

(19)中华人民共和国国家知识产权局



## (12)发明专利申请



(10)申请公布号 CN 107988259 A

(43)申请公布日 2018. 05. 04

(21)申请号 201810028508.0

(22)申请日 2018.01.12

(71)申请人 中国科学院生物物理研究所  
地址 100101 北京市朝阳区大屯路15号

(72)发明人 翟宇佳 孙飞

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245

代理人 关畅 张立娜

(51)Int. Cl.

C12N 15/866(2006.01)

C12N 15/65(2006.01)

权利要求书6页 说明书26页  
序列表76页 附图6页

### (54)发明名称

SmartBac杆状病毒表达系统及其应用

### (57)摘要

本发明公开了一种SmartBac杆状病毒表达系统及其应用。该系统可包括受体质粒(含片段A或者含片段B和C)和供体质粒(含片段D);片段A含启动子、蛋白酶编码序列、蛋白酶切割位点、待表达目的物编码基因的插入区和终止序列;片段B含启动子、蛋白酶编码序列和终止序列;片段C含启动子、待表达目的物编码基因的插入区和终止序列;片段D含启动子、待表达目的物编码基因的插入区和终止序列。本发明还提供三种克隆的策略以实现分子量小于600kDa的蛋白质复合物表达、分子量大于600kDa的蛋白质复合物表达及高效筛选最适添加纯化标签的亚基。本发明对于在昆虫细胞中重组表达组分复杂、分子量较大的蛋白质复合物具有重要意义。

1. SmartBac杆状病毒表达系统,为如下任一:

(A) SmartBac杆状病毒表达系统A,包括受体质粒和供体质粒;所述受体质粒和所述供体质粒能够发生重组融合为一个质粒;

所述受体质粒为受体质粒甲和/或受体质粒乙;

所述受体质粒甲上含有DNA片段A;所述DNA片段A自上游到下游依次包含:启动子A、蛋白酶的编码基因序列、所述蛋白酶的切割位点识别序列、待表达目的物的编码基因的插入区,以及终止序列A;

所述受体质粒乙上含有DNA片段B和DNA片段C;所述DNA片段B自上游到下游依次包含:启动子B、所述蛋白酶的编码基因序列,以及终止序列B;所述DNA片段C自上游到下游依次包含:启动子C、待表达目的物的编码基因的插入区,以及终止序列C;

所述供体质粒上含有DNA片段D;所述DNA片段D自上游到下游依次包含:启动子D、待表达目的物的编码基因的插入区,以及终止序列D;

所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段;

(B) SmartBac杆状病毒表达系统B,包括受体质粒和供体质粒;所述受体质粒和所述供体质粒能够发生重组融合为一个质粒;

所述受体质粒上含有DNA片段A;所述DNA片段A自上游到下游依次包含:启动子A、待表达目的物的编码基因的插入区,以及终止序列A;

所述供体质粒为供体质粒甲和/或供体质粒乙;

所述供体质粒甲上含有DNA片段B;所述DNA片段B自上游到下游依次包含:启动子B、蛋白酶的编码基因序列、所述蛋白酶的切割位点识别序列、待表达目的物的编码基因的插入区,以及终止序列B;

所述供体质粒乙上含有DNA片段C和DNA片段D;所述DNA片段C自上游到下游依次包含:启动子C、所述蛋白酶的编码基因序列,以及终止序列C;所述DNA片段D自上游到下游依次包含:启动子D、待表达目的物的编码基因的插入区,以及终止序列D;

所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段;

(C) SmartBac杆状病毒表达系统C,包括所述SmartBac杆状病毒表达系统A或所述SmartBac杆状病毒表达系统B中的所述受体质粒;

(D) SmartBac杆状病毒表达系统D,与所述SmartBac杆状病毒表达系统A或所述SmartBac杆状病毒表达系统B或所述SmartBac杆状病毒表达系统C相比,差别仅在于:将其中的所述待表达目的物的编码基因的插入区替换为两两之间由所述蛋白酶的切割位点识别序列间隔开来的若干个插入区,每个插入区用于插入一个所述待表达目的物的编码基因。

2. 根据权利要求1所述的SmartBac杆状病毒表达系统,其特征在于:在所述SmartBac杆状病毒表达系统A中,所述DNA片段A中,在所述待表达目的物的编码基因的插入区和所述终止序列A之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段B中,在所述蛋白酶的编码基因序列和所述终止序列B之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段D中,在所述待表达目的物的编码基因的插入区和所述终止序列D之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列。

3. 根据权利要求1或2所述的SmartBac杆状病毒表达系统,其特征在于:所述受体质粒和所述供体质粒上均含有位点特异性重组酶的识别序列;

进一步地,所述位点特异性重组酶为Cre重组酶;

和/或

所述受体质粒中含有非条件型复制起点;所述供体质粒中含有条件型复制起点;

进一步地,所述非条件型复制起点为p15A复制起点;所述条件型复制起点为R6K  $\gamma$  复制起点;

和/或

所述受体质粒和所述供体质粒上含有不同的抗性筛选标记基因。

4. 根据权利要求1-3中任一所述的SmartBac杆状病毒表达系统,其特征在于:在所述SmartBac杆状病毒表达系统A中,所述启动子A为p6.9启动子;所述启动子B为GP64启动子;所述启动子C为p6.9启动子;所述启动子D为p10启动子。

5. 根据权利要求1-4中任一所述的SmartBac杆状病毒表达系统,其特征在于:所述待表达目的物的编码基因的插入区自上游到下游依次含有多克隆位点1、LacZ- $\alpha$ 表达盒和多克隆位点2;

进一步地,在所述SmartBac杆状病毒表达系统A中,在所述DNA片段D中,所述待表达目的物的编码基因的插入区中的所述多克隆位点1和所述多克隆位点2之间还含有PUC复制起点。

6. 根据权利要求1-5中任一所述的SmartBac杆状病毒表达系统,其特征在于:所述蛋白酶为TEV蛋白酶。

7. 根据权利要求6所述的SmartBac杆状病毒表达系统,其特征在于:在所述SmartBac杆状病毒表达系统A中,所述DNA片段A自上游到下游依次由p6.9启动子、N端被HA标记的TEV蛋白酶的编码基因序列、TEV蛋白酶切割位点的识别序列、Twin-Strep标签编码序列、肠激酶切割位点的识别序列、多克隆位点1、LacZ- $\alpha$ 表达盒、多克隆位点2、TEV蛋白酶切割位点的识别序列、荧光蛋白的编码基因序列和SV40pA信号序列连接而成;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段B自上游到下游依次由GP64启动子、N端被HA标记的TEV蛋白酶的编码基因序列、TEV蛋白酶切割位点的识别序列、荧光蛋白的编码基因序列和IE1ter信号序列连接而成;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段C自上游到下游依次由p6.9启动子、Twin-Strep标签编码序列、肠激酶切割位点的识别序列、多克隆位点1、LacZ- $\alpha$ 表达盒、多克隆位点2和SV40pA信号序列连接而成;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段D自上游到下游依次由p10启动子、10 $\times$ His标签编码序列、肠激酶切割位点的识别序列、多克隆位点1、PUC复制起点、LacZ- $\alpha$ 表达盒、多克隆位点2、TEV蛋白酶切割位点的识别序列、荧光蛋白的编码基因序列和HSV tk pA信号序列连接而成。

8. 根据权利要求7所述的SmartBac杆状病毒表达系统,其特征在于:在所述SmartBac杆

状病毒表达系统A中,所述DNA片段A的序列为SEQ ID No.1的第1235-3971位或者为SEQ ID No.2的第1235-3971位;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段B的序列为SEQ ID No.3的第1208-3252位或者为SEQ ID No.4的第1208-3249位;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段C的序列为SEQ ID No.3的第3306-4521位或者为SEQ ID No.4的第3303-4518位;

在所述SmartBac杆状病毒表达系统A中,所述DNA片段D的序列为SEQ ID No.5的第259-3188位或者为SEQ ID No.6的第259-3188位。

9. 根据权利要求8所述的SmartBac杆状病毒表达系统,其特征在于:在所述SmartBac杆状病毒表达系统A中,所述受体质粒甲为4V1G质粒和/或4V1R质粒;所述4V1G质粒的全序列为SEQ ID No.1;所述4V1R质粒的全序列为SEQ ID No.2;

所述受体质粒乙为5V1TG质粒和/或5V1TR质粒;所述5V1TG质粒的全序列为SEQ ID No.3;所述5V1TR质粒的全序列为SEQ ID No.4;

所述供体质粒为4V2G质粒和/或4V2R质粒;所述4V2G质粒的全序列为SEQ ID No.5;所述4V2R质粒的全序列为SEQ ID No.6。

10. 生物材料、应用或方法;

所述生物材料为如下任一:

(a) DNA片段组;所述DNA片段组由权利要求1-9任一中所述的DNA片段A、所述DNA片段B、所述DNA片段C和所述DNA片段D中的全部或部分组成;

(b) 哺乳动物细胞表达系统,是通过将权利要求1-9中任一所述的SmartBac杆状病毒表达系统中的所述受体质粒和所述供体质粒中被昆虫细胞识别的用于表达目的蛋白的启动子、终止序列替换为能被哺乳动物细胞识别的启动子和终止序列后得到的;

(c) DNA片段组;所述DNA片段组由步骤(b)所述哺乳动物细胞表达系统中的所述DNA片段A、所述DNA片段B、所述DNA片段C和所述DNA片段D中的全部或部分组成;

所述应用为所述生物材料或权利要求1-9中任一所述的SmartBac杆状病毒表达系统同时在同时表达n个目的物中的应用;所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段;所述n为2以上的自然数;

所述方法为如下方法I或方法II:

方法I:利用权利要求1-9中任一所述的SmartBac杆状病毒表达系统表达蛋白质复合物的方法,为如下(1)或(2):

(1) 当待表达的蛋白质复合物的分子量小于600kDa时,所述方法包括如下步骤:

(a1) 将组成所述蛋白质复合物的全部亚基的编码基因分成总长度接近的两组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,两组得到两条融合基因;

(a2) 将所述两条融合基因中的一条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒;将所述两条融合基因中的另一条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒;

(a3) 将所述重组受体质粒和所述重组供体质粒在所述位点特异性重组酶的作用下融合为一个质粒,即得转移质粒;

(a4) 将所述转移质粒转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒;

(a5) 将所述重组杆状病毒穿梭质粒感染昆虫细胞,得到重组杆状病毒;

(a6) 将所述重组杆状病毒感染昆虫细胞,从而实现所述蛋白质复合物的表达;

(2) 当待表达的蛋白质复合物的分子量在600kDa以上时,所述方法为如下(B)或(C):

(B) 包括如下步骤:

(b1) 将组成所述蛋白质复合物的全部亚基的编码基因分成M组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,M组得到M条融合基因;其中, $M=2m$ , $m$ 为2以上的自然数;

(b2) 将所述M条融合基因中的第1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒1;将所述M条融合基因中的第2条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒1;将所述M条融合基因中的第3条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒2;将所述M条融合基因中的第4条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒2;以此类推;将所述M条融合基因中的第M-1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒m;将所述M条融合基因中的第M条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒m;

(b3) 将所述重组受体质粒1和所述重组供体质粒1在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒1;将所述重组受体质粒2和所述重组供体质粒2在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒2;以此类推;将所述重组受体质粒m和所述重组供体质粒m在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒m;

(b4) 将所述转移质粒1转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒1;将所述转移质粒2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒2;以此类推;将所述转移质粒m转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒m;

(b5) 将所述重组杆状病毒穿梭质粒1感染昆虫细胞,得到重组杆状病毒1;将所述重组杆状病毒穿梭质粒2感染昆虫细胞,得到重组杆状病毒2;以此类推;将所述重组杆状病毒穿梭质粒m感染昆虫细胞,得到重组杆状病毒m;

(b6) 将所述重组杆状病毒1、所述重组杆状病毒2、以此类推、所述重组杆状病毒m共同感染昆虫细胞,从而实现所述蛋白质复合物的表达;

(C) 包括如下步骤:

(c1) 将组成所述蛋白质复合物的全部亚基的编码基因分成M组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,M组得到M条融合基因;其中, $M=2m-1$ , $m$ 为2以上的自然数;

(c2) 将所述M条融合基因中的第1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒1;将所述M条融合基因中的第2条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒1;将所述M条融合基因中的第3

条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒2;将所述M条融合基因中的第4条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒2;以此类推;将所述M条融合基因中的第M-2条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒(M-1)/2;将所述M条融合基因中的第M-1条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒(M-1)/2;将所述M条融合基因中的第M条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒m;

(c3) 将所述重组受体质粒1和所述重组供体质粒1在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒1;将所述重组受体质粒2和所述重组供体质粒2在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒2;以此类推;将所述重组受体质粒(M-1)/2和所述重组供体质粒(M-1)/2在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒(M-1)/2;所述重组受体质粒m即为转移质粒m;

(c4) 将所述转移质粒1转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒1;将所述转移质粒2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒2;以此类推;将所述转移质粒(M-1)/2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒(M-1)/2;将所述转移质粒m转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒m;

(c5) 将所述重组杆状病毒穿梭质粒1感染昆虫细胞,得到重组杆状病毒1;将所述重组杆状病毒穿梭质粒2感染昆虫细胞,得到重组杆状病毒2;以此类推;将所述重组杆状病毒穿梭质粒(M-1)/2感染昆虫细胞,得到重组杆状病毒(M-1)/2;将所述重组杆状病毒穿梭质粒m感染昆虫细胞,得到重组杆状病毒m;

(c6) 将所述重组杆状病毒1、所述重组杆状病毒2、以此类推、所述重组杆状病毒m共同感染昆虫细胞,从而实现所述蛋白质复合物的表达;

方法II:一种利用权利要求1-9中任一所述的SmartBac杆状病毒表达系统筛选待表达蛋白复合物中适于添加纯化标签的亚基的方法,为如下(3)或(4):

(3) 当待表达的蛋白质复合物的分子量小于600kDa时,所述方法包括如下步骤:

(a' 1) 按照所述(a1)-(a3)的步骤构建得到一个转移质粒;所述转移质粒不含有纯化标签编码序列;

(a' 2) 将所述待表达蛋白复合物的全部亚基的编码基因分别克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,并能够与所述纯化标签编码序列融合表达,对应每个亚基各得到一个重组受体质粒,所述重组受体质粒即为转移质粒;

(a' 3) 将步骤(a' 1)和(a' 2)中的所有转移质粒分别转化含有杆状病毒穿梭质粒的受体菌,对应每个转移质粒均得到一个重组杆状病毒穿梭质粒;

(a' 4) 将步骤(a' 3)中的所有重组杆状病毒穿梭质粒分别感染昆虫细胞,得到各自对应的重组杆状病毒;

(a' 5) 将对应步骤(a' 2)中各转移质粒的所有所述重组杆状病毒中每一种分别与对应步骤(a' 1)中所述转移质粒的所述重组杆状病毒组合,共同感染昆虫细胞,然后根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化,从而确定所述蛋白复合物中适于添加所述纯化标签的亚基;

(4) 当待表达的蛋白质复合物的分子量在600kDa以上时,所述方法为包括如下步骤:

(b' 1) 按照所述 (b1) - (b3) 的步骤或者所述 (c1) - (c3) 的步骤构建得到m个转移质粒;所述转移质粒不含有纯化标签编码序列;

(b' 2) 将所述待表达蛋白复合物的全部亚基的编码基因分别克隆入所述受体载体的所述待表达目的物的编码基因的插入区,并能够与所述纯化标签编码序列融合表达,对应每个亚基各得到一个重组受体质粒,所述重组受体质粒即为转移质粒;

(b' 3) 将步骤 (b' 1) 和 (b' 2) 中的所有转移质粒分别转化含有杆状病毒穿梭质粒的受体菌,对应每个转移质粒均得到一个重组杆状病毒穿梭质粒;

(b' 4) 将步骤 (b' 3) 中的所有重组杆状病毒穿梭质粒分别感染昆虫细胞,得到各自对应的重组杆状病毒;

(b' 5) 将对应步骤 (b' 2) 中各转移质粒的所有所述重组杆状病毒中的每一种分别与对应步骤 (b' 1) 中所述m个转移质粒的m个所述重组杆状病毒组合,共同感染昆虫细胞,然后根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化,从而确定所述蛋白复合物中适于添加所述纯化标签的亚基。

## SmartBac杆状病毒表达系统及其应用

### 技术领域

[0001] 本发明属于生物技术领域,涉及一种杆状病毒表达系统及其应用,具体涉及一种新型的在昆虫细胞中同时表达多种蛋白质的杆状病毒表达系统SmartBac系统及其在表达超大分子复合物上的应用。

### 背景技术

[0002] 杆状病毒表达系统(Baculovirus Expression System,BVES)是在昆虫细胞中高效表达外源蛋白质的有力工具,具有安全性好、表达水平高、可进行翻译后加工等优点。由于杆状病毒的基因组庞大,外源基因的克隆不能通过酶切连接的方法直接插入,所以人们对杆状病毒基因组进行改造,并构建与之相匹配的转移载体,使两者重组为能够感染昆虫细胞的含外源基因的重组杆状病毒。在目前广泛使用的Bac to Bac系统中,杆状病毒穿梭载体(Bacmid)既可以在大肠杆菌中复制,又可以感染鳞翅目昆虫细胞,其在大肠杆菌中可以与含外源基因的匹配转移载体发生Tn7位点特异性重组。重组得到的杆状病毒穿梭载体能在大肠杆菌中高效复制,被提纯后可用于转染昆虫细胞。

[0003] 杆状病毒基因组容量大,可以在转移载体上插入多个开放阅读框(ORF,Open Reading Frame),继而使得到的重组杆状病毒穿梭载体在昆虫细胞中同时表达多个蛋白质。这也是目前在昆虫细胞中实现多种蛋白质共表达的通用思路。例如质粒pFastBac-Dual含有两个头对头放置的开放阅读框(ORF,Open Reading Frame),一个ORF以p10启动子起始,HSV tk polyadenylation(简称HSV tk pA)信号序列终止;另一个ORF以polyhedrin启动子起始,SV40polyadenylation(简称SV40pA)信号序列终止。质粒pFastBac-Dual可以作为转移载体,但只能同时表达两种蛋白质,若要表达两种以上的蛋白质,则需要另构建其它转移载体。例如要表达由四种不同亚基构成的蛋白质复合物,需要经过如下步骤:(1)构建重组质粒pFastBac-Dual-A-B(含A基因和B基因的质粒pFastBac-Dual)和重组质粒pFastBac-Dual-C-D(含C基因和D基因的质粒pFastBac-Dual);(2)将步骤(1)构建的重组质粒分别与Bacmid进行重组,得到两种重组Bacmid;(3)将步骤(2)得到的两种重组Bacmid分别转染昆虫细胞,得到两种病毒。(4)用步骤(3)得到的两种病毒同时感染昆虫细胞,实现四种亚基的共表达。这种多病毒共感染方法的蛋白表达量通常要低于单一病毒感染细胞的蛋白表达量。

[0004] 近几年流行的MultiBac系统考虑到这一点,只用一种重组Bacmid感染细胞,进行蛋白质复合体的表达。该系统的表达思路也是一个ORF表达一种蛋白质。通过受体质粒与供体质粒上LoxP位点介导的重组,实现质粒的融合,将不同来源的ORF整合到一个转移载体上,然后与Bacmid重组,实现多种蛋白质的共表达。这种方法的局限性为要构建多种含目的基因的供体质粒与受体质粒,供体质粒与受体质粒要经过多次整合与筛选,才能得到最终用于表达的转移载体,耗时耗力。

[0005] 上述两种方法除了分子克隆操作繁琐外,还含有以下几种缺陷:一是在表达过程中无法控制各个亚基的拷贝数,最终无法纯化到性质较为均一的蛋白质复合体;二是在病

毒感染过程中无法判断目的蛋白是否表达。

[0006] 为了解决上述问题,本发明的发明人所在团队开发了一种新型的在昆虫细胞中同时表达多种蛋白质的方法(后面简称方法1)并申请了专利(申请号:201610248592.8)。在方法1中,编码TEV酶的基因和编码各蛋白质的基因被串联在一个ORF里:首先以一条多肽链的形式被表达出来;然后,位于多肽链N端的TEV酶会通过各蛋白之间的TEV酶剪切位点(TCS),将各个蛋白质从多肽链上剪切下来,以实现多种蛋白质的同时表达和体内组装。为了在昆虫细胞中监测这些ORF的表达情况,将编码不同颜色荧光蛋白的基因通过TCS连接在多蛋白基因后面,与多个蛋白以一条长的多肽链共同表达。这一想法在发明人构建的pFBD-mCEG载体中得到了实现。使用上述方法和载体,发明人构建了重组质粒 pFBD-mCEG-COPI,在昆虫细胞中成功的表达了由7个不同亚基构成的,具有活性的人源 COPI蛋白。

[0007] 但是方法1仍然具有一定的使用局限性。首先,方法1要求将编码蛋白质亚基的各个基因都串联起来,插入在一个载体上进行蛋白质表达工作。而对于那些分子量较大的蛋白质复合体,其对应的DNA序列也较长。无论是通过基因合成的方法还是在实验室利用经典分子克隆的手段,构建含有较长基因片段的载体(载体大小在20kb左右)都是比较困难和耗时的。其次,将荧光蛋白接在多肽链末端融合表达,虽然可以很好的监测多肽链的表达情况,但是偶尔也会出现荧光蛋白与多肽链之间剪切不完全的情况。这必然会影响蛋白质复合体在表达细胞内的正常组装。第三,pFBD-mCEG载体在大肠杆菌中的复制起点是高拷贝数的复制起点。利用这个载体构建出来的较大质粒在大肠杆菌中进行复制时,由于其自身高拷贝数复制,会给菌体带来较大的压力,而细菌可能会启动其它代谢途径来对抗这种压力。最终的结果是质粒稳定性差,极有可能出现基因丢失的现象。

[0008] 此外,在体外重组表达含有多个亚基的蛋白质复合体,常常需要筛选最适于添加纯化标签的亚基。当我们重组表达蛋白质复合体时,通常将一种纯化标签加在某一个亚基上,利用亲和层析的方法对含有这一标签的亚基进行纯化。由于这一亚基与组成复合体的其它亚基之间有较为紧密的相互作用,可以将整个蛋白质复合体纯化出来。但是纯化标签加在不同的亚基上,对于纯化整个复合体常常有着不同的效果。如果不采用有效的克隆策略,会花费大量的时间和精力来构建表达载体。假设我们要用pFBD-mCEG载体表达一个含8个亚基的、分子量在600kDa左右蛋白质复合体,如果我们用简单的试错法,需要构建8种大小在25kb左右的表达载体,每种表达载体的区别就是带有纯化标签的亚基是不同的。这对于任何分子生物学实验室都不是非常轻松的工作。

## 发明内容

[0009] 为了有效的解决上述问题,本发明提供了一种新型的SmartBac杆状病毒表达系统,并提出了三种克隆的策略以实现分子量的在600kDa以下的蛋白质复合体表达、分子量在600kDa 以上的蛋白质复合体的表达以及高效筛选最适合添加纯化标签的亚基。

[0010] 本发明所提供的SmartBac杆状病毒表达系统,为如下(A)-(D)中任一:

[0011] (A) SmartBac杆状病毒表达系统A,包括受体质粒和供体质粒;所述受体质粒和所述供体质粒能够发生重组融合为一个质粒;

[0012] 所述受体质粒为受体质粒甲和/或受体质粒乙;

[0013] 所述受体质粒甲上含有DNA片段A;所述DNA片段A自上游到下游依次包含:启动子

A、蛋白酶的编码基因序列、所述蛋白酶的切割位点识别序列、待表达目的物的编码基因的插入区,以及终止序列A;

[0014] 所述受体质粒乙上含有DNA片段B和DNA片段C;所述DNA片段B自上游到下游依次包含:启动子B、所述蛋白酶的编码基因序列,以及终止序列B;所述DNA片段C自上游到下游依次包含:启动子C、待表达目的物的编码基因的插入区,以及终止序列C;

[0015] 所述供体质粒上含有DNA片段D;所述DNA片段D自上游到下游依次包含:启动子 D、待表达目的物的编码基因的插入区,以及终止序列D;

[0016] 所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段。

[0017] (B) SmartBac杆状病毒表达系统B,包括受体质粒和供体质粒;所述受体质粒和所述供体质粒能够发生重组融合为一个质粒;

[0018] 所述受体质粒上含有DNA片段A;所述DNA片段A自上游到下游依次包含:启动子 A、待表达目的物的编码基因的插入区,以及终止序列A;

[0019] 所述供体质粒为供体质粒甲和/或供体质粒乙;

[0020] 所述供体质粒甲上含有DNA片段B;所述DNA片段B自上游到下游依次包含:启动子 B、蛋白酶的编码基因序列、所述蛋白酶的切割位点识别序列、待表达目的物的编码基因的插入区,以及终止序列B;

[0021] 所述供体质粒乙上含有DNA片段C和DNA片段D;所述DNA片段C自上游到下游依次包含:启动子C、所述蛋白酶的编码基因序列,以及终止序列C;所述DNA片段D自上游到下游依次包含:启动子D、待表达目的物的编码基因的插入区,以及终止序列D;

[0022] 所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段。

[0023] (C) SmartBac杆状病毒表达系统C,包括所述SmartBac杆状病毒表达系统A或所述SmartBac杆状病毒表达系统B中的所述受体质粒;可不含供体质粒。

[0024] (D) SmartBac杆状病毒表达系统D,与所述SmartBac杆状病毒表达系统A或所述SmartBac杆状病毒表达系统B或所述SmartBac杆状病毒表达系统C相比,差别仅在于:将其中的所述待表达目的物的编码基因的插入区替换为两两之间由所述蛋白酶的切割位点识别序列间隔开来的若干个插入区,每个插入区用于插入一个所述待表达目的物的编码基因。

[0025] 所述受体质粒上还含有用于产生重组杆状病毒的小型Tn7元件侧翼。

[0026] 进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段A中,在所述待表达目的物的编码基因的插入区和所述终止序列A之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列。

[0027] 进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段B中,在所述蛋白酶的编码基因序列和所述终止序列B之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列。

[0028] 进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段D中,在所述待表达目的物的编码基因的插入区和所述终止序列D之间自上游到下游还依次含有所述蛋白酶的切割位点识别序列和荧光蛋白的编码基因序列。

[0029] 在本发明中,所述受体质粒和所述供体质粒上均含有位点特异性重组酶的识别序列;依靠所述位点特异性重组酶的识别序列,所述受体质粒和所述供体质粒可以融合为一

个大的质粒。

[0030] 进一步地,所述位点特异性重组酶具体为Cre重组酶;相应的,所述位点特异性重组酶的识别序列为loxP位点序列。当然所述位点特异性重组酶也可以为其他具有相同或类似功能的重组酶。

[0031] 进一步地,所述受体质粒中含有非条件型复制起点;所述供体质粒中含有条件型复制起点。

[0032] 更加具体地,在本发明中,所述非条件型复制起点具体为p15A复制起点(p15A复制起点允许质粒以低拷贝在普通大肠杆菌克隆菌株中繁殖,它更好的维持了大质粒的稳定性);所述条件型复制起点具体为R6K  $\gamma$  复制起点(质粒仅能在含有pir基因的细菌宿主中繁殖)。

[0033] 进一步地,所述受体质粒和所述供体质粒上含有不同的抗性筛选标记基因。

[0034] 更加具体地,在本发明中,所述受体质粒上携带有氨苄青霉素抗性基因和庆大霉素抗性基因;所述供体质粒上携带有卡那霉素抗性基因。

[0035] 进一步地,在本发明中,对于所述SmartBac杆状病毒表达系统A来说,所述启动子A为p6.9启动子;所述启动子B为GP64启动子;所述启动子C为p6.9启动子;所述启动子D 为p10启动子。

[0036] 其中,p6.9启动子与常规使用的极晚期多角体蛋白polyhedrin启动子相比,在感染的早期阶段驱动表达,这时的细胞状态与病毒感染晚期相比更好,可以避免表达的外源蛋白质的聚集。

[0037] 进一步地,在本发明中,对于所述SmartBac杆状病毒表达系统A来说,所述终止序列 A为SV40pA信号序列;所述终止序列B为IE1ter信号序列;所述终止序列C为SV40pA 信号序列;所述终止序列D为HSV tk pA信号序列。

[0038] 进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述待表达目的物的编码基因的插入区自上游到下游依次含有多克隆位点1、LacZ- $\alpha$ 表达盒和多克隆位点2。

[0039] 其中,对于所述SmartBac杆状病毒表达系统A来说,在所述DNA片段D中,所述待表达目的物的编码基因的插入区中的所述多克隆位点1和所述多克隆位点2之间还含有PUC复制起点。

[0040] 在本发明中,所述蛋白酶具体为TEV蛋白酶。当然所述蛋白酶也可以为具有相同或类似功能的其他蛋白酶。

[0041] 更进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段A自上游到下游依次由p6.9启动子(SEQ ID No.1的第1235-1329位或SEQ ID No.2的第1235-1329位)、N端被HA标记的TEV蛋白酶的编码基因序列(SEQ ID No.1的第1339-2097位或SEQ ID No.2的第1339-2097位)、TEV蛋白酶切割位点的识别序列(TCS)(SEQ ID No.1的第2098-2118 位或SEQ ID No.2的第2098-2118位)、Twin-Strep标签编码序列(SEQ ID No.1的第2119-2205位或SEQ ID No.2的第2119-2205位)、肠激酶切割位点的识别序列(SEQ ID No.1的第2206-2220位或SEQ ID No.2的第2206-2220位)、多克隆位点1(SEQ ID No.1的第2222-2249 位或SEQ ID No.2的第2222-2249位)、LacZ- $\alpha$ 表达盒(SEQ ID No.1的第2250-2790位或SEQ ID No.2的第2250-2790位)、多克隆位点2(SEQ ID No.1的第2799-2837位或SEQ ID No.2 的第2799-2843位)、TEV蛋白酶切割位点的识别序列(TCS)(SEQ ID No.1的第2839-2859 位或

SEQ ID No.2的第2845-2865位)、荧光蛋白的编码基因序列(SEQ ID No.1的第2869-3585位所示的EGFP或SEQ ID No.2的第2875-3585位所示的TagRFP)和SV40pA信号序列(SEQ ID No.1的第3731-3971位或SEQ ID No.2的第3731-3971位)连接而成。

[0042] 更进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段B自上游到下游依次由GP64启动子(SEQ ID No.3的第3047-3252位或SEQ ID No.4的第3044-3249位)、N端被HA标记的TEV蛋白酶的编码基因序列(SEQ ID No.3的第2275-3033位或SEQ ID No.4的第2272-3030位)、TEV蛋白酶切割位点的识别序列(TCS)(SEQ ID No.3的第2254-2274位或SEQ ID No.4的第2251-2271位)、荧光蛋白的编码基因序列(SEQ ID No.3的第1531-2247位所示的EGFP或SEQ ID No.4的第1531-2244位所示的TagRFP)和IE1ter信号序列(SEQ ID No.3的第1208-1514位或SEQ ID No.4的第1208-1514位)连接而成。

[0043] 更进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段C自上游到下游依次由p6.9启动子(SEQ ID No.3的第3306-3400位或SEQ ID No.4的第3303-3397位)、Twin-Strep标签编码序列(SEQ ID No.3的第3419-3505位或SEQ ID No.4的第3416-3502位)、肠激酶切割位点的识别序列(SEQ ID No.3的第3506-3520位或SEQ ID No.4的第3503-3517位)、多克隆位点1(SEQ ID No.3的第3522-2549位或SEQ ID No.4的第3519-3546位)、LacZ- $\alpha$ 表达盒(SEQ ID No.3的第3550-4090位或SEQ ID No.4的第3547-4087位)、多克隆位点2(SEQ ID No.3的第4099-4149位或SEQ ID No.4的第4096-4146位)和SV40pA信号序列(SEQ ID No.3的第4281-4521位或SEQ ID No.4的第4278-4528位)连接而成。

[0044] 更进一步地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段D自上游到下游依次由p10启动子(SEQ ID No.5的第259-368位或SEQ ID No.6的第259-368位)、10 $\times$ His标签编码序列(SEQ ID No.5的第397-426位或SEQ ID No.6的第397-426位)、肠激酶切割位点的识别序列(SEQ ID No.5的第436-450位或SEQ ID No.6的第436-450位)、多克隆位点1(SEQ ID No.5的第452-473位或SEQ ID No.6的第452-473位)、PUC复制起点(SEQ ID No.5的第644-1232位或SEQ ID No.6的第644-1232位)、LacZ- $\alpha$ 表达盒(SEQ ID No.5的第1443-1983位或SEQ ID No.6的第1443-1983位)、多克隆位点2(SEQ ID No.5的第1992-2030位或SEQ ID No.6的第1992-2036位)、TEV蛋白酶切割位点的识别序列(TCS)(SEQ ID No.5的第2032-2052位或SEQ ID No.6的第2038-2058位)、荧光蛋白的编码基因序列(SEQ ID No.5的第2062-2778位所示的EGFP或SEQ ID No.6的第2068-2778位所示的TagRFP)和HSV tk pA信号序列(SEQ ID No.5的第2907-3188位或SEQ ID No.6的第2907-3188位)连接而成。

[0045] 其中,所述荧光蛋白在本发明中具体为绿色荧光蛋白或红色荧光蛋白。

[0046] 更加具体地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段A的序列具体为SEQ ID No.1的第1235-3971位或者为SEQ ID No.2的第1235-3971位。

[0047] 更加具体地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段B的序列为SEQ ID No.3的第1208-3252位或者为SEQ ID No.4的第1208-3249位。

[0048] 更加具体地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段C的序列为SEQ ID No.3的第3306-4521位或者为SEQ ID No.4的第3303-4518位。

[0049] 更加具体地,对于所述SmartBac杆状病毒表达系统A来说,所述DNA片段D的序列为SEQ ID No.5的第259-3188位或者为SEQ ID No.6的第259-3188位。

[0050] 另外,对于所述SmartBac杆状病毒表达系统A来说,在所述受体质粒和所述供体质粒中的p6.9启动子和p10启动子区域的两侧分别存在几个单一的限制性位点,这样当需要时,可以将现有的启动子剪切下来,替换为其它的可以在昆虫细胞中启动蛋白质表达的启动子。

[0051] 在本发明的具体实施例中,对于所述SmartBac杆状病毒表达系统A来说,所述受体质粒甲为4V1G质粒和/或4V1R质粒;所述4V1G质粒的全序列为SEQ ID No.1;所述4V1R质粒的全序列为SEQ ID No.2。所述受体质粒乙为5V1TG质粒和/或5V1TR质粒;所述5V1TG质粒的全序列为SEQ ID No.3;所述5V1TR质粒的全序列为SEQ ID No.4。所述供体质粒为4V2G质粒和/或4V2R质粒;所述4V2G质粒的全序列为SEQ ID No.5;所述4V2R质粒的全序列为SEQ ID No.6。

[0052] 其中,SEQ ID No.1的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1224位为多克隆位点BP,第1235-1329位为p6.9启动子,第1339-2097位为N端被HA标记的TEV蛋白酶的编码基因序列,第2098-2118位为TEV蛋白酶切割位点的识别序列(TCS),第2119-2205位为Twin-Strep标签编码序列,第2206-2220位为肠激酶切割位点的识别序列,第2222-2249位为多克隆位点1,第2250-2790位为LacZ- $\alpha$ 表达盒,第2799-2837位为多克隆位点2,第2839-2859位为TEV蛋白酶切割位点的识别序列(TCS),第2862-2867位为SalI识别位点,第2869-3585位为荧光蛋白EGFP的编码基因序列,第3731-3971为SV40pA信号序列,第4000-4165位为Tn7L元件,第4349-4804位为f1复制起点,第4936-5796位为氨苄青霉素抗性蛋白的编码基因序列,第6067-6612位为p15A复制起点。

[0053] SEQ ID No.2的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1224位为多克隆位点BP,第1235-1329位为p6.9启动子,第1339-2097位为N端被HA标记的TEV蛋白酶的编码基因序列,第2098-2118位为TEV蛋白酶切割位点的识别序列(TCS),第2119-2205位为Twin-Strep标签编码序列,第2206-2220位为肠激酶切割位点的识别序列,第2222-2249位为多克隆位点1,第2250-2790位为LacZ- $\alpha$ 表达盒,第2799-2843位为多克隆位点2,第2845-2865位为TEV蛋白酶切割位点的识别序列(TCS),第2868-2873位为SalI识别位点,第2875-3585位为荧光蛋白TagRFP的编码基因序列,第3731-3971为SV40pA信号序列,第4000-4165位为Tn7L元件,第4349-4804位为f1复制起点,第4831-4935位为Amp启动子,第4936-5796位为氨苄青霉素抗性蛋白的编码基因序列,第6067-6612位为p15A复制起点。

[0054] SEQ ID No.3的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1193位为I-CeuI酶切位点,第1202-1207位为AatII酶切位点,第1208-1514位为IE1ter信号序列,第1515-1522位为FseI酶切位点,第1531-2247位为荧光蛋白EGFP的编码基因序列,第2248-2253位为SmaI酶切位点,第2254-2274位为TEV蛋白酶切割位点的识别序列,第2275-3033位为N端被HA标记的TEV蛋白酶的编码基因序列,第3034-3041位为SwaI酶切位点,第3047-3252位为GP64启动子,第3290-3295位为BspEI酶切位点,第3306-3400位为P6.9启动子,第3407-3418位为多克隆位点AP,第3419-3505位为Twin-Strep标签编码序列,第3506-3520位为肠激酶切割位点的识别序列,第3522-2549位为多克隆位点1,第3550-4090位为LacZ- $\alpha$ 表达盒,第4099-4149位为多

克隆位点2,第4281-4521位为SV40pA信号序列,第4550-4715位为Tn7L元件,第 4899-5354 位为f1复制起点,第5486-6346位为氨苄青霉素抗性蛋白的编码基因序列,第 6617-7162位为p15A复制起点。

[0055] SEQ ID No.4的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1193位为I-CeuI酶切位点,第1202-1207位为AatII酶切位点,第1208-1514位为IE1ter信号序列,第1515-1522位为FseI酶切位点,第1531-2244位为荧光蛋白TagRFP的编码基因序列,第2245-2250位为SmaI酶切位点,第2251-2271位为TEV蛋白酶切割位点的识别序列,第2272-3030位为N端被HA标记的TEV 蛋白酶的编码基因序列,第3031-3038位为SwaI酶切位点,第3044-3249位为GP64启动子,第3287-3292位为BspEI酶切位点,第3303-3397位为P6.9启动子,第3404-3415位为多克隆位点AP,第3416-3502位为Twin-Strep标签编码序列,第3503-3517位为肠激酶切割位点的识别序列,第3519-3546位为多克隆位点1,第3547-4087位为LacZ- $\alpha$ 表达盒,第4096-4146 位为多克隆位点2,第4278-4518位为SV40pA信号序列,第4547-4712位为Tn7L元件,第 4896-5351位为f1复制起点,第5483-6343位为氨苄青霉素抗性蛋白的编码基因序列,第 6614-7159位为p15A复制起点。

[0056] SEQ ID No.5的第169-202位为LoxP位点,第208-215位为NotI酶切位点,第259-368 位为p10启动子,第376-393位为多克隆位点AP,第397-426位为10 $\times$ His标签编码序列,第436-450位为肠激酶切割位点的识别序列,第452-473位为多克隆位点1,第644-1232位为PUC复制起点,第1443-1983位为LacZ- $\alpha$ 表达盒,第1992-2030位为多克隆位点2,第 2032-2052位为TEV蛋白酶切割位点的识别序列,第2055-2060位为SacI识别位点,第 2062-2778位为荧光蛋白EGFP的编码基因序列,第2907-3188位为HSV-tk pA信号序列,第 3224-3579位为R6K  $\gamma$  复制起点,第3940-4734位为卡那霉素抗性蛋白编码序列。

[0057] SEQ ID No.6的第169-202位为LoxP位点,第208-215位为NotI酶切位点,第259-368 位为p10启动子,第376-393位为多克隆位点AP,第397-426位为10 $\times$ His标签编码序列,第436-450位为肠激酶切割位点的识别序列,第452-473位为多克隆位点1,第644-1232位为PUC复制起点,第1443-1983位为LacZ- $\alpha$ 表达盒,第1992-2036位为多克隆位点2,第 2038-2058位为TEV蛋白酶切割位点的识别序列,第2061-2066位为SacI识别位点,第 2068-2778位为荧光蛋白TagRFP的编码基因序列,第2907-3188位为HSV-tk pA信号序列,第3224-3579位为R6K  $\gamma$  复制起点,第3940-4734位为卡那霉素抗性蛋白编码序列。

[0058] 所述SmartBac杆状病毒表达系统具体可由所述4V1G质粒、所述4V1R质粒、所述5V1TG 质粒、所述5V1TR质粒、所述4V2G质粒和所述4V2R质粒组成。

[0059] 当然,由前文所述的DNA片段A、所述DNA片段B、所述DNA片段C和所述DNA 片段D中的全部或部分组成的DNA片段组也属于本发明的保护范围。

[0060] 将所述SmartBac杆状病毒表达系统中的所述受体质粒和所述供体质粒中被昆虫细胞识别的用于表达目的蛋白的启动子、终止序列替换为能被哺乳动物细胞识别的启动子和终止序列后得到的哺乳动物细胞表达系统;以及由所述哺乳动物细胞表达系统中的所述DNA片段 A、所述DNA片段B、所述DNA片段C和所述DNA片段D中的全部或部分组成的DNA 片段组也属于本发明的保护范围。

[0061] 所述DNA片段组或前文所述的SmartBac杆状病毒表达系统在同时表达n个目的物

中的应用也属于本发明的保护范围。其中,所述目的物为蛋白或蛋白亚基或蛋白片段或多肽或多肽片段;所述n为2以上的自然数。

[0062] 本发明还要求保护利用前文所述的SmartBac杆状病毒表达系统表达蛋白质复合物的方法。

[0063] 本发明所提供的利用前文所述的SmartBac杆状病毒表达系统表达蛋白质复合物的方法,具体可为如下(1)或(2):

[0064] (1) 当待表达的蛋白质复合物的分子量小于600kDa时,所述方法包括如下步骤:

[0065] (a1) 将组成所述蛋白质复合物的全部亚基的编码基因分成总长度接近的两组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,两组得到两条融合基因。

[0066] (a2) 将所述两条融合基因中的一条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒;将所述两条融合基因中的另一条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒。

[0067] 该步骤中所使用的所述受体质粒和所述供体质粒可选择携带有不同颜色荧光蛋白编码基因的质粒,以便可以监测不同蛋白的表达情况。

[0068] (a3) 将所述重组受体质粒和所述重组供体质粒在所述位点特异性重组酶的作用下融合为一个质粒,即得转移质粒。

[0069] (a4) 将所述转移质粒转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒。

[0070] (a5) 将所述重组杆状病毒穿梭质粒感染昆虫细胞,得到重组杆状病毒;

[0071] (a6) 将所述重组杆状病毒感染昆虫细胞,从而实现所述蛋白质复合物的表达;

[0072] 在该方法中,所述受体质粒既可为所述受体质粒甲(具体如所述4V1G质粒或所述4V1R 质粒)也可可为所述受体质粒乙(具体如所述5V1TG质粒或所述5V1TR质粒)。

[0073] (2) 当待表达的蛋白质复合物的分子量在600kDa以上时,所述方法为如下(B)或(C):

[0074] (B) 包括如下步骤:

[0075] (b1) 将组成所述蛋白质复合物的全部亚基的编码基因分成M组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,M组得到M条融合基因;其中, $M=2m$ ,m为2以上的自然数。

[0076] (b2) 将所述M条融合基因中的第1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒1;将所述M条融合基因中的第2条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒1;将所述M条融合基因中的第3条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒2;将所述M条融合基因中的第4条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒2;以此类推;将所述M条融合基因中的第M-1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒m;将所述M条融合基因中的第M条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒m。该步骤中针对所述M条融合基因所使用的各所述受体质粒最好为携带有不同颜色荧光蛋白编码基因的质粒,以便可以监测不同蛋白的表达情况。另外,克隆入所述供体质粒

的所述待表达目的物的编码基因的插入区的所述融合基因最好是带有终止密码子的,这样保证后续所得的单个转移质粒具有单一颜色的荧光。

[0077] (b3) 将所述重组受体质粒1和所述重组供体质粒1在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒1;将所述重组受体质粒2和所述重组供体质粒2在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒2;以此类推;将所述重组受体质粒m和所述重组供体质粒m在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒m。

[0078] (b4) 将所述转移质粒1转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒1;将所述转移质粒2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒2;以此类推;将所述转移质粒m转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒m。

[0079] (b5) 将所述重组杆状病毒穿梭质粒1感染昆虫细胞,得到重组杆状病毒1;将所述重组杆状病毒穿梭质粒2感染昆虫细胞,得到重组杆状病毒2;以此类推;将所述重组杆状病毒穿梭质粒m感染昆虫细胞,得到重组杆状病毒m。

[0080] (b6) 将所述重组杆状病毒1、所述重组杆状病毒2、以此类推、所述重组杆状病毒m共同感染昆虫细胞,从而实现所述蛋白质复合物的表达。

[0081] (C) 包括如下步骤:

[0082] (c1) 将组成所述蛋白质复合物的全部亚基的编码基因分成M组,将每组中的各亚基编码基因以所述蛋白酶的切割位点识别序列为间隔融合成一条长的融合基因,M组得到M条融合基因;其中, $M=2m-1$ ,m为2以上的自然数。

[0083] (c2) 将所述M条融合基因中的第1条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒1;将所述M条融合基因中的第2条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒1;将所述M条融合基因中的第3条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒2;将所述M条融合基因中的第4条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒2;以此类推;将所述M条融合基因中的第M-2条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒 $(M-1)/2$ ;将所述M条融合基因中的第M-1条克隆入所述供体质粒的所述待表达目的物的编码基因的插入区,得到重组供体质粒 $(M-1)/2$ ;将所述M条融合基因中的第M条克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,得到重组受体质粒m。该步骤中针对所述M条融合基因所使用的各所述受体质粒最好为携带有不同颜色荧光蛋白编码基因的质粒,以便可以监测不同蛋白的表达情况。另外,克隆入所述供体质粒的所述待表达目的物的编码基因的插入区的所述融合基因最好是带有终止密码子的,这样保证后续所得的单个转移质粒具有单一颜色的荧光。

[0084] (c3) 将所述重组受体质粒1和所述重组供体质粒1在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒1;将所述重组受体质粒2和所述重组供体质粒2在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒2;以此类推;将所述重组受体质粒 $(M-1)/2$ 和所述重组供体质粒 $(M-1)/2$ 在所述位点特异性重组酶的作用下融合为一个大质粒,即得转移质粒 $(M-1)/2$ ;所述重组受体质粒m即为转移质粒m。

[0085] (c4) 将所述转移质粒1转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒1;将所述转移质粒2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒2;以此类推;将所述转移质粒(M-1)/2转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒(M-1)/2;将所述转移质粒m转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒m。

[0086] (c5) 将所述重组杆状病毒穿梭质粒1感染昆虫细胞,得到重组杆状病毒1;将所述重组杆状病毒穿梭质粒2感染昆虫细胞,得到重组杆状病毒2;以此类推;将所述重组杆状病毒穿梭质粒m感染昆虫细胞,得到重组杆状病毒m。

[0087] (c6) 将所述重组杆状病毒1、所述重组杆状病毒2、以此类推、所述重组杆状病毒m共同感染昆虫细胞,从而实现所述蛋白质复合物的表达。

[0088] 在方法(B)和(C)中,所述受体质粒最好为所述受体质粒乙(具体如所述5V1TG质粒或所述5V1TR质粒)。

[0089] 本发明还要求保护一种利用前文所述的SmartBac杆状病毒表达系统筛选待表达蛋白复合物中适于添加纯化标签的亚基的方法。

[0090] 本发明所提供的利用前文所述的SmartBac杆状病毒表达系统筛选待表达蛋白复合物中适于添加纯化标签的亚基的方法,具体可为如下(3)或(4):

[0091] (3) 当待表达的蛋白质复合物的分子量小于600kDa时,所述方法包括如下步骤:

[0092] (a' 1) 按照所述(a1)-(a3)的步骤构建得到一个转移质粒;所述转移质粒不含有纯化标签编码序列。

[0093] (a' 2) 将所述待表达蛋白复合物的全部亚基的编码基因分别克隆入所述受体质粒的所述待表达目的物的编码基因的插入区,并能够与所述纯化标签编码序列融合表达,对应每个亚基各得到一个重组受体质粒,所述重组受体质粒即为转移质粒。

[0094] (a' 3) 将步骤(a' 1)和(a' 2)中的所有转移质粒分别转化含有杆状病毒穿梭质粒的受体菌,对应每个转移质粒均得到一个重组杆状病毒穿梭质粒。

[0095] (a' 4) 将步骤(a' 3)中的所有重组杆状病毒穿梭质粒分别感染昆虫细胞,得到各自对应的重组杆状病毒。

[0096] (a' 5) 将对应步骤(a' 2)中各转移质粒的所有所述重组杆状病毒中每一种分别对应步骤(a' 1)中所述转移质粒的所述重组杆状病毒组合,共同感染昆虫细胞,然后根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化,从而确定所述蛋白复合物中适于添加所述纯化标签的亚基。

[0097] 其中,在根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化后,可进行SDS-PAGE,然后根据SDS-PAGE结果按照如下确定所述蛋白复合物中适于添加所述纯化标签的亚基:若某添加所述纯化标签的亚基A可以将组成复合物的其它亚基都捕捉到,且除了带所述纯化标签的亚基A之外,其他各亚基之间的比例也相对比较均匀,则可确定所述亚基A为所述蛋白复合物中适于添加所述纯化标签的亚基。

[0098] (4) 当待表达的蛋白质复合物的分子量在600kDa以上时,所述方法包括如下步骤:

[0099] (b' 1) 按照所述(b1)-(b3)的步骤或者所述(c1)-(c3)的步骤构建得到m个转移质粒;所述转移质粒不含有纯化标签编码序列。

[0100] (b' 2) 将所述待表达蛋白复合物的全部亚基的编码基因分别克隆入所述受体载体

的所述待表达目的物的编码基因的插入区,并能够与所述纯化标签编码序列融合表达,对应每个亚基各得到一个重组受体质粒,所述重组受体质粒即为转移质粒。

[0101] (b' 3) 将步骤(b' 1)和(b' 2)中的所有转移质粒分别转化含有杆状病毒穿梭质粒的受体菌,对应每个转移质粒均得到一个重组杆状病毒穿梭质粒。

[0102] (b' 4) 将步骤(b' 3)中的所有重组杆状病毒穿梭质粒分别感染昆虫细胞,得到各自对应的重组杆状病毒。

[0103] (b' 5) 将对应步骤(b' 2)中各转移质粒的所有所述重组杆状病毒中的每一种分别与对应步骤(b' 1)中所述m个转移质粒的m个所述重组杆状病毒组合,共同感染昆虫细胞,然后根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化,从而确定所述蛋白复合物中适于添加所述纯化标签的亚基。

[0104] 其中,在根据所述纯化标签对感染后的各组细胞所表达的蛋白复合物进行纯化后,可进行SDS-PAGE,然后根据SDS-PAGE结果按照如下确定所述蛋白复合物中适于添加所述纯化标签的亚基:若某添加所述纯化标签的亚基A可以将组成复合物的其它亚基都捕捉到,且除了带所述纯化标签的亚基A之外,其他各亚基之间的比例也相对比较均匀,则可确定所述亚基A为所述蛋白复合物中适于添加所述纯化标签的亚基。

[0105] 进一步地,在本发明中所述纯化标签为Twin-Strep标签。相应的,对所述待表达的蛋白复合物进行纯化时可用Strep亲和介质进行纯化。

[0106] 在本发明的一个实施例中,所述m具体为2。

[0107] 在本发明的一个具体实施例中,所述待表达蛋白复合物具体为人源exocyst复合物。所述人源exocyst复合物由如下8个蛋白亚基组成:EXOC1 (102kDa), EXOC2 (104kDa), EXOC3 (86kDa), EXOC4 (110kDa), EXOC5 (82kDa), EXOC6 (94kDa), EXOC7 (78kDa) 和EXOC8 (82kDa)。

[0108] 相应的,本发明提供的表达所述人源exocyst复合物的方法,具体包括如下步骤:

[0109] (1) 将重组受体质粒5V1TR-E47(全序列如SEQ ID No.16所示)和重组供体质粒4V2-E1S5(全序列如SEQ ID No.19所示)在Cre重组酶的作用下融合为一个质粒,即得转移质粒E1S547;将重组受体质粒5V1TG-E63(全序列如SEQ ID No.15所示)和重组供体质粒4V2-E28(全序列如SEQ ID No.18所示)在Cre重组酶的作用下融合为一个质粒,即得转移质粒E2863。

[0110] (2) 将所述转移质粒E1S547转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒BC-E1S547;将所述转移质粒E2863转化含有杆状病毒穿梭质粒的受体菌,得到重组杆状病毒穿梭质粒BC-E2863。

[0111] (3) 将所述重组杆状病毒穿梭质粒BC-E1S547转染昆虫细胞,可以得到重组杆状病毒 BV-E1S547;将所述重组杆状病毒BC-E2863转染昆虫细胞,可以得到重组杆状病毒 BV-E2863。

[0112] (4) 将所述重组杆状病毒BV-E1S547与所述重组杆状病毒BV-E2863共同感染昆虫细胞,从而实现所述人源exocyst复合物的表达。

[0113] 在本发明中,所述昆虫细胞具体为Sf9细胞。

[0114] 最后,本发明还要求保护利用前文所述的哺乳动物细胞表达系统表达蛋白质复合物的方法。该方法与本发明所提供的利用前文所述的SmartBac杆状病毒表达系统表达蛋白

质复合物的方法相比,差别仅在于将其中的所述SmartBac杆状病毒表达系统替换为所述哺乳动物细胞表达系统,并将昆虫细胞替换为哺乳动物细胞。

[0115] 针对方法1中所出现的三种问题,本发明设计了SmartBac系列载体来解决,并提出了三种克隆的策略以实现分子量的在600kDa以下的蛋白质复合物表达、分子量在600kDa以上的蛋白质复合物的表达以及高效筛选最适合添加纯化标签的亚基。实验证明,采用本发明所设计的新型SmartBac杆状病毒表达系统及具体克隆策略表达人源exocyst复合物,发现带有 Twin-strep标签的EXOC5亚基最适合于纯化整个Exocyst复合物,并且最终成功表达纯化出人源exocyst,其与从酵母中提取的exocyst有着相似的尺寸和形状,而且体外活性测定表明该重组表达的exocyst复合物具有生物学活性。本发明对于在昆虫细胞中重组表达组分复杂、分子量较大的蛋白质复合物具有重要意义。此外,本发明中各载体上被昆虫细胞识别的用于表达目的蛋白的启动子、终止序列可以替换为能被哺乳动物细胞识别的启动子和终止序列,则替换后的载体系统可以用于在哺乳动物细胞中表达复杂的蛋白质复合物,因此本发明也提供了一种在哺乳动物细胞中重组表达复杂蛋白质复合物的方法。

## 附图说明

[0116] 图1为SmartBac系统所含6种载体的质粒图谱。

[0117] 图2为表达蛋白质复合物的克隆策略(方案1和方案2)。a和b为表达分子量小于600kDa 的蛋白质复合物的克隆策略(方案1)。a为将编码8种亚基的基因分成总长度接近的两组。b 为表达分子量小于600kDa的蛋白质复合物的具体克隆示意图。c为表达分子量大于600kDa 的蛋白质复合物的克隆策略(方案2)。

[0118] 图3为筛选最佳的带有纯化标签的亚基的策略(方案3)。a为所构建的10个转移质粒;b为包装出的10个重组杆状病毒;c为感染昆虫细胞的8种病毒组合(每种组合都含有3种病毒);d为确定了适于带有纯化标签的亚基后继续表达蛋白复合物的流程图。

[0119] 图4为采用SmartBac杆状病毒表达系统表达人源exocyst复合物的表达效果鉴定。a为根据荧光蛋白表达情况初步判定蛋白亚基表达情况;b为SDS-PAGE鉴定哪个蛋白亚基适合于带有Twin-strep标签用于纯化整个exocyst复合物;c为SDS-PAGE鉴定纯化好的exocyst 复合物;d为exocyst复合物的电镜负染图;e为d所示电镜图进一步使用RELION 2进行二维分类的结果;f为使用EMAN2产生的exocyst复合物的初始模型。

[0120] 图5为昆虫细胞表达的人源exocyst复合物的生物学活性测定结果。a为四个组中实验组(脂质体A+Rab11Q70L+Exocyst+脂质体B)相对于其它三个对照组在585nm处,有明显的荧光发射峰,而只加入Rab11Q70L或exocyst复合物的对照组,相对于没有加入任何一种蛋白的对照组AB,均没有明显的荧光能量共振转移。b为三次独立实验的平均值及误差,纵坐标是三个对照组与实验组在585nm处荧光值与533nm处荧光值的比。

## 具体实施方式

[0121] 1、SmartBac载体系统描述

[0122] SmartBac系统含有6种载体,其中有4种是受体质粒,2种是供体质粒(图1)。本发明采用广泛适用的UPS(univector plasmid-fusion system)策略来克服用常规克隆方法构建大质粒的困难。该策略使用Cre-loxP位点特异性重组来催化供体质粒和受体质粒之间的

融合。具有卡那霉素抗性( $K^+$ )的供体质粒携带R6K  $\gamma$  条件复制起点,质粒仅能在含有pir基因的细菌宿主中繁殖。具有氨苄青霉素抗性( $A^+$ )的受体质粒携带非条件型复制起点,在所有菌株中都能复制。供体质粒和受体质粒上都有一个LoxP位点,可以被Cre重组酶识别发生分子间重组。将供体质粒和受体质粒进行融合反应后,将反应产物转化不含有pir基因的菌株(pir<sup>-</sup>菌株),用 $A^+K^+$ 双抗平板进行筛选,可以得到含有融合质粒的菌落。利用这种策略,本发明可以先制备含有几个目的基因的供体和受体质粒,再将供体、受体质粒融合来得到较大的质粒。

[0123] 4种受体质粒分别是4V1G,4V1R,5V1TG和5V1TR,2种供体质粒分别是4V2G和4V2R。受体质粒含有p15A复制起点,其允许质粒以低拷贝在普通大肠杆菌克隆菌株中繁殖,它更好的维持了大质粒的稳定性。受体质粒还含有氨苄青霉素和庆大霉素的抗性标记,以及用于产生重组杆状病毒的小型Tn7元件侧翼。受体质粒中启动转基因序列在昆虫细胞中表达的是 p6.9启动子。与常规使用的极晚期多角体蛋白polyhedrin启动子相比,p6.9启动子在感染的早期阶段驱动表达,这时的细胞状态与病毒感染晚期相比更好,可以避免表达的外源蛋白质的聚集。

[0124] 4V1G和4V1R受体质粒携带N端HA标记的TEV蛋白酶编码序列,随后是TEV蛋白酶切割位点(TCS)和Twin-Strep标签编码序列,随后是肠激酶切割位点的识别序列。在多克隆位点(MCS)1和2之间,存在LacZ- $\alpha$ 表达盒,其允许重组克隆的蓝/白选择。另一个 TCS和EGFP(4V1G)或TagRFP(4V1R)编码序列在MCS2的下游。荧光蛋白和靶蛋白可以表达为单个ORF。通过观察感染细胞的荧光,就可以很容易地确定目标蛋白是否已经表达。

[0125] 在5V1TG和5V1TR受体质粒中(不同于4V1G与4V1R受体质粒),将TEV蛋白酶和 EGFP(5V1TG)或TagRFP(5V1TR)编码序列融合并表达为GP64启动子驱动的ORF。这样可以完全避免4V1G与4V1R中可能出现的荧光蛋白与多肽链之间剪切不完全的情况。

[0126] 4V2G和4V2R供体质粒携带N-末端10 $\times$ His标签编码序列,随后是肠激酶切割位点的识别序列。两种载体均含有卡那霉素抗性标记。筛选区域由高拷贝PUC复制起点和侧翼为MCS1和MCS2的LacZ- $\alpha$ 表达盒组成。在MCS2的下游,存在TCS和荧光蛋白(4V2G中为 EGFP和4V2R中为TagRFP)编码序列。靶蛋白的表达是由非常晚的p10启动子驱动的。4V2 载体也含有条件复制起点R6K  $\gamma$ 。一旦筛选区被外源基因取代,供体载体仅包含R6K  $\gamma$  起点,并且只能在具有pir<sup>+</sup>基因型的大肠杆菌菌株中繁殖。

[0127] 在4V1/5V1受体质粒和4V2供体质粒中的p6.9和p10启动子区域的两侧分别存在几个单一的限制性酶切位点,这样当需要时,就可以将现有的启动子剪切下来,替换为其它的可以在昆虫细胞中启动蛋白质表达的启动子。

[0128] 2、SmartBac载体系统应用策略

[0129] 使用Smartbac载体系统可以在昆虫细胞中更容易和更快地表达大的蛋白质复合物。可以使用多种实验方案从Smartbac载体产生最终用于表达蛋白质复合物的转移质粒。这里本发明只提出两个经典方案来说明如何使用Smartbac系列载体。

[0130] 方案1、表达分子量小于600kDa的蛋白质复合物的克隆策略

[0131] 假设要在昆虫细胞中表达由8个不同亚基A,B,C,D,E,F,G和H构成的蛋白质复合物。如果该复合物的分子量小于600kDa,可以采用方案1。如图2中a所示,将编码8种亚基的基因分成总长度接近的两组,第一组含有基因A、B、C、D;第二组含有基因E、F、G和H。将每一

组的四条基因拼接成以TCS编码位点为间隔的一条长的融合基因。第一组的融合DNA片段为ABCD,第二组的融合DNA片段为EFGH。接下来,将长的ABCD片段进一步分成两个短的DNA片段AB和CD;而长的EFGH片段则分成EF和GH这两个较短的DNA片段。接下来设计相应的引物,利用重叠PCR的方法将A、B片段融合为AB,C、D片段融合为CD,E、F片段融合为EF,G、H片段融合为GH。然后通过Gibson组装反应,将片段AB和CD与线性化的Smartbac RFP表达受体质粒融合。同理,片段EF和GH也与线性化的Smartbac GFP表达供体质粒(4V2G)进行了融合。利用SmartBac系列载体,含有阳性重组质粒4V1R-ABCD与4V2G-EFGH重组菌可以很容易的通过蓝白筛选选择出来。最后提取出4V1R-ABCD与4V2G-EFGH这两种质粒,通过Cre-LoxP位点特异性重组将它们融合为最终转移质粒ABCD-EFGH。将该质粒转化到DH10Bac感受态细胞后,将获得重组Bacmid。将该Bacmid转染昆虫细胞中,就会产生用于表达目标复合物的高滴度杆状病毒BV-ABCD-EFGH。通过监测TagRFP的红色荧光,可以知道ABCD融合蛋白的表达情况;而监测EGFP的绿色荧光,可以判断EFGH融合蛋白的表达情况。图2中b所示的是使用4V1R与4V2G,但也可以使用5V1TR与4V2G,或使用4V1G与4V2R,或使用5V1TG与4V2R。

[0132] 方案2、表达分子量大于600kDa的蛋白质复合物的克隆策略

[0133] 如果要表达的蛋白质复合物的分子量大于600kDa,使用方案1就需要构建出长度大于25kb的最终转移质粒。在没有经验的情况下构建这么大的质粒通常是很困难的。即使构建成功,多蛋白复合物也可能不能在昆虫细胞中表达。这是因为大转移质粒产生的重组杆状病毒易于展现出内在的遗传不稳定性。在P2代病毒扩增期间可能发生外源基因的丢失。在这种情况下,使用方案2更加合适。如图2中c所示,片段A和B与线性化的5V1TR组装形成5V1TR-AB质粒,片段C和D与线性化的4V2载体融合成4V2-CD质粒。基因D的3'末端加入了终止密码子,这样存在于4V2G和4V2R载体上的荧光蛋白不会被表达。5V1TR-AB与4V2-CD会通过Cre-LoxP位点特异性重组融合为最终转移质粒ABCD(RFP),这个质粒转化DH10Bac会产生重组Bacmid-ABCD。同理制备装载片段E和F的5V1TG-EF质粒与装载片段G和H的4V2-GH质粒,两者可以融合为另一个最终转移质粒EFGH(GFP),并转化DH10Bac产生重组Bacmid-EFGH。用两种重组Bacmid分别感染昆虫细胞,会得到两种重组杆状病毒BV-ABCD与BV-EFGH。BV-ABCD会表达亚基A,B,C和D以及TagRFP,而BV-EFGH会表达亚基E,F,G和H以及EGFP。因此用这两种病毒同时感染杆状病毒的昆虫细胞将产生完整的蛋白质复合物,其中TagRFP和EGFP荧光的出现表明各亚基已经成功表达。图2中c所示的是使用5V1TG与5V1TR这两种受体质粒,也可使用以下的受体质粒组合:4V1G与4V1R,5V1TG与4V1R,5V1TR与4V1G。

[0134] 方案3、筛选最佳的带有纯化标签的亚基的策略

[0135] Smartbac系统的4种受体质粒都携带N端Twin-Strep标签,2种供体质粒携带N端10×His标签。每个标签可以与目的亚基融合表达,用来纯化整个复合物。假设要表达一个分子量大于600kDa,含8种不同亚基A,B,C,D,E,F,G和H构成的蛋白质复合物。由于先验知识的缺乏,不知道把纯化标签加在哪一种亚基上对纯化整个复合物更加有利。本发明要采用以下的克隆构建方案,筛选出最适合加纯化标签的亚基,用于整个复合物的纯化。首先根据方案2构建两个大的最终转移质粒ABCD(TagRFP)和EFGH(EGFP),这两个大质粒上表达的任何亚基都不含有纯化标签(图3中a)。然后用任意一种受体质粒(4V1G,4V1R,5V1TG或5V1TR)构建出8个较小的转移质粒(从V1-TSA到V1-TSH),每个质粒都表达N端带有Twin-Strep-标签的亚基。这样最终会获得10个重组杆状病毒,包括BV-ABCD(TagRFP),BV-EFGH

(EGFP) 和BV-TSn (其中n从A到H) (图3中b)。接下来用三种杆状病毒——BV-ABCD (TagRFP), BV-EFGH (EGFP) 和一种BV-TSn共感染昆虫细胞。这样我们一共有8种病毒组合 (每种组合都含有3种病毒) 要尝试 (图3中c)。在将用8种病毒组合感染的细胞进行裂解, 并用Strep亲和介质进行纯化后, 会分析出带亲和标记的亚基 H用于纯化整个复合物的效果最好。这时为了增加产量并获得更均匀的样品, 需要构建一个新的中间载体G-TSH, 在这个新载体中亚基H带有Twin-Strep亲和标签。利用新载体G-TSH 和原有的中间载体EF融合, 将产生一个新的转移载体——EFG-TSH (EGFP), 这个新的转移载体会产生新的重组杆状病毒BV-EFG-TSH (GFP)。用它和已有的重组杆状病毒BV-ABCD (TagRFP) 一起感染昆虫细胞, 将实现目的蛋白质复合物的表达 (图3中d)。病毒感染和蛋白质表达情况可以由被感染细胞的EGFP和TagRFP的荧光分布和强度来监测。

[0136] 实施例1、采用SmartBac杆状病毒表达系统表达人源exocyst复合物

[0137] 本实施例中所使用的实验方法如无特殊说明, 均为常规方法。

[0138] 本实施例中所用的材料、试剂等, 如无特殊说明, 均可从商业途径得到。

[0139] 1、本发明首先构建SmartBac系统的6种载体。6种载体在金唯智基因合成公司合成。

[0140] 四种受体质粒:

[0141] (1) 4V1G质粒

[0142] 4V1G质粒的全序列为SEQ ID No.1。其中, SEQ ID No.1的第20-243位为Tn7R元件, 第310-843位为庆大霉素抗性蛋白的编码基因序列, 第1131-1164位为loxP位点, 第1168-1224 位为多克隆位点BP, 第1235-1329位为p6.9启动子, 第1339-2097位为N端被HA标记的TEV蛋白酶的编码基因序列, 第2098-2118位为TEV蛋白酶切割位点的识别序列 (TCS), 第2119-2205位为Twin-Strep标签编码序列, 第2206-2220位为肠激酶切割位点的识别序列, 第2222-2249位为多克隆位点1, 第2250-2790位为LacZ- $\alpha$ 表达盒, 第2799-2837位为多克隆位点2, 第2839-2859位为TEV蛋白酶切割位点的识别序列 (TCS), 第2862-2867位为SalI 识别位点, 第2869-3585位为荧光蛋白EGFP的编码基因序列, 第3731-3971为SV40pA信号序列, 第4000-4165位为Tn7L元件, 第4349-4804位为f1复制起点, 第4936-5796位为氨苄青霉素抗性蛋白的编码基因序列, 第6067-6612位为p15A复制起点。

[0143] (2) 4V1R质粒

[0144] 4V1R质粒的全序列为SEQ ID No.2。其中, SEQ ID No.2的第20-243位为Tn7R元件, 第310-843位为庆大霉素抗性蛋白的编码基因序列, 第1131-1164位为loxP位点, 第1168-1224 位为多克隆位点BP, 第1235-1329位为p6.9启动子, 第1339-2097位为N端被HA标记的TEV蛋白酶的编码基因序列, 第2098-2118位为TEV蛋白酶切割位点的识别序列 (TCS), 第2119-2205位为Twin-Strep标签编码序列, 第2206-2220位为肠激酶切割位点的识别序列, 第2222-2249位为多克隆位点1, 第2250-2790位为LacZ- $\alpha$ 表达盒, 第2799-2843位为多克隆位点2, 第2845-2865位为TEV蛋白酶切割位点的识别序列 (TCS), 第2868-2873位为SalI 识别位点, 第2875-3585位为荧光蛋白TagRFP的编码基因序列, 第3731-3971为SV40pA 信号序列, 第4000-4165位为Tn7L元件, 第4349-4804位为f1复制起点, 第4831-4935位为 Amp启动子, 第4936-5796位为氨苄青霉素抗性蛋白的编码基因序列, 第6067-6612位为p15A 复制起点。

[0145] (3) 5V1TG质粒

[0146] 5V1TG质粒的全序列为SEQ ID No.3。其中,SEQ ID No.3的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1193 位为I-CeuI酶切位点,第1202-1207位为AatII酶切位点,第1208-1514位为IE1ter信号序列,第1515-1522位为FseI酶切位点,第1531-2247位为荧光蛋白EGFP的编码基因序列,第 2248-2253位为SmaI酶切位点,第2254-2274位为TEV蛋白酶切割位点的识别序列,第 2275-3033位为N端被HA标记的TEV蛋白酶的编码基因序列,第3034-3041位为SwaI酶切位点,第3047-3252位为GP64启动子,第3290-3295位为BspEI酶切位点,第3306-3400位为P6.9启动子,第3407-3418位为多克隆位点AP,第3419-3505位为Twin-Strep标签编码序列,第3506-3520位为肠激酶切割位点的识别序列,第3522-2549位为多克隆位点1,第3550-4090位为LacZ- $\alpha$ 表达盒,第4099-4149位为多克隆位点2,第4281-4521位为SV40pA 信号序列,第4550-4715位为Tn7L元件,第4899-5354位为f1复制起点,第5486-6346位为氨苄青霉素抗性蛋白的编码基因序列,第6617-7162位为p15A复制起点。

[0147] (4) 5V1TR质粒

[0148] 5V1TR质粒的全序列为SEQ ID No.4。其中,SEQ ID No.4的第20-243位为Tn7R元件,第310-843位为庆大霉素抗性蛋白的编码基因序列,第1131-1164位为loxP位点,第1168-1193 位为I-CeuI酶切位点,第1202-1207位为AatII酶切位点,第1208-1514位为IE1ter信号序列,第1515-1522位为FseI酶切位点,第1531-2244位为荧光蛋白TagRFP的编码基因序列,第 2245-2250位为SmaI酶切位点,第2251-2271位为TEV蛋白酶切割位点的识别序列,第 2272-3030位为N端被HA标记的TEV蛋白酶的编码基因序列,第3031-3038位为SwaI酶切位点,第3044-3249位为GP64启动子,第3287-3292位为BspEI酶切位点,第3303-3397位为P6.9启动子,第3404-3415位为多克隆位点AP,第3416-3502位为Twin-Strep标签编码序列,第3503-3517位为肠激酶切割位点的识别序列,第3519-3546位为多克隆位点1,第 3547-4087位为LacZ- $\alpha$ 表达盒,第4096-4146位为多克隆位点2,第4278-4518位为SV40pA 信号序列,第4547-4712位为Tn7L元件,第4896-5351位为f1复制起点,第5483-6343位为氨苄青霉素抗性蛋白的编码基因序列,第6614-7159位为p15A复制起点。

[0149] 两种供体质粒:

[0150] (1) 4V2G质粒

[0151] 4V2G质粒的全序列为SEQ ID No.5。其中,SEQ ID No.5的第169-202位为LoxP位点,第208-215位为NotI酶切位点,第259-368位为p10启动子,第376-393位为多克隆位点AP,第397-426位为10 $\times$ His标签编码序列,第436-450位为肠激酶切割位点的识别序列,第452-473位为多克隆位点1,第644-1232位为PUC复制起点,第1443-1983位为LacZ- $\alpha$ 表达盒,第1992-2030位为多克隆位点2,第2032-2052位为TEV蛋白酶切割位点的识别序列,第2055-2060位为SacI识别位点,第2062-2778位为荧光蛋白EGFP的编码基因序列,第 2907-3188位为HSV-tk pA信号序列,第3224-3579位为R6K复制起点,第3940-4734位为卡那霉素抗性蛋白编码序列。

[0152] (2) 4V2R质粒

[0153] 4V2R质粒的全序列为SEQ ID No.6。其中,SEQ ID No.6的第169-202位为LoxP位点,第208-215位为NotI酶切位点,第259-368位为p10启动子,第376-393位为多克隆位点

AP,第397-426位为10×His标签编码序列,第436-450位为肠激酶切割位点的识别序列,第452-473位为多克隆位点1,第644-1232位为PUC复制起点,第1443-1983位为LacZ-α表达盒,第1992-2036位为多克隆位点2,第2038-2058位为TEV蛋白酶切割位点的识别序列,第2061-2066位为SacI识别位点,第2068-2778位为荧光蛋白TagRFP的编码基因序列,第2907-3188位为HSV-tk pA信号序列,第3224-3579位为R6K复制起点,第3940-4734位为卡那霉素抗性蛋白编码序列。

[0154] 2、利用本发明的方案3,首次在昆虫细胞中表达了人源exocyst复合物,筛选到了最适于纯化该复合物的亚基。利用该亚基上的纯化标签,成功的纯化到较为均一的、具有生物学活性的exocyst复合物。对纯化后的样品使用负染色技术进行了电镜观察,发现该样品的形状与酿酒酵母中天然提取的exocyst复合物的形状相似。

[0155] Exocyst复合体负责将分泌小泡束缚在质膜上,为可溶性N-乙基马来酰亚胺敏感因子(NSF)附着蛋白受体(SNARE)介导的膜融合做准备。人源exocyst复合物包含8个在进化上非常保守的亚基——EXOC1 (102kDa),EXOC2 (104kDa),EXOC3 (86kDa),EXOC4 (110kDa),EXOC5 (82kDa),EXOC6 (94kDa),EXOC7 (78kDa)和EXOC8 (82kDa)。因为该复合物之前没有被重组表达过,所以不知道在哪个亚基上加纯化标签最适于纯化整个 exocyst复合物。为了确定最适和纯化整个复合物的亚基,本发明采用方案3进行载体构建工作。表1列出了本发明在昆虫细胞中重组表达exocyst所构建的所有载体。

[0156] 表1 使用SmartBac系统重组表达人源Exocyst复合物所构建的载体及使用的病毒

| [0157] | 含有的亚基                  | 中间载体      | 最终转移载体    | 重组杆状病毒    |
|--------|------------------------|-----------|-----------|-----------|
|        | TS-tagged EXOCn        |           | 5V1TG-SEn | BV-SEn    |
|        | EXOC2, EXOC8           | 4V2-E28   | E2863     | BV-E2863  |
|        | EXOC6, EXOC3           | 5V1TG-E63 |           |           |
|        | EXOC1, EXOC5           | 4V2-E15   | E1547     | BV-E1547  |
|        | EXOC4, EXOC7           | 5V1TR-E47 |           |           |
|        | EXOC1, TS-tagged EXOC5 | 4V2-E1SE5 | E1S547    | BV-E1S547 |

[0158] 编码exocyst各亚基的基因都来自Origene公司 (EXOC1-SC126966,EXOC2-SC111916,EXOC3-RC209413,EXOC4-SC102359,EXOC5-SC127665,EXOC6-SC100885,EXOC7-RC227511,EXOC8-RC207859)。Rab11 (1-173) Q70L基因在金唯智公司合成。所用到的引物在Invitrogen公司进行合成。所有克隆的测序都在北京博尚公司进行。限制性内切酶 BamHI-HF、EcoRI-HF、KpnI-HF、NdeI、NotI-HF、XhoI,碱性磷酸酶CIP,DNA聚合酶Q5 Premix,Cre重组酶及Gibson组装试剂盒NEBuilder®HiFi DNA Assembly Master Mix为NEB 公司产品。胶回收试剂盒为Qiagen公司产品。DNA聚合酶KOD-FX为Toyoba公司产品。GT115感受态细胞是Invivogen公司产品。Trans2-blue感受态细胞是全式金公司产品。DNA Ligation

Kit来自Takara公司。感受态细胞DH10Bac、昆虫细胞sf9、转染试剂cellfectin II和Grace's Insect Cell Culture Medium,Unsupplemented培养基均为Invitrogen公司产品,产品目录号分别为10361012、B82501、10362100和11595030。活性测定用磷脂Rhod B-DHPE 也购自Invitrogen公司,其它各种磷脂和制备脂质体所用的Extruder装置均购自Avanti Lipids Polar公司。昆虫细胞培养基ESF921为Expression System公司产品。Strep亲和介质为IBA 公司产品。蛋白酶抑制剂(Complete-EDTA Free)与Ni-NTA介质为Roche公司产品。分子筛Superdex 200 (10/300GL) 购自GE公司。GTP购自Sigma公司。其它试剂材料,均可从商业途径得到。

[0159] 裂解缓冲液1:含有50mM HEPES pH 8.0,150mM NaCl,10% (v/v) 甘油,1mM DTT (补充了Roche蛋白酶抑制剂)。

[0160] 洗脱缓冲液1:含有50mM HEPES pH 8.0,150mM NaCl,10% (v/v) 甘油,1mM DTT, 10mM脱巯生物素。

[0161] 裂解缓冲液2:含有20mM Tris-HCl pH 8.0,500mM NaCl,20mM imidazole,0.1mM PMSF。

[0162] 洗脱缓冲液2:含有20mM Tris-HCl pH 8.0,500mM NaCl,150mM imidazole,0.1mM PMSF。

[0163] 分子筛缓冲液:含有20mM Tris-HCl pH 8.0,150mM NaCl,0.1mM PMSF。

[0164] 测活缓冲液:50mM Hepes pH7.4,150mM NaCl。

[0165] 2.1双基因中间载体的构建

[0166] 双基因中间载体构建时用到的引物如表2所示

[0167] 表2 双基因中间载体构建时用到的引物

[0168]

| 引物名称 | 引物序列(5'-3')                                        |
|------|----------------------------------------------------|
| E1F1 | TATCCATATGGGATCCACAGCAATCAAGCATGCATTACAAAG         |
| E1F2 | ATCACTCGACACCGGTGATATCCATATGGGATCCACAGCAATCAAG     |
| E1R1 | CAGGTTTTCTACTCGAGCCGTGGGACTGTGCAATGCTGGAACAATAATC  |
| E1R2 | TAAGCTAGAGCTCTGGAAGTACAGGTTTTCTACTCGAGCCGTGGGACTG  |
| E5F1 | GCTTAAGCGCGGCCGCGACCACGGCCGAGTTGTTTCGAGGAGCCTTTTG  |
| E5F2 | ACCTGTACTTCCAGAGCTCTAGCTTAAGCGCGGCCGCGACCACGGCCG   |
| E5R1 | CCTTTCGGGTACCCTCGAGTTAGCTGAAGTGTCGAGCAAGGCGGGCAG   |
| E5R2 | CAGCAGCCAACTCAGCTTCCTTTCGGGTACCCTCGAGTTAGCTG       |
| E2F1 | TATCCATATGGCGGCCGCTAGCCGATCACGACAACCCCCCCTTG       |
| E2F2 | ATCACTCGACACCGGTGATATCCATATGGCGGCCGCTAGCCGATC      |
| E2R1 | ATACAAATTTTCTACTCGAGCCTGTTTTTCATCATGGTTGAAGAAGCTGC |
| E2R2 | GTACCACTAGTGCTTTGAAAATACAAATTTTCTACTCGAGCCTGTTTTTC |
| E8F1 | AGCACTAGTGGTACCCTTAAGATGGCGATGGCGATGTCGGACAGTGGG   |
| E8F2 | AGTGAAAATTTGTATTTTCAAAGCACTAGTGGTACCCTTAAGATGGCG   |
| E8R1 | CTTTCGGGTACCGAATTCTTAGACCACTGATGTTGTACTTTCAGG      |
| E8R2 | CAGCAGCCAACTCAGCTTCCTTTCGGGTACCGAATTCTTAGACCAC     |
| E3F1 | AGAGCGGTACCGCGGCCGCGATGAAGGAGACAGACCGGGAGGCCGTTG   |
| E3F2 | CGAGTGAGAATCTGTATTTCCAGAGCGGTACCGCGGCCGCGATGAAGG   |
| E3R1 | GTCAGTTAACTCGAGTTACTTGAGCAGCTTGGCCACGTTTCAG        |
| E3R2 | CCTTTCGAAGCTTTTAGTCAGTTAACTCGAGTTACTTGAGCAG        |
| E6F1 | AGCTCCATATGGGATCCGCGGAGAACAGCGAGAGTCTGGGCAC        |
| E6F2 | ATTTAAACGGATCGATGAGCTCCATATGGGATCCGCGGAGAAC        |
| E6R1 | ACAGATTCTCACTCGAGCCCATGTGCTGGGACATACCATTCACCAAAC   |
| E6R2 | CGCGGTACCGCTCTGGAATAACAGATTCTCACTCGAGCCCATGTGCTG   |
| E4F1 | GCTCCATATGGCGGCCGCGGAAGCAGCTGGTGGGAAATACAGAAG      |
| E4F2 | ATTTAAACGGATCGATGAGCTCCATATGGCGGCCGCGGAAGCAGC      |

[0169]

|      |                                                   |
|------|---------------------------------------------------|
| E4R1 | AGTACAAGTTCTCACTCGAGCCAACGGTAGTTATCTTCTTGTCCTTGG  |
| E4R2 | GTACCGCTTAAGGACTGAAAGTACAAGTTCTCACTCGAGCCAACGG    |
| E7F1 | TCAGTCCTTAAGCGGTACCATGATTCCCCACAGGAGGCATCCGCTC    |
| E7F2 | AGTGAGAACTTGTACTTTCAGTCCTTAAGCGGTACCATGATTCC      |
| E7R1 | TCAGTTAACTCGAGTTAGGCAGAGGTGTCAAAAAGGCGATCG        |
| E7R2 | CCTTTCGAAGCTTTTAGTCAGTTAACTCGAGTTAGGCAGAGGTG      |
| S5-F | TACTTCCAGAGCTCTAGCTTAAAGCATGGCCTGGAGCCATCCGCAATTG |
| S5-R | CAACTCAGCTTCCTTTCGGGTACCTTAGCTGAAGTGTCGAGCAAGGCGG |

[0170] (1) 各亚基片段的PCR反应

[0171] Exocyst八种亚基基因的PCR反应如表3所示。

[0172] 表3 Exocyst八种亚基基因的PCR反应

[0173]

| 模板      | 第一轮PCR反应用引物对 | 第二轮PCR反应用引物对 |
|---------|--------------|--------------|
| EXOC1基因 | E1F1和E1R1    | E1F2和E1R2    |
| EXOC2基因 | E2F1和E2R1    | E2F2和E2R2    |
| EXOC3基因 | E3F1和E3R1    | E3F2和E3R2    |
| EXOC4基因 | E4F1和E4R1    | E4F2和E4R2    |
| EXOC5基因 | E5F1和E5R1    | E5F2和E5R2    |
| EXOC6基因 | E6F1和E6R1    | E6F2和E6R2    |
| EXOC7基因 | E7F1和E7R1    | E7F2和E7R2    |
| EXOC8基因 | E8F1和E8R1    | E8F2和E8R2    |

[0174] 以含有目的基因的质粒为模板,进行表3中的PCR反应。第一轮PCR用NEB公司的Q5Premix完成。反应程序为:98℃,30s;(98℃10s,72℃3min,共30个循环),72℃10min。第一轮PCR产物回收后,作为模板进行第二轮PCR反应。反应程序为:98℃,30s;(98℃10s,72℃3min,共30个循环),72℃10min。回收PCR产物。

[0175] (2) 载体的线性化

[0176] 用限制性内切酶NdeI、KpnI-HF和EcoRI-HF酶切质粒4V2G,37℃处理1h;然后加入CIP,37℃处理1h;最后利用胶回收试剂盒回收约2.5kb的载体骨架4V2。

[0177] 用限制性内切酶NdeI、XhoI酶切质粒5V1TG,37℃处理1h;然后加入CIP,37℃处理1h;最后利用胶回收试剂盒回收约7kb的载体骨架。

[0178] 用限制性内切酶NdeI、XhoI酶切质粒5V1TR,37℃处理1h;然后加入CIP,37℃处理1h;最后利用胶回收试剂盒回收约7kb的载体骨架。

[0179] (3) PCR产物与线性化载体的Gibson组装反应

[0180] 按照NEBuilder®HiFi DNA Assembly Master Mix的说明书,将线性化的4V2载体骨架与EXOC1、EXOC5的第二轮PCR产物进行Gibson重组,以得到重组质粒4V2-E15;将线性化的4V2载体骨架与EXOC2、EXOC8的第二轮PCR产物进行重组,以得到重组质粒4V2-E28。重组产物转化GT115感受态细胞,通过抗生素筛选和蓝白斑筛选,挑取白色的重组菌落,进行

PCR鉴定。阳性菌落提取质粒并送测序。保存测序正确的克隆。

[0181] 同理,将线性化的5V1TG载体与EXOC6、EXOC3的第二轮PCR产物进行Gibson组装,以得到重组质粒5V1TG-E63;将线性化的5V1TR载体与EXOC4、EXOC7的第二轮PCR产物进行Gibson组装,以得到重组质粒5V1TR-E47。将这两种组装产物分别转化Trans2-blue 感受态细胞,通过抗生素筛选和蓝白斑筛选,挑取白色的重组菌落,进行PCR鉴定。阳性菌落提取质粒并送测序。保存测序正确的克隆。

[0182] 重组质粒4V2-E15的全序列如SEQ ID No.17所示;重组质粒4V2-E28的全序列如SEQ ID No.18所示;重组质粒5V1TG-E63的全序列如SEQ ID No.15所示;重组质粒5V1TR-E47的全序列如SEQ ID No.16所示。这四个重组质粒不表达Twin-Strep标签。

[0183] 2.2单基因中间载体的构建

[0184] (1) 各亚基片段的PCR反应及酶切处理

[0185] 以含有目的基因的质粒为模板,进行表3中的第一轮PCR反应。用NEB公司的Q5Premix 完成。反应程序为:98℃,30s;(98℃10s,72℃3min,共30个循环),72℃10min。回收PCR产物。EXOC1、EXOC6的PCR产物用BamHI-HF和XhoI双酶切,EXOC7的PCR产物用KpnI-HF和XhoI双酶切;EXOC2,EXOC3,EXOC4,EXOC5的PCR产物用NotI和 XhoI双酶切;EXOC8的PCR产物用EcoRI-HF和KpnI-HF双酶切。

[0186] (2) 载体的线性化

[0187] 用限制性内切酶BamHI-HF和XhoI双酶切质粒5V1TG,37℃处理1h;然后加入CIP,37℃处理1h;

[0188] 用限制性内切酶KpnI-HF和XhoI双酶切质粒5V1TG,37℃处理1h;然后加入CIP,37℃处理1h;

[0189] 用限制性内切酶NotI和XhoI双酶切质粒5V1TG,37℃处理1h;然后加入CIP,37℃处理1h;

[0190] 用限制性内切酶EcoRI-HF和KpnI-HF双酶切质粒5V1TG,37℃处理1h;然后加入CIP,37℃处理1h;

[0191] 利用胶回收试剂盒回收以上四种约7kb的载体骨架。

[0192] (3) 目的片段与载体骨架的连接反应

[0193] 按照Takara DNA ligation Kit的说明书进行酶切处理的PCR片段与线性化载体的连接反应,以生成5V1TG-SEn系列质粒(n为从1到8的自然数,对应于EXOC1到EXOC8)。连接反应物转化将这两种组装产物分别转化Trans2-blue感受态细胞,通过抗生素筛选和蓝白斑筛选,挑取白色的重组菌落,进行PCR鉴定。阳性菌落提取质粒并送测序。保存测序正确的克隆。

[0194] 重组质粒5V1TG-SE1的全序列如SEQ ID No.7所示;重组质粒5V1TG-SE2的全序列如SEQ ID No.8所示;重组质粒5V1TG-SE3的全序列如SEQ ID No.9所示;重组质粒5V1TG-SE4 的全序列如SEQ ID No.10所示;重组质粒5V1TG-SE5的全序列如SEQ ID No.11所示;重组质粒5V1TG-SE6的全序列如SEQ ID No.12所示;重组质粒5V1TG-SE7的全序列如SEQ ID No.13所示;重组质粒5V1TG-SE8的全序列如SEQ ID No.14所示。这八个重组质粒中EXOC 蛋白的N端与Twin-Strep标签融合表达。

[0195] 2.3最终转移载体E1547与E2863的构建

[0196] (1) 将0.1pmol的重组质粒5V1TR-E47与4V2-E15和1μl的Cre重组酶在20μl反应体系中混匀,30℃孵育1小时。用10μl反应混合物转化100μl Trans2-blue感受态细胞。42℃热休克30s后,加入500μl SOC培养基,30℃振荡孵育4h。将细胞悬浮液铺在含有50μg/ml 卡那霉素和100μg/ml氨苄青霉素的LB琼脂平板上,30℃倒置培养24小时。使用引物Loxp-F (5'-CCACTGCGCCGTTACCAC-3')和Loxp-R (5'-GCCGGTATGTACAGGAAG-3')对重组菌落进行PCR鉴定。阳性菌落可以扩增出375bp的PCR产物。从阳性克隆中提取最终转移质粒E1547。

[0197] (2) 由重组质粒5V1TG-E63与4V2-E28构建出最终转移质粒E2863的过程参考步骤(1)完成。

[0198] 2.4昆虫细胞中试表达人源Exocyst复合体并筛选最适于纯化整个复合物的亚基

[0199] (1) 将8种重组质粒5V1TG-SEn,2种最终转移质粒E1547、E2863转化DH10Bac感受态细胞,通过抗生素筛选和蓝白斑筛选,得到重组菌落。(注释:E1547和E2863为大于19kb的大质粒,因此它们的转化菌在30℃培养以避免基因丢失现象)

[0200] (2) 完成步骤(1)后,提取重组菌落的重组Bacmid DNA,并进行PCR鉴定。5V1TG-SEn的各重组Bacmid用经典的鉴定方法即可(详见Invitrogen Bac to Bac manual)。E1547、E2863质粒产生的Bacmid首先用下面三对引物进行鉴定。Tn7R引物对:5'-GTTTTCCAGTCACGAC-3'和5'-AAGTTTGAGCAGCCGCGTAG-3';Tn7L引物对:5'-CAGGAAACAGCTATGAC-3'和5'-ACCTCCCCCTGAACCTGAAA-3';Empty引物对:5'-GTTTTCCAGTCACGAC-3'和5'-CAGGAAACAGCTATGAC-3'。对于阳性Bacmid,使用“Tn7R”和“Tn7L”引物对,可以扩增出661bp和521bp的PCR产物。如果重组Bacmid被野生型杆粒污染,则使用“Empty引物对”会产生300bp的PCR产物。

[0201] (3) 完成步骤(2)后,进一步用PCR的方法鉴定各条基因在重组Bacmid中真实存在。

[0202] (4) 用Grace's Insect Cell Culture Medium,Unsupplemented昆虫细胞培养基稀释处于对数生长期的sf9细胞,得到稀释液,稀释液中sf9细胞的密度为 $5.0 \times 10^5$ 个/mL。

[0203] (5) 向直径为35mm的平皿中加入2mL步骤(4)得到的稀释液,28℃、培养4h。

[0204] (6) 向溶液B中加入溶液A,混匀,室温放置25min,得到混合物。其中,溶液A的制备方法为向100μL Grace's Insect Cell Culture Medium,Unsupplemented昆虫细胞培养基加入2.5μg步骤(2)提取的重组Bacmid DNA (PCR鉴定为阳性结果);溶液B的制备方法为取100μL Grace's Insect Cell Culture Medium,Unsupplemented昆虫细胞培养基加入8μL cellfectin II。

[0205] (7) 完成步骤(6)后,取所述平皿,逐滴加入步骤(6)得到的混合物,然后28℃、培养4h;弃上清,加入2mL Sf-900™II SFM培养基,28℃静置培养144h。

[0206] (8) 完成步骤(7)后,取所述平皿,用Nikon TS100倒置荧光显微镜观察(使用Nikon B-2A荧光模块观察,激发波长范围450~490nm;使用Nikon-2A荧光模块观察,激发波长范围510~560nm)。若绿色荧光与红色荧光都有较好的表达(图4中a),则收取上清,获得P1代病毒液。5V1TG-SEn质粒、E2863质粒和E1547质粒最终产生的病毒分别命名为BV-SEn、BV-E2863、BV-E1547。

[0207] (9) 完成步骤(8)后,取P1代病毒液200μL,接种于装有180mL、sf9细胞浓度为 $2.0 \times 10^6$ 个/mL的昆虫细胞培养基的2L三角摇瓶中,28℃、150rpm培养96h。然后4℃、2000g离心5min,取上清液,获得P2代病毒液。

[0208] (10) 取5mL BV-E1547P2代病毒液, 5mL BV-E2863-P2代病毒, 2.5mL BV-SEn病毒中的一种(因为BV-SEn病毒一共有8种, 因此这里一共有8种病毒组合, 每种组合含有三种病毒)接种于装有500mL、sf9细胞浓度为 $2.0 \times 10^6$ 个/mL的昆虫细胞培养基的2L三角摇瓶中, 28℃、150rpm培养72h。然后4℃、2000g离心5min, 收集沉淀, 即为SEn感染细胞。将收集的细胞冻存于-80℃。

[0209] (11) 完成步骤(10)后, 取所述感染细胞, 先用裂解缓冲液1重悬, 使用Dounce匀浆器匀浆40次(整个匀浆过程中, 细胞一直处于冰浴中)得到细胞裂解液; 然后4℃、18000rpm离心40min, 收集上清液。

[0210] (12) 完成步骤(11)后, 将收集的上清液上样至Strep自装柱, 先用50个柱体积的裂解缓冲液1洗脱以去除杂蛋白, 再用10个柱体积的洗脱缓冲液洗脱, 收集过柱后的洗脱液进行SDS-PAGE, 结果如图4中b所示。从SDS-PAGE结果上看, 用带有Twin-strep标签的EXOC5进行的纯化, 可以抓到较多的亚基, 且除了带Twin-strep标签的EXOC5之外, 各亚基之间的比例比较均匀, 因此判断从实验结果判断带有Twin-strep标签的EXOC5最适合于纯化整个exocyst复合物。

[0211] 为了得到更加均一的exocyst样品, 本发明进行了第二轮的分子克隆设计。主要是完成4V2-E1S5质粒的构建。在这个质粒中, EXOC5含有Twin-strep标签。具体方法如下:

[0212] 2.5中间转移载体4V2-E1S5的构建

[0213] (1) 以5V1TG-SE5质粒为模板, 引物S5-F与S5-R(序列见表1)及NEB Q5Premix进行PCR反应。反应程序为: 98℃, 30s; (98℃10s, 72℃3min, 共30个循环), 72℃10min。回收PCR产物, 得到S-EXOC5片段。

[0214] (2) 将线性化的4V2载体骨架与EXOC1的第二轮PCR产物、S-EXOC5片段进行Gibson重组, 得到重组质粒4V2-E1S5。

[0215] 重组质粒4V2-E1S5的全序列如SEQ ID No.19所示。

[0216] 2.6最终转移载体E1S547的构建

[0217] 将0.1pmol的重组质粒5V1TR-E47与4V2-E1S5和1μl的Cre重组酶在20μl反应体系中混匀, 30℃孵育1小时。用10μl反应混合物转化100μl的Trans2-blue感受态细胞。42℃热休克30s后, 加入500μl SOC培养基, 30℃振荡孵育4h。将细胞悬浮液铺在含有50μg/ml 卡那霉素和100μg/ml 氨苄青霉素的LB琼脂平板上, 30℃倒置培养24小时。使用引物Loxp-F (5'-CCACTGCGCCGTTACCAC-3') 和Loxp-R (5'-GCCGGTATGTACAGGAAG-3') 对重组菌落进行PCR鉴定。阳性菌落可以扩增出375bp的PCR产物。从阳性克隆中提取最终转移质粒E1S547。

[0218] 2.7在昆虫细胞中表达人源Exocyst复合体并进行电镜负染观察

[0219] (1) 将最终转移质粒E1S547转化DH10Bac感受态细胞, 通过抗生素筛选和蓝白斑筛选, 得到重组菌落。

[0220] (2) 完成步骤(1)后, 提取重组菌落的重组Bacmid DNA, 并进行PCR鉴定。鉴定方法同2.4的步骤(2)和步骤(3)。

[0221] (3) 取得E1S547的P2代病毒, 即BV-E1S547的第二代病毒。具体步骤同2.4的步骤(4)至步骤(9)。

[0222] (4) 完成步骤(3)后, 用5mL BV-E1S547P2代病毒液和5mL BV-E2863-P2代病毒共同感染培养在2L三角摇瓶中的500mL的sf9细胞, 细胞密度为 $2.0 \times 10^6$ 个/mL。细胞培养物在28

℃、150rpm培养72h。然后4℃、2000g离心5min,收集沉淀,冻存于-80℃。

[0223] (5) 完成步骤(4)后,进行蛋白质的纯化。步骤同2.4的步骤(11)和步骤(12)。纯化好的exocyst复合物进行SDS-PAGE,结果如图4中c所示,含有组成复合物的全部8种亚基,各亚基比例相当。

[0224] (6) 完成步骤(5)后,将纯化得到的exocyst复合物用裂解缓冲液1稀释至0.02mg/ml,再用经辉光放电处理的铺有薄碳膜的载网(载网为Life Trust公司产品)吸附样品1min。用裂解缓冲液1将载网洗涤两次,并用4% (w/v) 乙酸铀酰染色2分钟。

[0225] (7) 完成步骤(6)后,用FEI Talos F200C电子显微镜(ThermoFisher,USA)在200kv下观察载网。在28,000倍放大率下,使用4K×4K DE20照相机(Direct Electron,USA),像素尺寸为1.582埃采集图像(如图4中d所示)。离焦值范围从-2.5到-3.5μm。

[0226] (8) 完成步骤(7)后,使用Gctf程序估计衬度传递函数(CTF)。用Gautomatch(<http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/>)和RELION半自动地选取了微粒。共选取了379张显微照片和18669个微粒。使用RELION 2进行二维分类(图4中e)。使用EMAN2产生初始模型(图4中f)。从这些实验结果来看,重组表达的人源exocyst与从酵母中提取的exocyst有着相似的尺寸和形状。

[0227] 2.8对在昆虫细胞中表达的人源Exocyst复合体进行体外活性测定

[0228] Exocyst复合物被认为能够将来自于高尔基体分泌的小泡锚定到细胞质膜上,此活性依赖于GTP/GDP交换蛋白Rab11和细胞质膜上的磷脂PI(4,5)P2以及其它结合细胞质膜的蛋白组分。在体外构建两种脂质体,分别模拟高尔基体小泡和细胞质膜的磷脂成分,且在两种脂质体上分别加入荧光标记的磷脂NBD-PA(Ex:Es:460/534nm) RhodB-DHPE(Ex:Es:560/580)。用重组表达的方法取得人源Rab11(1-173)Q70L蛋白,此蛋白结合GTP,而不水解GTP,因此一直保持在可以结合exocyst的状态。将以上纯化出的人源exocyst复合体以及Rab11(1-173)Q70L蛋白和两种脂质体孵育,如果重组表达的exocyst具有生物学活性,将在Rab11(1-173)Q70L蛋白的协助下将两种脂质体拉近。当两种脂质体之间的距离小于10nm时,使用NBD-PA的激发光激发反应体系,会发生能量共振转移的现象,在580nm附近有明显的转移发射峰。

[0229] 2.8.1人源Rab11(1-173)Q70L蛋白的克隆、表达和纯化

[0230] (1) 原核表达载体pEXS-DH-Rab11Q70L的构建

[0231] 合成的人源Rab11(1-173)Q70L基因用引物Rab11F与Rab11R进行PCR反应。

[0232] Rab11F:5'-AAAACATATGGGCACCCGTGACGACGAGTA-3' ;

[0233] Rab11R:5'-ATTTCTCGAGCCGTAGATCTCGGTGAGGATGGTC-3' 。

[0234] PCR反应使用NEB公司的Q5Premix完成。反应程序为:98℃,30s;(98℃10s,72℃30s,共30个循环),72℃5min。回收PCR产物。所述PCR产物用限制性内切酶NdeI, XhoI处理,并回收。用限制性内切酶NdeI、XhoI酶切原核表达载体pEXS-DH,37℃处理1h;然后加入CIP,37℃处理1h;最后利用胶回收试剂盒回收约5kb的载体骨架。参照2.2的步骤(3)进行目的片段与载体的连接反应。将连接产物转化MT感受态细胞,涂氨苄青霉素抗性的LB平板,37℃过夜培养。挑取单菌落进行PCR鉴定,并送测序。重组质粒 pEXS-DH-Rab11Q70L的全序列如SEQ ID No.20所示。

[0235] (2) 目的蛋白Rab11(1-173)Q70L的表达纯化

[0236] 将测序正确的质粒pEXS-DH-Rab11Q70L转化大肠杆菌BL21 (DE3) 感受态细胞,涂氨苄青霉素抗性的LB平板,37℃过夜培养。挑取单菌落接种5ml带有氨苄青霉素抗性的LB培养基,37℃振荡培养过夜。次日以5ml接种量转接800ml带有氨苄青霉素抗性的LB培养基,37℃摇床培养到OD值为0.6-0.8之间。加入终浓度为0.25mM的IPTG,在16℃诱导表达 20小时。

[0237] 然后菌液在5000rpm,4℃离心10min,收集菌体。用裂解缓冲液2重悬,对细菌重悬液用超声的方法进行破碎。破碎后的菌液用18000rpm,4℃离心35min,上清重复挂用裂解缓冲液平衡好的Ni-NTA介质1ml两次,然后用20ml裂解缓冲液2清洗介质,最后用洗脱缓冲液2洗脱下目的蛋白。

[0238] 将浓缩后的目的蛋白用分子筛Superdex 200 (10/300GL) 进行进一步的纯化,过柱缓冲液为分子筛缓冲液。收集蛋白峰,浓缩后测定浓度,液氮冻存后,保存在-80℃冰箱待用。

[0239] 2.8.2对昆虫细胞表达的人源exocyst复合物进行活性测定

[0240] (1) 使用Extruder制备脂质体

[0241] 将各种磷脂组分用氯仿溶解,按表4中的量和比例加入后,得到A和B两种磷脂混合物。其中磷脂混合物A模拟高尔基体小泡的磷脂成分,而磷脂混合物B模拟细胞质膜的磷脂成分。使用真空泵抽真空4小时,抽去有机溶剂。加入一定体积的测活缓冲液到抽干的两种磷脂混合物中,使磷脂的终浓度为5mg/mL。然后,将磷脂悬液放入37℃水浴锅中,水合30min。之后,用液氮速冻磷脂悬液,再用37℃水浴解冻磷脂悬液。一共反复冻融5次。

[0242] 用装有0.05μm滤膜Extruder制备表4中的脂质体A,(具体仪器操作详见Avanti公司网页上对Extruder的说明);用装有0.2μm滤膜Extruder制备表4中的脂质体B。

[0243] 制备好的脂质体放到4℃,当天使用。

[0244] 表4 脂质体A、脂质体B所含的磷脂成分配方

| 脂质体 A       |         | 脂质体 B     |         |
|-------------|---------|-----------|---------|
| 磷脂名称        | 磷脂用量    | 磷脂名称      | 磷脂用量    |
| DOPC        | 0.5mg   | DOPC      | 0.5mg   |
| DOPE        | 0.2mg   | DOPE      | 0.2mg   |
| DOPS        | 0.07mg  | DOPS      | 0.07mg  |
| PI          | 0.145mg | PI        | 0.145mg |
| SM          | 0.075mg | SM        | 0.075mg |
| DOGS-NTA-Ni | 0.02mg  | PI(4,5)P2 | 0.02mg  |
| Rhod B-DHPE | 0.33mg  | NBD-PA    | 0.02mg  |

[0247] (2) 使用荧光分光光度计测定exocyst的活性

[0248] 表5 Exocyst复合物的体外测活反应体系

[0249]

|                        | 对照组<br>AB     | 对照组<br>AB-Rab11 | 对照组<br>AB-EXOC | 实验组<br>AB-Rab11-EXOC |
|------------------------|---------------|-----------------|----------------|----------------------|
| 5mg/ml 脂质体 A           | 24 $\mu$ l    | 24 $\mu$ l      | 24 $\mu$ l     | 24 $\mu$ l           |
| 5mg/ml 脂质体 B           | 24 $\mu$ l    | 24 $\mu$ l      | 24 $\mu$ l     | 24 $\mu$ l           |
| 157mg/ml Rab11Q70L     | 0             | 1 $\mu$ l       | 0              | 1 $\mu$ l            |
| 2mg/ml exocyst 复合物     | 0             | 0               | 22.5 $\mu$ l   | 22.5 $\mu$ l         |
| 测活缓冲液                  | 420.5 $\mu$ l | 420.5 $\mu$ l   | 420.5 $\mu$ l  | 420.5 $\mu$ l        |
| 0.2M MgCl <sub>2</sub> | 2.5 $\mu$ l   | 2.5 $\mu$ l     | 2.5 $\mu$ l    | 2.5 $\mu$ l          |
| 0.1M GTP               | 5 $\mu$ l     | 5 $\mu$ l       | 5 $\mu$ l      | 5 $\mu$ l            |
| 1M DTT                 | 0.5 $\mu$ l   | 0.5 $\mu$ l     | 0.5 $\mu$ l    | 0.5 $\mu$ l          |

[0250] 按表5体系准备对照组和实验组反应体系。此处以准备实验组反应体系为例,说明测活各组分的添加顺序。先向EP管中加入测活缓冲液、MgCl<sub>2</sub>,GTP与DTT,充分混匀;再向EP管内加入脂质体A和Rab11Q70L蛋白,在冰上孵育1-2h。Hitaiki F7000荧光分光光度计提前预热好,在测定实验即将开始前,向孵育好的脂质体A和Rab11Q70L混合物中加入 exocyst复合体和脂质体B并混匀。之后立即加入1mL石英比色皿,放入机器中进行测量。测定模式为波长扫描,激发光为460nm,发射扫描波长范围为470-700nm,电压500V。测定 37℃水浴条件下孵育30min后的荧光强度,先测对照组,再测定实验组,实验结果如图5 所示。可见昆虫细胞表达的人源exocyst复合物具有生物学活性。

[0251] 如图5中a所示,四个组中实验组(脂质体A+Rab11Q70L+Exocyst+脂质体B)相对于其它三个对照组在585nm处,有明显的荧光发射峰(见图5中a中的箭头处),表明有明显的荧光能量共振转移,即exocyst复合物将一定量A和B两种脂质体拉近且距离小于10nm,这表明体外纯化的exocyst复合物具有将两种膜锚定到一起的活性。而只加入Rab11Q70L或exocyst复合物的对照组,相对于没有加入任何一种蛋白的对照组AB,均没有明显的荧光能量共振转移,这表明Rab11Q70L和exocyst单独存在时均不能将两种脂质体拉近,也说明了Exocyst发挥功能依赖于结合GTP的Rab11。图5中b图显示了三次独立实验的平均值及误差,表明测活结果是可信的。纵坐标是三个对照组与实验组在585nm处荧光值与533nm处荧光值的比。533nm处的荧光是NBD-PA在460nm激发光下的发射波长。

<110> 中国科学院生物物理研究所

<120> SmartBac杆状病毒表达系统及其应用

<160> 20

<170> PatentIn version 3.5

<210> 1

<211> 7187

<212> DNA

<213> 人工序列

<220>

<223>

<400> 1

```

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60
taaactatga caataaagtc tttaaactaga cagaatagtt gtaaactgaa atcagtcag 120
ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180
gaagtgcaaa ttgcccgtcg tattaagag gggcgtggcc aaggcatgt aaagactata 240
ttcgcgcggt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300
aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatactt cttcccgtat 360
gccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420
cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480
gccctgcctc cggtgctcgc cggagactgc gagatcatag atatagatct cactacgcgg 540
ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600
ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660
ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720
ccgcatggat ttgacttgggt cagggccgag cctacatgtg cgaatgatgc ccatacttga 780
gccacctaac ttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840
catcgttgct gtccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900
atgcccaggg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960
cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgtc 1020
acttgcattha cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080
gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140
atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200
cactagtctg cgacctactc cggaatatta ataggttgcg gatatcggga gttcagtcgt 1260
cgaatgcaaa gcgtaaaaaa tattaataag gtaaaaatta cagctacata aattacacaa 1320
tttaaacgga tcgatcatat ggcttatcct tacgacgtgc ctgactacgc cggagagagc 1380
ttgtttaagg ggccgcgtga ttacaacct atatcgagca ccatttgtea ttgacgaat 1440
gaatctgatg ggcacacaac atcgttgtat ggtattggat ttggtccctt catcattaca 1500
aacaagcact tgtttagaag aaataatgga aactgttgg tccaatcact acatggtgta 1560
ttcaaggtea agaacaccac gactttgcaa caacacctca ttgatgggag ggacatgata 1620

```

attattcgca tgcctaagga tttcccacca tttcctcaaa agctgaaatt tagagagcca 1680  
 caaaggggaag agagaatatg tcttgtgaca accaacttcc aaactaagag catgtctagc 1740  
 atgggtgtcag acacttcttg cacattccct tcatctgatg gcatattctg gaagcattgg 1800  
 attcaaacca aggatgggca gtgtggcagt ccattagtat caactagaga tgggttcatt 1860  
 gttggtatac actcagcatc gaatttcacc aacacaaaca attatttcac aagcgtgccg 1920  
 aaaaacttca tggaattgtt gacaaatcag gaggcgcagc agtgggttag tggttggcga 1980  
 ttaaatgctg actcagtatt gtgggggggc cataaagttt tcatggtgaa acctgaagag 2040  
 ccttttcagc cagttaagga agcgactcaa ctcatgaatg aattggtgta ctcgcaagaa 2100  
 aacctgtact tccagtcagc ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc 2160  
 ggaggtagcg gcggaggttc ttggtctcac cctcagttcg agaaggatga cgatgataaa 2220  
 accatgggat ccctaggtac cgcggcccg cgcgttgccc gattcattaa tgcagctggc 2280  
 acgacaggtt tcccgaactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc 2340  
 tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa 2400  
 ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaaactat 2460  
 ttaggtgacg cgtagaata ctcaagctat gcatcatctt tggttccgtc atcggacca 2520  
 ttagtaacgg ccgccagtgt gctggagttt ttagataacc catcacactg gcgtccactg 2580  
 gaacatgcaa gtagagggcc caattcgccc tatagttagt cgtattacaa ttcactggcc 2640  
 gtcgttttac aacgtcgtga ctgggaaaac cctggcgtaa cccaacttaa tcgccttgca 2700  
 gcacatcccc ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttc 2760  
 caacagttgc gcagcctata cgtacggtaa ctgactaaga attccgatta caaagacgat 2820  
 gacgacaagg gctcgagtga aaatttgtat tttcaaagct cgtcgacggg gagcaagggc 2880  
 gaggagctgt tcaccggggg ggtgcccatc ctggtcgagc tggacggcga cgtaaacggc 2940  
 cacaagttca gcgtgtccgg cgaggcgag ggcatgcc cctacggcaa gctgaccctg 3000  
 aagttcatct gcaccaccgg caagctgccc gtgccctggc ccaccctcgt gaccaccctg 3060  
 acctacggcg tgcagtgtt cagccgtac cccgaccaca tgaagcagca cgacttcttc 3120  
 aagtcgcca tgcccgaagg ctacgtccag gagcgacca tcttcttcaa ggacgacggc 3180  
 aactacaaga cccgcgccga ggtgaagttc gagggcgaca ccctggtgaa ccgcatcgag 3240  
 ctgaagggca tcgacttcaa ggaggacggc aacatcctgg ggcaaaagct ggagtacaac 3300  
 tacaacagcc acaacgtcta tatcatggcc gacaagcaga agaacggcat caaggtgaac 3360  
 ttcaagatcc gccacaacat cgaggacggc agcgtgcagc tcgccacca ctaccagcag 3420  
 aacaccccc tggcgacgg cccgtgtctg ctgcccga accactacct gagcaccag 3480  
 tccgccctga gcaaagaccc caacgagaag cgcgatcaca tggctctgct ggagttcgtg 3540  
 accgccgccg ggatcactct cggcatggac gagctgtaca agtaactgac taaaagcttc 3600  
 gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa ccccttgggg 3660  
 cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatctc agggtcgaga 3720  
 agtactagag gatcataatc agccatacca catttgtaga ggttttactt gctttaaaaa 3780  
 acctcccaca cctccccctg aacctgaaac ataaaatgaa tgcaattgtt gttgttaact 3840  
 tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaat ttcacaaata 3900  
 aagcattttt ttcactgcat tctagttgtg gttgtccaa actcatcaat gtatcttctc 3960

atgtctggat ctgatcactg cttgagccta ggagatccga accagataag tgaaatctag 4020  
 ttccaaacta ttttgtcatt ttttaatttc gtattagctt acgacgctac acccagttcc 4080  
 catctatttt gtcactcttc cctaaataat ccttaaaaac tccatttcca cccctcccag 4140  
 ttcccaacta ttttgtccgc ccacagcggg gcatttttct tcctgttatg tttttaatca 4200  
 aacatcctgc caactccatg tgacaaaccg tcatcttcgg ctactttttc tctgtcacag 4260  
 aatgaaaatt tttctgtcat ctcttcggtt ttaatgtttg taattgactg aatatcaacg 4320  
 cttattttgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatg aagcgcggcg 4380  
 ggtgtgggtg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgtcctt 4440  
 ttgcgtttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat 4500  
 cgggggctcc ctttaggggtt ccgatttagt gctttacggc acctcgacct caaaaaactt 4560  
 gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg 4620  
 acgttggagt ccacgttctt taatagtga ctcttgttcc aaactggaac aacactcaac 4680  
 cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattgggtta 4740  
 aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca 4800  
 atttcagggtg gcacttttcg gggaaatgtg cgcggaacct ctatttgttt atttttctaa 4860  
 atacattcaa atatgtatcc gctcatgaga caataacctt gataaatgct tcaataatat 4920  
 tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg 4980  
 gcattttgcc ttctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa 5040  
 gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt 5100  
 gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt 5160  
 ggcgcggtat tatcccgat tgacgccggg caagagcaac tcggtcgccg catacactat 5220  
 tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg 5280  
 acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta 5340  
 cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaa catgggggat 5400  
 catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag 5460  
 cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa 5520  
 ctacttactc tagcttccc gcaacaatta atagactgga tggaggcgga taaagttgca 5580  
 ggaccacttc tgcgtcggc cttccggct ggctggttta ttgctgataa atctggagcc 5640  
 ggtgagcgtg ggtctcgcgg tatcattgca gcaactgggc cagatggtaa gccctcccgt 5700  
 atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc 5760  
 gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat 5820  
 atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt 5880  
 tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 5940  
 cgcggggcat gactaacatg agaattacaa cttatatcgt atggggctga cttcaggtgc 6000  
 tacatttgaa gagataaatt gcactgaaat ctagaaatat tttatctgat taataagatg 6060  
 atcttcttga gatcgttttg gtctgcgcgt aatctcttgc tctgaaaacg gaaaaaacg 6120  
 ccttgacagg cggtttttcg aaggttctct gagctaccaa ctctttgaac cgaggtaact 6180  
 ggcttgagg agcgcagtca ccaaaacttg tcctttcagt ttagccttaa ccggcgcatg 6240  
 acttcaagac taactcctct aaatcaatta ccagtggctg ctgccagtgg tgcttttgca 6300

tgtctttccg ggttggactc aagacgatag ttaccggata aggcgccagcg gtcggactga 6360  
 acgggggggtt cgtgcataca gtccagcttg gagcgaactg cctacccgga actgagtgtc 6420  
 aggcgtggaa tgagacaaac gcggccataa cagcggaatg acaccggtaa accgaaaggc 6480  
 aggaacagga gagcgcacga gggagccgcc aggggaaacg cctggtatct ttatagtcct 6540  
 gtcgggtttc gccaccactg atttgagcgt cagatttcgt gatgcttgtc aggggggcgg 6600  
 agcctatgga aaaacggctt tgccgcggcc ctctcacttc cctgttaagt atcttcctgg 6660  
 catcttccag gaaatctccg ccccgttcgt aagccatttc cgctcgccgc agtcgaacga 6720  
 ccgagcgtag cgagtcagtg agcgaggaag cggaatatat cctgtatcac atattctgct 6780  
 gacgcaccgg tgcagccttt tttctctgc cecatgaagc acttactga caccctcatc 6840  
 agtgccaaca tagtaagcca gtataactc cgctagcgt gatgtccggc ggtgcttttg 6900  
 ccgttacgca ccaccccgtc agtagctgaa caggaggac agctgataga aacagaagcc 6960  
 agttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttagagtga 7020  
 ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtga gcggaagcgg 7080  
 aagagcgcct gatgcggtat tttctcctta cgcactgtg cggtatttca caccgcatag 7140  
 accagccgcg taacctggca aaatcggtta cggttgagta ataatg 7187

<210> 2

<211> 7187

<212> DNA

<213> 人工序列

<220>

<223>

<400> 2

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtccag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgcccgtcg tattaaagag gggcgtggcc aaggcatgt aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatcactt ctcccgtat 360  
 gcccaacttt gtatagagag ccaactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cggtaggcaat 480  
 gccctgcctc cgggtgctgc cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcggtgc tgctgcgtaa 840  
 catcgttgct gctccataac atcaaacatc gaccacgcgc gtaacgcgct tgctgcttgg 900  
 atgcccaggc catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgtc 1020

acttgcatta cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatcc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200  
 cactagtctg cgacctactc cggaatatta ataggttgct gatatcggga gttcagtcgt 1260  
 cgaatgcaaa gcgtaaaaaa tattaataag gtaaaaatta cagctacata aattacacaa 1320  
 tttaaacgga tcgatcatat ggcttatect tacgacgtgc ctgactacgc cggagagagc 1380  
 ttgtttaagg ggccgcgtga ttacaacct atatecgagca ccatttgtca ttgacgaat 1440  
 gaatctgatg ggacacacaac atcgttgtat ggtattggat ttggtccctt catcattaca 1500  
 aacaagcact tgtttagaag aaataatgga acactgttgg tccaatcact acatggtgta 1560  
 ttcaagggtca agaacaccac gactttgcaa caacacctca ttgatgggag ggacatgata 1620  
 attattcgca tgcctaagga tttcccacca tttctcaaa agctgaaatt tagagagcca 1680  
 caaagggaag agagaatatg tcttgtgaca accaacttcc aaactaagag catgtctagc 1740  
 atggtgtcag acacttcttg cacattccct tcatctgatg gcattattctg gaagcattgg 1800  
 attcaaacca aggatgggca gtgtggcagt ccattagtat caactagaga tgggttcatt 1860  
 gttgttatac actcagcatc gaatttcacc aacacaaaca attatttcac aagcgtgccg 1920  
 aaaaacttca tggattgtt gacaaatcag gaggcgcagc agtgggttag tggttggcga 1980  
 ttaaatgctg actcagtatt gtgggggggc cataaagttt tcatggtgaa acctgaagag 2040  
 ccttttcagc cagttaagga agcgactcaa ctcatgaatg aattggtgta ctgcgaagaa 2100  
 aacctgtact tccagtcagc ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc 2160  
 ggaggtagcg gcggaggttc ttggtctcac cctcagttcg agaaggatga cgatgataaa 2220  
 accatgggat ccctaggtac cgcggccgcg cgcgttgccc gattcattaa tgcagctggc 2280  
 acgacagggt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc 2340  
 tcaactatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa 2400  
 ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat 2460  
 ttaggtgacg cgtagaata ctcaagctat gcattatctt tggttccgtc atcggaccca 2520  
 ttagtaacgg ccgccagtgt gctggagttt ttagataacc catcacactg gcgtccactg 2580  
 gaacatgcaa gtagagggcc caattcgccc tatagtgagt cgtattacaa ttcactggcc 2640  
 gtcgttttac aacgtcgtga ctgggaaaac cctggcgtaa cccaacttaa tcgcttgca 2700  
 gcacatcccc ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgccccttc 2760  
 caacagttgc gcagcctata cgtacggtaa ctgactaaga attccgagca aaagttgatt 2820  
 agcgaagaag acttaggtc gagtgaaaat ttgtattttc aaagctcgtc gacggtgtct 2880  
 aagggcgaag agctgattaa ggagaacatg cacatgaagc tgtacatgga gggcaccgtg 2940  
 aacaaccacc acttcaagtg cacatccgag ggccaaggca agccctacga gggcaccag 3000  
 accatgagaa tcaagtggt cgagggcggc cctctcccc tgccttcga catcctggt 3060  
 accagcttca tgtacggcag cagaaccttc atcaaccaca cccagggeat cccgacttc 3120  
 tttaagcagt ccttcctga gggttcaca tgggagagag tcaccacata cgaagacggg 3180  
 ggctgtctga ccgtaccca ggacaccagc ctccaggacg gctgcctcat ctacaacgtc 3240  
 aagatcagag ggggtgaactt cccatccaac ggccctgtga tgcagaagaa aacactcggc 3300  
 tgggaggcca acaccgagat gctgtacccc gctgacggcg gcctggaagg cagaagcgac 3360

atggccctga agctcgtggg cggggggccac ctgatctgca acttcaagac cacatacaga 3420  
 tccaagaaac ccgctaagaa cctcaagatg cccggcgctct actatgtgga ccacagactg 3480  
 gaaagaatca aggaggccga caaagagacc tacgtcgagc agcacgaggt ggctgtggcc 3540  
 agatactgcg acctccctag caaactgggg cacaaactta attaactgac taaaagcttc 3600  
 gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa ccccttgggg 3660  
 cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatcctc agggctcgaga 3720  
 agtactagag gatcataatc agccatacca cattttaga ggttttactt gctttaaaaa 3780  
 acctcccaca cctccccctg aacctgaaac ataaaatgaa tgcaattgtt gttgttaact 3840  
 tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaat ttcacaaata 3900  
 aagcattttt ttcactgcat tctagttgtg gtttgtccaa actcatcaat gtatcttate 3960  
 atgtctggat ctgatcactg cttgagccta ggagatccga accagataag tgaaatctag 4020  
 ttccaaacta ttttgtcatt ttttaatttc gtattagctt acgacgctac acccagttcc 4080  
 catctatttt gtcactcttc cctaaataat ccttaaaaac tccatttcca cccctcccag 4140  
 ttcccaacta ttttgtccgc ccacagcggg gcatttttct tctgtttatg tttttaatca 4200  
 aacatcctgc caactccatg tgacaaaccg tcatcttcgg ctacttttcc tctgtcacag 4260  
 aatgaaaatt tttctgtcat ctcttcgtta ttaatgtttg taattgactg aatatcaacg 4320  
 cttatttgca gcctgaatgg cgaatgggac gcgccctgta gcggcgctatt aagcgcggcg 4380  
 ggtgtgggtg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct 4440  
 ttgcgtttct tcccttcctt tctcgccacg ttcgcccgtt tccccgtca agctctaaat 4500  
 cgggggctcc ctttaggggt cggatttagt gctttacggc acctcgacct caaaaaactt 4560  
 gattaggggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg 4620  
 acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac 4680  
 cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctatttggtta 4740  
 aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca 4800  
 atttcagggtg gcacttttcg gggaaatgtg cgcggaacct ctatttgttt atttttctaa 4860  
 atacattcaa atatgtatcc gctcatgaga caataacctt gataaatgct tcaataatat 4920  
 tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg 4980  
 gcattttgcc ttccgttttt tgcacacca gaaacgctgg tgaaagtaaa agatgctgaa 5040  
 gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt 5100  
 gagagttttc gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt 5160  
 ggcgcggtat tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat 5220  
 tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg 5280  
 acagtaagag aattatgcag tgetgccata accatgagtg ataacactgc ggccaactta 5340  
 cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaa catgggggat 5400  
 catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag 5460  
 cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa 5520  
 ctacttactc tagcttcccc gcaacaatta atagactgga tggaggcgga taaagttgca 5580  
 ggaccacttc tgcgctcggc cttccggct ggctggttta ttgctgataa atctggagcc 5640  
 ggtgagcgtg ggtctcgcgg tatcattgca gcaactgggc cagatggtta gccctccgt 5700

atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc 5760  
 gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat 5820  
 atacttttaga ttgatttaaa acttcatittt taattttaaaa ggatctaggt gaagatcctt 5880  
 tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 5940  
 cgcggggcat gactaacatg agaattacaa cttatatcgt atggggctga cttcaggtgc 6000  
 tacatttgaa gagataaatt gcaactgaaat ctagaaatat tttatctgat taataagatg 6060  
 atcttcttga gatcgttttg gtctgcgcgt aatctcttgc tctgaaaacg gaaaaaacg 6120  
 ccttgcaagg cggtttttcg aaggttctct gagctaccaa ctctttgaac cgaggtaact 6180  
 ggcttgagg agcgagtc ccaaaacttg tcttttcagt ttagccttaa ccggcgcatg 6240  
 acttcaagac taactcctct aaatcaatta ccagtggctg ctgccagtgg tgcttttgca 6300  
 tgtctttccg ggttggaact aagacgatag ttaccggata aggcgcagcg gtcggactga 6360  
 acgggggggt cgtgcataca gtccagcttg gagcgaactg cctaccgga actgagtgtc 6420  
 aggcgtggaa tgagacaaac gcggccataa cagcggaatg acaccgtaa accgaaaggc 6480  
 aggaacagga gagcgcacga gggagccgcc aggggaaacg cctggtatct ttatagtcct 6540  
 gtcgggttcc gccaccactg atttgagcgt cagatttcgt gatgcttgc aggggggcg 6600  
 agcctatgga aaaacggctt tgccgcggcc ctctcacttc cctgttaagt atcttcttg 6660  
 catcttccag gaaatctccg ccccgttcgt aagccatttc cgctcgccgc agtcgaacga 6720  
 ccgagcgtag cgagtcagt agcgaggaag cggaatatat cctgtatcac atattctgct 6780  
 gacgcaccgg tgcagccttt tttctcctgc cacatgaage acttcaactga caccctcatc 6840  
 agtgccaaca tagtaagcca gtatacactc cgctagcgt gatgtccggc ggtgcttttg 6900  
 ccgttacgca ccacccgctc agtagctgaa caggaggac agctgataga aacagaagcc 6960  
 agttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagttag 7020  
 ctgataccgc tcgccgcagc cgaacgaccg agcgagcga gtcagtgagc gaggaagcgg 7080  
 aagagcgcct gatgcggtat tttctcctta cgcactctgt cggtatttca caccgcatag 7140  
 accagccgcg taacctggca aaatcggtta cggttgagta ataaatg 7187

<210> 3

<211> 7737

<212> DNA

<213> 人工序列

<220>

<223>

<400> 3

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc tttaaactaga cagaatagtt gtaaactgaa atcagtcag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcatTTTTT 180  
 gaagtgcaaa ttgccgctcg tattaaagag gggcgtggcc aagggcagtgt aaagactata 240  
 ttgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcactactt cttcccgat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420

cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggg cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840  
 catcgttgct gtcacataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccggagg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgtc 1020  
 acttgcatca cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatcc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacgtt cctaaggtag cgagttttaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgcgg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440  
 ggcccctcga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560  
 gatcccgccg gcggtcacga actccagcag gaccatgtga tcgcgttctt cgttgggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggttgctg ggcagcagca cggggccgtc 1680  
 gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgctgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac 1800  
 gttgtggctg ttgtagtgtt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860  
 gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctcggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040  
 ctgcacgccg taggtcaggg tggtcacgag gttgggccag ggcacgggca gcttgccgtt 2100  
 ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctcgc cctcgccgga 2160  
 cacgtgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220  
 ggtgaacagc tcctcgccct tgetcatecc gggtgactgg aagtacaggt tttcttgcca 2280  
 gtacaccaat tcattcatga gttgagtcgc ttcttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400  
 actaacccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaataa ttgttttgtt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520  
 atctctagtt gataactaat gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580  
 ccagaatatg ccatcagatg aagggaatgt gcaagaagtg tctgacacca tgctagacat 2640  
 gctcttagtt tggaagttag ttgtcacaag acatattctc tcttcccttt gtggctctct 2700  
 aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760

cctcccatca atgaggtggt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgcctgt ttgtaatgat 2880  
 gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaaaca agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggcctacgtg acacaaggca cgttggttaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttta acactagttc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540  
 cgcgcccgcg cgcgttggcc gattcattaa tgcagctggc acgacaggtt tcccgactgg 3600  
 aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcactcatta ggcacccag 3660  
 gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt 3720  
 cacacaggaa acagctatga ccatgattac gccaagctat ttaggtgacg cgttagaata 3780  
 ctcaagctat gcatcatctt tggttccgtc atcggacca ttagtaacgg ccgccagtgt 3840  
 gctggagttt tgtagatacc catcacactg gcgtccactg gaacatgcaa gtagagggcc 3900  
 caattcgccc tatagtgagt cgtattacaa ttcactggcc gtcgttttac aacgtcgtga 3960  
 ctgggaaaac cctggcgta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 4020  
 ctggcgtaat agcgaagagg cccgcaccga tcgcccttc caacagttgc gcagcctata 4080  
 cgtacggtaa ctgactaaga attccttaag cggaggcctg cagggtcga gttactgac 4140  
 taaaagcttc gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa 4200  
 ccccttgggg cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatcctc 4260  
 agggctcgaga agtactagag gatcataatc agccatacca catttgtaga ggttttactt 4320  
 gctttaaaaa acctcccaca cctccccctg aacctgaaac ataaaatgaa tgcaattgtt 4380  
 gttgttaact tgttttattg agcttataat ggttacaaat aaagcaatag catcacaaat 4440  
 ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa actcatcaat 4500  
 gtatcttata atgtctggat ctgatcactg cttgagccta ggagatccga accagataag 4560  
 tgaaatctag ttccaaacta ttttgtcatt tttattttc gtattagctt acgacgtac 4620  
 acccagttcc catctatttt gtcactcttc ctaataat ccttaaaaac tccatttcca 4680  
 cccctcccag ttcccaacta ttttgtccgc ccacagcggg gcatttttct tctgttatg 4740  
 tttttaatca aacatcctgc caactccatg tgacaaaccg tcattcttcgg ctacttttct 4800  
 tctgtcacag aatgaaaatt tttctgtcat ctcttcgtta ttaatgtttg taattgactg 4860  
 aatatcaacg cttatttgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt 4920  
 aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgt acacttgcca gcgccctagc 4980  
 gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcgcgggt tccccgtca 5040  
 agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc 5100

caaaaaactt gattaggggtg atggttcacg tagtggggcca tcgccctgat agacgggtttt 5160  
 tcgccctttg acgttggagt ccacgttctt taatagtgga ctcttggtcc aaactggaac 5220  
 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 5280  
 ctattgggta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt 5340  
 aacgtttaca atttcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt 5400  
 atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct 5460  
 tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc 5520  
 cttttttgcg gcattttgcc ttccgttttt tgctcaccca gaaacgctgg tgaaagtaaa 5580  
 agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg 5640  
 taagatcctt gagagtttcc gccccgaaga acgttttcca atgatgagca cttttaaaagt 5700  
 tctgctatgt ggcgcggtat tatcccgat tgacgcgggg caagagcaac tcggtcgccc 5760  
 catacactat tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac 5820  
 ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc 5880  
 ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaaa 5940  
 catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc 6000  
 aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaactatt 6060  
 aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga 6120  
 taaagttgca ggaccacttc tgcgctcggc cttccggtt ggctggttta ttgctgataa 6180  
 atctggagcc ggtgagcgtg ggtctcggcg tatcattgca gcactggggc cagatggtaa 6240  
 gccctcccg atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa 6300  
 tagacagatc gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt 6360  
 ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt 6420  
 gaagatcctt ttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg 6480  
 agcgtcagac cgcggggcat gactaacatg agaattacaa cttatatcgt atggggctga 6540  
 cttcaggtgc tacatttgaa gagataaatt gcactgaaat ctagaaatat tttatctgat 6600  
 taataagatg atcttcttga gatcgttttg gtctgcgctt aatctcttgc tctgaaaacg 6660  
 gaaaaaacgg ccttgcaggg cggtttttcg aaggttctct gagctaccaa ctctttgaac 6720  
 cgaggtaact ggcttggagg agcgcagtea ccaaaacttg tcctttcagt ttagccttaa 6780  
 ccggcgcatg acttcaagac taactcctct aatcaatta ccagtggctg ctgccagtgg 6840  
 tgcttttgca tgtctttccg ggttggactc aagacgatag ttaccggata aggcgcagcg 6900  
 gtcggactga acgggggggtt cgtgcataca gtccagcttg gagcgaactg cctaccggga 6960  
 actgagtgtc aggcgtggaa tgagacaaac gcggccataa cagcggaaatg acaccggtaa 7020  
 accgaaaggc aggaacagga gagcgcacga gggagccgcc aggggaaacg cctggtatct 7080  
 ttatagtcct gtcgggtttc gccaccactg atttgagcgt cagatttcgt gatgcttgtc 7140  
 agggggggcg agcctatgga aaaacggctt tgccgcggcc ctctcacttc cctgttaagt 7200  
 atcttctggt catcttccag gaaatctccg ccccgctcgt aagccatttc cgtcgcgcgc 7260  
 agtcgaacga ccgagcgtag cgagtcagt agcgaggaag cggaatatat cctgtatcac 7320  
 atattctgct gacgcaccgg tgcagccttt tttctcctgc cacatgaagc acttactga 7380  
 caccctcatc agtgccaaca tagtaagcca gtatacactc cgctagcgt gatgtccggc 7440

ggtgcttttg ccgttacgca ccaccccgtc agtagctgaa caggaggac agctgataga 7500  
 aacagaagcc agttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc 7560  
 tttgagttag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgage 7620  
 gaggaagcgg aagagcgcct gatgcggtat tttctcctta cgcattctgt cggtatttca 7680  
 caccgcatag accagccgcg taacctggca aaatcggta cggttgagta ataaatg 7737

<210> 4

<211> 7734

<212> DNA

<213> 人工序列

<220>

<223>

<400> 4

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc tttaaactaga cagaatagtt gtaaactgaa atcagtcacg 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgcccgtcg tattaaagag gggcgtggcc aagggcattg aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatcactt cttcccgat 360  
 gcccaacttt gtatagagag cactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcggttg tgctgcgtaa 840  
 catcggttgc gtcacataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccagag catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatga cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagttttaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagt gcgaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cttttgttg atagcgcggt tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt ttttctatg caaaaccca caccaacaac 1440  
 ggcccctcga taataaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttaattaagt ttgtgcccc gtttgctagg 1560  
 gaggtcgcag tatctggcca cagccacctc gtgctgctcg acgtaggtct ctttgcggc 1620

ctccttgatt ctttccagtc tgttggtccac atagtagacg ccgggcatct tgaggttctt 1680  
agcgggtttc ttggatctgt atgtggtctt gaagttgcag atcaggtggc ccccgccac 1740  
gagcttcagg gccatgtcgc ttctgccttc caggccgccg tcagcggggg acagcatctc 1800  
gggtgttgcc tcccagccga gtgttttctt ctgcatcaca gggccgttgg atgggaagtt 1860  
caccctctg atcttgacgt ttagatgag gcagccgtcc tggaggctgg tgcctgggt 1920  
agcggtcagc acgccccgt cttcgtatgt ggtgactctc tccatgtga agcctcagg 1980  
gaaggactgc ttaaagaagt cggggatgcc ctgggtgtgg ttgatgaagg ttctgctgcc 2040  
gtacatgaag ctggtagcca ggatgtcgaa ggcaagggg agaggccgc cctcgaccac 2100  
cttgattctc atggtctggg tgccctcgta gggcttgcct tcgccctcgg atgtgcactt 2160  
gaagtgggtg ttgttcacgg tgccctccat gtacagcttc atgtgcatgt tctccttaat 2220  
cagctcttcg cccttagaca ccateccggg tgactggaag tacaggtttt cttgcgagta 2280  
caccaattca ttcattgagt gagtcgcttc cttactggc tgaaaaggct cttcaggttt 2340  
caccatgaaa actttatggc cccccacaa tactgagtc gcatttaatc gccaccact 2400  
aaccactgc tgcgcctcct gatttgtcaa caattccatg aagtttttcg gcacgcttgt 2460  
gaaataattg tttgtgttg tgaaattcga tgctgagtgt ataccaacaa tgaaccatc 2520  
tctagttgat actaatggac tgccacactg cccatccttg gtttgaatcc aatgcttcca 2580  
gaatatgcca tcagatgaag ggaatgtgca agaagtgtct gacaccatgc tagacatgct 2640  
cttagtttg aagttggtg tcacaagaca tattctctct tccctttgtg gctctctaaa 2700  
tttcagcttt tgaggaaatg gtgggaaatc cttaggcatg cgaataatta tcatgtccct 2760  
cccatcaatg aggtgttgt gcaaagtcgt ggtgttcttg acctgaata caccatgtag 2820  
tgattggacc aacagtgttc cattatttct tctaaacaag tgcttgtttg taatgatgaa 2880  
gggaccaaatt ccaataccat acaacgatgt tgtgtgcccc tcagattcat tcgtcaaattg 2940  
acaaatgggt ctcgatatag ggttgtaatc acgcggcccc ttaaacaagc tctctccggc 3000  
gtagtcaggc acgtcgtgaa gataagccat atttaaatat atgcttgctt gtgtgttctt 3060  
tattgaagcc ttggtgtgac tgatttacta gtagcgttga ggcgtcttat atacccgacc 3120  
gttatctggc ctacgtgaca caaggcacgt tgttagatta ataactttat ctttttatct 3180  
taattgataa gattattttt atctggtgt tataaaaacg ggatcatgaa cacggacgct 3240  
cagtcgacag atctgtcgac ggtttaaaca ctagttcgcg acctactccg gaatattaat 3300  
aggttgctga tatcgggagt tcagtcgtcg aatgcaaagc gtaaaaaata ttaataaggt 3360  
aaaaattaca gctacataaa ttacacaatt taaacggatc gatgagctcc atatggcctg 3420  
gagccatccg caatttgaaa aaggtggcgg gtccggcgga ggtagcggcg gaggttcttg 3480  
gtctcaccct cagttcgaga aggatgacga tgataaaacc atgggatccc taggtaccgc 3540  
ggccgcgcgc gttggccgat tcattaatgc agctggcacg acaggtttcc cgactggaaa 3600  
gcgggcagtg agcgcaacgc aattaatgtg agttagctca ctcattagge acccagget 3660  
ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcggata acaatttcac 3720  
acaggaaaca gctatgacca tgattacgcc aagctattta ggtgacgct tagaatactc 3780  
aagctatgca tcacttttgg ttccgtcatc ggaccatta gtaacggccg ccagtgtgct 3840  
ggagttttgt agatacccat cacttgccg tccactggaa catgcaagta gagggcccaa 3900  
ttcgccctat agtgagtcgt attacaattc actggccgct gttttacaac gtcgtgactg 3960

ggaaaaccct ggcgttacct aacttaatcg ccttgcagca catccccctt tcgccagctg 4020  
 gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctatacgt 4080  
 acggtaactg actaagaatt ccttaagcgg aggccctgcag ggctcgagtt aactgactaa 4140  
 aagcttcgaa aggaagctga gttggctgct gccaccgctg agcaataact agcataaccc 4200  
 cttggggcct ctaaacgggt cttgaggggt tttttgctga aaggaggaaac tatcctcagg 4260  
 gtcgagaagt actagaggat cataatcagc cataccacat ttgtagaggt tttacttgct 4320  
 ttaaaaaacc tcccacacct cccctgaac ctgaaacata aatgaatgc aattgttggt 4380  
 gttaacttgt ttattgcagc ttataatggg tacaataaa gcaatagcat cacaaatttc 4440  
 acaataaaag catTTTTTt actgcattct agttgtgggt tgtccaaact catcaatgta 4500  
 tcttatcatg tctggatctg atcactgctt gaccctagga gatccgaacc agataagtga 4560  
 aatctagttc caaactatTT tgtcattttt aattttcgta ttagcttacg acgtacacc 4620  
 cagttcccat ctattttgtc actettccct aaataatcct taaaaactcc atttccaccc 4680  
 ctcccagttc ccaactatTT tgtccgccc cagcggggca tttttcttcc tgttatgttt 4740  
 ttaatcaaac atcctgccc ctccatgtga caaacgctc tcttcggcta cttttctct 4800  
 gtcacagaat gaaaattttt ctgtcatctc ttcgttatta atgtttgtaa ttgactgaat 4860  
 atcaacgctt atttgcagcc tgaatggcga atgggacgcg ccctgtagcg gcgcattaag 4920  
 cgcgcggggt gtggtgggta cgcgcgagct gaccgctaca cttgccagcg ccctagcgcc 4980  
 cgctccttcc gctttcttcc ctccctttct cgccacgttc gccggttcc cccgtcaagc 5040  
 tctaaatcgg gggctccctt tagggttccg atttagtget ttacggcacc tcgaccccaa 5100  
 aaaacttgat tagggtgatg gttcacgtag tgggcatcg ccctgataga cggtttttcg 5160  
 ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac 5220  
 actcaaccct atctcggctt attcttttga tttataaggg attttgccga tttcggccta 5280  
 ttggttaaaa aatgagctga tttacaaaa atttaacgcg aattttaaca aaatattaac 5340  
 gtttacaatt tcaggtggca cttttcgggg aaatgtgcgc ggaacccta tttgtttatt 5400  
 tttctaaata cattcaata tgtatccgct catgagacaa taaccctgat aaatgcttca 5460  
 ataatatgta aaaaggaaga gtatgagtat tcaacatttc cgtgtcgcgc ttattccctt 5520  
 ttttgcggca ttttgccttc ctgtttttgc tcaccagaa acgctgggtga aagtaaaaga 5580  
 tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa 5640  
 gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct 5700  
 gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg gtcgccgat 5760  
 acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga 5820  
 tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata acactgcggc 5880  
 caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat 5940  
 gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa 6000  
 cgacgagcgt gacaccacga tgctgtagc aatggcaaca acgttgcgca aactattaac 6060  
 tggcgaacta ctactctag cttcccggca acaattaata gactggatgg aggcggataa 6120  
 agttgcagga ccacttctgc gtcggccct tccggtggc tggtttattg ctgataaatc 6180  
 tggagccggt gagcgtgggt ctgcggtat cattgcagca ctggggccag atggtaagcc 6240  
 ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag 6300

acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag accaagttta 6360  
 ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa 6420  
 gatccttttt gataatctca tgacaaaaat cccttaacgt gagttttcgt tccactgagc 6480  
 gtcagaccgc ggggcatgac taacatgaga attacaactt atatcgatat gggctgactt 6540  
 caggtgctac atttgaagag ataaattgca ctgaaatcta gaaatatatt atctgattaa 6600  
 taagatgata ttcttgagat cgttttggtc tgcgcgtaat ctcttgctct gaaaacggaa 6660  
 aaaaccgcct tgcagggcgg tttttcgaag gttctctgag ctaccaactc tttgaaccga 6720  
 ggtaactggc ttggaggagc gcagtcacca aaacttgctc tttcagttta gccttaaccg 6780  
 gcgcatgact tcaagactaa ctctctataa tcaattacca gtggctgctg ccagtgggtg 6840  
 ttttgcatgt ctttccgggt tggactcaag acgatatgta cgggataagg cgcagcggtc 6900  
 ggactgaacg gggggttcgt gcatacagtc cagcttggag cgaactgcct acccggaact 6960  
 gagtgtcagg cgtggaatga gacaaacgcg gccataacag cggaatgaca ccggtaaacc 7020  
 gaaaggcagg aacaggagag cgcacgaggg agccgccagg ggaaacgcct ggtatcttta 7080  
 tagtcctgtc gggtttcgcc accactgatt tgagcgtcag atttcgtgat gcttgtcagg 7140  
 ggggcggagc ctatggaaaa acggctttgc cgcggccctc tcaattccct gttaagtatc 7200  
 ttcttggeat cttccaggaa atctccgcc cgttcgtaag ccatttccgc tcgccgcagt 7260  
 cgaacgaccg agcgtagcga gtcagtgagc gaggaagcgg aatatatcct gtatcacata 7320  
 ttctgctgac gcaccggtgc agcctttttt ctctgccac atgaagcact tcaatgacac 7380  
 cctcatcagt gccaacatag taagccagta taaactccgc tagcgtgat gtccggcgggt 7440  
 gcttttgccg ttacgcacca ccccgctagt agctgaacag gagggacagc tgatagaaac 7500  
 agaagccagt tctttcctgc gttatccctt gattctgtgg ataaccgtat taccgccttt 7560  
 gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag 7620  
 gaagcggaag agcgctgat gcggtatatt ctcttacgc atctgtgcgg tatttcacac 7680  
 cgcatagacc agccgcgtaa cctggcaaaa tcggttacgg ttgagtaata aatg 7734

<210> 5

<211> 4915

<212> DNA

<213> 人工序列

<220>

<223>

<400> 5

cgcgccggta tgtacaggaa gaggtttata ctaaactggt acattgcaaa cgtggtttcg 60  
 tgtgccaagt gtgaaaaccg atgtttaatc aaggtcttga cgcatttcta caaccacgac 120  
 tccaagtgtg tgggtgaagt cagatgttta aacctatgtg cctggcagat aacttcgtat 180  
 aatgtatgct atacgaagtt atggtacgcg gccgcgtaga ggatctgttg atcagcagtt 240  
 caacctgttg ataatacggc cctttaattc aaccaacac aatatattat agttaataaa 300  
 gaattattat caaatcattt gtatattaat taaaatacta tactgtaaat tacattttat 360  
 ttacaatcac tcgacaccgg tgatatccat atggcacatc accaccatca tcaccatcac 420  
 caccacggag gcagcgatga cgatgataaa accatgggat ccgctagctt aagctgtcag 480

accaagttta ctcatatata ctttagattg atttaaaact tcatttttta tttaaaagga 540  
 tctaggtgaa gatccttttt gataatctca tgaccaaagt cccttaacgt gagttttcgt 600  
 tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cttttttttc 660  
 tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttgc 720  
 cggatcaaga gctaccaact ctttttcgga aggttaactgg cttcagcaga gcgcagatac 780  
 caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac 840  
 cgcctacata cctcgtcttg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 900  
 cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgag cggtcgggct 960  
 gaacggggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 1020  
 acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaaag gcggacaggt 1080  
 atccggttaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg 1140  
 cctggtatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt cgatttttct 1200  
 gatgctctgc aggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt 1260  
 tcttgccctt ttgctggcct tttgtcaca tgttctttcc tgcgttatcc cctgattctg 1320  
 tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg 1380  
 agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc 1440  
 ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccagc tggaagcgg 1500  
 gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac 1560  
 actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa tttcacacag 1620  
 gaaacagcta tgacatgat tacgccaagc tatttaggtg acgcgttaga atactcaagc 1680  
 tatgcatcat ctttggttcc gtcacggac ccattagtaa cggccgccag tgtgctggag 1740  
 ttttgtagat acccatcaca ctggcgtcca ctggaacatg caagtagagg gcccattcgc 1800  
 ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg tgactgggaa 1860  
 aaccctggcg ttaccaact taatgcctt gcagcacatc cccctttcgc cagctggcgt 1920  
 aatagcgaag agggccgcac cgatcgccct tccaacagt tgccgcagcct atacgtacgg 1980  
 taactgacta agaattccga ttacaaagac gatgacgaca agggctcgag tgaattttg 2040  
 tattttcaaa gcgggagctc tgtgagcaag ggcgaggagc tgttcaccgg ggtggtgccc 2100  
 atcctggctg agctggacgg cgacgtaaac ggccacaagt tcagcgtgct cggcgagggc 2160  
 gagggcgatg ccacctacgg caagctgacc ctgaagtcca tctgcaccac cggcaagctg 2220  
 cccgtgccct ggcccacct cgtgaccacc ctgacctag cgtgacgtg cttcagccgc 2280  
 taccgacgacc acatgaagca gcacgaactt ttcaagtccg ccatgcccga aggctacgtc 2340  
 caggagcgca ccatcttctt caaggacgac ggcaactaca agaccgcgc cgaggtgaag 2400  
 ttcgagggcg acacctggt gaaccgcac gagctgaagg gcatcgactt caaggaggac 2460  
 ggcaacatcc tggggcaca gctggagtac aactacaaca gccacaacgt ctatatcatg 2520  
 gccgacaagc agaagaacgg catcaaggtg aacttcaaga tccgccaca catcgaggac 2580  
 ggacgcgtgc agctcgccga ccactaccag cagaacaccc ccatcggcga cggccccgtg 2640  
 ctgctgcccg acaaccacta cctgagcacc cagtccgcc tgagcaaaga cccaacgag 2700  
 aagcgcgatc acatggtcct gctggagttc gtgaccgcc cggggtacac tctcggtcatg 2760  
 gacgagctgt acaagtaact gactaaggtc ccgaaagga agctgagttg gctgctgcca 2820

```

ccgctgagca ataactagca taacccttg gggcctctaa acgggtcttg aggggttttt 2880
tgctgaaagg aggaactatc ctcaaggagg gatgggggag gctaactgaa acacggaagg 2940
agacaatacc ggaaggaacc cgcgctatga cggcaataaa aagacagaat aaaacgcacg 3000
gggtgttggt cgtttgttca taaacgcggg gtctgggtccc agggctggca ctctgtcgat 3060
acccccaccga gaccccatg ggaccaatac gcccgcttt cttccttttc cccaccccaa 3120
cccccaagtt cgggtgaagg ccagggctc gcagccaacg tcggggcggc aagccctgcc 3180
atagccacta cgggtacgtc tgaaagcatg ctttttggga atttacgtac taagctctca 3240
tgtttcacgt actaagctct catgtttaac gtactaagct ctcatgttta acgaactaaa 3300
ccctcatggc taacgtacta agctctcatg gctaactgac taagctctca tgtttcacgt 3360
actaagctct catgtttgaa caataaaatt aatataaatc agcaacttaa atagcctcta 3420
aggttttaag tttataaga aaaaaagaa tatataaggc ttttaaagct ttttaagttt 3480
aacggtgtg gacaacaagc cagggatgta acgactgag aagcccttag agcctctcaa 3540
agcaattttc agtgacacag gaacacttaa cggctgacag aattagcttc acgtgccgc 3600
aagcactcag ggcgcaagg ctgctaaagg aagcgaaca cgtagaaagc cagtccgcag 3660
aaacggtgct gaccccgat gaatgtcagc tactgggcta tctggacaag ggaaaacgca 3720
agcgcaaaga gaaagcaggt agcttgcagt gggcttacat ggcatagct agactgggcg 3780
gttttatgga cagcaagcga accggaattg ccagctgggg cgccctctgg taaggttggg 3840
aagccctgca aagtaaacg gatggcttcc ttgccgcaa ggatctgatg gcgcagggga 3900
tcaagatctg atcaagagac aggatgagga tcgtttcgca tgattgaaca agatggattg 3960
cacgcagggt ctccggccgc ttgggtggag aggtattcg gctatgactg ggcacaacag 4020
acaatcggct gctctgatgc cgccgtgttc cggctgtcag cgcaggggcg cccggttctt 4080
tttgtcaaga ccgacctgtc cgggtgccctg aatgaactgc aggacgaggc agcgcggtta 4140
tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg 4200
ggaagggact ggctgctatt gggcgaagtg ccggggcagg atctcctgtc atctcacctt 4260
gtcctgccg agaaagtatc catcatggct gatgcaatgc ggcggctgca tacgttgat 4320
ccggctacct gccattcga ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg 4380
atggaagccg gtcttgcga tcaggatgat ctggacgaag agcatcaggg gctcgcgcca 4440
gccgaactgt tcgccaggt caaggcgcgc atgcccagc gcgaggatct cgtcgtgaca 4500
catggcgatg cctgcttgc gaatatcatg gtggaaaatg gccgcttttc tggattcatc 4560
gactgtggcc ggctgggtgt ggcggaccgc tatcaggaca tagcgttggc taccgtgat 4620
attgctgaag agcttggcgg cgaatgggct gaccgttcc tcgtgcttta cggtatcgcc 4680
gtccccgatt cgcagcgcat cgcttctat cgcttcttg acgattctt ctgagcggga 4740
ctctgggggt cgaaatgacc gaccaagcga cgcccaacct gccatcacga gatttcgatt 4800
ccaccgccgc cttctatgaa aggttgggct tcggaatcgt ttccgggac gccggtgga 4860
tgatcctcca gcgcggggat ctcatgtgtg agttcttcgc ccacatcata tcgat 4915

```

<210> 6

<211> 4915

<212> DNA

<213> 人工序列

&lt;220&gt;

&lt;223&gt;

&lt;400&gt; 6

```

cgcgccggta tgtacaggaa gaggtttata ctaaactgtt acattgcaaa cgtgggtttcg 60
tgtgccaagt gtgaaaaccg atgtttaatc aaggctctga cgcatttcta caaccacgac 120
tccaagtgtg tgggtgaagt cagatgttta aacctatgtg cctggcagat aacttcgtat 180
aatgtatgct atacgaagtt atggtacgcg gccgcgtaga ggatctgttg atcagcagtt 240
caacctgttg ataatacggg cttttaattc aaccaacac aatatattat agttaaataa 300
gaattattat caaatcattt gtatattaat taaaatacta tactgtaaat tacattttat 360
ttacaatcac tcgacaccgg tgatatccat atggcacatc accaccatca tcaccatcac 420
caccacggag gcagcgatga cgatgataaa accatgggat ccgctagctt aagctgtcag 480
accaagttaa ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga 540
tctaggtgaa gatccttttt gataatctca tgacaaaat cccttaacgt gagttttcgt 600
tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat ctttttttcc 660
tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttgc 720
cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac 780
caaatactgt tcttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac 840
cgcttacata cctcgtctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 900
cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgcag cggtcgggct 960
gaacggggggg ttctgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 1020
acctacagcg tgagctatga gaaagcgcca cgcttccga agggagaaag gcggacaggt 1080
atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg 1140
cctggtatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt cgatttttgt 1200
gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc tttttacggt 1260
tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg 1320
tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg 1380
agcgcagcga gtcagtgage gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc 1440
ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg 1500
gcagtgagec caacgcaatt aatgtgagtt agctactca ttaggcaccc caggctttac 1560
actttatgct tccggtcgt atgttggtg gaattgtgag cggataacaa tttcacacag 1620
gaaacagcta tgaccatgat tacgccaagc tatttaggtg acgcgttaga atactcaagc 1680
tatgcatcat ctttggttcc gtcacgagc ccattagtaa cggccgccag tgtgctggag 1740
ttttgtagat acccatcaca ctggcgctca ctggaacatg caagtagagg gcccatttcg 1800
ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg tgactgggaa 1860
aaccctggcg ttacccaact taatgcctt gcagcacatc cccctttcgc cagctggcgt 1920
aatagcgaag agggccgcac cgatcgccct tccaacagt tgcgcagcct atacgtacgg 1980
taactgacta agaattccga gcaaaagttg attagcgaag aagacttagg ctcgagtga 2040
aatgtgtatt ttcaaagcgg gagctctgtg tctaaggcgc aagagctgat taaggagaac 2100
atgcacatga agctgtacat ggagggcacc gtgaacaacc accattcaa gtgcacatcc 2160

```

gagggcgaag gcaagcccta cgagggcacc cagacatga gaatcaaggt ggtcgagggc 2220  
 ggccctctcc ctttcgcctt cgacatcctg gctaccagct tcatgtacgg cagcagaacc 2280  
 ttcatcaacc acaccaggg catccccgac ttctttaagc agtccttccc tgagggttc 2340  
 acatgggaga gagtcaccac atacgaagac gggggcgtgc tgaccgctac ccaggacacc 2400  
 agcctccagg acggctgcct catctacaac gtcaagatca gaggggtgaa cttcccatcc 2460  
 aacggccctg tgatgcagaa gaaaacactc ggctgggagg ccaacaccga gatgctgtac 2520  
 cccgctgacg gcggcctgga aggcagaagc gacatggccc tgaagctcgt gggcgggggc 2580  
 cacctgatct gcaacttcaa gaccacatac agatccaaga aaccgctaa gaacctcaag 2640  
 atgcccggcg tctactatgt ggaccacaga ctggaaagaa tcaaggaggc cgacaaagag 2700  
 acctacgtcg agcagcacga ggtggctgtg gccagatact gcgacctccc tagcaaactg 2760  
 gggcacaaac ttaattaact gactaaggta cccgaaagga agctgagttg gctgctgcca 2820  
 ccgctgagca ataactagca taacccttg gggcctctaa acgggtcttg aggggttttt 2880  
 tgctgaaagg aggaactatc ctcaggggga gatgggggag gctaactgaa acacggaagg 2940  
 agacaatacc ggaaggaacc cgcgctatga cggcaataaa aagacagaat aaaacgcacg 3000  
 ggtgttgggt cgtttgttca taaacgcggg gtccggctcc agggctggca ctctgtcgat 3060  
 accccaccga gacccattg ggaccaatac gcccgcttt cttccttttc cccaccccaa 3120  
 ccccaagtt cgggtgaagg cccagggtc gcagccaacg tcggggcggc aagccctgcc 3180  
 atagccacta cgggtacgtc tgaaagcatg ctttttggga atttacgtac taagctctca 3240  
 tgtttcacgt actaagctct catgtttaac gtactaagct ctcatgttta acgaactaaa 3300  
 cctcatggc taacgtacta agctctcatg gtaacgtac taagctctca tgtttcacgt 3360  
 actaagctct catgtttgaa caataaaatt aatataaatc agcaacttaa atagcctcta 3420  
 aggttttaag tttataaga aaaaaagaa tatataaggc ttttaaagct ttttaaggttt 3480  
 aacggttggt gacaacaagc cagggatgta acgcaactag aagcccttag agcctctcaa 3540  
 agcaattttc agtgacacag gaacacttaa cggctgacag aattagcttc acgctgccgc 3600  
 aagcactcag ggcgcaaggg ctgctaaagg aagcgggaaca cgtagaaagc cagtccgcag 3660  
 aaacggtgct gaccccggt gatgtcagc tactgggcta tctggacaag ggaaaacgca 3720  
 agcgcaaaga gaaagcaggt agcttgcatg gggttacat ggcatagct agactgggcg 3780  
 gttttatgga cagcaagcga accggaattg ccagctgggg cgccctctgg taaggttggg 3840  
 aagccctgca aagtaaaact gatggctttc ttgccgcaa ggatctgat gcgcagggga 3900  
 tcaagatctg atcaagagac aggatgagga tcgtttcgca tgattgaaca agatggattg 3960  
 cacgcaggtt ctccggccgc ttgggtggag aggtattcg gctatgactg ggcacaacag 4020  
 acaatcggct gctctgatgc cgccgtgttc cggctgtcag cgcaggggcg cccggttctt 4080  
 tttgtcaaga ccgacctgtc cgtgccttg aatgaactgc aggacgagc agcgcggcta 4140  
 tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg 4200  
 ggaagggact ggctgctatt gggcgaagt cggggcagg atctcctgtc atctcactt 4260  
 gctcctgccg agaaagtatc catcatggct gatgcaatgc ggcggctgca tacgttgat 4320  
 ccggtacct gccattcga ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg 4380  
 atggaagccg gtcttgcga tcaggatgat ctggacgaag agcatcagg gctcgcgcca 4440  
 gccgaactgt tcgccaggct caaggcgcgc atgcccagc gcgaggatct cgtcgtgaca 4500

catggcgatg cctgcttgcc gaatatcatg gtggaaaatg gccgcttttc tggattcatc 4560  
 gactgtggcc ggctgggtgt ggccggaccgc tatcaggaca tagcgttggc taccctgat 4620  
 attgctgaag agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta cggatatgcc 4680  
 gctcccgatt cgcagcgcac cgccttctat cgccttcttg acgagttctt ctgagcggga 4740  
 ctctgggggtt cgaaatgacc gaccaagcga cgcccaacct gccatcacga gatttcgatt 4800  
 ccaccgccgc cttctatgaa aggttgggct tcggaatcgt tttccgggac gccggctgga 4860  
 tgatcctcca gcgcggggat ctcatgctgg agttcttcgc ccacatcata tcgat 4915

<210> 7

<211> 9846

<212> DNA

<213> 人工序列

<220>

<223>

<400> 7

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtcaaaa ttgcccgtcg tattaaagag gggcgtggcc aagggcattg aaagactata 240  
 ttgcggcgtg tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatacaag tgcatactt cttcccgat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccc taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcggttg tgctgcgtaa 840  
 catcggttgc gtcaccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccagag catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatga cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagt ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgcgg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaacccca caccaacaac 1440  
 ggcccctcga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500

aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560  
 gatcccgccg gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttgggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggccagcagca cggggccgtc 1680  
 gccgatgggg gtgttctgct ggtagtggtc ggccagctgc acgctgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac 1800  
 gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860  
 gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctcggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gacttgaaga agtcgtgctg cttcatgttg tcggggtagc ggctgaagca 2040  
 ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggccagggca gcttgccggt 2100  
 ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctgc cctcgccgga 2160  
 cacgtgaac ttgtggccgt ttacgtgcc gtccagctcg accaggatgg gcaccacccc 2220  
 ggtgaacagc tcctcgccct tgetcatccc ggtgactgg aagtacaggt tttcttgcca 2280  
 gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400  
 actaaccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaataa ttgttttgt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520  
 atctctagtt gatactaatg gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580  
 ccagaatatg ccatcagatg aagggaatgt gcaagaagt tctgacacca tgctagacat 2640  
 gctcttagtt tgggaagtgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700  
 aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760  
 cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgt ttgtaatgat 2880  
 gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaacaa agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttgggtg gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttta acactagttc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaaat acagctacat aaattacaca atttaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccacagcaat 3540  
 caagcatgca ttacaaagag acatttttac accaaatgat gaacgcctgc tgagcattgt 3600  
 gaatgtctgc aaagcaggaa aaaagaaaaa gaactgtttt ttatgtgcca cagtgacaac 3660  
 tgaacgcct gtgcaggta agtggttcaa agtcaagaaa tccgataagg gagatttcta 3720  
 caaaaggcag attgcatggg cccttcgaga tcttgtctgt gtagatgcca aagatgctat 3780  
 caaagaaaat cctgaatttg atttacactt tgaaaaaata tataaatggg ttgccagcag 3840

cactgctgaa aagaatgcat ttatttcatg catttggaaa ttgaatcagc gatatctccg 3900  
 gaagaaaatt gattttgtca atgttagctc acagcttttg gaagaactgc ctaaagttac 3960  
 agaagaatct gttccaagtg gagaaaatca gagtgtgaca ggaggtgatg aagaagtagt 4020  
 agatgaatac caagagttaa atgcaagaga agaacaggat atcgaaataa tgatggaagg 4080  
 ctgtgaatat gcaatctcga atgcggaagc ctttgcagaa aaattgtcca gagagctgca 4140  
 ggtgctagat ggggctaaca tccagtcaat catggcatct gaaaaacaag tcaacatcct 4200  
 gatgaaattg ctagatgagg ctctaaagga ggtagatcag attgaattga aactgagcag 4260  
 ttatgaggaa atgctccaaa gtgtaaaaga acaaatggat cagatctctg aaagcaacca 4320  
 cctaattcat cttagtaaca ctaataatgt aaaactccta tctgagatag agttccttgt 4380  
 gaaccacatg gacttggcca aaggtcatat aaaggccctt caggaaggag atcttgcttc 4440  
 ttccagaggc attgaggcct gcaccaatgc tgctgatgcc cttctgcagt gcatgaatgt 4500  
 agctcttcga ccaggccatg acttgettct ggcagtcaaa cagcaacagc agcgattcag 4560  
 tgatttgca gagctttttg cccggagact ggccagtcac ctcaacaatg tttttgttca 4620  
 acagggtcat gatcagagtt cgactcttgc ccaacactct gttgaactga ctttacccaa 4680  
 tcatcatcca ttcatagag atttgctccg atatgccaa ctgatggagt ggctaaagag 4740  
 tacagattat ggaaaaatg aaggactaac aaagaattac atggattatt tatcccgact 4800  
 atatgaaaga gaaatcaaag atttctttga agttgcaaag atcaagatga ctggcacaac 4860  
 taaagaaagc aagaagtttg ctacactgcc tcgaaaagaa agtgctgtca aacaggaaac 4920  
 agagagtctt catggaagtt cggggaaatt aactggatct acttctagtc taaataagct 4980  
 cagtgttcag agttcaggga atcgcagatc tcagtcatct tccctgttgg atatgggaaa 5040  
 catgtctgcc tctgatctcg atgttgctga caggaccaa tttgataaga tctttgaaca 5100  
 ggtactaagt gaactggagc ccctatgtct ggcagaacag gacttcataa gtaaattttt 5160  
 caaactacag caacatcaaa gtatgcctgg aactatggct gaagcagagg acctggatgg 5220  
 aggaacatta tcacggcaac ataattgttg cacaccactg cctgtttcat ctgagaaaga 5280  
 tatgatccgc caaatgatga ttaaaatatt tcgctgcatt gagccagagc tgaacaacct 5340  
 aattgcatta ggagacaaaa ttgatagctt taactctctt tatatgttag tcaaaatgag 5400  
 tcatcatgtg tggactgcac aaaatgtgga ccctgettct ttcctaagta ctacattggg 5460  
 aaatgttttg gtgactgtca aaaggaactt tgacaaatgc attagtaacc aaataaggca 5520  
 aatggaagaa gtaaagatct caaaaaagag taaagttgga attcttccat ttgttgctga 5580  
 atttgaagaa tttgctggac ttgcagaatc aatcttcaaa aatgctgagc gtcgtggaga 5640  
 cctggataaa gcatacacca aacttatcag aggagtatct gttaatgttg agaaagtagc 5700  
 aaatgaaagc cagaagacce ccagggatgt ggttatgatg gaaaactttc accatatctt 5760  
 tgcaactctt tctcgattga aaatctcatg tctagaagca gaaaaaaaag aagccaaaca 5820  
 aaaatacaca gatcaccttc agtcttatgt catttactct ttaggacaac ctcttgaaaa 5880  
 actaaatcat ttctttgaag gtgttgaagc tcgcgtggca cagggcataa gggaggagga 5940  
 agtaagtac caacttgcac ttaacaaaca agaacttcgt aaagtcatta aggagtaccc 6000  
 tggaaaggaa gtaaaaaaag gtctagataa cctctacaag aaagttgata aacatttatg 6060  
 tgaagaagag aacttacttc aggtggtgtg gcactccatg caagatgaat ttatacgcca 6120  
 gtataagcac tttgaagggt tgatagctcg ctgttatcct ggatctggtg ttacaatgga 6180

attcactatt caggacattc tggattattg ttccagcatt gcacagtccc acggctcgag 6240  
 ttaactgact aaaagcttcg aaaggaagct gagttggctg ctgccaccgc tgagcaataa 6300  
 ctagcataac cccttggggc ctctaaacgg gtcttgaggg gttttttgct gaaaggagga 6360  
 actatcctca gggtcgagaa gtactagagg atcataatca gccataccac atttgtagag 6420  
 gttttacttg ctttaaaaaa cctcccacac ctccccctga acctgaaaca taaaatgaat 6480  
 gcaattgttg ttgttaactt gtttattgca gcttataatg gttacaaata aagcaatagc 6540  
 atcacaaaatt tcacaaataa agcatttttt tcaactgcatt ctagttgttg tttgtccaaa 6600  
 ctcatcaatg tatcttatca tgtctggatc tgatcactgc ttgagcctag gagatccgaa 6660  
 ccagataagt gaaatctagt tccaaactat tttgtcattt ttaattttcg tattagctta 6720  
 cgacgctaca cccagttccc atctattttg tcaactcttc ctaaataatc cttaaaaact 6780  
 ccatttccac ccctcccagt tcccaactat tttgtccgcc cacagcgggg catttttctt 6840  
 cctgttatgt ttttaataca acatcctgcc aactccatgt gacaaaccgt catcttcggc 6900  
 tactttttct ctgtcacaga atgaaaattt ttctgtcate tcttcgttat taatgtttgt 6960  
 aattgactga atatcaacgc ttatttgcag cctgaatggc gaatgggacg cgccctgtag 7020  
 cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag 7080  
 cgccctagcg cccgctcctt tcgctttctt cccttccttt ctgccacgt tcgccggtt 7140  
 tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca 7200  
 cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggcat cgccctgata 7260  
 gacggttttt cgccctttga cgttgagtc cacgtttttt aatagtggac tcttgttcca 7320  
 aactggaaca acactcaacc ctatctcggc ctattctttt gatttataag ggattttgcc 7380  
 gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattttta 7440  
 caaaatatta acgtttacaa tttcaggttg cacttttcgg ggaaatgtgc gcggaacccc 7500  
 tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg 7560  
 ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc 7620  
 ccttattccc tttttgcgg cattttgcct tcctgttttt gctcaccag aaacgctggt 7680  
 gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct 7740  
 caacagcggc aagatccttg agagttttcg cccgaagaa cgttttccaa tgatgagcac 7800  
 ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc aagagcaact 7860  
 cggtcgccgc atacactatt ctcaaatga cttggttgag tactcaccag tcacagaaaa 7920  
 gcattcttac gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga 7980  
 taacactgcg gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt 8040  
 tttgcacaac atgggggatc atgtaactcg cttgatcgt tgggaaccgg agctgaatga 8100  
 agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg 8160  
 caaactatta actggcgaac tacttactct agcttcccgg caacaattaa tagactggat 8220  
 ggaggcggat aaagttgcag gaccacttct gcgctcgccc cttccggctg gctggtttat 8280  
 tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc 8340  
 agatggtaag ccctcccgtg tcgtagttat ctacacgacg gggagtcagg cactatgga 8400  
 tgaacgaaat agacagatcg ctgagatagg tgccctactg attaagcatt ggtaactgtc 8460  
 agaccaagtt tactcatata tacttttagat tgatttaaaa cttcattttt aatttaaaag 8520

gatctagggtg aagatccttt ttgataatct catgaccaa atcccttaac gtgagttttc 8580  
 gttccactga gcgtcagacc gcggggcatg actaacatga gaattacaac ttatatcgta 8640  
 tggggctgac ttcagggtgct acatttgaag agataaattg cactgaaatc tagaaatatt 8700  
 ttatctgatt aataagatga tcttcttgag atcgttttgg tctgcgcgta atctcttgct 8760  
 ctgaaaacgg aaaaaaccgc cttgcagggc ggtttttcga aggttctctg agctaccaac 8820  
 tctttgaacc gaggttaactg gcttggagga gcgcagtcac caaaacttgt cttttcagtt 8880  
 tagccctaac cggcgcgcatga cttcaagact aactcctcta aatcaattac cagtggctgc 8940  
 tgccagtggt gcttttgcgcat gtctttccgg gttggactca agacgatagt taccggataa 9000  
 ggcgcagcgg tcggactgaa cgggggggttc gtgcatacag tccagcttgg agcgaactgc 9060  
 ctacccgga ctgagtgtca ggcgtggaat gagacaaacg cggccataac agcggaatga 9120  
 caccggtaaa ccgaaaggca ggaacaggag agcgcacgag ggagccgcca ggggaaacgc 9180  
 ctggtatctt tatagtcttg tcgggtttcg ccaccactga tttgagcgtc agatttcgtg 9240  
 atgcttgta gggggggcgga gcctatggaa aaacggcttt gccgcggccc tctcacttcc 9300  
 ctgttaagta tcttcctggc atcttcagg aaatctccgc cccgttcgta agccatttcc 9360  
 gctcgccgca gtcgaacgac cgagcgtagc gagtcagtga gcgaggaagc ggaatatatc 9420  
 ctgtatcaca tattctgctg acgcaccggt gcagcctttt ttctcctgcc acatgaagca 9480  
 cttcactgac accctcatca gtgccaacat agtaagccag tatacactcc gctagcgtg 9540  
 atgtccggcg gtgcttttgc cgttacgcac caccctcgta gtagctgaac aggagggaca 9600  
 gctgatagaa acagaagcca gttctttcct gcgttatccc ctgattctgt ggataaccgt 9660  
 attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag 9720  
 tcagtgagcg aggaagcgga agagcgctg atgcggtatt ttctccttac gcatctgtgc 9780  
 ggtatttcac accgcataga ccagccgctg aacctggcaa aatcggttac ggttgagtaa 9840  
 taaatg 9846

<210> 8

<211> 9933

<212> DNA

<213> 人工序列

<220>

<223>

<400> 8

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcatTTTTT 180  
 gaagtcaaaa ttgcccgtcg tattaagag gggcgtggcc aagggcagtg aaagactata 240  
 ttgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatactt ctccccgtat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctgc cgagactgc gagatcatag atatagatct cactacgcgg 540

ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggg cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcggtgc tgctgcgtaa 840  
 catcggttgc gtccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccaggc catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgtc 1020  
 acttgcatta cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagttttaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgagg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440  
 ggcccctcga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560  
 gatcccgagg gcggtcacga actccagcag gaccatgtga tcgcgttct cgttggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggttgctg ggcagcagca cggggccgct 1680  
 gccgatgggg gtgttctgct ggtagtggtc ggcagctgc acgctgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcgcca tgatatagac 1800  
 gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860  
 gtcgatgcc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gactgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040  
 ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt 2100  
 ggtgcagatg aacttcaggg tcagcttgc gtaggtggca tcgccctgc cctcgccgga 2160  
 cacgtgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220  
 ggtgaacagc tcctcgccct tgctcatccc gggtgactgg aagtacaggt tttcttgcca 2280  
 gtacaccaat tcattcatga gttgagtcgc ttcttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400  
 actaaccac tgctgcgct cctgatttgc caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaataa ttgtttgtgt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520  
 atctctagtt gatactaatg gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580  
 ccagaatatg ccatcagatg aagggaatgt gcaagaagtg tctgacacca tgctagacat 2640  
 gctcttagtt tggaagtggg ttgtcacaag acatattctc tcttcccttt gtggtctct 2700  
 aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760  
 cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgt ttgtaatgat 2880

gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaaaca agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggcctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttaa acactagtcc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540  
 cgcgcccgct agccgatcac gacaaccccc ctttgtgacc ggcatctctc caaatgaagg 3600  
 gataccatgg acgaaggtca caatcagggg agaaaatctg gggactggcc ccaccgacct 3660  
 cataggcttg accatttgtg gacataattg cctcctgacg gcagaatgga tgtctgcaag 3720  
 taaaatagta tgtcgagtgg gacaagccaa aatgacaaa ggagacatta ttgtcaccac 3780  
 taagtcaggt ggcagaggaa cctcaacagt ctctttcaag ctactcaaac ctgagaaaat 3840  
 aggcattttg gatcagtcgt ctgtgtgggt tgatgaaatg aattattatg atatgcgcac 3900  
 tgacaggaac aaaggaattc cgcccttgct cttacgtcct gctaaccgcg ttggcattga 3960  
 gattgaaaaa agtaaatctt cgcaagga cttagaaatg ctattccatg gaatgagtgc 4020  
 tgattttaca agtgagaatt tctcagcagc ctggtatctt atagagaatc actcaaacac 4080  
 cagttttgag cagctcaaaa tggcagtcac caacctaaag agacaggcta acaagaagag 4140  
 tgagggcagc ctggcctatg tgaaaggcgg tctcagtaca ttcttcgaag cacaggatgc 4200  
 cctctcagcc atccatcaaa aactagaagc agatggaacg gaaaaagtag aaggatccat 4260  
 gacgcagaaa ctggagaatg ttctgaacag agcaagtaat actgcagaca cattgtttca 4320  
 agaagtatta ggtcggaaag acaaggcaga ttccactaga aatgcactca atgtgcttca 4380  
 gcgatttaag tttcttttca accttcctct aaatattgaa aggaatattc aaaagggtga 4440  
 ttatgatgtg gttattaatg attatgaaaa ggccaagtca ctttttggga aaacggaggt 4500  
 gcaagtttcc aagaaatatt atgtgaagt agaaacaagg attgaagctt taagagaatt 4560  
 acttctggat aaattgcttg agacaccatc aactttacat gacaaaaaac gttacataag 4620  
 gtacctgtct gaccttcatt cgtctggtga cctgcttgg caatgcattg gagcccaaca 4680  
 caagtggatc cttcagctca tgcacagttg caaagagggc tacgtgaaag atctgaaagg 4740  
 taaccaggc ctgcacagtc ccatgttgga tcttgataat gatacacgtc cctcagtggt 4800  
 gggccatctc agtcagacag cgtccctgaa gaggggcagc agctttcagt ctggtcgaga 4860  
 cgacacgtgg agatacaaaa ctccccacag ggtggccttt gttgaaaaat tgacaaaact 4920  
 cgtcttgagc cagctgccta acttctggaa actctggatc tcctacgtta atggaagcct 4980  
 cttcagtgag actgctgaga agtcaggcca gattgaaaga tcaaagaatg taaggcaaag 5040  
 acaaaatgat ttttaagaaaa tgattcagga agtaatgcac tccttggtga agcttaccgg 5100  
 cggagccctg cttccctca gcatccggga tggggaagcc aagcagtacg gaggctggga 5160  
 ggtgaagtgc gagctctccg gacagtggct cgctcacgcc atccagactg taagacttac 5220

tcacgaatcg ttgactgccc ttgaaattcc taatgacctg ttacagacta tccaggatct 5280  
 catcttggat ctccgagtac gttgcgtaat ggccacgttg cagcacacgg cggaagaaat 5340  
 aaagagatta gctgaaaaag aagactggat tgttgacaat gaaggactga cttctctacc 5400  
 atgtcagttt gaacagtgcg tcgtgtgttc tctgcagtca ctgaaggggg ttctggagtg 5460  
 caagccggga gaggccagtg tcttccaaca acctaaaaca caggaggagg ttgcccagct 5520  
 aagcatcaat ataatgcagg tttttatata ctgtctggaa cagttgagca ccaagcctga 5580  
 tgcagatata gatactacac atctctctgt tgatgtttct tcccctgact tgtttggaag 5640  
 tatccatgaa gacttcagct tgacctcaga acagcgcctt ttgatagtcc taagtaattg 5700  
 ctgctatcta gaacgtcaca ccttccataa tatcgcagaa ctttttgaag agcacaactt 5760  
 ccagggaata gaaaaaatca cacaggttag catggcctca ttgaaagaac tagatcaaag 5820  
 actctttgaa aattacatcg agttgaaagc agatcccatc gttggctcct tagaacctgg 5880  
 aatttatgca ggatattttg attggaagga ctgcctgcct ccaacagggtg tcagaaacta 5940  
 tttaaaagaa gcactggtga atataattgc cgtgcatgca gaggtgttca ccatttccaa 6000  
 agaactggtc cctcgggtac tatccaaggt gatagaagca gtttctgaag agctcagtcg 6060  
 actgatgcag tgtgtttcat ccttcagcaa aaatggagct ttacaggcga gacttgaaat 6120  
 ctgtgctttg aggacactg tggtgtttta cctgacaccc gaaagcaagt caagttttaa 6180  
 gcaggctttg gaagccctgc cccagctttc cagtggagca gataaaaagt tactggaaga 6240  
 gctcctgaac aagttcaaga gtagcatgca cttgcagctc acctgtttcc aagcagcttc 6300  
 ttcaaccatg atgaaaacag gctcgagtta actgactaaa agcttcgaaa ggaagctgag 6360  
 ttggctgctg ccaccgctga gcaataacta gcataacccc ttggggcctc taaacgggtc 6420  
 ttgaggggtt ttttgctgaa aggaggaact atcctcaggg tcgagaagta ctagaggatc 6480  
 ataatcagcc ataccacatt ttagaggtt ttacttgctt taaaaaacct cccacacctc 6540  
 cccctgaacc tgaaacataa aatgaatgca attgttgttg ttaacttggt tattgcagct 6600  
 tataatgggt acaataaag caatagcatc acaaatttca caaataaagc atttttttca 6660  
 ctgcattcta gttgtggtt gtccaaactc atcaatgtat cttatcatgt ctggatctga 6720  
 tcaactgctg agcctaggag atccgaacca gataagtga atctagtctc aaactatttt 6780  
 gtcattttta attttcgtat tagcttacga cgtacaccc agttcccatc tattttgtca 6840  
 ctcttcccta aataatcctt aaaaactcca tttccacccc tcccagttcc caactatttt 6900  
 gtccgcccac agcggggcat ttttcttct gttatgtttt taatcaaaca tctgccaac 6960  
 tccatgtgac aaaccgtcat cttcggttac ttttctctg tcacagaatg aaaatttttc 7020  
 tgtcatctct tcgttattaa tgtttgtaat tgactgaata tcaacgctta tttgcagcct 7080  
 gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcggtg tgggtggttac 7140  
 gcgcagcgtg accgctacac ttgccagcgc cctagcgcgc gctcctttcg ctttcttccc 7200  
 ttctttctc gccagttcg ccggtttcc ccgtcaagct ctaaatcggg ggtccctttt 7260  
 agggttccga tttagtgett tacggcacct cgaccccaaa aaacttgatt aggggtgatgg 7320  
 ttacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt tggagtccac 7380  
 gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta tctcggtcta 7440  
 ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa atgagctgat 7500  
 ttaacaaaaa ttaacgcga attttaacaa aatattaacg tttaacaattt caggtggcac 7560

```

ttttcgggga aatgtgcgcg gaacccttat ttgtttatatt ttctaaatac attcaaatat 7620
gtatccgctc atgagacaat aacctgata aatgcttcaa taatattgaa aaaggaagag 7680
tatgagtatt caacatttcc gtgtcgcctt tatcccttt tttgcggcat tttgccttcc 7740
tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc agttgggtgc 7800
acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga gttttcgccc 7860
cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg cggtattatc 7920
ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc agaatgactt 7980
ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag taagagaatt 8040
atgcagtgct gccataacca tgagtataa cactgcggcc aacttacttc tgacaacgat 8100
cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg taactcgcct 8160
tgatcgttgg gaaccggagc tgaatgaagc catacctaac gacgagcgtg acaccacgat 8220
gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgactac ttactctagc 8280
ttcccgcaa caattaatag actggatgga ggccgataaa gttgcaggac cacttctgcg 8340
ctcggccctt ccggttggtt ggtttattgc tgataaatct ggagccggtg agcgtgggtc 8400
tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgatcgc tagttatcta 8460
cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgtg agataggtgc 8520
ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac tttagattga 8580
tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg ataatctcat 8640
gacaaaaatc ccttaacgtg agttttcggt ccactgagcg tcagaccgcg gggcatgact 8700
aacatgagaa ttacaactta tatcgtatgg ggctgacttc aggtgctaca tttgaagaga 8760
taaattgcac tgaaatctag aaatatttta tctgattaat aagatgatct tcttgagatc 8820
gttttggtct gcgcgtaatc tcttgctctg aaaacggaaa aaaccgcctt gcaggcggtt 8880
ttttcgaagg ttctctgagc taccaactct ttgaaccgag gtaactggct tggaggagcg 8940
cagtcaccaa aacttgcctt ttcagtttag ccttaaccgg cgcatgactt caagactaac 9000
tcctctaaat caattaccag tggctgctgc cagtgggtgt tttgcatgtc tttccgggtt 9060
ggactcaaga cgatagttac cggataaggc gcagcggctc gactgaacgg ggggttcgtg 9120
catacagtcc agcttggagc gaactgccta cccggaactg agtgtcaggc gtggaatgag 9180
acaaacgcgg ccataacagc ggaatgacac cggtaaaccg aaaggcagga acaggagagc 9240
gcacgaggga gccgccaggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 9300
ccactgattt gagcgtcaga tttcgtgatg cttgtcaggg gggcgagacc tatggaaaaa 9360
cggctttgcc gcggccctct cacttccctg ttaagtatct tcctggcatc ttccaggaaa 9420
tctccgcccc gttcgtaage catttccgct cgccgcagtc gaacgaccga gcgtagcgag 9480
tcagtgagcg aggaagcgga atatatcctg tatcacatat tctgctgacg caccggtgca 9540
gccttttttc tcctgccaca tgaagcactt cactgacacc ctcactagtg ccaacatagt 9600
aagccagtat aactccgct agcgtgatg tccggcggtg cttttgccgt tacgcaccac 9660
cccgtcagta gctgaacagg agggacagct gatagaaaca gaagccagtt ctttctgcg 9720
ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc 9780
cgcagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcctgatg 9840
cggtattttc tccttacgca tctgtgcggt atttcacacc gcatagacca gccgcgtaac 9900

```

ctggcaaaat cggttacggt tgagtaataa atg 9933

<210> 9

<211> 9400

<212> DNA

<213> 人工序列

<220>

<223>

<400> 9

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgcccgtcg tattaaagag gggcgtggcc aagggcagtgt aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatactt cttcccgtat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840  
 catcgttgct gtcacataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccaggc catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgct 1020  
 acttgcatta cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatcc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgagg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440  
 ggccccctga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560  
 gatccccggc gcggtcacga actccagcag gaccatgtga tcgcgttct cgttggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgct 1680  
 gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgtgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagttca ccttgatgcc gttcttctgc ttgtcgcca tgatatagac 1800  
 gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860

gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctcggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gacttgaaga agtcgtgctg cttcatgttg tcggggtagc ggctgaagca 2040  
 ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt 2100  
 ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctcgc cctcgccgga 2160  
 cacgctgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220  
 ggtgaacagc tcctcgccct tgctcatccc gggtgactgg aagtacaggt tttcttgca 2280  
 gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400  
 actaaccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaaataa ttgttttgtt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520  
 atctctagtt gataactaatg gactgccaca ctgcccattc ttggtttgaa tccaatgctt 2580  
 ccagaatatg ccatcagatg aagggaatgt gcaagaagtg tctgacacca tgctagacat 2640  
 gctcttagtt tggaagttgg ttgtcacaag acatattctc tcttcccttt gtggtctctt 2700  
 aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcaataa ttatcatgtc 2760  
 cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgtt ttgtaatgat 2880  
 gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaacaa agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggcctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttta acactagttc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540  
 cgcgcccgcg atgaaggaga cagaccggga ggccgttgcg acagcagtcg aaagggttgc 3600  
 tgggatgctc cagcgcccg accagctgga caaggtggag cagtatcgca ggagagaagc 3660  
 gcggaagaag gcctccgtgg aggccagatt gaaggccgcc atccagtcac agttggacgg 3720  
 ggtgcgcaca ggctcagcc agctccacaa cgccctgaat gacgtcaaag acatccagca 3780  
 gtcgctggca gacgtcagca aggactggag gcagagcatc aacaccattg agagcctcaa 3840  
 ggacgtcaaa gacgccgtgg tgcagcacag ccagctcgcc gcagccgtgg agaacctcaa 3900  
 gaacatcttc tcagtgcctg agattgtgag ggagaccag gacctaatg aacaaggggc 3960  
 actcctgcaa gcccaccgga agctgatgga cctggaatgc tcccgggacg ggctgatgta 4020  
 cgagcagtac cgcatggaca gtgggaacac gcgtgacatg accctcatcc atggctactt 4080  
 tggcagcacg caggggctct ctgatgagct ggctaagcag ctgtggatgg tgctgcagag 4140  
 gtcactggtc actgtccgcc gtgacccac cttgctggte tcagttgtca ggatcattga 4200

aaggggaagag aaaattgaca ggcgcatact tgaccggaaa aagcaaactg gctttgttcc 4260  
tcctggggagg cccaagaatt ggaaggagaa aatgttcacc atcttggaga ggactgtgac 4320  
caccagaatt gagggcacac aggcagatac cagagagtct gacaagatgt ggcttgtccg 4380  
ccacctggaa attataagga agtacgtcct ggatgacctc attgtcgcca aaaacctgat 4440  
ggttcagtgc tttcctcccc actatgagat ctttaagaac ctctgaaca tgtaccacca 4500  
agccctgagc acgcggatgc aggacctgc atcggaagac ctggaagcca atgagatcgt 4560  
gagcctcttg acgtgggtct taaacaccta cacaagtact gagatgatga ggaacgtgga 4620  
gctggccccg gaagtggatg tcggcaccct ggagccattg ctttctccac acgtggctctc 4680  
tgagctgctt gacacgtaca tgtccacgct cacttcaaac atcatcgctt ggctgcggaa 4740  
agcgctggag acagacaaga aagactgggt caaagagaca gagccagaag ccgaccagga 4800  
cgggtactac cagaccacac tcctgccat tgtcttccag atgtttgaac agaattctca 4860  
agttgctgct cagataagtg aagatttgaa aacaaaggta ctagttttat gtcttcagca 4920  
gatgaattct ttcctaagca gatataaaga tgaagcgcag ctgtataaag aagagcacct 4980  
gaggaatcgg cagcaccctc actgctacgt tcagtacatg atcgccatca tcaacaactg 5040  
ccagaccttc aaggaatcca tagtcagttt aaaaagaaag tatttaaga atgaagtgga 5100  
agaggggtg tctccgagcc agcccagcat ggacgggatt ttagacgcca tcgcaagga 5160  
gggctgcagc ggtttgctgg aggaggtctt cctggacctg gagcaacatc tgaatgaatt 5220  
gatgacgaag aagtggctat tagggtcaaa cgctgtagac attatctgtg tcaccgtgga 5280  
agactatttc aacgattttg ccaaaattaa aaagccgtat aagaagagga tgacggccga 5340  
ggcgcaccgg cgcgtgggtg tggagtacct gcgggcggtc atgcagaagc gcatttcctt 5400  
ccggagcccc gaggagcgca aggagggtgc cgagaagatg gttaggaggagg cagagcagct 5460  
gcgcttcctg ttccggaagc tggcgtccgg tttcggggaa gacgtggacg gatactgcga 5520  
caccatcgtg gctgtggccg aagtgatcaa gctgacagac ctttctctgc tctacctgga 5580  
ggtctccact ctggtcagca agtatccaga catcagggat gaccacatcg gtgcgtgct 5640  
ggctgtgcgt ggggacgcca gccgtgacat gaagcagacc atcatggaga ccctggagca 5700  
gggcccagca caggccagcc ccagctacgt gcccctcttc aaggacattg tgggtgcccag 5760  
cctgaacgtg gccaagctgc tcaagtaact cgagttaact gactaaaagc ttcgaaagga 5820  
agctgagttg gctgctgcca ccgtgagca ataactagca taacccttg gggcctctaa 5880  
acgggtcttg aggggttttt tgctgaaagg aggaactatc ctcagggtcg agaagtacta 5940  
gaggatcata atcagccata ccacatttgt agaggtttta cttgctttta aaaacctccc 6000  
acacctcccc ctgaacctga aacataaaat gaatgcaatt gttgttgta acttgtttat 6060  
tgcagcttat aatggttaca aataaagcaa tagcatcaca aatttcacaa ataaagcatt 6120  
tttttcaactg cattctagtt gtggtttgtc caaactcacc aatgtatctt atcatgtctg 6180  
gatctgatca ctgcttgagc ctaggagatc cgaaccagat aagtgaatc tagttccaaa 6240  
ctattttgtc atttttaatt ttcgtattag cttacgacgc tacaccagat tccatctat 6300  
tttgtcactc ttcctaaat aatccttaaa aactccattt ccacctctc cagttcccaa 6360  
ctattttgtc cgccacagc ggggcatttt tcttctgtt atgttttta tcaaacatcc 6420  
tgccaactcc atgtgacaaa ccgtcatctt cggtactttt ttctctgtca cagaatgaaa 6480  
atttttctgt catctcttcg ttattaatgt ttgtaattga ctgaatatca acgcttattt 6540

gcagcctgaa tggcgaatgg gacgcgccct gtagcggcgc attagcgcg gcgggtgtgg 6600  
 tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct cttttcgctt 6660  
 tcttcccttc ctttctcgcc acgttcgccc gctttccccc tcaagctcta aatcgggggc 6720  
 tcccttttagg gttccgattt agtgctttac ggcacctega ccccaaaaaa cttgattagg 6780  
 gtgatgggtc acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg 6840  
 agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 6900  
 cggtctattc ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg 6960  
 agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag 7020  
 gtggcacatt tcgggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 7080  
 caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 7140  
 ggaagagtat gagtattcaa catttccttg tcgcccttat tccctttttt gcggcatttt 7200  
 gccttcctgt ttttgetcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 7260  
 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt 7320  
 ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgccg 7380  
 tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 7440  
 atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 7500  
 gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 7560  
 caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 7620  
 ctgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 7680  
 ccacgatgcc ttagcaatg gcaacaacgt tgcgcaaact attaacggc gaactactta 7740  
 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 7800  
 ttctgcgctc ggcccttcg gctggctggg ttattgctga taaatctgga gccggtgagc 7860  
 gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 7920  
 ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 7980  
 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 8040  
 agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 8100  
 atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccgcgggg 8160  
 catgactaac atgagaatta caacttatat cgtatggggc tgacttcagg tgctacattt 8220  
 gaagagataa attgcactga aatctagaaa tattttatct gattaataag atgatcttct 8280  
 tgagatcggt ttggtctgcg cgtaatctct tgctctgaaa acggaaaaaa ccgccttgca 8340  
 gggcggtttt tcgaagggtc tctgagctac caactctttg aaccgaggta actggcttgg 8400  
 aggagcgcag tcaccaaaac ttgtccttcc agtttagcct taaccggcgc atgacttcaa 8460  
 gactaactcc tctaaatcaa ttaccagtgg ctgctgccag tgggtgctttt gcatgtcttt 8520  
 ccgggttggg ctcaagacga tagttaccgg ataaggcgca gcggtcggac tgaacggggg 8580  
 gttcgtgcat acagtccagc ttggagcgaa ctgcctaccc ggaactgagt gtcaggcgtg 8640  
 gaatgagaca aacgcggcca taacagcgga atgacaccgg taaaccgaaa ggcaggaaca 8700  
 ggagagcgca cgaggagacc gccaggggaa acgcctggta tctttatagt cctgtcgggt 8760  
 ttcgccacca ctgatttgag cgtcagattt cgtgatgctt gtcagggggg cggagcctat 8820  
 ggaaaaacgg ctttgccgcg gccctctcac ttccctgtta agtatcttcc tggcatcttc 8880

caggaaatct ccgccccgtt cgtaagccat ttccgctcgc cgcagtcgaa cgaccgagcg 8940  
 tagcgagtca gtgagcgagg aagcggaata tatectgtat cacatattct gctgacgcac 9000  
 cgggtgcagcc ttttttctcc tgccacatga agcacttcac tgacaccctc atcagtgcca 9060  
 acatagtaag ccagtataca ctccgctagc gctgatgtcc ggcggtgctt ttgccgttac 9120  
 gcaccacccc gtcagtagct gaacaggagg gacagctgat agaaacagaa gccagttctt 9180  
 tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac 9240  
 cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg 9300  
 cctgatgcgg tattttctcc ttacgcctct gtgcggtatt tcacaccgca tagaccagcc 9360  
 gcgtaacctg gcaaaatcgg ttacggttga gtaataaatg 9400

<210> 10

<211> 10077

<212> DNA

<213> 人工序列

<220>

<223>

<400> 10

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcacg 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcatTTTTCT 180  
 gaagtgcaaa ttgcccgtcg tattaaagag gggcgtggcc aagggcattgt aaagactata 240  
 ttgcgcgctg tgtgacaatt tacccaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatacaag tgcatactt cttcccgtat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcgc cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggt cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840  
 catcgttgct gctccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccaggag catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgtgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatga cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatec 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgcgg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380

```

tggcgcagca acgtcgccaa ctccattgt tatttctatg caaaacccca caccaacaac 1440
ggccccctga taataaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500
aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560
gatccccggc gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc 1620
tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgctc 1680
gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgctgccgt cctcgatggt 1740
gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac 1800
gttggtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860
gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaact tcacctcggc 1920
gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980
gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040
ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt 2100
gggtcagatg aacttcaggg tcagcttgc gtaggtggca tcgccctcgc cctcgccgga 2160
cacgctgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220
gggtgaacagc tcctcgccct tgctcatccc gggtgactgg aagtacaggt tttcttgcca 2280
gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340
tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400
actaacccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460
tgtgaaataa ttgttttgtg tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520
atctctagtt gatactaatg gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580
ccagaatatg ccatcagatg aagggaatgt gcaagaagtg tctgacacca tgctagacat 2640
gctcttagtt tggaagttgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700
aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760
cctcccatca atgaggtggt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820
tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgtt ttgtaatgat 2880
gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940
atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaacaa agctctctcc 3000
ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060
ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120
accgttatct ggctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180
tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240
gctcagtcga cagatctgtc gacggtttta acactagttc gcgacctact ccggaatatt 3300
aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360
ggtaaaaaat acagctacat aaattacaca atttaacgg atcgatgagc tccatatggc 3420
ctggagccat ccgcaatttg aaaaagggtg cgggtccggc ggaggtagcg gcggaggttc 3480
ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540
cgcgcccgcg gaagcagctg gtgggaaata cagaagcaca gtcagcaaaa gcaaagaccc 3600
ctcggggctg ctcatctctg tgatcaggac tctgtctact agtgacgatg tcgaagacag 3660
ggaaaatgaa aagggtcgcc ttgaagaagc ctacgagaaa tgtgaccgtg acctggatga 3720

```

attgattgta cagcactaca cagaattgac gacagccatt cgcacatacc agagcatcac 3780  
 agagcgcatac actaactccc gaaataaaaat aaagcaggta aaagagaacc tgctttcatg 3840  
 caagatgctg ctgcactgca aacgggatga gcttcggaaa ctgtggattg aaggaattga 3900  
 gcataagcat gtctgaact tgttgatga aattgagaat atcaagcaag tgcctcaaaa 3960  
 gctggaacag tgcatggcca gcaagcacta tctcagtgcc actgacatgt tgggtgcagc 4020  
 agttgagtct ttggagggcc cctgctcca ggtggaagga ctgagtgacc ttcgactaga 4080  
 gcttcacagc aagaagatga accttcactt ggttctcata gatgaactac accggcacct 4140  
 gtacatcaaaa tcgactagcc gagttgtgca gcgtaacaag gaaaaaggga aaatcagctc 4200  
 cctcgtgaaa gatgcttctg ttcctctgat tgatgttaca aacctcccta ctctcgaaa 4260  
 attccttgat acctctcact attctactgc tggaagctca agtgtgaggg agataaatct 4320  
 gcaggacatc aaggaagatt tagaattgga tccagaggaa aacagcacc cgtttatggg 4380  
 tatcctcatt aagggettgg cgaaactgaa gaagatccca gaaacagtta aggcaatcat 4440  
 agagcgcttg gagcaggagt tgaagcaaat tgtgaagagg tctacaacc aggtggcaga 4500  
 cagtggctat cagcgggggg agaacgttac tgtggagaac caaccaaggt tgcttctaga 4560  
 actgctggag ttactgtttg acaagtttaa tgctgtagcc gctgcacact ctgtggctct 4620  
 gggataacctg caggacactg tagtgactcc actgactcag caggaagata tcaaactgta 4680  
 tgatatggca gatgtatggg tgaagatcca agatgttcta cagatgctat taactgagta 4740  
 cttggatatg aaaaatactc gtacggcctc tgaacatca gctcaactaa gctatgccag 4800  
 cactggacga gagtttgag ccttttttgc caagaagaaa cctcaaaggc caaaaaattc 4860  
 tcttttcaag ttcgatcgt cctcccatgc catcagtatg agcgcctatc tgcgagaaca 4920  
 gagaaggagg ctctatagtc ggagtggaga actgcaaggg ggtcctgatg acaacttaat 4980  
 tgaagggtgga ggaacaaaat ttgtctgcaa acctggagcc agaaacatta ccgtcatatt 5040  
 ccaccatta ctaagattta ttcaggagat tgagcatgct ctgggtcttg gccagccaa 5100  
 acagtgtcct cttcgagagt ttctcaccgt gtacatcaaa aacatctttc tcaatcaagt 5160  
 cttggctgag atcaacaagg agattgaagg agtcactaaa acatctgacc ctttgaagat 5220  
 tctggccaac gcagacacca tgaagggtgct gggagtgcag cggcctctcc tacagagcac 5280  
 aatcattgtg gagaagacag ttcaagacct cctgaacctg atgcatgact tgagtgcata 5340  
 ttcagatcaa ttctcaaca tgggtgtcgt gaagctccag gagtacaagg acacctgcac 5400  
 tgcagcttac aggggtattg tccagtcaga agaaaaactt gtcatcagtg catcctgggc 5460  
 aaaagatgat gatatcagca gactcttgaa atctctacca aactggatga atatggctca 5520  
 acccaaacag ctgaggccaa aaagagagga ggaagaagat ttcataaggg cagcttttgg 5580  
 caaggagtct gaagttctta ttgggaacct gggtgataaa ttaatccctc cacaagacat 5640  
 ccttcgtgac gtcagtgacc tcaaagcctt ggccaacatg catgaaagcc tggaatgggt 5700  
 ggcaagtcga acaaagtcag ctttctccaa tctttctaca tcccagatgc tttctctgc 5760  
 tcaagacagc cacacgaaca cggatctccc cccagtgtca gagcagatca tgcagactct 5820  
 cagtgaactt gccaaatcgt tccaggatat ggctgaccgc tgcttgcttg tcttacatct 5880  
 ggaagtgagg gttcactgtt tccactatct tatccctctt gcaaaggagg ggaactatgc 5940  
 cattgtggct aatgtggaaa gtatggatta tgacccctg gtggtcaagc tcaacaaaga 6000  
 tatcagcgcc attgaagagg ccatgagcgc cagccttcag cagcacaagt tccagtatat 6060

cttcgaaggc ctgggccacc tgatctcctg catcctcatt aatgggtgccc agtacttcag 6120  
 gcgcacacagt gaggctggca tcaagaaaat gtgttaggaac atttttgttc ttcagcagaa 6180  
 tttgaccaac atcaccatgt cgcggggaggc agacctggac tttgcaaggc agtactacga 6240  
 gatgctttac aacacagctg acgagctcct gaacctgggtg gtggaccagg gtgtgaagta 6300  
 cacggagctg gagtacatcc acgctctgac cctgctgcac cgcagccaga ctgggggtggg 6360  
 ggaactgacc acccagaaca cgaggctgca gaggctcaaa gagatcatct gcgagcaggc 6420  
 tgccatcaag caagccacca aggacaagaa gataactacc gttggctcga gttactgac 6480  
 taaaagcttc gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa 6540  
 ccccttgggg cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatcctc 6600  
 agggctcaga agtactagag gatcataatc agccatacca catttgtaga ggttttactt 6660  
 gctttaaaaa acctcccaca cctcccctg aacctgaaac ataaaatgaa tgcaattgtt 6720  
 gttgttaact tgtttattgc agcttataat ggttacaaat aaagcaatag catcacaaat 6780  
 ttcacaaata aagcattttt ttcactgcat tctagttgtg gtttgtccaa actcatcaat 6840  
 gtatcttata atgtctggat ctgatcactg cttgagccta ggagatccga accagataag 6900  
 tgaaatctag ttccaaacta ttttgtcatt ttttaatttc gtattagctt acgacgtac 6960  
 acccagttcc catctatttt gtcactcttc cctaaataat ccttaaaaac tccatttcca 7020  
 cccctcccag ttcccaacta ttttgtccgc ccacagcggg gcatttttct tctgttatg 7080  
 tttttaataca aacatcctgc caactccatg tgacaaaccg tcactctcgg ctactttttc 7140  
 tctgtcacag aatgaaaatt tttctgtcat ctcttcgtta ttaatgtttg taattgactg 7200  
 aatatcaacg cttatttgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt 7260  
 aagcgcggcg ggtgtgggtg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc 7320  
 gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcgccggct tccccgtca 7380  
 agctctaaat cgggggctcc ctttaggggt cggatttagt gctttacggc acctcgacct 7440  
 caaaaaactt gattaggggt atggttcacg tagtgggcca tcgccctgat agacggtttt 7500  
 tcgccctttg acgttggagt ccacgttctt taatagtga ctcttgttcc aaactggaac 7560  
 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 7620  
 ctattgggta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta aaaaaatatt 7680  
 aacgtttaca atttcaggtg gcaacttttc gggaaatgtg cgcggaaccc ctatttgttt 7740  
 atttttctaa atacattcaa atatgtatcc gtcacatgaga caataaccct gataaatgct 7800  
 tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc 7860  
 cttttttgcg gcattttgce ttctgtttt tgctcaccca gaaacgctgg tgaaagtaaa 7920  
 agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc tcaacagcgg 7980  
 taagatcctt gagagtttgc gccccgaaga acgttttcca atgatgagca cttttaaagt 8040  
 tctgctatgt ggcgcggtat tatcccgat tgacgcgggg caagagcaac tcggctcgccg 8100  
 catacactat tctcagaatg acttggttga gtactacca gtcacagaaa agcatcttac 8160  
 ggatggcatg acagtaagag aattatgcag tgctgccata accatgagtg ataacactgc 8220  
 ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt ttttgacaaa 8280  
 catgggggat catgtaactc gccttgatcg ttgggaaccg gagctgaatg aagccatacc 8340  
 aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc gcaaacattt 8400

aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga tggaggcgga 8460  
taaagttgca ggaccacttc tgcgctcggc cttccggct ggctggttta ttgctgataa 8520  
atctggagcc ggtgagcgtg ggtctcgcg tatcattgca gcactggggc cagatggtaa 8580  
gccctcccg atcgtagtta tctacacgac ggggagtcag gcaactatgg atgaacgaaa 8640  
tagacagatc gctgagatag gtgcctcact gattaagcat tggtaactgt cagaccaagt 8700  
ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa ggatctaggt 8760  
gaagatcctt ttgataaatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg 8820  
agcgtcagac cgcggggcat gactaacatg agaattacaa cttatatcgt atggggctga 8880  
cttcaggtgc tacatttgaa gagataaatt gcaactgaaat ctagaaatat tttatctgat 8940  
taataagatg atcttcttga gatcgttttg gtctgcgcgt aatctcttgc tctgaaaacg 9000  
gaaaaaacg ccttgcaagg cggtttttcg aaggttctct gagctaccaa ctctttgaac 9060  
cgaggtaact ggcttggagg agcgcagtca ccaaaacttg tcctttcagt ttagccttaa 9120  
ccggcgcag acttcaagac taactcctct aaatcaatta ccagtggctg ctgccagtgg 9180  
tgcttttgca tgtctttccg ggttggactc aagacgatag ttaccggata aggcgcagcg 9240  
gtcggactga acggggggtt cgtgcataca gtccagcttg gagcgaactg cctaccgga 9300  
actgagtgtc aggcgtggaa tgagacaaac gcggccataa cagcggaatg acaccgtaa 9360  
accgaaaggc aggaacagga gagcgcacga gggagccgcc aggggaaacg cctggtatct 9420  
ttatagtcct gtcgggtttc gccaccactg atttgagcgt cagatttcgt gatgcttgct 9480  
aggggggcg agcctatgga aaaacggctt tgccgcggcc ctctcacttc cctgttaagt 9540  
atcttcttgg catcttcag gaaatctccg ccccgttcgt aagccatttc cgctgcgcgc 9600  
agtcgaacga ccgagcgtag cgagtcagt agcgaggaag cggaatatat cctgtatcac 9660  
atattctgct gacgcaccgg tgcagccttt tttctctgc cacatgaagc acttactga 9720  
caccctcatc agtgccaaca tagtaagcca gtatacactc cgctagcgt gatgtccggc 9780  
ggtgcttttg ccgttacgca ccacccgctc agtagctgaa caggaggac agctgataga 9840  
aacagaagcc agttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc 9900  
tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc 9960  
gaggaagcgg aagagcgcct gatgcggtat tttctcctta cgcattctgt cggtatttca 10020  
caccgcatag accagccgcg taacctggca aaatcggtta cggttgagta ataaatg 10077

<210> 11

<211> 9283

<212> DNA

<213> 人工序列

<220>

<223>

<400> 11

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcacg 120  
ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
gaagtgcaaa ttgcccgtcg tattaaagag gggcgtggcc aaggcatgt aaagactata 240

ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggcttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatacaag tgcatacatt cttcccgtat 360  
 gcccacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cggtaggcaat 480  
 gccctgcctc cgggtgctgc cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagtgc cggaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggg cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac ttgttttag ggcgactgcc ctgctgcgta acatcggtgc tgctgcgtaa 840  
 catcggtgct gctccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccagg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatca cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatec 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggc cctaaggtag cgagtttaaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cttttgttg ataggcgcgg tcttgttttt 1320  
 aggtttaata ttttatatta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctccattgt ttttctatg caaaaccca caccaacaac 1440  
 ggcccccgca taataaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560  
 gatcccgccg gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggtgtgc ggacgcagca cggggccgtc 1680  
 gccgatgggg gtgttctgct ggtagtgtgc ggagagctgc acgctgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac 1800  
 gttgtggctg ttgtagtgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860  
 gtcgatgcc ttcagctcga tgcggttcac cagggtgtgc ccctcgaact tcacctggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gacttgaaga agtcgtgctg ctcatgtgg tcggggtagc ggctgaagca 2040  
 ctgcacgcc taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt 2100  
 ggtgcagatg aacttcaggg tcagcttgc gtaggtggca tcgccctgc cctcgccgga 2160  
 cacgtgaac ttgtggccgt ttacgtgcc gtccagctgc accaggatgg gcaccaccc 2220  
 ggtgaacagc tcctcgccct tgcateccc gggtagctgg aagtacaggt tttcttcgga 2280  
 gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caactagag tcagcattta atcgccaacc 2400  
 actaacccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaataa ttgttttgt tggtgaaatt cgatgctgag tgtataccaa caatgaacc 2520  
 atctctagtt gataactaat gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580

ccagaatatg ccatcagatg aagggaatgt gcaagaagtg tctgacacca tgctagacat 2640  
 gctcttagtt tggaagttgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700  
 aaatttcagc ttttgaggaa atgggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760  
 cctcccatca atgaggtggt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgtt ttgtaatgat 2880  
 gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaacaa agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggcctacgtg acacaaggca cgttggttaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttaa aactagtctc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540  
 cgcggcccg accacggccg agttgttcga ggagcctttt gtggcagatg aatatattga 3600  
 acgtcttgta tggagaaccc caggaggagg ctctagaggt ggacctgaag cttttgatcc 3660  
 taaaagatta ttagaagaat ttgtaaataa tattcaggaa ctccagataa tggatgaaag 3720  
 gattcagagg aaagtagaga aactagagca acaatgtcag aaagaagcca aggaatttgc 3780  
 caagaaggta caagagctgc agaaaagcaa tcaggttgcc ttccaacatt tccaagaact 3840  
 agatgagcac attagctatg tagcaactaa agtctgtcac cttggagacc agttagaggg 3900  
 ggtaaacaca cccagacaac gggcagtgga ggctcagaaa ttgatgaaat actttaatga 3960  
 gtttctagat ggagaattga aatctgatgt ttttacaat tctgaaaaga taaaggaagc 4020  
 agcagacatc attcagaagt tgcacctaat tgcccaagag ttaccttttg atagattttc 4080  
 agaagttaaa tccaaaattg caagtaaata ccatgattta gaatgccagc tgattcagga 4140  
 gtttaccagt gctcaaagaa gaggtgaaat ctccagaatg agagaagtag cagcagtttt 4200  
 acttcatttt aagggttatt cccattgtgt tgatgtttat ataaagcagt gccaggaggg 4260  
 tgcttatttg agaaatgata tatttgaaga cgctggaata ctctgtcaaa gagtgaacaa 4320  
 acaagttgga gatattctca gtaatccaga aacagtcttg gctaaactta ttcaaatgt 4380  
 atttgaaatc aaactacaga gttttgtgaa agagcagtta gaagaatgta ggaagtccga 4440  
 tgcagagcaa tatctcaaaa atctctatga tctgtataca agaaccacca atctttccag 4500  
 caagctgatg gagttaatt taggtactga taaacagact ttcttgtcta agcttatcaa 4560  
 atccattttc atttcctatt tggagaacta tattgaggtg gagactggat atttgaaaag 4620  
 cagaagtgt atgacctac agcgtatta tgattcgaaa aaccatcaaa agagatccat 4680  
 tggcacagga ggtattcaag atttgaagga aagaattaga cagcgtacca acttaccact 4740  
 tgggccaagt atcgatactc atggggagac ttttctatcc caagaagtgg tggttaatct 4800  
 tttacaagaa accaaacaag ctttgaaag atgtcatagg ctctctgac cttctgactt 4860  
 accaaggaat gccttcagaa tttttacat tcttgtgga tttttatgta ttgagcatat 4920

tgattatgct ttggaaacag gacttgctgg aattccctct tcagattcta ggaatgcaaa 4980  
 tctttatfff ttggacgttg tgcaacaggc caatactatt tttcatcttt ttgacaaaca 5040  
 gtttaatgat caccttatgc cactaataag ctcttctcct aagttatctg aatgccttca 5100  
 gaagaaaaaa gaaataattg aacaaatgga gatgaaattg gatactggca ttgataggac 5160  
 attaaattgt atgattggac agatgaagca tatfittggct gcagaacaga agaaaacaga 5220  
 ttttaagcca gaagatgaaa acaatgtfff gattcaatat actaatgcct gtgtaaaagt 5280  
 ctgtgcttac gtaagaaaaac aagtggagaa gattaaaaat tccatggatg ggaagaatgt 5340  
 ggatacagtt ttgatggaac ttggagtacg ttttcatcga cttatctatg agcatcttca 5400  
 acaatatfctc tacagttgta tgggtggcat gttggcaatt tgtgatgtag ccgaatatag 5460  
 gaagtgtgcc aaagacttca agattccaat ggtattacat ctttttgata ctctgcatgc 5520  
 tctttgcaat cttctggtag ttgccccaga taattttaag caagtctgct caggagaaca 5580  
 acttgctaat ctggacaaga atataactca ctcttctgta caacttctg ctgattatag 5640  
 atctgccccg cttgctcgac acttcageta actcgagtta actgactaaa agcttctgaaa 5700  
 ggaagctgag ttggctgctg ccaccgctga gcaataacta gcataacccc ttggggcctc 5760  
 taaacgggtc ttgaggggtt ttttgctgaa aggaggaact atcctcaggg tcgagaagta 5820  
 ctagaggatc ataatcagcc ataccacatt tgtagaggtt ttacttgctt taaaaaacct 5880  
 cccacacctc cccctgaacc tgaaacataa aatgaatgca attgttggtt ttaacttggt 5940  
 tattgcagct tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc 6000  
 atttttttca ctgcattcta gttgtggttt gtccaaacte atcaatgtat cttatcatgt 6060  
 ctggatctga tcactgcttg agcctaggag atccgaacca gataagtga atctagtctc 6120  
 aaactatfff gtcatfctta attttctgat tagcttacga cgctacaccc agttcccatc 6180  
 tatfittgtca ctcttcccta aataatcctt aaaaactcca tttccacccc tcccagttcc 6240  
 caactatfff gtccgcccac agcggggcat tttcttctc gttatgtfff taatcaaaca 6300  
 tcttgccaac tccatgtgac aaaccgtcat ctctggctac ttttctctg tcacagaatg 6360  
 aaaatfcttc tgtcatctct tctgtattaa tgtttgtaat tgactgaata tcaacgctta 6420  
 tttgcagcct gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcgggtg 6480  
 ttgttggttac gcgcagcgtg accgtacac ttgccagcgc cctagcgcgc gctcctttcg 6540  
 ctttcttccc ttccttctc gccacgttct ccgctttcc cgtcaagct ctaaateggg 6600  
 ggctcccttt aggtttccga tttagtgtt tacggcacct cgaccccaaa aaacttgatt 6660  
 aggttgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttctg cttttgacgt 6720  
 ttgagtccac gttctttaat agtggactct tgttccaaac ttgaacaaca ctcaacccta 6780  
 tctcggctca tttttttgat ttataaggga ttttgccgat ttcggcctat ttgttaaaaa 6840  
 atgagctgat ttaacaaaaa tttaacgcga atfittaaca aatattaacg tttacaatf 6900  
 caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttatf ttttaatac 6960  
 attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 7020  
 aaaggaagag tatgagtatt caacatttcc gtgtcgcctt tatccctf tttgcggcat 7080  
 tttgccttcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc 7140  
 agttgggtgc acgagtgggt tacatcgaa tggtctcaa cagcggtaag atccttgaga 7200  
 gttttcgcgc cgaagaacgt tttccaatga tgagcactf taaagttctg ctatgtggcg 7260

cggattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 7320  
 agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 7380  
 taagagaatt atgcagtgcg gccataacca tgagtataa cactgcggcc aacttacttc 7440  
 tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg 7500  
 taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 7560  
 acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact gggaactac 7620  
 ttactctagc ttcccggcaa caattaatag actggatgga ggcgataaa gttgcaggac 7680  
 cacttctgcg ctccggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 7740  
 agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggttaagccc tcccgatcg 7800  
 tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 7860  
 agatagggtc ctactgatt aagcattggg aactgtcaga ccaagtttac tcatatatac 7920  
 tttagattga tttaaaactt ctttttaaat ttaaaggat ctaggtgaag atcctttttg 7980  
 ataatctcat gacaaaaate ccttaacgtg agttttcgtt ccactgagcg tcagaccgcg 8040  
 gggcatgact aacatgagaa ttacaactta tatcgtatgg ggctgacttc aggtgctaca 8100  
 tttgaagaga taaattgcac tgaaatctag aaatatatta tctgattaat aagatgatct 8160  
 tcttgagatc gttttggtct gcgcgtaate tcttgctctg aaaacgaaa aaaccgcctt 8220  
 gcagggcggg ttttcgaagg ttctctgagc taccaactct ttgaaccgag gtaactggct 8280  
 tggaggagcg cagtcaccaa aacttgtcct ttcagtttag ccttaaccgg cgcattgactt 8340  
 caagactaac tcctctaaat caattaccag tggtctgtgc cagtgggtgct tttgcatgtc 8400  
 tttccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg gactgaacgg 8460  
 ggggttcgtg catacagtcc agcttgagc gaactgccta cccggaactg agtgtcaggc 8520  
 gtggaatgag acaaacgcgg ccataacagc ggaatgacac cggtaaaccg aaaggcagga 8580  
 acaggagagc gcacgaggga gccgccaggg gaaacgcctg gtatctttat agtcctgtcg 8640  
 ggtttcgcca ccactgattt gagcgtcaga tttcgtgatg cttgtcaggg gggcggagcc 8700  
 tatggaaaaa cggctttgcc gcggccctct cacttccctg ttaagtatct tcctggcatc 8760  
 ttccaggaaa tctccgcccc gttcgttaagc catttccgct cgccgcagtc gaacgaccga 8820  
 gcgtagcgag tcagtgcgag aggaagcgga atatatectg taccacatat tctgctgacg 8880  
 caccggtgca gccttttttc tctgccaca tgaagcactt cactgacacc ctcatcagt 8940  
 ccaacatagt aagccagtat aactccgct agcgtgatg tccggcgggtg cttttgccgt 9000  
 tacgaccac cccgtcagta gctgaacagg agggacagct gatagaaaca gaagccagtt 9060  
 ctttctgcg ttatccccctg attctgtgga taaccgtatt accgcctttg agtgagctga 9120  
 taccgctcgc cgcagccgaa cgaccgagc cagcgagtca gtgagcgagg aagcgggaaga 9180  
 gcgcctgatg cggtattttc tccttacgca tctgtgcggt atttcacacc gcatagacca 9240  
 gccgcgtaac ctggcaaaat cggttacggt tgagtaataa atg 9283

<210> 12

<211> 9555

<212> DNA

<213> 人工序列

<220>

&lt;223&gt;

&lt;400&gt; 12

```

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60
taaactatga caataaagtc tttaaactaga cagaatagtt gtaaactgaa atcagtcag 120
ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcatTTTTt 180
gaagtgcAAA ttgcccgtcg tattaagag gggcgtggcc aagggcattg aaagactata 240
ttcgcgcgct tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300
aacgaattgt taggtggcgg tacttgggtc gatataaag tgcatactt cttcccgat 360
gcccacttt gtatagagag ccaactgcggg atcgtcacg taatctgctt gcacgtagat 420
cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480
gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540
ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600
ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660
ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720
ccgcattgat ttgacttggg cagggccgag cctacatgtg cgaatgatgc ccatacttga 780
gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840
catcgttgct gctccataac atcaaacatc gaccacgcgc gtaacgcgct tgctgcttgg 900
atgcccaggc catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960
cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020
acttgcatga cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080
gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140
atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200
cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260
tgacagcgtc agcggcaagt taatgcctat catTTTgttg ataggcgcg tcttgttttt 1320
aggtttaata tttattttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380
tggcgcagca acgtcgcaa ctccattgt tatttctatg caaaaccca caccaacaac 1440
ggccccctga taataaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500
aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560
gatcccgcg gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc 1620
tttctcagg gcggactggg tgctcaggta gtggttgctg ggacgagca cggggccgtc 1680
gccgatgggg gtgttctgct ggtagtggtc ggagagctgc acgtgccgt cctcgatgtt 1740
gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcgcca tgatatagac 1800
gttgtggctg tttagttgt actccagctt gtccccagg atgttgccgt cctccttgaa 1860
gtcgatgccc ttcagtcga tgcggttcc cagggtgtcg cctcgaact tcacctcggc 1920
gcgggtcttg tagttgccgt cgtccttgaa gaagatggtg cgtccttga cgtagccttc 1980
gggcattggc gactgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040
ctgcacgcc taggtcaggg tggtcacgag ggtgggccag ggacgggca gcttgccggt 2100
ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgcctcgc cctcgccgga 2160
cacgtgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220

```

```

ggtgaacagc tcctcgccct tgctcatccc gggtagactgg aagtacaggt tttcttgcga 2280
gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340
tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400
actaaccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460
tgtgaaataa ttgttttgtt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520
atctctagtt gatactaagt gactgccaca ctgcccattc ttggtttgaa tccaatgctt 2580
ccagaatatg ccatcagatg aagggaatgt gcaagaagt tctgacacca tgctagacat 2640
gctcttagtt tggaagtggg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700
aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcaataa ttatcatgtc 2760
cctcccatca atgaggtgtt gttgcaaagt cgtgggtgtc ttgacctga atacaccatg 2820
tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgtt ttgtaatgat 2880
gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940
atgacaaatg gtgctcgata tagggttgta atcacgcgc cccttaacaa agctctctcc 3000
ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060
ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120
accgttatct ggcctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180
tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240
gctcagtcga cagatctgtc gacggtttaa aactagtctc gcgacctact ccggaatatt 3300
aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360
ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420
ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480
ttgggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccgcggagaa 3540
cagcgagagt ctgggcaccg tccccgagca cgagcggatc ttgcaggaga tcgagagcac 3600
cgacaccgcc tgtgtggggc ccaccctccg gtctgtgtat gatgaccaac caaatgcgca 3660
caagaagttt atggaaaagt tagatgcttg tatccgtaat catgacaagg aaattgaaaa 3720
gatgtgtaat tttcatcatc agggttttgt agatgctatt acagaactcc ttaaagtaag 3780
gactgatgca gaaaaactga aggtgcaagt tactgatacc aaccgaaggt ttcaagatgc 3840
tggaagagag gtgatatgcc acacagaaga tatcattcga tgtagaattc agcagagaaa 3900
tattacaact gtagtagaaa aattgcagtt atgccttcc tgcctagaaa tgtacagtaa 3960
gctgaaaagaa cagatgagtg ccaaaaggta ctattctgcc ctaaaaacta tggaacaatt 4020
agagaatgtg tactttccct gggtagtca ataccggttt tgcagctca tgatagaaaa 4080
tcttcccaaa ctccgtgagg atattaaaga aatctccatg tctgatctca aagacttttt 4140
ggaaagtatt cgaaaacatt ctgacaaaat aggtgaaaca gcaatgaaac aggcacagca 4200
tcagaaaacc ttcagtgttt ctctgcagaa acaaaataaa atgaaatttg ggaaaaatat 4260
gtatataaat cgtgatagaa ttccagagga aaggaatgaa actgtattga aacattcact 4320
tgaagaagag gatgagaatg aagaagagat cttaactgtt caggatcttg ttgatttttc 4380
ccctgtttat cgatgtttgc acatttatc tgttttgggt gacgaggaaa catttgaaaa 4440
ctattatcga aaacaaagaa agaaacaagc aagactggta ttgcaacccc agtcgaatat 4500
gcatgaaaca gttgatggct atagaagata tttcactcaa attgtagggt tctttgttgt 4560

```

agaagatcac attttacatg tgaccaaggg attagtaacc agggcataca ctgatgaact 4620  
 ttggaacatg gccctctcaa agataattgc tgtccttaga gctcattcat cctattgcac 4680  
 tgatcctgat cttgttctgg agctgaagaa tcttattgta atatttgcag atactttaca 4740  
 gggttatggt tttccagtga accgactttt tgacctttta tttgaaataa gagaccaata 4800  
 caatgaaaca ctgcttaaga aatgggctgg agttttcagg gacatttttg aagaagataa 4860  
 ttacagcccc atccctgttg tcaatgaaga agaataataa attgtcatca gcaaatttcc 4920  
 ctttcaagat ccagaccttg aaaagcagtc tttcccaaag aaattcccca tgtctcagtc 4980  
 agtgcctcat atttacattc aagttaaaga atttatttat gccagcctta aattttcaga 5040  
 gtcactacac cggagctcaa cagaaataga cgatatgctt agaaaatcaa caaatctgct 5100  
 gctgaccaga actttgagta gctgtttact gaaccttatt agaaaacctc atataggttt 5160  
 gacagagctg gtacaaatca tcataaacac aacacacctg gagcaagctt gtaaatatct 5220  
 tgaggacttt ataactaaca ttacaaatat ttcccaagaa actgttcata ctacaagact 5280  
 ttatggactt tctactttca aggatgctcg acatgcagca gaaggagaaa tatataccaa 5340  
 actgaatcaa aaaattgatg aatttgttca gcttgctgat tatgactgga caatgtctga 5400  
 gccagatgga agagctagtg gttatttaaat ggaccttata aattttttga gaagcatctt 5460  
 tcaagtgttt actcatttgc ctgggaaagt tgctcagaca gcttgcatgt cagcctgcc 5520  
 gcactgtca acatccttaa tgcagatgct actggacagt gagttaaac aaataagcat 5580  
 gggagctgtt cagcagttta acttagatgt catacagtgt gaattgtttg ccagctctga 5640  
 gcctgtgcc ggattccagg gggataccct gcagctagca ttcattgacc tcagacaact 5700  
 ccttgacctg tttatggttt gggattggtc tacttaccta gctgattatg ggcagccagc 5760  
 ttctaagtac cttcgggtga atccaaacac agcccttact cttttggaga agatgaagga 5820  
 tactagcaaa aagaacaata tatttgetca gttcaggaag aatgatcgag acaaacagaa 5880  
 gttgatagag acagtcgtga aacagctgag aagtttggtg aatggtatgt cccagcacat 5940  
 gggctcgagt taactgacta aaagcttcga aaggaagctg agttggctgc tgccaccgct 6000  
 gagcaataac tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg 6060  
 aaaggaggaa ctatcctcag ggtcgagaag tactagagga tcataatcag ccataccaca 6120  
 tttgtagagg ttttacttgc tttaaaaaac ctcccacacc tccccctgaa cctgaaacat 6180  
 aaaatgaatg caattgttgt tgtttaactt tttattgcag cttataatgg ttacaaataa 6240  
 agcaatagca tcacaaatct cacaataaaa gcattttttt cactgcattc tagttgtggt 6300  
 ttgtccaaac tcatcaatgt atcttateat gtctggatct gatcactgct tgagcctagg 6360  
 agatccgaac cagataagtg aaatctagtt ccaaactatt ttgtcatttt taattttcgt 6420  
 attagcttac gacgtacac ccagttccca tctattttgt cactcttccc taaataatcc 6480  
 ttaaaaactc catttccacc cctcccagtt cccaactatt ttgtccgccc acagcggggc 6540  
 atttttcttc ctgttatgtt tttaatcaaa catcctgcc acctcatgtg acaaaccgtc 6600  
 atcttcggct actttttctc tgtcacagaa tgaaaatttt tctgtcatct cttegttatt 6660  
 aatgtttgta attgactgaa tatcaacgct tatttgcagc ctgaatggcg aatgggacgc 6720  
 gccctgtagc ggcgcatcaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 6780  
 acttgccagc gccctagcgc ccgtctcttt cgttttcttc ccttcttttc tcgccaggtt 6840  
 cgccggcttt ccccgtaag ctctaaatcg ggggtctcct ttagggttcc gatttagtgc 6900

ttacggcac ctgacccca aaaaacttga ttagggatgat ggttcacgta gtgggccatc 6960  
 gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 7020  
 cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 7080  
 gattttgccg atttcggcct attgggttaa aaatgagctg atttaacaaa aatttaacgc 7140  
 gaattttaac aaaatatata cgtttacaat ttcaggtggc acttttcggg gaaatgtgcg 7200  
 cggaacccct atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca 7260  
 ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt 7320  
 ccggtgcgcc cttattccct tttttgcggc attttgcctt cctgtttttg ctcaccaga 7380  
 aacgctgggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga 7440  
 actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat 7500  
 gatgagcact tttaaagtgc tgctatgtgg cgcggtatta tcccgatttg acgccgggca 7560  
 agagcaactc ggtcgcgcga tacactatc tcagaatgac ttggttgagt actcaccagt 7620  
 cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 7680  
 catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 7740  
 aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcggt gggaaccgga 7800  
 gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 7860  
 aacgttgcc aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat 7920  
 agactggatg gaggcggata aagttgcagg accacttctg cgtcggccc ttcggctgg 7980  
 ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc 8040  
 actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc 8100  
 aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 8160  
 gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac ttcatTTTTA 8220  
 atttaaaagg atctaggtga agatcctttt tgataatctc atgacaaaaa tcccttaacg 8280  
 tgagttttcg ttccactgag cgtcagaccg cggggcatga ctaacatgag aattacaact 8340  
 tataatcgat ggggctgact tcaggtgcta catttgaaaga gataaattgc actgaaatct 8400  
 agaaaatatt tatctgatta ataagatgat cttcttgaga tcgttttggg ctgcgcgtaa 8460  
 tctcttgctc tgaacacgga aaaaaccgcc ttgcaggggc gtttttcgaa ggttctctga 8520  
 gctaccaact ctttgaaccg aggttaactgg cttggaggag cgcagtcacc aaaacttgct 8580  
 ctttcagttt agccttaacc ggcgcgatgac ttcaagacta actcctctaa atcaattacc 8640  
 agtggctgct gccagtgggt cttttgcatg tctttccggg ttggactcaa gacgatagtt 8700  
 accggataag gcgcagcggg cggactgaac ggggggttcg tgcatacagt ccagcttgga 8760  
 gcgaactgcc taccggaac tgagtgtcag gcgtggaatg agacaaacgc ggccataaca 8820  
 gcggaatgac accggtaaac cgaaaggcag gaacaggaga gcgcacgagg gagccgccag 8880  
 gggaaacgcc tggatatctt atagtccgtg cgggtttcgc caccactgat ttgagcgtca 8940  
 gatttcgtga tgcttgctag gggggcggag cctatggaaa aacggctttg ccgcggccct 9000  
 ctcacttccc tgtaagtat cttcctggca tcttcaggaa aatctccgcc ccgttcgtaa 9060  
 gccatttccg ctgcccgcag tcgaacgacc gagcgtagcg agtcagttag cgaggaagcg 9120  
 gaatatatcc tgtatcacat attctgtgta cgcaccgggt cagccttttt tctcctgcca 9180  
 catgaagcac ttactgaca ccctcatcag tgccaacata gtaagccagt atacactccg 9240

ctagcgctga tgtccggcgg tgcttttgcc gttacgcacc accccgtcag tagctgaaca 9300  
 ggaggacag ctgatagaaa cagaagccag ttctttcctg cgttatcccc tgattctgtg 9360  
 gataaccgta ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag 9420  
 cgcagcgagt cagtgagcga ggaagcggaa gagcgccctga tgcggtatct tctccttacg 9480  
 catctgtgcg gtatttcaca ccgcatagac cagccgcgta acctggcaaa atcggttacg 9540  
 gttgagtaat aaatg 9555

<210> 13

<211> 9208

<212> DNA

<213> 人工序列

<220>

<223>

<400> 13

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcacg 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgccgctcg tattaagag gggcgctggcc aagggcattg aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggcttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatacaag tgcatactt cttcccgat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac ttgttttag ggcgactgcc ctgctgcgta acatcgcttg tgctgcgtaa 840  
 catcggttgc gtcacataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccagg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatca cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacgtt cctaaggtag cgagttttaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat catcttgttg atagcgcgcg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440  
 ggcccctcga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560

gatcccgccg gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc 1620  
 tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgctc 1680  
 gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgctgccgt cctcgatgtt 1740  
 gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcggcca tgatatagac 1800  
 gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860  
 gtcgatgccc ttcagctcga tgcggttcac cagggtgtcg ccctcgaaact tcacctcggc 1920  
 gcgggtcttg tagttgccgt cgtccttgaa gaagatgggt cgctcctgga cgtagccttc 1980  
 gggcatggcg gacttgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040  
 ctgcacgccg taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccggt 2100  
 ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctgc cctcgccgga 2160  
 cacgtgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220  
 ggtgaacagc tcctcgccct tgetcatccc gggtgactgg aagtacaggt tttcttgca 2280  
 gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340  
 tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400  
 actaaccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460  
 tgtgaaataa ttgttttgt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520  
 atctctagtt gatactaag gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580  
 ccagaatatg ccatcagatg aagggaatgt gcaagaagt tctgacacca tgctagacat 2640  
 gctcttagtt tggaagttgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700  
 aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcgaataa ttatcatgtc 2760  
 cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820  
 tagtgattgg accaacagt ttccattatt tcttctaaac aagtgttgt ttgtaatgat 2880  
 gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940  
 atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaaaca agctctctcc 3000  
 ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060  
 ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120  
 accgttatct ggcttacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttaa acactagttc gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420  
 ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480  
 ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccttaggtac 3540  
 catgattccc ccacaggagg catccgctcg acggcgggag attgaggaca agctgaagca 3600  
 ggaggaggag actctgtcct tcatccgaga cagcctggag aagagcgacc agctcactaa 3660  
 gaacatgggtg tctatcttat catcctttga gaccgcctt atgaagctgg agaactccat 3720  
 catccctgtg cacaagcaga cggagaatct gcagcggctg caggagaatg ttgagaagac 3780  
 gctgtcctgc ctggaccatg tcatcagcta ctaccatgtg gccagtgaca ctgagaagat 3840  
 catcagagag ggccccacag gtaggctgga agagtacctg ggaagcatgg ccaagattca 3900

gaaggcagtg gagtatttcc aggacaacag cccagacagc ccggaactca acaaagtga 3960  
 actgctcttt gagcgcgga aggaggccct ggagtcgaa ttcgcagcc tgatgacgcg 4020  
 gcacagtaag gtcgtctcgc ccgtgctcat cttggatctg atcagtggtg acgatgatct 4080  
 ggaggcccag gaggacgtga ccctggagca cctgcccag agcgtgctcc aggatgtcat 4140  
 tcgcatctcc cgctggctgg tggaatatgg ccgcaacaa gatttcatga acgtctacta 4200  
 ccagatacgc tccagccagc tggaccgctc catcaaagga ctgaaggagc atttcataa 4260  
 gagcagttct tcctctgggg ttccctactc ccctgctatc cccaacaaga ggaaagacac 4320  
 acctaccaag aagccagtca agcggccagg gacgatccgt aaggctcaga accttctgaa 4380  
 acagtattcc cagcatggtc tagatgggaa aaaggggggc tctaacctca ttctctgga 4440  
 agggagagat gacatgctgg acgtggagac cgatgcctac atccactgcg tcagtgcctt 4500  
 cgtcaagctg gcgcagagcg agtaccagct gctggccgac atcatcccc agcaccacca 4560  
 gaagaagacc ttcgactccc tgatacagga tgccctggat gggctgatgc ttgaagggga 4620  
 gaacatcgtg tctgctgccc ggaaggccat tgtgcgacac gacttctcca cggctgctcac 4680  
 cgtcttcccc atcctgcgac acctcaagca gaccaagcct gagtttgacc aggtgctcca 4740  
 gggcacggct gccagcacia agaacaagct gcctggcctc atcacatcca tggagaccat 4800  
 cggtgccaaa gcgctggagg acttcgcaga caacatcaag aatgaccgga acaaggagta 4860  
 caacatgccg aaggacggca ccgtacacga gtcaccagc aatgccatcc tcttctgca 4920  
 gcagcttttg gacttccagg agacggcagg cgccatgctg gcctcccaag agaccagctc 4980  
 ttggccacc agctacagct ctgagttcag caagcggtg ctaagcacct atatctgtaa 5040  
 agtgctgggc aacctgcagt tgaacttgct gagcaagtcc aaggtgtacg aggaccagc 5100  
 tctgagcgcc atcttctgc acaacaacta caattacatc ctcaagtcct tggagaagtc 5160  
 tgaactgac cagctgggtg cagtgcaca gaagactgct gagcgctcct accgggagca 5220  
 cattgagcag cagatccaga cctaccagcg cagctgggta aaggtgactg attacatgc 5280  
 agagaagaat ctacctgtgt tccagccggg agtcaagctc cgggacaagg agcggcagat 5340  
 tatcaaggag cgttttaagg gcttcaatga tggcctcgaa gaactgtgca aaatccagaa 5400  
 ggcttgggct attccagaca cagagcagag ggacaggatt cgccaggccc agaagaccat 5460  
 tgtcaaggag acctacgggg ctttctaca gaagtttggc agcgtgccct tcaccaagaa 5520  
 cccggagaag tacatcaagt acggggtgga gcaggtgggc gacatgatcg atgcctttt 5580  
 cgacacctct gcctaactcg agttaactga ctaaaagctt cgaaaggaag ctgagttggc 5640  
 tgctgccacc gctgagcaat aactagcata accccttggg gcctctaaac gggctttgag 5700  
 gggttttttg ctgaaaggag gaactatcct cagggtcgag aagtactaga ggatcataat 5760  
 cagccatacc acattttag aggttttact tgctttaaaa aacctccac acctccccct 5820  
 gaacctgaaa cataaaatga atgcaattgt tgttgtaaac ttgtttattg cagcttataa 5880  
 tggttacaaa taaagcaata gcatcaciaa ttccacaaat aaagcatttt ttctactgca 5940  
 ttctagtgtt ggtttgtcca aactcatcaa tgtatcttat catgtctgga tctgatcact 6000  
 gcttgagcct aggagatccg aaccagataa gtgaaatcta gttccaaact attttgtcat 6060  
 ttttaatttt cgtattagct tacgacgcta caccagttc ccatctatct tgctactctt 6120  
 ccctaaataa tccttaaaaa ctccatttcc accctccca gttcccaact attttgtccg 6180  
 cccacagcgg ggcatttttc ttctgttat gtttttaate aaacatcctg ccaactccat 6240

gtgacaaacc gtcatcttcg gctacttttt ctctgtcaca gaatgaaaat ttttctgtca 6300  
 tctcttcgtt attaatgttt gtaattgact gaatatcaac gcttatttgc agcctgaatg 6360  
 gcgaatggga cgcgccctgt agcggcgcat taagcgcggc ggggtgtggtg gttacgcgca 6420  
 gcgtgaccgc tacacttgcc agcgccttag cgcccgtcc tttcgcttcc ttcccttcc 6480  
 ttctcgccac gttcgccggc ttccccgtc aagctctaaa tcggggggtc ctttaggggt 6540  
 tccgatttag tgctttacgg cacctcgacc caaaaaact tgattagggt gatggttcac 6600  
 gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttgag tccacgttct 6660  
 ttaatagtg actcttggtc caaactggaa caacactcaa ccctatctcg gtctattctt 6720  
 ttgatttata agggattttg ccgatttcgg cctattgggt aaaaaatgag ctgatttaac 6780  
 aaaaaattaa cgcgaatttt aacaaaatat taacgtttac aatttcaggt ggcacttttc 6840  
 ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca aatatgtatc 6900  
 cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaagg aagagtatga 6960  
 gtattcaaca tttccgtgtc gcccttatc ctttttttgc ggcattttgc ctctctgttt 7020  
 ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 7080  
 tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 7140  
 aacgttttcc aatgatgagc acttttaaaag ttctgctatg tggcgcggta ttatcccgt 7200  
 ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 7260  
 agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 7320  
 gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 7380  
 gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 7440  
 gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 7500  
 tagcaatggc aacaacgttg cgcaactat taactggcga actacttact ctagcttccc 7560  
 ggcaacaatt aatagactgg atggaggcgg ataaagtgtc aggaccactt ctgcgctcgg 7620  
 cccttccggc tggctggttt attgctgata aatctggagc cggtagcgt gggctctcgcg 7680  
 gtatcattgc agcactgggg ccagatggta agccctccc tatcgtagtt atctacacga 7740  
 cggggagtca ggcaactat gatgaacgaa atagacagat cgctgagata ggtgcctcac 7800  
 tgattaagca ttgtaactg tcagaccaag ttactcata tatacttttag attgatttaa 7860  
 aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 7920  
 aaatccctta acgtgagttt tcgttccact gacgctcaga ccgcggggca tgactaacat 7980  
 gagaattaca acttatatcg tatggggctg acttcaggtg ctacatttga agagataaat 8040  
 tgcactgaaa tctagaaata ttttatctga ttaataagat gatcttcttg agatcgtttt 8100  
 ggtctgcgcg taatctcttg ctctgaaaac ggaaaaaacc gccttgcagg gcggtttttc 8160  
 gaaggttctc tgagctacca actctttgaa ccgaggtaac tggcttgag gagcgcagtc 8220  
 accaaaactt gtcttttcag tttagcctta accggcgcat gacttcaaga ctaactctc 8280  
 taaatcaatt accagtggct gctgccagt gtgcttttgc atgtctttcc gggttggact 8340  
 caagacgata gttaccggat aaggcgcagc ggtcggactg aacggggggt tcgtgcatac 8400  
 agtccagctt ggagcgaact gcctaccgg aactgagtgt caggcgtgga atgagacaaa 8460  
 cgcggccata acagcggaat gacaccggt aaccgaaagg caggaacagg agagcgcacg 8520  
 agggagccgc caggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccaccact 8580

gatttgagcg tcagatttcg tgatgcttgt caggggggcg gagcctatgg aaaaacggct 8640  
 ttgccgcggc cctctcactt ccctgttaag tatcttcctg gcaccttcca ggaaatctcc 8700  
 gccccgttcg taagccattt ccgctcgccg cagtcgaacg accgagcgta gcgagtcagt 8760  
 gagcgaggaa gcggaatata tcctgtatca catattctgc tgacgcaccg gtgcagcctt 8820  
 ttttctcctg ccacatgaag cacttcactg acaccctcat cagtgcacaac atagtaagcc 8880  
 agtatacact ccgctagcgc tgatgtccgg cgggtgctttt gccgttacgc accaccccgt 8940  
 cagtagctga acaggaggga cagctgatag aaacagaagc cagttctttc ctgcgttatc 9000  
 ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctgcgcgcag 9060  
 ccgaacgacc gagcgagcgc agtcagttag cgaggaagcg gaagagcgcc tgatgcggta 9120  
 ttttctcctt acgcatctgt gcggtatttc acaccgcata gaccagccgc gtaacctggc 9180  
 aaaatcgggtt acggttgagt aataaatg 9208

<210> 14

<211> 9364

<212> DNA

<213> 人工序列

<220>

<223>

<400> 14

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcacg 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgcccgctg tattaaagag gggcgtggcc aagggcattg aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatatacaag tgcactcactt cttcccgat 360  
 gcccaacttt gtatagagag cactgcggg atcgtcaccc taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctgc cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggt cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcggttc tgctgcgtaa 840  
 catcgttgct gctccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900  
 atgcccgagg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgt 1020  
 acttgcatga cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcaccc 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260

```

tgacagcgtc agcggcaagt taatgcctat cttttgttg ataggcgcgg tcttgttttt 1320
aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380
tggcgcagca acgtcgccaa ctccattgt ttttctatg caaaacccca caccaacaac 1440
ggccccctga taataaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500
aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560
gatccccggcg gcggtcacga actccagcag gaccatgtga tcgcgcttct cgttggggtc 1620
tttgtcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgtc 1680
cccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgctgccgt cctcgatgtt 1740
gtggcggatc ttgaagtcca ccttgatgcc gttcttctgc ttgtcgcca tgatatagac 1800
gttgtggctg ttgtagttgt actccagctt gtgccccagg atgttgccgt cctccttgaa 1860
gtcgatgcc ttcagctcga tgcggttcac cagggtgtcg cctcgaact tcacctggc 1920
gcgggtcttg tagttgccgt cgtccttgaa gaagatggtg cgctcctgga cgtagccttc 1980
gggcatggcg gacttgaaga agtcgtgtg cttcatgtgg tcggggtagc ggctgaagca 2040
ctgcacgcc taggtcaggg tggtcacgag ggtgggccag ggcacgggca gcttgccgtt 2100
gggtcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctcgc cctcgccgga 2160
cacgtgaac ttgtggccgt ttacgtcgcc gtccagctcg accaggatgg gcaccacccc 2220
ggtaacagc tcctcgccct tgctcatccc gggtgactgg aagtacaggt tttcttgca 2280
gtacaccaat tcattcatga gttgagtcgc ttccttaact ggctgaaaag gctcttcagg 2340
tttcaccatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400
actaaccac tgctgcgcct cctgatttgt caacaattcc atgaagtttt tcggcacgct 2460
tgtgaaataa ttgttttgt tggtgaaatt cgatgctgag tgtataccaa caatgaacce 2520
atctctagtt gatactaatg gactgccaca ctgcccattc ttggtttgaa tccaatgctt 2580
ccagaatatg ccatcagatg aagggaatgt gcaagaagt tctgacacca tgctagacat 2640
gctcttagtt tggaagtgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700
aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcaataa ttatcatgtc 2760
cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820
tagtgattgg accaacagtg ttccattatt tcttctaaac aagtgttgt ttgtaatgat 2880
gaagggacca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940
atgacaaatg gtgctcgata tagggttgta atcacgggc cccttaaca agctctctcc 3000
ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060
ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120
accgttatct ggctacgtg acacaaggca cgttgtaga ttaataatct tatcttttta 3180
tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240
gctcagtcga cagatctgtc gacggtttaa acactagttc gcgacctact ccggaatatt 3300
aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360
ggtaaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggc 3420
ctggagccat ccgcaatttg aaaaaggtgg cgggtccggc ggaggtagcg gcggaggttc 3480
ttggtctcac cctcagttcg agaaggatga cgatgataaa accatgggat ccctaggtac 3540
ccttaagatg gcgatggcga tgtcggacag tggggcgagc cgcctgcgtc ggcagctgga 3600

```

gtcagggggt tttgaggcgc ggctgtacgt gaagcagctc tcgcagcagt cggatgggga 3660  
 cccggacctc caggagcacc ggcagcgcac ccaggcgcctg gcggaggaga cggcgagaa 3720  
 cctgaagcgc aacgtctacc agaactaccg gcagttcata gagacggccc gcgagatctc 3780  
 ctacctggag agcgagatgt accagctcag ccatttgctg accgagcaga aaagcagcct 3840  
 ggagagcatc ccgcttacgt tgctgcctgc cgctgctgcc gccggagccg ccgccgcctc 3900  
 tggaggggag gagggagtcg gtggggcggg gggccgagac cacctccgag gccaggccgg 3960  
 ctttttctcc acccccgggg gtgcctcccg cgacggctcc ggtccaggcg aggaaggaaa 4020  
 gcagcgcact ctaccaccc tgcttgagaa ggtggaaggc tgcaggcatc tgctggagac 4080  
 gccgggacag tacttggtgt acaatgggga cctagtggaa tacgatgcgg accacatggc 4140  
 ccaactgcag cgggtgcacg gctttctcat gaacgattgc ttgttggtgg ctacctggct 4200  
 gcctcagcgg cgtgggatgt atcgtacaa cgctctctat tccctagatg gtttgccgt 4260  
 agtcaatgtc aaggacaacc cgcccatgaa ggacatgttc aagctgctta tgttccccga 4320  
 gagccgtatt ttccaggccg aaaatgctaa aatcaaacga gagtggctgg aagtgcctga 4380  
 ggacaccaag agggccctca gtgagaaaag gcgaaggag caggaggagg cagcggcccc 4440  
 tcgagggcca cccaagtga cttccaaggc cactaacca tttgaggatg acgaagaaga 4500  
 agaaccagct gttcctgagg tagaggaaga gaaggtggac ctctccatgg aatggatcca 4560  
 ggagttacct gaagacctgg atgtctgcat tgcgcagaga gactttgaag gggcggttga 4620  
 cctgctggat aaattgaacc attacctgga agataaacct agcccacctc ctgtaaaaga 4680  
 actaagggcc aaagtggagg agcgagtctg acagctcact gaggtgctag ttttcgaact 4740  
 ctccccagat cgttcctga gaggtggtcc gaaggtact cgagagcag tttcgcaact 4800  
 gatccggctg ggccagtga cgaaggcctg tgagctatct ttgagaaaca gggcagccgc 4860  
 tgttcatact gcaattcgtc agcttcgcat cgaaggtgcc actttactct atattcataa 4920  
 gctgtgccat gtcttcttta ccagccttct cgagactgca agagaatttg agatcgattt 4980  
 tgcaggcact gacagcggct gctactctgc ctttgtggtc tgggcaagat cagccatggg 5040  
 catgttcgtg gatgcttita gcaagcaggt gtttgatagt aaggagagcc tctctacagc 5100  
 agctgagtgt gtaaaagtgg ctaaggagca ttgccagcaa ctgggtgata tcggactgga 5160  
 tctcaccttc atcatccatg ccttctggt gaaagacatc caaggggcct tgcacagtta 5220  
 caaagaaatc atcattgaag ccactaaaca tcgcaactct gaagagatgt ggaggaggat 5280  
 gaacttgatg acgccagaag ccctgggtaa gctcaaagaa gagatgaaaa gttgtggggt 5340  
 aagtaacttt gagcagtaca caggggatga ctgctgggtg aacctagtt acacagtgg 5400  
 tgctttcacc aaacagacca tgggtcttct ggaagaggcc ctgaagctgt atttcccaga 5460  
 gctgcacatg gtacttttgg agagcctggt ggaaatcatt ttggttgctg ttcagcatgt 5520  
 ggattatagt cttcgatgtg agcaggatcc agagaagaaa gcttttatca gacagaatgc 5580  
 atccttttta tatgaaacag tcctccctgt ggtggagaaa aggtttgaag aaggtgtggg 5640  
 gaaacctgcc aagcaactcc aagatctgag gaatgcact agacttattc gtgtgaatcc 5700  
 tgaaagtaca acatcagtgg tctaagaatt ccttaagcgg aggcctgcag ggctcgagtt 5760  
 aactgactaa aagcttcgaa aggaagctga gttggctgct gccaccgctg agcaataact 5820  
 agcataaccc cttggggcct ctaaagggt cttgaggggt tttttgctga aaggaggaa 5880  
 tatectcagg gtcgagaagt actagaggat cataatcage cataccacat ttgtagaggt 5940

tttacttgct ttaaaaaacc tcccacacct cccctgaac ctgaaacata aaatgaatgc 6000  
aattgttggt gttaacttgt ttattgcagc ttataatggt tacaataaaa gcaatagcat 6060  
cacaaatttc acaaataaag catTTTTTtT actgcattct agttgtgggt tgtccaaact 6120  
catcaatgta tcttatcatg tctggatctg atcactgctt gaggcctagga gatccgaacc 6180  
agataagtga aatctagtgc caaactatTT tgtcattttt aattttcgta ttagctttacg 6240  
acgctacacc cagttcccat ctattttgtc actcttccct aaataatcct taaaaactcc 6300  
atttccaccc ctcccagttc ccaactatTT tgtccgcccc cagcggggca tttttcttcc 6360  
tgttatgttt ttaatcaaac atcctgcccc ctccatgtga caaacgtca tcttcggcta 6420  
ctttttctct gtcacagaat gaaaattttt ctgtcatctc ttcgttatta atgtttgtaa 6480  
ttgactgaat atcaacgctt atttgcagcc tgaatggcga atgggacgcg ccctgtagcg 6540  
gcgcattaag cgcggcgggt gtggtgggtt cgcgcagcgt gaccgctaca cttgccagcg 6600  
ccctagcgcc cgtctcttcc gttttcttcc ctctcttctc cgccacgttc gccggcttcc 6660  
cccgtaagc tctaaatcgg gggtctctt tagggttccg atttagtgct ttacggcacc 6720  
tcgaccccaa aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga 6780  
cggtttttct ccttttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa 6840  
ctggaacaac actcaaccct atctcggctc attcttttga ttataaggg attttgccga 6900  
tttcggccta ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca 6960  
aaatattaac gtttacaatt tcaggtggca cttttcgggg aaatgtgcgc ggaaccctta 7020  
tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat 7080  
aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc 7140  
ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcaccagaa acgctggtga 7200  
aagtaaaaga tgctgaagat cagttgggtg cagcagtggt ttacatcgaa ctggatctca 7260  
acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt 7320  
ttaaagtctt gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg 7380  
gtcgccgcat acactattct cagaatgact tggtttagta ctcaccagtc acagaaaagc 7440  
atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata 7500  
acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt 7560  
tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag 7620  
ccataccaaa cgacgagcgt gacaccacga tgctgttagc aatggcaaca acgttgcgca 7680  
aactattaac tggcgaacta ctactctag ctcccgcca acaattaata gactggatgg 7740  
aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc tggttttattg 7800  
ctgataaaatc tggagccggt gagcgtgggt ctgcggtat cattgcagca ctggggccag 7860  
atggtaagcc ctcccgatc gtagttatct acacgacggg gactcaggca actatggatg 7920  
aacgaaaatg acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag 7980  
accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga 8040  
tctaggtgaa gatccttttt gataatctca tgacaaaaat cccttaacgt gagttttcgt 8100  
tccactgagc gtcagaccgc ggggcatgac taacatgaga attacaactt atatcgtatg 8160  
gggctgactt caggtgctac atttgaagag ataaattgca ctgaaatcta gaaatatttt 8220  
atctgattaa taagatgatc ttcttgagat cgttttggtc tgcgcgtaat ctcttgcctc 8280

gaaaacggaa aaaaccgcct tgcagggcgg tttttcgaag gttctctgag ctaccaactc 8340  
 tttgaaccga ggtaactggc ttggaggagc gcagtcacca aaacttgtcc tttcagttta 8400  
 gccttaaccg gcgcatgact tcaagactaa ctctctataa tcaattacca gtggctgctg 8460  
 ccagtgggtc ttttgcatgt ctttcggggt tggactcaag acgatagtta ccgataagg 8520  
 cgcagcggtc ggactgaacg ggggggttcgt gcatacagtc cagcttggag cgaactgcct 8580  
 acccggaact gagtgtcagg cgtggaatga gacaaacgcg gccataacag cggaatgaca 8640  
 ccggtaaacc gaaaggcagg aacaggagag cgcacgaggg agccgccagg ggaaacgcct 8700  
 ggtatcttta tagtcctgtc gggtttcgcc accactgatt tgagcgtcag atttcgtgat 8760  
 gcttgtcagg ggggcggagc ctatggaaaa acggctttgc cgcggccctc tcaattccct 8820  
 gttaagtatc ttctggcat cttccaggaa atctccgcc cgttcgtaag ccatttcgc 8880  
 tcgccgcagt cgaacgaccg agcgtagcga gtcagtgage gaggaagcgg aatatactct 8940  
 gtatcacata ttctgtgac gcaccgtgc agcctttttt ctctgccac atgaagcact 9000  
 tcaactgacac cctcatcagt gccaacatag taagccagta tacactccgc tagcgtgat 9060  
 gtccggcgggt gcttttgccg ttacgcacca ccccgtcagt agctgaacag gagggacagc 9120  
 tgatagaaac agaagccagt tctttctgc gttatccctt gattctgtgg ataaccgtat 9180  
 taccgccttt gagtgagctg ataccgtcg ccgcagccga acgaccgagc gcagcgagtc 9240  
 agtgagcgag gaagcggaag agcgctgat gcggtatfff ctccttacgc atctgtgcgg 9300  
 tatttcacac cgcatagacc agccgcgtaa cctggcaaaa tcggttacgg ttgagtaata 9360  
 aatg 9364

<210> 15

<211> 11728

<212> DNA

<213> 人工序列

<220>

<223>

<400> 15

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc tttaaactaga cagaatagtt gtaaactgaa atcagtccag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcatfctct 180  
 gaagtgcaaa ttgcccgtcg tattaagag gggcgtggcc aagggcattg aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggcttg 300  
 aacgaattgt taggtggcgg tacttgggtc gatataaag tgcatactt cttcccgat 360  
 gcccaacttt gtatagagag ccaactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcgatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgatggat ttgacttggc cagggccgag cctacatgtg cgaatgatgc ccatacttga 780

```

gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840
catcgttgct gtccataac atcaaacatc gaccacggc gtaacgcgct tgctgcttgg 900
atgcccagg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960
cgccgttacc accgctgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacgtc 1020
acttgcatta cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatcc 1080
gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140
atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagtttaaa 1200
cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260
tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgagg tcttgttttt 1320
aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380
tggcgcagca acgtcgccaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440
ggccccctga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500
aaatatatgt ataaggccgg ccttagtcag ttacttgtac agctcgtcca tgccgagagt 1560
gatcccgagg gcggtcacga actccagcag gaccatgtga tcgcgttct cgttggggtc 1620
tttgctcagg gcggactggg tgctcaggta gtggttgtcg ggcagcagca cggggccgtc 1680
gccgatgggg gtgttctgct ggtagtggtc ggcgagctgc acgtgccgt cctcgatgtt 1740
gtggcgatc ttgaagtta ccttgatgcc gttcttctgc ttgtcgcca tgatatagac 1800
gttggtgctg ttgtagttgt actccagctt gtgcccagg atgttgccgt cctccttgaa 1860
gtcgatgcc ttcagctga tgcggttcac cagggtgtcg ccctcgaact tcacctggc 1920
gcgggtcttg tagttgccgt cgtccttgaa gaagatggtg cgctcctgga cgtagccttc 1980
gggcatggcg gactgaaga agtcgtgctg cttcatgtgg tcggggtagc ggctgaagca 2040
ctgcacgcc taggtcaggg tggtcacgag gttgggccag ggcacgggca gcttgccgt 2100
ggtgcagatg aacttcaggg tcagcttgcc gtaggtggca tcgccctgc cctcgccga 2160
cacgtgaac ttgtggccgt ttacgtgcc gtccagctcg accaggatgg gcaccaccc 2220
ggtgaacagc tcctcgccct tgctcatccc ggtgactgg aagtacagg tttcttgca 2280
gtacaccaat tcattcatga gttagtcgc ttcttaact ggctgaaaag gctcttcagg 2340
tttcacatg aaaactttat ggcccccca caatactgag tcagcattta atcgccaacc 2400
actaacccac tgctgcgct cctgatttgt caacaattcc atgaagtttt tcggcacgtc 2460
tgtgaaataa ttgtttgtgt tggtgaaatt cgatgctgag tgtataccaa caatgaaccc 2520
atctctagtt gatactaatt gactgccaca ctgcccaccc ttggtttgaa tccaatgctt 2580
ccagaatatg ccatcagatg aagggaatgt gcaagaagt tctgacacca tgctagacat 2640
gctcttagtt tggaagtgg ttgtcacaag acatattctc tcttcccttt gtggctctct 2700
aaatttcagc ttttgaggaa atggtgggaa atccttaggc atgcaataa ttatcatgtc 2760
cctcccatca atgaggtgtt gttgcaaagt cgtggtgttc ttgacctga atacaccatg 2820
tagtgattgg accaacagt ttccattatt tcttctaaac aagtgttgt ttgtaatgat 2880
gaagggaaca aatccaatac catacaacga tgttgtgtgc ccatcagatt cattcgtcaa 2940
atgacaaatg gtgctcgata tagggttgta atcacgcggc cccttaacaa agctctctcc 3000
ggcgtagtca ggcacgtcgt aaggataagc catatttaaa tatatgcttg cttgtgtgtt 3060
ccttattgaa gccttggtgt gactgattta ctagtagcgt tgaggcgtct tatatacccg 3120

```

accgttatct ggcctacgtg acacaaggca cgttggttaga ttaataatct tatcttttta 3180  
 tcttaattga taagattatt tttatctggc tgttataaaa acgggatcat gaacacggac 3240  
 gctcagtcga cagatctgtc gacggtttaa acactagtct gcgacctact ccggaatatt 3300  
 aataggttgc tgatatcggg agttcagtcg tcgaatgcaa agcgtaaaaa atattaataa 3360  
 ggtaaaaatt acagctacat aaattacaca atttaaacgg atcgatgagc tccatatggg 3420  
 atccgcggag aacagcgaga gtctgggcac cgtccccgag cacgagcgga tcttgaggga 3480  
 gatcgagagc accgacaccg cctgtgtggg gccaccctc cggctctgtgt atgatgacca 3540  
 accaaatgcg cacaagaagt ttatggaaaa gttagatgct tgtatccgta atcatgacaa 3600  
 ggaaattgaa aagatgtgta attttcatca tcagggtttt gtagatgcta ttacagaact 3660  
 ccttaaagta aggactgatg cagaaaaact gaaggtgcaa gttactgata ccaaccgaag 3720  
 gtttcaagat gctggaaaag aggtgatagt ccacacagaa gatatcattc gatgtagaat 3780  
 tcagcagaga aatattacaa ctgtagtaga aaaattgcag ttatgccttc ctgtgctaga 3840  
 aatgtacagt aagctgaaag aacagatgag tgccaaaagg tactattctg ccctaaaaac 3900  
 tatggaacaa ttagagaatg tgtactttcc ctgggttagt caataccggt tttgtcagct 3960  
 catgatagaa aatcttccca aactccgtga ggatattaaa gaaatctcca tgtctgatct 4020  
 caaagacttt ttgaaaagta ttcgaaaaca ttctgacaaa ataggtgaaa cagcaatgaa 4080  
 acaggcacag catcagaaaa ccttcagtgt ttctctgcag aaacaaaata aaatgaaatt 4140  
 tgggaaaaat atgtatataa atcgtgatag aattccagag gaaaggaatg aaactgtatt 4200  
 gaaacattca cttgaagaag aggatgagaa tgaagaagag atcttaactg ttcaggatct 4260  
 tgttgatttt tcccctgttt atcgatgttt gcacatttat tctgttttgg gtgacgagga 4320  
 aacatttgaa aactattatc gaaaacaaag aaagaaacaa gcaagactgg tattgcaacc 4380  
 ccagtcgaat atgcatgaaa cagttgatgg ctatagaaga tatttcactc aaattgtagg 4440  
 gttctttgtg gtagaagatc acattttaca tgtgacccaa ggattagtaa ccagggcata 4500  
 cactgatgaa ctttggaaac tggccctctc aaagataatt gctgtcctta gagctcattc 4560  
 atcctattgc actgatcctg atcttgttct ggagctgaag aatcttattg taatatttgc 4620  
 agatacttta cagggttatg gttttccagt gaaccgactt tttgaccttt tatttgaaat 4680  
 aagagaccaa tacaatgaaa cactgcttaa gaaatgggct ggagttttca gggacatttt 4740  
 tgaagaagat aattacagcc ccacccctgt tgtcaatgaa gaagaatata aaattgtcat 4800  
 cagcaaatTT cctttcaag atccagacct tgaaaagcag tctttcccaa agaaattccc 4860  
 catgtctcag tcagtgcctc atattttacat tcaagttaaa gaattttatt atgccagcct 4920  
 taaattttca gagtcactac accggagctc aacagaaata gacgatatgc ttagaaaatc 4980  
 aacaaatctg ctgtgacca gaactttgag tagctgttta ctgaacctta ttagaaaacc 5040  
 tcatataggt ttgacagagc tggtaacaa catcataaac acaacacacc tggagcaagc 5100  
 ttgtaaatat cttgaggact ttataactaa cattacaaat atttcccaag aaactgttca 5160  
 tactacaaga ctttatggac tttctacttt caaggatgct cgacatgcag cagaaggaga 5220  
 aatatatacc aaactgaatc aaaaaattga tgaatttgtt cagcttgctg attatgactg 5280  
 gacaatgtct gagccagatg gaagagctag tggttattta atggacctta taaatttttt 5340  
 gagaagcatc tttcaagtgt ttactcattt gcctgggaaa gttgctcaga cagcttgcat 5400  
 gtcagcctgc cagcatctgt caacatcctt aatgcagatg ctactggaca gtgagttaaa 5460

acaaataagc atgggagctg ttcagcagtt taacttagat gtcatacagt gtgaattggt 5520  
 tgccagctct gagcctgtgc caggattcca gggggatacc ctgcagctag cattcattga 5580  
 cctcagacaa ctccctgacc tgtttatggg ttgggattgg tctacttacc tagctgatta 5640  
 tgggcagcca gcttctaagt accttcgggt gaatccaaac acagccctta ctcttttgga 5700  
 gaagatgaag gatactagca aaaagaacaa tatatttgct cagttcagga agaattgatcg 5760  
 agacaaacag aagttgatag agacagtcgt gaaacagctg agaagtttggt tgaatggtat 5820  
 gtcccagcac atgggctcga gtgagaatct gtatttccag agcgggtaccg cggcccgcat 5880  
 gaaggagaca gaccgggagg ccgttgcgac agcagtgcaa agggttgctg ggatgctcca 5940  
 gcgcccggac cagctggaca aggtggagca gtatcgcagg agagaagcgc ggaagaaggc 6000  
 ctccgtggag gccagattga aggccgcat ccagtcacag ttggacgggg tgcgcacagg 6060  
 cctcagccag ctccacaacg cctgaatga cgtcaaagac atccagcagt cgctggcaga 6120  
 cgtcagcaag gactggagge agagcatcaa caccattgag agcctcaagg acgtcaaaga 6180  
 cgccgtgggtg cagcacagcc agtcgcccgc agccgtggag aacctcaaga acatcttctc 6240  
 agtgcctgag attgtgaggg agaccagga cctaattgaa caaggggcac tcttgcaagc 6300  
 ccaccggaag ctgatggacc tggaatgctc ccgggacggg ctgatgtacg agcagtaccg 6360  
 catggacagt gggaacacgc gtgacatgac cctcatccat ggctactttg gcagcacgca 6420  
 ggggctctct gatgagctgg ctaagcagct gtggatgggt ctgcagaggt cactggtcac 6480  
 tgtccgccgt gacccacct tgctggtctc agttgtcagg atcattgaaa gggaagagaa 6540  
 aattgacagg cgcatacttg accggaaaaa gcaaaactggc tttgttctc ctgggaggcc 6600  
 caagaattgg aaggagaaaa tgttcacat cttggagagg actgtgacca ccagaattga 6660  
 gggcacacag gcagatacca gagagtctga caagatgtgg cttgtccgcc acctggaaat 6720  
 tataaggaag tacgtcctgg atgacctcat tgtcgccaaa aacctgatgg ttcagtgtt 6780  
 tctcccccac tatgagatct ttaagaacct cctgaacatg taccaccaag ccctgagcac 6840  
 gcggatgcag gacctcgcat cggaagacct ggaagccaat gagatcgtga gcctcttgac 6900  
 gtgggtctta aacacctaca caagtactga gatgatgagg aacgtggagc tggccccgga 6960  
 agtggatgtc ggcaccctgg agccattgct ttctccacac gtggtctctg agctgcttga 7020  
 cacgtacatg tccacgtca cttcaaakat catgcctgg ctgcggaaag cgctggagac 7080  
 agacaagaaa gactgggtca aagagacaga gccagaagcc gaccaggacg ggtactacca 7140  
 gaccacactc cctgccattg tcttccagat gtttgaacag aatcttcaag ttgctgtca 7200  
 gataagtga gatttgaaaa caaaggtact agttttatgt cttcagcaga tgaattcttt 7260  
 cctaagcaga tataaagatg aagcgcagct gtataaagaa gagcacctga ggaatcggca 7320  
 gcaccctcac tgctacgttc agtacatgat cgccatcatc aacaactgcc agaccttcaa 7380  
 ggaatccata gtcagtttaa aaagaaagta tttaaagaat gaagtggaaagg aggggtgtgtc 7440  
 tccgagccag cccagcatgg acgggatttt agacgccatc gcgaaggagg gctgcagcgg 7500  
 tttgctggag gaggtcttcc tggacctgga gcaacatctg aatgaattga tgacgaagaa 7560  
 gtggctatta gggtaaaacg ctgtagacat tatctgtgtc accgtggaag actatttcaa 7620  
 cgattttgcc aaaattaaaa agccgtataa gaagaggatg acggccgagg cgcaccggcg 7680  
 cgtggtgggtg gactacctgc gggcggtcat gcagaagcgc atttcttcc ggagcccgga 7740  
 ggagcgcaag gaggtgtccg agaagatggt tagggaggca gagcagctgc gcttctgtt 7800

ccggaagctg gcgtccggtt tcggggaaga cgtggacgga tactgcgaca ccatcgtggc 7860  
 tgtggccgaa gtgatcaagc tgacagaccc ttctctgctc tacctggagg tctccactct 7920  
 ggtcagcaag tatccagaca tcagggatga ccacatcggt gcgctgctgg ctgtgcgtgg 7980  
 ggacgccagc cgtgacatga agcagaccat catggagacc ctggagcagg gccagcaca 8040  
 ggccagcccc agctacgtgc cctcttcaa ggacattgtg gtgccagcc tgaacgtggc 8100  
 caagctgctc aagtaactcg agttaactga ctaaaagctt cgaaaggaag ctgagttggc 8160  
 tgctgccacc gctgagcaat aactagcata accccttggg gcctctaaac gggctttgag 8220  
 gggttttttg ctgaaaggag gaactatcct cagggtcgag aagtactaga ggatcataat 8280  
 cagccatacc acattttagt aggttttact tgctttaaaa aacctccac acctccccct 8340  
 gaacctgaaa cataaaatga atgcaattgt tgttgtaac ttgtttattg cagcttataa 8400  
 tggttacaaa taaagcaata gcatcaciaa ttccacaaat aaagcatttt ttctactgca 8460  
 ttctagtgtt ggtttgtcca aatcatcaa tgtatcttat catgtctgga tctgatcact 8520  
 gcttgagcct aggagatccg aaccagataa gtgaaatcta gttccaaact attttgtcat 8580  
 ttttaatttt cgtattagct tacgacgcta caccagttc ccatctattt tgtcactctt 8640  
 ccctaataaa tccttaaaaa ctccatttcc accctccca gttcccaact attttgtccg 8700  
 cccacagcgg ggcatttttc ttctgttat gtttttaac aaacatcctg ccaactccat 8760  
 gtgacaaacc gtcacttctg gctacttttt ctctgtcaca gaatgaaaat tttctgtca 8820  
 tctcttcgtt attaatgttt gtaattgact gaatatcaac gcttatttgc agcctgaatg 8880  
 gcgaatggga cgcgccctgt agcggcgcgt taagcgcggc ggggtgtggtg gttacgcgca 8940  
 gcgtgaccgc tacacttgcc agcgccttag cgcccgtcc tttcgcttcc ttccttctct 9000  
 ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc ctttaggggt 9060  
 tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac 9120  
 gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttgag tccacgttct 9180  
 ttaatagtgg actcttgttc caaactggaa caacactcaa ccctatctcg gtctattctt 9240  
 ttgatttata agggattttg ccgatttcgg cctatttggt aaaaaatgag ctgatttaac 9300  
 aaaaatttaa cgcgaatttt aacaaaatat taacgtttac aatttcagggt ggcacttttc 9360  
 ggggaaatgt gcgcggaacc cctatttggt tatttttcta aatacattca aatatgtatc 9420  
 cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 9480  
 gtattcaaca tttccgtgtc gcccttattc ctttttttgc ggcattttgc ctctctgttt 9540  
 ttgctcacc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 9600  
 tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 9660  
 aacgttttcc aatgatgagc acttttaaaag ttctgctatg tggcgcggta ttatcccgta 9720  
 ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 9780  
 agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 9840  
 gtgctgcat aaccatgagt gataaactg cggccaaact acttctgaca acgatcgag 9900  
 gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgcttgatc 9960  
 gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 10020  
 tagcaatggc aacaacgttg cgaaactat taactggcga actacttact ctagcttccc 10080  
 ggcaacaatt aatagactgg atggaggcgg ataaagtgc aggaccactt ctgcgctcgg 10140

cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggctctcgcg 10200  
 gtatcattgc agcactgggg ccagatggta agccctcccc tatcgtagtt atctacacga 10260  
 cggggagtcg ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 10320  
 tgattaagca ttgtaactg tcagaccaag tttactcata tatacttttag attgatttaa 10380  
 aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 10440  
 aaatccctta acgtgagttt tcgttccact gagcgtcaga ccgcggggca tgactaacat 10500  
 gagaattaca acttatatcg tatggggctg acttcaggtg ctacatttga agagataaat 10560  
 tgcaactgaaa tctagaaaata ttttatctga ttaataagat gatcttcttg agatcgtttt 10620  
 ggtctgcgcg taatctcttg ctctgaaaac ggaaaaaacc gccttgcagg gcggtttttc 10680  
 gaaggttctc tgagctacca actctttgaa ccgaggtaac tggcttggag gagcgcagtc 10740  
 accaaaaactt gtcttttcag tttagcetta accggcgcat gacttcaaga ctaactctc 10800  
 taaatcaatt accagtgget gtgccagtgt gtgcttttgc atgtctttcc gggttggact 10860  
 caagacgata gttaccggat aaggcgcagc ggtcggactg aacgggggggt tcgtgcatac 10920  
 agtccagctt ggagcgaaact gcctaccggg aactgagtgt caggcgtgga atgagacaaa 10980  
 cgcgggcata acagcggaat gacaccggtg aaccgaaagg caggaacagg agagcgcacg 11040  
 agggagccgc caggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccaccact 11100  
 gatttgagcg tcagatttcg tgatgcttgt caggggggcg gaggcctatgg aaaaacggct 11160  
 ttgccgcggc cctctcactt ccctgttaag tatcttcttg gcatcttcca ggaaatctcc 11220  
 gccccgttcg taagccattt ccgctcgccg cagtcgaacg accgagcgta gcgagtcagt 11280  
 gagcgaggaa gcggaatata tcctgtatca catattctgc tgacgcaccg gtgcagcctt 11340  
 ttttctcctg ccacatgaag cacttcaactg acaccctcat cagtgccaac atagtaagcc 11400  
 agtatacact ccgctagcgc tgatgtccgg cgggtgctttt gccgttacgc accaccccg 11460  
 cagtagctga acaggaggga cagctgatag aaacagaagc cagttctttc ctgcgttatc 11520  
 ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctcgccgcag 11580  
 ccgaacgacc gagcgcagcg agtcagttag cgaggaagcg gaagagcgcc tgatgcggta 11640  
 ttttctcctt acgcatctgt gcggtatttc acaccgcata gaccagccgc gtaacctggc 11700  
 aaaatcggtt acggttgagt aataaatg 11728

<210> 16

<211> 12046

<212> DNA

<213> 人工序列

<220>

<223>

<400> 16

gatgccctgc gtaagcgggt gtgggcggac aataaagtct taaactgaac aaaatagatc 60  
 taaactatga caataaagtc ttaaactaga cagaatagtt gtaaactgaa atcagtcag 120  
 ttatgctgtg aaaaagcata ctggactttt gttatggcta aagcaaactc ttcattttct 180  
 gaagtgcaaa ttgccgctcg tattaagag gggcgtggcc aaggcatgt aaagactata 240  
 ttcgcggcgt tgtgacaatt taccgaacaa ctccgcggcc gggaagccga tctcggttg 300

aacgaattgt taggtggcgg tacttgggtc gatatcaaag tgcatacactt cttcccgtat 360  
 gcccaacttt gtatagagag ccactgcggg atcgtcaccg taatctgctt gcacgtagat 420  
 cacataagca ccaagcgcgt tggcctcatg cttgaggaga ttgatgagcg cgggtggcaat 480  
 gccctgcctc cgggtgctcg cggagactgc gagatcatag atatagatct cactacgcgg 540  
 ctgctcaaac ttgggcagaa cgtaagccgc gagagcgcca acaaccgctt cttggtcgaa 600  
 ggcagcaagc gcatgaatg tcttactacg gagcaagttc ccgaggtaat cggagtccgg 660  
 ctgatgttgg gagtaggtgg ctacgtctcc gaactcacga ccgaaaagat caagagcagc 720  
 ccgcatggat ttgacttggg cagggccgag cctacatgtg cgaatgatgc ccatacttga 780  
 gccacctaac tttgttttag ggcgactgcc ctgctgcgta acatcgttgc tgctgcgtaa 840  
 catcgttgct gctccataac atcaaacatc gacccacggc gtaacgcgct tgctgcttgg 900  
 atgcccaggg catagactgt acaaaaaaac agtcataaca agccatgaaa accgccactg 960  
 cgccgttacc accgetgcgt tcggtcaagg ttctggacca gttgcgtgag cgcatacget 1020  
 acttgcatta cagtttacga accgaacagg cttatgtcaa ctgggttcgt gccttcatec 1080  
 gtttccacgg tgtgcgtcac ccggcaacct tgggcagcag cgaagtcgcc ataacttcgt 1140  
 atagcataca ttatacgaag ttatctgtaa ctataacggt cctaaggtag cgagttttaa 1200  
 cgacgtccta gattggttac tgggcgatga aggttttagtc ggcaaatcgt ccaacgttag 1260  
 tgacagcgtc agcggcaagt taatgcctat cattttgttg ataggcgcgg tcttgttttt 1320  
 aggtttaata ttttatttta tctacagata catgatgaaa ggagggaagg gaggtggtgt 1380  
 tggcgcagca acgtcgcaa ctcccattgt tatttctatg caaaaccca caccaacaac 1440  
 ggccccctga taataaaaaga caaaaataat ataaaatata tgtataatta attaaattca 1500  
 aaatatatgt ataaggccgg ccttagtcag ttaattaagt ttgtgcccc gtttgctagg 1560  
 gaggtcgcag tatctggcca cagccacctc gtgctgctcg acgtaggtct ctttgctggc 1620  
 ctccctgatt ctttccagtc tgttggtccac atagtagacg ccgggcatct tgaggttctt 1680  
 agcgggtttc ttggatctgt atgtggtctt gaagttgcag atcaggtggc ccccgccac 1740  
 gagcttcagg gccatgtcgc ttctgccttc caggccgccg tcagcggggt acagcatctc 1800  
 ggtgttggcc tcccagccga gtgttttctt ctgcatcaca gggccgttgg atgggaagtt 1860  
 caccctctg atcttgacgt tgtagatgag gcagccgtcc tggaggctgg tgctctgggt 1920  
 agcggtcagc acgccccgt cttcgtatgt ggtgactctc tccatgtga agccctcagg 1980  
 gaaggactgc ttaaagaagt cggggatgcc ctgggtgtgg ttgatgaagg ttctgctgcc 2040  
 gtacatgaag ctggtagcca ggatgtcgaa ggcaagggg agagggccgc cctcgaccac 2100  
 cttgattctc atggtctggg tgccctcgta gggcttgctt tcgcccctcg atgtgcactt 2160  
 gaagtgggtg ttgttcacgg tgccctccat gtacagcttc atgtgcatgt tctccttaat 2220  
 cagctcttcg cccttagaca ccatcccggg tgactggaag tacaggtttt cttgcgagta 2280  
 caccaattca ttcatgagtt gagtcgcttc cttactggc tgaagggtt cttcaggttt 2340  
 caccatgaaa actttatggc cccccacaa tactgagtca gcatttaate gccaccact 2400  
 aaccactgc tgcgctcctt gatttgtaaa caattccatg aagtttttcg gcacgcttgt 2460  
 gaaataattg tttgtgttgg tgaaattcga tgctgagtgt ataccaacaa tgaaccatc 2520  
 tctagttagt actaatggac tgccacactg cccatccttg gtttgaatcc aatgcttcca 2580  
 gaatatgcca tcagatgaag ggaatgtgca agaagtgtct gacaccatgc tagacatgct 2640

cttagtttgg aagttggttg tcacaagaca tattctctct tccctttgtg gctctctaaa 2700  
 tttagcgttt tgaggaaatg gtgggaaatc ctaggcatg cgaataatta tcatgtccct 2760  
 cccatcaatg aggtgttggt gcaaagtcgt ggtgttcttg acctgaata caccatgtag 2820  
 tgattggacc aacagtgttc cattatttct tctaaacaag tgcttgtttg taatgatgaa 2880  
 gggaccaaatt ccaataccat acaacgatgt tgtgtgcccc tcagattcat tcgtcaaattg 2940  
 acaaatggtg ctcgatatag ggttgtaatc acgcggcccc ttaaacaagc tctctccggc 3000  
 gtagtcaggc acgtcgtgag gataagccat atttaaatat atgcttgctt gtgtgttcct 3060  
 tattgaagcc ttggtgtgac tgatttacta gtagcgttga ggcgtcttat ataccgacc 3120  
 gttatctggc ctacgtgaca caaggcacgt tgttagatta ataactttat ctttttatct 3180  
 taattgataa gattatTTTT atctggctgt tataaaaacg ggatcatgaa cacggacgct 3240  
 cagtcgacag atctgtcgac ggtttaaaca ctagttcgcg acctactccg gaatattaat 3300  
 aggttgctga tatcgggagt tcagtcgtcg aatgcaaagc gtaaaaaata ttaataaggt 3360  
 aaaaaattaca gctacataaa ttacacaatt taaacggatc gatgagctcc atatggcggc 3420  
 cgcggaagca gctggtggga aatacagaag cacagtcagc aaaagcaaag acccctcggg 3480  
 gctgctcatc tctgtgatca ggactctgtc tactagtac gatgtcgaag acagggaata 3540  
 tgaaaagggc cgccttgag aagcctacga gaaatgtgac cgtgacctgg atgaattgat 3600  
 tgtacagcac tacacagaat tgacgacagc cattcgcaca taccagagca tcacagagcg 3660  
 catcactaac tcccgaata aaataaagca ggtaaaagag aacctgcttt catgcaagat 3720  
 gctgctgcac tgcaaacggg atgagcttcg gaaactgtgg attgaaggaa ttgagcataa 3780  
 gcatgtcctg aacttggttg atgaaattga gaatatcaag caagtgcctc aaaagctgga 3840  
 acagtgcag gccagcaagc actatctcag tgccactgac atgttggtgt cagcagttga 3900  
 gtctttggag ggccccctgc tccaggtgga aggactgagt gaccttcgac tagagcttca 3960  
 cagcaagaag atgaacctc acttggttct catagatgaa ctacaccggc acctgtacat 4020  
 caaatcgact agccgagttg tgcagcgtaa caaggaaaaa gggaaaatca gctccctcgt 4080  
 gaaagatgct tctgttcctc tgattgatgt tacaacctc cctactcctc gaaaattcct 4140  
 tgatacctct cactattcta ctgctggaag ctcaagtgtg agggagataa atctgcagga 4200  
 catcaaggaa gatttagaat tggatccaga ggaaaacagc accctgttta tgggtatcct 4260  
 cattaagggc ttggcgaaac tgaagaagat ccagaaaca gttaaggcaa tcatagagcg 4320  
 cttggagcag gaggttgaagc aaattgtgaa gaggtctaca acccaggtgg cagacagtgg 4380  
 ctatcagcgg ggggagaacg ttactgtgga gaaccaacca aggttgcttc tagaactgct 4440  
 ggagttactg ttgacaagt ttaatgtgt agccgctgca cactctgtgg tcctgggata 4500  
 cctgcaggac actgtagtga ctccactgac tcagcaggaa gatataaac tgtatgatat 4560  
 ggagatgta tgggtgaaga tccaagatgt tctacagatg ctattaactg agtacttgga 4620  
 tatgaaaaat actcgtacgg cctctgaacc atcagctcaa ctaagctatg ccagcactgg 4680  
 acgagagttt gcagcctttt ttgccaagaa gaaacctcaa aggccaaaaa attctctttt 4740  
 caagttcgaa tcgtcctccc atgccatcag tatgagcgcc tatctgcgag aacagagaag 4800  
 ggagctctat agtcggagtg gagaactgca aggggtcct gatgacaact taattgaagg 4860  
 tggaggaaca aaattgtct gcaaacctgg agccagaaac attaccgtca tattccacc 4920  
 attactaaga ttatttcagg agattgagca tgctctgggt cttggcccag ccaaacagt 4980

tcctcttcga gagtttctca ccgtgtacat caaaaacatc tttctcaatc aagtcttggc 5040  
 tgagatcaac aaggagattg aaggagtcac taaaacatct gaccctttga agattctggc 5100  
 caacgcagac accatgaagg tgctgggagt gcagcggcct ctctacaga gcacaatcat 5160  
 tgtggagaag acagttcaag acctcctgaa cctgatgcat gacttgagtg catattcaga 5220  
 tcaattcctc aacatgggtg gcgtgaagct ccaggagtac aaggacacct gcaactgcagc 5280  
 ttacaggggt attgtccagt cagaagaaaa acttgtcatc agtgcatacct gggcaaaaaga 5340  
 tgatgatatc agcagactct tgaaatctct accaaactgg atgaatatgg ctcaacccaa 5400  
 acagctgagg ccaaaaagag aggaggaaga agatttcata agggcagctt ttggcaagga 5460  
 gtctgaagtt cttattggga acctgggtga taaattaatc cctccacaag acatccttcg 5520  
 tgacgtcagt gacctcaaag ccttggccaa catgcatgaa agcctggaat ggttggcaag 5580  
 tcgaacaaaag tcagctttct ccaatcttct tacatcccag atgctttctc ctgctcaaga 5640  
 cagccacacg aacacggatc tccccccagt gtcagagcag atcatgcaga ctctcagtga 5700  
 acttgccaaa tcgttccagg atatggetga ccgctgcttg cttgtcttac atctggaagt 5760  
 gagggttcac tgtttccact atcttatccc tcttgcaaaag gaggggaact atgccattgt 5820  
 ggctaattgt gaaagtatgg attatgacct cctggtggtc aagctcaaca aagatatcag 5880  
 cgccattgaa gaggccatga gcgccagcct tcagcagcac aagttccagt atatcttcga 5940  
 aggctgggc cacctgatct cctgcatacct cattaatggt gccagctact tcaggcgcag 6000  
 cagtgagtct ggcatcaaga aaatgtgtag gaacattttt gttcttcagc agaatttgac 6060  
 caacatcacc atgtcgcggg aggcagacct ggactttgca aggcagctact acgagatgct 6120  
 ttacaacaca gctgacgagc tcctgaacct ggtggtggac caggggtgtga agtacacgga 6180  
 gctggagtac atccacgctc tgacctgct gcaccgcagc cagactgggg tgggggaact 6240  
 gaccaccag aacacgaggc tgcaaggct caaagagatc atctgcgagc aggctgcat 6300  
 caagcaagcc accaaggaca agaagataac taccgttggc tcgagtgaga acttgtactt 6360  
 tcagtcctta agcggtagca tgattcccc acaggaggca tccgctcgac ggcgggagat 6420  
 tgaggacaag ctgaagcagg aggaggagac tctgtccttc atccgagaca gcctggagaa 6480  
 gagcgaccag ctactaaga acatggtgtc tatcttatca tcctttgaga gccgccttat 6540  
 gaagctggag aactccatca tcctgtgca caagcagacg gagaatctgc agcggctgca 6600  
 ggagaatgtt gagaagacgc tgtcctgcct ggacctgtc atcagctact accatgtggc 6660  
 cagtgcact gagaagatca tcagagaggg cccacaggt aggctggaag agtacctggg 6720  
 aagcatggcc aagattcaga aggcagtgga gtatttccag gacaacagcc cagacagccc 6780  
 ggaaactcaac aaagtgaaac tgctctttga gcgcgggaag gaggccttg agtccgaatt 6840  
 tcgcagcctg atgacgcggc acagtaaggt cgtctcgccc gtgctcatct tggatctgat 6900  
 cagtggtgac gatgatctgg aggccagga ggacgtgacc ctggagcacc tgcccagagag 6960  
 cgtgctccag gatgtcatc gcatactccc ctggctgggtg gaatatggcc gcaaccaaga 7020  
 tttcatgaac gtctactacc agatacgtc cagccagctg gaccgtcca tcaaaggact 7080  
 gaaggagcat ttccataaga gcagttcttc ctctgggggtt ccctactccc ctgctatccc 7140  
 caacaagagg aaagacacac ctaccaagaa gccagtcaag cgccagggga cgatccgtaa 7200  
 ggctcagaac cttctgaaac agtattccca gcatggtcta gatgggaaaa aggggggctc 7260  
 taacctcatt cctctggaag ggagagatga catgctggac gtggagaccg atgcctacat 7320

ccactgcgtc agtgccttcg tcaagctggc gcagagcgag taccagctgc tggccgacat 7380  
 catccccgag caccaccaga agaagacctt cgactccctg atacaggatg ccctggatgg 7440  
 gctgatgctt gaaggggaga acatcgtgtc tgctgcccgg aaggccattg tgcgacacga 7500  
 cttctccacg gtgtccaccg tcttcccat cctgcgacac ctcaagcaga ccaagcctga 7560  
 gtttgaccag gtgtccagg gcacggctgc cagcacaag aacaagctgc ctggcctcat 7620  
 cacatccatg gagaccatcg gtgcaaagc gctggaggac ttgcgagaca acatcaagaa 7680  
 tgaccgggac aaggagtaca acatgccgaa ggacggcacc gtacacgagc tcaccagcaa 7740  
 tgccatcctc ttctgcagc agcttttggg cttccaggag acggcaggcg ccatgctggc 7800  
 ctcccaagag accagctctt cggccaccag ctacagctct gagttcagca agcggctgct 7860  
 aagcacctat atctgtaaag tgctgggcaa cctgcagttg aacttgctga gcaagtccaa 7920  
 ggtgtacgag gaccagctc tgagcgccat cttctgcac aacaactaca attacatcct 7980  
 caagtccctg gagaagtctg aactgatcca gctggtggca gtgacacaga agactgctga 8040  
 gcgctcctac cgggagcaca ttgagcagca gatccagacc taccagcgca gctggttaaa 8100  
 ggtgactgat tacatcgagc agaagaatct acctgtgttc cagccgggag tcaagctccg 8160  
 ggacaaggag cggcagatta tcaaggagcg ttttaagggc ttcaatgatg gcctcgaaga 8220  
 actgtgcaaa atccagaagg cctgggctat tccagacaca gagcagaggg acaggattcg 8280  
 ccaggcccag aagaccattg tcaaggagac ctacggggcc tttctacaga agtttggcag 8340  
 cgtgcccttc accaagaacc cggagaagta catcaagtac ggggtggagc aggtgggcga 8400  
 catgatcgat cgcttttctg acacctctgc ctaactcgag ttaactgact aaaagcttcg 8460  
 aaaggaagct gagttggctg ctgccaccgc tgagcaataa ctagcataac cccttggggc 8520  
 ctctaaacgg gtcttgaggg gttttttgct gaaaggagga actatcctca gggctcgagaa 8580  
 gtactagagg atcataatca gccataccac attttagag gttttacttg ctttaaaaaa 8640  
 cctccacac ctccccctga acctgaaaca taaaatgaat gcaattgttg ttgttaactt 8700  
 gtttattgca gcttataatg gttacaaata aagcaatagc atcacaatt tcacaaataa 8760  
 agcatttttt tcactgcatt ctagtgttg tttgtccaa ctcataatg tatcttatca 8820  
 tgtctggatc tgatcactgc ttgagcctag gagatccgaa ccagataagt gaaatctagt 8880  
 tccaaactat tttgtcattt ttaattttcg tattagetta cgacgtaca ccagttccc 8940  
 atctattttg tcactcttcc cttaaataatc cttaaaaact ccatttccac cctcccagt 9000  
 tcccaactat tttgtccgcc cacagcgggg catttttctt cctgttatgt ttttaatcaa 9060  
 acatcctgcc aactccatgt gacaaaccgt catcttcggc tactttttct ctgtcacaga 9120  
 atgaaaattt ttctgtcatc tcttcgttat taatgtttgt aattgactga atatcaacgc 9180  
 ttatttgag cctgaatggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg 9240  
 gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgtcctt 9300  
 tcgctttctt cccttcttt ctcgccacgt tcgcggctt tccccgtaa gctetaaatc 9360  
 gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg 9420  
 attagggtga tggttcacgt agtgggcat cgccctgata gacggttttt cgccctttga 9480  
 cgttgagtc cacgttcttt aatagtggac tcttgttcca aactggaaca aactcaacc 9540  
 ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa 9600  
 aaaatgagct gatttaacaa aaatttaacg cgaattttta caaatatta acgtttacaa 9660

```

tttcagggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 9720
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 9780
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg 9840
cattttgcct tcctgttttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag 9900
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg 9960
agagttttcg ccccgaaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 10020
gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt 10080
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcactttacg gatggcatga 10140
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac 10200
ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatc 10260
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 10320
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaa 10380
tacttactct agcttcccgg caacaattaa tagactggat ggagcgcat aaagttgcag 10440
gaccacttct gcgctcgcc cttccggctg gctggtttat tgctgataaa tctggagccg 10500
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgt 10560
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 10620
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 10680
tacttttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt 10740
ttgataatct catgacaaaa atcccttaac gtgagtttct gttccactga gcgtcagacc 10800
gcggggcatg actaacatga gaattacaac ttatatcgta tggggctgac ttcaggtgct 10860
acatttgaag agataaattg cactgaaatc tagaaatatt ttatctgatt aataagatga 10920
tcttcttgag atcgtttttg tctgcgcgta atctcttgct ctgaaaacgg aaaaaaccgc 10980
cttgacgggc ggtttttcga aggttctctg agctaccaac tctttgaacc gaggtaactg 11040
gcttgaggga gcgcagtcac caaaacttgt ctttcagtt tagccttaac cggcgcatga 11100
cttcaagact aactcctcta aatcaattac cagtggctgc tgccagtggg gcttttgc 11160
gtctttccgg gttggactca agacgatagt taccggataa ggcgagcgg tcggactgaa 11220
cgggggggtc gtgcatacag tccagcttgg agcgaactgc ctaccggaa ctgagtgtca 11280
ggcgtggaat gagacaaacg cggccataac agcggaaatga caccggtaaa ccgaaaggca 11340
ggaacaggag agcgcacgag ggagccgcca ggggaaacgc ctggtatctt tatagtcctg 11400
tcgggtttcg ccaccactga tttgagcgtc agatttcgtg atgcttgta ggggggcgga 11460
gcctatggaa aaacggcttt gccgcggccc tctcacttcc ctgttaagta tcttctggc 11520
atcttccagg aaatctccgc ccggttcgta agccatttcc gctcgccgca gtcgaacgac 11580
cgagcgtagc gagtcatgga gcgaggaagc ggaatatatc ctgtatcaca tattctgctg 11640
acgcaccggg gcagcctttt ttctctgccc acatgaagca cttcactgac accctcatca 11700
gtgccaacat agtaagccag tatacactcc gctagcgtg atgtccggcg gtgcttttgc 11760
cgttacgcac caccctgca gtagctgaac aggagggaca gctgatagaa acagaagcca 11820
gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgctt ttgagtgage 11880
tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga 11940
agagcgctg atgcggtatt ttctccttac gcactgtgct ggtatttcac accgcataga 12000

```

ccagccgcgt aacctggcaa aatcggttac ggttgagtaa taaatg 12046

<210> 17

<211> 7406

<212> DNA

<213> 人工序列

<220>

<223>

<400> 17

cgcgccggta tgtacaggaa gaggtttata ctaaactggtt acattgcaaa cgtggtttcg 60  
 tgtgccaaagt gtgaaaaccg atgtttaatc aaggctctga cgcatttcta caaccacgac 120  
 tccaagtgtg tgggtgaagt cagatgttta aacctatgtg cctggcagat aacttcgtat 180  
 aatgtatgct atacgaagtt atggtacgcg gccgcgtaga ggatctgttg atcagcagtt 240  
 caacctgttg ataatacggg cctttaattc aaccaacac aatatattat agttaataa 300  
 gaattattat caaatcattt gtatattaat taaaatacta tactgtaaat tacattttat 360  
 ttacaatcac tcgacaccgg tgatatccat atgggatcca cagcaatcaa gcatgcatta 420  
 caaagagaca tttttacacc aaatgatgaa cgctgctga gcattgtgaa tgtctgcaaa 480  
 gcaggaaaaa agaaaaagaa ctgtttttta tgtgccacag tgacaactga acgccctgtg 540  
 caggttaagg tgggtcaaagt caagaaatcc gataaggag atttctacaa aaggcagatt 600  
 gcatgggccc ttcgagatct tgctgtggta gatgccaaag atgctatcaa agaaaatcct 660  
 gaatttgatt tacactttga aaaaatatat aaatgggttg ccagcagcac tgctgaaaag 720  
 aatgcattta tttcatgcat ttggaaattg aatcagcgat atctccggaa gaaaattgat 780  
 tttgtcaatg ttagctcaca gcttttgga gaactgccta aagttacaga agaattctgtt 840  
 ccaagtggag aaaatcagag tgtgacagga ggtgatgaag aagtagtaga tgaataccaa 900  
 gagttaaatg caagagaaga acaggatatc gaaataatga tggaaggctg tgaatatgca 960  
 atctcgaatg cggaagcctt tgcagaaaaa ttgtccagag agctgcaggt gctagatggg 1020  
 gctaacatcc agtcaatcat ggcatctgaa aaacaagtca acatcctgat gaaattgcta 1080  
 gatgaggctc taaaggaggt agatcagatt gaattgaaac tgagcagtta tgaggaaatg 1140  
 ctccaaagtg taaaagaaca aatggatcag atctctgaaa gcaaccacct aattcatctt 1200  
 agtaacacta ataattgtaa actcctatct gagatagagt tccttgtgaa ccacatggac 1260  
 ttggccaaaag gtcatataaa ggcccttcag gaaggagatc ttgcttcttc cagaggcatt 1320  
 gaggcctgca ccaatgctgc tgatgccctt ctgcagtga tgaatgtagc tcttcgacca 1380  
 ggccatgact tgcttctggc agtcaaacag caacagcagc gattcagtga tttgcgagag 1440  
 ctttttgccc ggagactggc cagtcacctc aacaatgttt ttgttcaaca gggatcatgat 1500  
 cagagttcga ctcttgccca aactctgtt gaactgactt tacccaatca tcatccattt 1560  
 catagagatt tgctccgata tgccaagctg atggagtggc taaagagtac agattatgga 1620  
 aaatatgaag gactaacaaa gaattacatg gattatttat cccgactata tgaaagagaa 1680  
 atcaaagatt tctttgaagt tgcaaagatc aagatgactg gcacaactaa agaaagcaag 1740  
 aagtttgcta cactgcctcg aaaagaaagt gctgtcaaac aggaaacaga gagtcttcat 1800  
 ggaagttcgg ggaaattaac tggatctact tctagtctaa ataagctcag tgttcagagt 1860

tcagggaatc gcagatctca gtcattctcc ctgttggata tgggaaacat gtctgcctct 1920  
 gatctcgatg ttgtctgacag gaccaaattt gataagatct ttgaacaggt actaagttaa 1980  
 ctggagcccc tatgtctggc agaacaggac ttcataagta aatttttcaa actacagcaa 2040  
 catcaaagta tgcctggaac tatggctgaa gcagaggacc tggatggagg aacattatca 2100  
 cggcaacata attgtggcac accactgcct gtttcatctg agaaagatat gatccgccaa 2160  
 atgatgatta aaatattttc ctgcattgag ccagagctga acaacctaat tgcattagga 2220  
 gacaaaattg atagctttta ctctctttat atgttagtca aatgagtca tcatgtgtgg 2280  
 actgcacaaa atgtggacce tgcttctttc ctaagtacta cattgggaaa tgttttgggtg 2340  
 actgtcaaaa ggaactttga caaatgcatt agtaacaaa taaggcaa at ggaagaagta 2400  
 aagatctcaa aaaagagtaa agttggaatt cttccatttg ttgctgaatt tgaagaattt 2460  
 gctggacttg cagaatcaat cttcaaaaat gctgagcgtc gtggagacct ggataaagca 2520  
 tacacaaaac ttatcagagg agtatttgtt aatgtggaga aagtagcaaa tgaaagccag 2580  
 aagaccccca gggatgtggt tatgatggaa aactttcacc atatttttgc aactctttct 2640  
 cgattgaaaa tctcatgtct agaagcagaa aaaaaagaag ccaaacaaa atacacagat 2700  
 caccttcagt cttatgtcat ttactcttta ggacaacctc ttgaaaaact aaatcatttc 2760  
 tttgaagggtg ttgaagctcg cgtggcacag ggcataaggg aggaggaagt aagttacca 2820  
 cttgcattta acaaacaaga acttcgtaaa gtcattaagg agtaccctgg aaaggaagta 2880  
 aaaaaaggtc tagataacct ctacaagaaa gttgataaac atttatgtga agaagagaac 2940  
 ttacttcagg tgggtgtggca ctccatgcaa gatgaattta tacgccagta taagcacttt 3000  
 gaaggtttga tagctcgctg ttatcctgga tctggtgtta caatggaatt cactattcag 3060  
 gacattctgg attattgttc cagcattgca cagtcccacg gctcgagtga aaacctgtac 3120  
 ttccagagct ctagcttaag cgcggccgag accacggccg agttgttcga ggagcctttt 3180  
 gtggcagatg aatatattga acgtcttgta tggagaacct caggaggagg ctctagaggt 3240  
 ggacctgaag cttttgatcc taaaagatta ttagaagaat ttgtaaatca tattcaggaa 3300  
 ctccagataa tggatgaaag gattcagagg aaagtagaga aactagagca acaatgtcag 3360  
 aaagaagcca aggaatttgc caagaaggta caagagctgc agaaaagcaa tcaggttgcc 3420  
 ttccaacatt tccaagaact agatgagcac attagctatg tagcaactaa agtctgtcac 3480  
 cttggagacc agttagaggg ggtaaacaca cccagacaac gggcagtgga ggctcagaaa 3540  
 ttgatgaaat actttaatga gtttctagat ggagaattga aatctgatgt ttttacaat 3600  
 tctgaaaaga taaaggaagc agcagacatc attcagaagt tgcacctaat tgcccaagag 3660  
 ttaccttttg atagattttc agaagttaaa tccaaaattg caagtaata ccatgattta 3720  
 gaatgccagc tgattcagga gtttaccagt gctcaaagaa gaggtgaaat ctccagaatg 3780  
 agagaagtag cagcagtttt acttcatttt aagggttatt cccattgtgt tgatgtttat 3840  
 ataaagcagt gccaggaggg tgcttatttg agaaatgata tttttgaaga cgctggaata 3900  
 ctctgtcaaa gagtgaacaa acaagttgga gatattctca gtaatccaga aacagtcttg 3960  
 gctaaactta ttcaaaatgt atttgaaatc aaactacaga gttttgtgaa agagcagtta 4020  
 gaagaatgta ggaagtccga tgcagagcaa tatctcaaaa atctctatga tctgtataca 4080  
 agaaccacca atctttccag caagctgatg gagtttaatt taggtactga taaacagact 4140  
 ttcttgtcta agcttatcaa atccattttc atttcctatt tggagaacta tattgaggtg 4200

gagactggat atttgaagc cagaagtgc atgacccac agcgcctatta tgattcgaaa 4260  
 aaccatcaaa agagatccat tggcacagga ggtattcaag atttgaagga aagaattaga 4320  
 cagcgtacca acttaccact tgggccaagt atcgatactc atggggagac ttttctatcc 4380  
 caagaagtgg tggtaaatct tttaacaaga accaacaag cctttgaaag atgtcatagg 4440  
 ctctctgac cttctgactt accaaggaat gccttcagaa tttttaccat tcttgtggaa 4500  
 tttttatgta ttgagcatat tgattatgct ttggaaacag gacttgctgg aattccctct 4560  
 tcagattcta ggaatgcaaa tctttatatt ttggacgttg tgcaacaggc caatactatt 4620  
 tttcatcttt ttgacaaa gtttaattgat cacttatgc cactaataag ctcttctcct 4680  
 aagttatctg aatgccttca gaagaaaaa gaaataattg acaaatgga gatgaaattg 4740  
 gatactggca ttgataggac attaaattgt atgattggac agatgaagca tattttggct 4800  
 gcagaacaga agaaaacaga ttttaagcca gaagatgaa acaatgtttt gattcaatat 4860  
 actaatgcct gtgtaaaagt ctgtgcttac gtaagaaac aagtggagaa gattaataat 4920  
 tccatggatg ggaagaatgt ggatacagtt ttgatggaac ttggagtacg ttttcacga 4980  
 cttatctatg agcatcttca acaatattcc tacagttgta tgggtggcat gttggcaatt 5040  
 tgtgatgtag ccgaatatag gaagtgtgcc aaagacttca agattccaat ggtattacat 5100  
 ctttttgata ctctgcatgc tctttgcaat cttctggtag ttgccccaga taatttaaag 5160  
 caagtctgct caggagaaca acttgctaact ctggacaaga atatacttca ctcttcgta 5220  
 caacttcgtg ctgattatag atctgcccgc cttgctcgac acttcagcta actcgagggt 5280  
 acccgaaagg aagctgagtt ggctgctgcc accgctgagc aataactagc ataaccctt 5340  
 ggggcctcta aacgggtctt gaggggtttt ttgctgaaag gaggaactat cctcaggggg 5400  
 agatggggga ggctaactga aacacggaag gagacaatac cggaaggaa cgcgctatg 5460  
 acggcaataa aaagacagaa taaaacgcac ggggtgttggg tcgtttgttc ataaacgcgg 5520  
 ggttcggtcc cagggtggtc actctgtcga taccaccgc agaccctatt gggaccaata 5580  
 cgcgcgcgtt tcttcctttt cccacccca accccaagt tcgggtgaag gccaggggt 5640  
 cgcagccaac gtcggggcgg caagccctgc catagccact acgggtacgt ctgaaagcat 5700  
 gccttttttg aatttacgta ctaagctctc atgtttcacg tactaagctc tcatgtttta 5760  
 cgtactaagc tctcatgttt aacgaactaa accctcatgg ctaacgtact aagctctcat 5820  
 ggctaacgta ctaagctctc atgtttcacg tactaagctc tcatgtttga acaataaaat 5880  
 taatataaat cagcaactta aatagcctct aaggttttaa gttttataag aaaaaaaga 5940  
 atatataagg cttttaaaag ttttaagggt taacggttgt ggacaacaag ccagggtatg 6000  
 aacgcactga gaagccctta gagcctctca aagcaatttt cagtgcaca ggaacactta 6060  
 acggctgaca gaattagctt cacgtgccc caagcactca gggcgcaagg gctgctaaag 6120  
 gaagcggaac acgtagaaag ccagtccgca gaaacggtgc tgaccccgga tgaatgtcag 6180  
 ctactgggct atctggacaa gggaaaacgc aagcgcaag agaaagcagg tagcttgcag 6240  
 tgggcttaca tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt 6300  
 gccagctggg gcgccctctg gtaaggttgg gaagccctgc aaagtaaaact ggatggcttt 6360  
 cttgccgcca aggatctgat ggcgagggg atcaagatct gatcaagaga caggatgagg 6420  
 atcgtttcgc atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga 6480  
 gaggtattc ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt 6540

ccggctgtca gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct 6600  
 gaatgaactg caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg 6660  
 cgcagctgtg ctcgacgttg tcaactgaagc gggaaggagc tggctgctat tgggcgaagt 6720  
 gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc 6780  
 tgatgcaatg cggcgggtgc atacgttga tccggctacc tgccattcg accaccaagc 6840  
 gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgctg atcaggatga 6900  
 tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg 6960  
 catgcccagc ggcgaggatc tcgtcgtgac acatggcgat gcctgcttgc cgaatatcat 7020  
 ggtggaaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcgaccg 7080  
 ctatcaggac atagcgttgg ctaccctga tattgctgaa gagcttggcg gcgaatgggc 7140  
 tgaccgttc ctctgtcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta 7200  
 tcgccttctt gacgagttct tetgagcggg actctggggg tcgaaatgac cgaccaagcg 7260  
 acgccaacc tgccatcacg agatttcgat tccaccgccg cttctatga aaggttgggc 7320  
 ttcggaatcg tttccggga cgcggctgg atgacctcc agcgcgggga tctcatgctg 7380  
 gatttcttcg cccacatcat atcgat 7406

<210> 18

<211> 7532

<212> DNA

<213> 人工序列

<220>

<223>

<400> 18

cgcgccggta tgtacaggaa gaggtttata ctaaactgtt acattgcaaa cgtggtttcg 60  
 tgtgccaagt gtgaaaaccg atgtttaatc aaggctctga cgcatttcta caaccacgac 120  
 tccaagtgtg tgggtgaagt cagatgttta aacctatgtg cctggcagat aacttcgtat 180  
 aatgtatgct atacgaagtt atggtacgcg gccgcgtaga ggatctgttg atcagcagtt 240  
 caacctgttg ataatacggc cttttaattc aaccaacac aatatattat agttaataa 300  
 gaattattat caaatcattt gtatattaat taaaatacta tactgtaaat tacattttat 360  
 ttacaatcac tcgacaccgg tgatatccat atggcggccg ctagccgatc acgacaaccc 420  
 ccccttgtag cggcatctc tccaaatgaa gggataccat ggacgaaggt cacaatcagg 480  
 ggagaaaaatc tggggactgg cccaccgac ctcataggct tgaccatttg tggacataat 540  
 tgctctctga cggcagaatg gatgtctgca agtaaaatag tatgtcagat gggacaagcc 600  
 aaaaatgaca aaggagacat tattgtcacc actaagtcag gtggcagagg aacctcaaca 660  
 gtctctttca agctactcaa acctgagaaa ataggcattt tggatcagtc tgctgtgttg 720  
 gttgatgaaa tgaattatta tgatatgcgc actgacagga acaaaggaat tccgcccttg 780  
 tccttacgta ctgctaaccg gcttggcatt gagattgaaa aaagtaaatt ttcgcagaag 840  
 gacttagaaa tgctattcca tggaatgagt gctgatttta caagtgagaa tttctcagca 900  
 gcctggatc ttatagagaa tcaactcaaac accagttttg agcagctcaa aatggcagtc 960  
 accaacctaa agagacaggc taacaagaag agtgagggca gcctggccta tgtgaaaggc 1020

ggtctcagta cattcttcga agcacaggat gccctctcag ccatccatca aaaactagaa 1080  
 gcagatggaa cggaaaaagt agaaggatcc atgacgcaga aactggagaa tgttctgaac 1140  
 agagcaagta atactgcaga cacattgttt caagaagtat taggtcggaa agacaaggca 1200  
 gattccacta gaaatgcact caatgtgctt cagcgattta agtttctttt caaccttcct 1260  
 ctaaatattg aaaggaatat tcaaaagggt gattatgatg tggttattaa tgattatgaa 1320  
 aaggccaagt cactttttgg gaaaacggag gtgcaagttt tcaagaaata ttatgctgaa 1380  
 gtgaaacaa ggattgaagc ttttaagagaa ttacttctgg ataaattgct tgagacacca 1440  
 tcaactttac atgacaaaaa acgttacata aggtacctgt ctgaccttca tgcgtctggt 1500  
 gacctgctt ggcaatgcat tggagcccaa cacaagtgga tccttcagct catgcacagt 1560  
 tgcaaagagg gctacgtgaa agatctgaaa ggtaaccag gcctgcacag tcccatgttg 1620  
 gatcttgata atgatacacg tccctcagtg ttgggccatc tcagtcagac agcgtccctg 1680  
 aagaggggca gcagctttca gtctggtcga gacgacacgt ggagatacaa aactccccac 1740  
 aggggtggcct ttgttgaaaa attgacaaaa ctctgtctga gccagctgcc taacttctgg 1800  
 aaactctgga tctcctacgt taatggaagc ctcttcagt agactgctga gaagtcaggc 1860  
 cagattgaaa gatcaaagaa tgtaaggcaa agacaaaatg attttaagaa aatgattcag 1920  
 gaagtaatgc actccctggt gaagcttacc cgcggagccc tgcttcccct cagcatccgg 1980  
 gatggggaag ccaagcagta cggaggctgg gaggtgaagt gcgagctctc cggacagtgg 2040  
 ctctctcacg ccatccagac tgtaagactt actcacgaat cgttgactgc ccttgaaatt 2100  
 cctaatagacc tgttacagac tatccaggat ctcatcttgg atctccagat acgttgcgta 2160  
 atggccacgt tgcagcacac ggcggaagaa ataaagagat tagctgaaaa agaagactgg 2220  
 attgttgaca atgaaggact gacttctcta ccatgtcagt ttgaacagtg catcgtgtgt 2280  
 tctctgcagt cactgaaggg ggttctggag tgcaagccgg gagaggccag tgtcttccaa 2340  
 caacctaaaa cacaggagga ggtttgccag ctaagcatca atataatgca ggtttttata 2400  
 tactgtctgg aacagttgag caccaagcct gatgcagata tagatactac acatctctct 2460  
 gttgatgttt cttcccctga ctgttttgga agtatccatg aagacttcag cttgacctca 2520  
 gaacagcgcc ttttgatagt cctaagtaat tgctgctatc tagaacgtca caccttccta 2580  
 aatatcgag aacattttga aaagcacaac ttccagggaa tagaaaaaat cacacaggtt 2640  
 agcatggcct cattgaaaga actagatcaa agactctttg aaaattacat cgagttgaaa 2700  
 gcagatccca tcgttggctc cttagaacct ggaatttatg caggatattt tgattggaag 2760  
 gactgcctgc ctccaacagg tgtcagaaac tatttaaaag aagcactggg gaatataatt 2820  
 gccgtgcatg cagaggtgtt caccatttcc aaagaactgg tccctcgggt actatccaag 2880  
 gtgatagaag cagtttctga agagctcagt cgactgatgc agtgtgtttc atccttcagc 2940  
 aaaaaatggag ctttacagge gagacttgaa atctgtgctt tgagggacac tgtggctgtt 3000  
 tacctgacac ccgaaagcaa gtcaagtttt aagcaggctt tggaagccct gcccagctt 3060  
 tccagtggag cagataaaaa gttactggaa gagctcctga acaagttcaa gagtagcatg 3120  
 cacttgacgc tcacctgttt ccaagcagct tcttcaacca tgatgaaaac aggtctgagt 3180  
 gaaaatttgt attttcaaag cactagtggg acccttaaga tggcgatggc gatgtcggac 3240  
 agtggggcga gccgcctgct tcggcagctg gagtcagggg gttttgagge gcggctgtac 3300  
 gtgaagcagc tctcgcagca gtcggatggg gaccgggacc tccaggagca ccggcagcgc 3360

atccaggcgc tggcggagga gacggcgcag aacctgaagc gcaacgtcta ccagaactac 3420  
 cggcagttca tagagacggc ccgcgagatc tcctacctgg agagcgagat gtaccagctc 3480  
 agccatttgc tgaccgagca gaaaagcagc ctggagagca tcccgtttac gttgctgcct 3540  
 gccgctgctg ccgccggagc cgccgccgcc tctggagggg aggagggagt cgggtggggcg 3600  
 gggggccgag accacctccg aggccaggcc ggctttttct ccacccccgg ggggtgcctcc 3660  
 cgcgacggct ccggtccagg cgaggaagga aagcagcgca ctctcaccac cctgcttgag 3720  
 aaggtggaag gctgcaggca tctgctggag acgccgggac agtacttggg gtacaatggg 3780  
 gacctagtgg aatacgatgc ggaccacatg gccaactgc agcgggtgca cggctttctc 3840  
 atgaacgatt gcttgttggg ggctacctgg ctgcctcagc ggctggggat gtatcgctac 3900  
 aacgctctct attccctaga tggtttggcc gtagtcaatg tcaaggacaa cccgccccatg 3960  
 aaggacatgt tcaagctgct tatgttcccc gagagccgta ttttccaggc cgaaaatgct 4020  
 aaaatcaaac gagagtggct ggaagtgctg gaggacacca agagggccct cagtgagaaa 4080  
 aggcgaaagg agcaggagga ggcagcggcc cctcgagggc caccccaagt gacttccaag 4140  
 gccactaacc catttgagga tgacgaagaa gaagaaccag ctgttcctga ggtagaggaa 4200  
 gagaagggtg acctctccat ggaatggatc caggagttag ctgaagacct ggatgtctgc 4260  
 attgcgcaga gagactttga aggggcggtt gacctgctgg ataaattgaa ccattacctg 4320  
 gaagataaac ctagcccacc tcctgtaaaa gaactaaggg ccaaagtgga ggagcgagtt 4380  
 cgacagctca ctgagggtgct agttttcgaa ctctccccag atcgttccct gagaggtggg 4440  
 ccgaaggcta ctgcagagc agtttcgcaa ctgatccggc tgggccagtg cacgaaggcc 4500  
 tgtgagctat ttttgagaaa cagggcagcc gctgttcata ctgcaattcg tcagcttcgc 4560  
 atcgaagggtg ccactttact ctatattcat aagctgtgcc atgtcttctt taccagcctt 4620  
 ctcgagactg caagagaatt tgagatcgat ttgacaggca ctgacagcgg ctgctactct 4680  
 gcctttgtgg tctgggcaag atcagccatg ggcatgttcg tggatgcttt tagcaagcag 4740  
 gtgtttgata gtaaggagag cctctctaca gcagctgagt gtgtaaaagt ggctaaggag 4800  
 cattgccagc aactgggtga tatcggactg gatctcacct tcatcatcca tgcccttctg 4860  
 gtgaaagaca tccaaggggc cttgcacagt tacaaagaaa tcatcattga agccactaaa 4920  
 catcgcaact ctgaagagat gtggaggagg atgaacttga tgacgccaga agccctgggt 4980  
 aagctcaaag aagagatgaa aagttgtggg gtaagtaact ttgagcagta cacaggggat 5040  
 gactgctggg tgaacctaa ttacacagtg gttgctttca ccaaacagac catgggcttc 5100  
 ttggaagagg ccctgaagct gtatttccca gagctgcaca tgggtactttt ggagagcctg 5160  
 gtggaaaatca ttttggttgc tgttcagcat gtggattata gtcttcgatg tgagcaggat 5220  
 ccagagaaga aagcttttat cagacagaat gcatectttt tatatgaaac agtctctcct 5280  
 gtggtggaga aaaggtttga agaaggtgtg gggaaacctg ccaagcaact ccaagatctg 5340  
 aggaatgcat ctagacttat tcgtgtgaat cctgaaagta caacatcagt ggtctaagaa 5400  
 ttcggtaccc gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa 5460  
 ccccttgggg cctctaaacg ggtcttgagg ggttttttgc tgaaaggagg aactatcctc 5520  
 agggggagat gggggaggct aactgaaaca cggaaggaga caataccgga aggaacccgc 5580  
 gctatgacgg caataaaaag acagaataaa acgcacgggt gttgggtcgt ttgttcataa 5640  
 acgcgggggt cggtcccagg gctggcactc tgtcgatacc ccaccgagac cccattggga 5700

```

ccaatacgcc cgcgtttctt ccttttcccc accccaaccc ccaagttcgg gtgaaggccc 5760
agggctcgca gccaacgtcg gggcggcaag ccctgccata gccactacgg gtacgtctga 5820
aagcatgcct ttttggaatt tacgtactaa gctctcatgt ttcacgtact aagctctcat 5880
gtttaacgta ctaagctctc atgtttaacg aactaaaccc tcatggctaa cgtactaagc 5940
tctcatggct aacgtactaa gctctcatgt ttcacgtact aagctctcat gtttgaacaa 6000
taaaattaat ataaatcagc aacttaaata gcctctaagg ttttaagttt tataagaaaa 6060
aaaagaatat ataaggcttt taaagctttt aaggtttaac ggttggtggac aacaagccag 6120
ggatgtaacg cactgagaag cccttagagc ctctcaaagc aattttcagt gacacaggaa 6180
cacttaacgg ctgacagaat tagcttcacg ctgccgaag cactcagggc gcaagggctg 6240
ctaaaggaag cggaacacgt agaaagccag tccgcagaaa cgggtgctgac cccggatgaa 6300
tgtcagctac tgggctatct ggacaaggga aaacgcaagc gcaaagagaa agcaggtagc 6360
ttgcagtggg cttacatggc gatagctaga ctgggcggtt ttatggacag caagcgaacc 6420
ggaattgcca gctggggcgc cctctggtaa ggttggaag ccctgcaaag taaactggat 6480
ggctttcttg ccgccaagga tctgatggcg caggggatca agatctgac aagagacagg 6540
atgaggatcg tttcgcatga ttgaacaaga tggattgcac gcaggttctc cggccgcttg 6600
ggtggagagg ctattcggct atgactgggc acaacagaca atcggtgct ctgatgccgc 6660
cgtgttccgg ctgtcagcgc aggggcgcgc ggttcttttt gtcaagaccg acctgtccgg 6720
tgccctgaat gaactgcagg acgaggcagc gcggtatcg tggctggcca cgacgggcgt 6780
tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc tgctattggg 6840
cgaagtgccg gggcaggatc tcctgtcatc tcacettgct cctgccgaga aagtatccat 6900
catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc cattcgacca 6960
ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc ttgtcgatca 7020
ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg ccaggctcaa 7080
ggcgcgcatg cccgacggcg aggatctcgt cgtgacacat ggcatgcct gcttgccgaa 7140
tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc tgggtgtggc 7200
ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc ttggcggcga 7260
atgggctgac cgcttcctcg tgetttacgg tateccgct cccgattcgc agcgcatcgc 7320
cttctatcgc cttcttgacg agttcttctg agcgggactc tggggttcga aatgaccgac 7380
caagcgacgc ccaacctgcc atcacgagat ttgattcca ccgccgcctt ctatgaaagg 7440
ttgggcttcg gaatcgtttt ccgggacgcc ggctggatga tcctccagcg cggggatctc 7500
atgctggagt tcttcgccc catcatatcg at 7532

```

<210> 19

<211> 7526

<212> DNA

<213> 人工序列

<220>

<223>

<400> 19

```

cgcgccggtg tgtacaggaa gaggtttata ctaaactgtt acattgcaaa cgtggtttcg 60

```

tgtccaagt gtgaaaaccg atgtttaatc aaggctctga cgcatttcta caaccacgac 120  
 tccaagtgtg tgggtgaagt cagatgttta aacctatgtg cctggcagat aacttcgtat 180  
 aatgtatgct atacgaagtt atggtagcgc gccgcgtaga ggatctgttg atcagcagtt 240  
 caacctgttg ataatacggc cctttaattc aaccaacac aatatattat agttaaataa 300  
 gaattattat caaatcattt gtatattaat taaaatacta tactgtaaat tacattttat 360  
 ttacaatcac tcgacaccgg tgatatccat atgggatcca cagcaatcaa gcatgcatta 420  
 caaagagaca tttttacacc aaatgatgaa cgcctgctga gcattgtgaa tgtctgcaaa 480  
 gcaggaaaaa agaaaaagaa ctgtttttta tgtgccacag tgacaactga acgccctgtg 540  
 cagggttaagg tgggtcaaagt caagaaatcc gataaggagc atttctacaa aaggcagatt 600  
 gcatgggccc ttcgagatct tgctgtggta gatgccaaag atgctatcaa agaaaatcct 660  
 gaatttgatt tacactttga aaaaatataa aaatgggttg ccagcagcac tgctgaaaag 720  
 aatgcattta tttcatgcat ttggaaattg aatcagcgat atctccgga gaaaattgat 780  
 tttgtcaatg ttagctcaca gcttttgga gaactgccta aagttacaga agaattctgtt 840  
 ccaagtggag aaaatcagag tgtgacagga ggtgatgaag aagtagtaga tgaataccaa 900  
 gagttaaatg caagagaaga acaggatatc gaaataatga tggaaggctg tgaatatgca 960  
 atctcgaatg cggaagcctt tgcagaaaaa ttgtccagag agctgcaggt gctagatggg 1020  
 gctaacatcc agtcaatcat ggcatctgaa aaacaagtca acatcctgat gaaattgcta 1080  
 gatgaggctc taaaggaggt agatcagatt gaattgaaac tgagcagtta tgaggaaatg 1140  
 ctccaaagtg taaaagaaca aatggatcag atctctgaaa gcaaccacct aattcatctt 1200  
 agtaacacta ataattgaaa actcctatct gagatagagt tccttgtgaa ccacatggac 1260  
 ttggccaaag gtcatataaa ggcccttcag gaaggagatc ttgcttcttc cagaggcatt 1320  
 gaggcctgca ccaatgctgc tgatgccctt ctgcagtgca tgaatgtagc tcttcgacca 1380  
 ggccatgact tgcttctggc agtcaaacag caacagcagc gattcagtga tttgcgagag 1440  
 ctttttgccc ggagactggc cagtcacctc aacaatgttt ttgttcaaca gggcatgat 1500  
 cagagttcga ctcttgccca acactctgtt gaactgactt tacccaatca tcatccattt 1560  
 catagagatt tgctccgata tgccaagctg atggagtggc taaagagtac agattatgga 1620  
 aaatatgaag gactaacaaa gaattacatg gattatttat cccgactata tgaaagagaa 1680  
 atcaaagatt tctttgaagt tgcaaagatc aagatgactg gcacaactaa agaaagcaag 1740  
 aagtttgcta cactgcctcg aaaagaaagt gctgtcaaac aggaaacaga gagtcttcat 1800  
 ggaagttcgg ggaaattaac tggatctact tctagtctaa ataagctcag gtctcagagt 1860  
 tcagggaatc gcagatctca gtcattcttc ctgttggata tgggaaacat gtctgcctct 1920  
 gatctcgatg ttgtgacag gaccaaattt gataagatct ttgaacaggt actaagtga 1980  
 ctggagcccc tatgtctggc agaacaggac ttcataagta aatttttcaa actacagcaa 2040  
 catcaaagta tgcttggac tatggctgaa gcagaggacc tggatggagg aacattatca 2100  
 cggcaacata attgtggcac accactgcct gtttcatctg agaaagatat gatccgcaa 2160  
 atgatgatta aaatatttcg ctgcattgag ccagagctga acaacctaat tgcattagga 2220  
 gacaaaattg atagctttaa ctctctttat atgttagtca aaatgagtca tcatgtgtgg 2280  
 actgcacaaa atgtggaccc tgcttcttct ctaagtacta cattgggaaa tgttttgggtg 2340  
 actgtcaaaa ggaactttga caaatgcatt agtaacaaaa taaggcaaat ggaagaagta 2400

aagatctcaa aaaagagtaa agttggaatt cttccatttg ttgctgaatt tgaagaattt 2460  
gctggacttg cagaatcaat cttcaaaaat gctgagcgct gtggagacct ggataaagca 2520  
tacaccaaac ttatcagagg agtatttggt aatgtggaga aagtagcaaa tgaaagccag 2580  
aagaccccca gggatgtggt tatgatggaa aactttcacc atatttttgc aactctttct 2640  
cgattgaaaa tctcatgtct agaagcagaa aaaaaagaag ccaaacaaaa atacacagat 2700  
caccttcagt cttatgtcat ttactcttta ggacaacctc ttgaaaaact aaatcatttc 2760  
tttgaagggtg ttgaagctcg cgtggcacag ggcataaggg aggaggaagt aagttacca 2820  
cttgcattta acaaacaaga acttcgtaaa gtcattaagg agtaccctgg aaaggaagta 2880  
aaaaaagggtc tagataacct ctacaagaaa gttgataaac atttatgtga agaagagaac 2940  
ttacttcagg tgggtgtggca ctccatgcaa gatgaattta tacgccagta taagcacttt 3000  
gaaggtttga tagctcgtg ttatcctgga tctggtgtta caatggaatt cactattcag 3060  
gacattctgg attattgttc cagcattgca cagtcccacg gctcgagtga aaacctgtac 3120  
ttccagagct ctagcttaag catggcctgg agccatccgc aatttgaaaa aggtggcggg 3180  
tccggcggag gtagcggcgg aggttcttgg tctcacctc agttcgagaa ggatgacgat 3240  
gataaaacca tgggatccct aggtaccgcg gccgcgacca cggccgagtt gttcgaggag 3300  
ccttttgtgg cagatgaata tattgaacgt cttgtatgga gaacccagg aggaggctct 3360  
agagggtggac ctgaagcttt tgatcctaaa agattattag aagaatttgt aaatcatatt 3420  
caggaactcc agataatgga tgaaaggatt cagaggaaag tagagaaact agagcaacaa 3480  
tgtcagaaag aagccaagga atttgccaag aaggtacaag agctgcagaa aagcaatcag 3540  
gttgccttcc aacatttcca agaactagat gagcacatta gctatgtagc aactaaagtc 3600  
tgtcaccttg gagaccagtt agagggggtg aacacacca gacaacgggc agtggaggct 3660  
cagaaattga tgaaatactt taatgagttt ctagatggag aattgaaatc tgatgttttt 3720  
acaaattctg aaaagataaa ggaagcagca gacatcattc agaagttgca cctaattgcc 3780  
caagagttac cttttgatag attttcagaa gttaaatcca aaattgcaag taaataccat 3840  
gatttagaat gccagctgat tcaggagttt accagtgcctc aaagaagagg tgaaatctcc 3900  
agaatgagag aagtagcagc agttttactt cattttaagg gttattccca ttgtgttgat 3960  
gtttatataa agcagtgccg ggagggtgct tatttgagaa atgatataatt tgaagacgct 4020  
ggaatactct gtcaaagagt gaacaaacaa gttggagata tcttcagtaa tccagaaaca 4080  
gtcctggcta aacttattca aaatgtattt gaaatcaaac tacagagttt tgtgaaagag 4140  
cagttagaag aatgtaggaa gtccgatgca gagcaatate tcaaaaatct ctatgatctg 4200  
tatacaagaa ccaccaatct ttccagcaag ctgatggagt ttaatttagg tactgataaa 4260  
cagactttct tgtetaaget tatcaaatcc attttcattt cctatttggga gaactatatt 4320  
gaggtggaga ctggatattt gaaaagcaga agtgcctatga tcctacagcg ctattatgat 4380  
tcgaaaaacc atcaaaagag atccattggc acaggaggtg ttcaagattt gaaggaaaga 4440  
attagacagc gtaccaactt accacttggg ccaagtatcg atactcatgg ggagactttt 4500  
ctatcccaag aagtgggtgt taatctttta caagaaacca aacaagcctt tgaaagatgt 4560  
cataggctct ctgatccttc tgacttacca aggaatgcct tcagaatttt taccattctt 4620  
gtggaatttt tatgtattga gcatattgat tatgcttttg aaacaggact tgctggaatt 4680  
ccctcttcag attctaggaa tgcaaatctt tatttttttg acgttgtgca acaggccaat 4740

actatTTTTc atctTTTTga caaacagttt aatgatcacc ttatgccact aataagctct 4800  
 tctcctaagt tatctgaatg ccttcagaag aaaaaagaaa taattgaaca aatggagatg 4860  
 aaattggata ctggcattga taggacatta aattgtatga ttggacagat gaagcatatt 4920  
 ttggctgcag aacagaagaa aacagatttt aagccagaag atgaaaacaa tgTTTTgatt 4980  
 caatatacta atgcctgtgt aaaagtctgt gcttacgtaa gaaaacaagt ggagaagatt 5040  
 aaaaattcca tggatgggaa gaatgtggat acagTTTTga tggaaacttg agtacgtttt 5100  
 catcgactta tctatgagca tcttcaacaa tttcctaca gttgtatggg tggcatgttg 5160  
 gcaatttgtg atgtagccga atataggaag tgtgccaaag acttcaagat tccaatggta 5220  
 ttacatcttt ttgatactct gcatgctctt tgcaatcttc tggtagttgc cccagataat 5280  
 ttaaagcaag tctgctcagg agaacaactt gctaactctgg acaagaatat acttactctc 5340  
 ttcgtacaac ttcgtctga ttatagatct gcccgccttg ctgcacactt cagctaaggt 5400  
 acccgaaagg aagetgagtt ggetgctgcc accgctgagc aataactagc ataaccctt 5460  
 ggggcctcta aacgggtctt gaggggtttt ttgctgaaag gaggaactat cctcaggggg 5520  
 agatggggga ggctaactga aacacggaag gagacaatac cggaaggaaac ccgcgctatg 5580  
 acggcaataa aaagacagaa taaaacgcac ggggtgttggg tcgtttgttc ataaacgcgg 5640  
 ggttcggtcc cagggtctgc actctgtcga taccaccg agaccctt gggaccaata 5700  
 cgcccgctt tcttctttt cccacccca accccaagt tcgggtgaag gcccagggt 5760  
 cgcagccaac gtcggggcgg caagccctgc catagccact acgggtacgt ctgaaagcat 5820  
 gcctttttgg aatttacgta ctaagctctc atgtttcacg tactaagctc tcatgtttaa 5880  
 cgtactaagc tctcatgttt aacgaactaa accctcatgg ctaacgtact aagctctcat 5940  
 ggctaacgta ctaagctctc atgtttcacg tactaagctc tcatgtttga acaataaaat 6000  
 taatataaat cagcaactta aatagcctct aaggttttaa gttttataag aaaaaaaga 6060  
 atatataagg cttttaaagg ttttaagggt taacggttgt ggacaacaag ccagggatgt 6120  
 aacgcactga gaagccctta gagcctctca aagcaatttt cagtgcaca ggaacactta 6180  
 acggctgaca gaattagctt cacgtctccg caagcactca gggcgcaagg gctgctaaag 6240  
 gaagcggaac acgtagaaag ccagtccgca gaaacggtgc tgaccccgga tgaatgtcag 6300  
 ctactgggct atctggacaa gggaaaacgc aagcgcaaag agaaagcagg tagcttgcag 6360  
 tgggcttaca tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt 6420  
 gccagctggg gcgcccctct gtaaggttgg gaagccctgc aaagtaaact ggatggcttt 6480  
 cttgccgcca aggatctgat ggcgcagggg atcaagatct gatcaagaga caggatgagg 6540  
 atcgtttcgc atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga 6600  
 gaggtattc ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgcctgtt 6660  
 ccggctgtca gcgcaggggc gcccgttct ttttgtcaag accgacctgt ccggtgccct 6720  
 gaatgaactg caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttcttg 6780  
 cgcagctgtg ctgcacgttg tcaactgaagc gggaaggac tggtcttat tgggcgaagt 6840  
 gccggggcag gatctcctgt catctcact tgctcctgcc gagaaagtat ccatcatggc 6900  
 tgatgcaatg cggcggctgc atacgttga tccggtacc tgccattcg accaccaagc 6960  
 gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgcg atcaggatga 7020  
 tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg 7080

catgccccgac ggcgaggatc tcgtcgtgac acatggcgat gcctgcttgc cgaatatcat 7140  
 ggtggaaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg 7200  
 ctatcaggac atagcgttgg ctaccctga tattgctgaa gagcttggcg gcgaatgggc 7260  
 tgaccgcttc ctctgtcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta 7320  
 tcgccttctt gacgagttct tctgagcggg actctggggt tcgaaatgac cgaccaagcg 7380  
 acgccccaac tgccatcacg agatttcgat tccaccgccg ctttctatga aaggttgggc 7440  
 ttcggaatcg tttccggga cgccggtg atgacctcc agcgcgggga tctcatgctg 7500  
 gagtctctcg cccacatcat atcgat 7526

<210> 20

<211> 5911

<212> DNA

<213> 人工序列

<220>

<223>

<400> 20

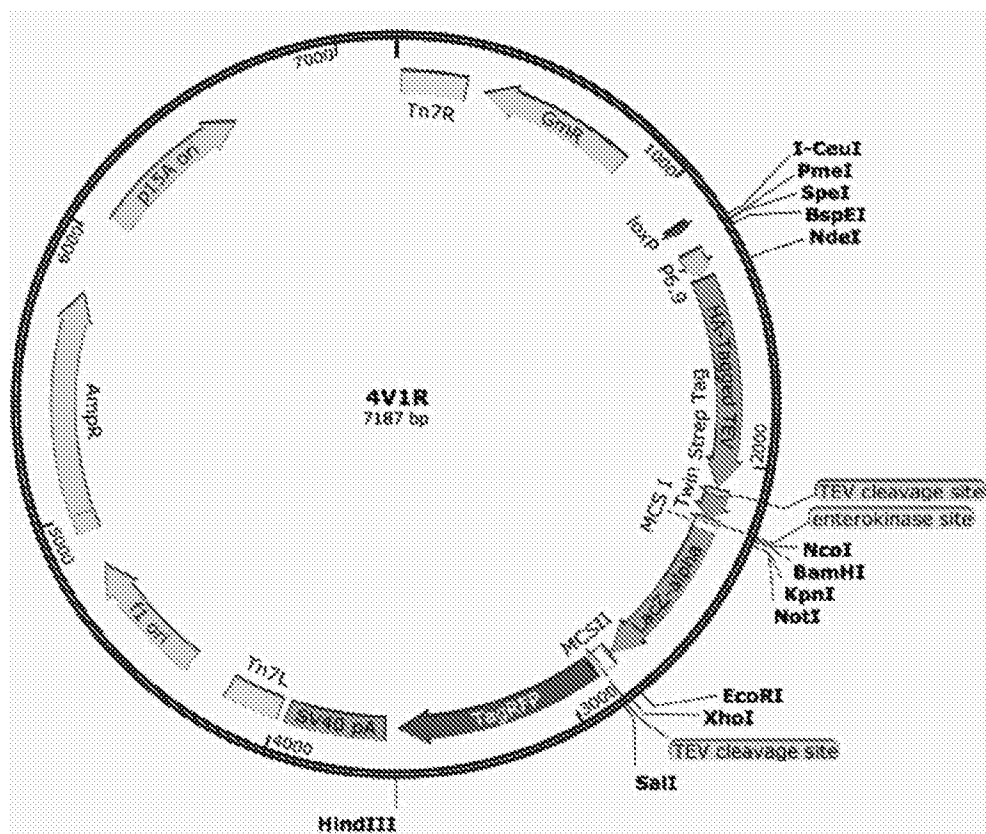
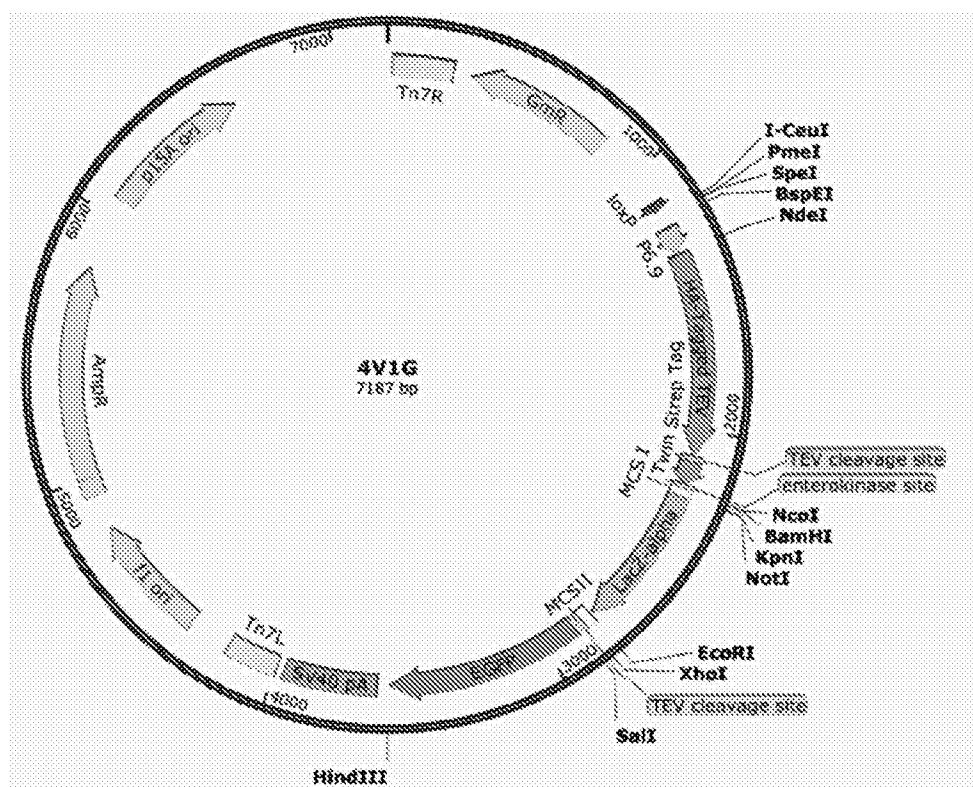
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60  
 cagcgtgacc gctacacttg ccagcgccct agcgcgccgt ctttctgctt tcttcccttc 120  
 ctttctcgcc acgttcgccg gctttccccg tcaagcteta aatcgggggc tccctttagg 180  
 gtccgattt agtgctttac ggcacctga ccccaaaaaa cttgattagg gtgatggttc 240  
 acgtagtggg ccatcgccct gatagacggt ttttgcctt ttgacgttgg agtccacgtt 300  
 ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatct cggtctatct 360  
 ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420  
 acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480  
 tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540  
 tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa ggaagagtat 600  
 gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt gccttcctgt 660  
 ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg 720  
 agtgggttac atcgaactgg atctcaacag cggtgaagatc cttgagagtt ttgcctccga 780  
 agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg tattatcccc 840  
 tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga atgacttggg 900  
 tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa gagaattatg 960  
 cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg 1020  
 aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa ctgccttga 1080  
 tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca ccacgatgcc 1140  
 tgcagcaatg gcaacaacgt tgcgcaaaact attactggc gaactactta ctctagcttc 1200  
 ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac ttctgcgctc 1260  
 ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg 1320  
 cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac 1380  
 gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga taggtgcctc 1440

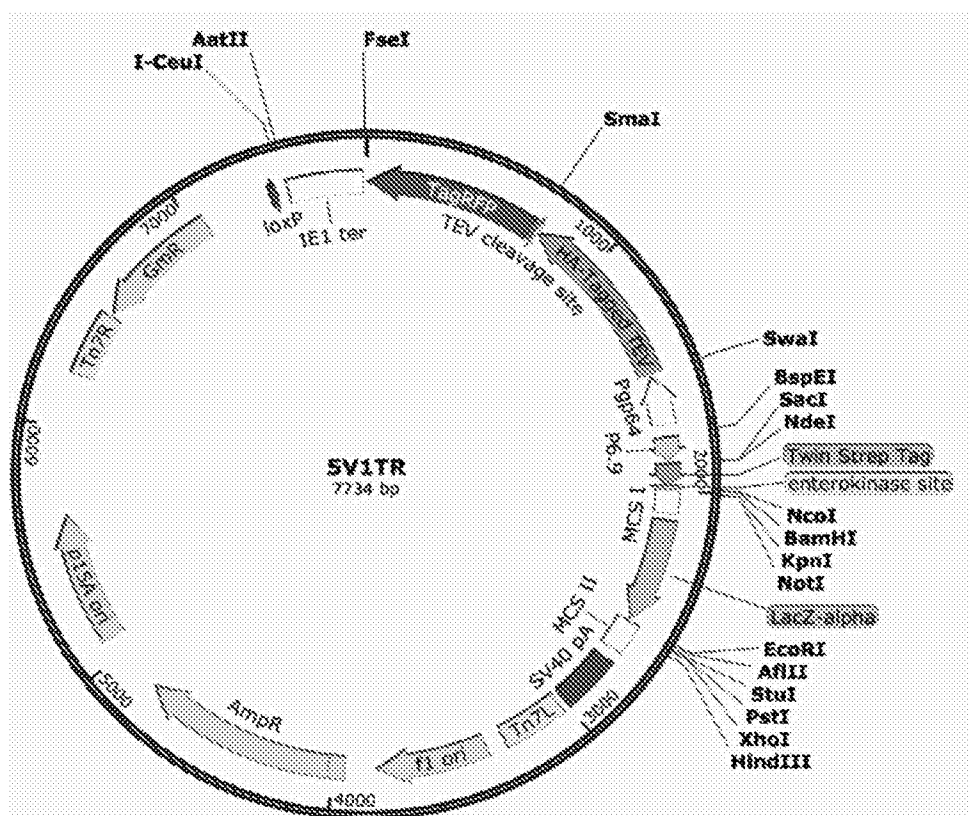
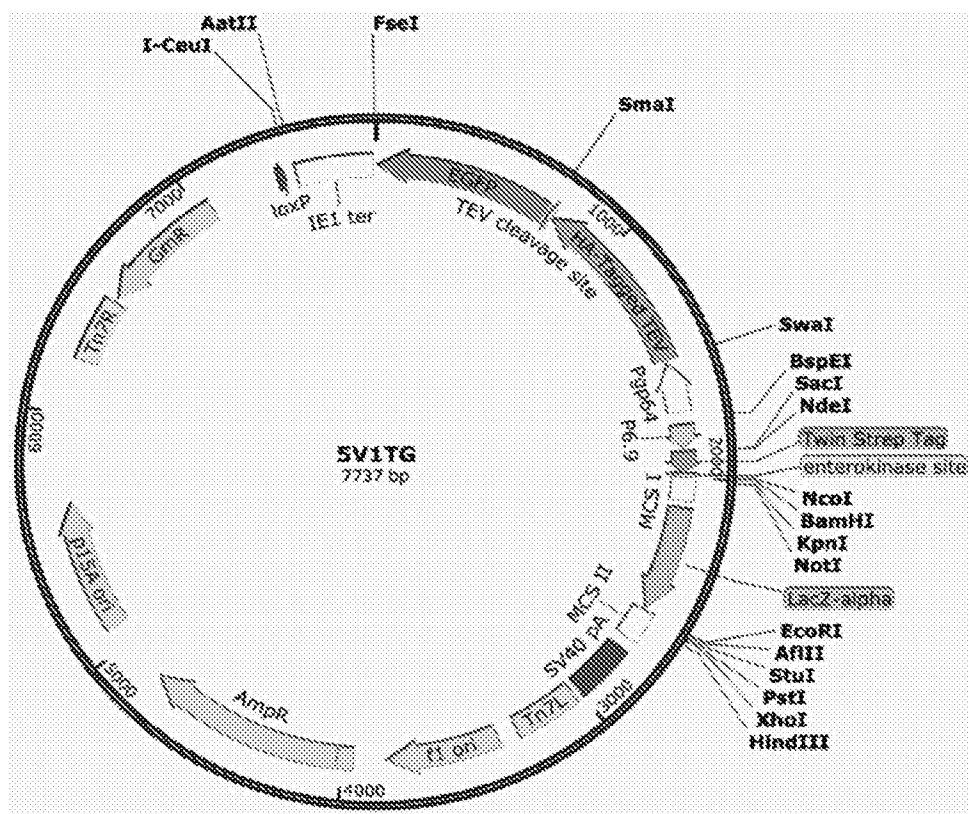
actgattaag cattggtaac tgtcagacca agtttactca tatatacttt agattgattt 1500  
 aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata atctcatgac 1560  
 caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa 1620  
 aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc 1680  
 accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt 1740  
 aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg 1800  
 ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa tcctgttacc 1860  
 agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt 1920  
 accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttgga 1980  
 gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct 2040  
 tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggctcgga caggagagcg 2100  
 cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca 2160  
 cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa 2220  
 cgccagcaac gcggcctttt tacggttccg ggcccttttg tgcccttttg ctcacatgtt 2280  
 ctttctgctg ttatccccctg attctgtgga taaccgtatt accgcctttg agtgagctga 2340  
 taccgctcgc cgcagccgaa cgaccgagcg cagcagtgca gtgagcgagg aagcggaaga 2400  
 gcgcctgatg cgggtattttc tccttacgca tctgtgcggt atttcacacc gcataatatgg 2460  
 tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat 2520  
 cgctacgtga ctgggtcatg gctgcgcccc gacacccgcc aacacccgct gacgcgccct 2580  
 gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct 2640  
 gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct 2700  
 catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcattccgcg tccagctcgt 2760  
 tgagtttctc cagaagcggt aatgtctggc ttctgataaa gcggggccatg ttaagggcgg 2820  
 ttttttccctg tttggctact gatgcctccg tgtaaggggg atttctgttc atgggggtaa 2880  
 tgataccgat gaaacgagag aggatgctca cgatacgggt tactgatgat gaacatgccc 2940  
 ggttactgga acgttgtgag ggtaaacac tggcgggtatg gatgcggcgg gaccagagaa 3000  
 aaatcactca ggggtcaatgc cagcgttcg ttaatacaga tgtaggtgtt ccacagggt 3060  
 gccagcagca tcctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg 3120  
 tttccagact ttacgaaaca cggaaaccga agaccattca tgttggtgct caggtcgcag 3180  
 acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cgggtgattca ttctgctaac 3240  
 cagtaaggca accccgccag cctagccggg tectcaacga caggagcacg atcatgcgca 3300  
 cccgtggggc cgccatgccg gcgataatgg cctgcttctc gccgaaacgt ttggtggcgg 3360  
 gaccagtgc gaaggttga gcgagggcgt gcaagattcc gaataccgca agcgacagge 3420  
 cgatcatcgt cgcgtccag cgaaagcggc cctcgcgaa aatgaccag agcgtgccg 3480  
 gcacctgtcc tacgagttgc atgataaaga agacagtcac aagtgcggcg acgatagtca 3540  
 tgccccgcgc ccaccggaag gagctgactg ggttgaaggc tctcaagggc atcggtcgag 3600  
 atcccgtgc ctaatgagtg agctaactta cattaattgc gttgcgtca ctgcccgtt 3660  
 tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 3720  
 gcggtttgcg tattgggcgc cagggtggtt tttcttttca ccagtgcagc gggcaacagc 3780

```

tgattgccct tcaccgcctg gccctgagag agttgcagca agcgggtccac gctgggtttgc 3840
cccagcaggc gaaaatcctg tttgatgggtg gttaacggcg ggatataaca tgagctgtct 3900
tcggtatcgt cgtatcccac taccgagata tccgcaccaa cgcgcagccc ggactcggta 3960
atggcgcgca ttgcgcccag cgccatctga tcgttggcaa ccagcatcgc agtgggaacg 4020
atgccctcat tcagcatttg catggtttgt tgaaaaccgg acatggcact ccagtcgcct 4080
tcccgttccg ctatcggtg aatttgattg cgagtgcgat atttatgccca gccagccaga 4140
cgcagacgcg ccgagacaga acttaatggg cccgctaaca gcgcgatttg ctgggtgacc 4200
aatgcgacca gatgctccac gccagtcgc gtaccgtctt catgggagaa aataatactg 4260
ttgatgggtg tctggtcaga gacatcaaga aataacgccg gaacattagt gcaggcagct 4320
tccacagcaa tggcatcctg gtcateccag ggatagttaa tgatcagccc actgacgcgt 4380
tgcgcgagaa gattgtgcac cgccgcttta caggcttcga cgccgcttcg ttctaccatc 4440
gacaccacca cgctggcacc cagttgatcg gcgcgagatt taatcgccgc gacaatttgc 4500
gacggcgcggt gcagggccag actggaggtg gcaacgcaa tcagcaacga ctgtttgccc 4560
gccagttgtt gtgccacgcg gttgggaatg taattcagct ccgccatcgc cgtttccact 4620
ttttcccgcg ttttcgcaga aacgtggctg gcctggttca ccacgcggga aacggtctga 4680
taagagacac cggcatactc tgcgacatcg tataacgtta ctggtttcac attcaccacc 4740
ctgaattgac tctcttcgg gcgctatcat gccataccgc gaaaggtttt gcgccattcg 4800
atggtgtccg ggatctcgac gctctccctt atgcgactcc tgcattagga agcagcccag 4860
tagtaggttg aggccgttga gcaccgccgc cgcaaggaat ggtgcatgca aggagatggc 4920
gccaacagt ccccgccga cggggcctgc caccataccc acgccgaaac aagcgtcat 4980
gagccgaag tggcgagccc gatcttcccc atcggtgatg tcggcgatat aggcgccagc 5040
aaccgcacct gtggcgccgg tgatgccggc cacgatgcgt ccggcgtaga ggatcgagat 5100
ctcgatcccc cgaaattaat acgactcact ataggggaat tgtgagcgga taacaattcc 5160
cctctagaaa taattttgtt taactttaag aaggagatat acatatgggc acccgtgacg 5220
acgagtacga ctacctgttc aaggtgggtg tgatcggtga cagcggcggtg ggtaaattcca 5280
acctgctgtc ccgcttcacc cgtaacgagt tcaacctgga gtccaagtcc accatcggcg 5340
tggagttcgc caccgttcc atccaggtgg acggtaaagc catcaaggct cagatctggg 5400
acaccgtgg tctggagcgc taccgcgcca tcacctccgc ttactaccgc ggtgccgtgg 5460
gtgctctcct ggtgtacgac atcgctaagc acctgacctg cgagaacgtg gagcgtggc 5520
tgaaggagct gcgcgaccac gctgactcca acatcgctcat catgctggtg ggcaacaagt 5580
ccgacctgcg ccacctgcgt gctgtgcta ctgacgaagc tcgcgctttc gctgagaaga 5640
acggcctgtc cttcatcgag accagcgtc tggactccac caacgtggag gccgctttcc 5700
agaccatcct caccgagatc tacggtcga gtgggagctc tggaggccac catcaccatc 5760
atcaccatca ctaagtaacc ggetgctaac aaagcccgaa aggaagctga gttggtgtgt 5820
gccaccgtg agcaataact agcataaccc cttggggcct ctaaaccgggt cttgaggggt 5880
tttttgcgtg aaggaggaa tatatccgga t 5911

```





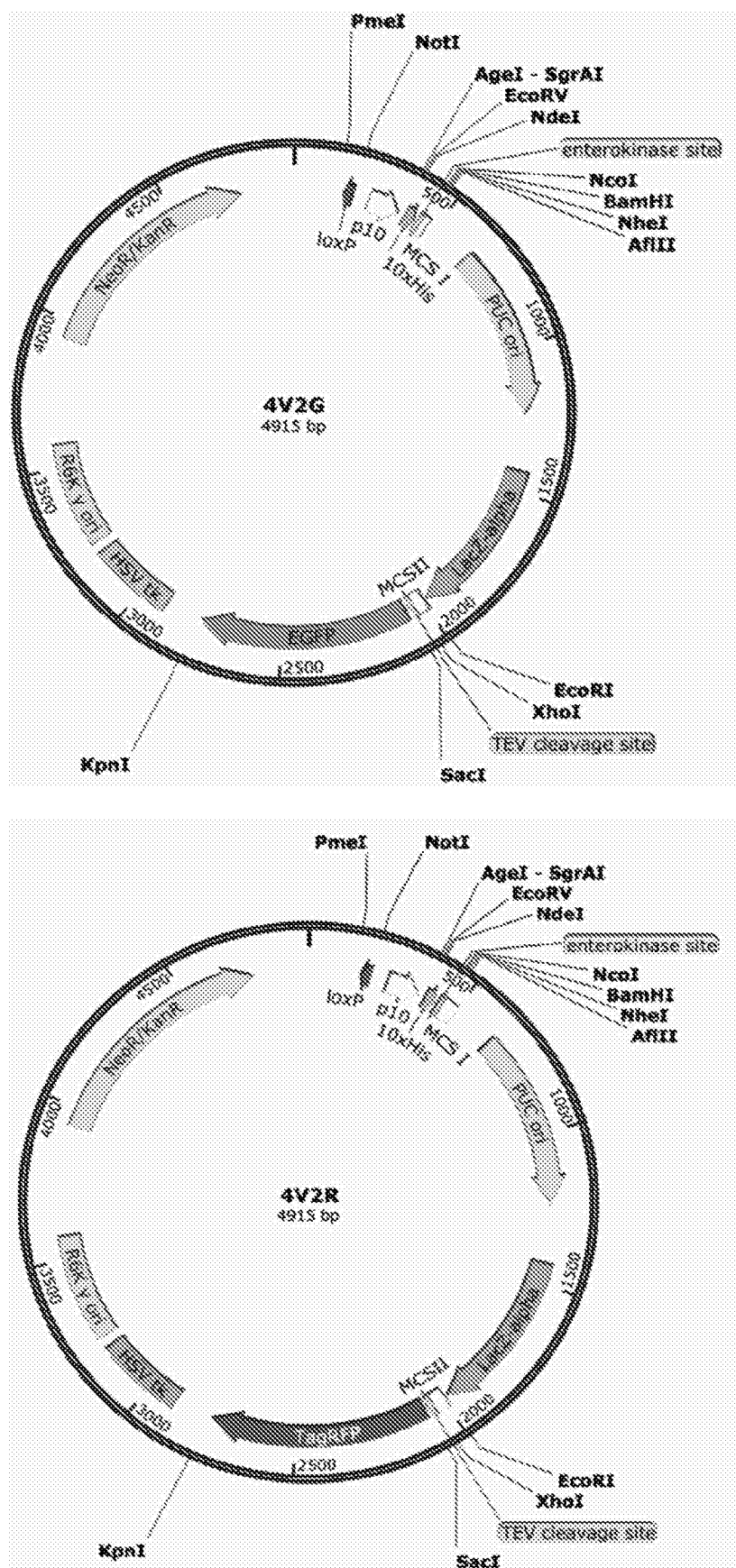


图1

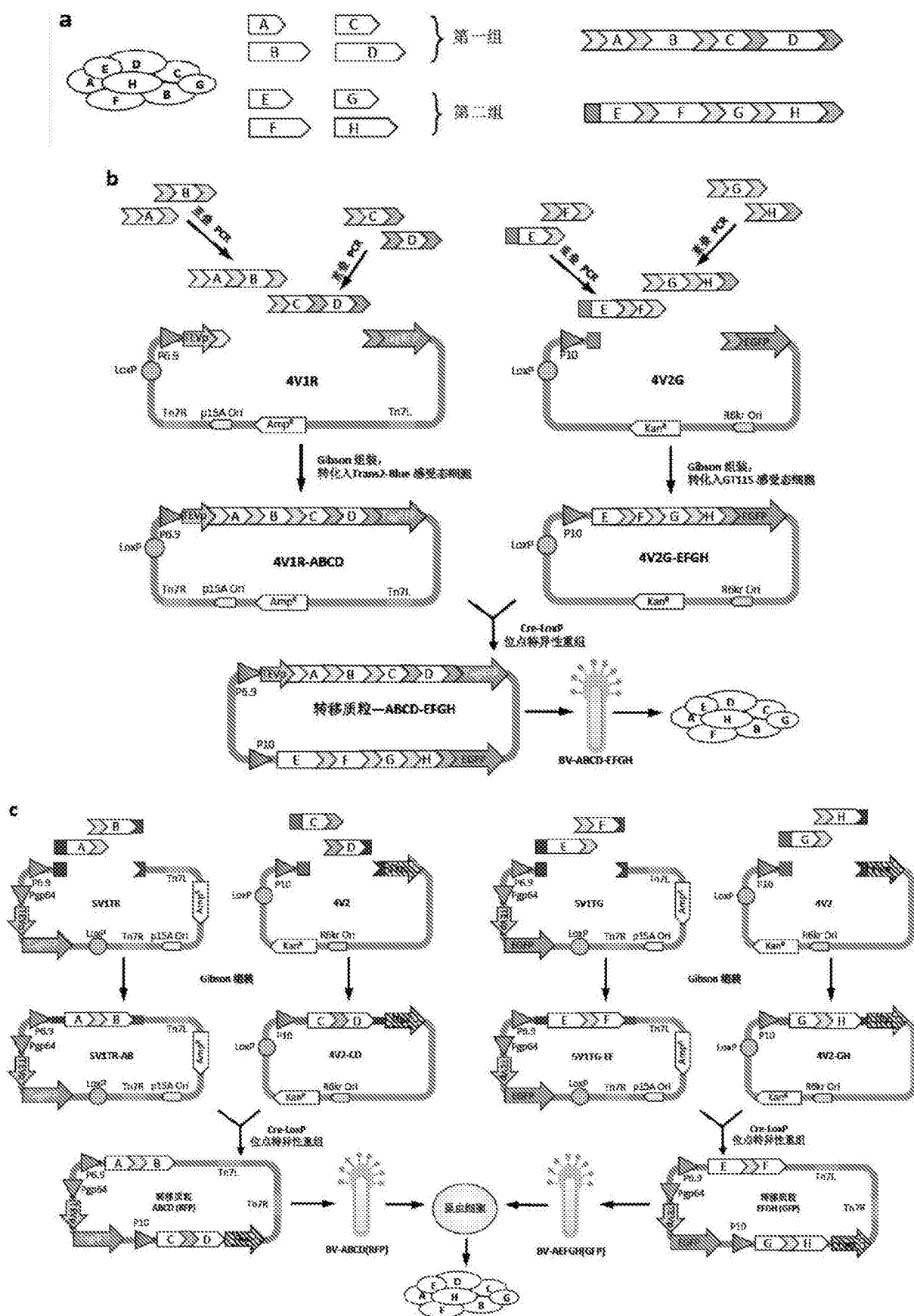


图2

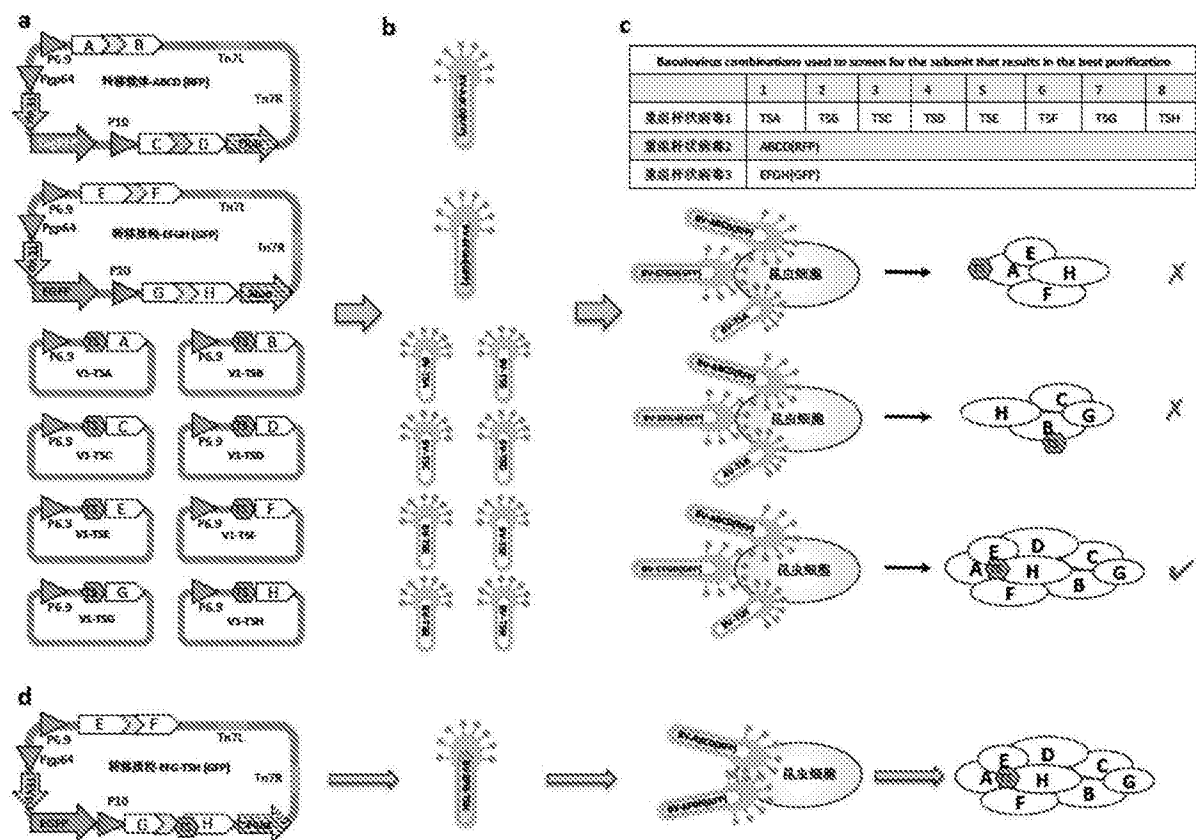


图3

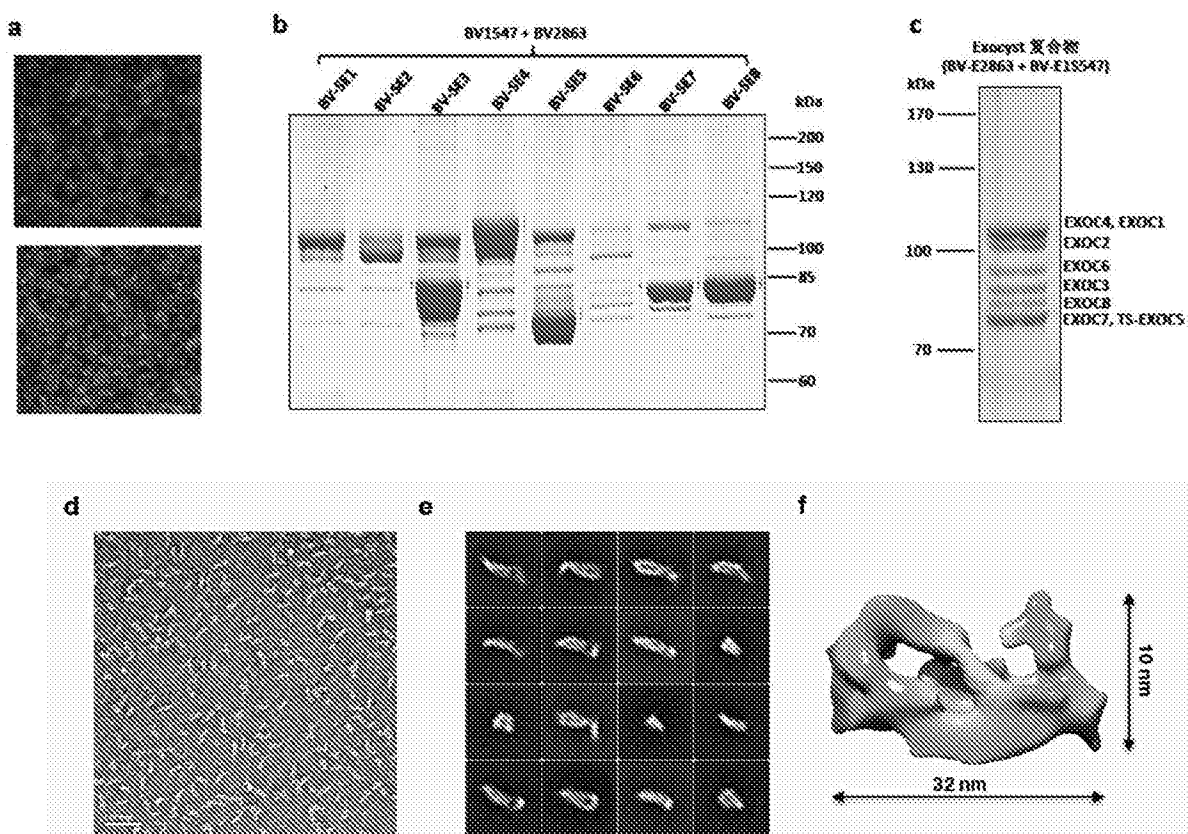


图4

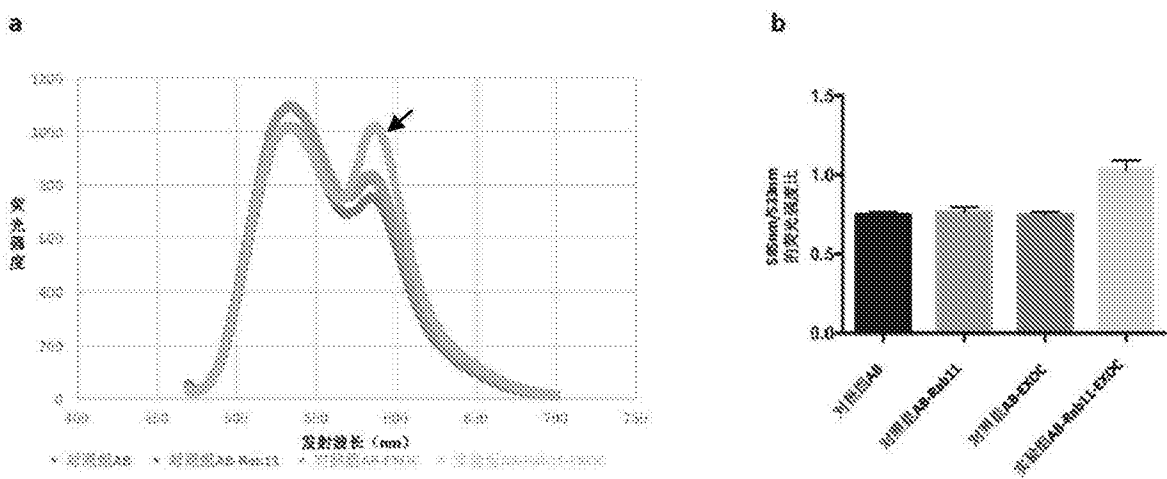


图5