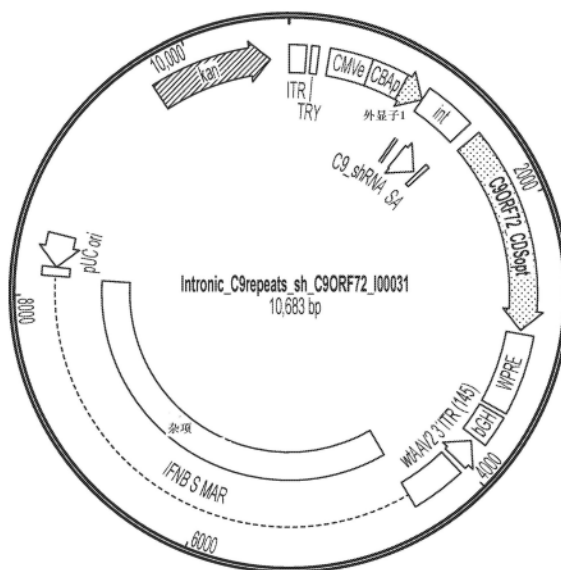


序列表57页 附图22页



1. 一种分离的核酸,其包含
  - (i) 包含编码(a)和(b)的转基因的表达构建体:(a)由SEQ ID NO:51中阐述的序列编码的C9orf72蛋白,以及(b)由SEQ ID NO:39中阐述的序列组成的抑制性核酸,以及
  - (ii) 在所述表达构建体两侧的两个腺相关病毒反向末端重复(ITR)序列。
2. 根据权利要求1所述的分离的核酸,其中所述转基因被可操作连接到启动子。
3. 根据权利要求2所述的分离的核酸,其中所述启动子是鸡 $\beta$ 肌动蛋白(CBA)启动子。
4. 根据权利要求2所述的分离的核酸,其还包含CMV增强子。
5. 根据权利要求1-4中任一项所述的分离的核酸,其还包含土拨鼠肝炎病毒转录后调控元件(WPRE)。
6. 根据权利要求1-4中任一项所述的分离的核酸,其还包含牛生长激素polyA信号尾部。
7. 根据权利要求1-4中任一项所述的分离的核酸,其中每个ITR序列是野生型AAV2 ITR序列。
8. 一种重组腺相关病毒(rAAV)载体,其包括包含编码(i)和(ii)的转基因的表达构建体:
  - (i) 由SEQ ID NO:51中阐述的序列编码的C9orf72蛋白;以及
  - (ii) 由SEQ ID NO:39中阐述的序列组成的抑制性核酸;其中所述表达构建体的两侧带有两个腺相关病毒(AAV)反向末端重复序列(ITR)。
9. 根据权利要求8所述的rAAV载体,其中所述转基因被可操作连接到启动子。
10. 根据权利要求9所述的rAAV载体,其中所述启动子是鸡 $\beta$ 肌动蛋白(CBA)启动子。
11. 根据权利要求9所述的rAAV载体,其还包含CMV增强子。
12. 根据权利要求8-11中任一项所述的rAAV载体,其还包含土拨鼠肝炎病毒转录后调控元件(WPRE)。
13. 根据权利要求8-11中任一项所述的rAAV载体,其还包含牛生长激素polyA信号尾部。
14. 根据权利要求8-11中任一项所述的rAAV载体,其中每个ITR序列是野生型AAV2 ITR序列。
15. 一种重组腺相关病毒(rAAV),其包含:
  - (i) AAV衣壳蛋白;和
  - (ii) 根据权利要求8-14中任一项所述的rAAV载体。
16. 根据权利要求15所述的rAAV,其中所述AAV衣壳蛋白是AAV9衣壳蛋白。
17. 一种重组腺相关病毒(rAAV)载体,其包含核酸,所述核酸以5'至3'顺序包含:
  - (a) 5' AAV ITR;
  - (b) CMV增强子;
  - (c) CBA启动子;
  - (d) 编码C9orf72蛋白和由SEQ ID NO:39中阐述的序列组成的抑制性核酸的转基因,其中所述C9orf72蛋白由SEQ ID NO:51中阐述的核酸序列编码;
  - (e) WPRE;
  - (f) 牛生长激素polyA信号尾部;和

(g) 3' AAVITR。

18. 一种重组腺相关病毒(rAAV), 其包含:

(i) AAV衣壳蛋白; 和

(ii) 根据权利要求17所述的rAAV载体。

19. 根据权利要求18所述的rAAV, 其中所述AAV衣壳蛋白是AAV9衣壳蛋白。

20. 一种质粒, 其包含根据权利要求8-14和17中任一项所述的rAAV载体。

21. 一种杆状病毒载体, 其包含 (i) SEQ ID NO:51中阐述的核酸序列, 以及 (ii) 由SEQ ID NO:39中阐述的序列组成的核酸序列。

22. 一种细胞, 其包含:

(i) 编码一种或多种腺相关病毒rep蛋白和/或一种或多种腺相关病毒cap蛋白的第一载体; 和

(ii) 包含编码由SEQ ID NO:51中阐述的序列编码的C9orf72蛋白和由SEQ ID NO:39中阐述的序列组成的抑制性核酸的表达构建体的第二载体。

23. 根据权利要求22所述的细胞, 其中所述第一载体是质粒, 并且所述第二载体是质粒。

24. 根据权利要求22或23所述的细胞, 其中所述细胞是哺乳动物细胞。

25. 根据权利要求24所述的细胞, 其中所述哺乳动物细胞是HEK293细胞。

26. 根据权利要求22所述的细胞, 其中所述第一载体是杆状病毒载体, 并且所述第二载体是杆状病毒载体。

27. 根据权利要求26所述的细胞, 其中所述细胞是昆虫细胞。

28. 根据权利要求27所述的细胞, 其中所述昆虫细胞是SF9细胞。

29. 一种生产根据权利要求15、16、18或19所述的rAAV的方法, 所述方法包括:

(i) 将编码一种或多种腺相关病毒rep蛋白和/或一种或多种腺相关病毒cap蛋白的第一载体以及包含编码由SEQ ID NO:51中阐述的序列编码的C9orf72蛋白和由SEQ ID NO:39中阐述的序列组成的抑制性核酸的表达构建体的重组AAV载体递送至细胞;

(ii) 在允许包装rAAV的条件下培养所述细胞; 以及

(iii) 收获培养的宿主细胞或培养基以收集所述rAAV。

30. 根据权利要求1-7中任一项所述的分离的核酸、根据权利要求8-14和17中任一项所述的rAAV载体、或根据权利要求15、16、18和19中任一项所述的rAAV在制备用于治疗患有或怀疑患有肌萎缩性脊髓侧索硬化症(ALS) 的对象的药物中的用途。

31. 根据权利要求30所述的用途, 其中所述药物被配制用于直接注射到所述对象的中枢神经系统(CNS)。

32. 根据权利要求31所述的用途, 其中所述直接注射是脑内注射、鞘内注射或其任何组合。

33. 根据权利要求31所述的用途, 其中所述直接注射是实质内注射。

34. 根据权利要求30所述的用途, 其中所述药物被配制用于直接注射到所述对象的脑脊液(CSF)。

35. 根据权利要求34所述的用途, 其中所述直接注射是脑池内注射、脑室内注射, 和/或腰椎内注射。

## 用于神经变性疾病的基因疗法

[0001] 相关申请

[0002] 本申请要求2018年10月8日提交的题为“用于神经变性疾病的基因疗法”(GENE THERAPIES FOR NEURODEGENERATIVE DISEASE)的美国临时申请系列号62/742,723和2017年10月23日提交的题为“用于神经变性疾病的基因疗法”(GENE THERAPIES FOR NEURODEGENERATIVE DISEASE)的62/575,795在35U.S.C.119(e)下的利益,每个所述临时申请的全部内容通过引用并入本文。

### 背景技术

[0003] 在人类中,肌萎缩性脊髓侧索硬化症(ALS)和额颞叶痴呆(FTD)是与C9orf72基因中的六核苷酸重复区的扩增相关的神经变性疾病。通常,与所述C9orf72重复区的扩增相关的疾病由C9orf72蛋白的表达降低和由于有毒RNA病灶的积累导致的功能获得所引起。目前,用于ALS/FTD的治疗选项是有限的。

### 发明内容

[0004] 本公开的多方面涉及可用于治疗神经变性疾病例如肌萎缩性脊髓侧索硬化症(ALS)和/或额颞叶痴呆(FTD)、阿兹海默氏病、高雪氏病、帕金森氏病、路易体痴呆或溶酶体贮积病的组合物和方法。在一些实施方案中,本文描述的方法和组合物用于治疗患有以C9orf72基因的二肽重复区的扩增为特征的ALS/FTD的对象。

[0005] 在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72和/或脊髓小脑性共济失调蛋白2(ATXN2)和/或核糖体蛋白25(RPS25)的表达或活性的抑制性核酸的表达盒。在一些实施方案中,靶向ATXN的抑制性核酸包含SEQ ID NO:10-25中的任一者中阐述的序列或由所述序列组成。在一些实施方案中,靶向C9orf72的抑制性核酸包含SEQ ID NO:37-50中的任一者中阐述的序列或由所述序列组成。

[0006] 在一些方面,本公开提供了一种分离的核酸,所述分离的核酸包含编码密码子优化的C9orf72蛋白(或其部分)的表达盒。在一些实施方案中,密码子优化的C9orf72蛋白包含SEQ ID NO:4中阐述的氨基酸序列。在一些实施方案中,密码子优化的C9orf72蛋白由具有SEQ ID NO:51中阐述的序列的核酸编码。

[0007] 在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72和/或ATXN2和/或RPS25的表达或活性的抑制性核酸,和野生型C9orf72蛋白(例如缺少致病性二肽重复序列(dipeptide repeat)扩增的C9orf72蛋白)的表达盒。在一些实施方案中,野生型C9orf72蛋白由SEQ ID NO:3或其部分编码。在一些实施方案中,野生型C9orf72蛋白包含SEQ ID NO:4中阐述的序列或其部分或由SEQ ID NO:4中阐述的序列或其部分组成。

[0008] 在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72的表达或活性的第一抑制性核酸,和抑制脊髓小脑性共济失调蛋白2(ATXN2)的表达或活性的第二抑制性核酸的表达盒。在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72的表达或活性的第一抑制性核酸,和抑制跨膜蛋白106B(TMEM106B)的表达或活性的第二抑制

性核酸的表达盒。在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72的表达或活性的第一抑制性核酸和抑制RPS25的表达或活性的第二抑制性核酸的表达盒。在一些实施方案中,分离的核酸还包含编码野生型C9orf72蛋白的核酸序列(例如在SEQ ID NO:3中阐述的)。

[0009] 在一些方面,本公开提供了一种分离的核酸,其包含编码抑制C9orf72的表达或活性的抑制性核酸和 $\beta$ -葡萄糖脑苷脂酶 (GBA) 蛋白的表达盒。在一些实施方案中,所述GBA蛋白是GBA1蛋白(例如由GBA1基因或其部分编码的蛋白)。在一些方面,本公开提供了一种分离的核酸,其包含编码抑制ATXN2的表达或活性的抑制性核酸和 $\beta$ -葡萄糖脑苷脂酶 (GBA) 蛋白的表达盒。在一些实施方案中,所述GBA蛋白是GBA1蛋白(例如由GBA1基因或其部分编码的蛋白)。在一些方面,本公开提供了一种分离的核酸,其包含编码抑制TMEM106B的表达或活性的抑制性核酸和 $\beta$ -葡萄糖脑苷脂酶 (GBA) 蛋白的表达盒。在一些实施方案中,所述GBA蛋白是GBA1蛋白(例如由GBA1基因或其部分编码的蛋白)。

[0010] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸、第二抑制性核酸、第三抑制性核酸等)结合到编码C9orf72的二肽重复区的核酸(例如包含二肽重复区的C9orf72 mRNA转录本)。在一些实施方案中,二肽重复区包含一个或多个GGGGCC重复序列(repeat),或一个或多个CCCCGG重复序列(例如C9orf72的二肽重复区)。在一些实施方案中,二肽重复区包含23个或更多的(例如23至10,000之间的任何整数,例如24、25、30、50、100、1000、5000或10,000个)GGGGCC重复序列(例如C9orf72的正义链二肽重复区)或23个或更多的(例如23至10,000之间的任何整数,例如24、25、30、50、100、1000、5000或10,000个)CCCCGG重复序列(例如C9orf72的反义链二肽重复区)。

[0011] 在一些实施方案中,抑制性核酸结合到编码C9orf72的、不是二肽重复区的区域的核酸(例如在C9orf72二肽重复区之外的核酸的部分)。在一些实施方案中,抑制性核酸结合到二肽重复区内的1个核苷酸(例如与二肽重复区相邻的)至约500个核苷酸之间的分离的核酸序列。在一些实施方案中,抑制性核酸靶向编码C9orf72蛋白的基因的内含子区。

[0012] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸、第二抑制性核酸、第三抑制性核酸等)结合到编码ATXN2的核酸序列(例如ATXN2 mRNA转录本),例如在SEQ ID NO:9中所阐述的。在一些实施方案中,靶向ATXN2的抑制性核酸结合到编码ATXN2的核酸序列的非翻译区(例如5' UTR、3' UTR等)。

[0013] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸、第二抑制性核酸、第三抑制性核酸等)结合到编码TMEM106B的核酸序列(例如TMEM106B mRNA转录本),例如在SEQ ID NO:7中所阐述的。在一些实施方案中,靶向TMEM106B的抑制性核酸结合到编码TMEM106B的核酸序列的非翻译区(例如5' UTR、3' UTR等)。

[0014] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸、第二抑制性核酸、第三抑制性核酸等)结合到编码RPS25的核酸序列(例如RPS25 mRNA转录本),例如在SEQ ID NO:60中所阐述的。在一些实施方案中,靶向RPS25的抑制性核酸结合到编码RPS25的核酸序列的非翻译区(例如5' UTR、3' UTR等)。

[0015] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸和/或第二抑制性核酸)是siRNA、shRNA、miRNA和dsRNA。在一些实施方案中,miRNA是人工miRNA(amiRNA),所述人工miRNA包含两侧带有miRNA支架序列例如miR-155支架序列的抑制性核酸序列。

[0016] 在一些实施方案中,抑制性核酸(例如第一抑制性核酸和/或第二抑制性核酸)位于所述表达构建体的非翻译区内。在一些实施方案中,非翻译区是内含子、5' 非翻译区(5' UTR)或3' 非翻译区(3' UTR)。

[0017] 在一些实施方案中,分离的核酸包含一个或多个启动子。在一些实施方案中,启动子是RNA pol III启动子(例如U6或H1)、RNA pol II启动子、鸡β肌动蛋白(CBA)启动子、CAG启动子、CD68启动子或JeT启动子。

[0018] 在一些实施方案中,表达构建体两侧带有两个腺相关病毒(AAV)反向末端重复(ITR)序列。在一些实施方案中,在表达构建体两侧的ITR序列之一缺少有功能的末端解链位点(terminal resolution site)。

[0019] 在一些方面,本公开涉及包含ITR的rAAV载体,所述ITR具有修饰的“D”区(例如相对于野生型AAV2 ITR即SEQ ID NO:32修饰的D序列)。在一些实施方案中,具有修饰的D区的ITR是所述rAAV载体的5' ITR。在一些实施方案中,修饰的“D”区包含“S”序列,例如在SEQ ID NO:29中所阐述的。在一些实施方案中,具有修饰的“D”区的ITR是所述rAAV载体的3' ITR。在一些实施方案中,修饰的“D”区包含3' ITR,其中所述“D”区位于所述ITR的3' 末端处(例如相对于所述载体的转基因插入片段在所述ITR的外侧或末端上)。在一些实施方案中,修饰的“D”区包含SEQ ID NO:29或30中所阐述的序列。

[0020] 在一些实施方案中,分离的核酸(例如rAAV载体)包含TRY区。在一些实施方案中,TRY区包含SEQ ID NO:31中阐述的序列。

[0021] 在一些实施方案中,分离的核酸包含SEQ ID NO:1-62中的任一者中阐述的序列或其部分(或编码SEQ ID NO:1-62中的任一者中阐述的氨基酸序列或其部分)。

[0022] 在一些方面,本公开提供了一种包含本公开所描述的分离的核酸的载体。在一些实施方案中,载体是质粒或病毒载体。在一些实施方案中,病毒载体是重组腺相关病毒载体(rAAV)(例如两侧带有AAV ITR的转基因,所述转基因包含编码一种或多种抑制性核酸的分离的核酸序列和/或编码一种或多种蛋白例如野生型C9orf72和/或GBA1的分离的核酸)或杆状病毒载体。在一些实施方案中,rAAV载体是单链的(例如单链DNA)。

[0023] 在一些方面,本公开提供了一种组合物,其包含本公开所描述的分离的核酸或载体。在一些实施方案中,组合物还包含可药用载剂(carrier)。

[0024] 在一些方面,本公开提供了一种宿主细胞,其包含本公开所描述的分离的核酸或载体。在一些实施方案中,宿主细胞是真核细胞(例如哺乳动物细胞、昆虫细胞等)或原核细胞(例如细菌细胞)。

[0025] 在一些方面,本公开提供了一种重组腺相关病毒(rAAV),其包含衣壳蛋白和本公开所描述的分离的核酸或载体。在一些实施方案中,衣壳蛋白能够跨过血脑屏障。在一些实施方案中,衣壳蛋白是AAV9衣壳蛋白、AAVrh.10衣壳蛋白或AAV-PHP.B衣壳蛋白。在一些实施方案中,rAAV转导中枢神经系统(CNS)的神经元细胞和/或非神经元细胞。

[0026] 在一些方面,本公开提供了一种治疗患有或怀疑患有神经变性障碍(例如肌萎缩性脊髓侧索硬化症(ALS)和/或额颞叶痴呆(FTD)、阿兹海默氏病、高雪氏病、帕金森氏病、路易体痴呆或溶酶体贮积病)的对象的方法,所述方法包括向所述对象施用本公开所描述的分离的核酸、载体、组合物或rAAV。

[0027] 在一些实施方案中,施用包括直接注射到对象的CNS。在一些实施方案中,直接注

射到所述CNS包括直接注射到对象的脑脊液(CSF),例如颅内注射、脑室内注射、腰椎内注射或其任何组合。在一些实施方案中,直接注射是脑内注射、实质内注射、鞘内注射、大池内注射或其任何组合。在一些实施方案中,直接注射包括对流增强递送(CED)。

[0028] 在一些实施方案中,所述对象是哺乳动物例如人类对象。在一些实施方案中,对象的特征在于具有约30至约5000个之间(例如30至5000之间(含)的任何整数)的GGGGCC二肽重复序列和/或约30至5000个之间(例如30至5000之间(含)的任何整数)的CCCCGG重复序列。在一些实施方案中,对象的特征在于具有超过5000个GGGGCC二肽重复序列和/或CCCCGG重复序列。

## 附图说明

[0029] 图1是描绘了包含rAAV载体的质粒的一个实施方案的示意图,所述rAAV载体包括编码靶向C9orf72的重复序列扩增的抑制性核酸、靶向跨膜蛋白106B(TMEM106B)的抑制性核酸和野生型C9orf72编码序列的表达构建体。所述rAAV载体还包含在所述表达构建体两侧的AAV反向末端重复序列。

[0030] 图2是描绘了包含rAAV的质粒的一个实施方案的示意图,所述rAAV包括编码靶向C9orf72的重复序列扩增的抑制性核酸和 $\beta$ -葡萄糖脑苷脂酶(GBA1)编码序列的表达构建体。所述rAAV载体还包含在所述表达构建体两侧的AAV反向末端重复序列。

[0031] 图3是描绘了包含rAAV载体的质粒的一个实施方案的示意图,所述rAAV载体包括编码靶向C9orf72的重复序列扩增的抑制性核酸和野生型C9orf72编码序列的表达构建体。

[0032] 图4是描绘了包含rAAV载体的质粒的一个实施方案的示意图,所述rAAV载体包括编码可操作连接到pol III(H1)启动子的靶向ATXN2(例如编码ATXN2的基因)的抑制性核酸的表达构建体。所述rAAV载体还包含在所述表达构建体两侧的AAV反向末端重复序列。所述3' UTR的“D”序列位于“外侧”位置中。

[0033] 图5是描绘了包含rAAV载体的质粒的一个实施方案的示意图,所述rAAV载体包括编码可操作连接到pol II(CBA)启动子的靶向ATXN2(例如编码ATXN2的基因)的抑制性核酸的表达构建体。所述rAAV载体还包含在所述表达构建体两侧的AAV反向末端重复序列。所述3' UTR的“D”序列位于“外侧”位置中。

[0034] 图6是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到pol II(CBA)启动子的靶向ATXN2(例如编码ATXN2的基因)的抑制性核酸的表达构建体。

[0035] 图7是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到pol II(CBA)启动子的各自靶向ATXN2(例如编码ATXN2的基因)的两个抑制性核酸的表达构建体。

[0036] 图8是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到pol II(CBA)启动子的靶向ATXN2(例如编码ATXN2的基因)的抑制性核酸的表达构建体。

[0037] 图9是描绘了包含rAAV载体的质粒的一个实施方案的示意图,所述rAAV载体包括编码可操作连接到pol II(CBA)启动子的靶向ATXN2(例如编码ATXN2的基因)的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。所述rAAV载体还包含在所述表达构建体两侧的AAV反向末端重复序列。所述3' UTR的“D”序列位于“外侧”位置中。

[0038] 图10是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到

pol II (CBA) 启动子的靶向C9orf72的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0039] 图11是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0040] 图12是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0041] 图13是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol III (例如H1) 启动子的靶向C9orf72的抑制性核酸的表达构建体。

[0042] 图14是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的抑制性核酸的表达构建体。

[0043] 图15是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的两个抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0044] 图16是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的两个抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0045] 图17是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向C9orf72的两个抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0046] 图18是描绘了包含位于ITR的“外侧”(例如相对于转基因插入片段或表达构建体在ITR的末端的近端)的“D”区的rAAV载体(上图)和具有在载体的“内侧”(例如在载体的转基因插入片段的近端)的ITR的野生型rAAV载体的示意图。

[0047] 图19A-19B示出了体外C9orf72表达和敲落测定的代表性数据。图19A示出的代表性数据指示了内源C9orf72被rAAV载体的统计学显著的沉默。图19B示出的代表性数据指示了在用rAAV载体转染后野生型C9orf72表达的统计学显著的提高。

[0048] 图20是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向RPS25的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

[0049] 图21是描绘了质粒的一个实施方案的示意图,所述质粒包含编码可操作连接到 pol II (CBA) 启动子的靶向RPS25的抑制性核酸和编码野生型C9orf72蛋白的密码子优化的核酸序列的表达构建体。

## 具体实施方式

[0050] 本公开的多方面涉及用于治疗神经变性疾病例如ALS/FTD、帕金森氏病、阿兹海默氏病、溶酶体贮积病和路易体痴呆的组合物和方法。本公开部分是基于对象中编码ALS/FTD相关基因产物(例如C9orf72、ATXN2、TMEM106B、靶向上述基因的抑制性核酸等)及其组合的表达构建体。基因产物可以是蛋白、蛋白的片段(例如部分)、抑制ALS/FTD相关基因的干扰

核酸等。在一些实施方案中,基因产物是由ALS/FTD相关基因编码的蛋白或蛋白片段。在一些实施方案中,基因产物是抑制ALS/FTD相关基因的干扰核酸(例如shRNA、siRNA、miRNA、amiRNA等)。

[0051] ALS/FTD相关基因是指编码在遗传、生物化学或功能上与肌萎缩性脊髓侧索硬化症(ALS)、额颞叶痴呆(FTD)或ALS和FTD(ALS/FTD)相关的基因产物的基因。例如,已观察到在C9orf72基因中具有超过23个GGGGCC六核苷酸重复序列的个体与没有重复区扩增的个体相比具有发生ALS/FTD的增加了的风险。在一些实施方案中,本文中描述的表达盒编码野生型或非突变形式的ALS/FTD相关基因(或其编码序列)。通常,基因的“野生型”或“非突变”形式是指编码与正常或非致病性活性相关的蛋白(例如缺少突变或变化例如导致神经变性疾病的发生或进展的重复区扩增的蛋白)的核酸。例如,在一些实施方案中,野生型C9orf72蛋白包含SEQ ID NO:4中阐述的序列或由所述序列组成。

[0052] 分离的核酸和载体

[0053] 分离的核酸可以是DNA或RNA。在一些方面,本公开提供了分离的核酸(例如rAAV载体),其编码靶向一种或多种ALS/FTD相关基因例如C9orf72(例如C9orf72的二肽重复区)、ATXN2、TMEM106B、RPS25等的一种或多种抑制性核酸。抑制性核酸可以靶向基因的正义链(例如从基因转录的mRNA)、基因的反义链(例如从基因转录的mRNA)或基因的正义和反义链两者(例如从基因转录的mRNA)。

[0054] 通常,本文中所描述的分离的核酸可以编码1、2、3、4、5、6、7、8、9、10种或更多种抑制性核酸(例如dsRNA、siRNA、shRNA、miRNA、amiRNA等)。在一些实施方案中,分离的核酸编码超过10种抑制性核酸。在一些实施方案中,所述一种或多种抑制性核酸中的每一者靶向不同基因或基因部分(例如第一miRNA靶向基因的第一靶序列,并且第二miRNA靶向所述基因的不同于所述第一靶序列的第二靶序列)。在一些实施方案中,所述一种或多种抑制性核酸中的每一者靶向同一基因的相同靶序列(例如分离的核酸编码同一miRNA的多个拷贝)。

[0055] 本公开的各方面涉及一种分离的核酸,其包含编码靶向C9orf72蛋白(例如C9orf72 mRNA转录本的二肽重复区)的一种或多种干扰核酸(例如dsRNA、siRNA、miRNA、amiRNA等)的表达构建体。在一些实施方案中,二肽重复区由六核苷酸重复序列GGGGCC的5个或更多个聚合物单元(例如包含所述GGGGCC重复序列的5、6、7、8、9、10、20、30、40、50、100、200、500、1000个或更多个重复序列的区)编码。

[0056] 通常,C9orf72蛋白是指在神经元的细胞质和突触前末梢中发现的一种蛋白,其据认为是小的GTP酶例如Rab的交换因子。在人类中,C9orf72基因位于9号染色体上。在一些实施方案中,所述C9orf72基因编码NCBI参考序列NP\_060795.1表示的肽。在一些实施方案中,C9orf72基因包含SEQ ID NO:3中阐述的序列或编码SEQ ID NO:4中阐述的氨基酸序列。

[0057] 靶向C9orf72的抑制性核酸可以包含长度在6至50个核苷酸之间的互补区(例如与靶基因例如C9orf72或靶基因的一部分例如C9orf72的二肽重复区或二肽重复区之外的区杂交的所述抑制性核酸的区)。在一些实施方案中,抑制性核酸包含长度在约6至30、约8至20或约10至19个核苷酸之间的与C9orf72互补的区。在一些实施方案中,抑制性核酸与C9orf72序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个毗连核苷酸互补。在一些实施方案中,被所述抑制性核酸靶向(例如结合)的C9orf72序列远离所述C9orf72的二肽重复区(相对于它在5'或3')1个核苷酸至500个核苷酸之间(例

如1至500(含)之间的任何整数个核苷酸)。在一些实施方案中,抑制性核酸靶向编码C9orf72蛋白的基因的内含子区(例如非蛋白编码区)。

[0058] 本公开的多方面涉及一种分离的核酸,其包含编码靶向TMEM106B蛋白(例如TMEM106B基因的基因产物)的一种或多种干扰核酸(例如dsRNA、siRNA、miRNA、amiRNA等)的表达构建体。TMEM106B蛋白是指跨膜蛋白106B,其是参与树突形态发生和溶酶体运输的调控的一种蛋白。在人类中,TMEM106B基因位于7号染色体上。在一些实施方案中,所述TMEM106B基因编码由NCBI参考序列NP\_060844.2表示的肽。在一些实施方案中,TMEM106B基因包含SEQ ID NO:7中阐述的序列或编码SEQ ID NO:6中阐述的氨基酸序列。

[0059] 靶向TMEM106B的抑制性核酸可以包含长度在6至50个核苷酸之间的互补区(例如与靶基因例如TMEM106B杂交的所述抑制性核酸的区)。在一些实施方案中,抑制性核酸包含长度在约6至30、约8至20或约10至19个核苷酸之间的与TMEM106B互补的区。在一些实施方案中,抑制性核酸与TMEM106B序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个毗连核苷酸互补。

[0060] 本公开的多方面涉及一种分离的核酸,其包含编码靶向ATXN2蛋白(例如ATXN2基因、也被称为SCA基因的基因产物)的一种或多种干扰核酸(例如dsRNA、siRNA、miRNA、amiRNA等)的表达构建体。ATXN2蛋白是指脊髓小脑性共济失调蛋白2,其是通过它与poly(A)结合蛋白的相互作用而参与调控mRNA翻译的一种蛋白。在人类中,ATXN2基因位于12号染色体上。在一些实施方案中,所述ATXN2基因编码由NCBI参考序列NP\_002964.3表示的肽。在一些实施方案中,ATXN2基因包含SEQ ID NO:9中阐述的序列或编码SEQ ID NO:8中阐述的氨基酸序列。

[0061] 靶向ATXN2的抑制性核酸可以包含长度在6至50个核苷酸之间的互补区(例如与靶基因例如ATXN2杂交的所述抑制性核酸的区)。在一些实施方案中,抑制性核酸包含长度在约6至30、约8至20或约10至19个核苷酸之间的与ATXN2互补的区。在一些实施方案中,抑制性核酸与ATXN2序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个毗连核苷酸互补。

[0062] 本公开的多方面涉及一种分离的核酸,其包含编码靶向核糖体蛋白s25(RPS25)(例如RPS25的基因产物)的一种或多种干扰核酸(例如dsRNA、siRNA、miRNA、amiRNA等)的表达构建体。RPS25蛋白是指一种核糖体蛋白,其是参与蛋白合成的蛋白复合体s40核糖体的亚基。在人类中,RPS25基因位于11号染色体上。在一些实施方案中,所述RPS25基因编码由NCBI参考序列NP\_001019.1表示的肽。在一些实施方案中,RPS25基因包含SEQ ID NO:60中阐述的序列。

[0063] 靶向RPS25的抑制性核酸可以包含长度在6至50个核苷酸之间的互补区(例如与靶基因例如RPS25杂交的所述抑制性核酸的区)。在一些实施方案中,抑制性核酸包含长度在约6至30、约8至20或约10至19个核苷酸之间的与RPS25互补的区。在一些实施方案中,抑制性核酸与RPS25序列的至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个毗连核苷酸互补。

[0064] 本公开的多方面涉及表达构建体,其包含编码一种或多种抑制性核酸(例如靶向C9orf72的二肽重复区的抑制性核酸、靶向C9orf72的非二肽重复区的抑制性核酸和/或靶向TMEM106B的抑制性核酸和/或靶向ATXN2的抑制性核酸和/或靶向RPS25的抑制性核酸等)

的第一基因产物和编码蛋白例如野生型C9orf72蛋白或GBA蛋白的第二基因产物。

[0065] 在一些实施方案中,分离的核酸包含表达盒,其编码抑制C9orf72的表达或活性的第一抑制性核酸和抑制TMEM106B的表达或活性的第二抑制性核酸。

[0066] 在一些实施方案中,分离的核酸包含表达盒,其编码抑制C9orf72的表达或活性的第一抑制性核酸和抑制ATXN2的表达或活性的第二抑制性核酸。

[0067] 在一些实施方案中,分离的核酸包含表达盒,其编码抑制C9orf72的表达或活性的第一抑制性核酸和抑制RPS25的表达或活性的第二抑制性核酸。

[0068] 在一些实施方案中,分离的核酸包含表达盒,其编码抑制C9orf72的表达或活性的抑制性核酸和 $\beta$ -葡萄糖脑苷脂酶 (GBA) 蛋白。在一些实施方案中,所述GBA蛋白是GBA1蛋白(例如由GBA1基因或其部分编码的蛋白)。

[0069] 在一些实施方案中,分离的核酸包含表达盒,其编码抑制C9orf72的表达或活性的抑制性核酸和野生型C9orf72蛋白(例如缺少致病性二肽重复序列扩增的C9orf72蛋白)。在一些实施方案中,编码野生型C9orf72蛋白或其部分的核酸序列是密码子优化的核酸序列。在一些实施方案中,野生型C9orf72蛋白由SEQ ID NO:3中阐述的核酸序列或其部分编码。在一些实施方案中,野生型C9orf72蛋白包含SEQ ID NO:4中阐述的序列或其部分或由SEQ ID NO:4中阐述的序列或其部分组成。在一些实施方案中,编码密码子优化的C9orf72的分离的核酸包含SEQ ID NO:51中阐述的序列或由SEQ ID NO:51中阐述的序列组成。

[0070] 专业技术人员认识到第一基因产物(例如编码C9orf72蛋白或GBA蛋白的核酸序列)和第二基因产物(例如靶向C9orf72、ATXN2、TMEM106B等的抑制性RNA)的表达顺序通常可以颠倒(例如所述抑制性RNA是第一基因产物,并且蛋白编码序列是第二基因产物)。在一些实施方案中,基因产物是基因(例如C9orf72、TMEM106B、ATXN2、GBA1等)的片段(例如部分)。蛋白片段可包含蛋白(例如C9orf72蛋白,GBA蛋白等)的约50%、约60%、约70%、约80%、约90%或约99%。在一些实施方案中,蛋白片段包含C9orf72蛋白或GBA蛋白的50%至99.9%之间(例如50%至99.9%之间的任何值)。在一些实施方案中,基因产物(例如抑制性RNA)杂交到靶基因的部分(例如与靶基因例如C9orf72、ATXN2或TMEM106B的5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21个或更多个毗连核苷酸互补)。

[0071] 在一些实施方案中,表达构建体是单顺反子的(例如所述表达构建体编码包含第一基因产物和第二基因产物的单一融合蛋白)。在一些实施方案中,表达构建体是多顺反子的(例如所述表达构建体编码两个不同的基因产物,例如两个不同的蛋白或蛋白片段)。

[0072] 多顺反子表达载体可以包含一个或多个(例如1、2、3、4、5个或更多个)启动子。可以使用任何适合的启动子,例如组成型启动子、诱导型启动子、内源启动子、组织特异性启动子(例如CNS特异性启动子)等。在一些实施方案中,启动子是鸡 $\beta$ -肌动蛋白启动子(CBA启动子)、CAG启动子(例如由Alexopoulou等,(2008)BMC Cell Biol.9:2;doi:10.1186/1471-2121-9-2所描述的)、CD68启动子或JeT启动子(例如由Tornøe等,(2002)Gene 297(1-2):21-32或Karumuthil-Melethil等,(2016)Human Gene Therapy 27(7):509-521所描述的)。在一些实施方案中,启动子被可操作连接到编码第一基因产物、第二基因产物或第一基因产物和第二基因产物的核酸序列。在一些实施方案中,表达盒包含一个或多个另外的调控序列,包括但不限于转录因子结合序列、内含子剪接位点、poly(A)添加位点、增强子序列、阻遏蛋白结合位点或上述序列的任何组合。

[0073] 在一些实施方案中,编码第一基因产物的核酸序列和编码第二基因产物的核酸序列被编码内部核糖体进入位点(IRES)的核酸序列分隔开。IRES位点的实例被例如Mokrejs等,(2006)Nucleic Acids Res.34(Database issue):D125-30描述。在一些实施方案中,编码第一基因产物的核酸序列和编码第二基因产物的核酸序列被编码自切割肽的核酸序列分隔开。自切割肽的实例包括但不限于T2A、P2A、E2A、F2A、BmCPV 2A和BmIFV 2A,以及被Liu等,(2017)Sci Rep.7:2193所描述的。在一些实施方案中,所述自切割肽是T2A肽。

[0074] 在病理学上,诸如ALS和FTD的障碍与主要由源自于C9orf72基因的与重复序列相关的非ATG(RAN)翻译的蛋白构成的蛋白聚集体的积累相关。因此,在一些实施方案中,本文中描述的分离的核酸包含降低或阻止C9orf72蛋白(例如由具有病原性二肽重复序列扩增的基因编码的C9orf72蛋白)的表达的抑制性核酸。编码抑制性核酸的序列可以被放置在所述表达构建体的非翻译区(例如内含子、5' UTR、3' UTR等)中。

[0075] 在一些实施方案中,抑制性核酸被放置在表达构建体的内含子中,例如在编码第一基因产物的序列上游的内含子中。抑制性核酸可以是双链RNA(dsRNA)、siRNA、micro RNA(miRNA)、人工miRNA(amiRNA)或RNA适体。通常,抑制性核酸结合到靶RNA(例如mRNA)的约6至约30个之间(例如6至30之间的任何整数(含))的毗连核苷酸(例如与所述毗连核苷酸杂交)。在一些实施方案中,所述抑制性核酸分子是miRNA或amiRNA,例如靶向C9orf72(编码病原性C9orf72蛋白的基因)的miRNA。在一些实施方案中,所述miRNA不包含与它所杂交的C9orf72 mRNA的区的任何错配(例如所述miRNA是“完美的”)。在一些实施方案中,miRNA包含与它所杂交的C9orf72 mRNA的区的2至20个之间(例如2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个)的错配,例如“凸起”。在一些实施方案中,miRNA包含与它所杂交的C9orf72mRNA的区的超过20个错配。

[0076] 在一些实施方案中,所述抑制性核酸是shRNA(例如靶向C9orf72的shRNA)。在一些实施方案中,所述抑制性核酸是miRNA(例如靶向C9orf72的miRNA)。在一些实施方案中,表达构建体的一个或多个抑制性RNA的表达由一个或多个RNA pol III启动子例如H1启动子或U6启动子驱动。每个抑制性RNA可以由不同的启动子或相同的启动子驱动。

[0077] 在一些实施方案中,抑制性核酸是人工microRNA(amiRNA)。microRNA(miRNA)通常是指在植物和动物中发现的、并在基因表达的转录和翻译后调控中起作用的小的非编码RNA。MiRNA被RNA聚合酶转录以形成被称为pri-miRNA的发夹-环结构,pri-miRNA随后被酶(例如Drosha、Pasha、剪接体等)加工以形成pre-miRNA发夹结构,其然后被Dicer加工以形成miRNA/miRNA\*双链体(其中\*指示所述miRNA双链体的随从链),其中一条链随后被并入到RNA诱导的沉默复合体(RISC)中。在一些实施方案中,本文中所描述的抑制性RNA是靶向C9orf72(例如C9orf72的二肽重复区或C9orf72的非二肽重复区)、ATXN2或TMEM106B的miRNA。

[0078] 在一些实施方案中,靶向C9orf72的抑制性核酸包含miRNA/miRNA\*双链体。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA链包含SEQ ID NO:24或25、37或38、40或41中的任一者中阐述的序列或其部分,或由SEQ ID NO:24或25、37或38、40或41中的任一者中阐述的序列或其部分组成。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA\*链包含SEQ ID NO:24或25、37或38、40或41中的任一者中阐述的序列或其部分,或由SEQ ID NO:24或25、37或38、40或41中的任一者中阐述的序列或其部分组成。

[0079] 在一些实施方案中,靶向TMEM106B的抑制性核酸包含miRNA/miRNA\*双链体。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA链包含SEQ ID NO:1或7中阐述的序列或其部分,或由SEQ ID NO:1或7中阐述的序列或其部分组成。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA\*链包含SEQ ID NO:1或7中阐述的序列或其部分,或由SEQ ID NO:1或7中阐述的序列或其部分组成。

[0080] 在一些实施方案中,靶向ATXN2的抑制性核酸包含miRNA/miRNA\*双链体。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA链包含SEQ ID NO:10-23中的任一者中阐述的序列或其部分,或由SEQ ID NO:10-23中的任一者中阐述的序列或其部分组成。在一些实施方案中,所述miRNA/miRNA\*双链体的miRNA\*链包含SEQ ID NO:10-23中的任一者中阐述的序列或其部分,或由SEQ ID NO:10-23中的任一者中阐述的序列或其部分组成。

[0081] 人工microRNA(amiRNA)通过修饰本源miRNA(native miRNA)以用感兴趣的靶向区代替pre-mRNA的天然靶向区来衍生。例如,可以使用天然存在的表达的miRNA作为支架或骨架(例如pri-miRNA支架),并将茎序列用靶向感兴趣的基因的miRNA的茎序列代替。人工前体microRNA(pre-amiRNA)通常被加工,以便优选地产生单一稳定的小RNA。在一些实施方案中,本文中描述的scAAV载体和scAAV包含编码amiRNA的核酸。在一些实施方案中,所述amiRNA的pri-miRNA支架源自于选自pri-MIR-21、pri-MIR-22、pri-MIR-26a、pri-MIR-30a、pri-MIR-33、pri-MIR-122、pri-MIR-375、pri-MIR-199、pri-MIR-99、pri-MIR-194、pri-MIR-155和pri-MIR-451的pri-miRNA。在一些实施方案中,amiRNA包含靶向C9orf72、ATXN2或TMEM106B的核酸序列和eSIBR amiRNA支架,例如在Fowler等,Nucleic Acids Res.2016Mar 18;44(5):e48中所描述的。

[0082] 在一些方面,本公开涉及表达构建体,其包含用于治疗神经变性疾病(例如ALS/FTD)的抑制性RNA的组合。例如,在一些实施方案中,由本公开描述的表达构建体包含靶向C9orf72的抑制性RNA和靶向跨膜蛋白106B(TMEM106B)的抑制性RNA。所述分离的核酸编码所述抑制性核酸的序列的顺序可以变化。例如,分离的核酸可以从5'末端到3'末端编码靶向C9orf72和TMEM106B,或靶向TMEM106B和C9orf72的shRNA。

[0083] 本文中描述的分离的核酸可以自身存在或作为载体的一部分存在。通常,载体可以是质粒、粘粒、噬菌粒、细菌人工染色体(BAC)或病毒载体(例如腺病毒载体、腺相关病毒(AAV)载体、反转录病毒载体、杆状病毒载体等)。在一些实施方案中,所述载体是质粒(例如包含本文中描述的分离的核酸的质粒)。在一些实施方案中,所述载体是重组AAV(rAAV)载体(例如两侧带有AAV ITR的编码转基因的表达构建体)。在一些实施方案中,rAAV载体是单链的(例如单链DNA)。在一些实施方案中,载体是杆状病毒载体(例如苜蓿银纹夜蛾核多角体病毒(AcNPV)载体)。

[0084] 通常,rAAV载体包含两侧带有两个AAV反向末端重复(ITR)序列的转基因。在一些实施方案中,所述rAAV载体的转基因包含本公开所描述的分离的核酸。在一些实施方案中,所述rAAV载体的两个ITR序列中的每一者是全长ITR(例如长度约为145bp,并含有有功能的Rep结合位点(RBS)和末端解链位点(trs))。在一些实施方案中,所述rAAV载体的ITR之一是截短的(例如缩短或非全长的)。在一些实施方案中,截短的ITR缺少有功能的末端解链位点(trs)并被用于生产自我互补的AAV载体(scAAV载体)。在一些实施方案中,截短的ITR是 $\Delta$ ITR,例如由McCarty等,(2003)Gene Ther.10(26):2112-8所描述的。

[0085] 本公开的多方面涉及分离的核酸(例如rAAV载体),其包含相对于野生型AAV ITR,例如相对于野生型AAV2 ITR(例如SEQ ID NO:32)具有一个或多个修饰(例如核酸添加、缺失、替换等)的ITR。野生型AAV2 ITR的结构示出在图18中。通常,野生型ITR包含125个核苷酸的区,其自我退火形成回文双链T形发夹结构,所述发夹结构由两个横臂(分别由被称为B/B'和C/C'的序列形成)、较长的茎区(由序列A/A'形成)和被称为“D”区的单链末端区组成(图18)。通常,所述ITR的“D”区位于所述由A/A'序列形成的茎区与含有所述rAAV载体的转基因的插入片段之间(例如相对于ITR的末端位于所述ITR的“内侧”或所述rAAV载体的转基因插入片段或表达构建体的近端)。在一些实施方案中,“D”区包含SEQ ID NO:30中阐述的序列。已观察到所述“D”区在rAAV载体被衣壳蛋白的衣壳化中发挥重要作用,例如由Ling等,(2015)J Mol Genet Med 9(3)所公开的。

[0086] 本公开部分是基于下述令人吃惊的发现,即包含位于ITR的“外侧”(例如相对于转基因插入片段或表达构建体位于ITR的末端的近端)上的“D”区的rAAV载体,与ITR具有未修饰的(例如野生型)ITR的rAAV载体相比高效地被AAV衣壳蛋白衣壳化。在一些实施方案中,具有修饰的“D”序列(例如在“外侧”位置中的“D”序列)的rAAV载体相对于具有野生型ITR序列的rAAV载体具有降低的毒性。

[0087] 在一些实施方案中,修饰的“D”序列相对于野生型“D”序列(例如SEQ ID NO:30)包含至少一个核苷酸替换。修饰的“D”序列相对于野生型“D”序列(例如SEQ ID NO:30)可以具有至少1、2、3、4、5、6、7、8、9、10个或超过10个核苷酸替换。在一些实施方案中,修饰的“D”序列相对于野生型“D”序列(例如SEQ ID NO:30)包含至少10、11、12、13、14、15、16、17、18或19个核苷酸替换。在一些实施方案中,修饰的“D”序列与野生型“D”序列(例如SEQ ID NO:30)具有约10%至约99%之间(例如10%、15%、20%、25%、30%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或99%)的一致性。在一些实施方案中,修饰的“D”序列包含SEQ ID NO:29中阐述的序列,其也被称为“S”序列,正如在Wang等,(1995)J Mol Biol 250(5):573-80中所描述的。

[0088] 本公开描述的分离的核酸或rAAV载体还可以包含“TRY”序列,例如在SEQ ID NO:31中所阐述的,正如由Francois等,2005,细胞的TATA结合蛋白为最小腺相关病毒2型p5元件的Rep依赖性复制所需(The Cellular TATA Binding Protein Is Required for Rep-Dependent Replication of a Minimal Adeno-Associated Virus Type 2p5 Element),J Virol.所描述的。在一些实施方案中,TRY序列位于分离的核酸或rAAV载体的ITR(例如5' ITR)与表达构建体(例如编码转基因的插入片段)之间。

[0089] 在一些方面,本公开涉及杆状病毒载体,其包含本公开所描述的分离的核酸或rAAV载体。在一些实施方案中,所述杆状病毒载体是苜蓿银纹夜蛾核多角体病毒(AcNPV)载体,例如由Urabe等,(2002)Hum Gene Ther13(16):1935-43和Smith等,(2009)Mol Ther17(11):1888-1896所描述的。

[0090] 在一些方面,本公开提供了一种宿主细胞,其包含本文中所描述的分离的核酸或载体。宿主细胞可以是原核细胞或真核细胞。例如,宿主细胞可以是哺乳动物细胞、细菌细胞、酵母细胞、昆虫细胞等。在一些实施方案中,宿主细胞是哺乳动物细胞例如HEK293T细胞。在一些实施方案中,宿主细胞是细菌细胞例如大肠杆菌细胞。

[0091] rAAV

[0092] 在一些方面,本公开涉及重组AAV(rAAV),其包含编码本文中所描述的核酸的转基因(例如本文中所描述的rAAV载体)。术语“rAAV”通常是指包含被一种或多种AAV衣壳蛋白衣壳化的rAAV载体的病毒粒子。本公开描述的rAAV可以包含具有选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9和AAV10的血清型的衣壳蛋白。在一些实施方案中,rAAV包含来自于非人类宿主的衣壳蛋白,例如恒河猴的AAV衣壳蛋白例如AAVrh.10、AAVrh.39等。在一些实施方案中,本公开描述的rAAV包含作为野生型衣壳蛋白的变体的衣壳蛋白,例如相对于它所源自的野生型AAV衣壳蛋白包括至少1、2、3、4、5、6、7、8、9、10个或超过10个(例如15、20、25、50、100个等)氨基酸替换(例如突变)的衣壳蛋白变体。

[0093] 在一些实施方案中,本公开描述的rAAV容易地通过CNS传播,特别是当被引入到CSF空间或直接引入到脑实质中时。因此,在一些实施方案中,本公开描述的rAAV包含能够跨过血脑屏障(BBB)的衣壳蛋白。例如,在一些实施方案中,rAAV包含具有AAV9或AAVrh.10血清型的衣壳蛋白。在一些实施方案中,rAAV包含跨过血脑屏障的AAV9变体例如AAV-PHP.B血清型,正如由Deverman等,(2016) *Nature Biotechnology* 34:204-209所描述的。一般来说,rAAV的生产被例如Samulski等,(1989) *J Virol* 63 (9):3822-8和Wright (2009) *Hum Gene Ther* 20 (7):698-706描述。

[0094] 在一些实施方案中,本公开描述的rAAV(例如包含被AAV衣壳蛋白衣壳化以形成rAAV衣壳粒子的重组rAAV基因组)在杆状病毒载体表达系统(BEVS)中生产。使用BEVS生产rAAV被例如Urabe等,(2002) *Hum Gene Ther* 13 (16):1935-43,Smith等,(2009) *Mol Ther* 17 (11):1888-1896,美国专利号8,945,918,美国专利号9,879,282和国际PCT公布W0 2017/184879描述。然而,rAAV可以使用任何适合的方法(例如使用重组rep和cap基因)来生产。

[0095] 药物组合物

[0096] 在一些方面,本公开提供了包含本文中所描述的分离的核酸或rAAV和可药用载剂的药物组合物。当在本文中使用时,术语“可药用”是指材料例如载剂或稀释剂,所述材料不废除化合物的生物活性或性质,并且是相对无毒性的,例如所述材料可以被施用到个体而不引起不想要的生物学效应或不以有害的方式与包含它的组合物的任何组分相互作用。

[0097] 当在本文中使用时,术语“可药用载剂”意味着可药用材料、组合物或载剂例如液体或固体填充剂、稳定剂、分散剂、悬浮剂、稀释剂、赋形剂、增稠剂、溶剂或包胶材料,它们参与将本发明中的有用化合物在患者内携带或运输,或携带或运输到患者,使得所述化合物可以执行其目标功能。可以包括在本发明的实践中使用的药物组合物中的其他成分在本领域中是已知的,并描述在例如《Remington制药学》(Remington's Pharmaceutical Sciences) (Genaro主编,Mack Publishing Co.,1985,Easton,PA)中,其通过引用并入本文。

[0098] 本文中提供的组合物(例如药物组合物)可以通过任何途径施用,包括肠(例如口服)、肠胃外、静脉内、肌肉内、动脉内、髓内、鞘内、大池内、皮下、脑室内、透皮、真皮内、直肠、阴道内、腹膜内、局部(如通过粉剂、软膏、霜剂和/或滴剂)、粘膜、鼻、颊、舌下,通过气管内滴注、支气管滴注和/或吸入,和/或作为口喷剂、鼻喷剂和/或气溶胶。具体设想的途径是口服施用、静脉内施用(例如系统性静脉内注射)、通过血液和/或淋巴供应的区域性施用和/或向患病部位的直接施用。通常,最适合的施途径取决于各种因素,包括所述药剂的本质(例如它在胃肠道环境中的稳定性)和/或对象的状况(例如所述对象是否能够耐受口

服施用)。

[0099] 方法

[0100] 本公开部分是基于在对象中一个或多个ALS-FTD相关基因产物(或其组合)的表达的组合物,所述一个或多个ALS-FTD相关基因产物(或其组合)一起作用(例如协同)以治疗神经变性疾病(例如ALS/FTD等)。当在本文中使用“治疗”(“treat”或者“treating”)是指(a)阻止或延迟神经变性疾病(例如ALS/FTD、阿兹海默氏病、高雪氏病、帕金森氏病、路易体痴呆、溶酶体贮积病等)的发作;(b)减轻神经变性疾病的严重性;(c)减轻或阻止神经变性疾病特有的症状的发生;(d)和/或阻止神经变性疾病特有的症状的恶化。例如,ALS/FTD的症状包括例如运动功能障碍(例如麻痹、摇晃、僵硬、运动缓慢、行走困难)、认知功能障碍(例如痴呆、抑郁、焦虑)、情绪和行为功能障碍。

[0101] 因此,在一些方面,本公开提供了一种治疗患有或怀疑患有神经变性疾病的对象的方法,所述方法包括向所述对象施用本公开所描述的组合物(例如包含分离的核酸或载体或rAAV的组合物)。在一些实施方案中,所述神经变性疾病是ALS/FTD、阿兹海默氏病、高雪氏病、帕金森氏病、路易体痴呆或溶酶体贮积病。

[0102] 在一些实施方案中,组合物被直接施用到所述对象的CNS,例如通过直接注射到所述对象的脑和/或脊髓中。CNS直接施用方式的实例包括但不限于脑内注射、脑室内注射、颅内注射、实质内注射、鞘内注射和上述方式的任何组合。在一些实施方案中,直接注射到对象的CNS中引起转基因在所述对象的中脑、纹状体和/或大脑皮层中的表达(例如第一基因产物、第二基因产物以及如果适用的话第三基因产物的表达)。

[0103] 在一些实施方案中,本公开所描述的组合物被直接施用到对象的脑脊液(CSF)。在一些实施方案中,直接注射到CNS中引起转基因在所述对象的脊髓和/或CSF中的表达(例如第一基因产物、第二基因产物以及如果适用的话第三基因产物的表达)。向对象的CSF的直接注射的实例包括但不限于颅内注射、脑室内注射、腰椎内注射或其任何组合。

[0104] 在一些实施方案中,直接注射到对象的CNS包括对流增强递送(CED)。对流增强递送是一种治疗策略,其包括将脑手术暴露并将小直径的导管直接放置在脑的靶区域中,然后将治疗剂(例如本文中所描述的组合物或rAAV)直接输注到所述对象的脑中。CED被例如Debinski等,(2009)Expert Rev Neurother.9(10):1519-27描述。

[0105] 在一些实施方案中,组合物被外周施用到对象,例如通过外周注射。外周注射的实例包括皮下注射、静脉内注射、动脉内注射、腹膜内注射或上述方式的任何组合。在一些实施方案中,所述外周注射是动脉内注射,例如注射到对象的颈动脉中。

[0106] 在一些实施方案中,本公开所描述的组合物(例如包含分离的核酸或载体或rAAV的组合物)通过外周和直接两种方式被施用到对象的CNS。例如,在一些实施方案中,对象通过动脉内注射(例如注射到颈动脉中)并通过实质内注射(例如通过CED的实质内注射)来施用组合物。在一些实施方案中,所述向CNS的直接注射和外周注射是同时的(例如在同一时间发生)。在一些实施方案中,所述直接注射发生在所述外周注射之前(例如1分钟至1周或更长时间之前)。在一些实施方案中,所述直接注射发生在所述外周注射之后(例如1分钟至1周或更长时间之后)。

[0107] 本公开所描述的组合物(例如包含分离的核酸或载体或rAAV的组合物)被施用到对象的量将随着施用方法而变。例如,在一些实施方案中,本文中所描述的rAAV以约 $10^9$ 个

基因组拷贝 (GC) /kg 至约  $10^{14}$  GC/kg 之间 (例如约  $10^9$  GC/kg、约  $10^{10}$  GC/kg、约  $10^{11}$  GC/kg、约  $10^{12}$  GC/kg、约  $10^{13}$  GC/kg 或约  $10^{14}$  GC/kg) 的滴度施用到对象。在一些实施方案中,对象通过注射到CSF空间或通过实质内注射被高滴度施用 (例如  $>10^{12}$  个基因组拷贝GC/kg的rAAV)。

[0108] 本公开所描述的组合物 (例如包含分离的核酸或载体或rAAV的组合物) 可以被一次或多次 (例如2、3、4、5、6、7、8、9、10、20次或更多次) 施用到对象。在一些实施方案中,组合物被连续 (例如长期) 施用到对象,例如通过输注泵。

[0109] 实施例

[0110] 实施例1:rAAV载体

[0111] AAV载体使用细胞例如用于三质粒转染的HEK293细胞来产生。ITR序列在表达构建体的两侧,所述表达构建体包含用于每个感兴趣的转基因的启动子/增强子元件、3' polyA信号和翻译后信号例如WPRE元件。多个基因产物例如C9orf72蛋白或GBA1蛋白和一种或多种抑制性核酸 (例如靶向C9orf72和/或TMEM106B的抑制性核酸) 可以被同时表达,例如通过使用单一表达盒或分开的表达盒来表达。在所述被表达基因上游的高效剪接的短内含子序列的存在可以提高表达水平。抑制性RNA (例如miRNA、shRNA等) 和其他调控性RNA可以潜在地被包含在这些序列中。本公开所描述的表达构建体的实例示出在图1-17和20-21以及下面的表1中。

[0112] 表1

[0113]

| 名称                                                                       | 启动子 1 | shRNA            | CDS1    | PolyA 1  | 启动子 2 | CDS2 | PolyA2 | ITR 之间的长度 |
|--------------------------------------------------------------------------|-------|------------------|---------|----------|-------|------|--------|-----------|
| PrevailVector_13_CMVe_CBAP_shRNAC9_mRNAiTMEM106B_C9ORF72_WPRE_bGH_4993nt | CBA   | C9重复序列和 TMEM106B | C9ORF72 | WPRE_bGH |       |      |        | 4993      |
| PrevailVector_13H_H1_C9sh_CMVe_CBAP_c9ORF72_WPRE_bGH_3944nt              | H1    | C9重复序列           | C9orf72 |          | CBA   |      |        | 3944      |
| PrevailVector_0_CMVe_CBAP_shRNAC9_GBA1_WPRE_bGH_3892nt                   | CBA   | C9orf72          | GBA1    | WPRE_bGH |       |      |        | 3892      |
| H1_ATXN2_sh_sen                                                          | H1    | ATXN2            |         |          |       |      |        | 899       |
| Intronic_ATXN2_sh_sen                                                    | CBA   | ATXN2            |         | WPRE_bGH |       |      |        | 2547      |
| Intronic_C9repeats_ATNX2_sh_C9ORF71_I0037                                | CBA   | C9重复序列_ATXN2     | C9orf72 | WPRE_bGH |       |      |        | 4145      |
| Intronic_eSIBR_Anti_C9_IntronicShL3_Intro nicShR3_CMVe_CBA               | CBA   | C9orf72          | C9orf72 | WPRE_bGH |       |      |        | 4140      |

[0114]

|                                                                                      |     |         |         |          |  |  |  |      |
|--------------------------------------------------------------------------------------|-----|---------|---------|----------|--|--|--|------|
| p_C9ORF72_WPRE_bGH_I00135                                                            |     |         |         |          |  |  |  |      |
| Intronic_eSIBR_Anti_C9_IntronicShL1_Intro nicShR1_CMVe_CBA p_C9ORF72_WPRE_bGH_I00133 | CBA | C9orf72 | C9orf72 | WPRE_bGH |  |  |  | 4143 |
| Intronic_C9repeatshRNA_I00066                                                        | CBA | C9orf72 |         | WPRE_bGH |  |  |  | 2250 |
| Intronic_eSIBR_Anti_C9_IntronicShL2_Intro nicShR2_CMVe_CBA p_C9ORF72_WPRE_bGH_I00134 | CBA | C9orf72 | C9orf72 | WPRE_bGH |  |  |  | 4142 |
| Intronic_C9repeats_sh_C9ORF72_I00031                                                 | CBA | C9orf72 | C9orf72 | WPRE_bGH |  |  |  | 3997 |
| Intronic_C9validated_sh_C9ORF72_I00032                                               | CBA | C9orf72 | C9orf72 | WPRE_bGH |  |  |  | 3994 |
| H1_C9repeatshRNA_I00069                                                              | H1  | C9orf72 |         |          |  |  |  | 902  |
| Intronic_C9conserved_sh_C9ORF72_I00030                                               | CBA | C9orf72 | C9orf72 | WPRE_bGH |  |  |  | 3994 |

[0115] 实施例2:病毒在ALS/FTD细胞中的转导的基于细胞的测定

[0116] 以C9orf72的重复序列扩增为特征的细胞作为例如来自于ALS/FTD患者的成纤维细胞、单核细胞或hES细胞或患者来源的诱导多能干细胞(iPSC)而被获得。这些细胞积累RNA病灶和RAN翻译的蛋白。

[0117] 使用这些细胞模型,根据蛋白聚集体例如RAN蛋白的聚集体的积累,使用抗RAN蛋白抗体,然后使用荧光显微术成像来定量细胞疾病。使用Western印迹和/或ELISA来定量RAN蛋白的异常积累。

[0118] 在AAV载体的转导的表达的背景下测量治疗终点(例如ALS/FTD相关疾病的减轻),以确认并定量活性和功能。内源(例如致病性的含有重复序列扩增的)C9orf72 mRNA水平的降低可以例如使用定量RT-PCT(qRT-PCR)来定量。

[0119] 实施例3:体外研究

[0120] 本实施例描述了本公开描述的C9orf72 rAAV载体的体外测试。在哺乳动物细胞中研究了C9orf72敲落和过表达的效果。测试的构建体的实例列于表2中。

[0121]

| ID     | 启动子          | 敲落               | 启动子 | 过表达    |
|--------|--------------|------------------|-----|--------|
| I00017 | H1           | C9_sh            | CMV | opt-C9 |
| I00018 | CMV          | C9_sh,TMEM_mi    | CMV | opt-C9 |
| I00030 | CMV_intronic | C9_sh(cons)      | CMV | opt-C9 |
| I00031 | CMV_intronic | C9repeat_sh      | CMV | opt-C9 |
| I00032 | CMV_intronic | C9_sh(validated) | CMV | opt-C9 |
| I00037 | CMV_intronic | C9r_sh,ATXN_sh   | CMV | opt-C9 |

[0122] 通过定量PCR(qPCR)和ELISA测定了基因敲落和过表达。图19A示出的代表性数据表明C9orf72被rAAV载体的统计学显著的沉默。图19B示出的代表性数据表明在用rAAV载体转染后野生型C9orf72表达的统计学显著的提高。

[0123] 序列

[0124] 在一些实施方案中,编码一个或多个基因产物(例如第一、第二和/或第三基因产物)的表达盒包含SEQ ID NO:1-62中的任一者中阐述的序列或由所述序列组成(或编码所

述序列)。在一些实施方案中,基因产物包含SEQ ID NO:1-62中的任一者的部分(例如片段)或由SEQ ID NO:1-62中的任一者的部分(例如片段)组成(或由SEQ ID NO:1-62中的任一者的部分(例如片段)编码)。在一些实施方案中,下述序列中的“T”核苷酸被“U”核苷酸代替,例如在RNA分子的情形中。

[0125] 专业技术人员将会认识到,上述序列的“部分”可以是所述表达盒的缺少质粒骨架(例如复制原点序列、选择标志物序列等)的序列(例如编码ITR的序列、干扰RNA、编码序列、调控序列等)。

|        |                                                                        |
|--------|------------------------------------------------------------------------|
| [0001] | 序列表                                                                    |
| [0002] | <110> 普利维尔治疗公司(Prevail Therapeutics, Inc. )                            |
| [0003] | <120> 用于神经变性疾病的基因疗法                                                    |
| [0004] | <130> P1094.70005W000                                                  |
| [0005] | <140> 尚未指派                                                             |
| [0006] | <141> 2018-10-23                                                       |
| [0007] | <150> 62/742723                                                        |
| [0008] | <151> 2018-10-08                                                       |
| [0009] | <150> 62/575795                                                        |
| [0010] | <151> 2017-10-23                                                       |
| [0011] | <160> 62                                                               |
| [0012] | <170> PatentIn version 3.5                                             |
| [0013] | <210> 1                                                                |
| [0014] | <211> 10953                                                            |
| [0015] | <212> DNA                                                              |
| [0016] | <213> 人工序列                                                             |
| [0017] | <220>                                                                  |
| [0018] | <223> 合成的多核苷酸                                                          |
| [0019] | <400> 1                                                                |
| [0020] | ttggccactc cctctctgcg cgctcgcctc ctcactgagg ccgggcgacc aaaggtcgcc 60   |
| [0021] | cgacgcccg gctttgccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg 120     |
| [0022] | gccaaactcca tctactagggg ttctctgtag ctctgggtat ttaagccga gtgagcacgc 180 |
| [0023] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc 240  |
| [0024] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag 300  |
| [0025] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagtgc 360  |
| [0026] | cgcgttacat aacttacggt aaatggccc cctggctgac cgcccaacga cccccgcca 420    |
| [0027] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt 480   |
| [0028] | caatgggtgg actatttacg gtaaatgcc cacttggcag tacatcaagt gtatcatatg 540   |
| [0029] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgcctggca ttatgccag 600   |
| [0030] | tacatgacct tatgggactt tctacttg cagtacatct acgtattagt categctatt 660    |
| [0031] | accatggctg aggtgagccc cagttctgc ttcactctcc ccatctcccc cccctcccca 720   |
| [0032] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg 780  |
| [0033] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 840  |
| [0034] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg 900   |
| [0035] | cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg 960  |
| [0036] | ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcgcc ggctctgact 1020 |
| [0037] | gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 1080  |
| [0038] | gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcgggg 1140  |
| [0039] | ccggtttcgg ggccttagtg aagccacaga tgtacggccc cgaaaccggc ccaggacac 1200  |
| [0040] | aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaagtgat 1260 |
| [0041] | atcacaaggt ccagggctg gggtcagaaa ttctctcccg agggaatgaa gccacaggag 1320  |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
| [0042] | ccaagagcag gaggaccaag gccctggcga aggccgtggc ctcgttcaag taaaagatcc | 1380 |
| [0043] | tagtacagtg caggtcccaa tgtgtactag gatcttttac ttgaacgggg acgccggcat | 1440 |
| [0044] | ccgggctcag gacccccctc tctgccagag gcaccaacac cagagttcac aaatcagtct | 1500 |
| [0045] | cctgcccttt gcatgtagca aatttctgtg gctgcgtgaa agccttgagg ggctccggga | 1560 |
| [0046] | gctagagcct ctgctaacca tgttcatgcc ttcttctttt tcctacagct cctgggcaac | 1620 |
| [0047] | gtgctgggta ttgtgctgtc tcatcatttt ggcaaagaat tcctcgaaga tccgaaggga | 1680 |
| [0048] | aagtcttcca cgactgtggg atccgttcga agatatcacc ggttgagcca ccatgagcac | 1740 |
| [0049] | cctgtgcccc cccccagcc ccgccgtggc caagaccgag atcgccctga gcggcaagag  | 1800 |
| [0050] | ccccctgtg gccgccacct tcgcctactg ggacaacatc ctgggcccc gcgtgcgcca   | 1860 |
| [0051] | catctgggcc cccaagaccg agcaggtgct gctgagcgac ggcgagatca ctttctggc  | 1920 |
| [0052] | caaccacacc ctgaacggcg agatcctgcg caacgccgag agcggcgcca tcgacgtgaa | 1980 |
| [0053] | gttcttcgtg ctgagcgaga agggcgtgat catcgtgagc ctgatcttcg acggcaactg | 2040 |
| [0054] | gaacggcgac cgcagcacct acggcctgag catcatcctg cccagaccg agctgagctt  | 2100 |
| [0055] | ctacctgcc ctgcaccgct tgtgcgtgga ccgcctgacc cacatcatcc gcaagggccg  | 2160 |
| [0056] | catctggatg cacaaggagc gccaggagaa cgtgcagaag atcatcctgg agggcaccga | 2220 |
| [0057] | gcgcatggag gaccagggcc agagcatcat ccccatgctg accggcgagg tgatccccgt | 2280 |
| [0058] | gatggagctg ctgagcagca tgaagagcca cagcgtgccc gaggagatcg acatcgccga | 2340 |
| [0059] | caccgtgctg aacgacgacg acatcggcga cagctgccac gagggttcc tgctgaacgc  | 2400 |
| [0060] | catcagcagc cacctgcaga cctgcggctg cagcgtggtg gtgggcagca gcgccgagaa | 2460 |
| [0061] | ggtgaacaag atcgtgcgca ccctgtgcct gttcctgacc cccgccgagc gcaagtgcag | 2520 |
| [0062] | ccgcctgtgc gaggccgaga gcagcttcaa gtacgagagc ggctgttcg tgcaggccct  | 2580 |
| [0063] | gctgaaggac agcaccggca gcttcgtgct gcccttccgc caggtgatgt acgcccccta | 2640 |
| [0064] | ccccaccacc cacatcgacg tggacgtgaa caccgtgaag cagatgcccc cctgccacga | 2700 |
| [0065] | gcacatctac aaccagcgcc gctacatgcg cagcgagctg accgccttct ggcgcgccac | 2760 |
| [0066] | cagcgaggag gacatggccc aggacaccat catctacacc gacgagagct tcacccccga | 2820 |
| [0067] | cctgaacatc ttccaggacg tgctgcaccg cgacaccctg gtgaaggcct tcctggacca | 2880 |
| [0068] | ggtgttccag ctgaagcccc gcctgagcct gcgcagcacc ttcttgcccc agttcctgct | 2940 |
| [0069] | ggtgctgcac cgcaaggccc tgacctgat caagtacatc gaggacgaca ccagaaggg   | 3000 |
| [0070] | caagaagccc ttcaagagcc tgcgcaacct gaagatcgac ctggacctga ccgccgaggg | 3060 |
| [0071] | cgacctgaac atcatcatgg ccctggccga gaagatcaag cccggcctgc acagcttcat | 3120 |
| [0072] | cttcggccgc cccttctaca ccagcgtgca ggagcgcgac gtgctgatga ctttctaatg | 3180 |
| [0073] | acaattgtta attaatgtaa aaccctcgag gccgcaagct tatcgataat caacctctgg | 3240 |
| [0074] | attacaaaat ttgtgaaaga ttgactggta ttcttaacta tgttgctcct ttacgctat  | 3300 |
| [0075] | gtggatacgc tgctttaatg cttttgtatc atgctattgc ttcccgtatg gctttcattt | 3360 |
| [0076] | tctcctcctt gtataaatcc tggttgctgt ctctttatga ggagttgtgg cccgttgtca | 3420 |
| [0077] | ggcaacgtgg cgtggtgtgc actgtgtttg ctgacgcaac cccactgggt tggggcattg | 3480 |
| [0078] | ccaccacctg tcagctcctt tccgggactt tcgctttccc cctccctatt gccacggcgg | 3540 |
| [0079] | aactcatgcg cgctgcctt gcccgctgct ggacaggggc tcggctgttg ggcactgaca  | 3600 |
| [0080] | attccgtggt gttgtcgggg aaatcatcgt ctttctcttg gctgctcgcc tgtgttgcca | 3660 |
| [0081] | cctggattct gcgcgggacg tccttctgct acgtcccttc ggccctcaat ccagcggacc | 3720 |
| [0082] | ttccttcccg cggcctgctg ccggctctgc ggcctcttcc gcgtcttcgc cttcgccctc | 3780 |
| [0083] | agacgagtcg gatctccctt tgggccgcct ccccgcatcg ataccgtcga ctagagctcg | 3840 |

|        |             |            |            |            |             |            |      |
|--------|-------------|------------|------------|------------|-------------|------------|------|
| [0084] | ctgatcagcc  | tcgactgtgc | cttctagttg | ccagccatct | gttgtttgcc  | cctccccctg | 3900 |
| [0085] | gccttccttg  | accctggaag | gtgccactcc | cactgtcctt | tcctaataaa  | atgaggaaat | 3960 |
| [0086] | tgcacgcgat  | tgtctgagta | ggtgtcattc | tattctgggg | ggtgggggtg  | ggcaggacag | 4020 |
| [0087] | caagggggag  | gattgggaag | acaatagcag | gcatgctggg | gagagatcca  | cgataacaaa | 4080 |
| [0088] | cagctttttt  | gggtgaaca  | tattgactga | attccctgca | ggttgggcac  | tccctctctg | 4140 |
| [0089] | cgcgctcgct  | cgctcactga | ggccgcccgg | gcaaagcccg | ggcgtcgggc  | gacctttggt | 4200 |
| [0090] | cgcccgccct  | cagtgagcga | gcgagcgcgc | agagaggag  | tggccaactc  | catcactagg | 4260 |
| [0091] | ggttcctcg   | gccgctcgta | cggctctgag | gaattcctgc | aggataactt  | gccaacctca | 4320 |
| [0092] | ttctaaaatg  | tatatagaag | cccaaaagac | aataacaaaa | atattcttgt  | agaacaaaat | 4380 |
| [0093] | gggaaagaat  | gttccactaa | atatcaagat | ttagagcaa  | gcatgagatg  | tgtggggata | 4440 |
| [0094] | gacagtgagg  | ctgataaaat | agagtagagc | tcagaaacag | accattgat   | atatgtaagt | 4500 |
| [0095] | gacctatgaa  | aaaaatatgg | cattttacaa | tgggaaaatg | atggtctttt  | tcttttttag | 4560 |
| [0096] | aaaaacaggg  | aaatatatgt | atatgtaaaa | aataaaagg  | aaccatag    | tcataccata | 4620 |
| [0097] | cacacaaaaa  | aattccagt  | aattataagt | ctaaatggag | aaggcaaac   | tttaaactct | 4680 |
| [0098] | ttagaaaata  | atatagaagc | atgcagacca | gcctggccaa | catgatgaaa  | ccctctctac | 4740 |
| [0099] | taataataaa  | atcagtagaa | ctactcagga | ctactttgag | tgggaagtcc  | ttttctatga | 4800 |
| [0100] | agacttcttt  | ggccaaaatt | aggctctaaa | tgcaaggaga | tagtgcatca  | tgctggctg  | 4860 |
| [0101] | cacttactga  | taaatgatgt | tatcaccatc | tttaaccaa  | tgacagagaa  | caagttatgg | 4920 |
| [0102] | tactgatgtg  | ctggattgag | aaggagctct | acttccttga | caggacacat  | ttgtatcaac | 4980 |
| [0103] | ttaaaaagc   | agatttttgc | cagcagaact | attcattcag | aggtaggaaa  | cttagaatag | 5040 |
| [0104] | atgatgtcac  | tgattagcat | ggcttcccca | tctccacagc | tgcttccac   | ccaggttgcc | 5100 |
| [0105] | cacagttgag  | tttgtccagt | gctcagggt  | gcccactctc | agtaagaagc  | cccacaccag | 5160 |
| [0106] | ccccctcca   | aatatgttgg | ctgttccttc | cattaaagt  | acccacttt   | agagcagcaa | 5220 |
| [0107] | gtggatttct  | gtttcttaca | gttcaggaag | gaggagtcag | ctgtgagaac  | ctggagcctg | 5280 |
| [0108] | agatgcttct  | aagtccact  | gctactggg  | tcagggaagc | cagactccag  | catcagcagt | 5340 |
| [0109] | caggagcact  | aagcccttgc | caacatcctg | tttctcagag | aaactgcttc  | cattataatg | 5400 |
| [0110] | gttgtccttt  | tttaagctat | caagccaaac | aaccagtgtc | taccattatt  | ctcatcacct | 5460 |
| [0111] | gaagccaagg  | gttctagcaa | aagtcaagct | gtcttgtaat | ggttgatgtg  | cctccagctt | 5520 |
| [0112] | ctgtcttcag  | tactccact  | cttagcctgc | tctgaatcaa | ctctgaccac  | agttccctgg | 5580 |
| [0113] | agccccctgcc | acctgctgcc | cctgccacct | tctccatctg | cagtgtctgtg | cagccttctg | 5640 |
| [0114] | cactcttgca  | gagctaatag | gtggagactt | gaaggaagag | gaggaaagt   | tctcataata | 5700 |
| [0115] | gccttgctgc  | aagctcaa   | gggaggtggg | cactgtgccc | aggagccttg  | gagcaaaggc | 5760 |
| [0116] | tgtgccaac   | ctctgactgc | atccaggttt | ggtcttgaca | gagataagaa  | gccctggctt | 5820 |
| [0117] | ttggagccaa  | aatctaggtc | agacttaggc | aggattctca | aagtttatca  | gcagaacatg | 5880 |
| [0118] | aggcagaaga  | cccttctgc  | tccagcttct | tcaggctcaa | ccttcatcag  | aatagataga | 5940 |
| [0119] | aagagaggct  | gtgagggttc | ttaaaacaga | agcaaactcg | actcagagaa  | taaacaacct | 6000 |
| [0120] | cctagtaaac  | tacagcttag | acagagcatc | tgggtgtgag | tgtgctcagt  | gtcctactca | 6060 |
| [0121] | actgtctggt  | atcagccctc | atgaggactt | ctcttcttct | cctcatagac  | ctccatctct | 6120 |
| [0122] | gttttcctta  | gcctgcagaa | atctggatgg | ctattcacag | aatgcctgtg  | ctttcagagt | 6180 |
| [0123] | tgcatttttt  | ctctggtatt | ctggttcaag | catttgaagg | taggaaaggt  | tctccaagt  | 6240 |
| [0124] | caagaaagcc  | agccctgagc | ctcaactgcc | tggctagtgt | ggtcagtagg  | atgcaaaggc | 6300 |
| [0125] | tgttgaatgc  | cacaaggcca | aactttaacc | tgtgtaccac | aagcctagca  | gcagaggcag | 6360 |

|        |             |            |             |             |            |             |      |
|--------|-------------|------------|-------------|-------------|------------|-------------|------|
| [0126] | ctctgctcac  | tggaactctc | tgtcttcttt  | ctcctgagcc  | ttttcttttc | ctgagttttc  | 6420 |
| [0127] | tagctctcct  | caaccttacc | tctgccctac  | ccaggacaaa  | cccaagagcc | actgtttctg  | 6480 |
| [0128] | tgatgtctc   | tccagcccta | attaggcatc  | atgacttcag  | cctgaccttc | catgctcaga  | 6540 |
| [0129] | agcagtgcta  | atccacttca | gatgagctgc  | tctatgcaac  | acaggcagag | cctacaaacc  | 6600 |
| [0130] | tttgaccag   | agccctccac | atatcagtgt  | ttgttcatac  | tcacttcaac | agcaaagtgt  | 6660 |
| [0131] | actgctgaga  | ttaagatttt | acacaagatg  | gtctgtaatt  | tcacagttag | ttttatccca  | 6720 |
| [0132] | ttaggtatga  | aagaattage | ataattcccc  | ttaaacaatga | atgaatctta | gattttttta  | 6780 |
| [0133] | taaatagttt  | tggaagtaaa | gacagagaca  | tcaggagcac  | aaggaatagc | ctgagaggac  | 6840 |
| [0134] | aaacagaaca  | agaaagagtc | tggaatatca  | caggatgttc  | ttggcctcct | caaagcaagt  | 6900 |
| [0135] | gcaagcagat  | agtaccagca | gccccaggct  | atcagagccc  | agtgaagaga | agtaccatga  | 6960 |
| [0136] | aagccacagc  | tctaaccacc | ctgttccaga  | gtgacagaca  | gtccccaaga | caagccagcc  | 7020 |
| [0137] | tgagccagag  | agagaactgc | aagagaaagt  | ttctaattta  | ggttctgtta | gattcagaca  | 7080 |
| [0138] | agtgcaggtc  | atcctctctc | cacagctact  | cacctctcca  | gcctaacaaa | gcctgcagtc  | 7140 |
| [0139] | cacactccaa  | ccctgggtgc | tcacctccta  | gcctctccca  | acatcctgct | ctctgaccat  | 7200 |
| [0140] | ctttctgcatc | tctcatctca | ccatctccca  | ctgtctacag  | cctactcttg | caactacat   | 7260 |
| [0141] | ctcattttct  | gacatcctgt | ctacatcttc  | tgccatactc  | tgccatctac | cataccacct  | 7320 |
| [0142] | cttaccatct  | accacacat  | cttttatctc  | catccctctc  | agaagcctcc | aagctgaatc  | 7380 |
| [0143] | ctgctttatg  | tgttcatctc | agccccgtca  | tggaagagtg  | accccagagg | cagaactatt  | 7440 |
| [0144] | cccagagagc  | ttggccaaga | aaaacaaaac  | taccagcctg  | gccaggctca | ggagtagtaa  | 7500 |
| [0145] | gctgcagtg   | ctgttggtgt | ctagcttcaa  | cagctgcagg  | agttccactc | tcaaagtctc  | 7560 |
| [0146] | cacatttctc  | acatcctcct | gattctggtc  | actaccatc   | ttcaaagaac | agaatatctc  | 7620 |
| [0147] | acatcagcat  | actgtgaagg | actagtcatg  | ggtgcagctg  | ctcagagctg | caaagtcatt  | 7680 |
| [0148] | ctggatgggtg | gagagcttac | aaacatttca  | tgatgtctcc  | cccgtctga  | tggtctggagc | 7740 |
| [0149] | ccaatcccta  | cacagactcc | tgtctgtatg  | gttttctctt  | cactctgagc | cacagccaga  | 7800 |
| [0150] | gggcaggcat  | tcagtctcct | cttcaggctg  | gggctggggc  | actgagaact | caccaaacac  | 7860 |
| [0151] | cttgctctca  | ctccttctgc | aaaacaagaa  | agagctttgt  | gctgcagtag | ccatgaagaa  | 7920 |
| [0152] | tgaaaggaag  | gctttaacta | aaaaatgtca  | gagattatit  | tcaaccctt  | actgtggatc  | 7980 |
| [0153] | accagcaagg  | aggaacacac | acacagagac  | attttttccc  | ctcaaattat | caaaagaatc  | 8040 |
| [0154] | actgcatttg  | ttaaagagag | caactgaatc  | aggaagcaga  | gttttgaaca | tatcagaagt  | 8100 |
| [0155] | taggaatctg  | catcagagac | aatgcagtc   | atggttggtt  | gctgcatacc | agccctaatc  | 8160 |
| [0156] | attagaagcc  | tcatggactt | caaacatcat  | tccctctgac  | aagatgctct | agcctaactc  | 8220 |
| [0157] | catgagataa  | aataaatctg | cctttcagag  | caaagaaga   | gtccaccagc | ttcttctcag  | 8280 |
| [0158] | tgtgaacaag  | agctccagtc | aggtttagtca | gtccagtgc   | gtagaggaga | ccagtctgca  | 8340 |
| [0159] | tcctctaatt  | ttcaaaggca | agaagatttg  | tttaccctgg  | acaccaggca | caagttaggt  | 8400 |
| [0160] | cacagagctc  | ttagatatgc | agtcctcatg  | agtgaggaga  | ctaaagcgca | tgccatcaag  | 8460 |
| [0161] | acttcagtgt  | agagaaaacc | tccaaaaaag  | cctcctcact  | acttctggaa | tagctcagag  | 8520 |
| [0162] | gccgaggcgg  | cctcggcctc | tgcataaata  | aaaaaaatta  | gtcagccatg | gggcggagaa  | 8580 |
| [0163] | tgggcggaac  | tgggcggagt | tagggcgagg  | atgggcggag  | ttagggcgag | gactatgggt  | 8640 |
| [0164] | gctgactaat  | tgagatgcat | gctttgcata  | cttctgcctg  | ctggggagcc | tggggacttt  | 8700 |
| [0165] | ccacacctgg  | ttgctgacta | attgagatgc  | atgctttgca  | tacttctgcc | tgctggggag  | 8760 |
| [0166] | cctggggact  | ttccacaccc | taactgacac  | acattccaca  | gctgcattaa | tgaatcggcc  | 8820 |
| [0167] | aacgcgcggg  | gagaggcggg | ttgcgtattg  | ggcgtctctc  | cgcttctctg | ctcactgact  | 8880 |

|        |                                                                    |       |
|--------|--------------------------------------------------------------------|-------|
| [0168] | cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcaactcaaag gcggtataac | 8940  |
| [0169] | ggttatccac agaatacagg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa  | 9000  |
| [0170] | aggccaggaa ccgtaaaaag gccgcgttgc tggcggtttt ccataggctc cccccccctg  | 9060  |
| [0171] | acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa  | 9120  |
| [0172] | gataccaggc gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc  | 9180  |
| [0173] | ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac  | 9240  |
| [0174] | gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac  | 9300  |
| [0175] | ccccgttca gcccgaccgc tgcgccttat ccgtaacta tcgtcttgag tccaacccgg    | 9360  |
| [0176] | taagacacga cttatcgcca ctggcagcag ccaactggtaa caggattagc agagcgaggt | 9420  |
| [0177] | atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa  | 9480  |
| [0178] | cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct  | 9540  |
| [0179] | cttgatccgg caaacaacc accgctggta gcggtggtt ttttgtttgc aagcagcaga    | 9600  |
| [0180] | ttacgcgcag aaaaaagga tctcaagaag atcctttgat cttttctacg ggtctgacg    | 9660  |
| [0181] | ctcagtggaa cgaaaactca cgtaagggga ttttggctcat gagattatca aaaaggatct | 9720  |
| [0182] | tcacctagat ctttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt  | 9780  |
| [0183] | aaacttggtc tgacagttac caatgcttaa tcagttaggc acctatctca gcgatctgtc  | 9840  |
| [0184] | tatttcgttc atccatagtt gcctgactcc tgcaaaccac gttgtgtctc aaaatctctg  | 9900  |
| [0185] | atgttacatt gcacaagata aaaatatatc atcatgaaca ataaaactgt ctgcttacat  | 9960  |
| [0186] | aaacagtaat acaaggggtg ttatgagcca tattcaacgg gaaacgtctt gctcgaggcc  | 10020 |
| [0187] | gcgattaaat tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt  | 10080 |
| [0188] | cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag cccgatgcgc cagagttgtt  | 10140 |
| [0189] | tctgaaacat ggcaaaggta gcgttgccaa tgatgttaca gatgagatgg tcagactaaa  | 10200 |
| [0190] | ctggctgacg gaatttatgc ctcttcgcac catcaagcat tttatccgta ctctgatga   | 10260 |
| [0191] | tgcatgggta ctaccactg cgatccccgg gaaaacagca ttccaggat tagaagaata    | 10320 |
| [0192] | tcctgattca ggtgaaaata ttgttgatgc gctggcagtg ttctgcgcc ggttgcatc    | 10380 |
| [0193] | gattcctgtt tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctccaggcca  | 10440 |
| [0194] | atcacgaatg aataacggtt tgggtgatgc gaggatgtt gatgacgagc gtaatggctg   | 10500 |
| [0195] | gcctgttgaa caagtctgga aagaaatgca taagcttttg ccattctcac cggattcagt  | 10560 |
| [0196] | cgtaactcat ggtgatttct cacttgataa cttattttt gacgaggga aattaatagg    | 10620 |
| [0197] | ttgtattgat gttggacgag tcggaatcgc agaccgatac caggatcttg ccacctatg   | 10680 |
| [0198] | gaactgcctc ggtgagtttt ctcttcatt acagaaacgg ctttttcaa aatatggtat    | 10740 |
| [0199] | tgataatcct gatatgaata aattgcagtt tcatattgat ctcatgagt ttttctaagg   | 10800 |
| [0200] | gcggcctgcc accataccca cgccgaaaca agcgtcatg agcccgaagt ggcgagccc    | 10860 |
| [0201] | atcttcccca tcggtgatgt cggcgatata ggcccgagca accgcacctg tggcgccggt  | 10920 |
| [0202] | gatgagggcg cgccaagtcg acgtccggca gtc                               | 10953 |
| [0203] | <210> 2                                                            |       |
| [0204] | <211> 10848                                                        |       |
| [0205] | <212> DNA                                                          |       |
| [0206] | <213> 人工序列                                                         |       |
| [0207] | <220>                                                              |       |
| [0208] | <223> 合成的多核苷酸                                                      |       |
| [0209] | <400> 2                                                            |       |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [0210] | ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc   | 60   |
| [0211] | cgacgcccgg gctttgcccg ggccgctca gtgagcgagc gagcgcgag agagggagt     | 120  |
| [0212] | gccaaactca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc  | 180  |
| [0213] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240  |
| [0214] | tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag  | 300  |
| [0215] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc  | 360  |
| [0216] | cgcgttacat aacttacggt aaatggccc cctggctgac cgcccaacga cccccgcca    | 420  |
| [0217] | ttgacgtcaa taatgacgta tgttcccata gtaacccaa tagggacttt ccattgacgt   | 480  |
| [0218] | caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg | 540  |
| [0219] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgcctggca ttatgccag   | 600  |
| [0220] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt  | 660  |
| [0221] | accatggctg aggtgagccc caggttctgc ttactctcc ccatctcccc cccctcccca   | 720  |
| [0222] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggccgggggg  | 780  |
| [0223] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggccggggcg ggccgaggcg  | 840  |
| [0224] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg  | 900  |
| [0225] | cggcgccggc ggccggcccta taaaaagcga agcgcgcggc ggccgggagt cgctgcgacg | 960  |
| [0226] | ctgccttgcg cccgtgcccc gctccgccgc cgctcgcgc cgcccgcccc ggctctgact   | 1020 |
| [0227] | gaccgcgtta ctcccacagg tgagcggcg ggacggccct tctcctccgg gctgtaatta   | 1080 |
| [0228] | gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcggg    | 1140 |
| [0229] | ccggtttcgg ggcttagtg aagccacaga tgtacggccc cgaaaccggc ccaggacac    | 1200 |
| [0230] | aaggcctgtt actagactc acatggaaca aatggccacc gtgggaggat gacaatttct   | 1260 |
| [0231] | gtggctgcgt gaaagccttg aggggctccg ggagctagag cctctgctaa ccatgttcat  | 1320 |
| [0232] | gccttcttct ttttctaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat    | 1380 |
| [0233] | tttgcaaag aattcctga agatccgaag ggaaagtctt ccacgactgt gggatccgtt    | 1440 |
| [0234] | cgaagatata ccggttgag ccacctgga attcagcagc ccagcagag aggaatgccc     | 1500 |
| [0235] | caagcctctg agccgggtgt caatcatggc cggatctctg acaggactgc tgctgcttca  | 1560 |
| [0236] | ggccgtgtct tgggcttctg gcgctagacc ttgcatcccc aagagcttcg gctacagcag  | 1620 |
| [0237] | cgtcgtgtgc gtgtgcaatg ccacctactg cgacagcttc gaccctccta ctttctctgc  | 1680 |
| [0238] | tctgggcacc ttacgagat acgagagcac cagatccggc agacggatgg aactgagcat   | 1740 |
| [0239] | gggacccatc caggccaatc acacaggcac tggcctgctg ctgacactgc agcctgagca  | 1800 |
| [0240] | gaaattccag aaagtgaag gcttcggcgg agccatgaca gatgccgccc ctctgaatat   | 1860 |
| [0241] | cctggctctg tctccaccag ctcagaacct gctgctcaag agctacttca gcgaggaagg  | 1920 |
| [0242] | catcggttac aacatcatca gagtgcccat ggccagctgc gacttcagca tcaggaccta  | 1980 |
| [0243] | cacctacgcc gacacaccg acgatttcca gctgcacaac ttacgcctgc ctgaagagga   | 2040 |
| [0244] | caccaagctg aagatccctc tgatccacag agccctgcag ctggcacaaa gaccctgttc  | 2100 |
| [0245] | actgctggcc tctccatgga catctccac ctggctgaaa acaaatggcg ccgtgaatgg   | 2160 |
| [0246] | caagggcagc ctgaaaggcc aacctggcga catctaccac cagacctggg ccagatactt  | 2220 |
| [0247] | cgtgaagttc ctggacgct atgccgagca caagctgcag ttttgggccc tgacagccga   | 2280 |
| [0248] | gaacgaacct tctgctggac tgctgagcgg ctaccccttt cagtgcctgg gctttacacc  | 2340 |
| [0249] | cgagcaccag cgggacttta tcgcccgtga tctgggacct aactggcca atagcaccca   | 2400 |
| [0250] | ccataatgtg cggctgctga tgctggacga ccagagactg cttctgcccc actgggctaa  | 2460 |
| [0251] | agtggctgct acagatcctg aggccgcca atacgtgcac ggaatcgccc tgactggta    | 2520 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [0252] | tctggacttt ctggccccctg ccaaggccac actgggagag acacacagac tgttcccaa  | 2580 |
| [0253] | caccatgctg ttgccagcg aagcctgtgt gggcagcaag ttttgggaac agagcgtgcg   | 2640 |
| [0254] | gctcggcagc tgggatatag gcatgcagta cagccacagc atcatcacca acctgctgta  | 2700 |
| [0255] | ccacgtcgtc ggctggaccg actggaatct ggccctgaat cctgaaggcg gccctaactg  | 2760 |
| [0256] | gggccgaaac ttctgtgaca gcccacatcat cgtggacatc accaaggaca ctttctacaa | 2820 |
| [0257] | gcagcccatg ttctaccacc tgggacactt cagcaagttc atccccgagg gctctcagcg  | 2880 |
| [0258] | cgttggactg gtggcttccc agaagaacga tctggacgcc gtggctctga tgcacctga   | 2940 |
| [0259] | tggatctgct gtgggtgtgg tctgaaccg cagcagcaaa gatgtgcccc tgaccatcaa   | 3000 |
| [0260] | ggatcccgcc gtgggattcc tggaacaat cagccctggc tactccatcc acacctacct   | 3060 |
| [0261] | gtggcgtaga cagtgacaat tgttaattaa gtttaaacc tggaggccgc aagcttatcg   | 3120 |
| [0262] | ataatcaacc tctggattac aaaatttgtg aaagattgac tggatttctt aactatgttg  | 3180 |
| [0263] | ctccttttac gctatgtgga tacgtgctt taatgcctt gtatcatgct attgcttccc    | 3240 |
| [0264] | gtatggcttt cattttctcc tcttgtata aatcctggtt gctgtctctt tatgaggagt   | 3300 |
| [0265] | tgtggcccg tgtcaggcaa cgtggcgtgg tgtgactgt gtttgctgac gcaacccca     | 3360 |
| [0266] | ctggttgggg cattgccacc acctgtcagc tctttccgg gactttcgt tccccctcc     | 3420 |
| [0267] | ctattgccac ggcggaactc atcgccgct gccttgccc ctgctggaca ggggctcggc    | 3480 |
| [0268] | tgttgggac tgacaattcc gtggtgtgt cggggaatc atcgtcctt ccttggctgc      | 3540 |
| [0269] | tcgctgtgt tgccacctg attctgcgc ggacgtcct ctgctacgtc cttcggccc       | 3600 |
| [0270] | tcaatccagc ggaccttct tccgcggcc tgcgtccgc tctgcggcct cttccgcgc      | 3660 |
| [0271] | ttcgcttcg ccctcagac agtcggatct cctttgggc cgcctcccc catcgatacc      | 3720 |
| [0272] | gtcgactaga gctcgtgat cagcctcgac tgtgccttct agttgccagc catctgttgt   | 3780 |
| [0273] | ttgccccctc cccgtgcctt ccttgacct ggaaggtgcc actccactg tcttttcta     | 3840 |
| [0274] | ataaaatgag gaaattgcat cgcattgtct gagtagtgt cattctattc tgggggtgtg   | 3900 |
| [0275] | ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggagag  | 3960 |
| [0276] | atccacgata acaaacagct tttttgggt gaacatatg actgaattcc ctgcaggttg    | 4020 |
| [0277] | gccactcct ctctgcgcgc tcgtcgtc actgaggccg cccgggcaaa gcccggcgt      | 4080 |
| [0278] | cgggcgacct ttggtcggc ggctcagtg agcgagcag cgcgagaga gggagtggc       | 4140 |
| [0279] | aactccatca ctaggggttc ctgcggccgc tcgtacggtc tcgaggaatt cctgcaggat  | 4200 |
| [0280] | aacttgcaa cctcattcta aaatgtatat agaagcccaa aagacaataa caaaaatatt   | 4260 |
| [0281] | ctttagaac aaaatgggaa agaattgtcc actaaatc aagatttaga gcaaagcatg     | 4320 |
| [0282] | agatgtgtgg ggatagacag tgaggctgat aaaatagagt agagctcaga aacagacca   | 4380 |
| [0283] | ttgatatatg taagtacct atgaaaaaa tatggcattt tacaatggga aaatgatgtg    | 4440 |
| [0284] | ctttttctt tttagaaaa cagggaata tatttatatg taaaaataa aagggaacct      | 4500 |
| [0285] | atatgtcata ccatacacac aaaaaattc cagtgaatta taagtctaaa tggagaaggc   | 4560 |
| [0286] | aaaacttta atcttttaga aaataatata gaagcatgca gaccagcctg gccaatga     | 4620 |
| [0287] | tgaaccctc tctactaata ataaaatcag tagaactact caggactact ttgagtggga   | 4680 |
| [0288] | agtcctttc tatgaagact tctttggcca aaattaggct ctaaagcaa ggagatagt     | 4740 |
| [0289] | catcatgct ggctgcactt actgataaat gatgttatca ccatcttta ccaaatgcac    | 4800 |
| [0290] | aggaacaagt tatgttactg atgtgctgga ttgagaagga gctctactc cttgacagga   | 4860 |
| [0291] | cacatttgta tcaacttaa aaagcagatt ttgcccagca gaactattca ttcagaggta   | 4920 |
| [0292] | ggaaacttag aatagatgat gtcactgatt agcatggctt ccccatctc acagctgctt   | 4980 |
| [0293] | cccaccagg ttgccacag ttgagttgt ccagtgtca gggctgcca ctctcagtaa       | 5040 |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
| [0294] | gaagccccac accagcccct ctccaaatat gttggctgtt cttccatta aagtgacccc  | 5100 |
| [0295] | actttagagc agcaagtgga tttctgtttc ttacagttca ggaaggagga gtcagctgtg | 5160 |
| [0296] | agaacctgga gcctgagatg cttctaagtc ccactgctac tggggtcagg gaagccagac | 5220 |
| [0297] | tccagcatca gcagtcagga gactaagcc cttgccaaca tcctgtttct cagagaaact  | 5280 |
| [0298] | gcttccatta taatggttgt ctttttttaa gctatcaagc caaacaacca gtgtctacca | 5340 |
| [0299] | ttattctcat cacctgaagc caagggttct agcaaaagtc aagctgtctt gtaatggttg | 5400 |
| [0300] | atgtgcctcc agcttctgtc ttacgtcact ccactcttag cctgctctga atcaactctg | 5460 |
| [0301] | accacagttc cctggagccc ctgccacctg ctgcccctgc caccttctcc atctgcagt  | 5520 |
| [0302] | ctgtgcagcc ttctgcactc ttgcagagct aataggtgga gacttgaagg aagaggagga | 5580 |
| [0303] | aagtttctca taatagcctt gctgcaagct caaatgggag gtgggactg tgcccaggag  | 5640 |
| [0304] | ccttgagca aaggctgtgc ccaacctctg actgcatcca ggtttggctc tgacagagat  | 5700 |
| [0305] | aagaagccct ggcttttga gccaaaatct aggtcagact taggcaggat tctcaaagtt  | 5760 |
| [0306] | tatcagcaga acatgaggca gaagaccctt tctgtccag cttcttcagg ctcaacctc   | 5820 |
| [0307] | atcagaatag atagaaagag aggctgtgag ggttcttaa acagaagcaa atctgactca  | 5880 |
| [0308] | gagaataaac aacctcctag taaactacag cttagacaga gcatctggtg gtgagtgtgc | 5940 |
| [0309] | tcagtgtcct actcaactgt ctggtatcag cctcatgag gacttctctt ctttccctca  | 6000 |
| [0310] | tagacctcca tctctgtttt ccttagcctg cagaaatctg gatggctatt cacagaatgc | 6060 |
| [0311] | ctgtgctttc agagttgcat ttttctctg gtattctggt tcaagcattt gaaggtagga  | 6120 |
| [0312] | aaggttctcc aagtgaaga aagccagccc tgagcctcaa ctgcctggct agtgtgtgca  | 6180 |
| [0313] | gtaggatgca aaggctgttg aatgccacaa ggccaaactt taacctgtgt accacaagcc | 6240 |
| [0314] | tagcagcaga ggcagctctg ctactggaa ctctctgtct tctttctcct gagccttttc  | 6300 |
| [0315] | ttttcctgag ttttctagct ctctcaacc ttacctctgc cctaccagg acaaaccxaa   | 6360 |
| [0316] | gagccactgt ttctgtgatg tcctctccag ccctaattag gcatcatgac ttcagcctga | 6420 |
| [0317] | ccttccatgc tcagaagcag tgctaatcca cttcagatga gctgctctat gcaacacagg | 6480 |
| [0318] | cagagcctac aaacctttgc accagagccc tccacatc agtgtttgtt catactcact   | 6540 |
| [0319] | tcaacagcaa atgtgactgc tgagattaag attttacaca agatggtctg taatttcaca | 6600 |
| [0320] | gtagtttta tccattagg tatgaaagaa ttagcataat tccccttaa catgaatgaa    | 6660 |
| [0321] | tcttagattt ttaataaat agttttgaa gtaaagacag agacatcagg agcacaagga   | 6720 |
| [0322] | atagcctgag aggacaaaca gaacaagaaa gagtctggaa atacacagga tgttcttggc | 6780 |
| [0323] | ctctcaaag caagtgaag cagatagtag cagcagcccc aggctatcag agcccagtga   | 6840 |
| [0324] | agagaagtac catgaaagcc acagctctaa ccacctgtt ccagagtga agacagtccc   | 6900 |
| [0325] | caagacaagc cagcctgagc cagagagaga actgcaagag aaagtttcta atttaggttc | 6960 |
| [0326] | tgtagattc agacaagtgc aggtcctcct ctctccacag ctactcacct ctccagccta  | 7020 |
| [0327] | acaaagcctg cagtcacac tccaacctg gtgtctcacc tctagcctc tccaacatc     | 7080 |
| [0328] | ctgctctctg accatcttct gcatctctca tctcaccatc tcccactgtc tacagcctac | 7140 |
| [0329] | tcttgaact accatctcat tttctgacat cctgtctaca tcttctgcca tactctgcca  | 7200 |
| [0330] | tctaccatac cacctcttac catctaccac accatctttt atctccatcc ctctcagaag | 7260 |
| [0331] | cctccaaget gaatcctgtt ttatgtgtc atctcagccc ctgcatggaa agctgacccc  | 7320 |
| [0332] | agaggcagaa ctattcccag agagcttggc caagaaaaac aaaactacca gcctggccag | 7380 |
| [0333] | gctcaggagt agtaagctgc agtgtctgtt gtgttctagc ttcaacagct gcaggagttc | 7440 |
| [0334] | cactctcaaa tgctccacat ttctcacatc ctctgattc tggctactac ccatcttcaa  | 7500 |
| [0335] | agaacagaat atctcacatc agcatactgt gaaggactag tcatgggtgc agctgctcag | 7560 |

|        |                                                                      |       |
|--------|----------------------------------------------------------------------|-------|
| [0336] | agctgcaaag tcattctgga tgggtggagag cttacaaaca tttcatgatg ctccccccgc   | 7620  |
| [0337] | tctgatggct ggagcccaat ccctacacag actcctgctg tatgtgtttt cctttcactc    | 7680  |
| [0338] | tgagccacag ccagagggea ggcattcagt ctctcttca ggctggggct ggggcactga     | 7740  |
| [0339] | gaactcacc c aacaccttgc tctcactcct tctgcaaaac aagaaagagc tttgtgctgc   | 7800  |
| [0340] | agtagccatg aagaatgaaa ggaaggcttt aactaaaaaa tgtcagagat tattttcaac    | 7860  |
| [0341] | cccttactgt ggatcaccag caaggaggaa acacaacaca gagacatttt ttcccctcaa    | 7920  |
| [0342] | attatcaaaa gaatcactgc atttgttaaa gagagcaact gaatcaggaa gcagagtttt    | 7980  |
| [0343] | gaacatatca gaagttagga atctgcatca gagacaaatg cagtcatggt tgtttgctgc    | 8040  |
| [0344] | ataccagccc taatcattag aagcctcatg gacttcaaac atcattccct ctgacaagat    | 8100  |
| [0345] | gctctagcct aactccatga gataaaataa atctgccttt cagagccaaa gaagagtcca    | 8160  |
| [0346] | ccagcttctt ctgagtgtga acaagagctc cagtcagggt agtcagtcca gtgcagtaga    | 8220  |
| [0347] | ggagaccagt ctgcatcctc taattttcaa aggcaagaag atttgtttac cctggacacc    | 8280  |
| [0348] | aggcacaagt gaggtcacag agctcttaga tatgcagtcc tcatgagtga ggagactaaa    | 8340  |
| [0349] | gcgcatgcca tcaagacttc agtgtagaga aaacctcaa aaaagcctcc tctactcttc     | 8400  |
| [0350] | tggaaatagct cagaggccga ggcgccctcg gcctctgcat aaataaaaaa aattagtcag   | 8460  |
| [0351] | ccatggggcg gagaatgggc ggaactgggc ggagttaggg gcgggatggg cggagttagg    | 8520  |
| [0352] | ggcgggacta tggttgctga ctaattgaga tgcattgctt gcatacttct gcctgctggg    | 8580  |
| [0353] | gagcctgggg actttccaca cctggttgc t gactaattga gatgcatgct ttgcatactt   | 8640  |
| [0354] | ctgcctgctg gggagcctgg ggactttcca caccctaact gacacacatt ccacagctgc    | 8700  |
| [0355] | attaatgaat cggccaacgc gcggggagag gcggtttgcg tattggggcg tcttccgctt    | 8760  |
| [0356] | cctcgctcac tgactcgctg cgctcggtcg ttcggtcgcg gcgagcggtc tcagctcact    | 8820  |
| [0357] | caaaggcggg aatacgggta tccacagaat caggggataa cgcaggaaag aacatgtgag    | 8880  |
| [0358] | caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata    | 8940  |
| [0359] | ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc    | 9000  |
| [0360] | cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgtctcctg     | 9060  |
| [0361] | ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc    | 9120  |
| [0362] | tttctcatag ctacgctgt aggtatctca gttcggtgta ggctgcttcgc tccaagctgg    | 9180  |
| [0363] | gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc     | 9240  |
| [0364] | ttgagtccaa cccggtgta cagacttat cgccactggc agcagccact ggtaacagga      | 9300  |
| [0365] | ttagcagagc gaggtatgta ggcggtgcta cagagtctt gaagtgtgg cctaactacg      | 9360  |
| [0366] | gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa    | 9420  |
| [0367] | aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggtttttttg    | 9480  |
| [0368] | tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt    | 9540  |
| [0369] | ctacggggtc tgacgctcag tggaacgaaa actcacgta agggattttg gtcatgagat     | 9600  |
| [0370] | tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct    | 9660  |
| [0371] | aaagtatatata tgagttaaact tgggtctgaca gttaccaatg cttaatcagt gaggcacct | 9720  |
| [0372] | tctcagcgat ctgtctatct cgttcattcca tagttgcctg actcctgcaa accacgttgt   | 9780  |
| [0373] | gtctcaaaaat ctctgatgtt acattgcaca agataaaaa atatcatcat gaacaataaa    | 9840  |
| [0374] | actgtctgct tacataaaca gtaatacaag ggggtgttatg agccatatc aacgggaaac    | 9900  |
| [0375] | gtcttgctcg aggccgcgat taaattccaa catggatgct gatttatatg ggtataaatg    | 9960  |
| [0376] | ggctcgcgat aatgtcgggc aatcagggtgc gacaatctat cgattgtatg ggaagccga    | 10020 |
| [0377] | tgcgccagag ttgtttctga aacatggcaa aggtagcgtt gccaatgatg ttacagatga    | 10080 |

|        |                                                                    |       |
|--------|--------------------------------------------------------------------|-------|
| [0378] | gatggtcaga ctaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat  | 10140 |
| [0379] | ccgtactcct gatgatgcat gggttactcac cactgcgac cccgggaaaa cagcattcca  | 10200 |
| [0380] | ggtattagaa gaatatcctg attcaggatga aaatattgtt gatgcgctgg cagtgttcct | 10260 |
| [0381] | gcgccggttg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg  | 10320 |
| [0382] | tctcgctcag gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg attttgatga  | 10380 |
| [0383] | cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa atgcataagc ttttgccatt  | 10440 |
| [0384] | ctcaccgat tcagtcgtca ctcatggatga tttctcactt gataacctta tttttgacga  | 10500 |
| [0385] | ggggaaatta ataggttgta ttgatgttg acgagtcgga atcgagacc gataccagga    | 10560 |
| [0386] | tcttgccatc ctatggaact gcctcggtga gttttctcct tcattacaga aacggctttt  | 10620 |
| [0387] | tcaaaaatat ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga  | 10680 |
| [0388] | tgagtttttc taaggcggc ctgccaccat acccagccg aaacaagcg tcatgagccc     | 10740 |
| [0389] | gaagtggcga gcccgatctt ccccatcggt gatgtcggcg atataggcg cagcaaccg    | 10800 |
| [0390] | acctgtggcg ccggtgatga gggcgcgcca agtcgacgtc cggcagtc               | 10848 |
| [0391] | <210> 3                                                            |       |
| [0392] | <211> 1446                                                         |       |
| [0393] | <212> DNA                                                          |       |
| [0394] | <213> 人工序列                                                         |       |
| [0395] | <220>                                                              |       |
| [0396] | <223> 合成的多核苷酸                                                      |       |
| [0397] | <400> 3                                                            |       |
| [0398] | atgagcacc tgtgcccccc cccagcccc gccgtggcca agaccgagat cgccctgagc    | 60    |
| [0399] | ggcaagagcc ccctgctggc cgccacctt gcctactggg acaacatcct gggccccgc    | 120   |
| [0400] | gtgcgccaca tctgggcccc caagaccgag caggtgctgc tgagcgacgg cgagatcacc  | 180   |
| [0401] | ttcttgcca accacacct gaacggcgag atcctgcgca acgccgagag cggcgccatc    | 240   |
| [0402] | gacgtgaagt tcttcgtgt gagcgagaag ggcgtgatca tcgtgagcct gatcttcgac   | 300   |
| [0403] | ggcaactgga acggcgaccg cagcacctac ggcctgagca tcctcctgcc ccagaccgag  | 360   |
| [0404] | ctgagcttct acctgcccc gcaccgctg tgcgtggacc gcctgacca catcatccgc     | 420   |
| [0405] | aagggccgca tctggatgca caaggagcg caggagaacg tgcagaagat catcctggag   | 480   |
| [0406] | ggcaccgagc gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgaggtg  | 540   |
| [0407] | atccccgtga tggagctgt gagcagcatg aagagccaca gcgtgcccga ggagatcgac   | 600   |
| [0408] | atcgccgaca ccgtgtgaa cgacgacgac atcgccgaca gctgccacga gggcttcctg   | 660   |
| [0409] | ctgaacgcca tcagcagcca cctgcagacc tgcggctgca gcgtggtggt gggcagcagc  | 720   |
| [0410] | gccgagaagg tgaacaagat cgtgcgcacc ctgtgcctgt tctgacccc cgccgagcgc   | 780   |
| [0411] | aagtgcagcc gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttcgtg  | 840   |
| [0412] | cagggcctgc tgaaggacag caccggcagc ttcgtgctgc cttccgcca ggtgatgtac   | 900   |
| [0413] | gccccctacc ccaccacca catcgacgtg gacgtgaaca ccgtgaagca gatgcccc     | 960   |
| [0414] | tgccacgagc acatctacaa ccagcgccgc tacatgcgca gcgagctgac cgccttctgg  | 1020  |
| [0415] | cgcgccacca gcgaggagga catggcccag gacaccatca tctacaccga cgagagcttc  | 1080  |
| [0416] | acccccgacc tgaacatctt ccaggacgtg ctgcaccgac acaccctggt gaaggccttc  | 1140  |
| [0417] | ctggaccagg tgttcagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag   | 1200  |
| [0418] | ttctgctggt tgctgcaccg caaggccctg accctgatca agtacatcga ggacgacacc  | 1260  |
| [0419] | cagaagggca agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc  | 1320  |

|        |                                                                         |
|--------|-------------------------------------------------------------------------|
| [0420] | gccgagggcg acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac 1380  |
| [0421] | agcttcattct tcggccgccc cttctacacc agcgtgcagg agcgcgacgt gctgatgacc 1440 |
| [0422] | ttctaa 1446                                                             |
| [0423] | <210> 4                                                                 |
| [0424] | <211> 481                                                               |
| [0425] | <212> PRT                                                               |
| [0426] | <213> 人工序列                                                              |
| [0427] | <220>                                                                   |
| [0428] | <223> 合成的多肽                                                             |
| [0429] | <400> 4                                                                 |
| [0430] | Met Ser Thr Leu Cys Pro Pro Pro Ser Pro Ala Val Ala Lys Thr Glu         |
| [0431] | 1 5 10 15                                                               |
| [0432] | Ile Ala Leu Ser Gly Lys Ser Pro Leu Leu Ala Ala Thr Phe Ala Tyr         |
| [0433] | 20 25 30                                                                |
| [0434] | Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro Lys         |
| [0435] | 35 40 45                                                                |
| [0436] | Thr Glu Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn         |
| [0437] | 50 55 60                                                                |
| [0438] | His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile         |
| [0439] | 65 70 75 80                                                             |
| [0440] | Asp Val Lys Phe Phe Val Leu Ser Glu Lys Gly Val Ile Ile Val Ser         |
| [0441] | 85 90 95                                                                |
| [0442] | Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu         |
| [0443] | 100 105 110                                                             |
| [0444] | Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His         |
| [0445] | 115 120 125                                                             |
| [0446] | Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile         |
| [0447] | 130 135 140                                                             |
| [0448] | Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Ile Leu Glu         |
| [0449] | 145 150 155 160                                                         |
| [0450] | Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu         |
| [0451] | 165 170 175                                                             |
| [0452] | Thr Gly Glu Val Ile Pro Val Met Glu Leu Leu Ser Ser Met Lys Ser         |
| [0453] | 180 185 190                                                             |
| [0454] | His Ser Val Pro Glu Glu Ile Asp Ile Ala Asp Thr Val Leu Asn Asp         |
| [0455] | 195 200 205                                                             |
| [0456] | Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile         |
| [0457] | 210 215 220                                                             |
| [0458] | Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser         |
| [0459] | 225 230 235 240                                                         |
| [0460] | Ala Glu Lys Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr         |
| [0461] | 245 250 255                                                             |

|        |                                                                   |     |     |
|--------|-------------------------------------------------------------------|-----|-----|
| [0462] | Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe   |     |     |
| [0463] | 260                                                               | 265 | 270 |
| [0464] | Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ser Thr   |     |     |
| [0465] | 275                                                               | 280 | 285 |
| [0466] | Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro   |     |     |
| [0467] | 290                                                               | 295 | 300 |
| [0468] | Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro   |     |     |
| [0469] | 305                                                               | 310 | 315 |
| [0470] | Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu   |     |     |
| [0471] | 325                                                               | 330 | 335 |
| [0472] | Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr   |     |     |
| [0473] | 340                                                               | 345 | 350 |
| [0474] | Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln   |     |     |
| [0475] | 355                                                               | 360 | 365 |
| [0476] | Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val   |     |     |
| [0477] | 370                                                               | 375 | 380 |
| [0478] | Phe Gln Leu Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu Ala Gln   |     |     |
| [0479] | 385                                                               | 390 | 395 |
| [0480] | Phe Leu Leu Val Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile   |     |     |
| [0481] | 405                                                               | 410 | 415 |
| [0482] | Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu Arg Asn   |     |     |
| [0483] | 420                                                               | 425 | 430 |
| [0484] | Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile   |     |     |
| [0485] | 435                                                               | 440 | 445 |
| [0486] | Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe   |     |     |
| [0487] | 450                                                               | 455 | 460 |
| [0488] | Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr   |     |     |
| [0489] | 465                                                               | 470 | 475 |
| [0490] | Phe                                                               |     | 480 |
| [0491] | <210> 5                                                           |     |     |
| [0492] | <211> 10900                                                       |     |     |
| [0493] | <212> DNA                                                         |     |     |
| [0494] | <213> 人工序列                                                        |     |     |
| [0495] | <220>                                                             |     |     |
| [0496] | <223> 合成的多核苷酸                                                     |     |     |
| [0497] | <400> 5                                                           |     |     |
| [0498] | ttggccactc cctctctgcg cgctcgcctc ctactgagg ccgggcgacc aaaggtcgcc  | 60  |     |
| [0499] | cgacgcccg gctttgcccg ggcgccctca gtgagcgagc gacgcgcag agagggagt    | 120 |     |
| [0500] | gccaaactca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180 |     |
| [0501] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc | 240 |     |
| [0502] | tccctaggtt ctagaaccgg tgacgtcttg tcatcctccc acggtggcca tttgttccat | 300 |     |
| [0503] | gtgagtgcga gtaacaggcc ttgtgtcctg gggccggttt cggggccgta catctgtggc | 360 |     |

|        |            |             |            |            |            |             |      |
|--------|------------|-------------|------------|------------|------------|-------------|------|
| [0504] | ttcactaagg | ccccgaaacc  | ggccccgaca | acagcataca | gccttcagca | agcctccagt  | 420  |
| [0505] | ggtctcatac | agaacttata  | agattcccaa | atccaaagac | atttcacgtt | tatggtgatt  | 480  |
| [0506] | tcccagaaca | catagcgaca  | tgcaaataat | gcaggcgcc  | actccctgt  | ccctcacagc  | 540  |
| [0507] | catcttctg  | ccaggcgca   | cgcgcgtg   | gtgttcccgc | ctagtacac  | tgggcccgcg  | 600  |
| [0508] | attccttga  | gcgggtgat   | gacgtcagcg | tttcccatgg | tgaagcttgg | atctgaattc  | 660  |
| [0509] | ggtaccctag | ttattaatag  | taatcaatta | cggggtcatt | agttcatagc | ccatatatgg  | 720  |
| [0510] | agttccgcgt | tacataactt  | acggtaaata | gcccgcctgg | ctgaccgccc | aacgaccccc  | 780  |
| [0511] | gccattgac  | gtcaataatg  | acgtatgttc | ccatagtaac | gccaataggg | acttttcatt  | 840  |
| [0512] | gacgtcaatg | ggtggactat  | ttacggtaaa | ctgcccactt | ggcagtacat | caagtgtatc  | 900  |
| [0513] | atatgccaag | tacgccccct  | attgacgtca | atgacggtaa | atggccccgc | tggcattatg  | 960  |
| [0514] | cccagtacat | gaccttatgg  | gactttccta | cttggcagta | catctacgta | ttagtcatcg  | 1020 |
| [0515] | ctattaccat | ggtcgagggtg | agccccacgt | tctgttcac  | tctccccatc | tccccccct   | 1080 |
| [0516] | ccccaccccc | aattttgtat  | ttatttat   | tttaattatt | ttgtgcagcg | atgggggcg   | 1140 |
| [0517] | gggggggggg | ggggcgcg    | ccaggcggg  | cggggcggg  | cgaggggcg  | ggcggggcga  | 1200 |
| [0518] | ggcgagagg  | tgccggcgca  | gccaatcaga | gcggcgcgct | ccgaaagttt | ccttttatgg  | 1260 |
| [0519] | cgaggcgcg  | gcggcgcg    | ccctataaaa | agcgaagcg  | gcggcgggcg | ggagtcgctg  | 1320 |
| [0520] | cgacgtgcc  | ttcgccccgt  | gccccgtcc  | gccgccgct  | cgcgccgccc | gccccggctc  | 1380 |
| [0521] | tgactgaccg | cgttactccc  | acaggtagc  | ggcggggacg | gcccttctcc | tccgggctgt  | 1440 |
| [0522] | aattagcgct | tggtttaatg  | acggcttgtt | ttctgtggct | gcgtgaaagc | cttgaggggc  | 1500 |
| [0523] | tccgggagct | agagcctctg  | ctaaccatgt | tcatgccttc | ttctttttcc | tacagctcct  | 1560 |
| [0524] | gggcaacgtg | ctggttat    | tgctgtctca | tcattttggc | aaagaattcc | tgaagatcc   | 1620 |
| [0525] | gaagggaag  | tcttcacga   | ctgtgggatc | cgttcgaaga | tatcacgggt | tgagccacca  | 1680 |
| [0526] | tgagaccct  | gtgccccccc  | cccagccccg | ccgtggccaa | gaccgagatc | gccctgagcg  | 1740 |
| [0527] | gcaagagccc | cctgttgcc   | gccacctcg  | cctactggga | caacatcctg | ggcccccg    | 1800 |
| [0528] | tgccacacat | ctgggcccc   | aagaccgagc | aggtgctgct | gagcgacggc | gagatcacct  | 1860 |
| [0529] | tcctggccaa | ccacacctg   | aacggcgaga | tcctgcgcaa | cgccgagagc | ggcgccatcg  | 1920 |
| [0530] | acgtgaagtt | cttcgtgctg  | agcgagaagg | gcgtgatcat | cgtgagcctg | atcttcgacg  | 1980 |
| [0531] | gcaactggaa | cggcgaccgc  | agcacctacg | gcctgagcat | catcctgccc | cagaccgagc  | 2040 |
| [0532] | tgagcttcta | cctgccccctg | caccgcgtgt | gcgtggaccg | cctgaccac  | atcatccga   | 2100 |
| [0533] | aggccgcac  | ctggatgcac  | aaggagcgcc | aggagaacgt | gcagaagatc | atcctggagg  | 2160 |
| [0534] | gcaccgagcg | catggaggac  | caggccaga  | gcacatccc  | catgctgacc | ggcgaggatga | 2220 |
| [0535] | tcccgtgat  | ggagctgctg  | agcagcatga | agagccacag | cgtgcccag  | gagatcgaca  | 2280 |
| [0536] | tcgccgacac | cgtgctgaac  | gacgacgaca | tcggcgacag | ctgccacgag | ggcttctctgc | 2340 |
| [0537] | tgaagccat  | cagcagccac  | ctgcagacct | gcggctgcag | cgtggtggtg | ggcagcagcg  | 2400 |
| [0538] | ccgagaaggt | gaacaagatc  | gtgcgcaccc | tgtgcctgtt | cctgaccccc | gccgagcgca  | 2460 |
| [0539] | agtcgacccg | cctgtgcgag  | gccgagagca | gcttcaagta | cgagagcggc | ctgttcgtgc  | 2520 |
| [0540] | aggccctgct | gaaggacagc  | accggcagct | tcgtgctgcc | cttccgccag | gtgatgtacg  | 2580 |
| [0541] | ccccctaccc | caccacccac  | atcgacgtgg | acgtgaacac | cgtgaagcag | atgccccct   | 2640 |
| [0542] | gccagagca  | catctacaac  | cagcgccgct | acatgcgcag | cgagctgacc | gccttctggc  | 2700 |
| [0543] | gcgccaccag | cgaggaggac  | atggcccagg | acaccatcat | ctacaccgac | gagagcttca  | 2760 |
| [0544] | ccccgacct  | gaacatcttc  | caggacgtgc | tgcaccgca  | caccctggtg | aaggccttc   | 2820 |
| [0545] | tggaccaggt | gttccagctg  | aagccccggc | tgagcctgcg | cagcaccttc | ctggcccagt  | 2880 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [0546] | tcctgctggt gctgcaccgc aagccctga ccctgatcaa gtacatcgag gacgacaccc   | 2940 |
| [0547] | agaagggcaa gaagcccttc aagagcctgc gcaacctgaa gatcgacctg gacctgaccg  | 3000 |
| [0548] | ccgagggega cctgaacatc atcatggccc tggccgagaa gatcaagccc ggccctgcaca | 3060 |
| [0549] | gcttcattctt cggccgcccc ttctacacca gcgtgcagga gcgcgacgtg ctgatgacct | 3120 |
| [0550] | tctaatagaca attgttaatt aagtttaaac cctcgaggcc gcaagcttat cgataatcaa | 3180 |
| [0551] | cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt tgctcctttt  | 3240 |
| [0552] | acgctatgtg gatacgtgc tttaatgcct ttgtatcatg ctattgcttc ccgtatggct   | 3300 |
| [0553] | ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga gttgtggccc  | 3360 |
| [0554] | gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc cactggttgg  | 3420 |
| [0555] | ggcattgcca ccacctgtca gctcctttcc gggactttcg ctttccccct ccctattgcc  | 3480 |
| [0556] | acggcggaac tcatgccgc ctgccttgcc cgctgctgga caggggctcg gctgttgggc   | 3540 |
| [0557] | actgacaatt ccgtggtgtt gtcggggaaa tcatcgtcct ttccttggct gctcgctgt   | 3600 |
| [0558] | gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca  | 3660 |
| [0559] | gcggaccttc cttcccgcgg cctgctgccg gctctgcggc ctcttcgcg tcttcgcctt   | 3720 |
| [0560] | cgccctcaga cgatcggat ctccctttgg gccgcctccc cgcacgata ccgtcgacta    | 3780 |
| [0561] | gagctcgtg atcagcctcg actgtgcctt ctagttgcca gccatctgtt gtttgccct    | 3840 |
| [0562] | ccccgtgcc ttccttgacc ctggaagggt ccaactccac tgcctttcc taataaaatg    | 3900 |
| [0563] | aggaaattgc atcgattgt ctgagtaggt gtcattctat tctgggggt ggggtgggc     | 3960 |
| [0564] | aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggag agatccacga  | 4020 |
| [0565] | taacaaacag cttttttggg gtgaacatat tgactgaatt ccctgcagg tggccactcc   | 4080 |
| [0566] | ctctctgcgc gctcgtcgc tcaactgagc cgccgggca aagccgggc gtcgggcgac     | 4140 |
| [0567] | ctttggtcgc ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaactccat  | 4200 |
| [0568] | cactaggggt tcctgcggc gctcgtacgg tctcgaggaa ttcctgcagg ataacttgc    | 4260 |
| [0569] | aacctcattc taaaatgtat atagaagccc aaaagacaat aacaaaaata ttctttaga   | 4320 |
| [0570] | acaaaatggg aaagaatgtt ccactaaata tcaagattta gagcaaagca tgagatgtgt  | 4380 |
| [0571] | gggatagac agtgaggctg ataaaataga gtagagctca gaaacagacc cattgatata   | 4440 |
| [0572] | tgtaagtac ctatgaaaa aatatggcat ttacaaatgg gaaaatgatg gtctttttct    | 4500 |
| [0573] | tttttagaaa aacagggaaa tatatttata tgtaaaaaat aaaagggaac ccatatgtca  | 4560 |
| [0574] | taccatacac acaaaaaaat tccagtgaat tataagtcta aatggagaag gcaaaacttt  | 4620 |
| [0575] | aaatctttta gaaaataata tagaagcatg cagaccagcc tggccaacat gatgaaaccc  | 4680 |
| [0576] | tcttactaa taataaaatc agtagaacta ctcaggacta ctttgagtgg gaagtccttt   | 4740 |
| [0577] | tctatgaaga cttctttggc caaaattagg ctctaaatgc aaggagatag tgcacatgc   | 4800 |
| [0578] | ctggtgcac ttactgataa atgatgttat caccatcttt aaccaaatgc acaggaacaa   | 4860 |
| [0579] | gttatggtac tgatgtgtg gattgagaag gagctctact tcttgacag gacacatttg    | 4920 |
| [0580] | tatcaactta aaaaagcaga tttttgccag cagaactatt cattcagagg taggaaactt  | 4980 |
| [0581] | agaatagatg atgtcactga ttagcatggc ttccccatct ccacagctgc ttcccacca   | 5040 |
| [0582] | ggttggccac agttgagttt gtccagtgtc cagggtgcc cactctcagt aagaagcccc   | 5100 |
| [0583] | acaccagccc ctctccaaat atgttggtg ttccttccat taaagtgacc ccactttaga   | 5160 |
| [0584] | gcagcaagtg gatttctgtt tcttacagtt caggaaggag gagtcagctg tgagaacctg  | 5220 |
| [0585] | gagcctgaga tgcttctaag tccactgct actgggtca gggaagccag actccagcat    | 5280 |
| [0586] | cagcagtcag gagcactaag cccttgccaa catcctgttt ctcagagaaa ctgcttccat  | 5340 |
| [0587] | tataatggtt gtctttttt aagctatcaa gccaaacaac cagtgtctac cattattctc   | 5400 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [0588] | atcacctgaa gccaaagggtt ctagcaaaag tcaagctgtc ttgtaatggt tgatgtgcct | 5460 |
| [0589] | ccagcttctg tcttcagtca ctccactctt agcctgctct gaatcaactc tgaccacagt  | 5520 |
| [0590] | tccttgagc cctgccacc tgctgccct gccaccttct ccatctgcag tgctgtgcag     | 5580 |
| [0591] | ccttctgcac tcttgacag ctaataggtg gagacttgaa ggaagaggag gaaagtttct   | 5640 |
| [0592] | cataatagcc ttgctgcaag ctcaaatggg aggtgggcac tgtgccagg agccttgag    | 5700 |
| [0593] | caaaggctgt gccaacctc tgactgcac caggtttggt cttgacagag ataagaagcc    | 5760 |
| [0594] | ctggcttttg gagccaaaat ctaggtcaga cttaggcagg attctcaaag tttatcagca  | 5820 |
| [0595] | gaacatgagg cagaagaccc tttctgctcc agcttcttca ggctcaacct tcatcagaat  | 5880 |
| [0596] | agatagaaag agaggctgtg agggttctta aaacagaagc aaatctgact cagagaataa  | 5940 |
| [0597] | acaacctcct agtaactac agcttagaca gagcatctgg tggtagtgt gctcagtgtc    | 6000 |
| [0598] | ctactcaact gtctggatc agccctcatg aggacttctc tttttccct catagacctc    | 6060 |
| [0599] | catctctgtt ttcttagcc tgcagaaatc tggatggcta ttcacagaat gcctgtgctt   | 6120 |
| [0600] | tcagagttgc atttttctc tggattctg gttcaagcat ttgaaggtag gaaagtttct    | 6180 |
| [0601] | ccaagtcaa gaaagccagc cctgagcctc aactgcctgg ctagtgtggt cagtaggatg   | 6240 |
| [0602] | caaaggctgt tgaatgccac aaggccaaac tttaacctgt gtaccacaag cctagcagca  | 6300 |
| [0603] | gaggcagctc tgctcactgg aactctctgt cttctttctc ctgagcctt tcttttctg    | 6360 |
| [0604] | agttttctag ctctcctcaa cttacctct gccctacca ggacaaacc aagagccact     | 6420 |
| [0605] | gtttctgtga tgcctctcc agccctaatt aggcacatc acttcagcct gaccttccat    | 6480 |
| [0606] | gctcagaagc agtgctaate cacttcagat gagctgctct atgcaacaca ggcagagcct  | 6540 |
| [0607] | acaaccttt gcaccagagc cctccacata tcagtgtttg ttcatactca cttcaacagc   | 6600 |
| [0608] | aaatgtgact gctgagatta agattttaca caagatggtc tgtaatttca cagttagttt  | 6660 |
| [0609] | tatccatta ggtatgaaag aattagcata attcccctta aacatgaatg aatcttagat   | 6720 |
| [0610] | ttttaataa atagttttg aagtaaagac agagacatca ggagcacaag gaatagcctg    | 6780 |
| [0611] | agaggacaaa cagaacaaga aagagtctgg aaatacacag gatgttcttg gcctcctcaa  | 6840 |
| [0612] | agcaagtgca agcagatagt accagcagcc ccaggctatc agagcccagt gaagagaagt  | 6900 |
| [0613] | accatgaaag ccacagctct aaccacctg ttccagagt acagacagtc cccaagacaa    | 6960 |
| [0614] | gccagcctga gccagagaga gaactgcaag agaaagtct taatttaggt tctgttagat   | 7020 |
| [0615] | tcagacaagt gcaggctac ctctctccac agctactcac ctctccagcc taacaaagcc   | 7080 |
| [0616] | tgcagtccac actccaacc tgggtgtctc cctctagcc tctcccaaca tctgtctctc    | 7140 |
| [0617] | tgaccatctt ctgcactct catctcacca tctccactg tctacagcct actcttgcaa    | 7200 |
| [0618] | ctaccatctc atttctgac atctgtctc catcttctgc catactctgc catctacat     | 7260 |
| [0619] | accacctctt accatctacc acaccatctt ttatctccat ccctctcaga agcctccaag  | 7320 |
| [0620] | ctgaatcctg ctttatgtgt tcatctcagc ccctgcatgg aaagctgacc ccagaggcag  | 7380 |
| [0621] | aactattccc agagagcttg gccaaagaaa acaaaactac cagcctggcc aggtctagga  | 7440 |
| [0622] | gtagtaagct gcagtgtctg ttgtgttcta gttcaacag ctgcaggagt tccactctca   | 7500 |
| [0623] | aatgctccac atttctcaca tctctctgat tctggtcact acccatctc aaagaacaga   | 7560 |
| [0624] | atatctcaca tcagcactgt gtgaaggact agtcatgggt gcagctgtc agagctgcaa   | 7620 |
| [0625] | agtcattctg gatggtggag agcttacaaa catttcatga tgcctcccc gctctgatgg   | 7680 |
| [0626] | ctggagccca atccctacac agactcctgc tgtatgtgt ttcctttcac tctgagccac   | 7740 |
| [0627] | agccagaggg caggcattca gtctcctct caggctgggg ctggggcact gagaactcac   | 7800 |
| [0628] | ccaacacctt gctctcactc cttctgcaa acaagaaaga gctttgtgct gcagtagcca   | 7860 |
| [0629] | tgaagaatga aaggaaggct ttaactaaaa aatgtcagag attattttca acccttact   | 7920 |

|        |                                                                    |       |
|--------|--------------------------------------------------------------------|-------|
| [0630] | gtggatcacc agcaaggagg aaacacaaca cagagacatt ttttccctc aaattatcaa   | 7980  |
| [0631] | aagaatcact gcatttgta aagagagcaa ctgaatcagg aagcagagtt ttgaacatat   | 8040  |
| [0632] | cagaagttag gaatctgcat cagagacaaa tgcagtcag gttgtttgct gcataccagc   | 8100  |
| [0633] | cctaatacatt agaagcctca tggacttcaa acatcattcc ctctgacaag atgctctagc | 8160  |
| [0634] | ctaactccat gagataaaat aaatctgcct ttcagagcca aagaagagtc caccagcttc  | 8220  |
| [0635] | ttctcagtg gaacaagagc tccagtcagg ttagtcagtc cagtcagta gaggagacca    | 8280  |
| [0636] | gtctgcatcc tctaattttc aaaggcaaga agatttggtt accctggaca ccaggcacia  | 8340  |
| [0637] | gtgaggtcac agagctctta gatatgcagt cctcatgagt gaggagacta aagcgcatgc  | 8400  |
| [0638] | catcaagact tcagtgtaga gaaaacctcc aaaaaagcct cctcactact tctggaatag  | 8460  |
| [0639] | ctcagaggcc gaggcggcct cggcctctgc ataaataaaa aaaattagtc agccatgggg  | 8520  |
| [0640] | cggagaatgg gcggaactgg gcggagttag gggcgggatg gcggagttg gggcggggac   | 8580  |
| [0641] | tatggttgc gactaattga gatgcagct ttgcatact ctgcctgctg gggagcctgg     | 8640  |
| [0642] | ggactttcca cacctggttg ctgactaatt gagatgcag ctttgcatc ttctgcctgc    | 8700  |
| [0643] | tggggagcct ggggactttc cacaccctaa ctgacacaca ttccacagct gcattaatga  | 8760  |
| [0644] | atcgccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc   | 8820  |
| [0645] | actgactgc tgcgtcggc cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg    | 8880  |
| [0646] | gtaatacggc tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc  | 8940  |
| [0647] | cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc  | 9000  |
| [0648] | ccccctgac agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa ccgacagga    | 9060  |
| [0649] | ctataaagat accaggcgtt tccccctgga agctccctcg tgcgtctcc tgttccgacc   | 9120  |
| [0650] | ctgccgtta ccgataacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat   | 9180  |
| [0651] | agctcacgct gtaggtatct cagttcggtg taggtcgtc gctccaagct gggctgtgtg   | 9240  |
| [0652] | cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtc   | 9300  |
| [0653] | aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga   | 9360  |
| [0654] | gcgaggtatg taggcggtgc tacagagttc ttgaagtgtt ggcctaacta cggctacact  | 9420  |
| [0655] | agaagaacag tatttggtat ctgcgtctg ctgaagccag ttaccttcgg aaaaagagtt   | 9480  |
| [0656] | gtagctctt gatccggcaa acaaacacc gctggtagcg gtggttttt tgtttgcaag     | 9540  |
| [0657] | cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg  | 9600  |
| [0658] | tctgacgctc agtggaaacg aaactcacgt taagggtatt tggctatgag attatcaaaa  | 9660  |
| [0659] | aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata  | 9720  |
| [0660] | tatgagtaaa cttggtctga cagttacca tgcctaatca gtgaggcacc tatctcagcg   | 9780  |
| [0661] | atctgtctat ttcgttcac catagttgcc tgactcctgc aaaccacgtt gtgtctcaaa   | 9840  |
| [0662] | atctctgatg ttacattgca caagataaaa atatatcatc atgaacaata aaactgtctg  | 9900  |
| [0663] | cttacataaa cagtaataca aggggtgta tgagccatat tcaacgggaa acgtcttgct   | 9960  |
| [0664] | cgaggccgcg attaaattcc aacatggatg ctgatttata tgggtataaa tgggctcgcg  | 10020 |
| [0665] | ataatgtcgg gcaatcaggt gcgacaatct atcgattgta tgggaagccc gatgcgccg   | 10080 |
| [0666] | agttgtttct gaaacatggc aaaggtagcg ttgccaatga tgttacagat gagatggtca  | 10140 |
| [0667] | gactaaactg gctgacggaa tttatgcctc ttccgaccat caagcatttt atccgtactc  | 10200 |
| [0668] | ctgatgatgc atggttactc accactgcga tccccgggaa aacagcattc caggtattag  | 10260 |
| [0669] | aagaatatcc tgattcaggt gaaaatattg ttgatgcgct ggcagtgctc ctgcgccggt  | 10320 |
| [0670] | tgcattcgat tcctgtttgt aattgtcctt ttaacagcga tcgcgtattt cgtctcgctc  | 10380 |
| [0671] | aggcgcaatc acgaatgaat aacggtttgg ttgatgcgag tgattttgat gacgagcgta  | 10440 |

[0672] atggctggcc tgttgaacaa gtctggaaag aaatgcataa gcttttgcca ttctcaccgg 10500  
 [0673] attcagtcgt cactcatggt gatttttcac ttgataacct tatttttgac gaggggaaat 10560  
 [0674] taataggttg tattgatgtt ggacgagtcg gaatcgcaga ccgataccag gatcttgcca 10620  
 [0675] tcctatggaa ctgcctcggg gagttttctc cttcattaca gaaacggctt tttcaaaaat 10680  
 [0676] atggatttga taatcctgat atgaataaat tgcagtttca ttgatgctc gatgagtttt 10740  
 [0677] tctaagggcg gcctgccacc atacccacgc cgaaacaagc gctcatgagc ccgaagtggc 10800  
 [0678] gagcccgatc ttcccatcgc gtgatgtcgg cgatataggc gccagcaacc gcacctgtgg 10860  
 [0679] cgccggtgat gagggcgcgc caagtcgacg tccggcagtc 10900  
 [0680] <210> 6  
 [0681] <211> 274  
 [0682] <212> PRT  
 [0683] <213> 人工序列  
 [0684] <220>  
 [0685] <223> 合成的多肽  
 [0686] <400> 6  
 [0687] Met Gly Lys Ser Leu Ser His Leu Pro Leu His Ser Ser Lys Glu Asp  
 [0688] 1 5 10 15  
 [0689] Ala Tyr Asp Gly Val Thr Ser Glu Asn Met Arg Asn Gly Leu Val Asn  
 [0690] 20 25 30  
 [0691] Ser Glu Val His Asn Glu Asp Gly Arg Asn Gly Asp Val Ser Gln Phe  
 [0692] 35 40 45  
 [0693] Pro Tyr Val Glu Phe Thr Gly Arg Asp Ser Val Thr Cys Pro Thr Cys  
 [0694] 50 55 60  
 [0695] Gln Gly Thr Gly Arg Ile Pro Arg Gly Gln Glu Asn Gln Leu Val Ala  
 [0696] 65 70 75 80  
 [0697] Leu Ile Pro Tyr Ser Asp Gln Arg Leu Arg Pro Arg Arg Thr Lys Leu  
 [0698] 85 90 95  
 [0699] Tyr Val Met Ala Ser Val Phe Val Cys Leu Leu Leu Ser Gly Leu Ala  
 [0700] 100 105 110  
 [0701] Val Phe Phe Leu Phe Pro Arg Ser Ile Asp Val Lys Tyr Ile Gly Val  
 [0702] 115 120 125  
 [0703] Lys Ser Ala Tyr Val Ser Tyr Asp Val Gln Lys Arg Thr Ile Tyr Leu  
 [0704] 130 135 140  
 [0705] Asn Ile Thr Asn Thr Leu Asn Ile Thr Asn Asn Asn Tyr Tyr Ser Val  
 [0706] 145 150 155 160  
 [0707] Glu Val Glu Asn Ile Thr Ala Gln Val Gln Phe Ser Lys Thr Val Ile  
 [0708] 165 170 175  
 [0709] Gly Lys Ala Arg Leu Asn Asn Ile Thr Ile Ile Gly Pro Leu Asp Met  
 [0710] 180 185 190  
 [0711] Lys Gln Ile Asp Tyr Thr Val Pro Thr Val Ile Ala Glu Glu Met Ser  
 [0712] 195 200 205  
 [0713] Tyr Met Tyr Asp Phe Cys Thr Leu Ile Ser Ile Lys Val His Asn Ile

|        |                                                                    |      |     |
|--------|--------------------------------------------------------------------|------|-----|
| [0714] | 210                                                                | 215  | 220 |
| [0715] | Val Leu Met Met Gln Val Thr Val Thr Thr Thr Tyr Phe Gly His Ser    |      |     |
| [0716] | 225                                                                | 230  | 235 |
| [0717] | Glu Gln Ile Ser Gln Glu Arg Tyr Gln Tyr Val Asp Cys Gly Arg Asn    |      |     |
| [0718] | 245                                                                | 250  | 255 |
| [0719] | Thr Thr Tyr Gln Leu Gly Gln Ser Glu Tyr Leu Asn Val Leu Gln Pro    |      |     |
| [0720] | 260                                                                | 265  | 270 |
| [0721] | Gln Gln                                                            |      |     |
| [0722] | <210> 7                                                            |      |     |
| [0723] | <211> 6514                                                         |      |     |
| [0724] | <212> DNA                                                          |      |     |
| [0725] | <213> 人工序列                                                         |      |     |
| [0726] | <220>                                                              |      |     |
| [0727] | <223> 合成的多核苷酸                                                      |      |     |
| [0728] | <400> 7                                                            |      |     |
| [0729] | aggcgcggac gcaggttaca gcagcgcttg gcctctgctg atgccgtcgt tatectaccc  | 60   |     |
| [0730] | ctcccccgtc ccagctctac ggcgggccgcg cgctccaggc cggtcgctcc accccccggc | 120  |     |
| [0731] | tcccgggact gtggactcca cgacctgtc ctcggccctg tccgcgccga agcagcccgg   | 180  |     |
| [0732] | gactgcgcag cgccccgcgt gccgatcttt tcctaattca gcagcgattt aaccaagagc  | 240  |     |
| [0733] | ctggaatatt ttaaggagta ataagagaca ttacaaaact attctctctg aagcctgcta  | 300  |     |
| [0734] | cctggaggca tcatctagat aatcagaacc ttggttcca catectctc ctttgtctta    | 360  |     |
| [0735] | actacaaaca tttctttctg ctgacttcaa ctctcagac atgggaaagt ctctttctca   | 420  |     |
| [0736] | tttgcttttg cattcaagca aagaagatgc ttatgatgga gtcacatctg aaaacatgag  | 480  |     |
| [0737] | gaatggactg gttaatagtg aagtccataa tgaagatgga agaaatggag atgtctctca  | 540  |     |
| [0738] | gtttccatat gtggaattta caggaagaga tagtgtcacc tgccctactt gtcagggaac  | 600  |     |
| [0739] | aggaagaatt cctagggggc aagaaaacca actggtggca ttgattccat atagtgatca  | 660  |     |
| [0740] | gagattaagg ccaagaagaa caaagctgta tgtgatggct tctgtgttg tctgtctact   | 720  |     |
| [0741] | cctttctgga ttggctgtgt ttttcctttt ccctcgctct atcgacgtga aatacattgg  | 780  |     |
| [0742] | tgtaaaatca gcctatgtca gttatgatgt tcagaagcgt acaatttatt taaatatcac  | 840  |     |
| [0743] | aaacacacta aatataacaa acaataacta ttactctgtc gaagttgaaa acatcactgc  | 900  |     |
| [0744] | ccaagttcaa ttttcaaaaa cagttatttg aaaggcacgc ttaaacaaca taaccattat  | 960  |     |
| [0745] | tggtccactt gatatgaaac aaattgatta cacagtacct accgttatag cagaggaaat  | 1020 |     |
| [0746] | gagttatatg tatgatttct gtactctgat atccatcaaa gtgcataaca tagtactcat  | 1080 |     |
| [0747] | gatgcaagtt actgtgacaa caacatactt tggccactct gaacagatat cccaggagag  | 1140 |     |
| [0748] | gtatcagtat gtcgactgtg gaagaaacac aacttatcag ttggggcagt ctgaatattt  | 1200 |     |
| [0749] | aaatgtactt cagccacaac agtaaaaact ggaagagatg gatttaaaga agaaatatct  | 1260 |     |
| [0750] | attgatattt cctatactct caatgaagag gtatttccta ataggagacc ttaaattgaa  | 1320 |     |
| [0751] | caaacctaaa gtttacactt ctaagagtac agttaaagt atgtggacct gcagttcttg   | 1380 |     |
| [0752] | taactctcca ctctgtgtta atgatatatt tgtactagga tcttttactt gaacttaaat  | 1440 |     |
| [0753] | ttactggttg atttccttct ccagcctatc ccctacaggg aaaagctgat acttccccta  | 1500 |     |
| [0754] | tagtacaata aataattatt taaaagtcac agctccagtc actactgaaa acataatttt  | 1560 |     |
| [0755] | ggtgataaac ataatttgag aaacttaatt tctgaatgtt tttatagaaa attactgaaa  | 1620 |     |

|        |                                                                     |      |
|--------|---------------------------------------------------------------------|------|
| [0756] | gtctattact catggaagac ttttaaagaa taaccttttt tcctgtttta taaattccca   | 1680 |
| [0757] | ttgttatatg gtagtatttc agctacacaa tatttttagct tttagctaga catttatagc  | 1740 |
| [0758] | ttttcatttg ttgaaatggt aatcatctgc atgtttttgt cacttatttc aggttagtga   | 1800 |
| [0759] | ttgcctaaca cttataagcc aaaataatct ttgcaaaatt ccatacctaa aattttgaaa   | 1860 |
| [0760] | gcccctaatag ttttcacaca tctttctgta ttagttatag ttttgtgaaa tctttgtgtg  | 1920 |
| [0761] | atcttcaaac attatcattt aatgtacaat actgtaaata aactgtgcat ggcttttata   | 1980 |
| [0762] | cagcttttagt aaatgtcaaa taaagtggta cagactcatt acaacaagtt tctcataaaa  | 2040 |
| [0763] | atacaataaa taggaaaatg aaattcagaa acccatagac tgggaatagg ttccagttac   | 2100 |
| [0764] | agcttggatc tggcataaaa taaatttgaa ataaaaatatt ttgatgctcc atttttttat  | 2160 |
| [0765] | gttgcttttc atactaaaga atgggtgtaga catgttttgc aactgttagg taccagttta  | 2220 |
| [0766] | tcaattttat caatgtttta gaggaggaaa ttattttttt ggtagaaatt gttcaagaaa   | 2280 |
| [0767] | tccttaattg aatgtcatta aatgatgggt gccaaaataa aacctattta gaaatttaat   | 2340 |
| [0768] | cactttgcac atcacttga atatgatgcc tctagtagtt acttttttat agttttctac    | 2400 |
| [0769] | ttttggtttt atttaaaatt gttttcaaat atagattatt gacttattca actttgctgt   | 2460 |
| [0770] | tttatatttt cagtatcatt ttctattttt tttttttttt gtcttttcac ttaccaagtt   | 2520 |
| [0771] | ctagggacat ttaaaatag tactaagtgt aggagtgggt atgataccaa aaaatgtagc    | 2580 |
| [0772] | tgggttgaga ttaatttcgt tctgttttct catgacagaa atcaggtttc cctttcccca   | 2640 |
| [0773] | cccctaagtg cctaacttag gtctgaaaca gcctgtttat tagtctgact ctctcaacca   | 2700 |
| [0774] | taaaacataa gctttattta attctgcctt taaacacact caggtttccc cttaattttc   | 2760 |
| [0775] | atattatttt ctgcaagttt tcttgagtat cttcaattcg ttgaatgtgg tttttggttt   | 2820 |
| [0776] | ttttttgttt taacactagt cttcccttaa ttcatgtcta actcaagcca tcttactat    | 2880 |
| [0777] | taaacccaaa tcagtccttt aagttcatta tggcctttct agtattttaa aaaaaaaaaa   | 2940 |
| [0778] | tcattttcat ttttctctg ctacgtttcc tgactactac tgcatacttc tctgatacag    | 3000 |
| [0779] | gttctgtttg tattttttat atcattctca ttttctcatt tgacatgac tatgtctata    | 3060 |
| [0780] | tatgatatag gtcccctttt gtctcaaaat ttttaattat gtgacttcaa aaatcacctg   | 3120 |
| [0781] | tatctgtagt agggcttcca aatctgcttc tccatatgtg accagtcacc tgtctgcttt   | 3180 |
| [0782] | cacatttagc tagtgaacta cacatttact aaaatgtgta aattttacac atttagtgac   | 3240 |
| [0783] | tgtgtaaaaat aaaaaaaaaag ttattttatc atatcctttc tattatgttc ccatcctgtc | 3300 |
| [0784] | ctcatgtccc atttacttta ttatcaccat tcatttcttc aaaattatct tttagatacg   | 3360 |
| [0785] | ctcatacaaa aatcaatcct tgttttcttg cttgtgtctt ttaaccttgg aaaattacat   | 3420 |
| [0786] | cgtgtaaaatt aaacagattt ttctgatgat ctgtgcttct tatatactat tagagtgc    | 3480 |
| [0787] | gatagtatct cctgaaaagg atggaaagta gaagcatttg ctttttagtca cttaattttg  | 3540 |
| [0788] | aatctttttt cttcatcttt tgaattaatt ttttttatta tatctacttt tagtggagtt   | 3600 |
| [0789] | tgagtcagaa aaaaacaaga atttgaaaca agtaaaaaga tagaagagaa ataaagatgg   | 3660 |
| [0790] | tatgtgacta ctttcagaga gagttaagta actgtcagaa taagcctgga acaaaacagg   | 3720 |
| [0791] | ctgtaaatta ataaaactac aaacacacat tcagggtgaag cagaagtata gccataaac   | 3780 |
| [0792] | atctagaaag agtgaatgag gcttttagct tttcttaggt caatgtccag tgtgcttttt   | 3840 |
| [0793] | tccatgggaa taggataggt attaatagc ttttctaaac tgcctctaga ccttatccag    | 3900 |
| [0794] | aggacatgggt aaagatatgt tacagaaatt tttctgatac ttcttgaat aactttaagt   | 3960 |
| [0795] | tacaccctag tagactggtc attctaataa aatccagtac tataacaaac ctctgtatgt   | 4020 |
| [0796] | tgatagcaca ttggcccttt ttagagttct ttcctatgtt tttcttacgt gatttccac    | 4080 |
| [0797] | agttccatga gtccaacaaa ggagagtgat aggctcctta tcttttagaa gaggaaggaa   | 4140 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [0798] | aggcatgaag aagttgaggg actggctgaa gatcacgtac ttactaagta gtacaactgg  | 4200 |
| [0799] | agcaagatca agtatctctg tctcccatat ctgtgttcta tcatttaaaa tatatattgg  | 4260 |
| [0800] | aatccctgc tgactcagat tggtagatt aaaaatgaga ggaaagttca aatagttagt    | 4320 |
| [0801] | agtgacaaac taatactgct ggactaagat tttggtagca ttgttttcta aaatatttta  | 4380 |
| [0802] | aatggagaat gaacacttat aaaatgcttt ggaacataat ctttagctta attttctgtt  | 4440 |
| [0803] | aaaatttagt accccttcat cattccaata aagataagac tgatccattg tctaaggaaa  | 4500 |
| [0804] | ttatttataa ataatagaga ttaatttatt tgagatttga aataagaata gtatgaaaat  | 4560 |
| [0805] | attagatacc acataaatg tttgaaatta ctgaataacc atcttaagta tggaacattt   | 4620 |
| [0806] | aaatggctat attttatttg tgtacagttt ttctgtgcct tgtaggcca gtgaagcaat   | 4680 |
| [0807] | tattttctct aagaaaatga caataaaata taacacactt cagattgtct gatttacagt  | 4740 |
| [0808] | ttggaaagga caccgcaatg ttcaaatagg taggagacca tcaaaaacac aattaaagta  | 4800 |
| [0809] | acatattagg agacttgaaa cttcagccta ataaatcctt catggttctt agccttatta  | 4860 |
| [0810] | ttgtgatata attctagata ttttcttga gggcatgtgc ccaactctcc cgcaccccat   | 4920 |
| [0811] | ttgtttgtc ttttaaagtt cttagaataa acagttcttt atataataat tatattttat   | 4980 |
| [0812] | ttaagaaaat agtttgttag gtacttttta aaagatgtaa atttttaaat ttacaaatac  | 5040 |
| [0813] | atatgggtct ttgataagca ataggaattg aattacaagt tactagggtt ataagcaaaa  | 5100 |
| [0814] | ggttgcttac cataatgtca ttaggtcacg attttttagct cacatctgga agcagcaact | 5160 |
| [0815] | acttggtcga agtacatata agagtaatta gttttattct ctctttttta taaaatcggg  | 5220 |
| [0816] | tttcagatga gatgtttatc ttagactatt ttagggaaaa atttttacatg tttagatgg  | 5280 |
| [0817] | tggagtaaaa agactgttaa acatttcttt taaaaatta tttttacatt acaacaatat   | 5340 |
| [0818] | atztatgatg tgttcagatc aaaaatttaa cttctgtgtc ccagatctac tttcaaagt   | 5400 |
| [0819] | agattttcac ttgtcagctt aaatttctga ctagaactaa catttgtgta tttttgtgct  | 5460 |
| [0820] | tagtcggaat acaaatttca cagtggattt ttgaagttg tccttaaatt ggataaaatc   | 5520 |
| [0821] | aagtatttaa agttactaaa gagataaaaa tggtaatttc cattttttaa agtaatttgg  | 5580 |
| [0822] | ttgtgtttat agttatttgt acaagtattt atcacagact ctaaattgaa aaatgtagta  | 5640 |
| [0823] | tgatctatat ttgaccctaa aaatgttga ttaatttaac aaatatggca gatttttcat   | 5700 |
| [0824] | aactaagtct taagtcttct aaaaggaagc tgttaccctt ctgtttttta ttacattaat  | 5760 |
| [0825] | tgaaatgtgt ttaagagat acaatttcag catattttat atattaaaaa aaaaaaagg    | 5820 |
| [0826] | attagtattg agccagtggc caaaaggtaa tattactacc atgtagactg ttatagtcca  | 5880 |
| [0827] | aattgtccca cttcaccag aatttttagaa actagaagtc tgggaggtac tatatcagct  | 5940 |
| [0828] | gtagtgggt aattccaagt gctgatagta ctattcatct tttttattat tgtgtcagat   | 6000 |
| [0829] | gaaacaaatg ccaagttgca aaatatgcag atttttatta tataatggtt ttaggcataa  | 6060 |
| [0830] | attattaaca agccatgcct tatgtgtttc atcttataatt tttctttaga actaaactat | 6120 |
| [0831] | aacagatttt ggaaaatgat ttgacgtgct tgctcacttg attgacttgg tcagatattt  | 6180 |
| [0832] | gaatgatggt attacctaga ttctaactct tgattctagt tatataataa ataatataga  | 6240 |
| [0833] | atatgaaaat atgtttgggc atttactgtt tatattatgt agtagcctcc atcatgacac  | 6300 |
| [0834] | acttactaca tttatgaatt gagcagttct gtaattgtaa ttattattgc tgttcatgta  | 6360 |
| [0835] | acaaaacatg cttataatag caaacaata gaaatgcccc caaaatgcta tttttttaat   | 6420 |
| [0836] | tcagttataa ctgttactct tgtagttgtg tatgacgcaa taaaatttgt aaaaaattt   | 6480 |
| [0837] | cagcatgaaa aataaaattt gtatcactta tgta                              | 6514 |
| [0838] | <210> 8                                                            |      |
| [0839] | <211> 1313                                                         |      |

|        |             |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
|--------|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| [0840] | <212> PRT   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| [0841] | <213> 人工序列  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| [0842] | <220>       |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| [0843] | <223> 合成的多肽 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| [0844] | <400> 8     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| [0845] | Met         | Arg | Ser | Ala | Ala | Ala | Ala | Pro | Arg | Ser | Pro | Ala | Val | Ala | Thr | Glu |  |  |  |  |
| [0846] | 1           |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |  |  |  |  |
| [0847] | Ser         | Arg | Arg | Phe | Ala | Ala | Ala | Arg | Trp | Pro | Gly | Trp | Arg | Ser | Leu | Gln |  |  |  |  |
| [0848] |             |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |  |  |  |  |
| [0849] | Arg         | Pro | Ala | Arg | Arg | Ser | Gly | Arg | Gly | Gly | Gly | Gly | Ala | Ala | Pro | Gly |  |  |  |  |
| [0850] |             |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |  |  |  |  |
| [0851] | Pro         | Tyr | Pro | Ser | Ala | Ala | Pro | Pro | Pro | Pro | Gly | Pro | Gly | Pro | Pro | Pro |  |  |  |  |
| [0852] |             |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |  |  |  |  |
| [0853] | Ser         | Arg | Gln | Ser | Ser | Pro | Pro | Ser | Ala | Ser | Asp | Cys | Phe | Gly | Ser | Asn |  |  |  |  |
| [0854] | 65          |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |  |  |  |  |
| [0855] | Gly         | Asn | Gly | Gly | Gly | Ala | Phe | Arg | Pro | Gly | Ser | Arg | Arg | Leu | Leu | Gly |  |  |  |  |
| [0856] |             |     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |  |  |  |  |
| [0857] | Leu         | Gly | Gly | Pro | Pro | Arg | Pro | Phe | Val | Val | Leu | Leu | Leu | Pro | Leu | Ala |  |  |  |  |
| [0858] |             |     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |  |  |  |  |
| [0859] | Ser         | Pro | Gly | Ala | Pro | Pro | Ala | Ala | Pro | Thr | Arg | Ala | Ser | Pro | Leu | Gly |  |  |  |  |
| [0860] |             |     |     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |  |  |  |  |
| [0861] | Ala         | Arg | Ala | Ser | Pro | Pro | Arg | Ser | Gly | Val | Ser | Leu | Ala | Arg | Pro | Ala |  |  |  |  |
| [0862] |             |     |     |     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |  |  |  |  |
| [0863] | Pro         | Gly | Cys | Pro | Arg | Pro | Ala | Cys | Glu | Pro | Val | Tyr | Gly | Pro | Leu | Thr |  |  |  |  |
| [0864] | 145         |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |  |  |  |  |
| [0865] | Met         | Ser | Leu | Lys | Pro | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln |  |  |  |  |
| [0866] |             |     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |  |  |  |  |
| [0867] | Gln         | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Gln | Pro | Pro | Pro | Ala |  |  |  |  |
| [0868] |             |     |     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |  |  |  |  |
| [0869] | Ala         | Ala | Asn | Val | Arg | Lys | Pro | Gly | Gly | Ser | Gly | Leu | Leu | Ala | Ser | Pro |  |  |  |  |
| [0870] |             |     |     |     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |  |  |  |  |
| [0871] | Ala         | Ala | Ala | Pro | Ser | Pro | Ser | Ser | Ser | Ser | Val | Ser | Ser | Ser | Ser | Ala |  |  |  |  |
| [0872] |             |     |     |     |     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |  |  |  |  |
| [0873] | Thr         | Ala | Pro | Ser | Ser | Val | Val | Ala | Ala | Thr | Ser | Gly | Gly | Gly | Arg | Pro |  |  |  |  |
| [0874] | 225         |     |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     | 240 |  |  |  |  |
| [0875] | Gly         | Leu | Gly | Arg | Gly | Arg | Asn | Ser | Asn | Lys | Gly | Leu | Pro | Gln | Ser | Thr |  |  |  |  |
| [0876] |             |     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |  |  |  |  |
| [0877] | Ile         | Ser | Phe | Asp | Gly | Ile | Tyr | Ala | Asn | Met | Arg | Met | Val | His | Ile | Leu |  |  |  |  |
| [0878] |             |     |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |  |  |  |  |
| [0879] | Thr         | Ser | Val | Val | Gly | Ser | Lys | Cys | Glu | Val | Gln | Val | Lys | Asn | Gly | Gly |  |  |  |  |
| [0880] |             |     |     |     |     | 275 |     |     |     |     |     | 280 |     |     |     | 285 |  |  |  |  |
| [0881] | Ile         | Tyr | Glu | Gly | Val | Phe | Lys | Thr | Tyr | Ser | Pro | Lys | Cys | Asp | Leu | Val |  |  |  |  |

|        |                                                                 |     |             |
|--------|-----------------------------------------------------------------|-----|-------------|
| [0882] | 290                                                             | 295 | 300         |
| [0883] | Leu Asp Ala Ala His Glu Lys Ser Thr Glu Ser Ser Ser Gly Pro Lys |     |             |
| [0884] | 305                                                             | 310 | 315 320     |
| [0885] | Arg Glu Glu Ile Met Glu Ser Ile Leu Phe Lys Cys Ser Asp Phe Val |     |             |
| [0886] |                                                                 | 325 | 330 335     |
| [0887] | Val Val Gln Phe Lys Asp Met Asp Ser Ser Tyr Ala Lys Arg Asp Ala |     |             |
| [0888] |                                                                 | 340 | 345 350     |
| [0889] | Phe Thr Asp Ser Ala Ile Ser Ala Lys Val Asn Gly Glu His Lys Glu |     |             |
| [0890] |                                                                 | 355 | 360 365     |
| [0891] | Lys Asp Leu Glu Pro Trp Asp Ala Gly Glu Leu Thr Ala Asn Glu Glu |     |             |
| [0892] |                                                                 | 370 | 375 380     |
| [0893] | Leu Glu Ala Leu Glu Asn Asp Val Ser Asn Gly Trp Asp Pro Asn Asp |     |             |
| [0894] |                                                                 | 385 | 390 395 400 |
| [0895] | Met Phe Arg Tyr Asn Glu Glu Asn Tyr Gly Val Val Ser Thr Tyr Asp |     |             |
| [0896] |                                                                 | 405 | 410 415     |
| [0897] | Ser Ser Leu Ser Ser Tyr Thr Val Pro Leu Glu Arg Asp Asn Ser Glu |     |             |
| [0898] |                                                                 | 420 | 425 430     |
| [0899] | Glu Phe Leu Lys Arg Glu Ala Arg Ala Asn Gln Leu Ala Glu Glu Ile |     |             |
| [0900] |                                                                 | 435 | 440 445     |
| [0901] | Glu Ser Ser Ala Gln Tyr Lys Ala Arg Val Ala Leu Glu Asn Asp Asp |     |             |
| [0902] |                                                                 | 450 | 455 460     |
| [0903] | Arg Ser Glu Glu Glu Lys Tyr Thr Ala Val Gln Arg Asn Ser Ser Glu |     |             |
| [0904] |                                                                 | 465 | 470 475 480 |
| [0905] | Arg Glu Gly His Ser Ile Asn Thr Arg Glu Asn Lys Tyr Ile Pro Pro |     |             |
| [0906] |                                                                 | 485 | 490 495     |
| [0907] | Gly Gln Arg Asn Arg Glu Val Ile Ser Trp Gly Ser Gly Arg Gln Asn |     |             |
| [0908] |                                                                 | 500 | 505 510     |
| [0909] | Ser Pro Arg Met Gly Gln Pro Gly Ser Gly Ser Met Pro Ser Arg Ser |     |             |
| [0910] |                                                                 | 515 | 520 525     |
| [0911] | Thr Ser His Thr Ser Asp Phe Asn Pro Asn Ser Gly Ser Asp Gln Arg |     |             |
| [0912] |                                                                 | 530 | 535 540     |
| [0913] | Val Val Asn Gly Gly Val Pro Trp Pro Ser Pro Cys Pro Ser Pro Ser |     |             |
| [0914] |                                                                 | 545 | 550 555 560 |
| [0915] | Ser Arg Pro Pro Ser Arg Tyr Gln Ser Gly Pro Asn Ser Leu Pro Pro |     |             |
| [0916] |                                                                 | 565 | 570 575     |
| [0917] | Arg Ala Ala Thr Pro Thr Arg Pro Pro Ser Arg Pro Pro Ser Arg Pro |     |             |
| [0918] |                                                                 | 580 | 585 590     |
| [0919] | Ser Arg Pro Pro Ser His Pro Ser Ala His Gly Ser Pro Ala Pro Val |     |             |
| [0920] |                                                                 | 595 | 600 605     |
| [0921] | Ser Thr Met Pro Lys Arg Met Ser Ser Glu Gly Pro Pro Arg Met Ser |     |             |
| [0922] |                                                                 | 610 | 615 620     |
| [0923] | Pro Lys Ala Gln Arg His Pro Arg Asn His Arg Val Ser Ala Gly Arg |     |             |

|        |                                                                 |     |     |     |
|--------|-----------------------------------------------------------------|-----|-----|-----|
| [0924] | 625                                                             | 630 | 635 | 640 |
| [0925] | Gly Ser Ile Ser Ser Gly Leu Glu Phe Val Ser His Asn Pro Pro Ser |     |     |     |
| [0926] |                                                                 | 645 | 650 | 655 |
| [0927] | Glu Ala Ala Thr Pro Pro Val Ala Arg Thr Ser Pro Ser Gly Gly Thr |     |     |     |
| [0928] |                                                                 | 660 | 665 | 670 |
| [0929] | Trp Ser Ser Val Val Ser Gly Val Pro Arg Leu Ser Pro Lys Thr His |     |     |     |
| [0930] |                                                                 | 675 | 680 | 685 |
| [0931] | Arg Pro Arg Ser Pro Arg Gln Asn Ser Ile Gly Asn Thr Pro Ser Gly |     |     |     |
| [0932] |                                                                 | 690 | 695 | 700 |
| [0933] | Pro Val Leu Ala Ser Pro Gln Ala Gly Ile Ile Pro Thr Glu Ala Val |     |     |     |
| [0934] | 705                                                             | 710 | 715 | 720 |
| [0935] | Ala Met Pro Ile Pro Ala Ala Ser Pro Thr Pro Ala Ser Pro Ala Ser |     |     |     |
| [0936] |                                                                 | 725 | 730 | 735 |
| [0937] | Asn Arg Ala Val Thr Pro Ser Ser Glu Ala Lys Asp Ser Arg Leu Gln |     |     |     |
| [0938] |                                                                 | 740 | 745 | 750 |
| [0939] | Asp Gln Arg Gln Asn Ser Pro Ala Gly Asn Lys Glu Asn Ile Lys Pro |     |     |     |
| [0940] |                                                                 | 755 | 760 | 765 |
| [0941] | Asn Glu Thr Ser Pro Ser Phe Ser Lys Ala Glu Asn Lys Gly Ile Ser |     |     |     |
| [0942] |                                                                 | 770 | 775 | 780 |
| [0943] | Pro Val Val Ser Glu His Arg Lys Gln Ile Asp Asp Leu Lys Lys Phe |     |     |     |
| [0944] | 785                                                             | 790 | 795 | 800 |
| [0945] | Lys Asn Asp Phe Arg Leu Gln Pro Ser Ser Thr Ser Glu Ser Met Asp |     |     |     |
| [0946] |                                                                 | 805 | 810 | 815 |
| [0947] | Gln Leu Leu Asn Lys Asn Arg Glu Gly Glu Lys Ser Arg Asp Leu Ile |     |     |     |
| [0948] |                                                                 | 820 | 825 | 830 |
| [0949] | Lys Asp Lys Ile Glu Pro Ser Ala Lys Asp Ser Phe Ile Glu Asn Ser |     |     |     |
| [0950] |                                                                 | 835 | 840 | 845 |
| [0951] | Ser Ser Asn Cys Thr Ser Gly Ser Ser Lys Pro Asn Ser Pro Ser Ile |     |     |     |
| [0952] |                                                                 | 850 | 855 | 860 |
| [0953] | Ser Pro Ser Ile Leu Ser Asn Thr Glu His Lys Arg Gly Pro Glu Val |     |     |     |
| [0954] | 865                                                             | 870 | 875 | 880 |
| [0955] | Thr Ser Gln Gly Val Gln Thr Ser Ser Pro Ala Cys Lys Gln Glu Lys |     |     |     |
| [0956] |                                                                 | 885 | 890 | 895 |
| [0957] | Asp Asp Lys Glu Glu Lys Lys Asp Ala Ala Glu Gln Val Arg Lys Ser |     |     |     |
| [0958] |                                                                 | 900 | 905 | 910 |
| [0959] | Thr Leu Asn Pro Asn Ala Lys Glu Phe Asn Pro Arg Ser Phe Ser Gln |     |     |     |
| [0960] |                                                                 | 915 | 920 | 925 |
| [0961] | Pro Lys Pro Ser Thr Thr Pro Thr Ser Pro Arg Pro Gln Ala Gln Pro |     |     |     |
| [0962] |                                                                 | 930 | 935 | 940 |
| [0963] | Ser Pro Ser Met Val Gly His Gln Gln Pro Thr Pro Val Tyr Thr Gln |     |     |     |
| [0964] | 945                                                             | 950 | 955 | 960 |
| [0965] | Pro Val Cys Phe Ala Pro Asn Met Met Tyr Pro Val Pro Val Ser Pro |     |     |     |

|        |                                                                 |      |  |      |  |      |
|--------|-----------------------------------------------------------------|------|--|------|--|------|
| [0966] |                                                                 | 965  |  | 970  |  | 975  |
| [0967] | Gly Val Gln Pro Leu Tyr Pro Ile Pro Met Thr Pro Met Pro Val Asn |      |  |      |  |      |
| [0968] |                                                                 | 980  |  | 985  |  | 990  |
| [0969] | Gln Ala Lys Thr Tyr Arg Ala Val Pro Asn Met Pro Gln Gln Arg Gln |      |  |      |  |      |
| [0970] |                                                                 | 995  |  | 1000 |  | 1005 |
| [0971] | Asp Gln His His Gln Ser Ala Met Met His Pro Ala Ser Ala Ala     |      |  |      |  |      |
| [0972] |                                                                 | 1010 |  | 1015 |  | 1020 |
| [0973] | Gly Pro Pro Ile Ala Ala Thr Pro Pro Ala Tyr Ser Thr Gln Tyr     |      |  |      |  |      |
| [0974] |                                                                 | 1025 |  | 1030 |  | 1035 |
| [0975] | Val Ala Tyr Ser Pro Gln Gln Phe Pro Asn Gln Pro Leu Val Gln     |      |  |      |  |      |
| [0976] |                                                                 | 1040 |  | 1045 |  | 1050 |
| [0977] | His Val Pro His Tyr Gln Ser Gln His Pro His Val Tyr Ser Pro     |      |  |      |  |      |
| [0978] |                                                                 | 1055 |  | 1060 |  | 1065 |
| [0979] | Val Ile Gln Gly Asn Ala Arg Met Met Ala Pro Pro Thr His Ala     |      |  |      |  |      |
| [0980] |                                                                 | 1070 |  | 1075 |  | 1080 |
| [0981] | Gln Pro Gly Leu Val Ser Ser Ser Ala Thr Gln Tyr Gly Ala His     |      |  |      |  |      |
| [0982] |                                                                 | 1085 |  | 1090 |  | 1095 |
| [0983] | Glu Gln Thr His Ala Met Tyr Ala Cys Pro Lys Leu Pro Tyr Asn     |      |  |      |  |      |
| [0984] |                                                                 | 1100 |  | 1105 |  | 1110 |
| [0985] | Lys Glu Thr Ser Pro Ser Phe Tyr Phe Ala Ile Ser Thr Gly Ser     |      |  |      |  |      |
| [0986] |                                                                 | 1115 |  | 1120 |  | 1125 |
| [0987] | Leu Ala Gln Gln Tyr Ala His Pro Asn Ala Thr Leu His Pro His     |      |  |      |  |      |
| [0988] |                                                                 | 1130 |  | 1135 |  | 1140 |
| [0989] | Thr Pro His Pro Gln Pro Ser Ala Thr Pro Thr Gly Gln Gln Gln     |      |  |      |  |      |
| [0990] |                                                                 | 1145 |  | 1150 |  | 1155 |
| [0991] | Ser Gln His Gly Gly Ser His Pro Ala Pro Ser Pro Val Gln His     |      |  |      |  |      |
| [0992] |                                                                 | 1160 |  | 1165 |  | 1170 |
| [0993] | His Gln His Gln Ala Ala Gln Ala Leu His Leu Ala Ser Pro Gln     |      |  |      |  |      |
| [0994] |                                                                 | 1175 |  | 1180 |  | 1185 |
| [0995] | Gln Gln Ser Ala Ile Tyr His Ala Gly Leu Ala Pro Thr Pro Pro     |      |  |      |  |      |
| [0996] |                                                                 | 1190 |  | 1195 |  | 1200 |
| [0997] | Ser Met Thr Pro Ala Ser Asn Thr Gln Ser Pro Gln Asn Ser Phe     |      |  |      |  |      |
| [0998] |                                                                 | 1205 |  | 1210 |  | 1215 |
| [0999] | Pro Ala Ala Gln Gln Thr Val Phe Thr Ile His Pro Ser His Val     |      |  |      |  |      |
| [1000] |                                                                 | 1220 |  | 1225 |  | 1230 |
| [1001] | Gln Pro Ala Tyr Thr Asn Pro Pro His Met Ala His Val Pro Gln     |      |  |      |  |      |
| [1002] |                                                                 | 1235 |  | 1240 |  | 1245 |
| [1003] | Ala His Val Gln Ser Gly Met Val Pro Ser His Pro Thr Ala His     |      |  |      |  |      |
| [1004] |                                                                 | 1250 |  | 1255 |  | 1260 |
| [1005] | Ala Pro Met Met Leu Met Thr Thr Gln Pro Pro Gly Gly Pro Gln     |      |  |      |  |      |
| [1006] |                                                                 | 1265 |  | 1270 |  | 1275 |
| [1007] | Ala Ala Leu Ala Gln Ser Ala Leu Gln Pro Ile Pro Val Ser Thr     |      |  |      |  |      |

|        |                                                                    |      |      |
|--------|--------------------------------------------------------------------|------|------|
| [1008] | 1280                                                               | 1285 | 1290 |
| [1009] | Thr Ala His Phe Pro Tyr Met Thr His Pro Ser Val Gln Ala His        |      |      |
| [1010] | 1295                                                               | 1300 | 1305 |
| [1011] | His Gln Gln Gln Leu                                                |      |      |
| [1012] | 1310                                                               |      |      |
| [1013] | <210> 9                                                            |      |      |
| [1014] | <211> 4712                                                         |      |      |
| [1015] | <212> DNA                                                          |      |      |
| [1016] | <213> 人工序列                                                         |      |      |
| [1017] | <220>                                                              |      |      |
| [1018] | <223> 合成的多核苷酸                                                      |      |      |
| [1019] | <400> 9                                                            |      |      |
| [1020] | acccccgaga aagcaaccca gcgcgccgcc cgctcctcac gtgtccctcc cggccccggg  | 60   |      |
| [1021] | gccacctcac gttctgcttc cgtctgacct ctccgacttc cggtaaagag tccctatccg  | 120  |      |
| [1022] | cacctccgct cccaccgggc gcctcggcgc gcccgccctc cgatgcgctc agcggccgca  | 180  |      |
| [1023] | gtcctcggga gtcccgcggt ggccaccgag tctcggcgtc tcgccgcagc caggtggccc  | 240  |      |
| [1024] | gggtggcgct cgctccagcg gccggcgccg cggagcgggc ggggcggcgg tggcgccggc  | 300  |      |
| [1025] | ccgggaccgt atccctccgc cgccctccc ccgcccggcc ccggccccc tccctcccgg    | 360  |      |
| [1026] | cagagctcgc ctccctccgc ctccagactgt tttggtagca acggcaacgg cggcgccgcg | 420  |      |
| [1027] | tttcggcccc gctcccgccg gctccttggt ctccggcgcc ctccccgcc cttcgtcgtc   | 480  |      |
| [1028] | ctccttctcc cctcggccag cccggcgcc cctccggccg cgccaacccg cgctccccg    | 540  |      |
| [1029] | ctcggcgccc gcgcgtcccc gccgcgttcc ggcgtctcct tggcgcgccc ggctccccgc  | 600  |      |
| [1030] | tgccccgcc cggcgtgcga gccggtgtat gggccctca ccatgtcgtc gaagccccag    | 660  |      |
| [1031] | cagcagcagc agcagcagca gcagcagcag cagcagcaac agcagcagca gcagcagcag  | 720  |      |
| [1032] | cagcagccgc cgcccgcggc tgccaatgtc cgcaagcccg gcggcagcgg ctttctagcg  | 780  |      |
| [1033] | tcgcccgcgc ccgcgccttc gccgtcctcg tctcgggtct cctcgtcctc ggccacggct  | 840  |      |
| [1034] | ccctcctcgg tggtcgcggc gacctccggc ggccggaggc ccggcctggg cagaggctga  | 900  |      |
| [1035] | aacagtaaca aaggactgcc tcagtctacg atttcttttg atggaatcta tgcaaatatg  | 960  |      |
| [1036] | aggatggttc atatacttac atcagttgtt ggctccaaat gtgaagtaca agtgaaaaat  | 1020 |      |
| [1037] | ggaggtatat atgaaggagt ttttaaaact tacagtccga agtgtgattt ggtacttgat  | 1080 |      |
| [1038] | gccgcacatg agaaaagtag agaatccagt tcggggccga aacgtgaaga aataatggag  | 1140 |      |
| [1039] | agtattttgt tcaaatgttc agactttgtt gtggtacagt ttaaagatat ggactccagt  | 1200 |      |
| [1040] | tatgcaaaaa gagatgcttt tactgactct gctatcagtg ctaaagtga tggcgaacac   | 1260 |      |
| [1041] | aaagagaagg acctggagcc ctgggatgca ggtgaactca cagccaatga ggaacttgag  | 1320 |      |
| [1042] | gctttgaaa atgacgtatc taatggatgg gatcccaatg atatgtttcg atataatgaa   | 1380 |      |
| [1043] | gaaaattatg gtgtagtgtc tacgtatgat agcagtttat cttcgtatac agtgccctta  | 1440 |      |
| [1044] | gaaagagata actcagaaga atttttaaaa cggaagcaa gggcaaacca gttagcagaa   | 1500 |      |
| [1045] | gaaattgagt caagtccca gtacaaagct cgagtggccc tggaaaatga tgataggagt   | 1560 |      |
| [1046] | gaggaagaaa aatacacagc agttcagaga aattccagtg aacgtgagg gcacagcata   | 1620 |      |
| [1047] | aacactaggg aaaataaata tattcctcct ggacaaagaa atagagaagt catatcctgg  | 1680 |      |
| [1048] | ggaagtggga gacagaattc accgcgtatg ggccagcctg gatcgggctc catgccatca  | 1740 |      |
| [1049] | agatccactt ctcacacttc agatttcaac ccgaattctg gttcagacca aagagtagtt  | 1800 |      |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [1050] | aatggaggtg ttccctggcc atcgcttgc ccatctcctt cctctcgccc accttctcgc   | 1860 |
| [1051] | taccagtcag gtcccaactc tttccacct cgggcagcca cccctacacg gccgcctcc    | 1920 |
| [1052] | aggcccccct cggggccatc cagacccccg tctcaccct ctgctcatgg ttctccagct   | 1980 |
| [1053] | cctgtctcta ctatgcctaa acgcatgtct tcagaagggc ctccaaggat gtccccaag   | 2040 |
| [1054] | gcccagcgac atcctcgaaa tcacagagtt tctgtctggga ggggttccat atccagtggc | 2100 |
| [1055] | ctagaatttg tatcccacaa cccaccaggt gaagcagcta ctctccagt agcaaggacc   | 2160 |
| [1056] | agtccctcgg ggggaacgtg gtcatcagtg gtcagtgggg ttccaagatt atcccctaaa  | 2220 |
| [1057] | actcatagac ccaggtctcc cagacagaac agtattggaa atacccccag tgggccagtt  | 2280 |
| [1058] | cttgcttctc cccaagctgg tattattcca actgaagctg ttgccatgcc tattccagct  | 2340 |
| [1059] | gcatctccta cgctgctag tcctgcatcg aacagagctg ttacccttc tagtgaggct    | 2400 |
| [1060] | aaagattcca ggcttcaaga tcagaggcag aactctcctg cagggaataa agaaaatatt  | 2460 |
| [1061] | aaacccaatg aaacatcacc tagcttctca aaagctgaaa acaaaggtat atcaccagtt  | 2520 |
| [1062] | gtttctgaac atagaaaaca gattgatgat ttaaagaaat ttaagaatga ttttaggtta  | 2580 |
| [1063] | cagccaagtt ctacttctga atctatggat caactactaa acaaaaatag agaggagaa   | 2640 |
| [1064] | aatcaagag atttgatcaa agacaaaatt gaaccaagtg ctaaggattc tttcattgaa   | 2700 |
| [1065] | aatagcagca gcaactgtac cagtggcagc agcaagccga atagccccag catttccct   | 2760 |
| [1066] | tcaatactta gtaacacgga gcacaagagg ggacctgagg tcacttcca aggggttcag   | 2820 |
| [1067] | acttccagcc cagcatgtaa acaagagaaa gacgataagg aagagaagaa agacgcagct  | 2880 |
| [1068] | gagcaagtta ggaaatcaac attgaatccc aatgcaaagg agttcaacc acgttcttc    | 2940 |
| [1069] | tctcagccaa agccttctac taccctcaact tcacctcggc ctcaagcaca acctagccca | 3000 |
| [1070] | tctatggtgg gtcatcaaca gccactcca gtttatactc agcctgtttg ttttgacca    | 3060 |
| [1071] | aatatgatgt atccagtccc agtgagccca ggcgtgcaac ctttataccc aatacctatg  | 3120 |
| [1072] | acgcccagtc cagtgaatca agccaagaca tatagagcag taccaaatat gccccaacag  | 3180 |
| [1073] | cggcaagacc agcatcatca gagtgccatg atgcaccag cgtcagcagc gggcccaccg   | 3240 |
| [1074] | attgcagcca ccccaccagc ttactccacg caatatgttg cctacagtcc tcagcagttc  | 3300 |
| [1075] | ccaaatcagc cccttgttca gcatgtgcca cattatcagt ctacagatcc tcatgtctat  | 3360 |
| [1076] | agtcctgtaa tacagggtaa tgctagaatg atggcaccac caacacacgc ccagcctggt  | 3420 |
| [1077] | ttagtatctt cttcagcaac tcagtacggg gtcctagc agacgcagtc gatgtatgca    | 3480 |
| [1078] | tgtcccaaat taccatacaa caaggagaca agccttctt tctactttgc catttccacg   | 3540 |
| [1079] | ggctcccttg ctacagcagta tgcgcacct aacgtaccc tgcaccaca tactccacac    | 3600 |
| [1080] | cctcagcctt cagctacccc cactggacag cagcaaagcc aacatggtgg aagtcacct   | 3660 |
| [1081] | gcaccagtc ctgttcagca ccatcagcac caggccgccc aggctctcca tctggccagt   | 3720 |
| [1082] | ccacagcagc agtcagccat ttaccacgcg gggcttgccg caactccacc ctccatgaca  | 3780 |
| [1083] | cctgcctcca acacgcagtc gccacagaat agtttcccag cagcacaaca gactgtcttt  | 3840 |
| [1084] | acgatccatc cttctcacgt tcagccggcg tataccaacc caccacacat ggcccacgta  | 3900 |
| [1085] | cctcaggctc atgtacagtc aggaatggtt ctttctcctc caactgcccc tgcgccaatg  | 3960 |
| [1086] | atgctaataa cgacacagcc acccgcggt cccaggccg ccctcgtc aagtgcacta      | 4020 |
| [1087] | cagcccatc cagtctcgac aacagcgc atttccctata tgacgcacc ttcagtacaa     | 4080 |
| [1088] | gcccaccacc aacagcagtt gtaagctgc cctggaggaa ccgaaaggcc aaattccctc   | 4140 |
| [1089] | ctcccttcta ctgcttctac caactggaag cacagaaaac tagaatttca tttattttgt  | 4200 |
| [1090] | ttttaaaata tatatgttga tttcttgtaa catccaatag gaatgctaac agttcacttg  | 4260 |
| [1091] | cagtgaaga tacttggacc gagtagaggc atttaggaac ttgggggcta ttccataatt   | 4320 |

|        |                                                                         |
|--------|-------------------------------------------------------------------------|
| [1092] | ccatatgctg tttcagagtc ccgcaggtag cccagctctg cttgccgaaa ctggaagtta 4380  |
| [1093] | tttatttttt aataaccctt gaaagtcattg aacacatcag ctagcaaaag aagtaacaag 4440 |
| [1094] | agtattctt gctgctatta ctgctaaaaa aaaaaaaaaa aaaaaatcaa gacttggaac 4500   |
| [1095] | gcccttttac taaacttgac aaagtttcag taaattctta ccgtcaaact gacggattat 4560  |
| [1096] | tatttataaa tcaagtttga tgaggtagtc actgtctaca gtggttcaac ttttaagtta 4620  |
| [1097] | agggaaaaac ttttactttg tagataatat aaaataaaaa cttaaaaaaa atttaaaaaa 4680  |
| [1098] | taaaaaaagt tttaaaaact gaaaaaaaaa aa 4712                                |
| [1099] | <210> 10                                                                |
| [1100] | <211> 19                                                                |
| [1101] | <212> DNA                                                               |
| [1102] | <213> 人工序列                                                              |
| [1103] | <220>                                                                   |
| [1104] | <223> 合成的多核苷酸                                                           |
| [1105] | <400> 10                                                                |
| [1106] | ttggcctttc ggttcctcc 19                                                 |
| [1107] | <210> 11                                                                |
| [1108] | <211> 19                                                                |
| [1109] | <212> DNA                                                               |
| [1110] | <213> 人工序列                                                              |
| [1111] | <220>                                                                   |
| [1112] | <223> 合成的多核苷酸                                                           |
| [1113] | <400> 11                                                                |
| [1114] | ggaggaaccg aaaggccaa 19                                                 |
| [1115] | <210> 12                                                                |
| [1116] | <211> 148                                                               |
| [1117] | <212> DNA                                                               |
| [1118] | <213> 人工序列                                                              |
| [1119] | <220>                                                                   |
| [1120] | <223> 合成的多核苷酸                                                           |
| [1121] | <400> 12                                                                |
| [1122] | tggaggcttg ctgaaggctg tatgctgttg tcttggcctt tcggttcctc cagtgaagcc 60    |
| [1123] | acagatggga ggaaccgaaa ggccaaagga cacaaggcct gttactagca ctcacatgga 120   |
| [1124] | acaaatggcc accgtgggag gatgacaa 148                                      |
| [1125] | <210> 13                                                                |
| [1126] | <211> 148                                                               |
| [1127] | <212> DNA                                                               |
| [1128] | <213> 人工序列                                                              |
| [1129] | <220>                                                                   |
| [1130] | <223> 合成的多核苷酸                                                           |
| [1131] | <400> 13                                                                |
| [1132] | ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctagtaacag gccttgtgtc 60    |
| [1133] | ctttggcctt tcggttcctc ccattctgtg cttcactgga ggaaccgaaa ggccaagaca 120   |

- [1134] acagcataca gccttcagca agcctcca 148  
[1135] <210> 14  
[1136] <211> 22  
[1137] <212> DNA  
[1138] <213> 人工序列  
[1139] <220>  
[1140] <223> 合成的多核苷酸  
[1141] <400> 14  
[1142] atagcagcaa gaatcactct tg 22  
[1143] <210> 15  
[1144] <211> 22  
[1145] <212> DNA  
[1146] <213> 人工序列  
[1147] <220>  
[1148] <223> 合成的多核苷酸  
[1149] <400> 15  
[1150] caagagtgat tcttgctgct at 22  
[1151] <210> 16  
[1152] <211> 151  
[1153] <212> DNA  
[1154] <213> 人工序列  
[1155] <220>  
[1156] <223> 合成的多核苷酸  
[1157] <400> 16  
[1158] ctggaggett gctttgggct gtatgctgat agcagcaaga atcactcttg ttttggcctc 60  
[1159] tgactgaaca agaggatctt gctgctatca ggacacaagg ccctttatca gcactcacat 120  
[1160] ggaacaaatg gccaccgtgg gaggatgaca a 151  
[1161] <210> 17  
[1162] <211> 151  
[1163] <212> DNA  
[1164] <213> 人工序列  
[1165] <220>  
[1166] <223> 合成的多核苷酸  
[1167] <400> 17  
[1168] ttgtcatcct cccacggtgg ccatttggtc catgtgagtg ctgataaagg gccttgtgtc 60  
[1169] ctgatagcag caagatcctc ttgttcagtc agaggccaaa acaagagtga ttcttgctgc 120  
[1170] tatcagcata cagcccaaag caagcctcca g 151  
[1171] <210> 18  
[1172] <211> 300  
[1173] <212> DNA  
[1174] <213> 人工序列  
[1175] <220>

- [1176] <223> 合成的多核苷酸  
[1177] <400> 18  
[1178] tggaggttg ctgaaggctg tatgctgttg tctactgctc tataatgtctt ggcagtgaag 60  
[1179] ccacagatgg ccaagacata tagagcagta aggacacaag gcctgttact agcactcaca 120  
[1180] tggaaacaaat ggccaccgtg ggaggatgac aatggaggct tgctgaaggc tgtatgctgt 180  
[1181] tgtcttggcc tttcggttcc tccagtgaag ccacagatgg gaggaaccga aaggccaaag 240  
[1182] gacacaaggc ctgttactag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa 300  
[1183] <210> 19  
[1184] <211> 300  
[1185] <212> DNA  
[1186] <213> 人工序列  
[1187] <220>  
[1188] <223> 合成的多核苷酸  
[1189] <400> 19  
[1190] ttgtcatcct cccacggttg ccatttgttc catgtgagtg ctagtaacag gccttgtgtc 60  
[1191] ctttggcctt tcggttcctc ccatctgttg cttcactgga ggaaccgaaa ggccaagaca 120  
[1192] acagcataca gccttcagea agcctccatt gtcacctcc caccgtggcc atttgttcca 180  
[1193] tgtgagtgtc agtaacaggc cttgtgtcct tactgtctta tatgtcttgg ccatctgttg 240  
[1194] cttcactgcc aagacatata gagcagtaga caacagcata cagccttcag caagcctcca 300  
[1195] <210> 20  
[1196] <211> 23  
[1197] <212> DNA  
[1198] <213> 人工序列  
[1199] <220>  
[1200] <223> 合成的多核苷酸  
[1201] <400> 20  
[1202] ttccagttgg tagaagcagt aga 23  
[1203] <210> 21  
[1204] <211> 23  
[1205] <212> DNA  
[1206] <213> 人工序列  
[1207] <220>  
[1208] <223> 合成的多核苷酸  
[1209] <400> 21  
[1210] tctactgctt ctaccaactg gaa 23  
[1211] <210> 22  
[1212] <211> 152  
[1213] <212> DNA  
[1214] <213> 人工序列  
[1215] <220>  
[1216] <223> 合成的多核苷酸  
[1217] <400> 22

[1218] ctggaggctt gctttgggct gtatgctgtt ccagttggta gaagcagtag attttggcct 60  
[1219] ctgactgata tactgttcac caactggaac aggacacaag gccctttatc agcactcaca 120  
[1220] tggaacaaat ggccaccgtg ggaggatgac aa 152  
[1221] <210> 23  
[1222] <211> 152  
[1223] <212> DNA  
[1224] <213> 人工序列  
[1225] <220>  
[1226] <223> 合成的多核苷酸  
[1227] <400> 23  
[1228] ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctgataaagg gccttgtgtc 60  
[1229] ctgttccagt tggatgaacag tagatcagtc agaggccaaa atctactgct tctaccaact 120  
[1230] ggaacagcat acagcccaaa gcaagcctcc ag 152  
[1231] <210> 24  
[1232] <211> 151  
[1233] <212> DNA  
[1234] <213> 人工序列  
[1235] <220>  
[1236] <223> 合成的多核苷酸  
[1237] <400> 24  
[1238] tggaggcttg ctgaagctg tatgctgttg tcggggccgg ttctggggcc ttagtgaagc 60  
[1239] cacagatgta cggccccgaa accggcccca ggacacaagg cctgttacta gcactcat 120  
[1240] ggaacaaatg gccaccgtgg gaggatgaca a 151  
[1241] <210> 25  
[1242] <211> 151  
[1243] <212> DNA  
[1244] <213> 人工序列  
[1245] <220>  
[1246] <223> 合成的多核苷酸  
[1247] <400> 25  
[1248] ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctagtaacag gccttgtgtc 60  
[1249] ctggggccgg ttctggggcc gtacatctgt ggcttacta aggccccgaa accggccccg 120  
[1250] acaacagcat acagccttca gcaagcctcc a 151  
[1251] <210> 26  
[1252] <211> 4145  
[1253] <212> DNA  
[1254] <213> 人工序列  
[1255] <220>  
[1256] <223> 合成的多核苷酸  
[1257] <400> 26  
[1258] ttggccactc cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc 60  
[1259] cgacggccgg gctttgcccg ggccggcctca gtgagcgagc gagcgcgag agaggagtg 120

|        |             |             |            |             |            |            |      |
|--------|-------------|-------------|------------|-------------|------------|------------|------|
| [1260] | gccaaactcca | tcactagggg  | ttcctgctag | ctctgggtat  | ttaagcccga | gtgagcacgc | 180  |
| [1261] | agggtctcca  | ttttgaagcg  | ggaggttacg | cgttcgtcga  | ctactagtgg | gtaccagagc | 240  |
| [1262] | tccttaggtt  | ctagaaccgg  | tgacgtctcc | catggtgaag  | cttgatctg  | aattcggtag | 300  |
| [1263] | cctagttatt  | aatagtaatc  | aattacgggg | tcattagttc  | atagcccata | tatggagttc | 360  |
| [1264] | cgcgttacat  | aacttacggt  | aaatggcccg | cctggctgac  | cgcccaacga | ccccgcgcca | 420  |
| [1265] | ttgacgtcaa  | taatgacgta  | tgttcccata | gtaacgcaa   | tagggacttt | ccattgacgt | 480  |
| [1266] | caatgggtgg  | actatttacg  | gtaaactgcc | cacttggcag  | tacatcaagt | gtatcatatg | 540  |
| [1267] | ccaagtacgc  | cccctattga  | cgtcaatgac | ggtaaattgc  | ccgcctggca | ttatgcccg  | 600  |
| [1268] | tacatgacct  | tatgggactt  | tcctacttgg | cagtacatct  | acgtattagt | catcgctatt | 660  |
| [1269] | accatggctg  | aggtgagccc  | cacgttctgc | ttcactctcc  | ccatctcccc | cccctcccca | 720  |
| [1270] | cccccaattt  | tgtatttatt  | tatttttta  | ttattttgtg  | cagcgatggg | ggcggggggg | 780  |
| [1271] | gggggggggc  | gcgcgccagg  | cggggcgggg | cggggcgagg  | ggcggggcgg | ggcgaggcgg | 840  |
| [1272] | agaggtgcgg  | cggcagccaa  | tcagagcggc | gcgctccgaa  | agtttccitt | tatggcgagg | 900  |
| [1273] | cggcggcggc  | ggcggcccta  | taaaaagcga | agcgcgcggc  | ggcggggagt | cgctgcgacg | 960  |
| [1274] | ctgccttcgc  | cccggtcccc  | gctccgccgc | cgcctcgcgc  | cgcccccccc | ggctctgact | 1020 |
| [1275] | gaccgcgtta  | ctcccacagg  | tgagcgggcg | ggacggccct  | tctctcagc  | gctgtaatta | 1080 |
| [1276] | gcgcttggtt  | taatgacggc  | ttgttgagg  | cttgctgaag  | gctgtatgct | gttgctgggg | 1140 |
| [1277] | ccggtttcgg  | ggccttagtg  | aagccacaga | tgtacggccc  | cgaaccggc  | cccaggacac | 1200 |
| [1278] | aaggcctgtt  | actagcactc  | acatggaaca | aatggccacc  | gtgggaggat | gacaatggag | 1260 |
| [1279] | gcttgctgaa  | ggctgtatgc  | tggtgtcttg | gcctttcggg  | tcctccagt  | aagccacaga | 1320 |
| [1280] | tgggaggaac  | cgaaggcca   | aaggacacaa | ggcctgttac  | tagcactcac | atggaacaaa | 1380 |
| [1281] | tggccaccgt  | gggaggatga  | caatttctgt | ggctgcgtga  | aagccttgag | gggctccggg | 1440 |
| [1282] | agctagagcc  | tctgctaacc  | atgttcatgc | cttcttcttt  | ttcctacagc | tcctgggcaa | 1500 |
| [1283] | cgtgctggtt  | attgtgctgt  | ctcatcat   | ttggcaaagaa | ttcctcgaag | atccgaaggg | 1560 |
| [1284] | aaagtcttcc  | acgactgtgg  | gatccgttcg | aagatatcac  | cggttgagcc | accatgagca | 1620 |
| [1285] | ccctgtgccc  | cccccccagc  | cccgccgtgg | ccaagaccga  | gatcgccctg | agcggcaaga | 1680 |
| [1286] | gccccctgct  | ggccgccacc  | ttgcctact  | gggacaacat  | cctgggcccc | cgcgtgcgcc | 1740 |
| [1287] | acatctgggc  | ccccaaagacc | gagcaggtgc | tgctgagcga  | cggcgagatc | accttctctg | 1800 |
| [1288] | ccaaccacac  | cctgaacggc  | gagatcctgc | gcaacgccga  | gagcggcgcc | atcgacgtga | 1860 |
| [1289] | agttcttcgt  | gctgagcgag  | aaggcgctga | tcacgtgag   | cctgatcttc | gacggcaact | 1920 |
| [1290] | ggaacggcga  | ccgcagcacc  | tacggcctga | gcatcatcct  | gccccagacc | gagctgagct | 1980 |
| [1291] | tctacctgcc  | cctgcaccgc  | gtgtgcgtgg | accgcctgac  | ccacatcatc | cgcaagggcc | 2040 |
| [1292] | gcatctggat  | gcacaaggag  | cgcaggaga  | acgtgcagaa  | gatcatcctg | gagggcaccg | 2100 |
| [1293] | agcgcattga  | ggaccagggc  | cagagcatca | tccccatgct  | gaccggcgag | gtgatccccg | 2160 |
| [1294] | tgatggagct  | gctgagcagc  | atgaagagcc | acagcgtgcc  | caggagatc  | gacatcgccg | 2220 |
| [1295] | acaccgtgct  | gaacgacgac  | gacatcggcg | acagctgcc   | caggggcttc | ctgctgaacg | 2280 |
| [1296] | ccatcagcag  | ccacctgcag  | acctgcggct | gcagcgtggt  | ggtgggcagc | agcgccgaga | 2340 |
| [1297] | aggtgaacaa  | gatcgtgcgc  | acctgtgcc  | tgttctgac   | ccccgccgag | cgcaagtga  | 2400 |
| [1298] | gccgcctgtg  | cgaggccgag  | agcagcttca | agtacgagag  | cggcctgttc | gtgcagggcc | 2460 |
| [1299] | tgctgaagga  | cagcaccggc  | agcttctgtc | tgcccttccg  | ccaggtgatg | tacgccccct | 2520 |
| [1300] | acccaccac   | ccacatcgac  | gtggacgtga | acaccgtgaa  | gcagatgccc | ccctgccacg | 2580 |
| [1301] | agcacatcta  | caaccagcgc  | cgctacatgc | gcagcgagct  | gaccgccttc | tggcgcgcca | 2640 |

|        |                                                                    |         |
|--------|--------------------------------------------------------------------|---------|
| [1302] | ccagcgagga ggacatggcc caggacacca tcatctacac cgacgagagc ttcacccccg  | 2700    |
| [1303] | acctgaacat cttccaggac gtgctgcacc gcgacaccct ggtgaaggcc ttcctggacc  | 2760    |
| [1304] | aggtgttcca gctgaagccc ggctgagcc tgcgcagcac cttcctggcc cagtctctgc   | 2820    |
| [1305] | tggtgctgca ccgcaaggcc ctgaccctga tcaagtacat cgaggacgac acccagaagg  | 2880    |
| [1306] | gcaagaagcc cttcaagagc ctgcgcaacc tgaagatcga cctggacctg accgccgagg  | 2940    |
| [1307] | gcgacctgaa catcatcatg gccctggccg agaagatcaa gcccggcctg cacagcttca  | 3000    |
| [1308] | tcttcggccg ccccttctac accagcgtgc aggagcgcga cgtgctgatg accttctaac  | 3060    |
| [1309] | aattgttaat taagttaaa ccctcgaggc cgcaagctta tcgataatca acctctggat   | 3120    |
| [1310] | tacaaaattt gtgaaagatt gactggtatt cttaactatg ttgctccttt tacgctatgt  | 3180    |
| [1311] | ggatacgtcg ctttaatgcc tttgtatcat gctattgctt cccgtatggc tttcattttc  | 3240    |
| [1312] | tcctccttgt ataaatcctg gttgctgtct ctttatgagg agttgtggcc cgttgtcagg  | 3300    |
| [1313] | caacgtggcg tggtgtgcac tgtgtttgct gacgcaaccc ccaactggtg gggcattgcc  | 3360    |
| [1314] | accacctgtc agctcctttc cgggaccttc gctttccccc tccctattgc cacggcgga   | 3420    |
| [1315] | ctcatcgccg cctgccttgc ccgctgctgg acaggggctc ggctgttggg cactgacaat  | 3480    |
| [1316] | tccgtggtgt tgctggggaa atcatcgctc tttccttggc tgctcgccg tggtgccacc   | 3540    |
| [1317] | tggattctgc gcgggacgtc cttctgtctac gtcccttcgg cctcaatcc agcggacctt  | 3600    |
| [1318] | ccttcccgcg gcctgctgcc ggctctgcgg cctcttccgc gtcttcgcct tcgccctcag  | 3660    |
| [1319] | acgagtcgga tctccctttg ggccgcctcc ccgcatcgat accgtcgact agagctcgct  | 3720    |
| [1320] | gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc tccccgtgc   | 3780    |
| [1321] | cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaattg   | 3840    |
| [1322] | catcgcatgt tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca  | 3900    |
| [1323] | agggggagga ttgggaagac aatagcaggc atgctgggga gagatccacg ataacaaaca  | 3960    |
| [1324] | gcttttttgg ggtgaacata ttgactgaat tccctgcagg ttggccactc cctctctgcg  | 4020    |
| [1325] | cgctcgctcg ctactgagg ccgcccgggc aaagcccggg cgtcggggcga cctttggtcg  | 4080    |
| [1326] | cccggcctca gtgagcgagc gagcgcgagc agaggagtg gccactcca tcaactaggg    | 4140    |
| [1327] | ttcct                                                              | 4145    |
| [1328] | <210>                                                              | 27      |
| [1329] | <211>                                                              | 899     |
| [1330] | <212>                                                              | DNA     |
| [1331] | <213>                                                              | 人工序列    |
| [1332] | <220>                                                              |         |
| [1333] | <223>                                                              | 合成的多核苷酸 |
| [1334] | <400>                                                              | 27      |
| [1335] | ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc   | 60      |
| [1336] | cgacccccgg gctttgcccg ggccgcctca gtgagcgagc gagcgcgagc agagggagtg  | 120     |
| [1337] | gccactcca tcaactagggg ttcctgctag ctctgggtat ttaagcccga gtgagcacgc  | 180     |
| [1338] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240     |
| [1339] | aaaaaaattg tcatctccc acggtggcca tttgttccat gtgagtgcta gtaacaggcc   | 300     |
| [1340] | ttgtgtcctt tgccctttcg gttcctccca tctgtggctt cactggagga accgaaaggc  | 360     |
| [1341] | caagacaaca gcatacagcc ttcagcaagc ctccagtggc ctacatacaga acttataaga | 420     |
| [1342] | ttcccaaatc caaagacatt tcacgtttat ggtgatttcc cagaacacat agcgacatgc  | 480     |
| [1343] | aaatattgca gggcgccact cccctgtccc tcacagccat cttcctgccg gggcgcacgc  | 540     |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [1344] | gcgctgggtg ttcccgcccta gtgacactgg gcccgcgatt ccttgagcgc ggttgatgac | 600  |
| [1345] | gtcagcggtt cccatgggtga agcttggatc tgatccctag gttctagaac cggtgaccaa | 660  |
| [1346] | ttgttaatta agtttaaac ctcgaggccg caagcagatc cagcataaca aacagctttt   | 720  |
| [1347] | ttggggtgaa catattgact gaattccctg caggttggtc actccctctc tgcgcgtcgc  | 780  |
| [1348] | ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg gtcgcccggc  | 840  |
| [1349] | ctcagtgagc gagcgagcgc gcagagagg agtggccaac tccatcacta ggggttcct    | 899  |
| [1350] | <210> 28                                                           |      |
| [1351] | <211> 2547                                                         |      |
| [1352] | <212> DNA                                                          |      |
| [1353] | <213> 人工序列                                                         |      |
| [1354] | <220>                                                              |      |
| [1355] | <223> 合成的多核苷酸                                                      |      |
| [1356] | <400> 28                                                           |      |
| [1357] | ttggccactc cctctctgcg cgctcgtcgc ctactgagg ccgcccgggc aaagcccggg   | 60   |
| [1358] | cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgag agagggagt    | 120  |
| [1359] | gccaaactca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc  | 180  |
| [1360] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240  |
| [1361] | tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag  | 300  |
| [1362] | ctagtattta atagtaatca attacggggg cattagtcca tagcccatat atggagttcc  | 360  |
| [1363] | gcgttacata acttacggtg aatggcccgc ctggctgacc gcccaacgac cccgcccac   | 420  |
| [1364] | tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc  | 480  |
| [1365] | aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagt tatcatatgc   | 540  |
| [1366] | caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgccag    | 600  |
| [1367] | acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgtatta   | 660  |
| [1368] | ccatggtega ggtgagcccc acgttctgct tcaactctcc catctcccc cctccccac    | 720  |
| [1369] | ccccaatttt gtatttattt attttttaat tattttgtgc agcgatgggg gcggggggg   | 780  |
| [1370] | ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga  | 840  |
| [1371] | gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttccttt atggcgaggc   | 900  |
| [1372] | ggcgcgggcg gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc  | 960  |
| [1373] | tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gccgccccg gctctgactg   | 1020 |
| [1374] | accgcgttac tcccacaggt gagcgggcgg gacggccctt ctctcagcg ctgtaattag   | 1080 |
| [1375] | cgcttgggtt aatgacggct tgttggaggc ttgctgaagg ctgtatgctg ttgtcttggc  | 1140 |
| [1376] | ctttcggttc ctccagtga gccacagatg ggaggaaccg aaaggccaaa ggacacaagg   | 1200 |
| [1377] | cctgttacta gcactcacat ggaacaaatg gccaccgtgg gaggatgaca atttctgtgg  | 1260 |
| [1378] | ctgcgtgaaa gccttgagg gctccgggag ctagagcctc tgctaaccat gttcatgcct   | 1320 |
| [1379] | tcttcttttt cctacagctc ctgggcaacg tgctggttat tgtgtgtct catcattttg   | 1380 |
| [1380] | gcaaagaatt cctcgaagat ccgaaggga agtccttcac gactgtggga tccgttcgaa   | 1440 |
| [1381] | gatatcacgc gttgagccac ccaattgtta attaatgta aaccctcgag gccgcaagct   | 1500 |
| [1382] | tatcgataat caacctctgg attacaaaat ttgtgaaaga ttgactggtg ttcttaacta  | 1560 |
| [1383] | tggtgtcct tttacgtat gtggatacgc tgctttaatg cttttgtatc atgctattgc    | 1620 |
| [1384] | ttcccgtagt gctttcattt tctcctcct gtataaatcc tgggtgtgt ctctttatga    | 1680 |
| [1385] | ggagttgtgg ccggtgtgca ggcaacgtgg cgtggtgtgc actgtgtttg ctgacgcaac  | 1740 |

|        |                                                                   |         |
|--------|-------------------------------------------------------------------|---------|
| [1386] | ccccactggt tggggcattg ccaccacctg tcagctcctt tccgggactt tcgctttccc | 1800    |
| [1387] | cctccctatt gccacggcgg aactcatcgc cgcctgcctt gcccgctgct ggacaggggc | 1860    |
| [1388] | tcggctgttg ggcactgaca attccgtggt gttgtcgggg aaatcatcgt cttttccttg | 1920    |
| [1389] | gctgctcgcc tgtgttgcca cctggattct gcgcgggacg tccttctgct acgtcccttc | 1980    |
| [1390] | ggccctcaat ccagcggacc ttcttcccgc cggcctgctg ccggctctgc ggcctcttc  | 2040    |
| [1391] | gcgtcttcgc cttcgccctc agacgagtcg gatctccctt tgggcgcct cccgcatcg   | 2100    |
| [1392] | ataccgtcga ctagagctcg ctgatcagcc tcgactgtgc cttctagttg ccagccatct | 2160    |
| [1393] | gttggttgcc cctccccctg gccttccttg accctggaag gtgccactcc cactgtcctt | 2220    |
| [1394] | tcctaataaa atgaggaaat tgcatcgcac tgtctgagta ggtgtcattc tattctgggg | 2280    |
| [1395] | ggtggggttg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg | 2340    |
| [1396] | gagagatcca cgataacaaa cagctttttt ggggtgaaca tattgactga attccctgca | 2400    |
| [1397] | ggttggccac tcctctctg cgcgtcgcct cgctcactga ggccgcccgg gcaaagcccg  | 2460    |
| [1398] | ggcgtcgggc gacctttggt cggccggcct cagtgagcga gcgagcgcgc agagaggag  | 2520    |
| [1399] | tggccaactc catcactagg ggttcct                                     | 2547    |
| [1400] | <210>                                                             | 29      |
| [1401] | <211>                                                             | 19      |
| [1402] | <212>                                                             | DNA     |
| [1403] | <213>                                                             | 人工序列    |
| [1404] | <220>                                                             |         |
| [1405] | <223>                                                             | 合成的多核苷酸 |
| [1406] | <400>                                                             | 29      |
| [1407] | tattagatct gatggccgc                                              | 19      |
| [1408] | <210>                                                             | 30      |
| [1409] | <211>                                                             | 20      |
| [1410] | <212>                                                             | DNA     |
| [1411] | <213>                                                             | 人工序列    |
| [1412] | <220>                                                             |         |
| [1413] | <223>                                                             | 合成的多核苷酸 |
| [1414] | <400>                                                             | 30      |
| [1415] | ctccatcact aggggttcct                                             | 20      |
| [1416] | <210>                                                             | 31      |
| [1417] | <211>                                                             | 60      |
| [1418] | <212>                                                             | DNA     |
| [1419] | <213>                                                             | 人工序列    |
| [1420] | <220>                                                             |         |
| [1421] | <223>                                                             | 合成的多核苷酸 |
| [1422] | <400>                                                             | 31      |
| [1423] | agctctgggt atttaagccc gagtgagcac gcagggtctc cattttgaag cgggaggtta | 60      |
| [1424] | <210>                                                             | 32      |
| [1425] | <211>                                                             | 145     |
| [1426] | <212>                                                             | DNA     |
| [1427] | <213>                                                             | 人工序列    |

- [1428] <220>  
[1429] <223> 合成的多核苷酸  
[1430] <400> 32  
[1431] aggaaccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg 60  
[1432] ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120  
[1433] gagcgcgcag agagggagtg gccaa 145  
[1434] <210> 33  
[1435] <211> 24  
[1436] <212> DNA  
[1437] <213> 人工序列  
[1438] <220>  
[1439] <223> 合成的多核苷酸  
[1440] <400> 33  
[1441] gtgtactagg atcttttact tgaa 24  
[1442] <210> 34  
[1443] <211> 24  
[1444] <212> DNA  
[1445] <213> 人工序列  
[1446] <220>  
[1447] <223> 合成的多核苷酸  
[1448] <400> 34  
[1449] ttcaagtaaa agatcctagt acac 24  
[1450] <210> 35  
[1451] <211> 267  
[1452] <212> DNA  
[1453] <213> 人工序列  
[1454] <220>  
[1455] <223> 合成的多核苷酸  
[1456] <400> 35  
[1457] gtgatatac aaggtcccag ggctggggtc agaaattctc tcccgaggga atgaagccac 60  
[1458] aggagccaag agcaggagga ccaaggcct ggcaaggcc gtggcctcgt tcaagtaaaa 120  
[1459] gatcctagta cagtcaggt cccaatgtgt actaggatct tttacttgaa cggggacgcc 180  
[1460] ggcatccggg ctgaggacc cctctctgc cagaggcacc aacaccagag ttcacaaatc 240  
[1461] agtctcctgc cttttgcatg tagcaaa 267  
[1462] <210> 36  
[1463] <211> 267  
[1464] <212> DNA  
[1465] <213> 人工序列  
[1466] <220>  
[1467] <223> 合成的多核苷酸  
[1468] <400> 36  
[1469] tttgtacat gcaaaggga ggagactgat ttgtgaactc tgggtgttgg gcctctggca 60

|        |                                                                        |
|--------|------------------------------------------------------------------------|
| [1470] | gagagggggg tcctgagccc ggatgccggc gtccccgttc aagtaaaaga tcctagtaca 120  |
| [1471] | cattgggacc tgcaactgtac taggatcttt tacttgaacg aggccacggc cttcgccagg 180 |
| [1472] | gccttggtec tcctgctctt ggctcctgtg gcttcattcc ctggggagag aatttctgac 240  |
| [1473] | cccagccctg ggacctgtg atatcac 267                                       |
| [1474] | <210> 37                                                               |
| [1475] | <211> 19                                                               |
| [1476] | <212> DNA                                                              |
| [1477] | <213> 人工序列                                                             |
| [1478] | <220>                                                                  |
| [1479] | <223> 合成的多核苷酸                                                          |
| [1480] | <400> 37                                                               |
| [1481] | atgcatccat attcttcct 19                                                |
| [1482] | <210> 38                                                               |
| [1483] | <211> 19                                                               |
| [1484] | <212> DNA                                                              |
| [1485] | <213> 人工序列                                                             |
| [1486] | <220>                                                                  |
| [1487] | <223> 合成的多核苷酸                                                          |
| [1488] | <400> 38                                                               |
| [1489] | aggaagaata tggatgcat 19                                                |
| [1490] | <210> 39                                                               |
| [1491] | <211> 148                                                              |
| [1492] | <212> DNA                                                              |
| [1493] | <213> 人工序列                                                             |
| [1494] | <220>                                                                  |
| [1495] | <223> 合成的多核苷酸                                                          |
| [1496] | <400> 39                                                               |
| [1497] | tggaggcttg ctgaaggctg tatgctgttg tcatgcatcc atattcttcc tagtgaagcc 60   |
| [1498] | acagatgagg aagaatatgg atgcatagga cacaaggcct gttactagca ctcacatgga 120  |
| [1499] | acaaatggcc accgtgggag gatgacaa 148                                     |
| [1500] | <210> 40                                                               |
| [1501] | <211> 19                                                               |
| [1502] | <212> DNA                                                              |
| [1503] | <213> 人工序列                                                             |
| [1504] | <220>                                                                  |
| [1505] | <223> 合成的多核苷酸                                                          |
| [1506] | <400> 40                                                               |
| [1507] | ttatttactt tctctgcac 19                                                |
| [1508] | <210> 41                                                               |
| [1509] | <211> 19                                                               |
| [1510] | <212> DNA                                                              |
| [1511] | <213> 人工序列                                                             |

- [1512] <220>  
[1513] <223> 合成的多核苷酸  
[1514] <400> 41  
[1515] gtgcagagaa agtaaataa 19  
[1516] <210> 42  
[1517] <211> 148  
[1518] <212> DNA  
[1519] <213> 人工序列  
[1520] <220>  
[1521] <223> 合成的多核苷酸  
[1522] <400> 42  
[1523] tggaggcttg ctgaaggctg tatgctgttg tcttatttac tttctctgca cagtgaagcc 60  
[1524] acagatggtg cagagaaagt aaataaagga cacaaggcct gttactagca ctacatgga 120  
[1525] acaaatggcc accgtgggag gatgacaa 148  
[1526] <210> 43  
[1527] <211> 151  
[1528] <212> DNA  
[1529] <213> 人工序列  
[1530] <220>  
[1531] <223> 合成的多核苷酸  
[1532] <400> 43  
[1533] tggaggcttg ctgaaggctg tatgctgttg tcggggccgg tttcggggcc ttagtgaagc 60  
[1534] cacagatgta cgccccgaa accggcccca ggacacaagg cctgttacta gcactcacat 120  
[1535] ggaacaaatg gccaccgtgg gaggatgaca a 151  
[1536] <210> 44  
[1537] <211> 151  
[1538] <212> DNA  
[1539] <213> 人工序列  
[1540] <220>  
[1541] <223> 合成的多核苷酸  
[1542] <400> 44  
[1543] tggaggcttg ctgaaggctg tatgctgttg tcggggccgg tttcggggcc ttagtgaagc 60  
[1544] cacagatgta cgccccgaa accggcccca ggacacaagg cctgttacta gcactcacat 120  
[1545] ggaacaaatg gccaccgtgg gaggatgaca a 151  
[1546] <210> 45  
[1547] <211> 296  
[1548] <212> DNA  
[1549] <213> 人工序列  
[1550] <220>  
[1551] <223> 合成的多核苷酸  
[1552] <400> 45  
[1553] ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctgataaagg gccttgtgtc 60

|        |                                                                        |
|--------|------------------------------------------------------------------------|
| [1554] | ctgtctggtt aatctttatc aggttcagtc agaggccaaa aacctgataa agattaacca 120  |
| [1555] | gacagcatatc agcccaaagc aagcctccag ttgtcatcct cccacggtgg ccatttggtc 180 |
| [1556] | catgtgagtg ctgataaagg gccttgtgtc ctgcggttgc ggtgcctgcg cctcagtcag 240  |
| [1557] | aggccaaaag gcgcaggcac cgcaaccgca gcatacagcc caaagcaagc ctccag 296      |
| [1558] | <210> 46                                                               |
| [1559] | <211> 296                                                              |
| [1560] | <212> DNA                                                              |
| [1561] | <213> 人工序列                                                             |
| [1562] | <220>                                                                  |
| [1563] | <223> 合成的多核苷酸                                                          |
| [1564] | <400> 46                                                               |
| [1565] | ctggaggett gctttgggct gtatgctgcg gttgcggtgc ctgcgccttt tggcctctga 60   |
| [1566] | ctgaggcgca ggcaccgcaa ccgcaggaca caaggccctt tatcagcact cacatggaac 120  |
| [1567] | aaatggccac cgtgggagga tgacaactgg aggcttgctt tgggctgtat gctgtctggt 180  |
| [1568] | taatctttat caggtttttg gcctctgact gaacctgata aagattaacc agacaggaca 240  |
| [1569] | caaggccctt tatcagcact cacatggaac aaatggccac cgtgggagga tgacaa 296      |
| [1570] | <210> 47                                                               |
| [1571] | <211> 297                                                              |
| [1572] | <212> DNA                                                              |
| [1573] | <213> 人工序列                                                             |
| [1574] | <220>                                                                  |
| [1575] | <223> 合成的多核苷酸                                                          |
| [1576] | <400> 47                                                               |
| [1577] | ttgtcatcct cccacggtgg ccatttggtc catgtgagtg ctgataaagg gccttgtgtc 60   |
| [1578] | ctgggttggt tccctccttg ttttcagtca gaggccaaaa taaacaagga gggaaacaac 120  |
| [1579] | ccagcatata gcccaaagca agcctccagt tgtcatcctc ccacggtgga catttggttc 180  |
| [1580] | atgtgagtg tgataaaggg ccttgtgtcc tggctgcggt tgcggtgcct gctcagtcag 240   |
| [1581] | aggccaaaag caggcaccgc aaccgcagcc agcatacagc ccaaagcaag cctccag 297     |
| [1582] | <210> 48                                                               |
| [1583] | <211> 297                                                              |
| [1584] | <212> DNA                                                              |
| [1585] | <213> 人工序列                                                             |
| [1586] | <220>                                                                  |
| [1587] | <223> 合成的多核苷酸                                                          |
| [1588] | <400> 48                                                               |
| [1589] | ctggaggctt gctttgggct gtatgctggc tgcggttgcg gtgcctgctt ttggcctctg 60   |
| [1590] | actgagcagg caccgcaacc gcagccagga cacaaggccc tttatcagca ctcacatgga 120  |
| [1591] | acaaatggcc accgtgggag gatgacaact ggaggettgc tttgggctgt atgctgggtt 180  |
| [1592] | gtttccctcc ttgtttattt tggcctctga ctgaaaacaa ggagggaac aaccaggac 240    |
| [1593] | acaaggccct ttatcagcac tcacatgga caaatggcca ccgtgggagg atgacaa 297      |
| [1594] | <210> 49                                                               |
| [1595] | <211> 294                                                              |

|        |                                                                        |
|--------|------------------------------------------------------------------------|
| [1596] | <212> DNA                                                              |
| [1597] | <213> 人工序列                                                             |
| [1598] | <220>                                                                  |
| [1599] | <223> 合成的多核苷酸                                                          |
| [1600] | <400> 49                                                               |
| [1601] | ttgtcatcct cccacggtgg ccatttgttc catgtgagtg ctgataaagg gccttgtgtc 60   |
| [1602] | ctgtttcttc tggttaatct ttatcagtca gaggccaaaa taaagattaa ccagaagaaa 120  |
| [1603] | cagcatacag cccaaagcaa gcctccagtt gtcacctcc cacggtggcc atttgttcca 180   |
| [1604] | tgtgagtgct gataaagggc cttgtgtcct gcggtgcctg cgcccgcggc tcagtcagag 240  |
| [1605] | gccaaaagcc gcgggcgcag gcaccgcagc atacagcca aagcaagcct ccag 294         |
| [1606] | <210> 50                                                               |
| [1607] | <211> 294                                                              |
| [1608] | <212> DNA                                                              |
| [1609] | <213> 人工序列                                                             |
| [1610] | <220>                                                                  |
| [1611] | <223> 合成的多核苷酸                                                          |
| [1612] | <400> 50                                                               |
| [1613] | ctggaggctt gctttgggct gtatgtctgcg gtgcctgcgc ccgcggttt tggcctctga 60   |
| [1614] | ctgagccgcg ggcgaggca ccgcaggaca caaggccctt tatcagcact cacatggaac 120   |
| [1615] | aaatggccac cgtgggagga tgacaactgg aggcttgcct tgggctgtat gctgtttctt 180  |
| [1616] | ctggttaatc tttattttgg cctctgactg ataaagatta accagaagaa acaggacaca 240  |
| [1617] | aggcccttta tcagcactca catggaacaa atggccaccg tgggaggatg acaa 294        |
| [1618] | <210> 51                                                               |
| [1619] | <211> 1446                                                             |
| [1620] | <212> DNA                                                              |
| [1621] | <213> 人工序列                                                             |
| [1622] | <220>                                                                  |
| [1623] | <223> 合成的多核苷酸                                                          |
| [1624] | <400> 51                                                               |
| [1625] | atgagcacc tgtgcccccc cccagcccc gccgtggcca agaccgagat cgccctgagc 60     |
| [1626] | ggcaagagcc ccctgtggc cgccaccttc gcctactggg acaacatcct gggccccgcg 120   |
| [1627] | gtgcgccaca tctgggcccc caagaccgag caggtgctgc tgagcgacgg cgagatcacc 180  |
| [1628] | ttcctggcca accacacct gaacggcgag atcctgcgca acgccgagag cggcgccatc 240   |
| [1629] | gacgtgaagt tcttcgtgct gagcgagaag ggcgtgatca tcgtgagcct gatcttcgac 300  |
| [1630] | ggcaactgga acggcgaccg cagcacctac ggcctgagca tcacctgcc ccagaccgag 360   |
| [1631] | ctgagcttct acctgcccct gcaccgcgtg tgcgtggacc gcctgaccca catcatccgc 420  |
| [1632] | aagggccgca tctggatgca caaggagcgc caggagaacg tgcagaagat catcctggag 480  |
| [1633] | ggcaccgagc gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgagggtg 540 |
| [1634] | atccccgtga tggagctgct gagcagcatg aagagccaca gcgtgcccga ggagatcgac 600  |
| [1635] | atgccgaca ccgtgctgaa cgacgacgac atcggcgaca gctgccacga gggcttcctg 660   |
| [1636] | ctgaacgcca tcagcagcca cctgcagacc tgcggctgca gcgtggtggt gggcagcagc 720  |
| [1637] | gccgagaagg tgaacaagat cgtgcgcacc ctgtgcctgt tcctgacccc cgccgagcgc 780  |

|        |                                                                     |         |
|--------|---------------------------------------------------------------------|---------|
| [1638] | aagtgcagcc gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttcgtg   | 840     |
| [1639] | cagggcctgc tgaaggacag caccggcagc ttctgtctgc cttccgcca ggtgatgtac    | 900     |
| [1640] | gccccctacc ccaccacca catcgacgtg gacgtgaaca ccgtgaagca gatgcccccc    | 960     |
| [1641] | tgccacgagc acatctacaa ccagcgccgc tacatgcgca gcgagctgac cgccttcttg   | 1020    |
| [1642] | cgcgccacca gcgaggagga catggcccag gacaccatca tctacaccga cgagagcttc   | 1080    |
| [1643] | acccccgacc tgaacatctt ccaggacgtg ctgcaccgcg acaccctggt gaaggccttc   | 1140    |
| [1644] | ctggaccagg tgttccagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag   | 1200    |
| [1645] | ttctgtctgg tgctgcaccg caaggccctg accctgatca agtacatcga ggacgacacc   | 1260    |
| [1646] | cagaagggca agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc   | 1320    |
| [1647] | gccgagggcg acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac   | 1380    |
| [1648] | agcttcatct tcggccgccc cttctacacc agcgtgcagg agcgcgacgt gctgatgacc   | 1440    |
| [1649] | ttctaa                                                              | 1446    |
| [1650] | <210>                                                               | 52      |
| [1651] | <211>                                                               | 3994    |
| [1652] | <212>                                                               | DNA     |
| [1653] | <213>                                                               | 人工序列    |
| [1654] | <220>                                                               |         |
| [1655] | <223>                                                               | 合成的多核苷酸 |
| [1656] | <400>                                                               | 52      |
| [1657] | ttggccactc cctctctgcg cgctcgtctg ctactgagg ccgggcgacc aaaggtcgcc    | 60      |
| [1658] | cgacgcccg gctttgccg ggcgccctca gtgagcgagc gagcgcgag agagggagt       | 120     |
| [1659] | gccaaactcca tctactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180     |
| [1660] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240     |
| [1661] | tcctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag    | 300     |
| [1662] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc   | 360     |
| [1663] | cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgcca    | 420     |
| [1664] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt    | 480     |
| [1665] | caatgggtgg actatttacg gtaaatgcc cacttggcag tacatcaagt gtatcatatg    | 540     |
| [1666] | ccaagtaagc cccctattga cgtcaatgac ggtaaattgc ccgctggca ttatgccag     | 600     |
| [1667] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt   | 660     |
| [1668] | accatggctg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccc    | 720     |
| [1669] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg   | 780     |
| [1670] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg   | 840     |
| [1671] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg   | 900     |
| [1672] | cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg   | 960     |
| [1673] | ctgccttcgc cccgtgcccc gtcgcccgcc gcctcgcgc cgcccgcgcc ggctctgact    | 1020    |
| [1674] | gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctcagc gctgtaatta   | 1080    |
| [1675] | gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcatgc    | 1140    |
| [1676] | atccatattc ttctagtga agccacagat gaggaagaat atggatgcat aggacacaag    | 1200    |
| [1677] | gcctgttact agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg   | 1260    |
| [1678] | gctgcgtgaa agccttgagg ggctccggga gctagagcct ctgctaacca tgttcatgcc   | 1320    |
| [1679] | ttcttctttt tcctacagct cctgggcaac gtgctggtta ttgtgctgtc tcatcatttt   | 1380    |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [1680] | ggcaaagaat tcctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga  | 1440 |
| [1681] | agatatcacc ggttgagcca ccatgagcac cctgtgcccc cccccagcc ccgccgtggc   | 1500 |
| [1682] | caagaccgag atcgccctga gcggcaagag cccctgtgtg gccgccacct tgcctactgt  | 1560 |
| [1683] | ggacaacatc ctgggcccc gcgtgcgcca catctgggcc cccaagaccg agcaggtgct   | 1620 |
| [1684] | gctgagcgac ggcgagatca ctttcttggc caaccacacc ctgaacggcg agatcctgcg  | 1680 |
| [1685] | caacgccgag agcggcgcca tcgacgtgaa gttcttctgt ctgagcgaga agggcgtgat  | 1740 |
| [1686] | catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagcacct acggcctgag  | 1800 |
| [1687] | catcatcctg cccagaccg agctgagctt ctacctgcc ctgcaccgcg tgtgctgga     | 1860 |
| [1688] | ccgcctgacc cacatcatcc gcaagggccg catctggatg cacaaggagc gccaggagaa  | 1920 |
| [1689] | cgtgcagaag atcatcctgg agggcaccga gcgcatggag gaccagggcc agagcatcat  | 1980 |
| [1690] | ccccatgtg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca   | 2040 |
| [1691] | cagcgtgcc gagagatcg acatcgccga caccgtgtg aacgacgac acatcgccga      | 2100 |
| [1692] | cagctgccac gagggcttcc tgctgaacgc catcagcagc cacctgcaga cctgcggctg  | 2160 |
| [1693] | cagcgtggtg gtggcgagca gcgccgagaa ggtgaacaag atcgtgcgca ccctgtgcct  | 2220 |
| [1694] | gttcttgacc cccgccgagc gcaagtgcag ccgcctgtgc gaggccgaga gcagcttcaa  | 2280 |
| [1695] | gtacgagagc ggctgttcg tgcaggcct gctgaaggac agcaccggca gttcgtgtct    | 2340 |
| [1696] | gcccttcgc caggtgatgt acgcccccta cccaccacc cacatcgac tggacgtgaa     | 2400 |
| [1697] | caccgtgaag cagatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg  | 2460 |
| [1698] | cagcgagctg accgccttct ggcgcgccac cagcgaggag gacatggccc aggacaccat  | 2520 |
| [1699] | catctacacc gacgagagct tcacccccga cctgaacatc ttccaggagc tgctgcaccg  | 2580 |
| [1700] | cgacacctg gtgaaggcct tcctggacca ggtgttccag ctgaagcccc gcctgagcct   | 2640 |
| [1701] | gcgcagcacc ttcttgcccc agttcttctg ggtgtgtcac cgcaaggccc tgaccctgat  | 2700 |
| [1702] | caagtacatc gaggacgaca ccagaaggg caagaagccc ttcaagagcc tgcgcaacct   | 2760 |
| [1703] | gaagatcgac ctggacctga ccgccgagg cgacctgaac atcatcatgg ccctggccga   | 2820 |
| [1704] | gaagatcaag cccggcctgc acagcttcat ctteggccgc cctttctaca ccagcgtgca  | 2880 |
| [1705] | ggagcgcgac gtgctgatga ctttctaaca attgttaatt aagtttaaac cctcgaggcc  | 2940 |
| [1706] | gcaagcttat cgataatcaa cctctggatt acaaaatttg tgaaagattg actggtattc  | 3000 |
| [1707] | ttaactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg   | 3060 |
| [1708] | ctattgttcc ccgtatggt ttcattttct cctccttgta taaatcctgg ttgtgtctc    | 3120 |
| [1709] | tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcaact gtgtttgctg | 3180 |
| [1710] | acgcaacccc cactggttgg ggcattgcc ccacctgtca gctcctttcc gggactttcg   | 3240 |
| [1711] | ctttccccct ccctattgcc acggcggaac tcatcgccgc ctgccttgcc cgctgctgga  | 3300 |
| [1712] | caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcggggaaa tcatcgtcct  | 3360 |
| [1713] | ttccttggt gctcgctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg    | 3420 |
| [1714] | tccttcggc cctcaatcca gcggaccttc cttcccggc cctgctgccg gctctgcggc    | 3480 |
| [1715] | ctcttcggc tcttcgctt cgccctcaga cgagtcggat ctccctttgg gccgcctccc    | 3540 |
| [1716] | cgcacgata ccgtcgacta gagctcgctg atcagcctcg actgtgcctt ctagttgcca   | 3600 |
| [1717] | gccatctgtt gtttggccct ccccggtgcc ttccctgacc ctggaagggt ccactccac   | 3660 |
| [1718] | tgctctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat   | 3720 |
| [1719] | tctggggggt ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca  | 3780 |
| [1720] | tgctggggag agatccacga taacaacag cttttttgg gtgaacatat tgactgaatt    | 3840 |
| [1721] | ccctgcaggt tggccactcc ctctctgcgc gctcgtcgc tcaactgagc cgcccgggca   | 3900 |

|        |                                                                      |         |
|--------|----------------------------------------------------------------------|---------|
| [1722] | aagccccgggc gtcgggcgac ctttggtcgc ccggcctcag tgagcgagcg agcgcgacaga  | 3960    |
| [1723] | gagggagtgg ccaactccat cactaggggt tcct                                | 3994    |
| [1724] | <210>                                                                | 53      |
| [1725] | <211>                                                                | 3997    |
| [1726] | <212>                                                                | DNA     |
| [1727] | <213>                                                                | 人工序列    |
| [1728] | <220>                                                                |         |
| [1729] | <223>                                                                | 合成的多核苷酸 |
| [1730] | <400>                                                                | 53      |
| [1731] | ttggccactc cctctctcgc cgctcgcctc ctcactgagg ccgggcgacc aaaggtcgcc    | 60      |
| [1732] | cgacccccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgagc agagggagtg    | 120     |
| [1733] | gccaaactcca tctactagggg ttcttgctag ctctgggtat ttaagccccga gtgagcacgc | 180     |
| [1734] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc    | 240     |
| [1735] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag    | 300     |
| [1736] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc    | 360     |
| [1737] | cgcgttacat aacttacggt aaatggcccc cctggctgac cgcccaacga cccccgcca     | 420     |
| [1738] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt     | 480     |
| [1739] | caatgggtgg actatttacg gtaaatgcc cacttggcag tacatcaagt gtatcatatg     | 540     |
| [1740] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgcctggca ttatgccag     | 600     |
| [1741] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt    | 660     |
| [1742] | accatggctg aggtgagccc caggttctgc ttcactctcc ccatctcccc cccctcccca    | 720     |
| [1743] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg    | 780     |
| [1744] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg    | 840     |
| [1745] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg    | 900     |
| [1746] | cggcggcggc ggcgcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg     | 960     |
| [1747] | ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcgcc ggctctgact    | 1020    |
| [1748] | gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctcagc gctgtaatta    | 1080    |
| [1749] | gcgcttggtt taatgacggc ttgttgaggg cttgctgaag gctgtatgct gttgtcgggg    | 1140    |
| [1750] | ccggtttcgg ggcttagtg aagccacaga tgtacggccc cgaaaccggc ccaggacac      | 1200    |
| [1751] | aaggcctgtt actagcactc acatggaaca aatggccacc gtgggaggat gacaatttct    | 1260    |
| [1752] | gtggctgcgt gaaagccttg aggggctccg ggagctagag cctctgctaa ccatgttcat    | 1320    |
| [1753] | gccttcttct ttttctaca gtcctgggc aacgtgctgg ttattgtgct gtctcatcat      | 1380    |
| [1754] | tttgcaaaag aattcctcga agatccgaag ggaaagtctt ccacgactgt gggatccgtt    | 1440    |
| [1755] | cgaagatata accggttgag ccacatgag caccctgtgc cccccccca gcccgccgt       | 1500    |
| [1756] | ggccaagacc gagatcgccc tgagcggcaa gagccccctg ctggccgcca ccttcgccta    | 1560    |
| [1757] | ctgggacaac atcctgggcc cccgcgtcgc ccacatctgg gcccccaaga ccgagcaggt    | 1620    |
| [1758] | gctgctgagc gacggcgaga tcaccttctt ggccaaccac accctgaacg gcgagatcct    | 1680    |
| [1759] | gcgcaacgcc gagagcggcg ccacgcagct gaagtcttct gtgctgagcg agaaggcgct    | 1740    |
| [1760] | gatcatctg agcctgatct tcgacggcaa ctggaacggc gaccgcagca cctacggcct     | 1800    |
| [1761] | gagcatcatc ctgccccaga ccgagctgag cttctacctg cccctgcacc gcgtgtgcgt    | 1860    |
| [1762] | ggaccgcctg acccacatca tccgcaaggg ccgcatctgg atgcacaagg agcgccagga    | 1920    |
| [1763] | gaacgtgcag aagatcatcc tggagggcac cgagcgcagt gaggaccagg gccagagcat    | 1980    |

|        |                                                                   |         |
|--------|-------------------------------------------------------------------|---------|
| [1764] | catcccatg ctgaccggcg aggtgatccc cgtgatggag ctgctgagca gcatgaagag  | 2040    |
| [1765] | ccacagcgtg cccgaggaga tcgacatcgc cgacaccgtg ctgaacgacg acgacatcgg | 2100    |
| [1766] | cgacagctgc cagcagggtt tctgtctgaa cgccatcagc agccacctgc agacctgcgg | 2160    |
| [1767] | ctgcagcgtg gtggtgggca gcagcgccga gaaggtgaac aagatcgtgc gcaccctgtg | 2220    |
| [1768] | cctgttcctg acccccgccg agcgcaagtg cagccgcctg tgcgaggccg agagcagctt | 2280    |
| [1769] | caagtacgag agcggcctgt tcgtgcaggc cctgctgaag gacagcaccg gcagcttcgt | 2340    |
| [1770] | gctgcccttc cgccagggtga tgtacgcccc ctacccacc accacatcg acgtggacgt  | 2400    |
| [1771] | gaacaccgtg aagcagatgc cccctgccg cgagcacatc tacaaccagc gccgtacat   | 2460    |
| [1772] | gcgcagcgag ctgaccgcct tctggcgcg caccagcgag gaggacatgg cccaggacac  | 2520    |
| [1773] | catcatctac accgacgaga gttcacccc cgacctgaac atcttcagg acgtgctgca   | 2580    |
| [1774] | ccgcgacacc ctggtgaagg ctttctgga ccagggttgc cagctgaagc ccggcctgag  | 2640    |
| [1775] | cctgcgcagc accttcctgg cccagttcct gctggtgctg caccgcaagg cctgacctt  | 2700    |
| [1776] | gatcaagtac atcgaggacg acaccagaa gggcaagaag cccttcaaga gcctgcgcaa  | 2760    |
| [1777] | cctgaagatc gacctggacc tgaccgccg gggcgacctg aacatcatca tggccctggc  | 2820    |
| [1778] | cgagaagatc aagcccgcc tgacacagct catcttcggc cgccccctt acaccagcgt   | 2880    |
| [1779] | gcaggagcgc gacgtgctga tgaccttcta acaattgtta attaatgta aacctcgag   | 2940    |
| [1780] | gccgcaagct tatcgataat caacctctgg attacaaaat ttgtgaaaga ttgactggta | 3000    |
| [1781] | ttcttaacta tgttgctcct tttagctat gtggatacgc tgctttaatg cttttgtatc  | 3060    |
| [1782] | atgctattgc ttccgctatg gctttcattt tctcctcct gtataaatcc tggttgctgt  | 3120    |
| [1783] | ctctttatga ggagttgtgg cccgttgta ggcaacgtgg cgtggtgtgc actgtgttg   | 3180    |
| [1784] | ctgacgcaac cccactggg tggggcattg ccaccacctg tcagctcctt tccgggactt  | 3240    |
| [1785] | tcgctttccc cctccctatt gccacggcgg aactcatcgc cgcctgcctt gcccgctgct | 3300    |
| [1786] | ggacaggggc tcggctgttg ggcactgaca attccgtggt gttgtcgggg aaatcatcgt | 3360    |
| [1787] | cctttccttg gctgctcgcc tgtgttgcca cctggattct gcgcgggacg tccttctgct | 3420    |
| [1788] | acgtcccttc ggccctcaat ccagcggacc ttcttcccg cggcctgctg ccggctctgc  | 3480    |
| [1789] | ggcctcttcc gcgtcttcgc ctteggcctc agacgagtcg gatctccctt tgggccgctt | 3540    |
| [1790] | ccccgcatcg ataccgtcga ctagagctcg ctgatcagcc tcgactgtgc cttctagtgt | 3600    |
| [1791] | ccagccatct gttgtttgcc cctccccctg gccttccttg accctggaag gtgccactcc | 3660    |
| [1792] | cactgtcctt tcttaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc | 3720    |
| [1793] | tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag acaatagcag | 3780    |
| [1794] | gcatgctggg gagagatcca cgataacaaa cagctttttt ggggtgaaca tattgactga | 3840    |
| [1795] | attccctgca ggttgccac tccctctctg cgcgctcgtc cgctcactga ggccgcccgg  | 3900    |
| [1796] | gcaaagcccg ggcgtcgggc gacctttggt cgccggcct cagtgagcga gcgagcgcgc  | 3960    |
| [1797] | agagagggag tggccaactc catcactagg ggttcct                          | 3997    |
| [1798] | <210>                                                             | 54      |
| [1799] | <211>                                                             | 3994    |
| [1800] | <212>                                                             | DNA     |
| [1801] | <213>                                                             | 人工序列    |
| [1802] | <220>                                                             |         |
| [1803] | <223>                                                             | 合成的多核苷酸 |
| [1804] | <400>                                                             | 54      |
| [1805] | ttggccactc cctctctgcg cgctcgtcgt ctcactgagg ccgggcgacc aaaggtcgcc | 60      |

|        |                                                                     |      |
|--------|---------------------------------------------------------------------|------|
| [1806] | cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agagggagtg    | 120  |
| [1807] | gccaaactcca tctactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| [1808] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240  |
| [1809] | tccctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag   | 300  |
| [1810] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc   | 360  |
| [1811] | cgcgttacat aacttacggg aaatggcccg cctggctgac cgcccaacga cccccccca    | 420  |
| [1812] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt    | 480  |
| [1813] | caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg  | 540  |
| [1814] | ccaagtacgc cccctattga cgtcaatgac ggtaaattggc ccgcctggca ttatgccag   | 600  |
| [1815] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt   | 660  |
| [1816] | accatggctg aggtgagccc caggttctgc ttactctcc ccatctcccc cccctcccca    | 720  |
| [1817] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg   | 780  |
| [1818] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgagcg     | 840  |
| [1819] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg   | 900  |
| [1820] | cggcgcgggc ggcgggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg  | 960  |
| [1821] | ctgccttcgc cccgtgcccc gctccgccgc cgctctgcgc cgcccgcccc ggctctgact   | 1020 |
| [1822] | gaccgcgtta ctcccacagg tgagcgggag ggacggccct tctctcagc gctgtaatta    | 1080 |
| [1823] | gcgcttggtt taatgacggc ttgttgagg cttgctgaag gctgtatgct gttgtcttat    | 1140 |
| [1824] | ttactttctc tgcacagtga agccacagat ggtgcagaga aagtaaataa aggacacaag   | 1200 |
| [1825] | gcctgttact agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg   | 1260 |
| [1826] | gctgcgtgaa agccttgagg ggctccggga gctagagcct ctgctaacca tgttcatgcc   | 1320 |
| [1827] | ttcttctttt tcctacagct cctgggcaac gtgctggta ttgtgctgtc tcattcattt    | 1380 |
| [1828] | ggcaagaat tcctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga    | 1440 |
| [1829] | agatatcacc ggttgagcca ccatgagcac cctgtgcccc ccccccagcc ccgccgtggc   | 1500 |
| [1830] | caagaccgag atcgccctga gcggcaagag cccctgctg gccgccacct tcgcctactg    | 1560 |
| [1831] | ggacaacatc ctgggcccc gcgtgcgcca catctgggcc cccaagaccg agcaggtgct    | 1620 |
| [1832] | gctgagcgac ggcgagatca ctttctggc caaccacacc ctgaacggcg agatcctgcg    | 1680 |
| [1833] | caacccgag agcggcgcca tcgacgtgaa gttcttctgt ctgagcgaga agggcgtgat    | 1740 |
| [1834] | catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagacct acggcctgag    | 1800 |
| [1835] | catcatcctg cccagaccg agctgagctt ctacctgcc ctgcaccgcg tgtgcgtgga     | 1860 |
| [1836] | ccgctgacc cacatcatcc gcaaggccg catctggatg cacaaggagc gccaggagaa     | 1920 |
| [1837] | cgtgcagaag atcatcctgg agggcaccga gcgcatggag gaccagggcc agagcatcat   | 1980 |
| [1838] | ccccatgctg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca   | 2040 |
| [1839] | cagcgtgccc gaggagatcg acatcgccga caccgtgctg aacgacgacg acatcgccga   | 2100 |
| [1840] | cagctgccac gagggcttcc tgctgaacgc catcagcagc cacctgcaga cctgcggctg   | 2160 |
| [1841] | cagcgtggtg gtgggcagca gcgccgagaa ggtgaacaag atcgtgcgca ccctgtgcct   | 2220 |
| [1842] | gttcttgacc cccgccgagc gcaagtgcag ccgcctgtgc gaggccgaga gcagcttcaa   | 2280 |
| [1843] | gtacgagagc ggctgttcg tgcaggccct gctgaaggac agcaccggca gcttcgtgct    | 2340 |
| [1844] | gcccttcgc caggtgatgt acgcccccta cccaccacc cacatcgacg tggacgtgaa     | 2400 |
| [1845] | caccgtgaag cagatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg   | 2460 |
| [1846] | cagcgagctg accgccttct ggcgcgccac cagcgaggag gacatggccc aggacacat    | 2520 |
| [1847] | catctacacc gacgagagct tcacccccga cctgaacatc ttccaggagc tgctgcaccg   | 2580 |

|        |                                                                   |         |
|--------|-------------------------------------------------------------------|---------|
| [1848] | cgacacctg gtgaaggcct tcctggacca ggtgttccag ctgaagccc gcctgagcct   | 2640    |
| [1849] | gcgagcacc ttcttgccc agttcttgc ggtgctgcac cgcaaggccc tgaccctgat    | 2700    |
| [1850] | caagtacatc gaggacgaca cccagaaggg caagaagccc ttcaagagcc tgcgcaacct | 2760    |
| [1851] | gaagatcgac ctggacctga ccgccgaggg cgacctgaac atcatcatgg ccctggccga | 2820    |
| [1852] | gaagatcaag cccggcctgc acagcttcat cttcgccgc cccttctaca ccagcgtgca  | 2880    |
| [1853] | ggagcgcgac gtgctgatga ctttctaaca attgttaatt aagtttaaac cctcgaggcc | 2940    |
| [1854] | gcaagcttat cgataatcaa cctctggatt acaaaatttg tgaaagattg actggtattc | 3000    |
| [1855] | ttactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg   | 3060    |
| [1856] | ctattgcttc ccgtatggct ttcatcttct cctccttgta taaatcctgg ttgctgtctc | 3120    |
| [1857] | tttatgagga gttgtgccc gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg  | 3180    |
| [1858] | acgcaacccc cactggttgg ggcattgcc cccactgtca gtcctttcc gggactttcg   | 3240    |
| [1859] | ctttccccc cctattgcc acggcggaac tcacgcccgc ctgccttgcc cgtgctgga    | 3300    |
| [1860] | caggggctcg gctgttggc actgacaatt ccgtggtgt gtcgggaaa tcacgtcct     | 3360    |
| [1861] | ttccttggt gctcgctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg   | 3420    |
| [1862] | tccttcggc cctcaatcca gcggacctt cttcccgcg cctgctgccg gctctgcggc    | 3480    |
| [1863] | ctcttcggc tcttcgctt cgccctcaga cgagtcggat ctccctttgg gccgctccc    | 3540    |
| [1864] | cgcacgata ccgtcgacta gagctcgtg atcagcctcg actgtgcct ctagttgcc     | 3600    |
| [1865] | gccatctgtt gtttgcctt ccccggtcc ttccttgacc ctggaagggt ccaactccac   | 3660    |
| [1866] | tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat  | 3720    |
| [1867] | tctgggggt ggggtgggc aggacagaa gggggaggat tgggaagaca atagcaggca    | 3780    |
| [1868] | tgctggggag agatccacga taacaacag cttttttgg gtgaacatat tgactgaatt   | 3840    |
| [1869] | ccctgcaggt tggccactcc ctctctgcgc gctcgtcgc tcaactgagg cgccgggca   | 3900    |
| [1870] | aagccgggc gtcggcgac ctttggtgc ccggcctcag tgagcgagcg agcgcgaga     | 3960    |
| [1871] | gagggagtg ccaactccat cactaggggt tcct                              | 3994    |
| [1872] | <210>                                                             | 55      |
| [1873] | <211>                                                             | 902     |
| [1874] | <212>                                                             | DNA     |
| [1875] | <213>                                                             | 人工序列    |
| [1876] | <220>                                                             |         |
| [1877] | <223>                                                             | 合成的多核苷酸 |
| [1878] | <400>                                                             | 55      |
| [1879] | ttggccactc cctctctgcg cgctcgtcg ctactgagg ccgggcgacc aaaggtcgcc   | 60      |
| [1880] | cgacgcccgg gctttgccc ggcgccctca gtgagcgagc gagcgcgag agaggagtg    | 120     |
| [1881] | gccaaactca tcaactaggg ttctgctag ctctgggtat ttaagcccga gtgagcacgc  | 180     |
| [1882] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc | 240     |
| [1883] | aaaaaaattg tcatctccc acggtggcca tttgttccat gtgagtgcta gtaacaggcc  | 300     |
| [1884] | ttgtgtctg gggccggtt cgggccgcta catctgtgac ttactaagg cccgaaacc     | 360     |
| [1885] | ggccccgaca acagcataca gccttcagca agcctccagt ggtctcatac agaacttata | 420     |
| [1886] | agattcccaa atccaaagac atttcacgtt tatggtgatt tcccagaaca catagcgaca | 480     |
| [1887] | tgcaaatatt gcaggcgcc actccctgt ccctcacagc catcttctg ccaggcgca     | 540     |
| [1888] | cgcgcgtgg gtgttcccgc ctagtgacac tgggcccgc attccttgga gcgggttgat   | 600     |
| [1889] | gacgtcagcg tttcccatgg tgaagcttg atctgatccc taggttctag aaccggtgac  | 660     |

|        |                                                                   |         |
|--------|-------------------------------------------------------------------|---------|
| [1890] | caattgttaa ttaagtttaa accctcgagg ccgcaagcag atccacgata acaaacagct | 720     |
| [1891] | tttttggggt gaacatattg actgaattcc ctgcaggttg gccactccct ctctgcgcgc | 780     |
| [1892] | tcgtcgcgtc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc | 840     |
| [1893] | ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc | 900     |
| [1894] | ct                                                                | 902     |
| [1895] | <210>                                                             | 56      |
| [1896] | <211>                                                             | 2550    |
| [1897] | <212>                                                             | DNA     |
| [1898] | <213>                                                             | 人工序列    |
| [1899] | <220>                                                             |         |
| [1900] | <223>                                                             | 合成的多核苷酸 |
| [1901] | <400>                                                             | 56      |
| [1902] | ttggccactc cctctctgcg cgctcgcgtc ctactgagg ccgcccgggc aaagcccggg  | 60      |
| [1903] | cgtcgggcga cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagt  | 120     |
| [1904] | gccaaactcca tactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180     |
| [1905] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc | 240     |
| [1906] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag | 300     |
| [1907] | ctagttatta atagtaatca attacggggg cattagttca tagcccatat atggagttcc | 360     |
| [1908] | gcgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat | 420     |
| [1909] | tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc | 480     |
| [1910] | aatgggtgga gtatttacgg taaactgcc acttggcagt acatcaagt tatcatatgc   | 540     |
| [1911] | caagtacgcc ccctattgac gtcaatgacg gtaaattggc cgcctggcat tatgccagc  | 600     |
| [1912] | acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgtatta  | 660     |
| [1913] | ccatggtcga ggtgagcccc acgttctgct tactctccc catctcccc ccctccccac   | 720     |
| [1914] | ccccaatttt gtatttattt attttttaat tattttgtgc agcgatgggg gcgggggggg | 780     |
| [1915] | ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg gcgaggcgga | 840     |
| [1916] | gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa gtttccttt atggcgaggc  | 900     |
| [1917] | ggcggcgggc gcggccctat aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc | 960     |
| [1918] | tgcttcgcc ccgtgccccg ctccgccgcc gctcgcgcc gccgccccg gctctgactg    | 1020    |
| [1919] | accgcgttac tcccacaggt gagcgggcgg gacggccctt ctctcagcg ctgtaattag  | 1080    |
| [1920] | cgcttggttt aatgacggct tggttgaggc ttgctgaagg ctgtatgctg ttgtcggggc | 1140    |
| [1921] | cggtttcggg gccttagtga agccacagat gtacggcccc gaaaccggcc ccaggacaca | 1200    |
| [1922] | aggcctgtta ctagactca catggaacaa atggccaccg tgggaggatg acaatttctg  | 1260    |
| [1923] | tggtcgcgtg aaagccttga ggggtccgg gagctagagc ctctgctaac catgttcatt  | 1320    |
| [1924] | ccttcttctt tttctacag ctctgggca acgtgctggt tattgtgctg tctcatcatt   | 1380    |
| [1925] | ttggcaaaga attcctcgaa gatccgaagg gaaagtctc cacgactgtg ggatccgttc  | 1440    |
| [1926] | gaagatatca ccggttgagc cacccaattg ttaattaagt ttaaaccctc gaggccgcaa | 1500    |
| [1927] | gcttatcgat aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa | 1560    |
| [1928] | ctatgttgc ctttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat  | 1620    |
| [1929] | tgcttcccgt atggctttca ttttctctc cttgtataaa tcctggttgc tgtctcttta  | 1680    |
| [1930] | tgaggagtgt tggcccgttg tcaggcaacg tggcgtggtg tgcactgtgt ttgctgacgc | 1740    |
| [1931] | aacccccact ggttggggca ttgccaccac ctgtcagctc ctttcggga ctttcgttt   | 1800    |

|        |                                                                    |         |
|--------|--------------------------------------------------------------------|---------|
| [1932] | ccccctccct attgccacgg cggaactcat cgccgcctgc cttgcccgt gctggacagg   | 1860    |
| [1933] | ggctcggctg ttgggcactg acaattccgt ggtgtgtcg gggaaatcat cgtcctttcc   | 1920    |
| [1934] | ttggctgctc gcctgtgttg ccacctggat tctgcgcggg acgtccttct gctacgtccc  | 1980    |
| [1935] | ttcggccctc aatccagcgg accttccttc ccgcggcctg ctgccggctc tgcggcctct  | 2040    |
| [1936] | tccgcgtctt cgccttcgcc ctacagacgag tcggatctcc ctttgggccg cctccccgca | 2100    |
| [1937] | tcgataccgt cgactagagc tcgctgatca gcctcgactg tgccttctag ttgccagcca  | 2160    |
| [1938] | tctgttgttt gccccctccc cgtgccttcc ttgacctg aggtgccac tcccactgtc     | 2220    |
| [1939] | ctttcctaataaaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg  | 2280    |
| [1940] | gggggtgggg tgggagcagga cagcaagggg gaggattggg aagacaatag caggcatgct | 2340    |
| [1941] | ggggagagat ccacgataac aaacagcttt tttgggtga acatattgac tgaattccct   | 2400    |
| [1942] | gcaggttggc cactccctct ctgcgcgtc gctcgtcac tgaggccgcc cgggcaaagc    | 2460    |
| [1943] | ccgggcgtcg ggcgacctt ggtcggcccg cctcagttag cgagcgagcg cgcagagagg   | 2520    |
| [1944] | gagtggccaa ctccatcact aggggttcct                                   | 2550    |
| [1945] | <210>                                                              | 57      |
| [1946] | <211>                                                              | 4142    |
| [1947] | <212>                                                              | DNA     |
| [1948] | <213>                                                              | 人工序列    |
| [1949] | <220>                                                              |         |
| [1950] | <223>                                                              | 合成的多核苷酸 |
| [1951] | <400>                                                              | 57      |
| [1952] | ttggccactc cctctctgcg cgctcgtcg ctactgagg ccgggcgacc aaaggtcgcc    | 60      |
| [1953] | cgacggccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgag agaggagtg    | 120     |
| [1954] | gccaaactcca tactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc  | 180     |
| [1955] | agggtctcca tttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240     |
| [1956] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac  | 300     |
| [1957] | cctagttatt aatagtaatc aattacggg tcattagttc atagcccata tatggagttc   | 360     |
| [1958] | cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga ccccccca    | 420     |
| [1959] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt   | 480     |
| [1960] | caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg | 540     |
| [1961] | ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgccag   | 600     |
| [1962] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt  | 660     |
| [1963] | accatggctg aggtgagccc cacgttctgc ttactctcc ccatctcccc cccctcccc    | 720     |
| [1964] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg  | 780     |
| [1965] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg  | 840     |
| [1966] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg  | 900     |
| [1967] | cggcgggcggc ggcgcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg  | 960     |
| [1968] | ctgccttcgc cccgtgcccc gctccggcgc cgcctcgcgc cgcccgcccc ggctctgact  | 1020    |
| [1969] | gaccgcgtta ctcccacagg tgagcgggag ggacggccct tctcctccgg gctgtaatta  | 1080    |
| [1970] | gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgcggttgcg  | 1140    |
| [1971] | gtgcctgcgc cttttggcct ctgactgagg cgcaggcacc gcaaccgcag gacacaaggc  | 1200    |
| [1972] | cctttatcag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa ctggaggctt  | 1260    |
| [1973] | gctttgggct gtatgctgtc tggttaatct ttatcaggtt tttggcctct gactgaacct  | 1320    |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
| [1974] | gataaagatt aaccagacag gacacaaggc cttttatcag cactcacatg gaacaaatgg | 1380 |
| [1975] | ccaccgtggg aggatgacaa tttctgtggc tgcgtgaaag ctttgagggg ctccgggagc | 1440 |
| [1976] | tagagcctct gctaaccatg ttcatgcctt cttcttttct ctacagctcc tgggcaacgt | 1500 |
| [1977] | gctggttatt gtgctgtctc atcatttttg caaagaattc ctcgaagatc cgaagggaaa | 1560 |
| [1978] | gtcttccacg actgtgggat ccgttcgaag atatcaccgg ttgagccacc atgagcacc  | 1620 |
| [1979] | tgtgcccccc cccagcccc gccgtggcca agaccgagat cgccctgagc ggcaagagcc  | 1680 |
| [1980] | ccctgctggc cgccaccttc gcctactggg acaacatcct gggccccgc gtgcgccaca  | 1740 |
| [1981] | tctgggcccc caagaccgag caggtgctgc tgagcgacgg cgagatcacc ttcttgacca | 1800 |
| [1982] | accacaccct gaacggcgag atcctgcgca acgccgagag cggcgccatc gacgtgaagt | 1860 |
| [1983] | tcttctgtct gagcgagaag ggcgtgatca tcgtgagcct gatcttcgac ggcaactgga | 1920 |
| [1984] | acggcgaccg cagcacctac ggcctgagca tcattctgcc ccagaccgag ctgagcttct | 1980 |
| [1985] | acctgccct gcaccgctg tgcgtggacc gctgacca catcatccgc aagggccgca     | 2040 |
| [1986] | tctgcatgca caaggagcg caggagaacg tgcagaagat catcctggag ggcaccgagc  | 2100 |
| [1987] | gcatggagga ccagggccag agcatcatcc ccatgctgac cggcgagggt atccccgtga | 2160 |
| [1988] | tggagctgct gagcagcatg aagagccaca gcgtgcccga ggagatcgac atcgccgaca | 2220 |
| [1989] | ccgtgctgaa cgacgacgac atcgccgaca gctgccacga gggcttctct ctgaacgcca | 2280 |
| [1990] | tcagcagcca cctgcagacc tgcggctgca gcgtggtggt gggcagcagc gccgagaagg | 2340 |
| [1991] | tgaacaagat cgtgcgcacc ctgtgcctgt tcctgacccc cgccgagcgc aagtgcagcc | 2400 |
| [1992] | gcctgtgcga ggccgagagc agcttcaagt acgagagcgg cctgttctgt cagggcctgc | 2460 |
| [1993] | tgaaggacag caccggcagc ttctgtctgc ctttccgcca ggtgatgtac gccccctacc | 2520 |
| [1994] | ccaccacca catcgacgtg gacgtgaaca ccgtgaagca gatgcccccc tgccacgagc  | 2580 |
| [1995] | acatctacaa ccagcgccgc tacatgcgca gcgagctgac cgccttcttg cgcgccacca | 2640 |
| [1996] | gcgaggagga catggcccag gacaccatca tctacaccga cgagagcttc acccccgacc | 2700 |
| [1997] | tgaacatctt ccaggacgtg ctgcaccgcg acacctggt gaaggccttc ctggaccagg  | 2760 |
| [1998] | tgttccagct gaagcccggc ctgagcctgc gcagcacctt cctggcccag ttctgtctgg | 2820 |
| [1999] | tgctgcaccg caaggccctg acctgatca agtacatcga ggacgacacc cagaagggca  | 2880 |
| [2000] | agaagccctt caagagcctg cgcaacctga agatcgacct ggacctgacc gccgaggcg  | 2940 |
| [2001] | acctgaacat catcatggcc ctggccgaga agatcaagcc cggcctgcac agcttcatct | 3000 |
| [2002] | tcggccgccc cttctacacc agcgtgcagg agcgcgacgt gctgatgacc ttctaacaat | 3060 |
| [2003] | tgtaatttaa gtttaaaccc tcgaggccgc aagcttatcg ataataacc tctggattac  | 3120 |
| [2004] | aaaatttgtg aaagattgac tggattctt aactatgttg ctccttttac gctatgtgga  | 3180 |
| [2005] | tacgtgctt taatgccttt gtatcatgct attgcttccc gtatggcttt cattttctcc  | 3240 |
| [2006] | tccttgata aatcctggtt gctgtctctt tatgaggagt tgtggcccggt tgtcaggcaa | 3300 |
| [2007] | cgtggcgtgg tgtgcaactgt gtttgtgac gcaaccccc ctggttgggg cattgccacc  | 3360 |
| [2008] | acctgtcagc tcctttccgg gactttcgt tccccctcc ctattgccac ggcggaactc   | 3420 |
| [2009] | atcgccgct gccttgcccg ctgctggaca ggggctcggc tggttggcac tgacaattcc  | 3480 |
| [2010] | gtggtgttgt cggggaaatc atcgtcctt ccttggtgc tcgcctgtgt tgccacctgg   | 3540 |
| [2011] | attctgcgag ggacgtcctt ctgctacgtc ccttcggccc tcaatccagc ggaccttct  | 3600 |
| [2012] | ccccgggcc tgctgccggc tctgcggcct cttccgctc ttgccttcg ccctcagacg    | 3660 |
| [2013] | agtcggatct ccctttgggc cgcctcccc catcgatacc gtcgactaga gctcgctgat  | 3720 |
| [2014] | cagcctcgac tgtgccttct agttgccagc catctgttgt ttgccccctc cccgtgcct  | 3780 |
| [2015] | ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat | 3840 |

|        |                                                                    |         |
|--------|--------------------------------------------------------------------|---------|
| [2016] | cgcatgtct gagtaggtgt cattctattc tgggggtgg ggtggggcag gacagcaagg    | 3900    |
| [2017] | gggaggattg ggaagacaat agcaggcatg ctggggagag atccacgata acaaacagct  | 3960    |
| [2018] | tttttgggt gaacatattg actgaattcc ctgcaggtg gccactccct ctctgcgcgc    | 4020    |
| [2019] | tcgctcgtc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct ttggtcgccc   | 4080    |
| [2020] | ggcctcagt agcgagcgag cgcgagaga gggagtggcc aactccatca ctaggggttc    | 4140    |
| [2021] | ct                                                                 | 4142    |
| [2022] | <210>                                                              | 58      |
| [2023] | <211>                                                              | 4143    |
| [2024] | <212>                                                              | DNA     |
| [2025] | <213>                                                              | 人工序列    |
| [2026] | <220>                                                              |         |
| [2027] | <223>                                                              | 合成的多核苷酸 |
| [2028] | <400>                                                              | 58      |
| [2029] | ttggccactc cctctctgcg cgctcgtcg ctactgagg ccgggcgacc aaaggtcgcc    | 60      |
| [2030] | cgacccccg gctttcccg ggcggcctca gtgagcgagc gagcgcgag agagggagt      | 120     |
| [2031] | gccaaactca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc  | 180     |
| [2032] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240     |
| [2033] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag  | 300     |
| [2034] | cctagtattt aatagtaatc aattacgggg tcattagttc atagcccata tatggagtgc  | 360     |
| [2035] | cggtttacat aacttacggg aaatggcccg cttggctgac cgcccaacga cccccgcca   | 420     |
| [2036] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt   | 480     |
| [2037] | caatgggtgg actatttacg gtaaaactgcc cacttggcag tacatcaagt gtatcatatg | 540     |
| [2038] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgcctggca ttatgcccg   | 600     |
| [2039] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt  | 660     |
| [2040] | accatggctg aggtgagccc caggttctgc ttactctcc ccctctccc cccctcccca    | 720     |
| [2041] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg  | 780     |
| [2042] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg  | 840     |
| [2043] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg  | 900     |
| [2044] | cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg  | 960     |
| [2045] | ctgccttgc cccgtgcccc gctccgccgc cgctcgcgc cgcccgcccc ggctctgact    | 1020    |
| [2046] | gaccgcgtta ctcccacagg tgagcggcg ggacggccct tctctccgg gctgtaatta    | 1080    |
| [2047] | gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tggctgcggt  | 1140    |
| [2048] | tgcggtgcct gcttttgcc tctgactgag caggcaccgc aaccgcagcc aggacacaag   | 1200    |
| [2049] | gccctttatc agcactcaca tggaacaaat ggccaccgtg ggaggatgac aactggaggc  | 1260    |
| [2050] | ttgctttggg ctgtatgctg ggttggttcc ctcttgttt attttggcct ctgactgaaa   | 1320    |
| [2051] | acaaggaggg aaacaacca ggacacaagg ccctttatca gcactcacat ggaacaaatg   | 1380    |
| [2052] | gccaccgtgg gaggatgaca atttctgtg ctgcgtgaaa gccttgaggg gctccgggag   | 1440    |
| [2053] | ctagagctc tgctaaccat gttcatgcct tcttctttt cctacagctc ctgggcaacg    | 1500    |
| [2054] | tgtgtgttat tgtgtgtct catcatttg gcaaagaatt cctcgaagat ccgaaggga     | 1560    |
| [2055] | agtcttcac gactgtggga tccgttcgaa gatatcaccg gttgagccac catgagcacc   | 1620    |
| [2056] | ctgtgcccc ccccagccc cgccgtggcc aagaccgaga tcgccctgag cggcaagagc    | 1680    |
| [2057] | ccctgctgg ccgccacctt cgctactgg gacaacatcc tgggccccg cgtgcgccac     | 1740    |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
| [2058] | atctgggccc ccaagaccga gcaggtgctg ctgagcgacg gcgagatcac cttcctggcc | 1800 |
| [2059] | aaccacaccc tgaacggcga gatcctgcgc aacgccgaga gcggcgccat cgacgtgaag | 1860 |
| [2060] | ttcttcgtgc tgagcgagaa gggcgtgac atcgtgagcc tgatcttcga cggcaactgg  | 1920 |
| [2061] | aacggcgacc gcagcaccta cggcctgagc atcatcctgc cccagaccga gctgagcttc | 1980 |
| [2062] | tacctgcccc tgcaccgcgt gtgcgtggac cgcctgaccc acatcatccg caagggccgc | 2040 |
| [2063] | atctggatgc acaaggagcg ccaggagaac gtgcagaaga tcatcctgga gggcaccgag | 2100 |
| [2064] | cgcatggagg accaggcca gagcatcatc cccatgctga cggcgaggt gatccccgtg   | 2160 |
| [2065] | atggagctgc tgagcagcat gaagagccac agcgtgcccg aggagatcga catcgccgac | 2220 |
| [2066] | accgtgctga acgacgacga catcgcgac agctgccacg agggcttcct gctgaacgcc  | 2280 |
| [2067] | atcagcagcc acctgcagac ctgcggctgc agcgtggtgg tgggcagcag cgccgagaag | 2340 |
| [2068] | gtgaacaaga tcgtgcgcac cctgtgcctg ttctgaccc ccgccgagcg caagtgcagc  | 2400 |
| [2069] | cgcctgtgcg aggccgagag cagcttcaag tacgagagcg gcctgttcgt gcagggcctg | 2460 |
| [2070] | ctgaaggaca gcaccggcag cttcgtgctg cccttcgcc aggtgatgta cggcccctac  | 2520 |
| [2071] | cccaccaccc acatcgacgt ggacgtgaac accgtgaagc agatgcccc ctgccacgag  | 2580 |
| [2072] | cacatctaca accagcgcg ctacatgcgc agcgagctga ccgccttctg gcgcgccacc  | 2640 |
| [2073] | agcgaggagg acatggccca ggacaccatc atctacaccg acgagagctt ccccccgac  | 2700 |
| [2074] | ctgaacatct tccaggacgt gctgcaccgc gacaccctgg tgaaggcctt cctggaccag | 2760 |
| [2075] | gtgttccagc tgaagcccgg cctgagcctg cgcagcacct tcctggccca gttcctgctg | 2820 |
| [2076] | gtgctgcacc gcaaggccct gaccctgac aagtacatcg aggacgacac ccagaagggc  | 2880 |
| [2077] | aagaagccct tcaagagcct gcgcaacctg aagatcgacc tggacctgac cgccgagggc | 2940 |
| [2078] | gacctgaaca tcatcatggc cctggccgag aagatcaagc ccggcctgca cagcttcac  | 3000 |
| [2079] | ttcgccgcc cttctacac cagcgtgcag gagcgcgacg tgctgatgac cttctaaca    | 3060 |
| [2080] | ttgttaatta agtttaaac ctcgagggcg caagcttacc gataatcaac ctctggatta  | 3120 |
| [2081] | caaaatttgt gaaagattga ctggtattct taactatgtt gtcctttta cgctatgtgg  | 3180 |
| [2082] | atacgtgct ttaatgcctt tgtatcatgc tattgcttcc cgtatggctt tcatcttctc  | 3240 |
| [2083] | ctccttgat aaatcctggt tgctgtctct ttatgaggag ttgtggcccg ttgtcaggca  | 3300 |
| [2084] | acgtggcgtg gtgtgcactg tgtttgctga cgcaaccccc actggttggg gcattgccac | 3360 |
| [2085] | cacctgtcag ctctttccg ggactttcgc tttccccctc cctattgcca cggcggaact  | 3420 |
| [2086] | catcgccgcc tgcttgccc gctgctggac aggggctcgg ctgttgggca ctgacaattc  | 3480 |
| [2087] | cgtggtgttg tcggggaaat catcgtcctt tccttggtg ctcgcctgtg ttgccacctg  | 3540 |
| [2088] | gattctgcgc gggacgtcct tctgtacgt cccttcggcc ctcaatccag cggaccttc   | 3600 |
| [2089] | ttcccgcggc ctgctgccg ctctgcggcc tcttcgcgt cttgccttc gccctcagac    | 3660 |
| [2090] | gagtcggatc tcccttggg ccgcctcccc gcatcgatac cgtcgactag agctcgtga   | 3720 |
| [2091] | tcagcctega ctgtgccttc tagttgccag ccactgtgtg ttgccccctc ccccgctcct | 3780 |
| [2092] | tccttgaccc tggaagggtg cactcccact gtcctttcct aataaaatga ggaaattgca | 3840 |
| [2093] | tcgcattgtc tgagtaggtg tcattctatt ctggggggtg ggggtgggca ggacagcaag | 3900 |
| [2094] | ggggaggatt gggaagacaa tagcaggcat gctggggaga gatccacgat aacaaacagc | 3960 |
| [2095] | tttttgggg tgaacatatt gactgaattc cctgcaggtt ggccactccc tctctgcgcg  | 4020 |
| [2096] | ctcgtcgtc cactgaggcc gcccgggcaa agcccgggcg tcgggcgacc ttgtgtgcc   | 4080 |
| [2097] | cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc actaggggtt | 4140 |
| [2098] | cct                                                               | 4143 |
| [2099] | <210>                                                             | 59   |

|        |                                                                     |      |
|--------|---------------------------------------------------------------------|------|
| [2100] | <211> 4140                                                          |      |
| [2101] | <212> DNA                                                           |      |
| [2102] | <213> 人工序列                                                          |      |
| [2103] | <220>                                                               |      |
| [2104] | <223> 合成的多核苷酸                                                       |      |
| [2105] | <400> 59                                                            |      |
| [2106] | ttggccactc cctctctcg cgctcgcctc ctactgagg cgggcgacc aaaggtcgcc      | 60   |
| [2107] | cgacgcccgg gctttgccg ggcgccctca gtgagcgagc gagcgcgag agagggagt      | 120  |
| [2108] | gccaaactcca tctactagggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| [2109] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc   | 240  |
| [2110] | tcctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac    | 300  |
| [2111] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagtgc   | 360  |
| [2112] | cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccccca    | 420  |
| [2113] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt    | 480  |
| [2114] | caatgggtgg actatttacg gtaaacgcc cacttggcag tacatcaagt gtatcatatg    | 540  |
| [2115] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgctggca ttatgccag     | 600  |
| [2116] | tacatgacct tatgggactt tctacttgg cagtacatct acgtattagt catcgctatt    | 660  |
| [2117] | accatggctg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctccca    | 720  |
| [2118] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcgggggg    | 780  |
| [2119] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggagg ggcgaggcgg   | 840  |
| [2120] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg    | 900  |
| [2121] | cggcggcggc ggcgcccta taaaaagcga agcgcgcggc gggcgaggat cgctgcgacg    | 960  |
| [2122] | ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact   | 1020 |
| [2123] | gaccgcgtta ctcccacagg tgagcgggag ggacggccct tctcctccgg gctgtaatta   | 1080 |
| [2124] | gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgcggtgcct   | 1140 |
| [2125] | gcgcccgcgg cttttggcct ctgactgagc cgcgggcgca ggcaccgcag gacacaaggc   | 1200 |
| [2126] | cctttatcag cactcacatg gaacaaatgg ccaccgtggg aggatgacaa ctggaggctt   | 1260 |
| [2127] | gctttgggct gtatgtgtt tcttctggtt aatctttatt ttggcctctg actgataaag    | 1320 |
| [2128] | attaaccaga agaaacagga cacaagcccc tttatcagca ctacatgga acaaatggcc    | 1380 |
| [2129] | accgtgggag gatgacaatt tctgtggtg cgtgaaagcc ttgaggggct ccgggagcta    | 1440 |
| [2130] | gagcctctgc taaccatgtt catgccttct tcttttctc acagctcctg ggcaacgtgc    | 1500 |
| [2131] | tggttattgt gctgtctcat cattttggca aagaattcct cgaagatccg aagggaagt    | 1560 |
| [2132] | cttccacgac tgtgggatcc gttcgaagat atcaccggtt gagccaccat gagcacctg    | 1620 |
| [2133] | tgcccccccc ccagccccgc cgtggccaag accgagatcg cctgagcgg caagagcccc    | 1680 |
| [2134] | ctgctggccg ccacctcgc ctactgggac aacatcctgg gccccgcgt gcgccacatc     | 1740 |
| [2135] | tgggccccca agaccgagca ggtgctgctg agcgacggcg agatcacctt cctggccaac   | 1800 |
| [2136] | cacaccctga acggcgagat cctgcgcaac gccgagagcg gcgccatcga cgtgaagtgc   | 1860 |
| [2137] | ttcgtgctga gcgagaaggc cgtgatcatc gtgagcctga tcttcgacgg caactggaac   | 1920 |
| [2138] | ggcgaccgca gcacctacgg cctgagcatc atcctgcccc agaccgagct gagcttctac   | 1980 |
| [2139] | ctgcccctgc accgcgtgtg cgtggaccgc ctgaccaca tcatccgcaa gggccgcatc    | 2040 |
| [2140] | tggatgcaca aggagcgcca ggagaactg cagaagatca tcctggaggg caccgagcgc    | 2100 |
| [2141] | atggaggacc agggccagag catcatcccc atgctgaccg gcgaggtgat ccccgatg     | 2160 |

|        |                                                                     |      |
|--------|---------------------------------------------------------------------|------|
| [2142] | gagctgctga gcagcatgaa gagccacagc gtgcccagg agatcgacat cgccgacacc    | 2220 |
| [2143] | gtgctgaacg acgacgacat cggcgacagc tgccacgagg gcttctctgct gaacgccatc  | 2280 |
| [2144] | agcagccacc tgcagacctg cggctgcagc gtgggtggtgg gcagcagcgc cgagaagggtg | 2340 |
| [2145] | aacaagatcg tgcgcaccct gtgcctgttc ctgacccccg ccgagcgcaa gtgcagccgc   | 2400 |
| [2146] | ctgtgcgagg ccgagagcag cttcaagtac gagagcggcc tgttcgtgca gggcctgctg   | 2460 |
| [2147] | aaggacagca ccggcagctt cgtgctgccc ttccgccagg tgatgtacgc cccctacccc   | 2520 |
| [2148] | accaccacaca tcgacgtgga cgtgaacacc gtgaagcaga tgccccctg ccacgagcac   | 2580 |
| [2149] | atctacaacc agcgccgcta catgcgcagc gagctgaccg ctttctggcg cgccaccagc   | 2640 |
| [2150] | gaggaggaca tggcccagga caccatcatc tacaccgacg agagcttcac ccccgacctg   | 2700 |
| [2151] | aacatcttcc aggacgtgct gcaccgcgac accctggtga aggccttcct ggaccaggtg   | 2760 |
| [2152] | ttccagctga agcccggcct gagcctgcgc agcaccttcc tggcccagtt cctgctggtg   | 2820 |
| [2153] | ctgcaccgca aggcctgac cctgatcaag tacatcgagg acgacacca gaagggaag      | 2880 |
| [2154] | aagcccttca agagcctgcg caacctgaag atcgacctgg acctgaccgc cgagggcgac   | 2940 |
| [2155] | ctgaacatca tcatggccct ggccgagaag atcaagccg gcctgcacag cttcatcttc    | 3000 |
| [2156] | ggccggccct tctacaccag cgtgcaggag cgcgacgtgc tgatgacctt ctaacaattg   | 3060 |
| [2157] | ttaattaagt ttaaaccctc gagggcgcaa gcttatcgat aatcaaccctc tggattacaa  | 3120 |
| [2158] | aatttgtgaa agattgactg gtattcttaa ctatgttget ccttttacgc tatgtggata   | 3180 |
| [2159] | cgctgcttta atgcctttgt atcatgctat tgcttcccgt atggctttca ttttctcctc   | 3240 |
| [2160] | cttgataaaa tcttggttgc tgtctcttta tgaggagtgt tggcccgttg tcaggcaacg   | 3300 |
| [2161] | tggcgtggtg tgcactgtgt ttgctgacgc aacccccact ggttggggca ttgccaccac   | 3360 |
| [2162] | ctgtcagctc ctttccggga ctttcgcttt cccctccct attgccacgg cggaactcat    | 3420 |
| [2163] | cgccgcctgc cttgcccgt gctggacagg ggctcggctg ttgggcactg acaattccgt    | 3480 |
| [2164] | ggtgttgtcg gggaaatcat cgtcctttcc ttggctgctc gcctgtgttg ccacctggat   | 3540 |
| [2165] | tctgcgcggg acgtccttct gctacgtccc ttcgccctc aatccagcgg accttccttc    | 3600 |
| [2166] | ccgcggcctg ctgccgctc tgcggcctct tccgcgtctt cgccttcgcc ctcagacgag    | 3660 |
| [2167] | tcgcatctcc ctttgggcg cctccccgca tcgataccgt cgactagagc tcgctgatca    | 3720 |
| [2168] | gcctcgactg tgccttctag ttgccagcca tctgttgttt gccctcccc cgtgccttcc    | 3780 |
| [2169] | ttgaccttg aagtgccac tcccactgtc ctttccctaat aaaatgagga aattgcatcg    | 3840 |
| [2170] | cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg   | 3900 |
| [2171] | gaggattggg aagacaatag caggcatgct ggggagagat ccacgataac aaacagcttt   | 3960 |
| [2172] | tttggggtga acatattgac tgaattccct gcaggttggc cactccctct ctgcgcgtc    | 4020 |
| [2173] | gctcgtcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgacctt ggtcgcccgg     | 4080 |
| [2174] | cctcagtgag cgagcgagcg cgcagagagg gaggggccaa ctccatcact aggggttcc    | 4140 |
| [2175] | <210> 60                                                            |      |
| [2176] | <211> 3793                                                          |      |
| [2177] | <212> DNA                                                           |      |
| [2178] | <213> 人工序列                                                          |      |
| [2179] | <220>                                                               |      |
| [2180] | <223> 合成的多核苷酸                                                       |      |
| [2181] | <400> 60                                                            |      |
| [2182] | tctcattaga agtgaggcgg ggccggccaa atcgaatgga caccgggtaa ttagcagggt   | 60   |
| [2183] | taccagata ctccagcacc ttttcccgt cggcgtgta cctgccattc acgtccatgc      | 120  |

|        |                                                                     |      |
|--------|---------------------------------------------------------------------|------|
| [2184] | cattgatggc cagcactgca tgaccactg cagaggtgaa gctaacggtc agcgaagggtg   | 180  |
| [2185] | cagccccggg attccgccga ggggacaagg gacccgacac aacccctttt cccccaaccc   | 240  |
| [2186] | cgcacctaca accagcccac ttctacagca ctggggccct cccacccccg caccgccac    | 300  |
| [2187] | gggcccagac ctagcccacc tcggatgccg tcccgtggc cgaaagcaac caacacacgc    | 360  |
| [2188] | tcatcgtgta gcttgagcag cagatccagc ggataactga aagttttctc agcctcagcc   | 420  |
| [2189] | cgtggcgcggt agctgtccaa ctggtaaattc aagccgccag ctttgttcac cacatacaca | 480  |
| [2190] | ctaaaaatcg ccatcgctgc cttgccgtc ggaaactggg attcagcctc taccgcagcg    | 540  |
| [2191] | ccccctcccg gaaccgcac acagcacttg ccgccggccc caccacagcc tctctctcct    | 600  |
| [2192] | cctcctctc ctcccgcgc ccccgctcag ccacctgctg cacttgcgca ctgggagcga     | 660  |
| [2193] | cacgctcggg cataagtagt gccgaaaagt tagctgccga gacctgggtg attgcttttc   | 720  |
| [2194] | gtttatcagt gcagggaaaac agcgtatag tactgcgtca caactagcgc agactccggc   | 780  |
| [2195] | agtatttagg cgggtcggtt tgggaactag aatccacttc ctgtcttccg cctcaggcta   | 840  |
| [2196] | gagggcgagc gcttcgccgt gggacttctt ctgcctggct ccgcctcttg ccccggaagt   | 900  |
| [2197] | actcacagcg gacggtgggt tttggggccg tttctgagca gcgcttcctt tttgtccgac   | 960  |
| [2198] | atcttgacga ggctgcggtg tctgctgcta ttctccgagc ttcgcaatgg taagcttcag   | 1020 |
| [2199] | gggtgtgaag tcgccggcgt tcttgggttt gaggactcag tggggagagc cttcggcggg   | 1080 |
| [2200] | agcgtcctt ggctgccgg cctcggttgc agggcgggcg cggttattgc ttggcccatg     | 1140 |
| [2201] | tgctctggtg gtggagtttg cgggggctga gggcgagta ttaggggact ttggcgctat    | 1200 |
| [2202] | ttgaggacct ggttgcatc ccgctgccct cctacagccg cctaaggacg acaagaagaa    | 1260 |
| [2203] | gaaggacgct ggaaagtcgg ccaagaaaaga caaagaccca gtgaacaaat ccgggggcaa  | 1320 |
| [2204] | ggccaaaaag aaggtagaaa taagacctct ctgaaagaga ctaggggtaa ctctctcgta   | 1380 |
| [2205] | atcctctagt aataggtaac ttgtatagta agtggttttt cagggtgtaga tttctagagt  | 1440 |
| [2206] | caaaatgtga gagtttatct tcccgctacc actcgttctt tttccatta ggatcatgaa    | 1500 |
| [2207] | aatgggtctg ttgtgcgaag tgtctgccgc tgtgcctgct gtgttatttt taactgatct   | 1560 |
| [2208] | agtggggctc ggccctgtt tgaaggccaa aaacgtgtcg gtgttttttt tttgtttttg    | 1620 |
| [2209] | ttttagtaat gtgtaattta tcttgataa cgggtgaaca gatttctctg acgcagatta    | 1680 |
| [2210] | ctcgagaggg aaagggtgct tctgccagaa atactaactt gtttctgttt tgttttggtg   | 1740 |
| [2211] | agcagaagtg gtccaaaggc aaagttcggg acaagctcaa taacttagtc ttgtttgaca   | 1800 |
| [2212] | aagctaccta tgataaactc tgtaaggaag ttcccaacta taaacttata accccagctg   | 1860 |
| [2213] | tggtctctga gagactgaag attcgaggt cctggccag ggcagccctt caggagctcc     | 1920 |
| [2214] | ttagtaaagg tgaggggtgt atcctacatg tgtgtttttg taggttaaatt tgtcttgacc  | 1980 |
| [2215] | atgttaagca tcttcagtgg ttttgcgtga aaagcagaat taaaaaaaaa aagcgtggct   | 2040 |
| [2216] | tgaccatttg ctgttagtaa tgtaattctg acgtcttact cctgatcctg agatgaattc   | 2100 |
| [2217] | tcagggttct tagccacttt tgtgccgtgg accctgtggc agtttagtga agcccaagga   | 2160 |
| [2218] | tcttttatgt ttcgagtaaa tggatgcata gaattacagg gacaaccgtt tttgaaataa   | 2220 |
| [2219] | ttagattact attttgaaac aactttgaaa atgtttaaaa cttttatggt aaatatattg   | 2280 |
| [2220] | ttgatgtatt aaattttaaa accagaaatt tagtacggtc tactcagtag tatggtctga   | 2340 |
| [2221] | ttaccataat tccacaataa taaggctcag ctaactatag tgactgaacg tctataattc   | 2400 |
| [2222] | tagcactttg ggaggccaag gcgggtgaat caacggaggt caggagttaa agaccagcct   | 2460 |
| [2223] | ggccaatatg gtgaaaacct gctctactga aagttagctg gacgtggggg cacacgtctg   | 2520 |
| [2224] | taatcccagc tactcaggat gctgagggat gaggatccct tgaaccagc agatggaggt    | 2580 |
| [2225] | ggcagtgagc cgagatgaca ccaactgcact ccagccttag tgacagcaaa agactgtctc  | 2640 |

|        |                                                                    |         |
|--------|--------------------------------------------------------------------|---------|
| [2226] | agaaaggggg ggggggtgga agataatgga gccctaattt aaaggaaaag taaggataga  | 2700    |
| [2227] | tgatccgtta aaaacttgga ttctcggtta ccgaacgtca gattaagcaa ttctggagcc  | 2760    |
| [2228] | aggtgcagtg gtacccttgt atttctagct acttgggagg ccaaagcagg aggatcattt  | 2820    |
| [2229] | gagccaagga gttttaagac cattctgggc acctctgaga gaactctgtc tttttgtttt  | 2880    |
| [2230] | ccttttcttt aaatagagat gcggttttgc catgttgccc aggctggtct cctgggctca  | 2940    |
| [2231] | agagatccac ctgtccaaag tgctgggatt acaggcatga gcctctgcac ccggccaaaa  | 3000    |
| [2232] | caaacccttac tagagtctca ttctgttgcc caggttggag tgcgaggagg cagtcttggc | 3060    |
| [2233] | tcaatgcaac caccaattcc tgggttcagg tggtcctcac ctacagctcc caagtagctg  | 3120    |
| [2234] | gaattacaag catgtgccac catgcccagc taatttttgt attttttgta gagatggggt  | 3180    |
| [2235] | ttcaccttgt tggccaggct ggtgtgcaac tccttacctc aagctatctg cccgtctcca  | 3240    |
| [2236] | cctcccaaag cagtgggatt ataagcatga gccaccgcgc ccagccaaaa accttactag  | 3300    |
| [2237] | tttctattgt agcatctgtt aagcatctca tcgtgctatt ctctccccct aggacttatt  | 3360    |
| [2238] | aaactggttt caaagcacag agctcaagta atttacacca gaaataccaa gggtaggat   | 3420    |
| [2239] | gctccagctg ctggtgaaga tgcatgaata ggtgagtagg aatgtgtggg ctcatggtgt  | 3480    |
| [2240] | aggaggtaga taaaaagctt tatggttctg attcttttaa tttttttta caggtccaac   | 3540    |
| [2241] | cagctgtaca tttgaaaaa taaacttta ttaaatcaaa tgaatgagta tgtctgtttc    | 3600    |
| [2242] | ctaagaaaga caatgataaa gaatttggtg gaaggataaa taggggtttg ttgactttgc  | 3660    |
| [2243] | ttttagctc atggtagtgt gtagagagca tgattagctt ttttctgtat gtgactgctt   | 3720    |
| [2244] | cttcattgct gcagcttcag ttttgaattg atgtctgaaa ggaaataaag ggttaacacg  | 3780    |
| [2245] | atgatgaagg gtg                                                     | 3793    |
| [2246] | <210>                                                              | 61      |
| [2247] | <211>                                                              | 3996    |
| [2248] | <212>                                                              | DNA     |
| [2249] | <213>                                                              | 人工序列    |
| [2250] | <220>                                                              |         |
| [2251] | <223>                                                              | 合成的多核苷酸 |
| [2252] | <400>                                                              | 61      |
| [2253] | ttggccactc cctctctgcg cgctcgcctc ctactgagg ccgggcgacc aaaggtcgcc   | 60      |
| [2254] | cgacgcccg gctttgccg ggcgccctca gtgagcgagc gagcgcgag agagggagt      | 120     |
| [2255] | gccaaactcca tctactaggg ttctgtctag ctctgggtat ttaagccga gtgagcacgc  | 180     |
| [2256] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240     |
| [2257] | tcctaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtac   | 300     |
| [2258] | cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc  | 360     |
| [2259] | cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccccca   | 420     |
| [2260] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt   | 480     |
| [2261] | caatgggtgg actatttacg gtaaacgcc cacttggcag tacatcaagt gtatcatatg   | 540     |
| [2262] | ccaagtacgc cccctattga cgtcaatgac ggtaaattgc ccgcctggca ttatgccag   | 600     |
| [2263] | tacatgacct tatgggactt tctacttg cagtacatct acgtattagt categtatt     | 660     |
| [2264] | accatggctg aggtgagccc caggttctgc ttactctcc ccatctcccc cccctccca    | 720     |
| [2265] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg  | 780     |
| [2266] | gggggggggc gcgcgccagg cggggcgagg cggggcgagg ggcggggcgg ggcgaggcgg  | 840     |
| [2267] | agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttcttt tatggcgagg   | 900     |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [2268] | cggcgggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgacg | 960  |
| [2269] | ctgccttcgc cccgtgcccc gctccgcccgc cgcctcgcgc cgcccgcccc ggctctgact | 1020 |
| [2270] | gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta  | 1080 |
| [2271] | gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgataccttc  | 1140 |
| [2272] | caccaaattc tttattttgg cctctgactg ataaagaatt gtggaaggta tcaggacaca  | 1200 |
| [2273] | aggcccttta tcagcactca catggaacaa atggccaccg tgggaggatg acaatttctg  | 1260 |
| [2274] | tggctgcgtg aaagccttga ggggctccgg gagctagagc ctctgctaac catgttcatg  | 1320 |
| [2275] | ccttcttctt tttcctacag ctctctggga acgtgctggt tattgtgctg tctcatcatt  | 1380 |
| [2276] | ttggcaaaga attcctcgaa gatccgaagg gaaagtcttc cacgactgtg ggatccgttc  | 1440 |
| [2277] | gaagatatca ccggttgagc caccatgagc accctgtgcc cccccccag ccccgccgtg   | 1500 |
| [2278] | gccaaagacc agatcgccct gagcggcaag agccccctgc tggccgccac cttcgcctac  | 1560 |
| [2279] | tgggacaaca tcttgggccc ccgctgcgc cacatctggg cccccaagac cgagcaggtg   | 1620 |
| [2280] | ctgctgagcg acggcgagat caccttctg gccaaaccaca ccctgaacgg cgagatcctg  | 1680 |
| [2281] | cgcaacgccg agagcggcgc catcgacgtg aagttcttcg tgctgagcga gaaggcgctg  | 1740 |
| [2282] | atcatcgtga gcctgatctt cgacggcaac tggaacggcg accgcagcac ctacggcctg  | 1800 |
| [2283] | agcatcatcc tgccccagac cgagctgagc ttctacctgc cctgcaccg cgtgtgcgtg   | 1860 |
| [2284] | gaccgcctga cccacatcat ccgcaagggc cgcactctga tgcacaagga gcgccaggag  | 1920 |
| [2285] | aacgtgcaga agatcatcct ggagggcacc gagcgcattg aggaccaggg ccagagcatc  | 1980 |
| [2286] | atcccatgc tgaccggcga ggtgatcccc gtgatggagc tgctgagcag catgaagagc   | 2040 |
| [2287] | cacagcgtgc ccgaggagat cgacatcgcc gacaccgtgc tgaacgacga cgacatcggc  | 2100 |
| [2288] | gacagctgcc acgagggtt cctgctgaac gccatcagca gccacctgca gacctgcggc   | 2160 |
| [2289] | tgcagcgtgg tgggtggcag cagcgccgag aaggtgaaca agatcgtgcg caccctgtgc  | 2220 |
| [2290] | ctgttcctga cccccccga gcgcaagtgc agccgcctgt gcgaggccga gagcagcttc   | 2280 |
| [2291] | aagtacgaga gcggcctggt cgtgcagggc ctgctgaagg acagcaccgg cagcttcgtg  | 2340 |
| [2292] | ctgcccttcc gccagtgat gtacgcccc taccaccca cccacatcga cgtggacgtg     | 2400 |
| [2293] | aacaccgtga agcagatgcc cccctgccac gagcacatct acaaccagcg ccgctacatg  | 2460 |
| [2294] | cgcagcgagc tgaccgcctt ctggcgccc accagcgagg aggacatggc ccaggacacc   | 2520 |
| [2295] | atcatctaca ccgacgagag cttcaccccc gacctgaaca tcttcagga cgtgctgcac   | 2580 |
| [2296] | cgcgacaccc tgggtgaagge cttcctggac caggtgttcc agctgaagcc cggcctgagc | 2640 |
| [2297] | ctgcgcagca cttcctggc ccagttcctg ctggtgctgc accgcaaggc cctgaccctg   | 2700 |
| [2298] | atcaagtaca tcgaggacga caccagaag ggcaagaagc ctttcaagag cctgcgcaac   | 2760 |
| [2299] | ctgaagatcg acctggacct gaccgccgag ggcgacctga acatcatcat ggccctggcc  | 2820 |
| [2300] | gagaagatca agcccgccct gcacagcttc atcttcggcc gccccttcta caccagcgtg  | 2880 |
| [2301] | caggagcgcg acgtgctgat gaccttctaa caattgttaa ttaagtttaa accctcgagg  | 2940 |
| [2302] | ccgcaagctt atcgataatc aacctctgga ttacaaaatt tgtgaaagat tgactggtat  | 3000 |
| [2303] | tcttaactat gttgctcctt ttacgctatg tggatacgt gctttaatgc ctttgtatca   | 3060 |
| [2304] | tgtatttgct tcccgatgg ctttcatttt ctctccttg tataaatcct ggttgctgtc    | 3120 |
| [2305] | tctttatgag gaggttgtgc ccgttgtcag gcaacgtggc gtggtgtgca ctgtgtttgc  | 3180 |
| [2306] | tgacgaacc cccactggtt ggggcattgc caccacctgt cagctccttt ccgggacttt   | 3240 |
| [2307] | cgctttcccc ctccctattg ccacggcgga actcatgcc gcctgccttg cccgctgctg   | 3300 |
| [2308] | gacaggggct cggctgttg gcactgacaa ttccgtgggt ttgtcgggga aatcatcgtc   | 3360 |
| [2309] | ctttccttgg ctgctcgct gtgttgccac ctggattctg cgcgggacgt ccttctgcta   | 3420 |

|        |                                                                    |      |
|--------|--------------------------------------------------------------------|------|
| [2310] | cgtcccttcg gccctcaatc cagcggacct tccttcccgc ggcctgctgc cggctctgcg  | 3480 |
| [2311] | gcctcttcgc cgtcttcgcc ttgcacctca gacgagtcgg atctcccttt gggccgcctc  | 3540 |
| [2312] | cccgcatcga taccgtcgac tagagctcgc tgatcagcct cgactgtgcc ttctagttagc | 3600 |
| [2313] | cagccatctg ttgtttgccc ctccccctgt ccttccttga ccctggaagg tgccactccc  | 3660 |
| [2314] | actgtccttt cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct  | 3720 |
| [2315] | attctggggg gtgggggtgg gcaggacagc aagggggagg attggaaga caatagcagg   | 3780 |
| [2316] | catgctgggg agagatccac gataacaaac agcttttttg gggatgaacat attgactgaa | 3840 |
| [2317] | ttccctgcag gttggccact ccctctctgc gcgtcgtcgc gctcactgag gccgcccggg  | 3900 |
| [2318] | caaagcccgg gcgtcgggag acctttgggc gcccgccctc agtgagcgag cgagcgcgca  | 3960 |
| [2319] | gagagggagt ggccaactcc atcactaggg gttcct                            | 3996 |
| [2320] | <210> 62                                                           |      |
| [2321] | <211> 3994                                                         |      |
| [2322] | <212> DNA                                                          |      |
| [2323] | <213> 人工序列                                                         |      |
| [2324] | <220>                                                              |      |
| [2325] | <223> 合成的多核苷酸                                                      |      |
| [2326] | <400> 62                                                           |      |
| [2327] | ttggccactc cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc   | 60   |
| [2328] | cgacgcccgg gctttgcccg ggcgccctca gtgagcgagc gagcgcgagc agagggagt   | 120  |
| [2329] | gccaaactcca tctactaggg ttctgtctag ctctgggtat ttaagcccga gtgagcacgc | 180  |
| [2330] | agggtctcca ttttgaagcg ggaggttacg cgttcgtcga ctactagtgg gtaccagagc  | 240  |
| [2331] | tccttaggtt ctagaaccgg tgacgtctcc catggtgaag cttggatctg aattcggtag  | 300  |
| [2332] | cctagttatt aatagtaatc aattacgggg tcattagtgc atagcccata tatggagtgc  | 360  |
| [2333] | cgcgttacat aacttacggg aaatggccc cctggctgac cgcccaacga ccccgccca    | 420  |
| [2334] | ttgacgtcaa taatgacgta tgttcccata gtaacgcaa tagggacttt ccattgacgt   | 480  |
| [2335] | caatgggtgg actatttacg gtaaatgcc cacttggcag tacatcaagt gtatcatatg   | 540  |
| [2336] | ccaagtacgc cccctattga cgtcaatgac ggtaaattggc ccgcctggca ttatgccag  | 600  |
| [2337] | tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt  | 660  |
| [2338] | accatgggtg aggtgagccc caggttctgc ttactctcc ccattctccc cccctccca    | 720  |
| [2339] | cccccaattt tgtatttatt tattttttaa ttattttgtg cagcgatggg ggcggggggg  | 780  |
| [2340] | gggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgagcg    | 840  |
| [2341] | agagggtcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg  | 900  |
| [2342] | cggcgggcgg gcggggccca taaaaagcga agcgcgcggc gggcgggagt cgctgcgagc  | 960  |
| [2343] | ctgccttcgc cccgtgcccc gctccgccgc cgctcgcgc cgcccgcgcc ggctctgact   | 1020 |
| [2344] | gaccgcgtta ctccacagg tgagcgggag ggacggccct tctcctccgg gctgtaatta   | 1080 |
| [2345] | gcgcttggtt taatgacggc ttgtctggag gcttgctttg ggctgtatgc tgataagtcc  | 1140 |
| [2346] | tttactaagg agcttttggc ctctgactga gctccttga aggacttatac aggacacaag  | 1200 |
| [2347] | gcccattatc agcactcaca tggaacaaat ggccaccgtg ggaggatgac aatttctgtg  | 1260 |
| [2348] | gctgcgtgaa agccttgagg ggctccggga gctagagcct ctgctaacca tgttcatgcc  | 1320 |
| [2349] | ttcttctttt tcctacagct cctgggcaac gtgctgggta ttgtgctgtc tcatcatttt  | 1380 |
| [2350] | ggcaaagaat tcctcgaaga tccgaaggga aagtcttcca cgactgtggg atccgttcga  | 1440 |
| [2351] | agatatcacc ggttgagcca ccatgagcac cctgtgcccc cccccagcc ccgcgtggc    | 1500 |

|        |                                                                   |      |
|--------|-------------------------------------------------------------------|------|
| [2352] | caagaccgag atcgccctga gcggaagag cccctgctg gccgccacct tgcctactg    | 1560 |
| [2353] | ggacaacatc ctgggcccc gctgcgcca catctgggcc cccaagaccg agcaggtgct   | 1620 |
| [2354] | gctgagcgac ggcgagatca ctttcttgge caaccacacc ctgaacggcg agatcctgcg | 1680 |
| [2355] | caacgccgag agcgggcgcca tcgacgtgaa gttcttctg ctgagcgaga agggcgtgat | 1740 |
| [2356] | catcgtgagc ctgatcttcg acggcaactg gaacggcgac cgcagcacct acggcctgag | 1800 |
| [2357] | catcatcctg cccagaccg agctgagctt ctacctgccc ctgcaccgcg tgtgctgga   | 1860 |
| [2358] | ccgctgacc cacatcatcc gcaaggccg catctggatg cacaaggagc gccaggagaa   | 1920 |
| [2359] | cgtgcagaag atcatcctgg agggcaccga gcgcatggag gaccagggcc agagcatcat | 1980 |
| [2360] | cccatgctg accggcgagg tgatccccgt gatggagctg ctgagcagca tgaagagcca  | 2040 |
| [2361] | cagcgtgccc gaggagatcg acatcgccga caccgtgctg aacgacgacg acatcgccga | 2100 |
| [2362] | cagctgccac gagggttcc tgctgaacgc catcagcagc cacctgcaga cctgcggtg   | 2160 |
| [2363] | cagcgtggtg gtgggcagca gcgccgagaa ggtgaacaag atcgtgcga cctgtgcct   | 2220 |
| [2364] | gttctgacc cccgccgagc gcaagtgcag ccgctgtgc gaggccgaga gcagcttcaa   | 2280 |
| [2365] | gtacgagagc ggcctgttcg tgcaggcct gctgaaggac agcaccggca gttcgtgct   | 2340 |
| [2366] | gcccttccgc caggtgatgt acgcccccta cccaccacc cacatcgacg tggacgtgaa  | 2400 |
| [2367] | caccgtgaag cagatgcccc cctgccacga gcacatctac aaccagcgcc gctacatgcg | 2460 |
| [2368] | cagcgagctg accgccttct ggcgcgccac cagcgaggag gacatggccc aggacacat  | 2520 |
| [2369] | catctacacc gacgagagct tcacccccga cctgaacatc ttccaggacg tgctgcaccg | 2580 |
| [2370] | cgacaccctg gtgaaggcct tcctggacca ggtgttccag ctgaagcccg gcctgagcct | 2640 |
| [2371] | gcgcagcacc ttcttgcccc agttctgct ggtgctgcac cgcaaggccc tgacctgat   | 2700 |
| [2372] | caagtacatc gaggacgaca ccagaaggg caagaagccc ttcaagagcc tgcgcaacct  | 2760 |
| [2373] | gaagatcgac ctggacctga ccgccgagg cgacctgaac atcatcatgg ccctggccga  | 2820 |
| [2374] | gaagatcaag cccggcctgc acagcttcat cttcgccgc cccttctaca ccagcgtgca  | 2880 |
| [2375] | ggagcgcgac gtgctgatga ctttctaaca attgttaatt aagtttaaac cctcaggcc  | 2940 |
| [2376] | gcaagcttat cgataatcaa cctctggatt acaaaattg tgaaagattg actggtattc  | 3000 |
| [2377] | ttaactatgt tgctcctttt acgctatgtg gatacgtgc tttaatgcct ttgtatcatg  | 3060 |
| [2378] | ctattgcttc ccgtatggct ttcatcttct cctccttgta taaatcctgg ttgctgtctc | 3120 |
| [2379] | tttatgagga gttgtggccc gttgtcaggc aacgtggcgt ggtgtgcact gtgtttgctg | 3180 |
| [2380] | acgcaacccc cactggttg ggcatcgcca ccacctgtca gctcctttcc gggactttcg  | 3240 |
| [2381] | ctttccccct ccctattgcc acggcggaac tcatcgccgc ctgccttgcc cgtgctgga  | 3300 |
| [2382] | caggggctcg gctgttgggc actgacaatt ccgtggtgtt gtcgggaaa tcatgctcct  | 3360 |
| [2383] | ttccttggtg gctcgctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg  | 3420 |
| [2384] | tccttcggc cctcaatcca gcggacctt cttcccgcg cctgctgccg gctctgaggc    | 3480 |
| [2385] | ctcttcggc tcttcgctt cgccctcaga cgagtcggat ctccctttgg gccgctccc    | 3540 |
| [2386] | cgcacgata ccgtcgacta gagctcgtg atcagcctcg actgtgcct ctagttgcca    | 3600 |
| [2387] | gccatctgtt gtttggccct ccccggtgcc ttccttgacc ctggaagggt ccaactccac | 3660 |
| [2388] | tgctctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat  | 3720 |
| [2389] | tctgggggt ggggtgggc aggacagca gggggaggat tgggaagaca atagcaggca    | 3780 |
| [2390] | tgctggggag agatccacga taacaacag cttttttgg gtgaacatat tgactgaatt   | 3840 |
| [2391] | ccctgcagg tggccactcc ctctctgcgc gctcgtcgc tctgagggc cgccgggca     | 3900 |
| [2392] | aagccgggc gtcggcgac ctttgctgc ccggcctcag tgagcgagcg agcgcgaga     | 3960 |
| [2393] | gaggagtg ccaactccat cactaggggt tcct                               | 3994 |

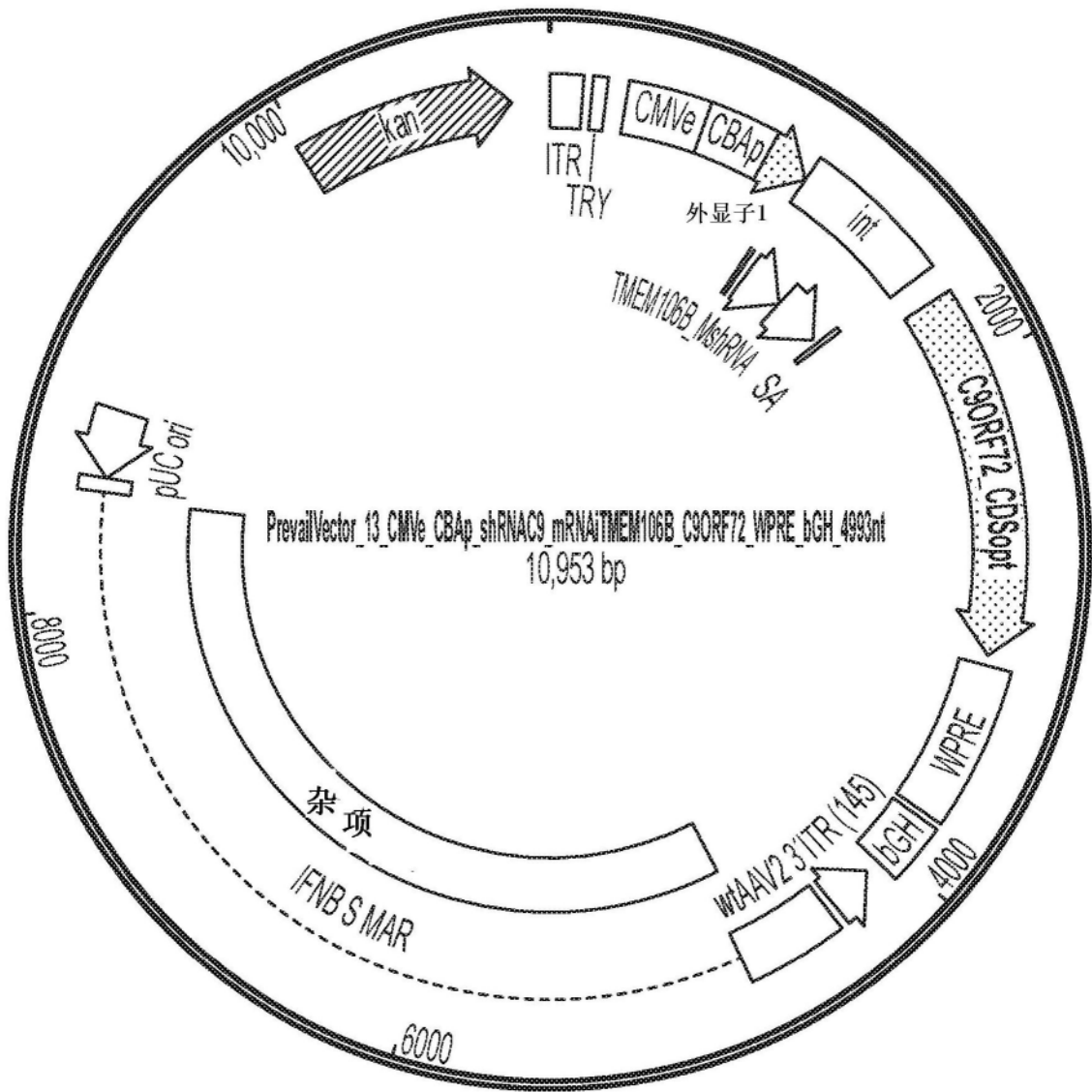


图1

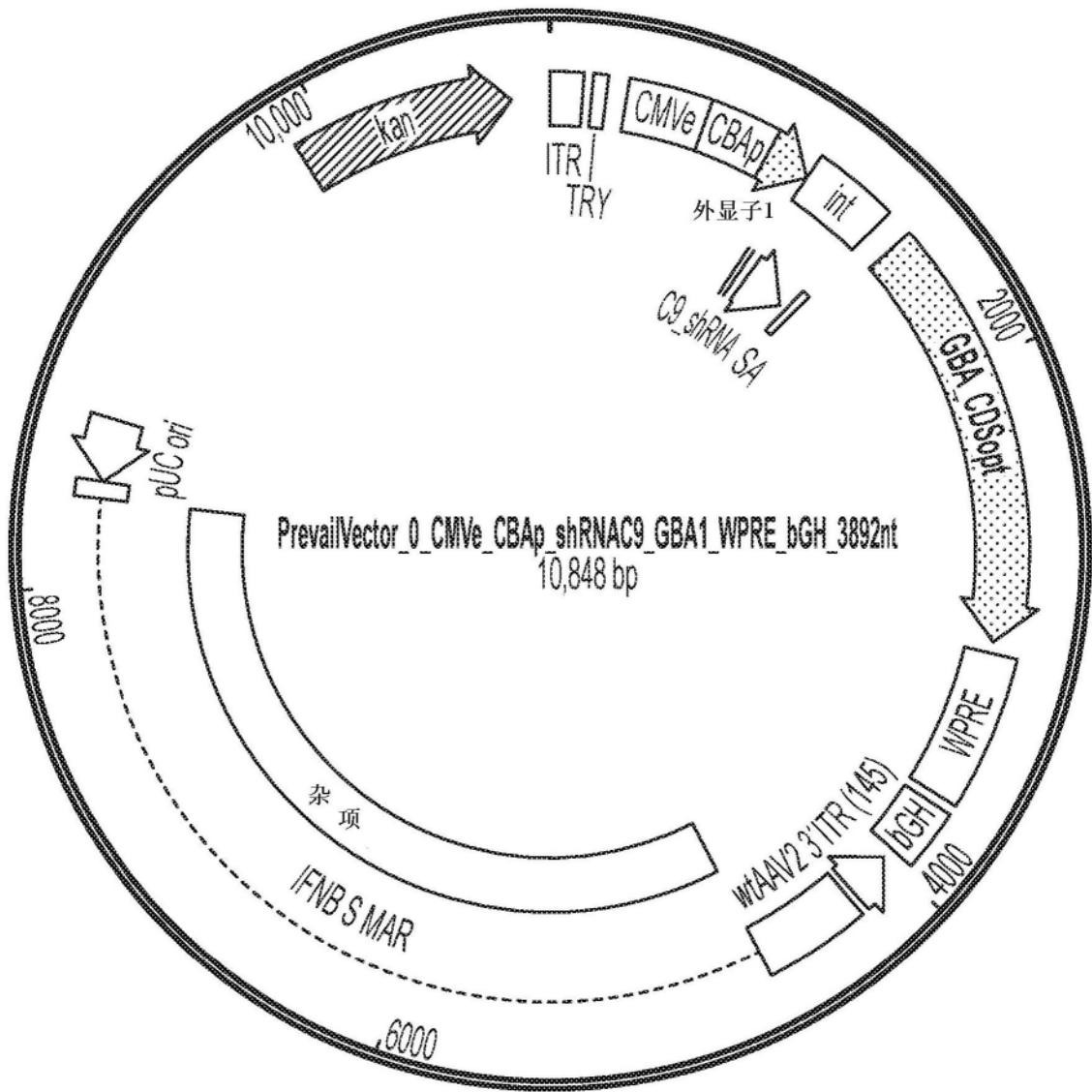


图2

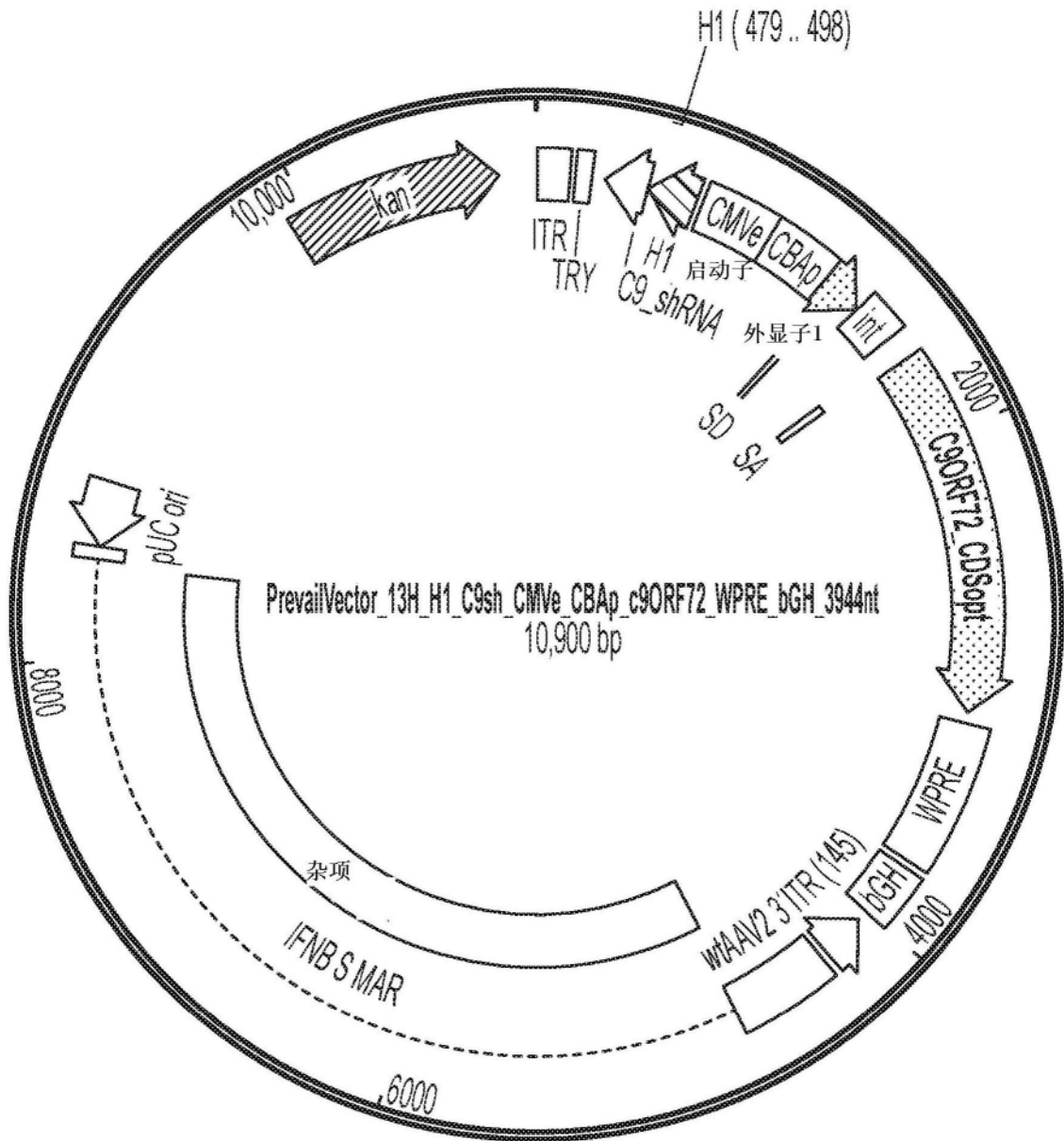


图3

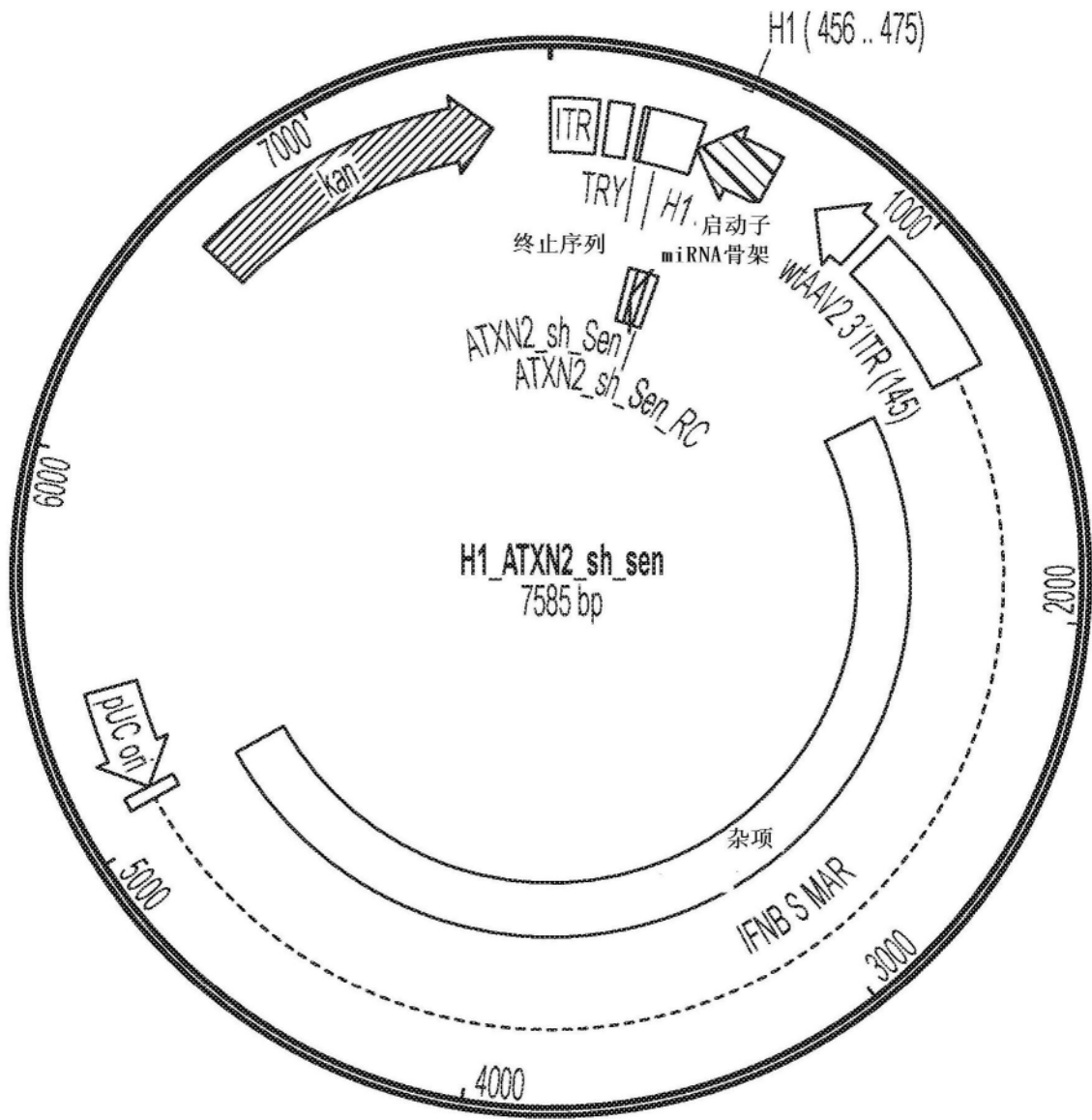


图4

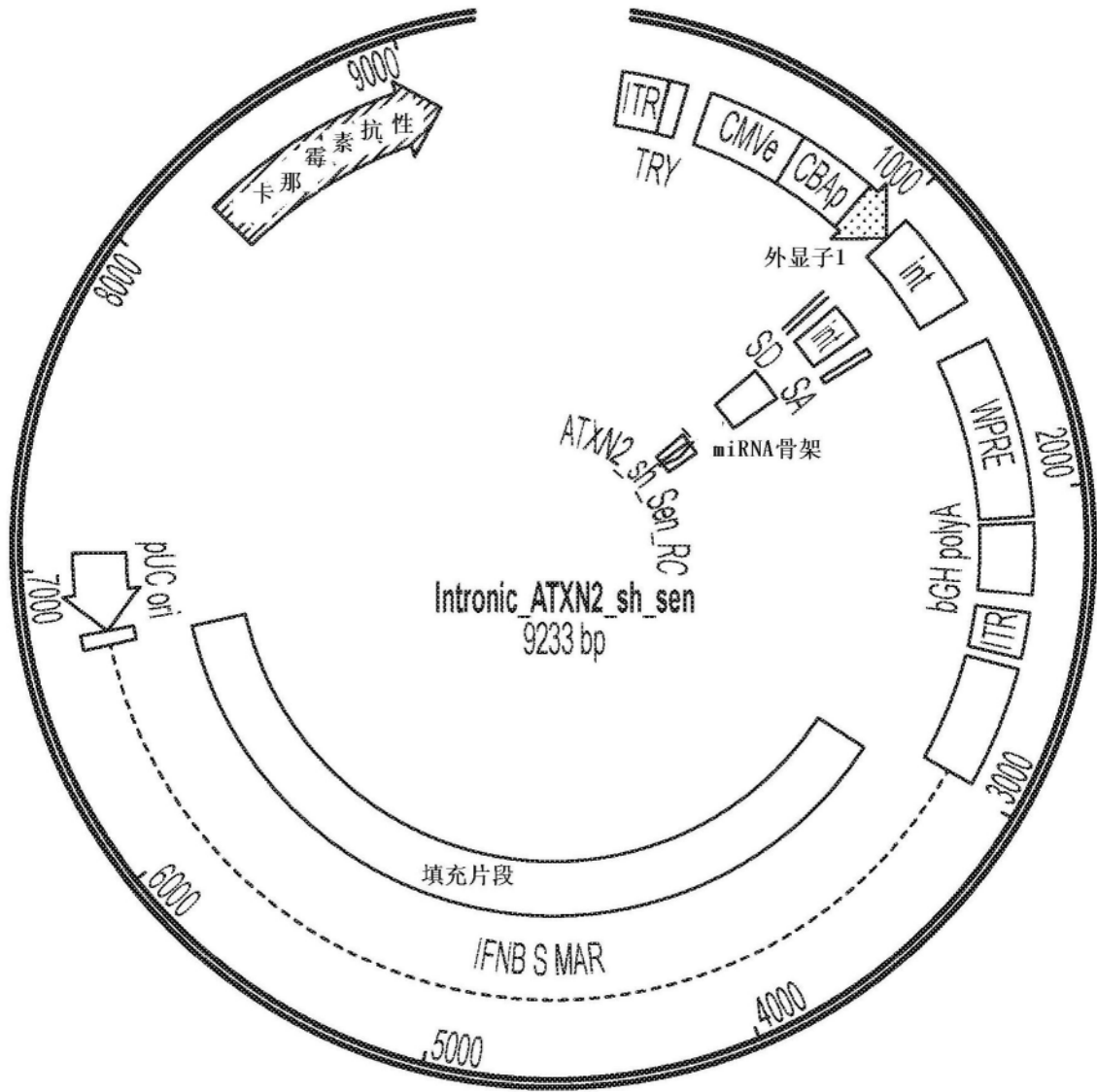


图5

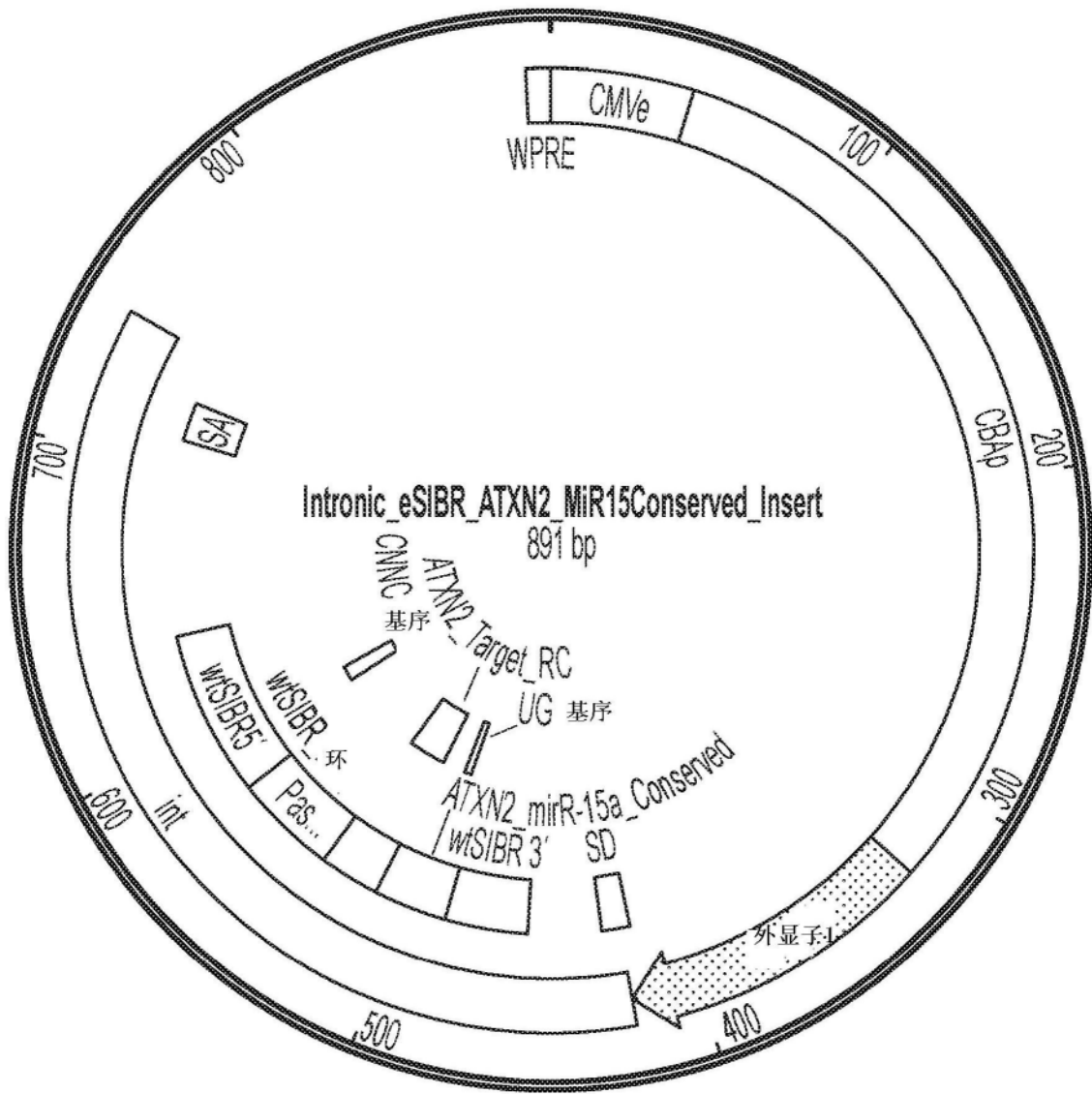


图6

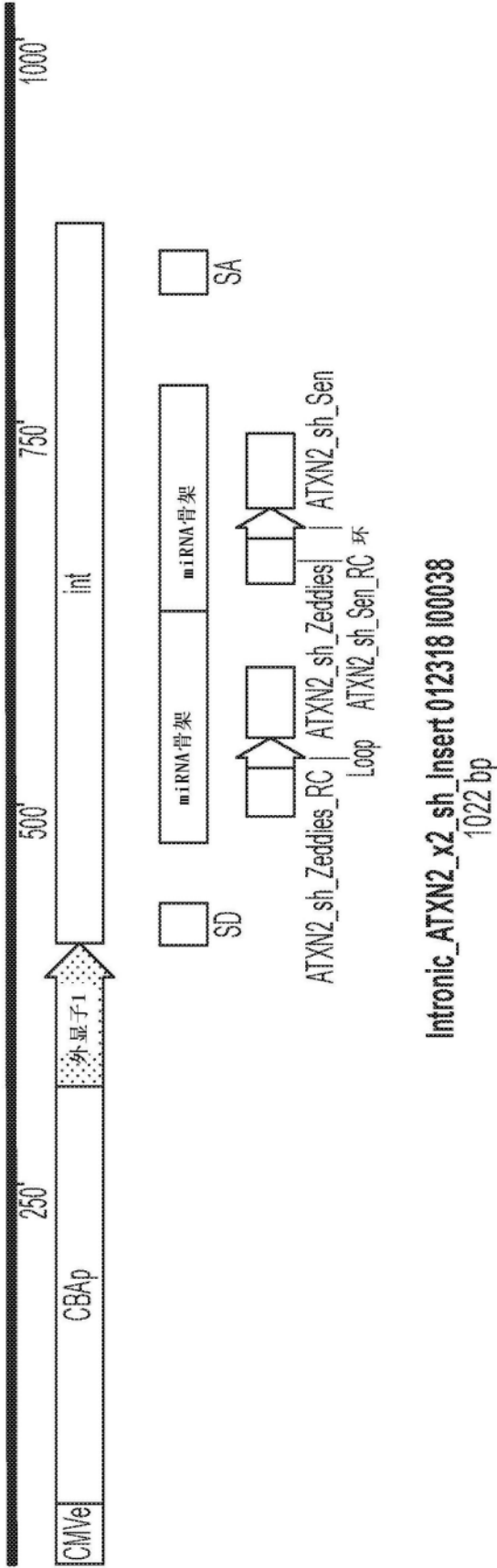


图7

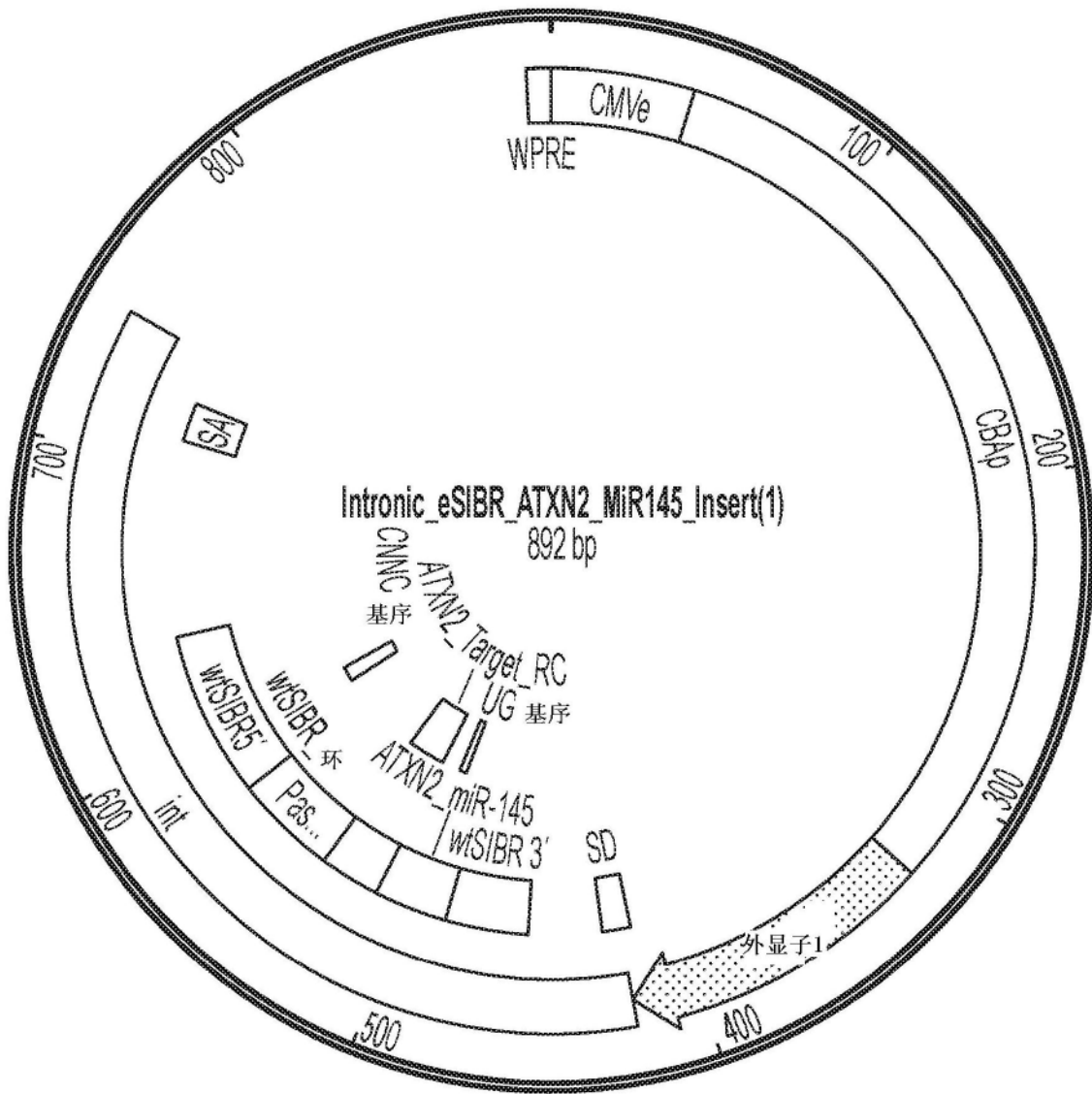


图8



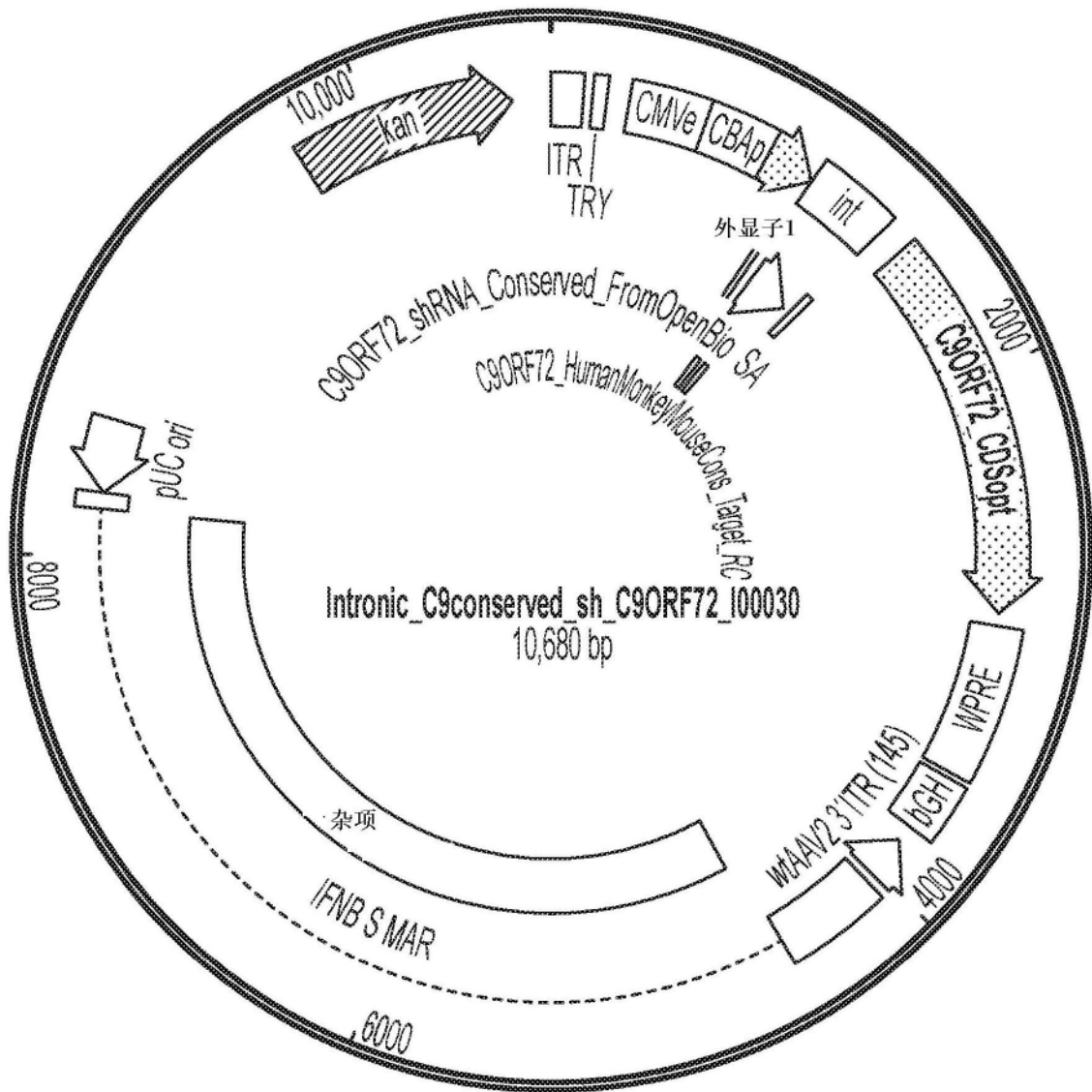


图10

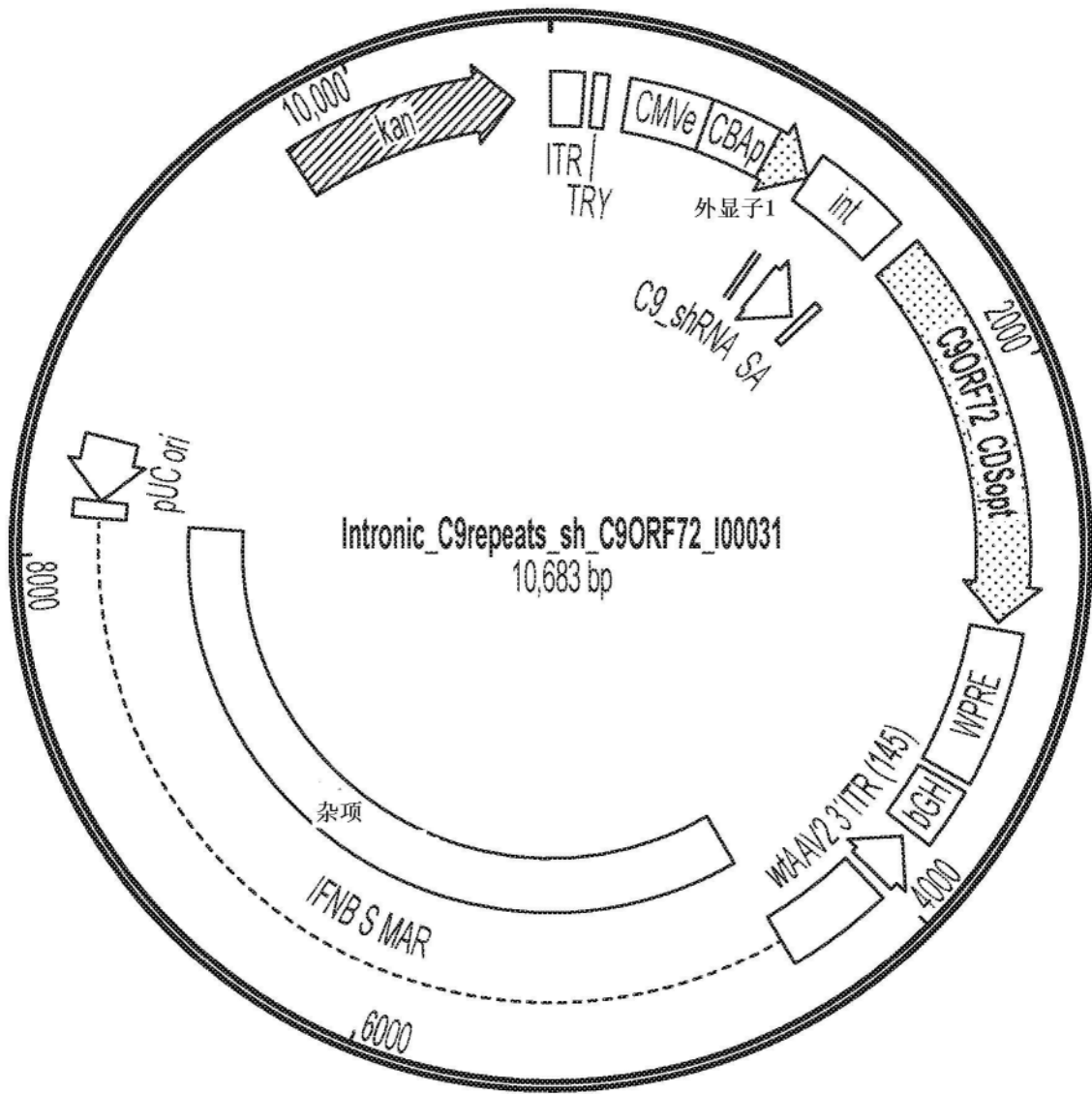


图11

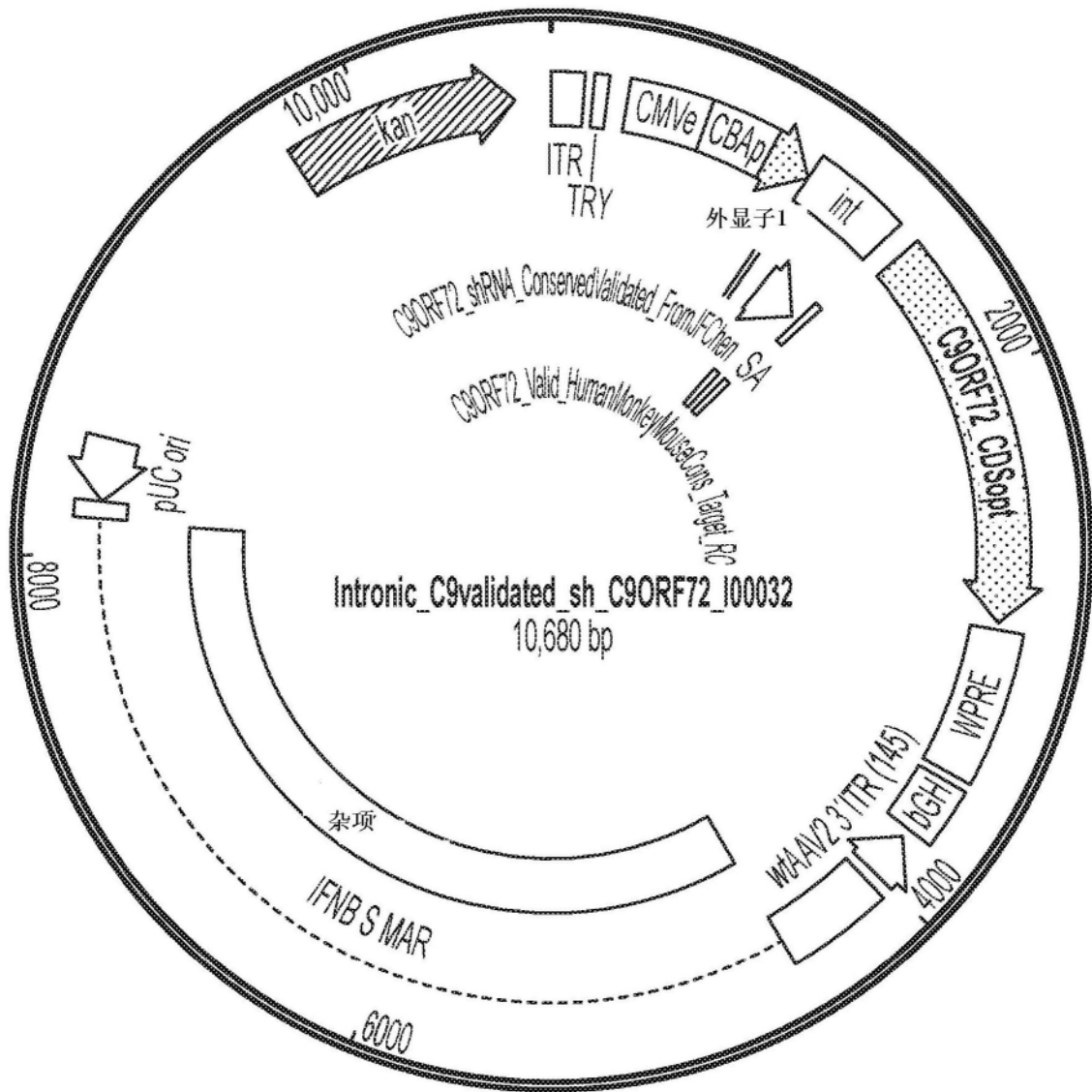


图12

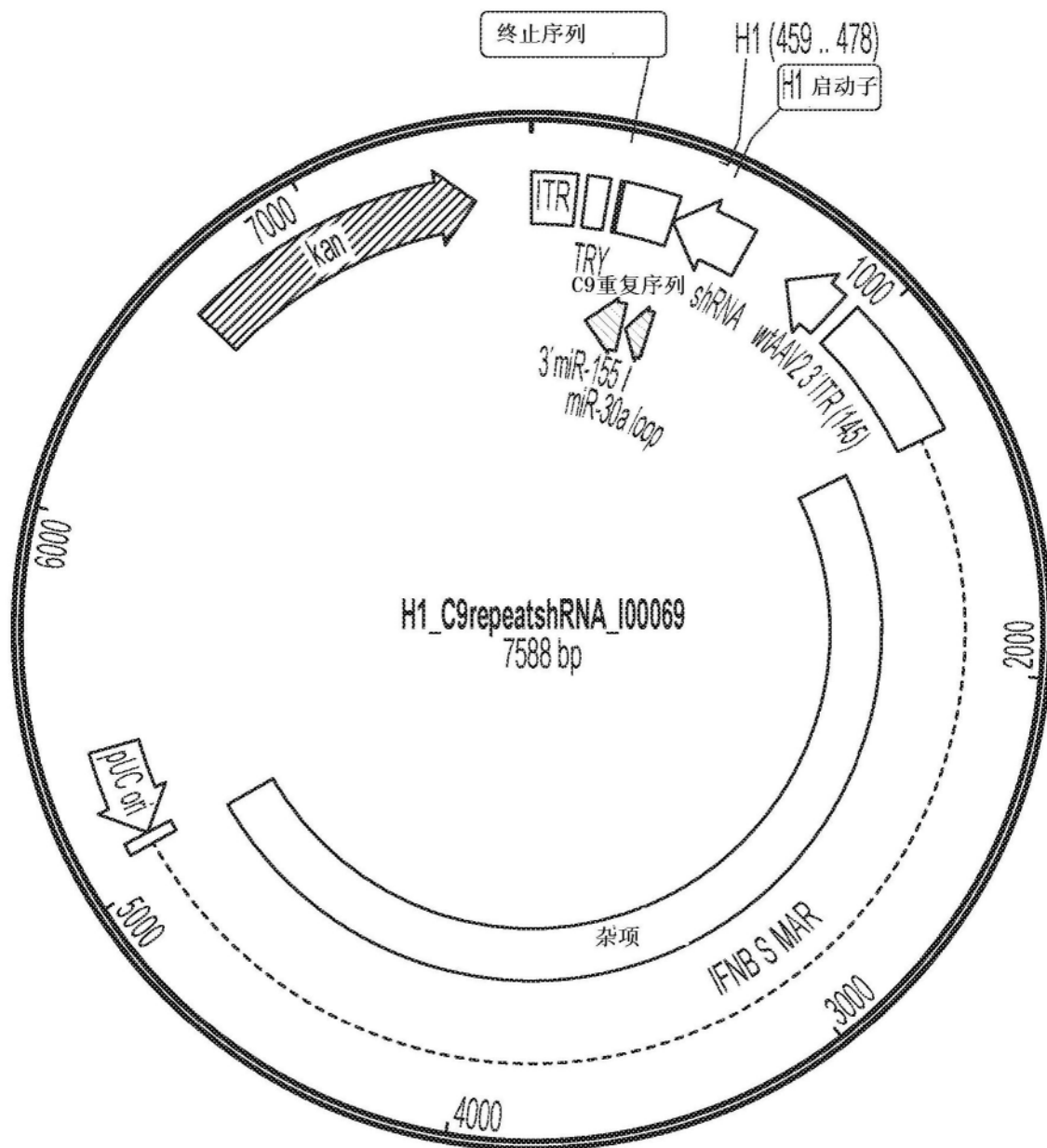


图13

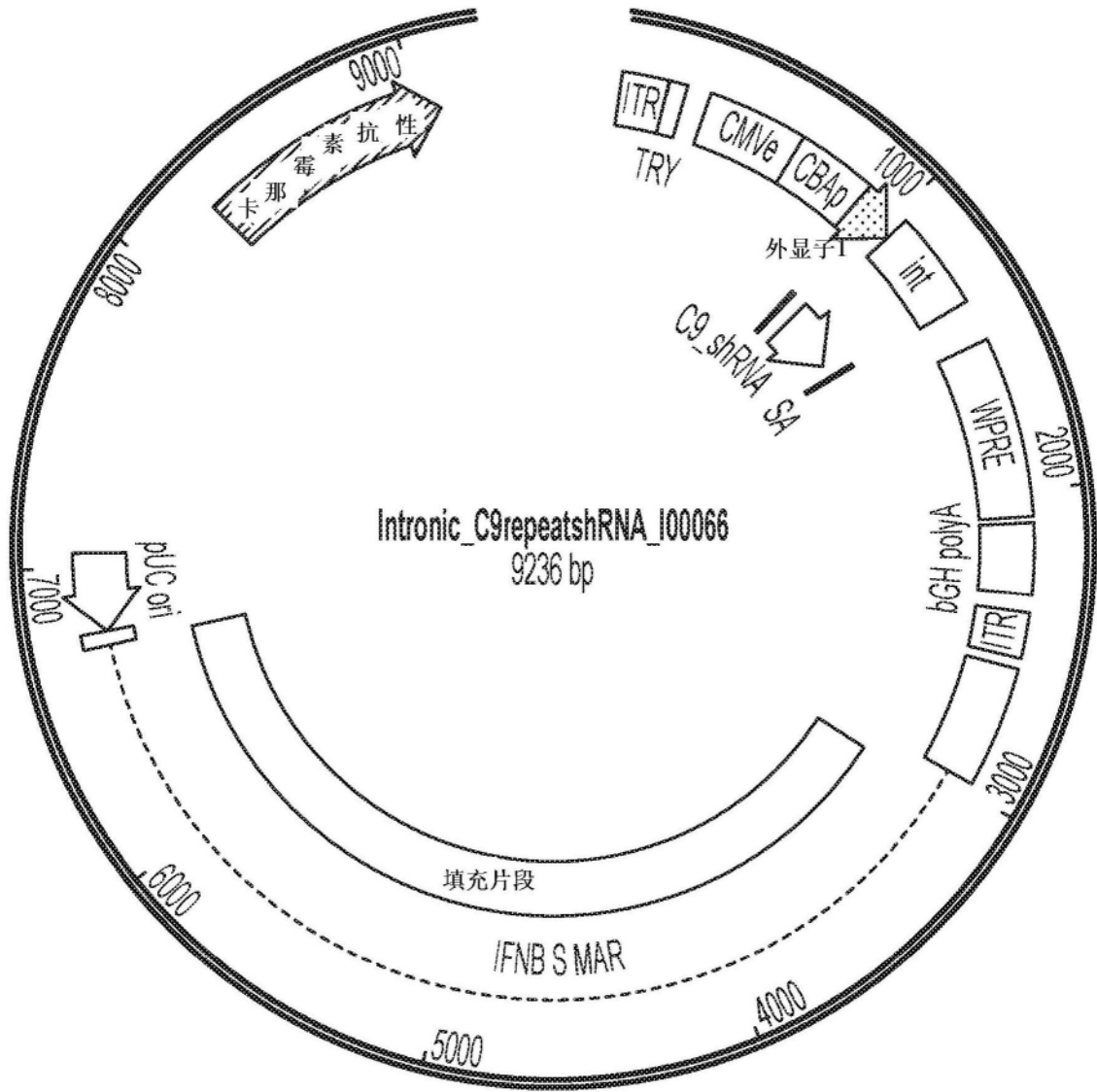
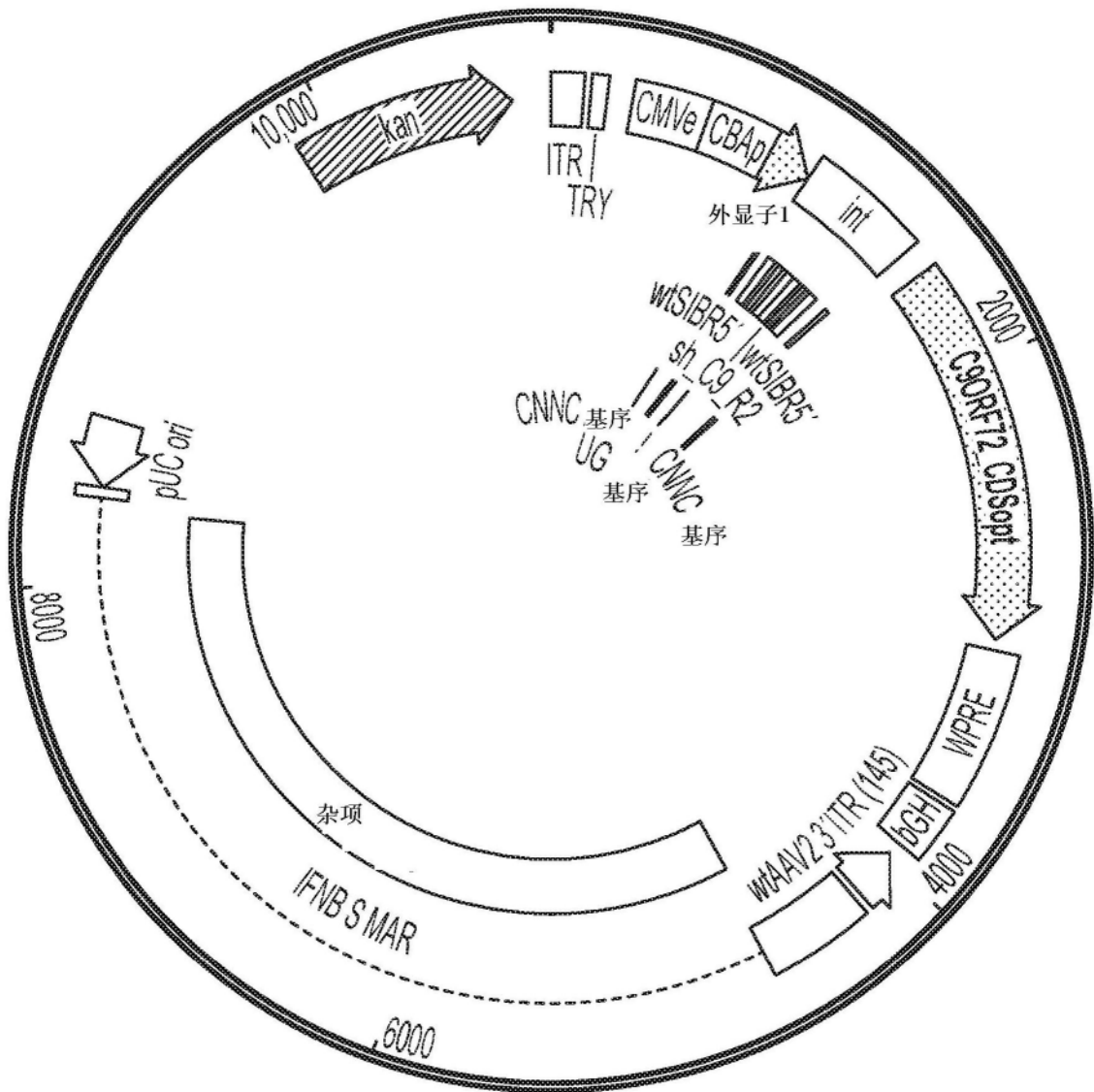
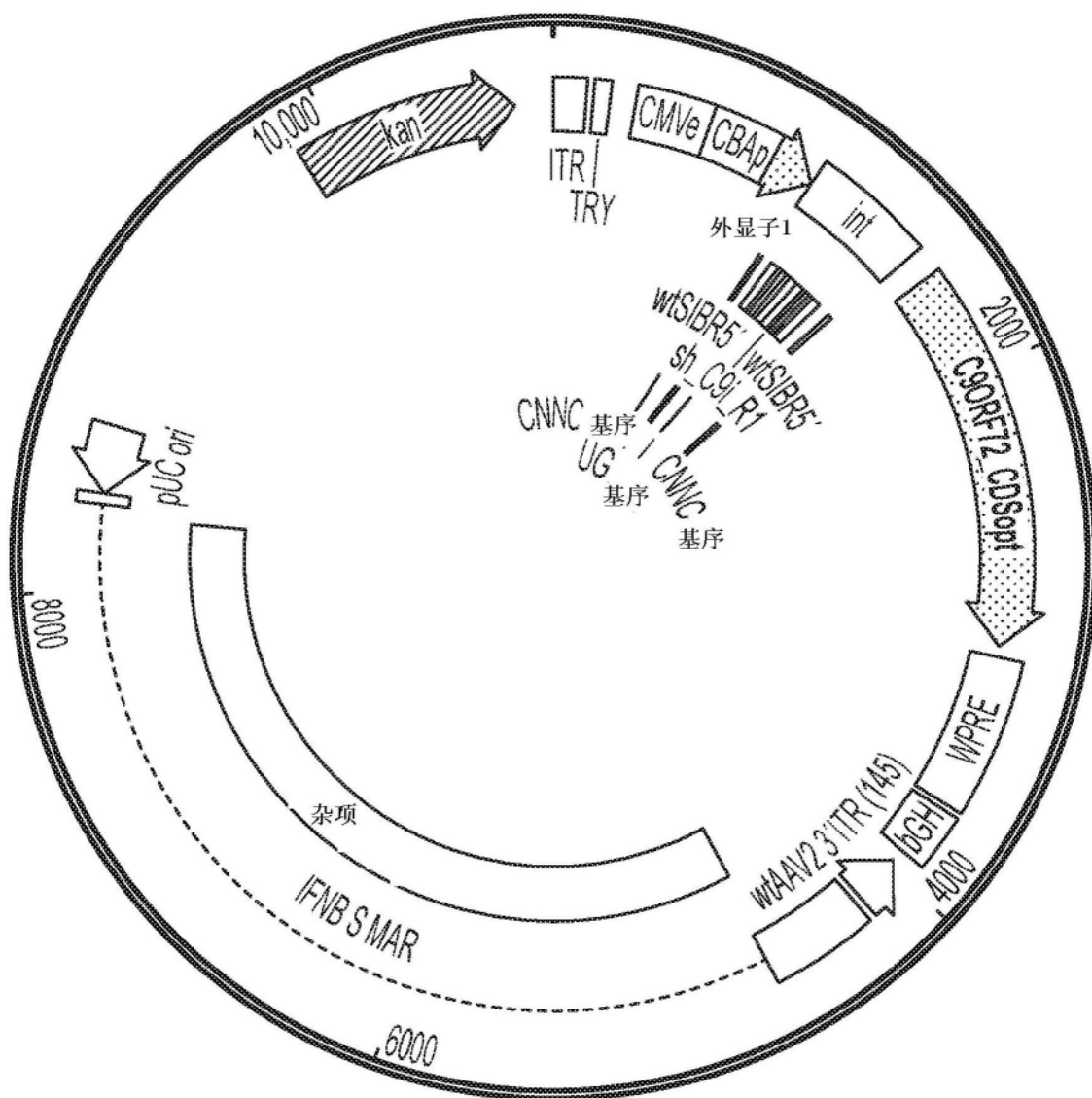


图14



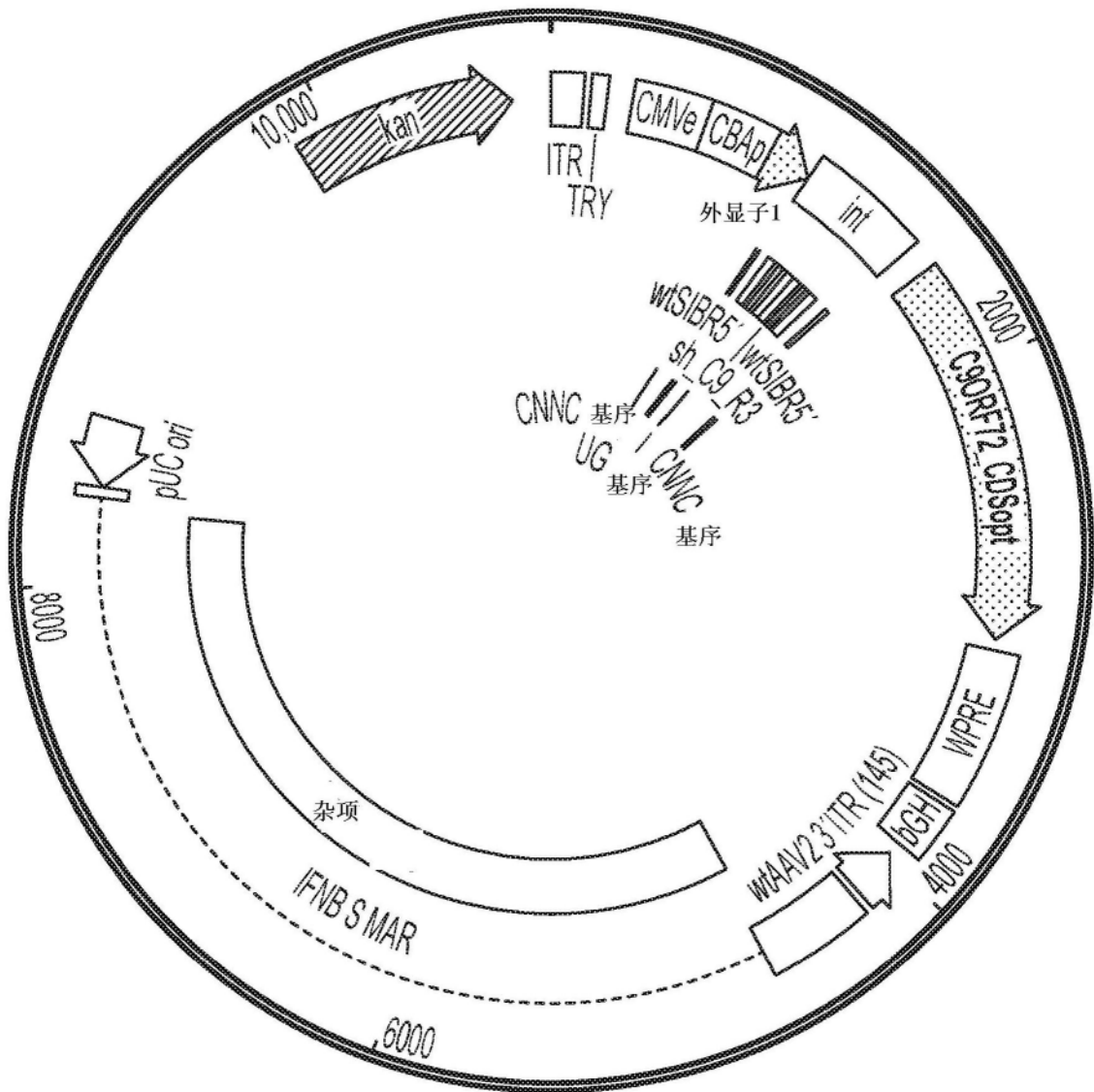
Intronic\_eSIBR\_Anti\_C9\_IntronicShL2\_IntronicShR2\_CMV\_e\_CBAP\_C9ORF72\_WPRE\_bGH\_I00134  
10,828 bp

图15



Intronic\_eSIBR\_Anti\_C9\_IntronicShL1\_IntronicShR1\_CMVe\_CBAP\_C9ORF72\_WPRE\_bGH\_I00133  
10,829 bp

图16



Intronic\_eSIBR\_Anti\_C9\_IntronicShL3\_IntronicShR3\_CMV\_e\_CBAp\_C9orf72\_WPRE\_bGH\_I00135  
10,826 bp

图17

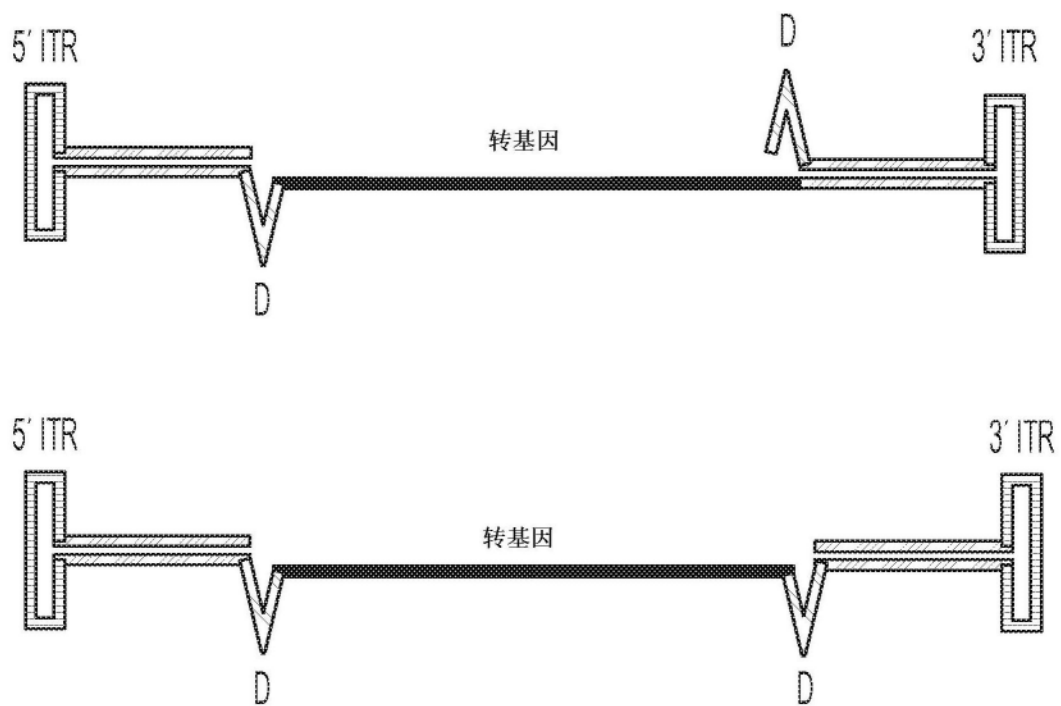


图18

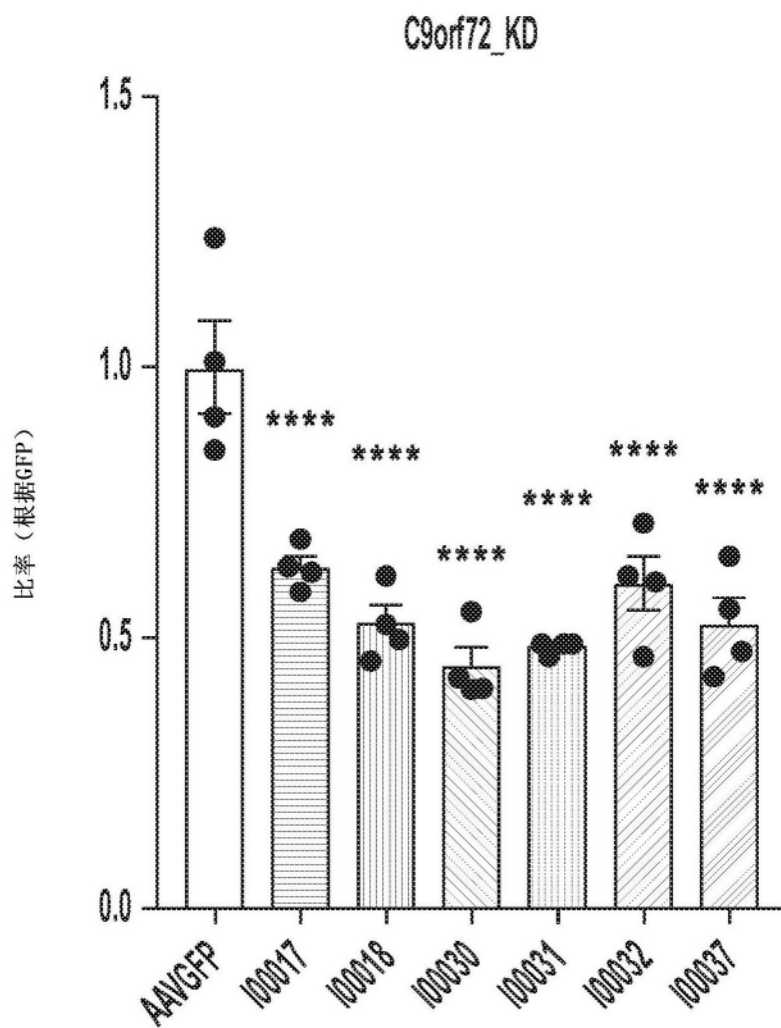


图19A

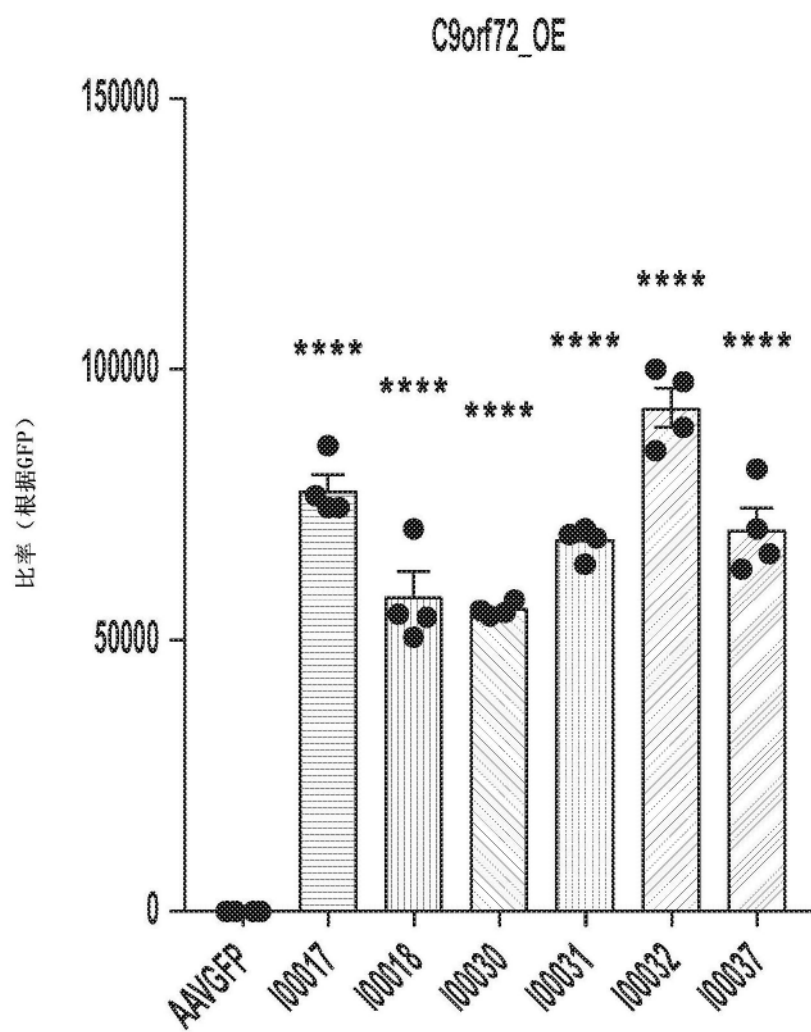
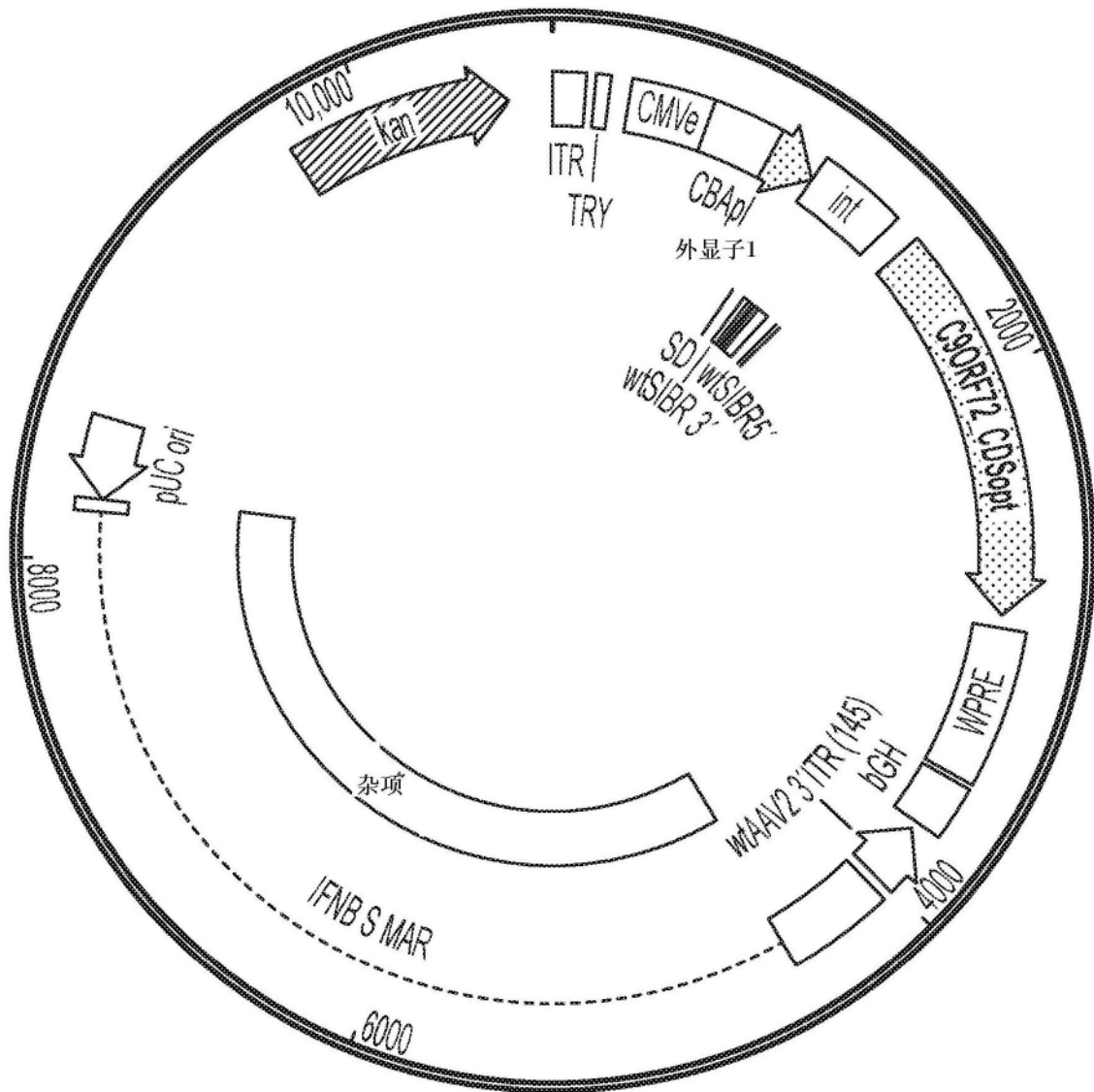
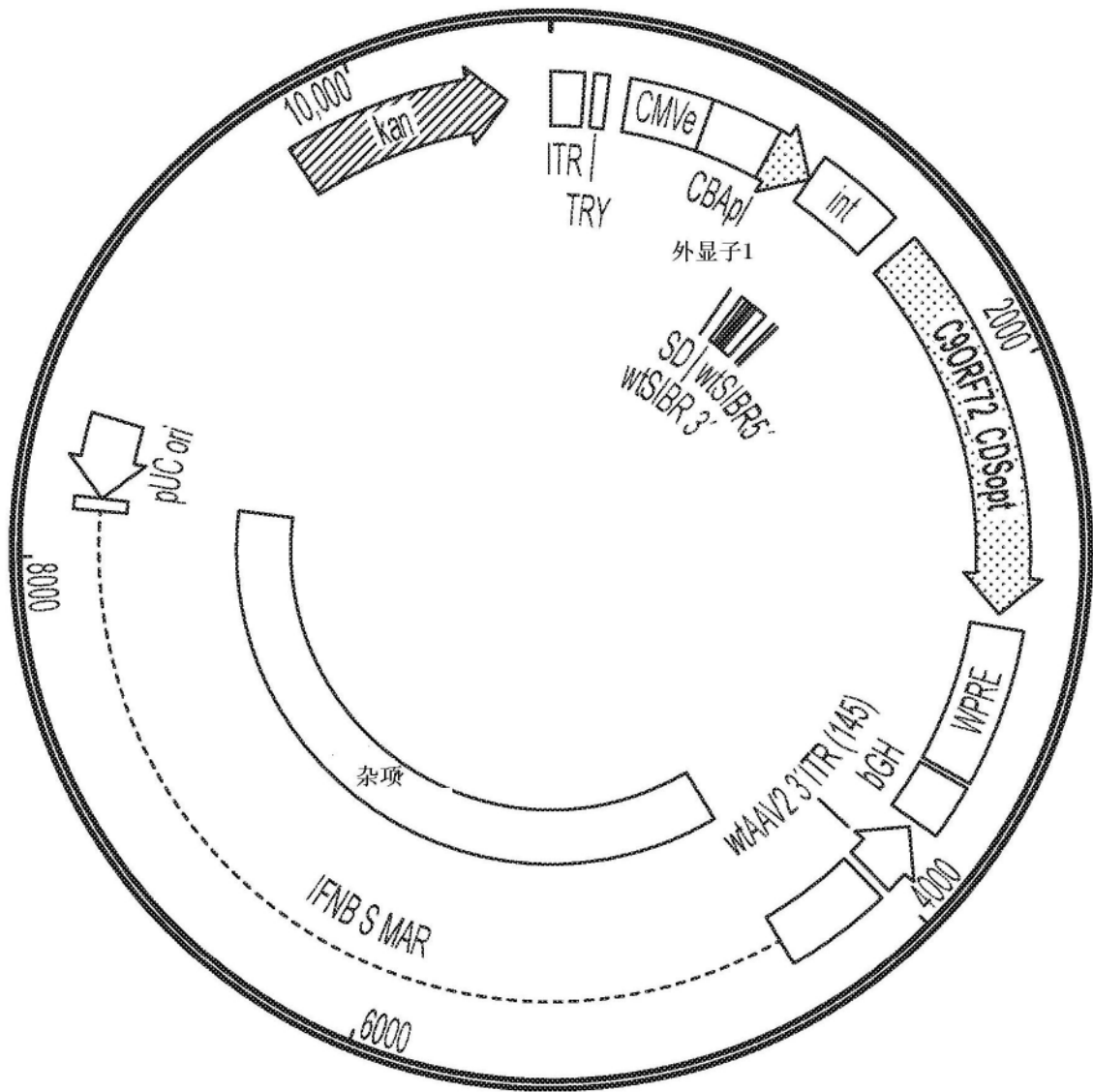


图19B



Intronic\_eSIBR\_RPS25\_MiR\_sh\_C9ORF72\_I00079  
10,682 bp

图20



Intronic\_eSIBR\_RPS25\_BroadConserved\_sh\_C9ORF72\_I00078  
10,680 bp

图21