



(12) 发明专利

(10) 授权公告号 CN 106103706 B

(45) 授权公告日 2021.08.24

(21) 申请号 201480076396.6
(22) 申请日 2014.09.18
(65) 同一申请的已公布的文献号
申请公布号 CN 106103706 A
(43) 申请公布日 2016.11.09
(30) 优先权数据
61/921,007 2013.12.26 US
61/930,782 2014.01.23 US
14/211,117 2014.03.14 US
PCT/US2014/029068 2014.03.14 US
PCT/US2014/029304 2014.03.14 US
PCT/US2014/028630 2014.03.14 US
PCT/US2014/035162 2014.04.23 US

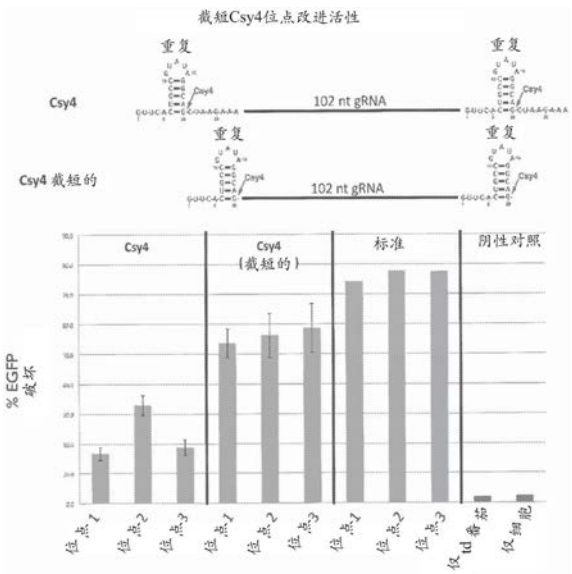
(86) PCT国际申请的申请数据
PCT/US2014/056416 2014.09.18
(87) PCT国际申请的公布数据
W02015/099850 EN 2015.07.02
(73) 专利权人 通用医疗公司
地址 美国马萨诸塞州
(72) 发明人 S·特赛 K·J·乔昂格
(74) 专利代理机构 北京林达刘知识产权代理事
务所(普通合伙) 11277
代理人 刘新宇 李茂家
(51) Int.Cl.
C12N 9/22 (2006.01)
审查员 翟羽佳

(85) PCT国际申请进入国家阶段日
2016.08.25

权利要求书3页 说明书15页
序列表26页 附图9页

(54) 发明名称
多重引导RNA

(57) 摘要
用于任选地在哺乳动物细胞中将高活性CRISPR引导RNA (gRNA) 从RNA聚合酶II和III启动子进行多重表达的方法和构建体。本发明至少部分地基于Csy4的发现,Csy4是一种内切核糖核酸酶,该内切核糖核酸酶识别短RNA发夹序列,可以用于切下在单个较长RNA转录物上编码的多功能gRNA (产生自RNA pol II或III启动子),在该较长RNA转录物中的各个gRNA由Csy4切割位点隔开。



1. 一种脱氧核糖核酸 (DNA) 分子, 包括:

编码引导RNA (gRNA) 的多个序列, 其中每个gRNA侧翼是至少一个Csy4切割序列, 该Csy4切割序列由GTTCAGTCCGTATAGGCAG组成, 并且该gRNA包括以下序列:

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCU;
 (X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCG (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGC (X_N) ,
 (X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGC;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC; 或
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC,

其中X₁₇₋₂₀是与靶序列的17-20个连续核苷酸的互补链互补的序列, 并且X_N是并不干扰核糖核酸与Cas9的结合的0-300个核苷酸序列。

2. 如权利要求1所述的DNA分子, 可操作地连接至启动子序列。

3. 如权利要求1所述的DNA分子, 包括两个、三个、或更多个gRNA序列。

4. 如权利要求2所述的DNA分子, 其中该启动子序列是RNA聚合酶II (Po1 II) 启动子或Po1 III启动子。

5. 如权利要求2所述的DNA分子, 其中该启动子序列是RNA Po1 II启动子。

6. 如权利要求5所述的DNA分子, 其中该Po1 II启动子选自下组, 该组由以下各项组成: CAG、EF1A、CAGGS、PGK、UbiC、CMV、B29、结蛋白、内皮因子、FLT-1、GFPA、和SYN1启动子。

7. 一种DNA分子, 包括与一个、两个、三个或更多个盒连接的启动子序列, 这些盒包括: 连接至Csy4切割位点的、编码引导RNA (gRNA) 的序列, 该Csy4切割位点由GTTCAGTCCGTATAGGCAG组成, 并且该gRNA包括以下序列:

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCU;
 (X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCG (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) ;
 (X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC GAGUCGGUGC (X_N) ,

(X₁₇₋₂₀) GUUUAAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGC;

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC;或

(X₁₇₋₂₀) GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC,

其中X₁₇₋₂₀是与靶序列的17-20个连续核苷酸的互补链互补的序列,并且X_N是并不干扰核糖核酸与Cas9的结合的0-300个核苷酸序列。

8.如权利要求7所述的DNA分子,包括:

Po1 II启动子,该Po1 II启动子可操作地连接至编码第一引导RNA的第一序列,该第一序列连接至一个Csy4切割位点,该Csy4切割位点连接至编码第二引导RNA的第二序列,该第二序列连接至一个Csy4切割位点,该Csy4切割位点连接至编码第三引导RNA的第三序列,该第三序列连接至Csy4切割位点。

9.如权利要求7所述的DNA分子,包括约100nt的引导RNA序列。

10.如权利要求1或7所述的DNA分子,进一步包括编码功能性Csy4酶的序列。

11.如权利要求1或权利要求7所述的DNA分子,其中

该靶序列是紧接前间区序列邻近基序(PAM)的5'的靶序列。

12.如权利要求1或权利要求7所述的DNA分子,包括在该分子的3'端的一个或多个U。

13.一种RNA分子,由如权利要求1-12所述的DNA分子转录。

14.一种载体,包括如权利要求1-12所述的DNA分子。

15.如权利要求14所述的载体,进一步包括编码功能性Csy4酶的序列。

16.一种在细胞中产生多个引导RNA的方法,该方法包括在该细胞中表达如权利要求1-12所述的DNA分子,或使该细胞与如权利要求13所述的RNA分子接触。

17.如权利要求16所述的方法,其中该细胞是哺乳动物细胞并且该细胞还表达外源功能性Csy4酶序列,或该方法进一步包括给予功能性Csy4酶或编码功能性Csy4酶的核酸。

18.一种改变细胞中多个靶基因的表达的方法,该方法包括表达DNA分子,该DNA分子包括编码引导RNA(gRNA)的多个序列,其中每个gRNA包括与靶基因的至少17-20nt互补的序列,并且每个gRNA侧翼是至少一个Csy4切割序列,该Csy4切割序列由GTTCAGTCCGTATAGGCAG组成,并且该gRNA包括以下序列:

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG;

(X₁₇₋₂₀) GUUUUAGAGCUAUGCU;

(X₁₇₋₂₀) GUUUUAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCG (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUC (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUC
(X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGC (X_N);

(X₁₇₋₂₀) GUUUAAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC

GAGUCGGUGC;

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC;或

(X₁₇₋₂₀) GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC,

其中X₁₇₋₂₀是与靶序列的17-20个连续核苷酸的互补链互补的序列,并且X_N是并不干扰核糖核酸与Cas9的结合的0-300个核苷酸序列。

19.一种改变细胞中靶基因的表达的方法,该方法包括表达DNA分子,该DNA分子包括编码引导RNA (gRNA) 的多个序列,其中每个gRNA包括与该靶基因的至少17-20nt互补的序列,并且每个gRNA侧翼是至少一个Csy4切割序列,该Csy4切割序列由GTTCAGTCCGTATAGGCAG组成,并且该gRNA包括以下序列:

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUG;

(X₁₇₋₂₀) GUUUUAGAGCUAAGCU;

(X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUCCG (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGAAAGCAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUC (X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUC
(X_N);

(X₁₇₋₂₀) GUUUUAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGC (X_N),

(X₁₇₋₂₀) GUUUAAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACC
GAGUCGGUGC;

(X₁₇₋₂₀) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC;或

(X₁₇₋₂₀) GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA
AAGUGGCACCGAGUCGGUGC,

其中X₁₇₋₂₀是与靶序列的17-20个连续核苷酸的互补链互补的序列,并且X_N是并不干扰核糖核酸与Cas9的结合的0-300个核苷酸序列。

20.如权利要求18或权利要求19所述的方法,其中该靶序列是紧接前间区序列邻近基序(PAM)的5'的靶序列。

多重引导RNA

[0001] 相关申请的交叉引用

[0002] 本申请要求于2013年12月26日提交的临时专利申请61/921,007;2014年3月14日提交的临时专利申请14/211,117;2014年3月14日提交的PCT/US 2014/029068;2014年3月14日提交的PCT/US 2014/028630;2014年3月23日提交的PCT/US 2014/035162;和2014年3月14日提交的PCT/US 2014/029304的权益。所有以上内容均通过引用结合在此。

[0003] 联邦资助的研究或进展

[0004] 本发明是由国立卫生研究院授予的在授权号DP1 GM105378下的政府支持所完成的。政府在本发明中享有一定的权利。

技术领域

[0005] 描述的是用于任选地在哺乳动物细胞中将高活性CRISPR引导RNA (gRNA) 从RNA聚合酶II和III启动子进行多重表达的方法和构建体。

[0006] 背景

[0007] Cas9核酸酶形成了可编程RNA引导的成簇规律间隔短回文重复 (CRISPR) /CRISPR 关联的 (Cas) 系统 (Wiedenheft (魏登赫夫特) 等人, 自然 (Nature) 482, 331-338 (2012); Horvath (霍瓦特) 等人, 科学 (Science) 327, 167-170 (2010); Terns (腾斯) 等人, 微生物学当前观点 (Curr Opin Microbiol) 14, 321-327 (2011)), 在体外、在哺乳动物细胞中、和在活模式生物, 例如斑马鱼中, 该系统可以用于产生靶DNA序列中的位点特异性断裂 (王 (Wang) 等人, 细胞 (Cell) 153, 910-918 (2013); 沈 (Shen) 等人, 细胞研究 (Cell Res) (2013); 迪卡罗 (Dicarlo) 等人, 核酸研究 (Nucleic Acids Res) (2013); 姜 (Jiang) 等人, 自然生物技术 (Nat Biotechnol) 31, 233-239 (2013); 季聂克 (Jinek) 等人, Elife 2, e00471 (2013); 黄 (Hwang) 等人, 自然生物技术 (Nat Biotechnol) 31, 227-229 (2013); 丛 (Cong) 等人, 科学 (Science) 339, 819-823 (2013); 马里 (Mali) 等人, 科学 (Science) 339, 823-826 (2013c); 丘 (Cho) 等人, 自然生物技术 (Nat Biotechnol) 31, 230-232 (2013); 格拉茨 (Gratz) 等人, 遗传学 (Genetics) 194 (4): 1029-35 (2013))。短的大约100nt引导RNA (gRNA) 与Cas9复合并且指导核酸酶至特异性靶DNA位点; 靶向是由在gRNA的5'端的至少17-20个核苷酸 (nt) 的序列介导的, 该序列被设计为经由工程化的gRNA的前17-20个核苷酸和紧接前间区序列邻近基序 (PAM), 例如匹配序列NGG或NAG的PAM定位的感兴趣的靶基因组DNA序列的互补链之间的简单碱基对互补性互补和相互作用 (沈 (Shen) 等人, 细胞研究 (Cell Res) (2013); 迪卡罗 (Dicarlo) 等人, 核酸研究 (Nucleic Acids Res) (2013); 姜 (Jiang) 等人, 自然生物技术 (Nat Biotechnol) 31, 233-239 (2013); 季聂克 (Jinek) 等人, Elife 2, e00471 (2013); 黄 (Hwang) 等人, 自然生物技术 (Nat Biotechnol) 31, 227-229 (2013); 丛 (Cong) 等人, 科学 (Science) 339, 819-823 (2013); 马里 (Mali) 等人, 科学 (Science) 339, 823-826 (2013c); 丘 (Cho) 等人, 自然生物技术 (Nat Biotechnol) 31, 230-232 (2013); 季聂克 (Jinek) 等人, 科学 (Science) 337, 816-821 (2012))。gRNA还可以指导催化失活的Cas9蛋白 (称为dCas9, 参见季聂克 (Jinek) 等人, 科学 (Science) 337: 816-821 (2012)), 该蛋白进而与效应物结构域 (例如

转录活化结构域)融合,参见例如2013年3月15日提交的USSN 61/799,647,和2013年6月21日提交的61/838,148,将这二者通过引用结合在此。后面的系统使得异源效应物结构域的RNA引导的募集能够到达感兴趣的基因组座位。

[0008] 概述

[0009] 本发明至少部分地基于Csy4的发现,Csy4是一种内切核糖核酸酶,该内切核糖核酸酶识别短RNA发夹序列,可以用于切下在单个较长RNA转录物上编码的多功能gRNA(产生自RNA pol II或III启动子),在该较长RNA转录物中的各个gRNA由Csy4切割位点隔开。

[0010] 因此,在一个第一方面,本发明提供了脱氧核糖核酸(DNA)分子,该DNA分子包括编码引导RNA(gRNA)的多个序列,其中每个gRNA侧翼是至少一个Csy4切割序列,该Csy4切割序列包括序列GTTCACTGCCGTATAGGCAG (SEQ ID NO:1) 或GTTCACTGCCGTATAGGCAGCTAAGAAA (SEQ ID NO:2) 或由其组成。

[0011] 在一些实施例中,该DNA分子可操作地连接至启动子序列。

[0012] 在一些实施例中,该DNA分子包括两个、三个、或更多个gRNA序列,每个序列侧翼是至少一个Csy4切割序列。

[0013] 在一些实施例中,该启动子序列是RNA聚合酶II (Pol II) 启动子或Pol III启动子,优选是RNA Pol II启动子。在一些实施例中,该Pol II启动子选自下组,该组由以下各项组成:CAG、EF1A、CAGGS、PGK、UbiC、CMV、B29、结蛋白、内皮因子、FLT-1、GFPA、和SYN1启动子。

[0014] 在另一方面,本发明提供了DNA分子,该DNA分子包括与一个、两个、三个或更多个盒连接的启动子序列,这些盒包括:编码引导RNA的序列,即约100nt的序列,例如95-300nt,例如对于基于酿脓链球菌的系统是95-105nt,该序列连接至Csy4切割位点,例如SEQ ID NO:1或2。

[0015] 在一些实施例中,该DNA分子包括Pol II启动子,该Pol II启动子可操作地连接至编码第一引导RNA的第一序列,该第一序列连接至一个Csy4切割位点,该Csy4切割位点连接至编码第二引导RNA的第二序列,该第二序列连接至一个Csy4切割位点,该Csy4切割位点连接至编码第三引导RNA的第三序列,该第三序列连接至Csy4切割位点。在一些实施例中,包括连接至Csy4切割位点的另外的引导RNA。例如,该DNA分子可以具有以下结构:启动子-C4-gRNA-C4-gRNA-C4-gRNA-C4启动子-C4-gRNA-C4-gRNA-C4-gRNA-C4-gRNA-C4启动子-C4-gRNA-C4-gRNA-C4-gRNA-C4-gRNA-C4-gRNA C4诸如此类。在这一例示中,C4是编码Csy4RNA切割位点的序列并且gRNA是编码引导RNA的序列。

[0016] 在一些实施例中,Cas9sgRNA包括以下序列:

[0017] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUG (X_N) (SEQ ID NO:4);

[0018] (X_{17-20}) GUUUUAGAGCUA (SEQ ID NO:5);

[0019] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:6);

[0020] (X_{17-20}) GUUUUAGAGCUAUGCU (SEQ ID NO:7);

[0021] (X_{17-20}) GUUUUAGAGCUAGAAAAGCAAGUAAAAUAAGGCUAGUCCG (X_N) (SEQ ID NO:8);

[0022] (X_{17-20}) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) (SEQ ID NO:9);

[0023] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCG

UUAUC (X_N) (SEQ ID NO:10) ;

[0024] (X_{17-20}) GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGC (X_N) (SEQ ID NO:11) ,

[0025] (X_{17-20}) GUUUAAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGG CACCGAGUCGGUGC (SEQ ID NO:12) ;

[0026] (X_{17-20}) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:13) ;或

[0027] (X_{17-20}) GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUU GAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:14) ,

[0028] 其中 X_{17-20} 是与靶序列的、优选是紧接前间区序列邻近基序(PAM)的5'的靶序列的17-20个连续核苷酸的互补链互补的序列,并且 X_N 是并不干扰核糖核酸与Cas9的结合的任何序列。尽管在此用酿脓链球菌Cas9系统例证了 X_{17-20} 的序列,但是还可以使用较长的序列,例如,如合乎其他系统的话。

[0029] 在一些实施例中,该DNA分子还包括编码功能性Csy4酶的序列。

[0030] 在此还提供的是包括在此描述的DNA分子的载体,例如,任选地包括编码功能性Csy4酶的序列的载体。在此还提供的是由这些DNA分子生产的多重转录物,例如仍未用Csy4切割的完整RNA。

[0031] 仍在另一方面,在此提供的是用于在细胞中生产多个引导RNA的方法。这些方法包括在细胞中表达在此描述的DNA分子。

[0032] 在一些实施例中,该细胞是哺乳动物细胞,并且该细胞还表达外源功能性Csy4酶序列,或该方法进一步包括给予功能性Csy4酶或编码功能性Csy4酶的核酸。

[0033] 在另一方面,本发明提供了用于在细胞中改变一个或多个靶基因的表达的方法。这些方法包括表达如在此描述的DNA分子,例如包括编码引导RNA(gRNA)的多个序列的DNA分子,其中每个gRNA包括与一个或多个靶基因的至少17-20nt互补的可变序列,并且每个gRNA侧翼是至少一个Csy4切割序列,该Csy4切割序列包括序列GTTCAGTCCGTATAGGCAG (SEQ ID NO:1) 或GTTCAGTCCGTATAGGCAGCTAAGAAA (SEQ ID NO:2) 或由其组成。

[0034] 在本发明的方法和组合物中,gRNA可以是,如在此描述的,包括融合的tracrRNA和crRNA的单个引导RNA,或可以只包括crRNA,并且可以从相同或不同的DNA分子表达tracrRNA。因此,在一些实施例中,在此描述的DNA分子还包括编码tracrRNA的序列。在一些实施例中,这些方法包括在细胞中表达分开的tracrRNA,例如使这些细胞与表达tracrRNA的载体或DNA分子接触。

[0035] 除非另外定义,在此所使用的全部技术术语和科学术语具有与本发明所属领域之内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。本文提及的所有出版物、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。

[0036] 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

[0037] 图1是用于如下初始多重实验的示意性例示构建体:

[0038] 1+2: 同向重复crRNA阵列和Cas9, 具有分开的tracrRNA

[0039] 3+4: 通过Csy4位点分开的短crRNA阵列, 具有Csy4、Cas9、和分开的tracrRNA

[0040] 5+6: 通过Csy4位点分开的全长嵌合gRNA。

[0041] 7: nls-FLAG标签化的Cas9

[0042] 图2是示出在表达图1中示出的构建体的细胞中的实验结果的柱状图。Csy4位点+完整gRNA (构建体5和6) 是最有效的多重框架。

[0043] 图3是示例性标准和基于多重Csy4的gRNA框架, 和它们生产的转录物的示意性概览和比较。注意到, Csy4使得能够使用RNA Pol II启动子 (例如CAG) 作为U6的替代物, U6是一种RNA Pol III启动子。

[0044] 图4是示出在人类细胞中, Csy4切割产生具有更高活性的gRNA的截短识别位点的柱状图。截短位点的加工还留下一个干净的5' 端, 有效除去了由U6启动子强加的对gRNA靶序列的5' G限制。

[0045] 图5A-C是示出在单个人类细胞中编辑的2-靶多重的证据的序列。在目的位点2或3观察到单个缺失。对于位点2和3, 观察到在同一序列上的多个缺失。还观察到跨位点2和3的缺失。

[0046] 图6是使用基于Csy4的系统, 示出三个gRNA的成功多重表达的示意图。

[0047] 图7是示出通过Csy4, 从RNA Pol II转录的mRNA切除的gRNA可以有效募集Cas9核酸酶至人类细胞中的特定靶标。在这些实验中, 在从RNA Pol II CAG启动子制成的较长mRNA转录物中表达gRNA。

[0048] 详述

[0049] Cas9系统的一个潜在优点是募集核酸酶活性或异源效应物结构域至细胞中的多于一个基因组座位或靶位点的能力。然而, 此类多重应用需要在细胞中有效表达多于一个gRNA的能力。对于哺乳动物细胞, 已经使用RNA聚合酶III启动子 (例如U6启动子) 来表达单个短gRNA。用来在人类细胞中从单个转录物表达多个gRNA组分的先前尝试未被证明有效。

[0050] 对于Cas9系统, 另外的所希望的能力是产生组分的可诱导版本并且实现组分的组织特异性表达。先前已经描述了可诱导的和/或组织特异性的RNA聚合酶II启动子。然而, 尽管Cas9或dCas9蛋白可以从此类RNA pol II启动子表达, 但是短的、限定的gRNA不能以此方式表达, 因为来自RNA pol II的转录的起始和停止位点是不精确的。事实上, 至今所有gRNA都已经从RNA聚合酶III启动子表达, 并且这些启动子都理想地适合短RNA的表达。

[0051] 如在此所证明, Csy4是一种内切核糖核酸酶, 该内切核糖核酸酶识别短RNA发夹序列, 可以用于切下在单个较长RNA转录物上编码的多功能gRNA (产生自RNA pol II或III启动子) 盒, 在该较长RNA转录物中的各个gRNA由Csy4切割位点隔开。可以从较长RNA转录物成功切割功能gRNA, 这些RNA转录物是从RNA pol II启动子表达的。

[0052] gRNA/Csy4多聚盒

[0053] 因此, 在此描述的是编码较长的RNA转录物的DNA分子, 该较长的RNA转录物在此称为多聚盒, 该多聚盒包括两个或更多个gRNA序列, 其中每个gRNA侧翼是Csy4切割序列。该DNA分子还可以包括启动子, 并且可以任选地包括一个或多个其他转录调节序列, 例如增强

子、沉默子、隔离子、和聚A序列。参见,例如,徐(Xu)等人,基因(Gene),2001年7月11日;272(1-2):149-56。

[0054] 启动子

[0055] 本领域已知可以用于本发明的方法的多个启动子。在一些实施例中,该启动子是Pol III或Pol III启动子,优选是Pol II启动子。已经描述了不同Pol II启动子,并且这些Pol II启动子可以用于本发明的组合物和方法,包括CAG启动子(参见例如亚历克斯奥波罗(Alexopoulou)等人,BMC细胞生物学(BMC Cell Biology) 9:2,2008;宫崎(Miyazaki)等人,基因(Gene) 79(2):269-77(1989);丹羽(Niwa)等人,基因(Gene) 108(2):193-9(1991));另外的启动子包括EF1A、CAGGS、PGK、UbiC和CMV启动子,连同组织特异性启动子,例如除其他之外,B29、结蛋白、内皮因子、FLT-1、GFPA、SYN1;多个启动子的序列是本领域已知的。例如,可以分别从pSicoR和pSicoR PGK扩增CMV和PGK启动子(凡吐拉(Ventura)等人,美国国家科学院院刊(Proc Natl Acad Sci U S A) 101:10380-10385(2004)),可以从pDSL_hpUGIH(ATCC)扩增UbiC启动子,可以从pCAGGS(BCCM)扩增CAGGS启动子,并且可以从pEF6载体(英杰公司(Invitrogen))扩增EF1A启动子。在巴特勒(Butler)和角永(Kadonaga),基因和发育(Genes&Dev.) 16:2583-2592(2002)中描述了Pol II核心启动子。由Pol II表达驱动的较大转录物切下gRNA使得人们能够生产gRNA,该gRNA具有在最5'-位置的任何核苷酸(来自U6或其他RNA聚合酶III启动子的标准表达对这一核苷酸的身份加以限制)。

[0056] 在一些实施例中,使用组织特异性启动子,并且可以离开RNA-Pol II转录物加工短的限定的gRNA序列。

[0057] 本领域已知多个Pol III启动子,包括U6小核(sn)RNA启动子、7SK启动子、和H1启动子。参见,例如,罗(Ro)等人,生物技术(BioTechniques) 38(4):625-627(2005)。

[0058] 引导RNA

[0059] 通过使用在其与基因组DNA靶位点互补的5'端携带至少17-20nt的单个gRNA,Cas9核酸酶可以被引导至携带序列NGG的另外的近端前间区序列邻近基序(PAM)的至少17-20nt的特定基因组靶。

[0060] 因此,在此描述的组合物可以包括编码单个引导RNA(sgRNA)的序列,该单个引导RNA(sgRNA)包括融合至正常反式编码的tracrRNA的crRNA,该单个引导RNA例如是单个Cas9引导RNA,如在马里(Mali)等人,科学(Science) 2013年2月15日;339(6121):823-6中描述,其中在5'端的序列与紧接前间区序列邻近基序(PAM),例如NGG、NAG或NNGG的5'的靶序列的17-20个核苷酸(nt)互补。

[0061] 设计和表达引导RNA的方法是本领域已知的。引导RNA总而言之归于两个不同系统:1)系统1使用分开的crRNA和一起发挥功能的tracrRNA,以便引导用Cas9切割;以及2)系统2使用组合在单个系统中的两个分开的引导RNA的嵌合crRNA-tracrRNA杂合体(季聂克(Jinek)等人2012)。tracrRNA可以被可变地截短,并且已经示出在分开系统(系统1)和嵌合gRNA系统(系统2)二者中发挥功能的长度范围。参见例如季聂克(Jinek)等人,科学(Science) 2012;337:816-821;马里(Mali)等人,科学(Science) 2013年2月15日;339(6121):823-6;丛(Cong)等人,科学(Science) 2013年2月15日;339(6121):819-23;以及黄(Hwang)和付(Fu)等人,自然生物技术(Nat Biotechnol.) 2013年三月;31(3):227-9;季聂克(Jinek)等人,Elife 2,e00471(2013))。对于系统2,通常,较长长度的嵌合gRNA已经示出

更大靶活性,但是,不同长度gRNA的相对特异性目前仍未限定,并且因此在某些实例中可能希望使用较短gRNA。在一些实施例中,gRNA与转录起始位点上游约100-800bp内的区域互补,该区域例如在转录起始位点上游约500bp,包括转录起始位点,或在转录起始位点下游约100-800bp,例如约500bp内。在一些实施例中,使用编码多于一个gRNA的载体(例如质粒),例如编码2、3、4、5、或更多个gRNA的质粒,这些gRNA被定向到靶基因的同区域中的不同位点。在题为针对RNA引导的基因组编辑的增加了的特异性(INCREASING SPECIFICITY FOR RNA-GUIDED GENOME EDITING)的临时专利申请序列号61/838,178中描述了增加基因组编辑的特异性的另外的引导RNA和方法。

[0062] 在一些实施例中,该gRNA包括以下序列或由其组成:

[0063] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUG (X_N) (SEQ ID NO:4);

[0064] (X_{17-20}) GUUUUAGAGCUA (SEQ ID NO:5);

[0065] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:6);

[0066] (X_{17-20}) GUUUUAGAGCUAUGC (SEQ ID NO:7);

[0067] (X_{17-20}) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCG (X_N) (SEQ ID NO:8);

[0068] (X_{17-20}) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUAAAAUAAGGCUA GUCCGUUAUC (X_N) (SEQ ID NO:9);

[0069] (X_{17-20}) GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUC (X_N) (SEQ ID NO:10);

[0070] (X_{17-20}) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (X_N) (SEQ ID NO:11);

[0071] (X_{17-20}) GUUUUAGAGCUAGAAUAGCAAGUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:12);

[0072] (X_{17-20}) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGC UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:13);或

[0073] (X_{17-20}) GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUAAAAUAAGGC UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:14);

[0074] 其中 X_{17-20} 是与靶序列的至少17-20个连续核苷酸的互补链互补的序列(尽管在一些实施例中,这一互补性区域可以是大于20nt,例如21、22、23、24、25、26、27、28、29、或30或更多个nt,例如17-30nt)的、优选紧接前间区序列邻近基序(PAM)的5'处的靶序列,例如NGG、NAG、或NNGG。 X_N 是任何序列,其中N(在RNA中)可以是并不干扰核糖核酸与Cas9的结合的0-300或0-200,例如0-100、0-50、或0-20。在一些实施例中,该RNA包括在该分子的3'端的一个或多个U,例如1至8个或更多个U(例如U、UU、UUU、UUUU、UUUUU、UUUUUU、UUUUUUU、UUUUUUUU),这是由于用作用来终止RNA PolIII转录的终止信号的一个或多个T的任选存在。在一些实施例中,该RNA包括在并不与靶序列互补的该RNA分子的5'端的一个或多个,例如高达3个,例如一个、两个、或三个、或更多个另外的核苷酸。任选地,该RNA核苷酸中的一个或多个被修饰,例如锁定(2'-O-4'-C亚甲基桥),是5'-甲基胞苷,是2'-O-甲基-假尿苷,或其中磷酸核糖主链已经被聚酰胺链,例如序列 X_{17-20} 内的核苷酸中的一个或多个,序列 X_N 内的核苷酸中的一个或多个,或该gRNA的任何序列内的核苷酸中的一个或多个替换。

[0075] 例如,在一些实施例中,可以使用在季聂克(Jinek)等人(科学(Science).337

(6096):816-21 (2012)) 中描述的嵌合引导RNA,例如, (X_{17-20}) GUUUUAGAGCUAGAAAUAGCAAGUU AAAUAAGGCUAGUCCG (SEQ ID NO:8); (X_{17-20}) GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGG (SEQ ID NO:9); 在一些实施例中,使用携带有与靶DNA序列互补的5'-末端17-20个核苷酸序列的sgRNA,以及Cas9结合必需的42个核苷酸的3'-末端茎环结构(描述在季聂克(Jinek)等人, Elife.2:e00471 (2013) 中),例如 (X_{17-20}) GUUUUAGAGCUAGAAAUAGCAAGUAAAAUAAGGCUAGUCCG (SEQ ID NO:8)。

[0076] 在一些实施例中,该引导RNA包括在3'端上的一个或多个腺嘌呤(A)或尿嘧啶(U)核苷酸。

[0077] 尽管在此描述的实例利用单个gRNA,但是还可以与双重gRNA(例如在天然存在的系统中发现的crRNA和tracrRNA)一起使用这些方法。在这一情况下,将结合使用本系统,例如以下项(注意对于RNA,在此T应被理解为U)表达的多个不同crRNA,使用单个tracrRNA:

[0078] crRNA序列: X_{17-20} -GTTT TAGAGCTAGAAA (SEQ ID NO:15)

[0079] tracrRNA序列:TAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTC GGTGC (SEQ ID NO:16)。在这一情况下,在此处描述的方法和分子中,crRNA被用作引导RNA,并且可以从相同或不同的DNA分子表达tracrRNA。

[0080] 此外,尽管在此例证了具有互补性的17-20个核苷酸的序列的引导RNA,但是在一些实施例中,可以使用较长的序列,例如21、22、23、24、25、26、27、28、29、或30或更多nt,例如17-30nt代替17-20nt。

[0081] Csy4切割序列

[0082] 在此处描述的方法和组合物中,将Csy4切割序列插入该DNA分子中,这样使得每个引导RNA侧翼是切割序列,其中一个或至少一个切割序列在每个引导RNA之间。示例性Csy4切割序列包括GTTCACTGCCGTATAGGCAG(截短的20nt)(SEQ ID NO:1)和GTTCACTGCCGTATAGGCAGCTAAGAAA(完整28nt)(SEQ ID NO:2)。如在此所证明,在人类细胞中,与使用标准位点相比,使用截短的Csy4切割位点(SEQ ID NO:1)更有效。据本发明的诸位发明人所知,这是在人类细胞中利用Csy4活性的第一个证明。

[0083] 功能性Csy4酶序列

[0084] 在此处描述的方法中,在细胞中还表达了能够在Csy4切割位点切割转录物的功能性Csy4酶。

[0085] 示例性Csy4序列来自Csy4同系物,这些Csy4同系物来自绿脓假单胞菌UCBPP-PA14 (Pa14),鼠疫杆菌AAM85295 (Yp),大肠杆菌UTI89 (Ec89),节瘤偶蹄菌(Dichelobacter nodosus) VCS1703A (Dn),鲍氏不动杆菌AB0057 (Ab),南极细菌属(Moritella sp.) PE36 (MP1、MP01),希瓦氏菌属W3-18-1 (SW),多杀性巴氏杆菌败血亚种Pm70 (Pm),植物腐败菌(Pectobacterium wasabiae) (Pw),和菊欧文氏菌(Dickeya dadantii) Ech703 (Dd),都列出在胡尔维茨(Haurwitz)等人,科学(Science) 329 (5997):1355-1358 (2010) 的图S6中。在优选的实施例中,Csy4来自绿脓假单胞菌。

[0086] 在一些实施例中,Csy4还用于共价连接异源效应物结构域至gRNA。Csy4被认为是单一转换酶,并且在切割后保留结合至其靶发夹序列(斯腾伯格(Sternberg)等人, RNA.2012年4月;18(4):661-72)。因此,预期Csy4保留结合至每个切割的gRNA的3'端。因为如在此所证明,切割的gRNA似乎在人类细胞中发挥功能,但是这一Csy4蛋白在gRNA的3'端

的存在似乎并不影响gRNA与Cas9复合的能力并且指导Cas9活性。因此,假定这些gRNA-Csy4融合还将能够指导携带突变的Cas9突变体,这些突变灭活其催化核酸酶活性(dCas9蛋白)。因此,人们可以将异源功能域(HFD)融合至Csy4(Csy4-HFD),并且然后dCas9:sgRNA:Csy4-HFD复合物可以指导此类结构域至特定基因组座位。此类HFD的实例可以包括其他核酸酶结构域,例如来自FokI的、转录激活因子的或阻遏蛋白的结构域的结构域,或修饰组蛋白或DNA甲基化状态的其他结构域。

[0087] 通过将异源功能域(例如转录活化结构域,例如来自VP64或NF- κ B p65)融合至Csy4的N-末端或C-末端,其中使用或不使用中介连接子,例如约5-20或13-18个核苷酸的连接子,来产生Csy4-HFD。可以将转录活化结构域融合在Csy4的N或C末端上。除了转录活化结构域,还可以使用如本领域中已知的其他异源功能域(例如转录阻遏蛋白(例如KRAB、SID、和其他阻遏蛋白)或沉默子,例如异染色质蛋白1(HP1,也称为swi6),例如HP1 α 或HP1 β ;可以募集融合至固定RNA结合序列的长非编码RNA(lncRNA)的蛋白或肽,例如通过MS2外壳蛋白、内切核糖核酸酶Csy4、或 λ N蛋白结合的那些;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),或组蛋白脱甲基化酶(例如用于赖氨酸或精氨酸残基的脱甲基作用)。此类结构域的多个序列是本领域已知的,例如催化DNA中甲基化的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,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。参见例如WO/2014/144761。

[0088] 人类TET1-3的序列是本领域中已知的,并且在下表中示出:

	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

[0090] *变体(1)表示较长转录物并且编码较长亚型(a)。与变体1相比,变体(2)差别在于5'UTR和在于3'UTR以及编码序列。与亚型a相比,所得亚型(b)更短并且具有不同C末端。

[0091] 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,例如包括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,例如包括氨基酸1580-2052的Tet1催化结构域,包括氨基酸1290-1905的Tet2,和包括氨基酸966-1678的Tet3。参见例如伊耶(Iyer)等人,细胞周期(Cell Cycle).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。

[0092] 其他催化模块可以例如来自伊耶(Iyer)等人,2009年鉴定的蛋白。

[0093] 在一些实施例中,融合蛋白包括Csy4和异源功能域之间的连接子。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)的连接子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,该连接子是短的,例如2-20个氨基酸,并且典型地是柔性的(即包括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,该连接子包括由GGGS (SEQ ID NO:3) 或GGGGS (SEQ ID NO:17) 组成的一个或多个单元,例如GGGS (SEQ ID NO:3) 或GGGGS (SEQ ID NO:17) 单元的两个、三个、四个、或更多个重复。还可以使用其他连接子序列,例如GGS、GGSG (SEQ ID NO:22)、SGSETPGTSESA (SEQ ID NO:23)、SGSETPGTSESATPES (SEQ ID NO:24)、或SGSETPGTSESATPEGGSGGS (SEQ ID NO:25)。

[0094] Cas9

[0095] 在此处描述的方法中,还在细胞中表达Cas9。多种细菌表达Cas9蛋白变体。目前最长使用来自酿脓链球菌的Cas9;一些其他Cas9蛋白具有与酿脓链球菌Cas9的高水平的序列一致性,并且使用相同引导RNA。其他Cas9是更多样化的,使用不同gRNA,并且也识别不同PAM序列(由邻近该RNA指定的序列的蛋白指定的2-5个核苷酸序列)。切林斯基 (Chylinski) 等人将来自大组的细菌的Cas9蛋白分类 (RNA生物学10:5,1-12;2013),并且在其补充图1和补充表1中列出了大量Cas9蛋白,将其通过引用结合在此。在此描述的构建体和方法可以包括使用那些Cas9蛋白中的任一种,及其相应引导RNA或其他相容的引导RNA。在丛 (Cong) 等人 (科学 (Science) 39,819 (2013)) 中,已经示出来自嗜热链球菌LMD-9 CRISPR1系统的Cas9在人类细胞中发挥功能。另外,季聂克 (Jinek) 等人示出在体外,来自嗜热链球菌和英诺克李斯特菌(但不是来自脑膜炎奈瑟氏球菌或空肠弯曲菌,它们可能使用不同引导RNA) 的Cas9直向同源物可以由双重酿脓链球菌gRNA引导,以便切割靶质粒DNA,尽管具有略微降低的效率。

[0096] 在一些实施例中,本系统利用来自酿脓链球菌的Cas9蛋白,作为在细菌中编码,或被密码子优化用于在哺乳动物细胞中表达。融合至HA表位(氨基酸序列DAYPYDVPDYASL (SEQ ID NO:18)) 和核定位信号(氨基酸序列PKKKRKVEDPKKKRKVD (SEQ ID NO:19)) 的酿脓链球菌Cas9的示例性序列(残基1-1368) 如下:

[0097] MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEPNINASGVDAKILSARLSKSRLENLIAQLPGKKNGLFGNLIALSLGLTPNFKSNFLAEDAKLQLSKDYYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGSLSHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVIDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG

GFDSP TVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELN
GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLAD
ANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRI
DLSQLGGDAYPYDVPDYASLGSGSPKKKRKVEDPKKKRKVD (SEQ ID NO:20)

[0098] 参见季聂克 (Jinek) 等人, 2013, 同上。

[0099] 在一些实施例中, 使用包含D10A和H840A突变之一的Cas9序列, 以便致使核酸酶成为切口酶, 或使用包含D10A和H840A突变这二者的Cas9序列, 以便致使蛋白的核酸酶部分催化失活。可以用于在此描述的方法和组合物的催化失活酿脓链球菌Cas9 (dCas9) 的序列如下; 突变处于粗体并且加下划线。

	10	20	30	40	50	60
	MDKKYSIGL A	IGTNSVGWAV	ITDEYKVPSK	KFKVLGNTDR	HSIKKNLIGA	LLFDSGETAE
	70	80	90	100	110	120
	ATRLKRTARR	RYTRRKNRIC	YLQEIFSNE	AKVDDSFHR	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
[0100]	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	HSLLYEYFTV	YNELTKVKYV	TEGMRKPAFL
	550	560	570	580	590	600
	SGEQKKAIVD	LLFKTNRKVT	VKQLKEDYFK	KIECFDSVEI	SGVEDRFNAS	LGTYHDLKI
	610	620	630	640	650	660
	IKDKDFLDNE	ENEDILEDIV	LTLTLFEDRE	MIEERLKTYA	HLFDDKVMKQ	LKRRRYTGWG
	670	680	690	700	710	720
	RLSRKLINGI	RDQSGKTIL	DFLKSDGFAN	RNFMQLIHDD	SLTFKEDIQK	AQVSGQGD SL
	730	740	750	760	770	780
	HEHIANLAGS	PAIKKGILQT	VKVVDLVKV	MGRHKPENIV	IEMARENQTT	QKGQKNSRER

790 800 810 820 830 840
 MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVA
 850 860 870 880 890 900
 IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKMK NYWRQLLNAK LITQRKFDNL
 910 920 930 940 950 960
 TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS
 970 980 990 1000 1010 1020
 KLVSDFRKDF QFYKVVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK
 1030 1040 1050 1060 1070 1080
 MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF
 [0101] 1090 1100 1110 1120 1130 1140
 ATVRKVLSP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPPK YGGFDSPTVA
 1150 1160 1170 1180 1190 1200
 YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK
 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 IDRKYRSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO:21)

[0102] 参见,例如马里(Mali)等人,2013,同上;和季聂克(Jinek)等人,2012,同上。可替代地,Cas9可以是dCas9-杂功能域融合(dCas9-HFD),如在2013年6月21日提交的题为至特定基因组座位的遗传学和表观遗传学调节蛋白的RNA引导的靶向(RNA-GUIDED TARGETING OF GENETIC AND EPIGENOMIC REGULATORY PROTEINS TO SPECIFIC GENOMIC LOCI)的美国临时专利申请,和指定序列号61/838,148中,和在PCT/US2014/027335中所描述。

[0103] 如在此所述,可以从表达载体,例如包括编码Cas9的序列,例如Cas9cDNA或基因的染色体外质粒或病毒载体表达Cas9;可以从已经整合进细胞的基因组的外源Cas9cDNA或基因表达Cas9;可以从编码Cas9的mRNA表达Cas9;Cas9可以是实际Cas9蛋白自身;或在非哺乳动物细胞的情况下,可以是外源Cas9。

[0104] 表达系统

[0105] 可以使用包括表达载体的核酸分子,例如用于在此描述的Csy4/引导RNA构建体的体内或体外表达。用于表达多个gRNA的载体(潜在地,以可诱导的或组织-/细胞-型特异性方式)可以用于研究和治疗应用。

[0106] 为了使用在此描述的融合蛋白和多聚引导RNA盒,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,可以将编码引导RNA盒或Csy4或Cas9蛋白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,用于编码融合蛋白的核酸的储存或操纵或用于生产融合蛋白。还可以将编码引导RNA或融合蛋白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。

[0107] 为了获得表达,典型地将编码引导RNA或融合蛋白的序列亚克隆进包含用来指导转录的启动子的表达载体。适合的细菌和真核启动子是本领域熟知的,并且例如在萨姆布鲁克(Sambrook)等人,分子克隆实验指南(Molecular Cloning,A Laboratory Manual)(第3版,2001);克里格勒(Kriegler),基因转移和表达:实验室手册(Gene Transfer and Expression:A Laboratory Manual)(1990);和分子生物学现代方法(Current Protocols in Molecular Biology)(奥苏贝尔(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、和沙门氏菌属中(帕尔瓦(Palva)等人,1983,基因(Gene)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。

[0108] 多个适合的载体是本领域已知的,例如病毒载体,包括重组反转录病毒、慢病毒、腺病毒、腺病毒伴随病毒、和单纯疱疹病毒1、腺病毒源载体,或重组细菌或真核质粒。例如,表达构建体可以包括:编码区;启动子序列,例如限制表达至所选择的细胞类型的启动子序列,条件启动子,或强一般启动子;增强子序列;未翻译调节序列,例如5'未翻译区(UTR)、3'UTR;聚腺苷酸化位点;和/或隔离子序列。此类序列是本领域已知的,并且技术人员将能够选择适合的序列。参见例如分子生物学现代方法(Current Protocols in Molecular Biology)(奥苏贝尔(Ausubel)F.M.等人(编辑)格林出版联合公司(Greene Publishing Associates), (1989), 9.10-9.14节段和其他标准实验室手册。在一些实施例中,如本领域已知,使用组织特异性启动子,表达可以被限制至具体细胞类型。

[0109] 如以上所描述,用于表达引导RNA的载体可以包括RNA Pol II或Pol III启动子,以便驱动引导RNA的表达。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gRNA。可替代地,可以使用T7启动子,例如用于体外转录,并且RNA可以在体外转录并且纯化。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,当体内给予融合蛋白用于基因调节时,可以使用组成型启动子或诱导型启动子,取决于融合蛋白的具体用途。此外,用于给予融合蛋白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,例如缺氧反应元件、Gal4反应元件、乳糖抑制体反应元件、和小分子控制系统,例如四环素调节系统和RU-486系统(参见例如哥森(Gossen)&布雅德(Bujard),1992,美国国家科学院院刊(Proc.Natl.Acad.Sci.USA),89:5547;奥利吉诺(Oligino)等人,1998,基因治疗(Gene Ther.),5:491-496;王(Wang)等人,1997,基因治疗(Gene Ther.),4:432-441;尼英(Neering)等人,1996,血液(Blood),88:1147-55;和伦德尔(Rendahl)等人,1998,自然生物技术(Nat.Biotechnol.),16:757-761)。

[0110] 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核酸在宿主细胞(原核的或真核的)中表达所需的所有另外的元件。因此,典型的表达盒包含可操作地连接至,例如编码融合蛋白,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。

[0111] 关于融合蛋白的预期用途选择用于运输遗传信息进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。优选的

标签融合蛋白是麦芽糖结合蛋白 (MBP)。此类标签融合蛋白可以用于纯化工程化的TALE重复蛋白。还可以添加表位标签至重组蛋白,以便提供分离的方便方法,用于监测表达,并且用于监测细胞的和亚细胞的定位,例如c-myc或FLAG。

[0112] 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。

[0113] 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融合蛋白编码序列。

[0114] 典型地包括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。

[0115] 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(参见例如科莱 (Colley) 等人,1989,生物化学杂志 (J.Biol.Chem.), 264:17619-22;在酶学方法中对蛋白纯化的引导 (Guide to Protein Purification, in Methods in Enzymology), 182卷 (多伊彻 (Deutscher), 编辑,1990))。根据标准技术,进行真核细胞和原核细胞的转化(参见例如莫里森 (Morrison), 1977,细菌学杂志 (J.Bacteriol.) 132:349-351;克拉克 (Clark) -柯蒂斯 (Curtiss) &柯蒂斯 (Curtiss), 酶学方法 (Methods in Enzymology) 101:347-362 (吴 (Wu) 等人,编辑,1983))。

[0116] 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(参见例如萨姆布鲁克 (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选择的蛋白的宿主细胞中的具体遗传工程化程序。

[0117] 在一些实施例中,Cas9或Csy4蛋白包括提供蛋白易位至细胞核的核定位结构域。已知若干核定位序列 (NLS),并且可以使用任何适合的NLS。例如,很多NLS具有多个碱性氨基酸,称为双向碱性重复(在加西亚 (Garcia) -比斯托 (Bustos) 等人,1991,生物化学与生物物理学学报 (Biochim.Biophys.Acta), 1071:83-101中综述)。包含双向碱性重复的NLS可以位于嵌合蛋白的任何部分中,并且导致嵌合蛋白定位在细胞核内。在优选实施例中,核定位结构域被整合进最终融合蛋白,因为在此描述的融合蛋白的最后功能将典型地需要将蛋白定位在细胞核中。然而,在其中蛋白具有固有的核易位功能的情况下,可能不需要添加分开的核定位结构域。

[0118] 本发明包括载体和包含载体的细胞。

[0119] 文库

[0120] 还在此提供的是gRNA的组合文库,例如在用于研究应用的可诱导的、组织或细胞

型特异性多重载体中,这些应用例如是用于筛选潜在药物靶标或用来限定在基因水平的相互作用。

[0121] 使用方法

[0122] 描述的方法可以包括在细胞中表达,或使细胞与如在此描述的多聚盒接触,加上可以由缩短的gRNA引导的核酸酶,例如,如以上描述的Cas9核酸酶,和如以上描述的Csy4核酸酶。

[0123] 对于同时修饰多个外源基因的表达,或对于靶向单个基因的多个部分,描述的系统是有用且通用的工具。为实现这一点,当前方法需要使用针对待靶向的每一位点的分开的gRNA编码转录物。对于细胞群的多重基因组编辑而言,分开的gRNA并不是最佳的,因为不可能保证每个细胞将表达每个gRNA;使用多个转录物,细胞得到gRNA的复杂并且不均匀的随机混合物。然而,本系统允许来自单个转录物的多个gRNA表达,这允许通过表达多个gRNA,在基因组中靶向多个位点。此外,使用单个转录物系统,每个细胞应表达具有类似化学计量的所有gRNA。因此,这一系统可以容易地用于同时改变大量基因的表达,或用于募集多个Cas9或HFD至单个基因、启动子、或增强子。这一能力具有广泛应用,例如用于基础生物学研究,其中它可以用于研究基因功能,并且操纵多个基因在单个通路中的表达,并且用于合成生物学,其中它将使得研究人员能够在细胞中产生响应于多个输入信号的回路。这一技术可以被实施并且适合于多重复用是相对容易的,这将使其成为具有很多广泛应用的广泛有用的技术。

[0124] 在此描述的方法包括使细胞与编码在此针对一个或多个基因描述的多聚gRNA盒的核酸、以及编码Csy4和Cas9的核酸接触,由此调整该一个或多个基因的表达。

[0125] 实例

[0126] 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。

[0127] 实例1.使用CRISPR/Cas9进行多重编辑

[0128] 使用从单个转录物表达的crRNA或sgRNA的阵列的CRISPR/Cas9,带着进行多重编辑的目标,尝试三种策略:

[0129] 1.同向重复侧翼的crRNA阵列和Cas9,具有分开的tracrRNA

[0130] 2.通过Csy4位点分开的短crRNA阵列,与Csy4、Cas9、和分开的tracrRNA一起表达

[0131] 3.通过Csy4位点分开的全长单个引导RNA (sgRNA)

[0132] 针对有效破坏U2OS-EGFP破坏阵列中的EGFP的能力,测试每组构建体(图1中例示的)。结果示出于表2中。在EGFP破坏阵列中,甚至是对于单个靶标,使用策略1和2设计的构建体显示最低活性;因此进一步实验(以下描述)集中在优化策略3上。

[0133] 实例2.在哺乳动物细胞中,高活性CRISPR引导RNA自RNA聚合酶II和III启动子的多重表达

[0134] 在图3中示出,使用Csy4核酸酶,用于从较长转录物切下gRNA的示例性策略的示意性概览。在用来证明概念验证的初始试验中,针对在人类细胞中的切割,测试Csy4切割的RNA发夹位点的两个版本。为了实现这一点,表达侧翼是两个Csy4切割位点之一的gRNA:

[0135] 1.GTTCACTGCCGTATAGGCAGCTAAGAAA(全28nt)(SEQ ID NO:2)

[0136] 2.GTTCACTGCCGTATAGGCAG(截短的20nt)(SEQ ID NO:1)

[0137] 结果示出,在哺乳动物细胞中,与侧翼是较长的28nt序列的那些相比,在其5'和3'端侧翼有截短的20nt序列的gRNA更具活性(图4)。据本发明的诸位发明人所知,这是在活人类细胞中,Csy4核酸酶可以用于加工RNA转录物的第一个证明。20nt截短位点的一个重要的另外的优点是,不像较长的28nt序列,它并不在从较长转录物加工的gRNA的5'端上留下任何另外的核苷酸(图4)。这使得能够表达在最5'位置具有任何希望的核苷酸的gRNA。相对于从具有在最5'位置对于特定的一个或多个核苷酸的需要的RNA聚合酶III启动子表达gRNA,这是一个改进。

[0138] 使用这一基于Csy4的系统,证明了两个和三个不同gRNA的有效表达(图5和6)。使用这一方法同时表达gRNA诱导在人类细胞中,在预期位点处的改变。

[0139] 这些结果还证明,这一基于Csy4的策略可以与在通过RNA Pol II启动子生产的较长mRNA上编码的gRNA一起使用(图7)。在这些实验中,在通过CAG启动子(一种RNA Pol II启动子)生产的mRNA上编码侧翼是截短Csy4位点的三种不同的单个gRNA之一。如在图7中所示,在人类细胞中,所有这三种构建体都能够产生可以指导Cas9核酸酶的功能gRNA,但是这只是在Csy4存在时才如此。观察到的目标Cas9活性的水平与当使用标准RNA Pol III启动子单独表达这些gRNA时,或作为来自RNA Pol III启动子的Csy4侧翼转录物时观察到的结果可比较(尽管在某种程度上更低)(图7)。

[0140] 总之,本发明的结果证明:1)当通过Csy4切割位点分开时并且在人类细胞中存在Csy4下,可以从单个RNA pol III转录物产生高达三种功能gRNA,2)多种Csy4加工的gRNA可以用于指导Cas9核酸酶以诱导在单个人类细胞中的多重变化,以及3)可以通过Csy4核酸酶,从由RNA聚合酶II启动子制成的较长mRNA转录物切除侧翼是Csy4切割位点的功能gRNA。

[0141] 其他实施方案

[0142] 应理解,虽然已经结合其详细描述,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、优点以及修改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
[0003] <120>多重引导RNA
[0004] <130> 40174-0012W01
[0005] <140> PCT/US14/56416
[0006] <141> 2014-09-18
[0007] <150> PCT/US2014/035162
[0008] <151> 2014-03-23
[0009] <150> 14/211,117
[0010] <151> 2014-03-14
[0011] <150> PCT/US2014/029068
[0012] <151> 2014-03-14
[0013] <150> PCT/US2014/028630
[0014] <151> 2014-03-14
[0015] <150> PCT/US2014/029304
[0016] <151> 2014-03-14
[0017] <150> 61/921,007
[0018] <151> 2013-12-26
[0019] <160> 54
[0020] <170> PatentIn version 3.5
[0021] <210> 1
[0022] <211> 20
[0023] <212> DNA
[0024] <213> 人工序列
[0025] <220>
[0026] <223> 人工序列的描述: 合成寡核苷酸
[0027] <400> 1
[0028] gttcactgcc gtataggcag 20
[0029] <210> 2
[0030] <211> 28
[0031] <212> DNA
[0032] <213> 人工序列
[0033] <220>
[0034] <223> 人工序列的描述: 合成寡核苷酸
[0035] <400> 2
[0036] gttcactgcc gtataggcag ctaagaaa 28
[0037] <210> 3
[0038] <211> 4

[0039]	<212> PRT
[0040]	<213> 人工序列
[0041]	<220>
[0042]	<223> 人工序列的描述: 合成肽
[0043]	<400> 3
[0044]	Gly Gly Gly Ser
[0045]	1
[0046]	<210> 4
[0047]	<211> 342
[0048]	<212> RNA
[0049]	<213> 人工序列
[0050]	<220>
[0051]	<223> 人工序列的描述: 合成寡核苷酸
[0052]	<220>
[0053]	<221> 修饰的碱基
[0054]	<222> (1) .. (20)
[0055]	<223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
[0056]	<220>
[0057]	<221> 修饰的碱基
[0058]	<222> (43) .. (342)
[0059]	<223> a、c、u或g、以及该区域可能包括0-300个核苷酸,其中某些位置可缺失
[0060]	<400> 4
[0061]	nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuguuu ugnnnnnnnn nnnnnnnnnn 60
[0062]	nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
[0063]	nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
[0064]	nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
[0065]	nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
[0066]	nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nn 342
[0067]	<210> 5
[0068]	<211> 32
[0069]	<212> RNA
[0070]	<213> 人工序列
[0071]	<220>
[0072]	<223> 人工序列的描述: 合成寡核苷酸
[0073]	<220>
[0074]	<221> 修饰的碱基
[0075]	<222> (1) .. (20)
[0076]	<223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
[0077]	<400> 5

- [0078] nnnnnnnnnn nnnnnnnnnn guuuuagagc ua 32
- [0079] <210> 6
- [0080] <211> 42
- [0081] <212> RNA
- [0082] <213> 人工序列
- [0083] <220>
- [0084] <223> 人工序列的描述: 合成寡核苷酸
- [0085] <220>
- [0086] <221> 修饰的碱基
- [0087] <222> (1) .. (20)
- [0088] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0089] <400> 6
- [0090] nnnnnnnnnn nnnnnnnnnn guuuuagagc uaucuguuu ug 42
- [0091] <210> 7
- [0092] <211> 36
- [0093] <212> RNA
- [0094] <213> 人工序列
- [0095] <220>
- [0096] <223> 人工序列的描述: 合成寡核苷酸
- [0097] <220>
- [0098] <221> 修饰的碱基
- [0099] <222> (1) .. (20)
- [0100] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0101] <400> 7
- [0102] nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcu 36
- [0103] <210> 8
- [0104] <211> 362
- [0105] <212> RNA
- [0106] <213> 人工序列
- [0107] <220>
- [0108] <223> 人工序列的描述: 合成寡核苷酸
- [0109] <220>
- [0110] <221> 修饰的碱基
- [0111] <222> (1) .. (20)
- [0112] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0113] <220>
- [0114] <221> 修饰的碱基
- [0115] <222> (63) .. (362)
- [0116] <223> a、c、u或g、以及该区域可能包括0-300个核苷酸,其中某些位置可缺失

[0117]	<400> 8	
[0118]	nnnnnnnnnnn nnnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
[0119]	cgnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	120
[0120]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	180
[0121]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	240
[0122]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	300
[0123]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	360
[0124]	nn	362
[0125]	<210> 9	
[0126]	<211> 375	
[0127]	<212> RNA	
[0128]	<213> 人工序列	
[0129]	<220>	
[0130]	<223> 人工序列的描述: 合成寡核苷酸	
[0131]	<220>	
[0132]	<221> 修饰的碱基	
[0133]	<222> (1) .. (20)	
[0134]	<223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失	
[0135]	<220>	
[0136]	<221> 修饰的碱基	
[0137]	<222> (76) .. (375)	
[0138]	<223> a、c、u或g、以及该区域可能包括0-300个核苷酸,其中某些位置可缺失	
[0139]	<400> 9	
[0140]	nnnnnnnnnnn nnnnnnnnnnn guuuuagagc uaugcugaaa agcauagcaa guuaaaauaa	60
[0141]	ggcuaguccg uuaucnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	120
[0142]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	180
[0143]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	240
[0144]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	300
[0145]	nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn	360
[0146]	nnnnnnnnnnn nnnnnn	375
[0147]	<210> 10	
[0148]	<211> 387	
[0149]	<212> RNA	
[0150]	<213> 人工序列	
[0151]	<220>	
[0152]	<223> 人工序列的描述: 合成寡核苷酸	
[0153]	<220>	
[0154]	<221> 修饰的碱基	
[0155]	<222> (1) .. (20)	

- [0156] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0157] <220>
- [0158] <221> 修饰的碱基
- [0159] <222> (88) .. (387)
- [0160] <223> a、c、u或g、以及该区域可能包括0-300个核苷酸,其中某些位置可缺失
- [0161] <400> 10
- [0162] nnnnnnnnnnnn nnnnnnnnnnnn guuuuagagc uaugcuguuu uggaaacaaa acagcauagc 60
- [0163] aaguuaaaaau aaggcuaguc cguuaucnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 120
- [0164] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 180
- [0165] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 240
- [0166] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 300
- [0167] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 360
- [0168] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnn 387
- [0169] <210> 11
- [0170] <211> 396
- [0171] <212> RNA
- [0172] <213> 人工序列
- [0173] <220>
- [0174] <223> 人工序列的描述: 合成寡核苷酸
- [0175] <220>
- [0176] <221> 修饰的碱基
- [0177] <222> (1) .. (20)
- [0178] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0179] <220>
- [0180] <221> 修饰的碱基
- [0181] <222> (97) .. (396)
- [0182] <223> a、c、u或g、以及该区域可能包括0-300个核苷酸,其中某些位置可缺失
- [0183] <400> 11
- [0184] nnnnnnnnnnnn nnnnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaaau aaggcuaguc 60
- [0185] cguuaucaac uugaaaaagu ggcaccgagu cggugcnnnn nnnnnnnnnnnn nnnnnnnnnnnn 120
- [0186] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 180
- [0187] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 240
- [0188] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 300
- [0189] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 360
- [0190] nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnn 396
- [0191] <210> 12
- [0192] <211> 96
- [0193] <212> RNA
- [0194] <213> 人工序列

- [0195] <220>
- [0196] <223> 人工序列的描述：合成寡核苷酸
- [0197] <220>
- [0198] <221> 修饰的碱基
- [0199] <222> (1) .. (20)
- [0200] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸，其中某些位置可缺失
- [0201] <400> 12
- [0202] nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuuaau aaggcuaguc 60
- [0203] cguuaucaac uugaaaaagu ggcaccgagu cggugc 96
- [0204] <210> 13
- [0205] <211> 106
- [0206] <212> RNA
- [0207] <213> 人工序列
- [0208] <220>
- [0209] <223> 人工序列的描述：合成寡核苷酸
- [0210] <220>
- [0211] <221> 修饰的碱基
- [0212] <222> (1) .. (20)
- [0213] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸，其中某些位置可缺失
- [0214] <400> 13
- [0215] nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuggaa acagcauagc aaguuuaau 60
- [0216] aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106
- [0217] <210> 14
- [0218] <211> 106
- [0219] <212> RNA
- [0220] <213> 人工序列
- [0221] <220>
- [0222] <223> 人工序列的描述：合成寡核苷酸
- [0223] <220>
- [0224] <221> 修饰的碱基
- [0225] <222> (1) .. (20)
- [0226] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸，其中某些位置可缺失
- [0227] <400> 14
- [0228] nnnnnnnnnn nnnnnnnnnn guuuuagagc uaugcuggaa acagcauagc aaguuuaau 60
- [0229] aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugc 106
- [0230] <210> 15
- [0231] <211> 36
- [0232] <212> DNA
- [0233] <213> 人工序列

- [0234] <220>
- [0235] <223> 人工序列的描述: 合成寡核苷酸
- [0236] <220>
- [0237] <223>组合DNA/RNA 分子 (Combined DNA/RNA Molecule) 的描述: 合成寡核苷酸
- [0238] <220>
- [0239] <221> 修饰的碱基
- [0240] <222> (1) .. (20)
- [0241] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0242] <400> 15
- [0243] nnnnnnnnnn nnnnnnnnnn gtttttagagc tagaaa 36
- [0244] <210> 16
- [0245] <211> 60
- [0246] <212> DNA
- [0247] <213> 人工序列
- [0248] <220>
- [0249] <223> 人工序列的描述: 合成寡核苷酸
- [0250] <400> 16
- [0251] tagcaagtta aaataaggct agtccgttat caacttgaaa aagtggcacc gagtcggtgc 60
- [0252] <210> 17
- [0253] <211> 5
- [0254] <212> PRT
- [0255] <213> 人工序列
- [0256] <220>
- [0257] <223> 人工序列的描述: 合成肽
- [0258] <400> 17
- [0259] Gly Gly Gly Gly Ser
- [0260] 1 5
- [0261] <210> 18
- [0262] <211> 13
- [0263] <212> PRT
- [0264] <213> 人工序列
- [0265] <220>
- [0266] <223> 人工序列的描述: 合成肽
- [0267] <400> 18
- [0268] Asp Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu
- [0269] 1 5 10
- [0270] <210> 19
- [0271] <211> 17
- [0272] <212> PRT

[0273]	<213> 人工序列															
[0274]	<220>															
[0275]	<223> 人工序列的描述: 合成肽															
[0276]	<400> 19															
[0277]	Pro	Lys	Lys	Lys	Arg	Lys	Val	Glu	Asp	Pro	Lys	Lys	Lys	Arg	Lys	Val
[0278]	1				5					10					15	
[0279]	Asp															
[0280]	<210> 20															
[0281]	<211> 1401															
[0282]	<212> PRT															
[0283]	<213> 人工序列															
[0284]	<220>															
[0285]	<223> 人工序列的描述: 合成多肽															
[0286]	<400> 20															
[0287]	Met	Asp	Lys	Lys	Tyr	Ser	Ile	Gly	Leu	Asp	Ile	Gly	Thr	Asn	Ser	Val
[0288]	1				5					10					15	
[0289]	Gly	Trp	Ala	Val	Ile	Thr	Asp	Glu	Tyr	Lys	Val	Pro	Ser	Lys	Lys	Phe
[0290]					20					25					30	
[0291]	Lys	Val	Leu	Gly	Asn	Thr	Asp	Arg	His	Ser	Ile	Lys	Lys	Asn	Leu	Ile
[0292]					35					40					45	
[0293]	Gly	Ala	Leu	Leu	Phe	Asp	Ser	Gly	Glu	Thr	Ala	Glu	Ala	Thr	Arg	Leu
[0294]					50					55					60	
[0295]	Lys	Arg	Thr	Ala	Arg	Arg	Arg	Tyr	Thr	Arg	Arg	Lys	Asn	Arg	Ile	Cys
[0296]	65					70					75					80
[0297]	Tyr	Leu	Gln	Glu	Ile	Phe	Ser	Asn	Glu	Met	Ala	Lys	Val	Asp	Asp	Ser
[0298]						85					90					95
[0299]	Phe	Phe	His	Arg	Leu	Glu	Glu	Ser	Phe	Leu	Val	Glu	Glu	Asp	Lys	Lys
[0300]						100					105					110
[0301]	His	Glu	Arg	His	Pro	Ile	Phe	Gly	Asn	Ile	Val	Asp	Glu	Val	Ala	Tyr
[0302]						115					120					125
[0303]	His	Glu	Lys	Tyr	Pro	Thr	Ile	Tyr	His	Leu	Arg	Lys	Lys	Leu	Val	Asp
[0304]						130					135					140
[0305]	Ser	Thr	Asp	Lys	Ala	Asp	Leu	Arg	Leu	Ile	Tyr	Leu	Ala	Leu	Ala	His
[0306]	145						150					155				160
[0307]	Met	Ile	Lys	Phe	Arg	Gly	His	Phe	Leu	Ile	Glu	Gly	Asp	Leu	Asn	Pro
[0308]						165						170				175
[0309]	Asp	Asn	Ser	Asp	Val	Asp	Lys	Leu	Phe	Ile	Gln	Leu	Val	Gln	Thr	Tyr
[0310]						180						185				190
[0311]	Asn	Gln	Leu	Phe	Glu	Glu	Asn	Pro	Ile	Asn	Ala	Ser	Gly	Val	Asp	Ala

[0312]	195	200	205
[0313]	Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn		
[0314]	210	215	220
[0315]	Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn		
[0316]	225	230	235
[0317]	Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe		
[0318]	245	250	255
[0319]	Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp		
[0320]	260	265	270
[0321]	Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp		
[0322]	275	280	285
[0323]	Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp		
[0324]	290	295	300
[0325]	Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser		
[0326]	305	310	315
[0327]	Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys		
[0328]	325	330	335
[0329]	Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe		
[0330]	340	345	350
[0331]	Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser		
[0332]	355	360	365
[0333]	Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp		
[0334]	370	375	380
[0335]	Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg		
[0336]	385	390	395
[0337]	Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu		
[0338]	405	410	415
[0339]	Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe		
[0340]	420	425	430
[0341]	Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile		
[0342]	435	440	445
[0343]	Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp		
[0344]	450	455	460
[0345]	Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu		
[0346]	465	470	475
[0347]	Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr		
[0348]	485	490	495
[0349]	Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser		
[0350]	500	505	510

[0351]	Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Lys
[0352]				515				520					525			
[0353]	Tyr	Val	Thr	Glu	Gly	Met	Arg	Lys	Pro	Ala	Phe	Leu	Ser	Gly	Glu	Gln
[0354]			530					535					540			
[0355]	Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr
[0356]	545					550					555					560
[0357]	Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp
[0358]					565					570					575	
[0359]	Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly
[0360]				580					585					590		
[0361]	Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp
[0362]			595					600						605		
[0363]	Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr
[0364]		610					615					620				
[0365]	Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala
[0366]	625					630					635					640
[0367]	His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr
[0368]				645						650					655	
[0369]	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp
[0370]			660						665					670		
[0371]	Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe
[0372]			675					680						685		
[0373]	Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe
[0374]		690					695					700				
[0375]	Lys	Glu	Asp	Ile	Gln	Lys	Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu
[0376]	705					710					715					720
[0377]	His	Glu	His	Ile	Ala	Asn	Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly
[0378]				725						730				735		
[0379]	Ile	Leu	Gln	Thr	Val	Lys	Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly
[0380]				740					745					750		
[0381]	Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln
[0382]			755						760					765		
[0383]	Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile
[0384]		770					775						780			
[0385]	Glu	Glu	Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro
[0386]	785					790					795					800
[0387]	Val	Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu
[0388]				805					810					815		
[0389]	Gln	Asn	Gly	Arg	Asp	Met	Tyr	Val	Asp	Gln	Glu	Leu	Asp	Ile	Asn	Arg

[0390]	820	825	830
[0391]	Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys		
[0392]	835	840	845
[0393]	Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg		
[0394]	850	855	860
[0395]	Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys		
[0396]	865	870	875
[0397]	Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys		
[0398]	885	890	895
[0399]	Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp		
[0400]	900	905	910
[0401]	Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr		
[0402]	915	920	925
[0403]	Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp		
[0404]	930	935	940
[0405]	Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser		
[0406]	945	950	955
[0407]	Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg		
[0408]	965	970	975
[0409]	Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val		
[0410]	980	985	990
[0411]	Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe		
[0412]	995	1000	1005
[0413]	Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala		
[0414]	1010	1015	1020
[0415]	Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe		
[0416]	1025	1030	1035
[0417]	Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala		
[0418]	1040	1045	1050
[0419]	Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu		
[0420]	1055	1060	1065
[0421]	Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val		
[0422]	1070	1075	1080
[0423]	Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr		
[0424]	1085	1090	1095
[0425]	Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys		
[0426]	1100	1105	1110
[0427]	Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro		
[0428]	1115	1120	1125

[0429]	Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val		
[0430]	1130	1135	1140
[0431]	Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys		
[0432]	1145	1150	1155
[0433]	Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser		
[0434]	1160	1165	1170
[0435]	Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys		
[0436]	1175	1180	1185
[0437]	Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu		
[0438]	1190	1195	1200
[0439]	Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly		
[0440]	1205	1210	1215
[0441]	Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val		
[0442]	1220	1225	1230
[0443]	Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser		
[0444]	1235	1240	1245
[0445]	Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys		
[0446]	1250	1255	1260
[0447]	His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys		
[0448]	1265	1270	1275
[0449]	Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala		
[0450]	1280	1285	1290
[0451]	Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn		
[0452]	1295	1300	1305
[0453]	Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala		
[0454]	1310	1315	1320
[0455]	Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser		
[0456]	1325	1330	1335
[0457]	Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr		
[0458]	1340	1345	1350
[0459]	Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp		
[0460]	1355	1360	1365
[0461]	Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Gly Ser Gly		
[0462]	1370	1375	1380
[0463]	Ser Pro Lys Lys Lys Arg Lys Val Glu Asp Pro Lys Lys Lys Arg		
[0464]	1385	1390	1395
[0465]	Lys Val Asp		
[0466]	1400		
[0467]	<210> 21		

[0468] <211> 1368
 [0469] <212> PRT
 [0470] <213> 人工序列
 [0471] <220>
 [0472] <223> 人工序列的描述: 合成多肽
 [0473] <400> 21
 [0474] Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
 [0475] 1 5 10 15
 [0476] Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 [0477] 20 25 30
 [0478] Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 [0479] 35 40 45
 [0480] Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 [0481] 50 55 60
 [0482] Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 [0483] 65 70 75 80
 [0484] Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
 [0485] 85 90 95
 [0486] Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
 [0487] 100 105 110
 [0488] His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 [0489] 115 120 125
 [0490] His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
 [0491] 130 135 140
 [0492] Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
 [0493] 145 150 155 160
 [0494] Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
 [0495] 165 170 175
 [0496] Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
 [0497] 180 185 190
 [0498] Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
 [0499] 195 200 205
 [0500] Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
 [0501] 210 215 220
 [0502] Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
 [0503] 225 230 235 240
 [0504] Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
 [0505] 245 250 255
 [0506] Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp

[0507]	260	265	270
[0508]	Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp		
[0509]	275	280	285
[0510]	Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp		
[0511]	290	295	300
[0512]	Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser		
[0513]	305	310	315
[0514]	Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys		
[0515]	325	330	335
[0516]	Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe		
[0517]	340	345	350
[0518]	Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser		
[0519]	355	360	365
[0520]	Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp		
[0521]	370	375	380
[0522]	Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg		
[0523]	385	390	395
[0524]	Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu		
[0525]	405	410	415
[0526]	Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe		
[0527]	420	425	430
[0528]	Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile		
[0529]	435	440	445
[0530]	Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp		
[0531]	450	455	460
[0532]	Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu		
[0533]	465	470	475
[0534]	Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr		
[0535]	485	490	495
[0536]	Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser		
[0537]	500	505	510
[0538]	Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys		
[0539]	515	520	525
[0540]	Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln		
[0541]	530	535	540
[0542]	Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr		
[0543]	545	550	555
[0544]	Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp		
[0545]	565	570	575

[0546]	Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly		
[0547]		580	585
[0548]	Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp		590
[0549]		595	600
[0550]	Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr		605
[0551]		610	615
[0552]	Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala		620
[0553]	625	630	635
[0554]	His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr		640
[0555]		645	650
[0556]	Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp		655
[0557]		660	665
[0558]	Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe		670
[0559]		675	680
[0560]	Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe		685
[0561]		690	695
[0562]	Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu		700
[0563]	705	710	715
[0564]	His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly		720
[0565]		725	730
[0566]	Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly		735
[0567]		740	745
[0568]	Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln		750
[0569]		755	760
[0570]	Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile		765
[0571]		770	775
[0572]	Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro		780
[0573]	785	790	795
[0574]	Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu		800
[0575]		805	810
[0576]	Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg		815
[0577]		820	825
[0578]	Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys		830
[0579]		835	840
[0580]	Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg		845
[0581]		850	855
[0582]	Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys		860
[0583]	865	870	875
[0584]	Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys		880

[0585]	885							890							895						
[0586]	Phe	Asp	Asn	Leu	Thr	Lys	Ala	Glu	Arg	Gly	Gly	Leu	Ser	Glu	Leu	Asp					
[0587]	900							905							910						
[0588]	Lys	Ala	Gly	Phe	Ile	Lys	Arg	Gln	Leu	Val	Glu	Thr	Arg	Gln	Ile	Thr					
[0589]	915							920							925						
[0590]	Lys	His	Val	Ala	Gln	Ile	Leu	Asp	Ser	Arg	Met	Asn	Thr	Lys	Tyr	Asp					
[0591]	930							935							940						
[0592]	Glu	Asn	Asp	Lys	Leu	Ile	Arg	Glu	Val	Lys	Val	Ile	Thr	Leu	Lys	Ser					
[0593]	945							950							955						
[0594]	Lys	Leu	Val	Ser	Asp	Phe	Arg	Lys	Asp	Phe	Gln	Phe	Tyr	Lys	Val	Arg					
[0595]	965							970							975						
[0596]	Glu	Ile	Asn	Asn	Tyr	His	His	Ala	His	Asp	Ala	Tyr	Leu	Asn	Ala	Val					
[0597]	980							985							990						
[0598]	Val	Gly	Thr	Ala	Leu	Ile	Lys	Lys	Tyr	Pro	Lys	Leu	Glu	Ser	Glu	Phe					
[0599]	995							1000							1005						
[0600]	Val	Tyr	Gly	Asp	Tyr	Lys	Val	Tyr	Asp	Val	Arg	Lys	Met	Ile	Ala						
[0601]	1010							1015							1020						
[0602]	Lys	Ser	Glu	Gln	Glu	Ile	Gly	Lys	Ala	Thr	Ala	Lys	Tyr	Phe	Phe						
[0603]	1025							1030							1035						
[0604]	Tyr	Ser	Asn	Ile	Met	Asn	Phe	Phe	Lys	Thr	Glu	Ile	Thr	Leu	Ala						
[0605]	1040							1045							1050						
[0606]	Asn	Gly	Glu	Ile	Arg	Lys	Arg	Pro	Leu	Ile	Glu	Thr	Asn	Gly	Glu						
[0607]	1055							1060							1065						
[0608]	Thr	Gly	Glu	Ile	Val	Trp	Asp	Lys	Gly	Arg	Asp	Phe	Ala	Thr	Val						
[0609]	1070							1075							1080						
[0610]	Arg	Lys	Val	Leu	Ser	Met	Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr						
[0611]	1085							1090							1095						
[0612]	Glu	Val	Gln	Thr	Gly	Gly	Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys						
[0613]	1100							1105							1110						
[0614]	Arg	Asn	Ser	Asp	Lys	Leu	Ile	Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro						
[0615]	1115							1120							1125						
[0616]	Lys	Lys	Tyr	Gly	Gly	Phe	Asp	Ser	Pro	Thr	Val	Ala	Tyr	Ser	Val						
[0617]	1130							1135							1140						
[0618]	Leu	Val	Val	Ala	Lys	Val	Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys						
[0619]	1145							1150							1155						
[0620]	Ser	Val	Lys	Glu	Leu	Leu	Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser						
[0621]	1160							1165							1170						
[0622]	Phe	Glu	Lys	Asn	Pro	Ile	Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys						
[0623]	1175							1180							1185						

[0624]	Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu
[0625]	1190 1195 1200
[0626]	Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly
[0627]	1205 1210 1215
[0628]	Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val
[0629]	1220 1225 1230
[0630]	Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
[0631]	1235 1240 1245
[0632]	Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys
[0633]	1250 1255 1260
[0634]	His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys
[0635]	1265 1270 1275
[0636]	Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala
[0637]	1280 1285 1290
[0638]	Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn
[0639]	1295 1300 1305
[0640]	Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
[0641]	1310 1315 1320
[0642]	Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
[0643]	1325 1330 1335
[0644]	Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
[0645]	1340 1345 1350
[0646]	Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
[0647]	1355 1360 1365
[0648]	<210> 22
[0649]	<211> 4
[0650]	<212> PRT
[0651]	<213> 人工序列
[0652]	<220>
[0653]	<223> 人工序列的描述: 合成肽
[0654]	<400> 22
[0655]	Gly Gly Ser Gly
[0656]	1
[0657]	<210> 23
[0658]	<211> 12
[0659]	<212> PRT
[0660]	<213> 人工序列
[0661]	<220>
[0662]	<223> 人工序列的描述: 合成肽

[0663] <400> 23
[0664] Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala
[0665] 1 5 10
[0666] <210> 24
[0667] <211> 16
[0668] <212> PRT
[0669] <213> 人工序列
[0670] <220>
[0671] <223> 人工序列的描述: 合成肽
[0672] <400> 24
[0673] Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Ser
[0674] 1 5 10 15
[0675] <210> 25
[0676] <211> 21
[0677] <212> PRT
[0678] <213> 人工序列
[0679] <220>
[0680] <223> 人工序列的描述: 合成肽
[0681] <400> 25
[0682] Ser Gly Ser Glu Thr Pro Gly Thr Ser Glu Ser Ala Thr Pro Glu Gly
[0683] 1 5 10 15
[0684] Gly Ser Gly Gly Ser
[0685] 20
[0686] <210> 26
[0687] <211> 24
[0688] <212> DNA
[0689] <213> 人工序列
[0690] <220>
[0691] <223> 人工序列的描述: 合成寡核苷酸
[0692] <400> 26
[0693] gccgaggtga agttcgaggg cgac 24
[0694] <210> 27
[0695] <211> 23
[0696] <212> DNA
[0697] <213> 人工序列
[0698] <220>
[0699] <223> 人工序列的描述: 合成寡核苷酸
[0700] <400> 27
[0701] cctacggcgt gcagtgccttc agc 23

[0702]	<210> 28
[0703]	<211> 332
[0704]	<212> DNA
[0705]	<213> 人工序列
[0706]	<220>
[0707]	<223> 人工序列的描述: 合成寡核苷酸
[0708]	<400> 28
[0709]	acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc acctacggca 60
[0710]	agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg cccaccctcg 120
[0711]	tgaccaccct gacctacggc gtgcagtgtc tcagccgcta ccccgaccac atgaagcagc 180
[0712]	acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc atcttcttca 240
[0713]	aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac accctggtga 300
[0714]	accgcatcga gctgaagggc atcgacttca ag 332
[0715]	<210> 29
[0716]	<211> 295
[0717]	<212> DNA
[0718]	<213> 人工序列
[0719]	<220>
[0720]	<223> 人工序列的描述: 合成寡核苷酸
[0721]	<400> 29
[0722]	gtccggcgag ggcgagggcg atgccaccta cggcaagctg accctgaagt tcattctgcac 60
[0723]	caccggcaag ctgcccgtgc cctggcccac cctcgtgacc accctgacct acggcgtgca 120
[0724]	gtgcttcagc cgctaccccg accacatgaa gcagcacgac ttcttcaagt ccgccatgcc 180
[0725]	cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact acaagacccg 240
[0726]	cgtcgagggc gacaccctgg tgaaccgcat cgagctgaag ggcacgcact tcaag 295
[0727]	<210> 30
[0728]	<211> 158
[0729]	<212> DNA
[0730]	<213> 人工序列
[0731]	<220>
[0732]	<223> 人工序列的描述: 合成寡核苷酸
[0733]	<400> 30
[0734]	tccggcgagg gcgagggcga tgccacctac ggcaagctga ccctgaagtt catctgcacc 60
[0735]	accggcaagc tgcccgtgcc ctggcccacc ctcgtgacca ccctgacgag ggcgacaccc 120
[0736]	tggtgaaccg catcgagctg aagggcacgc acttcaag 158
[0737]	<210> 31
[0738]	<211> 285
[0739]	<212> DNA
[0740]	<213> 人工序列

[0741] <220>
[0742] <223> 人工序列的描述: 合成寡核苷酸
[0743] <400> 31
[0744] tccggcgagg gcgaggcgga tgccacctac ggcaagctga ccctgaagtt catctgcacc 60
[0745] accggcaagc tgcccgtgcc ctggcccacc ctctgtaaca ccctgacctc cgtgcagtgc 120
[0746] ttcagccgct accccgacca catgaagcag cactcttct tcaagtccgc catgcccga 180
[0747] ggctacgtcc aggagcgcac catcttcttc aaggacgac gcaactacaa gttcgagggc 240
[0748] gacaccctgg tgaaccgcat cgagctgaag ggcatcgact tcaag 285
[0749] <210> 32
[0750] <211> 301
[0751] <212> DNA
[0752] <213> 人工序列
[0753] <220>
[0754] <223> 人工序列的描述: 合成寡核苷酸
[0755] <400> 32
[0756] ccggcgaggg cgagggcgat gccacctacg gcaagctgac cctgaagttc atctgcacca 60
[0757] ccggcaagct gcccgtgccc tggcccaccc tctgtaccac cctgacctac gtgcagtgtc 120
[0758] tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 180
[0759] gctacgtcca ggagcgacc atcttcttca aggacgacgg caactacaag acccgcgccg 240
[0760] aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 300
[0761] a 301
[0762] <210> 33
[0763] <211> 156
[0764] <212> DNA
[0765] <213> 人工序列
[0766] <220>
[0767] <223> 人工序列的描述: 合成寡核苷酸
[0768] <400> 33
[0769] tccggcgagg gcgaggcgga tgccacctac ggcaagctga ccctgaagtt catctgcacc 60
[0770] accggcaagc tgcccgtgcc ctggcccacc ctctgacca ccctgacctc cggacaccct 120
[0771] ggtgaaccgc atcgagctga agggcatcga cttcaa 156
[0772] <210> 34
[0773] <211> 197
[0774] <212> DNA
[0775] <213> 人工序列
[0776] <220>
[0777] <223> 人工序列的描述: 合成寡核苷酸
[0778] <400> 34
[0779] gtaaacggcc acaagttcag cgtgcagtgc ttcagccgct accccgacca catgaagcag 60

[0780] cacgacttct tcaagtccgc catgcccgaaggctacgtcc aggagcgcac catctttcttc 120
[0781] aaggacgacg gcaactacaa gacggcaact acaagttcga gggcgacacc ctggtgaacc 180
[0782] gcatcgagct gaagggc 197
[0783] <210> 35
[0784] <211> 283
[0785] <212> DNA
[0786] <213> 人工序列
[0787] <220>
[0788] <223> 人工序列的描述: 合成寡核苷酸
[0789] <400> 35
[0790] ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc acctacggca agctgaccct 60
[0791] gaagttcattc tgcaccaccg gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct 120
[0792] gacctacggc gtgcagtgtc tcagccgcta ccccgaccac atgaagcagc acgactttctt 180
[0793] caagtccgcc atgcccgaag gctacgtcca ggagcgcacc atctttcttca aggacgacgg 240
[0794] caactacaag acaccgcatt gagctgaagg gcatcgactt caa 283
[0795] <210> 36
[0796] <211> 175
[0797] <212> DNA
[0798] <213> 人工序列
[0799] <220>
[0800] <223> 人工序列的描述: 合成寡核苷酸
[0801] <400> 36
[0802] ggccacaagt tcagcgtgtc cggcgagggc gagggcgatg ccacctacgg caagctgacc 60
[0803] ctgaagttca tctgcaccac cggcaccatt aaggacgacg gcaactacaa gacccgcgcc 120
[0804] gaagttcgag ggcgacaccc tggtgaaccg catcgagctg aagggcattcg acttc 175
[0805] <210> 37
[0806] <211> 291
[0807] <212> DNA
[0808] <213> 人工序列
[0809] <220>
[0810] <223> 人工序列的描述: 合成寡核苷酸
[0811] <400> 37
[0812] taaacggcca caagttcagc gtgtccggcg agggcgaggg cgatgccacc tacggcaagc 60
[0813] tgaccctgaa gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga 120
[0814] ccaccctgac ctacggcgtg cagtgtctca gccgtaccc cgaccacatg aagcagcacg 180
[0815] actttcttcaa gtccgccatg cccgaaggct acgtccagga gcgcaccatt ttcttcaagg 240
[0816] acgacggcaa ctacaagacc tggtgaaccg catcgagctg aagggcattcg a 291
[0817] <210> 38
[0818] <211> 314

[0819] <212> DNA
[0820] <213> 人工序列
[0821] <220>
[0822] <223> 人工序列的描述: 合成寡核苷酸
[0823] <400> 38
[0824] cacaagttca gcgtgtccgg cgagggcgag ggcgatgcca cctacggcaa gctgaccctg 60
[0825] aagttcatct gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg 120
[0826] acctacggcg tgcagtgttt cagccgctac cccgaccaca tgaagcagca cgactttctt 180
[0827] aagtccgcca tgcccgaagg ctacgtccag gagegcacca tttttttcaa ggacgacggc 240
[0828] aactacaaga cccgcgccga ggttcgaggg cgacaccctg gtgaaccgca tcgagctgaa 300
[0829] gggcatcgac ttca 314
[0830] <210> 39
[0831] <211> 316
[0832] <212> DNA
[0833] <213> 人工序列
[0834] <220>
[0835] <223> 人工序列的描述: 合成寡核苷酸
[0836] <400> 39
[0837] taaacggcca caagttcagc gtgtccggcg agggcgaggg cgatgccacc tacggcaagc 60
[0838] tgaccctgaa gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga 120
[0839] ccaccctgac ctacggcggtg cagtgtttca gccgctaccc cgaccacatg aagcagcacg 180
[0840] actttttcaa gtccgccatg cccgaaggct acgtccagga gcgcaccatc ttcttcaagg 240
[0841] acgacggcaa ctacaagacc cgcgcggagg gcgacaccct ggtgaaccgc atcgagctga 300
[0842] agggcatcga cttcaa 316
[0843] <210> 40
[0844] <211> 318
[0845] <212> DNA
[0846] <213> 人工序列
[0847] <220>
[0848] <223> 人工序列的描述: 合成寡核苷酸
[0849] <400> 40
[0850] aaacggccac aagttcagcg tgtccggcga gggcgagggc gatgccacct acggcaagct 60
[0851] gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca ccctcgtgac 120
[0852] caccctgacc tacgtgcagt gtttcagccg ctaccccgac cacatgaagc agcacgactt 180
[0853] cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 240
[0854] cggcaactac aagaccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 300
[0855] cgagctgaag ggcacga 318
[0856] <210> 41
[0857] <211> 185

- [0858] <212> DNA
- [0859] <213> 人工序列
- [0860] <220>
- [0861] <223> 人工序列的描述: 合成寡核苷酸
- [0862] <400> 41
- [0863] acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac ggcaagctga 60
- [0864] ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc ctctgtgacca 120
- [0865] ccctgacctt cgtgaagtgc gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca 180
- [0866] tcgac 185
- [0867] <210> 42
- [0868] <211> 180
- [0869] <212> DNA
- [0870] <213> 人工序列
- [0871] <220>
- [0872] <223> 人工序列的描述: 合成寡核苷酸
- [0873] <400> 42
- [0874] aaacggccac aagttcagcg tgtccggcga gggcgagggc gatgccacct acggcaagct 60
- [0875] gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca ccctctgtgac 120
- [0876] caccctgacc tacgtgaagt tcgagggcga caccctggtg aaccgcatcg agctgaaggg 180
- [0877] <210> 43
- [0878] <211> 242
- [0879] <212> DNA
- [0880] <213> 人工序列
- [0881] <220>
- [0882] <223> 人工序列的描述: 合成寡核苷酸
- [0883] <220>
- [0884] <221> 修饰的碱基
- [0885] <222> (132) .. (132)
- [0886] <223> a、c、t、g未知或其他
- [0887] <220>
- [0888] <221> 修饰的碱基
- [0889] <222> (153) .. (153)
- [0890] <223> a、c、t、g未知或其他
- [0891] <220>
- [0892] <221> 修饰的碱基
- [0893] <222> (174) .. (174)
- [0894] <223> a、c、t、g未知或其他
- [0895] <400> 43
- [0896] aaacggccac aagttcagcg tgtccggcga gggcgagggc gatgccacct acggcaagct 60

[0897] gaccctgaag ttcatctgca ccaccggcaa gctgcccgtg ccctggccca cccttccgcc 120
[0898] atgcccgaag gntacgtcca ggagcgcacc atntttcttca aggacgacgg caantacaag 180
[0899] acccgcgccg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga agggcatcga 240
[0900] ct 242
[0901] <210> 44
[0902] <211> 175
[0903] <212> DNA
[0904] <213> 人工序列
[0905] <220>
[0906] <223> 人工序列的描述: 合成寡核苷酸
[0907] <400> 44
[0908] caccctgacc tacggcgtgc agtgcttcag ccgtacccc gaccacatga agcagcacga 60
[0909] cttcttcaag tccgcatgc ccgaaggcta cgtccaggag cgcaccatct tcttcaagga 120
[0910] cgacggcaac tacaagaccc gcgccgaggt gaagtctgag ggcgacaccc tgggtg 175
[0911] <210> 45
[0912] <211> 39
[0913] <212> DNA
[0914] <213> 人工序列
[0915] <220>
[0916] <223> 人工序列的描述: 合成寡核苷酸
[0917] <400> 45
[0918] caccctgacc tacgtgaagt tcgagggcga caccctggt 39
[0919] <210> 46
[0920] <211> 161
[0921] <212> DNA
[0922] <213> 人工序列
[0923] <220>
[0924] <223> 人工序列的描述: 合成寡核苷酸
[0925] <400> 46
[0926] caccctgacc tacggcgtgc agtgcttcag ccgtacccc gaccacatga agcagcacga 60
[0927] cttcttcaag tccgcatgc ccgaaggcta cgtccaggag cgcaccatct tcttcaagga 120
[0928] cgacggcaac tacaagaccc gcgccgaggg cgacaccctg g 161
[0929] <210> 47
[0930] <211> 145
[0931] <212> DNA
[0932] <213> 人工序列
[0933] <220>
[0934] <223> 人工序列的描述: 合成寡核苷酸
[0935] <400> 47

- [0936] caccctgacc tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc 60
[0937] atgcccgaag gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag 120
[0938] acccgcgccg aggcgacacc ctggt 145
[0939] <210> 48
[0940] <211> 153
[0941] <212> DNA
[0942] <213> 人工序列
[0943] <220>
[0944] <223> 人工序列的描述: 合成寡核苷酸
[0945] <400> 48
[0946] ttattataca tcggagccct gccaaaaaat caatgtgaag caaatcgcag cccgcctcct 60
[0947] gcctccgctc tactcaactgg tgttcatctt tggttttgtg ggcaacatgc tggtcaccc 120
[0948] catcctgata aactgcaaaa ggctgaagag cat 153
[0949] <210> 49
[0950] <211> 113
[0951] <212> DNA
[0952] <213> 人工序列
[0953] <220>
[0954] <223> 人工序列的描述: 合成寡核苷酸
[0955] <400> 49
[0956] ttattataca tcggagccct gccaaaaaat caatgtgaag caaatcgcag cccgcctccg 60
[0957] ctctactcac tgggtgttcat ctttggtttt gtgggcaaca tgctgggtcat cct 113
[0958] <210> 50
[0959] <211> 70
[0960] <212> DNA
[0961] <213> 人工序列
[0962] <220>
[0963] <223> 人工序列的描述: 合成寡核苷酸
[0964] <400> 50
[0965] ttattataca tcggagccct gccaaaaaat caatgtgaag caaatcgcag cccgcatgct 60
[0966] ggtcatcctc 70
[0967] <210> 51
[0968] <211> 62
[0969] <212> RNA
[0970] <213> 人工序列
[0971] <220>
[0972] <223> 人工序列的描述: 合成寡核苷酸
[0973] <220>
[0974] <221> 修饰的碱基

- [0975] <222> (1) .. (20)
- [0976] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0977] <400> 51
- [0978] nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaaau aaggcuaguc 60
- [0979] cg 62
- [0980] <210> 52
- [0981] <211> 54
- [0982] <212> RNA
- [0983] <213> 人工序列
- [0984] <220>
- [0985] <223> 人工序列的描述: 合成寡核苷酸
- [0986] <220>
- [0987] <221> 修饰的碱基
- [0988] <222> (1) .. (20)
- [0989] <223> a、c、u或g、以及该区域可能包括17-20个核苷酸,其中某些位置可缺失
- [0990] <400> 52
- [0991] nnnnnnnnnn nnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaaau aagg 54
- [0992] <210> 53
- [0993] <211> 28
- [0994] <212> RNA
- [0995] <213> 人工序列
- [0996] <220>
- [0997] <223> 人工序列的描述: 合成寡核苷酸
- [0998] <400> 53
- [0999] guucacugcc guauaggcag cuaagaaa 28
- [1000] <210> 54
- [1001] <211> 20
- [1002] <212> RNA
- [1003] <213> 人工序列
- [1004] <220>
- [1005] <223> 人工序列的描述: 合成寡核苷酸
- [1006] <400> 54
- [1007] guucacugcc guauaggcag 20

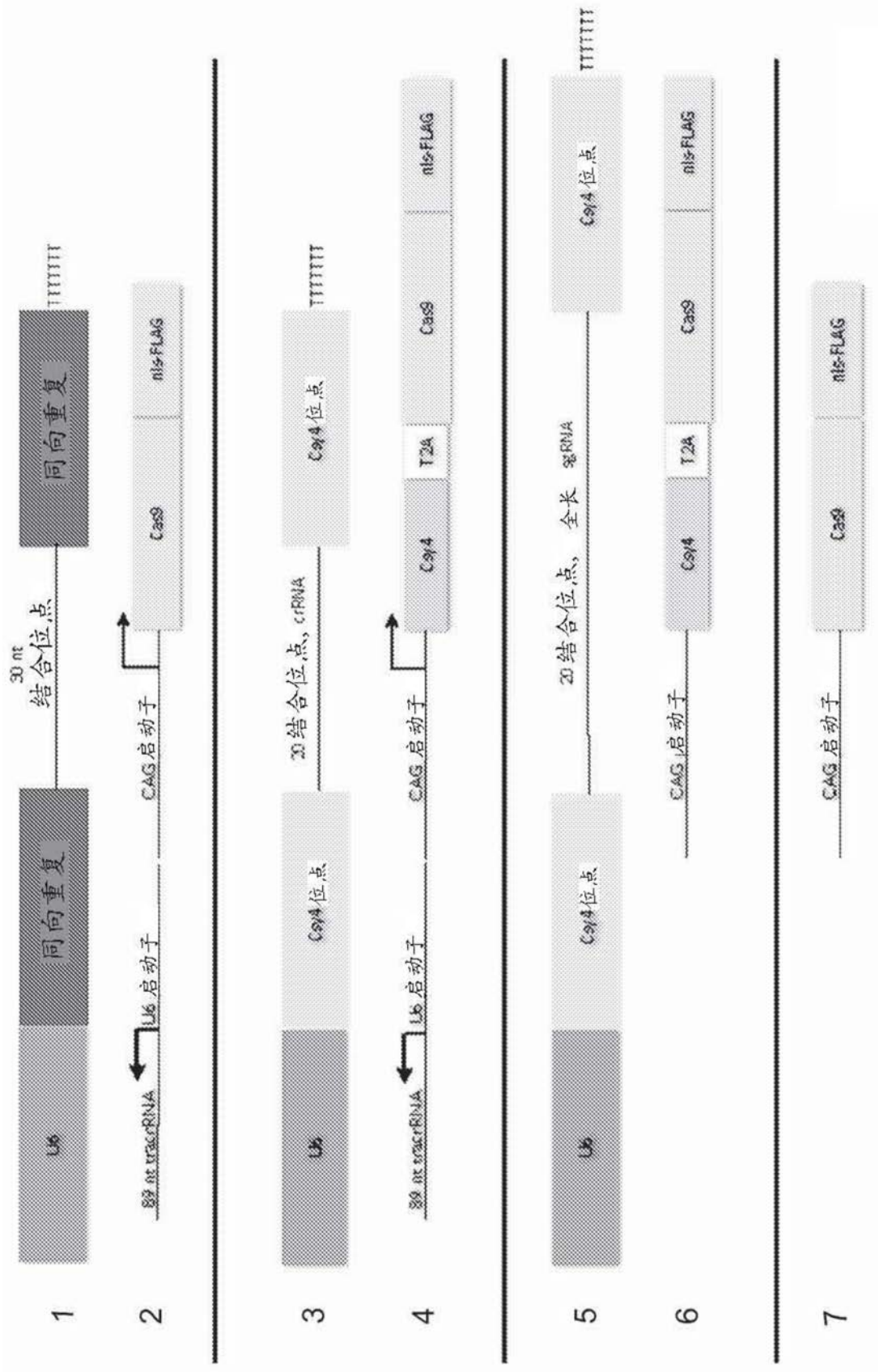


图1

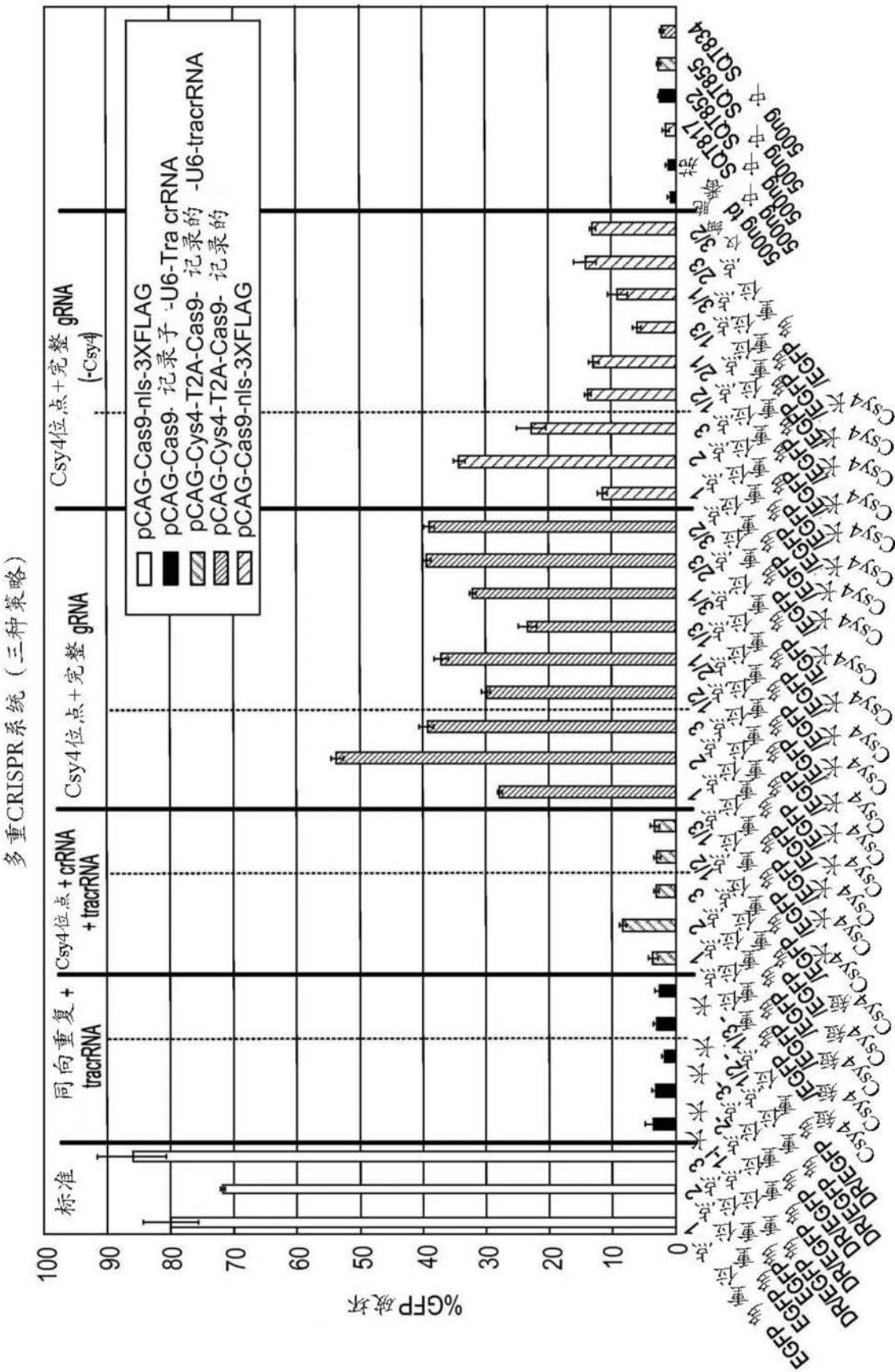


图2

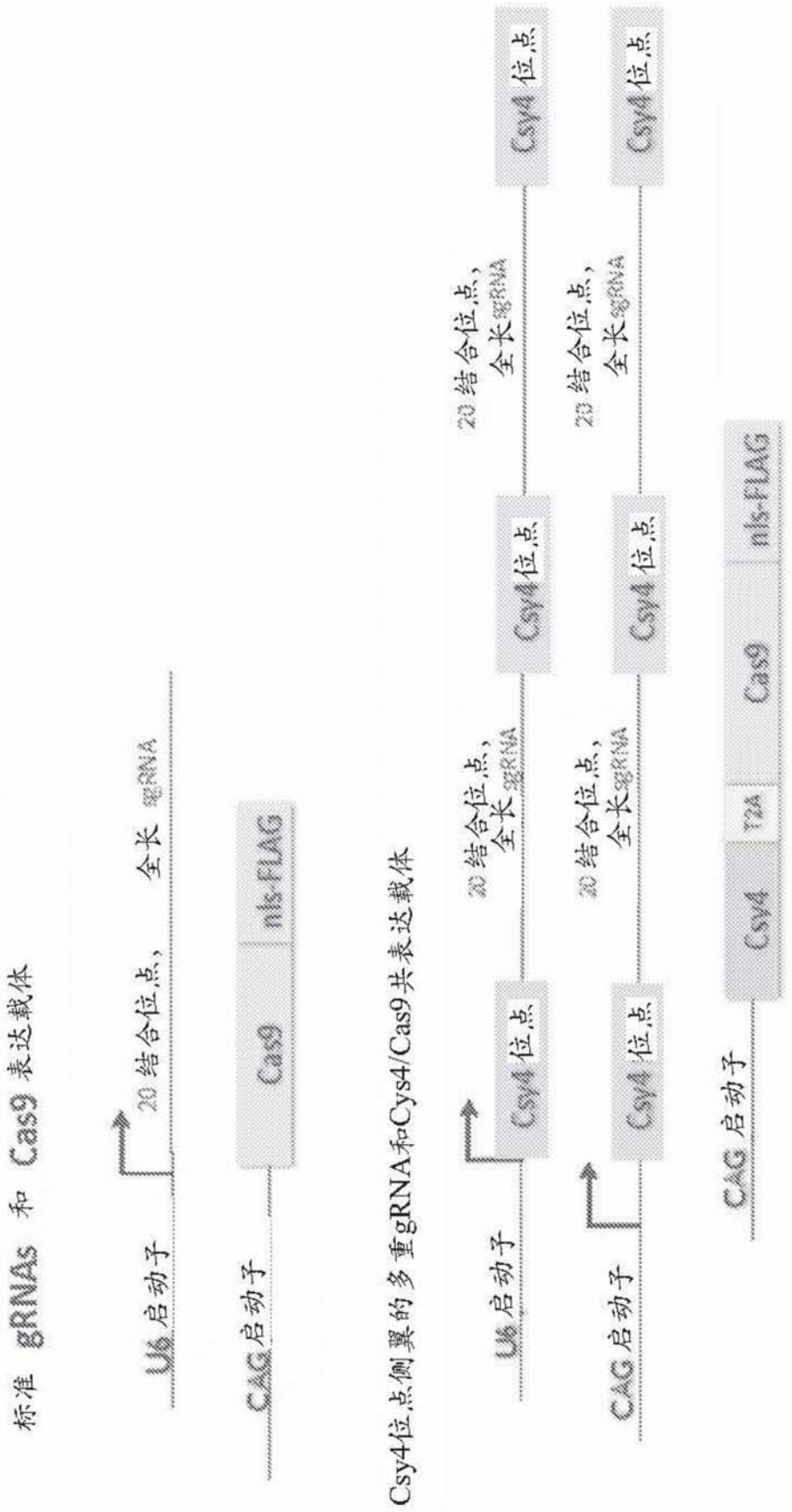


图3

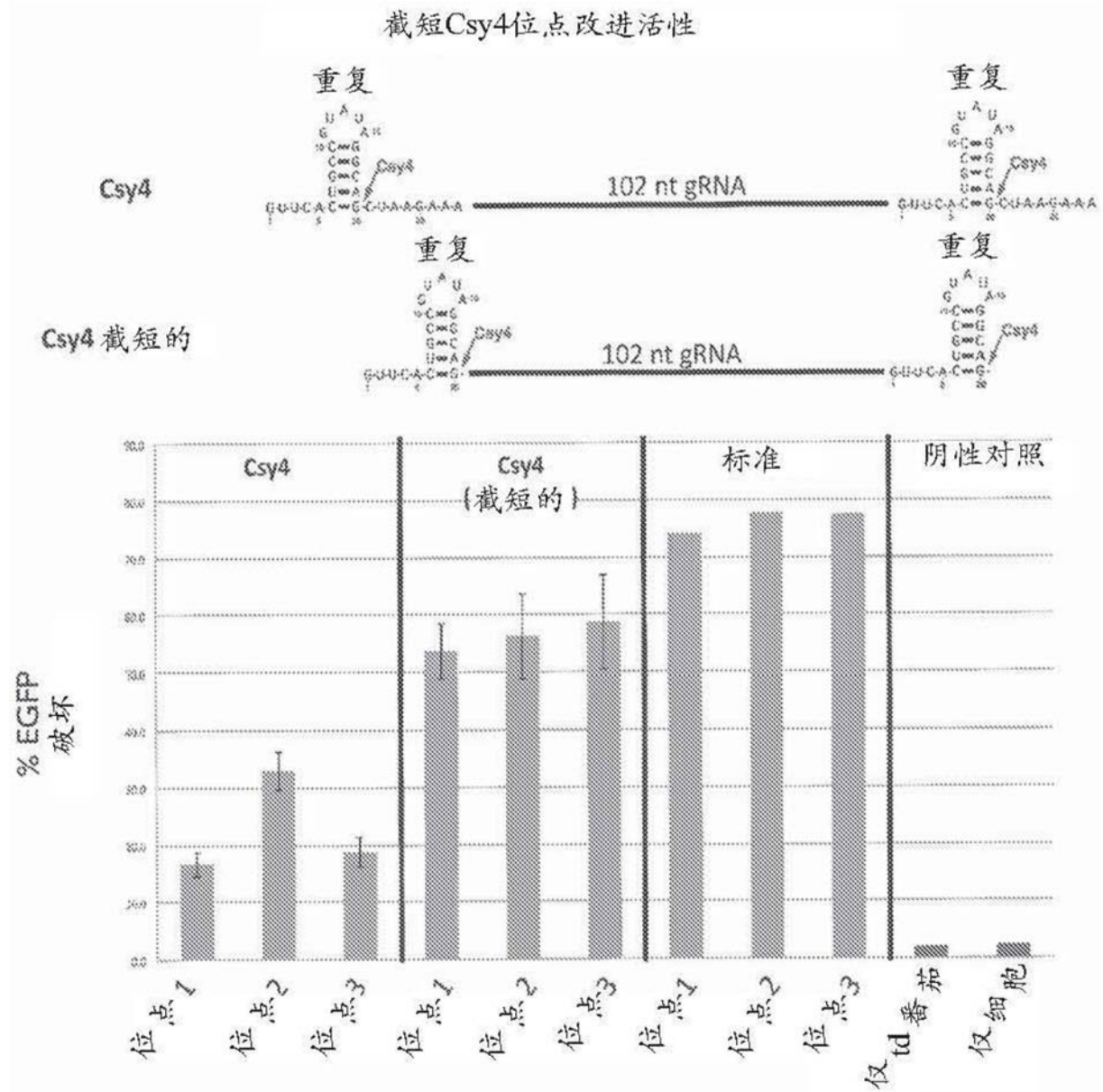


图4

位点 2 GCCGAGGTGAAGTTCGAGGGCGAC
 位点 3 CCTACGGCGTGCAGTGCTTCAGC
 (位点S3/2)

```

ACGTAAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
A1          GTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
A2          TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
A12         TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
B12         CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
C3          TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
C4GTAAACGGCCACAAGTTCAGCGTG-----
C8          CCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
D2          GGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
D4  TAAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
D12         CACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
E5  TAAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
G1  AAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
G7   ACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
G9  AAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT
G12 AAACGGCCACAAGTTCAGCGTGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT

GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGAC-----
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGAACACCCTGACCTACG--
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACG--
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGG--
-----
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAC-----
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGC
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACG--
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACG--
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACG--
GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCT-----

```

图5A

GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG

--TGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
-TGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG

---CAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG

GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG
-TGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAG

-----TCCGCCATGCCCCGAAG

GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCG-----T

GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG-----TT
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT

GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC-----GGCAACTACAAGTT
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC-----
-----CATCAAGGACGACGGCAACTACAAGACCCGCGCCGA-----AGTT
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACC-----
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGT-----T
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGG-----
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
-----TGAAGTT
-----TGAAGTT
GNTACGTCCAGGAGCGCACCATNTTCTTCAAGGACGACGGCAANTACAAGACCCGCGCCGA-----AGTT

图5B

CGAGGCGCACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG	SEQ ID NO:
-GAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA	SEQ ID NO:
-----ACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA	SEQ ID NO:
-----ACCGCATCGAGCTGAAGGGCATCGACTTCAA	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCA	SEQ ID NO:
-----TGGTGAACCGCATCGAGCTGAAGGGCATCGA	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCA	SEQ ID NO:
-----GCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAA	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGA	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGG	SEQ ID NO:
CGAGGCGCACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACT	SEQ ID NO:

图5C

eGFP 位点 3

CACCTGACCTACGGCGTGCAGTCTTCAGCCGTACCCGACCACATGAAGCAGCAGCAGCTCTTCAAGTCCGCAATGCCGGAAGCTACGTCCAGGAGCGCACCA
 CACCTGACCTAGC-----
 CACCTGACCTACGGCGTGCAGTCTTCAGCCGTACCCGACCACATGAAGCAGCAGCAGCTCTTCAAGTCCGCAATGCCGGAAGCTACGTCCAGGAGCGCACCA
 CACCTGACCT-----CAGCCGTACCCGACCACATGAAGCAGCAGCAGCTCTTCAAGTCCGCAATGCCGGAAGCTACGTCCAGGAGCGCACCA

eGFP 位点 2

TCCTTCTCAAGGACGACGGCAACTACAAGACCCGCGGAGGTGAAGTTCGAGGGCGACACCCCTGGTG
 -----TGAAGTTCGAGGGCGACACCCCTGGT
 TCCTTCTCAAGGACGACGGCAACTACAAGACCCGCGGAGG-----GCGACACCCCTGG
 TCCTTCTCAAGGACGACGGCAACTACAAGACCCGCGGAGG-----CGACACCCCTGGT

CCR5

TTATTATACATCGGAGCCCTGCCAAATAATCAATGTGAAGCAAAATCGCAGCCCGCCTCCTGCTCTACTCACTGGTGTTCATCTTTGTGGCAACA
 TTATTATACATCGGAGCCCTGCCAAATAATCAATGTGAAGCAAAATCGCAGCCCGCCTCC-----GCTCTACTCACTGGTGTTCATCTTTGTGGCAACA
 TTATTATACATCGGAGCCCTGCCAAATAATCAATGTGAAGCAAAATCGCAGCCCGC-----A

TGCTGGTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCAT
 TGCTGGTCATCCT
 TGCTGGTCATCCTC

SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:
 SEQ ID NO:

图6

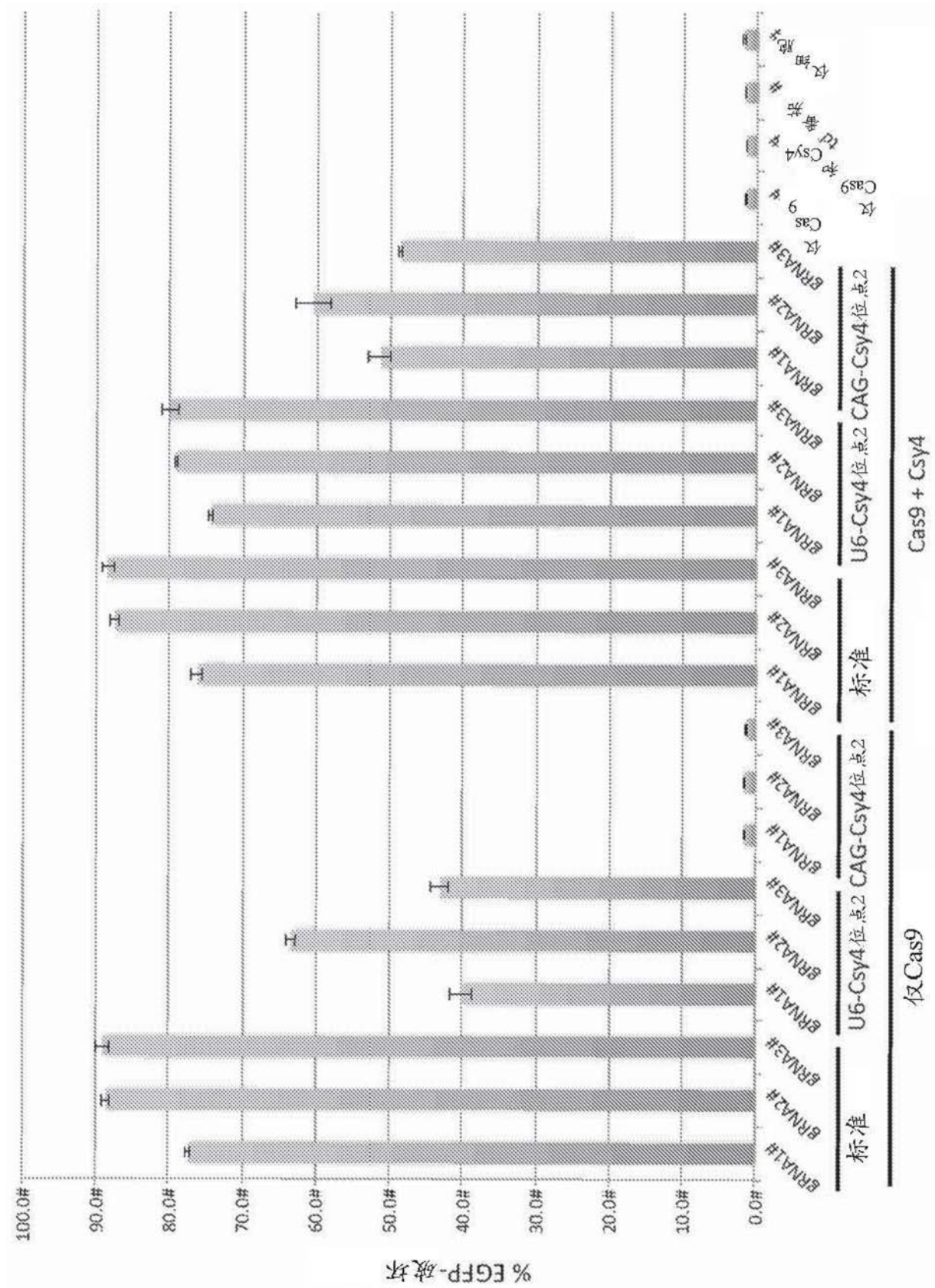


图7