



(12) 发明专利

(10) 授权公告号 CN 114875012 B

(45) 授权公告日 2025. 02. 18

(21) 申请号 202210550934.7

V·派特塔那雅克

(22) 申请日 2016.08.26

(74) 专利代理机构 北京林达刘知识产权代理事

(65) 同一申请的已公布的文献号

务所(普通合伙) 11277

申请公布号 CN 114875012 A

专利代理师 刘新宇 李茂家

(43) 申请公布日 2022.08.09

(51) Int.Cl.

(30) 优先权数据

C12N 9/22 (2006.01)

62/211,553 2015.08.28 US

C12N 15/55 (2006.01)

62/216,033 2015.09.09 US

C12N 15/113 (2010.01)

62/258,280 2015.11.20 US

62/271,938 2015.12.28 US

15/015,947 2016.02.04 US

(62) 分案原申请数据

201680063266.8 2016.08.26

(73) 专利权人 通用医疗公司

(56) 对比文件

NISHIMASU,H.等.Crystal structure of Staphylococcus aureus Cas9.《Cell》.2015,第162卷(第5期),第1117-1118页,图3-5, SUPPLEMENTAL EXPERIMENTAL PROCEDURES第5页.

地址 美国马萨诸塞州

审查员 张丽玮

(72) 发明人 J·K·乔昂格 B·克莱因斯蒂弗

权利要求书2页 说明书74页

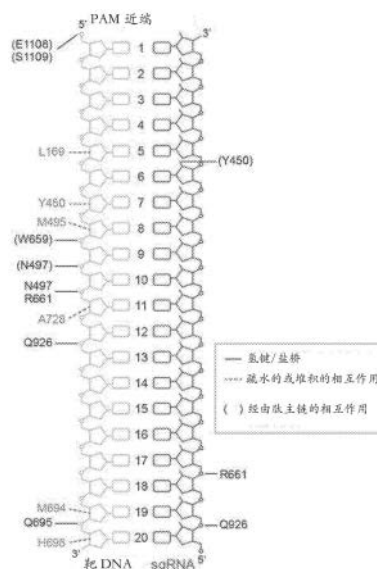
序列表88页 附图29页

(54) 发明名称

工程化的CRISPR-Cas9核酸酶

(57) 摘要

本发明涉及工程化的CRISPR-Cas9核酸酶。披露了具有改进的特异性的工程化CRISPR-Cas9核酸酶及其在基因组工程、表观基因组工程、基因组靶向、和基因组编辑中的用途。



1. 一种分离的金黄色葡萄球菌Cas9 (SaCas9) 蛋白, 该蛋白的氨基酸序列为在SEQ ID NO:2所示的氨基酸序列中具有以下突变中的任一个: Y211A、Y212A、W229A、Y230A、R245A、T392A和/或Y651A, 并且该蛋白任选地包含核定位序列、细胞穿透肽序列、和/或亲和标签中的一种或多种。

2. 如权利要求1所述的分离的蛋白, 其进一步包含以下突变中的一个或多个: N419A、L446A、R654A、R499A/Q500A和/或R499A/Q500A/R654A/G655A。

3. 如权利要求1或权利要求2所述的分离的蛋白, 其包含以下突变集中的一个: Y211A/R654A、Y211A/Y212A、Y211A/Y230A、Y211A/R245A、Y212A/Y230A、Y212A/R245A、Y230A/R245A、W229A/Y230A、W229A/R654A、Y211A/Y212A/Y230A、Y211A/Y212A/R245A、Y211A/Y212A/Y651A、Y211A/Y230A/R245A、Y211A/Y230A/Y651A、Y211A/R245A/Y651A、Y211A/R245A/R654A、Y211A/R245A/N419A、Y211A/N419A/R654A、Y212A/Y230A/R245A、Y212A/Y230A/Y651A、Y212A/R245A/Y651A、Y230A/R245A/Y651A、R245A/N419A/R654A、T392A/N419A/R654A、R245A/T392A/N419A/R654A、Y211A/R245A/N419A/R654A、W229A/R245A/N419A/R654A、Y211A/R245A/T392A/N419A/R654A、或Y211A/W229A/R245A/N419A/R654A。

4. 如权利要求1所述的分离的蛋白, 其进一步包含降低核酸酶活性的选自以下的一个或多个突变:

(i) D10A或D10N, 和/或

(ii) H557A、H557N、或H557Y, 和/或

(iii) N580A, 和/或

(iv) D556A。

5. 一种融合蛋白, 其包含使用任选的介入接头与异源功能结构域融合的、如权利要求1-4任一项中所述的分离的蛋白, 其中该接头不干扰该融合蛋白的活性。

6. 如权利要求5所述的融合蛋白, 其中该异源功能结构域是转录激活结构域。

7. 如权利要求6所述的融合蛋白, 其中该转录激活结构域来自VP64或NF- κ B p65。

8. 如权利要求5所述的融合蛋白, 其中该异源功能结构域是转录沉默子或转录阻遏结构域。

9. 如权利要求8所述的融合蛋白, 其中该转录阻遏结构域是克鲁贝尔相关盒 (KRAB) 结构域、ERF阻遏因子结构域 (ERD) 或mSin3A相互作用结构域 (SID)。

10. 如权利要求8所述的融合蛋白, 其中该转录沉默子是异染色质蛋白1 (HP1)。

11. 如权利要求10所述的融合蛋白, 其中该转录沉默子是HP1 α 或HP1 β 。

12. 如权利要求5所述的融合蛋白, 其中该异源功能结构域是修饰DNA的甲基化状态的酶。

13. 如权利要求12所述的融合蛋白, 其中该修饰DNA的甲基化状态的酶是DNA甲基转移酶 (DNMT) 或TET蛋白。

14. 如权利要求13所述的融合蛋白, 其中该TET蛋白是TET1。

15. 如权利要求5所述的融合蛋白, 其中该异源功能结构域是修饰组蛋白亚基的酶。

16. 如权利要求15所述的融合蛋白, 其中该修饰组蛋白亚基的酶是组蛋白乙酰转移酶 (HAT)、组蛋白脱乙酰酶 (HDAC)、组蛋白甲基转移酶 (HMT) 或组蛋白脱甲基酶。

17. 如权利要求5所述的融合蛋白, 其中该异源功能结构域是生物系链。

18. 如权利要求17所述的融合蛋白,其中该生物系链是MS2、Csy4或 λ N蛋白。
19. 如权利要求17所述的融合蛋白,其中该异源功能结构域是FokI。
20. 一种编码如权利要求1-4任一项所述的蛋白的分离的核酸。
21. 一种包含如权利要求20所述的分离的核酸的载体,该分离的核酸任选地有效地连接到一个或多个调节结构域,以表达如权利要求1-4任一项所述的蛋白。
22. 一种宿主细胞,该宿主细胞包含如权利要求20所述的核酸。
23. 如权利要求22所述的宿主细胞,其中该细胞为哺乳动物宿主细胞。
24. 根据权利要求22所述的宿主细胞,其中该细胞表达如权利要求1-4任一项所述的蛋白。
25. 一种改变细胞的基因组或表观基因组的离体或体外方法,该方法包括在该细胞中表达如权利要求1-4任一项所述的分离的蛋白或表达如权利要求5-19任一项所述的融合蛋白,以及具有与该细胞的基因组的所选部分互补的区域的指导RNA,或使该细胞与所述分离的蛋白或与所述融合蛋白以及所述指导RNA接触。
26. 如权利要求25所述的方法,其中该分离的蛋白或融合蛋白包含核定位序列、细胞穿透肽序列和/或亲和标签中的一种或多种。
27. 如权利要求25所述的方法,其中该细胞是干细胞。
28. 如权利要求27所述的方法,其中该干细胞是间充质干细胞或诱导多能干细胞。
29. 一种改变双链DNA(dsDNA)分子的离体或体外方法,该方法包括使该dsDNA分子与如权利要求1-4任一项所述的分离的蛋白或与权利要求5-19任一项所述的融合蛋白,以及具有与该dsDNA分子的所选部分互补的区域的指导RNA接触。
30. 如权利要求1-4中任一项所述的分离的蛋白或者如权利要求5-19中任一项所述的分离的融合蛋白在改变双链DNA(dsDNA)分子的方法中的用途,所述方法包括使该dsDNA分子与所述分离的蛋白或与所述融合蛋白,以及具有与该dsDNA分子的所选部分互补的区域的指导RNA接触。

工程化的CRISPR-Cas9核酸酶

[0001] 本申请是中国专利申请201680063266.8的分案申请,原申请201680063266.8的申请日为2016年08月26日,其名称为“工程化的CRISPR-Cas9核酸酶”。

[0002] 优先权声明

[0003] 本申请于35 USC§119(e)下要求以下美国专利申请序列号的优先权:62/211,553, 2015年8月28日提交;62/216,033, 2015年9月9日提交;62/258,280, 2015年11月20日提交; 62/271,938, 2015年12月28日提交;以及15/015,947, 2016年2月4日提交。以上文献的全部内容通过引用结合在此。

[0004] 序列表

[0005] 本申请包含一份已经以ASCII格式电子递交的序列表并且该序列表通过引用以其全文结合在此。创建于2016年8月26日的所述ASCII副本名称为SEQUENCE LISTING.txt并且大小是129,955字节。

[0006] 联邦资助的研究或开发

[0007] 本发明是根据由国立卫生研究院(National Institutes of Health)授予的授权号DP1 GM105378和R01 GM088040在政府支持下完成的。美国政府享有本发明的一定的权利。

技术领域

[0008] 本发明至少部分地涉及具有改变和改进的靶特异性的工程化规律间隔成簇短回文重复序列(CRISPR)/CRISPR相关蛋白9(Cas9)核酸酶及其在基因组工程、表观基因组工程、基因组靶向、基因组编辑、以及体外诊断中的用途。

背景技术

[0009] CRISPR-Cas9核酸酶使得在多种多样的生物体和细胞类型中进行有效的基因组编辑成为可能(Sander&Joung, Nat Biotechnol[自然生物技术]32,347-355(2014);Hsu等人, Cell[细胞]157,1262-1278(2014);Doudna&Charpentier, Science[科学]346,1258096(2014);Barrangou&May, Expert Opin Biol Ther[生物疗法的专家意见]15,311-314(2015))。Cas9的靶位点识别是由编码与靶原型间隔子互补的序列的嵌合单指导RNA(sgRNA)编程(Jinek等人, Science[科学]337,816-821(2012)),而且需要进行短的邻近PAM的识别(Mojica等人, Microbiology[微生物学]155,733-740(2009);Shah等人, RNA Biol[RNA生物学]10,891-899(2013);Jiang等人, Nat Biotechnol[自然生物技术]31,233-239(2013);Jinek等人, Science[科学]337,816-821(2012);Sternberg等人, Nature[自然]507,62-67(2014))。

发明内容

[0010] 如本文所述,理论上可以通过减小Cas9对DNA的结合亲和力,将Cas9蛋白工程化以显示增加的特异性。因此,本文描述了与野生型蛋白相比,具有增加的特异性(即在不完全

匹配的或错配的DNA位点,诱导显著更少的脱靶效应)的多个Cas9变体,连同使用它们的方法。

[0011] 在第一方面,本发明提供了分离的酿脓链球菌Cas9 (SpCas9) 蛋白,该蛋白具有在以下位置中的一个、两个、三个、四个、五个、六个、或全部七个处的突变:L169A、Y450、N497、R661、Q695、Q926、和/或D1135E,例如包含与SEQ ID NO:1的氨基酸序列至少80%一致的序列,所述序列具有在以下位置中的一个、两个、三个、四个、五个、六个、或七个处的突变:L169、Y450、N497、R661、Q695、Q926、D1135E,以及任选地核定位序列、细胞穿透肽序列、和/或亲和标签中的一种或多种。突变将氨基酸改变为不同于天然氨基酸的氨基酸(例如497是除了N之外的任何氨基酸)。在优选的实施例中,突变将氨基酸改变为不同于天然氨基酸(精氨酸或赖氨酸)的任何氨基酸;在一些实施例中,氨基酸是丙氨酸。

[0012] 在一些实施例中,变体SpCas9蛋白包含在以下位置中的一个、两个、三个、或全部四个处的突变:N497、R661、Q695、和Q926,例如以下突变中的一个、两个、三个、或全部四个:N497A、R661A、Q695A、和Q926A。

[0013] 在一些实施中,变体SpCas9蛋白包含在Q695和/或Q926处的突变,以及任选地L169、Y450、N497、R661和D1135E中的一个、两个、三个、四个或全部五个,例如包含但不限于:Y450A/Q695A、L169A/Q695A、Q695A/Q926A、Q695A/D1135E、Q926A/D1135E、Y450A/D1135E、L169A/Y450A/Q695A、L169A/Q695A/Q926A、Y450A/Q695A/Q926A、R661A/Q695A/Q926A、N497A/Q695A/Q926A、Y450A/Q695A/D1135E、Y450A/Q926A/D1135E、Q695A/Q926A/D1135E、L169A/Y450A/Q695A/Q926A、L169A/R661A/Q695A/Q926A、Y450A/R661A/Q695A/Q926A、N497A/Q695A/Q926A/D1135E、R661A/Q695A/Q926A/D1135E、和Y450A/Q695A/Q926A/D1135E。

[0014] 在一些实施例中,变体SpCas9蛋白包含在以下处的突变:N14;S15;S55;R63;R78;H160;K163;R165;L169;R403;N407;Y450;M495;N497;K510;Y515;W659;R661;M694;Q695;H698;A728;S730;K775;S777;R778;R780;K782;R783;K789;K797;Q805;N808;K810;R832;Q844;S845;K848;S851;K855;R859;K862;K890;Q920;Q926;K961;S964;K968;K974;R976;N980;H982;K1003;K1014;S1040;N1041;N1044;K1047;K1059;R1060;K1107;E1108;S1109;K1113;R1114;S1116;K1118;D1135;S1136;K1153;K1155;K1158;K1200;Q1221;H1241;Q1254;Q1256;K1289;K1296;K1297;R1298;K1300;H1311;K1325;K1334;T1337和/或S1216。

[0015] 在一些实施例中,变体SpCas9蛋白还包含以下突变中的一个或多个:N14A;S15A;S55A;R63A;R78A;R165A;R403A;N407A;N497A;Y450A;K510A;Y515A;R661A;Q695A;S730A;K775A;S777A;R778A;R780A;K782A;R783A;K789A;K797A;Q805A;N808A;K810A;R832A;Q844A;S845A;K848A;S851A;K855A;R859A;K862A;K890A;Q920A;Q926A;K961A;S964A;K968A;K974A;R976A;N980A;H982A;K1003A;K1014A;S1040A;N1041A;N1044A;K1047A;K1059A;R1060A;K1107A;E1108A;S1109A;K1113A;R1114A;S1116A;K1118A;D1135A;S1136A;K1153A;K1155A;K1158A;K1200A;Q1221A;H1241A;Q1254A;Q1256A;K1289A;K1296A;K1297A;R1298A;K1300A;H1311A;K1325A;K1334A;T1337A和/或S1216A。在一些实施例中,变体蛋白包含HF1 (N497A/R661A/Q695A/Q926A)+K810A、HF1+K848A、HF1+K855A、HF1+H982A、HF1+K848A/K1003A、HF1+K848A/R1060A、HF1+K855A/K1003A、HF1+K855A/R1060A、HF1+H982A/K1003A、HF1+H982A/R1060A、HF1+K1003A/R1060A、HF1+K810A/K1003A/R1060A、HF1+K848A/

K1003A/R1060A。在一些实施例中,变体蛋白包含HF1+K848A/K1003A、HF1+K848A/R1060A、HF1+K855A/K1003A、HF1+K855A/R1060A、HF1+K1003A/R1060A、HF1+K848A/K1003A/R1060A。在一些实施例中,变体蛋白包含Q695A/Q926A/R780A、Q695A/Q926A/R976A、Q695A/Q926A/H982A、Q695A/Q926A/K855A、Q695A/Q926A/K848A/K1003A、Q695A/Q926A/K848A/K855A、Q695A/Q926A/K848A/H982A、Q695A/Q926A/K1003A/R1060A、Q695A/Q926A/K848A/R1060A、Q695A/Q926A/K855A/H982A、Q695A/Q926A/K855A/K1003A、Q695A/Q926A/K855A/R1060A、Q695A/Q926A/H982A/K1003A、Q695A/Q926A/H982A/R1060A、Q695A/Q926A/K1003A/R1060A、Q695A/Q926A/K810A/K1003A/R1060A、Q695A/Q926A/K848A/K1003A/R1060A。在一些实施例中,变体包含N497A/R661A/Q695A/Q926A/K810A、N497A/R661A/Q695A/Q926A/K848A、N497A/R661A/Q695A/Q926A/K855A、N497A/R661A/Q695A/Q926A/R780A、N497A/R661A/Q695A/Q926A/K968A、N497A/R661A/Q695A/Q926A/H982A、N497A/R661A/Q695A/Q926A/K1003A、N497A/R661A/Q695A/Q926A/K1014A、N497A/R661A/Q695A/Q926A/K1047A、N497A/R661A/Q695A/Q926A/R1060A、N497A/R661A/Q695A/Q926A/K810A/K968A、N497A/R661A/Q695A/Q926A/K810A/K848A、N497A/R661A/Q695A/Q926A/K810A/K1003A、N497A/R661A/Q695A/Q926A/K810A/R1060A、N497A/R661A/Q695A/Q926A/K848A/K1003A、N497A/R661A/Q695A/Q926A/K848A/R1060A、N497A/R661A/Q695A/Q926A/K855A/K1003A、N497A/R661A/Q695A/Q926A/K855A/R1060A、N497A/R661A/Q695A/Q926A/K968A/K1003A、N497A/R661A/Q695A/Q926A/H982A/K1003A、N497A/R661A/Q695A/Q926A/H982A/R1060A、N497A/R661A/Q695A/Q926A/K1003A/R1060A、N497A/R661A/Q695A/Q926A/K810A/K1003A/R1060A、N497A/R661A/Q695A/Q926A/K848A/K1003A/R1060A、Q695A/Q926A/R780A、Q695A/Q926A/K810A、Q695A/Q926A/R832A、Q695A/Q926A/K848A、Q695A/Q926A/K855A、Q695A/Q926A/K968A、Q695A/Q926A/R976A、Q695A/Q926A/H982A、Q695A/Q926A/K1003A、Q695A/Q926A/K1014A、Q695A/Q926A/K1047A、Q695A/Q926A/R1060A、Q695A/Q926A/K848A/K968A、Q695A/Q926A/R976A、Q695A/Q926A/H982A、Q695A/Q926A/K855A、Q695A/Q926A/K848A/K1003A、Q695A/Q926A/K848A/K855A、Q695A/Q926A/K848A/H982A、Q695A/Q926A/K1003A/R1060A、Q695A/Q926A/R832A/R1060A、Q695A/Q926A/K968A/K1003A、Q695A/Q926A/K968A/R1060A、Q695A/Q926A/K848A/R1060A、Q695A/Q926A/K855A/H982A、Q695A/Q926A/K855A/K1003A、Q695A/Q926A/K855A/R1060A、Q695A/Q926A/H982A/K1003A、Q695A/Q926A/H982A/R1060A、Q695A/Q926A/K1003A/R1060A、Q695A/Q926A/K810A/K1003A/R1060A、Q695A/Q926A/K1003A/K1047A/R1060A、Q695A/Q926A/K968A/K1003A/R1060A、Q695A/Q926A/R832A/K1003A/R1060A、或Q695A/Q926A/K848A/K1003A/R1060A。

[0016] 还包含除了丙氨酸之外的氨基酸的突变,并且可以在本发明的方法和组合物中制造和应用。

[0017] 在一些实施例中,变体SpCas9蛋白包含以下另外的突变中的一个或多个:R63A、R66A、R69A、R70A、R71A、Y72A、R74A、R75A、K76A、N77A、R78A、R115A、H160A、K163A、R165A、L169A、R403A、T404A、F405A、N407A、R447A、N497A、I448A、Y450A、S460A、M495A、K510A、Y515A、R661A、M694A、Q695A、H698A、Y1013A、V1015A、R1122A、K1123A、K1124A、K1158A、K1185A、K1200A、S1216A、Q1221A、K1289A、R1298A、K1300A、K1325A、R1333A、K1334A、R1335A、和T1337A。

[0018] 在一些实施例中,变体SpCas9蛋白包含多重取代突变:N497/R661/Q695/Q926(四重变体突变体);Q695/Q926(二重突变体);R661/Q695/Q926和N497/Q695/Q926(三重突变体)。在一些实施例中,可以将在L169、Y450和/或D1135处的另外的取代突变添加至这些二重、三重、和四重突变体上,或添加至携带Q695或Q926处的突变的单重突变体上。在一些实施例中,突变体具有代替野生型氨基酸的丙氨酸。在一些实施例中,突变体具有除了精氨酸或赖氨酸(或天然氨基酸)之外的任何氨基酸。

[0019] 在一些实施例中,变体SpCas9蛋白还包含降低核酸酶活性的选自下组的一个或多个突变,该组由以下组成:在D10、E762、D839、H983、或D986处的突变;以及在H840或N863处的突变。在一些实施例中,突变是:(i)D10A或D10N以及(ii)H840A、H840N、或H840Y。

[0020] 在一些实施例中,SpCas9变体还可以包含以下突变组中的一个:D1135V/R1335Q/T1337R(VQR变体);D1135E/R1335Q/T1337R(EQR变体);D1135V/G1218R/R1335Q/T1337R(VRQR变体);或D1135V/G1218R/R1335E/T1337R(VRER变体)。

[0021] 本文还提供了分离的金黄色葡萄球菌Cas9(SaCas9)蛋白,该蛋白具有在以下位置中的一个、两个、三个、四个、五个、六个、或更多个位置处的突变:Y211、Y212、W229、Y230、R245、T392、N419、Y651、或R654,例如包含与SEQ ID NO:1的氨基酸序列至少80%一致的序列,所述序列具有在以下位置中的一个、两个、三个、四个、五个、或六个位置处的突变:Y211、Y212、W229、Y230、R245、T392、N419、Y651、或R654,并且任选地包含核定位序列、细胞穿透肽序列、和/或亲和标签中的一种或多种。在一些实施例中,本文描述的SaCas9变体包括SEQ ID NO:2的氨基酸序列,所述序列具有在以下位置中的一个、两个、三个、四个、五个、六个、或更多个位置处的突变:Y211、Y212、W229、Y230、R245、T392、N419、Y651和/或R654。在一些实施例中,变体包括以下突变中的一个或多个:Y211A、Y212A、W229、Y230A、R245A、T392A、N419A、Y651、和/或R654A。

[0022] 在一些实施例中,变体SaCas9蛋白包含在N419和/或R654处的突变,以及任选地以下另外的突变中的一个、两个、三个、四个或更多个:Y211、Y212、W229、Y230、R245和T392,优选N419A/R654A、Y211A/R654A、Y211A/Y212A、Y211A/Y230A、Y211A/R245A、Y212A/Y230A、Y212A/R245A、Y230A/R245A、W229A/R654A、Y211A/Y212A/Y230A、Y211A/Y212A/R245A、Y211A/Y212A/Y651A、Y211A/Y230A/R245A、Y211A/Y230A/Y651A、Y211A/R245A/Y651A、Y211A/R245A/R654A、Y211A/R245A/N419A、Y211A/N419A/R654A、Y212A/Y230A/R245A、Y212A/Y230A/Y651A、Y212A/R245A/Y651A、Y230A/R245A/Y651A、R245A/N419A/R654A、T392A/N419A/R654A、R245A/T392A/N419A/R654A、Y211A/R245A/N419A/R654A、W229A/R245A/N419A/R654A、Y211A/R245A/T392A/N419A/R654A、或Y211A/W229A/R245A/N419A/R654A。

[0023] 在一些实施例中,变体SaCas9蛋白包含在以下处的突变:Y211;Y212;W229;Y230;R245;T392;N419;L446;Q488;N492;Q495;R497;N498;R499;Q500;K518;K523;K525;H557;R561;K572;R634;Y651;R654;G655;N658;S662;N667;R686;K692;R694;H700;K751;D786;T787;Y789;T882;K886;N888;889;L909;N985;N986;R991;R1015;N44;R45;R51;R55;R59;R60;R116;R165;N169;R208;R209;Y211;T238;Y239;K248;Y256;R314;N394;Q414;K57;R61;H111;K114;V164;R165;L788;S790;R792;N804;Y868;K870;K878;K879;K881;Y897;R901;和/或K906。

[0024] 在一些实施例中,变体SaCas9蛋白包含以下突变中的一个或多个:Y211A;Y212A;W229A;Y230A;R245A;T392A;N419A;L446A;Q488A;N492A;Q495A;R497A;N498A;R499A;Q500A;K518A;K523A;K525A;H557A;R561A;K572A;R634A;Y651A;R654A;G655A;N658A;S662A;N667A;R686A;K692A;R694A;H700A;K751A;D786A;T787A;Y789A;T882A;K886A;N888A;A889A;L909A;N985A;N986A;R991A;R1015A;N44A;R45A;R51A;R55A;R59A;R60A;R116A;R165A;N169A;R208A;R209A;T238A;Y239A;K248A;Y256A;R314A;N394A;Q414A;K57A;R61A;H111A;K114A;V164A;R165A;L788A;S790A;R792A;N804A;Y868A;K870A;K878A;K879A;K881A;Y897A;R901A;K906A。

[0025] 在一些实施例中,变体SaCas9蛋白包含以下另外的突变中的一个或多个:Y211A、W229A、Y230A、R245A、T392A、N419A、L446A、Y651A、R654A、D786A、T787A、Y789A、T882A、K886A、N888A、A889A、L909A、N985A、N986A、R991A、R1015A、N44A、R45A、R51A、R55A、R59A、R60A、R116A、R165A、N169A、R208A、R209A、T238A、Y239A、K248A、Y256A、R314A、N394A、Q414A、K57A、R61A、H111A、K114A、V164A、R165A、L788A、S790A、R792A、N804A、Y868A、K870A、K878A、K879A、K881A、Y897A、R901A、K906A。

[0026] 在一些实施例中,变体SaCas9蛋白包含多重取代突变:R245/T392/N419/R654和Y221/R245/N419/R654(四重变体突变体);N419/R654,R245/R654,Y221/R654,和Y221/N419(二重突变体);R245/N419/R654,Y211/N419/R654,和T392/N419/R654(三重突变体)。在一些实施例中,突变体包含代替野生型氨基酸的丙氨酸。

[0027] 在一些实施例中,变体SaCas9蛋白还包含降低核酸酶活性的选自下组的一个或多个突变,该组由以下组成:在D10、E477、D556、H701、或D704处的突变;和在H557或N580处的突变。在一些实施例中,突变是:(i) D10A或D10N,(ii) H557A、H557N、或H557Y,(iii) N580A,和/或(iv) D556A。

[0028] 在一些实施例中,变体SaCas9蛋白包含以下突变中的一个或多个:E782K、K929R、N968K、或R1015H。特别是E782K/N968K/R1015H(KKH变体);E782K/K929R/R1015H(KRH变体);或E782K/K929R/N968K/R1015H(KRKH变体)。

[0029] 在一些实施例中,变体Cas9蛋白包含针对以下区域中的一个或多个的突变,以增加特异性:

[0030]

功能区域	SpCas9	SaCas9
接触间隔子区域的 DNA 的残基	L169; Y450; M495; N497; W659; R661; M694; Q695; H698; A728; Q926; E1108; V1015	Y211; W229; Y230; R245; T392; N419; L446; Y651; R654
潜在地接触非靶链的 DNA 的残基	N14; S15; S55; S730; K775; S777; R778; R780; K782; R783; K789; K797; Q805; N808; K810; R832; Q844; S845; K848; S851; K855; R859; K862; K890; Q920; K961; S964; K968; K974; R976; N980; H982; K1003; K1014; S1040; N1041; N1044; K1047; K1059; R1060; K1200; H1241; Q1254; Q1256; K1289; K1296; K1297; K1300; H1311; K1325	Q488A; N492A; Q495A; R497A; N498A; R499; Q500; K518; K523; K525; H557; R561; K572; R634; R654; G655; N658; S662; N667; R686; K692; R694; H700; K751
接触 PAM 区域的 DNA 的残基 (包括直接 PAM 接触)	R71; Y72; R78; R165; R403; T404; F405; K1107; S1109; R1114; S1116; K1118; D1135; S1136; K1200; S1216; E1219; R1333; R1335; T1337	D786; T787; Y789; T882; K886; N888; A889; L909; N985; N986; R991; R1015

[0031]

接触间隔子区域的 RNA 的残基	Y72 ; R75 ; K76 ; L101 ; S104 ; F105 ; R115 ; H116 ; I135 ; H160 ; K163 ; Y325 ; H328 ; R340 ; F351 ; D364 ; Q402 ; R403 ; I1110 ; K1113 ; R1122 ; Y1131	N44 ; R45 ; R51 ; R55 ; R59 ; R60 ; R116 ; R165 ; N169 ; R208 ; R209 ; Y211 ; T238 ; Y239 ; K248 ; Y256 ; R314 ; N394 ; Q414
接触重复区域/反 重复区域的 RNA 的残基	R63 ; R66 ; R70 ; R71 ; R74 ; R78 ; R403 ; T404 ; N407 ; R447 ; I448 ; Y450 ; K510 ; Y515 ; R661 ; V1009 ; Y1013	K57 ; R61 ; H111 ; K114 ; V164 ; R165 ; L788 ; S790 ; R792 ; N804 ; Y868 ; K870 ; K878 ; K879 ; K881 ; Y897 ; R901 ; K906
接触 RNA 茎环的 残基	K30 ; K33 ; N46 ; R40 ; K44 ; E57 ; T62 ; R69 ; N77 ; L455 ; S460 ; R467 ; T472 ; I473 ; H721 ; K742 ; K1097 ; V1100 ; T1102 ; F1105 ; K1123 ; K1124 ; E1225 ; Q1272 ; H1349 ; S1351 ; Y1356	R47 ; K50 ; R54 ; R58 ; H62 ; R209 ; E213 ; S219 ; R452 ; K459 ; R774 ; N780 ; R781 ; L783

[0032] 本文还提供了融合蛋白,其包含使用任选的介入接头(intervening linker)与异源功能结构域融合的、本文所述的分离的变体Cas9蛋白,其中该接头不干扰该融合蛋白的活性。在一些实施例中,异源功能结构域直接作用于DNA或蛋白,例如作用于染色质。在一些实施例中,异源功能结构域是转录激活结构域。在一些实施例中,转录激活结构域来自VP64或NF- κ B p65。在一些实施例中,异源功能结构域是转录沉默子或转录阻遏结构域。在一些实施例中,转录阻遏结构域是克鲁贝尔相关盒(Kruppel-associated box, KRAB)结构域、ERF阻遏因子域(ERD)或mSin3A相互作用结构域(SID)。在一些实施例中,转录沉默子是异染色质蛋白1(HP1),例如HP1 α 或HP1 β 。在一些实施例中,异源功能结构域是修饰DNA的甲基化状态的酶。在一些实施例中,修饰DNA的甲基化状态的酶是DNA甲基转移酶(DNMT)或TET蛋白的全部或双加氧酶结构域,例如包含富半胱氨酸延伸和由7个高度保守的外显子编码的20GFeD0结构域的催化模块,例如包含氨基酸1580-2052的Tet1催化结构域,包含氨基酸1290-1905的Tet2,和包含氨基酸966-1678的Tet3。在一些实施例中,TET蛋白或TET源双加氧酶结构域来自TET1。在一些实施例中,异源功能结构域是修饰组蛋白亚基的酶。在一些实施例中,修饰组蛋白亚基的酶是组蛋白乙酰转移酶(HAT)、组蛋白脱乙酰酶(HDAC)、组蛋白甲基转移酶(HMT)或组蛋白脱甲基酶。在一些实施例中,异源功能结构域是生物系链。在一些实施例中,生物系链是MS2、Csy4或 λ N蛋白。在一些实施例中,异源功能结构域是FokI。

[0033] 本文还提供了编码本文描述的变体Cas9蛋白的核酸、分离的核酸,连同载体,这些载体包含任选地有效地连接到一个或多个调节结构域的该分离的核酸,以用于表达本文描述的变体Cas9蛋白。本文还提供了宿主细胞,例如细菌的、酵母的、昆虫的、或哺乳动物的宿

主细胞或转基因动物(例如小鼠),其包含本文描述的核酸,并且任选地表达本文描述的变体Cas9蛋白。

[0034] 本文还提供了编码Cas9变体的分离的核酸,连同载体(这些载体包含任选地有效地连接到一个或多个调节结构域的该分离的核酸,以用于表达这些变体),以及包含核酸并任选表达这些变体蛋白的宿主细胞(例如,哺乳动物宿主细胞)。

[0035] 本文还提供了通过在细胞中表达如本文描述的变体Cas9蛋白或融合蛋白,以及具有与在基因组靶位点具有最佳核苷酸间距的细胞的基因组的所选部分互补的区域的至少一个指导RNA,或使该细胞与所述变体Cas9蛋白或融合蛋白以及所述指导RNA接触,来改变该细胞的基因组或表观基因组的方法。这些方法可以包括使细胞与编码Cas9蛋白和指导RNA的核酸(例如,在单个载体中)接触;使细胞与编码Cas9蛋白的核酸和编码指导RNA的核酸(例如在多重载体中)接触;并且尤其是使细胞与纯化的Cas9蛋白、和合成的或纯化的gRNA的复合物接触。在一些实施例中,细胞稳定地表达gRNA或变体蛋白/融合蛋白中的一者或二者,并且将其他元件转染进或引入细胞。例如,细胞可以稳定地表达如本文描述的变体蛋白或融合蛋白,并且这些方法可以包括使细胞与合成gRNA、纯化的重组产生的gRNA、或编码gRNA的核酸接触。在一些实施例中,变体蛋白或融合蛋白包含核定位序列、细胞穿透肽序列、和/或亲和标签中的一种或多种。

[0036] 本文还提供了用于在体外通过使分离的dsDNA与如本文描述的纯化的变体蛋白变体或融合蛋白,以及具有与该dsDNA分子的所选部分互补的区域的指导RNA接触,来改变(例如,选择性地改变)该dsDNA分子的方法。

[0037] 除非另外定义,否则所有在此使用的技术和科学术语具有与本发明所属领域的普通技术人员通常所理解的相同的意思。本文描述了用于本发明的方法和材料;还可使用本领域已知的其他合适的方法和材料。这些材料、方法、以及实例仅是说明性的,并且不旨在进行限制。此处提及的所有出版物、专利申请、专利、序列、数据库条目和其他参考文献通过引用以其整体结合。在有矛盾的情况下,将以本说明书(包括定义)为准。

[0038] 本发明的其他特征和优点将从下列详述和附图以及从权利要求中显而易见。

附图说明

[0039] 图1A-E|携带形成非特异性DNA接触的残基的突变的SpCas9变体的鉴定和表征。A,描绘基于PDB 400G和4UN3(分别改编自参考序列31和32),靶DNA:sgRNA双链体的野生型SpCas9识别的示意图。B,在与DNA主链形成氢键的位置包含丙氨酸取代的SpCas9变体的表征。当用完全匹配的sgRNA或编码与靶位点的错配的四种其他sgRNA进行编程时,使用人类细胞EGFP破坏测定对野生型SpCas9和变体进行评估。对于 $n=3$,误差条表示平均数标准误差(s.e.m.);用红色虚线表示背景EGFP损失的平均水平(对于此,小图和小图C)。C和D,通过EGFP破坏测定来评估横跨24个位点的野生型SpCas9和SpCas9-HF1的中靶活性(小图C),并且通过T7E1测定来评估13个内源位点的中靶活性(小图D)。对于 $n=3$,误差条表示平均数标准误差。E,SpCas9-HF1与野生型SpCas9的中靶活性的比率(来自小图C和D)。

[0040] 图2A-C|针对标准靶位点,具有sgRNA的野生型SpCas9和SpCas9-HF1的全基因组特异性。A,具有八个靶向内源人类基因的sgRNA的野生型SpCas9和SpCas9-HF1的脱靶位点,如通过GUIDE-seq确定。读数计数表示在给定位点的切割频率的度量;在间隔子或PAM内的错

配位置以颜色进行突出显示。B,通过GUIDE-seq,针对来自小图A中使用的八个sgRNA的野生型SpCas9和SpCas9-HF1鉴定的全基因组脱靶位点总数的总结。C,针对八个sgRNA,鉴定野生型SpCas9和SpCas9-HF1的脱靶位点,相对于中靶位点,根据错配(在原型间隔子和PAM内)的总数进行分级。

[0041] 图3A-C|通过由GUIDE-seq鉴定的脱靶位点的靶向深度测序进行的SpCas9-HF1特异性改进的验证。A,针对具有来自图2的六个sgRNA的野生型SpCas9和SpCas9-HF1,通过深度测序确定平均中靶百分比修饰。对于 $n=3$,误差条表示平均数标准误差。B,包含indel突变的深度测序的中靶位点和GUIDE-seq检测的脱靶位点的百分率。针对野生型SpCas9、SpCas9-HF1、和对照条件,绘制一式三份的实验的图。x轴之下的实心圆圈表示多个重复,针对其未观察到插入或缺失突变。在具有星号的红色文本中显示不能通过PCR扩增的脱靶位点。仅在EMX1-1脱靶1和FANCF-3脱靶1,针对SpCas9-HF1和对照条件之间的比较,使用具有汇集的读数计数的单侧Fisher精确检验的假设检验发现显著差异(在使用Benjamini-Hochberg方法调节多重比较后, $p<0.05$)。还发现在所有脱靶位点,野生型SpCas9和SpCas9-HF1之间存在显著差异,并且在所有脱靶位点(除了RUNX1-1脱靶2外),野生型SpCas9和对照条件之间存在显著差异。C,在具有野生型SpCas9的中靶和脱靶切割位点,GUIDE-seq读数计数(来自图2A)和通过深度测序确定的平均百分比修饰之间的相关性的散点图。

[0042] 图4A-C|具有sgRNA的野生型SpCas9和SpCas9-HF1的、非标准重复位点的全基因组特异性。A,使用已知用来切割大量脱靶位点的两种sgRNA,野生型SpCas9和SpCas9-HF1的GUIDE-seq特异性图(Fu等人,Nat Biotechnol[自然生物技术]31,822-826(2013);Tsai等人,Nat Biotechnol[自然生物技术]33,187-197(2015))。GUIDE-seq读数计数表示在给定位点的切割频率的度量;在间隔子或PAM内的错配位置以颜色进行突出显示;红圈指示在sgRNA-DNA界面,位点可能具有指示的凸起(Lin等人,Nucleic Acids Res[核酸研究]42,7473-7485(2014));蓝圈指示相对于显示的空位比对可以具有替代性空位比对的位点(参见图8)。B,针对来自小图A中使用的两个sgRNA的野生型SpCas9和SpCas9-HF1,通过GUIDE-seq鉴定的全基因组脱靶位点的总数的总结。C,针对VEGFA位点2和3,使用野生型SpCas9或SpCas9-HF1对脱靶位点进行鉴定,相对于中靶位点根据错配(在原型间隔子和PAM内)的总数进行分级。在小图A中标记有红圈的脱靶位点并不包括在这些计数中;用非空位比对的错配的数目,计数小图A中标记有蓝圈的位点。

[0043] 图5A-D|携带另外的取代的SpCas9-HF1衍生物的活性。A,具有八个sgRNA的野生型SpCas9、SpCas9-HF1、和SpCas9-HF1衍生物变体的人类细胞EGFP破坏活性。SpCas9-HF1具有N497A、R661A、Q695、和Q926A突变;HF2=HF1+D1135E;HF3=HF1+L169A;HF4=HF1+Y450A。对于 $n=3$,误差条表示平均数标准误差;用红色虚线表示背景EGFP损失的平均水平。B,与具有来自小图a的八个sgRNA的野生型SpCas9相比,当使用SpCas9-HF变体时,中靶活性的总结。显示了中位数和四分位数间距;将显示>70%的野生型活性的间隔用绿色突出显示。C,在对SpCas9-HF1的作用具有抗性的来自图2A和4A的FANCF位点2和VEGFA位点3中靶位点、连同脱靶位点,SpCas9和HF变体的平均百分比修饰。通过T7E1测定确定百分比修饰;针对所有实验,减去背景indel百分率。对于 $n=3$,误差条表示平均数标准误差。D,具有FANCF位点2或VEGFA位点3sgRNA的野生型SpCas9和HF变体的特异性比率,绘制为中靶与脱靶活性的比率(来自小图C)。

[0044] 图5E-F|具有sgRNA的SpCas9-HF1、-HF2、和-HF4的全基因组特异性,这些sgRNA具有对SpCas9-HF1的作用具有抗性的脱靶位点。E,对于小图F中的GUIDE-seq实验,在预期中靶位点处的平均GUIDE-seq标签整合。SpCas9-HF1=N497A/R661A/Q695A/Q926A;HF2=HF1+D1135E;HF4=HF1+Y450A。对于n=3,误差条表示平均数标准误差。F,具有FANCF位点2或VEGFA位点3sgRNA的SpCas9-HF1、-HF2、或-HF4的GUIDE-seq鉴定的脱靶位点。读数计数表示在给定位点的切割频率的度量;在间隔子或PAM内的错配位置以颜色进行突出显示。在SpCas9-HF变体之间的比较之前,通过将SpCas9-HF变体的脱靶读数计数归一化为在中靶位点的读数计数,计算脱靶辨别的倍数改进。

[0045] 图6A-B|SpCas9与sgRNA和靶DNA的相互作用。A,说明SpCas9:sgRNA复合物的示意图,其中sgRNA和靶DNA之间进行碱基配对。B, SpCas9:sgRNA复合物结合来自PDB:4UN3 (ref.32)的靶DNA的结构示意图。与靶链DNA主链形成氢键接触的四个残基用蓝色突出显示;出于可视化目的,隐藏HNH结构域。

[0046] 图7A-D|具有用于GUIDE-seq实验的不同sgRNA的野生型和SpCas9-HF1的中靶活性比较。A和C,通过限制性片段长度多态性测定量化,针对图2A和4A(分别是小图7A和7C)显示的GUIDE-seq实验,在预期中靶位点的平均GUIDE-seq标签整合。对于n=3,误差条表示平均数标准误差。b和d,通过T7E1测定检测,针对图2A和4A(分别是小图7B和7D)显示的GUIDE-seq实验,在预期中靶位点的平均百分比修饰。对于n=3,误差条表示平均数标准误差。

[0047] 图8|针对VEGFA位点2脱靶位点的潜在替代性比对。使用Geneious (Kearse等人, Bioinformatics[生物信息学]28,1647-1649(2012))版本8.1.6进行比对,用会潜在地被识别为包含单核苷酸空位(Lin等人, Nucleic Acids Res[核酸研究]42,7473-7485(2014)) (右)的脱靶位点的GUIDE-seq(左)鉴定的十个VEGFA位点2脱靶位点。

[0048] 图9|具有截短的sgRNAs14的野生型SpCas9和SpCas9-HF1的活性。使用靶向EGFP中的四个位点的全长或截短的sgRNA,野生型SpCas9和SpCas9-HF1的EGFP破坏。对于n=3,误差条表示平均数标准误差(s.e.m.);用红色虚线表示对照试验中背景EGFP损失的平均水平。

[0049] 图10|具有携带5' 错配的鸟嘌呤碱基的sgRNA的野生型SpCas9和SpCas9-HF1活性。具有靶向四个不同位点的sgRNA的野生型SpCas9和SpCas9-HF1的EGFP破坏活性。对于每个靶位点,sgRNA包含匹配的非鸟嘌呤5'-碱基或有意错配的5'-鸟嘌呤。

[0050] 图11|滴定野生型SpCas9和SpCas9-HF1表达质粒的量。用不同量的野生型和SpCas9-HF1表达质粒转染的人类细胞EGFP破坏活性。对于所有转染,将含sgRNA质粒的量固定在250ng。使用靶向单独位点的两个sgRNA;对于n=3,误差条表示平均数标准误差(s.e.m.);用红色虚线表示阴性对照中背景EGFP损失的平均水平。

[0051] 图12A-D|改变SpCas9-HF1的PAM识别特异性。A,用T7E1测定量化,使用8个sgRNA,通过SpCas9-VQR(ref.15)和改进的SpCas9-VRQR进行的中靶内源人类位点的平均百分比修饰的比较。将两种变体都工程化以识别NGAN PAM。对于n=2或3,误差条表示平均数标准误差。B,与使用八个sgRNA的SpCas9-HF1对应物相比,SpCas9-VQR和SpCas9-VRQR的中靶EGFP破坏活性。对于n=3,误差条表示平均数标准误差(s.e.m.);用红色虚线表示阴性对照中背景EGFP损失的平均水平。C,用T7E1测定量化,在八个内源人类基因位点,与SpCas9-HF1变体相比,通过SpCas9-VQR和SpCas9-VRQR进行的平均中靶百分比修饰的比较。对于n=3,误差

条表示平均数标准误差 (s.e.m.); ND, 未检测到。D, 与相应的SpCas9-HF1变体 (来自小图B和C) 相比, 当使用SpCas9-VQR或SpCas9-VRQR时, 中靶活性的倍数改变的总结。显示了中位数和四分位数间距; 将显示>70%的野生型活性的间隔用绿色突出显示。

[0052] 图13A-B|野生型SpCas9、SpCas9-HF1、和在可能潜在接触非靶DNA链的位置携带一个或多个丙氨酸取代的野生型SpCas9衍生物的活性。A和B, 采用与EGFP基因中的位点完全匹配的sgRNA, 连同有意在位置11和12 (小图A) 或位置9和10 (小图B) 错配的sgRNA, 使用EGFP破坏测定, 评估核酸酶。用为PAM最远端位置的位置20, 对错配的位置进行编号; 红色虚线表示EGFP破坏的背景水平; HF1=具有N497A/R661A/Q695A/Q926A取代的SpCas9。

[0053] 图14A-B|野生型SpCas9、SpCas9-HF1、和在可能潜在接触非靶DNA链的位置携带一个或多个丙氨酸取代的SpCas9-HF1衍生物的活性。A和B, 采用与EGFP基因中的位点完全匹配的sgRNA, 连同有意在位置11和12 (小图A) 或位置9和10 (小图B) 错配的sgRNA, 使用EGFP破坏测定, 评估核酸酶。用为PAM最远端位置的位置20, 对错配的位置进行编号; 红色虚线表示EGFP破坏的背景水平; HF1=具有N497A/R661A/Q695A/Q926A取代的SpCas9。

[0054] 图15|野生型SpCas9、SpCas9-HF1、和在可能潜在接触非靶DNA链的位置携带一个或多个丙氨酸取代的SpCas9 (Q695A/Q926A) 衍生物的活性。采用与EGFP基因中的位点完全匹配的sgRNA, 连同有意在位置11和12错配的sgRNA, 使用EGFP破坏测定, 评估核酸酶。用为PAM最远端位置的位置20, 对错配的位置进行编号; 红色虚线表示EGFP破坏的背景水平; HF1=具有N497A/R661A/Q695A/Q926A取代的SpCas9; Db1=具有Q695A/Q926A取代的SpCas9。

[0055] 图16|使用匹配的sgRNA和在间隔子中的每个位置处具有单错配的sgRNA, 野生型SpCas9、SpCas9-HF1、和eSpCas9-1.1的活性。采用与EGFP基因中的位点完全匹配的sgRNA (“匹配的”), 连同有意在指示的位置错配的sgRNA, 使用EGFP破坏测定, 评估核酸酶。用为PAM最远端位置的位置20, 对错配的位置进行编号。SpCas9-HF1=N497A/R661A/Q695A/Q926A, 并且eSP1.1=K848A/K1003A/R1060A。

[0056] 图17A-B|使用匹配的sgRNA和在间隔子中的不同位置处具有单错配的sgRNA, 野生型SpCas9和变体的活性。(A) 采用与EGFP基因中的位点完全匹配的sgRNA (“匹配的”), 连同有意在指示的间隔子位置错配的sgRNA, 使用EGFP破坏测定, 评估包含丙氨酸取代 (针对可能潜在接触靶或非靶DNA链的位置) 的组合的SpCas9核酸酶的活性。(B) 使用针对匹配的中靶位点的所有可能的单错配sgRNA的剩余者, 测试来自 (a) 的这些核酸酶的子集。用为PAM最远端位置的位置20, 对错配的位置进行编号。mm=错配, WT=野生型, Db=Q695A/Q926A, HF1=N497A/R661A/Q695A/Q926A, 1.0=K810A/K1003A/R1060A, 并且1.1=K848A/K1003A/R1060A。

[0057] 图18|使用匹配的sgRNA和在间隔子中的不同单个位置处具有错配的sgRNA, 野生型SpCas9和变体的活性。采用与EGFP基因中的位点完全匹配的sgRNA (“匹配的”), 连同有意在指示的位置错配的sgRNA, 使用EGFP破坏测定, 评估包含丙氨酸取代 (针对可能潜在接触靶或非靶DNA链的位置) 的组合的SpCas9核酸酶的活性。Db=Q695A/Q926A, HF1=N497A/R661A/Q695A/Q926A。

[0058] 图19A-B|使用匹配的sgRNA和在间隔子中的不同单个位置处具有错配的sgRNA, 野生型SpCas9和变体的活性。(A) 采用与EGFP基因中的位点完全匹配的两个sgRNA, 使用EGFP破坏测定, 评估包含丙氨酸取代 (针对可能潜在接触靶或非靶DNA链的位置) 的组合的

SpCas9核酸酶的中靶活性。(B) 采用在其间隔子序列中包含 (sgRNA“位点1”) 的位置12、14、16、或18的错配的sgRNA,测试来自 (a) 的这些核酸酶的子集,以确定是否对这些取代赋予的错配不耐受。Db=Q695A/Q926A, HF1=N497A/R661A/Q695A/Q926A。

[0059] 图20|SpCas9 (顶图) 和SaCas9 (底图) 的结构比较,例示了四重突变体构建体 (以黄色球形表示显示) 中突变的位置之间的相似性。而且,以粉色球形表示显示的是接触DNA主链的其他残基。

[0060] 图21A-B|携带一个或多个丙氨酸取代的野生SaCas9和SaCas9衍生物的活性。A和B, SaCas9取代针对可能潜在接触靶DNA链的位置 (小图A), 或先前已经显示影响PAM特异性的位置 (小图B)。采用与EGFP基因中的位点完全匹配的sgRNA,连同有意在位置11和12错配的sgRNA,使用EGFP破坏测定,评估核酸酶。用为PAM最远端位置的位置20,对错配的位置进行编号;红色虚线表示EGFP破坏的背景水平。

[0061] 图22A-B|野生型 (WT) SaCas9和在可能潜在接触靶DNA链的残基处携带一个或多个丙氨酸取代的SaCas9衍生物的活性。A和B,采用与EGFP基因中的位点完全匹配的sgRNA (“匹配的”),连同有意在位置19和20错配的sgRNA,使用EGFP破坏测定,评估核酸酶。用为PAM最远端位置的位置20,对错配的位置进行编号。

[0062] 图23|野生型 (WT) SaCas9和在可能潜在接触靶DNA链的残基处携带丙氨酸取代的三重组合的SaCas9变体的活性。使用EGFP破坏测定评估核酸酶。使用四个不同的sgRNA (匹配的#1-4),其中还用已知由野生型SaCas9有效使用的错配的sgRNA,对四个靶位点中的每一个进行测试。针对每个位点的错配的sgRNA显示在每个匹配的sgRNA的右侧 (例如,针对匹配位点3的唯一错配的sgRNA是mm 11&12)。用为PAM最远端位置的位置21,对错配的位置进行编号;mm,错配。

[0063] 图24A-B|野生型 (WT) SaCas9和在可能潜在接触靶DNA链的残基处携带一个或多个丙氨酸取代的SaCas9衍生物的活性。A和B,使用T7E1测定,针对匹配的和单错配的内源人类基因靶位点,对携带二重 (A) 或三重 (B) 组合取代的SaCas9变体进行评估。匹配的“中靶”位点根据其基因靶位点sgRNA数进行命名,根据Kleinstiver等人, Nature Biotechnology [自然生物技术] 2015进行。采用发生在位置21 (PAM最远端位置) 的错配,对错配的sgRNA进行编号;错配的sgRNA源自在错配的sgRNA的左侧列出的匹配的中靶位点。

具体实施方案

[0064] CRISPR-Cas9核酸酶的限制是其在不完全匹配靶位点处,诱导不希望的“脱靶”突变的潜能 (参见例如Tsai等人, Nat Biotechnol. [自然生物技术] 2015), 在一些情况下,其频率比得上在预期中靶位点处观察到的那些 (Fu等人, Nat Biotechnol. [自然生物技术] 2013)。用CRISPR-Cas9核酸酶进行的先前工作已经提示,减少指导RNA (gRNA) 和靶位点的间隔子区域之间的序列特异性相互作用的数目可以在人类细胞中减少切割的脱靶位点处的诱变效应 (Fu等人, Nat Biotechnol. [自然生物技术] 2014)。

[0065] 这早先是通过在其5'末端将gRNA截短2或3nt来实现,并且假设增加的特异性的机制是gRNA/Cas9复合物的相互作用能的降低,这样使得它达到平衡,具有刚好足够的能量来切割中靶位点,使它更不可能具有足够的能量来切割脱靶位点,在脱靶位点,由于靶DNA位点的错配,将可能存在能量惩罚 (WO 2015/099850)。

[0066] 假设可能通过减少与靶DNA位点的非特异性相互作用,将SpCas9的脱靶效应(在是与针对指导RNA的预期靶位点不完全匹配或错配的DNA位点处)最小化。SpCas9-sgRNA复合物切割由以下组成的靶位点:NGG PAM序列(由SpCas9识别)(Deltcheva,E.等人,Nature[自然]471,602-607(2011);Jinek,M.等人,Science[科学]337,816-821(2012);Jiang,W.等人,Nat Biotechnol[自然生物技术]31,233-239(2013);(Sternberg,S.H.等人,Nature[自然]507,62-67(2014))以及相邻的20bp原型间隔子序列(它与sgRNA的5'末端互补)(Jinek,M.等人,Science[科学]337,816-821(2012);Jinek,M.等人,Elife 2,e00471(2013);Mali,P.等人,Science[科学]339,823-826(2013);Cong,L.等人,Science[科学]339,819-823(2013))。先前建立理论:SpCas9-sgRNA复合物可以拥有比识别其预期靶DNA位点所需更多的能量,由此使得能够切割错配的脱靶位点(Fu,Y.,等人,Nat Biotechnol[自然生物技术]32,279-284(2014))。人们可以预想,对于Cas9在适应性细菌免疫中的预期作用,这一特性可能是有利的,赋予细菌切割可能会突变的外源序列的能力。先前研究也支持这一过剩能量模型,这些先前研究展示,通过以下,可以减小脱靶效应(但是未消除):通过减小SpCas9浓度(Hsu,P.D.等人,Nat Biotechnol[自然生物技术]31,827-832(2013);Pattanayak,V.等人,Nat Biotechnol[自然生物技术]31,839-843(2013)),或通过减小sgRNA的互补性长度(Fu,Y.,等人,Nat Biotechnol[自然生物技术]32,279-284(2014)),尽管也已经提出了针对此效应的其他解释(Josephs,E.A.等人,Nucleic Acids Res[核酸研究]43,8924-8941(2015);(Sternberg,S.H.等人,Nature[自然]527,110-113(2015);(Kiani,S.等人,Nat Methods[自然方法]12,1051-1054(2015)))。结构数据提示,可以通过若干SpCas9介导的DNA接触,来稳定SpCas9-sgRNA-靶DNA复合物,这些DNA接触包括由四个SpCas9残基(N497、R661、Q695、Q926)与靶DNA链的磷酸主链形成的直接氢键(Nishimasu,H.等人,Cell[细胞]156,935-949(2014);Anders,C.,等人,Nature[自然]513,569-573(2014))(图1a和图6a以及6b)。本发明的诸位发明人预想,这些接触中的一个或多个的破坏可能能量上将SpCas9-sgRNA复合物平衡在刚好足以保留稳健的中靶活性,但是对于切割错配的脱靶位点,具有降低的活性。

[0067] 如此处所述,理论上可以通过减小Cas9对DNA的结合亲和力,将Cas9蛋白工程化以显示增加的特异性。通过使用结构信息、基于细菌选择的定向进化、和组合设计,将单个丙氨酸取代引入预期可能与DNA主链上的磷酸相互作用的SpCas9中的不同残基,将广泛使用的酿脓链球菌Cas9(SpCas9)的若干变体工程化。可以使用稳健的基于大肠杆菌的筛选测定,进一步测试变体的细胞活性,以便评估这些变体的细胞活性;在此细菌系统中,细胞存活取决于包含针对毒性旋转酶毒物ccdB的基因和gRNA与SpCas9靶向的23个碱基对的序列的选择质粒的切割和随后的破坏,并且导致与保留地或损失的活性相关的残基的鉴定。此外,鉴定并且表征在人类细胞中表现出改进的靶特性的另一种SpCas9变体。

[0068] 此外,如在基于细菌细胞的系统中评估SpCas9的单丙氨酸取代突变体的活性,指示在50%-100%之间的存活百分率通常指示稳健的切割,而0%存活指示酶已经在功能上被损害。然后在细菌中测定SpCas9的另外的突变,包括:R63A、R66A、R69A、R70A、R71A、Y72A、R74A、R75A、K76A、N77A、R78A、R115A、H160A、K163A、R165A、L169A、R403A、T404A、F405A、N407A、R447A、N497A、I448A、Y450A、S460A、M495A、K510A、Y515A、R661A、M694A、Q695A、H698A、Y1013A、V1015A、R1122A、K1123A、K1124A、K1158A、K1185A、K1200A、S1216A、Q1221A、

K1289A、R1298A、K1300A、K1325A、R1333A、K1334A、R1335A、和T1337A。除了在细菌中具有<5%存活的2个突变体(R69A和F405A),所有这些另外的单突变似乎对SpCas9的中靶活性没有影响(在细菌筛选中>70%存活)。

[0069] 为了进一步确定在细菌筛选中鉴定的Cas9的变体在人类细胞中是否能有效发挥功能,使用基于人类U2OS细胞的EGFP破坏测定,测试不同丙氨酸取代Cas9突变体。在此测定中,在单个整合的、组成型地表达的EGFP基因的编码序列中的靶位点的成功切割导致indel突变的诱导和EGFP活性的破坏,这是通过流式细胞术定量评估的(参见例如Reyon等人,Nat Biotechnol.[自然生物技术]2012年5月;30(5):460-5)。

[0070] 这些实验显示,在基于细菌细胞的测定中获得的结果与人类细胞中的核酸酶活性很好地关联,提示这些工程化策略可以延伸至来自其他物种和不同细胞的Cas9。因此,这些发现提供了对SpCas9和SaCas9变体的支持,它们在此统称为“变体”或“这些变体”。

[0071] 本文描述的所有变体可以快速地并入现有的和广泛使用的载体中,例如通过简单的定点诱变,并且由于它们仅需要少量突变,所以这些变体应当还可以与对以下的其他先前描述的改进一起起作用:SpCas9平台(例如,截短的sgRNA(Tsai等人,Nat Biotechnol[自然生物技术]33,187-197(2015);Fu等人,Nat Biotechnol[自然生物技术]32,279-284(2014)),切口酶突变(Mali等人,Nat Biotechnol[自然生物技术]31,833-838(2013);Ran等人,Cell[细胞]154,1380-1389(2013)),FokI-dCas9融合(Guillinger等人,Nat Biotechnol[自然生物技术]32,577-582(2014);Tsai等人,Nat Biotechnol[自然生物技术]32,569-576(2014);WO 2014144288);以及具有改变的PAM特异性的工程化的CRISPR-Cas9核酸酶(Kleinstiver等人,Nature[自然].2015年7月,23;523(7561):481-5)。

[0072] 因此,此处提供的是Cas9变体,包括SpCas9变体。SpCas9野生型序列如下:

[0073]

10	20	30	40	50	60
MDKKYSIGLD	IGTNSVGWAV	ITDEYKVPSK	KFKVLGNTDR	HSIKKNLIGA	LLFDSGETAE
70	80	90	100	110	120
ATRLKRTARR	RYTRRKNRIC	YLQEIFSNE	AKVDDSFHHR	LEESFLVEED	KKHERHPIFG
130	140	150	160	170	180
NIVDEVAYHE	KYPTIYHLRK	KLVDSTDKAD	LRLIYLALAH	MIKFRGHFLI	EGDLNPDNSD
190	200	210	220	230	240
VDKLFIQLVQ	TYNQLFEENP	INASGVDAKA	ILSARLSKSR	RLENLIAQLP	GEKKNGLFGN
250	260	270	280	290	300
LIALSLGLTP	NFKSNFDLAE	DAKLQLSKDT	YDDDLNLLA	QIGDQYADLF	LAAKNLSDAI
310	320	330	340	350	360
LLSDILRVNT	EITKAPLSAS	MIKRYDEHHQ	DLTLLKALVR	QQLPEKYKEI	FFDQSKNGYA
370	380	390	400	410	420
GYIDGGASQE	EFYKFIKPIL	EKMDGTEELL	VKLNREDLLR	KQRTFDNGSI	PHQIHLGELH
430	440	450	460	470	480
AILRRQEDFY	PFLKDNREKI	EKILTFRIPY	YVGPLARGNS	RFAWMTRKSE	ETITPWNFEE
490	500	510	520	530	540
VVDKGASAQS	FIERMTNFDK	NLPNEKVLPK	HSLLEYEFTV	YNELTKVKYV	TEGMRKPAFL
550	560	570	580	590	600
SGEQKKAIVD	LLFKTNRKVT	VKQLKEDYFK	KIECFDSVEI	SGVEDRFNAS	LGTYHDLLKI
610	620	630	640	650	660
IKDKDFLDNE	ENEDILEDIV	LTLTLFEDRE	MIEERLKTYA	HLFDDKVMKQ	LKRRRYTGWG
670	680	690	700	710	720
RLSRKLINGI	RDKQSGKTIL	DFLKSDGFAN	RNFMQLIHDD	SLTFKEDIQK	AQVSGQGDSL

```

730      740      750      760      770      780
HEHIANLAGS PAIKKGILQT VKVVDDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER

790      800      810      820      830      840
MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDH

850      860      870      880      890      900
IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL

910      920      930      940      950      960
TKAERGGISE LDKAGFIKRQ LVETRQITKH VAQILD SRMN TKYDENDKLI REVKVITLKS

970      980      990      1000      1010      1020
KLVSDFRKDF QFYKVVREINN YHHAHDAYLN AVVG TALIKK YPKLESEFVY GDYKVYDVRK

[0074] 1030      1040      1050      1060      1070      1080
MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF

1090      1100      1110      1120      1130      1140
ATVRKVL SMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA

1150      1160      1170      1180      1190      1200
YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDLI IKLPK

1210      1220      1230      1240      1250      1260
YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE DNEQKQLFVE

1270      1280      1290      1300      1310      1320
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA

1330      1340      1350      1360
PAAFKYFDTT IDRKYTSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO:1)

```

[0075] 本文描述的SpCas9变体可以包括SEQ ID NO:1的氨基酸序列,所述氨基酸序列在以下位置中的一个或多个处具有突变(即用不同氨基酸,例如丙氨酸、甘氨酸、或丝氨酸替换天然氨基酸):N497,R661,Q695,Q926(或在与其类似的位置处)。在一些实施例中,SpCas9变体与SEQ ID NO:1的氨基酸序列是至少80%,例如至少85%、90%或95%一致的,例如,在被替换的SEQ ID NO:1的残基的高达5%、10%、15%或20%处具有差异,例如除了本文描述的突变,具有保守突变。在优选的实施例中,该变体保留亲本的所希望活性,例如核酸酶活性(亲本是切口酶或死Cas9除外),和/或与指导RNA和靶DNA相互作用的能力。

[0076] 为了测定两个核酸序列的百分比一致性,出于最佳比对的目的将序列比对(例如,在一个第一和第二氨基酸或核酸序列的一个或两个中引入空位以用于最佳比对,并且出于比较目的,非同源序列可以忽略)。为了比较目的,出于比较目的而比对的参考序列的长度是参考序列的长度的至少80%,并且在一些实施例中为至少90%或100%。然后比较相应的氨基酸位置或核苷酸位置处的核苷酸。当第一序列中的位置被与第二个序列中的对应位置相同的核苷酸占据时,那么分子在该位置是一致的(如此处所使用,核酸“一致性”等同于核酸“同源性”)。两个序列之间的百分比一致性是这些序列共享的多个一致性位置的一个函数,考虑了空位数目以及每个空位的长度,需要引入它以用于两个序列的最佳比对。两种多肽或核酸序列之间的百分比一致性以本领域技术范围内的各种方式确定,例如使用公开可获得的计算机软件,例如Smith Waterman Alignment (Smith,T.F.和M.S.Waterman (1981) J Mol Biol [分子生物学杂志] 147:195-7);“BestFit” (Smith和Waterman,Advances in

Applied Mathematics[应用数学进展],482-489(1981)),如结合于GeneMatcher PlusTM, Schwarz和Dayhof(1979)Atlas of Protein Sequence and Structure[蛋白序列和结构图谱],Dayhof,M.O.编辑,第353-358页;BLAST程序(基本局部比对搜索工具;(Altschul,S.F.,W.Gish,等人(1990)J Mol Biol[分子生物学杂志]215:403-10)、BLAST-2、BLAST-P、BLAST-N、BLAST-X、WU-BLAST-2、ALIGN、ALIGN-2、CLUSTAL、或Megalign (DNASTAR) 软件。另外,本领域技术人员可以确定用于测量比对的适当参数,包括为了在被比较的序列的长度上实现最大比对所需要的任何算法。通常,对于蛋白或核酸,比较长度可以是任何长度,直到并包括全长(例如5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%或100%)。为了本发明的组合物和方法的目的,对序列的全长的至少80%进行比对。

[0077] 为了本发明的目的,序列的比较和两个序列之间的百分比一致性的确定可以使用具有12的空位罚分、4的空位延伸罚分和5的移码空位罚分的Blossum 62评分矩阵来完成。

[0078] 保守取代典型地包括在以下组内的取代:甘氨酸,丙氨酸;缬氨酸,异亮氨酸,亮氨酸;天冬氨酸,谷氨酸,天冬酰胺,谷氨酰胺;丝氨酸,苏氨酸;赖氨酸,精氨酸;以及苯丙氨酸,酪氨酸。

[0079] 在一些实施例中,SpCas9变体包括以下组的突变中的一个:N497A/R661A/Q695/Q926A(四重丙氨酸突变体);Q695A/Q926A(二重丙氨酸突变体);R661A/Q695A/Q926A和N497A/Q695A/Q926A(三重丙氨酸突变体)。在一些实施例中,可以将L169和/或Y450处的另外的取代突变添加至这些二重、三重、和四重突变体,或添加至在Q695或Q926处携带突变的单重突变体。在一些实施例中,突变体具有代替野生型氨基酸的丙氨酸。在一些实施例中,突变体具有除了精氨酸或赖氨酸(或天然氨基酸)之外的任何氨基酸。

[0080] 在一些实施例中,SpCas9变体还包括以下突变之一,所述突变降低或破坏Cas9的核酸酶活性:D10、E762、D839、H983或D986和H840或N863,例如D10A/D10N和H840A/H840N/H840Y,以使蛋白的核酸酶部分无催化活性;在这些位置处的取代可以是丙氨酸(如在Nishimasu等人,Cell[细胞]156,935-949(2014)中的)或其他残基,例如谷氨酰胺、天冬酰胺、酪氨酸、丝氨酸或天冬氨酸,例如E762Q、H983N、H983Y、D986N、N863D、N863S、或N863H(参见WO 2014/152432)。在一些实施例中,该变体包括D10A或H840A处的突变(其产生单链切口酶)或D10A和H840A处的突变(其消除核酸酶活性;这种突变体称为死Cas9或dCas9)。

[0081] SpCas9 N497A/R661A/Q695A/R926A突变具有金黄色葡萄球菌Cas9(SaCas9)中的类似残基;参见图20。如我们对SpCas9已经观察到的那样,预期与DNA或RNA主链接触的残基的突变增加SaCas9的特异性。因此,本文还提供了SaCas9变体。

[0082] SaCas9野生型序列如下:

	10	20	30	40	50
	MKRNYILGLD	IGITSVGYGI	IDYETRDVID	AGVRLFKEAN	VENNEGRRSK
	60	70	80	90	100
	RGARRLKRRR	RHRIQRVKKL	LFDYNLLTDH	SELSGINPYE	ARVKGLSQKL
	110	120	130	140	150
[0083]	SEEEFSAALL	HLAKRRGVHN	VNEVEEDTGN	ELSTKEQISR	NSKALEEKYV
	160	170	180	190	200
	AELQLERLKK	DGEVRGSINR	FKTSDYVKEA	KQLLKVQKAY	HQLDQSFIDT
	210	220	230	240	250
	YIDLLETRRT	YYEGPGEOSP	FGWKDIKEWY	EMLMGHCTYF	PEELRSVKYA
	260	270	280	290	300
	YNADLYNALN	DLNNLVITRD	ENEKLEYEYK	FQIIENVFKQ	KKKPTLKQIA
	310	320	330	340	350
	KEILVNEEDI	KGYRVTSTGK	PEFTNLKVYH	DIKDITARKE	IIENAELLDQ
	360	370	380	390	400
	IAKILTIYQS	SEDIQEELTN	LNSELTQEEI	EQISNLKGYT	GTHNLSLKAI
	410	420	430	440	450
	NLILDELWHT	NDNQIAIFNR	LKLVPKKVDL	SQQKEIPTTL	VDDFILSPVV
	460	470	480	490	500
	KRSFIQSIKV	INAIKKYGL	PNDIIIELAR	EKNSKDAQKM	INEMQKRNRQ
	510	520	530	540	550
	TNERIEEIIR	TTGKENAKYL	IEKIKLHDMQ	EGKCLYSLEA	IPLEDLLNNP
	560	570	580	590	600
	FNYEVDHIIP	RSVSFDNSFN	NKVLVKQEEH	SKKGNRTPFQ	YLSSSDSKIS
	610	620	630	640	650
	YETFKKHILN	LAKGKGRISK	TKKEYLLEER	DINRFSVQKD	FINRNLVDTR
	660	670	680	690	700
[0084]	YATRGLMNL	RSYFRVNNLD	VKVKSINGGF	TSFLRRKWKF	KKERNKGYKH
	710	720	730	740	750
	HAEDALIIAN	ADFIFKEWKK	LDKAKKVMEN	QMFEEKQAES	MPEIETEQUEY
	760	770	780	790	800
	KEIFITPHQI	KHIKDFKDYK	YSHRVDKKPN	RELINDTLYS	TRKDDKGNTL
	810	820	830	840	850
	IVNNLNGLYD	KDNDKLKKLI	NKSPEKLLMY	HHDPQTYQKL	KLIMEQYGDE
	860	870	880	890	900
	KNPLYKYEE	TGNYLTKYSK	KDNGPVIKKI	KYYGNKLNH	LDITDDYPNS
	910	920	930	940	950
	RNKVVKLKSL	PYRFDVYLDN	GVYKFVTVKN	LDVIKKENYY	EVNSKCYEEA
	960	970	980	990	1000
	KKLKKISNQA	EFIASFYNN	LIKINGELYS	VIGVNNDLLN	RIEVNMIDIT
	1010	1020	1030	1040	1050
	YREYLENMND	KRPPRIIKTI	ASKTQSIKKY	STDILGNLYE	VSKKKHPQII
	KKG	(SEQ ID NO:2)			

[0085] 本文描述的SaCas9变体包括SEQ ID NO:2的氨基酸序列,所述氨基酸序列具有在以下位置中的一个、两个、三个、四个、五个、或全部六个位置处的突变:Y211、W229、R245、T392、N419、和/或R654,例如包含与SEQ ID NO:2的氨基酸序列至少80%一致的序列,所述

序列具有在以下位置中的一个、两个、三个、四个、五个或六个位置处的突变:Y211、W229、R245、T392、N419、和/或R654。

[0086] 在一些实施例中,变体SaCas9蛋白还包含以下突变中的一个或多个:Y211A;W229A;Y230A;R245A;T392A;N419A;L446A;Y651A;R654A;D786A;T787A;Y789A;T882A;K886A;N888A;A889A;L909A;N985A;N986A;R991A;R1015A;N44A;R45A;R51A;R55A;R59A;R60A;R116A;R165A;N169A;R208A;R209A;Y211A;T238A;Y239A;K248A;Y256A;R314A;N394A;Q414A;K57A;R61A;H111A;K114A;V164A;R165A;L788A;S790A;R792A;N804A;Y868A;K870A;K878A;K879A;K881A;Y897A;R901A;K906A。

[0087] 在一些实施例中,变体SaCas9蛋白包含以下另外的突变中的一个或多个:Y211A、W229A、Y230A、R245A、T392A、N419A、L446A、Y651A、R654A、D786A、T787A、Y789A、T882A、K886A、N888A、A889A、L909A、N985A、N986A、R991A、R1015A、N44A、R45A、R51A、R55A、R59A、R60A、R116A、R165A、N169A、R208A、R209A、Y211A、T238A、Y239A、K248A、Y256A、R314A、N394A、Q414A、K57A、R61A、H111A、K114A、V164A、R165A、L788A、S790A、R792A、N804A、Y868A、K870A、K878A、K879A、K881A、Y897A、R901A、K906A。

[0088] 在一些实施例中,变体SaCas9蛋白包含多重取代突变:R245/T392/N419/R654和Y221/R245/N419/R654(四重变体突变体);N419/R654,R245/R654,Y221/R654,和Y221/N419(二重突变体);R245/N419/R654,Y211/N419/R654,和T392/N419/R654(三重突变体)。在一些实施例中,突变体包含代替野生型氨基酸的丙氨酸。

[0089] 在一些实施例中,变体SaCas9蛋白还包含在E782K、K929R、N968K、和/或R1015H处的突变。例如,KKH变体(E782K/N968K/R1015H)、KRH变体(E782K/K929R/R1015H)、或KRKH变体(E782K/K929R/N968K/R1015H)。

[0090] 在一些实施例中,变体SaCas9蛋白还包含降低核酸酶活性的选自下组的一个或多个突变,该组由以下组成:在D10、E477、D556、H701、或D704处的突变;和在H557或N580处的突变。

[0091] 在一些实施例中,突变是:(i) D10A或D10N,(ii) H557A、H557N、或H557Y,(iii) N580A,和/或(iv) D556A。

[0092] 本文还提供了编码Cas9变体的分离的核酸,包含任选地有效地连接到一个或多个调节结构域的该分离的核酸的用于表达这些变体蛋白的载体,以及包含核酸并任选表达这些变体蛋白的宿主细胞(例如,哺乳动物宿主细胞)。

[0093] 本文描述的变体可用于改变细胞的基因组;方法通常包括在细胞中表达这些变体蛋白,以及具有与细胞基因组的所选部分互补的区域的指导RNA。选择性地改变细胞基因组的方法在本领域中是已知的,参见例如,US 8,993,233;US 20140186958;US 9,023,649;WO/2014/099744;WO 2014/089290;WO 2014/144592;WO 144288;WO 2014/204578;WO 2014/152432;WO 2115/099850;US 8,697,359;US 20160024529;US 20160024524;US 20160024523;US 20160024510;US 20160017366;US 20160017301;US 20150376652;US 20150356239;US 20150315576;US 20150291965;US 20150252358;US 20150247150;US 20150232883;US 20150232882;US 20150203872;US 20150191744;US 20150184139;US 20150176064;US 20150167000;US 20150166969;US 20150159175;US 20150159174;US 20150093473;US 20150079681;US 20150067922;US 20150056629;US 20150044772;US

20150024500;US 20150024499;US 20150020223;US 20140356867;US 20140295557;US 20140273235;US 20140273226;US 20140273037;US 20140189896;US 20140113376;US 20140093941;US 20130330778;US 20130288251;US 20120088676;US 20110300538;US 20110236530;US 20110217739;US 20110002889;US 20100076057;US 20110189776;US 20110223638;US 20130130248;US 20150050699;US 20150071899;US 20150050699;US 20150045546;US 20150031134;US 20150024500;US 20140377868;US 20140357530;US 20140349400;US 20140335620;US 20140335063;US 20140315985;US 20140310830;US 20140310828;US 20140309487;US 20140304853;US 20140298547;US 20140295556;US 20140294773;US 20140287938;US 20140273234;US 20140273232;US 20140273231;US 20140273230;US 20140271987;US 20140256046;US 20140248702;US 20140242702;US 20140242700;US 20140242699;US 20140242664;US 20140234972;US 20140227787;US 20140212869;US 20140201857;US 20140199767;US 20140189896;US 20140186958;US 20140186919;US 20140186843;US 20140179770;US 20140179006;US 20140170753;WO/2008/108989;WO/2010/054108;WO/2012/164565;WO/2013/098244;WO/2013/176772;US 20150071899;Makarova等人,“Evolution and classification of the CRISPR-Cas system[CRISPR-Cas系统的进化与分类]”9(6)Nature Reviews Microbiology[自然微生物学评论]467-477(1-23)(2011年6月);Wiedenheft等人,“RNA-guided genetic silencing systems in bacteria and archaea[在细菌和古生菌中RNA指导的基因沉默系统]”482Nature[自然]331-338(2012年2月16日);Gasiunas等人,“Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria[Cas9-crRNA核糖核蛋白复合体介导细菌中适应免疫的特异性DNA切割]”109(39)Proceedings of the National Academy of Sciences USA[美国国家科学院院刊]E2579-E2586(2012年9月4日);Jinek等人,“A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity[适应性细菌免疫中的一种可编程的双RNA指导DNA内切核酸酶]”337Science[科学]816-821(2012年8月17日);Carroll,“A CRISPR Approach to Gene Targeting[基因靶向的CRISPR途径]”20(9)Molecular Therapy[分子治疗]1658-1660(2012年9月);在2012年5月25日提交的美国申请号61/652,086;Al-Attar等人,Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs):The Hallmark of an Ingenious Antiviral Defense Mechanism in Prokaryotes[规律间隔成簇短回文重复序列(CRISPR):原核生物中独特的抗病毒防御机制的特点],Biol Chem.[生物化学](2011)第392卷,第4期,第277-289页;Hale等人,Essential Features and Rational Design of CRISPR RNAs That Function With the Cas RAMP Module Complex to Cleave RNAs[利用Cas RAMP模块复合物来切割RNA的CRISPR RNA的基本特征和合理设计],Molecular Cell[分子细胞],(2012)第45卷,第3期,292-302。

[0094] 本文描述的变体蛋白可以用于代替以上参考文献中描述的Cas9蛋白中的任一种,或除了以上参考文献中描述的Cas9蛋白中的任一种,可以使用本文描述的变体蛋白,或可以将本文描述的变体蛋白与本文描述的突变组合使用。此外,本文描述的变体可以用于融合蛋白,代替本领域已知的野生型Cas9或其他Cas9突变(例如上述的dCas9或Cas9切口酶),

例如具有异源功能结构域的融合蛋白,如以下文献中所述的:US 8,993,233;US 20140186958;US 9,023,649;WO/2014/099744;WO 2014/089290;WO 2014/144592;WO 144288;WO 2014/204578;WO 2014/152432;WO 2115/099850;US 8,697,359;US 2010/0076057;US 2011/0189776;US 2011/0223638;US 2013/0130248;WO/2008/108989;WO/2010/054108;WO/2012/164565;WO/2013/098244;WO/2013/176772;US 20150050699;US 20150071899和WO 2014/124284。例如,优选包含一个或多个核酸酶减少、改变、或杀死突变的变体可以在Cas9的N或C末端融合到转录激活结构域或其他异源功能结构域(例如,转录阻遏因子(例如KRAB、ERD、SID等,例如ets2阻遏因子(ERF)阻遏因子结构域(ERD)的氨基酸473-530,KOX1的KRAB结构域的氨基酸1-97或Mad mSIN3相互作用结构域(SID)的氨基酸1-36;参见Beerli等人,PNAS USA[美国国家科学院院刊]95:14628-14633(1998))或沉默子如异染色质蛋白1(HP1,也称为swi6),例如,HP1 α 或HP1 β ;也可以使用本领域中已知的可以募集与固定的RNA结合序列(例如由MS2外壳蛋白、内切核糖核酸酶Csy4或 λ N蛋白结合的那些)融合的长的非编码RNA(lncRNA)的蛋白或肽;修饰DNA的甲基化状态的酶(例如DNA甲基转移酶(DNMT)或TET蛋白));或修饰组蛋白亚基的酶(例如组蛋白乙酰转移酶(HAT)、组蛋白脱乙酰酶(HDAC)、组蛋白甲基转移酶(例如,用于赖氨酸或精氨酸残基的甲基化)或组蛋白脱甲基酶(例如,用于赖氨酸或精氨酸残基的脱甲基化))。此类结构域的多个序列是本领域已知的,例如催化DNA中甲基化胞嘧啶的羟基化作用的结构域。示例性蛋白包括十-十一易位(TET)1-3家族,这是在DNA中转化5-甲基胞嘧啶(5-mC)为5-羟甲基胞嘧啶(5-hmC)的酶。

[0095] 人类TET1-3的序列是本领域中已知的,并且在下表中的示出:

基因	基因库登录号	
	氨基酸	核酸
TET1	NP_085128.2	NM_030625.2
TET2*	NP_001120680.1 (变体 1) NP_060098.3 (变体 2)	NM_001127208.2 NM_017628.4
TET3	NP_659430.1	NM_144993.1

[0097] *变体(1)表示较长转录物并且编码较长同种型(a)。与变体1相比,变体(2)差别在于5'UTR且在于3'UTR以及编码序列。与同种型a相比,所得同种型(b)更短并且具有不同C末端。

[0098] 在一些实施例中,可以包括催化结构域的全长序列的全部或一部分,例如包含富半胱氨酸延伸和由7个高度保守的外显子编码的20GFeD0结构域的催化模块,例如包含氨基酸1580-2052的Tet1催化结构域,包含氨基酸1290-1905的Tet2,和包含氨基酸966-1678的Tet3。参见例如,Iyer等人,Cell Cycle[细胞周期].2009年6月1日;8(11):1698-710的图1。电子出版物2009年6月27日针对例示了所有三种Tet蛋白中的关键催化残基的比对,及其针对全长序列的补充材料(在ftp站点ftp.ncbi.nih.gov/pub/aravind/DONS/supplementary_material_DONS.html可得)(参见例如seq 2c);在一些实施例中,该序列包括Tet1的氨基酸1418-2136或Tet2/3中的相应区域。

[0099] 其他催化模块可以来自Iyer等人,2009鉴定的蛋白。

[0100] 在一些实施例中,异源功能结构域是生物系链,并且包含MS2外壳蛋白、内切核糖核酸酶Csy4或 λ N蛋白(例如,DNA结合结构域)的全部或部分。这些蛋白可用于将包含特定茎环结构的RNA分子募集到由dCas9 gRNA靶向序列指定的位置。例如,与MS2外壳蛋白、内切核

糖核酸酶Csy4或 λ N融合的dCas9变体可用于募集长的非编码RNA(lncRNA)如XIST或HOTAIR;参见例如,Keryer-Bibens等人,Biol.Cell[生物细胞]100:125-138(2008),其与Csy4、MS2或 λ N结合序列连接。可替代地,Csy4、MS2或 λ N蛋白结合序列可以连接到另一种蛋白,例如,如Keryer-Bibens等人,同上所述的,并且可以使用本文描述的方法和组合物将蛋白靶向dCas9变体结合位点。在一些实施例中,Csy4是无催化活性的。在一些实施例中,Cas9变体(优选dCas9变体)与FokI融合,如以下文献中所述:US 8,993,233;US 20140186958;US 9,023,649;WO/2014/099744;WO 2014/089290;WO 2014/144592;WO 144288;WO 2014/204578;WO 2014/152432;WO 2115/099850;US 8,697,359;US 2010/0076057;US 2011/0189776;US 2011/0223638;US 2013/0130248;WO/2008/108989;WO/2010/054108;WO/2012/164565;WO/2013/098244;WO/2013/176772;US 20150050699;US 20150071899和WO 2014/204578。

[0101] 在一些实施例中,融合蛋白包括dCas9变体和异源功能域之间的接头。可以用于这些融合蛋白(或在连接的结构中的融合蛋白之间)的接头可以包括并不干扰融合蛋白的功能的任何序列。在优选的实施例中,该接头是短的,例如2-20个氨基酸,并且典型地是柔性的(即包含具有高自由度的氨基酸,例如甘氨酸、丙氨酸、和丝氨酸)。在一些实施例中,该接头包含由GGGS(SEQ ID NO:3)或GGGGS(SEQ ID NO:4)组成的一个或多个单元,例如GGGS(SEQ ID NO:5)或GGGGS(SEQ ID NO:6)单元的两个、三个、四个、或更多个重复。也可以使用其他的接头序列。

[0102] 在一些实施例中,变体蛋白包括协助递送至细胞内空间的细胞穿透肽,例如HIV源的TAT肽、穿膜肽、转运肽、或hCT源的细胞穿透肽,参见例如Caron等人,(2001)Mol Ther.[分子治疗]3(3):310-8;Langel,Cell-Penetrating Peptides:Processes and Applications[细胞穿透肽:过程和应用](CRC出版社(CRC Press),博卡拉顿市(Boca Raton)FL 2002);El-Andaloussi等人,(2005),Curr Pharm Des.[当前药物设计]11(28):3597-611;以及Deshayes等人,(2005)Cell Mol Life Sci.[细胞分子生命科学]62(16):1839-49。

[0103] 细胞穿透肽(CPP)是协助广泛范围的生物分子横跨细胞膜移动进入细胞质或其他细胞器,例如线粒体和细胞核的短肽。可以由CPP递送的分子的实例包括治疗药物、质粒DNA、寡核苷酸、siRNA、肽核酸(PNA)、蛋白、肽、纳米颗粒、和脂质体。CPP一般是30个氨基酸或更小,天然地起源或是非天然存在的蛋白或嵌合序列,并且包含高相对丰度的正电荷氨基酸,例如赖氨酸或精氨酸,抑或交替模式的极性和非极性氨基酸。通常用于本领域的CPP包括Tat(Frankel等人,(1988)Cell.[细胞]55:1189-1193,Vives等人,(1997)J.Biol.Chem.[生物化学杂志]272:16010-16017)、穿膜肽(Derossi等人,(1994)J.Biol.Chem.[生物化学杂志]269:10444-10450)、聚精氨酸肽序列(Wender等人,(2000)Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]97:13003-13008,Futaki等人,(2001)J.Biol.Chem.[生物化学杂志]276:5836-5840)、以及转运素(transportan)(Pooga等人,(1998)Nat.Biotechnol.[自然生物技术]16:857-861)。

[0104] 通过共价的或非共价的策略,CPP可以连接有其货物。用于共价接合CPP及其货物的方法是本领域已知的,例如化学交联(Stetsenko等人,(2000)J.Org.Chem.[生物化学杂志]65:4900-4909,Gait等人,(2003)Cell.Mol.Life.Sci.[细胞与分子生命科学]60:844-

853) 或克隆融合蛋白 (Nagahara等人, (1998) Nat. Med. [自然医学] 4:1449-1453)。通过电的和疏水的相互作用, 建立货物和包含极性的和非极性的结构域的短两亲CPP之间的非共价连接。

[0105] 本领域已经利用CPP来潜在地递送治疗性生物分子到细胞中。实例包括用于免疫抑制的连接至聚精氨酸的环孢霉素 (Rothbard等人, (2000) Nature Medicine [自然医学] 6 (11):1253-1257), 用于抑制肿瘤发生的、针对连接至称为MPG的CPP的细胞周期蛋白B1的 siRNA (Crombez等人, (2007) Biochem Soc. Trans. [生物化学学会汇报] 35:44-46), 用来减少癌细胞生长的、连接至CPP的肿瘤抑制基因p53肽 (Takenobu等人, (2002) Mol. Cancer Ther. [分子癌症治疗学] 1 (12):1043-1049, Snyder等人, (2004) PLoS Biol. [公共科学图书馆生物学] 2:E36), 和用来治疗哮喘的、Ras或融合至Tat的磷酸肌醇3激酶 (PI3K) 的显性负效形式 (Myou等人, (2003) J. Immunol. [免疫学杂志] 171:4399-4405)。

[0106] 本领域已经利用CPP来运输造影剂到细胞中, 用于成像和生物传感应用。例如, 附接至Tat的绿色荧光蛋白 (GFP) 已经用于标记癌细胞 (Shokolenko等人, (2005) DNA Repair [DNA修复] 4 (4):511-518)。缀合至量子点的Tat已经被用于成功横跨血脑屏障以便将大鼠大脑可视化 (Santra等人, (2005) Chem. Commun. [化学通讯] 3144-3146)。CPP也已经结合磁共振成像技术用于细胞成像 (Liu等人, (2006) Biochem. and Biophys. Res. Comm. [生物化学和生物物理研究通讯] 347 (1):133-140)。还参见Ramsey和Flynn, Pharmacol Ther. [药理学和治疗学] 2015年7月22日. pii:S0163-7258(15)00141-2。

[0107] 可替代地, 或此外, 变体蛋白可以包括核定位序列, 例如SV40大T抗原NLS (PKKKRRV (SEQ ID NO:7)) 和核质蛋白NLS (KRPAATKKAGQAKKKK (SEQ ID NO:8))。其他NLS是本领域已知的; 参见例如Cokol等人, EMBO Rep. [欧洲分子生物学组织报告] 2000年11月15日; 1 (5):411-415; Freitas和Cunha, Curr Genomics. [现代基因组学] 2009年12月; 10 (8):550-557。

[0108] 在一些实施例中, 变体包括对于配体, 例如GST、FLAG或六组氨酸序列具有高亲和力的部分。此类亲和标签可以协助纯化重组变体蛋白。

[0109] 对于其中将变体蛋白递送至细胞的方法, 可以使用本领域已知的任何方法产生蛋白, 例如通过体外翻译, 或在适合的宿主细胞中, 自编码变体蛋白的核酸的表达; 本领域已知用于产生蛋白的多种方法。例如, 可以在以下中产生蛋白并且然后纯化: 酵母、大肠杆菌、昆虫细胞系、植物、转基因动物、或培养的哺乳动物细胞; 参见例如Palomares等人, "Production of Recombinant Proteins: Challenges and Solutions [重组蛋白的产生: 挑战和解决方案]," Methods Mol Biol. [分子生物学方法] 2004; 267:15-52。此外, 可以将变体蛋白连接至协助转移进细胞的部分, 例如脂质纳米颗粒, 任选地使用一旦蛋白进入细胞, 被切割掉的接头。参见例如LaFontaine等人, Int J Pharm. [国际药物杂志] 2015年8月13日; 494 (1):180-194。

[0110] 表达系统

[0111] 为了使用本文描述的Cas9变体, 可能希望从编码它们的核酸表达它们。这可以按多种方式进行。例如, 可以将编码Cas9变体的核酸克隆进中间载体, 用于转化进原核细胞或真核细胞, 以便复制和/或表达。中间载体典型地是原核生物载体, 例如质粒, 或穿梭载体, 或昆虫载体, 用于编码Cas9变体的核酸的储存或操纵或用于生产Cas9变体。还可以将编码Cas9变体的核酸克隆进表达载体, 用于给予至植物细胞、动物细胞, 优选是哺乳动物细胞或

人类细胞,真菌细胞,细菌细胞,或原生动物细胞。

[0112] 为了获得表达,典型地将编码Cas9变体的序列亚克隆进包含用来指导转录的启动子的表达载体中。适合的细菌和真核启动子是本领域熟知的,并且例如在Sambrook等人, *Molecular Cloning, A Laboratory Manual* [分子克隆实验指南] (第3版,2001);Kriegler, *Gene Transfer and Expression: A Laboratory Manual* [基因转移和表达:实验室手册] (1990);和Current Protocols in Molecular Biology [分子生物学现代方法] (Ausubel等人编辑,2010)中描述。用于表达工程蛋白的细菌表达系统是可得,例如在大肠杆菌、芽孢杆菌属物种、和沙门氏菌属中 (Palva等人,1983, *Gene* [基因] 22:229-235)。用于此类表达系统的试剂盒是可商购的。用于哺乳动物细胞、酵母、和昆虫细胞的真核生物表达系统是本领域熟知的,并且也是可商购的。

[0113] 用于指导核酸的表达的启动子取决于具体应用。例如,典型地,强组成型启动子用于融合蛋白的表达和纯化。相比之下,当体内给予Cas9变体用于基因调节时,可以使用组成型启动子或诱导型启动子,取决于Cas9变体的具体用途。此外,用于给予Cas9变体的优选启动子可以是弱启动子,例如HSV TK,或具有类似活性的启动子。该启动子还可以包括响应转录激活的元件,例如缺氧应答元件、Gal4应答元件、lac阻遏因子应答元件、和小分子控制系统,例如四环素调节系统和RU-486系统 (参见例如, Gossen & Bujard, 1992, *Proc. Natl. Acad. Sci. USA* [美国国家科学院院刊], 89:5547; Oligino等人, 1998, *Gene Ther.* [基因治疗], 5:491-496; Wang等人, 1997, *Gene Ther.* [基因治疗], 4:432-441; Neering等人, 1996, *Blood* [血液], 88:1147-55; 和Rendahl等人, 1998, *Nat. Biotechnol.* [自然生物技术], 16:757-761)。

[0114] 除了启动子,表达载体典型地包含转录单位或表达盒,该转录单位或表达盒包含在宿主细胞(原核的或真核的)中表达核酸所需的所有另外的元件。因此,典型的表达盒包含有效地连接到,例如编码Cas9变体,和例如转录物的有效聚腺苷酸化、转录终止、核糖体结合位点、或翻译终止所需的任何信号的核酸序列上的启动子。盒的另外的元件可以包括,例如增强子,和异源剪接固有信号。

[0115] 关于Cas9变体的预期用途,选择用于运输遗传信息进入细胞的具体表达载体,预期用途例如是在植物、动物、细菌、真菌、原生动物等中表达。标准细菌表达载体包括质粒,例如基于pBR322的质粒,pSKF、pET23D、和可商购的标签融合表达系统,例如GST和LacZ。

[0116] 包含来自真核病毒的调节元件的表达载体通常用于真核表达载体,例如SV40载体、乳头瘤病毒载体、和源自埃-巴二氏病毒的载体。其他示例性真核载体包括pMSG、pAV009/A+、pMT010/A+、pMAMneo-5、杆状病毒pDSVE以及允许在以下启动子的指导下表达蛋白的任何其他载体:SV40早期启动子、SV40晚期启动子、金属硫蛋白启动子、鼠类乳腺肿瘤病毒启动子、劳斯肉瘤病毒启动子、多角体蛋白启动子,或示出对在真核细胞中的表达有效的其他启动子。

[0117] 用于表达Cas9变体的载体可以包括RNA Pol III启动子,以驱动指导RNA的表达,例如H1、U6或7SK启动子。这些人类启动子允许在质粒转染后,在哺乳动物细胞中,表达Cas9变体。

[0118] 一些表达系统具有用于选择稳定转染细胞系的标记,例如胸苷激酶、潮霉素B磷酸转移酶、和二氢叶酸还原酶。高产率表达系统也是适合的,例如在昆虫细胞中使用杆状病毒

载体,具有在多角体蛋白启动子或其他强杆状病毒启动子的指导下的gRNA编码序列。

[0119] 典型地包含在表达载体中的元件还包括在大肠杆菌中发挥功能的复制子、编码用来允许选择容纳重组质粒的细菌的抗生素抗性的基因、和用来允许重组序列的插入的在质粒的非必需区中的独特限制酶切位点。

[0120] 使用标准转染方法来产生表达大量蛋白的细菌的、哺乳动物的、酵母的或昆虫的细胞系,然后使用标准技术纯化这些蛋白(参见例如Colley等人,1989,J.Biol.Chem.[生物化学杂志],264:17619-22;对蛋白纯化的引导[Guide to Protein Purification],Methods in Enzymology[酶学方法],第182卷(Deutscher编辑,1990))。根据标准技术,进行真核细胞和原核细胞的转化(参见例如,Morrison,1977,J.Bacteriol.[细菌学杂志]132:349-351;Clark-Curtiss&Curtiss,Methods in Enzymology[酶学方法]101:347-362(Wu等人编辑,1983))。

[0121] 用于将外源核苷酸序列引入宿主细胞的任何已知程序都是可以使用的。这些包括使用磷酸钙转染、凝聚胺(polybrene)、原生质体融合、电穿孔、核转染、脂质体、微注射、裸DNA、质粒载体、病毒载体,附加型和整合型二者,和用于将克隆的基因组DNA、cDNA、合成DNA、或其他外源基因材料引入宿主细胞的任何其他熟知方法(参见例如Sambrook等人,同上)。只需要使用能够成功地将至少一个基因引入能够表达Cas9变体的宿主细胞中的具体遗传工程化程序。

[0122] 本发明的方法还可以包括通过将纯化的Cas9蛋白与gRNA一起引入细胞作为核糖核蛋白(RNP)复合物,连同引入gRNA加上编码Cas9的mRNA,来修饰gDNA。gRNA可以是合成的gRNA或编码指导RNA的核酸(例如在表达载体中)。

[0123] 本发明还包括载体和包含载体的细胞。

[0124] 实施例

[0125] 本发明在以下实例中进一步说明,这些实例不得限制权利要求书中所描述的本发明的范围。

[0126] 方法

[0127] 用于进化SpCas9变体的基于细菌的阳性选择测定

[0128] 用Cas9/sgRNA编码质粒转化包含阳性选择质粒(具有嵌入靶位点)的感受态大肠杆菌BW25141(γ DE3)²³。在SOB培养基中恢复60分钟后,将转化体铺板在包含氯霉素(非选择性)或氯霉素+10mM阿拉伯糖(选择性)的LB板上。

[0129] 为了鉴定对全基因组靶特异性可能至关重要的另外的位置,我们调整了先前用于研究归巢内切核酸酶性质的细菌选择系统(以下称为阳性选择)(Chen&Zhao,Nucleic Acids Res[核酸研究]33,e154(2005);Doyon等人,J Am Chem Soc[美国化学会志]128,2477-2484(2006))。

[0130] 在对该系统的本调整中,Cas9介导的编码诱导型毒性基因的阳性选择质粒的切割使得细胞能够存活,这归因于线性化质粒的随后降解和损失。在确定SpCas9可以在阳性选择系统中起作用之后,测试了野生型和变体切割含选择自己知人类基因组的靶位点的选择质粒的能力。将这些变体用包含靶位点的阳性选择质粒引入细菌中并铺板在选择性培养基上。通过计算存活频率:选择性板上的菌落/非选择性板上的菌落,来估计阳性选择质粒的切割(参见图1,5-6)。

[0131] 在此研究中使用的质粒的子集(序列显示如下)

名称	Addgene ID	描述
JDS246	43861	CMV-T7-人类 SpCas9-NLS-3xFLAG
VP12	未决定的	CMV-T7-人类 SpCas9-HF1 (N497A , R661A , Q695A, Q926A) -NLS-3xFLAG
MSP2135	未决定的	CMV-T7-人类 SpCas9-HF2 (N497A , R661A , Q695A, Q926A, D1135E) -NLS-3xFLAG
MSP2133	未决定的	CMV-T7-人类 SpCas9-HF4 (Y450A , N497A , R661A, Q695A, Q926A) -NLS-3xFLAG
[0132] MSP469	65771	CMV-T7-人类 SpCas9-VQR (D1135V , R1335Q , T1337R) -NLS-3xFLAG
MSP2440	未决定的	CMV-T7-人类 SpCas9-VQR-HF1 (N497A, R661A, Q695A, Q926A, D1135V, R1335Q, T1337R) -NLS-3xFLAG
BPK2797	未决定的	CMV-T7-人类 SpCas9-VRQR (D1135V, G1218R, R1335Q, T1337R) -NLS-3xFLAG
MSP2443	未决定的	CMV-T7-人类 SpCas9-VRQR-HF1 (N497A , R661A , Q695A , Q926A , D1135V , G1218R , R1335Q, T1337R) -NLS-3xFLAG
BPK1520	65777	U6-BsmBI 盒-Sp-sgRNA

[0133] 人类细胞培养和转染

[0134] 在37℃, 5%CO₂下, 将含组成型地表达的EGFP-PEST报道基因¹⁵的单一整合拷贝的U2OS. EGFP细胞在补充有10%FBS、2mM GlutaMax(生命技术公司(Life Technologies))、青霉素/链霉素和400μg/ml的G418的高级DMEM培养基(生命技术公司)中进行培养。根据制造商的方案, 用Lonza 4D-核转染仪的DN-100程序, 将细胞与750ng的Cas9质粒和250ng的sgRNA质粒(除非另有注解)共转染。将与空U6启动子质粒一起转染的Cas9质粒用作所有人类细胞实验的阴性对照。(参见图2, 7-10)。

[0135] 人类细胞EGFP破坏测定

[0136] 如先前所述¹⁶进行EGFP破坏实验。使用Fortessa流式细胞仪(BD生物科学公司(BD Biosciences)), 转染后约52小时分析转染细胞的EGFP表达。所有实验的背景EGFP损失门控为大约2.5% (参见图2, 7)。

[0137] T7E1测定、靶向深度测序和GUIDE-seq来定量核酸酶诱导的突变率

[0138] 如先前针对人类细胞所述进行T7E1测定(Kleinstiver, B.P. 等人, Nature[自然] 523, 481-485 (2015))。对于U2OS. EGFP人类细胞, 在转染后约72小时, 使用Agencourt DNAdvance基因组DNA分离试剂盒(贝克曼库尔特基因组学公司), 从转染的细胞中提取基因组DNA。将大约200ng纯化的PCR产物变性、退火并用T7E1(新英格兰生物实验室)消化。使用Qiaxcel毛细管电泳仪(凯杰公司(QIAGEN))定量诱变频率, 如先前针对人类细胞所述(Kleinstiver等人, Nature[自然] 523, 481-485 (2015); Reyon等人, Nat Biotechnol[自然

生物技术]30,460-465(2012))。

[0139] 如先前所述进行GUIDE-seq实验(Tsai等人,Nat Biotechnol[自然生物技术]33,187-197(2015))。简言之,如上所述,将磷酸化的硫代磷酸酯修饰的双链寡脱氧核苷酸(dsODN)与Cas9核酸酶连同Cas9和sgRNA表达质粒一起转染入U2OS细胞。进行dsODN特异性扩增、高通量测序和作图,以鉴定含有DSB活性的基因组区间。对于野生型与二重或四重突变体实验,脱靶读数计数被归一化为中靶读数计数,以校正样本之间的测序深度差异。然后比较野生型和变体SpCas9的归一化比率,以计算脱靶位点处活性的倍数变化。为了确定GUIDE-seq的野生型和SpCas9变体样本在预期的靶位点是否具有相似的寡核苷酸标签整合率,通过用Phusion Hot-Start Flex从100ng基因组DNA(如上所述分离)扩增预期靶座位来进行限制性片段长度多态性(RFLP)测定。在使用Agencourt Ampure XP试剂盒进行清理之前,在37℃,将大约150ng PCR产物用20U NdeI(新英格兰生物实验室)消化3小时。使用Qiaxcel毛细管电泳仪(凯杰公司)对RFLP结果进行定量,以粗略估计寡核苷酸标签整合率。如上所述,为了类似目的进行T7E1测定。

[0140] 实施例1

[0141] 用来解决CRISPR-Cas9 RNA指导基因编辑的靶向特异性的一个潜在解决方案将是用新颖的突变将Cas9变体工程化。

[0142] 基于这些更早的结果,得出假设(不希望受理论束缚):通过降低Cas9对DNA的非特异性结合亲和力(由结合至DNA上的磷酸基团或与DNA堆积相互作用的疏水碱基介导),能显著降低CRISPR-Cas9核酸酶的特异性。这一方法将具有以下优点:如在先前所述的截短的gRNA方法中,不减小由gRNA/Cas9复合物识别的靶位点的长度。推断可能通过将靶DNA上的磷酸基团接触的氨基酸残基突变,来减小Cas9对DNA的非特异性结合亲和力。

[0143] 已经使用类似方法来产生非Cas9核酸酶(例如TALEN)的变体(参见例如Guilinger等人,Nat.Methods.[自然方法]11:429(2014))。

[0144] 在该假设的初始测试中,本发明的诸位发明人尝试通过将单个丙氨酸取代引入可能预期与DNA主链上的磷酸相互作用的SpCas9中的不同残基,来将广泛使用的酿脓链球菌Cas9(SpCas9)的减小亲和力变体工程化。使用基于大肠杆菌的筛选测定,来评估这些变体的活性(Kleinstiver等人,Nature[自然].2015年7月23日;523(7561):481-5)。在此细菌系统中,细胞存活取决于包含针对毒性旋转酶毒物ccdB的基因和gRNA与SpCas9靶向的23个碱基对的序列的选择质粒的切割(和随后的破坏)。此实验鉴定的保留或损失活性的残基的结果(表1)。

[0145] 表1:Cas9的单丙氨酸取代突变体的活性

[0146] 如在图1中显示的基于细菌细胞的系统中评估的。

[0147]

突变	%存活		突变	%存活		突变	%存活
R63A	84.2		Q926A	53.3		K1158A	46.5
R66A	0		K1107A	47.4		K1185A	19.3
R70A	0		E1108A	40.0		K1200A	24.5
R74A	0		S1109A	96.6		S1216A	100.4
R78A	56.4		K1113A	51.8		Q1221A	98.8
R165A	68.9		R1114A	47.3		K1289A	55.2

R403A	85.2		S1116A	73.8		R1298A	28.6
N407A	97.2		K1118A	48.7		K1300A	59.8
N497A	72.6		D1135A	67.2		K1325A	52.3
K510A	79.0		S1136A	69.2		R1333A	0
Y515A	34.1		K1151A	0		K1334A	87.5
R661A	75.0		K1153A	76.6		R1335A	0
Q695A	69.8		K1155A	44.6		T1337A	64.6

[0148] 在50%-100%之间的存活百分率通常指示稳健的切割,而0%存活指示酶已经在功能上被损害。在细菌中测定的另外的突变(但是未在上表中显示)包括:R69A、R71A、Y72A、R75A、K76A、N77A、R115A、H160A、K163A、L169A、T404A、F405A、R447A、I448A、Y450A、S460A、M495A、M694A、H698A、Y1013A、V1015A、R1122A、K1123A、和K1124A。除了R69A和F405A(在细菌中具有<5%存活),所有这些另外的单突变似乎对SpCas9的中靶活性没有影响(在细菌筛选中>70%存活)。

[0149] 构建携带N497A、R661A、Q695A、和Q926A突变的所有可能的单重、二重、三重和四重组的15种不同的SpCas9变体,以便测试由这些残基造成的接触是否可能对于中靶活性而言是无效的(图1b)。对于这些实验而言,使用先前所述的基于人类细胞的测定,其中通过单一整合的EGFP报道基因内的非同源末端接合(NHEJ)介导的修复的插入或缺失突变(indel)的切割和诱导导致细胞荧光的损失(Reyon, D.等人, Nat Biotechnol. [自然生物技术] 30, 460-465, 2012)。当与野生型SpCas9配对时,使用先前在人类细胞中显示有效破坏EGFP表达的EGFP靶向的sgRNA(Fu, Y.等人, Nat Biotechnol [自然生物技术] 31, 822-826 (2013)),所有15种SpCas9变体拥有与野生型SpCas9可比较的EGFP破坏活性(图1b, 灰色条)。因此,这些残基中的一个或所有的取代并不减小具有此EGFP靶向的sgRNA的SpCas9的中靶切割效率。

[0150] 接下来,进行实验来评估所有15种SpCas9变体在错配的靶位点处的相对活性。为做到这一点,使用在位置13和14、15和16、17和18、以及18和19处包含取代的碱基对的、在先前实验中使用的EGFP靶向的sgRNA的衍生物,重复EGFP破坏测定(如下进行编号,对于PAM最近端的碱基,用1开始,并且对于PAM最远端的碱基,用20结束;图1b)。这些分析揭示,三重突变体(R661A/Q695A/Q926A)和四重突变体(N497A/R661A/Q695A/Q926A)之一都显示等价于使用所有四种错配的sgRNA的背景的EGFP破坏的水平(图1b, 有颜色条)。值得注意地,在15种变体中,拥有与错配的sgRNA的最低活性的那些都含Q695A和Q926A突变。基于这些结果和来自使用针对另一EGFP靶位点的sgRNA的实验的类似数据,选择四重突变体(N497A/R661A/Q695A/Q926A)用于另外的分析并且将其指定为SpCas9-HF1(用于高保真性变体#1)。

[0151] SpCas9-HF1的中靶活性

[0152] 为了确定在更大数量的中靶位点处, SpCas9-HF1能够多稳健地起作用,使用另外的sgRNA,在此变体和野生型SpCas9之间进行直接比较。总计,测试了37种不同的sgRNA: 24种靶向至EGFP(用EGFP破坏测定进行测定),并且13种靶向内源人类基因靶(使用T7内切核酸酶I (T7EI) 错配测定进行测定)。用EGFP破坏测定测试的24种sgRNA中的20种(图1c)和在內源人类基因位点上测试的13种sgRNA中的12种(图1d)显示SpCas9-HF1的活性,这些活性为野生型SpCas9与相同sgRNA之间的活性的至少70%(图1e)。事实上, SpCas9-HF1显示与野生型SpCas9相比,与绝大部分sgRNA之间的高度可比较的活性(90-140%) (图1e)。测试的37

种sgRNA中的三种显示与SpCas9-HF1之间基本上没有活性,并且这些靶位点的检查并不提示,与观察到高活性的那些相比,这些序列的特征中的任何明显差异(表3)。总体而言,对于测试的86%(32/37)的sgRNA,SpCas9-HF1拥有可比较的活性(大于70%的野生型SpCas9活性)。

[0153] 表3:sgRNA靶的清单

[0154] 酿脓链球菌sgRNA

EGFP							
制品名称	名称	间隔子长度 (nt)	间隔子序列	SEQ ID NO:	具有延伸的 PAM 的序列	SEQ ID NO:	
FYF13 20	NGG 位点 1	20	GGGCACGGGCA GCTTGCCGG	9.	GGGCACGGGC AGCTTGCCGG TGGT	10.	
[0155] FYF16 41	NGG 位点 1	18	GCACGGGCAGCT TGCCGG	11.	GCACGGGCAG CTTGCCGGT GT	12.	
CK101 2	NGG 位点 1-13&14	20	GGGCACccGCAG CTTGCCGG	13.	GGGCACccGCA GCTTGCCGGT GGT	14.	
FYF14 29	NGG 位点 1-15&16	20	GGGctgGGGCAG CTTGCCGG	15.	GGGctgGGGCAG GCTTGCCGGT GGT	16.	
FYF14	NGG 位点	20	GGcgACGGGCAG	17.	GGcgACGGGCAG	18.	

[0156]

30	1-17&18		CTTGCCGG		GCTTGCCGGT GGT
FYF13 47	NGG 位点 1-18&19	20	GccCACGGGCAG CTTGCCGG	19.	GccCACGGGCA GCTTGCCGGT GGT
BPK13 45	NGG 位点 2	20	GTCGCCCTCGAA CTTCACCT	21.	GTCGCCCTCG AACTTCACCT CGGC
BPK13 50	NGG 位点 3	20	GTAGGTCAGGGT GGTCACGA	23.	GTAGGTCAGG GTGGTCACGA GGGT
BPK13 53	NGG 位点 4	20	GGCGAGGGCGA TGCCACCTA	25.	GGCGAGGGCG ATGCCACCTA CGGC
MSP79 2	NGG 位点 5	20	GGTCGCCACCAT GGTGAGCA	27.	GGTCGCCACC ATGGTGAGCA AGGG
MSP79 5	NGG 位点 6	20	GGTCAGGGTGGT CACGAGGG	29.	GGTCAGGGTG GTCACGAGGG TGGG
FYF13 28	NGG 位点 7	20	GGTGGTGCAGAT GAACTTCA	31.	GGTGGTGCAG ATGAACTTCA GGGT
JAF10 01	NGG 位点 7	17	GGTGCAGATGAA CTTCA	33.	GGTGCAGATG AACTTCAGGG T
BPK13 65	NGG 位点 8	20	GTTGGGGTCTTT GCTCAGGG	35.	GTTGGGGTCTT TGCTCAGGGC GGA
MSP79 4	NGG 位点 9	20	GGTGGTCACGAG GGTGGGCC	37.	GGTGGTCACG AGGGTGGGCC AGGG
FYF13 27	NGG 位点 10	20	GATGCCGTTCTT CTGCTTGT	39.	GATGCCGTTCT TCTGCTTGT CG GC
JAF99 7	NGG 位点 10	17	GCCGTTCTTCTG CTTGT	41.	GCCGTTCTTCT GCTTGTC CGGC
BPK13 47	NGG 位点 11	20	GTCGCCACCATG GTGAGCAA	43.	GTCGCCACCA TGGTGAGCAA GGGC
BPK13 69	NGG 位点 12	20	GCACTGCACGCC GTAGGTCA	45.	GCACTGCACG CCGTAGGTCA GGGT
MSP25 45	NGG 位点 13	20	GTGAACCGCATC GAGCTGAA	47.	GTGAACCGCA TCGAGCTGAA

[0157]	MSP25 46	NGG 位点 14	20	GAAGGGCATCG ACTTCAAGG	49.	GGGC GAAGGGCATC GACTTCAAGG AGGA	50.
	MSP25 47	NGG 位点 15	20	GCTTCATGTGGT CGGGGTAG	51.	GCTTCATGTG GTCGGGGTAG CGGC	52.
	MSP25 48	NGG 位点 16	20	GCTGAAGCACTG CACGCCGT	53.	GCTGAAGCAC TGCACGCCGT AGGT	54.
	MSP25 49	NGG 位点 17	20	GCCGTCGTCCTT GAAGAAGA	55.	GCCGTCGTCCT TGAAGAAGAT GGT	56.
	MSP25 50	NGG 位点 18	20	GACCAGGATGG GCACCACCC	57.	GACCAGGATG GGCACCACCC CGGT	58.
	MSP25 51	NGG 位点 19	20	GACGTAGCCTTC GGGCATGG	59.	GACGTAGCCT TCGGGCATGG CGGA	60.
	MSP25 53	NGG 位点 20	20	GAAGTTCGAGGG CGACACCC	61.	GAAGTTCGAG GGCGACACCC TGGT	62.
	MSP25 54	NGG 位点 21	20	GAGCTGGACGGC GACGTAAA	63.	GAGCTGGACG GCGACGTAAA CGGC	64.
	MSP25 55	NGG 位点 22	20	GGCATCGCCCTC GCCCTCGC	65.	GGCATCGCCC TCGCCCTCGC CGGA	66.
	MSP25 56	NGG 位点 23	20	GGCCACAAGTTC AGCGTGTC	67.	GGCCACAAGT TCAGCGTGTC CGGC	68.
	FYF13 31	NGG 位点 24	20	GGGCGAGGAGC TGTTACCG	69.	GGGCGAGGAG CTGTTACCG GGGT	70.
	FYF15 60	NGG 位点 24	18	GCGAGGAGCTGT TCACCG	71.	GCGAGGAGCT GTTACCGGG GT	72.
	BPK13 48	NGG 位点 25-无 5' G	20	CCTCGAACTTCA CCTCGGCG	73.	CCTCGAACTTC ACCTCGGCGC GGG	74.
	BPK13 49	NGG 位点 25-mm 5' G	20	GCTCGAACTTCA CCTCGGCG	75.	GCTCGAACTT CACCTCGGCG CGGG	76.
	BPK13 51	NGG 位点	20	CAACTACAAGAC CCGCGCCG	77.	CAACTACAAG ACCCGCGCCG	78.

[0158]

	26-无 5' G				AGGT	
BPK13 52	NGG 位点 26-mm 5' G	20	GAACTACAAGAC CCGCGCCG	79.	GAACTACAAG ACCCGCGCCG AGGT	80.
BPK13 73	NGG 位点 27-无 5' G	20	CGCTCCTGGACG TAGCCTTC	81.	CGCTCCTGGA CGTAGCCTTC GGGC	82.
BPK13 75	NGG 位点 27-mm 5' G	20	GGCTCCTGGACG TAGCCTTC	83.	CGCTCCTGGA CGTAGCCTTC GGGC	84.
BPK13 77	NGG 位点 28-无 5' G	20	AGGGCGAGGAG CTGTTCACC	85.	AGGGCGAGGA GCTGTTCACC GGGG	86.
BPK13 61	NGG 位点 28-mm 5' G	20	GGGGCGAGGAG CTGTTCACC	87.	GGGGCGAGGA GCTGTTCACC GGGG	88.
BPK14 68	NGAA 位 点 1	20	GTTTCGAGGGCGA CACCCTGG	89.	GTTTCGAGGGC GACACCCTGG TGAA	90.
MSP80 7	NGAA 位 点 2	20	GTTTCACCAGGGT GTCGCCCT	91.	GTTTCACCAGG GTGTCGCCCTC GAA	92.
MSP17 0	NGAC 位 点 1	20	GCCCACCCTCGT GACCACCC	93.	GCCCACCCTC GTGACCACCC TGAC	94.
MSP79 0	NGAC 位 点 2	20	GCCCTTGCTCAC CATGGTGG	95.	GCCCTTGCTCA CCATGGTGGC GAC	96.
MSP17 1	NGAT 位 点 1	20	GTCGCCGTCCAG CTCGACCA	97.	GTCGCCGTCC AGCTCGACCA GGAT	98.
MSP16 9	NGAT 位 点 2	20	GTGTCCGGCGAG GGCGAGGG	99.	GTGTCCGGCG AGGGCGAGGG CGAT	100.
MSP16 8	NGAG 位 点 1	20	GGGGTGGTGCCC ATCCTGGT	101.	GGGGTGGTGC CCATCCTGGTC GAG	102.
MSP36 6	NGAG 位 点 2	20	GCCACCATGGTG AGCAAGGG	103.	GCCACCATGG TGAGCAAGGG CGAG	104.

内源基因***EMXI***

制品名	名称	间隔子	间隔子序列	SEQ ID	具有延伸的 PAM	SEQ ID NO:
-----	----	-----	-------	--------	-----------	------------

称		长度 (nt)	NO: 的序列	
FYF15 48	NGG 位 点 1	20	GAGTCCGAGCA GAAGAAGAA	105. GAGTCCGAGCAG AAGAAGAAGGG C 106.
MSP80 9	NGG 位 点 2	20	GTCACCTCCAA TGACTAGGG	107. GTCACCTCCAAT GACTAGGGTGGG 108.
VC475	NGG 位 点 3	20	GGGAAGACTGA GGCTACATA	109. GGGAAGACTGA GGCTACATAGGG T 110.
MSP81 4 *1	NGA 位 点 1	20	GCCACGAAGCA GGCCAATGG	111. GCCACGAAGCAG GCCAATGGGGAG 112.
<i>FANCF</i>				
制品名 称	名称	间隔子 长度 (nt)	间隔子序列	SEQ ID NO: 具有延伸的 PAM 的序列 SEQ ID NO:
DR348	NGG 位 点 1	20	GGAATCCCTTC TGCAGCACC	113. GGAATCCCTTCT GCAGCACCTGGA 114.
[0159] MSP81 5	NGG 位 点 2	20	GCTGCAGAAGG GATTCCATG	115. GCTGCAGAAGGG ATTCCATGAGGT 116.
MSP81 6	NGG 位 点 3	20	GGCGGCTGCAC AACCAGTGG	117. GGCGGCTGCACA ACCAGTGGAGGC 118.
MSP81 7	NGG 位 点 4	20	GCTCCAGAGCC GTGCGAATG	119. GCTCCAGAGCCG TGCGAATGGGGC 120.
MSP81 8 *2	NGA 位 点 1	20	GAATCCCTTCT GCAGCACCT	121. GAATCCCTTCTG CAGCACCTGGAT 122.
MSP82 0 *3	NGA 位 点 2	20	GCGGCGGCTGC ACAACCAGT	123. GCGGCGGCTGCA CAACCAGTGGAG 124.
MSP88 5 *4	NGA 位 点 3	20	GGTTGTGCAGC CGCCGCTCC	125. GGTTGTGCAGCC GCCGCTCCAGAG 126.
<i>RUNXI</i>				
制品名 称	名称	间隔子 长度 (nt)	间隔子序列	SEQ ID NO: 具有延伸的 PAM 的序列 SEQ ID NO:
MSP82 2	NGG 位 点 1	20	GCATTTTCAGG AGGAAGCGA	127. GCATTTTCAGGA GGAAGCGATGGC 128.

[0160]

MSP82 5	NGG 位 点 2	20	GGGAGAAGAA AGAGAGATGT	129.	GGGAGAAGAAA GAGAGATGTAGG G	130.
MSP82 6 *5	NGA 位 点 1	20	GGTGCATTTTC AGGAGGAAG	131.	GGTGCATTTTCA GGAGGAAGCGA T	132.
MSP82 8 *6	NGA 位 点 2	20	GAGATGTAGGG CTAGAGGGG	133.	GAGATGTAGGGC TAGAGGGGTGAG	134.
MSP17 25	NGAA 位点 1	20	GGTATCCAGCA GAGGGGAGA	135.	GGTATCCAGCAG AGGGGAGAAGA A	136.
MSP17 26	NGAA 位点 2	20	GAGGCATCTCT GCACCGAGG	137.	GAGGCATCTCTG CACCGAGGTGAA	138.
MSP17 28	NGAC 位 点 1	20	GAGGGGTGAGG CTGAAACAG	139.	GAGGGGTGAGG CTGAAACAGTGA C	140.
MSP17 30	NGAC 位 点 2	20	GAGCAAAAGTA GATATTACA	141.	GAGCAAAAGTA GATATTACAAGA C	142.
MSP17 32	NGAT 位 点 1	20	GGAATTCAAAC TGAGGCATA	143.	GGAATTCAAAC GAGGCATATGAT	144.
MSP82 9	NGAT 位 点 2	20	GCAGAGGGGA GAAGAAAGAG	145.	GCAGAGGGGAG AAGAAAGAGAG AT	146.
MSP17 34	NGAG 位点 1	20	GCACCGAGGCA TCTCTGCAC	147.	GCACCGAGGCAT CTCTGCACCGAG	148.
MSP82 8	NGAG 位点 2	20	GAGATGTAGGG CTAGAGGGG	149.	GAGATGTAGGGC TAGAGGGGTGAG	150.

ZSCAN2

制品名 称	名称	间隔子 长度 (nt)	间隔子序列	SEQ ID NO:	具有延伸的 PAM 的序列	SEQ ID NO:
NN675	NGG 位 点	20	GTGCGGCAAGA GCTTCAGCC	151.	GTGCGGCAAGAG CTTCAGCCGGGG	152.

VEGFA

制品名 称	名称	间隔子 长度 (nt)	间隔子序列	SEQ ID NO:	具有延伸的 PAM 的序列	SEQ ID NO:
VC297	NGG 位	20	GGGTGGGGGGA	153.	GGGTGGGGGGA	154.

[0161]	点 1			GTTTGCTCC		GTTTGCTCCTGG A	
	VC299	NGG 位 点 2	20	GACCCCTCCA CCCCGCTC	155.	GACCCCTCCAC CCCGCTCCGGG	156.
	VC228	NGG 位 点 3	20	GGTGAGTGAGT GTGTGCGTG	157.	GGTGAGTGAGTG TGTGCGTGTGGG	158.
	BPK18 46 *7	NGA 位 点 1	20	GCGAGCAGCGT CTTCGAGAG	159.	GCGAGCAGCGTC TTCGAGAGTGAG	160.
	ZNF629						
	制品名 称	名称	间隔子 长度 (nt)	间隔子序列	SEQ ID NO:	具有延伸的 PAM 的序列	SEQ ID NO:
	NN675 *8	NGA 位 点	20	GTGCGGCAAGA GCTTCAGCC	161.	GTGCGGCAAGAG CTTCAGCCAGAG	162.

[0162] *1, NGA EMX1位点4, 来自Kleinstiver等人, Nature[自然]2015

[0163] *2, NGA FANCF位点1, 来自Kleinstiver等人, Nature[自然]2015

[0164] *3, NGA FANCF位点3, 来自Kleinstiver等人, Nature[自然]2015

[0165] *4, NGA FANCF位点4, 来自Kleinstiver等人, Nature[自然]2015

[0166] *5, NGA RUNX1位点1, 来自Kleinstiver等人, Nature[自然]2015

[0167] *6, NGA RUNX1位点3, 来自Kleinstiver等人, Nature[自然]2015

[0168] *7, NGA VEGFA位点1, 来自Kleinstiver等人, Nature[自然]2015

[0169] *8, NGA ZNF629位点, 来自Kleinstiver等人, Nature[自然]2015

[0170] SpCas9-HF1的全基因组特异性

[0171] 为了测试在人类细胞中, SpCas9-HF1是否表现出减小的脱靶效应, 使用通过测序(GUIDE-seq)方法实现的双链断裂的全基因组无偏鉴定。GUIDE-seq使用进入双链断裂的短双链寡脱氧核苷酸(dsODN)标签的整合, 以便实现相邻基因组序列的扩增和测序, 其中在任何给定位点的标签整合的数目提供切割效率的定量测量(Tsai, S.Q. 等人, Nat Biotechnol [自然生物技术] 33, 187-197 (2015))。采用靶向内源人类EMX1、FANCF、RUNX1、和ZSCAN2基因中的不同位点的八种不同的sgRNA, 使用GUIDE-seq来比较由野生型SpCas9和SpCas9-HF1诱导的脱靶效应的图谱。由这些sgRNA靶向的序列是独特的, 并且在参比人类基因组中具有可变量数的预测的错配位点(表2)。对于八种sgRNA, 中靶dsODN标签整合(通过限制性片段长度多态性(RFLP)测定)和indel形成(通过T7EI测定)的评估揭示与野生型SpCas9和SpCas9-HF1的可比较的中靶活性(分别为图7a和7b)。GUIDE-seq实验显示, 八种sgRNA中的七种诱导使用野生型SpCas9在多个全基因组脱靶位点(每个sgRNA范围为从2至25)处的切割, 而第八种sgRNA(对于FANCF位点4)并不产生任何可检测的脱靶位点(图2a和2b)。然而, 诱导与野生型SpCas9的indel的七种sgRNA中的六种显示与SpCas9-HF1的GUIDE-seq可检测脱靶事件的引人注目的完全不存在(图2a和2b); 并且剩余的第七种sgRNA(对于FANCF位点2)诱导在含原型间隔子种子序列内的一个错配的位点处的唯一一个单个可检测全基因组脱靶切割事件(图2a)。总而言之, 当使用SpCas9-HF1时未检测到的脱靶位点含原型间隔子和/或PAM序

列中的一至六个错配(图2c)。当使用SpCas9-HF1进行测试时,如与野生型SpCas9一起,第八种sgRNA(对于FANCF位点4)并未产生可检测的脱靶切割事件(图2a)。

[0172] 为了确认GUIDE-seq发现,将靶向扩增子测序用于更直接地测量由野生型SpCas9和SpCas9-HF1诱导的NHEJ介导的indel突变的频率。对于这些实验,仅用sgRNA-和Cas9编码质粒(即,无GUIDE-seq标签)转染人类细胞。然后将下一代测序用于检查对于GUIDE-seq实验中的六种sgRNA,已经用野生型SpCas9鉴定的40个脱靶位点中的36个(40个位点中的四个不能检查,因为它们不能从基因组DNA中特异性地扩增)。这些深度测序实验显示:(1)在六个sgRNA中靶位点中的每一个处,野生型SpCas9和SpCas9-HF1诱导可比较频率的indel(图3a和3b);(2)如预期,野生型SpCas9显示,在36个脱靶位点中的35个处,indel突变的统计上显著的证据(图3b),其频率与对于这些相同位点的GUIDE-seq读数计数很好地关联(图3c);以及(3)在36个脱靶位点中的34个处,由SpCas9-HF1诱导的indel的频率与在来自对照转染的样品中观察到的indel的背景水平不可区分(图3b)。对于相对于阴性对照,使用SpCas9-HF1似乎具有统计上显著的突变频率的两个脱靶位点,indel的平均频率为0.049%和0.037%,在这个水平难以确定这些是否是由于测序/PCR错误或是真实的核酸酶诱导的indel导致的。基于这些结果,结论是SpCas9-HF1可以完全地或接近完全地减少横跨使用野生型SpCas9的一系列不同频率发生的脱靶脱变至不可检测的水平。

[0173] 接下来,评估了SpCas9-HF1减少靶向非典型均聚的或重复的序列的sgRNA的全基因组脱靶效应的能力。尽管很多人现在试图避免具有这些特征的中靶位点由于其相对缺乏与基因组的正交性,仍希望探索SpCas9-HF1是否可能减少脱靶indel(甚至对于这些挑战性的靶)。因此,使用先前表征的sgRNA(Fu,Y.等人,Nat Biotechnol[自然生物技术]31,Tsai,S.Q.等人,Nat Biotechnol[自然生物技术]33,187-197(2015)),它们靶向在人类VEGFA基因中的富含胞嘧啶的均聚序列抑或包含多个TG重复的序列(分别是VEGFA位点2和VEGFA位点3)(表2)。在对照实验中,使用野生型SpCas9和SpCas9-HF1二者,这些sgRNA中的每一个诱导可比较水平的GUIDE-seq ds ODN标签合并(图7c)和indel突变(图7d),展示在与这些sgRNA中的任一者的中靶活性方面,SpCas9-HF1并未受损。重要的是,GUIDE-seq实验揭示,在减少这些sgRNA的脱靶位点方面,SpCas9-HF1高度有效,其中对于VEGFA位点2,123/144个位点未检测到,对于VEGFA位点3,31/32个位点未检测到(图4a和4b)。用SpCas9-HF1未检测到的这些脱靶位点的检查显示,它们各自在其原型间隔子和PAM序列内拥有一系列的总错配:对于VEGFA位点2sgRNA,是2至7个错配,并且对于VEGFA位点3sgRNA,是1至4个错配(图4c);而且,对于VEGFA位点2的这些脱靶位点中的九个可以在sgRNA-DNA界面具有潜在的凸起碱基(Lin,Y.等人,.Nucleic Acids Res[核酸研究]42,7473-7485(2014))(图4a和图8)。用SpCas9-HF1未检测到的位点拥有针对VEGFA位点2sgRNA的2至6个错配,和在针对VEGFA位点3sgRNA的单个位点中的2个错配(图4c),其中针对VEGFA位点2sgRNA的三个脱靶位点再次具有潜在的凸起(图8)。总而言之,这些结果展示SpCas9-HF1在减小靶向至简单重复序列的sgRNA的脱靶效应方面,可以是高度有效的,并且还可以具有对靶向至均聚序列的sgRNA的实质性影响。

[0174] 表2|针对由GUIDE-seq检查的十种sgRNA,在参比人类基因组中的潜在错配位点的总结

[0175]

位点	具有PAM的间隔子	与中靶位点的错配*						总计
		1	2	3	4	5	6	
EMX1-1	GAGTCCGAGCAGAAGAAGAAGGG (SEQ ID NO: 163)	0	1	18	273	2318	15831	18441
EMX1-2	GTCACCTCCAATGACTAGGGTGG (SEQ ID NO: 164)	0	0	3	68	780	6102	6953
FANCF-1	GGAATCCCTTCTGCAGCACCTGG (SEQ ID NO: 165)	0	1	18	288	1475	9611	11393
FANCF-2	GCTGCAGAAGGGATTCCATGAGG (SEQ ID NO: 166)	1	1	29	235	2000	13047	15313
FANCF-3	GGCGGCTGCACAACCAGTGGAGG (SEQ ID NO: 167)	0	0	11	79	874	6651	7615
FANCF-4	GCTCCAGAGCCGTGCGAATGGGG (SEQ ID NO: 168)	0	0	6	59	639	5078	5782
RUNX1-1	GCATTTTCAGGAGGAAGCGATGG (SEQ ID NO: 169)	0	2	6	189	1644	11546	13387
ZSCAN2	GTGCGGCAAGAGCTTCAGCCGGG (SEQ ID NO: 170)	0	3	12	127	1146	10687	11975
VEGFA2	GACCCCTCCACCCCGCCTCCGG (SEQ ID NO: 171)	0	2	35	456	3905	17576	21974
VEGFA3	GGTGAGTGAGTGTGTGCGTGTGG (SEQ ID NO: 172)	1	17	383	6089	13536	35901	55927

[0176] *使用Cas-OFFinder确定 (Bae等人,Bioinformatics[生物信息学]30,1473-1475 (2014))

[0177] 表4:研究中使用的寡核苷酸

	<u>T7E1 引物的描述</u>	<u>序列</u>	<u>SEQ ID NO:</u>
	在 U2OS 人类细胞中用来扩增 EMX1 的正向引物	GGAGCAGCTGGTCAGA GGGG	173.
	在 U2OS 人类细胞中用来扩增 EMX1 的反向引物	CCATAGGGAAGGGGG ACACTGG	174.
	在 U2OS 人类细胞中用来扩增 FANCF 的正向引物	GGGCCGGGAAAGAGT TGCTG	175.
	在 U2OS 人类细胞中用来扩增 FANCF 的反向引物	GCCCTACATCTGCTCT CCCTCC	176.
	在 U2OS 人类细胞中用来扩增 RUNX1 的正向引物	CCAGCACAACCTTACTC GCACTTGAC	177.
	在 U2OS 人类细胞中用来扩增 RUNX1 的反向引物	CATCACCAACCCACAG CCAAGG	178.
	在 U2OS 人类细胞中用来扩增 VEGFA 的正向引物	TCCAGATGGCACATTG TCAG	179.
	在 U2OS 人类细胞中用来扩增 VEGFA 的反向引物	AGGGAGCAGGAAAGT GAGGT	180.
[0178]	在 U2OS 人类细胞中用来扩增 VEGFA (NGG 位点 2) 的正向引物	CGAGGAAGAGAGAGA CGGGGTC	181.
	在 U2OS 人类细胞中用来扩增 VEGFA (NGG 位点 2) 的反向引物	CTCCAATGCACCCAAG ACAGCAG	182.
	在 U2OS 人类细胞中用来扩增 ZSCAN2 的正向引物	AGTGTGGGGTGTGTGG GAAG	183.
	在 U2OS 人类细胞中用来扩增 ZSCAN2 的反向引物	GCAAGGGGAAGACTCT GGCA	184.
	在 U2OS 人类细胞中用来扩增 ZNF629 的正向引物	TACGAGTGCCTAGAGT GCG	185.
	在 U2OS 人类细胞中用来扩增 ZNF629 的反向引物	GCAGATGTAGGTCTTG GAGGAC	186.
	<u>深度测序引物的描述</u>	<u>序列</u>	<u>SEQ ID NO:</u>
	用来扩增中靶的 EMX1-1 的正向引物	GGAGCAGCTGGTCAGA GGGG	187.
	用来扩增中靶的 EMX1-1 的反向引物	CGATGTCCTCCCCATT	188.

	GGCCTG	
用来扩增 EMX1-1-GUIDE_seq-OT#1 的 正向引物	GTGGGGAGATTTGCAT CTGTGGAGG	189.
用来扩增 EMX1-1-GUIDE_seq-OT#1 的 反向引物	GCTTTTATACCATCTTG GGGTTACAG	190.
用来扩增 EMX1-1-GUIDE_seq-OT#2 的 正向引物	CAATGTGCTTCAACCC ATCACGGC	191.
用来扩增 EMX1-1-GUIDE_seq-OT#2 的 反向引物	CCATGAATTTGTGATG GATGCAGTCTG	192.
用来扩增 EMX1-1-GUIDE_seq-OT#3 的 正向引物	GAGAAGGAGGTGCAG GAGCTAGAC	193.
用来扩增 EMX1-1-GUIDE_seq-OT#3 的 反向引物	CATCCCGACCTTCATC CCTCCTGG	194.
用来扩增 EMX1-1-GUIDE_seq-OT#4 的 正向引物	GTAGTTCTGACATTCC TCCTGAGGG	195.
用来扩增 EMX1-1-GUIDE_seq-OT#4 的 反向引物	TCAAACAAGGTGCAGA TACAGCA	196.
用来扩增 EMX1-1-GUIDE_seq-OT#5 的 正向引物	CAGGGTCGCTCAGTCT GTGTGG	197.
用来扩增 EMX1-1-GUIDE_seq-OT#5 的 反向引物	CCAGCGCACCATTAC TCCACCTG	198.
用来扩增 EMX1-1-GUIDE_seq-OT#6 的 正向引物	GGCTGAAGAGGAAGA CCAGACTCAG	199.
用来扩增 EMX1-1-GUIDE_seq-OT#6 的 反向引物	GGCCCCTCTGAATTCA ATTCTCTGC	200.
用来扩增 EMX1-1-GUIDE_seq-OT#7 的 正向引物	CCACAGCGAGGAGTG ACAGCC	201.
用来扩增 EMX1-1-GUIDE_seq-OT#7 的 反向引物	CCAAGTCTTTCCTAAC TCGACCTTGG	202.
用来扩增 EMX1-1-GUIDE_seq-OT#8 的 正向引物	CCCTAGGCCCCACACCA GCAATG	203.
用来扩增 EMX1-1-GUIDE_seq-OT#8 的 反向引物	GGGATGGGAATGGGA ATGTGAGGC	204.
用来扩增中靶的 EMX1-2 的正向引物	GCCCAGGTGAAGGTGT GGTTC	205.
用来扩增中靶的 EMX1-2 的反向引物	CCAAAGCCTGGCCAGG GAGTG	206.

[0179]

[0180]	用来扩增 EMX1-2-GUIDE_seq-OT#1 的正向引物	AGGCAAAGATCTAGG ACCTGGATGG	207.
	用来扩增 EMX1-2-GUIDE_seq-OT#1 的反向引物	CCATCTGAGTCAGCCA GCCTTGTC	208.
	用来扩增 EMX1-2-GUIDE_seq-OT#2 的正向引物	GGTTCCCTCCCTTCTG AGCCC	209.
	用来扩增 EMX1-2-GUIDE_seq-OT#2 的反向引物	GGATAGGAATGAAGA CCCCCTCTCC	210.
	用来扩增 EMX1-2-GUIDE_seq-OT#3 的正向引物	GGACTGGCTGGCTGTG TGTTTTGAG	211.
	用来扩增 EMX1-2-GUIDE_seq-OT#3 的反向引物	CTTATCCAGGGCTACC TCATTGCC	212.
	用来扩增 EMX1-2-GUIDE_seq-OT#4 的正向引物	GCTGCTGCTGCTTTGA TCACTCCTG	213.
	用来扩增 EMX1-2-GUIDE_seq-OT#4 的反向引物	CTCCTTAAACCCTCAG AAGCTGGC	214.
	用来扩增 EMX1-2-GUIDE_seq-OT#5 的正向引物	GCACTGTCAGCTGATC CTACAGG	215.
	用来扩增 EMX1-2-GUIDE_seq-OT#5 的反向引物	ACGTTGGAACAGTCGA GCTGTAGC	216.
	用来扩增 EMX1-2-GUIDE_seq-OT#6 的正向引物	TGTGCATAACTCATGT TGGCAAAT	217.
	用来扩增 EMX1-2-GUIDE_seq-OT#6 的反向引物	TCCACAACCTACCCTCA GCTGGAG	218.
	用来扩增 EMX1-2-GUIDE_seq-OT#7 的正向引物	CCACTGACAATTCACT CAACCCTGC	219.
	用来扩增 EMX1-2-GUIDE_seq-OT#7 的反向引物	AGGCAGACCAGTTATT TGGCAGTC	220.
	用来扩增 EMX1-2-GUIDE_seq-OT#9 的正向引物	ACAGGCGCAGTTCACT GAGAAG	221.
	用来扩增 EMX1-2-GUIDE_seq-OT#9 的反向引物	GGGTAGGCTGACTTTG GGCTCC	222.
	用来扩增中靶的 FANCF-1 的正向引物	GCCCTCTTGCCTCCAC TGGTTG	223.
	用来扩增中靶的 FANCF-1 的反向引物	CGCGGATGTTCCAATC AGTACGC	224.
	用来扩增 FANCF-1-GUIDE_seq-OT#1	GCGGGCAGTGGCGTCT	225.

	的正向引物	TAGTCG	
	用来扩增 FANCF-1-GUIDE_seq-OT#1	CCCTGGGTTTGGTTGG	226.
	的反向引物	CTGCTC	
	用来扩增 FANCF-1-GUIDE_seq-OT#2	CTCCTTGCCGCCCAGC	227.
	的正向引物	CGGTC	
	用来扩增 FANCF-1-GUIDE_seq-OT#2	CACTGGGGAAGAGGC	228.
	的反向引物	GAGGACAC	
	用来扩增 FANCF-1-GUIDE_seq-OT#3	CCAGTGTTTCCCATCC	229.
	的正向引物	CCAACAC	
	用来扩增 FANCF-1-GUIDE_seq-OT#3	GAATGGATCCCCCCT	230.
	的反向引物	AGAGCTC	
	用来扩增 FANCF-1-GUIDE_seq-OT#4	CAGGCCACAGGTCCT	231.
	的正向引物	TCTGGA	
	用来扩增 FANCF-1-GUIDE_seq-OT#4	CCACACGGAAGGCTGA	232.
	的反向引物	CCACG	
		GCGCAGAGAGAGCAG	233.
	用来扩增中靶的 FANCF-3 的正向引物	GACGTC	
		GCACCTCATGGAATCC	234.
	用来扩增中靶的 FANCF-3 的反向引物	CTTCTGC	
[0181]	用来扩增 FANCF-3-GUIDE_seq-OT#1	CAAGTGATGCGACTTC	235.
	的正向引物	CAACCTC	
	用来扩增 FANCF-3-GUIDE_seq-OT#1	CCCTCAGAGTTCAGCT	236.
	的反向引物	TAAAAAGACC	
	用来扩增 FANCF-3-GUIDE_seq-OT#2	TGCTTCTCATCCACTCT	237.
	的正向引物	AGACTGCT	
	用来扩增 FANCF-3-GUIDE_seq-OT#2	CACCAACCAGCCATGT	238.
	的反向引物	GCCATG	
	用来扩增 FANCF-3-GUIDE_seq-OT#3	CTGCCTGTGCTCCTCG	239.
	的正向引物	ATGGTG	
	用来扩增 FANCF-3-GUIDE_seq-OT#3	GGGTTCAAAGCTCATC	240.
	的反向引物	TGCCCC	
	用来扩增 FANCF-3-GUIDE_seq-OT#4	GCATGTGCCTTGAGAT	241.
	的正向引物	TGCCTGG	
	用来扩增 FANCF-3-GUIDE_seq-OT#4	GACATTCAGAGAAGCG	242.
	的反向引物	ACCATGTGG	
	用来扩增 FANCF-3-GUIDE_seq-OT#5	CCATCTTCCCCTTTGGC	243.
	的正向引物	CCACAG	

[0182]	用来扩增 FANCF-3-GUIDE_seq-OT#5 的反向引物	CCCCAAAAGTGGCCAA GAGCCTGAG	244.
	用来扩增 FANCF-3-GUIDE_seq-OT#6 的正向引物	GTTCTCCAAAGGAAGA GAGGGGAATG	245.
	用来扩增 FANCF-3-GUIDE_seq-OT#6 的反向引物	GGTGCTGTGTCCTCAT GCATCC	246.
	用来扩增 FANCF-3-GUIDE_seq-OT#7 的正向引物	CGGCTTGCCTAGGGTC GTTGAG	247.
	用来扩增 FANCF-3-GUIDE_seq-OT#7 的反向引物	CCTTCAGGGGCTCTTC CAGGTC	248.
	用来扩增中靶的 RUNX1-1 的正向引物	GGGAACTGGCAGGCA CCGAGG	249.
		GGGTGAGGCTGAAAC AGTGACC	250.
	用来扩增中靶的 RUNX1-1 的反向引物		
	用来扩增 RUNX1-1-GUIDE_seq-OT#1 的正向引物	GGGAGGATGTTGGTTT TAGGGAACTG	251.
	用来扩增 RUNX1-1-GUIDE_seq-OT#1 的反向引物	TCCAATCACTACATGC CATTTTGAAGA	252.
	用来扩增 RUNX1-1-GUIDE_seq-OT#2 的正向引物	CCACCCTCTTCCTTTGA TCCTCCC	253.
	用来扩增 RUNX1-1-GUIDE_seq-OT#2 的反向引物	TCCTCCCTACTCCTTCA CCCAGG	254.
	用来扩增中靶的 ZSCAN2 的正向引物	GAGTGCCTGACATGTG GGGAGAG	255.
		TCCAGCTAAAGCCTTT CCCACAC	256.
	用来扩增中靶的 ZSCAN2 的反向引物		
	用来扩增 ZSCAN2-GUIDE_seq-OT#1 的正向引物	GAACTCTCTGATGCAC CTGAAGGCTG	257.
	用来扩增 ZSCAN2-GUIDE_seq-OT#1 的反向引物	ACCGTATCAGTGTGAT GCATGTGGT	258.
	用来扩增 ZSCAN2-GUIDE_seq-OT#2 的正向引物	TGGGTTTAATCATGTG TTCTGCACTATG	259.
	用来扩增 ZSCAN2-GUIDE_seq-OT#2 的反向引物	CCCATCTTCCATTCTGC CCTCCAC	260.
	用来扩增 ZSCAN2-GUIDE_seq-OT#3 的正向引物	CAGCTAGTCCATTTGT TCTCAGACTGTG	261.
	用来扩增 ZSCAN2-GUIDE_seq-OT#3 的反向引物	GGCCAACATTGTGAAA CCCTGTCTC	262.

[0183]	用来扩增 ZSCAN2-GUIDE_seq-OT#4 的正向引物	CCAGGGACCTGTGCTT GGGTTC	263.
	用来扩增 ZSCAN2-GUIDE_seq-OT#4 的反向引物	CACCCCATGACCTGGC ACAAGTG	264.
	用来扩增 ZSCAN2-GUIDE_seq-OT#5 的正向引物	AAGTGTTTCCTCAGAAT GCCAGCCC	265.
	用来扩增 ZSCAN2-GUIDE_seq-OT#5 的反向引物	CAGGAGTGCAGTTGTG TTGGGAG	266.
	用来扩增 ZSCAN2-GUIDE_seq-OT#6 的正向引物	CTGATGAAGCACCAGA GAACCCACC	267.
	用来扩增 ZSCAN2-GUIDE_seq-OT#6 的反向引物	CACACCTGGCACCCAT ATGGC	268.
	用来扩增 ZSCAN2-GUIDE_seq-OT#7 的正向引物	GATCCACACTGGTGAG AAGCCTTAC	269.
	用来扩增 ZSCAN2-GUIDE_seq-OT#7 的反向引物	CTTCCCACACTCACAG CAGATGTAGG	270.

[0184] 将SpCas9-HF1的特异性细化

[0185] 先前所述的方法,例如截短的gRNA (Fu,Y.等人,Nat Biotechnol[自然生物技术] 32,279-284 (2014)) 和SpCas9-D1135E变体 (Kleinstiver,B.P.等人,Nature[自然]523,481-485 (2015)) 可以部分减小SpCas9脱靶效应,并且本发明的诸位发明人想知道这些是否可能与SpCas9-HF1组合,从而进一步改进其全基因组特异性。用在基于人类细胞的EGFP破坏测定中靶向四个位点的错配的全长的和截短的sgRNA测试SpCas9-HF1,揭示缩短sgRNA互补性长度实质上损害了中靶活性(图9)。相比之下,具有另外的D1135E突变的SpCas9-HF1(在此称为SpCas9-HF2的变体)保留了使用基于人类细胞的EGFP破坏测定测试的具有八个sgRNA中的六个的野生型SpCas9的70%的或更高的活性(图5a和5b)。还产生分别含L169A或Y450A突变的SpCas9-HF3和SpCas9-HF4变体,突变位置是在其侧链介导与靶DNA在其PAM近端的疏水非特异性相互作用处 (Nishimasu,H.等人,Cell[细胞]156,935-949 (2014); Jiang,F.等人,Science[科学]348,1477-1481 (2015))。SpCas9-HF3和SpCas9-HF4保留了用具有八个EGFP靶向的sgRNA中的相同六个的野生型SpCas9观察到的70%的或更高的活性(图5a和5b)。

[0186] 为了确定SpCas9-HF2、-HF3、和-HF4是否可以减小在对SpCas9-HF1有抗性的两个脱靶位点的indel频率(对于FANCF位点2和VEGFA位点3sgRNA),进行了进一步的实验。对于FANCF位点2脱靶,其携带在原型间隔子的种子序列中的单错配,如通过T7EI测定判断,SpCas9-HF4将indel突变频率降低至接近背景水平,同时还有益地增加了中靶活性(图5c),导致在这三个变体中的特异性方面的最大增加(图5d)。对于VEGFA位点3脱靶位点,其携带两个原型间隔子错配(一个在种子序列中,并且一个在PAM序列的最远端的核苷酸处),SpCas9-HF2显示在indel形成方面的最大减少,同时显示对中靶突变效率的仅适度的影响(图5c),导致在测试的三个变体中,特异性方面的最大增加(图5d)。总之,这些结果展示,通

过在介导非特异性DNA接触或可以改变PAM识别的其他残基处引入另外的突变,减少了对SpCas9-HF1有抗性的脱靶效应的潜能。

[0187] 为了分别对上述显示SpCas9-HF4和SpCas9-HF2具有针对FANCF位点2和VEGFA位点3sgRNA的脱靶的相对SpCas9-HF1的改进的辨别的T7E1测定发现进行概括,使用GUIDE-seq检查这些变体的全基因组特异性。使用RFLP测定,确定SpCas9-HF4和SpCas9-HF2具有如通过GUIDE-seq标签整合率测定,与SpCas9-HF1类似的中靶活性(图5E)。当分析GUIDE-seq数据时,对于SpCas9-HF2或SpCas9-HF4,没有鉴定到新的脱靶位点(图5F)。与SpCas9-HF1相比,在所有位点的脱靶活性都是通过GUIDE-seq不可检测抑或实质上降低的。相对于SpCas9-HF1,SpCas9-HF4具有针对单FANCF位点2脱靶位点(该位点保留了对SpCas9-HF1的特异性改进的不服从)的接近26倍更好的特异性(图5F)。对于高频率VEGFA位点3脱靶,相对于SpCas9-HF1,SpCas9-HF2具有接近4倍改进的特异性,同时还显著降低(>38倍)或消除了其他低频率脱靶位点处的GUIDE-seq可检测的事件。值得注意的是,针对SpCas9-HF1鉴定的这些低频率位点中的3个的基因组位置邻近先前表征的背景U2OS细胞断点热点。总而言之,这些结果提示,SpCas9-HF2和SpCas9-HF4变体可以改进SpCas9-HF1的全基因组特异性。

[0188] 当使用针对标准非重复靶序列设计的sgRNA时,SpCas9-HF1稳健并且一致地减少了脱靶突变。对SpCas9-HF1最具有抗性的两个脱靶位点具有原型间隔子中的唯一一个和两个错配。总之,这些观察结果提示,可能通过使用SpCas9-HF1,并且靶向并不具有携带基因组中其他位置处的一个或两个错配的紧密相关位点的非重复序列(其可以使用现有公开可得的软件程序实现),将脱靶突变最小化至不可检测的水平(Bae,S.等人,Bioinformatics[生物信息学]30,1473-1475(2014))。使用者应考虑到的一个参数是SpCas9-HF1可能是不与普通实践(使用与原型间隔子序列错配的gRNA的5'末端处的G)兼容的。携带与其靶位点错配的5' G的四个sgRNA的测试显示,与野生型SpCas9相比,四个中的三个具有与SpCas9-HF1之间的降低的活性(图10),可能反映了SpCas9-HF1更好地辨别了部分地错配的位点的能力。

[0189] 进一步的生物化学工作可以确认或阐明SpCas9-HF1借以实现其高的全基因组特异性的精确机制。引入的四个突变看起来并未改变SpCas9在细胞中的稳定性或稳态表达水平,因为降低浓度的表达质粒的滴定实验提示,野生型SpCas9和SpCas9-HF1表现为与其浓度降低是可比较的(图11)。替代地,最简单的机械解释是,这些突变减少了Cas9-sgRNA和靶DNA之间的相互作用的能量学,其中复合物的能量在刚好足以保留中靶活性的水平,而且将其充分降低,使得脱靶位点切割无效或不存在。这一机制符合结构数据中在突变的残基和靶DNA磷酸主链之间观察到的非特异性相互作用(Nishimasu,H.等人,Cell[细胞]156,935-949(2014),Anders,C等人,Nature[自然]513,569-573(2014))。已经提出在某种程度上类似的机制来解释携带在正电荷残基处的取代的类转录激活因子效应物核酸酶的增加的特异性(Guillinger,J.P.等人,Nat Methods[自然方法]11,429-435(2014))。

[0190] 可能是SpCas9-HF1也可能与已经显示改变Cas9功能的其他突变组合。例如,携带三个氨基酸取代的SpCas9突变体(D1135V/R1335Q/T1337R,也称为SpCas9-VQR变体)识别具有NGAN PAM的位点(具有以下相对效率,NGAG>NGAT=NGAA>NGAC)(Kleinstiver,B.P.等人,Nature[自然]523,481-485(2015)),并且最近鉴定的四重SpCas9突变体(D1135V/G1218R/R1335Q/T1337R,称为SpCas9-VRQR变体)相对于VQR,针对具有NGAH(H=A、C、或T)PAM的位

点,具有改进的活性(图12a)。分别将来自SpCas9-HF1的四个突变(N497A/R661A/Q695A/Q926A)引入SpCas9-VQR和SpCas9-VRQR,产生SpCas9-VQR-HF1和SpCas9-VRQR-HF1。这些核酸酶的两种HF版本显示出与其具有靶向EGFP报道基因的八个sgRNA中的五个的,以及具有靶向内源人类基因位点的八个sgRNA中的七个的非HF对应物的可比较的(即70%或更多)中靶活性(图12b-12d)。

[0191] 更广泛地,这些结果说明了用于CRISPR相关核酸酶的另外的高保真性变体的工程化的一般策略。在非特异性DNA接触残基处添加另外的突变进一步减少了SpCas9-HF1具有的、一些非常少量的残余脱靶位点。因此,取决于脱靶序列的性质,可以按自定义的方式利用变体,例如SpCas9-HF2、SpCas9-HF3、SpCas9-HF4等。此外,将SpCas9的高保真性变体工程化的成功提示,使非特异性DNA接触突变的方法可以扩展至其他天然存在的并且工程化的Cas9直系同源物(Ran,F.A.等人,Nature[自然]520,186-191(2015),Esvelt,K.M.等人,Nat Methods[自然方法]10,1116-1121(2013);Hou,Z.等人;Proc Natl Acad Sci U S A[美国国家科学院院刊](2013);Fonfara,I.等人;Nucleic Acids Res[核酸研究]42,2577-2590(2014);Kleinstiver,B.P.等人;Nat Biotechnol[自然生物技术](2015)),连同正以增加的频率被发现和表征的较新的CRISPR相关核酸酶(Zetsche,B.等人,Cell[细胞]163,759-771(2015);Shmakov,S等人,Molecular Cell[分子细胞]60,385-397)。

[0192] 实施例2

[0193] 在此描述的是在接触靶链DNA的残基中具有丙氨酸取代的SpCas9变体,包括N497A、Q695A、R661A、和Q926A。除了这些残基,本发明的诸位发明人试图确定是否可能通过在以下似乎与非靶DNA链接触的正电荷SpCas9残基中添加取代,进一步改进这些变体,例如SpCas9-HF1变体(N497A/R661A/Q695A/Q926A)的特异性:R780、K810、R832、K848、K855、K968、R976、H982、K1003、K1014、K1047、和/或R1060(参见Slaymaker等人,Science[科学].2016年1月1日;351(6268):84-8)。

[0194] 最初用设计为EGFP基因中的位点的完全匹配的sgRNA(以便评估中靶活性),以及携带在位置11和12(其中位置1是PAM最近端的碱基)处的有意错配的相同sgRNA(以便评估在错配位点处的活性,如将在脱靶位点处发现),使用EGFP破坏测定,测试携带在这些位置处的单丙氨酸取代的野生型SpCas9衍生物及其组合的活性(图13A)。(注意,携带三重取代K810A/K1003A/R1060A或K848A/K1003A/R1060A的衍生物与分别称为eSpCas9(1.0)和eSpCas9(1.1)的最近描述的变体是相同的;参见ref.1)。如预期,野生型SpCas9具有稳健的中靶和错配靶的活性。作为对照,在此实验中,我们还测试了SpCas9-HF1,并且发现,如预期,它维持了中靶活性,同时减少了错配靶活性(图13A)。携带在可能潜在接触非靶DNA链的位置处的一个或多个丙氨酸取代的所有野生型SpCas9衍生物显示出与野生型SpCas9可比较的中靶活性(图13A)。有趣的是,这些衍生物中的一些还显示出相对于用野生型SpCas9观察到的活性,错配的11/12sgRNA的切割减少了,提示在这些衍生物中的取代的子集赋予相对于野生型SpCas9,针对这一错配位点的增强的特异性(图13A)。然而,这些单取代或取代的组合中没有一个是足以完全消除用11/12错配的sgRNA观察到的活性。当我们使用携带在位置9和10处的错配的另外的sgRNA测试野生型SpCas9、SpCas9-HF1、和这些相同的野生型SpCas9衍生物时(图13B),对于多数衍生物,仅观察到错配靶活性方面的最小改变。这再一次展示了,在这些潜在非靶链接触残基处的单重、二重、或甚至三重取代(等价于先前所述

的eSpCas9(1.0)和(1.1)变体)不足以消除在不完全匹配DNA位点处的活性。总而言之,这些数据展示了,野生型SpCas9变体保留了与匹配的sgRNA的中靶活性,并且包含在这些衍生物中的取代靠它们自身(在野生型SpCas9的背景下)不足以消除针对两种不同错配的DNA位点的核酸酶活性(图13A和13B)。

[0195] 鉴于这些结果,假设携带在可能接触非靶DNA链的残基处的一个或多个另外的氨基酸取代的SpCas9-HF1衍生物相对于亲本SpCas9-HF1蛋白可能进一步改进了特异性。因此,在基于人类细胞的EGFP破坏测定中,使用完全匹配的sgRNA(以测试中靶活性),和携带在位置11和12处的错配的相同sgRNA(以评估在错配的靶位点处的活性,如将针对脱靶位点发现),测试携带单重、二重、或三重丙氨酸取代的组合的不同SpCas9-HF1衍生物。这些sgRNA是用于图13A-B的相同sgRNA。此实验揭示,我们测试的多数SpCas9-HF1衍生物变体显示了与用野生型SpCas9和SpCas9-HF1二者观察到的可比较的中靶活性(图14A)。用11/12错配的sgRNA,测试的SpCas9-HF1衍生物(例如SpCas9-HF1+R832A和SpCas9-HF1+K1014A)中的一些在用错配的sgRNA切割方面并未显示可感知的改变。然而,重要的是,与用SpCas9-HF1、eSpCas9(1.0)、或eSpCas9(1.1)观察到的相比,多数SpCas9-HF1衍生物具有与11/12错配的sgRNA的基本上更低的活性,提示这些新变体的某些组合已经减小了错配靶活性并且因此改进了特异性(图14A)。在将与11/12错配的sgRNA的错配靶活性减小至接近背景水平的16种SpCas9-HF1衍生物中,9种似乎对中靶活性具有仅最小的影响(使用完全匹配的sgRNA评估;图14A)。在EGFP破坏测定中,使用有意在位置9和10错配的sgRNA(图14B),这些SpCas9-HF1衍生物的子集的另外的测试也揭示,与用SpCas9-HF1(图14b),用eSpCas9(1.1)(图13A),抑或用添加至野生型SpCas9核酸酶的相同取代(图13B)观察到的相比,具有这一错配的sgRNA的这些变体拥有更低的活性。重要的是,在用9/10错配的sgRNA的此测定中,五种变体显示背景水平的脱靶活性。

[0196] 接下来,测试非靶链的这些丙氨酸取代是否可以与仅包含来自我们的SpCas9-HF1变体的Q695A和Q926A取代的SpCas9变体(在此为“二重”变体)组合。因为以上测试的很多HF1衍生物显示中靶活性方面的可观察到的(并且不希望的)减少,假设将来自SpCas9-HF1的仅有的两种最重要的取代(Q695A和Q926A;参见图1B)与接触取代的一个或多个非靶链结合可以拯救中靶活性,但是仍维持当这些取代被添加至SpCas9-HF1变体时观察到的特异性方面的增加。因此,使用用于图13A-B的靶向上述EGFP的相同完全匹配的sgRNA(以测试中靶活性)和携带在位置11和12处的错配的相同sgRNA(以评估在错配的靶位点处的活性,如将针对脱靶位点发现),在基于人类细胞的EGFP破坏测定中,测试携带在潜在非靶DNA链相互作用位置处的单重、二重、或三重丙氨酸取代的组合的不同SpCas9(Q695A/Q926A)衍生物。此实验揭示,多数测试的SpCas9(Q695A/Q926A)衍生物变体显示与用野生型SpCas9和SpCas9-HF1二者观察到的那些中靶活性可比较的中靶活性(图15)。重要的是,与用SpCas9-HF1、eSpCas9(1.0)或eSpCas9(1.1)观察到的相比,很多SpCas9-HF1衍生物具有与11/12错配的sgRNA的实质上更低的活性,提示这些新变体的某些组合具有减小的错配靶活性,并且因此具有改进的特异性(图15)。在将与11/12错配的sgRNA的错配靶活性减小至接近背景水平的13种SpCas9(Q695A/Q926A)衍生物中,仅1种似乎对中靶活性具有实质的影响(使用完全匹配的sgRNA评估;图15)。

[0197] 总体而言,这些数据展示了,在可能接触非靶DNA链的位置处,添加一个、两个、或

三个丙氨酸取代至SpCas9-HF1或SpCas9 (Q695A/Q926A) 可以导致具有改进的能力的新变体,以排斥错配的脱靶位点(相对于其亲本克隆或最近描述的eSpCas9 (1.0) 或 (1.1))。重要的是,在野生型SpCas9的背景下,这些相同取代似乎并不提供任何实质的特异性益处。

[0198] 为了更好地限定和比较SpCas9-HF1和eSpCas9-1.1对在sgRNA-靶DNA互补性界面处的错配的耐受性,使用包含在间隔子互补性区域中的所有可能位置处的单错配的sgRNA,检查其活性。当与野生型SpCas9比较时,SpCas9-HF1和eSpCas9-1.1变体二者具有对多数单错配的sgRNA的类似活性,有少量例外,其中SpCas9-HF1性能超过了eSpCas9-1.1 (图16)。

[0199] 接下来,我们测试了一些变体的单核苷酸错配耐受性,所述变体包含来自在接触靶链DNA或潜在接触非靶链DNA的残基中具有另外的丙氨酸取代的二重突变体 (Db=Q695A/Q926A)、SpCas9-HF1 (N497A/R661A/Q695A/Q926A)、eSpCas9-1.0 (1.0=K810A/K1003A/R1060A),抑或eSpCas9-1.1 (1.1=K848A/K1003A/R1060A) 的氨基酸取代的组合 (图17A-B)。使用完全匹配的sgRNA评估中靶活性,同时使用在间隔子序列中携带在位置4、8、12、或16处的此类错配的sgRNA,评估单核苷酸错配耐受性 (图17A)。多个这样的变体维持了中靶活性,在用错配的sgRNA观察到的活性方面,具有实质的减小。用剩余单错配sgRNA (包含在位置1-3、5-7、9-11、13-15、和17-20处的错配) 进一步测试了这些变体 (Q695A/K848A/Q926A/K1003A/R1060A、N497A/R661A/Q695A/K855A/Q926A/R1060A、和N497A/R661A/Q695A/Q926A/H982A/R1060A) 中的三个。这些变体展示了,与eSpCas9-1.1相比,对sgRNA中的单核苷酸取代的更稳健的不耐受,这展示了这些新变体的改进的特异性谱 (图17B)。使用在间隔子中包含在位置5、7、和9处的错配的sgRNA (因为更早的变体似乎耐受在这些位置处的错配,使用这些具体错配的sgRNA),测试包含氨基酸取代的替代性组合的另外的变体核酸酶 (图18)。多个这样的核酸酶具有针对错配位点的改进的特异性,具有中靶活性方面的仅微小的减小 (图18)。

[0200] 为了进一步确定突变的另外的组合是否可以传达特异性改进,测试了一个极大扩展的组的具有两种另外的匹配的sgRNA的核酸酶变体,以在EGFP破坏活性方面检查中靶活性 (图19A)。多个这样的变体维持了稳健的中靶活性,提示它们可以用于产生对特异性的进一步改进 (图19B)。用包含在位置12、14、16、或18处的单取代的sgRNA测试多个这样的变体,以确定是否可以观察到并且发现表现出对在这些位置处的单核苷酸错配的更大的不耐受的特异性改进 (图19B)。

[0201] 实施例3

[0202] 采用金黄色葡萄球菌Cas9 (SaCas9),如我们已经用SpCas9进行的那样类似的策略进行实验,以通过将丙氨酸取代引入已知接触靶DNA链的残基中 (图20和图21A),引入可以接触非靶DNA的残基中 (正在进行的实验),以及我们已经先前显示可以影响PAM特异性的残基中 (图21B),改进SaCas9的特异性。可以接触靶链DNA主链的残基包括:Y211,Y212,W229,Y230,R245,T392,N419,L446,Y651,和R654;可以接触非靶链DNA的残基包括:Q848,N492,Q495,R497,N498,R499,Q500,K518,K523,K525,H557,R561,K572,R634,R654,G655,N658,S662,N668,R686,K692,R694,H700,K751;并且接触PAM的残基包括:E782,D786,T787,Y789,T882,K886,N888,A889,L909,K929,N985,N986,R991,和R1015。在一个预实验中,在靶链DNA接触残基抑或PAM接触残基 (分别是图21A和B) 中,单丙氨酸取代 (或其一些组合) 具有对中靶EGFP破坏活性的可变影响 (使用完全匹配的sgRNA),并且不能消除脱靶切割 (当使用在位

置11和12处错配的sgRNA时)。有趣的是,在HF1中的SpCas9突变不能完全消除用类似错配的靶/sgrNA对的脱靶活性,提示包含靶链/非靶链取代的组合的变体对于改进在此类位点处的特异性而言是必需的(如我们用SpCas9观察到的)。

[0203] 为了进一步评估使潜在靶链DNA接触突变以改进SaCas9特异性的策略,检查了突变的单重、二重、三重、和四重组合耐受在sgRNA中的位置19和20处的错配的潜能(图22A和B)。这些组合揭示了,如通过更好地排斥错配位点的能力判断的那样,当与其他取代组合时,在Y230和R245处的丙氨酸取代可以增加特异性。

[0204] 接下来,在EGFP中的4个中靶位点处,检查了这些三重丙氨酸取代变体(Y211A/Y230A/R245A和Y212A/Y230A/R245A)中的两个的中靶基因破坏活性(匹配位点#1-4;图23)。对于匹配位点1和2,这些变体维持了稳健的中靶活性,但是用匹配位点3和4,显示了中靶活性的大约60%-70%的损失。如通过使用携带在靶位点1-4的间隔子中的不同位置处的二重错配的sgRNA判断的那样,相对于野生型SaCas9,这些三重丙氨酸取代变体中的两者显著改进了特异性(图23)。

[0205] 使用包含在位置21(预期PAM最远端位置对于排斥错配而言将是挑战)处的单错配的sgRNA,针对六个内源位点,测试了携带这些丙氨酸取代的二重和三重组的SaCas9变体(分别是图24A和B)的中靶活性和评估的特异性方面的改进。在一些情况下,用变体维持了匹配的sgRNA的中靶活性,同时消除了在位置21处错配的sgRNA的“脱靶”活性(图24A和B)。在其他情况下,用匹配的sgRNA观察到活性的微小至完全的损失。

[0206] 参考文献

[0207] 1.Sander,J.D.&Joung,J.K.CRISPR-Cas systems for editing,regulating and targeting genomes.Nat Biotechnol 32,347-355(2014) .

[0208] 2.Hsu,P.D.,Lander,E.S.&Zhang,F.Development and applications of CRISPR-Cas9 for genome engineering.Cell 157,1262-1278(2014) .

[0209] 3.Doudna,J.A.&Charpentier,E.Genome editing.The new frontier of genome engineering with CRISPR-Cas9.Science 346,1258096(2014) .

[0210] 4.Barrangou,R.&May,A.P.Unraveling the potential of CRISPR-Cas9 for gene therapy.Expert Opin Biol Ther 15,311-314(2015) .

[0211] 5.Jinek,M.et al.A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.Science 337,816-821(2012) .

[0212] 6.Sternberg,S.H.,Redding,S.,Jinek,M.,Greene,E.C.&Doudna,J.A.DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.Nature 507,62-67 (2014) .

[0213] 7.Hsu,P.D.et al.DNA targeting specificity of RNA-guided Cas9 nucleases.Nat Biotechnol 31,827-832(2013) .

[0214] 8.Tsai,S.Q.et al.GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases.Nat Biotechnol 33,187-197(2015) .

[0215] 9.Hou,Z.et al.Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis.Proc Natl Acad Sci U S A(2013) .

[0216] 10.Fonfara,I.et al.Phylogeny of Cas9 determines functional

exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res* 42,2577-2590(2014) .

[0217] 11.Esvelt,K.M.et al.Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods* 10,1116-1121(2013) .

[0218] 12.Cong,L.et al.Multiplex genome engineering using CRISPR/Cas systems. *Science* 339,819-823(2013) .

[0219] 13.Horvath,P.et al.Diversity,activity,and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* 190,1401-1412(2008) .

[0220] 14.Anders,C.,Niewoehner,O.,Duerst,A.&Jinek,M.Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513,569-573 (2014) .

[0221] 15.Reyon,D.et al.FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 30,460-465(2012) .

[0222] 16.Fu,Y.et al.High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31,822-826(2013) .

[0223] 17.Chen,Z.&Zhao,H.A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res* 33,e154(2005) .

[0224] 18.Doyon,J.B.,Pattanayak,V.,Meyer,C.B.&Liu,D.R.Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* 128,2477-2484(2006) .

[0225] 19.Jiang,W.,Bikard,D.,Cox,D.,Zhang,F.&Marraffini,L.A.RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31,233-239(2013) .

[0226] 20.Mali,P.et al.RNA-guided human genome engineering via Cas9. *Science* 339,823-826(2013) .

[0227] 21.Hwang,W.Y.et al.Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31,227-229(2013) .

[0228] 22.Chylinski,K.,Le Rhun,A.&Charpentier,E.The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol* 10,726-737(2013) .

[0229] 23.Kleinstiver,B.P.,Fernandes,A.D.,Gloor,G.B.&Edgell,D.R.A unified genetic,computational and experimental framework identifies functionally relevant residues of the homing endonuclease I-BmoI. *Nucleic Acids Res* 38, 2411-2427(2010) .

[0230] 24.Gagnon,J.A.et al.Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* 9,e98186(2014) .

[0231] 序列

[0232] SEQ ID NO:271-JDS246:CMV-T7-人类SpCas9-NLS-3xFLAG

[0233] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,NLS加双下划线,3xFLAG标

签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTGTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATG
[0234] AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA

ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA
CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCCCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAA
AAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTTACA
[0235] ATGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGC
CTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAG
ACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAG
AAAATTGAATGCTTCGATTCTGTTCGAGATCTCCGGGGTAGAAGATCGAT
TTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGA
TAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATA
GTGTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGAC
TAAAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAA
GAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAACTTATCAAC
GGGATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGA
GCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATGACTC
TTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGG
GACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCA
AAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGG
TCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCG
AAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGA

[0236]

AGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGG
AGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTA
TTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAAGTGGACATA
AACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTT
GAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAA
CCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAAT
GAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAG
AAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTT
GACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCCGCCAAATCA
CAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCA
AAATTGGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGG
AGATAAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGT
AGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTG
TATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCG
AACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACAT
TATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGG
GATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCC
AAGTCAACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAA
AGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAA
AAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTT
GCCTATTCTGTCTAGTAGTGCCAAAAGTTGAGAAGGGAAAATCCAAGA
AACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTC
GTCTTTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAG
GAAGTAAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTG
AGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTC
AAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTGA
TTTAGCGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAA
CAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCA
TAGAGCAAATTTGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAA

[0237] TCTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCAT
ACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCG
GCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACG
ATACACTTCTACCAAGGAGGTGCTAGACGCGGACACTGATTCACCAATCC
ATCACGGGATTATATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTG
ACGGATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGAC
CATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACG
ATGACAAGTGA

[0238] SEQ ID NO:272-VP12:CMV-T7-人类SpCas9-HF1 (N497A,R661A,Q695A,Q926A) - NLS-3xFLAG

[0239] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

[0240] ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGCGGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTGACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC

AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA
CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCCTGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACC_{gcc}TTTGACAAGAATTTACCGAACGAAA
AAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAAT
GAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCT
[0241] TTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC
CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAA
AATTGAATGCTTCGATTCTGTCTGAGATCTCCGGGGTAGAAGATCGATTT
AATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATA
AGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGT
GTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTA
AAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAAGA
GGCGTCGCTATACGGGCTGGGG_{Agcc}TTGTCTGCGGAAACTTATCAACGGG
ATAAGAGACAAGCAAAGTGGTAAACTATTCTCGATTTTCTAAAGAGCG
ACGGCTTCGCCAATAGGAACTTTATG_{gcc}CTGATCCATGATGACTCTTTAA
CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACT
CATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA
GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATG
GGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGGCAAAAAACAGTCGAGAGCGGATGAAGAG
AATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCA

TCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTAC
CTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAAC
CGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTTGAA
GGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGA
GGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAG
AACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAG
TTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACA
AGGCCGGATTTATTAACGTCAGCTCGTGGAAACCCGC_{gcc}ATCACAAAG
CATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGA
ACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATT
GGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA
AATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGA
CCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGG
TGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGA
[0242] ATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACG
ACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAA
GGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
AACATAGTAAAGAAAACCTGAGGTGCAGACCGGAGGGTTTTCAAAGGAA
TCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG
ACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTA
TTCTGTCCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAACTG
AAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT
TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT
AAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTA
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAG
GGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAG
CGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGA
AGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGA
GCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTG
GACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGT

GAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCG
CTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATA
CACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATC
[0243] ACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACG
GATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
ACAAGTGA

[0244] SEQ ID NO:273-MSP2135:CMV-T7-人类SpCas9-HF2 (N497A,R661A,Q695A,Q926A,
D1135E)-NLS-3xFLAG

[0245] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
[0246] TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA

CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA
CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTCTGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCCTGGAATTTTGAGGAAGTTGTCTGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACC_{gcc}TTTGACAAGAATTTACCGAACGAAA
AAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAAT
GAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCT
TTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC
[0247] CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAA
AATTGAATGCTTCGATTCTGTCTGAGATCTCCGGGGTAGAAGATCGATTT
AATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATA
AGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGT
GTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTA
AAAACATACGCTCACCTGTTCTGACGATAAGGTTATGAAACAGTTAAAGA
GGCGTCGCTATACGGGCTGGGGAG_{gcc}TTGTCTGCGGAAACTTATCAACGGG
ATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCG
ACGGCTTCGCCAATAGGAACTTTATG_{gcc}CTGATCCATGATGACTCTTTAA
CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACT
CATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA
GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATG
GGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGAG
AATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCA
TCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTAC

CTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAAC
CGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAA
GGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGA
GGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAG
AACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAG
TTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACA
AGGCCGGATTTATTAACGTCAGCTCGTGGAACCCCGC_{gcc}ATCACAAAG
CATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGA
ACGATAAGCTGATTCTGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATT
GGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA
AATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGA
CCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGG
TGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGA
ATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACG
[0248] ACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAA
GGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
AACATAGTAAAGAAAAGTACGGTGCAGACCGGAGGGTTTTCAAAGGAA
TCGATTCTTCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG
ACTGGGACCCGAAAAAGTACGGTGGCTTC_{gag}AGCCCTACAGTTGCCTAT
TCTGTCCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAACTG
AAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT
TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT
AAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTA
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAG
GGGAACGAACTCGCACTACCGTCTAAATACGTGAATTCCTGTATTTAG
CGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGA
AGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGA
GCAAATTTCTGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTG
GACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGT
GAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCG

[0249] CTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATA
CACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATC
ACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACG
GATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
ACAAGTGA

[0250] SEQ ID NO:274-MSP2133:CMV-T7-人类SpCas9-HF4(Y450A,N497A,R661A,Q695A,Q926A)-NLS-3xFLAG

[0251] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
[0252] AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA

[0253]

TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA
CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTg_{cc}TATGTGGGACCCCTGGCCCG
AGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGAT
TACTCCCTGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAA
TCGTTTCATCGAGAGGATGACCg_{cc}TTTGACAAGAATTTACCGAACGAAAA
AGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAATG
AACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTT
TCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC
CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAA
AATTGAATGCTTCGATTCTGTCGAGATCTCCGGGGTAGAAGATCGATTT
AATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATA
AGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGT
GTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTA
AAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGA
GGCGTCGCTATACGGGCTGGGGAg_{cc}TTGTCGCGGAAACTTATCAACGGG
ATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCG
ACGGCTTCGCCAATAGGAACCTTTATGg_{cc}CTGATCCATGATGACTCTTTAA
CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACT
CATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA
GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATG
GGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAG
AATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCA
TCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTAC
CTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAAC

CGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTTGAA
GGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGA
GGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAG
AACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAG
TTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACA
AGGCCGGATTTATTAACGTCAGCTCGTGGAACCCGC_{gcc}ATCACAAAG
CATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGA
ACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATT
GGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA
AATACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGA
CCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGG
TGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGA
ATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACG
ACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAA
[0254] GGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
AACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAAAGGAA
TCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG
ACTGGGACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTA
TTCTGTCCTAGTAGTGGC AAAAGTTGAGAAGGGAAAATCCAAGAACTG
AAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT
TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT
AAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTA
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAG
GGGAACGAACTCGCACTACCGTCTAAATACGTGAATTCCTGTATTTAG
CGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGA
AGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGA
GCAAATTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTG
GACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGT
GAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCG
CTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATA

CACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATC
ACGGGATTATATGAAACTCGGATAGATTTGTTCACAGCTTGGGGGTGACG
[0255] GATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
ACAAGTGA

[0256] SEQ ID NO:275-MSP469:CMV-T7-人类SpCas9-VQR(D1135V,R1335Q,T1337R)-NLS-3xFLAG

[0257] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
[0258] AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA

CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCATGGAATTTTGAGGAAGTTGTTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAA
AAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTTACAGTGTACA
ATGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGC
CTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAG
ACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAG
AAAATTGAATGCTTCGATTCTGTTCGAGATCTCCGGGGTAGAAGATCGAT
[0259] TTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGA
TAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATA
GTGTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGAC
TAAAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAA
GAGGCGTCGCTATACGGGCTGGGGACGATTGTCGCGGAAACTTATCAAC
GGGATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGA
GCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATGACTC
TTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGG
GACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCA
AAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGG
TCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCG
AAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGA
AGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGG
AGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTA
TTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATA
AACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTT

GAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAA
CCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAAT
GAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAG
AAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTT
GACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCCGCCAAATCA
CAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCA
AAATTGGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGG
AGATAAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGT
AGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTG
TATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCG
AACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACAT
TATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGG
GATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCC
[0260] AAGTCAACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAA
AGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAA
AAAGGACTGGGACCCGAAAAAGTACGGTGGCTTC_{gtg}AGCCCTACAGTTG
CCTATTCTGTCTAGTAGTGGAAGGTTGAGAAGGGAAAAATCCAAGAA
ACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCG
TCTTTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGG
AAGTAAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGA
GTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCA
AAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTAT
TTAGCGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAAC
AGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCAT
AGAGCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAAT
CTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATA
CGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGG
CGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAA_{cagT}
AC_{agaT}CTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATC

[0261] ACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACG
GATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATG
ACAAGTGA

[0262] SEQ ID NO:276-MSP2440:CMV-T7-人类SpCas9-VQR-HF1 (N497A,R661A,Q695A,
Q926A,D1135V,R1335Q,T1337R)-NLS-3xFLAG

[0263] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
[0264] AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA

[0265]

CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCCTGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACC_{gcc}TTTGACAAGAATTTACCGAACGAAA
AAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAAT
GAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCT
TTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC
CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAA
AATTGAATGCTTCGATTCTGTCTGAGATCTCCGGGGTAGAAGATCGATTT
AATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAGATA
AGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGT
GTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTA
AAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGA
GGCGTCGCTATACGGGCTGGGGAG_{gcc}TTGTCTGCGGAAACTTATCAACGGG
ATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCG
ACGGCTTCGCCAATAGGAACTTTATG_{gcc}CTGATCCATGATGACTCTTTAA
CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACT
CATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA
GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATG
GGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAG
AATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCA
TCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTAC
CTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAAC
CGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAA

GGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGA
GGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAG
AACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAGAAAG
TTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACA
AGGCCGGATTTATTAAACGTCAGCTCGTGGAACCCGC_{gcc}ATCACAAAG
CATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGA
ACGATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATT
GGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA
AATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGA
CCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGG
TGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGA
ATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACG
ACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAA
GGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
[0266] AACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAAAGGAA
TCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG
ACTGGGACCCGAAAAAGTACGGTGGCTTC_{gtg}AGCCCTACAGTTGCCTAT
TCTGTCCTAGTAGTGGAAGGTTGAGAAGGGGAAAATCCAAGAACTG
AAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT
TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT
AAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTA
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAG
GGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAG
CGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGA
AGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGA
GCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTG
GACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCATACGT
GAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCG
CTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAA_{cag}TAC
agaTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCAC

[0267] GGGATTATATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGA
TCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGA
CGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGAC
AAGTGA

[0268] SEQ ID NO:277-BPK2797:CMV-T7-人类SpCas9-VRQR (D1135V,G1218R,R1335Q,
T1337R)-NLS-3xFLAG

[0269] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
[0270] CGGATGTCGACAAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA
CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT

ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCATGGAATTTTGAGGAAGTTGTTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAA
AAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACA
ATGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGC
CTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAG
ACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAG
AAAATTGAATGCTTCGATTCTGTTCGAGATCTCCGGGGTAGAAGATCGAT
TTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGA
[0271] TAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATA
GTGTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGAC
TAAAAACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAA
GAGGCGTCGCTATACGGGCTGGGGACGATTGTTCGCGGAAACTTATCAAC
GGGATAAGAGACAAGCAAAGTGGTAAAACCTATTCTCGATTTTCTAAAGA
GCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATGACTC
TTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGG
GACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCA
AAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGG
TCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCG
AAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGA
AGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGG
AGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTA
TTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATA
AACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTT
GAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAA

CCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAAT
GAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAG
AAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTT
GACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAACCCGCCAAATCA
CAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGA
CGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCA
AAATTGGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGG
AGATAAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGT
AGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTG
TATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCG
AACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACAT
TATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGC
AAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGG
GATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCC
AAGTCAACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAA
[0272] AGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAA
AAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCgtgAGCCCTACAGTTG
CCTATTCTGTCCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAA
ACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCG
TCTTTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGG
AAGTAAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGA
GTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCcagaGAGCTTCAA
AAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATT
TAGCGTCCCATTACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACA
GAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATA
GAGCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATC
TGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATAC
GTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGC
GCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAAcagTA
CagaTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCA
CGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACGG

ATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATG

[0273] **ACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGA**
CAAGTGA

[0274] SEQ ID NO:278-MSP2443:CMV-T7-人类SpCas9-VRQR-HF1 (N497A,R661A,Q695A,Q926A,D1135V,G1218R,R1335Q,T1337R) -NLS-3xFLAG

[0275] 人类密码子优化的酿脓链球菌Cas9以常规字体显示,修饰的密码子以小写字母显示,NLS 加双下划线,3xFLAG标签以粗体显示:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCG
TTGGATGGGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATT
TAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATC
GGTGCCCTCCTATTTCGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGA
AACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATATGTT
ACTTACAAGAAATTTTGTAGCAATGAGATGGCCAAAGTTGACGATTCTTT
CTTTCACCGTTTGGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATG
AAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAAC
TGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATA
[0276] AAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAAC
CGGATGTCGACAACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTT
GTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATT
CTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCAC
AATTACCCGGAGAGAAGAAAAATGGGTTGTTTCGGTAACCTTATAGCGCT
CTCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAA
GATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACA
ATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTGGCTGCC
AAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATA
CTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGA
TGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA

CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGT
ACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTT
TATCAAACCCATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA
AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGACAACG
GTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAG
AAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGTGAAAAGATT
GAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCC
GAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGA
TTACTCCCTGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCA
ATCGTTCATCGAGAGGATGACC_{gcc}TTTGACAAGAATTTACCGAACGAAA
AAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTACAGTGTACAAT
GAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCT
TTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGAC
CAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAA
AATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTT
[0277] AATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATA
AGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGT
GTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTA
AAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAACAGTTAAAGA
GGCGTCGCTATACGGGCTGGGGAG_{gcc}TTGTGCGGAAACTTATCAACGGG
ATAAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCG
ACGGCTTCGCCAATAGGAACTTTATG_{gcc}CTGATCCATGATGACTCTTTAA
CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACT
CATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAA
GGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATG
GGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCGAAAAT
CAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAG
AATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCA
TCCTGTGGAAAATACCCAATTGCAGAACGAGAACTTTACCTCTATTAC
CTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAAC
CGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAA

GGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGA
GGGAAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAG
AACTATTGGCGGCAGCTCCTAAATGCGAAACTGATAACGCAAAGAAAG
TTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACA
AGGCCGGATTTATTAAACGTCAGCTCGTGGAACCCCGC_{gcc}ATCACAAAG
CATGTTGCGCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGA
ACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATT
GGTGTCTGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATA
AATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGA
CCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGG
TGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGA
ATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACG
ACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAA
GGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
[0278] AACATAGTAAAGAAAAGTGAAGGTGCAGACCGGAGGGTTTTCAAAGGAA
TCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGG
ACTGGGACCCGAAAAAGTACGGTGGCTTC_{gtg}AGCCCTACAGTTGCCTAT
TCTGTCCTAGTAGTGGAAGGTTGAGAAGGGGAAAATCCAAGAACTG
AAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT
TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGT
AAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTA
GAAAATGGCCGAAAACGGATGTTGGCTAGCGCC_{Caga}GAGCTTCAAAGG
GGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGC
GTCCCATACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAA
GCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGAG
CAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGG
ACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCATACGTG
AGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCT
CCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAA_{Acag}TAC_{Caga}
TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGG

[0279] GATTATATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATC
CCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACG
GTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAA
GTGA

[0280] SEQ ID NO:279-BPK1520:U6-BsmBI盒-Sp-sgRNA

[0281] U6启动子以常规字体显示,BsmBI位点以斜体显示,酿脓链球菌sgRNA以小写字母显示,U6终止子加双下划线:

[0282] TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGA
TCCGGTACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTT
CATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAA
TTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAG
TAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGA
CTATCATATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTGGCTTTAT
ATATCTTGTGGAAAGGACGAAACACCGGAGACGATTAATGCGTCTCCgttt
agagctagaaatagcaagttaaataaggctagtcggttatcaactgaaaaagtggcaccgagtcggtgctttttt

[0283] 其他实施方案

[0284] 应理解,虽然已经结合详细描述对本发明进行了描述,但前面的描述旨在说明而非限制本发明的范围,本发明的范围由所附的权利要求书的范围限定。其他方面、优点以及修饰都在以下权利要求书的范围之内。

[0001]	序列表
[0002]	<110> 通用医疗公司(THE GENERAL HOSPITAL CORPORATION)
[0003]	<120> 工程化的CRISPR-Cas9核酸酶
[0004]	<130> 3050-1809922D1
[0005]	<150> US 15/015,947
[0006]	<151> 2016-02-04
[0007]	<150> US 62/258,280
[0008]	<151> 2015-11-20
[0009]	<150> US 62/216,033
[0010]	<151> 2015-09-09
[0011]	<150> US 62/211,553
[0012]	<151> 2015-08-28
[0013]	<150> US 62/271,938
[0014]	<151> 2015-12-28
[0015]	<160> 279
[0016]	<170> SIPOSequenceListing 1.0
[0017]	<210> 1
[0018]	<211> 1368
[0019]	<212> PRT
[0020]	<213> 酿脓链球菌(Streptococcus pyogenes)
[0021]	<400> 1
[0022]	Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
[0023]	1 5 10 15
[0024]	Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
[0025]	20 25 30
[0026]	Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
[0027]	35 40 45
[0028]	Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
[0029]	50 55 60
[0030]	Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
[0031]	65 70 75 80
[0032]	Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
[0033]	85 90 95
[0034]	Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
[0035]	100 105 110
[0036]	His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
[0037]	115 120 125
[0038]	His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
[0039]	130 135 140
[0040]	Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
[0041]	145 150 155 160

[0042]	Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro		
[0043]		165	170
[0044]	Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr		175
[0045]		180	185
[0046]	Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala		190
[0047]		195	200
[0048]	Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn		205
[0049]		210	215
[0050]	Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn		220
[0051]		225	230
[0052]	Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe		235
[0053]		245	250
[0054]	Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp		255
[0055]		260	265
[0056]	Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp		270
[0057]		275	280
[0058]	Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp		285
[0059]		290	295
[0060]	Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser		300
[0061]		305	310
[0062]	Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys		315
[0063]		325	330
[0064]	Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe		335
[0065]		340	345
[0066]	Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser		350
[0067]		355	360
[0068]	Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp		365
[0069]		370	375
[0070]	Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg		380
[0071]		385	390
[0072]	Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu		395
[0073]		405	410
[0074]	Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe		415
[0075]		420	425
[0076]	Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile		430
[0077]		435	440
[0078]	Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp		445
[0079]		450	455
[0080]	Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu		460
[0081]		465	470
[0082]	Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr		475
[0083]		485	490

[0084]	Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser		
[0085]	500	505	510
[0086]	Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys		
[0087]	515	520	525
[0088]	Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln		
[0089]	530	535	540
[0090]	Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr		
[0091]	545	550	555
[0092]	Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp		
[0093]	565	570	575
[0094]	Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly		
[0095]	580	585	590
[0096]	Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp		
[0097]	595	600	605
[0098]	Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr		
[0099]	610	615	620
[0100]	Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala		
[0101]	625	630	635
[0102]	His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr		
[0103]	645	650	655
[0104]	Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp		
[0105]	660	665	670
[0106]	Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe		
[0107]	675	680	685
[0108]	Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe		
[0109]	690	695	700
[0110]	Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu		
[0111]	705	710	715
[0112]	His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly		
[0113]	725	730	735
[0114]	Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly		
[0115]	740	745	750
[0116]	Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln		
[0117]	755	760	765
[0118]	Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile		
[0119]	770	775	780
[0120]	Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro		
[0121]	785	790	795
[0122]	Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu		
[0123]	805	810	815
[0124]	Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg		
[0125]	820	825	830

[0126]	Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
[0127]	835 840 845
[0128]	Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
[0129]	850 855 860
[0130]	Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
[0131]	865 870 875 880
[0132]	Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
[0133]	885 890 895
[0134]	Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
[0135]	900 905 910
[0136]	Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
[0137]	915 920 925
[0138]	Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
[0139]	930 935 940
[0140]	Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
[0141]	945 950 955 960
[0142]	Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
[0143]	965 970 975
[0144]	Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
[0145]	980 985 990
[0146]	Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
[0147]	995 1000 1005
[0148]	Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys
[0149]	1010 1015 1020
[0150]	Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser
[0151]	1025 1030 1035 1040
[0152]	Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
[0153]	1045 1050 1055
[0154]	Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile
[0155]	1060 1065 1070
[0156]	Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
[0157]	1075 1080 1085
[0158]	Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
[0159]	1090 1095 1100
[0160]	Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
[0161]	1105 1110 1115 1120
[0162]	Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
[0163]	1125 1130 1135
[0164]	Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly
[0165]	1140 1145 1150
[0166]	Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
[0167]	1155 1160 1165

[0168]	Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala		
[0169]	1170	1175	1180
[0170]	Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys		
[0171]	1185	1190	1195
[0172]	Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser		
[0173]	1205	1210	1215
[0174]	Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr		
[0175]	1220	1225	1230
[0176]	Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
[0177]	1235	1240	1245
[0178]	Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His		
[0179]	1250	1255	1260
[0180]	Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val		
[0181]	1265	1270	1275
[0182]	Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys		
[0183]	1285	1290	1295
[0184]	His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu		
[0185]	1300	1305	1310
[0186]	Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp		
[0187]	1315	1320	1325
[0188]	Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp		
[0189]	1330	1335	1340
[0190]	Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile		
[0191]	1345	1350	1355
[0192]	Asp Leu Ser Gln Leu Gly Gly Asp		
[0193]	1365		
[0194]	<210> 2		
[0195]	<211> 1053		
[0196]	<212> PRT		
[0197]	<213> 金黄色葡萄球菌 (Staphylococcus aureus)		
[0198]	<400> 2		
[0199]	Met Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val		
[0200]	1	5	10
[0201]	Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly		
[0202]	20	25	30
[0203]	Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg		
[0204]	35	40	45
[0205]	Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile		
[0206]	50	55	60
[0207]	Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His		
[0208]	65	70	75
[0209]	Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu		

[0210]	85	90	95
[0211]	Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu		
[0212]	100	105	110
[0213]	Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr		
[0214]	115	120	125
[0215]	Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala		
[0216]	130	135	140
[0217]	Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys		
[0218]	145	150	155
[0219]	Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr		
[0220]	165	170	175
[0221]	Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln		
[0222]	180	185	190
[0223]	Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg		
[0224]	195	200	205
[0225]	Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys		
[0226]	210	215	220
[0227]	Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe		
[0228]	225	230	235
[0229]	Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr		
[0230]	245	250	255
[0231]	Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn		
[0232]	260	265	270
[0233]	Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe		
[0234]	275	280	285
[0235]	Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu		
[0236]	290	295	300
[0237]	Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys		
[0238]	305	310	315
[0239]	Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr		
[0240]	325	330	335
[0241]	Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala		
[0242]	340	345	350
[0243]	Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu		
[0244]	355	360	365
[0245]	Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser		
[0246]	370	375	380
[0247]	Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile		
[0248]	385	390	395
[0249]	Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala		
[0250]	405	410	415
[0251]	Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln		

[0252]	420	425	430
[0253]	Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro		
[0254]	435	440	445
[0255]	Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile		
[0256]	450	455	460
[0257]	Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg		
[0258]	465	470	475 480
[0259]	Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys		
[0260]	485	490	495
[0261]	Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr		
[0262]	500	505	510
[0263]	Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp		
[0264]	515	520	525
[0265]	Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu		
[0266]	530	535	540
[0267]	Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro		
[0268]	545	550	555 560
[0269]	Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys		
[0270]	565	570	575
[0271]	Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu		
[0272]	580	585	590
[0273]	Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile		
[0274]	595	600	605
[0275]	Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu		
[0276]	610	615	620
[0277]	Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp		
[0278]	625	630	635 640
[0279]	Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu		
[0280]	645	650	655
[0281]	Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys		
[0282]	660	665	670
[0283]	Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp		
[0284]	675	680	685
[0285]	Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp		
[0286]	690	695	700
[0287]	Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys		
[0288]	705	710	715 720
[0289]	Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys		
[0290]	725	730	735
[0291]	Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu		
[0292]	740	745	750
[0293]	Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp		

[0294]	755	760	765
[0295]	Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile		
[0296]	770	775	780
[0297]	Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu		
[0298]	785	790	795
[0299]	Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu		
[0300]	805	810	815
[0301]	Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His		
[0302]	820	825	830
[0303]	Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly		
[0304]	835	840	845
[0305]	Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr		
[0306]	850	855	860
[0307]	Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile		
[0308]	865	870	875
[0309]	Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp		
[0310]	885	890	895
[0311]	Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr		
[0312]	900	905	910
[0313]	Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val		
[0314]	915	920	925
[0315]	Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser		
[0316]	930	935	940
[0317]	Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala		
[0318]	945	950	955
[0319]	Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly		
[0320]	965	970	975
[0321]	Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile		
[0322]	980	985	990
[0323]	Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met		
[0324]	995	1000	1005
[0325]	Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr		
[0326]	1010	1015	1020
[0327]	Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr Glu		
[0328]	1025	1030	1035
[0329]	Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly		
[0330]	1045	1050	
[0331]	<210> 3		
[0332]	<211> 4		
[0333]	<212> PRT		
[0334]	<213> 人工序列(Artificial Sequence)		
[0335]	<220>		

[0336] <221> PEPTIDE
[0337] <222> (1) .. (4)
[0338] <223> 人工序列的说明: 合成 肽
[0339] <400> 3
[0340] Gly Gly Gly Ser
[0341] 1
[0342] <210> 4
[0343] <211> 5
[0344] <212> PRT
[0345] <213> 人工序列(Artificial Sequence)
[0346] <220>
[0347] <221> PEPTIDE
[0348] <222> (1) .. (5)
[0349] <223> 人工序列的说明: 合成 肽
[0350] <400> 4
[0351] Gly Gly Gly Gly Ser
[0352] 1 5
[0353] <210> 5
[0354] <211> 4
[0355] <212> PRT
[0356] <213> 人工序列(Artificial Sequence)
[0357] <220>
[0358] <221> PEPTIDE
[0359] <222> (1) .. (4)
[0360] <223> 人工序列的说明: 合成 肽
[0361] <400> 5
[0362] Gly Gly Gly Ser
[0363] 1
[0364] <210> 6
[0365] <211> 5
[0366] <212> PRT
[0367] <213> 人工序列(Artificial Sequence)
[0368] <220>
[0369] <221> PEPTIDE
[0370] <222> (1) .. (5)
[0371] <223> 人工序列的说明: 合成 肽
[0372] <400> 6
[0373] Gly Gly Gly Gly Ser
[0374] 1 5
[0375] <210> 7
[0376] <211> 7
[0377] <212> PRT

[0378] <213> 人工序列(Artificial Sequence)
[0379] <220>
[0380] <221> PEPTIDE
[0381] <222> (1) .. (7)
[0382] <223> 人工序列的说明: 合成 肽
[0383] <400> 7
[0384] Pro Lys Lys Lys Arg Arg Val
[0385] 1 5
[0386] <210> 8
[0387] <211> 16
[0388] <212> PRT
[0389] <213> 人工序列(Artificial Sequence)
[0390] <220>
[0391] <221> PEPTIDE
[0392] <222> (1) .. (16)
[0393] <223> 人工序列的说明: 合成 肽
[0394] <400> 8
[0395] Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
[0396] 1 5 10 15
[0397] <210> 9
[0398] <211> 20
[0399] <212> DNA
[0400] <213> 人工序列(Artificial Sequence)
[0401] <220>
[0402] <221> gene
[0403] <222> (1) .. (20)
[0404] <223> 人工序列的说明: 合成 寡核苷酸
[0405] <400> 9
[0406] gggcacgggc agcttgccgg 20
[0407] <210> 10
[0408] <211> 24
[0409] <212> DNA
[0410] <213> 人工序列(Artificial Sequence)
[0411] <220>
[0412] <221> gene
[0413] <222> (1) .. (24)
[0414] <223> 人工序列的说明: 合成 寡核苷酸
[0415] <400> 10
[0416] gggcacgggc agcttgccgg tggt 24
[0417] <210> 11
[0418] <211> 18
[0419] <212> DNA

[0420] <213> 人工序列(Artificial Sequence)
[0421] <220>
[0422] <221> gene
[0423] <222> (1) .. (18)
[0424] <223> 人工序列的说明: 合成 寡核苷酸
[0425] <400> 11
[0426] gcacgggcag cttgccgg 18
[0427] <210> 12
[0428] <211> 22
[0429] <212> DNA
[0430] <213> 人工序列(Artificial Sequence)
[0431] <220>
[0432] <221> gene
[0433] <222> (1) .. (22)
[0434] <223> 人工序列的说明: 合成 寡核苷酸
[0435] <400> 12
[0436] gcacgggcag cttgccgtg gt 22
[0437] <210> 13
[0438] <211> 20
[0439] <212> DNA
[0440] <213> 人工序列(Artificial Sequence)
[0441] <220>
[0442] <221> gene
[0443] <222> (1) .. (20)
[0444] <223> 人工序列的说明: 合成 寡核苷酸
[0445] <400> 13
[0446] gggcacccgc agcttgccgg 20
[0447] <210> 14
[0448] <211> 24
[0449] <212> DNA
[0450] <213> 人工序列(Artificial Sequence)
[0451] <220>
[0452] <221> gene
[0453] <222> (1) .. (24)
[0454] <223> 人工序列的说明: 合成 寡核苷酸
[0455] <400> 14
[0456] gggcacccgc agcttgccgg tggg 24
[0457] <210> 15
[0458] <211> 20
[0459] <212> DNA
[0460] <213> 人工序列(Artificial Sequence)
[0461] <220>

[0462] <221> gene
[0463] <222> (1) .. (20)
[0464] <223> 人工序列的说明: 合成 寡核苷酸
[0465] <400> 15
[0466] gggctggggc agcttgccgg 20
[0467] <210> 16
[0468] <211> 24
[0469] <212> DNA
[0470] <213> 人工序列(Artificial Sequence)
[0471] <220>
[0472] <221> gene
[0473] <222> (1) .. (24)
[0474] <223> 人工序列的说明: 合成 寡核苷酸
[0475] <400> 16
[0476] gggctggggc agcttgccgg tggt 24
[0477] <210> 17
[0478] <211> 20
[0479] <212> DNA
[0480] <213> 人工序列(Artificial Sequence)
[0481] <220>
[0482] <221> gene
[0483] <222> (1) .. (20)
[0484] <223> 人工序列的说明: 合成 寡核苷酸
[0485] <400> 17
[0486] ggcgacgggc agcttgccgg 20
[0487] <210> 18
[0488] <211> 24
[0489] <212> DNA
[0490] <213> 人工序列(Artificial Sequence)
[0491] <220>
[0492] <221> gene
[0493] <222> (1) .. (24)
[0494] <223> 人工序列的说明: 合成 寡核苷酸
[0495] <400> 18
[0496] ggcgacgggc agcttgccgg tggt 24
[0497] <210> 19
[0498] <211> 20
[0499] <212> DNA
[0500] <213> 人工序列(Artificial Sequence)
[0501] <220>
[0502] <221> gene
[0503] <222> (1) .. (20)

[0504] <223> 人工序列的说明: 合成 寡核苷酸
[0505] <400> 19
[0506] gcccacgggc agcttgccgg 20
[0507] <210> 20
[0508] <211> 24
[0509] <212> DNA
[0510] <213> 人工序列(Artificial Sequence)
[0511] <220>
[0512] <221> gene
[0513] <222> (1) .. (24)
[0514] <223> 人工序列的说明: 合成 寡核苷酸
[0515] <400> 20
[0516] gcccacgggc agcttgccgg tggt 24
[0517] <210> 21
[0518] <211> 20
[0519] <212> DNA
[0520] <213> 人工序列(Artificial Sequence)
[0521] <220>
[0522] <221> gene
[0523] <222> (1) .. (20)
[0524] <223> 人工序列的说明: 合成 寡核苷酸
[0525] <400> 21
[0526] gtcgccctcg aacttcacct 20
[0527] <210> 22
[0528] <211> 24
[0529] <212> DNA
[0530] <213> 人工序列(Artificial Sequence)
[0531] <220>
[0532] <221> gene
[0533] <222> (1) .. (24)
[0534] <223> 人工序列的说明: 合成 寡核苷酸
[0535] <400> 22
[0536] gtcgccctcg aacttcacct cggc 24
[0537] <210> 23
[0538] <211> 20
[0539] <212> DNA
[0540] <213> 人工序列(Artificial Sequence)
[0541] <220>
[0542] <221> gene
[0543] <222> (1) .. (20)
[0544] <223> 人工序列的说明: 合成 寡核苷酸
[0545] <400> 23

[0546] gtaggtcagg gtggtcacga 20
[0547] <210> 24
[0548] <211> 24
[0549] <212> DNA
[0550] <213> 人工序列(Artificial Sequence)
[0551] <220>
[0552] <221> gene
[0553] <222> (1) .. (24)
[0554] <223> 人工序列的说明: 合成 寡核苷酸
[0555] <400> 24
[0556] gtaggtcagg gtggtcacga gggt 24
[0557] <210> 25
[0558] <211> 20
[0559] <212> DNA
[0560] <213> 人工序列(Artificial Sequence)
[0561] <220>
[0562] <221> gene
[0563] <222> (1) .. (20)
[0564] <223> 人工序列的说明: 合成 寡核苷酸
[0565] <400> 25
[0566] ggcgagggcg atgccaccta 20
[0567] <210> 26
[0568] <211> 24
[0569] <212> DNA
[0570] <213> 人工序列(Artificial Sequence)
[0571] <220>
[0572] <221> gene
[0573] <222> (1) .. (20)
[0574] <223> 人工序列的说明: 合成 寡核苷酸
[0575] <400> 26
[0576] ggcgagggcg atgccaccta cggc 24
[0577] <210> 27
[0578] <211> 20
[0579] <212> DNA
[0580] <213> 人工序列(Artificial Sequence)
[0581] <220>
[0582] <221> gene
[0583] <222> (1) .. (20)
[0584] <223> 人工序列的说明: 合成 寡核苷酸
[0585] <400> 27
[0586] ggtcgccacc atggtgagca 20
[0587] <210> 28

[0588] <211> 24
[0589] <212> DNA
[0590] <213> 人工序列(Artificial Sequence)
[0591] <220>
[0592] <221> gene
[0593] <222> (1) .. (24)
[0594] <223> 人工序列的说明: 合成 寡核苷酸
[0595] <400> 28
[0596] ggtcgccacc atggtgagca aggg 24
[0597] <210> 29
[0598] <211> 20
[0599] <212> DNA
[0600] <213> 人工序列(Artificial Sequence)
[0601] <220>
[0602] <221> gene
[0603] <222> (1) .. (20)
[0604] <223> 人工序列的说明: 合成 寡核苷酸
[0605] <400> 29
[0606] ggtcagggtg gtcacgaggg 20
[0607] <210> 30
[0608] <211> 24
[0609] <212> DNA
[0610] <213> 人工序列(Artificial Sequence)
[0611] <220>
[0612] <221> gene
[0613] <222> (1) .. (24)
[0614] <223> 人工序列的说明: 合成 寡核苷酸
[0615] <400> 30
[0616] ggtcagggtg gtcacgaggg tggg 24
[0617] <210> 31
[0618] <211> 20
[0619] <212> DNA
[0620] <213> 人工序列(Artificial Sequence)
[0621] <220>
[0622] <221> gene
[0623] <222> (1) .. (20)
[0624] <223> 人工序列的说明: 合成 寡核苷酸
[0625] <400> 31
[0626] ggtggtgcag atgaacttca 20
[0627] <210> 32
[0628] <211> 24
[0629] <212> DNA

[0630] <213> 人工序列(Artificial Sequence)
[0631] <220>
[0632] <221> gene
[0633] <222> (1) .. (24)
[0634] <223> 人工序列的说明: 合成 寡核苷酸
[0635] <400> 32
[0636] ggtggtgcag atgaacttca ggg 24
[0637] <210> 33
[0638] <211> 17
[0639] <212> DNA
[0640] <213> 人工序列(Artificial Sequence)
[0641] <220>
[0642] <221> gene
[0643] <222> (1) .. (17)
[0644] <223> 人工序列的说明: 合成 寡核苷酸
[0645] <400> 33
[0646] ggtgcagatg aacttca 17
[0647] <210> 34
[0648] <211> 21
[0649] <212> DNA
[0650] <213> 人工序列(Artificial Sequence)
[0651] <220>
[0652] <221> gene
[0653] <222> (1) .. (21)
[0654] <223> 人工序列的说明: 合成 寡核苷酸
[0655] <400> 34
[0656] ggtgcagatg aacttcagg t 21
[0657] <210> 35
[0658] <211> 20
[0659] <212> DNA
[0660] <213> 人工序列(Artificial Sequence)
[0661] <220>
[0662] <221> gene
[0663] <222> (1) .. (20)
[0664] <223> 人工序列的说明: 合成 寡核苷酸
[0665] <400> 35
[0666] gttgggtct ttgctcagg 20
[0667] <210> 36
[0668] <211> 24
[0669] <212> DNA
[0670] <213> 人工序列(Artificial Sequence)
[0671] <220>

[0672] <221> gene
[0673] <222> (1) .. (24)
[0674] <223> 人工序列的说明: 合成 寡核苷酸
[0675] <400> 36
[0676] gttggggtct ttgctcaggg cgga 24
[0677] <210> 37
[0678] <211> 20
[0679] <212> DNA
[0680] <213> 人工序列(Artificial Sequence)
[0681] <220>
[0682] <221> gene
[0683] <222> (1) .. (20)
[0684] <223> 人工序列的说明: 合成 寡核苷酸
[0685] <400> 37
[0686] ggtgggtcacg aggggtgggcc 20
[0687] <210> 38
[0688] <211> 24
[0689] <212> DNA
[0690] <213> 人工序列(Artificial Sequence)
[0691] <220>
[0692] <221> gene
[0693] <222> (1) .. (24)
[0694] <223> 人工序列的说明: 合成 寡核苷酸
[0695] <400> 38
[0696] ggtgggtcacg aggggtgggcc aggg 24
[0697] <210> 39
[0698] <211> 20
[0699] <212> DNA
[0700] <213> 人工序列(Artificial Sequence)
[0701] <220>
[0702] <221> gene
[0703] <222> (1) .. (20)
[0704] <223> 人工序列的说明: 合成 寡核苷酸
[0705] <400> 39
[0706] gatgccgttc ttctgcttgt 20
[0707] <210> 40
[0708] <211> 24
[0709] <212> DNA
[0710] <213> 人工序列(Artificial Sequence)
[0711] <220>
[0712] <221> gene
[0713] <222> (1) .. (24)

[0714] <223> 人工序列的说明: 合成 寡核苷酸
[0715] <400> 40
[0716] gatgccgttc ttctgcttgt cggc 24
[0717] <210> 41
[0718] <211> 17
[0719] <212> DNA
[0720] <213> 人工序列(Artificial Sequence)
[0721] <220>
[0722] <221> gene
[0723] <222> (1) .. (17)
[0724] <223> 人工序列的说明: 合成 寡核苷酸
[0725] <400> 41
[0726] gccgttcttc tgcttgt 17
[0727] <210> 42
[0728] <211> 21
[0729] <212> DNA
[0730] <213> 人工序列(Artificial Sequence)
[0731] <220>
[0732] <221> gene
[0733] <222> (1) .. (21)
[0734] <223> 人工序列的说明: 合成 寡核苷酸
[0735] <400> 42
[0736] gccgttcttc tgcttgctgg c 21
[0737] <210> 43
[0738] <211> 20
[0739] <212> DNA
[0740] <213> 人工序列(Artificial Sequence)
[0741] <220>
[0742] <221> gene
[0743] <222> (1) .. (20)
[0744] <223> 人工序列的说明: 合成 寡核苷酸
[0745] <400> 43
[0746] gtcgccacca tggtagagcaa 20
[0747] <210> 44
[0748] <211> 24
[0749] <212> DNA
[0750] <213> 人工序列(Artificial Sequence)
[0751] <220>
[0752] <221> gene
[0753] <222> (1) .. (24)
[0754] <223> 人工序列的说明: 合成 寡核苷酸
[0755] <400> 44

[0756] gtcgccacca tggtagagcaa gggc 24
[0757] <210> 45
[0758] <211> 20
[0759] <212> DNA
[0760] <213> 人工序列(Artificial Sequence)
[0761] <220>
[0762] <221> gene
[0763] <222> (1) .. (20)
[0764] <223> 人工序列的说明: 合成 寡核苷酸
[0765] <400> 45
[0766] gcactgcacg ccgtaggtca 20
[0767] <210> 46
[0768] <211> 24
[0769] <212> DNA
[0770] <213> 人工序列(Artificial Sequence)
[0771] <220>
[0772] <221> gene
[0773] <222> (1) .. (24)
[0774] <223> 人工序列的说明: 合成 寡核苷酸
[0775] <400> 46
[0776] gcactgcacg ccgtaggtca gggt 24
[0777] <210> 47
[0778] <211> 20
[0779] <212> DNA
[0780] <213> 人工序列(Artificial Sequence)
[0781] <220>
[0782] <221> gene
[0783] <222> (1) .. (20)
[0784] <223> 人工序列的说明: 合成 寡核苷酸
[0785] <400> 47
[0786] gtgaaccgca tcgagctgaa 20
[0787] <210> 48
[0788] <211> 24
[0789] <212> DNA
[0790] <213> 人工序列(Artificial Sequence)
[0791] <220>
[0792] <221> gene
[0793] <222> (1) .. (24)
[0794] <223> 人工序列的说明: 合成 寡核苷酸
[0795] <400> 48
[0796] gtgaaccgca tcgagctgaa gggc 24
[0797] <210> 49

[0798] <211> 20
[0799] <212> DNA
[0800] <213> 人工序列(Artificial Sequence)
[0801] <220>
[0802] <221> gene
[0803] <222> (1) .. (20)
[0804] <223> 人工序列的说明: 合成 寡核苷酸
[0805] <400> 49
[0806] gaaggcatc gacttcaagg 20
[0807] <210> 50
[0808] <211> 24
[0809] <212> DNA
[0810] <213> 人工序列(Artificial Sequence)
[0811] <220>
[0812] <221> gene
[0813] <222> (1) .. (24)
[0814] <223> 人工序列的说明: 合成 寡核苷酸
[0815] <400> 50
[0816] gaaggcatc gacttcaagg agga 24
[0817] <210> 51
[0818] <211> 20
[0819] <212> DNA
[0820] <213> 人工序列(Artificial Sequence)
[0821] <220>
[0822] <221> gene
[0823] <222> (1) .. (20)
[0824] <223> 人工序列的说明: 合成 寡核苷酸
[0825] <400> 51
[0826] gcttcatgtg gtcgggtag 20
[0827] <210> 52
[0828] <211> 24
[0829] <212> DNA
[0830] <213> 人工序列(Artificial Sequence)
[0831] <220>
[0832] <221> gene
[0833] <222> (1) .. (24)
[0834] <223> 人工序列的说明: 合成 寡核苷酸
[0835] <400> 52
[0836] gcttcatgtg gtcgggtag cggc 24
[0837] <210> 53
[0838] <211> 20
[0839] <212> DNA

[0840] <213> 人工序列(Artificial Sequence)
[0841] <220>
[0842] <221> gene
[0843] <222> (1) .. (20)
[0844] <223> 人工序列的说明: 合成 寡核苷酸
[0845] <400> 53
[0846] gctgaagcac tgcacgccgt 20
[0847] <210> 54
[0848] <211> 24
[0849] <212> DNA
[0850] <213> 人工序列(Artificial Sequence)
[0851] <220>
[0852] <221> gene
[0853] <222> (1) .. (24)
[0854] <223> 人工序列的说明: 合成 寡核苷酸
[0855] <400> 54
[0856] gctgaagcac tgcacgccgt aggt 24
[0857] <210> 55
[0858] <211> 20
[0859] <212> DNA
[0860] <213> 人工序列(Artificial Sequence)
[0861] <220>
[0862] <221> gene
[0863] <222> (1) .. (20)
[0864] <223> 人工序列的说明: 合成 寡核苷酸
[0865] <400> 55
[0866] gccgtcgtcc ttgaagaaga 20
[0867] <210> 56
[0868] <211> 24
[0869] <212> DNA
[0870] <213> 人工序列(Artificial Sequence)
[0871] <220>
[0872] <221> gene
[0873] <222> (1) .. (24)
[0874] <223> 人工序列的说明: 合成 寡核苷酸
[0875] <400> 56
[0876] gccgtcgtcc ttgaagaaga tggt 24
[0877] <210> 57
[0878] <211> 20
[0879] <212> DNA
[0880] <213> 人工序列(Artificial Sequence)
[0881] <220>

[0882] <221> gene
[0883] <222> (1) .. (20)
[0884] <223> 人工序列的说明: 合成 寡核苷酸
[0885] <400> 57
[0886] gaccaggatg ggcaccaccc 20
[0887] <210> 58
[0888] <211> 24
[0889] <212> DNA
[0890] <213> 人工序列(Artificial Sequence)
[0891] <220>
[0892] <221> gene
[0893] <222> (1) .. (24)
[0894] <223> 人工序列的说明: 合成 寡核苷酸
[0895] <400> 58
[0896] gaccaggatg ggcaccaccc cggt 24
[0897] <210> 59
[0898] <211> 20
[0899] <212> DNA
[0900] <213> 人工序列(Artificial Sequence)
[0901] <220>
[0902] <221> gene
[0903] <222> (1) .. (20)
[0904] <223> 人工序列的说明: 合成 寡核苷酸
[0905] <400> 59
[0906] gacgtagcct tcgggcatgg 20
[0907] <210> 60
[0908] <211> 24
[0909] <212> DNA
[0910] <213> 人工序列(Artificial Sequence)
[0911] <220>
[0912] <221> gene
[0913] <222> (1) .. (24)
[0914] <223> 人工序列的说明: 合成 寡核苷酸
[0915] <400> 60
[0916] gacgtagcct tcgggcatgg cgga 24
[0917] <210> 61
[0918] <211> 20
[0919] <212> DNA
[0920] <213> 人工序列(Artificial Sequence)
[0921] <220>
[0922] <221> gene
[0923] <222> (1) .. (20)

- [0924] <223> 人工序列的说明: 合成 寡核苷酸
[0925] <400> 61
[0926] gaagttcgag ggcgacaccc 20
[0927] <210> 62
[0928] <211> 24
[0929] <212> DNA
[0930] <213> 人工序列(Artificial Sequence)
[0931] <220>
[0932] <221> gene
[0933] <222> (1) .. (24)
[0934] <223> 人工序列的说明: 合成 寡核苷酸
[0935] <400> 62
[0936] gaagttcgag ggcgacaccc tggc 24
[0937] <210> 63
[0938] <211> 20
[0939] <212> DNA
[0940] <213> 人工序列(Artificial Sequence)
[0941] <220>
[0942] <221> gene
[0943] <222> (1) .. (20)
[0944] <223> 人工序列的说明: 合成 寡核苷酸
[0945] <400> 63
[0946] gagctggacg ggcgacgtaaa 20
[0947] <210> 64
[0948] <211> 24
[0949] <212> DNA
[0950] <213> 人工序列(Artificial Sequence)
[0951] <220>
[0952] <221> gene
[0953] <222> (1) .. (24)
[0954] <223> 人工序列的说明: 合成 寡核苷酸
[0955] <400> 64
[0956] gagctggacg ggcgacgtaaa cgcc 24
[0957] <210> 65
[0958] <211> 20
[0959] <212> DNA
[0960] <213> 人工序列(Artificial Sequence)
[0961] <220>
[0962] <221> gene
[0963] <222> (1) .. (20)
[0964] <223> 人工序列的说明: 合成 寡核苷酸
[0965] <400> 65

[0966] ggcacgcgcc tcgccctcgc 20
[0967] <210> 66
[0968] <211> 24
[0969] <212> DNA
[0970] <213> 人工序列(Artificial Sequence)
[0971] <220>
[0972] <221> gene
[0973] <222> (1) .. (24)
[0974] <223> 人工序列的说明: 合成 寡核苷酸
[0975] <400> 66
[0976] ggcacgcgcc tcgccctcgc cgga 24
[0977] <210> 67
[0978] <211> 20
[0979] <212> DNA
[0980] <213> 人工序列(Artificial Sequence)
[0981] <220>
[0982] <221> gene
[0983] <222> (1) .. (20)
[0984] <223> 人工序列的说明: 合成 寡核苷酸
[0985] <400> 67
[0986] ggccacaagt tcagcgtgtc 20
[0987] <210> 68
[0988] <211> 24
[0989] <212> DNA
[0990] <213> 人工序列(Artificial Sequence)
[0991] <220>
[0992] <221> GC_signal
[0993] <222> (1) .. (24)
[0994] <223> 人工序列的说明: 合成 寡核苷酸
[0995] <400> 68
[0996] ggccacaagt tcagcgtgtc cggc 24
[0997] <210> 69
[0998] <211> 20
[0999] <212> DNA
[1000] <213> 人工序列(Artificial Sequence)
[1001] <220>
[1002] <221> gene
[1003] <222> (1) .. (20)
[1004] <223> 人工序列的说明: 合成 寡核苷酸
[1005] <400> 69
[1006] gggcgaggag ctgttcaccg 20
[1007] <210> 70

[1008] <211> 24
[1009] <212> DNA
[1010] <213> 人工序列(Artificial Sequence)
[1011] <220>
[1012] <221> gene
[1013] <222> (1) .. (24)
[1014] <223> 人工序列的说明: 合成 寡核苷酸
[1015] <400> 70
[1016] gggcgaggag ctgttcaccg gggt 24
[1017] <210> 71
[1018] <211> 18
[1019] <212> DNA
[1020] <213> 人工序列(Artificial Sequence)
[1021] <220>
[1022] <221> gene
[1023] <222> (1) .. (18)
[1024] <223> 人工序列的说明: 合成 寡核苷酸
[1025] <400> 71
[1026] gcgaggagct gttcaccg 18
[1027] <210> 72
[1028] <211> 22
[1029] <212> DNA
[1030] <213> 人工序列(Artificial Sequence)
[1031] <220>
[1032] <221> gene
[1033] <222> (1) .. (22)
[1034] <223> 人工序列的说明: 合成 寡核苷酸
[1035] <400> 72
[1036] gcgaggagct gttcaccggg gt 22
[1037] <210> 73
[1038] <211> 20
[1039] <212> DNA
[1040] <213> 人工序列(Artificial Sequence)
[1041] <220>
[1042] <221> gene
[1043] <222> (1) .. (20)
[1044] <223> 人工序列的说明: 合成 寡核苷酸
[1045] <400> 73
[1046] cctcgaactt cacctcggcg 20
[1047] <210> 74
[1048] <211> 24
[1049] <212> DNA

[1050] <213> 人工序列(Artificial Sequence)
[1051] <220>
[1052] <221> gene
[1053] <222> (1) .. (24)
[1054] <223> 人工序列的说明: 合成 寡核苷酸
[1055] <400> 74
[1056] cctcgaactt cacctcggcg cggg 24
[1057] <210> 75
[1058] <211> 20
[1059] <212> DNA
[1060] <213> 人工序列(Artificial Sequence)
[1061] <220>
[1062] <221> gene
[1063] <222> (1) .. (20)
[1064] <223> 人工序列的说明: 合成 寡核苷酸
[1065] <400> 75
[1066] gctcgaactt cacctcggcg 20
[1067] <210> 76
[1068] <211> 24
[1069] <212> DNA
[1070] <213> 人工序列(Artificial Sequence)
[1071] <220>
[1072] <221> gene
[1073] <222> (1) .. (24)
[1074] <223> 人工序列的说明: 合成 寡核苷酸
[1075] <400> 76
[1076] gctcgaactt cacctcggcg cggg 24
[1077] <210> 77
[1078] <211> 20
[1079] <212> DNA
[1080] <213> 人工序列(Artificial Sequence)
[1081] <220>
[1082] <221> gene
[1083] <222> (1) .. (20)
[1084] <223> 人工序列的说明: 合成 寡核苷酸
[1085] <400> 77
[1086] caactacaag acccgcgccg 20
[1087] <210> 78
[1088] <211> 24
[1089] <212> DNA
[1090] <213> 人工序列(Artificial Sequence)
[1091] <220>

[1092] <221> gene
[1093] <222> (1) .. (24)
[1094] <223> 人工序列的说明: 合成 寡核苷酸
[1095] <400> 78
[1096] caactacaag acccgcgccg aggt 24
[1097] <210> 79
[1098] <211> 20
[1099] <212> DNA
[1100] <213> 人工序列(Artificial Sequence)
[1101] <220>
[1102] <221> gene
[1103] <222> (1) .. (20)
[1104] <223> 人工序列的说明: 合成 寡核苷酸
[1105] <400> 79
[1106] gaactacaag acccgcgccg 20
[1107] <210> 80
[1108] <211> 24
[1109] <212> DNA
[1110] <213> 人工序列(Artificial Sequence)
[1111] <220>
[1112] <221> gene
[1113] <222> (1) .. (24)
[1114] <223> 人工序列的说明: 合成 寡核苷酸
[1115] <400> 80
[1116] gaactacaag acccgcgccg aggt 24
[1117] <210> 81
[1118] <211> 20
[1119] <212> DNA
[1120] <213> 人工序列(Artificial Sequence)
[1121] <220>
[1122] <221> gene
[1123] <222> (1) .. (20)
[1124] <223> 人工序列的说明: 合成 寡核苷酸
[1125] <400> 81
[1126] cgctcctgga cgtagccttc 20
[1127] <210> 82
[1128] <211> 24
[1129] <212> DNA
[1130] <213> 人工序列(Artificial Sequence)
[1131] <220>
[1132] <221> gene
[1133] <222> (1) .. (24)

- [1134] <223> 人工序列的说明: 合成 寡核苷酸
[1135] <400> 82
[1136] cgctcctgga cgtagccttc gggc 24
[1137] <210> 83
[1138] <211> 20
[1139] <212> DNA
[1140] <213> 人工序列(Artificial Sequence)
[1141] <220>
[1142] <221> gene
[1143] <222> (1) .. (20)
[1144] <223> 人工序列的说明: 合成 寡核苷酸
[1145] <400> 83
[1146] ggctcctgga cgtagccttc 20
[1147] <210> 84
[1148] <211> 24
[1149] <212> DNA
[1150] <213> 人工序列(Artificial Sequence)
[1151] <220>
[1152] <221> gene
[1153] <222> (1) .. (24)
[1154] <223> 人工序列的说明: 合成 寡核苷酸
[1155] <400> 84
[1156] cgctcctgga cgtagccttc gggc 24
[1157] <210> 85
[1158] <211> 20
[1159] <212> DNA
[1160] <213> 人工序列(Artificial Sequence)
[1161] <220>
[1162] <221> gene
[1163] <222> (1) .. (20)
[1164] <223> 人工序列的说明: 合成 寡核苷酸
[1165] <400> 85
[1166] agggcgagga gctgttcacc 20
[1167] <210> 86
[1168] <211> 24
[1169] <212> DNA
[1170] <213> 人工序列(Artificial Sequence)
[1171] <220>
[1172] <221> gene
[1173] <222> (1) .. (24)
[1174] <223> 人工序列的说明: 合成 寡核苷酸
[1175] <400> 86

[1176] agggcgagga gctgttcacc gggg 24
[1177] <210> 87
[1178] <211> 20
[1179] <212> DNA
[1180] <213> 人工序列(Artificial Sequence)
[1181] <220>
[1182] <221> gene
[1183] <222> (1) .. (20)
[1184] <223> 人工序列的说明: 合成 寡核苷酸
[1185] <400> 87
[1186] gggcgagga gctgttcacc 20
[1187] <210> 88
[1188] <211> 24
[1189] <212> DNA
[1190] <213> 人工序列(Artificial Sequence)
[1191] <220>
[1192] <221> gene
[1193] <222> (1) .. (24)
[1194] <223> 人工序列的说明: 合成 寡核苷酸
[1195] <400> 88
[1196] gggcgagga gctgttcacc gggg 24
[1197] <210> 89
[1198] <211> 20
[1199] <212> DNA
[1200] <213> 人工序列(Artificial Sequence)
[1201] <220>
[1202] <221> gene
[1203] <222> (1) .. (20)
[1204] <223> 人工序列的说明: 合成 寡核苷酸
[1205] <400> 89
[1206] gttcgagggc gacaccctgg 20
[1207] <210> 90
[1208] <211> 24
[1209] <212> DNA
[1210] <213> 人工序列(Artificial Sequence)
[1211] <220>
[1212] <221> gene
[1213] <222> (1) .. (24)
[1214] <223> 人工序列的说明: 合成 寡核苷酸
[1215] <400> 90
[1216] gttcgagggc gacaccctgg tgaa 24
[1217] <210> 91

[1218] <211> 20
[1219] <212> DNA
[1220] <213> 人工序列(Artificial Sequence)
[1221] <220>
[1222] <221> gene
[1223] <222> (1) .. (20)
[1224] <223> 人工序列的说明: 合成 寡核苷酸
[1225] <400> 91
[1226] gttcaccagg gtgtcgccct 20
[1227] <210> 92
[1228] <211> 24
[1229] <212> DNA
[1230] <213> 人工序列(Artificial Sequence)
[1231] <220>
[1232] <221> gene
[1233] <222> (1) .. (24)
[1234] <223> 人工序列的说明: 合成 寡核苷酸
[1235] <400> 92
[1236] gttcaccagg gtgtcgccct cgaa 24
[1237] <210> 93
[1238] <211> 20
[1239] <212> DNA
[1240] <213> 人工序列(Artificial Sequence)
[1241] <220>
[1242] <221> gene
[1243] <222> (1) .. (20)
[1244] <223> 人工序列的说明: 合成 寡核苷酸
[1245] <400> 93
[1246] gcccaccctc gtgaccaccc 20
[1247] <210> 94
[1248] <211> 24
[1249] <212> DNA
[1250] <213> 人工序列(Artificial Sequence)
[1251] <220>
[1252] <221> gene
[1253] <222> (1) .. (24)
[1254] <223> 人工序列的说明: 合成 寡核苷酸
[1255] <400> 94
[1256] gcccaccctc gtgaccaccc tgac 24
[1257] <210> 95
[1258] <211> 20
[1259] <212> DNA

[1260] <213> 人工序列(Artificial Sequence)
[1261] <220>
[1262] <221> gene
[1263] <222> (1) .. (20)
[1264] <223> 人工序列的说明: 合成 寡核苷酸
[1265] <400> 95
[1266] gcccttgctc accatggtgg 20
[1267] <210> 96
[1268] <211> 24
[1269] <212> DNA
[1270] <213> 人工序列(Artificial Sequence)
[1271] <220>
[1272] <221> gene
[1273] <222> (1) .. (24)
[1274] <223> 人工序列的说明: 合成 寡核苷酸
[1275] <400> 96
[1276] gcccttgctc accatggtgg cgac 24
[1277] <210> 97
[1278] <211> 20
[1279] <212> DNA
[1280] <213> 人工序列(Artificial Sequence)
[1281] <220>
[1282] <221> gene
[1283] <222> (1) .. (20)
[1284] <223> 人工序列的说明: 合成 寡核苷酸
[1285] <400> 97
[1286] gtcgccgtcc agctcgacca 20
[1287] <210> 98
[1288] <211> 24
[1289] <212> DNA
[1290] <213> 人工序列(Artificial Sequence)
[1291] <220>
[1292] <221> gene
[1293] <222> (1) .. (24)
[1294] <223> 人工序列的说明: 合成 寡核苷酸
[1295] <400> 98
[1296] gtcgccgtcc agctcgacca ggat 24
[1297] <210> 99
[1298] <211> 20
[1299] <212> DNA
[1300] <213> 人工序列(Artificial Sequence)
[1301] <220>

[1302] <221> gene
[1303] <222> (1) .. (20)
[1304] <223> 人工序列的说明: 合成 寡核苷酸
[1305] <400> 99
[1306] gtgtccggcg agggcgaggg 20
[1307] <210> 100
[1308] <211> 24
[1309] <212> DNA
[1310] <213> 人工序列(Artificial Sequence)
[1311] <220>
[1312] <221> gene
[1313] <222> (1) .. (24)
[1314] <223> 人工序列的说明: 合成 寡核苷酸
[1315] <400> 100
[1316] gtgtccggcg agggcgaggg cgat 24
[1317] <210> 101
[1318] <211> 20
[1319] <212> DNA
[1320] <213> 人工序列(Artificial Sequence)
[1321] <220>
[1322] <221> gene
[1323] <222> (1) .. (20)
[1324] <223> 人工序列的说明: 合成 寡核苷酸
[1325] <400> 101
[1326] ggggtggtgc ccatcctggt 20
[1327] <210> 102
[1328] <211> 24
[1329] <212> DNA
[1330] <213> 人工序列(Artificial Sequence)
[1331] <220>
[1332] <221> gene
[1333] <222> (1) .. (24)
[1334] <223> 人工序列的说明: 合成 寡核苷酸
[1335] <400> 102
[1336] ggggtggtgc ccatcctggt cgag 24
[1337] <210> 103
[1338] <211> 20
[1339] <212> DNA
[1340] <213> 人工序列(Artificial Sequence)
[1341] <220>
[1342] <221> gene
[1343] <222> (1) .. (20)

- [1344] <223> 人工序列的说明: 合成 寡核苷酸
[1345] <400> 103
[1346] gccaccatgg tgagcaaggg 20
[1347] <210> 104
[1348] <211> 24
[1349] <212> DNA
[1350] <213> 人工序列(Artificial Sequence)
[1351] <220>
[1352] <221> gene
[1353] <222> (1) .. (24)
[1354] <223> 人工序列的说明: 合成 寡核苷酸
[1355] <400> 104
[1356] gccaccatgg tgagcaaggg cgag 24
[1357] <210> 105
[1358] <211> 20
[1359] <212> DNA
[1360] <213> 人工序列(Artificial Sequence)
[1361] <220>
[1362] <221> gene
[1363] <222> (1) .. (20)
[1364] <223> 人工序列的说明: 合成 寡核苷酸
[1365] <400> 105
[1366] gagtccgagc agaagaagaa 20
[1367] <210> 106
[1368] <211> 24
[1369] <212> DNA
[1370] <213> 人工序列(Artificial Sequence)
[1371] <220>
[1372] <221> gene
[1373] <222> (1) .. (24)
[1374] <223> 人工序列的说明: 合成 寡核苷酸
[1375] <400> 106
[1376] gagtccgagc agaagaagaa gggc 24
[1377] <210> 107
[1378] <211> 20
[1379] <212> DNA
[1380] <213> 人工序列(Artificial Sequence)
[1381] <220>
[1382] <221> gene
[1383] <222> (1) .. (20)
[1384] <223> 人工序列的说明: 合成 寡核苷酸
[1385] <400> 107

[1386] gtcacctcca atgactaggg 20
[1387] <210> 108
[1388] <211> 24
[1389] <212> DNA
[1390] <213> 人工序列(Artificial Sequence)
[1391] <220>
[1392] <221> gene
[1393] <222> (1) .. (24)
[1394] <223> 人工序列的说明: 合成 寡核苷酸
[1395] <400> 108
[1396] gtcacctcca atgactaggg tggg 24
[1397] <210> 109
[1398] <211> 20
[1399] <212> DNA
[1400] <213> 人工序列(Artificial Sequence)
[1401] <220>
[1402] <221> gene
[1403] <222> (1) .. (20)
[1404] <223> 人工序列的说明: 合成 寡核苷酸
[1405] <400> 109
[1406] gggaagactg aggctacata 20
[1407] <210> 110
[1408] <211> 24
[1409] <212> DNA
[1410] <213> 人工序列(Artificial Sequence)
[1411] <220>
[1412] <221> gene
[1413] <222> (1) .. (24)
[1414] <223> 人工序列的说明: 合成 寡核苷酸
[1415] <400> 110
[1416] gggaagactg aggctacata gggt 24
[1417] <210> 111
[1418] <211> 20
[1419] <212> DNA
[1420] <213> 人工序列(Artificial Sequence)
[1421] <220>
[1422] <221> gene
[1423] <222> (1) .. (20)
[1424] <223> 人工序列的说明: 合成 寡核苷酸
[1425] <400> 111
[1426] gccacgaagc aggccaatgg 20
[1427] <210> 112

[1428] <211> 24
[1429] <212> DNA
[1430] <213> 人工序列(Artificial Sequence)
[1431] <220>
[1432] <221> gene
[1433] <222> (1) .. (24)
[1434] <223> 人工序列的说明: 合成 寡核苷酸
[1435] <400> 112
[1436] gccacgaagc aggccaatgg ggag 24
[1437] <210> 113
[1438] <211> 20
[1439] <212> DNA
[1440] <213> 人工序列(Artificial Sequence)
[1441] <220>
[1442] <221> gene
[1443] <222> (1) .. (20)
[1444] <223> 人工序列的说明: 合成 寡核苷酸
[1445] <400> 113
[1446] ggaatccctt ctgcagcacc 20
[1447] <210> 114
[1448] <211> 24
[1449] <212> DNA
[1450] <213> 人工序列(Artificial Sequence)
[1451] <220>
[1452] <221> gene
[1453] <222> (1) .. (24)
[1454] <223> 人工序列的说明: 合成 寡核苷酸
[1455] <400> 114
[1456] ggaatccctt ctgcagcacc tgga 24
[1457] <210> 115
[1458] <211> 20
[1459] <212> DNA
[1460] <213> 人工序列(Artificial Sequence)
[1461] <220>
[1462] <221> gene
[1463] <222> (1) .. (20)
[1464] <223> 人工序列的说明: 合成 寡核苷酸
[1465] <400> 115
[1466] gctgcagaag ggattccatg 20
[1467] <210> 116
[1468] <211> 24
[1469] <212> DNA

[1470] <213> 人工序列(Artificial Sequence)
[1471] <220>
[1472] <221> gene
[1473] <222> (1) .. (24)
[1474] <223> 人工序列的说明: 合成 寡核苷酸
[1475] <400> 116
[1476] gctgcagaag ggattccatg aggt 24
[1477] <210> 117
[1478] <211> 20
[1479] <212> DNA
[1480] <213> 人工序列(Artificial Sequence)
[1481] <220>
[1482] <221> gene
[1483] <222> (1) .. (20)
[1484] <223> 人工序列的说明: 合成 寡核苷酸
[1485] <400> 117
[1486] ggcggctgca caaccagtgg 20
[1487] <210> 118
[1488] <211> 24
[1489] <212> DNA
[1490] <213> 人工序列(Artificial Sequence)
[1491] <220>
[1492] <221> gene
[1493] <222> (1) .. (24)
[1494] <223> 人工序列的说明: 合成 寡核苷酸
[1495] <400> 118
[1496] ggcggctgca caaccagtgg aggc 24
[1497] <210> 119
[1498] <211> 20
[1499] <212> DNA
[1500] <213> 人工序列(Artificial Sequence)
[1501] <220>
[1502] <221> gene
[1503] <222> (1) .. (20)
[1504] <223> 人工序列的说明: 合成 寡核苷酸
[1505] <400> 119
[1506] gctccagagc cgtgcgaatg 20
[1507] <210> 120
[1508] <211> 24
[1509] <212> DNA
[1510] <213> 人工序列(Artificial Sequence)
[1511] <220>

[1512] <221> gene
[1513] <222> (1) .. (24)
[1514] <223> 人工序列的说明: 合成 寡核苷酸
[1515] <400> 120
[1516] gctccagagc cgtgcgaatg gggc 24
[1517] <210> 121
[1518] <211> 20
[1519] <212> DNA
[1520] <213> 人工序列(Artificial Sequence)
[1521] <220>
[1522] <221> gene
[1523] <222> (1) .. (20)
[1524] <223> 人工序列的说明: 合成 寡核苷酸
[1525] <400> 121
[1526] gaatcccttc tgcagcacct 20
[1527] <210> 122
[1528] <211> 24
[1529] <212> DNA
[1530] <213> 人工序列(Artificial Sequence)
[1531] <220>
[1532] <221> gene
[1533] <222> (1) .. (24)
[1534] <223> 人工序列的说明: 合成 寡核苷酸
[1535] <400> 122
[1536] gaatcccttc tgcagcacct ggat 24
[1537] <210> 123
[1538] <211> 20
[1539] <212> DNA
[1540] <213> 人工序列(Artificial Sequence)
[1541] <220>
[1542] <221> gene
[1543] <222> (1) .. (20)
[1544] <223> 人工序列的说明: 合成 寡核苷酸
[1545] <400> 123
[1546] gcggcggtg cacaaccagt 20
[1547] <210> 124
[1548] <211> 24
[1549] <212> DNA
[1550] <213> 人工序列(Artificial Sequence)
[1551] <220>
[1552] <221> gene
[1553] <222> (1) .. (24)

- [1554] <223> 人工序列的说明: 合成 寡核苷酸
[1555] <400> 124
[1556] gcggcggtg cacaaccagt ggag 24
[1557] <210> 125
[1558] <211> 20
[1559] <212> DNA
[1560] <213> 人工序列(Artificial Sequence)
[1561] <220>
[1562] <221> gene
[1563] <222> (1) .. (20)
[1564] <223> 人工序列的说明: 合成 寡核苷酸
[1565] <400> 125
[1566] ggttgtgcag ccgccgtcc 20
[1567] <210> 126
[1568] <211> 24
[1569] <212> DNA
[1570] <213> 人工序列(Artificial Sequence)
[1571] <220>
[1572] <221> gene
[1573] <222> (1) .. (24)
[1574] <223> 人工序列的说明: 合成 寡核苷酸
[1575] <400> 126
[1576] ggttgtgcag ccgccgtcc agag 24
[1577] <210> 127
[1578] <211> 20
[1579] <212> DNA
[1580] <213> 人工序列(Artificial Sequence)
[1581] <220>
[1582] <221> gene
[1583] <222> (1) .. (20)
[1584] <223> 人工序列的说明: 合成 寡核苷酸
[1585] <400> 127
[1586] gcattttcag gaggaagcga 20
[1587] <210> 128
[1588] <211> 24
[1589] <212> DNA
[1590] <213> 人工序列(Artificial Sequence)
[1591] <220>
[1592] <221> gene
[1593] <222> (1) .. (24)
[1594] <223> 人工序列的说明: 合成 寡核苷酸
[1595] <400> 128

[1596] gcattttcag gaggaagcga tggc 24
[1597] <210> 129
[1598] <211> 20
[1599] <212> DNA
[1600] <213> 人工序列(Artificial Sequence)
[1601] <220>
[1602] <221> gene
[1603] <222> (1) .. (20)
[1604] <223> 人工序列的说明: 合成 寡核苷酸
[1605] <400> 129
[1606] gggagaagaa agagagatgt 20
[1607] <210> 130
[1608] <211> 24
[1609] <212> DNA
[1610] <213> 人工序列(Artificial Sequence)
[1611] <220>
[1612] <221> gene
[1613] <222> (1) .. (24)
[1614] <223> 人工序列的说明: 合成 寡核苷酸
[1615] <400> 130
[1616] gggagaagaa agagagatgt aggg 24
[1617] <210> 131
[1618] <211> 20
[1619] <212> DNA
[1620] <213> 人工序列(Artificial Sequence)
[1621] <220>
[1622] <221> gene
[1623] <222> (1) .. (20)
[1624] <223> 人工序列的说明: 合成 寡核苷酸
[1625] <400> 131
[1626] ggtgcatttt caggaggaag 20
[1627] <210> 132
[1628] <211> 24
[1629] <212> DNA
[1630] <213> 人工序列(Artificial Sequence)
[1631] <220>
[1632] <221> gene
[1633] <222> (1) .. (24)
[1634] <223> 人工序列的说明: 合成 寡核苷酸
[1635] <400> 132
[1636] ggtgcatttt caggaggaag cgat 24
[1637] <210> 133

[1638] <211> 20
[1639] <212> DNA
[1640] <213> 人工序列(Artificial Sequence)
[1641] <220>
[1642] <221> gene
[1643] <222> (1) .. (20)
[1644] <223> 人工序列的说明: 合成 寡核苷酸
[1645] <400> 133
[1646] gagatgtagg gctagagggg 20
[1647] <210> 134
[1648] <211> 24
[1649] <212> DNA
[1650] <213> 人工序列(Artificial Sequence)
[1651] <220>
[1652] <221> gene
[1653] <222> (1) .. (24)
[1654] <223> 人工序列的说明: 合成 寡核苷酸
[1655] <400> 134
[1656] gagatgtagg gctagagggg tgag 24
[1657] <210> 135
[1658] <211> 20
[1659] <212> DNA
[1660] <213> 人工序列(Artificial Sequence)
[1661] <220>
[1662] <221> gene
[1663] <222> (1) .. (20)
[1664] <223> 人工序列的说明: 合成 寡核苷酸
[1665] <400> 135
[1666] ggtatccagc agaggggaga 20
[1667] <210> 136
[1668] <211> 24
[1669] <212> DNA
[1670] <213> 人工序列(Artificial Sequence)
[1671] <220>
[1672] <221> gene
[1673] <222> (1) .. (24)
[1674] <223> 人工序列的说明: 合成 寡核苷酸
[1675] <400> 136
[1676] ggtatccagc agaggggaga agaa 24
[1677] <210> 137
[1678] <211> 20
[1679] <212> DNA

[1680] <213> 人工序列(Artificial Sequence)
[1681] <220>
[1682] <221> gene
[1683] <222> (1) .. (20)
[1684] <223> 人工序列的说明: 合成 寡核苷酸
[1685] <400> 137
[1686] gaggcacatctc tgcaccgagg 20
[1687] <210> 138
[1688] <211> 24
[1689] <212> DNA
[1690] <213> 人工序列(Artificial Sequence)
[1691] <220>
[1692] <221> gene
[1693] <222> (1) .. (24)
[1694] <223> 人工序列的说明: 合成 寡核苷酸
[1695] <400> 138
[1696] gaggcacatctc tgcaccgagg tgaa 24
[1697] <210> 139
[1698] <211> 20
[1699] <212> DNA
[1700] <213> 人工序列(Artificial Sequence)
[1701] <220>
[1702] <221> gene
[1703] <222> (1) .. (20)
[1704] <223> 人工序列的说明: 合成 寡核苷酸
[1705] <400> 139
[1706] gaggggtgag gctgaaacag 20
[1707] <210> 140
[1708] <211> 24
[1709] <212> DNA
[1710] <213> 人工序列(Artificial Sequence)
[1711] <220>
[1712] <221> gene
[1713] <222> (1) .. (24)
[1714] <223> 人工序列的说明: 合成 寡核苷酸
[1715] <400> 140
[1716] gaggggtgag gctgaaacag tgac 24
[1717] <210> 141
[1718] <211> 20
[1719] <212> DNA
[1720] <213> 人工序列(Artificial Sequence)
[1721] <220>

[1722] <221> gene
[1723] <222> (1) .. (20)
[1724] <223> 人工序列的说明: 合成 寡核苷酸
[1725] <400> 141
[1726] gagcaaaagt agatattaca 20
[1727] <210> 142
[1728] <211> 24
[1729] <212> DNA
[1730] <213> 人工序列(Artificial Sequence)
[1731] <220>
[1732] <221> gene
[1733] <222> (1) .. (24)
[1734] <223> 人工序列的说明: 合成 寡核苷酸
[1735] <400> 142
[1736] gagcaaaagt agatattaca agac 24
[1737] <210> 143
[1738] <211> 20
[1739] <212> DNA
[1740] <213> 人工序列(Artificial Sequence)
[1741] <220>
[1742] <221> gene
[1743] <222> (1) .. (20)
[1744] <223> 人工序列的说明: 合成 寡核苷酸
[1745] <400> 143
[1746] ggaattcaaa ctgaggcata 20
[1747] <210> 144
[1748] <211> 24
[1749] <212> DNA
[1750] <213> 人工序列(Artificial Sequence)
[1751] <220>
[1752] <221> gene
[1753] <222> (1) .. (24)
[1754] <223> 人工序列的说明: 合成 寡核苷酸
[1755] <400> 144
[1756] ggaattcaaa ctgaggcata tgat 24
[1757] <210> 145
[1758] <211> 20
[1759] <212> DNA
[1760] <213> 人工序列(Artificial Sequence)
[1761] <220>
[1762] <221> gene
[1763] <222> (1) .. (20)

- [1764] <223> 人工序列的说明: 合成 寡核苷酸
[1765] <400> 145
[1766] gcagagggga gaagaaagag 20
[1767] <210> 146
[1768] <211> 24
[1769] <212> DNA
[1770] <213> 人工序列(Artificial Sequence)
[1771] <220>
[1772] <221> gene
[1773] <222> (1) .. (24)
[1774] <223> 人工序列的说明: 合成 寡核苷酸
[1775] <400> 146
[1776] gcagagggga gaagaaagag agat 24
[1777] <210> 147
[1778] <211> 20
[1779] <212> DNA
[1780] <213> 人工序列(Artificial Sequence)
[1781] <220>
[1782] <221> gene
[1783] <222> (1) .. (20)
[1784] <223> 人工序列的说明: 合成 寡核苷酸
[1785] <400> 147
[1786] gcaccgaggc atctctgcac 20
[1787] <210> 148
[1788] <211> 24
[1789] <212> DNA
[1790] <213> 人工序列(Artificial Sequence)
[1791] <220>
[1792] <221> gene
[1793] <222> (1) .. (24)
[1794] <223> 人工序列的说明: 合成 寡核苷酸
[1795] <400> 148
[1796] gcaccgaggc atctctgcac cgag 24
[1797] <210> 149
[1798] <211> 20
[1799] <212> DNA
[1800] <213> 人工序列(Artificial Sequence)
[1801] <220>
[1802] <221> gene
[1803] <222> (1) .. (20)
[1804] <223> 人工序列的说明: 合成 寡核苷酸
[1805] <400> 149

- [1806] gagatgtagg gctagagggg 20
[1807] <210> 150
[1808] <211> 24
[1809] <212> DNA
[1810] <213> 人工序列(Artificial Sequence)
[1811] <220>
[1812] <221> gene
[1813] <222> (1) .. (24)
[1814] <223> 人工序列的说明: 合成 寡核苷酸
[1815] <400> 150
[1816] gagatgtagg gctagagggg tgag 24
[1817] <210> 151
[1818] <211> 20
[1819] <212> DNA
[1820] <213> 人工序列(Artificial Sequence)
[1821] <220>
[1822] <221> gene
[1823] <222> (1) .. (20)
[1824] <223> 人工序列的说明: 合成 寡核苷酸
[1825] <400> 151
[1826] gtgcggcaag agcttcagcc 20
[1827] <210> 152
[1828] <211> 24
[1829] <212> DNA
[1830] <213> 人工序列(Artificial Sequence)
[1831] <220>
[1832] <221> gene
[1833] <222> (1) .. (24)
[1834] <223> 人工序列的说明: 合成 寡核苷酸
[1835] <400> 152
[1836] gtgcggcaag agcttcagcc gggg 24
[1837] <210> 153
[1838] <211> 20
[1839] <212> DNA
[1840] <213> 人工序列(Artificial Sequence)
[1841] <220>
[1842] <221> gene
[1843] <222> (1) .. (20)
[1844] <223> 人工序列的说明: 合成 寡核苷酸
[1845] <400> 153
[1846] ggggtggggg agtttgctcc 20
[1847] <210> 154

- [1848] <211> 24
[1849] <212> DNA
[1850] <213> 人工序列(Artificial Sequence)
[1851] <220>
[1852] <221> gene
[1853] <222> (1) .. (24)
[1854] <223> 人工序列的说明: 合成 寡核苷酸
[1855] <400> 154
[1856] gggtgggggg agtttgctcc tgga 24
[1857] <210> 155
[1858] <211> 20
[1859] <212> DNA
[1860] <213> 人工序列(Artificial Sequence)
[1861] <220>
[1862] <221> gene
[1863] <222> (1) .. (20)
[1864] <223> 人工序列的说明: 合成 寡核苷酸
[1865] <400> 155
[1866] gacccccctcc accccgcctc 20
[1867] <210> 156
[1868] <211> 24
[1869] <212> DNA
[1870] <213> 人工序列(Artificial Sequence)
[1871] <220>
[1872] <221> gene
[1873] <222> (1) .. (24)
[1874] <223> 人工序列的说明: 合成 寡核苷酸
[1875] <400> 156
[1876] gacccccctcc accccgcctc cggg 24
[1877] <210> 157
[1878] <211> 20
[1879] <212> DNA
[1880] <213> 人工序列(Artificial Sequence)
[1881] <220>
[1882] <221> gene
[1883] <222> (1) .. (20)
[1884] <223> 人工序列的说明: 合成 寡核苷酸
[1885] <400> 157
[1886] ggtgagtgag tgtgtgcgtg 20
[1887] <210> 158
[1888] <211> 24
[1889] <212> DNA

- [1890] <213> 人工序列(Artificial Sequence)
[1891] <220>
[1892] <221> gene
[1893] <222> (1) .. (24)
[1894] <223> 人工序列的说明: 合成 寡核苷酸
[1895] <400> 158
[1896] ggtgagtgag tgtgtgcgtg tggg 24
[1897] <210> 159
[1898] <211> 20
[1899] <212> DNA
[1900] <213> 人工序列(Artificial Sequence)
[1901] <220>
[1902] <221> gene
[1903] <222> (1) .. (20)
[1904] <223> 人工序列的说明: 合成 寡核苷酸
[1905] <400> 159
[1906] gcgagcagcg tcttcgagag 20
[1907] <210> 160
[1908] <211> 24
[1909] <212> DNA
[1910] <213> 人工序列(Artificial Sequence)
[1911] <220>
[1912] <221> gene
[1913] <222> (1) .. (24)
[1914] <223> 人工序列的说明: 合成 寡核苷酸
[1915] <400> 160
[1916] gcgagcagcg tcttcgagag tgag 24
[1917] <210> 161
[1918] <211> 20
[1919] <212> DNA
[1920] <213> 人工序列(Artificial Sequence)
[1921] <220>
[1922] <221> gene
[1923] <222> (1) .. (20)
[1924] <223> 人工序列的说明: 合成 寡核苷酸
[1925] <400> 161
[1926] gtgcggcaag agcttcagcc 20
[1927] <210> 162
[1928] <211> 24
[1929] <212> DNA
[1930] <213> 人工序列(Artificial Sequence)
[1931] <220>

- [1932] <221> gene
[1933] <222> (1) .. (24)
[1934] <223> 人工序列的说明: 合成 寡核苷酸
[1935] <400> 162
[1936] gtgcggcaag agcttcagcc agag 24
[1937] <210> 163
[1938] <211> 23
[1939] <212> DNA
[1940] <213> 人工序列(Artificial Sequence)
[1941] <220>
[1942] <221> gene
[1943] <222> (1) .. (23)
[1944] <223> 人工序列的说明: 合成 寡核苷酸
[1945] <400> 163
[1946] gagtccgagc agaagaagaa ggg 23
[1947] <210> 164
[1948] <211> 23
[1949] <212> DNA
[1950] <213> 人工序列(Artificial Sequence)
[1951] <220>
[1952] <221> gene
[1953] <222> (1) .. (23)
[1954] <223> 人工序列的说明: 合成 寡核苷酸
[1955] <400> 164
[1956] gtcacctcca atgactaggg tgg 23
[1957] <210> 165
[1958] <211> 23
[1959] <212> DNA
[1960] <213> 人工序列(Artificial Sequence)
[1961] <220>
[1962] <221> gene
[1963] <222> (1) .. (23)
[1964] <223> 人工序列的说明: 合成 寡核苷酸
[1965] <400> 165
[1966] ggaatccctt ctgcagcacc tgg 23
[1967] <210> 166
[1968] <211> 23
[1969] <212> DNA
[1970] <213> 人工序列(Artificial Sequence)
[1971] <220>
[1972] <221> gene
[1973] <222> (1) .. (23)

- [1974] <223> 人工序列的说明: 合成 寡核苷酸
[1975] <400> 166
[1976] gctgcagaag ggattccatg agg 23
[1977] <210> 167
[1978] <211> 23
[1979] <212> DNA
[1980] <213> 人工序列(Artificial Sequence)
[1981] <220>
[1982] <221> gene
[1983] <222> (1) .. (23)
[1984] <223> 人工序列的说明: 合成 寡核苷酸
[1985] <400> 167
[1986] ggcggctgca caaccagtgg agg 23
[1987] <210> 168
[1988] <211> 23
[1989] <212> DNA
[1990] <213> 人工序列(Artificial Sequence)
[1991] <220>
[1992] <221> gene
[1993] <222> (1) .. (23)
[1994] <223> 人工序列的说明: 合成 寡核苷酸
[1995] <400> 168
[1996] gctccagagc cgtgcgaatg ggg 23
[1997] <210> 169
[1998] <211> 23
[1999] <212> DNA
[2000] <213> 人工序列(Artificial Sequence)
[2001] <220>
[2002] <221> gene
[2003] <222> (1) .. (23)
[2004] <223> 人工序列的说明: 合成 寡核苷酸
[2005] <400> 169
[2006] gcattttcag gaggaagcga tgg 23
[2007] <210> 170
[2008] <211> 23
[2009] <212> DNA
[2010] <213> 人工序列(Artificial Sequence)
[2011] <220>
[2012] <221> gene
[2013] <222> (1) .. (23)
[2014] <223> 人工序列的说明: 合成 寡核苷酸
[2015] <400> 170

[2016] gtgcggcaag agcttcagcc ggg 23
[2017] <210> 171
[2018] <211> 23
[2019] <212> DNA
[2020] <213> 人工序列(Artificial Sequence)
[2021] <220>
[2022] <221> gene
[2023] <222> (1) .. (23)
[2024] <223> 人工序列的说明: 合成 寡核苷酸
[2025] <400> 171
[2026] gacccctcc accccgcctc cgg 23
[2027] <210> 172
[2028] <211> 23
[2029] <212> DNA
[2030] <213> 人工序列(Artificial Sequence)
[2031] <220>
[2032] <221> gene
[2033] <222> (1) .. (23)
[2034] <223> 人工序列的说明: 合成 寡核苷酸
[2035] <400> 172
[2036] ggtgagtgag tgtgtgcgtg tgg 23
[2037] <210> 173
[2038] <211> 20
[2039] <212> DNA
[2040] <213> 人工序列(Artificial Sequence)
[2041] <220>
[2042] <221> gene
[2043] <222> (1) .. (20)
[2044] <223> 人工序列的说明: 合成 引物
[2045] <400> 173
[2046] ggagcagctg gtcagagggg 20
[2047] <210> 174
[2048] <211> 22
[2049] <212> DNA
[2050] <213> 人工序列(Artificial Sequence)
[2051] <220>
[2052] <221> gene
[2053] <222> (1) .. (22)
[2054] <223> 人工序列的说明: 合成 引物
[2055] <400> 174
[2056] ccatagggaa gggggacact gg 22
[2057] <210> 175

[2058] <211> 20
[2059] <212> DNA
[2060] <213> 人工序列(Artificial Sequence)
[2061] <220>
[2062] <221> gene
[2063] <222> (1) .. (20)
[2064] <223> 人工序列的说明: 合成 引物
[2065] <400> 175
[2066] gggccgggaa agagttgctg 20
[2067] <210> 176
[2068] <211> 22
[2069] <212> DNA
[2070] <213> 人工序列(Artificial Sequence)
[2071] <220>
[2072] <221> gene
[2073] <222> (1) .. (22)
[2074] <223> 人工序列的说明: 合成 引物
[2075] <400> 176
[2076] gccctacatc tgctctccct cc 22
[2077] <210> 177
[2078] <211> 25
[2079] <212> DNA
[2080] <213> 人工序列(Artificial Sequence)
[2081] <220>
[2082] <221> gene
[2083] <222> (1) .. (25)
[2084] <223> 人工序列的说明: 合成 引物
[2085] <400> 177
[2086] ccagcacaac ttactcgcac ttgac 25
[2087] <210> 178
[2088] <211> 22
[2089] <212> DNA
[2090] <213> 人工序列(Artificial Sequence)
[2091] <220>
[2092] <221> gene
[2093] <222> (1) .. (22)
[2094] <223> 人工序列的说明: 合成 引物
[2095] <400> 178
[2096] catcaccaac ccacagccaa gg 22
[2097] <210> 179
[2098] <211> 20
[2099] <212> DNA

- [2100] <213> 人工序列(Artificial Sequence)
[2101] <220>
[2102] <221> gene
[2103] <222> (1) .. (20)
[2104] <223> 人工序列的说明: 合成 引物
[2105] <400> 179
[2106] tccagatggc acattgtcag 20
[2107] <210> 180
[2108] <211> 20
[2109] <212> DNA
[2110] <213> 人工序列(Artificial Sequence)
[2111] <220>
[2112] <221> gene
[2113] <222> (1) .. (20)
[2114] <223> 人工序列的说明: 合成 引物
[2115] <400> 180
[2116] agggagcagg aaagtgaggt 20
[2117] <210> 181
[2118] <211> 22
[2119] <212> DNA
[2120] <213> 人工序列(Artificial Sequence)
[2121] <220>
[2122] <221> gene
[2123] <222> (1) .. (22)
[2124] <223> 人工序列的说明: 合成 引物
[2125] <400> 181
[2126] cgaggaagag agagacgggg tc 22
[2127] <210> 182
[2128] <211> 23
[2129] <212> DNA
[2130] <213> 人工序列(Artificial Sequence)
[2131] <220>
[2132] <221> gene
[2133] <222> (1) .. (23)
[2134] <223> 人工序列的说明: 合成 引物
[2135] <400> 182
[2136] ctccaatgca cccaagacag cag 23
[2137] <210> 183
[2138] <211> 20
[2139] <212> DNA
[2140] <213> 人工序列(Artificial Sequence)
[2141] <220>

- [2142] <221> gene
[2143] <222> (1) .. (20)
[2144] <223> 人工序列的说明: 合成 引物
[2145] <400> 183
[2146] agtgtggggt gtgtgggaag 20
[2147] <210> 184
[2148] <211> 20
[2149] <212> DNA
[2150] <213> 人工序列(Artificial Sequence)
[2151] <220>
[2152] <221> gene
[2153] <222> (1) .. (20)
[2154] <223> 人工序列的说明: 合成 引物
[2155] <400> 184
[2156] gcaaggggaa gactctggca 20
[2157] <210> 185
[2158] <211> 19
[2159] <212> DNA
[2160] <213> 人工序列(Artificial Sequence)
[2161] <220>
[2162] <221> gene
[2163] <222> (1) .. (19)
[2164] <223> 人工序列的说明: 合成 引物
[2165] <400> 185
[2166] tacgagtgcc tagagtgcg 19
[2167] <210> 186
[2168] <211> 22
[2169] <212> DNA
[2170] <213> 人工序列(Artificial Sequence)
[2171] <220>
[2172] <221> gene
[2173] <222> (1) .. (22)
[2174] <223> 人工序列的说明: 合成 引物
[2175] <400> 186
[2176] gcagatgtag gtcttggagg ac 22
[2177] <210> 187
[2178] <211> 20
[2179] <212> DNA
[2180] <213> 人工序列(Artificial Sequence)
[2181] <220>
[2182] <221> gene
[2183] <222> (1) .. (20)

- [2184] <223> 人工序列的说明: 合成 引物
[2185] <400> 187
[2186] ggagcagctg gtcagagggg 20
[2187] <210> 188
[2188] <211> 22
[2189] <212> DNA
[2190] <213> 人工序列(Artificial Sequence)
[2191] <220>
[2192] <221> gene
[2193] <222> (1) .. (22)
[2194] <223> 人工序列的说明: 合成 引物
[2195] <400> 188
[2196] cgatgtcctc cccattggcc tg 22
[2197] <210> 189
[2198] <211> 25
[2199] <212> DNA
[2200] <213> 人工序列(Artificial Sequence)
[2201] <220>
[2202] <221> gene
[2203] <222> (1) .. (25)
[2204] <223> 人工序列的说明: 合成 引物
[2205] <400> 189
[2206] gtggggagat ttgcatctgt ggagg 25
[2207] <210> 190
[2208] <211> 26
[2209] <212> DNA
[2210] <213> 人工序列(Artificial Sequence)
[2211] <220>
[2212] <221> gene
[2213] <222> (1) .. (26)
[2214] <223> 人工序列的说明: 合成 引物
[2215] <400> 190
[2216] gcttttatac catcttgggg ttacag 26
[2217] <210> 191
[2218] <211> 24
[2219] <212> DNA
[2220] <213> 人工序列(Artificial Sequence)
[2221] <220>
[2222] <221> gene
[2223] <222> (1) .. (24)
[2224] <223> 人工序列的说明: 合成 引物
[2225] <400> 191

[2226] caatgtgctt caacccatca cggc 24
[2227] <210> 192
[2228] <211> 27
[2229] <212> DNA
[2230] <213> 人工序列(Artificial Sequence)
[2231] <220>
[2232] <221> gene
[2233] <222> (1) .. (27)
[2234] <223> 人工序列的说明: 合成 引物
[2235] <400> 192
[2236] ccatgaattt gtgatggatg cagtctg 27
[2237] <210> 193
[2238] <211> 24
[2239] <212> DNA
[2240] <213> 人工序列(Artificial Sequence)
[2241] <220>
[2242] <221> gene
[2243] <222> (1) .. (24)
[2244] <223> 人工序列的说明: 合成 引物
[2245] <400> 193
[2246] gagaaggagg tgcaggagct agac 24
[2247] <210> 194
[2248] <211> 24
[2249] <212> DNA
[2250] <213> 人工序列(Artificial Sequence)
[2251] <220>
[2252] <221> gene
[2253] <222> (1) .. (24)
[2254] <223> 人工序列的说明: 合成 引物
[2255] <400> 194
[2256] catcccgacc ttcatccctc ctgg 24
[2257] <210> 195
[2258] <211> 25
[2259] <212> DNA
[2260] <213> 人工序列(Artificial Sequence)
[2261] <220>
[2262] <221> gene
[2263] <222> (1) .. (25)
[2264] <223> 人工序列的说明: 合成 引物
[2265] <400> 195
[2266] gtagttctga cattcctcct gaggg 25
[2267] <210> 196

[2268] <211> 23
[2269] <212> DNA
[2270] <213> 人工序列(Artificial Sequence)
[2271] <220>
[2272] <221> gene
[2273] <222> (1) .. (23)
[2274] <223> 人工序列的说明: 合成 引物
[2275] <400> 196
[2276] tcaaacaagg tgcagataca gca 23
[2277] <210> 197
[2278] <211> 22
[2279] <212> DNA
[2280] <213> 人工序列(Artificial Sequence)
[2281] <220>
[2282] <221> gene
[2283] <222> (1) .. (22)
[2284] <223> 人工序列的说明: 合成 引物
[2285] <400> 197
[2286] cagggtcgct cagtctgtgt gg 22
[2287] <210> 198
[2288] <211> 24
[2289] <212> DNA
[2290] <213> 人工序列(Artificial Sequence)
[2291] <220>
[2292] <221> gene
[2293] <222> (1) .. (24)
[2294] <223> 人工序列的说明: 合成 引物
[2295] <400> 198
[2296] ccagcgcacc attcactcca cctg 24
[2297] <210> 199
[2298] <211> 25
[2299] <212> DNA
[2300] <213> 人工序列(Artificial Sequence)
[2301] <220>
[2302] <221> gene
[2303] <222> (1) .. (25)
[2304] <223> 人工序列的说明: 合成 引物
[2305] <400> 199
[2306] ggctgaagag gaagaccaga ctcag 25
[2307] <210> 200
[2308] <211> 25
[2309] <212> DNA

[2310] <213> 人工序列(Artificial Sequence)
[2311] <220>
[2312] <221> gene
[2313] <222> (1) .. (25)
[2314] <223> 人工序列的说明: 合成 引物
[2315] <400> 200
[2316] ggccccctctg aattcaattc tctgc 25
[2317] <210> 201
[2318] <211> 21
[2319] <212> DNA
[2320] <213> 人工序列(Artificial Sequence)
[2321] <220>
[2322] <221> gene
[2323] <222> (1) .. (21)
[2324] <223> 人工序列的说明: 合成 引物
[2325] <400> 201
[2326] ccacagcgag gaggacagc c 21
[2327] <210> 202
[2328] <211> 26
[2329] <212> DNA
[2330] <213> 人工序列(Artificial Sequence)
[2331] <220>
[2332] <221> gene
[2333] <222> (1) .. (26)
[2334] <223> 人工序列的说明: 合成 引物
[2335] <400> 202
[2336] ccaagtcttt cctaactcga ccttgg 26
[2337] <210> 203
[2338] <211> 22
[2339] <212> DNA
[2340] <213> 人工序列(Artificial Sequence)
[2341] <220>
[2342] <221> gene
[2343] <222> (1) .. (22)
[2344] <223> 人工序列的说明: 合成 引物
[2345] <400> 203
[2346] ccctaggccc acaccagcaa tg 22
[2347] <210> 204
[2348] <211> 24
[2349] <212> DNA
[2350] <213> 人工序列(Artificial Sequence)
[2351] <220>

- [2352] <221> gene
[2353] <222> (1) .. (24)
[2354] <223> 人工序列的说明: 合成 引物
[2355] <400> 204
[2356] gggatgggaa tgggaatgtg aggc 24
[2357] <210> 205
[2358] <211> 22
[2359] <212> DNA
[2360] <213> 人工序列(Artificial Sequence)
[2361] <220>
[2362] <221> gene
[2363] <222> (1) .. (22)
[2364] <223> 人工序列的说明: 合成 引物
[2365] <400> 205
[2366] gcccaggtga aggtgtggtt cc 22
[2367] <210> 206
[2368] <211> 21
[2369] <212> DNA
[2370] <213> 人工序列(Artificial Sequence)
[2371] <220>
[2372] <221> gene
[2373] <222> (1) .. (21)
[2374] <223> 人工序列的说明: 合成 引物
[2375] <400> 206
[2376] ccaaagcctg gccaggagt g 21
[2377] <210> 207
[2378] <211> 25
[2379] <212> DNA
[2380] <213> 人工序列(Artificial Sequence)
[2381] <220>
[2382] <221> gene
[2383] <222> (1) .. (25)
[2384] <223> 人工序列的说明: 合成 引物
[2385] <400> 207
[2386] aggcaaagat ctaggacctg gatgg 25
[2387] <210> 208
[2388] <211> 24
[2389] <212> DNA
[2390] <213> 人工序列(Artificial Sequence)
[2391] <220>
[2392] <221> gene
[2393] <222> (1) .. (24)

- [2394] <223> 人工序列的说明: 合成 引物
[2395] <400> 208
[2396] ccattctgagt cagccagcct tgtc 24
[2397] <210> 209
[2398] <211> 21
[2399] <212> DNA
[2400] <213> 人工序列(Artificial Sequence)
[2401] <220>
[2402] <221> gene
[2403] <222> (1) .. (21)
[2404] <223> 人工序列的说明: 合成 引物
[2405] <400> 209
[2406] ggttcctcc cttctgagcc c 21
[2407] <210> 210
[2408] <211> 25
[2409] <212> DNA
[2410] <213> 人工序列(Artificial Sequence)
[2411] <220>
[2412] <221> gene
[2413] <222> (1) .. (25)
[2414] <223> 人工序列的说明: 合成 引物
[2415] <400> 210
[2416] ggataggaat gaagaccccc tctcc 25
[2417] <210> 211
[2418] <211> 25
[2419] <212> DNA
[2420] <213> 人工序列(Artificial Sequence)
[2421] <220>
[2422] <221> gene
[2423] <222> (1) .. (25)
[2424] <223> 人工序列的说明: 合成 引物
[2425] <400> 211
[2426] ggactggctg gctgtgtgtt ttgag 25
[2427] <210> 212
[2428] <211> 24
[2429] <212> DNA
[2430] <213> 人工序列(Artificial Sequence)
[2431] <220>
[2432] <221> gene
[2433] <222> (1) .. (24)
[2434] <223> 人工序列的说明: 合成 引物
[2435] <400> 212

[2436] cttatccagg gctacctcat tgcc 24
[2437] <210> 213
[2438] <211> 25
[2439] <212> DNA
[2440] <213> 人工序列(Artificial Sequence)
[2441] <220>
[2442] <221> gene
[2443] <222> (1) .. (25)
[2444] <223> 人工序列的说明: 合成 引物
[2445] <400> 213
[2446] gctgctgctg ctttgatcac tcctg 25
[2447] <210> 214
[2448] <211> 24
[2449] <212> DNA
[2450] <213> 人工序列(Artificial Sequence)
[2451] <220>
[2452] <221> gene
[2453] <222> (1) .. (24)
[2454] <223> 人工序列的说明: 合成 引物
[2455] <400> 214
[2456] ctccttaaac cctcagaagc tggc 24
[2457] <210> 215
[2458] <211> 23
[2459] <212> DNA
[2460] <213> 人工序列(Artificial Sequence)
[2461] <220>
[2462] <221> gene
[2463] <222> (1) .. (23)
[2464] <223> 人工序列的说明: 合成 引物
[2465] <400> 215
[2466] gcactgtcag ctgacctac agg 23
[2467] <210> 216
[2468] <211> 24
[2469] <212> DNA
[2470] <213> 人工序列(Artificial Sequence)
[2471] <220>
[2472] <221> gene
[2473] <222> (1) .. (24)
[2474] <223> 人工序列的说明: 合成 引物
[2475] <400> 216
[2476] acgttggaac agtcgagctg tagc 24
[2477] <210> 217

[2478] <211> 25
[2479] <212> DNA
[2480] <213> 人工序列(Artificial Sequence)
[2481] <220>
[2482] <221> gene
[2483] <222> (1) .. (25)
[2484] <223> 人工序列的说明: 合成 引物
[2485] <400> 217
[2486] tgtgcataac tcatgttggc aaact 25
[2487] <210> 218
[2488] <211> 23
[2489] <212> DNA
[2490] <213> 人工序列(Artificial Sequence)
[2491] <220>
[2492] <221> gene
[2493] <222> (1) .. (23)
[2494] <223> 人工序列的说明: 合成 引物
[2495] <400> 218
[2496] tccacaacta ccctcagctg gag 23
[2497] <210> 219
[2498] <211> 25
[2499] <212> DNA
[2500] <213> 人工序列(Artificial Sequence)
[2501] <220>
[2502] <221> gene
[2503] <222> (1) .. (25)
[2504] <223> 人工序列的说明: 合成 引物
[2505] <400> 219
[2506] ccactgacaa ttactcaac cctgc 25
[2507] <210> 220
[2508] <211> 24
[2509] <212> DNA
[2510] <213> 人工序列(Artificial Sequence)
[2511] <220>
[2512] <221> gene
[2513] <222> (1) .. (24)
[2514] <223> 人工序列的说明: 合成 引物
[2515] <400> 220
[2516] aggcagacca gttatttggc agtc 24
[2517] <210> 221
[2518] <211> 22
[2519] <212> DNA

[2520] <213> 人工序列(Artificial Sequence)
[2521] <220>
[2522] <221> gene
[2523] <222> (1) .. (22)
[2524] <223> 人工序列的说明: 合成 引物
[2525] <400> 221
[2526] acaggcgcag ttactgaga ag 22
[2527] <210> 222
[2528] <211> 22
[2529] <212> DNA
[2530] <213> 人工序列(Artificial Sequence)
[2531] <220>
[2532] <221> gene
[2533] <222> (1) .. (22)
[2534] <223> 人工序列的说明: 合成 引物
[2535] <400> 222
[2536] gggtaggctg actttgggct cc 22
[2537] <210> 223
[2538] <211> 22
[2539] <212> DNA
[2540] <213> 人工序列(Artificial Sequence)
[2541] <220>
[2542] <221> gene
[2543] <222> (1) .. (22)
[2544] <223> 人工序列的说明: 合成 引物
[2545] <400> 223
[2546] gccctcttgc ctccactggt tg 22
[2547] <210> 224
[2548] <211> 23
[2549] <212> DNA
[2550] <213> 人工序列(Artificial Sequence)
[2551] <220>
[2552] <221> gene
[2553] <222> (1) .. (23)
[2554] <223> 人工序列的说明: 合成 引物
[2555] <400> 224
[2556] cgcggatgtt ccaatcagta cgc 23
[2557] <210> 225
[2558] <211> 22
[2559] <212> DNA
[2560] <213> 人工序列(Artificial Sequence)
[2561] <220>

[2562] <221> gene
[2563] <222> (1) .. (22)
[2564] <223> 人工序列的说明: 合成 引物
[2565] <400> 225
[2566] gcgggcagtg gcgtcttagt cg 22
[2567] <210> 226
[2568] <211> 22
[2569] <212> DNA
[2570] <213> 人工序列(Artificial Sequence)
[2571] <220>
[2572] <221> gene
[2573] <222> (1) .. (22)
[2574] <223> 人工序列的说明: 合成 引物
[2575] <400> 226
[2576] ccctgggttt gggttgctgc tc 22
[2577] <210> 227
[2578] <211> 21
[2579] <212> DNA
[2580] <213> 人工序列(Artificial Sequence)
[2581] <220>
[2582] <221> gene
[2583] <222> (1) .. (21)
[2584] <223> 人工序列的说明: 合成 引物
[2585] <400> 227
[2586] ctccttgccg cccagccggt c 21
[2587] <210> 228
[2588] <211> 23
[2589] <212> DNA
[2590] <213> 人工序列(Artificial Sequence)
[2591] <220>
[2592] <221> gene
[2593] <222> (1) .. (23)
[2594] <223> 人工序列的说明: 合成 引物
[2595] <400> 228
[2596] cactggggaa gaggcgagga cac 23
[2597] <210> 229
[2598] <211> 23
[2599] <212> DNA
[2600] <213> 人工序列(Artificial Sequence)
[2601] <220>
[2602] <221> gene
[2603] <222> (1) .. (23)

- [2604] <223> 人工序列的说明: 合成 引物
[2605] <400> 229
[2606] ccagtgtttc ccatccccaa cac 23
[2607] <210> 230
[2608] <211> 23
[2609] <212> DNA
[2610] <213> 人工序列(Artificial Sequence)
[2611] <220>
[2612] <221> gene
[2613] <222> (1) .. (23)
[2614] <223> 人工序列的说明: 合成 引物
[2615] <400> 230
[2616] gaatggatcc ccccctagag ctc 23
[2617] <210> 231
[2618] <211> 22
[2619] <212> DNA
[2620] <213> 人工序列(Artificial Sequence)
[2621] <220>
[2622] <221> gene
[2623] <222> (1) .. (22)
[2624] <223> 人工序列的说明: 合成 引物
[2625] <400> 231
[2626] caggcccaca ggtccttctg ga 22
[2627] <210> 232
[2628] <211> 21
[2629] <212> DNA
[2630] <213> 人工序列(Artificial Sequence)
[2631] <220>
[2632] <221> gene
[2633] <222> (1) .. (21)
[2634] <223> 人工序列的说明: 合成 引物
[2635] <400> 232
[2636] ccacacggaa ggctgaccac g 21
[2637] <210> 233
[2638] <211> 21
[2639] <212> DNA
[2640] <213> 人工序列(Artificial Sequence)
[2641] <220>
[2642] <221> gene
[2643] <222> (1) .. (21)
[2644] <223> 人工序列的说明: 合成 引物
[2645] <400> 233

[2646] gcgcagagag agcaggacgt c 21
[2647] <210> 234
[2648] <211> 23
[2649] <212> DNA
[2650] <213> 人工序列(Artificial Sequence)
[2651] <220>
[2652] <221> gene
[2653] <222> (1) .. (23)
[2654] <223> 人工序列的说明: 合成 引物
[2655] <400> 234
[2656] gcacctcatg gaatcccttc tgc 23
[2657] <210> 235
[2658] <211> 23
[2659] <212> DNA
[2660] <213> 人工序列(Artificial Sequence)
[2661] <220>
[2662] <221> gene
[2663] <222> (1) .. (23)
[2664] <223> 人工序列的说明: 合成 引物
[2665] <400> 235
[2666] caagtgatgc gacttccaac ctc 23
[2667] <210> 236
[2668] <211> 26
[2669] <212> DNA
[2670] <213> 人工序列(Artificial Sequence)
[2671] <220>
[2672] <221> gene
[2673] <222> (1) .. (26)
[2674] <223> 人工序列的说明: 合成 引物
[2675] <400> 236
[2676] ccctcagagt tcagcttaaa aagacc 26
[2677] <210> 237
[2678] <211> 25
[2679] <212> DNA
[2680] <213> 人工序列(Artificial Sequence)
[2681] <220>
[2682] <221> gene
[2683] <222> (1) .. (25)
[2684] <223> 人工序列的说明: 合成 引物
[2685] <400> 237
[2686] tgcttctcat ccactctaga ctgct 25
[2687] <210> 238

[2688] <211> 22
[2689] <212> DNA
[2690] <213> 人工序列(Artificial Sequence)
[2691] <220>
[2692] <221> gene
[2693] <222> (1) .. (22)
[2694] <223> 人工序列的说明: 合成 引物
[2695] <400> 238
[2696] caccaaccag ccatgtgcca tg 22
[2697] <210> 239
[2698] <211> 22
[2699] <212> DNA
[2700] <213> 人工序列(Artificial Sequence)
[2701] <220>
[2702] <221> gene
[2703] <222> (1) .. (22)
[2704] <223> 人工序列的说明: 合成 引物
[2705] <400> 239
[2706] ctgcctgtgc tcctcgatgg tg 22
[2707] <210> 240
[2708] <211> 22
[2709] <212> DNA
[2710] <213> 人工序列(Artificial Sequence)
[2711] <220>
[2712] <221> gene
[2713] <222> (1) .. (22)
[2714] <223> 人工序列的说明: 合成 引物
[2715] <400> 240
[2716] gggttcaaag ctcatctgcc cc 22
[2717] <210> 241
[2718] <211> 23
[2719] <212> DNA
[2720] <213> 人工序列(Artificial Sequence)
[2721] <220>
[2722] <221> gene
[2723] <222> (1) .. (23)
[2724] <223> 人工序列的说明: 合成 引物
[2725] <400> 241
[2726] gcatgtgcct tgagattgcc tgg 23
[2727] <210> 242
[2728] <211> 25
[2729] <212> DNA

- [2730] <213> 人工序列(Artificial Sequence)
[2731] <220>
[2732] <221> gene
[2733] <222> (1) .. (25)
[2734] <223> 人工序列的说明: 合成 引物
[2735] <400> 242
[2736] gacattcaga gaagcgacca tgtgg 25
[2737] <210> 243
[2738] <211> 23
[2739] <212> DNA
[2740] <213> 人工序列(Artificial Sequence)
[2741] <220>
[2742] <221> gene
[2743] <222> (1) .. (23)
[2744] <223> 人工序列的说明: 合成 引物
[2745] <400> 243
[2746] ccatcttccc ctttggccca cag 23
[2747] <210> 244
[2748] <211> 25
[2749] <212> DNA
[2750] <213> 人工序列(Artificial Sequence)
[2751] <220>
[2752] <221> gene
[2753] <222> (1) .. (25)
[2754] <223> 人工序列的说明: 合成 引物
[2755] <400> 244
[2756] ccccaaaagt ggccaagagc ctgag 25
[2757] <210> 245
[2758] <211> 26
[2759] <212> DNA
[2760] <213> 人工序列(Artificial Sequence)
[2761] <220>
[2762] <221> gene
[2763] <222> (1) .. (26)
[2764] <223> 人工序列的说明: 合成 引物
[2765] <400> 245
[2766] gttctccaaa ggaagagagg ggaatg 26
[2767] <210> 246
[2768] <211> 22
[2769] <212> DNA
[2770] <213> 人工序列(Artificial Sequence)
[2771] <220>

[2772] <221> gene
[2773] <222> (1) .. (22)
[2774] <223> 人工序列的说明: 合成 引物
[2775] <400> 246
[2776] ggtgctgtgt cctcatgcat cc 22
[2777] <210> 247
[2778] <211> 22
[2779] <212> DNA
[2780] <213> 人工序列(Artificial Sequence)
[2781] <220>
[2782] <221> gene
[2783] <222> (1) .. (22)
[2784] <223> 人工序列的说明: 合成 引物
[2785] <400> 247
[2786] cggccttcct agggctcgtt ag 22
[2787] <210> 248
[2788] <211> 22
[2789] <212> DNA
[2790] <213> 人工序列(Artificial Sequence)
[2791] <220>
[2792] <221> gene
[2793] <222> (1) .. (22)
[2794] <223> 人工序列的说明: 合成 引物
[2795] <400> 248
[2796] ccttcagggg ctcttcagg tc 22
[2797] <210> 249
[2798] <211> 21
[2799] <212> DNA
[2800] <213> 人工序列(Artificial Sequence)
[2801] <220>
[2802] <221> gene
[2803] <222> (1) .. (21)
[2804] <223> 人工序列的说明: 合成 引物
[2805] <400> 249
[2806] gggaactggc aggcaccgag g 21
[2807] <210> 250
[2808] <211> 22
[2809] <212> DNA
[2810] <213> 人工序列(Artificial Sequence)
[2811] <220>
[2812] <221> gene
[2813] <222> (1) .. (22)

- [2814] <223> 人工序列的说明: 合成 引物
[2815] <400> 250
[2816] gggtgaggct gaaacagtga cc 22
[2817] <210> 251
[2818] <211> 26
[2819] <212> DNA
[2820] <213> 人工序列(Artificial Sequence)
[2821] <220>
[2822] <221> gene
[2823] <222> (1) .. (26)
[2824] <223> 人工序列的说明: 合成 引物
[2825] <400> 251
[2826] gggaggatgt tggttttagg gaactg 26
[2827] <210> 252
[2828] <211> 27
[2829] <212> DNA
[2830] <213> 人工序列(Artificial Sequence)
[2831] <220>
[2832] <221> gene
[2833] <222> (1) .. (27)
[2834] <223> 人工序列的说明: 合成 引物
[2835] <400> 252
[2836] tccaatcact acatgccatt ttgaaga 27
[2837] <210> 253
[2838] <211> 24
[2839] <212> DNA
[2840] <213> 人工序列(Artificial Sequence)
[2841] <220>
[2842] <221> GC_signal
[2843] <222> (1) .. (24)
[2844] <223> 人工序列的说明: 合成 引物
[2845] <400> 253
[2846] ccaccctctt cctttgatcc tccc 24
[2847] <210> 254
[2848] <211> 23
[2849] <212> DNA
[2850] <213> 人工序列(Artificial Sequence)
[2851] <220>
[2852] <221> gene
[2853] <222> (1) .. (23)
[2854] <223> 人工序列的说明: 合成 引物
[2855] <400> 254

[2856] tcctccctac tccttcaccc agg 23
[2857] <210> 255
[2858] <211> 23
[2859] <212> DNA
[2860] <213> 人工序列(Artificial Sequence)
[2861] <220>
[2862] <221> gene
[2863] <222> (1) .. (23)
[2864] <223> 人工序列的说明: 合成 引物
[2865] <400> 255
[2866] gagtgcctga catgtgggga gag 23
[2867] <210> 256
[2868] <211> 23
[2869] <212> DNA
[2870] <213> 人工序列(Artificial Sequence)
[2871] <220>
[2872] <221> gene
[2873] <222> (1) .. (23)
[2874] <223> 人工序列的说明: 合成 引物
[2875] <400> 256
[2876] tccagctaaa gcctttccca cac 23
[2877] <210> 257
[2878] <211> 26
[2879] <212> DNA
[2880] <213> 人工序列(Artificial Sequence)
[2881] <220>
[2882] <221> gene
[2883] <222> (1) .. (26)
[2884] <223> 人工序列的说明: 合成 引物
[2885] <400> 257
[2886] gaactctctg atgcacctga aggctg 26
[2887] <210> 258
[2888] <211> 25
[2889] <212> DNA
[2890] <213> 人工序列(Artificial Sequence)
[2891] <220>
[2892] <221> gene
[2893] <222> (1) .. (25)
[2894] <223> 人工序列的说明: 合成 引物
[2895] <400> 258
[2896] accgtatcag tgtgatgcat gtggt 25
[2897] <210> 259

[2898] <211> 28
[2899] <212> DNA
[2900] <213> 人工序列(Artificial Sequence)
[2901] <220>
[2902] <221> gene
[2903] <222> (1) .. (28)
[2904] <223> 人工序列的说明: 合成 引物
[2905] <400> 259
[2906] tgggtttaat catgtgttct gcactatg 28
[2907] <210> 260
[2908] <211> 24
[2909] <212> DNA
[2910] <213> 人工序列(Artificial Sequence)
[2911] <220>
[2912] <221> gene
[2913] <222> (1) .. (24)
[2914] <223> 人工序列的说明: 合成 引物
[2915] <400> 260
[2916] cccatcttcc attctgccct ccac 24
[2917] <210> 261
[2918] <211> 28
[2919] <212> DNA
[2920] <213> 人工序列(Artificial Sequence)
[2921] <220>
[2922] <221> gene
[2923] <222> (1) .. (28)
[2924] <223> 人工序列的说明: 合成 引物
[2925] <400> 261
[2926] cagctagtcc atttgttctc agactgtg 28
[2927] <210> 262
[2928] <211> 25
[2929] <212> DNA
[2930] <213> 人工序列(Artificial Sequence)
[2931] <220>
[2932] <221> gene
[2933] <222> (1) .. (25)
[2934] <223> 人工序列的说明: 合成 引物
[2935] <400> 262
[2936] ggccaacatt gtgaaaccct gtctc 25
[2937] <210> 263
[2938] <211> 22
[2939] <212> DNA

[2940] <213> 人工序列(Artificial Sequence)
[2941] <220>
[2942] <221> gene
[2943] <222> (1) .. (22)
[2944] <223> 人工序列的说明: 合成 引物
[2945] <400> 263
[2946] ccaggacact gtgcttggt tc 22
[2947] <210> 264
[2948] <211> 23
[2949] <212> DNA
[2950] <213> 人工序列(Artificial Sequence)
[2951] <220>
[2952] <221> gene
[2953] <222> (1) .. (23)
[2954] <223> 人工序列的说明: 合成 引物
[2955] <400> 264
[2956] caccatga cctggcaca gtg 23
[2957] <210> 265
[2958] <211> 24
[2959] <212> DNA
[2960] <213> 人工序列(Artificial Sequence)
[2961] <220>
[2962] <221> gene
[2963] <222> (1) .. (24)
[2964] <223> 人工序列的说明: 合成 引物
[2965] <400> 265
[2966] aagtgtcct cagaatcca gccc 24
[2967] <210> 266
[2968] <211> 23
[2969] <212> DNA
[2970] <213> 人工序列(Artificial Sequence)
[2971] <220>
[2972] <221> gene
[2973] <222> (1) .. (23)
[2974] <223> 人工序列的说明: 合成 引物
[2975] <400> 266
[2976] caggagtga gttgtgttg gag 23
[2977] <210> 267
[2978] <211> 25
[2979] <212> DNA
[2980] <213> 人工序列(Artificial Sequence)
[2981] <220>

- [2982] <221> gene
[2983] <222> (1) .. (25)
[2984] <223> 人工序列的说明: 合成 引物
[2985] <400> 267
[2986] ctgatgaagc accagagaac ccacc 25
[2987] <210> 268
[2988] <211> 21
[2989] <212> DNA
[2990] <213> 人工序列(Artificial Sequence)
[2991] <220>
[2992] <221> gene
[2993] <222> (1) .. (21)
[2994] <223> 人工序列的说明: 合成 引物
[2995] <400> 268
[2996] cacacctggc acccatatgg c 21
[2997] <210> 269
[2998] <211> 25
[2999] <212> DNA
[3000] <213> 人工序列(Artificial Sequence)
[3001] <220>
[3002] <221> gene
[3003] <222> (1) .. (25)
[3004] <223> 人工序列的说明: 合成 引物
[3005] <400> 269
[3006] gatccacact ggtgagaagc cttac 25
[3007] <210> 270
[3008] <211> 26
[3009] <212> DNA
[3010] <213> 人工序列(Artificial Sequence)
[3011] <220>
[3012] <221> gene
[3013] <222> (1) .. (26)
[3014] <223> 人工序列的说明: 合成 引物
[3015] <400> 270
[3016] cttcccacac tcacagcaga tgtagg 26
[3017] <210> 271
[3018] <211> 4206
[3019] <212> DNA
[3020] <213> 人工序列(Artificial Sequence)
[3021] <220>
[3022] <221> gene
[3023] <222> (1) .. (4206)

[3024]	<223> 人工序列的说明: 合成 多核苷酸					
[3025]	<400> 271					
[3026]	atggataaaa	agtattctat	tggtttagac	atcggcacta	attccgttgg	atgggctgtc 60
[3027]	ataaccgatg	aatacaaagt	accttcaaag	aaatttaagg	tgttggggaa	cacagaccgt 120
[3028]	cattcgatta	aaaagaatct	tatcggtgcc	ctcctattcg	atagtggcga	aacggcagag 180
[3029]	gcgactcgcc	tgaaacgaac	cgctcggaga	aggtatacac	gtcgcaagaa	ccgaatatgt 240
[3030]	tacttacaag	aaatttttag	caatgagatg	gccaaagttg	acgattcttt	ctttcacctg 300
[3031]	ttggaagagt	ccttccttgt	cgaagaggac	aagaaacatg	aacggcaccc	catctttgga 360
[3032]	aacatagtag	atgaggtggc	atatcatgaa	aagtacccaa	cgatttatca	cctcagaaaa 420
[3033]	aagctagtgt	actcaactga	taaagcggac	ctgagggtta	tctacttggc	tcttgcccat 480
[3034]	atgataaagt	tccgtgggca	ctttctcatt	gagggtgac	taaatccgga	caactcggat 540
[3035]	gtcgacaaac	tgttcatcca	gttagtaca	acctataatc	agttgtttga	agagaacct 600
[3036]	ataaatgcaa	gtggcgtgga	tgcgaaggct	attcttagcg	cccgcctctc	taaatcccg 660
[3037]	cggctagaaa	acctgatcgc	acaattaccc	ggagagaaga	aaaatgggtt	gttcggtaac 720
[3038]	cttatagcgc	tctcactagg	cctgacacca	aattttaagt	cgaacttcga	cttagctgaa 780
[3039]	gatgccaaat	tgcagcttag	taaggacacg	tacgatgacg	atctcgacaa	tctactggca 840
[3040]	caaattggag	atcagtatgc	ggacttattt	ttggctgcca	aaaaccttag	cgatgcaatc 900
[3041]	ctcctatctg	acatactgag	agttaatact	gagattacca	aggcgccgtt	atccgcttca 960
[3042]	atgatcaaaa	ggtacgatga	acatcaccaa	gacttgacac	ttctcaaggc	cctagtccgt 1020
[3043]	cagcaactgc	ctgagaaata	taaggaaata	ttctttgac	agtcgaaaaa	cgggtacgca 1080
[3044]	ggttatattg	acggcggagc	gagtcaagag	gaattctaca	agttttatcaa	acccatatta 1140
[3045]	gagaagatgg	atgggacgga	agagttgctt	gtaaaactca	atcgcgaa	tctactgcga 1200
[3046]	aagcagcgga	ctttcgacaa	cggtagcatt	ccacatcaaa	tccacttagg	cgaattgcat 1260
[3047]	gctatactta	gaaggcagga	ggatttttat	ccgttctca	aagacaatcg	tgaaaagatt 1320
[3048]	gagaaaatcc	taacctttcg	cataccttac	tatgtgggac	ccctggcccg	agggaactct 1380
[3049]	cggttcgc	ggatgacaag	aaagtcgga	gaaacgatta	ctccatggaa	ttttgaggaa 1440
[3050]	gttgcgata	aagggtcgctc	agctcaatcg	ttcatcgaga	ggatgacca	ctttgacaag 1500
[3051]	aatttaccga	acgaaaaagt	attgcctaag	cacagtttac	tttacgagta	tttcacagt 1560
[3052]	tacaatgaac	tcacgaaagt	taagtatgtc	actgagggca	tgcgtaaacc	cgcctttcta 1620
[3053]	agcggagaac	agaagaaagc	aatagtagat	ctgttattca	agaccaaccg	caaagtgaca 1680
[3054]	gttaagcaat	tgaagagga	ctactttaag	aaaattgaat	gcttcgattc	tgctgagatc 1740
[3055]	tccgggtag	aagatcgatt	taatgcgtca	cttggtacgt	atcatgacct	cctaaagata 1800
[3056]	attaaagata	aggacttcct	ggataacgaa	gagaatgaag	atatcttaga	agatatagt 1860
[3057]	ttgactctta	ccctctttga	agatcgggaa	atgattgagg	aaagactaaa	aacatacgt 1920
[3058]	cacctgttcg	acgataaggt	tatgaaacag	ttaaagaggc	gtcgctatac	gggctgggga 1980
[3059]	cgattgtcgc	ggaaacttat	caacgggata	agagacaagc	aaagtggtaa	aactattctc 2040
[3060]	gattttctaa	agagcgacgg	cttcgccaat	aggaacttta	tgcagctgat	ccatgatgac 2100
[3061]	tctttaacct	tcaaagagga	tatacaaaag	gcacaggttt	ccggacaagg	ggactcattg 2160
[3062]	cacgaacata	ttgcgaatct	tgctggttcg	ccagccatca	aaaagggc	actccagaca 2220
[3063]	gtcaaagtag	tgatgagct	agttaaggtc	atgggacgtc	acaaaccgga	aaacattgta 2280
[3064]	atcgagatgg	cacgcgaaaa	tcaaacgact	cagaaggggc	aaaaaacag	tcgagagcgg 2340
[3065]	atgaagagaa	tagaagagg	tattaaagaa	ctgggcagcc	agatcttaaa	ggagcatcct 2400

[3066]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3067]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgctgatcac	2520
[3068]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgtc tacacgctcg	2580
[3069]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3070]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cggataactta	2700
[3071]	actaaagtct agaggggttg cttgtctgaa cttgacaagg ccgattttat taaacgtcag	2760
[3072]	ctcgtggaaa cccgccaaat cacaaagcat gttgcacaga tactagattc ccgaatgaat	2820
[3073]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3074]	aaattggtgt cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3075]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3076]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3077]	atgatcgca aaagcgaaca ggagatagc aaggctacag ccaataactt cttttattct	3120
[3078]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3079]	cctttaattg aaaccaatgg ggagacaggt gaaatcgat gggataaggg ccgggacttc	3240
[3080]	gcgacgggtgaaaaagtttt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3081]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360
[3082]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgatagccc tacagttgcc	3420
[3083]	tattctgtcc tagtagtggc aaaagttgag aagggaat ccaagaaact gaagtcagtc	3480
[3084]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3085]	ttccttgagg cgaaaggtta caaggaagta aaaaaggatc tcataattaa actaccaaag	3600
[3086]	tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt	3660
[3087]	caaaagggga acgaactcgc actaccgtct aaatacgtga atttcctgta tttagcgtcc	3720
[3088]	cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact tttgttgag	3780
[3089]	cagcacaac attatctcga cgaaatcata gagcaattt cggaattcag taagagagtc	3840
[3090]	atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa	3900
[3091]	cccatacgtg agcaggcgga aaatattatc catttggtta ctcttaccac cctcggcgt	3960
[3092]	ccagccgat tcaagtattt tgacacaacg atagatcgca aacgatacac ttctaccaag	4020
[3093]	gagtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3094]	gatttgcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3095]	tacaaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3096]	aagtga	4206
[3097]	<210>	272
[3098]	<211>	4206
[3099]	<212>	DNA
[3100]	<213>	人工序列(Artificial Sequence)
[3101]	<220>	
[3102]	<221>	gene
[3103]	<222>	(1) .. (4206)
[3104]	<223>	人工序列的说明: 合成 多核苷酸
[3105]	<400>	272
[3106]	atggataaaa agtattctat tggtttagac atcggcacta attccgttg atgggctgtc	60
[3107]	ataaccgatg aatacaaagt accttcaaag aaatttaagg tgttggggaa cacagaccgt	120

[3108]	cattcgatta aaaagaatct tatcgggtgcc ctctatttcg atagtggcga aacggcagag	180
[3109]	gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgcaagaa ccgaatatgt	240
[3110]	tacttacaag aaatTTTTtag caatgagatg gccaaagtgt acgattcttt ctttcaccgt	300
[3111]	ttggaagagt ctttccttgt cgaagaggac aagaaacatg aacggcaccc catcttttga	360
[3112]	aacatagtag atgaggtggc atatcatgaa aagtacccaa cgatttatca cctcagaaaa	420
[3113]	aagctagtgt actcaactga taaagcggac ctgagggttaa tctacttggc tcttgcccat	480
[3114]	atgataaagt tccgtgggca ctttctcatt gagggtgatc taaatccgga caactcggat	540
[3115]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaaccct	600
[3116]	ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg cccgcctctc taaatcccg	660
[3117]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3118]	cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa	780
[3119]	gatgcaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840
[3120]	caaattggag atcagtatgc ggacttatTT ttggctgcc	900
[3121]	ctcctatctg acatactgag agttaatact gagattacca aggcgcctgt atccgcttca	960
[3122]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc ctagtccgt	1020
[3123]	cagcaactgc ctgagaaata taaggaaata ttctttgatc agtcgaaaaa cgggtacgca	1080
[3124]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3125]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcgaaaga tctactgca	1200
[3126]	aagcagcggga ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3127]	gctatactta gaaggcagga ggatttttat ccgttctca aagacaatcg tgaaaagatt	1320
[3128]	gagaaaatcc taacctttcg cataccttac tatgtgggac ccctggcccc agggaaactct	1380
[3129]	cggttcgc	1440
[3130]	ggtgtcgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccgc ctttgacaag	1500
[3131]	aatttaccga acgaaaaagt attgcctaag cacagtttac tttacgagta tttcacagt	1560
[3132]	tacaatgaac tcacgaaagt taagtatgtc actgagggca tgcgtaaacc cgcctttcta	1620
[3133]	agcggagaac agaagaaagc aatagtagat ctgttattca agaccaaccg caaagtgaca	1680
[3134]	gttaagcaat tgaaagagga ctactttaag aaaattgaat gcttcgattc tgctgagatc	1740
[3135]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3136]	attaaagata aggacttctt ggataacgaa gagaatgaag atatcttaga agatatagtg	1860
[3137]	ttgactctta cctctttga agatcgggaa atgattgagg aaagactaaa aacatacgt	1920
[3138]	cacctgttcg acgataaggt tatgaaacag ttaaagaggc gtcgtatac gggctgggga	1980
[3139]	gccttgtcgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3140]	gattttctaa agagcgacgg cttcgccaat aggaacttta tggccctgat ccatgatgac	2100
[3141]	tctttaacct tcaaagagga tatacaaaag gcacaggttt ccggacaagg ggactcattg	2160
[3142]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaaggcat actccagaca	2220
[3143]	gtcaaagtag tggatgagct agttaaggtc atgggacgtc acaaaccgga aaacattgta	2280
[3144]	atcgagatgg cacgcgaaaa tcaaacgact cagaaggggc aaaaaaacag tcgagagcgg	2340
[3145]	atgaagagaa tagaagaggg tattaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3146]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3147]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgctgatcac	2520
[3148]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgtc tacacgctcg	2580
[3149]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640

[3150]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagtt cgataactta	2700
[3151]	actaaagctg agaggggtgg cttgtctgaa cttgacaagg ccgatttat taaacgtcag	2760
[3152]	ctcgtggaaa cccgcgccat cacaaagcat gttgcgcaga tactagattc ccgaatgaat	2820
[3153]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3154]	aaattggtgt cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3155]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3156]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3157]	atgatcgca aaagcgaaca ggagatagc aaggctacag ccaaatactt cttttattct	3120
[3158]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3159]	cctttaattg aaaccaatgg ggagacaggt gaaatcgtat gggataaggg ccgggacttc	3240
[3160]	gcgacggtga gaaaagtgtt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3161]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360
[3162]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgatagccc tacagttgcc	3420
[3163]	tattctgtcc tagtagtggc aaaagttgag aagggaat ccaagaaact gaagtcagtc	3480
[3164]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3165]	ttccttgagg cgaaagtta caaggaagta aaaaaggatc tcataattaa actaccaaag	3600
[3166]	tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt	3660
[3167]	caaaagggga acgaactcgc actaccgtct aaatacgtga atttcctgta tttagcgtcc	3720
[3168]	cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact tttgtttgag	3780
[3169]	cagcacaac attatctcga cgaaatcata gagcaaattt cggaattcag taagagagtc	3840
[3170]	atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa	3900
[3171]	cccatacgtg agcaggcgga aaatattatc catttggtta ctcttaccaa cctcggcgtc	3960
[3172]	ccagccgat tcaagtattt tgacacaacg atagatcgca aacgatacac ttctaccaag	4020
[3173]	gagtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3174]	gatttgcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3175]	tacaaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3176]	aagtga	4206
[3177]	<210>	273
[3178]	<211>	4206
[3179]	<212>	DNA
[3180]	<213>	人工序列(Artificial Sequence)
[3181]	<220>	
[3182]	<221>	gene
[3183]	<222>	(1) .. (4206)
[3184]	<223>	人工序列的说明: 合成 多核苷酸
[3185]	<400>	273
[3186]	atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc	60
[3187]	ataaccgatg aatacaaaag accttcaaag aaatttaagg tgttggggaa cacagaccgt	120
[3188]	cattcgatta aaaagaatct tatcggtgcc ctctatttcg atagtggcga aacggcagag	180
[3189]	gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgaagaa ccgaatatgt	240
[3190]	tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttcaccgt	300
[3191]	ttggaagagt ctttcctgt cgaagaggac aagaaacatg aacggcacc catctttgga	360

[3192]	aacatagtag atgaggtggc atatcatgaa aagtacccaa cgatttatca cctcagaaaa	420
[3193]	aagctagttag actcaactga taaagcggac ctgagggttaa tctacttggc tcttgcccat	480
[3194]	atgataaagt tccgtgggca ctttctcatt gaggggtgac taaatccgga caactcggat	540
[3195]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaaccct	600
[3196]	ataaatgcaa gtggcgtgga tgcgaaggct attccttagcg cccgcctctc taaatcccga	660
[3197]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3198]	cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa	780
[3199]	gatgccaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840
[3200]	caaattggag atcagtatgc ggacttattt ttggctgcca aaaaccttag cgatgcaatc	900
[3201]	ctcctatctg acatactgag agttaatact gagattacca aggcgccgtt atccgcttca	960
[3202]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc cctagtccgt	1020
[3203]	cagcaactgc ctgagaaata taaggaaata ttctttgac agtcgaaaaa cgggtacgca	1080
[3204]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3205]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcgaaga tctactgca	1200
[3206]	aagcagcggc ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3207]	gctatactta gaaggcagga ggatttttat cgttctctca aagacaatcg tgaagatt	1320
[3208]	gagaaaatcc taaccttctg cataccttac tatgtgggac ccctggcccg agggaaactct	1380
[3209]	cggttcgcac ggatgacaag aaagtcgaa gaaacgatta ctccctggaa ttttgaggaa	1440
[3210]	gttgcgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccgc ctttgacaag	1500
[3211]	aatttaccga acgaaaaagt attgcctaag cacagtttac ttacagta ttacacagt	1560
[3212]	tacaatgaac tcacgaaagt taagtatgac actgagggca tgcgtaaacc cgcctttcta	1620
[3213]	agcggagaac agaagaaagc aatagtagat ctgttattca agaccaaccg caaagtgaca	1680
[3214]	gttaagcaat tgaagagga ctactttaag aaaattgaat gcttcgattc tgctgagatc	1740
[3215]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3216]	attaaagata aggacttct ggataacgaa gagaatgaag atatcttaga agatatagt	1860
[3217]	ttgactctta ccctcttga agatcgggaa atgattgagg aaagactaaa aacatacgt	1920
[3218]	cacctgttcg acgataaggt tatgaaacag ttaaagaggc gtcgtatac gggctgggga	1980
[3219]	gccttgcgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3220]	gattttctaa agagcgacgg cttcgccaat aggaacttta tgccctgat ccatgatgac	2100
[3221]	tctttaacct tcaaagagga tatacaaaag gcacaggtt cggacaagg ggactcattg	2160
[3222]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaaggcat actccagaca	2220
[3223]	gtcaaagtag tggatgagct agttaaggtc atgggacgtc acaaaccgga aaacattgta	2280
[3224]	atcgagatgg cagcgaaaa tcaaacgact cagaaggggc aaaaaaacag tcgagagcgg	2340
[3225]	atgaagagaa tagaagaggg tattaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3226]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3227]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgtcgatcac	2520
[3228]	attgtacccc aatcctttt gaaggacgat tcaatcgaca ataaagtgt tacacgctcg	2580
[3229]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3230]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cgataactta	2700
[3231]	actaaagctg agaggggtgg cttgtctgaa cttgacaagg ccgatttat taaacgtcag	2760
[3232]	ctcgtggaac cccgcgccat cacaaagcat gttgcgcaga tactagattc ccgaatgaat	2820
[3233]	acgaaatagc acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880

[3234]	aaattggtgt cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3235]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3236]	tacccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3237]	atgatcgcga aaagcgaaca ggagataggc aaggctacag ccaaatactt cttttattct	3120
[3238]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3239]	cctttaattg aaaccaatgg ggagacaggt gaaatcgtat gggataaggg ccgggacttc	3240
[3240]	gcgacgggtg gaaaagtttt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3241]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360
[3242]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgagagccc tacagttgcc	3420
[3243]	tattctgtcc tagtagtggc aaaagttgag aagggaat ccaagaaact gaagtcagtc	3480
[3244]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3245]	ttccttgagg cgaaaggtta caaggaagta aaaaaggatc tcataattaa actaccaaag	3600
[3246]	tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt	3660
[3247]	caaaagggga acgaactcgc actaccgtct aaatactga atttcctgta tttagcgtcc	3720
[3248]	cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact tttgtttgag	3780
[3249]	cagcacaac attatctga cgaaatcata gagcaaattt cggaattcag taagagagtc	3840
[3250]	atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa	3900
[3251]	cccatactg agcaggcgga aaatattatc catttgttta ctcttacaa cctcggcgct	3960
[3252]	ccagccgcat tcaagtatit tgacacaacg atagatcgca aacgatacac ttctaccaag	4020
[3253]	gagtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3254]	gatttgteac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3255]	tacaaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3256]	aagtga	4206
[3257]	<210>	274
[3258]	<211>	4206
[3259]	<212>	DNA
[3260]	<213>	人工序列(Artificial Sequence)
[3261]	<220>	
[3262]	<221>	gene
[3263]	<222>	(1) .. (4206)
[3264]	<223>	人工序列的说明: 合成 多核苷酸
[3265]	<400>	274
[3266]	atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc	60
[3267]	ataaccgatg aatacaaaag accttcaaag aaatttaagg tgttggggaa cacagaccgt	120
[3268]	cattcgatta aaaagaatct tatcggtgcc ctctatttcg atagtggcga aacggcagag	180
[3269]	gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgcaagaa ccgaatatgt	240
[3270]	tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttcaccgt	300
[3271]	ttggaagagt ccttccttgt cgaagaggac aagaaacatg aacggcacc cactctttgga	360
[3272]	aacatagtag atgagtggtc atatcatgaa aagtacccaa cgatttatca cctcagaaaa	420
[3273]	aagctagttg actcaactga taaagcggac ctgagggtta tctacttggc tcttgcccat	480
[3274]	atgataaagt tccgtgggca ctttctcatt gagggtagtc taaatccgga caactcggat	540
[3275]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaacct	600

[3276]	ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg cccgcctctc taaatcccga	660
[3277]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3278]	cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa	780
[3279]	gatgccaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840
[3280]	caaattggag atcagtatgc ggacttattt ttggctgcca aaaaccttag cgatgcaatc	900
[3281]	ctcctatctg acatactgag agttaatact gagattacca aggcgcctgt atccgcttca	960
[3282]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc cctagtccgt	1020
[3283]	cagcaactgc ctgagaaata taaggaaata ttctttgatc agtcgaaaaa cgggtacgca	1080
[3284]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3285]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcgaaaga tctactgcga	1200
[3286]	aagcagcgga ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3287]	gctatactta gaaggcagga ggatttttat cegtctctca aagacaatcg tgaaaagatt	1320
[3288]	gagaaaatcc taacctttcg catacctgcc tatgtgggac ccctggcccc agggaaactct	1380
[3289]	cggttcgcgt ggatgacaag aaagtccgaa gaaacgatta ctccctggaa ttttgaggaa	1440
[3290]	gttgcgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccgc ctttgacaag	1500
[3291]	aatttaccga acgaaaaagt attgcctaag cacagtttac ttacagta ttacacagt	1560
[3292]	tacaatgaac tcacgaaagt taagtatgtc actgagggca tgcgtaaacc cgcctttcta	1620
[3293]	agcggagaac agaagaaagc aatagtagat ctgttattca agaccaaccg caaagtgaca	1680
[3294]	gttaagcaat tgaaagagga ctactttaag aaaattgaat gcttcgattc tgctgagatc	1740
[3295]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3296]	attaaagata aggacttctt ggataacgaa gagaatgaag atatcttaga agatatagtg	1860
[3297]	ttgactctta ccctctttga agatcgggaa atgattgagg aaagactaaa aacatacgct	1920
[3298]	cacctgttcg acgataaggt tatgaaacag ttaaagaggc gtcgtatac gggctgggga	1980
[3299]	gccttgcgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3300]	gattttctaa agagcgacgg cttcgccaat aggaacttta tggccctgat ccatgatgac	2100
[3301]	tctttaacct tcaaagagga tatacaaaag gcacaggtt cggacaagg ggactcattg	2160
[3302]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaaggcat actccagaca	2220
[3303]	gtcaaagtag tggatgagct agttaaggtc atgggacgtc acaaaccgga aaacattgta	2280
[3304]	atcgagatgg cacgcgaaaa tcaaacgact cagaaggggc aaaaaaacag tcgagagcgg	2340
[3305]	atgaagagaa tagaagaggg tattaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3306]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3307]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgtcgatcac	2520
[3308]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgt tacacgctcg	2580
[3309]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3310]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cgataactta	2700
[3311]	actaaagctg agagggttg cttgtctgaa cttgacaagg ccgatttat taaacgtcag	2760
[3312]	ctcgtggaaa cccgcgccat cacaaagcat gttgcgcaga tactagattc ccgaatgaat	2820
[3313]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3314]	aaattgggtg cgacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3315]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3316]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3317]	atgatcgcga aaagcgaaca ggagataggc aaggctacag ccaatactt cttttattct	3120

[3318]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3319]	cctttaattg aaaccaatgg ggagacaggt gaaatcgtat gggataaggg ccgggacttc	3240
[3320]	gcgacggtga gaaaagtfff gtccatgcc caagtcaaca tagtaaagaa aactgaggtg	3300
[3321]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360
[3322]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgatagccc tacagttgcc	3420
[3323]	tattctgtcc tagtagtggc aaaagttgag aagggaataa ccaagaaact gaagtcagtc	3480
[3324]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3325]	ttccttgagg cgaaaggta caaggaagta aaaaaggatc tcataattaa actaccaag	3600
[3326]	tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt	3660
[3327]	caaaagggga acgaactcgc actaccgtct aaatactga atttcctgta tttagcgtcc	3720
[3328]	cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact tttgtttgag	3780
[3329]	cagcacaac attatctcga cgaaatcata gagcaattt cggaattcag taagagagtc	3840
[3330]	atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa	3900
[3331]	cccatactg agcaggcgga aaatattatc catttgttta ctcttaccac cctcggcgct	3960
[3332]	ccagccgcat tcaagtattt tgacacaacg atagatcgca aacgatacac ttctaccaag	4020
[3333]	gaggtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3334]	gatttgteac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3335]	tacaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3336]	aagtga	4206
[3337]	<210>	275
[3338]	<211>	4206
[3339]	<212>	DNA
[3340]	<213>	人工序列(Artificial Sequence)
[3341]	<220>	
[3342]	<221>	gene
[3343]	<222>	(1) .. (4206)
[3344]	<223>	人工序列的说明: 合成 多核苷酸
[3345]	<400>	275
[3346]	atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc	60
[3347]	ataaccgatg aatacaaagt accttcaaag aaatttaagg tgttggggaa cacagaccgt	120
[3348]	cattcgatta aaaagaatct tatcggtgcc ctctattcgc atagtggcga aacggcagag	180
[3349]	gcgactcgcc tgaacgaac cgctcggaga aggtatacac gtcgcaagaa ccgaatatgt	240
[3350]	tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttcaccgt	300
[3351]	ttggaagagt ctttccttgt cgaagaggac aagaaacatg aacggcacc cctctttgga	360
[3352]	aacatagtag atgaggtggc atatcatgaa aagtacccaa cgatttatca cctcagaaaa	420
[3353]	aagctagtgt actcaactga taaagcggac ctgaggttaa tctacttggc tcttgcccat	480
[3354]	atgataaagt tccgtgggca ctttctcatt gagggtagtc taaatccgga caactcggat	540
[3355]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaacct	600
[3356]	ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg ccgcctctc taaatcccga	660
[3357]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3358]	cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa	780
[3359]	gatgccaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840

[3360]	caaattggag atcagtatgc ggacttattt ttggctgcca aaaaccttag cgatgcaatc	900
[3361]	ctcttatctg acatactgag agttaatact gagattacca aggcgccgtt atccgcttca	960
[3362]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc cctagtcctg	1020
[3363]	cagcaactgc ctgagaaata taaggaaata ttctttgatc agtcgaaaaa cgggtacgca	1080
[3364]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3365]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcgaaaga tctactgcga	1200
[3366]	aagcagcggg ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3367]	gctatactta gaaggcagga ggatttttat ccgttcctca aagacaatcg tgaaaagatt	1320
[3368]	gagaaaatcc taacctttcg cataccttac tatgtgggac ccctggcccc agggaaactct	1380
[3369]	cggttcgcac ggatgacaag aaagtcgaa gaaacgatta ctccatggaa ttttgaggaa	1440
[3370]	gttgctgata aagggtcgtc agctcaatcg ttcacgaga ggatgacaa ctttgacaag	1500
[3371]	aatttaccga acgaaaaagt attgcctaag cacagtttac tttacagata tttcacagt	1560
[3372]	tacaatgaac tcacgaaagt taagtatgtc actgagggca tgcgtaaacc cgcctttcta	1620
[3373]	agcggagAAC agaagaaagc aatagtagat ctgttatcca agaccaaccg caaagtgaca	1680
[3374]	gttaagcaat tgaaagagga ctactttaag aaaattgaat gcttcgattc tgctgagatc	1740
[3375]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3376]	attaaagata aggacttcct ggataacgaa gagaatgaag atatcttaga agatatagt	1860
[3377]	ttgactctta ccctctttga agatcgggaa atgattgagg aaagactaaa aacatacgt	1920
[3378]	cacctgttcg acgataaggt tatgaacag ttaaagaggc gtcgtatac gggctgggga	1980
[3379]	cgattgtcgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3380]	gattttctaa agagcgacgg cttcgccaat aggaacttta tgcagctgat ccatgatgac	2100
[3381]	tctttaacct tcaaagagga tatacaaaag gcacaggtt cggacaagg ggactcattg	2160
[3382]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaagggcat actccagaca	2220
[3383]	gtcaaagtag tggatgagct agttaaggtc atgggacgtc acaaaccgga aaacattgta	2280
[3384]	atcgagatgg cacgcgaaaa tcaaacgact cagaaggggc aaaaaaacag tcgagagcgg	2340
[3385]	atgaagagaa tagaagaggg tattaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3386]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3387]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgtcgatcac	2520
[3388]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgt tacacgtcg	2580
[3389]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3390]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cgataactta	2700
[3391]	actaaagctg agagggttg cttgtctgaa cttgacaagg ccgatttat taaacgtcag	2760
[3392]	ctcgtggaaa cccgccaaat cacaaagcat gttgcacaga tactagattc ccgaatgaat	2820
[3393]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3394]	aaattggtgt cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3395]	taccacatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3396]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3397]	atgatcgcca aaagcgaca ggagatagc aaggctacag ccaataactt cttttattct	3120
[3398]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3399]	cctttaattg aaaccaatgg ggagacaggt gaaatcgat gggataaggg ccgggacttc	3240
[3400]	gcgacggtga gaaaagtgtt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3401]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360

[3402] gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgtgagccc tacagttgcc 3420
 [3403] tattctgtcc tagtagtggc aaaagttgag aagggaat ccaagaaact gaagtcagtc 3480
 [3404] aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac 3540
 [3405] ttccttgagg cgaaagggtta caaggaagta aaaaaggatc tcataattaa actaccaag 3600
 [3406] tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt 3660
 [3407] caaaagggga acgaactcgc actaccgtct aaatacgtga atttcctgta tttagcgtcc 3720
 [3408] cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact ttttgttgag 3780
 [3409] cagcacaac attatctcga cgaaatcata gagcaattt cggaattcag taagagagtc 3840
 [3410] atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa 3900
 [3411] cccatacgtg agcaggcgga aaatattatc catttgttta ctcttaccac cctcggcgct 3960
 [3412] ccagccgcat tcaagtatct tgacacaacg atagatcgca aacagtacag atctaccaag 4020
 [3413] gaggtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata 4080
 [3414] gatttgtcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac 4140
 [3415] tacaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac 4200
 [3416] aagtga 4206
 [3417] <210> 276
 [3418] <211> 4206
 [3419] <212> DNA
 [3420] <213> 人工序列(Artificial Sequence)
 [3421] <220>
 [3422] <221> gene
 [3423] <222> (1) .. (4206)
 [3424] <223> 人工序列的说明: 合成 多核苷酸
 [3425] <400> 276
 [3426] atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc 60
 [3427] ataaccgatg aatacaaaag accttcaaag aaatttaagg tgttggggaa cacagaccgt 120
 [3428] cattcgatta aaaagaatct tatcgggtcc ctctattcgc atagtggcga aacggcagag 180
 [3429] gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgcaagaa ccgaatatgt 240
 [3430] tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttcaccgt 300
 [3431] ttggaagagt ctttccttgt cgaagaggac aagaaacatg aacggcacc cctctttgga 360
 [3432] aacatagtag atgaggtggc atatcatgaa aagtaccaa cgatttatca cctcagaaaa 420
 [3433] aagctagtgt actcaactga taaagcggac ctgagggtta tctacttggc tcttgcccat 480
 [3434] atgataaagt tccgtgggca ctttctcatt gagggtagc taaatccgga caactcggat 540
 [3435] gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaaccct 600
 [3436] ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg cccgcctctc taaatcccga 660
 [3437] cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac 720
 [3438] cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa 780
 [3439] gatgccaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca 840
 [3440] caaattggag atcagtatgc ggacttatct ttggctgcca aaaaccttag cgatgcaatc 900
 [3441] ctcttatctg acatactgag agttaatact gagattacca aggcgccgtt atccgcttca 960
 [3442] atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc ctagtccgt 1020
 [3443] cagcaactgc ctgagaaata taaggaaata ttctttgatc agtcgaaaaa cgggtacgca 1080

[3444]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3445]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcggaaga tctactgcga	1200
[3446]	aagcagcgga ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3447]	gctatactta gaaggcagga ggatttttat ccgttcctca aagacaatcg tgaaaagatt	1320
[3448]	gagaaaatcc taacctttcg cataccttac tatgtgggac ccctggcccc aggggaactct	1380
[3449]	cggttcgcga ggatgacaag aaagtcgga gaaacgatta ctccctggaa ttttgaggaa	1440
[3450]	gttgtcgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccgc ctttgacaag	1500
[3451]	aatttaccga acgaaaaagt attgcctaag cacagtttac ttacagta ttccacagt	1560
[3452]	tacaatgaac tcacgaaagt taagtatgtc actgagggca tgcgtaaacc cgcctttcta	1620
[3453]	agcggagaac agaagaaagc aatagtagat ctgttattca agaccaaccg caaagtga	1680
[3454]	gttaagcaat tgaaagagga ctactttaag aaaattgaat gcttcgattc tgcgagatc	1740
[3455]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3456]	attaaagata aggacttct ggataacgaa gagaatgaag atatcttaga agatatagt	1860
[3457]	ttgactctta ccctcttga agatcgggaa atgattgagg aaagactaaa aacatacgt	1920
[3458]	cacctgttcg acgataaggt tatgaaacag ttaaagagc gtcgtatac gggctgggga	1980
[3459]	gccttgctgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3460]	gattttctaa agagcgacgg cttcgccaat aggaacttta tggccctgat ccatgatgac	2100
[3461]	tctttaacct tcaaagagga tatacaaaag gcacaggtt cggacaagg ggactcattg	2160
[3462]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaaggcat actccagaca	2220
[3463]	gtcaaagtag tggatgagct agttaagtc atgggacgac acaaaccgga aaacattgta	2280
[3464]	atcgagatgg cacgcgaaaa tcaaacgact cagaaggggc aaaaaaacag tcgagagcgg	2340
[3465]	atgaagagaa tagaagaggg tattaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3466]	gtggaaaata cccaattgca gaacgagaaa ctttacctt attacctaca aaatggaagg	2460
[3467]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgtcgatcac	2520
[3468]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgt tacacgctcg	2580
[3469]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3470]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cgataactta	2700
[3471]	actaaagctg agagggttg cttgtctgaa cttgacaagg ccgatttat taaacgtcag	2760
[3472]	ctcgtggaaa cccgcgccat cacaagcat gttgcgcaga tactagattc ccgaatgaat	2820
[3473]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3474]	aaattgggtg cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3475]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3476]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3477]	atgatcgca aaagcgaaca ggagatagc aaggctacag ccaaatactt cttttattct	3120
[3478]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3479]	cctttaattg aaaccaatgg ggagacaggt gaaatcgtat gggataaggg ccgggacttc	3240
[3480]	gcgacggtag gaaaagtgtt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3481]	cagaccggag ggttttcaaa ggaatcgatt ctccaaaaa ggaatagtga taagctcatc	3360
[3482]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgtgagccc tacagttgcc	3420
[3483]	tattctgtcc tagtagtggc aaaagttgag aagggaatat ccaagaaact gaagtcagtc	3480
[3484]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3485]	ttccttgagg cgaaaggta caaggaagta aaaaaggatc tcataattaa actaccaaag	3600

[3486]	tatagtctgt ttgagttaga aaatggccga aaacggatgt tggctagcgc cggagagctt	3660
[3487]	caaaagggga acgaactcgc actaccgtct aaatacgtga atttcctgta tttagcgtcc	3720
[3488]	cattacgaga agttgaaagg ttcacctgaa gataacgaac agaagcaact ttttgttgag	3780
[3489]	cagcaciaaac attatctcga cgaaatcata gagcaaattt cggaattcag taagagagtc	3840
[3490]	atcctagctg atgccaatct ggacaaagta ttaagcgcac acaacaagca cagggataaa	3900
[3491]	cccatacgtg agcaggcgga aaatattatc catttggtta ctcttaccac cctcggcgt	3960
[3492]	ccagccgcat tcaagtatit tgacacaacg atagatcgca aacagtacag atctaccaag	4020
[3493]	gagtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3494]	gatttgtcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3495]	tacaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3496]	aagtga	4206
[3497]	<210>	277
[3498]	<211>	4206
[3499]	<212>	DNA
[3500]	<213>	人工序列(Artificial Sequence)
[3501]	<220>	
[3502]	<221>	gene
[3503]	<222>	(1) .. (4206)
[3504]	<223>	人工序列的说明: 合成 多核苷酸
[3505]	<400>	277
[3506]	atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc	60
[3507]	ataaccgatg aatacaaagt accttcaaag aaatttaagg tgttggggaa cacagaccgt	120
[3508]	cattcgatta aaaagaatct tatcggtgcc ctctatttcg atagtggcga aacggcagag	180
[3509]	gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgcaagaa ccgaatatgt	240
[3510]	tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttaccgt	300
[3511]	ttggaagagt ctttcctgt cgaagaggac aagaaacatg aacggcacc catctttgga	360
[3512]	aacatagtag atgaggtggc atatcatgaa aagtaccaa cgatttatca cctcagaaaa	420
[3513]	aagctagtgt actcaactga taaagcggac ctgaggttaa tctacttggc tcttgcccat	480
[3514]	atgataaagt tccgtggga ctttctcatt gaggtgac taaatccgga caactcggat	540
[3515]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaaccct	600
[3516]	ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg cccgcctctc taaatcccga	660
[3517]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3518]	cttatagcgc tctcactagg cctgacacca aattttaagt cgaacttcga cttagctgaa	780
[3519]	gatgccaaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840
[3520]	caaattggag atcagtatgc ggacttatit ttggctgcca aaaaccttag cgatgcaatc	900
[3521]	ctctatctg acatactgag agttaatact gagattacca aggcgccgtt atccgcttca	960
[3522]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc ctagtccgt	1020
[3523]	cagcaactgc ctgagaaata taaggaaata ttctttgac agtcgaaaaa cgggtacgca	1080
[3524]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3525]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcgaaaga tctactgca	1200
[3526]	aagcagcgga ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3527]	gctatactta gaaggcagga ggatttttat ccgttctca aagacaatcg tgaaaagatt	1320

[3528]	gagaaaatcc taacctttcg cataccttac tatgtgggac ccctggcccg agggaaactct	1380
[3529]	cggttcgcac ggatgacaag aaagtcgga gaaacgatta ctccatggaa ttttgaggaa	1440
[3530]	gttgctgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccaa ctttgacaag	1500
[3531]	aatttaccga acgaaaaagt attgcctaag cacagtttac tttacagta tttcacagt	1560
[3532]	tacaatgaac tcacgaaagt taagtatgtc actgagggca tgcgtaaacc cgcctttcta	1620
[3533]	agcggagaac agaagaaagc aatagtagat ctgttattca agaccaaccg caaagtgaca	1680
[3534]	gttaagcaat tgaaagagga ctactttaag aaaattgaat gcttcgattc tgcgagatc	1740
[3535]	tccgggtag aagatcgatt taatgcgtca cttggtacgt atcatgacct cctaaagata	1800
[3536]	attaaagata aggacttcct ggataacgaa gagaatgaag atatcttaga agatatagt	1860
[3537]	ttgactctta ccctctttga agatcgggaa atgattgagg aaagactaaa aacatacgt	1920
[3538]	cacctgttcg acgataaggt tatgaaacag ttaaagaggc gtcgtatac gggctgggga	1980
[3539]	cgattgtcgc ggaaacttat caacgggata agagacaagc aaagtggtaa aactattctc	2040
[3540]	gattttctaa agagcgacgg cttcgccaat aggaacttta tgcagctgat ccatgatgac	2100
[3541]	tctttaacct tcaaagagga tatacaaaag gcacaggttt cgggacaagg ggactcatt	2160
[3542]	cacgaacata ttgcgaatct tgctggttcg ccagccatca aaaagggcac actccagaca	2220
[3543]	gtcaaagtag tggatgagct agttaaggtc atgggacgtc acaaaccgga aaacattgta	2280
[3544]	atcgagatgg cacgcgaaaa tcaaacgact cagaaggggc aaaaaacag tcgagagcgg	2340
[3545]	atgaagagaa tagaagaggg tattaaagaa ctgggcagcc agatcttaaa ggagcatcct	2400
[3546]	gtggaaaata cccaattgca gaacgagaaa ctttacctct attacctaca aaatggaagg	2460
[3547]	gacatgtatg ttgatcagga actggacata aaccgtttat ctgattacga cgtcgatcac	2520
[3548]	attgtacccc aatccttttt gaaggacgat tcaatcgaca ataaagtgtc tacacgtcg	2580
[3549]	gataagaacc gagggaaaag tgacaatgtt ccaagcgagg aagtcgtaaa gaaaatgaag	2640
[3550]	aactattggc ggcagctcct aaatgcgaaa ctgataacgc aaagaaagt cgataactta	2700
[3551]	actaaagctg agagggttg cttgtctgaa cttgacaagg ccggtttat taaacgtcag	2760
[3552]	ctcgtggaaa cccgccaaat cacaaagcat gttgcacaga tactagattc ccgaatgaat	2820
[3553]	acgaaatacg acgagaacga taagctgatt cgggaagtca aagtaatcac tttaaagtca	2880
[3554]	aaattggtgt cggacttcag aaaggatttt caattctata aagttaggga gataaataac	2940
[3555]	taccaccatg cgcacgacgc ttatcttaat gccgtcgtag ggaccgcact cattaagaaa	3000
[3556]	taccgaagc tagaaagtga gtttgtgtat ggtgattaca aagtttatga cgtccgtaag	3060
[3557]	atgatcgcga aaagcgaaca ggagatagga aaggctacag ccaataactt cttttattct	3120
[3558]	aacattatga atttctttaa gacggaaatc actctggcaa acggagagat acgcaaacga	3180
[3559]	cctttaattg aaaccaatgg ggagacaggt gaaatcgat gggataaggg ccgggacttc	3240
[3560]	gcgacggtag gaaaagtgtt gtccatgccc caagtcaaca tagtaaagaa aactgaggtg	3300
[3561]	cagaccggag ggttttcaaa ggaatcgatt cttccaaaaa ggaatagtga taagctcatc	3360
[3562]	gctcgtaaaa aggactggga cccgaaaaag tacggtggct tcgtgagccc tacagttgcc	3420
[3563]	tattctgtcc tagtagtggc aaaagttgag aagggaat ccaagaaact gaagtcagtc	3480
[3564]	aaagaattat tggggataac gattatggag cgctcgtctt ttgaaaagaa ccccatcgac	3540
[3565]	ttccttgagg cgaaaggtta caaggaagta aaaaaggatc tcataattaa actaccaaag	3600
[3566]	tatagtctgt ttgagtaga aaatggccga aaacggatgt tggctagcgc cagagagctt	3660
[3567]	caaaagggga acgaactcgc actaccgtct aaatacgtga atttcctgta tttagcgtcc	3720
[3568]	cattacgaga agttgaaagg ttcactgaa gataacgaac agaagcaact tttgttgag	3780
[3569]	cagcacaac attatctcga cgaaatcata gagcaattt cggaattcag taagagagtc	3840

[3570]	atcctagctg atgccaatct ggacaaagta ttaagcgcat acaacaagca cagggataaa	3900
[3571]	cccatacgtg agcaggcgga aaatattatc catttgttta ctcttaccaa cctcggcgt	3960
[3572]	ccagccgcat tcaagtatit tgacacaacg atagatcgca aacagtacag atctaccaag	4020
[3573]	gaggtgctag acgcgacact gattcaccaa tccatcacgg gattatatga aactcggata	4080
[3574]	gatttgtcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac	4140
[3575]	tacaaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac	4200
[3576]	aagtga	4206
[3577]	<210>	278
[3578]	<211>	4206
[3579]	<212>	DNA
[3580]	<213>	人工序列(Artificial Sequence)
[3581]	<220>	
[3582]	<221>	gene
[3583]	<222>	(1) .. (4206)
[3584]	<223>	人工序列的说明: 合成 多核苷酸
[3585]	<400>	278
[3586]	atggataaaa agtattctat tggtttagac atcggcacta attccgttgg atgggctgtc	60
[3587]	ataaccgatg aatacaaaagt accttcaaag aaatttaagg tgttggggaa cacagaccgt	120
[3588]	cattcgatta aaaagaatct tatcggtgcc ctctattcg atagtggcga aacggcagag	180
[3589]	gcgactcgcc tgaaacgaac cgctcggaga aggtatacac gtcgaagaa ccgaatatgt	240
[3590]	tacttacaag aaatttttag caatgagatg gccaaagttg acgattcttt ctttcaccgt	300
[3591]	ttggaagagt ctttccttgt cgaagaggac aagaaacatg aacggcacc cactcttggga	360
[3592]	aacatagtag atgaggtggc atatcatgaa aagtaccaa cgatttatca cctcagaaaa	420
[3593]	aagctagttag actcaactga taaagcggac ctgagggtta tctacttggc tcttgcceat	480
[3594]	atgataaagt tccgtgggca ctttctcatt gaggggtgatc taaatccgga caactcgat	540
[3595]	gtcgacaaac tgttcatcca gttagtacaa acctataatc agttgtttga agagaaccct	600
[3596]	ataaatgcaa gtggcgtgga tgcgaaggct attcttagcg cccgcctctc taaatccga	660
[3597]	cggctagaaa acctgatcgc acaattaccc ggagagaaga aaaatgggtt gttcggtaac	720
[3598]	cttatagcgc tctcactagg cctgacacca aatttttaagt cgaacttcga cttagctgaa	780
[3599]	gatgccaat tgcagcttag taaggacacg tacgatgacg atctcgacaa tctactggca	840
[3600]	caaattggag atcagtatgc ggacttatit ttggctgcca aaaaccttag cgatgcaatc	900
[3601]	ctcctatctg acatactgag agttaatact gagattacca aggcgccgtt atccgcttca	960
[3602]	atgatcaaaa ggtacgatga acatcaccaa gacttgacac ttctcaaggc ctagtccgt	1020
[3603]	cagcaactgc ctgagaaata taaggaaata ttctttgatc agtcgaaaaa cgggtacgca	1080
[3604]	ggttatattg acggcggagc gagtcaagag gaattctaca agtttatcaa acccatatta	1140
[3605]	gagaagatgg atgggacgga agagttgctt gtaaaactca atcgcggaaga tctactgca	1200
[3606]	aagcagcgga ctttcgacaa cggtagcatt ccacatcaaa tccacttagg cgaattgcat	1260
[3607]	gctatactta gaaggcagga ggatttttat ccgttctca aagacaatcg tgaaaagatt	1320
[3608]	gagaaaatcc taaccttctg cataccttac tatgtgggac ccctggcccc agggaaactct	1380
[3609]	cggttcgcat ggatgacaag aaagtccgaa gaaacgatta ctccctggaa ttttgaggaa	1440
[3610]	gttgtcgata aaggtgcgtc agctcaatcg ttcacgaga ggatgaccgc ctttgacaag	1500
[3611]	aatttaccga acgaaaaagt attgcctaag cacagtttac tttacagta tttcacagt	1560

[3612]	tacaatgaac	tcacgaaagt	taagtatgtc	actgagggca	tgcgtaaacc	cgcctttcta	1620
[3613]	agcggagaac	agaagaaagc	aatagtagat	ctgttattca	agaccaaccg	caaagtgaca	1680
[3614]	gttaagcaat	tgaaagagga	ctactttaag	aaaattgaat	gcttcgattc	tgtecgagatc	1740
[3615]	tccggggtag	aagatcgatt	taatgcgtca	cttggtacgt	atcatgacct	cctaaagata	1800
[3616]	attaaagata	aggacttcct	ggataacgaa	gagaatgaag	atatcttaga	agatatagtg	1860
[3617]	ttgactctta	ccctctttga	agatcgggaa	atgattgagg	aaagactaaa	aacatacgct	1920
[3618]	cacctgttcg	acgataaggt	tatgaaacag	ttaaagaggc	gtcgctatac	gggctgggga	1980
[3619]	gccttgctgc	ggaaacttat	caacgggata	agagacaagc	aaagtggtaa	aactattctc	2040
[3620]	gattttctaa	agagcgacgg	cttcgccaat	aggaacttta	tggccctgat	ccatgatgac	2100
[3621]	tctttaacct	tcaaagagga	tatacaaaag	gcacaggttt	cggacaagg	ggactcattg	2160
[3622]	cacgaacata	ttgcgaatct	tgctggttcg	ccagccatca	aaaaggggcat	actccagaca	2220
[3623]	gtcaaagtag	tggatgagct	agttaaggtc	atgggacgtc	acaaaccgga	aaacattgta	2280
[3624]	atcgagatgg	cacgcgaaaa	tcaaacgact	cagaaggggc	aaaaaaacag	tcgagagcgg	2340
[3625]	atgaagagaa	tagaagaggg	tattaaagaa	ctgggcagcc	agatcttaaa	ggagcatcct	2400
[3626]	gtggaaaata	cccaattgca	gaacgagaaa	ctttacctct	attacctaca	aaatggaagg	2460
[3627]	gacatgtatg	ttgatcagga	actggacata	aaccgtttat	ctgattacga	cgtegatcac	2520
[3628]	attgtacccc	aatccttttt	gaaggacgat	tcaatcgaca	ataaagtgtc	tacacgtctg	2580
[3629]	gataagaacc	gagggaaaag	tgacaatgtt	ccaagcgagg	aagtcgtaaa	gaaaatgaag	2640
[3630]	aactattggc	ggcagctcct	aaatgcgaaa	ctgataacgc	aaagaaagtt	cgataactta	2700
[3631]	actaaaagctg	agaggggtgg	cttgtctgaa	cttgacaagg	cggattttat	taaacgtcag	2760
[3632]	ctcgtggaaa	cccgcgccat	cacaaagcat	gttgcgcaga	tactagattc	ccgaatgaat	2820
[3633]	acgaaatacg	acgagaacga	taagctgatt	cgggaagtca	aagtaatcac	tttaaagtca	2880
[3634]	aaattggtgt	cggacttcag	aaaggatttt	caattctata	aagttaggga	gataaataac	2940
[3635]	taccaccatg	cgcacgacgc	ttatcttaat	gccgtcgtag	ggaccgcact	cattaagaaa	3000
[3636]	taccgaagc	tagaaaagtg	gtttgtgtat	ggtgattaca	aagtttatga	cgtecgtaag	3060
[3637]	atgatcgcg	aaagcgaaca	ggagataggc	aaggctacag	caaataactt	ctttttattct	3120
[3638]	aacattatga	atttctttaa	gacggaaatc	actctggcaa	acggagagat	acgcaaacga	3180
[3639]	cctttaattg	aaaccaatgg	ggagacaggt	gaaatcgtat	gggataaggg	cggggacttc	3240
[3640]	gcgacggtga	gaaaagtttt	gtccatgccc	caagtcaaca	tagtaaagaa	aactgaggtg	3300
[3641]	cagaccggag	ggttttcaaa	ggaatcgatt	cttccaaaaa	ggaatagtga	taagctcatc	3360
[3642]	gctcgtaaaa	aggactggga	cccgaaaaag	tacggtggct	tcgtgagccc	tacagttgcc	3420
[3643]	tattctgtcc	tagtagtggc	aaaagttgag	aagggaat	ccaagaaact	gaagtcagtc	3480
[3644]	aaagaattat	tggggataac	gattatggag	cgctcgtctt	ttgaaaagaa	ccccatcgac	3540
[3645]	ttccttgagg	cgaaggtta	caaggaagta	aaaaaggatc	tcataattaa	actaccaaag	3600
[3646]	tatagtctgt	ttgagttaga	aaatggccga	aaacggatgt	tggctagcgc	cagagagctt	3660
[3647]	caaaagggga	acgaactcgc	actaccgtct	aaatacgtga	atttcctgta	tttagcgtcc	3720
[3648]	cattacgaga	agttgaaagg	ttcacctgaa	gataacgaac	agaagcaact	ttttgttgag	3780
[3649]	cagcacaac	attatctcga	cgaaatcata	gagcaaattt	cggaattcag	taagagagtc	3840
[3650]	atcctagctg	atgccaatct	ggacaaagta	ttaagcgc	acaacaagca	cagggataaa	3900
[3651]	cccatacgtg	agcaggcgga	aaatattatc	catttgttta	ctcttaccaa	cctcggcgct	3960
[3652]	ccagccgcat	tcaagtattt	tgacacaacg	atagatcgca	aacagtacag	atctaccaag	4020
[3653]	gaggtgctag	acgcgacact	gattcaccaa	tccatcacgg	gattatatga	aactcggata	4080

[3654] gatttgtcac agcttggggg tgacggatcc cccaagaaga agaggaaagt ctcgagcgac 4140
[3655] taaaagacc atgacggtga ttataaagat catgacatcg attacaagga tgacgatgac 4200
[3656] aagtga 4206
[3657] <210> 279
[3658] <211> 422
[3659] <212> DNA
[3660] <213> 人工序列(Artificial Sequence)
[3661] <220>
[3662] <221> gene
[3663] <222> (1) .. (422)
[3664] <223> 人工序列的说明: 合成 多核苷酸
[3665] <400> 279
[3666] tgtacaaaaa agcagccttt aaaggaacca attcagtcga ctggatccgg taccaaggtc 60
[3667] gggcaggaag agggcctatt tcccatgatt cttcatatt tgcataatcg atacaaggct 120
[3668] gttagagaga taattagaat taatttgact gtaaacacaa agatattagt acaaaatagc 180
[3669] tgacgtagaa agtaataatt tcttgggtag ttgacagttt taaaattatg ttttaaaatg 240
[3670] gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt atatatcttg 300
[3671] tggaaggac gaaacaccgg agacgattaa tgcgtctccg ttttagagct agaaatagca 360
[3672] agttaaata aggctagtc gttatcaact tgaaaaagtg gcaccgagtc ggtgcttttt 420
[3673] tt 422

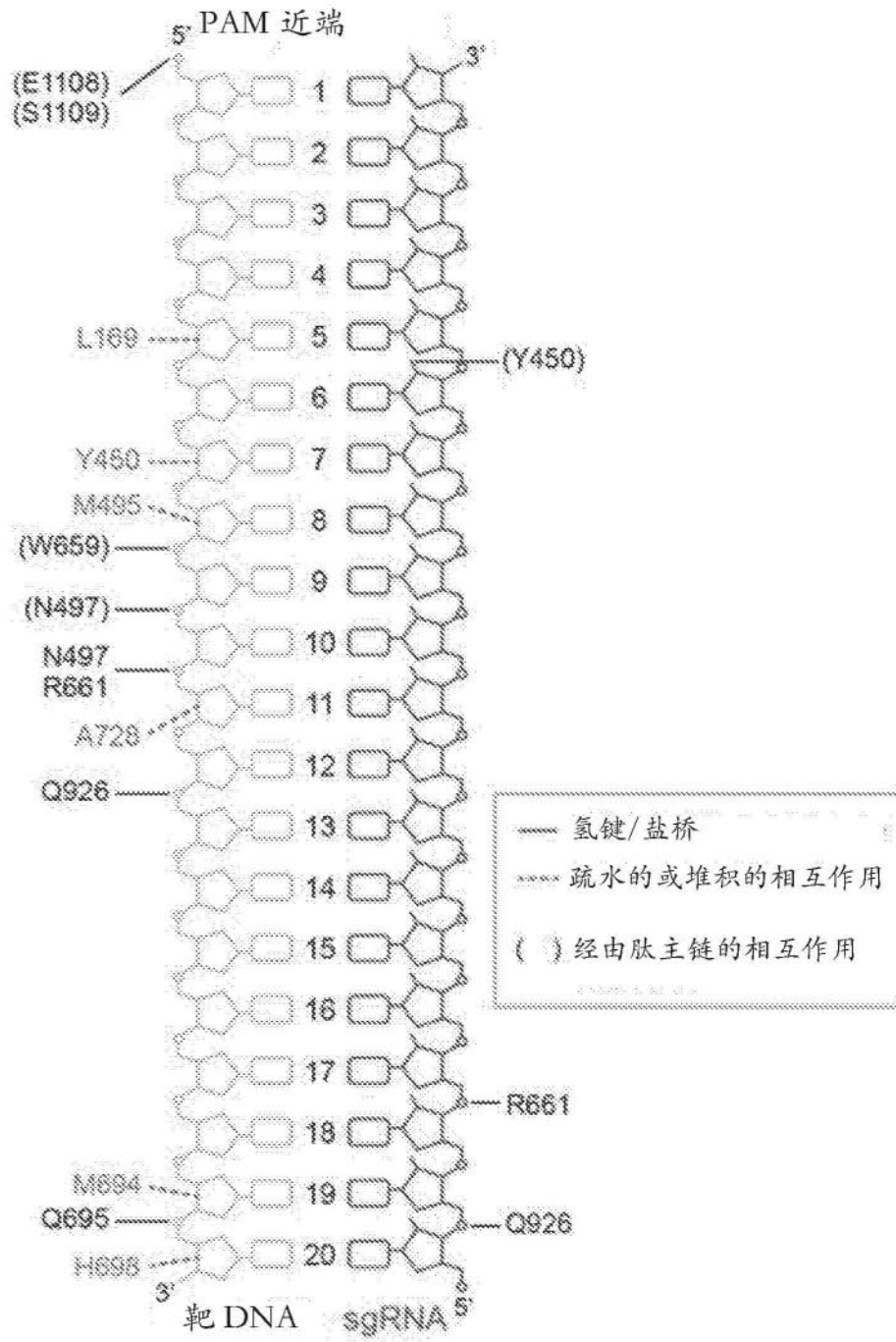


图1A

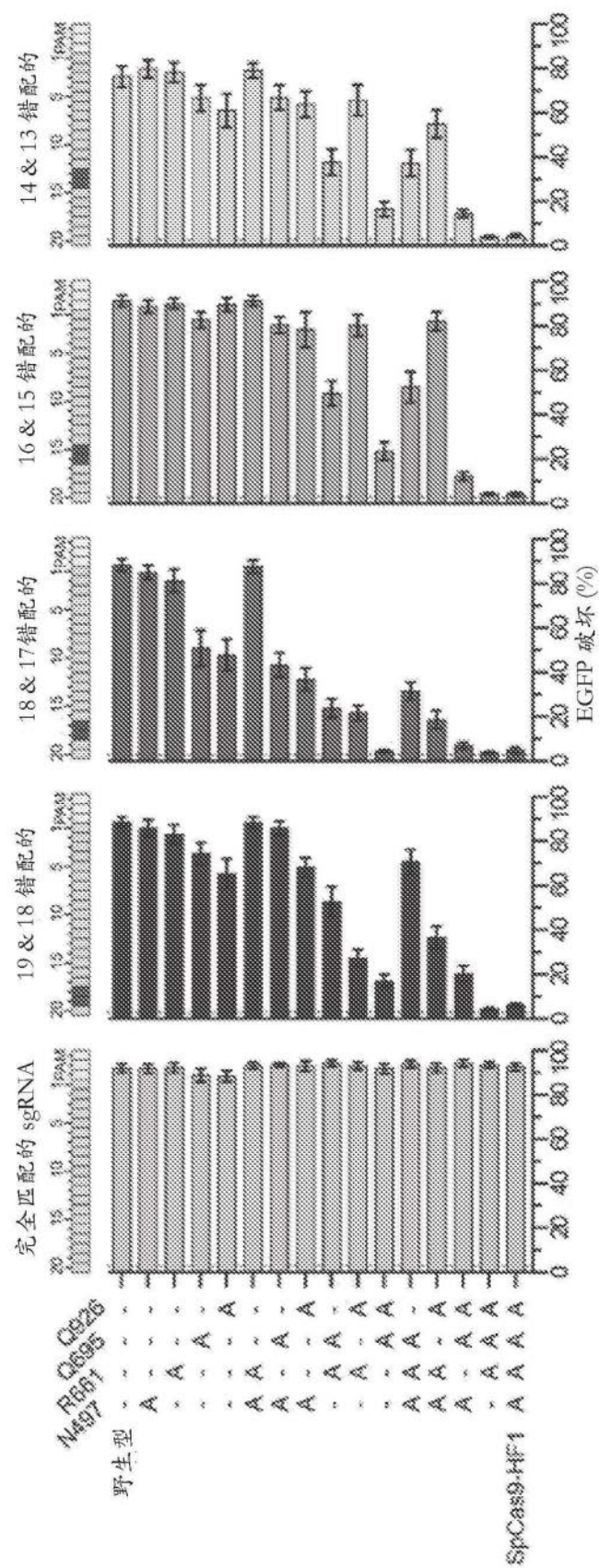


图1B

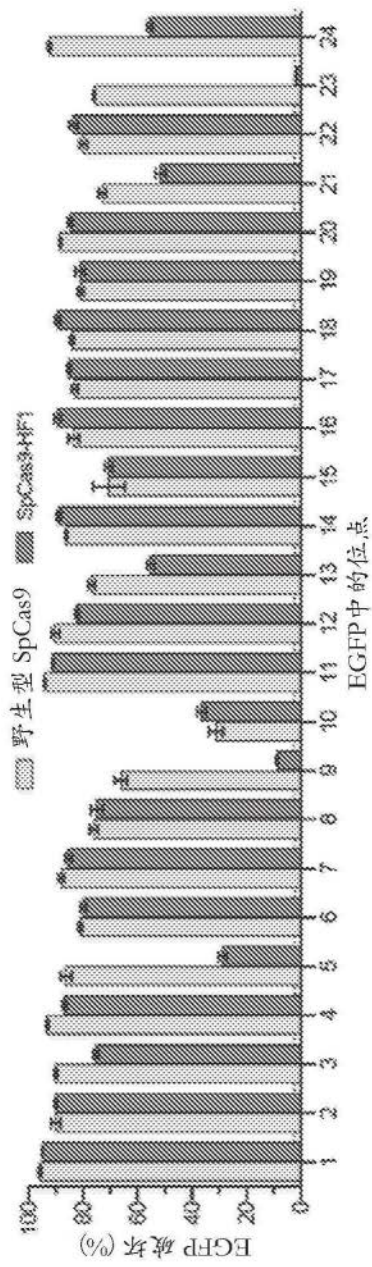


图1C

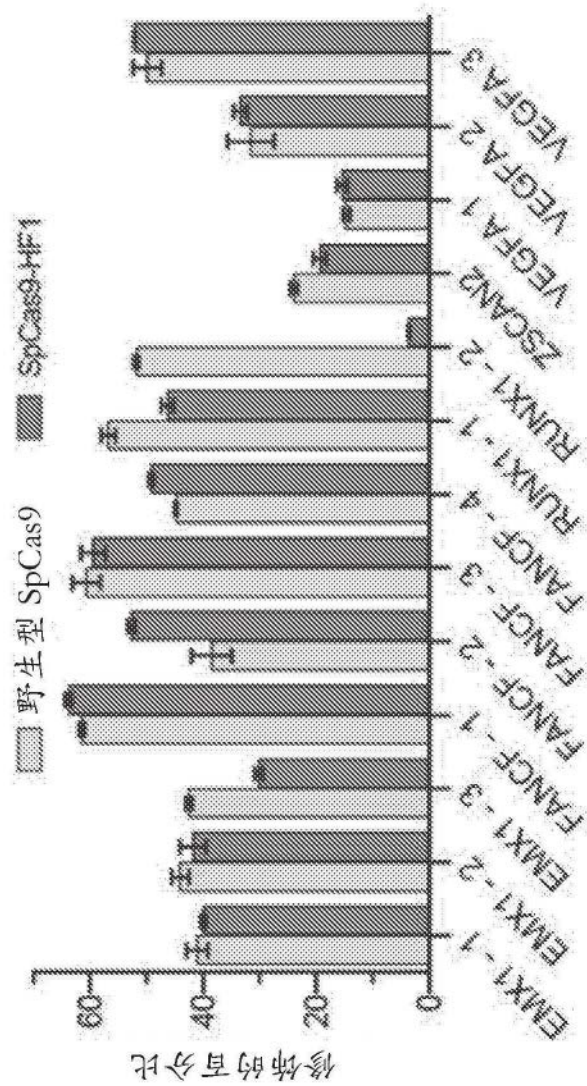


图1D

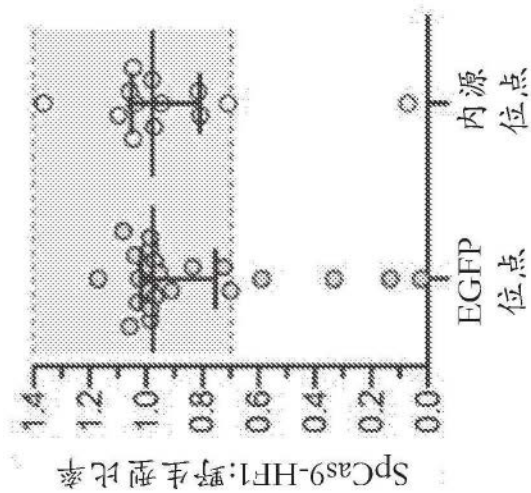


图1E

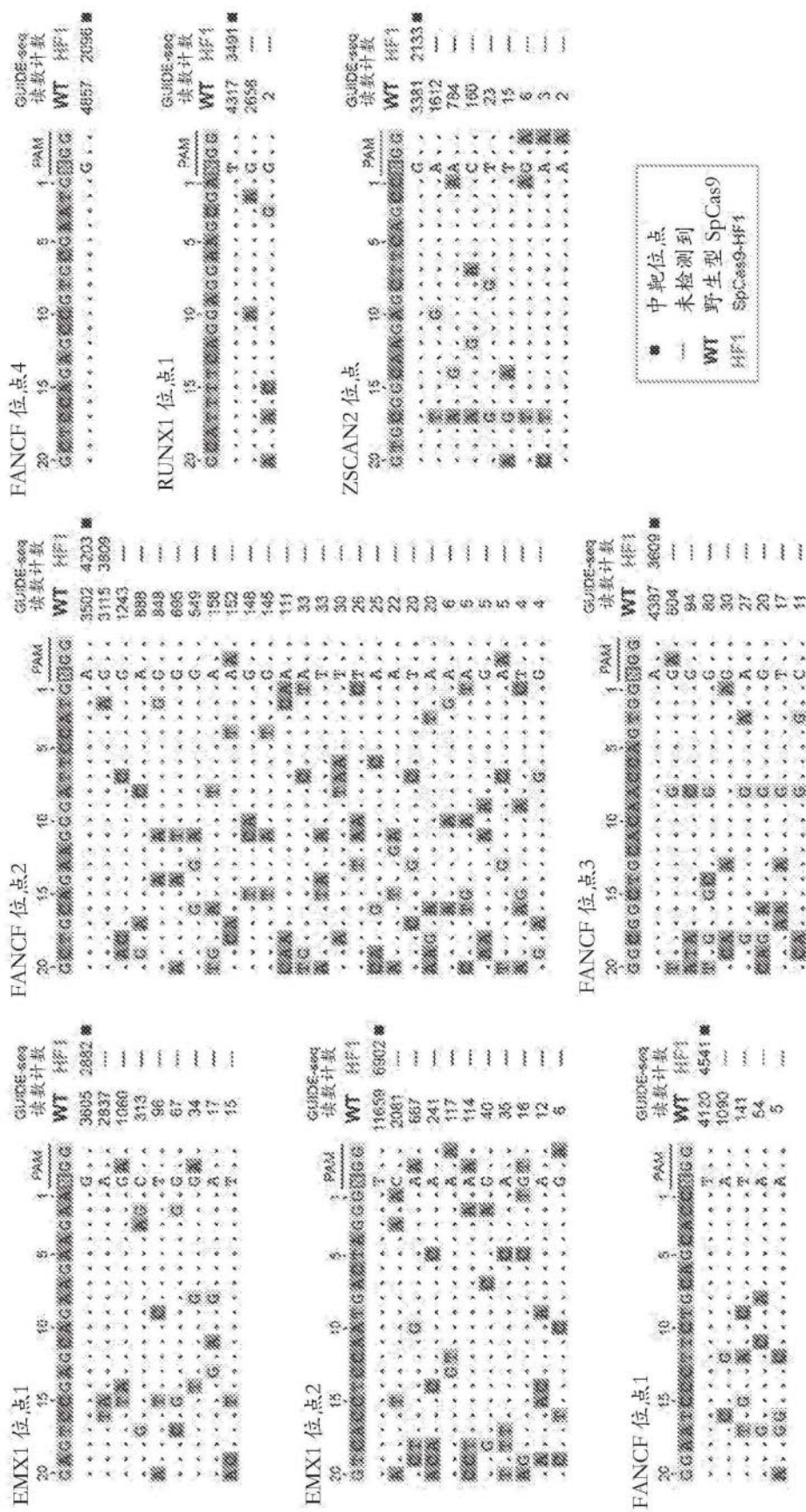


图2A

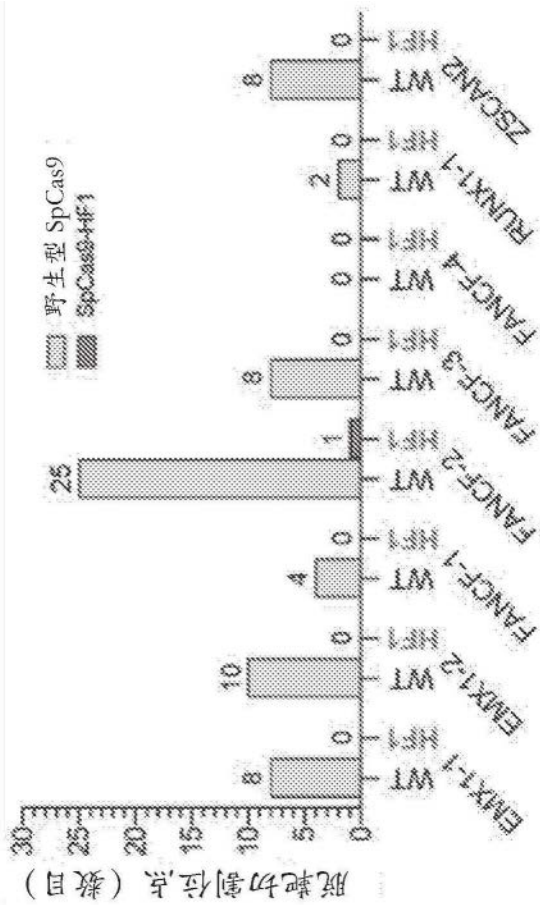


图2B

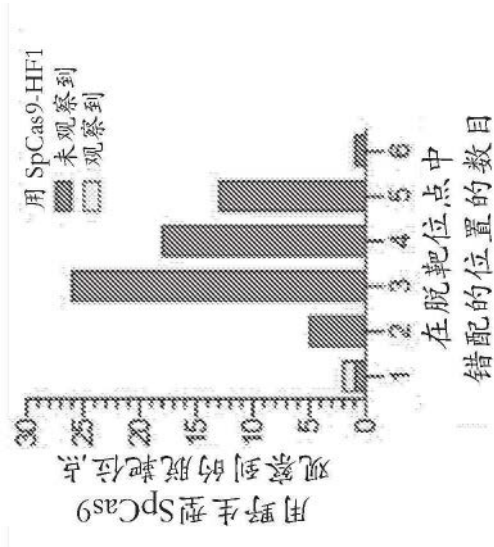


图2C

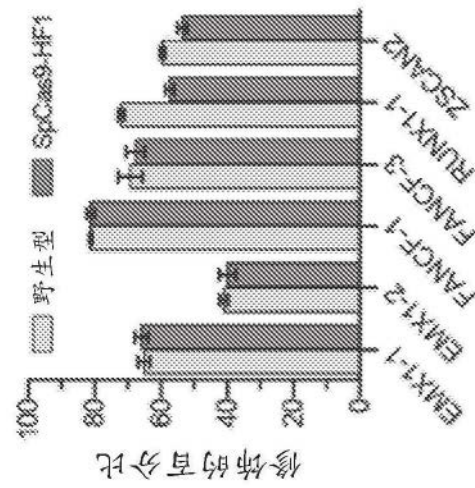


图3A

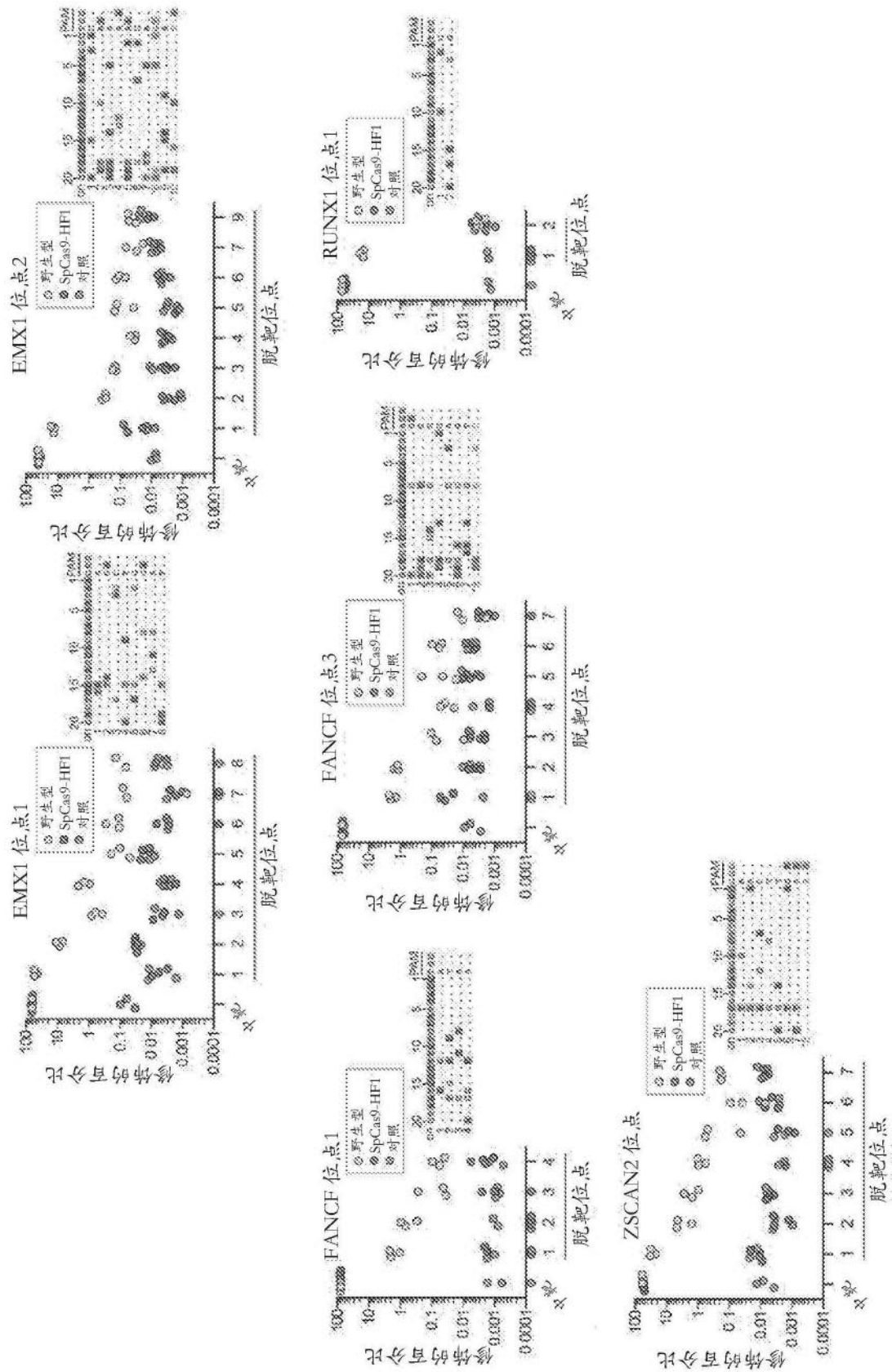


图3B

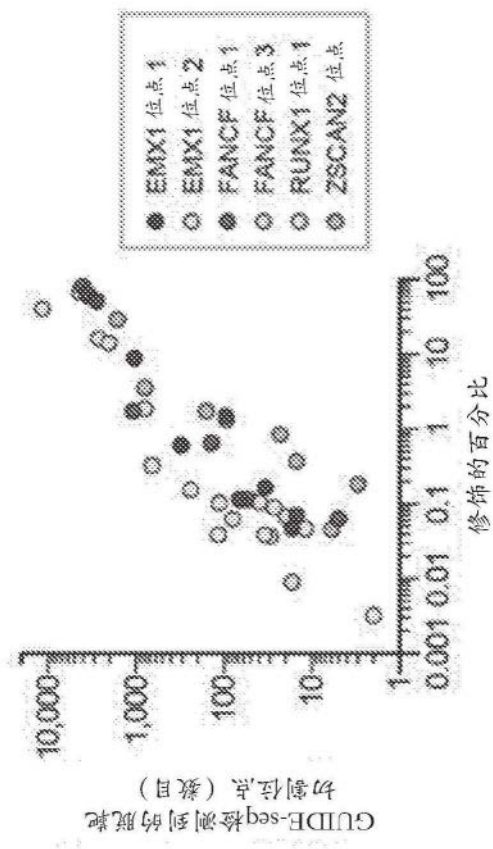


图3C



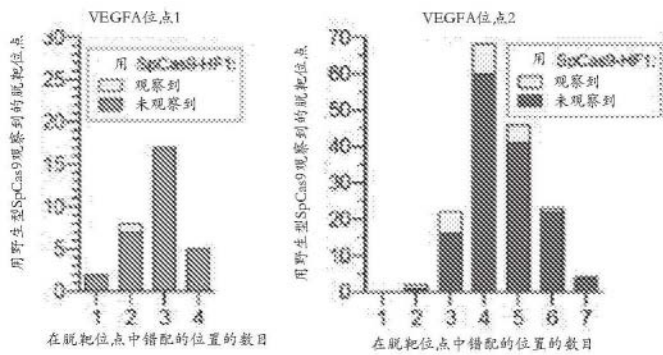


图4C

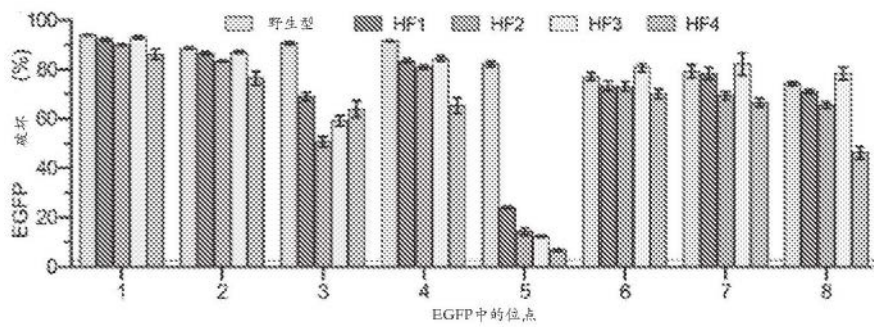


图5A

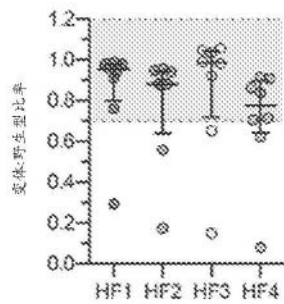


图5B

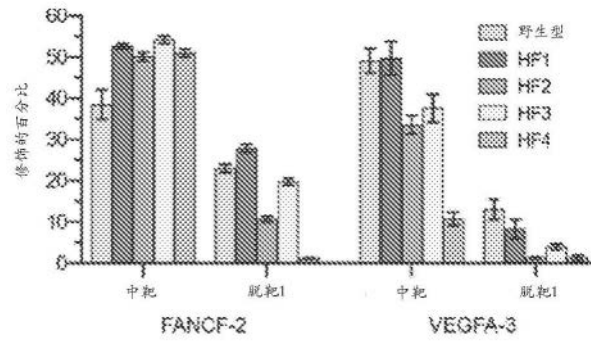


图5C

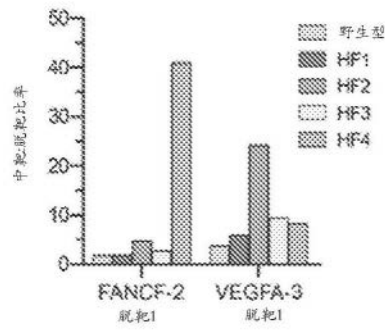


图5D

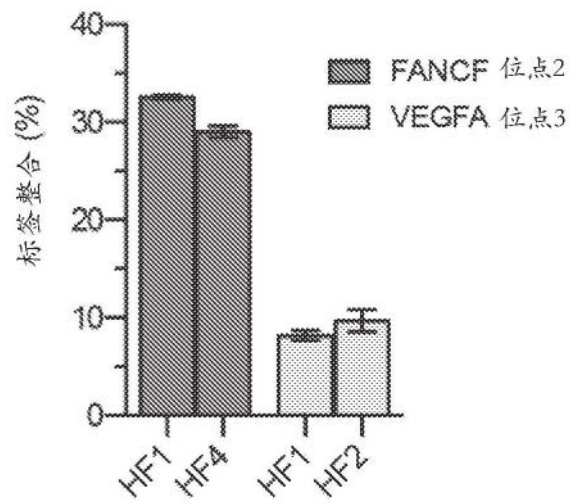


图5E

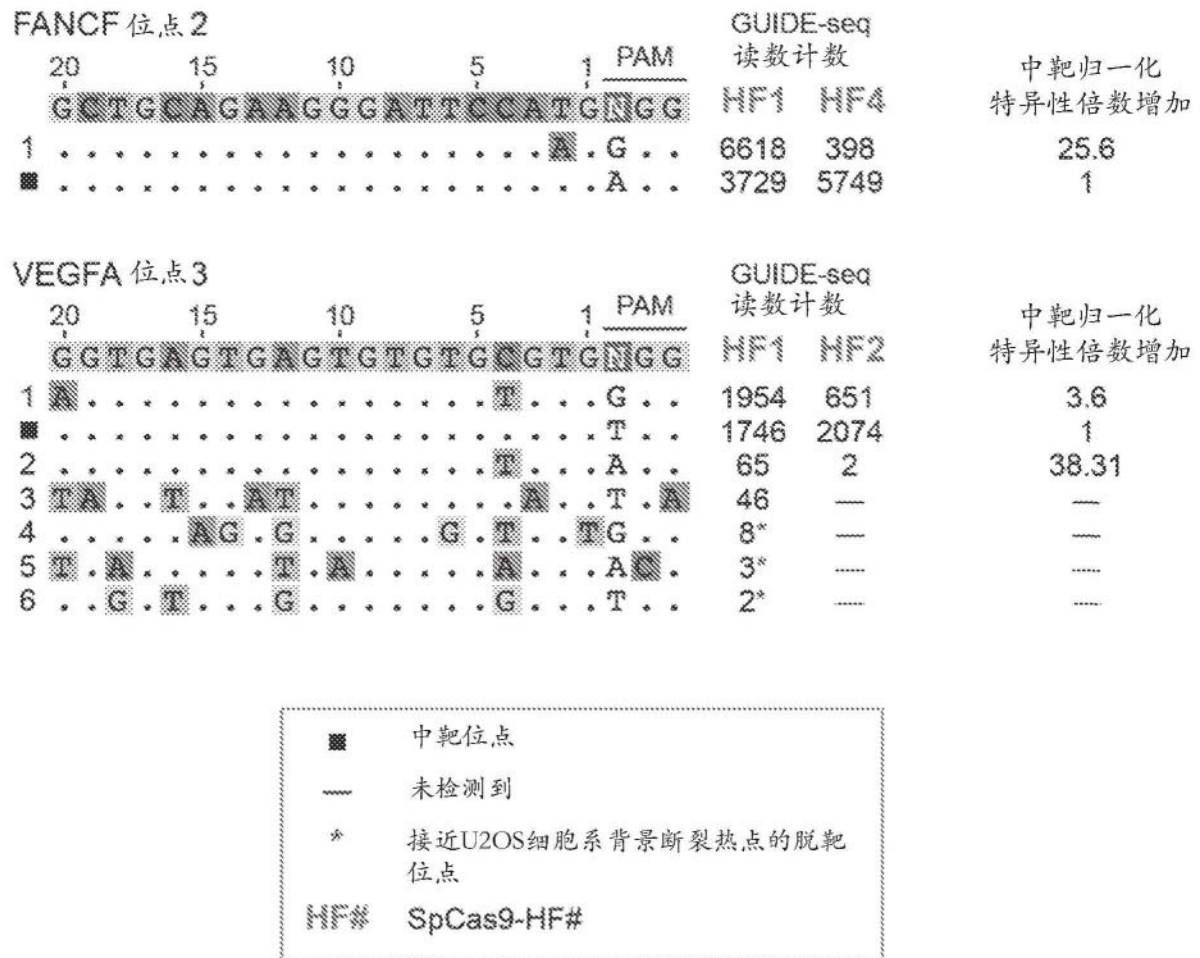


图5F

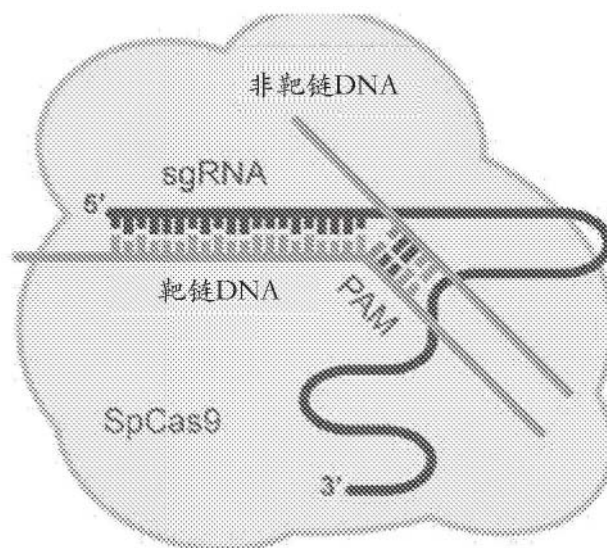


图6A

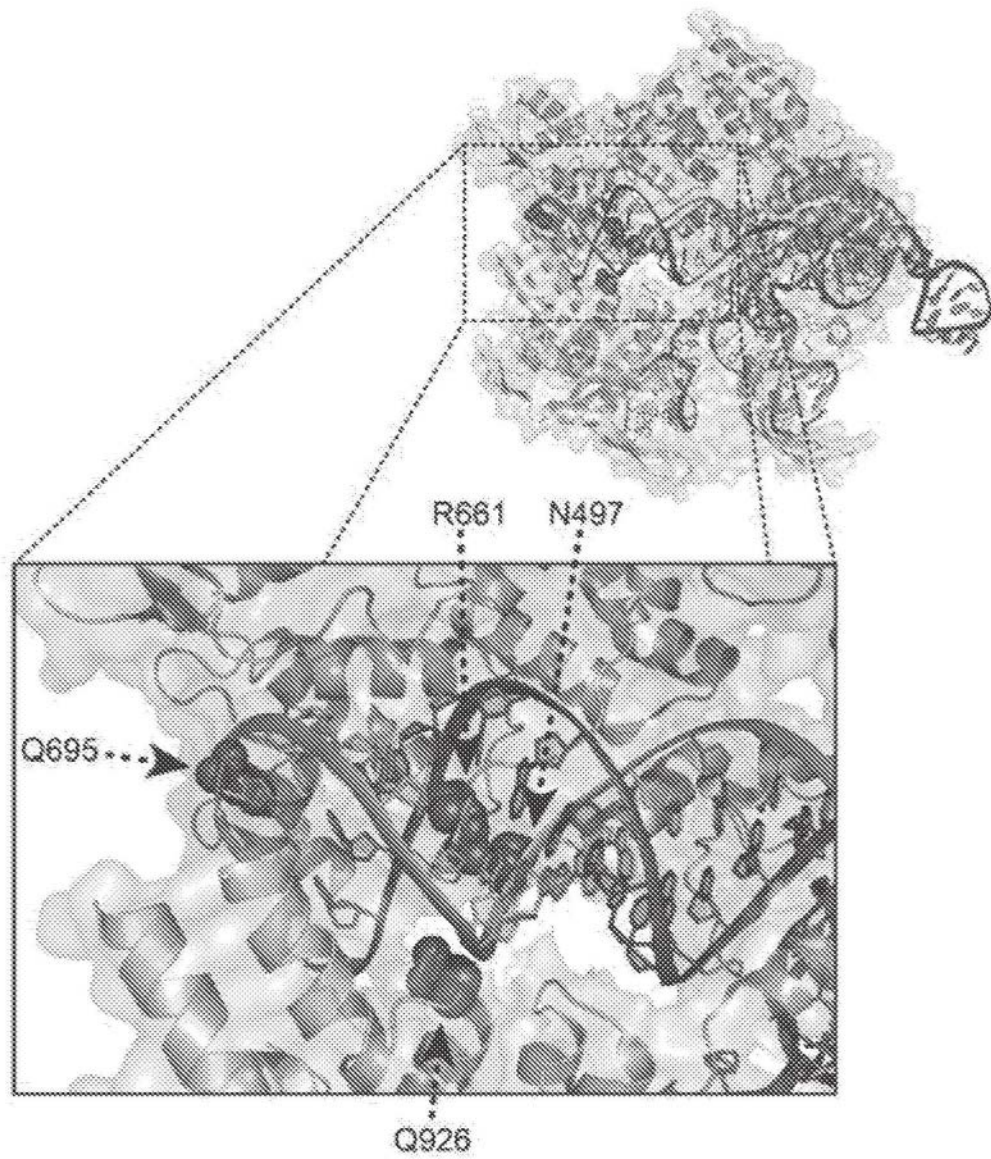


图6B

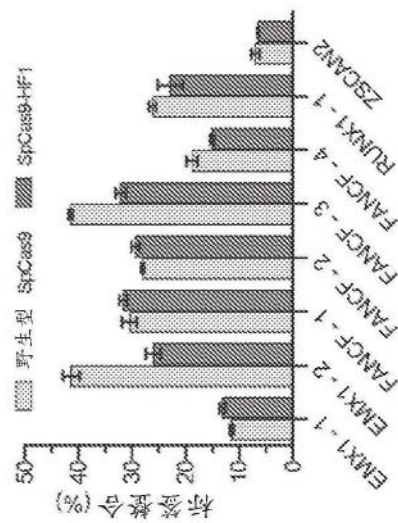


图7A

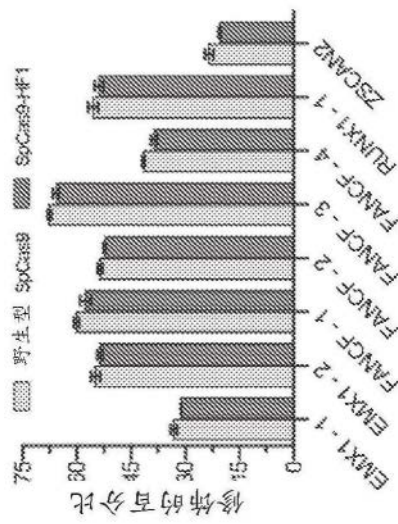


图7B

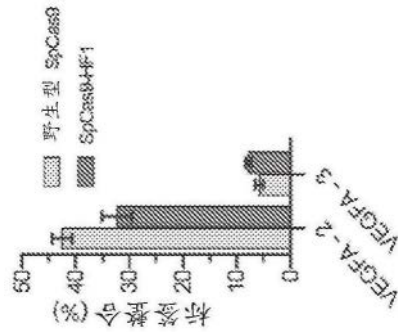


图7C

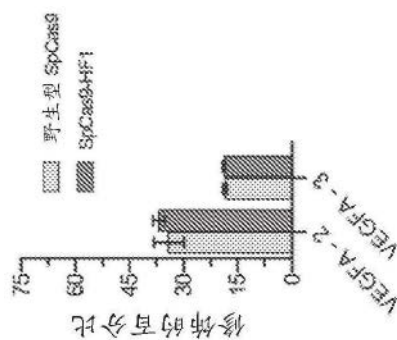


图7D

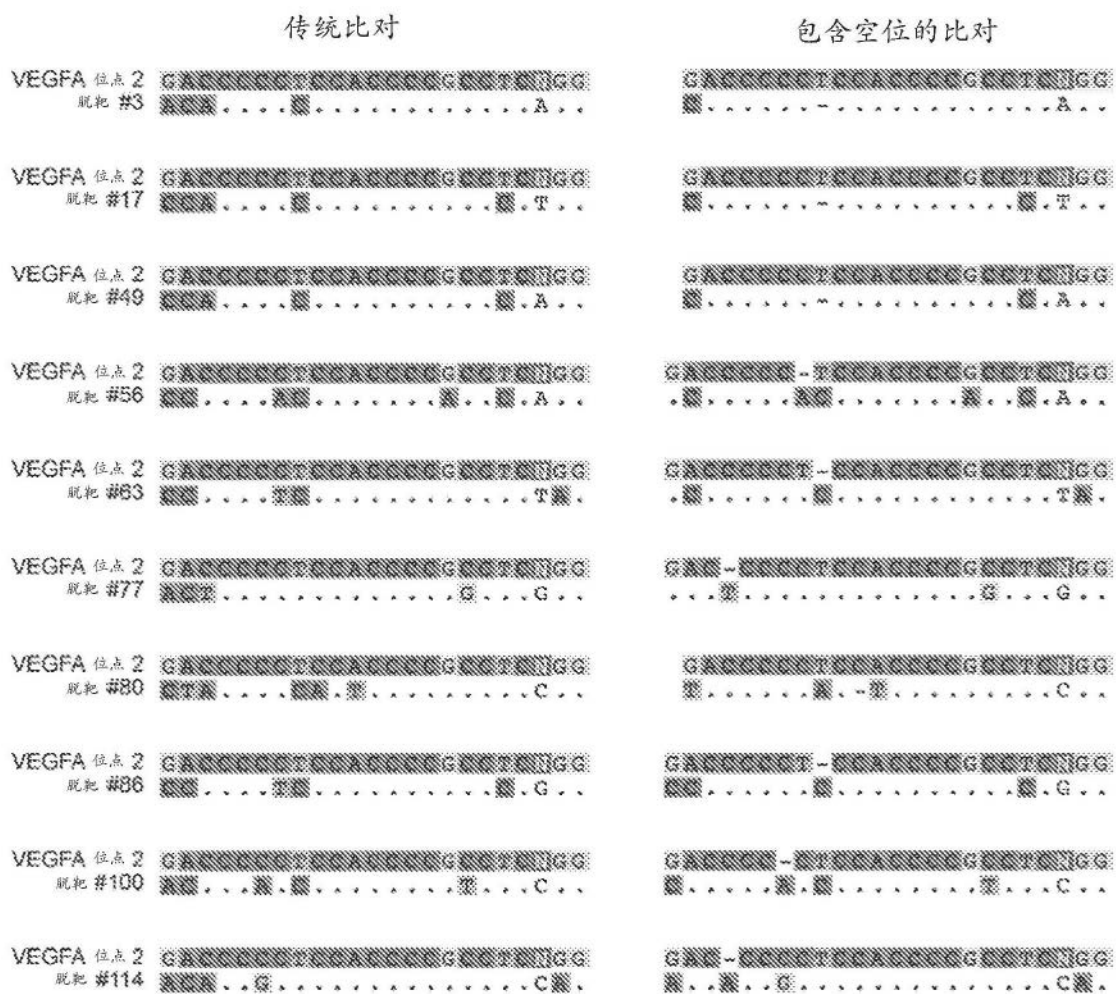


图8

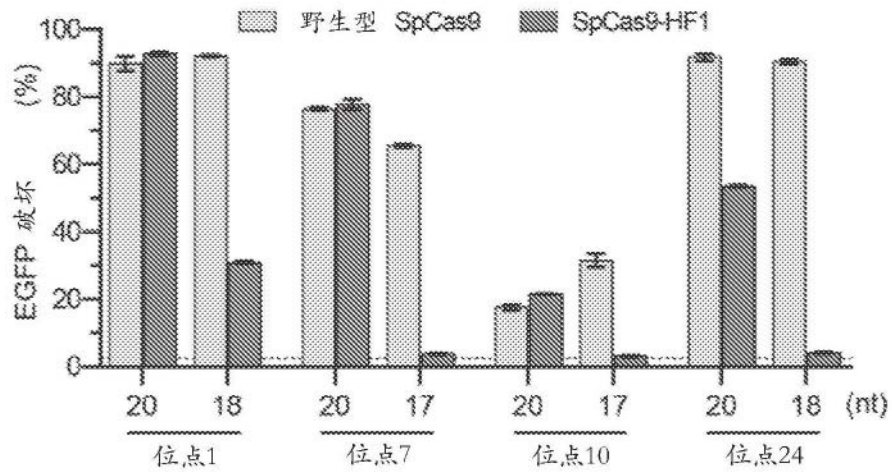


图9

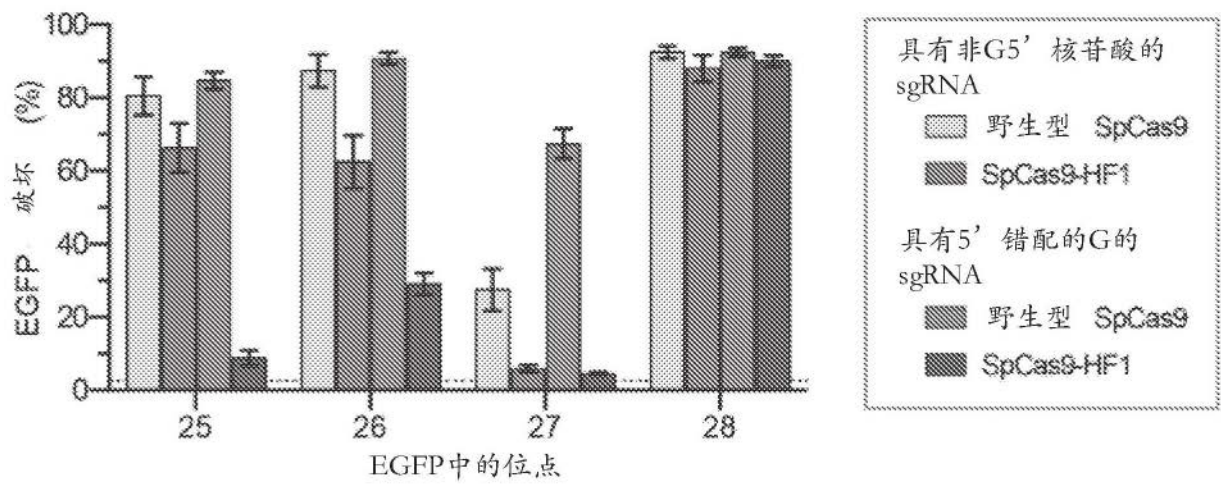


图10

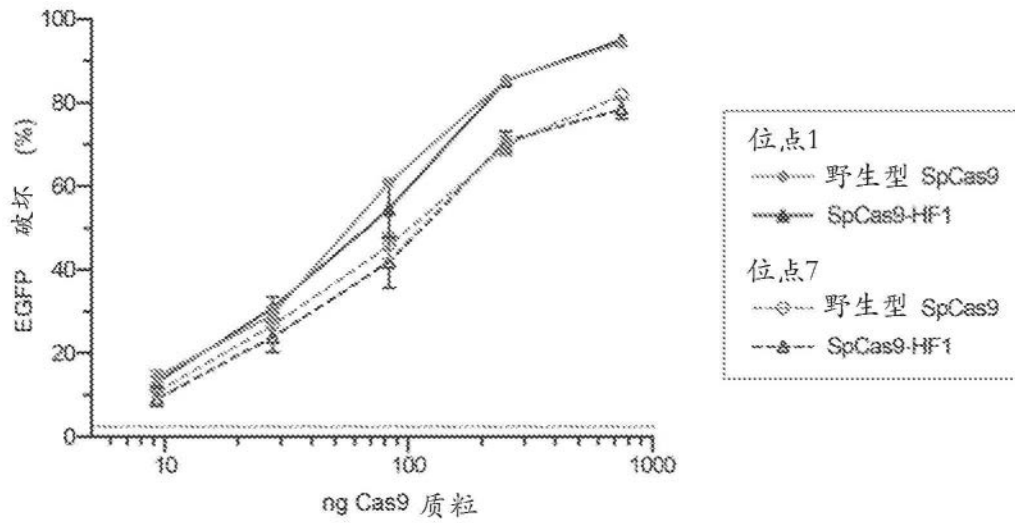


图11

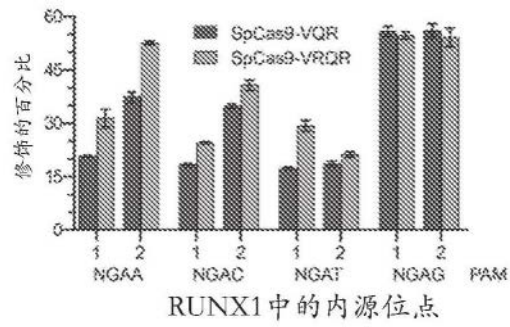


图12A

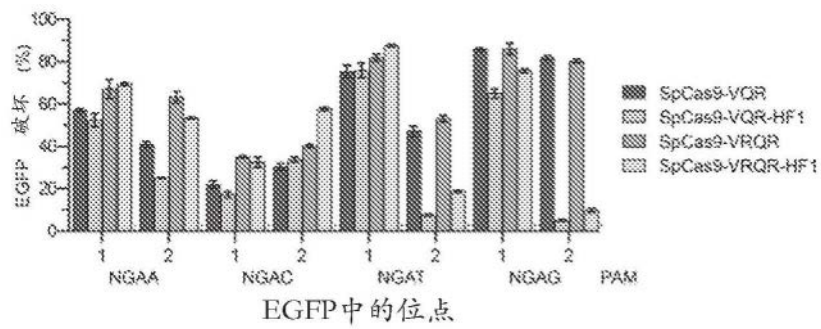


图12B

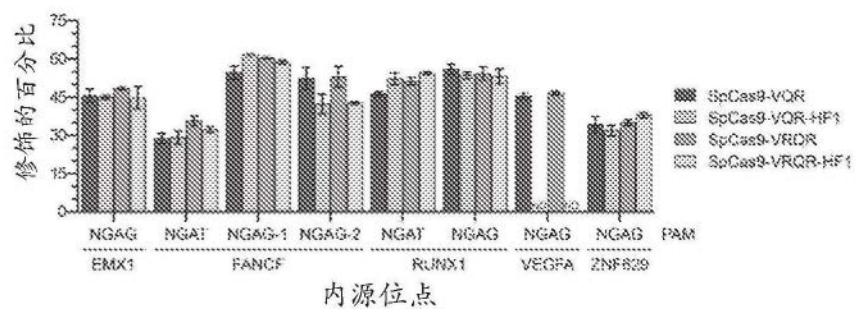


图12C

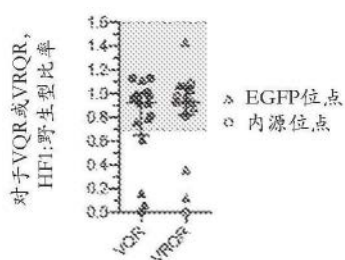


图12D

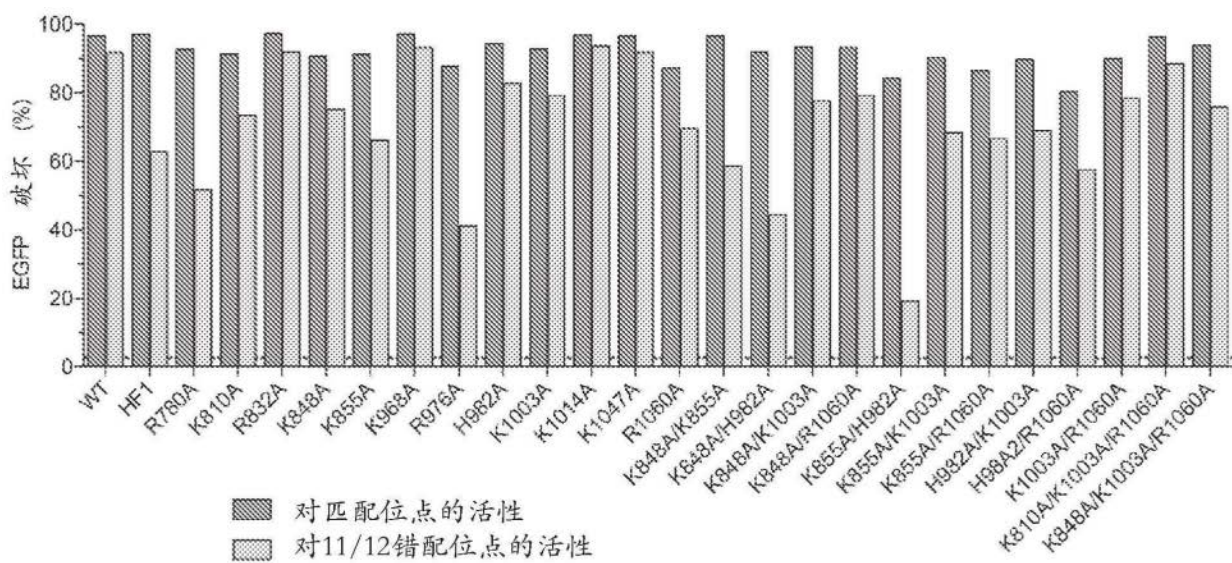


图13A

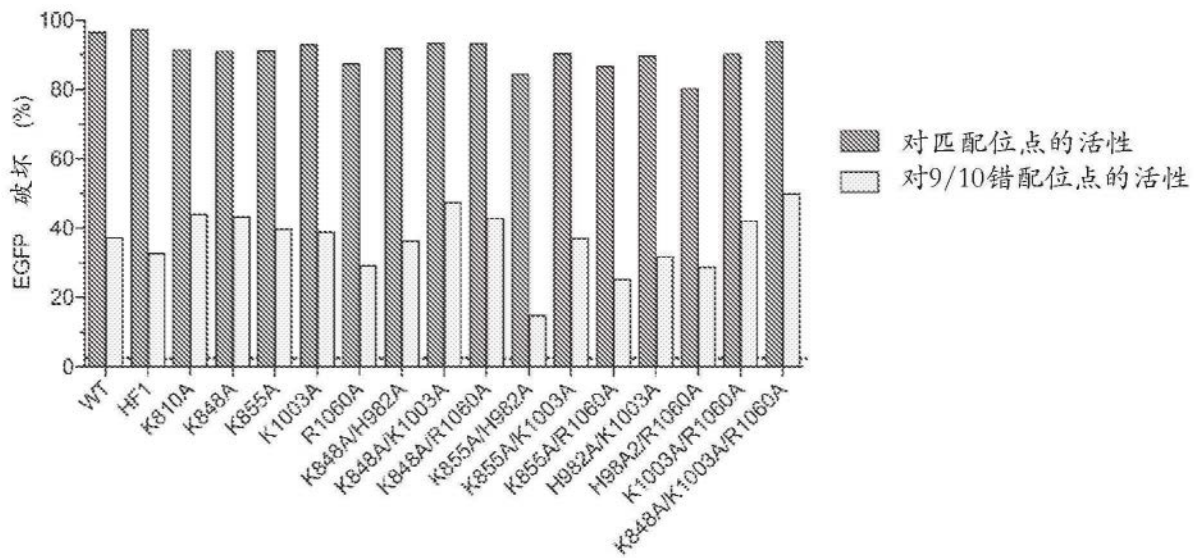


图13B

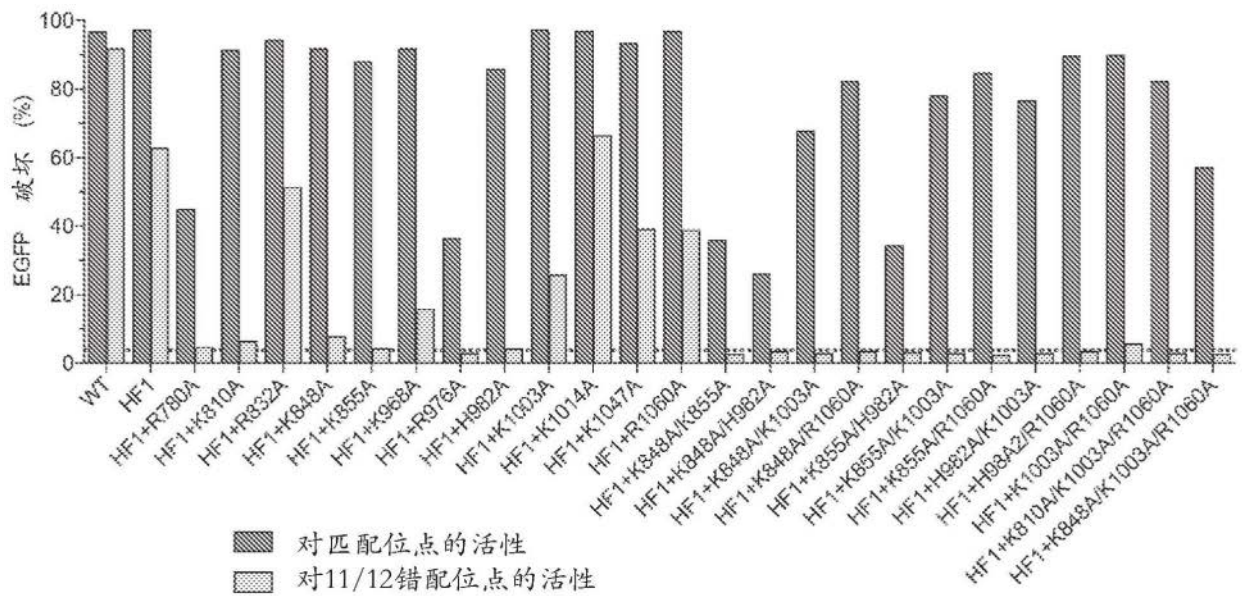


图14A

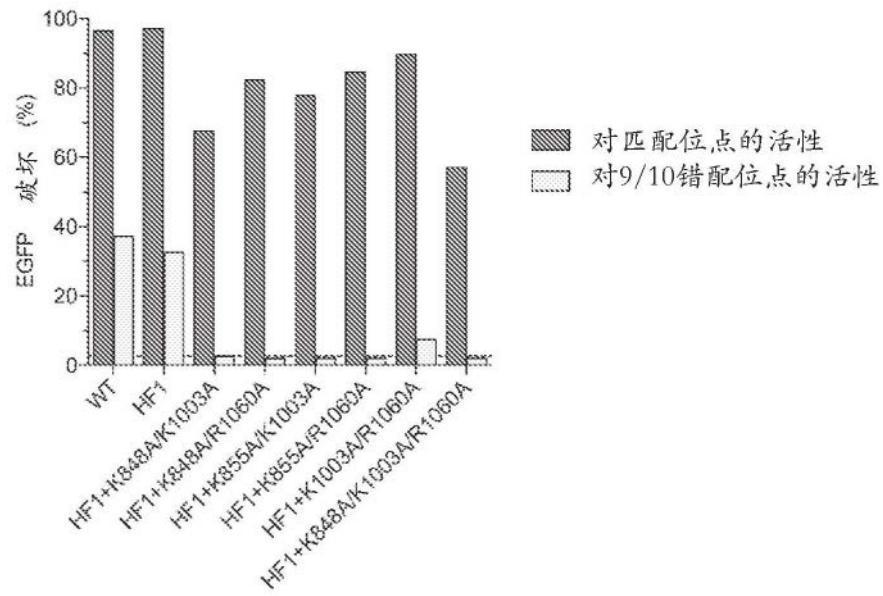


图14B

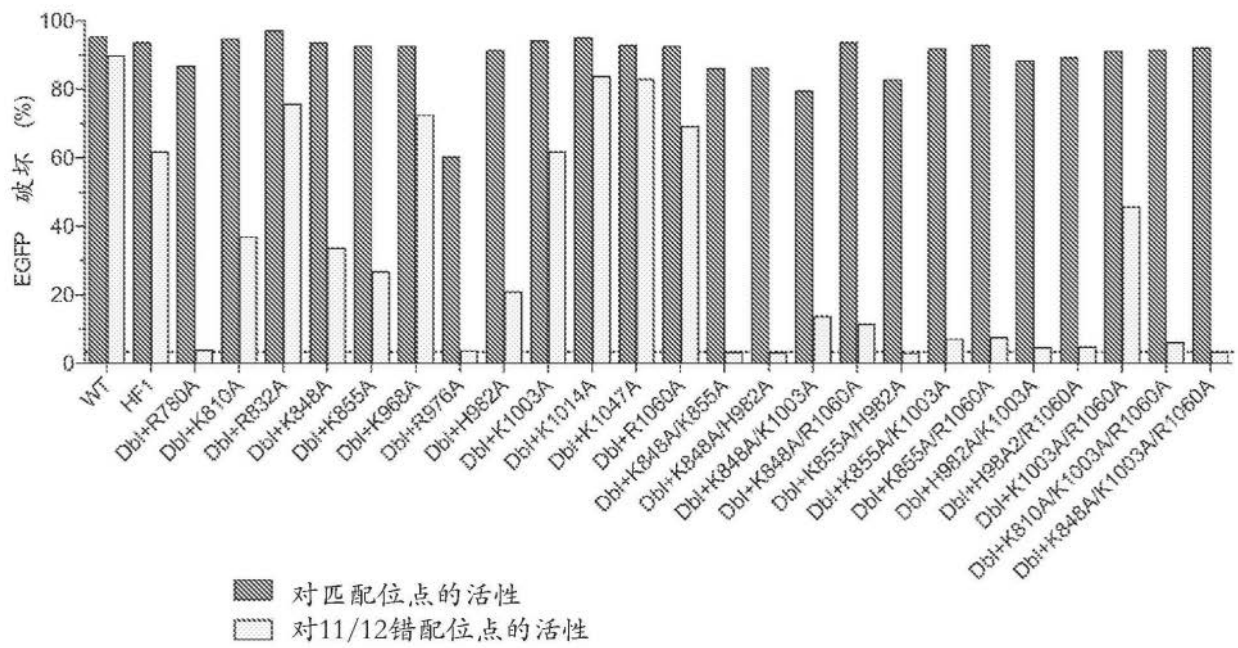


图15

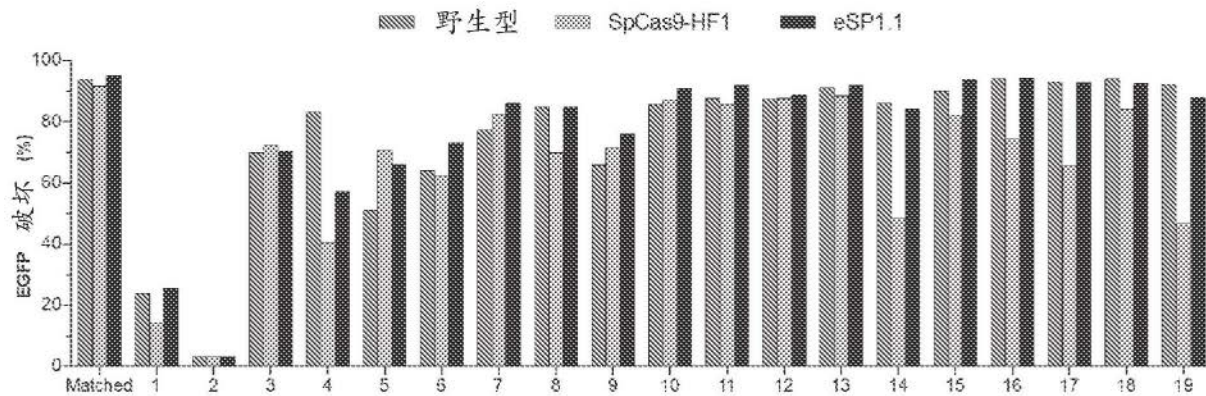


图16

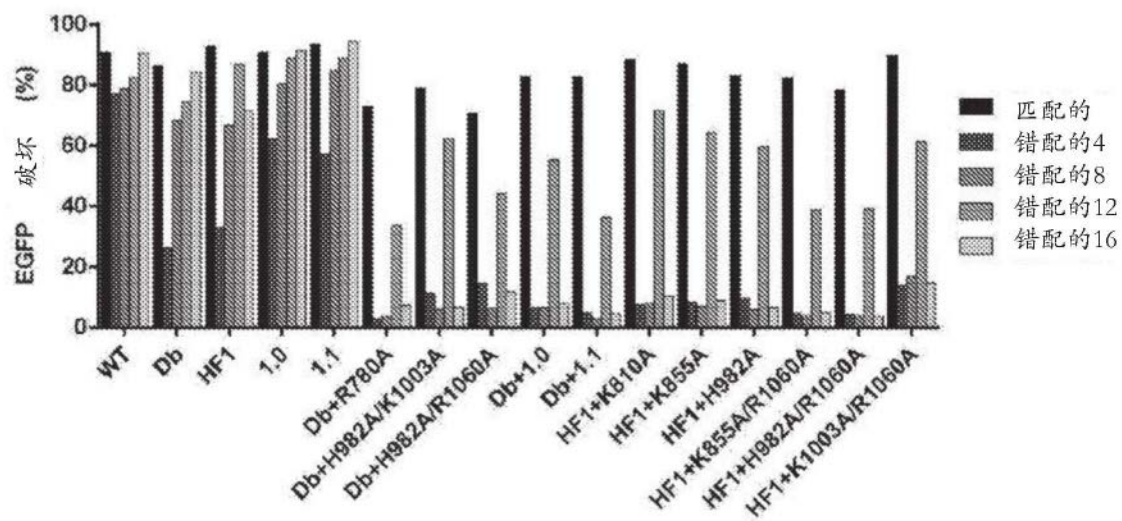


图17A

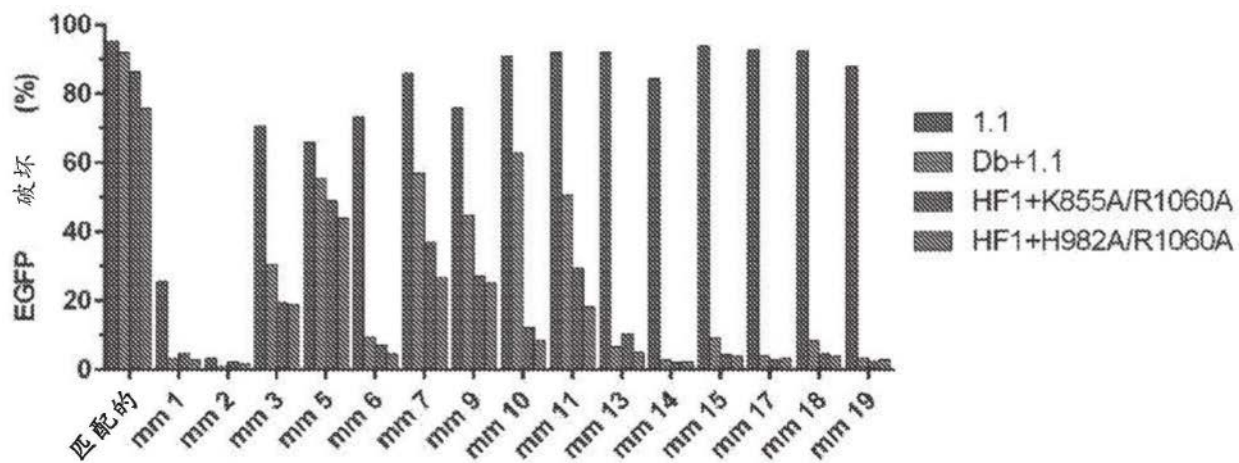


图17B

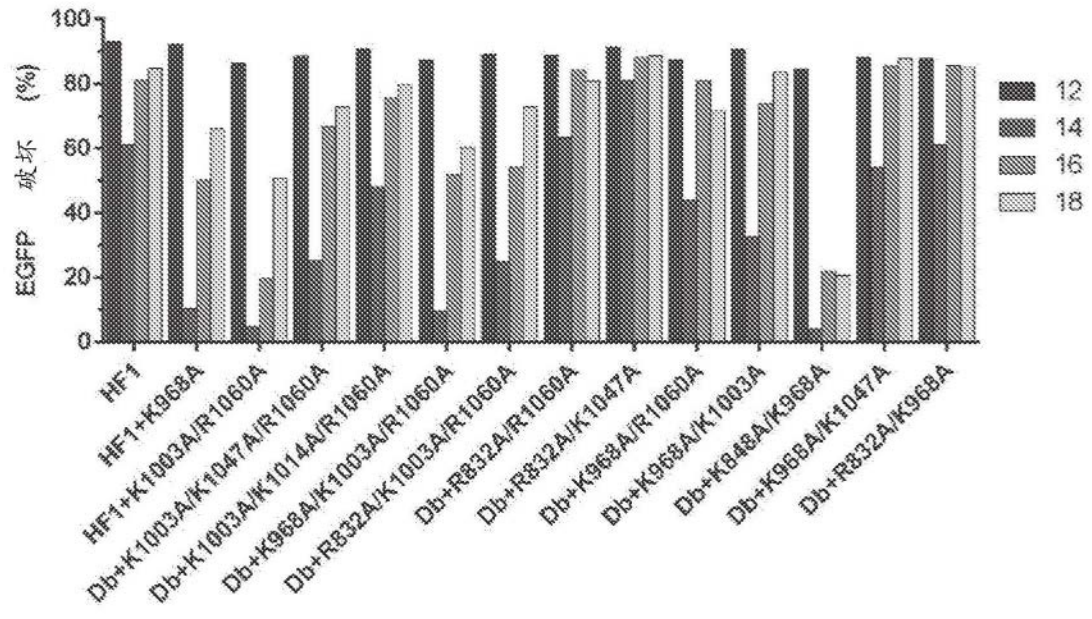
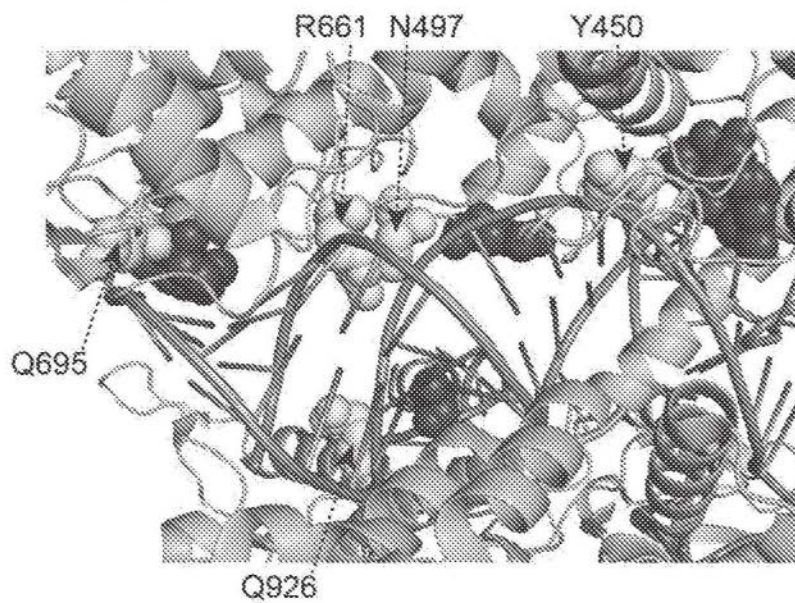


图19B

酿脓链球菌Cas9



金黄色葡萄球菌Cas9

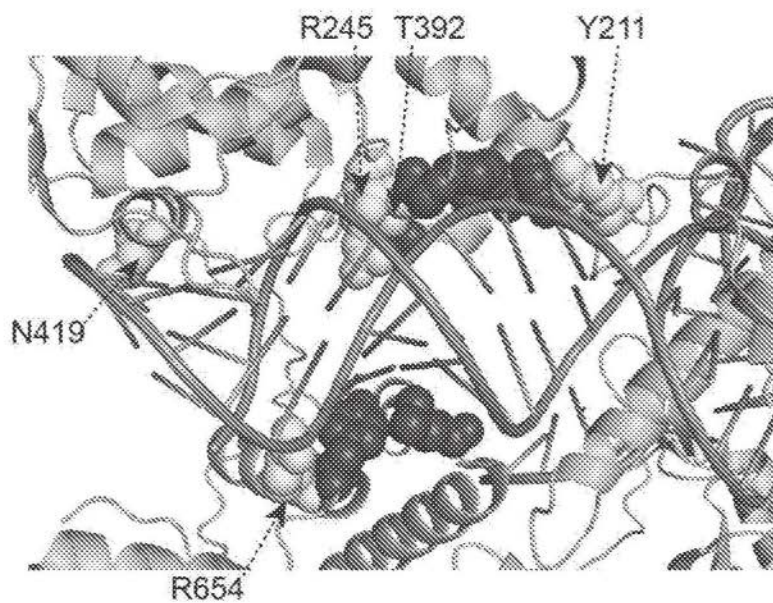


图20

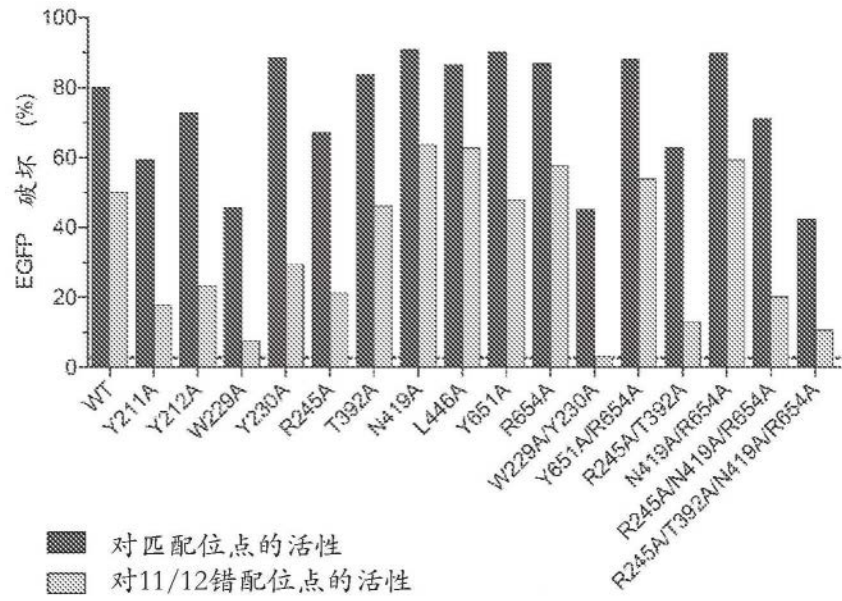


图21A

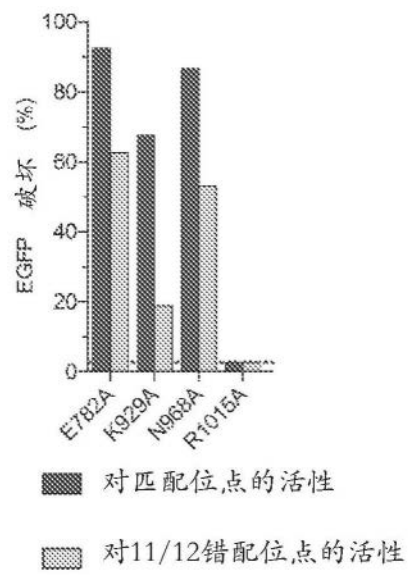


图21B

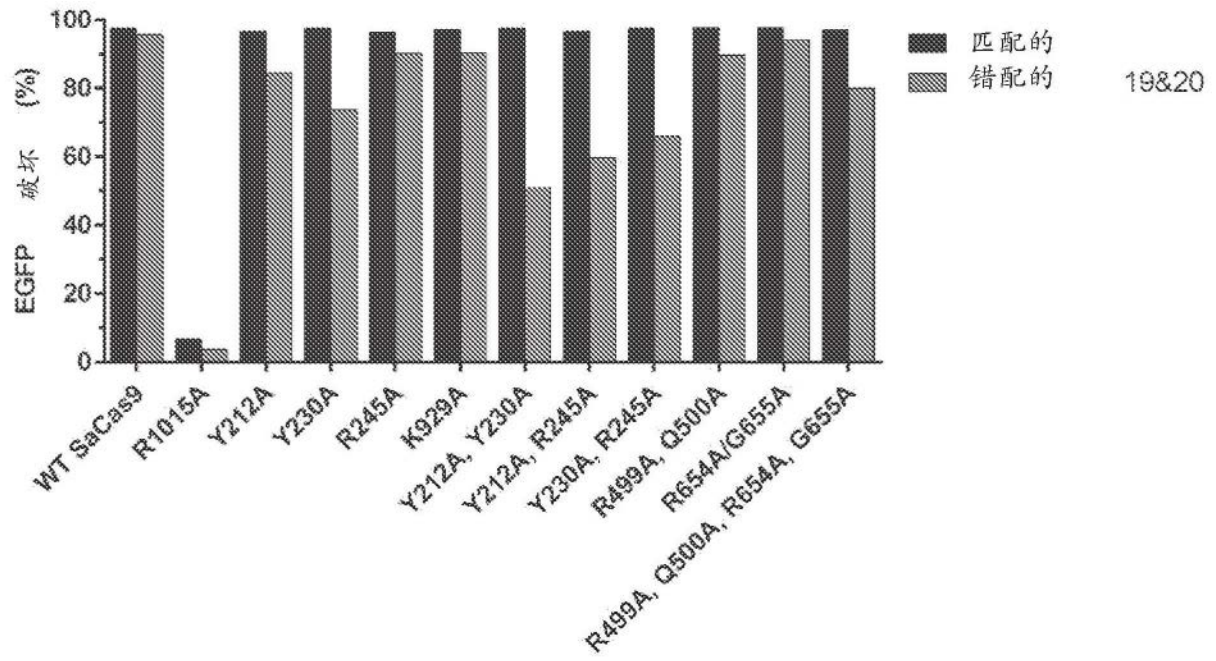


图22A

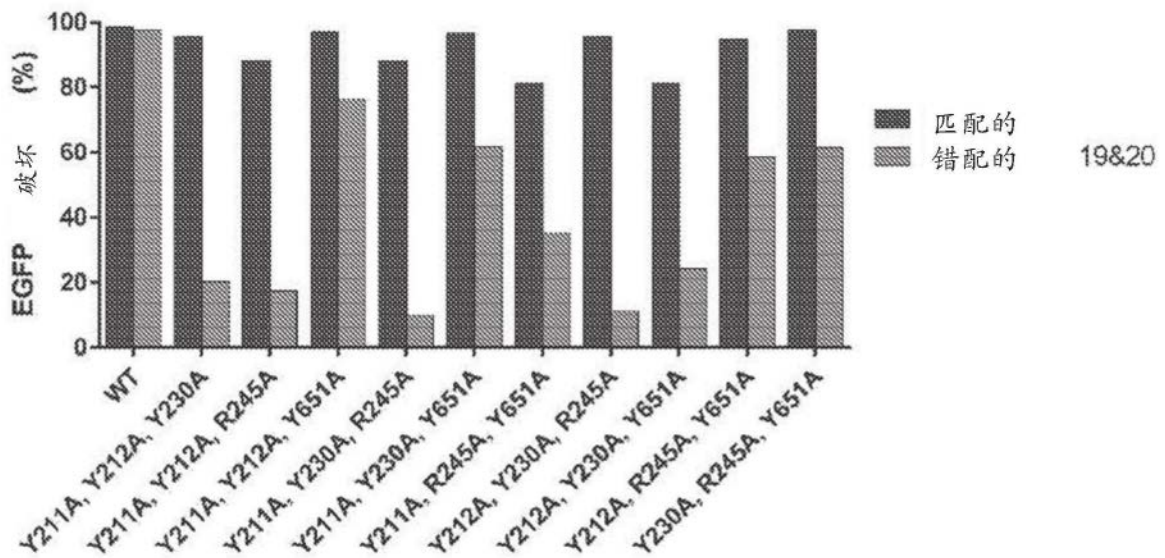


图22B

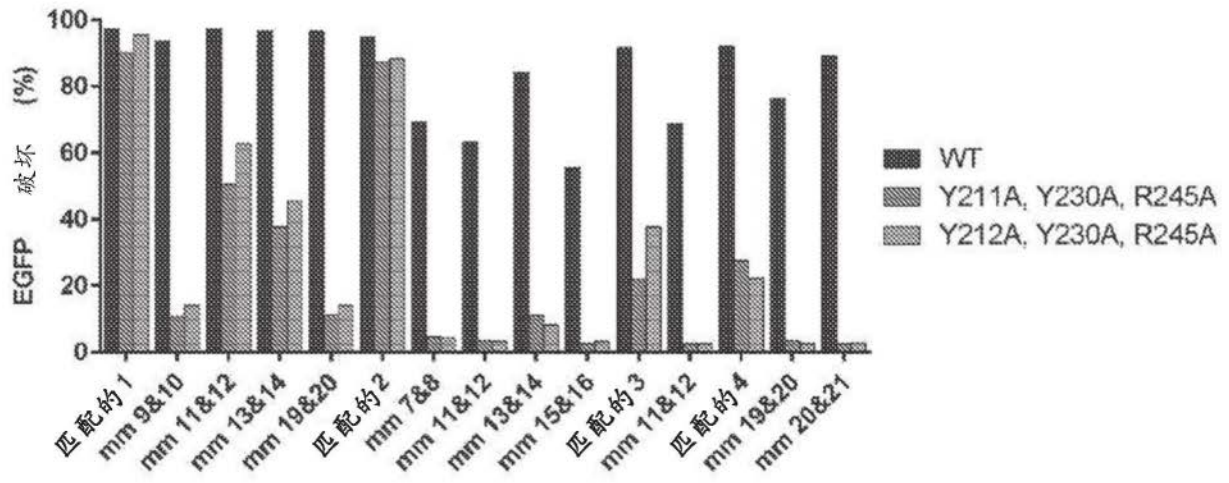


图23

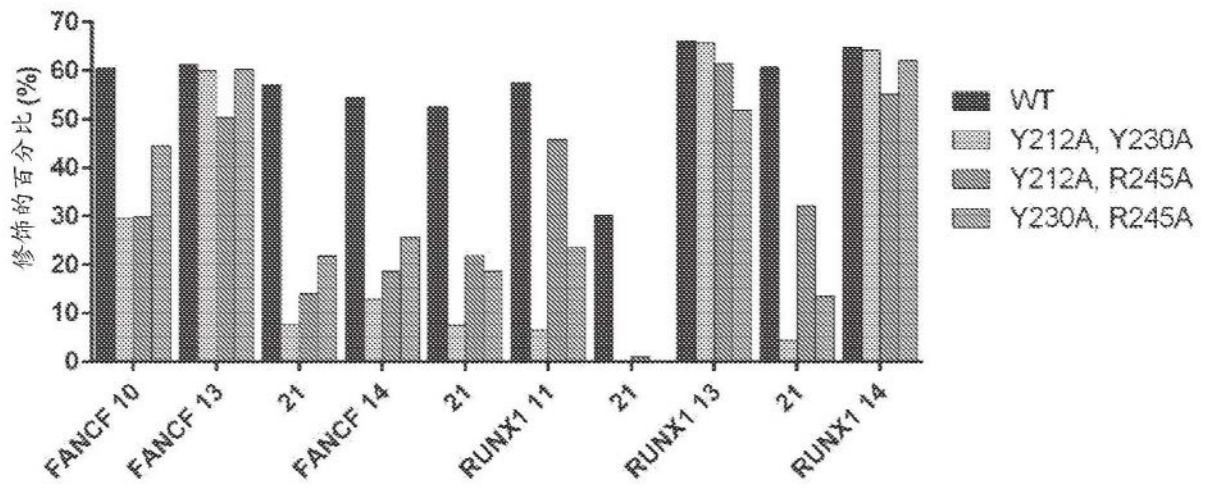


图24A

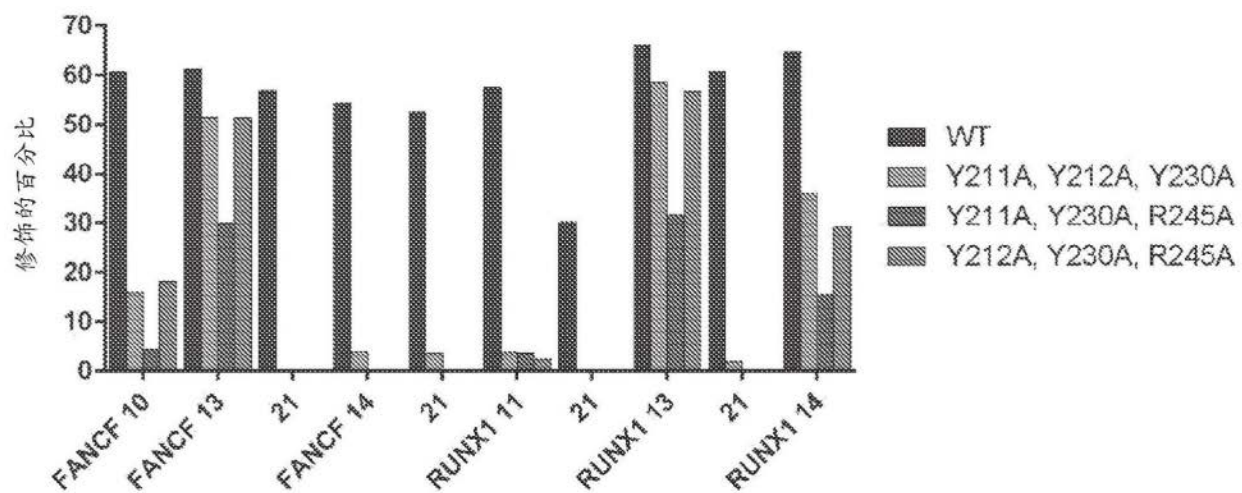


图24B