、纯红细胞再生障碍 (PRCA)、坏疽性脓皮病、雷诺现象 (Raynaud's phenomenon)、反应性关节炎、反射性交感神经营养不良、复发性多软骨炎、不宁腿综合征 (RLS)、腹膜后纤维化、风湿热、类风湿性关节炎、结节病、施密特综合征 (Schmidt syndrome)、巩膜炎、硬皮病、干燥综合征 (Sjögren's syndrome)、精子和睾丸自身免疫、僵人综合征 (SPS)、亚急性细菌性心内膜炎 (SBE)、苏萨克综合征 (Susac's syndrome)、交感性眼炎 (SO)、大动脉炎 (Takayasu's arteritis)、颞动脉炎/巨细胞动脉炎、血小板减少性紫癜 (TTP)、托洛萨-亨特综合征 (Tolosa-Hunt syndrome, THS)、横贯性脊髓炎、1型糖尿病、溃疡性结肠炎 (UC)、未分化结缔组织病 (UCTD)、葡萄膜炎、血管炎、白癜风、小柳原田病 (Vogt-Koyanagi-Harada Disease) 或韦氏肉芽肿病 (Wegener's granulomatosis)。

[0387] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于炎性疾病或障碍。在一些实施方案中,所述炎性疾病或障碍包括但不限于阿尔茨海默病 (Alzheimer's disease)、强直性脊柱炎、关节炎、骨关节炎、类风湿性关节炎、银屑病性关节炎、哮喘、动脉粥样硬化、克罗恩病、结肠炎、皮炎、憩室炎、纤维肌痛、肝炎、肠易激综合征 (IBS)、系统性红斑狼疮 (SLE)、肾炎、帕金森病 (Parkinson's disease)、溃疡性结肠炎、急性支气管炎、急性阑尾炎、扁桃体炎、感染性脑膜炎、鼻窦炎、哮喘、慢性消化性溃疡、结核病、类风湿性关节炎、牙周炎、痛风、硬皮病、血管炎和肌炎。

[0388] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于代谢性疾病或障碍。在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于退行性或进行性疾病或障碍。在一些实施方案中,所述退行性或进行性疾病或障碍包括但不限于肌萎缩侧索硬化 (ALS)、亨廷顿病、阿尔茨海默病和衰老。

[0389] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于感染性疾病或障碍。

[0390] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于儿科或发育性疾病或障碍。

[0391] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于心血管疾病或障碍。

[0392] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于增生性疾病或障碍。在一些实施方案中,所述增生性疾病或障碍是癌症。在一些实施方案中,所述癌症包括但不限于急性淋巴细胞性白血病 (ALL)、急性髓性白血病 (AML)、肾上腺皮质癌、AIDS相关性癌症、卡波西肉瘤 (Kaposi Sarcoma) (软组织肉瘤)、AIDS相关性淋巴瘤 (淋巴瘤)、原发性CNS淋巴瘤 (淋巴瘤)、肛门癌、阑尾癌、胃肠道类癌瘤、星形细胞瘤、非典型畸胎瘤/横纹肌样瘤、中枢神经系统 (脑瘤)、基底细胞癌、胆管癌、膀胱癌、骨癌、尤因肉瘤 (Ewing Sarcoma)、骨肉瘤、恶性纤维组织细胞瘤、脑瘤、乳腺癌、伯基特淋巴瘤 (Burkitt Lymphoma)、类癌瘤、癌、心脏 (Cardiac/Heart) 肿瘤、胚胎瘤、胚细胞瘤、原发性CNS淋巴瘤、宫颈癌、胆管细胞癌 (Cholangiocarcinoma)、脊索瘤、慢性淋巴细胞白血病 (CLL)、慢性髓细胞性白血病 (CML)、慢性骨髓增殖性肿瘤、结直肠癌、颅咽管瘤、皮肤T细胞淋巴瘤、原位导管

癌、胚胎瘤、子宫内膜癌(子宫癌)、室管膜瘤、食管癌、鼻腔神经胶质瘤(头颈癌)、尤因肉瘤(骨癌)、颅外胚细胞瘤、性腺外胚细胞瘤、眼癌、儿童眼内黑色素瘤、眼内黑色素瘤、视网膜母细胞瘤、输卵管癌、骨恶性纤维组织细胞瘤和骨肉瘤、胆囊癌、胃(Gastric/Stomach)癌、胃肠道类癌瘤、胃肠道间质瘤(GIST)(软组织肉瘤)、儿童胃肠道间质瘤、胚细胞瘤、儿童颅外胚细胞瘤、性腺外胚细胞瘤、卵巢胚细胞瘤、睾丸癌、妊娠滋养细胞疾病、毛细胞白血病、头颈癌、心脏肿瘤、肝细胞(肝)癌、组织细胞增多症、霍奇金淋巴瘤(Hodgkin Lymphoma)、下咽癌(头颈癌)、眼内黑色素瘤、胰岛细胞瘤、胰腺神经内分泌肿瘤、卡波西肉瘤(软组织肉瘤)、肾(肾细胞)癌、朗格汉斯细胞组织细胞增生症(Langerhans Cell Histiocytosis)、喉癌(Laryngeal Cancer)(头颈癌)、白血病、唇和口腔癌(Lip and Oral Cavity Cancer)(头颈癌)、肝癌、肺癌(非小细胞和小细胞)、儿童肺癌、淋巴瘤、男性乳腺癌、骨恶性纤维组织细胞瘤和骨肉瘤、黑色素瘤、梅克尔细胞癌(Merkel Cell Carcinoma)(皮肤癌)、间皮瘤、隐匿性原发性转移性鳞状颈癌(头颈癌)、具有NUT基因变化的中线道癌、口腔癌(Mouth Cancer)(头颈癌)、多发性内分泌肿瘤综合征、多发性骨髓瘤/浆细胞瘤、蕈样真菌病(淋巴瘤)、骨髓增生异常综合征、骨髓增生异常性/骨髓增生性肿瘤、鼻腔和鼻窦癌(头颈癌)、鼻咽癌(头颈癌)、神经母细胞瘤、非霍奇金淋巴瘤、非小细胞肺癌、口腔癌(Oral Cancer)、唇和口腔癌和口咽癌、骨肉瘤和骨恶性纤维组织细胞瘤、卵巢癌、胰腺癌、胰腺神经内分泌肿瘤(胰岛细胞瘤)、乳头状瘤病、副神经节瘤、甲状旁腺癌、阴茎癌、咽癌(头颈癌)、嗜铬细胞瘤、浆细胞瘤/多发性骨髓瘤、胸膜肺母细胞瘤、妊娠期乳腺癌、原发性中枢神经系统(CNS)淋巴瘤、原发性腹膜癌、前列腺癌、直肠癌、复发性癌症、肾细胞(肾)癌、视网膜母细胞瘤、横纹肌肉瘤、儿童(软组织肉瘤)、涎腺癌(头颈癌)、肉瘤、儿童横纹肌肉瘤(软组织肉瘤)、儿童血管瘤(软组织肉瘤)、尤因肉瘤(骨癌)、卡波西肉瘤(软组织肉瘤)、骨肉瘤(骨癌)、子宫肉瘤、塞扎里综合征(Sézary Syndrome)、淋巴瘤、皮肤癌、小细胞肺癌、小肠癌、软组织肉瘤、皮肤鳞状细胞癌、鳞状颈癌、胃(Stomach/Gastric)癌、T细胞淋巴瘤、睾丸癌、喉癌(Throat Cancer)(头颈癌)、鼻咽癌、口咽癌、下咽癌、胸腺瘤和胸腺癌、甲状腺癌、肾盂和输尿管移行细胞癌、肾细胞癌、尿道癌、子宫肉瘤、阴道癌、血管瘤(软组织肉瘤)、外阴癌、肾母细胞瘤(Wilms Tumor)和其他儿童肾脏肿瘤。

[0393] 在本公开文本的方法的一些实施方案中,本公开文本的受试者已经被诊断患有所述疾病或障碍。在一些实施方案中,本公开文本的受试者呈现所述疾病或障碍的至少一种体征或症状。在一些实施方案中,所述受试者具有预示患上所述疾病或障碍的风险的生物标记。在一些实施方案中,所述生物标记是基因突变。

[0394] 在本公开文本的方法的一些实施方案中,本公开文本的受试者是雌性。在本公开文本的方法的一些实施方案中,本公开文本的受试者是雄性。在一些实施方案中,本公开文本的受试者具有两个XX或XY染色体。在一些实施方案中,本公开文本的受试者具有两个XX或XY染色体和第三染色体(X或Y)。

[0395] 在本公开文本的方法的一些实施方案中,本公开文本的受试者是新生儿、婴儿、儿童、成人、年长成人或老年人。在本公开文本的方法的一些实施方案中,本公开文本的受试者为至少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或31日龄。在本公开文本的方法的一些实施方案中,本公开文本的受试者为至少1、2、3、4、5、6、7、8、9、10、11或12月龄。在本公开文本的方法的一些实施方案中,本公开文

本的受试者为至少1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100岁或之间的任何岁数或非整岁数。

[0396] 在本公开文本的方法的一些实施方案中,本公开文本的受试者是哺乳动物。在一些实施方案中,本公开文本的受试者是非人哺乳动物。

[0397] 在本公开文本的方法的一些实施方案中,本公开文本的受试者是人。

[0398] 在本公开文本的方法的一些实施方案中,治疗有效量包含本公开文本的组合物的单个剂量。在一些实施方案中,治疗有效量包含治疗有效量包含本公开文本的组合物的至少一个剂量。在一些实施方案中,治疗有效量包含治疗有效量包含本公开文本的组合物的一个或多个剂量。

[0399] 在本公开文本的方法的一些实施方案中,治疗有效量消除所述疾病或障碍的体征或症状。在一些实施方案中,治疗有效量降低所述疾病或障碍的体征或症状的严重程度。

[0400] 在本公开文本的方法的一些实施方案中,治疗有效量消除所述疾病或障碍。

[0401] 在本公开文本的方法的一些实施方案中,治疗有效量防止疾病或障碍的发作。在一些实施方案中,治疗有效量延迟疾病或障碍的发作。在一些实施方案中,治疗有效量降低所述疾病或障碍的体征或症状的严重程度。在一些实施方案中,治疗有效量改善所述受试者的预后。

[0402] 在本公开文本的方法的一些实施方案中,将本公开文本的组合物全身施用至所述受试者。在一些实施方案中,将本公开文本的组合物通过静脉内途径施用至所述受试者。在一些实施方案中,将本公开文本的组合物通过注射或输注施用至所述受试者。

[0403] 在本公开文本的方法的一些实施方案中,将本公开文本的组合物局部施用至所述受试者。在一些实施方案中,将本公开文本的组合物通过骨内、眼内、脑脊髓内或脊柱内途径施用至所述受试者。在一些实施方案中,将本公开文本的组合物直接施用至中枢神经系统的脑脊液。在一些实施方案中,将本公开文本的组合物直接施用至眼组织或流体,并且在眼结构外不具有生物利用性。在一些实施方案中,将本公开文本的组合物通过注射或输注施用至所述受试者。

[0404] 在一些实施方案中,将包含本文公开的RNA结合融合蛋白的组合物配制为药物组合物。简言之,如本文所公开使用的药物组合物可以包含与一种或多种药学上或生理上可接受的载体、稀释剂或赋形剂组合的一种或多种融合蛋白或编码所述一种或多种融合蛋白的多核苷酸,所述融合蛋白或多核苷酸任选地包含于AAV中,所述AAV任选地也是免疫正交的。此类组合物可以包含缓冲液,如中性缓冲盐水、磷酸盐缓冲盐水等;碳水化合物,如葡萄糖、甘露糖、蔗糖或葡聚糖、甘露醇;蛋白质;多肽或氨基酸,如甘氨酸;抗氧化剂;螯合剂,如EDTA或谷胱甘肽;佐剂(例如,氢氧化铝);和防腐剂。可以将本公开文本的组合物配制用于口服、静脉内、局部、肠内、眼内和/或肠胃外施用。在某些实施方案中,将本公开文本的组合物配制用于静脉内施用。

实施例

实施例1:病毒RNA分子的RNA指导的切割

[0405] 将A549细胞在含有10%FBS和1%青霉素/链霉素的DMEM(GIBCO)中培养,并以90%-100%汇合度进行传代。将细胞以 1×10^5 个细胞/孔接种于24孔板中用于RNA分离,或以 $.5 \times 10^5$ 个细胞/孔接种。将细胞用与基因NTHL1(残基31-312,E43)或CPSF4L(全长,E67)

融合的编码空肠弯曲杆菌Cas9 (CjeCas9) 的质粒与编码Zika NS5 RNA中的四个位点之一的质粒转染。CjeCas9是由EFS启动子驱动的,而指导RNA是由U6启动子驱动的。sgRNA的序列呈现于表8中。下文呈现了在此项研究中使用的构建体的序列(SEQ ID NO:13656和SEQ ID NO:13657)。

[0406] RNA分离是用RNAeasy柱(Qiagen)根据制造商的方案来进行。使用Nanodrop分光光度计来估计RNA质量和浓度。使用Superscript III (Thermo) 以随机引物根据制造商的方案进行cDNA制备。用如表7中所列的以下引物进行qPCR。

[0407] 图1显示了在具有含有如表8中指示的各种靶向NS5的间隔子序列的sgRNA的E43和E67内切核酸酶二者的存在下评估的Zika NS5的表达水平。将Zika NS5表达展示为相对于加载有含有对照(λ)间隔子序列的sgRNA的内切核酸酶的变化倍数。

[0408] 使用免疫荧光显微镜检查将在与CjeCas9融合的E43或E67内切核酸酶的存在下的Zika NS5表达可视化。图2A显示了用加载有含有靶向Zika NS5的间隔子序列的sgRNA的CjeCas9-内切核酸酶融合物转染的细胞的荧光显微镜检查图像。如与加载有不靶向Zika NS5的sgRNA的CjeCas9-内切核酸酶融合物相比,在加载有靶向Zika NS5的适当sgRNA的CjeCas9-内切核酸酶融合物的存在下,Zika NS5的表达显著降低(图2A和图2B)。图3是用于本公开文本的组合物中的示例性内切核酸酶的列表。

[0409] 表7:qPCR引物

GAPDH_F	CAGCCTCAAGATCATCAGCAA (SEQ ID NO:192)
GAPDH_R	TGTGGTCATGAGTCCTTCCA (SEQ ID NO:193)
NS5_F	GAGGAGAGTGCCAGAGTTGT (SEQ ID NO:194)
NS5_R	TCTCTCTCCCCATCCAGTGA (SEQ ID NO:195)

[0410] 表8:sgRNA序列

靶向NS5的间隔子1	gcaatgatcttcattgtgggagc (SEQ ID NO: 196)
靶向NS5的间隔子2	gaacctgttgatgaactcttc (SEQ ID NO: 197)
靶向NS5的间隔子3	gttggtgattagagcttcattc (SEQ ID NO: 198)
靶向NS5的间隔子4	gagtgatcctcgttcaagaatcc (SEQ ID NO: 199)
非靶向对照间隔子 (λ 2)	GTGATAAGTGGAATGCCATG (SEQ ID NO: 200)
sgRNA支架 (N指示间隔子)	GNNNNNNNNNNNNNNNNNNNNNGUUUAAGAGCUAUG CUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUUUUU (SEQ ID NO: 201)

[0411] E43-CjeCas9和sgRNA质粒可以包含以下序列或由其组成 (U6:N=sgRNA间隔子, E43,CjeCas9) :

gtttattacagggacagcagagatccagtttggttaattaaggtaccgagggcctatttcccatgatt
ccttcataatttgcataatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatat

tagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatgga
ctatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttGTGGAAGGACGAAACACC
NNNNNNNNNNNNNNNNNGTTTTAGTCCCTGAAGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGGTACAA
TCCCCTAAAACCGCTTTTTTTCCTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTC
CAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAGTGGGAATTGGCTC
CGGTGCCCCTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGTTCGGCAATTGATCC
GGTGCCTAGAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTTCCCGAGGGTG
GGGGAGAACCGTATATAAGTGCAAGTTCGCGTGAACGTTCTTTTTTCGCAACGGGTTGCCGCCAGAACACAGGA
CCGTTCTAGAGCGCTATTTAGAACCatgTGTCTCCCCAAGAATCTGGCATGACCGCTCTTCAGCGAGGATGTT
GACGCGAAGCAGATCCCTGGGACCTGGGGCCGGGCCACGAGGGTGTGCGGAAGAACCAGGACCGTTGCGACGGAGG
GAAGCAGCAGCGGAAGCTCGGAAATCCCATTCTCCGTTAAACGACCCCGCAAGGCACAACGGCTCAGGGTTGCTT
ACGAGGGGAGCGATTCCGAAAAGGGTGAAGGAGCAGAGCCCTTGAAGGTTCCAGTATGGGAACCCAGGATTGGCA
GCAGCAGCTTGTAACATCCGAGCAATGAGGAACAAAAAGATGCACCTGTTGATCACCTCGGAACCGAACATTGT
TATGATTCTAGTGCGCCGCCAAAAGTCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGG
ACCAGGTTACGGCCGGAGCAATGCAACGGCTTCGGGCACGGGGACTCACGGTCGATAGCATTTCGAGACCGATGA
CGCAACATTGGGTAAACTCATATATCCAGTTGGCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCC
ATTCTCCAACAACATTACGGAGGTGATATACCCGCAAGCGTAGCTGAACTGGTAGCACTGCCGGGCGTCGGTCCCA
AAATGGCACATCTGGCTATGGCGGTTGCTTGGGGAACGGTGTCTGGTATCGCAGTTGATACGCATGTCCACCGCAT
CGCCAATCGGCTGAGGTGGACTAAAAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGAGTGGCTT
CCACGAGAGCTGTGGCATGAAATCAATGGATTGCTGGTTGGTTTCGGGCAGCAGACATGCTTGCCCGTGACCCCC
GGTGTCTATGCTTGCTTGAACCAGGCTTTGTGCCAGCTGCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATC
TGAGTCTGCGACCCCGGAGAGCacaaacGCGCGAATCCTGGCCTTCGcgATTGGCATTAGCAGCATCGGCTGGGCA
TTCTCTGAAAACGACGAACTGAAGATTGCGGCGTGCGAATTTTCACTAAGGTGCGAAAATCCCAAACTGGTGAAT
CACTCGCTCTCCCTAGACGACTGGCACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCATCT
TAAACACCTTATTGCAAATGAGTTTAACTGAATTATGAGGACTACCAATCCTTTGACGAGTCTCTTGCTAAAGCC
TACAAAGGGAGCCTTATATCCCCGTATGAGCTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTTG
CTCGCGTGATTCTCCACATAGCGAAGAGGCGAGGATACGATGACATTAACAAACAGTGATGATAAGGAAAAAGGGGC
CATACTCAAAGCGATTAAGCAAAATGAAGAGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAG
TACTTCCAGAAGTTCAAAGAAAATAGCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGAAAGATGTA
TTGCGCAATCTTCTCAAGGACGAGCTCAAATTGATTTTCAAGAAACAAAGGGAATTTGGGTTTCAGCTTCTCAAA
AAAATTTGAGGAAGAGGTCTGAGCGTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAAT
TGTAGTTTCTTACCGATGAAAAACGGGCGCCAAAAATAGCCCTTTGGCTTTTATGTTTGTGCTCTGACTCGCA
TCATTAATCTGCTCAACAACCTTAAAAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAA
CGAAGTTTGAAGAACGGGACTTTGACCTACAAACAAACCAAAAAGCTTCTTGGTCTCAGTGATGACTACGAATTC
AAGGGAGAAAAAGGGACATATTTATCGAATTCAAGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACT
TGTCTCAAGATGATCTCAACGAAATTGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGCGTT
GGCGAAGTATGACCTTAACCAAAACCAAAATAGATAGCCTCAGCAAGTTGGAATTTAAAGATCACTTGAATATAAGT
TTCAAGGCCCTTAAGTTGGTCACCCCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATC
TCAAGGTTGCTATTAACGAAGACAAAAAAGATTTCTCCAGCTTTCAATGAGACTTACTATAAGGACGAGGTTAC

CAATCCTGTGGTGCTCCGAGCCATCAAAGAGTATCGAAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTA
CACAAAATAAATATTGAGCTCGCAAGGGAGGTCGTAAGAACCACTCCCAGCGCGCCAAAATAGAAAAGGAACAGA
ATGAAAATTACAAAGCGAAAAAGGACGCCGAGCTCGAGTGCGAAAAGCTGGGCCTGAAAATAAACAGCAAGAACAT
TCTCAAACCTCCGCCTCTTCAAAGAACAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAATCTCCGATCTT
CAAGACGAGAAGATGCTCGAAATAGACgcgATATATCCATATAGCAGGTCTTTTGACGATTCTTACATGAATAAAG
TGCTTGTTTTCACTAAGCAGAATCAGGAAAAGTTGAATCAGACCCCCCTTTGAGGCCTTTGGCAACGACTCAGCAAA
GTGGCAGAAGATCGAGGTCTTGGCTAAGAATCTTCTACTAAGAAACAGAAAAGGATATTGGATAAGAACTATAAA
GACAAAGAACAAAAGAACTTTAAAGACCGCAACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACA
CAAAAGATTATTTGGACTTCTTGCCGCTGTCTGATGATGAGAACACGAAAACCAACGACACGCAAAAGGGGTCTAA
AGTCCACGTGCAAGCTAAATCTGGGATGCTCACCTCAGCATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGA
AACAATCACCTGCACCATGCCATTGACGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGTTTAGCG
ACTTCAAGAAGGAACAAGAGTCCAACAGCGCCGAGCTCTACGCAAAAAAGATTAGTGAACTCGACTACAAAAACAA
AAGAAAATTCTTTGAGCCGTTTCAGCGATTTCGACAGAAGGTATTGGATAAAAATAGATGAAATTTTCGTGAGCAAA
CCCGAAAGGAAAAAGCCCTCAGGCGCCTTGACGAAGAGACTTTTCAGGAAGGAAGAGGAATTCTACCAAAGCTACG
GCGGAAAAGAGGGAGTTTGAAGGCTCTCGAAGTTGGAAGATTAGGAAGGTGAACGGCAAGATAGTGA AAAACGG
CGATATGTTCCGGGTTGATATCTTCAAACATAAAAAACGAATAAATTTTATGCTGTGCCTATATACACTATGGAC
TTCGCACTTAAGGTCCTGCCGAATAAGGCGGTAGCCCGATCTAAAAAAGGCGAAATTAAGGACTGGATTTTGATGG
ATGAAAATTACGAGTTCTGCTTTTCTCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACC
GGAATTTCGTGTATTACAACGCTTTTACTTCCTCTACGGTATCTTTGATTGTCTCCAAACATGACAACAAATTCGAA
ACACTCAGTAAAAACCAAAAGATTCTCTTTAAAAATGCGAACGAGAAAGAAGTAATTGCAAAATCAATTGGCATCC
AAAATTTGAAAGTTTTTGA AAAATATATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACAGCGAGA
GGACTTCAAAAAATCAGGTCCACCCAAGAAAAACGCAAGGTGGAAGATCCGAAGAAAAAGCGAAAAGTGGATGTG
taaCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAAC
TTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGCATTTTTTTTCAC
TGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGTATACCG (SEQ ID NO:202)。

[0412] E67-CjeCas9和sgRNA质粒可以包含以下序列或由其组成(U6:N=sgRNA间隔子, E67,CjeCas9):

gtttattacaggacagcagagatccagtttggttaattaaggtaccgagggcctatttcccatgatt
ccttcatatttgcataatacagatacaaggtggttagagagataattagaattaatttgactgtaaacacaaagatat
tagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaattatgttttaaaatgga
ctatcatatgcttaccgtaacttgaaagtatttgcagtttcttggctttatatatcttGTGGAAGGACGAAACACC
NNNNNNNNNNNNNNNNNGTTTAGTCCCTGAAGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGGTTACAA
TCCCTAAAACCGCTTTTTTCTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTC
CAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAGTGGAATTGGCTC
CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGATCC
GGTGCTAGAGAAGGTGGCGCGGGTAACTGGGAAAGTGATGTCTGTACTGGCTCCGCCTTTTTCCCGAGGGTG
GGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTTCGAACGGGTTTCCCGCCAGAACACAGGA
CCGTTCTAGAGCGCTATTTAGAACCatgCAGGAGGTAATAGCGGGCTTGAGCGATTACCTTTGCCTTCGAAAA
AGACGTAGAGATGCAGAAGGGAACCGGCCTGCTCCATTTCAAGGTATGGACAAATCAGCATCTGCCGTGTGAAT

TTTTTCACCAAGGGTCTGTGTGAAAAGGGGAAGCTCTGTCCATTTCCGCATGATCGCGGAGAGAAGATGGTGGTGT
GTAAGCACTGGCTGAGAGGGCTTTGCAAAAAAGGCGACCACTGCAAATTTCTTCACCAATATGACCTGACTCGAAT
GCCTGAGTGTTATTTTTACAGTAAGTTCGGTGACTGTAGCAACAAAGAATGCAGCTTCTTGCAATGTCAAACCAGCA
TTCAAGTCACAGGATTGCCCGTGGTACGATCAGGGTTTTTGCAAGGACGGTCCCCTCTGCAAATATCGACACGTAC
CCAGAATTATGTGCCCTAATTACCTGGTCGGCTTCTGTCTGAAGGGCCAAAAATGTCAGTTTGCTCAAAAAATTCTG
CGAGTTCAAATTGCTCCCTGGGTCTAAAATTTGGGAACCCAGGATTGGCAGCAGCAGCTTGTAACATCCGAGCA
ATGAGGAACAAAAAGATGCACCTGTTGATCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAAAAG
TCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGGACCAGGTTACGGCCGGAGCAATGCA
ACGGCTTCGGGCACGGGGACTCACGGTCGATAGCATTTTGCAGACCGATGACGCAACATTGGGTAAACTCATATAT
CCAGTTGGCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAACAACATTACGGAGGTG
ATATACCCGCAAGCGTAGCTGAAGTGGTAGCTGCGGGCGTCGGTCCCCAAATGGCACATCTGGCTATGGCGGT
TGCTTGGGGAACGGTGTCTGGTATCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGGTGGACTAAA
AAAGCCACTAAGTCTCTGAAGAAACACGGGTGCTCTGGAAGAGTGGCTTCCACGAGAGCTGTGGCATGAAATCA
ATGGATTGCTGGTTGGTTTCGGGCAGCAGACATGCTTGCCCGTGCACCCCCGGTGTCTGCTTGCTTGAACCAGGC
TTTGTGCCAGCTGCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATCTGAGTCTGCGACCCCGGAGAGCaca
aacGCGGAATCCTGGCCTTCGcgATTGGCATTAGCAGCATCGGCTGGGCATTCTCTGAAAACGACGAAGTGAAGG
ATTGCGGCGTGCGAATTTTCACTAAGGTCGAAAATCCCCAAACTGGTGAATCACTCGCTCTCCCTAGACGACTGGC
ACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCATCTTAAACACCTTATTGCAAATGAGTTT
AAACTGAATTATGAGGACTACCAATCCTTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCCCGT
ATGAGCTCCGGTTCAGAGCACTCAACGAAGTCTGTCCAAACAGGATTTTGCTCGCGTGATTCTCCACATAGCGAA
GAGGCGAGGATACGATGACATTAACAAACAGTGATGATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAGCAAAAT
GAAGAGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCCAGAAGTTCAAAGAAAATA
GCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGAAAGATGTATTGCGCAATCTTTCCTCAAGGACGA
GCTCAAATTGATTTTCAAGAAACAAAGGGAATTTGGGTTTCAGCTTCTCAAAAAAATTTGAGGAAGAGGTTCTGAGC
GTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAATTGTAGTTTCTTACCGATGAAAAAC
GGGCGCCAAAAAATAGCCCTTTGGCTTTTATGTTTGTGCTCTGACTCGCATCATTAATCTGCTCAACAACCTTAA
AAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAACGAAGTTTGAAGAACGGGACTTTG
ACCTACAAACAAACCAAAAAAGCTTCTTGGTCTCAGTGATGACTACGAATTCAGGGAGAAAAAGGGACATATTTCA
TCGAATTCAAGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACTTGCTCAAGATGATCTCAACGAAAT
TGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGCGTTGGCGAAGTATGACCTTAACCAAAAC
CAAATAGATAGCCTCAGCAAGTTGGAATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGTCACCC
CCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATCTCAAGGTTGCTATTAACGAAGACAA
AAAAGATTTCTCCCAGCTTTCAATGAGACTTACTATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCATC
AAAGAGTATCGAAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTACACAAAAATAAATATTGAGCTCGCAA
GGGAGGTGCGTAAGAACCACTCCCAGCGCGCCAAAAATAGAAAAGGAACAGAATGAAAATTACAAAGCGAAAAAGGA
CGCCGAGCTCGAGTGCGAAAAGCTGGGCCTGAAAATAAACAGCAAGAACATTCTCAAACCTCCGCTTCTCAAAGAA
CAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAAATCTCCGATCTTCAAGACGAGAAGATGCTCGAAATAG
ACcgATATATCCATATAGCAGGTCTTTTGACGATTCTTACATGAATAAAGTGCTTGTTCCTACTAAGCAGAATCA
GGAAAAGTTGAATCAGACCCCTTTGAGGCCTTTGGCAACGACTCAGCAAAGTGGCAGAAGATCGAGGTCTTGGCT

AAGAATCTTCTACTAAGAAACAGAAAAGGATATTGGATAAGAAGCTATAAAGACAAAGAACAAAAGAACTTTAAAG
ACCGCAACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACACAAAAGATTATTTGGACTTCTTGCC
GCTGTCTGATGATGAGAACACGAACTCAACGACACGCAAAAGGGGTCTAAAGTCCACGTCTGAAGCTAAATCTGGG
ATGCTCACCTCAGCATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACCATGCCATTG
ACGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGTTTAGCGACTTCAAGAAGGAACAAGAGTCCAA
CAGCGCCGAGCTCTACGCAAAAAAGATTAGTGAAGCTCGACTACAAAAACAAAAGAAAATTCTTTGAGCCGTTTCAGC
GGATTTTCGACAGAAGGTATTGGATAAAATAGATGAAATTTTCGTGAGCAAAACCCGAAAGGAAAAAGCCCTCAGGCG
CCTTGCACGAAGAGACTTTCAGGAAGGAAGAGGAATTCTACCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAGGC
TCTCGAAGTTGAAAAGATTAGGAAGGTGAACGGCAAGATAGTAAAAACGGCGATATGTTCCGGGTTGATATCTTC
AAACATAAAAAACGAATAAAATTTTATGCTGTGCCTATATACACTATGGACTTCGCACTTAAGGTCCTGCCGAATA
AGGCGGTAGCCCGATCTAAAAAAGGCGAAATTAAGGACTGGATTTTGATGGATGAAAATTACGAGTTCTGCTTTTC
TCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACCGGAATTCGTGTATTACAACGCTTTT
ACTTCCTCTACGGTATCTTTGATTGTCTCCAAACATGACAACAAATTGCAAACTCAGTAAAAACCAAAAGATTC
TCTTTAAAAATGCGAACGAGAAAGAAGTAATTGCAAAATCAATTGGCATCCAAAATTTGAAAGTTTTTGAAAAATA
TATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACAGCGAGAGGACTTCAAAAAATCAGGTCCACCC
AAGAAAAACGCAAGGTGGAAGATCCGAAGAAAAAGCGAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCGGCTGG
ATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA
ACTCATCAATGTATCTTATCATGTCTGTATACCG (SEQ ID NO:203)。

示例实施方案

实施方案1.一种组合物,其包含:

(a) 包含特异性结合RNA分子内的靶序列的第一指导RNA (gRNA) 的第一序列,其中所述靶序列包含编码适应性免疫应答的组分的序列;以及

(b) 编码融合蛋白的序列,所述序列包含编码第一RNA结合多肽的序列和编码第二RNA结合多肽的序列,

其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,

其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且

其中所述第二RNA结合多肽包含RNA-核酸酶活性。

实施方案2.一种组合物,其包含:(a) 包含特异性结合第一RNA分子内的第一靶序列的第一指导RNA (gRNA) 的第一序列,其中所述第一靶序列包含编码适应性免疫应答的组分的序列;和

(b) 包含特异性结合第二RNA分子内的第二靶序列的第二指导RNA (gRNA) 的第二序列;以及

(c) 编码融合蛋白的序列,所述序列包含编码第一RNA结合多肽的序列和编码第二RNA结合多肽的序列,

其中所述第一RNA结合多肽和所述第二RNA结合多肽都不包含显著DNA-核酸酶活性,

其中所述第一RNA结合多肽与所述第二RNA结合多肽不相同,并且

其中所述第二RNA结合多肽包含RNA-核酸酶活性。

实施方案3.根据实施方案2所述的组合物,其中所述第一靶序列或所述第二靶序列包含至少一个重复的序列。

实施方案4.根据实施方案2所述的组合物,其中包含所述第一gRNA的第一序列还包含能够在真核细胞中表达所述gRNA的第一启动子,和/或包含所述第二gRNA的第二序列还包含能够在真核细胞中表达所述gRNA的第二启动子。

实施方案5.根据实施方案2所述的组合物,其中包含含有所述第一gRNA的第一序列和含有所述第二gRNA的第二序列的序列包含能够在真核细胞中表达所述第一gRNA和所述第二gRNA的启动子。

实施方案6.根据实施方案4所述的组合物,其中所述第一启动子和所述第二启动子是相同的。

实施方案7.根据实施方案4所述的组合物,其中所述第一启动子和所述第二启动子是不同的。

实施方案8.根据实施方案4-7中任一项所述的组合物,其中所述真核细胞是动物细胞。

实施方案9.根据实施方案8所述的组合物,其中所述动物细胞是哺乳动物细胞。

实施方案10.根据实施方案9所述的组合物,其中所述动物细胞是人细胞。

实施方案11.根据实施方案5-10中任一项所述的组合物,其中所述启动子是组成型活性启动子。

实施方案12.根据实施方案5-11中任一项所述的组合物,其中所述启动子包含从能够驱动RNA聚合酶的表达的启动子分离或衍生的序列。

实施方案13.根据实施方案12所述的组合物,其中所述启动子包含从U6启动子分离或衍生的序列。

实施方案14.根据实施方案5-12中任一项所述的组合物,其中所述启动子包含从能够驱动转移RNA(tRNA)的表达的启动子分离或衍生的序列。

实施方案15.根据实施方案14所述的组合物,其中所述启动子包含从以下启动子分离或衍生的序列:丙氨酸tRNA启动子、精氨酸tRNA启动子、天冬酰胺tRNA启动子、天冬氨酸tRNA启动子、半胱氨酸tRNA启动子、谷氨酰胺tRNA启动子、谷氨酸tRNA启动子、甘氨酸tRNA启动子、组氨酸tRNA启动子、异亮氨酸tRNA启动子、亮氨酸tRNA启动子、赖氨酸tRNA启动子、甲硫氨酸tRNA启动子、苯丙氨酸tRNA启动子、脯氨酸tRNA启动子、丝氨酸tRNA启动子、苏氨酸tRNA启动子、色氨酸tRNA启动子、酪氨酸tRNA启动子或缬氨酸tRNA启动子。

实施方案16.根据实施方案14所述的组合物,其中所述启动子包含从缬氨酸tRNA启动子分离或衍生的序列。

实施方案17.根据实施方案2-16中任一项所述的组合物,其中包含所述第一gRNA的序列还包含与所述第一靶RNA序列特异性结合的第一间隔子序列。

实施方案18.根据实施方案17所述的组合物,其中所述第一间隔子序列与所述第一靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。

实施方案19.根据实施方案17所述的组合物,其中所述第一间隔子序列与所述靶

RNA序列具有100%互补性。

实施方案20.根据实施方案17-19中任一项所述的组合物,其中所述第一间隔子序列包含20个核苷酸或由其组成。

实施方案21.根据实施方案17-19中任一项所述的组合物,其中所述第一间隔子序列包含21个核苷酸或由其组成。

实施方案22.根据实施方案21所述的组合物,其中所述第一间隔子序列包含编码 β -2-微球蛋白(β 2M)蛋白的氨基酸序列的20个核苷酸或由其组成。

实施方案23.根据实施方案22所述的组合物,其中所述第一间隔子序列包含以下的氨基酸序列的20个核苷酸或由其组成:

MSRSVALAVL ALLSLSGLEA IQRTPKIQVY SRHPADIEVD LLKNGERIEK
VEHSDLSFSK DWSFYLLYYT EFTPTEKDEY ACRVNHVTLS QPKIVKWDRD
M (SEQ ID NO: 88)。

实施方案24.根据实施方案2-23中任一项所述的组合物,其中包含所述第一gRNA的序列还包含与所述第一RNA结合蛋白特异性结合的第一支架序列。

实施方案25.根据实施方案24所述的组合物,其中所述第一支架序列包含茎环结构。

实施方案26.根据实施方案24或25所述的组合物,其中所述支架序列包含90个核苷酸或由其组成。

实施方案27.根据实施方案24或25所述的组合物,其中所述支架序列包含93个核苷酸或由其组成。

实施方案28.根据实施方案27所述的组合物,其中所述支架序列包含以下序列:

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA
GUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:12) 或GUUUUAGAGCUAGAAAUAAGCAAGUUAAAUAAGGC
UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:13)。

实施方案29.根据实施方案1-28中任一项所述的组合物,其中包含所述第二gRNA的序列还包含与所述第二靶RNA序列特异性结合的第二间隔子序列。

实施方案30.根据实施方案29所述的组合物,其中所述第二间隔子序列与所述第一靶RNA序列具有至少50%、55%、60%、65%、70%、75%、80%、87%、90%、95%、97%、99%或之间的任何百分比的互补性。

实施方案31.根据实施方案29所述的组合物,其中所述第二间隔子序列与所述靶RNA序列具有100%互补性。

实施方案32.根据实施方案29-31中任一项所述的组合物,其中所述第二间隔子序列包含20个核苷酸或由其组成。

实施方案33.根据实施方案29-31中任一项所述的组合物,其中所述第二间隔子序列包含21个核苷酸或由其组成。

实施方案34.根据实施方案2-34中任一项所述的组合物,其中所述第二间隔子序列包含或还包含含有以下序列的至少1、2、3、4、5、6或7个重复的序列:CUG (SEQ ID NO:18)、CCUG (SEQ ID NO:19)、CAG (SEQ ID NO:80)、GGGGCC (SEQ ID NO:81)或其任何组合。

实施方案35.根据实施方案2-34中任一项所述的组合物,其中包含所述第二gRNA的序列还包含与所述第一RNA结合蛋白特异性结合的第二支架序列。

实施方案36.根据实施方案35所述的组合物,其中所述第二支架序列包含茎环结构。

实施方案37.根据实施方案35或36所述的组合物,其中所述第二支架序列包含85个核苷酸或由其组成。

实施方案38.根据实施方案37所述的组合物,其中所述第二支架序列包含以下序列:

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA
GUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:12) 或GUUUUAGAGCUAGAAAUAGCAAGUUAAAUAAGGC
UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:13)。

实施方案39.根据实施方案1所述的组合物,其中所述gRNA不结合或不选择性结合所述RNA分子内的第二序列。

实施方案40.根据实施方案2-38中任一项所述的组合物,其中所述第一gRNA不结合或不选择性结合所述第一RNA分子内的第二序列。

实施方案41.根据实施方案2-38中任一项所述的组合物,其中所述第二gRNA不结合或不选择性结合所述第二RNA分子内的第二序列。

实施方案42.根据实施方案39所述的组合物,其中RNA基因组或RNA转录组包含所述RNA分子。

实施方案43.根据实施方案40或41所述的组合物,其中RNA基因组或RNA转录组包含所述第一RNA分子或所述第二RNA分子。

实施方案44.根据实施方案1-43中任一项所述的组合物,其中所述第一RNA结合蛋白包含CRISPR-Cas蛋白。

实施方案45.根据实施方案44所述的组合物,其中所述CRISPR-Cas蛋白是II型CRISPR-Cas蛋白。

实施方案46.根据实施方案45所述的组合物,其中所述第一RNA结合蛋白包含Cas9多肽或其RNA结合部分。

实施方案47.根据实施方案44所述的组合物,其中所述CRISPR-Cas蛋白是V型CRISPR-Cas蛋白。

实施方案48.根据实施方案47所述的组合物,其中所述第一RNA结合蛋白包含Cpf1多肽或其RNA结合部分。

实施方案49.根据实施方案44所述的组合物,其中所述CRISPR-Cas蛋白是VI型CRISPR-Cas蛋白。

实施方案50.根据实施方案49所述的组合物,其中所述第一RNA结合蛋白包含Cas13多肽或其RNA结合部分。

实施方案51.根据实施方案44-50中任一项所述的组合物,其中所述CRISPR-Cas蛋白包含天然RNA核酸酶活性。

实施方案52.根据实施方案51所述的组合物,其中所述天然RNA核酸酶活性被降低或抑制。

实施方案53.根据实施方案52所述的组合物,其中所述天然RNA核酸酶活性被增加或诱导。

实施方案54.根据实施方案44-53中任一项所述的组合物,其中所述CRISPR-Cas蛋白包含天然DNA核酸酶活性,并且其中所述天然DNA核酸酶活性被抑制。

实施方案55.根据实施方案54所述的组合物,其中所述CRISPR-Cas蛋白包含突变。

实施方案56.根据实施方案54或55所述的组合物,其中所述CRISPR-Cas蛋白的核酸酶结构域包含所述突变。

实施方案57.根据实施方案54-56中任一项所述的组合物,其中所述突变发生在编码所述CRISPR-Cas蛋白的核酸中。

实施方案58.根据实施方案54-56中任一项所述的组合物,其中所述突变发生在编码所述CRISPR-Cas蛋白的氨基酸中。

实施方案59.根据实施方案54-58中任一项所述的组合物,其中所述突变包含取代、插入、缺失、移码、倒位或转座。

实施方案60.根据实施方案59所述的组合物,其中所述突变包含核酸酶结构域、所述核酸酶结构域内的结合位点、所述核酸酶结构域内的活性位点或所述核酸酶结构域内的至少一个必需氨基酸残基的缺失。

实施方案61.根据实施方案1-43中任一项所述的组合物,其中所述第一RNA结合蛋白包含Pumilio和FBF (PUF) 蛋白。

实施方案62.根据实施方案61所述的组合物,其中所述第一RNA结合蛋白包含基于Pumilio的联合体 (PUMBY) 蛋白。

实施方案63.根据实施方案1-56中任一项所述的组合物,其中所述第一RNA结合蛋白不需要多聚化以用于RNA结合活性。

实施方案64.根据实施方案63所述的组合物,其中所述第一RNA结合蛋白不是多聚体复合物的单体。

实施方案65.根据实施方案63所述的组合物,其中多聚体蛋白复合物不包含所述第一RNA结合蛋白。

实施方案66.根据实施方案1-65中任一项所述的组合物,其中所述第一RNA结合蛋白与所述RNA分子内的靶序列选择性结合。

实施方案67.根据实施方案66所述的组合物,其中所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的亲和力。

实施方案68.根据实施方案66或67所述的组合物,其中所述第一RNA结合蛋白不包含对所述RNA分子内的第二序列的高亲和力或不选择性结合所述第二序列。

实施方案69.根据实施方案68所述的组合物,其中RNA基因组或RNA转录组包含所述RNA分子。

实施方案70.根据实施方案1-69中任一项所述的组合物,其中所述第一RNA结合蛋白包含在2个与1300个之间的氨基酸,包括端点。

实施方案71.根据实施方案1-70中任一项所述的组合物,其中编码所述第一RNA结合蛋白的序列还包含核定位信号 (NLS) 。

实施方案72.根据实施方案71所述的组合物,其中编码核定位信号 (NLS) 的所述序

列定位于编码所述第一RNA结合蛋白的序列的3'。

实施方案73.根据实施方案72所述的组合物,其中所述第一RNA结合蛋白包含在所述蛋白质的C末端的NLS。

实施方案74.根据实施方案1-70中任一项所述的组合物,其中编码所述第一RNA结合蛋白的序列还包含编码第一NLS的第一序列和编码第二NLS的第二序列。

实施方案75.根据实施方案74所述的组合物,其中编码所述第一NLS或所述第二NLS的序列定位于编码所述第一RNA结合蛋白的序列的3'。

实施方案76.根据实施方案75所述的组合物,其中所述第一RNA结合蛋白包含在所述蛋白质的C末端的所述第一NLS或所述第二NLS。

实施方案77.根据实施方案1-76中任一项所述的组合物,其中所述第二RNA结合蛋白包含核酸酶结构域或由其组成。

实施方案78.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含RNA酶或由其组成。

实施方案79.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶1或由其组成。

实施方案80.根据实施方案79所述的组合物,其中所述RNA酶1蛋白包含SEQ ID NO:20或由其组成。

实施方案81.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶4或由其组成。

实施方案82.根据实施方案81所述的组合物,其中所述RNA酶4蛋白包含SEQ ID NO:21或由其组成。

实施方案83.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶6或由其组成。

实施方案84.根据实施方案83所述的组合物,其中所述RNA酶6蛋白包含SEQ ID NO:22或由其组成。

实施方案85.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶7或由其组成。

实施方案86.根据实施方案85所述的组合物,其中所述RNA酶7蛋白包含SEQ ID NO:23或由其组成。

实施方案87.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶8或由其组成。

实施方案88.根据实施方案87所述的组合物,其中所述RNA酶8蛋白包含SEQ ID NO:24或由其组成。

实施方案89.根据实施方案88所述的组合物,其中所述第二RNA结合蛋白包含RNA酶2或由其组成。

实施方案90.根据实施方案89所述的组合物,其中所述RNA酶2蛋白包含SEQ ID NO:25或由其组成。

实施方案91.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶6PL或由其组成。

实施方案92.根据实施方案91所述的组合物,其中所述RNA酶6PL蛋白包含SEQ ID NO:26或由其组成。

实施方案93.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶L或由其组成。

实施方案94.根据实施方案93所述的组合物,其中所述RNA酶L蛋白包含SEQ ID NO:27或由其组成。

实施方案95.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶T2或由其组成。

实施方案96.根据实施方案95所述的组合物,其中所述RNA酶T2蛋白包含SEQ ID NO:28或由其组成。

实施方案97.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶11或由其组成。

实施方案98.根据实施方案97所述的组合物,其中所述RNA酶11蛋白包含SEQ ID NO:29或由其组成。

实施方案99.根据实施方案78所述的组合物,其中所述第二RNA结合蛋白包含RNA酶T2样蛋白或由其组成。

实施方案100.根据实施方案99所述的组合物,其中所述RNA酶T2样蛋白包含SEQ ID NO:30或由其组成。

实施方案101.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含NOB1多肽或由其组成。

实施方案102.根据实施方案101所述的组合物,其中所述NOB1多肽包含SEQ ID NO:31或由其组成。

实施方案103.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含内切核酸酶或由其组成。

实施方案104.根据实施方案103所述的组合物,其中所述第二RNA结合蛋白包含内切核酸酶V (ENDO V) 多肽或由其组成。

实施方案105.根据实施方案104所述的组合物,其中所述ENDO V蛋白包含SEQ ID NO:32或由其组成。

实施方案106.根据实施方案103所述的组合物,其中所述第二RNA结合蛋白包含内切核酸酶G (ENDO G) 或由其组成。

实施方案107.根据实施方案106所述的组合物,其中所述ENDO G蛋白包含SEQ ID NO:33或由其组成。

实施方案108.根据实施方案103所述的组合物,其中所述第二RNA结合蛋白包含内切核酸酶D1 (ENDO D1) 多肽或由其组成。

实施方案109.根据实施方案108所述的组合物,其中所述ENDO D1包含SEQ ID NO:34或由其组成。

实施方案110.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含人瓣状内切核酸酶-1 (hFEN1) 多肽或由其组成。

实施方案111.根据实施方案110所述的组合物,其中所述hFEN1蛋白包含SEQ ID

NO:35或由其组成。

实施方案112.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含人斯库拉芬蛋白14(hSLFN14)多肽或由其组成。

实施方案113.根据实施方案112所述的组合物,其中所述hSLFN14多肽包含SEQ ID NO:36或由其组成。

实施方案114.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含人 β -内酰胺酶样蛋白2(hLACTB2)多肽或由其组成。

实施方案115.根据实施方案114所述的组合物,其中所述hLACTB2多肽包含SEQ ID NO:37或由其组成。

实施方案116.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含无嘌呤/无嘧啶(AP)内切脱氧核糖核酸酶(APEX2)多肽或由其组成。

实施方案117.根据实施方案116所述的组合物,其中所述APEX2多肽包含SEQ ID NO:38或由其组成。

实施方案118.根据实施方案116所述的组合物,其中所述APEX2多肽包含SEQ ID NO:39或由其组成。

实施方案119.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含血管生成素(ANG)多肽或由其组成。

实施方案120.根据实施方案119所述的组合物,其中所述ANG多肽包含SEQ ID NO:40或由其组成。

实施方案121.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含热反应蛋白12(HRSP12)多肽或由其组成。

实施方案122.根据实施方案121所述的组合物,其中所述HRSP12多肽包含SEQ ID NO:41或由其组成。

实施方案123.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含含锌指CCCH型12A(ZC3H12A)多肽或由其组成。

实施方案124.根据实施方案123所述的组合物,其中所述ZC3H12A多肽包含SEQ ID NO:42或由其组成。

实施方案125.根据实施方案124所述的组合物,其中所述ZC3H12A多肽包含SEQ ID NO:43或由其组成。

实施方案126.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含反应性中间亚胺脱氨酶A(RIDA)多肽或由其组成。

实施方案127.根据实施方案126所述的组合物,其中所述RIDA多肽包含SEQ ID NO:44或由其组成。

实施方案128.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含磷脂酶D家族成员6(PDL6)多肽或由其组成。

实施方案129.根据实施方案128所述的组合物,其中所述PDL6多肽包含SEQ ID NO:126或由其组成。

实施方案130.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含内切核酸酶III样蛋白1(NTHL)多肽或由其组成。

实施方案131.根据实施方案130所述的组合物,其中所述NTHL多肽包含SEQ ID NO:123或由其组成。

实施方案132.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含线粒体核糖核酸酶P催化亚基(KIAA0391)多肽或由其组成。

实施方案133.根据实施方案132所述的组合物,其中所述KIAA0391多肽包含SEQ ID NO:127或由其组成。

实施方案134.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含无嘌呤或无嘧啶位点裂解酶(APEX1)多肽或由其组成。

实施方案135.根据实施方案134所述的组合物,其中所述APEX1多肽包含SEQ ID NO:125或由其组成。

实施方案136.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含阿尔古蛋白2(AGO2)多肽或由其组成。

实施方案137.根据实施方案136所述的组合物,其中所述AGO2多肽包含SEQ ID NO:128或由其组成。

实施方案138.根据实施方案67所述的组合物,其中所述第二RNA结合蛋白包含线粒体核酸酶EXOG(EXOG)多肽或由其组成。

实施方案139.根据实施方案138所述的组合物,其中所述EXOG多肽包含SEQ ID NO:129或由其组成。

实施方案140.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含含锌指CCCH型12D(ZC3H12D)多肽或由其组成。

实施方案141.根据实施方案140所述的组合物,其中所述ZC3H12D多肽包含SEQ ID NO:130或由其组成。

实施方案142.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含内质网核信号转导蛋白2(ERN2)多肽或由其组成。

实施方案143.根据实施方案142所述的组合物,其中所述ERN2多肽包含SEQ ID NO:131或由其组成。

实施方案144.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含回力球mRNA监督和核糖体挽救因子(PELO)多肽或由其组成。

实施方案145.根据实施方案144所述的组合物,其中所述PELO多肽包含SEQ ID NO:132或由其组成。

实施方案146.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含YBEY金属肽酶(YBEY)多肽或由其组成。

实施方案147.根据实施方案146所述的组合物,其中所述YBEY多肽包含SEQ ID NO:133或由其组成。

实施方案148.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含剪切和多聚腺苷酸化特异性因子4样蛋白(CPSF4L)多肽或由其组成。

实施方案149.根据实施方案148所述的组合物,其中所述CPSF4L多肽包含SEQ ID NO:134或由其组成。

实施方案150.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含

hCG_2002731多肽或由其组成。

实施方案151.根据实施方案150所述的组合物,其中所述hCG_2002731多肽包含SEQ ID NO:135或由其组成。

实施方案152.根据实施方案150所述的组合物,其中所述hCG_2002731多肽包含SEQ ID NO:136或由其组成。

实施方案153.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含切除修复交叉互补组1 (ERCC1) 多肽或由其组成。

实施方案154.根据实施方案153所述的组合物,其中所述ERCC1多肽包含SEQ ID NO:137或由其组成。

实施方案155.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含ras相关C3肉毒杆菌毒素底物1亚型 (RAC1) 多肽或由其组成。

实施方案156.根据实施方案155所述的组合物,其中所述RAC1多肽包含SEQ ID NO:138或由其组成。

实施方案157.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含核糖核酸酶A A1 (RAA1) 多肽或由其组成。

实施方案158.根据实施方案157所述的组合物,其中所述RAA1多肽包含SEQ ID NO:139或由其组成。

实施方案159.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含Ras相关蛋白 (RAB1) 多肽或由其组成。

实施方案160.根据实施方案159所述的组合物,其中所述RAB1多肽包含SEQ ID NO:140或由其组成。

实施方案161.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含DNA复制解旋酶/核酸酶2 (DNA2) 多肽或由其组成。

实施方案162.根据实施方案161所述的组合物,其中所述DNA2多肽包含SEQ ID NO:141或由其组成。

实施方案163.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含FLJ35220多肽或由其组成。

实施方案164.根据实施方案163所述的组合物,其中所述FLJ35220多肽包含SEQ ID NO:142或由其组成。

实施方案165.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含FLJ13173多肽或由其组成。

实施方案166.根据实施方案165所述的组合物,其中所述FLJ13173多肽包含SEQ ID NO:143或由其组成。

实施方案167.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含DNA修复内切核酸酶XPF (ERCC4) 多肽或由其组成。

实施方案168.根据实施方案167所述的组合物,其中所述ERCC4多肽包含SEQ ID NO:124或由其组成。

实施方案169.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R)) 多肽或由其组成。

实施方案170.根据实施方案169所述的组合物,其中所述RNA酶1 (K41R) 多肽包含SEQ ID NO:116或由其组成。

实施方案171.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R、D121E)) 多肽或由其组成。

实施方案172.根据实施方案171所述的组合物,其中所述RNA酶1 (RNA酶1 (K41R、D121E)) 多肽包含SEQ ID NO:117或由其组成。

实施方案173.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (K41R、D121E、H119N)) 多肽或由其组成。

实施方案174.根据实施方案173所述的组合物,其中所述RNA酶1 (RNA酶1 (K41R、D121E、H119N)) 多肽包含SEQ ID NO:118或由其组成。

实施方案175.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (H119N)) 多肽或由其组成。

实施方案166.根据实施方案175所述的组合物,其中所述RNA酶1 (RNA酶1 (H119N)) 多肽包含SEQ ID NO:119或由其组成。

实施方案177.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)) 多肽或由其组成。

实施方案178.根据实施方案177所述的组合物,其中所述RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)) 多肽包含SEQ ID NO:120或由其组成。

实施方案179.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)) 多肽或由其组成。

实施方案180.根据实施方案179所述的组合物,其中所述RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N、K41R、D121E)) 多肽包含SEQ ID NO:121或由其组成。

实施方案181.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含突变的RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D、H119N)) 多肽或由其组成。

实施方案182.根据实施方案181所述的组合物,其中所述RNA酶1 (RNA酶1 (R39D、N67D、N88A、G89D、R91D)) 多肽包含SEQ ID NO:122或由其组成。

实施方案183.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含特诺伊林跨膜蛋白1 (TENM1) 多肽或由其组成。

实施方案184.根据实施方案173所述的组合物,其中所述TENM1多肽包含SEQ ID NO:144或由其组成。

实施方案185.根据实施方案77所述的组合物,其中所述第二RNA结合蛋白包含特诺伊林跨膜蛋白2 (TENM2) 多肽或由其组成。

实施方案186.根据实施方案185所述的组合物,其中所述TENM2多肽包含SEQ ID NO:145或由其组成。

实施方案187.根据实施方案1-77中任一项所述的组合物,其中所述第二RNA结合蛋白包含转录激活因子样效应物核酸酶 (TALEN) 多肽或所述多肽的核酸酶结构域或由其组成。

实施方案188.根据实施方案187所述的组合物,其中所述TALEN多肽包含以下序列或由其组成:

1 MRIGKSSGWL NESVSLEYEH VSPPTRPRDT RRRPRAAGDG GLAHLHRRRLA VGYAEDTPRT
61 EARSPAPRRP LPVAPASAPP APSLVPEPPM PVSLPAVSSP RFSAGSSAAI TDPFPSLPPT
121 PVLYAMAREL EALSDATWQP AVPLPAEPPT DARRGNTVFD EASASSPVIA SACPQAFASP
181 PRAPRSARAR RARTGGDAWP APTFLSRPSS SRIGRDVFGK LVALGYSREQ IRKLKQESLS
241 EIAKYHTTTLT GQGFTHADIC RISRRRQSLR VVARNYPELA AALPELTRAH IVDIARQSRG
301 DLALQALLPV ATALTAAPLR LSASQIATVA QYGERPAIQA LYRLRRKLTR APLHLTPQQV
361 VAIASNTGGK RALEAVCVQL PVLRAAPYRL STEQVVAIAS NKGGKQALEA VKAHLDDLGL
421 APYVLDTEQV VAIASHNGGK QALEAVKADL LDLRGAPYAL STEQVVAIAS HNGGKQALEA
481 VKADLLELRG APYALSTEQV VAIASHNGGK QALEAVKAHL LDLRGVPYAL STEQVVAIAS
541 HNGGKQALEA VKAQLLDLRG APYALSTAQV VAIASNGGGK QALEGIGEQL LKLRTAPYGL
601 STEQVVAIAS HDGGKQALEA VGAQLVALRA APYALSTEQV VAIASNKGGK QALEAVKAQL
661 LELRGAPYAL STAQVVAIAS HDGGNQALEA VGTQLVALRA APYALSTEQV VAIASHDGGK
721 QALEAVGAQL VALRAAPYAL NTEQVVAIAS SHGGKQALEA VRALFPDLRA APYALSTAQL
781 VAIASNPGGK QALEAVRALF RELRAAPYAL STEQVVAIAS NHGGKQALEA VRALFRGLRA
841 APYGLSTAQV VAIASSNGGK QALEAVWALL PVL RATPYDL NTAQIVAIAS HDGGKPALEA
901 VWAKLPVLRG APYALSTAQV VAIACISGQQ ALEAIEAHMP TLRQASHSLS PERVAIACI
961 GGRSAVEAVR QGLPVKAIRR IRREKAPVAG PPPASLGPTP QELVAVLHFF RAHQQPRQAF
1021 VDALAAFQAT RPALLRLLSS VGVTEIEALG GTIPDATERW QRLLGRLGFR PATGAAAPSP
1081 DSLQGFAQSL ERTLGSFGMA GQSACSPHRK RPAETAIAPR SIRRSPNNAG QPSEPWPDQL
1141 AWLQRRKRTA RSHIRADSAA SVPANLHLGT RAQFTPDRLR AEPGPIMQAH TSPASVSFGS
1201 HVAFEPGLPD PGTPTSADLA SFEAEFPGVG PLDFHLDWLL QILET (SEQ ID NO: 205)。

实施方案189.根据实施方案187所述的组合物,其中所述TALEN多肽包含以下序列
或由其组成:

1 mdpirsrtps parellpgpq pdrvqptadr ggappaggpl dglparrtms rtrlpsppap
 61 spafsagsfs dllrqfdpsl ldtsllldsmv avgtphataa paecdevqsg lraaddpppt
 121 vrvavtaarp prakpaprrr aaqpsdaspa aqvdlrtlgy sqqqqekikp kvgstvaqhh
 181 ealvghgfth ahivalsrhp aalgtvavky qdmiaalpea thedivgvkg qwsgaralea
 241 lltvagelrg pplqltdgql vkiakrggvt aveavhasrn altgaplnlt paqvaiasn
 301 nggkqaletv qrllpvlcqa hgltpaqvva iashdggkqa letmqrllpv lcqahglppd
 361 qvvaianasng gkqaletvqr llpvlcqhag ltpdqvvaia shgggkqale tvqrllpvlc
 421 qahgltpdqv vaiashdggk qaletvqrll pvlcqhaglt pdqvvaian gggkqaletv
 481 qrllpvlcqa hgltpdqvva iasnggkqal etvqrllpvl cqahgltpdq vvaishdgg
 541 kqaletvqrl lpvlcqhagl tpaqvvaian hdggkqalet vqqlpvlcqh ahgltpdqv
 601 aiasniggkq alatvqrllp vlcqhagltq dqvvaian gggkqaletv rllpvlcqh
 661 gltpdqvvai asngggkqal etvqrllpvl cqahgltpdq vvaianigg kqaletvqrl
 721 lpvlcqhagl tpaqvvaian hdggkqalet vqrllpvlcqh ahgltpdqv aiasngggkq
 781 aletvqrllp vlcqhagltq eqvvaian gggkqaletv rllpvlcqh gltpdqvvai
 841 asngggkqal etvqrllpvl cqahgltpaq vvaianigg kqaletvqrl lpvlcqhagl
 901 tlaqvvaian niggkqalet vqrllpvlcqh ahgltpdqv aiasniggkq aletvqrllp
 961 vlcqhagltq dqvvaian gggkqaletv rllpvlcqh gltpdqvvai asngggkqale
 1021 tvqrllpvlc qdhgltpdqv vaianssggk qaletvqrll pvlcqhaglt pnqvvaian
 1081 gggkqalesiv aqlsrpdal aaltndhlva lalggpam davkkglpha pelirrvnrr
 1141 igertshrva dyaqvrvle ffqchshpay afdeamtqfg msrnlvqlf rrvgvtelea
 1201 rggtlppasq rwdrlqasg mkrakpspts aqtpdqaslh afadslerdl dapspmhgd
 1261 qtgassrkrs rsdravtgps aqhsfevrp eqrdalhlpl swrvkrprtr igggldpdp
 1321 piaadlaass tvmweqdaap fagaaddfpa fneelawlm ellpqsgsvg gti (SEQ ID
 NO: 206)。

实施方案190.根据实施方案1-77中任一项所述的组合物,其中所述第二RNA结合蛋白包含锌指核酸酶多肽或所述多肽的核酸酶结构域或由其组成。

实施方案191.根据实施方案190所述的组合物,其中所述锌指核酸酶多肽包含以下序列或由其组成:

1 MSRPRENPRG DFPLQRPRAP NPSGMRPPGP FMRPGSMGLP RFYPAGRARG IPHRFAGHES
 61 YQNMGPQRMN VQVTQHRTDP RLTKEKLDHF EAQQKKGKPH GSRWDDEPHI SASVAVKQSS

121 VTQVTEQSPK VQSRYTEKSA SSILASFGLS NEDLEELSRV PDEQLTPENM PLILRDIRMR
 181 KMGRRLPNLP SQSRNKETLG SEAVSSNVID YGHASKYGYT EDPLEVRIYD PEIPTDEVEN
 241 EFQSQQNISA SVPNPNVICN SMFPVEDVFR QMDFPGESSN NRSFFSVESG TKMSGHLHISG
 301 GQSVLEPIKS VNQSINQTVS QTMSQSLIPP SMNQPPFSSE LISSVSQQER IPHEPVINSS
 361 NVHVGSRGSK KNYQSQADIP IRSPFGIVKA SWLPKFSHAD AQKMKRLPTP SMMNDYYAAS
 421 PRIFFPHLCSL CNVECSHLKD WIQHQNTSTH IESCRQLRQQ YPDWNPEILP SRRNEGNRKE
 481 NETPRRRSHS PSPRRSRRSS SSHRFRRSRS PMHYMYRPRS RSPRICHRFI SRYRSRSRSR
 541 SPYRIRNPFR GSPKCFRSVS PERMSRRSVR SDRKKALED VVQRSGHGTE FNKQKHLEAA
 601 DKGHSPAQKP KTSSGKPSV KPTSATKSDS NLGGHSIRCK SKNLEDDTLS ECKQVSDKAV
 661 SLQRKLRKEQ SLHYGSVLLI TELPEDGCTE EDVRKLFQPF GKVNDVLIVP YRKEAYLEME
 721 FKEAITAIMK YIETPLTIK GKSVKICVPG KKAQNKEVK KKTLESKKVS ASTLKRDATA
 781 SKAVEIVTST SAAKTQAKA SVAKVNKSTG KSASSVKSVV TVAVKGNKAS IKTAKSGGKK
 841 SLEAKKTGNV KNKDSNKPVT IPENSEIKTS IEVKATENCA KEAISDAALE ATENEPLNKE
 901 TEEMCVMLVS NLPNKGYSVE EVYDLAKPFG GLKDILILSS HKKAYIEINR KAAESMVKFY
 961 TCFPVLMGDN QLSISMAPEN MNIKDEEAI F ITLVKENDPE ANIDTIYDRF VHLNLPEDG
 1021 LQCVLCVGLQ FGKVDHVFIF SNRNKAILQL DSPESAQSMY SFLKQNPQNI GDHMLTCSLS
 1081 PKIDLPEVQI EHDPELEKES PGLKNSPIDE SEVQTATDSP SVKPNELEEE STPSIQTETL
 1141 VQQEEPCEEE AEKATCDSDF AVETLELETQ GEEVKEEIPV VASASVSIEQ FTENAEECAL
 1201 NQQMFNSDLE KKGAEIINPK TALLPSDSVF AEERNLKGIL EESPSEAEDE ISGITQTMVE
 1261 AVAEVEKNET VSEILPSTCI VTLVPGIPTG DEKTVDKKNI SEKKGNMDEK EEKEFNTKET
 1321 RMDLQIGTEK AEKNEGRMDA EKVEKMAAMK EKPAENTLFK AYPNKGVGQA NKPDETSKTS
 1381 ILAVSDVSSS KPSIKAVIVS SPKAKATVSK TENQKSFPS VPRDQINAEL KLSAKEFGLL
 1441 KPTSARSGLA ESSSKFKPTQ SSLTRGGSGR ISALQGKLSK LDYRDITKQS QETEARPSIM
 1501 KRDDSNNKTL AEQNTKNPKS TTGRSSKSKE EPLFPFNLDE FVTVDEVIEE VNPSQAKQNP
 1561 LKGKRKETLK NVPFSELNLK KKKGKTSTPR GVEGELSFTV LDEIGEEEDA AAHLAQALVT
 1621 VDEVIDEEEL NMEEMVKNSN SLFTLDELID QDDCISHSEP KDVTVLVAE EQDLLKQERL
 1681 VTVDEIGEVE ELPLNESADI TFATLNTKGN EGDTVRDSIG FISSQVPEDP STLVTVDEIQ
 1741 DDSSDLHLVT LDEVTEEDED SLADFNNLKE ELNFVTVDEV GEEEDGDNDL KVELAQSKND
 1801 HPTDKKGNRK KRAVDTKKTK LESLSQVGPV NENVMEEDLK TMIERHLTAK TPTKRVRIK
 1861 TLPSEKAVVT EPAKGEEAFQ MSEVDEESGL KDSEPERKRK KTEDSSSGKS VASDVPEELD
 1921 FLVPKAGFFC PICSLFYSGE KAMTNHCKST RHKQNTKFM AKQRKEKEQN EAEERSSR
 (SEQ ID NO: 207) 。

实施方案192.根据实施方案1-191中任一项所述的组合物,其中所述组合物还包

含(a)包含特异性结合于RNA分子内的gRNA的序列;以及
(b)编码核酸酶的序列。

实施方案193.根据实施方案192所述的组合物,其中所述核酸酶包含从CRISPR/Cas蛋白分离或衍生的序列。

实施方案194.根据实施方案193所述的组合物,其中所述CRISPR/Cas蛋白是从以下中的任一种分离或衍生的:I型、IA型、IB型、IC型、ID型、IE型、IF型、IU型、III型、IIIA型、IIIB型、IIIC型、IIID型、IV型、IVA型、IVB型、II型、IIA型、IIB型、IIC型、V型或VI型CRISPR/Cas蛋白。

实施方案195.根据实施方案192所述的组合物,其中所述核酸酶包含从TALEN或其核酸酶结构域分离或衍生的序列。

实施方案196.根据实施方案192所述的组合物,其中所述核酸酶包含从锌指核酸酶或其核酸酶结构域分离或衍生的序列。

实施方案197.根据实施方案191-196中任一项所述的组合物,其中所述靶序列包含编码适应性免疫应答的组分的序列。

实施方案198.一种载体,其包含根据实施方案1-197中任一项所述的组合物。

实施方案199.根据实施方案198所述的载体,其中所述载体是病毒载体。

实施方案200.根据实施方案199所述的载体,其中所述载体包含从慢病毒、腺病毒、腺相关病毒(AAV)载体或逆转录病毒分离或衍生的序列。

实施方案201.根据实施方案199或200所述的载体,其中所述载体无复制能力。

实施方案202.根据实施方案100-201中任一项的实施方案所述的载体,其中所述载体包含从腺相关载体(AAV)分离或衍生的序列。

实施方案203.根据实施方案202所述的载体,其中所述腺相关病毒(AAV)是分离的AAV。

实施方案204.根据实施方案202或203所述的载体,其中所述腺相关病毒(AAV)是自身互补的腺相关病毒(scAAV)。

实施方案205.根据实施方案202-204中任一项所述的载体,其中所述腺相关病毒(AAV)是重组腺相关病毒(rAAV)。

实施方案206.根据实施方案202-205中任一项所述的载体,其中所述腺相关病毒(AAV)包含从血清型AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11或AAV12的AAV分离或衍生的序列。

实施方案207.根据实施方案202-206中任一项所述的载体,其中所述腺相关病毒(AAV)包含从血清型AAV9的AAV分离或衍生的序列。

实施方案208.根据实施方案202-206中任一项所述的载体,其中所述腺相关病毒(AAV)包含从Anc80分离或衍生的序列。

实施方案209.根据实施方案100-201中任一项所述的载体,其中所述载体是逆转录病毒。

实施方案210.根据权利要求100-201中任一项的实施方案所述的载体,其中所述逆转录病毒是慢病毒。

实施方案211.根据实施方案198所述的载体,其中所述载体是非病毒载体。

实施方案212.根据实施方案211所述的载体,其中所述非病毒载体包含纳米颗粒、胶束、脂质体或阳离子脂质体/DNA复合物、聚合物囊泡、聚合物/DNA复合物或树枝状聚合物。

实施方案213.一种组合物,其包含根据实施方案198-212中任一项所述的载体。

实施方案214.一种细胞,其包含根据实施方案198-212中任一项所述的载体。

实施方案215.一种细胞,其包含根据实施方案214所述的组合物。

实施方案216.根据实施方案214或215所述的细胞,其中所述细胞是哺乳动物细胞。

实施方案217.根据实施方案216所述的细胞,其中所述细胞是人细胞。

实施方案218.根据实施方案215-217中任一项所述的细胞,其中所述细胞是免疫细胞。

实施方案219.根据实施方案218所述的细胞,其中所述免疫细胞是T淋巴细胞(T细胞)。

实施方案220.根据实施方案219所述的细胞,其中所述T细胞是效应T细胞、辅助T细胞、记忆T细胞、调节性T细胞、自然杀伤T细胞、粘膜相关恒定T细胞或 $\gamma\delta$ T细胞。

实施方案221.根据实施方案215-217中任一项所述的细胞,其中所述免疫细胞是抗原呈递细胞。

实施方案222.根据实施方案221所述的细胞,其中所述抗原呈递细胞是树突细胞、巨噬细胞或B细胞。

实施方案223.根据实施方案221所述的细胞,其中所述抗原呈递细胞是体细胞。

实施方案224.根据实施方案215-223中任一项所述的细胞,其中所述细胞是健康细胞。

实施方案225.根据实施方案215-223中任一项所述的细胞,其中所述细胞不是健康细胞。

实施方案226.根据实施方案225所述的细胞,其中所述细胞是从患有疾病或障碍的受试者分离或衍生的。

实施方案227.一种组合物,其包含根据实施方案215-226中任一项所述的细胞。

实施方案228.一种从适应性免疫应答中掩蔽细胞的方法,其包括使根据实施方案1-197、213或227中任一项所述的组合物与所述细胞接触以产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述RNA分子编码适应性免疫应答的组分。

实施方案229.根据实施方案228所述的方法,其中所述细胞是体内的、体外的、离体的或原位的。

实施方案230.根据实施方案228所述的方法,其中所述细胞是体外的或离体的。

实施方案231.根据实施方案228-230中任一项所述的方法,其中多个细胞包含所述细胞。

实施方案232.根据实施方案231所述的方法,其中使所述多个细胞中的每个细胞与所述组合物接触,从而产生多个修饰的细胞。

实施方案233.根据实施方案228-230中任一项所述的方法,其中所述方法还包括

向受试者施用所述修饰的细胞。

实施方案234.根据实施方案231-232中任一项所述的方法,其中所述方法还包括向受试者施用所述多个修饰的细胞。

实施方案235.根据实施方案233所述的方法,其中所述细胞是自体的。

实施方案236.根据实施方案233所述的方法,其中所述是同种异体的。

实施方案237.根据实施方案233所述的方法,其中所述多个修饰的细胞是自体的。

实施方案238.根据实施方案233所述的方法,其中所述多个修饰的细胞是同种异体的。

实施方案239.根据实施方案228-238中任一项所述的方法,其中适应性免疫应答的所述组分包含以下的组分或由其组成:I型主要组织相容性复合物(MHC I)、II型主要组织相容性复合物(MHC II)、T细胞受体(TCR)、共刺激分子或其组合。

实施方案240.根据实施方案239所述的方法,其中所述MHC I组分包含 α 1链、 α 2链、 α 3链或 β 2M蛋白。

实施方案241.根据实施方案228-238中任一项所述的方法,其中适应性免疫应答的所述组分包含MHC I β 2M蛋白或由其组成。

实施方案242.根据实施方案239所述的方法,其中所述MHC II组分包含 α 1链、 α 2链、 β 1链或 β 2链。

实施方案243.根据实施方案239所述的方法,其中所述TCR组分包含 α 链和 β 链。

实施方案244.根据实施方案239所述的方法,其中所述共刺激分子包含分化群28(CD28)、分化群80(CD80)、分化群86(CD86)、可诱导T细胞共刺激分子(ICOS)或ICOS配体(ICOSLG)蛋白。

实施方案245.一种防止或减少受试者的适应性免疫应答的方法,其包括向所述受试者施用治疗有效量的根据实施方案1-197、213或227中任一项所述的组合物,其中使所述组合物与所述受试者的至少一个细胞接触,产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述RNA分子编码适应性免疫应答的组分。

实施方案246.一种治疗受试者的疾病或障碍的方法,其包括向所述受试者施用治疗有效量的根据实施方案1-197、213或227中任一项所述的组合物,其中使所述组合物与所述受试者的至少一个细胞接触,产生修饰的细胞,其中所述组合物修饰所述修饰的细胞的RNA分子的表达水平,并且其中所述组合物防止或减少对所述修饰的细胞的适应性免疫应答。

实施方案247.根据实施方案246所述的方法,其中适应性免疫应答的所述组分包含以下的组分或由其组成:I型主要组织相容性复合物(MHC I)、II型主要组织相容性复合物(MHC II)、T细胞受体(TCR)、共刺激分子或其组合。

实施方案248.根据实施方案247所述的方法,其中所述MHC I组分包含 α 1链、 α 2链、 α 3链或 β 2M蛋白。

实施方案249.根据实施方案247或248所述的方法,其中适应性免疫应答的所述组分包含MHC I β 2M蛋白或由其组成。

实施方案250.根据实施方案249所述的方法,其中所述MHC II组分包含 α 1链、 α 2链、 β 1链或 β 2链。

实施方案251.根据实施方案247所述的方法,其中所述TCR组分包含 α 链和 β 链。

实施方案252.根据实施方案247所述的方法,其中所述共刺激分子包含分化群28 (CD28)、分化群80 (CD80)、分化群86 (CD86)、可诱导T细胞共刺激分子 (ICOS) 或ICOS配体 (ICOSLG) 蛋白。

实施方案253.根据实施方案246-252中任一项所述的方法,其中所述疾病或障碍是遗传疾病或障碍。

实施方案254.根据实施方案253所述的方法,其中所述疾病或障碍是单基因遗传疾病或障碍。

实施方案255.根据实施方案254所述的方法,其中所述疾病或障碍是由微卫星不稳定性引起的。

实施方案256.根据实施方案255所述的方法,其中所述微卫星不稳定性发生在DNA序列的至少1、2、3、4、5或6个重复的基序处。

实施方案257.根据实施方案256所述的方法,其中RNA分子包含所述DNA序列的转录物,并且其中所述组合物与所述RNA分子的靶序列结合,所述靶序列包含至少1、2、3、4、5或6个重复的基序。

实施方案258.根据实施方案246-257中任一项所述的方法,其中全身性施用所述组合物。

实施方案259.根据实施方案259所述的方法,其中静脉内施用所述组合物。

实施方案260.根据实施方案258或259所述的方法,其中通过注射或输注施用所述组合物。

实施方案261.根据实施方案246-257中任一项所述的方法,其中局部施用所述组合物。

实施方案262.根据实施方案261所述的方法,其中通过骨内、眼内、脑内或脊柱内途径施用所述组合物。

实施方案263.根据实施方案261或262所述的方法,其中通过注射或输注施用所述组合物。

实施方案264.根据实施方案265-263中任一项所述的方法,其中所述治疗有效量是单剂量。

实施方案265.根据实施方案265-264中任一项所述的方法,其中所述组合物是非基因组整合的。

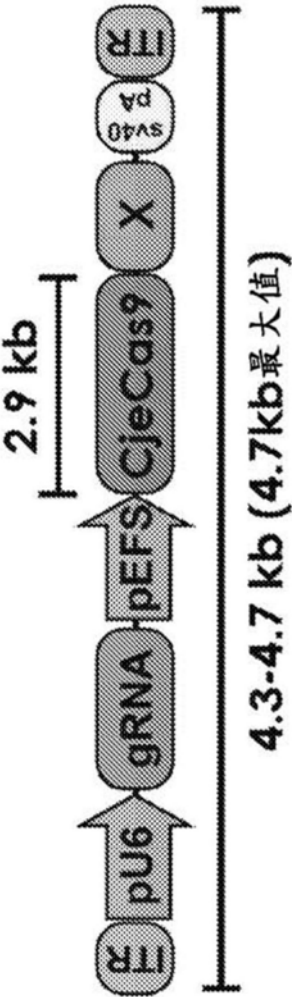
通过引用并入

[0413] 将在本文中引用的每个文件(包括任何交叉参考或相关的专利或申请)通过引用以其整体特此并入本文,除非明确排除或另有限制。引用任何文件并非承认,它是关于本文公开或要求保护的发明的现有技术,或者它单独地或与任何其他一个或多个参考文献的任何组合传授、表明或公开任何这种发明。此外,在本文件中术语的任何含义或定义与通过引用并入的文件中相同术语的任何含义或定义矛盾的方面来说,应当以分配给本文件中该术语的含义或定义为准。

其他实施方案

[0414] 虽然已经说明并描述了本公开文本的特定实施方案,但是可以在不背离本公开文

本的精神和范围的情况下做出各种其他变化和修改。所附权利要求的范围包括在本公开文本的范围内的所有此类变化和修改。



内切核酸酶
空肠弯曲杆菌 Cas9

图1A

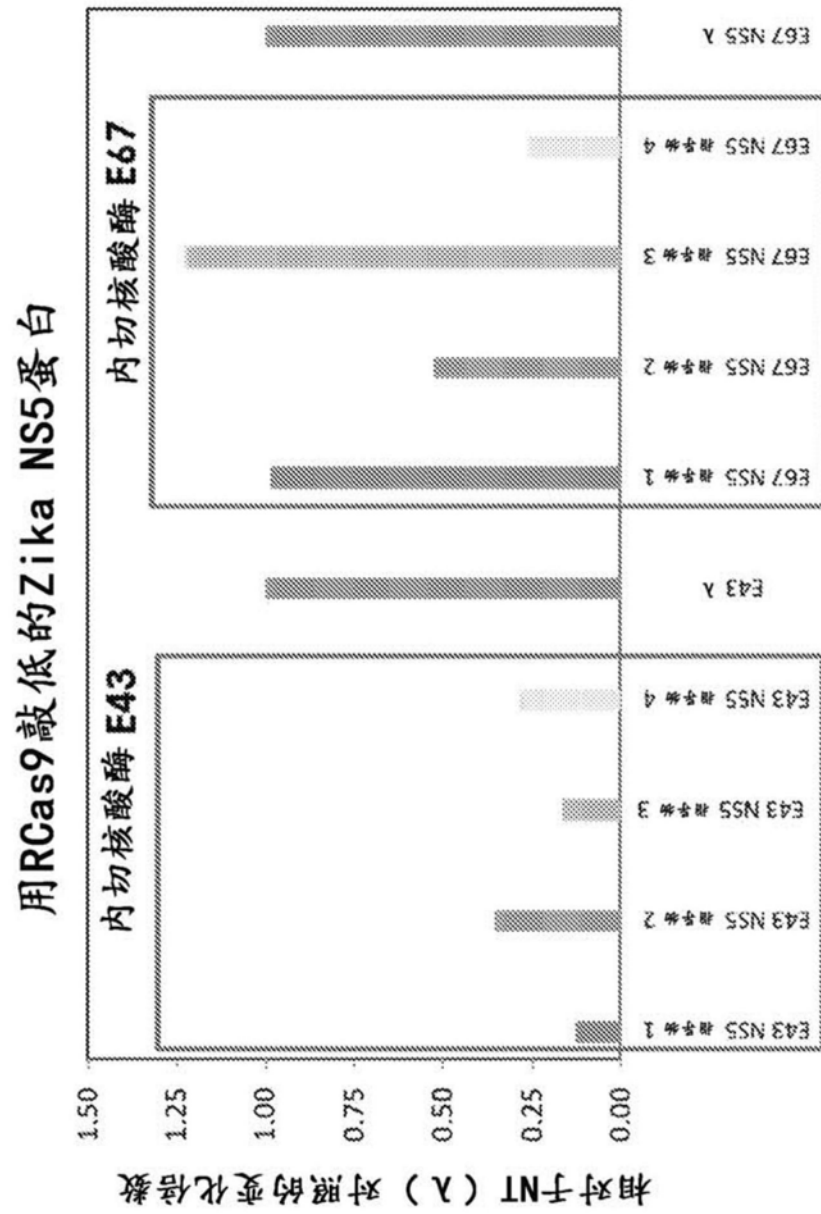


图1B

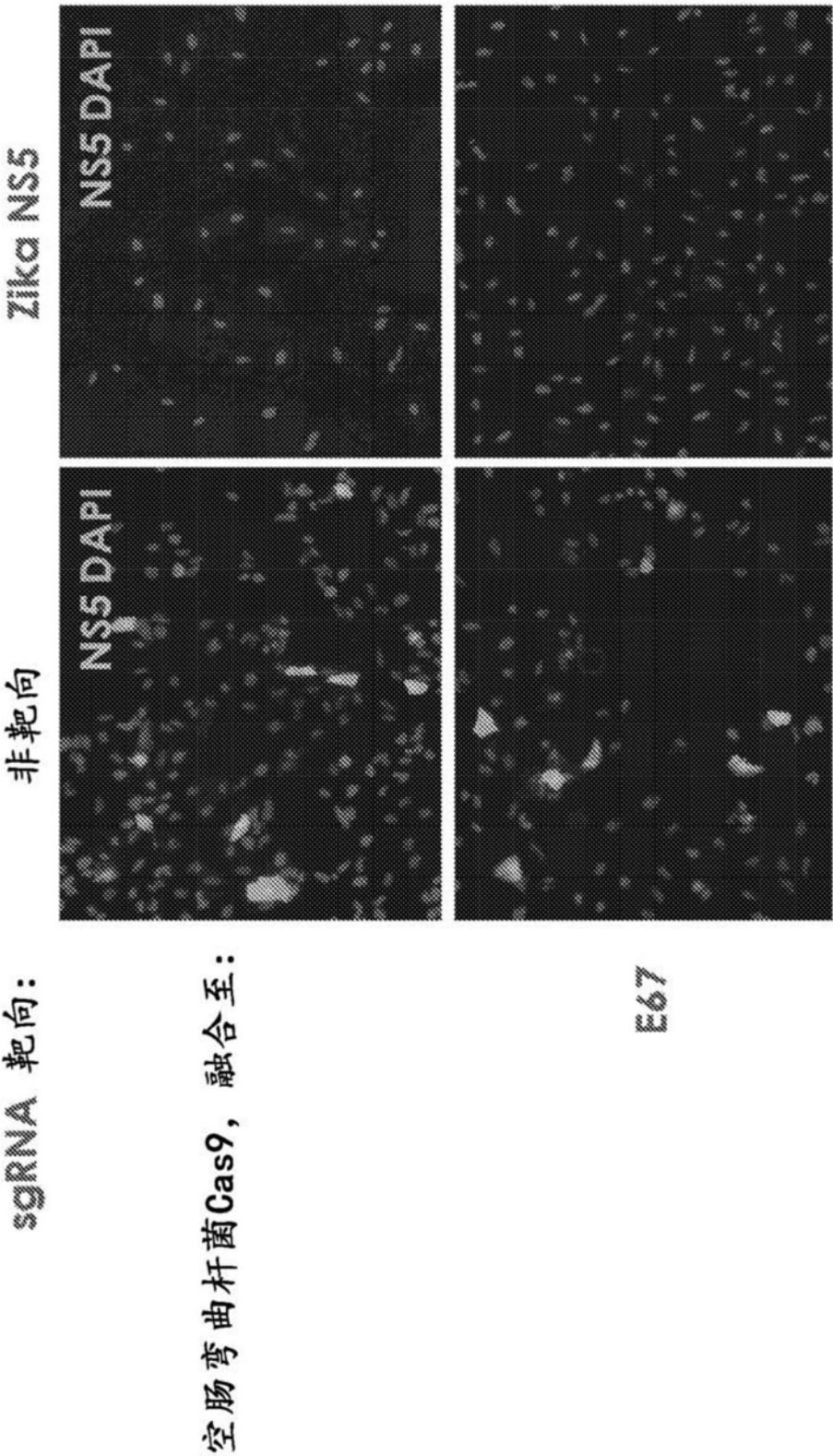


图2A

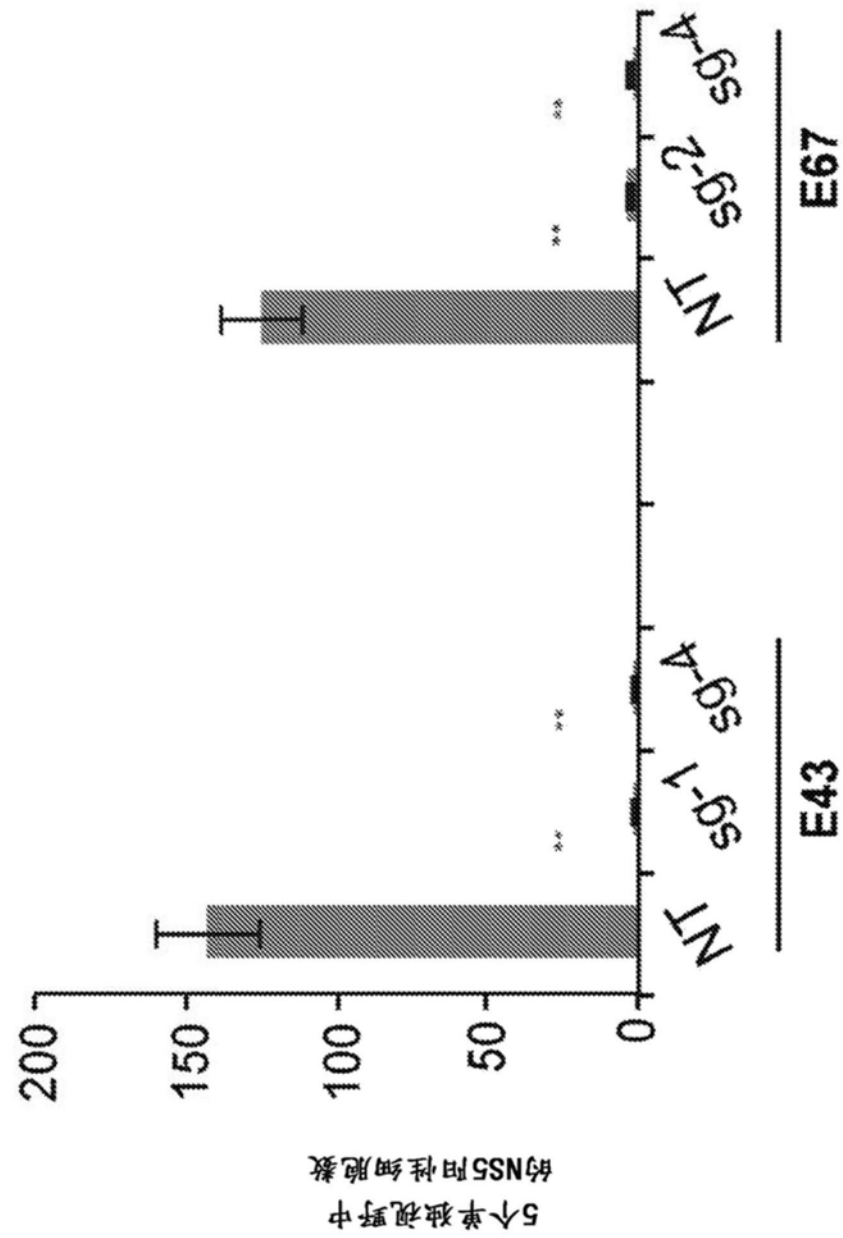


图2B

内切核酸酶名称关键词

E1	RNA酶1	E53	EXOG
E2	RNA酶4	E55	ZC3H13D
E3	RNA酶6	E57	ERN2
E4	NOB1	E59	ENDOG
E5	EndoV	E61	PELO
E6	FEN1	E63	Y8EY
E7	SLFN14	E65	ENDOD1
E8	LACTB2	E67	CPSF4L
E9	RNA酶7	E69	hCG_2002731
E10	RNA酶8	E71	hCG_2002731
E11	RNA酶2	E73	ERCC1
E12	ANG	E75	RAC1
E13	HRSP12	E77	APEX2
E14	RNA酶6PL	E79	APEX2_1-350
E15	RNA酶L	E81	R441_25-156
E16	RNA酶T2	E83	R481
E17	ZC3H12A	E85	RNA酶K
E37	RIDA	E87	DNA2_FL
E39	RNA酶11	E89	RNA酶1(K41R)
E41	PDL6	E91	RNA酶1(K41R, D121E)
E43	NTHL1	E93	RNA酶1(K41R, D121E, H119N)
E45	KIA0391	E95	RNA酶1(H119N)
E47	APEX1		
E49	Ago2	E97	RNA酶1(R39D, N67D, N88A, G89D, R91D, H119N)
E51	ZC3H12A	E99	RNA酶1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
		E100	RNA酶1(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
		E101	RNA酶1(R39D, N67D, N88A, G89D, R91D)

图3

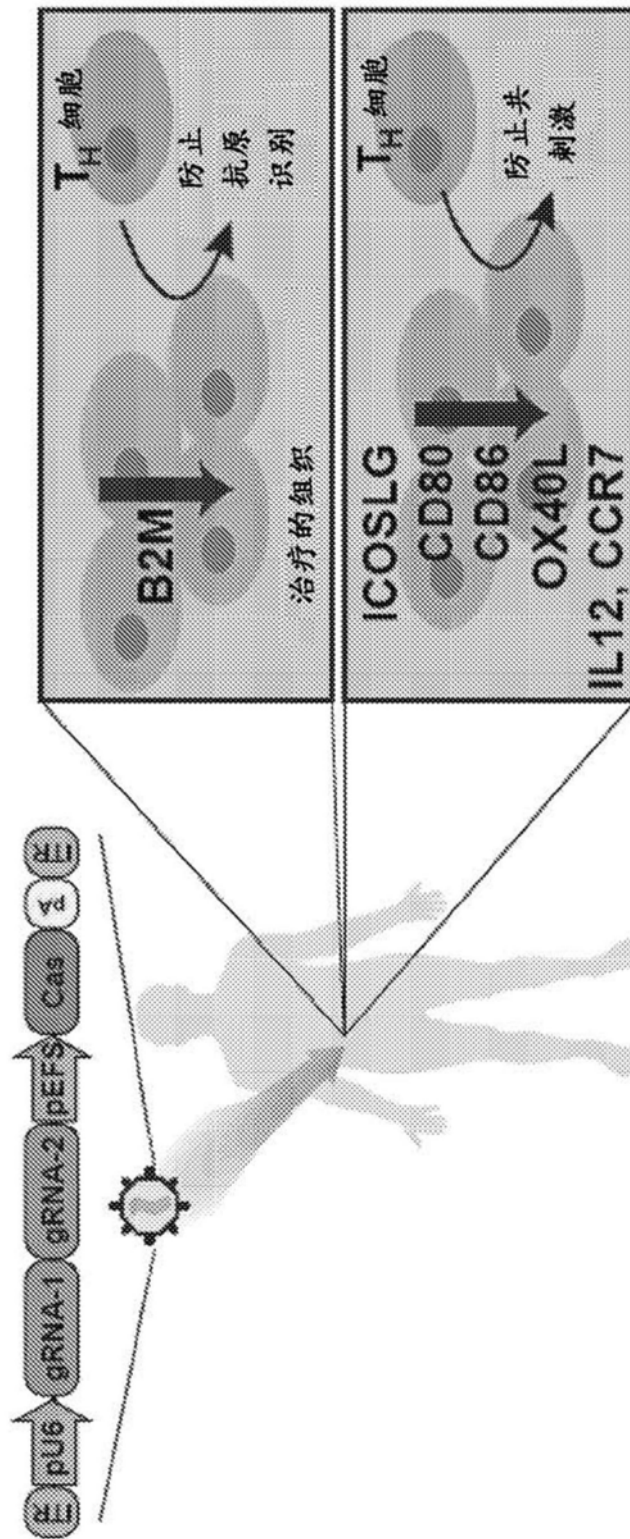


图4