



(12)发明专利申请

(10)申请公布号 CN 108431225 A

(43)申请公布日 2018.08.21

(21)申请号 201680074288.4

(22)申请日 2016.10.14

(30) 优先权数据

62/242,884 2015.10.16 US

(85)PCT国际申请进入国家阶段日

2018.06.15

(86)PCT国际申请的申请数据

PCT/IB2016/001902 2016.10.14

(87)PCT国际申请的公布数据

W02017/064566 EN 2017.04.20

(71)申请人 阿斯特拉捷利康股份公司

地址 瑞典南泰利耶

(72)发明人 M·马雷斯卡 M·博鲁利-叶甘内

H·杰斯瓦尔 L·M·迈尔

徐秀丰

(74) 专利代理机构 上海专利商标事务所有限公司 31100

代理人 余颖 陈扬扬

(51) Int.Cl.

C12N 15/11(2006.01)

C12N 15/79(2006.01)

权利要求书4页 说明书89页

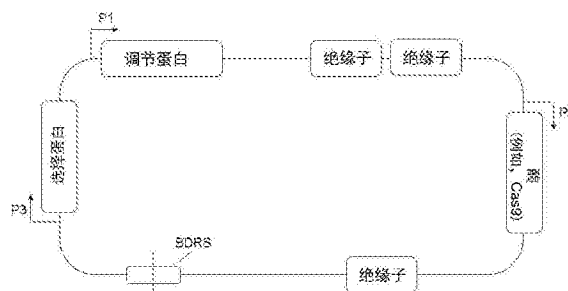
序列表85页 附图20页

(54)发明名称

细胞基因组的诱导型修饰

(57)摘要

本公开在一些实施方式中涉及用于细胞基因组的诱导型修饰的组合物和方法。



1. 一种工程改造的核酸构建体,其包括:
 - (a) 启动子,其可操作地连接编码调节蛋白的核酸;
 - (b) 诱导型启动子,其可操地连接编码切割核酸的酶的核酸,编码使核酸产生切口的酶的核酸,或编码催化核酸交换的酶的核酸,其中,所述诱导型启动子的活性通过所述调节蛋白调整;
 - (c) 至少两个绝缘子,其位于 (a) 的下游和 (b) 的上游;
 - (d) 至少一个绝缘子,其位于 (b) 的下游和 (a) 的上游;以及
 - (e) 至少一个脱氧核糖核酸 (DNA)-结合结构域识别序列,其位于 (b) 的下游和 (a) 的上游。
2. 如权利要求1所述的工程改造的核酸构建体,其包括至少两个DNA-结合结构域识别序列。
3. 如权利要求2所述的工程改造的核酸构建体,其中,所述至少两个DNA-结合结构域识别序列各自获自不同的物种。
4. 如权利要求3所述的工程改造的核酸构建体,其中,所述DNA-结合结构域识别序列中的至少一个对应人基因组AAVS1基因座中的核苷酸序列。
5. 如权利要求3所述的工程改造的核酸构建体,其中,所述DNA-结合结构域识别序列中的至少一个对应小鼠基因组Rosa26基因座中的核苷酸序列。
6. 如权利要求2所述的工程改造的核酸构建体,其中,至少一个DNA-结合结构域识别序列位于 (a) 的上游并且邻近 (a), 并且其中,至少一个DNA-结合结构域识别序列位于 (b) 的下游且邻近 (b)。
7. 如权利要求1-6中任一项所述的工程改造的核酸构建体,其还包括启动子,所述启动子可操作地连接编码选择性标志物蛋白的核酸。
8. 如权利要求1-7中任一项所述的工程改造的核酸构建体,其中,所述酶是核酸酶、切口酶或重组酶。
9. 如权利要求8所述的工程改造的核酸构建体,其中,所述酶是核酸酶。
10. 如权利要求9所述的工程改造的核酸构建体,其中,所述核酸酶是Cas核酸酶。
11. 如权利要求10所述的工程改造的核酸构建体,其中,所述Cas核酸酶是Cas9核酸酶。
12. 如权利要求11所述的工程改造的核酸构建体,其中,所述Cas9核酸酶是野生型Cas9核酸酶。
13. 如权利要求11所述的工程改造的核酸构建体,其中,所述Cas9核酸酶是催化失活的Cas9核酸酶。
14. 如权利要求13所述的工程改造的核酸构建体,其中,所述催化失活的Cas9核酸酶融合转录激活物肽、转录阻遏物肽或表观基因组调节物肽。
15. 如权利要求9所述的工程改造的核酸构建体,其中,所述核酸酶是Cpf1核酸酶。
16. 如权利要求10-15中任一项所述的工程改造的核酸构建体,其中,所述构建体还包括编码引导RNA的核酸。
17. 如权利要求16所述的工程改造的核酸构建体,其中,所述引导RNA位于 (b) 的下游和 (d) 的上游。

18. 如权利要求8所述的工程改造的核酸构建体,其中,所述酶是重组酶。
19. 如权利要求18所述的工程改造的核酸构建体,其中,所述重组酶是Cre重组酶。
20. 如权利要求1-19中任一项所述的工程改造的核酸构建体,其中,(b)的启动子是组织特异性诱导型启动子。
21. 如权利要求1-19中任一项所述的工程改造的核酸构建体,其中,(b)的启动子是发育特异性诱导型启动子。
22. 如权利要求1-21中任一项所述的工程改造的核酸构建体,其中,所述调节蛋白是四环素控制的反式激活因子(tTA)蛋白、反向四环素控制的反式激活因子(rtTA)蛋白或Lac阻遏物蛋白。
23. 如权利要求22所述的工程改造的核酸构建体,其中,所述调节蛋白是rtTA蛋白。
24. 如权利要求1-23中任一项所述的工程改造的核酸构建体,其中,(a)的启动子是诱导型启动子。
25. 如权利要求24所述的工程改造的核酸构建体,其中,(a)的启动子是组织特异性诱导型启动子。
26. 如权利要求24所述的工程改造的核酸构建体,其中,(a)的启动子是发育特异性诱导型启动子。
27. 如权利要求1-26中任一项所述的工程改造的核酸构建体,其中,所述至少一个DNA-结合结构域识别序列被具有FokI核酸酶结构域的核酸酶切割。
28. 如权利要求27所述的工程改造的核酸构建体,其中,所述核酸酶是锌指核酸酶、Tale核酸酶、Cpf1核酸酶或杂化dCas9-FokI核酸酶。
29. 一种载体,其包括权利要求1-28中任一项所述的工程改造的核酸构建体。
30. 如权利要求29所述的载体,其中,所述载体具有至少20千碱基的大小。
31. 如权利要求29或30所述的载体,其中所述载体是杆状病毒载体、细菌人工染色体、或酵母人工染色体。
32. 一种细胞,其包括权利要求1-31中任一项所述的工程改造的核酸构建体。
33. 一种细胞,其包括权利要求29-32中任一项所述的载体。
34. 如权利要求32或33所述的细胞,其中,所述细胞是干细胞。
35. 如权利要求34所述的细胞,其中,所述干细胞是多能干细胞。
36. 如权利要求35所述的方法,其中,所述多能干细胞是诱导型多能干细胞(iPSC)。
37. 一种修饰细胞基因组的方法,其包括:
向细胞导入工程改造的核酸构建体,所述工程改造的核酸构建体包括:
(a) 启动子,其可操作地连接编码调节蛋白的核酸;
(b) 诱导型启动子,其可操地连接编码切割核酸的酶的核酸,编码使核酸产生切口的酶的核酸,或编码催化核酸交换的酶的核酸,其中,所述诱导型启动子的活性通过所述调节蛋白调整;
(c) 至少两个绝缘子,其位于(a)的下游和(b)的上游;
(d) 至少一个绝缘子,其位于(b)的下游和(a)的上游;和
(e) 至少一个脱氧核糖核酸(DNA)-结合结构域识别序列,其位于(b)的下游和(a)的上游;

向所述细胞中导入工程改造的核酸,该工程改造的核酸包括可操作地连接编码引导RNA (gRNA) 的核酸的启动子,所述引导RNA靶向所述细胞基因组区域;

向所述细胞导入工程改造的核酸,该工程改造的核酸包括可操作地连接编码杂化核酸酶的核酸的启动子,所述杂化核酸酶结合至少一个DNA-结合结构域识别序列;并且

在存在效应物物质的情况下孵育所述细胞,所述效应物物质在导致细胞基因组修饰的条件下调整所述调节蛋白的活性。

38. 一种修饰细胞基因组的方法,其包括:

向细胞导入工程改造的核酸构建体,所述工程改造的核酸构建体包括:

(a) 启动子,其可操作地连接编码调节蛋白的核酸;

(b) 诱导型启动子,其可操地连接编码酶的核酸,所述酶切割核酸、使核酸产生切口、或催化核酸交换,其中,所述诱导型启动子的活性通过所述调节蛋白调整;

(c) 至少两个绝缘子,其位于 (a) 的下游和 (b) 的上游;

(d) 至少一个绝缘子,其位于 (b) 的下游和 (a) 的上游;

(e) 可操作地连接编码引导RNA (gRNA) 的核酸的启动子,所述引导RNA (gRNA) 靶向所述细胞基因组区域;和

(f) 至少一个脱氧核糖核酸 (DNA) -结合结构域识别序列,其位于 (b) 的下游和 (a) 的上游;

向所述细胞导入工程改造的核酸,所述工程改造的核酸包括可操作地连接编码杂化核酸酶的核酸的启动子,所述杂化核酸酶结合至少一个DNA-结合结构域识别序列;并且

在存在效应物物质的情况下孵育所述细胞,所述效应物物质在导致细胞基因组修饰的条件下调整所述调节蛋白的活性。

39. 一种工程改造的核酸,其包括:

(a) 启动子,其可操作地连接编码Cas9核酸酶的核酸,所述Cas9核酸酶不包括核定位信号,其中,编码所述酶的核酸侧接雌激素受体 (ERT2) 序列;和

(b) 脱氧核糖核酸 (DNA) -结合核酸酶识别序列。

40. 一种修饰细胞基因组的方法,其包括:

向细胞导入工程改造的核酸,所述工程改造的核酸包括:

(a) 启动子,其可操作地连接编码Cas9核酸酶的核酸,所述Cas9核酸酶不包括核定位信号,其中,编码所述酶的核酸侧接雌激素受体 (ERT2) 序列;和

(b) 脱氧核糖核酸 (DNA) -结合核酸酶识别序列;

向所述细胞中导入工程改造的核酸,所述工程改造的核酸包括可操作地连接编码引导RNA (gRNA) 的核酸的启动子,所述引导RNA靶向所述细胞基因组区域;并且

在存在他莫昔芬的情况下并在导致细胞基因组修饰的条件下孵育所述细胞。

41. 一种转基因小鼠,其在所述小鼠的基因组中包括权利要求1-31中任一项所述的工程改造的核酸构建体。

42. 如权利要求41所述的转基因小鼠,其中,所述工程改造的核酸构建体被整合到小鼠基因组的Rosa26基因座。

43. 如权利要求41或42所述的转基因小鼠,其中,所述小鼠具有免疫能力。

44. 如权利要求41-43中任一项所述的转基因小鼠,其中,在不存在诱导型启动子诱导

的情况下,所述酶的表达是无法测知的。

45.工程改造的核酸,其包括SEQ ID NO:8或12-20中任一项的序列。

46.一种转基因小鼠,所述小鼠的基因组中包括如权利要求45所述的工程改造的核酸构建体。

47.如权利要求46所述的转基因小鼠,其中,所述工程改造的核酸构建体被整合到小鼠基因组的Rosa26基因座。

48.如权利要求45或46所述的转基因小鼠,其中,所述小鼠具有免疫能力。

49.如权利要求46-48中任一项所述的转基因小鼠,其中,在不存在诱导型启动子诱导的情况下,所述酶的表达是无法测知的。

细胞基因组的诱导型修饰

[0001] 相关申请

[0002] 本申请根据35 U.S.C. §119 (e) 要求于2015年10月16日提交的美国临时申请序列号62/242,884的优先权,该申请内容通过引用全文纳入本文。

背景技术

[0003] 基因组编辑是一种遗传工程改造,其中使用工程改造的核酸酶修饰基因组(例如,导入、移除或替换DNA)。通常,核酸酶在基因组中所需位置产生特异性的双链断裂(DSB),并且利用细胞的同源性机制以通过同源重组(HR)和非同源末端连接(NHEJ)的自然过程修复诱导的断裂。当前,人们使用四个家族的工程改造的核酸酶:锌指核酸酶(ZFN)、转录激活物样效应物核酸酶(TALEN)、CRISPR/Cas系统、和经工程改造的兆核酸酶重新工程改造的寻靶核酸内切酶。

发明内容

[0004] 在一些实施方式中,本文提供了工程改造的核酸构建体(“工程改造的构建体”),其能够用于跨物种整合以及向基因组中导入诱导型基因组编辑系统。在一些实施方式中,在不另行导入载体(质粒)物质的情况下,将所述工程改造的构建体导入基因组。在一些实施方式中,本公开的工程改造的构建体在经非同源末端连接(NHEJ)途径位点特异性整合到基因组后,其允许空间控制且时间控制的靶基因表达活化(参见,例如,Maresca等.Genome Res. 2013年3月;23(3):539-46,通过引用纳入本文)。不同于当前可用的系统,其通常通过使用至少两个独立的构建体来调整细胞基因组的靶向修饰,本公开的基因组编辑系统依赖于单个构建体的活化,该构建体包括用于表达调节蛋白的遗传因子以及用于表达靶基因的遗传因子。该单个构建体的构造导致严格调控且基本上无渗漏的靶基因表达,因而提供相对于当前可用的系统更加精确且有效的基因组编辑能力。

[0005] 因此,本公开的一些实施方式提供了工程改造的核酸构建体,其包括(a)启动子,其可操作地连接编码调节蛋白(例如,诱导物蛋白或阻遏物蛋白)的核酸;(b)诱导型启动子,其可操作地连接编码切割核酸的酶(例如,Cas9核酸酶、Cpf1核酸酶或其功能性等同物)的核酸,编码使核酸产生切口的酶(例如,Cas9切口酶)的核酸,或编码催化核酸交换的酶(例如,Cre重组酶)的核酸,其中,诱导型启动子的活性通过调节蛋白调整;(c)至少两个绝缘子(例如,哺乳动物绝缘子),其位于(a)的下游且(b)的上游;至少一个绝缘子,其位于(b)的下游且(a)的上游;以及(e)至少一个脱氧核糖核酸(DNA)-结合结构域识别序列,其位于(b)的下游且(a)的上游。在一些实施方式中,至少两个DNA-结合结构域识别序列位于(b)的下游且(a)的上游。在一些实施方式中,诱导型启动子可操作地连接编码调节基因表达的酶的核酸(例如,融合KRAB的Cas9,融合VP64的Cas9,融合p300的Cas9),或编码修饰核苷酸碱基的酶的核酸(例如,融合AID/ApoBEC结构域并且融合尿嘧啶糖苷酶抑制剂的无核酸酶活性Cas9(Cas9 dead)或切口酶)。

[0006] 在一些实施方式中,本公开还提供包括工程改造的核酸构建体的载体,包括工程

改造的核酸构建体的细胞,或包括包含工程改造的核酸构建体的载体的细胞。

[0007] 在一些实施方式中,本文还提供了修饰细胞基因组的方法。例如,本公开的方法可以用于使感兴趣的基因缺失(敲除),导入(敲入)感兴趣的基因,或修饰感兴趣的基因。

[0008] 在一些实施方式中,酶是核酸酶、切口酶或重组酶。

[0009] 在一些实施方式中,诱导型启动子是组织特异性诱导型启动子或发育特异性诱导型启动子。

[0010] 在一些实施方式中,调节蛋白是四环素控制的反式激活因子(tTA)蛋白、反向四环素控制的反式激活因子(rtTA)蛋白或Lac阻遏物蛋白。

[0011] 在一些实施方式中,DNA-结合结构域识别序列被具有FokI核酸酶结构域的核酸酶切割。这样的“杂化核酸酶”的非限制性示例包括锌指核酸酶(ZFN)、转录激活物样效应物核酸酶(TALEN)和杂化Cas9-FokI核酸酶。

[0012] 在一些实施方式中,DNA-结合结构域识别序列被不具有FokI核酸酶结构域的核酸酶切割。例如,可以使用Cpf1核酸酶切割DNA-结合识别序列。与本文所述这些核酸酶在结构和功能相似的其它核酸酶也可用于本方法。

[0013] 本文提供了转基因小鼠,在该小鼠的基因组中包括本文提供的工程改造的核酸构建体(例如,T0ICas构建体)。在一些实施方式中,工程改造的核酸构建体被整合到小鼠基因组的Rosa26基因座。应当理解的是尽管Rosa26基因座在一些实施方式中详述,但是本公开并不限于位于Rosa26基因座的基因组整合。本公开的工程改造的构建体可以被整合到小鼠基因组的任何基因座中(或人基因组,当适用于诸如T0IC人iPSC系的T0IC细胞系的产生时)。

[0014] 一些实施方式中,小鼠具有免疫能力。在一些实施方式中,在不存在诱导型启动子诱导的情况下(例如,不存在Dox给予的情况下),无法检测到酶(例如,Cas9)的表达。

[0015] 在一些实施方式中,本公开的工程改造的核酸包括SEQ ID NO:8的序列。

[0016] 本文还提供了这样的转基因小鼠(例如,具有免疫能力的小鼠),在该小鼠的基因组中包括工程改造的核酸构建体,该核酸构建体包括SEQ ID NO:8(T0ICas9)的序列。在一些实施方式中,将包括SEQ ID NO:8序列的工程改造的核酸构建体整合到小鼠基因组的Rosa26基因座(或其它基因座)。

[0017] 附图简要说明

[0018] 本文附图并不是按照比例绘制。为了清楚的目的,并不是每一个组件可以在每一个附图中标记出来。

[0019] 图1是根据本公开工程改造的核酸构建体示例的示意图。启动子P1可操作地连接编码调节蛋白(例如,诱导物蛋白或阻遏物蛋白)的核酸,其位于两个(例如,至少两个)绝缘子的上游。绝缘子位于启动子P2的上游,其可操作地连接编码这样的酶的核酸,所述酶切割核酸(例如,Cas9核酸酶、Cpf1核酸酶,或其功能性等同物)或使核酸产生切口(例如,Cas9切口酶),或催化核酸交换(例如,Cre重组酶)。该酶位于附加绝缘子上游,附加绝缘子位于(例如,至少一个)脱氧核糖核酸(DNA)-结合结构域识别序列(BDRS)上游。工程改造的构建体的大小可以是例如至少15千碱基(kb)。

[0020] 图2是本公开工程改造的核酸构建体的示意图,其用于插入腺相关病毒整合位点1(AAVS1)基因座(人)或ROSA^{βgeo}26(ROSA26)基因座(小鼠)。工程改造的构建体的大小是

17575kb。

[0021] 图3A-3B显示了包含本公开的多四环素 (doxycycline) 诱导型系统的诱导型多能干细胞 (iPSC) 的成像,如图3所示,在不存在多四环素 (-dox) (图4A) 和存在多四环素 (+dox) 的情况下。绿色荧光蛋白 (GFP) 是核酸酶 (例如,Cas9核酸酶、Cpf1核酸酶或其功能性等同物) 或重组酶 (例如,Cre) 表达活化的指示物。GFP仅在多四环素存在的情况下表达,这表明该系统没有“渗漏” (图3B)。

[0022] 图4显示了表示DNA水平的诱导的电泳凝胶图像。使用Surveyor试验切除已知人基因Usp14 (注意:该概念适用于例如人基因组的任何序列)。该试验显示,只有当向细胞培养基添加多四环素并且存在引导RNA (gRNA) 时,诱导型系统在iPSC中活化。分析不同的细胞库或细胞的单个克隆 (C6、C25、C41),其各自包含本公开的诱导型Cas9系统。通过电泳凝胶中其它较低的条带的释放来指示切割。在所有分析的细胞中观测到相同的切割效率。

[0023] 图5是根据本公开工程改造的核酸构建体示例的示意图。启动子P1可操作地连接编码进行切割的酶的核酸,编码使核酸产生切口的酶的核酸,或编码催化核酸交换的酶的核酸,其中,编码所述酶的核酸侧接雌激素受体 (ERT2) 序列的修饰形式。该酶位于 (例如,至少一个) 脱氧核糖核酸 (DNA) -结合结构域识别序列 (BDRS) 上游。

[0024] 图6是用于靶向MTH1基因的工程改造的核酸构建体示例的示意图。启动子P1可操作地连接编码调节蛋白 (例如,诱导物蛋白或阻遏物蛋白) 的核酸,其位于 (例如,至少两个) 绝缘子的上游。绝缘子位于启动子P2的上游,其可操作地连接编码进行切割的酶 (例如,Cas9核酸酶、Cpf1核酸酶,或其功能性等同物) 的核酸,编码使核酸产生切口的酶 (例如,Cas9切口酶) 的核酸,或编码催化核酸交换的酶 (例如,Cre重组酶) 的核酸。该酶位于编码引导RNA (gRNA) 的核酸上游,该引导RNA (gRNA) 位于附加绝缘子上游,该附加绝缘子位于 (例如,至少一个) 脱氧核糖核酸 (DNA) -结合结构域识别序列 (BDRS) 上游。在该示例中,gRNA对MTH1基因具有特异性,其组成型表达并且结合Cas9以在Cas9表达后形成活性复合物。在该示例中的构建体包括DNA-结合结构域识别序列,其对应于位于人基因组AAVS1基因座中的序列。将构建体整合至AAVS1基因座,并且选择重组克隆。检测重组克隆的诱导后内源性MTH1基因座的切割。

[0025] 图7是显示通过TIDE软件确定的切割效率的图像。该分析显示,85%的细胞在MTH1基因中包含缺失 (敲除),然而在不存在诱导的情况下没有观测到缺失。来自该分析的数据显示,当与存在于测试的细胞系中的特定突变结合时,MTH1不是致命的。

[0026] 图8显示了不同的工程改造的核酸构建体的示意图,其中之一 (顶端图) 在用多四环素诱导后表达GFP。

[0027] 图9显示了确认图8顶端图所示工程改造的构建体的种系传递以及整合到Rosa26基因座的数据。

[0028] 图10显示了工程改造的TOICas构建体示例的示意图。Cas9通过T2A肽融合GFP,并且由绝缘子侧接,用于严格调控。

[0029] 图11是电泳凝胶图,其显示了在Dox和引导RNA存在的情况下DNA的高效切割。

[0030] 图12显示了使用AAV-U6gRNA (p53^{-/-}/KrasG12D) 构建体 (顶端图) 的基因靶向策略。Kras靶位点位于内含子中,以在HDR ssODN-介导的点突变“修复”失效的情况下使得对基因可能造成的损害较小。Trp53CRISPR按照“常规”策略 (底部图)。

[0031] 图13显示了在Dox刺激(48h)和移除(112h)细胞后,GFP曲线随时间的变化。

[0032] 图14是显示来自对用Dox刺激且用AAV-gRNA-Kras感染纤维母细胞以产生点突变和插入缺失(indel)的surveyor试验结果的电泳凝胶图。

[0033] 图15是显示了surveyor核酸酶试验结果的电泳凝胶图,其证明了气管内给予AAV-gRNA-Trp53(#37、#42、#52)的小鼠肺中Trp53基因的精确切割(红色星号),验证了TOICas的功能。

[0034] 图16显示来自Sanger测序分析的数据,其证明了在不存在Dox的情况下没有观测到TOICas9(诱导型Cas9)切割。

[0035] 图17是这样的图表,其显示了作为亲本群体和稳定库(stable pool)一部分的GFP表达和Cas9表达。

[0036] 图18是这样的图表,其显示了Cas9表达对急性DNA损伤的作用(gH2AX焦点形成)。

[0037] 图19是这样的图表,其显示了Cas9表达对持续性DNA损伤的作用(微核形成)。

[0038] 图20-28显示了本公开涵盖的TOICas构建体示例的质粒图。图20-28中的序列分别对应于SEQ ID NO:12-20。

具体实施方式

[0039] 本公开提供了工程改造的核酸构建体(例如,“TOICas构建体”),其能够用于跨物种整合以及向基因组中导入诱导型基因组编辑系统,在一些实施方式中,其不需要另行导入载体(质粒)物质。本公开的工程改造的构建体促进线性形式的构建体位点特异性地整合到基因组的单个基因座中,基因组的多个不同的基因座中,或不同基因组(来自不同的物种)的基因座中。这种直接整合依赖于构建体以及靶向的基因组基因组经协调的酶促切割,以及将线性的构建体连接到基因组基因座,例如,经非同源末端连接(NHEJ)途径(参见,例如,Maresca等.Genome Res.2013年3月;23(3):539-46,通过引用纳入本文)。

[0040] 定制的核酸酶,如Cas9、Cpf1、锌指核酸酶、Tale核酸酶和其功能性同等物,包含DNA切割结构域以及由最优化DNA结合模块组装的DNA结合结构域。在细胞中,这些核酸酶在基因组中位于或接近由核酸酶DNA结合结构域识别的序列(“DNA-结合结构域识别序列”)处生成双链断裂,并且诱导DNA损伤修复本公开的工程改造的构建体是(部分)基于这样的结果,其显示游离构建体可以连接到靶基因组基因座,如果该构建体和该基因座包含相同的DNA-结合结构域识别序列。

[0041] 在一些实施方式中,多个DNA-结合结构域识别序列位于构建体中,因此它们侧接不需要的载体(例如,细菌质粒)DNA。这样的结构导致构建体整合到基因组后载体DNA被移除。

[0042] 本公开的工程改造的构建体(例如,“TOICas构建体”)的多用性在于,在一些实施方式中,它们包括,用于以时间控制和空间控制的方式诱导基因表达的遗传因子,促进构建体位点特异性地整合到多个不同基因座的DNA-结合结构域识别序列的阵列(例如,两个或更多个),并且是基本无渗漏的。因此,相对于当前可用的基因组编辑系统,本公开的工程改造的构建体提供了更加精确和高效的基因组编辑能力。

[0043] TOICas构建体和转基因动物

[0044] 在一些实施方式中,本文所述工程改造的构建体指TOI、TOICas或TOICas9构建体,

其包括编码Cas9酶的核酸。应当理解的是,在TOIC、TOICas或TOICas9构建体的任一中,编码Cas9酶的核酸可以被替换成其它酶,其切割核酸,使核酸产生切口,催化核酸的交换,调节基因表达,或修饰核苷酸碱基。TOIC构建体的非限制性示例述于图20-28,其核酸序列分别由SEQ ID NO:12-20表示。

[0045] 在一些实施方式中,TOIC构建体包括图20-28中的任一种,例如,其分别具有SEQ ID NO:12-20中任一序列。如将在后续部分详细讨论的,本公开的TOICas构建体包括可操作地连接调节蛋白的启动子(P1)(例如,CMV、CAGG、CBh或EF1 α 或组织特异性启动子),其处于串联的两个绝缘子(绝缘子1和绝缘子2)的上游,这两个绝缘子处于可操作地连接酶(例如,Cas9)的启动子(P2)(例如,通过调节蛋白结合的顺序排列的序列),启动子(P2)处于绝缘子(绝缘子3)的上游,该绝缘子处于DNA-结合结构域识别位点(BDRS)(例如,核酸酶、重组酶或整合酶)的上游,该BDRS任选地位于可操作地连接编码选择蛋白(例如,药物选择蛋白或荧光标志物)的核酸的启动子(P3)的上游(参见,例如,图1和图10)。在一些实施方式中,TOIC构建体包括编码引导RNA(gRNA)的核酸,其可以位于例如酶和绝缘子3之间,或绝缘子3和BDRS之间。

[0046] 本文还提供了包括TOIC构建体的转基因动物,如转基因小鼠模型。尽管本文所述许多实施方式指涉及基因小鼠模型,但是应当理解的是,本公开涵盖各种转基因动物模型(无脊椎动物和脊椎动物),其包括但不限于:堡礁海绵(*Amphimedon queenslandica*)、阿巴海胆(*Arbacia punctulata*)、海兔(*Aplysia*)、佛罗里达文昌鱼(*Branchiostoma floridae*)、秀丽隐杆线虫、澳大利亚蝗虫(*Caledia captive*)、四纹豆象(*Callosobruchus maculatus*)、网翅蝗科雏蝗属雏(*Chorthippus parallelus*)、玻璃海鞘(*Ciona intestinalis*)、水蚤属(*Daphnia* spp.)、扁蝇科、突眼蝇、果蝇、夏威夷短尾鱿鱼(*Euprymna scolopes*)、大蜡螟(*Galleria mellonella*)、双斑蟋、水螅、大西洋乌贼(*Loligo pealei*)、利加诺大口虫(*Macrostomum lignano*)、梳状水母(*Mnemiopsis leidyi*)、新星海葵(*Nematostella vectensis*)、异体住囊虫(*Oikopleura dioica*)、黏液海绵(*Oscarella carmela*)、哈维宁思明钩虾(*parhyale hawaiiensis*)、沙蚕蠕虫(*Platynereis dumerilii*)、短角蚱蜢属(*Podisma* spp.)、太平洋蛔虫(*Pristionchus pacificus*)、黄粪蝇、地中海圆头涡虫(*Schmidtea mediterranea*)、口胃动物(stomatogastric)、紫色球海胆海胆(*Strongylocentrotus purpuratus*)、薄荷酱蠕虫(*Symsagittifera roscoffensis*)、赤拟谷盗(*Tribolium castaneum*)、以及,丝盘虫(*Trichoplax adhaerens*)、正颤蚓(*Tubifex tubifex*) (无脊椎动物);以及,铃蟾属(*Bombina*)、卡罗来纳蜥(*Carolina anole*) (安乐蜥(*Anolis carolinensis*))、猫(猫属樟子松猫(*Felis sylvestris catus*))、鸡(泰和鸡(*Gallus gallus domesticus*))、棉鼠(刚毛棉花鼠(*Sigmodon hispidus*))、狗(家犬(*Canis lupus familiaris*))、金黄仓鼠(叙利亚仓鼠(*Mesocricetus auratus*))、豚鼠(天竺鼠(*Cavia porcellus*))、小型棕蝙蝠((小棕蝠*Myotis lucifugus*))、青鳉(青鳉(*Oryzias latipes*)),或日本异鳉科(Japanese ricefish))、鼠((*Mus musculus*)小家鼠)、裸鼹鼠(Naked mole-rat)(裸鼹鼠(*Heterocephalus glaber*))、弗氏假鳃鳉(*Nothobranchius furzeri*)、鸽子(家鸽(*Columba livia domestica*))、孔雀鱼、鼠((*Rattus norvegicus*)褐家鼠)、恒河猴(或猕猴)(猕猴(*Macaca mulatta*))、七鳃鳗(Sea lamprey)(海七鳃鳗(*Petromyzon marinus*))、东方鲀(*Takifugu*) (红鳍东方鲀(*Takifugu rubripes*)),河豚)、三

刺鱼(三刺鱼(*Gasterosteus aculeatus*))、热带爪蟾(*Xenopus tropicalis*)和非洲爪蟾(*Xenopus laevis*) (非洲爪蛙)、斑胸草雀(珍珠鸟(*Taeniopygia guttata*))和斑马鱼(斑马鱼(*Danio rerio*),淡水鱼)。

[0047] 在一些实施方式中,转基因动物模型选自鱼、青蛙、鸟、小鼠、大鼠、仓鼠、猫、犬、猪、羊和猴。在一些实施方式中,转基因模型是小鼠模型。在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:8的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:8的序列。

[0048] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:12的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:12的序列。

[0049] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:13的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:13的序列。

[0050] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:14的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:14的序列。

[0051] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:15的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:15的序列。

[0052] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:16的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:16的序列。

[0053] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:17的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:17的序列。

[0054] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:18的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:18的序列。

[0055] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:19的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物

(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:19的序列。

[0056] 在一些实施方式中,本公开提供的工程改造的核酸包括SEQ ID NO:20的序列。本文还提供了转基因动物(例如,小鼠),例如,具有免疫能力的动物(例如,小鼠),在该动物(例如,小鼠)基因组中包括工程改造的核酸构建体,该工程改造的核酸构建体包括整合到例如动物(例如,小鼠)基因组Rosa26基因座的SEQ ID NO:29的序列。

[0057] 在动物模型中诱导酶(例如,Cas9)表达可以通过给予多四环素或其它适当的诱导剂(取决于用于TOIC构建体的特定诱导系统)实现。在一些实施方式中,诱导剂(直接或间接激活TOIC构建体的诱导型启动子的物质)经由注射(例如,尾静脉注射)或口服灌胃给予动物。

[0058] 本文提供的转基因动物,通过同源重组或通过非同源末端连接,可以用于产生敲除或敲减等位基因,或用于过表达基因或将基因敲入特定的基因座。例如,这可以通过向动物给予模板DNA(例如,包含感兴趣的修饰)以及核酸实现,所述核酸编码gRNA靶向的基因座和/或感兴趣的基因。感兴趣的基因的示例包括但不限于,致癌基因,如Pik3ca、Kras、Braf、Nras,和肿瘤抑制基因,如Pten、p53、Rb、Apc、p16/p19、Brcal、Brca2、Lkb1。通过将TOIC转基因小鼠与模板DNA以及靶向基因和靶向一种或多种感兴趣的基因的gRNA组合,可以产生各种疾病模型。这样的模型包括但不限于,肺癌(例如,Kras、Lkb、p53和/或Rb;例如,靶向MAPK,新陈代谢),胰腺癌(例如,Kras、p53、p16/p19和/或Pdx1;例如,靶向MAPK)、前列腺癌(例如,Pten、Brcal、Brca2和/或p53;例如,靶向PI3K、AR、ASO、DDR),乳腺癌(例如,Pik2ca、p53和/或Pten;例如,靶向PI3K、SERD),卵巢癌(例如,Brcal、Brca2、p53和/或Rb;例如,靶向Erk、MEK、Kras、ASO、modp53),黑素瘤(例如,Braf和/或Nras;例如,靶向MAPK)以及结直肠癌(例如,Pik3ca、Kras和/或Apc;例如,靶向MAPK、PI3K)。

[0059] 例如,编码gRNA以及相关模板DNA的核酸可以使用病毒(例如,腺相关病毒或腺病毒)经气管内、肾内或腹膜内转导给予动物。

[0060] 在一些实施方式中,将编码gRNA的核酸整合到转基因动物模型的基因组中。例如,组织特异性gRNA或组成型表达的gRNA可以被整合到转基因动物模型的基因组中。

[0061] 本文还提供了衍生自(获得自)本公开的转基因动物的类器官(organoid)(体外生长的三维器官芽(organ-bud))。因此,类器官可以包括任何TOIC构建体。

[0062] 例如,优选地,实例中描述的那些TOIC动物(例如,小鼠)可以具有免疫能力(能够在暴露于抗原后产生正常的免疫反应)。此外,本文开包括的是免疫功能不全的(具有弱化的免疫系统)TOIC动物。

[0063] 用于基因组整合的酶

[0064] 本公开的工程改造的构建体(例如,“TPICas构建体”)用于促进线性形式的构建体直接地、位点特异性地连接到单个基因组或多个不同的基因组的单个基因座或多个不同的基因座。这样的直接连接通过非同源的末端连接(NHEJ)途径发生(参见,例如Maresca等 Genome Res. 2013Mar; 23(3):539-46,通过引用纳入本文)。位点特异性连接依赖于杂化核酸酶的存在,所述杂化核酸酶包含DNA结合结构域和DNA切割结构域(通常是,FokI结构域),以及如下核酸的存在,所述核酸包含至少一个DNA结合结构域识别序列。“DNA结合结构域识别序列”是这样的核苷酸序列,核酸酶DNA结合结构域与之结合,并且核酸酶DNA切割结构域

将其切割。工程改造的构建体包含至少一个DNA结合结构域识别序列,其被杂化核酸酶识别和切割。对于工程改造的构建体的切割产生线性形式,然后可以将其以位点特异性的方式“连接”进入基因组。

[0065] 本公开的工程改造的构建体(例如,“T0ICas构建体”),在一些实施方式中,包括单个DNA-结合识别序列(BDRS)或DNA-结合结构域识别序列的阵列(例如,两个或更多个),其促进核酸的位点特异性基因组整合。优选地,本公开的工程改造的构建体可以用于促进线性形式的构建体位点特异性地连接到若干不同基因组的多个不同基因座,其能够用于相同构建体的跨物种整合。

[0066] 在一些实施方式中,工程改造的构建体的DNA-结合识别序列对应于位于Rosa26基因座中的序列,这样核酸可以被整合到小鼠基因组中。在一些实施方式中,工程改造的构建体的DNA-结合识别序列对应于位于AAVS1基因座中的序列,这样核酸可以被整合到人基因组中。位于其它基因组基因座中的其它DNA-结合结构域识别序列涵盖在本公开的范围内。

[0067] 用于使工程改造的构建体线性化的杂化核酸酶的示例包括但不限于,锌指核酸酶(ZFN)、Tale核酸酶(TALENs)、dCas9-FokI融合蛋白(融合至FokI的催化失活的Cas9)、Cas9、融合至FokI的Cas9切口酶、以及经进化以生成突出端的Cas9变体。

[0068] 在一些实施方式中,DNA-结合结构域识别序列是ZFN DNA结合结构域识别序列,其通过一个或多个锌指结合。个体ZFN的DNA-结合结构域可以包含3-6个单独的锌指重复,并且可以各自识别9-18个碱基对。如果锌指结构域对其意图的靶位置具有特异性,那么甚至识别总共18个碱基对的3-指ZFN对亦可以靶向哺乳动物基因组中的单个基因座。

[0069] 在一些实施方式中,DNA-结合结构域识别序列是TALEN DNA结合结构域识别序列,其通过一个或多个TAL效应物单元结合。TAL效应物是由黄单胞菌属(*Xanthomonas*)细菌分泌的蛋白质。DNA结合结构域通常包含重复的高度保守的33-34氨基酸序列,除了第12和第13个氨基酸。这两个位置是高度可变的(复可变双残基,RVD),并且显示出与特异性核苷酸识别的强相关性(Boch等*Science* 326 (5959):1509-12,2009;Moscou等*Science* 326 (5959):1501,2009,其各自通过应用纳入本文)。在一些实施方式中,通过对包含适当RVD的重复区段的组合进行选择来工程改造特异性DNA-结合结构域(Boch等*Nature Biotechnology* 29 (2):135-6,2011)。

[0070] 在一些实施方式中,DNA-结合结构域识别序列是与两个共表达的引导RNA互补(例如,100%互补性)的序列。在一些实施方式中,DNA-结合结构域识别序列是与两个共表达的引导RNA至少80%、至少85%、至少90%、至少95%或至少98%互补的序列。在这样的实施方式中,融合FokI核酸酶的催化失活的Cas9(dCas9)可以被用于在工程改造的核酸中生成双链断裂。

[0071] 应当理解的是,本公开的工程改造的构建体可以包括编码核酸酶(例如,Cas9核酸酶、Cpf1核酸酶或其功能性等同物)的核酸,并且包含不一定由该工程改造的核酸的核酸酶识别的DNA-结合核酸酶识别序列。例如,用于基因组编辑的工程改造的构建体可以编码Cas9(例如,野生型或在其他情况下具有催化活性的Cas9,出于编辑细胞基因组的目的)或Cpf1核酸酶,并且还可以包含DNA-结合核酸酶识别序列,其对锌指核酸酶或融合至FokI核酸酶的催化失活的Cas9(dCas)具有特异性。如果情况是这样的,那么可能需要向细胞导入编码特异性地识别并切割DNA-结合核酸酶识别的融合FokI的dCas9或锌指核酸酶的另一核

酸。

[0072] 用于基因组编辑的酶

[0073] 本文同样描述了编码这样酶的核酸,所述酶切割核酸,使核酸产生切口,或催化核酸的交换。切割核酸的酶被称为核酸酶。使核酸产生切口的酶被称为切口酶。催化核酸交换的酶被称为重组酶。

[0074] “切割”指这样的过程,其中核酸酶切开(水解)双链核酸的核酸主链(例如,糖-磷酸主链)。因此,该核酸酶产生两个切口:一个位于双链核酸一条链的核苷酸亚基之间的主链中,另一个位于双链核酸另一条链的核苷酸亚基之间的主链中。单个核酸分子的切割通常导致产生两个分开的核酸分子。与之相反,“使…产生切口”指这样的过程,其中切口酶仅切开双链核酸的一条链。“催化核酸的交换”指这样的过程,其中遗传物质经断裂,并且与其它遗传物质连接,并且包括基因重组。重组是将遗传物质通过例如染色体交叉或通过连接DNA区段来对遗传物质进行重组或重排。

[0075] 在一些实施方式中,工程改造的核酸编码核酸酶。本公开的核酸酶可以经工程改造以切开预定的核苷酸序列,以允许例如遗传信息的高效工程改造以及各种不同核酸修饰的产生。工程改造的核酸酶的示例包括但不限于,DNA-引导的内切酶,RNA-引导的内切酶(RGEN),诸如Cas9或Cpf1,锌指核酸酶(ZFN)(Kim等Proc Natl Acad Sci USA 93(3):1156-60,1996;Bitinaite等Proc Natl Acad Sci USA 95(18):10570-5,1998;以及Cathomen等Mol. Ther. 16(7):1200-7,2008),TAL效应物核酸酶(TALEN,转录激活物样效应物核酸内切酶)(Boch等Science 326(5959):1509-12,2009;Christian等.Genetics 186(2):757-61,2010);以及Miller等Nature Biotechnology 29(2):143-8,2011)(表1),以及其功能性等同物。核酸酶通常包括DNA结合结构域,其识别并结合特定DNA序列,以及DNA切割结构域,其在位于DNA结合结构域或接近DNA结合结构域(例如,在10个核苷酸内)的位置切割DNA。例如,ZFN包括锌指结构域,其结合DNA,以及Fok I结构域,其切割DNA(Kim等Natl Acad Sci USA 93(3):1156-60,1996)。相似地,TALEN包括TAL效应物单元,其结合A,以及Fok I结构域,其切割DNA。RNA-引导的Cas9核酸酶切割DNA,但是为了实现这样的切割,其必需首先由引导RNA引导至靶切割位点,其与DNA切割位点互补并结合,如本文其它地方所述。

[0076] 表1.工程改造的核酸酶系统

[0077]

核酸酶	DNA 结合结构域	DNA 切割结构域
RGEN	引导 RNA, 其与靶 DNA 杂交(1:1 核苷酸碱基配对)	Cas9 蛋白(包含两个核酸酶结构域); 或 或 Cpf1 蛋白
ZFN	锌指(各模块识别 3 bp 的靶序列)	Fok I 限制性酶核酸酶结构域(需要用于切割的二聚化)
TALEN	TAL 效应物单元(各模块识别 1 bp 的靶序列)	Fok I 限制性酶核酸酶结构域(需要用于切割的二聚化)

[0078] Cas9 (CRISPR相关蛋白9) 是RNA-引导的DNA核酸酶, 其与诸如酿脓链球菌 (*Streptococcus pyogenes*) 的细菌中的CRISPR (成簇且规律间隔的回文重复序列 (Clustered Regularly Interspersed Palindromic Repeats)) 获得性免疫系统相关联。用于编辑、调节和靶向基因组的CRISPR系统可以包括至少两个不同的组件: (1) 引导RNA (gRNA) 以及 (2) Cas9。gRNA是这样的单个嵌合转录本, 其将内源性细菌CRISPR靶向RNA (crRNA) 的靶向特异性与反式激活crRNA (tracrRNA) 的折叠性质相结合。通常, 用于基因组编辑的gRNA转录自细胞内的质粒或基因组基因座。gRNA转录本与Cas9形成复合物, 并且然后, 例如, 作为crRNA序列与其基因组DNA中互补靶序列之间碱基配对的结果, gRNA/Cas9复合物被招募至靶序列。

[0079] 在典型的合成CRISPR/Cas9基因组编辑系统中, 感兴趣的基因序列 (基因组靶序列) 通过使用与感兴趣的序列互补的gRNA进行修饰, 其将gRNA/Cas9复合物引导至靶标 (Sander JD等, 2014 *Nature Biotechnology* 32, 247-355, 通过引用纳入本文)。Cas9内切核酸酶切开前间区序列邻近基序 (PAM) 的基因组靶DNA上游, 产生双链断裂。双链断裂的修复常常导致双链断裂位点处的插入或缺失。CRISPR/Cas9系统常常用于编辑细胞的基因组, 各迭代需要设计和导入对感兴趣实施方式靶序列具有特异性的新的gRNA序列。

[0080] 在一些实施方式中, 本公开的工程改造的构建体包括启动子 (例如, 诱导型启动子), 其可操作地连接编码引导RNA的核酸 (例如, 位于编码Cas9核酸酶的核酸下游), 其将Cas9核酸酶引导至基因组靶 (修饰) 位点。与Cas9功能性相似的酶可以根据本公开使用。

[0081] Cpf1是2型CRISPR-Cas系统的单个RNA-引导的内切核酸酶 (Zetsche等, 2015, *Cell* 163:1-13, 通过引用纳入本文)。Cpf1与Cas9类似, 是一种两组件RNA可编程的DNA核酸酶。靶向的DNA被切割成位于5' T富集的前间区序列邻近基序 (PAM) 远端的5nt交错切口。两个Cpf1同源物在人细胞中展现出强健的核酸酶活性, 任何一种都可以在本文使用。与Cpf1功能性相似的酶可以根据本公开使用。

[0082] 在其他实施方式中, 将编码引导RNA的核酸导入这样的宿主细胞中, 该宿主细胞经修饰以在细胞基因组中表达 (例如, 稳定表达) 本公开的工程改造的构建体 (例如, 编码Cas9核酸酶的构建体), 并且将其用于替代在宿主细胞基因组的修饰期间导入宿主细胞的任何不需要的DNA。例如, 细胞, 如干细胞 (例如, 多能干细胞), 可以经修饰以在细胞基因组中表达 (例如, 稳定表达) 这样的工程改造的构建体, 其包括 (a) 可操作地连接编码调节蛋白的核酸的启动子, (b) 诱导型启动子, 其可操作地连接编码Cas9或Cpf1的核酸, 其中诱导型启动子的活性通过调节蛋白来调节, (c) 至少两个绝缘子, 其位于 (a) 的下游而在 (b) 的上游, 和 (d) 至少一个绝缘子, 其位于 (b) 的下游。对该宿主细胞的修饰可以使得在构建体的位点特异性整合后, 载体 (例如, 细菌质粒) DNA导入细胞的基因组, 以及其它不需要的DNA (例如, 编码选择标志物的序列) 的导入。该载体DNA和任何其它不需要的DNA可以被移除, 例如, 通过导入 (a) 编码由DNA-结合结构域识别序列侧接的引导RNA的工程改造的核酸, (b) 编码杂交核酸酶的工程改造的核酸, 所述杂交核酸酶识别并切割侧接引导RNA的DNA-结合结构域识别序列以及侧接位于细胞基因组中不需要的序列的DNA-结合结构域识别序列, 和 (c) 正交Cas9 (orthogonal Cas9) (Cas9来自与获得宿主细胞Cas9的物种不同的物种)。例如, 宿主细胞可以经工程改造以表达获自酿脓链球菌 (*Streptococcus pyogenes*) 的Cas9, 以及获得自金黄色葡萄球菌 (*Streptococcus aureus*)、嗜热链球菌 (*Streptococcus thermophilis*) 或

脑膜炎奈瑟菌 (*Neisseria meningitidis*) 的正交Cas9。其它正交Cas9核酸酶包括在本公开内。杂化核酸酶、引导RNA和正交Cas9可以包括在相同的构建体(例如,载体)或各自位于单独的构建体上。引导RNA替代载体DNA或宿主细胞基因组中任何不需要的DNA。

[0083] 在一些实施方式中,正交Cas9用于用感兴趣的引导RNA以及任选的选择标志物替代整合在宿主细胞基因组中的细菌质粒序列。在一些实施方式中,宿主细胞是多能干细胞(例如,人多能干细胞,如人诱导型多能干细胞)或胚胎干细胞(例如,小鼠胚胎干细胞,用于小鼠模型系统的产生)。在这些细胞类型中,细菌质粒序列具有使感兴趣的基因座失活的可能性,因此优选移除细菌质粒序列。

[0084] 在一些实施方式中,正交Cas9或Cpf1以及引导RNA被用于在多能状态后,在分化的状态中使感兴趣的基因的表达缺失(敲除)、降低其表达(敲减)或增加其表达(过表达)。

[0085] 在一些实施方式中,载体DNA或宿主细胞基因组中任何不需要的DNA通过使用位点特异性重组酶(例如,Cre、FLP、Dre、Vike或其组合)移除。

[0086] 将工程改造的核酸构建体导入细胞基因组中并移除载体DNA或不需要的DNA的方法可以单一步骤(例如,所有的构建体同时递送至宿主细胞)或以多个步骤(例如,各构建体顺序地递送至宿主细胞)进行。

[0087] 本公开还包括使用本文所述任何核酸酶的催化失活形式。例如,在一些实施方式中,可以使用如本文所提供的Cas9的催化失活形式(dCas9)或Cpf1的催化失活形式,其可以通过干扰转录来敲减基因表达。在一些实施方式中,dCas9(或Cpf1或其他核酸酶的催化失活形式)融合阻遏物肽(一种抑制转录的肽,例如,Cas9-KRAB (Urrutia 2003 *Genome Biol.* 4 (10):231))。在一些实施方式中,dCas9(或Cpf1或其他核酸酶的催化失活形式)融合激活物肽(一种激活或增强转录的肽,例如,Cas9-VP64 (Beerli等1998 *Proc Natl Acad Sci USA.* 95 (25):14628-33))。在一些实施方式中,dCas9(或Cpf1或其他核酸酶的催化失活形式)融合表观基因组调节物(epigenomic regulator)(例如,Cas9-DNMT或Cas9-p500)。在一些实施方式中,dCas9(或Cpf1或其他核酸酶的催化失活形式)融合FokI核酸酶,以在与两个共表达的gRNA同源的序列处生成双链断裂。

[0088] 在一些实施方式中,表达dCas9(或其它催化失活的核酸酶)的工程改造的核酸被用于对基因组中特定序列进行成像(参见,例如.,Chen B.,等(2013) *Cell* 155 (7):1479-1491,通过引用纳入本文)。

[0089] 在一些实施方式中,融合阻遏物肽(例如,Cas9-KRAB)或激活物肽(例如,Cas9-VP64)的野生型或未修饰的Cas9或Cpf1(或其它催化失活的核酸酶)与短gRNA联用以调节基因表达(参见,例如,Kiani S,等(2015)“工程改造Cas9gRNA用于基因组编辑、活化和抑制(Cas9gRNA engineering for genome editing,activation and repression).”*Nature Methods* (2015) [出版前电子公开],通过引用将其纳入本文)。

[0090] 在一些实施方式中,工程改造的核酸编码切口酶。切口酶是这样的酶,其在双链核酸中产生单链断裂。在一些实施方式中,切口酶是Cas9切口酶(Cong等2013 *Science* 339 (6121):819-823;Shen等2014 *Nature Methods* 11,399-402)。Cas9切口酶根据共表达的gRNA定义的靶序列上在特定位置产生单链DNA断裂(切口),而不是通过野生型酶产生双链DNA断裂(切割)。优选地,使用完整链作为模板,通过同源定向修复(HDR)在细胞中修复切口。HDR具有高保真度并且几乎不导致错误。两个相邻、相对链切口可以导致双链断裂

(DSB),并且引发易错非同源末端连接(NHEJ)修复;然而,在存在修复模板的情况下,可以通过HDR修复双切口。产生双切口通常降低不需要的脱靶作用。

[0091] 在一些实施方式中,工程改造的核酸编码重组酶。重组酶通常来自于细菌或真菌,其指向性地催化短(例如,30-40核苷酸)靶序列之间的敏感性DNA交换反应,所述短靶序列对各重组酶具有特异性。这些反应实现四个基本功能——缺失/插入、倒位、易位和盒交换,并且其可以单独或组合使用。如本文所提供使用的重组酶的示例包括但不限于,Cre重组酶、FLP重组酶、Hin重组酶(Dhar等2004 *Cell* 119(1):33-45;Sanders等2004 *Mol Biol* 340(4):753-66;Kamtekar等2006 *Proc Natl Acad Sci USA* 103(28):10642-7;Li等2005 *Science* 309(5738):1210-5)以及Tre重组酶(Sarkar等2007 *Science* 316(5833):1912-15)。

[0092] Cre是衍生自P1噬菌体的酪氨酸重组酶。该酶使用拓扑异构酶I样机制进行位点特异性重组。该酶(例如,38kDa)是位点特异性重组酶整合酶家族的成员,并且催化位于两个DNA识别位点(loxP位点)之间的位点特异性重组。这一约34碱基对(bp)的loxP识别位点包含两个约13bp的回文序列,其侧接约8bp的间隔物区域。Cre介导的重组在loxP位点的产物依赖于loxP位点的位置和相对方向。作为Cre介导的重组的结果,均包含loxP位点的两个单独的DNA物质可以进行融合。位于两个loxP位点之间的DNA序列被称为“floxed”。Cre介导的重组的产物依赖于loxP位点的取向。位于以相同方向取向的两个loxP位点的DNA被切成DNA圆环,而位于反相取向的两个loxP位点之间的间隔DNA被反转。Cre重组酶的功能不需要其辅因子(如ATP)或辅助蛋白。

[0093] FLP-FRT重组是与Cre-lox重组类似的定点重组技术。FLP-FRT重组包括通过衍生自面包酵母酿酒酵母2 μ m质粒的重组酶(FLP)重组位于短翻转酶识别靶标(FRT)位点之间的序列(Zhu等1995 *Journal of Biological Chemistry* 270(39):23044-54;Schlake等1994 *Biochemistry* 33(43):12746-12751;以及Turan等2010 *J.Mol.Biol.*402(1):52-69)。

[0094] 在一些实施方式中,工程改造的核酸编码碱基编辑酶。碱基编辑酶是这样的融合体,其融合DNA结合蛋白(例如,Cas9、TALE、ZF)至在DNA结合位点附近诱导碱基交换(例如,C交换成T)的特定效应物(参见,例如,Komor,A.C.,等*Nature*,2016;和Nishida,K.,等*Science* 353:6305,2016)。

[0095] 基因表达的控制

[0096] 本公开的工程改造的构建体(例如,“TOICas构建体”)允许基因组编辑的空间控制,基因组编辑的时间控制,以及空间和时间控制的组合。空间控制通常指在器官的特定组织中转录的激活。时间控制通常指在发育期间的特定时间激活转录。空间控制和/或时间控制可能是由例如以下所导致:来自使用细胞特异性或组织特异性启动子驱动核酸表达,来自对将效应物物质递送至细胞或器官以诱导或抑制核酸表达的时间的控制,或者来自其组合。在一些实施方式中,细胞特异性或组织特异性启动子驱动核酸的表达,以使其仅在细胞特化或细胞分化的特定阶段期间可操作地连接。在一些实施方式中,仅在细胞特化或细胞分化的特定阶段期间,向细胞或组织递送效应物物质(例如,Dox)。

[0097] 此外,本文公开的工程改造的构建体(例如,“TOICas构建体”)基本上降低渗漏基因表达。如果转录在不存在调节蛋白的情况下或以非受控的方式开始,基因表达系统可能会被认为是“渗漏的”。如本文其它地方所述,调节蛋白结合启动子以调节转录活性。如果表

达以不存在这样的调节蛋白的情况出现,基因表达被认为是“渗漏的”,所述调节蛋白旨在结合控制基因表达的启动子。如果不存在调节蛋白的基因表达的水平是存在条件蛋白的基因表达的水平的小于15% (例如,小于10%,小于5%,小于2%,小于1%,小于0.5%),基因表达被认为是“基本上无渗透的”。如图1和2所述,例如,至少一个绝缘子(例如,一个、两个或更多个绝缘子)位于编码调节蛋白的核酸以及控制酶(例如,核酸酶或重组蛋白酶)表达的下游启动子(P2) (调节蛋白与之结合)之间。“绝缘子”是阻断增强子和启动子之间相互作用的核苷酸序列。应当理解的是,在一些实施方式中,绝缘子可能位于驱动基因表达的任何两个启动子之间,从而防止在上游启动子转录激活时,下游启动子的转录激活。

[0098] 因此,在一些实施方式中,工程改造的构建体(例如,“TOICas构建体”)包括至少一个绝缘子。在一些实施方式中,工程改造的构建体包括至少2个、至少3个或至少4个绝缘子。在一些实施方式中,绝缘子包括哺乳动物绝缘子。例如,绝缘子可能包括(至少一个)人绝缘子,如5' HS5、DMD/ICR、BEAD-1、apoB (-57kb)、apoB (+43kb) 或DM1位点1或DM1位点2(表2)。在一些实施方式中,绝缘子可能包括(至少一个)小家鼠 (*Mus musculus*) 绝缘子,如BEAD-1、HS2-6或DMR/ICR。参见,例如,Bell等,Curr Opin Genet Dev.1999年4月;9(2):191-8; Science.2001年1月19日;291(5503):447-50;West等,Genes Dev.2002年2月1日;16(3):271-88;以及Ziebarth等,Nucleic Acids Research.2013;41(D1):D188-D194,其各自内容通过引用纳入本文。

[0099] 在一些实施方式中,绝缘子包括(至少一个)非哺乳动物绝缘子。例如,绝缘子可能包括(至少一个)黑腹果蝇 (*Drosophila melanogaster*) 绝缘子,如scs/scs'、gypsy、Fab-7、Fab-8、fa^{swb}或eve启动子。在一些实施方式中,绝缘体可能包括酿酒酵母 (*Scharomyces cerevisiae*) 绝缘子,包括HMR tRNA^{Thr},Chal UAS、UAS_{rpg}或STAR。在一些实施方式中,绝缘子可能包括(至少一个)原鸡 (*Gallus gallus*) 绝缘子,如Lys 5' A、HS4或3' HS。在一些实施方式中,绝缘子包括sns、普通海胆 (*Paracentrotus lividus*) 绝缘子、UR1、马粪海胆 (*Hemicentrotus pulcherrimus*) 绝缘子、或R0、非洲爪蟾 (*Xenopus laevis*) 绝缘子。

[0100] 表2

[0101]

5'HS5 (智人(<i>Homo</i>	CATCTTGGACCATTAGCTCCACAGGTATCT
------------------------	--------------------------------

[0102]

<i>sapiens</i>))	TCTTCCCTCTAGTGGTCATAACAGCAGCTT CAGCTACCTCTC (SEQ ID NO: 1)
apoB (-57 kb) (智人)	CAAATTATCCTGCCCCCTAGACATAACCTC CC (SEQ ID NO: 2)
BEAD-1(智人)	TGCATTGGCTGCCCAGGCCTGCACTGCCGC CTGCCGGCAGGGGTCCAGTCCACGAGACC CAGCTCCCTGC (SEQ ID NO: 3)
DM1 位点 1 (智人)	GCCGGCCGCGGACCCGGCCCCCTCCCTCCCC GGCCGCTAGGGGGCGGGCCCGGATCACAG GA (SEQ ID NO: 4)
DM1 位点 2 (智人)	CATGCACAAGAAAGCTTTGCACTTTGCGAA CCAACGATAGGTGGGGGTGCGTGGAGGAT GG (SEQ ID NO: 5)

[0103] “下游”和“上游”指核酸(例如,DNA或RNA)的相对位置。DNA或RNA的各条链具有5'端和3'端,据此命名脱氧核糖(或核糖)链上的碳位置。。按照惯例,上游和下游相对5'-3'方向处发生RNA转录。上游朝向RNA分子的5'端,下游朝向3'端。当考虑双链DNA时,上游朝向感兴趣基因编码链的5'端,下游朝向3'端。

[0104] 诱导型基因编辑系统

[0105] 诱导型基因编辑系统能够用于时间控制的、空间控制的、或时间控制且空间控制的基因修饰。一些实施方式利用Tet-关闭(Tet-Off)或Tet-开启(Tet-On)诱导型系统。其它实施方式利用他莫昔芬诱导型系统。而其它一些实施方式利用异丙基β-D-1-硫代半乳糖苷(IPTG)诱导型系统。

[0106] 四环素诱导型系统

[0107] Tet技术包括两个互补的控制电路,最初被描述为tTA依赖型(Gossen等Proc Natl Acad Sci USA.1992年6月15日;89(12):5547-51)以及rtTA依赖型(Gossen等Science.1995年6月23日;268(5218):1766-9)表达系统。它们现在常被称为Tet-关闭系统(tTA依赖型)和Tet-开启系统(rtTA依赖型)。在各系统中,重组四环素控制的转录因子(tTA或rtTA)与tTA/rtTA响应性启动子Ptet相互作用以驱动感兴趣基因的表达。表达受到效应物物质四环素(Tc)或其衍生物中的一种的调控。Tet-开启系统响应多四环素(Dox)。四环素在四环素控制的反式激活因子(tTA)与反式四环素控制的反式激活因子(rtTA)的转录因子的DNA结合水平作用。rtTA需要四环素配体用于DNA结合和转录。与之相反,tTA和DNA之间的相互作用被四环素阻止。因此,这两种形式的Tet系统差异性地响应四环素,并且可以互补的方式使用。

[0108] tTA是这样的杂化转录因子,其来自于原核Tet阻遏物TetR与真核转录反式激活结构域(例如,HSV VP16)的融合。TetR部分提供序列特异性DNA结合,对四环素的敏感性,以及与tTA融合蛋白的二聚化。相应地,TetR和tTA对四环素的响应是相似的:抗生素的结合低于其对其常见关联结合位点tet操纵子的亲和力。

[0109] rtTA与tTA在TetR中的几个点突变处不同。然而,这导致该转录因子四环素响应性的完全反转。rtTA需要四环素用于结合tetO。特定的四环素衍生物,诸如多四环素(Dox)或无水四环素(ATc)可以用于研究rtTA表型。

[0110] Ptet是响应tTA和rtTA两者的合成启动子。其包含融合至多聚化tetO序列的最小

RNA聚合酶II启动子(在不存在其它转录结合位点的情况下转录沉默)。这样的排布使得Ptet的活性依赖于tTA或rtTA的结合。设计这样的合成tTA/rtTA响应性启动子就以下而言是灵活的:最小启动子的起点以及操纵子的准确布置。例如,由融合至7个tetO序列阵列的CMV最小启动子组成的原始形式被称为Ptet-1。其被作为pTRE载体系列的部分销售(用于四环素响应性元件),一定程度上符合主流真核命名法。

[0111] 在一些实施方式中,多四环素(四环素衍生物)是用于Tet-开启或Tet-关闭系统的效应物物质。多四环素以高亲和力结合tTA以及rtTA,并且因此在例如tTA的情况中,其在低至1-2ng/ml的浓度时在Tet-关闭系统中完全有效。在Tet-开启系统中,在例如rtTA2-syn1的情况下,低至80ng/ml的浓度是有效的。

[0112] 在本公开的一些实施方式中,工程改造的构建体包括可操作地连接编码rtTA的核酸(例如,基因)的启动子(例如,CAG),其位于Ptet启动子的上游,所述Ptet启动子可操作地连接编码产生切割的酶(例如,Cas9核酸酶)的核酸,编码使核酸产生切口的酶(例如,Cas9切口酶)的核酸,或编码催化核酸交换的酶(例如,Cre重组酶)的核酸。通常,至少一个(例如,一个、两个或更多个)绝缘子位于编码rtTA的核酸和Ptet启动子之间,这样在不存在合适效应物物质(如多四环素)的情况下,激活编码rtTA的核酸的转录并不会另激活该酶的转录。

[0113] 在本公开的一些实施方式中,工程改造的构建体包括可操作地连接编码rTA的核酸(例如,基因)的启动子(例如,CAG),其位于Ptet启动子的上游,所述Ptet启动子可操作地连接编码产生切割的酶(例如,Cas9核酸酶)的核酸,编码使核酸产生切口的酶(例如,Cas9切口酶)的核酸,或编码催化核酸交换的酶(例如,Cre重组酶)的核酸。通常,至少一个(例如,一个、两个或更多个)绝缘子位于编码rtTA的核酸和Ptet启动子之间,这样在不存在合适效应物物质(如多四环素)的情况下,激活编码tTA的核酸的转录并不会另激活该酶的转录。

[0114] IPTG诱导型系统

[0115] 本公开的一些实施方式利用依赖于乳糖(lac)阻遏物蛋白和lac操纵子存在的系统。lac阻遏物是结合lac操纵子的DNA-结合蛋白,并且抑制可操作地连接lac操纵子的核酸的表达。异乳糖或异乳糖模拟物(如异丙基 β -D-1-硫代半乳糖苷(IPTG))的存在抑制lac阻遏物蛋白的DNA结合能力。通过lac阻遏物造成的DNA结合的缺失被用于lac操纵子的转录活化以及与该操纵子连接的任何核酸的表达。

[0116] lac操纵子包含三个结构基因,以及启动子,终止子,调控子,以及操纵子。三个结构基因是lacZ、lacY和lacA。lacZ编码 β -半乳糖苷酶(LacZ),其是将二糖乳糖切割成葡萄糖和半乳糖的胞内酶;lacY编码乳糖透性酶(LacY),其是以相同方向的质子梯度将 β -半乳糖苷泵入细胞的跨膜同向转移物;并且LacA编码半乳糖苷O-乙酰转移酶(LacA),其是将乙酰基团由乙酰辅酶A转移至 β -半乳糖苷的酶。

[0117] 在本公开的一些实施方式中,工程改造的构建体包括可操作地连接编码lac阻遏物蛋白的核酸(例如,基因)的启动子(例如,CAG),其位于lac操纵子的上游,所述lac操纵子可操作地连接编码产生切割的酶(例如,Cas9核酸酶)的核酸,编码使核酸产生切口的酶(例如,Cas9切口酶)的核酸,或编码催化核酸交换的酶(例如,Cre重组酶)的核酸。通常,至少一个(例如,一个、两个或更多个)绝缘子位于编码lac阻遏物蛋白的核酸和lac操纵子之间,这

样在不存在合适效应物物质(如IPTG)的情况下,激活编码lac阻遏物蛋白的核酸的转录并不会另激活该酶的转录。

[0118] 他莫昔芬诱导型系统

[0119] 本文还提供了工程改造的构建体,其包括(a)可操作地连接编码这样Cas9核酸酶的核酸的启动子,所述Cas9核酸酶并不包括核定位信号,其中,编码该酶的核酸侧接雌激素受体(ERT2)序列,以及(b)脱氧核糖核酸(DNA)-结合识别序列。核定位信号或序列(NLS)是这样的氨基酸序列,其“标记”通过核转运输入到细胞核的蛋白质。通常,该信号包含暴露于蛋白质表面的带正电荷的赖氨酸残基或精氨酸残基的一个或多个短序列。不同核定位的蛋白质可能共享相同的NLS。NLS具有与核输出信号相反的作用,其以将蛋白质输出核作为目标。不包括核定位信号的Cas9核酸酶是修饰的Cas9蛋白质,其没有通过核转运输入到细胞核中。

[0120] Cas9活性可以通过将核酸酶融合至雌激素受体(ERT2)经修饰的片段来进行调节。在一些实施方式中,ERT2是ER受体的修饰形式,其对他莫昔芬结合相对于内源性雌性激素具有高选择性。Cas9与经修饰的片段的融合被隔离在核的外部,不能指导重组。在存在雌激素受体拮抗剂(例如,他莫昔芬)的情况下,可以将Cas9重新定位到核中,其在那里能够发挥功能。在一些实施方式中,工程改造的核酸侧接ERT2片段(包括位于该蛋白质各末端的ERT2片段(ERT2Cas9ERT2;SEQ ID NO:7))。如果ERT2Cas9ERT2表达由阻止特异性启动子驱动(控制),感兴趣的基因可以在其发育的任何时间在感兴趣的组织中修饰。一旦记录该信息,可以在动物的整个生命周期监测该细胞。在一些实施方式中,工程改造的构建体包括编码不包括核定位信号的Cas9核酸酶的核酸,其中,编码该酶的核酸侧接雌激素受体(ERT2)序列(例如,SEQ ID NO:7)。

[0121] 核酸

[0122] “工程改造的构建体”指用于导入细胞的人工构建的核酸区段(线性或环状)。工程改造的构建体(例如,“TOICas构建体”)通常包含至少一个启动子,其可操作地连接编码感兴趣的蛋白质的核酸。本公开的实施方式提供了工程改造的核酸分子,其编码诱导型核酸表达系统的元件。“工程改造的核酸”是不存在于自然的核酸(例如,共价连接在一起的至少两个核苷酸,并且在一些实施方式中,包含磷酸二酯键,称之为磷酸二酯“主链”)。工程改造的核酸包括重组核酸和合成核酸。“重组核酸”是通过连接核酸(例如,分离的核酸、合成的核酸、或其组合)构建的分子,并且在一些实施方式中,其可以在活细胞中复制。“合成的核酸”是放大的或以化学方式或其它方法合成的分子。合成的核酸包括经化学修饰或以其他方式修饰,但是可以与天然产生的核酸分子碱基配对(也称为“结合”,例如,瞬时地或稳定的)的那些核酸。重组的和合成的核酸还包括产生自复制上述两种情况之一的那些分子。

[0123] 尽管工程改造的核酸作为一个整体不是天然产生的,但是其可以包括野生型核苷酸序列。在一些实施方式中,工程改造的核酸包括获得自不同生物(例如,获得自不同物种)的核苷酸序列。例如,在一些实施方式中,工程改造的核酸包括小鼠核苷酸序列、细菌核苷酸序列、人核苷酸序列、病毒核苷酸序列、或上述序列中任意两种或更多种的组合。

[0124] 在一些实施方式中,本公开的工程改造的核酸可以包括不同于磷酸二酯主链的主链。例如,在一些实施方式中,工程改造的核酸可以包括磷酰胺、硫代磷酸酯、二硫代磷酸酯、O-甲基亚磷酰胺连接、肽核酸或上述连接中任意两种或更多种的组合。工程改造的核酸

可以是单链或双链的,如本文具体所示,或者工程改造的核酸可以同时含有双链或单链序列的部分。在一些实施方式中,工程改造的核酸包含三链序列的部分。工程改造的核酸可以包括DNA(例如,基因组DNA、cDNA、或基因组DNA和cDNA的组合),RNA或杂化分子,例如,其中核酸包含脱氧核糖核苷酸和核糖核苷酸(例如,人工或天然的)的任意组合,以及两种或更多种碱基的任意组合,包括尿嘧啶、腺嘌呤、胸腺嘧啶、胞嘧啶、鸟嘌呤、肌苷、黄嘌呤、次黄嘌呤、异胞嘧啶、异鸟嘌呤等。

[0125] 本公开的工程改造的核酸可以使用标准分子生物学方法生产(参见,例如,Green和Sambrook,《分子克隆:实验室手册(Molecular Cloning, A Laboratory Manual)》,2012,冷泉港实验室出版社(Cold Spring Harbor Press))。在一些实施方式中,使用GIBSON ASSEMBLY®克隆生产核酸(参见,例如,Gibson, D.G.等Nature Methods, 343-345, 2009;和Gibson, D.G.等Nature Methods, 901-903, 2010,其各自通过引用纳入本文)。GIBSON ASSEMBLY®通常在单管反应中使用3种酶活性:5'外切核酸酶、DNA聚合酶的3'延伸活性以及DNA连接酶活性。5'外切核酸酶活性回嚼(chew back)5'短序列,并且将互补序列暴露用于退火。聚合酶活性随后填补退火区域的间隙。然后DNA连接物将切口封闭,并且将DNA片段共价连接在一起。邻近片段的重叠序列比在Golden Gate装配中使用的那些要长许多,并且因此导致较高比例的正确装配。生产工程改造的核酸的其它方法是本领域已知的并且可根据本公开使用。

[0126] 遗传因子

[0127] 本公开的工程改造的核酸可以包括一个或多个遗传因子。“遗传因子”指这样的核苷酸序列,其在核酸表达中起作用(例如,启动子、绝缘子、增强子、终止子和分子(例如,DNA或蛋白质)结合区域)或者编码核酸的产物(例如,编码调节蛋白的核苷酸序列或编码切割核酸的酶的核苷酸序列,编码使核酸产生缺口的酶的核酸,或编码催化核酸交换的酶的核酸)。

[0128] 工程改造的核酸的表达通常通过可操作地连接该工程改造的核酸的启动子驱动。“启动子”指这样一种核酸的控制序列,在其中控制核酸序列剩余物转录的起始和速率。启动子驱动转录或其调控的核酸序列的转录,因此,其通常位于或靠近基因转录起始位点。在一些实施方式中,启动子是100-1000个核苷酸长度。启动子还包括亚区域,其调节蛋白质和其它可能结合的分子,如RNA聚合酶和其它转录因子。启动子可以是组成型(例如,CAG启动子、巨细胞病毒(CMV)启动子)、诱导型(也称为激活型)、抑制型、组织特异型、发育阶段特异型或上述一种或多种的任意组合。

[0129] 当启动子相对于其调节(例如,控制(“驱动”)转录开始和/或表达序列)的核酸序列处于正确的功能性位置和方向时,可以认为该启动子是“可操作地连接”的。

[0130] 在一些实施方式中,启动子与核酸天然地相关联,并且可以通过分离位于给定核酸的编码区域上游的5'非编码序列获得。这种启动子称为“内源性”启动子。

[0131] 在一些实施方式中,启动子与核酸并不天然相关联。这种启动子称为“异源性”启动子,并且包括例如调节其它核酸的启动子以及获得自其它细胞的启动子。异源性启动子可以是合成的或重组的。在一些实施方式中,合成的异源性启动子包含获得自己知转录调控区域的各种元件。在一些实施方式中,合成的异源性启动子包含这样的突变,其通过本领域已知的遗传工程改造方法改变表达。在一些实施方式中,重组异源性启动子是这样产生

的：通过重组克隆、核酸扩增（例如，聚合酶链反应（PCR））、或重组克隆和核酸扩增的组合（参见，美国专利号4,683,202和美国专利号5,928,906）。本文考虑了产生合成和重组异源性启动子的其他方法。

[0132] 在一些实施方式中，启动子是诱导型启动子。当启动子受到相应调节蛋白的影响或接触时，“诱导型启动子”调节（例如，激活或灭活）其可操作地连接的核酸的转录活性。

[0133] 因此，本文所用“调节蛋白”是这样的蛋白质，其调节（例如，激活或灭活）来自启动子（例如，诱导型启动子）的转录活性。在一些所述方法中，调节蛋白直接结合诱导型启动子（例如，至启动子内的核苷酸的序列）。在一些实施方式中，调节蛋白结合位于诱导型启动子上游的区域（例如，位于诱导型启动子上游50-100个核苷酸）。在一些所述方法中，调节蛋白结合在诱导型启动子邻近（例如，附近）。调节蛋白的示例包括但不限于，四环素控制的反式激活因子（tTA）转录因子、反向四环素控制的反式激活因子（rtTA）转录因子、和Lac阻遏物蛋白。

[0134] 根据使用的系统，调整转录的调节蛋白可以激活或灭活转录。转录的激活可能涉及在启动子上直接作用以驱动转录，或通过灭活阻遏物元件（例如，阻遏物蛋白）来间接作用在启动子上，所述阻遏物元件阻止启动子驱动转录。相反，转录的灭活可能涉及在启动子上直接作用以阻止转录或通过激活作用于启动子的阻遏物元件来间接作用在启动子上。

[0135] 调节蛋白的活性常常有效应物物质调节。“效应物物质”是任何调整（例如，激活或灭活）调节蛋白活性的物质。效应物物质可以是内源性或外源性条件（例如，光或热）、化合物（例如，化学或非化学化合物）或调节（例如，直接或间接）调节蛋白活性的其它蛋白质。

[0136] 例如，在Tet-关闭系统（也称之为tTA依赖型系统）和Tet-开启系统（也称之为rtTA依赖型系统）中，重组四环素控制的转录因子（tTA或rtTA）（“调节蛋白”）与tTA/rtTA诱导型启动子Ptet相互作用以驱动可操作地连接启动子的基因的表达。基因表达受到效应物物质四环素或其衍生物中的一种的调控。四环素在tTA和rtTA转录因子的DNA结合水平作用。rtTA需要四环素配体用于DNA结合，并且因此转录。相反，tTA和DNA之间的相互作用被四环素阻止。因此，Tet系统的两种形式Tet-关闭系统和Tet-开启系统对于四环素（以及其衍生物，如多四环素）响应不同。

[0137] 通常，效应物物质的给予或去除将导致可操作连接的核酸序列的转录在激活或灭活之间转换。因此，可操作地连接核酸序列的启动子的活化状态指当该启动子积极地调节核酸序列转录（例如，表达连接的核酸序列）的状态。相反，可操作地连接核酸序列的启动子的灭活状态指当该启动子不积极地调节核酸序列转录（例如，不表达连接的核酸序列）的状态。

[0138] 调节诱导型启动子（例如，经由调节蛋白的调节）的效应物物质的示例包括但不限于，生理条件，如光、pH、温度、辐射、渗透压、盐梯度和细胞表面结合中的改变。诱导型启动子还可以通过改变外在或内在效应物物质的浓度来调节。外在效应物物质的示例包括但不限于，氨基酸和氨基酸类似物、糖和多糖、核酸、蛋白质转录激活物和阻遏物、细胞因子、毒素、基于石油的化合物、包含金属（例如，铜）的化合物、盐、离子、酶底物类似物、激素或上述任意两种或多种的组合。其它效应物物质是本领域已知的并且可根据本公开使用。

[0139] 诱导型启动子的示例包括但不限于，化学或生物化学调节的和物理调节的启动子，如醇调节的启动子，四环素调节的启动子（例如，无水四环素（aTc）响应性启动子和其他

四环素响应性启动子系统,这包括四环素阻遏物蛋白(tetR)、四环素操纵子序列(tetO)和四环素反式激活因子融合蛋白(Tta)),类固醇调节的启动子(例如,基于大鼠糖皮质激素受体的启动子,人雌激素受体,蛾蛻皮激素受体,以及来自类固醇/类维生素A/甲状腺受体超家族的启动子),金属调节的启动子(例如,来自酵母、小鼠和人的金属硫蛋白(结合和隔离金属离子的蛋白)基因的启动子),发病原调节的启动子(例如,由水杨酸、乙烯或苯并噻二唑(BTH)诱导的,温度/热诱导型启动子(例如,热休克启动子)和光调节的启动子(例如,来自植物细胞的光响应性启动子)。其它诱导型启动子是本领域已知的并且可根据本公开使用。

[0140] 增强子

[0141] 在一些实施方式中,工程改造的核酸包括增强子。“增强子”是顺式作用核苷酸调控序列,其涉及可操作地连接启动子的核酸序列的转录活化。增强子可以位于启动子上游或下游的任何功能性位置。

[0142] 终止子

[0143] 在一些实施方式中,工程改造的核酸包括终止子。“终止子”是导致转录停止的核苷酸序列。终止子可以是单向的或双向的。终止子包括这样的DNA序列,其涉及通过RNA聚合酶特异性终止RNA转录本,并且防止下游核酸序列被上游启动子转录活化。

[0144] 最常用类型的终止子是正向终止子。当将正向转录终止子置于通常转录的核酸序列下游时,正向转录终止子将会导致转录中止。在一些实施方式中,使用了双向转录终止子,其通常导致在正向和反相链上的转录终止。在一些实施方式中,提供了法向转录终止子,其通常仅终止在反相链上的转录。

[0145] 根据本公开使用的终止子的示例包括但不限于,基因的终止序列,例如,牛生长激素终止子,和病毒终止序列,例如,存在于细菌系统中的T0终止子、TE终止子、 λ T1和T1T2终止子。在一些实施方式中,终止信号可能是不可以被转录或翻译的信号,诸如来自序列截短的那些。

[0146] 选择性标志物

[0147] 在一些实施方式中,工程改造的构建体(例如,“TOICas构建体”)包括编码选择性标志物蛋白的核酸。选择性标志物是导入细胞的基因,其赋予适合人工选择的特征。例如,选择性标志物可以是抗生素抗性基因。抗生素抗性基因非限制性的示例包括编码对氨苄青霉素、氯霉素、四环素或卡那霉素抗性的基因。例如, β -内酰胺酶赋予对细菌宿主的氨苄青霉素抗性,获自Tn5的neo基因赋予细菌中的卡那霉素以及真核细胞中的遗传霉素的抗性,获自大肠杆菌基因组的突变FabI基因(mFabI)赋予对宿主的三氯生抗性,并且获自酵母的乳清苷-5'磷酸脱羧酶URA3是阳性和阴性选择性标记物。

[0148] 载体

[0149] 本公开的实施方式提供了包含工程改造的核酸分子的载体,所述工程改造的核酸编码诱导型核酸表达系统的元件。“载体”指用做载剂的核酸(例如,DNA)以将遗传物质(例如,工程改造的核酸)运送到细胞中,在其中其可以例如复制和/或表达。应当理解的是,本文所用属于“载体”不包括慢病毒载体。因此,在一些实施方式中,本公开的工程改造的核酸是被克隆到载体中,或经由载体(非慢病毒载体)递送至细胞。提供使用的载体通常是经工程改造的,并且包括附加型表达载体。附加型表达载体的示例包括但不限于,质粒,其是能

够在宿主细胞中自动复制的双链常环状的DNA序列,以及病毒载体,其可以基于来自DNA病毒的序列,如BK病毒、牛乳头瘤病毒1和EB病毒。

[0150] 本文也提供了杆状病毒载体(Kost等2005 Nat Biotechnol.2005年5月;23(5):567-575)。杆状病毒基因表达系统和基因递送系统是已知的(参见,例如,Makela AR等Cold Spring Harb Protoc.2010年3月;2010(3),通过引用纳入本文),并且可以根据本公开使用。在一些实施方式中,使用杆状病毒载体将如本文所提供的TOICas构建体递送至细胞,如哺乳动物细胞。

[0151] 载体还可能是工程改造的细菌人工染色体(BAC)(O'Conner M.等1989Science 244(4910):1307-1312;Shizuya H.等1992Proc Natl Acad Sci USA 89(18):8794-8797;和Shizuya H等2001 Keio J Med.50(1):26-30,其各自通过引用纳入本文)或酵母人工染色体(YAC)(Struhl K等1979 PNAS 76(3):1035-39,通过引用纳入本文),其中,例如编码调节蛋白的核酸、编码核酸酶的核酸或编码重组酶的核酸被置于诱导型启动子(例如,组织特异性启动子)或管家基因启动子控制下。在一些实施方式中,存在于BAC或YAC中的启动子可调节下游基因。

[0152] 在一些实施方式中,包括工程改造的核酸的载体(或工程改造的核酸本身)比典型的表达构建体大。在一些实施方式中,本公开的工程改造的核酸(或包括工程改造的核酸的载体)是至少12千碱基(kb)。例如,工程改造的核酸(或包括工程改造的核酸的载体)可以是至少13kb、至少14kb、至少15kb、至少16kb、至少17kb、至少18kb、至少19kb或至少20kb。在一些实施方式中,工程改造的核酸(或包括工程改造的核酸的载体)是15kb-20kb、15kb-30kb、15kb-40kb、20kb-30kb或20kb-40kb。

[0153] 细胞

[0154] 本公开的工程改造的构建体(例如,“TOICas构建体”)可以导入各种不同的细胞中。工程改造的构建体可以被导入的细胞的示例包括但不限于,哺乳动物细胞、昆虫细胞、细菌细胞和酵母细胞。例如,哺乳动物细胞可以是人细胞、灵长类动物细胞(例如,vero细胞)、大鼠细胞(例如,GH3细胞、C23细胞)或小鼠细胞(例如,MC3T3细胞)。存在各种人细胞系,其包括但不限于HEK细胞(例如,HEK 293或HEK 293T细胞)、HeLa细胞、来自国家癌症研究所60癌症细胞系(NCI60)的癌症细胞、DU145(前列腺癌症)细胞、Lncap(前列腺癌症)细胞、MCF-7(乳腺癌)细胞、MDA-MB-438(乳腺癌)细胞、PC3(前列腺癌症)细胞、T47D(乳腺癌)细胞、THP-1(急性骨髓性白血病)细胞、U87(恶性胶质瘤)细胞、SHSY5Y人神经母细胞瘤细胞(克隆自骨髓瘤)和Saos-2(骨癌)细胞。

[0155] 在一些实施方式中,工程改造的构建体表达在干细胞中(例如,人干细胞),如多能干细胞(例如,人多能干细胞,其包括人诱导型多能干细胞(hiPSCs))。“干细胞”指在培养中不确定的时期具有分裂能力且产生特化细胞的细胞。“多能干细胞”指这样类型的干细胞,其能够分化成器官的所有组织,但是其单独不能够持续完整器官发育。“人诱导型多能干细胞”指这样的体(例如,成熟或成体)细胞,其已经通过被强制表达对于维持胚胎干细胞限定特性重要的因子和基因来重编程至胚胎干细胞样状态(参见,例如,Takahashi和Yamanaka,2006 Cell 126(4):663-76,通过引用纳入本文)。人诱导型多能干细胞表达干细胞标志物,并且能够生产所有三胚层(外胚层、内胚层、中胚层)的细胞特性。

[0156] 可以更加本公开使用的细胞系的其它非限制性示例包括293-T、293-T、3T3、4T1、

721、9L、A-549、A172、A20、A253、A2780、A2780ADR、A2780cis、A431、ALC、B16、B35、BCP-1、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C2C12、C3H-10T1/2、C6、C6/36、Cal-27、CGR8、CHO、CML T1、CMT、COR-L23、COR-L23/5010、COR-L23/CPR、COR-L23/R23、COS-7、COV-434、CT26、D17、DH82、DU145、DuCaP、E14Tg2a、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、Hepalcl7、High Five细胞、HL-60、HMEC、HT-29、HUVEC、J558L细胞、Jurkat、JY细胞、K562细胞、KCL22、KG1、Ku812、KY01、LNCap、Ma-Me1 1、2、3...48、MC-38、MCF-10A、MCF-7、MDA-MB-231、MDA-MB-435、MDA-MB-468、MDCK II、MG63、MONO-MAC 6、MOR/0.2R、MRC5、MTD-1A、MyEnd、NALM-1、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NW-145、OPCN/OPCT Peer、PNT-1A/PNT 2、PTK2、Raji、RBL细胞、RenCa、RIN-5F、RMA/RMAS、S2、Saos-2细胞、Sf21、Sf9、SiHa、SKBR3、SKOV-3、T-47D、T2、T84、THP1、U373、U87、U937、VCaP、WM39、WT-49、X63、YAC-1和YAR细胞。

[0157] 在一些实施方式中，本公开的工程改造的构建体被导入人细胞，例如，淋巴细胞，如T细胞（例如，CD8⁺细胞、CD4⁺细胞）、B细胞或自然杀伤细胞（NK细胞）。

[0158] 在一些实施方式中，本公开的细胞是修饰的。修饰的细胞是这样的细胞，其包含外源性核酸或自然不存在的核酸。在一些实施方式中，修饰的细胞包含基因组核酸中的突变。在一些实施方式中，修饰的核酸包含外源性独立复制的核酸（例如，存在于附加型载体的工程改造的核酸）。在一些实施方式中，修饰的细胞通过将外来或外源性核酸导入细胞产生。

[0159] 工程改造的构建体可以通过诸如下述方法导入细胞：例如，电穿孔（参见，例如，Heiser W.C.转录因子方案：分子生物学方法TM 2000 (Transcription Factor Protocols: Methods in Molecular BiologyTM 2000)；130:117-134)、化学法（例如，磷酸钙或脂质法）、转染（参见，例如，Lewis W.H.,等,Somatic Cell Genet.1980年5月；6 (3) :333-47;Chen C.,等,Mol Cell Biol.1987年8月；7 (8) :2745-2752)、与包含重组质粒的细菌原生质体融合（参见，例如，Schaffner W.Proc Natl Acad Sci USA.1980年4月；77 (4) :2163-7)、或将纯化的DNA直接显微镜下注射到细胞的核内（参见，例如，Capecchi M.R.Cell.1980年11月；22 (2Pt 2) :479-88）。

[0160] 经修饰以包括本公开的工程改造的构建体的哺乳动物细胞（例如，人细胞）可以使用传统哺乳动物细胞培养方法培养（例如，维持在细胞培养中）（参见，例如，Phelan M.C.Curr Protoc Cell Biol.2007年9月；第1章:1.1单元，通过引用纳入本文）。例如，细胞可以在细胞培养箱中以适当的温度和气体混合物（例如，37℃、5%CO₂针对哺乳动物细胞）生长和维持。培养条件对于各种细胞类型可能不同。例如，细胞培养基的pH、葡萄糖浓度、生长因子和存在的其它营养素可能不同。用于补充培养基的生长因子通常来自自动物血液的血清，如胎牛血清（FBS）、小牛血清和/或猪血清。在一些实施方式中，本文所提供的使用的培养基可以是市售可得和/或已详细描述（参见，例如，Birch J.R.,R.G.Spier (Ed.)细胞技术百科全书 (Encyclopedia of Cell Technology), Wiley.411-424,2000;Keen M.J.Cytotechnology 17:125-132,1995;Zang,等Bio/Technology.13:389-392,1995)。在一些实施方式中，使用化学限定的培养基。

[0161] 本公开的诱导型基因组编辑系统允许空间控制的和时间控制的基因组修饰，这可以用于各种应用。例如，起可以被用于生成动物模型、细胞系和诱导型多能干细胞（iPSC），其中在分化阶段必需严格控制靶基因的调节。诱导型基因组编辑细胞还可以用于生长组织

特异性诱导型敲除或敲入。其它应用包括基因治疗。

[0162] 实施例

[0163] 实施例1:插入诱导型盒的方案

[0164] 1. 培养细胞,并且在转染前一天将约500000个细胞分至6孔板的新孔中。

[0165] 2. 以靶向核酸酶结合位点基因座的ZFN和TOICas诱导型构建体并使用适当的转染剂(通常是Lipofectamine LTX或Fugene HD(通常以3:1的ZFNs/TOICas比例)转染细胞。

[0166] 3. 转染后3天,将细胞分至20cm孔内,并利用适当的选择(G418针对dox诱导型、嘌呤霉素针对C9Ert2)进行选择。

[0167] 4. 选择后3周,挑取克隆并且针对安全港基因座中的整合进行检查。

[0168] 5. gRNA可以慢病毒感染或用诱导型构建体再次靶向基因座的方式插入或瞬时转染。

[0169] 6. 通过向培养基添加多四环素(10 μ g/ml)或他莫昔芬(0.5 μ M)进行诱导。

[0170] 7. 诱导后3天评价切割。

[0171] 实施例2:使用TOICas构建体生成细胞系

[0172] 人诱导型多能干细胞(hiPSC)是研究多步骤分化过程的宝贵工具,并且用于生成基于细胞的体外疾病模型。它们还被用于化学筛选和细胞治疗。这些应用中的每一种都可以通过基因组工程显著增强;特别是感兴趣基因的诱导型敲入或敲减促进了对于其功能在不同分化阶段或不需要使用动物模型的疾病进程的研究。

[0173] 使用严格调控的一个载体系统将TOICas构建体(图10)靶向位于hiPSC中的人安全港AAVS1基因座以获得多四环素(Dox)诱导型Cas9。在存在Dox的情况下,反式激活因子(3G)结合响应元件,并且驱动感兴趣的蛋白质Cas9T2AGFP的表达。

[0174] 生成的iPSC系显示在不存在Dox的情况下无可检测的GFP表达,而在存在Dox的情况下,检测到清楚的GFP信号(数据未显示)。可以用低至5ng/ml的Dox检测GFP,并且在40小时的周期内信号随着渐增的浓度(0ng/ml、1ng/ml、5ng/ml、10ng/ml、50ng/ml和100ng/ml)增强,这表明系统可以如所需调整(未显示荧光数据图像)。在不存在Dox或靶向引导RNA的情况下,gRNA靶基因座上并不存在可检测的基因组切割(通过入缺失标记指示),然而在存在Dox和引导RNA的情况下,观察到高效切割,这表明该系统是高效的并且没有Cas9的渗漏表达(图11)。

[0175] 可以使用本公开的方法插入构建体的单个或多个拷贝。较高拷贝数对于较高蛋白质(例如,GFP)表达水平。

[0176] 染色体组型在用于分化过程的细胞中特别重要,如诱导型多能干细胞(iPSC)和小鼠胚胎干细胞(mESC)。对来自人诱导型多能干细胞系的固定细胞悬浮液的染色体分析显示在检测的20个细胞中的明显正常染色体组型,这表明如本文所提供的整合方法并不干扰包含诱导型Cas9系统的诱导型多能干细胞中的多能性(数据未显示)。

[0177] Dox处理细胞后进行Cas9诱导的时间-过程,其在6小时后显示了可检测的GFP水平并且持续了至少40小时,在第30小时观察到最大表达(数据未显示)。

[0178] 使用相同的TOICas系统,还生成了表达Cas9变体的诱导型多能干细胞(iPSC)系,所述Cas9变体是酶失活的(Cas9死亡)。当与相同TOICas系统中的Cas9系比较时,仅在系统过表达时(100ng/mL,表达2周)观测到Cas9的毒性。毒性通过iPSC系分化的诱导来指示,如

通过形态学改变以及通过多能性基因(例如,Oct4)表达的降低所指示的。这表明Cas9的核酸酶活性与诱导型多能干细胞中的毒性相关联。

[0179] 接着,基因敲除效率在iPSC或衍生自iPSC的细胞中评估。高效基因敲除通过转染在表达质粒中编码的合成gRNA、合成crRNA/tracrRNA和gRNA而获得。通过使用靶向附近序列(相距约10-400bp)的2个或更多个gRNA构建体,效率和缺失精度有所增加。

[0180] 实施例3:人源性大麻素受体缺失

[0181] 内源性大麻素(eCB)在不同范围的生理或病理过程中有重要作用,其中包括神经发育、免疫功能、疼痛、能量平衡、脂质和葡萄糖代谢。尽管内源性大麻素需求以及在小鼠中详细地研究过,但是这些小信号转导脂质在人神经元发育和病理中的作用却极为少知。

[0182] 为了研究在人系统中内源性大麻素信号转导的需求,使用TOICas生成诱导型多能干细胞(hiPSC)库,其中人内源性大麻素受体1和2基因(CNR1和CNR2)被单独或组合靶向(分别为单个或两个敲除)。

[0183] hiPSC中进行单个DoxA诱导后,在Cas9存在的情况下使用双引导RNA在CNR1和2基因中实现精确缺失(数据未显示)。

[0184] 细胞培养1周后,传代细胞,并且提取基因组DNA。CNR1和CNR2编码序列中的精确缺失通过连接PCR、细胞试验和Sanger测序确认(数据未显示)。

[0185] 实施例4:针对CRISPR KO和CRISPRi的TOICAS-KRAB

[0186] 生成包含TOIC修饰形式的iPSC细胞系。正常的Cas9融合至KRAB结构域。使用具有正常/常规原型间隔子长度的gRNA或具有短原型间隔子长度的gRNA分别诱导CRISPR KO或CRISPRi(参见,例如,Kiani,S.,等Nature methods 12(11):1051-1054,2015中的方法)。该系统还用于靶向Oct4基因,并且通过使用如上所示的替代性gRNA实现Oct4基因的敲除或敲减。该系统一般适用于其它修饰,如CRISPRa,碱基编辑和适用于其它正交Cas9酶。

[0187] 实施例5:在具有免疫能力的小鼠中生成功能性小鼠模型

[0188] 用于离体原代细胞建模的Tet-开启诱导型Cas9(iCas9)小鼠。CRISPR-Cas9是用于遗传干扰的模块化且通用工具。其由核糖核蛋白内切核酸酶(Cas9)和引导RNA(gRNA)组成。培养的细胞对于CRISPR-Cas9工程是相对可修改的。然而,应用于活体动物产生了一些重大挑战,鉴于将所有组件精确且同时递送到感兴趣的组织中的复杂性。此外,常规的敲除策略会影响动物中的每种细胞,因此,在复制表型中区分主要和次要改变常常是不可能的。最后,任何基因的改变在小鼠发育上都具有潜在的后果,这使得在成体动物上的研究被妨碍或被复杂化(例如,胚胎致死性、畸型、代谢缺陷、癌症)。因此,存在对用于体内精确且时间编辑系统的日益增长的需求。

[0189] 本文提供的是一种严格的一组件系统,以在没有其它任何表型序列下在具有免疫能力的小鼠中实现Cas9的多四环素诱导型表达。将本公开的工程改造的TOICas构建体(图10)插入小鼠胚胎干细胞(mES)中的安全(safe) r26基因座中。以r26基因座中的单个或多个拷贝整合成功生成克隆。多四环素诱导且进一步确认Cas9表达后,选择不同的克隆注入胚胎以产生嵌合小鼠(数据未显示)。将dox处理后的Cas9的表达与r26基因座中TPICas构建体的拷贝数量关联。将多个拷贝导入特定基因座只有通过使用用于整合的基于NHEJ的方法才可以实现。这是通过该方法产生的第一个动物模型。

[0190] 为了证明TPICas在小鼠细胞中的功能性,分离来自初创小鼠耳成纤维细胞,并且

进行离体培养。然后使用Dox刺激细胞。在荧光显微镜下成像后,确认GFP表达(数据未显示)。

[0191] 用于AAV介导的体内基因组编辑和癌症建模的Tet-开启诱导型Cas9 (iCas9) 小鼠。CRISPR-Cas9将基因组编辑的边界移动至复杂的体内应用。CRISPR对于基因治疗的快速途径的前景受制于对更加可译且有效动物模型逐渐增长的需求。小鼠分子遗传学具有复制人疾病的潜能,但是常规的敲除策略不具有选择性,也不具有足以空间或时间方式控制的通用性。最近几年,对于更加灵活的体内基因工具的需求急剧增长,特别是在成体小鼠中复杂建模的领域。

[0192] Cas9表达在T0ICas小鼠中被严格调控,并且只有在多四环素处理后,才在大部分组织中观察到Cas9表达。T0IC小鼠具有免疫能力背景,因此可以用于例如免疫-肿瘤学应用。因此,向具有免疫能力的T0ICas小鼠(一组6只)给予溶于饮用水的2mg/ml的多四环素(Dox),自由摄取,持续3天。处理后,将T0ICas小鼠与未处理的(-Dox) T0ICas对照小鼠(一组6只)一起进行安乐死。立即分离来自两种小鼠的6个组织,并且用落射荧光立体显微镜成像,以便基准Cas9表达(数据未显示)。Dox诱导3天后,在心脏、肺、气管、肝脏、脾脏、胰腺和大脑中观测到体内T0ICas活化/GFP表达,然而并没有在来自未处理的对照小鼠的组织中检测到GFP(数据未显示)。组织病理学分析显示小鼠肿瘤与人肿瘤类似。可以在肺的所有区域观测到低分化腺瘤。生长模式包括显示乳头状、小梁状、腺状和紧凑生长模式的上皮肿瘤细胞,以及在肺泡腔和邻近组织中的局部浸润。观测到偶尔存在的近乎鳞状上皮样的转换。肿瘤组织中以及存在典型有丝分裂的少数支气管区域出现高有丝分裂活性。同样存在具有多倍性、各向异性和异型核型以及显著增加的肺泡巨噬细胞的高度分化的细胞。没有血管浸润或转移的迹象。观察到支气管肺泡的生长模式以及通过肺泡和细支气管间隙中多灶性结节所鉴定的小梁和腺体生长模式的区域。整体观察到高有丝分裂活性,特别是在所有肿瘤结节(基于针对增殖标志物Ki67的免疫组织化学试验)以及随机的支气管上皮细胞中。令人惊讶的是,被我们视为阴性对照的脑(通过药物治疗最难以到达的器官)也呈GFP阳性。

[0193] 为了生成体内癌症小鼠模型,使用了针对Trp53和Kras的gRNA以及模板以导入KrasG12D突变。用带有靶向Kras的gRNA的AAV感染细胞(图12,顶图)。通过GFP表达的时间-过程监测来评估Cas9内切核酸酶活性(图13)。在收集自感染4或9周后细胞的基因组DNA上进行的Surveyor试验确认由靶向策略所预期的独特模式表示的精确编辑(插入缺失和点突变-图14)。在纤维母细胞和成年小鼠中确认了p53敲除(KO)和kras诱变两者。

[0194] 然后,用AAV经由气管内注射来感染T0ICas小鼠,所述AAV具有靶向Trp53基因的引导RNA(图12,底图)。感染前24小时,在#37、#42和#52小鼠中激活Cas9,而不处理#98小鼠#99(-Dox)。感染4周后,从所有小鼠处收集肺部,并且提取DNA以进行Surveyor核酸酶试验。如图15所示,然后在经处理和感染的小鼠中确认Trp53基因中精确突变的存在。

[0195] 实施例6:组织特异性敲除

[0196] 针对本公开T0ICas系统中CRISPR的空间和时间调控研发了若干种策略。一种策略是在组织特异性启动子下驱动rtTA的表达。使用 α 肌球蛋白重链(α MHC)驱动rtTA的表达并且因此调节Cas9;该策略通常适合使用任何其它组织启动子,并且只在特定组织中诱导Cas9的表达。

[0197] 另一策略是引入由p0111启动子(例如,组织特异性)驱动的组织特异性转录本中

的gRNA盒。gRNA被整合到内含子中,并且可以具有核酶(rybozyme)或后续待由原始转录本切割的其它RNA加工序列。

[0198] 再一个策略是基于在组织特异性重组酶或组织特异性正交Cas9/CRISPR诱导后的gRNA激活。在该情况中,gRNA的表达是通过停止盒(stop cassette)阻止,该停止盒位于启动子和gRNA的功能性部分之间。正交Cas9/CRISPR或位点特异性重组酶的存在诱导停止盒的去除,因此使得gRNA仅在正交Cas9/CRISPR或位点特异性重组酶表达的组织中表达。

[0199] 上述实施例实现了空间和时间调节Cas9/CRISPR系统,因此,有可能在成体动物以及分化的细胞中产生组织特异性敲除、敲减和碱基编辑,由胚胎干细胞或诱导型多能干细胞开始。

[0200] TOICas系统的一个应用是感兴趣组织或细胞的诱导型消融,以研究细胞/组织的功能或模拟疾病的状态。在该情况中,TOICas系统中Cas9的时间表达(通过Dox调节)与gRNA的组织特异性表达组合,所述gRNA靶向靶基因组中多个重复。作为靶位置的重复的两个非限制示例是小鼠基因组中的重复B1 (CTCACTATGTAGACCAGGC (SEQ ID NO:10)),以及在人基因组中的重复A1uI (CCTGTAATCCCAGCACTTTCACTTTGGGAGGCCGAGGCGAGTCTCGCTCTGTCGCCC (SEQ ID NO:11))。系统的组织特异性和时间性活化促进在多个位点的切割,因此靶基因组的降解只在TOICas系统被表达的组织内且只在Dox处理的情况下。

[0201] 实施例7:gRNA克隆方法

[0202] 开发了三种策略以在表达Cas9的细胞或质粒中克隆针对特定基因的gRNA。

[0203] 第一种策略使用重组工程以插入这样的盒,所述盒包含位点特异性gRNA,以及细菌选择标志物和驱动gRNA表达的polIII启动子。通过在细菌中选择接受gRNA和选择标志物的细胞,gRNA可以高通量方式插入。

[0204] 第二种策略是使用来自脑膜炎奈瑟菌(*Neisseria Meningitis*)的Cas9以靶向包含TOICas构建体的细胞中的neo基因,并且导入gRNA,以及neo选择标志物内的哺乳动物选择标志物。该策略并不限于neo选择标志物以及Cas9NM,因为gRNA整合可以使用任何其它正交Cas9和插入点。

[0205] 第三种策略是使用Cas9核糖核蛋白(Cas9 RiboNucleoProtein)以通过同源重组方法或基于NHEJ的方法在精确的基因座中递送gRNA。其它方法是通过慢病毒和转座子插入。

[0206] 实施例8:TOICas系统的进一步分析

[0207] 如图16-19所示,尽管在不存在多四环素的情况下偶尔会有表达Cas9的细胞,但是Cas9诱导非常干净。仅在多四环素暴露24小时后,表达最大量的Cas9。此外,在A549细胞中,Cas9表达的诱导似乎以下作用:细胞生长、DNA损伤的发生率(通过gH2AX焦点测量)、或DNA损伤的持续性(通过微核形成测量)。通常,Cas9表达似乎不增强对于IR的敏感性,也不增强IR所导致的DNA损伤。

[0208] 序列

[0209] 用于根据本公开使用的哺乳动物绝缘子序列的其它示例:

[0210]

CTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCGCTAGGGGCAGCAGCGAGCCGC
CCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGA

AGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGA
AAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTGCAAAGGAGCACAGTGCTC
ATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGC
CTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGGGAA
AAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTATCAAGGGGGA
GTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGATCACGGAGAGGCAGATCTTGGGGA
TAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTT
CAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACA
AAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGG
AAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTCA
TACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCA
CTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGG
CAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCACCTCCCCTGGCACAAC
GTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCT (SEQ ID NO:6)

[0211] Cas9ERT2:

[0212]

ATGGCTCTCGAGCCATCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAGCCCGCTCATGATCAAACGCTCTAA
GAAGAACAGCCTGGCCTTGTCCCTGACGGCCGACCAGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCT
ATTCCGAGTATGATCCTACCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAG
CTGGTTCACATGATCAACTGGGCGAAGAGGGTGCCAGGCTTTGTGGATTGACCCTCCATGATCAGGTCCACCTTCT
GGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCTATGGAGCACCCAGTGAAGCTACTGTTTG
CTCCTAACTTGCTCTTGACAGGAACCAGGGAAAAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCT
ACATCATCTCGGTTCCGCATGATGAATCTGCAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAATTC
TGGAGTGACACATTTCTGTCCAGCACCCCTGAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCTGGACAAGA
TCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTGACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTC
CTCCTCATCCTCTCCACATCAGGCACATGAGTAACAAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGT
GGTGCCCCTCTATGACCTGCTGCTGGAGGCGGCGGACGCCCACCGCCTACATGCGCCCACTAGCCGTGGAGGGGCAT
CCGTGGAGGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTCTACTTCATCGCATTCCTTGCAAAAGTATTACATC
ACGGGGGAGGCAGAGGGTTTCCCTGCCACAGCTGACAAGAAATACTCAATCGGGCTGGACATCGGAACCTAACTCAGT
GGGGTGGGCAGTCATTACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACT
CCATCAAGAAAAATCTGATTGGGGCCCTGCTGTTGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACT
GCTAGAAGGCGCTATACCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGT
GGACGATAGTTTCTTTCATCGGCTGGAGGAATCATTCCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCT
TTGGAAACATTGTGGACGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGAC
TCTACAGATAAAGCAGACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGAT
CGAGGGAGATCTGAACCCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATAACAATCAGCTGT
TTGAGGAAAACCTATTAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGG
CTGGAGAACCTGATCGCTCAGCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGAAATCTGATTGCACTGTCACTGGG
ACTGACACCTAACTTCAAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACG

ATGACCTGGATAACCTGCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCCCTGGCCGCTAAGAATCTGTCTGAC
GCTATCCTGCTGAGTGATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCTCTGTCTAGCTAGCATGATCAAGAG
ATATGACGAGCACCATCAGGATCTGACCCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAA
TCTTCTTTGATCAGAGTAAGAACGGGTACGCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTT
ATCAAACCTATTCTGGAGAAGATGGACGGCACCGAGGAAGTCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAA
GCAGCGGACATTTGATAACGGCTCCATCCCCACCAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGG
AAGACTTCTACCCATTTCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCTTCCGGATCCCTTACTATGTG
GGGCCCCCTGGCAAGGGGTAATTCCCGCTTTGCCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGAACTT
CGAGGAAGTGGTCGATAAGGGAGCTTCCGCACAGTCTTTCATCGAGAGAATGACAAACTTCGACAAAAACCTGCCAA
ATGAGAAAAGTGCTGCCTAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATAC
GTCACCGAGGGGATGAGGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTTAAAC
CAATCGCAAGGTGACAGTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAATTGAATGTTTCGATTCTGTGGAGATCA
GTGGCGTCGAAGACAGATTTAACGCTTCTCTGGGAACCTACCAGATCTGCTGAAGATCATTAAGGATAAAGACTTC
CTGGACAACGAGGAAAAATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGAT
CGAGGAACGGCTGAAAACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCG
GCTGGGGACGACTGAGCAGAAAGCTGATCAACGGCATTCCGGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTG
AAATCAGATGGCTTCGCTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGATATCCA
GAAGGCACAGGTGTCCGGGCAGGGTGAATCTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGA
AAGGTATTCTGCAGACCGTGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTG
ATTGAGATGGCCCGCAAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGAGCGGATGAAAAGGATCGAGGA
AGGCATTAAGGAACTGGGATCCCAGATCCTGAAAGAGCACCCCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGT
ATCTGTACTATCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGAC
GTGGATCATATCGTCCCACAGTCATTCCCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAA
CCGAGGAAAGAGTGATAATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATG
CCAACTGATCACACAGCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTCTAGAACTGGACAAAGCA
GGCTTCATCAAGCGACAGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAA
CACAAAGTACGATGAGAATGACAAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACT
TCAGAAAGGATTTCCAGTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCA
GTGGTCGGGACCGCCCTGATTAAGAAAATACCCCAAACCTGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGA
TGTCAGAAAAATGATCGCCAAGAGTGAGCAGGAAATTTGAAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCA
TGAATTTCTTTAAGACTGAGATCACCCCTGGCAAATGGGGAAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAG
ACCGGAGAAAATCGTGTGGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCGTGCCATGCCTCAAGTGAATAT
TGTCAAGAAAACAGAGGTGCAGACTGGCGGATTCACTAAGGAATCAATTCTGCCCCAACGGAACCTCTGATAAGCTGA
TCGCCCCGAAAGAAAGACTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTG
GTCGCAAAGGTGGAGAAGGGGAAAAGCAAGAAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAG
GAGCTCCTTCGAGAAGAACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATT
AACTGCCAAAGTACAGCCTGTTTGAGCTGGAAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAG
GGTAATGAACTGGCCCTGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTC
CCCCGAGGATAACGAACAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTA

GCGAGTTCTCCAAAAGAGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGAT
AAGCCAATCAGGGAGCAGGCCGAAAAATATCATTCATCTGTTCACTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAA
GTATTTTGACACTACCATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCACCAGA
GCATTACCGGCCTGTATGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGCCATCTGCTGGAGACATG
AGAGCTGCCAACCTTTGGCCAAGCCGCTCATGATCAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCCTGACGGC
CGACCAGATGGTCAGTGCCTTGTGGATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTACCAGACCCTTCA
GTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTCACATGATCAACTGGGCGAAGAGG
GTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAGGTCCACCTTCTGGAATGTGCCTGGCTAGAGATCCTGATGAT
TGGTCTCGTCTGGCGCTCTATGGAGCACCCAGTGAAGCTACTGTTTGCTCCTAACTTGCTCTTGACAGGAACCAGG
GAAAAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTCCGCATGATGAATCTG
CAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAATTCTGGAGTGACACATTTCTGTCCAGCACCTT
GAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCCTGGACAAGATCACAGACACTTTGATCCACCTGATGGCCA
AGGCAGGCCTGACCCTGCAGCAGCAGCACCAGCGGCTGGCCAGCTCCTCCTCATCCTCTCCACATCAGGCACATG
AGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGGCCCTCTATGACCTGCTGCTGGAGGC
GGCGGACGCCACCGCTACATGCGCCCACTAGCCGTGGAGGGGCATCCGTGGAGGAGACGGACCAAGCCACTTGG
CCACTGCGGGCTCTACTTCATCGCATTCTTGCAAAAGTATTACATCACGGGGGAGGCAGAGGGTTTCCCTGCCACA
GCTTGA (SEQ ID NO:7)

[0213] 图2所示TOIC构建体的序列

[0214]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCTGATAAATGCTTCAATAATATGAAAAAGGAAGAGTATGAGTATTCACATTTCCGTGTGCG
CCTTATTCCTTTTTTTCGGCATTTCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAAGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAACATGGGGGATCATGTAACCTGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGAAGTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT

ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCCTGTGGGTTTCGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCCCTACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAATAAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTC
CTCATTTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTC
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCCGCAAGCAGGCATCGCCATGGGTGACGACGAGA
TCCTCGCCGTGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGGAGCCCTGATGCTCTTCGTCCAG
ATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTCAGTGACAACGTGAG
CACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTTATTACAGGGCAC
CGGACAGGTGCGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTAGTTCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTGAAAAGGCCCGGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGCAGAATGCCGGGCCTCCGGAGGACCTTCGGGCGCCCCGCCCCGCCCCCTGAGCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCAGTCTCT
GCACGACGCGAGCTGCGGGGCGGGGGGGAACCTCTGACTAGGGGAGGAGTAGAAGGTGGCGCGAAGGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTG
ACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT

GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCGGCTCTGAC
TGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCCTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTCGCGAGGGGAGCGCGCGGGGGCGGTGCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGTGAGCAGGGGTGTGGGCGCGCGGTGCG
GGCTGTAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGTGCGGCGAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGAGGGCTCGGGGAGGGGCGCGGCGCCCCGAGCGCCGCGGCTGTGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGACTTCCTTTGTCCCAAATCTGGCGAGCCGAAATCTGGGAGGC
GCCGCGCACCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGCGAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTGCGCGCGCCGCTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCCTGGAATAATCAGCTCGCGTTCCTGTGT
CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGAAGTCTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGTGACGCTCTTGACGATTTTGAAGTCTGACATGCTCCCCGGGTAAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTATTATGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGACGAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCGGGCACGGGAAGGTGGCACGGGATCGTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGC

GTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCTATCAAGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTTCAAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTATACAATCGTCAAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCAGGGACAGCCCGGGCACGG
GGAAGGTGGCACGGGATCGTTTCTCTGAACGCTTCTCGTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGG
GAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGCTCTATCAAGGG
GGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAAGACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCTCCCCTGGCAC
AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGTGACGGTGGGCGCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCTTACCCTCGTAAAAAGCTTGTCCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGACATCGGAACCTCACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTTCGACTCTGGAGAGACAGCTGAAGCACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA
CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGA
CGAGGTGCGTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT

TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGGAAATCTGATTGCACTGTCACCTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCCCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAACTGCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCACCAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCCTGGCAAGGG
GTAATTCCCGCTTTGCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGGAACTTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTTCATCGAGAGAATGACAAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGA
GGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTTCGGGACAAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGCAGCTGATTACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC
CGTGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCG
AAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTG
GGATCCCAGATCCTGAAAGAGCACCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCC
CACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTGCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCC
TGATTAAGAAATACCCCCAACTGGAGAGCGAATTCTGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTGGAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT
GGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATTTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTCGCAAGGTGGAGA
AGGGGAAAAGCAAGAAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACCTGCCAAAGTACAG

CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGGCCC
TGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATTCTCTGTTCACTCTGACCAACCTGGGAGCCCCCTGCAGCCTTCAAGTATTTTGACACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCCTGATTACCCAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACAT
GCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG
GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGACCA
CCTTACCTACGGCGTGAGTGCTTCGCCCCGTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACCTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTC
AAGACCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG
CCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCT
AGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCGTGCCTTCCT
TGACCCTGGAAGGTGCCACTCCCCTGTCTTCTTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
CATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCAT
ATGAATATCCTCCTTAGTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTAC
GTCCCTCCCCCGCTAGGGGCGACGAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGC
CGGCAGCGTGCGGGGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCT
TTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATA
GAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCA
GCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCA
ACCTGTTCAGTGCGTCACCACCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTG
TAAAGCCATTCCCCCTTGCTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAA
TTCAGTGATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTG
AGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGT
TTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTCAGTAGGACTGGGA
CAGGCAGCTGGAGAATTGCCATGTAGATGTTACATAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCC
CCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCAC
CTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTT
GCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGA
GGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCA
CGATATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGT

ATGAGTTAAATCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGA CTACGCGGTCGTTATAGTTCAAAATC
AGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCTAAATGCACAGCGA
CGGATTCGCGCTATTTAGAAAAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAG
GGTTAAAAAAGATTTGCGCTTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATT
GTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAG
TGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATC
GCCTTG CAGCACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGT TAAAAATTCGCGTTAAATTTTTGT TAAATCAG
CTCATTTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTG
TTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAA
CCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTG TAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCG
CCGCTACAGGGCGCGTCAG (SEQ IDNO:8)

[0215] AAV gRNA序列

[0216]

CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGC
CCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCTTGGGCCGCACGCG
TGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATT
TGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTAA
AAATTATGTTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTATATATCTT
GTGGAAAGGACGAAACACCGCAGCGTTACCTCTATCGTAGTTTTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAG
TCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTGGATCCGAGGGCCTATTTCCCATGATTCCTTCA
TATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACA
AAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTAAAAATTATGTTTTTAAAATGGACTATCATA
TGCTTACCGTAACTTGAAAGTATTTGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTGTAATAG
CTCCTGCATGGGTTTTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCC
GAGTCGGTGCTTTTTTTCTAGAAGGTACCAGGTCTGAAAGGAGTGGGAATTGGCTCCGGTGCCCGTCAGTGGGCAG
AGCGCACATCGCCACAGTCCCCGAGAAGTTGGGGGAGGGGTCCGCAATTGAACCGGTGCCTAGAGAAGGTGGCGC
GGGGTAAACTGGGAAAGTGATGTGCTGTACTGGCTCCGCTTTTTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGC
AGTAGTCGCGTGAAACGTTCTTTTTCGCAACGGGTTGCCGCCAGAACACAGGCGTACGGCCACCATGACTTCGAAA
GTTTATGATCCAGAACAAAGGAAACGGATGATAACTGGTCCGCAGTGGTGGGCCAGATGTAAACAAATGAATGTTCT
TGATTCATTTATTAATTATTATGATTCAGAAAAACATGCAGAAAAATGCTGTTATTTTTTTACATGGTAACGCGGCCT
CTTCTTATTTATGGCGACATGTTGTGCCACATATTGAGCCAGTAGCGCGGTGATTATACCAGACCTTATTGGTATG
GGCAAATCAGGCAAATCTGGTAATGGTTCTTATAGGTTACTTGATCATTACAAATATCTTACTGCATGGTTTGAAC
TCTTAATTTACCAAAGAAGATCATTTTTGTGCGCCATGATTGGGGTGCTTGTTTGGCATTTTATTATAGCTATGAGC
ATCAAGATAAGATCAAAGCAATAGTTCACGCTGAAAGTG TAGTAGATGTGATTGAATCATGGGATGAATGGCCTGAT
ATTGAAGAAGATATTGCGTTGATCAAATCTGAAGAAGGAGAAAAAATGGTTTTGGAGAATAACTTCTTCGTGGAAAC
CATGTTGCCATCAAAAATCATGAGAAAGTTAGAACCAGAAGAATTTGCAGCATATCTTGAACCATTCAAAGAGAAAG

GTGAAGTTCGTCGTCCAACATTATCATGGCCCTCGTGAAATCCCGTTAGTAAAAGGTGGTAAACCTGACGTTGTACAA
ATTGTTAGGAATTATAATGCTTATCTACGTGCAAGTGATGATTTACCAAAAATGTTTATTGAATCGGACCCAGGATT
CTTTTCCAATGCTATTGTTGAAGGTGCCAAGAAGTTTCCTAATACTGAATTTGTCAAAGTAAAAGGTCTTCATTTTT
CGCAAGAAGATGCACCTGATGAAATGGGAAAAATATATCAAAATCGTTCGTTGAGCGAGTTCTCAAAAATGAACAATAA
AGCGCTAATAAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTAAGCTTTGGCTCCAACACAGATG
TTCTTAGGCTACCTAACTTCTAACTTTTAATATCCAGTCAACAAAAGAATACCGCAAGGGTAGGTGTTGGGATAGCTG
TCGACAAGCTCATGCGGGTGTGTCCACAGGGTATAGCGTACTATGCAGAATATTTGTACTGAGTGAAGTCATGATAC
ATTCCTTTGAGAGCCATTAGCTGCTACAAAACAGTAATCTGGCTGTTTAGATCAACAAGCTAAATGATAGAAGATGA
AAGTACTGGTTTCCATGTATTTTTATTAAGTGTGATGAGAAAAGTTGTAAGTGACTTACAGGTTACTCTGTACATCT
GTAGTCACTGAATTCGGAATATCTTAGAGTTTTACACACAAAAGGTGAGTGTAAAAATATTGATAAAGTTTTTGATAA
TCTTGTGTGAGACATGTTCTAATTTAGTTGTATTTTATTATTTTTATTGTAAGGCCCTGCTGAAAATGACTGAGTATA
AACTTGTGGTCGTGGGCGCCGACGGCGTGGGCAAGAGCGCTTTGACGATACAGCTAATTCAGAATCACTTTGTGGAT
GAGTATGATCCAACCATCGAGGTAACGCTGCTCTACAGTCTGCGTGCGCTTGTAAGGACGGCAGCCAGCCGCTTTG
AAAAAGATATCATTTTTATATTTATTAGAAAATTATATTGAAAGTTATTTTCAGTTATATGTGATGTCCTTTAGTTCC
AAGGCTTTAACTGGGTGTTAGGGAACCATAGGTGCAAGAAAGTCCACTTCTCATGAGAGCTCACCACAGAGAAAGA
AAGTCCACTTCTCAGGTAACCACGTGCGGACCGAGCGGCCGAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCT
CTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCCTCAGT
GAGCGAGCGAGCGCGCAGCTGCCTGCAGGTATTTCTCCTTACGCATCTGTGCGGTATTTACACCCGCATACGTCAA
AGCAACCATAGTACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACA
CTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTCCCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCGTCA
AGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTTGG
GTGATGGTTACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAAT
AGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTATAAGGGATTTTGCC
GATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTA
CAATTTTATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCC
GCTGACGCGCCCTGACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCAT
GTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTA
ATGTCATGATAATAATGGTTTCTTAGACGTCAAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTA
TTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAG
GAAGAGTATGAGTATTCAACATTTCCGTGTGCCCTTATTCCTTTTTTGCGGCATTTTGCTTCCTGTTTTTGCTC
ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTC
AACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATG
TGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCCGCATACACTATTCTCAGAATGACTTGG
TTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACC
ATGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAA
CATGGGGGATCATGTAACCTGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACA
CCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAA
CAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCGGGCTGGCTGGTTTAT
TGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCC

GTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCC
TCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTAA
ATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACT
GAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAA
ACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTCGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGG
CTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAG
CACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGG
TTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTT
GGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAA
AGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGG
TATCTTTATAGTCTGTGCGGTTTTGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAG
CCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT (SEQ
ID NO:9) .

[0217] TOIC_ApoL1_wt_PgkPuro序列

[0218]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCTTTTTTTCGGCATTTCCTTCCCTGTTTTTGCTCACCAGAAACGCTGGTGAAAGTAAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCAACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCTGTGCGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA

TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGGTTTCC
CGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCCCTACTAAAGGGAACCTCCCCCTAGCTTAATTAACCCCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAATAAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAAACCTTTTAAACATTCTCTCTTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTTCG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTTC
CTCATTTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCCATTC
GCCGCCtcaggcaccgggcttgccgggtcatgcaccaggtgcgcgggtccttcgggcacctcgacgtcggcgggtgacgg
tgaagccgagccgctcgtagaaggggaggttgccggggcgcgaggtctccaggaaggcgggcaccgccggcgctcg
gcccctccactccggggagcacgacggcgctgcccagacccttgccctggtggtcgggcgagacgcccagcggtggc
caggaaccacgcgggctccttgggccgggtgcggcgccaggaggccttccatctgttgcgcggccagccgggaac
cgctcaactcgcccatgcgcggggcgatctcggcgaacaccgccccgcttcgacgctctccggcggtggtccagacc
gccaccgcgggcgccgctgctcgcgacccacaccttgccgatgtcgagcccgacgcgcgtgaggaagagtctcttcgag
ctcggtgacccgctcgatgtggcggtccgggtcgacgggtgtggcgctggcggggtagtcggcgaaacgcgggcgga
gggtgcgtacggccccgggggacgtcgctcgcggttgccgagggcgacccgtgggcttgtagtcgggtcatGGTTTAGTTC
CTCACCTTGTGCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAACACGTGCTGCTGCAGGTCGAAA
GGCCCGGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTGAAGCGTGAGAATGCCGGGCCTCCGGAG
GACCTTCGGGCGCCCGCCCCGCCCCCTGAGCCCGCCCCCTGAGCCCGCCCCCGGACCCACCCCTTCCCAGCCTCTGAGC
CCAGAAAGCGAAGGAGCAAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCTACCCGCTTCCATTGCTCAGCGGTGCTGT
CCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTACGTCCTGCACGACGCGAGCTGCGGGGCGGGGGG
AACTTCCTGACTAGGGGAGGAGTAGAAGGTGGCGGAAGGGGCCACCAAAGAACGGAGCCGTTGGCGCCTACCGGT
GGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGCCAGCGGGGCTGCTAAAGCGCATGCTCC
AGACTGCCTTGGGAAAAGCGCCTCCCCCTACCCGGTAGACACCCACAGTGGGTGGCCTAGGGACAGGATTGCAACTC
CAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCAT
ATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGACTATTTACGGTA
AACTGCCCCTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGC
CCGCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT
ATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGT
ATTTATTTATTTTTTAATTATTTTGTGACGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCG
GGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTT

TTATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGTTGCCTTCG
CCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGGTTACTCCCACAGGTGAGCGG
GCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTCGTTTCTTTTCTGTGGCTGCGTG
AAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTGCGT
GGGGAGCGCCGCTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCG
CGTGTGCGGAGGGGAGCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGCTGCGAGGGGAACAAAGGCTGCGTG
CGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTGCGGCTGTAACCCCCCTGCACCCCCCTC
CCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGC
GGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGG
CGGCCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGC
GCAGGGACTTCCTTTGTCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCGCCGCACCCCCCTCTAGCGGGCGCGG
GCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCTCCCCTTCT
CCATCTCCAGCCTCGGGGTGCCGAGGGGGACGGCTGCCTTCGGGGGGGACGGGCAGGGCGGGGTTCGGCTTCTG
GCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTTACAGCTCCTGGGCAACG
TGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCATGGTGCCCAAGAAGAAGAGGAAAGTCTCT
AGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGAAGGCCTGACGACAAG
GAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCACGTGAAGAACAAGCGGGCCCTGCTCGATG
CCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCTGGAAGGCGAGTCATGGCAAGACTTTCTG
CGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCATCTCGGCACCCGCCC
AACAGAGAAACAGTACGAAACCTGGAAAATCAGCTCGCGTTCCTGTGTGAGCAAGGCTTCTCCCTGGAGAACGCAC
TGTAAGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCATCAAGTAGCAAAAGAG
GAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGCTGTTGACCGGCAGGGAGC
CGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAGAAACAGCTAAAGTGCGAAAGCGGCGGGC
CGACCGACGCCCTTGACGATTTTGAAGTCTAGACATGCTCCCAGCCGATGCCCTTGACGACTTTGACCTTGATATGCTG
CCTGCTGACGCTCTTGACGATTTTGAAGTCTGACATGCTCCCCGGGTAAAGCGGCCGCGACTCTAGATCATAATCAGC
CATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAA
TGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAA
ATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTGTCCAAACTCATCAATGTATCTTAAGGGATCCCTAGAGGG
ACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCGGGGCT
CCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCA
CGGGATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTT
TAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGAT
CCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG
CCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCC
TCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCAAGGGGGAGTTTGCTG
TGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTTCAGTGTCATCACGGAGAGGCAGATCTTGGGGATAAGGAAG
TGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAG
CCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCA
CAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTT

GCCTGGAGAGATACAGCTGGGTGCTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCG
TCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCAT
GTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTG
CCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCCTCCCCTGGCACAACGTAAGGCC
ATTATCTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGACAGCCCCCCCCAAAG
CCCCAGGGATGTAATTACGTCCCTCCCCGCTAGGGGACAGCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGC
TCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGCACGGGGAAGGTGGCACGGGATCGCTTTCTCT
GAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGAT
TTAGAATGACAGAATCATAGAACGGCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGT
GCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCA
TCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCA
AACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGG
GGTGACACATGTTTGCCAAATTCAGTGTCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGAC
GTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGA
TAGGATAGAAGGACAAAGAGCAAGTTAAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCC
TGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGC
TGGGTTCAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTG
GAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACAT
CCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTC
TTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAAC
TCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACGTATGAAGAGTTTACT
CCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAAGGAGTTTACTCCCTA
TCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTACAGTTTACTCCCTATCAGT
GATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTATAAGCTTTAGGCGTGACGGTGGGCGCC
TATAAAGCAGAGCTCGTTTAGTGAAACCGTCAGATCGCCTGGAGCAATTCCACAACACTTTTGTCTTATACCAACTT
TCCGTACCACTTCCTACCCTCGTAAAAAGCTTGTCACCATGAGATTCAAAAGCCACACTGTGGAATTGAGGAGGCC
CTGCAGCGACATGGAGGGAGCTGCTTTGCTGAGAGTCTCTGTCTCTGCATCTGGATGAGTGCACTTTTCCTTGGTG
TGGGAGTGAGGGCAGAGGAAGCTGGAGCGAGGGTGCAACAAAACGTTCCAAGTGGGACAGATACTGGAGATCCTCAA
AGTAAGCCCCTCGGTGACTGGGCTGCTGGCACCATGGACCCAGAGAGCAGTATCTTTATTGAGGATGCCATTAAGTA
TTTCAAGGAAAAAGTGAGCACACAGAATCTGCTACTCCTGCTGACTGATAATGAGGCCTGGAACGGATTCTGGCTG
CTGCTGAACTGCCCAGGAATGAGGCAGATGAGCTCCGTAAAGCTCTGGACAACCTTGCAAGACAAATGATCATGAAA
GACAAAAACTGGCAGCATAAAGGCCAGCAGTACAGAAACTGGTTTCTGAAAGAGTTTCCTCGGTTGAAAAGTGAGCT
TGAGGATAACATAAGAAGGCTCCGTGCCCTTGACAGATGGGGTTCAGAAGGTCCACAAAGGCACCACCATCGCCAATG
TGGTGTCTGGCTCTCTCAGCATTTCTCTGGCATCCTGACCCTCGTCGGCATGGGTCTGGCACCCCTTCACAGAGGGA
GGCAGCCTTGTAATCTTGAACCTGGGATGGAGTTGGGAATCACAGCCGCTTTGACCGGGATTACCAGCAGTACCAT
GGACTACGGAAAGAAGTGGTGACACAAGCCCAAGCCACGACCTGGTCATCAAAAGCCTTGACAAATTGAAGGAGG
TGAGGGAGTTTTTGGGTGAGAACATATCCAATTTCTTTCTTAGCTGGCAATACTTACCAACTCACACGAGGCATT
GGGAAGGACATCCGTGCCCTCAGACGAGCCAGAGCCAATCTTCAGTCAGTACCGCATGCCTCAGCCTCACGCCACG
AGTCACTGAGCCAATCTCAGCTGAAAGCGGTGAACAGGTGGAGAGGGTTAATGAACCCAGCATCCTGGAAATGAGCA

GAGGAGTCAAGCTCACGGATGTGGCCCCTGTAAGCTTCTTTCTTGTGCTGGATGTAGTCTACCTCGTGTACGAATCA
AAGCACTTACATGAGGGGGCAAAGTCAGAGACAGCTGAGGAGCTGAAGAAGGTGGCTCAGGAGCTGGAGGAGAAGCT
AAACATTCTCAACAATAATTATAAGATTCTGCAGGCGGACCAAGAACTGTGAAATTCTAAAATACAGCATAGCAAAA
CTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGC
CAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTCATGGAGTTTAAGATATAGTGTATTTTCCCAAGGTTTGAAC
TAGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATTCCCTTTTTAGTAAAATATTTCAGAAATAATTT
AAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCCCTTCATAATATCCCC
CAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTCTAGCTCG
AGATGGTCCATATGAATATCCTCCTTAGTTCTTATCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGG
GATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCG
CATCCCCGAGCCGGCAGCGTGCGGGACAGCCCGGCACGGGGAAGTGGCACGGGATCGCTTTCCTCTGAACGCTT
CTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATG
ACAGAATCATAGAACGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTC
ATCAACCAGCAGCCCAGGCTGCCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGC
CTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCC
CTGTCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACAC
ATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACA
TGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAG
AAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTG
AGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAG
TAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCC
CTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAG
TTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGA
AGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGAC
GGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTA
TATATTAGTCACGATATCTATAACAAGAAAAATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAAT
ATAATTATCGTATGAGTTAAATCTTAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTGCTTAT
AGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAA
ATGCACAGCGACGGATTCGCGCTATTTAGAAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGAC
TATCTTTCTAGGGTTAAAAAGATTTGCGCTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTCGTTTGACA
AAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGG
TCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTAC
CCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT
CCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATGTAAAGCGTTAATATTTTGTAAAATTCGCGTTAAATTTT
TGTTAAATCAGCTCATTTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGAT
AGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAA
CCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCA
CTAAATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGG
GAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCG

CGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQ ID NO:12)

[0219] TOIC-Cas9_0b1_r26_AAVS_SANeo

[0220]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCGTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCCTTTTTTTCGGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCAACAATGGGGGATCATGTAACGCGCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCAGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTGC
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTT
ACGGTTCTTGGCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAACCGCCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAATTATACTTATTAATCGCATTGATTATTGACTAGTCACAATATGATTATCTTTCTAGGGTTAATTA
AGATATCTGAAGTTTCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAGCAAGCTAGAGACCATTA
AGGGTTCCGGATCAGCTTGATTGAGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCA

GCATGCCTGCTATTGTCTTCCCAATCCTCCCCCTTGCTGTCTGCCCCACCCACCCCCAGAATAGAATGACACCT
ACTCAGACAATGCGATGCAATTTCTCATTATTATTAGGAAAAGGACAGTGGGAGTGGCACCTTCCAGGTCAAGGAAG
GCACGGGGGAGGGGCAAACAACAGATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAAT
TGATCCCCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAG
CACGAGGAAGCGGTACGCCATTGCGCCGAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCTGATAGC
GGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTGGCAAGCAG
GCATCGCCATGGGTACGACGAGATCCTCGCCGTGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTTCGGCTGGCGC
GAGCCCCTGATGCTCTTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGC
GATGTTTTCGCTTGGTGGTGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTCATCAGCCATGATG
GATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCT
TCCCGCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCT
CGTCTGCAGTTCATTACGGGCACCGGACAGGTGCGTCTTGACAAAAAGAACCGGGCGCCCCGTCGCTGACAGCCGG
AACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCG
AGAACCTGCGTGCAATCCATCTTGTTCAATGGCCGATCCCATGGCGGTATCGATAAGCTAGCTTGGGCTGCAGGTGC
AGGGACCTAATTAAGGGTTCCGATCCACTAGTTCTAGAGCGGCCTCGACTCTACGATACCGTCGATCCCCACTGGA
AAGACCGGAAGAGTTTGTCTCAACCGCGAGCTGTGGAAAAAAGGGACAGGATAAGTATGACATCATCAAGGAA
ACCCTGGACTACTGCGCCCTACAGATCCCTGAAGTTCCCTATACTTTCTAGAGAATAGGAACCTTCGGAATAGGAACCT
CAAAGATGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCTACTAGTTATTAATAGTAATCAATTACGGGGTCAT
TAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGAC
CCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGT
GGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCA
ATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAC
GTATTAGTCATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCC
CACCCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGCGCGCGCC
AGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGC
TCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGGGAGTC
GCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACT
CCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTCGTTTCTT
TTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTG
CGTGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGGG
GCTTTGTGCGCTCCGCGTGTGCGGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGGGGGCTGCGAGGGG
AACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTGCGGCTGTAACCCCC
CCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGTGGCGGGG
CTCGCCGTGCCGGGCGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGAGGGCTC
GGGGGAGGGGCGGGCGGCCCCGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTGCCTTTTATGGTA
ATCGTGCGAGAGGGGCGAGGGACTTCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGCCGCCGACCCC
CTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGAGGGCCTTCGTGCGTCGCCGCG
CCGCCGTCCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCCTCGGGGGGGACGGGGCAGGGC
GGGGTTGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTCTTCTTTTCTAC

AGCTCCTGGGCAACGTGCTGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCCGCCACCATGGTGCCCAAGAAG
AAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGAGTCGGTATCGA
AGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCACGTGAAGAACAAGC
GGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCTGGAAGGCGAGTCA
TGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACGGGGCTAAAGTGCA
TCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGTCAGCAAGGCTTCT
CCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGGAACAGGAGCAT
CAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCAATTGAGCTGTT
CGACCGGCAGGGAGCCGAACCTGCCTTCTTTTCGGCCTGGAACTAATCATATGTGGCCTGGAGAAACAGCTAAAGT
GCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGCCCTTGACGACTTT
GACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAAGCGGCGCGGACTC
TAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCT
GAAACATAAAATGAATGCAATTGTTGTTGTTAACTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTAA
GGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGC
GAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGC
ACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGAT
ACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCAC
AGTGCTCATCCAGATCCAACCCCCCTGCTATGTGACAGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCA
GCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCT
GGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCA
AGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATC
TTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCA
CAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAACCCAGCATGGAGAGG
AGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAA
GGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAG
ATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACC
GTGCCCCTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCT
CCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTG
GCACAACGTAAGGCCATTATCTCTCATCCAACTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGACCTAGAGGGAC
AGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCGGGGCTCC
GCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCAGGGGAAGGTGGCACG
GGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTA
GGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCC
AACCCCTGCTATGTGACAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCC
TGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCCTC
CTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCAAGGGGGAGTTTGCTGTG
ACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTG
CAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCC

TTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACA
GACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGC
CTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTC
AAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGT
CCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCC
ACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCCTCCCTGGCACAACGTAAGGCCAT
TATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGTGATAGAGAACG
TATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAGAGAACGTATAA
GGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAACGTATCTACAGT
TTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTATAAGCTTTAGGCG
TGTACGGTGGGCGCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATTCCACAACACTTTT
GTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAAAAGCTTGTCACCATGGCTCCTAAGAAAAAGCGGAA
GGTGGACAAGAAATACTCAATCGGGCTGGACATCGGAACTAACTCAGTGGGGTGGGCAGTCATTACTGACGAGTACA
AAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGATTGGGGCCCTG
CTGTTGCACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACCCGGCGAAAGAA
TCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTTTCATCGGCTGGAGG
AATCATTCCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAACATTGTGGACGAGGTGCTTAT
CACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAGACCTGCGGCTGAT
CTATCTGGCCCTGGCTCACATGATTAAGTTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCCAGACAATAGCG
ATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTATTAATGCATCTGGC
GTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCTCAGCTGCCAGG
CGAAAAGAAAAACGGCCTGTTTGAAAATCTGATTGCACTGTCACTGGGACTGACACCTAACTTCAAGAGCAATTTTG
ATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCTGCTGGCTCAGATC
GGAGATCAGTACGCAGACCTGTTCTGCGCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTGATATTCTGCGGGT
GAACACCGAGATTACAAAAGCCCCCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCAGGATCTGACCC
TGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTAAGAACGGGTAC
GCCGGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGGAGAAGATGGACGG
CACCGAGGAACTGCTGGTGAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGATAACGGCTCCATCC
CCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATTTCTGAAGGATAAC
CGCGAGAAGATCGAAAAAATTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCCTGGCAAGGGTAATTCCCGCTT
TGCCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGATAAGGGAGCTTCCG
CACAGTCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCCTAAGCACAGTCTG
CTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGAGGAAGCCCGCCTT
CCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTAAAACCAATCGCAAGGTGACAGTCAAGCAGCTGA
AGGAGGACTACTTCAAGAAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAGATTTAACGCTTCT
CTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAAATGAGGATATCCT
GGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAACCTATGCCATC
TGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAGGTACACCGGCTGGGGACGACTGAGCAGAAAGCTGATC
AACGGCATTGGGACAAACAGAGTGAAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCGCTAACAGAAATTT

TATGCAGCTGATTACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCCGGGCAGGGTGACT
CTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCGTGAAGGTGGTC
GATGAGCTGGTGAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGAAAATCAGACCAC
ACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGGATCCCAGATCC
TGAAAGAGCACCCCTGTGGAAAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCAGAATGGGCGGGAT
ATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCCCACAGTCATTCT
GAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGATAATGTCCCTTCAG
AGGAAGTGGTCAAGAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGCGAAAGTTTGAT
AACCTGACTAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAGCTGGTGGAGAC
CAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAGAATGACAACTGA
TCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCAGTTCTACAAAGTC
AGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCCTGATTAAGAAATA
CCCCAACTGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATCGCCAAGAGTGAGC
AGGAAATTGAAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTGAGATCACCTG
GCAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGGGACAAGGGTAG
GGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAGGTGCAGACTGGCG
GATTTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTGATAAGCTGATCGCCCGAAAGAAAGACTGGGATCCCAAG
AAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTCGCAAGGTGGAGAAGGGGAAAAGCAA
GAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAACCCCATCGATT
TTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCCTGTTTGAGCTG
GAAAACGGAAGGAAGCGAATGCTGGCATCCGCGAGGAGCTGCAGAAGGTAATGAACTGGCCCTGCCTTCTAAGTA
CGTGAACCTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAACAGAAACAGCTGT
TTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAGAGTGATCCTGGCT
GACGCAAATCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAGGCCGAAAATAT
CATTTCATCTGTTCACTCTGACCAACCTGGGAGCCCCCTGCAGCCTTCAAGTATTTTGACACTACCATCGATCGGAAAC
GATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACCCAGAGCATTACCGGCCTGTATGAAACAAGGATT
GACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGTGGA
GGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGG
TCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAG
CTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTTCACCTACGG
CGTGCAGTGCTTCGCCCCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG
TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAA
CTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGACCCGCCACA
ACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTG
CCCACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCT
GGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGCAGCTCGCTGA
TCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCCTTCCTTGACCCTGGAAGG
TGCCACTCCCACTGTCTTTTCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGG

GGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCT
ATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCCTCC
TTAGTTCCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCGGGCAGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGC
GTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCGTCTCAGTGTAAGCCATTCCC
CCTTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTCAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGA
TCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAAC
AAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCT
TAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTGCTTATAGTTCAAAATCAGTGACACTTACC
GCATTGACAAGCACGCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTCGCGCTA
TTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAGAT
TTGCGCTTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACCGAG
CTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTAC
AATTCAGTGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTAATCGCCTTGACGACA
TCCCCCTTTGCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATG
GCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTTAAAC
CAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTG
GAACAAGAGTCCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCAC
TACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGC
CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCGGGCGC
TAGGGCGCTGGCAAGGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCG
CGTCAG (SEQ ID NO:13)

[0221] TOIC_无核酸酶活性Cas9_0b1_r26_AAV_PgkNeo

[0222]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCG
CCTTATTCCCTTTTTTTCGGCATTGTTGCTTCCGTGTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTG

AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAGTCTGCTTATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCTTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAAACCGCCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGACAGGTTTCC
CGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATTTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACCAAAACAAAACTCAAAATTTCTTCTATAAAAGTAACAAAATTTTAAACATTCTCTCTTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTTCG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTC
CTCATTTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTTC
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTGACGACGAGA
TCCTCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAG

ATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAG
CACAGCTGCGCAAGGAACGCCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTGCAGTTCATTACAGGGCAC
CGGACAGGTGCGTCTTGACAAAAAGAACCAGGGCGCCCCGTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTAGTTCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTGCAAAGGCCCGGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGAGAATGCCGGGCTCCGGAGGACCTTCGGGCGCCCCGCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCAGTCTCT
GCACGACGCGAGCTGCGGGGCGGGGGGAACTTCTTGAAGGAGGAGTAGAAGGTGGCGGAAGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTG
ACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGAC
TGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGAGGGCTCGGGGAGGGGCGCGGCGGCCCGAGCGCCGCGGCTGTGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGC
GCCGCGCACCCCCCTTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGCGAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTCCCGCGCCGCGTCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTCGGTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC

GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGCATCTCGGCACCCGCCAACAGAGAAACAGTACGAAACCCTGAAAAATCAGCTCGCGTTCCTGTGT
CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCAAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGC
GTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTTGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAAATCATGAAGGCTGGAAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGG
GAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGG
GGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAAGACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTTGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT

TCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCTCCCCTGGCAC
AACGTAAGGCCATTATCTCTCATCCAATCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCCTTCCCTACCCCTCGTAAAAAGCTTGTCCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGCCATCGGAACTAACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA
CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGAAACATTGTGGA
CGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT
TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGAAATCTGATTGCACTGTCAGTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGACAGCTGTTCCCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAACCTGCTGGTGAACCTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCTGGCAAGGG
GTAATTCCCGCTTTGCTGGATGACACGGAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAATACGTCACCGAGGGGATGA
GGAAGCCCGCCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCTGGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTCCGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGAAACCTGGCCGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC
CGTGAAGGTGGTCGATGAGCTGGTGAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCG

AAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAAC TG
GGATCCCAGATCCTGAAAGAGCACCCCTGTGGAAAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATGCCATCGTCC
CACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCC
TGATTAAGAAATACCCCCAACTGGAGAGCGAATTCGTGTACGGTGAATAAGGTGTACGATGTCAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTGAAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT
GGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCGTGCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTCGCAAAGGTGGAGA
AGGGGAAAAGCAAGAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAG
CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCGAGGAGCTGCAGAAGGGTAATGAACTGGCCC
TGCCTTCTAAGTACGTGAACCTTCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCCGTAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATTCTGTTCACTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCACCAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACAT
GCGGTGACGTCGAGGAGAATCCTGGCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG
GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGACCA
CCTTCACCTACGGCGTGAGTGCTTCGCCCCGTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACATAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTC
AAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG
CCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCT
AGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCGTGCCTTCT
TGACCCTGGAAGGTGCCACTCCCCTGTCTTCTTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
CATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCAT
ATGAATATCCTCCTTAGTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTAC

GTCCCTCCCCGCTAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGC
CGGCAGCGTGCGGGGACAGCCCGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCT
TTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATA
GAACGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCA
GCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCA
ACCTGTTCAGTGCGTCACCACCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTG
TAAAGCCATTCCCCCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAA
TTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTG
AGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGT
TTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCCTGGAGAGATACAGCTGGGTGCTAGGACTGGGA
CAGGCAGCTGGAGAATTGCCATGTAGATGTTTATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCC
CCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCAC
CTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTT
GCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGA
GGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCA
CGATATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGT
ATGAGTTAAATCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGAATCACGCGGTCGTTATAGTTCAAAATC
AGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGA
CGGATTCGCGCTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAG
GGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATT
GTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAG
TGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCACTTAATC
GCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGCAAAATTGTAAGCGTTAATATTTTGTTAAAATTTCGCGTTAAATTTTTGTTAAATCAG
CTCATTTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTG
TTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAA
CCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCGCGCTTAATGCG
CCGCTACAGGGCGCGTCAG (SEQ ID NO:14)

[0223] TOIC_Cas9_KRAB_0b1_r26_AAVS_PGKNeo

[0224]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCGTATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCCTTTTTTGCGGCATTTTGCCCTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTA

CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCTGTCAGACCA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCGAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTCT
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCCGCAAGCAGGCATCGCCATGGGTGACGACGAGA
TCCTCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCGTGATGCTCTTCGTCCAG
ATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTCAGTGACAACGTGAG

CACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCCTGCAGTTCATTACAGGGCAC
CGGACAGGTGGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTAGTTCCCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTCGAAAGGCCCGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGAGAATGCCGGGCCTCCGGAGGACCTTCGGGCGCCCCGCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCAGAAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTCTCT
GCACGACGCGAGCTGCGGGGCGGGGGGAACTTCTGACTAGGGGAGGAGTAGAAGGTGGCGCGAAGGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTG
ACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTGCCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGAC
TGACCGGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCCTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCGGCTCGGG
CCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGAATTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGC
GCCGCGCACCCCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTGCCGCGCCGCGTCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGT

CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCGAAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAAATAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGC
GTCACCACCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGG
GAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCAAGGG
GGAGTTTGTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAATCATGAAGGCTGGAAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCTCCCCTGGCAC

AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCCTCGTAAAAAGCTTGTCCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGACATCGGAACCTAACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTTCGACTCTGGAGAGACAGCTGAAGCACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA
CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGA
CGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT
TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAAATCTGATTGCACTGTCCTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTGTCTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAACTGCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCACCAGATTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGG
GTAATTCCCCTTTGCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGA
GGAAGCCCGCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCTTGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTCCGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC
CGTGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCG
AAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTG
GGATCCCAGATCCTGAAAGAGCACCTGTGGAACCACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCC

CACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCC
TGATTAAGAAATACCCCAAACTGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGATGTGAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTTGAAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT
GGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTCGCAAAGGTGGAGA
AGGGGAAAAGCAAGAAACTGAAATCCGTCAAGGAGCTGCTGGGTATCATTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAG
CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGGCCC
TGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAAATCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATTTCATCTGTTCCTGCTGACCAACCTGGGAGCCCCCTGCAGCCTTCAAGTATTTTGACACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCACCAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGATGGCGGTGGCGCGCTGTCCCCGCAGCACTCCG
CCGTGACCCAGGGGAGTATAATCAAAAACAAAGAGGGCATGGATGCTAAGAGCCTTACCGCCTGGTCCCGAACACTG
GTCACGTTTAAGGATGTGTTTCGTGCGATTTTACCCGGGAGGAGTGGAACCTGCTCGACACCGCGCAGCAGATCGTGTA
CCGGAATGTCATGCTCGAAAATTACAAAACCTTGGTCAGCCTCGGGTACCAATTGACCAAACCAGATGTCATACTGC
GACTGGAAAAAGGAGAGGAACCTGGCTCGTCGAGCGCGAAATTCATCAAGAAACACACCCGATTCTGAAACCGCC
TTCGAGATTAAGAGCAGTGTGCTAGGCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACATGCGGTGACGT
CGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCC
TGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC
AAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCACCTA
CGGCGTGAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCGAAGGCT
ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGAC
ACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGACCCGCC
ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTT
GCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCTAGCAGCTCGC
TGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCTTGACCCTGGA
AGGTGCCACTCCCCTGTCTTTCTTAATAAAATGAGGAAATGTCATCGCATTGTCTGAGTAGGTGTCTATTCTATTC
TGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGC

TCTATGGCTTCTGAGGCGGAAAGAACCCAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCC
TCCTTAGTTCTTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCC
CGCTAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTG
CGGGGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGC
AGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTG
GGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTG
CCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCA
TGCGTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATT
CCCCCTTGCTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCA
CACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGG
GACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAA
ACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGT
CTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTG
GAGAATTGCCATGTAGATGTTTATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAA
CCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGAC
GGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCA
GCCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCT
GGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTAT
AACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAA
TCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTATAGTTCAAAATCAGTGACACTT
ACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTTCGG
CTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAA
GATTTGCGCTTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGTACC
GAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTAT
TACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACCTTAATCGCCTTGCAGC
ACATCCCCCTTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGA
ATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTTT
AACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGT
TTGGAACAAGAGTCCACTATTAAGAAGCGTGGAAGTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCC
CACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGG
AGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCGGG
CGCTAGGGCGCTGGCAAGTGTAAGCGGTACGCTGCGCGTAACCACCACACCCGCGCGCTTAATGCGCCGCTACAGG
GCGCGTCAG (SEQ ID NO:15)

[0225] TOIC_Cas9_无核酸酶活性_VP64_0b1_r26_AAVS_SANeo

[0226]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGC
CCTTATTCCTTTTTTTCGGCATTGTTGCTTCCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC

GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTCGCCCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGCGT
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAAACGCCTCTCCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCCAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATT
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTGGCAAGCAGGCATCGCCATGGGTACGACGAGA
TCCTCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCTGATGCTCTTCGTCCAG
ATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCAATG

GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTCGAG
CACAGCTGCGCAAGGAACGCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCCTGCAGTTCATTACAGGGCAC
CGGACAGGTCGGTCTTGACAAAAAGAACCGGGCGCCCCGTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGCGGTATCGATAAGCTAGCTTGGGCTGCAGGTCGAGGGACCTAATTAAGGGTTCGGGA
TCCACTAGTTCTAGAGCGGCCTCGACTCTACGATACCGTCGATCCCCACTGGAAAGACCGCGAAGAGTTTGTCTCA
ACCGCGAGCTGTGGAAAAAAGGGACAGGATAAGTATGACATCATCAAGGAAACCCTGGACTACTGCGCCCTACAG
ATCCCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTCAAAGAACGCGTACCCACAGTGG
GTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCAAT
TACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
CGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGA
CGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
TATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTGGC
AGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTC
CCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGG
GGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCA
GAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGC
GGGCGGGAGTGCCTGCGTTGCCTTCGCCCCGTGCCCCGTCCGCGCCGCTCGCGCCGCCCGCCCGGCTCTGACTG
ACCGCGTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAAATGACG
GCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTGTGCGGGGGGAGCGGCTCG
GGGGTGCGTGCCTGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCCGCGTGCCCGCGGCTGTGAGCGCTGCGGG
CGCGGCGCGGGCTTTGTGCGCTCCGCGTGTGCGGAGGGGAGCGCGGCCGGGGCGGTGCCCGCGGTGCGGGGGG
GCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTGCGG
CTGTAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGG
GTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCC
GGGGAGGGCTCGGGGAGGGGCGCGGCGGCCCGGAGCGCCGCGGCTGTCGAGGCGCGGCGAGCCGACGCCATTGC
CTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCGC
CGCCGCACCCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGCGAGGAAGGAAATGGGCGGGGAGGGCCTTCGT
GCGTGCCTGCGCCGCGTCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGGGA
CGGGGCGAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTCTT
CTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATATTTTGGCAAAGAATTCGCCACCATGG
TGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATGGA
GTCCGTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCACGT
GAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCTGG
AAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACGGG
GCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAATCAGCTCGCGTTCCTGTGTCA
GCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTGTGCCCGTGGGCCACTTTACACTGGGCTGCGTATTGGAGG
AACAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAGCA

ATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAGAA
ACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGCCC
TTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAAGC
GGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTC
CCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATA
AAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCA
ATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCGCTA
GGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGG
ACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACA
CCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTG
CAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAG
AGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGT
CACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCC
TTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGG
AGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACAC
TCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCA
GCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGA
TGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAA
TTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCA
ACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGA
CCCCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCG
ACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTGGA
CCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCG
CCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGGGG
AAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGG
AAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCT
CATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGG
CCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTACCACCCTCTGGGGGA
AAAAGTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGGGG
AGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTGGGG
ATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGT
TCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACA
AAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTG
GAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTT
ATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCC
ACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGG
GCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCCTCCCCTGGCACA
CGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCAGT
GATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGATAG

AGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGAAC
GTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTATA
AGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATTCC
ACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCCTCGTAAAAAGCTTGTCCACCATGGCTCCTAAG
AAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGCCATCGGAACTAACTCAGTGGGGTGGGCAGTCATTAC
TGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCTGA
TTGGGGCCCTGCTGTTTCGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATACC
CGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTTCA
TCGGCTGGAGGAATCATTCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGACG
AGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAGAC
CTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAACCC
AGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTATTA
ATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCGCT
CAGCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGAAAATCTGATTGCACTGTCACTGGGACTGACACCTAACTTCAA
GAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCTGC
TGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTGAT
ATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCATCA
GGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAGTA
AGAACGGGTACGCCGTTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGGAG
AAGATGGACGGCACCGAGGAAGTCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGATAA
CGGCTCCATCCCCACAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATTTCT
TGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGGGT
AATTCCCCTTTGCTGCTGACACGGAATCTGAGGAAACAATCACTCCTTGGAACTTCGAGGAAGTGGTCGATAA
GGGAGCTTCCGCACAGTCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCCTA
AGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGAGG
AAGCCCGCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTAAAACCAATCGCAAGGTGACAGT
CAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAGAT
TTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCCTGGACAACGAGGAAAT
GAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAAAC
TTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAGCA
GAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCGCT
AACAGAAATTTTATGCAGCTGATTACGATGACAGCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCCGG
GCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGACCG
TGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCGAA
AATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTGGG
ATCCCAGATCCTGAAAGAGCACCTGTGGAAAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCAGA
ATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATGCCATCGTCCCA
CAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGAGTGACAAAAACCGAGGAAAGAGTGATAA
TGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACAGC

GAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGACAG
CTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAGAA
TGACAAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCAGT
TCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCCTG
ATTAAGAAATACCCCAAACTGGAGAGCGAATTTCGTGTACGGTGACTATAAGGTGTACGATGTGAGAAAAATGATCGC
CAAGAGTGAGCAGGAAATTGGAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGACTG
AGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGTGG
GACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAGGT
GCAGACTGGCGGATTTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCCGAAAGAAAGACT
GGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCCCTGGTGGTCGCAAAGGTGGAGAAG
GGGAAAAGCAAGAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAGAA
CCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAGCC
TGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGAGCTGCAGAAGGGTAATGAACTGGCCCTG
CCTTCTAAGTACGTGAACCTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAACA
GAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAGAG
TGATCCTGGCTGACGCAAATCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGCAG
GCCGAAAATATCATTTCATCTGTTCCTCTGACCAACCTGGGAGCCCCCTGCAGCCTTCAAGTATTTTGACACTACCAT
CGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCACCAGAGCATTACCGGCCTGTATG
AAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGCCGAAAAAGAAACGCAAAGTTGGGCGCGCCGACGCG
CTGGACGATTTTCGATCTCGACATGCTGGGCAGCGACGCCCTGGATGACTTCGACCTGGATATGCTGGGCTCTGATGC
CCTGGACGACTTTGACTTGACATGTTGGGATCCGACGCTCTCGATGATTTTGACCTTGACATGCTGATCAACGGCA
GCGGCGAGGGCAGAGGCAGCCTGCTAACATGCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTG
AGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAACCGGCCACAAGTT
CAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC
TGCCCGTGCCCTGGCCCCACCTCGTGACCACCTTACCTACGGCGTGAGTGCTTCGCCCGCTACCCCGACCACATG
AAGCAGCAGCACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGG
CAACTACAAGACCCGCGCCGAGGTGAAGTTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACT
TCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAAGGTCTATATCACCGCCGAC
AAGCAGAAGAACGGCATCAAGGTGAACCTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCA
CTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGTGCCCCACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGC
ATGGACGAGCTGTACAAGTAAACCTAATCTAGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCAT
CTGTTGTTTGCCCCCTCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTCTTAATAAAATGAG
GAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGA
TTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCT
CGATCCTCTAGTTGGCGCGTCATGGTCCATATGAATATCCTCCTTAGTTCCCTATTCGCTAGCCTAGAGGGACAGCC
CCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCCGGGGCTCCGCTC
CGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGAAGGTGGCACGGGAT
CGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCT

GAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACC
CCCTGCTATGTGCAGGGTCATCAACCAGCAGCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCA
GGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCCTCACCACCCTCTGGGGGAAAACTGCCTCCTCA
TATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTATCAAGGGGAGTTTGCTGTGACAT
TGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGG
ACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAA
GGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACA
CTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGG
AGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTCATACAATCGTCAAAT
CATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCT
CAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTG
CAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCCTGGCACAACGTAAGGCCATTATC
TCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAG
ACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAG
TAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAAGTCACGTAAGATAATCATGCGTCATTT
TGACTCACGCGGTCGTTATAGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGC
GGCGACTGAGATGTCCTAAATGCACAGCGACGGATTGCGGCTATTTAGAAAAGAGAGCAATATTTCAAGAATGCAT
GCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAACACGTC
ATAGAATCTTCGTTTGACAAAAACCATTTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTT
AGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAATTCAGTGCCGTCGTTTTACAACGTCGTGACT
GGGAAAACCTGGCGTTACCCAACCTTAATCGCCTTGACAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAG
GCCCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTT
AAAATTGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAAT
CAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCC
AACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGG
GTGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGAAAGCCGGCGA
ACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGTAGGGCGCTGGCAAGTGTAGCGGTACAGCTGCGC
GTAACCACACACCCGCGGCTTAATGCGCGCTACAGGGCGGTCAG (SEQ ID NO:16)

[0227] TOIC_Bsd_BE_Cas9_0b1_r26_AAVS_PgkNeo

[0228]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCCTTTTTTGCGGCATTTTGCTTCTGTGTTTGCTCACCAGAAACGCTGGTGAAAGTAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACACAACATGGGGGATCATGTAAGTGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC

GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCTGGTATCTTTATAGTCCTGTGGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAACCGCCTCTCCCCGCGCTTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGTTTCC
CGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCCCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAATTTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTTCG
AGCCCCAGCTGGTTCTTTCCGCCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCAAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTTATTAGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTC
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTGACGACGAGA
TCCTCGCCGTGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCCGGCTGGCGGAGCCCCCTGATGCTCTTCGTCCAG
ATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCACTGACAACGTGAG
CACAGCTGCGCAAGGAACGCCCCGTGCTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCCTGCAGTTTATTAGGGCAC
CGGACAGGTGGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG

TTCAATGGCCGATCCCATGGTTTAGTTCCCTCACCTTGTCTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTCGAAAGGCCCGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGAGAATGCCGGGCTCCGGAGGACCTTCGGGCGCCCGCCCCGCCCCTGAGCCCGCCCCTGAGCCCGCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCACGTCTT
GCACGACGCGAGCTGCGGGGCGGGGGGAACTTCCCTGACTAGGGGAGGAGTAGAAGGTGGCGCGAAGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTG
ACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGAC
TGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCCGCTGCGGCCCGCTGCCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGCGGTGCCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAACCCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGC
GCCGCCGACCCCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTCCCGCGCCGCGTCCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGCAGGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGT
CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAG

AAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGACAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGC
GTCACCAACCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCGTCTCAGTGTAAGCCATTCCC
CCTTGTCCTATCAAGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTTGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCTCCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCAACCTCTGGGG
GAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCGTCTCAGTGTAAGCCATTCCCCCTTGCTCTATCAAGGG
GGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTTGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCTCCCCTGGCAC
AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA

ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAAAAGCTTGTCCACTCGAGATTCTC
TAGACATCATTAATTCCTAATTTTTTGTGACACTCTATCATTGATAGAGTTATTTTACCCTCCCTATCAGTGATAG
AGAAAAGTGAAATGGCCAAGCCTTTGTCTCAAGAAGAATCCACCCTCATTGAAAGAGCAACGGCTACAATCAACAGC
ATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCTCTCTAGCGACGGCCGCATCTTCACTGGTGTCAATGT
ATATCATTTTTACTGGGGGACCTTGTGTCAGAACTCGTGGTGTCTGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGA
CTTGATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTGAGCCCCGCGGACGGTGTGCGACAGGTGCTTCTCGAT
CTGCATCCTGGGATCAAAGCGATAGTGAAGGACAGTGATGGACAGCCGACGGCAGTTGGGATTCTGTAATTGCTGCC
CTCTGGTTATGTGTGGGAGGGCTAACTCGAGATGAGCTCAGAGACTGGCCCAGTGGCTGTGGACCCACATTGAGAC
GGCGGATCGAGCCCCATGAGTTTGAGGTATTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCCTGCTTTACGAA
ATTAATTGGGGGGGCCGGCACTCCATTTGGCGACATACATCACAGAACCTAACAAGCACGTCGAAGTCAACTTCAT
CGAGAAGTTCACGACAGAAAGATATTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAGCCCAT
GCGGCGAATGTAGTAGGGCCATCACTGAATTCCTGTCAAGGTATCCCCACGTCACCTCTGTTTATTTACATCGCAAGG
CTGTACCACCACGCTGACCCCCGCAATCGACAAGGCCCTGCGGGATTTGATCTCTTCAGGTGTGACTATCCAAATTAT
GACTGAGCAGGAGTCAGGATACTGCTGGAGAACTTTGTGAATTATAGCCCGAGTAATGAAGCCCACTGGCCTAGGT
ATCCCCATCTGTGGGTACGACTGTACGTTCTTGAAGTGTACTGCATCATACTGGGCCTGCCTCCTTGTCTCAACATT
CTGAGAAGGAAGCAGCCACAGCTGACATTCTTTACCATCGCTCTTCAGTCTTGTCAATTACCAGCGACTGCCCCACA
CATTCTCTGGGCCACCGGGTTGAAAAGCGGCAGCGAGACTCCCGGGACCTCAGAGTCCGCCACACCCGAAAGTGATA
AAAAGTATTCTATTGGTTTAGCCATCGGCACTAATTCGGTTGGATGGGCTGTCATAACCGATGAATACAAAGTACCT
TCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCTGA
TAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCAAGAACCGAATAT
GTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGATTCTTCTTTCACCGTTTGGAAGAGTCCTTC
CTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAA
GTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGG
CTCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTGAC
AACTGTTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGC
GAAGGCTATTCTTAGCGCCCGCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGA
AAAATGGGTTGTTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAGTCGAACCTTCGACTTAGCT
GAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCACAATTTGGAGATCA
GTATGCGGACTTATTTTTGGCTGCCAAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTG
AGATTACCAAGGCGCGTTATCCGCTTCAATGATCAAAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAG
GCCCTAGTCCGTGAGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTA
TATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATGGATGGGACGGAAG
AGTTGCTTGTAAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTCGACAACGGTAGCATTCCACATCAA
ATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCCCTCAAAGACAATCGTAAAA
GATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACCTCTCGTTTCGCATGGA
TGACAAGAAAGTCCGAAGAAACGATTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCG
TTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTTACGA

GTATTTACAGTGTAACAATGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCTAAGCG
GAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGAC
TACTTTAAGAAAATTGAATGCTTCGATTCTGTGAGATCTCCGGGTAGAAGATCGATTTAATGCGTCACTTGGTAC
GTATCATGACCTCCTAAAGATAATTAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATA
TAGTGTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGAC
GATAAGGTTATGAAACAGTTAAAGAGGCGTCGTATACGGGCTGGGGACGATTGTGCGGAAACTTATCAACGGGAT
AAGAGACAAGCAAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGC
TGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCAC
GAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCT
AGTTAAGGTCATGGGACGTCACAAACCGGAAAAACATTGTAATCGAGATGGCACGCGAAAATCAAACGACTCAGAAGG
GGCAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAG
CATCCTGTGGAATAACCAATTGCAGAACGAGAACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGT
TGATCAGGAACTGGACATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAAGGACG
ATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGAGGAAGTC
GTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAACTGATAACGCAAAGAAAGTTGATAACTTAAC
TAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGAAACCCGCCAAA
TCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACGAGAACGATAAGCTGATTCCGGAA
GTCAAAGTAATCACTTTAAAGTCAAAATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGAT
AAATAACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAGAAATACCCGAAGC
TAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCGGTAAGATGATCGCGAAAAGCGAACAGGAGATA
GGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGAAAATCACTCTGGCAAACGG
AGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCG
CGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAAGAAAAGTGAAGTGCAGACCGGAGGGTTTTCA
AAGGAATCGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGG
TGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAACTGA
AGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAGAAGACCCCATCGACTTCCTTGAG
GCGAAAGGTTACAAGGAAGTAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGG
CCGAAACGGATGTTGGCTAGCGCCGAGAGCTTCAAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATT
TCCTGTATTTAGCGTCCCATACGAGAAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAG
CAGCACAAACATTATCTCGACGAAATCATAGAGCAAAATTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAA
TCTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATT
TGTTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACT
TCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAACTCGGATAGATTTGTC
ACAGCTTGGGGGTGACTCTGGTGGTTCTACTAATCTGTGAGATATTATTGAAAAGGAGACCGGTAAGCAACTGGTTA
TCCAGGAATCCATCCTCATGCTCCCAGAGGAGGTGGAAGAAGTCATTGGGAACAAGCCGGAAGCGATATACTCGTG
CACACCGCTACGACGAGAGCACCGACGAGAATGTCATGCTTCTGACTAGCGACGCCCTGAATACAAGCCTTGGGC
TCTGGTCATACAGGATAGCAACGGTGAGAACAAGATTAAGATGCTCTCTGGTGGTTCTCCCAAGAAGAAGAGGAAAG
TCTAAAAATTCTAAAATACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTGAATCCTTTTCTGAGGGA
TGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATGTGCATTAGCTGTTTGACGCTCACCTTCTTTCATGGAGTTT

AAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTCTTCATTTCTTTATGTTTTAAATGCACTGACCTCCCACATT
CCCTTTTTAGTAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAAT
CCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAA
TTGGACAGCAAGAAAGCGAGCTTCTAGATGGTCCATATGAATATCCTCCTTAGTTCCTATTCCGCTAGCCTAGAGGG
ACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCGGGGCT
CCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCCGGGCACGGGGAAGGTGGCA
CGGGATCGCTTTCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAGGCTT
TAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTGCAAAGGAGCACAGTGCTCATCCAGAT
CCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATG
CCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGGGAAAACTGCC
TCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGGGAGTTTGCTG
TGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGATCACGGAGAGGCAGATCTTGGGGATAAGGAAG
TGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAG
CCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCA
CAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTT
GCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTTATACAATCG
TCAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCAT
GTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTG
CCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCCTCCCCTGGCACAACGTAAGGCC
ATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCG
CTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATCTATAACAAGAAAATATATATATAATAAGTTATCA
CGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTTAAATCTTAAAAGTCACGTAAGATAATCATGCG
TCATTTTGACTCACGCGGTCGTTATAGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCT
CCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTTCGCGCTATTTAGAAAGAGAGAGCAATATTTCAAGA
ATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAA
CACGTCATAGAATCTTCGTTTGACAAAAACCACATTTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTT
CCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTC
GTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGC
GAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATAT
TTTGTTAAAATTCGCGTTAAATTTTTGTAAATCAGTCTATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTT
ATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAGAAGCGTG
GACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCATCACCTAATCAAGTTT
TTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGC
CGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACG
CTGCGCGTAACCAACACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCAG (SEQID NO: 17)

[0229] TOIC_Cas9_Bsd_p300_0b1_r26_AAVS_PgkNeo

[0230]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCG

76

TCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAG
ATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTCGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGCGAG
CACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTGCAGTTTCATTACAGGGCAC
CGGACAGGTGCGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTAGTTCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTGAAAGGCCCGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGCAGAATGCCGGGCCTCCGGAGGACCTTCGGGCGCCCGCCCCGCCCCTGAGCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCAGTCCCT
GCACGACGCGAGCTGCGGGGCGGGGGGAACTTCTGACTAGGGGAGGAGTAGAAGGTGGCGCGAAGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTG
ACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTCGTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGAC
TGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCCGCTGCGGCCCGCTGCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAACCCCCCCTGCACCCCCCTCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGAGGGCTCGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGGTGTCGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGAATCTCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGC
GCCGCCGACCCCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTGCCGCGCCGCGTCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTGCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG

GAGTCGGTATCGAAGGCCTGACGACAAGGAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGATCTCGGCACCCGCCAACAGAGAAACAGTACGAAACCCCTGGAATAATCAGCTCGCGTTCTGTGT
CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGAAGTTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGC
GTCACCACCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTTGAGAGATACAGCTGGGTCAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAAATCATGAAGGTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAAGTGCCTACACCCCTCTGGGG
GAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCAAGGG
GGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG

TGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTCTAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAATCTTGCTAAATCCAGCCCGACCCTCCCCTGGCAC
AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTAT
TAAGCTTTAGGCGTGACGGTGGGCGCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCTCGTAAAAAGCTTGTCCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGCCATCGGAACTAACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTGCACTCTGGAGAGACAGCTGAAGCACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA
CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCTGGTTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGA
CGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT
TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAAATCTGATTGCACTGTCCTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTGTCTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAACCTGCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCACCAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCTGGCAAGGG
GTAATTTCCGCTTTGCCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGA
GGAAGCCCGCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCTGGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTTCGGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGACGCTGATTACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC

CGTGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCCGCG
AAAATCAGACCACACAGAAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTG
GGATCCCAGATCCTGAAAGAGCACCCCTGTGGAAAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATGCCATCGTCC
CACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTGCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCC
TGATTAAGAAATACCCCCAACTGGAGAGCGAATTCTGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTGGAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT
GGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATTTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTTCGCAAAGGTGGAGA
AGGGGAAAAGCAAGAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAG
CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGAGCTGCAGAAGGGTAATGAAGTGGCC
TGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAAACTCTGGATAAGGTCCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATTTCATCTGTTCACTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTACACAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGATGGCCAAGCCCCTGAGCCAAGAGGAAAGCACCC
TGATCGAGCGGGCCACCGCCACCATCAACAGCATCCCCATCAGCGAGGACTACAGCGTGGCCTCTGCCGCCCTGAGC
AGCGACGGCAGAATCTTACCGGCGTGAACGTGTACCACTTCACAGGCGGGCCCTTGCGCCGAGCTGGTGGTGTCTGGG
AACAGCTGCCGCCGCTGCCGCTGGCAACCTGACCTGTATCGTGGCCATCGGCAACGAGAACCAGGGGCATCCTGAGCC
CCTGCGGCAGATGCAGACAGGTGCTGCTGGACCTGCACCCCGGCATCAAGGCCATCGTGAAGGACAGCGACGGCCAG
CCCACCGCCGTGGGCATTAGAGAGCTGCTGCCCAGCGGCTACGTGTGGGAGGGCTGACTCGAGATTTTCAAACCAGA
AGAACTACGACAGGCACTGATGCCACCCCTGGAAGCCCTGTACCGGCAGGACCCCGAGAGCCTGCCCTTCAGACAGC
CCGTGGATCCCCAGCTGCTGGGCATCCCCGACTACTTCGACATCGTGAAGTCCCCCATGGACCTGAGCACCATCAAG
CGGAAGCTGGACACCGGCCAGTACCAAGAGCCCTGGCAGTACGTGGACGACATCTGGCTGATGTTCAACAACGCCTG
GCTGTACAACAGAAAGACCAGCCGGGTGTACAAGTACTGCAGCAAGCTGAGCGAGGTGTTTCGAGCAAGAGATCGACC
CCGTGATGCAGAGCCTGGGCTACTGCTGCGGCAGAAAGCTGGAATTCAGCCCCCAGACCCTGTGCTGCTACGGCAAG
CAGCTGTGCACCATCCCCCGGACGCCACCTACTACAGTACCAGAACAGATACCACTTCTGCGAGAAGTGCTTCAA
CGAGATCCAGGGCGAGAGCGTGTCCCTGGGCGACGACCCTAGCCAGCCCCAGACCACAATCAACAAAGAGCAGTTCA
GCAAGCGGAAGAACGACACCCTGGACCCCGAGCTGTTCTGTGGAATGCACCGAGTGCGGGCCGAAGATGCACCAGATC
TGCGTGCTGCACCACGAGATCATCTGGCCTGCCGGCTTCGTGTGCGACGGCTGCCTGAAGAAGTCCGCCCGGACCCG

GAAAGAGAACAAGTTCAGCGCCAAGCGGCTGCCCTCTACCCGGCTGGGCACCTTCCTGGAAAACAGAGTGAACGACT
TCCTGCGGGCGGAGAACACCCCGAGTCCGGCGAAGTGACAGTGCAGGTGGTGCACGCCAGCGACAAGACCGTGGA
GTGAAGCCTGGCATGAAGGCCAGATTCTGTGGACAGCGGCGAGATGGCCGAGAGCTTCCCCTACCGGACCAAGGCCCT
GTTTCGCCTTCGAAGAGATCGATGGCGTGGACCTGTGCTTCTTCGGCATGCACGTGCAAGAGTACGGCAGCGACTGCC
CCCCACCAACCAGCGGGCGGGGTACATCAGCTACCTGGACAGCGTGCACCTTCTTCGGGCCAAGTGCCTGCGGACC
CCCGTGTATCACGAGATCCTGATCGGCTACCTGGAATACGTGAAGAAGCTGGGCTACACCACCGGCCACATCTGGGC
CTGTCTCTCCAGCGAGGGCGACGACTACATCTTCCACTGCCACCCCCCGACCAGAAGATCCCCAAGCCCAAGAGAC
TGCAAGAGTGGTACAAGAAGATGCTGGACAAGGCCGTGTCCGAGCGGATCGTGCACGACTACAAGGACATCTTCAAG
CAGGCCACCGAGGACCGGTGACCAGCGCCAAAGAGCTGCCCTACTTCGAGGGCGACTTCTGGCCCAACGTGCTGGA
AGAGAGCATCAAAGAGCTGGAACAAGAGGAAGAGGAACGCAAGCGGGAAGAGAACACCAGCAACGAGAGCACCGACG
TGACCAAGGGCGACAGCAAGAACGCCAAGAAGAACAACAAGAAAACCAGCAAGAACAAGAGCAGCCTGAGCCGG
GGAAACAAGAAAAAGCCCGCATGCCAACGTGTCCAACGACCTGAGCCAGAACTGTACGCCACCATGGAAAAGCA
CAAAGAGGTGTTCTTCGTTCATCCGGCTGATCGCCGACCTGCCGCCAACAGCCTGCCCCCATCGTGGACCCCGACC
CCCTGATCCCCTGCGACCTGATGGACGGCAGGGACGCCTTCTGACCCTGGCCCGGACAAGCACCTGGAATTCTCC
AGCCTGCGGAGAGCCAGTGGTCCACCATGTGCATGCTGGTGGAACTGCACACCCAGAGCCAGGACGAGGGCAGAGG
AAGTCTGCTAACATGCGGTGACGTGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGC
TGTTACCGGGGTGGTGCCATCCTGGTTCGAGCTGGACGGCGACGTAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCC
CACCTCGTGACCACCTTACCTACGGCGTGACGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCT
TCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGC
GCCGAGGTGAAGTTCGAGGGCGACACCCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAA
CATCTGGGGCACAAGCTGGAGTACAACACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCA
TCAAGGTGAAGTTCAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACC
CCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAA
CGAGAAGCGGATCACATGGTCTCTGCTGGAGTTCGTGACCGCCGCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAAACCTAATCTAGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTC
CCCCGTGCCTTCTTGACCCTGGAAGGTGCCACTCCCCTGTCTTCTTCTAATAAAATGAGGAAATTGCATCGCATT
GTCTGAGTAGGTGTCTTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGC
AGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATCCTCTAGTTGGC
GCGTCATGGTCCATATGAATATCCTCCTTAGTTCCTATTCGCTAGCCTAGAGGGACAGCCCCCCCCCAAAGCCCCC
AGGGATGTAATTACGTCCCTCCCCCGTAGGGGCAGCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCC
CCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCTCTGAACG
CTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGA
ATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGG
GTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCAC
AGCCTCCTTGGGCAACCTGTTCAGTGCGTCAACACCTCTGGGGGAAAACTGCCCTCCTCATATCCAACCCAAACCT
CCCCTGTCTCAGTGTAAGCCATTCCCCCTGTCTTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGA
CACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGG
ACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGA

TAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGT
CTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCCTGGAGAGATACAGCTGGGT
CAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGTTACATAAATCGTCAAATCATGAAGGCTGGAAAA
GCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCA
CAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGG
AGAAGGTAAATCTTGCTAAATCCAGCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAG
GACGGAGTCAGTGAGGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTA
TTATATATTAGTCACGATATCTATAACAAGAAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAAC
AATATAATTATCGTATGAGTTAAATCTTAAAAAGTCACGTAAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCTG
TATAGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCC
TAAATGCACAGCGACGGATTGCGCTATTTAGAAAAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCA
GACTATCTTTCTAGGGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAACACGTCATAGAATCTTCGTTTG
ACAAAAACCACATTGTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGC
GGGTGCGCCTATAGTGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCTGGCGT
TACCCAACCTTAATCGCCTTGACGACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCC
CTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCAAATGTAAAGCGTTAATATTTTGTTAAAATTCGCGTTAAAT
TTTTGTAAATCAGCTCATTTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAA
AAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAA
GCACTAAATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGA
AGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCG
CCGCGCTTAATGCGCGCTACAGGGCGGTCAG (SEQ ID NO:18)

[0231] TOIC_Cas9_0b1_r26_AAVS_PgkNeo_Bsd_tracR

[0232]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCTTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTA
CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAAGTGGCGCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCAACAACATGGGGGATCATGTAAGTGGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAAGTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTC
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT

CTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAAGCGCCACGCTTCCCGAAGGGAGAAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTCGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCTGCGCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAACCGCCTCTCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGACAGGTTTTCC
CGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACCTTATACAT
AATAGAAACAATTTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTTCG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCTGCCCCACCCCCACCCCAAGAATAGAAATGACACCTACTCAGACAATGCGATGCAATTC
CTCATTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATT
GCGGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCTGATAGCGGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTACGACGAGA
TCCTCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAG
ATCATCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTTCAGTGACAACGTGAG
CACAGCTGCGCAAGGAACGCCCCGTGCTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCTGCAGTTCATTACAGGGCAC
CGGACAGGTGCGTCTTGACAAAAAGAACCGGGCGCCCCGTGCGTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTTAGTTCCTCACCTGTGCTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTGCAAAGGCCCGGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGAGAATGCCGGGCTCCGGAGGACCTTCGGGCGCCCCGCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCCAGAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTACGTCCT
GCACGACGCGAGCTGCGGGGCGGGGGGGAACCTCCTGACTAGGGGAGGAGTAGAAGGTGGCGGAAGGGGGCCACCAA

AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTG
ACCGCCCAACGACCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCCGCCCGCCCGGCTCTGAC
TGACCGCGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCGTGTGTGTGCGTGGGGAGCGCCGCTGCGGCCCGCGCTGCCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAACCCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGAGGGCTCGGGGAGGGGCGCGCGGCCCGGAGCGCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGAGCCGAAATCTGGGAGGC
GCCGCCGACCCCCCTCTAGCGGGCGGGGCGAAGCGGTGCGGCGCCGCGAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTCCCGCGCCGCGTCCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTCTTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATATTTTGGCAAAGAATTGCGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTCGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGT
CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCGAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGAAGTGTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGAAGTGTAGACATGCTCCCAGCCGATGC
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTACGTCCCTCCCCCGC

TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCGGGCACGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAGGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTG
GTCACCACCCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCTATCAAGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTTCAGTGCGTCACCACCCTCTGGGG
GAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCCTATCAAGGG
GGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCGACCCTCCCCTGGCAC
AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGGTACGGTGGGCGCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCTACCTCGTAAAAAGCTTGTCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGACATCGGAACCTAACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTTCTGACTCTGGAGAGACAGCTGAAGCACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA

CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGA
CGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT
TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAAGAAAAACGGCCTGTTTGAAATCTGATTGCACTGTCAGTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGACAGCTGTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTCTGTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAAGTCTGGTGAAGTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCCACCAGATTCTCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCCTGGCAAGGG
GTAATTCGCTTTGCCTGGATGACACGGAAATCTGAGGAAACAATCACTCCTTGGAACTTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTTATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGA
GGAAGCCCGCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATTGAATGTTTCGATTCTGTGGAGATCAGTGGCGTGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAAGGATAAAGACTTCTTGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTCCGGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGCAGCTGATTACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCGGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC
CGTGAAGGTGGTTCGATGAGCTGGTGAAGTTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCG
AAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTG
GGATCCCAGATCCTGAAAGAGCACCTGTGGAAAACACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCC
CACAGTCATTCTGAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCCGGACCGCCC
TGATTAAGAAATACCCCAAACCTGGAGAGCGAATTCGTGTACGGTGAATAAGGTGTACGATGTCAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTGAAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT

GGGACAAGGGTAGGGATTTTGCCACAGTGC GCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATT CAGTAAGGAATCAATTCTGCCCCAACGGA ACTCTGATAAGCTGATCGCCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTGC GAAAGGTGGAGA
AGGGGAAAAGCAAGAAACTGAAATCCGTCAAGGAGCTGCTGGGTATC ACTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAAACTGCCAAAGTACAG
CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGCAGGAGAGCTGCAGAAGGGTAATGAACTGGCCC
TGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAATCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATT CATCTGTTC ACTCTGACCAACCTGGGAGCCCCTG CAGCCTTCAAGTATTTTGACTACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCTGATT CACCAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACAT
GCGGTGACGTGAGGAGAATCCTGGCCAGCACC GGGATCCATGGTGAGCAAGGGCGAGGAGCTGTT CACCGGGGTG
GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGACCA
CCTTACCTACGGCGTG CAGTGCTTCGCCCCGTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACA ACTACAACAGCCACAAGGTCTATATCACC GCGGACAAGCAGAAGAACGGCATCAAGGTGAAGTTC
AAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACC ACTACCAGCAGAACACCCCCATCGGCGACGG
CCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCT
AGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCTTCCT
TGACCCTGGAAGGTGCCACTCCC ACTGTCTTTCTTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
CATTCTATTCTGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCTGTACAAAA
AAGCAGGCTTTAAAGGAACCAATT CAGTCGACTGGATCCGGTACCAAGTCCGGGCAGGAAGAGGGCCTATTTCCCAT
GATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGA
TATTAGTACAAAATACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAA AATTATGTTTTAAATG
GACTATCATATGCTTACCGTAACTTGAAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACAC
CGAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTCTAG
ACCCAGCTTTCTTGTACAAAGTTGGCATTAA TCTCTAGACATCATTAATTCCTAATTTTTGTTGACTCTATCAT
TGATAGAGTTATTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAATGGCCAAGCCTTTGTCTCAAGAAGAATCC
ACCCTCATTGAAAGAGCAACGGCTACAATCAACAGCATCCCCATCTCTGAAGACTACAGCGTCGCCAGCGCAGCTCT
CTCTAGCGACGGCCGCATCTTCACTGGTGTC AATGTATATCATTTTACTGGGGGACCTTGTGCAGAACTCGTGGTG
TGGGCACTGCTGCTGCTGCGGCAGCTGGCAACCTGACTTGTATCGTCGCGATCGGAAATGAGAACAGGGGCATCTTG
AGCCCCTGCGGACGGTGTGACAGGTGCTTCTCGATCTGCATCCTGGGATCAAAGCGATAGTGAAGGACAGTGATGG
ACAGCCGACGGCAGTTGGGATTCTGTAATTGCTGCCCTCTGGTTATGTGTGGGAGGGCTAAATGGTCCATATGAATA
TCCTCCTTAGTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTC

CCCCGCTAGGGGACAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGC
GTGCGGGGACAGCCCGGGCACGGGAAGGTGGCAGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCC
TGCAGACACCTGGGGGATACGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGC
CTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGG
CTGCCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTT
CAGTGCGTCACCACCCTCTGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCC
ATTCCCCTTGTCTATCAAGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTG
CATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTC
TGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTT
AAAACCCAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGG
TGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTGGGACAGGCAG
CTGGAGAATTGCCATGTAGATGTTTATACAATCGTCAAATCATGAAGGCTGGAAGCCCTCCAAGATCCCCAAGAC
CAACCCCAACCCACCCACCGTGGCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGG
GACGGTGACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAAT
CCAGCCCGACCCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGG
GCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCACGATATC
TATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGTATGAGTT
AAATCTTAAAGTCACGTAAGATAATCATGCGTCATTTTGAATCAGCGGTGCTTATAGTTCAAATCAGTGACA
CTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTC
GCGCTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAGGGTTAAA
AAAGATTTGCGCTTTACTCGACCTAAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATTGTGGGGT
ACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAGTGAGTCG
TATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTAATCGCCTTGC
AGCACATCCCCCTTTGCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCAACAGTTGCGCAGCC
TGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCATTT
TTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCC
AGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATG
GCCCCTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAA
GGGAGCCCCGATTTAGAGCTTGACGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGC
GGGCGCTAGGGCGCTGGCAAGGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCGCTAC
AGGGCGCGTCAG (SEQ ID NO:19)

[0233] TOIC_Cas9_切口酶_0b1_r26_AAVS_PgkNeo

[0234]

GTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCGCTC
ATGAGACAATAACCCGTATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCG
CCTTATTCCCTTTTTTGCGGCATTGTCCTTCCGTGTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTG
AAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCC
GAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCA
AGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCAGTCACAGAAAAGCATCTTA

CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTG
GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT
TCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCTG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTGCCACCTC
TGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCG
GAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCC
CGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTCACACAGGAAACAGCTATGACCATGAT
TACGCCAAGCGCGCAATTAACCTCACTAAAGGGAACCTCCCCTAGCTTAATTAACCCTAGAAAGATAATCATATTG
TGACGTACGTAAAGATAATCATGCGTAAAAATTGACGCATGTGTTTTATCGATCTGTATATCGAGGTTTATTTATTA
ATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTAATAATAAATTCAACAAACAATTTATTTATGTT
TATTTATTTATTAACAAAAAACAACAACTCAAAATTTCTTCTATAAAAGTAACAAACTTTTAAACATTCTCTCTTT
ACAAAAATAAACTTATTTTGTACTTTAAAAACAGTCATGTTGTATTATAAAATAAGTAATTAGCTTAACTTATACAT
AATAGAAACAAATTATACTTATTAATCGCATTGATTATTGACTAGTCGTATTAAGGGTTCCGGATCAGCTTGATTCTG
AGCCCCAGCTGGTTCTTTCCGCCTCAGAAGCCATAGAGCCCACCGCATCCCCAGCATGCCTGCTATTGTCTTCCCAA
TCCTCCCCCTTGCTGTCTGCCCCACCCACCCCGAGAATAGAATGACACCTACTCAGACAATGCGATGCAATTTCT
CTCATTTTTATTAGGAAAGGACAGTGGGAGTGGCACCTTCCAGGGTCAAGGAAGGCACGGGGGAGGGGCAAACAACAG
ATGGCTGGCAACTAGAAGGCACAGTCGAGGCTGATCAGCGAGCTCTAGAGAATTGATCCCCTCAGAAGAACTCGTCA
AGAAGGCGATAGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTGAGCCATTCT
GCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGTCCGCCACACCCAGCCGGCCAC
AGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCCGCAAGCAGGCATCGCCATGGGTGACGACGAGA
TCCTCGCCGTGCGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCGTGATGCTCTTCGTCCAG
ATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCGCTTGGTGGTGAATG
GGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGG
TGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCCCTTCCCGCTCAGTGACAACGTGAG

CACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCCAGTTTCATTACAGGGCAC
CGGACAGGTGGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGTGACAGCCGGAACACGGCGGCATCAGAGCAGCCG
ATTGTCTGTTGTGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGAGAACCTGCGTGCAATCCATCTTG
TTCAATGGCCGATCCCATGGTTTAGTTCCCTCACCTTGTCGTATTATACTATGCCGATATACTATGCCGATGATTAAT
TGTCAACACGTGCTGCTGCAGGTGCGAAAGGCCCGAGATGAGGAAGAGGAGAACAGCGCGGCAGACGTGCGCTTTTG
AAGCGTGAGAATGCCGGGCCTCCGGAGGACCTTCGGGCGCCCCGCCCGCCCCCTGAGCCCCCCCCCTGAGCCCCCCCC
CGGACCCACCCCTTCCCAGCCTCTGAGCCAGAAAAGCGAAGGAGCAAAGCTGCTATTGGCCGCTGCCCCAAAGGCCT
ACCCGCTTCCATTGCTCAGCGGTGCTGTCCATCTGCACGAGACTAGTGAGACGTGCTACTTCCATTTGTCAGTCCCT
GCACGACGCGAGCTGCGGGGCGGGGGGAACTTCTGACTAGGGGAGGAGTAGAAGGTGGCGCGAAGGGGCCACCAA
AGAACGGAGCCGGTTGGCGCTACCGGTGGATGTGGAATGTGTGCGAGCCAGAGGCCACTTGTGTAGCGCCAAGTGC
CCAGCGGGGCTGCTAAAGCGCATGCTCCAGACTGCCTTGGGAAAAGCGCCTCCCCCTACCCGGTAGACACCCACAGT
GGGTGGCCTAGGGACAGGATTGCAACTCCAGTCTTTCTTCTTCTTGGGCGGGAGTCACTAGTTATTAATAGTAATCA
ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCCCCTGGCTG
ACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT
GACGTCAATGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCC
CCTATTGACGTCAATGACGGTAAATGGCCCCCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTGCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATC
TCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTTAAATTATTTTGTGCAGCGATGGGGGCGGGGGGGG
GGGGGCGCGGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAAT
CAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCG
GCGGGCGGGAGTGCCTGCGTTGCCTTCGCCCCGTGCCCCGCTCCGCGCCGCTCGCGCCGCCCGCCCCGGCTCTGAC
TGACCGGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGA
CGGCTCGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCCGGGAGGGCCCTTTGTGCGGGGGGAGCGGCT
CGGGGGGTGCGTGCCTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCCGCGCTGCCGGCGGCTGTGAGCGCTGCG
GGCGCGGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCGCGGTGCGGGG
GGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTGCG
GGCTGTAACCCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGG
GCGTGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGG
CCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGAGCCATT
GCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGC
GCCGCCGACCCCCCTCTAGCGGGCGCGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTC
GTGCGTGCCGCGCCGCCGTCCCCTTCTCCATCTCCAGCCTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGG
GACGGGGCAGGGCGGGGTTCCGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC
TTCTTTTTTCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTCGCCACCAT
GGTGCCCAAGAAGAAGAGGAAAGTCTCTAGACTGGACAAGAGCAAAGTCATAAACTCTGCTCTGGAATTACTCAATG
GAGTGGTATCGAAGGCCTGACGACAAGGAAACTCGCTCAAAAGCTGGGAGTTGAGCAGCCTACCCTGTACTGGCAC
GTGAAGAACAAGCGGGCCCTGCTCGATGCCCTGCCAATCGAGATGCTGGACAGGCATCATACCCACTCCTGCCCCCT
GGAAGGCGAGTCATGGCAAGACTTTCTGCGGAACAACGCCAAGTCATACCGCTGTGCTCTCCTCTCACATCGCGACG
GGGCTAAAGTGCATCTCGGCACCCGCCCAACAGAGAAACAGTACGAAACCCTGGAAAATCAGCTCGCGTTCCTGTGT

CAGCAAGGCTTCTCCCTGGAGAACGCACTGTACGCTCTGTCCGCCGTGGGCCACTTTACACTGGGCTGCGTATTGGA
GGAACAGGAGCATCAAGTAGCAAAAAGAGGAAAGAGAGACACCTACCACCGATTCTATGCCCCCACTTCTGAAACAAG
CAATTGAGCTGTTGACCGGCAGGGAGCCGAACCTGCCTTCCTTTTCGGCCTGGAACATAATCATATGTGGCCTGGAG
AAACAGCTAAAGTGCGAAAAGCGGCGGGCCGACCGACGCCCTTGACGATTTTGACTTAGACATGCTCCCAGCCGATGC
CCTTGACGACTTTGACCTTGATATGCTGCCTGCTGACGCTCTTGACGATTTTGACCTTGACATGCTCCCCGGGTAAA
GCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACC
TCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TAAAGCAATAGCATCACAAATTTACAAAATAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCAT
CAATGTATCTTAAGGGATCCCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGC
TAGGGGCAGCAGCGAGCCGCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGA
CACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGT
TGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCC
AGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGC
GTCACCACCCTCTGGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCC
CCTTGTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCAC
GGAGAGGCAGATCTTGGGGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGAC
ACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACC
CAGCATGGAGAGGAGCACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAG
AATTGCCATGTAGATGTTTCATACAATCGTCAAATCATGAAGGCTGGAAGGCCCTCCAAGATCCCCAAGACCAACCC
CAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGT
GACCCCCCACCTCCGTGGGCAGCTGTGCCACTGCAGCACCCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCC
CGACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTGTC
GACCTAGAGGGACAGCCCCCCCCAAAGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGCAGCAGCGAGC
CGCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGGACAGCCGGGCACGG
GGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGG
GGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTG
CTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCAGAGCCACATCCAGCCT
GGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCTGGGG
GAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAGCCATTCCCCCTTGTCTATCAAGGG
GGAGTTTGTCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAATTCAGTGCATCACGGAGAGGCAGATCTTGG
GGATAAGGAAGTGACAGGACAGCATGGACGTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGC
GTTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAGCA
CAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTGGATGCAAGCAGAAGGGG
TGGAAGAGCTTGCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGACAGGCAGCTGGAGAATTGCCATGTAGATGT
TCATACAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGC
CCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCGACCCTCCCCTGGCAC

AACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGAGGATGGGGCTCAATTGTTTACTCCCTATCA
GTGATAGAGAACGTATGAAGAGTTTACTCCCTATCAGTGATAGAGAACGTATGCAGACTTTACTCCCTATCAGTGAT
AGAGAACGTATAAGGAGTTTACTCCCTATCAGTGATAGAGAACGTATGACCAGTTTACTCCCTATCAGTGATAGAGA
ACGTATCTACAGTTTACTCCCTATCAGTGATAGAGAACGTATATCCAGTTTACTCCCTATCAGTGATAGAGAACGTA
TAAGCTTTAGGCGTGACGGTGGGCGCCTATAAAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGCAATT
CCACAACACTTTTGTCTTATACCAACTTTCCGTACCACTTCCTACCCCTCGTAAAAAGCTTGTCCACCATGGCTCCTA
AGAAAAAGCGGAAGGTGGACAAGAAATACTCAATCGGGCTGGCCATCGGAACCTAACTCAGTGGGGTGGGCAGTCATT
ACTGACGAGTACAAAGTGCCAAGCAAGAAATTTAAGGTCCTGGGCAACACCGATAGGCACTCCATCAAGAAAAATCT
GATTGGGGCCCTGCTGTTTCGACTCTGGAGAGACAGCTGAAGCAACTAGACTGAAAAGGACTGCTAGAAGGCGCTATA
CCCGGCGAAAGAATCGCATCTGCTACCTGCAGGAGATTTTCTCTAACGAAATGGCCAAGGTGGACGATAGTTTCTTT
CATCGGCTGGAGGAATCATTCTGGTCGAGGAAGATAAGAAACACGAGAGACATCCTATCTTTGGAAACATTGTGGA
CGAGGTGCTTATCACGAAAAATACCCACCATCTATCATCTGCGCAAGAACTGGTGGACTCTACAGATAAAGCAG
ACCTGCGGCTGATCTATCTGGCCCTGGCTCACATGATTAAGTTCAGAGGCCATTTTCTGATCGAGGGAGATCTGAAC
CCAGACAATAGCGATGTGGACAAGCTGTTTCATCCAGCTGGTCCAGACATACAATCAGCTGTTTGAGGAAAACCTAT
TAATGCATCTGGCGTGGACGCAAAAGCCATCCTGAGTGCCAGGCTGTCTAAGAGTAGAAGGCTGGAGAACCTGATCG
CTCAGCTGCCAGGCGAAAAGAAAAACGGCCTGTTTGGAAATCTGATTGCACTGTCCTGGGACTGACACCTAACTTC
AAGAGCAATTTTGATCTGGCCGAGGACGCTAAACTGCAGCTGAGCAAGGACACTTATGACGATGACCTGGATAACCT
GCTGGCTCAGATCGGAGATCAGTACGCAGACCTGTTCTGGCCGCTAAGAATCTGTCTGACGCTATCCTGCTGAGTG
ATATTCTGCGGGTGAACACCGAGATTACAAAAGCCCCCTGTCTCAGCTAGCATGATCAAGAGATATGACGAGCACCAT
CAGGATCTGACCCTGCTGAAGGCACTGGTGCGCCAGCAGCTGCCCCGAGAAGTACAAGGAAATCTTCTTTGATCAGAG
TAAGAACGGGTACGCCGTTTATATTGACGGCGGAGCTTCACAGGAGGAATTCTACAAGTTTATCAAACCTATTCTGG
AGAAGATGGACGGCACCGAGGAACTGCTGGTGAAACTGAATCGCGAGGACCTGCTGCGCAAGCAGCGGACATTTGAT
AACGGCTCCATCCCCACAGATTTCATCTGGGAGAGCTGCACGCAATCCTGCGACGACAGGAAGACTTCTACCCATT
TCTGAAGGATAACCGCGAGAAGATCGAAAAAATTTCTGACCTTCCGGATCCCTTACTATGTGGGGCCCCTGGCAAGGG
GTAATTCCCCTTTGCTGGATGACACGGAATCTGAGGAAACAATCACTCCTTGGAACCTCGAGGAAGTGGTCGAT
AAGGGAGCTTCCGCACAGTCTTTTCATCGAGAGAATGACAACTTCGACAAAAACCTGCCAAATGAGAAAGTGCTGCC
TAAGCACAGTCTGCTGTACGAGTATTTACAGTCTATAACGAACTGACTAAGGTGAAATACGTCACCGAGGGGATGA
GGAAGCCCGCTTCTGAGCGGTGAACAGAAGAAAGCTATCGTGACCTGCTGTTTAAACCAATCGCAAGGTGACA
GTCAAGCAGCTGAAGGAGGACTACTTCAAGAAAAATGAATGTTTCGATTCTGTGGAGATCAGTGGCGTCGAAGACAG
ATTTAACGCTTCTCTGGGAACCTACCACGATCTGCTGAAGATCATTAAGGATAAAGACTTCTTGACAACGAGGAAA
ATGAGGATATCCTGGAAGACATTGTGCTGACCCTGACACTGTTTGAGGATCGCGAAATGATCGAGGAACGGCTGAAA
ACTTATGCCCATCTGTTTCGATGACAAGGTGATGAAACAGCTGAAGCGAAGAAGGTACACCGGCTGGGGACGACTGAG
CAGAAAGCTGATCAACGGCATTCCGGACAAACAGAGTGGAAGACTATCCTGGACTTTCTGAAATCAGATGGCTTCG
CTAACAGAAATTTTATGCAGCTGATTCACGATGACAGCCTGACCTTCAAAGAGGATATCCAGAAGGCACAGGTGTCC
GGGCAGGGTGACTCTCTGCACGAGCATATCGCAAACCTGGCCGGTCCCCCGCCATCAAGAAAGGTATTCTGCAGAC
CGTGAAGGTGGTCGATGAGCTGGTGAAAGTCATGGGCAGGCATAAGCCAGAAAACATCGTGATTGAGATGGCCCGCG
AAAATCAGACCACACAGAAAGGACAGAAGAACAGCCGCGAGCGGATGAAAAGGATCGAGGAAGGCATTAAGGAACTG
GGATCCCAGATCCTGAAAGAGCACCTGTGGAACAACTCAGCTGCAGAATGAGAAGCTGTATCTGTACTATCTGCA
GAATGGGCGGGATATGTACGTGGACCAGGAGCTGGATATTAACCGACTGTCTGATTACGACGTGGATCATATCGTCC

CACAGTCATTCTGAAAAGATGACAGCATTGACAATAAGGTGCTGACCCGGAGTGACAAAAACCGAGGAAAGAGTGAT
AATGTCCCTTCAGAGGAAGTGGTCAAGAAAATGAAGAACTACTGGAGACAGCTGCTGAATGCCAACTGATCACACA
GCGAAAGTTTGATAACCTGACTAAAGCTGAGAGAGGGGGTCTGTCAGAACTGGACAAAGCAGGCTTCATCAAGCGAC
AGCTGGTGGAGACCAGACAGATCACAAAGCACGTCGCTCAGATTCTGGATAGCAGGATGAACACAAAGTACGATGAG
AATGACAACTGATCCGCGAAGTGAAGGTCATTACTCTGAAGTCAAACTTGTGAGCGACTTCAGAAAGGATTTCCA
GTTCTACAAAGTCAGGGAGATCAACAATTATCACCATGCTCATGACGCATACCTGAACGCAGTGGTCGGGACCGCCC
TGATTAAGAAATACCCCAAACTGGAGAGCGAATTCGTGTACGGTGACTATAAGGTGTACGATGTCAGAAAAATGATC
GCCAAGAGTGAGCAGGAAATTGGAAAAGCCACCGCTAAGTATTTCTTTTACTCAAACATCATGAATTTCTTTAAGAC
TGAGATCACCTGGCAAATGGGGAAATCCGAAAGAGACCACTGATTGAGACTAACGGCGAGACCGGAGAAATCGTGT
GGGACAAGGGTAGGGATTTTGCCACAGTGCGCAAGGTCCTGTCCATGCCTCAAGTGAATATTGTCAAGAAAACAGAG
GTGCAGACTGGCGGATTCAGTAAGGAATCAATTCTGCCCAAACGGAACCTCTGATAAGCTGATCGCCCGAAAGAAAGA
CTGGGATCCCAAGAAATATGGGGGTTTCGACTCCCCAACAGTGGCTTACTCTGTCTGGTGGTCGCAAAGGTGGAGA
AGGGGAAAAGCAAGAACTGAAATCCGTCAAGGAGCTGCTGGGTATCACTATTATGGAGAGGAGCTCCTTCGAGAAG
AACCCCATCGATTTTCTGGAGGCTAAAGGCTATAAGGAAGTGAAGAAAGACCTGATCATTAACTGCCAAAGTACAG
CCTGTTTGAGCTGGAAAACGGAAGGAAGCGAATGCTGGCATCCGAGGAGAGCTGCAGAAGGGTAATGAACTGGCCC
TGCCTTCTAAGTACGTGAACTTCCTGTATCTGGCTAGCCACTACGAGAAGCTGAAAGGCTCCCCGAGGATAACGAA
CAGAAACAGCTGTTTGTGGAGCAGCACAAGCATTATCTGGACGAGATCATTGAACAGATTAGCGAGTTCTCCAAAAG
AGTGATCCTGGCTGACGCAAATCTGGATAAGGTCCTGAGCGCATACAACAAACACAGAGATAAGCCAATCAGGGAGC
AGGCCGAAAATATCATTTCATCTGTTCACTCTGACCAACCTGGGAGCCCCTGCAGCCTTCAAGTATTTTGACACTACC
ATCGATCGGAAACGATACACATCCACTAAGGAGGTGCTGGACGCTACCCTGATTCACCAGAGCATTACCGGCCTGTA
TGAAACAAGGATTGACCTGTCTCAGCTGGGGGGCGACCTCGAGGGAAGCGGAGAGGGCAGAGGAAGTCTGCTAACAT
GCGGTGACGTCGAGGAGAATCCTGGCCCAGCACCGGGATCCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG
GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC
CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCTGCCCTGGCCCACCCTCGTGACCA
CCTTCACCTACGGCGTGAGTGCTTCGCCCCGCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTCCGCCATG
CCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTT
CGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACATAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTC
AAGACCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG
CCCCGTGCTGCTGCCCAGAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC
ACATGGTCTCTGCTGGAGTTCGTGACCGCCCGGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAACCTAATCT
AGCAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCTTCTCT
TGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGT
CATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACAGCTGGGGCTCGATCCTCTAGTTGGCGCGTCATGGTCCAT
ATGAATATCCTCCTTAGTTCTATTCCGCTAGCCTAGAGGGACAGCCCCCCCCAAAGCCCCAGGGATGTAATTAC
GTCCCTCCCCCGCTAGGGGCAGCAGCGAGCCGCCCCGGGGCTCCGCTCCGGTCCGGCGCTCCCCCGCATCCCCGAGC
CGGCAGCGTGCGGGGACAGCCCGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCT
TTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGAATCATA

GAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCTGCTATGTGCAGGGTCATCAACCAGCA
CCCCAGGCTGCCAGAGCCACATCCAGCCTGGCCTTGAATGCCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCA
ACCTGTTCAGTGCGTCAACCACCTCTGGGGGAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTG
TAAAGCCATTCCCCCTTGCTCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTGCCAA
TTCAGTGATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGACGTGGGACATGCAGGTGTTG
AGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCCTTAAGGATAAGAAGATAGGATAGAAGGACAAAGA
GCAAGTTAAACCCAGCATGGAGAGGAGCAGAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGT
TTGATGGTGTCTGGATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTGAGTAGGACTGGGA
CAGGCAGCTGGAGAATTGCCATGTAGATGTTTACATAATCGTCAAATCATGAAGGCTGGAAAAGCCCTCCAAGATCC
CCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCCCTCAGTGCCACATCCCCACAGTTCTTCATCAC
CTCCAGGGACGGTGACCCCCCACCTCCGTGGGACAGTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTT
GCTAAATCCAGCCCGACCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACCTCCAGGACGGAGTCAGTGA
GGATGGGGCTGGATCCGAAGCAGCTCCAGCCTACACAATCGCTCAAGACGTGTAATGCTTTTATTATATATTAGTCA
CGATATCTATAACAAGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAATTATCGT
ATGAGTTAAATCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGACTCACGCGGTCGTTATAGTTCAAAATC
AGTGACACTTACCGCATTGACAAGCACGCCTCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGA
CGGATTGCGCTATTTAGAAAGAGAGAGCAATATTTCAAGAATGCATGCGTCAATTTTACGCAGACTATCTTTCTAG
GGTTAAAAAAGATTTGCGCTTTACTCGACCTAACTTTAAACACGTCATAGAATCTTCGTTTGACAAAAACCACATT
GTGGGGTACCGAGCTCTTAATTAAGGCGCGCCGGGAGGTTCCCTTTAGTGAGGGTTAATTGCGGGTCGCCCTATAG
TGAGTCGTATTACAATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTAATC
GCCTTGACAGCACATCCCCCTTTGCGCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG
CGCAGCCTGAATGGCGAATGGCAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTAAATCAG
CTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTG
TTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCACTAAATCGGAA
CCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGA
AAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCG
CCGCTACAGGGCGCGTCAG (SEQ ID NO:20)

[0235] 本文公开的所有参考文献、专利和专利申请通过引用每一个被引用的主题纳入，其中在一些情况中，可以包括整个文件。

[0236] 在本申请的说明书和权利要求书中，除非有明确的相反说明，不定冠词“一个”和“一种”应理解为“至少一个(种)”。

[0237] 还应理解，除非清楚表明相反，在本文要求权利的包括多于一个步骤或行动的任何方法中，所述方法的步骤或行动的顺序不必限定是列举该方法步骤或行为的顺序。

[0238] 在权利要求以及上述说明书中，所有连接词例如“包括”、“含有”、“带有”、“具有”、“包含”、“涉及”、“持有”、“由……组成”等应被理解为是开放式的，即，指包括但不限于。如美国专利局的专利审查程序手册中2111.03部分所列举，仅连接词“由……组成”和“基本由……组成”分别应是封闭或半封闭的连接词。

序列表

<110> 阿斯特拉捷利康股份公司 (ASTRAZENECA AB)

<120> 细胞基因组的诱导型修饰

<130> A1106.70000W000

<140> Not Yet Assigned

<141> 2016-10-14

<150> US 62/242,884

<151> 2015-10-16

<160> 20

<170> PatentIn version 3.5

<210> 1

<211> 72

<212> DNA

<213> 智人 (Homo sapiens)

<400> 1

catcttggac cattagctcc acaggtatct tcttcctct agtggtcata acagcagctt 60
cagctacctc tc 72

<210> 2

<211> 32

<212> DNA

<213> 智人 (Homo sapiens)

<400> 2

caaattatcc tgccccctag acataacctc cc 32

<210> 3

<211> 70

<212> DNA

<213> 智人 (Homo sapiens)

<400> 3

tgcattggct gcccaggcct gactgccgc ctgccggcag gggtcagtc cagagaccc 60
agctccctgc 70

<210> 4

<211> 61

<212> DNA

<213> 智人 (Homo sapiens)

<400> 4

gccggccgcg gaccgggcc ctcctcccc ggccgctagg gggcgggcc ggatcacagg 60
a 61

<210> 5

<211> 61

<212> DNA

<213> 智人(Homo sapiens)

<400> 5

catgcacaag aaagctttgc actttgcgaa ccaacgatag gtgggggtgc gtggaggatg 60
g 61

<210> 6

<211> 1209

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 6

ctagagggac agcccccccc caaagcccc agggatgtaa ttacgtccct cccccgctag 60
gggcagcagc gagccgcccc gggtccgct ccggtccggc gtccccccg cateccccgag 120
ccggcagcgt gcggggacag cccgggcacg gggaaggtgg cacgggatcg ctttcctctg 180
aacgtttctc gctgctcttt gagcctgcag acacctgggg ggatacgggg aaaaagcttt 240
aggctgaaag agagatttag aatgacagaa tcatagaacg gcctgggttg caaaggagca 300
cagtgtcat ccagatccaa cccctgcta tgtgcagggt catcaaccag cagcccaggc 360
tgcccagagc cacatccagc ctggccttga atgcctgcag ggatggggca tccacagcct 420
ccttgggcaa cctgttcagt gcgtcaccac cctctggggg aaaaactgcc tcctcatatc 480
caacccaaac ctccccgtc tcagtgtaaa gccattcccc ctgttcctat caagggggag 540
tttctgtga cattgttggc ctggggtgac acatgtttgc caattcagt catcacggag 600
aggcagatct tggggataag gaagtgcagg acagcatgga cgtgggacat gcaggtgttg 660
agggtcttgg gacactctcc aagtcacagc gtccagaaca gccttaagga taagaagata 720
ggatagaagg acaaagagca agttaaaacc cagcatggag aggagcaca aaaggccaca 780
gacactgctg gtccctgtgt ctgagcctgc atgtttgatg gtgtctggat gcaagcagaa 840
ggggtggaag agcttgccct gagagataca gctgggtcag taggactggg acaggcagct 900
ggagaattgc catgtagatg ttcatacaat cgtcaaatca tgaaggctgg aaaagccctc 960
caagatcccc aagaccaacc ccaaccacc caccgtgcc actggccatg tccctcagt 1020
ccacatcccc acagttcttc atcacctcca gggacggtga cccccacc tccgtgggca 1080
gctgtgccac tgcagaccg ctctttggag aaggtaaate ttgctaaate cagcccagacc 1140
ctccccctggc acaacgtaag gccattatct ctcatccaac tccaggacgg agtcagtgag 1200
gatggggct 1209

<210> 7

<211> 6012

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 7

```

atggctctcg agccatctgc tggagacatg agagctgcc aacctttggcc aagccccgctc 60
atgatcaaac gctctaagaa gaacagcctg gccttgtccc tgacggccga ccagatggtc 120
agtgccttgt tggatgctga gccccccata ctctattccg agtatgatcc taccagaccc 180
ttcagtgaag cttcgatgat gggcttactg accaacctgg cagacaggga gctggttcac 240
atgatcaact gggcgaagag ggtgccaggc tttgtggatt tgacctcca tgatcaggtc 300
caccttctgg aatgtgcctg gctagagatc ctgatgattg gtctcgtctg gcgctctatg 360
gagcaccagc tgaagctact gtttgcctct aacttgcctt tggacaggaa ccagggaaaa 420
tgtgtagagg gcatggtgga gatcttcgac atgctgctgg ctacatcatc tcggttccgc 480
atgatgaatc tgcagggaga ggagtttgtg tgccctcaat ctattatttt gcttaattct 540
ggagtgtaca catttctgtc cagcacctg aagtctctgg aagagaagga ccatatccac 600
cgagtccagg acaagatcac agacaatttg atccacctga tggccaaggc aggctgacc 660
ctgcagcagc agcaccagcg gctggcccag ctctctctca tctctctcca catcaggcac 720
atgagtaaca aaggcatgga gcactctgtac agcatgaagt gcaagaacgt ggtgcccctc 780
tatgacctgc tgctggaggc ggcgagcgcc caccgcctac atgcgccac tagccgtgga 840
ggggcatccg tggaggagac ggaccaaagc cacttgccca ctgcgggctc tacttcatcg 900
cattccttgc aaaagtatta catcacgggg gaggcagagg gtttccttgc cacagctgac 960
aagaaatact caatcgggct ggacatcgga actaactcag tggggtgggc agtcattact 1020
gacgagtaca aagtccaag caagaaattt aagtccttgg gcaacaccga taggcactcc 1080
atcaagaaaa atctgattgg ggccctgctg ttcgactctg gagagacagc tgaagcaact 1140
agactgaaaa ggactgctag aaggcgctat acccgcgcaa agaategcat ctgctacctg 1200
caggagattt tctctaacga aatggccaag gtggacgata gtttctttca tcggctggag 1260
gaatcattcc tggctcagga agataagaaa cacgagagac atcctatctt tggaacatt 1320
gtggacgagg tcgcttatca cgaanaatac cccaccatct atcatctgcg caagaaactg 1380
gtggactcta cagataaagc agacctgcgg ctgatctatc tggccctggc tcacatgatt 1440
aagttcagag gccattttct gatcgaggga gatctgaacc cagacaatag cgatgtggac 1500
aagctgttca tccagctggt ccagacatac aatcagctgt ttgaggaaaa ccctattaat 1560
gcactctggcg tggacgcaaa agccatcctg agtgccaggc tgtctaagag tagaaggctg 1620
gagaacctga tcgctcagct gccaggcgaa aagaaaaacg gcctgtttgg aaatctgatt 1680
gcactgtcac tgggactgac acctaatctc aagagcaatt ttgatctggc cgaggacgct 1740
aaactgcagc tgagcaagga cacttatgac gatgacctgg ataacctgct ggctcagatc 1800
ggagatcagt acgcagacct gttcctggcc gctaagaatc tgtctgacgc tatectgctg 1860
agtgatattc tgcgggtgaa caccgagatt acaaaaagccc ctctgtcagc tagcatgac 1920
aagagatatg acgagacca tcaggatctg acctgctga aggcactggg gcgccagcag 1980
ctgcccgaga agtacaagga aatctttctt gatcagagta agaacgggta cgccggttat 2040
attgacggcg gagcttcaca ggaggaattc tacaagtta tcaaacctat tctggagaag 2100
atggacggca ccgaggaact gctggtgaaa ctgaatcgcg aggacctgct gcgcaagcag 2160
cggacatttg ataacggctc catccccac cagattcatc tgggagagct gcacgcaatc 2220

```

```

ctgcgacgac aggaagactt ctaccattt ctgaaggata accgcgagaa gatcgaaaaa 2280
attctgacct tccggatccc ttactatgtg gggcccttg caaggggtaa ttcccgttt 2340
gcctggatga cacggaaatc tgaggaaaca atcactcctt ggaacttcga ggaagtggtc 2400
gataagggag cttccgcaca gtctttcatc gagagaatga caaacttcga caaaaacctg 2460
ccaaatgaga aagtgtgcc taagcacagt ctgctgtacg agtatttcac agtctataac 2520
gaactgacta aggtgaaata cgtcaccgag gggatgagga agcccgctt cctgagcgg 2580
gaacagaaga aagctatcgt ggacctgtg tttaaaacca atcgcaagg gacagtcaag 2640
cagctgaagg aggactactt caagaaaatt gaatgtttcg attctgtgga gatcagtggc 2700
gtcgaagaca gatttaacgc ttctctggga acctaccacg atctgctgaa gatcattaag 2760
gataaagact tcctggacaa cgaggaaaat gaggatatcc tggaagacat tgtgtgacc 2820
ctgacactgt ttgaggatcg cgaaatgatc gaggaacggc tgaaaactta tgccatctg 2880
ttcgatgaca aggtgatgaa acagctgaag cgaagaagg acaccggctg gggacgactg 2940
agcagaaaagc tgatcaacgg cattcgggac aaacagagt gaaagactat cctggacttt 3000
ctgaaatcag atggcttcgc taacagaaat tttatgcagc tgattcacga tgacagcctg 3060
accttcaaag aggatatcca gaaggcacag gtgtccgggc aggtgactc tctgcacgag 3120
catatcgcaa acctggccgg gtccccgcc atcaagaaag gtattctgca gaccgtgaag 3180
gtggctgatg agctggtgaa agtcatgggc aggcataagc cagaaaacat cgtgattgag 3240
atggcccgcg aaaatcagac cacacagaaa ggacagaaga acagccgca gcggatgaaa 3300
aggatcgagg aaggcattaa ggaactggga tcccagatcc tgaaagagca ccctgtggaa 3360
aacactcagc tgcagaatga gaagctgtat ctgtactatc tgcagaatgg gcgggatatg 3420
tacgtggacc aggagctgga tattaaccga ctgtctgatt acgacgtgga tcatatcgtc 3480
ccacagtcac tcctgaaaga tgacagcatt gacaataagg tgctgacctg gactgacaaa 3540
aaccgaggaa agagtgataa tgtcccttca gaggaagtgg tcaagaaaat gaagaactac 3600
tgagacagc tgctgaatgc caaactgatc acacagcgaa agtttgataa cctgactaaa 3660
gctgagagag ggggtctgtc agaactggac aaagcaggct tcatcaagcg acagctggtg 3720
gagaccagac agatcacaaa gcacgtcgct cagattctgg atagcaggat gaacacaaag 3780
tacgatgaga atgacaaact gatccgcgaa gtgaaggtea ttactctgaa gtcaaaactt 3840
gtgagcgact tcagaaagga tttccagttc tacaaagtea gggagatcaa caattatcac 3900
catgctcatg acgcatacct gaacgcagt gtcgggaccg ccctgattaa gaaatacccc 3960
aaactggaga gcgaattcgt gtacggtgac tataaggtgt acgatgtcag aaaaatgatc 4020
gccaagagtg agcaggaaat tggaaaagcc accgctaagt atttctttta ctcaaaccatc 4080
atgaatttct ttaagactga gatcacctg gcaaatgggg aatccgaaa gagaccactg 4140
attgagacta acggcgagac cggagaaatc gtgtgggaca agggtaggga ttttgccaca 4200
gtgcgcaagg tcctgtccat gcctcaagt aatattgtca agaaaacaga ggtgcagact 4260
ggcggattca gtaaggaatc aattctgccc aaacggaact ctgataagct gatcgccga 4320
aagaaagact gggatcccaa gaaatatggg ggtttcgact cccaacagt ggcttactct 4380
gtcctgggtg tcgcaaaggt ggagaagggg aaaagcaaga aactgaaatc cgtcaaggag 4440
ctgctgggta tcaattat ggagaggagc tccttcgaga agaaccat cgattttctg 4500
gaggctaaag gctataagga agtgaagaaa gacctgatca ttaaactgcc aaagtacagc 4560

```

ctgtttgagc tggaaaacgg aaggaagcga atgctggcat ccgcaggaga gctgcagaag 4620
 ggtaatgaac tggccctgcc ttctaagtac gtgaacttcc tgtatctggc tagccactac 4680
 gagaagctga aaggctcccc cgaggataac gaacagaaac agctgtttgt ggagcagcac 4740
 aagcattatc tggacgagat cattgaacag attagcgagt tctccaaaag agtgatcctg 4800
 gctgacgcaa atctggataa ggtcctgagc gcatacaaca aacacagaga taagccaatc 4860
 agggagcagg ccgaaaatat cattcatctg ttcactctga ccaacctggg agccccctgca 4920
 gccttcaagt attttgacac taccatcgat cggaacgat acacatccac taaggaggtg 4980
 ctggacgcta ccctgattca ccagagcatt accggcctgt atgaaacaag gattgacctg 5040
 tctcagctgg ggggcgacct cgagccatct gctggagaca tgagagctgc caacctttgg 5100
 ccaagcccg ccatgatcaa acgctctaag aagaacagcc tggccttgte cctgacggcc 5160
 gaccagatgg tcagtgcctt gttggatgct gagccccca tactctatc cgagtatgat 5220
 cctaccagac ccttcagtga agcttcgatg atgggcttac tgaccaacct ggcagacagg 5280
 gagctggttc acatgatcaa ctgggcgaag aggggtccag gctttgtgga tttgacctc 5340
 catgatcagg tccaccttct ggaatgtgcc tggetagaga tctgatgat tggctctgct 5400
 tggcgtcta tggagcacc agtgaagcta ctgtttgct ctaacttget cttggacagg 5460
 aaccagggaa aatgtgtaga gggcatggtg gagatcttcg acatgctgct ggctacatca 5520
 tctcggttcc gcatgatgaa tctgcaggga gaggagtttg tgtgcctcaa atctattatt 5580
 ttgcttaatt ctggagtgt cacatttctg tccagcacc tgaagtctct ggaagagaag 5640
 gaccatatcc accgagtcct ggacaagatc acagacactt tgatccacct gatggccaag 5700
 gcaggcctga ccctgcagca gcagcaccag cggttgccc agctcctcct catcctctcc 5760
 cacatcaggc acatgagtaa caaaggcatg gagcatctgt acagcatgaa gtgcaagaac 5820
 gtggtgcccc tctatgacct gctgctggag gcggcggacg cccaccgct acatgcgccc 5880
 actagccgtg gaggggcatc cgtggaggag acggaccaa gccacttgcc cactgcgggc 5940
 tctacttcat cgcatcctt gcaaaagtat tacatcacgg gggaggcaga gggtttcct 6000
 gccacagctt ga 6012

<210> 8

<211> 17575

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 8

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
 caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
 ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
 gccttcctgt ttttgtctac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt 300
 ttgcctccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 360
 tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420

atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
 gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
 ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
 ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaaactggc gaactactta 720
 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
 ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 840
 gtgggtctcg cggatatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
 agattgatgtt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
 atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
 aaaagatcaa aggatcttct tgagatcett tttttctgcg cgtaatctgc tgccttgcaa 1200
 caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
 ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
 cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 1380
 tctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
 gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc 1500
 ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
 gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
 caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 1740
 tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg 1800
 ctcacatgtt ctttctgctg ttatcccctg attctgtgga taaccgtatt accgcctttg 1860
 agtgagctga taccgtctgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
 aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat 1980
 gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
 tgagttagct cactcattag gcacccagc ctttacactt tatgcttcg gctcgtatgt 2100
 tgtgtggaat tgtgagcgga taacaatttc acacaggaac cagctatgac catgattacg 2160
 ccaagcgcgc aattaaccct cactaaaggg aacctccct agcttaatta accctagaaa 2220
 gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
 atcgatctgt atatcgaggt ttattttatta atttgaatag atattaagtt ttattatatt 2340
 tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
 aaaaaaaaca aaaactcaaa attttcteta taaagtaaca aaacttttaa acattctctc 2460
 ttttacaaaa ataaacttat ttgtacttt aaaaacagtc atgttgatt ataaaataag 2520
 taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
 gactagtcgt attaagggtt ccggatcagc ttgattcgag ccccgctgg tttttccgc 2640
 ctcagaagcc atagagccca ccgcacccc agcatgcctg ctattgtctt cccaatcctc 2700
 ccccttgctg tctgccccca cccaccccc cagaatagaa tgacacctac tcagacaatg 2760

cgatgcaatt tcctcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
 ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcgaggc 2880
 tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
 ggcatgcgc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
 ttgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
 cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
 attcggcaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
 cttgagcctg gcgaacagtt cggtcggcgc gagccccctga tgctcttcgt ccagatcatc 3240
 ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
 gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgcattg catcagccat 3360
 gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc 3420
 gcccaatagc agccagtccc ttcccgettc agtgacaacg tcgagcacag ctgcgcaagg 3480
 aacgcccgtc gtggccagcc acgatagccg cgtgcctcg tcttcagtt cattcagggc 3540
 accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
 ggcgcatca gagcagccga ttgtctgttg tgcccagtea tagccgaata gcctctccac 3660
 ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720
 gttctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
 acgtgctgct gcaggtcgaa aggccccgag atgaggaaga ggagaacagc gcggcagacg 3840
 tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcgggcg cccgccccgc 3900
 ccttgagccc gcccctgagc ccgcccccg acccaccct tcccagctc tgagcccaga 3960
 aagcgaagga gcaaagctgc tattggccgc tgcccaaag gcctaccgc ttccattgct 4020
 cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttcatt tgtcacgtcc 4080
 tgcacgacgc gagctgcggg gcggggggga acttctgac taggggagga gtagaaggtg 4140
 gcgcgaaggg gccaccaaag aacggagccg gttggcgcct accggtggat gtggaatgtg 4200
 tgcgagccag aggccacttg ttagcgcga agtgcccagc ggggctgcta aagcgcgtgc 4260
 tccagactgc cttgggaaaa gcgcctcccc taccggtag acaccacaca gtgggtggcc 4320
 tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
 tagtaatcaa ttacggggtc attagtcat agcccatata tggagtccg cgttacataa 4440
 cttacggtaa atggcccgc tggtgaccg cccaacgacc cccgccatt gacgtcaata 4500
 atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
 tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgcc 4620
 cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 4680
 tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac catgggtcga 4740
 ggtgagcccc acgttctgct tcactctccc catctcccc cctctccac cccaatttt 4800
 gtatttattt attttttaat tattttgtgc agcatgggg gcgggggggg ggggggcgcg 4860
 cgccaggcgg gcgggggcgg gcgaggggc ggggcggggc gaggcggaga ggtgcggcgg 4920
 cagccaatca gagcggcgcg ctccgaaagt ttccttttat ggcgagcg cgcggcggc 4980
 ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgct tcgccccgtg 5040
 ccccgctccg cgccgcctcg cgccgcccgc cccggtctg actgaccgc ttactccac 5100

```

aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa ttagcgcttg gtttaatgac 5160
ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
gcggggggga gcggctcggg ggggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
ggccccgcgt gcccggcggc tgtgagcgct gcgggcgcgg cgcggggctt tgtgcgctcc 5340
gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg gggctgcgag 5400
gggaacaaag gctgcgtgcg ggggtgtgtc gtgggggggt gagcaggggg tgtgggcgcg 5460
gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc acggccccgc 5520
ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg gcgggggggt 5580
gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
ggggcgcggc ggccccggag cgccggcggc tgtcaggcgg cgcgagccg cagccattgc 5700
cttttatggt aatcgtgcga gagggcgag ggacttcctt tgtcccaaatt ctggcggagc 5760
cgaaatctgg gaggcgccgc cgcaccccc ctacggggcg cgggcgaagc ggtgcggcgc 5820
cggcaggaag gaaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgc gtcccccttct 5880
ccatctccag cctcggggct gccgcagggg gacggtgcc ttcggggggg acggggcagg 5940
gcgggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa catgttcat 6000
gccttcttct ttttctaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
tttgcaaaag aattcgccac catggtgcc aagaagaaga ggaaagtctc tagactggac 6120
aagagcaaag tcataaactc tgctctggaa ttactcaatg gagtcggtat cgaaggcctg 6180
acgacaagga aactcgctca aaagctggga gttgagcagc ctaccctgta ctggcacgtg 6240
aagaacaagc gggccctgct cgatgccctg ccaatcgaga tgctggacag gcatcatacc 6300
cactcctgcc ccctggaagg cgagtcatgg caagacttct tgcggaacaa cgccaagtca 6360
taccgctgtg ctctcctctc acatcgcgac ggggctaaag tgcatctcgg caccgcacca 6420
acagagaaac agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggcttc 6480
tccttgagga acgcactgta cgctctgtcc gccgtgggcc actttacact gggctgcgta 6540
ttggaggaac aggagcatca agtagcaaaa gaggaaagag agacacctac caccgattct 6600
atgcccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
ttccttttct gcctggaact aatcatatgt ggcttgaga aacagctaaa gtgcgaaagc 6720
ggcgggcccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
gacgactttg accttgatat gctgcctgct gacgtctttg acgattttga cttgacatg 6840
ctccccgggt aaagcggccg cgactctaga tcataatcag ccataaccaca tttgtagagg 6900
ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
tcacaaatct cacaaataaa gcattttttt cactgcattc tagttgttgt ttgtccaaac 7080
tcataatgt atcttaagg atccctagag ggacagcccc ccccaaaagc cccagggat 7140
gtaattacgt ccctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggtc 7200
cggcgctccc cccgcacccc cgagccggca gcgtgcgggg acagcccggg cagggggaag 7260
gtggcacggg atcgctttcc tctgaacgt tctcgtgct ctttgagcct gcagacacct 7320
gggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatacatg 7380
aacggcctgg gttgcaaagg agcacagtgc tcataccagat ccaaccccc gctatgtgca 7440

```

```

gggtcatcaa ccagcagccc aggctgccc gagccacatc cagcctggcc ttgaatgcct 7500
gcagggatgg ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccaccctctg 7560
ggggaaaaaac tgcctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
cccccttgtc ctatcaaggg ggagtttgct gtgacattgt tggctctgggg tgacacatgt 7680
ttgccaattc agtgcacac ggagaggcag atcttgggga taaggaagtg caggacagca 7740
tggaactggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
atcatgaagg ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt 8100
gcccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
gtgaccccc caccctcgtg ggcagctgtg ccaactgcag accgctcttt ggagaaggta 8220
aatcttgcta aatccagccc gaccctcccc tggcacaacg taaggccatt atctctcacc 8280
caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
ggctccgctc cggctccggc cccccccgc atccccgagc cggcagcgtg cggggacagc 8460
ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgttctctg ctgctctttg 8520
agcctgcaga cactggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
atgacagaat catagaacgg cctgggttgc aaaggagcac agtgcctacc cagatccaac 8640
cccctgctat gtgcagggc atcaaccagc agcccaggt gcccagagcc acatccagcc 8700
tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
cagtgtaaag ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
tggggtgaca catgtttgcc aattcagtc atcacggaga ggcagatctt ggggataagg 8940
aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000
agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
gttaaaaacc agcatggaga ggagcacaac aaggccacag aactgctgg tccctgtgtc 9120
tgagcctgca tgtttgatgg tgtctggatg caagcagaag gggtggaaga gcttgccctg 9180
agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
tcatacaatc gtcaaatcat gaaggttgga aaagcctcc aagatcccca agaccaaccc 9300
caaccacccc accgtgcccc ctggccatgt cctcagtgcc caccatccca cagttcttca 9360
tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
tctttggaga aggtaaatct tgctaaatcc agcccagccc tcccctggca caacgtaagg 9480
ccattatctc tcaccaact ccaggacgga gtcagtggag atggggctca attgtttact 9540
ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
tagagaacgt atgaccagtt tactccctat cagtataga gaacgtatct acagtttact 9720
ccctatcagt gatagagaac gtatatccag ttactccct atcagtata gagaacgtat 9780

```

aagcttttagg cgtgtacggt gggcgcctat aaaagcagag ctcgtttagt gaaccgtcag 9840
atcgccctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttccta 9900
ccctcgtaaa aagcttgtcc accatggctc ctaagaaaaa gcggaagggtg gacaagaaat 9960
actcaatcgg gctggacatc ggaactaact cagtgggggtg ggcagtcatt actgacgagt 10020
acaaagtgcc aagcaagaaa ttttaagggtcc tgggcaacac cgataggcac tccatcaaga 10080
aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
ttttctctaa cgaaatggcc aagggtggacg atagttttctt tcctcggtcg gaggaatcat 10260
tcctggctga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
aggctcgctta tcacgaaaaa taccaccacca tctatcattc gcgcaagaaa ctgggtggact 10380
ctacagataa agcagacctg cggtgatct atctggccct ggctccatg attaagttca 10440
gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
tcattccagct ggtccagaca tacaatcagc tgtttgagga aaacctatt aatgcattctg 10560
gcgtggacgc aaaagccatc ctgagtcca ggctgtctaa gagtagaagg ctggagaacc 10620
tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcaactgt 10680
cactgggact gacacctaac ttcaagagca attttgatct ggccgaggac gctaaactgc 10740
agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcggagatc 10800
agtacgcaga cctgttcctg gccgctaaga atctgtctga cgctatcctg ctgagtata 10860
ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagagat 10920
atgacgagca ccatcaggat ctgaccctgc tgaaggcact ggtgcgccag cagctgccc 10980
agaagtacaa ggaaatcttc ttgatcaga gtaagaacgg gtacgccggt tatattgacg 11040
gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
gcaccgagga actgctgggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat 11160
ttgataacgg ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac 11220
gacaggaaga cttctacca tttctgaagg ataaccgca gaagatcgaa aaaattctga 11280
ccttcaggat cccttactat gtggggcccc tggcaagggg taattccgc tttgcctgga 11340
tgacacggaa atctgaggaa acaatcactc cttggaactt cgaggaagtg gtcgataagg 11400
gagcttccgc acagtctttc atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg 11460
agaaagtgtc gcctaagcac agtctgtctg acgagtattt cacagtctat aacgaactga 11520
ctaagggtgaa atacgtcacc gaggggatga ggaagccgc cttcctgagc ggtgaacaga 11580
agaaagctat cgtggacctg ctgtttaaaa ccaatcgcaa ggtgacagtc aagcagctga 11640
aggaggacta cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag 11700
acagatttaa cgcttctctg ggaacctacc acgatctgt gaagatcatt aaggataaag 11760
acttcctgga caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac 11820
tgtttgagga tcgcgaaatg atcgaggaaac ggctgaaaac ttatgccc atgttcgatg 11880
acaagggtgat gaaacagctg aagcgaagaa ggtacaccgg ctggggacga ctgagcagaa 11940
agctgatcaa cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat 12000
cagatggctt cgctaacaga aattttatgc agctgattca cgatgacagc ctgaccttca 12060
aagaggatat ccagaaggca cagggtgtccg ggcagggtga ctctctgcac gagcatatcg 12120

```

caaacctggc cgggtccccc gccatcaaga aaggtattct gcagaccgtg aaggtggtcg 12180
atgagctggg gaaagtcatg ggcaggcata agccagaaaa catcgtgatt gagatggccc 12240
gcgaaaaatca gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg 12300
aggaaggcat taaggaactg ggatcccaga tcctgaaaga gcaccctgtg gaaaacactc 12360
agctgcagaa tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg 12420
accaggagct ggatattaac cgactgtctg attacgacgt ggatcatatc gtcccacagt 12480
cattcctgaa agatgacagc attgacaata aggtgctgac ccggagtgcg aaaaaccgag 12540
gaaagagtga taatgtccct tcagaggaag tggtaagaa aatgaagaac tactggagac 12600
agctgctgaa tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga 12660
gaggggggtct gtcagaactg gacaaagcag gcttcatcaa gcgacagctg gtggagacca 12720
gacagatcac aaagcacgtc gctcagattc tggatagcag gatgaacaca aagtacgatg 12780
agaatgacaa actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg 12840
acttcagaaa ggatttccag ttctacaaag tcaggagat caacaattat caccatgctc 12900
atgacgcata cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg 12960
agagcgaatt cgtgtacggt gactataagg tgtacgatgt cagaaaaatg atcgccaaga 13020
gtgagcagga aattggaaaa gccaccgcta agtatttctt ttactcaaac atcatgaatt 13080
tccttaagac tgagatcacc ctggcaaatg gggaaatccg aaagagacca ctgattgaga 13140
ctaacggcga gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcga 13200
aggtcctgtc catgcctcaa gtgaatattg tcaagaaaac agaggtgcag actggcggat 13260
tcagtaagga atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag 13320
actgggatcc caagaaatat gggggtttcg actccccaac agtggettac tctgtcctgg 13380
tggtcgcaaa ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg 13440
gtatcactat tatggagagg agtccttcg agaagaacct catcgatttt ctggaggcta 13500
aaggctataa ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg 13560
agctggaaaa cggaaggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg 13620
aactggccct gccttctaag tacgtgaact tcctgtatct ggctagccac tacgagaagc 13680
tgaaaggctc ccccaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt 13740
atctggacga gatcattgaa cagattagcg agttctccaa aagagtgatc ctggctgacg 13800
caaactctgga taaggtcctg agcgataca acaaacacag agataagcca atcagggagc 13860
aggccgaaaa tatcattcat ctgttcaact tgaccaacct gggagccct gcagccttca 13920
agtattttga cactaccatc gatcggaac gatacacatc cactaaggag gtgctggacg 13980
ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
tggggggcga cctcgaggga agcggagagg gcagaggaag tctgctaaca tgcggtgacg 14100
tcgaggagaa tcctggccca gcaccgggat ccatggtgag caaggcgag gagctgttca 14160
ccgggggtgg gccatcctg gtcgagctgg acggcgacgt aaacggccac aagttcagcg 14220
tgtccggcga gggcgagggc gatgccacct acggcaagct gacctgaag ttcattctga 14280
ccaccggcaa gctgcccgtg ccctggccca ccctcgtgac caccttacc tacggcgtgc 14340
agtgtctcgc ccgtacccc gaccacatga agcagcacga cttcttcaag tccgcatgc 14400
ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacct 14460

```

gcgcccaggt gaagttcgag ggcgacaccc tggatgaaccg catcgagctg aaggggcatcg 14520
 acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 14580
 aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaacttc aagacccgcc 14640
 acaacatcga ggacggcagc gtgcagctcg ccgaccacta ccagcagaac acccccacatcg 14700
 gcgacggccc cgtgctgctg cccgacaacc actacctgag caccagtcg gccctgagca 14760
 aagaccccaa cgagaagcgc gatcacatgg tctgctgga gttcgtgacc gccggccggga 14820
 tcaactctcg catggacgag ctgtacaagt aaacctaatc tagcagctcg ctgatcagcc 14880
 tcgactgtgc cttctagtgt ccagccatct gttgtttgcc cctccccctg gccttccttg 14940
 accctggaag gtgccactcc cactgtcctt tctaataaaa atgaggaaat tgcacgcat 15000
 tgtctgagta ggtgtcattc tattctgggg ggtgggggtgg ggcaggacag caagggggag 15060
 gattgggaag acaatagcag gcatgctggg gatgcgggtg gctctatggc ttctgaggcg 15120
 gaaagaacca gctggggctc gatcctctag ttggcgctc atggtccata tgaatatect 15180
 ccttagttcc tattccgcta gcctagaggg acagccccc cccaaagccc ccagggatgt 15240
 aattacgtcc ctccccgct aggggcagca gcgagccgc cggggctccg ctccggctccg 15300
 gcgctcccc cgcaccccc agccggcagc gtgcggggac agcccgggca cggggaaggt 15360
 ggcacgggat cgctttctc tgaacgctt tcgctgctct ttgagcctgc agacacctgg 15420
 ggggatacgg ggaaaaagct ttaggctgaa agagagattt agaatgacag aatcatagaa 15480
 cggcctgggt tgcaaaggag cacagtgtc atccagatcc aaccctctgc tatgtgcagg 15540
 gtcacacacc agcagcccag gctgcccaga gccacatcca gcctggcctt gaatgcctgc 15600
 agggatgggg catccacagc ctcttgggc aacctgttca gtgcgtcacc acctctggg 15660
 ggaaaaactg cctctcata tccaaccaa acctccctg tctcagtgt aagccattcc 15720
 ccttgtctct atcaaggggg agtttgtgt gacattgttg gtctggggtg acacatgttt 15780
 gccaatcag tgcacacgg agaggcagat cttggggata aggaagtgc ggacagcatg 15840
 gacgtgggac atgcagggtg tgagggtctt gggacactct ccaagtcaca gcgttcagaa 15900
 cagccttaag gataagaaga taggatagaa ggacaaagag caagttaaaa ccagcatgg 15960
 agaggagcac aaaaaggcca cagacactgc tggtcctgt gtctgagcct gcatgtttga 16020
 tgggtgtctg atgcaagcag aaggggtgga agagcttgcc tggagagata cagctgggtc 16080
 agtaggactg ggacaggcag ctggagaatt gccatgtaga tgttcataca atcgtcaaat 16140
 catgaaggct ggaaaaagccc tccaagatcc ccaagaccaa ccccaacca cccaccgtgc 16200
 ccaactggcca tgtccctcag tgccacatcc ccacagttct tcatcacctc caggacgggt 16260
 gaccccccca cctccgtggg cagctgtgcc actgcagcac cgctcttttg agaaggtaaa 16320
 tcttgctaaa tccagcccga cctccctctg gcacaacgta aggccattat ctctcatcca 16380
 actccaggac ggagtcagt aggatggggc tggatccgaa gcagctccag cctacacaat 16440
 cgctcaagac gtgtaatgct tttattatat attagtcacg atatctataa caagaaaata 16500
 tatatataat aagtatcac gtaagtagaa catgaaataa caatataatt atcgtatgag 16560
 ttaaatctta aaagtcagct aaaagataat catgcgtcat ttgactcac gcggtcgta 16620
 tagttcaaaa tcagtacac ttaccgcatt gacaagcac cctcacggga gctccaagcg 16680
 gcgactgaga tgcctaaat gcacagcgac ggattcgcgc tatttagaaa gagagagcaa 16740
 tatttcaaga atgcatgcgt caattttacg cagactatct ttctagggtt aaaaaagatt 16800

tgcgctttac tcgacctaaa ctttaaacac gtcataagaat cttcgtttga caaaaaccac 16860
 attgtgggggt accgagctct taattaaggc gcgccgggga ggttcccttt agtgagggtt 16920
 aattgcgggt cgccctatag tgagtcgtat tacaattcac tggccgtcgt ttacaacgt 16980
 cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tcccccttc 17040
 gccagctggc gtaatagcga agaggccgc accgatcgcc cttcccaaca gttgcgcagc 17100
 ctgaatggcg aatggcaa at tgtaagcgtt aatatittgt taaaattcgc gttaaatttt 17160
 tgttaaatca gctcatTTTT taaccaatag gccgaaatcg gcaaaatccc ttataaatca 17220
 aaagaataga ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 17280
 aagaacgtgg actccaacgt caaaggcgca aaaaccgtct atcaggcgca tggccacta 17340
 cgtgaaccat caccctaate aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 17400
 aaccctaaag ggagccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 17460
 aaggaaggga agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg 17520
 ctgcgcgtaa ccaccacacc cgccgcgctt aatgcgcgcg tacaggcggc gtcag 17575

<210> 9

<211> 5692

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 9

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgctc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgca cgctgtaggg cctatttccc atgattcctt 180
 catatttgca tatacgatac aaggctgtta gagagataat tggaattaat ttgactgtaa 240
 acacaaagat attagtacaa aatacgtgac gtagaaagta ataatttctt gggtagtttg 300
 cagttttaaa attatgtttt aaaatggact atcatatgct taccgtaact tgaaagtatt 360
 tcgatttctt ggctttatat atcttgtgga aaggacgaaa caccgcagcg ttacctctat 420
 cgtagtttta gagctagaaa tagcaagtta aaataagget agtccgttat caacttgaaa 480
 aagtggcacc gagtcggtgc ttttttggat ccgagggcct atttcccatg attccttcat 540
 atttgcatat acgatacaag gctgttagag agataatttg aattaatttg actgtaaaca 600
 caaagatatt agtacaaaat acgtgacgta gaaagtaata atttcttggg tagtttgcag 660
 ttttaaaatt atgttttaaa atggactatc atatgcttac cgtaacttga aagtatttcg 720
 atttcttggc tttatatatc ttgtggaaaag gacgaaacac cgtgtaatag ctctgcatg 780
 ggttttagag ctagaaatag caagttaaaa taaggetagt ccgttatcaa cttgaaaaag 840
 tggcaccgag tcggtgcttt ttttctagaa ggtaccaggt cttgaaagga gtgggaattg 900
 gctccggtgc ccgtcagtgg gcagagcgca catcggccac agtccccgag aagttggggg 960
 gaggggtcgg caattgaacc ggtgcctaga gaaggtggcg cggggtaa ac tgggaaagtg 1020
 atgtcgtgta ctggctccgc ctttttcccc aggggtggggg agaaccgtat ataagtgcag 1080
 tagtcgccgt gaacgttctt tttcgcaacg ggtttgccgc cagaacacag gcgtacggcc 1140

```

accatgactt cgaaagttaa tgatccagaa caaaggaaac ggatgataac tgggtccgcag 1200
tgggtgggcca gatgtaaaca aatgaatgtt cttgattcat ttattaatta ttatgattca 1260
gaaaaacatg cagaaaatgc tgttatTTTT ttacatggta acgcggcctc ttcttattta 1320
tggcgacatg ttgtgccaca tattgagcca gtagcgcggt gtattatacc agaccttatt 1380
ggtatgggca aatcaggcaa atctggtaat ggttcttata ggttacttga tcattacaaa 1440
tatcttactg catggtttga acttcttaat ttaccaaaga agatcatttt tgtcggccat 1500
gattgggggtg cttgtttggc atttcattat agctatgagc atcaagataa gatcaaagca 1560
atagttcacg ctgaaagtgt agtagatgtg attgaatcat gggatgaatg gcctgatatt 1620
gaagaagata ttgcgttgat caaatctgaa gaaggagaaa aaatggTTTT ggagaataac 1680
ttcttcgtgg aaacctgtt gccatcaaaa atcatgagaa agttagaacc agaagaattt 1740
gcagcatatc ttgaaccatt caaagagaaa ggtgaagtcc gtcgtccaac attatcatgg 1800
cctcgtgaaa tcccgttagt aaaaggtggg aaacctgacg ttgtacaaat tgtaggaat 1860
tataatgctt atctacgtgc aagtgatgat ttaccaaaaa tgtttattga atcggacca 1920
ggattctttt ccaatgctat tgttgaaggt gccagaagt ttcctaatac tgaatttgtc 1980
aaagtaaaaag gtcttcattt ttcgcaagaa gatgcacctg atgaaatggg aaaatatatc 2040
aaatcgttcg ttgagcgagt tctcaaaaat gaacaataaa gcgctaataa aagatcttta 2100
ttttcattag atctgtgtgt tggttttttg tgtaagcttt ggctccaaca cagatgttct 2160
taggtacct aacttctaac ttttaatatc cagtcaacaa agaataccgc aagggtaggt 2220
gttgggatag ctgtcgacaa gctcatgcgg gtgtgtccac agggatatgc gtactatgca 2280
gaatatTTgt actgagtga gtcagtatac attcctttga gagccattag ctgctacaaa 2340
acagtaatct ggctgtttag atcaacaagc taaatgatag aagatgaaag tactggtttc 2400
catgtatttt tattaagtgt tgatgagaaa gttgtaagt acttacaggt tactctgtac 2460
atctgtagtc actgaattcg gaatatctta gagttttaca cacaagggtg agtgttaaaa 2520
tattgataaa gtttttgata atcttgtgtg agacatgttc taatttagtt gtattttatt 2580
atttttattg taaggcctgc tgaaaatgac tgagtataaa cttgtggctg tggcgccga 2640
cggcgtgggc aagagcgctt tgacgataca gctaattcag aatcactttg tggatgagta 2700
tgatccaacc atcgaggtaa cgctgctcta cagtctgcgt gcgcttgtaa aggacggcag 2760
ccagccgctt tgaaaaagat atcattttta tatttattag aaaattatat tgaaagttat 2820
ttcagttata tgtgatgtcc tttagtcca aggttttaaa ctgggtgtta gggaaccata 2880
ggtgcaagaa agtccacttc tcatgagagc tcaccacaga gaaagaaagt ccacttctca 2940
ggtaaccacg tgcggaccga gcggccgcag gaaccctag tgatggagtt ggccactccc 3000
tctctgcgcg ctgcctcgt cactgaggcc gggcgaccaa aggtcgcccg acgccgggc 3060
tttggccggg cggcctcagt gagcgagcga gcgcgcagct gcctgcaggt attttctct 3120
tacgcatctg tgcggtattt cacaccgat acgtcaaage aaccatagta cgcgccctgt 3180
agcggcgcat taagcgcggc ggggtgtgtg gttacgcga gcgtgaccgc tacacttgcc 3240
agcgccctag cgcccgctcc tttcgtttc ttcccttctt ttctcgccac gttecgccgc 3300
tttccccgtc aagctctaaa tcgggggctc ctttaggggt tccgatttag tgctttacgg 3360
cacctcgacc ccaaaaaact tgatttgggt gatggttcac gtagtgggcc atcgccctga 3420
tagacggttt ttcgccctt gacgttgag tccacgttct ttaatagtgg actcttgctc 3480

```



```

caaactggaa caacactcaa ccctatctcg ggctattctt ttgatttata agggattttg 3540
ccgatttcgg cctattgggt aaaaaatgag ctgatttaac aaaaatttaa cggaattttt 3600
aacaaaatat taacgtttac aattttatgg tgcactctca gtacaatctg ctctgatgcc 3660
gcatagttaa gccagccccg acacccgcca acacccgctg acgcgccctg acgggcttgt 3720
ctgctccccg catccgctta cagacaagct gtgaccgtct ccgggagctg catgtgtcag 3780
aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg gcctcgtgat acgcctattt 3840
ttataggtta atgtcatgat aataatgggt tcttagacgt caggtggcac ttttcgggga 3900
aatgtgcgcg gaacccttat ttgtttattt ttctaaatac attcaaatat gtatccgctc 3960
atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt 4020
caacatttcc gtgtcgccct tattcccttt tttgcggcat tttgccttcc tgtttttgct 4080
caccagaaaa cgctggtgaa agtaaaaagat gctgaagatc agttgggtgc acgagtgggt 4140
tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt 4200
tttccaatga tgagcacttt taaagttctg ctatgtggcg cggattatc ccgtattgac 4260
gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt ggttgagtac 4320
tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgtc 4380
gccataacca tgagtataa cactgcggcc aacttacttc tgacaacgat cggaggaccg 4440
aaggagctaa ccgctttttt gcacaacatg ggggatcatg taactcgctt tgatcgttgg 4500
gaaccggagc tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca 4560
atggcaacaa cgttgcgcaa actattaact gggaactac ttactctagc ttcccgcaaa 4620
caattaatag actggatgga ggcgataaaa gttgcaggac cacttctgcg ctcgccctt 4680
ccggctggct gggttatgtc tgataaatct ggagccgggt agcgtgggtc tcgcggtatc 4740
attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta cagcagggg 4800
agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc ctactgatt 4860
aagcattggg aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt 4920
catttttaat ttaaaaggat ctaggtgaag atcctttttg ataatctcat gaccaaatac 4980
ccttaacgtg agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct 5040
tcttgagatc ctttttttct gcgcgtaate tgctgcttgc aaacaaaaaa accaccgcta 5100
ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc 5160
ttcagcagag cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac 5220
ttcaagaact ctgtagcacc gcctacatac ctgctctgc taatcctgtt accagtggct 5280
gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat 5340
aaggcgcagc ggtcgggctg aacgggggggt tcgtgcacac agcccagctt ggagcgaacg 5400
acctacaccg aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccga 5460
gggagaaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg 5520
gagcttccag ggggaaacgc ctggtatctt tatagtctg tcgggtttcg ccacctctga 5580
cttgagcgtc gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc 5640
aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat gt 5692

```

<210> 10

<211> 19

<212> DNA

<213> 盲鼯形鼠 (*Nannospalax galili*)

<400> 10

ctcactatgt agaccaggc 19

<210> 11

<211> 57

<212> DNA

<213> 智人 (*Homo sapiens*)

<400> 11

cctgtaatcc cagcactttc actttgggag gccgaggcga gtctcgtctt gtcgccc 57

<210> 12

<211> 13890

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 12

```

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
gccttctctgt ttttgtctac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 360
tattatccccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttggg ttagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc ttagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcett tttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320

```

cgtagttagg ccaccacttc aagaactctg tagcacccgc tacatacctc gctctgctaa 1380
 tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
 gacgatatgt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500
 ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
 gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
 caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 1740
 tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg 1800
 ctcacatgtt ctttctcgcg ttatcccttg attctgtgga taaccgtatt accgcctttg 1860
 agtgagctga taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
 aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgggcg attcattaat 1980
 gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
 tgagttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt 2100
 tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
 ccaagcgcgc aattaaccct cactaaaggg aacctccctt agcttaatta accctagaaa 2220
 gataatcata ttgtgacgta cgtaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
 atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
 tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
 aaaaaaaca aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
 ttttacaaaa ataaacttat ttgtacttt aaaaacagtc atgttgtatt ataaaataag 2520
 taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
 gactagtcgt attaagggtt ccggatcagc ttgattcgag cccagctgg tttttccgc 2640
 ctcagaagcc atagagccca ccgcatcccc agcatgcctg ctattgtctt cccaatcctc 2700
 ccccttgctg tcctgcccc aaccaccccc cagaatagaa tgacacctac tcagacaatg 2760
 cgatgcaatt tcctcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
 ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcgaggc 2880
 tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
 ggcatgcgc tgcaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
 ttcgccgct caggcaccgg gcttgccggg catgcaccag gtgcgcggtc cttcgggcac 3060
 ctcgacgtcg gcggtgacgg tgaagccgag ccgctcgtag aaggggaggt tgcggggcgc 3120
 ggaggtctcc aggaaggcgg gcaccccggc gcgctcggcc gcctccactc cggggagcac 3180
 gacggcgtg cccagaccct tgccctggtg gtcgggcgag acgccgacgg tggccaggaa 3240
 ccacgcgggc tccttgggcc ggtgcggcgc caggaggcct tccatctgtt gctgcgcggc 3300
 cagccgggaa ccgtcaact cgcccatgcg cgggccgate tcggcgaaca ccgccccgc 3360
 ttcgacgtc tccggcgtgg tccagaccgc caccgcggcg ccgtcgtccg cgaccacac 3420
 cttgccgatg tcgagccga gcgcgtgag gaagagttct tgcagctcgg tgaccgctc 3480
 gatgtggcgg tccgggtcga cgggtgtggc cgtggcgggg tagtcggcga acgcggcggc 3540
 gagggtcgt acggcccggg ggacgtcgtc gcgggtggcg aggcgcaccg tgggcttgta 3600
 ctcggtcatg gtttagttcc tcacctgtc gtattatact atgccgatat actatgccga 3660

```

tgattaattg tcaacacgtg ctgctgcagg tcgaaaggcc cggagatgag gaagaggaga 3720
acagcgcggc agacgtgcgc ttttgaagcg tgcagaatgc cgggcctccg gaggaccttc 3780
gggcgccccg cccgccccctg agcccccccc tgagccccgc cccggacca ccccttccca 3840
gcctctgagc ccagaaagcg aaggagcaaa gctgctattg gccgctgccc caaaggccta 3900
cccgttcca ttgctcagcg gtgctgtcca tctgcacgag actagtga cgtgctactt 3960
ccatttgtca cgtcctgcac gacgcgagct gcggggcggg ggggaacttc ctgactaggg 4020
gaggagtaga aggtggcgcg aagggggccac caaagaacgg agccggttgg cgcctaccgg 4080
tgatgtgga atgtgtgcga gccagaggcc acttgtgtag cgccaagtgc ccagcggggc 4140
tgctaaagcg catgctccag actgccttgg gaaaagcgcc tccccatccc ggtagacacc 4200
ccacagtggg tggcctaggg acaggattgc aactccagtc tttcttcttc ttgggcggga 4260
gtcactagtt attaatagta atcaattacg gggtcattag ttcatagccc atatatggag 4320
ttccgcgtta cataacttac ggtaaattgc ccgcttggt gaccgcccac cgacccccgc 4380
ccattgacgt caataatgac gtatgttccc atagtaacgc caataggac tttccattga 4440
cgtcaatggg tggactatth acggtaaaact gcccacttgg cagtacatca agtgtatcat 4500
atgccaagta cgccccctat tgacgtcaat gacggtaaat ggcccgctg gcattatgcc 4560
cagtacatga ccttatggga ctttctact tggcagtaca tctacgtatt agtcacgct 4620
attaccatgg gtcgaggtga gcccacggt ctgcttact ctcctcatct cccccctc 4680
cccaccccca attttgtatt tatttatttt ttaattattt tgtgcagcga tggggcggg 4740
gggggggggg gcgcgcgcca ggccggggcg ggccggggcg gggcggggc ggggcgaggc 4800
ggagaggtgc ggccgcagcc aatcagagcg gcgcgctccg aaagtctct tttatggcga 4860
ggccggcgcg gcggcgcccc tataaaaagc gaagcgcgcg gcggcgggga gtcgctgcgt 4920
tgcttccgcc ccgtgccccg ctccgcgccc cctcgcgccc cccgccccgg ctctgactga 4980
ccgcgttact cccacaggtg agcggggcggg acggcccttc tctccgggc tgtaattagc 5040
gcttgggtta atgacggctc gtttctttt tgtggctgcg tgaaagcctt aaagggtcc 5100
gggagggccc tttgtgcggg ggggagcggc tcgggggggt cgtgcgtgtg tgtgtgcgtg 5160
gggagcgccg cgtgcggccc gcgctgcccg gcggctgtga gcgctgcggg cgcgcgcgcg 5220
ggctttgtgc gctccgcgtg tgcgcgaggg gagcgcgccc gggggcggtg ccccgcggtg 5280
cggggggggt gcgaggggaa caaaggctgc gtgcgggggt tgtgcgtggg ggggtgagca 5340
gggggtgtgg gcgcggcggt cgggtgttaa cccccctg cccccctc cccgagttgc 5400
tgagcacggc ccggcttcgg gtgcgggggt ccgtgcgggg cgtggcgcgg ggctcgccgt 5460
gccgggcggg ggggtggcggc aggtgggggt gccgggcggg gcggggccgc ctcgggccgg 5520
ggagggtcgc ggggaggggc gcggcgcccc cggagcgccc gcggctgtcg aggcgcggcg 5580
agccgcagcc attgcctttt atggtaatcg tgcgagaggg cgcagggact tctttgtcc 5640
caaactctgg ggagccgaaa tctgggaggc gccgcgcac cccctctagc gggcgcgggc 5700
gaagcgggtc ggcccgggca ggaaggaaat gggcggggag ggcttctgtg cgtcgcccg 5760
ccgccgtccc cttctccatc tccagctcgc gggtgcgcgc agggggacgg ctgccttcgg 5820
gggggacggg gcagggcggg gttcggttc tggcgtgtga ccggcggtc tagagcctct 5880
gctaaccatg ttcattgcct cttctttttc ctacagctcc tgggcaacgt gctggttatt 5940
gtgctgtctc atcatthtgg caaagaattc gccaccatgg tgcccaagaa gaagaggaaa 6000

```

gtctctagac tggacaagag caaagtcata aactctgctc tggaattact caatggagtc 6060
 ggtatcgaag gcctgacgac aaggaaactc gctcaaaagc tgggagttga gcagcctacc 6120
 ctgtactggc acgtgaagaa caagcgggcc ctgctcgatg ccctgccaat cgagatgctg 6180
 gacaggcatc ataccactc ctgccccctg gaaggcgagt catggcaaga ctttctgcgg 6240
 aacaacgcca agtcataccg ctgtgctctc ctctcacatc gcgacggggc taaagtgcac 6300
 ctcggcaccc gcccaacaga gaaacagtac gaaaccctgg aaaatcagct cgcgttcctg 6360
 tgtcagcaag gcttctccct ggagaacgca ctgtacgctc tgtccgccgt gggccacttt 6420
 acactgggct gcgtattgga ggaacaggag catcaagtag caaaagagga aagagagaca 6480
 cctaccaccg attctatgcc ccactttctg aaacaagcaa ttgagctgtt cgaccggcag 6540
 ggagccgaac ctgccttcct tttcggcctg gaactaatca tatgtggcct ggagaaacag 6600
 ctaaagtgcg aaagcggcgg gccgaccgac gcccttgacg attttgactt agacatgctc 6660
 ccagccgatg cccttgacga ctttgacctt gatatgctgc ctgctgacgc tcttgacgat 6720
 tttgacctg acatgctccc cgggtaaagc ggccgcgact ctagatcata atcagccata 6780
 ccacatttgt agaggtttta cttgctttta aaaacctccc acacctccc ctgaacctga 6840
 aacataaaat gaatgcaatt gttgttgta acttgtttat tgcagcttat aatggttaca 6900
 aataaagcaa tagcatcaca aatttcacaa ataaagcatt tttttcactg cattctagtt 6960
 gtggtttgtc caaactcatc aatgtatctt aagggatccc tagagggaca gcccccccc 7020
 aaagccccc gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 7080
 ggctccgctc cggctccggc ctccccccgc atccccgagc cggcagcgtg cggggacagc 7140
 ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgcttctcg ctgctctttg 7200
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 7260
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgctcatc cagatccaac 7320
 cccctgctat gtgcagggtc atcaaccagc agcccagget gccagagcc acatccagcc 7380
 tggccttgaa tgccctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagt 7440
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 7500
 cagtgtaaag ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 7560
 tggggtgaca catgtttgcc aattcagtgc atcacggaga ggcagatctt ggggataagg 7620
 aagtgcagga cagcatggac gtgggacatg caggtgttga gggtcttggg acactctcca 7680
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 7740
 gttaaaacc agcatggaga ggagcacaaa aaggccacag aactgctgg tccctgtgtc 7800
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctg 7860
 agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 7920
 tcatacaatc gtcaaatcat gaaggetgga aaagccctcc aagatcccca agaccaaccc 7980
 caaccaccc accgtgccca ctggccatgt ccctcagtgc cacatcccca cagttcttca 8040
 tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcacgc 8100
 tctttggaga aggtaaatct tgctaaatcc agcccagccc tcccctggca caacgtaagg 8160
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atgggctgt cgacctagag 8220
 ggacagcccc ccccaaaagc cccagggat gtaattacgt ccctccccg ctaggggcag 8280
 cagcgagccg cccggggctc cgctccggtc cggcgtccc cccgcattccc cgagccggca 8340

gcgtgcgggg acagccccggg cacggggaag gtggcacggg atcgctttcc tctgaacgct 8400
 tctcgctgct ctttgagcct gcagacacct ggggggatac ggggaaaaag ctttaggctg 8460
 aaagagagat ttagaatgac agaatcatag aacggcctgg gttgcaaagg agcacagtgc 8520
 tcatccagat ccaacccccct gctatgtgca gggatcatca ccagcagccc aggctgcccc 8580
 gagccacatc cagcctggcc ttgaatgcct gcagggatgg ggcatccaca gcctccttgg 8640
 gcaacctgtt cagtgcgtca ccaccctctg ggggaaaaac tgcctcctca tatccaaccc 8700
 aaacctcccc tgtctcagtg taaagccatt ccccttctgc ctatcaaggg ggagtttgc 8760
 gtgacattgt tggctctggg tgacacatgt ttgccaatc agtgcacac ggagaggcag 8820
 atcttggggg taagggaagt caggacagca tggacgtggg acatgcaggt gttgagggct 8880
 ctgggacact ctccaagtca cagcgttcag aacagcctta aggataagaa gataggatag 8940
 aaggacaaaag agcaagttaa aaccagcat ggagaggagc acaaaaaggc cacagacact 9000
 gctgggccct gtgtctgagc ctgcatgttt gatggtgtct ggatgcaagc agaaggggtg 9060
 gaagagcttg cctggagaga tacagctggg tcagtaggac tgggacaggc agctggagaa 9120
 ttgccatgta gatgttcata caatcgtcaa atcatgaagg ctggaaaagc cctccaagat 9180
 cccaagacc aacccaacc caccaccgt gccactggc catgtccctc agtgccacat 9240
 cccacagtt ctcatcacc tccagggacg gtgaccccc cacctccgtg ggcagctgtg 9300
 ccaactgcagc accgctcttt ggagaaggta aatcttgcta aatccagccc gaccctcccc 9360
 tggcacaacg taaggccatt atctctcatc caactccagg acggagtcag tgaggatggg 9420
 gctcaattgt ttactcccta tcagtgatag agaacgtatg aagagtttac tccctatcag 9480
 tgatagagaa cgtatgcaga ctttactccc tatcagtgat agagaacgta taaggagttt 9540
 actccctatc agtgatagag aacgtatgac cagtttactc cctatcagtg atagagaacg 9600
 tatctacagt ttactcccta tcagtgatag agaacgtata tccagtttac tccctatcag 9660
 tgatagagaa cgtataagct ttaggcgtgt acgggtggcg cctataaaag cagagctcgt 9720
 ttagtgaacc gtcagatcgc ctggagcaat tccacaacac ttttgtctta taccaacttt 9780
 ccgtaccact tcctaccctc gtaaaaagct tgtccaccat gagattcaaa agccacactg 9840
 tggaattgag gaggccctgc agcgacatgg agggagctgc tttgctgaga gtctctgtcc 9900
 tctgcatctg gatgagtgc cttttccttg gtgtgggagt gagggcagag gaagctggag 9960
 cgagggtgca acaaaacgtt ccaagtggga cagatactgg agatcctcaa agtaagcccc 10020
 tcggtgactg ggctgctggc accatggacc cagagagcag tatctttatt gaggatgcca 10080
 ttaagtattt caaggaaaaa gtgagcacac agaactctgt actcctgctg actgataatg 10140
 aggcttgga cggattcgtg gctgctgctg aactgcccag gaatgaggca gatgagctcc 10200
 gtaaaactct ggacaacctt gcaagacaaa tgatcatgaa agacaaaaac tggcacgata 10260
 aaggccagca gtacagaaac tggtttctga aagagtttcc tcggttgaaa agtgagcttg 10320
 aggataacat aagaaggctc cgtgcccttg cagatggggt tcagaaggte cacaaggca 10380
 ccaccatcgc caatgtggtg tctggetctc tcagcatttc ctctggcatc ctgaccctcg 10440
 tcggcatggg tctggcacc ttacagagg gaggcagcct tgtactcttg gaacctggga 10500
 tggagttggg aatcacagcc gctttgaccg ggattaccag cagtaccatg gactacggaa 10560
 agaagtgggtg gacacaagcc caagcccacg acctggtcat caaaagcctt gacaaattga 10620
 aggaggtgag ggagtttttg ggtgagaaca tatccaactt tctttcctta gctggcaata 10680

cttaccaact cacacgaggc attgggaagg acatccgtgc cctcagacga gccagagcca 10740
 atcttcagtc agtaccgcat gcctcagcct cagccccacg agtcactgag ccaatctcag 10800
 ctgaaagcgg tgaacagggt gagagggtta atgaaccag catcctggaa atgagcagag 10860
 gagtcaagct cacggatgtg gcccctgtaa gcttctttct tgtgctggat gtagtctacc 10920
 tcgtgtacga atcaaagcac ttacatgagg gggcaaagtc agagacagct gaggagctga 10980
 agaagggtggc tcaggagctg gaggagaagc taaacattct caacaataat tataagattc 11040
 tgcaggcgga ccaagaactg tgaaattcta aaatacagca tagcaaaact ttaacctcca 11100
 aatcaagcct ctacttgaat ctttttctga gggatgaata aggcataggc atcaggggct 11160
 gttgccaatg tgcattagct gtttgcagcc tcaccttctt tcatggagtt taagatatag 11220
 tgtattttcc caaggtttga actagctctt catttcttta tgttttaaat gcactgacct 11280
 cccacattcc ctttttagta aaatattcag aaataattta aatacatcat tgcaatgaaa 11340
 ataaatgttt tttattagga agaatccaga tgetcaagge cttcataat atccccagt 11400
 ttagtagttg gacttaggga acaaaggaa ctttaataga aattggacag caagaaagcg 11460
 agcttctagc tcgagatggt ccatatgaat atcctcetta gtctctatc cgctagccta 11520
 gagggacagc ccccccccaa agccccaggg gatgtaatta cgctccctcc ccgctagggg 11580
 cagcagcgag ccgcccgggg ctccgctccg gtccggcgct cccccgcct ccccgagccg 11640
 gcagcgtgcg gggacagccc gggcacgggg aaggtggcac gggatcgctt tcctctgaac 11700
 gcttctcgct gctctttgag cctgcagaca cctgggggga tacggggaaa aagctttagg 11760
 ctgaaagaga gatttagaat gacagaatca tagaacggcc tgggttgcaa aggagcacag 11820
 tgetcatcca gatccaacc cctgctatgt gcagggtcat caaccagcag cccaggtgc 11880
 ccagagccac atccagcctg gccttgaatg cctgcaggga tggggcatcc acagcctcct 11940
 tgggcaacct gttcagtgcg tcaccaccct ctgggggaaa aactgcctcc tcatatccaa 12000
 cccaaacctc ccctgtctca gtgtaaagcc attccccctt gtcctatcaa gggggagt 12060
 gctgtgacat tgttggcttg gggtagacac tgtttgcaa ttcagtgcac cacggagagg 12120
 cagatcttgg ggataaggaa gtgcaggaca gcatggacgt gggacatgca ggtgttgagg 12180
 gctctgggac actctccaag tcacagcgtt cagaacagcc ttaaggataa gaagatagga 12240
 tagaaggaca aagagcaagt taaaaccag catggagagg agcacaaaaa ggccacagac 12300
 actgctggtc cctgtgtctg agcctgcatg tttgatgggt tctggatgca agcagaagg 12360
 gtggaagagc ttgcctggag agatacagct gggtcagtag gactgggaca ggcagctgga 12420
 gaattgccat gtagatgttc atacaatcgt caaatcatga aggttgaaa agccctccaa 12480
 gatccccaag accaaccaca acccaccac cgtgcccact ggccatgtcc ctcagtgcc 12540
 catccccaca gttcttcate acctccaggg acggtgacct cccacctcc gtgggcagct 12600
 gtgccactgc agcaccgctc tttggagaag gtaaattctt ctaaattccag cccgacctc 12660
 ccctggcaca acgtaaggcc attatctctc atccaactcc aggacggagt cagtgaggat 12720
 ggggctggat ccgaagcagc tccagcctac acaatcgctc aagacgtgta atgcttttat 12780
 tatatatag tcacgatatc tataacaaga aaatatatat ataataagtt atcacgtaag 12840
 tagaacatga aataacaata taattatcgt atgagttaaa tcttaaaagt cacgtaaaag 12900
 ataactatgc gtcattttga ctacgcgggt cgttatagtt caaaatcagt gacacttacc 12960
 gcattgacaa gcacgcctca cgggagctcc aagcggcgac tgagatgtcc taaatgcaca 13020

gcgacggatt cgcgctatTT agaaagagag agcaatatTT caagaatgca tgcgtcaatt 13080
 ttacgcagac tatctttcta gggTTaaaaa agatttgcgc ttactcgcac ctaaacttta 13140
 aacacgtcat agaatcttcg ttTgacaaaa accacattgt ggggtaccga gctcttaatt 13200
 aaggcgcgcc ggggaggTtc ctttagtga gggTTaattg cgggtcgccc tatagtgagt 13260
 cgtattacaa ttactggcc gtcgttttac aacgtcgtga ctgggaaaac cctggcgTTa 13320
 cccaacttaa tcgccttgca gcacatcccc ctttcgccag ctggcgtaat agcgaagagg 13380
 cccgcaccga tcgcccTtcc caacagTtgc gcagcctgaa tggcgaatgg caaattgtaa 13440
 gcgttaatat tttgtTaaaa ttcgcgTTaa atttttgtta aatcagctca ttttttaacc 13500
 aataggccga aatcggcaaa atcccttata aatcaaaaga atagaccgag atagggtTga 13560
 gtgttgttcc agtttggaaC aagagtccac tattaaagaa cgtggactcc aacgtcaaag 13620
 ggcgaaaaaC cgtctatcag ggcgatggcc cactacgtga accatcccc taatcaagtt 13680
 ttttggggTc gaggtgccgt aaagcaCTaa atcggaacCC taaagggagc ccccgattta 13740
 gagcttgacg gggaaaagccg gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag 13800
 cgggcgctag ggcgctggca agtgtagcgg tcacgtcgcg cgtaaccacc acaccgccg 13860
 cgcttaatgc gccgctacag ggcgcgtcag 13890

<210> 13

<211> 17331

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 13

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatTTTT taaatacatt 60
 caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
 ggaagagtat gagtattcaa catttcctg tcgcccttat tccctTTTT gcggcatttt 180
 gccttcctgt ttttgctcac ccagaaacgc tggTgaaagt aaaagatgct gaagatcagt 240
 tgggtgcacg agtgggttac atcgaaactgg atctcaacag cggtaagatc cttgagagtt 300
 ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
 tattatccccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
 atgacttggT tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
 gagaattatg cagtgtctgc ataaccatga gtgataaac tgcggccaaC ttacttctga 540
 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
 ctcgccTtga tcgttgggaa ccggagetga atgaagccat accaaacgac gagcgtgaca 660
 ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 720
 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
 ttctgcgtc ggcccttcg gctggetggT ttattgtga taaatctgga gccggtgagc 840
 gtgggtctcg cggtatcatt gcagcaactgg ggccagatgg taagccctcc cgtatcgtag 900
 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020


```

agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agatacaaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa 1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaaggcg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggagg cttccagggg gaaacgectg gtatctttat agtctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggctttttgc tggccttttg 1800
ctcacatgtt ctttctgcg ttatccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctccctt agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaa aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat ttgtactttt aaaaacagtc atgttgtatt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcac aatatgatta tctttctagg gtttaattaag atatctgaag ttctatact 2640
ttctagagaa taggaacttc ggaataggaa cttcaaagca agctagagac cattaagggt 2700
tccgcatcag cttgattcga gccccagctg gttctttccg cctcagaagc catagagccc 2760
accgcatccc cagcatgcct gctattgtct tcccaatcct ccccttgcg gtccctgccc 2820
acccaccccc ccagaataga atgacaccta ctcagacaat gcgatgcaat ttctctattt 2880
tattaggaaa ggacagtggg agtggcacct tccagggtca aggaaggcac gggggagggg 2940
caaacaacag atggctggca actagaaggc acagtcgagg ctgatcagcg agctctagag 3000
aattgatccc ctcagaagaa ctcgtcaaga aggcgataga aggcgatcgc ctgcgaatcg 3060
ggagcggcga taccgtaaag cacgaggaag cggtcagccc attcgcgcc aagctettca 3120
gcaatatcac gggtagccaa cgctatgtcc tgatageggt ccgccacacc cagecggcca 3180
cagtcgatga atccagaaaa gcggccattt tccacatga tattcggcaa gcaggcatcg 3240
ccatgggtca cgacgagatc ctcgccgtcg ggcagtcgcg ccttgagcct ggcaaacagt 3300
tcggctggcg cgagcccctg atgctcttcg tccagatcat cctgatcgac aagaccggct 3360

```

```

tccatccgag tacgtgctcg ctcgatgcga tgtttcgctt ggtggtcgaa tgggcaggta 3420
gccgatcaa gcgtatgcag ccgccgcatt gcatcagcca tgatggatac tttctcggca 3480
ggagcaaggt gagatgacag gagatcctgc cccggcactt cgcccaatag cagccagtcc 3540
cttcccgcctt cagtgacaac gtcgagcaca gctgcgcaag gaacgcccgt cgtggccagc 3600
cacgatatgcc gcgctgcctc gtcctgcagt tcattcaggg caccggacag gtcggtcttg 3660
acaaaaagaa ccgggcgccc ctgcgctgac agccggaaca cggcggcatc agagcagccg 3720
attgtctgtt gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct 3780
gcgtgcaatc catcttgctt aatggccgat cccatggcgg tatcgataag ctagcttggg 3840
ctgcaggctc agggacctaa ttaagggttc cggatccact agttctagag cggcctcgac 3900
tctacgatac cgtcgatccc cactggaaaag accgcgaaga gtttgtcctc aaccgcgagc 3960
tgtggaaaaa aaaggacag gataagtatg acatcatcaa ggaaacctg gactactgcg 4020
ccctacagat ccctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact 4080
tcaaagatgc aactccagtc tttcttcttc ttgggcggga gtctactagt tattaatagt 4140
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 4200
cggtaaatgg cccgcctggc tgaccgccc aacgaccccg cccattgacg tcaataatga 4260
cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggactatt 4320
tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccagt acgccccta 4380
ttgacgtcaa tgacggtaaa tggcccgctt ggcatatgc ccagtacatg acctatggg 4440
actttctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ggtcgaggtg 4500
agccccacgt tctgcttcac tctcccatc tccccccct cccaccccc aattttgtat 4560
ttattttatt tttaattatt ttgtgcagcg atgggggcgg gggggggggg ggcgcgcgcc 4620
aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg cggcggcagc 4680
caatcagagc ggcgcgctcc gaaagtctc ttttatggcg aggcggcggc ggcggcggcc 4740
ctataaaaag cgaagcgcg gcggggcggg agtcgctgcg ttgccttcgc cccgtgcccc 4800
gctccgcgcc gcctcgcgcc gcccgcccc gctctgactg accgcgttac tcccacaggt 4860
gagcggggcg gacggccctt ctctccggg ctgtaattag cgcttggttt aatgacggct 4920
cgtttctttt ctgtggctgc gtgaaagcct taaagggtc cgggagggcc ctttgtgcgg 4980
gggggagcgg ctcggggggt gcgtgcgtgt gtgtgtgcgt ggggagcgcc gcgtgcggcc 5040
cgcgctgccc ggcggtctg agcgtcgcgg gcgcggcgcg gggctttgtg cgctccgcgt 5100
gtgcgcgagg ggagcgcgcc cggggcggtt gcccccggt gcgggggggc tgcgaggga 5160
acaaaggctg cgtgcggggt gtgtgcgtgg gggggtgagc agggggtgtg ggcgcggcg 5220
tcgggctgta accccccct gcacccccct ccccgagttg ctgagcacgg cccggcttcg 5280
ggtgcggggc tccgtgcggg gcgtggcgcg gggtcgcgg tgccgggcgg ggggtggcgg 5340
caggtggggg tgccgggcgg ggccggggcc cctcggggcg gggagggtc gggggagggg 5400
cgcggcggcc ccggagcgcc ggcggtctg gaggcgcggc gagccgcagc cattgccttt 5460
tatggtaatc gtgcgagagg gcgcaggga ttcctttgtc ccaaactctg cggagccgaa 5520
atctgggagg cgccgccgca cccctctag cgggcgcggg cgaagcgtg cggcgccggc 5580
aggaaggaaa tgggcgggga gggccttcgt gcgtcgccgc gccgccgtc cttctccat 5640
ctccagcctc ggggctgccg cagggggacg gctgccttcg ggggggacgg ggcaggcg 5700

```

gggtcggcctt ctggcgtgtg accggcggct ctagagcctc tgctaaccat gttcatgcct 5760
 tcttctttttt cctacagctc ctgggcaacg tgctggttat tgtgtgtct catcattttg 5820
 gcaaagaatt cgccaccatg gtgccaaga agaagaggaa agtctctaga ctggacaaga 5880
 gcaaagtcac aaactctgct ctggaattac tcaatggagt cggatcga ggcctgacga 5940
 caaggaaact cgctcaaaag ctgggagttg agcagcctac cctgtactgg cacgtgaaga 6000
 acaagcgggc cctgctcgat gccctgccaa tcgagatgct ggacaggcat catacccact 6060
 cctgccccctt ggaaggcgag tcatggcaag actttctgcg gaacaacgcc aagtcatacc 6120
 gctgtgtctt cctctcacat cgcgacgggg ctaaagtga tctcggcacc cgcccaacag 6180
 agaaacagta cgaaaccctg gaaaatcagc tcgcgttctt gtgtcagcaa ggcttctccc 6240
 tggagaacgc actgtacgct ctgtccgccc tgggccaact tacactgggc tgcgtattgg 6300
 aggaacagga gcatcaagta gcaaaagagg aaagagagac acctaccacc gattctatgc 6360
 ccccacttct gaaacaagca attgagctgt tcgaccggca gggagccga cctgccttcc 6420
 ttttcggcct ggaactaatc atatgtggcc tggagaaaca gctaaagtgc gaaagcggcg 6480
 ggccgaccga cgcccttgac gattttgact tagacatgct ccagccgat gcccttgacg 6540
 actttgacct tgatatgctg cctgtgacg ctcttgacga ttttgacctt gacatgctcc 6600
 ccgggtaaag cggccgcgac tctagatcat aatcagccat accacatttg tagaggtttt 6660
 acttgcttta aaaaacctcc cacacctccc cctgaacctg aaacataaaa tgaatgcaat 6720
 tgttggtgtt aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac 6780
 aaatttcaca aataaagcat ttttttact gcattctagt tgtggtttgt ccaaaactcat 6840
 caatgtatct taagggatcc ctagagggac agccccccc caaagcccc aggatgtaa 6900
 ttacgtccct ccccgctag gggcagcagc gagccgccc gggctccgct ccggctccgc 6960
 gctccccccg catccccgag ccggcagcgt gcggggacag cccgggcacg gggaaggtgg 7020
 cacgggatcg ctttctctg aacgttctc gctgctcttt gagcctgcag acacctgggg 7080
 ggatacgggg aaaaagcttt aggctgaaag agagatttag aatgacagaa tcatagaacg 7140
 gcctgggttg caaaggagca cagtgtcat ccagatccaa cccctgcta tgtgcagggt 7200
 catcaaccag cagcccaggc tgcccagagc cacatccagc ctggccttga atgcctgcag 7260
 ggatggggca tccacagcct ccttgggcaa cctgttcagt gcgtcaccac cctctggggg 7320
 aaaaactgcc tcctcatatc caacccaaac ctccccgtc tcagtgtaaa gccattcccc 7380
 cttgtcctat caagggggag ttgtctgtga cattgttggc ctggggtgac acatgtttgc 7440
 caattcagtg catcacggag aggcagatct tggggataag gaagtgcagg acagcatgga 7500
 cgtgggacat gcagggtgtg agggctctgg gacactctcc aagtcacagc gttcagaaca 7560
 gccttaagga taagaagata ggatagaagg acaaagagca agttaaacc cagcatggag 7620
 aggagcaca aaaggccaca gacactgtg gtccctgtgt ctgagcctgc atgtttgatg 7680
 gtgtctggat gcaagcagaa ggggtggaag agcttgctg gagagataca gctgggtcag 7740
 taggactggg acaggcagct ggagaattgc catgtagatg ttcatacaat cgtcaaatca 7800
 tgaaggctgg aaaaagcctc caagatcccc aagaccaacc ccaaccacc caccgtgccc 7860
 actggccatg tccctcagtg ccacatcccc acagttcttc atcacctcca gggacggtga 7920
 cccccccacc tccgtgggca gctgtgccac tgcagaccg ctctttggag aaggtaaatc 7980
 ttgctaaatc cagcccgacc ctccccctggc acaacgtaag gccattatct ctcatccaac 8040

```

tccaggacgg agtcagttag gatggggctg tcgacctaga gggacagccc ccccccaaag 8100
ccccaggga tgtaattacg tccctcccc gctaggggca gcagcgagcc gcccggggct 8160
ccgctccggg cgggcgctcc ccccgcatcc ccgagccggc agcgtgcggg gacagcccgg 8220
gcacggggaa ggtggcacgg gatcgctttc ctctgaacgc ttctcgctgc tctttgagcc 8280
tgcagacacc tgggggggata cggggaaaaa gctttaggct gaaagagaga tttagaatga 8340
cagaatcata gaacggcctg ggttgcaaag gagcacagt ctcattccaga tccaaccccc 8400
tgctatgtgc aggggtcatca accagcagcc caggtgccc agagccacat ccagcctggc 8460
cttgaatgcc tgcagggatg gggcatccac agcctccttg ggcaacctgt tcagtgcgtc 8520
accacctctt gggggaaaaa ctgcctctc atatecaacc caaacctccc ctgtctcagt 8580
gtaaagccat tcccccttgt cctatcaagg gggagtttgc tgtgacattg ttggtctggg 8640
gtgacacatg tttgccaatt cagtgcata cggagaggca gatcttgggg ataaggaagt 8700
gcaggacagc atggacgtgg gacatgcagg tgttgagggc tctgggacac tctccaagtc 8760
acagcgttca gaacagcctt aaggataaga agataggata gaaggacaaa gagcaagtta 8820
aaaccagca tggagaggag cacaaaaagg ccacagacac tgctggtccc tgtgtctgag 8880
cctgcatgtt tgatggtgtc tggatgcaag cagaaggggt ggaagagctt gcctggagag 8940
atacagctgg gtcagtagga ctgggacagg cagctggaga attgccatgt agatgttcat 9000
acaatcgtca aatcatgaag gctggaaaag ccctccaaga tccccagac caaccccaac 9060
ccaccaccg tgccactgg ccatgtccct cagtgcaca tccccacagt tcttcatcac 9120
ctccagggac ggtgaccccc ccacctcgt gggcagctgt gccactgcag caccgtcttt 9180
tggagaagggt aaatcttgct aaatccagcc cgacctccc ctggcacaac gtaaggccat 9240
tatctctcat ccaactccag gacggagtca gtgaggatgg ggctcaattg tttactccct 9300
atcagtgata gagaacgtat gaagagtta ctccctatca gtgatagaga acgtatgcag 9360
actttactcc ctatcagtga tagagaacgt ataaggagtt tactccctat cagtgataga 9420
gaacgtatga ccagtttact ccctatcagt gatagagaac gtatctacag tttactccct 9480
atcagtgata gagaacgtat atccagttta ctccctatca gtgatagaga acgtataagc 9540
tttaggcgtg tacgggtgggc gcctataaaa gcagagctcg tttagtgaac cgtcagatcg 9600
cctggagcaa ttccacaaca cttttgtctt ataccaactt tccgtaccac ttcctaccct 9660
cgtaaaaagc ttgtccacca tggtcctaa gaaaaagcgg aaggtggaca agaaatactc 9720
aatcgggctg gacatcgga ctaactcagt ggggtgggca gtcattactg acgagtacaa 9780
agtgccaaagc aagaaattta aggtcctggg caacaccgat aggcactcca tcaagaaaaa 9840
tctgattggg gccctgctgt tcgactctgg agagacagct gaagcaacta gactgaaaag 9900
gactgctaga aggcgtata cccggcgaaa gaatcgcate tgctacctgc aggagatttt 9960
ctctaacgaa atggccaagg tggacgatag tttctttcat cggctggagg aatcattcct 10020
ggctcaggaa gataagaaac acgagagaca tcctatcttt ggaaacattg tggacgaggt 10080
cgcttatcac gaaaaatacc ccaccatcta tcatctgcgc aagaaactgg tggactctac 10140
agataaagca gacctgcggc tgatctatct ggccctggct cacatgatta agttcagagg 10200
ccattttctg atcaggggag atctgaacct agacaatagc gatgtggaca agctgttcat 10260
ccagctggtc cagacataca atcagctgtt tgaggaaaac cctattaatg catctggcgt 10320
ggacgcaaaa gccatcctga gtgccaggct gtctaagagt agaaggctgg agaacctgat 10380

```

cgctcagctg ccaggcgaaa agaaaaacgg cctgtttgga aatctgattg cactgtcact 10440
 gggactgaca cctaacttca agagcaattt tgatctggcc gaggacgcta aactgcagct 10500
 gagcaaggac acttatgacg atgacctgga taacctgctg gctcagatcg gagatcagta 10560
 cgcagacctg ttcctggccg ctaagaatct gtctgacgct atcctgctga gtgatattct 10620
 gcgggtgaac accgagatta caaaagcccc tctgtcagct agcatgatca agagatatga 10680
 cgagcaccat caggatctga cctgctgaa ggcactgggtg cgccagcagc tgcccagaaa 10740
 gtacaaggaa atcttctttg atcagagtaa gaacgggtac gccggttata ttgacggcgg 10800
 agcttcacag gaggaattct acaagtttat caaacctatt ctggagaaga tggacggcac 10860
 cgaggaactg ctggtgaaac tgaatcgcga ggacctgctg cgcaagcagc ggacatttga 10920
 taacggctcc atccccacc agattcatct gggagagctg cacgcaatcc tgcgacgaca 10980
 ggaagacttc taccatttc tgaaggataa ccgcgagaag atcgaaaaaa ttctgacctt 11040
 ccggatccct tactatgtgg ggccctggc aaggggtaat tcccgtttg cctggatgac 11100
 acggaaatct gaggaacaa tcaactcttg gaacttcgag gaagtggctg ataagggagc 11160
 ttccgcacag tctttcatcg agagaatgac aaacttcgac aaaaacctgc caaatgagaa 11220
 agtgtgcct aagcacagtc tgctgtacga gtatttcaca gtctataacg aactgactaa 11280
 ggtgaaatac gtcaccgagg ggatgaggaa gcccgccttc ctgagcgggtg aacagaagaa 11340
 agctatcgtg gacctgctgt ttaaaaccaa tcgcaaggtg acagtcaagc agctgaagga 11400
 ggactacttc aagaaaattg aatgtttcga ttctgtggag atcagtggcg tcgaagacag 11460
 atttaacgct tctctgggaa cctaccacga tctgctgaag atcattaagg ataaagactt 11520
 cctggacaac gaggaaaatg aggatatcct ggaagacatt gtgctgacct tgacactgtt 11580
 tgaggatcgc gaaatgatcg aggaacggct gaaaacttat gcccatctgt tcgatgacaa 11640
 ggtgatgaaa cagctgaagc gaagaaggta caccggctgg ggacgactga gcagaaagct 11700
 gatcaacggc attcgggaca aacagagtgg aaagactatc ctggactttc tgaaatcaga 11760
 tggcttcgct aacagaaatt ttatgcagct gattcacgat gacagcctga cttcaaaga 11820
 ggatatccag aaggcacagg tgtccgggca gggtgactct ctgcacgagc atatcgcaa 11880
 cctggccggg tcccccgcca tcaagaaagg tattctgcag accgtgaagg tggtcgatga 11940
 gctggtgaaa gtcatgggca ggcataagcc agaaaacatc gtgattgaga tggcccgca 12000
 aaatcagacc acacagaaag gacagaagaa cagccgcgag cggatgaaaa ggatcgagga 12060
 aggcattaag gaactgggat cccagatcct gaaagagcac cctgtggaaa aactcagct 12120
 gcagaatgag aagctgtatc tgtactatct gcagaatggg cgggatatgt acgtggacca 12180
 ggagctggat attaaccgac tgtctgatta cgacgtggat catatcgctc cacagtcatt 12240
 cctgaaagat gacagcattg acaataaggt gctgacctgg agtgacaaaa accgaggaaa 12300
 gagtgataat gtcccttcag aggaagtgggt caagaaaatg aagaactact ggagacagct 12360
 gctgaatgcc aaactgatca cacagcgaac gtttgataac ctgactaaag ctgagagagg 12420
 gggctctgtca gaactggaca aagcaggctt catcaagcga cagctgggtg agaccagaca 12480
 gatcacaag cacgtcgtc agattctgga tagcaggatg aacacaaagt acgatgagaa 12540
 tgacaaactg atccgcgaag tgaaggatcat tactctgaag tcaaaacttg tgagcgactt 12600
 cagaaaggat ttccagttct acaaagtcag ggagatcaac aattatcacc atgctcatga 12660
 cgcataacctg aacgcagtgg tcgggaccgc cctgattaag aaatacccca aactggagag 12720

cgaattcgtg tacggtgact ataaggtgta cgatgtcaga aaaatgatcg ccaagagtga 12780
 gcaggaaatt ggaaaagcca ccgctaagta tttcttttac tcaaacaatca tgaatttctt 12840
 taagactgag atcaccttgg caaatgggga aatccgaaag agaccactga ttgagactaa 12900
 cggcgagacc ggagaaatcg tgtgggacaa gggtagggat tttgccacag tgcgcaaggt 12960
 cctgtccatg cctcaagtga atattgtcaa gaaaacagag gtgcagactg gcggattcag 13020
 taaggaatca attctgcccc aacggaactc tgataagctg atcgccccgaa agaaagactg 13080
 ggatcccaag aaatatgggg gtttcgactc cccaacagtg gcttactctg tcctgggtgt 13140
 cgcaaaggtg gagaagggga aaagcaagaa actgaaatcc gtcaaggagc tgctgggtat 13200
 cactattatg gagaggagct ccttcgagaa gaaccccatc gattttctgg aggctaaagg 13260
 ctataaggaa gtgaagaaaag acctgatcat taaactgcca aagtacagcc tgtttgagct 13320
 ggaaaacgga aggaagcgaa tgctggcatc cgcaggagag ctgcagaagg gtaatgaact 13380
 ggccctgcct tctaagtacg tgaacttctt gtatctggtt agccactacg agaagctgaa 13440
 aggtccccc gaggataacg aacagaaaaca gctgtttgtg gagcagcaca agcattatct 13500
 ggacgagatc attgaacaga ttagcgagtt ctccaaaaga gtgatcctgg ctgacgcaaa 13560
 tctggataag gtctgagcg catacaacaa acacagagat aagccaatca gggagcaggc 13620
 cgaaaatata attcatctgt tcaacttgac caacctggga gccctgcag ctttcaagta 13680
 ttttgacact accatcgatc ggaaacgata cacatccact aaggaggtgc tggacgtac 13740
 cctgattcac cagagcatta ccggcctgta tgaacaagg attgacctgt ctcagctggg 13800
 gggcgacctc gaggggaagcg gagagggcag aggaagtctg ctaacatgcg gtgacgtcga 13860
 ggagaatcct ggcccagcac cgggatccat ggtgagcaag ggcgaggagc tgttcaccgg 13920
 ggtggtgccc atcctggtcg agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc 13980
 cggcgagggc gagggcgatg ccacctacgg caagctgacc ctgaagtcca tctgcaccac 14040
 cggcaagctg cccgtgccct ggcccaccct cgtgaccacc ttcacctacg gcgtgcagtg 14100
 cttcgccccg taccgacac acatgaagca gcacgacttc ttcaagtccg ccatgcccga 14160
 aggctacgtc caggagcgca ccatcttctt caaggacgac ggcaactaca agaccgcg 14220
 cgaggtgaag ttcgagggcg acacctggt gaaccgcatc gagctgaagg gcatcgactt 14280
 caaggaggac ggcaacatcc tggggcacia gctggagtac aactacaaca gccacaaggt 14340
 ctatatcacc gccgacaagc agaagaacgg catcaaggtg aacttcaaga cccgccacia 14400
 catcgaggac ggcagcgtgc agctcgccga ccaactaccg cagaacaccc ccatcggcga 14460
 cggccccgtg ctgctgcccc acaaccacta cctgagcacc cagtccgcc tgagcaaaga 14520
 cccaacgag aagcgcgatc acatggtcct gctggagtgc gtgaccgcc cgggatcac 14580
 tctcgcatg gacgagctgt acaagtaaac ctaatctagc agctcgctga tcagcctcga 14640
 ctgtgccttc tagttgccag ccatctgttg tttgcccctc cccgtgcct tccttgacct 14700
 tggaaggtgc cactcccaact gtcctttcct aataaaatga ggaaattgca tcgcatgtc 14760
 tgagtaggtg tcattctatt ctggggggtg gggtagggca ggacagcaag ggggaggatt 14820
 gggaagacia tagcaggcat gctggggatg cggtagggct tatggcttct gaggcggaaa 14880
 gaaccagctg gggctcgatc ctctagttgg cgcgtcatgg tccatatgaa tatectcctt 14940
 agttcctatt ccgctagcct agagggacag cccccccca aagccccag ggatgtaatt 15000
 acgtccctcc cccgctaggg gcagcagcga gccgccggg gctccgctcc ggtccgcg 15060

```

tccccccgca tccccgagcc ggcagcgtgc ggggacagcc cgggcacggg gaaggtggca 15120
cgggatcgct ttcctctgaa cgcttctcgc tgctctttga gcctgcagac acctgggggg 15180
atacggggaa aaagcttttag gctgaaagag agatttagaa tgacagaatc atagaacggc 15240
ctgggttgca aaggagcaca gtgctcatcc agatccaacc ccctgctatg tgcagggtca 15300
tcaaccagca gcccaggctg cccagagcca catccagcct ggccttgaat gcctgcaggg 15360
atggggcatc cacagcctcc ttgggcaacc tgttcagtgc gtcaccaccc tctgggggaa 15420
aaactgcctc ctcatatcca acccaaacct cccctgtctc agtgtaaagc cattccccct 15480
tgtcctatca agggggagtt tgctgtgaca ttgttggtct ggggtgacac atgtttgcc 15540
attcagtgc tcacggagag gcagatcttg gggataagga agtgcaggac agcatggacg 15600
tgggacatgc aggtgttgag ggctctggga cactctccaa gtcacagcgt tcagaacagc 15660
cttaaggata agaagatagg atagaaggac aaagagcaag ttaaaacca gcatggagag 15720
gagcacaanaa agccacaga cactgctggt cctgtgtct gagcctgcat gtttgatggt 15780
gtctggatgc aagcagaagg ggtggaagag cttgctgga gagatacagc tgggtcagta 15840
ggactgggac aggcagctgg agaattgcca tgtagatgtt catacaatcg tcaaatcatg 15900
aaggctggaa aagccctcca agatcccaaa gaccaacccc aaccaccca ccgtgcccac 15960
tggccatgtc cctcagtgcc acatccccac agttcttcat cacctccagg gacggtgacc 16020
ccccacctc cgtgggcagc tgtgccactg cagcacgct ctttgagaa ggtaaattctt 16080
gctaaatcca gcccgaccct cccctggcac aacgtaaggc cattatctct catccaactc 16140
caggacggag tcagtgagga tggggctgga tccgaagcag ctccagccta cacaatcgct 16200
caagacgtgt aatgctttta ttatatatta gtcacgatat ctataacaag aaaatatata 16260
tataataagt tatcacgtaa gtagaacatg aaataacaat ataattatcg tatgagttaa 16320
atcttaaaag tcacgtaaaa gataatcatg cgtcattttg actcacgcgg tcgttatagt 16380
tcaaaatcag tgacacttac cgcatlgaca agcacgcctc acgggagctc caagcggcga 16440
ctgagatgtc ctaaatgcac agcgacggat tcgcgctatt tagaaagaga gagcaatatt 16500
tcaagaatgc atgctcaat ttacgcaga ctatctttct agggttaaaa aagatttgcg 16560
ctttactcga cctaaacttt aaacacgtca tagaatcttc gtttgacaaa aaccacattg 16620
tggggtaccg agctcttaat taaggcgcgc cggggaggtt ccctttagt agggtttaatt 16680
gcgggtcgcc ctatagttag tcgtattaca attcactggc cgtcgtttta caacgtcgtg 16740
actgggaaaa ccctggcggt acccaactta atcgcttgc agcacatccc ctttcgcc 16800
gctggcgtaa tagcgaagag gcccgcaccg atcgccctc ccaacagttg cgcagcctga 16860
atggcgaatg gcaaattgta agcgtaata tttgttaaa attcgcgtta aatttttgtt 16920
aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag 16980
aatagaccga gatagggttg agtgttggtc cagtttgtaa caagagtcca ctattaaaga 17040
acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 17100
aaccatcacc ctaatcaagt tttttgggt cgaggtgccg taaagcacta aatcggaacc 17160
ctaaaggag ccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg 17220
aaggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgtgc 17280
gcgtaaccac cacaccgcc gcgcttaatg cgccgtaca gggcgctca g 17331

```

<210> 14

<211> 17575

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 14

```

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc ttagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttccg gctggctggg ttattgctga taaatctgga gccggtgagc 840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt ttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacttc gctctgctaa 1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaaggcg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtctgtcgc 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcgagacc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggcccttttg tgcccttttg 1800
ctcacatgtt ctttctgctg ttatcccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgtctgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgcccg attcattaat 1980

```


gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
 tgagtttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt 2100
 tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
 ccaagcgcgc aattaaccct cactaaaggg aacctcccct agcttaatta accctagaaa 2220
 gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
 atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
 tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
 aaaaaaaaca aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
 ttttacaaaa ataaacttat tttgtacttt aaaaacagtc atgttgtatt ataaaataag 2520
 taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
 gactagtcgt attaagggtt ccgcatcagc ttgattcgag cccagctgg ttctttccgc 2640
 ctcagaagcc atagagccca ccgcacccc agcatgcctg ctattgtctt cccaatcctc 2700
 ccccttgctg tcctgcccc cccaccccc cagaatagaa tgacacctac tcagacaatg 2760
 cgatgcaatt tcctcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
 ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcagggc 2880
 tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcgatagaa 2940
 ggcgatgcgc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
 ttgcgcgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
 cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
 attcggcaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
 cttgagcctg gcgaacagtt cggctggcgc gagccctga tgctcttcgt ccagatcatc 3240
 ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
 gtggctgaat gggcaggtag ccgcatcaag cgtatgcagc cgccgcattg catcagccat 3360
 gatggatact ttctcggcag gagcaagggt agatgacagg agatcctgcc ccggcacttc 3420
 gcccaatagc agccagtcct tccccgttc agtgacaacg tcgagcacag ctgcgcaagg 3480
 aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcagtt cattcagggc 3540
 accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
 ggcggcatca gagcagccga ttgtctgttg tgcccagtea tagccgaata gcctctccac 3660
 ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgata ccatggttta 3720
 gttcctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
 acgtgctgct gcaggtcgaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840
 tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ctttcgggcg cccgccccgc 3900
 ccctgagccc gcccctgagc ccgcccccg accacccct tcccagcctc tgagcccaga 3960
 aagcgaagga gcaaagctgc tattggccgc tgcccaaaag gcctaccgc ttccattgct 4020
 cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttccatt tgteacgtcc 4080
 tgcacgacgc gagctgcggg gcgggggggga acttctgac taggggagga gtagaagggtg 4140
 gcgcgaaggg gccaccaaag aacggagccg gttggcgct accggtggat gtggaatgtg 4200
 tgcgagccag aggccacttg ttagcgcca agtgcacagc ggggctgcta aagcgcacgc 4260
 tccagactgc cttgggaaaa gcgcctcccc taccggtag acaccacaca gtgggtggcc 4320

tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
 tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 4440
 cttacggtaa atggccccgc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 4500
 atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
 tattttacggg aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 4620
 cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 4680
 tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catgggtcga 4740
 ggtgagcccc acgttctgct tcaactctcc catctcccc ccctccccac cccaatttt 4800
 gtattttattt attttttaat tattttgtgc agcgatgggg gcgggggggg ggggggcgcg 4860
 cgccaggcgg ggccggggcg ggccgagggg ggggcggggc gaggcggaga ggtgcggcgg 4920
 cagccaatca gagcggcgcg ctccgaaagt ttctttttat ggccgagcgg cgccggcggc 4980
 ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcc tgcctccgtg 5040
 ccccgctccg cgccgcctcg cgccgcccgc cccgctctg actgaccgcg ttactccac 5100
 aggtgagcgg gcgggacggc cttctctctc cgggctgtaa ttagcgttg gtttaatgac 5160
 ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
 gcggggggga gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
 ggcccgctgc gcccggcggc tgtgagcgtc gcgggcgcgg cgccgggctt tgtgcgtcc 5340
 gcgtgtgcgc gaggggagcg cgccgggggg cggtgcccc cggtgcgggg gggctgcgag 5400
 gggaacaaag gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg 5460
 gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc acggcccgcg 5520
 ttccgggtgcg gggctccgtg cggggcgtgg cgccggggctc gccgtgccgg gcgggggggtg 5580
 gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
 ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cgccgagccg cagccattgc 5700
 cttttatggg aatcgtgcga gagggcgag ggacttcctt tgtcccaaatt ctggcggagc 5760
 cgaaatctgg gaggcgccgc cgcaccccct ctacggggcg cgggcgaagc ggtgcggcgc 5820
 cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccttct 5880
 ccatctccag cctcggggct gccgcagggg gacggctgcc ttcggggggg acggggcagg 5940
 gcgggggttc gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
 gccttcttct ttttcttaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
 tttggcaaag aattcgccac catggtgcc aagaagaaga ggaaagtctc tagactggac 6120
 aagagcaaag tcataaaactc tgctctggaa ttactcaatg gagtcgggtat cgaaggcctg 6180
 acgacaagga aactcgctca aaagctggga gttgagcagc ctacctgta ctggcacgtg 6240
 aagaacaagc gggccctgct cgatgcctg ccaatcgaga tgctggacag gcatacacc 6300
 cactcctgcc ccctggaagg cgagtcatgg caagacttct tgcggaacaa cgccaagtca 6360
 taccgtgtg ctctcctctc acatcgcgac ggggctaaag tgcatctcgg caccgcccc 6420
 acagagaaac agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggett 6480
 tccctggaga acgactgta cgtctgtcc gccgtgggccc actttacact gggctgcgta 6540
 ttggaggaac aggagcatca agtagcaaaa gaggaagag agacacctac caccgattct 6600
 atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660

```

ttccttttcg gcctggaact aatcatatgt ggcctggaga aacagctaaa gtgcgaaagc 6720
ggcggggccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
gacgactttg accttgatat gctgcctgct gacgctcttg acgattttga ccttgacatg 6840
ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
tcacaaatth cacaataaaa gcattttttt cactgcattc tagttgttgt ttgtccaaac 7080
tcatcaatgt atcttaaggg atccctagag ggacagcccc ccccaaagc ccccgaggat 7140
gtaattacgt ccctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggctc 7200
cggcgctccc cccgcacccc cgagccggca gcgtgcgggg acagcccggg cacggggaag 7260
gtggcacggg atcgctttcc tctgaacgct tctcgctgct ctttgagcct gcagacacct 7320
gggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag 7380
aacggcctgg gttgcaaagg agcacagtgc tcatccagat ccaaccccc gctatgtgca 7440
gggtcatcaa ccagcagccc aggetgccc gagccacatc cagcctggcc ttgaatgcct 7500
gcagggatgg ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg 7560
ggggaaaaaac tgctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
cccccttgct ctatcaaggg ggagtttgct gtgacattgt tggctctgggg tgacacatgt 7680
ttgccaattc agtgcacac ggagaggcag atcttgggga taaggaagtg caggacagca 7740
tggacgtggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
atcatgaagg ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt 8100
gcccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
gtgaccccc caccctcgtg ggcagctgtg ccaactgcagc accgctcttt ggagaaggta 8220
aatcttgcta aatccagccc gacctcccc tggcacaacg taaggccatt atctctcatc 8280
caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
ggctccgctc cggctccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgcttctcg ctgctctttg 8520
agcctgcaga cactggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
atgacagaat catagaacgg cctgggttgc aaaggagcac agtgctcatc cagatccaac 8640
cccctgctat gtgcagggtc atcaaccagc agcccagget gccagagcc acatccagcc 8700
tggccttgaa tgctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
cagtgtaaag ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
tggggtgaca catgtttgcc aattcagtgc atcacggaga ggcagatctt ggggataagg 8940
aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000

```

```

agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
gttaaaaccc agcatggaga ggagcacaaa aaggccacag acactgctgg tccctgtgtc 9120
tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctgg 9180
agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
tcatacaatc gtcaaatcat gaaggctgga aaagccctcc aagatcccca agaccaaccc 9300
caaccacccc accgtgcccc ctggccatgt ccctcagtgc cacatcccca cagttcttca 9360
tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcacgcg 9420
tctttggaga aggtaaatct tgctaaatcc agccccaccc tcccctggca caacgtaagg 9480
ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
ccctatcagt gatagagaac gtatgaagag ttactccct atcagtgata gagaacgtat 9600
gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
tagagaacgt atgaccagtt tactccctat cagtgataga gaacgtatct acagtttact 9720
ccctatcagt gatagagaac gtatatccag ttactccct atcagtgata gagaacgtat 9780
aagctttagg cgtgtacggg gggcgccctat aaaagcagag ctcgtttagt gaaccgtcag 9840
atgcctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttccca 9900
ccctcgtaaa aagcttgtcc accatggctc ctaagaaaaa gcggaagggtg gacaagaaat 9960
actcaatcgg gctggccatc ggaactaact cagtgggggtg ggcagtcatt actgacgagt 10020
acaaagtgcc aagcaagaaa ttaaggtcc tgggcaacac cgataggcac tccatcaaga 10080
aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
ttttctctaa cgaaatggcc aagggtggacg atagtttctt tcateggctg gaggaatcat 10260
tcttggtcga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
aggtcgctta tcacgaaaaa taccaccca tctatcatct gcgcaagaaa ctggtggact 10380
ctacagataa agcagacctg cggctgatct atctggccct ggctcacatg attaagttca 10440
gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
tcatccagct ggtccagaca tacaatcagc tgtttgagga aaaccctatt aatgcatctg 10560
gcgtggacgc aaaagccatc ctgagtgccg ggctgtctaa gagtagaagg ctggagaacc 10620
tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcaactgt 10680
cactgggact gacacctaac ttcaagagca attttgatct ggccgaggac gctaaactgc 10740
agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcgagatc 10800
agtacgcaga cctgttcctg gccgctaaga atctgtctga cgctatcctg ctgagtgata 10860
ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagagat 10920
atgacgagca ccatcaggat ctgacctgc tgaaggcact ggtgcgccag cagctgcccc 10980
agaagtacaa ggaaatcttc tttgatcaga gtaagaacgg gtacgccggt tatattgacg 11040
gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
gcaccgagga actgctgggt aaactgaatc gcgaggacct gctgcgcaag cagcgacat 11160
ttgataacgg ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac 11220
gacaggaaga cttctacca tttctgaagg ataaccgca gaagatcgaa aaaattctga 11280
ccttccggat cccttactat gtggggcccc tggcaagggg taattcccgc tttgcctgga 11340

```

tgacacggaa	atctgaggaa	acaatcactc	cttggaaactt	cgaggaagtg	gtcgataagg	11400
gagcttccgc	acagtctttc	atcgagagaa	tgacaaactt	cgacaaaaac	ctgccaaatg	11460
agaaagtgct	gcctaagcac	agtctgctgt	acgagtattt	cacagtctat	aacgaactga	11520
ctaaggtgaa	atacgtcacc	gaggggatga	ggaagcccg	cttcctgagc	ggtgaacaga	11580
agaaagctat	cgtggacctg	ctgtttaaaa	ccaatcgcaa	ggtgacagtc	aagcagctga	11640
aggaggacta	cttcaagaaa	attgaatgtt	tcgattctgt	ggagatcagt	ggcgtcgaag	11700
acagatttaa	cgcttctctg	ggaacctacc	acgatctgct	gaagatcatt	aaggataaag	11760
acttcctgga	caacgaggaa	aatgaggata	tcctggaaga	cattgtgctg	accctgacac	11820
tgtttgagga	tcgcgaaatg	atcgaggaa	ggctgaaaac	ttatgccccat	ctgttcgatg	11880
acaaggtgat	gaaacagctg	aagcgaagaa	ggtacaccgg	ctggggacga	ctgagcagaa	11940
agctgatcaa	cggcattcgg	gacaaacaga	gtggaaagac	tatcctggac	tttctgaaat	12000
cagatggctt	cgctaacaga	aattttatgc	agctgattca	cgatgacagc	ctgaccttca	12060
aagaggatat	ccagaaggca	caggtgtccg	ggcagggtga	ctctctgcac	gagcatatcg	12120
caaacctggc	cgggtcccc	gccatcaaga	aaggtattct	gcagaccgtg	aaggtggctg	12180
atgagctgg	gaaagtcag	ggcaggcata	agccagaaaa	catcgtgatt	gagatggccc	12240
gcgaaaatca	gaccacacag	aaaggacaga	agaacagccg	cgagcggatg	aaaaggatcg	12300
aggaaggcat	taaggaaactg	ggatcccaga	tcctgaaaga	gcaccctgtg	gaaaacactc	12360
agctgcagaa	tgagaagctg	tatctgtact	atctgcagaa	tgggcgggat	atgtacgtgg	12420
accaggagct	ggatattaac	cgactgtctg	attacgacgt	ggatgccatc	gtcccacagt	12480
cattcctgaa	agatgacagc	attgacaata	aggtgctgac	ccggagtgc	aaaaaccgag	12540
gaaagagtga	taatgtccct	tcagaggaag	tggatcaagaa	aatgaagaac	tactggagac	12600
agctgctgaa	tgccaaactg	atcacacagc	gaaagtttga	taacctgact	aaagctgaga	12660
gagggggtct	gtcagaactg	gacaaagcag	gcttcatcaa	gcgacagctg	gtggagacca	12720
gacagatcac	aaagcacgtc	gctcagattc	tggatagcag	gatgaacaca	aagtacgatg	12780
agaatgacaa	actgatccgc	gaagtgaagg	tcattactct	gaagtcaaaa	cttgtgagcg	12840
acttcagaaa	ggatttccag	ttctacaaag	tcaggggatg	caacaattat	caccatgctc	12900
atgacgcata	cctgaacgca	gtggctcgga	ccgccctgat	taagaaatac	cccaaactgg	12960
agagcgaatt	cgtgtacggt	gactataagg	tgtacgatgt	cagaaaaatg	atcgccaaga	13020
gtgagcagga	aattggaaaa	gccaccgcta	agtattttct	ttactcaaac	atcatgaatt	13080
tccttaagac	tgagatcacc	ctggcaaatg	gggaaatccg	aaagagacca	ctgattgaga	13140
ctaacggcga	gaccggagaa	atcgtgtggg	acaagggtag	ggattttgcc	acagtgcgca	13200
aggtcctgtc	catgcctcaa	gtgaatattg	tcaagaaaac	agaggtgcag	actggcggat	13260
tcagtaagga	atcaattctg	cccaaacgga	actctgataa	gctgatcgcc	cgaaagaaag	13320
actgggatcc	caagaaatat	gggggtttcg	actccccaac	agtggcttac	tctgtcctgg	13380
tggtcgcaaa	ggtggagaag	gggaaaagca	agaaactgaa	atccgtcaag	gagctgctgg	13440
gtatcactat	tatggagagg	agtccttcg	agaagaacce	catcgatttt	ctggaggcta	13500
aaggtataaa	ggaagtgaag	aaagacctga	tcattaaact	gccaaagtac	agcctgtttg	13560
agctggaaaa	cgggaaggaa	cgaatgctgg	catccgcagg	agagctgcag	aagggtaatg	13620
aactggccct	gccttctaag	tacgtgaact	tcctgtatct	ggctagccac	tacgagaagc	13680

```

tgaaaggctc ccccgaggat aacgaacaga aacagctggt tgtggagcag cacaagcatt 13740
atctggacga gatcattgaa cagattagcg agttctccaa aagagtgatc ctggctgacg 13800
caaatctgga taaggctcctg agcgcataca acaaacacag agataagcca atcagggagc 13860
aggccgaaaa tatcattcat ctgttcactc tgaccaacct gggagcccct gcagccttca 13920
agtattttga cactaccatc gatcggaac gatacacatc cactaaggag gtgctggacg 13980
ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
tgggggggcga cctcgaggga agcggagagg gcagaggaag tctgctaaca tgcggtgacg 14100
tcgaggagaa tcctggccca gcaccgggat ccatggtgag caagggcgag gagctgttca 14160
ccgggggtggt gcccacctcgt gtcgagctgg acggcgacgt aaacggccac aagttcagcg 14220
tgtccggcga gggcgagggc gatgccacct acggcaagct gacctgaag ttcactctga 14280
ccaccggcaa gctgcccgtg ccttggccca cctcgtgac caccttacc tacggcgtgc 14340
agtgtttcgc ccgtacccc gaccacatga agcagcacga cttcttcaag tccgcatgc 14400
ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacc 14460
gcgccgaggt gaagttcgag ggcgacacc tggtgaaccg catcgagctg aagggcacg 14520
acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 14580
aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaacttc aagaccgcc 14640
acaacatcga ggacggcagc gtgcagctcg ccgaccata ccagcagaac acccccatcg 14700
gcgacggccc cgtgctgctg cccgacaacc actacctgag caccagtcg gccctgagca 14760
aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgccggga 14820
tactctcgg catggacgag ctgtacaagt aaacctaatc tagcagctcg ctgatcagcc 14880
tcgactgtgc cttctagtgt ccagccatct gttgtttgcc cctccccctg gccttcttgc 14940
accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcac 15000
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 15060
gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg 15120
gaaagaacca gctggggctc gatcctctag ttggcgctc atggtccata tgaatacct 15180
ccttagttcc tattccgcta gcctagaggg acagcccccc ccaaagccc ccagggatgt 15240
aattacgtcc ctccccgct aggggcagca gcgagccgcc cggggctccg ctccggtccg 15300
gcgctcccc cgcatccccg agccggcagc gtgcggggac agcccgggca cggggaaggt 15360
ggcacgggat cgctttcctc tgaacgctt tcgtgctctt ttgagcctgc agacacctgg 15420
ggggatacgg ggaaaaagct ttaggctgaa agagagattt agaatgacag aatcatagaa 15480
cggcctgggt tgcaaggag cacagtgtc atccagatcc aacccctgc tatgtgcagg 15540
gtcatcaacc agcagcccag gctgccaga gccacatcca gctggcctt gaatgcctgc 15600
agggatgggg catccacagc ctcttgggc aacctgttca gtgcgtcacc acctctggg 15660
ggaaaaactg cctcctcata tccaacccaa acctccctg tctcagtgt aagccattcc 15720
cccttgcct atcaaggggg agtttgcgtg gacattgttg gtctggggtg acacatgttt 15780
gccaattcag tgcatcacgg agaggcagat cttggggata aggaagtgca ggacagcatg 15840
gacgtgggac atgcagggtg tgagggtctt gggacactct ccaagtcaca gcgttcagaa 15900
cagccttaag gataagaaga taggatagaa ggacaaagag caagttaaaa ccagcatgg 15960
agaggagcac aaaaaggcca cagacactgc tggtcctgt gtctgagcct gcatgtttga 16020

```

tgggtgtctgg atgcaagcag aaggggtgga agagcttgcc tggagagata cagctgggtc 16080
 agtaggactg ggacaggcag ctggagaatt gccatgtaga tgttcataca atcgtcaaat 16140
 catgaaggct ggaaaagccc tccaagatcc ccaagaccaa cccaaccca cccaccgtgc 16200
 ccaactggcca tgtccctcag tgccacatcc ccacagttct tcatcacctc cagggacggt 16260
 gaccccccca cctccgtggg cagctgtgcc actgcagcac cgctctttgg agaaggtaaa 16320
 tcttgctaaa tccagcccga ccctcccctg gcacaacgta aggccattat ctctcatcca 16380
 actccaggac ggagtcagtg aggatggggc tggatccgaa gcagctccag cctacacaat 16440
 cgctcaagac gtgtaatgct tttattatat attagtcacg atatctataa caagaaaata 16500
 tatatataat aagttatcac gtaagtagaa catgaaataa caatataatt atcgtatgag 16560
 ttaaattctta aaagtcacgt aaaagataat catgcgtcat tttgactcac gcggtcgta 16620
 tagttcaaaa tcagtgcacac ttaccgcatt gacaagcacg cctcacggga gctccaagcg 16680
 gcgactgaga tgcctaaat gcacagcgac ggattcgcgc tatttagaaa gagagagcaa 16740
 tattttcaaga atgcatgcgt caattttacg cagactatct ttctagggtt aaaaaagatt 16800
 tgcgctttac tcgacctaaa ctttaaacac gtcatagaat cttcgtttga caaaaaccac 16860
 attgtggggg accgagctct taattaagge gcgccgggga ggttcccttt agtgagggtt 16920
 aattgcgggt cgccctatag tgagtcgtat tacaattcac tggccgtcgt ttacaacgt 16980
 cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tcccccttc 17040
 gccagctggc gtaatagcga agaggccgc accgatcgcc cttcccaaca gttgcgcagc 17100
 ctgaatggcg aatggcaaat tgtaagcgtt aatattttgt taaaattcgc gttaaatttt 17160
 tgttaaatca gctcatTTTT taaccaatag gccgaaatcg gcaaaatccc ttataaatca 17220
 aaagaataga ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 17280
 aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta 17340
 cgtgaaccat caccctaate aagttttttg gggctcgaggt gccgtaaagc actaaatcgg 17400
 aaccctaaag ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 17460
 aaggaaggga agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg 17520
 ctgcgcgtaa ccaccacacc cgccgcgctt aatgcgcgcg tacagggcgc gtcag 17575

<210> 15

<211> 17950

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 15

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
 caaatatgta tccgctcatg agacaataac cctgataaat gtttcaataa tattgaaaaa 120
 ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
 gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt 300
 ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 360

tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttgggt tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtctgcc ataaccatga gtgataacac tgcggccaac ttactttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttccg gctggctgggt ttattgtctga taaatctgga gccggtgagc 840
gtgggtctcg cggatatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgettcaaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaagggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 1380
tctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcgggt cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg 1800
ctcacatggt ctttctgctg ttatccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccccata cgcaaaccgc ctctccccgc gcgttggccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcacccagg ctttacactt tatgcttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctccct agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaca aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat tttgtacttt aaaaacagtc atgttgattt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaagggtt ccggatcagc ttgattcgag cccagctgg ttctttccgc 2640
ctcagaagcc atagagccca ccgcacccc agcatgcctg ctattgtctt cccaatctc 2700


```

cccccttgctg tcctgcccc cccaccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tcctcatttt attaggaaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcgaggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
ggcgaatgcgc tgcaaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
attcggaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
cttgagcctg gcgaacagtt cggctggcgc gagccctga tgcctctcgt ccagatcatc 3240
ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgattg catcagccat 3360
gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtcct tcccccttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgtgcctcg tcctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcggcatca gagcagccga ttgtctgttg tgcccagtc tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720
gttctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
acgtgctgct gcaggctgaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840
tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcgggcg cccgccccgc 3900
ccctgagccc gcccctgagc ccgccccgg accacccct tcccagctc tgagcccaga 3960
aagcgaagga gcaaagctgc tattggccgc tgcccaaag gcctaccgc ttccattgct 4020
cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttcatt tgtcacgtcc 4080
tgcacgacgc gagctgcggg gcggggggga acttcctgac taggggagga gtagaagggtg 4140
gcgcgaaggg gccaccaaag aacggagccg gttggcgct accggtggat gtggaatgtg 4200
tgcgagccag aggccacttg ttagcgcca agtgcccagc ggggctgcta aagcgcatgc 4260
tccagactgc ctggggaaaa gcgcctcccc taccggtag acacccaca gtgggtggcc 4320
tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
tagtaatcaa ttacggggtc attagttcat agccatata tggagtccg cgttacataa 4440
cttacggtaa atggcccgc tggtgaccg ccaacgacc cccgccatt gacgtcaata 4500
atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
tatttacggt aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgcc 4620
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 4680
tgggactttc ctacttgga gtacatctac gtattagtc tcgtattac catgggtcga 4740
ggtgagcccc acgttctgct tcactctccc catctcccc cctccccac cccaatttt 4800
gtatttattt attttttaat tattttgtgc agcgtgggg gcgggggggg ggggggcgcg 4860
cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg 4920
cagccaatca gagcggcgcg ctccgaaagt ttcttttat ggcgaggcgg cggcggcggc 4980
ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttcct tcgccccgtg 5040

```

```

ccccgctccg cgccgcctcg cgccgcccgc cccggctctg actgaccgcg ttactcccac 5100
aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa ttagcgcttg gtttaatgac 5160
ggctcgtttc ttttctgttg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
gcggggggga gcggtcggg ggggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
ggcccgcgct gcccggcggc tgtgagcgct gcgggcgcgg cgcggggctt tgtgcgctcc 5340
gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg gggctgcgag 5400
gggaacaaag gctgcgtgcg ggggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg 5460
gcggtcgggc tgtaaccccc cctgcaccc cctcccccga gttgctgagc acggccccgc 5520
ttcgggtgcg gggctccgtg cggggcgttg cgcggggctc gccgtgccgg gcgggggggtg 5580
gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc 5700
cttttatggt aatcgtgcga gagggcgcag ggacttcctt tgtcccaaat ctggcggagc 5760
cgaaatctgg gaggcgccgc cgcaccccc ctacggggcg cgggcgaagc ggtgcggcgc 5820
cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgc gtecccttct 5880
ccatctccag cctcggggct gccgcagggg gacggctgcc ttcggggggg acggggcagg 5940
gcgggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
gccttcttct ttttcttaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
tttgcaaaag aattcgccac catggtgcc aagaagaaga ggaaagtctc tagactggac 6120
aagagcaaag tcataaactc tgctctggaa ttactcaatg gagtcggtat cgaaggcctg 6180
acgacaagga aactcgctca aaagctggga gttgagcagc ctacctgta ctggcacgtg 6240
aagaacaagc gggccctgct cgatgccctg ccaatcgaga tgctggacag gcatcatacc 6300
cactcctgcc cctggaagg cgagtcatgg caagacttct tgcggaacaa cgccaagtca 6360
taccgctgtg ctctcctctc acatcgcgac ggggctaaag tgcatctcgg caccgcacca 6420
acagagaaac agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggcttc 6480
tccttgaga acgcactgta cgctctgtcc gccgtgggcc actttacact gggctgcgta 6540
ttggaggaac aggagcatca agtagcaaaa gaggaagag agacacctac caccgattct 6600
atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
ttccttttcg gcctggaact aatcatatgt ggcttgaga aacagctaaa gtgcgaaagc 6720
ggcgggccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
gacgactttg accttgatat gctgcctgct gacgtctttg acgattttga cttgacatg 6840
ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgttgt ttgtccaaac 7080
tcatcaatgt atcttaagg atccctagag ggacagcccc ccccaaagc cccagggat 7140
gtaattacgt cctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggtc 7200
cggcgtccc ccgcacccc cgagccggca gcgtgcgggg acagcccggg cacggggaag 7260
gtggcacggg atcgctttcc tctgaacgct tctcgtgct ctttgagcct gcagacacct 7320
gggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag 7380

```

aacggcctgg gttgcaaagg agcacagtgc tcatccagat ccaaccccct gctatgtgca 7440
 ggggtcatcaa ccagcagccc aggctgccc gagccacatc cagcctggcc ttgaatgcct 7500
 gcagggatgg ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccaccctctg 7560
 ggggaaaaaac tgcctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
 ccccttgtgc ctatcaaggg ggagtttgct gtgacattgt tggcttgggg tgacacatgt 7680
 ttgccaatc agtgcacac ggagaggcag atcttgggga taaggaagtg caggacagca 7740
 tggacgtggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
 aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
 ggagaggagc acaaaaaggc cacagacact gctggtecc tgtctgagc ctgcatgttt 7920
 gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
 tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
 atcatgaagg ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt 8100
 gccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
 gtgaccccc caccctccgt ggagctgtg ccaactgcag accgctcttt ggagaaggta 8220
 aatcttgcta aatccagccc gaccctcccc tggeacaacg taaggccatt atctctcacc 8280
 caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
 aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
 ggctccgctc cggctccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
 ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgttctctg ctgctctttg 8520
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgcctcag cagatccaac 8640
 cccctgctat gtgcagggtc atcaaccagc agcccaggt gccagagcc acatccagcc 8700
 tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagt 8760
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
 cagtgtaaag ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
 tggggtgaca catgtttgcc aattcagtgc atcacggaga ggcagatctt ggggataagg 8940
 aagtgcagga cagcatggac gtgggacatg cagtggttga gggctctggg acactctcca 9000
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
 gttaaaaccc agcatggaga ggagcacaaa aaggccacag aactgctgg tccctgtgtc 9120
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctg 9180
 agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
 tcatacaatc gtcaaatcat gaaggtgga aaagccctcc aagatcccca agaccaaccc 9300
 caaccacccc accgtgccc ctggccatgt cctcagtgc cacatcccca cagttcttca 9360
 tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
 tctttggaga aggtaaatct tgctaaatcc agcccgacc tcccctggca caacgtaagg 9480
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
 ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
 gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
 tagagaacgt atgaccagtt tactccctat cagtgataga gaacgtatct acagtttact 9720

ccctatcagt gatagagaac gtatatccag ttactccct atcagtgata gagaacgtat 9780
 aagcttttagg cgtgtacggt gggcgccctat aaaagcagag ctgcttttagt gaaccgtcag 9840
 atcgccctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttccta 9900
 cccctcgtaaa aagcttgtcc accatggctc ctaagaaaaa gcggaagggtg gacaagaaat 9960
 actcaatcgg gctggacatc ggaactaact cagtgggggtg ggagtcatt actgacgagt 10020
 acaaagtgcc aagcaagaaa tttaagggtc tgggcaacac cgataggcac tccatcaaga 10080
 aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
 aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
 ttttctctaa cgaaatggcc aaggtggacg atagtctctt tcatcggtg gaggaatcat 10260
 tcctggctga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
 aggtcgctta tcacgaaaaa taccaccca tctatcctt gcgcaagaaa ctgggtggact 10380
 ctacagataa agcagacctg cggtgatct atctggcct ggctccatg attaagttca 10440
 gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
 tcatccagct ggtccagaca tacaatcagc tgtttgagga aaacctatt aatgcctctg 10560
 gcgtggacgc aaaagccatc ctgagtcca ggctgtctaa gagtagaagg ctggagaacc 10620
 tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcactgt 10680
 cactgggact gacacctaac ttcaagagca attttgatct ggccgaggac gctaaactgc 10740
 agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcgagatc 10800
 agtacgcaga cctgttctg gccgctaaga atctgtctga cgctatcctg ctgagtata 10860
 ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagatat 10920
 atgacgagca ccatcaggat ctgacctgc tgaaggcact ggtgcgccag cagctgccc 10980
 agaagtacaa ggaaatctt ttgatcaga gtaagaacgg gtacgccgt tatattgacg 11040
 gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
 gcaccgagga actgctggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat 11160
 ttgataacgg ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac 11220
 gacaggaaga cttctacca tttctgaagg ataaccgca gaagatcgaa aaaattctga 11280
 ccttccggat cccttactat gtggggcccc tggcaagggg taattcccgc tttgcctgga 11340
 tgacacggaa atctgaggaa acaatcactc cttggaactt cgaggaagtgt gtcgataagg 11400
 gagcttccgc acagtcttct atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg 11460
 agaaagtgtc gcctaagcac agtctgctgt acgagtattt cacagtctat aacgaactga 11520
 ctaagggtgaa atacgtcacc gaggggatga ggaagccgc cttctgagc ggtgaacaga 11580
 agaaagctat cgtggacctg ctgtttaaaa ccaatcgcaa ggtgacagtc aagcagctga 11640
 aggaggacta cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag 11700
 acagatttaa cgcttctctg ggaacctacc acgatctgt gaagatcatt aaggataaag 11760
 acttctgga caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac 11820
 tgtttgagga tcgcgaaatg atcgaggaaac ggctgaaaac ttatgcccct ctgttcgatg 11880
 acaaggatgat gaaacagctg aagcgaagaa ggtacaccgg ctggggacga ctgagcagaa 11940
 agctgatcaa cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat 12000
 cagatggctt cgctaacaga aattttatgc agctgattca cgatgacagc ctgaccttca 12060

```

aagaggatat ccagaaggca caggtgtccg ggcagggtga ctctctgcac gagcatatcg 12120
caaacctggc cgggtccccc gccatcaaga aaggtattct gcagaccgtg aaggtggctg 12180
atgagctggg gaaagtcatt ggcaggcata agccagaaaa catcgtgatt gagatggccc 12240
gcgaaaaatca gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg 12300
aggaaggcat taaggaactg ggatcccaga tcctgaaaga gcaccctgtg gaaaacactc 12360
agctgcagaa tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg 12420
accaggagct ggatattaac cgactgtctg attacgacgt ggatcatatc gtcccacagt 12480
cattcctgaa agatgacagc attgacaata aggtgctgac ccggagtgcg aaaaaccgag 12540
gaaagagtga taatgtccct tcagaggaag tggatcaaga aatgaagaac tactggagac 12600
agctgctgaa tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga 12660
gaggggggtct gtcagaactg gacaaagcag gcttcatcaa gcgacagctg gtggagacca 12720
gacagatcac aaagcacgtc gctcagatc tggatagcag gatgaacaca aagtacgatg 12780
agaatgacaa actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg 12840
acttcagaaa ggatttccag ttctacaaaag tcaggagat caacaattat caccatgctc 12900
atgacgcata cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg 12960
agagcgaatt cgtgtacggg gactataagg tgtacgatgt cagaaaaatg atcgccaaga 13020
gtgagcagga aattggaaaa gccaccgcta agtatttctt ttactcaaac atcatgaatt 13080
tctttaagac tgagatcacc ctggcaaatg gggaaatccg aaagagacca ctgattgaga 13140
ctaacggcga gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcgca 13200
aggtcctgtc catgcctcaa gtgaatattg tcaagaaaac agaggtgcag actggcggat 13260
tcagtaagga atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag 13320
actgggatcc caagaaatat gggggtttcg actccccaac agtggttac tctgtcctgg 13380
tggtcgcaaa ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg 13440
gtatcactat tatggagagg agtccttcg agaagaacct catcgatttt ctggaggcta 13500
aaggctataa ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg 13560
agctggaaaa cggaagggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg 13620
aactggccct gccttctaag tacgtgaact tcctgtatct ggctagccac tacgagaagc 13680
tgaaaggctc ccccgaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt 13740
atctggacga gatcattgaa cagattagcg agttctccaa aagagtgatc ctggctgacg 13800
caaatctgga taaggtcctg agcgcatata acaaacacag agataagcca atcaggagac 13860
aggccgaaaa tatcattcat ctgttcaact tgaccaacct gggagccct gcagccttca 13920
agtattttga cactaccatc gatcggaaac gatacacatc cactaaggag gtgctggacg 13980
ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
tggggggcca cctcgaggat ggcggtggcg cgtgtcccc gcagcactcc gccgtgacct 14100
aggggagtat aatcaaaaaa aaagagggca tggatgctaa gagccttacc gcctggctcc 14160
gaacactggg cacgtttaag gatgtgttcg tcgattttac ccgggaggag tggaaactgc 14220
tcgacaccgc gcagcagatc gtgtaccgga atgtcatgct cgaaaattac aaaaacttgg 14280
tcagcctcgg gtaccaattg accaaaccag atgtcatact gcgactggaa aaaggagagg 14340
aaccctggct cgtcgagcgc gaaattcatc aagaaacaca cccgattct gaaaccgct 14400

```

```

tcgagattaa gagcagtgtg cctaggctcg agggaagcgg agagggcaga ggaagtctgc 14460
taacatgcgg tgacgtcgag gagaatcctg gcccgacc gggatccatg gtgagcaagg 14520
gcgaggagct gttcaccggg gtggtgcca tcctggtcga gctggacggc gacgtaaacg 14580
gccacaagtt cagcgtgtcc ggcgagggcg agggcgatgc cacctacggc aagctgaccc 14640
tgaagttcat ctgcaccacc ggcaagctgc ccgtgccctg gccaccctc gtgaccacct 14700
tcacctacgg cgtgcagtgc ttgccccgt accccgacca catgaagcag cagcacttct 14760
tcaagtccgc catgcccga ggctacgtcc aggagcgcac catctttctt aaggacgacg 14820
gcaactacaa gacccgcgcc gaggtgaagt tcgagggcga caccctggtg aaccgcatcg 14880
agctgaaggg catcgacttc aaggaggacg gcaacatcct ggggcacaag ctggagtaca 14940
actacaacag ccacaaggtc tatatcaccg ccgacaagca gaagaacggc atcaaggtga 15000
acttcaagac ccgccacaac atcgaggacg gcagcgtgca gctcgccgac cactaccagc 15060
agaacacccc catcggcgac ggccccgtgc tgetgcccga caaccactac ctgagcacc 15120
agtccgccct gagcaaagac cccaacgaga agcgcgatca catggtcctg ctggagtctg 15180
tgaccgccgc cgggatcact ctcgcatgg acgagctgta caagtaaacc taatctagca 15240
gctcgtgat cagcctcgac tgtgccttct agttgccage catctgttgt ttgccccctc 15300
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttctta ataaaatgag 15360
gaaattgcat cgcattgtct gagtagtgt cattctatc tggggggtgg ggtggggcag 15420
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 15480
atggcttctg aggcggaaag aaccagctgg ggctcgatcc tctagttggc gcgtcatggt 15540
ccatatgaat atcctcctta gttcctatc cgctagccta gagggacagc cccccccaa 15600
agcccccagg gatgtaatta cgtccctccc ccgctagggg cagcagcgag ccgccccggg 15660
ctccgctccg gtccggcgct cccccgcat ccccgagccg gcagcgtgcg gggacagccc 15720
gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 15780
cctgcagaca cctgggggga tacggggaaa aagctttagg ctgaaagaga gatttagaat 15840
gacagaatca tagaacggcc tgggttgcaa aggagcacag tgctcatcca gatccaacc 15900
cctgctatgt gcagggtcat caaccagcag ccaggctgc ccagagccac atccagcctg 15960
gccttgaatg cctgcaggga tggggcatcc acagcctcct tgggcaacct gttcagtgcg 16020
tcaccaccct ctgggggaaa aactgcctcc tcatatccaa cccaaacctc ccctgtctca 16080
gtgtaaagcc attccccctt gtccatatca gggggagttt gctgtgacat tgttggctctg 16140
gggtgacaca tgtttgcaa ttcagtgcac cacggagagg cagatcttgg ggataaggaa 16200
gtgcaggaca gcatggacgt gggacatgca ggtgttgagg gctctgggac actctccaag 16260
tcacagcgtt cagaacagcc ttaaggataa gaagatagga tagaaggaca aagagcaagt 16320
taaaacccag catggagagg agcacaaaaa ggccacagac actgctggtc cctgtgtctg 16380
agcctgcatg tttgatggtg tctggatgca agcagaaggg gtggaagagc ttgcctggag 16440
agatacagct gggtcagtag gactgggaca ggcagctgga gaattgccat gtagatgttc 16500
atacaatcgt caaatcatga aggttgaaa agccctccaa gatecccaag accaaccaca 16560
accacccac cgtgcccact ggccatgtcc ctcaagtcca catcccaca gttcttcate 16620
acctccaggg acgtgaccc cccacctcc gtgggcagct gtgccactgc agcaccgctc 16680
tttgagaaag gtaaattctt ctaaattccag cccgaccctc ccctggcaca acgtaaggcc 16740

```

attatctctc atccaactcc aggacggagt cagtgaggat ggggctggat ccgaagcagc 16800
 tccagcctac acaatcgctc aagacgtgta atgcttttat tatatattag tcacgatatc 16860
 tataacaaga aaatatatat ataataagtt atcacgtaag tagaacatga aataacaata 16920
 taattatcgt atgagttaaa tcttaaaagt cacgtaaaag ataatcatgc gtcattttga 16980
 ctcacgcggt cggttatagtt caaaatcagt gacacttacc gcattgacaa gcacgcctca 17040
 cgggagctcc aagcggcgac tgagatgtcc taaatgcaca gcgacggatt cgcgctatct 17100
 agaaagagag agcaatatct caagaatgca tgcgtcaatt ttacgcagac tatctttcta 17160
 ggggttaaaaa agatttgctc tttactcgac ctaaacttta aacacgtcat agaactctcg 17220
 tttgacaaaa accacattgt ggggtaccga gctcttaatt aaggcgcgcc ggggaggttc 17280
 ccttttagtga ggggttaattg cgggtcgccc tatagtgagt cgtattacaa ttcactggcc 17340
 gtcgttttac aacgtcgtga ctgggaaaac cctggcggtta cccaacttaa tcgccttgca 17400
 gcacatcccc ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc 17460
 caacagttgc gcagcctgaa tggcgaatgg caaattgtaa gcgttaatat tttgttaaaa 17520
 ttcgcgttaa atttttgtta aatcagctca ttttttaacc aataggccga aatcgcaaaa 17580
 atcccttata aatcaaaaaga atagaccgag atagggttga gtgttggtcc agtttggaac 17640
 aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag 17700
 ggcgatggcc cactacgtga accatcccc taatcaagtt ttttggggtc gaggtgccgt 17760
 aaagcactaa atcggaaccc taaagggagc ccccgattta gagcttgacg gggaaagccg 17820
 gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag ggcgctggca 17880
 agtgtagcgg tcacgctgcg cgtaaccacc acaccgccg cgcttaatgc gccgctacag 17940
 ggcgcgtcag 17950

<210> 16

<211> 17451

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 16

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
 caaatatgta tccgctcatg agacaataac cctgataaat gtttcaataa tattgaaaaa 120
 ggaagagtat gagtattcaa catttcctgt tcgcccttat tccctttttt gcggcatttt 180
 gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
 tgggtgcacg agtgggttac atcgaaactgg atctcaacag cggtaagatc cttgagagtt 300
 ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
 tattatcccc tattgacgcc gggcaagagc aactcggctc ccgcatacac tattctcaga 420
 atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
 gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
 ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660

```

ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 720
ctctagcttc cgggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 840
gtgggtctcg cggatatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcacgcgc tacatactc gctctgctaa 1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggctttttgc tggccttttg 1800
ctcacatggt ctttctgctg ttatcccctg attctgtgga taaccgtatt accgcctttg 1860
agttagctga taccgctcgc gcgagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctccctt agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaa aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat ttgtactttt aaaaacagtc atgttgtatt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaagggtt ccggatcagc ttgattcgag cccagctgg ttctttccgc 2640
ctcagaagcc atagagccca ccgcacccc agcatgcctg ctattgtctt cccaatcctc 2700
ccccttgctg tcctgcccc acccccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tcctcatttt attaggaaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcgaggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
ggcgatgcgc tgccaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000

```



```

ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
cgccacacccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
attcggaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
cttgagcctg gcgaacagtt cggctggcgc gagcccccta tgctcttcgt ccagatcatc 3240
ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgcattg catcagccat 3360
gatggatact ttctcggcag gagcaagggt agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtccc tccccgttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcgcatca gagcagccga ttgtctgttg tgcccagtea tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggcgggt 3720
atcgataagc tagcttgggc tgcaggtcga gggacctaat taagggttcc ggatccacta 3780
gttctagagc ggctcgact ctacgatacc gtcgatcccc actggaaaga ccgcgaagag 3840
tttgtctca accgcgagct gtggaaaaaa aaggacagc ataagtatga catcatcaag 3900
gaaaccctgg actactgcgc cctacagatc cctgaagttc ctatactttc tagagaatag 3960
gaacttcgga ataggaactt caaagaacgc gtaccccaca gtgggtggcc tagggacagg 4020
attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa tagtaatcaa 4080
ttacggggtc attagtcat agcccatata tggagttccg cgttacataa cttacggtaa 4140
atggcccgcg tggctgaccg cccaacgacc cccgccatt gacgtcaata atgacgtatg 4200
tccccatagt aacgccaata gggactttcc attgacgtca atgggtggac tatttacggt 4260
aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgcc cctattgacg 4320
tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta tgggactttc 4380
ctacttggca gtacatctac gtattagtca tcgctattac catgggtcga ggtgagcccc 4440
acgttctgct tcaactctcc catctcccc cctccccac cccaatttt gtattttatt 4500
attttttaat tattttgtgc agcgatgggg gcgggggggg ggggggcgcg cgccaggcgg 4560
ggcggggcgg gcgagggggc ggggcggggc gaggcggaga ggtgcggcgg cagccaatca 4620
gagcggcgcg ctccgaaagt ttctttttat ggcgaggcgg cggcggcggc ggccctataa 4680
aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgctc tcgccccgtg ccccgctccg 4740
cgccgcctcg cgccggcgcg cccggctctg actgaccgcg ttactccac aggtgagcgg 4800
gcgggacggc cttctctc cgggtgttaa ttagcgttg gtttaatgac ggctcgtttc 4860
ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt gcggggggga 4920
gcggctcggg ggggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc ggccgcgct 4980
gcccggcggc tgtgagcgt gcggggcgcg cgcggggctt tgtgcgtcc gcgtgtgcgc 5040
gaggggagcg cgccgggggg cggtgccccg cggtgccggg gggctgcgag gggaacaaag 5100
gctgcgtgcg ggggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg gcggtcgggc 5160
tgtaaccccc ccctgcaccc cctccccga gttgctgagc acggcccggc ttcgggtgcg 5220
gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg gcggggggtg gcggcaggtg 5280
ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga ggggcgcggc 5340

```

```

ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc cttttatggt 5400
aatcgtgcga gagggcgcag ggacttcctt tgccccaaat ctggcgggagc cgaaatctgg 5460
gaggcgccgc cgcacccccct ctagcgggcg cgggcgaagc ggtgcggcgc cggcaggaag 5520
gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtcccccttct ccatctccag 5580
cctcggggct gccgcagggg gacggctgcc ttccggggggg acggggcagg gcggggttcg 5640
gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat gccttcttct 5700
ttttcctaca gctcctgggc aacgtgctgg ttattgtgct gtctcatcat tttggcaaag 5760
aatcgcacc catggtgccc aagaagaaga ggaaagtctc tagactggac aagagcaaag 5820
tcataaactc tgctctggaa ttactcaatg gagtcggtat cgaaggcctg acgacaagga 5880
aactcgctca aaagctggga gttgagcagc ctacctgta ctggcacgtg aagaacaagc 5940
gggccctgct cgatgccctg ccaatcgaga tgctggacag gcatcatacc cactcctgcc 6000
ccctggaagg cgagtcattg caagaatttc tgcggaacaa cgccaagtca taccgctgtg 6060
ctctcctctc acatcgcgac ggggctaaaag tgcatctcgg caccgcccac acagagaaac 6120
agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggttc tccctggaga 6180
acgcactgta cgctctgtcc gccgtgggcc actttacact gggctgcgta ttggaggaaac 6240
aggagcatca agtagcaaaa gaggaaagag agacacctac caccgattct atgccccac 6300
ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc ttctttttcg 6360
gcctggaact aatcatatgt ggcctggaga aacagctaaa gtgcgaaagc ggcgggccga 6420
ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt gacgactttg 6480
accttgatat gctgcctgct gacgctcttg acgattttga cettgacatg ctccccgggt 6540
aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg ttttacttgc 6600
tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg caattgttgt 6660
tgttaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt 6720
cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt 6780
atcttaaggg atccctagag ggacagcccc ccccaaaagc cccagggat gtaattacgt 6840
ccctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggtc cggcgctccc 6900
cccgcattcc cgagccggca gcgtgcgggg acagcccggg cacggggaag gtggcacggg 6960
atcgctttcc tctgaacgct tctcgtgct ctttgagcct gcagacacct ggggggatac 7020
ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag aacggcctgg 7080
gttgcaaagg agcacagtgc tcattccagat ccaaccccc gctatgtgca ggtcatcaa 7140
ccagcagccc aggctgccc gagccacatc cagcctggcc ttgaatgcct gcagggatgg 7200
ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg ggggaaaaac 7260
tgctcctca tatccaaccc aaacctcccc tgtctcagt taaagccatt ccccttgtc 7320
ctatcaaggg ggagtttgct gtgacattgt tggctgggg tgacacatgt ttgccaatc 7380
agtgcattac ggagaggcag atcttgggga taaggaagtg caggacagca tggacgtggg 7440
acatgcaggt gttgagggt ctgggacact ctcaaagtea cagcgttcag aacagcetta 7500
aggataagaa gataggatag aaggacaaa agcaagttaa aaccagcat ggagaggagc 7560
acaaaaaggc cacagacact gctgttcct gtgtctgagc ctgcatgtt gatggtgtct 7620
ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg tcagtaggac 7680

```

```

tgggacaggc agctggagaa ttgcatgta gatgttcata caatcgtcaa atcatgaagg 7740
ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt gccactggc 7800
catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg gtgaccccc 7860
cacctccgtg ggcagctgtg cactgcagc accgctcttt ggagaaggta aatcttgcta 7920
aatccagccc gaccctcccc tggcacaacg taaggccatt atctctcatc caactccagg 7980
acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc aaagcccca 8040
gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg ggctccgctc 8100
cggtccggcg cccccccgc atccccgagc cggcagcgtg cggggacagc ccgggcacgg 8160
ggaaggtggc acgggatcgc tttcctctga acgttctcg ctgctctttg agcctgcaga 8220
cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga atgacagaat 8280
catagaacgg cctgggttgc aaaggagcac agtgcctcag cagatccaac cccctgctat 8340
gtgcagggtc atcaaccagc agcccaggt gccagagcc acatccagcc tggccttgaa 8400
tgctgcagg gatggggcat ccacagctc cttgggcaac ctgttcagtg cgtcaccacc 8460
ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct cagtgtaaag 8520
ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc tggggtgaca 8580
catgtttgcc aattcagtgc atcacggaga ggcagatctt ggggataagg aagtgcagga 8640
cagcatggac gtgggacatg caggtgttga ggcctctggg acactctcca agtcacagcg 8700
ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa gttaaaacc 8760
agcatggaga ggagcacaaa aaggccacag acactgctgg tccctgtgtc tgagcctgca 8820
tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctg agagatacag 8880
ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt tcatacaatc 8940
gtcaaatcat gaaggctgga aaagccctcc aagatcccca agaccaacc caaccaccc 9000
accgtgcca ctggccatgt cctcagtgc cacatcccca cagttcttca tcacctcag 9060
ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcacgc tctttggaga 9120
aggtaaatct tgctaaatcc agcccgacc tcccctggca caacgtaagg ccattatctc 9180
tcaccaact ccaggacgga gtcagtgagg atggggctca attgtttact ccctatcagt 9240
gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat gcagacttta 9300
ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga tagagaacgt 9360
atgaccagtt tactccctat cagtataga gaacgtatct acagtttact ccctatcagt 9420
gatagagaac gtatatccag ttactccct atcagtata gagaacgtat aagctttagg 9480
cgtgtacggt gggcgccctat aaaagcagag ctgctttagt gaaccgtcag atcgccctgga 9540
gcaattccac aacacttttg tcttatacca actttccgta ccacttcta cctcgtaaa 9600
aagcttgctc accatggctc ctaagaaaaa gcggaagggt gacaagaaat actcaatcgg 9660
gctggccatc ggaactaact cagtggggtg ggcagtcatt actgacgagt acaagtgcc 9720
aagcaagaaa ttaaggtcc tgggcaacac cgataggcac tccatcaaga aaaatctgat 9780
tggggccctg ctgttcgact ctggagagac agctgaagca actagactga aaaggactgc 9840
tagaaggcgc tatacccggc gaaagaatcg catctgtac ctgcaggaga ttttctctaa 9900
cgaaatggcc aaggtggacg atagtttctt tcacggctg gaggaatcat tcctggctga 9960
ggaagataag aaacacgaga gacatcctat ctttgaaac attgtggacg aggtcgctta 10020

```

```

tcacgaaaaa tacccaccca tctatcatct gcgcaagaaa ctggtggact ctacagataa 10080
agcagacctg cggctgatct atctggccct ggctcacatg attaatgtca gaggccattt 10140
tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt tcatccagct 10200
ggtccagaca tacaatcagc tgtttgagga aaaccctatt aatgcatctg gcgtggacgc 10260
aaaagccatc ctgagtgccg ggctgtctaa gagtagaagg ctggagaacc tgatcgctca 10320
gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcactgt cactgggact 10380
gacacctaac ttcaagagca attttgatct ggccgaggac gctaaactgc agctgagcaa 10440
ggacacttat gacgatgacc tggataacct gctggctcag atcggagatc agtacgcaga 10500
cctgttcctg gccgctaaga atctgtctga cgctatcctg ctgagtataa ttctgcgggt 10560
gaacaccgag attacaaaaa cccctctgtc agctagcatg atcaagagat atgacgagca 10620
ccatcaggat ctgaccctgc tgaaggcact ggtgcgccag cagctgcccg agaagtacaa 10680
ggaaatcttc tttgatcaga gtaagaacgg gtacgccggt tatattgacg gcggagcttc 10740
acaggaggaa ttctacaagt ttatcaaac tattctggag aagatggacg gcaccgagga 10800
actgctggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat ttgataacgg 10860
ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac gacaggaaga 10920
cttctacca tttctgaagg ataaccgcga gaagatcgaa aaaattctga ctttccggat 10980
cccttactat gtggggcccc tggcaagggg taattccgc tttgcctgga tgacacggaa 11040
atctgaggaa acaatcactc cttggaactt cgaggaagtg gtcgataagg gagcttccgc 11100
acagtctttc atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg agaaagtgtc 11160
gcctaagcac agtctgctgt acgagtattt cacagtctat aacgaactga ctaaggtgaa 11220
atacgtcacc gaggggatga ggaagccgc cttcctgagc ggtgaacaga agaaagctat 11280
cgtggacctg ctgtttaaaa ccaatcgcaa ggtgacagtc aagcagctga aggaggacta 11340
cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag acagatttaa 11400
cgcttctctg ggaacctacc acgatctgct gaagatcatt aaggataaag acttcctgga 11460
caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac tgtttgagga 11520
tcgcgaaatg atcgaggaac ggctgaaaac ttatgcccat ctgttcgatg acaaggtgat 11580
gaaacagctg aagcgaaaga ggtacaccgg ctggggacga ctgagcagaa agctgatcaa 11640
cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat cagatggctt 11700
cgctaacaga aattttatgc agctgattca cgatgacagc ctgacctca aagaggatat 11760
ccagaaggca caggtgtccg ggcagggtga ctctctgcac gagcatatcg caaacctggc 11820
cgggtcccc gccatcaaga aaggtattct gcagaccgtg aaggtggctg atgagctggt 11880
gaaagtcatg ggcaggcata agccagaaaa catcgtgatt gagatggccc gcgaaaatca 11940
gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg aggaaggcat 12000
taaggaactg ggatcccaga tcctgaaaga gcacctgtg gaaaacactc agctgcagaa 12060
tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg accaggagct 12120
ggatattaac cgactgtctg attacgacgt ggatgccatc gtcccacagt cattcctgaa 12180
agatgacagc attgacaata aggtgctgac ccggagtgc aaaaaccgag gaaagagtga 12240
taatgtccct tcagaggaag tggtaagaa aatgaagaac tactggagac agctgctgaa 12300
tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga gaggggtct 12360

```

gtcagaactg gacaaagcag gcttcatcaa gcgacagctg gtggagacca gacagatcac 12420
 aaagcacgtc gctcagattc tggatagcag gatgaacaca aagtacgatg agaatgacaa 12480
 actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg acttcagaaa 12540
 ggatttccag ttctacaaag tcagggagat caacaattat caccatgctc atgacgcata 12600
 cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg agagcgaatt 12660
 cgtgtacggg gactataagg tgtacgatgt cagaaaaatg atcgccaaga gtgagcagga 12720
 aattggaaaa gccaccgcta agtattttctt ttactcaaac atcatgaatt tctttaagac 12780
 tgagatcacc ctggcaaatg gggaaatccg aaagagacca ctgattgaga ctaacggcga 12840
 gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcgca aggtcctgtc 12900
 catgcctcaa gtgaatattg tcaagaaaac agaggtgcag actggcggat tcagtaagga 12960
 atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag actgggatcc 13020
 caagaaatat ggggggtttcg actccccaac agtggcttac tctgtcctgg tggtcgcaaa 13080
 ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg gtatcactat 13140
 tatggagagg agctccttcg agaagaaccc catcgatttt ctggaggcta aaggctataa 13200
 ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg agctggaaaa 13260
 cggaaggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg aactggccct 13320
 gccttctaag tacgtgaact tcctgtatct ggctagccac tacgagaagc tgaaaggctc 13380
 ccccgaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt atctggacga 13440
 gatcattgaa cagattagcg agttctccaa aagagtgate ctggctgacg caaatctgga 13500
 taaggctctg agcgcataca acaaacacag agataagcca atcaggagac aggccgaaaa 13560
 tatcattcat ctgttcactc tgaccaacct gggagccctc gcagccttca agtattttga 13620
 cactaccatc gatcggaaac gatacacatc cactaaggag gtgctggacg ctaccctgat 13680
 tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc tggggggcga 13740
 cctcgagccg aaaaagaaac gcaaagttag gcgcgccgac gcgctggacg atttcgatct 13800
 cgacatgctg ggcagcgacg ccctggatga cttcgacctg gatatgctgg gctctgatgc 13860
 cctggacgac tttagcttgg acatgttggg atccgacgct ctcgatgatt ttgaccttga 13920
 catgctgate aacggcagcg gcgagggcag aggcagcctg ctaacatgcg gtgacgtcga 13980
 ggagaatcct ggcccagcac cgggatccat ggtgagcaag ggcgaggagc tgttcaccgg 14040
 ggtggtgccc atcctggctg agctggacgg cgacgtaaac ggccacaagt tcagcgtgtc 14100
 cggcgagggc gagggcgatg ccacctacgg caagctgacc ctgaagttca tctgcaccac 14160
 cggcaagctg cccgtgccct ggcccacct cgtgaccacc ttcacctacg gcgtgcagtg 14220
 cttcgcccg c taccgacg acatgaagca gcacgaactt ttcaagtccg ccatgcccga 14280
 aggctacgtc caggagcgca ccatcttctt caaggacgac ggcaactaca agaccgcgc 14340
 cgaggtgaag ttcgagggcg acacctggt gaaccgcac gagctgaagg gcatcgaact 14400
 caaggaggac ggcaacatcc tggggcacia gctggagtac aactacaaca gccacaaggt 14460
 ctatatcacc gccgacaagc agaagaacgg catcaaggtg aacttcaaga cccgccacia 14520
 catcgaggac ggcagcgtgc agctcgccga ccactaccag cagaacaccc ccatcggcga 14580
 cggccccgtg ctgtgcccc acaaccacta cctgagcacc cagtccgcc tgagcaaaga 14640
 cccaacgag aagcgcgatc acatggtcct gctggagttc gtgaccgccg ccgggatcac 14700

```

tctcggcatg gacgagctgt acaagtaaac ctaatctagc agctcgtga tcagcctcga 14760
ctgtgccttc tagttgccag ccatctgttg tttgccctc ccccggtgcct tccttgacct 14820
tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc 14880
tgagtaggtg tcattctatt ctggggggtg ggggtggggca ggacagcaag ggggaggatt 14940
gggaagacaa tagcaggcat gctggggatg cggtgggctc tatggcttct gaggcggaaa 15000
gaaccagctg gggctcgtac ctctagttag cgctcatgg tccatatgaa tatectcctt 15060
agttcctatt ccgctagcct agagggacag cccccccca aagccccag ggatgtaatt 15120
acgtccctcc cccgctaggg gcagcagcga gccgcccggg gctccgctcc ggtccggcgc 15180
tcccccgca tccccgagcc ggcagcgtgc ggggacagcc cgggcacggg gaaggtggca 15240
cgggatcgct ttctctgaa cgcttctcgc tgctctttga gcctgcagac acctgggggg 15300
atacggggaa aaagctttag gctgaaagag agatttagaa tgacagaatc atagaacggc 15360
ctgggttgca aaggagcaca gtgetcatcc agatccaacc cctgctatg tgcagggtca 15420
tcaaccagca gccaggtg cccagagcca catccagcct ggccttgaat gcctgcaggg 15480
atggggcatc cacagcctcc ttgggcaacc tgttcagtgc gtcaccacc tctgggggaa 15540
aaactgcctc ctcatatcca acccaaacct cccctgtctc agtgtaaagc cattccccct 15600
tgtcctatca agggggagtt tgctgtgaca ttgttggtct ggggtgacac atgtttgcca 15660
attcagtgca tcacggagag gcagatcttg gggataagga agtgcaggac agcatggacg 15720
tgggacatgc aggtgttag ggctctggga cactctccaa gtcacagcgt tcagaacagc 15780
cttaaggata agaagatagg atagaaggac aaagagcaag ttaaaacca gcatggagag 15840
gagcacaaaa aggccacaga cactgctggt ccctgtgtct gagcctgcat gtttgatggt 15900
gtctggatgc aagcagaagg ggtggaagag cttgcctgga gagatacagc tgggtcagta 15960
ggactgggac aggcagctgg agaattgcca ttagatgtt catacaatcg tcaaatcatg 16020
aaggctggaa aagccctcca agatcccaa gaccaaccc aaccaccca ccgtgcccac 16080
tggccatgtc cctcagtgc acatccccac agttcttcat cacctccagg gacggtgacc 16140
ccccacctc cgtgggcagc tgtgccactg cagcacgct ctttgagaa ggtaaattctt 16200
gctaaatcca gcccgaccct cccctggcac aacgtaaggc cattatctct catccaactc 16260
caggacggag tcagtgagga tggggctgga tccgaagcag ctccagccta cacaatcgct 16320
caagacgtgt aatgctttta ttatatatta gtcacgatat ctataacaag aaaatatata 16380
tataataagt tatcacgtaa gtagaacatg aaataacaat ataattatcg tatgagttaa 16440
atcttaaaaag tcacgtaaaa gataatcatg cgctattttg actcacggg tcgttatagt 16500
tcaaaatcag tgacacttac cgcattgaca agcacgctc acgggagctc caagcggcga 16560
ctgagatgtc ctaaatgcac agcgacggat tcgcgtatt tagaaagaga gagcaatatt 16620
tcaagaatgc atgcgtcaat ttacgcaga ctatctttct agggttaaaa aagatttgcg 16680
ctttactcga cctaaacttt aaacacgtca tagaatcttc gtttgacaaa aaccacattg 16740
tggggtaccg agctcttaat taaggcgcgc cggggagggt ccctttagtg agggttaatt 16800
gcgggtcgcc ctatagttag tcgtattaca attcaactgc cgtcgtttta caacgtcgtg 16860
actgggaaaa ccctggcggt acccaactta atcgccctgc agcacatccc ctttcgcca 16920
gctggcgtaa tagcgaagag gccgcaccg atcgccctc ccaacagttg cgcagcctga 16980
atggcgaatg gcaaattgta agcgtaata ttttgtaaa attcgcgtta aatttttgtt 17040

```

aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag 17100
 aatagaccga gatagggttg agtggtgttc cagtttggaa caagagtcca ctattaaaga 17160
 acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 17220
 aaccatcacc ctaatcaagt tttttggggt cgagggtgccg taaagcacta aatcggaacc 17280
 ctaaaggagg ccccccattt agagcttgac ggggaaagcc ggcgaaacgtg gcgagaaagg 17340
 aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc 17400
 gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgta g 17451

<210> 17

<211> 18381

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 17

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
 caaatatgta tccgctcatg agacaataac cctgataaat gtttcaataa tattgaaaaa 120
 ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
 gccttctctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggttaagatc cttgagagtt 300
 ttcccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
 tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
 atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
 gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
 ctccgcttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
 ccacgatgcc tntagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
 ctctagcttc ccggcaacaa ttaatagact ggatggagcg ggataaagtt gcaggaccac 780
 ttctgcgctc ggcccttccg gctggctggg ttattgctga taaatctgga gccggtgagc 840
 gtgggtctcg cggatatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
 agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
 atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
 aaaagatcaa aggatcttct tgagatcett tttttctgcg cgtaatctgc tgccttcaaa 1200
 caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
 ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
 cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 1380
 tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
 gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500

```

ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctctgcaggg gggcggagcc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg 1800
ctcacatgtt ctttcctgcg ttatccccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc cgacagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttggccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctcccct agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgttaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaa aaaaactcaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat ttgtactttt aaaaacagtc atgttgattt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaagggtt ccgcatcagc ttgattcgag cccagctggg ttctttccgc 2640
ctcagaagcc atagagccca ccgcattccc agcatgcctg ctattgtett cccaatctc 2700
ccccctgctg tctgccccca cccaccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tcttcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcgaggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
ggcgatgcgc tgcaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccacatgat 3120
attcggcaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
cttgagcctg gcgaacagtt cggtcggcgc gagccccctga tgctcttcgt ccagatcatc 3240
ctgatcgaca agaccgggtt ccatccgagt acgtgctcgc tcgatcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccgcatcaag cgtatgcagc cgccgcattg catcagccat 3360
gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtccc tttccgttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgtgcctcg tctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcgcatca gagcagccga ttgtctgttg tgcccagtea tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720
gttctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
acgtgctgct gcaggtcgaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840

```



```

tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcgggcg cccgccccgc 3900
ccctgagccc gcccttgagc ccgccccggg acccaccctc tcccagcctc tgagcccaga 3960
aagcgaagga gcaaagctgc tattggccgc tgccccaaag gcctaccgcg ttccattgct 4020
cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttccatt tgtcacgtcc 4080
tgcacgacgc gagctgcggg gcggggggga acttcctgac taggggagga gtagaaggtg 4140
gcgcgaaggg gccaccaaag aacggagccg gttggcgccct accggtggat gtggaatgtg 4200
tgcgagccag aggccacttg tgtagcgcca agtgcccagc ggggctgcta aagcgcatgc 4260
tccagactgc cttgggaaaa gcgcctcccc taccggtag acaccccaca gtgggtggcc 4320
tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
tagtaatcaa ttacggggtc attagttcat agcccatata tggagtcccg cgttacataa 4440
cttacggtaa atggccccgc tggtgaccg cccaacgacc cccgcccatt gacgtcaata 4500
atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
tattttacgt aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgcc 4620
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 4680
tgggactttc ctacttgga gtacatctac gtattagtea tcgctattac catgggtcga 4740
ggtgagcccc acgttctgct tcactctccc catctcccc cctccccac cccaatttt 4800
gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg ggggggcgcg 4860
cgccaggcgg ggccggggcg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg 4920
cagccaatca gagcggcgcg ctccgaaagt ttcttttat ggcgaggcgg cggcggcggc 4980
ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcc tgcccccgtg 5040
ccccgctccg cgccgcctcg cgccgccgcg cccggtctcg actgaccgcg ttactcccac 5100
aggtgagcgg gcgggacggc cttctcctc cgggctgtaa ttagcgcttg gtttaatgac 5160
ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
gcggggggga gcggctcggg ggggtgcgtg gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
ggccccgcgt gcccgcggcg tgtgagcgct gcgggcgcgg cgcggggctt tgtgcgtcc 5340
gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg gggctgcgag 5400
gggaacaaag gctgcgtgcg ggggtgtgtc gtgggggggt gagcaggggg tgtgggcgcg 5460
gcggtcgggc tgtaaccccc cctgcaccc cctccccga gttgctgagc acggccccgc 5520
ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg gcgggggggtg 5580
gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc 5700
cttttatggg aatcgtgcga gagggcgag ggaacttcct tgtcccaa tctggcgagc 5760
cgaaatctgg gaggcgccgc cgcacccct ctagcgggcg cgggcgaagc ggtgcggcgc 5820
cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgc gtecccttct 5880
ccatctccag cctcggggct gccgcagggg gacggtgcc ttcggggggg acggggcagg 5940
gcgggggttc gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
gccttcttct ttttctaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
tttgccaaag aattcgccac catggtgcc aagaagaaga ggaaagtctc tagactggac 6120
aagagcaaag tcataaactc tgctctggaa ttactcaatg gagtcggtat cgaaggcctg 6180

```

acgacaagga aactcgctca aaagctggga gttgagcagc ctaccctgta ctggcacgtg 6240
 aagaacaagc gggccctgct cgatgccctg ccaatcgaga tgctggacag gcatcatacc 6300
 cactcctgcc ccctggaagg cgagtcattg caagactttc tgcggaacaa cgccaagtca 6360
 taccgctgtg ctctcctctc acatcgcgac ggggctaaag tgcattctcg caccgcgcca 6420
 acagagaaac agtacgaaac cctggaaaat cagctcgctt tcctgtgtca gcaaggcttc 6480
 tccttgagga acgcaactgt cgctctgtcc gccgtgggcc actttacact gggctgcgta 6540
 ttggagggaac aggagcatca agtagcaaaa gaggaaagag agacacctac caccgattct 6600
 atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
 ttctttttcg gcctggaact aatcatatgt ggcttgaga aacagctaaa gtgcgaaagc 6720
 ggcgggccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
 gacgactttg accttgatat gctgcctgct gacgctcttg acgattttga cttgacatg 6840
 ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
 ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
 caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
 tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7080
 tcatcaatgt atcttaaggg atccctagag ggacagcccc cccccaagc cccagggat 7140
 gtaattacgt cctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggctc 7200
 cggcgctccc ccgcacccc cgagccggca gcgtgcgggg acagcccggg cagggggaag 7260
 gtggcacggg atcgctttcc tctgaacgct tctcgctgct ctttgagcct gcagacacct 7320
 ggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag 7380
 aacggccttg gttgcaaagg agcacagtgc tcatccagat ccaacccctt gctatgtgca 7440
 gggatcatca ccagcagccc aggctgccc gagccacatc cagcctggcc ttgaatgcct 7500
 gcagggatgg ggcattccca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg 7560
 ggggaaaaac tgcctcctca tatccaacc aaacctccc tgtctcagtg taaagccatt 7620
 ccccttgtc ctatcaaggg ggagtttctg gtgacattgt tggctctggg tgacacatgt 7680
 ttgccaattc agtgcatcac ggagaggcag atcttgggga taaggagtg caggacagca 7740
 tggacgtggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
 aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
 ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
 gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
 tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
 atcatgaagg ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt 8100
 gccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
 gtgaccccc cacctccgtg ggcagctgtg ccaactgcag accgctcttt ggagaaggta 8220
 aatcttgcta aatccagccc gacctcccc tggcacaacg taaggccatt atctctcacc 8280
 caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
 aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
 ggctccgctc cggctccggc ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
 ccgggcacgg ggaagggtgg acgggatcgc tttcctctga acgtttctcg ctgctctttg 8520

agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgctcatc cagatccaac 8640
 cccctgctat gtgcagggtc atcaaccagc agcccaggct gccagagcc acatccagcc 8700
 tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
 cagtgtaaag ccattccccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
 tggggtgaca catgtttgcc aattcagtg atcacggaga ggcagatctt ggggataagg 8940
 aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
 gttaaaaccc agcatggaga ggagcacaaa aaggccacag aactgtctgg tccctgtgtc 9120
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctgg 9180
 agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
 tcatacaatc gtcaaatcat gaaggetgga aaagccctcc aagatcccca agaccaaccc 9300
 caaccacccc accgtgccc a ctggccatgt ccctcagtg c acatcccca cagttcttca 9360
 tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
 tctttggaga aggtaaatct tgctaaatcc agcccgaccc tcccctggca caacgtaagg 9480
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
 ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
 gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
 tagagaacgt atgaccagtt tactccctat cagtataga gaacgtatct acagtttact 9720
 ccctatcagt gatagagaac gtatatccag ttactccct atcagtata gagaacgtat 9780
 aagctttagg cgtgtacggt gggcgcctat aaaagcagag ctcgtttagt gaaccgtcag 9840
 atcgccctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttccta 9900
 cccctgtaaa aagcttgtcc actcgagatt ctctagacat cattaattcc taatttttgt 9960
 tgacactcta tcattgatag agttatttta ccactcccta tcagtgatag agaaaagtga 10020
 aatggccaag cctttgtctc aagaagaatc caccctcatt gaaagagcaa cggctacaat 10080
 caacagcatc cccatctctg aagactacag cgtcgccagc gcagctctct ctagcgacgg 10140
 ccgcatcttc actggtgtca atgtatatca tttactggg ggaccttggt cagaactcgt 10200
 ggtgctgggc actgctgtg ctgcggcagc tggcaacctg acttgatctg tcgcgatcgg 10260
 aaatgagaac aggggcatct tgagccctg cggacggtgt cgacaggtgc ttctcgatct 10320
 gcacccctggg atcaaaagcga tagtgaagga cagtgtatga cagccgacgg cagttgggat 10380
 tcgtgaattg ctgccctctg gttatgtgtg ggagggctaa ctcgagatga gctcagagac 10440
 tggcccagtg gctgtggacc ccacattgag acggcggatc gagcccatg agtttgaggt 10500
 attcttcgat ccgagagagc tccgcaagga gacctgcctg ctttacgaaa ttaattgggg 10560
 gggccggcac tccatttggc gacatacatc acagaacact aacaagcacg tcgaagtcaa 10620
 cttcatcgag aagttcacga cagaaagata tttctgtccg aacacaaggt gcagcattac 10680
 ctggtttctc agctggagcc catgcggcga atgtagtagg gccatcactg aattcctgtc 10740
 aaggtatccc cacgtcactc tgtttattta catcgcaagg ctgtaccacc acgtgaccc 10800
 ccgcaatcga caaggcctgc gggatttgat ctcttcaggt gtgactatcc aaattatgac 10860

tgagcaggag	tcaggatact	gctggagaaa	ctttgtgaat	tatagcccga	gtaatgaagc	10920
ccactggcct	aggatatccc	atctgtgggt	acgactgtac	gttcttgaac	tgtactgcat	10980
catactgggc	ctgcctcctt	gtctcaacat	tctgagaagg	aagcagccac	agctgacatt	11040
ctttaccatc	gctcttcagt	cttgtcatta	ccagcgactg	ccccacaca	ttctctgggc	11100
caccgggttg	aaaagcggca	gcgagactcc	cgggacctca	gagtccgcca	cacccgaaag	11160
tgataaaaaag	tattctattg	gtttagccat	cggcactaat	tccgttggat	gggctgtcat	11220
aaccgatgaa	tacaaagtac	cttcaaagaa	atttaaggtg	ttggggaaca	cagaccgtca	11280
ttcgattaaa	aagaatctta	tcggtgccct	cctattcgat	agtggcgaaa	cggcagaggc	11340
gactcgcctg	aaacgaaccg	ctcggagaag	gtatacacgt	cgcaagaacc	gaatatgtta	11400
cttacaagaa	atttttagca	atgagatggc	caaagttgac	gattctttct	ttcacgttt	11460
ggaagagtcc	ttccttgtcg	aagaggacaa	gaaacatgaa	cggcacccca	tctttggaaa	11520
catagtagat	gaggtagcat	atcatgaaaa	gtacccaacg	atttatcacc	tcagaaaaaa	11580
gctagttagac	tcaactgata	aagcggacct	gaggttaatc	tacttggctc	ttgcccatat	11640
gataaagttc	cgtgggcact	ttctcattga	gggtgatcta	aatccggaca	actcgatgt	11700
cgacaaactg	ttcatccagt	tagtacaac	ctataatcag	ttgtttgaag	agaaccttat	11760
aaatgcaagt	ggcgtggatg	cgaaggetat	tcttagcgcc	cgctctctta	aatcccgacg	11820
gctagaaaac	ctgatcgcac	aattacccgg	agagaagaaa	aatgggttgt	tcggtaacct	11880
tatagcgctc	tcactaggcc	tgacaccaaa	ttttaagtcg	aacttcgact	tagctgaaga	11940
tgccaaattg	cagcttagta	aggacacgta	cgatgacgat	ctcgacaatc	tactggcaca	12000
aattggagat	cagtatgcgg	acttatTTTT	ggctgccaaa	aaccttagcg	atgcaatcct	12060
cctatctgac	atactgagag	ttaatactga	gattaccaag	gcgcggttat	ccgcttcaat	12120
gatcaaaagg	tacgatgaac	atcaccaaga	cttgacactt	ctcaaggccc	tagtccgtca	12180
gcaactgcct	gagaaatata	aggaaatatt	ctttgatcag	tcgaaaaacg	ggtacgcagg	12240
ttatattgac	ggcggagcga	gtcaagagga	attctacaag	tttatcaaac	ccatattaga	12300
gaagatggat	gggacggaag	agttgcttgt	aaaactcaat	cgcgaaagatc	tactgcgaaa	12360
gcagcggact	ttcgacaacg	gtagcattcc	acatcaaate	cacttaggcg	aattgcatgc	12420
tatacttaga	aggcaggagg	atttttatcc	gttcttcaaa	gacaatcgtg	aaaagattga	12480
gaaaatccta	acctttcgca	taccttacta	tgtgggaccc	ctggccccgag	ggaactctcg	12540
gttcgcatgg	atgacaagaa	agtccgaaga	aacgattact	ccatggaatt	ttgaggaagt	12600
tgtcgataaa	ggtgcgtcag	ctcaatcgtt	catcgagagg	atgaccaact	ttgacaagaa	12660
tttaccgaac	gaaaaagtat	tgcctaagca	cagtttactt	tacgagtatt	tcacagtgt	12720
caatgaactc	acgaaaagta	agtatgtcac	tgagggcatg	cgtaaaccgg	cctttctaag	12780
cggagaacag	aagaaaagcaa	tagtagatct	gttattcaag	accaaccgca	aagtgcagct	12840
taagcaattg	aaagaggact	actttaagaa	aattgaatgc	ttcgattctg	tcgagatctc	12900
cggggtagaa	gatcgattta	atgcgtcact	tggtacgtat	catgacctcc	taaagataat	12960
taaagataag	gacttcctgg	ataacgaaga	gaatgaagat	atcttagaag	atatagtgtt	13020
gactcttacc	ctctttgaag	atcgggaaat	gattgaggaa	agactaaaaa	catacgtca	13080
cctgttcgac	gataaggta	tgaacagtt	aaagaggcgt	cgctatacgg	gctggggacg	13140
attgtcgcgg	aaacttatca	acgggataag	agacaagcaa	agtggtaaaa	ctattctcga	13200

```

ttttctaaag agcgacggct tcgccaatag gaactttatg cagctgatcc atgatgactc 13260
ttaaaccctc aaagaggata tacaaaaggc acaggtttcc ggacaagggg actcattgca 13320
cgaacatatt gcgaatcttg ctggttcgcc agccatcaaa aagggcatac tccagacagt 13380
caaagtagtg gatgagctag ttaaggtcat gggacgtcac aaaccgaaa acattgtaat 13440
cgagatggca cgcgaaaatc aaacgactca gaagggggcaa aaaaacagtc gagagcggat 13500
gaagagaata gaagagggtg ttaaagaact gggcagccag atcttaaagg agcatcctgt 13560
ggaaaatacc caattgcaga acgagaaact ttacctctat tacctacaaa atggaaggga 13620
catgtatggt gatcaggaac tggacataaa ccgtttatct gattacgacg tcgatccat 13680
tgtaccccaa tcctttttga aggacgattc aatcgacaat aaagtgttta cacgctcgga 13740
taagaaccga gggaaaagtg acaatgttcc aagcgaggaa gtcgtaaaga aaatgaagaa 13800
ctattggcgg cagctcctaa atgcgaaact gataacgcaa agaaagtctg ataacttaac 13860
taaagctgag aggggtggct tgtctgaact tgacaaggcc ggatttatta aacgtcagct 13920
cgtggaaacc cgccaaatca caaagcatgt tgcacagata ctagattccc gaatgaatac 13980
gaaatacgac gagaacgata agctgattcg ggaagtcaaa gtaatcactt taaagtcaaa 14040
attggtgtcg gacttcagaa aggattttca attctataaa gttagggaga taaataacta 14100
ccaccatgcg cacgacgctt atcttaatgc cgtcgtaggg accgcactca ttaagaaata 14160
cccgaagcta gaaagtgagt ttgtgtatgg tgattacaaa gtttatgacg tccgtaagat 14220
gatcgcgaaa agcgaacagg agataggcaa ggctacagcc aaatacttct tttattctaa 14280
cattatgaat ttctttaaga cggaatcac tctggcaaac ggagagatac gcaaacgacc 14340
tttaattgaa accaatgggg agacaggtga aatcgtatgg gataagggcc gggacttcgc 14400
gacggtgaga aaagttttgt ccatgcccc agtcaacata gtaaagaaaa ctgaggtgca 14460
gaccggaggg ttttcaaagg aatcgattct tccaaaaagg aatagtata agctcatcgc 14520
tcgtaaaaag gactgggacc cgaaaaagta cgggtggcttc gatagcccta cagttgccta 14580
ttctgtccta gtagtggcaa aagttgagaa gggaaaatcc aagaaactga agtcagtcaa 14640
agaattattg gggataacga ttatggagcg ctctctttt gaaaagaacc ccatcgactt 14700
ccttgaggcg aaaggttaca aggaagtaaa aaaggatctc ataattaaac taccaaagta 14760
tagtctgttt gagttagaaa atggccgaaa acggatgttg gctagcgccg gagagcttca 14820
aaaggggaac gaactcgcac taccgtctaa atacgtgaat ttctgtatt tagcgtccca 14880
ttacgagaag ttgaaagggt cactgaaga taacgaacag aagcaacttt ttgttgagca 14940
gcacaaacat tatctcgacg aaatcataga gcaaatttcg gaattcagta agagagtcac 15000
cctagctgat gccaatctgg acaaagtatt aagcgcatac aacaagcaca gggataaacc 15060
catacgtgag caggcggaat atattatcca tttgtttact cttaccaacc tcggcgctcc 15120
agccgcattc aagtattttg acacaacgat agatcgcaaa cgatacactt ctaccaagga 15180
ggtgctagac gcgacactga ttcaccaatc catcacggga ttatatgaaa ctcggataga 15240
tttgtcacag cttgggggtg actctggttg ttctactaat ctgtcagata ttattgaaaa 15300
ggagaccggt aagcaactgg ttatccagga atccatctc atgctcccag aggaggtgga 15360
agaagtcatt gggaacaagc cggaaagcga tatactcgtg cacaccgct acgacgagag 15420
caccgacgag aatgtcatgc ttctgactag cgacgccct gaatacaagc cttgggctct 15480
ggtcatacag gatagcaacg gtgagaacaa gattaagatg ctctctggtg gttctcccaa 15540

```

```

gaagaagagg aaagtctaaa aattctaaaa tacagcatag caaaacttta acctccaaat 15600
caagcctcta cttgaatcct tttctgaggg atgaataagg catagggcatc aggggctgtt 15660
gccaatgtgc attagctgtt tgcagcctca ctttctttca tggagtttaa gatatagtgt 15720
attttcccaa ggtttgaact agctcttcat ttctttatgt tttaaatgca ctgacctccc 15780
acattccctt tttagtaaaa tattcagaaa taattttaat acatcattgc aatgaaaata 15840
aatgtttttt attaggcaga atccagatgc tcaaggccct tcataatata ccccagttta 15900
gtagttaggac ttagggaaca aaggaacctt taatagaaat tggacagcaa gaaagcgagc 15960
ttctagatgg tccatatgaa taccctcctt agttcctatt ccgctagcct agagggacag 16020
ccccccccca aagccccccag ggatgttaatt acgtccctcc cccgctaggg gcagcagcga 16080
gccgcccggg gctccgctcc ggtccggcgc tcccccgca tccccgagcc ggcagcgtgc 16140
ggggacagcc cgggcacggg gaaggtggca cgggatcgct ttctctgaa cgcttctcgc 16200
tgctctttga gcctgcagac acctggggggg atacggggaa aaagctttag gctgaaagag 16260
agatttagaa tgacagaatc atagaacggc ctgggttgca aaggagcaca gtgctcatcc 16320
agatccaacc ccctgctatg tgcagggtca tcaaccagca gcccagctg cccagagcca 16380
catccagcct ggcttgaat gcctgcaggg atggggcatc cacagcctcc ttgggcaacc 16440
tgttcagtgc gtcaccaccc tctgggggaa aaactgcctc ctcatatcca acccaaacct 16500
cccctgtctc agtgtaaagc cattccccct tgctctatca agggggagtt tgctgtgaca 16560
ttgttggctc ggggtgacac atgtttgcca attcagtga tcacggagag gcagatcttg 16620
gggataagga agtgcaggac agcatggacg tgggacatgc aggtgttgag ggctctggga 16680
cactctccaa gtcacagcgt tcagaacagc ctttaaggata agaagatagg atagaaggac 16740
aaagagcaag ttaaaaccca gcatggagag gagcacaana aggccacaga cactgctggt 16800
ccctgtgtct gagcctgcat gtttgatggt gtctggatgc aagcagaagg ggtggaagag 16860
cttgccctgga gagatacagc tgggtcagta ggactgggac aggcagctgg agaattgcca 16920
tgtagatgtt catacaatcg tcaaatcatg aaggctggaa aagccctcca agatcccca 16980
gaccaacccc aaccaccca ccgtgccac tggccatgtc cctcagtgcc acatccccac 17040
agttcttcat cacctccagg gacggtgacc cccccacctc cgtgggcagc tgtgccactg 17100
cagcaccgct ctttgagaaa ggtaaatctt gctaaatcca gcccgacct cccctggcac 17160
aacgtaaggc cattatctct catccaactc caggacggag tcagtgagga tggggctgga 17220
tccgaagcag ctccagccta cacaatcgct caagacgtgt aatgctttta ttatatatta 17280
gtcacgatat ctataacaag aaaatatata tataataagt taccacgtaa gtagaacatg 17340
aaataacaat ataattatcg tatgagttaa atctttaaag tcacgtaaaa gataatcatg 17400
cgtcattttg actcacgcgg tcgttatagt tcaaaatcag tgacacttac cgcattgaca 17460
agcacgcctc acgggagctc caagcggcga ctgagatgtc ctaaatgcac agcgacggat 17520
tcgcgctatt tagaaagaga gagcaatatt tcaagaatgc atgcgtcaat ttacgcaga 17580
ctatctttct agggttaaaa aagatttgcg ctttactcga cctaaacttt aaacacgtca 17640
tagaatcttc gtttgacaaa aaccacattg tggggtaccg agctcttaat taaggegcgc 17700
cggggagggt ccctttagtg agggttaatt gcgggtcgcc ctatagttag tcgtattaca 17760
attcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcggt acccaactta 17820
atcgccctgc agcacatccc ctttcgcca gctggcgtaa tagcgaagag gcccgcaccg 17880

```

atcgcccttc ccaacagttg cgcagcctga atggcgaaatg gcaaattgta agcgттаата 17940
ttttgttaaa attcgcgtta aattttttgtt aaatcagctc attttttaac caataggccg 18000
aaatcggcaa aatcccttat aaatcaaaaag aatagaccga gatagggttg agtgттгттс 18060
cagttttgaa caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa 18120
ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt 18180
cgagggtgccg taaagcacta aatcggaacc ctaaaggag ccccgattt agagcttgac 18240
ggggaaaagcc ggcgaacgtg gcgagaaaagg aagggaagaa agcgaaaagga gcgggcgcta 18300
gggcgctggc aagtgtagcg gtcacgtgc gcgtaaccac cacaccgcc gcgcttaatg 18360
cgccgctaca gggcgcgta g 18381

<210> 18

<211> 19822

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 18

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
gccttctctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tggtgtcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
tattatccccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttgggт tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgcccttgа tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttccg gctggctggт ttattgtgta taaatctgga gccggtgagc 840
gtgggtctcg cggтatcatt gcagcactgg ggccagatgg таgccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct таacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcett tttttctgcg cgtaatctgc tгcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 1380

```

tctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatatgtt accggataag gcgagcggt cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat tttgtgatg ctcgtcaggg gggcggagcc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg 1800
ctcacatgtt ctttcctgcg ttatccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgggcg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcacccagc ctttacactt tatgttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctccctt agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgtaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaa aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat ttgtactttt aaaaacagtc atgttgtatt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaaggggt ccgcatcagc ttgattcgag cccagctgg ttctttccgc 2640
ctcagaagcc atagagccca ccgcatcccc agcatgcctg ctattgtctt cccaatcctc 2700
ccccttgctg tctgccccca cccaccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tctcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcaggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
ggcgatgcgc tgcaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
ttcgccgcca agctcttcag caatatcagc ggtagccaac gctatgtcct gatagcggtc 3060
cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccacatgat 3120
attcggcaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgcgcgc 3180
cttgagcctg gcgaacagtt cggtggcgc gageccctga tgctcttcgt ccagatcatc 3240
ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgcattg catcagccat 3360
gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtccc tccccgttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgtgcctcg tctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcggcatca gagcagccga ttgtctgttg tgcccagtc tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720

```



```

gttctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
acgtgctgct gcaggctcaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840
tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcgggcg cccgccccgc 3900
ccctgagccc gccctgagc ccgcccccg acccaccct tcccagcctc tgagcccaga 3960
aagcgaagga gcaaagctgc tattggccgc tgcccaaag gcctaccgc ttccattgct 4020
cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttccatt tgtcacgtcc 4080
tgcacgacgc gagctgcggg gcggggggga acttcctgac taggggagga gtagaaggtg 4140
gcgcgaaggg gccaccaaag aacggagccg gttggcgccct accggtggat gtggaatgtg 4200
tgcgagccag aggccacttg tgtagcgcca agtgcccagc ggggctgcta aagcgcatgc 4260
tccagactgc cttgggaaaa gcgcctcccc taccggtag acaccacaca gtgggtggcc 4320
tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
tagtaatcaa ttacggggtc attagttcat agcccatata tggagtccg cgttacataa 4440
cttacgtaa atggcccgc tggtgaccg cccaacgacc cccgccatt gacgtcaata 4500
atgacgtatg tccccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
tatttacggg aaactgccc cttggcagta catcaagtgt atcatatgcc aagtacgcc 4620
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 4680
tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac catgggtcga 4740
ggtgagcccc acgttctgct tcactctccc catctcccc cctcccccac cccaatttt 4800
gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg ggggggcgcg 4860
cgccaggcgg gcgggggcgg gcgagggggc ggggcggggc gaggcggaga ggtgcggcgg 4920
cagccaatca gagcggcgcg ctccgaaagt ttcttttat ggcgaggcgg cggcggcggc 4980
ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcc tgcgccgtg 5040
ccccgctccg cgccgcctcg cgccggcgcc cccggtctg actgaccgcg ttactccac 5100
aggtgagcgg gcgggacggc cttctcctc cgggctgtaa ttagcgcttg gtttaatgac 5160
ggctcgtttc tttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
gcggggggga gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
ggcccgcgct gcccgcggc tgtgagcgt gcgggcgcgg cgcggggctt tgtgcgtcc 5340
gcgtgtgcgc gaggggagcg cggccggggg cggtgcccc cggtgcgggg gggctgcgag 5400
gggaacaaag gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg 5460
gcggtcgggc tgtaaccccc cctgcaccc cctccccga gttgctgagc acggcccggc 5520
ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg gcggggggtg 5580
gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc 5700
cttttatggg aatcgtgcga gagggcgag ggaattcctt tgtcccaaat ctggcgagc 5760
cgaaatctgg gaggcgccgc cgcacccct ctagcgggcg cgggcgaagc ggtgcggcgc 5820
cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgc gteccctct 5880
ccatctccag cctcggggct gccgcagggg gacggctgcc ttcggggggg acggggcagg 5940
gcgggggtcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
gccttctctt ttttctaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060

```

```

tttggcaaag aattcgccac catggtgccc aagaagaaga ggaaagtctc tagactggac 6120
aagagcaaag tcataaactc tgctctggaa ttactcaatg gagtcggtat cgaaggcctg 6180
acgacaagga aactcgctca aaagctggga gttgagcagc ctaccctgta ctggcacgtg 6240
aagaacaagc gggccctgct cgatgccctg ccaatcgaga tgctggacag gcatcatacc 6300
cactcctgcc ccctggaagg cgagtcattg caagactttc tgcggaacaa cgccaagtca 6360
taccgctgtg ctctcctctc acatcgcgac ggggctaaag tgcattctcg caccgcacca 6420
acagagaaac agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggcttc 6480
tccctggaga acgcactgta cgctctgtcc gccgtgggcc actttacact gggctgcgta 6540
ttggaggaaac aggagcatca agtagcaaaa gaggaagag agacacctac caccgattct 6600
atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
ttccttttcg gcctggaact aatcatatgt ggctggaga aacagctaaa gtgcgaaagc 6720
ggcgggccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
gacgactttg accttgatat gctgcctget gacgetcttg acgattttga cttgacatg 6840
ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7080
tcatcaatgt atcttaaggg atccctagag ggacagcccc ccccaaagc cccagggat 7140
gtaattacgt cctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggctc 7200
cggcgtccc cccgcatccc cgagccggca gcgtgcgggg acagcccggg caccgggaag 7260
gtggcacggg atcgctttcc tctgaacgct tctcgtctgt ctttgagcct gcagacacct 7320
gggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag 7380
aacggcctgg gttgcaaagg agcacagtgc tcatccagat ccaaccccct gctatgtgca 7440
gggtcatcaa ccagcagccc aggctgccc gagccacatc cagcctggcc ttgaatgcct 7500
gcagggatgg ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg 7560
ggggaaaaaac tgctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
cccccttgtc ctatcaaggg ggagtttget gtgacattgt tggctctggg tgacacatgt 7680
ttgccaattc agtgcacac ggagaggcag atcttgggga taaggaagtg caggacagca 7740
tggaactggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
atcatgaagg ctggaaaagc cctccaagat cccaagacc aacccaacc caccaccgt 8100
gccaactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
gtgaccccc caccctcgtg ggcagctgtg ccaactgcag accgctcttt ggagaaggta 8220
aatcttgcta aatccagccc gaccctcccc tggcacaacg taaggccatt atctctcacc 8280
caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
aaagcccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400

```

ggctccgctc cgggtccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
 ccgggcacagg ggaaggtggc acgggatcgc tttcctctga acgcttctcg ctgctctttg 8520
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgctcatc cagatccaac 8640
 cccctgctat gtgcagggtc atcaaccagc agcccaggct gccagagcc acatccagcc 8700
 tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
 cagtgtaaaag ccattccccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
 tggggtgaca catgtttgcc aattcagtg atcacggaga ggcagatctt ggggataagg 8940
 aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
 gttaaaaccc agcatggaga ggagcacaaa aaggccacag aactgctgg tccctgtgtc 9120
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctgg 9180
 agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
 tcatacaatc gtcaaatcat gaaggttgga aaagccctcc aagatcccca agaccaaccc 9300
 caaccacccc accgtgccc ctggccatgt ccctcagtg cacatcccca cagttcttca 9360
 tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
 tctttggaga aggtaaatct tgctaaatcc agcccgaccc tcccctggca caacgtaagg 9480
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
 ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
 gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
 tagagaacgt atgaccagtt tactccctat cagtataga gaacgtatct acagtttact 9720
 ccctatcagt gatagagaac gtatatccag ttactccct atcagtata gagaacgtat 9780
 aagctttagg cgtgtacggg gggcgcctat aaaagcagag ctcgtttagt gaaccgtcag 9840
 atcgccctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttccta 9900
 cccctgtaaa aagcttgtcc accatggctc ctaagaaaaa gcggaagggtg gacaagaaat 9960
 actcaatcgg gctggccatc ggaactaact cagtgggggtg ggcagtcatt actgacgagt 10020
 acaaagtgcc aagcaagaaa ttaaggtcc tgggcaacac cgataggcac tccatcaaga 10080
 aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
 aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
 ttttctctaa cgaaatggcc aaggtggacg atagtttctt tcateggctg gaggaatcat 10260
 tcctggctga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
 aggtcgctta tcacgaaaaa taccaccca tetatcatct gcgcaagaaa ctggtggact 10380
 ctacagataa agcagacctg cggtgatct atctggccct ggctcacatg attaagttca 10440
 gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
 tcattccagct ggtccagaca tacaatcagc tgtttgagga aaacctatt aatgcactgt 10560
 gcgtggacgc aaaagccatc ctgagtcca ggctgtctaa gagtagaagg ctggagaacc 10620
 tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcactgt 10680
 cactgggact gacacctaac ttcaagagca attttgatct ggccgaggac gctaaactgc 10740

```

agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcggagatc 10800
agtacgcaga cctgttcctg gccgctaaga atctgtctga cgctatcctg ctgagtata 10860
ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagagat 10920
atgacgagca ccatcaggat ctgacctgc tgaaggcact ggtgcgccag cagctgccc 10980
agaagtacaa ggaaatcttc ttgatcaga gtaagaacgg gtacgccgggt tatattgacg 11040
gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
gcaccgagga actgctggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat 11160
ttgataacgg ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac 11220
gacaggaaga cttctacceca tttctgaagg ataaccgca gaagatcgaa aaaattctga 11280
ccttccggat cccttactat gtggggcccc tggcaagggg taattcccgc tttgcctgga 11340
tgacacggaa atctgaggaa acaatcactc cttggaactt cgaggaagtg gtcgataagg 11400
gagcttccgc acagtctttc atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg 11460
agaaagtgtc gcctaagcac agtctgtgt acgagtattt cacagtctat aacgaactga 11520
ctaaggtgaa atacgtcacc gaggggatga ggaagccccgc cttctgagc ggtgaacaga 11580
agaaagctat cgtggacctg ctgtttaaaa ccaatcgcaa ggtgacagtc aagcagctga 11640
aggaggacta cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag 11700
acagatttaa cgcttctctg ggaacctacc acgatctgt gaagatcatt aaggataaag 11760
acttcttgga caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac 11820
tgtttgagga tcgcgaaatg atcgaggaaac ggctgaaaac ttatgccc atgttcgatg 11880
acaaggtgat gaaacagctg aagcgaagaa ggtacaccgg ctggggacga ctgagcagaa 11940
agctgatcaa cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat 12000
cagatggctt cgctaacaga aattttatgc agctgattca cgatgacagc ctgaccttca 12060
aagaggatat ccagaaggca caggtgtccg ggcagggtga ctctctgcac gagcatatcg 12120
caaacctggc cgggtcccc gccatcaaga aaggtattct gcagaccgtg aaggtggtcg 12180
atgagctggt gaaagtcacg ggcaggcata agccagaaaa catcgtgatt gagatggccc 12240
gcgaaaaatca gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg 12300
aggaaggcat taaggaactg ggatcccaga tcctgaaaga gcaccctgtg gaaaacactc 12360
agctgcagaa tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg 12420
accaggagct ggatattaac cgactgtctg attacgacgt ggatgccatc gtcccacagt 12480
cattcctgaa agatgacagc attgacaata aggtgctgac ccggagtgc aaaaaccgag 12540
gaaagagtga taatgtccct tcagaggaag tggtaagaa aatgaagaac tactggagac 12600
agctgctgaa tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga 12660
gagggggtct gtcagaactg gacaaagcag gttcatcaa gcgacagctg gtggagacca 12720
gacagatcac aaagcacgtc gctcagattc tggatagcag gatgaacaca aagtacgatg 12780
agaatgacaa actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg 12840
acttcagaaa ggatttccag ttctacaaag tcaggagat caacaattat caccatgctc 12900
atgacgcata cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg 12960
agagcgaatt cgtgtacggt gactataagg tgtacgatgt cagaaaaatg atcgccaaga 13020
gtgagcagga aattggaaaa gccaccgcta agtatttctt ttactcaaac atcatgaatt 13080

```

tctttaagac tgagatcacc ctggcaaattg gggaaatccg aaagagacca ctgattgaga 13140
 ctaacggcga gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcgca 13200
 aggtcctgtc catgcctcaa gtgaatatgt tcaagaaaac agaggtgcag actggcggat 13260
 tcagtaagga atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag 13320
 actgggatcc caagaaatat gggggtttcg actccccaac agtggcttac tctgtcctgg 13380
 tggtcgcaaa ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg 13440
 gtatcactat tatggagagg agctccttcg agaagaaccc catcgatttt ctggaggcta 13500
 aaggctataa ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg 13560
 agctggaaaa cggaagggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg 13620
 aactggccct gccttctaag tacgtgaact tectgtatct ggctagccac tacgagaagc 13680
 tgaaggctc ccccgaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt 13740
 atctggacga gatcattgaa cagattagcg agttctccaa aagagtgate ctggctgacg 13800
 caaatctgga taaggtcctg agcgcatata acaaacacag agataagcca atcagggagc 13860
 aggccgaaaa tatcattcat ctgttcactc tgaccaacct gggagccccct gcagcettca 13920
 agtatcttga cactaccatc gatcggaaac gatacacatc cactaaggag gtgctggacg 13980
 ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
 tggggggcga cctcgagatg gccaaagcccc tgagccaaga ggaaagcacc ctgatcgagc 14100
 gggccaccgc caccatcaac agcatcccca tcagcgagga ctacagcgtg gcctctgccg 14160
 ccctgagcag cgacggcaga atcttcaccg gcgtgaacgt gtaccacttc acaggcggcc 14220
 cttgcgccga gctgggtggg ctgggaacag ctgccgccgc tgccgctggc aacctgacct 14280
 gtatcgtggc catcggcaac gagaaccggg gcatcctgag cccctgcggc agatgcagac 14340
 aggtgctgct ggacctgcac cccggcatca aggccatcgt gaaggacagc gacggccagc 14400
 ccaccgccgt gggcattaga gagctgctgc ccagcggcta cgtgtgggag ggctgactcg 14460
 agatcttcaa accagaagaa ctacgacagg cactgatgcc caccctggaa gccctgtacc 14520
 ggcaggaccc cgagagcctg cccttcagac agcccgtgga tccccagctg ctgggcatcc 14580
 ccgactactt cgacatcgtg aagtccccca tggacctgag caccatcaag cggaagctgg 14640
 acaccggcca gtaccaagag ccctggcagt acgtggacga catctggctg atgttcaaca 14700
 acgcctggct gtacaacaga aagaccagcc ggggtgtacaa gtactgcagc aagctgagcg 14760
 aggtgttcga gcaagagatc gaccccgtga tgcagagcct gggctactgc tgccgcagaa 14820
 agctggaatt cagccccag accctgtgct gctacggcaa gcagctgtgc accatcccc 14880
 gggacgccac ctactacagc taccagaaca gataccactt ctgcgagaag tgcttcaacg 14940
 agatccaggg cgagagcgtg tccttgggag acgaccctag ccagccccag accacaatca 15000
 acaaagagca gttcagcaag cggaagaacg acacctgga ccccgagctg ttcgtggaat 15060
 gcaccgagtg cggccggaag atgcaccaga tetgctgct gcaccacgag atcatctggc 15120
 ctgccggctt cgtgtgcgac ggctgcctga agaagtccgc ccggaccgg aaagagaaca 15180
 agttcagcgc caagcggctg ccctctaccc ggttgggcac ctctctggaa aacagagtga 15240
 acgacttcct gcggcggcag aaccaccccc agtccggcga agtgacagtg cgggtggtgc 15300
 acgccagcga caagaccgtg gaagtgaagc ctggcatgaa ggccagattc gtggacagcg 15360
 gcgagatggc cgagagcttc ccctaccgga ccaagccct gtctgccttc gaagagatcg 15420

```

atggcgtgga cctgtgcttc ttcggcatgc acgtgcaaga gtacggcagc gactgcccc 15480
cacccaacca gcggcggttg tacatcagct acctggacag cgtgcacttc ttccggccca 15540
agtgcctgcg gaccgccgtg tatcacgaga tcctgatcgg ctacctggaa tacgtgaaga 15600
agctgggcta caccaccggc cacatctggg cctgtcctcc cagcgagggc gacgactaca 15660
tcttccactg ccaccccccc gaccagaaga tccccaagcc caagagactg caagagtgg 15720
acaagaagat gctggacaag gccgtgtccg agcggatcgt gcacgactac aaggacatct 15780
tcaagcaggc caccgaggac cggctgacca gcgccaaaga gctgccctac ttcgagggcg 15840
acttctggcc caacgtgtg gaagagagca tcaaagagct ggaacaagag gaagaggaac 15900
gcaagcggga agagaacacc agcaacgaga gcaccgacgt gaccaagggc gacagcaaga 15960
acgccaagaa gaagaacaac aagaaaacca gcaagaacaa gagcagcctg agccggggaa 16020
acaagaaaaa gcccggcatg cccaacgtgt ccaacgacct gagccagaaa ctgtacgcca 16080
ccatggaaaa gcacaaagag gtgtttcttc tcatecggct gatcgccgga cctgccgcca 16140
acagcctgcc ccccatcgtg gaccccgacc cctgatecc ctgcgacctg atggacggca 16200
gggacgcctt cctgaccctg gcccgggaca agcacctgga attctccagc ctgcggagag 16260
cccagtggtc caccatgtgc atgctggtgg aactgcacac ccagagccag gacgagggca 16320
gaggaagtct gctaacatgc ggtgacgtcg aggagaatcc tggcccagca ccgggatcca 16380
tggtgagcaa gggcgaggag ctgttcaccg ggggtggtgc catcctggtc gagctggacg 16440
gcgacgtaaa cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg 16500
gcaagctgac cctgaagtgc atctgcacca ccggcaagct gcccgtgccc tgggccaccc 16560
tcgtgaccac cttcacctac ggcgtgcagt gcttcgcccc ctaccccgac cacatgaagc 16620
agcacgactt cttcaagtcc gccatgcccc aaggtacgt ccaggagcgc accatcttct 16680
tcaaggacga cggcaactac aagaccgcg ccgaggtgaa gttcgagggc gacaccctgg 16740
tgaaccgcat cgagctgaag ggcacgact tcaaggagga cggcaacatc ctggggcaca 16800
agctggagta caactacaac agccacaagg tctatatcac cgccgacaag cagaagaacg 16860
gcatcaaggt gaacttcaag acccgccaca acatcgagga cggcagcgtg cagctcgccc 16920
accactacca gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact 16980
acctgagcac ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc 17040
tgctggagtt cgtgaccgcc gccgggatca ctctcggeat ggacgagctg tacaagtaaa 17100
cctaattctag cagctcgctg atcagcctcg actgtgctt ctagttagcca gccatctgtt 17160
gtttgccccct ccccgctgcc ttccttgacc ctggaaggtg ccactcccac tgtcctttcc 17220
taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat tctgggggg 17280
ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggat 17340
gcggtgggct ctatggcttc tgaggcggaa agaaccagct ggggctcgat cctctagttg 17400
gcgcgtcatg gtccatatga atatcctcct tagttcctat tccgctagcc tagagggaca 17460
gcccccccc aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg 17520
agccgccccg ggctccgctc cggctccggc cccccccgc atccccgagc cggcagcgtg 17580
cggggacagc ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgtttctcg 17640
ctgctctttg agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga 17700
gagatttaga atgacagaat catagaacgg cctgggttgc aaaggagcac agtgcctatc 17760

```

cagatccaac cccctgctat gtgcagggtc atcaaccage agcccaggct gcccagagcc 17820
 acatccagcc tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac 17880
 ctgttcagtg cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc 17940
 tcccctgtct cagtgtaaag ccattccccc ttgtcctatc aagggggagt ttgctgtgac 18000
 attgttggtc tggggtgaca catgtttgcc aattcagtgc atcacggaga ggcagatctt 18060
 ggggataagg aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg 18120
 acactctcca agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga 18180
 caaagagcaa gttaaaaccc agcatggaga ggagcacaaa aaggccacag acactgctgg 18240
 tccctgtgtc tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga 18300
 gcttgcctgg agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc 18360
 atgtagatgt tcatacaatc gtcaaatcat gaaggctgga aaagccctcc aagatcccca 18420
 agaccaaccc caacccaccc accgtgceca ctggccatgt ccctcagtgc cacatcccca 18480
 cagttcttca tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact 18540
 gcagcaccgc tctttggaga aggtaaatct tgctaaatcc agcccgacc tcccctggca 18600
 caacgtaagg ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctgg 18660
 atccgaagca gctccagcct acacaatcgc tcaagacgtg taatgctttt attatatatt 18720
 agtcacgata tctataacaa gaaaatatat atataataag ttatcacgta agtagaacat 18780
 gaaataacaa tataattatc gtatgagtta aatcttaaaa gtcacgtaaa agataatcat 18840
 gcgtcatttt gactcacgcg gtcgttatag ttcaaaatca gtgacactta ccgcattgac 18900
 aagcacgcct cacgggagct ccaagcggcg actgagatgt cctaaatgca cagcgacgga 18960
 ttcgcgctat ttagaaagag agagcaatat ttcaagaatg catgcgtcaa ttttacgcag 19020
 actatctttc taggggtaaa aaagatttgc gctttactcg acctaaactt taaacacgtc 19080
 atagaatctt cgtttgacaa aaaccacatt gtgggggtacc gagctcttaa ttaaggcgcg 19140
 ccggggaggt tccctttagt gagggttaat tgcgggtcgc cctatagtga gtcgtattac 19200
 aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt 19260
 aatcgccctg cagcacatcc ccctttcgcc agctggcgta atagcgaaga ggcccgacc 19320
 gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcaaattgt aagcgttaat 19380
 attttgttaa aatcgcggtt aaatttttgt taaatcagct ctttttttaa ccaataggcc 19440
 gaaatcggca aaatccctta taaatcaaaa gaatagaccg agatagggtt gagtgttggt 19500
 ccagtttgga acaagagtcc actattaaag aacgtggact ccaacgtcaa agggcgaaaa 19560
 accgtctatc agggcgatgg cccactacgt gaaccatcac cctaatcaag ttttttgggg 19620
 tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga gccccgatt tagagcttga 19680
 cggggaaaagc cggcgaacgt ggcgagaaaag gaagggaaga aagcgaaagg agcgggcgct 19740
 agggcgctgg caagtgtagc ggtcacgtg cgcgtaacca ccacaccgc cgcgcttaat 19800
 gcgccgtac agggcgcgctc ag 19822

<210> 19

<211> 18492

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 19

```

gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 180
gccttccctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tggtgtcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataacctga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagtaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc ttagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgtc ggcccttccg gctggctggg ttattgtctga taaatctgga gccggtgagc 840
gtgggtctcg cggatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgatgtt aaaacttcat tttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt ttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 1380
tcctgttacc agtggctgct gccagtggcg ataagtctg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaaggcg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgectg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgat ctcgtcaggg gggcgagacc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggccctttgc tggccttttg 1800
ctcacatgtt ctttctgctg ttatccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
aagcggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcacccagc ctttacactt tatgcttccg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160

```



```

ccaagcgcgc aattaaccct cactaaaggg aacctcccct agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgtaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaa aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat ttgtactttt aaaaacagtc atgttgtatt ataaaataag 2520
taattagctt aacttatata taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaagggtt ccggtacgag ttgattcgag ccccgctggg ttctttccgc 2640
ctcagaagcc atagagccca ccgcaccccc agcatgcctg ctattgtctt cccaatcctc 2700
ccccttgctg tcctgcccc acccccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tcctcatttt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggtctggca ctagaaggca cagtcgaggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcatagaa 2940
ggcgatgcgc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc ggtagccca 3000
ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
attcggaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcagtcgcgc 3180
cttgagcctg gcgaacagtt cggctggcgc gagccctga tgctcttcgt ccagatcatc 3240
ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccgatcaag cgtatgcagc cgccgcattg catcagccat 3360
gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtccc ttcccgttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcggcatca gagcagccga ttgtctgttg tgcccagtc tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720
gttcctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
acgtgctgct gcaggctcaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840
tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcggggc cccgccccgc 3900
ccctgagccc gcccctgagc ccgcccccg acccaccct tcccagctc tgagcccaga 3960
aagcgaagga gcaaagctgc tattggccgc tgccccaaag gcctaccgc ttccattgct 4020
cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttccatt tgtaacgtcc 4080
tgcacgacgc gagctgcggg gcggggggga acttctgac taggggagga gtagaagggtg 4140
gcgcgaaggg gccaccaaag aacggagccg gttggcgct accggtggat gtggaatgtg 4200
tgcgagccag aggccacttg ttagcgcca agtgcacgc ggggctgcta aagcgcgtgc 4260
tcagactgc cttgggaaaa gcgcctcccc taccggtag acaccacaca gtgggtggcc 4320
tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 4440
cttacggtaa atggcccgcc tggtgaccg cccaacgacc cccgccatt gacgtcaata 4500

```

atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
 tattttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 4620
 cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 4680
 tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catgggtcga 4740
 ggtgagcccc acgttctgct tcactctccc catctcccc ccctccccac cccaattttt 4800
 gtattttattt attttttaat tattttgtgc agcgatgggg gcgggggggg ggggggcgcg 4860
 cgccaggcgg gcgggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg 4920
 cagccaatca gagcggcgcg ctccgaaagt ttctttttat ggcgaggcgg cggcggcggc 4980
 ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcc tgcctccgtg 5040
 ccccgctccg cgccgcctcg cgccgcccgc cccggtctg actgaccgcg ttactccac 5100
 aggtgagcgg gcgggacggc cttctctctc cgggtgtaa ttagcgcttg gtttaatgac 5160
 ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
 gcggggggga gcggctcggg ggggtgcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
 ggcccgcgct gcccggcggc tgtgagcgt gcgggcgcgg cgccgggctt tgtgcgtcc 5340
 gcgtgtgcgc gaggggagcg cggccggggg cggtgcccc cggtgccggg gggctgcgag 5400
 gggaacaaag gctgcgtgcg ggggtgtgtgc gtgggggggt gagcagggg tgtgggcgcg 5460
 gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc acggcccggc 5520
 ttcgggtgcg gggctccgtg cggggcgtgg cgccggggctc gccgtgccgg gcgggggggtg 5580
 gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
 ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc 5700
 cttttatggt aatcgtgcga gagggcgcag ggacttcctt tgtcccaaatt ctggcggagc 5760
 cgaaatctgg gaggcgccgc cgcacccct ctagcgggcg cgggcgaagc ggtgcggcgc 5820
 cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccctct 5880
 ccatctccag cctcggggct gccgcagggg gacggctgcc ttcggggggg acggggcagg 5940
 gcgggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
 gccttcttct ttttcttaca gctcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
 tttggcaaag aattcgccac catggtgccc aagaagaaga ggaaagtctc tagactggac 6120
 aagagcaaag tcataaaact tgctctggaa ttactcaatg gartcgggtat cgaaggcctg 6180
 acgacaagga aactcgctca aaagctggga gttgagcagc ctaccctgta ctggcacgtg 6240
 aagaacaagc gggccctgct cgatgcctg ccaatcgaga tgctggacag gcatcatacc 6300
 cactcctgcc ccctggaagg cgagtcatgg caagactttc tgcggaacaa cgccaagtca 6360
 taccgctgtg ctctcctctc acatcgcgac ggggctaaag tgcattctcg caccgcccc 6420
 acagagaaac agtacgaaac cctggaaaat cagctcgcgt tctgtgtc gcaaggcttc 6480
 tccttggaag acgcactgta cgctctgtcc gccgtgggac actttacact gggctgcgta 6540
 ttggaggaaac aggagcatca agtagcaaaa gaggaagag agacacctac caccgattct 6600
 atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
 ttctttttcg gcctggaact aatcatatgt ggcttgaga aacagctaaa gtgcgaaagc 6720
 ggcgggccga ccgacgcct tgacgatttt gacttagaca tgctcccagc cgatgccttt 6780
 gacgactttg acctgatata gctgcctgct gacgctcttg acgattttga ccttgacatg 6840

ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
 ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
 caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
 tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7080
 tcatcaatgt atcttaaggg atccctagag ggacagcccc cccccaagc cccagggat 7140
 gtaattacgt cctcccccg ctaggggcag cagcgagccg cccggggctc cgctccggtc 7200
 cggcgctccc cccgcattccc cgagccggca gcgtgcgggg acagccccggg cacggggaag 7260
 gtggcacggg atcgctttcc tctgaacgct tctcgtgct ctttgagcct gcagacacct 7320
 ggggggatac ggggaaaaag ctttaggctg aaagagagat ttagaatgac agaatcatag 7380
 aacggcctgg gttgcaaagg agcacagtgc tcatccagat ccaaccccct gctatgtgca 7440
 gggcatcaa ccagcagccc aggtgcccga gagccacatc cagcctggcc ttgaatgcct 7500
 gcagggatgg ggcatccaca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg 7560
 ggggaaaaac tgcctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
 cccccctgtc ctatcaaggg ggagtttgc gtgacattgt tggctctgggg tgacacatgt 7680
 ttgccaatc agtgcattc ggagaggcag atcttgggga taaggaagtg caggacagca 7740
 tggacgtggg acatgcaggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
 aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
 ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
 gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
 tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
 atcatgaagg ctggaaaagc cctccaagat cccaagacc aaccccaacc caccaccgt 8100
 gccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
 gtgaccccc cacctccgtg ggcagctgtg cactgcagc accgctcttt ggagaaggta 8220
 aatcttgcta aatccagccc gacctcccc tggcacaacg taaggccatt atctctcatc 8280
 caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
 aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
 ggctccgctc cggctccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
 ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgcttctcg ctgctctttg 8520
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgcctcag catatccaac 8640
 cccctgctat gtgcagggtc atcaaccagc agcccaggt gccagagcc acatccagcc 8700
 tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
 cagtgtaaag ccattccccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
 tggggtgaca catgtttgcc aattcagtc atcacggaga ggcagatctt ggggataagg 8940
 aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
 gttaaaaccc agcatggaga ggagcacaaa aaggccacag aactgtctgg tcctgtgtc 9120
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtgaaga gcttgccctg 9180

agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
 tcatacaatc gtcaaatcat gaaggctgga aaagccctcc aagatcccca agaccaaccc 9300
 caaccacccc accgtgcccc ctggccatgt ccctcagtgc cacatcccca cagttcttca 9360
 tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
 tcttttgaga aggtaaatct tgctaaatcc agcccagacc tcccctggca caacgtaagg 9480
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
 ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
 gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
 tagagaacgt atgaccagtt tactccctat cagtataga gaacgtatct acagtttact 9720
 ccctatcagt gatagagaac gtatatccag ttactccct atcagtata gagaacgtat 9780
 aagctttagg cgtgtacggt gggcgccctat aaaagcagag ctcgtttagt gaaccgtcag 9840
 atcgccctgga gcaattccac aacacttttg tcttatacca actttccgta ccacttecta 9900
 cctcgtaaa aagcttgtcc accatggctc ctaagaaaa gcggaagggt gacaagaaat 9960
 actcaatcgg gctggacatc ggaactaact cagtgggggt ggcaatcatt actgacgagt 10020
 acaaagtgcc aagcaagaaa ttaaggtcc tgggcaacac cgataggcac tccatcaaga 10080
 aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
 aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
 ttttctctaa cgaaatggcc aagggtggacg atagtttctt tcatcggtg gaggaatcat 10260
 tcttggtcga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
 aggtcgctta tcacgaaaaa taccacacca tctatcatct gcgcaagaaa ctggtggact 10380
 ctacagataa agcagacctg cggtgatct atctggccct ggctcacatg attaagttca 10440
 gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
 tcatccagct ggtccagaca tacaatcagc tgtttgagga aaaccctatt aatgcatctg 10560
 gcgtggacgc aaaagccatc ctgagtgcc ggctgtctaa gagtagaagg ctggagaacc 10620
 tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgactgt 10680
 cactgggact gacacctaac ttcaagagca attttgatct ggccaggac gctaaactgc 10740
 agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcgagatc 10800
 agtacgcaga cctgttcctg gccgctaaga atctgtctga cgctatcctg ctgagtata 10860
 ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagagat 10920
 atgacgagca ccatcaggat ctgacctgc tgaaggcact ggtgcgccag cagctgccg 10980
 agaagtacaa ggaaatcttc ttgatcaga gtaagaacgg gtacgccgtt tatattgacg 11040
 gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
 gcaccgagga actgctggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat 11160
 ttgataacgg ctccatcccc caccagatc atctgggaga gctgcacgca atcctgcgac 11220
 gacaggaaga cttctacca tttctgaagg ataaccgcga gaagatcgaa aaaattctga 11280
 ccttcggat cccttactat gtggggcccc tggcaagggg taattcccgc tttgectgga 11340
 tgacacggaa atctgaggaa acaatcactc cttggaactt cgaggaagtg gtcgataagg 11400
 gagcttcgc acagtcttc atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg 11460
 agaaagtgt gcctaagcac agtctgtgt acgagtattt cacagtctat aacgaactga 11520

ctaaggtgaa atacgtcacc gaggggatga ggaagcccg cttcctgagc ggtgaacaga 11580
 agaaagctat cgtggacctg ctgttttaaaa ccaatcgcaa ggtgacagtc aagcagctga 11640
 aggaggacta cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag 11700
 acagatttaa cgcttctctg ggaacctacc acgatctgct gaagatcatt aaggataaag 11760
 acttcctgga caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac 11820
 tgtttgagga tcgcgaaatg atcgaggaac ggctgaaaac ttatgccccat ctgttcgatg 11880
 acaaggtgat gaaacagctg aagcgaagaa ggtacaccgg ctggggacga ctgagcagaa 11940
 agctgatcaa cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat 12000
 cagatggctt cgctaacaga aattttatgc agctgattca cgatgacagc ctgaccttca 12060
 aagaggatat ccagaaggca caggtgtccg ggcagggtga ctctctgcac gagcatatcg 12120
 caaacctggc cgggtccccc gccatcaaga aaggtattct gcagaccgtg aaggtggctg 12180
 atgagctggt gaaagtcatt ggccaggcata agccagaaaa catcgtgatt gagatggccc 12240
 gcgaaaatca gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg 12300
 aggaaggcat taaggaactg ggatcccaga tcctgaaaga gcaccctgtg gaaaacactc 12360
 agctgcagaa tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg 12420
 accaggagct ggatattaac cgactgtctg attacgacgt ggatcatatc gtcccacagt 12480
 cattcctgaa agatgacagc attgacaata aggtgctgac ccggagtgc aaaaaccgag 12540
 gaaagagtga taatgtccct tcagaggaag tggtaagaa aatgaagaac tactggagac 12600
 agctgctgaa tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga 12660
 gagggggtct gtcagaactg gacaaagcag gcttcatcaa gcgacagctg gtggagacca 12720
 gacagatcac aaagcacgtc gctcagattc tggatagcag gatgaacaca aagtacgatg 12780
 agaatgacaa actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg 12840
 acttcagaaa ggatttccag ttctacaaag tcaggggagat caacaattat caccatgctc 12900
 atgacgcata cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg 12960
 agagcgaatt cgtgtacggt gactataagg tgtacgatgt cagaaaaatg atcgccaaga 13020
 gtgagcagga aattggaaaa gccaccgcta agtatttctt ttactcaaac atcatgaatt 13080
 tctttaagac tgagatcacc ctggcaaatt gggaaatccg aaagagacca ctgattgaga 13140
 ctaacggcga gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcgca 13200
 aggtcctgtc catgcctcaa gtgaatattg tcaagaaaac agaggtgcag actggcggat 13260
 tcagtaagga atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag 13320
 actgggatcc caagaaatat gggggtttcg actceccaac agtggcttac tctgtcctgg 13380
 tggtcgcaaa ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg 13440
 gtatcactat tatggagagg agtccttcg agaagaaccc catcgatttt ctggaggcta 13500
 aaggtataa ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg 13560
 agctggaaaa cggaagggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg 13620
 aactggccct gccttctaag tacgtgaact tcctgtatct ggctagccac tacgagaagc 13680
 tgaaaggctc ccccgaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt 13740
 atctggacga gatcattgaa cagattagcg agttctccaa aagagtgate ctggctgacg 13800
 caaatctgga taaggctcctg agcgcataca acaaacacag agataagcca atcaggagc 13860

```

aggccgaaaa tatcattcat ctgttcactc tgaccaacct gggagcccct gcagcettca 13920
agtatttttga cactaccatc gatcggaac gatacacatc cactaaggag gtgctggacg 13980
ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
tgggggggcga cctcgaggga agcggagagg gcagaggaag tctgctaaca tgcggtgacg 14100
tcgaggagaa tcctggccca gcaccgggat ccatggtgag caagggcgag gagctgttca 14160
ccgggggtggt gcccacctg gtcgagctgg acggcgacgt aaacggccac aagttcagcg 14220
tgtccggcga gggcgagggc gatgccacct acggcaagct gacctgaag ttcatctgca 14280
ccaccggcaa gctgcccgtg ccttgccca cctcgtgac caccttacc tacggcgtgc 14340
agtgtttcgc ccgtacccc gaccacatga agcagcacga cttcttcaag tccgcatgc 14400
ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacct 14460
gcgccgaggt gaagtctgag ggcgacacct tggatgaacc catcgagctg aagggcacgc 14520
acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 14580
aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaacttc aagaccgcc 14640
acaacatcga ggacggcagc gtgcagctcg ccgaccata ccagcagaac acccccatcg 14700
gcgacggccc cgtgctgctg cccgacaacc actacctgag caccagtcg gccctgagca 14760
aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgccggga 14820
tactctcgg catggacgag ctgtacaagt aaacctaatc tagcagctcg ctgatcagcc 14880
tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctccccctg gccttccttg 14940
accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcat 15000
tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 15060
gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg 15120
gaaagaacca gctggggctc gatcctctag ttggcgctc tgtacaaaaa agcaggcttt 15180
aaaggaacca attcagtcga ctggatccgg taccaaggtc gggcaggaag agggcctatt 15240
tcccatgatt ccttcatatt tgcatatacg atacaaggct gttagagaga taattagaat 15300
taatttgact gtaaacacaa agatattagt acaaaatacg tgacgtagaa agtaataatt 15360
tcttgggtag tttgcagttt taaaattatg ttttaaatg gactatcata tgcttaccgt 15420
aacttgaaag tatttcgatt tcttggcttt atatatcttg tggaaaggac gaaacaccga 15480
gcatagcaag ttaaaataag gctagtccgt tatcaacttg aaaaagtggc accgagtcgg 15540
tgcttttttt ctagaccag ctttcttgta caaagttggc attaattctc tagacatcat 15600
taattcctaa tttttgttga cactctatca ttgatagagt tttttacca ctccctatca 15660
gtgatagaga aaagtgaat ggccaagcct ttgtctcaag aagaatccac cctcattgaa 15720
agagcaacgg ctacaatcaa cagcatcccc atctctgaag actacagcgt cgccagcgca 15780
gctctctcta gcgacggccg catcttcact ggtgtcaatg tatatcattt tactggggga 15840
ccttgtgcag aactcgtggt gctgggcact gctgctgctg cggcagctgg caacctgact 15900
tgtatcgtcg cgatcggaat tgagaacagg ggcatcttga gcccctgcgg acggtgtcga 15960
caggtgcttc tcgatctgca tcctgggatc aaagcgatag tgaaggacag tgatggacag 16020
ccgacggcag ttgggattcg tgaattgctg ccctctggtt atgtgtggga gggctaaatg 16080
gtccatatga atatcctcct tagttcctat tccgctagcc tagagggaca gcccccccc 16140
aaagcccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 16200

```

ggctccgctc cgggtccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 16260
 ccgggcacagg ggaaggtggc acgggatcgc tttcctctga acgcttctcg ctgctctttg 16320
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 16380
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgctcatc cagatccaac 16440
 cccctgctat gtgcagggtc atcaaccagc agcccaggct gccagagcc acatccagcc 16500
 tggccttgaa tgcctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 16560
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 16620
 cagtgtaaaag ccattccccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 16680
 tggggtgaca catgtttgcc aattcagtg atcacggaga ggcagatctt ggggataagg 16740
 aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 16800
 agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 16860
 gttaaaaccc agcatggaga ggagcacaaa aaggccacag aactgtctgg tcctgtgtc 16920
 tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctgg 16980
 agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 17040
 tcatacaatc gtcaaatcat gaaggttga aaagccctcc aagatcccca agaccaaccc 17100
 caaccacccc accgtgccc ctggccatgt ccctcagtg cacatcccca cagttcttca 17160
 tcacctcag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 17220
 tctttggaga aggtaaatct tgctaaatcc agcccgacct tcccctggca caacgtaagg 17280
 ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctgg atccgaagca 17340
 gctccagcct acacaatcgc tcaagacgtg taatgctttt attatatatt agtcacgata 17400
 tctataacaa gaaaatatat atataataag ttatcacgta agtagaacat gaaataacaa 17460
 tataattatc gtatgagtta aatcttaaaa gtcacgtaaa agataatcat gcgtcatttt 17520
 gactcacgcg gtcgttatag ttcaaaatca gtgacactta ccgcattgac aagcacgcct 17580
 cacgggagct ccaagcggcg actgagatgt cctaaatgca cagcgacgga ttcgcgctat 17640
 ttagaaagag agagcaatat ttcaagaatg catgcgtcaa ttttacgcag actatctttc 17700
 taggggttaa aaagatttgc gctttactcg acctaaactt taaacacgtc atagaatctt 17760
 cgtttgacaa aaaccacatt gtggggtacc gagctcttaa ttaaggcgcg ccggggaggt 17820
 tccccttagt gaggggtaat tgcgggtcgc cctatagtga gtcgtattac aattcactgg 17880
 ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccctg 17940
 cagcacatcc ccccttcgcc agctggcgta atagcgaaga ggccgcacc gatcgccctt 18000
 cccaacagtt gcgcagcctg aatggcgaat ggcaaattgt aagcgtaaat attttggtta 18060
 aattcgcggt aaatttttgt taaatcagct cattttttta ccaataggcc gaaatcgga 18120
 aaatccctta taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga 18180
 acaagagtcc actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc 18240
 agggcgatgg ccactacgt gaaccatcac cetaatcaag ttttttgggg tcgaggtgcc 18300
 gtaaaagcact aaatcggaac cctaaaggga gccccgatt tagagcttga cggggaaagc 18360
 cggcgaacgt ggcgagaaa gaagggaaga aagcgaaagg agcgggcgct agggcgctgg 18420
 caagtgtagc ggtcacgctg cgcgtaacca ccacaccgc cgcgcttaat gcgccgtac 18480
 agggcgcgctc ag 18492

<210> 20

<211> 17575

<212> DNA

<213> 人工序列

<220>

<223> 合成的多核苷酸

<400> 20

```
gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatTTTTc taaatacatt 60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 120
ggaagagtat gagtattcaa ctttccgtg tcgcccttat tccccTTTT gcggcatttt 180
gccttcctgt ttttgtcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 240
tgggtgcacg agtgggttac atcgaaactgg atctcaacag cggtaagatc cttgagagtt 300
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 360
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 420
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 480
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 540
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 600
ctcgccctga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 660
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attactggc gaactactta 720
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 780
ttctgcgctc ggcccttcgg gctggctggg ttattgtctga taaatctgga gccggtgagc 840
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 900
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 960
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 1020
agattgatatt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 1080
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 1140
aaaagatcaa aggatcttct tgagatcctt ttttctgcg cgtaatctgc tgcttgcaaa 1200
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 1260
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt ctagtgtagc 1320
cgtagttagg ccaccacttc aagaactctg tagcacgcc tacatactc gctctgctaa 1380
tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 1440
gacgatagtt accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc 1500
ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 1560
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 1620
caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 1680
ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcgagacc 1740
tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggcccttttg tgcccttttg 1800
ctcacatgtt ctttctgcg ttatccccctg attctgtgga taaccgtatt accgcctttg 1860
agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg 1920
```



```

aagcgggaaga gcgccaata cgcaaaccgc ctctccccgc gcgttgccg attcattaat 1980
gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac gcaattaatg 2040
tgagttagct cactcattag gcaccccagg ctttacactt tatgcttcg gctcgtatgt 2100
tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac catgattacg 2160
ccaagcgcgc aattaaccct cactaaaggg aacctccctt agcttaatta accctagaaa 2220
gataatcata ttgtgacgta cgtaaagat aatcatgcgt aaaattgacg catgtgtttt 2280
atcgatctgt atatcgaggt ttatttatta atttgaatag atattaagtt ttattatatt 2340
tacacttaca tactaataat aaattcaaca aacaatttat ttatgtttat ttatttatta 2400
aaaaaaaaaca aaaactcaaa atttcttcta taaagtaaca aaacttttaa acattctctc 2460
ttttacaaaa ataaacttat tttgtacttt aaaaacagtc atgttgtatt ataaaataag 2520
taattagctt aacttataca taatagaaac aaattatact tattaatcgc attgattatt 2580
gactagtcgt attaagggtt ccgcatcagc ttgattcgag cccagctgg ttctttccgc 2640
ctcagaagcc atagagccca ccgcacccc agcatgcctg ctattgtctt cccaatcctc 2700
ccccttgctg tcctgcccc acccccccc cagaatagaa tgacacctac tcagacaatg 2760
cgatgcaatt tcctcathtt attaggaaag gacagtggga gtggcacctt ccagggtcaa 2820
ggaaggcacg ggggaggggc aaacaacaga tggctggcaa ctagaaggca cagtcagggc 2880
tgatcagcga gctctagaga attgatcccc tcagaagaac tcgtcaagaa ggcgatagaa 2940
ggcgatgcgc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca 3000
ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc 3060
cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat 3120
attcggaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcagtcgcgc 3180
cttgagcctg gcgaacagtt cggctggcgc gagccctga tgctcttcgt ccagatcatc 3240
ctgatcgaca agaccggcct ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg 3300
gtggtcgaat gggcaggtag ccgatcaag cgtatgcagc cgccgcattg catcagccat 3360
gatggatact ttctcggcag gagcaagggt agatgacagg agatcctgcc ccggcacttc 3420
gccaatagc agccagtccc tttccgcttc agtgacaacg tcgagcacag ctgcgcaagg 3480
aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcagtt cattcagggc 3540
accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac 3600
ggcggcatca gagcagccga ttgtctgttg tgcccagtea tagccgaata gcctctccac 3660
ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atggccgatc ccatggttta 3720
gttcctcacc ttgtcgtatt atactatgcc gatatactat gccgatgatt aattgtcaac 3780
acgtgctgct gcaggtcgaa aggcccggag atgaggaaga ggagaacagc gcggcagacg 3840
tgcgcttttg aagcgtgcag aatgccgggc ctccggagga ccttcgggcg cccgccccgc 3900
ccctgagccc gccctgagc ccgcccccg acccaccctt tcccagctc tgagcccaga 3960
aagcgaagga gcaaagctgc tattggccgc tgcccaaag gcctaccgc ttccattgct 4020
cagcgggtgct gtccatctgc acgagactag tgagacgtgc tacttcatt tgtaacgtcc 4080
tgcacgacgc gagctgcggg gcggggggga acttctgac taggggagga gtagaaggtg 4140
gcgcgaaggg gccaccaaag aacggagccg gttggcgctt accggtggat gtggaatgtg 4200
tgcgagccag aggccacttg ttagcgcca agtgcccagc ggggctgcta aagcgcagtc 4260

```

```

tccagactgc cttgggaaaa ggcctcccc tacccggtag acaccccaca gtgggtggcc 4320
tagggacagg attgcaactc cagtctttct tcttcttggg cgggagtcac tagttattaa 4380
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 4440
cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 4500
atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 4560
tattttacgg aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgccc 4620
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 4680
tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catgggtcga 4740
ggtgagcccc acgttctgct tcactctccc catctcccc cctccccac cccaatttt 4800
gtatttatatt attttttaat tattttgtgc agcgatgggg gcgggggggg gggggcgcg 4860
cgccaggcgg gcggggcggg ggcgaggggc gggcgggggc gaggcggaga ggtgcggcgg 4920
cagccaatca gagcggcgcg ctccgaaagt ttcttttat ggcgaggcgg cggcggcggc 4980
ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgctt tcgccccgtg 5040
ccccgctccg cgccgcctcg cgccgcccgc cccgctctg actgaccgcg ttactcccac 5100
aggtgagcgg gcgggacggc cttctctc cgggtgtaa ttagcgcttg gtttaatgac 5160
ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag ggccctttgt 5220
gcggggggga gcggctcggg ggggtcgtgc gtgtgtgtgt gcgtggggag cgccgcgtgc 5280
ggcccgcgct gcccggcggc tgtgagcgct gcgggcgcgg cgcggggctt tgtgcgtcc 5340
gcgtgtgcgc gaggggagcg cggccggggg cggtgcccc cggtgcgggg gggctgcgag 5400
gggaacaaag gctgcgtgcg ggggtgtgtc gtgggggggt gagcaggggg tgtgggcgcg 5460
gcggtcgggc tgtaaccccc cctgcaccc cctccccga gttgctgagc acggcccggc 5520
ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg gcgggggggtg 5580
gcggcaggtg ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcggggga 5640
ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg cagccattgc 5700
cttttatggt aatcgtgcga gagggcgag ggacttcctt tgtcccaa atctggcgagc 5760
cgaaatctgg gaggcgccgc cgcacccct ctagcgggcg cgggcgaagc ggtgcggcgc 5820
cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccctct 5880
ccatctccag cctcggggct gccgcagggg gacggtgcc ttcggggggg acggggcagg 5940
gcgggggttc gcttctggcg tgtgaccggc ggctctagag cctctgctaa ccatgttcat 6000
gccttcttct ttttctaca gctcctgggc aacgtgctgg ttattgtgct gtctcatcat 6060
tttggcaaag aattcgccac catggtgcc aagaagaaga ggaaagtctc tagactggac 6120
aagagcaaag tcataaactc tgctctggaa ttactcaatg gagtccgtat cgaaggcctg 6180
acgacaagga aactcgtca aaagetggga gttgagcagc ctaccctgta ctggcacgtg 6240
aagaacaagc ggccctgct cgatgcctg ccaatcgaga tgctggacag gcatcatacc 6300
cactcctgcc ccctggaagg cgagtcattg caagacttct tgcggaacaa cgccaagtca 6360
taccgtgtg ctctcctctc acatcgcgac ggggctaaag tgcatctcgg caccgcccc 6420
acagagaaac agtacgaaac cctggaaaat cagctcgcgt tcctgtgtca gcaaggtctc 6480
tccttgaga acgactgta cgctctgtcc gccgtgggcc actttacact gggctgcgta 6540
ttggaggaa acaggagcatca agtagcaaaa gaggaagag agacacctac caccgattct 6600

```

atgccccac ttctgaaaca agcaattgag ctgttcgacc ggcagggagc cgaacctgcc 6660
 ttccttttcg gcctggaact aatcatatgt ggcctggaga aacagctaaa gtgcgaaagc 6720
 ggcggggccga ccgacgccct tgacgatttt gacttagaca tgctcccagc cgatgccctt 6780
 gacgactttg accttgatat gctgcctgct gacgctcttg acgattttga cttgacatg 6840
 ctccccgggt aaagcggccg cgactctaga tcataatcag ccataccaca tttgtagagg 6900
 ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat aaaatgaatg 6960
 caattgttgt tgtaacttg tttattgcag cttataatgg ttacaaataa agcaatagca 7020
 tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac 7080
 tcatcaatgt atcttaaggg atccctagag ggacagcccc ccccaaaagc cccagggat 7140
 gtaattacgt ccctcccccg ctagggggcag cagcagcccg cccggggctc cgctccggctc 7200
 cggcgctccc cccgcacccc cgagccggca gcgtgcgggg acagccccggg cacggggaag 7260
 gtggcacggg atcgctttcc tctgaacget tctcgetget ctttgagcct gcagacacct 7320
 ggggggatac ggggaaaaag ctttaggetg aaagagagat ttagaatgac agaatcatag 7380
 aacggcctgg gttgcaaagg agcacagtgc tcatccagat ccaaccccc gctatgtgca 7440
 gggcatcaa ccagcagccc aggetgccc gagccacatc cagcctggcc ttgaatgcct 7500
 gcagggatgg ggcattccaca gcctccttgg gcaacctgtt cagtgcgtca ccacctctg 7560
 ggggaaaaac tgctcctca tatccaaccc aaacctcccc tgtctcagtg taaagccatt 7620
 ccccttgtc ctatcaaggg ggagtttgct gtgacattgt tggctctgggg tgacacatgt 7680
 ttgccaatc agtgcatcac ggagaggcag atcttgggga taaggaagtg caggacagca 7740
 tggacgtggg acatgcagggt gttgagggt ctgggacact ctccaagtca cagcgttcag 7800
 aacagcctta aggataagaa gataggatag aaggacaaag agcaagttaa aaccagcat 7860
 ggagaggagc acaaaaaggc cacagacact gctggtccct gtgtctgagc ctgcatgttt 7920
 gatggtgtct ggatgcaagc agaaggggtg gaagagcttg cctggagaga tacagctggg 7980
 tcagtaggac tgggacaggc agctggagaa ttgccatgta gatgttcata caatcgtcaa 8040
 atcatgaagg ctggaaaagc cctccaagat cccaagacc aaccccaacc caccaccgt 8100
 gcccactggc catgtccctc agtgccacat cccacagtt cttcatcacc tccagggacg 8160
 gtgaccccc cacctccgtg ggcagctgtg ccaactgcagc accgctcttt ggagaaggta 8220
 aatcttgcta aatccagccc gacctcccc tggcacaacg taaggccatt atctctcatc 8280
 caactccagg acggagtcag tgaggatggg gctgtcgacc tagagggaca gcccccccc 8340
 aaagccccca gggatgtaat tacgtccctc ccccgctagg ggcagcagcg agccgcccgg 8400
 ggctccgctc cggctccggcg ctccccccgc atccccgagc cggcagcgtg cggggacagc 8460
 ccgggcacgg ggaaggtggc acgggatcgc tttcctctga acgttctctg ctgctctttg 8520
 agcctgcaga cacctggggg gatacgggga aaaagcttta ggctgaaaga gagatttaga 8580
 atgacagaat catagaacgg cctgggttgc aaaggagcac agtgcctatc cagatccaac 8640
 cccctgctat gtgcagggtc atcaaccagc agcccagget gccagagcc acatccagcc 8700
 tggccttgaa tgctgcagg gatggggcat ccacagcctc cttgggcaac ctgttcagtg 8760
 cgtcaccacc ctctggggga aaaactgcct cctcatatcc aacccaaacc tcccctgtct 8820
 cagtgtaaag ccattcccc ttgtcctatc aagggggagt ttgctgtgac attgttggtc 8880
 tggggtgaca catgtttgcc aattcagtgc atcacggaga ggcagatctt ggggataagg 8940

aagtgcagga cagcatggac gtgggacatg caggtgttga gggctctggg acactctcca 9000
agtcacagcg ttcagaacag ccttaaggat aagaagatag gatagaagga caaagagcaa 9060
gttaaaaccc agcatggaga ggagcacaaa aaggccacag acactgctgg tccctgtgtc 9120
tgagcctgca tgtttgatgg tgtctggatg caagcagaag ggggtggaaga gcttgccctgg 9180
agagatacag ctgggtcagt aggactggga caggcagctg gagaattgcc atgtagatgt 9240
tcatacaatc gtcaaatcat gaaggctgga aaagccctcc aagatcccca agaccaaccc 9300
caaccacccc accgtgcccc ctggccatgt ccctcagtgc cacatcccca cagttcttca 9360
tcacctccag ggacggtgac cccccacct ccgtgggcag ctgtgccact gcagcaccgc 9420
tctttggaga aggtaaatct tgctaaatcc agccccaccc tcccctggca caacgtaagg 9480
ccattatctc tcatccaact ccaggacgga gtcagtgagg atggggctca attgtttact 9540
ccctatcagt gatagagaac gtatgaagag ttactccct atcagtata gagaacgtat 9600
gcagacttta ctccctatca gtgatagaga acgtataagg agtttactcc ctatcagtga 9660
tagagaacgt atgaccagtt tactccctat cagtataga gaacgtatct acagtttact 9720
ccctatcagt gatagagaac gtatatccag ttactccct atcagtata gagaacgtat 9780
aagctttagg cgtgtacggt gggcgcttat aaaagcagag ctgctttagt gaaccgtcag 9840
atcgcttga gcaattccac aacacttttg tcttatacca actttccgta ccacttccta 9900
ccctcgtaaa aagcttgtcc accatggctc ctaagaaaaa gcggaagggt gacaagaaat 9960
actcaatcgg gctggccatc ggaactaact cagtgggggt ggcaagtcatt actgacgagt 10020
acaaagtgcc aagcaagaaa ttaaggtcc tgggcaacac cgataggcac tccatcaaga 10080
aaaatctgat tggggccctg ctgttcgact ctggagagac agctgaagca actagactga 10140
aaaggactgc tagaaggcgc tatacccggc gaaagaatcg catctgctac ctgcaggaga 10200
ttttctctaa cgaaatggcc aaggtggacg atagtttctt tcacggctg gaggaatcat 10260
tcctggctga ggaagataag aaacacgaga gacatcctat ctttggaac attgtggacg 10320
aggtcgctta tcacgaaaaa taccacacca tctatcatct gcgcaagaaa ctggtggact 10380
ctacagataa agcagacctg cggtgatct atctggccct ggctcacatg attaagttca 10440
gaggccattt tctgatcgag ggagatctga acccagacaa tagcgatgtg gacaagctgt 10500
tcattccagct ggtccagaca tacaatcagc tgtttgagga aaaccctatt aatgcatctg 10560
gcgtggacgc aaaagccatc ctgagtcca ggctgtctaa gagtagaagg ctggagaacc 10620
tgatcgctca gctgccaggc gaaaagaaaa acggcctgtt tggaaatctg attgcaactgt 10680
cactgggact gacacctaac ttcaagagca attttgatct ggccaggac gctaaactgc 10740
agctgagcaa ggacacttat gacgatgacc tggataacct gctggctcag atcggagatc 10800
agtacgcaga cctgttccct gccgctaaga atctgtctga cgctatcctg ctgagtata 10860
ttctgcgggt gaacaccgag attacaaaag cccctctgtc agctagcatg atcaagagat 10920
atgacgagca ccatcaggat ctgacctgc tgaaggcact ggtgcgccag cagctgccc 10980
agaagtacaa ggaaatcttc ttgatcaga gtaagaacgg gtacgccgt tatattgacg 11040
gcggagcttc acaggaggaa ttctacaagt ttatcaaacc tattctggag aagatggacg 11100
gcaccgagga actgctggtg aaactgaatc gcgaggacct gctgcgcaag cagcggacat 11160
ttgataacgg ctccatcccc caccagattc atctgggaga gctgcacgca atcctgcgac 11220
gacaggaaga cttctacca tttctgaagg ataaccgcga gaagatcgaa aaaattctga 11280

```

ccttccggat cccttactat gtggggcccc tggcaagggg taattcccgc tttgcctgga 11340
tgacacggaa atctgaggaa acaatcactc cttggaactt cgaggaagtg gtcgataagg 11400
gagcttccgc acagtctttc atcgagagaa tgacaaactt cgacaaaaac ctgccaaatg 11460
agaaagtgtc gcctaagcac agtctgtctg acgagtattt cacagtctat aacgaactga 11520
ctaagggtgaa atacgtcacc gaggggatga ggaagcccg cttcctgagc ggtgaacaga 11580
agaaagctat cgtggacctg ctgttttaaaa ccaatcgcaa ggtgacagtc aagcagctga 11640
aggaggacta cttcaagaaa attgaatgtt tcgattctgt ggagatcagt ggcgtcgaag 11700
acagatttaa cgcttctctg ggaacctacc acgatctgct gaagatcatt aaggataaag 11760
acttcctgga caacgaggaa aatgaggata tcctggaaga cattgtgctg accctgacac 11820
tgtttgagga tcgcgaaatg atcgaggaaac ggctgaaaac ttatgcccac ctgttcgatg 11880
acaaggtgat gaaacagctg aagcgaagaa ggtacaccgg ctggggacga ctgagcagaa 11940
agctgatcaa cggcattcgg gacaaacaga gtggaaagac tatcctggac tttctgaaat 12000
cagatggctt cgctaacaga aattttatgc agctgattca cgatgacagc ctgaccttca 12060
aagaggatat ccagaaggca caggtgtccg ggcagggtga ctctctgcac gagcatatcg 12120
caaacctggc cgggtccccc gccatcaaga aaggtattct gcagaccgtg aaggtggtcg 12180
atgagctggg gaaagtcatg ggcaggcata agccagaaaa catcgtgatt gagatggccc 12240
gcgaaaatca gaccacacag aaaggacaga agaacagccg cgagcggatg aaaaggatcg 12300
aggaaggcat taaggaaactg ggatcccaga tcctgaaaga gcacctgtg gaaaacactc 12360
agctgcagaa tgagaagctg tatctgtact atctgcagaa tgggcgggat atgtacgtgg 12420
accaggagct ggatattaac cgactgtctg attacgacgt ggatcatatc gtcccacagt 12480
cattcctgaa agatgacagc attgacaata aggtgctgac ccggagtgcac aaaaaccgag 12540
gaaagagtga taatgtccct tcagaggaag tggtaagaa aatgaagaac tactggagac 12600
agctgctgaa tgccaaactg atcacacagc gaaagtttga taacctgact aaagctgaga 12660
gaggggggtc gtcagaactg gacaaagcag gcttcatcaa gcgacagctg gtggagacca 12720
gacagatcac aaagcacgtc gctcagattc tggatagcag gatgaacaca aagtacgatg 12780
agaatgacaa actgatccgc gaagtgaagg tcattactct gaagtcaaaa cttgtgagcg 12840
acttcagaaa ggatttccag ttctacaaag tcaggggagat caacaattat caccatgctc 12900
atgacgcata cctgaacgca gtggtcggga ccgccctgat taagaaatac cccaaactgg 12960
agagcgaatt cgtgtacggg gactataagg tgtacgatgt cagaaaaatg atcgccaaga 13020
gtgagcagga aattggaaaa gccaccgcta agtattttct ttactcaaac atcatgaatt 13080
tcitttaagac tgagatcacc ctggcaaatg gggaaatccg aaagagacca ctgattgaga 13140
ctaacggcga gaccggagaa atcgtgtggg acaagggtag ggattttgcc acagtgcgca 13200
aggtcctgtc catgcctcaa gtgaatattg tcaagaaaac agaggtgcag actggcggat 13260
tcagtaagga atcaattctg cccaaacgga actctgataa gctgatcgcc cgaaagaaag 13320
actgggatcc caagaaatat gggggtttcg actccccaac agtggttac tctgtcctgg 13380
tggtcgcaaa ggtggagaag gggaaaagca agaaactgaa atccgtcaag gagctgctgg 13440
gtatcactat tatggagagg agtccttcg agaagaaccc catcgatttt ctggaggcta 13500
aaggctataa ggaagtgaag aaagacctga tcattaaact gccaaagtac agcctgtttg 13560
agctggaaaa cggaaggaag cgaatgctgg catccgcagg agagctgcag aagggtaatg 13620

```

aactggccct gccttctaag tacgtgaact tcctgtatct ggctagccac tacgagaagc 13680
tgaaggctc ccccgaggat aacgaacaga aacagctgtt tgtggagcag cacaagcatt 13740
atctggacga gatcattgaa cagattagcg agttctccaa aagagtgatc ctggctgacg 13800
caaactctgga taaggtcctg agcgcataca acaaacacag agataagcca atcagggagc 13860
aggccgaaaa tatcattcat ctgttcactc tgaccaacct gggagcccct gcagccttca 13920
agtattttga cactaccatc gatcggaac gatacacatc cactaaggag gtgctggacg 13980
ctaccctgat tcaccagagc attaccggcc tgtatgaaac aaggattgac ctgtctcagc 14040
tggggggcga cctcgaggga agcggagagg gcagaggaag tctgctaaca tgcggtgacg 14100
tcgaggagaa tcctggccca gcaccgggat ccatggtgag caagggcgag gagctgttca 14160
ccggggtggt gcccacctcgt gtcgagctgg acggcgacgt aaacggccac aagttcagcg 14220
tgtccggcga gggcgagggc gatgccacct acggcaagct gacctgaag ttcatctgca 14280
ccaccggcaa gctgcccgtg ccctggccca cctcgtgac caccttacc tacggcgtgc 14340
agtgtctcgc ccgtacccc gaccacatga agcagcacga cttcttcaag tccgcatgc 14400
ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacc 14460
gcgccgaggt gaagttcgag ggcgacaccc tggatgaacc catcgagctg aagggcacgc 14520
acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 14580
aggtctatat caccgccgac aagcagaaga acggcatcaa ggtgaactc aagacccgcc 14640
acaacatcga ggacggcagc gtgcagctcg ccgaccacta ccagcagaac acccccatcg 14700
gcgacggccc cgtgctgctg ccgcacaacc actacctgag caccagtcg gccctgagca 14760
aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgccggga 14820
tcaactctcg catggacgag ctgtacaagt aaacctaatc tagcagctcg ctgatcagcc 14880
tcgactgtgc cttctagtgt ccagccatct gttgtttgcc cctccccctg gccttccttg 14940
accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcgat 15000
tgtctgagta ggtgtcattc tattctgggg ggtgggggtgg ggcaggacag caagggggag 15060
gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg 15120
gaaagaacca gctggggctc gatcctctag ttggcgcgtc atggtccata tgaatattct 15180
ccttagttcc tattccgcta gcctagaggg acagccccc cccaaagccc ccagggatgt 15240
aattacgtcc ctcccccgct aggggcagca gcgagccgcc cggggctccg ctccgggtccg 15300
gcgctcccc cgcatccccg agccggcagc gtgcggggac agcccgggca cggggaaggt 15360
ggcacgggat cgctttcctc tgaacgctt tcgctgctct ttgagcctgc agacacctgg 15420
ggggatacgg ggaaaaagct ttaggctgaa agagagattt agaatgacag aatcatagaa 15480
cggcctgggt tgcaaaggag cacagtgtc atccagatcc aacccctgc tatgtgcagg 15540
gtcatcaacc agcagcccag gctgcccaga gccacatcca gcctggcctt gaatgcctgc 15600
agggatgggg catccacagc ctcttgggc aacctgttca gtgcgtcacc acctctggg 15660
ggaaaaactg cctcctcata tccaacccaa acctccccctg tctcagtgtg aagccattcc 15720
ccctgtcct atcaaggggg agtttgtgtg gacattgttg gtctgggggtg acacatgttt 15780
gccaattcag tgcatcacgg agaggcagat cttggggata aggaagtgc ggacagcatg 15840
gacgtgggac atgcagggtg tgagggtctt gggacactct ccaagtcaca gcgttcagaa 15900
cagccttaag gataagaaga taggatagaa ggacaaagag caagttaaaa cccagcatgg 15960

agaggagcac aaaaaggcca cagacactgc tggtcctgt gtctgagcct gcatgtttga 16020
tggtgtcttg atgcaagcag aaggggtgga agagcttgcc tggagagata cagctgggtc 16080
agtaggactg ggacaggcag ctggagaatt gccatgtaga tgttcataca atcgtcaaatt 16140
catgaaggct ggaaaagccc tccaagatcc ccaagaccaa ccccaaccca cccaccgtgc 16200
ccactggcca tgtccctcag tgccacatcc ccacagttct tcatcacctc cagggacggt 16260
gaccccccca cctccgtggg cagctgtgcc actgcagcac cgctcttttg agaaggtaaa 16320
tcttgctaaa tccagcccga ccctcccctg gcacaacgta aggccattat ctctcatcca 16380
actccaggac ggagtcagtg aggatggggc tggatccgaa gcagctccag cctacacaat 16440
cgctcaagac gtgtaatgct tttattatat attagtcacg atatctataa caagaaaata 16500
tatatataat aagttatcac gtaagtagaa catgaaataa caatataatt atcgtatgag 16560
ttaaatctta aaagtcacgt aaaagataat catgcgtcat ttgactcac gcggtcgta 16620
tagttcaaaa tcagtgcac ttaccgcatt gacaagcacg cctcacggga gctccaagcg 16680
gcgactgaga tgcctaaat gcacagcgac ggattcgcg tatttagaaa gagagagcaa 16740
tatttcaaga atgcatgcgt caattttacg cagactatct ttctagggtt aaaaaagatt 16800
tgcgtttac tcgacctaaa ctttaaacac gtcatagaat cttcgtttga caaaaaccac 16860
attgtggggg accgagctct taattaaggc gcgccgggga ggttcccttt agtgagggtt 16920
aattgcgggt cgccctatag tgagtcgtat tacaattcac tggccgtcgt ttacaacgt 16980
cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tcccccttc 17040
gccagctggc gtaatagcga agaggccgc accgatcgcc cttcccaaca gttgcgcagc 17100
ctgaatggcg aatggcaaatt tgtaagcgtt aatattttgt taaaattcgc gttaaatttt 17160
tgttaaatca gtcattttt taaccaatag gccgaaatcg gcaaaatccc ttataaatca 17220
aaagaataga ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta 17280
aagaacgtgg actccaacgt caaaggcgca aaaaccgtct atcaggcgca tggcccaacta 17340
cgtgaaccat caccctaact aagttttttg gggtcgaggt gccgtaaagc actaaatcgg 17400
aacccataag ggagccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 17460
aaggaaggga agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg 17520
ctgcgcgtaa ccaccacacc cgccgcgctt aatgcgcgcg tacaggcgcg gtcag 17575

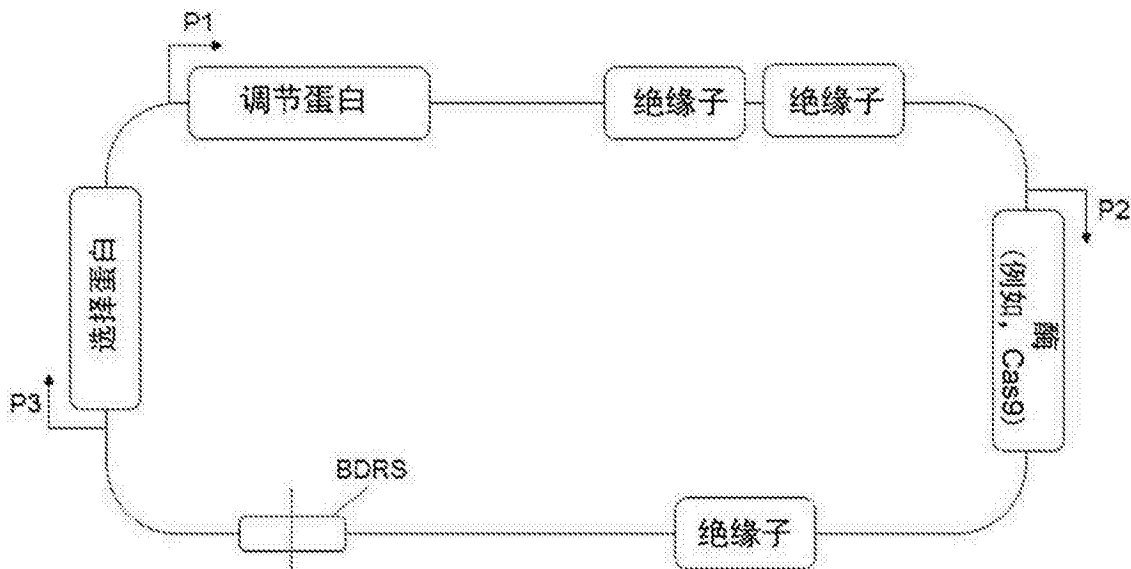


图1

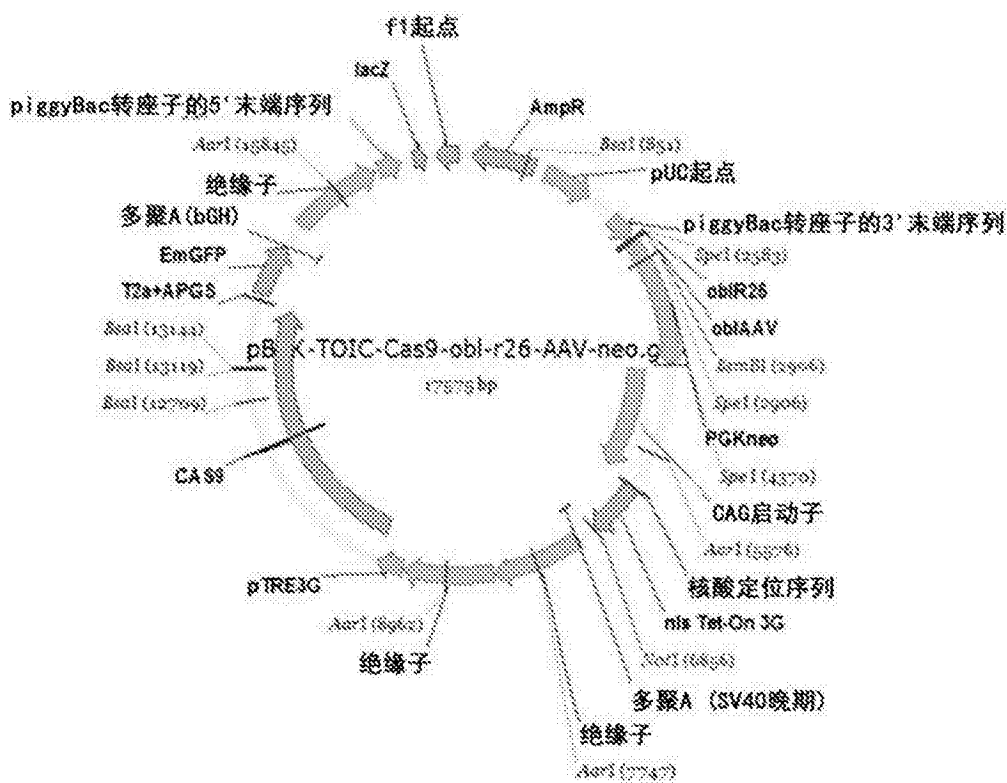


图2

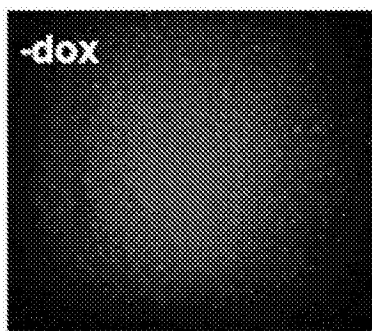


图3A

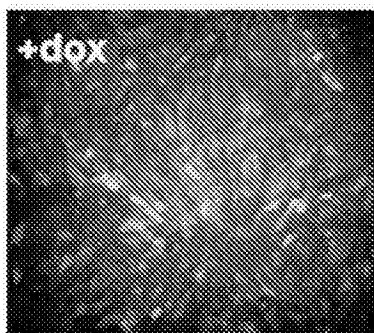


图3B

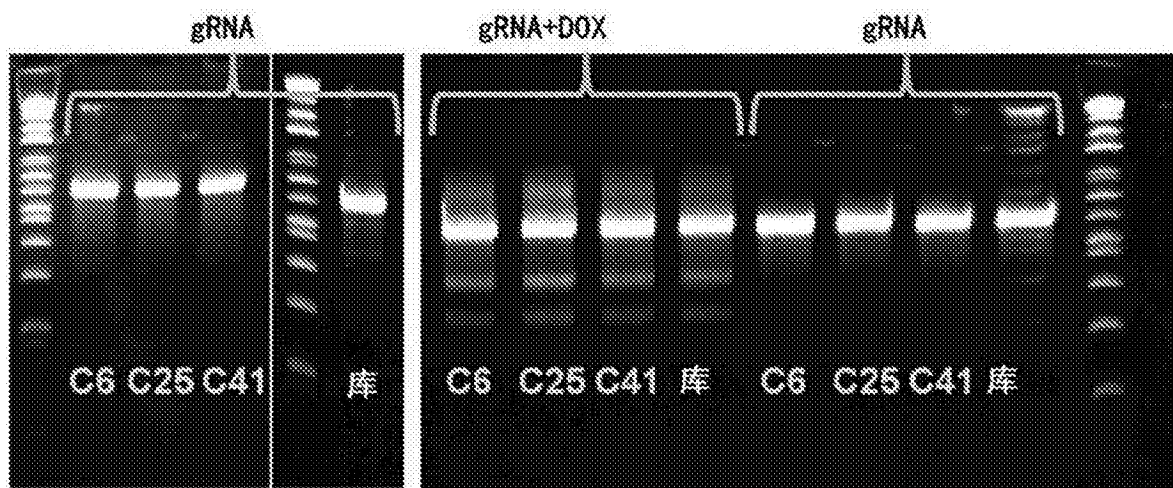


图4

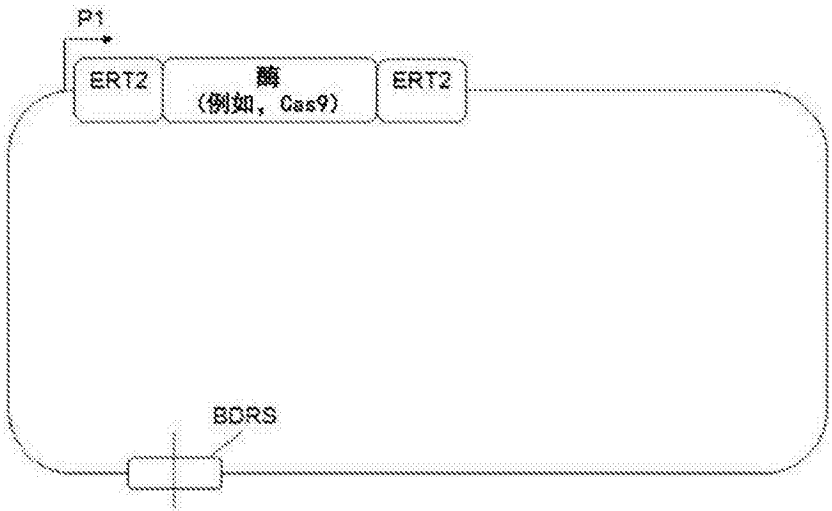


图5

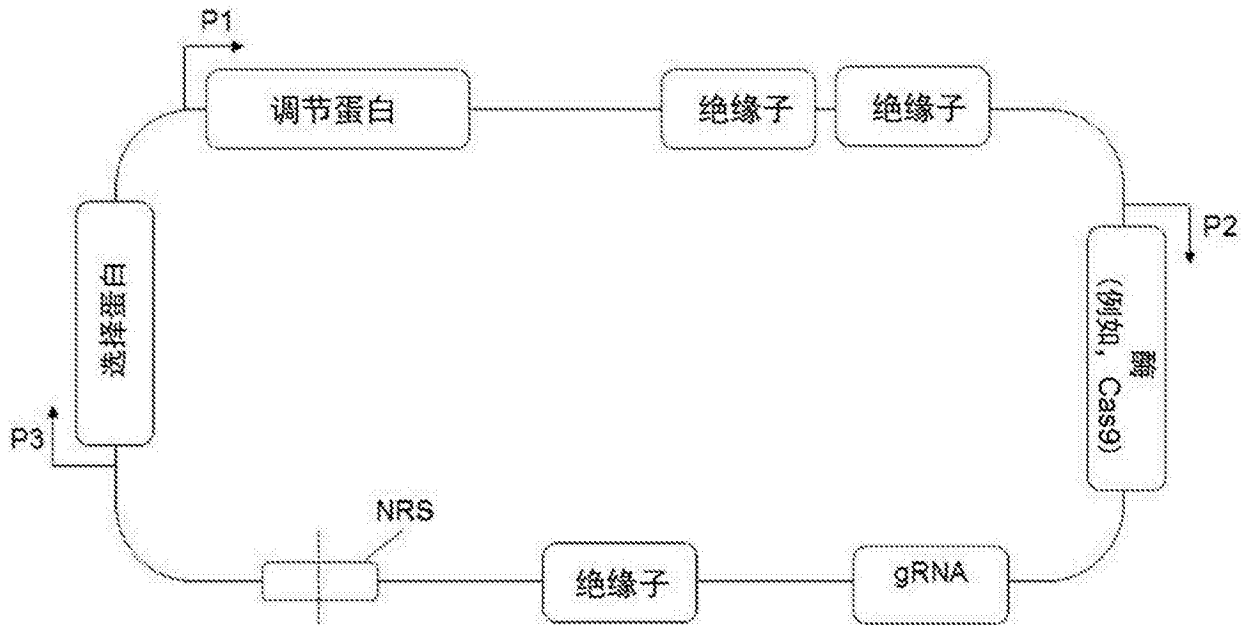


图6

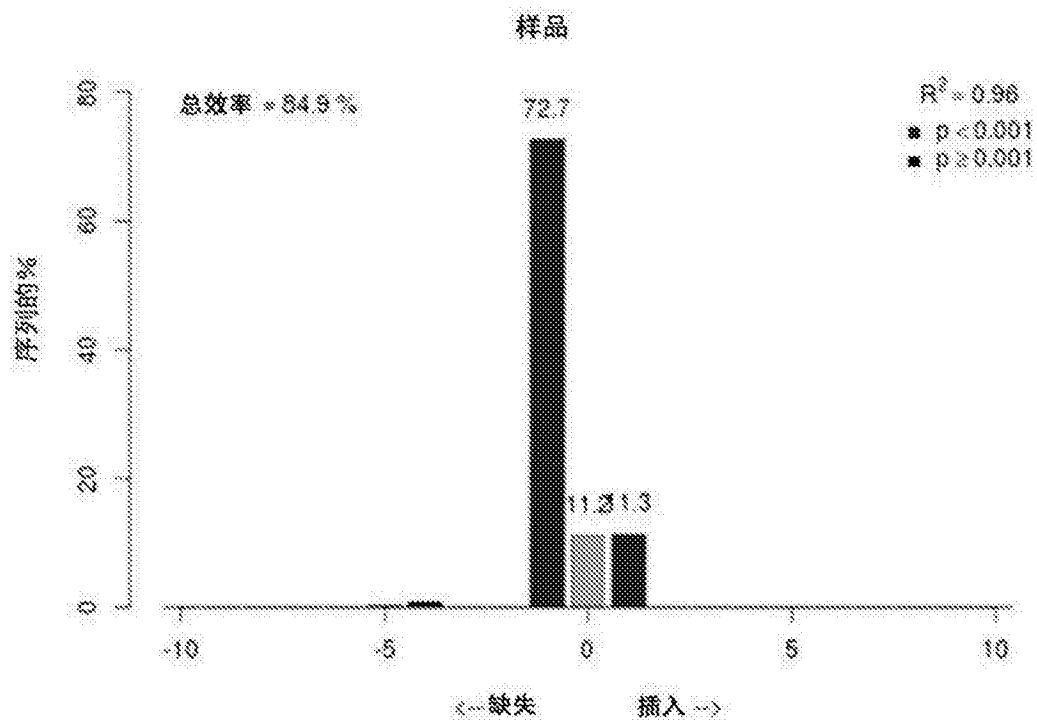


图7

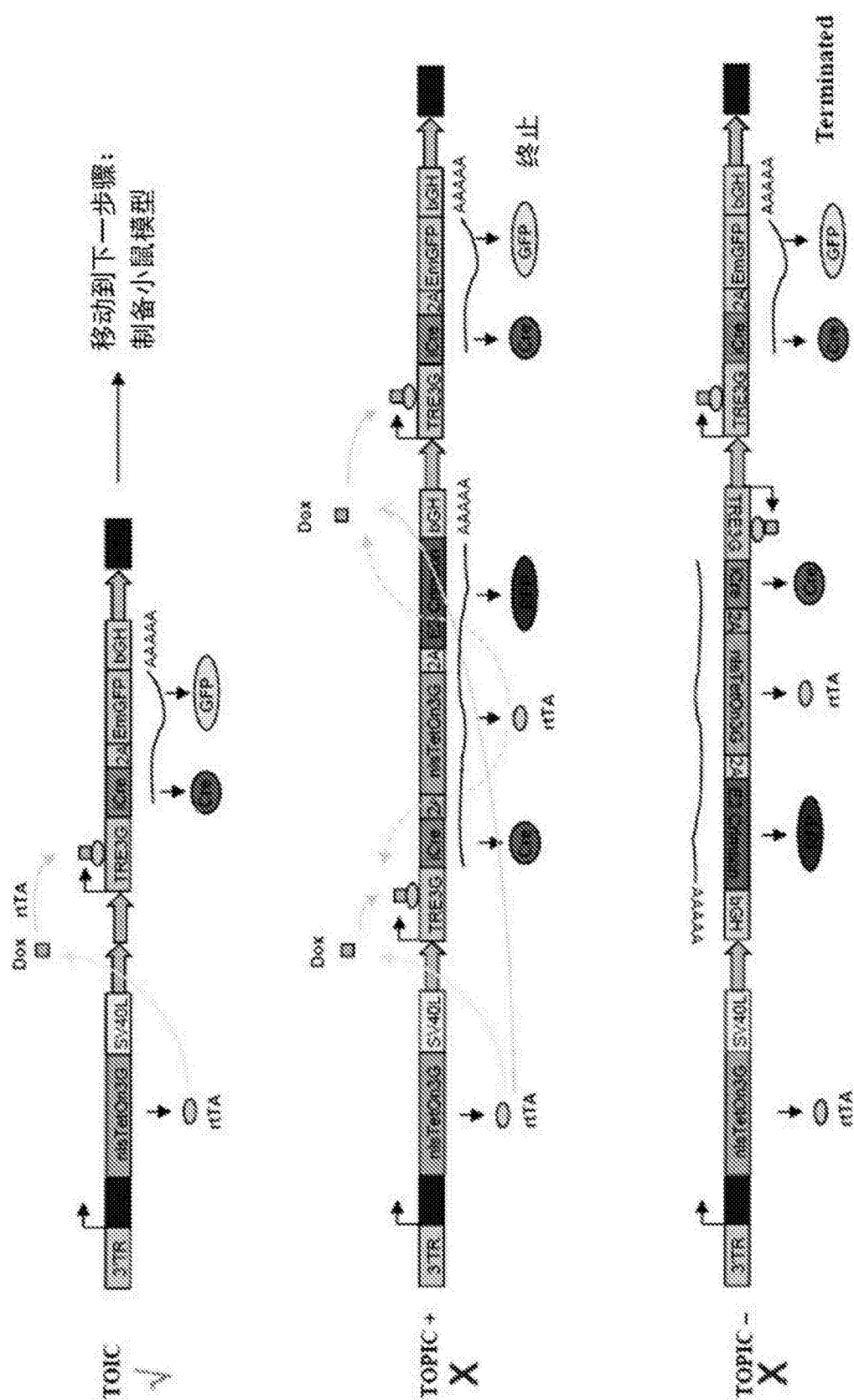


图8

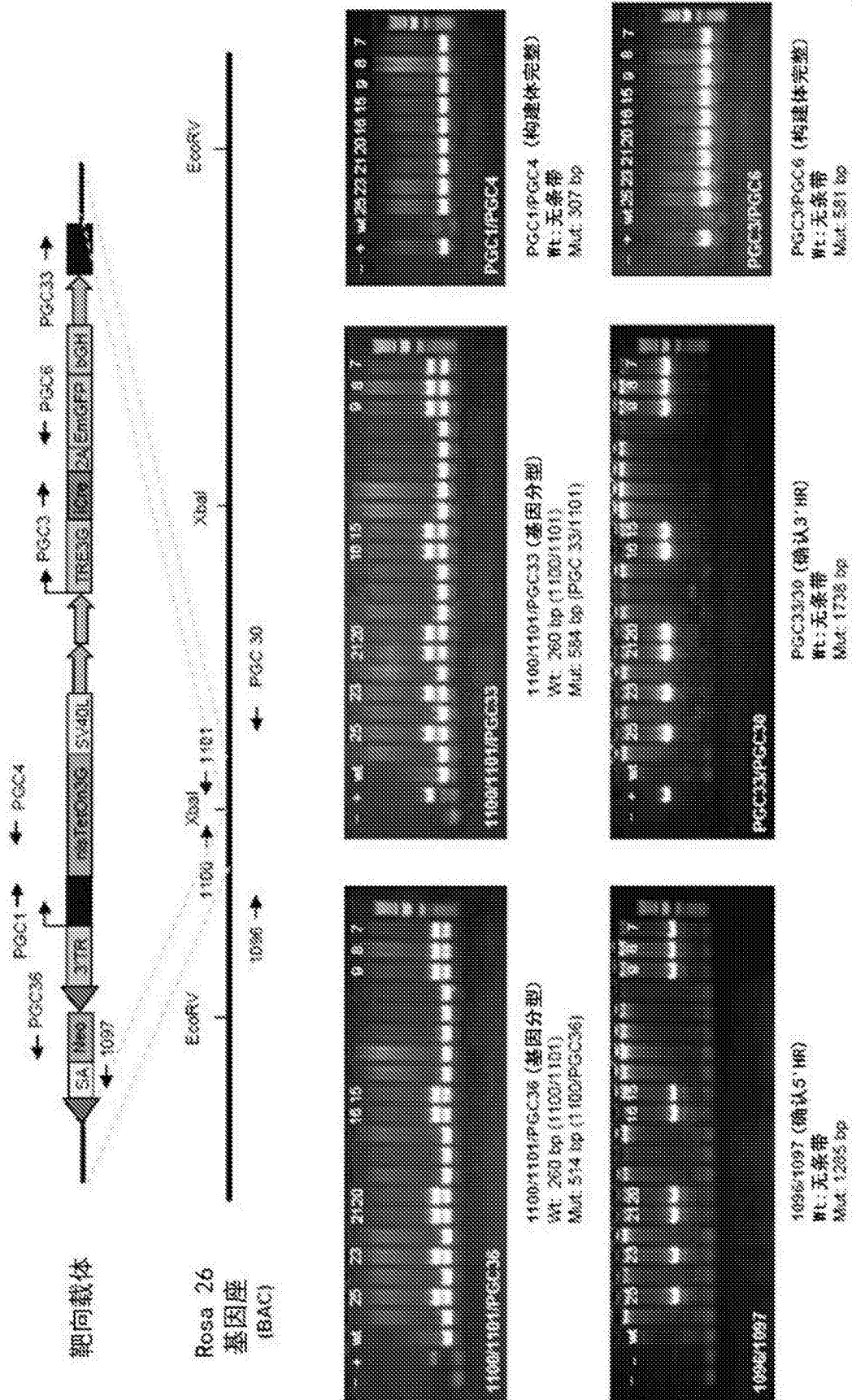


图9



图10

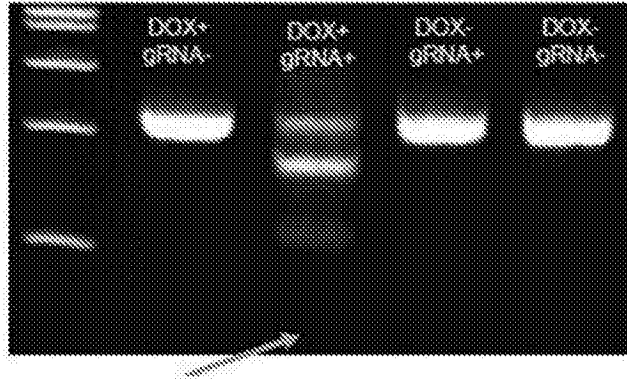


图11

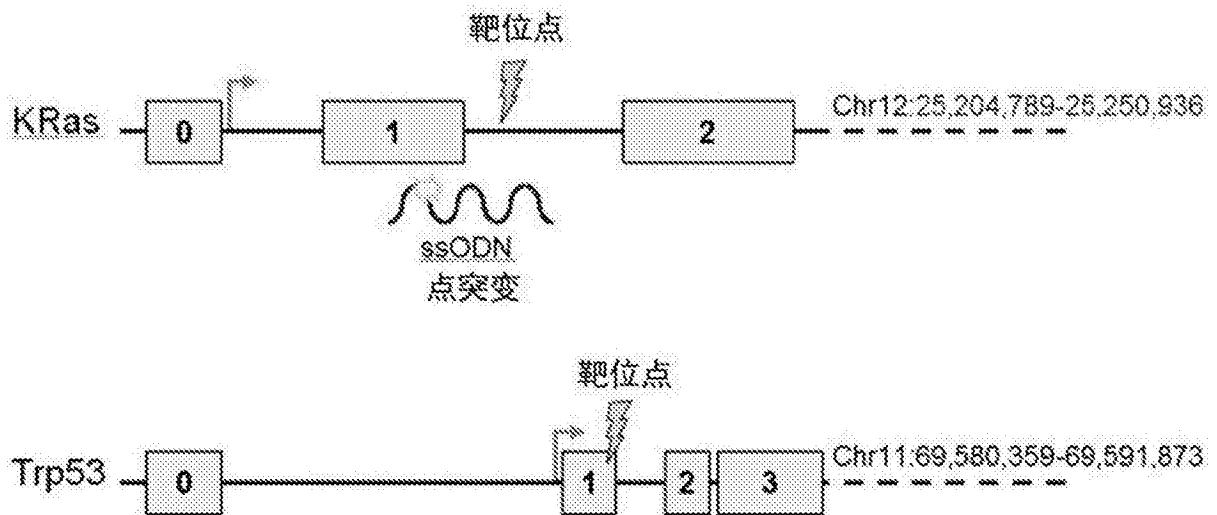


图12

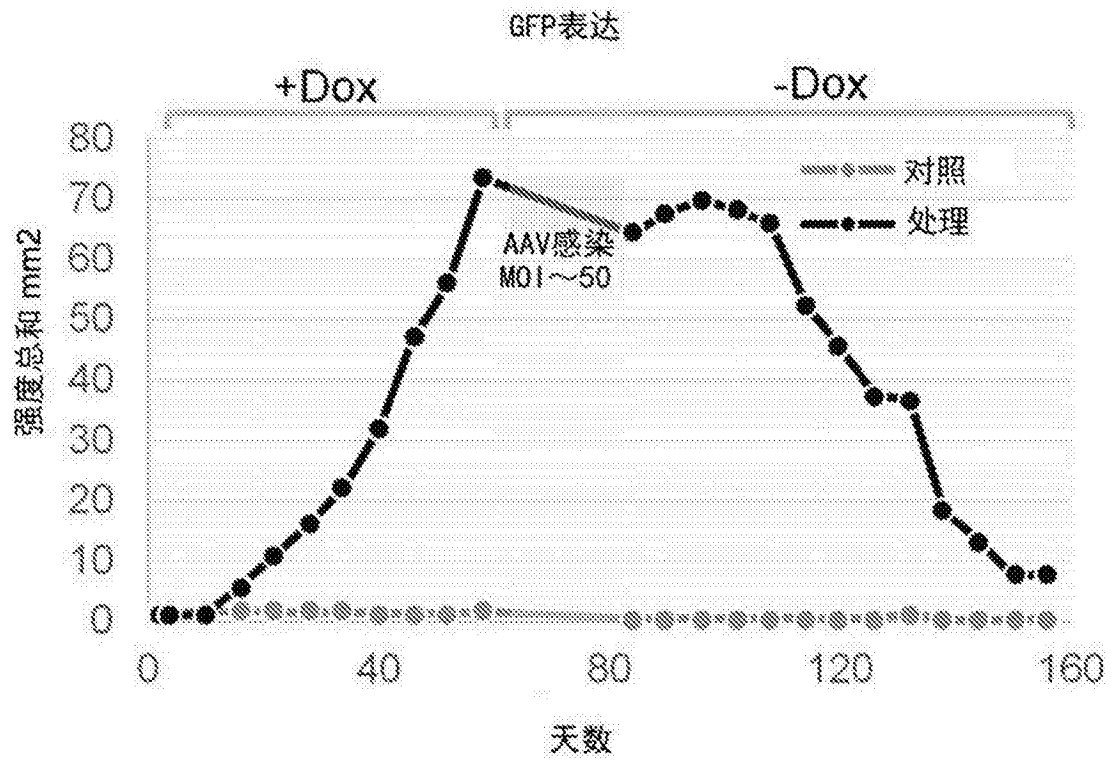


图13

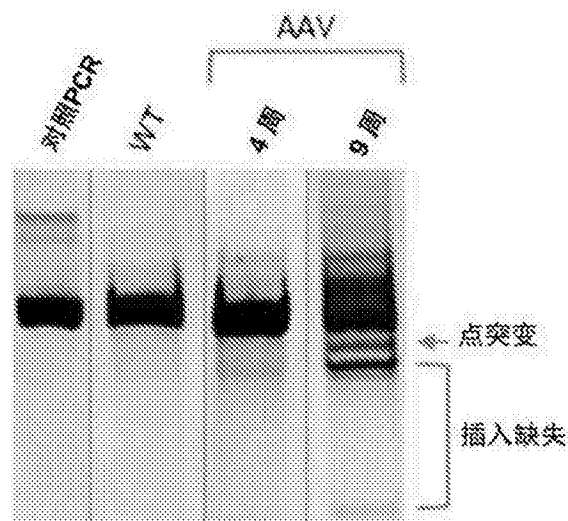


图14

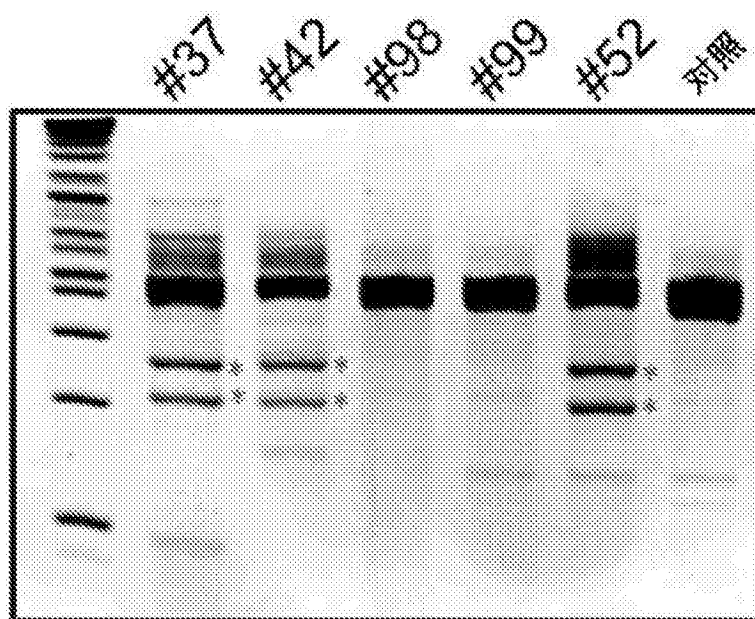


图15

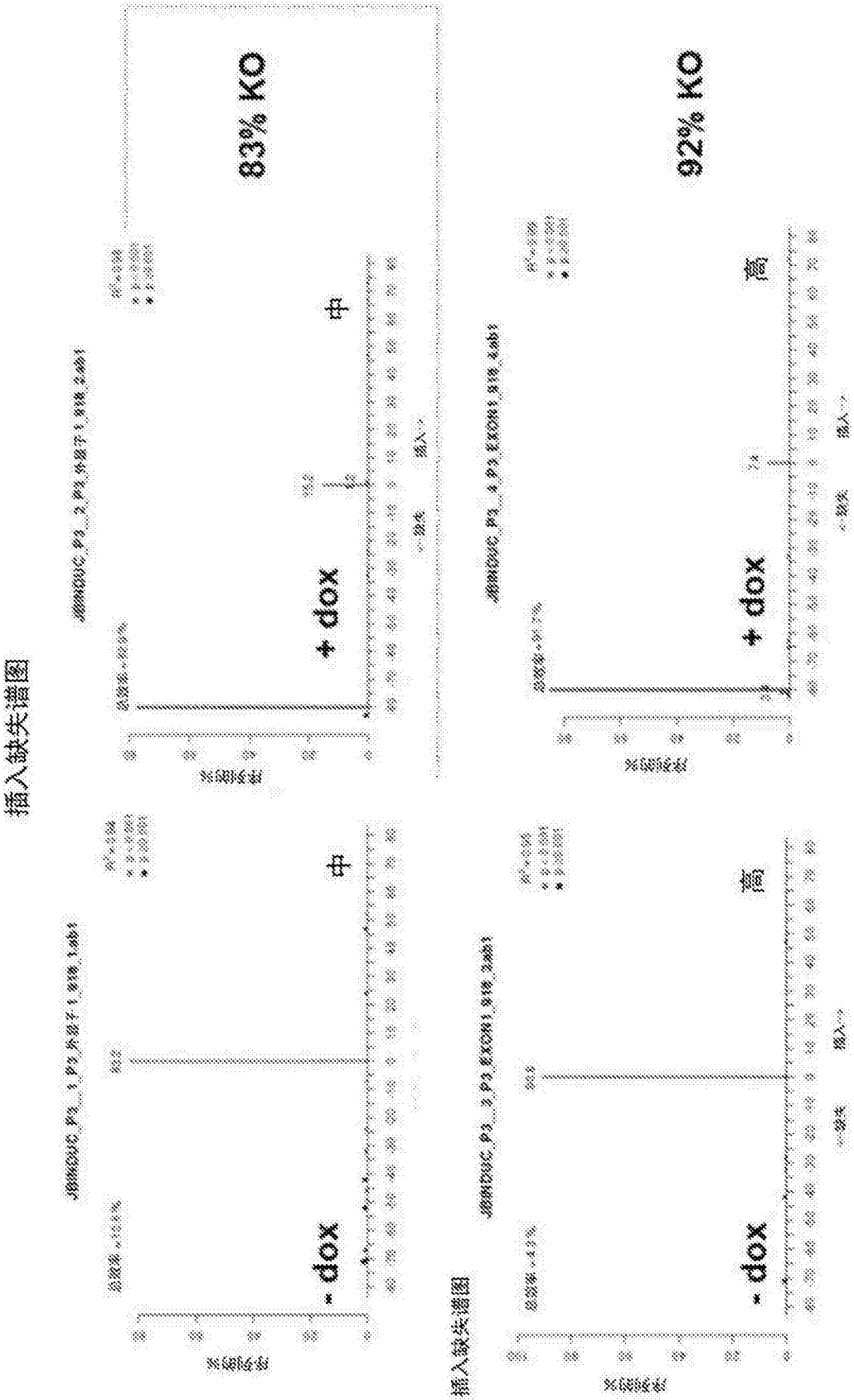


图16

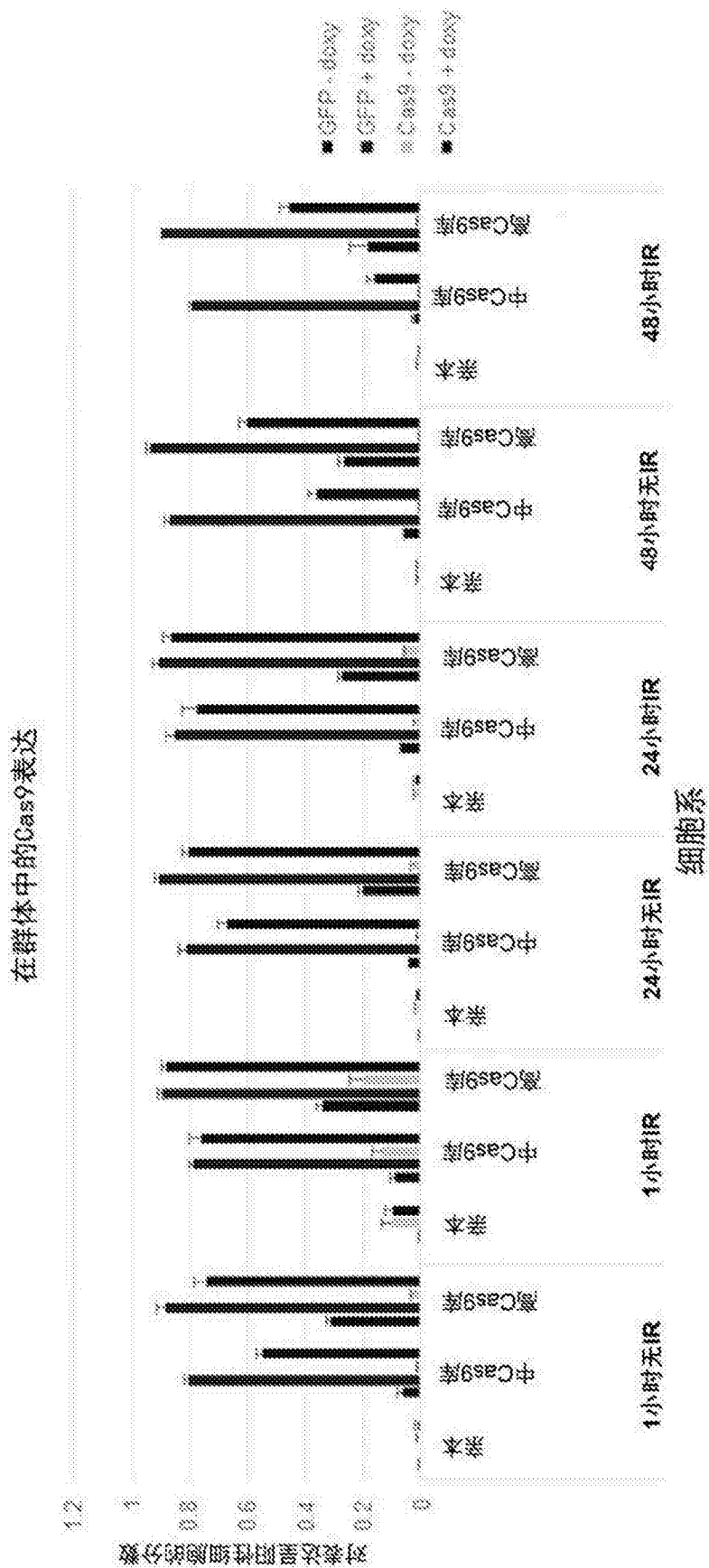


图17

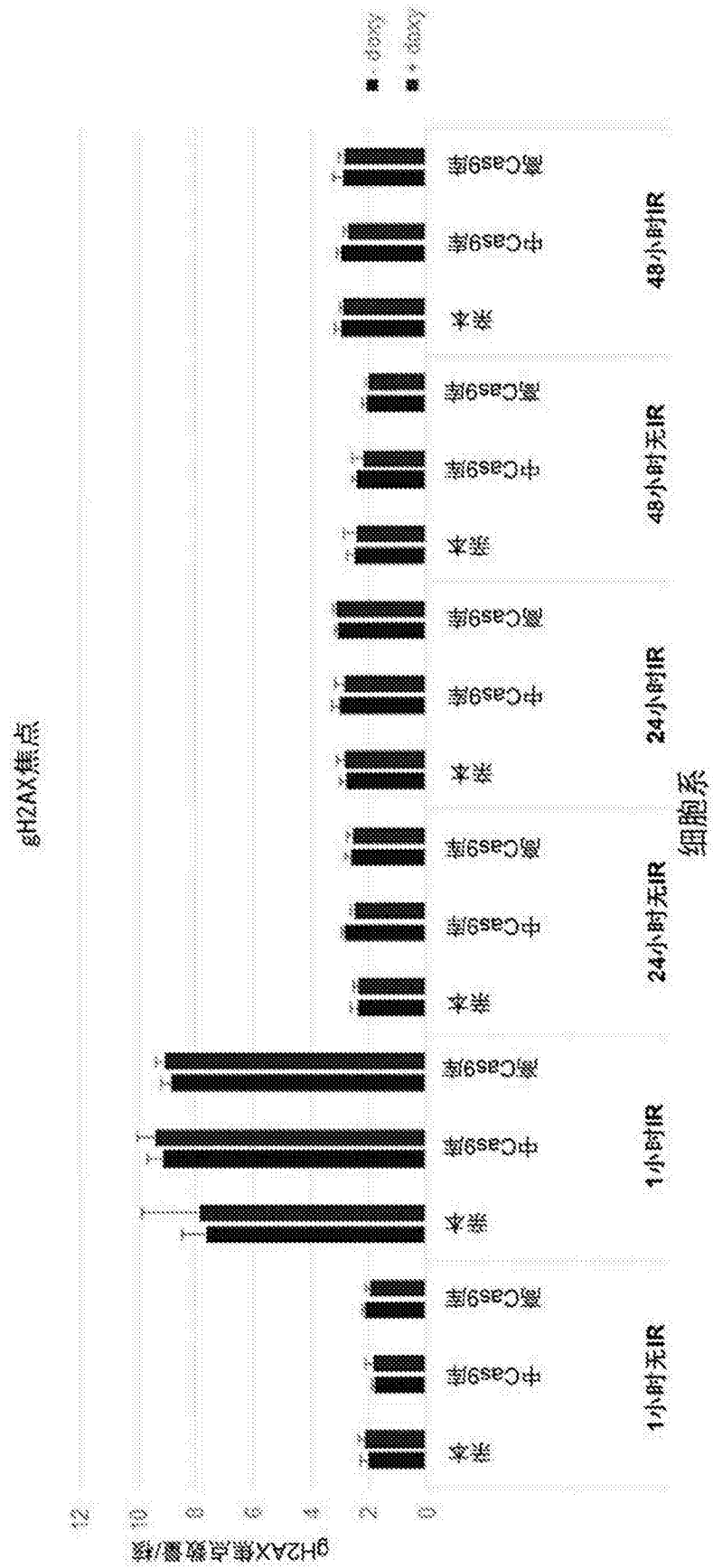


图18

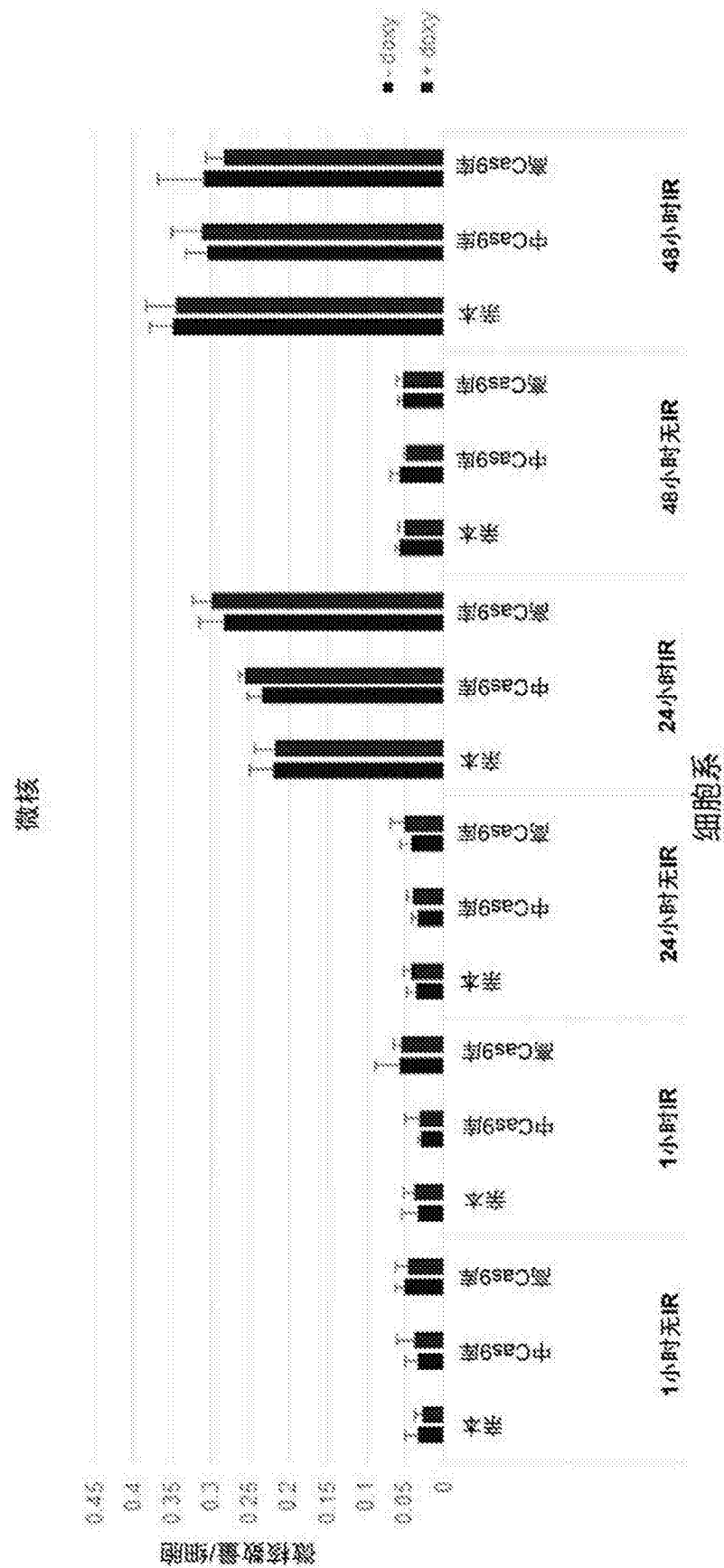


图19

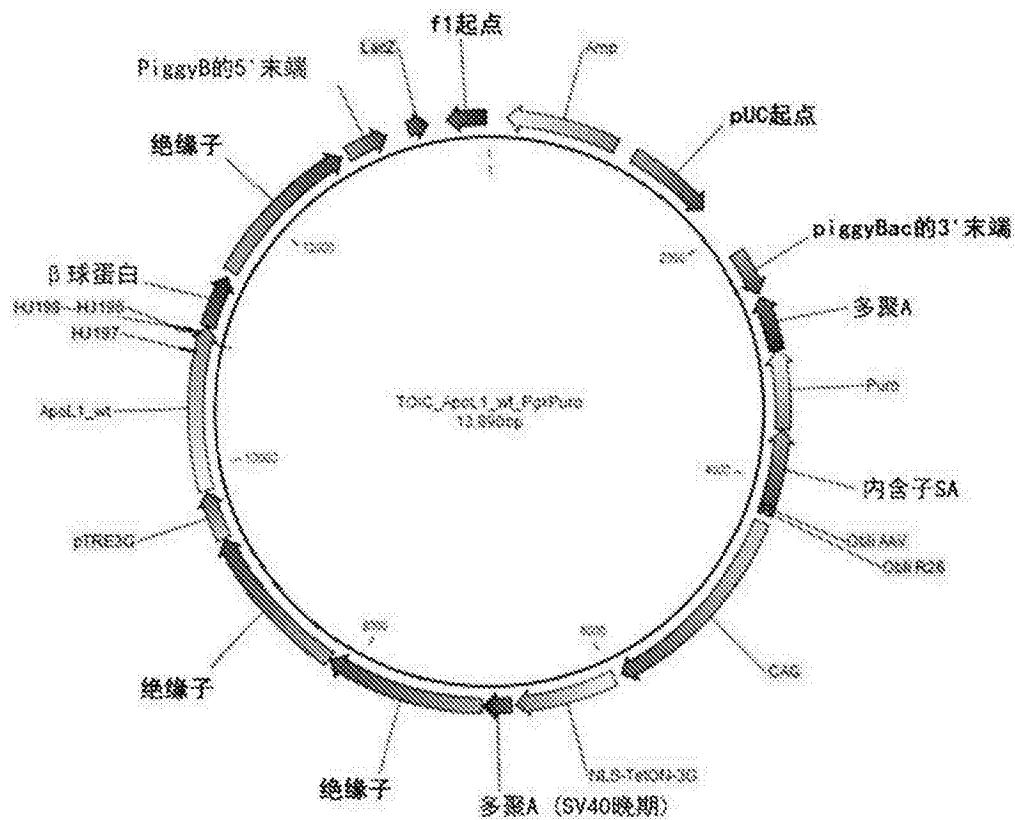


图20

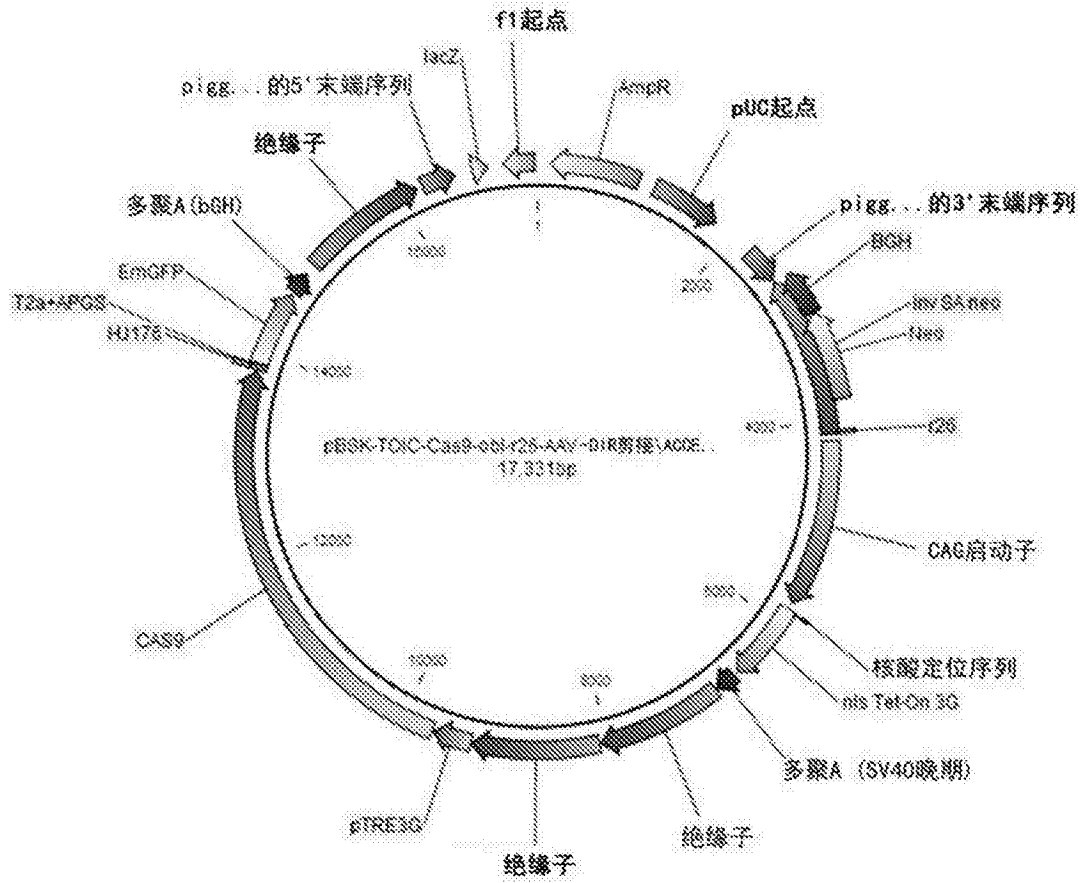
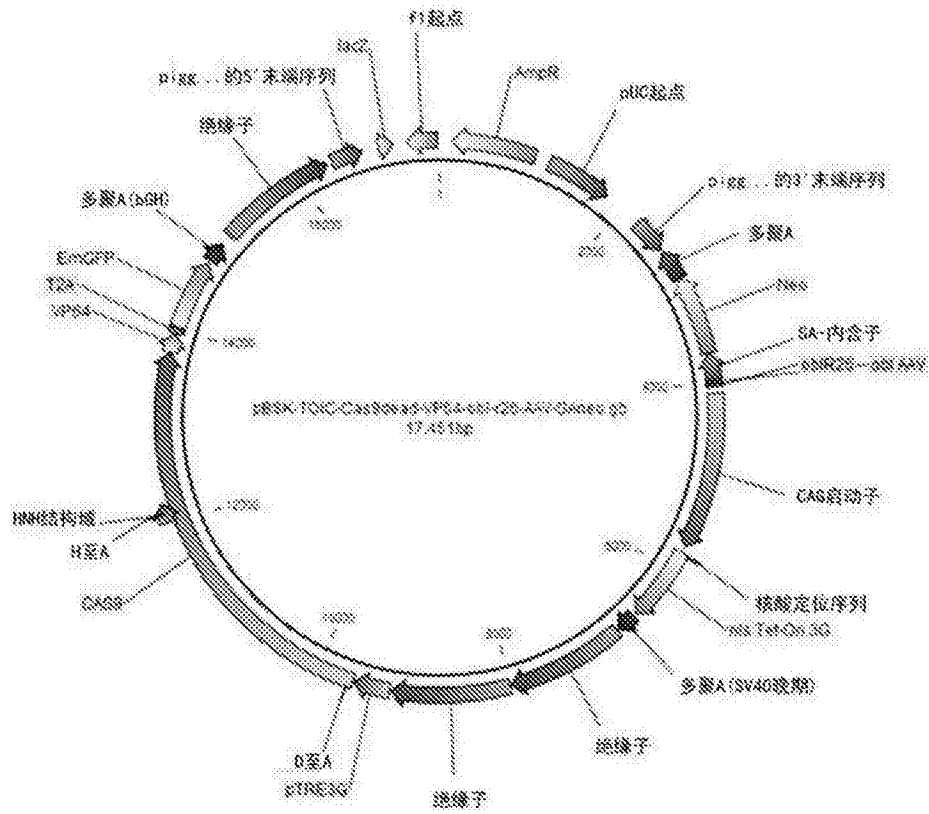


图21



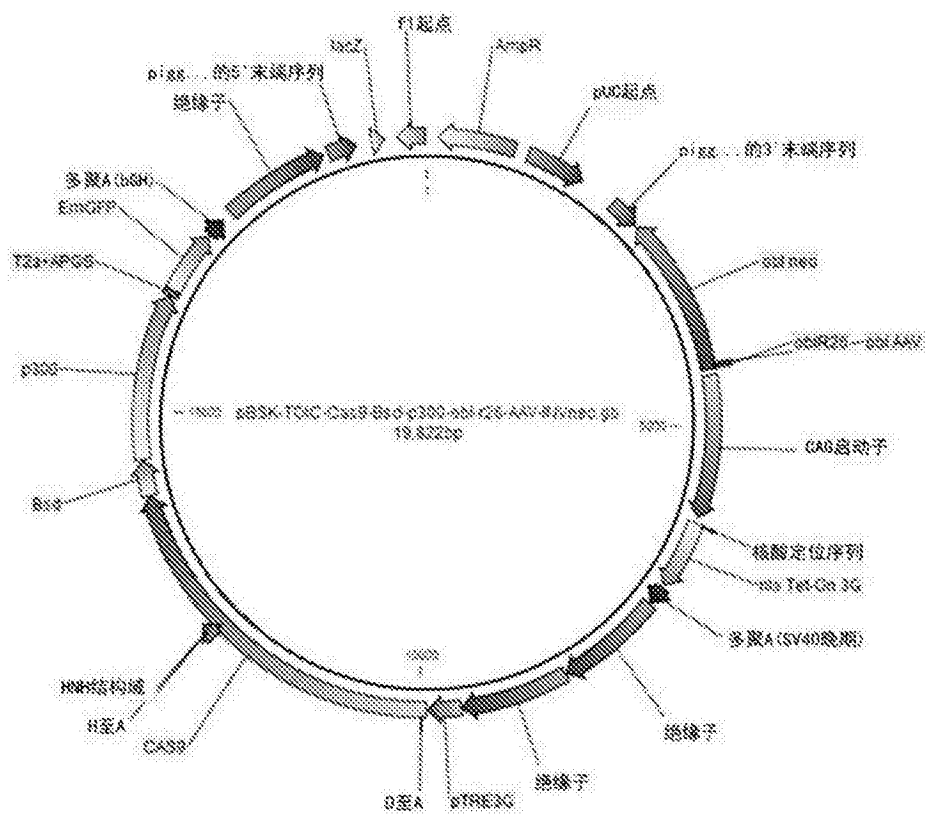


图 26

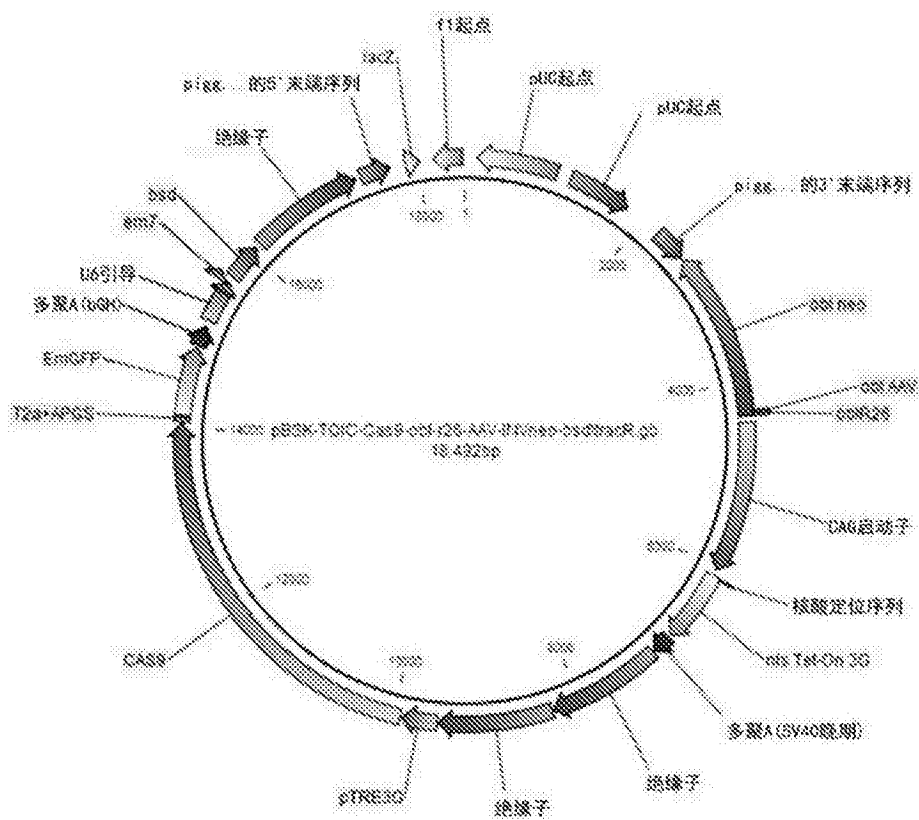


图27

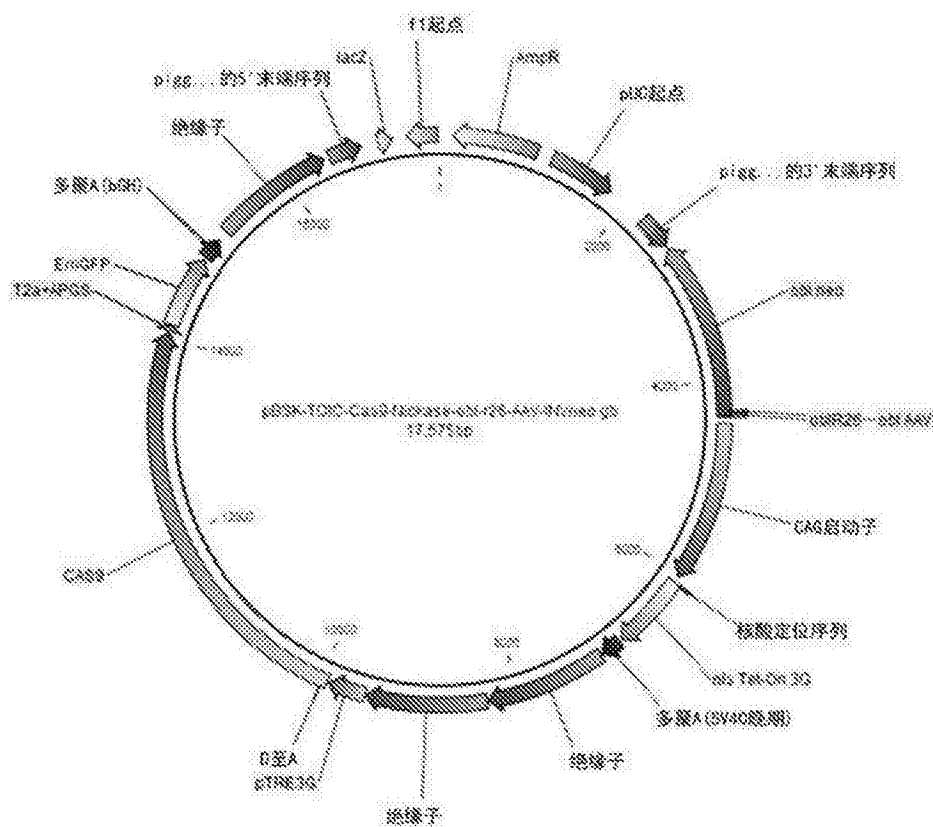


图 28