



(12) 发明专利

(10) 授权公告号 CN 113271982 B

(45) 授权公告日 2024. 07. 30

(21) 申请号 201980075173.0
(22) 申请日 2019.11.15
(65) 同一申请的已公布的文献号
申请公布号 CN 113271982 A
(43) 申请公布日 2021.08.17
(30) 优先权数据
62/768,474 2018.11.16 US
62/861,039 2019.06.13 US
62/931,925 2019.11.07 US
(85) PCT国际申请进入国家阶段日
2021.05.14
(86) PCT国际申请的申请数据
PCT/JP2019/045716 2019.11.15
(87) PCT国际申请的公布数据
W02020/101042 EN 2020.05.22
(73) 专利权人 摩大力斯医疗株式会社
地址 日本东京都
(72) 发明人 吉见英治 吉冈克郎 山形哲也
秦元宝 伊恩·罗伯特·汤普生

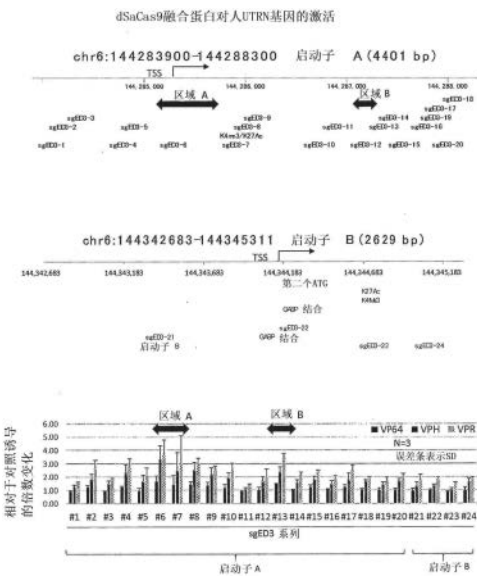
尼德西·康纳
(74) 专利代理机构 中原信达知识产权代理有限
责任公司 11219
专利代理师 金海霞 刘慧
(51) Int.Cl.
A61K 48/00 (2006.01)
C07K 14/47 (2006.01)
C12N 9/22 (2006.01)
C12N 15/113 (2006.01)
C12N 15/86 (2006.01)
(56) 对比文件
WO 2018169983 A1, 2018.09.20
Wojtal D等.Spell Checking Nature:
Versatility of CRISPR/Cas9 for Developing
Treatments for Inherited Disorders.《Am J
Hum Genet》.2015,第98卷(第01期),第90-101
页,特别是第92页右栏第24行-第93页右栏第4
行、图1.
审查员 胡雨嘉
权利要求书2页 说明书37页
序列表108页 附图13页

(54) 发明名称

通过靶向肌营养相关蛋白基因来治疗肌营
养不良的方法

(57) 摘要

预期可用于治疗或预防杜氏肌营养不良或
贝克肌营养不良的多核苷酸,所述多核苷酸包含
以下碱基序列:(a) 编码核酸酶缺陷型CRISPR效
应蛋白和转录激活子的融合蛋白的碱基序列,和
(b) 编码向导RNA的碱基序列,所述向导RNA靶向
人肌营养相关蛋白基因的表达调节区中SEQ ID
NO: 104、105、135、141、153、167或172所示区域中
长度为18至24个核苷酸的连续区域。



1. 一种多核苷酸,所述多核苷酸包含以下碱基序列:
 - (a) 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,和
 - (b) 编码向导RNA的碱基序列,所述向导RNA包含crRNA,其中编码crRNA的碱基序列如SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示,其中所述向导RNA靶向人肌营养相关蛋白基因的表达调节区,其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌(*Staphylococcus aureus*)的dCas9,并且其中所述转录激活子是VPR、miniVR或microVR。
2. 根据权利要求1所述的多核苷酸,所述多核苷酸包含至少两种不同的编码所述向导RNA的碱基序列。
3. 根据权利要求1所述的多核苷酸,所述多核苷酸进一步包含用于所述编码向导RNA的碱基序列的启动子序列和用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子序列。
4. 根据权利要求3所述的多核苷酸,其中用于所述编码向导RNA的碱基序列的启动子选自U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子和H1启动子组成的组。
5. 根据权利要求3所述的多核苷酸,其中用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子选自EFS启动子、EF-1 α 启动子、CMV启动子、CK8启动子、MHC启动子、Des启动子、CAG启动子和MYOD启动子组成的组。
6. 根据权利要求3所述的多核苷酸,其中所述编码crRNA的碱基序列如SEQ ID NO:45、46或59所示,所述转录激活子是miniVR或microVR,所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9,用于所述编码向导RNA的碱基序列的启动子是U6启动子,且用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子是CK8启动子。
7. 根据权利要求6所述的多核苷酸,其中所述编码crRNA的碱基序列如SEQ ID NO:59所示。
8. 一种载体,所述载体包含根据权利要求1至7中任一项所述的多核苷酸。
9. 根据权利要求8所述的载体,其中所述载体是质粒载体或病毒载体。
10. 根据权利要求9所述的载体,其中所述病毒载体选自AAV载体、腺病毒载体和慢病毒载体组成的组。
11. 根据权利要求10所述的载体,其中所述AAV载体选自AAV1、AAV2、AAV6、AAV7、AAV8、AAV9、AAV₅₈₇MTP、AAV₅₈₈MTP、AAV-B1、AAVM41、AAVrh74、AAVS1_P1和AAVS10_P1组成的组。
12. 一种药物组合物,所述药物组合物包含根据权利要求1至7中任一项所述的多核苷酸或根据权利要求8至11中任一项所述的载体。
13. 根据权利要求1至7中任一项所述的多核苷酸或根据权利要求8至11任一项所述的载体在制备用于治疗或预防杜氏肌营养不良或贝克肌营养不良的药物组合物中的用途。

14. 一种核糖核蛋白,所述核糖核蛋白包含以下:

(c) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,和

(d) 向导RNA,所述向导RNA包含crRNA,其中所述crRNA的碱基序列如SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示,

其中所述向导RNA靶向人肌营养相关蛋白基因的表达调节区,

其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌 (*Staphylococcus aureus*) 的dCas9,并且

其中所述转录激活子是VPR、miniVR或microVR。

15. 一种用于激活人肌营养相关蛋白基因的表达的组合物或试剂盒,所述组合物或试剂盒包含以下:

(e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

(f) 向导RNA,所述向导RNA包含crRNA,其中所述crRNA的碱基序列如SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示,

其中所述向导RNA靶向人肌营养相关蛋白基因的表达调节区,

其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌 (*Staphylococcus aureus*) 的dCas9,并且

其中所述转录激活子是VPR、miniVR或microVR。

通过靶向肌营养相关蛋白基因来治疗肌营养不良的方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2018年11月16日提交的美国临时专利申请第62/768,474号、2019年6月13日提交的美国临时专利申请第62/861,039号以及2019年11月7日提交的美国临时专利申请第62/931,925号的优先权,所述申请各自的内容通过引用全部并入本文。

技术领域

[0003] 本发明涉及通过靶向人肌营养相关蛋白 (Utrophin, UTRN) 基因来治疗肌营养不良的方法等。更特别地,本发明涉及通过使用靶向人UTRN基因的特定序列的向导RNA以及转录激活子和CRISPR效应蛋白的融合蛋白激活人UTRN基因的表达来治疗或预防肌营养不良的方法和药剂等。

背景技术

[0004] 肌营养不良是与进行性肌萎缩和肌无力相关的遗传性疾病的总称。在肌营养不良中,由X染色体上的肌营养不良蛋白基因突变引起的肌营养不良包括杜氏肌营养不良 (DUCHENNE muscular dystrophy, DMD) 及其轻度型、贝克肌营养不良 (BECKER muscular dystrophy, BMD)。

[0005] DMD是最常见的遗传性进行性肌肉疾病,每3,500名新出生的男性中就有1名患上DMD。其临床症状包括从约2至5岁开始的肌无力,此后肌无力的进一步发展,到约10岁前的步行不能,以及在二十多岁时因心力衰竭或呼吸衰竭而死亡 (参见WO 2009/044383,通过引用全部并入本文)。

[0006] 已知DMD是由肌营养不良蛋白基因的突变引起的。肌营养不良蛋白基因存在于X染色体上,是一个由约220万个DNA碱基组成的巨大基因。它从DNA转录成mRNA前体,通过剪接进一步去除内含子,并产生了由79个外显子组成的mRNA (约14kb)。该mRNA被翻译成3685个氨基酸以产生肌营养不良蛋白。肌营养不良蛋白参与维持肌肉细胞的膜稳定性。在DMD患者中,由于在肌营养不良蛋白基因中发生突变,因此肌营养不良蛋白几乎不表达并且不能维持肌肉细胞的结构,从而导致肌无力。

[0007] BMD也是由肌营养不良蛋白基因突变引起的;但是,与DMD相比,其症状通常较轻。DMD和BMD的临床症状之间的差异是基于DMD中功能性的肌营养不良蛋白几乎不表达,而BMD中却产生了不完整但功能性的肌营养不良蛋白。

[0008] 即使到现在,还没有有效的药物作为针对肌营养不良的病因治疗,只是进行了诸如施用类固醇的对症治疗。已经提出了多种治疗策略来治疗DMD和BMD,且作为策略之一的基因治疗方法已经引起关注。基因治疗的目的是通过向具有突变的肌肉细胞补充正常肌营养不良蛋白基因来实现正常肌营养不良蛋白的表达。然而,全长的肌营养不良蛋白cDNA相对较大,长度为约14kb;因此,对于某些载体,例如腺相关病毒 (AAV) 载体,可以被封包的DNA的尺寸限制可能是一个问题。作为这个问题的一种解决方案,已经提出了使用具有最小功能域的截短型肌营养不良蛋白基因 (小型/微型肌营养不良蛋白基因) 的方法 (参见

Sakamoto M.等人,《生物化学与生物物理学研究通讯(Biochem Biophys Res Commun)》,2002年5月17日;293(4):1265-72,通过引用全部并入本文)。考虑到向缺乏肌营养不良蛋白的DMD患者中引入肌营养不良蛋白可能会引起免疫反应,因此还报道了使用肌营养相关蛋白(在本说明书中有时也称为“UTRN”等)来降低这种免疫反应的手段(参见Gilbert R.等人,《人基因治疗(Hum Gene Ther)》,1999年5月20日;10(8):1299-310,通过引用全部并入本文)。肌营养相关蛋白是一种与肌营养不良蛋白高度同源的细胞骨架蛋白,存在于正常和DMD肌肉中,但含量很低。与肌营养不良蛋白一样,肌营养相关蛋白cDNA非常大(超过10kb)。已知肌营养相关蛋白能够补偿肌营养不良蛋白的肌肉细胞膜稳定功能(参见Gilbert R.等人,《人基因治疗》,1999年5月20日;10(8):1299-310和Liao H.等人,《细胞(Cell)》,2017年12月14日;171(7):1495-507,通过引用全部并入本文)。

[0009] 作为靶向肌营养相关蛋白的基因治疗,例如,W02015/018503公开了一项涉及用于在骨骼或心肌组织中表达基因的重组腺相关病毒(AAV)载体的发明,所述载体包含肌肉特异性启动子和编码融合蛋白的基因,其中所述融合蛋白包含:

[0010] a) 转录激活元件和

[0011] b) DNA结合元件,

[0012] 其中所述融合蛋白当在所述骨骼或心肌组织中表达时能够增加肌营养相关蛋白的表达(参见W02015/018503,通过引用全部并入本文)。在本发明中,锌指蛋白用作DNA结合元件。

[0013] 另一方面,近年来已经开发出一种使用核酸酶活性失活的Cas9(dCas9)和转录激活结构域或转录抑制结构域的组合物,其中通过使用向导RNA而不切割靶基因的DNA序列将所述蛋白质靶向靶基因来控制靶基因的表达(W02013/176772,通过引用全部并入本文)。预期其可应用于临床(参见Dominguez A.等人,《自然评论-分子细胞生物学(Nat Rev Mol Cell Biol)》,2016年1月;17(1):5-15,通过引用全部并入本文)。但是,存在一个问题,即编码dCas9、向导RNA和共转录激活子的复合物的序列超出了最常见的病毒载体(例如AAV)的容量,最常见的病毒载体代表了最有希望的体内基因递送方法(参见Liao H.等人,《细胞》,2017年12月14日;171(7):1495-507,通过引用全部并入本文)。

[0014] 2017年,报道了(a)通过向引入有Cas9基因的DMD模型小鼠(mdx小鼠)施用携带靶向小鼠UTRN并抑制Cas9的DNA切割能力的向导RNA(dgUtrn)和转录激活结构域的AAV,UTRN的表达水平增加,抓取力也得到了提高,以及(b)通过向mdx小鼠共同注射携带Cas9的AAV和携带上述dgUtrn和转录激活结构域的AAV,抓取力得到了提高(参见Liao H.等人,《细胞》,2017年12月14日;171(7):1495-507,通过引用全部并入本文)。

发明内容

[0015] 因此,本发明的一个目的是提供一种治疗肌营养不良(特别是DMD和BMD)的新方法。

[0016] 这一目的以及其它目的,在以下详细描述中将变得显而易见,这些目的已经通过本发明人的以下发现而实现,即,可以通过使用靶向人UTRN基因的特定序列的向导RNA以及转录激活子和核酸酶缺陷型CRISPR效应蛋白的融合蛋白来强烈激活人UTRN基因(基因ID: 7402)的表达。另外,本发明人发现,可以通过携带编码所述融合蛋白的碱基序列和编码所

述向导RNA的碱基序列的单个AAV载体,使用小型核酸酶缺陷型CRISPR效应蛋白和小型转录激活子来强烈激活人UTRN基因的表达。

[0017] 因此,本发明提供:

[0018] (1) 一种多核苷酸,所述多核苷酸包含以下碱基序列:

[0019] (a) 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,和

[0020] (b) 编码向导RNA的碱基序列,所述向导RNA靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域。

[0021] (2) 根据(1)所述的多核苷酸,其中所述编码向导RNA的碱基序列包含SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示的碱基序列,或SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0022] (3) 根据(1)或(2)所述的多核苷酸,所述多核苷酸包含至少两种不同的编码所述向导RNA的碱基序列。

[0023] (4) 根据(1)至(3)中任一项所述的多核苷酸,其中所述转录激活子是包含RTA的转录激活结构域和VP64的肽。

[0024] (5) 根据(4)所述的多核苷酸,其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列。

[0025] (6) 根据(1)至(5)中任一项所述的多核苷酸,其中所述核酸酶缺陷型CRISPR效应蛋白是dCas9。

[0026] (7) 根据(6)所述的多核苷酸,其中所述dCas9来源于金黄色葡萄球菌(*Staphylococcus aureus*)。

[0027] (8) 根据(1)至(7)中任一项所述的多核苷酸,所述多核苷酸进一步包含用于所述编码向导RNA的碱基序列的启动子序列和/或用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子序列。

[0028] (9) 根据(8)所述的多核苷酸,其中所述用于编码向导RNA的碱基序列的启动子序列选自U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子和H1启动子。

[0029] (10) 根据(8)或(9)所述的多核苷酸,其中所述用于编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子序列选自EFS启动子、EF-1 α 启动子、CMV启动子、CK8启动子、MHC启动子、Des启动子、CAG启动子和MYOD启动子。

[0030] (11) 根据(8)至(10)中任一项所述的多核苷酸,

[0031] 其中所述编码向导RNA的碱基序列包含SEQ ID NO:45、46或59所示的碱基序列,或SEQ ID NO:45、46或59所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列,

[0032] 所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列,

[0033] 所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9,

[0034] 所述用于编码向导RNA的碱基序列的启动子序列是U6启动子,且

[0035] 所述用于编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序

列的启动子序列是CK8启动子。

[0036] (12) 根据(11)所述的多核苷酸,

[0037] 其中所述编码向导RNA的碱基序列包含SEQ ID NO:59所示的碱基序列,或SEQ ID NO:59所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0038] (13) 一种载体,所述载体包含根据(1)至(12)中任一项所述的多核苷酸。

[0039] (14) 根据(13)所述的载体,其中所述载体是质粒载体或病毒载体。

[0040] (15) 根据(14)所述的载体,其中所述病毒载体选自腺相关病毒(AAV)载体、腺病毒载体和慢病毒载体。

[0041] (16) 根据(15)所述的载体,其中所述AAV载体选自AAV1、AAV2、AAV6、AAV7、AAV8、AAV9、AAV₅₈₇MTP、AAV₅₈₈MTP、AAV-B1、AAVM41、AAVrh74、AAVS1_P1和AAVS10_P1。

[0042] (17) 根据(16)所述的载体,其中所述AAV载体是AAV9。

[0043] (18) 一种药物组合物,所述药物组合物包含根据(1)至(12)中任一项所述的多核苷酸或根据(13)至(17)中任一项所述的载体。

[0044] (19) 根据(18)所述的药物组合物,所述药物组合物用于治疗或预防杜氏肌营养不良或贝克肌营养不良。

[0045] (20) 一种治疗或预防杜氏肌营养不良或贝克肌营养不良的方法,所述方法包括向有此需要的对象施用根据(1)至(12)中任一项所述的多核苷酸或根据(13)至(17)中任一项所述的载体。

[0046] (21) 根据(1)至(12)中任一项所述的多核苷酸或根据(13)至(17)中任一项所述的载体用于治疗或预防杜氏肌营养不良或贝克肌营养不良的用途。

[0047] (22) 根据(1)至(12)中任一项所述的多核苷酸或根据(13)至(17)中任一项所述的载体在制备用于治疗或预防杜氏肌营养不良或贝克肌营养不良的药物组合物中的用途。

[0048] (23) 一种核糖核蛋白,所述核糖核蛋白包含以下:

[0049] (c) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,和

[0050] (d) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA。

[0051] (24) 根据(23)所述的核糖核蛋白,其中所述向导RNA包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的碱基序列,或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0052] (25) 根据(23)或(24)所述的核糖核蛋白,其中所述转录激活子是包含RTA的转录激活结构域和VP64的肽。

[0053] (26) 根据(25)所述的核糖核蛋白,其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列。

[0054] (27) 根据(23)至(26)中任一项所述的核糖核蛋白,其中所述核酸酶缺陷型CRISPR效应蛋白是dCas9。

[0055] (28) 根据(27)所述的核糖核蛋白,其中所述dCas9来源于金黄色葡萄球菌。

[0056] (29) 根据(23)至(28)中任一项所述的核糖核蛋白,

[0057] 其中所述向导RNA包含SEQ ID NO:194、195或197所示的碱基序列,或SEQ ID NO:

194、195或197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列，

[0058] 其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列，或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列，且

[0059] 其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9。

[0060] (30) 根据(29)所述的核糖核蛋白，其中所述向导RNA包含SEQ ID NO:197所示的碱基序列或SEQ ID NO:197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0061] (31) 一种组合物或试剂盒，所述组合物或试剂盒包含以下用于激活人肌营养相关蛋白基因的表达：

[0062] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白，或编码所述融合蛋白的多核苷酸，和

[0063] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA，或编码所述向导RNA的多核苷酸。

[0064] (32) 根据(31)所述的组合物或试剂盒，其中所述向导RNA包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的碱基序列，或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0065] (33) 根据(31)或(32)所述的组合物或试剂盒，包含至少两种不同的向导RNA。

[0066] (34) 根据(31)至(33)中任一项所述的组合物或试剂盒，其中所述转录激活子是包含RTA的转录激活结构域和VP64的肽。

[0067] (35) 根据(34)所述的组合物或试剂盒，其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列，或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列。

[0068] (36) 根据(31)至(35)中任一项所述的组合物或试剂盒，其中所述核酸酶缺陷型CRISPR效应蛋白是dCas9。

[0069] (37) 根据(36)所述的组合物或试剂盒，其中所述dCas9来源于金黄色葡萄球菌。

[0070] (38) 根据(31)至(37)中任一项所述的组合物或试剂盒，

[0071] 其中所述组合物或试剂盒包含编码所述融合蛋白的多核苷酸和编码所述向导RNA的多核苷酸，且

[0072] 其中所述编码融合蛋白的多核苷酸进一步包含用于所述融合蛋白的启动子序列，且/或所述编码向导RNA的多核苷酸进一步包含用于所述向导RNA的启动子序列。

[0073] (39) 根据(38)所述的组合物或试剂盒，其中所述用于向导RNA的启动子序列选自U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子和H1启动子。

[0074] (40) 根据(38)或(39)所述的组合物或试剂盒，其中所述用于融合蛋白的启动子序列选自EFS启动子、EF-1 α 启动子、CMV启动子、CK8启动子、MHC启动子、Des启动子、CAG启动子和MYOD启动子。

[0075] (41) 根据(38)至(40)中任一项所述的组合物或试剂盒，其中所述向导RNA包含SEQ ID NO:194、195或197所示的碱基序列，或SEQ ID NO:194、195或197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列，

[0076] 其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列,

[0077] 其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9,

[0078] 其中所述用于向导RNA的启动子序列是U6启动子,且

[0079] 其中所述用于融合蛋白的启动子序列是CK8启动子。

[0080] (42) 根据(41)所述的组合物或试剂盒,其中所述向导RNA包含SEQ ID NO:197所示的碱基序列,或SEQ ID NO:197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0081] (43) 一种治疗或预防DUCCHENNE肌营养不良或贝克肌营养不良的方法,所述方法包括施用以下(e)和(f):

[0082] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

[0083] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA,或编码所述向导RNA的多核苷酸。

[0084] (44) 根据(43)所述的方法,其中所述向导RNA包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的碱基序列,或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0085] (45) 根据(43)或(44)所述的方法,包含至少两种不同的向导RNA。

[0086] (46) 根据(43)至(45)中任一项所述的方法,其中所述转录激活子是包含RTA的转录激活结构域和VP64的肽。

[0087] (47) 根据(46)所述的方法,其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列。

[0088] (48) 根据(43)至(47)中任一项所述的方法,其中所述核酸酶缺陷型CRISPR效应蛋白是dCas9。

[0089] (49) 根据(48)所述的方法,其中所述dCas9来源于金黄色葡萄球菌。

[0090] (50) 根据(43)至(49)中任一项所述的方法,

[0091] 其中所述方法包括施用编码所述融合蛋白的多核苷酸和编码所述向导RNA的多核苷酸,且

[0092] 其中所述编码融合蛋白的多核苷酸进一步包含用于所述融合蛋白的启动子序列,且/或所述编码向导RNA的多核苷酸进一步包含用于所述向导RNA的启动子序列。

[0093] (51) 根据(50)所述的方法,其中所述用于向导RNA的启动子序列选自U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子和H1启动子。

[0094] (52) 根据(50)或(51)所述的方法,其中所述用于融合蛋白的启动子序列选自EFS启动子、EF-1 α 启动子、CMV启动子、CK8启动子、MHC启动子、Des启动子、CAG启动子和MYOD启动子。

[0095] (53) 根据(50)至(52)中任一项所述的方法,其中所述向导RNA包含SEQ ID NO:194、195或197所示的碱基序列,或SEQ ID NO:194、195或197所示但其中缺失、替换、插入

和/或添加了1至3个碱基的碱基序列,

[0096] 其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列,

[0097] 其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9,

[0098] 其中所述用于向导RNA的启动子序列是U6启动子,且

[0099] 其中所述用于融合蛋白的启动子序列是CK8启动子。

[0100] (54) 根据(53)所述的方法,其中所述向导RNA包含SEQ ID NO:197所示的碱基序列,或SEQ ID NO:197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0101] (55) 以下(e)及(f)在制备用于治疗或预防杜氏肌营养不良或贝克肌营养不良的药物组合物中的用途:

[0102] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

[0103] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA,或编码所述向导RNA的多核苷酸。

[0104] (56) 根据(55)所述的用途,其中所述向导RNA包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的碱基序列,或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0105] (57) 根据(55)或(56)所述的用途,包含至少两种不同的向导RNA。

[0106] (58) 根据(55)至(57)中任一项所述的用途,其中所述转录激活子是包含RTA的转录激活结构域和VP64的肽。

[0107] (59) 根据(58)所述的用途,其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列。

[0108] (60) 根据(55)至(59)中任一项所述的用途,其中所述核酸酶缺陷型CRISPR效应蛋白是dCas9。

[0109] (61) 根据(60)所述的用途,其中所述dCas9来源于金黄色葡萄球菌。

[0110] (62) 根据(55)至(61)中任一项所述的用途,

[0111] 其中所述用途包括编码所述融合蛋白的多核苷酸的用途以及编码所述向导RNA的多核苷酸的用途,且

[0112] 其中所述编码融合蛋白的多核苷酸进一步包含用于所述融合蛋白的启动子序列,且/或所述编码向导RNA的多核苷酸进一步包含用于所述向导RNA的启动子序列。

[0113] (63) 根据(62)所述的用途,其中所述用于向导RNA的启动子序列选自U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子和H1启动子。

[0114] (64) 根据(62)或(63)所述的用途,其中所述用于融合蛋白的启动子序列选自EFS启动子、EF-1 α 启动子、CMV启动子、CK8启动子、MHC启动子、Des启动子、CAG启动子和MYOD启动子。

[0115] (65) 根据(62)至(64)中任一项所述的用途,其中所述向导RNA包含SEQ ID NO:194、195或197所示的碱基序列,或SEQ ID NO:194、195或197所示但其中缺失、替换、插入

和/或添加了1至3个碱基的碱基序列,

[0116] 其中所述转录激活子包含SEQ ID NO:117所示的氨基酸序列,或与SEQ ID NO:117所示的氨基酸序列具有至少90%同一性的氨基酸序列,

[0117] 其中所述核酸酶缺陷型CRISPR效应蛋白是来源于金黄色葡萄球菌的dCas9,

[0118] 其中所述用于向导RNA的启动子序列是U6启动子,且

[0119] 其中所述用于融合蛋白的启动子序列是CK8启动子。

[0120] (66)根据(65)所述的用途,其中所述向导RNA包含SEQ ID NO:197所示的碱基序列,或SEQ ID NO:197所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0121] 发明效果

[0122] 根据本发明,可以激活人肌营养相关蛋白基因的表达,因此,预期本发明能够治疗DMD和BMD。

附图说明

[0123] 当结合附图考虑以下详细描述时,将容易更完整地理解本发明及其许多附带优点,因为通过以下详细描述将使本发明得到更好的理解,其中:

[0124] 在图1中,上图示出了人UTRN基因中启动子A的区域,中图示出了启动子B的区域,且示出了相应区域中确定的24种靶向序列的位置(向导编号sgED3-1至sgED3-24(SEQ ID NO:129至152))。在图1中,下图示出了HEK293FT细胞中通过使用包含由SEQ ID NO:129至152所示的靶向序列编码的crRNA的sgRNA和3种不同的dSaCas9-转录激活子融合蛋白(dSaCas9-VP64(SEQ ID NO:188)、dSaCas9-VPH(SEQ ID NO:189)、dSaCas9-VPR(SEQ ID NO:190))的组合对人UTRN基因表达的激活(N=3,误差条表示标准偏差)。当分别使用特异性结合包含向导编号sgED3-6和sgED3-7(SEQ ID NO:134和135)的区域(区域A)和包含向导编号sgED3-13(SEQ ID NO:141)的另一个区域(区域B)的sgRNA时,与使用对照sgRNA的情况相比,人UTRN基因的表达被强烈激活。在3种dSaCas9-转录激活子融合蛋白中,使用dSaCas9-VPR融合蛋白时的激活效果最强。

[0125] 在图2中,上图示出了在人UTRN基因的启动子A的区域中确定的靶向序列(向导编号sgED3-1至sgED3-20和sgED3-25至sgED3-48(SEQ ID NO:129至148和153至176))的位置。在图2中,下图示出了HEK293FT细胞中通过使用包含由靶向序列向导编号sgED3-6、sgED3-13、sgED3-25至sgED3-48(SEQ ID NO:134、141、153至176)编码的crRNA的sgRNA和dSaCas9-VPR的组合对人UTRN基因表达的激活(N=3,误差条表示标准偏差)。当分别使用特异性结合包含靶向序列向导编号sgED3-6、sgED3-13、sgED3-25至sgED3-32、sgED3-39、sgED3-40、sgED3-44(SEQ ID NO:134、141、153至160、167、168和172)的区域的sgRNA时,与使用对照sgRNA的情况相比,人UTRN基因表达被激活了不少于两倍。

[0126] 图3示出了通过使用质粒载体对各sgRNA的功能的验证结果(N=1)。制备表达包含由靶向序列向导编号sgED3-6、sgED3-13、sgED3-25、sgED3-27、sgED3-30、sgED3-31、sgED3-39、sgED3-40或sgED3-44(SEQ ID NO:134、141、153、155、158、159、167、168或172)编码的crRNA的sgRNA的pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒,并将所述质粒转染到HEK293FT细胞中,并验证其功能。与对照sgRNA相比,当使用包含由靶向序列向导编号sgED3-6、sgED3-13、sgED3-25、sgED3-27、sgED3-30、sgED3-31、sgED3-39、sgED3-40或

sgED3-44 (SEQ ID NO:134、141、153、155、158、159、167、168或172) 编码的crRNA的sgRNA时, 观察到人UTRN基因表达的激活。

[0127] 图4示出了通过使用AAV载体对各sgRNA的功能的验证结果 (N=1)。使用表达包含由靶向序列向导编号sgED3-6、sgED3-30或sgED3-31 (SEQ ID NO:134、158或159) 编码的crRNA的sgRNA的pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒产生AAV2, 将其转导至HEK293FT细胞中。与对照sgRNA相比, 在包含由靶向序列向导编号sgED3-6、sgED3-30或sgED3-31 (SEQ ID NO:134、158或159) 编码的crRNA的所有sgRNA中, 都观察到了人UTRN基因的激活。

[0128] 图5示出了pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒的构建体。

[0129] 在图6中, 图A示出了人骨骼肌细胞中基因组的H3K4me3和H3K27Ac图谱, 以及人UTRN基因的推定的增强子区域E1、E2和E3以及推定的启动子区域P1和P2。图B至F示出了各向导编号所示的靶向序列 (SEQ ID NO:4至103所示的序列) 的位置。

[0130] 图7示出了HSMM细胞中通过使用包含由各向导编号所示的靶向序列 (SEQ ID NO:4至103所示的序列) 编码的crRNA的srRNA和dSaCas9-miniVR对人UTRN基因表达的激活的评估结果 (N=2)。

[0131] 在图8中, 上图示出了5种靶向序列向导编号145、146、205、208、210 (SEQ ID NO:45、46、58、59和60) 各自与食蟹猴 (*Macaca fascicularis*) 的同源性以及所定位的区域。下图示出了5种靶向序列的组合, 它们与食蟹猴的同源性和所定位的区域。

[0132] 图9示出了在5种不同的HSMM细胞中通过使用包含由靶向序列向导编号145、146、205、208、210 (SEQ ID NO:45、46、58、59和60) 编码的crRNA的sgRNA (分别或其组合) 和dSaCas9-miniVR, 对人UTRN基因表达的激活 (N=2)。

[0133] 图10示出了在pED260、pED261或pED263质粒骨架中包含由靶向序列向导编号145、146或208编码的crRNA的sgRNA上调了UTRN。相对mRNA表达是分别从瞬时表达pED260、pED261或pED263骨架中向导编号145、146或208的HEK293FT细胞确定的。数据表示为3次重复实验的平均值+标准偏差 (N=3, 误差条表示标准偏差)。

[0134] 在图11中, 左图示出了施加了各AAV9样品和标记的SDS-PAGE的泳道布局, 右图示出了SDS-PAGE的图像。泳道11旁的值表示分子量 (kDa)。从各AAV样品中检测到三种衣壳蛋白 (VP1、VP2和VP3, 分别为87kDa、72kDa和62kDa)。这些结果表明, 克隆到质粒 (pED261-145、pED261-146、pED261-208、pED263-145、pED263-146和pED263-208) 中的目标基因可以封包到AAV9中。

[0135] 图12示出了在HSMM细胞中通过使用3种AAV9 (AAV9-ED261-145、AAV9-ED261-208和AAV9-ED263-208) 对人UTRN基因表达的激活 (N=3-4, 对于AAV组; N=8, 对于非AAV对照。误差条表示标准误差)。这些AAV9激活了人UTRN基因的表达。

[0136] 图13示出了靶向性向导相对于非靶向性向导标准化的RNA测序结果, 绘制为log₂ 倍数变化对比标准化计数的平均值的图 (图A: 向导编号145对比NTg1; 图B: 向导编号146对比NTg1; 图C: 向导编号208对比NTg1)。

具体实施方式

[0137] 下面详细说明本发明的实施方式。

[0138] 1. 多核苷酸

[0139] 本发明提供了一种包含以下碱基序列的多核苷酸(下文中有时称为“本发明的多核苷酸”):

[0140] (a) 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,和

[0141] (b) 编码向导RNA的碱基序列,所述向导RNA靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域(即18至24个连续核苷酸)。

[0142] 将本发明的多核苷酸引入到所需细胞中并转录以产生核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,以及靶向人UTRN基因的表达调节区的特定区域的向导RNA。这些融合蛋白和向导RNA形成复合物(下文中,所述复合物有时称为“核糖核蛋白;RNP”)并协同作用于上述特定区域,从而激活人UTRN基因的转录。

[0143] (1) 定义

[0144] 在本说明书中,“人肌营养相关蛋白(UTRN)基因的表达调节区”是指其中RNP结合到所述区域可以激活人UTRN基因的表达的任何区域。即,人UTRN基因的表达调节区可以存在于人UTRN基因的任何区域,诸如启动子区、增强子区、内含子和外显子等,只要RNP结合激活人UTRN基因的表达即可。在本说明书中,当表达调节区由特定序列示出时,表达调节区在概念上包含有义链序列和反义链序列。

[0145] 在本发明中,核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白被向导RNA募集到人UTRN基因的表达调节区中的特定区域中。在本说明书中,“靶向.....的向导RNA”是指“将融合蛋白募集到.....中的向导RNA”。

[0146] 在本说明书中,“向导RNA(也称为“gRNA”)”是包含基因组特异性CRISPR-RNA(称为“crRNA”)的RNA。crRNA是与靶向序列(下文描述)的互补序列结合的RNA。当Cpf1用作CRISPR效应蛋白时,“向导RNA”是指包含由crRNA和与其5'末端连接的特异性序列组成的RNA的RNA(例如,在FnCpf 1的情况下,SEQ ID NO:106所示的RNA序列)。当Cas9用作CRISPR效应蛋白时,“向导RNA”是指嵌合体RNA(称为“单个向导RNA(sgRNA)”),包含crRNA和与其3'末端连接的反式激活crRNA(称为“tracrRNA”) (例如,参见Zhang F.等人,《人类分子遗传学(Hum Mol Genet)》,2014年9月15日;23(R1):R40-6和Zetsche B.等人,《细胞》,2015年10月22日;163(3):759-71,通过引用全部并入本文)。

[0147] 在本说明书中,将与在人UTRN基因的表达调节区中与crRNA结合的序列互补的序列称为“靶向序列”。即,在本说明书中,“靶向序列”是存在于人UTRN基因的表达调节区中且与PAM(前间隔子相邻基序)相邻的DNA序列。当Cpf1用作CRISPR效应蛋白时,PAM与靶向序列的5'侧相邻。当Cas9用作CRISPR效应蛋白时,PAM与靶向序列的3'侧相邻。靶向序列可以存在于人UTRN基因的表达调节区的有义链序列侧或反义链序列侧(参见,例如,上述Zhang F.等人,《人类分子遗传学》,2014年9月15日;23(R1):R40-6和Zetsche B.等人,《细胞》,2015年10月22日;163(3):759-71,通过引用全部并入本文)。

[0148] (2) 核酸酶缺陷型CRISPR效应蛋白

[0149] 在本发明中,使用核酸酶缺陷型CRISPR效应蛋白,将与其融合的转录激活子募集到人UTRN基因的表达调节区中。用于本发明的核酸酶缺陷型CRISPR效应蛋白(下文中简称为“CRISPR效应蛋白”)没有特别限制,只要它与gRNA形成复合物并被募集到人UTRN基因的

表达调节区中即可。例如,可以包括核酸酶缺陷型Cas9(下文中有时也称为“dCas9”)或核酸酶缺陷型Cpf1(下文中有时也称为“dCpf1”)。

[0150] 上述dCas9的实例包括但不限于以下各者的核酸酶缺陷型变体:酿脓链球菌(*Streptococcus pyogenes*)来源的Cas9(SpCas9;PAM序列:NGG(N为A、G、T或C,以下相同))、嗜热链球菌(*Streptococcus thermophilus*)来源的Cas9(StCas9;PAM序列:NNAGAAW(W为A或T,以下相同))、脑膜炎奈瑟氏球菌(*Neisseria meningitidis*)来源的Cas9(NmCas9;PAM序列:NNNGATT),或金黄色葡萄球菌(*Staphylococcus aureus*)来源的Cas9(SaCas9;PAM序列:NNGRRT(R为A或G,以下相同)等(例如,参见Nishimasu等人,《细胞》,2014年2月27日;156(5):935-49,Esvelt KM等人,《自然方法学(Nat Methods)》,2013年11月;10(11):1116-21,Zhang Y.《分子细胞(Mol Cell)》,2015年10月15日;60(2):242-55,以及Friedland AE等人,《基因组生物学(Genome Biol)》,2015年11月24日;16:257,通过引用全部并入本文)。例如,在SpCas9的情况下,可以使用其中第10位Asp残基被转化为Ala残基并且第840位His残基被转化为Ala残基的双突变体(有时称为“dSpCas9”) (参见,例如,上述Nishimasu等人,《细胞》,2014)。或者,在SaCas9的情况下,可以使用其中第10位Asp残基被转化为Ala残基并且第580位Asn残基被转化为Ala残基的双突变体(SEQ ID NO:107),或者其中第10位Asp残基被转化为Ala残基并且第557位His残基被转化为Ala残基的双突变体(SEQ ID NO:108)(下文中,这些双突变体中的任何一个有时称为“dSaCas9”) (参见,例如Friedland AE等人,《基因组生物学》,2015,通过引用全部并入本文)。

[0151] 另外,在本发明的一个实施方式中,作为dCas9,也可以使用通过修饰上述dCas9的氨基酸序列的一部分而得到的变体,所述变体与gRNA形成复合体并被募集到人UTRN基因的表达调节区中。此类变体的实例包括具有部分缺失的氨基酸序列的截短型变体。在本发明的一个实施方式中,作为dCas9,可以使用在PCT/JP2019/022795和PCT/JP2019/041751中公开的变体,通过引用全部并入本文。具体地,也可以使用通过从其中第10位Asp残基被转化为Ala残基且第580位Asn残基被转化为Ala残基的dSaCas9双突变体中缺失第721位至第745位氨基酸而获得的dSaCas9(SEQ ID NO:109),或其中缺失的部分被肽连接子替换的dSaCas9(例如,SEQ ID NO:111所示者,其中缺失的部分被GGSGGS连接子(SEQ ID NO:110)替换,和SEQ ID NO:214所示者,其中缺失的部分被SGGGS连接子(SEQ ID NO:213)替换,等)(下文中,这些双突变体中的任何一个有时称为“dSaCas9[-25]”),或通过从作为上述dSaCas9双突变体中缺失第482位至第648位氨基酸而获得的dSaCas9(SEQ ID NO:112),或其中缺失的部分被肽连接子替换的dSaCas9(SEQ ID NO:113所示者,其中缺失的部分被GGSGGS连接子替换)。

[0152] 上述dCpf1的实例包括但不限于以下各者的核酸酶缺陷型变体:来源于新凶手弗朗西丝氏菌(*Francisella novicida*)的Cpf1(FnCpf1;PAM序列:NTT)、来源于氨基球菌属(*Acidaminococcus* sp.)的Cpf1(AsCpf1;PAM序列:NTTT),或来源于毛螺科菌(*Lachnospiraceae* bacterium)的Cpf1(LbCpf1;PAM序列:NTTT)等(参见,例如,Zetsche B.等人,《细胞》,2015年10月22日;163(3):759-71,Yamano T等人,《细胞》,2016年5月5日;165(4):949-62,和Yamano T等人,《分子细胞》,2017年8月17日;67(4):633-45,通过引用全部并入本文)。例如,在FnCpf1的情况下,可以使用其中第917位Asp残基转化为Ala残基且第1006位Glu残基转化为Ala残基的双突变体(例如,参见上述Zetsche B等人,《细胞》,2015,

通过引用全部并入本文)。在本发明的一个实施方式中,作为dCpf1,也可以使用通过修饰上述dCpf1的氨基酸序列的一部分而获得的变体,所述变体与gRNA形成复合物并被募集到人UTRN基因的表达调节区中。

[0153] 在本发明的一个实施方式中,dCas9用作核酸酶缺陷型CRISPR效应蛋白。在一个实施方式中,dCas9是dSaCas9,且在特定的实施方式中,dSaCas9是dSaCas9[-25]。

[0154] 包含编码CRISPR效应蛋白的碱基序列的多核苷酸可以通过例如以下步骤克隆:基于其cDNA序列信息,合成覆盖编码所述蛋白质的所需部分的区域的寡DNA引物,并通过PCR方法使用从产生所述蛋白质的细胞制备的总RNA或mRNA级分作为模板扩增所述多核苷酸。另外,包含编码核酸酶缺陷型CRISPR效应蛋白的碱基序列的多核苷酸可以通过以已知的定点诱变方法将突变引入到编码克隆的CRISPR效应蛋白的核苷酸序列中以将在对DNA切割活性重要的位置处的氨基酸残基(例如,可以包括但不限于:在SaCas9的情况下,第10位Asp残基、第557位His残基和第580位Asn残基;在FnCpf1的情况下,第917位Asp残基和第1006位Glu残基等)转化成其它氨基酸而获得。

[0155] 或者,包含编码核酸酶缺陷型CRISPR效应蛋白的碱基序列的多核苷酸可以通过化学合成或化学合成与PCR方法或Gibson组装方法的组合,基于其cDNA序列信息而获得,且还可以进一步构建为经过密码子优化以产生适合在人体内表达的密码子的碱基序列。

[0156] (3) 转录激活子

[0157] 在本发明中,人UTRN基因表达通过与核酸酶缺陷型CRISPR效应蛋白融合的转录激活子的作用而被激活。在本说明书中,“转录激活子”是指能够激活人UTRN基因的基因转录的蛋白质或保留其功能的肽片段。用于本发明的转录激活子没有特别限制,只要其可以激活人UTRN基因的表达即可。例如,其包括VP64、VPH、VPR、miniVR和microVR,其具有转录激活能力的变体等。VP64是由SEQ ID NO:114所示的50个氨基酸组成的肽。VPH是VP64、p65和HSF1的融合蛋白,具体地,是由SEQ ID NO:115所示的376个氨基酸组成的肽。VPR是VP64、p65和爱泼斯坦-巴尔病毒(Epstein-Barr virus)的复制和转录激活子(RTA)的融合蛋白,例如,由SEQ ID NO:116所示的523个氨基酸组成的肽、由SEQ ID NO:216所示的519个氨基酸组成的肽等。VP64、VPH和VPR是已知的并且在例如以下文献中详细公开:Chavez A.等人,《自然方法学》,2016年7月;13(7):563-7和Chavez A.等人,《自然方法学》,2015年4月;12(4):326-8,通过引用全部并入本文。在本发明的一个实施方式中,作为转录激活子,可以使用包含RTA的转录激活结构域和VP64的肽。RTA的转录激活结构域是已知的并且在例如以下文献中公开:《病毒学杂志(J Virology)》,1992年9月;66(9):5500-8,通过引用全部并入本文。作为这种肽的序列,miniVR是由SEQ ID NO:117所示的167个氨基酸组成的肽,microVR是由SEQ ID NO:118所示的140个氨基酸组成的肽。SEQ ID NO:117所示的氨基酸序列由如下氨基酸序列组成,其中RTA的较短的转录激活结构域,即RTA的第493位至第605位氨基酸残基,和VP64以G-S-G-S连接子(SEQ ID NO:209)连接。SEQ ID NO:118所示的氨基酸序列由如下氨基酸序列组成,其中RTA的更短的转录激活结构域,即RTA的第520位至第605位氨基酸残基,和VP64以G-S-G-S连接子连接。miniVR和microVR的详情在PCT/JP2019/030972中进行了描述,通过引用全部并入本文。任何上述转录激活子都可以进行任何修饰和/或改变,只要其保持其转录激活能力即可。例如,作为本发明中的转录激活子,还可以使用(i) 包含SEQ ID NO:117所示的氨基酸序列的肽,(ii) 包含SEQ ID NO:117所示但其中缺失、替换、插入

和/或添加了1个或几个(例如2、3、4、5或更多个)氨基酸的氨基酸序列的肽,(iii)包含与SEQ ID NO:117所示的氨基酸序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的氨基酸序列的肽,(iv)由SEQ ID NO:117所示但其中缺失、替换、插入和/或添加了1个或几个(例如2、3、4、5或更多个)氨基酸的氨基酸序列组成的肽,或(v)由与SEQ ID NO:117所示的氨基酸序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的氨基酸序列组成的肽,只要其保持其转录激活能力即可。例如,作为本发明中的转录激活子,还可以使用(i)包含SEQ ID NO:118所示的氨基酸序列的肽,(ii)包含SEQ ID NO:118所示但其中缺失、替换、插入和/或添加了1个或几个(例如2、3、4、5或更多个)氨基酸的氨基酸序列的肽,(iii)由与SEQ ID NO:118中所示的氨基酸序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的氨基酸序列组成的肽,(iv)由SEQ ID NO:118所示但其中缺失、替换、插入和/或添加了1个或几个(例如2、3、4、5或更多个)氨基酸的氨基酸序列组成的肽,或(v)由与SEQ ID NO:118所示的氨基酸序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的氨基酸序列组成的肽,只要其保持其转录激活能力即可。

[0158] 包含编码转录激活子的碱基序列的多核苷酸可以通过化学合成或化学合成与PCR方法或Gibson组装方法的组合来构建。此外,包含编码转录激活子的碱基序列的多核苷酸还可以被构建为密码子优化的DNA序列,成为适合于在人体内表达的密码子。

[0159] 包含编码转录激活子和核酸酶缺陷型CRISPR效应蛋白的融合蛋白的碱基序列的多核苷酸可以通过将编码核酸酶缺陷型CRISPR效应蛋白的碱基序列与编码转录激活子的碱基序列直接或在添加编码连接子、NLS(核定位信号)、标签和/或其它的碱基序列后连接来制备。在本发明中,转录激活子可以与CRISPR效应蛋白的N-末端或C-末端融合。作为连接子,可以使用氨基酸数为约2至50的连接子,其具体实例包括但不限于其中甘氨酸(G)和丝氨酸(S)交替连接的G-S-G-S连接子,及类似物。

[0160] (4) 向导RNA

[0161] 在本发明中,核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白可以被向导RNA募集到人UTRN基因的表达调节区。如上述“(1)定义”中所述,向导RNA包含crRNA,且所述crRNA与靶向序列的互补序列结合。crRNA可能不与靶向序列的互补序列完全互补,只要向导RNA可以将融合蛋白募集到靶区域即可,且可以包含靶向序列中缺失、替换、插入和/或添加了至少1至3个碱基的碱基序列。

[0162] 当dCas9用作核酸酶缺陷型CRISPR效应蛋白时,例如,可以使用已发布的gRNA设计网站(CRISPR Design Tool、CRISPR direct等)确定靶向序列。具体而言,从靶基因(即人UTRN基因)的序列中,列出了长度为约20个核苷酸的候选靶向序列,其中PAM(例如,在SaCas9的情况下,NNGRRT)与其3'侧相邻,且可以将这些候选靶向序列中在人基因组中具有少量脱靶位点的序列用作靶向序列。靶向序列的碱基长度为18至24个核苷酸,优选为20至23个核苷酸,更优选为21至23个核苷酸。作为预测脱靶位点数目的初步筛选,许多生物信息学工具是已知的且可公开获得,可用于预测脱靶效应最低的靶向序列。其实例包括诸如Benchling(<https://benchling.com>)和COSMID(具有错配、插入和缺失的CRISPR脱靶位点)(可在互联网<https://crispr.bme.gatech.edu>上找到)的生物信息学工具。使用这些,可以

总结出与gRNA靶向的碱基序列的相似性。当要使用的gRNA设计软件不具有搜索靶基因组的脱靶位点的功能时,例如,可以通过对靶基因组进行关于候选靶向序列的3'侧的8至12个核苷酸(具有靶向核苷酸序列高识别能力的种子序列)的Blast搜索来搜索脱靶位点。

[0163] 在本发明的一个实施方式中,在存在于人6号染色体(Chr 6)的GRCh38.p12位置上的区域中,以下五个区域可能是人UTRN基因的表达调节区。组蛋白修饰图谱强力表明这些区域为表达调节区。因此,在本发明的一个实施方式中,在存在于人6号染色体(Chr 6)的GRCh38.p12位置上的以下五个区域中的至少一个区域中,靶向序列的长度可以为18至24个核苷酸,优选为20至23个核苷酸,更优选为21至23个核苷酸:

[0164] (1) 144,215,500-144,217,000,

[0165] (2) 144,248,500-144,249,800,

[0166] (3) 144,264,000-144,267,000,

[0167] (4) 144,283,900-144,288,300,

[0168] (5) 144,292,500-144,295,500。

[0169] 在本发明的一个实施方式中,在存在于上述区域(3)中的SEQ ID NO:104所示的区域中或存在于上述区域(4)中的SEQ ID NO:105、135、141、153、167或172所示的区域中,靶向序列的长度可以是连续的18至24个核苷酸,优选为20至23个核苷酸,更优选为21至23个核苷酸。

[0170] 在本发明的另一个实施方式中,靶向序列可以是SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示的碱基序列。SEQ ID NO:45和46所示的碱基序列是包含在上述SEQ ID NO:104所示区域中的靶向序列,且SEQ ID NO:58、59、60、155、156、157和159所示的碱基序列是包含在上述SEQ ID NO:105所示区域中的靶向序列。

[0171] 在本发明的一个实施方式中,编码crRNA的碱基序列可以是与靶向序列相同的碱基序列。例如,当将SEQ ID NO:4所示的靶向序列(AGAAAAGCGGCCCTAGGGGC)作为编码crRNA的碱基序列引入细胞时,从所述序列转录的crRNA是AGAAAAGCGGCCCUAGGGGC(SEQ ID NO:119),并与GCCCCTAGGGGCCGCTTTTCT(SEQ ID NO:120)结合,后者是与SEQ ID NO:4所示的碱基序列互补的序列,且存在于人UTRN基因的表达调节区中。在另一个实施方式中,可以使用靶向序列中缺失、替换、插入和/或添加了至少1至3个碱基的碱基序列作为编码crRNA的碱基序列,只要向导RNA可以募集融合蛋白至靶区域即可。因此,在本发明的一个实施方式中,作为编码crRNA的碱基序列,可以使用SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示的碱基序列,或SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0172] 在本发明的一个实施方式中,SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示的碱基序列可以用作编码crRNA的碱基序列以分别产生包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的crRNA的gRNA。在本发明的另一个实施方式中,gRNA可以包含SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示的碱基序列,或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0173] 当dCpf1用作核酸酶缺陷型CRISPR效应蛋白时,可以将编码gRNA的碱基序列设计

为编码在5'末端连接有特定RNA的crRNA的DNA序列。本领域普通技术人员可以根据要使用的dCpf1适当地选择所述与crRNA的5'末端连接的RNA和编码所述RNA的DNA序列。例如,当使用dFnCpf1时,其中SEQ ID NO:121,即AATTTCTACTGTTGTAGAT连接到靶向序列的5'侧的碱基序列可以用作编码gRNA的碱基序列(转录成RNA时,带下划线部分的序列形成碱基对以形成茎环结构)。要添加到5'末端的序列可以是通常用于各种Cpf1蛋白的序列,其中缺失、替换、插入和/或添加了至少1至6个碱基,只要转录后gRNA可以将融合蛋白募集到表达调节区即可。

[0174] 当dCas9用作CRISPR效应蛋白时,可以将编码gRNA的碱基序列设计为其中编码已知tracrRNA的DNA序列连接到编码crRNA的DNA序列的3'末端的DNA序列。本领域普通技术人员可以根据要使用的dCas9适当地选择这样的tracrRNA和编码所述tracrRNA的DNA序列。例如,当使用dSaCas9时,SEQ ID NO:122所示的碱基序列用作编码tracrRNA的DNA序列。编码tracrRNA的DNA序列可以是编码通常用于各种Cas9蛋白的tracrRNA的碱基序列,其中缺失、替换、插入和/或添加了至少1至6个碱基,只要转录后gRNA可以将融合蛋白募集到表达调节区即可。

[0175] 包含以此方式设计的编码gRNA的碱基序列的多核苷酸可以使用已知的DNA合成方法化学合成。

[0176] 在本发明的另一个实施方式中,本发明的多核苷酸可以包含至少两种不同的编码gRNA的碱基序列。例如,多核苷酸可以包含至少两种不同的编码向导RNA的碱基序列,其中所述至少两种不同的碱基序列选自包含SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示序列的碱基序列。在本发明的一个实施方式中,多核苷酸可以包含至少两种不同的编码向导RNA的碱基序列,其中所述至少两种不同的碱基序列选自包含SEQ ID NO:45、46或59所示序列的碱基序列。

[0177] (5) 启动子序列

[0178] 在本发明的一个实施方式中,启动子序列可以可操作地连接至编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列和/或编码gRNA的碱基序列各自的上游。可能连接的启动子没有特别限制,只要它在靶细胞中显示出启动子活性即可。可能连接至编码融合蛋白的碱基序列上游的启动子序列的实例包括但不限于EFS启动子、EF-1 α 启动子、CMV(巨细胞病毒)启动子、CK8启动子、MHC启动子、MLC启动子、Des启动子、cTnC启动子、MYOD启动子、hTERT启动子、SR α 启动子、SV40启动子、LTR启动子、CAG启动子、RSV(劳斯氏肉瘤病毒)启动子等。可能连接至编码gRNA的碱基序列上游的启动子序列的实例包括但不限于为po1 III启动子的U6启动子、SNR6启动子、SNR52启动子、SCR1启动子、RPR1启动子、U3启动子、H1启动子和tRNA启动子等。在本发明的一个实施方式中,当多核苷酸包含两种以上编码向导RNA的碱基序列时,单个启动子序列可以可操作地连接至所述两种以上碱基序列的上游。在另一个实施方式中,启动子序列可以可操作地连接至两种以上碱基序列各自的上游,其中可操作地连接至各碱基序列的启动子序列可以相同或不同。

[0179] 在本发明的一个实施方式中,可以使用肌肉特异性启动子作为连接至编码上述融合蛋白的碱基序列上游的启动子序列。肌肉特异性启动子的实例包括但不限于CK8启动子、CK6启动子、CK1启动子、CK7启动子、CK9启动子、心肌肌钙蛋白C启动子、 α 肌动蛋白启动子、肌球蛋白重链激酶(MHCK)启动子、肌球蛋白轻链2A启动子、肌营养不良蛋白启动子、肌肉肌

酸激酶启动子、dMCK启动子、tMCK启动子、enh348 MCK启动子、合成C5-12(Syn)启动子、Myf5启动子、MLC1/3f启动子、MYOD启动子、Myog启动子、Pax7启动子、Des启动子等(关于肌肉特异性启动子的详情,参见,例如,US2011/0212529A、McCarthy JJ等人,《骨骼肌(Skeletal Muscle)》,2012年5月;2(1):8,Wang B.等人,《基因治疗(Gene Ther)》(2008年11月;15(22):1489-99,通过引用全部并入本文)。

[0180] 在本发明的一个实施方式中,U6启动子可以用作编码gRNA的碱基序列的启动子序列,CK8启动子可以用作编码融合蛋白的碱基序列的启动子序列。具体地,对于U6启动子,可以使用以下碱基序列:(i)SEQ ID NO:128所示的碱基序列,(ii)SEQ ID NO:128所示但其中缺失、替换、插入和/或添加了1个或几个(例如2、3、4、5或更多个)碱基的碱基序列,在靶细胞中具有启动子活性,或(iii)与SEQ ID NO:128所示的碱基序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的碱基序列,在靶细胞中显示出启动子活性。对于CK8启动子,可以使用以下碱基序列:(i)SEQ ID NO:191所示的碱基序列,(ii)SEQ ID NO:191所示但其中缺失、替换、插入和/或添加了1个或几个(例如2、3、4、5或更多个)碱基的碱基序列,在靶细胞中具有启动子活性,或(iii)与SEQ ID NO:191所示的碱基序列具有不小于90%(例如90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或以上)同一性的碱基序列,在靶细胞中显示出启动子活性。

[0181] (6)其它碱基序列

[0182] 此外,本发明的多核苷酸除了上述提及的那些以外,可以进一步包含例如多腺苷酸化(polyA)信号、Kozak共有序列等已知序列,以达成提高由编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的转录产生的mRNA的翻译效率的目的。例如,本发明中的多腺苷酸化信号可以包括hGH polyA、bGH polyA、2x sNRP-1 polyA(参见US7557197B2,通过引用全部并入本文)等。另外,本发明的多核苷酸可以包含编码连接子序列的碱基序列、编码NLS的碱基序列和/或编码标签的碱基序列。

[0183] (7)本发明的示例性实施方式

[0184] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:

[0185] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,

[0186] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的启动子序列,

[0187] 一种或两种编码向导RNA的碱基序列,其中所述一种或两种碱基序列选自包含SEQ ID NO:45、46或59所示序列的碱基序列,或包含SEQ ID NO:45、46或59所示但其中缺失、替换、插入和/或添加了1至3个碱基的序列的碱基序列,以及

[0188] 用于所述编码向导RNA的碱基序列的U6启动子,

[0189] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9或dSaCas9[-25],

[0190] 其中所述转录激活子选自VP64、VPH、VPR、miniVR和microVR,且

[0191] 其中所述用于编码融合蛋白的碱基序列的启动子序列选自EF-1 α 启动子、EFS启动子和CK8启动子。

[0192] 所述多核苷酸可以进一步包含hGH polyA、bGH polyA或2x sNRP-1 polyA。

[0193] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:

[0194] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,

- [0195] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的CK8启动子,
- [0196] 一种或两种编码向导RNA的碱基序列,其中所述一种或两种碱基序列选自包含SEQ ID NO:45、46或59所示序列的碱基序列,或包含SEQ ID NO:45、46或59所示但其中缺失、替换、插入和/或添加了1至3个碱基的序列的碱基序列,以及
- [0197] 用于所述编码向导RNA的碱基序列的U6启动子,
- [0198] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9或dSaCas9[-25],且
- [0199] 其中所述转录激活子是miniVR或microVR。
- [0200] 所述多核苷酸可以进一步包含bGH polyA或2x sNRP-1 polyA。
- [0201] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:
- [0202] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,
- [0203] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的CK8启动子,
- [0204] 一种或两种编码向导RNA的碱基序列,其中所述一种或两种碱基序列选自包含SEQ ID NO:45、46或59所示序列的碱基序列,或包含SEQ ID NO:45、46或59所示但其中缺失、替换、插入和/或添加了1至3个碱基的序列的碱基序列,以及
- [0205] 用于所述编码向导RNA的碱基序列的U6启动子,
- [0206] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9,且
- [0207] 其中所述转录激活子是miniVR。
- [0208] 所述多核苷酸可以进一步包含bGH polyA或2x sNRP-1 polyA。
- [0209] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:
- [0210] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,
- [0211] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的CK8启动子,
- [0212] 一种或两种编码向导RNA的碱基序列,其中所述一种或两种碱基序列选自包含SEQ ID NO:45、46或59所示序列的碱基序列,或包含SEQ ID NO:45、46或59所示但其中缺失、替换、插入和/或添加了1至3个碱基的序列的碱基序列,以及
- [0213] 用于所述编码向导RNA的碱基序列的U6启动子,
- [0214] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9,且
- [0215] 其中所述转录激活子是microVR。
- [0216] 所述多核苷酸可以进一步包含bGH polyA或2x sNRP-1 polyA。
- [0217] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:
- [0218] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,
- [0219] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的CK8启动子,
- [0220] 编码向导RNA的碱基序列,包含SEQ ID NO:59所示的碱基序列,或SEQ ID NO:59所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列,以及
- [0221] 用于所述编码向导RNA的碱基序列的U6启动子,
- [0222] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9,且

[0223] 其中所述转录激活子是miniVR。

[0224] 所述多核苷酸可以进一步包含2x sNRP-1polyA。

[0225] 在本发明的一个实施方式中,提供了一种多核苷酸,所述多核苷酸包含:

[0226] 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列,

[0227] 用于所述编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列的CK8启动子,

[0228] 编码向导RNA的碱基序列,包含SEQ ID NO:59所示的碱基序列,或SEQ ID NO:59所示但其中缺失、替换、插入和/或添加了1至3个碱基的碱基序列,以及

[0229] 用于所述编码向导RNA的碱基序列的U6启动子,

[0230] 其中所述核酸酶缺陷型CRISPR效应蛋白是dSaCas9,且

[0231] 其中所述转录激活子是microVR。

[0232] 所述多核苷酸可以进一步包含2x sNRP-1polyA。

[0233] 在本发明的多核苷酸的一个实施方式中,多核苷酸从5'末端开始依次包含(i) 编码核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白的碱基序列和(ii) 编码gRNA的碱基序列。在另一个实施方式中,多核苷酸从5'末端开始依次包含(ii) 编码gRNA的碱基序列和(i) 编码核酸酶缺陷型CRISPR效应蛋白与转录激活子的融合蛋白的碱基序列。

[0234] 2. 载体

[0235] 本发明提供了一种包含本发明的多核苷酸的载体(下文中有时称为“本发明的载体”)。本发明的载体可以是质粒载体或病毒载体。

[0236] 当本发明的载体是质粒载体时,所使用的质粒载体没有特别限制,且可以是任何质粒载体,诸如克隆质粒载体和表达质粒载体。通过用已知方法将本发明的多核苷酸插入质粒载体中来制备质粒载体。

[0237] 当本发明的载体是病毒载体时,使用的病毒载体没有特别限制,其实例包括但不限于腺病毒载体、腺相关病毒(AAV)载体、慢病毒载体、逆转录病毒载体、仙台病毒载体等。在本说明书中,“病毒载体”还包括其衍生物。考虑到在基因治疗中的用途,优选使用AAV载体,原因在于它能够长时间表达转基因,且它衍生自非致病性病毒并具有高安全性。

[0238] 可以通过已知方法制备包含本发明的多核苷酸的病毒载体。简而言之,制备插入有本发明的多核苷酸的用于病毒表达的质粒载体,将所述载体转染到合适的宿主细胞中以允许瞬时产生包含本发明的多核苷酸的病毒载体,并收集病毒载体。

[0239] 在本发明的一个实施方式中,当使用AAV载体时,AAV载体的血清型没有特别限制,只要可以激活靶标中人UTRN基因的表达即可,且可以使用AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAVrh.10等中的任何一个(对于AAV的各种血清型,参见,例如,WO 2005/033321和EP2341068(A1),通过引用全部并入本文)。AAV变体的实例包括但不限于具有修饰的衣壳的新血清型(例如,WO 2012/057363,通过引用全部并入本文)等。例如,在本发明的一个实施方式中,可以使用具有修饰的衣壳以提高对肌肉细胞的感染性的新血清型,诸如AAV₅₈₇MTP、AAV₅₈₈MTP、AAV-B1、AAVM41、AAVS1_P1和AAVS10_P1等(参见Yu等人,《基因治疗》2009年4月;16(8):953-62,Choudhury等人,《分子治疗》,2016年4月;24(7):1247-57, Yang等人,《美国科学院院报(Proc Natl Acad Sci USA)》(2009年3月10日;106(10):3946-51和WO2019/207132,通过引用全部并入本文)。

[0240] 当制备AAV载体时,可以使用诸如以下的已知方法:(1)使用质粒的方法,(2)使用杆状病毒的方法,(3)使用单纯疱疹病毒的方法,(4)使用腺病毒的方法,或(5)使用酵母的方法(例如,《应用微生物学与生物技术(Appl Microbiol Biotechnol)》,2018;102(3):1045-1054等,通过引用全部并入本文)。例如,当通过使用质粒的方法制备AAV载体时,首先,制备在野生型AAV基因组序列的两个末端包含反向的末端重复序列(ITR)并插入有本发明的多核苷酸代替编码Rep蛋白和衣壳蛋白的DNA的载体质粒。另一方面,将编码形成病毒粒子所必需的Rep蛋白和衣壳蛋白的DNA插入其它质粒中。此外,将包含负责AAV增殖所必需的腺病毒的辅助作用的基因(E1A、E1B、E2A、VA和E4orf6)的质粒制备为腺病毒辅助质粒。将这三种质粒共转染到宿主细胞中使得在细胞中产生重组AAV(即,AAV载体)。作为宿主细胞,优选使用能够提供负责上述辅助作用的基因的一部分基因产物(蛋白质)的细胞(例如293细胞等)。当使用这种细胞时,不必在上述腺病毒辅助质粒中携带编码可以从宿主细胞提供的蛋白质的基因。产生的AAV载体存在于细胞核中。因此,通过以下步骤来制备所需的AAV载体:用冻融破坏宿主细胞,收集病毒,然后通过使用氯化铯的密度梯度超速离心法,柱法等对病毒级分进行分离和纯化。

[0241] AAV载体在安全性、基因转导效率等方面具有很大的优势,且用于基因治疗。然而,已知可以封包在AAV载体中的多核苷酸的大小是有限制的。例如,在本发明的一个实施方式中,包括包含编码dSaCas9和miniVR或microVR的融合蛋白的碱基序列、编码靶向人UTRN的表达调节区的gRNA的碱基序列以及作为启动子序列的EFS启动子序列或CK8启动子序列和U6启动子序列,以及ITR部分在内的多核苷酸的碱基长度的全长为约4.85kb,且它们可以被封包在单个AAV载体中。

[0242] 3. 药物组合物

[0243] 本发明还提供了一种包含本发明的多核苷酸或本发明的载体的药物组合物(下文有时称为“本发明的药物组合物”)。本发明的药物组合物可以用于治疗或预防DMD或BMD。

[0244] 本发明的药物组合物包含本发明的多核苷酸或本发明的载体作为活性成分,且可以制备成包含这种活性成分(即,本发明的多核苷酸或本发明的载体)和通常药学上可接受的载剂的制剂。

[0245] 本发明的药物组合物经肠胃外施用,且可以局部或全身施用。本发明的药物组合物可以通过但不限于例如静脉内施用、动脉内施用、皮下施用、腹膜内施用或肌肉内施用来施用。

[0246] 本发明的药物组合物向对象施用的剂量没有特别限制,只要它是用于治疗 and/或预防的有效量即可。可以根据活性成分、剂型、对象的年龄和体重、施用方案、施用方法等适当地优化。

[0247] 在本发明的一个实施方式中,本发明的药物组合物不仅可以施用于患有DMD或BMD的对象,而且还可以预防性地施用于基于遗传背景分析等将来可能发展为DMD或BMD的对象。本说明书中的术语“治疗”除了治愈疾病以外,还包括疾病缓解。另外,术语“预防”除了预防疾病发作以外,还可以包括延迟疾病发作。本发明的药物组合物也可以称为“本发明的药剂”等。

[0248] 4. 治疗或预防DMD或BMD的方法

[0249] 本发明还提供一种用于治疗或预防DMD或BMD的方法,所述方法包括向有此需要的

对象施用本发明的多核苷酸或本发明的载体(下文中有时称为“本发明的方法”)。另外,本发明包括用于治疗或预防DMD或BMD的本发明的多核苷酸或本发明的载体。此外,本发明包括本发明的多核苷酸或本发明的载体在制备用于治疗或预防DMD或BMD的药物组合物中的用途。

[0250] 本发明的方法可以通过向患有DMD或BMD的对象施用上述本发明的药物组合物来实践,且剂量、施用途径、对象等与上述相同。

[0251] 症状的测量可以在使用本发明的方法开始治疗之前以及在治疗之后的任何时间进行,以确定对象对治疗的反应。

[0252] 本发明的方法可以改善对象的骨骼肌和/或心肌的功能。要改善功能的肌肉没有特别限制,且可以列举任何肌肉和肌肉群。

[0253] 5.核糖核蛋白

[0254] 本发明提供了一种核糖核酸蛋白,所述核糖核酸蛋白包含以下(下文中有时称为“本发明的RNP”):

[0255] (c) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,和

[0256] (d) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA。

[0257] 作为包含在本发明的RNP中的核酸酶缺陷型CRISPR效应蛋白、转录激活子和向导RNA,可以使用在上述“1.多核苷酸”部分中详细描述的核酸酶缺陷型CRISPR效应蛋白、转录激活子和向导RNA。包含在本发明的RNP中的核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白可以通过例如将编码所述融合蛋白的多核苷酸引入到细胞、细菌或其它生物中以允许表达而产生,或由体外翻译系统通过使用所述多核苷酸而产生。另外,包含在本发明的RNP中的向导RNA可以通过例如化学合成或体外转录系统通过使用编码向导RNA的多核苷酸而产生。将由此制备的融合蛋白和向导RNA混合以制备本发明的RNP。必要时,可以混合其它物质,诸如金粒子。为了将本发明的RNP直接递送至靶细胞、组织等,可以通过已知方法将RNP封装在脂质纳米粒子(LNP)中。可以通过已知方法将本发明的RNP引入至靶细胞、组织等中。对于封装在LNP中和引入方法,可以参考例如, Lee K.等人,《自然-生物医学工程(Nat Biomed Eng)》,2017;1:889-901, W02016/153012,通过引用全部并入本文,等等。

[0258] 在本发明的一个实施方式中,本发明的RNP中所包含的向导RNA靶向存在于人6号染色体(Chr 6)的GRCh38.p12位置上的以下五个区域中至少一个区域中连续的18至24个核苷酸长度,优选20至23个核苷酸长度,更优选21至23个核苷酸长度:

[0259] (1) 144,215,500-144,217,000,

[0260] (2) 144,248,500-144,249,800,

[0261] (3) 144,264,000-144,267,000,

[0262] (4) 144,283,900-144,288,300,

[0263] (5) 144,292,500-144,295,500。

[0264] 在一个实施方式中,向导RNA靶向SEQ ID NO:104、105、135、141、153、167或172所示的DNA序列中连续的18至24个核苷酸长度,优选20至23个核苷酸长度,更优选为21至23个核苷酸长度的碱基序列。在一个实施方式中,向导RNA靶向包含SEQ ID NO:45、46、58、59、60、135、141、153、155、156、157、159、167或172所示的全部或部分序列的区域。在本发明

的一个实施方式中,可以使用包含SEQ ID NO:194、195、196、197、197、198、199、200、201、202、203、204、205、206或207所示的crRNA的向导RNA,或SEQ ID NO:194、195、196、197、198、199、200、201、202、203、204、205、206或207所示但其中分别缺失、替换、插入和/或添加了1至3个碱基的碱基序列。

[0265] 6.其它

[0266] 本发明还提供了一种组合物或试剂盒,所述组合物或试剂盒包含以下用于激活人肌营养相关蛋白基因的表达:

[0267] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

[0268] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA,或编码所述向导RNA的多核苷酸。

[0269] 本发明还提供了一种治疗或预防杜氏肌营养不良或贝克肌营养不良的方法,所述方法包括施用以下(e)和(f):

[0270] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

[0271] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA,或编码所述向导RNA的多核苷酸。

[0272] 本发明还提供以下(e)和(f)在制备用于治疗或预防杜氏肌营养不良或贝克肌营养不良的药物组合物中的用途:

[0273] (e) 核酸酶缺陷型CRISPR效应蛋白和转录激活子的融合蛋白,或编码所述融合蛋白的多核苷酸,和

[0274] (f) 靶向人肌营养相关蛋白基因的表达调节区中SEQ ID NO:104、105、135、141、153、167或172所示区域中长度为18至24个核苷酸的连续区域的向导RNA,或编码所述向导RNA的多核苷酸。

[0275] 作为本发明中的核酸酶缺陷型CRISPR效应蛋白、转录激活子、向导RNA以及编码它们的多核苷酸和携带它们的载体,可以使用在上述“1.多核苷酸”、“2.载体”和“5.核糖核蛋白”部分中详细说明的那些。上述(e)和(f)的剂量、施用途径、对象、制剂等与在“3.DMD或BMD的治疗或预防剂”中所述的相同。

[0276] 在以下示例性实施方式的描述过程中,本发明的其它特征将变得显而易见,所述示例性实施方式是为了说明本发明而给出的,而并不旨在限制本发明。

[0277] 实施例

[0278] 这些实施例描述了在人UTRN基因的指定的表达调节区中,使用dCas9与转录激活子的融合蛋白来增强基因表达,从而使得人UTRN基因表达被选择性激活。基因表达修饰的目的是增强野生型人UTRN基因产物的表达,补充了缺陷型肌营养不良蛋白基因产物的功能。实施例还描述了赋予人UTRN基因的选择性激活而对其它基因的表达影响最小的特异性基因组区域的定义。如本文所述和所示,本发明增强人UTRN基因表达的方法代表了一种新的治疗或预防策略,用于改善由缺陷型肌营养不良蛋白引起的缺陷型肌肉功能。

[0279] 实施例1.使用HEK293FT细胞筛选针对人肌营养相关蛋白基因的gRNA

[0280] 在这个实施例中,我们举例说明了使用本文所述的方法通过靶向UTRN基因的指定的表达调节区来实现UTRN基因的激活。所述方法通过设计合适的sgRNA序列,利用Cas9和sgRNA的复合物的性质,以将其募集到基因组的所需基因座。所述方法还利用SaCas9蛋白的核酸酶缺陷形式(D10A和N580A突变体(SEQ ID NO:107)或D10A和H557A突变体(SEQ ID NO:108);dSaCas9)来保持基因组序列完整,但将各种转录/表观遗传功能性结构域或基序拴系至dSaCas9,以实现sgRNA序列所靶向的预期基因座的所需修饰,如Gilbert LA等人,《细胞》,2013年7月;154(2):442-51,以及Gilbert LA等人,《细胞》,2014年10月;159(3):647-61所述,通过引用全部并入本文。

[0281] 在这个实施例中,我们说明了本文所述的方法可以用于激活野生型UTRN的表达。sgRNA被设计成靶向UTRN基因的表达调节区,赋予选择性和有效的基因激活。图1示出了人UTRN基因座和两个预测的转录起始位点(TSS)(上图和中图)。UTRN基因的TSS通过查询FANTOM5人promoterome数据库来鉴定(www.fantom.gsc.riken.jp,《自然》2014年3月;507(7493):462-70,通过引用全部并入本文)。报告了UTRN基因的两个启动子区域(启动子A和B),我们已经测试了这两个启动子的激活情况。向导RNA序列被设计成覆盖以上区域,以便确定这些区域内的有效和选择性的治疗序列。

[0282] (1) 实验方法

[0283] sgRNA序列的选择

[0284] 查看UTRN基因启动子区域周围的序列(启动子A,约4.4kb(Chr6:GRCh38/hg38;144,283,900-144,288,300)和启动子B,约2.6kb(Chr6:GRCh38/hg38;144,342,683-144,345,311)),以寻找dSaCas9和sgRNA复合体会结合的潜在识别序列。查看所述区域以寻找具有序列NNGRRT的原间隔子相邻基序(PAM)。鉴定了与PAM相邻的靶向序列。靶向序列(与靶DNA杂交的一部分gRNA)的长度被设为21个核苷酸。根据Benchling软件(<https://benchling.com>)生成的预测特异性和效率选择靶向序列,且其均匀分布在所选区域中。还参考了来自ENCODE研究(ENCODE项目联盟,《自然》,2012年9月;489:57-74,通过引用全部并入本文)中UTRN表达调节区周围的表观遗传信息,以选择与所述基因的功能元件结合的可能性高的gRNA。

[0285] 测试了表1中列出的24种靶向序列(向导编号sgED3-1至sgED3-24(SEQ ID NO:129至152))对UTRN基因表达的调节功能(下文中,表1中列出的靶向序列有时称为“sgED3系列”。

[0286] UTRN基因中靶向序列的位置也在图1中示出(上图和中图)。

[0287] 将选择的24种靶向序列和对照非靶向性靶向序列(SEQ ID NO:177)分别与编码tracrRNA的DNA序列(SEQ ID NO:122)融合以形成sgRNA序列,并将其克隆到来自Genecopoeia的pCRISPR-LvSG03载体(#pCRISPR-LvSG03)中。所获得的载体在本说明书中表示为pCRISPR-LvSG03 sgRNA表达载体。sgRNA的表达受U6启动子驱动,所述载体在SV40启动子下表达mCherry-IRES-嘌呤霉素基因,从而有助于追踪和选择sgRNA表达细胞。

[0288] 效应分子的克隆

[0289] 核酸酶缺陷型SaCas9蛋白(D10A和N580A,或D10A和H557A;dSaCas9)充当以直接融合蛋白形式拴系功能性结构域/基序的主要骨架。dSaCas9连接了两个核定位信号(NLS)在

其N末端(SEQ ID NO:178所示的氨基酸序列,SEQ ID NO:179所示的DNA序列)和C末端(SEQ ID NO:180所示的氨基酸序列,SEQ ID NO:181所示的DNA序列),以使效应分子能够有效定位至细胞核。

[0290] 在一个实施例中,将编码具有D10A和N580A突变的dSaCas9的DNA序列在其C末端与编码合成氨基酸转录激活部分VP64、VPH或VPR的DNA序列融合(参见Chavez A等人,《自然方法学》,2016年7月;13(7):563-67和Chavez A等人,《自然方法学》,2015年4月;12(4):326-8,通过引用全部并入本文)(SEQ ID NO:182、183或184)。所获得的融合蛋白在本说明书中分别表示为dSaCas9-VP64、dSaCas9-VPH或dSaCas9-VPR融合蛋白。

[0291] 融合蛋白被募集到UTRN基因的表达调节区,从而发挥其转录激活作用。结果,增强了UTRN基因的表达。

[0292] 在一个实施例中,将编码缺少氨基酸721-745的dSaCas9蛋白(dSaCas9[-25]),(SEQ ID NO:214))的DNA序列在其C端与编码合成氨基酸转录激活子miniVR的DNA序列融合(参见PCT/JP2019/030972,通过引用全部并入本文)(SEQ ID NO:185)。所获得的融合蛋白在本说明书中表示为dSaCas9[-25]-miniVR融合蛋白(SEQ ID NO:186)。

[0293] 为了表达dSaCas9-VP64、dSaCas9-VPH、dSaCas9-VPR和dSaCas9[-25]-miniVR融合蛋白,将编码所述融合蛋白的DNA片段克隆到来自Genecopoeia的CP-LvC9NU-09慢病毒表达载体(目录号CP-LvC9NU-09)中。用融合蛋白编码序列替换原始载体中的Cas9编码序列,从而产生包含编码dSaCas9-VP64、dSaCas9-VPH、dSaCas9-VPR或dSaCas9[-25]-miniVR四种融合蛋白之一的DNA片段的CP-LvC9NU-09慢病毒表达载体。在本说明书中,所得载体分别表示为CP-LvdSaCas9-VP64-09、CP-LvdSaCas9-VPH-09、CP-LvdSaCas9-VPR-09或CP-LvdSaCas9[-25]-miniVR-09质粒。所述载体使用EF1a启动子表达效应分子,并使用SV40启动子表达eGFP-IRES-新霉素基因。

[0294] 为了在腺相关病毒载体中表达,将编码dSaCas9[-25]-miniVR融合蛋白的DNA片段、U6启动子和sgRNA克隆到来自Takara的pAAV-CMV载体(#6234)中。将CMV启动子替换为EFS启动子(SEQ ID NO:187)。从原始载体中除去 β -珠蛋白内含子,并将hGH poly-A替换为牛GH polyA(bGH polyA)。所获得的载体从其5'末端到其3'末端依次包含ITR、EFS启动子、dCas9、miniVR、bGH polyA、U6启动子、sgRNA和ITR(图5),在本说明书中表示为pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒。

[0295] 细胞培养和转染

[0296] 转染前24小时,将HEK293FT细胞(Thermo Fisher#R70007)以每孔75,000个细胞的密度接种到24孔板(CORNING#351147)中,并在补充有10%FBS和2mM新鲜L-谷氨酰胺、1mM丙酮酸钠和非必需氨基酸的DMEM培养基(Thermo Fisher#11140050)中培养。为了在慢病毒表达载体中表达,根据制造商的说明,细胞用500ng CP-LvdSaCas9-VP64-09、CP-LvdSaCas9-VPH-09、CP-LvdSaCas9-VPR-09或CP-LvdSaCas9[-25]-miniVR-09质粒和500ng pCRISPR-LvSG03 sgRNA表达载体,使用1.5 μ l Lipofectamine 2000(Life Technologies#11668019)转染。用嘌呤霉素(1 μ g/ml)选择转染的细胞。为了在腺相关病毒载体中表达,根据制造商的说明,细胞用500ng pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒,使用1.5 μ l Lipofectamine 2000(Life Technologies#11668019)转染。不用嘌呤霉素选择转染的细胞。

[0297] 为了进行基因表达分析,将转染的细胞在37℃下5%CO₂中培养,并在转染后72h收获,并在RLT缓冲液(Qiagen#74104)中裂解,以使用RNeasy试剂盒(Qiagen#74104)提取总RNA。

[0298] 基因表达分析

[0299] 为了进行Taqman分析,使用TaqMan™高容量RNA转cDNA试剂盒(Applied Biosystems#4387406),以20μl的体积使用1.5μg的总RNA生成cDNA。用水将生成的cDNA稀释20倍,每个Taqman反应使用6.33μl。UTRN和HPRT基因的Taqman引物和探针获自Applied Biosystems。使用Taqman基因表达预混液(Thermo Fisher#4369016)在Roche LightCycler 96或LightCycler 480中运行Taqman反应,并使用LightCycler 96分析软件进行分析。UTRN基因的表达水平通过HPRT基因的表达水平进行标准化。

[0300] Taqman探针产品ID:

[0301] UTRN:HS01125994_m1 (FAM)

[0302] HPRT:Hs99999909_m1 (FAM,VIC)

[0303] Taqman QPCR条件:

[0304] 步骤1:95℃,10分钟

[0305] 步骤2:95℃,15秒

[0306] 步骤3:60℃,30秒

[0307] 重复步骤2和3:40次

[0308] 腺相关病毒(AAV)的产生

[0309] 腺相关病毒血清型2(AAV2)粒子是使用在150mm培养皿(Corning)中以每个培养皿9,000,000个细胞的密度接种,并在补充有10%FBS、2mM新鲜L-谷氨酰胺、1mM丙酮酸钠和非必需氨基酸的DMEM培养基(Thermo Fisher#11140050)中培养的AAVpro 293T细胞(Takara#632273)生成的。细胞用14.85μg pRC2-mi342和pHelper载体(Takara#6234)和含有81μl TransIT-VirusGen(Mirus Bio#MIR6703)的14.85μg pAAV-EFS-dsaCas9[-25]-miniVR-U6-sgRNA AIO质粒转染。72小时后,根据制造商在AAV2无辅助系统方案(Takara#6230)中的说明,以每个150mm培养皿550μl收获细胞并提取AAV2粗品。

[0310] 用AAV2转导细胞

[0311] 为了转导HEK293FT细胞(Thermo Fisher#R70007),将每孔75,000个细胞接种到24孔板(CORNING#351147)中,并在补充有10%FBS和2mM新鲜L-谷氨酰胺、1mM丙酮酸钠和非必需氨基酸的DMEM培养基(Thermo Fisher#11140050)中孵育16小时。将培养基替换为1000μl新鲜培养基,包含10μl或1μl(分别为1:100或1:1000稀释)的AAV2粗品。在随后的72h孵育后,将细胞裂解,并根据制造商的说明(Qiagen#74192)提取总RNA(RNeasy Plus 96试剂盒),并如Taqman在“基因表达分析”中所述确定肌营养相关蛋白的过表达。

[0312] 表1.用于筛选UTRN基因的表达调节区的靶向序列

[0313] 表1-1.

[0314]

SEQ ID NO.	向导编号	位置	链	靶向序列	PAM	特异性	效率
129	sgED3-1	144283943	1	CTTGTTAAATGAATGAATGAA	GTGAAT	21.67	24.07
130	sgED3-2	144284051	1	TGTCCTAGAAACCTTACAAGG	AAGAGT	81.73	47.49
131	sgED3-3	144284216	-1	GGTTTATTGCTGGCTTAATAT	TTGAGT	73.32	27.89
132	sgED3-4	144284644	1	ACGTCAGCAAACCTGAGATGGG	GTGAGT	72.28	29.99
133	sgED3-5	144284753	1	TTTTCGGATAATCTGAATAAG	GGGAAT	73.39	26.71
134	sgED3-6	144285129	-1	GGGGTCCGCTCTCCAGATGAG	AAGGGT	86.65	25.53
135	sgED3-7	144285744	-1	GGCTCCTCTAGGAGTTTGACA	CGGAGT	88.25	85.15
136	sgED3-8	144285873	1	TAATGTGACTACAGCCCCGA	GGGAAT	93.52	70.61
137	sgED3-9	144285972	1	CCAAGTCCCAGAGTCGAAGAT	GGGAGT	92.21	44.26
138	sgED3-10	144286550	-1	TCAGTTGCAGCAAGAGATCCC	CAGAGT	82.58	26.26
139	sgED3-11	144286736	-1	CCTCCTCCTCGAAAAACGCAC	TGGAAT	90.03	64.99
140	sgED3-12	144287009	-1	GGGAGGGTCGGCTCAGACCTA	GGGAAT	91.68	30.46
141	sgED3-13	144287207	1	GGGTAGTTCTGCGGTGACGGA	CAGGGT	92.71	23.34
142	sgED3-14	144287288	-1	ATTTTAGGTAAACACCCAAAG	GAGAGT	70.86	46.61
143	sgED3-15	144287397	-1	GAAACACAGTAAAAGAAACG	GTGAGT	51.32	53.15
144	sgED3-16	144287614	-1	TAAGATTTTAGGAATTATACA	ATGAAT	50.22	34.45

[0315] 表1-2.

145	sgED3-17	144287760	1	AGCGTTCTGAAGGGAGAGTTA	GTGAAT	75.62	42.44
146	sgED3-18	144287920	-1	CAGAAGGCTAGGTGAGAACT	GAGAAT	64.29	34.34
147	sgED3-19	144288078	-1	AATTTGAGTACACTTAAGGCA	AAGGAT	74.85	24.36
148	sgED3-20	144288193	-1	AGATACAGCAGAAAAGGTGAT	CAGAGT	59.61	52
149	sgED3-21	144343311	1	GACACATGCAGAAGTGACAGC	AGGAGT	62.51	64.83
150	sgED3-22	144344138	-1	AGCAGCCTTCGAACTGCACAC	TGGGAT	85.61	69.44
151	sgED3-23	144344637	-1	TCTAGATGGCAGTAAACAGCA	CAGAGT	72.98	81.01
152	sgED3-24	144345218	-1	GGCTGCTCCAATCATTTTGGT	TTGAAT	79.1	56.17
153	sgED3-25	144284787	-1	GAGTCCGGAGACCGAACCAGA	ATGGAT	91.54	23.9
154	sgED3-26	144284810	-1	GAACCGTGCGTGCCGGGAGCC	GGGAGT	86.09	1
155	sgED3-27	144284837	-1	GCTGGCCTGGGGCGCGCGCTC	CAGAGT	78.51	0.56
156	sgED3-28	144285003	-1	AAGATCAGCCCCACTACGTTC	CCGGGT	94.71	15.9
157	sgED3-29	144285172	1	CCGGAGGCGAGCCCCCTTCCCG	GGGGGT	82.7	14.59
158	sgED3-30	144285207	-1	GGAGGGTGGGGCGCAGGACCG	CTGGGT	68.11	4.19
159	sgED3-31	144285227	-1	GAGCGCTGGAGGCGGAGGAGG	GAGGGT	40.4	5.54
160	sgED3-32	144285325	1	CCTCTCTCGCGCACAAAGTTG	TGGAGT	92.3	13.5
161	sgED3-33	144285460	1	GGGAGCGCGCCCCCTTCTT	TTGGGT	92.82	3.72
162	sgED3-34	144285496	-1	CACCAACTTTGCCAAACGCTA	CAGAGT	90.69	15.32
163	sgED3-35	144285722	-1	GGAGTAACCGCGGGGGTGTGT	GCGAGT	90.76	15.84
164	sgED3-36	144285896	1	GAATGGGGCGGGGCGGGGAG	GAGGAT	47.73	3.79

[0316] 表1—3.

165	sgED3-37	144285926	1	TCTTCTGTGGTCTTCCGCC	TGGGAT	81.43	25.49
166	sgED3-38	144286089	-1	TTTGATCGTTCACAACACTAGT	ACGGAT	92.05	18.73
167	sgED3-39	144286240	1	AGAGGGGACGTGGCCTCTTAG	GAGAGT	83.03	23.82
168	sgED3-40	144286311	1	GTCCACAGGAGAGGGTGGGCA	GAGGGT	38.6	9.03
169	sgED3-41	144286418	1	GCTCCCAAGGGTGGGGCTCCG	GAGAGT	75.62	5.83
170	sgED3-42	144286683	1	TTTCAGATGGCAGGTTGTTCA	AAGGAT	84.92	0.55
171	sgED3-43	144286895	1	CTTCCACGCCCTTCAGGTCAG	CCGGAT	70.16	23.24
172	sgED3-44	144286993	1	GCGCGCGGAGCTCGGGGGAGG	CCGGAT	58.97	0.54
173	sgED3-45	144287068	-1	TGAGGCCGGTGCAACTTACAA	AGGAAT	94	33.46
174	sgED3-46	144287139	1	TGGGCGTGGGAGACGCAGCCT	GCGGAT	73.4	1.47
175	sgED3-47	144287184	1	AGGTGGAGGAATGCGAAGCTT	GTGGGT	87	21.29
176	sgED3-48	144287284	1	AGACAACTCTTTAACTCTCCT	TTGGGT	78.9	15.46

[0317] 在表1中,“位置”表示当使用SaCas9时,链中存在靶向序列的核苷酸的切割位置。

[0318] 在表1的“链”项中,1表示有义链,-1表示反义链。

[0319] (2) 结果

[0320] 图1示出了与对照sgRNA相比,三种不同的dSaCas9-激活子融合蛋白(dSaCas9-VP64、dSaCas9-VPH和dSaCas9-VPR)对UTRN基因表达的激活。对照sgRNA包含ACGGAGGCUAAGCGUCGCAAG(SEQ ID NO:215)和tracrRNA序列,且被设计为它在人基因组中的任何序列上均没有靶标。包含分别由向导编号sgED3-6、sgED3-7或sgED3-13(SEQ ID NO:134、135或141)编码的crRNA的sgRNA通过募集dSaCas9-激活子融合蛋白到UTRN基因的表达调节区来激活UTRN基因的表达。dSaCas9-VPR融合蛋白的激活作用最强。

[0321] 根据以上结果,由向导编号sgED3-6至sgED3-7(SEQ ID NO 134至135)(表1)覆盖的对应于Chr6:GRCh38/hg38;144,285,000-144,286,000(图1)的约1.0kb区域(区域A)以及位于向导编号sgED3-13(SEQ ID NO 141)周围的对应于Chr6:GRCh38/hg38;144,287,000-144,287,300的约0.3kb区域(区域B)赋予了UTRN基因表达的有效激活。与向导编号sgED3-6、sgED3-7和sgED3-13(SEQ ID NO:134、135和141)编码的crRNA相比,启动子B赋予的UTRN基因的激活相对较弱。

[0322] 在图2中,进一步用另外的24种sgRNA(表1,向导编号sgED3-25至sgED3-48(SEQ ID No:153至176))和dSaCas9-VPR筛选跨越区域A和B的对应于Chr6:GRCh38/hg38;144,284,750-144,287,300的区域,以发现更有效的UTRN基因激活。包含分别由向导编号sgED3-6、sgED3-13、sgED3-25至sgED3-32、sgED3-39、sgED3-40和sgED3-44(SEQ ID NO:134、141、153至160、167、168和172)编码的crRNA的sgRNA激活的UTRN基因表达是上述对照sgRNA的两倍以上。

[0323] 在图3中,对于一些包含分别由向导编号sgED3-6、sgED3-13、sgED3-25、sgED3-27、sgED3-30、sgED3-31、sgED3-39、sgED3-40和sgED3-44(SEQ ID NO:134、141、153、155、158、159、167、168和172)编码的crRNA的有效sgRNA,分别制备pAAV-EFS-dSaCas9[-25]-miniVR-U6-sgRNA A10质粒,然后转染到HEK293FT细胞中以验证功能。与上述对照sgRNA相比,观察到不同sgRNA对UTRN基因不同程度的诱导。

[0324] 在图4中,制得携带EFS-dSaCas9[-25]-miniVR-U6-sgRNA的AAV2,并转导HEK293FT细胞。作为sgRNA,使用包含分别由向导编号sgED3-6、sgED3-30或sgED3-31(SEQ ID NO:134、158或159)编码的crRNA的sgRNA。对于所有三种sgRNA,与上述对照sgRNA相比,观察到了UTRN基因诱导。

[0325] 实施例2.使用HSMC细胞筛选针对人肌营养相关蛋白基因的gRNA

[0326] (1) 实验方法

[0327] UTRN靶向序列的选择

[0328] 根据人骨骼肌细胞中基因组的H3K4me3和H3K27Ac图谱,查看人UTRN基因推定的增强子(称为E)和启动子(称为P)区域周围大约13.2kb的序列,以寻找可以被与gRNA复合的核酸酶缺陷型SaCas9(D10A和N580A突变体;dSaCas9[SEQ ID NO:123(蛋白质)])靶向的序列,在本文中定义为靶向序列。所靶向的基因组区域相对于UTRN基因的位置如图6所示,其坐标如下所示:

[0329] 1.Chr6:GRCh38.p12;144215500-144217000->约1.5kb(称为P2)

[0330] 2.Chr6:GRCh38.p12;144248500-144249800->约1.3kb(称为E1)

[0331] 3.Chr6:GRCh38.p12;144264000-144267000->约3.0kb(称为E2)

[0332] 4.Chr6:GRCh38.p12;144283900-144288300->约4.4kb(称为P1)

[0333] 5.Chr6:GRCh38.p12;144292500-144295500->约3.0kb(称为E3)

[0334] 靶向序列是由与具有序列NNGRRT的原间隔子相邻基序(PAM)相邻的21个核苷酸的片段指定(5'-21nt靶向序列-NNGRRT-3'),并经过过滤以包括大部分完美匹配食蟹猴(*Macaca fascicularis*)基因组的相应区域(靶向序列和PAM序列)的序列(在表3中列为“TRUE”)。

[0335] 慢病毒转移质粒(pED176)的构建

[0336] pLentiCRISPR v2是从Genscript(<https://www.genscript.com>)购得,并进行了以下修饰:SpCas9 gRNA骨架序列替换为SaCas9 gRNA骨架序列(SEQ ID NO:124);SpCas9替换为与密码子优化的VP64-迷你型miniRTA(也称为迷你型miniVR)[SEQ ID NO:125(DNA)和126(蛋白质)]融合的dSaCas9。迷你型miniVR转录激活结构域可以通过激活转录来激活基因表达。迷你型miniVR拴系在dSaCas9(D10A和N580A突变体)的C末端,下文中称为dSaCas9-迷你型miniVR(SEQ ID NO:192(DNA)和193(蛋白质)),并由在包含各靶向序列编码的crRNA的gRNA指导下靶向人UTRN基因的推定的增强子或启动子区域(图6)。所产生的骨架质粒被命名为pED176。

[0337] gRNA克隆

[0338] 将三种对照非靶向性靶向序列(表3,SEQ ID NO:1至3)和100种靶向序列(表3,SEQ ID NO:4至103)克隆到pED176中。正向和反向寡核苷酸是由Integrated DNA Technologies按以下格式合成:正向:5'CACC(G)-20至21个碱基对的靶向序列-3',以及反向:5'AAAC-20至21个碱基对的反向补体靶向序列-(C)-3',其中如果靶标不以G开头,则添加括号中的碱基。将寡核苷酸以100μM再悬浮于Tris-EDTA缓冲液(pH 8.0)中。在NE Buffer 3.1(New England Biolabs(NEB)#B7203S)中,在10μl反应物中混合1μl各互补寡核苷酸。将反应物加热至95℃,并在热循环仪中冷却至25℃,由此使寡核苷酸退火成具有与克隆至pED176相容的粘性突出末端。将退火的寡核苷酸与已用BsmBI消化并经凝胶纯化的慢病毒转移质粒pED176组合,并根据制造商的方案用T4 DNA连接酶(NEB#M0202S)连接。根据制造商的方案,将2μl的连接反应物转化到10μl的NEB稳定感受态细胞(NEB#C3040I)中。所得的构建体通过U6启动子(SEQ ID NO:128)驱动包含3'末端与tracrRNA(guuuuaguacucuggaaacagaaucuaacuaaacaaggcaaaaugccguguuuauaucugcaacuuguuggcgagauuuuuu(SEQ ID NO:127))融合的

由各个靶向序列编码的crRNA的sgRNA的表达,所述tracrRNA由添加有U6聚合酶TTTTTT末端信号的SaCas9 gRNA骨架序列编码。

[0339] 慢病毒的产生

[0340] 将HEK293TA细胞 (Genecopoeia#LT008) 以 0.75×10^6 个细胞/孔接种在6孔细胞培养皿 (VWR#10062-892) 中的2ml生长培养基 (补充有10%FBS和2mM新鲜L-谷氨酰胺、1mM丙酮酸钠和非必需氨基酸的DMEM培养基 (Thermo Fisher#11140050)) 中,并在37℃/5%CO₂下孵育24小时。第二天,根据制造商的方案,使用1.5μg封包质粒混合物[1μg封包质粒 (参见pCMV delta R8.2; addgene#12263) 和0.5μg包膜表达质粒 (参见pCMV-VSV-G; addgene#8454)] 和1μg含有编码dSaCas9-miniVR和指定sgRNA的序列的转移质粒设置TransIT-VirusGEN转染反应 (Mirus Bio#MIR6700)。转染后48小时,通过使培养基上清液通过0.45μm PES过滤器 (VWR#10218-488) 收获慢病毒。在准备使用之前,将经纯化且等分的慢病毒保存在-80℃的冰箱中。

[0341] HSMM细胞的转导

[0342] 如表2所示,来自5位年龄在0-35岁之间的不同人供体的原代骨骼肌成肌细胞 (HSMM) 获自Lonza Inc。

[0343] 表2.

[0344]

供体编号	批号	年龄(岁)	性别
1	650386	35	男
2	657512	34	女
3	542368	0	女
4	629287	19	女
5	655307	18	男

[0345] 细胞在来自Lonza的原代骨骼肌细胞生长培养基[SkGM™-2骨骼肌生长BulletKit培养基 (#CC-3245), 所述培养基包含有含骨骼肌成肌细胞生长所需的SkBM™-2基础培养基 (#CC-3246) 和SkGM™-2SingleQuots™补充剂 (#CC-3244) 的培养系统] 中培养。CC-3246包含1x SkBM™-2基础培养基, 500mL。1x SkGM™-2SingleQuots™补充包 (#CC-3244) 包含:

[0346] 1个红盖小瓶的GA-1000, 0.50mL

[0347] 1个绿盖小瓶的hEGF, 0.50mL

[0348] 1个天然盖小瓶的地塞米松, 0.50mL

[0349] 1瓶FBS, 50.00mL

[0350] 1瓶L-谷氨酰胺, 10.00mL

[0351] 根据制造商的说明, 将CC-3244的成分添加到500ml培养基 (#CC-3246) 中。

[0352] 为了进行转导, 将细胞以 $0.125-0.33 \times 10^6$ 个细胞/孔接种在含有生长培养基的6孔细胞培养皿 (VWR#10062-894) 中, 并在37℃/5%CO₂下孵育24小时。第二天, 将补充有8μg/ml聚凝胺的1.5ml生长培养基 (Sigma#TR-1003-G) 和1.0ml对应于包含与tracrRNA融合的由各个靶向序列 (表3) 编码的crRNA的各sgRNA的慢病毒上清液 (见上文) 添加到各孔中。通过使用Lenti-X™ qRT-PCR滴定试剂盒 (Clontech#631235) 测量, 慢病毒滴度范围为 10^8 至 10^9 个粒子/毫升。将细胞与慢病毒一起孵育6小时之后, 去除病毒培养基并替换为新鲜的生长培养基。转导后72小时, 向细胞供以选择培养基[补充有0.5μg/ml嘌呤霉素的生长培养基

(Sigma#P8833-100MG)]。每2至3天为细胞提供新鲜的选择培养基。将细胞置于选择培养基中7至10天后,收获细胞,并按照制造商的指导,使用RNeasy 96试剂盒(Qiagen#74182)提取RNA。

[0353] 以相同方式进行两种病毒的共转导实验,病毒总量与单一病毒转导相等。

[0354] 基因表达分析

[0355] 为了进行基因表达分析,根据高容量cDNA逆转录试剂盒(Applied Biosystems; Thermo Fisher#4368813)方案,以10 μ l的体积从约0.05至0.8 μ g的总RNA产生cDNA。将cDNA稀释10倍,并使用Taqman快速高级预混物(Thermo Fisher#4444557),根据制造商的方案进行分析。Taqman探针(UTRN:测定Id Hs01125994_m1 FAM;HPRT:测定Id Hs99999909_m1 VIC_PL)从Life Technologies获得。按照Taqman快速高级预混物方案的指导,由QuantStudio 5实时PCR系统对基于Taqman探针的实时PCR反应进行处理和分析。

[0356] 数据分析

[0357] 对于各样品和三个对照,通过将UTRN探针的3个技术重复样品的平均Ct值减去HPRT探针的3个技术重复样品的平均Ct值(平均Ct UTRN-平均Ct HPRT)来计算 Δ Ct值。使用式 $2^{-\Delta Ct}$ 确定各样品的表达值。然后将各实验的样品表达值(表3;SEQ ID NO:4至103)相对于3个对照表达值(表3;SEQ ID NO:1至3)的平均值标准化,以确定各样品的相对UTRN表达。分析每次筛选的两个生物学重复样品,并计算所有实验的平均值(表3)。

[0358] (2) 结果

[0359] RNP对UTRN基因表达的激活

[0360] 产生慢病毒,其将针对各靶向序列的dSaCas9-miniVR和sgRNA的表达盒递送给来自5个不同供体的原代HSMM细胞(表2)。由于细胞的生长速度,大多数测定是对来自供体#3的HSMM细胞(表2)进行的。针对嘌呤霉素抗性对转导细胞进行选择,且使用Taqman测定法对UTRN表达进行定量(表3)。将各样品的表达值相对于用对照sgRNA(表3;SEQ ID NO:1至3)转导的细胞中的UTRN表达的平均值标准化。对供体#3的重复样品测量平均表达水平(表3;和图7)。

[0361] 表3. 用于筛选UTRN基因的表达调节区的靶向序列。

[0362] 表3-1.

[0363]

SEQ ID	向导编号	坐标(hg38/Chr. 6)	核苷酸长度	有义或反义链	序列	食蟹猴匹配	HSMMd3 筛选 1	HSMMd3 筛选 2
1	对照 X3	NA	20	-	ACGGAGGCTAAGCGTCGCAA	-	1	1
2		NA	20	-	CGCTTCCGCGGCCCGTTCAA	-		
3		NA	20	-	GTAGCGCGCCGCTCTCTAC	-		
4	12	144216047	21	1	AGAAAAGCGGCCCTAGGGGC	TRUE	1.42	0.55
5	16	144216199	21	-1	CAAACACACACCAGCAAACCT	TRUE	1.27	0.53
6	17	144216257	21	1	TGAAAGCGCAACTGGAGGGCC	TRUE	0.99	0.73
7	24	144216593	21	-1	ACCCACGCGGACATATGTCCA	TRUE	0.82	0.59
8	25	144216602	21	1	ATCCAATGGACATATGTCCGC	TRUE	1.41	0.68
9	31	144216855	21	1	GAGGGGGAGGGCTGTGACCTG	TRUE	1.35	0.56
10	34	144248644	21	-1	ATTGGTGGTCAGGGAGCAAG	TRUE	1.71	0.57
11	35	144248677	21	1	AATGAAACCAAAGACAGCTTC	TRUE	1.32	0.51
12	44	144248973	21	-1	CCAAAATCCTTTAATGAATCA	TRUE	1.43	0.65
13	45	144248977	21	1	TACAGATTCCATGATTCATTA	TRUE	1.58	0.59
14	46	144248981	21	-1	GGAACAAACCAAATCCTTTA	TRUE	1.37	0.69
15	48	144249031	21	-1	ATCTGTTTGTGGGGAATCTT	TRUE	1.21	0.77
16	49	144249058	21	1	CAACAGATTTCAGTATTTTC	TRUE	1.41	0.64
17	51	144249159	21	1	GTGGTGATTTATGTTACTGGT	TRUE	1.18	0.77
18	52	144249181	21	1	TGAGTCTTTCAAGTTCCTTTC	TRUE	1.5	0.72
19	53	144249211	21	1	AGATCATTTTGGCTTCAAAC	TRUE	1.63	0.71
20	54	144249221	21	1	TGGGTTCAAAC TAGAATGTCC	TRUE	1.93	0.72
21	56	144249311	21	1	GATCTATCTATAGACACCAAA	TRUE	1.33	0.54
22	61	144249393	21	-1	TGCTTCTTCCAGGCTTGAGTG	TRUE	1.39	0.75
23	62	144249400	21	-1	ACCGCTTTGCTTCTTCCAGGC	TRUE	0.96	0.71
24	63	144249413	21	1	AAGCCTGGAAGAAGCAAAGCG	TRUE	1.51	0.98
25	73	144249669	21	-1	cttctgaatcagaattcctaa	TRUE	1.04	0.66

[0364] 表3-2.

[0365]

26	78	144249756	21	1	TGGTTCCAAGCTAGTACTTCA	TRUE	1.03	0.78
27	80	144264074	21	1	ATGTTACAAAAATAATTTAA	TRUE	0.99	0.75
28	86	144264238	21	1	CCTTTATGGTCACCTTCTCTG	TRUE	1.2	0.79
29	87	144264250	21	1	CCTTCTCTGCTGAGTAAAAAT	TRUE	1.09	1.07
30	88	144264297	21	1	AAGGTGGCCAAAAAGAACCC	FALSE	1.28	1.37
31	91	144264318	21	-1	AAGGAAGAGAGAGGCCAAGAA	TRUE	1.43	1.04
32	95	144264449	21	-1	TAAAGAATTCTAGCACTGGAA	TRUE	0.62	0.51
33	106	144264745	21	1	AAATGTGTCATGTGTTGGTTA	TRUE	0.89	0.8
34	114	144265048	21	1	AAAAATGAAATTGCAACTTC	TRUE	1.01	0.65
35	115	144265058	21	1	ATTGCAACTTCTAGAATTTAA	TRUE	0.79	0.57
36	121	144265214	21	1	CAGCTGGAGTGGGCCACGTAA	TRUE	1.19	1.25
37	123	144265304	21	-1	ATTTTTCATATTTCTTTGGT	TRUE	1.14	0.64
38	125	144265450	21	-1	AGTGACCTGCTGATTTCTCTA	TRUE	1.38	0.74
39	127	144265606	21	1	CTTTCCCATTTGTTCAGGACT	TRUE	1.1	0.82
40	135	144265764	21	1	TTGGTTGATAAATTTGTATAT	TRUE	1.41	0.82
41	136	144265795	21	-1	TCTCTAGTTCATTTTITAGCT	TRUE	1.17	0.82
42	139	144266101	21	-1	TCCTTCAACTTCAAGACAACA	TRUE	0.87	0.66
43	140	144266147	21	-1	GCTCCTCCTGCTGGATGGGGG	TRUE	1.36	0.8
44	141	144266158	21	-1	CTCTATTTCAGCTCCTCCTG	TRUE	1.08	0.86
45	145	144266243	21	1	GTACAGTTAGTGCTACTAGGA	TRUE	3.2	1.43
46	146	144266254	21	1	GCTACTAGGACAGGATGCTGG	TRUE	2.5	1.23
47	148	144266287	21	-1	CCCCAGCTGTGCCTCTGTTTT	TRUE	1.42	0.72
48	149	144266297	21	1	TTCCCAAAACAGAGGCACAGC	TRUE	1.36	0.87
49	151	144266338	21	-1	GTTTTGAAACTGGTAGCAGCT	TRUE	1.52	1.2
50	175	144283934	21	1	aaactgatgcttgttaaatga	TRUE	1.05	0.91
51	176	144283943	21	1	cttggttaaatgaatgaatGAA	TRUE	1.34	0.89
52	178	144283973	21	-1	AATCCAAAGGATTAAGTTGAA	TRUE	1.48	1.09
53	179	144283981	21	1	TACCCATTTCAGTTAATCCT	TRUE	1.3	1.02
54	183	144284099	21	1	TGCCCCCTCCCTGGAGCACTT	TRUE	1.39	0.65
55	192	144284640	21	1	AGCAACGTCAGCAAACCTGAGA	TRUE	1.06	0.96

表3-3.

[0366]

56	193	144284644	21	1	ACGTCAGCAAACCTGAGATGGG	TRUE	1.29	0.63
57	202	144284810	21	-1	GAACCGTGCCTGCCGGGAGCC	TRUE	1.15	0.95
58	205	144285129	21	-1	GGGGTCCGCTCTCCAGATGAG	FALSE	2.28	1.8
59	208	144285207	21	-1	GGAGGGTGGGGCGCAGGACCG	TRUE	2.29	1.59
60	210	144285325	21	1	CCTCTCTCGCGCACAAAGTTG	FALSE	1.88	1.76
61	211	144285429	21	1	TCTGGCTCCAGAAGCCGATTG	TRUE	1.11	1.01
62	214	144285603	21	1	ACAAGTAAGGGGCGTTTTCAG	TRUE	1.14	0.78
63	218	144285756	21	-1	GAGCTGGCCAAGGGCTCCTCT	TRUE	1.3	0.82
64	219	144285770	21	1	TAGAGGAGCCCTTGGCCAGCT	TRUE	1.34	0.86
65	224	144285972	21	1	CCAAGTCCCAGAGTCGAAGAT	TRUE	1.25	0.87
66	234	144286311	21	1	GTCCACAGGAGAGGGTGGGCA	TRUE	1.38	0.87
67	236	144286403	21	1	CTCTGGGTGGTTGCTGCTCCC	TRUE	1	0.73
68	239	144286550	21	-1	TCAGTTGCAGCAAGAGATCCC	TRUE	1.12	0.92
69	262	144287288	21	-1	ATTTTAGGTAAACACCCAAAG	TRUE	1.35	0.78
70	275	144287912	21	-1	taggtgagaaactgagaatca	TRUE	1.45	0.77
71	276	144287920	21	-1	cagaaggctaggtgagaaact	TRUE	1.3	0.69
72	283	144288096	21	-1	GCCATTATGGCCAGAGGAAT	TRUE	1.71	0.91
73	286	144288193	21	-1	AGATACAGCAGAAAAGGTGAT	TRUE	1.13	0.92
74	288	144288268	21	1	AATTGAAAAATCACCTTGAG	TRUE	1.47	0.81
75	289	144292526	21	-1	cagttgattcatctgtacacg	TRUE	1.1	1.01
76	290	144292529	21	1	tttttgactctggtgtacag	TRUE	1.26	1.21
77	291	144292541	21	1	gctgtacagatgaatcaactg	TRUE	2.12	1.08
78	295	144292639	21	-1	ATCTCCCTTTTGAGTTGTCT	TRUE	1.31	0.88
79	296	144292651	21	-1	CTGTTCAAAAATATCTCCCT	TRUE	1.31	1.03
80	297	144292708	21	1	AAAATTACACAGAACTCCACC	TRUE	1.69	1.06
81	300	144292779	21	-1	TTTTTTGCTTTAAAGTGACA	TRUE	1.12	0.7
82	303	144293063	21	1	TCTTGTTTTAAATATGCTTT	TRUE	1.15	1.11
83	308	144293185	21	-1	CTCTGTATATTACATATGT	TRUE	1.05	0.73
84	311	144293308	21	-1	TATAAATATCAAAGGCTTAC	TRUE	0.88	0.73
85	316	144293537	21	-1	cctaggggaaaaactctagaaa	TRUE	1.17	0.82

[0367] 表3-4.

[0368]

86	318	144293638	21	1	acaccatgaaaaatctaatt	TRUE	1.09	0.93
87	322	144293778	21	-1	agatgtgtctagagtaaagaaa	TRUE	1.25	0.93
88	323	144293791	21	-1	GTATGATCTGTTGagatgtgc	TRUE	1.55	1.17
89	330	144294147	21	1	TTTAAAGATTATCAAATTGCT	TRUE	1.23	0.71
90	332	144294262	21	1	ATATGAATCACATTCTTTTGG	TRUE	1.33	0.93
91	334	144294294	21	1	TGCAAAAGCCAGTAGATAAAT	TRUE	1.01	0.81
92	335	144294300	21	1	AGCCAGTAGATAAATTGGAT	TRUE	0.8	1.01
93	339	144294447	21	1	TTTGTAGTTAGATTAAGTCAT	TRUE	0.81	0.84
94	342	144294575	21	-1	AAGAAACCTGGAAGAGCAGAT	TRUE	1.07	1.1
95	343	144294603	21	1	GGTTCTTTTTTGGGGGAAA	TRUE	1.37	0.92
96	350	144294924	21	1	TATGGTTGTAGTATACTTGCC	TRUE	1.24	1.02
97	351	144294930	21	1	TGTAGTATACTTGCCCTGGGT	TRUE	1.07	0.83
98	352	144294934	21	1	GTATACCTTGCCCTGGGTTTGG	TRUE	1.11	0.93
99	358	144295231	21	1	ACATGAAATAATAAAATGGTT	TRUE	0.86	1.03
100	360	144295268	21	-1	ATTATTGAATGAAATAGCAGT	TRUE	0.86	1.06
101	363	144295330	21	-1	ACAACACTGACAGCAACAGAA	TRUE	0.89	0.97
102	366	144295418	21	1	AGTGTGTGCTGGCTCCATG	TRUE	1.09	1.14
103	367	144295435	21	1	CATGTGGAGTTCTTGACAGTT	TRUE	1.03	0.98

[0369] 在表3中,“坐标”表示当使用SaCas9时对于所有示出的gRNA可能的SaCas9切割位点。

[0370] 如图7所示,所测试的100种靶向序列当中,5种靶向序列示出了在供体HSMM#3细胞中对UTRN mRNA表达的一致上调(向导编号145、146、205、208和210(SEQ ID NO:45、46、58、59和60))。这些序列中有2种,即145号(SEQ ID NO:45)、146号(SEQ ID NO:46)聚集在增强子E2区域,而其余3种,即205号(SEQ ID NO:58)、208号(SEQ ID NO:59)和210号(SEQ ID NO:60)聚集在启动子P1区域。向导编号205、208和210分别与实施例1中的sgED3-6、sgED3-30和sgED3-32相同。

[0371] 在这5种靶向序列中,145号、146号和208号这3种序列与食蟹猴基因组的相应区域100%匹配。另一方面,这些序列中的2种,即205号和210号与食蟹猴基因组的相应区域不匹配(图8)。

[0372] 当单独测试时,这5种靶向序列一致地示出了在5个不同供体HSMM中UTRN mRNA表达上调约2至4倍(图9)。在启动子区中的向导编号205、208或210和增强子区中的向导编号145或146的组合(图8中示意性示出)中,2种组合,即向导编号205和145(205+145)以及向导编号208和145(208+145),使得在5种不同的HSMM供体中UTRN表达上调约3至7倍(图9)。

[0373] 实施例3.AAV顺式质粒的生成和评估

[0374] (1) 实验方法

[0375] AAV AIO顺式质粒的构建

[0376] 如表4中所示,所有测试的质粒骨架pED260(SEQ ID NO:210)、pED261(SEQ ID NO:211)和pED263(SEQ ID NO:212)含有相同的全长dSaCas9、CK8启动子和U6启动子的碱基序列,替换pAAV-CMV载体(Takara#6234)的ITR之间的序列。其区别在于激活子部分、polyA序列,即pED260、pED261包含miniVR,而pED263包含microVR作为激活子部分,且pED260具有bGH polyA,而pED261、pED263具有2个SNRP-1polyA序列(SEQ ID NO:208)。

[0377] 表4.

	启动子 (dCas9)	dSaCas9	激活子	polyA	启动子 (gRNA)	靶向序列
[0378]	质粒	CK8	全长	miniVR	bGH polyA	SEQ ID NO: 45 (#145)
	pED260	CK8	全长	miniVR	bGH polyA	SEQ ID NO: 46 (#146)
	(5171bp*)	CK8	全长	miniVR	bGH polyA	SEQ ID NO: 59 (#208)
	质粒	CK8	全长	miniVR	2×sNRP-1 polyA	SEQ ID NO: 45 (#145)
	pED261	CK8	全长	miniVR	2×sNRP-1 polyA	SEQ ID NO: 46 (#146)
	(4973bp*)	CK8	全长	miniVR	2×sNRP-1 polyA	SEQ ID NO: 59 (#208)
	质粒	CK8	全长	microVR	2×sNRP-1 polyA	SEQ ID NO: 45 (#145)
	pED263	CK8	全长	microVR	2×sNRP-1 polyA	SEQ ID NO: 46 (#146)
	(4883bp*)	CK8	全长	microVR	2×sNRP-1 polyA	SEQ ID NO: 59 (#208)

[0379] *ITR之间的核苷酸长度(含ITR核苷酸)

[0380] 将包含由靶向序列向导编号145、146或208编码的crRNA的sgRNA的各寡核苷酸克隆到这些骨架中的各者中,以创建用于测试的多合一(AIO)质粒。产生的各AIO质粒表示为pAAV-CK8-dSaCas9-miniVR-bGH polyA-U6-sgRNA#145(pED260-145)、pAAV-CK8-dSaCas9-miniVR-bGH polyA-U6-sgRNA#146(pED260-146)、pAAV-CK8-dSaCas9-miniVR-bGH polyA-U6-sgRNA#208(pED260-208)、pAAV-CK8-dSaCas9-miniVR-2×sNRP-1polyA-U6-sgRNA#145(pED261-145)、pAAV-CK8-dSaCas9-miniVR-2×sNRP-1 polyA-U6-sgRNA#146(pED261-146)、pAAV-CK8-dSaCas9-miniVR-2×sNRP-1polyA-U6-sgRNA#208(pED261-208)、pAAV-CK8-dSaCas9-microVR-2×sNRP-1 polyA-U6-sgRNA#145(pED263-145)、pAAV-CK8-dSaCas9-microVR-2×sNRP-1polyA-U6-sgRNA#146(pED263-146)或pAAV-CK8-dSaCas9-microVR-2×sNRP-1 polyA-U6-sgRNA#208(pED263-208),如表4所示。

[0381] 已知与人基因组的任何部分都不同源的两个不同序列被用作阴性对照,并被称为非靶向性向导(NTg1(SEQ ID NO:1)和NTg2(SEQ ID NO:2))。NTg1或NTg2的各寡核苷酸也被克隆到各自的骨架中,并用作对照质粒。

[0382] HEK293FT细胞的转染

[0383] 将HEK293FT细胞(Thermo Fisher#R70007)以 5×10^4 个细胞/孔接种在24孔细胞培养皿(CORNING#351147)中的0.5ml生长培养基(补充有10%FBS和2mM新鲜L-谷氨酰胺、1mM

丙酮酸钠和非必需氨基酸的DMEM培养基(Thermo Fisher#11140050))中,在37℃/5%CO₂下孵育24小时。第二天,根据制造商的方案,用0.5μg质粒设置lipofectamine-2000转染反应(Thermo Fisher#11668019),所述质粒包含编码dSaCas9-miniVR或dSaCas9-microVR和包含实施例2中选择的靶向序列向导编号145(SEQ ID NO:45)、146(SEQ ID NO:46)或208(SEQ ID NO:59))(表4)的sgRNA的序列。

[0384] 转染后48小时,按照制造商的指导,收获细胞并用RNeasy 96试剂盒(Qiagen#74182)提取RNA。

[0385] 基因表达分析

[0386] 为了进行基因表达分析,根据高容量cDNA逆转录试剂盒(Applied Biosystems; Thermo Fisher#4368813)方案,以10μl的体积从约0.5μg的总RNA产生cDNA。将cDNA稀释10倍,并使用Taqman快速高级预混物(Thermo Fisher#4444557),根据制造商的方案进行分析。Taqman探针(UTRN:测定Id Hs01125994_m1 FAM;HPRT:测定Id Hs99999909_m1 VIC_PL)从Thermo Fisher获得。按照Taqman快速高级预混物方案的指导,用QuantStudio 5实时PCR系统对基于Taqman探针的实时PCR反应进行处理和分析。

[0387] 数据分析

[0388] 对于含有NTg1或NTg2的质粒,结果的平均值示出为CtrlX2。

[0389] 对于各样品,通过将非靶向性向导对照的3个技术重复样品的平均Ct值减去各样品的3个技术重复样品的平均Ct值,计算出各探针的ΔCt值。

[0390] $\Delta Ct_{UTRN} = \text{平均对照Ct}_{UTRN} - \text{平均样本Ct}_{UTRN}$ 。

[0391] $\Delta Ct_{HPRT} = \text{平均对照Ct}_{HPRT} - \text{平均样本Ct}_{HPRT}$ 。

[0392] 然后将各样品的UTRN的Ct值减去HPRT的Ct值来计算Δ(ΔCt)值。

[0393] $\Delta(\Delta Ct) = \Delta Ct_{UTRN} - \Delta Ct_{HPRT}$ 。

[0394] 使用式 $2^{(\Delta(\Delta Ct))}$ 确定各样品的表达值。

[0395] (2) 结果

[0396] 存在包含由靶向序列向导编号145(SEQ ID NO:45)或146(SEQ ID NO:46)或208(SEQ ID NO:59)编码的crRNA的sgRNA时,所有3个受测骨架(pED260、261和263)能够上调HEK293FT细胞中的UTRN(图10)。

[0397] 实施例4.携带dSaCas9、转录激活子和sgRNA的重组AAV9的产生

[0398] (1) 实验方法

[0399] 腺相关病毒(AAV)的产生

[0400] 使用293T细胞(ATCC CRL-3216)产生腺相关病毒血清型9(AAV9)粒子,所述细胞以每个T225烧瓶(Corning) 0.96×10^7 - 1.8×10^7 个细胞的密度接种,并在补充有10%FBS的DMEM培养基(Thermo Fisher#11995-065)中培养。pRC9质粒的构建如下:将AAV9衣壳序列(参见JP5054975B)亚克隆到pRC2-mi342质粒(Takara#6230)中,以替换AAV2衣壳序列。用20μg pRC9质粒和pHelper载体(Takara#6230)和20μg实施例3中使用的6种AI0质粒之一,即pED261-145、pED261-146、pED261-208、pED263-145、pED263-146或pED263-208转染细胞,每个T225烧瓶用180μl TransIT-293转染试剂(Mirus Bio#MIR2700)。转染后一天,将培养基更换为补充有2%FBS的DMEM培养基。72小时后,收获细胞,并使用AAVpro纯化试剂盒(所有血清型)(Takara#6666)根据制造商的说明提取和纯化AAV。使用AAVpro滴定试剂盒(用于实

时PCR) (Takara#6233) 测量纯化的AAV的滴度。所产生的各AAV表示为AAV9-ED261-145、AAV9-ED261-146、AAV9-ED261-208、AAV9-ED263-145、AAV9-ED263-146或AAV9-ED263-208。

[0401] AAV的验证

[0402] 在用NuPAGE样品还原剂、抗氧化剂和缓冲液 (Thermo Fisher#NP0009、#NP0005、#NP0007) 制备AAV样品后,使用NuPAGE 4至12%Bis-Tris蛋白凝胶1.0mm×12孔 (Thermo Fisher#NP0322BOX) 通过SDS-PAGE检查AAV衣壳蛋白。各AAV的施加量为 1.0×10^{10} vg/泳道。用Oriole荧光凝胶染料溶液 (BioRad#161-0495) 对凝胶进行染色后,用具有紫外线激发光和580nm滤光片的ChemiDoc™Touch (BioRad) 捕获图像。

[0403] (2) 结果

[0404] 在T225烧瓶中产生的AAV9的滴度值计算如下。

[0405] 表5.

AAV名称	浓度 (vg/mL)
AAV9-ED261-145	1.82×10^{12}
AAV9-ED261-146	3.66×10^{12}
AAV9-ED261-208	2.11×10^{12}
AAV9-ED263-145	2.43×10^{12}
AAV9-ED263-146	6.00×10^{12}
AAV9-ED263-208	1.43×10^{12}

[0407] 在SDS-PAGE中,从各AAV样品中检测到3种衣壳蛋白 (VP1、VP2和VP3,分别为87kDa、72kDa和62kDa) (图11)。这些结果表明,克隆到AAV AIO质粒中的包括dSaCas9和转录激活子在目标基因可以被封包到AAV9中。

[0408] 实施例5.携带dSaCas9、转录激活子和sgRNA的重组AAV9对肌营养相关蛋白的上调的体外药理评估

[0409] (1) 实验方法

[0410] AAV9产生

[0411] 使用293T细胞 (ATCC#CRL-3216) 产生腺相关病毒血清型9 (AAV9) 粒子,所述细胞以 4.77×10^7 个细胞/700mL/细胞堆叠5烧瓶 (Corning) 的密度接种,并在补充有10%FBS (Hyclone#SH30070.03)、1%MEM (Sigma#M7145)、1%青霉素/链霉素 (Thermo Fisher#15070-063) 和2.5%HEPES (Sigma#H0887) 的DMEM培养基中培养。三天后,用227.9μg在实施例4中构建的pRC9质粒、pHelper载体 (Takara#6230) 和在实施例3中使用的3种AIO质粒之一,即pED261-145、pED261-208或pED263-208转染细胞,每瓶含683.7μl聚乙烯亚胺Max (2mg/mL) (Polysciences#24765-2)。转染后6天,用Triton X-100 (终浓度0.2%) (Roche#10789704001) 收获细胞。对AAV样品进行离心,过滤,浓缩,并使用色谱法 (AKTA avant 25, GE Healthcare和POROS CaptureSelect AAV Resins色谱柱, Thermo Fisher) 纯化,然后使用CsCl进行超速离心 (Optima XE-90, Beckman Coulter)。透析靶级分后,使用AAVpro滴定试剂盒 (用于实时PCR) (Takara#6233) 测量AAV的滴度。获得了AAV9-ED261-145、AAV9-ED261-208和AAV9-ED263-208。

[0412] 细胞培养和AAV感染

[0413] 将人骨骼肌成肌细胞 (HSMM, Lonza#CC-2580, 批号18TL211617) 以每孔100,000个

细胞的密度接种到覆盖有胶原蛋白I的24孔板(IWAKI#4820-010)中,并在补充有500U/mL青霉素/链霉素(Thermo Fisher#15070063)的SkGM™-2骨骼肌细胞生长培养基2BulletKit™(Lonza#CC-3245)中在37℃,5%CO₂下培养2天。将培养基替换为分化培养基(补充有2%FBS(GE Healthcare#SH30070.03)和500U/mL青霉素/链霉素的DMEM培养基(Sigma#D6429)),并在37℃,5%CO₂下将细胞培养3天。为了进行AAV感染,将培养基替换为500μL新鲜分化培养基,其中含有0.2、1.0或 5.0×10^{11} vg/mL AAV9-ED261-145、AAV9-ED261-208或AAV9-ED263-208。感染后,将感染的细胞在37℃,5%CO₂下培养3-4天,并使用RNeasy Plus Mini试剂盒(Qiagen#74134)根据制造商的说明提取总RNA。将未感染AAV的细胞中的RNA设为对照,并示出为AAV(-)。

[0414] 基因表达分析

[0415] 对于Taqman qPCR,使用SuperScript™VIL0™cDNA合成试剂盒(Thermo Fisher#11754250),以20μL反应体积将250ng总RNA转化为cDNA。用水将cDNA稀释5倍,并将2μL用于qPCR。用QuantStudio™12K Flex实时PCR系统(Thermo Fisher)进行qPCR,最终体积为10μL,包含针对UTRN的Taqman探针(Thermo Fisher#Hs01125994_m1,FAM)、针对HPRT1的Taqman探针(Thermo Fisher#Hs02800695_m1,FAM)和TaqMan™Universal PCR预混物(Thermo Fisher#4324018)。qPCR循环条件如下:在95℃下10分钟,然后进行45次在95℃下15秒和在60℃下1分钟的循环。使用QuantStudio™12K Flex软件(Thermo Fisher)分析数据。用各基因的标准曲线分析表达值,且将UTRN基因的表达水平相对于HPRT1基因的表达水平进行标准化。

[0416] (2) 结果

[0417] 通过将AAV9-ED261-145、AAV9-ED261-208或AAV9-ED263-208施加至HSMM细胞中,发现了肌营养相关蛋白mRNA上调,这表明携带dSaCas9、miniVR或microVR和包含向导编号145或208的sgRNA的转基因的AAV9对人肌肉细胞中的肌营养相关蛋白上调具有药理作用(图12)。

[0418] 实施例6:使用RNA测序分析的脱靶分析

[0419] (1) 实验方法

[0420] 慢病毒的产生

[0421] 将HEK293TA细胞(Genecopoeia#LT008)以 0.75×10^6 个细胞/孔接种在6孔细胞培养皿(VWR#10062-892)中的2mL生长培养基(补充有10%FBS和2mM新鲜L-谷氨酰胺、1mM丙酮酸钠和非必需氨基酸的DMEM培养基(Thermo Fisher#11140050))中并在37℃/5%CO₂下孵育24小时。第二天,按照制造商的方案,使用1.5μg封包质粒混合物[1μg封包质粒(参见pCMV delta R8.2;addgene#12263)和0.5μg包膜表达质粒(参见pCMV-VSV-G;addgene#8454)]和1μg转移质粒设置TransIT-VirusGEN转染反应,所述转移质粒包含编码dSaCas9-miniVR和包含实施例2中选择的靶向序列(即向导编号145(SEQ ID NO:45)、146(SEQ ID NO:46)、208(SEQ ID NO:59)或NTg1(非靶向性向导-1)(SEQ ID NO:1))的sgRNA的碱基序列。转染后48至72小时,通过使培养基上清液通过0.45μm PES过滤器(VWR#10218-488)收获慢病毒。在准备使用之前,将经纯化且等分的慢病毒保存在-80℃的冰箱中。

[0422] HSMM细胞的转导和RNA样品制备

[0423] 从Lonza Inc.获得年龄为0岁的人供体的原代骨骼肌成肌细胞(HSMM)(批号

542368)。将细胞在来自Lonza的原代骨骼肌细胞生长培养基[SkGM™-2骨骼肌生长BulletKit培养基(#CC-3245),其包含有含骨骼肌成肌细胞生长所需的SkBM™-2基础培养基(#CC-3246)和SkGMTM-2SingleQuots™补充剂(#CC-3244)的培养系统]中培养。为了进行转导,将细胞以 0.125×10^6 个细胞/孔接种在含有生长培养基的6孔细胞培养皿(VWR#10062-894)中,并在37℃/5%CO₂下孵育24小时。第二天,向各孔中添加补充有8μg/ml聚凝胺(Sigma#TR-1003-G)和1.0ml对应于包含由各个靶向序列(向导编号145(SEQ ID NO:45)、146(SEQ ID NO:46)或208(SEQ ID NO:59))编码的crRNA和tracrRNA的各sgRNA的慢病毒上清液(滴度范围为 0.2 至 2×10^9 拷贝/ml,通过使用Lenti-X™ qRT-PCR滴定试剂盒(Clontech#631235)测量)的1.5ml生长培养基。将细胞与慢病毒一起孵育6小时,然后去除病毒培养基并替换为新鲜的生长培养基。转导后72小时,向细胞供以选择培养基[补充有0.5μg/ml嘌呤霉素的生长培养基(Sigma#P8833-100MG)]。每2至3天为细胞提供新鲜的选择培养基。细胞在选择培养基中7至10天后,按照制造商的指导收获细胞并用RNeasy 96试剂盒(Qiagen#74182)提取RNA。用作对照的NTg1(非靶向性向导1)向导的序列为ACGGAGGCTAAGCGTCGCAA(SEQ ID NO:1)。

[0424] 脱靶分析

[0425] 由GeneWiz,LLC进行Illumina测序,其中RNA文库是使用NEBNext Ultra RNA文库制备试剂盒(Ipswich,MA,USA,NEB#E7530L)根据制造商的方案制备。测序文库聚集在Illumina HiSeq流动槽的三个泳道上,并使用2×150配对末端配置进行测序。将产生的原始序列数据(.bcl文件)转换为fastq文件,并使用Illumina的bcl2fastq2.17软件进行多路分解,其中索引序列识别允许一个不匹配项。使用STAR比对器将fastq文件与人基因组装体GRCh38.p12比对。使用DESeq2进行差异分析,并使用定制R脚本通过plotly(<https://plot.ly>)生成图。

[0426] (2) 结果

[0427] 相对于非靶向性向导1(NTg1)标准化各向导的mRNA水平的全基因组倍数变化。每个点代表一个基因。X轴示出了基因的平均表达水平。Y轴示出了相对于NTg1样品,基因表达的log₂倍数变化。水平Log₂=0以上的基因表示实验样品(例如,向导编号145)中的基因表达高于NTg1样品,水平Log₂=0以下的基因表明实验样品中的基因表达低于NTg1样品。示出了在实验样品中比在NTg1样品中高度上调或下调的基因的基因ID。不同的gRNA诱导了不同的基因表达变化(图13A:向导编号145,13B:向导编号146,13C:向导编号208)。向导编号208似乎触发了较少的其它基因表达变化,同时示出了出良好的UTRN基因上调。

[0428] 实施例7.对肌营养相关蛋白上调的药理作用的体内评估

[0429] (1) 实验方法

[0430] 动物和免疫抑制方案

[0431] 在本研究中使用AAV9血清反应阴性的食蟹猴(雄性)。驯化一周后,向所述食蟹猴口服施用0.75mg/kg/天的泼尼松龙磷酸钠(Abcam#ab142456)。AAV施用前14天开始给药,一直持续到处死。

[0432] AAV9治疗和肌肉组织取样

[0433] 向食蟹猴经由头静脉静脉内施用 1.0 或 6.0×10^{13} vg/kg的AAV9-ED261-208(SignaGen制)。为了进行股四头肌活检,通过肌肉内施用10mg/kg的氯胺酮盐酸盐和

0.08mg/kg的美托咪定盐酸盐将猴麻醉,并在AAV施用前19天和施用后28天获取50至200mg样品。施用AAV9后56天,处死猴子,并获取各肌肉和心脏样本。将样品在液氮中冷冻并应用于基因和蛋白质表达分析。

[0434] 肌肉组织样品的基因和蛋白质表达分析

[0435] 对于Taqman qPCR,使用RNeasy纤维组织Mini试剂盒(Qiagen#74704)从肌肉样品中提取总RNA,并使用SuperscriptTMVILOTMcDNA合成试剂盒(Thermo Fisher#11754250)转化为cDNA。用针对UTRN的Taqman探针(Thermo Fisher#Mf01126001_m1,FAM)、针对HPRT1的Taqman探针(Thermo Fisher,#Hs02800695_m1,FAM),和TaqManTM通用PCR预混物(Thermo Fisher,#4324018)与QuantStudioTM12K Flex实时PCR系统(Thermo Fisher)进行qPCR。UTRN基因的表达水平相对于HPRT1基因进行标准化。

[0436] 对于蛋白质表达分析,用含有蛋白酶和磷酸酶抑制剂混合物(Thermo Fisher#78441)的RIPA缓冲液(Millipore#20-188)制备全肌肉裂解物,并应用于SDS-PAGE和Western印迹法。使用针对肌营养相关蛋白的第一抗体(SantaCruz#SC-33700)和辣根过氧化物酶标记的第二抗体(Cell Signaling#7076)检测肌营养相关蛋白蛋白。

[0437] 工业适用性

[0438] 根据本发明,可以激活UTRN基因在人细胞中的表达。因此,本发明预期对治疗和/或预防DMD和BMD极其有用。

[0439] 若在本文中指定一个数值限制或范围,则包括端点。此外,数值限制或范围内的所有值和子范围被明确包括在内,如同明确地写出一一般。

[0440] 如本文中所使用,词“一个”和“一种”等有“一个(种)或多个(种)”的意思。

[0441] 显然,本发明的许多修改和变化根据上述教导是可能的。因此,应理解,在所附权利要求的范围内,本发明可以以不同于本文中具体描述的其它方式实施。

[0442] 上述所有专利和其它参考文献通过引入全部并入本文,如同详细描述一般。

- [0001] 序列表
- [0002] <110> 阿斯泰来制药株式会社 (Astellas Pharma Inc.)
- [0003] 摩大力斯医疗株式会社 (Modalis Therapeutics Corporation)
- [0004] <120> 通过靶向肌营养相关蛋白基因来治疗肌营养不良的方法
- [0005] <130> 092955
- [0006] <150> US62/768,474
- [0007] <151> 2018-11-16
- [0008] <150> US62/861,039
- [0009] <151> 2019-06-13
- [0010] <150> US62/931,925
- [0011] <151> 2019-11-07
- [0012] <160> 216
- [0013] <170> PatentIn version 3.5
- [0014] <210> 1
- [0015] <211> 20
- [0016] <212> DNA
- [0017] <213> 人工序列
- [0018] <220>
- [0019] <223> 对照非靶向性靶向序列
- [0020] <400> 1
- [0021] acggaggcta agcgtcgcaa 20
- [0022] <210> 2
- [0023] <211> 20
- [0024] <212> DNA
- [0025] <213> 人工序列
- [0026] <220>
- [0027] <223> 对照非靶向性靶向序列
- [0028] <400> 2
- [0029] cgcttccgcg gcccggtcaa 20
- [0030] <210> 3
- [0031] <211> 20
- [0032] <212> DNA
- [0033] <213> 人工序列
- [0034] <220>
- [0035] <223> 对照非靶向性靶向序列
- [0036] <400> 3
- [0037] gtaggcgcgc cgctctctac 20
- [0038] <210> 4
- [0039] <211> 21
- [0040] <212> DNA
- [0041] <213> 智人 (Homo sapiens)

[0042] <400> 4
[0043] agaaaagcgg cccctagggg c 21
[0044] <210> 5
[0045] <211> 21
[0046] <212> DNA
[0047] <213> 智人(Homo sapiens)
[0048] <400> 5
[0049] caaacacaca ccagcaaact t 21
[0050] <210> 6
[0051] <211> 21
[0052] <212> DNA
[0053] <213> 智人(Homo sapiens)
[0054] <400> 6
[0055] tgaaagcgca actggagggc c 21
[0056] <210> 7
[0057] <211> 21
[0058] <212> DNA
[0059] <213> 智人(Homo sapiens)
[0060] <400> 7
[0061] acccacgcgg acatatgtcc a 21
[0062] <210> 8
[0063] <211> 21
[0064] <212> DNA
[0065] <213> 智人(Homo sapiens)
[0066] <400> 8
[0067] atccaatgga catatgtccg c 21
[0068] <210> 9
[0069] <211> 21
[0070] <212> DNA
[0071] <213> 智人(Homo sapiens)
[0072] <400> 9
[0073] gagggggagg gctgtgacct g 21
[0074] <210> 10
[0075] <211> 21
[0076] <212> DNA
[0077] <213> 智人(Homo sapiens)
[0078] <400> 10
[0079] atttggtggt caggagcaa g 21
[0080] <210> 11
[0081] <211> 21
[0082] <212> DNA
[0083] <213> 智人(Homo sapiens)

[0084] <400> 11
[0085] aatgaaacca aagacagctt c 21
[0086] <210> 12
[0087] <211> 21
[0088] <212> DNA
[0089] <213> 智人(Homo sapiens)
[0090] <400> 12
[0091] ccaaaatcct ttaatgaatc a 21
[0092] <210> 13
[0093] <211> 21
[0094] <212> DNA
[0095] <213> 智人(Homo sapiens)
[0096] <400> 13
[0097] tacagattcc atgattcatt a 21
[0098] <210> 14
[0099] <211> 21
[0100] <212> DNA
[0101] <213> 智人(Homo sapiens)
[0102] <400> 14
[0103] ggaacaaacc aaaatccttt a 21
[0104] <210> 15
[0105] <211> 21
[0106] <212> DNA
[0107] <213> 智人(Homo sapiens)
[0108] <400> 15
[0109] atctgtttgt ggggaaatct t 21
[0110] <210> 16
[0111] <211> 21
[0112] <212> DNA
[0113] <213> 智人(Homo sapiens)
[0114] <400> 16
[0115] caaacagatt tcagtatttt c 21
[0116] <210> 17
[0117] <211> 21
[0118] <212> DNA
[0119] <213> 智人(Homo sapiens)
[0120] <400> 17
[0121] gtggtgattt atgttactgg t 21
[0122] <210> 18
[0123] <211> 21
[0124] <212> DNA
[0125] <213> 智人(Homo sapiens)

[0126] <400> 18
[0127] tgagtctttc aagttccttt c 21
[0128] <210> 19
[0129] <211> 21
[0130] <212> DNA
[0131] <213> 智人(Homo sapiens)
[0132] <400> 19
[0133] agatcatttt ttgcttcaaa c 21
[0134] <210> 20
[0135] <211> 21
[0136] <212> DNA
[0137] <213> 智人(Homo sapiens)
[0138] <400> 20
[0139] ttgcttcaaa ctagaatgtc c 21
[0140] <210> 21
[0141] <211> 21
[0142] <212> DNA
[0143] <213> 智人(Homo sapiens)
[0144] <400> 21
[0145] gatctatcta tagacaccaa a 21
[0146] <210> 22
[0147] <211> 21
[0148] <212> DNA
[0149] <213> 智人(Homo sapiens)
[0150] <400> 22
[0151] tgcttcttcc aggttgagt g 21
[0152] <210> 23
[0153] <211> 21
[0154] <212> DNA
[0155] <213> 智人(Homo sapiens)
[0156] <400> 23
[0157] accgctttgc ttcttcagg c 21
[0158] <210> 24
[0159] <211> 21
[0160] <212> DNA
[0161] <213> 智人(Homo sapiens)
[0162] <400> 24
[0163] aagcctggaa gaagcaaagc g 21
[0164] <210> 25
[0165] <211> 21
[0166] <212> DNA
[0167] <213> 智人(Homo sapiens)

[0168] <400> 25
[0169] cttctgaatc agaattccta a 21
[0170] <210> 26
[0171] <211> 21
[0172] <212> DNA
[0173] <213> 智人(Homo sapiens)
[0174] <400> 26
[0175] tggttccaag ctagtacttc a 21
[0176] <210> 27
[0177] <211> 21
[0178] <212> DNA
[0179] <213> 智人(Homo sapiens)
[0180] <400> 27
[0181] atgttcacaa aataaattta a 21
[0182] <210> 28
[0183] <211> 21
[0184] <212> DNA
[0185] <213> 智人(Homo sapiens)
[0186] <400> 28
[0187] cctttatggt caccttctct g 21
[0188] <210> 29
[0189] <211> 21
[0190] <212> DNA
[0191] <213> 智人(Homo sapiens)
[0192] <400> 29
[0193] ccttctctgc tgagtaaaaa t 21
[0194] <210> 30
[0195] <211> 21
[0196] <212> DNA
[0197] <213> 智人(Homo sapiens)
[0198] <400> 30
[0199] aaggtggcca aaaaagaacc c 21
[0200] <210> 31
[0201] <211> 21
[0202] <212> DNA
[0203] <213> 智人(Homo sapiens)
[0204] <400> 31
[0205] aaggaagaga gaggaagaa a 21
[0206] <210> 32
[0207] <211> 21
[0208] <212> DNA
[0209] <213> 智人(Homo sapiens)

[0210] <400> 32
[0211] taaagaattc tagcactgga a 21
[0212] <210> 33
[0213] <211> 21
[0214] <212> DNA
[0215] <213> 智人(Homo sapiens)
[0216] <400> 33
[0217] aaatgtgtca tgtgttggtt a 21
[0218] <210> 34
[0219] <211> 21
[0220] <212> DNA
[0221] <213> 智人(Homo sapiens)
[0222] <400> 34
[0223] aaaaatgaaa attgcaactt c 21
[0224] <210> 35
[0225] <211> 21
[0226] <212> DNA
[0227] <213> 智人(Homo sapiens)
[0228] <400> 35
[0229] attgcaactt ctagaattta a 21
[0230] <210> 36
[0231] <211> 21
[0232] <212> DNA
[0233] <213> 智人(Homo sapiens)
[0234] <400> 36
[0235] cagctggagt gggccacgta a 21
[0236] <210> 37
[0237] <211> 21
[0238] <212> DNA
[0239] <213> 智人(Homo sapiens)
[0240] <400> 37
[0241] atttttgcat atttctttgg t 21
[0242] <210> 38
[0243] <211> 21
[0244] <212> DNA
[0245] <213> 智人(Homo sapiens)
[0246] <400> 38
[0247] agtgacctgc tgatttctct a 21
[0248] <210> 39
[0249] <211> 21
[0250] <212> DNA
[0251] <213> 智人(Homo sapiens)

[0252] <400> 39
[0253] ctttcccat tgttcaggac t 21
[0254] <210> 40
[0255] <211> 21
[0256] <212> DNA
[0257] <213> 智人(Homo sapiens)
[0258] <400> 40
[0259] ttggttgata aatttgata t 21
[0260] <210> 41
[0261] <211> 21
[0262] <212> DNA
[0263] <213> 智人(Homo sapiens)
[0264] <400> 41
[0265] tctctagttc attttttagc t 21
[0266] <210> 42
[0267] <211> 21
[0268] <212> DNA
[0269] <213> 智人(Homo sapiens)
[0270] <400> 42
[0271] tccttcaact tcaagacaac a 21
[0272] <210> 43
[0273] <211> 21
[0274] <212> DNA
[0275] <213> 智人(Homo sapiens)
[0276] <400> 43
[0277] gctcctcctg ctggatgggg g 21
[0278] <210> 44
[0279] <211> 21
[0280] <212> DNA
[0281] <213> 智人(Homo sapiens)
[0282] <400> 44
[0283] ctctatttcc agctcctcct g 21
[0284] <210> 45
[0285] <211> 21
[0286] <212> DNA
[0287] <213> 智人(Homo sapiens)
[0288] <400> 45
[0289] gtacagttag tgctactagg a 21
[0290] <210> 46
[0291] <211> 21
[0292] <212> DNA
[0293] <213> 智人(Homo sapiens)

[0294] <400> 46
[0295] gctactagga caggatgctg g 21
[0296] <210> 47
[0297] <211> 21
[0298] <212> DNA
[0299] <213> 智人(Homo sapiens)
[0300] <400> 47
[0301] ccccagctgt gcctctgttt t 21
[0302] <210> 48
[0303] <211> 21
[0304] <212> DNA
[0305] <213> 智人(Homo sapiens)
[0306] <400> 48
[0307] ttcccaaac agaggcacag c 21
[0308] <210> 49
[0309] <211> 21
[0310] <212> DNA
[0311] <213> 智人(Homo sapiens)
[0312] <400> 49
[0313] gttttgaaac tggtagcage t 21
[0314] <210> 50
[0315] <211> 21
[0316] <212> DNA
[0317] <213> 智人(Homo sapiens)
[0318] <400> 50
[0319] aaactgatgc ttgttaaag a 21
[0320] <210> 51
[0321] <211> 21
[0322] <212> DNA
[0323] <213> 智人(Homo sapiens)
[0324] <400> 51
[0325] cttgttaaag gaatgaatga a 21
[0326] <210> 52
[0327] <211> 21
[0328] <212> DNA
[0329] <213> 智人(Homo sapiens)
[0330] <400> 52
[0331] aatccaaagg attaactga a 21
[0332] <210> 53
[0333] <211> 21
[0334] <212> DNA
[0335] <213> 智人(Homo sapiens)

[0336] <400> 53
[0337] taccattttc aagttaatcc t 21
[0338] <210> 54
[0339] <211> 21
[0340] <212> DNA
[0341] <213> 智人(Homo sapiens)
[0342] <400> 54
[0343] tgccccctcc ctggagcact t 21
[0344] <210> 55
[0345] <211> 21
[0346] <212> DNA
[0347] <213> 智人(Homo sapiens)
[0348] <400> 55
[0349] agcaacgtca gcaaactgag a 21
[0350] <210> 56
[0351] <211> 21
[0352] <212> DNA
[0353] <213> 智人(Homo sapiens)
[0354] <400> 56
[0355] acgtcagcaa actgagatgg g 21
[0356] <210> 57
[0357] <211> 21
[0358] <212> DNA
[0359] <213> 智人(Homo sapiens)
[0360] <400> 57
[0361] gaaccgtgcg tgccgggagc c 21
[0362] <210> 58
[0363] <211> 21
[0364] <212> DNA
[0365] <213> 智人(Homo sapiens)
[0366] <400> 58
[0367] ggggtccgct ctccagatga g 21
[0368] <210> 59
[0369] <211> 21
[0370] <212> DNA
[0371] <213> 智人(Homo sapiens)
[0372] <400> 59
[0373] ggagggtggg gcgcaggacc g 21
[0374] <210> 60
[0375] <211> 21
[0376] <212> DNA
[0377] <213> 智人(Homo sapiens)

[0378] <400> 60
[0379] cctctctcgc gcacaaagtt g 21
[0380] <210> 61
[0381] <211> 21
[0382] <212> DNA
[0383] <213> 智人(Homo sapiens)
[0384] <400> 61
[0385] tctggctcca gaagccgatt g 21
[0386] <210> 62
[0387] <211> 21
[0388] <212> DNA
[0389] <213> 智人(Homo sapiens)
[0390] <400> 62
[0391] acaagtaagg ggcgttttca g 21
[0392] <210> 63
[0393] <211> 21
[0394] <212> DNA
[0395] <213> 智人(Homo sapiens)
[0396] <400> 63
[0397] gagctggcca agggctcctc t 21
[0398] <210> 64
[0399] <211> 21
[0400] <212> DNA
[0401] <213> 智人(Homo sapiens)
[0402] <400> 64
[0403] tagaggagcc cttggccagc t 21
[0404] <210> 65
[0405] <211> 21
[0406] <212> DNA
[0407] <213> 智人(Homo sapiens)
[0408] <400> 65
[0409] ccaagtccca gagtcgaaga t 21
[0410] <210> 66
[0411] <211> 21
[0412] <212> DNA
[0413] <213> 智人(Homo sapiens)
[0414] <400> 66
[0415] gtccacagga gaggtgggc a 21
[0416] <210> 67
[0417] <211> 21
[0418] <212> DNA
[0419] <213> 智人(Homo sapiens)

[0420] <400> 67
[0421] ctctgggtgg ttgctgctcc c 21
[0422] <210> 68
[0423] <211> 21
[0424] <212> DNA
[0425] <213> 智人(Homo sapiens)
[0426] <400> 68
[0427] tcagttgcag caagagatcc c 21
[0428] <210> 69
[0429] <211> 21
[0430] <212> DNA
[0431] <213> 智人(Homo sapiens)
[0432] <400> 69
[0433] attttaggta aacacccaaa g 21
[0434] <210> 70
[0435] <211> 21
[0436] <212> DNA
[0437] <213> 智人(Homo sapiens)
[0438] <400> 70
[0439] taggtgagaa actgagaatc a 21
[0440] <210> 71
[0441] <211> 21
[0442] <212> DNA
[0443] <213> 智人(Homo sapiens)
[0444] <400> 71
[0445] cagaaggcta ggtgagaaac t 21
[0446] <210> 72
[0447] <211> 21
[0448] <212> DNA
[0449] <213> 智人(Homo sapiens)
[0450] <400> 72
[0451] gccattaatg gccagaggaa t 21
[0452] <210> 73
[0453] <211> 21
[0454] <212> DNA
[0455] <213> 智人(Homo sapiens)
[0456] <400> 73
[0457] agatacagca gaaaaggtga t 21
[0458] <210> 74
[0459] <211> 21
[0460] <212> DNA
[0461] <213> 智人(Homo sapiens)

[0462] <400> 74
[0463] aatttgaaaa atcaccttga g 21
[0464] <210> 75
[0465] <211> 21
[0466] <212> DNA
[0467] <213> 智人(Homo sapiens)
[0468] <400> 75
[0469] cagttgattc atctgtacag c 21
[0470] <210> 76
[0471] <211> 21
[0472] <212> DNA
[0473] <213> 智人(Homo sapiens)
[0474] <400> 76
[0475] tttttgactc tggctgtaca g 21
[0476] <210> 77
[0477] <211> 21
[0478] <212> DNA
[0479] <213> 智人(Homo sapiens)
[0480] <400> 77
[0481] gctgtacaga tgaatcaact g 21
[0482] <210> 78
[0483] <211> 21
[0484] <212> DNA
[0485] <213> 智人(Homo sapiens)
[0486] <400> 78
[0487] atctcccctt tgagtttgtc t 21
[0488] <210> 79
[0489] <211> 21
[0490] <212> DNA
[0491] <213> 智人(Homo sapiens)
[0492] <400> 79
[0493] ctgttcaaaa atatctcccc t 21
[0494] <210> 80
[0495] <211> 21
[0496] <212> DNA
[0497] <213> 智人(Homo sapiens)
[0498] <400> 80
[0499] aaaattacac agaactccac c 21
[0500] <210> 81
[0501] <211> 21
[0502] <212> DNA
[0503] <213> 智人(Homo sapiens)

[0504] <400> 81
[0505] ttttttgtct ttaaagtgac a 21
[0506] <210> 82
[0507] <211> 21
[0508] <212> DNA
[0509] <213> 智人(Homo sapiens)
[0510] <400> 82
[0511] tcttggttta aaatatgctt t 21
[0512] <210> 83
[0513] <211> 21
[0514] <212> DNA
[0515] <213> 智人(Homo sapiens)
[0516] <400> 83
[0517] ctctgttata ttacatatg t 21
[0518] <210> 84
[0519] <211> 21
[0520] <212> DNA
[0521] <213> 智人(Homo sapiens)
[0522] <400> 84
[0523] tataaatatc aaaggtctta c 21
[0524] <210> 85
[0525] <211> 21
[0526] <212> DNA
[0527] <213> 智人(Homo sapiens)
[0528] <400> 85
[0529] cctaggga aactctagaa a 21
[0530] <210> 86
[0531] <211> 21
[0532] <212> DNA
[0533] <213> 智人(Homo sapiens)
[0534] <400> 86
[0535] acaccatgaa aatctaatat t 21
[0536] <210> 87
[0537] <211> 21
[0538] <212> DNA
[0539] <213> 智人(Homo sapiens)
[0540] <400> 87
[0541] agatgtgcta gagtaaagaa a 21
[0542] <210> 88
[0543] <211> 21
[0544] <212> DNA
[0545] <213> 智人(Homo sapiens)

[0546] <400> 88
[0547] gtatgatctg ttcagatgtg c 21
[0548] <210> 89
[0549] <211> 21
[0550] <212> DNA
[0551] <213> 智人(Homo sapiens)
[0552] <400> 89
[0553] tttaaagatt atcaaattgc t 21
[0554] <210> 90
[0555] <211> 21
[0556] <212> DNA
[0557] <213> 智人(Homo sapiens)
[0558] <400> 90
[0559] atatgaatca cattcttttg g 21
[0560] <210> 91
[0561] <211> 21
[0562] <212> DNA
[0563] <213> 智人(Homo sapiens)
[0564] <400> 91
[0565] tgcaaaagcc agtagataaa t 21
[0566] <210> 92
[0567] <211> 21
[0568] <212> DNA
[0569] <213> 智人(Homo sapiens)
[0570] <400> 92
[0571] agccagtaga taaatttgga t 21
[0572] <210> 93
[0573] <211> 21
[0574] <212> DNA
[0575] <213> 智人(Homo sapiens)
[0576] <400> 93
[0577] ttttagttta gattaagtca t 21
[0578] <210> 94
[0579] <211> 21
[0580] <212> DNA
[0581] <213> 智人(Homo sapiens)
[0582] <400> 94
[0583] aagaaacctg gaagagcaga t 21
[0584] <210> 95
[0585] <211> 21
[0586] <212> DNA
[0587] <213> 智人(Homo sapiens)

[0588] <400> 95
[0589] ggtttctttt ttggggggaa a 21
[0590] <210> 96
[0591] <211> 21
[0592] <212> DNA
[0593] <213> 智人(Homo sapiens)
[0594] <400> 96
[0595] tatggttgta gtatacttgc c 21
[0596] <210> 97
[0597] <211> 21
[0598] <212> DNA
[0599] <213> 智人(Homo sapiens)
[0600] <400> 97
[0601] tgtagtatac ttgccttggg t 21
[0602] <210> 98
[0603] <211> 21
[0604] <212> DNA
[0605] <213> 智人(Homo sapiens)
[0606] <400> 98
[0607] gtatacttgc cttgggttg g 21
[0608] <210> 99
[0609] <211> 21
[0610] <212> DNA
[0611] <213> 智人(Homo sapiens)
[0612] <400> 99
[0613] acatgaaata ataaaatggt t 21
[0614] <210> 100
[0615] <211> 21
[0616] <212> DNA
[0617] <213> 智人(Homo sapiens)
[0618] <400> 100
[0619] attattgaat gaaatagcag t 21
[0620] <210> 101
[0621] <211> 21
[0622] <212> DNA
[0623] <213> 智人(Homo sapiens)
[0624] <400> 101
[0625] acaacactga cagcaacaga a 21
[0626] <210> 102
[0627] <211> 21
[0628] <212> DNA
[0629] <213> 智人(Homo sapiens)

[0630] <400> 102
 [0631] agtgtgtcag ctggctccat g 21
 [0632] <210> 103
 [0633] <211> 21
 [0634] <212> DNA
 [0635] <213> 智人(Homo sapiens)
 [0636] <400> 103
 [0637] catgtggagt tcttgacagt t 21
 [0638] <210> 104
 [0639] <211> 32
 [0640] <212> DNA
 [0641] <213> 智人(Homo sapiens)
 [0642] <400> 104
 [0643] gtacagttag tgctactagg acaggatgct gg 32
 [0644] <210> 105
 [0645] <211> 494
 [0646] <212> DNA
 [0647] <213> 智人(Homo sapiens)
 [0648] <400> 105
 [0649] gagcgcgcgc cccaggccag ccaagcgccg agccgggctg ctgcgggctg ggagggcgcg 60
 [0650] cagggccggc gctgattgac ggggcgcgca gtcaggtgac ttggggcgcc aagttcccga 120
 [0651] cgcggtggcc gcggtgaccg ccgaggcccg gcagacgctg acccggaac gtagtggggc 180
 [0652] tgatcttccg gaacaaagtt gctgggcccg cggcggcggg gcgagagcgc cgagggggag 240
 [0653] ccggagcgct gcagagcgcg gggccggagg gctggcgctg atctgcaccc ttctcatctg 300
 [0654] gagagcggac ccctggctgc ccgagggcga gccccttccc ggggggtggg ggcggaacg 360
 [0655] cgcgaccag cggtcctgcg cccaccctc cctcctccgc ctccagcgt cggtccaac 420
 [0656] aaaggggcag gcccgagcg gggaggagga ggaggagccg ccgaaggagc ggcctctct 480
 [0657] cgcgacaaa gttg 494
 [0658] <210> 106
 [0659] <211> 19
 [0660] <212> RNA
 [0661] <213> 新凶手弗朗西丝氏菌(Francisella novicida)
 [0662] <220>
 [0663] <221> misc_feature
 [0664] <222> (1) .. (19)
 [0665] <223> crRNA的5'-柄
 [0666] <400> 106
 [0667] aaauucuacu guuguagau 19
 [0668] <210> 107
 [0669] <211> 1053
 [0670] <212> PRT
 [0671] <213> 金黄色葡萄球菌(Staphylococcus aureus)

[0672]	<220>
[0673]	<221> 变体
[0674]	<222> (10) .. (10)
[0675]	<223> Asp残基转化成Ala残基
[0676]	<220>
[0677]	<221> 变体
[0678]	<222> (580) .. (580)
[0679]	<223> Asn残基转化成Ala残基
[0680]	<400> 107
[0681]	Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
[0682]	1 5 10 15
[0683]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
[0684]	20 25 30
[0685]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
[0686]	35 40 45
[0687]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
[0688]	50 55 60
[0689]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
[0690]	65 70 75 80
[0691]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
[0692]	85 90 95
[0693]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
[0694]	100 105 110
[0695]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
[0696]	115 120 125
[0697]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
[0698]	130 135 140
[0699]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
[0700]	145 150 155 160
[0701]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
[0702]	165 170 175
[0703]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
[0704]	180 185 190
[0705]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
[0706]	195 200 205
[0707]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
[0708]	210 215 220
[0709]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
[0710]	225 230 235 240
[0711]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
[0712]	245 250 255
[0713]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn

[0714]	260	265	270
[0715]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[0716]	275	280	285
[0717]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[0718]	290	295	300
[0719]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[0720]	305	310	315
[0721]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[0722]	325	330	335
[0723]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[0724]	340	345	350
[0725]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[0726]	355	360	365
[0727]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[0728]	370	375	380
[0729]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[0730]	385	390	395
[0731]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[0732]	405	410	415
[0733]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[0734]	420	425	430
[0735]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[0736]	435	440	445
[0737]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[0738]	450	455	460
[0739]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[0740]	465	470	475
[0741]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[0742]	485	490	495
[0743]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[0744]	500	505	510
[0745]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[0746]	515	520	525
[0747]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[0748]	530	535	540
[0749]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[0750]	545	550	555
[0751]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[0752]	565	570	575
[0753]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[0754]	580	585	590
[0755]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		

[0756]	595	600	605
[0757]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[0758]	610	615	620
[0759]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[0760]	625	630	635
[0761]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[0762]	645	650	655
[0763]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[0764]	660	665	670
[0765]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[0766]	675	680	685
[0767]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[0768]	690	695	700
[0769]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[0770]	705	710	715
[0771]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys		
[0772]	725	730	735
[0773]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu		
[0774]	740	745	750
[0775]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp		
[0776]	755	760	765
[0777]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile		
[0778]	770	775	780
[0779]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu		
[0780]	785	790	795
[0781]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu		
[0782]	805	810	815
[0783]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His		
[0784]	820	825	830
[0785]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly		
[0786]	835	840	845
[0787]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr		
[0788]	850	855	860
[0789]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile		
[0790]	865	870	875
[0791]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp		
[0792]	885	890	895
[0793]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr		
[0794]	900	905	910
[0795]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val		
[0796]	915	920	925
[0797]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser		

[0798]	930	935	940
[0799]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala		
[0800]	945	950	955 960
[0801]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly		
[0802]	965	970	975
[0803]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile		
[0804]	980	985	990
[0805]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met		
[0806]	995	1000	1005
[0807]	Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys		
[0808]	1010	1015	1020
[0809]	Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu		
[0810]	1025	1030	1035
[0811]	Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly		
[0812]	1040	1045	1050
[0813]	<210> 108		
[0814]	<211> 1053		
[0815]	<212> PRT		
[0816]	<213> 金黄色葡萄球菌 (Staphylococcus aureus)		
[0817]	<220>		
[0818]	<221> 变体		
[0819]	<222> (10) .. (10)		
[0820]	<223> Asp残基转化成Ala残基		
[0821]	<220>		
[0822]	<221> 变体		
[0823]	<222> (557) .. (557)		
[0824]	<223> His残基转化成Ala残基		
[0825]	<400> 108		
[0826]	Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val		
[0827]	1 5 10 15		
[0828]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly		
[0829]	20 25 30		
[0830]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg		
[0831]	35 40 45		
[0832]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile		
[0833]	50 55 60		
[0834]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His		
[0835]	65 70 75 80		
[0836]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu		
[0837]	85 90 95		
[0838]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu		
[0839]	100 105 110		

[0840]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr		
[0841]	115	120	125
[0842]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[0843]	130	135	140
[0844]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[0845]	145	150	155
[0846]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[0847]	165	170	175
[0848]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[0849]	180	185	190
[0850]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[0851]	195	200	205
[0852]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[0853]	210	215	220
[0854]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[0855]	225	230	235
[0856]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[0857]	245	250	255
[0858]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[0859]	260	265	270
[0860]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[0861]	275	280	285
[0862]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[0863]	290	295	300
[0864]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[0865]	305	310	315
[0866]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[0867]	325	330	335
[0868]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[0869]	340	345	350
[0870]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[0871]	355	360	365
[0872]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[0873]	370	375	380
[0874]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[0875]	385	390	395
[0876]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[0877]	405	410	415
[0878]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[0879]	420	425	430
[0880]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[0881]	435	440	445

[0882]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
[0883]	450 455 460
[0884]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
[0885]	465 470 475 480
[0886]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
[0887]	485 490 495
[0888]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
[0889]	500 505 510
[0890]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
[0891]	515 520 525
[0892]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
[0893]	530 535 540
[0894]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp Ala Ile Ile Pro
[0895]	545 550 555 560
[0896]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
[0897]	565 570 575
[0898]	Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
[0899]	580 585 590
[0900]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
[0901]	595 600 605
[0902]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
[0903]	610 615 620
[0904]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
[0905]	625 630 635 640
[0906]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
[0907]	645 650 655
[0908]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
[0909]	660 665 670
[0910]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
[0911]	675 680 685
[0912]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
[0913]	690 695 700
[0914]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
[0915]	705 710 715 720
[0916]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
[0917]	725 730 735
[0918]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
[0919]	740 745 750
[0920]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
[0921]	755 760 765
[0922]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
[0923]	770 775 780

[0924]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
[0925]	785 790 795 800
[0926]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
[0927]	805 810 815
[0928]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
[0929]	820 825 830
[0930]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
[0931]	835 840 845
[0932]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
[0933]	850 855 860
[0934]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
[0935]	865 870 875 880
[0936]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
[0937]	885 890 895
[0938]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
[0939]	900 905 910
[0940]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
[0941]	915 920 925
[0942]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
[0943]	930 935 940
[0944]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
[0945]	945 950 955 960
[0946]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
[0947]	965 970 975
[0948]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
[0949]	980 985 990
[0950]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
[0951]	995 1000 1005
[0952]	Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
[0953]	1010 1015 1020
[0954]	Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
[0955]	1025 1030 1035
[0956]	Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
[0957]	1040 1045 1050
[0958]	<210> 109
[0959]	<211> 1028
[0960]	<212> PRT
[0961]	<213> 人工序列
[0962]	<220>
[0963]	<223> 氨基酸残基 (dSaCas9的第721位至第745位氨基酸残基) 缺失突变体
[0964]	<220>
[0965]	<221> 变体

[0966]	<222> (10) .. (10)																
[0967]	<223> Asp残基转化成Ala残基																
[0968]	<220>																
[0969]	<221> 变体																
[0970]	<222> (580) .. (580)																
[0971]	<223> Asn残基转化成Ala残基																
[0972]	<400> 109																
[0973]	Met	Lys	Arg	Asn	Tyr	Ile	Leu	Gly	Leu	Ala	Ile	Gly	Ile	Thr	Ser	Val	
[0974]	1				5					10						15	
[0975]	Gly	Tyr	Gly	Ile	Ile	Asp	Tyr	Glu	Thr	Arg	Asp	Val	Ile	Asp	Ala	Gly	
[0976]				20					25					30			
[0977]	Val	Arg	Leu	Phe	Lys	Glu	Ala	Asn	Val	Glu	Asn	Asn	Glu	Gly	Arg	Arg	
[0978]				35					40					45			
[0979]	Ser	Lys	Arg	Gly	Ala	Arg	Arg	Leu	Lys	Arg	Arg	Arg	Arg	His	Arg	Ile	
[0980]		50						55					60				
[0981]	Gln	Arg	Val	Lys	Lys	Leu	Leu	Phe	Asp	Tyr	Asn	Leu	Leu	Thr	Asp	His	
[0982]	65					70						75				80	
[0983]	Ser	Glu	Leu	Ser	Gly	Ile	Asn	Pro	Tyr	Glu	Ala	Arg	Val	Lys	Gly	Leu	
[0984]					85					90					95		
[0985]	Ser	Gln	Lys	Leu	Ser	Glu	Glu	Glu	Phe	Ser	Ala	Ala	Leu	Leu	His	Leu	
[0986]				100					105					110			
[0987]	Ala	Lys	Arg	Arg	Gly	Val	His	Asn	Val	Asn	Glu	Val	Glu	Glu	Asp	Thr	
[0988]				115					120					125			
[0989]	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Ile	Ser	Arg	Asn	Ser	Lys	Ala	
[0990]		130						135					140				
[0991]	Leu	Glu	Glu	Lys	Tyr	Val	Ala	Glu	Leu	Gln	Leu	Glu	Arg	Leu	Lys	Lys	
[0992]	145					150						155				160	
[0993]	Asp	Gly	Glu	Val	Arg	Gly	Ser	Ile	Asn	Arg	Phe	Lys	Thr	Ser	Asp	Tyr	
[0994]					165					170					175		
[0995]	Val	Lys	Glu	Ala	Lys	Gln	Leu	Leu	Lys	Val	Gln	Lys	Ala	Tyr	His	Gln	
[0996]				180					185					190			
[0997]	Leu	Asp	Gln	Ser	Phe	Ile	Asp	Thr	Tyr	Ile	Asp	Leu	Leu	Glu	Thr	Arg	
[0998]				195					200					205			
[0999]	Arg	Thr	Tyr	Tyr	Glu	Gly	Pro	Gly	Glu	Gly	Ser	Pro	Phe	Gly	Trp	Lys	
[1000]		210						215					220				
[1001]	Asp	Ile	Lys	Glu	Trp	Tyr	Glu	Met	Leu	Met	Gly	His	Cys	Thr	Tyr	Phe	
[1002]	225					230					235					240	
[1003]	Pro	Glu	Glu	Leu	Arg	Ser	Val	Lys	Tyr	Ala	Tyr	Asn	Ala	Asp	Leu	Tyr	
[1004]					245					250					255		
[1005]	Asn	Ala	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Val	Ile	Thr	Arg	Asp	Glu	Asn	
[1006]				260					265					270			
[1007]	Glu	Lys	Leu	Glu	Tyr	Tyr	Glu	Lys	Phe	Gln	Ile	Ile	Glu	Asn	Val	Phe	

[1008]	275	280	285
[1009]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[1010]	290	295	300
[1011]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[1012]	305	310	315
[1013]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[1014]	325	330	335
[1015]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[1016]	340	345	350
[1017]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[1018]	355	360	365
[1019]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[1020]	370	375	380
[1021]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[1022]	385	390	395
[1023]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[1024]	405	410	415
[1025]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[1026]	420	425	430
[1027]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[1028]	435	440	445
[1029]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[1030]	450	455	460
[1031]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[1032]	465	470	475
[1033]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[1034]	485	490	495
[1035]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[1036]	500	505	510
[1037]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[1038]	515	520	525
[1039]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[1040]	530	535	540
[1041]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[1042]	545	550	555
[1043]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[1044]	565	570	575
[1045]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[1046]	580	585	590
[1047]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[1048]	595	600	605
[1049]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		

[1050]	610	615	620
[1051]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[1052]	625	630	635 640
[1053]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[1054]	645	650	655
[1055]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[1056]	660	665	670
[1057]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[1058]	675	680	685
[1059]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[1060]	690	695	700
[1061]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[1062]	705	710	715 720
[1063]	Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile Thr Pro His Gln Ile Lys		
[1064]	725	730	735
[1065]	His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr Ser His Arg Val Asp Lys		
[1066]	740	745	750
[1067]	Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr Leu Tyr Ser Thr Arg Lys		
[1068]	755	760	765
[1069]	Asp Asp Lys Gly Asn Thr Leu Ile Val Asn Asn Leu Asn Gly Leu Tyr		
[1070]	770	775	780
[1071]	Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu Ile Asn Lys Ser Pro Glu		
[1072]	785	790	795 800
[1073]	Lys Leu Leu Met Tyr His His Asp Pro Gln Thr Tyr Gln Lys Leu Lys		
[1074]	805	810	815
[1075]	Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys Asn Pro Leu Tyr Lys Tyr		
[1076]	820	825	830
[1077]	Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys Tyr Ser Lys Lys Asp Asn		
[1078]	835	840	845
[1079]	Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr Gly Asn Lys Leu Asn Ala		
[1080]	850	855	860
[1081]	His Leu Asp Ile Thr Asp Asp Tyr Pro Asn Ser Arg Asn Lys Val Val		
[1082]	865	870	875 880
[1083]	Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp Val Tyr Leu Asp Asn Gly		
[1084]	885	890	895
[1085]	Val Tyr Lys Phe Val Thr Val Lys Asn Leu Asp Val Ile Lys Lys Glu		
[1086]	900	905	910
[1087]	Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr Glu Glu Ala Lys Lys Leu		
[1088]	915	920	925
[1089]	Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile Ala Ser Phe Tyr Asn Asn		
[1090]	930	935	940
[1091]	Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr Arg Val Ile Gly Val Asn		

[1092]	945	950	955	960
[1093]	Asn Asp Leu Leu Asn Arg Ile Glu Val Asn Met Ile Asp Ile Thr Tyr			
[1094]		965	970	975
[1095]	Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys Arg Pro Pro Arg Ile Ile			
[1096]		980	985	990
[1097]	Lys Thr Ile Ala Ser Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp			
[1098]		995	1000	1005
[1099]	Ile Leu Gly Asn Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln			
[1100]		1010	1015	1020
[1101]	Ile Ile Lys Lys Gly			
[1102]	1025			
[1103]	<210> 110			
[1104]	<211> 6			
[1105]	<212> PRT			
[1106]	<213> 人工序列			
[1107]	<220>			
[1108]	<223> GSGGS连接子			
[1109]	<400> 110			
[1110]	Gly Gly Ser Gly Gly Ser			
[1111]	1	5		
[1112]	<210> 111			
[1113]	<211> 1034			
[1114]	<212> PRT			
[1115]	<213> 人工序列			
[1116]	<220>			
[1117]	<223> 具有GSGGS连接子的氨基酸残基 (dSaCas9的第721位至第745位氨基酸残基) 缺失突变体			
[1118]	<220>			
[1119]	<221> 变体			
[1120]	<222> (10) .. (10)			
[1121]	<223> Asp残基转化成Ala残基			
[1122]	<220>			
[1123]	<221> 变体			
[1124]	<222> (580) .. (580)			
[1125]	<223> Asn残基转化成Ala残基			
[1126]	<220>			
[1127]	<221> 位点			
[1128]	<222> (721) .. (726)			
[1129]	<223> GSGGS连接子			
[1130]	<400> 111			
[1131]	Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val			
[1132]	1	5	10	15

[1133]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
[1134]	20 25 30
[1135]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
[1136]	35 40 45
[1137]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
[1138]	50 55 60
[1139]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
[1140]	65 70 75 80
[1141]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
[1142]	85 90 95
[1143]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
[1144]	100 105 110
[1145]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
[1146]	115 120 125
[1147]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
[1148]	130 135 140
[1149]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
[1150]	145 150 155 160
[1151]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
[1152]	165 170 175
[1153]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
[1154]	180 185 190
[1155]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
[1156]	195 200 205
[1157]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
[1158]	210 215 220
[1159]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
[1160]	225 230 235 240
[1161]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
[1162]	245 250 255
[1163]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
[1164]	260 265 270
[1165]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
[1166]	275 280 285
[1167]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
[1168]	290 295 300
[1169]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
[1170]	305 310 315 320
[1171]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
[1172]	325 330 335
[1173]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
[1174]	340 345 350

[1175]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[1176]	355	360	365
[1177]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[1178]	370	375	380
[1179]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[1180]	385	390	395
[1181]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[1182]	405	410	415
[1183]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[1184]	420	425	430
[1185]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[1186]	435	440	445
[1187]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[1188]	450	455	460
[1189]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[1190]	465	470	475
[1191]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[1192]	485	490	495
[1193]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[1194]	500	505	510
[1195]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[1196]	515	520	525
[1197]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[1198]	530	535	540
[1199]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[1200]	545	550	555
[1201]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[1202]	565	570	575
[1203]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[1204]	580	585	590
[1205]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[1206]	595	600	605
[1207]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[1208]	610	615	620
[1209]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[1210]	625	630	635
[1211]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[1212]	645	650	655
[1213]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[1214]	660	665	670
[1215]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[1216]	675	680	685

[1217]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
[1218]	690 695 700
[1219]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
[1220]	705 710 715 720
[1221]	Gly Gly Ser Gly Gly Ser Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile
[1222]	725 730 735
[1223]	Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr
[1224]	740 745 750
[1225]	Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr
[1226]	755 760 765
[1227]	Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile Val Asn
[1228]	770 775 780
[1229]	Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu
[1230]	785 790 795 800
[1231]	Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp Pro Gln
[1232]	805 810 815
[1233]	Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys
[1234]	820 825 830
[1235]	Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys
[1236]	835 840 845
[1237]	Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr
[1238]	850 855 860
[1239]	Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr Pro Asn
[1240]	865 870 875 880
[1241]	Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp
[1242]	885 890 895
[1243]	Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys Asn Leu
[1244]	900 905 910
[1245]	Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr
[1246]	915 920 925
[1247]	Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile
[1248]	930 935 940
[1249]	Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr
[1250]	945 950 955 960
[1251]	Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu Val Asn
[1252]	965 970 975
[1253]	Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys
[1254]	980 985 990
[1255]	Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr Gln Ser Ile
[1256]	995 1000 1005
[1257]	Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu Val Lys
[1258]	1010 1015 1020

[1259]	Ser Lys Lys His Pro Gln IleIle Lys Lys Gly															
[1260]	1025				1030											
[1261]	<210> 112															
[1262]	<211> 886															
[1263]	<212> PRT															
[1264]	<213> 人工序列															
[1265]	<220>															
[1266]	<223> 氨基酸残基 (dSaCas9的第482位至第648位氨基酸残基) 缺失突变体															
[1267]	<220>															
[1268]	<221> 变体															
[1269]	<222> (10) .. (10)															
[1270]	<223> Asp残基转化成Ala残基															
[1271]	<400> 112															
[1272]	Met	Lys	Arg	Asn	Tyr	Ile	Leu	Gly	Leu	Ala	Ile	Gly	Ile	Thr	Ser	Val
[1273]	1	5				10				15						
[1274]	Gly	Tyr	Gly	Ile	Ile	Asp	Tyr	Glu	Thr	Arg	Asp	Val	Ile	Asp	Ala	Gly
[1275]	20				25				30							
[1276]	Val	Arg	Leu	Phe	Lys	Glu	Ala	Asn	Val	Glu	Asn	Asn	Glu	Gly	Arg	Arg
[1277]	35				40				45							
[1278]	Ser	Lys	Arg	Gly	Ala	Arg	Arg	Leu	Lys	Arg	Arg	Arg	Arg	His	Arg	Ile
[1279]	50				55				60							
[1280]	Gln	Arg	Val	Lys	Lys	Leu	Leu	Phe	Asp	Tyr	Asn	Leu	Leu	Thr	Asp	His
[1281]	65				70				75				80			
[1282]	Ser	Glu	Leu	Ser	Gly	Ile	Asn	Pro	Tyr	Glu	Ala	Arg	Val	Lys	Gly	Leu
[1283]	85				90				95							
[1284]	Ser	Gln	Lys	Leu	Ser	Glu	Glu	Glu	Phe	Ser	Ala	Ala	Leu	Leu	His	Leu
[1285]	100				105				110							
[1286]	Ala	Lys	Arg	Arg	Gly	Val	His	Asn	Val	Asn	Glu	Val	Glu	Glu	Asp	Thr
[1287]	115				120				125							
[1288]	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Ile	Ser	Arg	Asn	Ser	Lys	Ala
[1289]	130				135				140							
[1290]	Leu	Glu	Glu	Lys	Tyr	Val	Ala	Glu	Leu	Gln	Leu	Glu	Arg	Leu	Lys	Lys
[1291]	145				150				155				160			
[1292]	Asp	Gly	Glu	Val	Arg	Gly	Ser	Ile	Asn	Arg	Phe	Lys	Thr	Ser	Asp	Tyr
[1293]	165				170				175							
[1294]	Val	Lys	Glu	Ala	Lys	Gln	Leu	Leu	Lys	Val	Gln	Lys	Ala	Tyr	His	Gln
[1295]	180				185				190							
[1296]	Leu	Asp	Gln	Ser	Phe	Ile	Asp	Thr	Tyr	Ile	Asp	Leu	Leu	Glu	Thr	Arg
[1297]	195				200				205							
[1298]	Arg	Thr	Tyr	Tyr	Glu	Gly	Pro	Gly	Glu	Gly	Ser	Pro	Phe	Gly	Trp	Lys
[1299]	210				215				220							
[1300]	Asp	Ile	Lys	Glu	Trp	Tyr	Glu	Met	Leu	Met	Gly	His	Cys	Thr	Tyr	Phe

[1301]	225	230	235	240
[1302]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr			
[1303]		245	250	255
[1304]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn			
[1305]		260	265	270
[1306]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe			
[1307]		275	280	285
[1308]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu			
[1309]		290	295	300
[1310]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys			
[1311]	305	310	315	320
[1312]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr			
[1313]		325	330	335
[1314]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala			
[1315]		340	345	350
[1316]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu			
[1317]		355	360	365
[1318]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser			
[1319]		370	375	380
[1320]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile			
[1321]	385	390	395	400
[1322]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala			
[1323]		405	410	415
[1324]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln			
[1325]		420	425	430
[1326]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro			
[1327]		435	440	445
[1328]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile			
[1329]		450	455	460
[1330]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg			
[1331]	465	470	475	480
[1332]	Glu Thr Arg Tyr Ala Thr Arg Gly Leu Met Asn Leu Leu Arg Ser Tyr			
[1333]		485	490	495
[1334]	Phe Arg Val Asn Asn Leu Asp Val Lys Val Lys Ser Ile Asn Gly Gly			
[1335]		500	505	510
[1336]	Phe Thr Ser Phe Leu Arg Arg Lys Trp Lys Phe Lys Lys Glu Arg Asn			
[1337]		515	520	525
[1338]	Lys Gly Tyr Lys His His Ala Glu Asp Ala Leu Ile Ile Ala Asn Ala			
[1339]		530	535	540
[1340]	Asp Phe Ile Phe Lys Glu Trp Lys Lys Leu Asp Lys Ala Lys Lys Val			
[1341]	545	550	555	560
[1342]	Met Glu Asn Gln Met Phe Glu Glu Lys Gln Ala Glu Ser Met Pro Glu			

[1343]		565		570		575
[1344]	Ile Glu Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile Thr Pro His Gln					
[1345]		580		585		590
[1346]	Ile Lys His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr Ser His Arg Val					
[1347]		595		600		605
[1348]	Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr Leu Tyr Ser Thr					
[1349]		610		615		620
[1350]	Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile Val Asn Asn Leu Asn Gly					
[1351]		625		630		635
[1352]	Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu Ile Asn Lys Ser					
[1353]		645		650		655
[1354]	Pro Glu Lys Leu Leu Met Tyr His His Asp Pro Gln Thr Tyr Gln Lys					
[1355]		660		665		670
[1356]	Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys Asn Pro Leu Tyr					
[1357]		675		680		685
[1358]	Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys Tyr Ser Lys Lys					
[1359]		690		695		700
[1360]	Asp Asn Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr Gly Asn Lys Leu					
[1361]		705		710		715
[1362]	Asn Ala His Leu Asp Ile Thr Asp Asp Tyr Pro Asn Ser Arg Asn Lys					
[1363]		725		730		735
[1364]	Val Val Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp Val Tyr Leu Asp					
[1365]		740		745		750
[1366]	Asn Gly Val Tyr Lys Phe Val Thr Val Lys Asn Leu Asp Val Ile Lys					
[1367]		755		760		765
[1368]	Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr Glu Glu Ala Lys					
[1369]		770		775		780
[1370]	Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile Ala Ser Phe Tyr					
[1371]		785		790		795
[1372]	Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr Arg Val Ile Gly					
[1373]		805		810		815
[1374]	Val Asn Asn Asp Leu Leu Asn Arg Ile Glu Val Asn Met Ile Asp Ile					
[1375]		820		825		830
[1376]	Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys Arg Pro Pro Arg					
[1377]		835		840		845
[1378]	Ile Ile Lys Thr Ile Ala Ser Lys Thr Gln Ser Ile Lys Lys Tyr Ser					
[1379]		850		855		860
[1380]	Thr Asp Ile Leu Gly Asn Leu Tyr Glu Val Lys Ser Lys Lys His Pro					
[1381]		865		870		875
[1382]	Gln Ile Ile Lys Lys Gly					880
[1383]		885				
[1384]	<210> 113					

[1385]	<211>	892
[1386]	<212>	PRT
[1387]	<213>	人工序列
[1388]	<220>	
[1389]	<223>	具有GGSGGS连接子的氨基酸残基 (dSaCas9的第482位至第648位氨基酸残基) 缺失突变体
[1390]	<220>	
[1391]	<221>	变体
[1392]	<222>	(10) .. (10)
[1393]	<223>	Asp残基转化成Ala残基
[1394]	<220>	
[1395]	<221>	位点
[1396]	<222>	(482) .. (487)
[1397]	<223>	GGSGGS连接子
[1398]	<400>	113
[1399]	Met	Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
[1400]	1	5 10 15
[1401]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly	
[1402]	20 25 30	
[1403]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg	
[1404]	35 40 45	
[1405]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile	
[1406]	50 55 60	
[1407]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His	
[1408]	65 70 75 80	
[1409]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu	
[1410]	85 90 95	
[1411]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu	
[1412]	100 105 110	
[1413]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr	
[1414]	115 120 125	
[1415]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala	
[1416]	130 135 140	
[1417]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys	
[1418]	145 150 155 160	
[1419]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr	
[1420]	165 170 175	
[1421]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln	
[1422]	180 185 190	
[1423]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg	
[1424]	195 200 205	
[1425]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys	

[1426]	210	215	220
[1427]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[1428]	225	230	235
[1429]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[1430]		245	250
[1431]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[1432]		260	265
[1433]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[1434]		275	280
[1435]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[1436]	290	295	300
[1437]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[1438]	305	310	315
[1439]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[1440]		325	330
[1441]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[1442]		340	345
[1443]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[1444]		355	360
[1445]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[1446]	370	375	380
[1447]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[1448]	385	390	395
[1449]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[1450]		405	410
[1451]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[1452]		420	425
[1453]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[1454]		435	440
[1455]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[1456]	450	455	460
[1457]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[1458]	465	470	475
[1459]	Glu Gly Gly Ser Gly Gly Ser Thr Arg Tyr Ala Thr Arg Gly Leu Met		
[1460]		485	490
[1461]	Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys Val		
[1462]		500	505
[1463]	Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp Lys		
[1464]		515	520
[1465]	Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp Ala		
[1466]	530	535	540
[1467]	Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys Leu		

[1468]	545	550	555	560
[1469]	Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys Gln			
[1470]		565	570	575
[1471]	Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu Ile			
[1472]		580	585	590
[1473]	Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr			
[1474]		595	600	605
[1475]	Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn			
[1476]		610	615	620
[1477]	Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile			
[1478]		625	630	635
[1479]	Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys			
[1480]		645	650	655
[1481]	Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp			
[1482]		660	665	670
[1483]	Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp			
[1484]		675	680	685
[1485]	Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu			
[1486]		690	695	700
[1487]	Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys			
[1488]		705	710	715
[1489]	Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr			
[1490]		725	730	735
[1491]	Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg			
[1492]		740	745	750
[1493]	Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys			
[1494]		755	760	765
[1495]	Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys			
[1496]		770	775	780
[1497]	Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu			
[1498]		785	790	795
[1499]	Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu			
[1500]		805	810	815
[1501]	Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu			
[1502]		820	825	830
[1503]	Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn			
[1504]		835	840	845
[1505]	Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr Gln			
[1506]		850	855	860
[1507]	Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu Val			
[1508]		865	870	875
[1509]	Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly			880

[1510]		885		890
[1511]	<210>	114		
[1512]	<211>	50		
[1513]	<212>	PRT		
[1514]	<213>	人工序列		
[1515]	<220>			
[1516]	<223>	VP64		
[1517]	<400>	114		
[1518]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu			
[1519]	1	5	10	15
[1520]	Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe			
[1521]		20	25	30
[1522]	Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp			
[1523]		35	40	45
[1524]	Met Leu			
[1525]	50			
[1526]	<210>	115		
[1527]	<211>	376		
[1528]	<212>	PRT		
[1529]	<213>	人工序列		
[1530]	<220>			
[1531]	<223>	VPH		
[1532]	<400>	115		
[1533]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu			
[1534]	1	5	10	15
[1535]	Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe			
[1536]		20	25	30
[1537]	Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp			
[1538]		35	40	45
[1539]	Met Leu Ser Ser Gly Ser Pro Lys Lys Lys Arg Lys Val Gly Ser Pro			
[1540]	50	55	60	
[1541]	Ser Gly Gln Ile Ser Asn Gln Ala Leu Ala Leu Ala Pro Ser Ser Ala			
[1542]	65	70	75	80
[1543]	Pro Val Leu Ala Gln Thr Met Val Pro Ser Ser Ala Met Val Pro Leu			
[1544]		85	90	95
[1545]	Ala Gln Pro Pro Ala Pro Ala Pro Val Leu Thr Pro Gly Pro Pro Gln			
[1546]		100	105	110
[1547]	Ser Leu Ser Ala Pro Val Pro Lys Ser Thr Gln Ala Gly Glu Gly Thr			
[1548]		115	120	125
[1549]	Leu Ser Glu Ala Leu Leu His Leu Gln Phe Asp Ala Asp Glu Asp Leu			
[1550]		130	135	140
[1551]	Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Gly Val Phe Thr Asp Leu			

[1552]	145	150	155	160
[1553]	Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Val			
[1554]		165	170	175
[1555]	Ser Met Ser His Ser Thr Ala Glu Pro Met Leu Met Glu Tyr Pro Glu			
[1556]		180	185	190
[1557]	Ala Ile Thr Arg Leu Val Thr Gly Ser Gln Arg Pro Pro Asp Pro Ala			
[1558]		195	200	205
[1559]	Pro Thr Pro Leu Gly Thr Ser Gly Leu Pro Asn Gly Leu Ser Gly Asp			
[1560]		210	215	220
[1561]	Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Ser			
[1562]	225	230	235	240
[1563]	Gln Ile Ser Ser Ser Gly Gln Gly Gly Gly Ser Gly Phe Ser Val			
[1564]		245	250	255
[1565]	Asp Thr Ser Ala Leu Leu Asp Leu Phe Ser Pro Ser Val Thr Val Pro			
[1566]		260	265	270
[1567]	Asp Met Ser Leu Pro Asp Leu Asp Ser Ser Leu Ala Ser Ile Gln Glu			
[1568]		275	280	285
[1569]	Leu Leu Ser Pro Gln Glu Pro Pro Arg Pro Pro Glu Ala Glu Asn Ser			
[1570]		290	295	300
[1571]	Ser Pro Asp Ser Gly Lys Gln Leu Val His Tyr Thr Ala Gln Pro Leu			
[1572]	305	310	315	320
[1573]	Phe Leu Leu Asp Pro Gly Ser Val Asp Thr Gly Ser Asn Asp Leu Pro			
[1574]		325	330	335
[1575]	Val Leu Phe Glu Leu Gly Glu Gly Ser Tyr Phe Ser Glu Gly Asp Gly			
[1576]		340	345	350
[1577]	Phe Ala Glu Asp Pro Thr Ile Ser Leu Leu Thr Gly Ser Glu Pro Pro			
[1578]		355	360	365
[1579]	Lys Ala Lys Asp Pro Thr Val Ser			
[1580]		370	375	
[1581]	<210> 116			
[1582]	<211> 523			
[1583]	<212> PRT			
[1584]	<213> 人工序列			
[1585]	<220>			
[1586]	<223> VPR			
[1587]	<400> 116			
[1588]	Glu Ala Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu			
[1589]	1	5	10	15
[1590]	Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu			
[1591]		20	25	30
[1592]	Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp			
[1593]		35	40	45

[1594]	Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser
[1595]	50 55 60
[1596]	Gly Ser Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu
[1597]	65 70 75 80
[1598]	Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys
[1599]	85 90 95
[1600]	Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile
[1601]	100 105 110
[1602]	Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln
[1603]	115 120 125
[1604]	Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe
[1605]	130 135 140
[1606]	Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu
[1607]	145 150 155 160
[1608]	Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro
[1609]	165 170 175
[1610]	Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro
[1611]	180 185 190
[1612]	Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys
[1613]	195 200 205
[1614]	Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu
[1615]	210 215 220
[1616]	Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp
[1617]	225 230 235 240
[1618]	Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln
[1619]	245 250 255
[1620]	Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro
[1621]	260 265 270
[1622]	Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala
[1623]	275 280 285
[1624]	Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu
[1625]	290 295 300
[1626]	Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp
[1627]	305 310 315 320
[1628]	Met Asp Phe Ser Ala Leu Leu Gly Ser Gly Ser Gly Ser Arg Asp Ser
[1629]	325 330 335
[1630]	Arg Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser
[1631]	340 345 350
[1632]	Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro
[1633]	355 360 365
[1634]	Phe His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser
[1635]	370 375 380

[1636]	Leu	Ala	Pro	Thr	Pro	Thr	Gly	Pro	Val	His	Glu	Pro	Val	Gly	Ser	Leu
[1637]	385						390				395					400
[1638]	Thr	Pro	Ala	Pro	Val	Pro	Gln	Pro	Leu	Asp	Pro	Ala	Pro	Ala	Val	Thr
[1639]							405				410					415
[1640]	Pro	Glu	Ala	Ser	His	Leu	Leu	Glu	Asp	Pro	Asp	Glu	Glu	Thr	Ser	Gln
[1641]							420				425					430
[1642]	Ala	Val	Lys	Ala	Leu	Arg	Glu	Met	Ala	Asp	Thr	Val	Ile	Pro	Gln	Lys
[1643]							435				440					445
[1644]	Glu	Glu	Ala	Ala	Ile	Cys	Gly	Gln	Met	Asp	Leu	Ser	His	Pro	Pro	Pro
[1645]							450				455					460
[1646]	Arg	Gly	His	Leu	Asp	Glu	Leu	Thr	Thr	Thr	Leu	Glu	Ser	Met	Thr	Glu
[1647]	465						470				475					480
[1648]	Asp	Leu	Asn	Leu	Asp	Ser	Pro	Leu	Thr	Pro	Glu	Leu	Asn	Glu	Ile	Leu
[1649]							485				490					495
[1650]	Asp	Thr	Phe	Leu	Asn	Asp	Glu	Cys	Leu	Leu	His	Ala	Met	His	Ile	Ser
[1651]							500				505					510
[1652]	Thr	Gly	Leu	Ser	Ile	Phe	Asp	Thr	Ser	Leu	Phe					
[1653]							515				520					
[1654]	<210> 117															
[1655]	<211> 167															
[1656]	<212> PRT															
[1657]	<213> 人工序列															
[1658]	<220>															
[1659]	<223> miniVR															
[1660]	<400> 117															
[1661]	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu
[1662]	1					5					10					15
[1663]	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe
[1664]						20					25					30
[1665]	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp
[1666]						35					40					45
[1667]	Met	Leu	Gly	Ser	Gly	Ser	Pro	Ala	Pro	Ala	Val	Thr	Pro	Glu	Ala	Ser
[1668]						50					55					60
[1669]	His	Leu	Leu	Glu	Asp	Pro	Asp	Glu	Glu	Thr	Ser	Gln	Ala	Val	Lys	Ala
[1670]	65										70					80
[1671]	Leu	Arg	Glu	Met	Ala	Asp	Thr	Val	Ile	Pro	Gln	Lys	Glu	Glu	Ala	Ala
[1672]						85					90					95
[1673]	Ile	Cys	Gly	Gln	Met	Asp	Leu	Ser	His	Pro	Pro	Pro	Arg	Gly	His	Leu
[1674]						100					105					110
[1675]	Asp	Glu	Leu	Thr	Thr	Thr	Leu	Glu	Ser	Met	Thr	Glu	Asp	Leu	Asn	Leu
[1676]						115					120					125
[1677]	Asp	Ser	Pro	Leu	Thr	Pro	Glu	Leu	Asn	Glu	Ile	Leu	Asp	Thr	Phe	Leu

[1678]	130	135	140
[1679]	Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly Leu Ser		
[1680]	145	150	155
[1681]	Ile Phe Asp Thr Ser Leu Phe		160
[1682]		165	
[1683]	<210> 118		
[1684]	<211> 140		
[1685]	<212> PRT		
[1686]	<213> 人工序列		
[1687]	<220>		
[1688]	<223> microVR		
[1689]	<400> 118		
[1690]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu		
[1691]	1	5	10
[1692]	Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe		15
[1693]		20	25
[1694]	Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp		30
[1695]		35	40
[1696]	Met Leu Gly Ser Gly Ser Arg Glu Met Ala Asp Thr Val Ile Pro Gln		45
[1697]		50	55
[1698]	Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro		60
[1699]	65	70	75
[1700]	Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr		80
[1701]		85	90
[1702]	Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile		95
[1703]		100	105
[1704]	Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile		110
[1705]		115	120
[1706]	Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe		125
[1707]		130	135
[1708]	<210> 119		140
[1709]	<211> 21		
[1710]	<212> RNA		
[1711]	<213> 智人(Homo sapiens)		
[1712]	<220>		
[1713]	<221> misc_feature		
[1714]	<222> (1) .. (21)		
[1715]	<223> 对应于靶序列 (SEQ ID NO:4) 的crRNA		
[1716]	<400> 119		
[1717]	agaaaagcgg ccccuagggg c 21		
[1718]	<210> 120		
[1719]	<211> 21		

[1720]	<212> DNA
[1721]	<213> 智人(Homo sapiens)
[1722]	<220>
[1723]	<221> misc_feature
[1724]	<222> (1) .. (21)
[1725]	<223> 与靶序列(SEQ ID NO:4)互补的序列
[1726]	<400> 120
[1727]	gcccctaggg gccgcttttc t 21
[1728]	<210> 121
[1729]	<211> 19
[1730]	<212> DNA
[1731]	<213> 新凶手弗朗西丝氏菌(Francisella novicida)
[1732]	<220>
[1733]	<221> misc_feature
[1734]	<222> (1) .. (19)
[1735]	<223> crRNA的5'-柄
[1736]	<400> 121
[1737]	aatttctact gttgtagat 19
[1738]	<210> 122
[1739]	<211> 83
[1740]	<212> DNA
[1741]	<213> 金黄色葡萄球菌(Staphylococcus aureus)
[1742]	<220>
[1743]	<221> misc_feature
[1744]	<222> (1) .. (83)
[1745]	<223> 编码tracrRNA的序列
[1746]	<400> 122
[1747]	gttttagtac tctggaaca gaatctacta aaacaaggca aaatgccgtg tttatctcgt 60
[1748]	caacttggtg gcgagat ttt 83
[1749]	<210> 123
[1750]	<211> 1053
[1751]	<212> PRT
[1752]	<213> 金黄色葡萄球菌(Staphylococcus aureus)
[1753]	<220>
[1754]	<221> 变体
[1755]	<222> (10) .. (10)
[1756]	<223> Asp残基转化成Ala残基
[1757]	<220>
[1758]	<221> 变体
[1759]	<222> (580) .. (580)
[1760]	<223> Asn残基转化成Ala残基
[1761]	<400> 123

[1762]	Met	Lys	Arg	Asn	Tyr	Ile	Leu	Gly	Leu	Ala	Ile	Gly	Ile	Thr	Ser	Val
[1763]	1			5					10					15		
[1764]	Gly	Tyr	Gly	Ile	Ile	Asp	Tyr	Glu	Thr	Arg	Asp	Val	Ile	Asp	Ala	Gly
[1765]				20					25					30		
[1766]	Val	Arg	Leu	Phe	Lys	Glu	Ala	Asn	Val	Glu	Asn	Asn	Glu	Gly	Arg	Arg
[1767]			35					40					45			
[1768]	Ser	Lys	Arg	Gly	Ala	Arg	Arg	Leu	Lys	Arg	Arg	Arg	Arg	His	Arg	Ile
[1769]			50					55					60			
[1770]	Gln	Arg	Val	Lys	Lys	Leu	Leu	Phe	Asp	Tyr	Asn	Leu	Leu	Thr	Asp	His
[1771]	65					70					75				80	
[1772]	Ser	Glu	Leu	Ser	Gly	Ile	Asn	Pro	Tyr	Glu	Ala	Arg	Val	Lys	Gly	Leu
[1773]					85					90					95	
[1774]	Ser	Gln	Lys	Leu	Ser	Glu	Glu	Glu	Phe	Ser	Ala	Ala	Leu	Leu	His	Leu
[1775]				100					105						110	
[1776]	Ala	Lys	Arg	Arg	Gly	Val	His	Asn	Val	Asn	Glu	Val	Glu	Glu	Asp	Thr
[1777]				115				120						125		
[1778]	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Ile	Ser	Arg	Asn	Ser	Lys	Ala
[1779]			130					135						140		
[1780]	Leu	Glu	Glu	Lys	Tyr	Val	Ala	Glu	Leu	Gln	Leu	Glu	Arg	Leu	Lys	Lys
[1781]	145					150					155				160	
[1782]	Asp	Gly	Glu	Val	Arg	Gly	Ser	Ile	Asn	Arg	Phe	Lys	Thr	Ser	Asp	Tyr
[1783]				165						170					175	
[1784]	Val	Lys	Glu	Ala	Lys	Gln	Leu	Leu	Lys	Val	Gln	Lys	Ala	Tyr	His	Gln
[1785]				180						185					190	
[1786]	Leu	Asp	Gln	Ser	Phe	Ile	Asp	Thr	Tyr	Ile	Asp	Leu	Leu	Glu	Thr	Arg
[1787]			195					200						205		
[1788]	Arg	Thr	Tyr	Tyr	Glu	Gly	Pro	Gly	Glu	Gly	Ser	Pro	Phe	Gly	Trp	Lys
[1789]		210					215					220				
[1790]	Asp	Ile	Lys	Glu	Trp	Tyr	Glu	Met	Leu	Met	Gly	His	Cys	Thr	Tyr	Phe
[1791]	225					230					235				240	
[1792]	Pro	Glu	Glu	Leu	Arg	Ser	Val	Lys	Tyr	Ala	Tyr	Asn	Ala	Asp	Leu	Tyr
[1793]				245						250					255	
[1794]	Asn	Ala	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Val	Ile	Thr	Arg	Asp	Glu	Asn
[1795]			260						265						270	
[1796]	Glu	Lys	Leu	Glu	Tyr	Tyr	Glu	Lys	Phe	Gln	Ile	Ile	Glu	Asn	Val	Phe
[1797]			275						280					285		
[1798]	Lys	Gln	Lys	Lys	Lys	Pro	Thr	Leu	Lys	Gln	Ile	Ala	Lys	Glu	Ile	Leu
[1799]			290					295						300		
[1800]	Val	Asn	Glu	Glu	Asp	Ile	Lys	Gly	Tyr	Arg	Val	Thr	Ser	Thr	Gly	Lys
[1801]	305					310					315				320	
[1802]	Pro	Glu	Phe	Thr	Asn	Leu	Lys	Val	Tyr	His	Asp	Ile	Lys	Asp	Ile	Thr
[1803]				325						330					335	

[1804]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[1805]		340	345 350
[1806]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[1807]		355	360 365
[1808]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[1809]		370	375 380
[1810]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[1811]		385	390 395 400
[1812]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[1813]		405	410 415
[1814]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[1815]		420	425 430
[1816]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[1817]		435	440 445
[1818]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[1819]		450	455 460
[1820]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[1821]		465	470 475 480
[1822]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[1823]		485	490 495
[1824]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[1825]		500	505 510
[1826]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[1827]		515	520 525
[1828]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[1829]		530	535 540
[1830]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[1831]		545	550 555 560
[1832]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[1833]		565	570 575
[1834]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[1835]		580	585 590
[1836]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[1837]		595	600 605
[1838]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[1839]		610	615 620
[1840]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[1841]		625	630 635 640
[1842]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[1843]		645	650 655
[1844]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[1845]		660	665 670

[1846]	Val	Lys	Ser	Ile	Asn	Gly	Gly	Phe	Thr	Ser	Phe	Leu	Arg	Arg	Lys	Trp
[1847]				675				680					685			
[1848]	Lys	Phe	Lys	Lys	Glu	Arg	Asn	Lys	Gly	Tyr	Lys	His	His	Ala	Glu	Asp
[1849]			690				695					700				
[1850]	Ala	Leu	Ile	Ile	Ala	Asn	Ala	Asp	Phe	Ile	Phe	Lys	Glu	Trp	Lys	Lys
[1851]	705					710					715					720
[1852]	Leu	Asp	Lys	Ala	Lys	Lys	Val	Met	Glu	Asn	Gln	Met	Phe	Glu	Glu	Lys
[1853]					725					730					735	
[1854]	Gln	Ala	Glu	Ser	Met	Pro	Glu	Ile	Glu	Thr	Glu	Gln	Glu	Tyr	Lys	Glu
[1855]				740						745					750	
[1856]	Ile	Phe	Ile	Thr	Pro	His	Gln	Ile	Lys	His	Ile	Lys	Asp	Phe	Lys	Asp
[1857]			755							760					765	
[1858]	Tyr	Lys	Tyr	Ser	His	Arg	Val	Asp	Lys	Lys	Pro	Asn	Arg	Glu	Leu	Ile
[1859]			770							775					780	
[1860]	Asn	Asp	Thr	Leu	Tyr	Ser	Thr	Arg	Lys	Asp	Asp	Lys	Gly	Asn	Thr	Leu
[1861]	785					790						795				800
[1862]	Ile	Val	Asn	Asn	Leu	Asn	Gly	Leu	Tyr	Asp	Lys	Asp	Asn	Asp	Lys	Leu
[1863]					805						810					815
[1864]	Lys	Lys	Leu	Ile	Asn	Lys	Ser	Pro	Glu	Lys	Leu	Leu	Met	Tyr	His	His
[1865]					820						825				830	
[1866]	Asp	Pro	Gln	Thr	Tyr	Gln	Lys	Leu	Lys	Leu	Ile	Met	Glu	Gln	Tyr	Gly
[1867]			835								840				845	
[1868]	Asp	Glu	Lys	Asn	Pro	Leu	Tyr	Lys	Tyr	Tyr	Glu	Glu	Thr	Gly	Asn	Tyr
[1869]		850									855				860	
[1870]	Leu	Thr	Lys	Tyr	Ser	Lys	Lys	Asp	Asn	Gly	Pro	Val	Ile	Lys	Lys	Ile
[1871]	865						870					875				880
[1872]	Lys	Tyr	Tyr	Gly	Asn	Lys	Leu	Asn	Ala	His	Leu	Asp	Ile	Thr	Asp	Asp
[1873]					885						890					895
[1874]	Tyr	Pro	Asn	Ser	Arg	Asn	Lys	Val	Val	Lys	Leu	Ser	Leu	Lys	Pro	Tyr
[1875]				900							905					910
[1876]	Arg	Phe	Asp	Val	Tyr	Leu	Asp	Asn	Gly	Val	Tyr	Lys	Phe	Val	Thr	Val
[1877]			915												925	
[1878]	Lys	Asn	Leu	Asp	Val	Ile	Lys	Lys	Glu	Asn	Tyr	Tyr	Glu	Val	Asn	Ser
[1879]		930													935	
[1880]	Lys	Cys	Tyr	Glu	Glu	Ala	Lys	Lys	Leu	Lys	Lys	Ile	Ser	Asn	Gln	Ala
[1881]	945											950				955
[1882]	Glu	Phe	Ile	Ala	Ser	Phe	Tyr	Asn	Asn	Asp	Leu	Ile	Lys	Ile	Asn	Gly
[1883]						965										970
[1884]	Glu	Leu	Tyr	Arg	Val	Ile	Gly	Val	Asn	Asn	Asp	Leu	Leu	Asn	Arg	Ile
[1885]						980										985
[1886]	Glu	Val	Asn	Met	Ile	Asp	Ile	Thr	Tyr	Arg	Glu	Tyr	Leu	Glu	Asn	Met
[1887]				995												1000

[1888]	Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
[1889]	1010 1015 1020
[1890]	Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
[1891]	1025 1030 1035
[1892]	Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
[1893]	1040 1045 1050
[1894]	<210> 124
[1895]	<211> 76
[1896]	<212> DNA
[1897]	<213> 金黄色葡萄球菌(Staphylococcus aureus)
[1898]	<220>
[1899]	<221> misc_feature
[1900]	<222> (1) .. (76)
[1901]	<223> SaCas9 gRNA骨架序列
[1902]	<400> 124
[1903]	gttttagtac tctggaaaca gaatctacta aaacaaggca aaatgccgtg tttatctcgt 60
[1904]	caacttggtg gcgaga 76
[1905]	<210> 125
[1906]	<211> 501
[1907]	<212> DNA
[1908]	<213> 人工序列
[1909]	<220>
[1910]	<223> miniVR DNA
[1911]	<400> 125
[1912]	gatgcactcg atgattttga cctcgatatg cttgggagtg atgcgctcga tgacttcgat 60
[1913]	ttggatatgc ttggatctga tgccctcgac gatttcgacc ttgatatgct cgggtcagac 120
[1914]	gctttggatg actttgacct tgacatgctg gggagcggct cccccgcacc ggcagttaca 180
[1915]	ccggaggcca gtcacttgct cgaagatcct gacgaggaaa ccagccaggc cgtaaaggcg 240
[1916]	ttgcgggaga tggtgacac agtaataccc caaaaagagg aggctgcgat ttgtgggcag 300
[1917]	atggatttgt cccaccctcc accgagaggt catcttgacg aattgacaac gacgctcgaa 360
[1918]	tccatgaccg aggacctgaa cctcgatagc ccgctcacc cagagttgaa tgagatcctg 420
[1919]	gatacatctt ttaatgatga gtgtttgctt cagcaatgc atatttctac gggctcttagt 480
[1920]	atcttcgaca cgagcctgtt t 501
[1921]	<210> 126
[1922]	<211> 167
[1923]	<212> PRT
[1924]	<213> 人工序列
[1925]	<220>
[1926]	<223> miniVR蛋白
[1927]	<400> 126
[1928]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
[1929]	1 5 10 15

[1930] Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 [1931] 20 25 30
 [1932] Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 [1933] 35 40 45
 [1934] Met Leu Gly Ser Gly Ser Pro Ala Pro Ala Val Thr Pro Glu Ala Ser
 [1935] 50 55 60
 [1936] His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala Val Lys Ala
 [1937] 65 70 75 80
 [1938] Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu Glu Ala Ala
 [1939] 85 90 95
 [1940] Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg Gly His Leu
 [1941] 100 105 110
 [1942] Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu
 [1943] 115 120 125
 [1944] Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu
 [1945] 130 135 140
 [1946] Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly Leu Ser
 [1947] 145 150 155 160
 [1948] Ile Phe Asp Thr Ser Leu Phe
 [1949] 165
 [1950] <210> 127
 [1951] <211> 82
 [1952] <212> RNA
 [1953] <213> 金黄色葡萄球菌 (Staphylococcus aureus)
 [1954] <220>
 [1955] <221> misc_feature
 [1956] <222> (1) .. (82)
 [1957] <223> tracrRNA
 [1958] <400> 127
 [1959] guuuuaguac ucuggaaaca gaauacuacua aaacaaggca aaaugccgug uuuaucucgu 60
 [1960] caacuuguug gcgagauuuu uu 82
 [1961] <210> 128
 [1962] <211> 241
 [1963] <212> DNA
 [1964] <213> 人工序列
 [1965] <220>
 [1966] <223> U6启动子
 [1967] <400> 128
 [1968] gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60
 [1969] ataattagaa ttaatttgac tgtaaacaca aagatattag taaaaatac gtgacgtaga 120
 [1970] aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180
 [1971] atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

[1972] c 241
[1973] <210> 129
[1974] <211> 21
[1975] <212> DNA
[1976] <213> 智人 (Homo sapiens)
[1977] <400> 129
[1978] cttgttaa at gaatgaatga a 21
[1979] <210> 130
[1980] <211> 21
[1981] <212> DNA
[1982] <213> 智人 (Homo sapiens)
[1983] <400> 130
[1984] tgtcctagaa accttacaag g 21
[1985] <210> 131
[1986] <211> 21
[1987] <212> DNA
[1988] <213> 智人 (Homo sapiens)
[1989] <400> 131
[1990] ggtttattgc tggcttaata t 21
[1991] <210> 132
[1992] <211> 21
[1993] <212> DNA
[1994] <213> 智人 (Homo sapiens)
[1995] <400> 132
[1996] acgtcagcaa actgagatgg g 21
[1997] <210> 133
[1998] <211> 21
[1999] <212> DNA
[2000] <213> 智人 (Homo sapiens)
[2001] <400> 133
[2002] ttttcggata atctgaataa g 21
[2003] <210> 134
[2004] <211> 21
[2005] <212> DNA
[2006] <213> 智人 (Homo sapiens)
[2007] <400> 134
[2008] ggggtccgct ctccagatga g 21
[2009] <210> 135
[2010] <211> 21
[2011] <212> DNA
[2012] <213> 智人 (Homo sapiens)
[2013] <400> 135

[2014] ggctcctcta ggagtttgac a 21
[2015] <210> 136
[2016] <211> 21
[2017] <212> DNA
[2018] <213> 智人(Homo sapiens)
[2019] <400> 136
[2020] taatgtgact acagcccccg a 21
[2021] <210> 137
[2022] <211> 21
[2023] <212> DNA
[2024] <213> 智人(Homo sapiens)
[2025] <400> 137
[2026] ccaagtccca gagtgaaga t 21
[2027] <210> 138
[2028] <211> 21
[2029] <212> DNA
[2030] <213> 智人(Homo sapiens)
[2031] <400> 138
[2032] tcagttgcag caagagatcc c 21
[2033] <210> 139
[2034] <211> 21
[2035] <212> DNA
[2036] <213> 智人(Homo sapiens)
[2037] <400> 139
[2038] cctcctcctc gaaaaacgca c 21
[2039] <210> 140
[2040] <211> 21
[2041] <212> DNA
[2042] <213> 智人(Homo sapiens)
[2043] <400> 140
[2044] gggagggtcg gctcagacct a 21
[2045] <210> 141
[2046] <211> 21
[2047] <212> DNA
[2048] <213> 智人(Homo sapiens)
[2049] <400> 141
[2050] gggtagttct gcggtgacgg a 21
[2051] <210> 142
[2052] <211> 21
[2053] <212> DNA
[2054] <213> 智人(Homo sapiens)
[2055] <400> 142

[2056] attttaggta aacacccaaa g 21
[2057] <210> 143
[2058] <211> 21
[2059] <212> DNA
[2060] <213> 智人 (Homo sapiens)
[2061] <400> 143
[2062] gaaacacagt aaaagaaaac g 21
[2063] <210> 144
[2064] <211> 21
[2065] <212> DNA
[2066] <213> 智人 (Homo sapiens)
[2067] <400> 144
[2068] taagatttta ggaattatac a 21
[2069] <210> 145
[2070] <211> 21
[2071] <212> DNA
[2072] <213> 智人 (Homo sapiens)
[2073] <400> 145
[2074] agcgttctga agggagagtt a 21
[2075] <210> 146
[2076] <211> 21
[2077] <212> DNA
[2078] <213> 智人 (Homo sapiens)
[2079] <400> 146
[2080] cagaaggcta ggtgagaaac t 21
[2081] <210> 147
[2082] <211> 21
[2083] <212> DNA
[2084] <213> 智人 (Homo sapiens)
[2085] <400> 147
[2086] aatttgagta cacttaaggc a 21
[2087] <210> 148
[2088] <211> 21
[2089] <212> DNA
[2090] <213> 智人 (Homo sapiens)
[2091] <400> 148
[2092] agatacagca gaaaaggtga t 21
[2093] <210> 149
[2094] <211> 21
[2095] <212> DNA
[2096] <213> 智人 (Homo sapiens)
[2097] <400> 149

[2098] gacacatgca gaagtgcag c 21
[2099] <210> 150
[2100] <211> 21
[2101] <212> DNA
[2102] <213> 智人 (Homo sapiens)
[2103] <400> 150
[2104] agcagccttc gaactgcaca c 21
[2105] <210> 151
[2106] <211> 21
[2107] <212> DNA
[2108] <213> 智人 (Homo sapiens)
[2109] <400> 151
[2110] tctagatggc agtaaacagc a 21
[2111] <210> 152
[2112] <211> 21
[2113] <212> DNA
[2114] <213> 智人 (Homo sapiens)
[2115] <400> 152
[2116] ggctgctcca atcatTTtTg t 21
[2117] <210> 153
[2118] <211> 21
[2119] <212> DNA
[2120] <213> 智人 (Homo sapiens)
[2121] <400> 153
[2122] gagtccggag accgaaccag a 21
[2123] <210> 154
[2124] <211> 21
[2125] <212> DNA
[2126] <213> 智人 (Homo sapiens)
[2127] <400> 154
[2128] gaaccgtgcg tgccgggagc c 21
[2129] <210> 155
[2130] <211> 21
[2131] <212> DNA
[2132] <213> 智人 (Homo sapiens)
[2133] <400> 155
[2134] gctggcctgg ggcgcgcgct c 21
[2135] <210> 156
[2136] <211> 21
[2137] <212> DNA
[2138] <213> 智人 (Homo sapiens)
[2139] <400> 156

[2140] aagatcagcc ccactacgtt c 21
[2141] <210> 157
[2142] <211> 21
[2143] <212> DNA
[2144] <213> 智人 (Homo sapiens)
[2145] <400> 157
[2146] ccggaggcga gccccttccc g 21
[2147] <210> 158
[2148] <211> 21
[2149] <212> DNA
[2150] <213> 智人 (Homo sapiens)
[2151] <400> 158
[2152] ggagggtggg gcgcaggacc g 21
[2153] <210> 159
[2154] <211> 21
[2155] <212> DNA
[2156] <213> 智人 (Homo sapiens)
[2157] <400> 159
[2158] gagcgctgga ggcgaggag g 21
[2159] <210> 160
[2160] <211> 21
[2161] <212> DNA
[2162] <213> 智人 (Homo sapiens)
[2163] <400> 160
[2164] cctctctcgc gcacaaagtt g 21
[2165] <210> 161
[2166] <211> 21
[2167] <212> DNA
[2168] <213> 智人 (Homo sapiens)
[2169] <400> 161
[2170] gggagcggcg ccccccttct t 21
[2171] <210> 162
[2172] <211> 21
[2173] <212> DNA
[2174] <213> 智人 (Homo sapiens)
[2175] <400> 162
[2176] caccaacttt gccaaacgct a 21
[2177] <210> 163
[2178] <211> 21
[2179] <212> DNA
[2180] <213> 智人 (Homo sapiens)
[2181] <400> 163

[2182] ggagtaaccg cgggggtgtg t 21
[2183] <210> 164
[2184] <211> 21
[2185] <212> DNA
[2186] <213> 智人 (Homo sapiens)
[2187] <400> 164
[2188] gaatggggcg ggggccggga g 21
[2189] <210> 165
[2190] <211> 21
[2191] <212> DNA
[2192] <213> 智人 (Homo sapiens)
[2193] <400> 165
[2194] tctttctgtg gttcttccgc c 21
[2195] <210> 166
[2196] <211> 21
[2197] <212> DNA
[2198] <213> 智人 (Homo sapiens)
[2199] <400> 166
[2200] tttggatcgt tcacaactag t 21
[2201] <210> 167
[2202] <211> 21
[2203] <212> DNA
[2204] <213> 智人 (Homo sapiens)
[2205] <400> 167
[2206] agaggggacg tggcctctta g 21
[2207] <210> 168
[2208] <211> 21
[2209] <212> DNA
[2210] <213> 智人 (Homo sapiens)
[2211] <400> 168
[2212] gtccacagga gaggtgggc a 21
[2213] <210> 169
[2214] <211> 21
[2215] <212> DNA
[2216] <213> 智人 (Homo sapiens)
[2217] <400> 169
[2218] gctcccaagg gtggggctcc g 21
[2219] <210> 170
[2220] <211> 21
[2221] <212> DNA
[2222] <213> 智人 (Homo sapiens)
[2223] <400> 170

[2224] tttcagatgg caggttggtc a 21
[2225] <210> 171
[2226] <211> 21
[2227] <212> DNA
[2228] <213> 智人 (Homo sapiens)
[2229] <400> 171
[2230] ctttcccagc cttcaggtca g 21
[2231] <210> 172
[2232] <211> 21
[2233] <212> DNA
[2234] <213> 智人 (Homo sapiens)
[2235] <400> 172
[2236] gcgcgcggag ctcgggggag g 21
[2237] <210> 173
[2238] <211> 21
[2239] <212> DNA
[2240] <213> 智人 (Homo sapiens)
[2241] <400> 173
[2242] tgaggccggt gcaacttaca a 21
[2243] <210> 174
[2244] <211> 21
[2245] <212> DNA
[2246] <213> 智人 (Homo sapiens)
[2247] <400> 174
[2248] tgggcgtggg agacgcagcc t 21
[2249] <210> 175
[2250] <211> 21
[2251] <212> DNA
[2252] <213> 智人 (Homo sapiens)
[2253] <400> 175
[2254] aggtggagga atgcgaagct t 21
[2255] <210> 176
[2256] <211> 21
[2257] <212> DNA
[2258] <213> 智人 (Homo sapiens)
[2259] <400> 176
[2260] agacaactct ttaactctcc t 21
[2261] <210> 177
[2262] <211> 21
[2263] <212> DNA
[2264] <213> 人工序列
[2265] <220>

- [2266] <223> 编码对照sgRNA的crRNA的序列
- [2267] <400> 177
- [2268] acggaggcta agcgtcgcaa g 21
- [2269] <210> 178
- [2270] <211> 16
- [2271] <212> PRT
- [2272] <213> 人工序列
- [2273] <220>
- [2274] <223> 核定位信号
- [2275] <400> 178
- [2276] Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala Ala
- [2277] 1 5 10 15
- [2278] <210> 179
- [2279] <211> 48
- [2280] <212> DNA
- [2281] <213> 人工序列
- [2282] <220>
- [2283] <223> 编码SEQ ID NO 178的NLS的DNA序列
- [2284] <400> 179
- [2285] gccccaaaga agaagcggaa ggtcggatc cacggagtcc cagcagcc 48
- [2286] <210> 180
- [2287] <211> 16
- [2288] <212> PRT
- [2289] <213> 人工序列
- [2290] <220>
- [2291] <223> 核定位信号
- [2292] <400> 180
- [2293] Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
- [2294] 1 5 10 15
- [2295] <210> 181
- [2296] <211> 48
- [2297] <212> DNA
- [2298] <213> 人工序列
- [2299] <220>
- [2300] <223> 编码SEQ ID NO 180的NLS的DNA序列
- [2301] <400> 181
- [2302] aaaaggccgg cggccacgaa aaaggccggc caggcaaaaa agaaaaag 48
- [2303] <210> 182
- [2304] <211> 3414
- [2305] <212> DNA
- [2306] <213> 人工序列
- [2307] <220>

[2308]	<223> 编码dSaCas9 (D10A, N580A) -VP64的DNA序列					
[2309]	<400> 182					
[2310]	atggcccaaa	agaagaagcg	gaaggtcgg	atccacggag	tcccagcagc	caagcggaaac 60
[2311]	tacatcctgg	gcctggccat	cggcatcacc	agcgtgggct	acggcatcat	cgactacgag 120
[2312]	acacgggacg	tgatcgatgc	cggcgtgcgg	ctgttcaaag	aggccaacgt	ggaaaacaac 180
[2313]	gagggcaggc	ggagcaagag	aggcgccaga	aggctgaagc	ggcggaggcg	gcatagaatc 240
[2314]	cagagagtga	agaagctgct	gttcgactac	aacctgctga	ccgaccacag	cgagctgagc 300
[2315]	ggcatcaacc	cctacgaggc	cagagtgaag	ggcctgagcc	agaagctgag	cgaggaagag 360
[2316]	ttctctgccg	ccctgctgca	cctggccaag	agaagaggcg	tgcacaacgt	gaacgagggtg 420
[2317]	gaagaggaca	ccggcaacga	gctgtccacc	aaagagcaga	tcagccggaa	cagcaaggcc 480
[2318]	ctggaagaga	aatacgtggc	cgaactgcag	ctggaacggc	tgaagaaaga	cggcgaagtg 540
[2319]	cggggcagca	tcaacagatt	caagaccagc	gactacgtga	aagaagccaa	acagctgctg 600
[2320]	aaggtgcaga	aggcctacca	ccagctggac	cagagcttca	tcgacaccta	catcgacctg 660
[2321]	ctggaaaccc	ggcggaccta	ctatgaggga	cctggcgagg	gcagcccctt	cggctggaag 720
[2322]	gacatcaaag	aatggtacga	gatgctgatg	ggccactgca	cctacttccc	cgaggaactg 780
[2323]	cggagcgtga	agtacgcta	caacgccgac	ctgtacaacg	ccctgaacga	cctgaacaat 840
[2324]	ctcgtgatca	ccaggacga	gaacgagaag	ctggaatatt	acgagaagtt	ccagatcatc 900
[2325]	gagaacgtgt	tcaagcagaa	gaagaagccc	accctgaagc	agatcgccaa	agaaatcctc 960
[2326]	gtgaacgaag	aggatattaa	gggctacaga	gtgaccagca	ccggcaagcc	cgagttcacc 1020
[2327]	aacctgaagg	tgtaccacga	catcaaggac	attaccgccc	ggaaagagat	tattgagaac 1080
[2328]	gccgagctgc	tggatcagat	tgccaagatc	ctgaccatct	accagagcag	cgaggacatc 1140
[2329]	caggaagaac	tgaccaatct	gaactccgag	ctgaccagg	aagagatcga	gcagatctct 1200
[2330]	aatctgaagg	gctataccgg	caccacaac	ctgagcctga	aggccatcaa	cctgatcctg 1260
[2331]	gacgagctgt	ggcacacca	cgacaaccag	atcgctatct	tcaaccggct	gaagctgggtg 1320
[2332]	cccaagaagg	tggacctgtc	ccagcagaaa	gagatcccca	ccacctgggt	ggacgacttc 1380
[2333]	atcctgagcc	ccgtcgtgaa	gagaagcttc	atccagagca	tcaaagtgat	caacgccatc 1440
[2334]	atcaagaagt	acggcctgcc	caacgacatc	attatcgagc	tggcccgcga	gaagaactcc 1500
[2335]	aaggacgccc	agaaaatgat	caacgagatg	cagaagcgga	accggcagac	caacgagcgg 1560
[2336]	atcgaggaaa	tcatccggac	caccggcaaa	gagaacgcca	agtacctgat	cgagaagatc 1620
[2337]	aagctgcacg	acatgcagga	aggcaagtgc	ctgtacagcc	tggaagccat	ccctctggaa 1680
[2338]	gatctgctga	acaaccctt	caactatgag	gtggaccaca	tcatccccag	aagcgtgtcc 1740
[2339]	ttcgacaaca	gcttcaacaa	caaggtgctc	gtgaagcagg	aagaagccag	caagaagggc 1800
[2340]	aaccggaccc	cattccagta	cctgagcagc	agcgacagca	agatcagcta	cgaaaccttc 1860
[2341]	aagaagcaca	tcttgaatct	ggccaagggc	aagggcagaa	tcagcaagac	caagaaagag 1920
[2342]	tatctgctgg	aagaacggga	catcaacagg	ttctccgtgc	agaaagactt	catcaaccgg 1980
[2343]	aacctgggtg	ataccagata	cgccaccaga	ggcctgatga	acctgctgcg	gagctacttc 2040
[2344]	agagtgaaca	acctggacgt	gaaagtgaag	tccatcaatg	gcggcttcac	cagctttctg 2100
[2345]	cggcggaagt	ggaagttaa	gaaagagcgg	aacaaggggt	acaagcacca	cgccgaggac 2160
[2346]	gccctgatca	ttgccaacgc	cgatttcate	ttcaaagagt	ggaagaaact	ggacaaggcc 2220
[2347]	aaaaaagtga	tggaaaacca	gatgttcgag	gaaaagcagg	ccgagagcat	gcccagatc 2280
[2348]	gaaaccgagc	aggagtacaa	agagatcttc	atcaccccc	accagatcaa	gcacattaag 2340
[2349]	gacttcaagg	actacaagta	cagccaccgg	gtggacaaga	agcctaatag	agagctgatt 2400

[2350]	aacgacaccc tgtactccac ccggaaggac gacaaggga acaccctgat cgtgaacaat	2460
[2351]	ctgaacggcc tgtacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc	2520
[2352]	gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg	2580
[2353]	gaacagtacg gcgacgagaa gaatccccctg tacaagtact acgaggaaac cgggaactac	2640
[2354]	ctgaccaagt actccaaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc	2700
[2355]	aacaaactga acgcccattt ggacatcacc gacgactacc ccaacagcag aaacaaggctc	2760
[2356]	gtgaagctgt ccctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
[2357]	ttcgtgaccg tgaagaattt ggatgtgatc aaaaaagaaa actactacga agtgaatagc	2880
[2358]	aagtgtctatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940
[2359]	tccttctaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg	3000
[2360]	aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac	3060
[2361]	ctggaaaaca tgaacgacaa gaggcccccc aggatcatta agacaatcgc ctccaagacc	3120
[2362]	cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag	3180
[2363]	aagcaccctc agatcatcaa aaagggcaaa aggccggcgg ccacgaaaaa ggccggccag	3240
[2364]	gcaaaaaaga aaaagtctag agatgcttta gacgattttg acttagatat gcttggttca	3300
[2365]	gacgcgttag acgacttga cctagacatg ttaggtcag atgcattgga cgacttcgat	3360
[2366]	ttagatatgt tgggctccga tgccctagat gactttgatt tggatatgct ataa	3414
[2367]	<210>	183
[2368]	<211>	4392
[2369]	<212>	DNA
[2370]	<213>	人工序列
[2371]	<220>	
[2372]	<223>	编码dSaCas9 (D10A, N580A) -VPH的DNA序列
[2373]	<400>	183
[2374]	atggcccaaa agaagaagcg gaaggtcggc atccacggag tcccagcagc caagcggaac	60
[2375]	tacatcctgg gcctggccat cggcatcacc agcgtgggct acggcatcat cgactacgag	120
[2376]	acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac	180
[2377]	gaggcgagcg ggagcaagag aggcgccaga aggctgaagc ggcgaggcgc gcatagaatc	240
[2378]	cagagagtga agaagctgct gttcgactac aacctgctga ccgaccacag cgagctgagc	300
[2379]	ggcatcaacc cctacgagcg cagagtgaag ggcctgagcc agaagctgag cgaggaagag	360
[2380]	ttctctgccg ccctgtgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
[2381]	gaagaggaca ccggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
[2382]	ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagt	540
[2383]	cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
[2384]	aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
[2385]	ctggaaaccc ggcggaacct ctatgaggga cctggcgagg gcagccccctt cggctggaag	720
[2386]	gacatcaaag aatggtacga gatgctgatg ggccactgca cctacttccc cgaggaactg	780
[2387]	cggagcgtga agtacgcta caacgccgac ctgtacaacg ccctgaacga cctgaacaat	840
[2388]	ctcgtgatca ccaggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
[2389]	gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatcctc	960
[2390]	gtgaacgaag aggatattaa gggctacaga gtgaccagca ccggcaagcc cgagttcacc	1020
[2391]	aacctgaagg tgtaccacga catcaaggac attaccgccc ggaaagagat tattgagaac	1080

[2392]	gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc	1140
[2393]	caggaagaac tgaccaatct gaactccgag ctgaccagg aagagatcga gcagatctct	1200
[2394]	aatctgaagg gctataccgg caccacaac ctgagcctga aggccatcaa cctgatcctg	1260
[2395]	gacgagctgt ggcacaccaa cgacaaccag atcgtatct tcaaccggct gaagctggtg	1320
[2396]	cccaagaagg tggacctgtc ccagcagaaa gagatcccca ccaccctggt ggacgacttc	1380
[2397]	atcctgagcc ccgtcgtgaa gagaagcttc atccagagca tcaaagtgat caacgccatc	1440
[2398]	atcaagaagt acggcctgcc caacgacatc attatcgagc tggcccgcga gaagaactcc	1500
[2399]	aaggacgccc agaaaatgat caacgagatg cagaagcgga accggcagac caacgagcgg	1560
[2400]	atcgaggaaa tcatccggac caccggcaaa gagaacgcca agtacctgat cgagaagatc	1620
[2401]	aagctgcacg acatgcagga aggcaagtgc ctgtacagcc tggaagccat ccctctggaa	1680
[2402]	gatctgctga acaacccctt caactatgag gtggaccaca tcatccccag aagcgtgtcc	1740
[2403]	ttcgacaaca gttcaacaa caaggtgctc gtgaagcagg aagaagccag caagaagggc	1800
[2404]	aaccggaccc cattccagta cctgagcagc agcgacagca agatcagcta cgaaaccttc	1860
[2405]	aagaagcaca tcctgaatct ggccaagggc aagggcagaa tcagcaagac caagaaagag	1920
[2406]	tatctgctgg aagaacggga catcaacagg ttctccgtgc agaaagactt catcaaccgg	1980
[2407]	aacctggtgg ataccagata cgccaccaga ggctgatga acctgctgcg gagctacttc	2040
[2408]	agagtgaaca acctggacgt gaaagtgaag tccatcaatg gcggcttcac cagctttctg	2100
[2409]	cggcggaagt ggaagtttaa gaaagagcgg aacaaggggt acaagcacca cgccgaggac	2160
[2410]	gccctgatca ttgccaacgc cgatttcac ttcaaagagt ggaagaaact ggacaaggcc	2220
[2411]	aaaaaagtga tggaaaacca gatgttcgag gaaaagcagg ccgagagcat gcccagatc	2280
[2412]	gaaaccgagc aggagtacaa agagatcttc atcaccccc accagatcaa gcacattaag	2340
[2413]	gacttcaagg actacaagta cagccaccgg gtggacaaga agcctaatag agagctgatt	2400
[2414]	aacgacaccc tgtactccac ccggaaggac gacaagggca acaccctgat cgtgaacaat	2460
[2415]	ctgaacggcc tgtacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc	2520
[2416]	gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg	2580
[2417]	gaacagtacg gcgacgagaa gaatccccg tacaagtact acgaggaaac cgggaactac	2640
[2418]	ctgaccaagt actccaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc	2700
[2419]	aacaaactga acgcccctct ggacatcacc gacgactacc ccaacagcag aaacaaggctc	2760
[2420]	gtgaagctgt ccctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
[2421]	ttcgtgaccg tgaagaatct ggatgtgatc aaaaaagaaa actactacga agtgaatagc	2880
[2422]	aagtgtatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940
[2423]	tccttctaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg	3000
[2424]	aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac	3060
[2425]	ctggaaaaca tgaacgacaa gagggcccc aggatcatta agacaatcgc ctccaagacc	3120
[2426]	cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag	3180
[2427]	aagcaccctc agatcatcaa aaagggcaaa aggcggcgccg ccacgaaaaa ggccggccag	3240
[2428]	gcaaaaaaga aaaagtctag agatgcttta gacgattttg acttagatat gcttggttca	3300
[2429]	gacgcgttag acgacttga cctagacatg ttaggtcag atgcattgga cgacttcgat	3360
[2430]	ttagatatgt tgggtccga tgccctagat gactttgatt tggatatgct aagttccgga	3420
[2431]	tctccgaaaa agaaacgcaa agttggtagc cttcagggc agatcagcaa ccaggccctg	3480
[2432]	gctctggccc ctagctccgc tccagtgtg gccagacta tgggtgccctc tagtgctatg	3540
[2433]	gtgcctctgg ccagccacc tgctccagcc cctgtgtgta cccagggacc accccagtca	3600

[2434]	ctgagcgctc cagtgcccaa gtctacacag gccggcgagg ggactctgag tgaagctctg	3660
[2435]	ctgcacctgc agttcgacgc tgatgaggac ctgggagctc tgctggggaa cagcaccgat	3720
[2436]	cccggagtgt tcacagacct ggctccgtg gacaactctg agtttcagca gctgctgaat	3780
[2437]	cagggcgtgt ccatgtctca tagtacagcc gaaccaatgc tgatggagta ccccgagcc	3840
[2438]	attaccggc tggtagccgg cagccagcgg ccccccagacc ccgctccaac tcccctggga	3900
[2439]	accagcggcc tgcctaattg gctgtccgga gatgaagact tctcaagcat cgctgatatg	3960
[2440]	gacttttagtg ccctgctgtc acagatttcc tctagtgggc agggaggagg tggaagcggc	4020
[2441]	ttcagcgtgg acaccagtgc cctgctggac ctgttcagcc cctcggtgac cgtgcccgc	4080
[2442]	atgagcctgc ctgacctga cagcagcctg gccagtatcc aagagctcct gtctccccag	4140
[2443]	gagccccca ggctcccgga ggcagagaac agcagcccgg attcagggaa gcagctggtg	4200
[2444]	cactacacag cgcagccgct gttcctgctg gaccccggct ccgtggacac cgggagcaac	4260
[2445]	gacctgccgg tgctgtttga gctgggagag ggctcctact tctccgaagg ggacggcttc	4320
[2446]	gccgaggacc ccaccatctc cctgctgaca ggctcggagc ctcccaaagc caaggacccc	4380
[2447]	actgtctcct ga	4392
[2448]	<210>	184
[2449]	<211>	4821
[2450]	<212>	DNA
[2451]	<213>	人工序列
[2452]	<220>	
[2453]	<223>	编码dSaCas9(D10A, N580A)-VPR的DNA序列
[2454]	<400>	184
[2455]	atggcccaa agaagaagcg gaaggtcggg atccacggag tcccagcagc caagcggaac	60
[2456]	tacatcctgg gcctggccat cggcatcacc agcgtgggct acggcatcat cgactacgag	120
[2457]	acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac	180
[2458]	gagggcaggc ggagcaagag aggcgccaga aggctgaagc ggcgaggcgc gcatagaatc	240
[2459]	cagagagtga agaagctgct gttcgactac aacctgctga ccgaccacag cgagctgagc	300
[2460]	ggcatcaacc cctacgaggc cagagtgaag ggctgagcc agaagctgag cgaggaagag	360
[2461]	ttctctgccg ccctgctgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
[2462]	gaagaggaca ccggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
[2463]	ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagtg	540
[2464]	cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
[2465]	aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
[2466]	ctggaaaccc ggcggaacct ctatgaggga cctggcgagg gcagccccctt cggctggaag	720
[2467]	gacatcaaag aatggtacga gatgctgatg ggccactgca cctacttccc cgaggaactg	780
[2468]	cggagcgtga agtacgccta caacgccgac ctgtacaacg ccctgaacga cctgaacaat	840
[2469]	ctcgtgatca ccagggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
[2470]	gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatcctc	960
[2471]	gtgaacgaag aggatattaa gggctacaga gtgaccagca ccggcaagcc cgagttcacc	1020
[2472]	aacctgaagg tgtaccacga catcaaggac attaccgccc ggaaagagat tattgagaac	1080
[2473]	gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc	1140
[2474]	caggaagaac tgaccaatct gaactccgag ctgaccaggg aagagatcga gcagatctct	1200
[2475]	aatctgaagg gctataccgg caccacaac ctgagcctga aggccatcaa cctgatcctg	1260

[2476]	gacgagctgt	ggcacacca	cgacaaccag	atcgctatct	tcaaccggct	gaagctggtg	1320
[2477]	cccaagaagg	tggacctgtc	ccagcagaaa	gagatcccca	ccaccctggt	ggacgacttc	1380
[2478]	atctgagcc	ccgtcgtgaa	gagaagcttc	atccagagca	tcaaagtgat	caacgccatc	1440
[2479]	atcaagaagt	acggcctgcc	caacgacatc	attatcgagc	tggcccgcga	gaagaactcc	1500
[2480]	aaggacgccc	agaaaatgat	caacgagatg	cagaagcgga	accggcagac	caacgagcgg	1560
[2481]	atcgaggaaa	tcatccggac	caccggcaaa	gagaacgcca	agtacctgat	cgagaagatc	1620
[2482]	aagctgcacg	acatgcagga	aggcaagtgc	ctgtacagcc	tggaagccat	ccctctggaa	1680
[2483]	gatctgctga	acaacccctt	caactatgag	gtggaccaca	tcatccccag	aagcgtgtcc	1740
[2484]	ttcgacaaca	gcttcaacaa	caaggtgctc	gtgaagcagg	aagaagccag	caagaagggc	1800
[2485]	aaccggaccc	cattccagta	cctgagcagc	agcgacagca	agatcagcta	cgaaaccttc	1860
[2486]	aagaagcaca	tctgaatct	ggccaagggc	aagggcagaa	tcagcaagac	caagaaagag	1920
[2487]	tatctgctgg	aagaacggga	catcaacagg	ttctccgtgc	agaaagactt	catcaaccgg	1980
[2488]	aacctggtgg	ataccagata	cgccaccaga	ggcctgatga	acctgctgcg	gagctacttc	2040
[2489]	agagtgaaca	acctggacgt	gaaagtgaag	tccatcaatg	gcggcttcac	cagctttctg	2100
[2490]	cggcggaagt	ggaagtttaa	gaaagacggg	aacaaggggt	acaagcacca	cgccgaggac	2160
[2491]	gccctgatca	ttgccaacgc	cgatttcac	ttcaaagagt	ggaagaaact	ggacaaggcc	2220
[2492]	aaaaaagtga	tggaaaacca	gatgttcgag	gaaaagcagg	ccgagagcat	gcccagagatc	2280
[2493]	gaaaccgagc	aggagtacaa	agagatcttc	atcaccccc	accagatcaa	gcacattaag	2340
[2494]	gacttcaagg	actacaagta	cagccaccgg	gtggacaaga	agcctaatag	agagctgatt	2400
[2495]	aacgacaccc	tgtactccac	ccggaaggac	gacaagggca	acaccctgat	cgtgaacaat	2460
[2496]	ctgaacggcc	tgtacgacaa	ggacaatgac	aagctgaaaa	agctgatcaa	caagagcccc	2520
[2497]	gaaaagctgc	tgatgtacca	ccacgacccc	cagacctacc	agaaactgaa	gctgattatg	2580
[2498]	gaacagtacg	gcgacgagaa	gaatcccctg	tacaagtact	acgaggaaac	cggggaactac	2640
[2499]	ctgaccaagt	actccaaaaa	ggacaacggc	cccgtgatca	agaagattaa	gtattacggc	2700
[2500]	aacaaactga	acgcccattc	ggacatcacc	gacgactacc	ccaacagcag	aaacaaggtc	2760
[2501]	gtgaagctgt	ccctgaagcc	ctacagattc	gacgtgtacc	tggacaatgg	cgtgtacaag	2820
[2502]	ttcgtgaccg	tgaagaatct	ggatgtgac	aaaaaagaaa	actactacga	agtgaatagc	2880
[2503]	aagtgtctatg	aggaagctaa	gaagctgaag	aagatcagca	accaggccga	gtttatcgcc	2940
[2504]	tccttctaca	acaacgatct	gatcaagatc	aacggcgagc	tgtatagagt	gateggcgtg	3000
[2505]	aacaacgacc	tgctgaaccg	gatcgaagtg	aacatgatcg	acatcaccta	ccgcgagtac	3060
[2506]	ctggaaaaca	tgaacgacaa	gaggcccccc	aggatcatta	agacaatcgc	ctccaagacc	3120
[2507]	cagagcatta	agaagtacag	cacagacatt	ctgggcaacc	tgtatgaagt	gaaatctaag	3180
[2508]	aagcaccctc	agatcatcaa	aaagggcaaa	aggccggcgg	ccacgaaaaa	ggccggccag	3240
[2509]	gcaaaaaaga	aaaagtctag	agatgcttta	gacgattttg	acttagatat	gcttggttca	3300
[2510]	gacgcgttag	acgacttcga	cctagacatg	ttaggctcag	atgcattgga	cgacttcgat	3360
[2511]	ttagatatgt	tgggctccga	tgccctagat	gactttgatt	tggatatgct	aagttccgga	3420
[2512]	tctccgaaaa	agaaacgcaa	agttggtagc	cagtacctgc	ccgacaccga	cgaccggcac	3480
[2513]	cggatcgagg	aaaagcggaa	gcggacctac	gagacattca	agagcatcat	gaagaagtcc	3540
[2514]	cccttcagcg	gccccaccga	ccctagacct	ccacctagaa	gaatcgccgt	gcccagcaga	3600
[2515]	tccagcgcca	gcgtgccaaa	acctgcccc	cagccttacc	ccttcaccag	cagcctgagc	3660
[2516]	accatcaact	acgacgagtt	ccctaccatg	gtgttcccca	gcggccagat	ctctcaggcc	3720
[2517]	tctgtctctg	ctccagcccc	tcctcaggtg	ctgcctcagg	ctctgtctcc	tgcaccagct	3780

[2518]	ccagccatgg tgtctgcact ggctcaggca ccagcacccg tgcctgtgct ggctcctgga	3840
[2519]	cctccacagg ctgtggctcc accagcccct aaacctacac aggccggcga gggcacactg	3900
[2520]	tctgaagctc tgctgcagct gcagttcgac gacgaggatc tgggagccct gctgggaaac	3960
[2521]	agcaccgatc ctgccgtgtt caccgacctg gccagcgtgg acaacagcga gttccagcag	4020
[2522]	ctgtcgaacc agggcatccc tgtggcccct cacaccaccg agcccatgct gatggaatac	4080
[2523]	cccaggcca tcacccggct cgtgacaggc gctcagaggc ctcctgatcc agctcctgcc	4140
[2524]	cctctgggag caccaggcct gcctaattga ctgctgtctg gcgacagga cttcagctct	4200
[2525]	atcgccgata tggatttctc agccttgctg ggctctggca gcggcagccg ggattccagg	4260
[2526]	gaagggatgt ttttgccgaa gcctgaggcc ggctccgcta ttagtgacgt gtttgagggc	4320
[2527]	cgcgagggtg gccagcaaaa acgaatccgg ccatttcac ctcaggaag tccatgggcc	4380
[2528]	aaccgccac tccccgccag cctcgacca acaccaaccg gtccagtaca tgagccagtc	4440
[2529]	gggtcactga ccccgccacc agtccctcag ccactcgatc cagcggccgc agtgactccc	4500
[2530]	gagggcagtc acctgttga agatcccgat gaagagacca gccaggctgt caaagccctt	4560
[2531]	cgggagatgg ccgatactgt gattccccag aaggaagagg ctgcaatctg tggccaaatg	4620
[2532]	gacctttccc atccgcccc aaggggccat ctggatgagc tgacaaccac atttgagtcc	4680
[2533]	atgaccgagg atctgaacct ggactcacc ctgaccccg aattgaacga gattctggat	4740
[2534]	accttctga acgacgagtg cctcttgcat gccatgcata tcagcacagg actgtccatc	4800
[2535]	ttcgacacat ctctgttttg a	4821
[2536]	<210>	185
[2537]	<211>	3726
[2538]	<212>	DNA
[2539]	<213>	人工序列
[2540]	<220>	
[2541]	<223>	编码具有SGGS连接子的dSaCas9[-25] (D10A, N580A) -miniVR的DNA序列
[2542]	<400>	185
[2543]	atggcccaa agaagaagcg gaaggtcggg atccacggag tcccagcagc caagcggaac	60
[2544]	tacatcctgg gcctggccat cggcatcacc agcgtgggct acggcatcat cgactacgag	120
[2545]	acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac	180
[2546]	gagggcaggc ggagcaagag aggcgccaga aggctgaagc ggcgaggcgc gcatagaatc	240
[2547]	cagagagtga agaagctgct gttcgactac aacctgctga ccgaccacag cgagctgagc	300
[2548]	ggcatcaacc cctacaggc cagagtgaag ggcctgagcc agaagctgag cgaggaagag	360
[2549]	ttctctgccg ccctgctgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
[2550]	gaagaggaca cgggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
[2551]	ctggaagaga aatacgtggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagtg	540
[2552]	cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
[2553]	aaggctcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
[2554]	ctggaaaccc ggcggaacct ctatgaggga cctggcgagg gcagccccct cggttggaag	720
[2555]	gacatcaaag aatggtacga gatgctgatg ggccactgca cctacttccc cgaggaactg	780
[2556]	cggagcgtga agtacgccta caacgccgac ctgtacaacg ccctgaacga cctgaacaat	840
[2557]	ctcgtgatca ccaggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
[2558]	gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatcctc	960
[2559]	gtgaacgaag aggatattaa gggctacaga gtgaccagca ccggcaagcc cgagttcacc	1020

[2560]	aacctgaagg	tgtaccacga	catcaaggac	attaccgccc	ggaaagagat	tattgagaac	1080
[2561]	gccgagctgc	tggatcagat	tgccaagatc	ctgaccatct	accagagcag	cgaggacatc	1140
[2562]	caggaagaac	tgaccaatct	gaactccgag	ctgaccagg	aagagatcga	gcagatctct	1200
[2563]	aatctgaagg	gctataccgg	caccacaaac	ctgagcctga	aggccatcaa	cctgatcctg	1260
[2564]	gacgagctgt	ggcacaccaa	cgacaaccag	atcgctatct	tcaaccggct	gaagctgggt	1320
[2565]	cccaagaagg	tggacctgtc	ccagcagaaa	gagatcccca	ccaccctgg	ggacgacttc	1380
[2566]	atctgagcc	ccgtcgtgaa	gagaagcttc	atccagagca	tcaaagtgat	caacgccatc	1440
[2567]	atcaagaagt	acggcctgcc	caacgacatc	attatcgagc	tggcccgca	gaagaactcc	1500
[2568]	aaggacgccc	agaaaatgat	caacgagatg	cagaagcgga	accggcagac	caacgagcgg	1560
[2569]	atcgaggaaa	tcatccggac	caccggcaaa	gagaacgcca	agtacctgat	cgagaagatc	1620
[2570]	aagctgcacg	acatgcagga	aggcaagtgc	ctgtacagcc	tggaagccat	ccctctggaa	1680
[2571]	gatctgctga	acaacccctt	caactatgag	gtggaccaca	tcatccccag	aagcgtgtcc	1740
[2572]	ttcgacaaca	gcttcaacaa	caaggtgtc	gtgaagcagg	aagaagccag	caagaagggc	1800
[2573]	aaccggaccc	cattccagta	cctgagcagc	agcgacagca	agatcagcta	cgaaaccttc	1860
[2574]	aagaagcaca	tctgaatct	ggccaagggc	aagggcagaa	tcagcaagac	caagaaagag	1920
[2575]	tatctgctgg	aagaacggga	catcaacagg	ttctccgtgc	agaaagactt	catcaaccgg	1980
[2576]	aacctgggtg	ataccagata	cgccaccaga	ggcctgatga	acctgctgcg	gagctacttc	2040
[2577]	agagtgaaca	acctggacgt	gaaagtgaag	tccatcaatg	gcggcttcac	cagctttctg	2100
[2578]	cggcggaagt	ggaagttaa	gaaagagcgg	aacaaggggt	acaagcacca	cgccgaggac	2160
[2579]	gccctgatca	ttgccaacgc	cgatttcac	ttcaaagagt	ggaagaaatc	cggcgggcgt	2220
[2580]	tcgaccgagc	aggagtacaa	agagatcttc	atccccccc	accagatcaa	gcacattaag	2280
[2581]	gacttcaagg	actacaagta	cagccaccgg	gtggacaaga	agcctaatag	agagctgatt	2340
[2582]	aacgacaccc	tgtactccac	ccggaaggac	gacaagggca	acaccctgat	cgtgaacaat	2400
[2583]	ctgaacggcc	tgtacgacaa	ggacaatgac	aagctgaaaa	agctgatcaa	caagagcccc	2460
[2584]	gaaaagctgc	tgatgtacca	ccacgacccc	cagacctacc	agaaactgaa	gctgattatg	2520
[2585]	gaacagtacg	gcgacgagaa	gaatccccctg	tacaagtact	acgaggaaac	cgggaactac	2580
[2586]	ctgaccaagt	actccaaaaa	ggacaacggc	cccgtgatca	agaagattaa	gtattacggc	2640
[2587]	aacaaactga	acgcccctct	ggacatcacc	gacgactacc	ccaacagcag	aaacaaggct	2700
[2588]	gtgaagctgt	ccctgaagcc	ctacagattc	gacgtgtacc	tggaacaatg	cgtgtacaag	2760
[2589]	ttcgtgaccg	tgaagaatct	ggatgtgatc	aaaaaagaaa	actactacga	agtgaatagc	2820
[2590]	aagtgtctatg	aggaagctaa	gaagctgaag	aagatcagca	accaggccga	gtttatcgcc	2880
[2591]	tccttctaca	acaacgatct	gatcaagatc	aacggcgagc	tgtatagagt	gatcggcgtg	2940
[2592]	aacaacgacc	tgctgaaccg	gatcgaagtg	aacatgatcg	acatcaccta	ccgcgagtac	3000
[2593]	ctggaaaaca	tgaacgacaa	gaggcccccc	aggatcatta	agacaatcgc	ctccaagacc	3060
[2594]	cagagcatta	agaagtacag	cacagacatt	ctgggcaacc	tgtatgaagt	gaaatctaag	3120
[2595]	aagcaccctc	agatcatcaa	aaagggcaaa	aggccggcgg	ccacgaaaaa	ggccggccag	3180
[2596]	gcaaaaaaga	aaaagggtac	cggtgctaga	aagcttgatg	cttttagacga	ttttgactta	3240
[2597]	gatatgcttg	gttcagacgc	gttagacgac	ttcgacctag	acatgttagg	ctcagatgca	3300
[2598]	ttggacgact	tcgatttaga	tatgttgggc	tccgatgccc	tagatgactt	tgattttggat	3360
[2599]	atgctaggat	ctggtagccc	agcggccgca	gtgactccc	aggccagtca	cctgtttggaa	3420
[2600]	gatcccgatg	aagaaaccag	ccaggctgtc	aaagcccttc	gggagatggc	cgatactgtg	3480
[2601]	attccccaga	aggaagaggc	tgcaatctgt	ggccaaatgg	acctttccca	tccgccccca	3540

103

[2644]	20	25	30
[2645]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly		
[2646]	35	40	45
[2647]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg		
[2648]	50	55	60
[2649]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile		
[2650]	65	70	75
[2651]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His		
[2652]	85	90	95
[2653]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu		
[2654]	100	105	110
[2655]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu		
[2656]	115	120	125
[2657]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr		
[2658]	130	135	140
[2659]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[2660]	145	150	155
[2661]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[2662]	165	170	175
[2663]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[2664]	180	185	190
[2665]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[2666]	195	200	205
[2667]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[2668]	210	215	220
[2669]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[2670]	225	230	235
[2671]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[2672]	245	250	255
[2673]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[2674]	260	265	270
[2675]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[2676]	275	280	285
[2677]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[2678]	290	295	300
[2679]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[2680]	305	310	315
[2681]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[2682]	325	330	335
[2683]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[2684]	340	345	350
[2685]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		

[2686]	355	360	365
[2687]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[2688]	370	375	380
[2689]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[2690]	385	390	395
[2691]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[2692]	405	410	415
[2693]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[2694]	420	425	430
[2695]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[2696]	435	440	445
[2697]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[2698]	450	455	460
[2699]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[2700]	465	470	475
[2701]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[2702]	485	490	495
[2703]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[2704]	500	505	510
[2705]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[2706]	515	520	525
[2707]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[2708]	530	535	540
[2709]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[2710]	545	550	555
[2711]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[2712]	565	570	575
[2713]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[2714]	580	585	590
[2715]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[2716]	595	600	605
[2717]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[2718]	610	615	620
[2719]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[2720]	625	630	635
[2721]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[2722]	645	650	655
[2723]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[2724]	660	665	670
[2725]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[2726]	675	680	685
[2727]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		

[2728]	690	695	700
[2729]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[2730]	705	710	715
[2731]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[2732]		725	730
[2733]	Ser Gly Gly Gly Ser Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile Thr		
[2734]		740	745
[2735]	Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr Ser		
[2736]		755	760
[2737]	His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr Leu		
[2738]	770	775	780
[2739]	Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile Val Asn Asn		
[2740]	785	790	795
[2741]	Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu Ile		
[2742]		805	810
[2743]	Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp Pro Gln Thr		
[2744]		820	825
[2745]	Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys Asn		
[2746]		835	840
[2747]	Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys Tyr		
[2748]	850	855	860
[2749]	Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr Gly		
[2750]	865	870	875
[2751]	Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr Pro Asn Ser		
[2752]		885	890
[2753]	Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp Val		
[2754]		900	905
[2755]	Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys Asn Leu Asp		
[2756]		915	920
[2757]	Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr Glu		
[2758]	930	935	940
[2759]	Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile Ala		
[2760]	945	950	955
[2761]	Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr Arg		
[2762]		965	970
[2763]	Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu Val Asn Met		
[2764]		980	985
[2765]	Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys Arg		
[2766]		995	1000
[2767]	Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr Gln Ser Ile		
[2768]	1010	1015	1020
[2769]	Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu Val Lys		

[2770]	1025	1030	1035
[2771]	Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly Lys Arg Pro Ala		
[2772]	1040	1045	1050
[2773]	Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Gly Ser Gly		
[2774]	1055	1060	1065
[2775]	Ala Arg Lys Leu Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu		
[2776]	1070	1075	1080
[2777]	Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser		
[2778]	1085	1090	1095
[2779]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala		
[2780]	1100	1105	1110
[2781]	Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Gly Ser Pro Ala		
[2782]	1115	1120	1125
[2783]	Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp		
[2784]	1130	1135	1140
[2785]	Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp		
[2786]	1145	1150	1155
[2787]	Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met		
[2788]	1160	1165	1170
[2789]	Asp Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr		
[2790]	1175	1180	1185
[2791]	Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro		
[2792]	1190	1195	1200
[2793]	Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp		
[2794]	1205	1210	1215
[2795]	Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly Leu Ser Ile		
[2796]	1220	1225	1230
[2797]	Phe Asp Thr Ser Leu Phe Val Asp		
[2798]	1235	1240	
[2799]	<210> 187		
[2800]	<211> 256		
[2801]	<212> DNA		
[2802]	<213> 人工序列		
[2803]	<220>		
[2804]	<223> EFS启动子序列		
[2805]	<400> 187		
[2806]	taggtcttga aaggagtggg aattggctcc ggtgcccgtc agtgggcaga gcgcacatcg 60		
[2807]	cccacagtcc ccgagaagtt ggggggaggg gtcggcaatt gatccggtgc ctagagaagg 120		
[2808]	tggcgcgggg taaactggga aagtgatgtc gtgtactggc tccgcctttt tcccgagggt 180		
[2809]	gggggagaac cgtatataag tgcagtagtc gccgtgaacg ttcttttttcg caacgggttt 240		
[2810]	gccgccagaa cacagg 256		
[2811]	<210> 188		

[2812]	<211>	1137
[2813]	<212>	PRT
[2814]	<213>	人工序列
[2815]	<220>	
[2816]	<223>	dSaCas9 (D10A和N580A) -VP64
[2817]	<220>	
[2818]	<221>	信号
[2819]	<222>	(1) .. (17)
[2820]	<223>	核定位信号
[2821]	<220>	
[2822]	<221>	肽
[2823]	<222>	(18) .. (1069)
[2824]	<223>	dSaCas9 (D10A和N580A)
[2825]	<220>	
[2826]	<221>	变体
[2827]	<222>	(26) .. (26)
[2828]	<223>	Asp残基转化成Ala残基
[2829]	<220>	
[2830]	<221>	变体
[2831]	<222>	(596) .. (596)
[2832]	<223>	Asn残基转化成Ala残基
[2833]	<220>	
[2834]	<221>	信号
[2835]	<222>	(1070) .. (1085)
[2836]	<223>	核定位信号
[2837]	<220>	
[2838]	<221>	肽
[2839]	<222>	(1088) .. (1137)
[2840]	<223>	VP64
[2841]	<400>	188
[2842]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala	
[2843]	1	5 10 15
[2844]	Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val	
[2845]	20	25 30
[2846]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly	
[2847]	35	40 45
[2848]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg	
[2849]	50	55 60
[2850]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile	
[2851]	65	70 75 80
[2852]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His	
[2853]	85	90 95

[2854]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu		
[2855]	100	105	110
[2856]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu		
[2857]	115	120	125
[2858]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr		
[2859]	130	135	140
[2860]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[2861]	145	150	155
[2862]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[2863]	165	170	175
[2864]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[2865]	180	185	190
[2866]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[2867]	195	200	205
[2868]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[2869]	210	215	220
[2870]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[2871]	225	230	235
[2872]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[2873]	245	250	255
[2874]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[2875]	260	265	270
[2876]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[2877]	275	280	285
[2878]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[2879]	290	295	300
[2880]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[2881]	305	310	315
[2882]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[2883]	325	330	335
[2884]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[2885]	340	345	350
[2886]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[2887]	355	360	365
[2888]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[2889]	370	375	380
[2890]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[2891]	385	390	395
[2892]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[2893]	405	410	415
[2894]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[2895]	420	425	430

[2896]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[2897]	435	440	445
[2898]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[2899]	450	455	460
[2900]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[2901]	465	470	475
[2902]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[2903]	485	490	495
[2904]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[2905]	500	505	510
[2906]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[2907]	515	520	525
[2908]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[2909]	530	535	540
[2910]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[2911]	545	550	555
[2912]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[2913]	565	570	575
[2914]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[2915]	580	585	590
[2916]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[2917]	595	600	605
[2918]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[2919]	610	615	620
[2920]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[2921]	625	630	635
[2922]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[2923]	645	650	655
[2924]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[2925]	660	665	670
[2926]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[2927]	675	680	685
[2928]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[2929]	690	695	700
[2930]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[2931]	705	710	715
[2932]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[2933]	725	730	735
[2934]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys		
[2935]	740	745	750
[2936]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu		
[2937]	755	760	765

[2938]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp		
[2939]	770	775	780
[2940]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile		
[2941]	785	790	795 800
[2942]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu		
[2943]	805	810	815
[2944]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu		
[2945]	820	825	830
[2946]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His		
[2947]	835	840	845
[2948]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly		
[2949]	850	855	860
[2950]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr		
[2951]	865	870	875 880
[2952]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile		
[2953]	885	890	895
[2954]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp		
[2955]	900	905	910
[2956]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr		
[2957]	915	920	925
[2958]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val		
[2959]	930	935	940
[2960]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser		
[2961]	945	950	955 960
[2962]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala		
[2963]	965	970	975
[2964]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly		
[2965]	980	985	990
[2966]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile		
[2967]	995	1000	1005
[2968]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn		
[2969]	1010	1015	1020
[2970]	Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser		
[2971]	1025	1030	1035
[2972]	Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn		
[2973]	1040	1045	1050
[2974]	Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys		
[2975]	1055	1060	1065
[2976]	Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys		
[2977]	1070	1075	1080
[2978]	Lys Lys Ser Arg Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu		
[2979]	1085	1090	1095

[2980]	Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
[2981]	1100 1105 1110
[2982]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala
[2983]	1115 1120 1125
[2984]	Leu Asp Asp Phe Asp Leu Asp Met Leu
[2985]	1130 1135
[2986]	<210> 189
[2987]	<211> 1463
[2988]	<212> PRT
[2989]	<213> 人工序列
[2990]	<220>
[2991]	<223> dSaCas9 (D10A和N580A) -VPH
[2992]	<220>
[2993]	<221> 信号
[2994]	<222> (1) .. (17)
[2995]	<223> 核定位信号
[2996]	<220>
[2997]	<221> 肽
[2998]	<222> (18) .. (1069)
[2999]	<223> dSaCas9 (D10A和N580A)
[3000]	<220>
[3001]	<221> 变体
[3002]	<222> (26) .. (26)
[3003]	<223> Asp残基转化成Ala残基
[3004]	<220>
[3005]	<221> 变体
[3006]	<222> (596) .. (596)
[3007]	<223> Asp残基转化成Ala残基
[3008]	<220>
[3009]	<221> 信号
[3010]	<222> (1070) .. (1085)
[3011]	<223> 核定位信号
[3012]	<220>
[3013]	<221> 肽
[3014]	<222> (1088) .. (1463)
[3015]	<223> VPH
[3016]	<400> 189
[3017]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
[3018]	1 5 10 15
[3019]	Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
[3020]	20 25 30
[3021]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly

[3022]	35	40	45
[3023]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg		
[3024]	50	55	60
[3025]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile		
[3026]	65	70	75
[3027]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His		
[3028]	85	90	95
[3029]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu		
[3030]	100	105	110
[3031]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu		
[3032]	115	120	125
[3033]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr		
[3034]	130	135	140
[3035]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[3036]	145	150	155
[3037]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[3038]	165	170	175
[3039]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[3040]	180	185	190
[3041]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[3042]	195	200	205
[3043]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[3044]	210	215	220
[3045]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[3046]	225	230	235
[3047]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[3048]	245	250	255
[3049]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[3050]	260	265	270
[3051]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[3052]	275	280	285
[3053]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[3054]	290	295	300
[3055]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[3056]	305	310	315
[3057]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[3058]	325	330	335
[3059]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[3060]	340	345	350
[3061]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[3062]	355	360	365
[3063]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		

[3064]	370	375	380
[3065]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[3066]	385	390	395 400
[3067]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[3068]	405	410	415
[3069]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[3070]	420	425	430
[3071]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[3072]	435	440	445
[3073]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[3074]	450	455	460
[3075]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[3076]	465	470	475 480
[3077]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[3078]	485	490	495
[3079]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[3080]	500	505	510
[3081]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[3082]	515	520	525
[3083]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[3084]	530	535	540
[3085]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[3086]	545	550	555 560
[3087]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[3088]	565	570	575
[3089]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[3090]	580	585	590
[3091]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[3092]	595	600	605
[3093]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[3094]	610	615	620
[3095]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[3096]	625	630	635 640
[3097]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[3098]	645	650	655
[3099]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[3100]	660	665	670
[3101]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[3102]	675	680	685
[3103]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[3104]	690	695	700
[3105]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		

[3106]	705	710	715	720
[3107]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys			
[3108]		725	730	735
[3109]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys			
[3110]		740	745	750
[3111]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu			
[3112]		755	760	765
[3113]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp			
[3114]		770	775	780
[3115]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile			
[3116]	785	790	795	800
[3117]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu			
[3118]		805	810	815
[3119]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu			
[3120]		820	825	830
[3121]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His			
[3122]		835	840	845
[3123]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly			
[3124]		850	855	860
[3125]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr			
[3126]	865	870	875	880
[3127]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile			
[3128]		885	890	895
[3129]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp			
[3130]		900	905	910
[3131]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr			
[3132]		915	920	925
[3133]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val			
[3134]		930	935	940
[3135]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser			
[3136]	945	950	955	960
[3137]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala			
[3138]		965	970	975
[3139]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly			
[3140]		980	985	990
[3141]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile			
[3142]		995	1000	1005
[3143]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn			
[3144]		1010	1015	1020
[3145]	Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser			
[3146]		1025	1030	1035
[3147]	Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn			

[3148]	1040	1045	1050
[3149]	Leu Tyr Glu Val Lys Ser	Lys Lys His Pro Gln Ile Ile Lys Lys	
[3150]	1055	1060	1065
[3151]	Gly Lys Arg Pro Ala Ala	Thr Lys Lys Ala Gly Gln Ala Lys Lys	
[3152]	1070	1075	1080
[3153]	Lys Lys Ser Arg Asp Ala	Leu Asp Asp Phe Asp Leu Asp Met Leu	
[3154]	1085	1090	1095
[3155]	Gly Ser Asp Ala Leu Asp	Asp Phe Asp Leu Asp Met Leu Gly Ser	
[3156]	1100	1105	1110
[3157]	Asp Ala Leu Asp Asp Phe	Asp Leu Asp Met Leu Gly Ser Asp Ala	
[3158]	1115	1120	1125
[3159]	Leu Asp Asp Phe Asp Leu	Asp Met Leu Ser Ser Gly Ser Pro Lys	
[3160]	1130	1135	1140
[3161]	Lys Lys Arg Lys Val Gly	Ser Pro Ser Gly Gln Ile Ser Asn Gln	
[3162]	1145	1150	1155
[3163]	Ala Leu Ala Leu Ala Pro	Ser Ser Ala Pro Val Leu Ala Gln Thr	
[3164]	1160	1165	1170
[3165]	Met Val Pro Ser Ser Ala	Met Val Pro Leu Ala Gln Pro Pro Ala	
[3166]	1175	1180	1185
[3167]	Pro Ala Pro Val Leu Thr	Pro Gly Pro Pro Gln Ser Leu Ser Ala	
[3168]	1190	1195	1200
[3169]	Pro Val Pro Lys Ser Thr	Gln Ala Gly Glu Gly Thr Leu Ser Glu	
[3170]	1205	1210	1215
[3171]	Ala Leu Leu His Leu Gln	Phe Asp Ala Asp Glu Asp Leu Gly Ala	
[3172]	1220	1225	1230
[3173]	Leu Leu Gly Asn Ser Thr	Asp Pro Gly Val Phe Thr Asp Leu Ala	
[3174]	1235	1240	1245
[3175]	Ser Val Asp Asn Ser Glu	Phe Gln Gln Leu Leu Asn Gln Gly Val	
[3176]	1250	1255	1260
[3177]	Ser Met Ser His Ser Thr	Ala Glu Pro Met Leu Met Glu Tyr Pro	
[3178]	1265	1270	1275
[3179]	Glu Ala Ile Thr Arg Leu	Val Thr Gly Ser Gln Arg Pro Pro Asp	
[3180]	1280	1285	1290
[3181]	Pro Ala Pro Thr Pro Leu	Gly Thr Ser Gly Leu Pro Asn Gly Leu	
[3182]	1295	1300	1305
[3183]	Ser Gly Asp Glu Asp Phe	Ser Ser Ile Ala Asp Met Asp Phe Ser	
[3184]	1310	1315	1320
[3185]	Ala Leu Leu Ser Gln Ile	Ser Ser Ser Gly Gln Gly Gly Gly Gly	
[3186]	1325	1330	1335
[3187]	Ser Gly Phe Ser Val Asp	Thr Ser Ala Leu Leu Asp Leu Phe Ser	
[3188]	1340	1345	1350
[3189]	Pro Ser Val Thr Val Pro	Asp Met Ser Leu Pro Asp Leu Asp Ser	

[3190]	1355	1360	1365
[3191]	Ser Leu Ala Ser Ile Gln Glu Leu Leu Ser Pro Gln Glu Pro Pro		
[3192]	1370	1375	1380
[3193]	Arg Pro Pro Glu Ala Glu Asn Ser Ser Pro Asp Ser Gly Lys Gln		
[3194]	1385	1390	1395
[3195]	Leu Val His Tyr Thr Ala Gln Pro Leu Phe Leu Leu Asp Pro Gly		
[3196]	1400	1405	1410
[3197]	Ser Val Asp Thr Gly Ser Asn Asp Leu Pro Val Leu Phe Glu Leu		
[3198]	1415	1420	1425
[3199]	Gly Glu Gly Ser Tyr Phe Ser Glu Gly Asp Gly Phe Ala Glu Asp		
[3200]	1430	1435	1440
[3201]	Pro Thr Ile Ser Leu Leu Thr Gly Ser Glu Pro Pro Lys Ala Lys		
[3202]	1445	1450	1455
[3203]	Asp Pro Thr Val Ser		
[3204]	1460		
[3205]	<210> 190		
[3206]	<211> 1606		
[3207]	<212> PRT		
[3208]	<213> 人工序列		
[3209]	<220>		
[3210]	<223> dSaCas9 (D10A和N580A) -VPR		
[3211]	<220>		
[3212]	<221> 信号		
[3213]	<222> (1) .. (17)		
[3214]	<223> 核定位信号		
[3215]	<220>		
[3216]	<221> 肽		
[3217]	<222> (18) .. (1069)		
[3218]	<223> dSaCas9 (D10A和N580A)		
[3219]	<220>		
[3220]	<221> 变体		
[3221]	<222> (26) .. (26)		
[3222]	<223> Asp残基转化成Ala残基		
[3223]	<220>		
[3224]	<221> 变体		
[3225]	<222> (596) .. (596)		
[3226]	<223> Asn残基转化成Ala残基		
[3227]	<220>		
[3228]	<221> 信号		
[3229]	<222> (1070) .. (1085)		
[3230]	<223> 核定位信号		
[3231]	<220>		

[3232]	<221>	肽
[3233]	<222>	(1088) .. (1606)
[3234]	<223>	VPR
[3235]	<400>	190
[3236]	Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala	
[3237]	1	5 10 15
[3238]	Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val	
[3239]		20 25 30
[3240]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly	
[3241]		35 40 45
[3242]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg	
[3243]		50 55 60
[3244]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile	
[3245]	65	70 75 80
[3246]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His	
[3247]		85 90 95
[3248]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu	
[3249]		100 105 110
[3250]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu	
[3251]		115 120 125
[3252]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr	
[3253]		130 135 140
[3254]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala	
[3255]	145	150 155 160
[3256]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys	
[3257]		165 170 175
[3258]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr	
[3259]		180 185 190
[3260]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln	
[3261]		195 200 205
[3262]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg	
[3263]		210 215 220
[3264]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys	
[3265]	225	230 235 240
[3266]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe	
[3267]		245 250 255
[3268]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr	
[3269]		260 265 270
[3270]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn	
[3271]		275 280 285
[3272]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe	
[3273]		290 295 300

[3274]	Lys	Gln	Lys	Lys	Lys	Pro	Thr	Leu	Lys	Gln	Ile	Ala	Lys	Glu	Ile	Leu
[3275]	305					310				315						320
[3276]	Val	Asn	Glu	Glu	Asp	Ile	Lys	Gly	Tyr	Arg	Val	Thr	Ser	Thr	Gly	Lys
[3277]					325					330						335
[3278]	Pro	Glu	Phe	Thr	Asn	Leu	Lys	Val	Tyr	His	Asp	Ile	Lys	Asp	Ile	Thr
[3279]				340					345					350		
[3280]	Ala	Arg	Lys	Glu	Ile	Ile	Glu	Asn	Ala	Glu	Leu	Leu	Asp	Gln	Ile	Ala
[3281]			355					360					365			
[3282]	Lys	Ile	Leu	Thr	Ile	Tyr	Gln	Ser	Ser	Glu	Asp	Ile	Gln	Glu	Glu	Leu
[3283]	370					375				380						
[3284]	Thr	Asn	Leu	Asn	Ser	Glu	Leu	Thr	Gln	Glu	Glu	Ile	Glu	Gln	Ile	Ser
[3285]	385				390					395						400
[3286]	Asn	Leu	Lys	Gly	Tyr	Thr	Gly	Thr	His	Asn	Leu	Ser	Leu	Lys	Ala	Ile
[3287]				405					410							415
[3288]	Asn	Leu	Ile	Leu	Asp	Glu	Leu	Trp	His	Thr	Asn	Asp	Asn	Gln	Ile	Ala
[3289]			420					425						430		
[3290]	Ile	Phe	Asn	Arg	Leu	Lys	Leu	Val	Pro	Lys	Lys	Val	Asp	Leu	Ser	Gln
[3291]		435					440					445				
[3292]	Gln	Lys	Glu	Ile	Pro	Thr	Thr	Leu	Val	Asp	Asp	Phe	Ile	Leu	Ser	Pro
[3293]	450					455						460				
[3294]	Val	Val	Lys	Arg	Ser	Phe	Ile	Gln	Ser	Ile	Lys	Val	Ile	Asn	Ala	Ile
[3295]	465				470					475						480
[3296]	Ile	Lys	Lys	Tyr	Gly	Leu	Pro	Asn	Asp	Ile	Ile	Ile	Glu	Leu	Ala	Arg
[3297]				485				490						495		
[3298]	Glu	Lys	Asn	Ser	Lys	Asp	Ala	Gln	Lys	Met	Ile	Asn	Glu	Met	Gln	Lys
[3299]			500					505					510			
[3300]	Arg	Asn	Arg	Gln	Thr	Asn	Glu	Arg	Ile	Glu	Glu	Ile	Ile	Arg	Thr	Thr
[3301]		515					520						525			
[3302]	Gly	Lys	Glu	Asn	Ala	Lys	Tyr	Leu	Ile	Glu	Lys	Ile	Lys	Leu	His	Asp
[3303]		530					535						540			
[3304]	Met	Gln	Glu	Gly	Lys	Cys	Leu	Tyr	Ser	Leu	Glu	Ala	Ile	Pro	Leu	Glu
[3305]	545				550					555						560
[3306]	Asp	Leu	Leu	Asn	Asn	Pro	Phe	Asn	Tyr	Glu	Val	Asp	His	Ile	Ile	Pro
[3307]				565				570						575		
[3308]	Arg	Ser	Val	Ser	Phe	Asp	Asn	Ser	Phe	Asn	Asn	Lys	Val	Leu	Val	Lys
[3309]			580					585						590		
[3310]	Gln	Glu	Glu	Ala	Ser	Lys	Lys	Gly	Asn	Arg	Thr	Pro	Phe	Gln	Tyr	Leu
[3311]		595					600						605			
[3312]	Ser	Ser	Ser	Asp	Ser	Lys	Ile	Ser	Tyr	Glu	Thr	Phe	Lys	Lys	His	Ile
[3313]		610				615						620				
[3314]	Leu	Asn	Leu	Ala	Lys	Gly	Lys	Gly	Arg	Ile	Ser	Lys	Thr	Lys	Lys	Glu
[3315]	625				630					635						640

[3316]	Tyr	Leu	Leu	Glu	Glu	Arg	Asp	Ile	Asn	Arg	Phe	Ser	Val	Gln	Lys	Asp
[3317]					645					650					655	
[3318]	Phe	Ile	Asn	Arg	Asn	Leu	Val	Asp	Thr	Arg	Tyr	Ala	Thr	Arg	Gly	Leu
[3319]					660					665					670	
[3320]	Met	Asn	Leu	Leu	Arg	Ser	Tyr	Phe	Arg	Val	Asn	Asn	Leu	Asp	Val	Lys
[3321]					675					680					685	
[3322]	Val	Lys	Ser	Ile	Asn	Gly	Gly	Phe	Thr	Ser	Phe	Leu	Arg	Arg	Lys	Trp
[3323]					690					695					700	
[3324]	Lys	Phe	Lys	Lys	Glu	Arg	Asn	Lys	Gly	Tyr	Lys	His	His	Ala	Glu	Asp
[3325]	705					710					715				720	
[3326]	Ala	Leu	Ile	Ile	Ala	Asn	Ala	Asp	Phe	Ile	Phe	Lys	Glu	Trp	Lys	Lys
[3327]					725					730					735	
[3328]	Leu	Asp	Lys	Ala	Lys	Lys	Val	Met	Glu	Asn	Gln	Met	Phe	Glu	Glu	Lys
[3329]					740					745					750	
[3330]	Gln	Ala	Glu	Ser	Met	Pro	Glu	Ile	Glu	Thr	Glu	Gln	Glu	Tyr	Lys	Glu
[3331]					755					760					765	
[3332]	Ile	Phe	Ile	Thr	Pro	His	Gln	Ile	Lys	His	Ile	Lys	Asp	Phe	Lys	Asp
[3333]					770					775					780	
[3334]	Tyr	Lys	Tyr	Ser	His	Arg	Val	Asp	Lys	Lys	Pro	Asn	Arg	Glu	Leu	Ile
[3335]	785					790					795				800	
[3336]	Asn	Asp	Thr	Leu	Tyr	Ser	Thr	Arg	Lys	Asp	Asp	Lys	Gly	Asn	Thr	Leu
[3337]					805					810					815	
[3338]	Ile	Val	Asn	Asn	Leu	Asn	Gly	Leu	Tyr	Asp	Lys	Asp	Asn	Asp	Lys	Leu
[3339]					820					825					830	
[3340]	Lys	Lys	Leu	Ile	Asn	Lys	Ser	Pro	Glu	Lys	Leu	Leu	Met	Tyr	His	His
[3341]					835					840					845	
[3342]	Asp	Pro	Gln	Thr	Tyr	Gln	Lys	Leu	Lys	Leu	Ile	Met	Glu	Gln	Tyr	Gly
[3343]					850					855					860	
[3344]	Asp	Glu	Lys	Asn	Pro	Leu	Tyr	Lys	Tyr	Tyr	Glu	Glu	Thr	Gly	Asn	Tyr
[3345]	865					870					875				880	
[3346]	Leu	Thr	Lys	Tyr	Ser	Lys	Lys	Asp	Asn	Gly	Pro	Val	Ile	Lys	Lys	Ile
[3347]					885					890					895	
[3348]	Lys	Tyr	Tyr	Gly	Asn	Lys	Leu	Asn	Ala	His	Leu	Asp	Ile	Thr	Asp	Asp
[3349]					900					905					910	
[3350]	Tyr	Pro	Asn	Ser	Arg	Asn	Lys	Val	Val	Lys	Leu	Ser	Leu	Lys	Pro	Tyr
[3351]					915					920					925	
[3352]	Arg	Phe	Asp	Val	Tyr	Leu	Asp	Asn	Gly	Val	Tyr	Lys	Phe	Val	Thr	Val
[3353]					930					935					940	
[3354]	Lys	Asn	Leu	Asp	Val	Ile	Lys	Lys	Glu	Asn	Tyr	Tyr	Glu	Val	Asn	Ser
[3355]	945					950					955				960	
[3356]	Lys	Cys	Tyr	Glu	Glu	Ala	Lys	Lys	Leu	Lys	Lys	Ile	Ser	Asn	Gln	Ala
[3357]					965					970					975	

[3358]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly		
[3359]	980	985	990
[3360]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile		
[3361]	995	1000	1005
[3362]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn		
[3363]	1010	1015	1020
[3364]	Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser		
[3365]	1025	1030	1035
[3366]	Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn		
[3367]	1040	1045	1050
[3368]	Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys		
[3369]	1055	1060	1065
[3370]	Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys		
[3371]	1070	1075	1080
[3372]	Lys Lys Ser Arg Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu		
[3373]	1085	1090	1095
[3374]	Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser		
[3375]	1100	1105	1110
[3376]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala		
[3377]	1115	1120	1125
[3378]	Leu Asp Asp Phe Asp Leu Asp Met Leu Ser Ser Gly Ser Pro Lys		
[3379]	1130	1135	1140
[3380]	Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp Asp		
[3381]	1145	1150	1155
[3382]	Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe		
[3383]	1160	1165	1170
[3384]	Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro		
[3385]	1175	1180	1185
[3386]	Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser Ser Ala		
[3387]	1190	1195	1200
[3388]	Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser		
[3389]	1205	1210	1215
[3390]	Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro		
[3391]	1220	1225	1230
[3392]	Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro		
[3393]	1235	1240	1245
[3394]	Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met		
[3395]	1250	1255	1260
[3396]	Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala		
[3397]	1265	1270	1275
[3398]	Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr		
[3399]	1280	1285	1290

[3400]	Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln
[3401]	1295 1300 1305
[3402]	Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp
[3403]	1310 1315 1320
[3404]	Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe
[3405]	1325 1330 1335
[3406]	Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr
[3407]	1340 1345 1350
[3408]	Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val
[3409]	1355 1360 1365
[3410]	Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly
[3411]	1370 1375 1380
[3412]	Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe
[3413]	1385 1390 1395
[3414]	Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
[3415]	1400 1405 1410
[3416]	Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro
[3417]	1415 1420 1425
[3418]	Glu Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val
[3419]	1430 1435 1440
[3420]	Cys Gln Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro
[3421]	1445 1450 1455
[3422]	Trp Ala Asn Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr
[3423]	1460 1465 1470
[3424]	Gly Pro Val His Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val
[3425]	1475 1480 1485
[3426]	Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr Pro Glu Ala Ser
[3427]	1490 1495 1500
[3428]	His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala Val Lys
[3429]	1505 1510 1515
[3430]	Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu Glu
[3431]	1520 1525 1530
[3432]	Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg
[3433]	1535 1540 1545
[3434]	Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu
[3435]	1550 1555 1560
[3436]	Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile
[3437]	1565 1570 1575
[3438]	Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His
[3439]	1580 1585 1590
[3440]	Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
[3441]	1595 1600 1605

[3442]	<210> 191	
[3443]	<211> 450	
[3444]	<212> DNA	
[3445]	<213> 人工序列	
[3446]	<220>	
[3447]	<223> CK8启动子序列	
[3448]	<400> 191	
[3449]	ctagactagc atgtctgcca tgtaaggagg caaggcctgg ggacacccga gatgcctggt	60
[3450]	tataattaac ccagacatgt ggctgcccc cccccccaa cacctgctgc ctctaaaaat	120
[3451]	aaccctgcat gccatgttcc cggcgaagg ccagctgtcc cccgccagct agactcagca	180
[3452]	cttagtttag gaaccagtga gcaagtcagc ctttggggca gccatacaa ggccatgggg	240
[3453]	ctgggcaagc tgcacgctg ggtccgggt gggcacggtg cccgggcaac gagctgaaag	300
[3454]	ctcatctgct ctacggggcc cctccctggg gacagcccct cctggctagt cacaccctgt	360
[3455]	aggctcctct atataaccca ggggcacagg ggctgccctc attctaccac cacctccaca	420
[3456]	gcacagacag aactcagga gccagccagc	450
[3457]	<210> 192	
[3458]	<211> 3762	
[3459]	<212> DNA	
[3460]	<213> 人工序列	
[3461]	<220>	
[3462]	<223> 编码dSaCas9(D10A, N580A)-miniVR的DNA序列	
[3463]	<400> 192	
[3464]	atggcccaaa agaagaagcg gaaggtcggg atccacggag tcccagcagc caagcggaac	60
[3465]	tacatcctgg gcctggccat cggcatcacc agcgtgggct acggcatcat cgactacgag	120
[3466]	acacgggacg tgatcgatgc cggcgtgcgg ctgttcaaag aggccaacgt ggaaaacaac	180
[3467]	gagggcaggc ggagcaagag aggcgccaga aggtgaagc ggcggaggcg gcatagaatc	240
[3468]	cagagagtga agaagctgct gttcgactac aacctgctga ccgaccacag cgagctgagc	300
[3469]	ggcatcaacc cctacgaggc cagagtgaag ggcctgagcc agaagctgag cgaggaagag	360
[3470]	ttctctgccc cctgtgca cctggccaag agaagaggcg tgcacaacgt gaacgaggtg	420
[3471]	gaagaggaca ccggcaacga gctgtccacc aaagagcaga tcagccggaa cagcaaggcc	480
[3472]	ctggaagaga aatactggc cgaactgcag ctggaacggc tgaagaaaga cggcgaagt	540
[3473]	cggggcagca tcaacagatt caagaccagc gactacgtga aagaagccaa acagctgctg	600
[3474]	aaggtgcaga aggcctacca ccagctggac cagagcttca tcgacaccta catcgacctg	660
[3475]	ctggaaaccc ggcggacctt ctatgaggga cctggcgagg gcagccccctt cggctggaag	720
[3476]	gacatcaaag aatggtacga gatgctgatg ggccactgca cctacttccc cgaggaactg	780
[3477]	cggagcgtga agtacgccta caacgccgac ctgtacaacg ccctgaacga cctgaacaat	840
[3478]	ctcgtgatca ccaggacga gaacgagaag ctggaatatt acgagaagtt ccagatcatc	900
[3479]	gagaacgtgt tcaagcagaa gaagaagccc accctgaagc agatcgccaa agaaatctc	960
[3480]	gtgaacgaag aggatattaa gggctacaga gtgaccagca ccggcaagcc cgagttcacc	1020
[3481]	aacctgaagg tgtaccacga catcaaggac attaccgccc ggaaagagat tattgagaac	1080
[3482]	gccgagctgc tggatcagat tgccaagatc ctgaccatct accagagcag cgaggacatc	1140
[3483]	caggaagaac tgaccaatct gaactccgag ctgaccagc aagagatcga gcagatctct	1200

[3484]	aatctgaagg gctataccgg caccacaac ctgagcctga aggccatcaa cctgatcctg	1260
[3485]	gacgagctgt ggcacaccaa cgacaaccag atcgctatct tcaaccggct gaagctgggtg	1320
[3486]	cccaagaagg tggacctgtc ccagcagaaa gagatcccca ccaccctggt ggacgacttc	1380
[3487]	atcctgagcc ccgtcgtgaa gagaagcttc atccagagca tcaaagtgat caacgccatc	1440
[3488]	atcaagaagt acggcctgcc caacgacatc attatcgagc tggcccgga gaagaactcc	1500
[3489]	aaggacgccc agaaaatgat caacgagatg cagaagcgga accggcagac caacgagcgg	1560
[3490]	atcgaggaaa tcatccggac caccggcaaa gagaacgcca agtacctgat cgagaagatc	1620
[3491]	aagctgcacg acatgcagga aggcaagtgc ctgtacagcc tggaagccat ccctctggaa	1680
[3492]	gatctgctga acaaccctt caactatgag gtggacgcca tcatccccag aagcgtgtcc	1740
[3493]	ttcgacaaca gttcaacaa caaggtgtc gtgaagcagg aagaagccag caagaagggc	1800
[3494]	aaccggaccc cattccagta cctgagcagc agcgacagca agatcagcta cgaaaccttc	1860
[3495]	aagaagcaca tctgaatct ggccaagggc aagggcagaa tcagcaagac caagaaagag	1920
[3496]	tatctgctgg aagaacggga catcaacagg ttctccgtgc agaaagactt catcaaccgg	1980
[3497]	aacctgggtg ataccagata cgccaccaga ggcctgatga acctgctgcg gagctacttc	2040
[3498]	agagtgaaca acctggacgt gaaagtgaag tccatcaatg gcggcttcac cagctttctg	2100
[3499]	cggcggaagt ggaagtttaa gaaagagcgg aacaaggggt acaagcacca cgccgaggac	2160
[3500]	gccctgatca ttgccaacgc cgatttcac ttcaaagagt ggaagaaact ggacaaggcc	2220
[3501]	aaaaaagtga tggaaaacca gatgttcgag gaaaagcagg ccgagagcat gcccgagatc	2280
[3502]	gaaaccgagc aggagtacaa agagatcttc atcaccccc accagatcaa gcacattaag	2340
[3503]	gacttcaagg actacaagta cagccaccgg gtggacaaga agcctaatag agagctgatt	2400
[3504]	aacgacaccc tgtactccac ccggaaggac gacaagggca acaccctgat cgtgaacaat	2460
[3505]	ctgaacggcc tgtacgacaa ggacaatgac aagctgaaaa agctgatcaa caagagcccc	2520
[3506]	gaaaagctgc tgatgtacca ccacgacccc cagacctacc agaaactgaa gctgattatg	2580
[3507]	gaacagtacg gcgacgagaa gaatccccg tacaagtact acgaggaaac cggaactac	2640
[3508]	ctgaccaagt actccaaaaa ggacaacggc cccgtgatca agaagattaa gtattacggc	2700
[3509]	aacaaactga acgcccctc ggacatcacc gacgactacc ccaacagcag aaacaaggtc	2760
[3510]	gtgaagctgt ccctgaagcc ctacagattc gacgtgtacc tggacaatgg cgtgtacaag	2820
[3511]	ttctgtaccg tgaagaatct ggatgtgatc aaaaaagaaa actactacga agtgaatagc	2880
[3512]	aagtgtatg aggaagctaa gaagctgaag aagatcagca accaggccga gtttatcgcc	2940
[3513]	tccttctaca acaacgatct gatcaagatc aacggcgagc tgtatagagt gatcggcgtg	3000
[3514]	aacaacgacc tgctgaaccg gatcgaagtg aacatgatcg acatcaccta ccgcgagtac	3060
[3515]	ctggaaaaca tgaacgacaa gagggccccc aggatcatta agacaatcgc ctccaagacc	3120
[3516]	cagagcatta agaagtacag cacagacatt ctgggcaacc tgtatgaagt gaaatctaag	3180
[3517]	aagcaccctc agatcatcaa aaagggcaaa aggccggcgg ccacgaaaaa ggccggccag	3240
[3518]	gcaaaaaaga aaaaggctag cgatgcactc gatgattttg acctcgatat gcttgggagt	3300
[3519]	gatgcgctcg atgacttga tttggatatg cttggatctg atgccctcga cgatttcgac	3360
[3520]	cttgatatgc tcgggtcaga cgctttgatg gactttgacc ttgacatgct ggggagcggc	3420
[3521]	tccccgcac cggcagttac accggaggcc agtcacttgc tcgaagatcc tgacgaggaa	3480
[3522]	accagccagg ccgtaaaggc gttgcgggag atggctgaca cagtaatacc caaaaagag	3540
[3523]	gaggctgcga tttgtgggca gatggatttg tcccacctc caccgagagg tcatcttgac	3600
[3524]	gaattgacaa cgacgtcga atccatgacc gaggacctga acctcgatag cccgctcacc	3660
[3525]	cccgagttga atgagatcct ggatacatctt cttaatgatg agtgtttgc tcaacgaatg	3720

[3526] catatttcta cgggtcttag tattttcgac acgagcctgt tt 3762
 [3527] <210> 193
 [3528] <211> 1254
 [3529] <212> PRT
 [3530] <213> 人工序列
 [3531] <220>
 [3532] <223> dSaCas9(D10A, N580A) -miniVR
 [3533] <400> 193
 [3534] Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
 [3535] 1 5 10 15
 [3536] Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
 [3537] 20 25 30
 [3538] Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
 [3539] 35 40 45
 [3540] Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
 [3541] 50 55 60
 [3542] Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
 [3543] 65 70 75 80
 [3544] Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
 [3545] 85 90 95
 [3546] Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
 [3547] 100 105 110
 [3548] Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
 [3549] 115 120 125
 [3550] Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
 [3551] 130 135 140
 [3552] Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
 [3553] 145 150 155 160
 [3554] Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
 [3555] 165 170 175
 [3556] Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
 [3557] 180 185 190
 [3558] Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
 [3559] 195 200 205
 [3560] Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
 [3561] 210 215 220
 [3562] Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
 [3563] 225 230 235 240
 [3564] Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
 [3565] 245 250 255
 [3566] Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
 [3567] 260 265 270

[3568]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[3569]	275	280	285
[3570]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[3571]	290	295	300
[3572]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[3573]	305	310	315
[3574]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[3575]	325	330	335
[3576]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[3577]	340	345	350
[3578]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[3579]	355	360	365
[3580]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[3581]	370	375	380
[3582]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[3583]	385	390	395
[3584]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[3585]	405	410	415
[3586]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[3587]	420	425	430
[3588]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[3589]	435	440	445
[3590]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[3591]	450	455	460
[3592]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[3593]	465	470	475
[3594]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[3595]	485	490	495
[3596]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[3597]	500	505	510
[3598]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[3599]	515	520	525
[3600]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[3601]	530	535	540
[3602]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[3603]	545	550	555
[3604]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp Ala Ile Ile Pro		
[3605]	565	570	575
[3606]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[3607]	580	585	590
[3608]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[3609]	595	600	605

[3610]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
[3611]	610 615 620
[3612]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
[3613]	625 630 635 640
[3614]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
[3615]	645 650 655
[3616]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
[3617]	660 665 670
[3618]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
[3619]	675 680 685
[3620]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
[3621]	690 695 700
[3622]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
[3623]	705 710 715 720
[3624]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
[3625]	725 730 735
[3626]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
[3627]	740 745 750
[3628]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
[3629]	755 760 765
[3630]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
[3631]	770 775 780
[3632]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
[3633]	785 790 795 800
[3634]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
[3635]	805 810 815
[3636]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
[3637]	820 825 830
[3638]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
[3639]	835 840 845
[3640]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
[3641]	850 855 860
[3642]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
[3643]	865 870 875 880
[3644]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
[3645]	885 890 895
[3646]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
[3647]	900 905 910
[3648]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
[3649]	915 920 925
[3650]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
[3651]	930 935 940

[3652]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
[3653]	945 950 955 960
[3654]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
[3655]	965 970 975
[3656]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
[3657]	980 985 990
[3658]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
[3659]	995 1000 1005
[3660]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn
[3661]	1010 1015 1020
[3662]	Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser
[3663]	1025 1030 1035
[3664]	Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn
[3665]	1040 1045 1050
[3666]	Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys
[3667]	1055 1060 1065
[3668]	Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys
[3669]	1070 1075 1080
[3670]	Lys Lys Ala Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
[3671]	1085 1090 1095
[3672]	Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
[3673]	1100 1105 1110
[3674]	Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala
[3675]	1115 1120 1125
[3676]	Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Gly Ser Pro Ala
[3677]	1130 1135 1140
[3678]	Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp
[3679]	1145 1150 1155
[3680]	Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
[3681]	1160 1165 1170
[3682]	Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met
[3683]	1175 1180 1185
[3684]	Asp Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr
[3685]	1190 1195 1200
[3686]	Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro
[3687]	1205 1210 1215
[3688]	Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp
[3689]	1220 1225 1230
[3690]	Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly Leu Ser Ile
[3691]	1235 1240 1245
[3692]	Phe Asp Thr Ser Leu Phe
[3693]	1250

[3694] <210> 194
[3695] <211> 21
[3696] <212> RNA
[3697] <213> 智人(Homo sapiens)
[3698] <220>
[3699] <221> misc_feature
[3700] <222> (1) .. (21)
[3701] <223> 对应于靶序列 (SEQ ID NO:45) 的crRNA
[3702] <400> 194
[3703] guacaguuag ugcua cuagg a 21
[3704] <210> 195
[3705] <211> 21
[3706] <212> RNA
[3707] <213> 智人(Homo sapiens)
[3708] <220>
[3709] <221> misc_feature
[3710] <222> (1) .. (21)
[3711] <223> 对应于靶序列 (SEQ ID NO:46) 的crRNA
[3712] <400> 195
[3713] gcuacuagga caggaugcug g 21
[3714] <210> 196
[3715] <211> 21
[3716] <212> RNA
[3717] <213> 智人(Homo sapiens)
[3718] <220>
[3719] <221> misc_feature
[3720] <222> (1) .. (21)
[3721] <223> 对应于靶序列 (SEQ ID NO:58) 的crRNA
[3722] <400> 196
[3723] gggguccgcu cuccagauga g 21
[3724] <210> 197
[3725] <211> 21
[3726] <212> RNA
[3727] <213> 智人(Homo sapiens)
[3728] <220>
[3729] <221> misc_feature
[3730] <222> (1) .. (21)
[3731] <223> 对应于靶序列 (SEQ ID NO:59) 的crRNA
[3732] <400> 197
[3733] ggaggguggg gcgcaggacc g 21
[3734] <210> 198
[3735] <211> 21

[3736] <212> RNA
[3737] <213> 智人(Homo sapiens)
[3738] <220>
[3739] <221> misc_feature
[3740] <222> (1) .. (21)
[3741] <223> 对应于靶序列 (SEQ ID NO:60) 的crRNA
[3742] <400> 198
[3743] ccucucucgc gcacaaaguu g 21
[3744] <210> 199
[3745] <211> 21
[3746] <212> RNA
[3747] <213> 智人(Homo sapiens)
[3748] <220>
[3749] <221> misc_feature
[3750] <222> (1) .. (21)
[3751] <223> 对应于靶序列 (SEQ ID NO:135) 的crRNA
[3752] <400> 199
[3753] ggcuuccucua ggaguuugac a 21
[3754] <210> 200
[3755] <211> 21
[3756] <212> RNA
[3757] <213> 智人(Homo sapiens)
[3758] <220>
[3759] <221> misc_feature
[3760] <222> (1) .. (21)
[3761] <223> 对应于靶序列 (SEQ ID NO:141) 的crRNA
[3762] <400> 200
[3763] ggguguuucu gcggugacgg a 21
[3764] <210> 201
[3765] <211> 21
[3766] <212> RNA
[3767] <213> 智人(Homo sapiens)
[3768] <220>
[3769] <221> misc_feature
[3770] <222> (1) .. (21)
[3771] <223> 对应于靶序列 (SEQ ID NO:153) 的crRNA
[3772] <400> 201
[3773] gaguccggag accgaaccag a 21
[3774] <210> 202
[3775] <211> 21
[3776] <212> RNA
[3777] <213> 智人(Homo sapiens)

[3778] <220>
[3779] <221> misc_feature
[3780] <222> (1) .. (21)
[3781] <223> 对应于靶序列 (SEQ ID NO:155) 的crRNA
[3782] <400> 202
[3783] gcugggccugg ggcgcgcgcgc c 21
[3784] <210> 203
[3785] <211> 21
[3786] <212> RNA
[3787] <213> 智人 (Homo sapiens)
[3788] <220>
[3789] <221> misc_feature
[3790] <222> (1) .. (21)
[3791] <223> 对应于靶序列 (SEQ ID NO:156) 的crRNA
[3792] <400> 203
[3793] aagaucagcc ccacuacguu c 21
[3794] <210> 204
[3795] <211> 21
[3796] <212> RNA
[3797] <213> 智人 (Homo sapiens)
[3798] <220>
[3799] <221> misc_feature
[3800] <222> (1) .. (21)
[3801] <223> 对应于靶序列 (SEQ ID NO:157) 的crRNA
[3802] <400> 204
[3803] ccggaggcga gccccuuccc g 21
[3804] <210> 205
[3805] <211> 21
[3806] <212> RNA
[3807] <213> 智人 (Homo sapiens)
[3808] <220>
[3809] <221> misc_feature
[3810] <222> (1) .. (21)
[3811] <223> 对应于靶序列 (SEQ ID NO:159) 的crRNA
[3812] <400> 205
[3813] gagcgcugga ggcggaggag g 21
[3814] <210> 206
[3815] <211> 21
[3816] <212> RNA
[3817] <213> 智人 (Homo sapiens)
[3818] <220>
[3819] <221> misc_feature

[3820]	<222> (1) .. (21)
[3821]	<223> 对应于靶序列 (SEQ ID NO:167) 的crRNA
[3822]	<400> 206
[3823]	agaggggacg uggccucuua g 21
[3824]	<210> 207
[3825]	<211> 21
[3826]	<212> RNA
[3827]	<213> 智人 (Homo sapiens)
[3828]	<220>
[3829]	<221> misc_feature
[3830]	<222> (1) .. (21)
[3831]	<223> 对应于靶序列 (SEQ ID NO:172) 的crRNA
[3832]	<400> 207
[3833]	gcgcgcggag cucgggggag g 21
[3834]	<210> 208
[3835]	<211> 34
[3836]	<212> DNA
[3837]	<213> 人工序列
[3838]	<220>
[3839]	<223> 2x sNRP-1 polyA序列
[3840]	<400> 208
[3841]	aaataaaata cgaaatgaaa taaaatacga aatg 34
[3842]	<210> 209
[3843]	<211> 4
[3844]	<212> PRT
[3845]	<213> 人工序列
[3846]	<220>
[3847]	<223> GSGS连接子
[3848]	<400> 209
[3849]	Gly Ser Gly Ser
[3850]	1
[3851]	<210> 210
[3852]	<211> 8288
[3853]	<212> DNA
[3854]	<213> 人工序列
[3855]	<220>
[3856]	<223> pED260质粒
[3857]	<400> 210
[3858]	cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg 60
[3859]	cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga 120
[3860]	caggtttccc gactggaaag cgggcagtga gcgcaacgca attaattgtga gttagctcac 180
[3861]	tcattaggea ccccaggctt tacactttat gcttcggct cgtatgttgt gtggaattgt 240

[3862]	gagcggataa caatttcaca caggaaacag ctatgaccat gattacgaat tgcctgcagg	300
[3863]	cagctgcgcg ctgcctcgct cactgaggcc gcccgggcaa agccccggcg tcgggcgacc	360
[3864]	tttggtcgcc cggcctcagt gagcgcgcga gcgcgcagag agggagtggc caactccatc	420
[3865]	actaggggtt cctatcgata tctactcgagg cgttgctaga ctagcatgct gcccatgtaa	480
[3866]	ggaggcaagg cctggggaca cccgagatgc ctggttataa ttaaccaga catgtggctg	540
[3867]	ccccccccc cccaacacct gctgcctcta aaaataaccc tgcattgcat gttcccggcg	600
[3868]	aagggccagc tgtccccgc cagctagact cagcacttag tttaggaacc agtgagcaag	660
[3869]	tcagcccttg gggcagccca tacaaggcca tggggctggg caagctgcac gcctgggtcc	720
[3870]	ggggtgggca cgggtcccgg gcaacgagct gaaagctcat ctgctctcag gggccctcc	780
[3871]	ctggggacag cccctcctgg ctagtcacac cctgtaggct cctctatata acccaggggc	840
[3872]	acaggggctg ccctcattct accaccacct ccacagcaca gacagacact caggagccag	900
[3873]	ccagcagagc tctctggcta actaccggtg ccaccatggc cccaaagaag aagcgggaag	960
[3874]	tcggtatcca cggagtccca gcagccaagc ggaactacat cctgggcctg gccatcgga	1020
[3875]	tcaccagcgt gggctacggc atcatcgact acgagacacg ggacgtgatc gatgccggcg	1080
[3876]	tgcggtgtt caaagaggcc aacgtggaaa acaacgaggc caggcggagc aagagaggcg	1140
[3877]	ccagaaggct gaagcggcgg aggcggcata gaatccagag agtgaagaag ctgctgttcg	1200
[3878]	actacaacct gctgaccgac cacagcgagc tgagcggcat caaccctac gaggccagag	1260
[3879]	tgaaggcct gagccagaag ctgagcgagg aagagtctc tgccgccctg ctgcacctgg	1320
[3880]	ccaagagaag aggcgtgcac aacgtgaacg aggtggaaga ggacaccggc aacgagctgt	1380
[3881]	ccaccaaaga gcagatcagc cggaacagca aggcctgga agagaaatac gtggccgaac	1440
[3882]	tgcagctgga acggctgaag aaagacggcg aagtgcgggg cagcatcaac agattcaaga	1500
[3883]	ccagcgacta cgtgaaagaa gccaaacagc tgctgaaggt gcagaaggcc taccaccagc	1560
[3884]	tggaccagag ctcatcgac acctacatcg acctgctgga aaccggcg accctactatg	1620
[3885]	agggacctgg cgagggcagc cccttcggct ggaaggacat caaagaatgg tacgagatgc	1680
[3886]	tgatgggcca ctgcacctac ttccccagg aactgcggag cgtgaagtac gcctacaacg	1740
[3887]	ccgacctgta caacgccctg aacgacctga acaatctcgt gatcaccagg gacgagaacg	1800
[3888]	agaagctgga atattacgag aagttccaga tcatcgagaa cgtgttcaag cagaagaaga	1860
[3889]	agcccacct gaagcagatc gccaaagaaa tctcgtgaa cgaagaggat attaagggt	1920
[3890]	acagagtgac cagcaccggc aagcccgagt tcaccaacct gaaggtgtac cagcagatca	1980
[3891]	aggacattac cgcgggaaa gagattattg agaacgccga gctgctggat cagattgcca	2040
[3892]	agatcctgac catctaccag agcagcgagg acatccagga agaactgacc aatctgaact	2100
[3893]	ccgagctgac ccaggaagag atcgagcaga tctctaattc gaagggtat accggcacc	2160
[3894]	acaacctgag cctgaaggcc atcaacctga tcttgacga gctgtggcac accaacgaca	2220
[3895]	accagatgc tatcttcaac cggctgaagc tgggtgcccga gaaggtggac ctgtcccagc	2280
[3896]	agaaagagat cccaccacc ctggtggacg acttcatcct gagccccgtc gtgaagagaa	2340
[3897]	gcttcatcca gagcatcaaa gtgatcaacg ccatcatcaa gaagtacggc ctgcccacg	2400
[3898]	acatcattat cgagctggcc cgcgagaaga actccaagga cgcccagaaa atgatcaacg	2460
[3899]	agatgcagaa gcggaaccgg cagaccaacg agcggatcga ggaaatcatc cggaccaccg	2520
[3900]	gcaaagagaa cgccaagtac ctgatcgaga agatcaagct gcacgacatg caggaaggca	2580
[3901]	agtgctgta cagcctggaa gccatccctc tggaagatct gctgaacaac cccttcaact	2640
[3902]	atgaggtgga cgccatcatc ccagaaagcg tgtccttcga caacagctt aacaacaagg	2700
[3903]	tgctcgtgaa gcaggaagaa gccagcaaga agggcaaccg gacccattc cagtacctga	2760

[3904]	gcagcagcga cagcaagatc agctacgaaa ccttcaagaa gcacatcctg aatctggcca	2820
[3905]	agggcaaggg cagaatcagc aagaccaaga aagagtatct gctggaagaa cgggacatca	2880
[3906]	acaggttctc cgtgcagaaa gacttcatca accggaacct ggtggatacc agatacgcca	2940
[3907]	ccagaggcct gatgaacctg ctgcggagct acttcagagt gaacaacctg gacgtgaaag	3000
[3908]	tgaagtccat caatggcggc ttcaccagct ttctgcggcg gaagtggaag ttttaagaaag	3060
[3909]	agcggaaaca ggggtacaag caccacgccg aggacgccct gatcattgcc aacgccgatt	3120
[3910]	tcattcttaa agagtggaag aaactggaca aggccaaaaa agtgatggaa aaccagatgt	3180
[3911]	tcgaggaaaa gcaggccgag agcatgcccg agatcgaaac cgagcaggag taaaaagaga	3240
[3912]	tcttcatcac cccccaccag atcaagcaca ttaaggactt caaggactac aagtacagcc	3300
[3913]	accgggtgga caagaagcct aatagagagc tgattaacga caccctgtac tccacccgga	3360
[3914]	aggacgacaa gggcaacacc ctgatctgta acaatctgaa cggcctgtac gacaaggaca	3420
[3915]	atgacaagct gaaaaagctg atcaacaaga gccccgaaaa gctgctgatg taccaccacg	3480
[3916]	acccccagac ctaccagaaa ctgaagctga ttatggaaca gtacggcgac gagaagaatc	3540
[3917]	ccctgtacaa gtactacgag gaaaccggga actacctgac caagtactcc aaaaaggaca	3600
[3918]	acggccccgt gatcaagaag attaagtatt acggcaacaa actgaacgcc catctggaca	3660
[3919]	tcaccgacga ctaccccaac agcagaaaca aggtcgtgaa gctgtccctg aagccctaca	3720
[3920]	gattcgacgt gtacctggac aatggcgtgt acaagtctgt gaccgtgaag aatctggatg	3780
[3921]	tgatcaaaaa agaaaactac tacgaagtga atagcaagtg ctatgaggaa gctaagaagc	3840
[3922]	tgaagaagat cagcaaccag gccgagtta tcgcctcctt ctacaacaac gatctgatca	3900
[3923]	agatcaacgg cgagctgtat agagtgatcg gcgtgaacaa cgacctgtg aaccggatcg	3960
[3924]	aagtgaacat gatcgacatc acctaccgag agtacctgga aaacatgaac gacaagaggc	4020
[3925]	cccccaggat cattaagaca atgcctcca agaccagag cattaagaag tacagcacag	4080
[3926]	acattctggg caacctgtat gaagtgaat ctaagaagca ccctcagatc atcaaaaagg	4140
[3927]	gcaaaaaggcc ggcgccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccggtg	4200
[3928]	ctagagatgc actcgatgat tttagacctg atatgcttgg gagtgatgag ctcgatgact	4260
[3929]	tcgatttggg tatgcttggg tctgatgccc tcgacgattt cgaccttgat atgctcgggt	4320
[3930]	cagacgcttt ggatgacttt gaccttgaca tgctggggag cggctcccc gcaccggcag	4380
[3931]	ttacaccgga ggccagtcac ttgctcgaag atcctgacga ggaaaccagc caggccgtaa	4440
[3932]	aggcgttgag ggagatggct gacacagtaa taccceaaaa agaggagget gcgatttgtg	4500
[3933]	ggcagatgga tttgtccac cctccaccga gaggtcatct tgacgaattg acaacgacgc	4560
[3934]	tcgaatccat gaccgaggac ctgaacctcg atagcccgt ccccccgag ttgaatgaga	4620
[3935]	tcctggatac atttcttaat gatgagtgtt tgcttcacgc aatgcatatt tctacgggtc	4680
[3936]	ttagtatttt cgacacgagc ctgtttgtcg actgagaatt cctagagctc gctgatcagc	4740
[3937]	ctcgactgtg ctttctagtt gccagccatc tgttgtttgc cctcccccg tgccttcctt	4800
[3938]	gaccttgaa ggtgccactc ccaactgtcct ttcctaataa aatgaggaaa ttgcatcgca	4860
[3939]	ttgtctgagt aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga	4920
[3940]	ggattgggaa gagaatagca ggcattgctg ggaggtaccg agggcctatt tccatgatt	4980
[3941]	ccttcatatt tgcatatacg atacaaggct gttagagaga taattggaat taatttgact	5040
[3942]	gtaaacacaa agatattagt acaaaatagc tgacgtagaa agtaataatt tcttgggtag	5100
[3943]	tttgagttt taaaattatg ttttaaaatg gactatcata tgcttaccgt aacttgaaag	5160
[3944]	tatttcgatt tcttggcttt atatatcttg tggaaaggac gaaacaccgg agaccacggc	5220
[3945]	aggtctcagt tttagtactc tggaaacaga atctactaaa acaaggcaaa atgccgtgtt	5280

[3946]	tatctcgtca	acttgttggc	gagatttttg	cggccgcgat	ctaggaaccc	ctagtgatgg	5340
[3947]	agttggccac	tccctctctg	cgcgctcgct	cgctcactga	ggccgggcca	ccaaaggtcg	5400
[3948]	cccgcgccc	gggttttgc	cgggcggcct	cagtgcgcga	gcgagcgcg	agctgcctgc	5460
[3949]	aggcagcttg	gcaactggcg	tcgttttaca	acgtcgtgac	tgggaaaacc	ctggcgttac	5520
[3950]	ccaacttaat	cgccttgacg	cacatccccc	tttcgccagc	tggcgtaata	gcgaagaggc	5580
[3951]	ccgcaccgat	cgccttccc	aacagttgcg	cagcctgaat	ggcgaatggc	gcctgatgcg	5640
[3952]	gtattttctc	cttacgcac	tgtgcggtat	ttcacaccgc	atacgtcaaa	gcaaccatag	5700
[3953]	tacgcgcct	gtagcggcg	attaagcgcg	gcgggtgtgg	tggttacgcg	cagcgtgacc	5760
[3954]	gtacacattg	ccagcgcct	agcgcgcct	cctttcgctt	tcttcccttc	ctttctcgcc	5820
[3955]	acgttcgcg	gctttcccc	tcaagctcta	aatcgggggc	tccctttagg	gttccgattt	5880
[3956]	agtgccttac	ggcacctcga	ccccaaaaa	cttgatttgg	gtgatgggtc	acgtagtggg	5940
[3957]	ccatcgccct	gatagacggt	ttttcgccct	ttgacgttgg	agtccacgtt	ctttaatagt	6000
[3958]	ggactcttgt	tccaaactgg	aacaacactc	aaccctatct	cgggctattc	ttttgattta	6060
[3959]	taagggattt	tgccgatttc	ggcctattgg	ttaaaaaatg	agctgattta	acaaaaattt	6120
[3960]	aacgcgaatt	ttaacaaaat	attaacgttt	acaattttat	ggtgcactct	cagtacaatc	6180
[3961]	tgtctgatg	ccgcatagtt	aagccagccc	cgacacccgc	caacacccgc	tgacgcgccc	6240
[3962]	tgacgggctt	gtctgtccc	ggcatccgct	tacagacaag	ctgtgaccgt	ctccgggagc	6300
[3963]	tgcatgtgtc	agaggttttc	accgtcatca	cgaacacgcg	cgagacgaaa	gggcctcgtg	6360
[3964]	atacgcctat	ttttataggt	taatgtcatg	ataataatgg	tttcttagac	gtcaggtggc	6420
[3965]	acttttcggg	gaaatgtgcg	cggaaacccct	atttgtttat	ttttctaaat	acattcaaat	6480
[3966]	atgtatccgc	tcatgagaca	ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	6540
[3967]	agtatgagta	ttcaacattt	ccgtgtcgcc	cttattccct	tttttgcggc	attttgcctt	6600
[3968]	cctgtttttg	ctcaccaga	aacgctggtg	aaagtaaaag	atgctgaaga	tcagttgggt	6660
[3969]	gcacgagtgg	gttacatcga	actggatctc	aacagcggta	agatccttga	gagttttcgc	6720
[3970]	cccgaagaac	gttttccaat	gatgagcact	tttaaagtgc	tgctatgtgg	cgcggtatta	6780
[3971]	tcccgatttg	acgcgggcca	agagcaactc	ggtcgcccga	tacactattc	tcagaatgac	6840
[3972]	ttgggtgagt	actcaccagt	cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	6900
[3973]	ttatgcagtg	ctgccataac	catgagtgat	aacactgcgg	ccaacttact	tctgacaacg	6960
[3974]	atcggaggac	cgaaggagct	aaccgctttt	ttgcacaaca	tgggggatca	tgtaactcgc	7020
[3975]	cttgatcggt	gggaaccgga	gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	7080
[3976]	atgcctgtag	caatggcaac	aacgttgcgc	aaactattaa	ctggcgaaact	acttactcta	7140
[3977]	gcttcccggc	aacaattaat	agactggatg	gaggcggata	aagttgcagg	accacttctg	7200
[3978]	cgtcggcccc	ttccggctgg	ctggtttatt	gctgataaat	ctggagccgg	tgagcgtgga	7260
[3979]	agccgcggta	tcattgcagc	actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	7320
[3980]	tacacgacgg	ggagtcaggc	aactatggat	gaacgaaata	gacagatcgc	tgagataggt	7380
[3981]	gcctcactga	ttaagcattg	gtaactgtca	gaccaagttt	actcatatat	acttttagatt	7440
[3982]	gatttaaaac	ttcattttta	atttaaaagg	atctaggtga	agatcctttt	tgataatctc	7500
[3983]	atgacaaaa	tcccttaacg	tgagttttcg	ttccactgag	cgtcagaccc	cgtagaaaag	7560
[3984]	atcaaaggat	cttcttgaga	tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	7620
[3985]	aaaccaccgc	taccagcggg	ggtttgtttg	ccgatcaag	agctaccaac	tctttttccg	7680
[3986]	aaggtaactg	gcttcagcag	agcgcagata	ccaaatactg	ttcttctagt	gtagccgtag	7740
[3987]	ttaggccacc	acttcaagaa	ctctgtagca	ccgcctacat	acctcgtctt	gctaactctg	7800

[3988]	ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga	7860
[3989]	tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc	7920
[3990]	ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc	7980
[3991]	acgcttcccg aagggagaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga	8040
[3992]	gagcgcacga gggagcttcc agggggaaac gcctggatc tttatagtcc tgtcgggttt	8100
[3993]	cgccacctct gacttgagcg tcgatttttg tgatgctcgt caggggggcg gaggcctatgg	8160
[3994]	aaaaacgcca gcaacgcggc ctttttacgg ttcttggcct ttgtctggcc ttttgctcac	8220
[3995]	atgttctttc ctgcgttata ccctgattct gtggataacc gtattaccgc ctttgagtga	8280
[3996]	gctgatac	8288
[3997]	<210>	211
[3998]	<211>	8090
[3999]	<212>	DNA
[4000]	<213>	人工序列
[4001]	<220>	
[4002]	<223>	pED261质粒
[4003]	<400>	211
[4004]	cgctcgccgc agccgaacga ccgagcgag cgagtcagtg agcgaggaag cggaagagcg	60
[4005]	cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga	120
[4006]	caggtttccc gactggaaag cgggcagtga gcgcaacgca attaattgta gttagctcac	180
[4007]	tcattaggea cccagagctt tacactttat gcttccggt cgtatgttgt gtggaattgt	240
[4008]	gagcggataa caatttcaca caggaaacag ctatgaccat gattacgaat tgcttcgagg	300
[4009]	cagctgcgcg ctgcctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc	360
[4010]	tttggtcgcc cggcctcagt gagcgcgcga gcgcgcagag agggagtggt caactccatc	420
[4011]	actaggggtt cctatcgata tactcgagg cgttgctaga ctagcatgct gcccatgtaa	480
[4012]	ggaggcaagg cctggggaca cccgagatgc ctggttataa ttaaccaga catgtggctg	540
[4013]	ccccccccc cccaacacct gctgcctcta aaaataaccc tgcatgcat gttcccggcg	600
[4014]	aagggccagc tgtccccgc cagctagact cagcacttag tttaggaacc agtgagcaag	660
[4015]	tcagcccttg ggcagccca tacaaggcca tggggctggg caagctgcac gcctgggtcc	720
[4016]	ggggtgggca cgggtcccg gcaacgagct gaaagctcat ctgctctcag gggeccctcc	780
[4017]	ctggggacag cccctcctgg ctagtacac cctgtaggct cctctatata acccaggggc	840
[4018]	acaggggctg ccctcattct accaccacct ccacagcaca gacagacact caggagccag	900
[4019]	ccagcagagc tctctggcta actaccgtg ccaccatggc ccaaagaag aagcggaag	960
[4020]	tcggtatcca cggagtccca gcagccaagc ggaactacat cctgggcctg gccatcgga	1020
[4021]	tcaccagcgt gggctacggc atcatcgact acgagacacg ggacgtgatc gatgccggcg	1080
[4022]	tgcggtgtt caaagaggcc aacgtggaaa acaacgaggg caggcggagc aagagaggcg	1140
[4023]	ccagaaggct gaagcggcgg aggcggcata gaatccagag agtgaagaag ctgctgttcg	1200
[4024]	actacaacct gctgaccgac cacagcgagc tgagcggcat caaccctac gaggccagag	1260
[4025]	tgaaggcct gagccagaag ctgagcgagg aagagtctc tgccgccctg ctgcacctgg	1320
[4026]	ccaagagaag aggcgtgcac aacgtgaac aggtggaaga ggacaccggc aacgagctgt	1380
[4027]	ccaccaaaga gcagatcagc cggaacagca aggccttga agagaaatac gtggccgaac	1440
[4028]	tgcagctgga acggctgaag aaagacggcg aagtgcgggg cagcatcaac agattcaaga	1500
[4029]	ccagcgacta cgtgaaagaa gccaaacagc tgctgaaggt gcagaaggcc taccaccagc	1560

[4030]	tggaccagag	cttcatcgac	acctacatcg	acctgctgga	aacccggcgg	acctactatg	1620
[4031]	agggacctgg	cgagggcagc	cccttcggct	ggaaggacat	caaagaatgg	tacgagatgc	1680
[4032]	tgatgggcca	ctgcacctac	ttccccgagg	aactgcggag	cgtgaagtac	gcctacaacg	1740
[4033]	ccgacctgta	caacgccctg	aacgacctga	acaatctcgt	gatcaccagg	gacgagaacg	1800
[4034]	agaagctgga	atattacgag	aagttccaga	tcatcgagaa	cgtgttcaag	cagaagaaga	1860
[4035]	agccccacct	gaagcagatc	gccaaagaaa	tcctcgtgaa	cgaagaggat	attaagggtc	1920
[4036]	acagagtgc	cagcaccggc	aagccccagt	tcaccaacct	gaaggtgtac	cacgacatca	1980
[4037]	aggacattac	cgccccgaaa	gagattattg	agaacgccga	gctgctggat	cagattgcc	2040
[4038]	agatcctgac	catctaccag	agcagcgagg	acatccagga	agaactgacc	aatctgaact	2100
[4039]	ccgagctgac	ccaggaagag	atcgagcaga	tctctaattc	gaagggctat	accggcacc	2160
[4040]	acaacctgag	cctgaaggcc	atcaacctga	tcctggacga	gctgtggcac	accaacgaca	2220
[4041]	accagatgc	tatcttcaac	cggctgaagc	tggtgcccc	gaaggtggac	ctgtcccagc	2280
[4042]	agaaagagat	ccccaccacc	ctggtggacg	acttcatcct	gagccccgtc	gtgaagagaa	2340
[4043]	gcttcatcca	gagcatcaaa	gtgatcaacg	ccatcatcaa	gaagtacggc	ctgccccaacg	2400
[4044]	acatcattat	cgagctggcc	cgcgagaaga	actccaagga	cgcccagaaa	atgatcaacg	2460
[4045]	agatgcagaa	gcggaaccgg	cagaccaacg	agcggatcga	ggaaatcatc	cggaccaccg	2520
[4046]	gcaaagagaa	cgccaagtac	ctgatcgaga	agatcaagct	gcacgacatg	caggaaggca	2580
[4047]	agtgctgta	cagcctggaa	gccatccctc	tggaagatct	gctgaacaac	cccttcaact	2640
[4048]	atgaggtgga	cgccatcatc	cccagaagcg	tgctcttcga	caacagcttc	aacaacaagg	2700
[4049]	tgctcgtgaa	gcaggaagaa	gccagcaaga	agggcaaccg	gacccattc	cagtacctga	2760
[4050]	gcagcagcga	cagcaagatc	agctacgaaa	ccttcaagaa	gcacatcctg	aatctggcca	2820
[4051]	agggaaggg	cagaatcagc	aagaccaaga	aagagtatct	gctggaagaa	cgggacatca	2880
[4052]	acaggttctc	cgtgcagaaa	gacttcatca	accggaacct	ggtggatacc	agatacgcca	2940
[4053]	ccagaggcct	gatgaacctg	ctgcggagct	acttcagagt	gaacaacctg	gacgtgaaag	3000
[4054]	tgaagtccat	caatggcggc	ttcaccagct	ttctgcggcg	gaagtggaag	tttaagaaag	3060
[4055]	agcgaacaa	gggttacaag	caccacgccg	aggacgccct	gatcattgcc	aacgccgatt	3120
[4056]	tcattctcaa	agagtggaag	aaactggaca	aggccaaaa	agtgatggaa	aaccagatgt	3180
[4057]	tcgagaaaa	gcaggccgag	agcatgcccc	agatcgaaac	cgagcaggag	tacaaagaga	3240
[4058]	ttttcatcac	ccccaccag	atcaagcaca	ttaaggactt	caaggactac	aagtacagcc	3300
[4059]	accgggtgga	caagaagcct	aatagagagc	tgattaacga	cacctgtac	tccaccggga	3360
[4060]	aggacgacaa	gggcaacacc	ctgatcgtga	acaatctgaa	cggcctgtac	gacaaggaca	3420
[4061]	atgacaagct	gaaaaagctg	atcaacaaga	gccccgaaaa	gctgctgatg	taccaccacg	3480
[4062]	acccccagac	ctaccagaaa	ctgaagctga	ttatggaaca	gtacggcgac	gagaagaatc	3540
[4063]	ccctgtacaa	gtactacgag	gaaaccggga	actacctgac	caagtactcc	aaaaaggaca	3600
[4064]	acggccccgt	gatcaagaag	attaagtatt	acggcaacaa	actgaacgcc	catctggaca	3660
[4065]	tcaccgacga	ctacccaac	agcagaaaca	aggtcgtgaa	gctgtccctg	aagccctaca	3720
[4066]	gattcgacgt	gtacctggac	aatggcgtgt	acaagtctgt	gaccgtgaag	aatctggatg	3780
[4067]	tgatcaaaaa	agaaaactac	tacgaagtga	atagcaagtg	ctatgaggaa	gctaagaagc	3840
[4068]	tgaagaagat	cagcaaccag	gccgagttta	tcgcctcctt	ctacaacaac	gatctgatca	3900
[4069]	agatcaacgg	cgagctgtat	agagtgatcg	gcgtgaacaa	cgacctgctg	aaccggatcg	3960
[4070]	aagtgaacat	gatcgacatc	acctaccgcg	agtacctgga	aaacatgaac	gacaagaggc	4020
[4071]	cccccaggat	cattaagaca	atcgcttcca	agaccagag	cattaagaag	tacagcacag	4080

[4072]	acattctggg	caacctgtat	gaagtgaat	ctaagaagca	ccctcagatc	atcaaaaagg	4140
[4073]	gcaaaaaggcc	ggcggccacg	aaaaaggccg	gccaggcaaa	aaagaaaaag	ggatccgggtg	4200
[4074]	ctagagatgc	actcgatgat	tttgacctcg	atatgcttgg	gagtgatgcg	ctcgatgact	4260
[4075]	tcgatttggg	tatgcttggg	tctgatgccc	tcgacgattt	cgaccttgat	atgctcgggt	4320
[4076]	cagacgcttt	ggatgacttt	gaccttgaca	tgctggggag	cggctcccc	gcaccggcag	4380
[4077]	ttacaccgga	ggccagtcac	ttgctcgaag	atcctgacga	ggaaaccagc	caggccgtaa	4440
[4078]	aggcgttgcg	ggagatggct	gacacagtaa	taccccaaaa	agaggaggct	gcgatttgtg	4500
[4079]	ggcagatgga	tttgtccac	cctccaccga	gaggtcatct	tgacgaattg	acaacgacgc	4560
[4080]	tcgaatccat	gaccgaggac	ctgaacctcg	atagcccgt	cacccccgag	ttgaatgaga	4620
[4081]	tcctggatac	atttcttaat	gatgagtgtt	tgcttcacgc	aatgcatatt	tctacgggtc	4680
[4082]	ttagtatttt	cgacacgagc	ctgtttgtcg	actgagaatt	caaataaaat	acgaaatgaa	4740
[4083]	ataaaatacg	aatgggtac	cgagggccta	tttcccatga	tcccttcata	tttgcatata	4800
[4084]	cgatacaagg	ctgttagaga	gataattgga	attaatttga	ctgtaaacac	aaagatatta	4860
[4085]	gtacaaaata	cgtgacgtag	aaagtaataa	tttcttgggt	agtttgagc	tttaaaatta	4920
[4086]	tgttttaaaa	tggactatca	tatgcttacc	gtaacttgaa	agtatttcga	tttcttggct	4980
[4087]	ttatatatct	tgtggaaagg	acgaaacacc	ggagaccacg	gcaggtctca	gttttagtac	5040
[4088]	tctggaaaca	gaatctacta	aaacaaggca	aatgccgtg	tttatctcgt	caacttgttg	5100
[4089]	gcgagatttt	tgcgcccgcg	atctaggaac	ccctagtgat	ggagttggcc	actccctctc	5160
[4090]	tgcgcgctcg	ctcgctcact	gaggccgggc	gaccaaagg	cgcccgcgc	ccgggctttg	5220
[4091]	ccggggcggc	ctcagtgage	gagcagcgcg	gcagctgcct	gcaggcagct	tggcactggc	5280
[4092]	cgtcgtttta	caacgtcgtg	actgggaaaa	ccctggcgtt	acccaactta	atcgcccttg	5340
[4093]	agcacatccc	cctttcgcca	gctggcgtaa	tagcgaagag	gcccgcaccg	atcgcccttc	5400
[4094]	ccaacagttg	cgcagcctga	atggcgaatg	gcgcctgatg	cggtattttc	tccttacgca	5460
[4095]	tctgtgcggg	atttcacacc	gcatacgtca	aagcaaccat	agtacgcgcc	ctgtagcggc	5520
[4096]	gcattaaagc	cggcgggtgt	ggtggttacg	cgcagcgtga	ccgctacact	tgccagcgcc	5580
[4097]	ctagcgcccg	ctcctttcgc	tttcttccct	tcctttctcg	ccacgttcgc	cggctttccc	5640
[4098]	cgtcaagctc	taaatcgggg	gctcccttta	gggttccgat	ttagtgcttt	acggcacctc	5700
[4099]	gaccccaaaa	aacttgattt	gggtgatggt	tcacgtagt	ggccatcgcc	ctgatagacg	5760
[4100]	gtttttcgcc	ctttgacgtt	ggagtccacg	ttctttaata	gtggactctt	gttccaaact	5820
[4101]	ggaacaacac	tcaaccctat	ctcgggctat	tccttttgatt	tataagggat	tttgccgatt	5880
[4102]	tcggcctatt	ggttaaaaaa	tgagctgatt	taacaaaaat	ttaacgcgaa	ttttaacaaa	5940
[4103]	atattaacgt	ttacaatttt	atggtgcact	ctcagtacaa	tctgctctga	tgccgcatag	6000
[4104]	ttaagccagc	cccgcacccc	gccaacaccc	gctgacgcgc	cctgacgggc	ttgtctgctc	6060
[4105]	ccggcatccg	cttacagaca	agctgtgacc	gtctccggga	gctgcatgtg	tcagaggttt	6120
[4106]	tcaccgtcat	caccgaaacg	cgcgagacga	aaggccctcg	tgatacgctt	atttttatag	6180
[4107]	gttaatgtca	tgataataat	ggtttcttag	acgtcagggtg	gcacttttcg	gggaaatgtg	6240
[4108]	cgcggaaccc	ctatttgttt	atttttctaa	atacattcaa	atatgtatcc	gctcatgaga	6300
[4109]	caataaccct	gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	6360
[4110]	ttccgtgtcg	cccttattec	cttttttgcg	gcattttgcc	ttcctgtttt	tgctcaccca	6420
[4111]	gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	6480
[4112]	gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgtttttcca	6540
[4113]	atgatgagca	cttttaaaagt	tctgctatgt	ggcgcggtat	tatcccgat	tgacgccggg	6600

[4114]	caagagcaac tcggtcgccg catacactat tctcagaatg acttggttga gtactcacca	6660
[4115]	gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata	6720
[4116]	accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag	6780
[4117]	ctaaccgctt ttttgacaa catgggggat catgtaactc gccttgatcg ttgggaaccg	6840
[4118]	gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca	6900
[4119]	acaacgttgc gcaaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta	6960
[4120]	atagactgga tggaggcgga taaagttgca ggaccacttc tgcgctcgcc ctttccggct	7020
[4121]	ggctggttta ttgctgataa atctggagcc ggtgagcgtg gaagccgcgg tatcattgca	7080
[4122]	gcactggggc cagatggtaa gccctcccgat atcgtagtta tctacacgac ggggagtcag	7140
[4123]	gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat	7200
[4124]	tggttaactgt cagaccaagt ttactcatat atacctttaga ttgattttaa acttcatttt	7260
[4125]	taatttataa ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa	7320
[4126]	cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga	7380
[4127]	gacctttttt ttctgcgcgt aatctgctgc ttgcaacaa aaaaaccacc gctaccagcg	7440
[4128]	gtggtttgtt tgccggatca agagctacca actcttttcc cgaaggtaac tggcttcagc	7500
[4129]	agagcgcaga taccaaatac tgttcttcta gtgtagccgt agttaggcca ccacttcaag	7560
[4130]	aactctgtag caccgcctac atacctcgtc ctgctaatac tgttaccagt ggctgctgcc	7620
[4131]	agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg	7680
[4132]	cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttgagcgc aacgacctac	7740
[4133]	accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga	7800
[4134]	aagcgggaca ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt	7860
[4135]	ccagggggaa acgcctggta tctttatagt cctgtcgggt ttgccacct ctgacttgag	7920
[4136]	cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg	7980
[4137]	gcctttttac ggttcttggc cttttgctgg cttttgctc acatgttctt tctgcgtta	8040
[4138]	tccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac	8090
[4139]	<210> 212	
[4140]	<211> 8000	
[4141]	<212> DNA	
[4142]	<213> 人工序列	
[4143]	<220>	
[4144]	<223> pED263质粒	
[4145]	<400> 212	
[4146]	cgctcgccgc agccgaacga ccgagcgag cgagtcagt agcgaggaag cggaagagcg	60
[4147]	cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga	120
[4148]	caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga gttagctcac	180
[4149]	tcattaggca cccagcgtt tacactttat gcttccggct cgtatgttgt gtggaattgt	240
[4150]	gagcggataa caatttcaca caggaaacag ctatgaccat gattacgaat tgcctgcagg	300
[4151]	cagctgcgcg ctcgctcgt cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc	360
[4152]	tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc	420
[4153]	actaggggtt cctatcgata tctctgagg cgttgctaga ctagcatgct gcccatgtaa	480
[4154]	ggaggcaagg cctggggaca cccgagatgc ctggttataa ttaaccaga catgtggctg	540
[4155]	ccccccccc cccaacacct gctgcctcta aaaataaccc tgcatgcat gttcccgcg	600

[4156]	aagggccagc	tgtcccccgc	cagctagact	cagcacttag	tttaggaacc	agtgagcaag	660
[4157]	tcagcccttg	gggcagccca	tacaaggcca	tggggctggg	caagctgcac	gcctgggtcc	720
[4158]	gggggtggga	cgggtcccgg	gcaacgagct	gaaagctcat	ctgctctcag	gggcccctcc	780
[4159]	ctggggacag	cccctcctgg	ctagtcacac	cctgtaggct	cctctatata	accaggggc	840
[4160]	acaggggctg	ccctcattct	accaccacct	ccacagcaca	gacagacact	caggagccag	900
[4161]	ccagcagagc	tctctggcta	actaccggtg	ccaccatggc	cccaaagaag	aagcggaagg	960
[4162]	tcggtatcca	cggagtccca	gcagccaagc	ggaactacat	cctggggcctg	gccatcgcca	1020
[4163]	tcaccagcgt	gggctacggc	atcatcgact	acgagacacg	ggacgtgac	gatgccggcg	1080
[4164]	tgcggctgtt	caaagaggcc	aacgtggaaa	acaacgaggg	caggcggagc	aagagaggcg	1140
[4165]	ccagaaggct	gaagcggcgg	aggcggcata	gaatccagag	agtgaagaag	ctgctgttcg	1200
[4166]	actacaacct	gctgaccgac	cacagcgagc	tgagcggcat	caaccctac	gaggccagag	1260
[4167]	tgaaggcct	gagccagaag	ctgagcgagg	aagagttctc	tgccgccctg	ctgcacctgg	1320
[4168]	ccaagagaag	aggcgtgcac	aacgtgaacg	aggtggaaga	ggacaccggc	aacgagctgt	1380
[4169]	ccaccaaaga	gcagatcagc	cggaacagca	aggccctgga	agagaaatac	gtggccgaac	1440
[4170]	tgcagctgga	acggctgaag	aaagacggcg	aagtgcgggg	cagcatcaac	agattcaaga	1500
[4171]	ccagcgacta	cgtgaaagaa	gccaaacagc	tgctgaaggt	gcagaaggcc	taccaccagc	1560
[4172]	tggaccagag	cttcatcgac	acctacatcg	acctgctgga	aaccggcg	acctactatg	1620
[4173]	agggacctgg	cgagggcagc	cccttcggct	ggaaggacat	caaagaatgg	tacgagatgc	1680
[4174]	tgatgggcca	ctgcacctac	ttccccgagg	aactgcggag	cgtgaagtac	gcctacaacg	1740
[4175]	ccgacctgta	caacgccctg	aacgacctga	acaatctcgt	gatcaccagg	gacgagaacg	1800
[4176]	agaagctgga	atattacgag	aagttccaga	tcatcgagaa	cgtgttcaag	cagaagaaga	1860
[4177]	agcccacct	gaagcagatc	gccaaagaaa	tcctcgtgaa	cgaagaggat	attaagggt	1920
[4178]	acagagtgac	cagcaccggc	aagcccaggt	tcaccaacct	gaaggtgtac	cacgacatca	1980
[4179]	aggacattac	cgcccggaaa	gagattattg	agaacgccga	gctgctggat	cagattgcca	2040
[4180]	agatcctgac	catctaccag	agcagcgagg	acatccagga	agaactgacc	aatctgaact	2100
[4181]	ccgagctgac	ccaggaagag	atcgagcaga	tctctaattc	gaagggtat	accggcacc	2160
[4182]	acaacctgag	cctgaaggcc	atcaacctga	tcctggacga	gctgtggcac	accaacgaca	2220
[4183]	accagatcgc	tatcttcaac	cggctgaagc	tggtgcccga	gaaggtggac	ctgtcccagc	2280
[4184]	agaaagagat	ccccaccacc	ctggtggacg	acttcatcct	gagccccgtc	gtgaagagaa	2340
[4185]	gcttcatcca	gagcatcaaa	gtgatcaacg	ccatcatcaa	gaagtacggc	ctgcccacg	2400
[4186]	acatcattat	cgagctggcc	cgcgagaaga	actccaagga	cgcccagaaa	atgatcaacg	2460
[4187]	agatgcagaa	gcggaaccgg	cagaccaacg	agcggatcga	ggaaatcatc	cggaccaccg	2520
[4188]	gcaaagagaa	cgccaagtac	ctgatcgaga	agatcaagct	gcacgacatg	caggaaggca	2580
[4189]	agtgcctgta	cagcctggaa	gccatccctc	tggaagatct	gctgaacaac	cccttcaact	2640
[4190]	atgaggtgga	cgccatcatc	cccagaagcg	tgtccttcga	caacagcttc	aacaacaagg	2700
[4191]	tgctcgtgaa	gcaggaagaa	gccagcaaga	agggcaaccg	gacccattc	cagtacctga	2760
[4192]	gcagcagcga	cagcaagatc	agctacgaaa	ccttcaagaa	gcacatcctg	aatctggcca	2820
[4193]	agggcaagg	cagaatcagc	aagaccaaga	aagagtatct	gctggaagaa	cgggacatca	2880
[4194]	acaggttctc	cgtgcagaaa	gacttcatca	accggaacct	ggtggatacc	agatacgcca	2940
[4195]	ccagaggcct	gatgaacctg	ctgcggagct	acttcagagt	gaacaacctg	gacgtgaaag	3000
[4196]	tgaagtccat	caatggcggc	ttcaccagct	ttctgcggcg	gaagtggaag	tttaagaaag	3060
[4197]	agcggaacaa	ggggtacaag	caccacgccg	aggacgccct	gatcattgcc	aacgccgatt	3120

[4198]	tcattcttcaa agagtggag aaactggaca aggccaaaaa agtgatggaa aaccagatgt	3180
[4199]	tcgaggaaaa gcaggccgag agcatgccc agatcgaaac cgagcaggag taaaaagaga	3240
[4200]	tcttcatcac cccccaccag atcaagcaca ttaaggactt caaggactac aagtacagcc	3300
[4201]	accgggtgga caagaagcct aatagagagc tgattaacga caccctgtac tccacccgga	3360
[4202]	aggacgacaa gggcaacacc ctgatcgtga acaatctgaa cggcctgtac gacaaggaca	3420
[4203]	atgacaagct gaaaaagctg atcaacaaga gccccgaaaa gctgctgatg taccaccacg	3480
[4204]	acccccagac ctaccagaaa ctgaagctga ttatggaaac gtacggcgac gagaagaatc	3540
[4205]	ccctgtacaa gtactacgag gaaaccggga actacctgac caagtactcc aaaaaggaca	3600
[4206]	acggccccgt gatcaagaag attaatgtatt acggcaacaa actgaacgcc catctggaca	3660
[4207]	tcaccgacga ctaccccaac agcagaaaca aggtcgtgaa gctgtccctg aagccctaca	3720
[4208]	gattcgacgt gtacctggac aatggcgtgt acaagttcgt gaccgtgaag aatctggatg	3780
[4209]	tgatcaaaaa agaaaactac tacgaagtga atagcaagt ctatgaggaa gctaagaagc	3840
[4210]	tgaagaagat cagcaaccag gccgagtta tcgcctcctt ctacaacaac gatctgatca	3900
[4211]	agatcaacgg cgagctgtat agagtgatcg gcgtgaacaa cgacctgctg aaccggatcg	3960
[4212]	aagtgaacat gatcgacatc acctaccgag agtacctgga aaacatgaac gacaagaggc	4020
[4213]	cccccaggat cattaagaca atcgctcca agaccagag cattaagaag tacagcacag	4080
[4214]	acattctggg caacctgtat gaagtgaat ctaagaagca ccctcagatc atcaaaaagg	4140
[4215]	gcaaaaaggcc ggcgccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgatg	4200
[4216]	cactcgatga ttttgacctc gatatgcttg ggagtgatgc gctcgatgac ttcgatttgg	4260
[4217]	atatgcttgg atctgatgcc ctgacgatt tcgacctga tatgctcggg tcagacgctt	4320
[4218]	tgatgactt tgacctgac atgctgggga gcggctccc ggagatggct gacacagtaa	4380
[4219]	taccccaaaa agaggaggct gcgatttgtg ggcagatgga tttgtccac cctccaccga	4440
[4220]	gaggtcatct tgacgaattg acaacgacgc tcgaatccat gaccgaggac ctgaacctcg	4500
[4221]	atagcccgct ccccccgag ttgaatgaga tcctggatac atttcttaat gatgagtgtt	4560
[4222]	tgcttcacgc aatgcatatt tctacgggtc ttagtatatt cgacacgagc ctgtttgtcg	4620
[4223]	actgagaatt caaataaaat acgaaatgaa ataaaatacg aaatgggtac cgagggccta	4680
[4224]	tttcccatga ttccttcata tttgcatata cgatacaagg ctgttagaga gataattgga	4740
[4225]	attaatttga ctgtaaacac aaagatatta gtacaaaata cgtgacgtag aaagtaataa	4800
[4226]	tttcttgggt agtttgcagt tttaaaatta tgttttaaaa tggactatca tatgcttacc	4860
[4227]	gtaacttgaa agtatttgcg tttcttggct ttatatatct tgtggaaagg acgaaacacc	4920
[4228]	ggagaccacg gcaggtctca gtttttagtac tctggaaaca gaatctacta aaacaaggca	4980
[4229]	aaatgccgtg tttatctcgt caacttgttg gcgagatttt tgcggccgag atctaggaac	5040
[4230]	ccctagtgat ggagttggcc actccctctc tgcgcgtcgc ctgcctcact gaggccgggc	5100
[4231]	gaccaaaggt cgcggcagc ccgggctttg ccggggcggc ctgagtgagc gagcgagcgc	5160
[4232]	gcagctgcct gcaggcagct tggcactggc cgtcgtttta caacgtcgtg actgggaaaa	5220
[4233]	ccctggcggt acccaactta atcgcttgc agcacatccc ctttctgcca gctggcgtaa	5280
[4234]	tagcgaagag gccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg	5340
[4235]	gcgcctgatg cggatatttc tcttacgca tctgtcgggt atttcacacc gcatacgtca	5400
[4236]	aagcaaccat agtacgcgc ctgtagcggc gcattaagcg cggcggtgt ggtggttacg	5460
[4237]	cgcagcgtga ccgtacact tgccagcgcc ctacgcccc ctcctttcgc tttcttcctt	5520
[4238]	tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta	5580
[4239]	gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgattt gggatgatgt	5640

[4240]	tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg	5700
[4241]	ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat ctctgggctat	5760
[4242]	tccttttgatt tataagggat ttgtccgatt tgggcctatt ggttaaaaaa tgagctgatt	5820
[4243]	taacaaaaat ttaacgcgaa ttttaacaaa atattaacgt ttacaatttt atgggtgact	5880
[4244]	ctcagtacaa tctgctctga tgccgcatag ttaagccagc cccgacaccc gccaacaccc	5940
[4245]	gctgacgcgc cctgacgggc ttgtctgctc ccggcatccg cttacagaca agctgtgacc	6000
[4246]	gtctccggga gctgcatgtg tcagaggttt tcaccgtcat caccgaaacg cgcgagacga	6060
[4247]	aagggcctcg tgatacgctt atttttatag gttaatgtca tgataataat ggtttcttag	6120
[4248]	acgtcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa	6180
[4249]	atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat	6240
[4250]	tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg	6300
[4251]	gcattttgcc ttctgtttt tgctcaccca gaaacgttg tgaaagtaaa agatgctgaa	6360
[4252]	gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt	6420
[4253]	gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaaagt tctgctatgt	6480
[4254]	ggcgcggtat tatcccgat tgacgccggg caagagcaac tcggtcgccg catacactat	6540
[4255]	tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg	6600
[4256]	acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta	6660
[4257]	cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaa catgggggat	6720
[4258]	catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag	6780
[4259]	cgtgacacca cgatgcctgt agcaatggca acaacgttg gcaaactatt aactggcgaa	6840
[4260]	ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca	6900
[4261]	ggaccacttc tgcgctcggc ctttccggtt ggctggttta ttgctgataa atctggagcc	6960
[4262]	ggtgagcgtg gaagcccgcg tatcatgtca gcaactgggc cagatggtaa gccctcccg	7020
[4263]	atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc	7080
[4264]	gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat	7140
[4265]	atactttaga ttgatttaaa acttcatttt taattttaaa ggatctaggt gaagatcctt	7200
[4266]	tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac	7260
[4267]	cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgctt aatctgctgc	7320
[4268]	ttgcaaacia aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca	7380
[4269]	actctttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta	7440
[4270]	gtgtagccgt agttaggcca ccaattcaag aactctgtag caccgcctac atacctcgct	7500
[4271]	ctgctaatac tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg	7560
[4272]	gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc	7620
[4273]	acacagccca gcttgagcgc aacgacctac accgaactga gatacctaca gcgtgagcta	7680
[4274]	tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg	7740
[4275]	gtcggaaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt	7800
[4276]	cctgtcgggt ttgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg	7860
[4277]	cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttcttgcc cttttgctgg	7920
[4278]	ccttttgctc acatgttctt tctgctgta tccctgatt ctgtggataa ccgtattacc	7980
[4279]	gcctttgagt gagctgatac	8000
[4280]	<210>	213
[4281]	<211>	5

[4282]	<212>	PRT
[4283]	<213>	人工序列
[4284]	<220>	
[4285]	<223>	SGGGS连接子
[4286]	<400>	213
[4287]	Ser Gly Gly Gly Ser	
[4288]	1	5
[4289]	<210>	214
[4290]	<211>	1033
[4291]	<212>	PRT
[4292]	<213>	人工序列
[4293]	<220>	
[4294]	<223>	具有SGGGS连接子的氨基酸残基 (dSaCas9的第721位至第745位氨基酸残基) 缺失突变体
[4295]	<220>	
[4296]	<221>	变体
[4297]	<222>	(10) .. (10)
[4298]	<223>	Asp残基转化成Ala残基
[4299]	<220>	
[4300]	<221>	变体
[4301]	<222>	(580) .. (580)
[4302]	<223>	Asn残基转化成Ala残基
[4303]	<220>	
[4304]	<221>	位点
[4305]	<222>	(721) .. (725)
[4306]	<223>	SGGGS连接子
[4307]	<400>	214
[4308]	Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val	
[4309]	1	5 10 15
[4310]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly	
[4311]	20	25 30
[4312]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg	
[4313]	35	40 45
[4314]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile	
[4315]	50	55 60
[4316]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His	
[4317]	65	70 75 80
[4318]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu	
[4319]	85	90 95
[4320]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu	
[4321]	100	105 110
[4322]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr	

[4323]	115	120	125
[4324]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[4325]	130	135	140
[4326]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[4327]	145	150	155
[4328]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[4329]	165	170	175
[4330]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[4331]	180	185	190
[4332]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[4333]	195	200	205
[4334]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[4335]	210	215	220
[4336]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[4337]	225	230	235
[4338]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[4339]	245	250	255
[4340]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[4341]	260	265	270
[4342]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[4343]	275	280	285
[4344]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[4345]	290	295	300
[4346]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[4347]	305	310	315
[4348]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[4349]	325	330	335
[4350]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[4351]	340	345	350
[4352]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[4353]	355	360	365
[4354]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[4355]	370	375	380
[4356]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[4357]	385	390	395
[4358]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[4359]	405	410	415
[4360]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		
[4361]	420	425	430
[4362]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[4363]	435	440	445
[4364]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		

[4365]	450	455	460
[4366]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[4367]	465	470	475 480
[4368]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[4369]	485	490	495
[4370]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[4371]	500	505	510
[4372]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[4373]	515	520	525
[4374]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[4375]	530	535	540
[4376]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[4377]	545	550	555 560
[4378]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[4379]	565	570	575
[4380]	Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[4381]	580	585	590
[4382]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[4383]	595	600	605
[4384]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[4385]	610	615	620
[4386]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[4387]	625	630	635 640
[4388]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[4389]	645	650	655
[4390]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[4391]	660	665	670
[4392]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[4393]	675	680	685
[4394]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[4395]	690	695	700
[4396]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[4397]	705	710	715 720
[4398]	Ser Gly Gly Gly Ser Thr Glu Gln Glu Tyr Lys Glu Ile Phe Ile Thr		
[4399]	725	730	735
[4400]	Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr Lys Tyr Ser		
[4401]	740	745	750
[4402]	His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn Asp Thr Leu		
[4403]	755	760	765
[4404]	Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile Val Asn Asn		
[4405]	770	775	780
[4406]	Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys Lys Leu Ile		

[4407]	785	790	795	800
[4408]	Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp Pro Gln Thr			
[4409]		805	810	815
[4410]	Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp Glu Lys Asn			
[4411]		820	825	830
[4412]	Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu Thr Lys Tyr			
[4413]		835	840	845
[4414]	Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys Tyr Tyr Gly			
[4415]		850	855	860
[4416]	Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr Pro Asn Ser			
[4417]	865	870	875	880
[4418]	Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg Phe Asp Val			
[4419]		885	890	895
[4420]	Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys Asn Leu Asp			
[4421]		900	905	910
[4422]	Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys Cys Tyr Glu			
[4423]		915	920	925
[4424]	Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu Phe Ile Ala			
[4425]		930	935	940
[4426]	Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu Leu Tyr Arg			
[4427]	945	950	955	960
[4428]	Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu Val Asn Met			
[4429]		965	970	975
[4430]	Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn Asp Lys Arg			
[4431]		980	985	990
[4432]	Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr Gln Ser Ile Lys			
[4433]		995	1000	1005
[4434]	Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu Val Lys Ser			
[4435]		1010	1015	1020
[4436]	Lys Lys His Pro Gln Ile Ile Lys Lys Gly			
[4437]		1025	1030	
[4438]	<210> 215			
[4439]	<211> 21			
[4440]	<212> RNA			
[4441]	<213> 人工序列			
[4442]	<220>			
[4443]	<223> 对照sgRNA中包含的crRNA			
[4444]	<400> 215			
[4445]	acggaggcua agcgucgcaa g 21			
[4446]	<210> 216			
[4447]	<211> 519			
[4448]	<212> PRT			

[4449]	<213> 人工序列															
[4450]	<220>															
[4451]	<223> VPR															
[4452]	<400> 216															
[4453]	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu
[4454]	1				5					10					15	
[4455]	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe
[4456]				20					25					30		
[4457]	Asp	Leu	Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp
[4458]			35					40					45			
[4459]	Met	Leu	Ser	Ser	Gly	Ser	Pro	Lys	Lys	Lys	Arg	Lys	Val	Gly	Ser	Gln
[4460]		50					55					60				
[4461]	Tyr	Leu	Pro	Asp	Thr	Asp	Asp	Arg	His	Arg	Ile	Glu	Glu	Lys	Arg	Lys
[4462]	65					70					75				80	
[4463]	Arg	Thr	Tyr	Glu	Thr	Phe	Lys	Ser	Ile	Met	Lys	Lys	Ser	Pro	Phe	Ser
[4464]					85					90					95	
[4465]	Gly	Pro	Thr	Asp	Pro	Arg	Pro	Pro	Pro	Arg	Arg	Ile	Ala	Val	Pro	Ser
[4466]				100					105					110		
[4467]	Arg	Ser	Ser	Ala	Ser	Val	Pro	Lys	Pro	Ala	Pro	Gln	Pro	Tyr	Pro	Phe
[4468]			115					120					125			
[4469]	Thr	Ser	Ser	Leu	Ser	Thr	Ile	Asn	Tyr	Asp	Glu	Phe	Pro	Thr	Met	Val
[4470]		130					135					140				
[4471]	Phe	Pro	Ser	Gly	Gln	Ile	Ser	Gln	Ala	Ser	Ala	Leu	Ala	Pro	Ala	Pro
[4472]	145					150					155				160	
[4473]	Pro	Gln	Val	Leu	Pro	Gln	Ala	Pro	Ala	Pro	Ala	Pro	Ala	Pro	Ala	Met
[4474]				165					170						175	
[4475]	Val	Ser	Ala	Leu	Ala	Gln	Ala	Pro	Ala	Pro	Val	Pro	Val	Leu	Ala	Pro
[4476]			180						185					190		
[4477]	Gly	Pro	Pro	Gln	Ala	Val	Ala	Pro	Pro	Ala	Pro	Lys	Pro	Thr	Gln	Ala
[4478]			195						200					205		
[4479]	Gly	Glu	Gly	Thr	Leu	Ser	Glu	Ala	Leu	Leu	Gln	Leu	Gln	Phe	Asp	Asp
[4480]		210					215					220				
[4481]	Glu	Asp	Leu	Gly	Ala	Leu	Leu	Gly	Asn	Ser	Thr	Asp	Pro	Ala	Val	Phe
[4482]	225					230					235				240	
[4483]	Thr	Asp	Leu	Ala	Ser	Val	Asp	Asn	Ser	Glu	Phe	Gln	Gln	Leu	Leu	Asn
[4484]				245						250					255	
[4485]	Gln	Gly	Ile	Pro	Val	Ala	Pro	His	Thr	Thr	Glu	Pro	Met	Leu	Met	Glu
[4486]			260						265					270		
[4487]	Tyr	Pro	Glu	Ala	Ile	Thr	Arg	Leu	Val	Thr	Gly	Ala	Gln	Arg	Pro	Pro
[4488]			275					280					285			
[4489]	Asp	Pro	Ala	Pro	Ala	Pro	Leu	Gly	Ala	Pro	Gly	Leu	Pro	Asn	Gly	Leu
[4490]		290					295					300				

[4491]	Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser
[4492]	305 310 315 320
[4493]	Ala Leu Leu Gly Ser Gly Ser Gly Ser Arg Asp Ser Arg Glu Gly Met
[4494]	325 330 335
[4495]	Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser Asp Val Phe Glu
[4496]	340 345 350
[4497]	Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro Phe His Pro Pro
[4498]	355 360 365
[4499]	Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr
[4500]	370 375 380
[4501]	Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu Thr Pro Ala Pro
[4502]	385 390 395 400
[4503]	Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr Pro Glu Ala Ser
[4504]	405 410 415
[4505]	His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala Val Lys Ala
[4506]	420 425 430
[4507]	Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu Glu Ala Ala
[4508]	435 440 445
[4509]	Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg Gly His Leu
[4510]	450 455 460
[4511]	Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu
[4512]	465 470 475 480
[4513]	Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu
[4514]	485 490 495
[4515]	Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly Leu Ser
[4516]	500 505 510
[4517]	Ile Phe Asp Thr Ser Leu Phe
[4518]	515

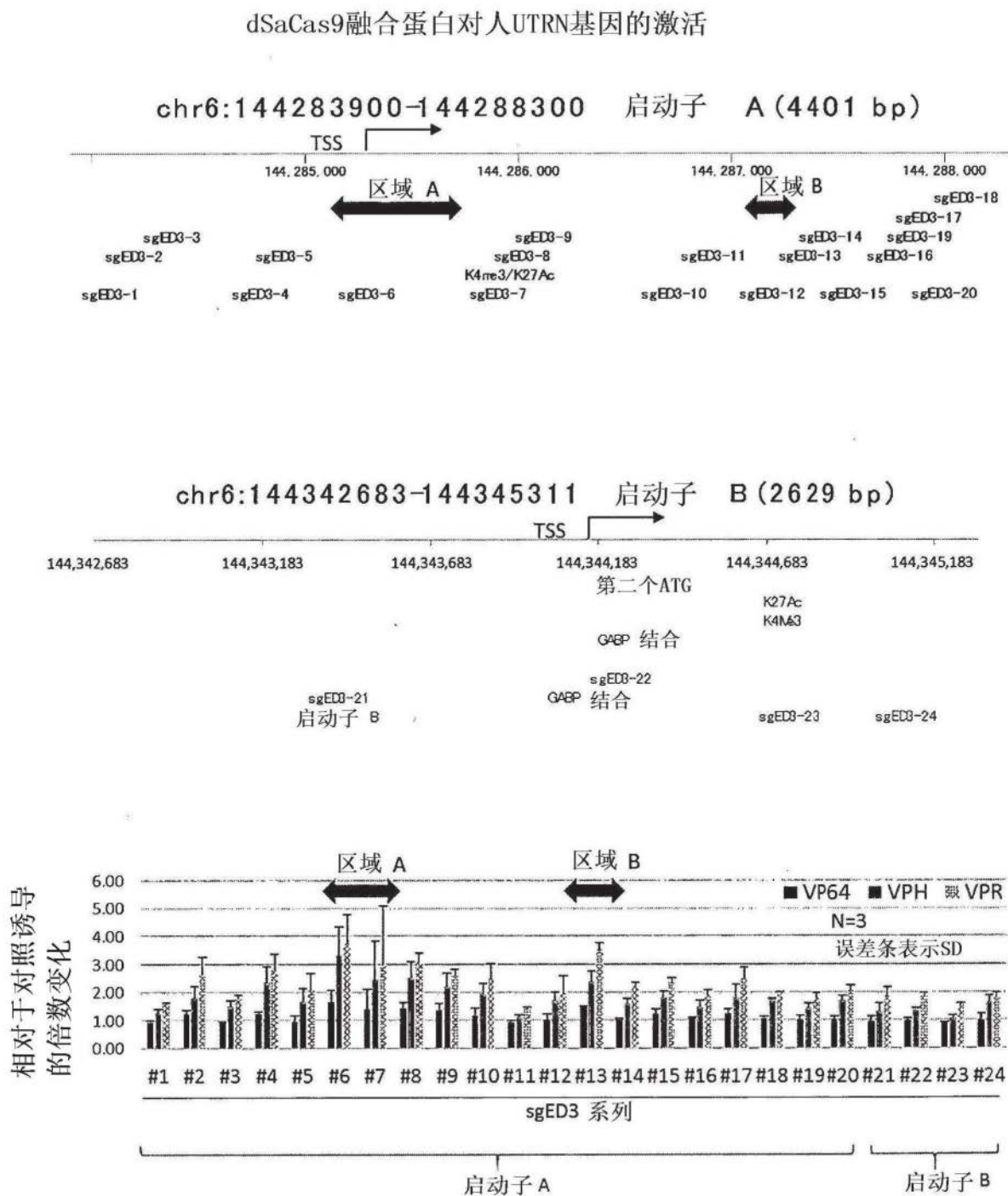


图1

dSaCas9-VPR融合蛋白对人UTRN基因的激活

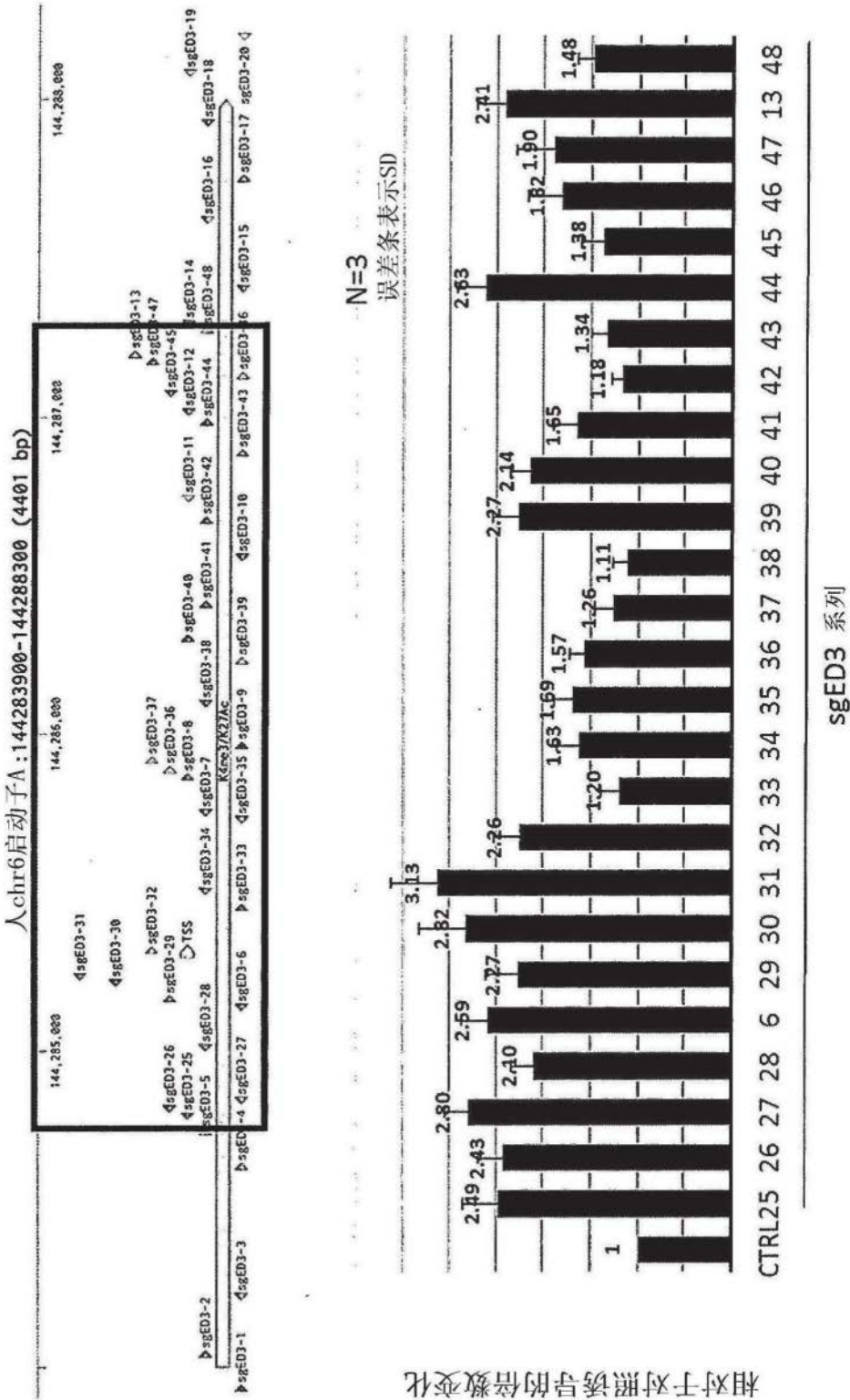


图2

克隆到pAAV-dSaCas9[-25]-miniVR质粒中的sgRNA的验证

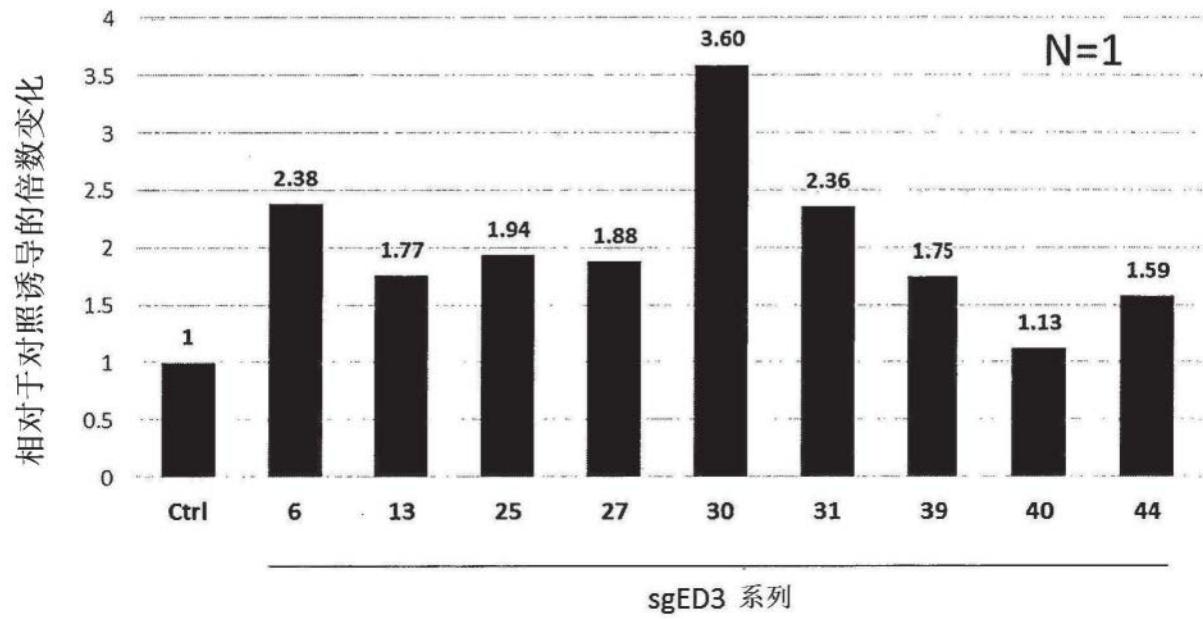


图3

表达dSaCas9[-25]-miniVR和sgRNA的AAV2对UTRN基因的激活

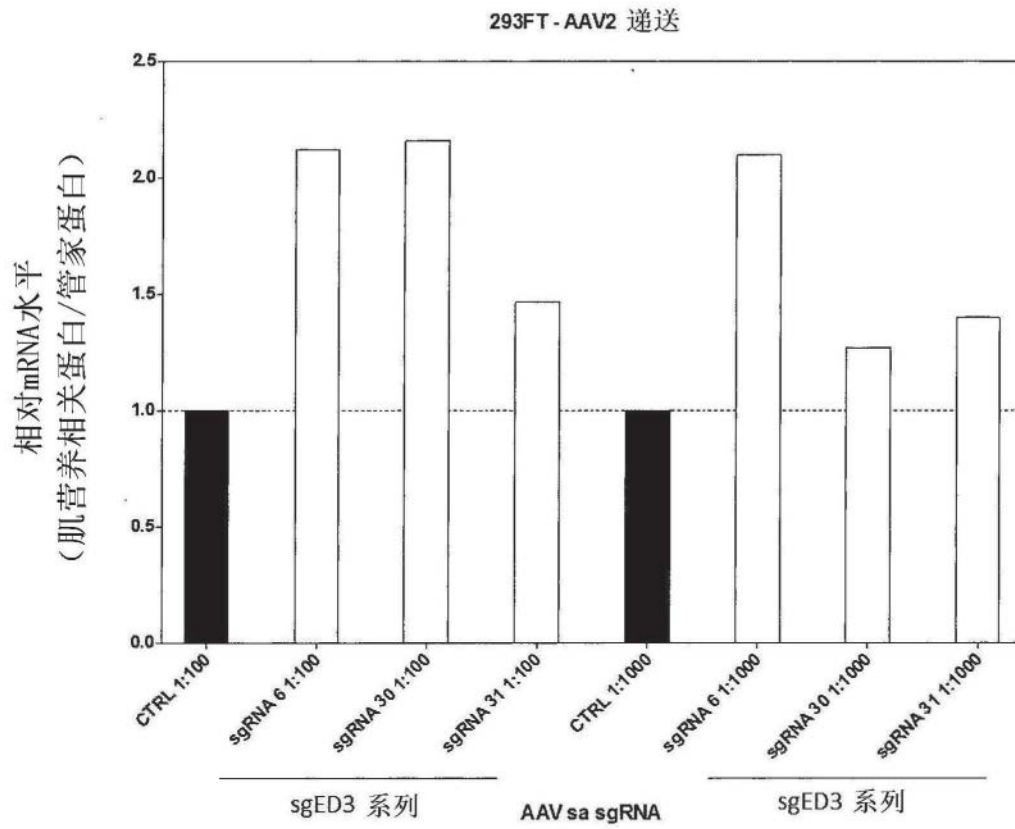


图4



图5

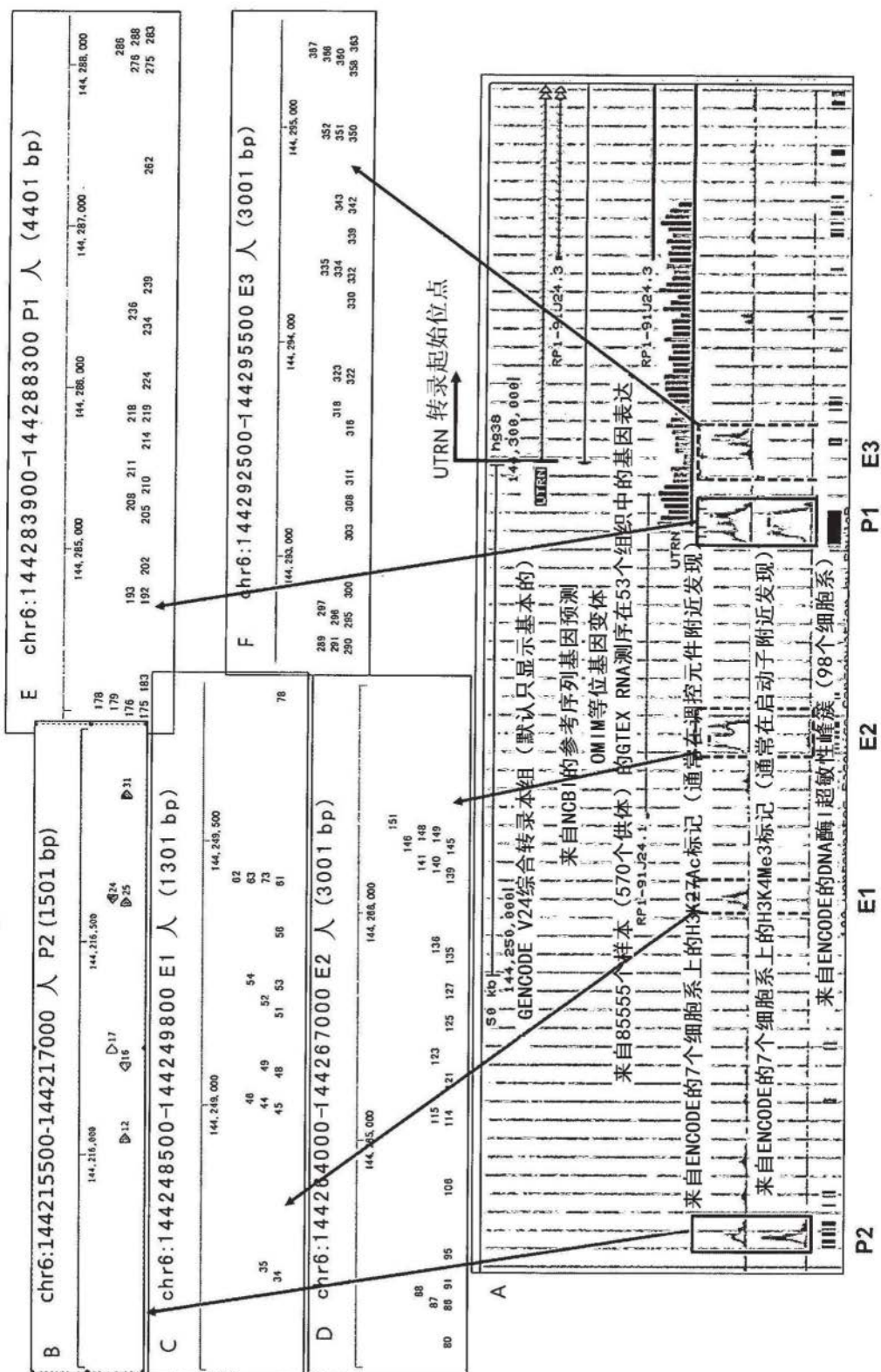


图6

单个sgRNA及其组合的实验设置

5种靶向序列

向导编号	食蟹猴匹配	靶标
145	是	增强子 (E2)
146	是	增强子 (E2)
205	否	启动子 (P1)
208	是	启动子 (P1)
210	否	启动子 (P1)

组合

组合	食蟹猴匹配	靶标
205+145	否	增强子 (E2) +启动子 (P1)
208+145	是	
210+145	否	
205+146	否	
208+146	是	
210+146	否	

图8

在5个HSMM供体中的UTRN上调-单个sgRNA对比其组合

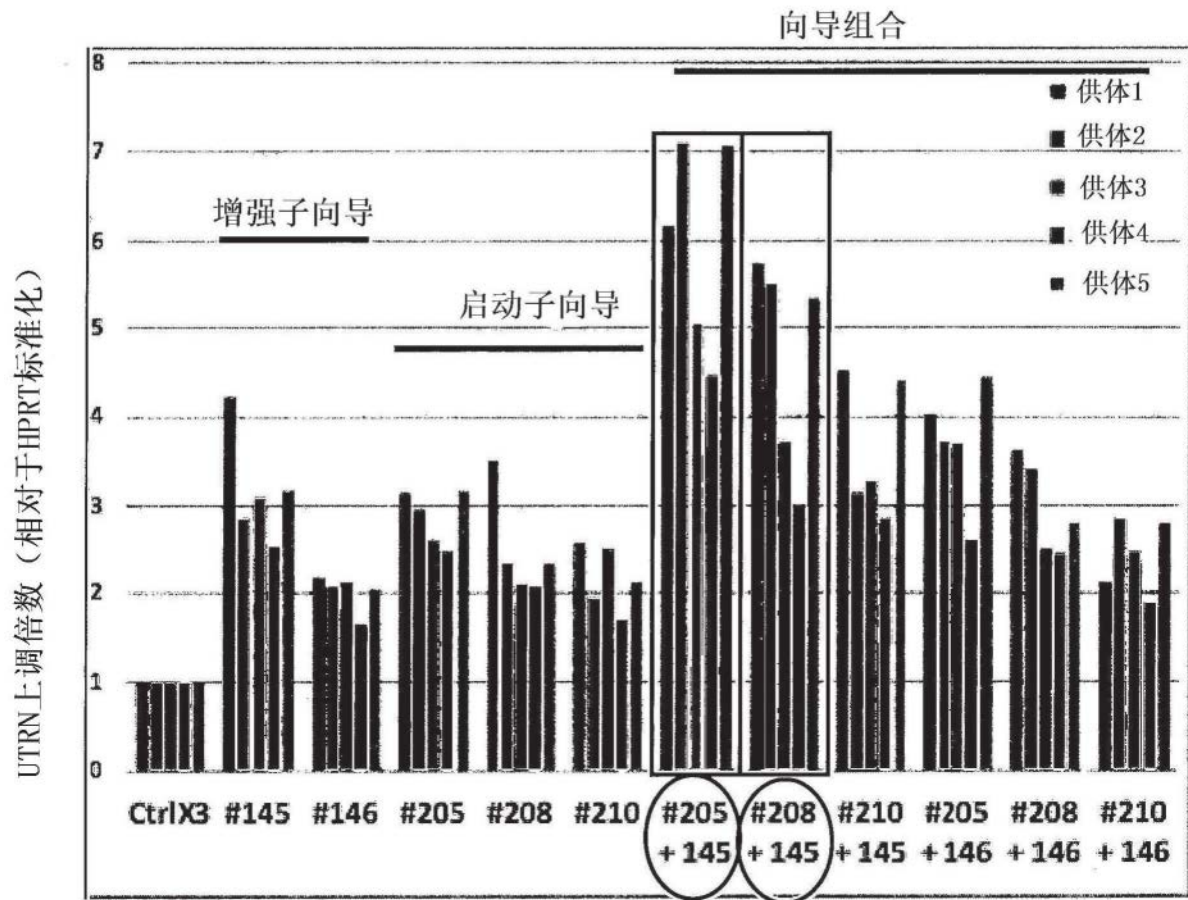


图9

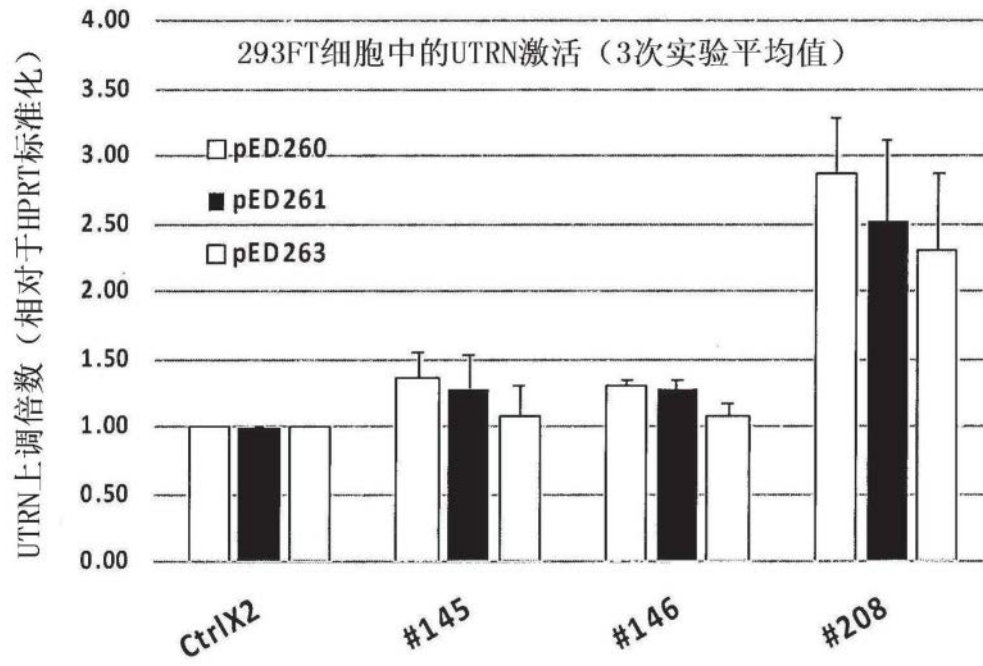


图10

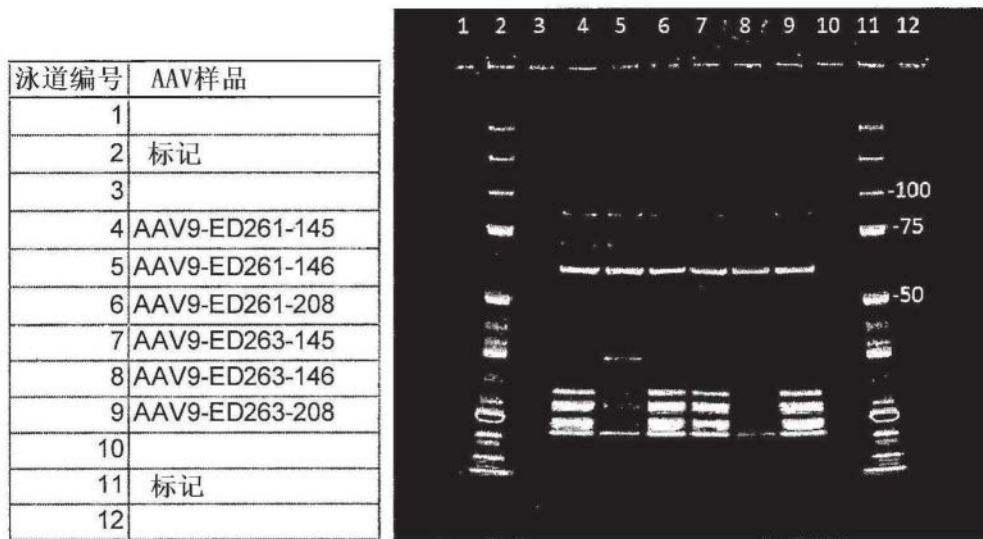


图11

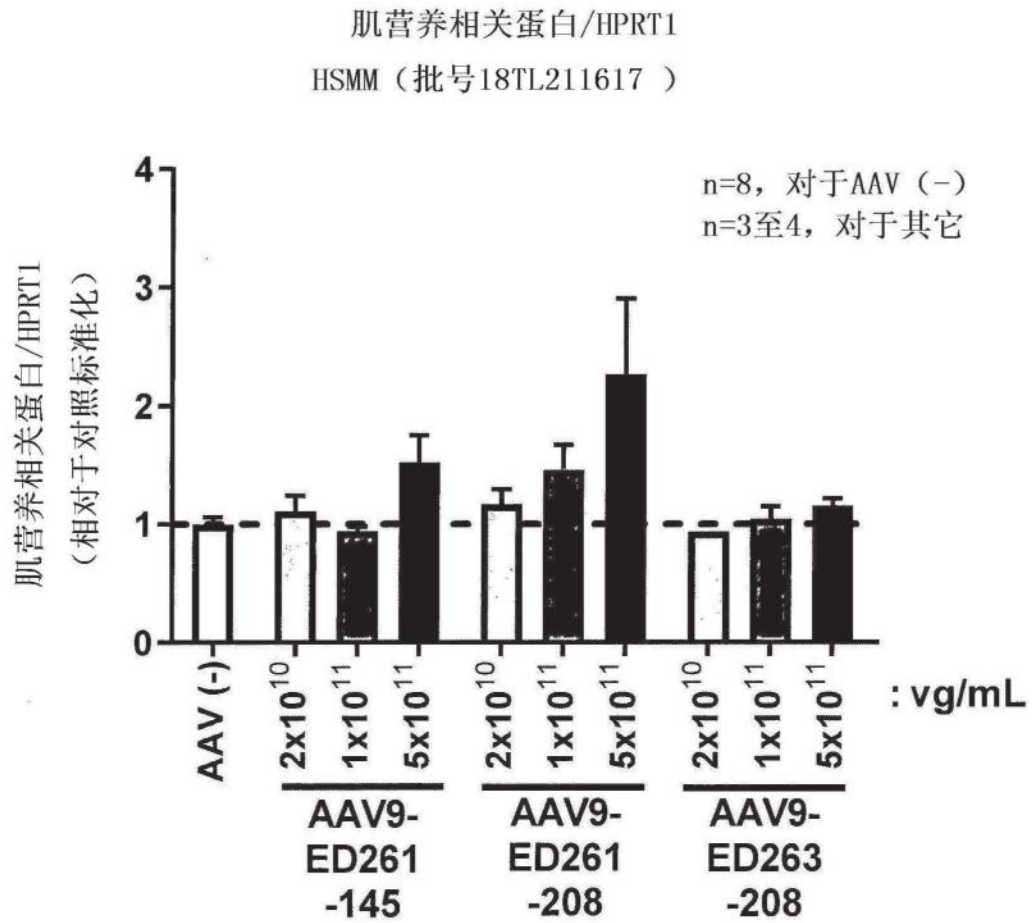


图12

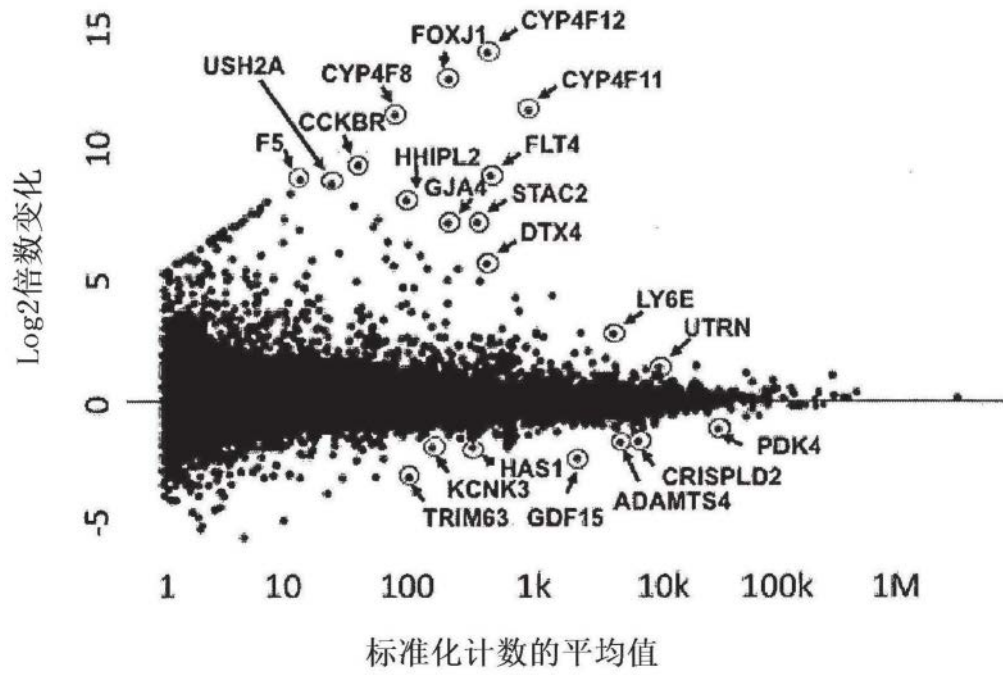


图13A

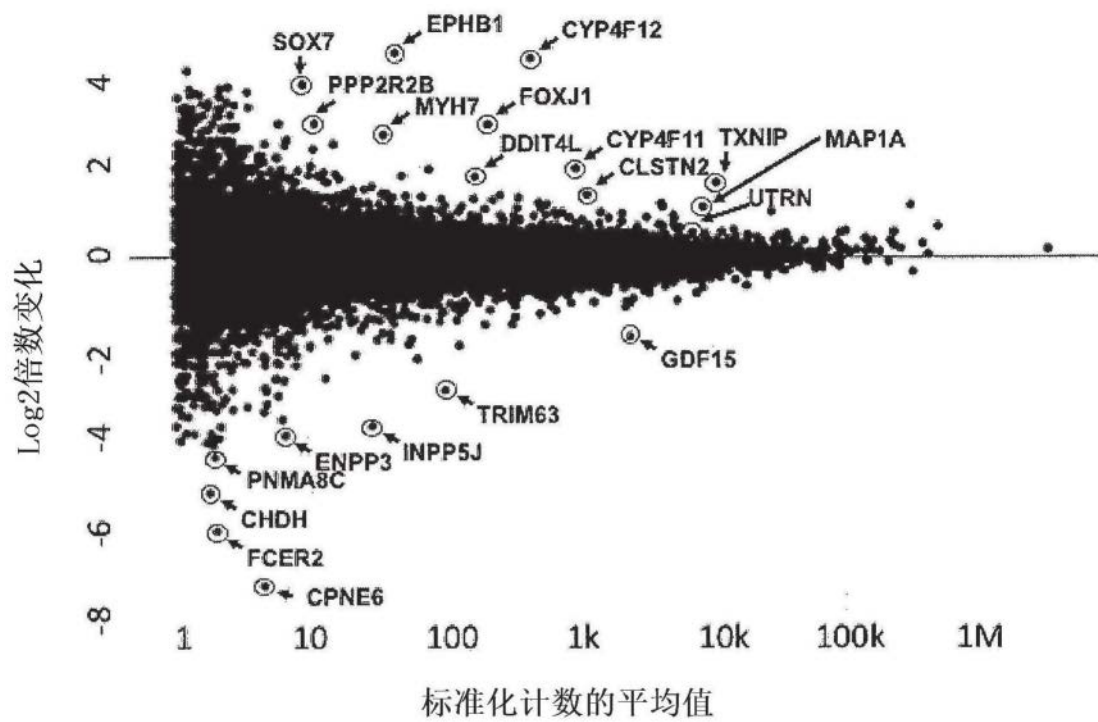


图13B

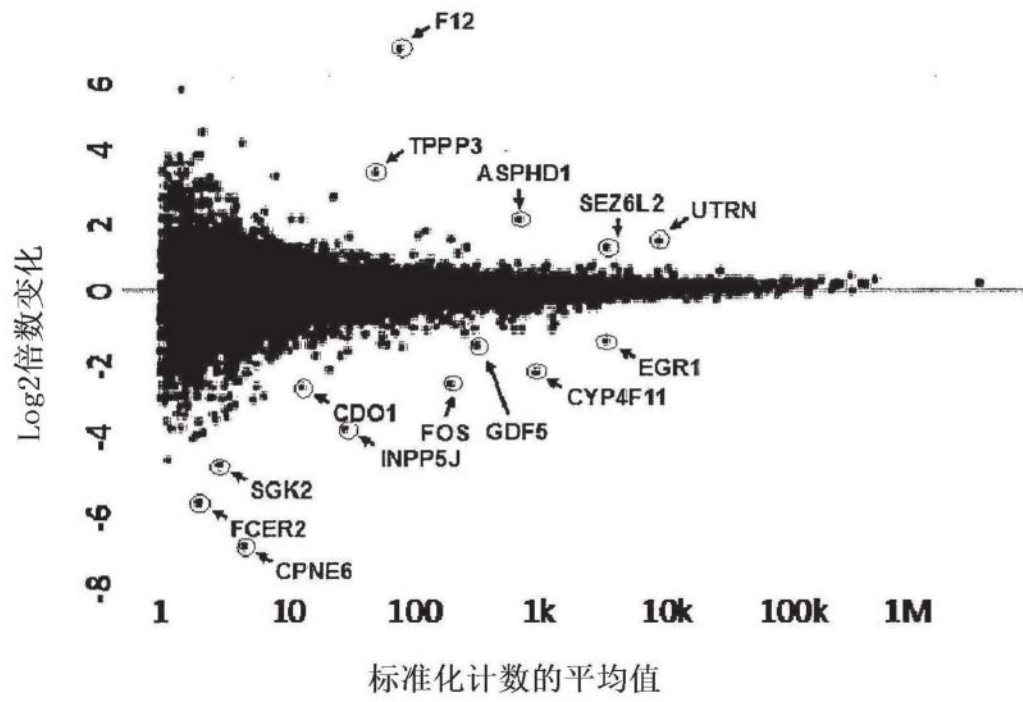


图13C