



(12) 发明专利申请

(10) 申请公布号 CN 112930395 A

(43) 申请公布日 2021.06.08

(21) 申请号 201980050249.4

(22) 申请日 2019.06.07

(30) 优先权数据

62/682,271 2018.06.08 US

(85) PCT国际申请进入国家阶段日

2021.01.27

(86) PCT国际申请的申请数据

PCT/US2019/036021 2019.06.07

(87) PCT国际申请的公布数据

W02019/236982 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

(74) 专利代理机构 北京坤瑞律师事务所 11494

代理人 封新琴

(51) Int.Cl.

C12N 9/22 (2006.01)

C12N 15/113 (2006.01)

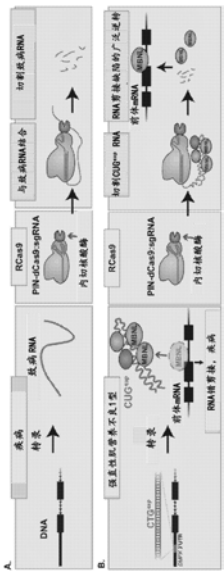
权利要求书2页 说明书103页
序列表327页 附图14页

(54) 发明名称

靶向RNA的融合蛋白组合物和使用方法

(57) 摘要

公开了组合物,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。还提供了本公开文本的组合物的制备方法和使用方法。例如,本公开文本的组合合物可以用于治疗受试者的疾病或障碍。本公开文本的示例性疾病或障碍包括遗传和表观遗传疾病或障碍。



1. 一种组合物,所述组合物包含编码RNA指导的靶RNA结合融合蛋白的核酸序列,所述融合蛋白包含(a)第一RNA结合多肽或其部分;以及(b)第二RNA结合多肽,其中在通过gRNA序列指导时,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。

2. 根据权利要求1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。

3. 根据权利要求2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。

4. 根据权利要求1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。

5. 根据权利要求4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。

6. 根据权利要求1所述的组合物,其中所述核酸序列包含启动子。

7. 根据权利要求6所述的组合物,其中所述启动子是组成型启动子或组织特异性启动子。

8. 根据权利要求1所述的组合物,其中所述核酸序列还包含gRNA序列,其中所述gRNA序列包含特异性结合RNA分子内的靶序列的间隔子序列和与所述第一RNA结合多肽特异性结合的支架序列。

9. 根据权利要求8所述的组合物,其中所述间隔子序列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。

10. 根据权利要求8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。

11. 根据权利要求9所述的组合物,其中所述启动子是聚合酶III启动子。

12. 根据权利要求10所述的组合物,其中所述聚合酶III启动子是U6启动子。

13. 根据权利要求1或9所述的组合物,其中所述启动子是tRNA启动子。

14. 根据权利要求1或9所述的组合物,其中所述融合蛋白包含NLS、NES或标签。

15. 一种载体,其包含根据权利要求1或8所述的组合物。

16. 根据权利要求15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。

17. 一种细胞,其包含根据权利要求15所述的载体。

18. 根据权利要求1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、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和ZNF638。

19. 一种组合物,其包含:

(a) 指导RNA (gRNA) 序列,所述指导RNA (gRNA) 序列包含特异性结合RNA分子内的靶序列的间隔子序列和与所述第一RNA结合多肽特异性结合的支架序列;

(b) 编码融合蛋白的核酸序列,所述融合蛋白包含第一RNA结合多肽和编码第二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。

20. 一种用于修饰靶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融合蛋白或其部分与所述RNA分子结合的条件下使根据权利要求19所述的组合物与所述RNA分子接触。

靶向RNA的融合蛋白组合物和使用方法

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62/682,271的优先权,将其内容通过引用以其整体并入本文。将2018年6月8日提交的美国专利申请号62/682,276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773KB的名为“LOCN_002_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] 对于特异性结合靶RNA分子以修饰所述RNA分子或由所述RNA分子编码的蛋白质的表达或活性的方法,在本领域中存在长期但未满足的需求。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,所述序列特异性方式进一步排除DNA序列的修饰。

发明内容

[0005] 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。

[0006] 本公开文本还提供了一种组合物,所述组合物包含编码RNA指导的靶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结合多肽结合由gRNA序列指导的靶RNA,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0007] 本公开文本另外提供了一种组合物,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结合多肽在没有gRNA序列的情况下结合靶RNA,并且其中所述第二RNA结合多肽包含RNA-核酸酶活性。

[0008] 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。

[0009] 在本公开文本的组合物的一些实施方案中,包含所述gRNA的序列还包含编码能够在真核细胞中表达所述gRNA的启动子的序列。

[0010] 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。

[0011] 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,所述启动子是从能够驱动转移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,所述启动子是从缬氨酸tRNA启动子分离或衍生的。

[0012] 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:UGGAGCGAGCAUCCCCCAA (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:12)、GGGGCCGGGGCCGGGGCCGGG (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)。

[0013] 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。

[0014] 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)。

[0015] 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,所述序列包含序列CUG (SEQ ID NO:18)、CCUG (SEQ ID NO:19)、CAG (SEQ ID NO:80)、GGGCC (SEQ ID NO:81)或其任何组合的至少1、2、3、4、5、6或7个重复。

[0016] 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:GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCUUUUUUU (SEQ ID NO:13)。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:GGACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCUUUUUUU (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:GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCUUUUUUU (SEQ ID NO:82)或GUUUUAGAGCUAGAAAUAAGCAAGUUUAAAUAAGGCUAGUCGUAUCAACUUGAAAAAGUGGCACCGAGUCGUGCUUUUUUU (SEQ ID NO:83)。

[0017] 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。

[0018] 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。

[0019] 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。

[0020] 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。

[0021] 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CasRx/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。

[0022] 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。

[0023] 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。

[0024] 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和力或不选择性结合所述第二序列。

[0025] 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。

[0026] 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。

[0027] 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核定位信号(NLS)、核输出信号(NES)或标签的序列。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。

[0028] 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。

[0029] 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。

[0030] 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或由其组成。在一些实施方案中,所述RNA酶1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蛋白包含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。

[0031] 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,所述RNA酶1 (K41R) 多肽包含SEQ ID NO:116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、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:66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、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。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,所述RNA酶1 (RNA酶1 (H119N)) 多肽包含SEQ ID NO:119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 (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。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,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。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,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。

[0032] 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,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,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蛋白包含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ENDO G蛋白包含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 (ENDO D1) 或由其组成。在一些实施方案中,所述ENDO D1蛋白包含SEQ ID NO:34或由其组成。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) 或由其组成。在一些实施方案中,所述hFEN1蛋白包含SEQ ID NO:35或由其组成。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,所述ERCC4蛋白包含SEQ ID NO:64或由其组成。

[0034] 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,所述NTHL多肽包含SEQ ID NO:123或由其组成。

[0035] 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 (Schlafen) 14 (hSLFN14) 多肽或由其组成。在一些实施方案中,所述hSLFN14多肽包含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,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多肽包含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,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,所述APEX2多肽包含SEQ ID NO:38或由其组成。在一些实施方案中,所述APEX2多肽包含SEQ ID NO:39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。

[0038] 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,所述ANG多肽包含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,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多肽包含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,所述第二RNA结合蛋白包含含锌指CCCH型12A (Zinc Finger CCCH-Type Containing 12A, ZC3H12A) 多肽或由其组成。在一些实施方案中,所述ZC3H12A多肽包含SEQ ID NO:42或由其组成。在一些实施方案中,所述ZC3H12A多肽包含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,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,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,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。

[0044] 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(argonaute)2(AGO2)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,所述AGO2多肽包含SEQ ID NO:128或由其组成。

[0045] 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,所述EXOG多肽包含SEQ ID NO:129或由其组成。

[0046] 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,所述ZC3H12D多肽包含SEQ ID NO:130或由其组成。

[0047] 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,所述ERN2多肽包含SEQ ID NO:131或由其组成。

[0048] 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(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。

[0049] 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,所述YBEY多肽包含SEQ ID NO:133或由其组成。

[0050] 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,所述CPSF4L多肽包含SEQ ID NO:134或由其组成。

[0051] 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,所述hCG_2002731包含SEQ ID NO:135或由其组成。在一些实施方案中,所述hCG_2002731多肽包含SEQ ID NO:136或由其组成。

[0052] 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,所述ERCC1多肽包含SEQ ID NO:137或由其组成。

[0053] 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,所述RAC1多肽包含

SEQ ID NO:138或由其组成。

[0054] 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,所述RAA1多肽包含SEQ ID NO:139或由其组成。

[0055] 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,所述RAB1多肽包含SEQ ID NO:140或由其组成。

[0056] 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,所述DNA2多肽包含SEQ ID NO:141或由其组成。

[0057] 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,所述FLJ35220多肽包含SEQ ID NO:142或由其组成。

[0058] 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,所述FLJ13173多肽包含SEQ ID NO:143或由其组成。

[0059] 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 (Teneurin) 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,所述TENM1多肽包含SEQ ID NO:144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2 (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。

[0060] 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,所述RNA酶K多肽包含SEQ ID NO:204或由其组成。

[0061] 在一些实施方案中,将本公开文本的融合蛋白用于治疗有需要的受试者的方法中,所述方法包括使靶RNA与融合蛋白或编码所述融合蛋白的序列接触。

附图说明

[0062] 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。

[0063] 图1A-图1B是本公开文本的组合物示例性实施方案的示意图。(A) 与内切核酸酶融合的靶向RNA的Cas9系统靶向并切割引起疾病的RNA。(B) 描绘了(A) 在强直性肌营养不良1型情况下的应用,其中与内切核酸酶融合的靶向RNA的Cas9系统靶向并切割由重复CUG单元构成的重复性RNA。在靶向RNA的Cas9系统不存在的情况下,由重复CUG单元构成的重复性RNA与剪接因子MBNL结合,并通过功能障碍性RNA剪接引起病状。该重复性RNA的切割改善疾病。

[0064] 图2是描绘了用于通过靶向RNA分子治疗遗传疾病的示例性模块化治疗平台的示意图。

[0065] 图3A-图3B是一对示意图,其描绘了(A) “高表达”对照系统(也称为“pos对照”),所述对照系统包含双质粒系统,所述质粒系统包含驱动RNA内切核酸酶/Cas9融合物的表达的

巨细胞病毒启动子;以及(B)“低表达”对照系统(也称为“P13”),所述对照系统包含单质粒系统,所述质粒系统包含驱动RNA内切核酸酶/Cas9融合物的表达的较低表达启动子(pEFS)。

[0066] 图4A是一对示意图,其描绘了示例性RNA内切核酸酶-空肠弯曲杆菌(*C. jejuni*) Cas9融合蛋白(左)和包含示例性RNA内切核酸酶-酿脓链球菌(*S. pyogenes*) Cas9融合蛋白的载体(右)。

[0067] 图4B是描绘了如图4A中所示的包含空肠弯曲杆菌Cas9或酿脓链球菌Cas9的多种融合蛋白切割重复性RNA分子的能力的图。

[0068] 图5A是一对示意图,其描绘了示例性RNA内切核酸酶-空肠弯曲杆菌Cas9融合蛋白(左)和包含示例性RNA内切核酸酶-酿脓链球菌Cas9融合蛋白的载体(右)。

[0069] 图5B是描绘了如图5A中所示的包含空肠弯曲杆菌Cas9或酿脓链球菌Cas9的多种融合蛋白切割编码萤光素酶蛋白的mRNA分子的能力的图。

[0070] 图6是提供了图4B、图5B和图9中所示的内切核酸酶的关键码的表格。

[0071] 图7A是描绘了示例性RNA内切核酸酶-空肠弯曲杆菌Cas9融合蛋白的示意图。

[0072] 图7B是描绘了在E43和E67CjeCas9-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2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。

[0073] 图8A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。

[0074] 图8B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-内切核酸酶融合物存在下Zika NS5的表达的变化的图。

[0075] 图9是描绘了各种示例性融合蛋白(与注释的内切核酸酶融合的SpyCas9)的切割效率的图。

具体实施方式

[0076] 本公开文本提供了RNA指导的融合蛋白,其选择性结合并且任选地切割RNA分子。本公开文本提供了包含所述RNA指导的融合蛋白的载体、组合物和细胞。本公开文本提供了使用本公开文本的RNA指导的融合蛋白、载体、组合物和细胞治疗疾病或障碍的方法。

指导RNA

[0077] 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。

[0078] 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不连续的。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,编码本公开文本的指导RNA或单一指导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。

[0079] 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。

[0080] 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)。

[0081] 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)。

[0082] 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。

[0083] 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。

[0084] 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。

[0085] 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。

[0086] 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, 本公开文本的间隔子序列包含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。

[0087] 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)。

[0088] 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。

[0089] 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'。

[0090] 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(包括其中指导RNA或其部分包含与PFS互补的序

列的那些实施方案)中,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(包括其中指导RNA或其部分包含与PFS互补的序列的那些实施方案)中,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。

[0091] 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的表达的启动子序列。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的表达的启动子序列。在一些实施方案中,用于驱动指导RNA的表达的启动子是组成型启动子。在一些实施方案中,所述启动子序列是诱导型启动子。在一些实施方案中,所述启动子是序列是组织特异性和/或细胞类型特异性启动子。在一些实施方案中,所述启动子是杂合启动子或重组启动子。在一些实施方案中,所述启动子是能够在哺乳动物细胞中表达指导RNA的启动子。在一些实施方案中,所述启动子是能够在人细胞中表达指导RNA的启动子。在一些实施方案中,所述启动子是能够表达指导RNA并将指导RNA限制于细胞核的启动子。在一些实施方案中,所述启动子是人RNA聚合酶启动子或从编码人RNA聚合酶启动子的序列分离或衍生的序列。在一些实施方案中,所述启动子是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,所述启动子是人tRNA启动子或从编码人tRNA启动子的序列分离或衍生的序列。在一些实施方案中,所述启动子是人缬氨酸tRNA启动子或从编码人缬氨酸tRNA启动子的序列分离或衍生的序列。

[0092] 在本公开文本的组合物的一些实施方案中,用于驱动指导RNA的表达的启动子还包含调节元件。在一些实施方案中,包含用于驱动指导RNA的表达的启动子序列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。

[0093] 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编码指导RNA的序列、用于驱动指导RNA的表达的启动子序列和编码调节元件的序列中的一种或多种。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序列。

融合蛋白

[0094] 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'。

[0095] 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。

[0096] 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。

[0097] 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。

[0098] 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。

[0099] 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。

[0100] 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  IGTSVGVAV  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  EKILTFRIPY  YVGPLARGNS  RFAWMTRKSE  ETITPNWFEE
481 VVDKGASQAS  FIERMTNFDK  NLPNEKVLPK  HSLLEYEFTV  YNELTKVKYV  TEGMRKPAFL
541 SGEQKKAIVD  LLFKTNRKVT  VKQLKEDYFK  KIECFDSVEI  SGVEDRFNAS  LGTYHDLKI
601 IKDKDFLDNE  ENEDILEDIV  LTLTLFEDRE  MIEERLKTYA  HLFDDKVMKQ  LKRRRYTGWG
661 RLSRKLINGI  RDKQSGKTIL  DFLKSDGFAN  RNFMQLIHDD  SLTFKEDIQK  AQVSGQGDLS
721 HEHIANLAGS  PAIKKGILQT  VKVVDLVKV  MGRHKPENIV  IEMARENQTT  QKGQKNSRER
781 MKRIEEGIKE  LGSQILKEHP  VENTQLQNEK  LYLYYLQNGR  DMYVDQELDI  NRLSDYDVDH
841 IVPQSFLKDD  SIDNKVLTRS  DKNRGKSDNV  PSEEVVKMK  NYWRQLLNAK  LITQRKFDNL
901 TKAERGGLSE  LDKAGFIKRQ  LVETRQITKH  VAQILDSRMN  TKYDENDKLI  REVKVITLKS
961 KLVSDFRKDF  QFYKVREINN  YHHAHDAYLN  AVVGTAIIK  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  LSAYNKHDK  PIREQAENII  HLFTLTNLGA
1321 PAAFKYFDTT  IDRKYTSTK  EVLDATLIHQ  SITGLYETRI  DLSQLGGD (SEQ ID NO: 147) 。
```

[0101] 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  IGTSVGVAV  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 DLTLKALVR QQLPEKYKEI FFDQSKNGYA
361 GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH
421 AILRRQEDFY PFLKDNREKI EKILTFRIPI YVGPLARGNS RFAWMTRKSE ETITPWNFEE
481 VVDKGASAQS FIERMTNFDK NLPNEKVLPK HSLLEYFTV YNELTKVKYV TEGMRKPAFL
541 SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLKI
601 IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMQK LKRRRYTGWG
661 RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL
721 HEHIANLAGS PAIKKGILQT VKVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER
781 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDA
841 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL
901 TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
961 KLVSDFRKDF QFYKVFREINN YHHAHDAYLN AVVGTAIIK YPKLESEFVY GDYKVYDVRK
1021 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGAIRKR PLIETNGETG EIVWDKGRDF
1081 ATVRKVLSP QVNIVKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPPK YGGFDSPTVA
1141 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK
1201 YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE
1261 QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHDK PIREQAENII HLFTLTNLGA
1321 PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO: 148)。

[0102] 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。

[0103] 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。

[0104] 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:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSVDKLFILVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLLAAKNL
SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF
YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
PIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNEL
TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI

IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVDELVKVMG
RHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE
RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPPKYG
GFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN
GRKRMLASAGELQKGNELALPSKYVNFYLASHYEKLKGSPEDEKQQLFVEQHKHYLDEIEQISEFSKRVLAD
ANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRI
DLSQLGGD (SEQ ID NO:149)

[0105] 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:

MKRNILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVK
KLLFDYNLLTDHSELGINPYEARVKGLSQKLSSEEFSAALLHLAKRRGVHNVNEVEDTGNELSTKEQISRNKA
LEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYIDLLETRRTYYEGPGEOSP
FGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYKFQI IENVFKQKKKPTL
KQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEI IENAELLDQIAKILTIYQSSEDIQEELTNLN
SELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSP
VVKRSFIQSIKVINAIIKKYGLPNDII IELAREKNSKDAQKMINEMQKRNRQTNERIEEII RTTGKENAKYLIEKI
KLHDMQEGKCLYSLEAIPLEDLLNPNFYEVNDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFYQLSSSDSKIS
YETFKKHILNLA KGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKS
NGGFTSFLRRKWKFKKERNKGYKHHADALIIANADFIKWKKLDKAKKVMENQMFEKQAESMPEIETEYQYKE
IFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLINKSPEKLL
MYHHPQTYQKLKLIMEQYGDENPLYKYEEETGNLYTKYSKKDNGPVIKKIKYYGNKLNALHDITDDYPNSRNKV
VKLSLKPYPYFDVYLDNGVYKFVTVKNLDVIKENYEEVNSKCYEEAKKLLKISNQAEFIASFYNNDLIKINGELYR
VIGVNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG
(SEQID NO:150)

[0106] 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:

MSDLVLGLDIGISVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNR
LFEEGLITDFTKISINLPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDASDDGNSSVGDYAIVKENS
KQLETKTPGQIQLERYQTYGQLRGDFTVEKDGGKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILT
GKRKYHHPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPPDEFRAAKASYTAQEFNLLNDLNLTVPTETKLL
SKEQKNQIIINYVNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMTLETLDIEQMDRETL
DKLAYVLTNTEREGIQEALHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMELIPELYETSEEQMT
ILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIIKIVNAAIKEYGDFDNIVIAMARETNEDDEKKAIQ
KIQKANKDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGTISIHDLINNSNQFEVDHI
LPLSITFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTLNKKKEYLLTEEDISKFDVR

KKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLN
LWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNKISDATIY
ATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINDKGKE
VPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGHNHIDITPKDSNNKVVLQSVSPWRADVFNKTTGKYEI
LGLKYADLQFDKGTGYKISQEKYNDIKKKEGVSDSEFKFTLYKNDLLLVKDTETKEQQFLRFLSRTMPKQKHVY
ELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNI SIYKVRTDVLGNQHI IKNEGDKPKLDF (SEQ ID NO:
151)

[0107] 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
RRLTRRAHRLLRARLLKREGVLQAADFENGLIKSLPNTPWQLRAAALDRKLTPEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKSGHIRNQRGDYSHTFSRKDLQAEILLFEKQK
EFGNPHVSGGLKEGIEITLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTCLNNLRILEQGSERP
LTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALKKEGLKDKKSPL
NLSPQLQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTEEKIYLPPIPADEIRNPVVLRLSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLVLGSENQKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDFKERNLNDTRYVNR
LCQFVADRMRLTGKGGKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPL
FVSRAPNRKMSGQGHMETVKSARKLDEGVSVLRVPLTQLKLKDLEKMNREREPPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKAVRVEQVQKTGVVVRNHNGIADNATMVRVDVFEKGDYKYLVIYSWQVAKGILPDRVV
QGKDEEDWQLIDDSFNFKFSLHPNDLVEVITTKARMFGYFASCHRG TGNNIRIHDLDHKIGKNGILEGIGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:152)

[0108] 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:

MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIFPEARDPDGTPLNQQRRQKRMRRQLRRRI
RRKALNETLHEAGFLPAYGSADWPVMADEPYELRRRGLEGLSAYEFGRAIYHLAQHRHFKGRELEESDTPDPDV
DDEKEAANERAATLKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVAEEFERLWEVQSKFHPALKSEEMRARI
SDTIFAQRPVFWRKNTLGEGRFMPGEPLCPKGSWLSQRRMLEKLNNLAITAGGNARPLDAEERDAILSKLQQQASMSW
PGVRSALKALYKQRGEPGAEKSLKFNLELGESKLLGNALEAKLADMFGPDWPAHPRKQEI RHAVHERLWAADYGE
TPDKKRVII LSEKDRKAHREAAANSFVADFGITGEQAAQLQALKLPTGWEPYSIPALNFLAELEKGERFGALVNG
PDWEGWRRTNFPHRNQPTGEILDKLPSASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYGKPDRI RIEVGRD
VGKSKREEREI IQSGIRRNEKQRKATEDLIKNGIANPSRDDVEKWILWKEGQERCPTYTGDQIGFNALFREGRYEVE
HIWPRSRSFDNSPRNKTLCKRDVNI EKGNRMPFEAFGHDEDRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMP
EDFAARQLNDTRYAAKQILAQLKRLWPDMPPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDA
LTVACTHPGMTNKL SRYWQLRDDPRAEKPALTPPWDTIRADA EKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDI
KTKSGTYRQFVTRKKIESLSKGELDEIRDPRIKEIVA AHVAGRGGDPKKAFFPPYPCVSPGGPEIRKVRLTSKQQLN
LMAQTGNGYADLGSNHIIAIYRLPDGKADFEIVSLFDASRRLAQRNP IVQRTRADGASFVMSLAAGEAIMIPEGSK

KGIWIVQGVWASGQVVLERDTHSTTRPMPNPILKDDAKKVSIDPIGRVRPSND (SEQ ID NO:153)

[0109] 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
KRRRLQQLDKFIQRQGWPIELEDYSDPLYPWKVRAELAASYIADEKERGEKLSVALRHIARHRGWRNPYAKVSS
LYLPDGPDAFKAIREEIKRASGQVPETATVGQMVTLCELGTLKLRGEGVLSARLQQSDYAREIQEICRMQEIG
QELYRKIIDVVFAAESP KGSASSRVGKDPLQPGKNRALKASDAFQRYRIAALIGNLRVRVDGEKRILSVEEKNLVF
DHLVNLTPKKEPEWVTIAEILGIDRGQLIGTATMTDDGERAGARPPTHDTNRSIVNSRIAPLVDWWKTASALEQHA
MVKALSNAEVDDFSPEGAKVQAFFADLDDDVHAKLDSLHLPGRAAYSEDTLVRLTRRMLSDGVDLYTARLQEFQ
IEPSWTPPTPRIGEPVGNPAVDRVLKTVSRWLESATKTWGAPERVIIIEHVREGFVTEKRAREMDGDMRRRAARNAK
LFQEMQEKLNVQGKPSRADLWRYQSVQRQCAYCGSPITFSNSEMDHIVPRAGQGSTNTRENLVAVCHRCNQSK
GNTPF AIWAKNTSIEGVSVKEAVERTRHWVTDGMRSTDFKKFTKAVVERFQRATMDEEIDARSMESVAWMANELR
SRVAQHFASHGTTVRVYRGS LTAEARRASGISGKLKFFDGVGKSRLDRRHHAIDAAVIAFTSDYVAETLAVRSNLK
QSQAHRQEAPQWREFTGKDAEHRAAWRVWCQKMEKLSALLTEDLRDDR VVMSNVRLRLGNGSAHKETIGKLSKVK
LSSQLSVSDIDKASSEALWCALTREPGFDPKEGLPANPERHIRVNGTHVYAGDNIGLFPVSAGSIALRGGYAELGS
SFHHARVYKITSGKKPAFAMLRVYTIDLLPYRNQDLFSVELKPQTMSMRQAEKKLRDALATGNAEYLGWLVDDEL
VVDTSKIATDQVKA VEALGTIRRWRVDGFFSPSKLRLRLQMSKEG IKKESAPELSKIIDRPGWLPV NKLFS DG
NVTVVRDLSLGRVRLESTAHLPVTWKVQ (SEQ ID NO:154)

[0110] 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:

MTNGKILGLDIGIASVGVIIEAKTGKVVHANSRLFSAANAENNAERRGFRGSRRLNRRKKHRVKRVR
DLFEKYGIVTDFRNLNLPYELRVKGLTEQLKNEELFAALRTISKRRGISYLD DAEDDSTGSTDYAKSIDENRRL
KNKTPGQIQLEKLEKYGQLRGNFTVYDENGEAHLINVFSTSDYEKEARKILETQADYNKKITAEIFIDDYVEILTQ
KRKYHGPNEKSRTDYGRFRDGTTLNIFGILIGKCNFY PDEYRASKASYTAQEYNFLNDLNLKVSTETGKLS
TEQKESLVEFAKNTATLGPALLKEIAKILDCKVDEIKGYREDDKGKPD LHTFEPYRKLKFNLESINIDDL SREVI
DKLADILTLNTEREGIEDAIKRNLPNQFTTEEQISEIIKVRKSQSTAFNKGWHSFSAKLMNELIPELYATSDEQMTI
LTRLEKFKVNKKSSKNTKTIDEKEVTDEIYNPVVAKSVRQTIKIINAAVKKYGDFDKI VIEMPRDKNADDEKKFID
KRNKENKKEKDDALKRAAYLYNSSDKLPDEFVHGNKQLETKIRLWYQQGERCLYSGKPISIQELVHNSNNFEIDHI
LPLSLSFDDSLANKVLVYAWTNQEGGKTPYQVIDSMDAAWSFREMKDYVLKQKGLGKKR DYLLTTENIDKIEVK
KKFIERNLVDTRYASRVVLSLQ SALRELGKDTKVS VVRGQFTSQLRRKWKIDKSRETYHHH AVDALIIAASSQLK
LWEKQDNPMFVDY GKNQVVDKQTGEILSVSDDEYKELVFQPPYQGFVNTISSKGFED EILFSYQVDSKYNRKVSDA
TIYSTRKAKIGKDKKEETYVLGKIKDIYSQNGFDTFIKKYNKDKTQFLMYQKDSL TWENVIEVILRDYPTTKKSED
GKNDVKCNPFEEYRRENGLICKYSKKGKGTPIKSLKYYDKKLGNCIDITPEESRNKVILQSINPWRADVYFNPETL
KYELMGLKYSDLSFEKGTGNYHISQEKYDAIKEGIGKKSEFKFTLYRNDLILIKDIASGEQE IYRFLSRTMPNV
NHYVELKPYDKEKFDNVQELVEALGEADKVGRCIKGLNKPNI SIYKVRTDVLGNKYFVKKKGDKPKLDFKNNKK
(SEQ ID NO:155)

[0111] 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:

MAAFKPNPMNYILGLDIGIASVGWAIVEIDEEENPIRLIDLGV RVFERAEV PKTGDSLAAARRLARSV

RRLTRRRARHLLRARRLLKREGVLQAADFDEGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVDNTHALQTGDFRTPAELALNKFEEKESGHIRNQRGDYSHTFNRKDLQAELNLLFEKQK
EFGNPHVSDGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPTPKAAKNTYTAERFVWLTKLNNLRILEQGSERP
LTDTERATLMDEPYRKSCLTYAQARKLLDLDATAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPL
NLSPQLQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGNRYDEACTEI
YGDHYGKKNTTEKIYLPPIPADEIRNPVVLRLASQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKSAAKFREYFPNPFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLALGSENQNGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDFGKERNLNDTRYINRF
LCQFVADHMLLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTIAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKAHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHKYVTPL
FISRAPNRKMSGQGHMETVKSARKLDEGISVLRVPLTQLKLKDEKMNREREPPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKAVRVEQVQKTGVVWHNHNGIADNATIVRVDVFEKGGKYLYLPIYSWQVAKGILPDRVV
QKGDEEDWTVMDDSFQFVLYANDLIKLTAKKNEFLGYFVSLNRATGAIDIRTHDSTDSTKGKNGIFQSVGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:156)

[0112] 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
RILAKELKLNKYDYAADGELPKAYEGSLASVYELRYKALTQNLETKDLARVILHIAKHRGYMKNKNEKKSNDKKG
KILSALKNNALKLENYQSVGEYFYKEFFQKYKNTKNFIKIRNTKDNYNVCVSSDLEKELKILEKQKEFGYNYS
EDFINEILKVAFFQRPLKDFSHLVGACTFFEEKKRACKNSYSWAFVALTKIINEIKSLEKISGEIVPTQTINEVL
NLILDKGSITYKKFRSCINLHESISFKSLKYDKENAENAKLIDFRKLVEFKKALGVHSLSRQELDQISTHITLIKD
NVKLKTVLEKYNLSNEQINNLEIEFNDYINLSFKALGMILPLMREGKRYDEACEIANLKPKTVDEKKDFLPAFCD
SIFAHELSNPVVNRAISEYRKVLNALLKKYGVHKKIHLELARDVGLSKKAREKIEKEQKENQAVNAWALKECENIG
LKASAKNILKLKLWKEQKEICISYGNKISIEHLKDEKALEVDHIYPYSRSFDDSFINKVLVFTKENQEKLNKTPFE
AFGKNIKWSKIQTALQNLPHYKKKNKILDENFKDKQKQEDFISRNLNDTRYIATLIAKYTKEYLNFLLLSENENANL
KSGEKGSKIHVQTISGMLTSVLRHTWGFDDKDRNNHLHHALDAIIVAYSTNSIIKAFSDFRKNQELLKARFYAKEL
TSDNYKHQVKFEPFKSFREKILSKIDEIFVSKPPRKRARRALHKDTFHSENKIIDKCSYNSKEGLQIALSCGRVR
KIGTKYVENDTIVRVDIFKKQNKFYAIIPIYAMDFALGILPNKIVITGDKNNNPQWQTIDESYEFCSLYKNDLI
LLQKKNMQEPEFAYYNDFSISTSSICVEKHNDKNFENLTSNQKLLFSNAKEGSVKVESLGIQNLKVFEKYIITPLGD
KIKADFQPRENISLKTSKKYGLR (SEQ ID NO:157)

[0113] 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:

MKKEIKDYFLGLDVGTSVGWAVTDTDYKLLKANRKDLWGMRCFETAETAEVRRRLHRGARRRIERRKK
RIKLLQELFSQEIAKTDEGFFQRMKESPFYAEDKTILQENTLFNDKDFADKTYHKAYPTINHLIKAWIENKVKPDP
RLLYLACHNIIKKRGHFLFEGDFSENQFDTSIQALFEYLREDMEVDIDADSQKVKEILKDSSLKNSEKQSRLNKI
LGLKPSDKQKKAITNLISGNKINFADLYDNPDLKDAEKNSISFSKDDFDALSDDLASILGDSFELLLKAKAVYNCS
VLSKVIQDEQYLSFAKVKIYEKHKTDLTKLKNVIKKHFPKDYKKVFGYNKNEKNNNYSYGVGVCKTKSKKLIINN
SVNQEDFYKFLKTIKSAKSEIKEVNDILTEIETGTFLPKQISKSNAEIPYQLRKMELEKILSNAEKHFSFLKQKDE
KGLSHSEKIIMLLTFKIPYYIGPINDNHKKFFPDRCWVVKKEKSPSGKTPWNFFDHIDKEKTAEAFITSRNTNCT

YLVGESVLPKSSLLYSEYTVLNEINNLIIDGKNICDIKLKQKIYEDLFKKYKKITQKQISTFIKHEGICNKTDE
VIILGIDKECTSSLKSYIELKNIFGKQVDEISTKNMLEEIRWATIYDEGEKGTILKTKIKAIEYGYCSDEQIKKI
LNLKFSGWGRLSRKFLFETVTSEMPGFSEPVNIITAMRETQNNLMELLSSEFTFTENIKKINSGFEDAQKQFSYDGL
VKPLFLSPSVKKMLWQTLKLVKEISHITQAPPKKIFIEMAKGALEPARTKTRLKILQDLYNNCKNDADAFSSEIK
DLSGKIENEDNLRRLSDKLYLYYTQLGKCMYCGKPIEIGHVFDTSNYDIDHIYPQSKIKDDISISNRVLCSSCNKN
KEDKYPLKSEIQSKQRGFWNFLQRNNFISLEKLNRLTRATPISDDDETAKFARQLVETRQATKVAKVLEKMPET
KIVYSKAETVSMFRNKFIVKCREINDFHHAHDAYLNIIVGVNVTNTKFTNNPWNFIKEKRDNPKIADTYNYYKVFD
YDVKRNNITAWEEKGTIITVKDMLKRNTPIYTRQAACKKGELFNQTIMKKGLGQHPLKKEGPFNSISKYGGYNKVS
AAYYTLIEYEEKGNKIRSLETIPLYLVKDIQKDQDVLSYLTDLGKKEFKILVPKIKINSLLKINGFPCHITGKT
NDSFLLRPAVQFCCSNNEVLVFKKIRFSEIRSQREKIGKTISPYEDLSFRSYIKENLWKKTKNDEIGEKEFYDLL
QKKNLEIYDMLLTCHKDTIYKKRPN SATIDILVKGKEKFKSLIENQFEVILEILKLSATRNVSDLQHIGGSKYS
GVAKIGNKISSLDNCILYQSITGIFEKRIDLKLV (SEQ ID NO:158)

[0114] 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:

MKKPYSIGLDIGTNSVGWAVVTDDYKVPKMMKVLGNTDKSHIEKNLLGALLFDSGNTAEDRRLKRTA
RRRYTRRRNRILYLQEIFSEEMGKVDDSFHRLSDSFLVTEDKRGERHPIFGNLEEEVKYHENFPTIYHLRQYLAD
NPEKVDLRLVYLALAHIIKFRGHFLIEGKFDRNNVDQRLFQEFLLAVYDNTFENSSSLQEQNQVEEILTDKISKSA
KKDRVLKLPNEKSNGRFAEFLKLVGNQADFKKHFELEEKAPLQFSKDTYEEELVLLAQIGDNYAELFLSAKKL
YDSILLSGILTVDVGTAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSVSKDGYAGYIDGKTNQEAF
YKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEFYPLADNQDRIEKLTFRI
PYVVGPLARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPHSLLYEKFTVYNEL
TKVKYKTEQGKTAFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKNKVFNASYGTYHDL
KILDKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLNYSDDLTKQVKKLERRHYTGWGRLSAELIHGIRNKES
RKTILDYLI DDGNSNRNFMQLINDDALSFKEEIAKAQVIGETDNLNQVVS DIAGSPA IKKGI LQSLKIVDELVKIM
GHQPENIVVEMARENQFTNQGRNSQQLKGLTDSIKEFGSQILKEHPVENSQQLQNDRLFLYYLQNGRDMYTGEEL
DIDYLSQYDIDHIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYWSKLLSAKLITQRKFDNLTKAE
RGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTETDENNKIRQVKIVTLKSNLVSNFRKEFELYKVREIND
YHHAHDAYLNAVIGKALLGVYPQLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKD
EHISNIKKVLSYPQVNIVKKVEEQTGGSKE SILPKGNSDKLIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIE
KGKSKKLKTVKALVGVTIMEKMTFERDPVAFLEKGYRNVQEENIIKLPKYSLFKLENGRKRLLASARELQKGNEI
VLPNHLGTLTYHAKNIHKVDEPKHLDYVDKHKDEFKELLDVVS NFSKKYTLAEGNLEKIKELYAQNNGEDLKELAS
SFINLLTFTAIGAPATFKFFDKNIDRKRYTSTTEILNATLIHQ SITGLYETRIDLNKLGGD (SEQ ID NO:159)

[0115] 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:

MTKPYSIGLDIGTNSVGWAVTTDNYKVPKMMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLKRTA
RRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDSDSFLVPDDKRDSKYP IFGNLVEEKAYHDEFPTIYHLRKYLD
STKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLE
KKDRILKLPGEKNSGIFSEFLKLVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKL
YDAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKYTYNEVFKDDTKNGYAGYIDGKTNQEDF

YVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYFPFLAKNKERIEKILTFRI
PYVVGPLARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPPEEKVLPKHSLLYETFNVYNEL
TKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLLNII
NDKEFLDDSSNEAIIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGLSAKLINGIRDEKSGN
TILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVKVM
GGRKPESIVVEMARENQYTNQGKSNSQQRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKD
MYTGDDLIDRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKRRTFWYQLLKKSLISQRKF
DNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRVTRVKIITLKSTLVSQFRKDFELY
KVREINDFHHAHDAYLNAVVASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKISISLADGR
VIERPLIEVNEETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGPKGLFNANLSSKPKPNSNENLVG
AKEYLDPKKYGGYAGISNSFTVLVKGITIEKGAKKKITNVLEFQGISILDRINRDKLNFLEKGYKDIELIIELP
KYSLFELSDGSRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKFEELFYI
LEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGSEKGLFELTSRGSAADFEFLGVKIPRYRDYTPSS
LLKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO:160)

[0116] 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:

MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLARRKARLNHL
KHLIANEFKLNIEDYQSFDSESLAKAYKGSLSIPYELRFRALNELLSKQDFARVILHIAKRRGYDDIKNSDDKEKGA
ILKAIKQNEEKLANYQSVGEYLKEYFQKFKENSKEFTNVRNKESYERCIASFLKDELKLIFKKQREFGFSFSK
KFEEEVLSVAFYKRALKDFSHLVGNCFFTEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGILYTKDDLNALLN
EVLKNGTLTYKQTKKLLGLSDDYEFKGEKGYFIEFKKYKEFIKALGEHNSQDDLNEIAKDITLIKDEIKLKKAL
AKYDLNQNQIDSLSKLEFKDHLNISFKALKLVTPMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVT
NPVVLRAIKEYRKVLNALLKKYGVHKKINIELAREVGKNHSQRAKIEKEQENYKAKKDAELECEKLGKINSKNI
LKLRLFKKEQKEFCAYSGEKIKISDLQDEKMLEDHIYPYSRSDDSYMNKVLVFTKQKQEKLNQTPFEAFGNDSAK
WQKIEVLAKNLPKKQKRILDKNYKDKQKFNFKDRNLNDTRYIARLVNLYTKDYLDLPLSDDENTKLNQTKGSK
VHVEAKSGMLTSALRHTWGFSKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNK
RKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVKALELGKIRKVNGKIVKNG
DMFRVDIFKHKKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCSLYKDSLILIQTKDMQEP
EFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQRE
DFKK (SEQ ID NO:161)

[0117] 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:

MQTTNLSYILGLDLGIASVGWAVEINENEDPIGLIDVGVRIFERAEPKTTGESLALSRRRLARSTRRL
IRRRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEWGAVLLHLIKHRGYLSKRKNESQTNN
KELGALLSGVAQNHQLLQSDDYRTPAELALKKFAKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCK
EHIQQYMTPELLMWQKPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWTKLNNLRILEDGAERALNEERQL
LINHPYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIRKALENQGLKDTWQDLAKKPDLLD
EIGTAFSLYKTDEDIQQYLTNKPNSVINALLVSLNFDKFIELSLKSLRKILPLMEQKRYDQACREIYGHYGEA
NQKTSQLLPAIPAQEIRNPVVLRTLSQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQKQEDNRTKRES

AVQKFKELFSDFSSEPKSKDILKFRLYEQQHGKCLYSGKEINIHRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLA
SENQNKGNQTPYEWLQGKINSERWKNFVALVLGSQCSAAKKQRLLTQVIDDNKFIDRNLNDTRYIARFLSNYIQEN
LLLVGKNKKNVFTPNGQITALLRSRWGLIKARENNNRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENTRYE
MVDQESGEIISPHFPEPWAYFRQEVNIRVFDNHPDVTVLKEMPLDRPQANHQFVQPLFVSRAPTRKMSGQGHMETIK
SAKRLAEGISVLRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPAKAFATPFYKQGGQVKAIRVEQVQ
KSGVLVRENGVADNASIVRTDVFIKNNKFFLVPIYTWQVAKGILPNKAIVAHKNEDEWEEMDEGAFFKFSLFPN
LVELKTKKEYFFGYIIGLDRATGNISLKEHDGEISKGKGVYRVGVKLALSFEKYQVDELGKNRQICRPQQRQPV
(SEQ ID NO:162)

[0118] 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:

MNFKILPIAIDLGVKNTGVFSAFYQKGTSLERLDNKNKGVYELSKDSYTLLMNNRTARRHQRRGIDRK
QLVKRLFKLIWTEQLNLEWDKDTQQAISFLFNRRGFSFITDGYSPEYLNIVPEQVKAILMDIFDDYNGEDDLDSYL
KLATEQESKISEIYNKLMQKILEFKLMKLCTDIKDDKVSTKTKEITSYEFELLADYLANYESLKTQKFSYTDKQ
GNLKELSYHHDKYNIQEFLLKRHATINDRILDTLLTDDLDIWNFNFEKFDKNEEKLQNQEDKDHIQAHLLHVFV
AVNLIKSEMASGGRHRSQYFQEITNVLDENNHQEGYLKNFCENLHNKKYSNLSVKNLVNLIGNLSNLELKPLRKYF
NDKIHAKADHWDEQKFTETTYCHWILGEWRVGVKDQDKKDGAKYSYKDLCNELKQKVTAGLVDFLLELDPCRTIPP
YLDNNNRKPPKCQSLILNPKFLDNQYPNWQQYLQELKKLSIQNYLDSFETDLKVLKSSKDQPYFVEYKSSNQQA
SGQRDYKDLARILQFIFDRVKASDELLNEIYFQAKKLKQKASSELEKLESSKKLDEVIANSQLSQILKSQHTNG
IFEQGTFLHLVCKYYYKQRQARDSRLYIMPEYRYDKKLHKYNNTRGFDDDNQLLTYCNHKPRQKRYQLLNDLAGVL
QVSPNFKDKIGSDDDLFISKWLVEHIRGFKKACEDSLKIQKDNRGLLNHKINIARNTKGKCEKEIFNLICKIEGS
EDKKGNKXKHLAYELGVLLFGEPNEASKPEFDRKIKKFNSIYSFAQIQQAFAERKGNANTCAVCSADNAHRMQQI
KITEPVEDNKDKIILSAKAQRLPAIPTRIVDGAVKMATILAKNIVDDNWQNIKQVLSAKHQLHIPIITESNAFEF
EPALADVKGKSLKDRRKALERISPENIFKDKNNRIKEFAKGISAYSGANLTDGDFDGAKEELDHIIPRSHKKYGT
LNDEANLICVTRGDNKNKGNRIFCLRDADNYKLKQFETTDDLEIEKKIADTIWDANKKDFKFGNYRSFINLTPQE
QKAFRHALFLADENPIKQAVIRAINNRNRTFVNGTQRYFAEVLANNIYLAKKENLNTDKISFDYFGIPTIGNRG
IAEIRQLYEKVDSDIQAYAKGDKPQASYSHLIDAMLAFCIAADEHRNDGSIGLEIDKNYSLYPLDKNTGEVFTKDI
FSQIKITDNEFSDDKLVRKKAIEGFNTHRQMTDRGIIYAENYLPILIHKELNEVRKGYTWKNSEEIKIFKGKKYDIQ
QLNNLVYCLKFVDKPIISIDIQISTLEELRNILTTNIAATAEYYYINLKTQKLHEYYIENYNTALGYKKYSKEMEF
LRSLAYRSERVKIKSIDDVKQVLDKDSNFIIGKITLPFKKEWQRLYREWQNTTIKDDYEFLKSFFNVKSITKLHKK
VRKDFSLPISTNEGKFLVKKRTWDNNFIYQILNDSDSRADGTPFIPAFDISKNEIVEAIDSFTSKNIFWLPKNI
ELQKVDNKNIFAIDTSKWFEVETPSDLRDIGIATIQYKIDNNSRPKVRVKLDYVIDDDSKINYFMNHSLLKSRYPD
KVLEILKQSTIIEFESSGFNKTIKEMLGKLAGIYNETSNN (SEQ ID NO:163)

[0119] 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:

MKVNNYHIGLDIGTSSIGWVAIGKDGKPLRVKGKTAIGARLFQEGNPAADRRMFRTTRRRLSRRKWRL
KLEEIFDPYITPVDSTFFARLKQSNLSPKDSRKEFKGSMLFPDLTDMQYHKNYPTIYHLRHALMTQDKKFDIRMV
YLAIHHIVKYRGNFLNSTPVDSFKASKVDFVDQFKKLNELYAAINPEESFKINLANSEDIGHQFLDPSIRKFDKKK
QIPKIVPVMNDKVTDRNLNGKIASIIEHAILGYKAKLDVVLQCTPVDSKPPWALKFDDDEDIDAKLEKILPEMDENQQ
SIVAILQNLYSQVTLNQIVPNGMSLSESMIEKYNDHHDHLKLYKKLIDQLADPKKKAVLKKAYSQYVGGDGKIVIEQ

AEFWSSVKKNLDDSELSKQIMDLIDAEKFMKQRTSQNGVIPHLHQRELDEIEHQSKYYPWLVEINPNKHDHL
AKYKIEQLVAFRVPYVVGPMITPKDQAESAETVFSWMERKGTETGQITPWNFDEKVDKASANRFIKRMTTKDYL
IGEDVLPDESLLYEKFKVLNLMVRVNGKLLKVADKQAIQDLFENYKHVSVKKLQNYIKAKTGLPSDPEISGLS
DPEHFNNSLGTYNDFKKLFGSKVDEPDLQDDFEKIVEWSTVFEDKKILREKLNEITWLSDDQKQDVLESSRYQGWGR
LSKKLLTGIVNDQGERIIDKLWNTNKNFMQIQSDDDFAKRIHEANADQMCAVDVEDVLADAYTSPQNKKAIRQVVK
VVDDIQKAMGGVAPKYISIEFTRSEDRNPRRTISRQRQLENTLKDTAKSLAKSINPELLSELDNAASKKGLTDRL
YLYFTQLGKDIYTGEPINIDELNKYDIDHILPQAFIKDNSLDNRVLVLTAVNNGKSDNVPLRMFGAKMGHFWKQLA
EAGLISKRKLKNLQTDPTISKYAMHGFIRRLVETSQVIKLVANILGDKYRNDDTKIEITARMNHQMRDEFGFI
KNREINDYHHAFFDAYLTAFLGRYLYHRYIKLRPFYVYGDFKKFREDKVTMRNFNFLHDLTDDTQEKIADAETGEVI
WDRENSIQQLKDVYHYKFMILISHEVYTLRGAMFNQTVYPASDAGKRKLIPVKADRPVNVYGGYSGSADAYMAIVRI
HNKKGDKYRVVGVPMRALDRLDAAKNVSDADFDRALKDVLAPQLTKTKSRKTGEITQVIEDFEIVLGKVMYRQLM
IDGDKKFMGLSSTYQYNAKQLVLSQSVKTLASKGRLDPLQESMDYNNVYTEILDKVNQYFSLYDMNKRFRHKLNLG
FSKFISFPNHNVLGNTKVSSGKREILQEILNGLHANPTFGNLKDVGITTPFGQLQQPNGILLSDETKIRYQSPTG
LFERTVSLKDL (SEQ ID NO:164)

[0120] 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:

MKKPYTIGLDIGTNSVGWAVLTDQYDLVKRKMKIAGDSEKKQIKKNFWGVRLFDEGQTAADRRMARTA
RRRIERRRNRI SYLQGI FAEEMSKTDANFFCRLSDSFYVDNEKRNSRHPFFATIEEEVEYHKNYPTIYHLREELVN
SSEKADLRLVYLALAHIIKYRGNFLIEGALDTQNTSVDGIYKQFIQTYNQVFASGIEDGSLKKLEDNKDVAKILVE
KVTRKEKLERILKLYPGEKSAGMFAQFISLIVGSKGNFQKPFDLIEKSDIECAKDSYEEDLESLLALIGDEYAEFL
VAAKNAYSAVVLSSIIITVAETETNAKLSASMIERFDTHEEDLGELKAFIKLHLPKHYYEIFSNTKEKHGYAGYIDGK
TKQADFYKYMKMTLENIEGADYFIAKIEKENFLRKQRTFDNGAIPHQLHLEEELEAILHQQAKYYPFLKENYDKIKS
LVTFRIPIYFVGPLANGQSEFAWLTRKADGEIRPWNIEEKVDFGKSAVDFIEKMTNKDYLKENVLPKHSLEYQKY
LVYNELTKVRYINDQGKTSYFSGQEKEQIFNDLFKQKRKVKKDLELFLRNMSHVESPTIEGLEDSFNSSYSTYHD
LLKVGIKQEILDNPVNTEMLENIVKILTVFEDKRMIKEQLQQFSDVLDGVVLKKLERRHYTGWGRLSAKLLMGIRD
KQSHLTILDYLMNDDGLNRNLMQLINDSNLSFKSIIIEKEQVTTADKDIQSIVADLAGSPAIIKKGILQSLKIVDELV
SVMGYPPQTIVVEMARENQTTGKGKNSRPRYKSLEKAIKEFGSQILKEHPTDNQELRNRLYLYLQNGKDMYTG
QDLDIHNLSNYDIDHIVPQSFITDNSIDNLVLTSSAGNREKGGDVPPLEIVRKRKVFWEKLYQGNLSKRKFDYLT
KAERGGLTEADKARFIHRQLVETRQITKNVANILHQRFNIEKDDHGNTMKQVRIVTLKSALVSQFRKQFQLYKVRD
VNDYHHAHDAYLNGVVANTLLKVYPQLEPEFVYGDYHQFDWFKANKATAKKQFYTNIMLFFAQKDRIIDENGEILW
DKKYLDTVKKVMSYRQMNIVKKTEIQKGEFSKATIKPKGNSSKLIPRKTNWDPMKYGGLDSPNMAYAVVIEYAKGK
NKLVFEEKIIRVTIMERKAFKDEKAFLEEQGYRQPKVLAKLPKYTLYECEEGRRRMLASANEAKGNQQVLPNHL
VTLLHHAANCEVSDGKSLDYIESNREMFAELLAHVSEFAKRYTLAEANLNKINQLFEQNKEGDIKAIASQSFVDLMA
FNAMGAPASFKFFETTIERKRYNNLKELLNSTIIYQSITGLYESRKRLDD (SEQ ID NO:165)

[0121] 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:

MESSQILSPIGIDLGGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRVRNKK
RNQFVKRVALQLFQHILSRDLNAKEETALCHYLNRRGYTYVDTDLDEYIKDETTINLLKELLPSESEHNFIWFLQ
KMQSSEFRKILVSKVEEKDDKELKNAVKNIKNFITGFEKNSVEGHRHRKVYFENIKSDITKDNQLDSIKKKIPSV

CLSNLLGHLNQLQWKNLHRYLAKNPKQFDEQTFGNEFLRMLKNFRHLKGSQESLAVRNLIQQLEQSQDYISILEKT
PPEITIPPYEARTNTGMEKDQSLLLNPEKLNLYPNWRNLIPGIIDAHPFLEKDLEHTKLRDRKRIISPSKQDEKR
DSYILQRYLDLNNKIDKFKIKKQLSFLGQKQLPANLIETQKEMETHFNSSLVSVLIQIASAYNKEREDAAQGIWF
DNAFSLCELSNINPPRKQKILPLLVGAILSEDFINNKKDWAKFKIFWNTHKIGRTSLKSKCKEIEEARKNSGNAFK
IDYEEALNHPEHSNNKALIKIIQTIPDIIQAIQSHLGHNSQALIIYHNPFSLSQLYITILETKRDGFHKNCVAVTCE
NYWRSQKTEIDPEISYASRLPADSVRPFDFGLARMMQRLAYEIAMAKWEQIKHIPDNSSLLIPIYLEQNRFEFEES
FKKIKGSSSDKTLEQAIEKQNIQWEEKFQRIINASMNICPYKGASIGGQGEIDHIYPRSLSKKHFGVIFNSEVNLI
YCSSLQGNREKKEEHYLLLEHLSPLYLKHQFGTDNVSDIKNFISQNVANIKKYISFHLLTPEQQKAARHALFLDYDDE
AFKTITKFLMSQQKARVNGTQKFLGKQIMEFLSTLADSKQLQLEFSIKQITAEVVDHRELLSKQEPKLVKSRQSS
FPSHAIDATLTMSIGLKEFPQFSQELDNSWFINHLMPDEVHLNPVRSKEKYKNPNISSTPLFKDSLAEFIPVWV
KGETFAIGFSEKDLFEIKPSNKEKLFLLKTYSTKNPGESLQELQAKSKAKWLYFPINKTLALEFLHHYFHKEIVT
PDDTTVCHFINSRLYYTKKESITVKILKEPMPVLSVKFESSKKNVLGSFKHTIALPATKDWERLFNHPNFLALKAN
PAPNPKEFNEFIRKYFLSDNNPNSDIPNNGHNKPKQKHAKVRKVFSLPVIPGNAGTMMRIRKDNKGQPLYQLQTI
DDTPSMGIQINEDRLVKQEVLMDAYKTRNLSTIDGINNSEGQAYATFDNWLTPVSTFKPEIIKLEMKPHSKTRRY
IRITQSLADFIKTIDEALMIKPSDSIDDPLNMPNEIVCKNKLFGNELKPRDGMKIVSTGKIVTYEFESDSTPQWI
QTLYVTQLKKQP (SEQ ID NO:166)

[0122] 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:MAAFKPNPMNYILGLDIGIASVGWAMVEVDEEENPIRLIDL
GVRVFERAEPKTDGSLAMARRLARSVRRLTRRRARHLLRARRLLKREGVLQDADFDENGLVKSLPNTPWQLRAAA
LDRKLTCLWSAVLLHLVKHRGYLSQRKNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKSGHIR
NQRGDYSHTFSRKDLQAEINLLFEKQKEFGNPHVSDGLKEDIETLLMAQRPALSGDAVQKMLGHCTFEPAEPKAAK
NTYTAERFIWLTKLNNLRILEQGSRPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEAS
TLMEMKAYHAISRLEKEGLKDKKSPLNLSTELQDEIGTAFSLFKTDKIDITGRLKDRVQPEILEALLKHISFDKQV
QISLKALRRIVPLMEQGKRYDEACAEIYGDHYCKKNAEEKIYLPPIPADEIRNPVVLRLALSQARKVINCVVRRYGS
PARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAFREYFPNFVGEPSKDILKLRLYEQQHKGCLYSGKEI
NLVRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKK
QRILLQKFDEEGFKERNLNDTRYVNRFLCQFVADHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRTENDRRHAL
DAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQAHPFPQWEFFAQEVMIRVFGKPDGKPEFEEAD
TPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETVKSARKLDEGISVLRVPLTQKLKGLEKVMN
REREPKLYDALKAQLETHKDDPAKFAEPFYKYDKAGSRTQQVKAVRIEQVQKTGVWVRNHNGIADNATMVRVDVF
EKGGKYYLVPIYSWQVAKGILPDRAVFAFKDEEDWTVMDDSFEFRFVLYANDLIKLTAKKNEFLGYFVSLNRATGA
IDIRTHDSTDSTGKNGIFQSVGVKTALSFKQNKIDELGKEIRPCRLKKRPPVR (SEQ ID NO:167)

[0123] 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:

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRFERAEPKTDGSLAMARRLARS
RRLTRRRARHLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPEWSAVLLHLIKHRGYLSQR
KNEGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEEKSGHIRNQRGDYSHTFSRKDLQAEILLFEKQK
EFGNPHVSGGLKEGIIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSER
PLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRLEKEGLKDKKSPL

NLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEI
YGDHYGKKNTTEEKIYLPPIPADEIRNPVVLRLASQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQE
ENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLEKGYVEIDHALPFSRTWDDSF
NNKVLVLGSENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDEGFKERNLNDTRYVNR
LCQFVADRMRLTGKGGKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNA
FDGKTIDKETGEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRLLAEKLSSRPEAVHEYVTPL
FVSRAPNRKMSGQGHEMETVKSARKLDEGVSVLRVPLTQLKLKDLEKMNREREPPLYEALKARLEAHKDDPAKAF
EPFYKYDKAGNRTQQVKAVRVEVQVKTGVWVRNHNGIADNATMVRVDVFEKGDKYLVPIYSWQVAKGILPDRVV
QKGDEEDWQLIDDSFNFKFSLHPNDLVEVITTKARMFYGFASCHRG TGNNIRIHDL DHKIGKNGILEGIGVKTAL
SFQKYQIDELGKEIRPCRLKKRPPVR (SEQ ID NO:168)

[0124] 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:MLSRQLLGASHLARPVSYSYNVQDNDVHCSYGERCFMRGKRYRIGID
VGLNSVGLAAVEVSDENSPVRLNAQSVIHGGVDPQKNKEAITRKNMSGVARRTRMRKRERLHKLDMLLGKF
GYPVIEPESLDKPFEEWHVRAELATRYIEDDELRRRESISIALRHMARHRGWRNPYRQVDSLISDNPYSKYGELKE
KAKAYNDATAAEESTPAQLVVAMLDAGYAEAPRLRWRTGSKKPDAEGYLPVRLMQEDNANELKQIFRVQRPAD
EWKPLFRSVFYAVSPKGSAEQRVGQDPLAPEQARALKASLAFQEYRIANVITNLRIKDASAE LRKLTVD EKQSIYD
QLVSPSSEDITWSDLCDFLGFKRSQ LKGVGSLTEDGEERISSRPRLTSVQRIYESDNKIRKPLVAWWKSASDNEH
EAMIRLLSNTVDIDKVREDVAYASAEFIDGLDDALTKLDSVDLPSGRAAYSVETLQKLTRQMLTTDDDLHEARK
TLFNVTDSWRPPADPIGEPLGNPSVDRVLKVNRYLMNCQQRWGNPVSVNIEHVRSSSFSSVAFARKDKREYEKNE
KRSIFRSSLSEQLRADEQMEKVRESDLRRLAEAIQRQNGQCLYCGRTITFRTCMDHIVPRKGVGSTNTRTNFAAVC
AECNRMKSNTPFATWARSEDAQTRGVSLAEAKKRVMTFTFNPKSYAPREVKAFAKQAVIARLQQTEDDAAIDNRSIE
SVAWMADLHRRIDWYFNAKQYVNSASIDDAEATMKTTVSFVQGRVTASARRAAGIEGKIHF IGQQSKTRLDRRH
HAVDASVIAMMNTAAQTLMERESLRESQRLIGLMPGERSWKEYPYEGTSRYESFHLWLDNMDVLELLNDALDND
RIAVMQSQRYVLGNSIAHDATIHPLEKVPLGSAMSADLIRRASTPALWCALTRLPDYDEKEGLPEDSHREIRVHDT
RYSADDEMGGFASQAAQIAVQEGSADIGSAIHARVYRCWKTNAKGVRKYFYGMIRVFQTDLLRACHDDLFTVPLP
PQSISMRYGEPRVVQALQSGNAQYLGSLVVGDEIEMDFSSLDVDGQIGEYLQFFSQFSGGNLAWKHWVDGFFNQ
TLRIRPRYLAAEGLAKAFSDDVVPDGVQKIVTKQGWLPVNTASKTAVRIVRRNAFGEPRLSSAHMPCSWQRHE
(SEQ ID NO:169)

[0125] 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:

MSRSLTFSFDIGYASIGWAVIASASHDDADPSVCGCTVLFPKDDCQAFKRREYRRLRRNIRSRRVRI
ERIGRLLVQAQIITPEMKETSGHPAPFYLAASEALKGHRTLAPIELWHVLRWYAHNRGYDNNASWSNSLSEGGNGE
DTERVKHAQDLMDKHGTATMAETICRELKLEEGKADAPMEVSTPAYKNLNTAFPRLIVEKEVRRILELSAPLIPGL
TAEIIELIAQHHPLTTEQRGVLLQHGIKLARRYRGSLLFGQLIPRFDNRIISRCPVTAQVYEAELKKGNSEQSAR
ERAEKLSKVPTANCPEFYEYRMARILCNIRADGEPLSAEIRRELMNQARQEGKLTASLEKAISSRLGKETETNVS
NYFTLHPDSEEALYNPAVEVLQRSGIGQILSPSVYRIANRLRRGKSVTPNYLLNLLKSRGESGEALEKKIEKES
KKKEADYADTPLKPKYATGRAPYARTVLKKVVEEILDGEDPTRPARGEAHPDGELKAHDGCLYCLLDTDSSVNQH
KERRLDTMTNNHLVRHRMLILDRLLKDLIQDFADGQKDRISRVCEVVGKELTTFSAMDSKKIQRELTLRQKSHTDA
VNRLKRKLPGKALSANLIRKCRIAMDMNWTCPFTGATYGDHELENLELEHIVPHSFRQSNALSSLVLTWPGVNRMK

GQRTGYDFVEQEENPVPDPKPNLHICSLNNYRELVEKLDDKKGHEDDRRRKKKRKALLMVRGLSHKHQSQNHEAMK
EIGMTEGMMTQSSHLMKLACKSIKTSPLDAHIDMIPGAVTAEVKAWDVFGVFKELCPEAADPDGKILKENLRSL
THLHHALDACVLGLIPYIIPAHHNGLLRVLAMRRIPEKLIPQVRPVANQRHYVLNDDGRMMLRDLASLKENIRE
QLMEQRVIQHVPADMGGALLKETMQRVLSVDGSGEDAMVSLSKKKDGKKEKNQVKASKLVGVFPEGPSKLKALKAA
IEIDGNYGVALDPKPVIRHIKVFKRIMALKEQNGGKPVIRILKKGMLIHLTSSKDPKHAGVWRIESIQDSKGGVKL
DLQRAHCAVPKNKTHECNWREVDLISLLKKYQMKRYPTSYTGTPR (SEQ ID NO:170)

[0126] 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
KKLRKAKLLELLIAYDMCPLKPEDVRRWKNWDKQQKSTVRQFPDTPAFREWLKQNPYELRKQAVTEDVTRPELGRI
LYQMIQRRGFLSSRKKEEGKIFTGKDRMVGIDETRKNLQKQTLGAYLYDIAPKNGEKYRFRTERVRARYTLRDMY
IREFEIIWQRQAGHLGLAHEQATRKNIFLEGSATNVRNSKLITHLQAKYGRGHVLIEDTRITVTFQLPLKEVLGG
KIEIEEEQLKFKNESVLFWRPLRSQKSLSSKCVFEGRNFYDPVHQWIIAGPTPAPLSHPEFEEFRAYQFINNI
IYGKNEHLTAIQREAVFELMCTESKDFNFEKIPKHLKLFKFNFDTTKVPACTTISQLRKLFPHPVWEEKREEIW
HCFYFYDDNTLLFEKLQKDYLQTNDELEKIKIRLSESYGNVSLKAIRRNINPYLKKGYAYSTAVLLGGIRNSFGKR
FEYFKEYEPEIEKAVCRILKEKNAEGEVIRKIKDYLVHNRFGFAKNDRAFQKLYHHSQAITTQAQKERLPETGNLR
NPVQQGLNELRRRTVNKLLATCREKYGPSFKFDHIHVMGRELRSSKTEREKQSRQIRENEKKNEAAKVLAEYGL
KAYRDNIQKYLLYKEIEEKGGTVCCPYTGKTLNISHTLGSDNSVQIEHIIPYSISLDDSLANKTLCDATFNREKGE
LTPYDFYQKDPSPKKGASSWEEIEDRAFRLLPYAKAQRFIRRKPQESNEFISRQLNDTRYISKKAWEYLSAICSD
VKAFFGQLTAELRHLWGLNNILQSAPDITFPLPVSATENHREYYVITNEQNEVIRLFPKQGETPRTEKGELLTGE
VERKVFRCKGMQEFQTDVSDGKYWRIKLSSSVTWSPLFAPKPIADGQIVLKGRIEKGVFVCNQLKQKLKTGLPD
GSYWISLPVISQTFKEGESVNNSKLTSSQVQLFGRVREGIFRCHNYQCPASGADGNFWCTLDTDTAQPAFTPIKNA
PPGVGGGQIILTGDVDDKGIHADDLHYELPASLPKGKYYGIFTVESCDPTLIPIELSAPKTSKGENLIEGNIWV
DEHTGEVRFPKKNREDQRHHAIDAIVIALSSQSLFQRLSTYNARRENKKRGLDSTEHPSPWPFGAQDVRQSVVP
LLVSYKQNPKTLCISKTLTKDGGKIHSCGNAVVRGQLHKETVYGGRTAPGATEKSYHIRKDIRELKTSHKIGKVVD
ITIRQMLLKHLQENYHIDITQEFNIPSNAFFKEGVYRIFLPNKHGEPVPIKKIRMKEELGNAERLKDNIQYVNP
NNHHVMIYQDADGNLKEEIVSFWSVIERQNQGQPIYQLPREGRNIVSILQINDTFLIGLKEEEPEVYRNDLSTLSK
HLYRVQKLSGMYTFRHHLASTLNNEREEFRIQSLEAWKRANPVKVQIDEIGRITFLNGPLC (SEQ ID NO:
171)。

[0127]

[0128] 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蛋白或其部分分离或衍生的序列。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。

[0129] 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 YLSKTLRFE LIPQGTLEN IKARGLILDD EKRAKDYKKA KQIIDKYHQF
61 FIEEILSSVC ISEDLLQNY S DVYFKLKKSD DDNLQKDFKS AKDTIKKQIS EYIKDSEKFK
121 NLFNQNLIDA KKGQESDLIL WLKQSKDNGI ELFKANS DIT DIDEALEIIK SFGWTTYFK
181 GFHENRKNVY SSNDIPT SII YRIVDDNLPK FLENKAKYES LKDKAPEAIN YEQIKKDLAE
241 ELTFDIDYKT SEVNQRV FSL DEVFEIANFN NYLNQSGITK FNTIIGGKFV NGENTKRKGI
301 NEYINLYSQQ INDKTLKKYK MSVLFKQILS DTESKSFVID KLEDDSDVVT TMQSFYEQIA
361 AFKTVEEKSI KETLSLLFDD LKAQKLDLSK IYFKNKSLT DLSQQVFDDY SVIGTAVLEY
421 ITQQIAPKNL DNPSKKEQEL IAKKTEKAKY LSLETIKLAL EEFNKHRDID KQCRFEEILA
481 NFAAIPMIFD EIAQNKD NLA QISIKYQNGG KKDLLQASAE DDVKA IKDLL DQTNNLLHKL
541 KIFHISQSED KANILDKDEH FYLVFEECYF ELANIVPLYN KIRNYITQKP YSDEKFKLNF
601 ENSTLANGWD KNKEPDNTAI LFIKDDKY YL GVMNKKNNKI FDDKA IKENK GEGYKKIVYK
661 LLPGANKMLP KVFFSAKSIK FYNPSEDILR IRNHSTHTKN GSPQKG YEF EFNIEDCRKF
721 IDFYKQSISK HPEWKDFGFR FSDTQRYNSI DEFYREVENQ GYKLTFENIS ESYIDSVVNQ
781 GKLYLFQIYN KDFSAYS KGR PNLHTLYWKA LFDERNLQDV VYKLNGEAE L FYRKQSIPKK
841 ITHPAKEAIA NKNKDNPKKE SVFEYDLIKD KRFTEDKFFF HCPITINFKS SGANKFNDEI
901 NLLLKEKAND VHILSIDRGE RHLAYYTLVD GKGNI IKQDT FNIIGNDRMK TNYHDKLAAI
961 EKDRDSARKD WKKINNIKEM KEGYLSQVVH EIAKL VIEYN AIVVFEDLNF GFKRGRFKVE
1021 KQVYQKLEKM LIEKLN YLVF KDNEFDKTGG VLRAYQLTAP FETFKKMGKQ TGIIYYVPAG
1081 FTSKICPVTG FVNQLYPKYE SVSKSQEFFS KFDKICYNLD KGYFEFSFDY KNFGDKAAKG
1141 KWTIASFGSR LINFRNSDN HNWDTREVPY TKELEKLLKD YSIEYGHGEC IKA AICGESD
1201 KKFFAKLTSV LNTILQMRNS KTGTELD YLI SPVADVNGNF FDSRQAPKNM PQDADANGAY
1261 HIGLKGLMLL GRIKNNQEGK KLNLVIKNEE YFEFVQNRNN (SEQ ID NO: 172)。

```

[0130] 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 AASKLEFTN CYLSKTLRF KAIPVGKTQE NIDNKRLLE DEKRAEDYKG VKLLLD RYYL
61 SFINDVLHSI KLKNLN NYIS LFRKKTRTEK ENKELENLEI NLRKEIAKAF KGAAGYKSLF
121 KKDIIETILP EAADDKDEIA LVNSFNGFTT AFTGFFDNRE NMFSEEAKST SIAFR CINEN
181 LTRYISNMDI FEKVDAIFDK HEVQEIKEKI LNSDYDVEDF FEGEFFNFVL TQEGIDVYNA
241 IIGGFVTESG EKI KGLNEYI NLYNAKTKQA LPKF KPLYKQ VLSDRSLSF YGEGYTSDEE
301 VLEVFRNTLN KNSEIFSSIK KLEKLFKNFD EYSSAGIFVK NGPAISTISK DIFGEWNLIR

```

361 DKWNAEYDDI HLKKKAVVTE KYEDDRRKSF KKIGSFSLEQ LQEYADADLS VVEKLKEIII
421 QKVDEIYKVY GSSEKLFDAD 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 GGAELFMRRR SLKKEELVH PANSPIANKN PDNPKKTTTL
781 SYDVYKDKRF SEDQYELHIP IAINKCPKNI FKINTEVRVL LKHDDNPYVI GIDRGERNLL
841 YIVVVDGKGN IVEQYSLNEI INNFNIGIRIK TDYHSLDDKK EKERFEARQN WTSIENIKEL
901 KAGYISQVVH KICELVEKYD AVIALEDLNS GFKNSRVKVE KQVYQKFEKM LIDKLNVMVD
961 KKSNPCATGG ALKGYQITNK FESFKSMSTQ NGFIFYIPAW LTSKIDPSTG FVNLLKTKYT
1021 SIADSKKFIS SFDRIMYVPE EDLFEFALDY KNFSRTDADY IKKWKLYSYG NRIRIFAAAK
1081 KNNVFAWEEV CLTSAYKELF NKYGINYQQG DIRALLCEQS DKAFYSSFMA LMSLMLQMRN
1141 SITGRDQVDF LISPVKNSDG IFYDSRNYEA QENAILPKNA DANGAYNIAR KVLWAIGQFK
1201 KAEDEKLDKV KIAISNKEWL EYAQTSVK (SEQ ID NO: 173)。

[0131] 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 HRIVQDNFPK FKENCHIFTR LITAVPSLRE HFENVKKAIG IFVSTSIEEV
241 FSFPFYNQLL TQTQIDLYNQ LLGGISREAG TEKIKGLNEV LNLAIQKNDE TAHIIASLPH
301 RFIPLFKQIL SDRNTLSFIL EEFKSDEEVI QSFCKYKTLL RNENVLETAE ALFNELNSID
361 LTHIFISHKK LETISSALCD HWDTLRNALY ERRISELTGK ITKSAKEKVQ RSLKHEDINL
421 QEIIISAAGKE LSEAFKQKTS EILSHAAAL DQPLPTTLKK QEEKEILKSQ LDSLLGLYHL
481 LDWFAVDESN EVDPFESARL TGIKLEMEPS LSFYNKARNY ATKKPYSVEK FKLNFQMPTL
541 ASGWDVNKEK NNGAILFVKN GLYYLGIMPK QKGRYKALSF EPTEKTSEGF DKMYDYFPD
601 AAKMIPKCST QLKAVTAHFQ THTTPILLSN NFIEPLEITK EIYDLNNPEK EPKKFQTAYA
661 KKTGDQKGYR EALCKWIDFT RDFSLSKYTKT TSIDLSSLRP SSQYKDLGEY YAEINPLLYH
721 ISFQRIAEKE IMDAVETGKL YLFQIYNKDF AKGHHGKPNL HTLYWTGLFS PENLAKTSIK
781 LNGQAELFYR PKSRMKRMAH RLGEKMLNKK LKDQKTPIPD TLYQELYDYV NHRLSHDLSD

841 EARALLPNVI TKEVSHEIIK DRRFTSDKFF FHVPIITLNYQ AANSPSKFNQ RVNAYLKEHP
 901 ETPIIGIDRG ERNLIYITVI DSTGKILEQR SLNTIQQFDY QKKLDNREKE RVAARQAWSV
 961 VGTIKDLKQG YLSQVIHEIV DLMIHQAVV VLENLNFQFK SKRTGIAEKA VYQQFEKMLI
 1021 DKLNCLVLKD YPAEKVGGVL NPYQLTDQFT SFAKMGTQSG FLFYVPAPYT SKIDPLTGFV
 1081 DPFVWKTIKN HESRKHFLG FDFLHYDVKT GDFILHFKMN RNLSFQRGLP GFMPAWDIVF
 1141 EKNETQFQAK GTPFIAGKRI VPVIENHRFT GRYRDLYPAN ELIALLEEKQ IVFRDGSNIL
 1201 PKLLENDSSH AIDTMVALIR SVLQMRNSNA ATGEDYINSP VRDLNGVCFD SRFQNPWPM
 1261 DADANGAYHI ALKGQLLLNH LKESKDLKLQ NGISNQDWLA YIQELRN (SEQ ID NO: 174)。

[0132] 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*) DSM 10710、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FSLM6-0635 (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。

[0133] 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>Leptotrichi</i>)	WP_018451595.1	CCACCCCAATATCGAAGGGGAC TAA AAC (SEQ ID NO: 175)

		a shahii)		
Cas13a2	LwaCas13a	韦德纤毛菌	WP_021746774.1	GATTTAGACTACCCCAAAAACG AAGGGGACTAAAAC (SEQ ID NO: 176)
Cas13a3	LseCas13a	西尔李斯特 菌	WP_012985477.1	GTAAGAGACTACCTCTATATGAA AGAGGACTAAAAC (SEQ ID NO: 177)
Cas13a4	LbmCas13a	毛螺科细菌 MA2020	WP_044921188.1	GTATTGAGAAAAGCCAGATATA GTTGGCAATAGAC (SEQ ID NO: 178)
Cas13a5	LbnCas13a	毛螺科细菌 NK4A179	WP_022785443.1	GTTGATGAGAAGAGCCCAAGAT AGAGGGCAATAAC (SEQ ID NO: 179)
Cas13a6	CamCas13a	嗜氨基[梭菌] DSM 10710	WP_031473346.1	GTCTATTGCCCTCTATATCGGGC TGTTCTCCAAAC (SEQ ID NO: 180)
Cas13a7	CgaCas13a	鸡肉杆菌 DSM 4847	WP_034560163.1	ATTAAAGACTACCTCTAAATGTA AGAGGACTATAAC (SEQ ID NO: 181)
Cas13a8	Cga2Cas13a	鸡肉杆菌 DSM 4847	WP_034563842.1	AATATAAACTACCTCTAAATGTA AGAGGACTATAAC (SEQ ID NO: 182)
Cas13a9	Pprcas13a	产丙酸沼杆 菌WB4	WP_013443710.1	CTTGTGGATTATCCCAAATTGA AGGGA ACTACAAC (SEQ ID NO: 183)
Cas13a10	LweCas13a	韦氏李斯特 菌 FSL R9-0317	WP_036059185.1	GATTTAGAGTACCTCAAAATAG AAGAGGTCTAAAAC (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	GATTTAGAGTACCTCAAAACAA AAGAGGACTAAAAC (SEQ ID NO: 185)
Cas13a12	Lwa2cas13a	韦德纤毛菌 F0279	WP_021746774.1	GATATAGATAACCCCAAAAACG AAGGGATCTAAAAC (SEQ ID NO: 186)
Cas13a13	RcsCas13a	荚膜红细菌 SB 1003	WP_013067728.1	GCCTCACATCACCGCCAAGACG ACGGCGGACTGAAC (SEQ ID NO: 187)
Cas13a14	RcrCas13a	荚膜红细菌 R121	WP_023911507.1	GCCTCACATCACCGCCAAGACG ACGGCGGACTGAAC (SEQ ID NO: 188)
Cas13a15	RcdCas13a	荚膜红细菌 DE442	WP_023911507.1	GCCTCACATCACCGCCAAGACG ACGGCGGACTGAAC (SEQ ID NO: 189)

[0134] 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  MGNLFGHKRW YEVRDKKDFK IKRKVKVKRN YDGNKYILNI NENNNKEKID NNFIRKYIN
61  YKKNDNILKE FTRKFHAGNI LFKLKGKEGI IRIENDDFL ETEEVLVYIE AYKSEKLKA
121 LGITKKKIID EAIRQGITKD DKKIEIKRQE NEEIEIDIR DEYTNKTLND CSIILRIEN
181 DELETKSIY EIFKNINMSL YKIEEKIEN ETEKVFENRY YEEHLREKLL KDDKIDVILT
241 NFMEIREKIK SNLEILGFVK FYLNVGGDKK KSKNKKMLVE KILNINVDLT VEDIADFIK
301 ELEFWNITKR IEKVKKVNNE FLEKRRNRTY IKSIVLLDKH EKFKIERENK KDKIVKFFVE
361 NIKNNSIKEK IEKILAEFKI DELIKKLEKE LKKGNCDEI FGIFKKHYKV NFDSKKFSK
421 SDEEKELYKI IYRYLKGRIE KILVNEQKVR LKKMEKIEIE KILNESILSE KILKRVQYT
481 LEHIMYLGKL RHNDIDMTTV NTDDFSRLHA KEELDLELIT FFASTNMELN KIFSRENINN
541 DENIDFFGGD REKNYVLDKK ILNSKIKIIR DLDFIDNKN ITNNFIRKFT KIGTNERNRI
601 LHAISKERDL QGTQDDYNKV INIIQNLKIS DEEVSKALNL DVVFKDKNI ITKINDIKIS

```

661 EENNNDIKYL PSFSKVLPEI LNLYRNNPKN EPFDTIETEK IVLNALIYVN KELYKKLILE
 721 DDLEENESKN IFLQELKCTL GNIDEIDENI IENYYKNAQI SASKGNNKAI KKYQKKVIEC
 781 YIGYLRKNYE ELDFDSDFKM NIQEIKKQIK DINDNKTYER ITVKTSDKTI VINDDFEYII
 841 SIFALLNSNA VINKIRNRFF ATSVWLNTSE YQNIIDILDE IMQLNTRLNE CITENWNLNL
 901 EEFIQKMKEI EKDFDDFKIQ TKKEIFNNYY EDIKNNILTE FKDDINGCDV LEKKLEKIVI
 961 FDDETKFEID KKSNIQDEQ RKLSNINKKD LKKKVDQYIK DKDQEIKSKI LCRIIFNSDF
 1021 LKKYKKEIDN LIEDMESENE NKFQEIYYPK ERKNELYIYK KNLFNLNIGNP NFDKIYGLIS
 1081 NDIKMADAKF LFNIDGKNIR KNKISEIDAI LKNLNDKLNQ YSKEYKEYI KKLKENDDF
 1141 AKNIQNKNYK SFEKDYNRV SFEKDYNRV EYKKIRDLVE FNYLNKIESY LIDINWKLAI QMARFERDMH
 1201 YIVNGLRELQ IIKLSGYNTG ISRAYPKRNG SDGFYTTTAY YKFFDEESYK KFEKICYGFG
 1261 IDLSENSEIN KPENESIRNY ISHFYIVRNP FADYSIAEQI DRVSNLLSYS TRYNNSTYAS
 1321 VFEVFKKDVN LDYDELKKKF KLIGNNDILE RLMKPKKVSF LELESYNSDY IKNLIIELLT
 1381 KIENTNDTL (SEQ ID NO: 190)。

[0135] 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 (Prevotella sp.) 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 (Porphyromonas gingivalis) 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 (Porphyromonas gulae)	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 (Myroides odoratimimus) 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 (Bergeyella zoohelcum) 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 (<i>Chryseobacterium</i> 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 (<i>Prevotella buccae</i>) 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 (<i>Prevotella pallens</i>) 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.)	WP_039428968.1	1176

COT-052 OH4946		
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

[0136] 本公开文本的示例性野生型动物溃疡伯格菌ATCC 43767 Cas13b (BzCas13b) 蛋白可以包含以下氨基酸序列或由其组成:

```

1  menktslgnn iyympfkpqd ksyfagyfna amentdsvfr elgkrllgke ytsenffdai
61  fkenislvey eryvklldsy fpmarllldkk evpikerken fkknfkgiik avrdlnrfyt
121 hkehgeveit deifgvldem lkstvlvtvkk kkvktldktke ilkksiekql dilcqqkley
181 lrdtarkiee krrnqrerge kelvapfkys dkrddliaai yndafdvaid kkkdskless
241 kakyntksdp qqeegdlkip isknvgvfl1 slfltkqeih afkskiagfk atvideatvs
301 eatvshgkns icfmatheif shlaykklkr kvrtaeinyg eaenaeqlsv yaketlmmqm
361 ldelskvpdv vyqnlsedvq ktfiedwney lkenngdvgt meeeqvihpv irkryedkfn
421 yfairfldef aqfptlrfqv hlgnylhdsr pkenlisdr ikekitvfgr lselehkkal
481 fikntetned rehyweifpn pnydfpkeni svndkdfpia gsildrekqp vagkigikvk
541 llngqyvsev dkavkahqlk qrkaskpsiq niieevpin esnpkeaivf ggqptaylsm
601 ndihsilyef fdkwekkkek lekkgekelr keigkelekk ivgkiquaqiq qiidkdnak

```

```

661 ilkpyqdgns taidkeklik dlkqeqnilq klkdeqtvre keyndfiayq dknreinkvr
721 drnhkqylkd nlkrkypeap arkevlyyre kgkvavwlan dikrfmptdf knewkgeqhs
781 llqkslayye qckeelknll pekvfqhlpf klgyyfqgky lyqfytcyld krleyisglv
841 qqaenfksen kvfkkvenec fkflkkqnyt hkeldarvqs ilgypifler gfmdekptii
901 kgkftkgnea lfadwfryyk eyqnfqtfyd tenyplvele kkgadrkrkt kiyqqkknvd
961 ftllmakhif ksvfkqdsid qfsledlyqs reerlgnqr arqtgerntn yiwnktvdlk
1021 lcdgkitven vklknvgdfi kyeydqrva flkyeeniew qafilkeske eenypyvver
1081 eieqyekvrr eellkevhli eeyilekvkd keilkkgdnq nfkyiylngl lkqlknedve
1141 sykvfnlnte pedvninglk qeatdlegka fvltyirnkf ahnqlpkkef wdycqekygk
1201 ektyaey faevfkkeke alik (SEQ ID NO: 191)。

```

[0137] 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。关于CasRX/Cas13d蛋白的其他例子和序列还参见W0公开号W0 2019/040664和US 2019/0062724,将其通过引用以其整体并入本文,在没有限制的情况下,具体参考

[0138] 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 ITPSITVKA TDESVRFRVK RKIAQKDEFI      60
ADNPMHEGRH RIEPSAGSDM LGLKTKLEKY YFGKEFDDNL HIQIIYNILD IEKILAVYST      120
NITA                                              124

```

(SEQ ID NO:54)。

[0139] 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
KKIAFDFFRN VRQKTRTREE SLEKERFKAV IGLYLTVVYL LIKNLVNVNS RYVMAFHCLC      300
RDAKLYGINI GKNYIELTED LCRENENSRS AYLARNKRLR DCVKQIDNA KNMKSKEK      358

```

(SEQ ID NO:57)。

[0140] 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 YQLENTPDLD NEYRDTLDHF FDERFNEINE HFVTQNATNL CIMKEVFPDE      60
DFKSIADLYY DFIVVKSYKN IGFSIKKLRE KMLELPEAKR VTSTEMDSVR SKLYKLIDFC      120
IFKHHEKPE TVEMIVSMLR AYTSEDME      149
```

(SEQ ID NO:61)。

[0141] 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 YHNVVNGKF LEDQLGARSE LEKKVFGRTF      120
DDNVHIQLIH NILDIEKIMA QYVSDIVYLL HNTIKRDMND DIMGYISIRN SFDDFCHPER      180
IPDRKAKDNL QKQHDIFFDE ILKCGRLAYF GNAFFEDGSD NKEIAKLKRY KEIYHIIALM      240
GSLRQSYFHG ENSDKNFQGP TWAYTLESNL TGKYKEFKDT LDKTFDERYE MISKDFGSTN      300
MVNLQILEEL LKMLYGNVSP      320
```

(SEQ ID NO:67)。

[0142] 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
KLTQKFITDK YLNDGAQGCS KKVGNLYSHN ITCCSDELK EYRNQVDHFA VVRMIGKYAA      120
DIGKFSTWFE LYHYVMQRII FDKRNPLSET ERTYKQLIAK HHTYCKDLVK ALNTPFGYNL      180
ARYKNLSIGE LFDRNNYNAK TKET      204
```

(SEQ ID NO:69)。

[0143] 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13552000311:

```
LIDFLIYDLY YNRKPARIEE IVDKLRESVN DEEKESIYSA ETKYVYEALG KVLVRSLLKY      60
LNGATIRDLK NRYDAKTANR IWDISEHSKS GHVNCFCCLI YMTLMMDGK EINDLLTTLV      120
NKFDNIASFI DVMDELGLEH SFTDNYKMFA DSKAICLDLQ FINSFARMSK IDDEKSKRQL      180
FRDALVVLDI GDKNEDWIEK YLTSDIFKRD ENGKNIDGK RDFRNFIANN VIKSARFKYL      240
VKYSSADGMI KLKKNEKLIS FVLEQLPETQ IDRYYESCGL DCAVADRKVR IEKLTGLIRD      300
MRFDNFRGVN YSNDACKKDK QAKAKYQAI SLYLMVLYQI VKNMIYVNSR YVIAFHCLER      360
DLLFFNIELD NSYQYSNCNE LTEKFIKDKY MKEGALGFNM KAGRYLTKNI GNCSNELRKI      420
YRNQVDHFAV VRKIGNYAAD IASVGSWFE      449
```

(SEQ ID NO:71)。

[0144] 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	AWAIHVYRNK	IQHLDVVRHA	DMYIGDIREF	HSWFELYHYI	IQRRIIDQYA	60
YESTPGSSRD	GSAIIDEERL	NPATRRYFRL	ITTYKT			96

(SEQ ID NO:72)。

[0145] 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	LPKIHRLLPR	GQKISNFSKL	MYLLTMFLDG	120
KEINDLLTTL	INKFENIQGF	LDIMPEINVN	AKFEPEYVFF	NKSHEIAGEL	KLIKGFQMG	180
EPAATLKLEM	TADAIKILGT	EKEDAEIKL	AESLFKDENG	KLLGNKQHGM	RNFIGNNVIK	240
SKRFHYLIRY	GDPAHLHKIA	TNKNVVRFL	GRIADMQKKQ	GQKGKNQIDR	YYEVCVGND	300
IKKTIEEKID	ALTDIIVNMN	YDQFEKKKAV	IENQNRGKTF	EEKNKYKRDN	AEREKFKKII	360
SLYLTVIYHI	LKNIVNVNSR	YILGFHCLE	RDKQLYIEKYN	KDKLDGFVAL	TKFCLGDEER	420
YEDLKAKAQA	SIQALETANP	KLYAKYMNYS	DEEKKEEFKK	QLNRERVKNA	RNAYLKNIKN	480
YIMIRLQLRD	QTDSSGYLCG	EFRDKVAHLE	VARHAHEYI			519

(SEQ ID NO:73)。

[0146] 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)。

[0147] 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	GGNDAKLEKV	VENSEIVSLA	EKESFSAEIF	60
KKNIGCKIEN	KKFKHPKRYD	VIADNPLYKG	SVRQDMLGLK	ETLEKRYFNS	ADGTDNVCIQ	120
VIHNILDIEK	ILAEYITNAV	YSFDNIAGFG	EDIIGMGGFK	PIYTYKQFKE	PKYNNKFDD	180
ILNNSRLGYY	GKAFFEKNDL	KHNPNKKKRD	KNPYILKYDN	ECYIIALLS	GLRHWNIHSH	240
AKDDLVSYRW	LYNLDSILNR	EYISTLNYLY	DDIADELTES	FSKNSSANVN	YIAETLNIDP	300
SEFAQQYFRF	SIMKEQKNMG	FNVSKLREIM	LDRKELSDIR	DNHRVFDISIR	SKLYTMMDFV	360
IYRYYIEEAA	KTEAENRNLP	ENEKKISEKD	FFVINLRGSF	DENQKEKLYI	EEAKRLWEKL	420
KDMLKIKEF	RGEKVKEYKK					440

(SEQ ID NO:85)。

[0148] 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  ADELTRTFSEK  NSAANVNYIA  ETLNIDPNKF  AEQYFRFSIM      60
KEQKNLGFNL  TKLRESMLDR  RELSDIRDNH  NVFDSIRPKL  YTMMDFVIYK  HYIDEAKKTE     120
AENKSLPDDR  KNLSEKD                               137
(SEQ ID NO:86)。
```

[0149] 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  RVMMDIAVKI  LGTDLSDDEL  KEMADSFFKD  SDGNLLKKGK  HGMRNFITNN      60
VIKNKRFHYL  IRYGDPAPHL  EIAKNEA                               87
(SEQ ID NO:87)。
```

[0150] 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  KYTWLYNLDK  YLDAEYITTL  NYMYNDIGDE  LTDSFSKNSA  ANINYIAETL      60
GIDPKTFAEQ  YFRFSIMKEQ  KNLGFNLTKL  REVMLDRKDM  SEIRENHND  DSIRAKVYTM     120
MDFVIYRYEI  EEAQVNAAN  KSLPDNEKSL  SEKDFVISL  RGSFNEDQKD  RLYYDEAQL      180
WSKVGKMLLK  IKKFRGKDTR  KYKNMGTPRI  RRLIPEGRDI  STFSKLMYAL  TMFLDGKEIN     240
DLLTTLINKF  DNIQSFLKVM  PLIGVNAKFA  EEYSFFNNSE  KIADELRLIK  SFARMGEPVA     300
DARRAMYIDA  IRILGTDLSD  DELKALADSF  SLDENGKLG  KGKHGMRNFI  INNVITNKRF     360
HYLIRYGNPV  HLHEIAKNEA  VVKFVLGRIA  DIQKKQGQNG  KNQIDRYYET  CIGK            414
(SEQ ID NO:88)。
```

[0151] 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  VKGLGLKSTL  VSDSKVYLT  FADGSNAKLE  KCVENNKIIC  ISNDKEAF      60
SIANKNVGYK  IKNDEKFRHP  KGYDIISNP  LLHNSVQQD  MLGLKNVLEK  RYFGKSSGGD     120
NNLCIQIIHN  IIDIEKILSE  YIPNVVYAFN  NIAGFKDEHN  NIIDIIGTQT  YNSSYTYADF     180
SKDKSDKKYI  EFQKLLKNKR  LGYWGKAFFT  GQGNNAKVRQ  ENQCFHIIAL  LISLRNWATH     240
SNELDKHTKR  TWLYKLDDTN  ILNAEYVKTL  NYLYDTIADE  LTKSFSKNGA  VNVNVLAKKY     300
NIKDDLPGFS  EQYFRFSIMK  EQKNLGFNIS  KLRENMLDFK  DMSVI              345
(SEQ ID NO:89)。
```

[0152] 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:

```
KKISLTKEFC  LGESDEKKLK  ALAKSLEEL  KTTNSKLYEN  YIKYSDERKA  EEAKRQINRE      60
RAKTAMNAHL  RNTKWNIMY  GQLKDLADSK  SRICSEFRNK  AAHLEVARYA  HMYINDISEV     120
KSYFRLYHYI  MQRRIIDVIE  NNPKAKYEGK  VKVYFEDVKK  NKKYNKNLLK  LMCVPFGYCI     180
PRFKNLSIEQ  MFDNETDNS  DKKKEK                               206
(SEQ ID NO:90)。
```

[0153] 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)。

[0154] 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 PSESDFS VIR LRGSFNENQK EELYIEEAER      60
LWKKFGEIML KIKEFRGEKV KEYKKEVPRI ERILPHGKDI SAFSKLMYML SMFLD          115
```

(SEQ ID NO:92)。

[0155] 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 AASAKLTMR DALITLGIDD KITDDRISSEI LKLKEKGKGI     120
HGLRNFITNN VIESSRFVYL IKYANAQKIR EVAKNEKVVM FVLGGIPDTQ IERYKSCVE      180
FPDMNSSLEA KRSELARMIK NISFDDFKNV KQQAAGRENV AKERAKAVIG LYLT           234
```

(SEQ ID NO:93)。

[0156] 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 LEAIXIARM ENKIENAPQE PLLKDALLSL	600
GVSDDTKVL NTKYFYFDSK EKTDKQSQKV STFLMNNVIN NNRFKYVIKY INPADINGLA	660
KNRYLVKFVL SKIPEEQIDS YYKLFSNEEE PGCEEKIKLL TKKISKLNQ TLFENNKIPN	720
VEKEKKKAI TLYFTIVYIL VKNLVNINGL YTLALYFVER DGYFYKDICG KKDKKKSYND	780
VDYLLLPEIF SGSKYREETK NLKLPKEKDR DIMKKYLPND KDREKYNKFF TAYRNNIVHL	840
NIIAKLSELT KNIDKDINSY FDIYHYCTQR VMFNYCKEKN DVVLAKMKDL AHKSDCNEF	900
SSKHTYPFSS AVLRFMNLPF AYNVPRFKNL SYKKFFDKQ	939

(SEQ ID NO:94)。

[0157] 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 SDHVNFWLY	240
QLEDQLSDEF KETLSILWEE VTERIDSEFL KTNTVNLHIL CHVFPKESKE TIVRAYEFL	300
IKKSFKNMGF SIKKLREIML EQSDLKSFKE DKYNSVRACL YKLFDIITY YYDHHAFEKE	360
ALVSSLRSSL TEENKEEIIYI KTARTLASAL GADFKKAAAD VNAKNIRDYQ KKANDYRISF	420
EDIKIGNGTI GYFSELIYML TLLLDGKEIN DLLTTLINKF DNIISFIDIL KKLNLFEKFK	480
PEYADFFNMT NCRYTLEELR VINSIARMQK PSADARKIMY RDALRILGMD NRPDEEIDRE	540
LERTMPVGAD GKFIKKGQGF RNFIASNVIE SSRFHLYVRY NNPHKTRTLV KNPVVKFVL	600
EGIPETQIKR YFDVCKGQEI PPTSDKSAQI DVLARISSV DYKIFEDVPQ SAKINKDDPS	660
RNFSDALKKQ RYQAIVSLYL TVMYLITKNL VYVNSRYVIA FHCLERDAFL HGVTLPKMNK	720
KIVYSQLTTH LLTDKNYTTY GHLKNQKGRH KQYVLVKNL QNSDITAVSS FRNIVAHISV	780
VRNSNEYISG IGEHLSYFEL YHYLVQSMIA KNNWYDTSHQ PKTAEYLNNL KKHHTYCKDF	840
VKAYCIPFGY VVPYKNLTI NELFDRNNPN PEPKEEV	877

(SEQ ID NO:95)。

[0158] 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

[0159] (SEQ ID NO:96)。

[0160] 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 ILDEKILAE YTTNAVFAID NVSGCSDDL SNFSTRNQWD EFQNPQHRE	240
HFGNKDNVIC SVKKQQDLFF NFFKNNRIGY FGKAFFHAES ERKIVKKTEK EVYHILTLLIG	300
SLRQWITHST EGGISRLWLY QLEDALSREY QETMNNCYNs TIYGLQKDFE KTNAPNLNLF	360

```

AEILGKNASE LAEPYFRFII TKEYKNLGFS IKTLREMLLD QPDLQEIREN HNVYDSIRSK      420
LYKMIDFVLV YAYSNERKSK ADALASNLRS AITEDAKKRI YQNEADQLWT SYQELFKRIR      480
GFKGAQVKEY SSKNMPPIPIQ 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
LCKILTNSY DQFQDVQSG AEITAEERK KERYKAIISL YLTVLYQLIK NLVNINARIYI      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)。

[0161] 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 VKDYDFVVR KNYKHLGFSV      420
KELRELLFAN HDSNKYIKEF DKISNKKFDS VRSRLNRLAD YIIYDYNNKN NAKVSDLVKY      480
LRAAADDEQK KKIYLNESIN LVKSGILERI KKILPKLNGK IIGNMQPDST ITASMLHNTG      540
KDWHPISENA HYFTKWIYTL TLFMDGKEIN DLVTTLINKF DNIAFIEVL KSQSVCTHFS      600
EERKMFIOSA EICSELSAMN SFARMEAPGA SSKRAMFVEA ARILGDNRSK EELEEFDTL      660
FDKSASKKEK GFRNFIRNNV VDSNRFKYLT RYTDTSVKA FSNKALVKF AIKDIPQEQI      720
LRYYNCSFGA SERYNDGMS DKLVEAIGKI NLMQFNGVIQ QADRNLPEE KKKANAQKEK      780
YKSIIRLYLT VCYLFFKNLV YVNSRYSAF YNLEKDRSLF EINGELKPTG KFDEGHYTG      840
VKLFIDNGWI NPRASAYLTV NLANSDETAI RTFRNTAEHL EALRNADKYL NDLKQFDSYF      900
EIIHYITQRN IKEKCEMLKE QTVKYNNDLL KYHGYSKDFV KALCVPFPGYN LPRFKNLSID      960
ALFDKNDKRE KLKKGFE

```

(SEQ ID NO:98)。

[0162] 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 REEKERLKNQ KWAKNDSVII VPETKEEIKT GEIQDNNRKR SRQKSQAKAM	60
GLKAVLSFDN KIAIASFVSS KNAKSSHIER ITDKEGTTIS VNSKMFESSV NKRDNIEKR	120
ITIEEPQQDG TIKKEEKGVK STTCNPYFKV GSKDYIGIKE IAEHFFFGRA 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 IIPGKEDPV KFSHGKLDKK EIESFCLTTK NTEDITKVIY FLCKFLDGKE	540
INELCCAMMN KLDGISDLIE TAKQCGEDVE FVDQFKCLSK CATMSNQIRI VKNISRMKKE	600
MTIDNDTIFL DALELLGRKI EKYQKDKNGD YVKDEKGGKV YTKDYNFQD MFFEGKNHRV	660
RNFVSNNVIK SKWFSYVVRY NKPAECQALM RNSKLVKFAL DELPDSQIEK YYISVFGEKS	720
SSSNEEMRRE LLKKLCDFSV RGFLDEIVLL SEDEMKQKDK FSEKEKKKSL IRLYLTIYVL	780
ITKSMVKINT RFSIACATYE RDYILLQSE KAERAWEKGA TAFALTRKFL NHDKPTFEQY	840
YTREREISAM PQEKRKELRK ENDQLLKKTH YSKHAYCYIV DNVNNLTGAV ANDNGRGLPC	900
LSEKNDNANL FLEMRNKIVH LNVVHDMVKY INEIKNITSY YAFFCYVLQR MIIGNNSNEQ	960
NKFKAKYSKT LQEFGTYSKD LMWVLNLPFA YNLPRYKNLS NEQLFYDEEE RMEKIVGRKN	1020
DSR	1023

(SEQ ID NO:99)。

[0163] 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):

MTETPKKRED IAKTPAAKSR SKAAGLKSTF AVNGSVLLTS FGRGNDAVPE KLITEKAVSE	60
INTVKPRFSV EKPATSYSSS FGIKSHISAT ADNPLAGRAP VGEDAIHAKE VLEQRVFGKT	120
FSDDNIIHQL 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 QNFYDFRVL	360
QDKNLGFSIK KLREKLLDSP ALSVVRSKKY DTMRSKIYSL IDFMIRKFS ENHVAVDFFV	420
EELRSLLED EKESAYSRA ETLLNDGFAQ EILVKLLPQT DPAVIGKIKG KKLNDISIAG	480
IKLKKDASFF TKIINVLCMF QDGKEINELV SSLVNKFANI QSFVDVMRSQ GIDSGFTADY	540
AMFAESGRIS RELHILKGIA RMQHSIAGLG DVKIYGSDDK FHGVSRRVYT DAAYILGFGE	600
RSEDNDGYVD DYVSSKLLGG ADKNLRNFIT NNVIKNRRFL YTVRYMNPGR AKKLVQNDAL	660
VVLALSGIPE TQIDRYKSC IEKRSFNPDL NEKIAALSEM ITTLKIDDFE DVKQNPKEKA	720
NYEAKKNQRI SKERYKACIG LYLTVLYLIC KNLVKINARY SIAIGCLERD TQLHGVDFFG	780
AAYMTRDVFI AKGWINPKKP TVKSIKEQYA FLTPYIFTTY RNMIHAAV TNAYKYIPQM	840
DRFKSWFHLY HTVIQHSLIQ QYEDRDYGR KGAPVVSERV LQLLEQCREH SNYSRDLHI	900
LNLPFGYNLP RYLNLSSEKY FDANAI	926

[0164] (SEQ ID NO:100)。

[0165] 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)：

MAKKITAKQK REEKERLKNQ KWAKQDTPVV PKSKTEEPV AASDDKLLKT TQVKKVQTKS	60
KAKAMGLKTV LSFDDKIAIA SFVNDKKTKL PHIERITDKS GTTIHENARM FDSSVDEQNV	120
NIEKRMTIEE KQNDGTFKKD EKDVKATICN PYFKTCGKDY IGIKDVAEKY FFGKTFPNE	180
LRVQIAYNVF DIQKILGTYV NNIIYSFYNL RRDGKSDVDI IGSLYAFADF DNQLKDKPAF	240
REAKDLLKNT EAYFSYFGDV FKKSCKGKKD ENNEDYEKNL RHNFNVLRLV SFLRQICTHA	300
YVKCTGGAKN NGDSTKVEAE SLDALFNITE YFAKTAPELS KTINEIYKEG IDRINNDFT	360
NGKNNLYILS KVYPDMQRNE LVKKYYQFVV CKEGNNVGIN TRKLKESIIS QHPWITTPQD	420
NNKANDYESC RHKLYTIMCF ILVAELDAHE SIRDNMVAEL RANMDGDDGR DAIYEKYAKD	480
IYHIVKDKLL AMQKVFDEEL VPKVVEGKND PQQFTHGKLG KKEIESFCLS DKNTSDIAKV	540
VYFLCNFLDG KEINELCCAM MNKFDGIGDL IDTAKQCGEE VKFIEEFACL SNCRKITNDI	600
RVAKSISKMK NKVNIDNDII YLDAIELLGR KIEKYQKDEN GKILLGTDGK RLYTQEYKYF	660
NDMFFNAGNH KVRNFIANNV MQSKWFFYVV RYNKPAECQI IMRNKTLVKF TLDDLPMQI	720
QRYYSVFGD NNMPAVDEMR KRLLDKINQF 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
VLQRRVLGKD PNCQNKFKAK YAKELNDYGT YNKNLMWMLN LPFAYNLPRY KNLSSEFLFY	1020
DMEYNKKDDE	1030

(SEQ ID NO:101)。

[0166] 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 NGKTIREDVR MFDSNVDECS	120
INLEKRMTVE ERQKDGTIKK DEKDVKSTIC NPYSNECGKD YIGIKSVAEE LFFGRTPND	180
NLRVQIAYNI FDIQKILGTI INNIIYSFY LSRDESQSDN DVIGTLYMLK DFDGQKETDT	240
FRQARALLER TEAYYSYFDN VFKKIDKNKK KSDDCKRERN EILRYNFNVL RVL SFLRQIC	300
AHAQVKISNE HDREKGGGLV DSLDALFNIS RFFDAVAPEL NEVINSVYSK GIDDINDNFV	360
KNGKNNFYIL SKIYPEVARE DLLREYYYFV VSKEGNNIGI STKKLKEAII VQDMSYIKSE	420
DYD TYRNKLY TVLCFILVKE LNERTTIREQ MVADLRANMN GDIGREDIYS KYAKIIYAQV	480
KPRFDTMKSA FEEEA KDVIV PDKKKPVKFS HGKLDKNEIE RFCITSANTD SVAKIIYFLC	540
KFLDGKEINE LCCAMMNKLD GINDLIETAE QCGAKVEFVD KFSVLSNCET ISDQIRIVKS	600
ISKMKKEIAI DNDTIFLDAL ELLGRKIDKY KKDATGKYLK DENGKYLYSK EYDDFQYMEF	660
KDSHRVRNFI SNSVIKSKWF SYIVRYNQPS ECRAIMKNKT LVKFALDELP DLQIQRYFVA	720
LYGDEDLPSY GEMRKILLKK LHDFSIGFL DEIVLLSDLD MESQDKYCEK EQKKSFLRLY	780
LTIAYLITKS MVKINTRFSI ACATYERDIA LLCASNKQER AWSSGATALA LTRRFLNQDK	840
LIFEKHYARE GEISKLPKEE RKAMRKVNDQ LLKRTHFSKH SYCYIVDNVN RLTGGECDT	900
KRVLPVLNEK NDNAGILLDF RKTIAHLNVV HKMVDYVDEI KGITSYYAFF CYVLQRMLVG	960
NNLNEKNAIK EKYSATVKSF GTYSKDFMWL INLPFAYNLP RYKNLSNEQL FYDEEERNET	1020
EEQIDRL	1027

(SEQ ID NO:102)。

[0167] 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 VTADDPYRR KDGGFVPGMD ALRRKNVLEQ RFFGKSFADN IHIQMIYSIL	180
DIHKILAAAS GHIVHLLNIV NGSKDRDFIG MLAHVLYNE LNEEAKRSIA DFCKSPRLIY	240
YSAAFYETLD NGKSERRSNE DIFNILALMT CLRNFSSSHHS IAIKVKDYSA AGLYNLRRLG	300
PDMKKMLDTF YTEAFIQLNQ SFQDHNTTNL TCLFDILNIS DSARQKQLAE ERYRYVVFKE	360
QKNLGFSVRK LREEMLLLPD AAVIADKRYD TCRSKLYNLM DFLILRVYRT GRADRCDKLP	420
EALRAALTDE EKAVVYHKEA LSLWNEMRTL ILDGLLPQMT PENLSRLSGQ KRKGELSDD	480
AMLKECLYEP GPVPEDAPE EANAIFYCRM IYLATLFMDG KEINTLLTTL ISKFENIAAF	540
LQTMEQLNIE AELGPEYAMF TRSRAVAEQL RVINSFALMK KPQVNAKQQL YRAAVTLLGT	600
EDPDGVTDEM LCIDPVTGKM LPPNQRHHGD TGLRNFIANN VVESRRFQYL IRYSDPAQLH	660
QLASNKKLVR FVLSSIPDTQ INRYETCGQ TRLAGRAAKV EFLTDMIAAI RFDQFRDVNQ	720
KERGANTQKE RYKAMLGLYQ TVLYLAVKNL VNINARYVMA FHCVERDMFL YDGELTDPKG	780
ESVSAFLAVN GKKGVQPQYL LLTQLFIRRD YLKRSACEQI QHNMENISDR LLREYRNAVA	840
HLNVIAHLAD YSADMREITS YGGLYHYLMQ RHLFKRHAWQ IRQPERPTEE EQKLIEQEQK	900
QLAWEKALFD KTLQYHSYNK DLVKALNAPF GYNLARYKNL SIEPLFSKEA APAAEIKATH	960
A	961

(SEQ ID NO:103)。

[0168] 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 GAAAELESIL TGGEIRKLSD	60
KAILEITDDT QKNAYNVKS SRIPNLART DKLSKSGMD DLGFKRELEL EVFGQCFDD	120
IHIQIAHAVF DIQKSLAAVI PNVLTYLNNL DRSYSTDNST DKKDIIGNTL NYQHSYESFN	180
VEKRGFEFTEY YNAAKDRFSY FPDILCVLEK VNGKDRYQPK SEKDAFNVLS SVNMLRNSLF	240
HFAPKSNDBG ARIAVFKNQF DSDFSHITST VNKIYSAKIA GVNENFLNNE GNNLYIILKA	300
TNWDIKKIVP QLYRFSVLKS DKNMGFNMRK LREFAVESKN IDLSRLNDKF LTNNRKKLYK	360
VIDFIIYYHL NKVLKDSFVD DFVAALRASQ SEEEKEKLYA QYSERLFADE GLKSAIKKAV	420
DMISDTKSNL FKMKTPLDKA LIENIKVNSD ASDFCCLIYV FTRFLDGKEI NILLNSLIKK	480
FQDIHSFNNT VKKLSENLI INADYVDDYS LFEQSGTVAR ELMLIKSISK MDFGLDNINL	540
SFMYDDALRT LGVSDENLPE VKREYFGGTK NLSAYIRNNV LENRRFKYVI KYIHPSDVQK	600
IACNKAIAGF VLNRMPTQI KRYYDSLIN KATDIQAQAK ALLDCITGIS FDAIKDDKHL	660
HKSKEKSPQR SADRERKKAM LTLYYTIVYI FVKQMLHINS LYTIGFFYLE RDQRFIYSRA	720

KKENKNPSKN SYLNDFRSVT AYFIPSEIMK RIEKNENKGF LEDFEALWNS CGKTSRLRKE	780
DVLLYARYIS PDHALKNYKM ILNSYRNKIA HINVIMSAGK YTGGIKRMDS YFSVFQHLVQ	840
CDILSNPNNK GKCFESESLEK PLLLDMMKFDG TDEKLYSKRL TRALNIPFGY NVPRYKNLTF	900
EKIYKSSIN E	911

(SEQ ID NO:104)。

[0169] 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 NFSAHYLDNR	240
RAKSSDAEDF DGTSVIMYNL DNEELYKKLY NKKVHMALTG MKKVLDANFN KKVEHLNNSF	300
IKNSAKDFVI LCEVLGIKSR DEKTKFVKDY YDFVVRKNYK HLGFSVKELR ELLFANHDSN	360
KYIKEFDKIS NKKFDSVRSR LNRLADYIIY DYNNKNNKV SLDVKYLRAA 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 TSSVKAFFSN KALVKFAIKD IPQEQILRYN NSCFGASERY	660
YNDGMSDKLV EAIGKINLMQ FNGVIQQADR NMLPEEKKA NAQKEKYKSI IRLYLTVCYL	720
FFKNLVVYNS RYYSAFYNLE KDRSLFEING ELKPTGKFDE GHYTGVLVKLF IDNGWINPRA	780
SAYLTVNLAN SDETAIRTFR NTAEHLEALR NADKYLNDLK QFDSYFEIYH YITQRNIKEK	840
CEMLKEQTVK YNNDLLKYHG YSKDFVKALC VPFGYNLPRF KNLSIDALFD KNDKREKLKK	900
GFED	904

(SEQ ID NO:105)。

[0170] 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 EKDIDENGVV NDCRNPAAYD	60
AVYGTDSIRV KKTNNNIRAK VNNPLAKSNI RSEESALFRT RVNEYKREKQ DKYETLFFGK	120
TFDDNIHIQL ISKILDIEKT FSVVIGNIVY AINNLSLEQS IDRPIDIFGD KNTQGSLRE	180

DNDYLKTMPL RCEYLFHNIL NSDSDNNSKM NYNKNVNGKE EKDNRRNENI EKLKKALEVI	240
KIIRVDSFHG VDGIGDQKF PRSKYNLAVN YNEEIQKTIS EPNRNVVEEV QQDFYRNSCV	300
NIDFLKEIMY GSNYTDGRSD SLECSYFNFA ILKQKNMGF SITSIRECLL DLYELNFESM	360
QNLPRANSF CDFLIYDYC KNESERANLV DCLRSAASEE EKKNIFYQTA ERVKEKFRNA	420
FNRISRFDAS YIKNSREKNL SGGSSLPKYS FIEGFTKRSK KINDNDEKNA DLFCNMLYYL	480
AQFLDGKEIN IFLTSIHNI QNIDSFLKVM KEGMECKFQ KDFKMFHAG HVAKKIEIVI	540
SLAKMKKTL D FYNAQALKDA VTILGVSKKH QYLDMNSYLD FYMFDNRSGA TGKNAGKDH	600
LRNFLVSNVI RSRKFNYLSR YSNLAEVKKL AQNPSLVQFV LSRIEPLSLIC RYESSQGIS	660
SEGITIDEQI KKLGTGIIVDM NIDSFENINN GEIGMRYSKA TPQSIERRNQ MRVCVGLYLN	720
VLYQIEKNLM NVNARYVLAF AFAERDALML NFTLEECKKN KKRSSGGFSF IEMTQFFIDK	780
KLFKVATEAI KKNVLKYNGN PESLNHIPGE YICKNMEGYH ENTVRNFRNM VAHLTAVARV	840
PLYISEVTQI DSYIALYHYC MQMNLQIE QSGKILDNIK LKNALENARV HRTYSKDAVK	900
YLCLPFAYNI SRYKALTIKD LFDWTEYSCK KDE	933

(SEQ ID NO:106)。

[0171] 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:

MKRQKTFAGR IGKSTVAYG QGKYAITTFG KGSKAIEIAVR SADPPEETLP TESDATLSIH	60
AKFAKAGRDG REFKCGDVDE TRIHTSRSEY ESLISNPAES PREDYLGLKG TLERKFFGDE	120
YPKDNLRIQI IYSILDIQKI LGLYVEDILH FVDGLQDEPE DLVGLGLGDE KMQKLLSKAL	180
PYMGFFGSTD VFVKVTKKREE RAAADEHNAK VFRALGAIRQ KLAHFKWKES LAIFGANANM	240
PIRFFQGATG GRQLWNDVIA PLWKKRIERV RKSFLSNSAK NLWVLYQVFK DDTDEKKKAR	300
ARQYYHFSVL KEGKNLGFNL TKTREYFLDK FFPIFHSSAP DVKRKVDTFR SKFYAILDFI	360
IYEASVSVAN SGQMGKVAPW KGAIIDNALVK LREAPDEEAK EKIYNVLAAS IRNDSLFLRL	420
KSACDKFGAE QNRPVFPNEL RNNRDIRNVR SEWLEATQDV DAAAFVQLIA FLCNFLEGKE	480
INELVTALIK KFEGIQALID LLRNLEGVDS IRFENEALF NDDKGNMAGR IARQLRLLAS	540
VGKMKPDMTD AKRVLYKSAL EILGAPPDEV SDEWLAENIL LDKSNNDYQK AKKTVNPFRN	600
YIAKNVITSR SFYYLVRYAK PTAVRKLSN PKIVRYVLKR LPEKQVASYY SAIWTQSESN	660
SNEMVKLIEM IDRLTTEIAG FSFAVLKDKK DSIVSASRES RAVNLEVERL KKLTTLYMSI	720
AYIAVKSLVK VNARYFIAYS ALERDLYFFN EKYGEEFRLH FIPYELNGKT CQFEYLAILK	780
YYLARDEETL KRKCEICEEI KVGCEKHKK ANPPYEYDQE WIDKKKALNS ERKACERRLH	840

FSTHWAQYAT KRDENMAKHP QKWYDILASH YDELLALQAT GWLATQARND AEHLNPNVNEF	900
DVYIEDLRRY PEGTPKNKDY HIGSYFEIYH YIRQRAYLEE VLAKRKEYRD SGSFTDEQLD	960
KLQKILDDIR ARGSYDKNLL KLEYLPFAYN LPRYKNLTTE ALFDDDSVSG KKRVAEWREER	1020
EKTREAEREQ RRQR	1034

(SEQ ID NO:107)。

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

(SEQ ID NO:108)。

[0172] 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 KSVFVLSEK ELLTSFGRGN EAVPEKRVTG	60
GTIANARTDN KEAFSAALQN KRFEVFGRTA GSSDDPLAVS RAPGQDLIGA KTALEERYFG	120
RAFADNIHQ VIYAIQDINK ILAVHANNIV YTLNNLDREA DPETDDFIGS GYLTLKNTFE	180
TYCDPAALNE REREKVTVSK QHFDAFMQNP RLAYYGN AFF RKLSKAERLA RGREIFDKES	240
PERRQEILGS RGKNKSVDDE 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
GGLVEHADEL REYLDNKMLD KTKLRLLPDG KVDTGFRNFI ISNVTESRRF RYLVRycePR	660
AVRDYMSCRP LIRLTLRDMP DTILRRYYEQ SVGAATVDRE RILDTLADKL LSLRFTDFEN	720
VNQRANAERN REKQKMMGII SLYLNVAYQI VKNLVYVNAR YTMAYHCAER DTELLLNAAAG	780
EGNLLRRDRS WPARLHLPRR ALARRRDRVE VMERDVARGP EAYNRDEWLK LVRTLREKR	840
VCDNLHNNYA YLCGADAEPG DASLSLLFVY RNKAAHLSVL NKGGRLSGDL KEAKSWFYVY	900
HFLMQRVLEE EFRNTQALPE RLRELLMAE RYRGCSKDLI KVLNLTFAYN LPRYKNLSID	960
GRFDKNHPDP SDE	973

(SEQ ID NO:109)。

[0173] 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 AKNPTHKKDN EQKNTRQDML GLKSTLEKFY FGSTFDDNIH	120
IQIIHNIQDI AKILAAHSNN AGYALDNMLA YQGVEFSDMI GYMGTSRTFD NYDPNHKNNK	180
DDFRFLKLPR LGYFGSAFYS QKGKDFEKRS DEEVYNICAL MGQIRQCCFH GKQEKYQLKW	240
LYNFHNFKSN KPFLDTLDKH FDEMIDRINK NFIKNNTPD L ILSGLYPDM AKKELVRLFY	300
DFTTVKEYKN MGFSVKKLRE KMLESEEEASD FRDKDYDSVR RKLYKLMDFC IYYLYYSDSE	360
RNENLVSRRL 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 DLRHLISEG	720
DSSCNHFISH NRRMRDCIEE NVKNSEQLIF GKEDAVIRFR NNVAHLSAIR NANEYIGDIR	780
EITSYFALYH YLMQRKLIDD CKVNDTAHKY FEQLTKYKTY VMDMVKALCS PFGYNLPRFK	840
NLSIEGKFDM HESK	854

(SEQ ID NO:110)。

[0174] 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 YYEVCVGNKD IKKTIEEKID	660
ALTDIIIVNMN 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)。

[0175] 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 AKAAGVKSVE AIGNDLLMTG FGPNDATIE KRVEQNRAIE TLSSPEQYSA 120
 EFQNKQFKIK GNIKVLNHST QKMEEIQTEL QDNYNRPHFD LLGCKNVLEQ KYFGRTFSDN 180
 IHVQIAYNIM DIEKLLTPYI NNIIYTLNEL MRDNSKDDFF GCDSHFSVAY LYDELKAGYS 240
 DRLKTKPNLS KNIDRIWNNF CNYMNSDSGN TEARLAYFGE LFYKPKETGD AKSDYKTHLS 300
 NNQKEEWELK SDKEVYNIFA ILCDLRHFT HGESITPSGK PFPYNLEKNL FPEAKQVLNS 360
 LFEEKAESLG AEAFGKTAGK TDVSILLKVF EKEQASQKEQ QALLKEYYDF KVQKTYKNMG 420
 FSIKKLREAI MEIPDAAKFK DDLYSSLRHK LYGLDFDILV KHFLDTSDE NLQNNDFRQ 480
 LRACRCEEEK DQVYRSIAVK VWEKVKKKEL NMFQVQVVIP SLSKDELKQM EMTKNTELLS 540
 SIETISTQAS LFSEMIFMMT YLLDGKEINL LCTSLIEKFE NIASFNEVLK SPQIGYETKY 600
 TEGYAFFKNA DKTAKELRQV NNMARMTKPL GGVNTKCVMY NEAAKILGAK PMSKAELESV 660
 FNLDNHDYTY SPSGKKIPNK NFRNFIINNV ITSRRFLYLI RYGNPEKIRK IAINPSIISF 720
 VLKQIPDEQI KRYPPCIGK 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)。

[0176] 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：

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 tgtaggggtc taacac 36

(SEQ ID NO:113)。

[0177] 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 LARSEEYDIF YNSQSENSGR DCDVIGSLYY QASYRNQDAN RFEKDGKKKA	240
IDSLLDDTRA YYTYFDGLFS VPKREDDGKI KESEKEKAKD QNFDVLRLLS VGRQLTFHSD	300
KSNNEAYLFD LSKLTRAAQD ENRRQDIQSL LNILNSTCRS NLEGVNGDFV KHAKNNLYVL	360
NQLYPSLKAN DLIGEYYNFI VKKENRNIGI RLITVRELII EHNYTNLKDS KYDTYRNKIY	420
TVLNFILFRE IQENSIAIKN FREKLRSTEK AEQPALYQAF ANKIYPMVQA KFAKAIDLFE	480
EQYKTKFKSE FKGGISIE NM 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 RYYYTITGNE	720
ALGDAESLKG VIIEQLHAFS IKNTLLSIKN MGEGEYKIQQ IGSSKEKLKA IVNLYLTWAY	780
LLTKSLVKVN IRFSIAFGCL ERDLVLQKKS EKKFDAINE ILLEDDKIRK ECDKERAQAK	840
TLPRELAQER FAQIKRRESG CYFKSYHVD YLSKNSNEFK QNHIDFAVTS YRNNVEHLNV	900
VHCMTKYFSE VKDVKSYYGV YCYIMQRMLC DELIKNQDK PDVRQTFEEY NRLLKDHGTY	960
SKNLMWLLNF PFAYNLARYK NLSNEDLFNA KNNDQKSK	998

(SEQ ID NO:114)。

[0178] 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 RLEQEFFGKT 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 FFVCKFLDGK EINELLCALI      540
NKFDNIADIL DAAAQCGSSV WFVDSYRFFE RSRRISAQIR IVKNIASKDF KKSCKDSDES      600
YPEQLYLDAL ALLGDVISKY KQNRDGSVVI DDQGNVLTQ QYKRFYEFF EEIKRDESGG      660
IKYKKSCKPE YNHQRRNFIL NNVLKSCKWF YVVKYNRPSS CRELMKNKEI LRFVLRDIPD      720
SQVRRYFKAV QGEEAYASAE AMRTRLVDAL SQFSVTACLD EVGGMTDKEF ASQRAVDSKE      780
KLRAIIRLYL TVAYLITKSM VKVNTREFSIA FSVLERDYLL LIDGKKKSSD YTGEDMLALT      840
RKFDVGEDAGL YREWKEKNAE AKDKYFDKAE RKKVLRQNDK MIRKMHFTPH SLNYVQKNLE      900
SVQSNGLAAV IKEYRNAVAH LNIINRLDEY IGSARADSY SLYCYCLQMY LSKNFSVGYL      960
INVQKQLEEH HTYMKDLMWL LNIPFAYNLA RYKNLSNEKL FYDEEAAAEEK ADKAENERGE     1020

```

(SEQ ID NO:115)。

[0179] 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 W0 2019/006471 (Cas13d), 将其通过引用以其整体并入本文。

[0180] 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:

[0181] Cas13d (生黄瘤胃球菌 (*Ruminococcus flavefaciens*) XPD3002) 序列:

```

1 IEKKKSFAKG MGVKSTLVSG SKVYMTTFAE GSDARLEKIV EGDSIRSUNE GEAFSAEMAD
61 KNAGYKIGNA KFSHPKGYAV VANNPLYTGP VQQDMLGLKE TLEKRYFGES ADGNDNICIQ
121 VIHNLIDIEK ILAEYITNAA YAVNNISGLD KDIIGFGKFS TVYTYDEFKD PEHHRAAFNN
181 NDKLINAIAK QYDEFDNFLD NPRLGYFGQA FFSKEGRNYI INYGNCEYDI 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 ADIQKKQGQN GKNQIDRYYE TCIGKDKGKS
661 VSEKVDALTK IITGMNYDQF DKKRSVIEDT GRENAEREKF KKIISLYLTV IYHILKNIVN

```

721 INARYVIGFH CVERDAQLYK EKGYDINLKK LEEKGFSSVT KLCAGIDETA PDKRKDVEKE
781 MAERAKESID SLESANPKLY ANYIKYSDEK KAEETFTRQIN REKAKTALNA YLRNTKWNVI
841 IREDLLRIDN KTCTLFANKA VALEVARYVH AYINDIAEVN SYFQLYHYIM QRIIMNERYE
901 KSSGKVSEYF DAVNDEKKYN DRLLKLLCVP FGYCIPRFKN LSIEALFDRN EAAKFDKEKK
961 KVSGNS (SEQ ID NO: 45)。

[0182] 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：

[0183] Cas13d(重叠群e-k87_11092736)：

MKRQKTFAKRIGIKSTVAYGQGYAITTFGKSGSKAEIAVRSDPPEETLPTESDATLSIHAKFAKAGRDGREFKCG
DVDTRIHTSRSEYESLISNPAESPREDYLGKGTLEKFFGDEYPKDNLRIQIIYSILDIQKILGLYVEDILHFV
DGLQDEPEDLVGLGLGDEKMQKLLSKALPYMGFFGSTDVFKVTKKREERAADEHNAKVFRALGAIRQKLAHFKWK
ESLAIFGANANMPIRFFQGATGGRQLWNDVIAPLWKKRIERVRKSFLSNSAKNLWVLYQVFKDDTDEKKKARARQY
YHFSVLKEGKNLGFNLTKTREYFLDKFFPIFHSSAPDVKRKVDTFRSKFYAILDFFIYEASVSVANSQMGKVAPW
KGAIIDNALVKLREAPDEEAKKINYNVLAASIRNDSLFLRLKSACDKFGAEQNRPVFPNELRNNRDIRNVRSEWLEA
TQDVDAAAFVQLIAFLCNFLEGKEINELVTALIKKFEGIQALIDLLRNLEGVDSIRFENEFALFNDDKGNMAGRIA
RQLRLLASVGKMKPDMTDAKRVLYKSALEILGAPPDEVSEWLAENILDKSNNDYQKAKKTVNPFERNYIAKNVIT
SRSFYYLVRYAKPTAVRKLMSPKIVRYVLKRLPEKQVASYYSAIWTQSESNSNEMVKLIEMIDRLTTEIAGFSFA
VLKDKDSIVSASRESRAVNLEVERLKKLTLYMSIAYIAVKSIVKVNARYFIAYSALERDLYFFNEKYGEEFRLH
FIPYELNGKTCQFEYLAILKYLLARDEETLKRKCEICEEIKVGCEKHKNANPPYEYDQEWIDKKKALNSERKACE
RRLHFSTHWAQYATKRDENMAKHPQKWYDILASHYDELLALQATGWLATQARNDAEHLNPNVNEFDVYIEDLRRYPE
GTPKNKDYHIGSYFEIYHYIRQRAYLEEVLAKRKEYRDSGSFTDEQLDKLQKILDDIRARGSYDKNLLKLEYLPFA
YNLPRYKNLTTEALFDDDSVSGKKRVAEWREREKTTREAREQRRQR (SEQ ID NO: 46)。

[0184] 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：Cas13d(重叠群e-k87_11092736)同向重复序列：GTGAGAAGTCTCCTTATGGGAGATGCTAC (SEQ ID NO:47)。

[0185] 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：

[0186] Cas13d(160582958_基因49834)：

MKNSVTFKLIQAQENKEAARKKAKDIAEQARIAKRNGVVKKEENRINRIQIEIQTQKKSNTQNAYHLKSLAKAAGV
KSVFAIGNDLLMTGFGPGNDATIEKRVFQNRAIETLSSPEQYSAEFQNKQFKIKGNIKVLNHSTQKMEEIQTTELQD

NYNRPHFDLLGCKNVLEQKYFGRTFSDNIHVQIAYNIMDIEKLLTPYINNIIYTLNELMRDNSKDDFFGCDSHFSV
AYLYDELKAGYSDRCLKTPNLSKNIDRIWNNFCNYMNSDSGNTTEARLAYFGELFYKPKETGDAKSDYKTHLSNNQK
EEWELKSDKEVYNIFAILCDLRHFCTHGESITPSGKPFYPYNLEKNLFPEAKQVLNSLFEEKAESLGAEAFGKTAGK
TDVSILLKVFEKEQASQKEQQALLKEYYDFKVQKTYKNMGFSIKKLREAIMEIPDAAKFKDDLYSSLRHKLYGLFD
FILVKHFLDTSSENLONNDFRQLRACRCEEKQVYRSIAVKVWEKVKKELNMFQVQVVIPSLSKDELKQMEM
TKNTELLSSIETISTQASLFSEMIFFMTYLLDGKEINLLCTSLIEKFENIASFNEVLKSPQIGYETKYTEGYAFFK
NADKTAKELRQVNNMARMTKPLGGVNTKCVMYNEAAKILGAKPMSKAELESVFNLDNHDTYSPSGKKIPNKNFRN
FIINNVTSSRRFLYLIRYGNPEKIRKIAINPSIISFVLKQIPDEQIKRYPPCIGKRTDDVTLMRDELGKMLQSVN
FEQFSRVNNKQNAKQNPNGEKARLQACVRLYLTVPYLFIKNMVNINARYVLAHFCLERDHALCFNSRKLNDDSYNE
MANKFQMRKAKKEQYEKEYCKKQETGTAHTKKIEKLNQQIAYIDKDIKNMHSYTCRNYRNLVAHLNVVSKLQNY
VSELPNDYQITSYFSFYHYCMQLGLMEKVSSKNIPLVESLKNEANDAQSYSAKKTLEYFDLIEKNRTYCKDFLKAL
NAPFSYNLPRFKNLSIEALFDKNIVYEQADLKKE (SEQ ID NO: 48)。

[0187] 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:

[0188] Cas13d (160582958_基因49834) 同向重复序列:GAACACACCCCTCTGTTCTTGTAGGGG TCTAACAC (SEQ ID NO:49)。

[0189] 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:

[0190] 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):

MKKQSKKTVSKTSGLEALSVQGTVMITSFGKGNMANLSYKIPSSQKPQNLNSSAGLKNVEVSGKKIKFQGRHPK
IATTDNPLFKPQPGMDLLCLKDKLEMHYFGKTFDDNIHQIQLIYQILDIEKILAVHVNINIVFTLDNVLHPQKEELTE
DFIGAGGWRINLDYQTLRGQTNKYDRFKNYIKRKELLYFGEAFYHENERRYEDIFAILTLLSALRQFCFHSDDLSS
DESDHVNSFWLYQLEDQLSDEFKETLSILWEEVTERIDSEFLKTNVTNLHILCHVFPKESKETIVRAYYEFLIKKS
FKNMGFSIKKLREIMLEQSDLKSFKEDKYNVRAKLYKLDFDIITYYYDHHAFEKEALVSSSLRSSLTEENKEEIIYI
KTARTLASALGADFKKAAADVNAKNIRDYQKKANDYRISFEDIKIGNTGIGYFSELIYMLTLLLDGKEINDLLTTL
INKFDNIISFIDILKKNLEFKFKPEYADFFNMTCRYTLEELRVINSIARMQKPSADARKIMYRDALRILGMDNR
PDEEIDRELERTMPVGADGKFIKKGQGFNFNFIASNVISSRFHYLVRYNNPHKTRTLVKPNPVVKFVLEGIPETQI
KRYFDVCKGQEIPPTSDKSAQIDVLARIISSVDYKIFEDVPQSAKINKDDPSRNFSDALKKQRYQAIVSLYLTVMY
LITKNLVYVNSRYVIAFHCLERDAFLHGVTLPKMNKKIVYSQLTTHLLTDKNYTTYGHLKNQKGHRKWYVLVKNNL
QNSDITAVSSFRNIVAHISVVRNSNEYISGIGELHSYFELYHYLVQSMIAKNNWYDTSHQPKTAEYLNLLKKHHTY
CKDFVKAYCIPFGYVVPYKYNLTINELFDRNPNPEPKKEEV (SEQ ID NO: 50)。

[0191] 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: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) 同向重复序列:CAACTACAACCCCGTAAAAATACGGGGTTCTGAAAC (SEQ ID NO:51)。

gRNA靶序列

[0192] 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。

[0193] 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。

[0194] 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。

[0195] 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。

[0196] 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。

[0197] 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。

[0198] 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。

[0199] 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。

[0200] 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。

[0201] 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。

[0202] 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。

[0203] 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。

[0204] 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

[0205] 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。

[0206] 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。

[0207] 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。

[0208] 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。

[0209] 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。

[0210] 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。

[0211] 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）、细胞外或分泌体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）。

[0212] 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。

[0213] 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 (polymersome)、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。

[0214] 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(包括其中所述双链序列包含一个RNA分子和一个DNA分子的那些实施方案)中,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。

RNA结合内切核酸酶

[0215] 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。

[0216] 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RNA酶或由其组成。

[0217] 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, RNA酶1蛋白包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGLCKPVNTFVHEPLVDVQNVCFQEKVTCCKNGQGNCKYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHIIIVACEGSPYVPVHFDASVEDST (SEQ ID NO:20)。

[0218] 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, RNA酶4蛋白包含以下序列或由其组成:

QDGMVQRFLRQHVHPEETGGSDRYCDLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICSTTNIQCKNGKMNCHGVVKTDCRDTGSSRAPNCRYRAIASTRRVIIACEGNPQVPVHFDG (SEQ ID NO:21)。

[0219] 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, RNA酶6蛋白包含以下序列或由其组成:

WPKRLTKAHWFEIQHIQPSPLQCNRAMSGINNYTQHCKHQNTFLHDSFQNVAAVCDLLSIVCKNRRHNCHQSSKPVNMTDCRLTSGKYPQCRYSAQAQYKFFIVACDPPQKSDPPYKLVPHLDSIL (SEQ ID NO:22)。

[0220] 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, RNA酶7蛋白包含以下序列或由其组成:

APARAGFCPLLLLLLLGLWVAEIPVSAKPKGMTSSQWFKIQHMQPSPQACNSAMKNINKHTKRCKDLN

TFLHEPFSSVAATCQTPKIACKNGDKNCHQSHGVPVSLTMCKLTSGKYPNCRYKEKRQNKSYVVACKPPQKKDSQQFHLVPVHLDRVL (SEQ ID NO:23)。

[0221] 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,RNA酶8蛋白包含以下序列或由其组成:

TSSQWFKTQHVQPSPQACNSAMSIINKYTERCKDLNTFLHEPFSSVAITCQTPNIACKNSCKNCHQSHGPMSLTMGELTSGKYPNCRYKEKHLNTPYIVACDPPQQGDPGYPLVPVHLDKVV (SEQ ID NO:24)。

[0222] 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,RNA酶2蛋白包含以下序列或由其组成:

KPPQFTWAQWFETQHINMTSQQCTNAMQVINNYQRRCKNQNTFLLTTFANVVNVCNPNMTCPNKTCKNCHHSGSQVPLIHCNLTTPSPQNISNCRYAQTANMFYIVACDNRDQRRDPPQYPVVPVHLDRII (SEQ ID NO:25)。

[0223] 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,RNA酶6PL蛋白包含以下序列或由其组成:

DKRLRDNHEWKKLIMVQHPETVCEKIQNDCRDPPDYWTIHGLWPKSEGCNRSWPFNLEEIKKNWMEITDSSLPSPSMGPAPPRWMRSTPRRSTLAEAWNSTGSWTSTGGCALPPAALPSGDLCCRPSTAGSRGVGVDLTALHQLLVHYSATGIIPCECSEPTKPFQIILHHDHTEWVQSIGMPIWGTISSSESAIGKNEESQPACAVLSHDS (SEQ ID NO:26)。

[0224] 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,RNA酶L蛋白包含以下序列或由其组成:

AAVEDNHLLIKAVQNEDVDLVQQLEGGANVNFQEEEGGWTPHNAVQMSREDIVELLRLRHGADPVLRLKNGATPFILAAIAGSVKdLLKLFLSKGADVNECDFYGFATFMEAAYGKVKALKFLYKRGANVNLRRKTKEDQERLRKGGATALMDAAEKHVEVLKILLDEMGADVACDNMGRNALIHALLSSDDSDVEAITHLLLDHGADVNVRRGERGKTPLILAVEKKHLGLVQRLLQEHEIINDTSDGKTALLLAVELKLKKIAELLCKRGASTDCGDLVMTARRNYDHS LVKVLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLKFFIDEKYKIADTSEGGIYLGFIYEKQEVAVKTFCEGSPRAQREVSCQSRENSHLVTFYGSSEHRGHLFVCVTLCSEQTLEACLDVHRGEDVENEDEFARNVLSSIFKAVQELHLSCGYTHQDLQPQNILIDSKKAAHLADFDKSIKWAGDPQEVKRDLEDLGRVLVYVVKKGSISFEDLK AQSNEEVVQLSPDEETKDLIHRLFHPGEHVRDCLSDLLGHPFFWTWESRYRTLNRVGNESDIKTRKSESEILRLLQPGPSEHSKSFDKWTTKINCEVMKKMNKFYEKRGNFYQNTVGDLLKFI RNLGEHIDEKHKMKLKIIGDPSLYFQKTFPDLVIYVYTKLQNTTEYRKHFQTHSPNKPQCDGAGGASGLASPGC (SEQ ID NO:27)。

[0225] 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,RNA酶T2蛋白包含以下序列或由其组成:

VQHPETVCEKIQNDCRDPPDYWTIHGLWPKSEGCNRSWPFNLEEIKDLLPEMRAVWPDIHSFPNRSRFWKHEWEKHGTCAAQVDALNSQKKYFGRSLELYRELDLNSVLLKLGKPSINYQVADFKDALARYGVIPKIQCLPPSQDEEVQTIGQIELCLTKQDQQLQNCTEPGEQSPKQEVWLANGAAESRGLRVCEGDPVFYPPPKTKH (SEQ ID NO:28)。

[0226] 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,RNA酶11蛋白包含以下序列或由其组成:

EASESTMKIIKEEFTDEEMQYDMAKSGQEKQTIEILMNPILLVKNTSLSMSKDDMSSTLLTFRSLHYN DPKGNSSGNDKECCNDMTVWRKVSEANGSCKWSNNFIRSSTEVMRRVHRAPSCKFVQNPGISCCESLELENTVCQF

TTGKQFPRCQYHSVTSLEKILTVLTGHSLMSWLVCQSKL (SEQ ID NO:29)。

[0227] 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,RNA酶T2样蛋白包含以下序列或由其组成:

XLGGADKRLRDNHEWKKLIMVQHPETVCEKIQNDCRPPDYWTIHGLWDPKSEGCNRSWPFNLEEIK
DLLPEMAYWPDVHSFPNRSRFRWKHEWEKHGTCAAQVDALNSQKKYFGRSLELYRELDLNSVLLKLGIKPSINYY
QTTEEDLNL DVEPTTEDTAEVETIHVLLHSALFGEIGPRRW (SEQ ID NO:30)。

[0228] 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。

[0229] 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:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHII VACEGSPYVPVHFDASVEDST (SEQ ID NO:116)。

[0230] 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)) 多肽包含以下序列或由其组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHII VACEGSPYVPVHFEASVEDST (SEQ ID NO:117)。

[0231] 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:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGRCPVNTFVHEPLVDVQNVCFQEKVTCKNG
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHII VACEGSPYVPVNFEDASVEDST (SEQ ID NO:118)。

[0232] 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
QGNCYKSNSMHIITDCRLTNGSRYPNCAYRTSPKERHII VACEGSPYVPVNFEDASVEDST (SEQ ID NO:119)。

[0233] 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
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHII VACEGSPYVPVNFEDASVEDST (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:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGD CRPVNTFVHEPLVDVQNVCFQEKVTCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHII VACEGSPYVPVNFEDASVEDST (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:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCPVNTFVHEPLVDVQNVCFQEKVTCCKDG
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所述序列组成:

KESRAKKFQRQHMDSDSSPSSSSTYCNQMMRRRNMTQGDCRPVNTFVHEPLVDVQNVCFQEKVTCCKDG
QGNCYKSNSMHIITDCRLTADSDYPNCAYRTSPKERHIIIVACEGSPYVPVNFASVEDST (SEQ ID NO:208)。

[0234] 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:

APVEHVADAGAFRLHAALQDIGKNIYTIREVVTEIRDKATRRRLAVLPYELRFKEPLPEYVRLVTEF
SKKTGDYPSLSATDIQVLALTYQLEAEFVGVSHLKQEPQKVKVSSSIQHPETPLHISGFHLPYKPKPPQETEKGHS
ACEPENLEFSSFMFWRNPLPNIDHELQELLIDRGEDVPSEEEEEENGFEKDDSDDDGGGWITPSNIKQIQQEL
EQCDVPEDVRVGCLTTDFAMQNVLLQMGLHVLAVNGMLIREARSYILRCHGCFKTTSDMSRVFCSHCNKTLLKKVS
VTV (SEQ ID NO:31)。

[0235] 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 (ENDO V) 或由其组成。在一些实施方案中,ENDO V蛋白包含以下序列或由其组成:

AFSGLQRVGGVDVSFVKGDSVRACASLVLSFPELEVVEESRMVSLTAPYVSGFLAFREVPFLLELV
QQLREKEPGLMPQVLLVDGNGVLHHRGFGVACHLGVLTDLPCVGVAKKLLQVDGLENNALHKEKIRLLQTRGDSFP
LLGDSGTVLGMALRSHDRSTRPLYISVGHRMSLEAAVRLTCCCCRFRIPEPVRQADICSREHIRKS (SEQ ID NO:
32)。

[0236] 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,ENDO G蛋白包含以下序列或由其组成:

AELPPVPGGPRGPGLAKYGLPGLAQLKSRESYVLCYDPRTRGALWVVEQLRPERLRGDGRRECDFR
EDDSVHAYHRATNADYRGSGFDRGHLAAAANHRWSQKAMDDTFYLSNVAPQVPHLNQNAWNNLEKYSRSLTRSYQN
VYVCTGPLFLPRTEADGKSYVKYQVIGKNHVAVPTHFFKVLILEAAGGQIELRTYVMPNAPVDEAIPLERFLVPIE
SIERASGLLFVPNILARAGSLKAITAGSK (SEQ ID NO:33)。

[0237] 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 (ENDO D1) 或由其组成。在一些实施方案中,ENDO D1蛋白包含以下序列或由其组成:

RLVGEEAGFGCECDKFFYAGTPPAGLAADSHVKICQRAEGAERFATLYSTRDRIPVYSAFRAPRPAPG
GAEQRWLVEPQIDDPNSNLEEAINAEAITSVNSLGSKQALNTDYLDSDYQRGQLYPFSLSSDVQVATFTLTNSAP
MTQSFQERWYVNLHSLMDRALTPQCGSGEDLYILTGTVPDYSYRVKDKVAVPEFVWLAACCAVPGGGWAMGFVKHTR
DSDI IEDVMVKDLQKLLPFNPQLFQNNCGETEQTDEKMKKILEVVNQIQDEERMVQSQSSSPLSSTRSKRSTLLP
PEASEGSSSFLGKLMGFIATPFIKLFQLIYYLVVAILKNIVYFLWCVTKQVINGIESCLYRLGSATISYFMAIGEE
LVSIPWKVLKVVAKVIRALLRILCCLLKAICRVLSIPVRVLVDVATFPVYTMGAIPVCKDIALGLGGTVSLLFDT
AFGTLGGLFQVVSFVCKRIGYKVTFDMSGEL (SEQ ID NO:34)。

[0238] 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) 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:

MGIIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQNEEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDCKHL LSLMGIPYLDAPSEAEASCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQELGLNQE QFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFLEPEVLDPESVEL KWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGRLLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGA AGKFKRGK (SEQ ID NO:35)。

[0239] 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:

MESGQPARRIAMAPLLEYERQLVLELLDGLVVCARGLGADRLLYHFLQLHCHPACLVVLNTQPAE EEFYFINQLKIEGVEHLPRRVITNEITSNSRYEVYTQGGVIFATSRILVVDFLTDRIPSDLITGILVYRAHRIIESCQ EAFILRLFRQKNKRGFIKAFTDNAFAFDTGCHVERVMRNLFVRKLYLWPRFHVAVNSFLEQHKPEVVEIHVSMTPTMLAIQTAILDILNACLKELKCHNPSLEVEDLSLENAIGKPFDKTIRHYLDPLWHQLGAKTKSLVQDLKILRTLQ YLSQYDCVTFLNLLLESLRATEKAFGQNSGWFLDSSSTSMFINARARVYHLPDAKMSKKEKISEKMEIKEGEGILWG (SEQ ID NO:124)。

[0240] 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:

CSPQESGMTALSARMLTRSRLGPGAGPRGCREEPGLRRREAAAEARKSHSPVKRPRKAQRLRVAYE GSDSEKGEAEPLKVPVWEPQDWQQQLVNIRAMRNKKDAPVDHLGTEHCYDSSAPPKVRRYQVLLSLMLSSQTKDQ VTAGAMQRLRARGLTVDLSILQTDDATLGKLIYPVGFWRSKVKYIKQTSAILQQHYGGDIPASVAELVALPGVGPKM AHLAMAVAWGTVSGIAVDTHVHRIANRLRWTKKATKSPEETRAALEEWLPRELWHEINGLLVGFGQQTCLPVHPRC HACLNQALCPAAQGL (SEQ ID NO:123)。

[0241] 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 CIEKLPTFHFCCEKPKVNFTTKILNVYQKDVLGDGYVCVIQVEPFCCVVFAEAPDSWIMKDNSVTRLTAEQWVVMML DTQSAPPSLVTDYNSCLISSASSARKSPGYPIKVHKFKEALQ (SEQ ID NO:36)。

[0242] 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:

TLQGTNTYLVGTGPRRILIDTGEPAPIEYISCLKQALTEFNIAIQEIVVTHWHRDHSGGIGDICKSIN NDTTYCIKKLPRNPQREEIIGNGEQQYVYLKGDVIKTEGATLRVLTPGHTDDHMLLLEENAIIFSGDCILGEG TTVFEDLYDYMNSLCELLKIKADIIYPGHGPVIHNAEAKIQQYISHRNIREQQILTLFRENFEKSFTVMELVKI IY KNTPENLHEMAKHNNLLHLKKLEKEGKIFSNTDPDKKWK AHL (SEQ ID NO:37)。

[0243] 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 I ILGDLNTAHRPIDHWDVNLECFEEDP
GRKWMDSLLSNLGCQSASHVGPFI DSYRCFQPKQEGAF TCWSAVTGARHLNYGSRLDYVLGDR TLVIDTFQASFL
PEVMGSDHCPVGAVLSVSSVPAKQCPPLCTRFLPEFAGTQLKILRFLVPLEQSPVLEQSTLQHNNQTRVQTCQNK
QVRSTRPQPSQVGSSRGQKNLKS YFQPS PSCPQASPDIELPSLPLMSALMTPKTPEEKAVAKVVKGQAKTSEAKDE
KELRTSFWKS VLAGPLRTPLCGGHREPCVMRTVKKPGPNLGRRFYMCARPRGPPTDPSSRCNFFLWSRPS (SEQ ID
NO:38)。

[0244] 在一些实施方案中, APEX2多肽包含以下序列或由其组成: MLRVVSWNINGIRRPLQGV
ANQEPSNCAAVAVGRILDELADIVCLQETKVTRDALTEPLAIVEGYNSYFSFSRNRSGYSGVATFCKDNATPVAA
EEGLSGLFATQNGDVGCYGNMDEFTQEELRALDSEGRALLTQHKIRTWEGKEKTLTLINVYCPHADPGRPERLVFK
MRFYRLLQIRAEALLAAGSHV I ILGDLNTAHRPIDHWDVNLECFEEDPGRKWMDSLLSNLGCQSASHVGPFI DSY
RCFQPKQEGAF TCWSAVTGARHLNYGSRLDYVLGDR TLVIDTFQASFLLEVMGSDHCPVGAVLSVSSVPAKQCP
LCTRFLPEFAGTQLKILRFLVPLEQSP (SEQ ID NO:39)。

[0245] 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
DGLRAWIKKKGLDWVKEEAPDILCLQETKCS ENKLPAELQELPGLSHQYWSAPSDKEGYSGVGLLSRQCPLKVSYG
IGDEEHDQEGRVIVA EFDSFVLVTAYVPNAGRGLVRLEYRQRWDEAFRKF LKGLASRKPLVLCGDLNVAHEEIDLR
NPKGKNKNAGFTPQERQGF GELLQAVPLADSFRLHY PNTPYAYTFWTYMMNARSKNVGWRLDYFLLS (SEQ ID
NO:125)。

[0246] 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:

QDNSRYTHFLTQH YDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHG NKRSIKAICENKNGNPHREN
LRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVACENGLPVHLDQSIFRRP (SEQ ID NO:40)。

[0247] 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:

SSLIRRVISTAKAPGAIGPYSQAVLVDR TIYISGQIGMDPSSGQLVSGGVAEEAKQALKNMGEILKAA
GCDFTNVVKTTVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGSRIEIEAVAIQGPLTTASL (SEQ ID
NO:41)。

[0248] 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:

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

[0249] 在一些实施方案中, ZC3H12A多肽包含以下序列或由其组成: SGPCGEKPVLEASPTMS
LWEFEDSHSRQGTTPRPGQELAAEEASALELQMKVDFFRKLGYSSTEIHSVLQKLG VQADTNTVLGELVKHGTATER
ERQTSPDPCPQLPLVPRGGGTPKAPNLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFLER

GHTDITVFVPSWRKEQPRPDVPTDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIVKLAYESDGI VVSNDTY
RDLQGERQEWRKFI EERLLMYSFVNDKFMPDDPLGRHGPSLDNFLRKKPLTLEHRKQPCPYGRKCTYGIKCRFFH
PERPSCPQRSVADEL RANALLSPPRAPSKDKNGRRPSPSSQSSSLLTESEQCSLDGKKLGAQASPGSRQEGLTQTY
APSGRSLAPSGGSGSSFGPTDWLPQTLD SLPYVSQDCLDSGIGSLESQMSELWGVRRGGGPGEPPRAPYTGYSPY
GSELPATAAFSAFGRAMGAGHFSVPADYPPAPPAPFPREYWSEPYPLPPPTSVLQEPPVQSPGAGRSPWGRAGSLA
KEQASVYTKLCGVFPFHLVEAVMGRFPQLLDPQQLA AEILSYKSQHPSE (SEQ ID NO:43)。

[0250] 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:

SSLIRRVISTAKAPGAIGPYSQAVLVDR TIYISGQIGMDPSSGQLVSGGVAEEAKQALKNMGEILKAA
GCDFTNVVKT TVLLADINDFNTVNEIYKQYFKSNFPARAAYQVAALPKGSRIEIEAVAIQGPLTTASL (SEQ ID
NO:44)。

[0251] 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 LPEGCPCGLPHGESALSRLLRALLAARASLDLCLFAFSSPQLG
RAVQLLHQRGVVRVVTDCDYMALNGSQIGLLRKAGIQVRHDQDPGYMHKFAIVDKRVLITGSLNWTQAIQNNR
ENVLITEDDEYVRLFLEEFERIWEQFNPTKYTFPPPKSHGSCAPPVSRAGGRLLSWHRTC GTSSSESQT (SEQ ID
NO:126)。

[0252] 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:

KARYKTLEPRGYSLIRGLIHS DRWREALLLLEDIKKVITPSKKNYNDCIQGALLHQDVNTAWNLYQE
LLGHDI VPMLETLKAFFDFGKD IKDDNYSNKLLDILSYLRNNQLYPGESFAHSIKTWFE SVPGKQWKQGFTTVRKS
GQCSGCGKTIESIQLSPEEYECLKGKIMRDVIDGGDQYRKTT PQELKRFENFIKSRPPFDVVIDGLNVAKMFPKVR
ESQLLLNVVSQ LAKRNLRLVLGRKHMLRRSSQWSRDEMEEVQKQASCFFADDISEDDPFLLYATLHSGNHCRFIT
RDLMRDHKACL PDAKTQRLFFKWQQGHQLAIVNRFP GSKLTFQRILSYDTVVQTTGDSWHIPYDEDLVERCSCEVP
TKWLCLHQKT (SEQ ID NO:127)。

[0253] 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:

SVEPMFRHLKNTYAGLQLVVVILPGKTPVYAEVKRVGDTV LGMATQCVQMKNVQRTPQTLSNLCLKI
NVKLGGVNNILLPQGRPPVFQQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDAHPNRYCATVRVQQHRQEIIQDLA
AMVRELLIQFYKSTRFKPTRIIFYRDGVSEGQFQQVLHHELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTD
KNERVGKSGNIPAGTTVDTKITHPT EFD FYLC SHAGIQGTSRPSHYHVLWDDNRFSSDELQILTYQLCHTYVRCTR
SVSIPAPAYYAHLVAFRARYHLVDKEHDSAEGSHTSGQSNGRDHQALAKAVQVHQDTLR TMYFA (SEQ ID NO:
128)。

[0254] 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
TERFEDVWVVSGLPLTLPTQTRGDGKKIVSYQVIGEDNVAVPSHLYKVILARRSSVSTEPLALGAFVVPNEAIGFQPQ
LTEFQVSLQDLEKLSGLVFFPHLDRTSDIRNICSVDTCKLDFQEFTLYLSTRKIEGARSVLRLEKIMENLNKNAEI
EPDDYFMSRYEKKLEELKAKEQSGTQIRKPS (SEQ ID NO:129)。

[0255] 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:

EHPSKMEFFQKLGYDREDVLRVLGKLGEALVNDVLQELIRTGSRPGALEHPAAPRLVPRGSCGVPDS
AQRGPGTALEEDFRTLASSLRPIVIDGSNVAMSHGNKETFSCRGIKLAVDWFRDRGHTYIKVFVPSWRKDPPrADT
PIREQHVLAELERQAVLVYTPSRKVHGKRLVCYDDRYIVKVAYEQDGVIVSNDNYRDLQSENPEWKWFIEQRLLMF
SFVNDRFMPDDPLGRHGPSLSNFLSRKPKPEPSWQHCPYGGKCTYGIKCKFYHPERPHHAQLAVADELRAKTGA
RPGAGAEERPPRAPGGSAGARAAPREPFHSLPPARGSPDLAALRGSFSRLAFSDDLGLPLGPPLVPACSLTPRL
GGPDWVSAGGRVPGPLSLPSPESQFSPGDLPPPPGLQLQPRGEHRPRDLHGDLLSPRRPPDDPWARPPRSDFRPGR
SVWAEPAWGDGATGGLSVYATEDDEGDARARARIALYSVFPRDQVDRVMAAFPELSDLARLILLVQRCQSAGAPLG
KP (SEQ ID NO:130)。

[0256] 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
FNPKDVLGRGAGGTFFVRGQFEGRAVAVKRLRECFGLVRREVQLLQESDRHPNVLRYFCTERGPQFHYIALELCR
ASLQEYVENPDLDRGGLEPEVVLQQLMSGLAHLHSLHIVHRDLKPGNILITGPDSSQLGRVVLSDFGCLKKLPAGR
CSFSLHSGIPGTEGWMAPELLQLLPDSPTS AVDIFSAGCVFYVLSGGSHPFGDSLYRQANILTGAPCLAHLEEE
VHDKVVARDLVGAMLSPLPQPRPSAPQVLAHPFFWSRAKQLQFFQDVSDWLEKESEQEPLVRALEAGGCAVVRDNW
HEHISMLQTDLRKFRSYKGTSVRDLLRAVRNKKHHYRELPEVRQALGQVPDGFVQYFTNRFPRLLLHTHRMR
CASESLFLPYPPDSEARRPCPGATGR (SEQ ID NO:131)。

[0257] 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
LAHICLVTPSMTLTRAKEVNIPKRKGNCQHDRLERFYEQVVQAIQRHIHFDVVKCILVASPGFVREQFCDYL
FQQAVKTDNKLLENRSKFLQVHASSGHKYSLEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRAFYGLK
QVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSLHVSQEQLSQTGVAAILRFVPPEL
SDQEGDSSSEED (SEQ ID NO:132)。

[0258] 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:

SLVIRNLQRVIPIRRAPLRSKIEIVRRILGVQKFDLGIICVDNKNIQHINRIYDRNVPTDVLSPFPFH
EHLKAGEFPQPDFDDYNLGDIFLGVEYIFHQCKENEDYNDVLTVTATHGLCHLLGFTHGTEAEWQQMFQKEKAVL
DELGRRTGTRLQPLTRGLFGGS (SEQ ID NO:133)。

[0259] 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多肽包含以下序列或由其组成:QEVIAGLERFTFAFEKDVEMQKGTGLLPFQGMDSASAVCNFFTKGLCEKGKLCPPFRHDRGEKMMVCKHWLRGLCKKGDHCKFLHQYDLTRMEPCYFYSKFGDCSNKECSFLHVKPAFKSQDCPWYDQG FCKDGPLCKYRHVPRIMCLNYLVGFCPEGPKCQFAQKIREFKLLPGSKI (SEQ ID NO:134)。

[0260] 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:

KLVRKNIEKDNAGQVTLVPEEPEDMWHTYNLVQVGDLSRASTIRKVQTESSTGSGVSNRVRTTLTLCV
EAIDFDSQACQLRVKGTNIQENEYVKMGAYHTIELEPNRQFTLAKKQWDSVVLIERIEQACDPAWSADVAAVVMQEG
LAHICLVTPSMTLTRAKVEVNIPRKRKGNCQHDRLERFYEQVVQAIQRHIHFVVKCILVASPGFVREQFCDYM
FQQAVKTDNKLLENRSKFLQVHASSGHKYSLEALCDPTVASRLSDTKAAGEVKALDDFYKMLQHEPDRAFYGLK
QVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSLHVSQEQLSQLTGVAAILRFPVPEL
SDQEGDSSSEED (SEQ ID NO:135)。

[0261] 在一些实施方案中,hCG_2002731多肽包含以下序列或由其组成:

DPAWSADVAAVVMQEGLAHICLVTPSMTLTRAKVEVNIPRKRKGNCQHDRLERFYEQVVQAIQRHI
HFVVKCILVASPGFVREQFCDYMFQQAVKTDNKLLENRSKFLQVHASSGHKYSLEALCDPTVASRLSDTKAAG
EVKALDDFYKMLQHEPDRAFYGLKQVEKANEAMAIDTLLISDELFRHQDVATRSRYVRLVDSVKENAGTVRIFSSL
HVSQEQLSQLTGVAAILRFPVPELSDQEGDSSSEED (SEQ ID NO:136)。

[0262] 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:

MDPGKDKEGVPQPSGPPARKKFVIPLDEDEVPPGVRGNPVLKFVRNVPWEFGDVIPDYVLGQSTCALF
LSLRYHNLHPDIYIHGRLQSLGKNFALRVLLVQVDVKDPQQALKELAKMCILADCTLILAWSPEEAGRYLETYKAYE
QKPADLLMEKLEQDFVSRVTECLTTVKS VNKTDSQTLTTFGSLEQLIAASREDLALCPGLGPQK (SEQ ID NO:
137)。

[0263] 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:KESRAKKFQRQHMDSDSSPSSSTYCNQMMRRRNMTQGRCKPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNCYKSNSSMHITDCRLTNGSRYPNCAYRTSPKERHII VACEGSPYVPVHFDASVEDST (SEQ ID NO:138)。

[0264] 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:

QDNTRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLTSPCKDINTFIHGKRSIKAICENKNGNPHREN
LRISKSSFQVTTCKLHGGSPWPPCQYRATAGFRNVVACENGLPVHLDQSIFRRP (SEQ ID NO:139)。

[0265] 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:

GLGLVQPSYGQDGMVQRFLRQHVHPEETGGSDRYCNLMMQRRKMTLYHCKRFNTFIHEDIWNIRSICS
TTNIQCKNGKMNCHGVVKTDCRDTGSSRAPNCRYRAIASTRRVVACEGNPQVPVHFDG (SEQ ID NO:140)。

[0266] 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:

XSAVDNILLKLAKFKIGFLRLGQIQKVHPAIQQFTEQEICRSKSIKSLALLEELYNSQLIVATTTCMGI
NHPIFSRKIFDFCIVDEASQISQPICLGPLFFSRRFVLVGHDHQLPPLVLNREARALGMSESLFKRLEQNKSAVVQ
LTVQYRMNSKIMSLSNKLTYEGKLECGSDKVANAVINLRHFKDVKLELEFYADYSDNPWLMGVFEPNNPVCFLNTD
KVPAPQVEKGGVSNVTEAKLIVFLTSIFVKAGCSPSDIGIIAPYRQQLKIINDLLARSIGMVEVNTVDKYQGRDK
SIVLVSFVRSNKDGTVGELLKDWRRRLNVAITRAKHKLILLGCVP SLNCPPEKLLNHLNSEKLISFFFCIWSHLI
ALL (SEQ ID NO:141)。

[0267] 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:

MALRSHDRSTRPLYISVGHRMSLEAAVRLTCCCFRIFPEPVRQADICSREHIRKSLGLPGPPTPRSP
KAQRPVACPKGDSGESSALC (SEQ ID NO:142)。

[0268] 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:

CYTNHALSYDQAKRVPRWVLEHISKSKIMGDADRKHKCFKPDNPPTPSAFNEDYVSGSWRSGHMAP
AGNNKFSSKAMAETFYLSNIVPQDFDNNSGYWNRIEMYCRELTERFEDVWVVSGLTLPQTRGDGKKIVSYQVIGE
DNVAVPSHLYKVLARRSSVSTEPLALGAFVVPNEAIGFQPQLTEFQVSLQDLEKLSGLVFFPHLDRT (SEQ ID
NO:143)。

[0269] 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
EGIRAWTEGEKQQLSTGRVQGYDGYFVLSVEQYLELSDSANNIHFMRQSEIGRR (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:

TVSQPTLLVNGKTRRFTNIEFQYSTLLLSIRYGLTPDTLDEEKARVLDQARQALGTAWAKEQQKARD
GREGSRLWTEGEKQQLSTGRVQGYEGYYVLPVEQYPELADSSSNIQFLRQNMGKR (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 κ (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:

MGWLRPGPRPLCPPARASWAFSHRFPSPLAPRRSPTPFFMASLLCCGPKLAACGIVLSAWGVIMLIML
GIFFNVHSAVLIEDVPFTEKDFENGPNQNIYNLYEQVSYNCFIAAGLYLLLGGFSFCQVRLNKRKEYMVR (SEQ ID
NO:204)。

[0270] 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。在一些实施方案中, 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 IVDIARQSG
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 VAIASNKGKQALEAVKAQL
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 SFEAEFGVG PLDFHLDWLL QILET (SEQ ID NO: 205)。

在一些实施方案中,TALEN多肽包含以下序列或由其组成:

1 mdpirsrtps parellpgpq pdrvqptadr ggappaggl dglparrtms rtrlpsspap
61 spafsagsfs dllrqfdpsl ldtslldsmp avgtphataa paecdevqsg lraaddpppt
121 vrvavtaarp prakpaprrr aaqpsdaspa aqvdlrtlgy sqqqqekikp kvgstvaqhh
181 ealvghgfth ahivalsrhp aalgtvavky qdmiaalpea thedivgvvk qwsgaralea
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 vvaishdgg
541 kqaletvqrl lpvlcqhthgl tpaqvvaia hdggkqalet vqllpvlcqhahgltpdqvv
601 aiasniggkq alatvqrllp vlcqhahgltp dqvvaiaasn gggkqaletvq rllpvlcqhah
661 gltpdqvvai asngggkqal etvqrllpvl cqahgltpdqv vvaiaasnigg kqaletvqrl
721 lpvlcqhahgl tpaqvvaia hdggkqalet vqrllpvlcqhahgltpdqvv aiasngggkq
781 aletvqrllp vlcqhahgltp eqvvaiaasn gggkqaletvq rllpvlcqhah gltpdqvvai
841 asngggkqal etvqrllpvl cqahgltpaq vvaiaasnigg kqaletvqrl lpvlcqhahgl
901 tlaqvvaia niggkqalet 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 eqrdaahlpl swrvkrprtr igggldpdpgt
1321 piaadlaass tvmweqdaap fagaaddfpa fneelawlm ellpqsgsvg gti (SEQ ID
NO: 206)。

[0271] 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。在一些实施方案中,ZNF638多肽多肽包含以下序列或由其组成:

1 MSRPRFNPRG DFPLQRPRAP NPSGMRPPGP FMRPGSMGLP RFYPAGRARG IPHRFAGHES
61 YQNMGPQRMN VQVTQHRDP RLTKEKLDHFH EAQQKKGKPH 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 VVQSRGHGTE FNKQKHLEAA
601 DKGHSPAQKP KTSSGTKPSV KPTSATKSDS NLGGHSIRCK SKNLEDDTLS ECKQVSDKAV
661 SLQRKLRKEQ SLHYGSVLLI TELPEDGCTE EDVRKLFQPF GKVNDVLIVP YRKEAYLEME
721 FKEAITAIMK YIETTPLETK GSKSVKICVPG KKKAQNKEVK KKTLESKKVS ASTLKRDATA
781 SKAVEIVTST SAAKTGQAKA SVAKVNKSTG KSASSVKSVM 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 AEKATCDSD F AVETLELETQ GEEVKEEIP L VASASVSIEQ FTENAEECAL
1201 NQQMFNSDLE KKGAEIINPK TALLPSDSVF AEERNLKGIL EESPSEAEDF ISGITQTMVE
1261 AVAEVEKNET VSEILPSTCI VTLVPGIPTG DEKTVDKKNI SEKKGNMDEK EEKEFNKTET
1321 RMDLQIGTEK AEKNEGRMDA EKVEKMAAMK EKPAENTLFK AYPNKGVGQA NKPDETSKTS
1381 ILAVSDVSSS KPSIKAVIVS SPKAKATVSK TENQKSFPKS VPRDQINAEK KLSAKEFGLL
1441 KPTSARSGLA ESSSKFKPTQ SSLTRGGSGR ISALQGKLSK LDYRDITKQS QETEARPSIM

```

1501 KRDDSNNKTL 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 TPTKRVIRIG
1861 TLPSEKAVVT EPAKGEEAFQ MSEVDEESGL KDSEPERKRK KTEDSSSGKS VASDVPEELD
1921 FLVPKAGFFC PICSIFYSGE KAMTNHCKST RHKQNTKFM AKQRKEKEQN EAEERSSR

```

(SEQ ID NO: 207)。

[0272] 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在本文中以Cas融合蛋白的形式并作为内部对照使用,例如但不限于参见图9,其显示了PIN-dSauCas9、PIN-dSauCas9dHNNH、PIN-dSPCas9和dcjeCas9-PIN。

[0273] 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,核酸酶包含从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,核酸酶包含从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,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。

融合蛋白

[0274] 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。

[0275] 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。

[0276] 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。

[0277] 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US 2016/0238593,将其通过引用以其整体并入本文。

[0278] 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 LQQLQPTTSN NTTNSNTLND YSTKTAYFDN      240
MVSTSGSQMA DNKMNTNNLA IPNSVWSNTR QRSQSNASSI YTDAPLYEQP ARASISSHYT      300
IPTQESPLIA DEIDPQSINW VTMDPTVPSI NQISNLLPTN TISISNVFPL QHQQPQLNNA      360
INLTSTSLAT LCSKYGEVIS ARTLRNLNMA LVEFSSVESA VKALDSLQ GK EVSMIGAPSK      420
ISFAKILPMH QQPPQFLLNS QGLPLGLENN NLQPQPLLQE QLFGNAVTFQ QQGNVSIPVF      480
NQQSQQSQHQ NHSSGSAGFS NVLHGYNNNN SMHGNNNNNSA NEKEQCPFPPL PPPNVNEKED      540
LLREIIELEF ANSDEYQINS LIKKSLNHKG TSDTQNF GPL PEPLSGREFD PPKLRELKRS      600
IDSNAFSDLE IEQLAIAMLD ELPELSSDYL GNTIVQKLFE HSSDIKDIM LRKTSKYLTS      660
MGVHKNGTWA CQKMITMAHT PRQIMQVTQG VKDYCTPLIN DQFGNYVIQC VLKFGFPWNQ      720
FIFESIIANF WVIVQNR YGA RAVRACLEAH DIVTPEQSIV LSAMIVTYAE YLSTNSNGAL      780
LVTWFLDTSV LPNRHSILAP RLTKRIVELC GHRLASLTIL KVLNRYGDDN ARKIILDSL F      840
GNVNAHDSSP PKELTKLLCE TNYGPTFVHK VLAMPLLEDD LRAHIKQVR KVLTDSTQIQ      900
PSRRLLEEVG LASPSSTHNK TKQQQQQHNN SSISHMFATP DTSGQHMRGL SVSSVKSGGS      960
KHTTMNTTTT NGSSASTLSP GQPLNANSNS SMGYFSYPGV FPGVSGFSGNA SNGYAMNND D      1020
LSSQFDMLEF NNGTRLSLPQ LSLTNHNNNTT MELVNVVGSS QPHTNNNNNN NNTNYND DNT      1080
VFETLTLHSA N                                1091
```

(SEQ ID NO: 209)。

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 MEMNMDMDMD MELASIVSSL SALSHSNNNG GQAAAAGIVN GGAAGSQQIG GFRRSSFTTA
61 NEVDSEILLL HGSSESSPIF KKTALSVGTA PPFSTNSKKF FGNGGNYQY RSTD TASLSS
121 ASYNNYHTHH TAANLGKNNK VNHLLGQYSA SIAGPVYYNG NDNNSGGEG FFEKFGKSLI
181 DGTRELESQD RPDVNTQSQ FISKSVSNAS LDTQNTFEQN VESDKNFNKL NRNTTNSGSL
241 YHSSSNSGSS ASLESENAHY PKRNIWNVAN TPVFRPSNNP AAVGATNVAL PNQDGPANN
301 NFPPYMNGFP PNQFHQGPY QNFPPNYLIGS PSNFISQMIS VQIPANEDTE DSNKKKKKA
361 NRPSSVSSPS SPPNNSPFPF AYPNPMFMF PPPLSAPQQQ QQQQQQQQOE DQQQQQQQEN
421 PYIYYPTPNP IPVKMPKDEK TFKKRNNKNH PANNSNNANK QANPYLENSI PTKNTSKKNA
481 SSKSNESTAN NHKSHSHSHP HSQSLQQQQQ TYHRSPLLEQ LRNSSSDKNS NSNMSLKDIF
541 GHSLEFCKDQ HGSRFIQREL ATSPASEKEV IFNEIRDDAI ELSNDVFGNY VIQKFFFEFGS
601 KIQKNTLVDQ FKGNMKQLSL QMYACRVIQK ALEYIDSNQR IELVLELSDS VLQMIKDQNG
661 NHVIQKAIET IPIEKLPFIL SSLTGHIYHL STHSYGCRVI QRLLEFGSSE DQESILNELK
721 DFIPYLIQDQ YGNYVIQYVL QDQFTNKEM VDIKQEIET VANNVVEYSK HKFASNVVEK
781 SILYGSKNQK DLIISKILPR DKNHALNLED DSPMILMIKD QFANYVIQKL VNVSEGEKK
841 LIVIAIRAYL DKLKNSNSLG NRHLASVEKL AALVENAEV (SEQ ID NO: 210)。
```

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 DDVPSVDPVV SETVNSALEQ LQLDDPEENA TSNAFANKVS QDSQFANGPP
61 SQMFPHQMM GGMGFMPYSQ MMQVPHNCP FFPPPDFNDP TAPLSSSPLN AGGPPMLFKN
121 DSLPFQMLSS GAAVATQGGQ NLNPLINDNS MKVLPIASAD PLWTHSNVPG SASVAIEETT
181 ATLQESLPSK GRESNNKASS FRRQTFHALS PTDLINAANN VTLSKDFQSD MQNFSKAKKP
241 SVGANNTAKT RTQISFDNT PSSTSFPPT NSVSEKLSDF KIETSKEDLI NKTAPAKKES
301 PTTYGAAPY GGPLLQPNPI MPGHPHNISS PIYGIRSPFP NSYEMGAQFQ PFSPILNPTS
361 HSLNANSPIP LTQSPIHLAP VLNPSNSVA FSDMKNDGGK PTTDNDKAGP NVRMDLINPN
421 LGPSMQPFHI LPPQQNTPPP PWLYSTPPPF NAMVPPHLLA QNHMPLMNSA NNKHHGRNNN
481 SMSSHNDNDN IGNSNYNNKD TGRSNVGKMK NMKNSYHGYY NNNNNNNNNN NNNNNSNATN
541 SNSAEQQRKI EESSRFADAV LDQYIGSIHS LCKDQHGRF LQKQLDILGS KAADAIFEET
601 KDYTVELMTD SFGNYLIQKL LEEVTTEQRI VLTKISSPHF VEISLNPHGT RALQKLIETI
661 KTDEEAQIVV DSLRPYTVQL SKDLNGNHVI QKCLQRLKPE NFQFIFDAIS DSCIDIATHR
721 HGCCVLQRCL DHGTTEQCDN LCDKLLALVD KLTLDPFQNY VVQYIITKEA EKNKYDYTHK
781 IVHLLKPRAI ELSIHKFGSN VIEKILKTAI VSEPMILEIL NNGGETGIQS LLNDSYGNVY
841 LQTALDISHK QNDYLYKRLS EIVAPLLVGP IRNTPHGKRI IGMHLHDS (SEQ ID NO:
211)。
```

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 GSFDGGAFFS SNIPLFNFMG TGNQRFQYSP
61 HPFAKSSDPC RLAALTPSTP KGPLNLTPAD FGLADFSVGN ESFADFTANN TSFVGNVQSN
121 VRSTRLLPAW AVDNNGNIRD DLTLDQDVVSN GSLIDFAMDR TGVKFLERHF PEDHDNEMHF
181 VLFDKLTEQG AVFTSLCRSA AGNFIIQKFV EHATLDEQER LVRKMC DNGL IEMCLDKFAC
241 RVVQMSIQKF DVSIAMKLVE KISSLDLFLPL CTDQCAIHVL QKVVKLLPIS AWSFFVKFLC
301 RDDNLMTVCQ DKYGCRLVQQ TIDKLSDNPK LHCFNTRLQL LHGLMTSVAR NCFRLSSNEF
361 ANYVVQYVIK SSGVMEMYRD TIEKCLLRN ILSMSQDKYA SHVVEGAFLF APPLLLSEMM
421 DEIFDGYVKD QETNRDALDI LLFHQYGNV VQQMISICIS ALLGKEERKM VASEMRLYAK
481 WFDRIKNRVN RHSGRLERFS SGKKIIESLQ KLNVPMTMTN EPMPYWAMPT PLMDISAHFM
541 NKLNFQKNSV FDE (SEQ ID NO: 212)。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 MTPNRRSTD S YNMLGASFDF DPDFSLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDL DNY
61 IFNSGSGSSD DETPPPA API FISLEEVL LN GLLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKLTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMIDGGL 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 WYEKMKQ RVL QHASRLERFS SGKKIIDSVM RHGVPTAAAI NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 213)。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 MTPNRRSTD S YNMLGASFDF DPDFSLLSNK THKNKNPKPP VKLLPYRHGS NTTSSDSDSY
61 IFNSGSGSSD AETPAPVAPI FISLEDVLLN GQLIDFAIDP SGVKFLEANY PLDSEDQIRK
121 AVFEKFTEST TLFVGLCHSR NGNFIVQKLV ELATPAEQRE LLRQMIDGGL 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 WYEKMKQ RVL QHASRLERFS SGKKIIDSVM RHGVPTAAAV NAQAAPSLME LTAQFDAMFP
481 SFLAR (SEQ ID NO: 214)。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 LQMHAHSQS IRVPVQPSAT GIFSNPYREP TTTDDLTRY
 181 RANPAMMKNL KLSDIRGALL KFAKDQVGSR FIQQELASSK DRFEKDSIFD EVVSNADLV
 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: 215)。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 PIPIPQDGVP QGTPDFQMTF
 121 FLQGGHLLIG GSPNGPVQVS GNWYSGGAGI FSTMQQADPS NGMPGMAAEF VNNENGMPGP
 181 NGMHQQAMIS GSPPFPYQNM MNLTTSFGAM GLGPQQIQQR DPQMFQQPIL HEPIQGMAQN
 241 GFGQQVFFFTQ MQNQOHPQGO AQOQLQQLAQ 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 QPVLDAIQIH LKQLVLDQYG
 541 NYVIQHVIEH GSPSDKEQIV QDVISDDLK FAQHKFASNV IEKCLTFGGH AERNLIIDKV
 601 CGDPNDPSPP LLQMMKDPFA NYVVQKMLDV ADPQHRKKIT LTIKPHIATL RKYNFGKHIL
 661 LKLEKYFAKQ APANSSNSSS NDQIYEHSPF DIPLGADFSN HPF (SEQ ID NO: 216)。

[0279] 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W0 2013/058404通过引用以其整体并入本文。

[0280] 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GGS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。

[0281] 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。

[0282] 在一些实施方案中,本文公开的融合蛋白的所述至少一种RNA结合蛋白还包含编码核定位信号(NLS)的序列。在一些实施方案中,核定位信号(NLS)定位于所述RNA结合蛋白的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,所述至少一种RNA结合蛋白还包含编码第一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,所述第一NLS或所述第二NLS定位于所述RNA结合蛋白的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。

[0283] 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。

[0284] 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。例如,一种这样的示例性融合蛋白是E99,其被配置为使得RNA酶1(R39D、N67D、N88A、G89D、R19D、H119N、K41R)位于SpyCas9的N末端;而另一种示例性融合蛋白E100被配置为使得RNA酶1(R39D、N67D、N88A、G89D、R19D、H119N、K41R)位于SpyCas9的C末端。参见图6。

载体

[0285] 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。

[0286] 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。

[0287] 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序列杂交或结合的RNA序列(或间隔子序列);和在其3'端,能够与所述融合蛋白的CRISPR/Cas蛋白结合或缔合的RNA序列(或支架序列);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在单一载体中。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。

[0288] 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。

[0289] 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)。

[0290] 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(lipoplex)、聚合物囊泡、聚合物/DNA复合物(polyplex)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。

[0291] 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,启动子对表达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。

[0292] 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表达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。

[0293] 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、白颌白眉猴(sooty mangabey)猿猴免疫缺陷病毒(SIV_{SM})载体、修饰的白颌白眉猴猿猴免疫缺陷病毒(SIV_{SM})载体、非洲绿猴猿猴免疫缺陷病毒(SIV_{AGM})载体、修饰的非洲绿猴猿猴免疫缺陷病毒(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

[0294] 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。

[0295] 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. Freeman and Co., NY）。

[0296] “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PCR反应的起始或核酶对多核苷酸的酶促切割）中的步骤。

[0297] 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。

[0298] “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

[0299] 在本公开文本的组合物和方法的一些实施方案中，本公开文本的细胞是原核细胞。

[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。

[0301] 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。

[0302] 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）。

[0303] 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。

[0304] 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。

[0305] 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。

[0306] 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。

[0307] 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。

[0308] 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。

[0309] 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。

[0310] 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。

使用方法

[0311] 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。

[0312] 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。

[0313] 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。

[0314] 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或单一指导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。

[0315] 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。

[0316] 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。

[0317] 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。

[0318] 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或单一指导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。

[0319] 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。

[0320] 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。

[0321] 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。

[0322] 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)。

[0323] 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(DeVick'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)。

[0324] 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。

[0325] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于代谢性疾病或障碍。

[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 (Alzheimer's disease) 和衰老。

[0327] 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。

[0328] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于儿科或发育性疾病或障碍。

[0329] 在本公开文本的组合物和方法的一些实施方案中,本公开文本的疾病或障碍包括但不限于心血管疾病或障碍。

[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(脑癌)、基底细胞癌、胆管癌、膀胱癌、骨癌、尤因肉瘤(Ewing Sarcoma)、骨肉瘤、恶化纤维组织细胞瘤、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、鼻腔和鼻旁窦癌(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。

[0331] 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。

[0332] 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)。

[0333] 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。

[0334] 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。

[0335] 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。

[0336] 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。

[0337] 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。

[0338] 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。

[0339] 在本公开文本的方法的一些实施方案中,治疗有效量预防疾病或障碍的发作。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。

[0340] 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。

[0341] 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。

[0342] 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。

示例实施方案:

[0343] 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;

或者

一种组合物,其包含编码融合蛋白的核酸序列,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结合多肽不是指导的RNA结合多肽,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. 根据实施方案1所述的组合物,其中所述靶序列包含至少一个重复的序列。

实施方案3. 根据实施方案1或2所述的组合物,其中包含所述gRNA的序列包含能够在真核细胞中表达所述gRNA的启动子。

实施方案4. 根据实施方案3所述的组合物,其中所述真核细胞是动物细胞。

实施方案5. 根据实施方案4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6. 根据实施方案5所述的组合物,其中所述动物细胞是人细胞。

实施方案7. 根据实施方案1-6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8. 根据实施方案1-7中任一项所述的组合物,其中所述启动子是从能够驱动RNA聚合酶的表达的启动子分离或衍生的。

实施方案9. 根据实施方案8所述的组合物,其中所述启动子是从U6启动子分离或衍生的。

实施方案10. 根据实施方案1-7中任一项所述的组合物,其中所述启动子是从能够驱动转移RNA (tRNA) 的表达的启动子分离或衍生的。

实施方案11. 根据实施方案10所述的组合物,其中所述启动子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。

实施方案12. 根据实施方案10所述的组合物, 其中所述启动子是从缬氨酸tRNA启动子分离或衍生的。

实施方案13. 根据实施方案1-12中任一项所述的组合物, 其中包含所述gRNA的序列包含与所述靶RNA序列特异性结合的间隔子序列。

实施方案14. 根据实施方案13所述的组合物, 其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。

实施方案15. 根据实施方案13所述的组合物, 其中所述间隔子序列与所述靶RNA序列具有100%互补性。

实施方案16. 根据实施方案13-15中任一项所述的组合物, 其中所述间隔子序列包含20个核苷酸或由其组成。

实施方案17. 根据实施方案13-15中任一项所述的组合物, 其中所述间隔子序列包含21个核苷酸或由其组成。

实施方案18. 根据实施方案17所述的组合物, 其中所述间隔子序列包含序列 UGGAGCGAGCAUCCCCAAA (SEQ ID NO:1)、GUUUGGGGAUGCUCGCUCCA (SEQ ID NO:2)、CCCUCACUGCUGGGGAGUCC (SEQ ID NO:3)、GGACUCCCCAGCAGUGAGGG (SEQ ID NO:4)、GCAACUGGAUCAUUUGCUG (SEQ ID NO:5)、GCAGCAAUUGAUCCAGUUGC (SEQ ID NO:6)、GCAUUCUUAUCUGGUCAGUGC (SEQ ID NO:7)、GCACUGACCAGAUAAAGAAUG (SEQ ID NO:8)、GAGCAGCAGCAGCAGCAGCAG (SEQ ID NO:9)、GCAGGCAGGCAGGCAGGCAGG (SEQ ID NO:10)、GCCCCGCCCCGCCCCGGC (SEQ ID NO:11)、或GCTGCTGCTGCTGCTGCTGC (SEQ ID NO:12)、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)、CGGGGCCGGGGCCGGGGCCG (SEQ ID NO:79)。

实施方案19. 根据实施方案1-18中任一项所述的组合物, 其中包含所述gRNA的序列包含与所述第一RNA结合蛋白特异性结合的支架序列。

实施方案20. 根据实施方案19所述的组合物, 其中所述支架序列包含茎环结构。

实施方案21. 根据实施方案19或20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22. 根据实施方案19或20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23. 根据实施方案22所述的组合物, 其中所述支架序列包含序列 GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA GUGGCACCGAGUCGGUGCUUUUUU (SEQ ID NO:13)。

实施方案24. 根据实施方案16所述的组合物, 其中所述间隔子序列包含序列 GUGAUAAGUGGAAUGCCAUG (SEQ ID NO:14)、CUGGUGAACUCCGAUAGUG (SEQ ID NO:15) 或 GAGATATAGCCTGGTGGTTC (SEQ ID NO:16)。

实施方案25. 根据实施方案19或24所述的组合物, 其中所述支架序列包含茎环结构。

实施方案26. 根据实施方案25所述的组合物, 其中所述支架序列包含85个核苷酸或由其组成。

实施方案27. 根据实施方案26所述的组合物, 其中所述支架序列包含序列
GGACAGCAUAGCAAGUUAUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
UUUUU (SEQ ID NO:17)。

实施方案28. 根据实施方案16所述的组合物, 其中所述间隔子序列包含以下序列:
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。

实施方案29. 根据实施方案28所述的组合物, 其中包含所述gRNA的序列包含与所述第一RNA结合蛋白特异性结合的支架序列。

实施方案30. 根据实施方案29所述的组合物, 其中所述支架序列包含茎环结构。

实施方案31. 根据实施方案29或30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32. 根据实施方案30或31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33. 根据实施方案32所述的组合物, 其中所述支架序列包含序列
GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAA
GUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:82) 或GUUUUAGAGCUAGAAAUAGCAAGUUAUAAAAUAAGGC
UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU (SEQ ID NO:83)。

实施方案34. 根据实施方案1-33中任一项所述的组合物, 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35. 根据实施方案34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36. 根据实施方案1-35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37. 根据实施方案36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38. 根据实施方案37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39. 根据实施方案36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40. 根据实施方案39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41. 根据实施方案36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42. 根据实施方案41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43.根据实施方案36-42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44.根据实施方案43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45.根据实施方案43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46.根据实施方案36-45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47.根据实施方案46所述的组合物,其中所述CRISPR-Cas蛋白包含突变。

实施方案48.根据实施方案47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49.根据实施方案47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50.根据实施方案47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51.根据实施方案47-50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52.根据实施方案47-50中任一项所述的组合物,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53.根据实施方案1-35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54.根据实施方案53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55.根据实施方案1-54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56.根据实施方案55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57.根据实施方案55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58.根据实施方案1-57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59.根据实施方案58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60.根据实施方案58或59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61.根据实施方案60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62.根据实施方案1-61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63.根据实施方案1-62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核定位信号(NLS)的序列。

实施方案64.根据实施方案63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65.根据实施方案64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66.根据实施方案1-62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67.根据实施方案66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68.根据实施方案67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69.根据实施方案1-68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70.根据实施方案69所述的组合物,其中编码所述第二RNA结合蛋白的序列包含RNA酶或由其组成。

实施方案71.根据实施方案70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72.根据实施方案71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73.根据实施方案72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74.根据实施方案73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75.根据实施方案70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76.根据实施方案75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77.根据实施方案70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78.根据实施方案77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79.根据实施方案70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80.根据实施方案79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81.根据实施方案70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82.根据实施方案81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83.根据实施方案70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84.根据实施方案83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85.根据实施方案70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86.根据实施方案85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87.根据实施方案70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88.根据实施方案87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89.根据实施方案70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90.根据实施方案89所述的组合物,其中所述RNA酶11包含SEQ ID NO:29或由其组成。

实施方案91.根据实施方案70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92.根据实施方案91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93.根据实施方案69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94.根据实施方案93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95.根据实施方案69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96.根据实施方案95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 (ENDOV) 或由其组成。

实施方案97.根据实施方案96所述的组合物,其中所述ENDOV蛋白包含SEQ ID NO:32或由其组成。

实施方案98.根据实施方案95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99.根据实施方案98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100.根据实施方案95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 (ENDOD1) 或由其组成。

实施方案101.根据实施方案100所述的组合物,其中所述ENDOD1蛋白包含SEQ ID

NO:34或由其组成。

实施方案102.根据实施方案95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)或由其组成。

实施方案103.根据实施方案102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104.根据实施方案69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105.根据实施方案104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106.根据实施方案69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107.根据实施方案106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108.根据实施方案69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109.根据实施方案108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110.根据实施方案108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111.根据实施方案69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112.根据实施方案111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113.根据实施方案69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114.根据实施方案113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115.根据实施方案69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116.根据实施方案115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117.根据实施方案115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118.根据实施方案69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119.根据实施方案118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120.根据实施方案69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121.根据实施方案120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122.根据实施方案69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123.根据实施方案122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124.根据实施方案69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125.根据实施方案124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126.根据实施方案69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127.根据实施方案126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128.根据实施方案69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129.根据实施方案128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130.根据实施方案69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131.根据实施方案130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132.根据实施方案69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133.根据实施方案132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134.根据实施方案69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135.根据实施方案134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136.根据实施方案69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137.根据实施方案136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138.根据实施方案69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139.根据实施方案138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140.根据实施方案69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141.根据实施方案140所述的组合物,其中所述CPSF4L包含SEQ ID NO:134或由其组成。

实施方案142.根据实施方案69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143.根据实施方案142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144.根据实施方案142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145.根据实施方案69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146.根据实施方案145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147.根据实施方案69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148.根据实施方案147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149.根据实施方案69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150.根据实施方案149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151.根据实施方案69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152.根据实施方案151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153.根据实施方案69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154.根据实施方案153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155.根据实施方案69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156.根据实施方案155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157.根据实施方案69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158.根据实施方案157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159.根据实施方案69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160.根据实施方案159所述的组合物,其中所述ERCC4多肽包含SEQ ID NO:64或由其组成。

实施方案161.根据实施方案69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162.根据实施方案161所述的组合物,其中所述RNA酶1(RNA酶1(K41R))多肽包含SEQ ID NO:116或由其组成。

实施方案163.根据实施方案69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164.根据实施方案163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165.根据实施方案69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166.根据实施方案165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167.根据实施方案69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168.根据实施方案167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169.根据实施方案69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170.根据实施方案169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171.根据实施方案69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172.根据实施方案171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173.根据实施方案69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174.根据实施方案173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175.根据实施方案69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176.根据实施方案175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177.根据实施方案69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178.根据实施方案177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179.一种组合物,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结合多肽结合并非由gRNA序列指导的靶RNA,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180.根据实施方案179所述的组合物,其中所述第一RNA结合多肽或其部分是PUF、PUMBY或PPR多肽或其部分。

实施方案181.一种用于修饰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融合蛋白或其部分与所述RNA分子结合的条件下使根据实施方案1或179所述的组合物与所述RNA分子接触。

实施例

实施例1:方法

[0344] 将HEK-293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/孔接种于96孔板中用于萤光素酶测定。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。以与报告基因质粒中的CTG重复相邻的序列中的引物使用以下引物来进行qPCR:

正向引物	TetCTG_DMPK_E15_F	TCGGAGCGGTTGTGAACT	SEQ ID NO:83
反向引物	TetCTG_DMPK_E15_R	GTTCCCGTTGTTCTGTC	SEQ ID NO:84

[0345] 通过针对GAPDH进行归一化来确定CTG重复报告基因的相对丰度。接下来,将靶向CTG的sgRNA的水平针对非靶向sgRNA进行归一化,以产生在相关数据包中报告的最终值。

靶向CTG 的间隔子	AGCAGCAGCAGCAGCAGCAG	SEQ ID NO: 85
非靶向对 照间隔子 (λ2)	GTGATAAGTGGAATGCCATG	SEQ ID NO: 86
sgRNA 支 架(N指示 间隔子)	GNNNNNNNNNNNNNNNNNNNGUUUAAGAGCUAUGCUG GAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU	SEQ ID NO: 87

[0346] 用Promega双重萤光素酶试剂盒根据制造商的说明进行萤光素酶测定。所报告的值是萤火虫与海肾萤光素酶发光读数的比率。

实施例2:重复性RNA分子和mRNA分子的RNA指导的切割

[0347] 实验设计:构建具有注释的RNA内切核酸酶活性的人蛋白与Cas9 (酿脓链球菌或空肠弯曲杆菌) 的各种融合物。将编码上文融合物的质粒与含有重复序列的质粒或萤光素酶测定质粒 (包含编码萤光素酶蛋白的mRNA序列) 共转染。在将RNA内切核酸酶/Cas9融合物与重复性RNA共转染的条件下用qPCR测量含有CTG重复的RNA的水平。在将RNA内切核酸酶/Cas9融合物与萤光素酶测定质粒共转染的条件下使用发光测定测量萤光素酶蛋白的水平。

将所有测量值针对非靶向sgRNA对照构建体进行归一化(图3A-图5和图9)。

实施例3:病毒RNA分子的RNA指导的切割

[0348] 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1中。此项研究中使用的构建体的序列呈现于下文中。

[0349] 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2中所列的以下引物进行qPCR。

[0350] 图7显示了在E43和E67内切核酸酶二者的存在下用含有如表2中指示的各种靶向NS5的间隔子序列的sgRNA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。

[0351] 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8A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8A和图8B)。图6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。

[0352] 表1: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	TCTCTCTCCCATCCAGTGA (SEQ ID NO:195)

[0353] 表2: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)	GTGATAAGTGAATGCCATG (SEQ ID NO: 200)
sgRNA支架 (N指示间隔子)	GNNNNNNNNNNNNNNNNNNNNNGUUUAAGAGCU AUGCUGGAAACAGCAUAGCAAGUUUAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGCUUUUUUU (SEQ ID NO: 201)

[0354] 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
ccttcatatttgcataacgatacaaggctgtagagagataattagaattaatttgactgtaaacacaaagatat
tagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatgga
ctatcatatgcttaccgtaacttgaaagtatttgcatttcttggctttatatatcttGTGGAAGGACGAAACACC
NNNNNNNNNNNNNNNNNGTTTTAGTCCCTGAAGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGGTTACAA
TCCCTAAAACCGCTTTTTTTCCTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTC
CAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAGTGGGAATTGGCTC
CGGTGCCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGATCC
GGTGCCTAGAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCCTTTTCCCGAGGGTG
GGGGAGAACCGTATATAAGTGCAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTGCCGCCAGAACACAGGA
CCGTTCTAGAGCGCTATTTAGAACCatgTGTCTCCCCAAGAATCTGGCATGACCGCTCTTCAGCGAGGATGTT
GACGCGAAGCAGATCCCTGGGACCTGGGGCCGGGCCACGAGGGTGTCGGGAAGAACCAGGACCGTTGCGACGGAGG
GAAGCAGCAGCGGAAGCTCGGAAATCCCATTCTCCGTTAAACGACCCCGCAAGGCACAACGGCTCAGGGTTGCTT
ACGAGGGGAGCGATTCCGAAAAGGGTGAAGGAGCAGAGCCCTTGAAGGTTCCAGTATGGGAACCCAGGATTGGCA
GCAGCAGCTTGTAACATCCGAGCAATGAGGAACAAAAAGATGCACCTGTTGATCACCTCGGAACCGAACATTGT
TATGATTCTAGTGCGCCGCCAAAAGTCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGG
ACCAGGTTACGGCCGGAGCAATGCAACGGCTTCGGGCACGGGGACTCACGGTCGATAGCATTTCGAGACCGATGA
CGCAACATTGGGTAACTCATATATCCAGTTGGCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCC
ATTCTCCAACAACATTACGGAGGTGATATACCCGCAAGCGTAGCTGAACTGGTAGCACTGCCGGGCGTCGGTCCCA
AAATGGCACATCTGGCTATGGCGGTTGCTTGGGGAACGGTGTCTGGTATCGCAGTTGATACGCATGTCCACCGCAT
CGCCAATCGGCTGAGGTGGACTAAAAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGAGTGGCTT
CCACGAGAGCTGTGGCATGAAATCAATGGATTGCTGGTTGGTTTCGGGCAGCAGACATGCTTGCCCGTGACCCCC
GGTGTCTATGCTTGCTTGAACCAGGCTTTGTGCCAGCTGCCAGGGCTGAGTGGAAGTGAGACACCGGGAACATC
TGAGTCTGCGACCCCGGAGAGCacaaacGCGCGAATCCTGGCCTTCGcgATTGGCATTAGCAGCATCGGCTGGGCA
TTCTCTGAAAACGACGAACTGAAGGATTGCGGCGTGCGAATTTTCACTAAGGTCGAAAATCCCAAACTGGTGAAT
CACTCGCTCTCCCTAGACGACTGGCACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCATCT
TAAACACCTTATTGCAAATGAGTTTAACTGAATTATGAGGACTACCAATCCTTTGACGAGTCTCTTGCTAAAGCC
TACAAAGGGAGCCTTATATCCCCGTATGAGCTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTTG
CTCGCGTGATTCTCCACATAGCGAAGAGGCGAGGATACGATGACATTAACAAACAGTGATGATAAGGAAAAAGGGGC
CATACTCAAAGCGATTAAGCAAAATGAAGAGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAG
TACTTCCAGAAGTTCAAAGAAAATAGCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGAAAGATGTA
TTGCGCAATCTTCTCAAGGACGAGCTCAAATTGATTTTCAAGAAACAAAGGGAATTTGGGTTTACGCTTCTCAA
AAAATTTGAGGAAGAGGTTCTGAGCGTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAAT
TGTAGTTTCTTACCGATGAAAAACGGGCGCCAAAAATAGCCCTTTGGCTTTTATGTTTGTGCTCTGACTCGCA
TCATTAATCTGCTCAACAACCTTAAAAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAA
CGAAGTTTGAAGAACGGGACTTTGACCTACAAACAAACCAAAAAGCTTCTTGGTCTCAGTGATGACTACGAATTC
AAGGGAGAAAAAGGGACATATTTATCGAATTCAAGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACT

TGTCTCAAGATGATCTCAACGAAATTGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGCGTT
GGCGAAGTATGACCTTAACCAAAACCAAATAGATAGCCTCAGCAAGTTGGAATTTAAAGATCACTTGAATATAAGT
TTCAAGGCCCTTAAGTTGGTCACCCCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATC
TCAAGGTTGCTATTAACGAAGACAAAAAGATTTCTCCAGCTTTCAATGAGACTTACTATAAGGACGAGGTTAC
CAATCCTGTGGTGCTCCGAGCCATCAAAGAGTATCGAAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTA
CACAAAATAAATATTGAGCTCGCAAGGGAGGTCGGTAAGAACCACTCCAGCGCGCCAAAATAGAAAAGGAACAGA
ATGAAAATTACAAAGCGAAAAAGGACGCCGAGCTCGAGTGCAGAAAGCTGGGCTGAAAATAAACAGCAAGAACAT
TCTCAAACCTCCGCTCTTCAAAGAACAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAATCTCCGATCTT
CAAGACGAGAAGATGCTCGAAATAGACgcgATATATCCATATAGCAGGTCTTTTGACGATTCTTACATGAATAAAG
TGCTTGTTTTCTACTAAGCAGAATCAGGAAAAGTTGAATCAGACCCCTTTGAGGCCTTTGGCAACGACTCAGCAAA
GTGGCAGAAGATCGAGGTCTTGGCTAAGAATCTTCTACTAAGAAACAGAAAAGGATATTGGATAAGAACTATAAA
GACAAAGAACAAAAGAACTTTAAAGACCGCAACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACA
CAAAAGATTATTTGGACTTCTTGCCGCTGTCTGATGATGAGAACACGAAACTCAACGACACGCAAAAGGGGTCTAA
AGTCCACGTGCAAGCTAAATCTGGGATGCTCACCTCAGCATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGA
AACATCACCTGCACCATGCCATTGACGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGTTTAGCG
ACTTCAAGAAGGAACAAGAGTCCAACAGCGCCGAGCTCTACGCAAAAAAGATTAGTGAACTCGACTACAAAAACAA
AAGAAAATTCTTTGAGCCGTTTCAGCGGATTTTCGACAGAAGGTATTGGATAAAAATAGATGAAATTTTCGTGAGCAAA
CCCGAAAGGAAAAAGCCCTCAGGCGCCTTGACGAAGAGACTTTTCAGGAAGGAAGAGGAATTCTACCAAAGCTACG
GCGGAAAAGAGGGAGTTTGAAGGCTCTCGAAGTTGGAAGATTAGGAAGGTGAACGGCAAGATAGTAAAAACGG
CGATATGTTCCGGGTGATATCTTCAAACATAAAAAACGAATAAATTTTATGCTGTGCCTATATACACTATGGAC
TTCGCACTTAAGGTCCTGCCGAATAAGGCGGTAGCCCGATCTAAAAAGGCGAAATTAAGGACTGGATTTTGATGG
ATGAAAATTACGAGTTCTGCTTTTCTCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACC
GGAATTTCGTGTATTACAACGCTTTTACTTCCTCTACGGTATCTTTGATTGTCTCCAAACATGACAACAAATTCGAA
ACACTCAGTAAAAACCAAAGATTCTCTTTAAAAATGCGAACGAGAAAGAAGTAATTGCAAAATCAATTGGCATCC
AAAATTTGAAAGTTTTTGAAGAAATATATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACAGCGAGA
GGACTTCAAAAAATCAGGTCCACCCAAGAAAAACGCAAGGTGGAAGATCCGAAGAAAAAGCGAAAAGTGATGTG
taaCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCAAC
TTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAATTTACAAATAAAGCATTTTTTTTCAC
TGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGTATACCG (SEQ ID NO:202)。

[0355] 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) :

gtttattacagggacagcagagatccagtttggttaattaaggtaccgagggcctatttcccatgatt
ccttcataatttgcataatacgatacaaggtggttagagagataattagaattaatttgactgtaaacacaaagatat
tagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatgga
ctatcatatgcttaccgtaacttgaaagtatttctgatttcttggctttatatatcttGTGGAAAGGACGAAACACC
NNNNNNNNNNNNNNNNNNNGTTTAGTCCCTGAAGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGTTACAA
TCCCTAAAACCGCTTTTTTCTGCAGCCCGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTC
CAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGAATTCGCTAGCTAGGTCTTGAAAGGAGTGGAATTGGCTC
CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCGGCAATTGATCC

GGTGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTTTTCCCGAGGGTG
GGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGCCGCCAGAACACAGGA
CCGGTTCTAGAGCGCTATTTAGAACCatgCAGGAGGTAATAGCGGGGCTTGAGCGATTTACCTTTGCCTTCGAAAA
AGACGTAGAGATGCAGAAGGGAACCGGCCTGCTCCCATTTCAGGTATGGACAAATCAGCATCTGCCGTGTGCAAT
TTTTTCACCAAGGGTCTGTGTGAAAAGGGGAAGCTCTGTCCATTTCCGCATGATCGCGGAGAGAAGATGGTGGTGT
GTAAGCACTGGCTGAGAGGGCTTTGCAAAAAAGGCGACCACTGCAAATTTCTTCACCAATATGACCTGACTCGAAT
GCCTGAGTGTTATTTTTACAGTAAGTTCGGTGACTGTAGCAACAAAGAATGCAGCTTCTTGCATGTCAAACCAGCA
TTCAAGTCACAGGATTGCCCCGTGGTACGATCAGGGTTTTTGCAAGGACGGTCCCCTCTGCAAATATCGACACGTAC
CCAGAATTATGTGCCTTAATTACCTGGTCGGCTTCTGTCTGAAGGGCCAAAAATGTCAGTTTGCTCAAAAAATTCTG
CGAGTTCAAATTGCTCCCTGGGTCTAAAATTTGGGAACCCAGGATTGGCAGCAGCAGCTTGTAACATCCGAGCA
ATGAGGAACAAAAAGATGCACCTGTTGATCACCTCGGAACCGAACATTGTTATGATTCTAGTGCGCCGCCAAAAAG
TCCGCCGGTATCAGGTTCTGTTGAGTTTGATGCTGAGTAGTCAGACTAAGGACCAGGTTACGGCCGGAGCAATGCA
ACGGCTTCGGGCACGGGGACTCACGGTCGATAGCATTTTGCAGACCGATGACGCAACATTGGGTAACTCATATAT
CCAGTTGGCTTCTGGCGGAGCAAAGTGAAGTACATCAAGCAGACCTCAGCCATTCTCCAACAACATTACGGAGGTG
ATATACCCGCAAGCGTAGCTGAAGTGGTAGCTGCCGGGCGTCGGTCCCAAAATGGCACATCTGGCTATGGCGGT
TGCTTGGGGAACGGTGTCTGGTATCGCAGTTGATACGCATGTCCACCGCATCGCCAATCGGCTGAGGTGGACTAAA
AAAGCCACTAAGTCTCCTGAAGAAACACGGGCTGCTCTGGAAGAGTGGCTTCCACGAGAGCTGTGGCATGAAATCA
ATGGATTGCTGGTTGGTTTCGGGCAGCAGACATGCTTGCCCGTGACCCCCGGTGTCTGCTTGAACCAGGC
TTTGTGCCAGCTGCCAGGGCCTGAGTGGAAGTGAGACACCGGGAACATCTGAGTCTGCGACCCCGGAGAGCaca
aacGCGGAATCCTGGCCTTCGcgATTGGCATTAGCAGCATCGGCTGGGCATTCTCTGAAAACGACGAACTGAAGG
ATTGCGGCGTGCGAATTTTCACTAAGGTGCAAAATCCCAAACTGGTGAATCACTCGCTCTCCCTAGACGACTGGC
ACGCTCCGCACGAAAGAGGCTTGCCCGCCGCAAGGCACGCTTGAACCATCTTAAACACCTTATTGCAAATGAGTTT
AAACTGAATTATGAGGACTACCAATCCTTTGACGAGTCTCTTGCTAAAGCCTACAAAGGGAGCCTTATATCCCCGT
ATGAGCTCCGGTTCAGAGCACTCAACGAACTGCTGTCCAAACAGGATTTTGCTCGCGTGATTCTCCACATAGCGAA
GAGGCGAGGATACGATGACATTAACAAACAGTGATGATAAGGAAAAAGGGGCCATACTCAAAGCGATTAAAGCAAAAT
GAAGAGAAGCTCGCTAACTATCAATCAGTAGGGGAGTATCTCTATAAAGAGTACTTCCAGAAGTTCAAAGAAAATA
GCAAGGAATTTACTAATGTCCGGAATAAAAAGGAGTCTTACGAAAGATGTATTGCGCAATCTTTCCTCAAGGACGA
GCTCAAATTGATTTTCAAGAAACAAAGGGAATTTGGGTTTCAGCTTCTCAAAAAAATTTGAGGAAGAGGTTCTGAGC
GTTGCCTTTTACAAACGCGCCCTTAAGGACTTCTCACATCTCGTAGGGAATTGTAGTTTCTTACCGATGAAAAAC
GGGCGCAAAAAATAGCCCTTTGGCTTTTATGTTTGTGCTCTGACTCGCATCATTAATCTGCTCAACAACCTTAA
AAACACGGAAGGGATTCTGTACACAAAGGATGATCTGAACGCTCTGCTTAACGAAGTTTGAAGAACGGGACTTTG
ACCTACAAACAAACCAAAAAAGCTTCTTGGTCTCAGTGATGACTACGAATTCAGGGAGAAAAAGGGACATATTTCA
TCGAATTCAAGAAGTATAAGGAGTTCATCAAAGCCTTGGGCGAGCACAACCTTGCTCAAGATGATCTCAACGAAAT
TGCTAAGGATATCACTCTGATTAAAGACGAGATCAAGCTCAAAAAGGCGTTGGCGAAGTATGACCTTAACCAAAAC
CAAATAGATAGCCTCAGCAAGTTGGAATTTAAAGATCACTTGAATATAAGTTTCAAGGCCCTTAAGTTGGTCACCC
CCTTGATGCTTGAAGGAAAGAAATATGATGAGGCATGTAATGAGCTGAATCTCAAGGTGCTATTAACGAAGACAA
AAAAGATTTCTCCCAGCTTCAATGAGACTTACTATAAGGACGAGGTTACCAATCCTGTGGTGCTCCGAGCCATC
AAAGAGTATCGAAAGGTCCTGAATGCTTTGCTCAAAAAATACGGTAAGGTACACAAAAATAAATATTGAGCTCGCAA
GGGAGGTGCGTAAGAACCACTCCAGCGCGCAAAATAGAAAAGGAACAGAATGAAAATTACAAAGCGAAAAAGGA

CGCCGAGCTCGAGTGCAGAAAAGCTGGGCCTGAAAATAAACAGCAAGAACATTCTCAAACCTCCGCCTCTTCAAAGAA
CAAAAAGAATTTTGTGCTTATAGTGGTGAGAAAATAAAAATCTCCGATCTTCAAGACGAGAAGATGCTCGAAATAG
ACgcgATATATCCATATAGCAGGTCTTTTGACGATTCTTACATGAATAAAGTGCTTGTTTTCACTAAGCAGAATCA
GGAAAAGTTGAATCAGACCCCCTTTGAGGCCTTTGGCAACGACTCAGCAAAGTGGCAGAAGATCGAGGTCTTGGCT
AAGAATCTTCCTACTAAGAAACAGAAAAGGATATTGGATAAAGAACTATAAAGACAAAGAACAAAAGAACTTTAAAG
ACCGCAACCTCAATGACACCAGATACATAGCAAGATTGGTTCTGAACTACACAAAAGATTATTTGGACTTCTTGCC
GCTGTCTGATGATGAGAACACGAAACTCAACGACACGCAAAAGGGGTCTAAAGTCCACGTCAAGCTAAATCTGGG
ATGCTCACCTCAGCATTGAGGCATACGTGGGGATTCTCAGCAAAGGACCGAAACAATCACCTGCACCATGCCATTG
ACGCAGTTATCATAGCGTATGCCAATAATTCAATAGTAAAAGCGTTTAGCGACTTCAAGAAGGAACAAGAGTCCAA
CAGCGCCGAGCTCTACGCAAAAAAGATTAGTGAACCTCGACTACAAAAACAAAAGAAAATTCTTTGAGCCGTTTCAGC
GGATTTTCGACAGAAGGTATTGGATAAAATAGATGAAATTTTCGTGAGCAAAACCCGAAAGGAAAAAGCCCTCAGGCG
CCTTGCACGAAGAGACTTTCAGGAAGGAAGAGGAATTCTACCAAAGCTACGGCGGAAAAGAGGGAGTTTTGAAGGC
TCTCGAACTTGAAAGATTAGGAAGGTGAACGGCAAGATAGTAAAAACGGCGATATGTTCCGGGTGATATCTTC
AAACATAAAAAACGAATAAAATTTTATGCTGTGCCTATATACACTATGGACTTCGCACTTAAGGTCCTGCCGAATA
AGGCGGTAGCCCGATCTAAAAAAGGCGAAATTAAGGACTGGATTTTGATGGATGAAAATTACGAGTTCTGCTTTTC
TCTCTACAAGGATTCCCTTATATTGATACAGACGAAAGATATGCAGGAACCGGAATTCGTGTATTACAACGCTTTT
ACTTCCTCTACGGTATCTTTGATTGTCTCCAAACATGACAACAAATTCGAAACACTCAGTAAAAACAAAAGATTC
TCTTTAAAAATGCGAACGAGAAAAGTAATTGCAAAATCAATTGGCATCCAAAATTTGAAAGTTTTTGAAAAATA
TATAGTATCTGCCCTCGGAGAGGTTACTAAAGCGGAATTTAGACAGCGAGAGGACTTCAAAAAATCAGGTCCACCC
AAGAAAAACGCAAGGTGGAAGATCCGAAGAAAAAGCGAAAAGTGGATGTGtaaCGTTTTCCGGGACGCCGGCTGG
ATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTT
ACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAA
ACTCATCAATGTATCTTATCATGTCTGTATACCG (SEQ ID NO:203)。

通过引用并入

[0356] 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具体化的任何发明的现有技术,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

[0357] 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所有此类变化和修改。

序列表

<110> 洛卡纳生物股份有限公司

<120> 靶向RNA的融合蛋白组合物和使用方法

<130> LOCN-002/001W0 330675-2007

<150> US 62/682,271

<151> 2018-06-08

<160> 208

<170> PatentIn3.5版

<210> 1

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 1

uggagcgagc auccccaaa 20

<210> 2

<211> 21

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 2

guuuggggga ugcucgcucc a 21

<210> 3

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 3

cccucacugc uggggagucc 20

<210> 4

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 4
ggacucccca gcagugagg 20
<210> 5
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 5
gcaacuggau caauuugcug 20
<210> 6
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 6
gcagcaaaau gauccaguug c 21
<210> 7
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 7
gcauucuuau cuggucagug c 21
<210> 8
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 8
gcacugacca gauaagaaug 20
<210> 9
<211> 21
<212> RNA
<213> 人工序列
<220>

<223> 间隔序列
<400> 9
gagcagcagc agcagcagca g 21
<210> 10
<211> 21
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 10
gcaggcaggc aggcaggcag g 21
<210> 11
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 11
gccccggccc cggccccggc 20
<210> 12
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 12
gctgctgctg ctgctgctgc 20
<210> 13
<211> 93
<212> RNA
<213> 人工序列
<220>
<223> 支架序列
<400> 13
guuuuagagc uaugcuggaa acagcauagc aaguuuuuuu aaggcuaguc cguuaucaac 60
uugaaaaagu ggcaccgagu cggugcuuuu uuu 93
<210> 14
<211> 20
<212> RNA

<213> 人工序列
<220>
<223> 间隔序列
<400> 14
gugauaagug gaaugccaug 20
<210> 15
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 15
cuggugaacu uccgauagug 20
<210> 16
<211> 20
<212> DNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 16
gagatatagc ctggtggttc 20
<210> 17
<211> 73
<212> RNA
<213> 人工序列
<220>
<223> 支架序列
<400> 17
ggacagcaua gcaaguuaaa auaaggcuag uccguuauca acuugaaaaa guggcaccga 60
gucggugcuu uuu 73
<210> 18
<211> 3
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 18
cug 3
<210> 19

<211> 4

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 19

ccug 4

<210> 20

<211> 128

<212> PRT

<213> 未知的

<220>

<223> RNA酶1

<400> 20

Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp

1 5 10 15

Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg

20 25 30

Arg Asn Met Thr Gln Gly Leu Cys Lys Pro Val Asn Thr Phe Val His

35 40 45

Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr

50 55 60

Cys Lys Asn Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His

65 70 75 80

Ile Thr Asp Cys Arg Leu Thr Asn Gly Ser Arg Tyr Pro Asn Cys Ala

85 90 95

Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly

100 105 110

Ser Pro Tyr Val Pro Val His Phe Asp Ala Ser Val Glu Asp Ser Thr

115 120 125

<210> 21

<211> 119

<212> PRT

<213> 未知的

<220>

<223> RNA酶4

<400> 21

Gln Asp Gly Met Tyr Gln Arg Phe Leu Arg Gln His Val His Pro Glu

1 5 10 15

Glu Thr Gly Gly Ser Asp Arg Tyr Cys Asp Leu Met Met Gln Arg Arg
 20 25 30
 Lys Met Thr Leu Tyr His Cys Lys Arg Phe Asn Thr Phe Ile His Glu
 35 40 45
 Asp Ile Trp Asn Ile Arg Ser Ile Cys Ser Thr Thr Asn Ile Gln Cys
 50 55 60
 Lys Asn Gly Lys Met Asn Cys His Glu Gly Val Val Lys Val Thr Asp
 65 70 75 80
 Cys Arg Asp Thr Gly Ser Ser Arg Ala Pro Asn Cys Arg Tyr Arg Ala
 85 90 95
 Ile Ala Ser Thr Arg Arg Val Val Ile Ala Cys Glu Gly Asn Pro Gln
 100 105 110
 Val Pro Val His Phe Asp Gly
 115

<210> 22

<211> 127

<212> PRT

<213> 未知的

<220>

<223> RNA酶6

<400> 22

Trp Pro Lys Arg Leu Thr Lys Ala His Trp Phe Glu Ile Gln His Ile
 1 5 10 15
 Gln Pro Ser Pro Leu Gln Cys Asn Arg Ala Met Ser Gly Ile Asn Asn
 20 25 30
 Tyr Thr Gln His Cys Lys His Gln Asn Thr Phe Leu His Asp Ser Phe
 35 40 45
 Gln Asn Val Ala Ala Val Cys Asp Leu Leu Ser Ile Val Cys Lys Asn
 50 55 60
 Arg Arg His Asn Cys His Gln Ser Ser Lys Pro Val Asn Met Thr Asp
 65 70 75 80
 Cys Arg Leu Thr Ser Gly Lys Tyr Pro Gln Cys Arg Tyr Ser Ala Ala
 85 90 95
 Ala Gln Tyr Lys Phe Phe Ile Val Ala Cys Asp Pro Pro Gln Lys Ser
 100 105 110
 Asp Pro Pro Tyr Lys Leu Val Pro Val His Leu Asp Ser Ile Leu
 115 120 125

<210> 23

<211> 155

<212> PRT

<213> 未知的

<220>

<223> RNA酶7

<400> 23

Ala	Pro	Ala	Arg	Ala	Gly	Phe	Cys	Pro	Leu	Leu	Leu	Leu	Leu	Leu	Leu
1				5					10					15	
Gly	Leu	Trp	Val	Ala	Glu	Ile	Pro	Val	Ser	Ala	Lys	Pro	Lys	Gly	Met
			20					25					30		
Thr	Ser	Ser	Gln	Trp	Phe	Lys	Ile	Gln	His	Met	Gln	Pro	Ser	Pro	Gln
			35				40				45				
Ala	Cys	Asn	Ser	Ala	Met	Lys	Asn	Ile	Asn	Lys	His	Thr	Lys	Arg	Cys
			50				55				60				
Lys	Asp	Leu	Asn	Thr	Phe	Leu	His	Glu	Pro	Phe	Ser	Ser	Val	Ala	Ala
65					70				75					80	
Thr	Cys	Gln	Thr	Pro	Lys	Ile	Ala	Cys	Lys	Asn	Gly	Asp	Lys	Asn	Cys
				85				90					95		
His	Gln	Ser	His	Gly	Pro	Val	Ser	Leu	Thr	Met	Cys	Lys	Leu	Thr	Ser
			100					105					110		
Gly	Lys	Tyr	Pro	Asn	Cys	Arg	Tyr	Lys	Glu	Lys	Arg	Gln	Asn	Lys	Ser
			115					120				125			
Tyr	Val	Val	Ala	Cys	Lys	Pro	Pro	Gln	Lys	Lys	Asp	Ser	Gln	Gln	Phe
			130				135				140				
His	Leu	Val	Pro	Val	His	Leu	Asp	Arg	Val	Leu					
145					150					155					

<210> 24

<211> 122

<212> PRT

<213> 未知的

<220>

<223> RNA酶8

<400> 24

Thr	Ser	Ser	Gln	Trp	Phe	Lys	Thr	Gln	His	Val	Gln	Pro	Ser	Pro	Gln
1				5				10						15	
Ala	Cys	Asn	Ser	Ala	Met	Ser	Ile	Ile	Asn	Lys	Tyr	Thr	Glu	Arg	Cys
			20					25					30		
Lys	Asp	Leu	Asn	Thr	Phe	Leu	His	Glu	Pro	Phe	Ser	Ser	Val	Ala	Ile
			35				40				45				
Thr	Cys	Gln	Thr	Pro	Asn	Ile	Ala	Cys	Lys	Asn	Ser	Cys	Lys	Asn	Cys

50	55	60			
His Gln Ser His Gly Pro Met Ser Leu Thr Met Gly Glu Leu Thr Ser					
65	70	75	80		
Gly Lys Tyr Pro Asn Cys Arg Tyr Lys Glu Lys His Leu Asn Thr Pro					
	85	90	95		
Tyr Ile Val Ala Cys Asp Pro Pro Gln Gln Gly Asp Pro Gly Tyr Pro					
	100	105	110		
Leu Val Pro Val His Leu Asp Lys Val Val					
115	120				

<210> 25

<211> 134

<212> PRT

<213> 未知的

<220>

<223> RNA酶2

<400> 25

Lys Pro Pro Gln Phe Thr Trp Ala Gln Trp Phe Glu Thr Gln His Ile					
1	5	10	15		
Asn Met Thr Ser Gln Gln Cys Thr Asn Ala Met Gln Val Ile Asn Asn					
	20	25	30		
Tyr Gln Arg Arg Cys Lys Asn Gln Asn Thr Phe Leu Leu Thr Thr Phe					
	35	40	45		
Ala Asn Val Val Asn Val Cys Gly Asn Pro Asn Met Thr Cys Pro Ser					
50	55	60			
Asn Lys Thr Arg Lys Asn Cys His His Ser Gly Ser Gln Val Pro Leu					
65	70	75	80		
Ile His Cys Asn Leu Thr Thr Pro Ser Pro Gln Asn Ile Ser Asn Cys					
	85	90	95		
Arg Tyr Ala Gln Thr Pro Ala Asn Met Phe Tyr Ile Val Ala Cys Asp					
	100	105	110		
Asn Arg Asp Gln Arg Arg Asp Pro Pro Gln Tyr Pro Val Val Pro Val					
	115	120	125		
His Leu Asp Arg Ile Ile					
130					

<210> 26

<211> 216

<212> PRT

<213> 未知的

<220>

<223> RNA酶6PL

<400> 26

```

Asp Lys Arg Leu Arg Asp Asn His Glu Trp Lys Lys Leu Ile Met Val
1           5           10           15
Gln His Trp Pro Glu Thr Val Cys Glu Lys Ile Gln Asn Asp Cys Arg
          20           25           30
Asp Pro Pro Asp Tyr Trp Thr Ile His Gly Leu Trp Pro Asp Lys Ser
          35           40           45
Glu Gly Cys Asn Arg Ser Trp Pro Phe Asn Leu Glu Glu Ile Lys Lys
          50           55           60
Asn Trp Met Glu Ile Thr Asp Ser Ser Leu Pro Ser Pro Ser Met Gly
65           70           75           80
Pro Ala Pro Pro Arg Trp Met Arg Ser Thr Pro Arg Arg Ser Thr Leu
          85           90           95
Ala Glu Ala Trp Asn Ser Thr Gly Ser Trp Thr Ser Thr Gly Gly Cys
          100          105          110
Ala Leu Pro Pro Ala Ala Leu Pro Ser Gly Asp Leu Cys Cys Arg Pro
          115          120          125
Ser Leu Thr Ala Gly Ser Arg Gly Val Gly Val Asp Leu Thr Ala Leu
          130          135          140
His Gln Leu Leu His Val His Tyr Ser Ala Thr Gly Ile Ile Pro Glu
145          150          155          160
Glu Cys Ser Glu Pro Thr Lys Pro Phe Gln Ile Ile Leu His His Asp
          165          170          175
His Thr Glu Trp Val Gln Ser Ile Gly Met Pro Ile Trp Gly Thr Ile
          180          185          190
Ser Ser Ser Glu Ser Ala Ile Gly Lys Asn Glu Glu Ser Gln Pro Ala
          195          200          205
Cys Ala Val Leu Ser His Asp Ser
          210          215

```

<210> 27

<211> 722

<212> PRT

<213> 未知的

<220>

<223> RNA酶L

<400> 27

```

Ala Ala Val Glu Asp Asn His Leu Leu Ile Lys Ala Val Gln Asn Glu
1           5           10           15

```

Asp	Val	Asp	Leu	Val	Gln	Gln	Leu	Leu	Glu	Gly	Gly	Ala	Asn	Val	Asn
20				25				30							
Phe	Gln	Glu	Glu	Glu	Gly	Gly	Trp	Thr	Pro	Leu	His	Asn	Ala	Val	Gln
35				40				45							
Met	Ser	Arg	Glu	Asp	Ile	Val	Glu	Leu	Leu	Leu	Arg	His	Gly	Ala	Asp
50				55				60							
Pro	Val	Leu	Arg	Lys	Lys	Asn	Gly	Ala	Thr	Pro	Phe	Ile	Leu	Ala	Ala
65				70				75				80			
Ile	Ala	Gly	Ser	Val	Lys	Asp	Leu	Leu	Lys	Leu	Phe	Leu	Ser	Lys	Gly
85				90				95							
Ala	Asp	Val	Asn	Glu	Cys	Asp	Phe	Tyr	Gly	Phe	Thr	Ala	Phe	Met	Glu
100				105				110							
Ala	Ala	Val	Tyr	Gly	Lys	Val	Lys	Ala	Leu	Lys	Phe	Leu	Tyr	Lys	Arg
115				120				125							
Gly	Ala	Asn	Val	Asn	Leu	Arg	Arg	Lys	Thr	Lys	Glu	Asp	Gln	Glu	Arg
130				135				140							
Leu	Arg	Lys	Gly	Gly	Ala	Thr	Ala	Leu	Met	Asp	Ala	Ala	Glu	Lys	Gly
145				150				155				160			
His	Val	Glu	Val	Leu	Lys	Ile	Leu	Leu	Asp	Glu	Met	Gly	Ala	Asp	Val
165				170				175							
Asn	Ala	Cys	Asp	Asn	Met	Gly	Arg	Asn	Ala	Leu	Ile	His	Ala	Leu	Leu
180				185				190							
Ser	Ser	Asp	Asp	Ser	Asp	Val	Glu	Ala	Ile	Thr	His	Leu	Leu	Leu	Asp
195				200				205							
His	Gly	Ala	Asp	Val	Asn	Val	Arg	Gly	Glu	Arg	Gly	Lys	Thr	Pro	Leu
210				215				220							
Ile	Leu	Ala	Val	Glu	Lys	Lys	His	Leu	Gly	Leu	Val	Gln	Arg	Leu	Leu
225				230				235				240			
Glu	Gln	Glu	His	Ile	Glu	Ile	Asn	Asp	Thr	Asp	Ser	Asp	Gly	Lys	Thr
245				250				255							
Ala	Leu	Leu	Leu	Ala	Val	Glu	Leu	Lys	Leu	Lys	Lys	Ile	Ala	Glu	Leu
260				265				270							
Leu	Cys	Lys	Arg	Gly	Ala	Ser	Thr	Asp	Cys	Gly	Asp	Leu	Val	Met	Thr
275				280				285							
Ala	Arg	Arg	Asn	Tyr	Asp	His	Ser	Leu	Val	Lys	Val	Leu	Leu	Ser	His
290				295				300							
Gly	Ala	Lys	Glu	Asp	Phe	His	Pro	Pro	Ala	Glu	Asp	Trp	Lys	Pro	Gln
305				310				315				320			
Ser	Ser	His	Trp	Gly	Ala	Ala	Leu	Lys	Asp	Leu	His	Arg	Ile	Tyr	Arg

				325					330					335			
Pro	Met	Ile	Gly	Lys	Leu	Lys	Phe	Phe	Ile	Asp	Glu	Lys	Tyr	Lys	Ile		
				340					345					350			
Ala	Asp	Thr	Ser	Glu	Gly	Gly	Ile	Tyr	Leu	Gly	Phe	Tyr	Glu	Lys	Gln		
				355					360					365			
Glu	Val	Ala	Val	Lys	Thr	Phe	Cys	Glu	Gly	Ser	Pro	Arg	Ala	Gln	Arg		
				370					375					380			
Glu	Val	Ser	Cys	Leu	Gln	Ser	Ser	Arg	Glu	Asn	Ser	His	Leu	Val	Thr		
385					390					395				400			
Phe	Tyr	Gly	Ser	Glu	Ser	His	Arg	Gly	His	Leu	Phe	Val	Cys	Val	Thr		
				405					410					415			
Leu	Cys	Glu	Gln	Thr	Leu	Glu	Ala	Cys	Leu	Asp	Val	His	Arg	Gly	Glu		
				420					425					430			
Asp	Val	Glu	Asn	Glu	Glu	Asp	Glu	Phe	Ala	Arg	Asn	Val	Leu	Ser	Ser		
				435					440					445			
Ile	Phe	Lys	Ala	Val	Gln	Glu	Leu	His	Leu	Ser	Cys	Gly	Tyr	Thr	His		
				450					455					460			
Gln	Asp	Leu	Gln	Pro	Gln	Asn	Ile	Leu	Ile	Asp	Ser	Lys	Lys	Ala	Ala		
465					470					475				480			
His	Leu	Ala	Asp	Phe	Asp	Lys	Ser	Ile	Lys	Trp	Ala	Gly	Asp	Pro	Gln		
				485					490					495			
Glu	Val	Lys	Arg	Asp	Leu	Glu	Asp	Leu	Gly	Arg	Leu	Val	Leu	Tyr	Val		
				500					505					510			
Val	Lys	Lys	Gly	Ser	Ile	Ser	Phe	Glu	Asp	Leu	Lys	Ala	Gln	Ser	Asn		
				515					520					525			
Glu	Glu	Val	Val	Gln	Leu	Ser	Pro	Asp	Glu	Glu	Thr	Lys	Asp	Leu	Ile		
				530					535					540			
His	Arg	Leu	Phe	His	Pro	Gly	Glu	His	Val	Arg	Asp	Cys	Leu	Ser	Asp		
545					550					555				560			
Leu	Leu	Gly	His	Pro	Phe	Phe	Trp	Thr	Trp	Glu	Ser	Arg	Tyr	Arg	Thr		
				565					570					575			
Leu	Arg	Asn	Val	Gly	Asn	Glu	Ser	Asp	Ile	Lys	Thr	Arg	Lys	Ser	Glu		
				580					585					590			
Ser	Glu	Ile	Leu	Arg	Leu	Leu	Gln	Pro	Gly	Pro	Ser	Glu	His	Ser	Lys		
				595					600					605			
Ser	Phe	Asp	Lys	Trp	Thr	Thr	Lys	Ile	Asn	Glu	Cys	Val	Met	Lys	Lys		
				610					615					620			
Met	Asn	Lys	Phe	Tyr	Glu	Lys	Arg	Gly	Asn	Phe	Tyr	Gln	Asn	Thr	Val		
625					630					635				640			

Gly Asp Leu Leu Lys Phe Ile Arg Asn Leu Gly Glu His Ile Asp Glu
 645 650 655
 Glu Lys His Lys Lys Met Lys Leu Lys Ile Gly Asp Pro Ser Leu Tyr
 660 665 670
 Phe Gln Lys Thr Phe Pro Asp Leu Val Ile Tyr Val Tyr Thr Lys Leu
 675 680 685
 Gln Asn Thr Glu Tyr Arg Lys His Phe Pro Gln Thr His Ser Pro Asn
 690 695 700
 Lys Pro Gln Cys Asp Gly Ala Gly Gly Ala Ser Gly Leu Ala Ser Pro
 705 710 715 720
 Gly Cys
 <210> 28
 <211> 217
 <212> PRT
 <213> 未知的
 <220>
 <223> RNA酶T2
 <400> 28
 Val Gln His Trp Pro Glu Thr Val Cys Glu Lys Ile Gln Asn Asp Cys
 1 5 10 15
 Arg Asp Pro Pro Asp Tyr Trp Thr Ile His Gly Leu Trp Pro Asp Lys
 20 25 30
 Ser Glu Gly Cys Asn Arg Ser Trp Pro Phe Asn Leu Glu Glu Ile Lys
 35 40 45
 Asp Leu Leu Pro Glu Met Arg Ala Tyr Trp Pro Asp Val Ile His Ser
 50 55 60
 Phe Pro Asn Arg Ser Arg Phe Trp Lys His Glu Trp Glu Lys His Gly
 65 70 75 80
 Thr Cys Ala Ala Gln Val Asp Ala Leu Asn Ser Gln Lys Lys Tyr Phe
 85 90 95
 Gly Arg Ser Leu Glu Leu Tyr Arg Glu Leu Asp Leu Asn Ser Val Leu
 100 105 110
 Leu Lys Leu Gly Ile Lys Pro Ser Ile Asn Tyr Tyr Gln Val Ala Asp
 115 120 125
 Phe Lys Asp Ala Leu Ala Arg Val Tyr Gly Val Ile Pro Lys Ile Gln
 130 135 140
 Cys Leu Pro Pro Ser Gln Asp Glu Glu Val Gln Thr Ile Gly Gln Ile
 145 150 155 160
 Glu Leu Cys Leu Thr Lys Gln Asp Gln Gln Leu Gln Asn Cys Thr Glu

119

<211> 185

<212> PRT

<213> 未知的

<220>

<223> RNA酶T2样蛋白

<220>

<221> 尚未归类的特征

<222> (1) .. (1)

<223> Xaa可以是任何天然存在的氨基酸

<400> 30

Xaa	Leu	Gly	Gly	Ala	Asp	Lys	Arg	Leu	Arg	Asp	Asn	His	Glu	Trp	Lys
1				5				10					15		
Lys	Leu	Ile	Met	Val	Gln	His	Trp	Pro	Glu	Thr	Val	Cys	Glu	Lys	Ile
			20				25					30			
Gln	Asn	Asp	Cys	Arg	Asp	Pro	Pro	Asp	Tyr	Trp	Thr	Ile	His	Gly	Leu
		35				40					45				
Trp	Pro	Asp	Lys	Ser	Glu	Gly	Cys	Asn	Arg	Ser	Trp	Pro	Phe	Asn	Leu
	50					55				60					
Glu	Glu	Ile	Lys	Asp	Leu	Leu	Pro	Glu	Met	Arg	Ala	Tyr	Trp	Pro	Asp
65				70				75					80		
Val	Ile	His	Ser	Phe	Pro	Asn	Arg	Ser	Arg	Phe	Trp	Lys	His	Glu	Trp
			85				90					95			
Glu	Lys	His	Gly	Thr	Cys	Ala	Ala	Gln	Val	Asp	Ala	Leu	Asn	Ser	Gln
		100				105						110			
Lys	Lys	Tyr	Phe	Gly	Arg	Ser	Leu	Glu	Leu	Tyr	Arg	Glu	Leu	Asp	Leu
	115					120					125				
Asn	Ser	Val	Leu	Leu	Lys	Leu	Gly	Ile	Lys	Pro	Ser	Ile	Asn	Tyr	Tyr
	130					135				140					
Gln	Thr	Thr	Glu	Glu	Asp	Leu	Asn	Leu	Asp	Val	Glu	Pro	Thr	Thr	Glu
145				150				155					160		
Asp	Thr	Ala	Glu	Glu	Val	Thr	Ile	His	Val	Leu	Leu	His	Ser	Ala	Leu
			165				170					175			
Phe	Gly	Glu	Ile	Gly	Pro	Arg	Arg	Trp							
		180					185								

<210> 31

<211> 299

<212> PRT

<213> 未知的

<220>

<223> NOB1

<400> 31

Ala	Pro	Val	Glu	His	Val	Val	Ala	Asp	Ala	Gly	Ala	Phe	Leu	Arg	His
1				5					10					15	
Ala	Ala	Leu	Gln	Asp	Ile	Gly	Lys	Asn	Ile	Tyr	Thr	Ile	Arg	Glu	Val
		20					25						30		
Val	Thr	Glu	Ile	Arg	Asp	Lys	Ala	Thr	Arg	Arg	Arg	Leu	Ala	Val	Leu
		35					40					45			
Pro	Tyr	Glu	Leu	Arg	Phe	Lys	Glu	Pro	Leu	Pro	Glu	Tyr	Val	Arg	Leu
	50					55					60				
Val	Thr	Glu	Phe	Ser	Lys	Lys	Thr	Gly	Asp	Tyr	Pro	Ser	Leu	Ser	Ala
65					70				75					80	
Thr	Asp	Ile	Gln	Val	Leu	Ala	Leu	Thr	Tyr	Gln	Leu	Glu	Ala	Glu	Phe
			85						90					95	
Val	Gly	Val	Ser	His	Leu	Lys	Gln	Glu	Pro	Gln	Lys	Val	Lys	Val	Ser
			100						105				110		
Ser	Ser	Ile	Gln	His	Pro	Glu	Thr	Pro	Leu	His	Ile	Ser	Gly	Phe	His
		115					120					125			
Leu	Pro	Tyr	Lys	Pro	Lys	Pro	Pro	Gln	Glu	Thr	Glu	Lys	Gly	His	Ser
	130						135					140			
Ala	Cys	Glu	Pro	Glu	Asn	Leu	Glu	Phe	Ser	Ser	Phe	Met	Phe	Trp	Arg
145					150					155					160
Asn	Pro	Leu	Pro	Asn	Ile	Asp	His	Glu	Leu	Gln	Glu	Leu	Leu	Ile	Asp
				165					170					175	
Arg	Gly	Glu	Asp	Val	Pro	Ser	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Asn	Gly
			180					185						190	
Phe	Glu	Asp	Arg	Lys	Asp	Asp	Ser	Asp	Asp	Asp	Gly	Gly	Gly	Trp	Ile
		195					200					205			
Thr	Pro	Ser	Asn	Ile	Lys	Gln	Ile	Gln	Gln	Glu	Leu	Glu	Gln	Cys	Asp
	210					215					220				
Val	Pro	Glu	Asp	Val	Arg	Val	Gly	Cys	Leu	Thr	Thr	Asp	Phe	Ala	Met
225					230					235				240	
Gln	Asn	Val	Leu	Leu	Gln	Met	Gly	Leu	His	Val	Leu	Ala	Val	Asn	Gly
			245						250					255	
Met	Leu	Ile	Arg	Glu	Ala	Arg	Ser	Tyr	Ile	Leu	Arg	Cys	His	Gly	Cys
			260					265					270		
Phe	Lys	Thr	Thr	Ser	Asp	Met	Ser	Arg	Val	Phe	Cys	Ser	His	Cys	Gly
		275						280					285		
Asn	Lys	Thr	Leu	Lys	Lys	Val	Ser	Val	Thr	Val					

290 295
 <210> 32
 <211> 210
 <212> PRT
 <213> 未知的
 <220>
 <223> ENDOV
 <400> 32
 Ala Phe Ser Gly Leu Gln Arg Val Gly Gly Val Asp Val Ser Phe Val
 1 5 10 15
 Lys Gly Asp Ser Val Arg Ala Cys Ala Ser Leu Val Val Leu Ser Phe
 20 25 30
 Pro Glu Leu Glu Val Val Tyr Glu Glu Ser Arg Met Val Ser Leu Thr
 35 40 45
 Ala Pro Tyr Val Ser Gly Phe Leu Ala Phe Arg Glu Val Pro Phe Leu
 50 55 60
 Leu Glu Leu Val Gln Gln Leu Arg Glu Lys Glu Pro Gly Leu Met Pro
 65 70 75 80
 Gln Val Leu Leu Val Asp Gly Asn Gly Val Leu His His Arg Gly Phe
 85 90 95
 Gly Val Ala Cys His Leu Gly Val Leu Thr Asp Leu Pro Cys Val Gly
 100 105 110
 Val Ala Lys Lys Leu Leu Gln Val Asp Gly Leu Glu Asn Asn Ala Leu
 115 120 125
 His Lys Glu Lys Ile Arg Leu Leu Gln Thr Arg Gly Asp Ser Phe Pro
 130 135 140
 Leu Leu Gly Asp Ser Gly Thr Val Leu Gly Met Ala Leu Arg Ser His
 145 150 155 160
 Asp Arg Ser Thr Arg Pro Leu Tyr Ile Ser Val Gly His Arg Met Ser
 165 170 175
 Leu Glu Ala Ala Val Arg Leu Thr Cys Cys Cys Cys Arg Phe Arg Ile
 180 185 190
 Pro Glu Pro Val Arg Gln Ala Asp Ile Cys Ser Arg Glu His Ile Arg
 195 200 205
 Lys Ser
 210
 <210> 33
 <211> 249
 <212> PRT

<213> 未知的

<220>

<223> ENDOG

<400> 33

Ala	Glu	Leu	Pro	Pro	Val	Pro	Gly	Gly	Pro	Arg	Gly	Pro	Gly	Glu	Leu
1				5					10					15	
Ala	Lys	Tyr	Gly	Leu	Pro	Gly	Leu	Ala	Gln	Leu	Lys	Ser	Arg	Glu	Ser
			20					25					30		
Tyr	Val	Leu	Cys	Tyr	Asp	Pro	Arg	Thr	Arg	Gly	Ala	Leu	Trp	Val	Val
			35				40					45			
Glu	Gln	Leu	Arg	Pro	Glu	Arg	Leu	Arg	Gly	Asp	Gly	Asp	Arg	Arg	Glu
			50				55				60				
Cys	Asp	Phe	Arg	Glu	Asp	Asp	Ser	Val	His	Ala	Tyr	His	Arg	Ala	Thr
65					70					75					80
Asn	Ala	Asp	Tyr	Arg	Gly	Ser	Gly	Phe	Asp	Arg	Gly	His	Leu	Ala	Ala
				85					90					95	
Ala	Ala	Asn	His	Arg	Trp	Ser	Gln	Lys	Ala	Met	Asp	Asp	Thr	Phe	Tyr
			100					105					110		
Leu	Ser	Asn	Val	Ala	Pro	Gln	Val	Pro	His	Leu	Asn	Gln	Asn	Ala	Trp
			115					120				125			
Asn	Asn	Leu	Glu	Lys	Tyr	Ser	Arg	Ser	Leu	Thr	Arg	Ser	Tyr	Gln	Asn
			130				135				140				
Val	Tyr	Val	Cys	Thr	Gly	Pro	Leu	Phe	Leu	Pro	Arg	Thr	Glu	Ala	Asp
145					150					155					160
Gly	Lys	Ser	Tyr	Val	Lys	Tyr	Gln	Val	Ile	Gly	Lys	Asn	His	Val	Ala
				165					170					175	
Val	Pro	Thr	His	Phe	Phe	Lys	Val	Leu	Ile	Leu	Glu	Ala	Ala	Gly	Gly
			180					185					190		
Gln	Ile	Glu	Leu	Arg	Thr	Tyr	Val	Met	Pro	Asn	Ala	Pro	Val	Asp	Glu
			195					200				205			
Ala	Ile	Pro	Leu	Glu	Arg	Phe	Leu	Val	Pro	Ile	Glu	Ser	Ile	Glu	Arg
			210				215				220				
Ala	Ser	Gly	Leu	Leu	Phe	Val	Pro	Asn	Ile	Leu	Ala	Arg	Ala	Gly	Ser
225					230					235					240
Leu	Lys	Ala	Ile	Thr	Ala	Gly	Ser	Lys							
					245										

<210> 34

<211> 479

<212> PRT

<213> 未知的

<220>

<223> ENDOD1

<400> 34

Arg	Leu	Val	Gly	Glu	Glu	Glu	Ala	Gly	Phe	Gly	Glu	Cys	Asp	Lys	Phe
1				5					10					15	
Phe	Tyr	Ala	Gly	Thr	Pro	Pro	Ala	Gly	Leu	Ala	Ala	Asp	Ser	His	Val
			20					25					30		
Lys	Ile	Cys	Gln	Arg	Ala	Glu	Gly	Ala	Glu	Arg	Phe	Ala	Thr	Leu	Tyr
		35					40					45			
Ser	Thr	Arg	Asp	Arg	Ile	Pro	Val	Tyr	Ser	Ala	Phe	Arg	Ala	Pro	Arg
	50				55					60					
Pro	Ala	Pro	Gly	Gly	Ala	Glu	Gln	Arg	Trp	Leu	Val	Glu	Pro	Gln	Ile
65					70				75					80	
Asp	Asp	Pro	Asn	Ser	Asn	Leu	Glu	Glu	Ala	Ile	Asn	Glu	Ala	Glu	Ala
				85					90					95	
Ile	Thr	Ser	Val	Asn	Ser	Leu	Gly	Ser	Lys	Gln	Ala	Leu	Asn	Thr	Asp
			100					105					110		
Tyr	Leu	Asp	Ser	Asp	Tyr	Gln	Arg	Gly	Gln	Leu	Tyr	Pro	Phe	Ser	Leu
		115					120					125			
Ser	Ser	Asp	Val	Gln	Val	Ala	Thr	Phe	Thr	Leu	Thr	Asn	Ser	Ala	Pro
	130					135					140				
Met	Thr	Gln	Ser	Phe	Gln	Glu	Arg	Trp	Tyr	Val	Asn	Leu	His	Ser	Leu
145					150					155				160	
Met	Asp	Arg	Ala	Leu	Thr	Pro	Gln	Cys	Gly	Ser	Gly	Glu	Asp	Leu	Tyr
			165						170				175		
Ile	Leu	Thr	Gly	Thr	Val	Pro	Ser	Asp	Tyr	Arg	Val	Lys	Asp	Lys	Val
			180					185					190		
Ala	Val	Pro	Glu	Phe	Val	Trp	Leu	Ala	Ala	Cys	Cys	Ala	Val	Pro	Gly
		195					200					205			
Gly	Gly	Trp	Ala	Met	Gly	Phe	Val	Lys	His	Thr	Arg	Asp	Ser	Asp	Ile
	210					215					220				
Ile	Glu	Asp	Val	Met	Val	Lys	Asp	Leu	Gln	Lys	Leu	Leu	Pro	Phe	Asn
225					230					235				240	
Pro	Gln	Leu	Phe	Gln	Asn	Asn	Cys	Gly	Glu	Thr	Glu	Gln	Asp	Thr	Glu
			245						250				255		
Lys	Met	Lys	Lys	Ile	Leu	Glu	Val	Val	Asn	Gln	Ile	Gln	Asp	Glu	Glu
			260						265				270		
Arg	Met	Val	Gln	Ser	Gln	Lys	Ser	Ser	Ser	Pro	Leu	Ser	Ser	Thr	Arg

275	280	285
Ser Lys Arg Ser Thr Leu Leu Pro Pro Glu Ala Ser Glu Gly Ser Ser		
290	295	300
Ser Phe Leu Gly Lys Leu Met Gly Phe Ile Ala Thr Pro Phe Ile Lys		
305	310	315
Leu Phe Gln Leu Ile Tyr Tyr Leu Val Val Ala Ile Leu Lys Asn Ile		
325	330	335
Val Tyr Phe Leu Trp Cys Val Thr Lys Gln Val Ile Asn Gly Ile Glu		
340	345	350
Ser Cys Leu Tyr Arg Leu Gly Ser Ala Thr Ile Ser Tyr Phe Met Ala		
355	360	365
Ile Gly Glu Glu Leu Val Ser Ile Pro Trp Lys Val Leu Lys Val Val		
370	375	380
Ala Lys Val Ile Arg Ala Leu Leu Arg Ile Leu Cys Cys Leu Leu Lys		
385	390	395
Ala Ile Cys Arg Val Leu Ser Ile Pro Val Arg Val Leu Val Asp Val		
405	410	415
Ala Thr Phe Pro Val Tyr Thr Met Gly Ala Ile Pro Ile Val Cys Lys		
420	425	430
Asp Ile Ala Leu Gly Leu Gly Gly Thr Val Ser Leu Leu Phe Asp Thr		
435	440	445
Ala Phe Gly Thr Leu Gly Gly Leu Phe Gln Val Val Phe Ser Val Cys		
450	455	460
Lys Arg Ile Gly Tyr Lys Val Thr Phe Asp Asn Ser Gly Glu Leu		
465	470	475
<210> 35		
<211> 380		
<212> PRT		
<213> 未知的		
<220>		
<223> hFEN1		
<400> 35		
Met Gly Ile Gln Gly Leu Ala Lys Leu Ile Ala Asp Val Ala Pro Ser		
1	5	10
Ala Ile Arg Glu Asn Asp Ile Lys Ser Tyr Phe Gly Arg Lys Val Ala		
20	25	30
Ile Asp Ala Ser Met Ser Ile Tyr Gln Phe Leu Ile Ala Val Arg Gln		
35	40	45
Gly Gly Asp Val Leu Gln Asn Glu Glu Gly Glu Thr Thr Ser His Leu		

50	55	60
Met Gly Met Phe Tyr Arg Thr Ile Arg Met Met Glu Asn Gly Ile Lys		
65	70	75
Pro Val Tyr Val Phe Asp Gly Lys Pro Pro Gln Leu Lys Ser Gly Glu		
85	90	95
Leu Ala Lys Arg Ser Glu Arg Arg Ala Glu Ala Glu Lys Gln Leu Gln		
100	105	110
Gln Ala Gln Ala Ala Gly Ala Glu Gln Glu Val Glu Lys Phe Thr Lys		
115	120	125
Arg Leu Val Lys Val Thr Lys Gln His Asn Asp Glu Cys Lys His Leu		
130	135	140
Leu Ser Leu Met Gly Ile Pro Tyr Leu Asp Ala Pro Ser Glu Ala Glu		
145	150	155
Ala Ser Cys Ala Ala Leu Val Lys Ala Gly Lys Val Tyr Ala Ala Ala		
165	170	175
Thr Glu Asp Met Asp Cys Leu Thr Phe Gly Ser Pro Val Leu Met Arg		
180	185	190
His Leu Thr Ala Ser Glu Ala Lys Lys Leu Pro Ile Gln Glu Phe His		
195	200	205
Leu Ser Arg Ile Leu Gln Glu Leu Gly Leu Asn Gln Glu Gln Phe Val		
210	215	220
Asp Leu Cys Ile Leu Leu Gly Ser Asp Tyr Cys Glu Ser Ile Arg Gly		
225	230	235
Ile Gly Pro Lys Arg Ala Val Asp Leu Ile Gln Lys His Lys Ser Ile		
245	250	255
Glu Glu Ile Val Arg Arg Leu Asp Pro Asn Lys Tyr Pro Val Pro Glu		
260	265	270
Asn Trp Leu His Lys Glu Ala His Gln Leu Phe Leu Glu Pro Glu Val		
275	280	285
Leu Asp Pro Glu Ser Val Glu Leu Lys Trp Ser Glu Pro Asn Glu Glu		
290	295	300
Glu Leu Ile Lys Phe Met Cys Gly Glu Lys Gln Phe Ser Glu Glu Arg		
305	310	315
Ile Arg Ser Gly Val Lys Arg Leu Ser Lys Ser Arg Gln Gly Ser Thr		
325	330	335
Gln Gly Arg Leu Asp Asp Phe Phe Lys Val Thr Gly Ser Leu Ser Ser		
340	345	350
Ala Lys Arg Lys Glu Pro Glu Pro Lys Gly Ser Thr Lys Lys Lys Ala		
355	360	365

Lys Thr Gly Ala Ala Gly Lys Phe Lys Arg Gly Lys
 370 375 380

<210> 36

<211> 186

<212> PRT

<213> 智人(Homo Sapiens)

<400> 36

Glu Ser Thr His Val Glu Phe Lys Arg Phe Thr Thr Lys Lys Val Ile
 1 5 10 15

Pro Arg Ile Lys Glu Met Leu Pro His Tyr Val Ser Ala Phe Ala Asn
 20 25 30

Thr Gln Gly Gly Tyr Val Leu Ile Gly Val Asp Asp Lys Ser Lys Glu
 35 40 45

Val Val Gly Cys Lys Trp Glu Lys Val Asn Pro Asp Leu Leu Lys Lys
 50 55 60

Glu Ile Glu Asn Cys Ile Glu Lys Leu Pro Thr Phe His Phe Cys Cys
 65 70 75 80

Glu Lys Pro Lys Val Asn Phe Thr Thr Lys Ile Leu Asn Val Tyr Gln
 85 90 95

Lys Asp Val Leu Asp Gly Tyr Val Cys Val Ile Gln Val Glu Pro Phe
 100 105 110

Cys Cys Val Val Phe Ala Glu Ala Pro Asp Ser Trp Ile Met Lys Asp
 115 120 125

Asn Ser Val Thr Arg Leu Thr Ala Glu Gln Trp Val Val Met Met Leu
 130 135 140

Asp Thr Gln Ser Ala Pro Pro Ser Leu Val Thr Asp Tyr Asn Ser Cys
 145 150 155 160

Leu Ile Ser Ser Ala Ser Ser Ala Arg Lys Ser Pro Gly Tyr Pro Ile
 165 170 175

Lys Val His Lys Phe Lys Glu Ala Leu Gln
 180 185

<210> 37

<211> 262

<212> PRT

<213> 智人(Homo Sapiens)

<400> 37

Thr Leu Gln Gly Thr Asn Thr Tyr Leu Val Gly Thr Gly Pro Arg Arg
 1 5 10 15

Ile Leu Ile Asp Thr Gly Glu Pro Ala Ile Pro Glu Tyr Ile Ser Cys

20	25	30
Leu Lys Gln Ala Leu Thr Glu Phe Asn Thr Ala Ile Gln Glu Ile Val		
35	40	45
Val Thr His Trp His Arg Asp His Ser Gly Gly Ile Gly Asp Ile Cys		
50	55	60
Lys Ser Ile Asn Asn Asp Thr Thr Tyr Cys Ile Lys Lys Leu Pro Arg		
65	70	75
Asn Pro Gln Arg Glu Glu Ile Ile Gly Asn Gly Glu Gln Gln Tyr Val		
85	90	95
Tyr Leu Lys Asp Gly Asp Val Ile Lys Thr Glu Gly Ala Thr Leu Arg		
100	105	110
Val Leu Tyr Thr Pro Gly His Thr Asp Asp His Met Ala Leu Leu Leu		
115	120	125
Glu Glu Glu Asn Ala Ile Phe Ser Gly Asp Cys Ile Leu Gly Glu Gly		
130	135	140
Thr Thr Val Phe Glu Asp Leu Tyr Asp Tyr Met Asn Ser Leu Lys Glu		
145	150	155
Leu Leu Lys Ile Lys Ala Asp Ile Ile Tyr Pro Gly His Gly Pro Val		
165	170	175
Ile His Asn Ala Glu Ala Lys Ile Gln Gln Tyr Ile Ser His Arg Asn		
180	185	190
Ile Arg Glu Gln Gln Ile Leu Thr Leu Phe Arg Glu Asn Phe Glu Lys		
195	200	205
Ser Phe Thr Val Met Glu Leu Val Lys Ile Ile Tyr Lys Asn Thr Pro		
210	215	220
Glu Asn Leu His Glu Met Ala Lys His Asn Leu Leu Leu His Leu Lys		
225	230	235
Lys Leu Glu Lys Glu Gly Lys Ile Phe Ser Asn Thr Asp Pro Asp Lys		
245	250	255
Lys Trp Lys Ala His Leu		
260		
<210> 38		
<211> 518		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 38		
Met Leu Arg Val Val Ser Trp Asn Ile Asn Gly Ile Arg Arg Pro Leu		
1	5	10
Gln Gly Val Ala Asn Gln Glu Pro Ser Asn Cys Ala Ala Val Ala Val		

20	25	30
Gly Arg Ile Leu Asp Glu Leu Asp Ala Asp Ile Val Cys Leu Gln Glu		
35	40	45
Thr Lys Val Thr Arg Asp Ala Leu Thr Glu Pro Leu Ala Ile Val Glu		
50	55	60
Gly Tyr Asn Ser Tyr Phe Ser Phe Ser Arg Asn Arg Ser Gly Tyr Ser		
65	70	75
Gly Val Ala Thr Phe Cys Lys Asp Asn Ala Thr Pro Val Ala Ala Glu		
85	90	95
Glu Gly Leu Ser Gly Leu Phe Ala Thr Gln Asn Gly Asp Val Gly Cys		
100	105	110
Tyr Gly Asn Met Asp Glu Phe Thr Gln Glu Glu Leu Arg Ala Leu Asp		
115	120	125
Ser Glu Gly Arg Ala Leu Leu Thr Gln His Lys Ile Arg Thr Trp Glu		
130	135	140
Gly Lys Glu Lys Thr Leu Thr Leu Ile Asn Val Tyr Cys Pro His Ala		
145	150	155
Asp Pro Gly Arg Pro Glu Arg Leu Val Phe Lys Met Arg Phe Tyr Arg		
165	170	175
Leu Leu Gln Ile Arg Ala Glu Ala Leu Leu Ala Ala Gly Ser His Val		
180	185	190
Ile Ile Leu Gly Asp Leu Asn Thr Ala His Arg Pro Ile Asp His Trp		
195	200	205
Asp Ala Val Asn Leu Glu Cys Phe Glu Glu Asp Pro Gly Arg Lys Trp		
210	215	220
Met Asp Ser Leu Leu Ser Asn Leu Gly Cys Gln Ser Ala Ser His Val		
225	230	235
Gly Pro Phe Ile Asp Ser Tyr Arg Cys Phe Gln Pro Lys Gln Glu Gly		
245	250	255
Ala Phe Thr Cys Trp Ser Ala Val Thr Gly Ala Arg His Leu Asn Tyr		
260	265	270
Gly Ser Arg Leu Asp Tyr Val Leu Gly Asp Arg Thr Leu Val Ile Asp		
275	280	285
Thr Phe Gln Ala Ser Phe Leu Leu Pro Glu Val Met Gly Ser Asp His		
290	295	300
Cys Pro Val Gly Ala Val Leu Ser Val Ser Ser Val Pro Ala Lys Gln		
305	310	315
Cys Pro Pro Leu Cys Thr Arg Phe Leu Pro Glu Phe Ala Gly Thr Gln		
325	330	335

Leu Lys Ile Leu Arg Phe Leu Val Pro Leu Glu Gln Ser Pro Val Leu
 340 345 350
 Glu Gln Ser Thr Leu Gln His Asn Asn Gln Thr Arg Val Gln Thr Cys
 355 360 365
 Gln Asn Lys Ala Gln Val Arg Ser Thr Arg Pro Gln Pro Ser Gln Val
 370 375 380
 Gly Ser Ser Arg Gly Gln Lys Asn Leu Lys Ser Tyr Phe Gln Pro Ser
 385 390 395 400
 Pro Ser Cys Pro Gln Ala Ser Pro Asp Ile Glu Leu Pro Ser Leu Pro
 405 410 415
 Leu Met Ser Ala Leu Met Thr Pro Lys Thr Pro Glu Glu Lys Ala Val
 420 425 430
 Ala Lys Val Val Lys Gly Gln Ala Lys Thr Ser Glu Ala Lys Asp Glu
 435 440 445
 Lys Glu Leu Arg Thr Ser Phe Trp Lys Ser Val Leu Ala Gly Pro Leu
 450 455 460
 Arg Thr Pro Leu Cys Gly Gly His Arg Glu Pro Cys Val Met Arg Thr
 465 470 475 480
 Val Lys Lys Pro Gly Pro Asn Leu Gly Arg Arg Phe Tyr Met Cys Ala
 485 490 495
 Arg Pro Arg Gly Pro Pro Thr Asp Pro Ser Ser Arg Cys Asn Phe Phe
 500 505 510
 Leu Trp Ser Arg Pro Ser
 515
 <210> 39
 <211> 350
 <212> PRT
 <213> 智人(Homo Sapiens)
 <400> 39
 Met Leu Arg Val Val Ser Trp Asn Ile Asn Gly Ile Arg Arg Pro Leu
 1 5 10 15
 Gln Gly Val Ala Asn Gln Glu Pro Ser Asn Cys Ala Ala Val Ala Val
 20 25 30
 Gly Arg Ile Leu Asp Glu Leu Asp Ala Asp Ile Val Cys Leu Gln Glu
 35 40 45
 Thr Lys Val Thr Arg Asp Ala Leu Thr Glu Pro Leu Ala Ile Val Glu
 50 55 60
 Gly Tyr Asn Ser Tyr Phe Ser Phe Ser Arg Asn Arg Ser Gly Tyr Ser
 65 70 75 80

Gly Val Ala Thr Phe Cys Lys Asp Asn Ala Thr Pro Val Ala Ala Glu		
85	90	95
Glu Gly Leu Ser Gly Leu Phe Ala Thr Gln Asn Gly Asp Val Gly Cys		
100	105	110
Tyr Gly Asn Met Asp Glu Phe Thr Gln Glu Glu Leu Arg Ala Leu Asp		
115	120	125
Ser Glu Gly Arg Ala Leu Leu Thr Gln His Lys Ile Arg Thr Trp Glu		
130	135	140
Gly Lys Glu Lys Thr Leu Thr Leu Ile Asn Val Tyr Cys Pro His Ala		
145	150	155
Asp Pro Gly Arg Pro Glu Arg Leu Val Phe Lys Met Arg Phe Tyr Arg		
165	170	175
Leu Leu Gln Ile Arg Ala Glu Ala Leu Leu Ala Ala Gly Ser His Val		
180	185	190
Ile Ile Leu Gly Asp Leu Asn Thr Ala His Arg Pro Ile Asp His Trp		
195	200	205
Asp Ala Val Asn Leu Glu Cys Phe Glu Glu Asp Pro Gly Arg Lys Trp		
210	215	220
Met Asp Ser Leu Leu Ser Asn Leu Gly Cys Gln Ser Ala Ser His Val		
225	230	235
Gly Pro Phe Ile Asp Ser Tyr Arg Cys Phe Gln Pro Lys Gln Glu Gly		
245	250	255
Ala Phe Thr Cys Trp Ser Ala Val Thr Gly Ala Arg His Leu Asn Tyr		
260	265	270
Gly Ser Arg Leu Asp Tyr Val Leu Gly Asp Arg Thr Leu Val Ile Asp		
275	280	285
Thr Phe Gln Ala Ser Phe Leu Leu Pro Glu Val Met Gly Ser Asp His		
290	295	300
Cys Pro Val Gly Ala Val Leu Ser Val Ser Ser Val Pro Ala Lys Gln		
305	310	315
Cys Pro Pro Leu Cys Thr Arg Phe Leu Pro Glu Phe Ala Gly Thr Gln		
325	330	335
Leu Lys Ile Leu Arg Phe Leu Val Pro Leu Glu Gln Ser Pro		
340	345	350

<210> 40

<211> 123

<212> PRT

<213> 智人(Homo Sapiens)

<400> 40

Ser	Ser	Leu	Ile	Arg	Arg	Val	Ile	Ser	Thr	Ala	Lys	Ala	Pro	Gly	Ala
1				5					10					15	
Ile	Gly	Pro	Tyr	Ser	Gln	Ala	Val	Leu	Val	Asp	Arg	Thr	Ile	Tyr	Ile
			20					25					30		
Ser	Gly	Gln	Ile	Gly	Met	Asp	Pro	Ser	Ser	Gly	Gln	Leu	Val	Ser	Gly
		35					40					45			
Gly	Val	Ala	Glu	Glu	Ala	Lys	Gln	Ala	Leu	Lys	Asn	Met	Gly	Glu	Ile
	50					55					60				
Leu	Lys	Ala	Ala	Gly	Cys	Asp	Phe	Thr	Asn	Val	Val	Lys	Thr	Thr	Val
65					70					75					80
Leu	Leu	Ala	Asp	Ile	Asn	Asp	Phe	Asn	Thr	Val	Asn	Glu	Ile	Tyr	Lys
				85					90					95	
Gln	Tyr	Phe	Lys	Ser	Asn	Phe	Pro	Ala	Arg	Ala	Ala	Tyr	Gln	Val	Ala
			100					105					110		
Ala	Leu	Pro	Lys	Gly	Ser	Arg	Ile	Glu	Ile	Glu	Ala	Val	Ala	Ile	Gln
		115					120					125			
Gly	Pro	Leu	Thr	Thr	Ala	Ser	Leu								
	130					135									

<210> 42

<211> 189

<212> PRT

<213> 智人(Homo Sapiens)

<400> 42

Gly Gly Gly Thr Pro Lys Ala Pro Asn Leu Glu Pro Pro Leu Pro Glu
 1 5 10 15
 Glu Glu Lys Glu Gly Ser Asp Leu Arg Pro Val Val Ile Asp Gly Ser
 20 25 30
 Asn Val Ala Met Ser His Gly Asn Lys Glu Val Phe Ser Cys Arg Gly
 35 40 45
 Ile Leu Leu Ala Val Asn Trp Phe Leu Glu Arg Gly His Thr Asp Ile
 50 55 60
 Thr Val Phe Val Pro Ser Trp Arg Lys Glu Gln Pro Arg Pro Asp Val
 65 70 75 80
 Pro Ile Thr Asp Gln His Ile Leu Arg Glu Leu Glu Lys Lys Lys Ile
 85 90 95
 Leu Val Phe Thr Pro Ser Arg Arg Val Gly Gly Lys Arg Val Val Cys
 100 105 110
 Tyr Asp Asp Arg Phe Ile Val Lys Leu Ala Tyr Glu Ser Asp Gly Ile
 115 120 125
 Val Val Ser Asn Asp Thr Tyr Arg Asp Leu Gln Gly Glu Arg Gln Glu
 130 135 140
 Trp Lys Arg Phe Ile Glu Glu Arg Leu Leu Met Tyr Ser Phe Val Asn
 145 150 155 160
 Asp Lys Phe Met Pro Pro Asp Asp Pro Leu Gly Arg His Gly Pro Ser
 165 170 175
 Leu Asp Asn Phe Leu Arg Lys Lys Pro Leu Thr Leu Glu
 180 185

<210> 43

<211> 598

<212> PRT

<213> 智人(Homo Sapiens)

<400> 43

Ser Gly Pro Cys Gly Glu Lys Pro Val Leu Glu Ala Ser Pro Thr Met
 1 5 10 15
 Ser Leu Trp Glu Phe Glu Asp Ser His Ser Arg Gln Gly Thr Pro Arg
 20 25 30
 Pro Gly Gln Glu Leu Ala Ala Glu Glu Ala Ser Ala Leu Glu Leu Gln

	35					40					45					
Met	Lys	Val	Asp	Phe	Phe	Arg	Lys	Leu	Gly	Tyr	Ser	Ser	Thr	Glu	Ile	
	50					55					60					
His	Ser	Val	Leu	Gln	Lys	Leu	Gly	Val	Gln	Ala	Asp	Thr	Asn	Thr	Val	
65					70					75					80	
Leu	Gly	Glu	Leu	Val	Lys	His	Gly	Thr	Ala	Thr	Glu	Arg	Glu	Arg	Gln	
				85					90					95		
Thr	Ser	Pro	Asp	Pro	Cys	Pro	Gln	Leu	Pro	Leu	Val	Pro	Arg	Gly	Gly	
		100					105					110				
Gly	Thr	Pro	Lys	Ala	Pro	Asn	Leu	Glu	Pro	Pro	Leu	Pro	Glu	Glu	Glu	
	115					120					125					
Lys	Glu	Gly	Ser	Asp	Leu	Arg	Pro	Val	Val	Ile	Asp	Gly	Ser	Asn	Val	
	130					135					140					
Ala	Met	Ser	His	Gly	Asn	Lys	Glu	Val	Phe	Ser	Cys	Arg	Gly	Ile	Leu	
145					150					155					160	
Leu	Ala	Val	Asn	Trp	Phe	Leu	Glu	Arg	Gly	His	Thr	Asp	Ile	Thr	Val	
				165					170					175		
Phe	Val	Pro	Ser	Trp	Arg	Lys	Glu	Gln	Pro	Arg	Pro	Asp	Val	Pro	Ile	
	180					185					190					
Thr	Asp	Gln	His	Ile	Leu	Arg	Glu	Leu	Glu	Lys	Lys	Lys	Ile	Leu	Val	
	195					200					205					
Phe	Thr	Pro	Ser	Arg	Arg	Val	Gly	Gly	Lys	Arg	Val	Val	Cys	Tyr	Asp	
	210					215					220					
Asp	Arg	Phe	Ile	Val	Lys	Leu	Ala	Tyr	Glu	Ser	Asp	Gly	Ile	Val	Val	
225					230					235					240	
Ser	Asn	Asp	Thr	Tyr	Arg	Asp	Leu	Gln	Gly	Glu	Arg	Gln	Glu	Trp	Lys	
				245					250					255		
Arg	Phe	Ile	Glu	Glu	Arg	Leu	Leu	Met	Tyr	Ser	Phe	Val	Asn	Asp	Lys	
	260					265					270					
Phe	Met	Pro	Pro	Asp	Asp	Pro	Leu	Gly	Arg	His	Gly	Pro	Ser	Leu	Asp	
	275					280					285					
Asn	Phe	Leu	Arg	Lys	Lys	Pro	Leu	Thr	Leu	Glu	His	Arg	Lys	Gln	Pro	
	290					295					300					
Cys	Pro	Tyr	Gly	Arg	Lys	Cys	Thr	Tyr	Gly	Ile	Lys	Cys	Arg	Phe	Phe	
305					310					315					320	
His	Pro	Glu	Arg	Pro	Ser	Cys	Pro	Gln	Arg	Ser	Val	Ala	Asp	Glu	Leu	
				325					330					335		
Arg	Ala	Asn	Ala	Leu	Leu	Ser	Pro	Pro	Arg	Ala	Pro	Ser	Lys	Asp	Lys	
	340					345					350					

Asn Gly Arg Arg Pro Ser Pro Ser Ser Gln Ser Ser Ser Leu Leu Thr
 355 360 365
 Glu Ser Glu Gln Cys Ser Leu Asp Gly Lys Lys Leu Gly Ala Gln Ala
 370 375 380
 Ser Pro Gly Ser Arg Gln Glu Gly Leu Thr Gln Thr Tyr Ala Pro Ser
 385 390 395 400
 Gly Arg Ser Leu Ala Pro Ser Gly Gly Ser Gly Ser Ser Phe Gly Pro
 405 410 415
 Thr Asp Trp Leu Pro Gln Thr Leu Asp Ser Leu Pro Tyr Val Ser Gln
 420 425 430
 Asp Cys Leu Asp Ser Gly Ile Gly Ser Leu Glu Ser Gln Met Ser Glu
 435 440 445
 Leu Trp Gly Val Arg Gly Gly Gly Pro Gly Glu Pro Gly Pro Pro Arg
 450 455 460
 Ala Pro Tyr Thr Gly Tyr Ser Pro Tyr Gly Ser Glu Leu Pro Ala Thr
 465 470 475 480
 Ala Ala Phe Ser Ala Phe Gly Arg Ala Met Gly Ala Gly His Phe Ser
 485 490 495
 Val Pro Ala Asp Tyr Pro Pro Ala Pro Pro Ala Phe Pro Pro Arg Glu
 500 505 510
 Tyr Trp Ser Glu Pro Tyr Pro Leu Pro Pro Pro Thr Ser Val Leu Gln
 515 520 525
 Glu Pro Pro Val Gln Ser Pro Gly Ala Gly Arg Ser Pro Trp Gly Arg
 530 535 540
 Ala Gly Ser Leu Ala Lys Glu Gln Ala Ser Val Tyr Thr Lys Leu Cys
 545 550 555 560
 Gly Val Phe Pro Pro His Leu Val Glu Ala Val Met Gly Arg Phe Pro
 565 570 575
 Gln Leu Leu Asp Pro Gln Gln Leu Ala Ala Glu Ile Leu Ser Tyr Lys
 580 585 590
 Ser Gln His Pro Ser Glu
 595

<210> 44

<211> 136

<212> PRT

<213> 智人(Homo Sapiens)

<400> 44

Ser Ser Leu Ile Arg Arg Val Ile Ser Thr Ala Lys Ala Pro Gly Ala
 1 5 10 15

Ile Gly Pro Tyr Ser Gln Ala Val Leu Val Asp Arg Thr Ile Tyr Ile		
20	25	30
Ser Gly Gln Ile Gly Met Asp Pro Ser Ser Gly Gln Leu Val Ser Gly		
35	40	45
Gly Val Ala Glu Glu Ala Lys Gln Ala Leu Lys Asn Met Gly Glu Ile		
50	55	60
Leu Lys Ala Ala Gly Cys Asp Phe Thr Asn Val Val Lys Thr Thr Val		
65	70	75
Leu Leu Ala Asp Ile Asn Asp Phe Asn Thr Val Asn Glu Ile Tyr Lys		
85	90	95
Gln Tyr Phe Lys Ser Asn Phe Pro Ala Arg Ala Ala Tyr Gln Val Ala		
100	105	110
Ala Leu Pro Lys Gly Ser Arg Ile Glu Ile Glu Ala Val Ala Ile Gln		
115	120	125
Gly Pro Leu Thr Thr Ala Ser Leu		
130	135	
<210> 45		
<211> 966		
<212> PRT		
<213> 生黄瘤胃球菌(Ruminococcus flavefaciens)		
<400> 45		
Ile Glu Lys Lys Lys Ser Phe Ala Lys Gly Met Gly Val Lys Ser Thr		
1	5	10
Leu Val Ser Gly Ser Lys Val Tyr Met Thr Thr Phe Ala Glu Gly Ser		
20	25	30
Asp Ala Arg Leu Glu Lys Ile Val Glu Gly Asp Ser Ile Arg Ser Val		
35	40	45
Asn Glu Gly Glu Ala Phe Ser Ala Glu Met Ala Asp Lys Asn Ala Gly		
50	55	60
Tyr Lys Ile Gly Asn Ala Lys Phe Ser His Pro Lys Gly Tyr Ala Val		
65	70	75
Val Ala Asn Asn Pro Leu Tyr Thr Gly Pro Val Gln Gln Asp Met Leu		
85	90	95
Gly Leu Lys Glu Thr Leu Glu Lys Arg Tyr Phe Gly Glu Ser Ala Asp		
100	105	110
Gly Asn Asp Asn Ile Cys Ile Gln Val Ile His Asn Ile Leu Asp Ile		
115	120	125
Glu Lys Ile Leu Ala Glu Tyr Ile Thr Asn Ala Ala Tyr Ala Val Asn		
130	135	140

Asn	Ile	Ser	Gly	Leu	Asp	Lys	Asp	Ile	Ile	Gly	Phe	Gly	Lys	Phe	Ser
145						150				155					160
Thr	Val	Tyr	Thr	Tyr	Asp	Glu	Phe	Lys	Asp	Pro	Glu	His	His	Arg	Ala
						165				170					175
Ala	Phe	Asn	Asn	Asn	Asp	Lys	Leu	Ile	Asn	Ala	Ile	Lys	Ala	Gln	Tyr
						180				185					190
Asp	Glu	Phe	Asp	Asn	Phe	Leu	Asp	Asn	Pro	Arg	Leu	Gly	Tyr	Phe	Gly
						195				200					205
Gln	Ala	Phe	Phe	Ser	Lys	Glu	Gly	Arg	Asn	Tyr	Ile	Ile	Asn	Tyr	Gly
						210				215					220
Asn	Glu	Cys	Tyr	Asp	Ile	Leu	Ala	Leu	Leu	Ser	Gly	Leu	Ala	His	Trp
225						230				235					240
Val	Val	Ala	Asn	Asn	Glu	Glu	Glu	Ser	Arg	Ile	Ser	Arg	Thr	Trp	Leu
						245				250					255
Tyr	Asn	Leu	Asp	Lys	Asn	Leu	Asp	Asn	Glu	Tyr	Ile	Ser	Thr	Leu	Asn
						260				265					270
Tyr	Leu	Tyr	Asp	Arg	Ile	Thr	Asn	Glu	Leu	Thr	Asn	Ser	Phe	Ser	Lys
						275				280					285
Asn	Ser	Ala	Ala	Asn	Val	Asn	Tyr	Ile	Ala	Glu	Thr	Leu	Gly	Ile	Asn
						290				295					300
Pro	Ala	Glu	Phe	Ala	Glu	Gln	Tyr	Phe	Arg	Phe	Ser	Ile	Met	Lys	Glu
305						310				315					320
Gln	Lys	Asn	Leu	Gly	Phe	Asn	Ile	Thr	Lys	Leu	Arg	Glu	Val	Met	Leu
						325				330					335
Asp	Arg	Lys	Asp	Met	Ser	Glu	Ile	Arg	Lys	Asn	His	Lys	Val	Phe	Asp
						340				345					350
Ser	Ile	Arg	Thr	Lys	Val	Tyr	Thr	Met	Met	Asp	Phe	Val	Ile	Tyr	Arg
						355				360					365
Tyr	Tyr	Ile	Glu	Glu	Asp	Ala	Lys	Val	Ala	Ala	Ala	Asn	Lys	Ser	Leu
						370				375					380
Pro	Asp	Asn	Glu	Lys	Ser	Leu	Ser	Glu	Lys	Asp	Ile	Phe	Val	Ile	Asn
385						390				395					400
Leu	Arg	Gly	Ser	Phe	Asn	Asp	Asp	Gln	Lys	Asp	Ala	Leu	Tyr	Tyr	Asp
						405				410					415
Glu	Ala	Asn	Arg	Ile	Trp	Arg	Lys	Leu	Glu	Asn	Ile	Met	His	Asn	Ile
						420				425					430
Lys	Glu	Phe	Arg	Gly	Asn	Lys	Thr	Arg	Glu	Tyr	Lys	Lys	Lys	Asp	Ala
						435				440					445
Pro	Arg	Leu	Pro	Arg	Ile	Leu	Pro	Ala	Gly	Arg	Asp	Val	Ser	Ala	Phe

450	455	460
Ser Lys Leu Met Tyr Ala Leu Thr Met Phe Leu Asp Gly Lys Glu Ile		
465	470	475
Asn Asp Leu Leu Thr Thr Leu Ile Asn Lys Phe Asp Asn Ile Gln Ser		480
	485	490
Phe Leu Lys Val Met Pro Leu Ile Gly Val Asn Ala Lys Phe Val Glu		495
	500	505
Glu Tyr Ala Phe Phe Lys Asp Ser Ala Lys Ile Ala Asp Glu Leu Arg		510
	515	520
Leu Ile Lys Ser Phe Ala Arg Met Gly Glu Pro Ile Ala Asp Ala Arg		525
	530	535
Arg Ala Met Tyr Ile Asp Ala Ile Arg Ile Leu Gly Thr Asn Leu Ser		540
545	550	555
Tyr Asp Glu Leu Lys Ala Leu Ala Asp Thr Phe Ser Leu Asp Glu Asn		560
	565	570
Gly Asn Lys Leu Lys Lys Gly Lys His Gly Met Arg Asn Phe Ile Ile		575
	580	585
Asn Asn Val Ile Ser Asn Lys Arg Phe His Tyr Leu Ile Arg Tyr Gly		590
	595	600
Asp Pro Ala His Leu His Glu Ile Ala Lys Asn Glu Ala Val Val Lys		605
	610	615
Phe Val Leu Gly Arg Ile Ala Asp Ile Gln Lys Lys Gln Gly Gln Asn		620
625	630	635
Gly Lys Asn Gln Ile Asp Arg Tyr Tyr Glu Thr Cys Ile Gly Lys Asp		640
	645	650
Lys Gly Lys Ser Val Ser Glu Lys Val Asp Ala Leu Thr Lys Ile Ile		655
	660	665
Thr Gly Met Asn Tyr Asp Gln Phe Asp Lys Lys Arg Ser Val Ile Glu		670
	675	680
Asp Thr Gly Arg Glu Asn Ala Glu Arg Glu Lys Phe Lys Lys Ile Ile		685
	690	695
Ser Leu Tyr Leu Thr Val Ile Tyr His Ile Leu Lys Asn Ile Val Asn		700
705	710	715
Ile Asn Ala Arg Tyr Val Ile Gly Phe His Cys Val Glu Arg Asp Ala		720
	725	730
Gln Leu Tyr Lys Glu Lys Gly Tyr Asp Ile Asn Leu Lys Lys Leu Glu		735
	740	745
Glu Lys Gly Phe Ser Ser Val Thr Lys Leu Cys Ala Gly Ile Asp Glu		750
	755	760
		765

Thr Ala Pro Asp Lys Arg Lys Asp Val Glu Lys Glu Met Ala Glu Arg		
770	775	780
Ala Lys Glu Ser Ile Asp Ser Leu Glu Ser Ala Asn Pro Lys Leu Tyr		
785	790	795 800
Ala Asn Tyr Ile Lys Tyr Ser Asp Glu Lys Lys Ala Glu Glu Phe Thr		
805	810	815
Arg Gln Ile Asn Arg Glu Lys Ala Lys Thr Ala Leu Asn Ala Tyr Leu		
820	825	830
Arg Asn Thr Lys Trp Asn Val Ile Ile Arg Glu Asp Leu Leu Arg Ile		
835	840	845
Asp Asn Lys Thr Cys Thr Leu Phe Ala Asn Lys Ala Val Ala Leu Glu		
850	855	860
Val Ala Arg Tyr Val His Ala Tyr Ile Asn Asp Ile Ala Glu Val Asn		
865	870	875 880
Ser Tyr Phe Gln Leu Tyr His Tyr Ile Met Gln Arg Ile Ile Met Asn		
885	890	895
Glu Arg Tyr Glu Lys Ser Ser Gly Lys Val Ser Glu Tyr Phe Asp Ala		
900	905	910
Val Asn Asp Glu Lys Lys Tyr Asn Asp Arg Leu Leu Lys Leu Leu Cys		
915	920	925
Val Pro Phe Gly Tyr Cys Ile Pro Arg Phe Lys Asn Leu Ser Ile Glu		
930	935	940
Ala Leu Phe Asp Arg Asn Glu Ala Ala Lys Phe Asp Lys Glu Lys Lys		
945	950	955 960
Lys Val Ser Gly Asn Ser		
965		

<210> 46

<211> 1034

<212> PRT

<213> 未知的

<220>

<223> Cas13d(重叠群e-k87 11092736)

<400> 46

Met Lys Arg Gln Lys Thr Phe Ala Lys Arg Ile Gly Ile Lys Ser Thr		
1	5	10 15
Val Ala Tyr Gly Gln Gly Lys Tyr Ala Ile Thr Thr Phe Gly Lys Gly		
20	25	30
Ser Lys Ala Glu Ile Ala Val Arg Ser Ala Asp Pro Pro Glu Glu Thr		
35	40	45

Leu Pro Thr Glu Ser Asp Ala Thr Leu Ser Ile His Ala Lys Phe Ala		
50	55	60
Lys Ala Gly Arg Asp Gly Arg Glu Phe Lys Cys Gly Asp Val Asp Glu		
65	70	75 80
Thr Arg Ile His Thr Ser Arg Ser Glu Tyr Glu Ser Leu Ile Ser Asn		
	85	90 95
Pro Ala Glu Ser Pro Arg Glu Asp Tyr Leu Gly Leu Lys Gly Thr Leu		
	100	105 110
Glu Arg Lys Phe Phe Gly Asp Glu Tyr Pro Lys Asp Asn Leu Arg Ile		
	115	120 125
Gln Ile Ile Tyr Ser Ile Leu Asp Ile Gln Lys Ile Leu Gly Leu Tyr		
	130	135 140
Val Glu Asp Ile Leu His Phe Val Asp Gly Leu Gln Asp Glu Pro Glu		
145	150	155 160
Asp Leu Val Gly Leu Gly Leu Gly Asp Glu Lys Met Gln Lys Leu Leu		
	165	170 175
Ser Lys Ala Leu Pro Tyr Met Gly Phe Phe Gly Ser Thr Asp Val Phe		
	180	185 190
Lys Val Thr Lys Lys Arg Glu Glu Arg Ala Ala Ala Asp Glu His Asn		
	195	200 205
Ala Lys Val Phe Arg Ala Leu Gly Ala Ile Arg Gln Lys Leu Ala His		
	210	215 220
Phe Lys Trp Lys Glu Ser Leu Ala Ile Phe Gly Ala Asn Ala Asn Met		
225	230	235 240
Pro Ile Arg Phe Phe Gln Gly Ala Thr Gly Gly Arg Gln Leu Trp Asn		
	245	250 255
Asp Val Ile Ala Pro Leu Trp Lys Lys Arg Ile Glu Arg Val Arg Lys		
	260	265 270
Ser Phe Leu Ser Asn Ser Ala Lys Asn Leu Trp Val Leu Tyr Gln Val		
	275	280 285
Phe Lys Asp Asp Thr Asp Glu Lys Lys Lys Ala Arg Ala Arg Gln Tyr		
	290	295 300
Tyr His Phe Ser Val Leu Lys Glu Gly Lys Asn Leu Gly Phe Asn Leu		
305	310	315 320
Thr Lys Thr Arg Glu Tyr Phe Leu Asp Lys Phe Phe Pro Ile Phe His		
	325	330 335
Ser Ser Ala Pro Asp Val Lys Arg Lys Val Asp Thr Phe Arg Ser Lys		
	340	345 350
Phe Tyr Ala Ile Leu Asp Phe Ile Ile Tyr Glu Ala Ser Val Ser Val		

355	360	365
Ala Asn Ser Gly Gln Met Gly Lys Val Ala Pro Trp Lys Gly Ala Ile		
370	375	380
Asp Asn Ala Leu Val Lys Leu Arg Glu Ala Pro Asp Glu Glu Ala Lys		
385	390	395
Glu Lys Ile Tyr Asn Val Leu Ala Ala Ser Ile Arg Asn Asp Ser Leu		
405	410	415
Phe Leu Arg Leu Lys Ser Ala Cys Asp Lys Phe Gly Ala Glu Gln Asn		
420	425	430
Arg Pro Val Phe Pro Asn Glu Leu Arg Asn Asn Arg Asp Ile Arg Asn		
435	440	445
Val Arg Ser Glu Trp Leu Glu Ala Thr Gln Asp Val Asp Ala Ala Ala		
450	455	460
Phe Val Gln Leu Ile Ala Phe Leu Cys Asn Phe Leu Glu Gly Lys Glu		
465	470	475
Ile Asn Glu Leu Val Thr Ala Leu Ile Lys Lys Phe Glu Gly Ile Gln		
485	490	495
Ala Leu Ile Asp Leu Leu Arg Asn Leu Glu Gly Val Asp Ser Ile Arg		
500	505	510
Phe Glu Asn Glu Phe Ala Leu Phe Asn Asp Asp Lys Gly Asn Met Ala		
515	520	525
Gly Arg Ile Ala Arg Gln Leu Arg Leu Leu Ala Ser Val Gly Lys Met		
530	535	540
Lys Pro Asp Met Thr Asp Ala Lys Arg Val Leu Tyr Lys Ser Ala Leu		
545	550	555
Glu Ile Leu Gly Ala Pro Pro Asp Glu Val Ser Asp Glu Trp Leu Ala		
565	570	575
Glu Asn Ile Leu Leu Asp Lys Ser Asn Asn Asp Tyr Gln Lys Ala Lys		
580	585	590
Lys Thr Val Asn Pro Phe Arg Asn Tyr Ile Ala Lys Asn Val Ile Thr		
595	600	605
Ser Arg Ser Phe Tyr Tyr Leu Val Arg Tyr Ala Lys Pro Thr Ala Val		
610	615	620
Arg Lys Leu Met Ser Asn Pro Lys Ile Val Arg Tyr Val Leu Lys Arg		
625	630	635
Leu Pro Glu Lys Gln Val Ala Ser Tyr Tyr Ser Ala Ile Trp Thr Gln		
645	650	655
Ser Glu Ser Asn Ser Asn Glu Met Val Lys Leu Ile Glu Met Ile Asp		
660	665	670

Arg Leu Thr Thr Glu Ile Ala Gly Phe Ser Phe Ala Val Leu Lys Asp		
675	680	685
Lys Lys Asp Ser Ile Val Ser Ala Ser Arg Glu Ser Arg Ala Val Asn		
690	695	700
Leu Glu Val Glu Arg Leu Lys Lys Leu Thr Thr Leu Tyr Met Ser Ile		
705	710	715
Ala Tyr Ile Ala Val Lys Ser Leu Val Lys Val Asn Ala Arg Tyr Phe		
	725	730
		735
Ile Ala Tyr Ser Ala Leu Glu Arg Asp Leu Tyr Phe Phe Asn Glu Lys		
	740	745
		750
Tyr Gly Glu Glu Phe Arg Leu His Phe Ile Pro Tyr Glu Leu Asn Gly		
	755	760
		765
Lys Thr Cys Gln Phe Glu Tyr Leu Ala Ile Leu Lys Tyr Tyr Leu Ala		
	770	775
		780
Arg Asp Glu Glu Thr Leu Lys Arg Lys Cys Glu Ile Cys Glu Glu Ile		
785	790	795
		800
Lys Val Gly Cys Glu Lys His Lys Lys Asn Ala Asn Pro Pro Tyr Glu		
	805	810
		815
Tyr Asp Gln Glu Trp Ile Asp Lys Lys Lys Ala Leu Asn Ser Glu Arg		
	820	825
		830
Lys Ala Cys Glu Arg Arg Leu His Phe Ser Thr His Trp Ala Gln Tyr		
	835	840
		845
Ala Thr Lys Arg Asp Glu Asn Met Ala Lys His Pro Gln Lys Trp Tyr		
	850	855
		860
Asp Ile Leu Ala Ser His Tyr Asp Glu Leu Leu Ala Leu Gln Ala Thr		
865	870	875
		880
Gly Trp Leu Ala Thr Gln Ala Arg Asn Asp Ala Glu His Leu Asn Pro		
	885	890
		895
Val Asn Glu Phe Asp Val Tyr Ile Glu Asp Leu Arg Arg Tyr Pro Glu		
	900	905
		910
Gly Thr Pro Lys Asn Lys Asp Tyr His Ile Gly Ser Tyr Phe Glu Ile		
	915	920
		925
Tyr His Tyr Ile Arg Gln Arg Ala Tyr Leu Glu Glu Val Leu Ala Lys		
	930	935
		940
Arg Lys Glu Tyr Arg Asp Ser Gly Ser Phe Thr Asp Glu Gln Leu Asp		
945	950	955
		960
Lys Leu Gln Lys Ile Leu Asp Asp Ile Arg Ala Arg Gly Ser Tyr Asp		
	965	970
		975
Lys Asn Leu Leu Lys Leu Glu Tyr Leu Pro Phe Ala Tyr Asn Leu Pro		

980	985	990
Arg Tyr Lys Asn Leu Thr Thr Glu Ala Leu Phe Asp Asp Asp Ser Val		
995	1000	1005
Ser Gly Lys Lys Arg Val Ala Glu Trp Arg Glu Arg Glu Lys Thr		
1010	1015	1020
Arg Glu Ala Glu Arg Glu Gln Arg Arg Gln Arg		
1025	1030	
<210> 47		
<211> 30		
<212> DNA		
<213> 未知的		
<220>		
<223> 重叠群e-k87 11092736		
<400> 47		
gtgagaagtc tccttatggg gagatgctac 30		
<210> 48		
<211> 1022		
<212> PRT		
<213> 未知的		
<220>		
<223> Cas13d(160582958 基因49834)		
<400> 48		
Met Lys Asn Ser Val Thr Phe Lys Leu Ile Gln Ala Gln Glu Asn Lys		
1 5 10 15		
Glu Ala Ala Arg Lys Lys Ala Lys Asp Ile Ala Glu Gln Ala Arg Ile		
20 25 30		
Ala Lys Arg Asn Gly Val Val Lys Lys Glu Glu Asn Arg Ile Asn Arg		
35 40 45		
Ile Gln Ile Glu Ile Gln Thr Gln Lys Lys Ser Asn Thr Gln Asn Ala		
50 55 60		
Tyr His Leu Lys Ser Leu Ala Lys Ala Ala Gly Val Lys Ser Val Phe		
65 70 75 80		
Ala Ile Gly Asn Asp Leu Leu Met Thr Gly Phe Gly Pro Gly Asn Asp		
85 90 95		
Ala Thr Ile Glu Lys Arg Val Phe Gln Asn Arg Ala Ile Glu Thr Leu		
100 105 110		
Ser Ser Pro Glu Gln Tyr Ser Ala Glu Phe Gln Asn Lys Gln Phe Lys		
115 120 125		
Ile Lys Gly Asn Ile Lys Val Leu Asn His Ser Thr Gln Lys Met Glu		

130	135	140
Glu Ile Gln Thr Glu Leu Gln Asp Asn Tyr Asn Arg Pro His Phe Asp		
145	150	155
Leu Leu Gly Cys Lys Asn Val Leu Glu Gln Lys Tyr Phe Gly Arg Thr		160
165	170	175
Phe Ser Asp Asn Ile His Val Gln Ile Ala Tyr Asn Ile Met Asp Ile		
180	185	190
Glu Lys Leu Leu Thr Pro Tyr Ile Asn Asn Ile Ile Tyr Thr Leu Asn		
195	200	205
Glu Leu Met Arg Asp Asn Ser Lys Asp Asp Phe Phe Gly Cys Asp Ser		
210	215	220
His Phe Ser Val Ala Tyr Leu Tyr Asp Glu Leu Lys Ala Gly Tyr Ser		
225	230	235
Asp Arg Leu Lys Thr Lys Pro Asn Leu Ser Lys Asn Ile Asp Arg Ile		
245	250	255
Trp Asn Asn Phe Cys Asn Tyr Met Asn Ser Asp Ser Gly Asn Thr Glu		
260	265	270
Ala Arg Leu Ala Tyr Phe Gly Glu Leu Phe Tyr Lys Pro Lys Glu Thr		
275	280	285
Gly Asp Ala Lys Ser Asp Tyr Lys Thr His Leu Ser Asn Asn Gln Lys		
290	295	300
Glu Glu Trp Glu Leu Lys Ser Asp Lys Glu Val Tyr Asn Ile Phe Ala		
305	310	315
Ile Leu Cys Asp Leu Arg His Phe Cys Thr His Gly Glu Ser Ile Thr		
325	330	335
Pro Ser Gly Lys Pro Phe Pro Tyr Asn Leu Glu Lys Asn Leu Phe Pro		
340	345	350
Glu Ala Lys Gln Val Leu Asn Ser Leu Phe Glu Glu Lys Ala Glu Ser		
355	360	365
Leu Gly Ala Glu Ala Phe Gly Lys Thr Ala Gly Lys Thr Asp Val Ser		
370	375	380
Ile Leu Leu Lys Val Phe Glu Lys Glu Gln Ala Ser Gln Lys Glu Gln		
385	390	395
Gln Ala Leu Leu Lys Glu Tyr Tyr Asp Phe Lys Val Gln Lys Thr Tyr		
405	410	415
Lys Asn Met Gly Phe Ser Ile Lys Lys Leu Arg Glu Ala Ile Met Glu		
420	425	430
Ile Pro Asp Ala Ala Lys Phe Lys Asp Asp Leu Tyr Ser Ser Leu Arg		
435	440	445

His	Lys	Leu	Tyr	Gly	Leu	Phe	Asp	Phe	Ile	Leu	Val	Lys	His	Phe	Leu
450							455					460			
Asp	Thr	Ser	Asp	Ser	Glu	Asn	Leu	Gln	Asn	Asn	Asp	Ile	Phe	Arg	Gln
465					470				475					480	
Leu	Arg	Ala	Cys	Arg	Cys	Glu	Glu	Glu	Lys	Asp	Gln	Val	Tyr	Arg	Ser
				485					490					495	
Ile	Ala	Val	Lys	Val	Trp	Glu	Lys	Val	Lys	Lys	Lys	Glu	Leu	Asn	Met
			500					505						510	
Phe	Lys	Gln	Val	Val	Val	Ile	Pro	Ser	Leu	Ser	Lys	Asp	Glu	Leu	Lys
		515					520					525			
Gln	Met	Glu	Met	Thr	Lys	Asn	Thr	Glu	Leu	Leu	Ser	Ser	Ile	Glu	Thr
	530					535					540				
Ile	Ser	Thr	Gln	Ala	Ser	Leu	Phe	Ser	Glu	Met	Ile	Phe	Met	Met	Thr
545					550					555					560
Tyr	Leu	Leu	Asp	Gly	Lys	Glu	Ile	Asn	Leu	Leu	Cys	Thr	Ser	Leu	Ile
				565					570					575	
Glu	Lys	Phe	Glu	Asn	Ile	Ala	Ser	Phe	Asn	Glu	Val	Leu	Lys	Ser	Pro
			580					585						590	
Gln	Ile	Gly	Tyr	Glu	Thr	Lys	Tyr	Thr	Glu	Gly	Tyr	Ala	Phe	Phe	Lys
	595						600					605			
Asn	Ala	Asp	Lys	Thr	Ala	Lys	Glu	Leu	Arg	Gln	Val	Asn	Asn	Met	Ala
610						615						620			
Arg	Met	Thr	Lys	Pro	Leu	Gly	Gly	Val	Asn	Thr	Lys	Cys	Val	Met	Tyr
625					630					635					640
Asn	Glu	Ala	Ala	Lys	Ile	Leu	Gly	Ala	Lys	Pro	Met	Ser	Lys	Ala	Glu
				645					650					655	
Leu	Glu	Ser	Val	Phe	Asn	Leu	Asp	Asn	His	Asp	Tyr	Thr	Tyr	Ser	Pro
			660					665						670	
Ser	Gly	Lys	Lys	Ile	Pro	Asn	Lys	Asn	Phe	Arg	Asn	Phe	Ile	Ile	Asn
		675				680						685			
Asn	Val	Ile	Thr	Ser	Arg	Arg	Phe	Leu	Tyr	Leu	Ile	Arg	Tyr	Gly	Asn
690						695					700				
Pro	Glu	Lys	Ile	Arg	Lys	Ile	Ala	Ile	Asn	Pro	Ser	Ile	Ile	Ser	Phe
705					710					715					720
Val	Leu	Lys	Gln	Ile	Pro	Asp	Glu	Gln	Ile	Lys	Arg	Tyr	Tyr	Pro	Pro
				725					730					735	
Cys	Ile	Gly	Lys	Arg	Thr	Asp	Asp	Val	Thr	Leu	Met	Arg	Asp	Glu	Leu
			740					745						750	
Gly	Lys	Met	Leu	Gln	Ser	Val	Asn	Phe	Glu	Gln	Phe	Ser	Arg	Val	Asn

755	760	765
Asn Lys Gln Asn Ala Lys Gln Asn Pro Asn Gly Glu Lys Ala Arg Leu		
770	775	780
Gln Ala Cys Val Arg Leu Tyr Leu Thr Val Pro Tyr Leu Phe Ile Lys		
785	790	795
Asn Met Val Asn Ile Asn Ala Arg Tyr Val Leu Ala Phe His Cys Leu		800
805	810	815
Glu Arg Asp His Ala Leu Cys Phe Asn Ser Arg Lys Leu Asn Asp Asp		
820	825	830
Ser Tyr Asn Glu Met Ala Asn Lys Phe Gln Met Val Arg Lys Ala Lys		
835	840	845
Lys Glu Gln Tyr Glu Lys Glu Tyr Lys Cys Lys Lys Gln Glu Thr Gly		
850	855	860
Thr Ala His Thr Lys Lys Ile Glu Lys Leu Asn Gln Gln Ile Ala Tyr		
865	870	875
Ile Asp Lys Asp Ile Lys Asn Met His Ser Tyr Thr Cys Arg Asn Tyr		
885	890	895
Arg Asn Leu Val Ala His Leu Asn Val Val Ser Lys Leu Gln Asn Tyr		
900	905	910
Val Ser Glu Leu Pro Asn Asp Tyr Gln Ile Thr Ser Tyr Phe Ser Phe		
915	920	925
Tyr His Tyr Cys Met Gln Leu Gly Leu Met Glu Lys Val Ser Ser Lys		
930	935	940
Asn Ile Pro Leu Val Glu Ser Leu Lys Asn Glu Ala Asn Asp Ala Gln		
945	950	955
Ser Tyr Ser Ala Lys Lys Thr Leu Glu Tyr Phe Asp Leu Ile Glu Lys		
965	970	975
Asn Arg Thr Tyr Cys Lys Asp Phe Leu Lys Ala Leu Asn Ala Pro Phe		
980	985	990
Ser Tyr Asn Leu Pro Arg Phe Lys Asn Leu Ser Ile Glu Ala Leu Phe		
995	1000	1005
Asp Lys Asn Ile Val Tyr Glu Gln Ala Asp Leu Lys Lys Glu		
1010	1015	1020

<210> 49

<211> 36

<212> DNA

<213> 未知的

<220>

<223> Cas13d(160582958 基因49834)

<400> 49

gaactacacc cctctgttct ttaggggtc taacac 36

<210> 50

<211> 877

<212> PRT

<213> 未知的

<220>

<223> Cas13d(重叠群tpg DJXD01000002.1)

<400> 50

Met	Lys	Lys	Gln	Lys	Ser	Lys	Lys	Thr	Val	Ser	Lys	Thr	Ser	Gly	Leu
1				5					10					15	
Lys	Glu	Ala	Leu	Ser	Val	Gln	Gly	Thr	Val	Ile	Met	Thr	Ser	Phe	Gly
			20					25					30		
Lys	Gly	Asn	Met	Ala	Asn	Leu	Ser	Tyr	Lys	Ile	Pro	Ser	Ser	Gln	Lys
		35				40					45				
Pro	Gln	Asn	Leu	Asn	Ser	Ser	Ala	Gly	Leu	Lys	Asn	Val	Glu	Val	Ser
		50				55					60				
Gly	Lys	Lys	Ile	Lys	Phe	Gln	Gly	Arg	His	Pro	Lys	Ile	Ala	Thr	Thr
65				70				75					80		
Asp	Asn	Pro	Leu	Phe	Lys	Pro	Gln	Pro	Gly	Met	Asp	Leu	Leu	Cys	Leu
			85					90					95		
Lys	Asp	Lys	Leu	Glu	Met	His	Tyr	Phe	Gly	Lys	Thr	Phe	Asp	Asp	Asn
		100						105					110		
Ile	His	Ile	Gln	Leu	Ile	Tyr	Gln	Ile	Leu	Asp	Ile	Glu	Lys	Ile	Leu
		115						120					125		
Ala	Val	His	Val	Asn	Asn	Ile	Val	Phe	Thr	Leu	Asp	Asn	Val	Leu	His
		130				135					140				
Pro	Gln	Lys	Glu	Glu	Leu	Thr	Glu	Asp	Phe	Ile	Gly	Ala	Gly	Gly	Trp
145				150				155					160		
Arg	Ile	Asn	Leu	Asp	Tyr	Gln	Thr	Leu	Arg	Gly	Gln	Thr	Asn	Lys	Tyr
			165					170					175		
Asp	Arg	Phe	Lys	Asn	Tyr	Ile	Lys	Arg	Lys	Glu	Leu	Leu	Tyr	Phe	Gly
		180						185					190		
Glu	Ala	Phe	Tyr	His	Glu	Asn	Glu	Arg	Arg	Tyr	Glu	Glu	Asp	Ile	Phe
		195				200							205		
Ala	Ile	Leu	Thr	Leu	Leu	Ser	Ala	Leu	Arg	Gln	Phe	Cys	Phe	His	Ser
		210				215					220				
Asp	Leu	Ser	Ser	Asp	Glu	Ser	Asp	His	Val	Asn	Ser	Phe	Trp	Leu	Tyr
225				230				235					240		

Gln Leu Glu Asp Gln Leu Ser Asp Glu Phe Lys Glu Thr Leu Ser Ile		
245	250	255
Leu Trp Glu Glu Val Thr Glu Arg Ile Asp Ser Glu Phe Leu Lys Thr		
260	265	270
Asn Thr Val Asn Leu His Ile Leu Cys His Val Phe Pro Lys Glu Ser		
275	280	285
Lys Glu Thr Ile Val Arg Ala Tyr Tyr Glu Phe Leu Ile Lys Lys Ser		
290	295	300
Phe Lys Asn Met Gly Phe Ser Ile Lys Lys Leu Arg Glu Ile Met Leu		
305	310	315
Glu Gln Ser Asp Leu Lys Ser Phe Lys Glu Asp Lys Tyr Asn Ser Val		
325	330	335
Arg Ala Lys Leu Tyr Lys Leu Phe Asp Phe Ile Ile Thr Tyr Tyr Tyr		
340	345	350
Asp His His Ala Phe Glu Lys Glu Ala Leu Val Ser Ser Leu Arg Ser		
355	360	365
Ser Leu Thr Glu Glu Asn Lys Glu Glu Ile Tyr Ile Lys Thr Ala Arg		
370	375	380
Thr Leu Ala Ser Ala Leu Gly Ala Asp Phe Lys Lys Ala Ala Ala Asp		
385	390	395
Val Asn Ala Lys Asn Ile Arg Asp Tyr Gln Lys Lys Ala Asn Asp Tyr		
405	410	415
Arg Ile Ser Phe Glu Asp Ile Lys Ile Gly Asn Thr Gly Ile Gly Tyr		
420	425	430
Phe Ser Glu Leu Ile Tyr Met Leu Thr Leu Leu Leu Asp Gly Lys Glu		
435	440	445
Ile Asn Asp Leu Leu Thr Thr Leu Ile Asn Lys Phe Asp Asn Ile Ile		
450	455	460
Ser Phe Ile Asp Ile Leu Lys Lys Leu Asn Leu Glu Phe Lys Phe Lys		
465	470	475
Pro Glu Tyr Ala Asp Phe Phe Asn Met Thr Asn Cys Arg Tyr Thr Leu		
485	490	495
Glu Glu Leu Arg Val Ile Asn Ser Ile Ala Arg Met Gln Lys Pro Ser		
500	505	510
Ala Asp Ala Arg Lys Ile Met Tyr Arg Asp Ala Leu Arg Ile Leu Gly		
515	520	525
Met Asp Asn Arg Pro Asp Glu Glu Ile Asp Arg Glu Leu Glu Arg Thr		
530	535	540
Met Pro Val Gly Ala Asp Gly Lys Phe Ile Lys Gly Lys Gln Gly Phe		

545	550	555	560
Arg Asn Phe Ile Ala Ser Asn Val Ile Glu Ser Ser Arg Phe His Tyr			
	565	570	575
Leu Val Arg Tyr Asn Asn Pro His Lys Thr Arg Thr Leu Val Lys Asn			
	580	585	590
Pro Asn Val Val Lys Phe Val Leu Glu Gly Ile Pro Glu Thr Gln Ile			
	595	600	605
Lys Arg Tyr Phe Asp Val Cys Lys Gly Gln Glu Ile Pro Pro Thr Ser			
	610	615	620
Asp Lys Ser Ala Gln Ile Asp Val Leu Ala Arg Ile Ile Ser Ser Val			
625	630	635	640
Asp Tyr Lys Ile Phe Glu Asp Val Pro Gln Ser Ala Lys Ile Asn Lys			
	645	650	655
Asp Asp Pro Ser Arg Asn Phe Ser Asp Ala Leu Lys Lys Gln Arg Tyr			
	660	665	670
Gln Ala Ile Val Ser Leu Tyr Leu Thr Val Met Tyr Leu Ile Thr Lys			
	675	680	685
Asn Leu Val Tyr Val Asn Ser Arg Tyr Val Ile Ala Phe His Cys Leu			
	690	695	700
Glu Arg Asp Ala Phe Leu His Gly Val Thr Leu Pro Lys Met Asn Lys			
705	710	715	720
Lys Ile Val Tyr Ser Gln Leu Thr Thr His Leu Leu Thr Asp Lys Asn			
	725	730	735
Tyr Thr Thr Tyr Gly His Leu Lys Asn Gln Lys Gly His Arg Lys Trp			
	740	745	750
Tyr Val Leu Val Lys Asn Asn Leu Gln Asn Ser Asp Ile Thr Ala Val			
	755	760	765
Ser Ser Phe Arg Asn Ile Val Ala His Ile Ser Val Val Arg Asn Ser			
	770	775	780
Asn Glu Tyr Ile Ser Gly Ile Gly Glu Leu His Ser Tyr Phe Glu Leu			
785	790	795	800
Tyr His Tyr Leu Val Gln Ser Met Ile Ala Lys Asn Asn Trp Tyr Asp			
	805	810	815
Thr Ser His Gln Pro Lys Thr Ala Glu Tyr Leu Asn Asn Leu Lys Lys			
	820	825	830
His His Thr Tyr Cys Lys Asp Phe Val Lys Ala Tyr Cys Ile Pro Phe			
	835	840	845
Gly Tyr Val Val Pro Arg Tyr Lys Asn Leu Thr Ile Asn Glu Leu Phe			
	850	855	860

Asp Arg Asn Asn Pro Asn Pro Glu Pro Lys Glu Glu Val
 865 870 875
 <210> 51
 <211> 36
 <212> PRT
 <213> 智人(Homo Sapiens)
 <400> 51
 Cys Ala Ala Cys Thr Ala Cys Ala Ala Cys Cys Cys Cys Gly Thr Ala
 1 5 10 15
 Ala Ala Ala Ala Thr Ala Cys Gly Gly Gly Gly Thr Thr Cys Thr Gly
 20 25 30
 Ala Ala Ala Cys
 35
 <210> 52
 <400> 52
 000
 <210> 53
 <400> 53
 000
 <210> 54
 <211> 124
 <212> PRT
 <213> 未知的
 <220>
 <223> CasRX/Cas13d肠道宏基因组重叠群6049000251
 <400> 54
 Leu Tyr Leu Thr Ser Phe Gly Lys Gly Asn Ala Ala Val Ile Glu Gln
 1 5 10 15
 Lys Ile Glu Pro Glu Asn Gly Tyr Arg Val Thr Gly Met Gln Ile Thr
 20 25 30
 Pro Ser Ile Thr Val Asn Lys Ala Thr Asp Glu Ser Val Arg Phe Arg
 35 40 45
 Val Lys Arg Lys Ile Ala Gln Lys Asp Glu Phe Ile Ala Asp Asn Pro
 50 55 60
 Met His Glu Gly Arg His Arg Ile Glu Pro Ser Ala Gly Ser Asp Met
 65 70 75 80
 Leu Gly Leu Lys Thr Lys Leu Glu Lys Tyr Tyr Phe Gly Lys Glu Phe
 85 90 95
 Asp Asp Asn Leu His Ile Gln Ile Ile Tyr Asn Ile Leu Asp Ile Glu

	100		105		110
Lys Ile Leu Ala Val Tyr Ser Thr Asn Ile Thr Ala					
	115		120		
<210> 55					
<400> 55					
000					
<210> 56					
<400> 56					
000					
<210> 57					
<211> 358					
<212> PRT					
<213> 未知的					
<220>					
<223> CasRX/Cas13d肠道宏基因组重叠群546000275					
<400> 57					
Met Asp Ser Tyr Arg Pro Lys Leu Tyr Lys Leu Ile Asp Phe Cys Ile					
1	5		10		15
Phe Lys His Tyr His Glu Tyr Thr Glu Ile Ser Glu Lys Asn Val Asp					
	20		25		30
Thr Leu Arg Ala Ala Val Ser Glu Glu Gln Lys Glu Ser Phe Tyr Ala					
	35		40		45
Asp Glu Ala Lys Arg Leu Trp Gly Ile Phe Asp Lys Gln Phe Leu Gly					
	50		55		60
Phe Cys Lys Lys Ile Asn Val Trp Val Asn Gly Ser His Glu Lys Glu					
65		70		75	
Ile Leu Gly Tyr Ile Asp Lys Asp Ala Tyr Arg Lys Lys Ser Asp Val					
	85		90		95
Ser Tyr Phe Ser Lys Phe Leu Tyr Ala Met Ser Phe Phe Leu Asp Gly					
	100		105		110
Lys Glu Ile Asn Asp Leu Leu Thr Thr Leu Ile Asn Lys Phe Asp Asn					
	115		120		125
Ile Ala Ser Phe Ile Ser Thr Ala Lys Glu Leu Asp Ala Glu Ile Asp					
	130		135		140
Arg Ile Leu Glu Lys Lys Leu Asp Pro Val Thr Gly Lys Pro Leu Lys					
145		150		155	
Gly Lys Asn Ser Phe Arg Asn Phe Ile Ala Asn Asn Val Ile Glu Asn					
	165		170		175
Lys Arg Phe Ile Tyr Val Ile Lys Phe Cys Asn Pro Lys Asn Val Leu					

180	185	190
Lys Leu Val Lys Asn Thr Lys Val Thr Glu Phe Val Leu Lys Arg Met		
195	200	205
Pro Glu Ser Gln Ile Asp Arg Tyr Tyr Ser Ser Cys Ile Asp Thr Glu		
210	215	220
Lys Asn Pro Ser Val Asp Lys Lys Ile Ser Asp Leu Ala Glu Met Ile		
225	230	235
Lys Lys Ile Ala Phe Asp Asp Phe Arg Asn Val Arg Gln Lys Thr Arg		
245	250	255
Thr Arg Glu Glu Ser Leu Glu Lys Glu Arg Phe Lys Ala Val Ile Gly		
260	265	270
Leu Tyr Leu Thr Val Val Tyr Leu Leu Ile Lys Asn Leu Val Asn Val		
275	280	285
Asn Ser Arg Tyr Val Met Ala Phe His Cys Leu Glu Arg Asp Ala Lys		
290	295	300
Leu Tyr Gly Ile Asn Ile Gly Lys Asn Tyr Ile Glu Leu Thr Glu Asp		
305	310	315
Leu Cys Arg Glu Asn Glu Asn Ser Arg Ser Ala Tyr Leu Ala Arg Asn		
325	330	335
Lys Arg Leu Arg Asp Cys Val Lys Gln Asn Ile Asp Asn Ala Lys Asn		
340	345	350
Met Lys Ser Lys Glu Lys		
355		

<210> 58

<400> 58

000

<210> 59

<400> 59

000

<210> 60

<400> 60

000

<210> 61

<211> 149

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群4114000374

<400> 61

50 55 60

Gln	Pro	Ala	Glu	Glu	Glu	Tyr	Phe	Ile	Asn	Gln	Leu	Lys	Ile	Glu	Gly
65					70				75					80	
Val	Glu	His	Leu	Pro	Arg	Arg	Val	Thr	Asn	Glu	Ile	Thr	Ser	Asn	Ser
			85						90					95	
Arg	Tyr	Glu	Val	Tyr	Thr	Gln	Gly	Gly	Val	Ile	Phe	Ala	Thr	Ser	Arg
			100					105					110		
Ile	Leu	Val	Val	Asp	Phe	Leu	Thr	Asp	Arg	Ile	Pro	Ser	Asp	Leu	Ile
			115					120					125		
Thr	Gly	Ile	Leu	Val	Tyr	Arg	Ala	His	Arg	Ile	Ile	Glu	Ser	Cys	Gln
			130				135					140			
Glu	Ala	Phe	Ile	Leu	Arg	Leu	Phe	Arg	Gln	Lys	Asn	Lys	Arg	Gly	Phe
145					150					155				160	
Ile	Lys	Ala	Phe	Thr	Asp	Asn	Ala	Val	Ala	Phe	Asp	Thr	Gly	Phe	Cys
					165					170				175	
His	Val	Glu	Arg	Val	Met	Arg	Asn	Leu	Phe	Val	Arg	Lys	Leu	Tyr	Leu
			180					185					190		
Trp	Pro	Arg	Phe	His	Val	Ala	Val	Asn	Ser	Phe	Leu	Glu	Gln	His	Lys
			195					200					205		
Pro	Glu	Val	Val	Glu	Ile	His	Val	Ser	Met	Thr	Pro	Thr	Met	Leu	Ala
			210				215					220			
Ile	Gln	Thr	Ala	Ile	Leu	Asp	Ile	Leu	Asn	Ala	Cys	Leu	Lys	Glu	Leu
225					230					235				240	
Lys	Cys	His	Asn	Pro	Ser	Leu	Glu	Val	Glu	Asp	Leu	Ser	Leu	Glu	Asn
					245					250				255	
Ala	Ile	Gly	Lys	Pro	Phe	Asp	Lys	Thr	Ile	Arg	His	Tyr	Leu	Asp	Pro
			260					265					270		
Leu	Trp	His	Gln	Leu	Gly	Ala	Lys	Thr	Lys	Ser	Leu	Val	Gln	Asp	Leu
			275					280					285		
Lys	Ile	Leu	Arg	Thr	Leu	Leu	Gln	Tyr	Leu	Ser	Gln	Tyr	Asp	Cys	Val
			290				295				300				
Thr	Phe	Leu	Asn	Leu	Leu	Glu	Ser	Leu	Arg	Ala	Thr	Glu	Lys	Ala	Phe
305					310					315				320	
Gly	Gln	Asn	Ser	Gly	Trp	Leu	Phe	Leu	Asp	Ser	Ser	Thr	Ser	Met	Phe
					325					330				335	
Ile	Asn	Ala	Arg	Ala	Arg	Val	Tyr	His	Leu	Pro	Asp	Ala	Lys	Met	Ser
			340					345					350		
Lys	Lys	Glu	Lys	Ile	Ser	Glu	Lys	Met	Glu	Ile	Lys	Glu	Gly	Glu	Gly
			355					360					365		
Ile	Leu	Trp	Gly												

370
 <210> 65
 <400> 65
 000
 <210> 66
 <400> 66
 000
 <210> 67
 <211> 320
 <212> PRT
 <213> 未知的
 <220>
 <223> CasRX/Cas13d肠道宏基因组重叠群721000619
 <400> 67
 Lys Glu Gly Ser Thr Met Ala Lys Asn Glu Lys Lys Lys Ser Thr Ala
 1 5 10 15
 Lys Ala Leu Gly Leu Lys Ser Ser Phe Val Val Asn Asn Asp Ile Tyr
 20 25 30
 Met Thr Ser Phe Gly Lys Gly Asn Lys Ala Val Leu Glu Lys Lys Ile
 35 40 45
 Thr Glu Asn Thr Ile Glu Asn Lys Ser Asp Thr Thr Tyr Phe Asp Val
 50 55 60
 Ile Asn Arg Asp Pro Lys Gly Phe Thr Leu Glu Gly Arg Arg Ile Ala
 65 70 75 80
 Asp Met Thr Ala Phe Ser Asn Asp Pro Lys Tyr His Val Asn Val Val
 85 90 95
 Asn Gly Lys Phe Leu Glu Asp Gln Leu Gly Ala Arg Ser Glu Leu Glu
 100 105 110
 Lys Lys Val Phe Gly Arg Thr Phe Asp Asp Asn Val His Ile Gln Leu
 115 120 125
 Ile His Asn Ile Leu Asp Ile Glu Lys Ile Met Ala Gln Tyr Val Ser
 130 135 140
 Asp Ile Val Tyr Leu Leu His Asn Thr Ile Lys Arg Asp Met Asn Asp
 145 150 155 160
 Asp Ile Met Gly Tyr Ile Ser Ile Arg Asn Ser Phe Asp Asp Phe Cys
 165 170 175
 His Pro Glu Arg Ile Pro Asp Arg Lys Ala Lys Asp Asn Leu Gln Lys
 180 185 190
 Gln His Asp Ile Phe Phe Asp Glu Ile Leu Lys Cys Gly Arg Leu Ala

195	200	205
Tyr Phe Gly Asn Ala Phe Phe Glu Asp Gly Ser Asp Asn Lys Glu Ile		
210	215	220
Ala Lys Leu Lys Arg Tyr Lys Glu Ile Tyr His Ile Ile Ala Leu Met		
225	230	235
Gly Ser Leu Arg Gln Ser Tyr Phe His Gly Glu Asn Ser Asp Lys Asn		
245	250	255
Phe Gln Gly Pro Thr Trp Ala Tyr Thr Leu Glu Ser Asn Leu Thr Gly		
260	265	270
Lys Tyr Lys Glu Phe Lys Asp Thr Leu Asp Lys Thr Phe Asp Glu Arg		
275	280	285
Tyr Glu Met Ile Ser Lys Asp Phe Gly Ser Thr Asn Met Val Asn Leu		
290	295	300
Gln Ile Leu Glu Glu Leu Leu Lys Met Leu Tyr Gly Asn Val Ser Pro		
305	310	315
320		
<210> 68		
<400> 68		
000		
<210> 69		
<211> 204		
<212> PRT		
<213> 未知的		
<220>		
<223> CasRX/Cas13d肠道宏基因组重叠群2002000411		
<400> 69		
Glu Lys Gln Asn Lys Ala Lys Tyr Gln Ala Ile Ile Ser Leu Tyr Leu		
1	5	10
15		
Met Val Met Tyr Gln Ile Val Lys Asn Met Ile Tyr Val Asn Ser Arg		
20	25	30
Tyr Val Ile Ala Phe His Cys Leu Glu Arg Asp Ser Asn Gln Leu Leu		
35	40	45
Gly Arg Phe Asn Ser Arg Asp Ala Ser Met Tyr Asn Lys Leu Thr Gln		
50	55	60
Lys Phe Ile Thr Asp Lys Tyr Leu Asn Asp Gly Ala Gln Gly Cys Ser		
65	70	75
80		
Lys Lys Val Gly Asn Tyr Leu Ser His Asn Ile Thr Cys Cys Ser Asp		
85	90	95
Glu Leu Arg Lys Glu Tyr Arg Asn Gln Val Asp His Phe Ala Val Val		
100	105	110

Arg Met Ile Gly Lys Tyr Ala Ala Asp Ile Gly Lys Phe Ser Thr Trp		
115	120	125
Phe Glu Leu Tyr His Tyr Val Met Gln Arg Ile Ile Phe Asp Lys Arg		
130	135	140
Asn Pro Leu Ser Glu Thr Glu Arg Thr Tyr Lys Gln Leu Ile Ala Lys		
145	150	155
His His Thr Tyr Cys Lys Asp Leu Val Lys Ala Leu Asn Thr Pro Phe		
165	170	175
Gly Tyr Asn Leu Ala Arg Tyr Lys Asn Leu Ser Ile Gly Glu Leu Phe		
180	185	190
Asp Arg Asn Asn Tyr Asn Ala Lys Thr Lys Glu Thr		
195	200	
<210> 70		
<400> 70		
000		
<210> 71		
<211> 449		
<212> PRT		
<213> 未知的		
<220>		
<223> CasRX/Cas13d肠道宏基因组重叠群13552000311		
<400> 71		
Leu Ile Asp Phe Leu Ile Tyr Asp Leu Tyr Tyr Asn Arg Lys Pro Ala		
1	5	10
Arg Ile Glu Glu Ile Val Asp Lys Leu Arg Glu Ser Val Asn Asp Glu		
20	25	30
Glu Lys Glu Ser Ile Tyr Ser Ala Glu Thr Lys Tyr Val Tyr Glu Ala		
35	40	45
Leu Gly Lys Val Leu Val Arg Ser Leu Lys Lys Tyr Leu Asn Gly Ala		
50	55	60
Thr Ile Arg Asp Leu Lys Asn Arg Tyr Asp Ala Lys Thr Ala Asn Arg		
65	70	75
Ile Trp Asp Ile Ser Glu His Ser Lys Ser Gly His Val Asn Cys Phe		
85	90	95
Cys Lys Leu Ile Tyr Met Met Thr Leu Met Leu Asp Gly Lys Glu Ile		
100	105	110
Asn Asp Leu Leu Thr Thr Leu Val Asn Lys Phe Asp Asn Ile Ala Ser		
115	120	125
Phe Ile Asp Val Met Asp Glu Leu Gly Leu Glu His Ser Phe Thr Asp		

130	135	140
Asn Tyr Lys Met Phe Ala Asp Ser Lys Ala Ile Cys Leu Asp Leu Gln		
145	150	155
Phe Ile Asn Ser Phe Ala Arg Met Ser Lys Ile Asp Asp Glu Lys Ser		
165	170	175
Lys Arg Gln Leu Phe Arg Asp Ala Leu Val Val Leu Asp Ile Gly Asp		
180	185	190
Lys Asn Glu Asp Trp Ile Glu Lys Tyr Leu Thr Ser Asp Ile Phe Lys		
195	200	205
Arg Asp Glu Asn Gly Asn Lys Ile Asp Gly Glu Lys Arg Asp Phe Arg		
210	215	220
Asn Phe Ile Ala Asn Asn Val Ile Lys Ser Ala Arg Phe Lys Tyr Leu		
225	230	235
Val Lys Tyr Ser Ser Ala Asp Gly Met Ile Lys Leu Lys Lys Asn Glu		
245	250	255
Lys Leu Ile Ser Phe Val Leu Glu Gln Leu Pro Glu Thr Gln Ile Asp		
260	265	270
Arg Tyr Tyr Glu Ser Cys Gly Leu Asp Cys Ala Val Ala Asp Arg Lys		
275	280	285
Val Arg Ile Glu Lys Leu Thr Gly Leu Ile Arg Asp Met Arg Phe Asp		
290	295	300
Asn Phe Arg Gly Val Asn Tyr Ser Asn Asp Ala Cys Lys Lys Asp Lys		
305	310	315
Gln Ala Lys Ala Lys Tyr Gln Ala Ile Ile Ser Leu Tyr Leu Met Val		
325	330	335
Leu Tyr Gln Ile Val Lys Asn Met Ile Tyr Val Asn Ser Arg Tyr Val		
340	345	350
Ile Ala Phe His Cys Leu Glu Arg Asp Leu Leu Phe Phe Asn Ile Glu		
355	360	365
Leu Asp Asn Ser Tyr Gln Tyr Ser Asn Cys Asn Glu Leu Thr Glu Lys		
370	375	380
Phe Ile Lys Asp Lys Tyr Met Lys Glu Gly Ala Leu Gly Phe Asn Met		
385	390	395
Lys Ala Gly Arg Tyr Leu Thr Lys Asn Ile Gly Asn Cys Ser Asn Glu		
405	410	415
Leu Arg Lys Ile Tyr Arg Asn Gln Val Asp His Phe Ala Val Val Arg		
420	425	430
Lys Ile Gly Asn Tyr Ala Ala Asp Ile Ala Ser Val Gly Ser Trp Phe		
435	440	445

Glu

<210> 72

<211> 96

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群10037000527

<400> 72

Tyr Met Asp Gln Asn Phe Ala Asn Ser Asp Ala Trp Ala Ile His Val

1 5 10 15

Tyr Arg Asn Lys Ile Gln His Leu Asp Ala Val Arg His Ala Asp Met

20 25 30

Tyr Ile Gly Asp Ile Arg Glu Phe His Ser Trp Phe Glu Leu Tyr His

35 40 45

Tyr Ile Ile Gln Arg Arg Ile Ile Asp Gln Tyr Ala Tyr Glu Ser Thr

50 55 60

Pro Gly Ser Ser Arg Asp Gly Ser Ala Ile Ile Asp Glu Glu Arg Leu

65 70 75 80

Asn Pro Ala Thr Arg Arg Tyr Phe Arg Leu Ile Thr Thr Tyr Lys Thr

85 90 95

<210> 73

<211> 519

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群238000329

<400> 73

Arg Tyr Asp Lys Asp Arg Ser Lys Ile Tyr Thr Met Met Asp Phe Val

1 5 10 15

Ile Tyr Arg Tyr Tyr Ile Asp Asn Asn Asn Asp Ser Ile Asp Phe Ile

20 25 30

Asn Lys Leu Arg Ser Ser Ile Asp Glu Lys Ser Lys Glu Lys Leu Tyr

35 40 45

Asn Glu Glu Ala Asn Arg Leu Trp Asn Lys Leu Lys Glu Tyr Met Leu

50 55 60

Tyr Ile Lys Glu Phe Asn Gly Lys Leu Ala Ser Arg Thr Pro Asp Arg

65 70 75 80

Asp Gly Asn Ile Ser Glu Phe Val Glu Ser Leu Pro Lys Ile His Arg

85 90 95

Leu	Leu	Pro	Arg	Gly	Gln	Lys	Ile	Ser	Asn	Phe	Ser	Lys	Leu	Met	Tyr
100				105				110							
Leu	Leu	Thr	Met	Phe	Leu	Asp	Gly	Lys	Glu	Ile	Asn	Asp	Leu	Leu	Thr
115				120				125							
Thr	Leu	Ile	Asn	Lys	Phe	Glu	Asn	Ile	Gln	Gly	Phe	Leu	Asp	Ile	Met
130				135				140							
Pro	Glu	Ile	Asn	Val	Asn	Ala	Lys	Phe	Glu	Pro	Glu	Tyr	Val	Phe	Phe
145				150				155				160			
Asn	Lys	Ser	His	Glu	Ile	Ala	Gly	Glu	Leu	Lys	Leu	Ile	Lys	Gly	Phe
165				170				175							
Ala	Gln	Met	Gly	Glu	Pro	Ala	Ala	Thr	Leu	Lys	Leu	Glu	Met	Thr	Ala
180				185				190							
Asp	Ala	Ile	Lys	Ile	Leu	Gly	Thr	Glu	Lys	Glu	Asp	Ala	Glu	Leu	Ile
195				200				205							
Lys	Leu	Ala	Glu	Ser	Leu	Phe	Lys	Asp	Glu	Asn	Gly	Lys	Leu	Leu	Gly
210				215				220							
Asn	Lys	Gln	His	Gly	Met	Arg	Asn	Phe	Ile	Gly	Asn	Asn	Val	Ile	Lys
225				230				235				240			
Ser	Lys	Arg	Phe	His	Tyr	Leu	Ile	Arg	Tyr	Gly	Asp	Pro	Ala	His	Leu
245				250				255							
His	Lys	Ile	Ala	Thr	Asn	Lys	Asn	Val	Val	Arg	Phe	Val	Leu	Gly	Arg
260				265				270							
Ile	Ala	Asp	Met	Gln	Lys	Lys	Gln	Gly	Gln	Lys	Gly	Lys	Asn	Gln	Ile
275				280				285							
Asp	Arg	Tyr	Tyr	Glu	Val	Cys	Val	Gly	Asn	Lys	Asp	Ile	Lys	Lys	Thr
290				295				300							
Ile	Glu	Glu	Lys	Ile	Asp	Ala	Leu	Thr	Asp	Ile	Ile	Val	Asn	Met	Asn
305				310				315				320			
Tyr	Asp	Gln	Phe	Glu	Lys	Lys	Lys	Ala	Val	Ile	Glu	Asn	Gln	Asn	Arg
325				330				335							
Gly	Lys	Thr	Phe	Glu	Glu	Lys	Asn	Lys	Tyr	Lys	Arg	Asp	Asn	Ala	Glu
340				345				350							
Arg	Glu	Lys	Phe	Lys	Lys	Ile	Ile	Ser	Leu	Tyr	Leu	Thr	Val	Ile	Tyr
355				360				365							
His	Ile	Leu	Lys	Asn	Ile	Val	Asn	Val	Asn	Ser	Arg	Tyr	Ile	Leu	Gly
370				375				380							
Phe	His	Cys	Leu	Glu	Arg	Asp	Lys	Gln	Leu	Tyr	Ile	Glu	Lys	Tyr	Asn
385				390				395				400			
Lys	Asp	Lys	Leu	Asp	Gly	Phe	Val	Ala	Leu	Thr	Lys	Phe	Cys	Leu	Gly

405	410	415
Asp Glu Glu Arg Tyr Glu Asp Leu Lys Ala Lys Ala Gln Ala Ser Ile		
420	425	430
Gln Ala Leu Glu Thr Ala Asn Pro Lys Leu Tyr Ala Lys Tyr Met Asn		
435	440	445
Tyr Ser Asp Glu Glu Lys Lys Glu Glu Phe Lys Lys Gln Leu Asn Arg		
450	455	460
Glu Arg Val Lys Asn Ala Arg Asn Ala Tyr Leu Lys Asn Ile Lys Asn		
465	470	475
Tyr Ile Met Ile Arg Leu Gln Leu Arg Asp Gln Thr Asp Ser Ser Gly		
485	490	495
Tyr Leu Cys Gly Glu Phe Arg Asp Lys Val Ala His Leu Glu Val Ala		
500	505	510
Arg His Ala His Glu Tyr Ile		
515		

<210> 74

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 74

ggggccgggg ccggggccgg 20

<210> 75

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 75

gggccggggc cggggccggg 20

<210> 76

<211> 20

<212> RNA

<213> 人工序列

<220>

<223> 间隔序列

<400> 76

ggccggggcc ggggccgggg 20

<210> 77
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 77
gccggggccg gggccggggc 20
<210> 78
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 78
ccggggccgg ggccggggcc 20
<210> 79
<211> 20
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 79
cggggccggg gccggggccg 20
<210> 80
<211> 3
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 80
cag 3
<210> 81
<211> 6
<212> RNA
<213> 人工序列
<220>
<223> 间隔序列
<400> 81

ggggcc 6

<210> 82

<211> 93

<212> RNA

<213> 人工序列

<220>

<223> 支架序列

<400> 82

guuuuagagc uaugcuggaa acagcauagc aaguuuuuuu aaggcuaguc cguuaucaac 60

uugaaaaaagu ggcaccgagu cggugcuuuu uuu 93

<210> 83

<211> 83

<212> RNA

<213> 人工序列

<220>

<223> 支架序列

<400> 83

guuuuagagc uagaaauagc aaguuuuuuu aaggcuaguc cguuaucaac uugaaaaaagu 60

ggcaccgagu cggugcuuuu uuu 83

<210> 84

<211> 181

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群2643000492

<400> 84

Asn Gly Glu Ile Val Ser Leu Ala Glu Lys Glu Ala Phe Ser Ala Lys

1 5 10 15

Ile Ala Asp Lys Asn Ile Gly Cys Lys Ile Glu Asn Lys Gln Phe Arg

20 25 30

His Pro Lys Gly Tyr Asp Val Ile Ala Asp Asn Pro Ile Tyr Lys Gly

35 40 45

Ser Pro Arg Gln Asp Met Leu Gly Leu Lys Glu Thr Leu Glu Lys Arg

50 55 60

Tyr Phe Ser Pro Ser Asp Ser Ile Asp Asn Val Arg Val Gln Val Ala

65 70 75 80

His Asn Ile Leu Asp Ile Glu Lys Ile Leu Ala Glu Tyr Ile Thr Asn

85 90 95

Ala Val Tyr Ser Phe Asp Asn Ile Ala Gly Phe Gly Lys Asp Ile Ile

100							105					110						
Gly	Asp	Asp	Phe	Ser	Pro	Val	Tyr	Thr	Tyr	Asp	Lys	Phe	Glu	Lys	Ser			
115							120					125						
Asp	Arg	Tyr	Glu	Tyr	Phe	Lys	Asn	Leu	Leu	Asn	Asn	Ser	Arg	Leu	Gly			
130							135					140						
Tyr	Tyr	Gly	Gln	Ala	Phe	Phe	Glu	Cys	Asp	Asp	Ser	Lys	Glu	Asn	Lys			
145							150					155					160	
Lys	Lys	Lys	Asp	Ala	Ile	Lys	Cys	Tyr	Asn	Ile	Ile	Ala	Leu	Leu	Ser			
165							170					175						
Gly	Leu	Arg	His	Trp														
180																		
<210> 85																		
<211> 440																		
<212> PRT																		
<213> 未知的																		
<220>																		
<223> CasRX/Cas13d肠道宏基因组重叠群874000057																		
<400> 85																		
Met	Ser	Lys	Asn	Lys	Glu	Ser	Tyr	Ala	Lys	Gly	Met	Gly	Leu	Lys	Ser			
1				5				10				15						
Ala	Leu	Val	Ser	Gly	Ser	Lys	Val	Tyr	Met	Thr	Ser	Phe	Glu	Gly	Gly			
20							25					30						
Asn	Asp	Ala	Lys	Leu	Glu	Lys	Val	Val	Glu	Asn	Ser	Glu	Ile	Val	Ser			
35				40						45								
Leu	Ala	Glu	Lys	Glu	Ser	Phe	Ser	Ala	Glu	Ile	Phe	Lys	Lys	Asn	Ile			
50				55						60								
Gly	Cys	Lys	Ile	Glu	Asn	Lys	Lys	Phe	Lys	His	Pro	Lys	Arg	Tyr	Asp			
65					70					75					80			
Val	Ile	Ala	Asp	Asn	Pro	Leu	Tyr	Lys	Gly	Ser	Val	Arg	Gln	Asp	Met			
85							90					95						
Leu	Gly	Leu	Lys	Glu	Thr	Leu	Glu	Lys	Arg	Tyr	Phe	Asn	Ser	Ala	Asp			
100							105					110						
Gly	Thr	Asp	Asn	Val	Cys	Ile	Gln	Val	Ile	His	Asn	Ile	Leu	Asp	Ile			
115							120					125						
Glu	Lys	Ile	Leu	Ala	Glu	Tyr	Ile	Thr	Asn	Ala	Val	Tyr	Ser	Phe	Asp			
130							135					140						
Asn	Ile	Ala	Gly	Phe	Gly	Glu	Asp	Ile	Ile	Gly	Met	Gly	Gly	Phe	Lys			
145							150					155					160	
Pro	Ile	Tyr	Thr	Tyr	Lys	Gln	Phe	Lys	Glu	Pro	Asp	Lys	Tyr	Asn	Lys			

	165		170		175
Lys Phe Asp Asp Ile Leu Asn Asn Ser Arg Leu Gly Tyr Tyr Gly Lys					
	180		185		190
Ala Phe Phe Glu Lys Asn Asp Leu Lys His Asn Pro Asn Lys Lys Lys					
	195		200		205
Arg Asp Lys Asn Pro Tyr Ile Leu Lys Tyr Asp Asn Glu Cys Tyr Tyr					
	210		215		220
Ile Ile Ala Leu Leu Ser Gly Leu Arg His Trp Asn Ile His Ser His					
225		230		235	240
Ala Lys Asp Asp Leu Val Ser Tyr Arg Trp Leu Tyr Asn Leu Asp Ser					
	245		250		255
Ile Leu Asn Arg Glu Tyr Ile Ser Thr Leu Asn Tyr Leu Tyr Asp Asp					
	260		265		270
Ile Ala Asp Glu Leu Thr Glu Ser Phe Ser Lys Asn Ser Ser Ala Asn					
	275		280		285
Val Asn Tyr Ile Ala Glu Thr Leu Asn Ile Asp Pro Ser Glu Phe Ala					
	290		295		300
Gln Gln Tyr Phe Arg Phe Ser Ile Met Lys Glu Gln Lys Asn Met Gly					
305		310		315	320
Phe Asn Val Ser Lys Leu Arg Glu Ile Met Leu Asp Arg Lys Glu Leu					
	325		330		335
Ser Asp Ile Arg Asp Asn His Arg Val Phe Asp Ser Ile Arg Ser Lys					
	340		345		350
Leu Tyr Thr Met Met Asp Phe Val Ile Tyr Arg Tyr Tyr Ile Glu Glu					
	355		360		365
Ala Ala Lys Thr Glu Ala Glu Asn Arg Asn Leu Pro Glu Asn Glu Lys					
	370		375		380
Lys Ile Ser Glu Lys Asp Phe Phe Val Ile Asn Leu Arg Gly Ser Phe					
385		390		395	400
Asp Glu Asn Gln Lys Glu Lys Leu Tyr Ile Glu Glu Ala Lys Arg Leu					
	405		410		415
Trp Glu Lys Leu Lys Asp Ile Met Leu Lys Ile Lys Glu Phe Arg Gly					
	420		425		430
Glu Lys Val Lys Glu Tyr Lys Lys					
	435		440		

<210> 86

<211> 137

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群4781000489

<400> 86

```

Leu Asp Lys Gln Leu Asp Tyr Glu Tyr Ile Arg Thr Leu Asn Tyr Met
1           5           10           15
Phe Asn Asp Ile Ala Asp Glu Leu Thr Arg Thr Phe Ser Lys Asn Ser
           20           25           30
Ala Ala Asn Val Asn Tyr Ile Ala Glu Thr Leu Asn Ile Asp Pro Asn
           35           40           45
Lys Phe Ala Glu Gln Tyr Phe Arg Phe Ser Ile Met Lys Glu Gln Lys
           50           55           60
Asn Leu Gly Phe Asn Leu Thr Lys Leu Arg Glu Ser Met Leu Asp Arg
65           70           75           80
Arg Glu Leu Ser Asp Ile Arg Asp Asn His Asn Val Phe Asp Ser Ile
           85           90           95
Arg Pro Lys Leu Tyr Thr Met Met Asp Phe Val Ile Tyr Lys His Tyr
           100          105          110
Ile Asp Glu Ala Lys Lys Thr Glu Ala Glu Asn Lys Ser Leu Pro Asp
           115          120          125
Asp Arg Lys Asn Leu Ser Glu Lys Asp
           130          135

```

<210> 87

<211> 87

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群12144000352

<400> 87

```

Arg Met Gly Glu Pro Val Ala Asn Thr Lys Arg Val Met Met Ile Asp
1           5           10           15
Ala Val Lys Ile Leu Gly Thr Asp Leu Ser Asp Asp Glu Leu Lys Glu
           20           25           30
Met Ala Asp Ser Phe Phe Lys Asp Ser Asp Gly Asn Leu Leu Lys Lys
           35           40           45
Gly Lys His Gly Met Arg Asn Phe Ile Thr Asn Asn Val Ile Lys Asn
           50           55           60
Lys Arg Phe His Tyr Leu Ile Arg Tyr Gly Asp Pro Ala His Leu His
65           70           75           80
Glu Ile Ala Lys Asn Glu Ala

```

85

<210> 88

<211> 414

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群5590000448

<400> 88

Val	His	Asn	Asn	Glu	Glu	Lys	Asp	Leu	Ile	Lys	Tyr	Thr	Trp	Leu	Tyr
1				5					10					15	
Asn	Leu	Asp	Lys	Tyr	Leu	Asp	Ala	Glu	Tyr	Ile	Thr	Thr	Leu	Asn	Tyr
			20					25					30		
Met	Tyr	Asn	Asp	Ile	Gly	Asp	Glu	Leu	Thr	Asp	Ser	Phe	Ser	Lys	Asn
		35					40					45			
Ser	Ala	Ala	Asn	Ile	Asn	Tyr	Ile	Ala	Glu	Thr	Leu	Gly	Ile	Asp	Pro
	50				55					60					
Lys	Thr	Phe	Ala	Glu	Gln	Tyr	Phe	Arg	Phe	Ser	Ile	Met	Lys	Glu	Gln
65				70				75						80	
Lys	Asn	Leu	Gly	Phe	Asn	Leu	Thr	Lys	Leu	Arg	Glu	Val	Met	Leu	Asp
			85					90					95		
Arg	Lys	Asp	Met	Ser	Glu	Ile	Arg	Glu	Asn	His	Asn	Asp	Phe	Asp	Ser
		100						105					110		
Ile	Arg	Ala	Lys	Val	Tyr	Thr	Met	Met	Asp	Phe	Val	Ile	Tyr	Arg	Tyr
	115						120					125			
Tyr	Ile	Glu	Glu	Ala	Ala	Lys	Val	Asn	Ala	Ala	Asn	Lys	Ser	Leu	Pro
	130					135					140				
Asp	Asn	Glu	Lys	Ser	Leu	Ser	Glu	Lys	Asp	Ile	Phe	Val	Ile	Ser	Leu
145				150					155					160	
Arg	Gly	Ser	Phe	Asn	Glu	Asp	Gln	Lys	Asp	Arg	Leu	Tyr	Tyr	Asp	Glu
			165					170						175	
Ala	Gln	Arg	Leu	Trp	Ser	Lys	Val	Gly	Lys	Leu	Met	Leu	Lys	Ile	Lys
		180						185					190		
Lys	Phe	Arg	Gly	Lys	Asp	Thr	Arg	Lys	Tyr	Lys	Asn	Met	Gly	Thr	Pro
	195					200						205			
Arg	Ile	Arg	Arg	Leu	Ile	Pro	Glu	Gly	Arg	Asp	Ile	Ser	Thr	Phe	Ser
	210				215						220				
Lys	Leu	Met	Tyr	Ala	Leu	Thr	Met	Phe	Leu	Asp	Gly	Lys	Glu	Ile	Asn
225				230					235					240	
Asp	Leu	Leu	Thr	Thr	Leu	Ile	Asn	Lys	Phe	Asp	Asn	Ile	Gln	Ser	Phe

245	250	255
Leu Lys Val Met Pro Leu Ile Gly Val Asn Ala Lys Phe Ala Glu Glu		
260	265	270
Tyr Ser Phe Phe Asn Asn Ser Glu Lys Ile Ala Asp Glu Leu Arg Leu		
275	280	285
Ile Lys Ser Phe Ala Arg Met Gly Glu Pro Val Ala Asp Ala Arg Arg		
290	295	300
Ala Met Tyr Ile Asp Ala Ile Arg Ile Leu Gly Thr Asp Leu Ser Asp		
305	310	315
Asp Glu Leu Lys Ala Leu Ala Asp Ser Phe Ser Leu Asp Glu Asn Gly		
325	330	335
Asn Lys Leu Gly Lys Gly Lys His Gly Met Arg Asn Phe Ile Ile Asn		
340	345	350
Asn Val Ile Thr Asn Lys Arg Phe His Tyr Leu Ile Arg Tyr Gly Asn		
355	360	365
Pro Val His Leu His Glu Ile Ala Lys Asn Glu Ala Val Val Lys Phe		
370	375	380
Val Leu Gly Arg Ile Ala Asp Ile Gln Lys Lys Gln Gly Gln Asn Gly		
385	390	395
Lys Asn Gln Ile Asp Arg Tyr Tyr Glu Thr Cys Ile Gly Lys		
405	410	

<210> 89

<211> 345

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群525000349

<400> 89

Met Ser Lys Lys Glu Asn Arg Lys Ser Tyr Val Lys Gly Leu Gly Leu		
1	5	10
Lys Ser Thr Leu Val Ser Asp Ser Lys Val Tyr Leu Thr Thr Phe Ala		
20	25	30
Asp Gly Ser Asn Ala Lys Leu Glu Lys Cys Val Glu Asn Asn Lys Ile		
35	40	45
Ile Cys Ile Ser Asn Asp Lys Glu Ala Phe Ala Ala Ser Ile Ala Asn		
50	55	60
Lys Asn Val Gly Tyr Lys Ile Lys Asn Asp Glu Lys Phe Arg His Pro		
65	70	75
Lys Gly Tyr Asp Ile Ile Ser Asn Asn Pro Leu Leu His Asn Asn Ser		

85	90	95
Val Gln Gln Asp Met Leu Gly Leu Lys Asn Val Leu Glu Lys Arg Tyr		
100	105	110
Phe Gly Lys Ser Ser Gly Gly Asp Asn Asn Leu Cys Ile Gln Ile Ile		
115	120	125
His Asn Ile Ile Asp Ile Glu Lys Ile Leu Ser Glu Tyr Ile Pro Asn		
130	135	140
Val Val Tyr Ala Phe Asn Asn Ile Ala Gly Phe Lys Asp Glu His Asn		
145	150	155
Asn Ile Ile Asp Ile Ile Gly Thr Gln Thr Tyr Asn Ser Ser Tyr Thr		
165	170	175
Tyr Ala Asp Phe Ser Lys Asp Lys Ser Asp Lys Lys Tyr Ile Glu Phe		
180	185	190
Gln Lys Leu Leu Lys Asn Lys Arg Leu Gly Tyr Trp Gly Lys Ala Phe		
195	200	205
Phe Thr Gly Gln Gly Asn Asn Ala Lys Val Arg Gln Glu Asn Gln Cys		
210	215	220
Phe His Ile Ile Ala Leu Leu Ile Ser Leu Arg Asn Trp Ala Thr His		
225	230	235
Ser Asn Glu Leu Asp Lys His Thr Lys Arg Thr Trp Leu Tyr Lys Leu		
245	250	255
Asp Asp Thr Asn Ile Leu Asn Ala Glu Tyr Val Lys Thr Leu Asn Tyr		
260	265	270
Leu Tyr Asp Thr Ile Ala Asp Glu Leu Thr Lys Ser Phe Ser Lys Asn		
275	280	285
Gly Ala Val Asn Val Asn Tyr Leu Ala Lys Lys Tyr Asn Ile Lys Asp		
290	295	300
Asp Leu Pro Gly Phe Ser Glu Gln Tyr Phe Arg Phe Ser Ile Met Lys		
305	310	315
Glu Gln Lys Asn Leu Gly Phe Asn Ile Ser Lys Leu Arg Glu Asn Met		
325	330	335
Leu Asp Phe Lys Asp Met Ser Val Ile		
340	345	

<210> 90

<211> 206

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群7229000302

<400> 90

Lys Lys Ile Ser Ser Leu Thr Lys Phe Cys Leu Gly Glu Ser Asp Glu
 1 5 10 15
 Lys Lys Leu Lys Lys Ala Leu Ala Lys Lys Ser Leu Glu Glu Leu Lys Thr
 20 25 30
 Thr Asn Ser Lys Leu Tyr Glu Asn Tyr Ile Lys Tyr Ser Asp Glu Arg
 35 40 45
 Lys Ala Glu Glu Ala Lys Arg Gln Ile Asn Arg Glu Arg Ala Lys Thr
 50 55 60
 Ala Met Asn Ala His Leu Arg Asn Thr Lys Trp Asn Asp Ile Met Tyr
 65 70 75 80
 Gly Gln Leu Lys Asp Leu Ala Asp Ser Lys Ser Arg Ile Cys Ser Glu
 85 90 95
 Phe Arg Asn Lys Ala Ala His Leu Glu Val Ala Arg Tyr Ala His Met
 100 105 110
 Tyr Ile Asn Asp Ile Ser Glu Val Lys Ser Tyr Phe Arg Leu Tyr His
 115 120 125
 Tyr Ile Met Gln Arg Arg Ile Ile Asp Val Ile Glu Asn Asn Pro Lys
 130 135 140
 Ala Lys Tyr Glu Gly Lys Val Lys Val Tyr Phe Glu Asp Val Lys Lys
 145 150 155 160
 Asn Lys Lys Tyr Asn Lys Asn Leu Leu Lys Leu Met Cys Val Pro Phe
 165 170 175
 Gly Tyr Cys Ile Pro Arg Phe Lys Asn Leu Ser Ile Glu Gln Met Phe
 180 185 190
 Asp Met Asn Glu Thr Asp Asn Ser Asp Lys Lys Lys Glu Lys
 195 200 205

<210> 91

<211> 95

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群3227000343

<400> 91

Ile Gly Asp Ile Ser Glu Val Asn Ser Tyr Phe Gln Leu Tyr His Tyr
 1 5 10 15
 Ile Met Gln Arg Ile Leu Ile Asp Lys Ile Gly Ser Lys Thr Thr Gly
 20 25 30
 Lys Ala Lys Glu Tyr Phe Asp Ser Val Ile Val Asn Lys Lys Tyr Asp

35	40	45
Asp Arg Leu Leu Lys Leu Leu Cys Ser Pro Leu Gly Tyr Cys Leu Thr		
50	55	60
Arg Tyr Lys Asp Leu Ser Ile Glu Ala Leu Phe Asp Met Asn Glu Ala		
65	70	75
Ala Lys Tyr Asp Lys Leu Asn Lys Glu Arg Lys Asn Lys Lys Lys		
85	90	95

<210> 92

<211> 115

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组重叠群7030000469

<400> 92

Ser Ile Arg Ser Lys Leu Tyr Thr Met Met Asp Phe Val Ile Tyr Arg		
1	5	10
Tyr Tyr Ile Glu Glu Ser Ala Lys Ala Ala Glu Asn Lys Pro Ser		
20	25	30
Glu Ser Asp Ser Phe Val Ile Arg Leu Arg Gly Ser Phe Asn Glu Asn		
35	40	45
Gln Lys Glu Glu Leu Tyr Ile Glu Glu Ala Glu Arg Leu Trp Lys Lys		
50	55	60
Phe Gly Glu Ile Met Leu Lys Ile Lys Glu Phe Arg Gly Glu Lys Val		
65	70	75
Lys Glu Tyr Lys Lys Glu Val Pro Arg Ile Glu Arg Ile Leu Pro His		
85	90	95
Gly Lys Asp Ile Ser Ala Phe Ser Lys Leu Met Tyr Met Leu Ser Met		
100	105	110

Phe Leu Asp

115

<210> 93

<211> 234

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d肠道宏基因组P17E0k2120140920, c87000043

<400> 93

Met Tyr Phe Ser Lys Met Ile Tyr Met Leu Thr Tyr Phe Leu Asp Gly		
1	5	10
		15

Lys Glu Ile Asn Asp Leu Leu Thr Thr Leu Ile Ser Lys Phe Asp Asn
 20 25 30
 Ile Lys Glu Phe Leu Lys Ile Met Lys Ser Ser Ala Val Asp Val Glu
 35 40 45
 Cys Glu Leu Thr Ala Gly Tyr Lys Leu Phe Asn Asp Ser Gln Arg Ile
 50 55 60
 Thr Asn Glu Leu Phe Ile Val Lys Asn Ile Ala Ser Met Arg Lys Pro
 65 70 75 80
 Ala Ala Ser Ala Lys Leu Thr Met Phe Arg Asp Ala Leu Thr Ile Leu
 85 90 95
 Gly Ile Asp Asp Lys Ile Thr Asp Asp Arg Ile Ser Glu Ile Leu Lys
 100 105 110
 Leu Lys Glu Lys Gly Lys Gly Ile His Gly Leu Arg Asn Phe Ile Thr
 115 120 125
 Asn Asn Val Ile Glu Ser Ser Arg Phe Val Tyr Leu Ile Lys Tyr Ala
 130 135 140
 Asn Ala Gln Lys Ile Arg Glu Val Ala Lys Asn Glu Lys Val Val Met
 145 150 155 160
 Phe Val Leu Gly Gly Ile Pro Asp Thr Gln Ile Glu Arg Tyr Tyr Lys
 165 170 175
 Ser Cys Val Glu Phe Pro Asp Met Asn Ser Ser Leu Glu Ala Lys Arg
 180 185 190
 Ser Glu Leu Ala Arg Met Ile Lys Asn Ile Ser Phe Asp Asp Phe Lys
 195 200 205
 Asn Val Lys Gln Gln Ala Lys Gly Arg Glu Asn Val Ala Lys Glu Arg
 210 215 220
 Ala Lys Ala Val Ile Gly Leu Tyr Leu Thr
 225 230
 <210> 94
 <211> 939
 <212> PRT
 <213> 未知的
 <220>
 <223> 重叠群emb OBVH01003037.1, 人肠道宏基因组序列
 <400> 94
 Met Ala Lys Lys Lys Arg Ile Thr Ala Lys Glu Arg Lys Gln Asn His
 1 5 10 15
 Arg Glu Leu Leu Met Lys Lys Ala Asp Ser Asn Ala Glu Lys Glu Lys
 20 25 30

Ala Lys Lys Pro Val Val Glu Asn Lys Pro Asp Thr Ala Ile Ser Lys		
35	40	45
Asp Asn Thr Pro Lys Pro Asn Lys Glu Ile Lys Lys Ser Lys Ala Lys		
50	55	60
Leu Ala Gly Val Lys Trp Val Ile Lys Ala Asn Asp Asp Val Ala Tyr		
65	70	75
Ile Ser Ser Phe Gly Lys Gly Asn Asn Ser Val Leu Glu Lys Arg Ile		
85	90	95
Met Gly Asp Val Ser Ser Asn Val Asn Lys Asp Ser His Met Tyr Val		
100	105	110
Asn Pro Lys Tyr Thr Lys Lys Asn Tyr Glu Ile Lys Asn Gly Phe Ser		
115	120	125
Ser Gly Ser Ser Leu Val Thr Tyr Pro Asn Lys Pro Asp Lys Asn Ser		
130	135	140
Gly Met Asp Ala Leu Cys Leu Lys Pro Tyr Phe Glu Lys Asp Phe Phe		
145	150	155
Gly His Ile Phe Thr Asp Asn Met His Ile Gln Ala Ile Tyr Asn Ile		
165	170	175
Phe Asp Ile Glu Lys Ile Leu Ala Lys His Ile Thr Asn Ile Ile Tyr		
180	185	190
Thr Val Asn Ser Phe Asp Arg Asn Tyr Asn Gln Ser Gly Asn Asp Thr		
195	200	205
Ile Gly Phe Gly Leu Asn Tyr Arg Val Pro Tyr Ser Glu Tyr Gly Gly		
210	215	220
Gly Lys Asp Ser Asn Gly Glu Pro Lys Asn Gln Ser Lys Trp Glu Lys		
225	230	235
Arg Asp Asn Phe Ile Lys Phe Tyr Asn Glu Ser Lys Pro His Leu Gly		
245	250	255
Tyr Tyr Glu Asn Ile Phe Tyr Asp His Gly Glu Pro Ile Ser Glu Glu		
260	265	270
Lys Phe Tyr Asn Tyr Leu Asn Ile Leu Asn Phe Ile Arg Asn Asn Thr		
275	280	285
Phe His Tyr Lys Asp Asp Asp Ile Glu Leu Tyr Ser Glu Asn Tyr Ser		
290	295	300
Glu Glu Phe Val Phe Ile Asn Cys Leu Asn Lys Phe Val Lys Asn Lys		
305	310	315
Phe Lys Asn Val Asn Lys Asn Phe Ile Ser Asn Glu Lys Asn Asn Leu		
325	330	335
Tyr Ile Ile Leu Asn Ala Tyr Gly Lys Asp Thr Glu Asn Val Glu Val		

340	345	350
Val Lys Lys Tyr Ser Lys Glu Leu Tyr Lys Leu Ser Val Leu Lys Thr		
355	360	365
Asn Lys Asn Leu Gly Val Asn Val Lys Lys Leu Arg Glu Ser Ala Ile		
370	375	380
Glu Tyr Gly Tyr Cys Pro Leu Pro Tyr Asp Lys Glu Lys Glu Val Ala		
385	390	395
Lys Leu Ser Ser Val Lys His Lys Leu Tyr Lys Thr Tyr Asp Phe Val		
405	410	415
Ile Thr His Tyr Leu Asn Ser Asn Asp Lys Leu Leu Leu Glu Ile Val		
420	425	430
Glu Thr Leu Arg Leu Ser Lys Asn Asp Asp Glu Lys Glu Asn Val Tyr		
435	440	445
Lys Lys Tyr Ala Glu Lys Leu Phe Lys Ala Asp Asp Val Ile Asn Pro		
450	455	460
Ile Lys Ala Ile Ser Lys Leu Phe Ala Arg Lys Gly Asn Lys Leu Phe		
465	470	475
Lys Glu Lys Ile Ile Ile Lys Lys Glu Tyr Ile Glu Asp Val Ser Ile		
485	490	495
Asp Lys Asn Ile Tyr Asp Phe Thr Lys Val Ile Phe Phe Met Thr Cys		
500	505	510
Phe Leu Asp Gly Lys Glu Ile Asn Asp Leu Leu Thr Asn Ile Ile Ser		
515	520	525
Lys Leu Gln Val Ile Glu Asp His Asn Asn Val Ile Lys Phe Ile Ser		
530	535	540
Asn Asn Lys Asp Ala Val Tyr Lys Asp Tyr Ser Asp Lys Tyr Ala Ile		
545	550	555
Phe Arg Asn Ala Gly Lys Ile Ala Thr Glu Leu Glu Ala Ile Lys Ser		
565	570	575
Ile Ala Arg Met Glu Asn Lys Ile Glu Asn Ala Pro Gln Glu Pro Leu		
580	585	590
Leu Lys Asp Ala Leu Leu Ser Leu Gly Val Ser Asp Asp Thr Lys Val		
595	600	605
Leu Glu Asn Thr Tyr Asn Lys Tyr Phe Asp Ser Lys Glu Lys Thr Asp		
610	615	620
Lys Gln Ser Gln Lys Val Ser Thr Phe Leu Met Asn Asn Val Ile Asn		
625	630	635
Asn Asn Arg Phe Lys Tyr Val Ile Lys Tyr Ile Asn Pro Ala Asp Ile		
645	650	655

Asn Gly Leu Ala Lys Asn Arg Tyr Leu Val Lys Phe Val Leu Ser Lys		
660	665	670
Ile Pro Glu Glu Gln Ile Asp Ser Tyr Tyr Lys Leu Phe Ser Asn Glu		
675	680	685
Glu Glu Pro Gly Cys Glu Glu Lys Ile Lys Leu Leu Thr Lys Lys Ile		
690	695	700
Ser Lys Leu Asn Phe Gln Thr Leu Phe Glu Asn Asn Lys Ile Pro Asn		
705	710	715
Val Glu Lys Glu Lys Lys Lys Ala Ile Ile Thr Leu Tyr Phe Thr Ile		
725	730	735
Val Tyr Ile Leu Val Lys Asn Leu Val Asn Ile Asn Gly Leu Tyr Thr		
740	745	750
Leu Ala Leu Tyr Phe Val Glu Arg Asp Gly Tyr Phe Tyr Lys Asp Ile		
755	760	765
Cys Gly Lys Lys Asp Lys Lys Lys Ser Tyr Asn Asp Val Asp Tyr Leu		
770	775	780
Leu Leu Pro Glu Ile Phe Ser Gly Ser Lys Tyr Arg Glu Glu Thr Lys		
785	790	795
Asn Leu Lys Leu Pro Lys Glu Lys Asp Arg Asp Ile Met Lys Lys Tyr		
805	810	815
Leu Pro Asn Asp Lys Asp Arg Glu Lys Tyr Asn Lys Phe Phe Thr Ala		
820	825	830
Tyr Arg Asn Asn Ile Val His Leu Asn Ile Ile Ala Lys Leu Ser Glu		
835	840	845
Leu Thr Lys Asn Ile Asp Lys Asp Ile Asn Ser Tyr Phe Asp Ile Tyr		
850	855	860
His Tyr Cys Thr Gln Arg Val Met Phe Asn Tyr Cys Lys Glu Lys Asn		
865	870	875
Asp Val Val Leu Ala Lys Met Lys Asp Leu Ala His Ile Lys Ser Asp		
885	890	895
Cys Asn Glu Phe Ser Ser Lys His Thr Tyr Pro Phe Ser Ser Ala Val		
900	905	910
Leu Arg Phe Met Asn Leu Pro Phe Ala Tyr Asn Val Pro Arg Phe Lys		
915	920	925
Asn Leu Ser Tyr Lys Lys Phe Phe Asp Lys Gln		
930	935	
<210> 95		
<211> 877		
<212> PRT		

<213> 未知的

<220>

<223> 重叠群tpg DJXD01000002.1

<400> 95

```

Met Lys Lys Gln Lys Ser Lys Lys Thr Val Ser Lys Thr Ser Gly Leu
1           5           10           15
Lys Glu Ala Leu Ser Val Gln Gly Thr Val Ile Met Thr Ser Phe Gly
           20           25           30
Lys Gly Asn Met Ala Asn Leu Ser Tyr Lys Ile Pro Ser Ser Gln Lys
           35           40           45
Pro Gln Asn Leu Asn Ser Ser Ala Gly Leu Lys Asn Val Glu Val Ser
           50           55           60
Gly Lys Lys Ile Lys Phe Gln Gly Arg His Pro Lys Ile Ala Thr Thr
65           70           75           80
Asp Asn Pro Leu Phe Lys Pro Gln Pro Gly Met Asp Leu Leu Cys Leu
           85           90           95
Lys Asp Lys Leu Glu Met His Tyr Phe Gly Lys Thr Phe Asp Asp Asn
           100          105          110
Ile His Ile Gln Leu Ile Tyr Gln Ile Leu Asp Ile Glu Lys Ile Leu
           115          120          125
Ala Val His Val Asn Asn Ile Val Phe Thr Leu Asp Asn Val Leu His
           130          135          140
Pro Gln Lys Glu Glu Leu Thr Glu Asp Phe Ile Gly Ala Gly Gly Trp
145          150          155          160
Arg Ile Asn Leu Asp Tyr Gln Thr Leu Arg Gly Gln Thr Asn Lys Tyr
           165          170          175
Asp Arg Phe Lys Asn Tyr Ile Lys Arg Lys Glu Leu Leu Tyr Phe Gly
           180          185          190
Glu Ala Phe Tyr His Glu Asn Glu Arg Arg Tyr Glu Glu Asp Ile Phe
           195          200          205
Ala Ile Leu Thr Leu Leu Ser Ala Leu Arg Gln Phe Cys Phe His Ser
           210          215          220
Asp Leu Ser Ser Asp Glu Ser Asp His Val Asn Ser Phe Trp Leu Tyr
225          230          235          240
Gln Leu Glu Asp Gln Leu Ser Asp Glu Phe Lys Glu Thr Leu Ser Ile
           245          250          255
Leu Trp Glu Glu Val Thr Glu Arg Ile Asp Ser Glu Phe Leu Lys Thr
           260          265          270
Asn Thr Val Asn Leu His Ile Leu Cys His Val Phe Pro Lys Glu Ser

```

275	280	285
Lys Glu Thr Ile Val Arg Ala Tyr Tyr Glu Phe Leu Ile Lys Lys Ser		
290	295	300
Phe Lys Asn Met Gly Phe Ser Ile Lys Lys Leu Arg Glu Ile Met Leu		
305	310	315
Glu Gln Ser Asp Leu Lys Ser Phe Lys Glu Asp Lys Tyr Asn Ser Val		
325	330	335
Arg Ala Lys Leu Tyr Lys Leu Phe Asp Phe Ile Ile Thr Tyr Tyr Tyr		
340	345	350
Asp His His Ala Phe Glu Lys Glu Ala Leu Val Ser Ser Leu Arg Ser		
355	360	365
Ser Leu Thr Glu Glu Asn Lys Glu Glu Ile Tyr Ile Lys Thr Ala Arg		
370	375	380
Thr Leu Ala Ser Ala Leu Gly Ala Asp Phe Lys Lys Ala Ala Ala Asp		
385	390	395
Val Asn Ala Lys Asn Ile Arg Asp Tyr Gln Lys Lys Ala Asn Asp Tyr		
405	410	415
Arg Ile Ser Phe Glu Asp Ile Lys Ile Gly Asn Thr Gly Ile Gly Tyr		
420	425	430
Phe Ser Glu Leu Ile Tyr Met Leu Thr Leu Leu Leu Asp Gly Lys Glu		
435	440	445
Ile Asn Asp Leu Leu Thr Thr Leu Ile Asn Lys Phe Asp Asn Ile Ile		
450	455	460
Ser Phe Ile Asp Ile Leu Lys Lys Leu Asn Leu Glu Phe Lys Phe Lys		
465	470	475
Pro Glu Tyr Ala Asp Phe Phe Asn Met Thr Asn Cys Arg Tyr Thr Leu		
485	490	495
Glu Glu Leu Arg Val Ile Asn Ser Ile Ala Arg Met Gln Lys Pro Ser		
500	505	510
Ala Asp Ala Arg Lys Ile Met Tyr Arg Asp Ala Leu Arg Ile Leu Gly		
515	520	525
Met Asp Asn Arg Pro Asp Glu Glu Ile Asp Arg Glu Leu Glu Arg Thr		
530	535	540
Met Pro Val Gly Ala Asp Gly Lys Phe Ile Lys Gly Lys Gln Gly Phe		
545	550	555
Arg Asn Phe Ile Ala Ser Asn Val Ile Glu Ser Ser Arg Phe His Tyr		
565	570	575
Leu Val Arg Tyr Asn Asn Pro His Lys Thr Arg Thr Leu Val Lys Asn		
580	585	590

Pro Asn Val Val Lys Phe Val Leu Glu Gly Ile Pro Glu Thr Gln Ile		
595	600	605
Lys Arg Tyr Phe Asp Val Cys Lys Gly Gln Glu Ile Pro Pro Thr Ser		
610	615	620
Asp Lys Ser Ala Gln Ile Asp Val Leu Ala Arg Ile Ile Ser Ser Val		
625	630	635 640
Asp Tyr Lys Ile Phe Glu Asp Val Pro Gln Ser Ala Lys Ile Asn Lys		
	645	650 655
Asp Asp Pro Ser Arg Asn Phe Ser Asp Ala Leu Lys Lys Gln Arg Tyr		
	660	665 670
Gln Ala Ile Val Ser Leu Tyr Leu Thr Val Met Tyr Leu Ile Thr Lys		
	675	680 685
Asn Leu Val Tyr Val Asn Ser Arg Tyr Val Ile Ala Phe His Cys Leu		
690	695	700
Glu Arg Asp Ala Phe Leu His Gly Val Thr Leu Pro Lys Met Asn Lys		
705	710	715 720
Lys Ile Val Tyr Ser Gln Leu Thr Thr His Leu Leu Thr Asp Lys Asn		
	725	730 735
Tyr Thr Thr Tyr Gly His Leu Lys Asn Gln Lys Gly His Arg Lys Trp		
	740	745 750
Tyr Val Leu Val Lys Asn Asn Leu Gln Asn Ser Asp Ile Thr Ala Val		
	755	760 765
Ser Ser Phe Arg Asn Ile Val Ala His Ile Ser Val Val Arg Asn Ser		
770	775	780
Asn Glu Tyr Ile Ser Gly Ile Gly Glu Leu His Ser Tyr Phe Glu Leu		
785	790	795 800
Tyr His Tyr Leu Val Gln Ser Met Ile Ala Lys Asn Asn Trp Tyr Asp		
	805	810 815
Thr Ser His Gln Pro Lys Thr Ala Glu Tyr Leu Asn Asn Leu Lys Lys		
	820	825 830
His His Thr Tyr Cys Lys Asp Phe Val Lys Ala Tyr Cys Ile Pro Phe		
	835	840 845
Gly Tyr Val Val Pro Arg Tyr Lys Asn Leu Thr Ile Asn Glu Leu Phe		
850	855	860
Asp Arg Asn Asn Pro Asn Pro Glu Pro Lys Glu Glu Val		
865	870	875

<210> 96

<211> 36

<212> DNA

<213> 未知的

<220>

<223> CasRX/Cas13d DR

<400> 96

caactacaac cccgtaaaaa tacgggggttc tgaaac 36

<210> 97

<211> 984

<212> PRT

<213> 未知的

<220>

<223> 重叠群OGZC01000639.1

<400> 97

Met	Lys	Lys	Lys	Asn	Ile	Arg	Ala	Thr	Arg	Glu	Ala	Leu	Lys	Ala	Gln
1				5					10					15	
Lys	Ile	Lys	Lys	Ser	Gln	Glu	Asn	Glu	Ala	Leu	Lys	Lys	Gln	Lys	Leu
				20				25					30		
Ala	Glu	Glu	Ala	Ala	Gln	Lys	Arg	Arg	Glu	Glu	Leu	Glu	Lys	Lys	Asn
				35				40					45		
Leu	Ala	Gln	Trp	Glu	Glu	Thr	Ser	Ala	Glu	Gly	Arg	Arg	Ser	Arg	Val
				50				55				60			
Lys	Ala	Val	Gly	Val	Lys	Ser	Val	Phe	Val	Val	Gly	Asp	Asp	Leu	Tyr
65					70				75					80	
Leu	Ala	Thr	Phe	Gly	Asn	Gly	Asn	Glu	Thr	Val	Leu	Glu	Lys	Lys	Ile
				85				90						95	
Thr	Pro	Asp	Gly	Lys	Ile	Thr	Thr	Phe	Pro	Glu	Glu	Glu	Thr	Phe	Thr
				100				105					110		
Ala	Lys	Leu	Lys	Phe	Ala	Gln	Thr	Glu	Pro	Thr	Val	Ala	Thr	Ser	Ile
				115				120					125		
Gly	Ile	Ser	Asn	Gly	Arg	Ile	Val	Leu	Pro	Glu	Ile	Ser	Val	Asp	Asn
				130				135					140		
Pro	Leu	His	Thr	Thr	Met	Gln	Lys	Asn	Thr	Ile	Lys	Arg	Ser	Ala	Gly
145					150				155					160	
Glu	Asp	Ile	Leu	Gln	Leu	Lys	Asp	Val	Leu	Glu	Asn	Arg	Tyr	Phe	Asp
				165					170					175	
Arg	Ser	Phe	Asn	Asp	Asp	Leu	His	Ile	Arg	Leu	Ile	Tyr	Asn	Ile	Leu
				180				185					190		
Asp	Ile	Glu	Lys	Ile	Leu	Ala	Glu	Tyr	Thr	Thr	Asn	Ala	Val	Phe	Ala
				195				200					205		
Ile	Asp	Asn	Val	Ser	Gly	Cys	Ser	Asp	Asp	Phe	Leu	Ser	Asn	Phe	Ser

210	215	220
Thr Arg Asn Gln Trp Asp Glu Phe Gln Asn Pro Glu Gln His Arg Glu		
225	230	235
His Phe Gly Asn Lys Asp Asn Val Ile Cys Ser Val Lys Lys Gln Gln		240
245	250	255
Asp Leu Phe Phe Asn Phe Phe Lys Asn Asn Arg Ile Gly Tyr Phe Gly		
260	265	270
Lys Ala Phe Phe His Ala Glu Ser Glu Arg Lys Ile Val Lys Lys Thr		
275	280	285
Glu Lys Glu Val Tyr His Ile Leu Thr Leu Ile Gly Ser Leu Arg Gln		
290	295	300
Trp Ile Thr His Ser Thr Glu Gly Gly Ile Ser Arg Leu Trp Leu Tyr		
305	310	315
Gln Leu Glu Asp Ala Leu Ser Arg Glu Tyr Gln Glu Thr Met Asn Asn		
325	330	335
Cys Tyr Asn Ser Thr Ile Tyr Gly Leu Gln Lys Asp Phe Glu Lys Thr		
340	345	350
Asn Ala Pro Asn Leu Asn Phe Leu Ala Glu Ile Leu Gly Lys Asn Ala		
355	360	365
Ser Glu Leu Ala Glu Pro Tyr Phe Arg Phe Ile Ile Thr Lys Glu Tyr		
370	375	380
Lys Asn Leu Gly Phe Ser Ile Lys Thr Leu Arg Glu Met Leu Leu Asp		
385	390	395
Gln Pro Asp Leu Gln Glu Ile Arg Glu Asn His Asn Val Tyr Asp Ser		
405	410	415
Ile Arg Ser Lys Leu Tyr Lys Met Ile Asp Phe Val Leu Val Tyr Ala		
420	425	430
Tyr Ser Asn Glu Arg Lys Ser Lys Ala Asp Ala Leu Ala Ser Asn Leu		
435	440	445
Arg Ser Ala Ile Thr Glu Asp Ala Lys Lys Arg Ile Tyr Gln Asn Glu		
450	455	460
Ala Asp Gln Leu Trp Thr Ser Tyr Gln Glu Leu Phe Lys Arg Ile Arg		
465	470	475
Gly Phe Lys Gly Ala Gln Val Lys Glu Tyr Ser Ser Lys Asn Met Pro		
485	490	495
Ile Pro Ile Gln Lys Gln Ile Gln Asn Ile Leu Lys Pro Ala Glu Gln		
500	505	510
Val Thr Tyr Phe Thr Lys Leu Met Tyr Leu Leu Thr Met Phe Leu Asp		
515	520	525

Gly Lys Glu Ile Asn Asp Leu Leu Thr Thr Leu Ile Asn Lys Phe Asp		
530	535	540
Asn Ile Ser Ser Leu Leu Lys Thr Met Glu Gln Leu Glu Leu Gln Thr		
545	550	555
Thr Phe Lys Glu Asp Tyr Thr Phe Phe Gln Gln Ser Ser Arg Leu Cys		
	565	570
Lys Glu Ile Thr Gln Leu Lys Ser Phe Ala Arg Met Gly Asn Pro Ile		
	580	585
Ser Asn Leu Lys Glu Val Met Met Val Asp Ala Ile Gln Ile Leu Gly		
	595	600
Thr Glu Lys Ser Glu Gln Glu Leu Gln Ser Met Ala Cys Phe Phe Phe		
	610	615
Arg Asp Lys Asn Gly Lys Lys Leu Asn Thr Gly Glu His Gly Met Arg		
625	630	635
Asn Phe Ile Gly Asn Asn Val Ile Ser Asn Thr Arg Phe Gln Tyr Leu		
	645	650
Ile Arg Tyr Gly Asn Pro Gln Lys Leu His Thr Leu Ser Gln Asn Glu		
	660	665
Thr Val Val Arg Phe Val Leu Ser Arg Ile Ala Lys Asn Gln Arg Val		
	675	680
Gln Gly Met Asn Gly Lys Asn Gln Ile Asp Arg Tyr Tyr Glu Thr Cys		
690	695	700
Gly Gly Thr Asn Ser Trp Ser Val Ser Glu Glu Glu Lys Ile Asn Phe		
705	710	715
Leu Cys Lys Ile Leu Thr Asn Met Ser Tyr Asp Gln Phe Gln Asp Val		
	725	730
Lys Gln Ser Gly Ala Glu Ile Thr Ala Glu Glu Lys Arg Lys Lys Glu		
	740	745
Arg Tyr Lys Ala Ile Ile Ser Leu Tyr Leu Thr Val Leu Tyr Gln Leu		
	755	760
Ile Lys Asn Leu Val Asn Ile Asn Ala Arg Tyr Ile Ile Ala Phe His		
	770	775
Cys Leu Glu Arg Asp Ala Ile Leu Tyr Ser Ser Lys Phe Asn Thr Ser		
785	790	795
Ile Asn Leu Lys Lys Arg Tyr Thr Ala Leu Thr Glu Met Ile Leu Gly		
	805	810
Tyr Glu Thr Asp Glu Lys Ala Arg Arg Lys Asp Thr Arg Thr Val Tyr		
	820	825
Glu Lys Ala Glu Ala Ala Lys Asn Arg His Leu Lys Asn Val Lys Trp		
		830

65	70	75	80
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	85	90	95
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	100	105	110
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	115	120	125
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	130	135	140
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
145	150	155	160
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	165	170	175
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	180	185	190
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa			
	195	200	205
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Pro			
210	215	220	
Leu Gln Lys Arg Tyr Arg Tyr Leu Thr Ser Thr Asn Leu Lys Ser Phe			
225	230	235	240
Glu Thr Tyr Lys Asn Asn Leu Val Asn Lys Lys Lys Phe Asp Leu Asp			
	245	250	255
Arg Val Lys Lys Ile Pro Gln Leu Ala Tyr Phe Gly Ser Ala Phe Tyr			
	260	265	270
Asn Thr Pro Glu Asp Thr Ser Ala Lys Ile Thr Lys Thr Lys Ile Lys			
	275	280	285
Ser Asn Glu Glu Ile Tyr Tyr Thr Phe Met Leu Leu Ser Thr Ala Arg			
	290	295	300
Asn Phe Ser Ala His Tyr Leu Asp Arg Asn Arg Ala Lys Ser Ser Asp			
305	310	315	320
Ala Glu Asp Phe Asp Gly Thr Ser Val Ile Met Tyr Asn Leu Asp Asn			
	325	330	335
Glu Glu Leu Tyr Lys Lys Leu Tyr Asn Lys Lys Val His Met Ala Leu			
	340	345	350
Thr Gly Met Lys Lys Val Leu Asp Ala Asn Phe Asn Lys Lys Val Glu			
	355	360	365
His Leu Asn Asn Ser Phe Ile Lys Asn Ser Ala Lys Asp Phe Val Ile			
370	375	380	

Leu Cys Glu Val Leu Gly Ile Lys Ser Arg Asp Glu Lys Thr Lys Phe			
385	390	395	400
Val Lys Asp Tyr Tyr Asp Phe Val Val Arg Lys Asn Tyr Lys His Leu			
	405	410	415
Gly Phe Ser Val Lys Glu Leu Arg Glu Leu Leu Phe Ala Asn His Asp			
	420	425	430
Ser Asn Lys Tyr Ile Lys Glu Phe Asp Lys Ile Ser Asn Lys Lys Phe			
	435	440	445
Asp Ser Val Arg Ser Arg Leu Asn Arg Leu Ala Asp Tyr Ile Ile Tyr			
	450	455	460
Asp Tyr Tyr Asn Lys Asn Asn Ala Lys Val Ser Asp Leu Val Lys Tyr			
465	470	475	480
Leu Arg Ala Ala Ala Asp Asp Glu Gln Lys Lys Lys Ile Tyr Leu Asn			
	485	490	495
Glu Ser Ile Asn Leu Val Lys Ser Gly Ile Leu Glu Arg Ile Lys Lys			
	500	505	510
Ile Leu Pro Lys Leu Asn Gly Lys Ile Ile Gly Asn Met Gln Pro Asp			
	515	520	525
Ser Thr Ile Thr Ala Ser Met Leu His Asn Thr Gly Lys Asp Trp His			
	530	535	540
Pro Ile Ser Glu Asn Ala His Tyr Phe Thr Lys Trp Ile Tyr Thr Leu			
545	550	555	560
Thr Leu Phe Met Asp Gly Lys Glu Ile Asn Asp Leu Val Thr Thr Leu			
	565	570	575
Ile Asn Lys Phe Asp Asn Ile Ala Ser Phe Ile Glu Val Leu Lys Ser			
	580	585	590
Gln Ser Val Cys Thr His Phe Ser Glu Glu Arg Lys Met Phe Ile Asp			
	595	600	605
Ser Ala Glu Ile Cys Ser Glu Leu Ser Ala Met Asn Ser Phe Ala Arg			
	610	615	620
Met Glu Ala Pro Gly Ala Ser Ser Lys Arg Ala Met Phe Val Glu Ala			
625	630	635	640
Ala Arg Ile Leu Gly Asp Asn Arg Ser Lys Glu Glu Leu Glu Glu Tyr			
	645	650	655
Phe Asp Thr Leu Phe Asp Lys Ser Ala Ser Lys Lys Glu Lys Gly Phe			
	660	665	670
Arg Asn Phe Ile Arg Asn Asn Val Val Asp Ser Asn Arg Phe Lys Tyr			
	675	680	685
Leu Thr Arg Tyr Thr Asp Thr Ser Ser Val Lys Ala Phe Ser Asn Asn			

690	695	700
Lys Ala Leu Val Lys Phe Ala Ile Lys Asp Ile Pro Gln Glu Gln Ile		
705	710	715
Leu Arg Tyr Tyr Asn Ser Cys Phe Gly Ala Ser Glu Arg Tyr Tyr Asn		
	725	730
Asp Gly Met Ser Asp Lys Leu Val Glu Ala Ile Gly Lys Ile Asn Leu		
	740	745
Met Gln Phe Asn Gly Val Ile Gln Gln Ala Asp Arg Asn Met Leu Pro		
	755	760
Glu Glu Lys Lys Lys Ala Asn Ala Gln Lys Glu Lys Tyr Lys Ser Ile		
	770	780
Ile Arg Leu Tyr Leu Thr Val Cys Tyr Leu Phe Phe Lys Asn Leu Val		
785	790	795
Tyr Val Asn Ser Arg Tyr Tyr Ser Ala Phe Tyr Asn Leu Glu Lys Asp		
	805	810
Arg Ser Leu Phe Glu Ile Asn Gly Glu Leu Lys Pro Thr Gly Lys Phe		
	820	825
Asp Glu Gly His Tyr Thr Gly Leu Val Lys Leu Phe Ile Asp Asn Gly		
	835	840
Trp Ile Asn Pro Arg Ala Ser Ala Tyr Leu Thr Val Asn Leu Ala Asn		
	850	855
Ser Asp Glu Thr Ala Ile Arg Thr Phe Arg Asn Thr Ala Glu His Leu		
865	870	875
Glu Ala Leu Arg Asn Ala Asp Lys Tyr Leu Asn Asp Leu Lys Gln Phe		
	885	890
Asp Ser Tyr Phe Glu Ile Tyr His Tyr Ile Thr Gln Arg Asn Ile Lys		
	900	905
Glu Lys Cys Glu Met Leu Lys Glu Gln Thr Val Lys Tyr Asn Asn Asp		
	915	920
Leu Leu Lys Tyr His Gly Tyr Ser Lys Asp Phe Val Lys Ala Leu Cys		
	930	935
Val Pro Phe Gly Tyr Asn Leu Pro Arg Phe Lys Asn Leu Ser Ile Asp		
945	950	955
Ala Leu Phe Asp Lys Asn Asp Lys Arg Glu Lys Leu Lys Lys Gly Phe		
	965	970
		975
Glu Asp		
<210> 99		
<211> 1023		
<212> PRT		

<213> 未知的

<220>

<223> 重叠群emb OHCP01000044.1

<400> 99

```

Met Ala Lys Lys Ile Thr Ala Lys Gln Lys Arg Glu Glu Lys Glu Arg
1           5           10           15
Leu Asn Lys Gln Lys Trp Ala Lys Asn Asp Ser Val Ile Ile Val Pro
          20           25           30
Glu Thr Lys Glu Glu Ile Lys Thr Gly Glu Ile Gln Asp Asn Asn Arg
          35           40           45
Lys Arg Ser Arg Gln Lys Ser Gln Ala Lys Ala Met Gly Leu Lys Ala
          50           55           60
Val Leu Ser Phe Asp Asn Lys Ile Ala Ile Ala Ser Phe Val Ser Ser
65           70           75           80
Lys Asn Ala Lys Ser Ser His Ile Glu Arg Ile Thr Asp Lys Glu Gly
          85           90           95
Thr Thr Ile Ser Val Asn Ser Lys Met Phe Glu Ser Ser Val Asn Lys
          100          105          110
Arg Asp Ile Asn Ile Glu Lys Arg Ile Thr Ile Glu Glu Pro Gln Gln
          115          120          125
Asp Gly Thr Ile Lys Lys Glu Glu Lys Gly Val Lys Ser Thr Thr Cys
          130          135          140
Asn Pro Tyr Phe Lys Val Gly Gly Lys Asp Tyr Ile Gly Ile Lys Glu
145          150          155          160
Ile Ala Glu Glu His Phe Phe Gly Arg Ala Phe Pro Asn Glu Asn Leu
          165          170          175
Arg Val Gln Ile Ala Tyr Asn Ile Phe Asp Val Gln Lys Ile Leu Gly
          180          185          190
Thr Phe Val Asn Asn Ile Ile Tyr Ser Phe Tyr Asn Leu Ser Arg Asp
          195          200          205
Glu Val Gln Ser Asp Asn Asp Val Ile Gly Met Leu Tyr Ser Ile Ser
          210          215          220
Asp Tyr Asp Arg Gln Lys Glu Thr Glu Thr Phe Leu Gln Ala Lys Ser
225          230          235          240
Leu Leu Lys Gln Thr Glu Ala Tyr Tyr Ala Tyr Phe Asp Asp Val Phe
          245          250          255
Lys Lys Asn Lys Lys Pro Asp Lys Asn Lys Glu Gly Asp Asn Ser Lys
          260          265          270
Gln Tyr Gln Glu Asn Leu Arg His Asn Phe Asn Ile Leu Arg Val Leu

```

275	280	285
Ser Phe Leu Arg Gln Ile Cys Met His Ala Glu Val His Val Ser Asp		
290	295	300
Asp Glu Gly Cys Thr Arg Thr Gln Asn Tyr Thr Asp Ser Leu Glu Ala		
305	310	315
Leu Phe Asn Ile Ser Lys Ala Phe Gly Lys Lys Met Pro Glu Leu Lys		
325	330	335
Thr Leu Ile Asp Asn Ile Tyr Ser Lys Gly Ile Asn Ala Ile Asn Asp		
340	345	350
Glu Phe Val Lys Asn Gly Lys Asn Asn Leu Tyr Ile Leu Ser Lys Val		
355	360	365
Tyr Pro Asn Glu Lys Arg Glu Val Leu Leu Arg Glu Tyr Tyr Asn Phe		
370	375	380
Val Val Cys Lys Glu Gly Ser Asn Ile Gly Ile Ser Thr Arg Lys Leu		
385	390	395
Lys Glu Thr Met Ile Ala Gln Asn Met Pro Ser Leu Lys Glu Glu Asn		
405	410	415
Thr Tyr Arg Asn Lys Leu Tyr Thr Val Met Asn Phe Ile Leu Val Arg		
420	425	430
Glu Leu Lys Asn Cys Ala Thr Ile Arg Glu Gln Met Ile Lys Glu Leu		
435	440	445
Arg Ala Asn Met Asp Glu Glu Gly Arg Asp Arg Ile Tyr Ser Lys		
450	455	460
Tyr Ala Lys Glu Ile Tyr Leu Tyr Val Lys Asp Lys Leu Lys Leu Met		
465	470	475
Leu Asn Val Phe Lys Glu Glu Ala Glu Gly Ile Ile Ile Pro Gly Lys		
485	490	495
Glu Asp Pro Val Lys Phe Ser His Gly Lys Leu Asp Lys Lys Glu Ile		
500	505	510
Glu Ser Phe Cys Leu Thr Thr Lys Asn Thr Glu Asp Ile Thr Lys Val		
515	520	525
Ile Tyr Phe Leu Cys Lys Phe Leu Asp Gly Lys Glu Ile Asn Glu Leu		
530	535	540
Cys Cys Ala Met Met Asn Lys Leu Asp Gly Ile Ser Asp Leu Ile Glu		
545	550	555
Thr Ala Lys Gln Cys Gly Glu Asp Val Glu Phe Val Asp Gln Phe Lys		
565	570	575
Cys Leu Ser Lys Cys Ala Thr Met Ser Asn Gln Ile Arg Ile Val Lys		
580	585	590

Asn Ile Ser Arg Met Lys Lys Glu Met Thr Ile Asp Asn Asp Thr Ile		
595	600	605
Phe Leu Asp Ala Leu Glu Leu Leu Gly Arg Lys Ile Glu Lys Tyr Gln		
610	615	620
Lys Asp Lys Asn Gly Asp Tyr Val Lys Asp Glu Lys Gly Lys Lys Val		
625	630	635
Tyr Thr Lys Asp Tyr Asn Asn Phe Gln Asp Met Phe Phe Glu Gly Lys		
	645	650
		655
Asn His Arg Val Arg Asn Phe Val Ser Asn Asn Val Ile Lys Ser Lys		
	660	665
		670
Trp Phe Ser Tyr Val Val Arg Tyr Asn Lys Pro Ala Glu Cys Gln Ala		
	675	680
		685
Leu Met Arg Asn Ser Lys Leu Val Lys Phe Ala Leu Asp Glu Leu Pro		
	690	695
		700
Asp Ser Gln Ile Glu Lys Tyr Tyr Ile Ser Val Phe Gly Glu Lys Ser		
705	710	715
		720
Ser Ser Ser Asn Glu Glu Met Arg Arg Glu Leu Leu Lys Lys Leu Cys		
	725	730
		735
Asp Phe Ser Val Arg Gly Phe Leu Asp Glu Ile Val Leu Leu Ser Glu		
	740	745
		750
Asp Glu Met Lys Gln Lys Asp Lys Phe Ser Glu Lys Glu Lys Lys Lys		
	755	760
		765
Ser Leu Ile Arg Leu Tyr Leu Thr Ile Val Tyr Leu Ile Thr Lys Ser		
	770	775
		780
Met Val Lys Ile Asn Thr Arg Phe Ser Ile Ala Cys Ala Thr Tyr Glu		
785	790	795
		800
Arg Asp Tyr Ile Leu Leu Cys Gln Ser Glu Lys Ala Glu Arg Ala Trp		
	805	810
		815
Glu Lys Gly Ala Thr Ala Phe Ala Leu Thr Arg Lys Phe Leu Asn His		
	820	825
		830
Asp Lys Pro Thr Phe Glu Gln Tyr Tyr Thr Arg Glu Arg Glu Ile Ser		
	835	840
		845
Ala Met Pro Gln Glu Lys Arg Lys Glu Leu Arg Lys Glu Asn Asp Gln		
	850	855
		860
Leu Leu Lys Lys Thr His Tyr Ser Lys His Ala Tyr Cys Tyr Ile Val		
865	870	875
		880
Asp Asn Val Asn Asn Leu Thr Gly Ala Val Ala Asn Asp Asn Gly Arg		
	885	890
		895
Gly Leu Pro Cys Leu Ser Glu Lys Asn Asp Asn Ala Asn Leu Phe Leu		

900	905	910
Glu Met Arg Asn Lys Ile Val His Leu Asn Val Val His Asp Met Val		
915	920	925
Lys Tyr Ile Asn Glu Ile Lys Asn Ile Thr Ser Tyr Tyr Ala Phe Phe		
930	935	940
Cys Tyr Val Leu Gln Arg Met Ile Ile Gly Asn Asn Ser Asn Glu Gln		
945	950	955
Asn Lys Phe Lys Ala Lys Tyr Ser Lys Thr Leu Gln Glu Phe Gly Thr		
965	970	975
Tyr Ser Lys Asp Leu Met Trp Val Leu Asn Leu Pro Phe Ala Tyr Asn		
980	985	990
Leu Pro Arg Tyr Lys Asn Leu Ser Asn Glu Gln Leu Phe Tyr Asp Glu		
995	1000	1005
Glu Glu Arg Met Glu Lys Ile Val Gly Arg Lys Asn Asp Ser Arg		
1010	1015	1020
<210> 100		
<211> 926		
<212> PRT		
<213> 未知的		
<220>		
<223> 重叠群emb OGDF01008514.1		
<400> 100		
Met Thr Glu Thr Lys Pro Lys Arg Glu Asp Ile Ala Lys Thr Pro Ala		
1	5	10
Ala Lys Ser Arg Ser Lys Ala Ala Gly Leu Lys Ser Thr Phe Ala Val		
20	25	30
Asn Gly Ser Val Leu Leu Thr Ser Phe Gly Arg Gly Asn Asp Ala Val		
35	40	45
Pro Glu Lys Leu Ile Thr Glu Lys Ala Val Ser Glu Ile Asn Thr Val		
50	55	60
Lys Pro Arg Phe Ser Val Glu Lys Pro Ala Thr Ser Tyr Ser Ser Ser		
65	70	75
Phe Gly Ile Lys Ser His Ile Ser Ala Thr Ala Asp Asn Pro Leu Ala		
85	90	95
Gly Arg Ala Pro Val Gly Glu Asp Ala Ile His Ala Lys Glu Val Leu		
100	105	110
Glu Gln Arg Val Phe Gly Lys Thr Phe Ser Asp Asp Asn Ile His Ile		
115	120	125
Gln Leu Ile Tyr Asn Ile Leu Asp Ile Arg Lys Ile Leu Ser Thr Tyr		

130	135	140
Ala Asn Asn Val Val Phe Thr Ile Asn Ser Met Arg Arg Leu Asp Glu		
145	150	155
Tyr Asp Arg Glu Gln Asp Tyr Leu Gly Tyr Leu Tyr Thr Gly Asn Ser		
165	170	175
Tyr Glu Arg Leu Leu Asp Ile Ala Asp Lys Tyr Ala Val Asp Gly Glu		
180	185	190
Asp Trp Arg Asn Thr Ala Ala Gly Ile Ser Asn Asp Phe Glu Lys Lys		
195	200	205
Gln Phe Gln Thr Ile Asn Gly Phe Trp Asp Leu Leu Asp Met Ile Glu		
210	215	220
Pro Tyr Met Cys Tyr Phe Ser Glu Ala Phe Phe Cys Glu Thr Thr Val		
225	230	235
Lys Asp Pro Asp Ser Gly Arg Ile Val Pro Cys Leu Glu Gln Arg Ser		
245	250	255
Asp Gly Asp Ile Tyr Asn Ile Leu Arg Ile Leu Ser Ile Val Arg Gln		
260	265	270
Thr Cys Met His Asp Asn Ala Ser Met Arg Thr Val Met Phe Thr Leu		
275	280	285
Gly Gln Asn Ser Val Arg Asp Arg Lys Asn Gly Phe Asp Glu Leu Ala		
290	295	300
Glu Leu Leu Asp Tyr Leu Tyr Asp Glu Lys Ile Asp Ile Val Asn Arg		
305	310	315
Asp Phe Leu Arg Asn Gln Lys Asn Asn Ile Glu Leu Leu Ser Arg Ile		
325	330	335
Tyr Gly Ser Ser Ala Asp Ser Pro Glu Arg Asp Arg Leu Val Gln Asn		
340	345	350
Phe Tyr Asp Phe Arg Val Leu Ser Gln Asp Lys Asn Leu Gly Phe Ser		
355	360	365
Ile Lys Lys Leu Arg Glu Lys Leu Leu Asp Ser Pro Ala Leu Ser Val		
370	375	380
Val Arg Ser Lys Lys Tyr Asp Thr Met Arg Ser Lys Ile Tyr Ser Leu		
385	390	395
Ile Asp Phe Met Ile Tyr Arg Lys Phe Ser Glu Asn His Val Ala Val		
405	410	415
Asp Asp Phe Val Glu Glu Leu Arg Ser Leu Leu Thr Glu Asp Glu Lys		
420	425	430
Glu Ser Ala Tyr Ser Arg Trp Ala Glu Thr Leu Ile Asn Asp Gly Phe		
435	440	445

Ala Gln Glu Ile Leu Val Lys Leu Leu Pro Gln Thr Asp Pro Ala Val		
450	455	460
Ile Gly Lys Ile Lys Gly Lys Lys Leu Leu Asn Asp Ser Ile Ala Gly		
465	470	475 480
Ile Lys Leu Lys Lys Asp Ala Ser Phe Phe Thr Lys Ile Ile Asn Val		
	485	490 495
Leu Cys Met Phe Gln Asp Gly Lys Glu Ile Asn Glu Leu Val Ser Ser		
	500	505 510
Leu Val Asn Lys Phe Ala Asn Ile Gln Ser Phe Val Asp Val Met Arg		
	515	520 525
Ser Gln Gly Ile Asp Ser Gly Phe Thr Ala Asp Tyr Ala Met Phe Ala		
	530	535 540
Glu Ser Gly Arg Ile Ser Arg Glu Leu His Ile Leu Lys Gly Ile Ala		
545	550	555 560
Arg Met Gln His Ser Ile Ala Gly Leu Gly Asp Val Lys Ile Tyr Gly		
	565	570 575
Ser Asp Asp Lys Phe His Gly Val Ser Arg Arg Val Tyr Thr Asp Ala		
	580	585 590
Ala Tyr Ile Leu Gly Phe Gly Glu Arg Ser Glu Asp Asn Asp Gly Tyr		
	595	600 605
Val Asp Asp Tyr Val Ser Ser Lys Leu Leu Gly Gly Ala Asp Lys Asn		
610	615	620
Leu Arg Asn Phe Ile Thr Asn Asn Val Ile Lys Asn Arg Arg Phe Leu		
625	630	635 640
Tyr Thr Val Arg Tyr Met Asn Pro Lys Arg Ala Lys Lys Leu Val Gln		
	645	650 655
Asn Asp Ala Leu Val Val Leu Ala Leu Ser Gly Ile Pro Glu Thr Gln		
	660	665 670
Ile Asp Arg Tyr Tyr Lys Ser Cys Ile Glu Lys Arg Ser Phe Asn Pro		
	675	680 685
Asp Leu Asn Glu Lys Ile Ala Ala Leu Ser Glu Met Ile Thr Thr Leu		
690	695	700
Lys Ile Asp Asp Phe Glu Asp Val Lys Gln Asn Pro Glu Lys Asn Ala		
705	710	715 720
Asn Tyr Glu Ala Lys Lys Asn Gln Arg Ile Ser Lys Glu Arg Tyr Lys		
	725	730 735
Ala Cys Ile Gly Leu Tyr Leu Thr Val Leu Tyr Leu Ile Cys Lys Asn		
	740	745 750
Leu Val Lys Ile Asn Ala Arg Tyr Ser Ile Ala Ile Gly Cys Leu Glu		

755	760	765
Arg Asp Thr Gln Leu His Gly Val Asp Phe Lys Gly Ala Ala Tyr Met		
770	775	780
Thr Arg Asp Val Phe Ile Ala Lys Gly Trp Ile Asn Pro Lys Lys Pro		
785	790	795
Thr Val Lys Ser Ile Lys Glu Gln Tyr Ala Phe Leu Thr Pro Tyr Ile		800
805	810	815
Phe Thr Thr Tyr Arg Asn Met Ile Ala His Leu Ala Ala Val Thr Asn		
820	825	830
Ala Tyr Lys Tyr Ile Pro Gln Met Asp Arg Phe Lys Ser Trp Phe His		
835	840	845
Leu Tyr His Thr Val Ile Gln His Ser Leu Ile Gln Gln Tyr Glu Tyr		
850	855	860
Asp Arg Asp Tyr Gly Arg Lys Gly Ala Pro Val Val Ser Glu Arg Val		
865	870	875
Leu Gln Leu Leu Glu Gln Cys Arg Glu His Ser Asn Tyr Ser Arg Asp		880
885	890	895
Leu Leu His Ile Leu Asn Leu Pro Phe Gly Tyr Asn Leu Pro Arg Tyr		
900	905	910
Leu Asn Leu Ser Ser Glu Lys Tyr Phe Asp Ala Asn Ala Ile		
915	920	925
<210> 101		
<211> 1030		
<212> PRT		
<213> 未知的		
<220>		
<223> 重叠群emb OGPN01002610.1		
<400> 101		
Met Ala Lys Lys Ile Thr Ala Lys Gln Lys Arg Glu Glu Lys Glu Arg		
1	5	10
Leu Asn Lys Gln Lys Trp Ala Lys Gln Asp Thr Pro Val Val Pro Lys		
20	25	30
Ser Lys Thr Glu Glu Lys Pro Val Ala Ala Ser Asp Asp Lys Leu Leu		
35	40	45
Lys Thr Thr Gln Val Lys Lys Val Gln Thr Lys Ser Lys Ala Lys Ala		
50	55	60
Met Gly Leu Lys Thr Val Leu Ser Phe Asp Asp Lys Ile Ala Ile Ala		
65	70	75
Ser Phe Val Asn Asp Lys Lys Thr Lys Leu Pro His Ile Glu Arg Ile		80

				85					90					95			
Thr	Asp	Lys	Ser	Gly	Thr	Thr	Ile	His	Glu	Asn	Ala	Arg	Met	Phe	Asp		
				100					105					110			
Ser	Ser	Val	Asp	Glu	Gln	Asn	Val	Asn	Ile	Glu	Lys	Arg	Met	Thr	Ile		
				115					120					125			
Glu	Glu	Lys	Gln	Asn	Asp	Gly	Thr	Phe	Lys	Lys	Asp	Glu	Lys	Asp	Val		
				130					135					140			
Lys	Ala	Thr	Ile	Cys	Asn	Pro	Tyr	Phe	Lys	Thr	Cys	Gly	Lys	Asp	Tyr		
145						150					155				160		
Ile	Gly	Ile	Lys	Asp	Val	Ala	Glu	Lys	Tyr	Phe	Phe	Gly	Lys	Thr	Phe		
				165					170					175			
Pro	Asn	Glu	Asn	Leu	Arg	Val	Gln	Ile	Ala	Tyr	Asn	Val	Phe	Asp	Ile		
				180					185					190			
Gln	Lys	Ile	Leu	Gly	Thr	Tyr	Val	Asn	Asn	Ile	Ile	Tyr	Ser	Phe	Tyr		
				195					200					205			
Asn	Leu	Arg	Arg	Asp	Gly	Lys	Ser	Asp	Val	Asp	Ile	Ile	Gly	Ser	Leu		
				210					215					220			
Tyr	Ala	Phe	Ala	Asp	Phe	Asp	Asn	Gln	Leu	Lys	Asp	Lys	Pro	Ala	Phe		
225						230					235				240		
Arg	Glu	Ala	Lys	Asp	Leu	Leu	Lys	Asn	Thr	Glu	Ala	Tyr	Phe	Ser	Tyr		
				245							250				255		
Phe	Gly	Asp	Val	Phe	Lys	Lys	Ser	Lys	Lys	Gly	Lys	Lys	Asp	Glu	Asn		
				260					265					270			
Asn	Glu	Asp	Tyr	Glu	Lys	Asn	Leu	Arg	His	Asn	Phe	Asn	Val	Leu	Arg		
				275					280					285			
Val	Leu	Ser	Phe	Leu	Arg	Gln	Ile	Cys	Thr	His	Ala	Tyr	Val	Lys	Cys		
				290					295					300			
Thr	Gly	Gly	Ala	Lys	Asn	Asn	Gly	Asp	Ser	Thr	Lys	Val	Glu	Ala	Glu		
305						310					315				320		
Ser	Leu	Asp	Ala	Leu	Phe	Asn	Ile	Thr	Glu	Tyr	Phe	Ala	Lys	Thr	Ala		
				325							330				335		
Pro	Glu	Leu	Ser	Lys	Thr	Ile	Asn	Glu	Ile	Tyr	Lys	Glu	Gly	Ile	Asp		
				340					345					350			
Arg	Ile	Asn	Asn	Asp	Phe	Val	Thr	Asn	Gly	Lys	Asn	Asn	Leu	Tyr	Ile		
				355					360					365			
Leu	Ser	Lys	Val	Tyr	Pro	Asp	Met	Gln	Arg	Asn	Glu	Leu	Val	Lys	Lys		
				370					375					380			
Tyr	Tyr	Gln	Phe	Val	Val	Cys	Lys	Glu	Gly	Asn	Asn	Val	Gly	Ile	Asn		
385						390					395				400		

Thr Arg Lys Leu Lys Glu Ser Ile Ile Ser Gln His Pro Trp Ile Thr		
405	410	415
Thr Pro Gln Asp Asn Asn Lys Ala Asn Asp Tyr Glu Ser Cys Arg His		
420	425	430
Lys Leu Tyr Thr Ile Met Cys Phe Ile Leu Val Ala Glu Leu Asp Ala		
435	440	445
His Glu Ser Ile Arg Asp Asn Met Val Ala Glu Leu Arg Ala Asn Met		
450	455	460
Asp Gly Asp Asp Gly Arg Asp Ala Ile Tyr Glu Lys Tyr Ala Lys Asp		
465	470	475
Ile Tyr His Ile Val Lys Asp Lys Leu Leu Ala Met Gln Lys Val Phe		
485	490	495
Asp Glu Glu Leu Val Pro Val Lys Val Glu Gly Lys Asn Asp Pro Gln		
500	505	510
Gln Phe Thr His Gly Lys Leu Gly Lys Lys Glu Ile Glu Ser Phe Cys		
515	520	525
Leu Ser Asp Lys Asn Thr Ser Asp Ile Ala Lys Val Val Tyr Phe Leu		
530	535	540
Cys Asn Phe Leu Asp Gly Lys Glu Ile Asn Glu Leu Cys Cys Ala Met		
545	550	555
Met Asn Lys Phe Asp Gly Ile Gly Asp Leu Ile Asp Thr Ala Lys Gln		
565	570	575
Cys Gly Glu Glu Val Lys Phe Ile Glu Glu Phe Ala Cys Leu Ser Asn		
580	585	590
Cys Arg Lys Ile Thr Asn Asp Ile Arg Val Ala Lys Ser Ile Ser Lys		
595	600	605
Met Lys Asn Lys Val Asn Ile Asp Asn Asp Ile Ile Tyr Leu Asp Ala		
610	615	620
Ile Glu Leu Leu Gly Arg Lys Ile Glu Lys Tyr Gln Lys Asp Glu Asn		
625	630	635
Gly Lys Ile Leu Leu Gly Thr Asp Gly Lys Arg Leu Tyr Thr Gln Glu		
645	650	655
Tyr Lys Tyr Phe Asn Asp Met Phe Phe Asn Ala Gly Asn His Lys Val		
660	665	670
Arg Asn Phe Ile Ala Asn Asn Val Met Gln Ser Lys Trp Phe Phe Tyr		
675	680	685
Val Val Arg Tyr Asn Lys Pro Ala Glu Cys Gln Ile Ile Met Arg Asn		
690	695	700
Lys Thr Leu Val Lys Phe Thr Leu Asp Asp Leu Pro Asp Met Gln Ile		

705	710	715	720
Gln Arg Tyr Tyr Ser Ser Val Phe Gly Asp Asn Asn Met Pro Ala Val			
	725	730	735
Asp Glu Met Arg Lys Arg Leu Leu Asp Lys Ile Asn Gln Phe Ser Val			
	740	745	750
Arg Gly Phe Leu Asp Glu Leu Asp Glu Ile Val Leu Met Ser Asp Glu			
	755	760	765
Glu Ser Lys Arg Asn Lys Ser Ser Glu Lys Glu Gln Lys Lys Ser Leu			
	770	775	780
Ile Arg Leu Tyr Leu Thr Ile Ala Tyr Leu Ile Thr Lys Ser Met Val			
785	790	795	800
Lys Ile Asn Thr Arg Phe Ser Ile Ala Cys Ala Met Tyr Glu Arg Asp			
	805	810	815
Tyr Ala Leu Leu Cys Gln Ser Glu Met Lys Gly Gly Pro Trp Asp Gly			
	820	825	830
Gly Ala Gln Ala Leu Ala Val Thr Arg Lys Phe Leu Asn His Asp Arg			
	835	840	845
Glu Val Phe Asp Arg Tyr Cys Ala Arg Glu Ala Glu Ile Ala Arg Leu			
850	855	860	
Pro Ser Glu Glu Arg Lys Pro Leu Arg Lys Ala Asn Asp Lys Leu Leu			
865	870	875	880
Lys Gln Thr His Tyr Thr Asn His Ser Tyr Thr Tyr Ile Val Asn Asn			
	885	890	895
Leu Asn Ser Phe Thr Asp Ile Asp Tyr Cys Ala Lys Asp Val Gly Leu			
	900	905	910
Pro Ala Pro Asn Asp Lys Asn Asp Asn Ala Ser Ile Leu Gly Glu Met			
	915	920	925
Arg Asn Asp Ile Ala His Leu Asn Ile Val His Asp Met Val Lys Tyr			
930	935	940	
Ile Glu Glu Leu Lys Asp Ile Ser Ser Tyr Tyr Ala Phe Tyr Cys Tyr			
945	950	955	960
Val Leu Gln Arg Arg Leu Val Gly Lys Asp Pro Asn Cys Gln Asn Lys			
	965	970	975
Phe Lys Ala Lys Tyr Ala Lys Glu Leu Asn Asp Tyr Gly Thr Tyr Asn			
	980	985	990
Lys Asn Leu Met Trp Met Leu Asn Leu Pro Phe Ala Tyr Asn Leu Pro			
	995	1000	1005
Arg Tyr Lys Asn Leu Ser Ser Glu Phe Leu Phe Tyr Asp Met Glu			
1010	1015	1020	

Tyr Asn Lys Lys Asp Asp Glu
 1025 1030
 <210> 102
 <211> 1027
 <212> PRT
 <213> 未知的
 <220>
 <223> 重叠群emb OBLI01020244
 <400> 102
 Met Ala Lys Lys Ile Thr Ala Lys Gln Arg Arg Glu Glu Arg Glu Arg
 1 5 10 15
 Gln Asn Lys Gln Lys Trp Ala Lys Lys Gln Ala Asp Ala Thr Ala Val
 20 25 30
 Phe Glu Cys Glu Ala Asp Ile Lys Pro Ala Asp Ser Lys Asp Glu Asp
 35 40 45
 Cys Thr Asn Ile Tyr Ile Lys Arg Glu Lys Lys Lys Thr Gln Ala Lys
 50 55 60
 Ala Met Gly Leu Lys Thr Val Leu Gly Phe Asp Asn Lys Ile Ala Ile
 65 70 75 80
 Ala Ser Phe Met Ser Ser Lys Asp Ser Lys Ser Ser His Ile Glu Arg
 85 90 95
 Ile Thr Asp Pro Asn Gly Lys Thr Ile Arg Glu Asp Val Arg Met Phe
 100 105 110
 Asp Ser Asn Val Asp Glu Cys Ser Ile Asn Leu Glu Lys Arg Met Thr
 115 120 125
 Val Glu Glu Arg Gln Lys Asp Gly Thr Ile Lys Lys Asp Glu Lys Asp
 130 135 140
 Val Lys Ser Thr Ile Cys Asn Pro Tyr Ser Asn Glu Cys Gly Lys Asp
 145 150 155 160
 Tyr Ile Gly Ile Lys Ser Val Ala Glu Glu Leu Phe Phe Gly Arg Thr
 165 170 175
 Phe Pro Asn Asp Asn Leu Arg Val Gln Ile Ala Tyr Asn Ile Phe Asp
 180 185 190
 Ile Gln Lys Ile Leu Gly Thr Tyr Ile Asn Asn Ile Ile Tyr Ser Phe
 195 200 205
 Tyr Asn Leu Ser Arg Asp Glu Ser Gln Ser Asp Asn Asp Val Ile Gly
 210 215 220
 Thr Leu Tyr Met Leu Lys Asp Phe Asp Gly Gln Lys Glu Thr Asp Thr
 225 230 235 240

Phe Arg Gln Ala Arg Ala Leu Leu Glu Arg Thr Glu Ala Tyr Tyr Ser		
245	250	255
Tyr Phe Asp Asn Val Phe Lys Lys Ile Asp Lys Asn Lys Lys Lys Ser		
260	265	270
Asp Asp Cys Lys Arg Glu Arg Asn Glu Ile Leu Arg Tyr Asn Phe Asn		
275	280	285
Val Leu Arg Val Leu Ser Phe Leu Arg Gln Ile Cys Ala His Ala Gln		
290	295	300
Val Lys Ile Ser Asn Glu His Asp Arg Glu Lys Gly Gly Gly Leu Val		
305	310	315
Asp Ser Leu Asp Ala Leu Phe Asn Ile Ser Arg Phe Phe Asp Ala Val		
325	330	335
Ala Pro Glu Leu Asn Glu Val Ile Asn Ser Val Tyr Ser Lys Gly Ile		
340	345	350
Asp Asp Ile Asn Asp Asn Phe Val Lys Asn Gly Lys Asn Asn Phe Tyr		
355	360	365
Ile Leu Ser Lys Ile Tyr Pro Glu Val Ala Arg Glu Asp Leu Leu Arg		
370	375	380
Glu Tyr Tyr Tyr Phe Val Val Ser Lys Glu Gly Asn Asn Ile Gly Ile		
385	390	395
Ser Thr Lys Lys Leu Lys Glu Ala Ile Ile Val Gln Asp Met Ser Tyr		
405	410	415
Ile Lys Ser Glu Asp Tyr Asp Thr Tyr Arg Asn Lys Leu Tyr Thr Val		
420	425	430
Leu Cys Phe Ile Leu Val Lys Glu Leu Asn Glu Arg Thr Thr Ile Arg		
435	440	445
Glu Gln Met Val Ala Asp Leu Arg Ala Asn Met Asn Gly Asp Ile Gly		
450	455	460
Arg Glu Asp Ile Tyr Ser Lys Tyr Ala Lys Ile Ile Tyr Ala Gln Val		
465	470	475
Lys Pro Arg Phe Asp Thr Met Lys Ser Ala Phe Glu Glu Glu Ala Lys		
485	490	495
Asp Val Ile Val Pro Asp Lys Lys Lys Pro Val Lys Phe Ser His Gly		
500	505	510
Lys Leu Asp Lys Asn Glu Ile Glu Arg Phe Cys Ile Thr Ser Ala Asn		
515	520	525
Thr Asp Ser Val Ala Lys Ile Ile Tyr Phe Leu Cys Lys Phe Leu Asp		
530	535	540
Gly Lys Glu Ile Asn Glu Leu Cys Cys Ala Met Met Asn Lys Leu Asp		

545	550	555	560
Gly Ile Asn Asp Leu Ile Glu Thr Ala Glu Gln Cys Gly Ala Lys Val			
	565	570	575
Glu Phe Val Asp Lys Phe Ser Val Leu Ser Asn Cys Glu Thr Ile Ser			
	580	585	590
Asp Gln Ile Arg Ile Val Lys Ser Ile Ser Lys Met Lys Lys Glu Ile			
	595	600	605
Ala Ile Asp Asn Asp Thr Ile Phe Leu Asp Ala Leu Glu Leu Leu Gly			
	610	615	620
Arg Lys Ile Asp Lys Tyr Lys Lys Asp Ala Thr Gly Lys Tyr Leu Lys			
625	630	635	640
Asp Glu Asn Gly Lys Tyr Leu Tyr Ser Lys Glu Tyr Asp Asp Phe Gln			
	645	650	655
Tyr Met Phe Phe Lys Asp Ser His Arg Val Arg Asn Phe Ile Ser Asn			
	660	665	670
Ser Val Ile Lys Ser Lys Trp Phe Ser Tyr Ile Val Arg Tyr Asn Gln			
	675	680	685
Pro Ser Glu Cys Arg Ala Ile Met Lys Asn Lys Thr Leu Val Lys Phe			
	690	695	700
Ala Leu Asp Glu Leu Pro Asp Leu Gln Ile Gln Arg Tyr Phe Val Ala			
705	710	715	720
Leu Tyr Gly Asp Glu Asp Leu Pro Ser Tyr Gly Glu Met Arg Lys Ile			
	725	730	735
Leu Leu Lys Lys Leu His Asp Phe Ser Ile Lys Gly Phe Leu Asp Glu			
	740	745	750
Ile Val Leu Leu Ser Asp Leu Asp Met Glu Ser Gln Asp Lys Tyr Cys			
	755	760	765
Glu Lys Glu Gln Lys Lys Ser Leu Phe Arg Leu Tyr Leu Thr Ile Ala			
	770	775	780
Tyr Leu Ile Thr Lys Ser Met Val Lys Ile Asn Thr Arg Phe Ser Ile			
785	790	795	800
Ala Cys Ala Thr Tyr Glu Arg Asp Tyr Ala Leu Leu Cys Ala Ser Asn			
	805	810	815
Lys Gln Glu Arg Ala Trp Ser Ser Gly Ala Thr Ala Leu Ala Leu Thr			
	820	825	830
Arg Arg Phe Leu Asn Gln Asp Lys Leu Ile Phe Glu Lys His Tyr Ala			
	835	840	845
Arg Glu Gly Glu Ile Ser Lys Leu Pro Lys Glu Glu Arg Lys Ala Met			
	850	855	860

Arg Lys Val Asn Asp Gln Leu Leu Lys Arg Thr His Phe Ser Lys His
 865 870 875 880
 Ser Tyr Cys Tyr Ile Val Asp Asn Val Asn Arg Leu Thr Gly Gly Glu
 885 890 895
 Cys Arg Thr Asp Lys Arg Val Leu Pro Val Leu Asn Glu Lys Asn Asp
 900 905 910
 Asn Ala Gly Ile Leu Leu Asp Phe Arg Lys Thr Ile Ala His Leu Asn
 915 920 925
 Val Val His Lys Met Val Asp Tyr Val Asp Glu Ile Lys Gly Ile Thr
 930 935 940
 Ser Tyr Tyr Ala Phe Phe Cys Tyr Val Leu Gln Arg Met Leu Val Gly
 945 950 955 960
 Asn Asn Leu Asn Glu Lys Asn Ala Ile Lys Glu Lys Tyr Ser Ala Thr
 965 970 975
 Val Lys Ser Phe Gly Thr Tyr Ser Lys Asp Phe Met Trp Leu Ile Asn
 980 985 990
 Leu Pro Phe Ala Tyr Asn Leu Pro Arg Tyr Lys Asn Leu Ser Asn Glu
 995 1000 1005
 Gln Leu Phe Tyr Asp Glu Glu Glu Arg Asn Glu Thr Glu Glu Gln
 1010 1015 1020
 Ile Asp Arg Leu
 1025
 <210> 103
 <211> 961
 <212> PRT
 <213> 未知的
 <220>
 <223> 重叠群OIZX01000427.1
 <400> 103
 Met Ala Lys Lys Lys Lys Thr Ala Arg Gln Leu Arg Glu Glu Met Gln
 1 5 10 15
 Gln Gln Arg Lys Gln Ala Ile Gln Lys Gln Gln Glu Gln Arg Gln Glu
 20 25 30
 Lys Ala Ala Ala Ala Arg Glu Thr Ala Ala Pro Glu Gln Pro Ala Ala
 35 40 45
 Ala Pro Val Pro Lys Arg Gln Arg Lys Ser Leu Ala Lys Ala Ala Gly
 50 55 60
 Leu Lys Ser Asn Phe Ile Leu Asp Pro Gln Arg Arg Thr Thr Val Met
 65 70 75 80

Thr	Ala	Phe	Gly	Gln	Gly	Ser	Thr	Ala	Ile	Leu	Glu	Lys	Gln	Ile	Val		
85						90						95					
Asp	Arg	Ala	Ile	Ser	Asp	Leu	Gln	Pro	Val	Gln	Gln	Phe	Gln	Val	Glu		
100						105						110					
Pro	Ala	Ser	Ala	Ala	Lys	Tyr	Arg	Leu	Lys	Asn	Ser	Arg	Val	Arg	Phe		
115						120						125					
Pro	Asn	Val	Thr	Ala	Asp	Asp	Pro	Leu	Tyr	Arg	Arg	Lys	Asp	Gly	Gly		
130						135						140					
Phe	Val	Pro	Gly	Met	Asp	Ala	Leu	Arg	Arg	Lys	Asn	Val	Leu	Glu	Gln		
145						150						155				160	
Arg	Phe	Phe	Gly	Lys	Ser	Phe	Ala	Asp	Asn	Ile	His	Ile	Gln	Met	Ile		
165						170						175					
Tyr	Ser	Ile	Leu	Asp	Ile	His	Lys	Ile	Leu	Ala	Ala	Ala	Ser	Gly	His		
180						185						190					
Ile	Val	His	Leu	Leu	Asn	Ile	Val	Asn	Gly	Ser	Lys	Asp	Arg	Asp	Phe		
195						200						205					
Ile	Gly	Met	Leu	Ala	Ala	His	Val	Leu	Tyr	Asn	Glu	Leu	Asn	Glu	Glu		
210						215						220					
Ala	Lys	Arg	Ser	Ile	Ala	Asp	Phe	Cys	Lys	Ser	Pro	Arg	Leu	Ile	Tyr		
225						230						235				240	
Tyr	Ser	Ala	Ala	Phe	Tyr	Glu	Thr	Leu	Asp	Asn	Gly	Lys	Ser	Glu	Arg		
245						250						255					
Arg	Ser	Asn	Glu	Asp	Ile	Phe	Asn	Ile	Leu	Ala	Leu	Met	Thr	Cys	Leu		
260						265						270					
Arg	Asn	Phe	Ser	Ser	His	His	Ser	Ile	Ala	Ile	Lys	Val	Lys	Asp	Tyr		
275						280						285					
Ser	Ala	Ala	Gly	Leu	Tyr	Asn	Leu	Arg	Arg	Leu	Gly	Pro	Asp	Met	Lys		
290						295						300					
Lys	Met	Leu	Asp	Thr	Phe	Tyr	Thr	Glu	Ala	Phe	Ile	Gln	Leu	Asn	Gln		
305						310						315				320	
Ser	Phe	Gln	Asp	His	Asn	Thr	Thr	Asn	Leu	Thr	Cys	Leu	Phe	Asp	Ile		
325						330						335					
Leu	Asn	Ile	Ser	Asp	Ser	Ala	Arg	Gln	Lys	Gln	Leu	Ala	Glu	Glu	Phe		
340						345						350					
Tyr	Arg	Tyr	Val	Val	Phe	Lys	Glu	Gln	Lys	Asn	Leu	Gly	Phe	Ser	Val		
355						360						365					
Arg	Lys	Leu	Arg	Glu	Glu	Met	Leu	Leu	Leu	Pro	Asp	Ala	Ala	Val	Ile		
370						375						380					
Ala	Asp	Lys	Arg	Tyr	Asp	Thr	Cys	Arg	Ser	Lys	Leu	Tyr	Asn	Leu	Met		

385	390	395	400
Asp Phe Leu Ile Leu Arg Val Tyr Arg Thr Gly Arg Ala Asp Arg Cys			
	405	410	415
Asp Lys Leu Pro Glu Ala Leu Arg Ala Ala Leu Thr Asp Glu Glu Lys			
	420	425	430
Ala Val Val Tyr His Lys Glu Ala Leu Ser Leu Trp Asn Glu Met Arg			
	435	440	445
Thr Leu Ile Leu Asp Gly Leu Leu Pro Gln Met Thr Pro Glu Asn Leu			
	450	455	460
Ser Arg Leu Ser Gly Gln Lys Arg Lys Gly Glu Leu Ser Leu Asp Asp			
465	470	475	480
Ala Met Leu Lys Glu Cys Leu Tyr Glu Pro Gly Pro Val Pro Glu Asp			
	485	490	495
Ala Ala Pro Glu Glu Ala Asn Ala Glu Tyr Phe Cys Arg Met Ile Tyr			
	500	505	510
Leu Ala Thr Leu Phe Met Asp Gly Lys Glu Ile Asn Thr Leu Leu Thr			
	515	520	525
Thr Leu Ile Ser Lys Phe Glu Asn Ile Ala Ala Phe Leu Gln Thr Met			
	530	535	540
Glu Gln Leu Asn Ile Glu Ala Glu Leu Gly Pro Glu Tyr Ala Met Phe			
545	550	555	560
Thr Arg Ser Arg Ala Val Ala Glu Gln Leu Arg Val Ile Asn Ser Phe			
	565	570	575
Ala Leu Met Lys Lys Pro Gln Val Asn Ala Lys Gln Gln Leu Tyr Arg			
	580	585	590
Ala Ala Val Thr Leu Leu Gly Thr Glu Asp Pro Asp Gly Val Thr Asp			
	595	600	605
Glu Met Leu Cys Ile Asp Pro Val Thr Gly Lys Met Leu Pro Pro Asn			
	610	615	620
Gln Arg His His Gly Asp Thr Gly Leu Arg Asn Phe Ile Ala Asn Asn			
625	630	635	640
Val Val Glu Ser Arg Arg Phe Gln Tyr Leu Ile Arg Tyr Ser Asp Pro			
	645	650	655
Ala Gln Leu His Gln Leu Ala Ser Asn Lys Lys Leu Val Arg Phe Val			
	660	665	670
Leu Ser Ser Ile Pro Asp Thr Gln Ile Asn Arg Tyr Tyr Glu Thr Cys			
	675	680	685
Gly Gln Thr Arg Leu Ala Gly Arg Ala Ala Lys Val Glu Phe Leu Thr			
	690	695	700

Asp Met Ile Ala Ala Ile Arg Phe Asp Gln Phe Arg Asp Val Asn Gln
705 710 715 720
Lys Glu Arg Gly Ala Asn Thr Gln Lys Glu Arg Tyr Lys Ala Met Leu
725 730 735
Gly Leu Tyr Gln Thr Val Leu Tyr Leu Ala Val Lys Asn Leu Val Asn
740 745 750
Ile Asn Ala Arg Tyr Val Met Ala Phe His Cys Val Glu Arg Asp Met
755 760 765
Phe Leu Tyr Asp Gly Glu Leu Thr Asp Pro Lys Gly Glu Ser Val Ser
770 775 780
Ala Phe Leu Ala Val Asn Gly Lys Lys Gly Val Gln Pro Gln Tyr Leu
785 790 795 800
Leu Leu Thr Gln Leu Phe Ile Arg Arg Asp Tyr Leu Lys Arg Ser Ala
805 810 815
Cys Glu Gln Ile Gln His Asn Met Glu Asn Ile Ser Asp Arg Leu Leu
820 825 830
Arg Glu Tyr Arg Asn Ala Val Ala His Leu Asn Val Ile Ala His Leu
835 840 845
Ala Asp Tyr Ser Ala Asp Met Arg Glu Ile Thr Ser Tyr Tyr Gly Leu
850 855 860
Tyr His Tyr Leu Met Gln Arg His Leu Phe Lys Arg His Ala Trp Gln
865 870 875 880
Ile Arg Gln Pro Glu Arg Pro Thr Glu Glu Glu Gln Lys Leu Ile Glu
885 890 895
Gln Glu Gln Lys Gln Leu Ala Trp Glu Lys Ala Leu Phe Asp Lys Thr
900 905 910
Leu Gln Tyr His Ser Tyr Asn Lys Asp Leu Val Lys Ala Leu Asn Ala
915 920 925
Pro Phe Gly Tyr Asn Leu Ala Arg Tyr Lys Asn Leu Ser Ile Glu Pro
930 935 940
Leu Phe Ser Lys Glu Ala Ala Pro Ala Ala Glu Ile Lys Ala Thr His
945 950 955 960
Ala

<210> 104

<211> 911

<212> PRT

<213> 未知的

<220>

<223> 重叠群OCTW011587266.1

<400> 104

Met	Lys	Gln	Asn	Asp	Arg	Glu	Asn	Asn	Asn	Lys	Ile	Lys	Lys	Ser	Ala
1				5						10				15	
Ala	Lys	Ala	Val	Gly	Val	Lys	Ser	Leu	Ala	Arg	Leu	Ser	Asp	Gly	Ser
			20					25					30		
Thr	Val	Val	Ser	Ser	Phe	Gly	Lys	Gly	Ala	Ala	Ala	Glu	Leu	Glu	Ser
			35					40					45		
Leu	Ile	Thr	Gly	Gly	Glu	Ile	Arg	Lys	Leu	Ser	Asp	Lys	Ala	Ile	Leu
			50					55				60			
Glu	Ile	Thr	Asp	Asp	Thr	Gln	Asn	Lys	Asn	Ala	Tyr	Asn	Val	Lys	Ser
65					70					75				80	
Ser	Arg	Ile	Pro	Asn	Leu	Thr	Ala	Arg	Thr	Asp	Lys	Leu	Ser	Asp	Lys
				85					90					95	
Ser	Gly	Met	Asp	Asp	Leu	Gly	Phe	Lys	Arg	Glu	Leu	Glu	Leu	Glu	Val
			100						105					110	
Phe	Gly	Gln	Cys	Phe	Asp	Asp	Ser	Ile	His	Ile	Gln	Ile	Ala	His	Ala
			115					120					125		
Val	Phe	Asp	Ile	Gln	Lys	Ser	Leu	Ala	Ala	Val	Ile	Pro	Asn	Val	Leu
			130					135					140		
Tyr	Thr	Leu	Asn	Asn	Leu	Asp	Arg	Ser	Tyr	Ser	Thr	Asp	Asn	Thr	Ser
145					150					155				160	
Asp	Lys	Lys	Asp	Ile	Ile	Gly	Asn	Thr	Leu	Asn	Tyr	Gln	His	Ser	Tyr
			165						170					175	
Glu	Ser	Phe	Asn	Val	Glu	Lys	Arg	Gly	Glu	Phe	Thr	Glu	Tyr	Tyr	Asn
			180						185					190	
Ala	Ala	Lys	Asp	Arg	Phe	Ser	Tyr	Phe	Pro	Asp	Ile	Leu	Cys	Val	Leu
			195					200					205		
Glu	Lys	Val	Asn	Gly	Lys	Asp	Arg	Tyr	Gln	Pro	Lys	Ser	Glu	Lys	Asp
			210					215					220		
Ala	Phe	Asn	Val	Leu	Ser	Ser	Val	Asn	Met	Leu	Arg	Asn	Ser	Leu	Phe
225					230					235				240	
His	Phe	Ala	Pro	Lys	Ser	Asn	Asp	Gly	Lys	Ala	Arg	Ile	Ala	Val	Phe
			245						250					255	
Lys	Asn	Gln	Phe	Asp	Ser	Asp	Phe	Ser	His	Ile	Thr	Ser	Thr	Val	Asn
			260					265						270	
Lys	Ile	Tyr	Ser	Ala	Lys	Ile	Ala	Gly	Val	Asn	Glu	Asn	Phe	Leu	Asn
			275					280					285		
Asn	Glu	Gly	Asn	Asn	Leu	Tyr	Ile	Ile	Leu	Lys	Ala	Thr	Asn	Trp	Asp
			290					295					300		

Ile Lys Lys Ile Val Pro Gln Leu Tyr Arg Phe Ser Val Leu Lys Ser			
305	310	315	320
Asp Lys Asn Met Gly Phe Asn Met Arg Lys Leu Arg Glu Phe Ala Val			
	325	330	335
Glu Ser Lys Asn Ile Asp Leu Ser Arg Leu Asn Asp Lys Phe Leu Thr			
	340	345	350
Asn Asn Arg Lys Lys Leu Tyr Lys Val Ile Asp Phe Ile Ile Tyr Tyr			
	355	360	365
His Leu Asn Lys Val Leu Lys Asp Ser Phe Val Asp Asp Phe Val Ala			
	370	375	380
Ala Leu Arg Ala Ser Gln Ser Glu Glu Glu Lys Glu Lys Leu Tyr Ala			
385	390	395	400
Gln Tyr Ser Glu Arg Leu Phe Ala Asp Glu Gly Leu Lys Ser Ala Ile			
	405	410	415
Lys Lys Ala Val Asp Met Ile Ser Asp Thr Lys Ser Asn Ile Phe Lys			
	420	425	430
Met Lys Thr Pro Leu Asp Lys Ala Leu Ile Glu Asn Ile Lys Val Asn			
	435	440	445
Ser Asp Ala Ser Asp Phe Cys Lys Leu Ile Tyr Val Phe Thr Arg Phe			
450	455	460	
Leu Asp Gly Lys Glu Ile Asn Ile Leu Leu Asn Ser Leu Ile Lys Lys			
465	470	475	480
Phe Gln Asp Ile His Ser Phe Asn Thr Thr Val Lys Lys Leu Ser Glu			
	485	490	495
Asn Asn Leu Ile Ile Asn Ala Asp Tyr Val Asp Asp Tyr Ser Leu Phe			
	500	505	510
Glu Gln Ser Gly Thr Val Ala Arg Glu Leu Met Leu Ile Lys Ser Ile			
	515	520	525
Ser Lys Met Asp Phe Gly Leu Asp Asn Ile Asn Leu Ser Phe Met Tyr			
	530	535	540
Asp Asp Ala Leu Arg Thr Leu Gly Val Ser Asp Glu Asn Leu Pro Glu			
545	550	555	560
Val Lys Arg Glu Tyr Phe Gly Lys Thr Lys Asn Leu Ser Ala Tyr Ile			
	565	570	575
Arg Asn Asn Val Leu Glu Asn Arg Arg Phe Lys Tyr Val Ile Lys Tyr			
	580	585	590
Ile His Pro Ser Asp Val Gln Lys Ile Ala Cys Asn Lys Ala Ile Ala			
	595	600	605
Gly Phe Val Leu Asn Arg Met Pro Asp Thr Gln Ile Lys Arg Tyr Tyr			

610	615	620
Asp Ser Leu Ile Asn Lys Gly Ala Thr Asp Ile Gln Ala Gln Ala Lys		
625	630	635
Ala Leu Leu Asp Cys Ile Thr Gly Ile Ser Phe Asp Ala Ile Lys Asp		640
645	650	655
Asp Lys His Leu His Lys Ser Lys Glu Lys Ser Pro Gln Arg Ser Ala		
660	665	670
Asp Arg Glu Arg Lys Lys Ala Met Leu Thr Leu Tyr Tyr Thr Ile Val		
675	680	685
Tyr Ile Phe Val Lys Gln Met Leu His Ile Asn Ser Leu Tyr Thr Ile		
690	695	700
Gly Phe Phe Tyr Leu Glu Arg Asp Gln Arg Phe Ile Tyr Ser Arg Ala		
705	710	715
Lys Lys Glu Asn Lys Asn Pro Ser Lys Asn Ser Tyr Leu Asn Asp Phe		720
725	730	735
Arg Ser Val Thr Ala Tyr Phe Ile Pro Ser Glu Ile Met Lys Arg Ile		
740	745	750
Glu Lys Asn Glu Asn Lys Gly Phe Leu Glu Asp Phe Glu Ala Leu Trp		
755	760	765
Asn Ser Cys Gly Lys Thr Ser Arg Leu Arg Lys Glu Asp Val Leu Leu		
770	775	780
Tyr Ala Arg Tyr Ile Ser Pro Asp His Ala Leu Lys Asn Tyr Lys Met		
785	790	795
Ile Leu Asn Ser Tyr Arg Asn Lys Ile Ala His Ile Asn Val Ile Met		800
805	810	815
Ser Ala Gly Lys Tyr Thr Gly Gly Ile Lys Arg Met Asp Ser Tyr Phe		
820	825	830
Ser Val Phe Gln His Leu Val Gln Cys Asp Ile Leu Ser Asn Pro Asn		
835	840	845
Asn Lys Gly Lys Cys Phe Glu Ser Glu Ser Leu Lys Pro Leu Leu Leu		
850	855	860
Asp Met Lys Phe Asp Gly Thr Asp Glu Lys Leu Tyr Ser Lys Arg Leu		
865	870	875
Thr Arg Ala Leu Asn Ile Pro Phe Gly Tyr Asn Val Pro Arg Tyr Lys		880
885	890	895
Asn Leu Thr Phe Glu Lys Ile Tyr Leu Lys Ser Ser Ile Asn Glu		
900	905	910
<210> 105		
<211> 904		

<212> PRT

<213> 未知的

<220>

<223> 重叠群emb OGNF01009141.1

<400> 105

```

Met Ala Asp Ile Asp Lys Lys Lys Ser Ser Ala Lys Ala Ala Gly Leu
1           5           10           15
Lys Ser Thr Phe Val Leu Glu Asn Asn Lys Leu Leu Met Thr Ser Phe
           20           25           30
Gly Asn Gly Asn Lys Ala Val Ile Glu Lys Ile Ile Asp Glu Lys Val
           35           40           45
Asp Ser Ile Asn Glu Pro Glu Val Phe Ser Val Thr Pro Cys Asp Lys
           50           55           60
Lys Phe Glu Leu Gln Pro Ala Lys Arg Gly Leu Ala Ala Asp Ser Leu
65           70           75           80
Val Asp Asn Pro Leu Lys Ser Lys Lys Thr Ala Gly Asp Asp Ala Ile
           85           90           95
His Ser Arg Lys Phe Leu Glu Arg Gln Phe Phe Asp Gly Asn Thr Phe
           100          105          110
Asn Asp Asn Ile His Ile Gln Leu Ile Tyr Asn Ile Leu Asp Ile Glu
           115          120          125
Lys Ile Leu Ser Val His Val Asn Asp Ile Val Tyr Ser Val Asn Asn
           130          135          140
Ile Leu Ser Arg Gly Glu Gly Met Glu Tyr Asn Asp Tyr Ile Gly Thr
145          150          155          160
Leu Asn Leu Lys Ser Phe Glu Thr Tyr Lys Asn Asn Leu Val Asn Lys
           165          170          175
Lys Lys Phe Asp Leu Asp Arg Val Lys Lys Ile Pro Gln Leu Ala Tyr
           180          185          190
Phe Gly Ser Ala Phe Tyr Asn Thr Pro Glu Asp Thr Ser Ala Lys Ile
           195          200          205
Thr Lys Thr Lys Ile Lys Ser Asn Glu Glu Ile Tyr Tyr Thr Phe Met
           210          215          220
Leu Leu Ser Thr Ala Arg Asn Phe Ser Ala His Tyr Leu Asp Arg Asn
225          230          235          240
Arg Ala Lys Ser Ser Asp Ala Glu Asp Phe Asp Gly Thr Ser Val Ile
           245          250          255
Met Tyr Asn Leu Asp Asn Glu Glu Leu Tyr Lys Lys Leu Tyr Asn Lys
           260          265          270

```

Lys Val His Met Ala Leu Thr Gly Met Lys Lys Val Leu Asp Ala Asn		
275	280	285
Phe Asn Lys Lys Val Glu His Leu Asn Asn Ser Phe Ile Lys Asn Ser		
290	295	300
Ala Lys Asp Phe Val Ile Leu Cys Glu Val Leu Gly Ile Lys Ser Arg		
305	310	315
Asp Glu Lys Thr Lys Phe Val Lys Asp Tyr Tyr Asp Phe Val Val Arg		
	325	330
		335
Lys Asn Tyr Lys His Leu Gly Phe Ser Val Lys Glu Leu Arg Glu Leu		
	340	345
		350
Leu Phe Ala Asn His Asp Ser Asn Lys Tyr Ile Lys Glu Phe Asp Lys		
	355	360
		365
Ile Ser Asn Lys Lys Phe Asp Ser Val Arg Ser Arg Leu Asn Arg Leu		
370	375	380
Ala Asp Tyr Ile Ile Tyr Asp Tyr Tyr Asn Lys Asn Asn Ala Lys Val		
385	390	395
		400
Ser Asp Leu Val Lys Tyr Leu Arg Ala Ala Ala Asp Asp Glu Gln Lys		
	405	410
		415
Lys Lys Ile Tyr Leu Asn Glu Ser Ile Asn Leu Val Lys Ser Gly Ile		
	420	425
		430
Leu Glu Arg Ile Lys Lys Ile Leu Pro Lys Leu Asn Gly Lys Ile Ile		
	435	440
		445
Gly Asn Met Gln Pro Asp Ser Thr Ile Thr Ala Ser Met Leu His Asn		
450	455	460
Thr Gly Lys Asp Trp His Pro Ile Ser Glu Asn Ala His Tyr Phe Thr		
465	470	475
		480
Lys Trp Ile Tyr Thr Leu Thr Leu Phe Met Asp Gly Lys Glu Ile Asn		
	485	490
		495
Asp Leu Val Thr Thr Leu Ile Asn Lys Phe Asp Asn Ile Ala Ser Phe		
	500	505
		510
Ile Glu Val Leu Lys Ser Gln Ser Val Cys Thr His Phe Ser Glu Glu		
	515	520
		525
Arg Lys Met Phe Ile Asp Ser Ala Glu Ile Cys Ser Glu Leu Ser Ala		
	530	535
		540
Met Asn Ser Phe Ala Arg Met Glu Ala Pro Gly Ala Ser Ser Lys Arg		
545	550	555
		560
Ala Met Phe Val Glu Ala Ala Arg Ile Leu Gly Asp Asn Arg Ser Lys		
	565	570
		575
Glu Glu Leu Glu Glu Tyr Phe Asp Thr Leu Phe Asp Lys Ser Ala Ser		

580	585	590
Lys Lys Glu Lys Gly Phe Arg Asn Phe Ile Arg Asn Asn Val Val Asp		
595	600	605
Ser Asn Arg Phe Lys Tyr Leu Thr Arg Tyr Thr Asp Thr Ser Ser Val		
610	615	620
Lys Ala Phe Ser Asn Asn Lys Ala Leu Val Lys Phe Ala Ile Lys Asp		
625	630	635
Ile Pro Gln Glu Gln Ile Leu Arg Tyr Tyr Asn Ser Cys Phe Gly Ala		
645	650	655
Ser Glu Arg Tyr Tyr Asn Asp Gly Met Ser Asp Lys Leu Val Glu Ala		
660	665	670
Ile Gly Lys Ile Asn Leu Met Gln Phe Asn Gly Val Ile Gln Gln Ala		
675	680	685
Asp Arg Asn Met Leu Pro Glu Glu Lys Lys Lys Ala Asn Ala Gln Lys		
690	695	700
Glu Lys Tyr Lys Ser Ile Ile Arg Leu Tyr Leu Thr Val Cys Tyr Leu		
705	710	715
Phe Phe Lys Asn Leu Val Tyr Val Asn Ser Arg Tyr Tyr Ser Ala Phe		
725	730	735
Tyr Asn Leu Glu Lys Asp Arg Ser Leu Phe Glu Ile Asn Gly Glu Leu		
740	745	750
Lys Pro Thr Gly Lys Phe Asp Glu Gly His Tyr Thr Gly Leu Val Lys		
755	760	765
Leu Phe Ile Asp Asn Gly Trp Ile Asn Pro Arg Ala Ser Ala Tyr Leu		
770	775	780
Thr Val Asn Leu Ala Asn Ser Asp Glu Thr Ala Ile Arg Thr Phe Arg		
785	790	795
Asn Thr Ala Glu His Leu Glu Ala Leu Arg Asn Ala Asp Lys Tyr Leu		
805	810	815
Asn Asp Leu Lys Gln Phe Asp Ser Tyr Phe Glu Ile Tyr His Tyr Ile		
820	825	830
Thr Gln Arg Asn Ile Lys Glu Lys Cys Glu Met Leu Lys Glu Gln Thr		
835	840	845
Val Lys Tyr Asn Asn Asp Leu Leu Lys Tyr His Gly Tyr Ser Lys Asp		
850	855	860
Phe Val Lys Ala Leu Cys Val Pro Phe Gly Tyr Asn Leu Pro Arg Phe		
865	870	875
Lys Asn Leu Ser Ile Asp Ala Leu Phe Asp Lys Asn Asp Lys Arg Glu		
885	890	895

Lys Leu Lys Lys Gly Phe Glu Asp

900

<210> 106

<211> 933

<212> PRT

<213> 未知的

<220>

<223> 重叠群emb 0IEN01002196.1

<400> 106

Met Glu Arg Gln Lys Arg Lys Met Lys Ser Lys Ser Lys Met Ala Gly

1 5 10 15

Val Lys Ser Val Phe Val Ile Gly Asp Glu Leu Leu Met Thr Ser Phe

20 25 30

Gly Asp Gly Asp Asp Ala Val Leu Glu Lys Asp Ile Asp Glu Asn Gly

35 40 45

Val Val Asn Asp Cys Arg Asn Pro Ala Ala Tyr Asp Ala Val Tyr Gly

50 55 60

Thr Asp Ser Ile Arg Val Lys Lys Thr Asn Asn Asn Ile Arg Ala Lys

65 70 75 80

Val Asn Asn Pro Leu Ala Lys Ser Asn Ile Arg Ser Glu Glu Ser Ala

85 90 95

Leu Phe Arg Thr Arg Val Asn Glu Tyr Lys Arg Glu Gln Lys Asp Lys

100 105 110

Tyr Glu Thr Leu Phe Phe Gly Lys Thr Phe Asp Asp Asn Ile His Ile

115 120 125

Gln Leu Ile Ser Lys Ile Leu Asp Ile Glu Lys Thr Phe Ser Val Val

130 135 140

Ile Gly Asn Ile Val Tyr Ala Ile Asn Asn Leu Ser Leu Glu Gln Ser

145 150 155 160

Ile Asp Arg Pro Ile Asp Ile Phe Gly Asp Lys Asn Thr Gln Gly Ile

165 170 175

Ser Leu Arg Glu Asp Asn Asp Tyr Leu Lys Thr Met Leu Pro Arg Cys

180 185 190

Glu Tyr Leu Phe His Asn Ile Leu Asn Ser Asp Ser Asp Asn Asn Ser

195 200 205

Lys Met Asn Tyr Asn Lys Val Asn Lys Gly Lys Glu Glu Lys Asp Asn

210 215 220

Arg Asn Asn Glu Asn Ile Glu Lys Leu Lys Lys Ala Leu Glu Val Ile

225 230 235 240

Lys Ile Ile Arg Val Asp Ser Phe His Gly Val Asp Gly Ile Lys Gly		
245	250	255
Asp Gln Lys Phe Pro Arg Ser Lys Tyr Asn Leu Ala Val Asn Tyr Asn		
260	265	270
Glu Glu Ile Gln Lys Thr Ile Ser Glu Pro Phe Asn Arg Lys Val Glu		
275	280	285
Glu Val Gln Gln Asp Phe Tyr Arg Asn Ser Cys Val Asn Ile Asp Phe		
290	295	300
Leu Lys Glu Ile Met Tyr Gly Ser Asn Tyr Thr Asp Arg Gly Ser Asp		
305	310	315
Ser Leu Glu Cys Ser Tyr Phe Asn Phe Ala Ile Leu Lys Gln Asn Lys		
325	330	335
Asn Met Gly Phe Ser Ile Thr Ser Ile Arg Glu Cys Leu Leu Asp Leu		
340	345	350
Tyr Glu Leu Asn Phe Glu Ser Met Gln Asn Leu Arg Pro Arg Ala Asn		
355	360	365
Ser Phe Cys Asp Phe Leu Ile Tyr Asp Tyr Tyr Cys Lys Asn Glu Ser		
370	375	380
Glu Arg Ala Asn Leu Val Asp Cys Leu Arg Ser Ala Ala Ser Glu Glu		
385	390	395
Glu Lys Lys Asn Ile Tyr Phe Gln Thr Ala Glu Arg Val Lys Glu Lys		
405	410	415
Phe Arg Asn Ala Phe Asn Arg Ile Ser Arg Phe Asp Ala Ser Tyr Ile		
420	425	430
Lys Asn Ser Arg Glu Lys Asn Leu Ser Gly Gly Ser Ser Leu Pro Lys		
435	440	445
Tyr Ser Phe Ile Glu Gly Phe Thr Lys Arg Ser Lys Lys Ile Asn Asp		
450	455	460
Asn Asp Glu Lys Asn Ala Asp Leu Phe Cys Asn Met Leu Tyr Tyr Leu		
465	470	475
Ala Gln Phe Leu Asp Gly Lys Glu Ile Asn Ile Phe Leu Thr Ser Ile		
485	490	495
His Asn Ile Phe Gln Asn Ile Asp Ser Phe Leu Lys Val Met Lys Glu		
500	505	510
Lys Gly Met Glu Cys Lys Phe Gln Lys Asp Phe Lys Met Phe Ser His		
515	520	525
Ala Gly His Val Ala Lys Lys Ile Glu Ile Val Ile Ser Leu Ala Lys		
530	535	540
Met Lys Lys Thr Leu Asp Phe Tyr Asn Ala Gln Ala Leu Lys Asp Ala		

545	550	555	560
Val Thr Ile Leu Gly	Val Ser Lys Lys His	Gln Tyr Leu Asp Met	Asn
565	570	575	
Ser Tyr Leu Asp Phe	Tyr Met Phe Asp Asn	Arg Ser Gly Ala Thr	Gly
580	585	590	
Lys Asn Ala Gly Lys	Asp His Asn Leu Arg	Asn Phe Leu Val Ser	Asn
595	600	605	
Val Ile Arg Ser Arg	Lys Phe Asn Tyr Leu	Ser Arg Tyr Ser Asn	Leu
610	615	620	
Ala Glu Val Lys Lys	Leu Ala Gln Asn Pro	Ser Leu Val Gln Phe	Val
625	630	635	640
Leu Ser Arg Ile Glu	Pro Ser Leu Ile Cys	Arg Tyr Tyr Glu Ser	Ser
645	650	655	
Gln Gly Ile Ser Ser	Glu Gly Ile Thr Ile	Asp Glu Gln Ile Lys	Lys
660	665	670	
Leu Thr Gly Ile Ile	Val Asp Met Asn Ile	Asp Ser Phe Glu Asn	Ile
675	680	685	
Asn Asn Gly Glu Ile	Gly Met Arg Tyr Ser	Lys Ala Thr Pro Gln	Ser
690	695	700	
Ile Glu Arg Arg Asn	Gln Met Arg Val Cys	Val Gly Leu Tyr Leu	Asn
705	710	715	720
Val Leu Tyr Gln Ile	Glu Lys Asn Leu Met	Asn Val Asn Ala Arg	Tyr
725	730	735	
Val Leu Ala Phe Ala	Phe Ala Glu Arg Asp	Ala Leu Met Leu Asn	Phe
740	745	750	
Thr Leu Glu Glu Cys	Lys Lys Asn Lys Lys	Arg Ser Ser Gly Gly	Phe
755	760	765	
Ser Phe Ile Glu Met	Thr Gln Phe Phe Ile	Asp Lys Lys Leu Phe	Lys
770	775	780	
Val Ala Thr Glu Ala	Ile Lys Lys Asn Val	Leu Lys Tyr Asn Gly	Asn
785	790	795	800
Pro Glu Ser Leu Asn	His Ile Pro Gly Glu	Tyr Ile Cys Lys Asn	Met
805	810	815	
Glu Gly Tyr His Glu	Asn Thr Val Arg Asn	Phe Arg Asn Met Val	Ala
820	825	830	
His Leu Thr Ala Val	Ala Arg Val Pro Leu	Tyr Ile Ser Glu Val	Thr
835	840	845	
Gln Ile Asp Ser Tyr	Tyr Ala Leu Tyr His	Tyr Cys Met Gln Met	Asn
850	855	860	

Ile Leu Gln Gly Ile Glu Gln Ser Gly Lys Ile Leu Asp Asn Ile Lys
 865 870 875 880
 Leu Lys Asn Ala Leu Glu Asn Ala Arg Val His Arg Thr Tyr Ser Lys
 885 890 895
 Asp Ala Val Lys Tyr Leu Cys Leu Pro Phe Ala Tyr Asn Ile Ser Arg
 900 905 910
 Tyr Lys Ala Leu Thr Ile Lys Asp Leu Phe Asp Trp Thr Glu Tyr Ser
 915 920 925
 Cys Lys Lys Asp Glu
 930
 <210> 107
 <211> 1034
 <212> PRT
 <213> 未知的
 <220>
 <223> 重叠群e-k87 11092736
 <400> 107
 Met Lys Arg Gln Lys Thr Phe Ala Lys Arg Ile Gly Ile Lys Ser Thr
 1 5 10 15
 Val Ala Tyr Gly Gln Gly Lys Tyr Ala Ile Thr Thr Phe Gly Lys Gly
 20 25 30
 Ser Lys Ala Glu Ile Ala Val Arg Ser Ala Asp Pro Pro Glu Glu Thr
 35 40 45
 Leu Pro Thr Glu Ser Asp Ala Thr Leu Ser Ile His Ala Lys Phe Ala
 50 55 60
 Lys Ala Gly Arg Asp Gly Arg Glu Phe Lys Cys Gly Asp Val Asp Glu
 65 70 75 80
 Thr Arg Ile His Thr Ser Arg Ser Glu Tyr Glu Ser Leu Ile Ser Asn
 85 90 95
 Pro Ala Glu Ser Pro Arg Glu Asp Tyr Leu Gly Leu Lys Gly Thr Leu
 100 105 110
 Glu Arg Lys Phe Phe Gly Asp Glu Tyr Pro Lys Asp Asn Leu Arg Ile
 115 120 125
 Gln Ile Ile Tyr Ser Ile Leu Asp Ile Gln Lys Ile Leu Gly Leu Tyr
 130 135 140
 Val Glu Asp Ile Leu His Phe Val Asp Gly Leu Gln Asp Glu Pro Glu
 145 150 155 160
 Asp Leu Val Gly Leu Gly Leu Gly Asp Glu Lys Met Gln Lys Leu Leu
 165 170 175

Ser Lys Ala Leu Pro Tyr Met Gly Phe Phe Gly Ser Thr Asp Val Phe		
180	185	190
Lys Val Thr Lys Lys Arg Glu Glu Arg Ala Ala Ala Asp Glu His Asn		
195	200	205
Ala Lys Val Phe Arg Ala Leu Gly Ala Ile Arg Gln Lys Leu Ala His		
210	215	220
Phe Lys Trp Lys Glu Ser Leu Ala Ile Phe Gly Ala Asn Ala Asn Met		
225	230	235
Pro Ile Arg Phe Phe Gln Gly Ala Thr Gly Gly Arg Gln Leu Trp Asn		
245	250	255
Asp Val Ile Ala Pro Leu Trp Lys Lys Arg Ile Glu Arg Val Arg Lys		
260	265	270
Ser Phe Leu Ser Asn Ser Ala Lys Asn Leu Trp Val Leu Tyr Gln Val		
275	280	285
Phe Lys Asp Asp Thr Asp Glu Lys Lys Lys Ala Arg Ala Arg Gln Tyr		
290	295	300
Tyr His Phe Ser Val Leu Lys Glu Gly Lys Asn Leu Gly Phe Asn Leu		
305	310	315
Thr Lys Thr Arg Glu Tyr Phe Leu Asp Lys Phe Phe Pro Ile Phe His		
325	330	335
Ser Ser Ala Pro Asp Val Lys Arg Lys Val Asp Thr Phe Arg Ser Lys		
340	345	350
Phe Tyr Ala Ile Leu Asp Phe Ile Ile Tyr Glu Ala Ser Val Ser Val		
355	360	365
Ala Asn Ser Gly Gln Met Gly Lys Val Ala Pro Trp Lys Gly Ala Ile		
370	375	380
Asp Asn Ala Leu Val Lys Leu Arg Glu Ala Pro Asp Glu Glu Ala Lys		
385	390	395
Glu Lys Ile Tyr Asn Val Leu Ala Ala Ser Ile Arg Asn Asp Ser Leu		
405	410	415
Phe Leu Arg Leu Lys Ser Ala Cys Asp Lys Phe Gly Ala Glu Gln Asn		
420	425	430
Arg Pro Val Phe Pro Asn Glu Leu Arg Asn Asn Arg Asp Ile Arg Asn		
435	440	445
Val Arg Ser Glu Trp Leu Glu Ala Thr Gln Asp Val Asp Ala Ala Ala		
450	455	460
Phe Val Gln Leu Ile Ala Phe Leu Cys Asn Phe Leu Glu Gly Lys Glu		
465	470	475
Ile Asn Glu Leu Val Thr Ala Leu Ile Lys Lys Phe Glu Gly Ile Gln		

				485				490				495			
Ala	Leu	Ile	Asp	Leu	Leu	Arg	Asn	Leu	Glu	Gly	Val	Asp	Ser	Ile	Arg
				500				505				510			
Phe	Glu	Asn	Glu	Phe	Ala	Leu	Phe	Asn	Asp	Asp	Lys	Gly	Asn	Met	Ala
				515				520				525			
Gly	Arg	Ile	Ala	Arg	Gln	Leu	Arg	Leu	Leu	Ala	Ser	Val	Gly	Lys	Met
				530				535				540			
Lys	Pro	Asp	Met	Thr	Asp	Ala	Lys	Arg	Val	Leu	Tyr	Lys	Ser	Ala	Leu
545				550				555				560			
Glu	Ile	Leu	Gly	Ala	Pro	Pro	Asp	Glu	Val	Ser	Asp	Glu	Trp	Leu	Ala
				565				570				575			
Glu	Asn	Ile	Leu	Leu	Asp	Lys	Ser	Asn	Asn	Asp	Tyr	Gln	Lys	Ala	Lys
				580				585				590			
Lys	Thr	Val	Asn	Pro	Phe	Arg	Asn	Tyr	Ile	Ala	Lys	Asn	Val	Ile	Thr
				595				600				605			
Ser	Arg	Ser	Phe	Tyr	Tyr	Leu	Val	Arg	Tyr	Ala	Lys	Pro	Thr	Ala	Val
610				615				620							
Arg	Lys	Leu	Met	Ser	Asn	Pro	Lys	Ile	Val	Arg	Tyr	Val	Leu	Lys	Arg
625				630				635				640			
Leu	Pro	Glu	Lys	Gln	Val	Ala	Ser	Tyr	Tyr	Ser	Ala	Ile	Trp	Thr	Gln
				645				650				655			
Ser	Glu	Ser	Asn	Ser	Asn	Glu	Met	Val	Lys	Leu	Ile	Glu	Met	Ile	Asp
				660				665				670			
Arg	Leu	Thr	Thr	Glu	Ile	Ala	Gly	Phe	Ser	Phe	Ala	Val	Leu	Lys	Asp
				675				680				685			
Lys	Lys	Asp	Ser	Ile	Val	Ser	Ala	Ser	Arg	Glu	Ser	Arg	Ala	Val	Asn
690				695				700							
Leu	Glu	Val	Glu	Arg	Leu	Lys	Lys	Leu	Thr	Thr	Leu	Tyr	Met	Ser	Ile
705				710				715				720			
Ala	Tyr	Ile	Ala	Val	Lys	Ser	Leu	Val	Lys	Val	Asn	Ala	Arg	Tyr	Phe
				725				730				735			
Ile	Ala	Tyr	Ser	Ala	Leu	Glu	Arg	Asp	Leu	Tyr	Phe	Phe	Asn	Glu	Lys
				740				745				750			
Tyr	Gly	Glu	Glu	Phe	Arg	Leu	His	Phe	Ile	Pro	Tyr	Glu	Leu	Asn	Gly
				755				760				765			
Lys	Thr	Cys	Gln	Phe	Glu	Tyr	Leu	Ala	Ile	Leu	Lys	Tyr	Tyr	Leu	Ala
770				775				780							
Arg	Asp	Glu	Glu	Thr	Leu	Lys	Arg	Lys	Cys	Glu	Ile	Cys	Glu	Glu	Ile
785				790				795				800			

Lys Val Gly Cys Glu Lys His Lys Lys Asn Ala Asn Pro Pro Tyr Glu		
805	810	815
Tyr Asp Gln Glu Trp Ile Asp Lys Lys Lys Ala Leu Asn Ser Glu Arg		
820	825	830
Lys Ala Cys Glu Arg Arg Leu His Phe Ser Thr His Trp Ala Gln Tyr		
835	840	845
Ala Thr Lys Arg Asp Glu Asn Met Ala Lys His Pro Gln Lys Trp Tyr		
850	855	860
Asp Ile Leu Ala Ser His Tyr Asp Glu Leu Leu Ala Leu Gln Ala Thr		
865	870	875
Gly Trp Leu Ala Thr Gln Ala Arg Asn Asp Ala Glu His Leu Asn Pro		
885	890	895
Val Asn Glu Phe Asp Val Tyr Ile Glu Asp Leu Arg Arg Tyr Pro Glu		
900	905	910
Gly Thr Pro Lys Asn Lys Asp Tyr His Ile Gly Ser Tyr Phe Glu Ile		
915	920	925
Tyr His Tyr Ile Arg Gln Arg Ala Tyr Leu Glu Glu Val Leu Ala Lys		
930	935	940
Arg Lys Glu Tyr Arg Asp Ser Gly Ser Phe Thr Asp Glu Gln Leu Asp		
945	950	955
Lys Leu Gln Lys Ile Leu Asp Asp Ile Arg Ala Arg Gly Ser Tyr Asp		
965	970	975
Lys Asn Leu Leu Lys Leu Glu Tyr Leu Pro Phe Ala Tyr Asn Leu Pro		
980	985	990
Arg Tyr Lys Asn Leu Thr Thr Glu Ala Leu Phe Asp Asp Asp Ser Val		
995	1000	1005
Ser Gly Lys Lys Arg Val Ala Glu Trp Arg Glu Arg Glu Lys Thr		
1010	1015	1020
Arg Glu Ala Glu Arg Glu Gln Arg Arg Gln Arg		
1025	1030	
<210> 108		
<211> 30		
<212> DNA		
<213> 未知的		
<220>		
<223> CasRX/Cas13d同向重复1		
<400> 108		
gtgagaagtc tccttatggg gagatgctac 30		
<210> 109		

<211> 973

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d Ga0129306 1000735

<400> 109

Met	Gln	Lys	Gln	Arg	Glu	Gln	Gln	Thr	Val	Thr	Asp	Glu	Ser	Glu	Arg
1				5					10					15	
Lys	Lys	Lys	Pro	Leu	Lys	Ser	Gly	Ala	Lys	Ala	Ala	Gly	Leu	Lys	Ser
			20					25					30		
Val	Phe	Val	Leu	Ser	Glu	Gly	Lys	Glu	Leu	Leu	Thr	Ser	Phe	Gly	Arg
		35					40					45			
Gly	Asn	Glu	Ala	Val	Pro	Glu	Lys	Arg	Val	Thr	Gly	Gly	Thr	Ile	Ala
	50					55					60				
Asn	Ala	Arg	Thr	Asp	Asn	Lys	Glu	Ala	Phe	Ser	Ala	Ala	Leu	Gln	Asn
65				70					75					80	
Lys	Arg	Phe	Glu	Val	Phe	Gly	Arg	Thr	Ala	Gly	Ser	Ser	Asp	Asp	Pro
			85					90					95		
Leu	Ala	Val	Ser	Arg	Ala	Pro	Gly	Gln	Asp	Leu	Ile	Gly	Ala	Lys	Thr
		100						105					110		
Ala	Leu	Glu	Glu	Arg	Tyr	Phe	Gly	Arg	Ala	Phe	Ala	Asp	Asn	Ile	His
		115					120					125			
Met	Gln	Val	Ile	Tyr	Ala	Ile	Gln	Asp	Ile	Asn	Lys	Ile	Leu	Ala	Val
	130					135					140				
His	Ala	Asn	Asn	Ile	Val	Tyr	Thr	Leu	Asn	Asn	Leu	Asp	Arg	Glu	Ala
145				150					155					160	
Asp	Pro	Glu	Thr	Asp	Asp	Phe	Ile	Gly	Ser	Gly	Tyr	Leu	Thr	Leu	Lys
			165					170				175			
Asn	Thr	Phe	Glu	Thr	Tyr	Cys	Asp	Pro	Ala	Ala	Leu	Asn	Glu	Arg	Glu
		180					185				190				
Arg	Glu	Lys	Val	Thr	Val	Ser	Lys	Gln	His	Phe	Asp	Ala	Phe	Met	Gln
		195					200				205				
Asn	Pro	Arg	Leu	Ala	Tyr	Tyr	Gly	Asn	Ala	Phe	Phe	Arg	Lys	Leu	Ser
	210					215				220					
Lys	Ala	Glu	Arg	Leu	Ala	Arg	Gly	Arg	Glu	Ile	Phe	Asp	Lys	Glu	Ser
225				230					235					240	
Pro	Glu	Arg	Arg	Gln	Glu	Ile	Leu	Gly	Ser	Arg	Gly	Lys	Asn	Lys	Ser
			245				250				255				
Val	Asp	Asp	Glu	Ile	Arg	Ala	Leu	Ala	Pro	Glu	Trp	Val	Lys	Arg	Glu

260	265	270
Glu Arg Asp Val Tyr Ser Glu Leu Val Leu Met Ser Glu Leu Arg Gln		
275	280	285
Ser Cys Phe His Gly Gln Gln Lys Asn Ser Ala Arg Ile Phe Arg Leu		
290	295	300
Asp Asn Asp Leu Gly Pro Gly Val Asp Gly Ala Arg Glu Leu Leu Asp		
305	310	315
Arg Leu Tyr Ala Glu Lys Ile Asn Asp Leu Arg Ser Phe Asp Lys Thr		
325	330	335
Ser Ala Ser Ser Asn Phe Arg Leu Leu Phe Asn Ala Tyr His Ala Asp		
340	345	350
Asn Glu Lys Lys Lys Glu Leu Ala Gln Glu Phe Tyr Arg Phe Ser Val		
355	360	365
Leu Lys Val Ser Lys Asn Thr Gly Phe Ser Ile Arg Thr Leu Arg Glu		
370	375	380
Lys Ile Ile Glu Asp His Ala Ala Gln Tyr Arg Asp Lys Ile Tyr Asp		
385	390	395
Ser Met Arg Lys Lys Leu Phe Ser Thr Phe Asp Phe Phe Leu Trp Arg		
405	410	415
Phe Tyr Glu Glu Arg Glu Asp Glu Ala Glu Glu Leu Arg Ala Cys Leu		
420	425	430
Arg Ala Ala Arg Ser Asp Glu Glu Lys Glu Gln Ile Tyr Ala Glu Ala		
435	440	445
Ala Ala Ser Cys Trp Pro Ser Val Lys Pro Phe Val Glu Ser Val Ala		
450	455	460
Ala Thr Leu Cys Asp Val Val Lys Gly Arg Thr Lys Leu Asn Lys Leu		
465	470	475
Lys Leu Ser Ala Asp Glu Ser Thr Leu Val Arg Asn Ala Ile Asp Gly		
485	490	495
Val Arg Ile Ser Pro Arg Ala Ser Tyr Phe Thr Lys Leu Ile Tyr Leu		
500	505	510
Met Thr Leu Phe Leu Asp Gly Lys Glu Ile Asn Asp Leu Leu Thr Thr		
515	520	525
Leu Ile His Ala Phe Glu Asn Ile Asp Ser Phe Leu Ser Val Leu Gly		
530	535	540
Ser Glu Arg Leu Glu Arg Thr Phe Asp Ala Asn Tyr Arg Ile Phe Ala		
545	550	555
Asp Ser Gly Val Ile Ala Gln Glu Leu Arg Ala Val Asn Ser Phe Ala		
565	570	575

Arg Met Thr Thr Glu Pro Phe Asn Ser Lys Leu Val Met Phe Glu Asp		
580	585	590
Ala Ala Gln Leu Phe Gly Met Ser Gly Gly Leu Val Glu His Ala Glu		
595	600	605
Glu Leu Arg Glu Tyr Leu Asp Asn Lys Met Leu Asp Lys Thr Lys Leu		
610	615	620
Arg Leu Leu Pro Asp Gly Lys Val Asp Thr Gly Phe Arg Asn Phe Ile		
625	630	635
Ile Ser Asn Val Thr Glu Ser Arg Arg Phe Arg Tyr Leu Val Arg Tyr		
645	650	655
Cys Glu Pro Arg Ala Val Arg Asp Tyr Met Ser Cys Arg Pro Leu Ile		
660	665	670
Arg Leu Thr Leu Arg Asp Met Pro Asp Thr Ile Leu Arg Arg Tyr Tyr		
675	680	685
Glu Gln Ser Val Gly Ala Ala Thr Val Asp Arg Glu Arg Ile Leu Asp		
690	695	700
Thr Leu Ala Asp Lys Leu Leu Ser Leu Arg Phe Thr Asp Phe Glu Asn		
705	710	715
Val Asn Gln Arg Ala Asn Ala Glu Arg Asn Arg Glu Lys Gln Lys Met		
725	730	735
Met Gly Ile Ile Ser Leu Tyr Leu Asn Val Ala Tyr Gln Ile Val Lys		
740	745	750
Asn Leu Val Tyr Val Asn Ala Arg Tyr Thr Met Ala Tyr His Cys Ala		
755	760	765
Glu Arg Asp Thr Glu Leu Leu Leu Asn Ala Ala Gly Glu Gly Asn Leu		
770	775	780
Leu Arg Arg Asp Arg Ser Trp Pro Ala Arg Leu His Leu Pro Arg Arg		
785	790	795
Ala Leu Ala Arg Arg Arg Asp Arg Val Glu Val Met Glu Arg Asp Val		
805	810	815
Ala Arg Gly Pro Glu Ala Tyr Asn Arg Asp Glu Trp Leu Gly Leu Val		
820	825	830
Arg Thr Leu Arg Arg Glu Lys Arg Val Cys Asp Asn Leu His Asn Asn		
835	840	845
Tyr Ala Tyr Leu Cys Gly Ala Asp Ala Glu Pro Gly Asp Ala Ser Leu		
850	855	860
Ser Leu Leu Phe Val Tyr Arg Asn Lys Ala Ala His Leu Ser Val Leu		
865	870	875
Asn Lys Gly Gly Arg Leu Ser Gly Asp Leu Lys Glu Ala Lys Ser Trp		

				885					890					895			
Phe	Tyr	Val	Tyr	His	Phe	Leu	Met	Gln	Arg	Val	Leu	Glu	Glu	Glu	Phe		
				900					905					910			
Arg	Asn	Thr	Gln	Ala	Leu	Pro	Glu	Arg	Leu	Arg	Glu	Leu	Leu	Met	Met		
			915					920						925			
Ala	Glu	Arg	Tyr	Arg	Gly	Cys	Ser	Lys	Asp	Leu	Ile	Lys	Val	Leu	Asn		
			930					935						940			
Leu	Thr	Phe	Ala	Tyr	Asn	Leu	Pro	Arg	Tyr	Lys	Asn	Leu	Ser	Ile	Asp		
945						950					955				960		
Gly	Arg	Phe	Asp	Lys	Asn	His	Pro	Asp	Pro	Ser	Asp	Glu					
				965					970								
<210> 110																	
<211> 854																	
<212> PRT																	
<213> 未知的																	
<220>																	
<223> CasRX/Cas13d Ga0129317 1008067																	
<400> 110																	
Met	Lys	Lys	Gln	Lys	Lys	Ser	Leu	Val	Lys	Ala	Ala	Gly	Leu	Lys	Ser		
1				5					10					15			
Ala	Phe	Val	Val	Gly	Asp	Ser	Val	Tyr	Leu	Thr	Ser	Phe	Gly	Lys	Gly		
			20					25					30				
Asn	Ala	Ala	Arg	Leu	Asp	Thr	Lys	Ile	Asn	Pro	Asp	Asn	Ser	Thr	Glu		
			35				40					45					
Arg	Tyr	Val	Ser	Asp	Ser	Glu	Lys	His	Thr	Leu	Lys	Ile	Asn	Ser	Ile		
			50				55					60					
Thr	Asp	Thr	Glu	Leu	Arg	Leu	Ser	Gly	Pro	Phe	Pro	Lys	Gln	Ala	Glu		
65					70				75						80		
Ala	Lys	Asn	Pro	Thr	His	Lys	Lys	Asp	Asn	Glu	Gln	Lys	Asn	Thr	Arg		
				85					90					95			
Gln	Asp	Met	Leu	Gly	Leu	Lys	Ser	Thr	Leu	Glu	Lys	Phe	Tyr	Phe	Gly		
			100					105						110			
Ser	Thr	Phe	Asp	Asp	Asn	Ile	His	Ile	Gln	Ile	Ile	His	Asn	Ile	Gln		
			115				120						125				
Asp	Ile	Ala	Lys	Ile	Leu	Ala	Ala	His	Ser	Asn	Asn	Ala	Gly	Tyr	Ala		
			130				135					140					
Leu	Asp	Asn	Met	Leu	Ala	Tyr	Gln	Gly	Val	Glu	Phe	Ser	Asp	Met	Ile		
145					150					155					160		
Gly	Tyr	Met	Gly	Thr	Ser	Arg	Thr	Phe	Asp	Asn	Tyr	Asp	Pro	Asn	His		

	165		170		175
Lys Asn Asn Lys Asp Phe Phe Arg Phe Leu Lys Leu Pro Arg Leu Gly					
	180		185		190
Tyr Phe Gly Ser Ala Phe Tyr Ser Gln Lys Gly Lys Asp Phe Glu Lys					
	195		200		205
Arg Ser Asp Glu Glu Val Tyr Asn Ile Cys Ala Leu Met Gly Gln Ile					
	210		215		220
Arg Gln Cys Cys Phe His Gly Lys Gln Glu Lys Tyr Gln Leu Lys Trp					
225		230		235	240
Leu Tyr Asn Phe His Asn Phe Lys Ser Asn Lys Pro Phe Leu Asp Thr					
	245		250		255
Leu Asp Lys His Phe Asp Glu Met Ile Asp Arg Ile Asn Lys Asn Phe					
	260		265		270
Ile Lys Asn Asn Thr Pro Asp Leu Ile Ile Leu Ser Gly Leu Tyr Pro					
	275		280		285
Asp Met Ala Lys Lys Glu Leu Val Arg Leu Phe Tyr Asp Phe Thr Thr					
	290		295		300
Val Lys Glu Tyr Lys Asn Met Gly Phe Ser Val Lys Lys Leu Arg Glu					
305		310		315	320
Lys Met Leu Glu Ser Glu Glu Ala Ser Asp Phe Arg Asp Lys Asp Tyr					
	325		330		335
Asp Ser Val Arg Arg Lys Leu Tyr Lys Leu Met Asp Phe Cys Ile Tyr					
	340		345		350
Tyr Leu Tyr Tyr Ser Asp Ser Glu Arg Asn Glu Asn Leu Val Ser Arg					
	355		360		365
Leu Arg Glu Ser Leu Thr Asp Glu Asn Lys Asp Ile Ile Tyr Ser Lys					
	370		375		380
Glu Ala Lys Ile Val Trp Asn Glu Leu Arg Lys Lys Phe Ser Thr Ile					
385		390		395	400
Leu Asp Asn Val Lys Gly Ser Asn Ile Lys Lys Leu Glu Asn Val Lys					
	405		410		415
Glu Lys Phe Ile Ser Glu Asp Glu Phe Asp Asp Ile Lys Leu Asp Ile					
	420		425		430
Asp Ile Ser Tyr Phe Ser Lys Leu Met Tyr Val Met Cys Tyr Phe Leu					
	435		440		445
Asp Gly Lys Glu Ile Asn Asp Leu Leu Thr Thr Leu Val Ser Lys Phe					
	450		455		460
Asp Asn Ile Gly Ser Ile Ile Glu Ala Ala Thr Gln Ile Gly Ile Asn					
465		470		475	480

Ile Glu Phe Ile Asp Asp Phe Lys Phe Phe Asp Arg Ser Lys Asp Ile		
	485	490 495
Ser Val Glu Leu Asn Ile Ile Arg Asn Phe Ala Arg Met Gln Ala Pro		
	500	505 510
Val Pro Asn Ala Lys Arg Ala Met Gln Glu Asp Ala Ile Arg Ile Leu		
	515	520 525
Gly Gly Ser Glu Glu Asp Ile Phe Ser Ile Leu Asp Asp Met Thr Gly		
	530	535 540
Tyr Asp Lys Ser Gly Lys Lys Leu Ala Gln Ser Lys Lys Gly Phe Arg		
545	550	555 560
Asn Phe Ile Ile Asn Asn Val Val Glu Ser Ser Arg Phe Lys Tyr Ile		
	565	570 575
Val Arg Tyr Ser Asn Pro Gln Lys Ile Arg Lys Leu Ala Asn Asn Ser		
	580	585 590
Val Val Val Gly Phe Val Leu Gly Lys Leu Pro Asp Ala Gln Ile Glu		
	595	600 605
Ser Tyr Phe Asn Ser Cys Leu Pro Asn Arg Val Tyr Ser Thr Pro Asp		
610	615	620
Lys Ala Arg Glu Ser Leu Arg Asp Met Leu His Asn Ile Ser Phe Asn		
625	630	635 640
Asp Phe Ala Asp Val Lys Gln Asp Asp Arg Arg Ala Thr Pro Glu Glu		
	645	650 655
Lys Val Glu Lys Glu Arg Tyr Lys Ala Ile Ile Gly Leu Tyr Leu Thr		
	660	665 670
Val Met Tyr His Leu Val Lys Asn Leu Val Tyr Val Asn Ser Arg Tyr		
	675	680 685
Val Met Ala Phe His Cys Leu Glu Arg Asp Ala Met His Tyr Asp Val		
	690	695 700
Ser Leu Asp Asn Tyr Arg Asp Leu Ile Arg His Leu Ile Ser Glu Gly		
705	710	715 720
Asp Ser Ser Cys Asn His Phe Ile Ser His Asn Arg Arg Met Arg Asp		
	725	730 735
Cys Ile Glu Glu Asn Val Lys Asn Ser Glu Gln Leu Ile Phe Gly Lys		
	740	745 750
Glu Asp Ala Val Ile Arg Phe Arg Asn Asn Val Ala His Leu Ser Ala		
	755	760 765
Ile Arg Asn Ala Asn Glu Tyr Ile Gly Asp Ile Arg Glu Ile Thr Ser		
	770	775 780
Tyr Phe Ala Leu Tyr His Tyr Leu Met Gln Arg Lys Leu Ile Asp Asp		

785	790	795	800
Cys Lys Val Asn Asp Thr Ala His Lys Tyr Phe Glu Gln Leu Thr Lys			
	805	810	815
Tyr Lys Thr Tyr Val Met Asp Met Val Lys Ala Leu Cys Ser Pro Phe			
	820	825	830
Gly Tyr Asn Leu Pro Arg Phe Lys Asn Leu Ser Ile Glu Gly Lys Phe			
	835	840	845
Asp Met His Glu Ser Lys			
850			
<210> 111			
<211> 965			
<212> PRT			
<213> 未知的			
<220>			
<223> CasRX/Cas13d Ga0224415 10048792			
<400> 111			
Met Ser Lys Lys Glu Asn Arg Lys Ser Tyr Val Lys Gly Leu Gly Leu			
1	5	10	15
Lys Ser Thr Leu Val Ser Asp Ser Lys Val Tyr Leu Thr Thr Phe Ala			
	20	25	30
Asp Gly Ser Asn Ala Lys Leu Glu Lys Cys Val Glu Asn Asn Lys Ile			
	35	40	45
Ile Cys Ile Ser Asn Asp Lys Glu Ala Phe Ala Ala Ser Ile Ala Asn			
	50	55	60
Lys Asn Val Gly Tyr Lys Ile Lys Asn Asp Glu Lys Phe Arg His Pro			
65	70	75	80
Lys Gly Tyr Asp Ile Ile Ser Asn Asn Pro Leu Leu His Asn Asn Ser			
	85	90	95
Val Gln Gln Asp Met Leu Gly Leu Lys Asn Val Leu Glu Lys Arg Tyr			
	100	105	110
Phe Gly Lys Ser Ser Gly Gly Asp Asn Asn Leu Cys Ile Gln Ile Ile			
	115	120	125
His Asn Ile Ile Asp Ile Glu Lys Ile Leu Ser Glu Tyr Ile Pro Asn			
	130	135	140
Val Val Tyr Ala Phe Asn Asn Ile Ala Gly Phe Lys Asp Glu His Asn			
145	150	155	160
Asn Ile Ile Asp Ile Ile Gly Thr Gln Thr Tyr Asn Ser Ser Tyr Thr			
	165	170	175
Tyr Ala Asp Phe Ser Lys Asp Lys Ser Asp Lys Lys Tyr Ile Glu Phe			

			180				185				190					
Gln	Lys	Leu	Leu	Lys	Asn	Lys	Arg	Leu	Gly	Tyr	Trp	Gly	Lys	Ala	Phe	
			195				200				205					
Phe	Thr	Gly	Gln	Gly	Asn	Asn	Ala	Lys	Val	Arg	Gln	Glu	Asn	Gln	Cys	
			210				215				220					
Phe	His	Ile	Ile	Ala	Leu	Leu	Ile	Ser	Leu	Arg	Asn	Trp	Ala	Thr	His	
225							230				235				240	
Ser	Asn	Glu	Leu	Asp	Lys	His	Thr	Lys	Arg	Thr	Trp	Leu	Tyr	Lys	Leu	
			245								250				255	
Asp	Asp	Thr	Asn	Ile	Leu	Asn	Ala	Glu	Tyr	Val	Lys	Thr	Leu	Asn	Tyr	
			260								265				270	
Leu	Tyr	Asp	Thr	Ile	Ala	Asp	Glu	Leu	Thr	Lys	Ser	Phe	Ser	Lys	Asn	
			275								280				285	
Gly	Ala	Val	Asn	Val	Asn	Tyr	Leu	Ala	Lys	Lys	Tyr	Asn	Ile	Lys	Asp	
290							295				300					
Asp	Leu	Pro	Gly	Phe	Ser	Glu	Gln	Tyr	Phe	Arg	Phe	Ser	Ile	Met	Lys	
305							310				315				320	
Glu	Gln	Lys	Asn	Leu	Gly	Phe	Asn	Ile	Ser	Lys	Leu	Arg	Glu	Asn	Met	
			325								330				335	
Leu	Asp	Phe	Lys	Asp	Met	Ser	Val	Ile	Arg	Asp	Asp	His	Asn	Arg	Tyr	
			340								345				350	
Asp	Lys	Asp	Arg	Ser	Lys	Ile	Tyr	Thr	Met	Met	Asp	Phe	Val	Ile	Tyr	
			355								360				365	
Arg	Tyr	Tyr	Ile	Asp	Asn	Asn	Asn	Asp	Ser	Ile	Asp	Phe	Ile	Asn	Lys	
370							375				380					
Leu	Arg	Ser	Ser	Ile	Asp	Glu	Lys	Ser	Lys	Glu	Lys	Leu	Tyr	Asn	Glu	
385							390				395				400	
Glu	Ala	Asn	Arg	Leu	Trp	Asn	Lys	Leu	Lys	Glu	Tyr	Met	Leu	Tyr	Ile	
			405								410				415	
Lys	Glu	Phe	Asn	Gly	Lys	Leu	Ala	Ser	Arg	Thr	Pro	Asp	Arg	Asp	Gly	
			420								425				430	
Asn	Ile	Ser	Glu	Phe	Val	Glu	Ser	Leu	Pro	Lys	Ile	His	Arg	Leu	Leu	
			435								440				445	
Pro	Arg	Gly	Gln	Lys	Ile	Ser	Asn	Phe	Ser	Lys	Leu	Met	Tyr	Leu	Leu	
450							455				460					
Thr	Met	Phe	Leu	Asp	Gly	Lys	Glu	Ile	Asn	Asp	Leu	Leu	Thr	Thr	Leu	
465							470				475				480	
Ile	Asn	Lys	Phe	Glu	Asn	Ile	Gln	Gly	Phe	Leu	Asp	Ile	Met	Pro	Glu	
			485								490				495	

Ile Asn Val Asn Ala Lys Phe Glu Pro Glu Tyr Val Phe Phe Asn Lys		
500	505	510
Ser His Glu Ile Ala Gly Glu Leu Lys Leu Ile Lys Gly Phe Ala Gln		
515	520	525
Met Gly Glu Pro Ala Ala Thr Leu Lys Leu Glu Met Thr Ala Asp Ala		
530	535	540
Ile Lys Ile Leu Gly Thr Glu Lys Glu Asp Ala Glu Leu Ile Lys Leu		
545	550	555
Ala Glu Ser Leu Phe Lys Asp Glu Asn Gly Lys Leu Leu Gly Asn Lys		
565	570	575
Gln His Gly Met Arg Asn Phe Ile Gly Asn Asn Val Ile Lys Ser Lys		
580	585	590
Arg Phe His Tyr Leu Ile Arg Tyr Gly Asp Pro Ala His Leu His Lys		
595	600	605
Ile Ala Thr Asn Lys Asn Val Val Arg Phe Val Leu Gly Arg Ile Ala		
610	615	620
Asp Met Gln Lys Lys Gln Gly Gln Lys Gly Lys Asn Gln Ile Asp Arg		
625	630	635
Tyr Tyr Glu Val Cys Val Gly Asn Lys Asp Ile Lys Lys Thr Ile Glu		
645	650	655
Glu Lys Ile Asp Ala Leu Thr Asp Ile Ile Val Asn Met Asn Tyr Asp		
660	665	670
Gln Phe Glu Lys Lys Lys Ala Val Ile Glu Asn Gln Asn Arg Gly Lys		
675	680	685
Thr Phe Glu Glu Lys Asn Lys Tyr Lys Arg Asp Asn Ala Glu Arg Glu		
690	695	700
Lys Phe Lys Lys Ile Ile Ser Leu Tyr Leu Thr Val Ile Tyr His Ile		
705	710	715
Leu Lys Asn Ile Val Asn Val Asn Ser Arg Tyr Ile Leu Gly Phe His		
725	730	735
Cys Leu Glu Arg Asp Lys Gln Leu Tyr Ile Glu Lys Tyr Asn Lys Asp		
740	745	750
Lys Leu Asp Gly Phe Val Ala Leu Thr Lys Phe Cys Leu Gly Asp Glu		
755	760	765
Glu Arg Phe Glu Asp Leu Lys Ala Lys Ala Gln Ala Ser Ile Gln Ala		
770	775	780
Leu Glu Thr Ala Asn Pro Lys Leu Tyr Ala Lys Tyr Met Asn Tyr Ser		
785	790	795
Asp Glu Glu Lys Lys Glu Glu Phe Lys Lys Gln Leu Asn Arg Glu Arg		

					805					810					815	
Val	Lys	Asn	Ala	Arg	Asn	Ala	Tyr	Leu	Lys	Asn	Ile	Lys	Asn	Tyr	Ile	
			820					825					830			
Met	Ile	Arg	Leu	Gln	Leu	Arg	Asp	Gln	Thr	Asp	Ser	Ser	Gly	Tyr	Leu	
		835					840					845				
Cys	Gly	Glu	Phe	Arg	Asp	Lys	Val	Ala	His	Leu	Glu	Val	Ala	Arg	His	
	850					855					860					
Ala	His	Glu	Tyr	Ile	Gly	Asn	Ile	Lys	Glu	Val	Asn	Ser	Tyr	Phe	Gln	
865					870					875					880	
Leu	Tyr	His	Tyr	Ile	Met	Gln	Cys	Arg	Leu	Tyr	Asp	Val	Leu	Lys	Asn	
			885					890					895			
Asn	Thr	Lys	Ala	Glu	Ala	Met	Val	Lys	Gly	Lys	Ala	Lys	Glu	Tyr	Phe	
		900						905				910				
Glu	Ala	Leu	Glu	Lys	Glu	Gly	Thr	Tyr	Asn	Asp	Lys	Leu	Leu	Lys	Ile	
		915					920					925				
Ala	Cys	Val	Pro	Phe	Gly	Tyr	Cys	Ile	Pro	Arg	Tyr	Lys	Asn	Leu	Ser	
	930					935					940					
Met	Glu	Glu	Leu	Phe	Asp	Met	Asn	Glu	Glu	Lys	Lys	Phe	Lys	Lys	Lys	
945					950					955					960	
Ala	Pro	Glu	Asn	Thr												
			965													
<210> 112																
<211> 1022																
<212> PRT																
<213> 未知的																
<220>																
<223> CasRX/Cas13d 160582958 基因49834																
<400> 112																
Met	Lys	Asn	Ser	Val	Thr	Phe	Lys	Leu	Ile	Gln	Ala	Gln	Glu	Asn	Lys	
1			5					10					15			
Glu	Ala	Ala	Arg	Lys	Lys	Ala	Lys	Asp	Ile	Ala	Glu	Gln	Ala	Arg	Ile	
			20					25				30				
Ala	Lys	Arg	Asn	Gly	Val	Val	Lys	Lys	Glu	Glu	Asn	Arg	Ile	Asn	Arg	
		35					40					45				
Ile	Gln	Ile	Glu	Ile	Gln	Thr	Gln	Lys	Lys	Ser	Asn	Thr	Gln	Asn	Ala	
	50					55					60					
Tyr	His	Leu	Lys	Ser	Leu	Ala	Lys	Ala	Ala	Gly	Val	Lys	Ser	Val	Phe	
65					70					75					80	
Ala	Ile	Gly	Asn	Asp	Leu	Leu	Met	Thr	Gly	Phe	Gly	Pro	Gly	Asn	Asp	

				85					90					95			
Ala	Thr	Ile	Glu	Lys	Arg	Val	Phe	Gln	Asn	Arg	Ala	Ile	Glu	Thr	Leu		
				100					105					110			
Ser	Ser	Pro	Glu	Gln	Tyr	Ser	Ala	Glu	Phe	Gln	Asn	Lys	Gln	Phe	Lys		
				115					120					125			
Ile	Lys	Gly	Asn	Ile	Lys	Val	Leu	Asn	His	Ser	Thr	Gln	Lys	Met	Glu		
				130					135					140			
Glu	Ile	Gln	Thr	Glu	Leu	Gln	Asp	Asn	Tyr	Asn	Arg	Pro	His	Phe	Asp		
145						150				155					160		
Leu	Leu	Gly	Cys	Lys	Asn	Val	Leu	Glu	Gln	Lys	Tyr	Phe	Gly	Arg	Thr		
				165					170					175			
Phe	Ser	Asp	Asn	Ile	His	Val	Gln	Ile	Ala	Tyr	Asn	Ile	Met	Asp	Ile		
				180					185					190			
Glu	Lys	Leu	Leu	Thr	Pro	Tyr	Ile	Asn	Asn	Ile	Ile	Tyr	Thr	Leu	Asn		
				195					200					205			
Glu	Leu	Met	Arg	Asp	Asn	Ser	Lys	Asp	Asp	Phe	Phe	Gly	Cys	Asp	Ser		
				210					215					220			
His	Phe	Ser	Val	Ala	Tyr	Leu	Tyr	Asp	Glu	Leu	Lys	Ala	Gly	Tyr	Ser		
225						230				235					240		
Asp	Arg	Leu	Lys	Thr	Lys	Pro	Asn	Leu	Ser	Lys	Asn	Ile	Asp	Arg	Ile		
				245					250					255			
Trp	Asn	Asn	Phe	Cys	Asn	Tyr	Met	Asn	Ser	Asp	Ser	Gly	Asn	Thr	Glu		
				260					265					270			
Ala	Arg	Leu	Ala	Tyr	Phe	Gly	Glu	Leu	Phe	Tyr	Lys	Pro	Lys	Glu	Thr		
				275					280					285			
Gly	Asp	Ala	Lys	Ser	Asp	Tyr	Lys	Thr	His	Leu	Ser	Asn	Asn	Gln	Lys		
				290					295					300			
Glu	Glu	Trp	Glu	Leu	Lys	Ser	Asp	Lys	Glu	Val	Tyr	Asn	Ile	Phe	Ala		
305						310				315					320		
Ile	Leu	Cys	Asp	Leu	Arg	His	Phe	Cys	Thr	His	Gly	Glu	Ser	Ile	Thr		
				325					330					335			
Pro	Ser	Gly	Lys	Pro	Phe	Pro	Tyr	Asn	Leu	Glu	Lys	Asn	Leu	Phe	Pro		
				340					345					350			
Glu	Ala	Lys	Gln	Val	Leu	Asn	Ser	Leu	Phe	Glu	Glu	Lys	Ala	Glu	Ser		
				355					360					365			
Leu	Gly	Ala	Glu	Ala	Phe	Gly	Lys	Thr	Ala	Gly	Lys	Thr	Asp	Val	Ser		
				370					375					380			
Ile	Leu	Leu	Lys	Val	Phe	Glu	Lys	Glu	Gln	Ala	Ser	Gln	Lys	Glu	Gln		
385						390				395					400		

Gln Ala Leu Leu Lys Glu Tyr Tyr Asp Phe Lys Val Gln Lys Thr Tyr		
405	410	415
Lys Asn Met Gly Phe Ser Ile Lys Lys Leu Arg Glu Ala Ile Met Glu		
420	425	430
Ile Pro Asp Ala Ala Lys Phe Lys Asp Asp Leu Tyr Ser Ser Leu Arg		
435	440	445
His Lys Leu Tyr Gly Leu Phe Asp Phe Ile Leu Val Lys His Phe Leu		
450	455	460
Asp Thr Ser Asp Ser Glu Asn Leu Gln Asn Asn Asp Ile Phe Arg Gln		
465	470	475
Leu Arg Ala Cys Arg Cys Glu Glu Glu Lys Asp Gln Val Tyr Arg Ser		
485	490	495
Ile Ala Val Lys Val Trp Glu Lys Val Lys Lys Lys Glu Leu Asn Met		
500	505	510
Phe Lys Gln Val Val Val Ile Pro Ser Leu Ser Lys Asp Glu Leu Lys		
515	520	525
Gln Met Glu Met Thr Lys Asn Thr Glu Leu Leu Ser Ser Ile Glu Thr		
530	535	540
Ile Ser Thr Gln Ala Ser Leu Phe Ser Glu Met Ile Phe Met Met Thr		
545	550	555
Tyr Leu Leu Asp Gly Lys Glu Ile Asn Leu Leu Cys Thr Ser Leu Ile		
565	570	575
Glu Lys Phe Glu Asn Ile Ala Ser Phe Asn Glu Val Leu Lys Ser Pro		
580	585	590
Gln Ile Gly Tyr Glu Thr Lys Tyr Thr Glu Gly Tyr Ala Phe Phe Lys		
595	600	605
Asn Ala Asp Lys Thr Ala Lys Glu Leu Arg Gln Val Asn Asn Met Ala		
610	615	620
Arg Met Thr Lys Pro Leu Gly Gly Val Asn Thr Lys Cys Val Met Tyr		
625	630	635
Asn Glu Ala Ala Lys Ile Leu Gly Ala Lys Pro Met Ser Lys Ala Glu		
645	650	655
Leu Glu Ser Val Phe Asn Leu Asp Asn His Asp Tyr Thr Tyr Ser Pro		
660	665	670
Ser Gly Lys Lys Ile Pro Asn Lys Asn Phe Arg Asn Phe Ile Ile Asn		
675	680	685
Asn Val Ile Thr Ser Arg Arg Phe Leu Tyr Leu Ile Arg Tyr Gly Asn		
690	695	700
Pro Glu Lys Ile Arg Lys Ile Ala Ile Asn Pro Ser Ile Ile Ser Phe		

705	710	715	720
Val Leu Lys Gln Ile Pro Asp Glu Gln Ile Lys Arg Tyr Tyr Pro Pro			
	725	730	735
Cys Ile Gly Lys Arg Thr Asp Asp Val Thr Leu Met Arg Asp Glu Leu			
	740	745	750
Gly Lys Met Leu Gln Ser Val Asn Phe Glu Gln Phe Ser Arg Val Asn			
	755	760	765
Asn Lys Gln Asn Ala Lys Gln Asn Pro Asn Gly Glu Lys Ala Arg Leu			
	770	775	780
Gln Ala Cys Val Arg Leu Tyr Leu Thr Val Pro Tyr Leu Phe Ile Lys			
785	790	795	800
Asn Met Val Asn Ile Asn Ala Arg Tyr Val Leu Ala Phe His Cys Leu			
	805	810	815
Glu Arg Asp His Ala Leu Cys Phe Asn Ser Arg Lys Leu Asn Asp Asp			
	820	825	830
Ser Tyr Asn Glu Met Ala Asn Lys Phe Gln Met Val Arg Lys Ala Lys			
	835	840	845
Lys Glu Gln Tyr Glu Lys Glu Tyr Lys Cys Lys Lys Gln Glu Thr Gly			
	850	855	860
Thr Ala His Thr Lys Lys Ile Glu Lys Leu Asn Gln Gln Ile Ala Tyr			
865	870	875	880
Ile Asp Lys Asp Ile Lys Asn Met His Ser Tyr Thr Cys Arg Asn Tyr			
	885	890	895
Arg Asn Leu Val Ala His Leu Asn Val Val Ser Lys Leu Gln Asn Tyr			
	900	905	910
Val Ser Glu Leu Pro Asn Asp Tyr Gln Ile Thr Ser Tyr Phe Ser Phe			
	915	920	925
Tyr His Tyr Cys Met Gln Leu Gly Leu Met Glu Lys Val Ser Ser Lys			
	930	935	940
Asn Ile Pro Leu Val Glu Ser Leu Lys Asn Glu Ala Asn Asp Ala Gln			
945	950	955	960
Ser Tyr Ser Ala Lys Lys Thr Leu Glu Tyr Phe Asp Leu Ile Glu Lys			
	965	970	975
Asn Arg Thr Tyr Cys Lys Asp Phe Leu Lys Ala Leu Asn Ala Pro Phe			
	980	985	990
Ser Tyr Asn Leu Pro Arg Phe Lys Asn Leu Ser Ile Glu Ala Leu Phe			
	995	1000	1005
Asp Lys Asn Ile Val Tyr Glu Gln Ala Asp Leu Lys Lys Glu			
1010	1015	1020	

<210> 113

<211> 36

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d DR

<400> 113

Gly Ala Ala Cys Thr Ala Cys Ala Cys Cys Cys Cys Thr Cys Thr Gly
 1 5 10 15
 Thr Thr Cys Thr Thr Gly Thr Ala Gly Gly Gly Gly Thr Cys Thr Ala
 20 25 30
 Ala Cys Ala Cys
 35

<210> 114

<211> 998

<212> PRT

<213> 未知的

<220>

<223> CasRX/Cas13d 250twins 35838 GL0110300

<400> 114

Met Gly Asn Lys Gln Arg Val Ser Ala Gln Lys Arg Arg Glu Asn Ala
 1 5 10 15
 Lys Leu Cys Asn Gln Gln Lys Ala Arg Gln Ala Glu Ser Gln Arg Asp
 20 25 30
 Lys Ile Lys Asn Met Asn Val Glu Lys Met Lys Asn Ile Asn Thr Asn
 35 40 45
 Asp Ile Lys His Thr Lys Thr Thr Ala Lys Lys Leu Gly Leu Lys Ser
 50 55 60
 Thr Ile Ile Ala Asp Lys Lys Ile Ile Leu Thr Ser Phe Ile Asn Glu
 65 70 75 80
 Gln Ser Ser Lys Thr Ala Asn Ile Glu Lys Val Ala Gly Phe Lys Gly
 85 90 95
 Asp Thr Ile Asp Thr Ile Ser Tyr Thr Pro Arg Met Phe Arg Ser Glu
 100 105 110
 Ile Asn Pro Gly Glu Ile Val Ile Ser Lys Gly Asp Asp Leu Ser Glu
 115 120 125
 Phe Ala Asn Pro Ala Asn Phe Pro Ile Gly Arg Asp Tyr Val Lys Ile
 130 135 140
 Arg Ser Ala Leu Glu Lys Gln Tyr Phe Gly Lys Glu Phe Pro Glu Asp

145	150	155	160
Asn Leu His Val Gln Ile Ala Tyr Asn Val Ala Asp Ile Lys Lys Ile			
	165	170	175
Leu Ser Val Tyr Ile Asn Asn Ile Ile Tyr Met Phe Tyr Asn Leu Ala			
	180	185	190
Arg Ser Glu Glu Tyr Asp Ile Phe Tyr Asn Ser Gln Ser Glu Asn Ser			
	195	200	205
Gly Arg Asp Cys Asp Val Ile Gly Ser Leu Tyr Tyr Gln Ala Ser Tyr			
	210	215	220
Arg Asn Gln Asp Ala Asn Arg Phe Glu Lys Asp Gly Lys Lys Lys Ala			
225	230	235	240
Ile Asp Ser Leu Leu Asp Asp Thr Arg Ala Tyr Tyr Thr Tyr Phe Asp			
	245	250	255
Gly Leu Phe Ser Val Pro Lys Arg Glu Asp Asp Gly Lys Ile Lys Glu			
	260	265	270
Ser Glu Lys Glu Lys Ala Lys Asp Gln Asn Phe Asp Val Leu Arg Leu			
	275	280	285
Leu Ser Val Gly Arg Gln Leu Thr Phe His Ser Asp Lys Ser Asn Asn			
	290	295	300
Glu Ala Tyr Leu Phe Asp Leu Ser Lys Leu Thr Arg Ala Ala Gln Asp			
305	310	315	320
Glu Asn Arg Arg Gln Asp Ile Gln Ser Leu Leu Asn Ile Leu Asn Ser			
	325	330	335
Thr Cys Arg Ser Asn Leu Glu Gly Val Asn Gly Asp Phe Val Lys His			
	340	345	350
Ala Lys Asn Asn Leu Tyr Val Leu Asn Gln Leu Tyr Pro Ser Leu Lys			
	355	360	365
Ala Asn Asp Leu Ile Gly Glu Tyr Tyr Asn Phe Ile Val Lys Lys Glu			
	370	375	380
Asn Arg Asn Ile Gly Ile Arg Leu Ile Thr Val Arg Glu Leu Ile Ile			
385	390	395	400
Glu His Asn Tyr Thr Asn Leu Lys Asp Ser Lys Tyr Asp Thr Tyr Arg			
	405	410	415
Asn Lys Ile Tyr Thr Val Leu Asn Phe Ile Leu Phe Arg Glu Ile Gln			
	420	425	430
Glu Asn Ser Ile Ala Ile Lys Asn Phe Arg Glu Lys Leu Arg Ser Thr			
	435	440	445
Glu Lys Ala Glu Gln Pro Ala Leu Tyr Gln Ala Phe Ala Asn Lys Ile			
450	455	460	

Tyr	Pro	Met	Val	Gln	Ala	Lys	Phe	Ala	Lys	Ala	Ile	Asp	Leu	Phe	Glu	465	470	475	480
Glu	Gln	Tyr	Lys	Thr	Lys	Phe	Lys	Ser	Glu	Phe	Lys	Gly	Gly	Ile	Ser	485	490	495	
Ile	Glu	Asn	Met	Gln	Gln	Gln	Asn	Ile	Leu	Leu	Gln	Thr	Glu	Asn	Ile	500	505	510	
Asp	Tyr	Phe	Ser	Lys	Tyr	Val	Leu	Phe	Leu	Thr	Lys	Phe	Leu	Asp	Gly	515	520	525	
Lys	Glu	Ile	Asn	Glu	Leu	Leu	Cys	Ala	Leu	Ile	Asn	Lys	Phe	Asp	Asn	530	535	540	
Ile	Ala	Asp	Leu	Leu	Asp	Ile	Ser	Lys	Gln	Ile	Gly	Thr	Pro	Val	Val	545	550	555	560
Phe	Cys	Ala	Asp	Tyr	Glu	Ser	Leu	Asn	Asp	Ala	Ala	Lys	Ile	Ala	Glu	565	570	575	
Asn	Ile	Arg	Leu	Ile	Lys	Asn	Ile	Ala	His	Leu	Arg	Pro	Ala	Ile	Gln	580	585	590	
Glu	Ala	Gln	Ser	Ser	Lys	Asp	Asn	Ala	Asp	Ala	Ala	Gly	Thr	Pro	Ala	595	600	605	
Thr	Leu	Leu	Ile	Asp	Ala	Tyr	Asn	Met	Leu	Asn	Thr	Asp	Ile	Gln	Leu	610	615	620	
Val	Tyr	Gly	Glu	Ala	Ala	Tyr	Glu	Glu	Leu	Arg	Lys	Asp	Leu	Phe	Glu	625	630	635	640
Arg	Lys	Asn	Gly	Thr	Lys	Tyr	Asn	Lys	Lys	Gly	Lys	Lys	Val	Asp	Val	645	650	655	
Tyr	Asp	His	Lys	Phe	Arg	Asn	Phe	Leu	Ile	Asn	Asn	Val	Ile	Lys	Ser	660	665	670	
Lys	Trp	Phe	Phe	Tyr	Ile	Ala	Lys	Tyr	Val	Lys	Pro	Ala	Asp	Cys	Ala	675	680	685	
Lys	Met	Met	Ser	Asn	Lys	Lys	Met	Ile	Glu	Phe	Ala	Leu	Arg	Asp	Leu	690	695	700	
Pro	Glu	Thr	Gln	Ile	Lys	Arg	Tyr	Tyr	Tyr	Thr	Ile	Thr	Gly	Asn	Glu	705	710	715	720
Ala	Leu	Gly	Asp	Ala	Glu	Ser	Leu	Lys	Gly	Val	Ile	Ile	Glu	Gln	Leu	725	730	735	
His	Ala	Phe	Ser	Ile	Lys	Asn	Thr	Leu	Leu	Ser	Ile	Lys	Asn	Met	Gly	740	745	750	
Glu	Gly	Glu	Tyr	Lys	Ile	Gln	Gln	Ile	Gly	Ser	Ser	Lys	Glu	Lys	Leu	755	760	765	
Lys	Ala	Ile	Val	Asn	Leu	Tyr	Leu	Thr	Val	Ala	Tyr	Leu	Leu	Thr	Lys				

770	775	780
Ser Leu Val Lys Val Asn Ile Arg Phe Ser Ile Ala Phe Gly Cys Leu		
785	790	795
Glu Arg Asp Leu Val Leu Gln Lys Lys Ser Glu Lys Lys Phe Asp Ala		800
805	810	815
Ile Ile Asn Glu Ile Leu Leu Glu Asp Asp Lys Ile Arg Lys Glu Cys		
820	825	830
Asp Lys Glu Arg Ala Gln Ala Lys Thr Leu Pro Arg Glu Leu Ala Gln		
835	840	845
Glu Arg Phe Ala Gln Ile Lys Arg Arg Glu Ser Gly Cys Tyr Phe Lys		
850	855	860
Ser Tyr His Val Tyr Asp Tyr Leu Ser Lys Asn Ser Asn Glu Phe Lys		
865	870	875
Gln Asn His Ile Asp Phe Ala Val Thr Ser Tyr Arg Asn Asn Val Glu		
885	890	895
His Leu Asn Val Val His Cys Met Thr Lys Tyr Phe Ser Glu Val Lys		
900	905	910
Asp Val Lys Ser Tyr Tyr Gly Val Tyr Cys Tyr Ile Met Gln Arg Met		
915	920	925
Leu Cys Asp Glu Leu Ile Ile Lys Asn Gln Asp Lys Pro Asp Val Arg		
930	935	940
Gln Thr Phe Glu Glu Tyr Asn Arg Leu Leu Lys Asp His Gly Thr Tyr		
945	950	955
Ser Lys Asn Leu Met Trp Leu Leu Asn Phe Pro Phe Ala Tyr Asn Leu		
965	970	975
Ala Arg Tyr Lys Asn Leu Ser Asn Glu Asp Leu Phe Asn Ala Lys Asn		
980	985	990
Asn Asp Gln Lys Ser Lys		
995		
<210> 115		
<211> 1020		
<212> PRT		
<213> 未知的		
<220>		
<223> CasRX/Cas13d 250twins 36050 GL0158985		
<400> 115		
Met Lys Lys Lys His Gln Ser Ala Ala Glu Lys Arg Gln Val Lys Lys		
1	5	10
Leu Lys Asn Gln Glu Lys Ala Gln Lys Tyr Ala Ser Glu Pro Ser Pro		15

20	25	30
Leu Gln Ser Asp Thr Ala Gly Val	Glu Cys Ser Gln Lys Lys Thr Val	
35	40	45
Val Ser His Ile Ala Ser Ser Lys Thr	Leu Ala Lys Ala Met Gly Leu	
50	55	60
Lys Ser Thr Leu Val Met Gly Asp Lys	Leu Val Ile Thr Ser Phe Ala	
65	70	75
Ala Ser Lys Ala Val Gly Gly Ala Gly	Tyr Lys Ser Ala Asn Ile Glu	
85	90	95
Lys Ile Thr Asp Leu Gln Gly Arg Val	Ile Glu Glu His Glu Arg Met	
100	105	110
Phe Ser Ala Asp Val Gly Glu Lys Asn	Ile Glu Leu Ser Lys Asn Asp	
115	120	125
Cys His Thr Asn Val Asn Asn Pro Val	Val Thr Asn Ile Gly Lys Asp	
130	135	140
Tyr Ile Gly Leu Lys Ser Arg Leu Glu	Gln Glu Phe Phe Gly Lys Thr	
145	150	155
Phe Glu Asn Asp Asn Leu His Val Gln	Leu Ala Tyr Asn Ile Leu Asp	
165	170	175
Ile Lys Lys Ile Leu Gly Thr Tyr Val	Asn Asn Ile Ile Tyr Ile Phe	
180	185	190
Tyr Asn Leu Asn Arg Ala Gly Thr Gly	Arg Asp Glu Arg Met Tyr Asp	
195	200	205
Asp Leu Ile Gly Thr Leu Tyr Ala Tyr	Lys Pro Met Glu Ala Gln Gln	
210	215	220
Thr Tyr Leu Leu Lys Gly Asp Lys Asp	Met Arg Arg Phe Glu Glu Val	
225	230	235
Lys Gln Leu Leu Gln Asn Thr Ser Ala	Tyr Tyr Val Tyr Tyr Gly Thr	
245	250	255
Leu Phe Glu Lys Val Lys Ala Lys Ser	Lys Lys Glu Gln Arg Ala Lys	
260	265	270
Glu Ala Glu Ile Asp Ala Cys Thr Ala	His Asn Tyr Asp Val Leu Arg	
275	280	285
Leu Leu Ser Leu Met Arg Gln Leu Cys	Met His Ser Val Ala Gly Thr	
290	295	300
Ala Phe Lys Leu Ala Glu Ser Ala Leu	Phe Asn Ile Glu Asp Val Leu	
305	310	315
Ser Ala Asp Leu Lys Glu Ile Leu Asp	Glu Ala Phe Ser Gly Ala Val	
325	330	335

Asn Lys Leu Asn Asp Gly Phe Val Gln His Ser Gly Asn Asn Leu Tyr		
340	345	350
Val Leu Gln Gln Leu Tyr Pro Asn Glu Thr Ile Glu Arg Ile Ala Glu		
355	360	365
Lys Tyr Tyr Arg Leu Thr Val Arg Lys Glu Asp Leu Asn Met Gly Val		
370	375	380
Asn Ile Lys Lys Leu Arg Glu Leu Ile Val Gly Gln Tyr Phe Pro Glu		
385	390	395
Val Leu Asp Lys Glu Tyr Asp Leu Ser Lys Asn Gly Asp Ser Val Val		
405	410	415
Thr Tyr Arg Ser Lys Ile Tyr Thr Val Met Asn Tyr Ile Leu Leu Tyr		
420	425	430
Tyr Leu Glu Asp His Asp Ser Ser Arg Glu Ser Met Val Glu Ala Leu		
435	440	445
Arg Gln Asn Arg Glu Gly Asp Glu Gly Lys Glu Glu Ile Tyr Arg Gln		
450	455	460
Phe Ala Lys Lys Val Trp Asn Gly Val Ser Gly Leu Phe Gly Val Cys		
465	470	475
Leu Asn Leu Phe Lys Thr Glu Lys Arg Asn Lys Phe Arg Ser Lys Val		
485	490	495
Ala Leu Pro Asp Val Ser Gly Ala Ala Tyr Met Leu Ser Ser Glu Asn		
500	505	510
Ile Asp Tyr Phe Val Lys Met Leu Phe Phe Val Cys Lys Phe Leu Asp		
515	520	525
Gly Lys Glu Ile Asn Glu Leu Leu Cys Ala Leu Ile Asn Lys Phe Asp		
530	535	540
Asn Ile Ala Asp Ile Leu Asp Ala Ala Ala Gln Cys Gly Ser Ser Val		
545	550	555
Trp Phe Val Asp Ser Tyr Arg Phe Phe Glu Arg Ser Arg Arg Ile Ser		
565	570	575
Ala Gln Ile Arg Ile Val Lys Asn Ile Ala Ser Lys Asp Phe Lys Lys		
580	585	590
Ser Lys Lys Asp Ser Asp Glu Ser Tyr Pro Glu Gln Leu Tyr Leu Asp		
595	600	605
Ala Leu Ala Leu Leu Gly Asp Val Ile Ser Lys Tyr Lys Gln Asn Arg		
610	615	620
Asp Gly Ser Val Val Ile Asp Asp Gln Gly Asn Ala Val Leu Thr Glu		
625	630	635
Gln Tyr Lys Arg Phe Arg Tyr Glu Phe Phe Glu Glu Ile Lys Arg Asp		

				645					650					655			
Glu	Ser	Gly	Gly	Ile	Lys	Tyr	Lys	Lys	Ser	Gly	Lys	Pro	Glu	Tyr	Asn		
				660					665					670			
His	Gln	Arg	Arg	Asn	Phe	Ile	Leu	Asn	Asn	Val	Leu	Lys	Ser	Lys	Trp		
				675					680					685			
Phe	Phe	Tyr	Val	Val	Lys	Tyr	Asn	Arg	Pro	Ser	Ser	Cys	Arg	Glu	Leu		
				690					695					700			
Met	Lys	Asn	Lys	Glu	Ile	Leu	Arg	Phe	Val	Leu	Arg	Asp	Ile	Pro	Asp		
705						710				715				720			
Ser	Gln	Val	Arg	Arg	Tyr	Phe	Lys	Ala	Val	Gln	Gly	Glu	Glu	Ala	Tyr		
				725					730					735			
Ala	Ser	Ala	Glu	Ala	Met	Arg	Thr	Arg	Leu	Val	Asp	Ala	Leu	Ser	Gln		
				740					745					750			
Phe	Ser	Val	Thr	Ala	Cys	Leu	Asp	Glu	Val	Gly	Gly	Met	Thr	Asp	Lys		
				755					760					765			
Glu	Phe	Ala	Ser	Gln	Arg	Ala	Val	Asp	Ser	Lys	Glu	Lys	Leu	Arg	Ala		
				770					775					780			
Ile	Ile	Arg	Leu	Tyr	Leu	Thr	Val	Ala	Tyr	Leu	Ile	Thr	Lys	Ser	Met		
785						790				795				800			
Val	Lys	Val	Asn	Thr	Arg	Phe	Ser	Ile	Ala	Phe	Ser	Val	Leu	Glu	Arg		
				805					810					815			
Asp	Tyr	Tyr	Leu	Leu	Ile	Asp	Gly	Lys	Lys	Lys	Ser	Ser	Asp	Tyr	Thr		
				820					825					830			
Gly	Glu	Asp	Met	Leu	Ala	Leu	Thr	Arg	Lys	Phe	Val	Gly	Glu	Asp	Ala		
				835					840					845			
Gly	Leu	Tyr	Arg	Glu	Trp	Lys	Glu	Lys	Asn	Ala	Glu	Ala	Lys	Asp	Lys		
				850					855					860			
Tyr	Phe	Asp	Lys	Ala	Glu	Arg	Lys	Lys	Val	Leu	Arg	Gln	Asn	Asp	Lys		
865						870				875				880			
Met	Ile	Arg	Lys	Met	His	Phe	Thr	Pro	His	Ser	Leu	Asn	Tyr	Val	Gln		
				885					890					895			
Lys	Asn	Leu	Glu	Ser	Val	Gln	Ser	Asn	Gly	Leu	Ala	Ala	Val	Ile	Lys		
				900					905					910			
Glu	Tyr	Arg	Asn	Ala	Val	Ala	His	Leu	Asn	Ile	Ile	Asn	Arg	Leu	Asp		
				915					920					925			
Glu	Tyr	Ile	Gly	Ser	Ala	Arg	Ala	Asp	Ser	Tyr	Tyr	Ser	Leu	Tyr	Cys		
				930					935					940			
Tyr	Cys	Leu	Gln	Met	Tyr	Leu	Ser	Lys	Asn	Phe	Ser	Val	Gly	Tyr	Leu		
945						950				955				960			

Ile Asn Val Gln Lys Gln Leu Glu Glu His His Thr Tyr Met Lys Asp
 965 970 975
 Leu Met Trp Leu Leu Asn Ile Pro Phe Ala Tyr Asn Leu Ala Arg Tyr
 980 985 990
 Lys Asn Leu Ser Asn Glu Lys Leu Phe Tyr Asp Glu Glu Ala Ala Ala
 995 1000 1005
 Glu Lys Ala Asp Lys Ala Glu Asn Glu Arg Gly Glu
 1010 1015 1020
 <210> 116
 <211> 128
 <212> PRT
 <213> 人工序列
 <220>
 <223> RNA酶1 (K41R)
 <400> 116
 Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp
 1 5 10 15
 Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg
 20 25 30
 Arg Asn Met Thr Gln Gly Arg Cys Arg Pro Val Asn Thr Phe Val His
 35 40 45
 Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr
 50 55 60
 Cys Lys Asn Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His
 65 70 75 80
 Ile Thr Asp Cys Arg Leu Thr Asn Gly Ser Arg Tyr Pro Asn Cys Ala
 85 90 95
 Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly
 100 105 110
 Ser Pro Tyr Val Pro Val His Phe Asp Ala Ser Val Glu Asp Ser Thr
 115 120 125
 <210> 117
 <211> 128
 <212> PRT
 <213> 人工序列
 <220>
 <223> RNA酶1 (RNA酶1 (K41R, D121E))
 <400> 117
 Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp

1	5	10	15
Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg			
20	25	30	
Arg Asn Met Thr Gln Gly Arg Cys Arg Pro Val Asn Thr Phe Val His			
35	40	45	
Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr			
50	55	60	
Cys Lys Asn Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His			
65	70	75	80
Ile Thr Asp Cys Arg Leu Thr Asn Gly Ser Arg Tyr Pro Asn Cys Ala			
85	90	95	
Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly			
100	105	110	
Ser Pro Tyr Val Pro Val His Phe Glu Ala Ser Val Glu Asp Ser Thr			
115	120	125	
<210> 118			
<211> 128			
<212> PRT			
<213> 人工序列			
<220>			
<223> RNA酶1 (RNA酶1 (K41R, D121E, H119N))			
<400> 118			
Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp			
1	5	10	15
Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg			
20	25	30	
Arg Asn Met Thr Gln Gly Arg Cys Arg Pro Val Asn Thr Phe Val His			
35	40	45	
Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr			
50	55	60	
Cys Lys Asn Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His			
65	70	75	80
Ile Thr Asp Cys Arg Leu Thr Asn Gly Ser Arg Tyr Pro Asn Cys Ala			
85	90	95	
Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly			
100	105	110	
Ser Pro Tyr Val Pro Val Asn Phe Glu Ala Ser Val Glu Asp Ser Thr			
115	120	125	
<210> 119			

<211> 128

<212> PRT

<213> 人工序列

<220>

<223> (RNA酶1 (H119N))

<400> 119

Lys	Glu	Ser	Arg	Ala	Lys	Lys	Phe	Gln	Arg	Gln	His	Met	Asp	Ser	Asp
1				5				10						15	
Ser	Ser	Pro	Ser	Ser	Ser	Ser	Thr	Tyr	Cys	Asn	Gln	Met	Met	Arg	Arg
			20					25					30		
Arg	Asn	Met	Thr	Gln	Gly	Arg	Cys	Lys	Pro	Val	Asn	Thr	Phe	Val	His
		35					40					45			
Glu	Pro	Leu	Val	Asp	Val	Gln	Asn	Val	Cys	Phe	Gln	Glu	Lys	Val	Thr
	50					55					60				
Cys	Lys	Asn	Gly	Gln	Gly	Asn	Cys	Tyr	Lys	Ser	Asn	Ser	Ser	Met	His
65					70				75					80	
Ile	Thr	Asp	Cys	Arg	Leu	Thr	Asn	Gly	Ser	Arg	Tyr	Pro	Asn	Cys	Ala
			85					90					95		
Tyr	Arg	Thr	Ser	Pro	Lys	Glu	Arg	His	Ile	Ile	Val	Ala	Cys	Glu	Gly
			100					105					110		
Ser	Pro	Tyr	Val	Pro	Val	Asn	Phe	Asp	Ala	Ser	Val	Glu	Asp	Ser	Thr
		115					120					125			

<210> 120

<211> 128

<212> PRT

<213> 人工序列

<220>

<223> (RNA酶1 (R39D, N67D, N88A, G89D, R91D, H119N))

<400> 120

Lys	Glu	Ser	Arg	Ala	Lys	Lys	Phe	Gln	Arg	Gln	His	Met	Asp	Ser	Asp
1				5				10						15	
Ser	Ser	Pro	Ser	Ser	Ser	Ser	Thr	Tyr	Cys	Asn	Gln	Met	Met	Arg	Arg
			20					25					30		
Arg	Asn	Met	Thr	Gln	Gly	Asp	Cys	Lys	Pro	Val	Asn	Thr	Phe	Val	His
		35					40					45			
Glu	Pro	Leu	Val	Asp	Val	Gln	Asn	Val	Cys	Phe	Gln	Glu	Lys	Val	Thr
	50					55					60				
Cys	Lys	Asp	Gly	Gln	Gly	Asn	Cys	Tyr	Lys	Ser	Asn	Ser	Ser	Met	His
65					70				75					80	

Ile Thr Asp Cys Arg Leu Thr Ala Asp Ser Asp Tyr Pro Asn Cys Ala
 85 90 95
 Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly
 100 105 110
 Ser Pro Tyr Val Pro Val Asn Phe Asp Ala Ser Val Glu Asp Ser Thr
 115 120 125

<210> 121

<211> 128

<212> PRT

<213> 人工序列

<220>

<223> RNA酶1 (RNA酶1 (R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E))

<400> 121

Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp
 1 5 10 15
 Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg
 20 25 30
 Arg Asn Met Thr Gln Gly Asp Cys Arg Pro Val Asn Thr Phe Val His
 35 40 45
 Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr
 50 55 60
 Cys Lys Asp Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His
 65 70 75 80
 Ile Thr Asp Cys Arg Leu Thr Ala Asp Ser Asp Tyr Pro Asn Cys Ala
 85 90 95
 Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly
 100 105 110
 Ser Pro Tyr Val Pro Val Asn Phe Glu Ala Ser Val Glu Asp Ser Thr
 115 120 125

<210> 122

<211> 128

<212> PRT

<213> 人工序列

<220>

<223> RNA酶1 (RNA酶1 (R39D, N67D, N88A, G89D, R91D))

<400> 122

Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp
 1 5 10 15
 Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg

			20						25						30		
Arg	Asn	Met	Thr	Gln	Gly	Asp	Cys	Lys	Pro	Val	Asn	Thr	Phe	Val	His		
35						40						45					
Glu	Pro	Leu	Val	Asp	Val	Gln	Asn	Val	Cys	Phe	Gln	Glu	Lys	Val	Thr		
50						55						60					
Cys	Lys	Asp	Gly	Gln	Gly	Asn	Cys	Tyr	Lys	Ser	Asn	Ser	Ser	Met	His		
65						70						75			80		
Ile	Thr	Asp	Cys	Arg	Leu	Thr	Ala	Asp	Ser	Asp	Tyr	Pro	Asn	Cys	Ala		
			85						90						95		
Tyr	Arg	Thr	Ser	Pro	Lys	Glu	Arg	His	Ile	Ile	Val	Ala	Cys	Glu	Gly		
100						105						110					
Ser	Pro	Tyr	Val	Pro	Val	His	Phe	Asp	Ala	Ser	Val	Glu	Asp	Ser	Thr		
115						120						125					
<210>			123														
<211>			311														
<212>			PRT														
<213>			智人(Homo Sapiens)														
<400>			123														
Cys	Ser	Pro	Gln	Glu	Ser	Gly	Met	Thr	Ala	Leu	Ser	Ala	Arg	Met	Leu		
1			5						10						15		
Thr	Arg	Ser	Arg	Ser	Leu	Gly	Pro	Gly	Ala	Gly	Pro	Arg	Gly	Cys	Arg		
20						25						30					
Glu	Glu	Pro	Gly	Pro	Leu	Arg	Arg	Arg	Glu	Ala	Ala	Ala	Glu	Ala	Arg		
35						40						45					
Lys	Ser	His	Ser	Pro	Val	Lys	Arg	Pro	Arg	Lys	Ala	Gln	Arg	Leu	Arg		
50						55						60					
Val	Ala	Tyr	Glu	Gly	Ser	Asp	Ser	Glu	Lys	Gly	Glu	Gly	Ala	Glu	Pro		
65						70						75			80		
Leu	Lys	Val	Pro	Val	Trp	Glu	Pro	Gln	Asp	Trp	Gln	Gln	Gln	Leu	Val		
			85						90						95		
Asn	Ile	Arg	Ala	Met	Arg	Asn	Lys	Lys	Asp	Ala	Pro	Val	Asp	His	Leu		
100						105						110					
Gly	Thr	Glu	His	Cys	Tyr	Asp	Ser	Ser	Ala	Pro	Pro	Lys	Val	Arg	Arg		
115						120						125					
Tyr	Gln	Val	Leu	Leu	Ser	Leu	Met	Leu	Ser	Ser	Gln	Thr	Lys	Asp	Gln		
130						135						140					
Val	Thr	Ala	Gly	Ala	Met	Gln	Arg	Leu	Arg	Ala	Arg	Gly	Leu	Thr	Val		
145						150						155			160		
Asp	Ser	Ile	Leu	Gln	Thr	Asp	Asp	Ala	Thr	Leu	Gly	Lys	Leu	Ile	Tyr		

				165						170					175		
Pro	Val	Gly	Phe	Trp	Arg	Ser	Lys	Val	Lys	Tyr	Ile	Lys	Gln	Thr	Ser		
				180						185					190		
Ala	Ile	Leu	Gln	Gln	His	Tyr	Gly	Gly	Asp	Ile	Pro	Ala	Ser	Val	Ala		
				195						200					205		
Glu	Leu	Val	Ala	Leu	Pro	Gly	Val	Gly	Pro	Lys	Met	Ala	His	Leu	Ala		
				210						215					220		
Met	Ala	Val	Ala	Trp	Gly	Thr	Val	Ser	Gly	Ile	Ala	Val	Asp	Thr	His		
225						230					235				240		
Val	His	Arg	Ile	Ala	Asn	Arg	Leu	Arg	Trp	Thr	Lys	Lys	Ala	Thr	Lys		
				245						250					255		
Ser	Pro	Glu	Glu	Thr	Arg	Ala	Ala	Leu	Glu	Glu	Trp	Leu	Pro	Arg	Glu		
				260						265					270		
Leu	Trp	His	Glu	Ile	Asn	Gly	Leu	Leu	Val	Gly	Phe	Gly	Gln	Gln	Thr		
				275						280					285		
Cys	Leu	Pro	Val	His	Pro	Arg	Cys	His	Ala	Cys	Leu	Asn	Gln	Ala	Leu		
				290						295					300		
Cys	Pro	Ala	Ala	Gln	Gly	Leu											
305						310											
<210>	124																
<211>	372																
<212>	PRT																
<213>	智人 (Homo Sapiens)																
<400>	124																
Met	Glu	Ser	Gly	Gln	Pro	Ala	Arg	Arg	Ile	Ala	Met	Ala	Pro	Leu	Leu		
1				5					10					15			
Glu	Tyr	Glu	Arg	Gln	Leu	Val	Leu	Glu	Leu	Leu	Asp	Thr	Asp	Gly	Leu		
				20					25					30			
Val	Val	Cys	Ala	Arg	Gly	Leu	Gly	Ala	Asp	Arg	Leu	Leu	Tyr	His	Phe		
				35					40					45			
Leu	Gln	Leu	His	Cys	His	Pro	Ala	Cys	Leu	Val	Leu	Val	Leu	Asn	Thr		
				50					55					60			
Gln	Pro	Ala	Glu	Glu	Glu	Tyr	Phe	Ile	Asn	Gln	Leu	Lys	Ile	Glu	Gly		
65						70				75					80		
Val	Glu	His	Leu	Pro	Arg	Arg	Val	Thr	Asn	Glu	Ile	Thr	Ser	Asn	Ser		
				85						90					95		
Arg	Tyr	Glu	Val	Tyr	Thr	Gln	Gly	Gly	Val	Ile	Phe	Ala	Thr	Ser	Arg		
				100						105					110		
Ile	Leu	Val	Val	Asp	Phe	Leu	Thr	Asp	Arg	Ile	Pro	Ser	Asp	Leu	Ile		

115	120	125
Thr Gly Ile Leu Val Tyr Arg	Ala His Arg Ile Ile Glu Ser Cys Gln	
130	135	140
Glu Ala Phe Ile Leu Arg Leu Phe Arg Gln Lys Asn Lys Arg Gly Phe		
145	150	155
Ile Lys Ala Phe Thr Asp Asn Ala Val Ala Phe Asp Thr Gly Phe Cys		
165	170	175
His Val Glu Arg Val Met Arg Asn Leu Phe Val Arg Lys Leu Tyr Leu		
180	185	190
Trp Pro Arg Phe His Val Ala Val Asn Ser Phe Leu Glu Gln His Lys		
195	200	205
Pro Glu Val Val Glu Ile His Val Ser Met Thr Pro Thr Met Leu Ala		
210	215	220
Ile Gln Thr Ala Ile Leu Asp Ile Leu Asn Ala Cys Leu Lys Glu Leu		
225	230	235
Lys Cys His Asn Pro Ser Leu Glu Val Glu Asp Leu Ser Leu Glu Asn		
245	250	255
Ala Ile Gly Lys Pro Phe Asp Lys Thr Ile Arg His Tyr Leu Asp Pro		
260	265	270
Leu Trp His Gln Leu Gly Ala Lys Thr Lys Ser Leu Val Gln Asp Leu		
275	280	285
Lys Ile Leu Arg Thr Leu Leu Gln Tyr Leu Ser Gln Tyr Asp Cys Val		
290	295	300
Thr Phe Leu Asn Leu Leu Glu Ser Leu Arg Ala Thr Glu Lys Ala Phe		
305	310	315
Gly Gln Asn Ser Gly Trp Leu Phe Leu Asp Ser Ser Thr Ser Met Phe		
325	330	335
Ile Asn Ala Arg Ala Arg Val Tyr His Leu Pro Asp Ala Lys Met Ser		
340	345	350
Lys Lys Glu Lys Ile Ser Glu Lys Met Glu Ile Lys Glu Gly Glu Gly		
355	360	365
Ile Leu Trp Gly		
370		
<210> 125		
<211> 287		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 125		
Pro Lys Arg Gly Lys Lys Gly Ala Val Ala Glu Asp Gly Asp Glu Leu		

1	5	10	15
Arg Thr Glu Pro Glu Ala Lys Lys Ser Lys Thr Ala Ala Lys Lys Asn			
20	25	30	
Asp Lys Glu Ala Ala Gly Glu Gly Pro Ala Leu Tyr Glu Asp Pro Pro			
35	40	45	
Asp Gln Lys Thr Ser Pro Ser Gly Lys Pro Ala Thr Leu Lys Ile Cys			
50	55	60	
Ser Trp Asn Val Asp Gly Leu Arg Ala Trp Ile Lys Lys Lys Gly Leu			
65	70	75	80
Asp Trp Val Lys Glu Glu Ala Pro Asp Ile Leu Cys Leu Gln Glu Thr			
85	90	95	
Lys Cys Ser Glu Asn Lys Leu Pro Ala Glu Leu Gln Glu Leu Pro Gly			
100	105	110	
Leu Ser His Gln Tyr Trp Ser Ala Pro Ser Asp Lys Glu Gly Tyr Ser			
115	120	125	
Gly Val Gly Leu Leu Ser Arg Gln Cys Pro Leu Lys Val Ser Tyr Gly			
130	135	140	
Ile Gly Asp Glu Glu His Asp Gln Glu Gly Arg Val Ile Val Ala Glu			
145	150	155	160
Phe Asp Ser Phe Val Leu Val Thr Ala Tyr Val Pro Asn Ala Gly Arg			
165	170	175	
Gly Leu Val Arg Leu Glu Tyr Arg Gln Arg Trp Asp Glu Ala Phe Arg			
180	185	190	
Lys Phe Leu Lys Gly Leu Ala Ser Arg Lys Pro Leu Val Leu Cys Gly			
195	200	205	
Asp Leu Asn Val Ala His Glu Glu Ile Asp Leu Arg Asn Pro Lys Gly			
210	215	220	
Asn Lys Lys Asn Ala Gly Phe Thr Pro Gln Glu Arg Gln Gly Phe Gly			
225	230	235	240
Glu Leu Leu Gln Ala Val Pro Leu Ala Asp Ser Phe Arg His Leu Tyr			
245	250	255	
Pro Asn Thr Pro Tyr Ala Tyr Thr Phe Trp Thr Tyr Met Met Asn Ala			
260	265	270	
Arg Ser Lys Asn Val Gly Trp Arg Leu Asp Tyr Phe Leu Leu Ser			
275	280	285	

<210> 126

<211> 213

<212> PRT

<213> 智人(Homo Sapiens)

<400> 126

Glu Ala Leu Phe Phe Pro Ser Gln Val Thr Cys Thr Glu Ala Leu Leu
 1 5 10 15
 Arg Ala Pro Gly Ala Glu Leu Ala Glu Leu Pro Glu Gly Cys Pro Cys
 20 25 30
 Gly Leu Pro His Gly Glu Ser Ala Leu Ser Arg Leu Leu Arg Ala Leu
 35 40 45
 Leu Ala Ala Arg Ala Ser Leu Asp Leu Cys Leu Phe Ala Phe Ser Ser
 50 55 60
 Pro Gln Leu Gly Arg Ala Val Gln Leu Leu His Gln Arg Gly Val Arg
 65 70 75 80
 Val Arg Val Val Thr Asp Cys Asp Tyr Met Ala Leu Asn Gly Ser Gln
 85 90 95
 Ile Gly Leu Leu Arg Lys Ala Gly Ile Gln Val Arg His Asp Gln Asp
 100 105 110
 Pro Gly Tyr Met His His Lys Phe Ala Ile Val Asp Lys Arg Val Leu
 115 120 125
 Ile Thr Gly Ser Leu Asn Trp Thr Thr Gln Ala Ile Gln Asn Asn Arg
 130 135 140
 Glu Asn Val Leu Ile Thr Glu Asp Asp Glu Tyr Val Arg Leu Phe Leu
 145 150 155 160
 Glu Glu Phe Glu Arg Ile Trp Glu Gln Phe Asn Pro Thr Lys Tyr Thr
 165 170 175
 Phe Phe Pro Pro Lys Lys Ser His Gly Ser Cys Ala Pro Pro Val Ser
 180 185 190
 Arg Ala Gly Gly Arg Leu Leu Ser Trp His Arg Thr Cys Gly Thr Ser
 195 200 205
 Ser Glu Ser Gln Thr
 210

<210> 127

<211> 382

<212> PRT

<213> 智人(Homo Sapiens)

<400> 127

Lys Ala Arg Tyr Lys Thr Leu Glu Pro Arg Gly Tyr Ser Leu Leu Ile
 1 5 10 15
 Arg Gly Leu Ile His Ser Asp Arg Trp Arg Glu Ala Leu Leu Leu Leu
 20 25 30
 Glu Asp Ile Lys Lys Val Ile Thr Pro Ser Lys Lys Asn Tyr Asn Asp

35	40	45
Cys Ile Gln Gly Ala Leu	Leu His Gln Asp Val	Asn Thr Ala Trp Asn
50	55	60
Leu Tyr Gln Glu Leu Leu	Gly His Asp Ile Val	Pro Met Leu Glu Thr
65	70	75
Leu Lys Ala Phe Phe Asp	Phe Gly Lys Asp Ile	Lys Asp Asp Asn Tyr
85	90	95
Ser Asn Lys Leu Leu Asp	Ile Leu Ser Tyr Leu	Arg Asn Asn Gln Leu
100	105	110
Tyr Pro Gly Glu Ser Phe	Ala His Ser Ile Lys	Thr Trp Phe Glu Ser
115	120	125
Val Pro Gly Lys Gln Trp	Lys Gly Gln Phe Thr	Thr Val Arg Lys Ser
130	135	140
Gly Gln Cys Ser Gly Cys	Gly Lys Thr Ile Glu	Ser Ile Gln Leu Ser
145	150	155
Pro Glu Glu Tyr Glu Cys	Leu Lys Gly Lys Ile	Met Arg Asp Val Ile
165	170	175
Asp Gly Gly Asp Gln Tyr	Arg Lys Thr Thr Pro	Gln Glu Leu Lys Arg
180	185	190
Phe Glu Asn Phe Ile Lys	Ser Arg Pro Pro Phe	Asp Val Val Ile Asp
195	200	205
Gly Leu Asn Val Ala Lys	Met Phe Pro Lys Val	Arg Glu Ser Gln Leu
210	215	220
Leu Leu Asn Val Val Ser	Gln Leu Ala Lys Arg	Asn Leu Arg Leu Leu
225	230	235
Val Leu Gly Arg Lys His	Met Leu Arg Arg Ser	Ser Ser Gln Trp Ser Arg
245	250	255
Asp Glu Met Glu Glu Val	Gln Lys Gln Ala Ser	Cys Phe Phe Ala Asp
260	265	270
Asp Ile Ser Glu Asp Asp	Pro Phe Leu Leu Tyr	Ala Thr Leu His Ser
275	280	285
Gly Asn His Cys Arg Phe	Ile Thr Arg Asp Leu	Met Arg Asp His Lys
290	295	300
Ala Cys Leu Pro Asp Ala	Lys Thr Gln Arg Leu	Phe Phe Lys Trp Gln
305	310	315
Gln Gly His Gln Leu Ala	Ile Val Asn Arg Phe	Pro Gly Ser Lys Leu
325	330	335
Thr Phe Gln Arg Ile Leu	Ser Tyr Asp Thr Val	Val Gln Thr Thr Gly
340	345	350

Asp Ser Trp His Ile Pro Tyr Asp Glu Asp Leu Val Glu Arg Cys Ser
 355 360 365
 Cys Glu Val Pro Thr Lys Trp Leu Cys Leu His Gln Lys Thr
 370 375 380
 <210> 128
 <211> 360
 <212> PRT
 <213> 智人(Homo Sapiens)
 <400> 128
 Ser Val Glu Pro Met Phe Arg His Leu Lys Asn Thr Tyr Ala Gly Leu
 1 5 10 15
 Gln Leu Val Val Val Ile Leu Pro Gly Lys Thr Pro Val Tyr Ala Glu
 20 25 30
 Val Lys Arg Val Gly Asp Thr Val Leu Gly Met Ala Thr Gln Cys Val
 35 40 45
 Gln Met Lys Asn Val Gln Arg Thr Thr Pro Gln Thr Leu Ser Asn Leu
 50 55 60
 Cys Leu Lys Ile Asn Val Lys Leu Gly Gly Val Asn Asn Ile Leu Leu
 65 70 75 80
 Pro Gln Gly Arg Pro Pro Val Phe Gln Gln Pro Val Ile Phe Leu Gly
 85 90 95
 Ala Asp Val Thr His Pro Pro Ala Gly Asp Gly Lys Lys Pro Ser Ile
 100 105 110
 Ala Ala Val Val Gly Ser Met Asp Ala His Pro Asn Arg Tyr Cys Ala
 115 120 125
 Thr Val Arg Val Gln Gln His Arg Gln Glu Ile Ile Gln Asp Leu Ala
 130 135 140
 Ala Met Val Arg Glu Leu Leu Ile Gln Phe Tyr Lys Ser Thr Arg Phe
 145 150 155 160
 Lys Pro Thr Arg Ile Ile Phe Tyr Arg Asp Gly Val Ser Glu Gly Gln
 165 170 175
 Phe Gln Gln Val Leu His His Glu Leu Leu Ala Ile Arg Glu Ala Cys
 180 185 190
 Ile Lys Leu Glu Lys Asp Tyr Gln Pro Gly Ile Thr Phe Ile Val Val
 195 200 205
 Gln Lys Arg His His Thr Arg Leu Phe Cys Thr Asp Lys Asn Glu Arg
 210 215 220
 Val Gly Lys Ser Gly Asn Ile Pro Ala Gly Thr Thr Val Asp Thr Lys
 225 230 235 240

Ile Thr His Pro Thr Glu Phe Asp Phe Tyr Leu Cys Ser His Ala Gly		
	245	250 255
Ile Gln Gly Thr Ser Arg Pro Ser His Tyr His Val Leu Trp Asp Asp		
	260	265 270
Asn Arg Phe Ser Ser Asp Glu Leu Gln Ile Leu Thr Tyr Gln Leu Cys		
	275	280 285
His Thr Tyr Val Arg Cys Thr Arg Ser Val Ser Ile Pro Ala Pro Ala		
	290	295 300
Tyr Tyr Ala His Leu Val Ala Phe Arg Ala Arg Tyr His Leu Val Asp		
305	310	315 320
Lys Glu His Asp Ser Ala Glu Gly Ser His Thr Ser Gly Gln Ser Asn		
	325	330 335
Gly Arg Asp His Gln Ala Leu Ala Lys Ala Val Gln Val His Gln Asp		
	340	345 350
Thr Leu Arg Thr Met Tyr Phe Ala		
	355	360
<210> 129		
<211> 327		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 129		
Gln Gly Ala Glu Gly Ala Leu Thr Gly Lys Gln Pro Asp Gly Ser Ala		
1	5	10 15
Glu Lys Ala Val Leu Glu Gln Phe Gly Phe Pro Leu Thr Gly Thr Glu		
	20	25 30
Ala Arg Cys Tyr Thr Asn His Ala Leu Ser Tyr Asp Gln Ala Lys Arg		
	35	40 45
Val Pro Arg Trp Val Leu Glu His Ile Ser Lys Ser Lys Ile Met Gly		
	50	55 60
Asp Ala Asp Arg Lys His Cys Lys Phe Lys Pro Asp Pro Asn Ile Pro		
65	70	75 80
Pro Thr Phe Ser Ala Phe Asn Glu Asp Tyr Val Gly Ser Gly Trp Ser		
	85	90 95
Arg Gly His Met Ala Pro Ala Gly Asn Asn Lys Phe Ser Ser Lys Ala		
	100	105 110
Met Ala Glu Thr Phe Tyr Leu Ser Asn Ile Val Pro Gln Asp Phe Asp		
	115	120 125
Asn Asn Ser Gly Tyr Trp Asn Arg Ile Glu Met Tyr Cys Arg Glu Leu		
	130	135 140

Thr	Glu	Arg	Phe	Glu	Asp	Val	Trp	Val	Val	Ser	Gly	Pro	Leu	Thr	Leu
145					150					155					160
Pro	Gln	Thr	Arg	Gly	Asp	Gly	Lys	Lys	Ile	Val	Ser	Tyr	Gln	Val	Ile
				165					170					175	
Gly	Glu	Asp	Asn	Val	Ala	Val	Pro	Ser	His	Leu	Tyr	Lys	Val	Ile	Leu
			180					185					190		
Ala	Arg	Arg	Ser	Ser	Val	Ser	Thr	Glu	Pro	Leu	Ala	Leu	Gly	Ala	Phe
			195				200					205			
Val	Val	Pro	Asn	Glu	Ala	Ile	Gly	Phe	Gln	Pro	Gln	Leu	Thr	Glu	Phe
	210					215				220					
Gln	Val	Ser	Leu	Gln	Asp	Leu	Glu	Lys	Leu	Ser	Gly	Leu	Val	Phe	Phe
225				230					235						240
Pro	His	Leu	Asp	Arg	Thr	Ser	Asp	Ile	Arg	Asn	Ile	Cys	Ser	Val	Asp
			245					250						255	
Thr	Cys	Lys	Leu	Leu	Asp	Phe	Gln	Glu	Phe	Thr	Leu	Tyr	Leu	Ser	Thr
			260					265					270		
Arg	Lys	Ile	Glu	Gly	Ala	Arg	Ser	Val	Leu	Arg	Leu	Glu	Lys	Ile	Met
		275					280					285			
Glu	Asn	Leu	Lys	Asn	Ala	Glu	Ile	Glu	Pro	Asp	Asp	Tyr	Phe	Met	Ser
	290					295				300					
Arg	Tyr	Glu	Lys	Lys	Leu	Glu	Glu	Leu	Lys	Ala	Lys	Glu	Gln	Ser	Gly
305				310					315						320
Thr	Gln	Ile	Arg	Lys	Pro	Ser									
				325											

<210> 130

<211> 526

<212> PRT

<213> 智人(Homo Sapiens)

<400> 130

Glu	His	Pro	Ser	Lys	Met	Glu	Phe	Phe	Gln	Lys	Leu	Gly	Tyr	Asp	Arg
1				5					10					15	
Glu	Asp	Val	Leu	Arg	Val	Leu	Gly	Lys	Leu	Gly	Glu	Gly	Ala	Leu	Val
			20					25					30		
Asn	Asp	Val	Leu	Gln	Glu	Leu	Ile	Arg	Thr	Gly	Ser	Arg	Pro	Gly	Ala
		35					40					45			
Leu	Glu	His	Pro	Ala	Ala	Pro	Arg	Leu	Val	Pro	Arg	Gly	Ser	Cys	Gly
	50					55				60					
Val	Pro	Asp	Ser	Ala	Gln	Arg	Gly	Pro	Gly	Thr	Ala	Leu	Glu	Glu	Asp
65				70					75						80

Phe Arg Thr Leu Ala Ser Ser Leu Arg Pro Ile Val Ile Asp Gly Ser		
85	90	95
Asn Val Ala Met Ser His Gly Asn Lys Glu Thr Phe Ser Cys Arg Gly		
100	105	110
Ile Lys Leu Ala Val Asp Trp Phe Arg Asp Arg Gly His Thr Tyr Ile		
115	120	125
Lys Val Phe Val Pro Ser Trp Arg Lys Asp Pro Pro Arg Ala Asp Thr		
130	135	140
Pro Ile Arg Glu Gln His Val Leu Ala Glu Leu Glu Arg Gln Ala Val		
145	150	155
Leu Val Tyr Thr Pro Ser Arg Lys Val His Gly Lys Arg Leu Val Cys		
165	170	175
Tyr Asp Asp Arg Tyr Ile Val Lys Val Ala Tyr Glu Gln Asp Gly Val		
180	185	190
Ile Val Ser Asn Asp Asn Tyr Arg Asp Leu Gln Ser Glu Asn Pro Glu		
195	200	205
Trp Lys Trp Phe Ile Glu Gln Arg Leu Leu Met Phe Ser Phe Val Asn		
210	215	220
Asp Arg Phe Met Pro Pro Asp Asp Pro Leu Gly Arg His Gly Pro Ser		
225	230	235
Leu Ser Asn Phe Leu Ser Arg Lys Pro Lys Pro Pro Glu Pro Ser Trp		
245	250	255
Gln His Cys Pro Tyr Gly Lys Lys Cys Thr Tyr Gly Ile Lys Cys Lys		
260	265	270
Phe Tyr His Pro Glu Arg Pro His His Ala Gln Leu Ala Val Ala Asp		
275	280	285
Glu Leu Arg Ala Lys Thr Gly Ala Arg Pro Gly Ala Gly Ala Glu Glu		
290	295	300
Gln Arg Pro Pro Arg Ala Pro Gly Gly Ser Ala Gly Ala Arg Ala Ala		
305	310	315
Pro Arg Glu Pro Phe Ala His Ser Leu Pro Pro Ala Arg Gly Ser Pro		
325	330	335
Asp Leu Ala Ala Leu Arg Gly Ser Phe Ser Arg Leu Ala Phe Ser Asp		
340	345	350
Asp Leu Gly Pro Leu Gly Pro Pro Leu Pro Val Pro Ala Cys Ser Leu		
355	360	365
Thr Pro Arg Leu Gly Gly Pro Asp Trp Val Ser Ala Gly Gly Arg Val		
370	375	380
Pro Gly Pro Leu Ser Leu Pro Ser Pro Glu Ser Gln Phe Ser Pro Gly		

385	390	395	400
Asp Leu Pro Pro Pro Pro Gly Leu Gln Leu Gln Pro Arg Gly Glu His			
	405	410	415
Arg Pro Arg Asp Leu His Gly Asp Leu Leu Ser Pro Arg Arg Pro Pro			
	420	425	430
Asp Asp Pro Trp Ala Arg Pro Pro Arg Ser Asp Arg Phe Pro Gly Arg			
	435	440	445
Ser Val Trp Ala Glu Pro Ala Trp Gly Asp Gly Ala Thr Gly Gly Leu			
	450	455	460
Ser Val Tyr Ala Thr Glu Asp Asp Glu Gly Asp Ala Arg Ala Arg Ala			
465	470	475	480
Arg Ile Ala Leu Tyr Ser Val Phe Pro Arg Asp Gln Val Asp Arg Val			
	485	490	495
Met Ala Ala Phe Pro Glu Leu Ser Asp Leu Ala Arg Leu Ile Leu Leu			
	500	505	510
Val Gln Arg Cys Gln Ser Ala Gly Ala Pro Leu Gly Lys Pro			
	515	520	525
<210> 131			
<211> 475			
<212> PRT			
<213> 智人(Homo Sapiens)			
<400> 131			
Arg Gln Gln Gln Pro Gln Val Val Glu Lys Gln Gln Glu Thr Pro Leu			
1	5	10	15
Ala Pro Ala Asp Phe Ala His Ile Ser Gln Asp Ala Gln Ser Leu His			
	20	25	30
Ser Gly Ala Ser Arg Arg Ser Gln Lys Arg Leu Gln Ser Pro Ser Lys			
	35	40	45
Gln Ala Gln Pro Leu Asp Asp Pro Glu Ala Glu Gln Leu Thr Val Val			
	50	55	60
Gly Lys Ile Ser Phe Asn Pro Lys Asp Val Leu Gly Arg Gly Ala Gly			
65	70	75	80
Gly Thr Phe Val Phe Arg Gly Gln Phe Glu Gly Arg Ala Val Ala Val			
	85	90	95
Lys Arg Leu Leu Arg Glu Cys Phe Gly Leu Val Arg Arg Glu Val Gln			
	100	105	110
Leu Leu Gln Glu Ser Asp Arg His Pro Asn Val Leu Arg Tyr Phe Cys			
	115	120	125
Thr Glu Arg Gly Pro Gln Phe His Tyr Ile Ala Leu Glu Leu Cys Arg			

130	135	140
Ala Ser Leu Gln Glu Tyr Val	Glu Asn Pro Asp Leu Asp Arg Gly Gly	
145	150	155
Leu Glu Pro Glu Val Val	Leu Gln Gln Leu Met Ser Gly Leu Ala His	160
	165	170
Leu His Ser Leu His Ile Val	His Arg Asp Leu Lys Pro Gly Asn Ile	175
	180	185
Leu Ile Thr Gly Pro Asp Ser	Gln Gly Leu Gly Arg Val Val Leu Ser	190
	195	200
Asp Phe Gly Leu Cys Lys Lys	Leu Pro Ala Gly Arg Cys Ser Phe Ser	205
	210	220
Leu His Ser Gly Ile Pro Gly	Thr Glu Gly Trp Met Ala Pro Glu Leu	225
	230	235
Leu Gln Leu Leu Pro Pro Asp	Ser Pro Thr Ser Ala Val Asp Ile Phe	240
	245	250
Ser Ala Gly Cys Val Phe Tyr	Tyr Val Leu Ser Gly Gly Ser His Pro	255
	260	265
Phe Gly Asp Ser Leu Tyr Arg	Gln Ala Asn Ile Leu Thr Gly Ala Pro	270
	275	280
Cys Leu Ala His Leu Glu Glu	Glu Val His Asp Lys Val Val Ala Arg	285
	290	295
Asp Leu Val Gly Ala Met Leu	Ser Pro Leu Pro Gln Pro Arg Pro Ser	300
	305	310
Ala Pro Gln Val Leu Ala His	Pro Phe Phe Trp Ser Arg Ala Lys Gln	315
	320	325
Leu Gln Phe Phe Gln Asp Val	Ser Asp Trp Leu Glu Lys Glu Ser Glu	330
	335	340
Gln Glu Pro Leu Val Arg Ala	Leu Glu Ala Gly Gly Cys Ala Val Val	345
	350	355
Arg Asp Asn Trp His Glu His	Ile Ser Met Pro Leu Gln Thr Asp Leu	360
	365	370
Arg Lys Phe Arg Ser Tyr Lys	Gly Thr Ser Val Arg Asp Leu Leu Arg	375
	380	385
Ala Val Arg Asn Lys Lys His	His Tyr Arg Glu Leu Pro Val Glu Val	390
	395	400
	405	410
Arg Gln Ala Leu Gly Gln Val	Pro Asp Gly Phe Val Gln Tyr Phe Thr	415
	420	425
Asn Arg Phe Pro Arg Leu Leu	Leu His Thr His Arg Ala Met Arg Ser	430
	435	440
	445	

Cys Ala Ser Glu Ser Leu Phe Leu Pro Tyr Tyr Pro Pro Asp Ser Glu
 450 455 460
 Ala Arg Arg Pro Cys Pro Gly Ala Thr Gly Arg
 465 470 475
 <210> 132
 <211> 384
 <212> PRT
 <213> 智人(Homo Sapiens)
 <400> 132
 Lys Leu Val Arg Lys Asn Ile Glu Lys Asp Asn Ala Gly Gln Val Thr
 1 5 10 15
 Leu Val Pro Glu Glu Pro Glu Asp Met Trp His Thr Tyr Asn Leu Val
 20 25 30
 Gln Val Gly Asp Ser Leu Arg Ala Ser Thr Ile Arg Lys Val Gln Thr
 35 40 45
 Glu Ser Ser Thr Gly Ser Val Gly Ser Asn Arg Val Arg Thr Thr Leu
 50 55 60
 Thr Leu Cys Val Glu Ala Ile Asp Phe Asp Ser Gln Ala Cys Gln Leu
 65 70 75 80
 Arg Val Lys Gly Thr Asn Ile Gln Glu Asn Glu Tyr Val Lys Met Gly
 85 90 95
 Ala Tyr His Thr Ile Glu Leu Glu Pro Asn Arg Gln Phe Thr Leu Ala
 100 105 110
 Lys Lys Gln Trp Asp Ser Val Val Leu Glu Arg Ile Glu Gln Ala Cys
 115 120 125
 Asp Pro Ala Trp Ser Ala Asp Val Ala Ala Val Val Met Gln Glu Gly
 130 135 140
 Leu Ala His Ile Cys Leu Val Thr Pro Ser Met Thr Leu Thr Arg Ala
 145 150 155 160
 Lys Val Glu Val Asn Ile Pro Arg Lys Arg Lys Gly Asn Cys Ser Gln
 165 170 175
 His Asp Arg Ala Leu Glu Arg Phe Tyr Glu Gln Val Val Gln Ala Ile
 180 185 190
 Gln Arg His Ile His Phe Asp Val Val Lys Cys Ile Leu Val Ala Ser
 195 200 205
 Pro Gly Phe Val Arg Glu Gln Phe Cys Asp Tyr Leu Phe Gln Gln Ala
 210 215 220
 Val Lys Thr Asp Asn Lys Leu Leu Leu Glu Asn Arg Ser Lys Phe Leu
 225 230 235 240

Gln Val His Ala Ser Ser Gly His Lys Tyr Ser Leu Lys Glu Ala Leu			
	245	250	255
Cys Asp Pro Thr Val Ala Ser Arg Leu Ser Asp Thr Lys Ala Ala Gly			
	260	265	270
Glu Val Lys Ala Leu Asp Asp Phe Tyr Lys Met Leu Gln His Glu Pro			
	275	280	285
Asp Arg Ala Phe Tyr Gly Leu Lys Gln Val Glu Lys Ala Asn Glu Ala			
	290	295	300
Met Ala Ile Asp Thr Leu Leu Ile Ser Asp Glu Leu Phe Arg His Gln			
305	310	315	320
Asp Val Ala Thr Arg Ser Arg Tyr Val Arg Leu Val Asp Ser Val Lys			
	325	330	335
Glu Asn Ala Gly Thr Val Arg Ile Phe Ser Ser Leu His Val Ser Gly			
	340	345	350
Glu Gln Leu Ser Gln Leu Thr Gly Val Ala Ala Ile Leu Arg Phe Pro			
	355	360	365
Val Pro Glu Leu Ser Asp Gln Glu Gly Asp Ser Ser Ser Glu Glu Asp			
	370	375	380
<210> 133			
<211> 166			
<212> PRT			
<213> 智人(Homo Sapiens)			
<400> 133			
Ser Leu Val Ile Arg Asn Leu Gln Arg Val Ile Pro Ile Arg Arg Ala			
1	5	10	15
Pro Leu Arg Ser Lys Ile Glu Ile Val Arg Arg Ile Leu Gly Val Gln			
	20	25	30
Lys Phe Asp Leu Gly Ile Ile Cys Val Asp Asn Lys Asn Ile Gln His			
	35	40	45
Ile Asn Arg Ile Tyr Arg Asp Arg Asn Val Pro Thr Asp Val Leu Ser			
	50	55	60
Phe Pro Phe His Glu His Leu Lys Ala Gly Glu Phe Pro Gln Pro Asp			
65	70	75	80
Phe Pro Asp Asp Tyr Asn Leu Gly Asp Ile Phe Leu Gly Val Glu Tyr			
	85	90	95
Ile Phe His Gln Cys Lys Glu Asn Glu Asp Tyr Asn Asp Val Leu Thr			
	100	105	110
Val Thr Ala Thr His Gly Leu Cys His Leu Leu Gly Phe Thr His Gly			
	115	120	125

Thr Glu Ala Glu Trp Gln Gln Met Phe Gln Lys Glu Lys Ala Val Leu
 130 135 140
 Asp Glu Leu Gly Arg Arg Thr Gly Thr Arg Leu Gln Pro Leu Thr Arg
 145 150 155 160
 Gly Leu Phe Gly Gly Ser
 165

<210> 134

<211> 178

<212> PRT

<213> 智人(Homo Sapiens)

<400> 134

Gln Glu Val Ile Ala Gly Leu Glu Arg Phe Thr Phe Ala Phe Glu Lys
 1 5 10 15
 Asp Val Glu Met Gln Lys Gly Thr Gly Leu Leu Pro Phe Gln Gly Met
 20 25 30
 Asp Lys Ser Ala Ser Ala Val Cys Asn Phe Phe Thr Lys Gly Leu Cys
 35 40 45
 Glu Lys Gly Lys Leu Cys Pro Phe Arg His Asp Arg Gly Glu Lys Met
 50 55 60
 Val Val Cys Lys His Trp Leu Arg Gly Leu Cys Lys Lys Gly Asp His
 65 70 75 80
 Cys Lys Phe Leu His Gln Tyr Asp Leu Thr Arg Met Pro Glu Cys Tyr
 85 90 95
 Phe Tyr Ser Lys Phe Gly Asp Cys Ser Asn Lys Glu Cys Ser Phe Leu
 100 105 110
 His Val Lys Pro Ala Phe Lys Ser Gln Asp Cys Pro Trp Tyr Asp Gln
 115 120 125
 Gly Phe Cys Lys Asp Gly Pro Leu Cys Lys Tyr Arg His Val Pro Arg
 130 135 140
 Ile Met Cys Leu Asn Tyr Leu Val Gly Phe Cys Pro Glu Gly Pro Lys
 145 150 155 160
 Cys Gln Phe Ala Gln Lys Ile Arg Glu Phe Lys Leu Leu Pro Gly Ser
 165 170 175

Lys Ile

<210> 135

<211> 384

<212> PRT

<213> 智人(Homo Sapiens)

<400> 135

Lys	Leu	Val	Arg	Lys	Asn	Ile	Glu	Lys	Asp	Asn	Ala	Gly	Gln	Val	Thr
1				5					10					15	
Leu	Val	Pro	Glu	Glu	Pro	Glu	Asp	Met	Trp	His	Thr	Tyr	Asn	Leu	Val
			20					25					30		
Gln	Val	Gly	Asp	Ser	Leu	Arg	Ala	Ser	Thr	Ile	Arg	Lys	Val	Gln	Thr
		35					40					45			
Glu	Ser	Ser	Thr	Gly	Ser	Val	Gly	Ser	Asn	Arg	Val	Arg	Thr	Thr	Leu
	50					55					60				
Thr	Leu	Cys	Val	Glu	Ala	Ile	Asp	Phe	Asp	Ser	Gln	Ala	Cys	Gln	Leu
65					70				75					80	
Arg	Val	Lys	Gly	Thr	Asn	Ile	Gln	Glu	Asn	Glu	Tyr	Val	Lys	Met	Gly
				85					90					95	
Ala	Tyr	His	Thr	Ile	Glu	Leu	Glu	Pro	Asn	Arg	Gln	Phe	Thr	Leu	Ala
			100					105					110		
Lys	Lys	Gln	Trp	Asp	Ser	Val	Val	Leu	Glu	Arg	Ile	Glu	Gln	Ala	Cys
		115					120					125			
Asp	Pro	Ala	Trp	Ser	Ala	Asp	Val	Ala	Ala	Val	Val	Met	Gln	Glu	Gly
	130					135					140				
Leu	Ala	His	Ile	Cys	Leu	Val	Thr	Pro	Ser	Met	Thr	Leu	Thr	Arg	Ala
145					150				155					160	
Lys	Val	Glu	Val	Asn	Ile	Pro	Arg	Lys	Arg	Lys	Gly	Asn	Cys	Ser	Gln
				165					170				175		
His	Asp	Arg	Ala	Leu	Glu	Arg	Phe	Tyr	Glu	Gln	Val	Val	Gln	Ala	Ile
			180					185					190		
Gln	Arg	His	Ile	His	Phe	Asp	Val	Val	Lys	Cys	Ile	Leu	Val	Ala	Ser
	195						200					205			
Pro	Gly	Phe	Val	Arg	Glu	Gln	Phe	Cys	Asp	Tyr	Met	Phe	Gln	Gln	Ala
	210					215					220				
Val	Lys	Thr	Asp	Asn	Lys	Leu	Leu	Leu	Glu	Asn	Arg	Ser	Lys	Phe	Leu
225					230				235					240	
Gln	Val	His	Ala	Ser	Ser	Gly	His	Lys	Tyr	Ser	Leu	Lys	Glu	Ala	Leu
				245					250				255		
Cys	Asp	Pro	Thr	Val	Ala	Ser	Arg	Leu	Ser	Asp	Thr	Lys	Ala	Ala	Gly
		260						265				270			
Glu	Val	Lys	Ala	Leu	Asp	Asp	Phe	Tyr	Lys	Met	Leu	Gln	His	Glu	Pro
	275						280					285			
Asp	Arg	Ala	Phe	Tyr	Gly	Leu	Lys	Gln	Val	Glu	Lys	Ala	Asn	Glu	Ala
	290					295					300				
Met	Ala	Ile	Asp	Thr	Leu	Leu	Ile	Ser	Asp	Glu	Leu	Phe	Arg	His	Gln

305	310	315	320
Asp Val Ala Thr Arg Ser Arg Tyr Val Arg Leu Val Asp Ser Val Lys			
	325	330	335
Glu Asn Ala Gly Thr Val Arg Ile Phe Ser Ser Leu His Val Ser Gly			
	340	345	350
Glu Gln Leu Ser Gln Leu Thr Gly Val Ala Ala Ile Leu Arg Phe Pro			
	355	360	365
Val Pro Glu Leu Ser Asp Gln Glu Gly Asp Ser Ser Ser Glu Glu Asp			
	370	375	380
<210> 136			
<211> 256			
<212> PRT			
<213> 智人(Homo Sapiens)			
<400> 136			
Asp Pro Ala Trp Ser Ala Asp Val Ala Ala Val Val Met Gln Glu Gly			
1	5	10	15
Leu Ala His Ile Cys Leu Val Thr Pro Ser Met Thr Leu Thr Arg Ala			
	20	25	30
Lys Val Glu Val Asn Ile Pro Arg Lys Arg Lys Gly Asn Cys Ser Gln			
	35	40	45
His Asp Arg Ala Leu Glu Arg Phe Tyr Glu Gln Val Val Gln Ala Ile			
	50	55	60
Gln Arg His Ile His Phe Asp Val Val Lys Cys Ile Leu Val Ala Ser			
65	70	75	80
Pro Gly Phe Val Arg Glu Gln Phe Cys Asp Tyr Met Phe Gln Gln Ala			
	85	90	95
Val Lys Thr Asp Asn Lys Leu Leu Leu Glu Asn Arg Ser Lys Phe Leu			
	100	105	110
Gln Val His Ala Ser Ser Gly His Lys Tyr Ser Leu Lys Glu Ala Leu			
	115	120	125
Cys Asp Pro Thr Val Ala Ser Arg Leu Ser Asp Thr Lys Ala Ala Gly			
	130	135	140
Glu Val Lys Ala Leu Asp Asp Phe Tyr Lys Met Leu Gln His Glu Pro			
145	150	155	160
Asp Arg Ala Phe Tyr Gly Leu Lys Gln Val Glu Lys Ala Asn Glu Ala			
	165	170	175
Met Ala Ile Asp Thr Leu Leu Ile Ser Asp Glu Leu Phe Arg His Gln			
	180	185	190
Asp Val Ala Thr Arg Ser Arg Tyr Val Arg Leu Val Asp Ser Val Lys			

195	200	205
Glu Asn Ala Gly Thr Val Arg Ile Phe Ser Ser Leu His Val Ser Gly		
210	215	220
Glu Gln Leu Ser Gln Leu Thr Gly Val Ala Ala Ile Leu Arg Phe Pro		
225	230	235
Val Pro Glu Leu Ser Asp Gln Glu Gly Asp Ser Ser Ser Glu Glu Asp		
245	250	255
<210> 137		
<211> 209		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 137		
Met Asp Pro Gly Lys Asp Lys Glu Gly Val Pro Gln Pro Ser Gly Pro		
1	5	10
Pro Ala Arg Lys Lys Phe Val Ile Pro Leu Asp Glu Asp Glu Val Pro		
20	25	30
Pro Gly Val Arg Gly Asn Pro Val Leu Lys Phe Val Arg Asn Val Pro		
35	40	45
Trp Glu Phe Gly Asp Val Ile Pro Asp Tyr Val Leu Gly Gln Ser Thr		
50	55	60
Cys Ala Leu Phe Leu Ser Leu Arg Tyr His Asn Leu His Pro Asp Tyr		
65	70	75
Ile His Gly Arg Leu Gln Ser Leu Gly Lys Asn Phe Ala Leu Arg Val		
85	90	95
Leu Leu Val Gln Val Asp Val Lys Asp Pro Gln Gln Ala Leu Lys Glu		
100	105	110
Leu Ala Lys Met Cys Ile Leu Ala Asp Cys Thr Leu Ile Leu Ala Trp		
115	120	125
Ser Pro Glu Glu Ala Gly Arg Tyr Leu Glu Thr Tyr Lys Ala Tyr Glu		
130	135	140
Gln Lys Pro Ala Asp Leu Leu Met Glu Lys Leu Glu Gln Asp Phe Val		
145	150	155
Ser Arg Val Thr Glu Cys Leu Thr Thr Val Lys Ser Val Asn Lys Thr		
165	170	175
Asp Ser Gln Thr Leu Leu Thr Thr Phe Gly Ser Leu Glu Gln Leu Ile		
180	185	190
Ala Ala Ser Arg Glu Asp Leu Ala Leu Cys Pro Gly Leu Gly Pro Gln		
195	200	205
Lys		

<210> 138

<211> 128

<212> PRT

<213> 智人(Homo Sapiens)

<400> 138

```

Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp
1           5           10           15
Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg
           20           25           30
Arg Asn Met Thr Gln Gly Arg Cys Lys Pro Val Asn Thr Phe Val His
           35           40           45
Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr
           50           55           60
Cys Lys Asn Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His
65           70           75           80
Ile Thr Asp Cys Arg Leu Thr Asn Gly Ser Arg Tyr Pro Asn Cys Ala
           85           90           95
Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly
           100          105          110
Ser Pro Tyr Val Pro Val His Phe Asp Ala Ser Val Glu Asp Ser Thr
           115          120          125

```

<210> 139

<211> 123

<212> PRT

<213> 智人(Homo Sapiens)

<400> 139

```

Gln Asp Asn Ser Arg Tyr Thr His Phe Leu Thr Gln His Tyr Asp Ala
1           5           10           15
Lys Pro Gln Gly Arg Asp Asp Arg Tyr Cys Glu Ser Ile Met Arg Arg
           20           25           30
Arg Gly Leu Thr Ser Pro Cys Lys Asp Ile Asn Thr Phe Ile His Gly
           35           40           45
Asn Lys Arg Ser Ile Lys Ala Ile Cys Glu Asn Lys Asn Gly Asn Pro
           50           55           60
His Arg Glu Asn Leu Arg Ile Ser Lys Ser Ser Phe Gln Val Thr Thr
65           70           75           80
Cys Lys Leu His Gly Gly Ser Pro Trp Pro Pro Cys Gln Tyr Arg Ala
           85           90           95
Thr Ala Gly Phe Arg Asn Val Val Val Ala Cys Glu Asn Gly Leu Pro

```

100	105	110
Val His Leu Asp Gln Ser Ile Phe Arg Arg Pro		
115	120	
<210> 140		
<211> 129		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 140		
Gly Leu Gly Leu Val Gln Pro Ser Tyr Gly Gln Asp Gly Met Tyr Gln		
1 5 10 15		
Arg Phe Leu Arg Gln His Val His Pro Glu Glu Thr Gly Gly Ser Asp		
20 25 30		
Arg Tyr Cys Asn Leu Met Met Gln Arg Arg Lys Met Thr Leu Tyr His		
35 40 45		
Cys Lys Arg Phe Asn Thr Phe Ile His Glu Asp Ile Trp Asn Ile Arg		
50 55 60		
Ser Ile Cys Ser Thr Thr Asn Ile Gln Cys Lys Asn Gly Lys Met Asn		
65 70 75 80		
Cys His Glu Gly Val Val Lys Val Thr Asp Cys Arg Asp Thr Gly Ser		
85 90 95		
Ser Arg Ala Pro Asn Cys Arg Tyr Arg Ala Ile Ala Ser Thr Arg Arg		
100 105 110		
Val Val Ile Ala Cys Glu Gly Asn Pro Gln Val Pro Val His Phe Asp		
115 120 125		
Gly		
<210> 141		
<211> 375		
<212> PRT		
<213> 智人(Homo Sapiens)		
<220>		
<221> 尚未归类的特征		
<222> (1) .. (1)		
<223> Xaa可以是任何天然存在的氨基酸		
<400> 141		
Xaa Ser Ala Val Asp Asn Ile Leu Leu Lys Leu Ala Lys Phe Lys Ile		
1 5 10 15		
Gly Phe Leu Arg Leu Gly Gln Ile Gln Lys Val His Pro Ala Ile Gln		
20 25 30		
Gln Phe Thr Glu Gln Glu Ile Cys Arg Ser Lys Ser Ile Lys Ser Leu		

35	40	45
Ala Leu Leu Glu Glu Leu Tyr Asn Ser Gln Leu Ile Val Ala Thr Thr		
50	55	60
Cys Met Gly Ile Asn His Pro Ile Phe Ser Arg Lys Ile Phe Asp Phe		
65	70	75
Cys Ile Val Asp Glu Ala Ser Gln Ile Ser Gln Pro Ile Cys Leu Gly		
85	90	95
Pro Leu Phe Phe Ser Arg Arg Phe Val Leu Val Gly Asp His Gln Gln		
100	105	110
Leu Pro Pro Leu Val Leu Asn Arg Glu Ala Arg Ala Leu Gly Met Ser		
115	120	125
Glu Ser Leu Phe Lys Arg Leu Glu Gln Asn Lys Ser Ala Val Val Gln		
130	135	140
Leu Thr Val Gln Tyr Arg Met Asn Ser Lys Ile Met Ser Leu Ser Asn		
145	150	155
Lys Leu Thr Tyr Glu Gly Lys Leu Glu Cys Gly Ser Asp Lys Val Ala		
165	170	175
Asn Ala Val Ile Asn Leu Arg His Phe Lys Asp Val Lys Leu Glu Leu		
180	185	190
Glu Phe Tyr Ala Asp Tyr Ser Asp Asn Pro Trp Leu Met Gly Val Phe		
195	200	205
Glu Pro Asn Asn Pro Val Cys Phe Leu Asn Thr Asp Lys Val Pro Ala		
210	215	220
Pro Glu Gln Val Glu Lys Gly Gly Val Ser Asn Val Thr Glu Ala Lys		
225	230	235
Leu Ile Val Phe Leu Thr Ser Ile Phe Val Lys Ala Gly Cys Ser Pro		
245	250	255
Ser Asp Ile Gly Ile Ile Ala Pro Tyr Arg Gln Gln Leu Lys Ile Ile		
260	265	270
Asn Asp Leu Leu Ala Arg Ser Ile Gly Met Val Glu Val Asn Thr Val		
275	280	285
Asp Lys Tyr Gln Gly Arg Asp Lys Ser Ile Val Leu Val Ser Phe Val		
290	295	300
Arg Ser Asn Lys Asp Gly Thr Val Gly Glu Leu Leu Lys Asp Trp Arg		
305	310	315
Arg Leu Asn Val Ala Ile Thr Arg Ala Lys His Lys Leu Ile Leu Leu		
325	330	335
Gly Cys Val Pro Ser Leu Asn Cys Tyr Pro Pro Leu Glu Lys Leu Leu		
340	345	350

Asn His Leu Asn Ser Glu Lys Leu Ile Ser Phe Phe Phe Cys Ile Trp
 355 360 365

Ser His Leu Ile Ala Leu Leu
 370 375

<210> 142

<211> 88

<212> PRT

<213> 智人(Homo Sapiens)

<400> 142

Met Ala Leu Arg Ser His Asp Arg Ser Thr Arg Pro Leu Tyr Ile Ser
 1 5 10 15
 Val Gly His Arg Met Ser Leu Glu Ala Ala Val Arg Leu Thr Cys Cys
 20 25 30
 Cys Cys Arg Phe Arg Ile Pro Glu Pro Val Arg Gln Ala Asp Ile Cys
 35 40 45
 Ser Arg Glu His Ile Arg Lys Ser Leu Gly Leu Pro Gly Pro Pro Thr
 50 55 60
 Pro Arg Ser Pro Lys Ala Gln Arg Pro Val Ala Cys Pro Lys Gly Asp
 65 70 75 80
 Ser Gly Glu Ser Ser Ala Leu Cys
 85

<210> 143

<211> 212

<212> PRT

<213> 智人(Homo Sapiens)

<400> 143

Cys Tyr Thr Asn His Ala Leu Ser Tyr Asp Gln Ala Lys Arg Val Pro
 1 5 10 15
 Arg Trp Val Leu Glu His Ile Ser Lys Ser Lys Ile Met Gly Asp Ala
 20 25 30
 Asp Arg Lys His Cys Lys Phe Lys Pro Asp Pro Asn Ile Pro Pro Thr
 35 40 45
 Phe Ser Ala Phe Asn Glu Asp Tyr Val Gly Ser Gly Trp Ser Arg Gly
 50 55 60
 His Met Ala Pro Ala Gly Asn Asn Lys Phe Ser Ser Lys Ala Met Ala
 65 70 75 80
 Glu Thr Phe Tyr Leu Ser Asn Ile Val Pro Gln Asp Phe Asp Asn Asn
 85 90 95
 Ser Gly Tyr Trp Asn Arg Ile Glu Met Tyr Cys Arg Glu Leu Thr Glu

100	105	110
Arg Phe Glu Asp Val Trp Val Val Ser Gly Pro Leu Thr Leu Pro Gln		
115	120	125
Thr Arg Gly Asp Gly Lys Lys Ile Val Ser Tyr Gln Val Ile Gly Glu		
130	135	140
Asp Asn Val Ala Val Pro Ser His Leu Tyr Lys Val Ile Leu Ala Arg		
145	150	155
Arg Ser Ser Val Ser Thr Glu Pro Leu Ala Leu Gly Ala Phe Val Val		
165	170	175
Pro Asn Glu Ala Ile Gly Phe Gln Pro Gln Leu Thr Glu Phe Gln Val		
180	185	190
Ser Leu Gln Asp Leu Glu Lys Leu Ser Gly Leu Val Phe Phe Pro His		
195	200	205
Leu Asp Arg Thr		
210		
<210> 144		
<211> 123		
<212> PRT		
<213> 智人(Homo Sapiens)		
<400> 144		
Val Thr Val Ser Gln Met Thr Ser Val Leu Asn Gly Lys Thr Arg Arg		
1	5	10
Phe Ala Asp Ile Gln Leu Gln His Gly Ala Leu Cys Phe Asn Ile Arg		
20	25	30
Tyr Gly Thr Thr Val Glu Glu Glu Lys Asn His Val Leu Glu Ile Ala		
35	40	45
Arg Gln Arg Ala Val Ala Gln Ala Trp Thr Lys Glu Gln Arg Arg Leu		
50	55	60
Gln Glu Gly Glu Glu Gly Ile Arg Ala Trp Thr Glu Gly Glu Lys Gln		
65	70	75
Gln Leu Leu Ser Thr Gly Arg Val Gln Gly Tyr Asp Gly Tyr Phe Val		
85	90	95
Leu Ser Val Glu Gln Tyr Leu Glu Leu Ser Asp Ser Ala Asn Asn Ile		
100	105	110
His Phe Met Arg Gln Ser Glu Ile Gly Arg Arg		
115	120	
<210> 145		
<211> 125		
<212> PRT		

<213> 智人 (Homo Sapiens)

<400> 145

```

Thr Val Ser Gln Pro Thr Leu Leu Val Asn Gly Lys Thr Arg Arg Phe
1           5           10           15
Thr Asn Ile Glu Phe Gln Tyr Ser Thr Leu Leu Leu Ser Ile Arg Tyr
          20           25           30
Gly Leu Thr Pro Asp Thr Leu Asp Glu Glu Lys Ala Arg Val Leu Asp
          35           40           45
Gln Ala Arg Gln Arg Ala Leu Gly Thr Ala Trp Ala Lys Glu Gln Gln
          50           55           60
Lys Ala Arg Asp Gly Arg Glu Gly Ser Arg Leu Trp Thr Glu Gly Glu
65           70           75           80
Lys Gln Gln Leu Leu Ser Thr Gly Arg Val Gln Gly Tyr Glu Gly Tyr
          85           90           95
Tyr Val Leu Pro Val Glu Gln Tyr Pro Glu Leu Ala Asp Ser Ser Ser
          100          105          110
Asn Ile Gln Phe Leu Arg Gln Asn Glu Met Gly Lys Arg
          115          120          125

```

<210> 146

<400> 146

000

<210> 147

<211> 1368

<212> PRT

<213> 酿脓链球菌 (Streptococcus pyogenes)

<400> 147

```

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1           5           10           15
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
          20           25           30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
          35           40           45
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
          50           55           60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65           70           75           80
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
          85           90           95
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

```

100	105	110
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr		
115	120	125
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp		
130	135	140
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His		
145	150	155
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro		
165	170	175
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr		
180	185	190
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala		
195	200	205
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn		
210	215	220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn		
225	230	235
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe		
245	250	255
Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp		
260	265	270
Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp		
275	280	285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp		
290	295	300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser		
305	310	315
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys		
325	330	335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe		
340	345	350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser		
355	360	365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp		
370	375	380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg		
385	390	395
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu		
405	410	415

Gly	Glu	Leu	His	Ala	Ile	Leu	Arg	Arg	Gln	Glu	Asp	Phe	Tyr	Pro	Phe	420	425	430
Leu	Lys	Asp	Asn	Arg	Glu	Lys	Ile	Glu	Lys	Ile	Leu	Thr	Phe	Arg	Ile	435	440	445
Pro	Tyr	Tyr	Val	Gly	Pro	Leu	Ala	Arg	Gly	Asn	Ser	Arg	Phe	Ala	Trp	450	455	460
Met	Thr	Arg	Lys	Ser	Glu	Glu	Thr	Ile	Thr	Pro	Trp	Asn	Phe	Glu	Glu	465	470	475
Val	Val	Asp	Lys	Gly	Ala	Ser	Ala	Gln	Ser	Phe	Ile	Glu	Arg	Met	Thr	485	490	495
Asn	Phe	Asp	Lys	Asn	Leu	Pro	Asn	Glu	Lys	Val	Leu	Pro	Lys	His	Ser	500	505	510
Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Lys	515	520	525
Tyr	Val	Thr	Glu	Gly	Met	Arg	Lys	Pro	Ala	Phe	Leu	Ser	Gly	Glu	Gln	530	535	540
Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr	545	550	555
Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp	565	570	575
Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly	580	585	590
Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	595	600	605
Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr	610	615	620
Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala	625	630	635
His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr	645	650	655
Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp	660	665	670
Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe	675	680	685
Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe	690	695	700
Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu	705	710	715
His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly			

	725		730		735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly					
	740		745		750
Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln					
	755		760		765
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile					
	770		775		780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro					
785		790		795	800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu					
	805		810		815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg					
	820		825		830
Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys					
	835		840		845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg					
	850		855		860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys					
865		870		875	880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys					
	885		890		895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp					
	900		905		910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr					
	915		920		925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp					
	930		935		940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser					
945		950		955	960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg					
	965		970		975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val					
	980		985		990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe					
	995		1000		1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala					
	1010		1015		1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe					
	1025		1030		1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala		
1040	1045	1050
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu		
1055	1060	1065
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val		
1070	1075	1080
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr		
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys		
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro		
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		
1130	1135	1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly		
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val		
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys		
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys		
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
1295	1300	1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		

1325	1330	1335
Thr Lys Glu Val Leu Asp	Ala Thr Leu Ile His	Gln Ser Ile Thr
1340	1345	1350
Gly Leu Tyr Glu Thr Arg	Ile Asp Leu Ser Gln	Leu Gly Gly Asp
1355	1360	1365
<210> 148		
<211> 1368		
<212> PRT		
<213> 酿脓链球菌 (Streptococcus pyogenes)		
<400> 148		
Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val		
1	5	10
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe		
20	25	30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile		
35	40	45
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu		
50	55	60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys		
65	70	75
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser		
85	90	95
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys		
100	105	110
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr		
115	120	125
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp		
130	135	140
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His		
145	150	155
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro		
165	170	175
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr		
180	185	190
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala		
195	200	205
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn		
210	215	220
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn		

225	230	235	240
Leu Ile Ala Leu Ser	Leu Gly Leu Thr Pro	Asn Phe Lys Ser	Asn Phe
245	250	255	
Asp Leu Ala Glu Asp	Ala Lys Leu Gln Leu Ser	Lys Asp Thr Tyr	Asp
260	265	270	
Asp Asp Leu Asp Asn	Leu Leu Ala Gln Ile	Gly Asp Gln Tyr	Ala Asp
275	280	285	
Leu Phe Leu Ala Ala	Lys Asn Leu Ser Asp	Ala Ile Leu Leu	Ser Asp
290	295	300	
Ile Leu Arg Val Asn	Thr Glu Ile Thr Lys	Ala Pro Leu Ser	Ala Ser
305	310	315	320
Met Ile Lys Arg Tyr	Asp Glu His His Gln	Asp Leu Thr Leu	Leu Lys
325	330	335	
Ala Leu Val Arg Gln	Gln Leu Pro Glu Lys	Tyr Lys Glu Ile	Phe Phe
340	345	350	
Asp Gln Ser Lys Asn	Gly Tyr Ala Gly Tyr	Ile Asp Gly Gly	Ala Ser
355	360	365	
Gln Glu Glu Phe Tyr	Lys Phe Ile Lys Pro	Ile Leu Glu Lys	Met Asp
370	375	380	
Gly Thr Glu Glu Leu	Leu Val Lys Leu Asn	Arg Glu Asp Leu	Leu Arg
385	390	395	400
Lys Gln Arg Thr Phe	Asp Asn Gly Ser Ile	Pro His Gln Ile	His Leu
405	410	415	
Gly Glu Leu His Ala	Ile Leu Arg Arg Gln	Glu Asp Phe Tyr	Pro Phe
420	425	430	
Leu Lys Asp Asn Arg	Glu Lys Ile Glu Lys	Ile Leu Thr Phe	Arg Ile
435	440	445	
Pro Tyr Tyr Val Gly	Pro Leu Ala Arg Gly	Asn Ser Arg Phe	Ala Trp
450	455	460	
Met Thr Arg Lys Ser	Glu Glu Thr Ile Thr	Pro Trp Asn Phe	Glu Glu
465	470	475	480
Val Val Asp Lys Gly	Ala Ser Ala Gln Ser	Phe Ile Glu Arg	Met Thr
485	490	495	
Asn Phe Asp Lys Asn	Leu Pro Asn Glu Lys	Val Leu Pro Lys	His Ser
500	505	510	
Leu Leu Tyr Glu Tyr	Phe Thr Val Tyr Asn	Glu Leu Thr Lys	Val Lys
515	520	525	
Tyr Val Thr Glu Gly	Met Arg Lys Pro Ala	Phe Leu Ser Gly	Glu Gln
530	535	540	

Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr
545					550				555					560	
Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp
				565					570					575	
Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly
			580					585						590	
Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp
	595					600							605		
Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr
	610					615							620		
Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala
625				630						635					640
His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr
			645						650						655
Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp
			660					665						670	
Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe
	675							680						685	
Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe
	690					695								700	
Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu
705					710					715					720
His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly
			725						730						735
Ile	Leu	Gln	Thr	Val	Lys	Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly
			740						745						750
Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln
		755					760								765
Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile
	770						775								780
Glu	Glu	Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro
785					790					795					800
Val	Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu
			805						810						815
Gln	Asn	Gly	Arg	Asp	Met	Tyr	Val	Asp	Gln	Glu	Leu	Asp	Ile	Asn	Arg
			820						825					830	
Leu	Ser	Asp	Tyr	Asp	Val	Asp	Ala	Ile	Val	Pro	Gln	Ser	Phe	Leu	Lys
		835					840								845
Asp	Asp	Ser	Ile	Asp	Asn	Lys	Val	Leu	Thr	Arg	Ser	Asp	Lys	Asn	Arg

850	855	860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys		
865	870	875
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys		880
	885	890
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp		895
	900	905
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr		910
	915	920
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp		925
	930	935
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser		940
945	950	955
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg		960
	965	970
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val		975
	980	985
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe		990
	995	1000
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala		1005
	1010	1015
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe		1020
	1025	1030
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala		1035
	1040	1045
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu		1050
	1055	1060
Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val		1065
	1070	1075
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr		1080
	1085	1090
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys		1095
	1100	1105
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro		1110
	1115	1120
Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		1125
	1130	1135
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		1140
	1145	1150
		1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly		
1205	1210	1215
Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val		
1220	1225	1230
Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
1235	1240	1245
Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys		
1250	1255	1260
His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys		
1265	1270	1275
Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
1295	1300	1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr		
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
1355	1360	1365

<210> 149

<211> 1368

<212> PRT

<213> 酿脓链球菌(*Streptococcus pyogenes*)

<400> 149

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val		
1	5	10
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe		
20	25	30
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile		
35	40	45

Gly	Ala	Leu	Leu	Phe	Asp	Ser	Gly	Glu	Thr	Ala	Glu	Ala	Thr	Arg	Leu
50						55					60				
Lys	Arg	Thr	Ala	Arg	Arg	Arg	Tyr	Thr	Arg	Arg	Lys	Asn	Arg	Ile	Cys
65					70					75				80	
Tyr	Leu	Gln	Glu	Ile	Phe	Ser	Asn	Glu	Met	Ala	Lys	Val	Asp	Asp	Ser
				85					90					95	
Phe	Phe	His	Arg	Leu	Glu	Glu	Ser	Phe	Leu	Val	Glu	Glu	Asp	Lys	Lys
			100					105					110		
His	Glu	Arg	His	Pro	Ile	Phe	Gly	Asn	Ile	Val	Asp	Glu	Val	Ala	Tyr
		115					120				125				
His	Glu	Lys	Tyr	Pro	Thr	Ile	Tyr	His	Leu	Arg	Lys	Lys	Leu	Val	Asp
		130				135					140				
Ser	Thr	Asp	Lys	Ala	Asp	Leu	Arg	Leu	Ile	Tyr	Leu	Ala	Leu	Ala	His
145					150					155					160
Met	Ile	Lys	Phe	Arg	Gly	His	Phe	Leu	Ile	Glu	Gly	Asp	Leu	Asn	Pro
				165					170					175	
Asp	Asn	Ser	Asp	Val	Asp	Lys	Leu	Phe	Ile	Gln	Leu	Val	Gln	Thr	Tyr
			180					185					190		
Asn	Gln	Leu	Phe	Glu	Glu	Asn	Pro	Ile	Asn	Ala	Ser	Gly	Val	Asp	Ala
		195					200					205			
Lys	Ala	Ile	Leu	Ser	Ala	Arg	Leu	Ser	Lys	Ser	Arg	Arg	Leu	Glu	Asn
		210				215					220				
Leu	Ile	Ala	Gln	Leu	Pro	Gly	Glu	Lys	Lys	Asn	Gly	Leu	Phe	Gly	Asn
225					230					235					240
Leu	Ile	Ala	Leu	Ser	Leu	Gly	Leu	Thr	Pro	Asn	Phe	Lys	Ser	Asn	Phe
				245					250					255	
Asp	Leu	Ala	Glu	Asp	Ala	Lys	Leu	Gln	Leu	Ser	Lys	Asp	Thr	Tyr	Asp
			260					265					270		
Asp	Asp	Leu	Asp	Asn	Leu	Leu	Ala	Gln	Ile	Gly	Asp	Gln	Tyr	Ala	Asp
		275					280					285			
Leu	Phe	Leu	Ala	Ala	Lys	Asn	Leu	Ser	Asp	Ala	Ile	Leu	Leu	Ser	Asp
		290				295					300				
Ile	Leu	Arg	Val	Asn	Thr	Glu	Ile	Thr	Lys	Ala	Pro	Leu	Ser	Ala	Ser
305					310					315					320
Met	Ile	Lys	Arg	Tyr	Asp	Glu	His	His	Gln	Asp	Leu	Thr	Leu	Leu	Lys
				325					330					335	
Ala	Leu	Val	Arg	Gln	Gln	Leu	Pro	Glu	Lys	Tyr	Lys	Glu	Ile	Phe	Phe
			340					345					350		
Asp	Gln	Ser	Lys	Asn	Gly	Tyr	Ala	Gly	Tyr	Ile	Asp	Gly	Gly	Ala	Ser

355	360	365
Gln Glu Glu Phe Tyr Lys Phe	Ile Lys Pro Ile Leu Glu Lys Met Asp	
370	375	380
Gly Thr Glu Glu Leu Leu Val Lys	Leu Asn Arg Glu Asp Leu Leu Arg	
385	390	395
Lys Gln Arg Thr Phe Asp Asn Gly	Ser Ile Pro His Gln Ile His Leu	
405	410	415
Gly Glu Leu His Ala Ile Leu Arg	Arg Gln Glu Asp Phe Tyr Pro Phe	
420	425	430
Leu Lys Asp Asn Arg Glu Lys Ile	Glu Lys Ile Leu Thr Phe Arg Ile	
435	440	445
Pro Tyr Tyr Val Gly Pro Leu Ala	Arg Gly Asn Ser Arg Phe Ala Trp	
450	455	460
Met Thr Arg Lys Ser Glu Glu Thr	Ile Thr Pro Trp Asn Phe Glu Glu	
465	470	475
Val Val Asp Lys Gly Ala Ser Ala	Gln Ser Phe Ile Glu Arg Met Thr	
485	490	495
Asn Phe Asp Lys Asn Leu Pro Asn	Glu Lys Val Leu Pro Lys His Ser	
500	505	510
Leu Leu Tyr Glu Tyr Phe Thr Val	Tyr Asn Glu Leu Thr Lys Val Lys	
515	520	525
Tyr Val Thr Glu Gly Met Arg Lys	Pro Ala Phe Leu Ser Gly Glu Gln	
530	535	540
Lys Lys Ala Ile Val Asp Leu Leu	Phe Lys Thr Asn Arg Lys Val Thr	
545	550	555
Val Lys Gln Leu Lys Glu Asp Tyr	Phe Lys Lys Ile Glu Cys Phe Asp	
565	570	575
Ser Val Glu Ile Ser Gly Val Glu	Asp Arg Phe Asn Ala Ser Leu Gly	
580	585	590
Thr Tyr His Asp Leu Leu Lys Ile	Ile Lys Asp Lys Asp Phe Leu Asp	
595	600	605
Asn Glu Glu Asn Glu Asp Ile Leu	Glu Asp Ile Val Leu Thr Leu Thr	
610	615	620
Leu Phe Glu Asp Arg Glu Met Ile	Glu Glu Arg Leu Lys Thr Tyr Ala	
625	630	635
His Leu Phe Asp Asp Lys Val Met	Lys Gln Leu Lys Arg Arg Arg Tyr	
645	650	655
Thr Gly Trp Gly Arg Leu Ser Arg	Lys Leu Ile Asn Gly Ile Arg Asp	
660	665	670

Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe
675				680				685							
Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe
690				695				700							
Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu
705				710				715				720			
His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly
				725				730				735			
Ile	Leu	Gln	Thr	Val	Lys	Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly
				740				745				750			
Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln
				755				760				765			
Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile
				770				775				780			
Glu	Glu	Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro
785				790				795				800			
Val	Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu
				805				810				815			
Gln	Asn	Gly	Arg	Asp	Met	Tyr	Val	Asp	Gln	Glu	Leu	Asp	Ile	Asn	Arg
				820				825				830			
Leu	Ser	Asp	Tyr	Asp	Val	Asp	His	Ile	Val	Pro	Gln	Ser	Phe	Leu	Lys
				835				840				845			
Asp	Asp	Ser	Ile	Asp	Asn	Lys	Val	Leu	Thr	Arg	Ser	Asp	Lys	Asn	Arg
850				855				860							
Gly	Lys	Ser	Asp	Asn	Val	Pro	Ser	Glu	Glu	Val	Val	Lys	Lys	Met	Lys
865				870				875				880			
Asn	Tyr	Trp	Arg	Gln	Leu	Leu	Asn	Ala	Lys	Leu	Ile	Thr	Gln	Arg	Lys
				885				890				895			
Phe	Asp	Asn	Leu	Thr	Lys	Ala	Glu	Arg	Gly	Gly	Leu	Ser	Glu	Leu	Asp
				900				905				910			
Lys	Ala	Gly	Phe	Ile	Lys	Arg	Gln	Leu	Val	Glu	Thr	Arg	Gln	Ile	Thr
915				920				925							
Lys	His	Val	Ala	Gln	Ile	Leu	Asp	Ser	Arg	Met	Asn	Thr	Lys	Tyr	Asp
930				935				940							
Glu	Asn	Asp	Lys	Leu	Ile	Arg	Glu	Val	Lys	Val	Ile	Thr	Leu	Lys	Ser
945				950				955				960			
Lys	Leu	Val	Ser	Asp	Phe	Arg	Lys	Asp	Phe	Gln	Phe	Tyr	Lys	Val	Arg
				965				970				975			
Glu	Ile	Asn	Asn	Tyr	His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn	Ala	Val

980	985	990
Val Gly Thr Ala Leu Ile Lys	Lys Tyr Pro Lys Leu Glu Ser Glu Phe	
995	1000	1005
Val Tyr Gly Asp Tyr Lys	Val Tyr Asp Val Arg Lys Met Ile Ala	
1010	1015	1020
Lys Ser Glu Gln Glu Ile Gly	Lys Ala Thr Ala Lys Tyr Phe Phe	
1025	1030	1035
Tyr Ser Asn Ile Met Asn Phe	Phe Lys Thr Glu Ile Thr Leu Ala	
1040	1045	1050
Asn Gly Glu Ile Arg Lys	Arg Pro Leu Ile Glu Thr Asn Gly Glu	
1055	1060	1065
Thr Gly Glu Ile Val Trp	Asp Lys Gly Arg Asp Phe Ala Thr Val	
1070	1075	1080
Arg Lys Val Leu Ser Met	Pro Gln Val Asn Ile Val Lys Lys Thr	
1085	1090	1095
Glu Val Gln Thr Gly Gly	Phe Ser Lys Glu Ser Ile Leu Pro Lys	
1100	1105	1110
Arg Asn Ser Asp Lys Leu	Ile Ala Arg Lys Lys Asp Trp Asp Pro	
1115	1120	1125
Lys Lys Tyr Gly Gly Phe	Asp Ser Pro Thr Val Ala Tyr Ser Val	
1130	1135	1140
Leu Val Val Ala Lys Val	Glu Lys Gly Lys Ser Lys Lys Leu Lys	
1145	1150	1155
Ser Val Lys Glu Leu Leu	Gly Ile Thr Ile Met Glu Arg Ser Ser	
1160	1165	1170
Phe Glu Lys Asn Pro Ile	Asp Phe Leu Glu Ala Lys Gly Tyr Lys	
1175	1180	1185
Glu Val Lys Lys Asp Leu	Ile Ile Lys Leu Pro Lys Tyr Ser Leu	
1190	1195	1200
Phe Glu Leu Glu Asn Gly	Arg Lys Arg Met Leu Ala Ser Ala Gly	
1205	1210	1215
Glu Leu Gln Lys Gly Asn	Glu Leu Ala Leu Pro Ser Lys Tyr Val	
1220	1225	1230
Asn Phe Leu Tyr Leu Ala	Ser His Tyr Glu Lys Leu Lys Gly Ser	
1235	1240	1245
Pro Glu Asp Asn Glu Gln	Lys Gln Leu Phe Val Glu Gln His Lys	
1250	1255	1260
His Tyr Leu Asp Glu Ile	Ile Glu Gln Ile Ser Glu Phe Ser Lys	
1265	1270	1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
1280	1285	1290
Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
1295	1300	1305
Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
1310	1315	1320
Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		
1325	1330	1335
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr		
1340	1345	1350
Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
1355	1360	1365

<210> 150

<211> 1053

<212> PRT

<213> 金黄色葡萄球菌(Staphylococcus aureus)

<400> 150

Met Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val			
1	5	10	15
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly			
	20	25	30
Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg			
	35	40	45
Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile			
	50	55	60
Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His			
65	70	75	80
Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu			
	85	90	95
Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu			
	100	105	110
Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr			
	115	120	125
Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala			
	130	135	140
Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys			
145	150	155	160
Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr			
	165	170	175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
180	185	190
Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
195	200	205
Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
210	215	220
Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
225	230	235
Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
245	250	255
Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
260	265	270
Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
275	280	285
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
290	295	300
Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
305	310	315
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
325	330	335
Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
340	345	350
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
355	360	365
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
370	375	380
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
385	390	395
Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
405	410	415
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
420	425	430
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
435	440	445
Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
450	455	460
Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
465	470	475
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		

	485		490		495
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr					
	500		505		510
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp					
	515		520		525
Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu					
	530		535		540
Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro					
545		550		555	560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys					
	565		570		575
Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu					
	580		585		590
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile					
	595		600		605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu					
	610		615		620
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp					
625		630		635	640
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu					
	645		650		655
Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys					
	660		665		670
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp					
	675		680		685
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp					
	690		695		700
Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys					
705		710		715	720
Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys					
	725		730		735
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu					
	740		745		750
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp					
	755		760		765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile					
	770		775		780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu					
785		790		795	800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu			
805	810	815	
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His			
820	825	830	
Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly			
835	840	845	
Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr			
850	855	860	
Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile			
865	870	875	880
Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp			
885	890	895	
Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr			
900	905	910	
Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val			
915	920	925	
Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser			
930	935	940	
Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala			
945	950	955	960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly			
965	970	975	
Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile			
980	985	990	
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met			
995	1000	1005	
Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys			
1010	1015	1020	
Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu			
1025	1030	1035	
Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly			
1040	1045	1050	
<210> 151			
<211> 1121			
<212> PRT			
<213> 嗜热链球菌 (Streptococcus thermophilus)			
<400> 151			
Met Ser Asp Leu Val Leu Gly Leu Asp Ile Gly Ile Gly Ser Val Gly			
1	5	10	15

Val	Gly	Ile	Leu	Asn	Lys	Val	Thr	Gly	Glu	Ile	Ile	His	Lys	Asn	Ser
20				25				30							
Arg	Ile	Phe	Pro	Ala	Ala	Gln	Ala	Glu	Asn	Asn	Leu	Val	Arg	Arg	Thr
35				40				45							
Asn	Arg	Gln	Gly	Arg	Arg	Leu	Ala	Arg	Arg	Lys	Lys	His	Arg	Arg	Val
50				55				60							
Arg	Leu	Asn	Arg	Leu	Phe	Glu	Glu	Ser	Gly	Leu	Ile	Thr	Asp	Phe	Thr
65				70				75				80			
Lys	Ile	Ser	Ile	Asn	Leu	Asn	Pro	Tyr	Gln	Leu	Arg	Val	Lys	Gly	Leu
85				90				95							
Thr	Asp	Glu	Leu	Ser	Asn	Glu	Glu	Leu	Phe	Ile	Ala	Leu	Lys	Asn	Met
100				105				110							
Val	Lys	His	Arg	Gly	Ile	Ser	Tyr	Leu	Asp	Asp	Ala	Ser	Asp	Asp	Gly
115				120				125							
Asn	Ser	Ser	Val	Gly	Asp	Tyr	Ala	Gln	Ile	Val	Lys	Glu	Asn	Ser	Lys
130				135				140							
Gln	Leu	Glu	Thr	Lys	Thr	Pro	Gly	Gln	Ile	Gln	Leu	Glu	Arg	Tyr	Gln
145				150				155				160			
Thr	Tyr	Gly	Gln	Leu	Arg	Gly	Asp	Phe	Thr	Val	Glu	Lys	Asp	Gly	Lys
165				170				175							
Lys	His	Arg	Leu	Ile	Asn	Val	Phe	Pro	Thr	Ser	Ala	Tyr	Arg	Ser	Glu
180				185				190							
Ala	Leu	Arg	Ile	Leu	Gln	Thr	Gln	Gln	Glu	Phe	Asn	Pro	Gln	Ile	Thr
195				200				205							
Asp	Glu	Phe	Ile	Asn	Arg	Tyr	Leu	Glu	Ile	Leu	Thr	Gly	Lys	Arg	Lys
210				215				220							
Tyr	Tyr	His	Gly	Pro	Gly	Asn	Glu	Lys	Ser	Arg	Thr	Asp	Tyr	Gly	Arg
225				230				235				240			
Tyr	Arg	Thr	Ser	Gly	Glu	Thr	Leu	Asp	Asn	Ile	Phe	Gly	Ile	Leu	Ile
245				250				255							
Gly	Lys	Cys	Thr	Phe	Tyr	Pro	Asp	Glu	Phe	Arg	Ala	Ala	Lys	Ala	Ser
260				265				270							
Tyr	Thr	Ala	Gln	Glu	Phe	Asn	Leu	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Thr
275				280				285							
Val	Pro	Thr	Glu	Thr	Lys	Lys	Leu	Ser	Lys	Glu	Gln	Lys	Asn	Gln	Ile
290				295				300							
Ile	Asn	Tyr	Val	Lys	Asn	Glu	Lys	Ala	Met	Gly	Pro	Ala	Lys	Leu	Phe
305				310				315				320			
Lys	Tyr	Ile	Ala	Lys	Leu	Leu	Ser	Cys	Asp	Val	Ala	Asp	Ile	Lys	Gly

	325		330		335
Tyr Arg Ile Asp Lys Ser Gly Lys Ala Glu Ile His Thr Phe Glu Ala					
	340		345		350
Tyr Arg Lys Met Lys Thr Leu Glu Thr Leu Asp Ile Glu Gln Met Asp					
	355		360		365
Arg Glu Thr Leu Asp Lys Leu Ala Tyr Val Leu Thr Leu Asn Thr Glu					
	370		375		380
Arg Glu Gly Ile Gln Glu Ala Leu Glu His Glu Phe Ala Asp Gly Ser					
385		390		395	400
Phe Ser Gln Lys Gln Val Asp Glu Leu Val Gln Phe Arg Lys Ala Asn					
	405		410		415
Ser Ser Ile Phe Gly Lys Gly Trp His Asn Phe Ser Val Lys Leu Met					
	420		425		430
Met Glu Leu Ile Pro Glu Leu Tyr Glu Thr Ser Glu Glu Gln Met Thr					
	435		440		445
Ile Leu Thr Arg Leu Gly Lys Gln Lys Thr Thr Ser Ser Ser Asn Lys					
	450		455		460
Thr Lys Tyr Ile Asp Glu Lys Leu Leu Thr Glu Glu Ile Tyr Asn Pro					
465		470		475	480
Val Val Ala Lys Ser Val Arg Gln Ala Ile Lys Ile Val Asn Ala Ala					
	485		490		495
Ile Lys Glu Tyr Gly Asp Phe Asp Asn Ile Val Ile Glu Met Ala Arg					
	500		505		510
Glu Thr Asn Glu Asp Asp Glu Lys Lys Ala Ile Gln Lys Ile Gln Lys					
	515		520		525
Ala Asn Lys Asp Glu Lys Asp Ala Ala Met Leu Lys Ala Ala Asn Gln					
	530		535		540
Tyr Asn Gly Lys Ala Glu Leu Pro His Ser Val Phe His Gly His Lys					
545		550		555	560
Gln Leu Ala Thr Lys Ile Arg Leu Trp His Gln Gln Gly Glu Arg Cys					
	565		570		575
Leu Tyr Thr Gly Lys Thr Ile Ser Ile His Asp Leu Ile Asn Asn Ser					
	580		585		590
Asn Gln Phe Glu Val Asp His Ile Leu Pro Leu Ser Ile Thr Phe Asp					
	595		600		605
Asp Ser Leu Ala Asn Lys Val Leu Val Tyr Ala Thr Ala Asn Gln Glu					
	610		615		620
Lys Gly Gln Arg Thr Pro Tyr Gln Ala Leu Asp Ser Met Asp Asp Ala					
625		630		635	640

Trp Ser Phe Arg Glu Leu Lys Ala Phe Val Arg Glu Ser Lys Thr Leu		
645	650	655
Ser Asn Lys Lys Lys Glu Tyr Leu Leu Thr Glu Glu Asp Ile Ser Lys		
660	665	670
Phe Asp Val Arg Lys Lys Phe Ile Glu Arg Asn Leu Val Asp Thr Arg		
675	680	685
Tyr Ala Ser Arg Val Val Leu Asn Ala Leu Gln Glu His Phe Arg Ala		
690	695	700
His Lys Ile Asp Thr Lys Val Ser Val Val Arg Gly Gln Phe Thr Ser		
705	710	715
Gln Leu Arg Arg His Trp Gly Ile Glu Lys Thr Arg Asp Thr Tyr His		
725	730	735
His His Ala Val Asp Ala Leu Ile Ile Ala Ala Ser Ser Gln Leu Asn		
740	745	750
Leu Trp Lys Lys Gln Lys Asn Thr Leu Val Ser Tyr Ser Glu Asp Gln		
755	760	765
Leu Leu Asp Ile Glu Thr Gly Glu Leu Ile Ser Asp Asp Glu Tyr Lys		
770	775	780
Glu Ser Val Phe Lys Ala Pro Tyr Gln His Phe Val Asp Thr Leu Lys		
785	790	795
Ser Lys Glu Phe Glu Asp Ser Ile Leu Phe Ser Tyr Gln Val Asp Ser		
805	810	815
Lys Phe Asn Arg Lys Ile Ser Asp Ala Thr Ile Tyr Ala Thr Arg Gln		
820	825	830
Ala Lys Val Gly Lys Asp Lys Ala Asp Glu Thr Tyr Val Leu Gly Lys		
835	840	845
Ile Lys Asp Ile Tyr Thr Gln Asp Gly Tyr Asp Ala Phe Met Lys Ile		
850	855	860
Tyr Lys Lys Asp Lys Ser Lys Phe Leu Met Tyr Arg His Asp Pro Gln		
865	870	875
Thr Phe Glu Lys Val Ile Glu Pro Ile Leu Glu Asn Tyr Pro Asn Lys		
885	890	895
Gln Ile Asn Asp Lys Gly Lys Glu Val Pro Cys Asn Pro Phe Leu Lys		
900	905	910
Tyr Lys Glu Glu His Gly Tyr Ile Arg Lys Tyr Ser Lys Lys Gly Asn		
915	920	925
Gly Pro Glu Ile Lys Ser Leu Lys Tyr Tyr Asp Ser Lys Leu Gly Asn		
930	935	940
His Ile Asp Ile Thr Pro Lys Asp Ser Asn Asn Lys Val Val Leu Gln		

945	950	955	960
Ser Val Ser Pro Trp Arg Ala Asp Val Tyr Phe Asn Lys Thr Thr Gly			
	965	970	975
Lys Tyr Glu Ile Leu Gly Leu Lys Tyr Ala Asp Leu Gln Phe Asp Lys			
	980	985	990
Gly Thr Gly Thr Tyr Lys Ile Ser Gln Glu Lys Tyr Asn Asp Ile Lys			
	995	1000	1005
Lys Lys Glu Gly Val Asp Ser Asp Ser Glu Phe Lys Phe Thr Leu			
	1010	1015	1020
Tyr Lys Asn Asp Leu Leu Leu Val Lys Asp Thr Glu Thr Lys Glu			
	1025	1030	1035
Gln Gln Leu Phe Arg Phe Leu Ser Arg Thr Met Pro Lys Gln Lys			
	1040	1045	1050
His Tyr Val Glu Leu Lys Pro Tyr Asp Lys Gln Lys Phe Glu Gly			
	1055	1060	1065
Gly Glu Ala Leu Ile Lys Val Leu Gly Asn Val Ala Asn Ser Gly			
	1070	1075	1080
Gln Cys Lys Lys Gly Leu Gly Lys Ser Asn Ile Ser Ile Tyr Lys			
	1085	1090	1095
Val Arg Thr Asp Val Leu Gly Asn Gln His Ile Ile Lys Asn Glu			
	1100	1105	1110
Gly Asp Lys Pro Lys Leu Asp Phe			
	1115	1120	
<210> 152			
<211> 1082			
<212> PRT			
<213> 脑膜炎奈瑟球菌(<i>Neisseria meningitidis</i>)			
<400> 152			
Met Ala Ala Phe Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Asp			
1	5	10	15
Ile Gly Ile Ala Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Asp			
	20	25	30
Glu Asn Pro Ile Cys Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg			
	35	40	45
Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu			
	50	55	60
Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu			
65	70	75	80
Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp			

				85					90					95			
Phe	Asp	Glu	Asn	Gly	Leu	Ile	Lys	Ser	Leu	Pro	Asn	Thr	Pro	Trp	Gln		
				100					105					110			
Leu	Arg	Ala	Ala	Ala	Leu	Asp	Arg	Lys	Leu	Thr	Pro	Leu	Glu	Trp	Ser		
				115					120					125			
Ala	Val	Leu	Leu	His	Leu	Ile	Lys	His	Arg	Gly	Tyr	Leu	Ser	Gln	Arg		
				130					135					140			
Lys	Asn	Glu	Gly	Glu	Thr	Ala	Asp	Lys	Glu	Leu	Gly	Ala	Leu	Leu	Lys		
145						150					155				160		
Gly	Val	Ala	Asp	Asn	Ala	His	Ala	Leu	Gln	Thr	Gly	Asp	Phe	Arg	Thr		
				165					170					175			
Pro	Ala	Glu	Leu	Ala	Leu	Asn	Lys	Phe	Glu	Lys	Glu	Ser	Gly	His	Ile		
				180					185					190			
Arg	Asn	Gln	Arg	Gly	Asp	Tyr	Ser	His	Thr	Phe	Ser	Arg	Lys	Asp	Leu		
				195					200					205			
Gln	Ala	Glu	Leu	Ile	Leu	Leu	Phe	Glu	Lys	Gln	Lys	Glu	Phe	Gly	Asn		
				210					215					220			
Pro	His	Val	Ser	Gly	Gly	Leu	Lys	Glu	Gly	Ile	Glu	Thr	Leu	Leu	Met		
225						230				235					240		
Thr	Gln	Arg	Pro	Ala	Leu	Ser	Gly	Asp	Ala	Val	Gln	Lys	Met	Leu	Gly		
				245					250					255			
His	Cys	Thr	Phe	Glu	Pro	Ala	Glu	Pro	Lys	Ala	Ala	Lys	Asn	Thr	Tyr		
				260					265					270			
Thr	Ala	Glu	Arg	Phe	Ile	Trp	Leu	Thr	Lys	Leu	Asn	Asn	Leu	Arg	Ile		
				275					280					285			
Leu	Glu	Gln	Gly	Ser	Glu	Arg	Pro	Leu	Thr	Asp	Thr	Glu	Arg	Ala	Thr		
				290					295					300			
Leu	Met	Asp	Glu	Pro	Tyr	Arg	Lys	Ser	Lys	Leu	Thr	Tyr	Ala	Gln	Ala		
305						310				315					320		
Arg	Lys	Leu	Leu	Gly	Leu	Glu	Asp	Thr	Ala	Phe	Phe	Lys	Gly	Leu	Arg		
				325					330					335			
Tyr	Gly	Lys	Asp	Asn	Ala	Glu	Ala	Ser	Thr	Leu	Met	Glu	Met	Lys	Ala		
				340					345					350			
Tyr	His	Ala	Ile	Ser	Arg	Ala	Leu	Glu	Lys	Glu	Gly	Leu	Lys	Asp	Lys		
				355					360					365			
Lys	Ser	Pro	Leu	Asn	Leu	Ser	Pro	Glu	Leu	Gln	Asp	Glu	Ile	Gly	Thr		
				370					375					380			
Ala	Phe	Ser	Leu	Phe	Lys	Thr	Asp	Glu	Asp	Ile	Thr	Gly	Arg	Leu	Lys		
385						390				395					400		

Asp	Arg	Ile	Gln	Pro	Glu	Ile	Leu	Glu	Ala	Leu	Leu	Lys	His	Ile	Ser	
					405				410						415	
Phe	Asp	Lys	Phe	Val	Gln	Ile	Ser	Leu	Lys	Ala	Leu	Arg	Arg	Ile	Val	
					420				425						430	
Pro	Leu	Met	Glu	Gln	Gly	Lys	Arg	Tyr	Asp	Glu	Ala	Cys	Ala	Glu	Ile	
					435				440						445	
Tyr	Gly	Asp	His	Tyr	Gly	Lys	Lys	Asn	Thr	Glu	Glu	Lys	Ile	Tyr	Leu	
					450										460	
Pro	Pro	Ile	Pro	Ala	Asp	Glu	Ile	Arg	Asn	Pro	Val	Val	Leu	Arg	Ala	
465					470					475					480	
Leu	Ser	Gln	Ala	Arg	Lys	Val	Ile	Asn	Gly	Val	Val	Arg	Arg	Tyr	Gly	
					485				490						495	
Ser	Pro	Ala	Arg	Ile	His	Ile	Glu	Thr	Ala	Arg	Glu	Val	Gly	Lys	Ser	
					500				505						510	
Phe	Lys	Asp	Arg	Lys	Glu	Ile	Glu	Lys	Arg	Gln	Glu	Glu	Asn	Arg	Lys	
					515				520						525	
Asp	Arg	Glu	Lys	Ala	Ala	Ala	Lys	Phe	Arg	Glu	Tyr	Phe	Pro	Asn	Phe	
					530				535						540	
Val	Gly	Glu	Pro	Lys	Ser	Lys	Asp	Ile	Leu	Lys	Leu	Arg	Leu	Tyr	Glu	
545					550					555					560	
Gln	Gln	His	Gly	Lys	Cys	Leu	Tyr	Ser	Gly	Lys	Glu	Ile	Asn	Leu	Gly	
					565				570						575	
Arg	Leu	Asn	Glu	Lys	Gly	Tyr	Val	Glu	Ile	Asp	His	Ala	Leu	Pro	Phe	
					580				585						590	
Ser	Arg	Thr	Trp	Asp	Asp	Ser	Phe	Asn	Asn	Lys	Val	Leu	Val	Leu	Gly	
					595				600						605	
Ser	Glu	Asn	Gln	Asn	Lys	Gly	Asn	Gln	Thr	Pro	Tyr	Glu	Tyr	Phe	Asn	
					610				615						620	
Gly	Lys	Asp	Asn	Ser	Arg	Glu	Trp	Gln	Glu	Phe	Lys	Ala	Arg	Val	Glu	
625					630					635					640	
Thr	Ser	Arg	Phe	Pro	Arg	Ser	Lys	Lys	Gln	Arg	Ile	Leu	Leu	Gln	Lys	
					645				650						655	
Phe	Asp	Glu	Asp	Gly	Phe	Lys	Glu	Arg	Asn	Leu	Asn	Asp	Thr	Arg	Tyr	
					660				665						670	
Val	Asn	Arg	Phe	Leu	Cys	Gln	Phe	Val	Ala	Asp	Arg	Met	Arg	Leu	Thr	
					675				680						685	
Gly	Lys	Gly	Lys	Lys	Arg	Val	Phe	Ala	Ser	Asn	Gly	Gln	Ile	Thr	Asn	
					690				695						700	
Leu	Leu	Arg	Gly	Phe	Trp	Gly	Leu	Arg	Lys	Val	Arg	Ala	Glu	Asn	Asp	

705	710	715	720
Arg His His Ala Leu Asp Ala Val Val Val Ala Cys Ser Thr Val Ala			
	725	730	735
Met Gln Gln Lys Ile Thr Arg Phe Val Arg Tyr Lys Glu Met Asn Ala			
	740	745	750
Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr Gly Glu Val Leu His Gln			
	755	760	765
Lys Thr His Phe Pro Gln Pro Trp Glu Phe Phe Ala Gln Glu Val Met			
	770	775	780
Ile Arg Val Phe Gly Lys Pro Asp Gly Lys Pro Glu Phe Glu Glu Ala			
785	790	795	800
Asp Thr Pro Glu Lys Leu Arg Thr Leu Leu Ala Glu Lys Leu Ser Ser			
	805	810	815
Arg Pro Glu Ala Val His Glu Tyr Val Thr Pro Leu Phe Val Ser Arg			
	820	825	830
Ala Pro Asn Arg Lys Met Ser Gly Gln Gly His Met Glu Thr Val Lys			
	835	840	845
Ser Ala Lys Arg Leu Asp Glu Gly Val Ser Val Leu Arg Val Pro Leu			
	850	855	860
Thr Gln Leu Lys Leu Lys Asp Leu Glu Lys Met Val Asn Arg Glu Arg			
865	870	875	880
Glu Pro Lys Leu Tyr Glu Ala Leu Lys Ala Arg Leu Glu Ala His Lys			
	885	890	895
Asp Asp Pro Ala Lys Ala Phe Ala Glu Pro Phe Tyr Lys Tyr Asp Lys			
	900	905	910
Ala Gly Asn Arg Thr Gln Gln Val Lys Ala Val Arg Val Glu Gln Val			
	915	920	925
Gln Lys Thr Gly Val Trp Val Arg Asn His Asn Gly Ile Ala Asp Asn			
	930	935	940
Ala Thr Met Val Arg Val Asp Val Phe Glu Lys Gly Asp Lys Tyr Tyr			
945	950	955	960
Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp			
	965	970	975
Arg Ala Val Val Gln Gly Lys Asp Glu Glu Asp Trp Gln Leu Ile Asp			
	980	985	990
Asp Ser Phe Asn Phe Lys Phe Ser Leu His Pro Asn Asp Leu Val Glu			
	995	1000	1005
Val Ile Thr Lys Lys Ala Arg Met Phe Gly Tyr Phe Ala Ser Cys			
1010	1015	1020	

His Arg Gly Thr Gly Asn Ile Asn Ile Arg Ile His Asp Leu Asp
 1025 1030 1035
 His Lys Ile Gly Lys Asn Gly Ile Leu Glu Gly Ile Gly Val Lys
 1040 1045 1050
 Thr Ala Leu Ser Phe Gln Lys Tyr Gln Ile Asp Glu Leu Gly Lys
 1055 1060 1065
 Glu Ile Arg Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg
 1070 1075 1080
 <210> 153
 <211> 1037
 <212> PRT
 <213> 食清洁剂细小棒菌(Parvibaculum lavamentivorans)
 <400> 153
 Met Glu Arg Ile Phe Gly Phe Asp Ile Gly Thr Thr Ser Ile Gly Phe
 1 5 10 15
 Ser Val Ile Asp Tyr Ser Ser Thr Gln Ser Ala Gly Asn Ile Gln Arg
 20 25 30
 Leu Gly Val Arg Ile Phe Pro Glu Ala Arg Asp Pro Asp Gly Thr Pro
 35 40 45
 Leu Asn Gln Gln Arg Arg Gln Lys Arg Met Met Arg Arg Gln Leu Arg
 50 55 60
 Arg Arg Arg Ile Arg Arg Lys Ala Leu Asn Glu Thr Leu His Glu Ala
 65 70 75 80
 Gly Phe Leu Pro Ala Tyr Gly Ser Ala Asp Trp Pro Val Val Met Ala
 85 90 95
 Asp Glu Pro Tyr Glu Leu Arg Arg Arg Gly Leu Glu Glu Gly Leu Ser
 100 105 110
 Ala Tyr Glu Phe Gly Arg Ala Ile Tyr His Leu Ala Gln His Arg His
 115 120 125
 Phe Lys Gly Arg Glu Leu Glu Glu Ser Asp Thr Pro Asp Pro Asp Val
 130 135 140
 Asp Asp Glu Lys Glu Ala Ala Asn Glu Arg Ala Ala Thr Leu Lys Ala
 145 150 155 160
 Leu Lys Asn Glu Gln Thr Thr Leu Gly Ala Trp Leu Ala Arg Arg Pro
 165 170 175
 Pro Ser Asp Arg Lys Arg Gly Ile His Ala His Arg Asn Val Val Ala
 180 185 190
 Glu Glu Phe Glu Arg Leu Trp Glu Val Gln Ser Lys Phe His Pro Ala
 195 200 205

Leu	Lys	Ser	Glu	Glu	Met	Arg	Ala	Arg	Ile	Ser	Asp	Thr	Ile	Phe	Ala	210	215	220
Gln	Arg	Pro	Val	Phe	Trp	Arg	Lys	Asn	Thr	Leu	Gly	Glu	Cys	Arg	Phe	225	230	235
Met	Pro	Gly	Glu	Pro	Leu	Cys	Pro	Lys	Gly	Ser	Trp	Leu	Ser	Gln	Gln	245	250	255
Arg	Arg	Met	Leu	Glu	Lys	Leu	Asn	Asn	Leu	Ala	Ile	Ala	Gly	Gly	Asn	260	265	270
Ala	Arg	Pro	Leu	Asp	Ala	Glu	Glu	Arg	Asp	Ala	Ile	Leu	Ser	Lys	Leu	275	280	285
Gln	Gln	Gln	Ala	Ser	Met	Ser	Trp	Pro	Gly	Val	Arg	Ser	Ala	Leu	Lys	290	295	300
Ala	Leu	Tyr	Lys	Gln	Arg	Gly	Glu	Pro	Gly	Ala	Glu	Lys	Ser	Leu	Lys	305	310	315
Phe	Asn	Leu	Glu	Leu	Gly	Gly	Glu	Ser	Lys	Leu	Leu	Gly	Asn	Ala	Leu	325	330	335
Glu	Ala	Lys	Leu	Ala	Asp	Met	Phe	Gly	Pro	Asp	Trp	Pro	Ala	His	Pro	340	345	350
Arg	Lys	Gln	Glu	Ile	Arg	His	Ala	Val	His	Glu	Arg	Leu	Trp	Ala	Ala	355	360	365
Asp	Tyr	Gly	Glu	Thr	Pro	Asp	Lys	Lys	Arg	Val	Ile	Ile	Leu	Ser	Glu	370	375	380
Lys	Asp	Arg	Lys	Ala	His	Arg	Glu	Ala	Ala	Ala	Asn	Ser	Phe	Val	Ala	385	390	395
Asp	Phe	Gly	Ile	Thr	Gly	Glu	Gln	Ala	Ala	Gln	Leu	Gln	Ala	Leu	Lys	405	410	415
Leu	Pro	Thr	Gly	Trp	Glu	Pro	Tyr	Ser	Ile	Pro	Ala	Leu	Asn	Leu	Phe	420	425	430
Leu	Ala	Glu	Leu	Glu	Lys	Gly	Glu	Arg	Phe	Gly	Ala	Leu	Val	Asn	Gly	435	440	445
Pro	Asp	Trp	Glu	Gly	Trp	Arg	Arg	Thr	Asn	Phe	Pro	His	Arg	Asn	Gln	450	455	460
Pro	Thr	Gly	Glu	Ile	Leu	Asp	Lys	Leu	Pro	Ser	Pro	Ala	Ser	Lys	Glu	465	470	475
Glu	Arg	Glu	Arg	Ile	Ser	Gln	Leu	Arg	Asn	Pro	Thr	Val	Val	Arg	Thr	485	490	495
Gln	Asn	Glu	Leu	Arg	Lys	Val	Val	Asn	Asn	Leu	Ile	Gly	Leu	Tyr	Gly	500	505	510
Lys	Pro	Asp	Arg	Ile	Arg	Ile	Glu	Val	Gly	Arg	Asp	Val	Gly	Lys	Ser			

515	520	525
Lys Arg Glu Arg Glu Glu Ile Gln Ser Gly Ile Arg Arg Asn Glu Lys		
530	535	540
Gln Arg Lys Lys Ala Thr Glu Asp Leu Ile Lys Asn Gly Ile Ala Asn		
545	550	555
Pro Ser Arg Asp Asp Val Glu Lys Trp Ile Leu Trp Lys Glu Gly Gln		
565	570	575
Glu Arg Cys Pro Tyr Thr Gly Asp Gln Ile Gly Phe Asn Ala Leu Phe		
580	585	590
Arg Glu Gly Arg Tyr Glu Val Glu His Ile Trp Pro Arg Ser Arg Ser		
595	600	605
Phe Asp Asn Ser Pro Arg Asn Lys Thr Leu Cys Arg Lys Asp Val Asn		
610	615	620
Ile Glu Lys Gly Asn Arg Met Pro Phe Glu Ala Phe Gly His Asp Glu		
625	630	635
Asp Arg Trp Ser Ala Ile Gln Ile Arg Leu Gln Gly Met Val Ser Ala		
645	650	655
Lys Gly Gly Thr Gly Met Ser Pro Gly Lys Val Lys Arg Phe Leu Ala		
660	665	670
Lys Thr Met Pro Glu Asp Phe Ala Ala Arg Gln Leu Asn Asp Thr Arg		
675	680	685
Tyr Ala Ala Lys Gln Ile Leu Ala Gln Leu Lys Arg Leu Trp Pro Asp		
690	695	700
Met Gly Pro Glu Ala Pro Val Lys Val Glu Ala Val Thr Gly Gln Val		
705	710	715
Thr Ala Gln Leu Arg Lys Leu Trp Thr Leu Asn Asn Ile Leu Ala Asp		
725	730	735
Asp Gly Glu Lys Thr Arg Ala Asp His Arg His His Ala Ile Asp Ala		
740	745	750
Leu Thr Val Ala Cys Thr His Pro Gly Met Thr Asn Lys Leu Ser Arg		
755	760	765
Tyr Trp Gln Leu Arg Asp Asp Pro Arg Ala Glu Lys Pro Ala Leu Thr		
770	775	780
Pro Pro Trp Asp Thr Ile Arg Ala Asp Ala Glu Lys Ala Val Ser Glu		
785	790	795
Ile Val Val Ser His Arg Val Arg Lys Lys Val Ser Gly Pro Leu His		
805	810	815
Lys Glu Thr Thr Tyr Gly Asp Thr Gly Thr Asp Ile Lys Thr Lys Ser		
820	825	830

Gly Thr Tyr Arg Gln Phe Val Thr Arg Lys Lys Ile Glu Ser Leu Ser		
835	840	845
Lys Gly Glu Leu Asp Glu Ile Arg Asp Pro Arg Ile Lys Glu Ile Val		
850	855	860
Ala Ala His Val Ala Gly Arg Gly Gly Asp Pro Lys Lys Ala Phe Pro		
865	870	875
Pro Tyr Pro Cys Val Ser Pro Gly Gly Pro Glu Ile Arg Lys Val Arg		
	885	890
		895
Leu Thr Ser Lys Gln Gln Leu Asn Leu Met Ala Gln Thr Gly Asn Gly		
	900	905
		910
Tyr Ala Asp Leu Gly Ser Asn His His Ile Ala Ile Tyr Arg Leu Pro		
	915	920
		925
Asp Gly Lys Ala Asp Phe Glu Ile Val Ser Leu Phe Asp Ala Ser Arg		
	930	935
		940
Arg Leu Ala Gln Arg Asn Pro Ile Val Gln Arg Thr Arg Ala Asp Gly		
945	950	955
		960
Ala Ser Phe Val Met Ser Leu Ala Ala Gly Glu Ala Ile Met Ile Pro		
	965	970
		975
Glu Gly Ser Lys Lys Gly Ile Trp Ile Val Gln Gly Val Trp Ala Ser		
	980	985
		990
Gly Gln Val Val Leu Glu Arg Asp Thr Asp Ala Asp His Ser Thr Thr		
	995	1000
		1005
Thr Arg Pro Met Pro Asn Pro Ile Leu Lys Asp Asp Ala Lys Lys		
	1010	1015
		1020
Val Ser Ile Asp Pro Ile Gly Arg Val Arg Pro Ser Asn Asp		
	1025	1030
		1035
<210> 154		
<211> 1084		
<212> PRT		
<213> 白喉棒状杆菌(Corynebacter diphtheria)		
<400> 154		
Met Lys Tyr His Val Gly Ile Asp Val Gly Thr Phe Ser Val Gly Leu		
1	5	10
		15
Ala Ala Ile Glu Val Asp Asp Ala Gly Met Pro Ile Lys Thr Leu Ser		
	20	25
		30
Leu Val Ser His Ile His Asp Ser Gly Leu Asp Pro Asp Glu Ile Lys		
	35	40
		45
Ser Ala Val Thr Arg Leu Ala Ser Ser Gly Ile Ala Arg Arg Thr Arg		
	50	55
		60

Arg	Leu	Tyr	Arg	Arg	Lys	Arg	Arg	Arg	Leu	Gln	Gln	Leu	Asp	Lys	Phe
65					70				75						80
Ile	Gln	Arg	Gln	Gly	Trp	Pro	Val	Ile	Glu	Leu	Glu	Asp	Tyr	Ser	Asp
			85						90						95
Pro	Leu	Tyr	Pro	Trp	Lys	Val	Arg	Ala	Glu	Leu	Ala	Ala	Ser	Tyr	Ile
			100						105						110
Ala	Asp	Glu	Lys	Glu	Arg	Gly	Glu	Lys	Leu	Ser	Val	Ala	Leu	Arg	His
			115						120						125
Ile	Ala	Arg	His	Arg	Gly	Trp	Arg	Asn	Pro	Tyr	Ala	Lys	Val	Ser	Ser
			130						135						140
Leu	Tyr	Leu	Pro	Asp	Gly	Pro	Ser	Asp	Ala	Phe	Lys	Ala	Ile	Arg	Glu
145					150					155					160
Glu	Ile	Lys	Arg	Ala	Ser	Gly	Gln	Pro	Val	Pro	Glu	Thr	Ala	Thr	Val
			165						170						175
Gly	Gln	Met	Val	Thr	Leu	Cys	Glu	Leu	Gly	Thr	Leu	Lys	Leu	Arg	Gly
			180						185						190
Glu	Gly	Gly	Val	Leu	Ser	Ala	Arg	Leu	Gln	Gln	Ser	Asp	Tyr	Ala	Arg
			195						200						205
Glu	Ile	Gln	Glu	Ile	Cys	Arg	Met	Gln	Glu	Ile	Gly	Gln	Glu	Leu	Tyr
			210						215						220
Arg	Lys	Ile	Ile	Asp	Val	Val	Phe	Ala	Ala	Glu	Ser	Pro	Lys	Gly	Ser
225					230					235					240
Ala	Ser	Ser	Arg	Val	Gly	Lys	Asp	Pro	Leu	Gln	Pro	Gly	Lys	Asn	Arg
			245						250						255
Ala	Leu	Lys	Ala	Ser	Asp	Ala	Phe	Gln	Arg	Tyr	Arg	Ile	Ala	Ala	Leu
			260						265						270
Ile	Gly	Asn	Leu	Arg	Val	Arg	Val	Asp	Gly	Glu	Lys	Arg	Ile	Leu	Ser
			275						280						285
Val	Glu	Glu	Lys	Asn	Leu	Val	Phe	Asp	His	Leu	Val	Asn	Leu	Thr	Pro
			290						295						300
Lys	Lys	Glu	Pro	Glu	Trp	Val	Thr	Ile	Ala	Glu	Ile	Leu	Gly	Ile	Asp
305					310					315					320
Arg	Gly	Gln	Leu	Ile	Gly	Thr	Ala	Thr	Met	Thr	Asp	Asp	Gly	Glu	Arg
			325						330						335
Ala	Gly	Ala	Arg	Pro	Pro	Thr	His	Asp	Thr	Asn	Arg	Ser	Ile	Val	Asn
			340						345						350
Ser	Arg	Ile	Ala	Pro	Leu	Val	Asp	Trp	Trp	Lys	Thr	Ala	Ser	Ala	Leu
			355						360						365
Glu	Gln	His	Ala	Met	Val	Lys	Ala	Leu	Ser	Asn	Ala	Glu	Val	Asp	Asp

370	375	380
Phe Asp Ser Pro Glu Gly Ala Lys Val Gln Ala Phe Phe Ala Asp Leu		
385	390	395
Asp Asp Asp Val His Ala Lys Leu Asp Ser Leu His Leu Pro Val Gly		400
	405	410
Arg Ala Ala Tyr Ser Glu Asp Thr Leu Val Arg Leu Thr Arg Arg Met		415
	420	425
Leu Ser Asp Gly Val Asp Leu Tyr Thr Ala Arg Leu Gln Glu Phe Gly		430
	435	440
Ile Glu Pro Ser Trp Thr Pro Pro Thr Pro Arg Ile Gly Glu Pro Val		445
	450	455
Gly Asn Pro Ala Val Asp Arg Val Leu Lys Thr Val Ser Arg Trp Leu		460
465	470	475
Glu Ser Ala Thr Lys Thr Trp Gly Ala Pro Glu Arg Val Ile Ile Glu		480
	485	490
His Val Arg Glu Gly Phe Val Thr Glu Lys Arg Ala Arg Glu Met Asp		495
	500	505
Gly Asp Met Arg Arg Arg Ala Ala Arg Asn Ala Lys Leu Phe Gln Glu		510
	515	520
Met Gln Glu Lys Leu Asn Val Gln Gly Lys Pro Ser Arg Ala Asp Leu		525
	530	535
Trp Arg Tyr Gln Ser Val Gln Arg Gln Asn Cys Gln Cys Ala Tyr Cys		540
545	550	555
Gly Ser Pro Ile Thr Phe Ser Asn Ser Glu Met Asp His Ile Val Pro		560
	565	570
Arg Ala Gly Gln Gly Ser Thr Asn Thr Arg Glu Asn Leu Val Ala Val		575
	580	585
Cys His Arg Cys Asn Gln Ser Lys Gly Asn Thr Pro Phe Ala Ile Trp		590
	595	600
Ala Lys Asn Thr Ser Ile Glu Gly Val Ser Val Lys Glu Ala Val Glu		605
	610	615
Arg Thr Arg His Trp Val Thr Asp Thr Gly Met Arg Ser Thr Asp Phe		620
625	630	635
Lys Lys Phe Thr Lys Ala Val Val Glu Arg Phe Gln Arg Ala Thr Met		640
	645	650
Asp Glu Glu Ile Asp Ala Arg Ser Met Glu Ser Val Ala Trp Met Ala		655
	660	665
Asn Glu Leu Arg Ser Arg Val Ala Gln His Phe Ala Ser His Gly Thr		670
	675	680
		685

Thr Val Arg Val Tyr Arg Gly Ser Leu Thr Ala Glu Ala Arg Arg Ala			
690	695	700	
Ser Gly Ile Ser Gly Lys Leu Lys Phe Phe Asp Gly Val Gly Lys Ser			
705	710	715	720
Arg Leu Asp Arg Arg His His Ala Ile Asp Ala Ala Val Ile Ala Phe			
	725	730	735
Thr Ser Asp Tyr Val Ala Glu Thr Leu Ala Val Arg Ser Asn Leu Lys			
	740	745	750
Gln Ser Gln Ala His Arg Gln Glu Ala Pro Gln Trp Arg Glu Phe Thr			
	755	760	765
Gly Lys Asp Ala Glu His Arg Ala Ala Trp Arg Val Trp Cys Gln Lys			
	770	775	780
Met Glu Lys Leu Ser Ala Leu Leu Thr Glu Asp Leu Arg Asp Asp Arg			
785	790	795	800
Val Val Val Met Ser Asn Val Arg Leu Arg Leu Gly Asn Gly Ser Ala			
	805	810	815
His Lys Glu Thr Ile Gly Lys Leu Ser Lys Val Lys Leu Ser Ser Gln			
	820	825	830
Leu Ser Val Ser Asp Ile Asp Lys Ala Ser Ser Glu Ala Leu Trp Cys			
	835	840	845
Ala Leu Thr Arg Glu Pro Gly Phe Asp Pro Lys Glu Gly Leu Pro Ala			
	850	855	860
Asn Pro Glu Arg His Ile Arg Val Asn Gly Thr His Val Tyr Ala Gly			
865	870	875	880
Asp Asn Ile Gly Leu Phe Pro Val Ser Ala Gly Ser Ile Ala Leu Arg			
	885	890	895
Gly Gly Tyr Ala Glu Leu Gly Ser Ser Phe His His Ala Arg Val Tyr			
	900	905	910
Lys Ile Thr Ser Gly Lys Lys Pro Ala Phe Ala Met Leu Arg Val Tyr			
	915	920	925
Thr Ile Asp Leu Leu Pro Tyr Arg Asn Gln Asp Leu Phe Ser Val Glu			
	930	935	940
Leu Lys Pro Gln Thr Met Ser Met Arg Gln Ala Glu Lys Lys Leu Arg			
945	950	955	960
Asp Ala Leu Ala Thr Gly Asn Ala Glu Tyr Leu Gly Trp Leu Val Val			
	965	970	975
Asp Asp Glu Leu Val Val Asp Thr Ser Lys Ile Ala Thr Asp Gln Val			
	980	985	990
Lys Ala Val Glu Ala Glu Leu Gly Thr Ile Arg Arg Trp Arg Val Asp			

995	1000	1005
Gly Phe Phe Ser Pro Ser Lys Leu Arg Leu Arg Pro Leu Gln Met		
1010	1015	1020
Ser Lys Glu Gly Ile Lys Lys Glu Ser Ala Pro Glu Leu Ser Lys		
1025	1030	1035
Ile Ile Asp Arg Pro Gly Trp Leu Pro Ala Val Asn Lys Leu Phe		
1040	1045	1050
Ser Asp Gly Asn Val Thr Val Val Arg Arg Asp Ser Leu Gly Arg		
1055	1060	1065
Val Arg Leu Glu Ser Thr Ala His Leu Pro Val Thr Trp Lys Val		
1070	1075	1080
Gln		
<210> 155		
<211> 1130		
<212> PRT		
<213> 巴氏链球菌(Streptococcus pasteurianus)		
<400> 155		
Met Thr Asn Gly Lys Ile Leu Gly Leu Asp Ile Gly Ile Ala Ser Val		
1	5	10
Gly Val Gly Ile Ile Glu Ala Lys Thr Gly Lys Val Val His Ala Asn		
20	25	30
Ser Arg Leu Phe Ser Ala Ala Asn Ala Glu Asn Asn Ala Glu Arg Arg		
35	40	45
Gly Phe Arg Gly Ser Arg Arg Leu Asn Arg Arg Lys Lys His Arg Val		
50	55	60
Lys Arg Val Arg Asp Leu Phe Glu Lys Tyr Gly Ile Val Thr Asp Phe		
65	70	75
Arg Asn Leu Asn Leu Asn Pro Tyr Glu Leu Arg Val Lys Gly Leu Thr		
85	90	95
Glu Gln Leu Lys Asn Glu Glu Leu Phe Ala Ala Leu Arg Thr Ile Ser		
100	105	110
Lys Arg Arg Gly Ile Ser Tyr Leu Asp Asp Ala Glu Asp Asp Ser Thr		
115	120	125
Gly Ser Thr Asp Tyr Ala Lys Ser Ile Asp Glu Asn Arg Arg Leu Leu		
130	135	140
Lys Asn Lys Thr Pro Gly Gln Ile Gln Leu Glu Arg Leu Glu Lys Tyr		
145	150	155
Gly Gln Leu Arg Gly Asn Phe Thr Val Tyr Asp Glu Asn Gly Glu Ala		
165	170	175

His Arg Leu Ile Asn Val Phe Ser Thr Ser Asp Tyr Glu Lys Glu Ala		
180	185	190
Arg Lys Ile Leu Glu Thr Gln Ala Asp Tyr Asn Lys Lys Ile Thr Ala		
195	200	205
Glu Phe Ile Asp Asp Tyr Val Glu Ile Leu Thr Gln Lys Arg Lys Tyr		
210	215	220
Tyr His Gly Pro Gly Asn Glu Lys Ser Arg Thr Asp Tyr Gly Arg Phe		
225	230	235
Arg Thr Asp Gly Thr Thr Leu Glu Asn Ile Phe Gly Ile Leu Ile Gly		
245	250	255
Lys Cys Asn Phe Tyr Pro Asp Glu Tyr Arg Ala Ser Lys Ala Ser Tyr		
260	265	270
Thr Ala Gln Glu Tyr Asn Phe Leu Asn Asp Leu Asn Asn Leu Lys Val		
275	280	285
Ser Thr Glu Thr Gly Lys Leu Ser Thr Glu Gln Lys Glu Ser Leu Val		
290	295	300
Glu Phe Ala Lys Asn Thr Ala Thr Leu Gly Pro Ala Lys Leu Leu Lys		
305	310	315
Glu Ile Ala Lys Ile Leu Asp Cys Lys Val Asp Glu Ile Lys Gly Tyr		
325	330	335
Arg Glu Asp Asp Lys Gly Lys Pro Asp Leu His Thr Phe Glu Pro Tyr		
340	345	350
Arg Lys Leu Lys Phe Asn Leu Glu Ser Ile Asn Ile Asp Asp Leu Ser		
355	360	365
Arg Glu Val Ile Asp Lys Leu Ala Asp Ile Leu Thr Leu Asn Thr Glu		
370	375	380
Arg Glu Gly Ile Glu Asp Ala Ile Lys Arg Asn Leu Pro Asn Gln Phe		
385	390	395
Thr Glu Glu Gln Ile Ser Glu Ile Ile Lys Val Arg Lys Ser Gln Ser		
405	410	415
Thr Ala Phe Asn Lys Gly Trp His Ser Phe Ser Ala Lys Leu Met Asn		
420	425	430
Glu Leu Ile Pro Glu Leu Tyr Ala Thr Ser Asp Glu Gln Met Thr Ile		
435	440	445
Leu Thr Arg Leu Glu Lys Phe Lys Val Asn Lys Lys Ser Ser Lys Asn		
450	455	460
Thr Lys Thr Ile Asp Glu Lys Glu Val Thr Asp Glu Ile Tyr Asn Pro		
465	470	475
Val Val Ala Lys Ser Val Arg Gln Thr Ile Lys Ile Ile Asn Ala Ala		

				485				490				495			
Val	Lys	Lys	Tyr	Gly	Asp	Phe	Asp	Lys	Ile	Val	Ile	Glu	Met	Pro	Arg
				500				505				510			
Asp	Lys	Asn	Ala	Asp	Asp	Glu	Lys	Lys	Phe	Ile	Asp	Lys	Arg	Asn	Lys
				515				520				525			
Glu	Asn	Lys	Lys	Glu	Lys	Asp	Asp	Ala	Leu	Lys	Arg	Ala	Ala	Tyr	Leu
				530				535				540			
Tyr	Asn	Ser	Ser	Asp	Lys	Leu	Pro	Asp	Glu	Val	Phe	His	Gly	Asn	Lys
545				550				555				560			
Gln	Leu	Glu	Thr	Lys	Ile	Arg	Leu	Trp	Tyr	Gln	Gln	Gly	Glu	Arg	Cys
				565				570				575			
Leu	Tyr	Ser	Gly	Lys	Pro	Ile	Ser	Ile	Gln	Glu	Leu	Val	His	Asn	Ser
				580				585				590			
Asn	Asn	Phe	Glu	Ile	Asp	His	Ile	Leu	Pro	Leu	Ser	Leu	Ser	Phe	Asp
				595				600				605			
Asp	Ser	Leu	Ala	Asn	Lys	Val	Leu	Val	Tyr	Ala	Trp	Thr	Asn	Gln	Glu
610				615				620							
Lys	Gly	Gln	Lys	Thr	Pro	Tyr	Gln	Val	Ile	Asp	Ser	Met	Asp	Ala	Ala
625				630				635				640			
Trp	Ser	Phe	Arg	Glu	Met	Lys	Asp	Tyr	Val	Leu	Lys	Gln	Lys	Gly	Leu
				645				650				655			
Gly	Lys	Lys	Lys	Arg	Asp	Tyr	Leu	Leu	Thr	Thr	Glu	Asn	Ile	Asp	Lys
				660				665				670			
Ile	Glu	Val	Lys	Lys	Lys	Phe	Ile	Glu	Arg	Asn	Leu	Val	Asp	Thr	Arg
675				680				685							
Tyr	Ala	Ser	Arg	Val	Val	Leu	Asn	Ser	Leu	Gln	Ser	Ala	Leu	Arg	Glu
690				695				700							
Leu	Gly	Lys	Asp	Thr	Lys	Val	Ser	Val	Val	Arg	Gly	Gln	Phe	Thr	Ser
705				710				715				720			
Gln	Leu	Arg	Arg	Lys	Trp	Lys	Ile	Asp	Lys	Ser	Arg	Glu	Thr	Tyr	His
				725				730				735			
His	His	Ala	Val	Asp	Ala	Leu	Ile	Ile	Ala	Ala	Ser	Ser	Gln	Leu	Lys
740				745				750							
Leu	Trp	Glu	Lys	Gln	Asp	Asn	Pro	Met	Phe	Val	Asp	Tyr	Gly	Lys	Asn
755				760				765							
Gln	Val	Val	Asp	Lys	Gln	Thr	Gly	Glu	Ile	Leu	Ser	Val	Ser	Asp	Asp
770				775				780							
Glu	Tyr	Lys	Glu	Leu	Val	Phe	Gln	Pro	Pro	Tyr	Gln	Gly	Phe	Val	Asn
785				790				795				800			

Thr	Ile	Ser	Ser	Lys	Gly	Phe	Glu	Asp	Glu	Ile	Leu	Phe	Ser	Tyr	Gln	
				805					810						815	
Val	Asp	Ser	Lys	Tyr	Asn	Arg	Lys	Val	Ser	Asp	Ala	Thr	Ile	Tyr	Ser	
			820					825						830		
Thr	Arg	Lys	Ala	Lys	Ile	Gly	Lys	Asp	Lys	Lys	Glu	Glu	Thr	Tyr	Val	
			835					840						845		
Leu	Gly	Lys	Ile	Lys	Asp	Ile	Tyr	Ser	Gln	Asn	Gly	Phe	Asp	Thr	Phe	
			850					855						860		
Ile	Lys	Lys	Tyr	Asn	Lys	Asp	Lys	Thr	Gln	Phe	Leu	Met	Tyr	Gln	Lys	
865					870					875					880	
Asp	Ser	Leu	Thr	Trp	Glu	Asn	Val	Ile	Glu	Val	Ile	Leu	Arg	Asp	Tyr	
				885					890						895	
Pro	Thr	Thr	Lys	Lys	Ser	Glu	Asp	Gly	Lys	Asn	Asp	Val	Lys	Cys	Asn	
			900						905						910	
Pro	Phe	Glu	Glu	Tyr	Arg	Arg	Glu	Asn	Gly	Leu	Ile	Cys	Lys	Tyr	Ser	
			915						920						925	
Lys	Lys	Gly	Lys	Gly	Thr	Pro	Ile	Lys	Ser	Leu	Lys	Tyr	Tyr	Asp	Lys	
			930						935						940	
Lys	Leu	Gly	Asn	Cys	Ile	Asp	Ile	Thr	Pro	Glu	Glu	Ser	Arg	Asn	Lys	
945					950						955				960	
Val	Ile	Leu	Gln	Ser	Ile	Asn	Pro	Trp	Arg	Ala	Asp	Val	Tyr	Phe	Asn	
				965						970					975	
Pro	Glu	Thr	Leu	Lys	Tyr	Glu	Leu	Met	Gly	Leu	Lys	Tyr	Ser	Asp	Leu	
			980						985						990	
Ser	Phe	Glu	Lys	Gly	Thr	Gly	Asn	Tyr	His	Ile	Ser	Gln	Glu	Lys	Tyr	
			995						1000						1005	
Asp	Ala	Ile	Lys	Glu	Lys	Glu	Gly	Ile	Gly	Lys	Lys	Ser	Glu	Phe		
			1010						1015					1020		
Lys	Phe	Thr	Leu	Tyr	Arg	Asn	Asp	Leu	Ile	Leu	Ile	Lys	Asp	Ile		
			1025						1030					1035		
Ala	Ser	Gly	Glu	Gln	Glu	Ile	Tyr	Arg	Phe	Leu	Ser	Arg	Thr	Met		
			1040						1045					1050		
Pro	Asn	Val	Asn	His	Tyr	Val	Glu	Leu	Lys	Pro	Tyr	Asp	Lys	Glu		
			1055						1060					1065		
Lys	Phe	Asp	Asn	Val	Gln	Glu	Leu	Val	Glu	Ala	Leu	Gly	Glu	Ala		
			1070						1075					1080		
Asp	Lys	Val	Gly	Arg	Cys	Ile	Lys	Gly	Leu	Asn	Lys	Pro	Asn	Ile		
			1085						1090					1095		
Ser	Ile	Tyr	Lys	Val	Arg	Thr	Asp	Val	Leu	Gly	Asn	Lys	Tyr	Phe		

1100	1105	1110
Val Lys Lys Lys Gly Asp	Lys Pro Lys Leu Asp	Phe Lys Asn Asn
1115	1120	1125
Lys Lys		
1130		
<210> 156		
<211> 1082		
<212> PRT		
<213> 灰色奈瑟球菌(<i>Neisseria cinerea</i>)		
<400> 156		
Met Ala Ala Phe Lys Pro Asn Pro Met Asn Tyr Ile Leu Gly Leu Asp		
1	5	10
Ile Gly Ile Ala Ser Val Gly Trp Ala Ile Val Glu Ile Asp Glu Glu		15
	20	25
Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg		30
	35	40
Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Ala Ala Arg Arg Leu		45
	50	55
Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu		60
65	70	75
Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp		80
	85	90
Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln		95
	100	105
Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser		110
	115	120
Ala Val Leu Leu His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg		125
	130	135
Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys		140
145	150	155
Gly Val Ala Asp Asn Thr His Ala Leu Gln Thr Gly Asp Phe Arg Thr		160
	165	170
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile		175
	180	185
Arg Asn Gln Arg Gly Asp Tyr Ser His Thr Phe Asn Arg Lys Asp Leu		190
	195	200
Gln Ala Glu Leu Asn Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn		205
	210	215
Pro His Val Ser Asp Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met		220

225	230	235	240
Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly			
	245	250	255
His Cys Thr Phe Glu Pro Thr Glu Pro Lys Ala Ala Lys Asn Thr Tyr			
	260	265	270
Thr Ala Glu Arg Phe Val Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile			
	275	280	285
Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr			
	290	295	300
Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala			
305	310	315	320
Arg Lys Leu Leu Asp Leu Asp Asp Thr Ala Phe Phe Lys Gly Leu Arg			
	325	330	335
Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala			
	340	345	350
Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys			
	355	360	365
Lys Ser Pro Leu Asn Leu Ser Pro Glu Leu Gln Asp Glu Ile Gly Thr			
	370	375	380
Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys			
385	390	395	400
Asp Arg Val Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser			
	405	410	415
Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val			
	420	425	430
Pro Leu Met Glu Gln Gly Asn Arg Tyr Asp Glu Ala Cys Thr Glu Ile			
	435	440	445
Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu			
	450	455	460
Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala			
465	470	475	480
Leu Ser Gln Ala Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly			
	485	490	495
Ser Pro Ala Arg Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser			
	500	505	510
Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys			
	515	520	525
Asp Arg Glu Lys Ser Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe			
	530	535	540

Val Gly Glu Pro Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu			
545	550	555	560
Gln Gln His Gly Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Gly			
	565	570	575
Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe			
	580	585	590
Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn Lys Val Leu Ala Leu Gly			
	595	600	605
Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn			
	610	615	620
Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu			
625	630	635	640
Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys			
	645	650	655
Phe Asp Glu Asp Gly Phe Lys Glu Arg Asn Leu Asn Asp Thr Arg Tyr			
	660	665	670
Ile Asn Arg Phe Leu Cys Gln Phe Val Ala Asp His Met Leu Leu Thr			
	675	680	685
Gly Lys Gly Lys Arg Arg Val Phe Ala Ser Asn Gly Gln Ile Thr Asn			
	690	695	700
Leu Leu Arg Gly Phe Trp Gly Leu Arg Lys Val Arg Ala Glu Asn Asp			
705	710	715	720
Arg His His Ala Leu Asp Ala Val Val Val Ala Cys Ser Thr Ile Ala			
	725	730	735
Met Gln Gln Lys Ile Thr Arg Phe Val Arg Tyr Lys Glu Met Asn Ala			
	740	745	750
Phe Asp Gly Lys Thr Ile Asp Lys Glu Thr Gly Glu Val Leu His Gln			
	755	760	765
Lys Ala His Phe Pro Gln Pro Trp Glu Phe Phe Ala Gln Glu Val Met			
	770	775	780
Ile Arg Val Phe Gly Lys Pro Asp Gly Lys Pro Glu Phe Glu Glu Ala			
785	790	795	800
Asp Thr Pro Glu Lys Leu Arg Thr Leu Leu Ala Glu Lys Leu Ser Ser			
	805	810	815
Arg Pro Glu Ala Val His Lys Tyr Val Thr Pro Leu Phe Ile Ser Arg			
	820	825	830
Ala Pro Asn Arg Lys Met Ser Gly Gln Gly His Met Glu Thr Val Lys			
	835	840	845
Ser Ala Lys Arg Leu Asp Glu Gly Ile Ser Val Leu Arg Val Pro Leu			

850	855	860
Thr Gln Leu Lys Leu Lys Asp Leu Glu Lys Met Val Asn Arg Glu Arg		
865	870	875
Glu Pro Lys Leu Tyr Glu Ala Leu Lys Ala Arg Leu Glu Ala His Lys		880
	885	890
Asp Asp Pro Ala Lys Ala Phe Ala Glu Pro Phe Tyr Lys Tyr Asp Lys		895
	900	905
Ala Gly Asn Arg Thr Gln Gln Val Lys Ala Val Arg Val Glu Gln Val		910
	915	920
Gln Lys Thr Gly Val Trp Val His Asn His Asn Gly Ile Ala Asp Asn		925
	930	935
Ala Thr Ile Val Arg Val Asp Val Phe Glu Lys Gly Gly Lys Tyr Tyr		940
945	950	955
Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp		960
	965	970
Arg Ala Val Val Gln Gly Lys Asp Glu Glu Asp Trp Thr Val Met Asp		975
	980	985
Asp Ser Phe Glu Phe Lys Phe Val Leu Tyr Ala Asn Asp Leu Ile Lys		990
	995	1000
Leu Thr Ala Lys Lys Asn Glu Phe Leu Gly Tyr Phe Val Ser Leu		1005
	1010	1015
Asn Arg Ala Thr Gly Ala Ile Asp Ile Arg Thr His Asp Thr Asp		1020
	1025	1030
Ser Thr Lys Gly Lys Asn Gly Ile Phe Gln Ser Val Gly Val Lys		1035
	1040	1045
Thr Ala Leu Ser Phe Gln Lys Tyr Gln Ile Asp Glu Leu Gly Lys		1050
	1055	1060
Glu Ile Arg Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg		1065
	1070	1075
		1080
<210> 157		
<211> 1003		
<212> PRT		
<213> 红嘴鸥弯曲杆菌 (Campylobacter lari)		
<400> 157		
Met Arg Ile Leu Gly Phe Asp Ile Gly Ile Asn Ser Ile Gly Trp Ala		
1	5	10
Phe Val Glu Asn Asp Glu Leu Lys Asp Cys Gly Val Arg Ile Phe Thr		15
	20	25
Lys Ala Glu Asn Pro Lys Asn Lys Glu Ser Leu Ala Leu Pro Arg Arg		30

35					40					45						
Asn	Ala	Arg	Ser	Ser	Arg	Arg	Arg	Leu	Lys	Arg	Arg	Lys	Ala	Arg	Leu	
50					55					60						
Ile	Ala	Ile	Lys	Arg	Ile	Leu	Ala	Lys	Glu	Leu	Lys	Leu	Asn	Tyr	Lys	
65					70					75					80	
Asp	Tyr	Val	Ala	Ala	Asp	Gly	Glu	Leu	Pro	Lys	Ala	Tyr	Glu	Gly	Ser	
85					90					95						
Leu	Ala	Ser	Val	Tyr	Glu	Leu	Arg	Tyr	Lys	Ala	Leu	Thr	Gln	Asn	Leu	
100					105					110						
Glu	Thr	Lys	Asp	Leu	Ala	Arg	Val	Ile	Leu	His	Ile	Ala	Lys	His	Arg	
115					120					125						
Gly	Tyr	Met	Asn	Lys	Asn	Glu	Lys	Lys	Ser	Asn	Asp	Ala	Lys	Lys	Gly	
130					135					140						
Lys	Ile	Leu	Ser	Ala	Leu	Lys	Asn	Asn	Ala	Leu	Lys	Leu	Glu	Asn	Tyr	
145					150					155					160	
Gln	Ser	Val	Gly	Glu	Tyr	Phe	Tyr	Lys	Glu	Phe	Phe	Gln	Lys	Tyr	Lys	
165					170					175						
Lys	Asn	Thr	Lys	Asn	Phe	Ile	Lys	Ile	Arg	Asn	Thr	Lys	Asp	Asn	Tyr	
180					185					190						
Asn	Asn	Cys	Val	Leu	Ser	Ser	Asp	Leu	Glu	Lys	Glu	Leu	Lys	Leu	Ile	
195					200					205						
Leu	Glu	Lys	Gln	Lys	Glu	Phe	Gly	Tyr	Asn	Tyr	Ser	Glu	Asp	Phe	Ile	
210					215					220						
Asn	Glu	Ile	Leu	Lys	Val	Ala	Phe	Phe	Gln	Arg	Pro	Leu	Lys	Asp	Phe	
225					230					235					240	
Ser	His	Leu	Val	Gly	Ala	Cys	Thr	Phe	Phe	Glu	Glu	Glu	Lys	Arg	Ala	
245					250					255						
Cys	Lys	Asn	Ser	Tyr	Ser	Ala	Trp	Glu	Phe	Val	Ala	Leu	Thr	Lys	Ile	
260					265					270						
Ile	Asn	Glu	Ile	Lys	Ser	Leu	Glu	Lys	Ile	Ser	Gly	Glu	Ile	Val	Pro	
275					280					285						
Thr	Gln	Thr	Ile	Asn	Glu	Val	Leu	Asn	Leu	Ile	Leu	Asp	Lys	Gly	Ser	
290					295					300						
Ile	Thr	Tyr	Lys	Lys	Phe	Arg	Ser	Cys	Ile	Asn	Leu	His	Glu	Ser	Ile	
305					310					315					320	
Ser	Phe	Lys	Ser	Leu	Lys	Tyr	Asp	Lys	Glu	Asn	Ala	Glu	Asn	Ala	Lys	
325					330					335						
Leu	Ile	Asp	Phe	Arg	Lys	Leu	Val	Glu	Phe	Lys	Lys	Ala	Leu	Gly	Val	
340					345					350						

His Ser Leu Ser Arg Gln Glu Leu Asp Gln Ile Ser Thr His Ile Thr		
355	360	365
Leu Ile Lys Asp Asn Val Lys Leu Lys Thr Val Leu Glu Lys Tyr Asn		
370	375	380
Leu Ser Asn Glu Gln Ile Asn Asn Leu Leu Glu Ile Glu Phe Asn Asp		
385	390	395
400		
Tyr Ile Asn Leu Ser Phe Lys Ala Leu Gly Met Ile Leu Pro Leu Met		
405	410	415
Arg Glu Gly Lys Arg Tyr Asp Glu Ala Cys Glu Ile Ala Asn Leu Lys		
420	425	430
Pro Lys Thr Val Asp Glu Lys Lys Asp Phe Leu Pro Ala Phe Cys Asp		
435	440	445
Ser Ile Phe Ala His Glu Leu Ser Asn Pro Val Val Asn Arg Ala Ile		
450	455	460
Ser Glu Tyr Arg Lys Val Leu Asn Ala Leu Leu Lys Lys Tyr Gly Lys		
465	470	475
480		
Val His Lys Ile His Leu Glu Leu Ala Arg Asp Val Gly Leu Ser Lys		
485	490	495
Lys Ala Arg Glu Lys Ile Glu Lys Glu Gln Lys Glu Asn Gln Ala Val		
500	505	510
Asn Ala Trp Ala Leu Lys Glu Cys Glu Asn Ile Gly Leu Lys Ala Ser		
515	520	525
Ala Lys Asn Ile Leu Lys Leu Lys Leu Trp Lys Glu Gln Lys Glu Ile		
530	535	540
Cys Ile Tyr Ser Gly Asn Lys Ile Ser Ile Glu His Leu Lys Asp Glu		
545	550	555
560		
Lys Ala Leu Glu Val Asp His Ile Tyr Pro Tyr Ser Arg Ser Phe Asp		
565	570	575
Asp Ser Phe Ile Asn Lys Val Leu Val Phe Thr Lys Glu Asn Gln Glu		
580	585	590
Lys Leu Asn Lys Thr Pro Phe Glu Ala Phe Gly Lys Asn Ile Glu Lys		
595	600	605
Trp Ser Lys Ile Gln Thr Leu Ala Gln Asn Leu Pro Tyr Lys Lys Lys		
610	615	620
Asn Lys Ile Leu Asp Glu Asn Phe Lys Asp Lys Gln Gln Glu Asp Phe		
625	630	635
640		
Ile Ser Arg Asn Leu Asn Asp Thr Arg Tyr Ile Ala Thr Leu Ile Ala		
645	650	655
Lys Tyr Thr Lys Glu Tyr Leu Asn Phe Leu Leu Leu Ser Glu Asn Glu		

660	665	670
Asn Ala Asn Leu Lys Ser Gly Glu Lys Gly Ser Lys Ile His Val Gln		
675	680	685
Thr Ile Ser Gly Met Leu Thr Ser Val Leu Arg His Thr Trp Gly Phe		
690	695	700
Asp Lys Lys Asp Arg Asn Asn His Leu His His Ala Leu Asp Ala Ile		
705	710	715
Ile Val Ala Tyr Ser Thr Asn Ser Ile Ile Lys Ala Phe Ser Asp Phe		
725	730	735
Arg Lys Asn Gln Glu Leu Leu Lys Ala Arg Phe Tyr Ala Lys Glu Leu		
740	745	750
Thr Ser Asp Asn Tyr Lys His Gln Val Lys Phe Phe Glu Pro Phe Lys		
755	760	765
Ser Phe Arg Glu Lys Ile Leu Ser Lys Ile Asp Glu Ile Phe Val Ser		
770	775	780
Lys Pro Pro Arg Lys Arg Ala Arg Arg Ala Leu His Lys Asp Thr Phe		
785	790	795
His Ser Glu Asn Lys Ile Ile Asp Lys Cys Ser Tyr Asn Ser Lys Glu		
805	810	815
Gly Leu Gln Ile Ala Leu Ser Cys Gly Arg Val Arg Lys Ile Gly Thr		
820	825	830
Lys Tyr Val Glu Asn Asp Thr Ile Val Arg Val Asp Ile Phe Lys Lys		
835	840	845
Gln Asn Lys Phe Tyr Ala Ile Pro Ile Tyr Ala Met Asp Phe Ala Leu		
850	855	860
Gly Ile Leu Pro Asn Lys Ile Val Ile Thr Gly Lys Asp Lys Asn Asn		
865	870	875
Asn Pro Lys Gln Trp Gln Thr Ile Asp Glu Ser Tyr Glu Phe Cys Phe		
885	890	895
Ser Leu Tyr Lys Asn Asp Leu Ile Leu Leu Gln Lys Lys Asn Met Gln		
900	905	910
Glu Pro Glu Phe Ala Tyr Tyr Asn Asp Phe Ser Ile Ser Thr Ser Ser		
915	920	925
Ile Cys Val Glu Lys His Asp Asn Lys Phe Glu Asn Leu Thr Ser Asn		
930	935	940
Gln Lys Leu Leu Phe Ser Asn Ala Lys Glu Gly Ser Val Lys Val Glu		
945	950	955
Ser Leu Gly Ile Gln Asn Leu Lys Val Phe Glu Lys Tyr Ile Ile Thr		
965	970	975

Pro Leu Gly Asp Lys Ile Lys Ala Asp Phe Gln Pro Arg Glu Asn Ile
 980 985 990
 Ser Leu Lys Thr Ser Lys Lys Tyr Gly Leu Arg
 995 1000
 <210> 158
 <211> 1395
 <212> PRT
 <213> 齿垢密螺旋体(Treponema denticola)
 <400> 158
 Met Lys Lys Glu Ile Lys Asp Tyr Phe Leu Gly Leu Asp Val Gly Thr
 1 5 10 15
 Gly Ser Val Gly Trp Ala Val Thr Asp Thr Asp Tyr Lys Leu Leu Lys
 20 25 30
 Ala Asn Arg Lys Asp Leu Trp Gly Met Arg Cys Phe Glu Thr Ala Glu
 35 40 45
 Thr Ala Glu Val Arg Arg Leu His Arg Gly Ala Arg Arg Arg Ile Glu
 50 55 60
 Arg Arg Lys Lys Arg Ile Lys Leu Leu Gln Glu Leu Phe Ser Gln Glu
 65 70 75 80
 Ile Ala Lys Thr Asp Glu Gly Phe Phe Gln Arg Met Lys Glu Ser Pro
 85 90 95
 Phe Tyr Ala Glu Asp Lys Thr Ile Leu Gln Glu Asn Thr Leu Phe Asn
 100 105 110
 Asp Lys Asp Phe Ala Asp Lys Thr Tyr His Lys Ala Tyr Pro Thr Ile
 115 120 125
 Asn His Leu Ile Lys Ala Trp Ile Glu Asn Lys Val Lys Pro Asp Pro
 130 135 140
 Arg Leu Leu Tyr Leu Ala Cys His Asn Ile Ile Lys Lys Arg Gly His
 145 150 155 160
 Phe Leu Phe Glu Gly Asp Phe Asp Ser Glu Asn Gln Phe Asp Thr Ser
 165 170 175
 Ile Gln Ala Leu Phe Glu Tyr Leu Arg Glu Asp Met Glu Val Asp Ile
 180 185 190
 Asp Ala Asp Ser Gln Lys Val Lys Glu Ile Leu Lys Asp Ser Ser Leu
 195 200 205
 Lys Asn Ser Glu Lys Gln Ser Arg Leu Asn Lys Ile Leu Gly Leu Lys
 210 215 220
 Pro Ser Asp Lys Gln Lys Lys Ala Ile Thr Asn Leu Ile Ser Gly Asn
 225 230 235 240

Lys Ile Asn Phe Ala Asp Leu Tyr Asp Asn Pro Asp Leu Lys Asp Ala		
245	250	255
Glu Lys Asn Ser Ile Ser Phe Ser Lys Asp Asp Phe Asp Ala Leu Ser		
260	265	270
Asp Asp Leu Ala Ser Ile Leu Gly Asp Ser Phe Glu Leu Leu Lys		
275	280	285
Ala Lys Ala Val Tyr Asn Cys Ser Val Leu Ser Lys Val Ile Gly Asp		
290	295	300
Glu Gln Tyr Leu Ser Phe Ala Lys Val Lys Ile Tyr Glu Lys His Lys		
305	310	315
Thr Asp Leu Thr Lys Leu Lys Asn Val Ile Lys Lys His Phe Pro Lys		
325	330	335
Asp Tyr Lys Lys Val Phe Gly Tyr Asn Lys Asn Glu Lys Asn Asn Asn		
340	345	350
Asn Tyr Ser Gly Tyr Val Gly Val Cys Lys Thr Lys Ser Lys Lys Leu		
355	360	365
Ile Ile Asn Asn Ser Val Asn Gln Glu Asp Phe Tyr Lys Phe Leu Lys		
370	375	380
Thr Ile Leu Ser Ala Lys Ser Glu Ile Lys Glu Val Asn Asp Ile Leu		
385	390	395
Thr Glu Ile Glu Thr Gly Thr Phe Leu Pro Lys Gln Ile Ser Lys Ser		
405	410	415
Asn Ala Glu Ile Pro Tyr Gln Leu Arg Lys Met Glu Leu Glu Lys Ile		
420	425	430
Leu Ser Asn Ala Glu Lys His Phe Ser Phe Leu Lys Gln Lys Asp Glu		
435	440	445
Lys Gly Leu Ser His Ser Glu Lys Ile Ile Met Leu Leu Thr Phe Lys		
450	455	460
Ile Pro Tyr Tyr Ile Gly Pro Ile Asn Asp Asn His Lys Lys Phe Phe		
465	470	475
Pro Asp Arg Cys Trp Val Val Lys Lys Glu Lys Ser Pro Ser Gly Lys		
485	490	495
Thr Thr Pro Trp Asn Phe Phe Asp His Ile Asp Lys Glu Lys Thr Ala		
500	505	510
Glu Ala Phe Ile Thr Ser Arg Thr Asn Phe Cys Thr Tyr Leu Val Gly		
515	520	525
Glu Ser Val Leu Pro Lys Ser Ser Leu Leu Tyr Ser Glu Tyr Thr Val		
530	535	540
Leu Asn Glu Ile Asn Asn Leu Gln Ile Ile Ile Asp Gly Lys Asn Ile		

545	550	555	560
Cys Asp Ile Lys Leu Lys Gln Lys Ile Tyr Glu Asp Leu Phe Lys Lys			
	565	570	575
Tyr Lys Lys Ile Thr Gln Lys Gln Ile Ser Thr Phe Ile Lys His Glu			
	580	585	590
Gly Ile Cys Asn Lys Thr Asp Glu Val Ile Ile Leu Gly Ile Asp Lys			
	595	600	605
Glu Cys Thr Ser Ser Leu Lys Ser Tyr Ile Glu Leu Lys Asn Ile Phe			
	610	615	620
Gly Lys Gln Val Asp Glu Ile Ser Thr Lys Asn Met Leu Glu Glu Ile			
625	630	635	640
Ile Arg Trp Ala Thr Ile Tyr Asp Glu Gly Glu Gly Lys Thr Ile Leu			
	645	650	655
Lys Thr Lys Ile Lys Ala Glu Tyr Gly Lys Tyr Cys Ser Asp Glu Gln			
	660	665	670
Ile Lys Lys Ile Leu Asn Leu Lys Phe Ser Gly Trp Gly Arg Leu Ser			
	675	680	685
Arg Lys Phe Leu Glu Thr Val Thr Ser Glu Met Pro Gly Phe Ser Glu			
	690	695	700
Pro Val Asn Ile Ile Thr Ala Met Arg Glu Thr Gln Asn Asn Leu Met			
705	710	715	720
Glu Leu Leu Ser Ser Glu Phe Thr Phe Thr Glu Asn Ile Lys Lys Ile			
	725	730	735
Asn Ser Gly Phe Glu Asp Ala Glu Lys Gln Phe Ser Tyr Asp Gly Leu			
	740	745	750
Val Lys Pro Leu Phe Leu Ser Pro Ser Val Lys Lys Met Leu Trp Gln			
	755	760	765
Thr Leu Lys Leu Val Lys Glu Ile Ser His Ile Thr Gln Ala Pro Pro			
	770	775	780
Lys Lys Ile Phe Ile Glu Met Ala Lys Gly Ala Glu Leu Glu Pro Ala			
785	790	795	800
Arg Thr Lys Thr Arg Leu Lys Ile Leu Gln Asp Leu Tyr Asn Asn Cys			
	805	810	815
Lys Asn Asp Ala Asp Ala Phe Ser Ser Glu Ile Lys Asp Leu Ser Gly			
	820	825	830
Lys Ile Glu Asn Glu Asp Asn Leu Arg Leu Arg Ser Asp Lys Leu Tyr			
	835	840	845
Leu Tyr Tyr Thr Gln Leu Gly Lys Cys Met Tyr Cys Gly Lys Pro Ile			
	850	855	860

Glu	Ile	Gly	His	Val	Phe	Asp	Thr	Ser	Asn	Tyr	Asp	Ile	Asp	His	Ile
865						870					875				880
Tyr	Pro	Gln	Ser	Lys	Ile	Lys	Asp	Asp	Ser	Ile	Ser	Asn	Arg	Val	Leu
				885						890					895
Val	Cys	Ser	Ser	Cys	Asn	Lys	Asn	Lys	Glu	Asp	Lys	Tyr	Pro	Leu	Lys
				900					905					910	
Ser	Glu	Ile	Gln	Ser	Lys	Gln	Arg	Gly	Phe	Trp	Asn	Phe	Leu	Gln	Arg
				915					920					925	
Asn	Asn	Phe	Ile	Ser	Leu	Glu	Lys	Leu	Asn	Arg	Leu	Thr	Arg	Ala	Thr
				930					935					940	
Pro	Ile	Ser	Asp	Asp	Glu	Thr	Ala	Lys	Phe	Ile	Ala	Arg	Gln	Leu	Val
945					950					955					960
Glu	Thr	Arg	Gln	Ala	Thr	Lys	Val	Ala	Ala	Lys	Val	Leu	Glu	Lys	Met
				965						970					975
Phe	Pro	Glu	Thr	Lys	Ile	Val	Tyr	Ser	Lys	Ala	Glu	Thr	Val	Ser	Met
				980					985					990	
Phe	Arg	Asn	Lys	Phe	Asp	Ile	Val	Lys	Cys	Arg	Glu	Ile	Asn	Asp	Phe
				995					1000					1005	
His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn	Ile	Val	Val	Gly	Asn	Val	
				1010					1015					1020	
Tyr	Asn	Thr	Lys	Phe	Thr	Asn	Asn	Pro	Trp	Asn	Phe	Ile	Lys	Glu	
				1025					1030					1035	
Lys	Arg	Asp	Asn	Pro	Lys	Ile	Ala	Asp	Thr	Tyr	Asn	Tyr	Tyr	Lys	
				1040					1045					1050	
Val	Phe	Asp	Tyr	Asp	Val	Lys	Arg	Asn	Asn	Ile	Thr	Ala	Trp	Glu	
				1055					1060					1065	
Lys	Gly	Lys	Thr	Ile	Ile	Thr	Val	Lys	Asp	Met	Leu	Lys	Arg	Asn	
				1070					1075					1080	
Thr	Pro	Ile	Tyr	Thr	Arg	Gln	Ala	Ala	Cys	Lys	Lys	Gly	Glu	Leu	
				1085					1090					1095	
Phe	Asn	Gln	Thr	Ile	Met	Lys	Lys	Gly	Leu	Gly	Gln	His	Pro	Leu	
				1100					1105					1110	
Lys	Lys	Glu	Gly	Pro	Phe	Ser	Asn	Ile	Ser	Lys	Tyr	Gly	Gly	Tyr	
				1115					1120					1125	
Asn	Lys	Val	Ser	Ala	Ala	Tyr	Tyr	Thr	Leu	Ile	Glu	Tyr	Glu	Glu	
				1130					1135					1140	
Lys	Gly	Asn	Lys	Ile	Arg	Ser	Leu	Glu	Thr	Ile	Pro	Leu	Tyr	Leu	
				1145					1150					1155	
Val	Lys	Asp	Ile	Gln	Lys	Asp	Gln	Asp	Val	Leu	Lys	Ser	Tyr	Leu	

1160	1165	1170
Thr Asp Leu Leu Gly Lys	Lys Glu Phe Lys Ile	Leu Val Pro Lys
1175	1180	1185
Ile Lys Ile Asn Ser Leu	Leu Lys Ile Asn Gly	Phe Pro Cys His
1190	1195	1200
Ile Thr Gly Lys Thr Asn	Asp Ser Phe Leu Leu	Arg Pro Ala Val
1205	1210	1215
Gln Phe Cys Cys Ser Asn	Asn Glu Val Leu Tyr	Phe Lys Lys Ile
1220	1225	1230
Ile Arg Phe Ser Glu Ile	Arg Ser Gln Arg Glu	Lys Ile Gly Lys
1235	1240	1245
Thr Ile Ser Pro Tyr Glu	Asp Leu Ser Phe Arg	Ser Tyr Ile Lys
1250	1255	1260
Glu Asn Leu Trp Lys Lys	Thr Lys Asn Asp Glu	Ile Gly Glu Lys
1265	1270	1275
Glu Phe Tyr Asp Leu Leu	Gln Lys Lys Asn Leu	Glu Ile Tyr Asp
1280	1285	1290
Met Leu Leu Thr Lys His	Lys Asp Thr Ile Tyr	Lys Lys Arg Pro
1295	1300	1305
Asn Ser Ala Thr Ile Asp	Ile Leu Val Lys Gly	Lys Glu Lys Phe
1310	1315	1320
Lys Ser Leu Ile Ile Glu	Asn Gln Phe Glu Val	Ile Leu Glu Ile
1325	1330	1335
Leu Lys Leu Phe Ser Ala	Thr Arg Asn Val Ser	Asp Leu Gln His
1340	1345	1350
Ile Gly Gly Ser Lys Tyr	Ser Gly Val Ala Lys	Ile Gly Asn Lys
1355	1360	1365
Ile Ser Ser Leu Asp Asn	Cys Ile Leu Ile Tyr	Gln Ser Ile Thr
1370	1375	1380
Gly Ile Phe Glu Lys Arg	Ile Asp Leu Leu Lys	Val
1385	1390	1395
<210> 159		
<211> 1345		
<212> PRT		
<213> 变形链球菌 (Streptococcus mutans)		
<400> 159		
Met Lys Lys Pro Tyr Ser	Ile Gly Leu Asp Ile	Gly Thr Asn Ser Val
1	5	10
Gly Trp Ala Val Val Thr	Asp Asp Tyr Lys Val	Pro Ala Lys Lys Met

			20						25						30			
Lys	Val	Leu	Gly	Asn	Thr	Asp	Lys	Ser	His	Ile	Glu	Lys	Asn	Leu	Leu			
		35					40					45						
Gly	Ala	Leu	Leu	Phe	Asp	Ser	Gly	Asn	Thr	Ala	Glu	Asp	Arg	Arg	Leu			
	50					55					60							
Lys	Arg	Thr	Ala	Arg	Arg	Arg	Tyr	Thr	Arg	Arg	Arg	Asn	Arg	Ile	Leu			
65				70					75					80				
Tyr	Leu	Gln	Glu	Ile	Phe	Ser	Glu	Glu	Met	Gly	Lys	Val	Asp	Asp	Ser			
			85					90					95					
Phe	Phe	His	Arg	Leu	Glu	Asp	Ser	Phe	Leu	Val	Thr	Glu	Asp	Lys	Arg			
		100					105					110						
Gly	Glu	Arg	His	Pro	Ile	Phe	Gly	Asn	Leu	Glu	Glu	Glu	Val	Lys	Tyr			
	115					120					125							
His	Glu	Asn	Phe	Pro	Thr	Ile	Tyr	His	Leu	Arg	Gln	Tyr	Leu	Ala	Asp			
	130					135					140							
Asn	Pro	Glu	Lys	Val	Asp	Leu	Arg	Leu	Val	Tyr	Leu	Ala	Leu	Ala	His			
145				150					155					160				
Ile	Ile	Lys	Phe	Arg	Gly	His	Phe	Leu	Ile	Glu	Gly	Lys	Phe	Asp	Thr			
		165					170					175						
Arg	Asn	Asn	Asp	Val	Gln	Arg	Leu	Phe	Gln	Glu	Phe	Leu	Ala	Val	Tyr			
	180					185					190							
Asp	Asn	Thr	Phe	Glu	Asn	Ser	Ser	Leu	Gln	Glu	Gln	Asn	Val	Gln	Val			
	195					200					205							
Glu	Glu	Ile	Leu	Thr	Asp	Lys	Ile	Ser	Lys	Ser	Ala	Lys	Lys	Asp	Arg			
	210					215					220							
Val	Leu	Lys	Leu	Phe	Pro	Asn	Glu	Lys	Ser	Asn	Gly	Arg	Phe	Ala	Glu			
225				230					235					240				
Phe	Leu	Lys	Leu	Ile	Val	Gly	Asn	Gln	Ala	Asp	Phe	Lys	Lys	His	Phe			
		245					250					255						
Glu	Leu	Glu	Glu	Lys	Ala	Pro	Leu	Gln	Phe	Ser	Lys	Asp	Thr	Tyr	Glu			
	260					265					270							
Glu	Glu	Leu	Glu	Val	Leu	Leu	Ala	Gln	Ile	Gly	Asp	Asn	Tyr	Ala	Glu			
	275					280					285							
Leu	Phe	Leu	Ser	Ala	Lys	Lys	Leu	Tyr	Asp	Ser	Ile	Leu	Leu	Ser	Gly			
	290					295					300							
Ile	Leu	Thr	Val	Thr	Asp	Val	Gly	Thr	Lys	Ala	Pro	Leu	Ser	Ala	Ser			
305				310					315					320				
Met	Ile	Gln	Arg	Tyr	Asn	Glu	His	Gln	Met	Asp	Leu	Ala	Gln	Leu	Lys			
		325					330					335						

Gln Phe Ile Arg Gln Lys Leu Ser Asp Lys Tyr Asn Glu Val Phe Ser		
340	345	350
Asp Val Ser Lys Asp Gly Tyr Ala Gly Tyr Ile Asp Gly Lys Thr Asn		
355	360	365
Gln Glu Ala Phe Tyr Lys Tyr Leu Lys Gly Leu Leu Asn Lys Ile Glu		
370	375	380
Gly Ser Gly Tyr Phe Leu Asp Lys Ile Glu Arg Glu Asp Phe Leu Arg		
385	390	395
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu		
405	410	415
Gln Glu Met Arg Ala Ile Ile Arg Arg Gln Ala Glu Phe Tyr Pro Phe		
420	425	430
Leu Ala Asp Asn Gln Asp Arg Ile Glu Lys Leu Leu Thr Phe Arg Ile		
435	440	445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Lys Ser Asp Phe Ala Trp		
450	455	460
Leu Ser Arg Lys Ser Ala Asp Lys Ile Thr Pro Trp Asn Phe Asp Glu		
465	470	475
Ile Val Asp Lys Glu Ser Ser Ala Glu Ala Phe Ile Asn Arg Met Thr		
485	490	495
Asn Tyr Asp Leu Tyr Leu Pro Asn Gln Lys Val Leu Pro Lys His Ser		
500	505	510
Leu Leu Tyr Glu Lys Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys		
515	520	525
Tyr Lys Thr Glu Gln Gly Lys Thr Ala Phe Phe Asp Ala Asn Met Lys		
530	535	540
Gln Glu Ile Phe Asp Gly Val Phe Lys Val Tyr Arg Lys Val Thr Lys		
545	550	555
Asp Lys Leu Met Asp Phe Leu Glu Lys Glu Phe Asp Glu Phe Arg Ile		
565	570	575
Val Asp Leu Thr Gly Leu Asp Lys Glu Asn Lys Val Phe Asn Ala Ser		
580	585	590
Tyr Gly Thr Tyr His Asp Leu Cys Lys Ile Leu Asp Lys Asp Phe Leu		
595	600	605
Asp Asn Ser Lys Asn Glu Lys Ile Leu Glu Asp Ile Val Leu Thr Leu		
610	615	620
Thr Leu Phe Glu Asp Arg Glu Met Ile Arg Lys Arg Leu Glu Asn Tyr		
625	630	635
Ser Asp Leu Leu Thr Lys Glu Gln Val Lys Lys Leu Glu Arg Arg His		

645																650						655					
Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Ala	Glu	Leu	Ile	His	Gly	Ile	Arg												
660																665						670					
Asn	Lys	Glu	Ser	Arg	Lys	Thr	Ile	Leu	Asp	Tyr	Leu	Ile	Asp	Asp	Gly												
675																680						685					
Asn	Ser	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	Asn	Asp	Asp	Ala	Leu	Ser												
690																695						700					
Phe	Lys	Glu	Glu	Ile	Ala	Lys	Ala	Gln	Val	Ile	Gly	Glu	Thr	Asp	Asn												
705																710						715					
Leu	Asn	Gln	Val	Val	Ser	Asp	Ile	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys												
725																730						735					
Gly	Ile	Leu	Gln	Ser	Leu	Lys	Ile	Val	Asp	Glu	Leu	Val	Lys	Ile	Met												
740																745						750					
Gly	His	Gln	Pro	Glu	Asn	Ile	Val	Val	Glu	Met	Ala	Arg	Glu	Asn	Gln												
755																760						765					
Phe	Thr	Asn	Gln	Gly	Arg	Arg	Asn	Ser	Gln	Gln	Arg	Leu	Lys	Gly	Leu												
770																775						780					
Thr	Asp	Ser	Ile	Lys	Glu	Phe	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro												
785																790						795					
Val	Glu	Asn	Ser	Gln	Leu	Gln	Asn	Asp	Arg	Leu	Phe	Leu	Tyr	Tyr	Leu												
805																810						815					
Gln	Asn	Gly	Arg	Asp	Met	Tyr	Thr	Gly	Glu	Glu	Leu	Asp	Ile	Asp	Tyr												
820																825						830					
Leu	Ser	Gln	Tyr	Asp	Ile	Asp	His	Ile	Ile	Pro	Gln	Ala	Phe	Ile	Lys												
835																840						845					
Asp	Asn	Ser	Ile	Asp	Asn	Arg	Val	Leu	Thr	Ser	Ser	Lys	Glu	Asn	Arg												
850																855						860					
Gly	Lys	Ser	Asp	Asp	Val	Pro	Ser	Lys	Asp	Val	Val	Arg	Lys	Met	Lys												
865																870						875					
Ser	Tyr	Trp	Ser	Lys	Leu	Leu	Ser	Ala	Lys	Leu	Ile	Thr	Gln	Arg	Lys												
885																890						895					
Phe	Asp	Asn	Leu	Thr	Lys	Ala	Glu	Arg	Gly	Gly	Leu	Thr	Asp	Asp	Asp												
900																905						910					
Lys	Ala	Gly	Phe	Ile	Lys	Arg	Gln	Leu	Val	Glu	Thr	Arg	Gln	Ile	Thr												
915																920						925					
Lys	His	Val	Ala	Arg	Ile	Leu	Asp	Glu	Arg	Phe	Asn	Thr	Glu	Thr	Asp												
930																935						940					
Glu	Asn	Asn	Lys	Lys	Ile	Arg	Gln	Val	Lys	Ile	Val	Thr	Leu	Lys	Ser												
945																950						955					

Asn	Leu	Val	Ser	Asn	Phe	Arg	Lys	Glu	Phe	Glu	Leu	Tyr	Lys	Val	Arg
				965				970						975	
Glu	Ile	Asn	Asp	Tyr	His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn	Ala	Val
				980				985						990	
Ile	Gly	Lys	Ala	Leu	Leu	Gly	Val	Tyr	Pro	Gln	Leu	Glu	Pro	Glu	Phe
				995				1000						1005	
Val	Tyr	Gly	Asp	Tyr	Pro	His	Phe	His	Gly	His	Lys	Glu	Asn	Lys	
				1010				1015						1020	
Ala	Thr	Ala	Lys	Lys	Phe	Phe	Tyr	Ser	Asn	Ile	Met	Asn	Phe	Phe	
				1025				1030						1035	
Lys	Lys	Asp	Asp	Val	Arg	Thr	Asp	Lys	Asn	Gly	Glu	Ile	Ile	Trp	
				1040				1045						1050	
Lys	Lys	Asp	Glu	His	Ile	Ser	Asn	Ile	Lys	Lys	Val	Leu	Ser	Tyr	
				1055				1060						1065	
Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Val	Glu	Glu	Gln	Thr	Gly	Gly	
				1070				1075						1080	
Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys	Gly	Asn	Ser	Asp	Lys	Leu	
				1085				1090						1095	
Ile	Pro	Arg	Lys	Thr	Lys	Lys	Phe	Tyr	Trp	Asp	Thr	Lys	Lys	Tyr	
				1100				1105						1110	
Gly	Gly	Phe	Asp	Ser	Pro	Ile	Val	Ala	Tyr	Ser	Ile	Leu	Val	Ile	
				1115				1120						1125	
Ala	Asp	Ile	Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys	Thr	Val	Lys	
				1130				1135						1140	
Ala	Leu	Val	Gly	Val	Thr	Ile	Met	Glu	Lys	Met	Thr	Phe	Glu	Arg	
				1145				1150						1155	
Asp	Pro	Val	Ala	Phe	Leu	Glu	Arg	Lys	Gly	Tyr	Arg	Asn	Val	Gln	
				1160				1165						1170	
Glu	Glu	Asn	Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu	Phe	Lys	Leu	
				1175				1180						1185	
Glu	Asn	Gly	Arg	Lys	Arg	Leu	Leu	Ala	Ser	Ala	Arg	Glu	Leu	Gln	
				1190				1195						1200	
Lys	Gly	Asn	Glu	Ile	Val	Leu	Pro	Asn	His	Leu	Gly	Thr	Leu	Leu	
				1205				1210						1215	
Tyr	His	Ala	Lys	Asn	Ile	His	Lys	Val	Asp	Glu	Pro	Lys	His	Leu	
				1220				1225						1230	
Asp	Tyr	Val	Asp	Lys	His	Lys	Asp	Glu	Phe	Lys	Glu	Leu	Leu	Asp	
				1235				1240						1245	
Val	Val	Ser	Asn	Phe	Ser	Lys	Lys	Tyr	Thr	Leu	Ala	Glu	Gly	Asn	

1250	1255	1260
Leu Glu Lys Ile Lys Glu	Leu Tyr Ala Gln Asn	Asn Gly Glu Asp
1265	1270	1275
Leu Lys Glu Leu Ala Ser	Ser Phe Ile Asn Leu	Leu Thr Phe Thr
1280	1285	1290
Ala Ile Gly Ala Pro Ala	Thr Phe Lys Phe Phe	Asp Lys Asn Ile
1295	1300	1305
Asp Arg Lys Arg Tyr Thr	Ser Thr Thr Glu Ile	Leu Asn Ala Thr
1310	1315	1320
Leu Ile His Gln Ser Ile	Thr Gly Leu Tyr Glu	Thr Arg Ile Asp
1325	1330	1335
Leu Asn Lys Leu Gly Gly	Asp	
1340	1345	
<210> 160		
<211> 1388		
<212> PRT		
<213> 嗜热链球菌 (Streptococcus thermophilus)		
<400> 160		
Met Thr Lys Pro Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val		
1	5	10
Gly Trp Ala Val Thr Thr Asp Asn Tyr Lys Val Pro Ser Lys Lys Met		
20	25	30
Lys Val Leu Gly Asn Thr Ser Lys Lys Tyr Ile Lys Lys Asn Leu Leu		
35	40	45
Gly Val Leu Leu Phe Asp Ser Gly Ile Thr Ala Glu Gly Arg Arg Leu		
50	55	60
Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Arg Asn Arg Ile Leu		
65	70	75
Tyr Leu Gln Glu Ile Phe Ser Thr Glu Met Ala Thr Leu Asp Asp Ala		
85	90	95
Phe Phe Gln Arg Leu Asp Asp Ser Phe Leu Val Pro Asp Asp Lys Arg		
100	105	110
Asp Ser Lys Tyr Pro Ile Phe Gly Asn Leu Val Glu Glu Lys Ala Tyr		
115	120	125
His Asp Glu Phe Pro Thr Ile Tyr His Leu Arg Lys Tyr Leu Ala Asp		
130	135	140
Ser Thr Lys Lys Ala Asp Leu Arg Leu Val Tyr Leu Ala Leu Ala His		
145	150	155
Met Ile Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Glu Phe Asn Ser		160

	165		170		175
Lys Asn Asn Asp	Ile Gln Lys Asn Phe	Gln Asp Phe Leu	Asp Thr Tyr		
	180		185		190
Asn Ala Ile Phe	Glu Ser Asp Leu	Ser Leu Glu Asn	Ser Lys Gln Leu		
	195		200		205
Glu Glu Ile Val	Lys Asp Lys Ile	Ser Lys Leu Glu	Lys Lys Asp Arg		
	210		215		220
Ile Leu Lys Leu	Phe Pro Gly Glu	Lys Asn Ser Gly	Ile Phe Ser Glu		
225		230		235	240
Phe Leu Lys Leu	Ile Val Gly Asn	Gln Ala Asp Phe	Arg Lys Cys Phe		
	245		250		255
Asn Leu Asp Glu	Lys Ala Ser Leu	His Phe Ser Lys	Glu Ser Tyr Asp		
	260		265		270
Glu Asp Leu Glu	Thr Leu Leu Gly	Tyr Ile Gly Asp	Asp Tyr Ser Asp		
	275		280		285
Val Phe Leu Lys	Ala Lys Lys Leu	Tyr Asp Ala Ile	Leu Leu Ser Gly		
	290		295		300
Phe Leu Thr Val	Thr Asp Asn Glu	Thr Glu Ala Pro	Leu Ser Ser Ala		
305		310		315	320
Met Ile Lys Arg	Tyr Asn Glu His	Lys Glu Asp Leu	Ala Leu Leu Lys		
	325		330		335
Glu Tyr Ile Arg	Asn Ile Ser Leu	Lys Thr Tyr Asn	Glu Val Phe Lys		
	340		345		350
Asp Asp Thr Lys	Asn Gly Tyr Ala	Gly Tyr Ile Asp	Gly Lys Thr Asn		
	355		360		365
Gln Glu Asp Phe	Tyr Val Tyr Leu	Lys Lys Leu Leu	Ala Glu Phe Glu		
	370		375		380
Gly Ala Asp Tyr	Phe Leu Glu Lys	Ile Asp Arg Glu	Asp Phe Leu Arg		
385		390		395	400
Lys Gln Arg Thr	Phe Asp Asn Gly	Ser Ile Pro Tyr	Gln Ile His Leu		
	405		410		415
Gln Glu Met Arg	Ala Ile Leu Asp	Lys Gln Ala Lys	Phe Tyr Pro Phe		
	420		425		430
Leu Ala Lys Asn	Lys Glu Arg Ile	Glu Lys Ile Leu	Thr Phe Arg Ile		
	435		440		445
Pro Tyr Tyr Val	Gly Pro Leu Ala	Arg Gly Asn Ser	Asp Phe Ala Trp		
	450		455		460
Ser Ile Arg Lys	Arg Asn Glu Lys	Ile Thr Pro Trp	Asn Phe Glu Asp		
465		470		475	480

Val	Ile	Asp	Lys	Glu	Ser	Ser	Ala	Glu	Ala	Phe	Ile	Asn	Arg	Met	Thr			
																485	490	495
Ser	Phe	Asp	Leu	Tyr	Leu	Pro	Glu	Glu	Lys	Val	Leu	Pro	Lys	His	Ser			
																500	505	510
Leu	Leu	Tyr	Glu	Thr	Phe	Asn	Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Arg			
																515	520	525
Phe	Ile	Ala	Glu	Ser	Met	Arg	Asp	Tyr	Gln	Phe	Leu	Asp	Ser	Lys	Gln			
																530	535	540
Lys	Lys	Asp	Ile	Val	Arg	Leu	Tyr	Phe	Lys	Asp	Lys	Arg	Lys	Val	Thr			
																545	550	555
Asp	Lys	Asp	Ile	Ile	Glu	Tyr	Leu	His	Ala	Ile	Tyr	Gly	Tyr	Asp	Gly			
																565	570	575
Ile	Glu	Leu	Lys	Gly	Ile	Glu	Lys	Gln	Phe	Asn	Ser	Ser	Leu	Ser	Thr			
																580	585	590
Tyr	His	Asp	Leu	Leu	Asn	Ile	Ile	Asn	Asp	Lys	Glu	Phe	Leu	Asp	Asp			
																595	600	605
Ser	Ser	Asn	Glu	Ala	Ile	Ile	Glu	Glu	Ile	Ile	His	Thr	Leu	Thr	Ile			
																610	615	620
Phe	Glu	Asp	Arg	Glu	Met	Ile	Lys	Gln	Arg	Leu	Ser	Lys	Phe	Glu	Asn			
																625	630	635
Ile	Phe	Asp	Lys	Ser	Val	Leu	Lys	Lys	Leu	Ser	Arg	Arg	His	Tyr	Thr			
																645	650	655
Gly	Trp	Gly	Lys	Leu	Ser	Ala	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp	Glu			
																660	665	670
Lys	Ser	Gly	Asn	Thr	Ile	Leu	Asp	Tyr	Leu	Ile	Asp	Asp	Gly	Ile	Ser			
																675	680	685
Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ala	Leu	Ser	Phe	Lys			
																690	695	700
Lys	Lys	Ile	Gln	Lys	Ala	Gln	Ile	Ile	Gly	Asp	Glu	Asp	Lys	Gly	Asn			
																705	710	715
Ile	Lys	Glu	Val	Val	Lys	Ser	Leu	Pro	Gly	Ser	Pro	Ala	Ile	Lys	Lys			
																725	730	735
Gly	Ile	Leu	Gln	Ser	Ile	Lys	Ile	Val	Asp	Glu	Leu	Val	Lys	Val	Met			
																740	745	750
Gly	Gly	Arg	Lys	Pro	Glu	Ser	Ile	Val	Val	Glu	Met	Ala	Arg	Glu	Asn			
																755	760	765
Gln	Tyr	Thr	Asn	Gln	Gly	Lys	Ser	Asn	Ser	Gln	Gln	Arg	Leu	Lys	Arg			
																770	775	780
Leu	Glu	Lys	Ser	Leu	Lys	Glu	Leu	Gly	Ser	Lys	Ile	Leu	Lys	Glu	Asn			

785	790	795	800
Ile Pro Ala Lys Leu Ser Lys Ile Asp Asn Asn Ala Leu Gln Asn Asp			
	805	810	815
Arg Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Lys Asp Met Tyr Thr Gly			
	820	825	830
Asp Asp Leu Asp Ile Asp Arg Leu Ser Asn Tyr Asp Ile Asp His Ile			
	835	840	845
Ile Pro Gln Ala Phe Leu Lys Asp Asn Ser Ile Asp Asn Lys Val Leu			
	850	855	860
Val Ser Ser Ala Ser Asn Arg Gly Lys Ser Asp Asp Val Pro Ser Leu			
865	870	875	880
Glu Val Val Lys Lys Arg Lys Thr Phe Trp Tyr Gln Leu Leu Lys Ser			
	885	890	895
Lys Leu Ile Ser Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg			
	900	905	910
Gly Gly Leu Ser Pro Glu Asp Lys Ala Gly Phe Ile Gln Arg Gln Leu			
	915	920	925
Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Arg Leu Leu Asp Glu			
	930	935	940
Lys Phe Asn Asn Lys Lys Asp Glu Asn Asn Arg Ala Val Arg Thr Val			
945	950	955	960
Lys Ile Ile Thr Leu Lys Ser Thr Leu Val Ser Gln Phe Arg Lys Asp			
	965	970	975
Phe Glu Leu Tyr Lys Val Arg Glu Ile Asn Asp Phe His His Ala His			
	980	985	990
Asp Ala Tyr Leu Asn Ala Val Val Ala Ser Ala Leu Leu Lys Lys Tyr			
	995	1000	1005
Pro Lys Leu Glu Pro Glu Phe Val Tyr Gly Asp Tyr Pro Lys Tyr			
	1010	1015	1020
Asn Ser Phe Arg Glu Arg Lys Ser Ala Thr Glu Lys Val Tyr Phe			
	1025	1030	1035
Tyr Ser Asn Ile Met Asn Ile Phe Lys Lys Ser Ile Ser Leu Ala			
	1040	1045	1050
Asp Gly Arg Val Ile Glu Arg Pro Leu Ile Glu Val Asn Glu Glu			
	1055	1060	1065
Thr Gly Glu Ser Val Trp Asn Lys Glu Ser Asp Leu Ala Thr Val			
	1070	1075	1080
Arg Arg Val Leu Ser Tyr Pro Gln Val Asn Val Val Lys Lys Val			
	1085	1090	1095

Glu Glu Gln Asn His Gly Leu Asp Arg Gly Lys Pro Lys Gly Leu		
1100	1105	1110
Phe Asn Ala Asn Leu Ser Ser Lys Pro Lys Pro Asn Ser Asn Glu		
1115	1120	1125
Asn Leu Val Gly Ala Lys Glu Tyr Leu Asp Pro Lys Lys Tyr Gly		
1130	1135	1140
Gly Tyr Ala Gly Ile Ser Asn Ser Phe Thr Val Leu Val Lys Gly		
1145	1150	1155
Thr Ile Glu Lys Gly Ala Lys Lys Lys Ile Thr Asn Val Leu Glu		
1160	1165	1170
Phe Gln Gly Ile Ser Ile Leu Asp Arg Ile Asn Tyr Arg Lys Asp		
1175	1180	1185
Lys Leu Asn Phe Leu Leu Glu Lys Gly Tyr Lys Asp Ile Glu Leu		
1190	1195	1200
Ile Ile Glu Leu Pro Lys Tyr Ser Leu Phe Glu Leu Ser Asp Gly		
1205	1210	1215
Ser Arg Arg Met Leu Ala Ser Ile Leu Ser Thr Asn Asn Lys Arg		
1220	1225	1230
Gly Glu Ile His Lys Gly Asn Gln Ile Phe Leu Ser Gln Lys Phe		
1235	1240	1245
Val Lys Leu Leu Tyr His Ala Lys Arg Ile Ser Asn Thr Ile Asn		
1250	1255	1260
Glu Asn His Arg Lys Tyr Val Glu Asn His Lys Lys Glu Phe Glu		
1265	1270	1275
Glu Leu Phe Tyr Tyr Ile Leu Glu Phe Asn Glu Asn Tyr Val Gly		
1280	1285	1290
Ala Lys Lys Asn Gly Lys Leu Leu Asn Ser Ala Phe Gln Ser Trp		
1295	1300	1305
Gln Asn His Ser Ile Asp Glu Leu Cys Ser Ser Phe Ile Gly Pro		
1310	1315	1320
Thr Gly Ser Glu Arg Lys Gly Leu Phe Glu Leu Thr Ser Arg Gly		
1325	1330	1335
Ser Ala Ala Asp Phe Glu Phe Leu Gly Val Lys Ile Pro Arg Tyr		
1340	1345	1350
Arg Asp Tyr Thr Pro Ser Ser Leu Leu Lys Asp Ala Thr Leu Ile		
1355	1360	1365
His Gln Ser Val Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ala		
1370	1375	1380
Lys Leu Gly Glu Gly		

1385
 <210> 161
 <211> 984
 <212> PRT
 <213> 空肠弯曲杆菌 (Campylobacter jejuni)
 <400> 161
 Met Ala Arg Ile Leu Ala Phe Asp Ile Gly Ile Ser Ser Ile Gly Trp
 1 5 10 15
 Ala Phe Ser Glu Asn Asp Glu Leu Lys Asp Cys Gly Val Arg Ile Phe
 20 25 30
 Thr Lys Val Glu Asn Pro Lys Thr Gly Glu Ser Leu Ala Leu Pro Arg
 35 40 45
 Arg Leu Ala Arg Ser Ala Arg Lys Arg Leu Ala Arg Arg Lys Ala Arg
 50 55 60
 Leu Asn His Leu Lys His Leu Ile Ala Asn Glu Phe Lys Leu Asn Tyr
 65 70 75 80
 Glu Asp Tyr Gln Ser Phe Asp Glu Ser Leu Ala Lys Ala Tyr Lys Gly
 85 90 95
 Ser Leu Ile Ser Pro Tyr Glu Leu Arg Phe Arg Ala Leu Asn Glu Leu
 100 105 110
 Leu Ser Lys Gln Asp Phe Ala Arg Val Ile Leu His Ile Ala Lys Arg
 115 120 125
 Arg Gly Tyr Asp Asp Ile Lys Asn Ser Asp Asp Lys Glu Lys Gly Ala
 130 135 140
 Ile Leu Lys Ala Ile Lys Gln Asn Glu Glu Lys Leu Ala Asn Tyr Gln
 145 150 155 160
 Ser Val Gly Glu Tyr Leu Tyr Lys Glu Tyr Phe Gln Lys Phe Lys Glu
 165 170 175
 Asn Ser Lys Glu Phe Thr Asn Val Arg Asn Lys Lys Glu Ser Tyr Glu
 180 185 190
 Arg Cys Ile Ala Gln Ser Phe Leu Lys Asp Glu Leu Lys Leu Ile Phe
 195 200 205
 Lys Lys Gln Arg Glu Phe Gly Phe Ser Phe Ser Lys Lys Phe Glu Glu
 210 215 220
 Glu Val Leu Ser Val Ala Phe Tyr Lys Arg Ala Leu Lys Asp Phe Ser
 225 230 235 240
 His Leu Val Gly Asn Cys Ser Phe Phe Thr Asp Glu Lys Arg Ala Pro
 245 250 255
 Lys Asn Ser Pro Leu Ala Phe Met Phe Val Ala Leu Thr Arg Ile Ile

260	265	270
Asn Leu Leu Asn Asn Leu Lys	Asn Thr Glu Gly Ile Leu Tyr Thr Lys	
275	280	285
Asp Asp Leu Asn Ala Leu Leu	Asn Glu Val Leu Lys Asn Gly Thr Leu	
290	295	300
Thr Tyr Lys Gln Thr Lys Lys	Leu Leu Gly Leu Ser Asp Asp Tyr Glu	
305	310	315
Phe Lys Gly Glu Lys Gly Thr Tyr Phe	Ile Glu Phe Lys Lys Tyr Lys	
325	330	335
Glu Phe Ile Lys Ala Leu Gly Glu His	Asn Leu Ser Gln Asp Asp Leu	
340	345	350
Asn Glu Ile Ala Lys Asp Ile Thr Leu	Ile Lys Asp Glu Ile Lys Leu	
355	360	365
Lys Lys Ala Leu Ala Lys Tyr Asp Leu	Asn Gln Asn Gln Ile Asp Ser	
370	375	380
Leu Ser Lys Leu Glu Phe Lys Asp His	Leu Asn Ile Ser Phe Lys Ala	
385	390	395
Leu Lys Leu Val Thr Pro Leu Met Leu	Glu Gly Lys Lys Tyr Asp Glu	
405	410	415
Ala Cys Asn Glu Leu Asn Leu Lys Val	Ala Ile Asn Glu Asp Lys Lys	
420	425	430
Asp Phe Leu Pro Ala Phe Asn Glu Thr	Tyr Tyr Lys Asp Glu Val Thr	
435	440	445
Asn Pro Val Val Leu Arg Ala Ile Lys	Glu Tyr Arg Lys Val Leu Asn	
450	455	460
Ala Leu Leu Lys Lys Tyr Gly Lys Val	His Lys Ile Asn Ile Glu Leu	
465	470	475
Ala Arg Glu Val Gly Lys Asn His Ser	Gln Arg Ala Lys Ile Glu Lys	
485	490	495
Glu Gln Asn Glu Asn Tyr Lys Ala Lys	Lys Asp Ala Glu Leu Glu Cys	
500	505	510
Glu Lys Leu Gly Leu Lys Ile Asn Ser	Lys Asn Ile Leu Lys Leu Arg	
515	520	525
Leu Phe Lys Glu Gln Lys Glu Phe Cys	Ala Tyr Ser Gly Glu Lys Ile	
530	535	540
Lys Ile Ser Asp Leu Gln Asp Glu Lys	Met Leu Glu Ile Asp His Ile	
545	550	555
Tyr Pro Tyr Ser Arg Ser Phe Asp Asp	Ser Tyr Met Asn Lys Val Leu	
565	570	575

Val Phe Thr Lys Gln Asn Gln Glu Lys Leu Asn Gln Thr Pro Phe Glu		
580	585	590
Ala Phe Gly Asn Asp Ser Ala Lys Trp Gln Lys Ile Glu Val Leu Ala		
595	600	605
Lys Asn Leu Pro Thr Lys Lys Gln Lys Arg Ile Leu Asp Lys Asn Tyr		
610	615	620
Lys Asp Lys Glu Gln Lys Asn Phe Lys Asp Arg Asn Leu Asn Asp Thr		
625	630	635
Arg Tyr Ile Ala Arg Leu Val Leu Asn Tyr Thr Lys Asp Tyr Leu Asp		
645	650	655
Phe Leu Pro Leu Ser Asp Asp Glu Asn Thr Lys Leu Asn Asp Thr Gln		
660	665	670
Lys Gly Ser Lys Val His Val Glu Ala Lys Ser Gly Met Leu Thr Ser		
675	680	685
Ala Leu Arg His Thr Trp Gly Phe Ser Ala Lys Asp Arg Asn Asn His		
690	695	700
Leu His His Ala Ile Asp Ala Val Ile Ile Ala Tyr Ala Asn Asn Ser		
705	710	715
Ile Val Lys Ala Phe Ser Asp Phe Lys Lys Glu Gln Glu Ser Asn Ser		
725	730	735
Ala Glu Leu Tyr Ala Lys Lys Ile Ser Glu Leu Asp Tyr Lys Asn Lys		
740	745	750
Arg Lys Phe Phe Glu Pro Phe Ser Gly Phe Arg Gln Lys Val Leu Asp		
755	760	765
Lys Ile Asp Glu Ile Phe Val Ser Lys Pro Glu Arg Lys Lys Pro Ser		
770	775	780
Gly Ala Leu His Glu Glu Thr Phe Arg Lys Glu Glu Glu Phe Tyr Gln		
785	790	795
Ser Tyr Gly Gly Lys Glu Gly Val Leu Lys Ala Leu Glu Leu Gly Lys		
805	810	815
Ile Arg Lys Val Asn Gly Lys Ile Val Lys Asn Gly Asp Met Phe Arg		
820	825	830
Val Asp Ile Phe Lys His Lys Lys Thr Asn Lys Phe Tyr Ala Val Pro		
835	840	845
Ile Tyr Thr Met Asp Phe Ala Leu Lys Val Leu Pro Asn Lys Ala Val		
850	855	860
Ala Arg Ser Lys Lys Gly Glu Ile Lys Asp Trp Ile Leu Met Asp Glu		
865	870	875
Asn Tyr Glu Phe Cys Phe Ser Leu Tyr Lys Asp Ser Leu Ile Leu Ile		

				885					890					895			
Gln	Thr	Lys	Asp	Met	Gln	Glu	Pro	Glu	Phe	Val	Tyr	Tyr	Asn	Ala	Phe		
				900					905					910			
Thr	Ser	Ser	Thr	Val	Ser	Leu	Ile	Val	Ser	Lys	His	Asp	Asn	Lys	Phe		
				915					920					925			
Glu	Thr	Leu	Ser	Lys	Asn	Gln	Lys	Ile	Leu	Phe	Lys	Asn	Ala	Asn	Glu		
				930					935					940			
Lys	Glu	Val	Ile	Ala	Lys	Ser	Ile	Gly	Ile	Gln	Asn	Leu	Lys	Val	Phe		
945					950					955					960		
Glu	Lys	Tyr	Ile	Val	Ser	Ala	Leu	Gly	Glu	Val	Thr	Lys	Ala	Glu	Phe		
				965					970						975		
Arg	Gln	Arg	Glu	Asp	Phe	Lys	Lys										
				980													
<210> 162																	
<211> 1056																	
<212> PRT																	
<213> 多杀巴斯德菌(Pasteurella multocida)																	
<400> 162																	
Met	Gln	Thr	Thr	Asn	Leu	Ser	Tyr	Ile	Leu	Gly	Leu	Asp	Leu	Gly	Ile		
1				5					10					15			
Ala	Ser	Val	Gly	Trp	Ala	Val	Val	Glu	Ile	Asn	Glu	Asn	Glu	Asp	Pro		
				20					25					30			
Ile	Gly	Leu	Ile	Asp	Val	Gly	Val	Arg	Ile	Phe	Glu	Arg	Ala	Glu	Val		
				35					40					45			
Pro	Lys	Thr	Gly	Glu	Ser	Leu	Ala	Leu	Ser	Arg	Arg	Leu	Ala	Arg	Ser		
				50					55					60			
Thr	Arg	Arg	Leu	Ile	Arg	Arg	Arg	Ala	His	Arg	Leu	Leu	Leu	Ala	Lys		
65					70					75					80		
Arg	Phe	Leu	Lys	Arg	Glu	Gly	Ile	Leu	Ser	Thr	Ile	Asp	Leu	Glu	Lys		
				85					90						95		
Gly	Leu	Pro	Asn	Gln	Ala	Trp	Glu	Leu	Arg	Val	Ala	Gly	Leu	Glu	Arg		
				100					105					110			
Arg	Leu	Ser	Ala	Ile	Glu	Trp	Gly	Ala	Val	Leu	Leu	His	Leu	Ile	Lys		
				115					120					125			
His	Arg	Gly	Tyr	Leu	Ser	Lys	Arg	Lys	Asn	Glu	Ser	Gln	Thr	Asn	Asn		
				130					135					140			
Lys	Glu	Leu	Gly	Ala	Leu	Leu	Ser	Gly	Val	Ala	Gln	Asn	His	Gln	Leu		
145					150					155					160		
Leu	Gln	Ser	Asp	Asp	Tyr	Arg	Thr	Pro	Ala	Glu	Leu	Ala	Leu	Lys	Lys		

				165				170				175			
Phe	Ala	Lys	Glu	Glu	Gly	His	Ile	Arg	Asn	Gln	Arg	Gly	Ala	Tyr	Thr
				180				185				190			
His	Thr	Phe	Asn	Arg	Leu	Asp	Leu	Leu	Ala	Glu	Leu	Asn	Leu	Leu	Phe
				195				200				205			
Ala	Gln	Gln	His	Gln	Phe	Gly	Asn	Pro	His	Cys	Lys	Glu	His	Ile	Gln
				210				215				220			
Gln	Tyr	Met	Thr	Glu	Leu	Leu	Met	Trp	Gln	Lys	Pro	Ala	Leu	Ser	Gly
225				230				235				240			
Glu	Ala	Ile	Leu	Lys	Met	Leu	Gly	Lys	Cys	Thr	His	Glu	Lys	Asn	Glu
				245				250				255			
Phe	Lys	Ala	Ala	Lys	His	Thr	Tyr	Ser	Ala	Glu	Arg	Phe	Val	Trp	Leu
				260				265				270			
Thr	Lys	Leu	Asn	Asn	Leu	Arg	Ile	Leu	Glu	Asp	Gly	Ala	Glu	Arg	Ala
				275				280				285			
Leu	Asn	Glu	Glu	Glu	Arg	Gln	Leu	Leu	Ile	Asn	His	Pro	Tyr	Glu	Lys
				290				295				300			
Ser	Lys	Leu	Thr	Tyr	Ala	Gln	Val	Arg	Lys	Leu	Leu	Gly	Leu	Ser	Glu
305				310				315				320			
Gln	Ala	Ile	Phe	Lys	His	Leu	Arg	Tyr	Ser	Lys	Glu	Asn	Ala	Glu	Ser
				325				330				335			
Ala	Thr	Phe	Met	Glu	Leu	Lys	Ala	Trp	His	Ala	Ile	Arg	Lys	Ala	Leu
				340				345				350			
Glu	Asn	Gln	Gly	Leu	Lys	Asp	Thr	Trp	Gln	Asp	Leu	Ala	Lys	Lys	Pro
				355				360				365			
Asp	Leu	Leu	Asp	Glu	Ile	Gly	Thr	Ala	Phe	Ser	Leu	Tyr	Lys	Thr	Asp
				370				375				380			
Glu	Asp	Ile	Gln	Gln	Tyr	Leu	Thr	Asn	Lys	Val	Pro	Asn	Ser	Val	Ile
385				390				395				400			
Asn	Ala	Leu	Leu	Val	Ser	Leu	Asn	Phe	Asp	Lys	Phe	Ile	Glu	Leu	Ser
				405				410				415			
Leu	Lys	Ser	Leu	Arg	Lys	Ile	Leu	Pro	Leu	Met	Glu	Gln	Gly	Lys	Arg
				420				425				430			
Tyr	Asp	Gln	Ala	Cys	Arg	Glu	Ile	Tyr	Gly	His	His	Tyr	Gly	Glu	Ala
				435				440				445			
Asn	Gln	Lys	Thr	Ser	Gln	Leu	Leu	Pro	Ala	Ile	Pro	Ala	Gln	Glu	Ile
				450				455				460			
Arg	Asn	Pro	Val	Val	Leu	Arg	Thr	Leu	Ser	Gln	Ala	Arg	Lys	Val	Ile
465				470				475				480			

Asn	Ala	Ile	Ile	Arg	Gln	Tyr	Gly	Ser	Pro	Ala	Arg	Val	His	Ile	Glu			
																485	490	495
Thr	Gly	Arg	Glu	Leu	Gly	Lys	Ser	Phe	Lys	Glu	Arg	Arg	Glu	Ile	Gln			
																500	505	510
Lys	Gln	Gln	Glu	Asp	Asn	Arg	Thr	Lys	Arg	Glu	Ser	Ala	Val	Gln	Lys			
																515	520	525
Phe	Lys	Glu	Leu	Phe	Ser	Asp	Phe	Ser	Ser	Glu	Pro	Lys	Ser	Lys	Asp			
																530	535	540
Ile	Leu	Lys	Phe	Arg	Leu	Tyr	Glu	Gln	Gln	His	Gly	Lys	Cys	Leu	Tyr			
																545	550	555
Ser	Gly	Lys	Glu	Ile	Asn	Ile	His	Arg	Leu	Asn	Glu	Lys	Gly	Tyr	Val			
																565	570	575
Glu	Ile	Asp	His	Ala	Leu	Pro	Phe	Ser	Arg	Thr	Trp	Asp	Asp	Ser	Phe			
																580	585	590
Asn	Asn	Lys	Val	Leu	Val	Leu	Ala	Ser	Glu	Asn	Gln	Asn	Lys	Gly	Asn			
																595	600	605
Gln	Thr	Pro	Tyr	Glu	Trp	Leu	Gln	Gly	Lys	Ile	Asn	Ser	Glu	Arg	Trp			
																610	615	620
Lys	Asn	Phe	Val	Ala	Leu	Val	Leu	Gly	Ser	Gln	Cys	Ser	Ala	Ala	Lys			
																625	630	635
Lys	Gln	Arg	Leu	Leu	Thr	Gln	Val	Ile	Asp	Asp	Asn	Lys	Phe	Ile	Asp			
																645	650	655
Arg	Asn	Leu	Asn	Asp	Thr	Arg	Tyr	Ile	Ala	Arg	Phe	Leu	Ser	Asn	Tyr			
																660	665	670
Ile	Gln	Glu	Asn	Leu	Leu	Leu	Val	Gly	Lys	Asn	Lys	Lys	Asn	Val	Phe			
																675	680	685
Thr	Pro	Asn	Gly	Gln	Ile	Thr	Ala	Leu	Leu	Arg	Ser	Arg	Trp	Gly	Leu			
																690	695	700
Ile	Lys	Ala	Arg	Glu	Asn	Asn	Asn	Arg	His	His	Ala	Leu	Asp	Ala	Ile			
																705	710	715
Val	Val	Ala	Cys	Ala	Thr	Pro	Ser	Met	Gln	Gln	Lys	Ile	Thr	Arg	Phe			
																725	730	735
Ile	Arg	Phe	Lys	Glu	Val	His	Pro	Tyr	Lys	Ile	Glu	Asn	Arg	Tyr	Glu			
																740	745	750
Met	Val	Asp	Gln	Glu	Ser	Gly	Glu	Ile	Ile	Ser	Pro	His	Phe	Pro	Glu			
																755	760	765
Pro	Trp	Ala	Tyr	Phe	Arg	Gln	Glu	Val	Asn	Ile	Arg	Val	Phe	Asp	Asn			
																770	775	780
His	Pro	Asp	Thr	Val	Leu	Lys	Glu	Met	Leu	Pro	Asp	Arg	Pro	Gln	Ala			

785	790	795	800
Asn His Gln Phe Val	Gln Pro Leu Phe Val	Ser Arg Ala Pro Thr	Arg
805	810	815	
Lys Met Ser Gly Gln Gly	His Met Glu Thr Ile	Lys Ser Ala Lys	Arg
820	825	830	
Leu Ala Glu Gly Ile Ser	Val Leu Arg Ile Pro	Leu Thr Gln Leu	Lys
835	840	845	
Pro Asn Leu Leu Glu Asn	Met Val Asn Lys Glu	Arg Glu Pro Ala	Leu
850	855	860	
Tyr Ala Gly Leu Lys Ala	Arg Leu Ala Glu Phe	Asn Gln Asp Pro	Ala
865	870	875	880
Lys Ala Phe Ala Thr Pro	Phe Tyr Lys Gln Gly	Gly Gln Gln Val	Lys
885	890	895	
Ala Ile Arg Val Glu Gln	Val Gln Lys Ser Gly	Val Leu Val Arg	Glu
900	905	910	
Asn Asn Gly Val Ala Asp	Asn Ala Ser Ile Val	Arg Thr Asp Val	Phe
915	920	925	
Ile Lys Asn Asn Lys Phe	Phe Leu Val Pro Ile	Tyr Thr Trp Gln	Val
930	935	940	
Ala Lys Gly Ile Leu Pro	Asn Lys Ala Ile Val	Ala His Lys Asn	Glu
945	950	955	960
Asp Glu Trp Glu Glu Met	Asp Glu Gly Ala Lys	Phe Lys Phe Ser	Leu
965	970	975	
Phe Pro Asn Asp Leu Val	Glu Leu Lys Thr Lys	Lys Glu Tyr Phe	Phe
980	985	990	
Gly Tyr Tyr Ile Gly Leu	Asp Arg Ala Thr Gly	Asn Ile Ser Leu	Lys
995	1000	1005	
Glu His Asp Gly Glu Ile	Ser Lys Gly Lys Asp	Gly Val Tyr Arg	
1010	1015	1020	
Val Gly Val Lys Leu Ala	Leu Ser Phe Glu Lys	Tyr Gln Val Asp	
1025	1030	1035	
Glu Leu Gly Lys Asn Arg	Gln Ile Cys Arg Pro	Gln Gln Arg Gln	
1040	1045	1050	
Pro Val Arg			
1055			
<210> 163			
<211> 1629			
<212> PRT			
<213> 新凶手弗朗西丝菌 (Francisella novicida)			

<400> 163

Met	Asn	Phe	Lys	Ile	Leu	Pro	Ile	Ala	Ile	Asp	Leu	Gly	Val	Lys	Asn
1				5					10					15	
Thr	Gly	Val	Phe	Ser	Ala	Phe	Tyr	Gln	Lys	Gly	Thr	Ser	Leu	Glu	Arg
			20					25					30		
Leu	Asp	Asn	Lys	Asn	Gly	Lys	Val	Tyr	Glu	Leu	Ser	Lys	Asp	Ser	Tyr
		35					40					45			
Thr	Leu	Leu	Met	Asn	Asn	Arg	Thr	Ala	Arg	Arg	His	Gln	Arg	Arg	Gly
	50					55					60				
Ile	Asp	Arg	Lys	Gln	Leu	Val	Lys	Arg	Leu	Phe	Lys	Leu	Ile	Trp	Thr
65				70					75					80	
Glu	Gln	Leu	Asn	Leu	Glu	Trp	Asp	Lys	Asp	Thr	Gln	Gln	Ala	Ile	Ser
			85					90					95		
Phe	Leu	Phe	Asn	Arg	Arg	Gly	Phe	Ser	Phe	Ile	Thr	Asp	Gly	Tyr	Ser
			100					105					110		
Pro	Glu	Tyr	Leu	Asn	Ile	Val	Pro	Glu	Gln	Val	Lys	Ala	Ile	Leu	Met
	115					120					125				
Asp	Ile	Phe	Asp	Asp	Tyr	Asn	Gly	Glu	Asp	Asp	Leu	Asp	Ser	Tyr	Leu
	130					135					140				
Lys	Leu	Ala	Thr	Glu	Gln	Glu	Ser	Lys	Ile	Ser	Glu	Ile	Tyr	Asn	Lys
145				150					155					160	
Leu	Met	Gln	Lys	Ile	Leu	Glu	Phe	Lys	Leu	Met	Lys	Leu	Cys	Thr	Asp
			165					170					175		
Ile	Lys	Asp	Asp	Lys	Val	Ser	Thr	Lys	Thr	Leu	Lys	Glu	Ile	Thr	Ser
		180						185					190		
Tyr	Glu	Phe	Glu	Leu	Leu	Ala	Asp	Tyr	Leu	Ala	Asn	Tyr	Ser	Glu	Ser
	195					200						205			
Leu	Lys	Thr	Gln	Lys	Phe	Ser	Tyr	Thr	Asp	Lys	Gln	Gly	Asn	Leu	Lys
	210					215					220				
Glu	Leu	Ser	Tyr	Tyr	His	His	Asp	Lys	Tyr	Asn	Ile	Gln	Glu	Phe	Leu
225				230					235					240	
Lys	Arg	His	Ala	Thr	Ile	Asn	Asp	Arg	Ile	Leu	Asp	Thr	Leu	Leu	Thr
			245					250					255		
Asp	Asp	Leu	Asp	Ile	Trp	Asn	Phe	Asn	Phe	Glu	Lys	Phe	Asp	Phe	Asp
		260						265					270		
Lys	Asn	Glu	Glu	Lys	Leu	Gln	Asn	Gln	Glu	Asp	Lys	Asp	His	Ile	Gln
		275					280					285			
Ala	His	Leu	His	His	Phe	Val	Phe	Ala	Val	Asn	Lys	Ile	Lys	Ser	Glu
	290					295					300				

Met	Ala	Ser	Gly	Gly	Arg	His	Arg	Ser	Gln	Tyr	Phe	Gln	Glu	Ile	Thr
305						310				315					320
Asn	Val	Leu	Asp	Glu	Asn	Asn	His	Gln	Glu	Gly	Tyr	Leu	Lys	Asn	Phe
				325					330					335	
Cys	Glu	Asn	Leu	His	Asn	Lys	Lys	Tyr	Ser	Asn	Leu	Ser	Val	Lys	Asn
			340					345					350		
Leu	Val	Asn	Leu	Ile	Gly	Asn	Leu	Ser	Asn	Leu	Glu	Leu	Lys	Pro	Leu
		355					360					365			
Arg	Lys	Tyr	Phe	Asn	Asp	Lys	Ile	His	Ala	Lys	Ala	Asp	His	Trp	Asp
370						375					380				
Glu	Gln	Lys	Phe	Thr	Glu	Thr	Tyr	Cys	His	Trp	Ile	Leu	Gly	Glu	Trp
385					390				395						400
Arg	Val	Gly	Val	Lys	Asp	Gln	Asp	Lys	Lys	Asp	Gly	Ala	Lys	Tyr	Ser
				405					410					415	
Tyr	Lys	Asp	Leu	Cys	Asn	Glu	Leu	Lys	Gln	Lys	Val	Thr	Lys	Ala	Gly
		420						425					430		
Leu	Val	Asp	Phe	Leu	Leu	Glu	Leu	Asp	Pro	Cys	Arg	Thr	Ile	Pro	Pro
		435						440				445			
Tyr	Leu	Asp	Asn	Asn	Asn	Arg	Lys	Pro	Pro	Lys	Cys	Gln	Ser	Leu	Ile
450						455					460				
Leu	Asn	Pro	Lys	Phe	Leu	Asp	Asn	Gln	Tyr	Pro	Asn	Trp	Gln	Gln	Tyr
465					470				475						480
Leu	Gln	Glu	Leu	Lys	Lys	Leu	Gln	Ser	Ile	Gln	Asn	Tyr	Leu	Asp	Ser
				485					490					495	
Phe	Glu	Thr	Asp	Leu	Lys	Val	Leu	Lys	Ser	Ser	Lys	Asp	Gln	Pro	Tyr
			500					505					510		
Phe	Val	Glu	Tyr	Lys	Ser	Ser	Asn	Gln	Gln	Ile	Ala	Ser	Gly	Gln	Arg
		515					520					525			
Asp	Tyr	Lys	Asp	Leu	Asp	Ala	Arg	Ile	Leu	Gln	Phe	Ile	Phe	Asp	Arg
530						535					540				
Val	Lys	Ala	Ser	Asp	Glu	Leu	Leu	Leu	Asn	Glu	Ile	Tyr	Phe	Gln	Ala
545					550				555						560
Lys	Lys	Leu	Lys	Gln	Lys	Ala	Ser	Ser	Glu	Leu	Glu	Lys	Leu	Glu	Ser
				565					570					575	
Ser	Lys	Lys	Leu	Asp	Glu	Val	Ile	Ala	Asn	Ser	Gln	Leu	Ser	Gln	Ile
			580					585					590		
Leu	Lys	Ser	Gln	His	Thr	Asn	Gly	Ile	Phe	Glu	Gln	Gly	Thr	Phe	Leu
		595					600					605			
His	Leu	Val	Cys	Lys	Tyr	Tyr	Lys	Gln	Arg	Gln	Arg	Ala	Arg	Asp	Ser

610	615	620
Arg Leu Tyr Ile Met Pro Glu Tyr Arg Tyr Asp Lys Lys Leu His Lys		
625	630	635
Tyr Asn Asn Thr Gly Arg Phe Asp Asp Asp Asn Gln Leu Leu Thr Tyr		
645	650	655
Cys Asn His Lys Pro Arg Gln Lys Arg Tyr Gln Leu Leu Asn Asp Leu		
660	665	670
Ala Gly Val Leu Gln Val Ser Pro Asn Phe Leu Lys Asp Lys Ile Gly		
675	680	685
Ser Asp Asp Asp Leu Phe Ile Ser Lys Trp Leu Val Glu His Ile Arg		
690	695	700
Gly Phe Lys Lys Ala Cys Glu Asp Ser Leu Lys Ile Gln Lys Asp Asn		
705	710	715
Arg Gly Leu Leu Asn His Lys Ile Asn Ile Ala Arg Asn Thr Lys Gly		
725	730	735
Lys Cys Glu Lys Glu Ile Phe Asn Leu Ile Cys Lys Ile Glu Gly Ser		
740	745	750
Glu Asp Lys Lys Gly Asn Tyr Lys His Gly Leu Ala Tyr Glu Leu Gly		
755	760	765
Val Leu Leu Phe Gly Glu Pro Asn Glu Ala Ser Lys Pro Glu Phe Asp		
770	775	780
Arg Lys Ile Lys Lys Phe Asn Ser Ile Tyr Ser Phe Ala Gln Ile Gln		
785	790	795
Gln Ile Ala Phe Ala Glu Arg Lys Gly Asn Ala Asn Thr Cys Ala Val		
805	810	815
Cys Ser Ala Asp Asn Ala His Arg Met Gln Gln Ile Lys Ile Thr Glu		
820	825	830
Pro Val Glu Asp Asn Lys Asp Lys Ile Ile Leu Ser Ala Lys Ala Gln		
835	840	845
Arg Leu Pro Ala Ile Pro Thr Arg Ile Val Asp Gly Ala Val Lys Lys		
850	855	860
Met Ala Thr Ile Leu Ala Lys Asn Ile Val Asp Asp Asn Trp Gln Asn		
865	870	875
Ile Lys Gln Val Leu Ser Ala Lys His Gln Leu His Ile Pro Ile Ile		
885	890	895
Thr Glu Ser Asn Ala Phe Glu Phe Glu Pro Ala Leu Ala Asp Val Lys		
900	905	910
Gly Lys Ser Leu Lys Asp Arg Arg Lys Lys Ala Leu Glu Arg Ile Ser		
915	920	925

Pro Glu Asn Ile Phe Lys Asp Lys Asn Asn Arg Ile Lys Glu Phe Ala		
930	935	940
Lys Gly Ile Ser Ala Tyr Ser Gly Ala Asn Leu Thr Asp Gly Asp Phe		
945	950	955 960
Asp Gly Ala Lys Glu Glu Leu Asp His Ile Ile Pro Arg Ser His Lys		
965	970	975
Lys Tyr Gly Thr Leu Asn Asp Glu Ala Asn Leu Ile Cys Val Thr Arg		
980	985	990
Gly Asp Asn Lys Asn Lys Gly Asn Arg Ile Phe Cys Leu Arg Asp Leu		
995	1000	1005
Ala Asp Asn Tyr Lys Leu Lys Gln Phe Glu Thr Thr Asp Asp Leu		
1010	1015	1020
Glu Ile Glu Lys Lys Ile Ala Asp Thr Ile Trp Asp Ala Asn Lys		
1025	1030	1035
Lys Asp Phe Lys Phe Gly Asn Tyr Arg Ser Phe Ile Asn Leu Thr		
1040	1045	1050
Pro Gln Glu Gln Lys Ala Phe Arg His Ala Leu Phe Leu Ala Asp		
1055	1060	1065
Glu Asn Pro Ile Lys Gln Ala Val Ile Arg Ala Ile Asn Asn Arg		
1070	1075	1080
Asn Arg Thr Phe Val Asn Gly Thr Gln Arg Tyr Phe Ala Glu Val		
1085	1090	1095
Leu Ala Asn Asn Ile Tyr Leu Arg Ala Lys Lys Glu Asn Leu Asn		
1100	1105	1110
Thr Asp Lys Ile Ser Phe Asp Tyr Phe Gly Ile Pro Thr Ile Gly		
1115	1120	1125
Asn Gly Arg Gly Ile Ala Glu Ile Arg Gln Leu Tyr Glu Lys Val		
1130	1135	1140
Asp Ser Asp Ile Gln Ala Tyr Ala Lys Gly Asp Lys Pro Gln Ala		
1145	1150	1155
Ser Tyr Ser His Leu Ile Asp Ala Met Leu Ala Phe Cys Ile Ala		
1160	1165	1170
Ala Asp Glu His Arg Asn Asp Gly Ser Ile Gly Leu Glu Ile Asp		
1175	1180	1185
Lys Asn Tyr Ser Leu Tyr Pro Leu Asp Lys Asn Thr Gly Glu Val		
1190	1195	1200
Phe Thr Lys Asp Ile Phe Ser Gln Ile Lys Ile Thr Asp Asn Glu		
1205	1210	1215
Phe Ser Asp Lys Lys Leu Val Arg Lys Lys Ala Ile Glu Gly Phe		

1220	1225	1230
Asn Thr His Arg Gln Met	Thr Arg Asp Gly Ile	Tyr Ala Glu Asn
1235	1240	1245
Tyr Leu Pro Ile Leu Ile	His Lys Glu Leu Asn	Glu Val Arg Lys
1250	1255	1260
Gly Tyr Thr Trp Lys Asn	Ser Glu Glu Ile Lys	Ile Phe Lys Gly
1265	1270	1275
Lys Lys Tyr Asp Ile Gln	Gln Leu Asn Asn Leu	Val Tyr Cys Leu
1280	1285	1290
Lys Phe Val Asp Lys Pro	Ile Ser Ile Asp Ile	Gln Ile Ser Thr
1295	1300	1305
Leu Glu Glu Leu Arg Asn	Ile Leu Thr Thr Asn	Asn Ile Ala Ala
1310	1315	1320
Thr Ala Glu Tyr Tyr Tyr	Ile Asn Leu Lys Thr	Gln Lys Leu His
1325	1330	1335
Glu Tyr Tyr Ile Glu Asn	Tyr Asn Thr Ala Leu	Gly Tyr Lys Lys
1340	1345	1350
Tyr Ser Lys Glu Met Glu	Phe Leu Arg Ser Leu	Ala Tyr Arg Ser
1355	1360	1365
Glu Arg Val Lys Ile Lys	Ser Ile Asp Asp Val	Lys Gln Val Leu
1370	1375	1380
Asp Lys Asp Ser Asn Phe	Ile Ile Gly Lys Ile	Thr Leu Pro Phe
1385	1390	1395
Lys Lys Glu Trp Gln Arg	Leu Tyr Arg Glu Trp	Gln Asn Thr Thr
1400	1405	1410
Ile Lys Asp Asp Tyr Glu	Phe Leu Lys Ser Phe	Phe Asn Val Lys
1415	1420	1425
Ser Ile Thr Lys Leu His	Lys Lys Val Arg Lys	Asp Phe Ser Leu
1430	1435	1440
Pro Ile Ser Thr Asn Glu	Gly Lys Phe Leu Val	Lys Arg Lys Thr
1445	1450	1455
Trp Asp Asn Asn Phe Ile	Tyr Gln Ile Leu Asn	Asp Ser Asp Ser
1460	1465	1470
Arg Ala Asp Gly Thr Lys	Pro Phe Ile Pro Ala	Phe Asp Ile Ser
1475	1480	1485
Lys Asn Glu Ile Val Glu	Ala Ile Ile Asp Ser	Phe Thr Ser Lys
1490	1495	1500
Asn Ile Phe Trp Leu Pro	Lys Asn Ile Glu Leu	Gln Lys Val Asp
1505	1510	1515

Asn Lys Asn Ile Phe Ala Ile Asp Thr Ser Lys Trp Phe Glu Val
 1520 1525 1530
 Glu Thr Pro Ser Asp Leu Arg Asp Ile Gly Ile Ala Thr Ile Gln
 1535 1540 1545
 Tyr Lys Ile Asp Asn Asn Ser Arg Pro Lys Val Arg Val Lys Leu
 1550 1555 1560
 Asp Tyr Val Ile Asp Asp Asp Ser Lys Ile Asn Tyr Phe Met Asn
 1565 1570 1575
 His Ser Leu Leu Lys Ser Arg Tyr Pro Asp Lys Val Leu Glu Ile
 1580 1585 1590
 Leu Lys Gln Ser Thr Ile Ile Glu Phe Glu Ser Ser Gly Phe Asn
 1595 1600 1605
 Lys Thr Ile Lys Glu Met Leu Gly Met Lys Leu Ala Gly Ile Tyr
 1610 1615 1620

Asn Glu Thr Ser Asn Asn
 1625

<210> 164

<211> 1371

<212> PRT

<213> 布氏乳杆菌(Lactobacillus buchneri)

<400> 164

Met Lys Val Asn Asn Tyr His Ile Gly Leu Asp Ile Gly Thr Ser Ser
 1 5 10 15
 Ile Gly Trp Val Ala Ile Gly Lys Asp Gly Lys Pro Leu Arg Val Lys
 20 25 30
 Gly Lys Thr Ala Ile Gly Ala Arg Leu Phe Gln Glu Gly Asn Pro Ala
 35 40 45
 Ala Asp Arg Arg Met Phe Arg Thr Thr Arg Arg Arg Leu Ser Arg Arg
 50 55 60
 Lys Trp Arg Leu Lys Leu Leu Glu Glu Ile Phe Asp Pro Tyr Ile Thr
 65 70 75 80
 Pro Val Asp Ser Thr Phe Phe Ala Arg Leu Lys Gln Ser Asn Leu Ser
 85 90 95
 Pro Lys Asp Ser Arg Lys Glu Phe Lys Gly Ser Met Leu Phe Pro Asp
 100 105 110
 Leu Thr Asp Met Gln Tyr His Lys Asn Tyr Pro Thr Ile Tyr His Leu
 115 120 125
 Arg His Ala Leu Met Thr Gln Asp Lys Lys Phe Asp Ile Arg Met Val
 130 135 140

Tyr	Leu	Ala	Ile	His	His	Ile	Val	Lys	Tyr	Arg	Gly	Asn	Phe	Leu	Asn
145						150				155					160
Ser	Thr	Pro	Val	Asp	Ser	Phe	Lys	Ala	Ser	Lys	Val	Asp	Phe	Val	Asp
				165						170					175
Gln	Phe	Lys	Lys	Leu	Asn	Glu	Leu	Tyr	Ala	Ala	Ile	Asn	Pro	Glu	Glu
				180						185				190	
Ser	Phe	Lys	Ile	Asn	Leu	Ala	Asn	Ser	Glu	Asp	Ile	Gly	His	Gln	Phe
			195					200						205	
Leu	Asp	Pro	Ser	Ile	Arg	Lys	Phe	Asp	Lys	Lys	Lys	Gln	Ile	Pro	Lys
			210					215						220	
Ile	Val	Pro	Val	Met	Met	Asn	Asp	Lys	Val	Thr	Asp	Arg	Leu	Asn	Gly
225						230						235			240
Lys	Ile	Ala	Ser	Glu	Ile	Ile	His	Ala	Ile	Leu	Gly	Tyr	Lys	Ala	Lys
				245						250					255
Leu	Asp	Val	Val	Leu	Gln	Cys	Thr	Pro	Val	Asp	Ser	Lys	Pro	Trp	Ala
				260						265					270
Leu	Lys	Phe	Asp	Asp	Glu	Asp	Ile	Asp	Ala	Lys	Leu	Glu	Lys	Ile	Leu
			275							280					285
Pro	Glu	Met	Asp	Glu	Asn	Gln	Gln	Ser	Ile	Val	Ala	Ile	Leu	Gln	Asn
			290											300	
Leu	Tyr	Ser	Gln	Val	Thr	Leu	Asn	Gln	Ile	Val	Pro	Asn	Gly	Met	Ser
305						310									320
Leu	Ser	Glu	Ser	Met	Ile	Glu	Lys	Tyr	Asn	Asp	His	His	Asp	His	Leu
				325						330					335
Lys	Leu	Tyr	Lys	Lys	Leu	Ile	Asp	Gln	Leu	Ala	Asp	Pro	Lys	Lys	Lys
				340						345					350
Ala	Val	Leu	Lys	Lys	Ala	Tyr	Ser	Gln	Tyr	Val	Gly	Asp	Asp	Gly	Lys
				355						360					365
Val	Ile	Glu	Gln	Ala	Glu	Phe	Trp	Ser	Ser	Val	Lys	Lys	Asn	Leu	Asp
				370											380
Asp	Ser	Glu	Leu	Ser	Lys	Gln	Ile	Met	Asp	Leu	Ile	Asp	Ala	Glu	Lys
385						390									400
Phe	Met	Pro	Lys	Gln	Arg	Thr	Ser	Gln	Asn	Gly	Val	Ile	Pro	His	Gln
				405						410					415
Leu	His	Gln	Arg	Glu	Leu	Asp	Glu	Ile	Ile	Glu	His	Gln	Ser	Lys	Tyr
				420						425					430
Tyr	Pro	Trp	Leu	Val	Glu	Ile	Asn	Pro	Asn	Lys	His	Asp	Leu	His	Leu
				435						440					445
Ala	Lys	Tyr	Lys	Ile	Glu	Gln	Leu	Val	Ala	Phe	Arg	Val	Pro	Tyr	Tyr

450	455	460
Val Gly Pro Met Ile Thr Pro Lys Asp Gln Ala Glu Ser Ala Glu Thr		
465	470	475
Val Phe Ser Trp Met Glu Arg Lys Gly Thr Glu Thr Gly Gln Ile Thr		
485	490	495
Pro Trp Asn Phe Asp Glu Lys Val Asp Arg Lys Ala Ser Ala Asn Arg		
500	505	510
Phe Ile Lys Arg Met Thr Thr Lys Asp Thr Tyr Leu Ile Gly Glu Asp		
515	520	525
Val Leu Pro Asp Glu Ser Leu Leu Tyr Glu Lys Phe Lys Val Leu Asn		
530	535	540
Glu Leu Asn Met Val Arg Val Asn Gly Lys Leu Leu Lys Val Ala Asp		
545	550	555
Lys Gln Ala Ile Phe Gln Asp Leu Phe Glu Asn Tyr Lys His Val Ser		
565	570	575
Val Lys Lys Leu Gln Asn Tyr Ile Lys Ala Lys Thr Gly Leu Pro Ser		
580	585	590
Asp Pro Glu Ile Ser Gly Leu Ser Asp Pro Glu His Phe Asn Asn Ser		
595	600	605
Leu Gly Thr Tyr Asn Asp Phe Lys Lys Leu Phe Gly Ser Lys Val Asp		
610	615	620
Glu Pro Asp Leu Gln Asp Asp Phe Glu Lys Ile Val Glu Trp Ser Thr		
625	630	635
Val Phe Glu Asp Lys Lys Ile Leu Arg Glu Lys Leu Asn Glu Ile Thr		
645	650	655
Trp Leu Ser Asp Gln Gln Lys Asp Val Leu Glu Ser Ser Arg Tyr Gln		
660	665	670
Gly Trp Gly Arg Leu Ser Lys Lys Leu Leu Thr Gly Ile Val Asn Asp		
675	680	685
Gln Gly Glu Arg Ile Ile Asp Lys Leu Trp Asn Thr Asn Lys Asn Phe		
690	695	700
Met Gln Ile Gln Ser Asp Asp Asp Phe Ala Lys Arg Ile His Glu Ala		
705	710	715
Asn Ala Asp Gln Met Gln Ala Val Asp Val Glu Asp Val Leu Ala Asp		
725	730	735
Ala Tyr Thr Ser Pro Gln Asn Lys Lys Ala Ile Arg Gln Val Val Lys		
740	745	750
Val Val Asp Asp Ile Gln Lys Ala Met Gly Gly Val Ala Pro Lys Tyr		
755	760	765

Ile Ser Ile Glu Phe Thr Arg Ser Glu Asp Arg Asn Pro Arg Arg Thr		
770	775	780
Ile Ser Arg Gln Arg Gln Leu Glu Asn Thr Leu Lys Asp Thr Ala Lys		
785	790	795 800
Ser Leu Ala Lys Ser Ile Asn Pro Glu Leu Leu Ser Glu Leu Asp Asn		
805	810	815
Ala Ala Lys Ser Lys Lys Gly Leu Thr Asp Arg Leu Tyr Leu Tyr Phe		
820	825	830
Thr Gln Leu Gly Lys Asp Ile Tyr Thr Gly Glu Pro Ile Asn Ile Asp		
835	840	845
Glu Leu Asn Lys Tyr Asp Ile Asp His Ile Leu Pro Gln Ala Phe Ile		
850	855	860
Lys Asp Asn Ser Leu Asp Asn Arg Val Leu Val Leu Thr Ala Val Asn		
865	870	875 880
Asn Gly Lys Ser Asp Asn Val Pro Leu Arg Met Phe Gly Ala Lys Met		
885	890	895
Gly His Phe Trp Lys Gln Leu Ala Glu Ala Gly Leu Ile Ser Lys Arg		
900	905	910
Lys Leu Lys Asn Leu Gln Thr Asp Pro Asp Thr Ile Ser Lys Tyr Ala		
915	920	925
Met His Gly Phe Ile Arg Arg Gln Leu Val Glu Thr Ser Gln Val Ile		
930	935	940
Lys Leu Val Ala Asn Ile Leu Gly Asp Lys Tyr Arg Asn Asp Asp Thr		
945	950	955 960
Lys Ile Ile Glu Ile Thr Ala Arg Met Asn His Gln Met Arg Asp Glu		
965	970	975
Phe Gly Phe Ile Lys Asn Arg Glu Ile Asn Asp Tyr His His Ala Phe		
980	985	990
Asp Ala Tyr Leu Thr Ala Phe Leu Gly Arg Tyr Leu Tyr His Arg Tyr		
995	1000	1005
Ile Lys Leu Arg Pro Tyr Phe Val Tyr Gly Asp Phe Lys Lys Phe		
1010	1015	1020
Arg Glu Asp Lys Val Thr Met Arg Asn Phe Asn Phe Leu His Asp		
1025	1030	1035
Leu Thr Asp Asp Thr Gln Glu Lys Ile Ala Asp Ala Glu Thr Gly		
1040	1045	1050
Glu Val Ile Trp Asp Arg Glu Asn Ser Ile Gln Gln Leu Lys Asp		
1055	1060	1065
Val Tyr His Tyr Lys Phe Met Leu Ile Ser His Glu Val Tyr Thr		

1070	1075	1080
Leu Arg Gly Ala Met Phe	Asn Gln Thr Val Tyr	Pro Ala Ser Asp
1085	1090	1095
Ala Gly Lys Arg Lys Leu	Ile Pro Val Lys Ala	Asp Arg Pro Val
1100	1105	1110
Asn Val Tyr Gly Gly Tyr	Ser Gly Ser Ala Asp	Ala Tyr Met Ala
1115	1120	1125
Ile Val Arg Ile His Asn	Lys Lys Gly Asp Lys	Tyr Arg Val Val
1130	1135	1140
Gly Val Pro Met Arg Ala	Leu Asp Arg Leu Asp	Ala Ala Lys Asn
1145	1150	1155
Val Ser Asp Ala Asp Phe	Asp Arg Ala Leu Lys	Asp Val Leu Ala
1160	1165	1170
Pro Gln Leu Thr Lys Thr	Lys Lys Ser Arg Lys	Thr Gly Glu Ile
1175	1180	1185
Thr Gln Val Ile Glu Asp	Phe Glu Ile Val Leu	Gly Lys Val Met
1190	1195	1200
Tyr Arg Gln Leu Met Ile	Asp Gly Asp Lys Lys	Phe Met Leu Gly
1205	1210	1215
Ser Ser Thr Tyr Gln Tyr	Asn Ala Lys Gln Leu	Val Leu Ser Asp
1220	1225	1230
Gln Ser Val Lys Thr Leu	Ala Ser Lys Gly Arg	Leu Asp Pro Leu
1235	1240	1245
Gln Glu Ser Met Asp Tyr	Asn Asn Val Tyr Thr	Glu Ile Leu Asp
1250	1255	1260
Lys Val Asn Gln Tyr Phe	Ser Leu Tyr Asp Met	Asn Lys Phe Arg
1265	1270	1275
His Lys Leu Asn Leu Gly	Phe Ser Lys Phe Ile	Ser Phe Pro Asn
1280	1285	1290
His Asn Val Leu Asp Gly	Asn Thr Lys Val Ser	Ser Gly Lys Arg
1295	1300	1305
Glu Ile Leu Gln Glu Ile	Leu Asn Gly Leu His	Ala Asn Pro Thr
1310	1315	1320
Phe Gly Asn Leu Lys Asp	Val Gly Ile Thr Thr	Pro Phe Gly Gln
1325	1330	1335
Leu Gln Gln Pro Asn Gly	Ile Leu Leu Ser Asp	Glu Thr Lys Ile
1340	1345	1350
Arg Tyr Gln Ser Pro Thr	Gly Leu Phe Glu Arg	Thr Val Ser Leu
1355	1360	1365

Lys Asp Leu

1370

<210> 165

<211> 1334

<212> PRT

<213> 无害李斯特菌(*Listeria innocua*)

<400> 165

Met Lys Lys Pro Tyr Thr Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Leu Thr Asp Gln Tyr Asp Leu Val Lys Arg Lys Met

20 25 30

Lys Ile Ala Gly Asp Ser Glu Lys Lys Gln Ile Lys Lys Asn Phe Trp

35 40 45

Gly Val Arg Leu Phe Asp Glu Gly Gln Thr Ala Ala Asp Arg Arg Met

50 55 60

Ala Arg Thr Ala Arg Arg Arg Ile Glu Arg Arg Arg Asn Arg Ile Ser

65 70 75 80

Tyr Leu Gln Gly Ile Phe Ala Glu Glu Met Ser Lys Thr Asp Ala Asn

85 90 95

Phe Phe Cys Arg Leu Ser Asp Ser Phe Tyr Val Asp Asn Glu Lys Arg

100 105 110

Asn Ser Arg His Pro Phe Phe Ala Thr Ile Glu Glu Glu Val Glu Tyr

115 120 125

His Lys Asn Tyr Pro Thr Ile Tyr His Leu Arg Glu Glu Leu Val Asn

130 135 140

Ser Ser Glu Lys Ala Asp Leu Arg Leu Val Tyr Leu Ala Leu Ala His

145 150 155 160

Ile Ile Lys Tyr Arg Gly Asn Phe Leu Ile Glu Gly Ala Leu Asp Thr

165 170 175

Gln Asn Thr Ser Val Asp Gly Ile Tyr Lys Gln Phe Ile Gln Thr Tyr

180 185 190

Asn Gln Val Phe Ala Ser Gly Ile Glu Asp Gly Ser Leu Lys Lys Leu

195 200 205

Glu Asp Asn Lys Asp Val Ala Lys Ile Leu Val Glu Lys Val Thr Arg

210 215 220

Lys Glu Lys Leu Glu Arg Ile Leu Lys Leu Tyr Pro Gly Glu Lys Ser

225 230 235 240

Ala Gly Met Phe Ala Gln Phe Ile Ser Leu Ile Val Gly Ser Lys Gly

245 250 255

Asn Phe Gln Lys Pro Phe Asp Leu Ile Glu Lys Ser Asp Ile Glu Cys		
260	265	270
Ala Lys Asp Ser Tyr Glu Glu Asp Leu Glu Ser Leu Leu Ala Leu Ile		
275	280	285
Gly Asp Glu Tyr Ala Glu Leu Phe Val Ala Ala Lys Asn Ala Tyr Ser		
290	295	300
Ala Val Val Leu Ser Ser Ile Ile Thr Val Ala Glu Thr Glu Thr Asn		
305	310	315
Ala Lys Leu Ser Ala Ser Met Ile Glu Arg Phe Asp Thr His Glu Glu		
325	330	335
Asp Leu Gly Glu Leu Lys Ala Phe Ile Lys Leu His Leu Pro Lys His		
340	345	350
Tyr Glu Glu Ile Phe Ser Asn Thr Glu Lys His Gly Tyr Ala Gly Tyr		
355	360	365
Ile Asp Gly Lys Thr Lys Gln Ala Asp Phe Tyr Lys Tyr Met Lys Met		
370	375	380
Thr Leu Glu Asn Ile Glu Gly Ala Asp Tyr Phe Ile Ala Lys Ile Glu		
385	390	395
Lys Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ala Ile		
405	410	415
Pro His Gln Leu His Leu Glu Glu Leu Glu Ala Ile Leu His Gln Gln		
420	425	430
Ala Lys Tyr Tyr Pro Phe Leu Lys Glu Asn Tyr Asp Lys Ile Lys Ser		
435	440	445
Leu Val Thr Phe Arg Ile Pro Tyr Phe Val Gly Pro Leu Ala Asn Gly		
450	455	460
Gln Ser Glu Phe Ala Trp Leu Thr Arg Lys Ala Asp Gly Glu Ile Arg		
465	470	475
Pro Trp Asn Ile Glu Glu Lys Val Asp Phe Gly Lys Ser Ala Val Asp		
485	490	495
Phe Ile Glu Lys Met Thr Asn Lys Asp Thr Tyr Leu Pro Lys Glu Asn		
500	505	510
Val Leu Pro Lys His Ser Leu Cys Tyr Gln Lys Tyr Leu Val Tyr Asn		
515	520	525
Glu Leu Thr Lys Val Arg Tyr Ile Asn Asp Gln Gly Lys Thr Ser Tyr		
530	535	540
Phe Ser Gly Gln Glu Lys Glu Gln Ile Phe Asn Asp Leu Phe Lys Gln		
545	550	555
Lys Arg Lys Val Lys Lys Lys Asp Leu Glu Leu Phe Leu Arg Asn Met		

				565					570					575			
Ser	His	Val	Glu	Ser	Pro	Thr	Ile	Glu	Gly	Leu	Glu	Asp	Ser	Phe	Asn		
				580					585					590			
Ser	Ser	Tyr	Ser	Thr	Tyr	His	Asp	Leu	Leu	Lys	Val	Gly	Ile	Lys	Gln		
				595				600						605			
Glu	Ile	Leu	Asp	Asn	Pro	Val	Asn	Thr	Glu	Met	Leu	Glu	Asn	Ile	Val		
				610				615						620			
Lys	Ile	Leu	Thr	Val	Phe	Glu	Asp	Lys	Arg	Met	Ile	Lys	Glu	Gln	Leu		
625					630				635					640			
Gln	Gln	Phe	Ser	Asp	Val	Leu	Asp	Gly	Val	Val	Leu	Lys	Lys	Leu	Glu		
				645					650					655			
Arg	Arg	His	Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Ala	Lys	Leu	Leu	Met		
				660				665						670			
Gly	Ile	Arg	Asp	Lys	Gln	Ser	His	Leu	Thr	Ile	Leu	Asp	Tyr	Leu	Met		
				675				680						685			
Asn	Asp	Asp	Gly	Leu	Asn	Arg	Asn	Leu	Met	Gln	Leu	Ile	Asn	Asp	Ser		
				690				695						700			
Asn	Leu	Ser	Phe	Lys	Ser	Ile	Ile	Glu	Lys	Glu	Gln	Val	Thr	Thr	Ala		
705					710				715					720			
Asp	Lys	Asp	Ile	Gln	Ser	Ile	Val	Ala	Asp	Leu	Ala	Gly	Ser	Pro	Ala		
				725					730					735			
Ile	Lys	Lys	Gly	Ile	Leu	Gln	Ser	Leu	Lys	Ile	Val	Asp	Glu	Leu	Val		
				740				745						750			
Ser	Val	Met	Gly	Tyr	Pro	Pro	Gln	Thr	Ile	Val	Val	Glu	Met	Ala	Arg		
				755				760						765			
Glu	Asn	Gln	Thr	Thr	Gly	Lys	Gly	Lys	Asn	Asn	Ser	Arg	Pro	Arg	Tyr		
				770				775						780			
Lys	Ser	Leu	Glu	Lys	Ala	Ile	Lys	Glu	Phe	Gly	Ser	Gln	Ile	Leu	Lys		
785					790				795					800			
Glu	His	Pro	Thr	Asp	Asn	Gln	Glu	Leu	Arg	Asn	Asn	Arg	Leu	Tyr	Leu		
				805					810					815			
Tyr	Tyr	Leu	Gln	Asn	Gly	Lys	Asp	Met	Tyr	Thr	Gly	Gln	Asp	Leu	Asp		
				820				825						830			
Ile	His	Asn	Leu	Ser	Asn	Tyr	Asp	Ile	Asp	His	Ile	Val	Pro	Gln	Ser		
				835				840						845			
Phe	Ile	Thr	Asp	Asn	Ser	Ile	Asp	Asn	Leu	Val	Leu	Thr	Ser	Ser	Ala		
				850				855						860			
Gly	Asn	Arg	Glu	Lys	Gly	Asp	Asp	Val	Pro	Pro	Leu	Glu	Ile	Val	Arg		
865					870				875					880			

340

1175	1180	1185
Met Leu Ala Ser Ala Asn	Glu Ala Gln Lys Gly	Asn Gln Gln Val
1190	1195	1200
Leu Pro Asn His Leu Val	Thr Leu Leu His His	Ala Ala Asn Cys
1205	1210	1215
Glu Val Ser Asp Gly Lys	Ser Leu Asp Tyr Ile	Glu Ser Asn Arg
1220	1225	1230
Glu Met Phe Ala Glu Leu	Leu Ala His Val Ser	Glu Phe Ala Lys
1235	1240	1245
Arg Tyr Thr Leu Ala Glu	Ala Asn Leu Asn Lys	Ile Asn Gln Leu
1250	1255	1260
Phe Glu Gln Asn Lys Glu	Gly Asp Ile Lys Ala	Ile Ala Gln Ser
1265	1270	1275
Phe Val Asp Leu Met Ala	Phe Asn Ala Met Gly	Ala Pro Ala Ser
1280	1285	1290
Phe Lys Phe Phe Glu Thr	Thr Ile Glu Arg Lys	Arg Tyr Asn Asn
1295	1300	1305
Leu Lys Glu Leu Leu Asn	Ser Thr Ile Ile Tyr	Gln Ser Ile Thr
1310	1315	1320
Gly Leu Tyr Glu Ser Arg	Lys Arg Leu Asp Asp	
1325	1330	
<210> 166		
<211> 1372		
<212> PRT		
<213> 嗜肺军团菌(<i>Legionella pneumophila</i>)		
<400> 166		
Met Glu Ser Ser Gln Ile	Leu Ser Pro Ile Gly	Ile Asp Leu Gly Gly
1	5	10
Lys Phe Thr Gly Val Cys	Leu Ser His Leu Glu	Ala Phe Ala Glu Leu
20	25	30
Pro Asn His Ala Asn Thr	Lys Tyr Ser Val Ile	Leu Ile Asp His Asn
35	40	45
Asn Phe Gln Leu Ser Gln	Ala Gln Arg Arg Ala	Thr Arg His Arg Val
50	55	60
Arg Asn Lys Lys Arg Asn	Gln Phe Val Lys Arg	Val Ala Leu Gln Leu
65	70	75
Phe Gln His Ile Leu Ser	Arg Asp Leu Asn Ala	Lys Glu Glu Thr Ala
85	90	95
Leu Cys His Tyr Leu Asn	Asn Arg Gly Tyr Thr	Tyr Val Asp Thr Asp

100	105	110
Leu Asp Glu Tyr Ile Lys Asp Glu Thr Thr Ile Asn Leu Leu Lys Glu		
115	120	125
Leu Leu Pro Ser Glu Ser Glu His Asn Phe Ile Asp Trp Phe Leu Gln		
130	135	140
Lys Met Gln Ser Ser Glu Phe Arg Lys Ile Leu Val Ser Lys Val Glu		
145	150	155
Glu Lys Lys Asp Asp Lys Glu Leu Lys Asn Ala Val Lys Asn Ile Lys		
165	170	175
Asn Phe Ile Thr Gly Phe Glu Lys Asn Ser Val Glu Gly His Arg His		
180	185	190
Arg Lys Val Tyr Phe Glu Asn Ile Lys Ser Asp Ile Thr Lys Asp Asn		
195	200	205
Gln Leu Asp Ser Ile Lys Lys Lys Ile Pro Ser Val Cys Leu Ser Asn		
210	215	220
Leu Leu Gly His Leu Ser Asn Leu Gln Trp Lys Asn Leu His Arg Tyr		
225	230	235
Leu Ala Lys Asn Pro Lys Gln Phe Asp Glu Gln Thr Phe Gly Asn Glu		
245	250	255
Phe Leu Arg Met Leu Lys Asn Phe Arg His Leu Lys Gly Ser Gln Glu		
260	265	270
Ser Leu Ala Val Arg Asn Leu Ile Gln Gln Leu Glu Gln Ser Gln Asp		
275	280	285
Tyr Ile Ser Ile Leu Glu Lys Thr Pro Pro Glu Ile Thr Ile Pro Pro		
290	295	300
Tyr Glu Ala Arg Thr Asn Thr Gly Met Glu Lys Asp Gln Ser Leu Leu		
305	310	315
Leu Asn Pro Glu Lys Leu Asn Asn Leu Tyr Pro Asn Trp Arg Asn Leu		
325	330	335
Ile Pro Gly Ile Ile Asp Ala His Pro Phe Leu Glu Lys Asp Leu Glu		
340	345	350
His Thr Lys Leu Arg Asp Arg Lys Arg Ile Ile Ser Pro Ser Lys Gln		
355	360	365
Asp Glu Lys Arg Asp Ser Tyr Ile Leu Gln Arg Tyr Leu Asp Leu Asn		
370	375	380
Lys Lys Ile Asp Lys Phe Lys Ile Lys Lys Gln Leu Ser Phe Leu Gly		
385	390	395
Gln Gly Lys Gln Leu Pro Ala Asn Leu Ile Glu Thr Gln Lys Glu Met		
405	410	415

Glu Thr His Phe Asn Ser Ser Leu Val Ser Val Leu Ile Gln Ile Ala		
420	425	430
Ser Ala Tyr Asn Lys Glu Arg Glu Asp Ala Ala Gln Gly Ile Trp Phe		
435	440	445
Asp Asn Ala Phe Ser Leu Cys Glu Leu Ser Asn Ile Asn Pro Pro Arg		
450	455	460
Lys Gln Lys Ile Leu Pro Leu Leu Val Gly Ala Ile Leu Ser Glu Asp		
465	470	475
Phe Ile Asn Asn Lys Asp Lys Trp Ala Lys Phe Lys Ile Phe Trp Asn		
485	490	495
Thr His Lys Ile Gly Arg Thr Ser Leu Lys Ser Lys Cys Lys Glu Ile		
500	505	510
Glu Glu Ala Arg Lys Asn Ser Gly Asn Ala Phe Lys Ile Asp Tyr Glu		
515	520	525
Glu Ala Leu Asn His Pro Glu His Ser Asn Asn Lys Ala Leu Ile Lys		
530	535	540
Ile Ile Gln Thr Ile Pro Asp Ile Ile Gln Ala Ile Gln Ser His Leu		
545	550	555
Gly His Asn Asp Ser Gln Ala Leu Ile Tyr His Asn Pro Phe Ser Leu		
565	570	575
Ser Gln Leu Tyr Thr Ile Leu Glu Thr Lys Arg Asp Gly Phe His Lys		
580	585	590
Asn Cys Val Ala Val Thr Cys Glu Asn Tyr Trp Arg Ser Gln Lys Thr		
595	600	605
Glu Ile Asp Pro Glu Ile Ser Tyr Ala Ser Arg Leu Pro Ala Asp Ser		
610	615	620
Val Arg Pro Phe Asp Gly Val Leu Ala Arg Met Met Gln Arg Leu Ala		
625	630	635
Tyr Glu Ile Ala Met Ala Lys Trp Glu Gln Ile Lys His Ile Pro Asp		
645	650	655
Asn Ser Ser Leu Leu Ile Pro Ile Tyr Leu Glu Gln Asn Arg Phe Glu		
660	665	670
Phe Glu Glu Ser Phe Lys Lys Ile Lys Gly Ser Ser Ser Asp Lys Thr		
675	680	685
Leu Glu Gln Ala Ile Glu Lys Gln Asn Ile Gln Trp Glu Glu Lys Phe		
690	695	700
Gln Arg Ile Ile Asn Ala Ser Met Asn Ile Cys Pro Tyr Lys Gly Ala		
705	710	715
Ser Ile Gly Gly Gln Gly Glu Ile Asp His Ile Tyr Pro Arg Ser Leu		

			725					730					735				
Ser	Lys	Lys	His	Phe	Gly	Val	Ile	Phe	Asn	Ser	Glu	Val	Asn	Leu	Ile		
			740					745					750				
Tyr	Cys	Ser	Ser	Gln	Gly	Asn	Arg	Glu	Lys	Lys	Glu	Glu	His	Tyr	Leu		
			755					760					765				
Leu	Glu	His	Leu	Ser	Pro	Leu	Tyr	Leu	Lys	His	Gln	Phe	Gly	Thr	Asp		
			770					775					780				
Asn	Val	Ser	Asp	Ile	Lys	Asn	Phe	Ile	Ser	Gln	Asn	Val	Ala	Asn	Ile		
785			790					795					800				
Lys	Lys	Tyr	Ile	Ser	Phe	His	Leu	Leu	Thr	Pro	Glu	Gln	Gln	Lys	Ala		
			805					810					815				
Ala	Arg	His	Ala	Leu	Phe	Leu	Asp	Tyr	Asp	Asp	Glu	Ala	Phe	Lys	Thr		
			820					825					830				
Ile	Thr	Lys	Phe	Leu	Met	Ser	Gln	Gln	Lys	Ala	Arg	Val	Asn	Gly	Thr		
			835					840					845				
Gln	Lys	Phe	Leu	Gly	Lys	Gln	Ile	Met	Glu	Phe	Leu	Ser	Thr	Leu	Ala		
850			855					860									
Asp	Ser	Lys	Gln	Leu	Gln	Leu	Glu	Phe	Ser	Ile	Lys	Gln	Ile	Thr	Ala		
865			870					875					880				
Glu	Glu	Val	His	Asp	His	Arg	Glu	Leu	Leu	Ser	Lys	Gln	Glu	Pro	Lys		
			885					890					895				
Leu	Val	Lys	Ser	Arg	Gln	Gln	Ser	Phe	Pro	Ser	His	Ala	Ile	Asp	Ala		
			900					905					910				
Thr	Leu	Thr	Met	Ser	Ile	Gly	Leu	Lys	Glu	Phe	Pro	Gln	Phe	Ser	Gln		
			915					920					925				
Glu	Leu	Asp	Asn	Ser	Trp	Phe	Ile	Asn	His	Leu	Met	Pro	Asp	Glu	Val		
930			935					940									
His	Leu	Asn	Pro	Val	Arg	Ser	Lys	Glu	Lys	Tyr	Asn	Lys	Pro	Asn	Ile		
945			950					955					960				
Ser	Ser	Thr	Pro	Leu	Phe	Lys	Asp	Ser	Leu	Tyr	Ala	Glu	Arg	Phe	Ile		
			965					970					975				
Pro	Val	Trp	Val	Lys	Gly	Glu	Thr	Phe	Ala	Ile	Gly	Phe	Ser	Glu	Lys		
			980					985					990				
Asp	Leu	Phe	Glu	Ile	Lys	Pro	Ser	Asn	Lys	Glu	Lys	Leu	Phe	Thr	Leu		
995			1000					1005									
Leu	Lys	Thr	Tyr	Ser	Thr	Lys	Asn	Pro	Gly	Glu	Ser	Leu	Gln	Glu			
1010			1015					1020									
Leu	Gln	Ala	Lys	Ser	Lys	Ala	Lys	Trp	Leu	Tyr	Phe	Pro	Ile	Asn			
1025			1030					1035									

Lys Thr Leu Ala Leu Glu Phe Leu His His Tyr Phe His Lys Glu		
1040	1045	1050
Ile Val Thr Pro Asp Asp Thr Thr Val Cys His Phe Ile Asn Ser		
1055	1060	1065
Leu Arg Tyr Tyr Thr Lys Lys Glu Ser Ile Thr Val Lys Ile Leu		
1070	1075	1080
Lys Glu Pro Met Pro Val Leu Ser Val Lys Phe Glu Ser Ser Lys		
1085	1090	1095
Lys Asn Val Leu Gly Ser Phe Lys His Thr Ile Ala Leu Pro Ala		
1100	1105	1110
Thr Lys Asp Trp Glu Arg Leu Phe Asn His Pro Asn Phe Leu Ala		
1115	1120	1125
Leu Lys Ala Asn Pro Ala Pro Asn Pro Lys Glu Phe Asn Glu Phe		
1130	1135	1140
Ile Arg Lys Tyr Phe Leu Ser Asp Asn Asn Pro Asn Ser Asp Ile		
1145	1150	1155
Pro Asn Asn Gly His Asn Ile Lys Pro Gln Lys His Lys Ala Val		
1160	1165	1170
Arg Lys Val Phe Ser Leu Pro Val Ile Pro Gly Asn Ala Gly Thr		
1175	1180	1185
Met Met Arg Ile Arg Arg Lys Asp Asn Lys Gly Gln Pro Leu Tyr		
1190	1195	1200
Gln Leu Gln Thr Ile Asp Asp Thr Pro Ser Met Gly Ile Gln Ile		
1205	1210	1215
Asn Glu Asp Arg Leu Val Lys Gln Glu Val Leu Met Asp Ala Tyr		
1220	1225	1230
Lys Thr Arg Asn Leu Ser Thr Ile Asp Gly Ile Asn Asn Ser Glu		
1235	1240	1245
Gly Gln Ala Tyr Ala Thr Phe Asp Asn Trp Leu Thr Leu Pro Val		
1250	1255	1260
Ser Thr Phe Lys Pro Glu Ile Ile Lys Leu Glu Met Lys Pro His		
1265	1270	1275
Ser Lys Thr Arg Arg Tyr Ile Arg Ile Thr Gln Ser Leu Ala Asp		
1280	1285	1290
Phe Ile Lys Thr Ile Asp Glu Ala Leu Met Ile Lys Pro Ser Asp		
1295	1300	1305
Ser Ile Asp Asp Pro Leu Asn Met Pro Asn Glu Ile Val Cys Lys		
1310	1315	1320
Asn Lys Leu Phe Gly Asn Glu Leu Lys Pro Arg Asp Gly Lys Met		

1325	1330	1335
Lys Ile Val Ser Thr Gly	Lys Ile Val Thr Tyr	Glu Phe Glu Ser
1340	1345	1350
Asp Ser Thr Pro Gln Trp	Ile Gln Thr Leu Tyr	Val Thr Gln Leu
1355	1360	1365
Lys Lys Gln Pro		
1370		
<210> 167		
<211> 1082		
<212> PRT		
<213> 嗜乳糖奈瑟球菌(<i>Neisseria lactamica</i>)		
<400> 167		
Met Ala Ala Phe Lys Pro Asn Pro Met Asn Tyr Ile Leu Gly Leu Asp		
1	5	10
Ile Gly Ile Ala Ser Val Gly Trp Ala Met Val Glu Val Asp Glu Glu		
20	25	30
Glu Asn Pro Ile Arg Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg		
35	40	45
Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu		
50	55	60
Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu		
65	70	75
Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Asp Ala Asp		
85	90	95
Phe Asp Glu Asn Gly Leu Val Lys Ser Leu Pro Asn Thr Pro Trp Gln		
100	105	110
Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Cys Leu Glu Trp Ser		
115	120	125
Ala Val Leu Leu His Leu Val Lys His Arg Gly Tyr Leu Ser Gln Arg		
130	135	140
Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys		
145	150	155
Gly Val Ala Asp Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr		
165	170	175
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile		
180	185	190
Arg Asn Gln Arg Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu		
195	200	205
Gln Ala Glu Leu Asn Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn		

210	215	220
Pro His Val Ser Asp Gly Leu Lys Glu Asp Ile Glu Thr Leu Leu Met		
225	230	235
Ala Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly		240
	245	250
His Cys Thr Phe Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr		255
	260	265
Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile		270
	275	280
Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr		285
290	295	300
Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala		
305	310	315
Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg		320
	325	330
Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala		335
	340	345
Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys		350
	355	360
Lys Ser Pro Leu Asn Leu Ser Thr Glu Leu Gln Asp Glu Ile Gly Thr		365
370	375	380
Ala Phe Ser Leu Phe Lys Thr Asp Lys Asp Ile Thr Gly Arg Leu Lys		
385	390	395
Asp Arg Val Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser		400
	405	410
Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val		415
	420	425
Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile		430
	435	440
Tyr Gly Asp His Tyr Cys Lys Lys Asn Ala Glu Glu Lys Ile Tyr Leu		445
	450	455
Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala		460
465	470	475
Leu Ser Gln Ala Arg Lys Val Ile Asn Cys Val Val Arg Arg Tyr Gly		480
	485	490
Ser Pro Ala Arg Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser		495
	500	505
Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys		510
	515	520
		525

Asp	Arg	Glu	Lys	Ala	Ala	Ala	Lys	Phe	Arg	Glu	Tyr	Phe	Pro	Asn	Phe
530							535					540			
Val	Gly	Glu	Pro	Lys	Ser	Lys	Asp	Ile	Leu	Lys	Leu	Arg	Leu	Tyr	Glu
545					550					555					560
Gln	Gln	His	Gly	Lys	Cys	Leu	Tyr	Ser	Gly	Lys	Glu	Ile	Asn	Leu	Val
				565					570					575	
Arg	Leu	Asn	Glu	Lys	Gly	Tyr	Val	Glu	Ile	Asp	His	Ala	Leu	Pro	Phe
				580					585					590	
Ser	Arg	Thr	Trp	Asp	Asp	Ser	Phe	Asn	Asn	Lys	Val	Leu	Val	Leu	Gly
				595			600					605			
Ser	Glu	Asn	Gln	Asn	Lys	Gly	Asn	Gln	Thr	Pro	Tyr	Glu	Tyr	Phe	Asn
						610		615				620			
Gly	Lys	Asp	Asn	Ser	Arg	Glu	Trp	Gln	Glu	Phe	Lys	Ala	Arg	Val	Glu
625					630					635					640
Thr	Ser	Arg	Phe	Pro	Arg	Ser	Lys	Lys	Gln	Arg	Ile	Leu	Leu	Gln	Lys
					645				650					655	
Phe	Asp	Glu	Glu	Gly	Phe	Lys	Glu	Arg	Asn	Leu	Asn	Asp	Thr	Arg	Tyr
				660				665						670	
Val	Asn	Arg	Phe	Leu	Cys	Gln	Phe	Val	Ala	Asp	His	Ile	Leu	Leu	Thr
				675				680						685	
Gly	Lys	Gly	Lys	Arg	Arg	Val	Phe	Ala	Ser	Asn	Gly	Gln	Ile	Thr	Asn
				690			695					700			
Leu	Leu	Arg	Gly	Phe	Trp	Gly	Leu	Arg	Lys	Val	Arg	Thr	Glu	Asn	Asp
705					710					715					720
Arg	His	His	Ala	Leu	Asp	Ala	Val	Val	Val	Ala	Cys	Ser	Thr	Val	Ala
					725					730				735	
Met	Gln	Gln	Lys	Ile	Thr	Arg	Phe	Val	Arg	Tyr	Lys	Glu	Met	Asn	Ala
				740				745						750	
Phe	Asp	Gly	Lys	Thr	Ile	Asp	Lys	Glu	Thr	Gly	Glu	Val	Leu	His	Gln
				755			760					765			
Lys	Ala	His	Phe	Pro	Gln	Pro	Trp	Glu	Phe	Phe	Ala	Gln	Glu	Val	Met
				770			775					780			
Ile	Arg	Val	Phe	Gly	Lys	Pro	Asp	Gly	Lys	Pro	Glu	Phe	Glu	Glu	Ala
785					790					795					800
Asp	Thr	Pro	Glu	Lys	Leu	Arg	Thr	Leu	Leu	Ala	Glu	Lys	Leu	Ser	Ser
					805					810				815	
Arg	Pro	Glu	Ala	Val	His	Glu	Tyr	Val	Thr	Pro	Leu	Phe	Val	Ser	Arg
				820						825				830	
Ala	Pro	Asn	Arg	Lys	Met	Ser	Gly	Gln	Gly	His	Met	Glu	Thr	Val	Lys

835	840	845
Ser Ala Lys Arg Leu Asp Glu Gly Ile Ser Val Leu Arg Val Pro Leu		
850	855	860
Thr Gln Leu Lys Leu Lys Gly Leu Glu Lys Met Val Asn Arg Glu Arg		
865	870	875
Glu Pro Lys Leu Tyr Asp Ala Leu Lys Ala Gln Leu Glu Thr His Lys		
885	890	895
Asp Asp Pro Ala Lys Ala Phe Ala Glu Pro Phe Tyr Lys Tyr Asp Lys		
900	905	910
Ala Gly Ser Arg Thr Gln Gln Val Lys Ala Val Arg Ile Glu Gln Val		
915	920	925
Gln Lys Thr Gly Val Trp Val Arg Asn His Asn Gly Ile Ala Asp Asn		
930	935	940
Ala Thr Met Val Arg Val Asp Val Phe Glu Lys Gly Gly Lys Tyr Tyr		
945	950	955
Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp		
965	970	975
Arg Ala Val Val Ala Phe Lys Asp Glu Glu Asp Trp Thr Val Met Asp		
980	985	990
Asp Ser Phe Glu Phe Arg Phe Val Leu Tyr Ala Asn Asp Leu Ile Lys		
995	1000	1005
Leu Thr Ala Lys Lys Asn Glu Phe Leu Gly Tyr Phe Val Ser Leu		
1010	1015	1020
Asn Arg Ala Thr Gly Ala Ile Asp Ile Arg Thr His Asp Thr Asp		
1025	1030	1035
Ser Thr Lys Gly Lys Asn Gly Ile Phe Gln Ser Val Gly Val Lys		
1040	1045	1050
Thr Ala Leu Ser Phe Gln Lys Asn Gln Ile Asp Glu Leu Gly Lys		
1055	1060	1065
Glu Ile Arg Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg		
1070	1075	1080
<210> 168		
<211> 1082		
<212> PRT		
<213> 脑膜炎奈瑟球菌(<i>Neisseria meningitidis</i>)		
<400> 168		
Met Ala Ala Phe Lys Pro Asn Pro Ile Asn Tyr Ile Leu Gly Leu Asp		
1	5	10
Ile Gly Ile Ala Ser Val Gly Trp Ala Met Val Glu Ile Asp Glu Asp		

20	25	30
Glu Asn Pro Ile Cys Leu Ile Asp Leu Gly Val Arg Val Phe Glu Arg		
35	40	45
Ala Glu Val Pro Lys Thr Gly Asp Ser Leu Ala Met Ala Arg Arg Leu		
50	55	60
Ala Arg Ser Val Arg Arg Leu Thr Arg Arg Arg Ala His Arg Leu Leu		
65	70	75
Arg Ala Arg Arg Leu Leu Lys Arg Glu Gly Val Leu Gln Ala Ala Asp		
85	90	95
Phe Asp Glu Asn Gly Leu Ile Lys Ser Leu Pro Asn Thr Pro Trp Gln		
100	105	110
Leu Arg Ala Ala Ala Leu Asp Arg Lys Leu Thr Pro Leu Glu Trp Ser		
115	120	125
Ala Val Leu Leu His Leu Ile Lys His Arg Gly Tyr Leu Ser Gln Arg		
130	135	140
Lys Asn Glu Gly Glu Thr Ala Asp Lys Glu Leu Gly Ala Leu Leu Lys		
145	150	155
Gly Val Ala Asp Asn Ala His Ala Leu Gln Thr Gly Asp Phe Arg Thr		
165	170	175
Pro Ala Glu Leu Ala Leu Asn Lys Phe Glu Lys Glu Ser Gly His Ile		
180	185	190
Arg Asn Gln Arg Gly Asp Tyr Ser His Thr Phe Ser Arg Lys Asp Leu		
195	200	205
Gln Ala Glu Leu Ile Leu Leu Phe Glu Lys Gln Lys Glu Phe Gly Asn		
210	215	220
Pro His Val Ser Gly Gly Leu Lys Glu Gly Ile Glu Thr Leu Leu Met		
225	230	235
Thr Gln Arg Pro Ala Leu Ser Gly Asp Ala Val Gln Lys Met Leu Gly		
245	250	255
His Cys Thr Phe Glu Pro Ala Glu Pro Lys Ala Ala Lys Asn Thr Tyr		
260	265	270
Thr Ala Glu Arg Phe Ile Trp Leu Thr Lys Leu Asn Asn Leu Arg Ile		
275	280	285
Leu Glu Gln Gly Ser Glu Arg Pro Leu Thr Asp Thr Glu Arg Ala Thr		
290	295	300
Leu Met Asp Glu Pro Tyr Arg Lys Ser Lys Leu Thr Tyr Ala Gln Ala		
305	310	315
Arg Lys Leu Leu Gly Leu Glu Asp Thr Ala Phe Phe Lys Gly Leu Arg		
325	330	335

Tyr Gly Lys Asp Asn Ala Glu Ala Ser Thr Leu Met Glu Met Lys Ala		
340	345	350
Tyr His Ala Ile Ser Arg Ala Leu Glu Lys Glu Gly Leu Lys Asp Lys		
355	360	365
Lys Ser Pro Leu Asn Leu Ser Pro Glu Leu Gln Asp Glu Ile Gly Thr		
370	375	380
Ala Phe Ser Leu Phe Lys Thr Asp Glu Asp Ile Thr Gly Arg Leu Lys		
385	390	395
Asp Arg Ile Gln Pro Glu Ile Leu Glu Ala Leu Leu Lys His Ile Ser		
405	410	415
Phe Asp Lys Phe Val Gln Ile Ser Leu Lys Ala Leu Arg Arg Ile Val		
420	425	430
Pro Leu Met Glu Gln Gly Lys Arg Tyr Asp Glu Ala Cys Ala Glu Ile		
435	440	445
Tyr Gly Asp His Tyr Gly Lys Lys Asn Thr Glu Glu Lys Ile Tyr Leu		
450	455	460
Pro Pro Ile Pro Ala Asp Glu Ile Arg Asn Pro Val Val Leu Arg Ala		
465	470	475
Leu Ser Gln Ala Arg Lys Val Ile Asn Gly Val Val Arg Arg Tyr Gly		
485	490	495
Ser Pro Ala Arg Ile His Ile Glu Thr Ala Arg Glu Val Gly Lys Ser		
500	505	510
Phe Lys Asp Arg Lys Glu Ile Glu Lys Arg Gln Glu Glu Asn Arg Lys		
515	520	525
Asp Arg Glu Lys Ala Ala Ala Lys Phe Arg Glu Tyr Phe Pro Asn Phe		
530	535	540
Val Gly Glu Pro Lys Ser Lys Asp Ile Leu Lys Leu Arg Leu Tyr Glu		
545	550	555
Gln Gln His Gly Lys Cys Leu Tyr Ser Gly Lys Glu Ile Asn Leu Gly		
565	570	575
Arg Leu Asn Glu Lys Gly Tyr Val Glu Ile Asp His Ala Leu Pro Phe		
580	585	590
Ser Arg Thr Trp Asp Asp Ser Phe Asn Asn Lys Val Leu Val Leu Gly		
595	600	605
Ser Glu Asn Gln Asn Lys Gly Asn Gln Thr Pro Tyr Glu Tyr Phe Asn		
610	615	620
Gly Lys Asp Asn Ser Arg Glu Trp Gln Glu Phe Lys Ala Arg Val Glu		
625	630	635
Thr Ser Arg Phe Pro Arg Ser Lys Lys Gln Arg Ile Leu Leu Gln Lys		

				645					650					655			
Phe	Asp	Glu	Asp	Gly	Phe	Lys	Glu	Arg	Asn	Leu	Asn	Asp	Thr	Arg	Tyr		
				660					665					670			
Val	Asn	Arg	Phe	Leu	Cys	Gln	Phe	Val	Ala	Asp	Arg	Met	Arg	Leu	Thr		
				675					680					685			
Gly	Lys	Gly	Lys	Lys	Arg	Val	Phe	Ala	Ser	Asn	Gly	Gln	Ile	Thr	Asn		
				690					695					700			
Leu	Leu	Arg	Gly	Phe	Trp	Gly	Leu	Arg	Lys	Val	Arg	Ala	Glu	Asn	Asp		
705						710				715					720		
Arg	His	His	Ala	Leu	Asp	Ala	Val	Val	Val	Ala	Cys	Ser	Thr	Val	Ala		
				725					730					735			
Met	Gln	Gln	Lys	Ile	Thr	Arg	Phe	Val	Arg	Tyr	Lys	Glu	Met	Asn	Ala		
				740					745					750			
Phe	Asp	Gly	Lys	Thr	Ile	Asp	Lys	Glu	Thr	Gly	Glu	Val	Leu	His	Gln		
				755					760					765			
Lys	Thr	His	Phe	Pro	Gln	Pro	Trp	Glu	Phe	Phe	Ala	Gln	Glu	Val	Met		
				770					775					780			
Ile	Arg	Val	Phe	Gly	Lys	Pro	Asp	Gly	Lys	Pro	Glu	Phe	Glu	Glu	Ala		
785						790				795					800		
Asp	Thr	Pro	Glu	Lys	Leu	Arg	Thr	Leu	Leu	Ala	Glu	Lys	Leu	Ser	Ser		
				805					810					815			
Arg	Pro	Glu	Ala	Val	His	Glu	Tyr	Val	Thr	Pro	Leu	Phe	Val	Ser	Arg		
				820					825					830			
Ala	Pro	Asn	Arg	Lys	Met	Ser	Gly	Gln	Gly	His	Met	Glu	Thr	Val	Lys		
				835					840					845			
Ser	Ala	Lys	Arg	Leu	Asp	Glu	Gly	Val	Ser	Val	Leu	Arg	Val	Pro	Leu		
				850					855					860			
Thr	Gln	Leu	Lys	Leu	Lys	Asp	Leu	Glu	Lys	Met	Val	Asn	Arg	Glu	Arg		
865						870				875					880		
Glu	Pro	Lys	Leu	Tyr	Glu	Ala	Leu	Lys	Ala	Arg	Leu	Glu	Ala	His	Lys		
				885					890					895			
Asp	Asp	Pro	Ala	Lys	Ala	Phe	Ala	Glu	Pro	Phe	Tyr	Lys	Tyr	Asp	Lys		
				900					905					910			
Ala	Gly	Asn	Arg	Thr	Gln	Gln	Val	Lys	Ala	Val	Arg	Val	Glu	Gln	Val		
				915					920					925			
Gln	Lys	Thr	Gly	Val	Trp	Val	Arg	Asn	His	Asn	Gly	Ile	Ala	Asp	Asn		
				930					935					940			
Ala	Thr	Met	Val	Arg	Val	Asp	Val	Phe	Glu	Lys	Gly	Asp	Lys	Tyr	Tyr		
945						950				955					960		

Leu Val Pro Ile Tyr Ser Trp Gln Val Ala Lys Gly Ile Leu Pro Asp		
965	970	975
Arg Ala Val Val Gln Gly Lys Asp Glu Glu Asp Trp Gln Leu Ile Asp		
980	985	990
Asp Ser Phe Asn Phe Lys Phe Ser Leu His Pro Asn Asp Leu Val Glu		
995	1000	1005
Val Ile Thr Lys Lys Ala Arg Met Phe Gly Tyr Phe Ala Ser Cys		
1010	1015	1020
His Arg Gly Thr Gly Asn Ile Asn Ile Arg Ile His Asp Leu Asp		
1025	1030	1035
His Lys Ile Gly Lys Asn Gly Ile Leu Glu Gly Ile Gly Val Lys		
1040	1045	1050
Thr Ala Leu Ser Phe Gln Lys Tyr Gln Ile Asp Glu Leu Gly Lys		
1055	1060	1065
Glu Ile Arg Pro Cys Arg Leu Lys Lys Arg Pro Pro Val Arg		
1070	1075	1080
<210> 169		
<211> 1187		
<212> PRT		
<213> 长双歧杆菌(Bifidobacterium longum)		
<400> 169		
Met Leu Ser Arg Gln Leu Leu Gly Ala Ser His Leu Ala Arg Pro Val		
1	5	10
Ser Tyr Ser Tyr Asn Val Gln Asp Asn Asp Val His Cys Ser Tyr Gly		
20	25	30
Glu Arg Cys Phe Met Arg Gly Lys Arg Tyr Arg Ile Gly Ile Asp Val		
35	40	45
Gly Leu Asn Ser Val Gly Leu Ala Ala Val Glu Val Ser Asp Glu Asn		
50	55	60
Ser Pro Val Arg Leu Leu Asn Ala Gln Ser Val Ile His Asp Gly Gly		
65	70	75
Val Asp Pro Gln Lys Asn Lys Glu Ala Ile Thr Arg Lys Asn Met Ser		
85	90	95
Gly Val Ala Arg Arg Thr Arg Arg Met Arg Arg Arg Lys Arg Glu Arg		
100	105	110
Leu His Lys Leu Asp Met Leu Leu Gly Lys Phe Gly Tyr Pro Val Ile		
115	120	125
Glu Pro Glu Ser Leu Asp Lys Pro Phe Glu Glu Trp His Val Arg Ala		
130	135	140

Glu	Leu	Ala	Thr	Arg	Tyr	Ile	Glu	Asp	Asp	Glu	Leu	Arg	Arg	Glu	Ser
145						150				155					160
Ile	Ser	Ile	Ala	Leu	Arg	His	Met	Ala	Arg	His	Arg	Gly	Trp	Arg	Asn
				165						170					175
Pro	Tyr	Arg	Gln	Val	Asp	Ser	Leu	Ile	Ser	Asp	Asn	Pro	Tyr	Ser	Lys
			180							185					190
Gln	Tyr	Gly	Glu	Leu	Lys	Glu	Lys	Ala	Lys	Ala	Tyr	Asn	Asp	Asp	Ala
			195							200					205
Thr	Ala	Ala	Glu	Glu	Glu	Ser	Thr	Pro	Ala	Gln	Leu	Val	Val	Ala	Met
			210							215					220
Leu	Asp	Ala	Gly	Tyr	Ala	Glu	Ala	Pro	Arg	Leu	Arg	Trp	Arg	Thr	Gly
225						230				235					240
Ser	Lys	Lys	Pro	Asp	Ala	Glu	Gly	Tyr	Leu	Pro	Val	Arg	Leu	Met	Gln
				245						250					255
Glu	Asp	Asn	Ala	Asn	Glu	Leu	Lys	Gln	Ile	Phe	Arg	Val	Gln	Arg	Val
			260							265					270
Pro	Ala	Asp	Glu	Trp	Lys	Pro	Leu	Phe	Arg	Ser	Val	Phe	Tyr	Ala	Val
			275							280					285
Ser	Pro	Lys	Gly	Ser	Ala	Glu	Gln	Arg	Val	Gly	Gln	Asp	Pro	Leu	Ala
			290							295					300
Pro	Glu	Gln	Ala	Arg	Ala	Leu	Lys	Ala	Ser	Leu	Ala	Phe	Gln	Glu	Tyr
305						310				315					320
Arg	Ile	Ala	Asn	Val	Ile	Thr	Asn	Leu	Arg	Ile	Lys	Asp	Ala	Ser	Ala
				325						330					335
Glu	Leu	Arg	Lys	Leu	Thr	Val	Asp	Glu	Lys	Gln	Ser	Ile	Tyr	Asp	Gln
			340							345					350
Leu	Val	Ser	Pro	Ser	Ser	Glu	Asp	Ile	Thr	Trp	Ser	Asp	Leu	Cys	Asp
			355							360					365
Phe	Leu	Gly	Phe	Lys	Arg	Ser	Gln	Leu	Lys	Gly	Val	Gly	Ser	Leu	Thr
			370							375					380
Glu	Asp	Gly	Glu	Glu	Arg	Ile	Ser	Ser	Arg	Pro	Pro	Arg	Leu	Thr	Ser
385						390				395					400
Val	Gln	Arg	Ile	Tyr	Glu	Ser	Asp	Asn	Lys	Ile	Arg	Lys	Pro	Leu	Val
				405						410					415
Ala	Trp	Trp	Lys	Ser	Ala	Ser	Asp	Asn	Glu	His	Glu	Ala	Met	Ile	Arg
			420							425					430
Leu	Leu	Ser	Asn	Thr	Val	Asp	Ile	Asp	Lys	Val	Arg	Glu	Asp	Val	Ala
			435							440					445
Tyr	Ala	Ser	Ala	Ile	Glu	Phe	Ile	Asp	Gly	Leu	Asp	Asp	Asp	Ala	Leu

450	455	460
Thr Lys Leu Asp Ser Val	Asp Leu Pro Ser Gly Arg Ala Ala Tyr Ser	
465	470	475
Val Glu Thr Leu Gln Lys	Leu Thr Arg Gln Met Leu Thr Thr Asp Asp	480
	485	490
Asp Leu His Glu Ala Arg	Lys Thr Leu Phe Asn Val Thr Asp Ser Trp	495
	500	505
Arg Pro Pro Ala Asp Pro	Ile Gly Glu Pro Leu Gly Asn Pro Ser Val	510
	515	520
Asp Arg Val Leu Lys Asn	Val Asn Arg Tyr Leu Met Asn Cys Gln Gln	525
	530	535
Arg Trp Gly Asn Pro Val	Ser Val Asn Ile Glu His Val Arg Ser Ser	540
545	550	555
Phe Ser Ser Val Ala Phe	Ala Arg Lys Asp Lys Arg Glu Tyr Glu Lys	560
	565	570
Asn Asn Glu Lys Arg Ser	Ile Phe Arg Ser Ser Leu Ser Glu Gln Leu	575
	580	585
Arg Ala Asp Glu Gln Met	Glu Lys Val Arg Glu Ser Asp Leu Arg Arg	590
	595	600
Leu Glu Ala Ile Gln Arg	Gln Asn Gly Gln Cys Leu Tyr Cys Gly Arg	605
	610	615
Thr Ile Thr Phe Arg Thr	Cys Glu Met Asp His Ile Val Pro Arg Lys	620
625	630	635
Gly Val Gly Ser Thr Asn	Thr Arg Thr Asn Phe Ala Ala Val Cys Ala	640
	645	650
Glu Cys Asn Arg Met Lys	Ser Asn Thr Pro Phe Ala Ile Trp Ala Arg	655
	660	665
Ser Glu Asp Ala Gln Thr	Arg Gly Val Ser Leu Ala Glu Ala Lys Lys	670
	675	680
Arg Val Thr Met Phe Thr	Phe Asn Pro Lys Ser Tyr Ala Pro Arg Glu	685
	690	695
Val Lys Ala Phe Lys Gln	Ala Val Ile Ala Arg Leu Gln Gln Thr Glu	700
705	710	715
Asp Asp Ala Ala Ile Asp	Asn Arg Ser Ile Glu Ser Val Ala Trp Met	720
	725	730
Ala Asp Glu Leu His Arg	Arg Ile Asp Trp Tyr Phe Asn Ala Lys Gln	735
	740	745
Tyr Val Asn Ser Ala Ser	Ile Asp Asp Ala Glu Ala Glu Thr Met Lys	750
	755	760
		765

Thr	Thr	Val	Ser	Val	Phe	Gln	Gly	Arg	Val	Thr	Ala	Ser	Ala	Arg	Arg
770						775					780				
Ala	Ala	Gly	Ile	Glu	Gly	Lys	Ile	His	Phe	Ile	Gly	Gln	Gln	Ser	Lys
785					790					795					800
Thr	Arg	Leu	Asp	Arg	Arg	His	His	Ala	Val	Asp	Ala	Ser	Val	Ile	Ala
				805					810					815	
Met	Met	Asn	Thr	Ala	Ala	Ala	Gln	Thr	Leu	Met	Glu	Arg	Glu	Ser	Leu
			820					825					830		
Arg	Glu	Ser	Gln	Arg	Leu	Ile	Gly	Leu	Met	Pro	Gly	Glu	Arg	Ser	Trp
		835					840					845			
Lys	Glu	Tyr	Pro	Tyr	Glu	Gly	Thr	Ser	Arg	Tyr	Glu	Ser	Phe	His	Leu
	850					855					860				
Trp	Leu	Asp	Asn	Met	Asp	Val	Leu	Leu	Glu	Leu	Leu	Asn	Asp	Ala	Leu
865				870					875						880
Asp	Asn	Asp	Arg	Ile	Ala	Val	Met	Gln	Ser	Gln	Arg	Tyr	Val	Leu	Gly
				885					890					895	
Asn	Ser	Ile	Ala	His	Asp	Ala	Thr	Ile	His	Pro	Leu	Glu	Lys	Val	Pro
			900					905						910	
Leu	Gly	Ser	Ala	Met	Ser	Ala	Asp	Leu	Ile	Arg	Arg	Ala	Ser	Thr	Pro
		915						920					925		
Ala	Leu	Trp	Cys	Ala	Leu	Thr	Arg	Leu	Pro	Asp	Tyr	Asp	Glu	Lys	Glu
	930					935					940				
Gly	Leu	Pro	Glu	Asp	Ser	His	Arg	Glu	Ile	Arg	Val	His	Asp	Thr	Arg
945				950						955					960
Tyr	Ser	Ala	Asp	Asp	Glu	Met	Gly	Phe	Phe	Ala	Ser	Gln	Ala	Ala	Gln
				965					970					975	
Ile	Ala	Val	Gln	Glu	Gly	Ser	Ala	Asp	Ile	Gly	Ser	Ala	Ile	His	His
			980					985					990		
Ala	Arg	Val	Tyr	Arg	Cys	Trp	Lys	Thr	Asn	Ala	Lys	Gly	Val	Arg	Lys
	995						1000					1005			
Tyr	Phe	Tyr	Gly	Met	Ile	Arg	Val	Phe	Gln	Thr	Asp	Leu	Leu	Arg	
	1010					1015					1020				
Ala	Cys	His	Asp	Asp	Leu	Phe	Thr	Val	Pro	Leu	Pro	Pro	Gln	Ser	
	1025					1030					1035				
Ile	Ser	Met	Arg	Tyr	Gly	Glu	Pro	Arg	Val	Val	Gln	Ala	Leu	Gln	
	1040					1045					1050				
Ser	Gly	Asn	Ala	Gln	Tyr	Leu	Gly	Ser	Leu	Val	Val	Gly	Asp	Glu	
	1055					1060					1065				
Ile	Glu	Met	Asp	Phe	Ser	Ser	Leu	Asp	Val	Asp	Gly	Gln	Ile	Gly	

1070	1075	1080
Glu Tyr Leu Gln Phe Phe	Ser Gln Phe Ser Gly	Gly Asn Leu Ala
1085	1090	1095
Trp Lys His Trp Val Val	Asp Gly Phe Phe Asn	Gln Thr Gln Leu
1100	1105	1110
Arg Ile Arg Pro Arg Tyr	Leu Ala Ala Glu Gly	Leu Ala Lys Ala
1115	1120	1125
Phe Ser Asp Asp Val Val	Pro Asp Gly Val Gln	Lys Ile Val Thr
1130	1135	1140
Lys Gln Gly Trp Leu Pro	Pro Val Asn Thr Ala	Ser Lys Thr Ala
1145	1150	1155
Val Arg Ile Val Arg Arg	Asn Ala Phe Gly Glu	Pro Arg Leu Ser
1160	1165	1170
Ser Ala His His Met Pro	Cys Ser Trp Gln Trp	Arg His Glu
1175	1180	1185
<210> 170		
<211> 1101		
<212> PRT		
<213> 嗜粘蛋白艾克曼菌(Akkermansia muciniphila)		
<400> 170		
Met Ser Arg Ser Leu Thr	Phe Ser Phe Asp Ile	Gly Tyr Ala Ser Ile
1	5	10
Gly Trp Ala Val Ile Ala	Ser Ala Ser His Asp	Asp Ala Asp Pro Ser
20	25	30
Val Cys Gly Cys Gly Thr	Val Leu Phe Pro Lys	Asp Asp Cys Gln Ala
35	40	45
Phe Lys Arg Arg Glu Tyr	Arg Arg Leu Arg Arg	Asn Ile Arg Ser Arg
50	55	60
Arg Val Arg Ile Glu Arg	Ile Gly Arg Leu Leu	Val Gln Ala Gln Ile
65	70	75
Ile Thr Pro Glu Met Lys	Glu Thr Ser Gly His	Pro Ala Pro Phe Tyr
85	90	95
Leu Ala Ser Glu Ala Leu	Lys Gly His Arg Thr	Leu Ala Pro Ile Glu
100	105	110
Leu Trp His Val Leu Arg	Trp Tyr Ala His Asn	Arg Gly Tyr Asp Asn
115	120	125
Asn Ala Ser Trp Ser Asn	Ser Leu Ser Glu Asp	Gly Gly Asn Gly Glu
130	135	140
Asp Thr Glu Arg Val Lys	His Ala Gln Asp Leu	Met Asp Lys His Gly

145	150	155	160
Thr Ala Thr Met Ala Glu Thr Ile Cys Arg Glu Leu Lys Leu Glu Glu			
	165	170	175
Gly Lys Ala Asp Ala Pro Met Glu Val Ser Thr Pro Ala Tyr Lys Asn			
	180	185	190
Leu Asn Thr Ala Phe Pro Arg Leu Ile Val Glu Lys Glu Val Arg Arg			
	195	200	205
Ile Leu Glu Leu Ser Ala Pro Leu Ile Pro Gly Leu Thr Ala Glu Ile			
	210	215	220
Ile Glu Leu Ile Ala Gln His His Pro Leu Thr Thr Glu Gln Arg Gly			
225	230	235	240
Val Leu Leu Gln His Gly Ile Lys Leu Ala Arg Arg Tyr Arg Gly Ser			
	245	250	255
Leu Leu Phe Gly Gln Leu Ile Pro Arg Phe Asp Asn Arg Ile Ile Ser			
	260	265	270
Arg Cys Pro Val Thr Trp Ala Gln Val Tyr Glu Ala Glu Leu Lys Lys			
	275	280	285
Gly Asn Ser Glu Gln Ser Ala Arg Glu Arg Ala Glu Lys Leu Ser Lys			
	290	295	300
Val Pro Thr Ala Asn Cys Pro Glu Phe Tyr Glu Tyr Arg Met Ala Arg			
305	310	315	320
Ile Leu Cys Asn Ile Arg Ala Asp Gly Glu Pro Leu Ser Ala Glu Ile			
	325	330	335
Arg Arg Glu Leu Met Asn Gln Ala Arg Gln Glu Gly Lys Leu Thr Lys			
	340	345	350
Ala Ser Leu Glu Lys Ala Ile Ser Ser Arg Leu Gly Lys Glu Thr Glu			
	355	360	365
Thr Asn Val Ser Asn Tyr Phe Thr Leu His Pro Asp Ser Glu Glu Ala			
	370	375	380
Leu Tyr Leu Asn Pro Ala Val Glu Val Leu Gln Arg Ser Gly Ile Gly			
385	390	395	400
Gln Ile Leu Ser Pro Ser Val Tyr Arg Ile Ala Ala Asn Arg Leu Arg			
	405	410	415
Arg Gly Lys Ser Val Thr Pro Asn Tyr Leu Leu Asn Leu Leu Lys Ser			
	420	425	430
Arg Gly Glu Ser Gly Glu Ala Leu Glu Lys Lys Ile Glu Lys Glu Ser			
	435	440	445
Lys Lys Lys Glu Ala Asp Tyr Ala Asp Thr Pro Leu Lys Pro Lys Tyr			
	450	455	460

Ala Thr Gly Arg Ala Pro Tyr Ala Arg Thr Val Leu Lys Lys Val Val			
465	470	475	480
Glu Glu Ile Leu Asp Gly Glu Asp Pro Thr Arg Pro Ala Arg Gly Glu			
	485	490	495
Ala His Pro Asp Gly Glu Leu Lys Ala His Asp Gly Cys Leu Tyr Cys			
	500	505	510
Leu Leu Asp Thr Asp Ser Ser Val Asn Gln His Gln Lys Glu Arg Arg			
	515	520	525
Leu Asp Thr Met Thr Asn Asn His Leu Val Arg His Arg Met Leu Ile			
	530	535	540
Leu Asp Arg Leu Leu Lys Asp Leu Ile Gln Asp Phe Ala Asp Gly Gln			
545	550	555	560
Lys Asp Arg Ile Ser Arg Val Cys Val Glu Val Gly Lys Glu Leu Thr			
	565	570	575
Thr Phe Ser Ala Met Asp Ser Lys Lys Ile Gln Arg Glu Leu Thr Leu			
	580	585	590
Arg Gln Lys Ser His Thr Asp Ala Val Asn Arg Leu Lys Arg Lys Leu			
	595	600	605
Pro Gly Lys Ala Leu Ser Ala Asn Leu Ile Arg Lys Cys Arg Ile Ala			
	610	615	620
Met Asp Met Asn Trp Thr Cys Pro Phe Thr Gly Ala Thr Tyr Gly Asp			
625	630	635	640
His Glu Leu Glu Asn Leu Glu Leu Glu His Ile Val Pro His Ser Phe			
	645	650	655
Arg Gln Ser Asn Ala Leu Ser Ser Leu Val Leu Thr Trp Pro Gly Val			
	660	665	670
Asn Arg Met Lys Gly Gln Arg Thr Gly Tyr Asp Phe Val Glu Gln Glu			
	675	680	685
Gln Glu Asn Pro Val Pro Asp Lys Pro Asn Leu His Ile Cys Ser Leu			
	690	695	700
Asn Asn Tyr Arg Glu Leu Val Glu Lys Leu Asp Asp Lys Lys Gly His			
705	710	715	720
Glu Asp Asp Arg Arg Arg Lys Lys Lys Arg Lys Ala Leu Leu Met Val			
	725	730	735
Arg Gly Leu Ser His Lys His Gln Ser Gln Asn His Glu Ala Met Lys			
	740	745	750
Glu Ile Gly Met Thr Glu Gly Met Met Thr Gln Ser Ser His Leu Met			
	755	760	765
Lys Leu Ala Cys Lys Ser Ile Lys Thr Ser Leu Pro Asp Ala His Ile			

770	775	780
Asp Met Ile Pro Gly Ala Val Thr Ala Glu Val Arg Lys Ala Trp Asp		
785	790	795
Val Phe Gly Val Phe Lys Glu Leu Cys Pro Glu Ala Ala Asp Pro Asp		
805	810	815
Ser Gly Lys Ile Leu Lys Glu Asn Leu Arg Ser Leu Thr His Leu His		
820	825	830
His Ala Leu Asp Ala Cys Val Leu Gly Leu Ile Pro Tyr Ile Ile Pro		
835	840	845
Ala His His Asn Gly Leu Leu Arg Arg Val Leu Ala Met Arg Arg Ile		
850	855	860
Pro Glu Lys Leu Ile Pro Gln Val Arg Pro Val Ala Asn Gln Arg His		
865	870	875
Tyr Val Leu Asn Asp Asp Gly Arg Met Met Leu Arg Asp Leu Ser Ala		
885	890	895
Ser Leu Lys Glu Asn Ile Arg Glu Gln Leu Met Glu Gln Arg Val Ile		
900	905	910
Gln His Val Pro Ala Asp Met Gly Gly Ala Leu Leu Lys Glu Thr Met		
915	920	925
Gln Arg Val Leu Ser Val Asp Gly Ser Gly Glu Asp Ala Met Val Ser		
930	935	940
Leu Ser Lys Lys Lys Asp Gly Lys Lys Glu Lys Asn Gln Val Lys Ala		
945	950	955
Ser Lys Leu Val Gly Val Phe Pro Glu Gly Pro Ser Lys Leu Lys Ala		
965	970	975
Leu Lys Ala Ala Ile Glu Ile Asp Gly Asn Tyr Gly Val Ala Leu Asp		
980	985	990
Pro Lys Pro Val Val Ile Arg His Ile Lys Val Phe Lys Arg Ile Met		
995	1000	1005
Ala Leu Lys Glu Gln Asn Gly Gly Lys Pro Val Arg Ile Leu Lys		
1010	1015	1020
Lys Gly Met Leu Ile His Leu Thr Ser Ser Lys Asp Pro Lys His		
1025	1030	1035
Ala Gly Val Trp Arg Ile Glu Ser Ile Gln Asp Ser Lys Gly Gly		
1040	1045	1050
Val Lys Leu Asp Leu Gln Arg Ala His Cys Ala Val Pro Lys Asn		
1055	1060	1065
Lys Thr His Glu Cys Asn Trp Arg Glu Val Asp Leu Ile Ser Leu		
1070	1075	1080

Leu Lys Lys Tyr Gln Met Lys Arg Tyr Pro Thr Ser Tyr Thr Gly
 1085 1090 1095
 Thr Pro Arg
 1100
 <210> 171
 <211> 1498
 <212> PRT
 <213> 兰氏臭杆菌 (*Odoribacter laneus*)
 <400> 171
 Met Glu Thr Thr Leu Gly Ile Asp Leu Gly Thr Asn Ser Ile Gly Leu
 1 5 10 15
 Ala Leu Val Asp Gln Glu Glu His Gln Ile Leu Tyr Ser Gly Val Arg
 20 25 30
 Ile Phe Pro Glu Gly Ile Asn Lys Asp Thr Ile Gly Leu Gly Glu Lys
 35 40 45
 Glu Glu Ser Arg Asn Ala Thr Arg Arg Ala Lys Arg Gln Met Arg Arg
 50 55 60
 Gln Tyr Phe Arg Lys Lys Leu Arg Lys Ala Lys Leu Leu Glu Leu Leu
 65 70 75 80
 Ile Ala Tyr Asp Met Cys Pro Leu Lys Pro Glu Asp Val Arg Arg Trp
 85 90 95
 Lys Asn Trp Asp Lys Gln Gln Lys Ser Thr Val Arg Gln Phe Pro Asp
 100 105 110
 Thr Pro Ala Phe Arg Glu Trp Leu Lys Gln Asn Pro Tyr Glu Leu Arg
 115 120 125
 Lys Gln Ala Val Thr Glu Asp Val Thr Arg Pro Glu Leu Gly Arg Ile
 130 135 140
 Leu Tyr Gln Met Ile Gln Arg Arg Gly Phe Leu Ser Ser Arg Lys Gly
 145 150 155 160
 Lys Glu Glu Gly Lys Ile Phe Thr Gly Lys Asp Arg Met Val Gly Ile
 165 170 175
 Asp Glu Thr Arg Lys Asn Leu Gln Lys Gln Thr Leu Gly Ala Tyr Leu
 180 185 190
 Tyr Asp Ile Ala Pro Lys Asn Gly Glu Lys Tyr Arg Phe Arg Thr Glu
 195 200 205
 Arg Val Arg Ala Arg Tyr Thr Leu Arg Asp Met Tyr Ile Arg Glu Phe
 210 215 220
 Glu Ile Ile Trp Gln Arg Gln Ala Gly His Leu Gly Leu Ala His Glu
 225 230 235 240

Gln Ala Thr Arg Lys Lys Asn Ile Phe Leu Glu Gly Ser Ala Thr Asn		
245	250	255
Val Arg Asn Ser Lys Leu Ile Thr His Leu Gln Ala Lys Tyr Gly Arg		
260	265	270
Gly His Val Leu Ile Glu Asp Thr Arg Ile Thr Val Thr Phe Gln Leu		
275	280	285
Pro Leu Lys Glu Val Leu Gly Gly Lys Ile Glu Ile Glu Glu Glu Gln		
290	295	300
Leu Lys Phe Lys Ser Asn Glu Ser Val Leu Phe Trp Gln Arg Pro Leu		
305	310	315
Arg Ser Gln Lys Ser Leu Leu Ser Lys Cys Val Phe Glu Gly Arg Asn		
325	330	335
Phe Tyr Asp Pro Val His Gln Lys Trp Ile Ile Ala Gly Pro Thr Pro		
340	345	350
Ala Pro Leu Ser His Pro Glu Phe Glu Glu Phe Arg Ala Tyr Gln Phe		
355	360	365
Ile Asn Asn Ile Ile Tyr Gly Lys Asn Glu His Leu Thr Ala Ile Gln		
370	375	380
Arg Glu Ala Val Phe Glu Leu Met Cys Thr Glu Ser Lys Asp Phe Asn		
385	390	395
Phe Glu Lys Ile Pro Lys His Leu Lys Leu Phe Glu Lys Phe Asn Phe		
405	410	415
Asp Asp Thr Thr Lys Val Pro Ala Cys Thr Thr Ile Ser Gln Leu Arg		
420	425	430
Lys Leu Phe Pro His Pro Val Trp Glu Glu Lys Arg Glu Glu Ile Trp		
435	440	445
His Cys Phe Tyr Phe Tyr Asp Asp Asn Thr Leu Leu Phe Glu Lys Leu		
450	455	460
Gln Lys Asp Tyr Ala Leu Gln Thr Asn Asp Leu Glu Lys Ile Lys Lys		
465	470	475
Ile Arg Leu Ser Glu Ser Tyr Gly Asn Val Ser Leu Lys Ala Ile Arg		
485	490	495
Arg Ile Asn Pro Tyr Leu Lys Lys Gly Tyr Ala Tyr Ser Thr Ala Val		
500	505	510
Leu Leu Gly Gly Ile Arg Asn Ser Phe Gly Lys Arg Phe Glu Tyr Phe		
515	520	525
Lys Glu Tyr Glu Pro Glu Ile Glu Lys Ala Val Cys Arg Ile Leu Lys		
530	535	540
Glu Lys Asn Ala Glu Gly Glu Val Ile Arg Lys Ile Lys Asp Tyr Leu		

545	550	555	560
Val His Asn Arg Phe Gly Phe Ala Lys Asn Asp Arg Ala Phe Gln Lys			
	565	570	575
Leu Tyr His His Ser Gln Ala Ile Thr Thr Gln Ala Gln Lys Glu Arg			
	580	585	590
Leu Pro Glu Thr Gly Asn Leu Arg Asn Pro Ile Val Gln Gln Gly Leu			
	595	600	605
Asn Glu Leu Arg Arg Thr Val Asn Lys Leu Leu Ala Thr Cys Arg Glu			
	610	615	620
Lys Tyr Gly Pro Ser Phe Lys Phe Asp His Ile His Val Glu Met Gly			
625	630	635	640
Arg Glu Leu Arg Ser Ser Lys Thr Glu Arg Glu Lys Gln Ser Arg Gln			
	645	650	655
Ile Arg Glu Asn Glu Lys Lys Asn Glu Ala Ala Lys Val Lys Leu Ala			
	660	665	670
Glu Tyr Gly Leu Lys Ala Tyr Arg Asp Asn Ile Gln Lys Tyr Leu Leu			
	675	680	685
Tyr Lys Glu Ile Glu Glu Lys Gly Gly Thr Val Cys Cys Pro Tyr Thr			
	690	695	700
Gly Lys Thr Leu Asn Ile Ser His Thr Leu Gly Ser Asp Asn Ser Val			
705	710	715	720
Gln Ile Glu His Ile Ile Pro Tyr Ser Ile Ser Leu Asp Asp Ser Leu			
	725	730	735
Ala Asn Lys Thr Leu Cys Asp Ala Thr Phe Asn Arg Glu Lys Gly Glu			
	740	745	750
Leu Thr Pro Tyr Asp Phe Tyr Gln Lys Asp Pro Ser Pro Glu Lys Trp			
	755	760	765
Gly Ala Ser Ser Trp Glu Glu Ile Glu Asp Arg Ala Phe Arg Leu Leu			
	770	775	780
Pro Tyr Ala Lys Ala Gln Arg Phe Ile Arg Arg Lys Pro Gln Glu Ser			
785	790	795	800
Asn Glu Phe Ile Ser Arg Gln Leu Asn Asp Thr Arg Tyr Ile Ser Lys			
	805	810	815
Lys Ala Val Glu Tyr Leu Ser Ala Ile Cys Ser Asp Val Lys Ala Phe			
	820	825	830
Pro Gly Gln Leu Thr Ala Glu Leu Arg His Leu Trp Gly Leu Asn Asn			
	835	840	845
Ile Leu Gln Ser Ala Pro Asp Ile Thr Phe Pro Leu Pro Val Ser Ala			
	850	855	860

Thr	Glu	Asn	His	Arg	Glu	Tyr	Tyr	Val	Ile	Thr	Asn	Glu	Gln	Asn	Glu
865					870					875					880
Val	Ile	Arg	Leu	Phe	Pro	Lys	Gln	Gly	Glu	Thr	Pro	Arg	Thr	Glu	Lys
			885						890					895	
Gly	Glu	Leu	Leu	Leu	Thr	Gly	Glu	Val	Glu	Arg	Lys	Val	Phe	Arg	Cys
		900						905					910		
Lys	Gly	Met	Gln	Glu	Phe	Gln	Thr	Asp	Val	Ser	Asp	Gly	Lys	Tyr	Trp
	915						920					925			
Arg	Arg	Ile	Lys	Leu	Ser	Ser	Ser	Val	Thr	Trp	Ser	Pro	Leu	Phe	Ala
	930					935					940				
Pro	Lys	Pro	Ile	Ser	Ala	Asp	Gly	Gln	Ile	Val	Leu	Lys	Gly	Arg	Ile
945				950					955						960
Glu	Lys	Gly	Val	Phe	Val	Cys	Asn	Gln	Leu	Lys	Gln	Lys	Leu	Lys	Thr
			965					970						975	
Gly	Leu	Pro	Asp	Gly	Ser	Tyr	Trp	Ile	Ser	Leu	Pro	Val	Ile	Ser	Gln
		980						985					990		
Thr	Phe	Lys	Glu	Gly	Glu	Ser	Val	Asn	Asn	Ser	Lys	Leu	Thr	Ser	Gln
	995						1000					1005			
Gln	Val	Gln	Leu	Phe	Gly	Arg	Val	Arg	Glu	Gly	Ile	Phe	Arg	Cys	
	1010					1015					1020				
His	Asn	Tyr	Gln	Cys	Pro	Ala	Ser	Gly	Ala	Asp	Gly	Asn	Phe	Trp	
	1025					1030					1035				
Cys	Thr	Leu	Asp	Thr	Asp	Thr	Ala	Gln	Pro	Ala	Phe	Thr	Pro	Ile	
	1040					1045					1050				
Lys	Asn	Ala	Pro	Pro	Gly	Val	Gly	Gly	Gly	Gln	Ile	Ile	Leu	Thr	
	1055					1060					1065				
Gly	Asp	Val	Asp	Asp	Lys	Gly	Ile	Phe	His	Ala	Asp	Asp	Asp	Leu	
	1070					1075					1080				
His	Tyr	Glu	Leu	Pro	Ala	Ser	Leu	Pro	Lys	Gly	Lys	Tyr	Tyr	Gly	
	1085					1090					1095				
Ile	Phe	Thr	Val	Glu	Ser	Cys	Asp	Pro	Thr	Leu	Ile	Pro	Ile	Glu	
	1100					1105					1110				
Leu	Ser	Ala	Pro	Lys	Thr	Ser	Lys	Gly	Glu	Asn	Leu	Ile	Glu	Gly	
	1115					1120					1125				
Asn	Ile	Trp	Val	Asp	Glu	His	Thr	Gly	Glu	Val	Arg	Phe	Asp	Pro	
	1130					1135					1140				
Lys	Lys	Asn	Arg	Glu	Asp	Gln	Arg	His	His	Ala	Ile	Asp	Ala	Ile	
	1145					1150					1155				
Val	Ile	Ala	Leu	Ser	Ser	Gln	Ser	Leu	Phe	Gln	Arg	Leu	Ser	Thr	

1160	1165	1170
Tyr Asn Ala Arg Arg Glu Asn Lys Lys Arg Gly	Leu Asp Ser Thr	
1175	1180	1185
Glu His Phe Pro Ser Pro Trp Pro Gly Phe Ala	Gln Asp Val Arg	
1190	1195	1200
Gln Ser Val Val Pro Leu Leu Val Ser Tyr Lys	Gln Asn Pro Lys	
1205	1210	1215
Thr Leu Cys Lys Ile Ser Lys Thr Leu Tyr Lys	Asp Gly Lys Lys	
1220	1225	1230
Ile His Ser Cys Gly Asn Ala Val Arg Gly Gln	Leu His Lys Glu	
1235	1240	1245
Thr Val Tyr Gly Gln Arg Thr Ala Pro Gly Ala	Thr Glu Lys Ser	
1250	1255	1260
Tyr His Ile Arg Lys Asp Ile Arg Glu Leu Lys	Thr Ser Lys His	
1265	1270	1275
Ile Gly Lys Val Val Asp Ile Thr Ile Arg Gln	Met Leu Leu Lys	
1280	1285	1290
His Leu Gln Glu Asn Tyr His Ile Asp Ile Thr	Gln Glu Phe Asn	
1295	1300	1305
Ile Pro Ser Asn Ala Phe Phe Lys Glu Gly Val	Tyr Arg Ile Phe	
1310	1315	1320
Leu Pro Asn Lys His Gly Glu Pro Val Pro Ile	Lys Lys Ile Arg	
1325	1330	1335
Met Lys Glu Glu Leu Gly Asn Ala Glu Arg Leu	Lys Asp Asn Ile	
1340	1345	1350
Asn Gln Tyr Val Asn Pro Arg Asn Asn His His	Val Met Ile Tyr	
1355	1360	1365
Gln Asp Ala Asp Gly Asn Leu Lys Glu Glu Ile	Val Ser Phe Trp	
1370	1375	1380
Ser Val Ile Glu Arg Gln Asn Gln Gly Gln Pro	Ile Tyr Gln Leu	
1385	1390	1395
Pro Arg Glu Gly Arg Asn Ile Val Ser Ile Leu	Gln Ile Asn Asp	
1400	1405	1410
Thr Phe Leu Ile Gly Leu Lys Glu Glu Glu Pro	Glu Val Tyr Arg	
1415	1420	1425
Asn Asp Leu Ser Thr Leu Ser Lys His Leu Tyr	Arg Val Gln Lys	
1430	1435	1440
Leu Ser Gly Met Tyr Tyr Thr Phe Arg His His	Leu Ala Ser Thr	
1445	1450	1455

Leu Asn Asn Glu Arg Glu Glu Phe Arg Ile Gln Ser Leu Glu Ala
 1460 1465 1470
 Trp Lys Arg Ala Asn Pro Val Lys Val Gln Ile Asp Glu Ile Gly
 1475 1480 1485
 Arg Ile Thr Phe Leu Asn Gly Pro Leu Cys
 1490 1495
 <210> 172
 <211> 1300
 <212> PRT
 <213> 土拉热弗朗西丝菌新凶手亚种 (*Francisella tularensis* subsp. *Novicida*)
 <400> 172
 Met Ser Ile Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr
 1 5 10 15
 Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys
 20 25 30
 Ala Arg Gly Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys
 35 40 45
 Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu
 50 55 60
 Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser
 65 70 75 80
 Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys
 85 90 95
 Asp Phe Lys Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr
 100 105 110
 Ile Lys Asp Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile
 115 120 125
 Asp Ala Lys Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln
 130 135 140
 Ser Lys Asp Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr
 145 150 155 160
 Asp Ile Asp Glu Ala Leu Glu Ile Ile Lys Ser Phe Lys Gly Trp Thr
 165 170 175
 Thr Tyr Phe Lys Gly Phe His Glu Asn Arg Lys Asn Val Tyr Ser Ser
 180 185 190
 Asn Asp Ile Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp Asp Asn Leu
 195 200 205
 Pro Lys Phe Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys
 210 215 220

Ala Pro Glu Ala Ile Asn Tyr Glu Gln Ile Lys Lys Asp Leu Ala Glu			
225	230	235	240
Glu Leu Thr Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg			
	245	250	255
Val Phe Ser Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr			
	260	265	270
Leu Asn Gln Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys			
	275	280	285
Phe Val Asn Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile			
	290	295	300
Asn Leu Tyr Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys			
305	310	315	320
Met Ser Val Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser			
	325	330	335
Phe Val Ile Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met			
	340	345	350
Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys			
	355	360	365
Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln			
	370	375	380
Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr			
385	390	395	400
Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala			
	405	410	415
Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn			
	420	425	430
Pro Ser Lys Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala			
	435	440	445
Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn			
	450	455	460
Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala			
465	470	475	480
Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys			
	485	490	495
Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys			
	500	505	510
Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp			
	515	520	525
Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His			

530	535	540
Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His		
545	550	555
Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val		
565	570	575
Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser		
580	585	590
Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly		
595	600	605
Trp Asp Lys Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys		
610	615	620
Asp Asp Lys Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile		
625	630	635
Phe Asp Asp Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys		
645	650	655
Ile Val Tyr Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val		
660	665	670
Phe Phe Ser Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile		
675	680	685
Leu Arg Ile Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln		
690	695	700
Lys Gly Tyr Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe		
705	710	715
Ile Asp Phe Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp		
725	730	735
Phe Gly Phe Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu		
740	745	750
Phe Tyr Arg Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn		
755	760	765
Ile Ser Glu Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr		
770	775	780
Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg		
785	790	795
Pro Asn Leu His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn		
805	810	815
Leu Gln Asp Val Val Tyr Lys Leu Asn Gly Glu Ala Glu Leu Phe Tyr		
820	825	830
Arg Lys Gln Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala		
835	840	845

Ile Ala Asn Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu		
850	855	860
Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe		
865	870	875 880
His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe		
	885	890 895
Asn Asp Glu Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His		
	900	905 910
Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu		
	915	920 925
Val Asp Gly Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile		
	930	935 940
Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile		
945	950	955 960
Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn		
	965	970 975
Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile		
	980	985 990
Ala Lys Leu Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu		
	995	1000 1005
Asn Phe Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val		
1010	1015	1020
Tyr Gln Lys Leu Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu		
1025	1030	1035
Val Phe Lys Asp Asn Glu Phe Asp Lys Thr Gly Gly Val Leu Arg		
1040	1045	1050
Ala Tyr Gln Leu Thr Ala Pro Phe Glu Thr Phe Lys Lys Met Gly		
1055	1060	1065
Lys Gln Thr Gly Ile Ile Tyr Tyr Val Pro Ala Gly Phe Thr Ser		
1070	1075	1080
Lys Ile Cys Pro Val Thr Gly Phe Val Asn Gln Leu Tyr Pro Lys		
1085	1090	1095
Tyr Glu Ser Val Ser Lys Ser Gln Glu Phe Phe Ser Lys Phe Asp		
1100	1105	1110
Lys Ile Cys Tyr Asn Leu Asp Lys Gly Tyr Phe Glu Phe Ser Phe		
1115	1120	1125
Asp Tyr Lys Asn Phe Gly Asp Lys Ala Ala Lys Gly Lys Trp Thr		
1130	1135	1140
Ile Ala Ser Phe Gly Ser Arg Leu Ile Asn Phe Arg Asn Ser Asp		

1145	1150	1155
Lys Asn His Asn Trp Asp Thr Arg Glu Val Tyr Pro Thr Lys Glu		
1160	1165	1170
Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu Tyr Gly His Gly		
1175	1180	1185
Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp Lys Lys Phe		
1190	1195	1200
Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln Met Arg		
1205	1210	1215
Asn Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro Val		
1220	1225	1230
Ala Asp Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys		
1235	1240	1245
Asn Met Pro Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly		
1250	1255	1260
Leu Lys Gly Leu Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu		
1265	1270	1275
Gly Lys Lys Leu Asn Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu		
1280	1285	1290
Phe Val Gln Asn Arg Asn Asn		
1295	1300	
<210> 173		
<211> 1228		
<212> PRT		
<213> 毛螺科细菌物种(Lachnospiraceae bacterium sp.)ND2006		
<400> 173		
Ala Ala Ser Lys Leu Glu Lys Phe Thr Asn Cys Tyr Ser Leu Ser Lys		
1 5 10 15		
Thr Leu Arg Phe Lys Ala Ile Pro Val Gly Lys Thr Gln Glu Asn Ile		
20 25 30		
Asp Asn Lys Arg Leu Leu Val Glu Asp Glu Lys Arg Ala Glu Asp Tyr		
35 40 45		
Lys Gly Val Lys Lys Leu Leu Asp Arg Tyr Tyr Leu Ser Phe Ile Asn		
50 55 60		
Asp Val Leu His Ser Ile Lys Leu Lys Asn Leu Asn Asn Tyr Ile Ser		
65 70 75 80		
Leu Phe Arg Lys Lys Thr Arg Thr Glu Lys Glu Asn Lys Glu Leu Glu		
85 90 95		
Asn Leu Glu Ile Asn Leu Arg Lys Glu Ile Ala Lys Ala Phe Lys Gly		

100	105	110
Ala Ala Gly Tyr Lys Ser Leu Phe	Lys Lys Asp Ile Ile Glu Thr Ile	
115	120	125
Leu Pro Glu Ala Ala Asp Asp Lys	Asp Glu Ile Ala Leu Val Asn Ser	
130	135	140
Phe Asn Gly Phe Thr Thr Ala Phe	Thr Gly Phe Phe Asp Asn Arg Glu	
145	150	155
Asn Met Phe Ser Glu Glu Ala Lys	Ser Thr Ser Ile Ala Phe Arg Cys	
165	170	175
Ile Asn Glu Asn Leu Thr Arg Tyr	Ile Ser Asn Met Asp Ile Phe Glu	
180	185	190
Lys Val Asp Ala Ile Phe Asp Lys	His Glu Val Gln Glu Ile Lys Glu	
195	200	205
Lys Ile Leu Asn Ser Asp Tyr Asp	Val Glu Asp Phe Phe Glu Gly Glu	
210	215	220
Phe Phe Asn Phe Val Leu Thr Gln	Glu Gly Ile Asp Val Tyr Asn Ala	
225	230	235
Ile Ile Gly Gly Phe Val Thr Glu	Ser Gly Glu Lys Ile Lys Gly Leu	
245	250	255
Asn Glu Tyr Ile Asn Leu Tyr Asn	Ala Lys Thr Lys Gln Ala Leu Pro	
260	265	270
Lys Phe Lys Pro Leu Tyr Lys Gln	Val Leu Ser Asp Arg Glu Ser Leu	
275	280	285
Ser Phe Tyr Gly Glu Gly Tyr Thr	Ser Asp Glu Glu Val Leu Glu Val	
290	295	300
Phe Arg Asn Thr Leu Asn Lys Asn	Ser Glu Ile Phe Ser Ser Ile Lys	
305	310	315
Lys Leu Glu Lys Leu Phe Lys Asn	Phe Asp Glu Tyr Ser Ser Ala Gly	
325	330	335
Ile Phe Val Lys Asn Gly Pro Ala	Ile Ser Thr Ile Ser Lys Asp Ile	
340	345	350
Phe Gly Glu Trp Asn Leu Ile Arg	Asp Lys Trp Asn Ala Glu Tyr Asp	
355	360	365
Asp Ile His Leu Lys Lys Lys Ala	Val Val Thr Glu Lys Tyr Glu Asp	
370	375	380
Asp Arg Arg Lys Ser Phe Lys Lys	Ile Gly Ser Phe Ser Leu Glu Gln	
385	390	395
Leu Gln Glu Tyr Ala Asp Ala Asp	Leu Ser Val Val Glu Lys Leu Lys	
405	410	415

Glu Ile Ile Ile Gln Lys Val Asp Glu Ile Tyr Lys Val Tyr Gly Ser			
420	425	430	
Ser Glu Lys Leu Phe Asp Ala Asp Phe Val Leu Glu Lys Ser Leu Lys			
435	440	445	
Lys Asn Asp Ala Val Val Ala Ile Met Lys Asp Leu Leu Asp Ser Val			
450	455	460	
Lys Ser Phe Glu Asn Tyr Ile Lys Ala Phe Phe Gly Glu Gly Lys Glu			
465	470	475	480
Thr Asn Arg Asp Glu Ser Phe Tyr Gly Asp Phe Val Leu Ala Tyr Asp			
485	490	495	
Ile Leu Leu Lys Val Asp His Ile Tyr Asp Ala Ile Arg Asn Tyr Val			
500	505	510	
Thr Gln Lys Pro Tyr Ser Lys Asp Lys Phe Lys Leu Tyr Phe Gln Asn			
515	520	525	
Pro Gln Phe Met Gly Gly Trp Asp Lys Asp Lys Glu Thr Asp Tyr Arg			
530	535	540	
Ala Thr Ile Leu Arg Tyr Gly Ser Lys Tyr Tyr Leu Ala Ile Met Asp			
545	550	555	560
Lys Lys Tyr Ala Lys Cys Leu Gln Lys Ile Asp Lys Asp Asp Val Asn			
565	570	575	
Gly Asn Tyr Glu Lys Ile Asn Tyr Lys Leu Leu Pro Gly Pro Asn Lys			
580	585	590	
Met Leu Pro Lys Val Phe Phe Ser Lys Lys Trp Met Ala Tyr Tyr Asn			
595	600	605	
Pro Ser Glu Asp Ile Gln Lys Ile Tyr Lys Asn Gly Thr Phe Lys Lys			
610	615	620	
Gly Asp Met Phe Asn Leu Asn Asp Cys His Lys Leu Ile Asp Phe Phe			
625	630	635	640
Lys Asp Ser Ile Ser Arg Tyr Pro Lys Trp Ser Asn Ala Tyr Asp Phe			
645	650	655	
Asn Phe Ser Glu Thr Glu Lys Tyr Lys Asp Ile Ala Gly Phe Tyr Arg			
660	665	670	
Glu Val Glu Glu Gln Gly Tyr Lys Val Ser Phe Glu Ser Ala Ser Lys			
675	680	685	
Lys Glu Val Asp Lys Leu Val Glu Glu Gly Lys Leu Tyr Met Phe Gln			
690	695	700	
Ile Tyr Asn Lys Asp Phe Ser Asp Lys Ser His Gly Thr Pro Asn Leu			
705	710	715	720
His Thr Met Tyr Phe Lys Leu Leu Phe Asp Glu Asn Asn His Gly Gln			

	725		730		735
Ile Arg Leu Ser Gly Gly Ala Glu Leu Phe Met Arg Arg Ala Ser Leu					
	740		745		750
Lys Lys Glu Glu Leu Val Val His Pro Ala Asn Ser Pro Ile Ala Asn					
	755		760		765
Lys Asn Pro Asp Asn Pro Lys Lys Thr Thr Thr Leu Ser Tyr Asp Val					
	770		775		780
Tyr Lys Asp Lys Arg Phe Ser Glu Asp Gln Tyr Glu Leu His Ile Pro					
785		790		795	800
Ile Ala Ile Asn Lys Cys Pro Lys Asn Ile Phe Lys Ile Asn Thr Glu					
	805		810		815
Val Arg Val Leu Leu Lys His Asp Asp Asn Pro Tyr Val Ile Gly Ile					
	820		825		830
Asp Arg Gly Glu Arg Asn Leu Leu Tyr Ile Val Val Val Asp Gly Lys					
	835		840		845
Gly Asn Ile Val Glu Gln Tyr Ser Leu Asn Glu Ile Ile Asn Asn Phe					
	850		855		860
Asn Gly Ile Arg Ile Lys Thr Asp Tyr His Ser Leu Leu Asp Lys Lys					
865		870		875	880
Glu Lys Glu Arg Phe Glu Ala Arg Gln Asn Trp Thr Ser Ile Glu Asn					
	885		890		895
Ile Lys Glu Leu Lys Ala Gly Tyr Ile Ser Gln Val Val His Lys Ile					
	900		905		910
Cys Glu Leu Val Glu Lys Tyr Asp Ala Val Ile Ala Leu Glu Asp Leu					
	915		920		925
Asn Ser Gly Phe Lys Asn Ser Arg Val Lys Val Glu Lys Gln Val Tyr					
	930		935		940
Gln Lys Phe Glu Lys Met Leu Ile Asp Lys Leu Asn Tyr Met Val Asp					
945		950		955	960
Lys Lys Ser Asn Pro Cys Ala Thr Gly Gly Ala Leu Lys Gly Tyr Gln					
	965		970		975
Ile Thr Asn Lys Phe Glu Ser Phe Lys Ser Met Ser Thr Gln Asn Gly					
	980		985		990
Phe Ile Phe Tyr Ile Pro Ala Trp Leu Thr Ser Lys Ile Asp Pro Ser					
	995		1000		1005
Thr Gly Phe Val Asn Leu Leu Lys Thr Lys Tyr Thr Ser Ile Ala					
	1010		1015		1020
Asp Ser Lys Lys Phe Ile Ser Ser Phe Asp Arg Ile Met Tyr Val					
	1025		1030		1035

Pro Glu Glu Asp Leu Phe Glu Phe Ala Leu Asp Tyr Lys Asn Phe		
1040	1045	1050
Ser Arg Thr Asp Ala Asp Tyr Ile Lys Lys Trp Lys Leu Tyr Ser		
1055	1060	1065
Tyr Gly Asn Arg Ile Arg Ile Phe Ala Ala Ala Lys Lys Asn Asn		
1070	1075	1080
Val Phe Ala Trp Glu Glu Val Cys Leu Thr Ser Ala Tyr Lys Glu		
1085	1090	1095
Leu Phe Asn Lys Tyr Gly Ile Asn Tyr Gln Gln Gly Asp Ile Arg		
1100	1105	1110
Ala Leu Leu Cys Glu Gln Ser Asp Lys Ala Phe Tyr Ser Ser Phe		
1115	1120	1125
Met Ala Leu Met Ser Leu Met Leu Gln Met Arg Asn Ser Ile Thr		
1130	1135	1140
Gly Arg Thr Asp Val Asp Phe Leu Ile Ser Pro Val Lys Asn Ser		
1145	1150	1155
Asp Gly Ile Phe Tyr Asp Ser Arg Asn Tyr Glu Ala Gln Glu Asn		
1160	1165	1170
Ala Ile Leu Pro Lys Asn Ala Asp Ala Asn Gly Ala Tyr Asn Ile		
1175	1180	1185
Ala Arg Lys Val Leu Trp Ala Ile Gly Gln Phe Lys Lys Ala Glu		
1190	1195	1200
Asp Glu Lys Leu Asp Lys Val Lys Ile Ala Ile Ser Asn Lys Glu		
1205	1210	1215
Trp Leu Glu Tyr Ala Gln Thr Ser Val Lys		
1220	1225	

<210> 174

<211> 1307

<212> PRT

<213> 氨基酸球菌属物种(Acidaminococcus sp.)BV3L6

<400> 174

Met Thr Gln Phe Glu Gly Phe Thr Asn Leu Tyr Gln Val Ser Lys Thr		
1	5	10
Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Lys His Ile Gln		
20	25	30
Glu Gln Gly Phe Ile Glu Glu Asp Lys Ala Arg Asn Asp His Tyr Lys		
35	40	45
Glu Leu Lys Pro Ile Ile Asp Arg Ile Tyr Lys Thr Tyr Ala Asp Gln		
50	55	60

Cys	Leu	Gln	Leu	Val	Gln	Leu	Asp	Trp	Glu	Asn	Leu	Ser	Ala	Ala	Ile
65					70				75						80
Asp	Ser	Tyr	Arg	Lys	Glu	Lys	Thr	Glu	Glu	Thr	Arg	Asn	Ala	Leu	Ile
				85					90					95	
Glu	Glu	Gln	Ala	Thr	Tyr	Arg	Asn	Ala	Ile	His	Asp	Tyr	Phe	Ile	Gly
			100					105					110		
Arg	Thr	Asp	Asn	Leu	Thr	Asp	Ala	Ile	Asn	Lys	Arg	His	Ala	Glu	Ile
			115					120					125		
Tyr	Lys	Gly	Leu	Phe	Lys	Ala	Glu	Leu	Phe	Asn	Gly	Lys	Val	Leu	Lys
			130				135					140			
Gln	Leu	Gly	Thr	Val	Thr	Thr	Thr	Glu	His	Glu	Asn	Ala	Leu	Leu	Arg
145					150					155					160
Ser	Phe	Asp	Lys	Phe	Thr	Thr	Tyr	Phe	Ser	Gly	Phe	Tyr	Glu	Asn	Arg
			165							170				175	
Lys	Asn	Val	Phe	Ser	Ala	Glu	Asp	Ile	Ser	Thr	Ala	Ile	Pro	His	Arg
			180					185					190		
Ile	Val	Gln	Asp	Asn	Phe	Pro	Lys	Phe	Lys	Glu	Asn	Cys	His	Ile	Phe
			195					200					205		
Thr	Arg	Leu	Ile	Thr	Ala	Val	Pro	Ser	Leu	Arg	Glu	His	Phe	Glu	Asn
			210					215					220		
Val	Lys	Lys	Ala	Ile	Gly	Ile	Phe	Val	Ser	Thr	Ser	Ile	Glu	Glu	Val
225					230					235					240
Phe	Ser	Phe	Pro	Phe	Tyr	Asn	Gln	Leu	Leu	Thr	Gln	Thr	Gln	Ile	Asp
			245							250				255	
Leu	Tyr	Asn	Gln	Leu	Leu	Gly	Gly	Ile	Ser	Arg	Glu	Ala	Gly	Thr	Glu
			260					265					270		
Lys	Ile	Lys	Gly	Leu	Asn	Glu	Val	Leu	Asn	Leu	Ala	Ile	Gln	Lys	Asn
			275					280					285		
Asp	Glu	Thr	Ala	His	Ile	Ile	Ala	Ser	Leu	Pro	His	Arg	Phe	Ile	Pro
			290				295					300			
Leu	Phe	Lys	Gln	Ile	Leu	Ser	Asp	Arg	Asn	Thr	Leu	Ser	Phe	Ile	Leu
305					310					315					320
Glu	Glu	Phe	Lys	Ser	Asp	Glu	Glu	Val	Ile	Gln	Ser	Phe	Cys	Lys	Tyr
			325							330				335	
Lys	Thr	Leu	Leu	Arg	Asn	Glu	Asn	Val	Leu	Glu	Thr	Ala	Glu	Ala	Leu
			340					345					350		
Phe	Asn	Glu	Leu	Asn	Ser	Ile	Asp	Leu	Thr	His	Ile	Phe	Ile	Ser	His
			355					360					365		
Lys	Lys	Leu	Glu	Thr	Ile	Ser	Ser	Ala	Leu	Cys	Asp	His	Trp	Asp	Thr

370	375	380
Leu Arg Asn Ala Leu Tyr Glu Arg Arg Ile Ser Glu Leu Thr Gly Lys		
385	390	395
Ile Thr Lys Ser Ala Lys Glu Lys Val Gln Arg Ser Leu Lys His Glu		400
405	410	415
Asp Ile Asn Leu Gln Glu Ile Ile Ser Ala Ala Gly Lys Glu Leu Ser		
420	425	430
Glu Ala Phe Lys Gln Lys Thr Ser Glu Ile Leu Ser His Ala His Ala		
435	440	445
Ala Leu Asp Gln Pro Leu Pro Thr Thr Leu Lys Lys Gln Glu Glu Lys		
450	455	460
Glu Ile Leu Lys Ser Gln Leu Asp Ser Leu Leu Gly Leu Tyr His Leu		
465	470	475
Leu Asp Trp Phe Ala Val Asp Glu Ser Asn Glu Val Asp Pro Glu Phe		
485	490	495
Ser Ala Arg Leu Thr Gly Ile Lys Leu Glu Met Glu Pro Ser Leu Ser		
500	505	510
Phe Tyr Asn Lys Ala Arg Asn Tyr Ala Thr Lys Lys Pro Tyr Ser Val		
515	520	525
Glu Lys Phe Lys Leu Asn Phe Gln Met Pro Thr Leu Ala Ser Gly Trp		
530	535	540
Asp Val Asn Lys Glu Lys Asn Asn Gly Ala Ile Leu Phe Val Lys Asn		
545	550	555
Gly Leu Tyr Tyr Leu Gly Ile Met Pro Lys Gln Lys Gly Arg Tyr Lys		
565	570	575
Ala Leu Ser Phe Glu Pro Thr Glu Lys Thr Ser Glu Gly Phe Asp Lys		
580	585	590
Met Tyr Tyr Asp Tyr Phe Pro Asp Ala Ala Lys Met Ile Pro Lys Cys		
595	600	605
Ser Thr Gln Leu Lys Ala Val Thr Ala His Phe Gln Thr His Thr Thr		
610	615	620
Pro Ile Leu Leu Ser Asn Asn Phe Ile Glu Pro Leu Glu Ile Thr Lys		
625	630	635
Glu Ile Tyr Asp Leu Asn Asn Pro Glu Lys Glu Pro Lys Lys Phe Gln		640
645	650	655
Thr Ala Tyr Ala Lys Lys Thr Gly Asp Gln Lys Gly Tyr Arg Glu Ala		
660	665	670
Leu Cys Lys Trp Ile Asp Phe Thr Arg Asp Phe Leu Ser Lys Tyr Thr		
675	680	685

Lys Thr Thr Ser Ile Asp Leu Ser Ser Leu Arg Pro Ser Ser Gln Tyr		
690	695	700
Lys Asp Leu Gly Glu Tyr Tyr Ala Glu Leu Asn Pro Leu Leu Tyr His		
705	710	715 720
Ile Ser Phe Gln Arg Ile Ala Glu Lys Glu Ile Met Asp Ala Val Glu		
	725	730 735
Thr Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ala Lys		
	740	745 750
Gly His His Gly Lys Pro Asn Leu His Thr Leu Tyr Trp Thr Gly Leu		
	755	760 765
Phe Ser Pro Glu Asn Leu Ala Lys Thr Ser Ile Lys Leu Asn Gly Gln		
	770	775 780
Ala Glu Leu Phe Tyr Arg Pro Lys Ser Arg Met Lys Arg Met Ala His		
785	790	795 800
Arg Leu Gly Glu Lys Met Leu Asn Lys Lys Leu Lys Asp Gln Lys Thr		
	805	810 815
Pro Ile Pro Asp Thr Leu Tyr Gln Glu Leu Tyr Asp Tyr Val Asn His		
	820	825 830
Arg Leu Ser His Asp Leu Ser Asp Glu Ala Arg Ala Leu Leu Pro Asn		
	835	840 845
Val Ile Thr Lys Glu Val Ser His Glu Ile Ile Lys Asp Arg Arg Phe		
850	855	860
Thr Ser Asp Lys Phe Phe Phe His Val Pro Ile Thr Leu Asn Tyr Gln		
865	870	875 880
Ala Ala Asn Ser Pro Ser Lys Phe Asn Gln Arg Val Asn Ala Tyr Leu		
	885	890 895
Lys Glu His Pro Glu Thr Pro Ile Ile Gly Ile Asp Arg Gly Glu Arg		
	900	905 910
Asn Leu Ile Tyr Ile Thr Val Ile Asp Ser Thr Gly Lys Ile Leu Glu		
	915	920 925
Gln Arg Ser Leu Asn Thr Ile Gln Gln Phe Asp Tyr Gln Lys Lys Leu		
	930	935 940
Asp Asn Arg Glu Lys Glu Arg Val Ala Ala Arg Gln Ala Trp Ser Val		
945	950	955 960
Val Gly Thr Ile Lys Asp Leu Lys Gln Gly Tyr Leu Ser Gln Val Ile		
	965	970 975
His Glu Ile Val Asp Leu Met Ile His Tyr Gln Ala Val Val Val Leu		
	980	985 990
Glu Asn Leu Asn Phe Gly Phe Lys Ser Lys Arg Thr Gly Ile Ala Glu		

995	1000	1005
Lys Ala Val Tyr Gln Gln Phe Glu Lys Met Leu Ile Asp Lys Leu		
1010	1015	1020
Asn Cys Leu Val Leu Lys Asp Tyr Pro Ala Glu Lys Val Gly Gly		
1025	1030	1035
Val Leu Asn Pro Tyr Gln Leu Thr Asp Gln Phe Thr Ser Phe Ala		
1040	1045	1050
Lys Met Gly Thr Gln Ser Gly Phe Leu Phe Tyr Val Pro Ala Pro		
1055	1060	1065
Tyr Thr Ser Lys Ile Asp Pro Leu Thr Gly Phe Val Asp Pro Phe		
1070	1075	1080
Val Trp Lys Thr Ile Lys Asn His Glu Ser Arg Lys His Phe Leu		
1085	1090	1095
Glu Gly Phe Asp Phe Leu His Tyr Asp Val Lys Thr Gly Asp Phe		
1100	1105	1110
Ile Leu His Phe Lys Met Asn Arg Asn Leu Ser Phe Gln Arg Gly		
1115	1120	1125
Leu Pro Gly Phe Met Pro Ala Trp Asp Ile Val Phe Glu Lys Asn		
1130	1135	1140
Glu Thr Gln Phe Asp Ala Lys Gly Thr Pro Phe Ile Ala Gly Lys		
1145	1150	1155
Arg Ile Val Pro Val Ile Glu Asn His Arg Phe Thr Gly Arg Tyr		
1160	1165	1170
Arg Asp Leu Tyr Pro Ala Asn Glu Leu Ile Ala Leu Leu Glu Glu		
1175	1180	1185
Lys Gly Ile Val Phe Arg Asp Gly Ser Asn Ile Leu Pro Lys Leu		
1190	1195	1200
Leu Glu Asn Asp Asp Ser His Ala Ile Asp Thr Met Val Ala Leu		
1205	1210	1215
Ile Arg Ser Val Leu Gln Met Arg Asn Ser Asn Ala Ala Thr Gly		
1220	1225	1230
Glu Asp Tyr Ile Asn Ser Pro Val Arg Asp Leu Asn Gly Val Cys		
1235	1240	1245
Phe Asp Ser Arg Phe Gln Asn Pro Glu Trp Pro Met Asp Ala Asp		
1250	1255	1260
Ala Asn Gly Ala Tyr His Ile Ala Leu Lys Gly Gln Leu Leu Leu		
1265	1270	1275
Asn His Leu Lys Glu Ser Lys Asp Leu Lys Leu Gln Asn Gly Ile		
1280	1285	1290

Ser Asn Gln Asp Trp Leu Ala Tyr Ile Gln Glu Leu Arg Asn
1295 1300 1305
<210> 175
<211> 28
<212> DNA
<213> 沙氏纤毛菌(*Leptotrichia shahii*)
<400> 175
ccacccaat atcgaagggg actaaaac 28
<210> 176
<211> 36
<212> DNA
<213> 韦德纤毛菌(*Leptotrichia wadei*)
<400> 176
gatttagact accccaacaaa cgaaggggac taaaac 36
<210> 177
<211> 36
<212> DNA
<213> 西尔李斯特菌(*Listeria seeligeri*)
<400> 177
gtaagagact acctctatat gaaagaggac taaaac 36
<210> 178
<211> 35
<212> DNA
<213> 毛螺科(*Lachnospiraceae*)细菌MA2020
<400> 178
gtattgagaa aagccagata tagttggcaa tagac 35
<210> 179
<211> 35
<212> DNA
<213> 毛螺科(*Lachnospiraceae*)细菌NK4A179
<400> 179
gttgatgaga agagcccaag atagagggca ataac 35
<210> 180
<211> 35
<212> DNA
<213> 嗜氨基梭菌(*Clostridium aminophilum*)DSM 10710
<400> 180
gtctattgcc ctctatatcg ggctgttctc caaac 35
<210> 181

<211> 36
<212> DNA
<213> 鸡肉杆菌 (*Carnobacterium gallinarum*) DSM 4847
<400> 181
atttaaagact acctctaaat gtaagaggac tataac 36
<210> 182
<211> 36
<212> DNA
<213> 鸡肉杆菌 (*Carnobacterium gallinarum*) DSM 4847
<400> 182
aatataaact acctctaaat gtaagaggac tataac 36
<210> 183
<211> 36
<212> DNA
<213> 产丙酸沼杆菌 (*Paludibacter propionicigenes*) WB4
<400> 183
cttgtggatt atcccaaaat tgaagggaac tacaac 36
<210> 184
<211> 36
<212> DNA
<213> 韦氏李斯特菌 (*Listeria weihenstephanensis*) FSL R9-0317
<400> 184
gatttagagt acctcaaaat agaagaggac taaaac 36
<210> 185
<211> 36
<212> DNA
<213> 李斯特菌科 (*Listeriaceae*) 细菌 FSL M6-0635
<400> 185
gatttagagt acctcaaaaac aaaagaggac taaaac 36
<210> 186
<211> 36
<212> DNA
<213> 韦德纤毛菌 (*Leptotrichia wadei*) F0279
<400> 186
gatatagata accccaaaaa cgaaggatc taaaac 36
<210> 187
<211> 36
<212> DNA
<213> 荚膜红细菌 (*Rhodobacter capsulatus*) SB 1003

<400> 187

gcctcacatc accgccaaga cgacggcgga ctgaac 36

<210> 188

<211> 36

<212> DNA

<213> 荚膜红细菌 (Rhodobacter capsulatus) R121

<400> 188

gcctcacatc accgccaaga cgacggcgga ctgaac 36

<210> 189

<211> 36

<212> DNA

<213> 荚膜红细菌 (Rhodobacter capsulatus) DE442

<400> 189

gcctcacatc accgccaaga cgacggcgga ctgaac 36

<210> 190

<211> 1389

<212> PRT

<213> 未知的

<220>

<223> Cas13a

<400> 190

Met Gly Asn Leu Phe Gly His Lys Arg Trp Tyr Glu Val Arg Asp Lys

1 5 10 15

Lys Asp Phe Lys Ile Lys Arg Lys Val Lys Val Lys Arg Asn Tyr Asp

20 25 30

Gly Asn Lys Tyr Ile Leu Asn Ile Asn Glu Asn Asn Asn Lys Glu Lys

35 40 45

Ile Asp Asn Asn Lys Phe Ile Arg Lys Tyr Ile Asn Tyr Lys Lys Asn

50 55 60

Asp Asn Ile Leu Lys Glu Phe Thr Arg Lys Phe His Ala Gly Asn Ile

65 70 75 80

Leu Phe Lys Leu Lys Gly Lys Glu Gly Ile Ile Arg Ile Glu Asn Asn

85 90 95

Asp Asp Phe Leu Glu Thr Glu Glu Val Val Leu Tyr Ile Glu Ala Tyr

100 105 110

Gly Lys Ser Glu Lys Leu Lys Ala Leu Gly Ile Thr Lys Lys Lys Ile

115 120 125

Ile Asp Glu Ala Ile Arg Gln Gly Ile Thr Lys Asp Asp Lys Lys Ile

130 135 140

Glu Ile Lys Arg Gln Glu Asn Glu Glu Glu Ile Glu Ile Asp Ile Arg			
145	150	155	160
Asp Glu Tyr Thr Asn Lys Thr Leu Asn Asp Cys Ser Ile Ile Leu Arg			
	165	170	175
Ile Ile Glu Asn Asp Glu Leu Glu Thr Lys Lys Ser Ile Tyr Glu Ile			
	180	185	190
Phe Lys Asn Ile Asn Met Ser Leu Tyr Lys Ile Ile Glu Lys Ile Ile			
	195	200	205
Glu Asn Glu Thr Glu Lys Val Phe Glu Asn Arg Tyr Tyr Glu Glu His			
	210	215	220
Leu Arg Glu Lys Leu Leu Lys Asp Asp Lys Ile Asp Val Ile Leu Thr			
225	230	235	240
Asn Phe Met Glu Ile Arg Glu Lys Ile Lys Ser Asn Leu Glu Ile Leu			
	245	250	255
Gly Phe Val Lys Phe Tyr Leu Asn Val Gly Gly Asp Lys Lys Lys Ser			
	260	265	270
Lys Asn Lys Lys Met Leu Val Glu Lys Ile Leu Asn Ile Asn Val Asp			
	275	280	285
Leu Thr Val Glu Asp Ile Ala Asp Phe Val Ile Lys Glu Leu Glu Phe			
	290	295	300
Trp Asn Ile Thr Lys Arg Ile Glu Lys Val Lys Lys Val Asn Asn Glu			
305	310	315	320
Phe Leu Glu Lys Arg Arg Asn Arg Thr Tyr Ile Lys Ser Tyr Val Leu			
	325	330	335
Leu Asp Lys His Glu Lys Phe Lys Ile Glu Arg Glu Asn Lys Lys Asp			
	340	345	350
Lys Ile Val Lys Phe Phe Val Glu Asn Ile Lys Asn Asn Ser Ile Lys			
	355	360	365
Glu Lys Ile Glu Lys Ile Leu Ala Glu Phe Lys Ile Asp Glu Leu Ile			
	370	375	380
Lys Lys Leu Glu Lys Glu Leu Lys Lys Gly Asn Cys Asp Thr Glu Ile			
385	390	395	400
Phe Gly Ile Phe Lys Lys His Tyr Lys Val Asn Phe Asp Ser Lys Lys			
	405	410	415
Phe Ser Lys Lys Ser Asp Glu Glu Lys Glu Leu Tyr Lys Ile Ile Tyr			
	420	425	430
Arg Tyr Leu Lys Gly Arg Ile Glu Lys Ile Leu Val Asn Glu Gln Lys			
	435	440	445
Val Arg Leu Lys Lys Met Glu Lys Ile Glu Ile Glu Lys Ile Leu Asn			

450	455	460
Glu Ser Ile Leu Ser Glu Lys Ile Leu Lys Arg Val Lys Gln Tyr Thr		
465	470	475
Leu Glu His Ile Met Tyr Leu Gly Lys Leu Arg His Asn Asp Ile Asp		
485	490	495
Met Thr Thr Val Asn Thr Asp Asp Phe Ser Arg Leu His Ala Lys Glu		
500	505	510
Glu Leu Asp Leu Glu Leu Ile Thr Phe Phe Ala Ser Thr Asn Met Glu		
515	520	525
Leu Asn Lys Ile Phe Ser Arg Glu Asn Ile Asn Asn Asp Glu Asn Ile		
530	535	540
Asp Phe Phe Gly Gly Asp Arg Glu Lys Asn Tyr Val Leu Asp Lys Lys		
545	550	555
Ile Leu Asn Ser Lys Ile Lys Ile Ile Arg Asp Leu Asp Phe Ile Asp		
565	570	575
Asn Lys Asn Asn Ile Thr Asn Asn Phe Ile Arg Lys Phe Thr Lys Ile		
580	585	590
Gly Thr Asn Glu Arg Asn Arg Ile Leu His Ala Ile Ser Lys Glu Arg		
595	600	605
Asp Leu Gln Gly Thr Gln Asp Asp Tyr Asn Lys Val Ile Asn Ile Ile		
610	615	620
Gln Asn Leu Lys Ile Ser Asp Glu Glu Val Ser Lys Ala Leu Asn Leu		
625	630	635
Asp Val Val Phe Lys Asp Lys Lys Asn Ile Ile Thr Lys Ile Asn Asp		
645	650	655
Ile Lys Ile Ser Glu Glu Asn Asn Asn Asp Ile Lys Tyr Leu Pro Ser		
660	665	670
Phe Ser Lys Val Leu Pro Glu Ile Leu Asn Leu Tyr Arg Asn Asn Pro		
675	680	685
Lys Asn Glu Pro Phe Asp Thr Ile Glu Thr Glu Lys Ile Val Leu Asn		
690	695	700
Ala Leu Ile Tyr Val Asn Lys Glu Leu Tyr Lys Lys Leu Ile Leu Glu		
705	710	715
Asp Asp Leu Glu Glu Asn Glu Ser Lys Asn Ile Phe Leu Gln Glu Leu		
725	730	735
Lys Lys Thr Leu Gly Asn Ile Asp Glu Ile Asp Glu Asn Ile Ile Glu		
740	745	750
Asn Tyr Tyr Lys Asn Ala Gln Ile Ser Ala Ser Lys Gly Asn Asn Lys		
755	760	765

Ala Ile Lys Lys Tyr Gln Lys Lys Val Ile Glu Cys Tyr Ile Gly Tyr		
770	775	780
Leu Arg Lys Asn Tyr Glu Glu Leu Phe Asp Phe Ser Asp Phe Lys Met		
785	790	795 800
Asn Ile Gln Glu Ile Lys Lys Gln Ile Lys Asp Ile Asn Asp Asn Lys		
805	810	815
Thr Tyr Glu Arg Ile Thr Val Lys Thr Ser Asp Lys Thr Ile Val Ile		
820	825	830
Asn Asp Asp Phe Glu Tyr Ile Ile Ser Ile Phe Ala Leu Leu Asn Ser		
835	840	845
Asn Ala Val Ile Asn Lys Ile Arg Asn Arg Phe Phe Ala Thr Ser Val		
850	855	860
Trp Leu Asn Thr Ser Glu Tyr Gln Asn Ile Ile Asp Ile Leu Asp Glu		
865	870	875 880
Ile Met Gln Leu Asn Thr Leu Arg Asn Glu Cys Ile Thr Glu Asn Trp		
885	890	895
Asn Leu Asn Leu Glu Glu Phe Ile Gln Lys Met Lys Glu Ile Glu Lys		
900	905	910
Asp Phe Asp Asp Phe Lys Ile Gln Thr Lys Lys Glu Ile Phe Asn Asn		
915	920	925
Tyr Tyr Glu Asp Ile Lys Asn Asn Ile Leu Thr Glu Phe Lys Asp Asp		
930	935	940
Ile Asn Gly Cys Asp Val Leu Glu Lys Lys Leu Glu Lys Ile Val Ile		
945	950	955 960
Phe Asp Asp Glu Thr Lys Phe Glu Ile Asp Lys Lys Ser Asn Ile Leu		
965	970	975
Gln Asp Glu Gln Arg Lys Leu Ser Asn Ile Asn Lys Lys Asp Leu Lys		
980	985	990
Lys Lys Val Asp Gln Tyr Ile Lys Asp Lys Asp Gln Glu Ile Lys Ser		
995	1000	1005
Lys Ile Leu Cys Arg Ile Ile Phe Asn Ser Asp Phe Leu Lys Lys		
1010	1015	1020
Tyr Lys Lys Glu Ile Asp Asn Leu Ile Glu Asp Met Glu Ser Glu		
1025	1030	1035
Asn Glu Asn Lys Phe Gln Glu Ile Tyr Tyr Pro Lys Glu Arg Lys		
1040	1045	1050
Asn Glu Leu Tyr Ile Tyr Lys Lys Asn Leu Phe Leu Asn Ile Gly		
1055	1060	1065
Asn Pro Asn Phe Asp Lys Ile Tyr Gly Leu Ile Ser Asn Asp Ile		

1070	1075	1080
Lys Met Ala Asp Ala Lys	Phe Leu Phe Asn Ile	Asp Gly Lys Asn
1085	1090	1095
Ile Arg Lys Asn Lys Ile	Ser Glu Ile Asp Ala	Ile Leu Lys Asn
1100	1105	1110
Leu Asn Asp Lys Leu Asn	Gly Tyr Ser Lys Glu	Tyr Lys Glu Lys
1115	1120	1125
Tyr Ile Lys Lys Leu Lys	Glu Asn Asp Asp Phe	Phe Ala Lys Asn
1130	1135	1140
Ile Gln Asn Lys Asn Tyr	Lys Ser Phe Glu Lys	Asp Tyr Asn Arg
1145	1150	1155
Val Ser Glu Tyr Lys Lys	Ile Arg Asp Leu Val	Glu Phe Asn Tyr
1160	1165	1170
Leu Asn Lys Ile Glu Ser	Tyr Leu Ile Asp Ile	Asn Trp Lys Leu
1175	1180	1185
Ala Ile Gln Met Ala Arg	Phe Glu Arg Asp Met	His Tyr Ile Val
1190	1195	1200
Asn Gly Leu Arg Glu Leu	Gly Ile Ile Lys Leu	Ser Gly Tyr Asn
1205	1210	1215
Thr Gly Ile Ser Arg Ala	Tyr Pro Lys Arg Asn	Gly Ser Asp Gly
1220	1225	1230
Phe Tyr Thr Thr Thr Ala	Tyr Tyr Lys Phe Phe	Asp Glu Glu Ser
1235	1240	1245
Tyr Lys Lys Phe Glu Lys	Ile Cys Tyr Gly Phe	Gly Ile Asp Leu
1250	1255	1260
Ser Glu Asn Ser Glu Ile	Asn Lys Pro Glu Asn	Glu Ser Ile Arg
1265	1270	1275
Asn Tyr Ile Ser His Phe	Tyr Ile Val Arg Asn	Pro Phe Ala Asp
1280	1285	1290
Tyr Ser Ile Ala Glu Gln	Ile Asp Arg Val Ser	Asn Leu Leu Ser
1295	1300	1305
Tyr Ser Thr Arg Tyr Asn	Asn Ser Thr Tyr Ala	Ser Val Phe Glu
1310	1315	1320
Val Phe Lys Lys Asp Val	Asn Leu Asp Tyr Asp	Glu Leu Lys Lys
1325	1330	1335
Lys Phe Lys Leu Ile Gly	Asn Asn Asp Ile Leu	Glu Arg Leu Met
1340	1345	1350
Lys Pro Lys Lys Val Ser	Val Leu Glu Leu Glu	Ser Tyr Asn Ser
1355	1360	1365

Asp Tyr Ile Lys Asn Leu Ile Ile Glu Leu Leu Thr Lys Ile Glu
 1370 1375 1380
 Asn Thr Asn Asp Thr Leu
 1385
 <210> 191
 <211> 1221
 <212> PRT
 <213> 动物溃疡伯格菌 (*Bergeyella zoohelcum*) ATCC 43767
 <400> 191
 Met Glu Asn Lys Thr Ser Leu Gly Asn Asn Ile Tyr Tyr Asn Pro Phe
 1 5 10 15
 Lys Pro Gln Asp Lys Ser Tyr Phe Ala Gly Tyr Phe Asn Ala Ala Met
 20 25 30
 Glu Asn Thr Asp Ser Val Phe Arg Glu Leu Gly Lys Arg Leu Lys Gly
 35 40 45
 Lys Glu Tyr Thr Ser Glu Asn Phe Phe Asp Ala Ile Phe Lys Glu Asn
 50 55 60
 Ile Ser Leu Val Glu Tyr Glu Arg Tyr Val Lys Leu Leu Ser Asp Tyr
 65 70 75 80
 Phe Pro Met Ala Arg Leu Leu Asp Lys Lys Glu Val Pro Ile Lys Glu
 85 90 95
 Arg Lys Glu Asn Phe Lys Lys Asn Phe Lys Gly Ile Ile Lys Ala Val
 100 105 110
 Arg Asp Leu Arg Asn Phe Tyr Thr His Lys Glu His Gly Glu Val Glu
 115 120 125
 Ile Thr Asp Glu Ile Phe Gly Val Leu Asp Glu Met Leu Lys Ser Thr
 130 135 140
 Val Leu Thr Val Lys Lys Lys Lys Val Lys Thr Asp Lys Thr Lys Glu
 145 150 155 160
 Ile Leu Lys Lys Ser Ile Glu Lys Gln Leu Asp Ile Leu Cys Gln Lys
 165 170 175
 Lys Leu Glu Tyr Leu Arg Asp Thr Ala Arg Lys Ile Glu Glu Lys Arg
 180 185 190
 Arg Asn Gln Arg Glu Arg Gly Glu Lys Glu Leu Val Ala Pro Phe Lys
 195 200 205
 Tyr Ser Asp Lys Arg Asp Asp Leu Ile Ala Ala Ile Tyr Asn Asp Ala
 210 215 220
 Phe Asp Val Tyr Ile Asp Lys Lys Lys Asp Ser Leu Lys Glu Ser Ser
 225 230 235 240

Lys	Ala	Lys	Tyr	Asn	Thr	Lys	Ser	Asp	Pro	Gln	Gln	Glu	Glu	Gly	Asp		
245						250					255						
Leu	Lys	Ile	Pro	Ile	Ser	Lys	Asn	Gly	Val	Val	Phe	Leu	Leu	Ser	Leu		
260						265					270						
Phe	Leu	Thr	Lys	Gln	Glu	Ile	His	Ala	Phe	Lys	Ser	Lys	Ile	Ala	Gly		
275						280					285						
Phe	Lys	Ala	Thr	Val	Ile	Asp	Glu	Ala	Thr	Val	Ser	Glu	Ala	Thr	Val		
290						295					300						
Ser	His	Gly	Lys	Asn	Ser	Ile	Cys	Phe	Met	Ala	Thr	His	Glu	Ile	Phe		
305						310					315					320	
Ser	His	Leu	Ala	Tyr	Lys	Lys	Leu	Lys	Arg	Lys	Val	Arg	Thr	Ala	Glu		
325						330					335						
Ile	Asn	Tyr	Gly	Glu	Ala	Glu	Asn	Ala	Glu	Gln	Leu	Ser	Val	Tyr	Ala		
340						345					350						
Lys	Glu	Thr	Leu	Met	Met	Gln	Met	Leu	Asp	Glu	Leu	Ser	Lys	Val	Pro		
355						360					365						
Asp	Val	Val	Tyr	Gln	Asn	Leu	Ser	Glu	Asp	Val	Gln	Lys	Thr	Phe	Ile		
370						375					380						
Glu	Asp	Trp	Asn	Glu	Tyr	Leu	Lys	Glu	Asn	Asn	Gly	Asp	Val	Gly	Thr		
385						390					395					400	
Met	Glu	Glu	Glu	Gln	Val	Ile	His	Pro	Val	Ile	Arg	Lys	Arg	Tyr	Glu		
405						410					415						
Asp	Lys	Phe	Asn	Tyr	Phe	Ala	Ile	Arg	Phe	Leu	Asp	Glu	Phe	Ala	Gln		
420						425					430						
Phe	Pro	Thr	Leu	Arg	Phe	Gln	Val	His	Leu	Gly	Asn	Tyr	Leu	His	Asp		
435						440					445						
Ser	Arg	Pro	Lys	Glu	Asn	Leu	Ile	Ser	Asp	Arg	Arg	Ile	Lys	Glu	Lys		
450						455					460						
Ile	Thr	Val	Phe	Gly	Arg	Leu	Ser	Glu	Leu	Glu	His	Lys	Lys	Ala	Leu		
465						470					475					480	
Phe	Ile	Lys	Asn	Thr	Glu	Thr	Asn	Glu	Asp	Arg	Glu	His	Tyr	Trp	Glu		
485						490					495						
Ile	Phe	Pro	Asn	Pro	Asn	Tyr	Asp	Phe	Pro	Lys	Glu	Asn	Ile	Ser	Val		
500						505					510						
Asn	Asp	Lys	Asp	Phe	Pro	Ile	Ala	Gly	Ser	Ile	Leu	Asp	Arg	Glu	Lys		
515						520					525						
Gln	Pro	Val	Ala	Gly	Lys	Ile	Gly	Ile	Lys	Val	Lys	Leu	Leu	Asn	Gln		
530						535					540						
Gln	Tyr	Val	Ser	Glu	Val	Asp	Lys	Ala	Val	Lys	Ala	His	Gln	Leu	Lys		

545	550	555	560
Gln Arg Lys Ala Ser Lys Pro Ser Ile Gln Asn Ile Ile Glu Glu Ile			
	565	570	575
Val Pro Ile Asn Glu Ser Asn Pro Lys Glu Ala Ile Val Phe Gly Gly			
	580	585	590
Gln Pro Thr Ala Tyr Leu Ser Met Asn Asp Ile His Ser Ile Leu Tyr			
	595	600	605
Glu Phe Phe Asp Lys Trp Glu Lys Lys Lys Glu Lys Leu Glu Lys Lys			
	610	615	620
Gly Glu Lys Glu Leu Arg Lys Glu Ile Gly Lys Glu Leu Glu Lys Lys			
625	630	635	640
Ile Val Gly Lys Ile Gln Ala Gln Ile Gln Gln Ile Ile Asp Lys Asp			
	645	650	655
Thr Asn Ala Lys Ile Leu Lys Pro Tyr Gln Asp Gly Asn Ser Thr Ala			
	660	665	670
Ile Asp Lys Glu Lys Leu Ile Lys Asp Leu Lys Gln Glu Gln Asn Ile			
	675	680	685
Leu Gln Lys Leu Lys Asp Glu Gln Thr Val Arg Glu Lys Glu Tyr Asn			
	690	695	700
Asp Phe Ile Ala Tyr Gln Asp Lys Asn Arg Glu Ile Asn Lys Val Arg			
705	710	715	720
Asp Arg Asn His Lys Gln Tyr Leu Lys Asp Asn Leu Lys Arg Lys Tyr			
	725	730	735
Pro Glu Ala Pro Ala Arg Lys Glu Val Leu Tyr Tyr Arg Glu Lys Gly			
	740	745	750
Lys Val Ala Val Trp Leu Ala Asn Asp Ile Lys Arg Phe Met Pro Thr			
	755	760	765
Asp Phe Lys Asn Glu Trp Lys Gly Glu Gln His Ser Leu Leu Gln Lys			
	770	775	780
Ser Leu Ala Tyr Tyr Glu Gln Cys Lys Glu Glu Leu Lys Asn Leu Leu			
785	790	795	800
Pro Glu Lys Val Phe Gln His Leu Pro Phe Lys Leu Gly Gly Tyr Phe			
	805	810	815
Gln Gln Lys Tyr Leu Tyr Gln Phe Tyr Thr Cys Tyr Leu Asp Lys Arg			
	820	825	830
Leu Glu Tyr Ile Ser Gly Leu Val Gln Gln Ala Glu Asn Phe Lys Ser			
	835	840	845
Glu Asn Lys Val Phe Lys Lys Val Glu Asn Glu Cys Phe Lys Phe Leu			
	850	855	860

Lys	Lys	Gln	Asn	Tyr	Thr	His	Lys	Glu	Leu	Asp	Ala	Arg	Val	Gln	Ser
865						870				875					880
Ile	Leu	Gly	Tyr	Pro	Ile	Phe	Leu	Glu	Arg	Gly	Phe	Met	Asp	Glu	Lys
						885				890					895
Pro	Thr	Ile	Ile	Lys	Gly	Lys	Thr	Phe	Lys	Gly	Asn	Glu	Ala	Leu	Phe
						900				905					910
Ala	Asp	Trp	Phe	Arg	Tyr	Tyr	Lys	Glu	Tyr	Gln	Asn	Phe	Gln	Thr	Phe
						915				920					925
Tyr	Asp	Thr	Glu	Asn	Tyr	Pro	Leu	Val	Glu	Leu	Glu	Lys	Lys	Gln	Ala
						930									940
Asp	Arg	Lys	Arg	Lys	Thr	Lys	Ile	Tyr	Gln	Gln	Lys	Lys	Asn	Asp	Val
945						950									960
Phe	Thr	Leu	Leu	Met	Ala	Lys	His	Ile	Phe	Lys	Ser	Val	Phe	Lys	Gln
						965									975
Asp	Ser	Ile	Asp	Gln	Phe	Ser	Leu	Glu	Asp	Leu	Tyr	Gln	Ser	Arg	Glu
						980									990
Glu	Arg	Leu	Gly	Asn	Gln	Glu	Arg	Ala	Arg	Gln	Thr	Gly	Glu	Arg	Asn
						995									1005
Thr	Asn	Tyr	Ile	Trp	Asn	Lys	Thr	Val	Asp	Leu	Lys	Leu	Cys	Asp	
						1010									1020
Gly	Lys	Ile	Thr	Val	Glu	Asn	Val	Lys	Leu	Lys	Asn	Val	Gly	Asp	
						1025									1035
Phe	Ile	Lys	Tyr	Glu	Tyr	Asp	Gln	Arg	Val	Gln	Ala	Phe	Leu	Lys	
						1040									1050
Tyr	Glu	Glu	Asn	Ile	Glu	Trp	Gln	Ala	Phe	Leu	Ile	Lys	Glu	Ser	
						1055									1065
Lys	Glu	Glu	Glu	Asn	Tyr	Pro	Tyr	Val	Val	Glu	Arg	Glu	Ile	Glu	
						1070									1080
Gln	Tyr	Glu	Lys	Val	Arg	Arg	Glu	Glu	Leu	Leu	Lys	Glu	Val	His	
						1085									1095
Leu	Ile	Glu	Glu	Tyr	Ile	Leu	Glu	Lys	Val	Lys	Asp	Lys	Glu	Ile	
						1100									1110
Leu	Lys	Lys	Gly	Asp	Asn	Gln	Asn	Phe	Lys	Tyr	Tyr	Ile	Leu	Asn	
						1115									1125
Gly	Leu	Leu	Lys	Gln	Leu	Lys	Asn	Glu	Asp	Val	Glu	Ser	Tyr	Lys	
						1130									1140
Val	Phe	Asn	Leu	Asn	Thr	Glu	Pro	Glu	Asp	Val	Asn	Ile	Asn	Gln	
						1145									1155
Leu	Lys	Gln	Glu	Ala	Thr	Asp	Leu	Glu	Gln	Lys	Ala	Phe	Val	Leu	

1160	1165	1170
Thr Tyr Ile Arg Asn Lys	Phe Ala His Asn Gln	Leu Pro Lys Lys
1175	1180	1185
Glu Phe Trp Asp Tyr Cys	Gln Glu Lys Tyr Gly	Lys Glu Lys Thr
1190	1195	1200
Tyr Ala Glu Tyr Phe Ala	Glu Val Phe Lys Lys	Glu Lys Glu Ala
1205	1210	1215
Leu Ile Lys		

1220

<210> 192

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> GAPDH F引物

<400> 192

cagcctcaag atcatcagca a 21

<210> 193

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> GAPDH R引物

<400> 193

tgttggtcatg agtccttcca 20

<210> 194

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> NS5 F引物

<400> 194

gaggagagtg ccagagttgt 20

<210> 195

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> NS5 R引物

<400> 195
tctctctccc catccagtga 20
<210> 196
<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 靶向NS5的间隔子1
<400> 196
gcaatgatct tcatgttggg agc 23
<210> 197
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 靶向NS5的间隔子2
<400> 197
gaaccttggt gatgaactct tc 22
<210> 198
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 靶向NS5的间隔子3
<400> 198
gttggtgatt agagcttcat tc 22
<210> 199
<211> 23
<212> DNA
<213> 人工序列
<220>
<223> 靶向NS5的间隔子4
<400> 199
gagtgatcct cgttcaagaa tcc 23
<210> 200
<211> 20
<212> DNA
<213> 人工序列
<220>

<223> 非靶向对照间隔子(Δ2)

<400> 200

gtgataagtg gaatgccatg 20

<210> 201

<211> 114

<212> DNA

<213> 人工序列

<220>

<223> sgRNA支架

<220>

<221> 尚未归类的特征

<222> (2) .. (21)

<223> n是a、c、g、t或u

<400> 201

gnnnnnnnnnn nnnnnnnnnnn nguuaaagag cuaugcugga aacagcauag caaguuuaaa 60

uaaggcuagu ccguuaucac cuugaaaaag uggcaccgag ucggugcuuu uuuu 114

<210> 202

<211> 4991

<212> DNA

<213> 人工序列

<220>

<223> E43-CjeCas9和sgRNA质粒

<220>

<221> 尚未归类的特征

<222> (297) .. (315)

<223> n是a、c、g、t或u

<400> 202

gtttattaca gggacagcag agatccagtt tggttaatta aggtaccgag ggcctatttc 60

ccatgattcc ttcatatttg catatacgat acaagctgt tagagagata attagaatta 120

atttgactgt aaacacaaaag atattagtagt aaaatacgtg acgtagaaaag taataatttc 180

ttgggtagtt tgcagtttta aaattatggt ttaaaatgga ctatcatatg cttaccgtaa 240

cttgaaagta tttcgatttc ttggctttat atatcttggt gaaaggacga aacaccnnnn 300

nnnnnnnnnn nnnnnngtttt agtcctgaa gggactaaaa taaagagttt gcgggactct 360

gcgggggttac aatcccctaa aaccgctttt tttcctgcag cccgggggat ccactagtgc 420

tagagcggcc gccaccgcgg tggagctcca gcttttggtt ccttttagtga gggttaattg 480

cgcgaattcg ctagctaggt cttgaaagga gtgggaattg gctccggtgc ccgtcagtgg 540

gcagagcgca catcgccac agtccccgag aagttggggg gaggggtcgg caattgatcc 600

gggtgcctaga gaaggtggcg cggggtaaag tgggaaagt atgtcgtgta ctggctccgc 660

ctttttcccg aggggtggggg agaaccgtat ataagtgcag tagtcgccgt gaacgttctt 720

```

tttcgcaacg ggtttgccgc cagaacacag gaccggttct agagcgctat ttagaaccat 780
gtgttctccc caagaatctg gcatgaccgc tctttcagcg aggatgttga cgcgaaagcag 840
atccctggga cctggggccg ggccacgagg gtgtcgggaa gaaccaggac cgttgcgacg 900
gaggggaagca gcagcggaag ctcggaatc ccattctccg gttaaagcag cccgcaaggc 960
acaacggctc agggttgctt acgaggggag cgattccgaa aagggtgaag gagcagagcc 1020
cttgaagggtt ccagtatggg aaccccagga ttggcagcag cagcttgtaa acatccgagc 1080
aatgaggaac aaaaaagatg cacctgttga tcacctcgga accgaacatt gttatgattc 1140
tagtgcgccg ccaaaaagtcc gccggtatca gtttctgttg agtttgatgc tgagtagtca 1200
gactaaggac caggttacgg ccggagcaat gcaacggctt cgggcacggg gactcacggt 1260
cgatagcatt ttgcagaccg atgacgcaac attgggtaaa ctcatatata cagttggctt 1320
ctggcggagc aaagtgaagt acatcaagca gacctcagcc attctccaac aacattacgg 1380
aggtgatata cccgcaagcg tagctgaact ggtagcactg cggggcgctg gtcccaaaat 1440
ggcacatctg gctatggcgg ttgcttgggg aacggtgtct ggtatcgagc ttgatacgca 1500
tgtccaccgc atcgccaatc ggctgaggtg gactaaaaaa gccactaagt ctctgaaga 1560
aacacgggct gctctggaag agtggttcc acgagagctg tggcatgaaa tcaatggatt 1620
gctggttggg ttcgggcagc agacatgctt gcccggtgcac ccccggtgtc atgcttgctt 1680
gaaccaggct ttgtgcccag ctgcccaggg cctgagtgga agtgagacac cgggaacatc 1740
tgagtctgag accccggaga gcacaaacgc gcgaatcctg gccttcgca ttggcattag 1800
cagcatcggc tgggcattct ctgaaaacga cgaactgaag gattgcggcg tgcgaatttt 1860
cactaaggct gaaaatccca aaactggtga atcactcgct ctccctagac gactggcacg 1920
ctccgcacga aagaggcttg cccgccgcaa ggcacgcttg aaccatctta aacaccttat 1980
tgcaaatgag tttaaactga attatgagga ctaccaatcc tttgacgagt ctcttgctaa 2040
agcctacaaa gggagcctta tatccccgta tgagctccgg ttcagagcac tcaacgaact 2100
gctgtccaaa caggattttg ctgcgctgat tctccacata gcgaagaggc gaggatacga 2160
tgacattaaa aacagtgatg ataaggaaaa agggggcata ctcaaagcga ttaagcaaaa 2220
tgaagagaag ctgcctaact atcaatcagt aggggagtat ctctataaag agtacttcca 2280
gaagttcaaa gaaaatagca aggaatttac taatgtccgg aataaaaagg agtcttacga 2340
aagatgtatt gcgcaatctt tcctcaagga cgagctcaaa ttgattttca agaaacaaag 2400
ggaatttggg ttcagcttct caaaaaaatt tgaggaagag gttctgagcg ttgcctttta 2460
caaacgcgcc cttaaggact tctcacatct cgtagggaat tgtagtttct tcaccgatga 2520
aaaacgggcg ccaaaaaata gccctttggc ttttatgttt gtcgctctga ctgcgcatcat 2580
taatctgctc aacaacctta aaaacacgga agggattctg tacacaaagg atgatctgaa 2640
cgctctgctt aacgaagttt tgaagaacgg gactttgacc tacaacaaa ccaaaaagct 2700
tcttggctc agtgatgact acgaattcaa gggagaaaaa gggacatatt tcatcgaatt 2760
caagaagtat aaggagtcca tcaaagcctt gggcgagcac aacttgctc aagatgatct 2820
caacgaaatt gctaaggata tcaactctgat taaagacgag atcaagctca aaaaggcgtt 2880
ggcgaagtat gaccttaacc aaaaccaa atagatgcctc agcaagttgg aatttaaaga 2940
tcaactgaat ataagtttca aggcccttaa gttggtcacc cccttgatgc ttgaaggaaa 3000
gaaatatgat gaggcattga atgagctgaa tctcaagggt gctattaacg aagacaaaaa 3060

```

```

agatttcctc ccagctttca atgagactta ctataaggac gaggttacca atcctgtggt 3120
gctccgagcc atcaaagagt atcgaaaggt cctgaatgct ttgctcaaaa aatacggtaa 3180
ggtacacaaa ataaatattg agctcgcaag ggaggtcggt aagaaccact cccagcgcg 3240
caaaaatagaa aaggaacaga atgaaaatta caaagcgaaa aaggacgccg agctcgagt 3300
cgaaaagctg ggcctgaaaa taaacagcaa gaacattctc aaactccgcc tcttcaaaga 3360
acaaaaagaa ttttgtgctt atagtgggtg gaaaataaaa atctccgatc ttcaagacga 3420
gaagatgctc gaaaatagacg cgatatatcc atatagcagg tcttttgacg attcttacat 3480
gaataaagt  cttgttttca ctaagcagaa tcaggaaaag ttgaatcaga ccccttttga 3540
ggccttttggc aacgactcag caaagtggca gaagatcgag gtcttggcta agaactcttc 3600
tactaagaaa cagaaaagga tattggataa gaactataaa gacaaagaac aaaagaactt 3660
taaagaccgc aacctcaatg acaccagata catagcaaga ttggttctga actacacaaa 3720
agattatttg gacttcttgc cgtgtcttga tgatgagaac acgaaactca acgacacgca 3780
aaaggggtct aaagtccacg tcgaagctaa atctgggatg ctcacctcag cattgaggca 3840
tacgtgggga ttctcagcaa aggaccgaaa caatcacctg caccatgcca ttgacgcagt 3900
tatcatagcg tatgccaata attcaatagt aaaagcgttt agcgacttca agaaggaaca 3960
agagtccaac agcgccgagc tctacgcaaa aaagattagt gaactcgact acaaaaacaa 4020
aagaaaattc tttgagccgt tcagcgatt tcgacagaag gtattggata aaatagatga 4080
aattttcgtg agcaaaccg aaaggaaaaa gccctcaggc gccttgacg aagagacttt 4140
caggaaggaa gaggaattct accaaagcta cggcgaaaaa gagggagttt tgaaggctct 4200
cgaacttgga aagattagga aggtgaacgg caagatagtg aaaaacggcg atatgttccg 4260
ggttgatata ttcaaacata aaaaaacgaa taaattttat gctgtgccta tatacactat 4320
ggacttcgca ctttaaggctc tgccgaataa ggcggtagcc cgatctaaaa aaggcgaaat 4380
taaggactgg attttgatgg atgaaaatta cgagtcttgc ttttctctct acaaggattc 4440
ccttatattg atacagacga aagatatgca ggaaccggaa ttcgtgtatt acaacgcttt 4500
tacttcctct acggtatctt tgattgtctc caaacatgac aacaaattcg aaacactcag 4560
taaaaaccaa aagattctct ttaaaaatgc gaacgagaaa gaagtaattg caaaatcaat 4620
tggtcatcaa aatttgaaag tttttgaaaa atatatagta tctgccctcg gagaggttac 4680
taaagcggaa tttagacagc gagaggactt caaaaaatca ggtccacca agaaaaaacg 4740
caaggtggaa gatccgaaga aaaagcgaaa agtggatgtg taacgttttc cgggacgccg 4800
gctggatgat cctccagcg ggggatctca tgctggagtt cttcgccac cccaacttgt 4860
ttattgcagc ttataatggt tacaaataaa gcaatagcat cacaaatttc acaataaag 4920
catttttttc actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg 4980
tctgtatacc g 4991

```

<210> 203

<211> 5270

<212> DNA

<213> 人工序列

<220>

<223> E67-CjeCas9和sgRNA质粒

<220>

<221> 尚未归类的特征

<222> (297) .. (315)

<223> n是a、c、g、t或u

<400> 203

```
gtttattaca gggacagcag agatccagtt tggtaatta aggtaccgag ggcctatttc 60
ccatgattcc ttcatatttg catatacgat acaaggctgt tagagagata attagaatta 120
atttgactgt aaacacaaaag atattagtac aaaatacgtg acgtagaaaag taataatttc 180
ttgggtagtt tgcagtttta aaattatgtt ttaaaatgga ctatcatatg cttaccgtaa 240
cttgaaaagta tttcgatttc ttggctttat atatcttggtg gaaaggacga aacaccnnnn 300
nnnnnnnnnn nnnnnngtttt agtcctgaa gggactaaaa taaagagttt gcgggactct 360
gcgggggttac aatcccctaa aaccgctttt tttcctgcag cccgggggat ccactagttc 420
tagagcggcc gccaccgcgg tggagctcca gcttttggtc ccttttagtga gggtaattg 480
cgcaattcg ctagctaggt cttgaaagga gtgggaattg gctccggtgc ccgtcagtgg 540
gcagagcgca catcgccac agtccccgag aagttggggg gaggggtcgg caattgatcc 600
ggtgcctaga gaagggtggcg cggggtaaac tgggaaagtg atgtcgtgta ctggctccgc 660
ctttttcccg aggggtgggg agaaccgtat ataagtgcag tagtcgccgt gaacgttctt 720
tttcgcaacg ggtttgccgc cagaacacag gaccggttct agagcgctat ttagaacat 780
gcaggaggta atagcggggc ttgagcgatt tacctttgcc ttcgaaaaag acgtagagat 840
gcagaaggga accggcctgc tcccatttca aggtatggac aaatcagcat ctgccgtgtg 900
caattttttc accaagggtc tgtgtgaaaa ggggaagctc tgtccatttc gccatgatcg 960
cggagagaag atggtggtgt gtaagcactg gctgagaggg ctttgcaaaa aaggcgacca 1020
ctgcaaatTT cttaccaat atgacctgac tcgaatgcct gagtgttatt tttacagtaa 1080
gttcggtgac ttagcaaca aagaatgcag cttcttgcat gtcaaaccag cattcaagtc 1140
acaggattgc ccgtggtacg atcagggttt ttgcaaggac ggtccccctc gcaaatatcg 1200
acacgtaccc agaattatgt gccttaatta cctggtcggc ttctgtcctg aagggccaaa 1260
atgtcagttt gctcaaaaaa ttcgcgagtt caaattgctc cctgggtcta aaatttgga 1320
accccaggat tggcagcagc agcttgtaaa catccgagca atgaggaaca aaaaagatgc 1380
acctgttgat cacctcggaa ccgaacattg ttatgattct agtgcgccgc caaaagtccg 1440
ccggtatcag gttctgttga gtttgatgct gagtagtcag actaaggacc aggttacggc 1500
cggagcaatg caacggcttc gggcacgggg actcacggtc gatagcattt tgcagaccga 1560
tgacgcaaca ttgggtaaac tcatatatcc agttggcttc tggcggagca aagtgaagta 1620
catcaagcag acctcagcca ttctccaaca acattacgga ggtgatatac ccgcaagcgt 1680
agctgaactg gtagcactgc cgggcgtcgg tcccaaatg gcacatctgg ctatggcggt 1740
tgcttgggga acggtgtctg gtatcgcagt tgatacgcgt gtccaccgca tcgccaatcg 1800
gctgagggtg actaaaaaag ccactaagtc tcctgaagaa acacgggctg ctctggaaga 1860
gtggcttcca cgagagctgt ggcatgaaat caatggattg ctggttggtt tcgggcagca 1920
gacatgcttg cccgtgcacc cccggtgtca tgcttgcttg aaccaggctt tgtgccagc 1980
tgcccagggc ctgagtggaa gtgagacacc gggaacatct gagtctgcga ccccgagag 2040
```

cacaaacgcg cgaatcctgg ccttcgcat tggcattagc agcatcggct gggcattctc 2100
tgaaaacgac gaactgaagg attgcggcgt gcgaattttc actaaggtcg aaaatcccaa 2160
aactgggtgaa tcactcgctc tccctagacg actggcacgc tccgcacgaa agaggcttgc 2220
ccgccgcaag gcacgcttga accatcttaa acaccttatt gcaaagtagt ttaaactgaa 2280
ttatgaggac taccaatcct ttgacgagtc tcttgctaaa gcctacaaag ggagccttat 2340
atccccgtat gagctccggt tcagagcact caacgaactg ctgtccaaac aggattttgc 2400
tcgcgtgatt ctccacatag cgaagaggcg aggatacgat gacattaaaa acagtgatga 2460
taaggaaaaa ggggccatac tcaaagcgat taagcaaaat gaagagaagc tcgctaacta 2520
tcaatcagta ggggagtatc tctataaaga gtacttccag aagttcaaag aaaatagcaa 2580
ggaattttact aatgtccgga ataaaaagga gtcttacgaa agatgtattg cgcaatcttt 2640
cctcaaggac gagctcaaat tgattttcaa gaaacaaagg gaatttgggt tcagcttctc 2700
aaaaaaattt gaggaagagg ttctgagcgt tgctttttac aaacgcgcc ttaaggactt 2760
ctcacatctc gtagggaatt gtagtttctt caccgatgaa aaacgggcgc caaaaaatag 2820
ccctttggct tttatgtttg tcgctctgac tcgcatcatt aatctgctca acaaccttaa 2880
aaacacggaa gggattctgt acacaaagga tgatctgaac gctctgctta acgaagtttt 2940
gaagaacggg actttgacct acaaacaaac caaaaagctt cttggtctca gtgatgacta 3000
cgaattcaag ggagaaaaag ggacatattt catcgaattc aagaagtata aggagttcat 3060
caaagccttg ggcgagcaca acttgctca agatgatctc aacgaaattg ctaaggatat 3120
cactctgatt aaagacgaga tcaagctcaa aaaggcgttg gcgaagtatg accttaacca 3180
aaaccaaata gatagcctca gcaagttgga atttaaagat cacttgaata taagtttcaa 3240
ggcccttaag ttggtcaccc ccttgatgct tgaaggaaag aaatatgatg aggcatgtaa 3300
tgagctgaat ctcaagggtg ctattaacga agacaaaaaa gatttcctcc cagctttcaa 3360
tgagacttac tataaggacg aggttaccaa tcctgtggtg ctccgagcca tcaaagagta 3420
tcgaaaggtc ctgaatgctt tgctcaaaaa atacggttaag gtacacaaaa taaatattga 3480
gctcgcaagg gaggtcggta agaaccactc ccagcgcgcc aaaatagaaa aggaacagaa 3540
tgaaaattac aaagcgaaaa aggacgccga gctcgagtgc gaaaagctgg gcctgaaaat 3600
aaacagcaag aacattctca aactccgcct cttcaaagaa caaaaagaat tttgtgctta 3660
tagtggtgag aaaataaaaa tctccgatct tcaagacgag aagatgctcg aaatagacgc 3720
gatatatcca tatagcaggt cttttgacga ttcttacatg aataaagtgc ttgttttcac 3780
taagcagaat caggaaaaag tgaatcagac cccctttgag gcctttggca acgactcagc 3840
aaagtggcag aagatcgagg tcttggctaa gaatcttctt actaagaaac agaaaaggat 3900
attggataag aactataaag acaaagaaca aaagaacttt aaagaccgca acctcaatga 3960
caccagatac atagcaagat tggttctgaa ctacacaaaa gattatttgg acttcttgcc 4020
gctgtctgat gatgagaaca cgaaactcaa cgacacgcaa aagggttcta aagtccacgt 4080
cgaagctaaa tctgggatgc tcacctcagc attgaggcat acgtggggat tctcagcaaa 4140
ggaccgaaac aatcacctgc accatgccat tgacgcagtt atcatagcgt atgccataa 4200
ttcaatagta aaagcgttta gcgacttcaa gaaggaacaa gagtccaaca gcgccgagct 4260
ctacgcaaaa aagattagtg aactcgacta caaaaacaaa agaaaattct ttgagccgtt 4320
cagcggattt cgacagaagg tattggataa aatagatgaa attttcgtga gcaaaccgga 4380

aaggaaaaag ccctcaggcg ccttgcacga agagactttc aggaaggaag aggaattcta 4440
ccaaagctac ggcggaagaggagtttt gaaggctctc gaacttggaag agattaggaa 4500
ggtgaacggc aagatagtga aaaacggcga tatgttccgg gttgatattct tcaaacataa 4560
aaaaacgaat aaattttatg ctgtgcctat atacactatg gacttcgcac ttaaggtcct 4620
gccgaataag gcggtagccc gatctaaaaa aggcgaaatt aaggactgga ttttgatgga 4680
tgaaaattac gagttctgct tttctctcta caaggattcc cttatattga tacagacgaa 4740
agatatgcag gaaccggaat tcgtgtatta caacgctttt acttcctcta cggtatcttt 4800
gattgtctcc aaacatgaca acaaattcga aacactcagt aaaaaccaa agattctctt 4860
taaaaatgcg aacgagaaaag aagtaattgc aaaatcaatt ggcatccaaa atttgaaagt 4920
ttttgaaaaa tatatagtat ctgccctcgg agaggttact aaagcggaat ttagacagcg 4980
agaggacttc aaaaaatcag gtccacccaa gaaaaaacgc aaggtggaag atccgaagaa 5040
aaagcgaaaa gtggatgtgt aacgttttcc gggacgccgg ctggatgatc ctccagcgcg 5100
gggatctcat gctggagttc ttcgccacc ccaacttggt tattgcagct tataatggtt 5160
acaaataaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta 5220
gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgtataccg 5270

<210> 204

<211> 137

<212> PRT

<213> 未知的

<220>

<223> RNA酶K

<400> 204

Met	Gly	Trp	Leu	Arg	Pro	Gly	Pro	Arg	Pro	Leu	Cys	Pro	Pro	Ala	Arg
1				5				10						15	
Ala	Ser	Trp	Ala	Phe	Ser	His	Arg	Phe	Pro	Ser	Pro	Leu	Ala	Pro	Arg
			20					25						30	
Arg	Ser	Pro	Thr	Pro	Phe	Phe	Met	Ala	Ser	Leu	Leu	Cys	Cys	Gly	Pro
			35				40					45			
Lys	Leu	Ala	Ala	Cys	Gly	Ile	Val	Leu	Ser	Ala	Trp	Gly	Val	Ile	Met
			50				55					60			
Leu	Ile	Met	Leu	Gly	Ile	Phe	Phe	Asn	Val	His	Ser	Ala	Val	Leu	Ile
65					70				75					80	
Glu	Asp	Val	Pro	Phe	Thr	Glu	Lys	Asp	Phe	Glu	Asn	Gly	Pro	Gln	Asn
				85					90					95	
Ile	Tyr	Asn	Leu	Tyr	Glu	Gln	Val	Ser	Tyr	Asn	Cys	Phe	Ile	Ala	Ala
			100					105					110		
Gly	Leu	Tyr	Leu	Leu	Leu	Gly	Gly	Phe	Ser	Phe	Cys	Gln	Val	Arg	Leu
			115				120						125		
Asn	Lys	Arg	Lys	Glu	Tyr	Met	Val	Arg							

130 135
 <210> 205
 <211> 1245
 <212> PRT
 <213> 未知的
 <220>
 <223> TALEN
 <400> 205
 Met Arg Ile Gly Lys Ser Ser Gly Trp Leu Asn Glu Ser Val Ser Leu
 1 5 10 15
 Glu Tyr Glu His Val Ser Pro Pro Thr Arg Pro Arg Asp Thr Arg Arg
 20 25 30
 Arg Pro Arg Ala Ala Gly Asp Gly Gly Leu Ala His Leu His Arg Arg
 35 40 45
 Leu Ala Val Gly Tyr Ala Glu Asp Thr Pro Arg Thr Glu Ala Arg Ser
 50 55 60
 Pro Ala Pro Arg Arg Pro Leu Pro Val Ala Pro Ala Ser Ala Pro Pro
 65 70 75 80
 Ala Pro Ser Leu Val Pro Glu Pro Pro Met Pro Val Ser Leu Pro Ala
 85 90 95
 Val Ser Ser Pro Arg Phe Ser Ala Gly Ser Ser Ala Ala Ile Thr Asp
 100 105 110
 Pro Phe Pro Ser Leu Pro Pro Thr Pro Val Leu Tyr Ala Met Ala Arg
 115 120 125
 Glu Leu Glu Ala Leu Ser Asp Ala Thr Trp Gln Pro Ala Val Pro Leu
 130 135 140
 Pro Ala Glu Pro Pro Thr Asp Ala Arg Arg Gly Asn Thr Val Phe Asp
 145 150 155 160
 Glu Ala Ser Ala Ser Ser Pro Val Ile Ala Ser Ala Cys Pro Gln Ala
 165 170 175
 Phe Ala Ser Pro Pro Arg Ala Pro Arg Ser Ala Arg Ala Arg Arg Ala
 180 185 190
 Arg Thr Gly Gly Asp Ala Trp Pro Ala Pro Thr Phe Leu Ser Arg Pro
 195 200 205
 Ser Ser Ser Arg Ile Gly Arg Asp Val Phe Gly Lys Leu Val Ala Leu
 210 215 220
 Gly Tyr Ser Arg Glu Gln Ile Arg Lys Leu Lys Gln Glu Ser Leu Ser
 225 230 235 240
 Glu Ile Ala Lys Tyr His Thr Thr Leu Thr Gly Gln Gly Phe Thr His

	245		250		255
Ala Asp Ile Cys Arg Ile Ser Arg Arg Arg Gln Ser Leu Arg Val Val					
	260		265		270
Ala Arg Asn Tyr Pro Glu Leu Ala Ala Ala Leu Pro Glu Leu Thr Arg					
	275		280		285
Ala His Ile Val Asp Ile Ala Arg Gln Arg Ser Gly Asp Leu Ala Leu					
	290		295		300
Gln Ala Leu Leu Pro Val Ala Thr Ala Leu Thr Ala Ala Pro Leu Arg					
305		310		315	320
Leu Ser Ala Ser Gln Ile Ala Thr Val Ala Gln Tyr Gly Glu Arg Pro					
	325		330		335
Ala Ile Gln Ala Leu Tyr Arg Leu Arg Arg Lys Leu Thr Arg Ala Pro					
	340		345		350
Leu His Leu Thr Pro Gln Gln Val Val Ala Ile Ala Ser Asn Thr Gly					
	355		360		365
Gly Lys Arg Ala Leu Glu Ala Val Cys Val Gln Leu Pro Val Leu Arg					
	370		375		380
Ala Ala Pro Tyr Arg Leu Ser Thr Glu Gln Val Val Ala Ile Ala Ser					
385		390		395	400
Asn Lys Gly Gly Lys Gln Ala Leu Glu Ala Val Lys Ala His Leu Leu					
	405		410		415
Asp Leu Leu Gly Ala Pro Tyr Val Leu Asp Thr Glu Gln Val Val Ala					
	420		425		430
Ile Ala Ser His Asn Gly Gly Lys Gln Ala Leu Glu Ala Val Lys Ala					
	435		440		445
Asp Leu Leu Asp Leu Arg Gly Ala Pro Tyr Ala Leu Ser Thr Glu Gln					
	450		455		460
Val Val Ala Ile Ala Ser His Asn Gly Gly Lys Gln Ala Leu Glu Ala					
465		470		475	480
Val Lys Ala Asp Leu Leu Glu Leu Arg Gly Ala Pro Tyr Ala Leu Ser					
	485		490		495
Thr Glu Gln Val Val Ala Ile Ala Ser His Asn Gly Gly Lys Gln Ala					
	500		505		510
Leu Glu Ala Val Lys Ala His Leu Leu Asp Leu Arg Gly Val Pro Tyr					
	515		520		525
Ala Leu Ser Thr Glu Gln Val Val Ala Ile Ala Ser His Asn Gly Gly					
	530		535		540
Lys Gln Ala Leu Glu Ala Val Lys Ala Gln Leu Leu Asp Leu Arg Gly					
545		550		555	560

Ala Pro Tyr Ala Leu Ser Thr Ala Gln Val Val Ala Ile Ala Ser Asn		
565	570	575
Gly Gly Gly Lys Gln Ala Leu Glu Gly Ile Gly Glu Gln Leu Leu Lys		
580	585	590
Leu Arg Thr Ala Pro Tyr Gly Leu Ser Thr Glu Gln Val Val Ala Ile		
595	600	605
Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Ala Val Gly Ala Gln		
610	615	620
Leu Val Ala Leu Arg Ala Ala Pro Tyr Ala Leu Ser Thr Glu Gln Val		
625	630	635
Val Ala Ile Ala Ser Asn Lys Gly Gly Lys Gln Ala Leu Glu Ala Val		
645	650	655
Lys Ala Gln Leu Leu Glu Leu Arg Gly Ala Pro Tyr Ala Leu Ser Thr		
660	665	670
Ala Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Asn Gln Ala Leu		
675	680	685
Glu Ala Val Gly Thr Gln Leu Val Ala Leu Arg Ala Ala Pro Tyr Ala		
690	695	700
Leu Ser Thr Glu Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys		
705	710	715
Gln Ala Leu Glu Ala Val Gly Ala Gln Leu Val Ala Leu Arg Ala Ala		
725	730	735
Pro Tyr Ala Leu Asn Thr Glu Gln Val Val Ala Ile Ala Ser Ser His		
740	745	750
Gly Gly Lys Gln Ala Leu Glu Ala Val Arg Ala Leu Phe Pro Asp Leu		
755	760	765
Arg Ala Ala Pro Tyr Ala Leu Ser Thr Ala Gln Leu Val Ala Ile Ala		
770	775	780
Ser Asn Pro Gly Gly Lys Gln Ala Leu Glu Ala Val Arg Ala Leu Phe		
785	790	795
Arg Glu Leu Arg Ala Ala Pro Tyr Ala Leu Ser Thr Glu Gln Val Val		
805	810	815
Ala Ile Ala Ser Asn His Gly Gly Lys Gln Ala Leu Glu Ala Val Arg		
820	825	830
Ala Leu Phe Arg Gly Leu Arg Ala Ala Pro Tyr Gly Leu Ser Thr Ala		
835	840	845
Gln Val Val Ala Ile Ala Ser Ser Asn Gly Gly Lys Gln Ala Leu Glu		
850	855	860
Ala Val Trp Ala Leu Leu Pro Val Leu Arg Ala Thr Pro Tyr Asp Leu		

865	870	875	880
Asn Thr Ala Gln Ile Val Ala Ile Ala Ser His Asp Gly Gly Lys Pro			
	885	890	895
Ala Leu Glu Ala Val Trp Ala Lys Leu Pro Val Leu Arg Gly Ala Pro			
900	905	910	
Tyr Ala Leu Ser Thr Ala Gln Val Val Ala Ile Ala Cys Ile Ser Gly			
915	920	925	
Gln Gln Ala Leu Glu Ala Ile Glu Ala His Met Pro Thr Leu Arg Gln			
930	935	940	
Ala Ser His Ser Leu Ser Pro Glu Arg Val Ala Ala Ile Ala Cys Ile			
945	950	955	960
Gly Gly Arg Ser Ala Val Glu Ala Val Arg Gln Gly Leu Pro Val Lys			
965	970	975	
Ala Ile Arg Arg Ile Arg Arg Glu Lys Ala Pro Val Ala Gly Pro Pro			
980	985	990	
Pro Ala Ser Leu Gly Pro Thr Pro Gln Glu Leu Val Ala Val Leu His			
995	1000	1005	
Phe Phe Arg Ala His Gln Gln Pro Arg Gln Ala Phe Val Asp Ala			
1010	1015	1020	
Leu Ala Ala Phe Gln Ala Thr Arg Pro Ala Leu Leu Arg Leu Leu			
1025	1030	1035	
Ser Ser Val Gly Val Thr Glu Ile Glu Ala Leu Gly Gly Thr Ile			
1040	1045	1050	
Pro Asp Ala Thr Glu Arg Trp Gln Arg Leu Leu Gly Arg Leu Gly			
1055	1060	1065	
Phe Arg Pro Ala Thr Gly Ala Ala Ala Pro Ser Pro Asp Ser Leu			
1070	1075	1080	
Gln Gly Phe Ala Gln Ser Leu Glu Arg Thr Leu Gly Ser Pro Gly			
1085	1090	1095	
Met Ala Gly Gln Ser Ala Cys Ser Pro His Arg Lys Arg Pro Ala			
1100	1105	1110	
Glu Thr Ala Ile Ala Pro Arg Ser Ile Arg Arg Ser Pro Asn Asn			
1115	1120	1125	
Ala Gly Gln Pro Ser Glu Pro Trp Pro Asp Gln Leu Ala Trp Leu			
1130	1135	1140	
Gln Arg Arg Lys Arg Thr Ala Arg Ser His Ile Arg Ala Asp Ser			
1145	1150	1155	
Ala Ala Ser Val Pro Ala Asn Leu His Leu Gly Thr Arg Ala Gln			
1160	1165	1170	

Phe Thr Pro Asp Arg Leu Arg Ala Glu Pro Gly Pro Ile Met Gln		
1175	1180	1185
Ala His Thr Ser Pro Ala Ser Val Ser Phe Gly Ser His Val Ala		
1190	1195	1200
Phe Glu Pro Gly Leu Pro Asp Pro Gly Thr Pro Thr Ser Ala Asp		
1205	1210	1215
Leu Ala Ser Phe Glu Ala Glu Pro Phe Gly Val Gly Pro Leu Asp		
1220	1225	1230
Phe His Leu Asp Trp Leu Leu Gln Ile Leu Glu Thr		
1235	1240	1245
<210> 206		
<211> 1373		
<212> PRT		
<213> 未知的		
<220>		
<223> TALEN		
<400> 206		
Met Asp Pro Ile Arg Ser Arg Thr Pro Ser Pro Ala Arg Glu Leu Leu		
1	5	10
Pro Gly Pro Gln Pro Asp Arg Val Gln Pro Thr Ala Asp Arg Gly Gly		
20	25	30
Ala Pro Pro Ala Gly Gly Pro Leu Asp Gly Leu Pro Ala Arg Arg Thr		
35	40	45
Met Ser Arg Thr Arg Leu Pro Ser Pro Pro Ala Pro Ser Pro Ala Phe		
50	55	60
Ser Ala Gly Ser Phe Ser Asp Leu Leu Arg Gln Phe Asp Pro Ser Leu		
65	70	75
Leu Asp Thr Ser Leu Leu Asp Ser Met Pro Ala Val Gly Thr Pro His		
85	90	95
Thr Ala Ala Ala Pro Ala Glu Cys Asp Glu Val Gln Ser Gly Leu Arg		
100	105	110
Ala Ala Asp Asp Pro Pro Pro Thr Val Arg Val Ala Val Thr Ala Ala		
115	120	125
Arg Pro Pro Arg Ala Lys Pro Ala Pro Arg Arg Arg Ala Ala Gln Pro		
130	135	140
Ser Asp Ala Ser Pro Ala Ala Gln Val Asp Leu Arg Thr Leu Gly Tyr		
145	150	155
Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Gly Ser Thr Val		
165	170	175

Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His		
180	185	190
Ile Val Ala Leu Ser Arg His Pro Ala Ala Leu Gly Thr Val Ala Val		
195	200	205
Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Asp		
210	215	220
Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala		
225	230	235
240		
Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp		
245	250	255
Thr Gly Gln Leu Val Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val		
260	265	270
Glu Ala Val His Ala Ser Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn		
275	280	285
Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys		
290	295	300
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Ala		
305	310	315
320		
His Gly Leu Thr Pro Ala Gln Val Val Ala Ile Ala Ser His Asp Gly		
325	330	335
Gly Lys Gln Ala Leu Glu Thr Met Gln Arg Leu Leu Pro Val Leu Cys		
340	345	350
Gln Ala His Gly Leu Pro Pro Asp Gln Val Val Ala Ile Ala Ser Asn		
355	360	365
Ile Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val		
370	375	380
Leu Cys Gln Ala His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala		
385	390	395
400		
Ser His Gly Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg Leu Leu		
405	410	415
Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Asp Gln Val Val Ala		
420	425	430
Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg		
435	440	445
Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Asp Gln Val		
450	455	460
Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala Leu Glu Thr Val		
465	470	475
480		
Gln Arg Leu Leu Pro Val Leu Cys Gln Ala His Gly Leu Thr Pro Asp		

				485					490					495			
Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr		
				500					505					510			
Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro		
				515					520					525			
Asp	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu		
				530					535					540			
Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Thr	His	Gly	Leu		
545					550					555					560		
Thr	Pro	Ala	Gln	Val	Val	Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln		
				565					570					575			
Ala	Leu	Glu	Thr	Val	Gln	Gln	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His		
				580					585					590			
Gly	Leu	Thr	Pro	Asp	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly		
				595					600					605			
Lys	Gln	Ala	Leu	Ala	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln		
				610					615					620			
Ala	His	Gly	Leu	Thr	Pro	Asp	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly		
625					630					635					640		
Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu		
				645					650					655			
Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Asp	Gln	Val	Val	Ala	Ile	Ala	Ser		
				660					665					670			
Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro		
				675					680					685			
Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Gln	Val	Gln	Val	Val	Ala	Ile		
				690					695					700			
Ala	Ser	Asn	Ile	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu		
705					710					715					720		
Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Ala	Gln	Val	Val		
				725					730					735			
Ala	Ile	Ala	Ser	His	Asp	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln		
				740					745					750			
Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Pro	Asp	Gln		
				755					760					765			
Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr		
				770					775					780			
Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu	Thr	Gln		
785					790					795					800		

Glu	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Asn	Gly	Gly	Lys	Gln	Ala	Leu
				805			810				815				
Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His	Gly	Leu
				820			825				830				
Thr	Pro	Asp	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Gly	Lys	Gln
				835			840				845				
Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Ala	His
				850			855				860				
Gly	Leu	Thr	Pro	Ala	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile	Gly	Gly
865				870			875				880				
Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln
				885			890				895				
Asp	His	Gly	Leu	Thr	Leu	Ala	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ile
				900			905				910				
Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu
				915			920				925				
Cys	Gln	Ala	His	Gly	Leu	Thr	Gln	Asp	Gln	Val	Val	Ala	Ile	Ala	Ser
930				935			940								
Asn	Ile	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro
945				950			955				960				
Val	Leu	Cys	Gln	Asp	His	Gly	Leu	Thr	Pro	Asp	Gln	Val	Val	Ala	Ile
				965			970				975				
Ala	Ser	Asn	Ile	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	Arg	Leu
				980			985				990				
Leu	Pro	Val	Leu	Cys	Gln	Asp	His	Gly	Leu	Thr	Leu	Asp	Gln	Val	Val
				995			1000				1005				
Ala	Ile	Ala	Ser	Asn	Gly	Gly	Lys	Gln	Ala	Leu	Glu	Thr	Val	Gln	
1010				1015			1020								
Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Asp	His	Gly	Leu	Thr	Pro	Asp	
1025				1030			1035								
Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Ser	Gly	Gly	Lys	Gln	Ala	Leu	
1040				1045			1050								
Glu	Thr	Val	Gln	Arg	Leu	Leu	Pro	Val	Leu	Cys	Gln	Asp	His	Gly	
1055				1060			1065								
Leu	Thr	Pro	Asn	Gln	Val	Val	Ala	Ile	Ala	Ser	Asn	Gly	Gly	Lys	
1070				1075			1080								
Gln	Ala	Leu	Glu	Ser	Ile	Val	Ala	Gln	Leu	Ser	Arg	Pro	Asp	Pro	
1085				1090			1095								
Ala	Leu	Ala	Ala	Leu	Thr	Asn	Asp	His	Leu	Val	Ala	Leu	Ala	Cys	

1100	1105	1110
Leu Gly Gly Arg Pro Ala	Met Asp Ala Val Lys	Lys Gly Leu Pro
1115	1120	1125
His Ala Pro Glu Leu Ile	Arg Arg Val Asn Arg	Arg Ile Gly Glu
1130	1135	1140
Arg Thr Ser His Arg Val	Ala Asp Tyr Ala Gln	Val Val Arg Val
1145	1150	1155
Leu Glu Phe Phe Gln Cys	His Ser His Pro Ala	Tyr Ala Phe Asp
1160	1165	1170
Glu Ala Met Thr Gln Phe	Gly Met Ser Arg Asn	Gly Leu Val Gln
1175	1180	1185
Leu Phe Arg Arg Val Gly	Val Thr Glu Leu Glu	Ala Arg Gly Gly
1190	1195	1200
Thr Leu Pro Pro Ala Ser	Gln Arg Trp Asp Arg	Ile Leu Gln Ala
1205	1210	1215
Ser Gly Met Lys Arg Ala	Lys Pro Ser Pro Thr	Ser Ala Gln Thr
1220	1225	1230
Pro Asp Gln Ala Ser Leu	His Ala Phe Ala Asp	Ser Leu Glu Arg
1235	1240	1245
Asp Leu Asp Ala Pro Ser	Pro Met His Glu Gly	Asp Gln Thr Gly
1250	1255	1260
Ala Ser Ser Arg Lys Arg	Ser Arg Ser Asp Arg	Ala Val Thr Gly
1265	1270	1275
Pro Ser Ala Gln His Ser	Phe Glu Val Arg Val	Pro Glu Gln Arg
1280	1285	1290
Asp Ala Leu His Leu Pro	Leu Ser Trp Arg Val	Lys Arg Pro Arg
1295	1300	1305
Thr Arg Ile Gly Gly Gly	Leu Pro Asp Pro Gly	Thr Pro Ile Ala
1310	1315	1320
Ala Asp Leu Ala Ala Ser	Ser Thr Val Met Trp	Glu Gln Asp Ala
1325	1330	1335
Ala Pro Phe Ala Gly Ala	Ala Asp Asp Phe Pro	Ala Phe Asn Glu
1340	1345	1350
Glu Glu Leu Ala Trp Leu	Met Glu Leu Leu Pro	Gln Ser Gly Ser
1355	1360	1365
Val Gly Gly Thr Ile		
1370		
<210> 207		
<211> 1978		

<212> PRT

<213> 未知的

<220>

<223> ZNF638

<400> 207

Met	Ser	Arg	Pro	Arg	Phe	Asn	Pro	Arg	Gly	Asp	Phe	Pro	Leu	Gln	Arg
1				5					10					15	
Pro	Arg	Ala	Pro	Asn	Pro	Ser	Gly	Met	Arg	Pro	Pro	Gly	Pro	Phe	Met
			20					25					30		
Arg	Pro	Gly	Ser	Met	Gly	Leu	Pro	Arg	Phe	Tyr	Pro	Ala	Gly	Arg	Ala
			35					40					45		
Arg	Gly	Ile	Pro	His	Arg	Phe	Ala	Gly	His	Glu	Ser	Tyr	Gln	Asn	Met
			50				55					60			
Gly	Pro	Gln	Arg	Met	Asn	Val	Gln	Val	Thr	Gln	His	Arg	Thr	Asp	Pro
65					70					75					80
Arg	Leu	Thr	Lys	Glu	Lys	Leu	Asp	Phe	His	Glu	Ala	Gln	Gln	Lys	Lys
				85					90					95	
Gly	Lys	Pro	His	Gly	Ser	Arg	Trp	Asp	Asp	Glu	Pro	His	Ile	Ser	Ala
			100					105					110		
Ser	Val	Ala	Val	Lys	Gln	Ser	Ser	Val	Thr	Gln	Val	Thr	Glu	Gln	Ser
			115					120					125		
Pro	Lys	Val	Gln	Ser	Arg	Tyr	Thr	Lys	Glu	Ser	Ala	Ser	Ser	Ile	Leu
			130					135					140		
Ala	Ser	Phe	Gly	Leu	Ser	Asn	Glu	Asp	Leu	Glu	Glu	Leu	Ser	Arg	Tyr
145					150					155					160
Pro	Asp	Glu	Gln	Leu	Thr	Pro	Glu	Asn	Met	Pro	Leu	Ile	Leu	Arg	Asp
				165					170					175	
Ile	Arg	Met	Arg	Lys	Met	Gly	Arg	Arg	Leu	Pro	Asn	Leu	Pro	Ser	Gln
				180				185						190	
Ser	Arg	Asn	Lys	Glu	Thr	Leu	Gly	Ser	Glu	Ala	Val	Ser	Ser	Asn	Val
			195					200						205	
Ile	Asp	Tyr	Gly	His	Ala	Ser	Lys	Tyr	Gly	Tyr	Thr	Glu	Asp	Pro	Leu
			210				215					220			
Glu	Val	Arg	Ile	Tyr	Asp	Pro	Glu	Ile	Pro	Thr	Asp	Glu	Val	Glu	Asn
225					230					235					240
Glu	Phe	Gln	Ser	Gln	Gln	Asn	Ile	Ser	Ala	Ser	Val	Pro	Asn	Pro	Asn
				245					250					255	
Val	Ile	Cys	Asn	Ser	Met	Phe	Pro	Val	Glu	Asp	Val	Phe	Arg	Gln	Met
			260					265						270	

Asp Phe Pro Gly Glu Ser Ser Asn Asn Arg Ser Phe Phe Ser Val Glu		
275	280	285
Ser Gly Thr Lys Met Ser Gly Leu His Ile Ser Gly Gly Gln Ser Val		
290	295	300
Leu Glu Pro Ile Lys Ser Val Asn Gln Ser Ile Asn Gln Thr Val Ser		
305	310	315
Gln Thr Met Ser Gln Ser Leu Ile Pro Pro Ser Met Asn Gln Gln Pro		
	325	330
Phe Ser Ser Glu Leu Ile Ser Ser Val Ser Gln Gln Glu Arg Ile Pro		
	340	345
His Glu Pro Val Ile Asn Ser Ser Asn Val His Val Gly Ser Arg Gly		
	355	360
Ser Lys Lys Asn Tyr Gln Ser Gln Ala Asp Ile Pro Ile Arg Ser Pro		
370	375	380
Phe Gly Ile Val Lys Ala Ser Trp Leu Pro Lys Phe Ser His Ala Asp		
385	390	395
Ala Gln Lys Met Lys Arg Leu Pro Thr Pro Ser Met Met Asn Asp Tyr		
	405	410
Tyr Ala Ala Ser Pro Arg Ile Phe Pro His Leu Cys Ser Leu Cys Asn		
	420	425
Val Glu Cys Ser His Leu Lys Asp Trp Ile Gln His Gln Asn Thr Ser		
435	440	445
Thr His Ile Glu Ser Cys Arg Gln Leu Arg Gln Gln Tyr Pro Asp Trp		
450	455	460
Asn Pro Glu Ile Leu Pro Ser Arg Arg Asn Glu Gly Asn Arg Lys Glu		
465	470	475
Asn Glu Thr Pro Arg Arg Arg Ser His Ser Pro Ser Pro Arg Arg Ser		
	485	490
Arg Arg Ser Ser Ser Ser His Arg Phe Arg Arg Ser Arg Ser Pro Met		
	500	505
His Tyr Met Tyr Arg Pro Arg Ser Arg Ser Pro Arg Ile Cys His Arg		
	515	520
Phe Ile Ser Arg Tyr Arg Ser Arg Ser Arg Ser Arg Ser Pro Tyr Arg		
530	535	540
Ile Arg Asn Pro Phe Arg Gly Ser Pro Lys Cys Phe Arg Ser Val Ser		
545	550	555
Pro Glu Arg Met Ser Arg Arg Ser Val Arg Ser Ser Asp Arg Lys Lys		
	565	570
Ala Leu Glu Asp Val Val Gln Arg Ser Gly His Gly Thr Glu Phe Asn		

580	585	590
Lys Gln Lys His Leu Glu Ala Ala Asp Lys Gly His Ser Pro Ala Gln		
595	600	605
Lys Pro Lys Thr Ser Ser Gly Thr Lys Pro Ser Val Lys Pro Thr Ser		
610	615	620
Ala Thr Lys Ser Asp Ser Asn Leu Gly Gly His Ser Ile Arg Cys Lys		
625	630	635
Ser Lys Asn Leu Glu Asp Asp Thr Leu Ser Glu Cys Lys Gln Val Ser		
645	650	655
Asp Lys Ala Val Ser Leu Gln Arg Lys Leu Arg Lys Glu Gln Ser Leu		
660	665	670
His Tyr Gly Ser Val Leu Leu Ile Thr Glu Leu Pro Glu Asp Gly Cys		
675	680	685
Thr Glu Glu Asp Val Arg Lys Leu Phe Gln Pro Phe Gly Lys Val Asn		
690	695	700
Asp Val Leu Ile Val Pro Tyr Arg Lys Glu Ala Tyr Leu Glu Met Glu		
705	710	715
Phe Lys Glu Ala Ile Thr Ala Ile Met Lys Tyr Ile Glu Thr Thr Pro		
725	730	735
Leu Thr Ile Lys Gly Lys Ser Val Lys Ile Cys Val Pro Gly Lys Lys		
740	745	750
Lys Ala Gln Asn Lys Glu Val Lys Lys Lys Thr Leu Glu Ser Lys Lys		
755	760	765
Val Ser Ala Ser Thr Leu Lys Arg Asp Ala Asp Ala Ser Lys Ala Val		
770	775	780
Glu Ile Val Thr Ser Thr Ser Ala Ala Lys Thr Gly Gln Ala Lys Ala		
785	790	795
Ser Val Ala Lys Val Asn Lys Ser Thr Gly Lys Ser Ala Ser Ser Val		
805	810	815
Lys Ser Val Val Thr Val Ala Val Lys Gly Asn Lys Ala Ser Ile Lys		
820	825	830
Thr Ala Lys Ser Gly Gly Lys Lys Ser Leu Glu Ala Lys Lys Thr Gly		
835	840	845
Asn Val Lys Asn Lys Asp Ser Asn Lys Pro Val Thr Ile Pro Glu Asn		
850	855	860
Ser Glu Ile Lys Thr Ser Ile Glu Val Lys Ala Thr Glu Asn Cys Ala		
865	870	875
Lys Glu Ala Ile Ser Asp Ala Ala Leu Glu Ala Thr Glu Asn Glu Pro		
885	890	895

Leu Asn Lys Glu Thr Glu Glu Met Cys Val Met Leu Val Ser Asn Leu		
900	905	910
Pro Asn Lys Gly Tyr Ser Val Glu Glu Val Tyr Asp Leu Ala Lys Pro		
915	920	925
Phe Gly Gly Leu Lys Asp Ile Leu Ile Leu Ser Ser His Lys Lys Ala		
930	935	940
Tyr Ile Glu Ile Asn Arg Lys Ala Ala Glu Ser Met Val Lys Phe Tyr		
945	950	955
Thr Cys Phe Pro Val Leu Met Asp Gly Asn Gln Leu Ser Ile Ser Met		
965	970	975
Ala Pro Glu Asn Met Asn Ile Lys Asp Glu Glu Ala Ile Phe Ile Thr		
980	985	990
Leu Val Lys Glu Asn Asp Pro Glu Ala Asn Ile Asp Thr Ile Tyr Asp		
995	1000	1005
Arg Phe Val His Leu Asp Asn Leu Pro Glu Asp Gly Leu Gln Cys		
1010	1015	1020
Val Leu Cys Val Gly Leu Gln Phe Gly Lys Val Asp His His Val		
1025	1030	1035
Phe Ile Ser Asn Arg Asn Lys Ala Ile Leu Gln Leu Asp Ser Pro		
1040	1045	1050
Glu Ser Ala Gln Ser Met Tyr Ser Phe Leu Lys Gln Asn Pro Gln		
1055	1060	1065
Asn Ile Gly Asp His Met Leu Thr Cys Ser Leu Ser Pro Lys Ile		
1070	1075	1080
Asp Leu Pro Glu Val Gln Ile Glu His Asp Pro Glu Leu Glu Lys		
1085	1090	1095
Glu Ser Pro Gly Leu Lys Asn Ser Pro Ile Asp Glu Ser Glu Val		
1100	1105	1110
Gln Thr Ala Thr Asp Ser Pro Ser Val Lys Pro Asn Glu Leu Glu		
1115	1120	1125
Glu Glu Ser Thr Pro Ser Ile Gln Thr Glu Thr Leu Val Gln Gln		
1130	1135	1140
Glu Glu Pro Cys Glu Glu Glu Ala Glu Lys Ala Thr Cys Asp Ser		
1145	1150	1155
Asp Phe Ala Val Glu Thr Leu Glu Leu Glu Thr Gln Gly Glu Glu		
1160	1165	1170
Val Lys Glu Glu Ile Pro Leu Val Ala Ser Ala Ser Val Ser Ile		
1175	1180	1185
Glu Gln Phe Thr Glu Asn Ala Glu Glu Cys Ala Leu Asn Gln Gln		

1190	1195	1200
Met Phe Asn Ser Asp Leu	Glu Lys Lys Gly Ala	Glu Ile Ile Asn
1205	1210	1215
Pro Lys Thr Ala Leu Leu	Pro Ser Asp Ser Val	Phe Ala Glu Glu
1220	1225	1230
Arg Asn Leu Lys Gly Ile	Leu Glu Glu Ser Pro	Ser Glu Ala Glu
1235	1240	1245
Asp Phe Ile Ser Gly Ile	Thr Gln Thr Met Val	Glu Ala Val Ala
1250	1255	1260
Glu Val Glu Lys Asn Glu	Thr Val Ser Glu Ile	Leu Pro Ser Thr
1265	1270	1275
Cys Ile Val Thr Leu Val	Pro Gly Ile Pro Thr	Gly Asp Glu Lys
1280	1285	1290
Thr Val Asp Lys Lys Asn	Ile Ser Glu Lys Lys	Gly Asn Met Asp
1295	1300	1305
Glu Lys Glu Glu Lys Glu	Phe Asn Thr Lys Glu	Thr Arg Met Asp
1310	1315	1320
Leu Gln Ile Gly Thr Glu	Lys Ala Glu Lys Asn	Glu Gly Arg Met
1325	1330	1335
Asp Ala Glu Lys Val Glu	Lys Met Ala Ala Met	Lys Glu Lys Pro
1340	1345	1350
Ala Glu Asn Thr Leu Phe	Lys Ala Tyr Pro Asn	Lys Gly Val Gly
1355	1360	1365
Gln Ala Asn Lys Pro Asp	Glu Thr Ser Lys Thr	Ser Ile Leu Ala
1370	1375	1380
Val Ser Asp Val Ser Ser	Ser Lys Pro Ser Ile	Lys Ala Val Ile
1385	1390	1395
Val Ser Ser Pro Lys Ala	Lys Ala Thr Val Ser	Lys Thr Glu Asn
1400	1405	1410
Gln Lys Ser Phe Pro Lys	Ser Val Pro Arg Asp	Gln Ile Asn Ala
1415	1420	1425
Glu Lys Lys Leu Ser Ala	Lys Glu Phe Gly Leu	Leu Lys Pro Thr
1430	1435	1440
Ser Ala Arg Ser Gly Leu	Ala Glu Ser Ser Ser	Lys Phe Lys Pro
1445	1450	1455
Thr Gln Ser Ser Leu Thr	Arg Gly Gly Ser Gly	Arg Ile Ser Ala
1460	1465	1470
Leu Gln Gly Lys Leu Ser	Lys Leu Asp Tyr Arg	Asp Ile Thr Lys
1475	1480	1485

Gln Ser Gln Glu Thr Glu Ala Arg Pro Ser Ile Met Lys Arg Asp		
1490	1495	1500
Asp Ser Asn Asn Lys Thr Leu Ala Glu Gln Asn Thr Lys Asn Pro		
1505	1510	1515
Lys Ser Thr Thr Gly Arg Ser Ser Lys Ser Lys Glu Glu Pro Leu		
1520	1525	1530
Phe Pro Phe Asn Leu Asp Glu Phe Val Thr Val Asp Glu Val Ile		
1535	1540	1545
Glu Glu Val Asn Pro Ser Gln Ala Lys Gln Asn Pro Leu Lys Gly		
1550	1555	1560
Lys Arg Lys Glu Thr Leu Lys Asn Val Pro Phe Ser Glu Leu Asn		
1565	1570	1575
Leu Lys Lys Lys Lys Gly Lys Thr Ser Thr Pro Arg Gly Val Glu		
1580	1585	1590
Gly Glu Leu Ser Phe Val Thr Leu Asp Glu Ile Gly Glu Glu Glu		
1595	1600	1605
Asp Ala Ala Ala His Leu Ala Gln Ala Leu Val Thr Val Asp Glu		
1610	1615	1620
Val Ile Asp Glu Glu Glu Leu Asn Met Glu Glu Met Val Lys Asn		
1625	1630	1635
Ser Asn Ser Leu Phe Thr Leu Asp Glu Leu Ile Asp Gln Asp Asp		
1640	1645	1650
Cys Ile Ser His Ser Glu Pro Lys Asp Val Thr Val Leu Ser Val		
1655	1660	1665
Ala Glu Glu Gln Asp Leu Leu Lys Gln Glu Arg Leu Val Thr Val		
1670	1675	1680
Asp Glu Ile Gly Glu Val Glu Glu Leu Pro Leu Asn Glu Ser Ala		
1685	1690	1695
Asp Ile Thr Phe Ala Thr Leu Asn Thr Lys Gly Asn Glu Gly Asp		
1700	1705	1710
Thr Val Arg Asp Ser Ile Gly Phe Ile Ser Ser Gln Val Pro Glu		
1715	1720	1725
Asp Pro Ser Thr Leu Val Thr Val Asp Glu Ile Gln Asp Asp Ser		
1730	1735	1740
Ser Asp Leu His Leu Val Thr Leu Asp Glu Val Thr Glu Glu Asp		
1745	1750	1755
Glu Asp Ser Leu Ala Asp Phe Asn Asn Leu Lys Glu Glu Leu Asn		
1760	1765	1770
Phe Val Thr Val Asp Glu Val Gly Glu Glu Glu Asp Gly Asp Asn		

1775	1780	1785
Asp Leu Lys Val Glu Leu	Ala Gln Ser Lys Asn	Asp His Pro Thr
1790	1795	1800
Asp Lys Lys Gly Asn Arg	Lys Lys Arg Ala Val	Asp Thr Lys Lys
1805	1810	1815
Thr Lys Leu Glu Ser Leu	Ser Gln Val Gly Pro	Val Asn Glu Asn
1820	1825	1830
Val Met Glu Glu Asp Leu	Lys Thr Met Ile Glu	Arg His Leu Thr
1835	1840	1845
Ala Lys Thr Pro Thr Lys	Arg Val Arg Ile Gly	Lys Thr Leu Pro
1850	1855	1860
Ser Glu Lys Ala Val Val	Thr Glu Pro Ala Lys	Gly Glu Glu Ala
1865	1870	1875
Phe Gln Met Ser Glu Val	Asp Glu Glu Ser Gly	Leu Lys Asp Ser
1880	1885	1890
Glu Pro Glu Arg Lys Arg	Lys Lys Thr Glu Asp	Ser Ser Ser Gly
1895	1900	1905
Lys Ser Val Ala Ser Asp	Val Pro Glu Glu Leu	Asp Phe Leu Val
1910	1915	1920
Pro Lys Ala Gly Phe Phe	Cys Pro Ile Cys Ser	Leu Phe Tyr Ser
1925	1930	1935
Gly Glu Lys Ala Met Thr	Asn His Cys Lys Ser	Thr Arg His Lys
1940	1945	1950
Gln Asn Thr Glu Lys Phe	Met Ala Lys Gln Arg	Lys Glu Lys Glu
1955	1960	1965
Gln Asn Glu Ala Glu Glu	Arg Ser Ser Arg	
1970	1975	
<210> 208		
<211> 128		
<212> PRT		
<213> 人工序列		
<220>		
<223> RNA酶1 (R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)		
<400> 208		
Lys Glu Ser Arg Ala Lys Lys Phe Gln Arg Gln His Met Asp Ser Asp		
1	5	10
Ser Ser Pro Ser Ser Ser Ser Thr Tyr Cys Asn Gln Met Met Arg Arg		
20	25	30
Arg Asn Met Thr Gln Gly Asp Cys Arg Pro Val Asn Thr Phe Val His		

35	40	45
Glu Pro Leu Val Asp Val Gln Asn Val Cys Phe Gln Glu Lys Val Thr		
50	55	60
Cys Lys Asp Gly Gln Gly Asn Cys Tyr Lys Ser Asn Ser Ser Met His		
65	70	75
Ile Thr Asp Cys Arg Leu Thr Ala Asp Ser Asp Tyr Pro Asn Cys Ala		
85	90	95
Tyr Arg Thr Ser Pro Lys Glu Arg His Ile Ile Val Ala Cys Glu Gly		
100	105	110
Ser Pro Tyr Val Pro Val Asn Phe Glu Ala Ser Val Glu Asp Ser Thr		
115	120	125
<210> 209		
<211> 1091		
<212> PRT		
<213> 未知的		
<220>		
<223> PUMBY		
<400> 209		
Met Asp Lys Ser Lys Gln Met Asn Ile Asn Asn Leu Ser Asn Ile Pro		
1	5	10
Glu Val Ile Asp Pro Gly Ile Thr Ile Pro Ile Tyr Glu Glu Glu Tyr		
20	25	30
Glu Asn Asn Gly Glu Ser Asn Ser Gln Leu Gln Gln Gln Pro Gln Lys		
35	40	45
Leu Gly Ser Tyr Arg Ser Arg Ala Gly Lys Phe Ser Asn Thr Leu Ser		
50	55	60
Asn Leu Leu Pro Ser Ile Ser Ala Lys Leu His His Ser Lys Lys Asn		
65	70	75
Ser His Gly Lys Asn Gly Ala Glu Phe Ser Ser Ser Asn Asn Ser Ser		
85	90	95
Gln Ser Thr Val Ala Ser Lys Thr Pro Arg Ala Ser Pro Ser Arg Ser		
100	105	110
Lys Met Met Glu Ser Ser Ile Asp Gly Val Thr Met Asp Arg Pro Gly		
115	120	125
Ser Leu Thr Pro Pro Gln Asp Met Glu Lys Leu Val His Phe Pro Asp		
130	135	140
Ser Ser Asn Asn Phe Leu Ile Pro Ala Pro Arg Gly Ser Ser Asp Ser		
145	150	155
Phe Asn Leu Pro His Gln Ile Ser Arg Thr Arg Asn Asn Thr Met Ser		

			165					170					175					
Ser	Gln	Ile	Thr	Ser	Ile	Ser	Ser	Ile	Ala	Pro	Lys	Pro	Arg	Thr	Ser			
			180					185					190					
Ser	Gly	Ile	Trp	Ser	Ser	Asn	Ala	Ser	Ala	Asn	Asp	Pro	Met	Gln	Gln			
			195					200					205					
His	Leu	Leu	Gln	Gln	Leu	Gln	Pro	Thr	Thr	Ser	Asn	Asn	Thr	Thr	Asn			
			210					215					220					
Ser	Asn	Thr	Leu	Asn	Asp	Tyr	Ser	Thr	Lys	Thr	Ala	Tyr	Phe	Asp	Asn			
225				230					235					240				
Met	Val	Ser	Thr	Ser	Gly	Ser	Gln	Met	Ala	Asp	Asn	Lys	Met	Asn	Thr			
			245					250					255					
Asn	Asn	Leu	Ala	Ile	Pro	Asn	Ser	Val	Trp	Ser	Asn	Thr	Arg	Gln	Arg			
			260					265					270					
Ser	Gln	Ser	Asn	Ala	Ser	Ser	Ile	Tyr	Thr	Asp	Ala	Pro	Leu	Tyr	Glu			
			275					280					285					
Gln	Pro	Ala	Arg	Ala	Ser	Ile	Ser	Ser	His	Tyr	Thr	Ile	Pro	Thr	Gln			
			290					295					300					
Glu	Ser	Pro	Leu	Ile	Ala	Asp	Glu	Ile	Asp	Pro	Gln	Ser	Ile	Asn	Trp			
305				310					315					320				
Val	Thr	Met	Asp	Pro	Thr	Val	Pro	Ser	Ile	Asn	Gln	Ile	Ser	Asn	Leu			
			325					330					335					
Leu	Pro	Thr	Asn	Thr	Ile	Ser	Ile	Ser	Asn	Val	Phe	Pro	Leu	Gln	His			
			340					345					350					
Gln	Gln	Pro	Gln	Leu	Asn	Asn	Ala	Ile	Asn	Leu	Thr	Ser	Thr	Ser	Leu			
			355					360					365					
Ala	Thr	Leu	Cys	Ser	Lys	Tyr	Gly	Glu	Val	Ile	Ser	Ala	Arg	Thr	Leu			
			370					375					380					
Arg	Asn	Leu	Asn	Met	Ala	Leu	Val	Glu	Phe	Ser	Ser	Val	Glu	Ser	Ala			
385				390					395					400				
Val	Lys	Ala	Leu	Asp	Ser	Leu	Gln	Gly	Lys	Glu	Val	Ser	Met	Ile	Gly			
			405					410					415					
Ala	Pro	Ser	Lys	Ile	Ser	Phe	Ala	Lys	Ile	Leu	Pro	Met	His	Gln	Gln			
			420					425					430					
Pro	Pro	Gln	Phe	Leu	Leu	Asn	Ser	Gln	Gly	Leu	Pro	Leu	Gly	Leu	Glu			
			435					440					445					
Asn	Asn	Asn	Leu	Gln	Pro	Gln	Pro	Leu	Leu	Gln	Glu	Gln	Leu	Phe	Asn			
			450					455					460					
Gly	Ala	Val	Thr	Phe	Gln	Gln	Gln	Gly	Asn	Val	Ser	Ile	Pro	Val	Phe			
465				470					475					480				

Asn	Gln	Gln	Ser	Gln	Gln	Ser	Gln	His	Gln	Asn	His	Ser	Ser	Gly	Ser			
																485	490	495
Ala	Gly	Phe	Ser	Asn	Val	Leu	His	Gly	Tyr	Asn	Asn	Asn	Asn	Ser	Met			
																500	505	510
His	Gly	Asn	Asn	Asn	Asn	Ser	Ala	Asn	Glu	Lys	Glu	Gln	Cys	Pro	Phe			
																515	520	525
Pro	Leu	Pro	Pro	Pro	Asn	Val	Asn	Glu	Lys	Glu	Asp	Leu	Leu	Arg	Glu			
																530	535	540
Ile	Ile	Glu	Leu	Phe	Glu	Ala	Asn	Ser	Asp	Glu	Tyr	Gln	Ile	Asn	Ser			
																545	550	555
Leu	Ile	Lys	Lys	Ser	Leu	Asn	His	Lys	Gly	Thr	Ser	Asp	Thr	Gln	Asn			
																565	570	575
Phe	Gly	Pro	Leu	Pro	Glu	Pro	Leu	Ser	Gly	Arg	Glu	Phe	Asp	Pro	Pro			
																580	585	590
Lys	Leu	Arg	Glu	Leu	Arg	Lys	Ser	Ile	Asp	Ser	Asn	Ala	Phe	Ser	Asp			
																595	600	605
Leu	Glu	Ile	Glu	Gln	Leu	Ala	Ile	Ala	Met	Leu	Asp	Glu	Leu	Pro	Glu			
																610	615	620
Leu	Ser	Ser	Asp	Tyr	Leu	Gly	Asn	Thr	Ile	Val	Gln	Lys	Leu	Phe	Glu			
																625	630	635
His	Ser	Ser	Asp	Ile	Ile	Lys	Asp	Ile	Met	Leu	Arg	Lys	Thr	Ser	Lys			
																645	650	655
Tyr	Leu	Thr	Ser	Met	Gly	Val	His	Lys	Asn	Gly	Thr	Trp	Ala	Cys	Gln			
																660	665	670
Lys	Met	Ile	Thr	Met	Ala	His	Thr	Pro	Arg	Gln	Ile	Met	Gln	Val	Thr			
																675	680	685
Gln	Gly	Val	Lys	Asp	Tyr	Cys	Thr	Pro	Leu	Ile	Asn	Asp	Gln	Phe	Gly			
																690	695	700
Asn	Tyr	Val	Ile	Gln	Cys	Val	Leu	Lys	Phe	Gly	Phe	Pro	Trp	Asn	Gln			
																705	710	715
Phe	Ile	Phe	Glu	Ser	Ile	Ile	Ala	Asn	Phe	Trp	Val	Ile	Val	Gln	Asn			
																725	730	735
Arg	Tyr	Gly	Ala	Arg	Ala	Val	Arg	Ala	Cys	Leu	Glu	Ala	His	Asp	Ile			
																740	745	750
Val	Thr	Pro	Glu	Gln	Ser	Ile	Val	Leu	Ser	Ala	Met	Ile	Val	Thr	Tyr			
																755	760	765
Ala	Glu	Tyr	Leu	Ser	Thr	Asn	Ser	Asn	Gly	Ala	Leu	Leu	Val	Thr	Trp			
																770	775	780
Phe	Leu	Asp	Thr	Ser	Val	Leu	Pro	Asn	Arg	His	Ser	Ile	Leu	Ala	Pro			

785	790	795	800
Arg Leu Thr Lys Arg	Ile Val Glu Leu Cys Gly His Arg	Leu Ala Ser	
805	810	815	
Leu Thr Ile Leu Lys Val	Leu Asn Tyr Arg Gly Asp Asp	Asn Ala Arg	
820	825	830	
Lys Ile Ile Leu Asp Ser	Leu Phe Gly Asn Val Asn Ala His Asp	Ser	
835	840	845	
Ser Pro Pro Lys Glu Leu Thr	Lys Leu Leu Cys Glu Thr Asn Tyr Gly		
850	855	860	
Pro Thr Phe Val His Lys Val	Leu Ala Met Pro Leu Leu Glu Asp Asp		
865	870	875	880
Leu Arg Ala His Ile Ile	Lys Gln Val Arg Lys Val Leu Thr Asp Ser		
885	890	895	
Thr Gln Ile Gln Pro Ser Arg	Arg Leu Leu Glu Glu Val Gly Leu Ala		
900	905	910	
Ser Pro Ser Ser Thr His Asn	Lys Thr Lys Gln Gln Gln Gln Gln His		
915	920	925	
His Asn Ser Ser Ile Ser His	Met Phe Ala Thr Pro Asp Thr Ser Gly		
930	935	940	
Gln His Met Arg Gly Leu Ser	Val Ser Ser Val Lys Ser Gly Gly Ser		
945	950	955	960
Lys His Thr Thr Met Asn Thr	Thr Thr Thr Asn Gly Ser Ser Ala Ser		
965	970	975	
Thr Leu Ser Pro Gly Gln Pro	Leu Asn Ala Asn Ser Asn Ser Ser Met		
980	985	990	
Gly Tyr Phe Ser Tyr Pro Gly	Val Phe Pro Val Ser Gly Phe Ser Gly		
995	1000	1005	
Asn Ala Ser Asn Gly Tyr Ala	Met Asn Asn Asp Asp Leu Ser Ser		
1010	1015	1020	
Gln Phe Asp Met Leu Asn Phe	Asn Asn Gly Thr Arg Leu Ser Leu		
1025	1030	1035	
Pro Gln Leu Ser Leu Thr Asn	His Asn Asn Thr Thr Met Glu Leu		
1040	1045	1050	
Val Asn Asn Val Gly Ser Ser	Gln Pro His Thr Asn Asn Asn Asn		
1055	1060	1065	
Asn Asn Asn Asn Thr Asn Tyr	Asn Asp Asp Asn Thr Val Phe Glu		
1070	1075	1080	
Thr Leu Thr Leu His Ser Ala	Asn		
1085	1090		

<210> 210

<211> 879

<212> PRT

<213> 未知的

<220>

<223> PUF3

<400> 210

```

Met Glu Met Asn Met Asp Met Asp Met Asp Met Glu Leu Ala Ser Ile
1           5           10           15
Val Ser Ser Leu Ser Ala Leu Ser His Ser Asn Asn Asn Gly Gly Gln
          20           25           30
Ala Ala Ala Ala Gly Ile Val Asn Gly Gly Ala Ala Gly Ser Gln Gln
          35           40           45
Ile Gly Gly Phe Arg Arg Ser Ser Phe Thr Thr Ala Asn Glu Val Asp
          50           55           60
Ser Glu Ile Leu Leu Leu His Gly Ser Ser Glu Ser Ser Pro Ile Phe
65           70           75           80
Lys Lys Thr Ala Leu Ser Val Gly Thr Ala Pro Pro Phe Ser Thr Asn
          85           90           95
Ser Lys Lys Phe Phe Gly Asn Gly Gly Asn Tyr Tyr Gln Tyr Arg Ser
          100          105          110
Thr Asp Thr Ala Ser Leu Ser Ser Ala Ser Tyr Asn Asn Tyr His Thr
          115          120          125
His His Thr Ala Ala Asn Leu Gly Lys Asn Asn Lys Val Asn His Leu
          130          135          140
Leu Gly Gln Tyr Ser Ala Ser Ile Ala Gly Pro Val Tyr Tyr Asn Gly
145          150          155          160
Asn Asp Asn Asn Asn Ser Gly Gly Glu Gly Phe Phe Glu Lys Phe Gly
          165          170          175
Lys Ser Leu Ile Asp Gly Thr Arg Glu Leu Glu Ser Gln Asp Arg Pro
          180          185          190
Asp Ala Val Asn Thr Gln Ser Gln Phe Ile Ser Lys Ser Val Ser Asn
          195          200          205
Ala Ser Leu Asp Thr Gln Asn Thr Phe Glu Gln Asn Val Glu Ser Asp
          210          215          220
Lys Asn Phe Asn Lys Leu Asn Arg Asn Thr Thr Asn Ser Gly Ser Leu
225          230          235          240
Tyr His Ser Ser Ser Asn Ser Gly Ser Ser Ala Ser Leu Glu Ser Glu
          245          250          255

```

Asn Ala His Tyr Pro Lys Arg Asn Ile Trp Asn Val Ala Asn Thr Pro		
260	265	270
Val Phe Arg Pro Ser Asn Asn Pro Ala Ala Val Gly Ala Thr Asn Val		
275	280	285
Ala Leu Pro Asn Gln Gln Asp Gly Pro Ala Asn Asn Asn Phe Pro Pro		
290	295	300
Tyr Met Asn Gly Phe Pro Pro Asn Gln Phe His Gln Gly Pro His Tyr		
305	310	315
Gln Asn Phe Pro Asn Tyr Leu Ile Gly Ser Pro Ser Asn Phe Ile Ser		
325	330	335
Gln Met Ile Ser Val Gln Ile Pro Ala Asn Glu Asp Thr Glu Asp Ser		
340	345	350
Asn Gly Lys Lys Lys Lys Lys Ala Asn Arg Pro Ser Ser Val Ser Ser		
355	360	365
Pro Ser Ser Pro Pro Asn Asn Ser Pro Phe Pro Phe Ala Tyr Pro Asn		
370	375	380
Pro Met Met Phe Met Pro Pro Pro Pro Leu Ser Ala Pro Gln Gln Gln		
385	390	395
Gln Gln Gln Gln Gln Gln Gln Gln Gln Glu Asp Gln Gln Gln Gln Gln		
405	410	415
Gln Gln Glu Asn Pro Tyr Ile Tyr Tyr Pro Thr Pro Asn Pro Ile Pro		
420	425	430
Val Lys Met Pro Lys Asp Glu Lys Thr Phe Lys Lys Arg Asn Asn Lys		
435	440	445
Asn His Pro Ala Asn Asn Ser Asn Asn Ala Asn Lys Gln Ala Asn Pro		
450	455	460
Tyr Leu Glu Asn Ser Ile Pro Thr Lys Asn Thr Ser Lys Lys Asn Ala		
465	470	475
Ser Ser Lys Ser Asn Glu Ser Thr Ala Asn Asn His Lys Ser His Ser		
485	490	495
His Ser His Pro His Ser Gln Ser Leu Gln Gln Gln Gln Gln Thr Tyr		
500	505	510
His Arg Ser Pro Leu Leu Glu Gln Leu Arg Asn Ser Ser Ser Asp Lys		
515	520	525
Asn Ser Asn Ser Asn Met Ser Leu Lys Asp Ile Phe Gly His Ser Leu		
530	535	540
Glu Phe Cys Lys Asp Gln His Gly Ser Arg Phe Ile Gln Arg Glu Leu		
545	550	555
Ala Thr Ser Pro Ala Ser Glu Lys Glu Val Ile Phe Asn Glu Ile Arg		

				565					570					575			
Asp	Asp	Ala	Ile	Glu	Leu	Ser	Asn	Asp	Val	Phe	Gly	Asn	Tyr	Val	Ile		
				580					585					590			
Gln	Lys	Phe	Phe	Glu	Phe	Gly	Ser	Lys	Ile	Gln	Lys	Asn	Thr	Leu	Val		
				595					600					605			
Asp	Gln	Phe	Lys	Gly	Asn	Met	Lys	Gln	Leu	Ser	Leu	Gln	Met	Tyr	Ala		
				610					615					620			
Cys	Arg	Val	Ile	Gln	Lys	Ala	Leu	Glu	Tyr	Ile	Asp	Ser	Asn	Gln	Arg		
625						630					635				640		
Ile	Glu	Leu	Val	Leu	Glu	Leu	Ser	Asp	Ser	Val	Leu	Gln	Met	Ile	Lys		
				645						650					655		
Asp	Gln	Asn	Gly	Asn	His	Val	Ile	Gln	Lys	Ala	Ile	Glu	Thr	Ile	Pro		
				660					665						670		
Ile	Glu	Lys	Leu	Pro	Phe	Ile	Leu	Ser	Ser	Leu	Thr	Gly	His	Ile	Tyr		
				675					680						685		
His	Leu	Ser	Thr	His	Ser	Tyr	Gly	Cys	Arg	Val	Ile	Gln	Arg	Leu	Leu		
				690					695						700		
Glu	Phe	Gly	Ser	Ser	Glu	Asp	Gln	Glu	Ser	Ile	Leu	Asn	Glu	Leu	Lys		
705						710					715				720		
Asp	Phe	Ile	Pro	Tyr	Leu	Ile	Gln	Asp	Gln	Tyr	Gly	Asn	Tyr	Val	Ile		
				725						730					735		
Gln	Tyr	Val	Leu	Gln	Gln	Asp	Gln	Phe	Thr	Asn	Lys	Glu	Met	Val	Asp		
				740					745						750		
Ile	Lys	Gln	Glu	Ile	Ile	Glu	Thr	Val	Ala	Asn	Asn	Val	Val	Glu	Tyr		
				755					760						765		
Ser	Lys	His	Lys	Phe	Ala	Ser	Asn	Val	Val	Glu	Lys	Ser	Ile	Leu	Tyr		
				770					775						780		
Gly	Ser	Lys	Asn	Gln	Lys	Asp	Leu	Ile	Ile	Ser	Lys	Ile	Leu	Pro	Arg		
785						790					795				800		
Asp	Lys	Asn	His	Ala	Leu	Asn	Leu	Glu	Asp	Asp	Ser	Pro	Met	Ile	Leu		
				805						810					815		
Met	Ile	Lys	Asp	Gln	Phe	Ala	Asn	Tyr	Val	Ile	Gln	Lys	Leu	Val	Asn		
				820					825						830		
Val	Ser	Glu	Gly	Glu	Gly	Lys	Lys	Leu	Ile	Val	Ile	Ala	Ile	Arg	Ala		
				835					840						845		
Tyr	Leu	Asp	Lys	Leu	Asn	Lys	Ser	Asn	Ser	Leu	Gly	Asn	Arg	His	Leu		
				850					855						860		
Ala	Ser	Val	Glu	Lys	Leu	Ala	Ala	Leu	Val	Glu	Asn	Ala	Glu	Val			
865						870									875		

<210> 211

<211> 888

<212> PRT

<213> 未知的

<220>

<223> PUF4

<400> 211

```

Met Ser Thr Lys Gly Leu Lys Glu Glu Ile Asp Asp Val Pro Ser Val
1           5           10           15
Asp Pro Val Val Ser Glu Thr Val Asn Ser Ala Leu Glu Gln Leu Gln
          20           25           30
Leu Asp Asp Pro Glu Glu Asn Ala Thr Ser Asn Ala Phe Ala Asn Lys
          35           40           45
Val Ser Gln Asp Ser Gln Phe Ala Asn Gly Pro Pro Ser Gln Met Phe
          50           55           60
Pro His Pro Gln Met Met Gly Gly Met Gly Phe Met Pro Tyr Ser Gln
65           70           75           80
Met Met Gln Val Pro His Asn Pro Cys Pro Phe Phe Pro Pro Pro Asp
          85           90           95
Phe Asn Asp Pro Thr Ala Pro Leu Ser Ser Ser Pro Leu Asn Ala Gly
          100          105          110
Gly Pro Pro Met Leu Phe Lys Asn Asp Ser Leu Pro Phe Gln Met Leu
          115          120          125
Ser Ser Gly Ala Ala Val Ala Thr Gln Gly Gly Gln Asn Leu Asn Pro
          130          135          140
Leu Ile Asn Asp Asn Ser Met Lys Val Leu Pro Ile Ala Ser Ala Asp
145          150          155          160
Pro Leu Trp Thr His Ser Asn Val Pro Gly Ser Ala Ser Val Ala Ile
          165          170          175
Glu Glu Thr Thr Ala Thr Leu Gln Glu Ser Leu Pro Ser Lys Gly Arg
          180          185          190
Glu Ser Asn Asn Lys Ala Ser Ser Phe Arg Arg Gln Thr Phe His Ala
          195          200          205
Leu Ser Pro Thr Asp Leu Ile Asn Ala Ala Asn Asn Val Thr Leu Ser
          210          215          220
Lys Asp Phe Gln Ser Asp Met Gln Asn Phe Ser Lys Ala Lys Lys Pro
225          230          235          240
Ser Val Gly Ala Asn Asn Thr Ala Lys Thr Arg Thr Gln Ser Ile Ser
          245          250          255

```

Phe Asp Asn Thr Pro Ser Ser Thr Ser Phe Ile Pro Pro Thr Asn Ser		
260	265	270
Val Ser Glu Lys Leu Ser Asp Phe Lys Ile Glu Thr Ser Lys Glu Asp		
275	280	285
Leu Ile Asn Lys Thr Ala Pro Ala Lys Lys Glu Ser Pro Thr Thr Tyr		
290	295	300
Gly Ala Ala Tyr Pro Tyr Gly Gly Pro Leu Leu Gln Pro Asn Pro Ile		
305	310	315
Met Pro Gly His Pro His Asn Ile Ser Ser Pro Ile Tyr Gly Ile Arg		
325	330	335
Ser Pro Phe Pro Asn Ser Tyr Glu Met Gly Ala Gln Phe Gln Pro Phe		
340	345	350
Ser Pro Ile Leu Asn Pro Thr Ser His Ser Leu Asn Ala Asn Ser Pro		
355	360	365
Ile Pro Leu Thr Gln Ser Pro Ile His Leu Ala Pro Val Leu Asn Pro		
370	375	380
Ser Ser Asn Ser Val Ala Phe Ser Asp Met Lys Asn Asp Gly Gly Lys		
385	390	395
Pro Thr Thr Asp Asn Asp Lys Ala Gly Pro Asn Val Arg Met Asp Leu		
405	410	415
Ile Asn Pro Asn Leu Gly Pro Ser Met Gln Pro Phe His Ile Leu Pro		
420	425	430
Pro Gln Gln Asn Thr Pro Pro Pro Pro Trp Leu Tyr Ser Thr Pro Pro		
435	440	445
Pro Phe Asn Ala Met Val Pro Pro His Leu Leu Ala Gln Asn His Met		
450	455	460
Pro Leu Met Asn Ser Ala Asn Asn Lys His His Gly Arg Asn Asn Asn		
465	470	475
Ser Met Ser Ser His Asn Asp Asn Asp Asn Ile Gly Asn Ser Asn Tyr		
485	490	495
Asn Asn Lys Asp Thr Gly Arg Ser Asn Val Gly Lys Met Lys Asn Met		
500	505	510
Lys Asn Ser Tyr His Gly Tyr Tyr Asn Asn Asn Asn Asn Asn Asn		
515	520	525
Asn Asn Asn Asn Asn Asn Asn Ser Asn Ala Thr Asn Ser Asn Ser Ala		
530	535	540
Glu Lys Gln Arg Lys Ile Glu Glu Ser Ser Arg Phe Ala Asp Ala Val		
545	550	555
Leu Asp Gln Tyr Ile Gly Ser Ile His Ser Leu Cys Lys Asp Gln His		

				565					570					575			
Gly	Cys	Arg	Phe	Leu	Gln	Lys	Gln	Leu	Asp	Ile	Leu	Gly	Ser	Lys	Ala		
				580					585					590			
Ala	Asp	Ala	Ile	Phe	Glu	Glu	Thr	Lys	Asp	Tyr	Thr	Val	Glu	Leu	Met		
				595				600						605			
Thr	Asp	Ser	Phe	Gly	Asn	Tyr	Leu	Ile	Gln	Lys	Leu	Leu	Glu	Glu	Val		
				610				615						620			
Thr	Thr	Glu	Gln	Arg	Ile	Val	Leu	Thr	Lys	Ile	Ser	Ser	Pro	His	Phe		
625					630				635					640			
Val	Glu	Ile	Ser	Leu	Asn	Pro	His	Gly	Thr	Arg	Ala	Leu	Gln	Lys	Leu		
				645					650					655			
Ile	Glu	Cys	Ile	Lys	Thr	Asp	Glu	Glu	Ala	Gln	Ile	Val	Val	Asp	Ser		
				660					665					670			
Leu	Arg	Pro	Tyr	Thr	Val	Gln	Leu	Ser	Lys	Asp	Leu	Asn	Gly	Asn	His		
				675					680					685			
Val	Ile	Gln	Lys	Cys	Leu	Gln	Arg	Leu	Lys	Pro	Glu	Asn	Phe	Gln	Phe		
				690					695					700			
Ile	Phe	Asp	Ala	Ile	Ser	Asp	Ser	Cys	Ile	Asp	Ile	Ala	Thr	His	Arg		
705					710				715					720			
His	Gly	Cys	Cys	Val	Leu	Gln	Arg	Cys	Leu	Asp	His	Gly	Thr	Thr	Glu		
				725					730					735			
Gln	Cys	Asp	Asn	Leu	Cys	Asp	Lys	Leu	Leu	Ala	Leu	Val	Asp	Lys	Leu		
				740					745					750			
Thr	Leu	Asp	Pro	Phe	Gly	Asn	Tyr	Val	Val	Gln	Tyr	Ile	Ile	Thr	Lys		
				755					760					765			
Glu	Ala	Glu	Lys	Asn	Lys	Tyr	Asp	Tyr	Thr	His	Lys	Ile	Val	His	Leu		
				770					775					780			
Leu	Lys	Pro	Arg	Ala	Ile	Glu	Leu	Ser	Ile	His	Lys	Phe	Gly	Ser	Asn		
785					790				795					800			
Val	Ile	Glu	Lys	Ile	Leu	Lys	Thr	Ala	Ile	Val	Ser	Glu	Pro	Met	Ile		
				805					810					815			
Leu	Glu	Ile	Leu	Asn	Asn	Gly	Gly	Glu	Thr	Gly	Ile	Gln	Ser	Leu	Leu		
				820					825					830			
Asn	Asp	Ser	Tyr	Gly	Asn	Tyr	Val	Leu	Gln	Thr	Ala	Leu	Asp	Ile	Ser		
				835					840					845			
His	Lys	Gln	Asn	Asp	Tyr	Leu	Tyr	Lys	Arg	Leu	Ser	Glu	Ile	Val	Ala		
				850					855					860			
Pro	Leu	Leu	Val	Gly	Pro	Ile	Arg	Asn	Thr	Pro	His	Gly	Lys	Arg	Ile		
865					870				875					880			

Ile Gly Met Leu His Leu Asp Ser

885

<210> 212

<211> 553

<212> PRT

<213> 未知的

<220>

<223> PUF5

<400> 212

Met Ser Asp Ser Thr Gly Arg Ile Asn Ser Lys Ala Ser Asp Ser Ser

1 5 10 15

Ser Ile Ser Asp His Gln Thr Ala Asp Leu Ser Ile Phe Asn Gly Ser

20 25 30

Phe Asp Gly Gly Ala Phe Ser Ser Ser Asn Ile Pro Leu Phe Asn Phe

35 40 45

Met Gly Thr Gly Asn Gln Arg Phe Gln Tyr Ser Pro His Pro Phe Ala

50 55 60

Lys Ser Ser Asp Pro Cys Arg Leu Ala Ala Leu Thr Pro Ser Thr Pro

65 70 75 80

Lys Gly Pro Leu Asn Leu Thr Pro Ala Asp Phe Gly Leu Ala Asp Phe

85 90 95

Ser Val Gly Asn Glu Ser Phe Ala Asp Phe Thr Ala Asn Asn Thr Ser

100 105 110

Phe Val Gly Asn Val Gln Ser Asn Val Arg Ser Thr Arg Leu Leu Pro

115 120 125

Ala Trp Ala Val Asp Asn Ser Gly Asn Ile Arg Asp Asp Leu Thr Leu

130 135 140

Gln Asp Val Val Ser Asn Gly Ser Leu Ile Asp Phe Ala Met Asp Arg

145 150 155 160

Thr Gly Val Lys Phe Leu Glu Arg His Phe Pro Glu Asp His Asp Asn

165 170 175

Glu Met His Phe Val Leu Phe Asp Lys Leu Thr Glu Gln Gly Ala Val

180 185 190

Phe Thr Ser Leu Cys Arg Ser Ala Ala Gly Asn Phe Ile Ile Gln Lys

195 200 205

Phe Val Glu His Ala Thr Leu Asp Glu Gln Glu Arg Leu Val Arg Lys

210 215 220

Met Cys Asp Asn Gly Leu Ile Glu Met Cys Leu Asp Lys Phe Ala Cys

225 230 235 240

Arg Val Val Gln Met Ser Ile Gln Lys Phe Asp Val Ser Ile Ala Met		
	245	250 255
Lys Leu Val Glu Lys Ile Ser Ser Leu Asp Phe Leu Pro Leu Cys Thr		
	260	265 270
Asp Gln Cys Ala Ile His Val Leu Gln Lys Val Val Lys Leu Leu Pro		
	275	280 285
Ile Ser Ala Trp Ser Phe Phe Val Lys Phe Leu Cys Arg Asp Asp Asn		
	290	295 300
Leu Met Thr Val Cys Gln Asp Lys Tyr Gly Cys Arg Leu Val Gln Gln		
305	310	315 320
Thr Ile Asp Lys Leu Ser Asp Asn Pro Lys Leu His Cys Phe Asn Thr		
	325	330 335
Arg Leu Gln Leu Leu His Gly Leu Met Thr Ser Val Ala Arg Asn Cys		
	340	345 350
Phe Arg Leu Ser Ser Asn Glu Phe Ala Asn Tyr Val Val Gln Tyr Val		
	355	360 365
Ile Lys Ser Ser Gly Val Met Glu Met Tyr Arg Asp Thr Ile Ile Glu		
	370	375 380
Lys Cys Leu Leu Arg Asn Ile Leu Ser Met Ser Gln Asp Lys Tyr Ala		
385	390	395 400
Ser His Val Val Glu Gly Ala Phe Leu Phe Ala Pro Pro Leu Leu Leu		
	405	410 415
Ser Glu Met Met Asp Glu Ile Phe Asp Gly Tyr Val Lys Asp Gln Glu		
	420	425 430
Thr Asn Arg Asp Ala Leu Asp Ile Leu Leu Phe His Gln Tyr Gly Asn		
	435	440 445
Tyr Val Val Gln Gln Met Ile Ser Ile Cys Ile Ser Ala Leu Leu Gly		
	450	455 460
Lys Glu Glu Arg Lys Met Val Ala Ser Glu Met Arg Leu Tyr Ala Lys		
465	470	475 480
Trp Phe Asp Arg Ile Lys Asn Arg Val Asn Arg His Ser Gly Arg Leu		
	485	490 495
Glu Arg Phe Ser Ser Gly Lys Lys Ile Ile Glu Ser Leu Gln Lys Leu		
	500	505 510
Asn Val Pro Met Thr Met Thr Asn Glu Pro Met Pro Tyr Trp Ala Met		
	515	520 525
Pro Thr Pro Leu Met Asp Ile Ser Ala His Phe Met Asn Lys Leu Asn		
	530	535 540
Phe Gln Lys Asn Ser Val Phe Asp Glu		

545 550
 <210> 213
 <211> 485
 <212> PRT
 <213> 未知的
 <220>
 <223> PUF6
 <400> 213
 Met Thr Pro Asn Arg Arg Ser Thr Asp Ser Tyr Asn Met Leu Gly Ala
 1 5 10 15
 Ser Phe Asp Phe Asp Pro Asp Phe Ser Leu Leu Ser Asn Lys Thr His
 20 25 30
 Lys Asn Lys Asn Pro Lys Pro Pro Val Lys Leu Leu Pro Tyr Arg His
 35 40 45
 Gly Ser Asn Thr Thr Ser Ser Asp Leu Asp Asn Tyr Ile Phe Asn Ser
 50 55 60
 Gly Ser Gly Ser Ser Asp Asp Glu Thr Pro Pro Pro Ala Ala Pro Ile
 65 70 75 80
 Phe Ile Ser Leu Glu Glu Val Leu Leu Asn Gly Leu Leu Ile Asp Phe
 85 90 95
 Ala Ile Asp Pro Ser Gly Val Lys Phe Leu Glu Ala Asn Tyr Pro Leu
 100 105 110
 Asp Ser Glu Asp Gln Ile Arg Lys Ala Val Phe Glu Lys Leu Thr Glu
 115 120 125
 Ser Thr Thr Leu Phe Val Gly Leu Cys His Ser Arg Asn Gly Asn Phe
 130 135 140
 Ile Val Gln Lys Leu Val Glu Leu Ala Thr Pro Ala Glu Gln Arg Glu
 145 150 155 160
 Leu Leu Arg Gln Met Ile Asp Gly Gly Leu Leu Val Met Cys Lys Asp
 165 170 175
 Lys Phe Ala Cys Arg Val Val Gln Leu Ala Leu Gln Lys Phe Asp His
 180 185 190
 Ser Asn Val Phe Gln Leu Ile Gln Glu Leu Ser Thr Phe Asp Leu Ala
 195 200 205
 Ala Met Cys Thr Asp Gln Ile Ser Ile His Val Ile Gln Arg Val Val
 210 215 220
 Lys Gln Leu Pro Val Asp Met Trp Thr Phe Phe Val His Phe Leu Ser
 225 230 235 240
 Ser Gly Asp Ser Leu Met Ala Val Cys Gln Asp Lys Tyr Gly Cys Arg

245								250					255				
Leu	Val	Gln	Gln	Val	Ile	Asp	Arg	Leu	Ala	Glu	Asn	Pro	Lys	Leu	Pro		
260								265					270				
Cys	Phe	Lys	Phe	Arg	Ile	Gln	Leu	Leu	His	Ser	Leu	Met	Thr	Cys	Ile		
275								280					285				
Val	Arg	Asn	Cys	Tyr	Arg	Leu	Ser	Ser	Asn	Glu	Phe	Ala	Asn	Tyr	Val		
290								295					300				
Ile	Gln	Tyr	Val	Ile	Lys	Ser	Ser	Gly	Ile	Met	Glu	Met	Tyr	Arg	Asp		
305								310					315				
Thr	Ile	Ile	Asp	Lys	Cys	Leu	Leu	Arg	Asn	Leu	Leu	Ser	Met	Ser	Gln		
325								330					335				
Asp	Lys	Tyr	Ala	Ser	His	Val	Ile	Glu	Gly	Ala	Phe	Leu	Phe	Ala	Pro		
340								345					350				
Pro	Ala	Leu	Leu	His	Glu	Met	Met	Glu	Glu	Ile	Phe	Ser	Gly	Tyr	Val		
355								360					365				
Lys	Asp	Val	Glu	Leu	Asn	Arg	Asp	Ala	Leu	Asp	Ile	Leu	Leu	Phe	His		
370								375					380				
Gln	Tyr	Gly	Asn	Tyr	Val	Val	Gln	Gln	Met	Ile	Ser	Ile	Cys	Thr	Ala		
385								390					395				
Ala	Leu	Ile	Gly	Lys	Glu	Glu	Arg	Gln	Leu	Pro	Pro	Ala	Ile	Leu	Leu		
405								410					415				
Leu	Tyr	Ser	Gly	Trp	Tyr	Glu	Lys	Met	Lys	Gln	Arg	Val	Leu	Gln	His		
420								425					430				
Ala	Ser	Arg	Leu	Glu	Arg	Phe	Ser	Ser	Gly	Lys	Lys	Ile	Ile	Asp	Ser		
435								440					445				
Val	Met	Arg	His	Gly	Val	Pro	Thr	Ala	Ala	Ala	Ile	Asn	Ala	Gln	Ala		
450								455					460				
Ala	Pro	Ser	Leu	Met	Glu	Leu	Thr	Ala	Gln	Phe	Asp	Ala	Met	Phe	Pro		
465								470					475				
Ser	Phe	Leu	Ala	Arg													
485																	

<210> 214

<211> 485

<212> PRT

<213> 未知的

<220>

<223> PUF7

<400> 214

Met Thr Pro Asn Arg Arg Ser Thr Asp Ser Tyr Asn Met Leu Gly Ala

1	5	10	15
Ser Phe Asp Phe Asp Pro Asp Phe Ser Leu Leu Ser Asn Lys Thr His			
20	25	30	
Lys Asn Lys Asn Pro Lys Pro Pro Val Lys Leu Leu Pro Tyr Arg His			
35	40	45	
Gly Ser Asn Thr Thr Ser Ser Asp Ser Asp Ser Tyr Ile Phe Asn Ser			
50	55	60	
Gly Ser Gly Ser Ser Asp Ala Glu Thr Pro Ala Pro Val Ala Pro Ile			
65	70	75	80
Phe Ile Ser Leu Glu Asp Val Leu Leu Asn Gly Gln Leu Ile Asp Phe			
85	90	95	
Ala Ile Asp Pro Ser Gly Val Lys Phe Leu Glu Ala Asn Tyr Pro Leu			
100	105	110	
Asp Ser Glu Asp Gln Ile Arg Lys Ala Val Phe Glu Lys Phe Thr Glu			
115	120	125	
Ser Thr Thr Leu Phe Val Gly Leu Cys His Ser Arg Asn Gly Asn Phe			
130	135	140	
Ile Val Gln Lys Leu Val Glu Leu Ala Thr Pro Ala Glu Gln Arg Glu			
145	150	155	160
Leu Leu Arg Gln Met Ile Asp Gly Gly Leu Leu Ala Met Cys Lys Asp			
165	170	175	
Lys Phe Ala Cys Arg Val Val Gln Leu Ala Leu Gln Lys Phe Asp His			
180	185	190	
Ser Asn Val Phe Gln Leu Ile Gln Glu Leu Ser Thr Phe Asp Leu Ala			
195	200	205	
Ala Met Cys Thr Asp Gln Ile Ser Ile His Val Ile Gln Arg Val Val			
210	215	220	
Lys Gln Leu Pro Val Asp Met Trp Thr Phe Phe Val His Phe Leu Ser			
225	230	235	240
Ser Gly Asp Ser Leu Met Ala Val Cys Gln Asp Lys Tyr Gly Cys Arg			
245	250	255	
Leu Val Gln Gln Val Ile Asp Arg Leu Ala Glu Asn Pro Lys Leu Pro			
260	265	270	
Cys Phe Lys Phe Arg Ile Gln Leu Leu His Ser Leu Met Thr Cys Ile			
275	280	285	
Val Arg Asn Cys Tyr Arg Leu Ser Ser Asn Glu Phe Ala Asn Tyr Val			
290	295	300	
Ile Gln Tyr Val Ile Lys Ser Ser Gly Ile Met Glu Met Tyr Arg Asp			
305	310	315	320

Thr Ile Ile Asp Lys Cys Leu Leu Arg Asn Leu Leu Ser Met Ser Gln
 325 330 335
 Asp Lys Tyr Ala Ser His Val Ile Glu Gly Ala Phe Leu Phe Ala Pro
 340 345 350
 Pro Ala Leu Leu His Glu Met Met Glu Glu Ile Phe Ser Gly Tyr Val
 355 360 365
 Lys Asp Val Glu Ser Asn Arg Asp Ala Leu Asp Ile Leu Leu Phe His
 370 375 380
 Gln Tyr Gly Asn Tyr Val Val Gln Gln Met Ile Ser Ile Cys Thr Ala
 385 390 395 400
 Ala Leu Ile Gly Lys Glu Glu Arg Glu Leu Pro Pro Ala Ile Leu Leu
 405 410 415
 Leu Tyr Ser Gly Trp Tyr Glu Lys Met Lys Gln Arg Val Leu Gln His
 420 425 430
 Ala Ser Arg Leu Glu Arg Phe Ser Ser Gly Lys Lys Ile Ile Asp Ser
 435 440 445
 Val Met Arg His Gly Val Pro Thr Ala Ala Ala Val Asn Ala Gln Ala
 450 455 460
 Ala Pro Ser Leu Met Glu Leu Thr Ala Gln Phe Asp Ala Met Phe Pro
 465 470 475 480
 Ser Phe Leu Ala Arg
 485

<210> 215

<211> 535

<212> PRT

<213> 未知的

<220>

<223> PUF8

<400> 215

Met Ser Arg Pro Ile Ser Ile Gly Asn Thr Cys Thr Phe Asp Pro Ser
 1 5 10 15
 Ala Ser Pro Ile Glu Ser Leu Gly Arg Ser Ile Gly Ala Gln Lys Ile
 20 25 30
 Val Asp Ser Val Cys Gly Ser Pro Ile Arg Ser Tyr Gly Arg His Ile
 35 40 45
 Ser Thr Asn Pro Lys Asn Glu Arg Leu Pro Asp Thr Pro Glu Phe Gln
 50 55 60
 Phe Ala Thr Tyr Met His Gln Gly Gly Lys Val Ile Gly Gln Asn Thr
 65 70 75 80

Leu His Met Phe Gly Thr Pro Pro Ser Cys Tyr Cys Ala Gln Glu Asn		
85	90	95
Ile Pro Ile Ser Ser Asn Val Gly His Val Leu Ser Thr Ile Asn Asn		
100	105	110
Asn Tyr Met Asn His Gln Tyr Asn Gly Ser Asn Met Phe Ser Asn Gln		
115	120	125
Met Thr Gln Met Leu Gln Ala Gln Ala Tyr Asn Asp Leu Gln Met His		
130	135	140
Gln Ala His Ser Gln Ser Ile Arg Val Pro Val Gln Pro Ser Ala Thr		
145	150	155
Gly Ile Phe Ser Asn Pro Tyr Arg Glu Pro Thr Thr Thr Asp Asp Leu		
165	170	175
Leu Thr Arg Tyr Arg Ala Asn Pro Ala Met Met Lys Asn Leu Lys Leu		
180	185	190
Ser Asp Ile Arg Gly Ala Leu Leu Lys Phe Ala Lys Asp Gln Val Gly		
195	200	205
Ser Arg Phe Ile Gln Gln Glu Leu Ala Ser Ser Lys Asp Arg Phe Glu		
210	215	220
Lys Asp Ser Ile Phe Asp Glu Val Val Ser Asn Ala Asp Glu Leu Val		
225	230	235
Asp Asp Ile Phe Gly Asn Tyr Val Val Gln Lys Phe Phe Glu Tyr Gly		
245	250	255
Glu Glu Arg His Trp Ala Arg Leu Val Asp Ala Ile Ile Asp Arg Val		
260	265	270
Pro Glu Tyr Ala Phe Gln Met Tyr Ala Cys Arg Val Leu Gln Lys Ala		
275	280	285
Leu Glu Lys Ile Asn Glu Pro Leu Gln Ile Lys Ile Leu Ser Gln Ile		
290	295	300
Arg His Val Ile His Arg Cys Met Lys Asp Gln Asn Gly Asn His Val		
305	310	315
Val Gln Lys Ala Ile Glu Lys Val Ser Pro Gln Tyr Val Gln Phe Ile		
325	330	335
Val Asp Thr Leu Leu Glu Ser Ser Asn Thr Ile Tyr Glu Met Ser Val		
340	345	350
Asp Pro Tyr Gly Cys Arg Val Val Gln Arg Cys Leu Glu His Cys Ser		
355	360	365
Pro Ser Gln Thr Lys Pro Val Ile Gly Gln Ile His Lys Arg Phe Asp		
370	375	380
Glu Ile Ala Asn Asn Gln Tyr Gly Asn Tyr Val Val Gln His Val Ile		

385	390	395	400
Glu His Gly Ser Glu Glu Asp Arg Met Val Ile Val Thr Arg Val Ser			
	405	410	415
Asn Asn Leu Phe Glu Phe Ala Thr His Lys Tyr Ser Ser Asn Val Ile			
	420	425	430
Glu Lys Cys Leu Glu Gln Gly Ala Val Tyr His Lys Ser Met Ile Val			
	435	440	445
Gly Ala Ala Cys His His Gln Glu Gly Ser Val Pro Ile Val Val Gln			
	450	455	460
Met Met Lys Asp Gln Tyr Ala Asn Tyr Val Val Gln Lys Met Phe Asp			
465	470	475	480
Gln Val Thr Ser Glu Gln Arg Arg Glu Leu Ile Leu Thr Val Arg Pro			
	485	490	495
His Ile Pro Val Leu Arg Gln Phe Pro His Gly Lys His Ile Leu Ala			
	500	505	510
Lys Leu Glu Lys Tyr Phe Gln Lys Pro Ala Val Met Ser Tyr Pro Tyr			
	515	520	525
Gln Asp Met Gln Gly Ser His			
	530	535	
<210> 216			
<211> 703			
<212> PRT			
<213> 未知的			
<220>			
<223> PUF9			
<400> 216			
Met Ala Asp Pro Asn Trp Ala Tyr Ala Pro Pro Thr Asn Tyr Tyr Ala			
1	5	10	15
Asp His Ser Ile Ala Lys Pro Ile Met Ile Ser Gly Gly His Pro Ser			
	20	25	30
Gln Asp Gln Gly His Ser Pro Lys Ser Glu Ser Phe Gly Gln Ser Val			
	35	40	45
Thr Thr Ala Phe Asn Gly Met Val Asp Asn Leu Val Gly Ser Pro Ser			
	50	55	60
Ser Ser Val Gln Gln Arg Asn Tyr Phe Thr Thr Thr Pro Phe Pro Ile			
65	70	75	80
Ser Arg Ser Pro Asn Asp Arg Asn Asp Asp Lys Ile Met Gly Asn Gly			
	85	90	95
Ser Tyr Gly Val Pro Ile Pro Ile Pro Gln Asp Gly Val Pro Gln Gly			

100	105	110
Thr Pro Asp Phe Gln Met Thr	Pro Phe Leu Gln Gln Gly	Gly His Leu
115	120	125
Ile Gly Gly Ser Pro Asn Gly	Pro Val Gln Val Ser Gly	Asn Trp Tyr
130	135	140
Ser Gly Gly Ala Gly Ile Phe	Ser Thr Met Gln Gln Ala	Asp Pro Ser
145	150	155
Asn Gly Met Pro Gly Met Ala	Ala Glu Phe Val Asn Asn	Glu Asn Gly
165	170	175
Met Pro Gly Pro Asn Gly Met	His Gln Gln Ala Met Ile	Ser Gly Ser
180	185	190
Pro Pro Phe Pro Tyr Gln Asn	Met Met Asn Leu Thr Thr	Ser Phe Gly
195	200	205
Ala Met Gly Leu Gly Pro Gln	Gln Ile Gln Gln Arg Asp	Pro Gln Met
210	215	220
Phe Gln Gln Pro Ile Leu His	Glu Pro Ile Gln Gly Met	Ala Gln Asn
225	230	235
Gly Phe Gly Gln Gln Val Phe	Phe Thr Gln Met Gln Asn	Gln Gln His
245	250	255
Pro Gln Gly Gln Ala Gln Gln	Gln Leu Gln Gln Leu Ala	Gln Gln His
260	265	270
Gln Gln Gln Gln Asn Ser Gln	Gln Phe Phe Gly Gln Gly	Pro Asn Gly
275	280	285
Met Gly Asn Gly Gly Val Met	Asn Asp Trp Ser Gln Arg	Ser Phe Gly
290	295	300
Met Pro Gln Gln Gln Ala Gln	Gln Asn Gly Leu Pro Pro	Asn Phe Ser
305	310	315
Gln Asn Pro Pro Arg Arg Arg	Gly Pro Glu Asp Pro Asn	Gly Gln Thr
325	330	335
Pro Lys Thr Leu Gln Asp Ile	Lys Asn Asn Val Ile Glu	Phe Ala Lys
340	345	350
Asp Gln His Gly Ser Arg Phe	Ile Gln Gln Lys Leu Glu	Arg Ala Ser
355	360	365
Leu Arg Asp Lys Ala Ala Ile	Phe Thr Pro Val Leu Glu	Asn Ala Glu
370	375	380
Glu Leu Met Thr Asp Val Phe	Gly Asn Tyr Val Ile Gln	Lys Phe Phe
385	390	395
Glu Phe Gly Asn Asn Glu Gln	Arg Asn Gln Leu Val Gly	Thr Ile Arg
405	410	415

Gly Asn Val Met Lys Leu Ala Leu Gln Met Tyr Gly Cys Arg Val Ile		
420	425	430
Gln Lys Ala Leu Glu Tyr Val Glu Glu Lys Tyr Gln His Glu Ile Leu		
435	440	445
Gly Glu Met Glu Gly Gln Val Leu Lys Cys Val Lys Asp Gln Asn Gly		
450	455	460
Asn His Val Ile Gln Lys Val Ile Glu Arg Val Glu Pro Glu Arg Leu		
465	470	475
Gln Phe Ile Ile Asp Ala Phe Thr Lys Asn Asn Ser Asp Asn Val Tyr		
485	490	495
Thr Leu Ser Val His Pro Tyr Gly Cys Arg Val Ile Gln Arg Val Leu		
500	505	510
Glu Tyr Cys Asn Glu Glu Gln Lys Gln Pro Val Leu Asp Ala Leu Gln		
515	520	525
Ile His Leu Lys Gln Leu Val Leu Asp Gln Tyr Gly Asn Tyr Val Ile		
530	535	540
Gln His Val Ile Glu His Gly Ser Pro Ser Asp Lys Glu Gln Ile Val		
545	550	555
Gln Asp Val Ile Ser Asp Asp Leu Leu Lys Phe Ala Gln His Lys Phe		
565	570	575
Ala Ser Asn Val Ile Glu Lys Cys Leu Thr Phe Gly Gly His Ala Glu		
580	585	590
Arg Asn Leu Ile Ile Asp Lys Val Cys Gly Asp Pro Asn Asp Pro Ser		
595	600	605
Pro Pro Leu Leu Gln Met Met Lys Asp Pro Phe Ala Asn Tyr Val Val		
610	615	620
Gln Lys Met Leu Asp Val Ala Asp Pro Gln His Arg Lys Lys Ile Thr		
625	630	635
Leu Thr Ile Lys Pro His Ile Ala Thr Leu Arg Lys Tyr Asn Phe Gly		
645	650	655
Lys His Ile Leu Leu Lys Leu Glu Lys Tyr Phe Ala Lys Gln Ala Pro		
660	665	670
Ala Asn Ser Ser Asn Ser Ser Ser Asn Asp Gln Ile Tyr Glu His Ser		
675	680	685
Pro Phe Asp Ile Pro Leu Gly Ala Asp Phe Ser Asn His Pro Phe		
690	695	700

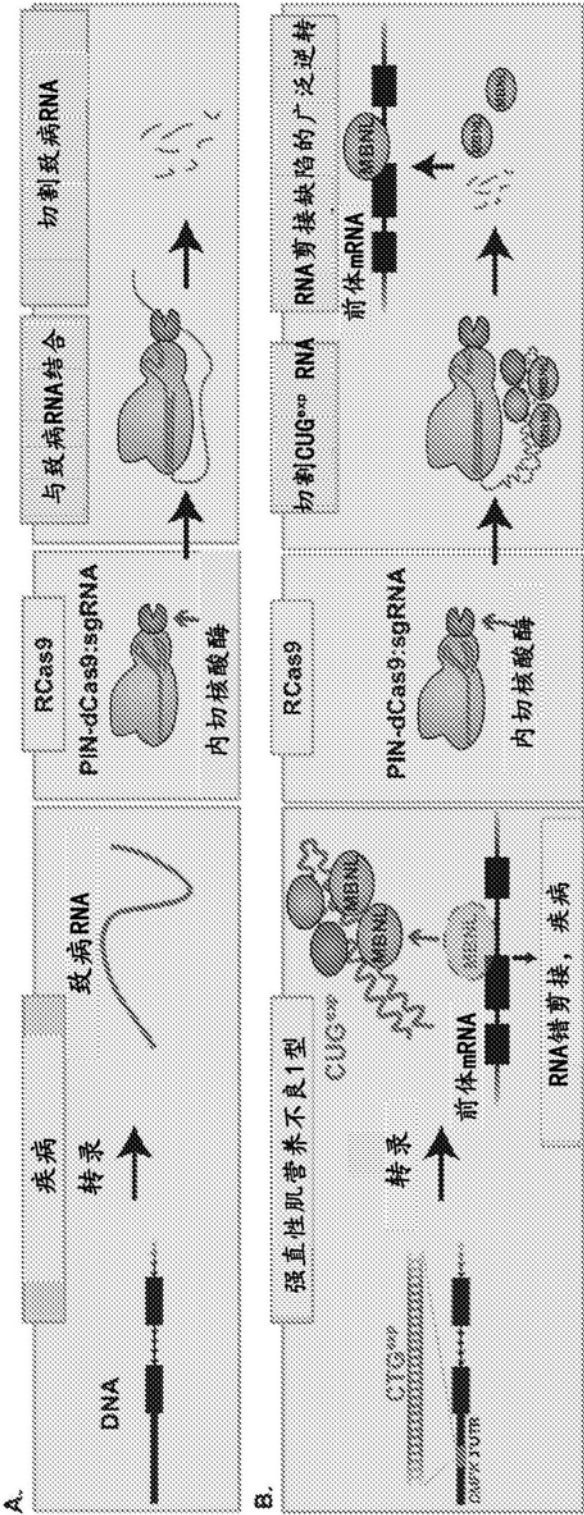


FIGURE 1

模块化治疗平台

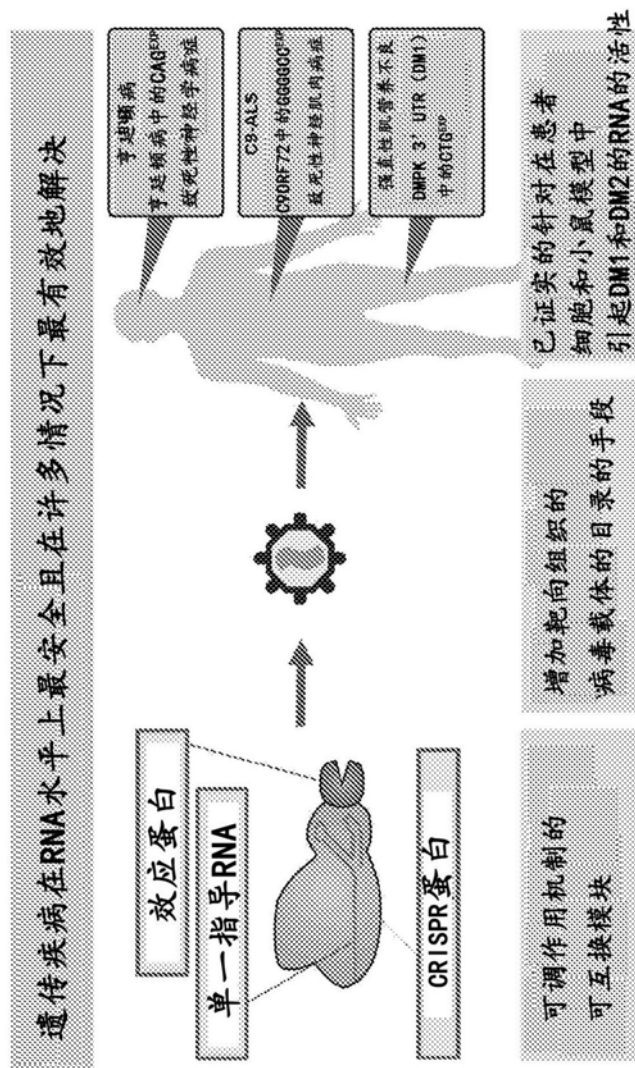


FIGURE 2

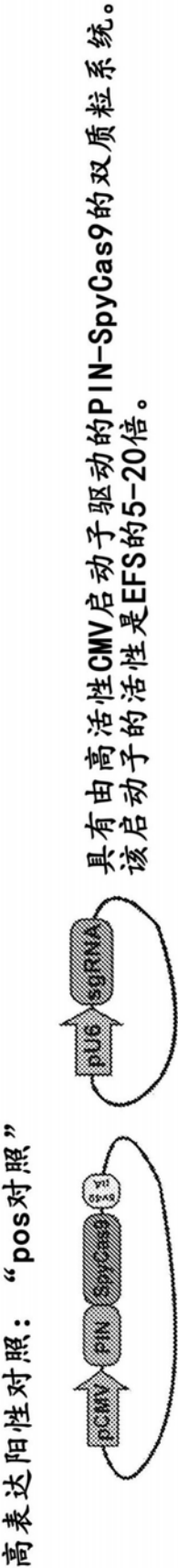
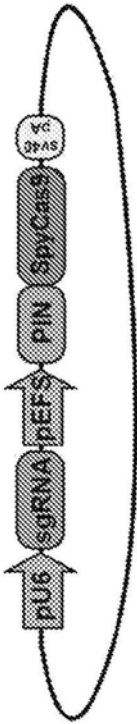


FIGURE 3A

低表达阳性对照：“P13”



融合物表达较低的单质核酶系统。融合物相同
架构与包括新型内切核酶的融合物相同

FIGURE 3B

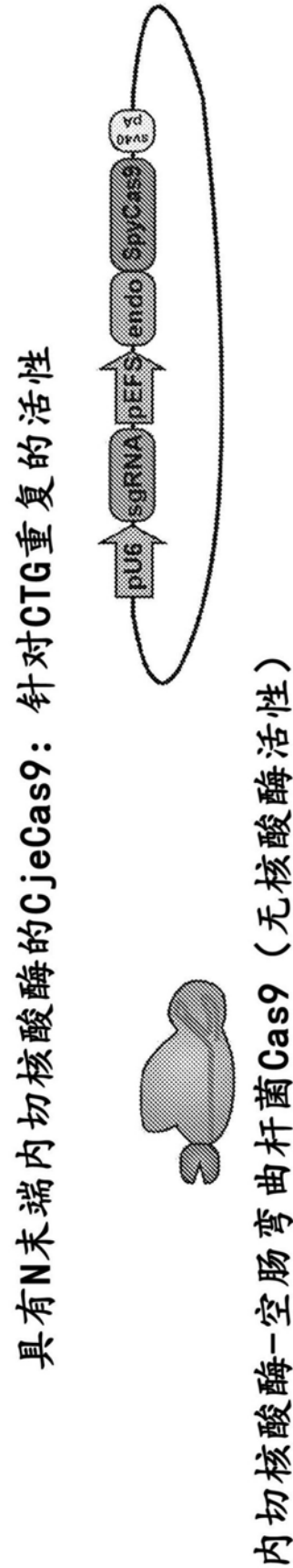


FIGURE 4A

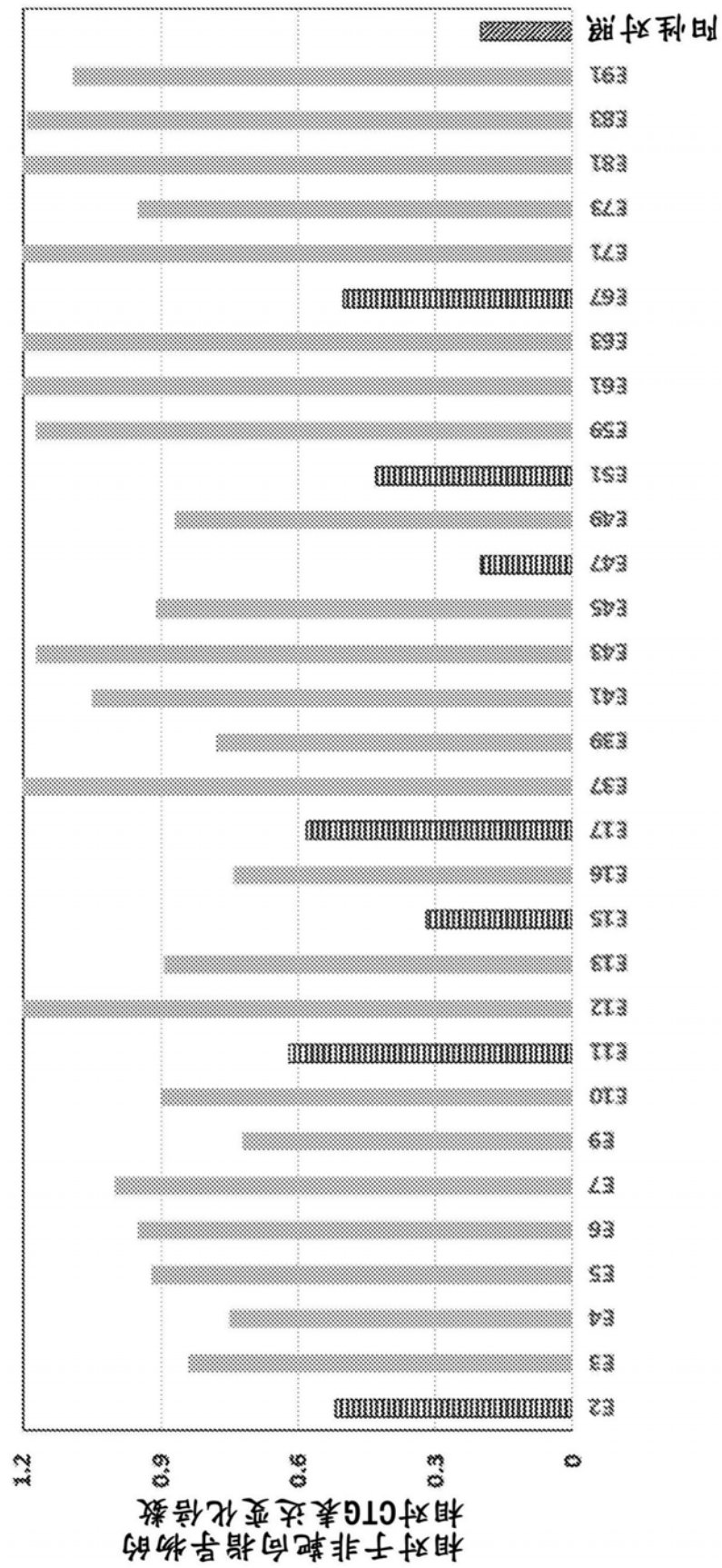


FIGURE 4B

具有N末端内切核酸酶的SpyCas9：针对萤光素酶mRNA的活性

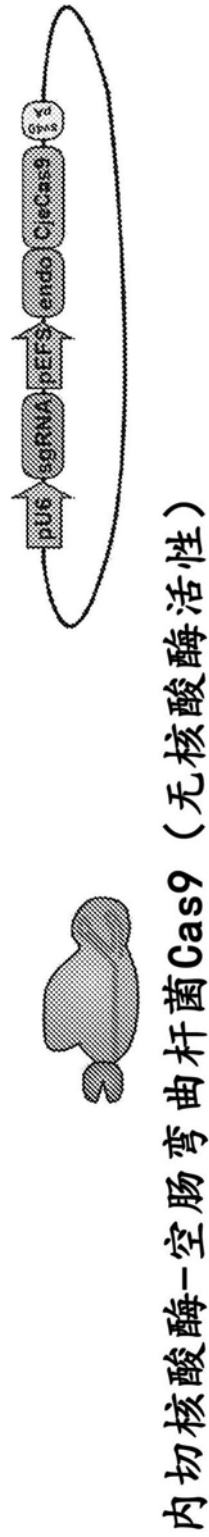


FIGURE 5A

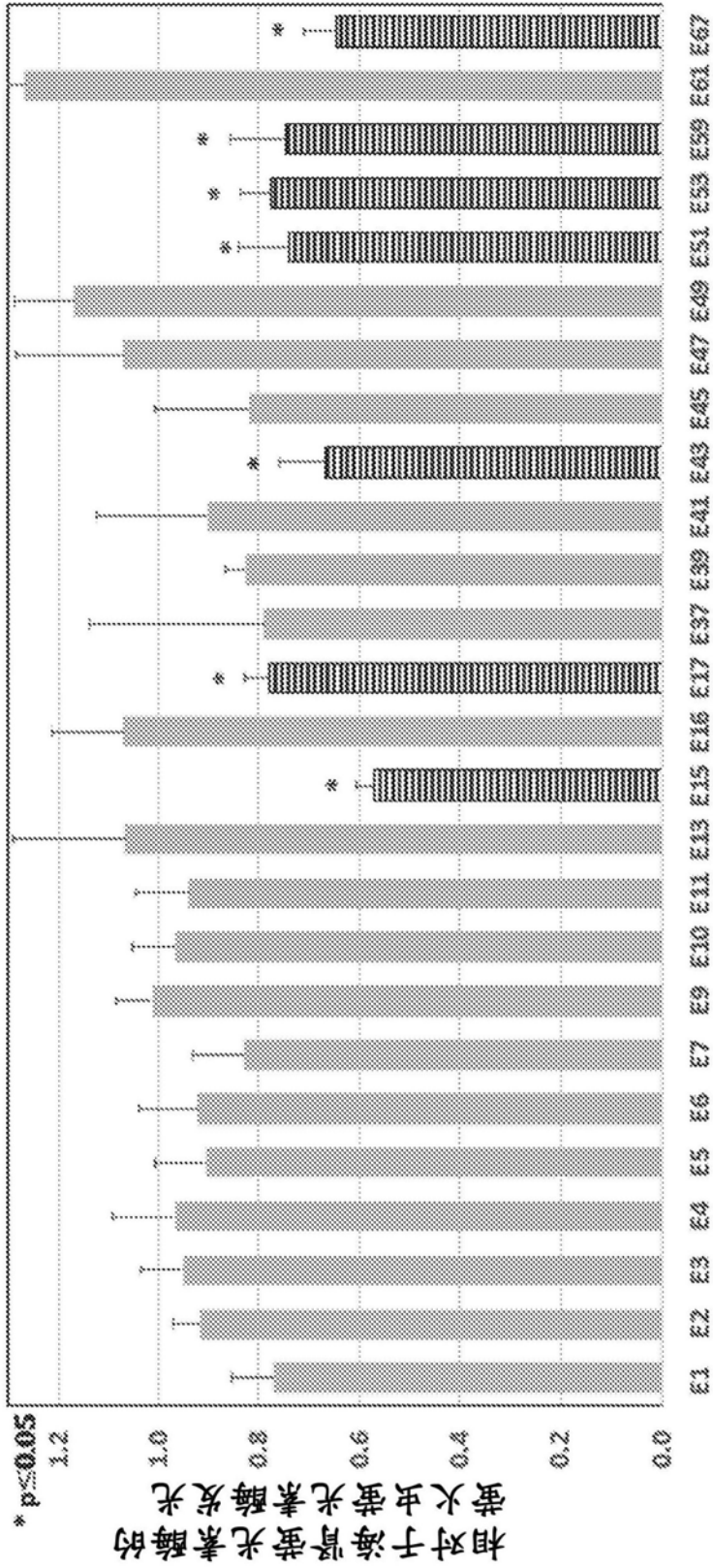
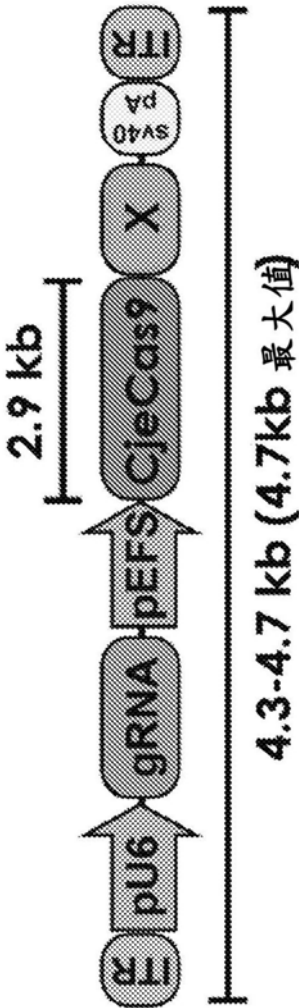


FIGURE 5B

内切核酸酶名称关键词

E1	RNA酶I	E53	EXO6
E2	RNA酶4	E55	ZC3H12D
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	OPSF4L
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6P1	E79	APEX2_1-350
E15	RNA酶L	E81	R4A1_25-136
E16	RNA酶T2	E83	RAB1
E17	ZC3H12A	E85	RNA酶K
E37	RIDA	E87	DNA2_FL
E39	RNA酶11	E89	RNA酶I(K41R)
E41	PDL6	E91	RNA酶(K41R, D121E)
E43	NTHL1	E93	RNA酶(K41R, D121E, H119N)
E45	KIA0391	E95	RNA酶(H119N)
E47	APEX1	E97	RNA酶I(R39D, N67D, N88A, G89D, R91D, H119N)
E49	Ago2	E99	RNA酶I(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
E51	ZC3H12A	E100	RNA酶I(R39D, N67D, N88A, G89D, R91D, H119N, K41R, D121E)
		E101	RNA酶I(R39D, N67D, N88A, G89D, R91D)

FIGURE 6



内切核酸酶-空肠弯曲杆菌Cas9

FIGURE 7A

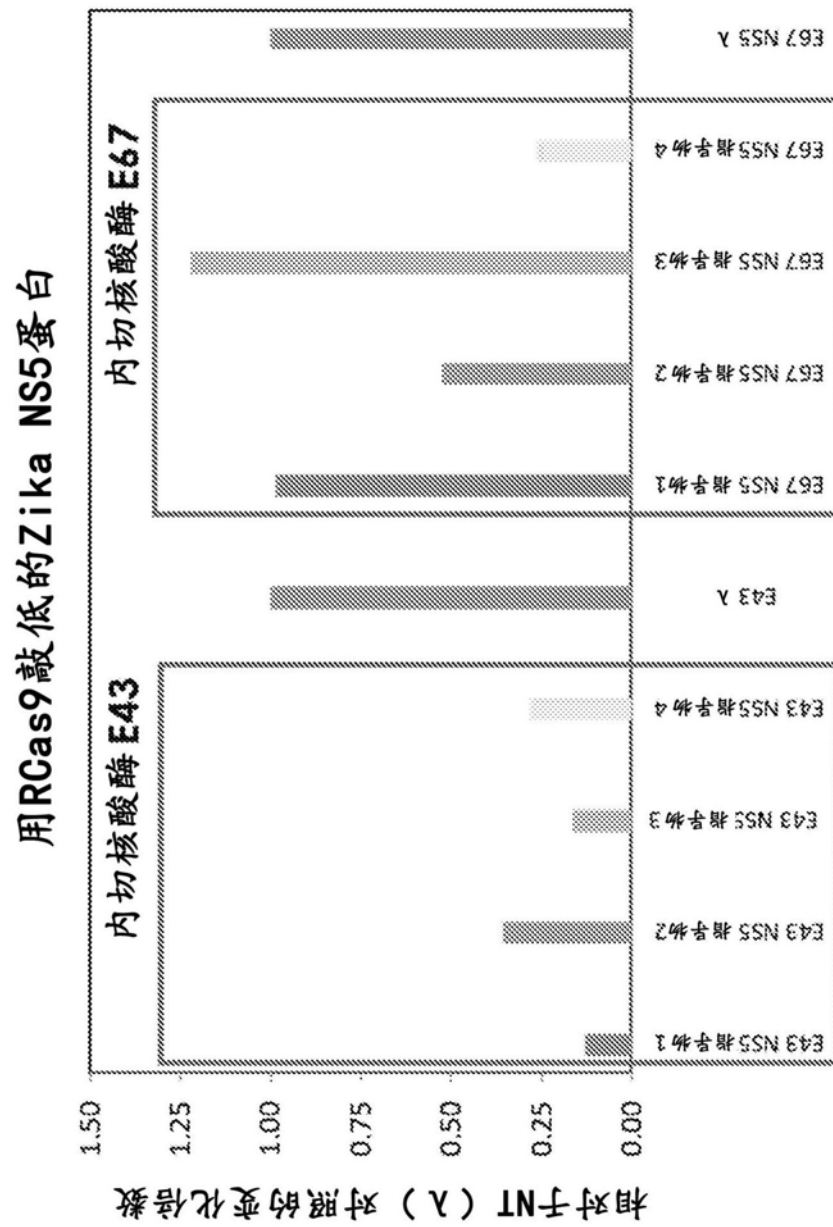


FIGURE 7B

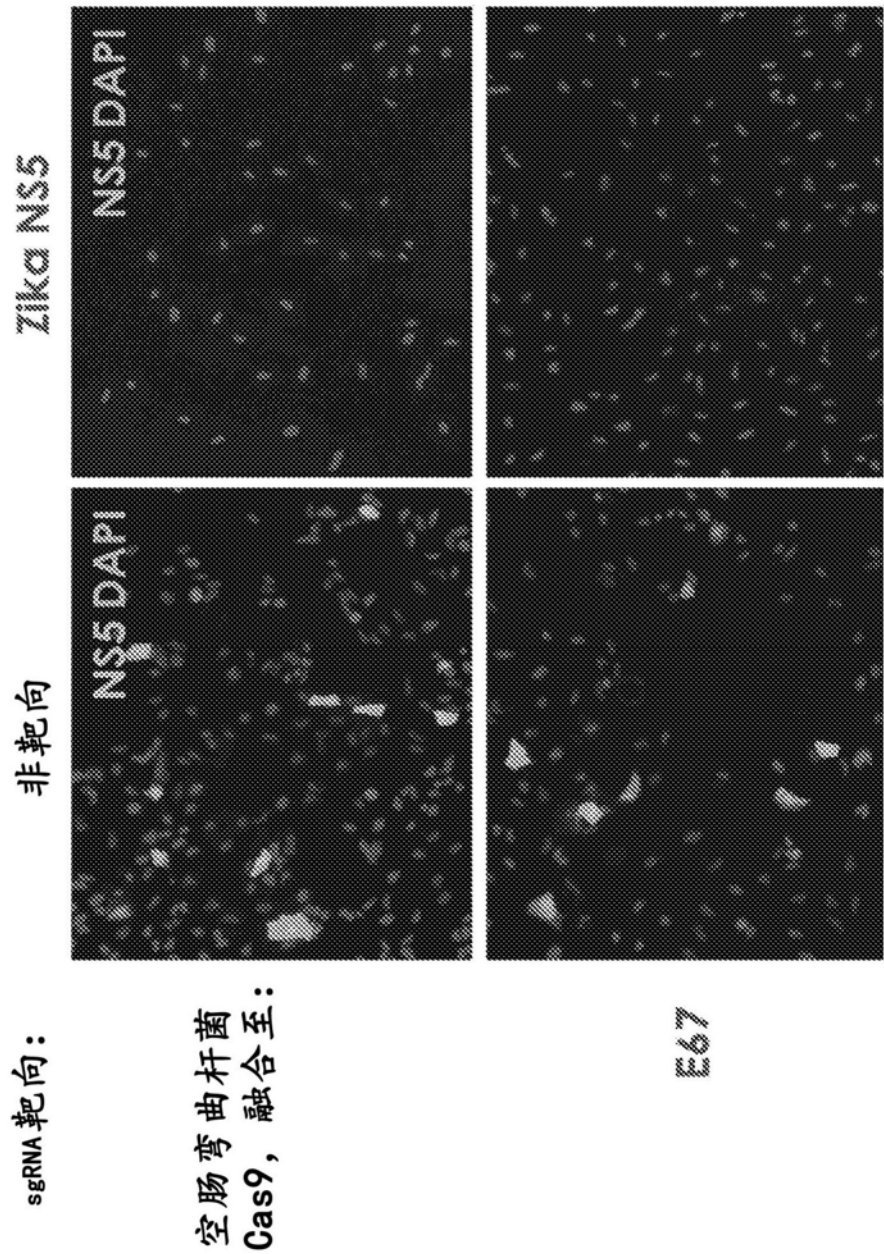


FIGURE 8A

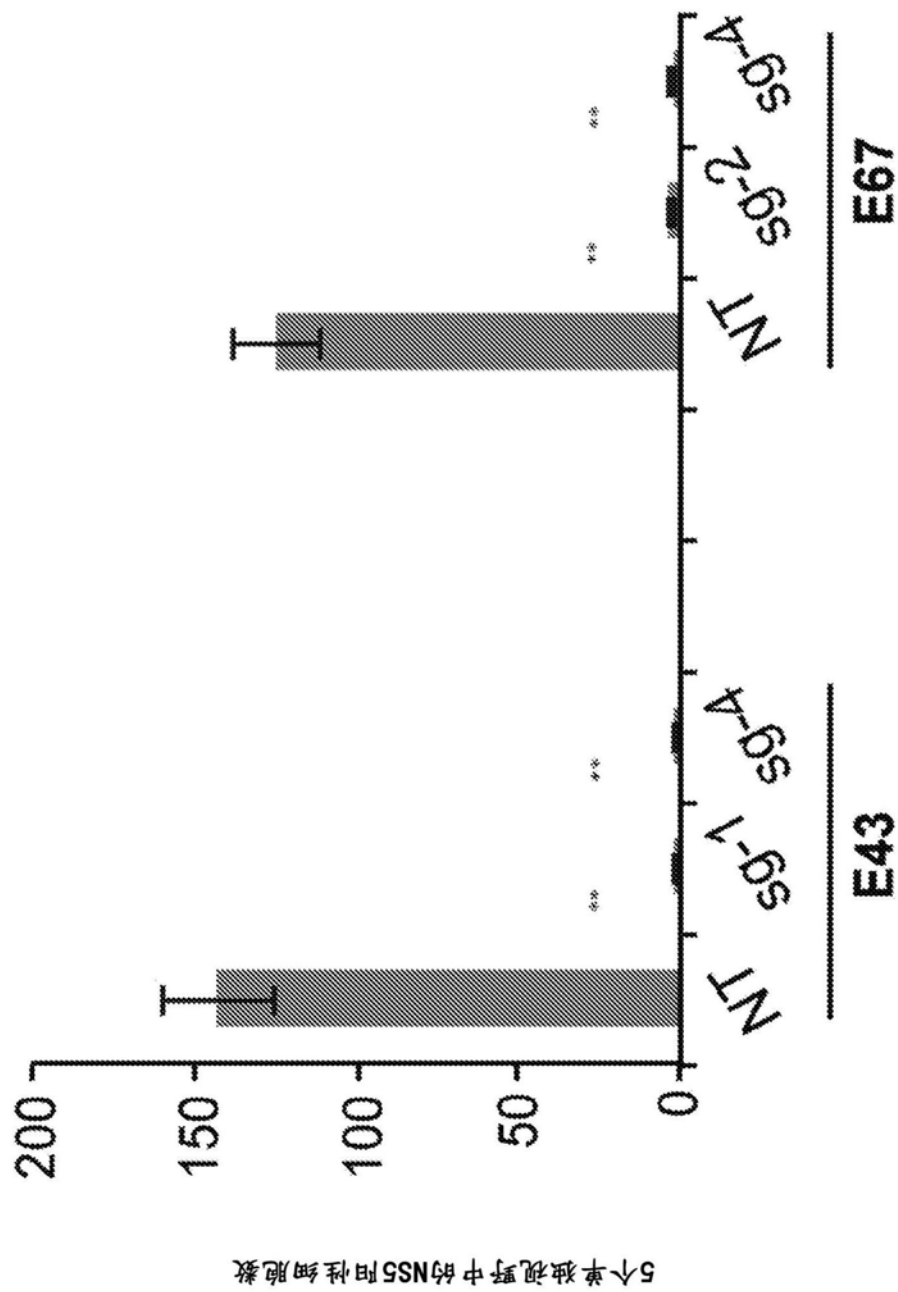


FIGURE 8B

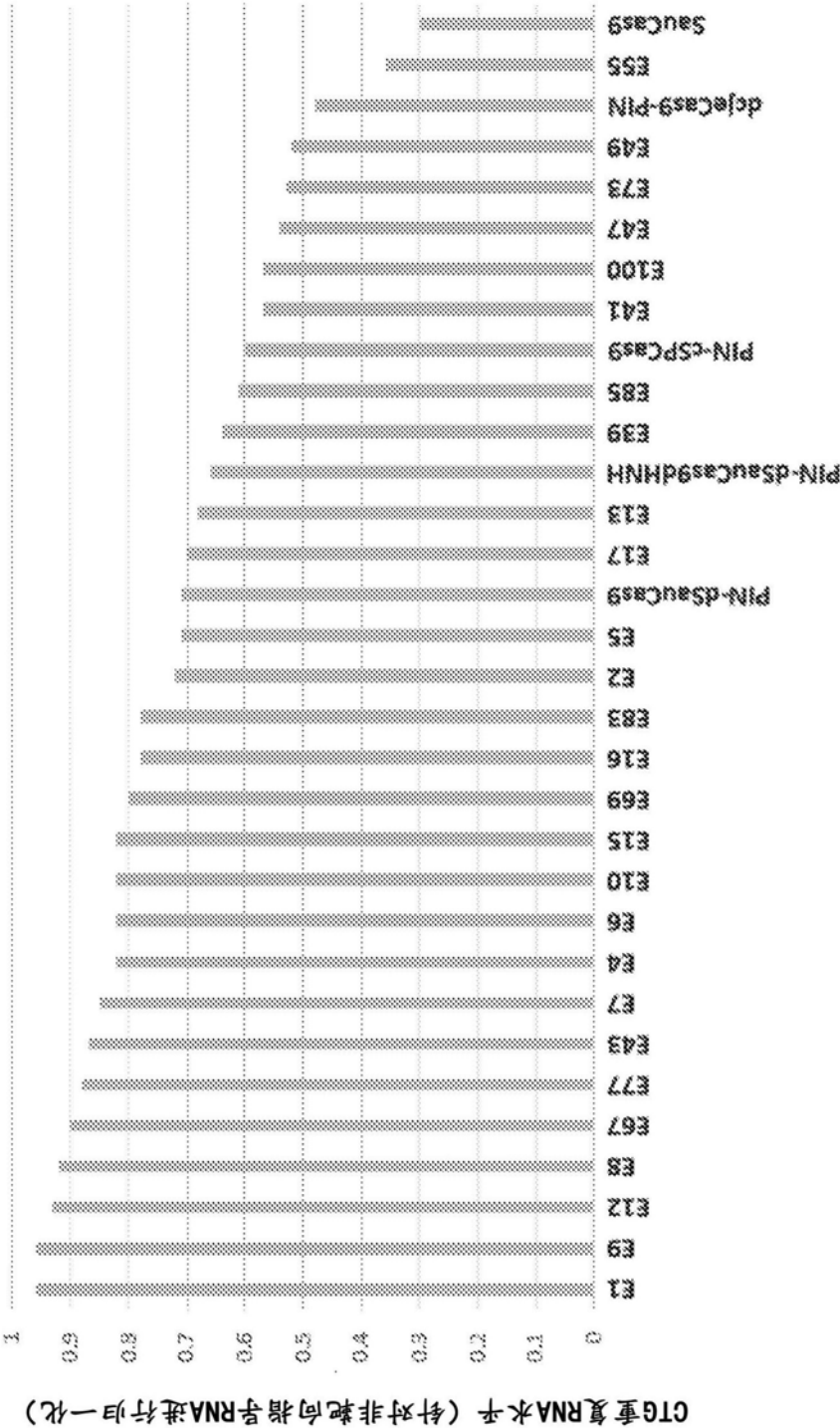


FIGURE 9