(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 109265562 B
(45)授权公告日 2021.03.30

(21)申请号 201811122909.9

(22)申请日 2018.09.26

(65)同一申请的已公布的文献号
申请公布号 CN 109265562 A

(43)申请公布日 2019.01.25

(73)专利权人 北京市农林科学院
地址 100097 北京市海淀区曙光花园中路9号北京市农林科学院玉米研究中心

(72)发明人 杨进孝 杨永星 吕欣欣 赵思 冯峰

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245
代理人 关畅 秦梦楠

(51)Int.Cl.
C07K 19/00(2006.01)
C12N 15/62(2006.01)
C12N 15/82(2006.01)
C12N 9/22(2006.01)

(56)对比文件
US 2014179006 A1,2014.06.26
US 2017058271 A1,2017.03.02
CN 108350449 A,2018.07.31
WO 2018156818 A1,2018.08.30
CN 108513575 A,2018.09.07
Janice S.Chen等.Enhanced proofreading governs CRISPR-Cas9 targeting accuracy.《Nature》.2017,第550卷(第7676期),407-410.

审查员 李翠莹

权利要求书1页 说明书6页
序列表18页 附图1页

(54)发明名称

一种切刻酶及其在基因组碱基替换中的应用

(57)摘要

本发明公开了切刻酶及其在基因组碱基替换中的应用。本发明首次将切刻酶HypaCas9n与PmCDA1和UGI融合构建了碱基编辑系统,发现HypaCas9n&PmCDA1&UGI与SpCas9n&PmCDA1&UGI相比,在基本不影响C·T碱基替换效率的情况下,脱靶效率能够降低。

CN 109265562 B

1.融合蛋白,由切刻酶、胞嘧啶核苷脱氨酶PmCDA1和尿嘧啶DNA糖化酶抑制剂UGI组成;所述融合蛋白中所述切刻酶、所述胞嘧啶核苷脱氨酶PmCDA1和所述尿嘧啶DNA糖化酶抑制剂UGI自N端起依次排列;所述切刻酶如序列表的序列13自N端第1-1423位氨基酸所示。

2.权利要求1所述融合蛋白的编码基因。

3.一种碱基编辑系统,包括权利要求1所述的融合蛋白。

4.如权利要求3所述的碱基编辑系统,其特征在于:所述系统还包括sgRNA。

5.表达权利要求3或4所述的碱基编辑系统的重组表达载体、表达盒或重组菌。

6.一种用于基因组碱基替换的重组表达载体,包括表达盒甲和表达盒乙;所述表达盒甲表达权利要求1所述的融合蛋白;所述表达盒乙包括n个元件乙;所述元件乙包括sgRNA和靶序列;所述重组表达载体可靶向n个不同的靶序列进行碱基替换。

7.一种植物基因组碱基替换的方法,包括如下步骤:利用权利要求3或4所述的碱基编辑系统完成植物基因组碱基替换。

8.一种植物基因组碱基替换的方法,包括如下步骤:将权利要求6所述的重组表达载体导入目的植物,实现植物基因组碱基替换。

9.切刻酶,如序列表的序列13自N端第1-1423位氨基酸所示。

10.权利要求1所述的融合蛋白或权利要求3或4所述的碱基编辑系统或权利要求5所述的重组表达载体、表达盒或重组菌或权利要求6所述的重组表达载体或权利要求9所述的切刻酶在植物基因组碱基替换中的应用;所述碱基替换为将碱基C替换为T。

一种切刻酶及其在基因组碱基替换中的应用

技术领域

[0001] 本发明涉及一种切刻酶及其在基因组碱基替换中的应用。

背景技术

[0002] CRISPR-Cas9(the clustered regularly interspaced short palindromic repeats-CRISPR-associatedprotein 9)技术的出现和发展,已经成为强有力的基因组编辑手段,被广泛应用到很多组织和细胞中。CRISPR/Cas9protein-RNA复合物通过向导RNA(guide RNA)定位于靶点上,切割产生DNA双链断裂(DSB,dsDNAbreak),而后,生物体会本能的启动DNA修复机制修复DSB。一般有两种修复机制,一种是非同源末端连接(NHEJ,non-homologous endjoining),一种是同源重组(HDR,homology-directed repair),通常情况下,NHEJ修复占大多数,因此,修复产生的随机的indels(insertions or deletions)比精确修复高很多。对于碱基精确替换,因为HDR效率低以及需要DNA模板,所以使用HDR实现碱基精确替换的应用受到很大的限制。

[0003] 2016年,David Liu和Akihiko Kondo两个实验室分别独立报道了两个不同类型的胞嘧啶碱基编辑器(CBE,cytosine base editor),原理都是通过使用胞苷脱氨酶实现直接对单个的C(胞嘧啶,Cytosine)碱基进行编辑,而不再通过产生DSB和启动HDR修复,大大提高了C替换为T(胸腺嘧啶,Thymine)的碱基编辑效率。PmCDA1(activation-induced cytidine deaminase(AID)ortholog from sea lamprey)就是其中使用的一种胞苷脱氨酶。在所测试的PmCDA1编辑器中,SpCas9n(D10A)&PmCDA1&UGI(尿嘧啶DNA糖化酶抑制剂,uracil DNA glycosylase inhibitor)碱基编辑系统的平均突变率较高,一是因为UGI可以抑制UDG(尿嘧啶DNA糖化酶,uracil DNAglycosylase)催化清除DNA中U(尿嘧啶,Uracil),二是因为SpCas9n(D10A)在非编辑链上产生切口,诱导真核错配修复机制或long-patch BER(base-excision repair)修复机制,促使U:G错配更多的偏好性修复成U:A。SpCas9n(D10A)连带着PmCDA1通过sgRNA定位到靶点,PmCDA1催化非配对的单链DNA上的C发生胞嘧啶脱氨反应变成U,通过DNA的修复使得U与A(腺嘌呤,Adenine)配对,又通过DNA复制,最终使得T与A配对,从而实现了C到T的转换。

[0004] 植物中使用SpCas9进行基因组编辑时具有一定的脱靶效应,SpCas9n(D10A)与PmCDA1融合使用进行碱基编辑,即使用SpCas9n(D10A)&PmCDA1&UGI碱基编辑系统时可能存在潜在的脱靶风险。虽然植物不同于动物,在后代可以通过遗传分离去除脱靶位点,但是由于一些潜在的脱靶位点可能是未知的,所以,后代中就很难有目的性的去除。因此,降低脱靶效应也是植物中一直以来的技术方向。

发明内容

[0005] 本发明的目的是提供一种切刻酶及其在基因组碱基替换中的应用。

[0006] 本发明首先提供了融合蛋白,由切刻酶、胞嘧啶核苷脱氨酶PmCDA1和尿嘧啶DNA糖化酶抑制剂UGI组成;所述切刻酶如序列表的序列13自N端第1-1423位氨基酸所示。

[0007] 所述胞嘧啶核苷脱氨酶PmCDA1如序列表的序列13自N端第1521-1728位氨基酸所示。

[0008] 所述尿嘧啶DNA糖化酶抑制剂UGI如序列表的序列13自N端第1736-1833位氨基酸所示。

[0009] 本发明还保护所述融合蛋白的编码基因。

[0010] 所述融合蛋白的编码基因具体可如序列表的序列2自5'端第1721-7222位所示(其中,第1721-5989位为切刻酶的编码基因,第6281-6904位为胞嘧啶核苷脱氨酶PmCDA1的编码基因,第6926-7222位为尿嘧啶DNA糖化酶抑制剂UGI的编码基因)。

[0011] 本发明还保护一种碱基编辑系统,包括所述融合蛋白。

[0012] 所述系统还包括sgRNA。

[0013] 所述sgRNA的核苷酸序列如序列表的序列1自5'端第571-646所示。

[0014] 本发明还保护表达所述碱基编辑系统的重组表达载体、表达盒、重组细胞或重组菌。表达所述融合蛋白的表达盒具体可为表达盒甲。表达所述sgRNA的表达盒具体可为表达盒乙。

[0015] 本发明还保护一种用于基因组碱基替换的重组表达载体,包括表达盒甲和表达盒乙;所述表达盒甲表达所述融合蛋白;所述表达盒乙包括n个元件乙;所述元件乙包括sgRNA和靶序列;所述重组表达载体可靶向n个不同的靶序列进行碱基替换。

[0016] 所述sgRNA的核苷酸序列如序列表的序列1自5'端第571-646所示。

[0017] 所述元件乙还包括pre-tRNA。所述pre-tRNA的核苷酸序列如序列表的序列1自5'端第第474-550位所示。所述元件乙自5'端依次为pre-tRNA、靶点序列和sgRNA的核苷酸序列。

[0018] 以上任一所述表达盒甲由启动子甲启动切刻酶的编码基因、胞嘧啶核苷脱氨酶PmCDA1的编码基因和尿嘧啶DNA糖化酶抑制剂UGI的编码基因表达。所述表达盒甲自5'端依次包括启动子甲、切刻酶的编码基因、胞嘧啶核苷脱氨酶PmCDA1的编码基因、尿嘧啶DNA糖化酶抑制剂UGI的编码基因和终止子甲。所述启动子甲具体可为OsUbq3启动子。所述OsUbq3启动子的核苷酸序列如序列表的序列2自5'端第1-1714位所示。所述终止子甲具体可为CaMV35S终止子。所述CaMV35S终止子的核苷酸序列如序列表的序列2自5'端第7229-7423位所示。所述切刻酶的编码基因如序列表的序列2自5'端第1721-5989位所示。所述胞嘧啶核苷脱氨酶PmCDA1的编码基因如序列表的序列2自5'端第6281-6904位所示。所述尿嘧啶DNA糖化酶抑制剂UGI编码基因如序列表的序列2自5'端第6926-7222位所示。所述表达盒甲具体可如序列表的序列2所示。

[0019] 以上任一所述表达盒乙由启动子乙启动元件乙表达。所述表达盒乙自5'端依次为启动子乙、元件乙和终止子乙。所述启动子乙具体可为OsU3启动子。所述OsU3启动子的核苷酸序列如序列表的序列1自5'端第131-467位所示。所述终止子乙具体可为OsU3终止子。所述OsU3终止子的核苷酸序列如序列表的序列1自5'端第993-1283位所示。当靶序列如表1所示时,所述表达盒乙具体可如序列表的序列1自5'端第131位-1283位所示。

[0020] 当靶序列如表1所示时,以上任一所述重组表达载体具体可为将序列表的序列1自5'端第1290-8712替换为序列表的序列2得到的环状质粒。

[0021] 本发明还保护一种植物基因组碱基替换的方法,包括如下步骤:利用以上任一所

述的碱基编辑系统完成植物基因组碱基替换。

[0022] 本发明还保护一种植物基因组碱基替换的方法，包括如下步骤：将以上任一所述的重组表达载体导入目的植物，实现植物基因组碱基替换。

[0023] 本发明还保护切刻酶，如序列表的序列13自N端第1-1423位氨基酸所示。

[0024] 本发明还保护所述融合蛋白，或，所述碱基编辑系统，或，以上任一所述重组表达载体、表达盒、重组细胞或重组菌，或，所述切刻酶在植物基因组碱基替换中的应用。

[0025] 以上任一所述碱基替换为将碱基C替换为T。

[0026] 以上任一所述所述植物具体可为水稻，更具体可为日本晴水稻。

[0027] 本发明首次将切刻酶HypaCas9n与PmCDA1和UGI融合构建了碱基编辑系统，发现HypaCas9n&PmCDA1&UGI与SpCas9n&PmCDA1&UGI相比，在基本不影响C·T碱基替换效率的情况下，脱靶效率能够降低。

附图说明

[0028] 图1为SpCas9n&PmCDA1&UGI与HypaCas9n&PmCDA1&UGI的C·T碱基替换效率。

[0029] 图2为SpCas9n&PmCDA1&UGI与HypaCas9n&PmCDA1&UGI的脱靶效率。

具体实施方式

[0030] 以下的实施例便于更好地理解本发明，但并不限定本发明。下述实施例中的实验方法，如无特殊说明，均为常规方法。下述实施例中所用的试验材料，如无特殊说明，均为自常规生化试剂商店购买得到的。

[0031] 日本晴水稻：参考文献：梁卫红，王高华，杜京尧，等.硝普钠及其光解产物对日本晴水稻幼苗生长和5种激素标记基因表达的影响[J].河南师范大学学报(自然版)，2017(2):48-52.；公众可以从北京市农林科学院获得。

[0032] 下述实施例中靶标基因、靶点名称和序列如表1所示。

[0033] 表1

[0034]

靶标基因	靶点名称	靶点序列
OsALS	CS650	cgcgtccatggagatccacc
OsCDC48	CS651	gaccagccagcgtctggcgc
OsNRT1.1B	CS652	cggcgacggcgagcaagtgg

[0035] 实施例1、HypaCas9n&PmCDA1&UGI系统C·T碱基替换效率

[0036] 一、基因组编辑载体的构建

[0037] 1、SpCas9n&PmCDA1&UGI载体：人工合成序列表的序列1所示的环状质粒。

[0038] 序列表的序列1包括如下三个表达盒：

[0039] 序列1自5’端第131位-1283位为表达盒I，其中，第131-467位为OsU3启动子的核苷酸序列，第474-550位为pre-tRNA的核苷酸序列，第551-570位为CS650靶点的核苷酸序列，第571-646位为sgRNA的核苷酸序列，第647-723位为pre-tRNA的核苷酸序列，第724-743位为CS651靶点的核苷酸序列，第744-819位为sgRNA的核苷酸序列，第820-896位为pre-tRNA的核苷酸序列，第897-916位为CS652靶点的核苷酸序列，第917-992位为sgRNA的核苷酸序列，第993-1283位为OsU3终止子的核苷酸序列。

[0040] 序列1自5’端第1290-8712位为表达盒Ⅱ，其中，第1290-3003位为OsUbq3启动子的核苷酸序列，第3010-7278位为SpCas9n的核苷酸序列(不含有终止密码子)，第7570-8193位为PmCDA1的核苷酸序列，第8215-8511位为UGI的核苷酸序列，第8518-8712位为CaMV35S终止子的核苷酸序列。

[0041] 序列1自5’端第8787-12064位为表达盒Ⅲ，其中，第8787-10779为ZmUbi1启动子的核苷酸序列，第10786-11811位为潮霉素的核苷酸序列，第11812-12064位为Nos终止子的核苷酸序列。

[0042] 2、HypaCas9n&PmCDA1&UGI载体：将序列表的序列1自5’端第1290-8712(表达盒Ⅱ)替换为序列表的序列2得到的环状质粒。HypaCas9n&PmCDA1&UGI载体与SpCas9n&PmCDA1&UGI载体的区别仅在于将SpCas9n的核苷酸序列(不含有终止密码子)替换为HypaCas9n的核苷酸序列(不含有终止密码子)。

[0043] 序列表的序列2中，自5’端第1-1714位为OsUbq3启动子的核苷酸序列，第1721-5989位为HypaCas9n的核苷酸序列(不含有终止密码子)，第6281-6904位为PmCDA1的核苷酸序列，第6926-7222位为UGI的核苷酸序列，第7229-7423位CaMV35S终止子的核苷酸序列。

[0044] 二、在水稻愈伤组织中进行基因编辑

[0045] 将步骤一构建的SpCas9n&PmCDA1&UGI载体和HypaCas9n&PmCDA1&UGI载体分别按照如下步骤1-5进行操作：

[0046] 1、将载体导入农杆菌LBA4404(唯地生物，上海，CAT#:AC1030)，得到重组菌，采用YEP培养基培养重组菌，得到菌液OD_{600nm}为0.2的侵染液。

[0047] 2、选取日本晴水稻种子，剥去种皮，灭菌洗涤后，均匀的点入在N6培养基中，28℃暗培养4-6周以诱导愈伤组织的产生。

[0048] 3、将步骤2得到的水稻愈伤组织浸泡在步骤1得到的侵染液中浸泡10min，取愈伤组织接种于含有两层滤纸的培养皿上，25℃黑暗下培养3天(培养基为含有100mg/L的特美汀的N6培养基)，然后愈伤组织在筛选培养基中(含有50mg/L潮霉素的N6培养基，pH5.7)28℃黑暗下筛选培养2周，转入新配置的筛选培养基中再次进行筛选培养2周，获得抗性愈伤。

[0049] 4、提取步骤3得到的抗性愈伤的基因组DNA，采用引物F(5’-attatgtagcttgtgcgtttcg-3’)和引物R(5’-gatgaagagcttatcgacgt-3’)组成的引物对进行PCR扩增；将获得的扩增产物进行琼脂糖凝胶电泳，如电泳图显示有1150bp大小的条带，说明对应的愈伤为阳性抗性愈伤。

[0050] 5、对步骤4得到的阳性抗性愈伤DNA(每个载体选择15块)，分别采用CS650靶点引物(CS650-F和CS650-R)、CS651靶点引物(CS651-F和CS651-R)和CS652靶点引物(CS652-F和CS652-R)进行PCR扩增，然后将扩增产物测序。

[0051] CS650-F:5’-taagaaccaccagcgacacc-3’；

[0052] CS650-R:5’-ggtaattgtgcttggtgatggag-3’；

[0053] CS651-F:5’-acatcgagatggagaagcgg-3’；

[0054] CS651-R:5’-ccatgctccaatcgatgaatac-3’；

[0055] CS652-F:5’-ttacgaactttataactttgtcgg-3’；

[0056] CS652-R:5’-atggaggcgatgaggaagac-3’。

[0057] 对于CS650靶点，发生C·T碱基替换的位点对应为：cgTgtccatggagatccacc；

[0058] 对于CS651靶点,发生C·T碱基替换的位点对应为:gaTTagccagcgtctggcgc;

[0059] 对于CS652靶点,发生C·T碱基替换的位点对应为:cggTgacggcgagcaagtgg。

[0060] 三个靶点分别统计发生C·T碱基替换的阳性抗性愈伤数,碱基替换效率为15块阳性抗性愈伤中发生C·T碱基替换的阳性抗性愈伤所占的比例。

[0061] 结果如图1所示。实验结果表明,HypaCas9n&PmCDA1&UGI与SpCas9n&PmCDA1&UGI的C·T碱基替换效率相差不大或略有降低。

[0062] 实施例2、HypaCas9n&PmCDA1&UGI系统脱靶效应

[0063] 一、基因组编辑载体的构建

[0064] SpCas9n&PmCDA1&UGI-T1:将序列表的序列1自5'端第551-916位替换为序列表的序列3得到的环状质粒。

[0065] SpCas9n&PmCDA1&UGI-T2:将序列表的序列1自5'端第551-916位替换为序列表的序列4得到的环状质粒。

[0066] SpCas9n&PmCDA1&UGI-T3:将序列表的序列1自5'端第551-916位替换为序列表的序列5得到的环状质粒。

[0067] SpCas9n&PmCDA1&UGI-T4:将序列表的序列1自5'端第551-916位替换为序列表的序列6得到的环状质粒。

[0068] SpCas9n&PmCDA1&UGI-T5:将序列表的序列1自5'端第551-916位替换为序列表的序列7得到的环状质粒。

[0069] SpCas9n&PmCDA1&UGI-T6:将序列表的序列1自5'端第551-916位替换为序列表的序列8得到的环状质粒。

[0070] SpCas9n&PmCDA1&UGI-T7:将序列表的序列1自5'端第551-916位替换为序列表的序列9得到的环状质粒。

[0071] SpCas9n&PmCDA1&UGI-T8:将序列表的序列1自5'端第551-916位替换为序列表的序列10得到的环状质粒。

[0072] SpCas9n&PmCDA1&UGI-T9:将序列表的序列1自5'端第551-916位替换为序列表的序列11得到的环状质粒。

[0073] SpCas9n&PmCDA1&UGI-T10:将序列表的序列1自5'端第551-916位替换为序列表的序列12得到的环状质粒。

[0074] HypaCas9n&PmCDA1&UGI-T1:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替换为序列表的序列3得到的环状质粒。

[0075] HypaCas9n&PmCDA1&UGI-T2:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替换为序列表的序列4得到的环状质粒。

[0076] HypaCas9n&PmCDA1&UGI-T3:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替换为序列表的序列5得到的环状质粒。

[0077] HypaCas9n&PmCDA1&UGI-T4:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替换为序列表的序列6得到的环状质粒。

[0078] HypaCas9n&PmCDA1&UGI-5:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替换为序列表的序列7得到的环状质粒。

[0079] HypaCas9n&PmCDA1&UGI-T6:将HypaCas9n&PmCDA1&UGI载体自5'端第551-916位替

换为序列表的序列8得到的环状质粒。

[0080] HypaCas9n&PmCDA1&UGI-T7:将HypaCas9n&PmCDA1&UGI载体自5’端第551-916位替换为序列表的序列9得到的环状质粒。

[0081] HypaCas9n&PmCDA1&UGI-T8:将HypaCas9n&PmCDA1&UGI载体自5’端第551-916位替换为序列表的序列10得到的环状质粒。

[0082] HypaCas9n&PmCDA1&UGI-T9:将HypaCas9n&PmCDA1&UGI载体自5’端第551-916位替换为序列表的序列11得到的环状质粒。

[0083] HypaCas9n&PmCDA1&UGI-T10:将HypaCas9n&PmCDA1&UGI载体自5’端第551-916位替换为序列表的序列12得到的环状质粒。

[0084] 二、在水稻愈伤组织中进行基因编辑

[0085] 1、将步骤一得到的载体分别按照实施例1中步骤二的1-4进行操作,得到阳性抗性愈伤。

[0086] 2、每个载体从步骤1得到的阳性抗性愈伤中随机选取8块,将其DNA进行混合,分别采用CS652靶点引物(CS652-F和CS652-R)进行第一轮PCR扩增,以第一轮PCR产物为模板,将正向和反向条码加入PCR产物末端进行文库构建,使用IlluminaHiSeq2500高通量测序平台测序,测序深度为10000X(石家庄博瑞迪生物技术有限公司),目标靶点区检测C·T碱基替换和indels,靶点区任何C·T碱基替换和indels均计入脱靶突变,脱靶效率为检测到的有突变的reads数占总reads数的比例。

[0087] 结果如图2所示。实验结果表明,当5’端2bp存在差异时,会发生脱靶效应;与SpCas9n&PmCDA1&UGI相比,HypaCas9n&PmCDA1&UGI发生的脱靶效率降低约4倍。

[0088] 上述结果表明,在HypaCas9中引入D10A突变,变为切刻酶(HypaCas9n),并与PmCDA1和UGI融合构建了碱基编辑系统。HypaCas9n&PmCDA1&UGI与SpCas9n&PmCDA1&UGI相比,在基本不影响C·T碱基替换效率的情况下,脱靶效率能够降低。

[0001] 序列表

[0002] <110> 北京市农林科学院

[0003] <120> 一种切刻酶及其在基因组碱基替换中的应用

[0004] <160> 13

[0005] <170> SIPOSequenceListing 1.0

[0006] <210> 1

[0007] <211> 18476

[0008] <212> DNA

[0009] <213> 人工序列(Artificial Sequence)

[0010] <400> 1

[0011] ggtggcagga tatattgtgg tgtaaacatg gcactagcct caccgtcttc gcagacgagg 60

[0012] ccgctaagtc gcagctacgc tctcaacggc actgactagg tagtttaaac gtgcacttaa 120

[0013] ttaaggtacc gaagcaactt aaagttatca ggcatgcatg gatcttggag gaatcagatg 180

[0014] tgcagtcagg gaccatagca caagacaggc gtcttctact ggtgctacca gcaaatgctg 240

[0015] gaagccggga acactgggta cgttggaaac cacgtgatgt gaagaagtaa gataaactgt 300

[0016] aggagaaaag catttcgtag tgggccatga agcctttcag gacatgtatt gcagtatggg 360

[0017] ccggcccatt acgcaattgg acgacaacaa agactagtat tagtaccacc tcggctatcc 420

[0018] acatagatca aagctgattt aaaagagttg tgcagatgat ccgtggcgga tccaacaaag 480

[0019] caccagtggt ctagtggtag aatagtaccc tgccacggta cagacccggg ttcgattccc 540

[0020] ggctggtgca cgcgtccatg gagatccacc gttttagagc tagaaatagc aagttaaaat 600

[0021] aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgcaaca aagcaccagt 660

[0022] ggtctagtgg tagaatagta ccctgccacg gtacagaccc gggttcgatt cccggctggt 720

[0023] gcagaccagc cagcgtctgg cgcgttttag agctagaaat agcaagttaa aataaggcta 780

[0024] gtccgttatc aacttgaaaa agtggcaccg agtcggtgca acaaagcacc agtggtctag 840

[0025] tggtagaata gtaccctgcc acggtacaga cccgggttcg attcccggct ggtgcacggc 900

[0026] gacggcgagc aagtgggttt tagagctaga aatagcaagt taaaataagg ctagtccgtt 960

[0027] atcaacttga aaaagtggca ccgagtcggt gctttttttt ttcgttttgc attgagtttt 1020

[0028] ctccgtcgca tgtttgcagt tttattttcc gttttgcatt gaaatttctc cgtctcatgt 1080

[0029] ttgcagcgtg ttcaaaaagt acgcagctgt atttcactta tttacggcgc cacattttca 1140

[0030] tgccgtttgt gccaactatc ccgagctagt gaatacagct tggcttcaca caacactggt 1200

[0031] gacccgctga cctgctcgta cctcgtaccg tcgtacggca cagcatttgg aattaaaggg 1260

[0032] tgtgatcgat actgcttgct gctaagctta caaattcggg tcaaggcgga agccagcgcg 1320

[0033] ccaccccacg tcagcaaata cggaggcgcg gggttgacgg cgtcacccgg tcctaacggc 1380

[0034] gaccaacaaa ccagccagaa gaaattacag taaaaaaaaa gtaaattgca ctttgatcca 1440

[0035] ccttttatta cctaagtctc aatttggatc acccttaaac ctatcttttc aatttgggcc 1500

[0036] gggttgtggt ttggactacc atgaacaact tttcgtcatg tctaacttcc ctttcagcaa 1560

[0037] acatatgaac catatataga ggagatcggc cgtatactag agctgatgtg tttaaggtcg 1620

[0038] ttgattgcac gagaaaaaaa aatccaaatc gcaacaatag caaatttatc tggttcaaag 1680

[0039] tgaaaagata tgtttaaagg tagtccaaag taaaacttat agataataaa atgtggtcca 1740

[0040] aagcgtaatt cactcaaaaa aaatcaacga gacgtgtacc aaacggagac aaacggcatc 1800

[0041] ttctcgaaat ttcccaaccg ctcgctcgcc cgcctcgtct tcccggaaac cgcggtggtt 1860

[0042] tcagcgtggc ggattctcca agcagacgga gacgtcacgg cacgggactc ctcccaccac 1920

[0043] ccaaccgcca taaataccag ccccctcatc tcctctcctc gcatcagctc cacccccgaa 1980

[0044] aaatttctcc ccaatctcgc gaggctctcg tcgtcgaatc gaatcctctc gcgtcctcaa 2040

[0045] ggtacgctgc ttctcctctc ctcgcttcgt ttcgattcga tttcggacgg gtgaggttgt 2100

[0046] tttgttgcta gatccgattg gtggttaggg ttgtcgatgt gattatcgtg agatgtttag 2160

[0047] gggttgtaga tctgatggtt gtgatttggg cacggttggt tcgataggtg gaatcgtggt 2220

[0048] taggttttgg gattggatgt tggttctgat gattgggggg aatttttacg gttagatgaa 2280

[0049] ttgttggatg attcgattgg ggaaatcggt gtagatctgt tggggaattg tggaactagt 2340

[0050] catgcctgag tgattggtgc gatttgtagc gtgttccatc ttgtaggcct tgttgcgagc 2400

[0051] atgttcagat ctactgttcc gctcttgatt gagttattgg tgccatgggt tggtgcaaac 2460

[0052] acaggcttta atatgttata tctgttttgt gtttgatgta gatctgtagg gtagttcttc 2520

[0053] ttagacatgg ttcaattatg tagcttgtgc gtttcgattt gatttcatat gttcacagat 2580

[0054] tagataatga tgaactcttt taattaattg tcaatggtaa ataggaagtc ttgtcgctat 2640

[0055] atctgtcata atgatctcat gttactatct gccagtaatt tatgctaaga actatattag 2700

[0056] aatatcatgt tacaatctgt agtaatatca tgttacaatc tgtagttcat ctatataatc 2760

[0057] tattgtggta atttcttttt actatctgtg tgaagattat tgccactagt tcattctact 2820

[0058] tatttctgaa gttcaggata cgtgtgctgt tactacctat ctgaatacat gtgtgatgtg 2880

[0059] cctgttacta tctttttgaa tacatgtatg ttctgttgga atatgtttgc tgtttgatcc 2940

[0060] gttgttgtgt ccttaatctt gtgctagttc ttaccctatc tgtttggtga ttatttcttg 3000

[0061] cagtacgtaa tggactacaa ggaccacgac ggggattaca aagaccacga catagactac 3060

[0062] aaggatgacg atgacaaaat ggcaccgaag aaaaaaagga aggtcggaat ccatggcgtt 3120

[0063] ccagctgccg ataagaaata ttccatcgga ctcgccattg gcacgaatag cgtcggatgg 3180

[0064] gctgttatta ctgatgagta caaagttccg tctaagaagt tcaaggtgct gggcaacaca 3240

[0065] gaccgccaca gcataaagaa aaatctcatc ggtgcactcc ttttcgatag tggggagact 3300

[0066] gcagaagcga caagattgaa aaggactgcg agaaggcgct atacacggcg taagaataga 3360

[0067] atctgctacc ttcaggagat tttctctaac gaaatggcta aggtcgatga cagtttcttt 3420

[0068] catagacttg aggaatcgtt cttggttgag gaggataaga aacatgagag gcacccgata 3480

[0069] tttggaaaca tcgtggatga ggtcgcatat catgaaaagt accccacaat ctaccacctg 3540

[0070] agaaagaaac tcgttgattc caccgacaaa gcggatttga gactcatcta cctcgctctt 3600

[0071] gcccatatga taaagttccg cggacacttt ctgatcgagg gcgacctcaa ccctgataat 3660

[0072] agcgacgtcg ataagctctt catccagttg gttcaaacct acaatcagct ctttgaggaa 3720

[0073] aacccaatta atgctagtgg agtggatgca aaagcgatac tgtcggccag actctccaag 3780

[0074] agcagaaggt tggagaacct gatcgctcaa cttcctggag aaaagaaaaa cggtcttttt 3840

[0075] gggaatttga ttgccttgtc tctgggcctc acaccaaact tcaagtcaaa ttttgacctc 3900

[0076] gctgaggatg ccaaacttca gttgtctaag gatacctatg atgacgatct tgacaatttg 3960

[0077] ctggcacaaa ttggcgacca gtacgcggat ctgttcctcg cagcgaagaa tctgagtgat 4020

[0078] gctattctcc tttcggacat actcagggtt aacactgaga tcacaaaagc acctttgagt 4080

[0079] gcgtcgatga ttaagcgcta tgatgaacat caccaagacc tcactttgct gaaggccctt 4140

[0080] gtgcggcagc aattgccaga gaagtacaaa gaaatcttct ttgaccaatc taagaacgga 4200

[0081] tacgctggct atattgatgg aggagcttct caggaggaat tctataagtt tatcaaacct 4260

[0082] atacttgaga agatggatgg tacagaggaa ctccttgtta aattgaacag agaagatttg 4320

[0083] ctgcgcaagc aacggacctt tgacaacgga tcaattccgc atcagataca cctcggcgag 4380

[0084] cttcatgcca tccttcgccg gcaggaagat ttctacccct ttttgaagga caaccgcgag 4440

[0085] aagatagaaa aaatccttac gttccggatt ccttactatg tgggtccatt ggcaaggggg 4500

[0086] aattcccgct ttgcgtggat gactcggaaa agcgaggaaa ctatcacacc gtggaacttc 4560

[0087] gaggaagttg tggacaaggg agcttctgcc caatcattca ttgagaggat gactaacttc 4620

[0088] gataagaacc tgccgaacga gaaagttctc cccaagcact ccctccttta cgagtatttc 4680

[0089] accgtgtata acgaacttac gaaggttaaa tacgtgactg agggtatgag gaagccagca 4740

[0090] ttcttgagcg gggaacaaaa gaaagcgatt gttgatttgc tgtttaaaac taatcgcaag 4800

[0091] gtgacagtca agcagctcaa agaggattat ttcaagaaaa ttgaatgttt cgactctgtg 4860

[0092] gagatatcag gagtcgaaga taggtttaac gcttcccttg gcacatacca tgacctcctt 4920

[0093] aagatcatta aggacaaaga tttcctggat aacgaggaaa atgaggacat cctcgaagat 4980

[0094] attgttctta ccttgacgct gtttgaggat cgcgaaatga tcgaggaacg gcttaagacg 5040

[0095] tatgctcact tgttcgacga taaggttatg aagcagctca agcgtagaag gtacactgga 5100

[0096] tggggccgtc tgtctagaaa gctcatcaac ggaatacgtg ataaacaaag tggcaagaca 5160

[0097] attttggatt ttctgaagtc ggacggattc gccaacagaa attttatgca gctgattcat 5220

[0098] gacgatagtc tcaccttcaa agaggacata cagaaggctc aagtgagtgg tcaaggggat 5280

[0099] tcgctgcatg aacacatcgc aaacctcgcg ggttcaccgg ccataaagaa aggaatcctt 5340

[0100] caaactgtta aggtcgttga tgagttggtt aaagtgatgg gtaggcacaa gcccgaaaac 5400

[0101] atagtgatcg agatggctcg cgaaaatcag actacacaaa aagggcagaa gaactctcgc 5460

[0102] gagcggatga aaaggattga ggaaggaatc aaggaactgg gctcacagat tctcaaagag 5520

[0103] catccagtcg aaaacacaca gctgcaaaat gagaagctct atctttacta tctccaaaat 5580

[0104] ggccgggaca tgtatgttga tcaggagctt gacatcaacc gtttgtccga ctatgatgtg 5640

[0105] gaccacattg tcccgcaatc tttccttaag gacgattcaa tcgataataa ggtgttgacc 5700

[0106] cggagcgata aaaaccgtgg aaagtctgac aatgtccctt cagaggaagt ggttaagaag 5760

[0107] atgaagaact actggagaca attgctgaat gcaaaactga tcacacagag aaagttcgac 5820

[0108] aacctcacca aagcagagag aggtgggctc agtgaacttg ataaagcggg cttcattaag 5880

[0109] cgtcagctcg ttgagactag acagatcacg aagcatgtcg cgcagatttt ggattcgcgg 5940

[0110] atgaacacga agtacgacga gaatgataaa ctgatacgtg aagtcaaggt tatcactctt 6000

[0111] aagtccaaat tggtgagcga tttcagaaag gacttccaat tctataaggt cagggagatc 6060

[0112] aacaattatc atcacgctca cgatgcctac cttaatgctg ttgtggggac cgcccttatt 6120

[0113] aagaaatacc ctaaattgga gtctgaattc gtttacgggg attataaggt ctacgacgtt 6180

[0114] aggaaaatga tagctaagag tgagcaggag atcggtaaag caactgcgaa gtatttcttt 6240

[0115] tactcgaaca tcatgaattt ctttaagacc gagataacgc tggcaaatgg cgaaattaga 6300

[0116] aagaggcctc tcatagagac taacggtgag acaggggaaa tcgtctggga taagggtagg 6360

[0117] gactttgcga cagtgcgcaa ggtcctctct atgccgcaag ttaatattgt gaagaaaacc 6420

[0118] gaggtgcaga cgggaggctt ctccaaggaa agcatacttc ccaaacggaa ctctgataag 6480

[0119] ttgatcgctc gtaagaaaga ttgggaccct aagaaatatg gtgggttcga ttccccaact 6540

[0120] gttgcttaca gcgtgctggt cgttgccaag gtcgagaagg gtaaatccaa gaaactcaaa 6600

[0121] agcgttaagg aactccttgg gattactatc atggagagat cttcattcga aaagaatcct 6660

[0122] atcgactttc ttgaggccaa aggatataag gaagttaaga aagatctgat aatcaaactc 6720

[0123] ccaaagtact cattgtttga gctggaaaac ggcaggaagc gcatgcttgc ttccgccgga 6780

[0124] gagttgcaga aagggaacga gttggctctg ccttctaagt atgttaactt cctctatctt 6840

[0125] gcctctcatt acgagaagct caaaggctca ccagaggaca acgaacagaa acaacttttt 6900

[0126] gtcgagcaac ataagcacta tttggatgag attatagaac agatcagtga attctcgaaa 6960
[0127] agggttatcc ttgcagatgc gaatcttgac aaggtgttgt ctgcatacaa caaacataga 7020
[0128] gataagccga tcagggagca agcggaaaat atcattcacc tcttcactct tacaaacttg 7080
[0129] ggtgctcccg ctgccttcaa gtattttgat accacgattg accggaaacg ttacacctca 7140
[0130] acgaaggagg tgctggatgc caccctcatc caccaatcta ttaccggact ctacgagact 7200
[0131] agaatcgatc tctcacagct cggcggggat aaaagaccag cagcgacgaa aaaggcagga 7260
[0132] caggctaaga agaagaaaga gctcggagga ggaggcacgg gaggaggagg ctccgccgag 7320
[0133] tatgtgcgcg cgctcttcga cttcaacggc aatgacgagg aggatctccc tttcaagaag 7380
[0134] ggcgacatcc tccgcatccg cgataagccg gaggagcagt ggtggaacgc agaggactcc 7440
[0135] gagggcaagc ggggcatgat cctggtgcca tacgtcgaga agtacagcgg cgattacaag 7500
[0136] gaccacgatg gcgactacaa ggatcatgac atcgattaca aggacgatga cgataagtcc 7560
[0137] ggcgtcgaca tgacggacgc ggagtatgtg cgcatccacg agaagctcga tatctacacc 7620
[0138] ttcaagaagc agttcttcaa caataagaag tcggtgtccc atcggtgcta cgtcctcttc 7680
[0139] gagctgaagc gcaggggaga gcgccgcgcc tgcttctggg gctacgcggt gaataagccg 7740
[0140] cagtcaggca cagagcgcgg catccacgcc gagatcttct cgatccggaa ggtcgaggag 7800
[0141] tacctccgcg acaacccagg ccagttcacg atcaattggt actccagctg gtccccttgc 7860
[0142] gcagattgcg cagagaagat cctcgagtgg tacaaccagg agctgagggg caatggccat 7920
[0143] accctcaaga tctgggcctg caagctgtac tacgagaaga acgcgaggaa tcagatcggc 7980
[0144] ctctggaacc tgcgggataa tggcgtgggc ctcaacgtga tggtgtccga gcactaccag 8040
[0145] tgctgccgca agatcttcat ccagtcctcc cacaatcagc tgaacgagaa taggtggctc 8100
[0146] gaaaagaccc tgaagcgcgc cgagaagtgg aggagcgagc tgtctatcat gatccaggtc 8160
[0147] aagatcctgc acaccacaaa gtcaccggcg gtgggcggcg gcggcagcga attctccggc 8220
[0148] ggcagcacga acctcagcga catcatcgag aaggagacag gcaagcagct cgtgatccag 8280
[0149] gagtctatcc tcatgctgcc tgaggaggtg gaggaggtca tcggcaacaa gccggagtcc 8340
[0150] gatatcctcg tgcacaccgc ctacgacgag tcgacagatg agaatgtcat gctcctgacc 8400
[0151] tccgacgcac cagagtacaa gccatgggcg ctcgtgatcc aggattccaa cggcgagaat 8460
[0152] aagatcaaga tgctgtctgg cggctccccg aagaagaagc gcaaggtcta gactagtctg 8520
[0153] aaatcaccag tctctctcta caaatctatc tctctctata ataatgtgtg agtagttccc 8580
[0154] agataaggga attagggttc ttatagggtt tcgctcatgt gttgagcata taagaaaccc 8640
[0155] ttagtatgta tttgtatttg taaaatactt ctatcaataa aatttctaat tcctaaaacc 8700
[0156] aaaatccagt ggggcgcccg acctgtactc gcgaaggtta acttacagag agtgtccggg 8760
[0157] cgcgcctggt ggatcgtccg cctaggctgc agtgcagcgt gacccggtcg tgcccctctc 8820
[0158] tagagataat gagcattgca tgtctaagtt ataaaaaatt accacatatt ttttttgtca 8880
[0159] cacttgtttg aagtgcagtt tatctatctt tatacatata tttaaacttt actctacgaa 8940
[0160] taatataatc tatagtacta caataatatc agtgttttag agaatcatat aaatgaacag 9000
[0161] ttagacatgg tctaaaggac aattgagtat tttgacaaca ggactctaca gttttatctt 9060
[0162] tttagtgtgc atgtgttctc ctttttttttt gcaaatagct tcacctatat aatacttcat 9120
[0163] ccattttatt agtacatcca tttagggttt agggttaatg gtttttatag actaattttt 9180
[0164] ttagtacatc tattttattc tattttagcc tctaaattaa gaaaactaaa actctatttt 9240
[0165] agtttttta tttaataatt tagatataaa atagaataaa ataaagtgac taaaaattaa 9300
[0166] acaaataccc tttaagaaat taaaaaaact aaggaaacat ttttcttgtt tcgagtagat 9360
[0167] aatgccagcc tgttaaacgc cgtcgacgag tctaacggac accaaccagc gaaccagcag 9420

[0168] cgtcgcgtcg ggccaagcga agcagacggc acggcatctc tgtcgctgcc tctggacccc 9480
[0169] tctcgagagt tccgctccac cgttggactt gctccgctgt cggcatccag aaattgcgtg 9540
[0170] gcggagcggc agacgtgagc cggcacggca ggcggcctcc tcctcctctc acggcaccgg 9600
[0171] cagctacggg ggattccttt cccaccgctc cttcgctttc ccttcctcgc ccgccgtaat 9660
[0172] aaatagacac cccctccaca ccctctttcc ccaacctcgt gttgttcgga gcgcacacac 9720
[0173] acacaaccag atctccccca aatccacccg tcggcacctc cgcttcaagg tacgccgctc 9780
[0174] gtcctccccc ccccccccctc tctaccttct ctagatcggc gttccggtcc atggttaggg 9840
[0175] cccggtagtt ctacttctgt tcatgtttgt gttagatccg tgtttgtgtt agatccgtgc 9900
[0176] tgctagcgtt cgtacacgga tgcgacctgt acgtcagaca cgttctgatt gctaacttgc 9960
[0177] cagtgtttct ctttggggaa tcctgggatg gctctagccg ttccgcagac gggatcgatt 10020
[0178] tcatgatttt ttttgtttcg ttgcataggg tttggtttgc ccttttcctt tatttcaata 10080
[0179] tatgccgtgc acttgtttgt cgggtcatct tttcatgctt ttttttgtct tggttgtgat 10140
[0180] gatgtggtct ggttgggcgg tcgttctaga tcggagtaga attctgtttc aaactacctg 10200
[0181] gtggatttat taattttgga tctgtatgtg tgtgccatac atattcatag ttacgaattg 10260
[0182] aagatgatgg atggaaatat cgatctagga taggtataca tgttgatgcg ggttttactg 10320
[0183] atgcatatac agagatgctt tttgttcgct tggttgtgat gatgtggtgt ggttgggcgg 10380
[0184] tcgttcattc gttctagatc ggagtagaat actgtttcaa actacctggt gtatttatta 10440
[0185] attttggaac tgtatgtgtg tgtcatacat cttcatagtt acgagtttaa gatggatgga 10500
[0186] aatatcgatc taggataggt atacatgttg atgtgggttt tactgatgca tatacatgat 10560
[0187] ggcatatgca gcatctattc atatgctcta accttgagta cctatctatt ataataaaca 10620
[0188] agtatgtttt ataattattt tgatcttgat atacttggat gatggcatat gcagcagcta 10680
[0189] tatgtggatt tttttagccc tgccttcata cgctatttat ttgcttggta ctgtttcttt 10740
[0190] tgtcgatgct caccctgttg tttggtgtta cttctgcagg agctcatgaa aaagcctgaa 10800
[0191] ctcaccgcga cgtctgtcga gaagtttctg atcgaaaagt tcgacagcgt ctccgacctg 10860
[0192] atgcagctct cggagggcga agaatctcgt gctttcagct tcgatgtagg agggcgtgga 10920
[0193] tatgtcctgc gggtaaatag ctgcgccgat ggtttctaca aagatcgtta tgtttatcgg 10980
[0194] cactttgcat cggccgcgct cccgattccg gaagtgcttg acattgggga gtttagcgag 11040
[0195] agcctgacct attgcatctc ccgccgttca cagggtgtca cgttgcaaga cctgcctgaa 11100
[0196] accgaactgc ccgctgttct acaaccggtc gcggaggcta tggatgcgat cgctgcggcc 11160
[0197] gatcttagcc agacgagcgg gttcggccca ttcggaccgc aaggaatcgg tcaatacact 11220
[0198] acatggcgtg atttcatatg cgcgattgct gatccccatg tgtatcactg gcaaactgtg 11280
[0199] atggacgaca ccgtcagtgc gtccgtcgcg caggctctcg atgagctgat gctttgggcc 11340
[0200] gaggactgcc ccgaagtccg gcacctcgtg cacgcggatt tcggctccaa caatgtcctg 11400
[0201] acggacaatg gccgcataac agcggtcatt gactggagcg aggcgatgtt cggggattcc 11460
[0202] caatacgagg tcgccaacat cttcttctgg aggccgtggt tggcttgtat ggagcagcag 11520
[0203] acgcgctact tcgagcggag gcatccggag cttgcaggat cgccacgact ccgggcgtat 11580
[0204] atgctccgca ttggtcttga ccaactctat cagagcttgg ttgacggcaa tttcgatgat 11640
[0205] gcagcttggg cgcagggtcg atgcgacgca atcgtccgat ccggagccgg gactgtcggg 11700
[0206] cgtacacaaa tcgcccgcag aagcgcggcc gtctggaccg atggctgtgt agaagtactc 11760
[0207] gccgatagtg gaaaccgacg ccccagcact cgtccgaggg caaagaaata ggatcgttca 11820
[0208] aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 11880
[0209] atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 11940

[0210] tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 12000

[0211] aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 12060

[0212] gatctgtagc cctgcaggac gcgtttaatt aagtgcacgc ggccgcctac ttagtcaaga 12120

[0213] gcctcgcacg cgactgtcac gcggccagga tcgcctcgtg agcctcgcaa tctgtaccta 12180

[0214] gtgtttaaac tatcagtgtt tgacaggata tattggcggg taaacctaag agaaaagagc 12240

[0215] gtttattaga ataacggata tttaaaaggg cgtgaaaagg tttatccgtt cgtccatttg 12300

[0216] tatgtgcatg ccaaccacag ggttcccctc gggatcaaag tactttgatc caacccctcc 12360

[0217] gctgctatag tgcagtcggc ttctgacgtt cagtgcagcc gtcttctgaa aacgacatgt 12420

[0218] cgcacaagtc ctaagttacg cgacaggctg ccgccctgcc cttttcctgg cgttttcttg 12480

[0219] tcgcgtgttt tagtcgcata aagtagaata cttgcgacta gaaccggaga cattacgcca 12540

[0220] tgaacaagag cgccgccgct ggcctgctgg gctatgcccg cgtcagcacc gacgaccagg 12600

[0221] acttgaccaa ccaacgggcc gaactgcacg cggccggctg caccaagctg ttttccgaga 12660

[0222] agatcaccgg caccaggcgc gaccgcccgg agctggccag gatgcttgac cacctacgcc 12720

[0223] ctggcgacgt tgtgacagtg accaggctag accgcctggc ccgcagcacc cgcgacctac 12780

[0224] tggacattgc cgagcgcatc caggaggccg gcgcgggcct gcgtagcctg gcagagccgt 12840

[0225] gggccgacac caccacgccg gccggccgca tggtgttgac cgtgttcgcc ggcattgccg 12900

[0226] agttcgagcg ttccctaatc atcgaccgca cccggagcgg gcgcgaggcc gccaaggccc 12960

[0227] gaggcgtgaa gtttggcccc cgccctaccc tcaccccggc acagatcgcg cacgcccgcg 13020

[0228] agctgatcga ccaggaaggc cgcaccgtga aagaggcggc tgcactgctt ggcgtgcatc 13080

[0229] gctcgaccct gtaccgcgca cttgagcgca gcgaggaagt gacgcccacc gaggccaggc 13140

[0230] ggcgcggtgc cttccgtgag gacgcattga ccgaggccga cgccctggcg gccgccgaga 13200

[0231] atgaacgcca agaggaacaa gcatgaaacc gcaccaggac ggccaggacg aaccgttttt 13260

[0232] cattaccgaa gagatcgagg cggagatgat cgcggccggg tacgtgttcg agccgcccgc 13320

[0233] gcacgtctca accgtgcggc tgcatgaaat cctggccggt ttgtctgatg ccaagctggc 13380

[0234] ggcctggccg gccagcttgg ccgctgaaga aaccgagcgc cgccgtctaa aaaggtgatg 13440

[0235] tgtatttgag taaaacagct tgcgtcatgc ggtcgctgcg tatatgatgc gatgagtaaa 13500

[0236] taaacaaata cgcaagggga acgcatgaag gttatcgctg tacttaacca gaaaggcggg 13560

[0237] tcaggcaaga cgaccatcgc aacccatcta gcccgcgccc tgcaactcgc cggggccgat 13620

[0238] gttctgttag tcgattccga tccccagggc agtgcccgcg attgggcggc cgtgcgggaa 13680

[0239] gatcaaccgc taaccgttgt cggcatcgac cgcccgacga ttgaccgcga cgtgaaggcc 13740

[0240] atcggccggc gcgacttcgt agtgatcgac ggagcgcccc aggcggcgga cttggctgtg 13800

[0241] tccgcgatca aggcagccga cttcgtgctg attccggtgc agccaagccc ttacgacata 13860

[0242] tgggccaccg ccgacctggt ggagctggtt aagcagcgca ttgaggtcac ggatggaagg 13920

[0243] ctacaagcgg cctttgtcgt gtcgcgggcg atcaaaggca cgcgcatcgg cggtgaggtt 13980

[0244] gccgaggcgc tggccgggta cgagctgccc attcttgagt cccgtatcac gcagcgcgtg 14040

[0245] agctacccag gcactgccgc cgccggcaca accgttcttg aatcagaacc cgagggcgac 14100

[0246] gctgcccgcg aggtccaggc gctggccgct gaaattaaat caaaactcat ttgagttaat 14160

[0247] gaggtaaaga gaaaatgagc aaaagcacaa acacgctaag tgccggccgt ccgagcgcac 14220

[0248] gcagcagcaa ggctgcaacg ttggccagcc tggcagacac gccagccatg aagcgggtca 14280

[0249] actttcagtt gccggcggag gatcacacca agctgaagat gtacgcggta cgccaaggca 14340

[0250] agaccattac cgagctgcta tctgaataca tcgcgcagct accagagtaa atgagcaaat 14400

[0251] gaataaatga gtagatgaat tttagcggct aaaggaggcg gcatggaaaa tcaagaacaa 14460

[0252] ccaggcaccg acgccgtgga atgccccatg tgtggaggaa cgggcggttg gccaggcgta 14520
[0253] agcggctggg ttgtctgccg gccctgcaat ggcactggaa cccccaagcc cgaggaatcg 14580
[0254] gcgtgacggt cgcaaaccat ccggcccggt acaaatcggc gcggcgctgg gtgatgacct 14640
[0255] ggtggagaag ttgaaggccg cgcaggccgc ccagcggcaa cgcatcgagg cagaagcacg 14700
[0256] ccccggtgaa tcgtggcaag cggccgctga tcgaatccgc aaagaatccc ggcaaccgcc 14760
[0257] ggcagccggt gcgccgtcga ttaggaagcc gcccaagggc gacgagcaac cagatttttt 14820
[0258] cgttccgatg ctctatgacg tgggcacccg cgatagtcgc agcatcatgg acgtggccgt 14880
[0259] tttccgtctg tcgaagcgtg accgacgagc tggcgaggtg atccgctacg agcttccaga 14940
[0260] cgggcacgta gaggtttccg cagggccggc cggcatggcc agtgtgtggg attacgacct 15000
[0261] ggtactgatg gcggtttccc atctaaccga atccatgaac cgataccggg aagggaaggg 15060
[0262] agacaagccc ggccgcgtgt tccgtccaca cgttgcggac gtactcaagt tctgccggcg 15120
[0263] agccgatggc ggaaagcaga aagacgacct ggtagaaacc tgcattcggt taaacaccac 15180
[0264] gcacgttgcc atgcagcgta cgaagaaggc caagaacggc cgcctggtga cggtatccga 15240
[0265] gggtgaagcc ttgattagcc gctacaagat cgtaaagagc gaaaccgggc ggccggagta 15300
[0266] catcgagatc gagctagctg attggatgta ccgcgagatc acagaaggca agaacccgga 15360
[0267] cgtgctgacg gttcaccccg attacttttt gatcgatccc ggcatcggcc gttttctcta 15420
[0268] ccgcctggca cgccgcgccg caggcaaggc agaagccaga tggttgttca agacgatcta 15480
[0269] cgaacgcagt ggcagcgccg gagagttcaa gaagttctgt ttcaccgtgc gcaagctgat 15540
[0270] cgggtcaaat gacctgccgg agtacgattt gaaggaggag gcggggcagg ctggcccgat 15600
[0271] cctagtcatg cgctaccgca acctgatcga gggcgaagca tccgccggtt cctaatgtac 15660
[0272] ggagcagatg ctagggcaaa ttgccctagc aggggaaaaa ggtcgaaaag gtctctttcc 15720
[0273] tgtggatagc acgtacattg ggaacccaaa gccgtacatt gggaaccgga acccgtacat 15780
[0274] tgggaaccca aagccgtaca ttgggaaccg gtcacacatg taagtgactg atataaaaga 15840
[0275] gaaaaaaggc gatttttccg cctaaaactc tttaaaactt attaaaactc ttaaaacccg 15900
[0276] cctggcctgt gcataactgt ctggccagcg cacagccgaa gagctgcaaa aagcgcctac 15960
[0277] ccttcggtcg ctgcgctccc tacgccccgc cgcttcgcgt cggcctatcg cggccgctgg 16020
[0278] ccgctcaaaa atggctggcc tacggccagg caatctacca gggcgcggac aagccgcgcc 16080
[0279] gtcgccactc gaccgccggc gcccacatca aggcaccctg cctcgcgcgt ttcggtgatg 16140
[0280] acggtgaaaa cctctgacac atgcagctcc cggagacggt cacagcttgt ctgtaagcgg 16200
[0281] atgccgggag cagacaagcc cgtcagggcg cgtcagcggg tgttggcggg tgtcggggcg 16260
[0282] cagccatgac ccagtcacgt agcgatagcg gagtgtatac tggcttaact atgcggcatc 16320
[0283] agagcagatt gtactgagag tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag 16380
[0284] gagaaaatac cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 16440
[0285] cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 16500
[0286] atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 16560
[0287] taaaaaggcc gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa 16620
[0288] aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 16680
[0289] tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 16740
[0290] gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 16800
[0291] cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 16860
[0292] cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 16920
[0293] atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 16980

[0294] tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat 17040

[0295] ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 17100

[0296] acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 17160

[0297] aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 17220

[0298] aaactcacgt taagggattt tggtcatgca ttctaggtac taaaacaatt catccagtaa 17280

[0299] aatataatat tttattttct cccaatcagg cttgatcccc agtaagtcaa aaaatagctc 17340

[0300] gacatactgt tcttccccga tatcctccct gatcgaccgg acgcagaagg caatgtcata 17400

[0301] ccacttgtcc gccctgccgc ttctcccaag atcaataaag ccacttactt tgccatcttt 17460

[0302] cacaaagatg ttgctgtctc ccaggtcgcc gtgggaaaag acaagttcct cttcgggctt 17520

[0303] ttccgtcttt aaaaaatcat acagctcgcg cggatcttta aatggagtgt cttcttccca 17580

[0304] gttttcgcaa tccacatcgg ccagatcgtt attcagtaag taatccaatt cggctaagcg 17640

[0305] gctgtctaag ctattcgtat agggacaatc cgatatgtcg atggagtgaa agagcctgat 17700

[0306] gcactccgca tacagctcga taatcttttc agggctttgt tcatcttcat actcttccga 17760

[0307] gcaaaggacg ccatcggcct cactcatgag cagattgctc cagccatcat gccgttcaaa 17820

[0308] gtgcaggacc tttggaacag gcagctttcc ttccagccat agcatcatgt ccttttcccg 17880

[0309] ttccacatca taggtggtcc ctttataccg gctgtccgtc atttttaaat ataggttttc 17940

[0310] attttctccc accagcttat ataccttagc aggagacatt ccttccgtat cttttacgca 18000

[0311] gcggtatttt tcgatcagtt ttttcaattc cggtgatatt ctcatttag ccatttatta 18060

[0312] tttccttcct cttttctaca gtatttaaag ataccccaag aagctaatta taacaagacg 18120

[0313] aactccaatt cactgttcct tgcattctaa aaccttaaat accagaaaac agctttttca 18180

[0314] aagttgtttt caaagttggc gtataacata gtatcgacgg agccgatttt gaaaccgcgg 18240

[0315] tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct ccgcgagatc 18300

[0316] atccgtgttt caaacccggc agcttagttg ccgttcttcc gaatagcatc ggtaacatga 18360

[0317] gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 18420

[0318] ctgtatcgag tggtgatttt gtgccgagct gccggtcggg gagctgttgg ctggct 18476

[0319] <210> 2

[0320] <211> 7423

[0321] <212> DNA

[0322] <213> 人工序列(Artificial Sequence)

[0323] <400> 2

[0324] acaaattcgg gtcaaggcgg aagccagcgc gccaccccac gtcagcaaat acggaggcgc 60

[0325] ggggttgacg gcgtcacccg gtcctaacgg cgaccaacaa accagccaga agaaattaca 120

[0326] gtaaaaaaaa agtaaattgc actttgatcc acctttttatt acctaagtct caatttggat 180

[0327] cacccttaaa cctatctttt caatttgggc cgggttgtgg tttggactac catgaacaac 240

[0328] ttttcgtcat gtctaacttc cctttcagca aacatatgaa ccatatatag aggagatcgg 300

[0329] ccgtatacta gagctgatgt gtttaaggtc gttgattgca cgagaaaaaa aaatccaaat 360

[0330] cgcaacaata gcaaatttat ctggttcaaa gtgaaaagat atgtttaaag gtagtccaaa 420

[0331] gtaaaactta tagataataa aatgtggtcc aaagcgtaat tcactcaaaa aaaatcaacg 480

[0332] agacgtgtac caaacggaga caaacggcat cttctcgaaa tttcccaacc gctcgctcgc 540

[0333] ccgcctcgtc ttcccggaaa ccgcggtggt ttcagcgtgg cggattctcc aagcagacgg 600

[0334] agacgtcacg gcacgggact cctcccacca cccaaccgcc ataaatacca gccccctcat 660

[0335] ctcctctcct cgcatcagct ccacccccga aaaatttctc cccaatctcg cgaggctctc 720

[0336] gtcgtcgaat cgaatcctct cgcgtcctca aggtacgctg cttctcctct cctcgcttcg 780

[0337] tttcgattcg atttcggacg ggtgaggttg ttttgttgct agatccgatt ggtggttagg 840

[0338] gttgtcgatg tgattatcgt gagatgttta ggggttgtag atctgatggt tgtgatttgg 900

[0339] gcacggttgg ttcgataggt ggaatcgtgg ttaggttttg ggattggatg ttggttctga 960

[0340] tgattggggg gaatttttac ggttagatga attgttggat gattcgattg gggaaatcgg 1020

[0341] tgtagatctg ttggggaatt gtggaactag tcatgcctga gtgattggtg cgatttgtag 1080

[0342] cgtgttccat cttgtaggcc ttgttgcgag catgttcaga tctactgttc cgctcttgat 1140

[0343] tgagttattg gtgccatggg ttggtgcaaa cacaggcttt aatatgttat atctgttttg 1200

[0344] tgtttgatgt agatctgtag ggtagttctt cttagacatg gttcaattat gtagcttgtg 1260

[0345] cgtttcgatt tgatttcata tgttcacaga ttagataatg atgaactctt ttaattaatt 1320

[0346] gtcaatggta aataggaagt cttgtcgcta tatctgtcat aatgatctca tgttactatc 1380

[0347] tgccagtaat ttatgctaag aactatatta gaatatcatg ttacaatctg tagtaatatc 1440

[0348] atgttacaat ctgtagttca tctatataat ctattgtggt aatttctttt tactatctgt 1500

[0349] gtgaagatta ttgccactag ttcattctac ttatttctga agttcaggat acgtgtgctg 1560

[0350] ttactaccta tctgaataca tgtgtgatgt gcctgttact atctttttga atacatgtat 1620

[0351] gttctgttgg aatatgtttg ctgtttgatc cgttgttgtg tccttaatct tgtgctagtt 1680

[0352] cttaccctat ctgtttggtg attatttctt gcagtacgta atggactaca aggaccacga 1740

[0353] cggggattac aaagaccacg acatagacta caaggatgac gatgacaaaa tggcaccgaa 1800

[0354] gaaaaaaagg aaggtcggaa tccatggcgt tccagctgcc gataagaaat attccatcgg 1860

[0355] actcgccatt ggcacgaata gcgtcggatg ggctgttatt actgatgagt acaaagttcc 1920

[0356] gtctaagaag ttcaaggtgc tgggcaacac agaccgccac agcataaaga aaaatctcat 1980

[0357] cggtgcactc cttttcgata gtggggagac tgcagaagcg acaagattga aaaggactgc 2040

[0358] gagaaggcgc tatacacggc gtaagaatag aatctgctac cttcaggaga ttttctctaa 2100

[0359] cgaaatggct aaggtcgatg acagtttctt tcatagactt gaggaatcgt tcttggttga 2160

[0360] ggaggataag aaacatgaga ggcacccgat atttggaaac atcgtggatg aggtcgcata 2220

[0361] tcatgaaaag taccccacaa tctaccacct gagaaagaaa ctcgttgatt ccaccgacaa 2280

[0362] agcggatttg agactcatct acctcgctct tgcccatatg ataaagttcc gcggacactt 2340

[0363] tctgatcgag ggcgacctca accctgataa tagcgacgtc gataagctct tcatccagtt 2400

[0364] ggttcaaacc tacaatcagc tctttgagga aaacccaatt aatgctagtg gagtggatgc 2460

[0365] aaaagcgata ctgtcggcca gactctccaa gagcagaagg ttggagaacc tgatcgctca 2520

[0366] acttcctgga gaaaagaaaa acggtctttt tgggaatttg attgccttgt ctctgggcct 2580

[0367] cacaccaaac ttcaagtcaa attttgacct cgctgaggat gccaaacttc agttgtctaa 2640

[0368] ggatacctat gatgacgatc ttgacaattt gctggcacaa attggcgacc agtacgcgga 2700

[0369] tctgttcctc gcagcgaaga atctgagtga tgctattctc ctttcggaca tactcagggt 2760

[0370] taacactgag atcacaaaag caccttttgag tgcgtcgatg attaagcgct atgatgaaca 2820

[0371] tcaccaagac ctcactttgc tgaaggccct tgtgcggcag caattgccag agaagtacaa 2880

[0372] agaaatcttc tttgaccaat ctaagaacgg atacgctggc tatattgatg gaggagcttc 2940

[0373] tcaggaggaa ttctataagt ttatcaaacc tatacttgag aagatggatg gtacagagga 3000

[0374] actccttgtt aaattgaaca gagaagattt gctgcgcaag caacggacct ttgacaacgg 3060

[0375] atcaattccg catcagatac acctcggcga gcttcatgcc atccttcgcc ggcaggaaga 3120

[0376] tttctacccc tttttgaagg acaaccgcga gaagatagaa aaaatcctta cgttccggat 3180

[0377] tccttactat gtgggtccat tggcaagggg gaattcccgc tttgcgtgga tgactcggaa 3240

[0378] aagcgaggaa actatcacac cgtggaactt cgaggaagtt gtggacaagg gagcttctgc 3300

[0379] ccaatcattc attgagagga tgactaactt cgataagaac ctgccgaacg agaaagttct 3360

[0380] ccccaagcac tccctccttt acgagtattt caccgtgtat aacgaactta cgaaggttaa 3420

[0381] atacgtgact gagggtatga ggaagccagc attcttgagc ggggaacaaa agaaagcgat 3480

[0382] tgttgatttg ctgtttaaaa ctaatcgcaa ggtgacagtc aagcagctca aagaggatta 3540

[0383] tttcaagaaa attgaatgtt tcgactctgt ggagatatca ggagtcgaag ataggtttaa 3600

[0384] cgcttccctt ggcacatacc atgacctcct taagatcatt aaggacaaag atttcctgga 3660

[0385] taacgaggaa aatgaggaca tcctcgaaga tattgttctt accttgacgc tgtttgagga 3720

[0386] tcgcgaaatg atcgaggaac ggcttaagac gtatgctcac ttgttcgacg ataaggttat 3780

[0387] gaagcagctc aagcgtagaa ggtacactgg atggggccgt ctgtctagaa agctcatcaa 3840

[0388] cggaatacgt gataaacaaa gtggcaagac aattttggat tttctgaagt cggacggatt 3900

[0389] cgccaacaga gcttttgcgg cactgattgc tgacgatagt ctcaccttca aagaggacat 3960

[0390] acagaaggct caagtgagtg gtcaagggga ttcgctgcat gaacacatcg caaacctcgc 4020

[0391] gggttcaccg gccataaaga aaggaatcct tcaaactgtt aaggtcgttg atgagttggt 4080

[0392] taaagtgatg ggtaggcaca agcccgaaaa catagtgatc gagatggctc gcgaaaatca 4140

[0393] gactacacaa aaagggcaga agaactctcg cgagcggatg aaaaggattg aggaaggaat 4200

[0394] caaggaactg ggctcacaga ttctcaaaga gcatccagtc gaaaacacac agctgcaaaa 4260

[0395] tgagaagctc tatctttact atctccaaaa tggccgggac atgtatgttg atcaggagct 4320

[0396] tgacatcaac cgtttgtccg actatgatgt ggaccacatt gtcccgcaat ctttccttaa 4380

[0397] ggacgattca atcgataata aggtgttgac ccggagcgat aaaaaccgtg gaaagtctga 4440

[0398] caatgtccct tcagaggaag tggttaagaa gatgaagaac tactggagac aattgctgaa 4500

[0399] tgcaaaactg atcacacaga gaaagttcga caacctcacc aaagcagaga gaggtgggct 4560

[0400] cagtgaactt gataaagcgg gcttcattaa gcgtcagctc gttgagacta gacagatcac 4620

[0401] gaagcatgtc gcgcagattt tggattcgcg gatgaacacg aagtacgacg agaatgataa 4680

[0402] actgatacgt gaagtcaagg ttatcactct taagtccaaa ttggtgagcg atttcagaaa 4740

[0403] ggacttccaa ttctataagg tcagggagat caacaattat catcacgctc acgatgccta 4800

[0404] ccttaatgct gttgtgggga ccgcccttat taagaaatac cctaaattgg agtctgaatt 4860

[0405] cgtttacggg gattataagg tctacgacgt taggaaaatg atagctaaga gtgagcagga 4920

[0406] gatcggtaaa gcaactgcga agtatttctt ttactcgaac atcatgaatt tctttaagac 4980

[0407] cgagataacg ctggcaaatg gcgaaattag aaagaggcct ctcatagaga ctaacggtga 5040

[0408] gacaggggaa atcgtctggg ataagggtag ggactttgcg acagtgcgca aggtcctctc 5100

[0409] tatgccgcaa gttaatattg tgaagaaaac cgaggtgcag acgggaggct tctccaagga 5160

[0410] aagcatactt cccaaacgga actctgataa gttgatcgct cgtaagaaag attgggaccc 5220

[0411] taagaaatat ggtgggttcg attccccaac tgttgcttac agcgtgctgg tcgttgccaa 5280

[0412] ggtcgagaag ggtaaatcca agaaactcaa aagcgttaag gaactccttg ggattactat 5340

[0413] catggagaga tcttcattcg aaaagaatcc tatcgacttt cttgaggcca aaggatataa 5400

[0414] ggaagttaag aaagatctga taatcaaact cccaaagtac tcattgtttg agctggaaaa 5460

[0415] cggcaggaag cgcatgcttg cttccgccgg agagttgcag aaagggaacg agttggctct 5520

[0416] gccttctaag tatgttaact tcctctatct tgcctctcat tacgagaagc tcaaaggctc 5580

[0417] accagaggac aacgaacaga aacaactttt tgtcgagcaa cataagcact atttggatga 5640

[0418] gattatagaa cagatcagtg aattctcgaa aagggttatc cttgcagatg cgaatcttga 5700

[0419] caaggtgttg tctgcataca acaaacatag agataagccg atcagggagc aagcggaaaa 5760

[0420] tatcattcac ctcttcactc ttacaaactt gggtgctccc gctgccttca agtattttga 5820

[0421] taccacgatt gaccggaaac gttacacctc aacgaaggag gtgctggatg ccaccctcat 5880

[0422] ccaccaatct attaccggac tctacgagac tagaatcgat ctctcacagc tcggcgggga 5940

[0423] taaaagacca gcagcgacga aaaaggcagg acaggctaag aagaagaaag agctcggagg 6000

[0424] aggaggcacg ggaggaggag gctccgccga gtatgtgcgc gcgctcttcg acttcaacgg 6060

[0425] caatgacgag gaggatctcc ctttcaagaa gggcgacatc ctccgcatcc gcgataagcc 6120

[0426] ggaggagcag tggtggaacg cagaggactc cgagggcaag cggggcatga tcctggtgcc 6180

[0427] atacgtcgag aagtacagcg gcgattacaa ggaccacgat ggcgactaca aggatcatga 6240

[0428] catcgattac aaggacgatg acgataagtc cggcgtcgac atgacggacg cggagtatgt 6300

[0429] gcgcatccac gagaagctcg atatctacac cttcaagaag cagttcttca acaataagaa 6360

[0430] gtcggtgtcc catcggtgct acgtcctctt cgagctgaag cgcaggggag agcgccgcgc 6420

[0431] ctgcttctgg ggctacgcgg tgaataagcc gcagtcaggc acagagcgcg gcatccacgc 6480

[0432] cgagatcttc tcgatccgga aggtcgagga gtacctccgc gacaacccag gccagttcac 6540

[0433] gatcaattgg tactccagct ggtccccttg cgcagattgc gcagagaaga tcctcgagtg 6600

[0434] gtacaaccag gagctgaggg gcaatggcca taccctcaag atctgggcct gcaagctgta 6660

[0435] ctacgagaag aacgcgagga atcagatcgg cctctggaac ctgcgggata atggcgtggg 6720

[0436] cctcaacgtg atggtgtccg agcactacca gtgctgccgc aagatcttca tccagtcctc 6780

[0437] ccacaatcag ctgaacgaga ataggtggct cgaaaagacc ctgaagcgcg ccgagaagtg 6840

[0438] gaggagcgag ctgtctatca tgatccaggt caagatcctg cacaccacaa agtcaccggc 6900

[0439] ggtgggcggc ggcggcagcg aattctccgg cggcagcacg aacctcagcg acatcatcga 6960

[0440] gaaggagaca ggcaagcagc tcgtgatcca ggagtctatc ctcatgctgc ctgaggaggt 7020

[0441] ggaggaggtc atcggcaaca agccggagtc cgatatcctc gtgcacaccg cctacgacga 7080

[0442] gtcgacagat gagaatgtca tgctcctgac ctccgacgca ccagagtaca agccatgggc 7140

[0443] gctcgtgatc caggattcca acggcgagaa taagatcaag atgctgtctg gcggctcccc 7200

[0444] gaagaagaag cgcaaggtct agactagtct gaaatcacca gtctctctct acaaatctat 7260

[0445] ctctctctat aataatgtgt gagtagttcc cagataaggg aattagggtt cttatagggt 7320

[0446] ttcgctcatg tgttgagcat ataagaaacc cttagtatgt atttgtattt gtaaaatact 7380

[0447] tctatcaata aaatttctaa ttcctaaaac caaaatccag tgg 7423

[0448] <210> 3

[0449] <211> 20

[0450] <212> DNA

[0451] <213> 人工序列(Artificial Sequence)

[0452] <400> 3

[0453] tagcgacggc gagcaagtgg 20

[0454] <210> 4

[0455] <211> 20

[0456] <212> DNA

[0457] <213> 人工序列(Artificial Sequence)

[0458] <400> 4

[0459] cgatgacggc gagcaagtgg 20

[0460] <210> 5

[0461] <211> 20

[0462] <212> DNA

[0463] <213> 人工序列(Artificial Sequence)

[0464] <400> 5

[0465] cggcagcggc gagcaagtgg 20

[0466] <210> 6

[0467] <211> 20

[0468] <212> DNA

[0469] <213> 人工序列(Artificial Sequence)

[0470] <400> 6

[0471] cggcgatagc gagcaagtgg 20

[0472] <210> 7

[0473] <211> 20

[0474] <212> DNA

[0475] <213> 人工序列(Artificial Sequence)

[0476] <400> 7

[0477] cggcgacgat gagcaagtgg 20

[0478] <210> 8

[0479] <211> 20

[0480] <212> DNA

[0481] <213> 人工序列(Artificial Sequence)

[0482] <400> 8

[0483] cggcgacggc aggcaagtgg 20

[0484] <210> 9

[0485] <211> 20

[0486] <212> DNA

[0487] <213> 人工序列(Artificial Sequence)

[0488] <400> 9

[0489] cggcgacggc gaataagtgg 20

[0490] <210> 10

[0491] <211> 20

[0492] <212> DNA

[0493] <213> 人工序列(Artificial Sequence)

[0494] <400> 10

[0495] cggcgacggc gagcgggtgg 20

[0496] <210> 11

[0497] <211> 20

[0498] <212> DNA

[0499] <213> 人工序列(Artificial Sequence)

[0500] <400> 11

[0501] cggcgacggc gagcaaacgg 20

[0502] <210> 12

[0503] <211> 20

[0504] <212> DNA

[0505] <213> 人工序列(Artificial Sequence)

[0506] <400> 12

[0507] cggcgacggc gagcaagtaa 20

[0508] <210> 13

[0509] <211> 1833

[0510] <212> PRT

[0511] <213> 人工序列(Artificial Sequence)

[0512] <400> 13

[0513] Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

[0514] 1 5 10 15

[0515] Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

[0516] 20 25 30

[0517] Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu

[0518] 35 40 45

[0519] Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr

[0520] 50 55 60

[0521] Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His

[0522] 65 70 75 80

[0523] Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu

[0524] 85 90 95

[0525] Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr

[0526] 100 105 110

[0527] Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu

[0528] 115 120 125

[0529] Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe

[0530] 130 135 140

[0531] Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn

[0532] 145 150 155 160

[0533] Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His

[0534] 165 170 175

[0535] Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu

[0536] 180 185 190

[0537] Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu

[0538] 195 200 205

[0539] Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe

[0540] 210 215 220

[0541] Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile

[0542] 225 230 235 240

[0543] Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser

[0544] 245 250 255

[0545] Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys

[0546] 260 265 270

[0547] Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr

[0548] 275 280 285

[0549] Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln

[0550] 290 295 300

[0551] Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln

[0552] 305 310 315 320

[0553] Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser

[0554] 325 330 335

[0555] Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr

[0556] 340 345 350

[0557] Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His

[0558] 355 360 365

[0559] Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu

[0560] 370 375 380

[0561] Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly

[0562] 385 390 395 400

[0563] Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys

[0564] 405 410 415

[0565] Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu

[0566] 420 425 430

[0567] Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser

[0568] 435 440 445

[0569] Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg

[0570] 450 455 460

[0571] Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu

[0572] 465 470 475 480

[0573] Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg

[0574] 485 490 495

[0575] Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile

[0576] 500 505 510

[0577] Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln

[0578] 515 520 525

[0579] Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu

[0580] 530 535 540

[0581] Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr

[0582] 545 550 555 560

[0583] Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro

[0584] 565 570 575

[0585] Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe

[0586] 580 585 590

[0587] Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe

[0588] 595 600 605

[0589] Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp

[0590] 610 615 620

[0591] Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile

[0592] 625 630 635 640

[0593] Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu

[0594] 645 650 655

[0595] Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu

[0596] 660 665 670

[0597] Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys

[0598] 675 680 685

[0599] Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys

[0600] 690 695 700

[0601] Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp

[0602] 705 710 715 720

[0603] Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Ala Phe Ala Ala Leu Ile

[0604] 725 730 735

[0605] Ala Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val

[0606] 740 745 750

[0607] Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly

[0608] 755 760 765

[0609] Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp

[0610] 770 775 780

[0611] Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile

[0612] 785 790 795 800

[0613] Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser

[0614] 805 810 815

[0615] Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser

[0616] 820 825 830

[0617] Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu

[0618] 835 840 845

[0619] Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp

[0620] 850 855 860

[0621] Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile

[0622] 865 870 875 880

[0623] Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu

[0624] 885 890 895

[0625] Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu

[0626] 900 905 910

[0627] Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala

[0628] 915 920 925

[0629] Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg

[0630] 930 935 940

[0631] Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu

[0632] 945 950 955 960

[0633] Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser

[0634] 965 970 975

[0635] Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val

[0636] 980 985 990

[0637] Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp

[0638] 995 1000 1005

[0639] Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His

[0640] 1010 1015 1020

[0641] Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr

[0642] 1025 1030 1035 1040

[0643] Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp

[0644] 1045 1050 1055

[0645] Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr

[0646] 1060 1065 1070

[0647] Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu

[0648] 1075 1080 1085

[0649] Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr

[0650] 1090 1095 1100

[0651] Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala

[0652] 1105 1110 1115 1120

[0653] Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys

[0654] 1125 1130 1135

[0655] Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys

[0656] 1140 1145 1150

[0657] Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys

[0658] 1155 1160 1165

[0659] Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val

[0660] 1170 1175 1180

[0661] Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys

[0662] 1185 1190 1195 1200

[0663] Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn

[0664] 1205 1210 1215

[0665] Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp

[0666] 1220 1225 1230

[0667] Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly

[0668] 1235 1240 1245

[0669] Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu

[0670] 1250 1255 1260

[0671] Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His

[0672] 1265 1270 1275 1280

[0673] Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu

[0674] 1285 1290 1295

[0675] Phe Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile

[0676] 1300 1305 1310

[0677] Ser Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys

[0678] 1315 1320 1325

[0679] Val Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln

[0680] 1330 1335 1340

[0681] Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro

[0682] 1345 1350 1355 1360

[0683] Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr

[0684] 1365 1370 1375

[0685] Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr

[0686] 1380 1385 1390

[0687] Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Lys

[0688] 1395 1400 1405

[0689] Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys Glu

[0690] 1410 1415 1420

[0691] Leu Gly Gly Gly Gly Thr Gly Gly Gly Gly Ser Ala Glu Tyr Val Arg

[0692] 1425 1430 1435 1440

[0693] Ala Leu Phe Asp Phe Asn Gly Asn Asp Glu Glu Asp Leu Pro Phe Lys

[0694] 1445 1450 1455

[0695] Lys Gly Asp Ile Leu Arg Ile Arg Asp Lys Pro Glu Glu Gln Trp Trp

[0696] 1460 1465 1470

[0697] Asn Ala Glu Asp Ser Glu Gly Lys Arg Gly Met Ile Leu Val Pro Tyr

[0698] 1475 1480 1485

[0699] Val Glu Lys Tyr Ser Gly Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys

[0700] 1490 1495 1500

[0701] Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys Ser Gly Val Asp

[0702] 1505 1510 1515 1520

[0703] Met Thr Asp Ala Glu Tyr Val Arg Ile His Glu Lys Leu Asp Ile Tyr

[0704] 1525 1530 1535

[0705] Thr Phe Lys Lys Gln Phe Phe Asn Asn Lys Lys Ser Val Ser His Arg

[0706] 1540 1545 1550

[0707] Cys Tyr Val Leu Phe Glu Leu Lys Arg Arg Gly Glu Arg Arg Ala Cys

[0708] 1555 1560 1565

[0709] Phe Trp Gly Tyr Ala Val Asn Lys Pro Gln Ser Gly Thr Glu Arg Gly

[0710] 1570 1575 1580

[0711] Ile His Ala Glu Ile Phe Ser Ile Arg Lys Val Glu Glu Tyr Leu Arg

[0712] 1585 1590 1595 1600

[0713] Asp Asn Pro Gly Gln Phe Thr Ile Asn Trp Tyr Ser Ser Trp Ser Pro

[0714] 1605 1610 1615

[0715] Cys Ala Asp Cys Ala Glu Lys Ile Leu Glu Trp Tyr Asn Gln Glu Leu

[0716] 1620 1625 1630

[0717] Arg Gly Asn Gly His Thr Leu Lys Ile Trp Ala Cys Lys Leu Tyr Tyr

[0718] 1635 1640 1645

[0719] Glu Lys Asn Ala Arg Asn Gln Ile Gly Leu Trp Asn Leu Arg Asp Asn

[0720] 1650 1655 1660

[0721] Gly Val Gly Leu Asn Val Met Val Ser Glu His Tyr Gln Cys Cys Arg

[0722] 1665 1670 1675 1680

[0723] Lys Ile Phe Ile Gln Ser Ser His Asn Gln Leu Asn Glu Asn Arg Trp

[0724] 1685 1690 1695

[0725] Leu Glu Lys Thr Leu Lys Arg Ala Glu Lys Trp Arg Ser Glu Leu Ser

[0726] 1700 1705 1710

[0727] Ile Met Ile Gln Val Lys Ile Leu His Thr Thr Lys Ser Pro Ala Val

[0728] 1715 1720 1725

[0729] Gly Gly Gly Gly Ser Glu Phe Ser Gly Gly Ser Thr Asn Leu Ser Asp

[0730] 1730 1735 1740

[0731] Ile Ile Glu Lys Glu Thr Gly Lys Gln Leu Val Ile Gln Glu Ser Ile

[0732] 1745 1750 1755 1760

[0733] Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile Gly Asn Lys Pro Glu

[0734] 1765 1770 1775

[0735] Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu Ser Thr Asp Glu Asn

[0736] 1780 1785 1790

[0737] Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr Lys Pro Trp Ala Leu

[0738] 1795 1800 1805

[0739] Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile Lys Met Leu Ser Gly

[0740] 1810 1815 1820

[0741] Gly Ser Pro Lys Lys Lys Arg Lys Val

[0742] 1825 1830

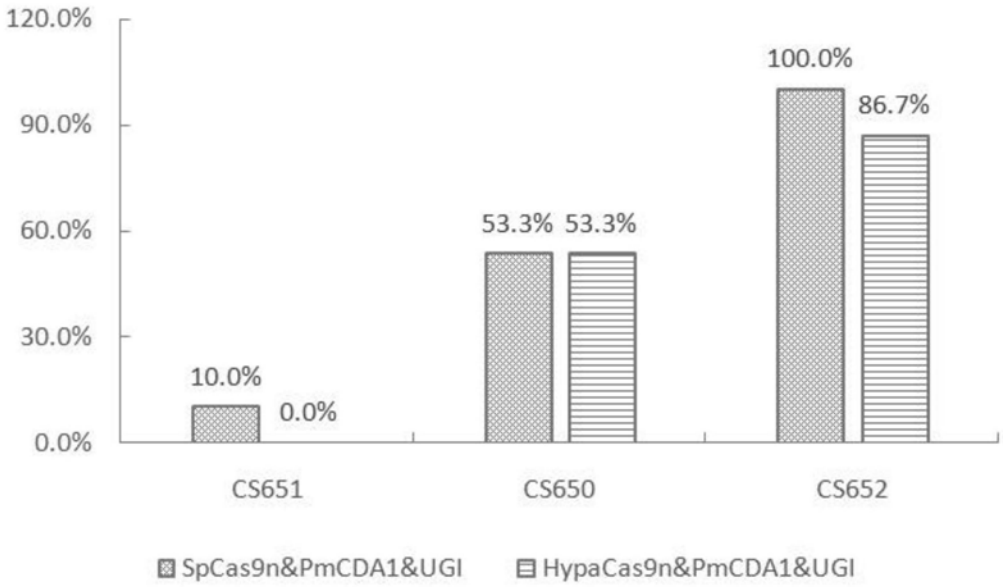
120.0%
90.0%
60.0%
30.0%
0.0%
10.0%
0.0%
53.3%
53.3%
100.0%
86.7%
CS651
CS650
CS652
SpCas9n&PmCDA1&UGI
HypaCas9n&PmCDA1&UGI

图1

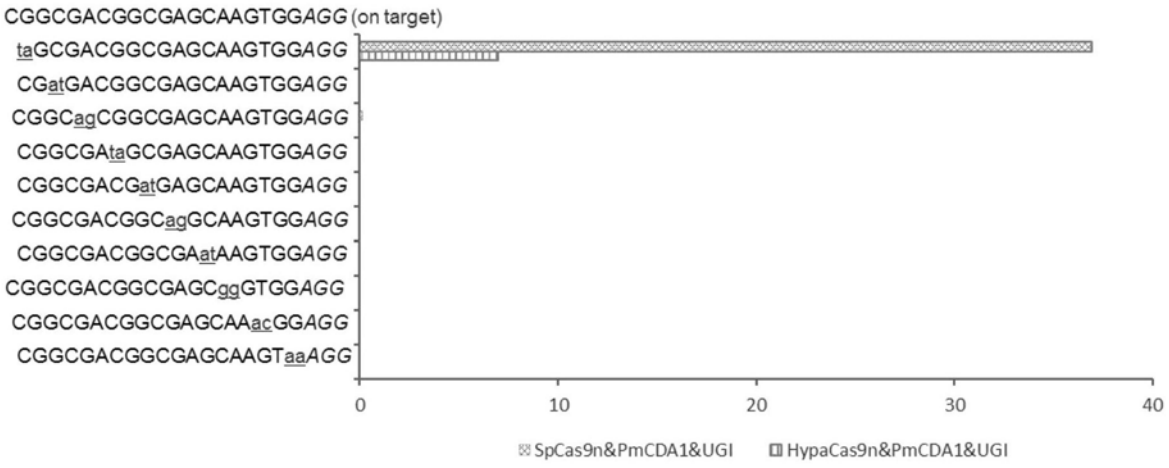
CGGCGACGGCGAGCAAGTGGAGG (on target)
taGCGACGGCGAGCAAGTGGAGG
CGatGACGGCGAGCAAGTGGAGG
CGGCagCGGCGAGCAAGTGGAGG
CGGCGAtaGCGAGCAAGTGGAGG
CGGCGACGatGAGCAAGTGGAGG
CGGCGACGGCagGCAAGTGGAGG
CGGCGACGGCGAatAAGTGGAGG
CGGCGACGGCGAGCggGTGGAGG
CGGCGACGGCGAGCAAacGGAGG
CGGCGACGGCGAGCAAGTaaAGG
0
10
20
30
40
SpCas9n&PmCDA1&UGI
HypaCas9n&PmCDA1&UGI

图2