(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111587119 A
(43)申请公布日 2020.08.25

(21)申请号 201880086289.X

(22)申请日 2018.12.05

(30)优先权数据
62/594,811 2017.12.05 US

(85)PCT国际申请进入国家阶段日
2020.07.09

(86)PCT国际申请的申请数据
PCT/US2018/064091 2018.12.05

(87)PCT国际申请的公布数据
WO2019/113225 EN 2019.06.13

(71)申请人 华盛顿大学
地址 美国华盛顿

(72)发明人 M·尼茨 J·尼茨

(74)专利代理机构 北京市中伦律师事务所 11410
代理人 杨黎峰 王奕勋

(51)Int.Cl.
A61K 35/761(2006.01)
C12N 15/67(2006.01)
C12N 15/861(2006.01)
C12N 15/86(2006.01)
C07K 14/72(2006.01)

权利要求书7页 说明书51页
序列表156页 附图8页

(54)发明名称

用于增强治疗性基因在光感受器中的功能性表达的组合物和方法

(57)摘要

本文公开了核酸调节元件、表达盒、表达载体以及使用所述核酸调节元件、表达盒、表达载体治疗视锥细胞病症的方法。

CN 111587119 A

1.一种分离的核酸,其包括序列:

GGGAGGAG GAGGTCTAAG TCCCAGGCCC

AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG

CCCGGGCTGA TCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGAC

GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC(X) (SEQ ID NO: 1);

其中X不存在,或者选自由以下组成的组:ATAGCCGGTACC(ATG)(SEQ ID NO:104)和GCCGCCACC。

2.根据权利要求1所述的分离的核酸,其中所述核酸序列包括SEQ ID NO:3或SEQ ID NO:4的序列。

3.根据权利要求1所述的分离的核酸,其中所述核酸序列包括SEQ ID NO:5的序列。

4.根据权利要求1至3中任一项所述的分离的核酸,其中所述核酸序列不包含SEQ ID NO:70的序列。

5.一种核酸表达盒,其包括:

(a)基因座控制区(LCR),所述LCR包括SEQ ID NO:2的序列、基本上由所述序列组成或由所述序列组成;以及

(b)启动子,所述启动子包括根据权利要求1至4中任一项所述的核酸序列、基本上由所述核酸序列组成或由所述核酸序列组成,其中所述启动子定位于所述LCR的3'端。

6.根据权利要求5所述的核酸表达盒,其中所述启动子包括SEQ ID NO:4或SEQ ID NO:5的序列或由所述序列组成。

7.根据权利要求6所述的核酸表达盒,其中所述启动子包括SEQ ID NO:4的序列或由所述序列组成,其中所述盒进一步包括:

(c)内含子,所述内含子包括SEQ ID NO:6的序列、基本上由所述序列组成或由所述序列组成,其中所述内含子定位于所述启动子的3'端;以及

(d)Kozak序列,如包括序列GCCGCCACC、基本上由所述序列组成或由所述序列组成的Kozak序列,其中所述Kozak序列定位于所述内含子的3'端。

8.根据权利要求7所述的核酸表达盒,其进一步包括

(e)克隆位点,所述克隆位点定位于所述Kozak序列的3'端。

9.根据权利要求7所述的核酸表达盒,其进一步包括

(e)编码区,所述编码区定位于所述Kozak序列的3'端。

10.根据权利要求9所述的核酸表达盒,其中所述编码区不包含任何内含子。

11.根据权利要求7所述的核酸表达盒,其中所述启动子包括SEQ ID NO:5的序列或由所述序列组成。

12.根据权利要求11所述的核酸表达盒,其中所述盒进一步包括定位于所述启动子的3'端的克隆位点。

13.根据权利要求11所述的核酸表达盒,其中所述盒进一步包括定位于所述启动子的3'端的编码区。

14.根据权利要求13所述的核酸表达盒,其中所述编码区包含一个或多个内含子。

15.根据权利要求9、10、13或14所述的核酸表达盒,其中所述编码区包括编码选自由以

下组成的组的多肽的核酸序列:OPN1LW(例如OPN1LW的外显子1-6)、OPN1MW(例如OPN1MW的外显子1-6)、序列14(经过修饰的阿柏西普)或其功能片段、衍生物或变体;RS1(视网膜劈裂蛋白前体)、针对VEGFA的抗体(如单克隆抗体)、抗血管内皮生长因子(VEGF)抗体或其抗体片段如兰尼单抗和/或贝伐单抗、色素上皮衍生因子(PEDF)、可溶性fms样酪氨酸激酶-1(FLT1)、CD59、视锥视蛋白、视锥离子通道,或者选自由以下组成的组的多肽:

(a)SEQ ID NO:36智人视蛋白1(视锥色素),对短波敏感(OPN1SW);

(b)SEQ ID NO:37智人视蛋白1(视锥色素),对中波敏感(OPN1MW);

(c)SEQ ID NO:38智人视蛋白1(视锥色素),对长波敏感(OPN1LW);

(d)SEQ ID NO:39ATP结合盒视网膜基因(ABCR)基因(NM_000350);

(e)SEQ ID NO:40视网膜色素上皮特异性65kD蛋白基因(RPE65)(NM_000329);

(f)SEQ ID NO:41视网膜结合蛋白1基因(RLBP1)(NM_000326);

(g)SEQ ID NO:42(外周蛋白/视网膜变性慢基因,(NM_000322);

(h)SEQ ID NO:43抑制蛋白(SAG)(NM_000541);

(i)SEQ ID NO:44α-转导蛋白(GNAT1)(NM_000172);

(j)SEQ ID NO:45鸟苷酸环化酶活化蛋白1A(GUCA1A)(NP_000400.2);

(k)SEQ ID NO:46视网膜特异性鸟苷酸环化酶(GUCY2D),(NP_000171.1);

(l)SEQ ID NO:47和/或48视锥环核苷酸门控阳离子通道的α亚基(CNGA3)(NP_001073347.1或NP_001289.1);

(m)SEQ ID NO:49人视锥转导蛋白α亚基(不完全性全色盲);

(n)SEQ ID NO:50视锥cGMP特异性3',5'-环磷酸二酯酶亚基α',蛋白质(4型视锥营养不良);

(o)SEQ ID NO:51视网膜视锥视紫红质敏感性cGMP 3',5'-环磷酸二酯酶亚基γ,蛋白质(3A型视网膜视锥营养不良);

(p)SEQ ID NO:52视锥视杆同源盒,蛋白质(视锥-视杆营养不良);

(q)SEQ ID NO:53视锥光感受器环核苷酸门控通道β亚基,蛋白质(全色盲);

(r)SEQ ID NO:54视锥光感受器cGMP门控阳离子通道β亚基,蛋白质(全色盲,例如见于平格拉普环礁岛民);

(s)SEQ ID NO:55视网膜色素变性1(常染色体显性)(RP1);

(t)SEQ ID NO:57视网膜色素变性GTP酶调节因子相互作用蛋白1(RPGRIP1);

(u)SEQ ID NO:59PRP8;

(v)SEQ ID NO:61中心体蛋白290kDa(CEP290);

(w)SEQ ID NO:63IMP(肌苷5'-单磷酸)脱氢酶1(IMPDH1),转录变体1;

(x)SEQ ID NO:65芳香烃受体相互作用蛋白样蛋白1(AIPL1),转录变体1;

(y)SEQ ID NO:66视黄醇脱氢酶12(全反式/9-顺式/11-顺式)(RDH12);

(z)SEQ ID NO:67莱伯氏先天性黑蒙症5(LCA5),转录变体1;以及

(aa)示例性OPN1LW/OPN1MW2多晶型物(与OPN1LW(L视蛋白)多肽序列相比);数字左侧的氨基酸是L视蛋白序列中存在的残基;数字是L视蛋白中的残基编号,并且数字右侧的残基是来自L视蛋白的变体:

(i)Thr65Ile;

(ii)Ile111Val;

(iii)Ser116Tyr;

(iv)Leu153Met;

(v)Ile171Val;

(vi)Ala174Val;

(vii)Ile178Val;

(viii)Ser180Ala;

(ix)Ile230Thr;

(x)Ala233Ser;

(xi)Val236Met;

(xii)Ile274Val;

(xiii)Phe275Leu;

(xiv)Tyr277Phe;

(xv)Val279Phe;

(xvi)Thr285Ala;

(xvii)Pro298Ala;以及

(xviii)Tyr309Phe。

16.根据权利要求13至15中任一项所述的核酸表达盒,其中所述编码区包括编码OPN1LW(外显子1-6)或OPN1MW(外显子1-6)的核酸,并且其中:

(a)所述编码区不包含内源性OPN1LW/MW内含子;并且

(b)所述一个或多个内含子包括:

(i)包括在OPN1LW/MW外显子3上游的SEQ ID NO:7的序列的第一内含子,和包括在OPN1LW/MW外显子3下游的SEQ ID NO:8的序列的第二内含子;或者

(ii)包括在OPN1LW/MW外显子3上游的SEQ ID NO:10的序列的第一内含子,和包括在OPN1LW/MW外显子3下游的SEQ ID NO:12的序列的第二内含子。

17.根据权利要求9至10和13至16中任一项所述的核酸表达盒,其中所述编码区包括编码在N端处具有信号序列的多肽的核酸序列。

18.根据权利要求9至10和13至17中任一项所述的核酸表达盒,其中所述编码区包括编码在C端处包括外层细胞膜靶向序列的多肽的核酸序列。

19.根据权利要求18所述的核酸表达盒,其中所述外层细胞膜靶向序列包括VXPX或由VXPX组成,其中X是任何氨基酸。

20.根据权利要求9至10和13至19中任一项所述的核酸表达盒,其中所述编码区包括(i)治疗性基因编码区和(ii)荧光蛋白编码区。

21.根据权利要求20所述的核酸表达盒,其中所述荧光蛋白编码区存在于所述治疗性基因编码区的3'端。

22.根据权利要求20至21中任一项所述的核酸表达盒,其中所述治疗性基因编码区包括编码OPN1LW/MW蛋白的核酸序列,并且所述荧光蛋白编码区包括编码绿色荧光蛋白的核酸序列。

23.根据权利要求22所述的核酸表达盒,其中:

(a)(i)用编码S视蛋白的最后12个氨基酸的TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72)替换OPN1LW/MW编码区的最后27个核苷酸(TCATCTGTGTCCTCGGTATCGCCTGCA(SEQ ID NO:71));并且(ii)在GFP的最后一个氨基酸编码密码子后插入序列TCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID NO:73),所述序列指定所述OPN1LW/MW C端的最后10个氨基酸;或者

(b)紧跟在GFP的最后一个氨基酸编码密码子之后且在GFP终止密码子之前插入S视蛋白的最后12个氨基酸的编码区(TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72))。

24.根据权利要求20至21中任一项所述的核酸表达盒,其中所述治疗性基因编码区包括编码阿柏西普或其功能片段、衍生物或变体的核酸序列;并且所述荧光蛋白编码区包括编码柠檬黄的核酸序列。

25.根据权利要求20至24中任一项所述的核酸表达盒,其中所述编码区包括编码在所述治疗性基因编码区与所述荧光蛋白编码区之间的连接子的核酸序列。

26.根据权利要求25所述的核酸表达盒,其中所述连接子由SEQ ID NO:76的核酸序列编码。

27.根据权利要求8至26中任一项所述的核酸表达盒,其进一步包括:

(f)调节元件,所述调节元件包括SEQ ID NO:15的序列、基本上由所述序列组成或由所述序列组成,其中所述调节元件定位于所述克隆位点或所述编码区的3'端;以及

(g)非翻译区核酸,所述非翻译区核酸包括SEQ ID NO:16、基本上由其组成或由其组成,其中所述非翻译区核酸定位于所述调节元件的3'端。

28.一种核酸表达盒,其包括编码视蛋白多肽的核酸,所述视蛋白多肽与启动子可操作地连接,其中所述编码视蛋白多肽的核酸包括一个或多个内含子,所述一个或多个内含子包括SEQ ID NO:10和/或SEQ ID NO:12的核酸序列、基本上由所述核酸序列组成或由所述核酸序列组成。

29.根据权利要求28所述的核酸表达盒,其中编码OPN1LW/MW的核酸编码第一内含子和第二内含子,所述第一内含子包括在OPN1LW/MW外显子3上游的SEQ ID NO:10的序列,并且所述第二内含子包括在OPN1LW/MW外显子3下游的SEQ ID NO:12的序列。

30.根据权利要求28至29中任一项所述的核酸表达盒,其进一步包括:

(a)基因座控制区(LCR),所述LCR包括SEQ ID NO:2的序列、基本上由所述序列组成或由所述序列组成;以及

(b)启动子,所述启动子包括核酸、基本上由所述核酸组成或由所述核酸组成,所述核酸包括以下序列、基本上由以下序列组成或由以下序列组成:

GGGAGGAG GAGGTCTAAG TCCCAGGCCC

AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG

CCCGGGCTGA TCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGAC

GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC(X) (SEQ ID NO: 1);

其中X不存在,或者选自由以下组成的组:ATAGCCGGTACCATG(SEQ ID NO:104)和GCCGCCACC;

其中所述启动子定位于所述LCR的3'端,并且其中所述启动子定位于所述编码视蛋白

多肽的核酸的5'端。

31.根据权利要求30所述的核酸表达盒,其中所述启动子包括SEQ ID NO:4或SEQ ID NO:5的序列、或由所述序列组成。

32.根据权利要求44至47中任一项所述的核酸表达盒,其中所述编码视蛋白多肽的核酸在其N端处编码信号序列。

33.根据权利要求28至32中任一项所述的核酸表达盒,其中所述编码视蛋白多肽的核酸包括在其C端处编码外层细胞膜靶向序列。

34.根据权利要求33所述的核酸表达盒,其中所述外层细胞膜靶向序列包括VXPX或由VXPX组成,其中X是任何氨基酸。

35.根据权利要求28至34中任一项所述的核酸表达盒,其中编码所述视蛋白多肽或其功能片段、衍生物或变体的核酸编码具有荧光蛋白的融合物。

36.根据权利要求35所述的核酸表达盒,其中荧光蛋白编码区存在于所述治疗性基因编码区的3'端。

37.根据权利要求28至36中任一项所述的核酸表达盒,其中所述编码视蛋白多肽的核酸包括编码OPN1LW/MW蛋白,并且所述荧光蛋白编码区包括编码绿色荧光蛋白的核酸序列。

38.根据权利要求28至37中任一项所述的核酸表达盒,其进一步包括:

调节元件,所述调节元件包括SEQ ID NO:15的序列、基本上由所述序列组成或由所述序列组成,其中所述调节元件定位于所述编码视蛋白多肽的核酸的3'端;以及

非翻译区核酸,所述非翻译区核酸包括SEQ ID NO:16、基本上由其组成或由其组成,其中所述非翻译区核酸定位于所述调节元件的3'端。

39.根据权利要求5至38中任一项所述的核酸表达盒,其中所述盒的5'端和3'端包括反向末端重复序列。

40.根据权利要求39所述的核酸表达盒,其中所述反向末端重复序列是功能性腺相关病毒(AAV)反向末端重复序列。

41.根据权利要求37至38中任一项所述的核酸表达盒,其中5'ITR包括(SEQ ID NO:79)的核酸序列或由所述核酸序列组成,并且3'ITR包括(SEQ ID NO:80)的核酸序列或由所述核酸序列组成。

42.根据权利要求5至41中任一项所述的核酸表达盒,其中所述核酸表达盒的长度小于约5kb。

43.根据权利要求5至42中任一项所述的核酸表达盒,其中所述核酸表达盒包括选自由SEQ ID NO:91-95组成的组的序列或由所述序列组成。

44.一种核酸表达载体,其包括根据权利要求5至43中任一项所述的核酸表达盒。

45.一种重组宿主细胞,其包括根据权利要求44所述的核酸表达载体。

46.一种重组腺相关病毒(rAAV)颗粒,其包括

(a)AAV衣壳蛋白;以及

(b)根据权利要求5至43中任一项所述的核酸表达盒或根据权利要求44所述的核酸表达载体。

47.一种药物组合物,其包括

(a)根据权利要求5至43中任一项所述的核酸表达盒、根据权利要求44所述的核酸表达

载体或根据权利要求46所述的rAAV颗粒;以及

(b)药学上可接受的载剂。

48.一种配制物,其包括经过包装的病毒颗粒,所述经过包装的病毒颗粒包括根据权利要求5至43中任一项所述的核酸表达载体或根据权利要求44所述的核酸表达载体。

49.一种用于在视锥细胞中表达基因产物如蛋白质的方法,其包括使一个或多个视锥细胞与有效量的根据权利要求5至43中任一项所述的核酸表达盒、根据权利要求44所述的核酸表达载体、根据权利要求45所述的重组宿主细胞、根据权利要求46所述的rAAV颗粒、根据权利要求47所述的药物组合物或根据权利要求48所述的配制物接触,其中由编码区编码的所述基因产物如蛋白质在所述一个或多个视锥细胞中以可检测的水平表达。

50.一种用于在需要治疗或预防视锥细胞病症的哺乳动物中治疗或预防视锥细胞病症的方法,其包括向所述哺乳动物的眼睛施用有效量的根据权利要求5至43中任一项所述的核酸表达盒、根据权利要求44所述的核酸表达载体、根据权利要求45所述的重组宿主细胞、根据权利要求46所述的rAAV颗粒、根据权利要求47所述的药物组合物或根据权利要求48所述的配制物,其中编码区包括编码治疗性基因产物的核酸序列。

51.根据权利要求50所述的方法,其中所述视锥细胞病症选择由以下组成的组:黄斑营养不良、色觉障碍或中央黄斑视力障碍。

52.根据权利要求50或51所述的方法,其中所述视锥细胞病症是全色盲、蓝视锥单色视、红绿色盲、红色觉缺陷、绿色觉缺陷、蓝色觉缺陷、黄斑营养不良如斯特格氏黄斑营养不良、视锥营养不良、视锥-视杆营养不良、X连锁视锥营养不良、7型脊髓小脑性共济失调和巴尔得-别德尔氏综合征1、年龄相关性黄斑变性、黄斑毛细管扩张症、视网膜色素变性、糖尿病性视网膜病变、视网膜静脉阻塞、青光眼、索斯比氏眼底营养不良、成人型卵黄样黄斑营养不良、贝斯特氏病、视杆-视锥营养不良、莱伯氏先天性黑蒙症和X连锁视网膜劈裂症、博恩霍姆氏眼病和伴有近视的X连锁视锥功能障碍综合征。

53.根据权利要求50或51所述的方法,其中所述视锥细胞病症是红绿色盲、蓝视锥单色视、博恩霍姆氏眼病、伴有近视的X连锁视锥功能障碍综合征和X连锁视锥营养不良、X连锁视网膜劈裂症、常染色体隐性遗传性视网膜色素变性和年龄相关性黄斑变性(AMD;湿性或干性)。

54.根据权利要求49至53中任一项所述的方法,其中所述方法包括将根据权利要求5至43中任一项所述的核酸表达盒、根据权利要求44所述的核酸表达载体、根据权利要求45所述的重组宿主细胞、根据权利要求46所述的rAAV颗粒、根据权利要求47所述的药物组合物或根据权利要求48所述的配制物通过玻璃体内注射到受试者体内。

55.一种分离的核酸序列,其包括选自由以下组成的组的序列:

(i)SEQ ID NO:10;以及

(ii)SEQ ID NO:12。

56.一种分离的核酸,其包括通式A-B的核酸序列,其中A编码包括SEQ ID NO:21的氨基酸序列、基本上由所述氨基酸序列组成或由所述氨基酸序列组成的多肽,并且其中B编码用于治疗视锥细胞病症的基因产物。

57.根据权利要求56所述的核酸,其中所述用于治疗视锥细胞病症的基因产物是本文所公开的基因产物。

58.根据权利要求56或57所述的核酸，其中B包括序列SEQ ID NO:101或由所述序列组成。

59.一种分离的多肽，其包括SEQ ID NO:103的氨基酸序列、基本上由所述氨基酸序列组成或由所述氨基酸序列组成。

60.根据权利要求15所述的核酸表达盒，其中所述编码区包括编码针对VEGFA的抗体、抗血管内皮生长因子(VEGF)蛋白或其抗体片段的核酸序列。

61.根据权利要求60所述的核酸表达盒，其中所述抗体包括单克隆抗体或其片段。

62.根据权利要求60或61所述的核酸表达盒，其中所述抗体包括贝伐单抗或其片段。

63.根据权利要求60或61所述的核酸表达盒，其中所述抗体包括兰尼单抗或其片段。

用于增强治疗性基因在光感受器中的功能性表达的组合物和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求于2017年12月5日提交的美国临时专利申请序列号62/594811的优先权,所述美国临时专利申请通过引用整体并入本文。

背景技术

[0003] 以显著的视力丧失为症状的各种临床病症(包含蓝视锥单色视、博恩霍尔姆氏眼病(Bornholm eye disease)、伴有近视的X连锁视锥功能障碍综合征和X连锁视锥营养不良的形式)均涉及影响L和M视锥感光色素基因的表达或功能的突变。L和M视锥共同构成人类视网膜中视锥光感受器的约94%,并且最终导致的视觉障碍的严重性取决于特定的基础突变以及表达所述突变的L/M视锥的相对数量。尽管与这些突变相关的表型和疾病进展存在很大的变异性,但患有严重视觉障碍的人却没有或很少有功能性L或M视锥光感受器,并且视觉敏锐度为约20/60到20/200。通常,在所有L/M视锥中表达突变视蛋白的男性中,由于L/M视蛋白突变引起的视力丧失更为严重。许多病状是静止的,但其它病状是进行性的,并且存在与视锥营养不良相关的特定视蛋白突变,其时间进程可预测。目前,没有针对这些病症的疗法;然而,自适应光学成像的结果表明,这些病症中的许多病症与存活的但非功能性的视锥相关,所述视锥可能适用于涉及病毒介导的功能性视蛋白基因转移的治疗。然而,目前尚无用于眼部病症的有效基因疗法的合适构建体和方法。

发明内容

[0004] 在第一方面,本公开提供了分离的核酸,所述分离的核酸包括SEQ ID NO:1的序列,其中X不存在,或者选自由以下组成的组:ATAGCCGGTACC(ATG)(SEQ ID NO:104)和GCCGCCACC。在各个实施例中,

[0005] 所述核酸包括SEQ ID NO:4或SEQ ID NO:5的序列。

[0006] 在第二方面,本公开提供了核酸表达盒,所述核酸表达盒包括:

[0007] (a)基因座控制区(LCR),所述LCR包括SEQ ID NO:2的序列、基本上由所述序列组成或由所述序列组成;以及

[0008] (b)启动子,所述启动子包括根据本公开的第一方面的任何实施例或实施例组合所述的核酸序列、基本上由所述核酸序列组成或由所述核酸序列组成,其中所述启动子定位于所述LCR的3'端。在一个实施例中,所述启动子包括SEQ ID NO:4的序列或由所述序列组成,其中所述盒进一步包括:

[0009] (c)内含子,所述内含子包括SEQ ID NO:6的序列、基本上由所述序列组成或由所述序列组成,其中所述内含子定位于所述启动子的3';以及

[0010] (d)Kozak序列,如包括序列GCCGCCACC、基本上由所述序列组成或由所述序列组成的Kozak序列,其中所述Kozak序列定位于所述内含子的3'端。在各个另外的实施例中,所述盒进一步包括

[0011] (e)克隆位点,所述克隆位点定位于所述Kozak序列的3',或编码区,所述编码区定位于所述Kozak序列的3'端。

[0012] 在所述盒的另一个实施例中,所述启动子包括SEQ ID NO:5的序列、基本上由所述序列组成或由所述序列组成。在此实施例中,所述盒可以进一步包括定位于所述启动子的3'端的克隆位点或定位于所述启动子的3'端的编码区。在存在定位于所述启动子的3'端的编码区的实施例中,所述编码区可以包括多个内含子中的一个。

[0013] 在所述盒包含编码区的所有实施例中,所述编码区可以包括适合在视锥细胞中表达的任何编码区,包含但不限于本文所公开的编码区。在具体实施例中,所述编码区可以编码视蛋白基因或其变体,或者可以包括阿柏西普(Aflibercept)或其功能片段、衍生物或变体。

[0014] 在各个另外的实施例中,所述表达盒可以进一步包括

[0015] (f)调节元件,所述调节元件包括SEQ ID NO:15的序列、基本上由所述序列组成或由所述序列组成,其中所述调节元件定位于所述克隆位点或所述编码区的3'端;以及

[0016] (g)非翻译区核酸,所述非翻译区核酸包括SEQ ID NO:16、基本上由其组成或由其组成,其中所述非翻译区核酸定位于所述调节元件的3'端。

[0017] 在第三方面,本公开提供了核酸表达盒,所述核酸表达盒包括编码视蛋白多肽的核酸,所述视蛋白多肽与启动子可操作地连接,其中所述编码视蛋白多肽的核酸包括一个或多个内含子,所述一个或多个内含子包括SEQ ID NO:10和/或SEQ ID NO:12的核酸序列、基本上由所述核酸序列组成或由所述核酸序列组成。在一个实施例中,所述核酸编码包含第一内含子和第二内含子的OPN1LW/MW,所述第一内含子包括在OPN1LW/MW外显子3上游的SEQ ID NO:10的序列,并且所述第二内含子包括在OPN1LW/MW外显子3下游的SEQ ID NO:12的序列。在另外的实施例中,所述盒可以进一步包括

[0018] (a)基因座控制区(LCR),所述LCR包括SEQ ID NO:2的序列、基本上由所述序列组成或由所述序列组成;以及

[0019] (b)启动子,所述启动子包括本公开的第一方面的任何实施例中的核酸、基本上由所述核酸组成或由所述核酸组成,

[0020] 其中所述启动子定位于所述LCR的3'端,并且其中所述启动子定位于所述编码视蛋白多肽的核酸的5'端。

[0021] 在本公开的任何表达盒的一个实施例中,所述盒的5'端和3'端可以包括反向末端重复序列,包含但不限于功能性腺相关病毒(AAV)反向末端重复序列。在另一个具体实施例中,所述表达盒可以包括选自以下组成的组的序列:SEQ ID NO:91-95。

[0022] 在其它方面,本公开提供了(a)核酸表达载体,其包括根据本文所公开的任何实施例或实施例组合所述的核酸表达盒;(b)重组宿主细胞,其包括根据本公开所述的核酸表达载体;(c)重组腺相关病毒(rAAV)颗粒,其包括(i)AAV衣壳蛋白和(ii)根据本文所公开的任何实施例或实施例组合所述的核酸表达盒或核酸表达载体;(d)药物组合物,其包括(i)根据本文所公开的任何实施例或实施例组合所述的核酸表达盒、核酸表达载体或rAAV颗粒;以及(ii)药学上可接受的载剂;以及(e)配制物,其包括经过包装的病毒颗粒,所述经过包装的病毒颗粒包括根据本文所公开的任何实施例或实施例组合所述的核酸表达盒或核酸表达载体。

[0023] 在另一个方面，本公开提供了用于在视锥细胞中表达如蛋白质等基因产物的方法，所述方法包括使一个或多个视锥细胞与有效量的根据本文所公开的任何实施例或实施例组合所述的核酸表达盒、核酸表达载体、重组宿主细胞、rAAV颗粒、药物组合物或配制物接触，其中由编码区编码的所述基因产物如蛋白质在所述一个或多个视锥细胞中以可检测的水平表达。

[0024] 在另外的方面，本公开提供了用于在需要治疗或预防视锥细胞病症的哺乳动物中治疗或预防视锥细胞病症的方法，所述方法包括向所述哺乳动物的眼睛施用有效量的根据本文所公开的任何实施例或实施例组合所述的核酸表达盒、核酸表达载体、重组宿主细胞、rAAV颗粒、药物组合物或配制物，其中所述编码区包括编码治疗性基因产物的核酸序列。在各个实施例中，所述视锥细胞病症可以选自由以下组成的组：黄斑营养不良、色觉障碍、中央黄斑视力障碍、全色盲、蓝视锥单色视、红绿色盲、红色觉缺陷、绿色觉缺陷、蓝色觉缺陷、黄斑营养不良如斯特格氏黄斑营养不良(Stargardt's macular dystrophy)、视锥营养不良、视锥-视杆营养不良、X连锁视锥营养不良、7型脊髓小脑性共济失调和巴尔得-别德尔氏综合征(Bardet-Biedl syndrome)1，年龄相关性黄斑变性、黄斑毛细管扩张症、视网膜色素变性、糖尿病性视网膜病变、视网膜静脉阻塞、青光眼、索斯比氏眼底营养不良(Sorsby's fundus dystrophy)、成人型卵黄样黄斑营养不良、贝斯特氏病(Best's disease)、视杆-视锥营养不良、莱伯氏先天性黑蒙症、X连锁视网膜劈裂症、博恩霍姆氏眼病、伴有近视的X连锁视锥功能障碍综合征、常染色体隐性遗传性视网膜色素变性和年龄相关性黄斑变性(AMD;湿性或干性)。

[0025] 在其它方面，本公开提供了：(a)一种分离的核酸，其包括序列：(i)SEQ ID NO:10；或(ii)SEQ ID NO:12；(b)一种分离的核酸，其包括通式A-B的核酸序列，其中A编码包括SEQ ID NO:21的氨基酸序列、基本上由所述氨基酸序列组成或由所述氨基酸序列组成的多肽，并且其中B编码用于治疗视锥细胞病症的基因产物；以及(c)一种分离的多肽，其包括SEQ ID NO:103的氨基酸序列、基本上由所述氨基酸序列组成或由所述氨基酸序列组成。

附图说明

[0026] 图1：平整固定的小鼠视网膜的共聚焦显微镜荧光图像。将SEQ ID NO:91包装到HEK293细胞中的AAV2_7m8衣壳中，并通过碘克沙醇密度梯度离心法进行纯化。将含有约5×10^{11}个病毒基因组的三微升病毒溶液注入到小鼠眼睛的玻璃体中。使用纳升注射泵进行玻璃体内注射，同时在显微镜下观察注射。注射后一个月，用过剂量的戊巴比妥(pentobarbital)钠处死小鼠，并收获眼睛用于组织学。此处的图像示出了来自L视蛋白上的GFP标签的荧光。小鼠M视蛋白表达表现出特征性空间表达模式，其中表达集中在上侧视网膜中，而在下侧视网膜中很少表达或没有表达。(A)携带SEQ ID NO:91的AAV2_7m8的空间表达模式视蛋白-GFP遵循小鼠M视蛋白的原生表达模式，这表明SEQ ID NO:91不驱动小鼠S视锥中的表达，所述视锥在下侧小鼠视网膜中占主导地位。(B)示出了由白框勾勒出的A中的区域的放大图，并表明了来自SEQ ID NO:91的视蛋白基因正确定位在被称为外部区段的细胞的专门区室中。

[0027] 图2视网膜电图(ERG)结果，其示出在玻璃体内注射包装在AAV2_7m8衣壳中的SEQ ID NO:92一个月后，人类长波长(L)敏感性视蛋白在小鼠视锥中表达。将SEQ ID NO:92包装

到HEK293细胞中的AAV2_7m8衣壳中，并通过碘克沙醇密度梯度超速离心法进行纯化。使用注射泵将含有约5×10^{11}个病毒基因组的三微升病毒溶液注入到玻璃体中，并用显微镜进行观察。对于ERG，用氯胺酮/甲苯噻嗪麻醉小鼠，并响应于一系列交替的634nm(红色)和535nm(绿色)发光二极管(LED)光脉冲，从放置在眼睛角膜上的电极记录ERG电位。图的Y轴表示与记录的对红色LED响应的振幅相匹配所需的绿色LED的光强度。野生型小鼠不具有长波长敏感性(L)感光色素，并且其内源性M视锥介导了对634nm光的所有敏感性。因此，未经过处理的小鼠的基线值平均为5的相对较低强度值。如此处所示，平均而言，经过处理的动物需要约6倍的绿光(值接近30)才能匹配由红光引起的响应。此处显示的对红光敏感性的急剧增加表明，在通过玻璃体内注射携带SEQ ID NO:92的AAV2-7m8进行处理后，人L-视蛋白的功能性表达水平很高。使用这项基于ERG的测定，先前使用表达盒pR2.1和pMNTC进行玻璃体内处理不会产生任何显著的功能性L-视蛋白表达。

[0028] 图3在玻璃体内注射携带SEQ ID SEQ ID NO:91的AAV2_7m8后1年，在体内拍摄的灵长类动物视网膜黄斑的Retcam图像。使用0.33ml的结核菌素注射器和25号针头，将含有约3×10^{12}个病毒基因组的三十微升病毒溶液注射到用氯胺酮/甲苯噻嗪麻醉的灵长类动物的玻璃体中。将SEQ ID NO:91包装到HEK293细胞中的AAV2_7m8衣壳中，并通过碘克沙醇密度梯度超速离心法进行纯化。(A)经过处理的眼底白光下的图像示出了黄斑区的解剖学标志。左上方可见视神经和视网膜血管的小新月形。右中下部的较暗区域是视网膜黄斑，中央凹(fovea)处在正中(B)。在蓝光下拍摄A中所示的眼底相同区域，以允许检测GFP荧光。中央凹中心的亮白点(图像右中下部)证明，注入到灵长类动物眼玻璃体中的AAV2_7m8中的SEQ ID NO:91产生持久、稳健的表达，所述表达遵循原生L视蛋白的空间表达，后者处于集中于中央凹的视锥光感受器中。

[0029] 图4拯救灵长类视网膜疾病模型的视锥中的L-视蛋白表达。在玻璃体内注射携带SEQ ID NO:91的AAV2_7m8一年后，图3的灵长类动物眼睛中的GFP表达的共聚焦显微镜图像。穿过眼中央凹的横截面证明，单次玻璃体内注射导致GFP在中央黄斑区在近100％的L/M视锥中稳健表达，然而，所述单次玻璃体内注射不会转导S视锥或视杆光感受器，所述处理也不转导任何其它细胞类型，如视网膜中的神经节细胞。图像在中央凹上从右到左居中。白色标记是来自在视锥光感受器中特异性表达的L-视蛋白-GFP融合转基因的原生GFP荧光。灰色标记是用于使视网膜中所有细胞类型的细胞核视觉化的DAPI。所述结果证明了对天然存在的灵长类动物视网膜疾病模型的成功的基因疗法处理。

[0030] 图5用携带与柠檬黄(Citrine)融合的VEGF-TRAP的合成基因的质粒转染的HEK293T细胞分泌VEGF-TRAP-柠檬黄融合蛋白。用X轴上指示的三个质粒之一转染HEK293T细胞。一种质粒含有在人突触蛋白1启动子控制下的柠檬黄基因(AAV2柠檬黄)，另一种质粒含有图7B所示的构建体(VEGF-Trap)，并且第三种质粒也含有图7B所示的构建体，不同之处在于用RS1信号序列替换“sflt信号序列”以产生RS1-VEGF-Trap。转染两天后，收集培养基并通过离心法除去细胞碎片。使用BioRad Glomax发光计测量培养基的荧光强度。每个质粒转染六份重复培养物。测量六个重复的“无转染”对照，以给出背景荧光水平。不应分泌AAV2-Citrine，并且因此培养基不应示出高的表达水平。如在此图中所观察的，预计一些荧光来自裂解的细胞。已经用VEGF-Trap或RS1-VEGF-Trap转染的HEK293T细胞的荧光强度最高，如经过转染的细胞将VEGF-柠檬黄融合蛋白分泌到培养基中所预期的那样。为了使

VEGF-Trap对眼部疾病具有治疗效果，必须从表达VEGF-Trap的细胞中将其分泌出来，并且此实验证明VEGF-Trap构建体(SEQ ID NO:93)是从细胞中分泌的。

[0031] 图6经过修饰的β-珠蛋白内含子被称为pFLARE内含子，当插入OPN1LW cDNA的特定位置处时，所述内含子提供强大的剪接信号，并且可以通过玻璃体内注射携带视蛋白基因的AAV来增加视蛋白基因的表达。OPN1LW外显子3的两个变体(被称为LIAVA和LVAVA)表现出完整的或接近完整的剪接缺陷，其中最终信使RNA(mRNA)不包含外显子3。如图8所展示的，在这里我们已经将pFLARE内含子插入了外显子3的两侧，并通过转染携带图8的构建体区段的质粒进行了剪接测定，所述区段从OPN1LW x1/x2穿过短的WPRE延伸到HEK293细胞中。转染后24小时到48小时，从细胞中分离mRNA，并通过凝胶电泳和对凝胶上观察到的条带的直接测序进行检查。凝胶上的每对通道均示出重复转染的结果。通道1a和1b示出了来自对照构建体的mRNA，所述对照构建体携带外显子3的未表现出剪接缺陷的版本。pFLARE内含子对此构建体中的剪接没有影响，因为观察到的唯一条带对应于包含外显子3的全长mRNA。通道2a和2b显示，pFLARE内含子抑制了OPN1LW外显子3的LIAVA版本的剪接缺陷，所述版本通常表现出完整的剪接缺陷，所述完整的剪接缺陷仅产生缺少外显子3的mRNA。同样，通道3a和3b显示，pFLARE内含子完全抑制了通常针对外显子3的LVAVA版本观察到的剪接缺陷。通道4是示出了未经过剪接的构建体的对照，并且mw是分子量标志物。*表示对应于全长mRNA的条带，并且**表示缺少外显子3的mRNA的条带，所述条带比全长mRNA短169bp。这些观察结果与存在原生内含子时，针对LIAVA和LVAVA OPN1LW外显子变体的剪接观察到的结果相反。这些数据表明，pFLARE内含子携带非常强的剪接信号，并且因此成为用于视蛋白基因疗法的理想内含子。

[0032] 图7具有VEGF-trap(上图)和VEGF-trap柠檬黄融合物的新载体。VEGF-Trap构建体是SEQ ID NO:94，而VEGF-Trap柠檬黄融合物是SEQ ID NO:93。在图5中，标记为“VEGF-Trap”的数据使用SEQ ID NO:93中的构建体，而标记为“RS1-VEGF-Trap”的数据使用SEQ ID NO:93中的构建体，不同之处在于标记为sflt信号序列的区域被来自另一个分泌分子(RS1)的信号序列替换。图9中的数据使用包装在AAV2 7m8衣壳中的B(SEQ ID NO:93)中的构建体。

[0033] 图8具有源自人β珠蛋白基因的内含子(被称为pFLARE内含子)的新型视锥视蛋白载体。此构建体为SEQ ID NO:95。将这个构建体的从OPN1LW x1/x2延伸穿过WPRE短246的区段克隆到质粒(被称为小基因)中，并用于产生图6中的数据。

[0034] 图9从接受玻璃体内注射包装在AAV2_7m8衣壳中的SEQ ID NO:93(图7B)的小鼠的视锥光感受器中分泌VEGF-Trap-柠檬黄。使用HEK293细胞产生病毒，并通过碘克沙醇梯度超速离心法将所述病毒纯化。经过麻醉的小鼠接受玻璃体内注射约8×10^{11}个病毒基因组，体积为3微升。大约7周后，通过过剂量的戊巴比妥处死小鼠，并对眼睛进行加工以用于组织学。整个固定的视网膜(a)示出了来自SEQ ID NO:93的柠檬黄荧光的从上至下的图案，其遵循M-视蛋白表达的正常梯度。柠檬黄荧光可见于通过视锥内部区段的光学部分(b)。11周后，处死小鼠，并对眼睛进行加工以用于组织学。注射后11周，通过视网膜的冰冻切片(c,d)示出了VEGF-Trap-柠檬黄表达。在上侧视网膜看见最高浓度。用花生凝集素(d)染色视锥鞘，并用DAPI(a,b,d)染色细胞核。

具体实施方式

[0035] 所引用的所有参考文献均通过引用整体并入本文。在本申请中,除非另有说明,否则所使用的技术可以在以下若干个众所周知的参考文献中找到:如《分子克隆:实验室手册(Molecular Cloning:A Laboratory Manual)》(Sambrook等人,1989,冷泉港实验室出版社(Cold Spring Harbor Laboratory Press))、《基因表达技术(Gene Expression Technology)》(《酶学方法(Methods in Enzymology)》,第185卷,D.Goeddel编辑,1991学术出版社(Academic Press),加利福尼亚州圣地亚哥)、《酶学方法中》的"蛋白质纯化指南(Guide to Protein Purification)"(编辑者M.P.Deutshcer,(1990),学术出版社公司(Academic Press,Inc.));《PCR方案:方法和应用指南(PCR Protocols:A Guide to Methods and Applications)》(Innis等人,1990学术出版社,加利福尼亚州圣地亚哥)、《动物细胞培养:基本技术指南(Culture of Animal Cells:A Manual of Basic Technique)》,第2版(R.I.Freshney.1987利斯有限公司(Liss,Inc.)纽约州纽约)、《基因转移和表达方案(Gene Transfer and Expression Protocols)》(第109-128页,编辑者E.J.Murray,胡马纳出版社公司(The Humana Press Inc.),新泽西州克利夫顿)以及Ambion 1998目录(德克萨斯州奥斯汀的Ambion公司)。

[0036] 如本文所使用的,除非上下文另有明确指示,否则单数形式"一个/种(a/an)"和"所述(the)"可以包含复数个提及物。除非另有明确说明,否则本文中使用的"和"可与"或"互换使用。

[0037] 如本文所使用的,氨基酸残基如下进行缩写:丙氨酸(Ala;A)、天冬酰胺(Asn;N)、天冬氨酸(Asp;D)、精氨酸(Arg;R)、半胱氨酸(Cys;C)、谷氨酸(Glu;E)、谷氨酰胺(Gln;Q)、甘氨酸(Gly;G),组氨酸(His;H)、异亮氨酸(Ile;I)、亮氨酸(Leu;L)、赖氨酸(Lys;K)、蛋氨酸(Met;M)、苯丙氨酸(Phe;F)、脯氨酸(Pro;P)、丝氨酸(Ser;S)、苏氨酸(Thr;T)、色氨酸(Trp;W)、酪氨酸(Tyr;Y)和缬氨酸(Val;V)。

[0038] 除非上下文另外明确指出,否则可以组合使用本公开的任何方面的所有实施例。

[0039] 在第一方面,本公开提供了分离的核酸,所述分离的核酸包括以下序列、基本上由以下序列组成或由以下序列组成:

[0040] GGGAGGAG GAGGTCTAAG TCCCAGGCCC AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG CCCGGGCTGATCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGAC GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC(X)(SEQ ID NO:1);

[0041] 其中X不存在,或者选自由以下组成的组:ATAGCCGGTACC(ATG)(SEQ ID NO:104,其中括号中的残基是任选的)和GCCGCCACC。当X为SEQ ID NO:104时,全长序列为SEQ ID NO:3;当X为GCCGCCACC时,全长序列为SEQ ID NO:5。

[0042] 本公开的此方面的分离的核酸是发明人在本文中已示出的经过修饰的视蛋白启动子序列,所述经过修饰的视蛋白启动子序列对于表达所关注的基因产物并将所述基因产物递送至刀视锥细胞特别有用。例如,高亮显示的T>C变化(增大的字体大小)被示出为显著改善了转录。"X"是经过修饰的Kozak序列,所述经过修饰的Kozak序列可以在改善来自包含本公开的分离的核酸的表达盒的基因产物的翻译方面提供另外的益处。

[0043] 在一个实施例中，核酸包括序列4(SEQ ID NO:4)的序列、基本上由所述序列组成或由所述序列组成，当在内含子可以存在于分离的核酸与有待表达的编码区之间的表达盒中使用时，所述序列可能特别有用。

[0044] 序列4:L启动子版本1.0(V1.0)(SEQ ID NO:4)

[0045]
```
                                    GGGAGGAG GAGGTCTAAG TCCCAGGCCC
AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG
CCCGGGCTGATCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGAC
GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC ATA GCCGGTACC
```

[0046] 在另一个实施例中，核酸包括序列5(SEQ ID NO:5)的序列、基本上由所述序列组成或由所述序列组成，当在分离的核酸紧邻有待表达的基因的翻译起始信号的上游存在的表达盒中使用时，所述序列可能特别有用。

[0047] 序列5:L启动子版本2.0(SEQ ID NO:5)

[0048]
```
                                    GGGAGGAG GAGGTCTAAG TCCCAGGCCC
AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG
CCCGGGCTGA TCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGAC
GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC GCCGCCACC
```

[0049] 在序列4(V1.0)中，通过紧邻ATG翻译起始信号的上游插入GGTACC来破坏原生Kozak，所述翻译起始信号是Kozak序列的一部分。这样做是为了防止在原生Kozak处开始转录，当此启动子用于内含子中时，在原生Kozak处开始转录会降低自载体中的下一个元件以来的表达水平。使用此启动子的载体可以包含紧邻蛋白质编码区上游的Kozak序列。

[0050] 在序列5(V2.0)(SEQ ID NO:5)中，原生Kozak被共有Kozak序列替换，不同之处在于所述共有Kozak序列缺少ATG起始密码子。当紧邻此启动子克隆蛋白质编码区时，所述蛋白质编码区的ATG起始密码子加上启动子中的GCCGCCACC产生完整的Kozak序列。

[0051] 在另一个实施例中，核酸不包含形成L视蛋白启动子的一部分的SEQ ID NO:70。因此，在此实施例中，分离的核酸包括截短的L视蛋白突变体启动子：

[0052]
```
GATCCGGTTC CAGGCCTCGG CCCTAAATAG TCTCCCTGGG CTTTCAAGAG
AACCACATGA GAAAGGAGGA TTCGGGCTCT GAGCAGTTTC ACCACCCACC
CCCCAGTCTG CAAATCCTGA CCCGTGGGTC CACCTGCCCC AAAGGCGGAC
GCAGGACAGT AGAAGGGAAC AGAGAACACA TAAACACAGA GAGGGCCACA
GCGGCTCCCA CAGTCACCGC CACCTTCCTG GCGGGGATGG GTGGGGCGTC
TGAGTTTGGT TCCCAGCAAA TCCCTCTGAG CCGCCCTTGC GGGCTCGCCT
CAGGAGCAGG GGAGCAAGAG GT (SEQ ID NO: 70)
```

[0053] 在第二方面，本公开提供了一种核酸表达盒，所述核酸表达盒包括：

[0054] (a)基因座控制区(LCR)，所述LCR包括序列2(SEQ ID NO:2)的序列、基本上由所述序列组成或由所述序列组成，以及

[0055] (b)启动子，所述启动子包括本公开的第一方面的任何实施例中的分离的核酸序

列、基本上由所述分离的核酸序列组成或由所述分离的核酸序列组成，其中所述启动子定位于所述LCR的3'端。

[0056] 序列2：通过缺失LCR(上文)的前325个bp并缺失3'端的G残基来产生截短的LCR

GGAGG CTGAGGGGTG GGGAAAGGGC

ATGGGTGTTT CATGAGGACA GAGCTTCCGT TTCATGCAAT GAAAAGAGTT

TGGAGACGGA TGGTGGTGAC TGGACTATAC ACTTACACAC GGTAGCGATG

GTACACTTTG TATTATGTAT ATTTTACCAC GATCTTTTTA AAGTGTCAAA

GGCAAATGGC CAAATGGTTC CTTGTCCTAT AGCTGTAGCA GCCATCGGCT

[0057]

GTTAGTGACA AAGCCCCTGA GTCAAGATGA CAGCAGCCCC CATAACTCCT

AATCGGCTCT CCCGCGTGGA GTCATTTAGG AGTAGTCGCA TTAGAGACAA

GTCCAACATC TAATCTTCCA CCCTGGCCAG GGCCCCAGCT GGCAGCGAGG

GTGGGAGACT CCGGGCAGAG CAGAGGGCGC TGACATTGGG GCCCGGCCTG

GCTTGGGTCC CTCTGGCCTT TCCCCAGGGG CCCTCTTTCC TTGGGGCTTT

CTTGGGCCGC CACTGCTCCC GCTCCTCTCC CCCCATCCCA CCCCCTCACC

CCCTCGTTCT TCATATCCTT CTCTAGTGCT CCCTCCACTT TCATCCACCC

TTCTGCAAGA GTGTGGGACC ACAAATGAGT TTTCACCTGG CCTGGGGACA

CACGTGCCCC CACAGGTGCT GAGTGACTTT CTAGGACAGT AATCTGCTTT

AGGCTAAAAT GGGACTTGAT CTTCTGTTAG CCCTAATCAT CAATTAGCAG

AGCCGGTGAA GGTGCAGAAC CTACCGCCTT TCCAGGCCTC CTCCCACCTC

TGCCACCTCC ACTCTCCTTC CTGGGATGTG GGGGCTGGCA CACGTGTGGC

CCAGGGCATT GGTGGGATTG CACTGAGCTG GGTCATTAGC GTAATCCTGG

ACAAGGGCAG ACAGGGCGAG CGGAGGGCCA GCTCCGGGGC TCAGGCAAGG

[0058]

CTGGGGGCTT CCCCCAGACA CCCCACTCCT CCTCTGCTGG ACCCCCACTT

CATAGGGCAC TTCGTGTTCT CAAAGGGCTT CCAAATAGCA TGGTGGCCTT

GGATGCCCAG GGAAGCCTCA GAGTTGCTTA TCTCCCTCTA GACAGAAGGG

GAATCTCGGT CAAGAGGGAG AGGTCGCCCT GTTCAAGGCC ACCCAGCCAG

CTCATGGCGG TAATGGGACA AGGCTGGCCA GCCATCCCAC CCTCAGAAGG

GACCCGGTGG GGCAGGTGAT CTCAGAGGAG GCTCACTTCT GGGTCTCACA

TTCTT (SEQ ID NO: 2)

[0059] 发明人已经表明，本公开的此方面的核酸表达盒可以驱动视锥细胞中经过编码的基因的表达显著改善。

[0060] 序列2(SEQ ID NO:2)的LCR是L/M最小视蛋白增强子的截短版本，其中LCR的前325个bp和3'端的G残基被缺失。发明人在本文中已经表明，与本领域中已报道的相反，序列2

(SEQ ID NO:2)的LCR不促进灵长类动物S视锥中的基因的表达。因此,序列2(SEQ ID NO:2)的LCR是有用的,例如在旨在治疗红绿色视觉缺陷的载体中是有用的,在所述载体中,S视锥中的表达是不期望的。

[0061] 如本文所使用的,“表达盒”是包括彼此可操作地连接的两个或更多个多核苷酸序列(例如启动子、LCR等)的多核苷酸。同样,“用于在视锥细胞中表达基因产物的表达盒”是两个或更多个多核苷酸序列(例如启动子、LCR等)的组合,所述多核苷酸序列促进视锥细胞中基因产物的表达。

[0062] “视锥细胞”,也可以被称为“视锥光感受器”或“视锥”,是眼睛的视网膜中在相对明亮的光线下发挥最佳作用的光感受器细胞的亚型。视锥对特定波长的光敏感,并且因此支持颜色感知。另外,视锥比视杆光感受器对刺激的反应更快,感知的细节比视杆更精细,并且比视杆对图像的变化感知更快,并且因此,视锥支持用于视觉细节至关重要的活动(如阅读和驾驶)的高敏锐视力。通过视网膜外区段的圆锥样形状,在视网膜的横截面中可容易地识别视锥。通过其在视网膜中的位置也可容易地识别视锥,存在于1.5mm凹穴处的视锥的密度最高,所述凹穴定位于视网膜黄斑的中心,被称为“中央凹”或“中央窝”。

[0063] 在表达盒的一个实施例中,启动子包括序列4(SEQ ID NO:4)或5(SEQ ID NO:5)的序列,或由所述序列组成。在另一个实施例中,启动子包括序列4(SEQ ID NO:4)的序列或由所述序列组成,其中所述盒进一步包括内含子,所述内含子包括序列6(SEQ ID NO:6)中的序列、基本上由所述序列组成或由所述序列组成,其中所述内含子定位于所述启动子的3’端。

[0064] 序列6:SV40微小内含子(SEQ ID NO:6)

[0065] CTAGAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCC

[0066] 当将序列4(SEQ ID NO:4)用作启动子时,可以修饰表达盒中使用的蛋白质编码区,使其在5’端包含完整的Kozak序列。当将序列5(SEQ ID NO:5)用作启动子时,通过将蛋白质编码区(包含起始密码子和内含子)3’克隆到启动子上,可以产生完整的Kozak。

[0067] 根据本文的教导,启动子中的任何启动子都可以被使用者认为适当地使用。在一个实施例中,何时使用序列4(SEQ ID NO:4)与序列5(SEQ ID NO:5)之间的区别基于表达盒中元件的顺序。序列4(SEQ ID NO:4)被设计成使得克隆到启动子的3’端的下一个元件是内含子。序列5(SEQ ID NO:5)被设计成使得克隆到启动子的3’端的下一个元件是蛋白质编码区。在后一种情况下,蛋白质编码区可以包含一个或多个内含子。在一个实施例中,凭经验鉴别内含子的序列,以验证内含子被正确剪接。

[0068] 在另一个实施例中,表达盒可以进一步包括定位于Kozak序列3’端的克隆位点。在此实施例中,内含子可以定位于启动子与克隆位点之间。在这些实施例中的任何实施例中,表达盒可以用作可插入所关注的编码区(如编码多肽或其功能片段、衍生物或变体的编码区)的克隆载体,所述多肽或其功能片段、衍生物或变体可以用于治疗视锥细胞病症。此类多肽的示例性实施例在下文中讨论。

[0069] 在另外的实施例中,表达盒可以进一步包括定位于Kozak序列的3’处并与启动子和LCR(例如,截短的LCR)可操作地连接的编码区。如本文所使用的,术语“可操作地连接”是

指两个或更多个遗传元件的并置,其中所述元件处于允许其以预期方式操作的关系中。例如,如果启动子有助于启动编码区的转录,则所述启动子可操作地与编码区连接。只要保持这种功能关系,在启动子与编码区之间就可能存在间插残基。在一个实施例中,编码区可以编码可用于治疗视锥细胞病症的蛋白质。此类蛋白质的示例性实施例在下文中讨论。在另一个实施例中,编码区不包含任何内含子。在此实施例中,盒包含定位于内含子与编码区之间的启动子。

[0070] 如本文所使用的,术语“基因”或“编码区”是指编码基因产物的体外或体内核苷酸序列。在一些情况下,基因由编码区组成或基本上由编码区组成,所述编码区即编码基因产物的序列。在其它情况下,基因包括另外的非编码序列。例如,基因可以包含或可以不包含在编码区域之前和之后的区域,例如5’非翻译序列(5’UTR)或“前导”序列和3’UTR或“尾”序列、以及各个编码区段(外显子)之间的间插序列(内含子)。

[0071] 如本文所使用的,术语“基因产物”是指多核苷酸序列的期望的表达产物,所述多核苷酸序列包含但不限于多肽、肽、蛋白质或包含短干扰RNA(siRNA)、miRNA或小发夹RNA(shRNA)的干扰RNA。

[0072] 在本公开的核酸表达盒的另一个实施例中,启动子包括序列5(SEQ ID NO:5)的序列或由所述序列组成。在一个此类实施例中,表达盒可以进一步包括定位于启动子的3’端的克隆位点。在此实施例中,表达盒可以用作可插入所关注的编码区(如编码蛋白质的编码区)的克隆载体,所述蛋白质可以用于治疗视锥细胞病症。此类蛋白质的示例性实施例在下文中讨论。在另一个实施例中,盒进一步包括定位于启动子的3’端的编码区。在此实施例中,编码区可以包含一个或多个内含子。一个或多个内含子的位置可以由本领域技术人员根据本文的教导和所关注的编码区确定。

[0073] 对于本公开的所有实施例,编码区可以是任何多核苷酸序列,例如编码基因产物(例如基于多肽或RNA的治疗剂(siRNA、反义、核酶、shRNA等))的基因或cDNA。基因产物可以内在地作用于视锥细胞,或者其可以外在地起作用,例如,其可能是分泌性的。例如,当基因产物是治疗性基因时,编码区可以是编码所期望的基因产物或其功能片段或变体的任何基因,所期望的基因产物或其功能片段或变体可以用作治视锥细胞病症的治疗剂或用作用于以其它方式增强视觉的手段,包含但不限于促进四色颜色视觉。

[0074] 在一个实施例中,编码区包括序列(SEQ ID NO:13)或序列14(SEQ ID NO:14),或由所述序列13或序列14组成。

[0075] 序列13(SEQ ID NO:13是全长构建体):具有外显子的OPN1LW mRNA序列,并且翻译起始密码子(ATG)用黑体加下划线标出。在外显子3内,有8个多态性位置,由带下划线的稍大字体表示。还突出了外显子2、4和5中引起OPN1MW基因产生的其它多态性。

[0076] 外显子1和2

[0077]
ATGGCCCAGCAGTGGAGCCTCCAAAGGCTCGCAGGCCGCCATCCGCAGGACAGC
TATGAGGACAGCACCCAGTCCAGCATCTTCACCTACACCAACAGCAACTCCACCAGA
GGCCCCTTCGAAGGCCCGAATTACCACATCGCTCCCAGATGGGTGTACCACCTCACC
AGTGTCTGGATGATCTTTGTGGTCA**C**TGCATCCGTCTTCACAAATGGGCTTGTGCTG

[0078] GCGGCCACCATGAAGTTCAAGAAGCTGCGCCACCCGCTGAACTGGATCCTGGTGAAC CTGGCGGTCGCTGACCTAGCAGAGACCGTCATCGCCAGCACTATCAGCATTGTGAAC CAGGTCTCTGGCTACTTCGTGCTGGGCCACCCTATGTGTGTCCTGGAGGGCTACACC GTCTCCCTGTGTG (SEQ ID NO: 9)

[0079] 外显子3

[0080] GGATCACAGGTCTCTGGTCTCTGGCCATCATTTCCTGGGAGAGGTGGCTGGTGGTGTGCAAGCCCTTT GGCAATGTGAGATTTGATGCCAAGCTGGCCATCGTGGGCATTGCCTTCTCCTGGATCTGGTCTGCTGTGTGGACAG CCCCGCCCATCTTTGGTTGGAGCAG(SEQ ID NO:11)

[0081] 外显子4、5和6

[0082] GTACTGGCCCCACGGCCTGAAGACTTCATGCGGCCCAGACGTGTTCAGCGGCAG CTCGTACCCCGGGGTGCAGTCTTACATGATTGTCCTCATGGTCACCTGCTGCATCAT CCCACTCGCTATCATCATGCTCTGCTACCTCCAAGTGTGGCTGGCCATCCGAGCGGT GGCAAAGCAGCAGAAAGAGTCTGAATCCACCCAGAAGGCAGAGAAGGAAGTGACGCG CATGGTGGTGGTGATGATCTTTGCGTACTGCGTCTGCTGGGGACCCTACACCTTCTT CGCATGCTTTGCTGCTGCCAACCCTGGTTACGCCTTCCACCCTTTGATGGCTGCCCT GCCGGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCGTTATCTATGTCTTTAT GAACCGGCAGTTTCGAAACTGCATCTTGCAGCTTTTCGGGAAGAAGGTTGACGATGG CTCTGAACTCTCCAGCGCCTCCAAAACGGAGGTCTCATCTGTGTCCTCGGTATCGCC TGCA(**TGA**) (**SEQ ID NO: 77**),

[0083] 其中括号中的“TGA”可以是TGA或任何合适的终止密码子,或在与荧光蛋白或其它蛋白的编码区的5’端融合时不存在。

[0084] 序列14(SEQ ID NO:14):经过修饰的阿柏西普,其中阿柏西普的前78个核苷酸被黑体、带下划线的文字突出显示的核苷酸取代。突出显示的序列来自RS1基因,并且编码负责由光感受器分泌视黄酰胺的信号肽。

[0085] ***ATGTCACGCAAGATAGAAGGCTTTTTGTTATTACTTCTCTTTGGCTATGAAGCC ACATTGGGATTATCGTCT***

[0086] AGTGATACCGGTAGACCTTTCGTAGAGATGTACAGTGAAATCCCCGAAATTATA

[0087] CACATGACTGAAGGAAGGGAGCTCGTCATTCCCTGCCGGGTTACGTCACCTAACATC
ACTGTTACTTTAAAAAAGTTTCCACTTGACACTTTGATCCCTGATGGAAAACGCATA
ATCTGGGACAGTAGAAAGGGCTTCATCATATCAAATGCAACGTACAAAGAAATAGGG
CTTCTGACCTGTGAAGCAACAGTCAATGGGCATTTGTATAAGACAAACTATCTCACA
CATCGACAAACCAATACAATCATAGATGTGGTTCTGAGTCCGTCTCATGGAATTGAA
CTATCTGTTGGAGAAAAGCTTGTCTTAAATTGTACAGCAAGAACTGAACTAAATGTG
GGGATTGACTTCAACTGGGAATACCCTTCTTCGAAGCATCAGCATAAGAAACTTGTA
AACCGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAGAAATTTTTGAGCACCTTA
ACTATAGATGGTGTAACCCGGAGTGACCAAGGATTGTACACCTGTGCAGCATCCAGT
GGGCTGATGACCAAGAAGAACAGCACATTTGTCAGGGTCCATGAAAAGGACAAAACT
CACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTC
TTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGC
GTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGAC
GGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACG
TACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCC
AAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGAT
GAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGC
GACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACG
CCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGAC
AAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTG
CACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAA(TAG),

[0088] 其中括号中的“TGA”可以是TGA或任何合适的终止密码子,或在与荧光蛋白或其它蛋白的编码区的5’端融合时不存在。

[0089] 在其它实施例中,编码区包括编码选自由以下所列出的多肽组成的组的多肽的核酸序列:

[0090] • OPN1LW(例如,OPN1LW的外显子1-6)、OPN1MW(例如,OPN1MW的外显子1-6);

[0091] • 阿柏西普(SEQ ID NO:115是全长构建体)、分为4部分的Vegf trap(sflt信号序列、sflt结构域2、VEGFR2结构域3、IgG1fc)或其功能片段、衍生物或变体:

[0092] Sflt信号序列:

[0093] ATGGTCAGCTACTGGGACACCGGGGTCCTGCTGTGCGCGCTGCTCAGCTGTCTGCTTCTCACAGGATC TAGTTCCGGA(SEQ ID NO:78)

[0094] Sflt结构域2

[0095] AGTGATACCGGTAGACCTTTCGTAGAGATGTACAGTGAAATCCCCGAAATTATACACATGACTGAAGG AAGGGAGCTCGTCATTCCCTGCCGGGTTACGTCACCTAACATCACTGTTACTTTAAAAAAGTTTCCACTTGACACT

TTGATCCCTGATGGAAAACGCATAATCTGGGACAGTAGAAAGGGCTTCATCATATCAAATGCAACGTACAAAGAAA TAGGGCTTCTGACCTGTGAAGCAACAGTCAATGGGCATTTGTATAAGACAAACTATCTCACACATCGACAAACCAA TACAATCATAGATGTG(SEQ ID NO:17)

[0096] VEGFR2结构域3

[0097] GTTCTGAGTCCGTCTCATGGAATTGAACTATCTGTTGGAGAAAAGCTTGTCTTAAATTGTACAGCAAG AACTGAACTAAATGTGGGGATTGACTTCAACTGGGAATACCCTTCTTCGAAGCATCAGCATAAGAAACTTGTAAAC CGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAGAAATTTTTGAGCACCTTAACTATAGATGGTGTAACCCGGA GTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCTGATGACCAAGAAGAACAGCACATTTGTCAGGGTCCA TGAAAAG(SEQ ID NO:18)

[0098] IgG1 fc

[0099] GACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTT CCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCAC GAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGG AGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTA CAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGA GAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCA AAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCC TCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGG AACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG GTAAA(SEQ ID NO:19)；

[0100] •经过修饰的阿柏西普(SEQ ID NO:20是全长构建体)，经过修饰的阿柏西普的AA序列(示出了不同部分的起源)或其功能片段、衍生物或变体：

[0101] MSRKIEGFLLLLLFGYEATLGLSS(SEQ ID NO:21)(RS1信号序列)

[0102] SDTGR(SEQ ID NO:22)(sFLT1先前结构域2)

[0103] PFVEMYSEIPEIIHMTEGRELVIPCRVTSPNITVTLKKFPLDTLIPDGKRIIWDSRKGFIISNATYKE IGLLTCEATVNGHLYKTNYLTHRQTNTIIDV(SEQ ID NO:23)(来自Sflt1的IgG样结构域

[0104] VLSPSHGIELSVGEKLVLNCTARTELNVGIDFNWEYPSSKHQHKKLVNRDLKTQSGSEMKKFLSTLTI DGVTRSDQGLYTCAASSGLMTKKNSTFVRVHEK(SEQ ID NO:24)(VEGFR2结构域3)

[0105] DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVS LTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS LSLSPGK(SEQ ID NO:25)(IgG1 FC)；

[0106] •RS1(视网膜劈裂蛋白前体)；

[0107] >sp|O15537|XLRS1_HUMAN视网膜劈裂蛋白OS＝智人GN＝RS1 PE＝1SV＝2

[0108] MSRKIEGFLLLLLFGYEATLGLSSTEDEGEDPWYQKACKCDCQGGPNALWSAGATSLDCIPECPYHKP LGFESGEVTPDQITCSNPEQYVGWYSSWTANKARLNSQGFGCAWLSKFQDSSQWLQIDLKEIKVISGILTQGRCDI DEWMTKYSVQYRTDERLNWIYYKDQTGNNRVFYGNSDRTSTVQNLLRPPI ISRFIRLI PLGWHVRIAIRMELLECVSKCA(SEQ ID NO:26)；

[0109] •针对VEGFA的抗体(如单克隆抗体)、抗血管内皮生长因子(VEGF)蛋白或其抗体

片段(如兰尼单抗和/或贝伐单抗)；

[0110] >兰尼单抗轻链

[0111] DIQLTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSG TDFTLTISSLQPEDFATYYCQQYSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC(SEQ ID NO:27)

[0112] >兰尼单抗重链

[0113] EVQLVESGGGLVQPGGSLRLSCAASGYDFTHYGMNWVRQAPGKGLEWVGWINTYTGEPTYAADFKRRF TFSLDTSKSTAYLQMNSLRAEDTAVYYCAKYPYYYGTSHWYFDVWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGG TAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKK VEPKSCDKTHL(SEQ ID NO:28)

[0114] >“贝伐单抗轻链”

[0115] DIQMTQSPSSLSASVGDRVTITCSASQDISNYLNWYQQKPGKAPKVLIYFTSSLHSGVPSRFSGSGSG TDFTLTISSLQPEDFATYYCQQYSTVPWTFGQGTKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC(SEQ ID NO:29)

[0116] >“贝伐单抗重链”

[0117] EVQLVESGGGLVQPGGSLRLSCAASGYTFTNYGMNWVRQAPGKGLEWVGWINTYTGEPTYAADFKRRF TFSLDTSKSTAYLQMNSLRAEDTAVYYCAKYPHYYGSSHWYFDVWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGG TAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKK VEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTK PREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLT CLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLS LSPGK(SEQ ID NO:30)；

[0118] •色素上皮衍生因子(PEDF)

MQALVLLLCI GALLGHSSCQ NPASPPEEGS PDPDSTGALV EEEDPFFKVP

VNKLAAAVSN FGYDLYRVRS STSPTTNVLL SPLSVATALS ALSLGAEQRT

ESIIHRALYY DLISSPDIHG TYKELLDTVT APQKNLKSAS RIVFEKKLRI

KSSFVAPLEK SYGTRPRVLT GNPRLDLQEI NNWVQAQMKG KLARSTKEIP

[0119] DEISILLLGV AHFKGQWVTK FDSRKTSLED FYLDEERTVR VPMMSDPKAV

LRYGLDSDLS CKIAQLPLTG SMSIIFFLPL KVTQNLTLIE ESLTSEFIHD

IDRELKTVQA VLTVPKLKLS YEGEVTKSLQ EMKLQSLFDS PDFSKITGKP

IKLTQVEHRA GFEWNEDGAG TTPSPGLQPA HLTFPLDYHL NQPFIFVLRD

TDTGALLFIG KILDPRGP （SEQ ID NO: 31）

[0120] •可溶性fms样酪氨酸激酶-1(FLT1),例如,同种型sflt1；

[0121] 同种型sflt1

[0122] sp|P17948-2|VGFR1_HUMAN血管内皮生长因子受体1的同种型2OS＝智人GN＝FLT1

[0123] MVSYWDTGVLLCALLSCLLLTGSSSGSKLKDPELSLKGTQHIMQAGQTLHLQCRGEAAHKWSLPEMVS KESERLSITKSACGRNGKQFCSTLTLNTAQANHTGFYSCKYLAVPTSKKKETESAIYIFISDTGRPFVEMYSEIPE IIHMTEGRELVIPCRVTSPNITVTLKKFPLDTLIPDGKRIIWDSRKGFIISNATYKEIGLLTCEATVNGHLYKTNY LTHRQTNTIIDVQISTPRPVKLLRGHTLVLNCTATTPLNTRVQMTWSYPDEKNKRASVRRRIDQSNSHANIFYSVL TIDKMQNKDKGLYTCRVRSGPSFKSVNTSVHIYDKAFITVKHRKQQVLETVAGKRSYRLSMKVKAFPSPEVVWLKD GLPATEKSARYLTRGYSLIIKDVTEEDAGNYTILLSIKQSNVFKNLTATLIVNVKPQIYEKAVSSFPDPALYPLGS RQILTCTAYGIPQPTIKWFWHPCNHNHSEARCDFCSNNEESFILDADSNMGNRIESITQRMAIIEGKNKMASTLVV ADSRISGIYICIASNKVGTVGRNISFYITDVPNGFHVNLEKMPTEGEDLKLSCTVNKFLYRDVTWILLRTVNNRTM HYSISKQKMAITKEHSITLNLTIMNVSLQDSGTYACRARNVYTGEEILQKKEITIRGEHCNKKAVFSRISKFKSTR NDCTTQSNVKH(SEQ ID NO:32)；

[0124] • CD59；

[0125] >sp|P13987|CD59_HUMAN CD59糖蛋白OS＝智人GN＝CD59 PE＝1SV＝1

[0126] MGIQGGSVLFGLLLVLAVFCHSGHSLQCYNCPNPTADCKTAVNCSSDFDACLITKAGLQVYNKCWKFE HCNFNDVTTRLRENELTYYCCKKDLCNFNEQLENGGTSLSEKTVLLLVTPFLAAAWSLHP(SEQ ID NO:33)

[0127] • (CNGA3)长度＝698(人环核苷酸门控离子通道亚基A;版本1)；或者

[0128] METRGLADSGQGSFTGQGIARFGRIQKKSQPEKVVRAASRGRPLIGWTQWCAEDGGDESEMALAGSPG CSSGPQGRLSRLIFLLRRWAARHVHHQDQGPDSFPDRFRGAELKEVSSQESNAQANVGSQEPADRGRSAWPLAKCN TNTSNNTEEEKKTKKKDAIVVDPSSNLYYRWLTAIALPVFYNWYLLICRACFDELQSEYLMLWLVLDYSADVLYVL DVLVRARTGFLEQGLMVSDTNRLWQHYKTTTQFKLDVLSLVPTDLAYLKVGTNYPEVRFNRLLKFSRLFEFFDRTE TRTNYPNMFRIGNLVLYILIIIHWNACIYFAISKFIGFGTDSWVYPNISIPEHGRLSRKYIYSLYWSTLTLTTIGE TPPPVKDEEYLFVVVDFLVGVLIFATIVGNVGSMISNMNASRAEFQAKIDSIKQYMQFRKVTKDLETRVIRWFDYL WANKKTVDEKEVLKSLPDKLKAEIAINVHLDTLKKVRIFQDCEAGLLVELVLKLRPTVFSPGDYICKKGDIGKEMY IINEGKLAVVADDGVTQFVVLSDGSYFGEISILNIKGSKSGNRRTANIRSIGYSDLFCLSKDDLMEALTEYPEAKK ALEEKGRQILMKDNLIDEELARAGADPKDLEEKVEQLGSSLDTLQTRFARLLAEYNATQMKMKQRLSQLESQVKGG GDKPLADGEVPGDATKTEDKQQ(SEQ ID NO:34)；

[0129] • 人环核苷酸门控离子通道亚基A；

[0130] >sp|Q16281|CNGA3_HUMAN环核苷酸门控离子通道α-3OS＝智人GN＝CNGA3 PE＝ 1SV＝2(人环核苷酸门控离子通道亚基A;版本2)

[0131] MAKINTQYSHPSRTHLKVKTSDRDLNRAENGLSRAHSSSEETSSVLQPGIAMETRGLADSGQGSFTGQ GIARLSRLIFLLRRWAARHVHHQDQGPDSFPDRFRGAELKEVSSQESNAQANVGSQEPADRGRSAWPLAKCNTNTS NNTEEEKKTKKKDAIVVDPSSNLYYRWLTAIALPVFYNWYLLICRACFDELQSEYLMLWLVLDYSADVLYVLDVLV RARTGFLEQGLMVSDTNRLWQHYKTTTQFKLDVLSLVPTDLAYLKVGTNYPEVRFNRLLKFSRLFEFFDRTETRTN YPNMFRIGNLVLYILIIIHWNACIYFAISKFIGFGTDSWVYPNISIPEHGRLSRKYIYSLYWSTLTLTTIGETPPP VKDEEYLFVVVDFLVGVLIFATIVGNVGSMISNMNASRAEFQAKIDSIKQYMQFRKVTKDLETRVIRWFDYLWANK KTVDEKEVLKSLPDKLKAEIAINVHLDTLKKVRIFQDCEAGLLVELVLKLRPTVFSPGDYICKKGDIGKEMYIINE GKLAVVADDGVTQFVVLSDGSYFGEISILNIKGSKSGNRRTANIRSIGYSDLFCLSKDDLMEALTEYPEAKKALEE KGRQILMKDNLIDEELARAGADPKDLEEKVEQLGSSLDTLQTRFARLLAEYNATQMKMKQRLSQLESQVKGGGDKP LADGEVPGDATKTEDKQQ(SEQ ID NO:35)。

[0132] 在其它实施例中,编码区包括编码选自由以下所列出的多肽组成的组的多肽的核

酸序列：

[0133] (a)SEQ ID NO:36智人视蛋白1(视锥色素),对短波敏感(OPN1SW)

[0134] (b)SEQ ID NO:37智人视蛋白1(视锥色素),对中波敏感(OPN1MW)

[0135] (c)SEQ ID NO:38智人视蛋白1(视锥色素),对长波敏感(OPN1LW)

[0136] (d)SEQ ID NO:39ATP结合盒视网膜基因(ABCR)基因(NM_000350)

[0137] (e)SEQ ID NO:40视网膜色素上皮特异性65kD蛋白基因(RPE65)(NM_000329)

[0138] (f)SEQ ID NO:41视网膜结合蛋白1基因(RLBP1)(NM_000326)

[0139] (g)SEQ ID NO:42(外周蛋白/视网膜变性慢基因,(NM_000322)

[0140] (h)SEQ ID NO:43抑制蛋白(SAG)(NM_000541)

[0141] (i)SEQ ID NO:44α-转导蛋白(GNAT1)(NM_000172)

[0142] (j)SEQ ID NO:45鸟苷酸环化酶活化蛋白1A(GUCA1A)(NP_000400.2)

[0143] (k)SEQ ID NO:46视网膜特异性鸟苷酸环化酶(GUCY2D)(NP_000171.1)

[0144] (l)SEQ ID NO:47和48视锥环核苷酸门控阳离子通道的α亚基(CNGA3)(NP_001073347.1或NP_001289.1)

[0145] (m)SEQ ID NO:49人视锥转导蛋白α亚基(不完全性全色盲)；

[0146] (n)SEQ ID NO:50视锥cGMP特异性3',5'-环磷酸二酯酶亚基α',蛋白质(4型视锥营养不良)；

[0147] (o)SEQ ID NO:51视网膜视锥视紫红质敏感性cGMP 3',5'-环磷酸二酯酶亚基γ，蛋白质(3A型视网膜视锥营养不良)；

[0148] (p)SEQ ID NO:52视锥视杆同源盒,蛋白质(视锥-视杆营养不良)；

[0149] (q)SEQ ID NO:53视锥光感受器环核苷酸门控通道β亚基,蛋白质(全色盲)；

[0150] (r)SEQ ID NO:54视锥光感受器cGMP门控阳离子通道β亚基,蛋白质(全色盲,例如见于平格拉普环礁岛民(Pingelapese Islanders))；

[0151] (s)SEQ ID NO:55(SEQ ID NO:56)视网膜色素变性1(常染色体显性)(RP1)；

[0152] (t)SEQ ID NO:57(SEQ ID NO:58)视网膜色素变性GTP酶调节因子相互作用蛋白1(RPGRIP1)；

[0153] (u)SEQ ID NO:59(SEQ ID NO:60)PRP8；

[0154] (v)SEQ ID NO:61(SEQ ID NO:62)中心体蛋白290kDa(CEP290)；

[0155] (w)SEQ ID NO:63(SEQ ID NO:64)IMP(肌苷5'-单磷酸)脱氢酶1(IMPDH1),转录变体1；

[0156] (x)SEQ ID NO:65芳香烃受体相互作用蛋白样蛋白1(AIPL1),转录变体1；

[0157] (y)SEQ ID NO:66视黄醇脱氢酶12(全反式/9-顺式/11-顺式)(RDH12)；

[0158] (z)SEQ ID NO:67莱伯氏先天性黑蒙症5(LCA5),转录变体1;以及

[0159] (aa)示例性OPN1LW/OPN1MW2多晶型物(与OPN1LW(L视蛋白)多肽序列相比);数字左侧的氨基酸是L视蛋白序列中存在的残基;数字是L视蛋白中的残基编号,并且数字右侧的残基是来自L视蛋白的变体。根据这些实施例的多晶型物可以包括下表1中的氨基酸取代中的一种或多种氨基酸取代：

[0160] 表1

[0161] (i)Thr65Ile

[0162] (ii)Ile111Val

[0163] (iii)Ser116Tyr

[0164] (iv)Leu153Met

[0165] (v)Ile171Val

[0166] (vi)Ala174Val

[0167] (vii)Ile178Val

[0168] (viii)Ser180Ala

[0169] (ix)Ile230Thr

[0170] (x)Ala233Ser

[0171] (xi)Val236Met

[0172] (xii)Ile274Val

[0173] (xiii)Phe275Leu

[0174] (xiv)Tyr277Phe

[0175] (xv)Val279Phe

[0176] (xvi)Thr285Ala

[0177] (xvii)Pro298Ala

[0178] (xviii)Tyr309Phe。

[0179] (a)-(c)和(aa)中所述的蛋白质均涉及色觉。示例性多晶型物包含在位置65、116、180、230、233、277、285和309处的影响表达其的视锥细胞中色素的光谱的多晶型物。位置274、275、277、279、285、298和309一起将L视蛋白与M视蛋白区分开。

[0180] (d)-(z)中所述的蛋白质是与以下示例性眼病相关的基因:如视网膜色素变性(多肽“e”-“l”、“s”-“y”)、不完全性全色盲(多肽“m”)、斯特格氏(多肽“d”);莱伯氏先天性黑蒙(多肽“z”);视锥营养不良,如4型视锥营养不良(多肽“n”);视网膜视锥营养不良,例如3A型视网膜视锥营养不良(多肽“o”);视锥-视杆营养不良(多肽“p”);全色盲(多肽“q”);以及例如见于平格拉普环礁岛民的全色盲(多肽“r”)。

[0181] 在各个实施例中,编码区包括编码针对VEGFA的抗体、抗血管内皮生长因子(VEGF)蛋白或其抗体片段的核酸序列。在一个此类实施例中,抗体包括单克隆抗体或其片段。在另一个实施例中,抗体包括贝伐单抗或其片段。在另外的实施例中,抗体包括兰尼单抗或其片段。

[0182] 在一个具体实施例中,编码区包括编码OPN1LW(外显子1-6—SEQ ID NO:38)、OPN1MW(外显子1-6—SEQ ID NO:387)或OPN1LW/MW(SEQ ID NO:68)(下文)或其多态性(如相对于上述编码视蛋白cDNA示出的外显子2、3和4中的多态性)的核酸,并且其中:

[0183] (a)所述编码区不包含内源性OPN1LW/MW内含子;并且

[0184] (b)所述一个或多个内含子包括:

[0185] (i)包括在OPN1LW/MW外显子3上游的序列7(SEQ ID NO:7)的序列的第一内含子,和包括在OPN1LW/MW外显子3下游的序列8(SEQ ID NO:8)的序列的第二内含子(发明人已经表明这些内含子具有非常强的剪接信号);或者

[0186] (ii)包括在OPN1LW/MW外显子3上游的序列10(SEQ ID NO:10)的序列的第一内含子,和包括在OPN1LW/MW外显子3下游的序列12(SEQ ID NO:12)的序列的第二内含子。

[0187] OPN1LW/MW

[0188] MAQQWSLQRLAGRHPQDSYEDSTQSSIFTYTNSNSTRGPFEGPNYHIAPRWVYHLTSVWMIFVV[T/I]ASVFTNGLVLAATMKFKKLRHPLNWILVNLAVADLAETVIASTIS[I/V]VNQV[S/Y]GYFVLGHPMCVLEGYTVSLCGITGLWSLAIISWERW[L/M]VVCKPFGNVRFDAKLAI[V/I]GI[A/V]FSW[I/V]W[S/A]AVWTAPPIFGWSRYWPHGLKTSCGPDVFSGSSYPGVQSYMIVLMVTCCI[I/T]PL[A/S]II[M/V]LCYLQVWLAIRAVAKQQKESESTQKAEKEVTRMVVVM[I/V][F/L]A[Y/F]C[V/F]CWGPY[T/A]FFACFAAANPGY[A/P]FHPLMAALPAYFAKSATIYNPVIYVFMNRQ FRNCILQLFGKKVDDGSELSSASKTEVSSVSSVSPA(SEQ ID NO:68)

[0189] 其中,在OPN1LW与OPN1MW之间以及OPN1LW中和OPN1MW中不同的位置处用括号给出,其中给出了两个可能的氨基酸,并用斜杠(/)分开。

[0190] 序列7:pFLARE内含子1(SEQ ID NO:7)

[0191] GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCACGCGTTTCTGAATTCACTGACTCTCTCTGCCTATTGGTCTATTTTCTCACCCTTAG

[0192] 序列8:pFLARE内含子2(SEQ ID NO:8)

[0193] GTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATGGATCCATAGTCGACCACCATGGTGGCTTAGATCCGGGCATGTGGAGACAGAGAAGACTGTTGAGTTTGTGATAAGCACTGACTCTCTCTGCCTATTGGTCTATTTTCCCTCCCTCAG

[0194] 发明人已经表明,在视蛋白基因中包含这些内含子提供了OPN1LW/MW基因产物的显著增强的转录和翻译。

[0195] 在本文的所有实施例中的一个实施例中,编码区包括编码在N端处具有视锥细胞信号序列的多肽的核酸序列,以允许经过编码的多肽分泌。在一些实施例中,此视锥细胞信号序列可以替换以其它方式存在于多肽中的信号序列。可以使用任何合适的视锥细胞信号序列,包含自然产生视锥细胞分泌的多肽的信号序列,或包含但不限于RS1信号序列(MSRKIEGFLLLLLFGYEATLGLSS(SEQ ID NO:21)的异源信号序列。在另一个实施例中,编码区包括编码在C端处包括外层细胞膜靶向序列的多肽的核酸序列。可以使用任何合适的外层细胞膜靶向序列,包含天然存在于经过编码的多肽中(如在OPN1LW或OPN1MW中)的外层细胞膜靶向序列,或包含但不限于VXPX基序的异源信号序列,其中X是任何氨基酸残基。

[0196] 在另一个实施例中,编码区包括(i)治疗性基因编码区和(ii)荧光蛋白编码区。当期望追踪基因产物的表达/定位时,此实施例特别有用。荧光蛋白编码区可以编码任何合适的荧光蛋白(包含但不限于绿色荧光蛋白、蓝色荧光蛋白、柠檬黄等)。

[0197] 在一个实施例中,荧光蛋白编码区位于治疗性基因编码区的3'端。在此实施例中(即:荧光蛋白编码区位于治疗性基因编码区的3'端),外区段膜靶向序列(当存在时)位于经过编码的融合多肽的C端。

[0198] 在一个特定实施例,治疗性基因编码区包括编码OPN1LW/MW蛋白的核酸序列,并且荧光蛋白编码区包括编码绿色荧光蛋白的核酸序列。在两个示例性的此类实施例中:

[0199] (a)(i)用编码S视蛋白的最后12个氨基酸的TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72)替换OPN1LW/MW编码区的最后27个核苷酸(TCATCTGTGTCCTCGGTATCGCCTGCA(SEQ ID NO:71));并且(ii)在GFP的最后一个氨基酸编码密码子后插入序列TCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID NO:73--),所述序列指定所

述OPN1LW/MW C端的最后10个氨基酸;或者

[0200] (b)紧跟在GFP的最后一个氨基酸编码密码子之后且在GFP终止密码子之前插入S视蛋白的最后12个氨基酸的编码区(TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72))。

[0201] 在另一个具体实施例中,治疗性基因编码区包括编码阿柏西普(SEQ ID NO:115)或经过修饰的阿柏西普(SEQ ID NO:14)或其功能片段、衍生物或变体的核酸序列。在另外的实施例中,治疗性基因编码阿柏西普或经过修饰的阿柏西普与柠檬黄(SEQ ID NO:75)的融合多肽或其功能片段、衍生物或变体。

[0202] 柠檬黄

[0203] ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGATGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG(TAG)(SEQ ID NO:74),

[0204] 其中3'端TAG终止密码子是任选的,因为所述终止密码子可以是任何终止密码子,或者如果柠檬黄定位于融合多肽中的N端,则所述终止密码子可以不存在。

[0205] 柠檬黄多肽:

[0206] MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLMCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK(SEQ ID NO:75)

[0207] 在上述实施例中的任何实施例中,编码区可以编码治疗性基因编码区与荧光蛋白编码区之间的氨基酸连接子。根据本公开,如本领域技术人员将理解的,可以编码任何合适的连接子。在一个非限制性实施例中,连接子由核酸序列

GGAGGTGGAGGTTCTGGTGGAGGAGGTTCC(SEQ ID NO:76)编码。

[0208] 如本领域技术人员将理解的,核酸表达盒可以含有被认为适合于给定目的的其它合适的调节元件。在可以与本文任何其它实施例组合的一个实施例中,核酸表达盒可以进一步包括:

[0209] (f)调节元件,所述调节元件包括序列15(SEQ ID NO:15)的序列、基本上由所述序列组成或由所述序列组成,其中所述调节元件定位于所述克隆位点或所述编码区的3'端;以及

[0210] (g)非翻译区核酸,所述非翻译区核酸包括序列16(SEQ ID NO:16)、基本上由其组成或由其组成,其中所述非翻译区核酸定位于所述调节元件的3'端。

[0211] 序列15:短(246 bp)WPRE(土拨鼠肝炎病毒翻译后响应元件)

[0212] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0213] 序列16:从OPN1LW基因延伸的3'非翻译区。粗体文字是在mRNA中发现的3'UTR,并且包含聚腺苷酸化信号(AAAATAAA)和添加了A'的位点(3'最C端残基)。任选的序列对应于紧邻基因组DNA中的3'UTR下游的序列(X染色体坐标154159,003到154,159,224)

[0214] **GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATCCCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAA**

[0215] (CAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACCTGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACTTTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC)(SEQ ID NO:16),其中括号中的残基是任选的。

[0216] 调节元件和非翻译区核酸用于进一步促进治疗性基因/蛋白质的有效转录和/或翻译。

[0217] 在第三方面,本公开提供了核酸表达盒,所述核酸表达盒包括编码视蛋白多肽(如OPN1LW/MW蛋白)的核酸,所述视蛋白多肽与启动子可操作地连接,其中所述编码视蛋白多肽的核酸包括一个或多个内含子,所述一个或多个内含子包括序列10(SEQ ID NO:10)和/或序列12(SEQ ID NO:12)的核酸序列、基本上由所述核酸序列组成或由所述核酸序列组成。上文讨论了这些内含子。在一个实施例中,编码视蛋白多肽的核酸编码OPN1LW/MW,其中编码OPN1LW/MW的核酸编码第一内含子和第二内含子,所述第一内含子包括在OPN1LW/MW外显子3上游的序列10(SEQ ID NO:10)的序列,并且所述第二内含子包括在OPN1LW/MW外显子3下游的序列12(SEQ ID NO:12)的序列。如上文所描述的,发明人已发现,序列10(SEQ ID NO:10)和/或12(SEQ ID NO:12)的突变内含子可以用于显著改善OPN1LW/MW的表达。在一个实施例中,启动子包括核酸或由所述核酸组成,所述核酸包括以下序列、或基本上由以下序列组成或由以下序列组成:

[0218] GGGAGGAG GAGGTCTAAG TCCCAGGCCC
AATTAAGAGA TCAGGTAGTG TAGGGTTTGG GAGCTTTTAA GGTGAAGAGG
CCCGGGCTGA TCCCACAGGC CAGTATAAAG CGCCGTGACC CTCAGGTGA***C***
GCGCCAGGGC CGGCTGCCGT CGGGGACAGG GCTTTCC(X) (SEQ ID NO: 1);

[0219] 其中X不存在,或者选自由以下组成的组:ATAGCC***GGTACC***ATG(SEQ ID NO: 104)和 ***GCCGCCACC***;

[0220] 并且核酸表达盒进一步包括基因座控制区(LCR),所述LCR包括序列2(SEQ ID NO:2)的序列、基本上由所述序列组成或由所述序列组成,其中启动子定位于所述LCR的3'端,并且其中启动子定位于编码视蛋白多肽的核酸的5'端。上文讨论了此实施例中的启动子和LCR。在一个此类实施例中,启动子包括序列4(SEQ ID NO:4)或5(SEQ ID NO:5)的序列或由所述序列组成。在另一个实施例中,编码视蛋白多肽的核酸在其N端处编码视锥细胞信号序列(如RS1信号序列(MSRKIEGFLLLLLFGYEATLGLSS(SEQ ID NO:21))。在另一个实施例中,编

码视蛋白多肽的核酸在其C端处编码外区段膜靶向序列(如VXPX结构域),其中X是任何氨基酸。在另一个实施例中,编码视蛋白多肽的核酸编码具有荧光蛋白的融合物(如其中荧光蛋白编码区存在于治疗性基因编码区的3'端的构建体)。在一个此类实施例中,编码视蛋白多肽的核酸包括编码OPN1LW/MW蛋白,并且荧光蛋白编码区包括编码绿色荧光蛋白的核酸序列。

[0221] 在另一个实施例中,核酸表达盒可以进一步包括:

[0222] 调节元件,所述调节元件包括序列15(SEQ ID NO:15)的序列、基本上由所述序列组成或由所述序列组成,其中所述调节元件定位于所述编码视蛋白多肽的核酸的3'端;和/或

[0223] 非翻译区核酸,所述非翻译区核酸包括序列16(SEQ ID NO:16)、基本上由其组成或由其组成,其中所述非翻译区核酸定位于所述调节元件的3'端。序列15的调节元件和序列16的非翻译区如上进行了讨论。

[0224] 在本公开的任何方面的表达盒的实施例中,核酸表达盒可以包括另外的调节元件,包含但不限于增强子序列、终止信号序列等。

[0225] 在本公开的任何方面的表达盒的实施例中,所述盒的5'端和3'端可以包括反向末端重复序列,如可用于允许在病毒递送媒剂中进行拯救、复制和包装的反向末端重复序列。在一个此类实施例中,反向末端重复序列是功能性腺相关病毒(AAV)反向末端重复序列。"功能性AAV ITR序列"意味着ITR序列旨在用于拯救、复制和包装AAV病毒粒子。因此,本公开的载体中使用的AAV ITR不需要具有野生型核苷酸序列并且可以通过核苷酸的插入、缺失或取代来改变,或者所述AAV ITR可以源自几种AAV血清型中的任何一种。在一个实施例中,5'ITR包括以下或由以下组成:

[0226] AAV2 5'ITR

[0227] Tgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgattaacccgccatgctacttatctacg(SEQ ID NO:79);

[0228] 并且3'ITR包括以下或由以下组成:

[0229] AAV2 3'ITR

[0230] GTTAATCATTAACTACAAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC(SEQ ID NO:80)。

[0231] 核酸表达盒可以具有被认为适合于特定目的的任何合适的长度。在本公开的任何方面的表达盒的一个非限制性实施例中,核酸表达盒的长度为约5kb或更短。

[0232] 在具体实施例中,本公开的核酸表达盒可以包括选自由序列91-95组成的组的序列、基本上由所述序列组成或由所述序列组成。

[0233] 序列#91表达盒L视蛋白GFP融合物的组分(包含ITR的总长度=4387(全长序列为SEQ ID NO:91;组分在下面分别说明)。

[0234] 5'ITR

[0235] tgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgatta

acccgccatgctacttatctacg(SEQ ID NO:79)。

[0236] 5'间隔子

[0237] TAGCCATGCTCTAGGAAGATCT(SEQ ID NO:81)。

[0238] 1.2kb LCR(从5'端缺失325个bp)

[0239] GGAGGCTGAGGGGTGGGGAAAGGGCATGGGTGTTTCATGAGGACAGAGCTTCCGTTTCATGCAATGAA
AAGAGTTTGGAGACGGATGGTGGTGACTGGACTATACACTTACACACGGTAGCGATGGTACACTTTGTATTATGTA
TATTTTACCACGATCTTTTTAAAGTGTCAAAGGCAAATGGCCAAATGGTTCCTTGTCCTATAGCTGTAGCAGCCAT
CGGCTGTTAGTGACAAAGCCCCTGAGTCAAGATGACAGCAGCCCCCATAACTCCTAATCGGCTCTCCCGCGTGGAG
TCATTTAGGAGTAGTCGCATTAGAGACAAGTCCAACATCTAATCTTCCACCCTGGCCAGGGCCCCAGCTGGCAGCG
AGGGTGGGAGACTCCGGGCAGAGCAGAGGGCGCTGACATTGGGGCCCGGCCTGGCTTGGGTCCCTCTGGCCTTTCC
CCAGGGGCCCTCTTTCCTTGGGGCTTTCTTGGGCCGCCACTGCTCCCGCTCCTCTCCCCCCATCCCACCCCCTCAC
CCCCTCGTTCTTCATATCCTTCTCTAGTGCTCCCTCCACTTTCATCCACCCTTCTGCAAGAGTGTGGGACCACAAA
TGAGTTTTCACCTGGCCTGGGGACACACGTGCCCCCACAGGTGCTGAGTGACTTTCTAGGACAGTAATCTGCTTTA
GGCTAAAATGGGACTTGATCTTCTGTTAGCCCTAATCATCAATTAGCAGAGCCGGTGAAGGTGCAGAACCTACCGC
CTTTCCAGGCCTCCTCCCACCTCTGCCACCTCCACTCTCCTTCCTGGGATGTGGGGGCTGGCACACGTGTGGCCCA
GGGCATTGGTGGGATTGCACTGAGCTGGGTCATTAGCGTAATCCTGGACAAGGGCAGACAGGGCGAGCGGAGGGCC
AGCTCCGGGGCTCAGGCAAGGCTGGGGGCTTCCCCCAGACACCCCACTCCTCCTCTGCTGGACCCCCACTTCATAG
GGCACTTCGTGTTCTCAAAGGGCTTCCAAATAGCATGGTGGCCTTGGATGCCCAGGGAAGCCTCAGAGTTGCTTAT
CTCCCTCTAGACAGAAGGGGAATCTCGGTCAAGAGGGAGAGGTCGCCCTGTTCAAGGCCACCCAGCCAGCTCATGG
CGGTAATGGGACAAGGCTGGCCAGCCATCCCACCCTCAGAAGGGACCCGGTGGGGCAGGTGATCTCAGAGGAGGCT
CACTTCTGGGTCTCACATTCTT(SEQ ID NO:2)。

[0240] 171+9个经过优化的L启动子,T>C,Kpn,置换了ATG(缺失了-190到-130区域,所述区域同时存在于抑制转录的495L启动子和M启动子两者中)

[0241] GGGGAGGAGGAGGTCTAAGTCCCAGGCCCAATTAAGAGATCAGGTAGTGTAGGG
TTTGGGAGCTTTTAAGGTGAAGAGGCCCGGGCTGATCCCACAGGCCAGTATAAAGCG
CCGTGACCCTCAGGTGA**C**GCGCCAGGGCCGGCTGCCGTCGGGGACAGGGCTTTCCAT
AGCC**GGTACC**ATG（SEQ ID NO: 82）

[0242] 突出显示的说明：**GGTACC** 被插入KpnI位点,以取代ATG并破坏诱饵Kozak

[0243] **C** =单核苷酸取代增加了表达（通常这是 T 残基）

[0244] SV40微小内含子

[0245] CTAGAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTT
TAGTCTTTTTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAA
CTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTG
CTCTAAAAGCTGCGGAATTGTACCC**GCCGCCACC**（SEQ ID NO: 83）

[0246] 突出显示的说明：**GCCGCCACC** =经过优化的 Kozak（ATG 除外）

[0247] L视蛋白/GFP融合物(L视蛋白部分、S尾、GFP、原生尾)

[0248] L视蛋白部分：

[0249] ATGGCCCAGCAGTGGAGCCTCCAAAGGCTCGCAGGCCGCCATCCGCAGGACAGCTATGAGGACAGCAC
CCAGTCCAGCATCTTCACCTACACCAACAGCAACTCCACCAGAGGCCCCTTCGAAGGCCCGAATTACCACATCGCT
CCCAGATGGGTGTACCACCTCACCAGTGTCTGGATGATCTTTGTGGTCACTGCATCCGTCTTCACAAATGGGCTTG
TGCTGGCGGCCACCATGAAGTTCAAGAAGCTGCGCCACCCGCTGAACTGGATCCTGGTGAACCTGGCGGTCGCTGA
CCTAGCAGAGACCGTCATCGCCAGCACTATCAGCATTGTGAACCAGGTCTCTGGCTACTTCGTGCTGGGCCACCCT
ATGTGTGTCCTGGAGGGCTACACCGTCTCCCTGTGTGGGATCACAGGTCTCTGGTCTCTGGCCATCATTTCCTGGG
AGAGGTGGCTGGTGGTGTGCAAGCCCTTTGGCAATGTGAGATTTGATGCCAAGCTGGCCATCGTGGGCATTGCCTT
CTCCTGGATCTGGTCTGCTGTGTGGACAGCCCCGCCCATCTTTGGTTGGAGCAGGTACTGGCCCCACGGCCTGAAG
ACTTCATGCGGCCCAGACGTGTTCAGCGGCAGCTCGTACCCCGGGGTGCAGTCTTACATGATTGTCCTCATGGTCA
CCTGCTGCATCATCCCACTCGCTATCATCATGCTCTGCTACCTCCAAGTGTGGCTGGCCATCCGAGCGGTGGCAAA
GCAGCAGAAAGAGTCTGAATCCACCCAGAAGGCAGAGAAGGAAGTGACGCGCATGGTGGTGGTGATGATCTTTGCG
TACTGCGTCTGCTGGGGACCCTACACCTTCTTCGCATGCTTTGCTGCTGCCAACCCTGGTTACGCCTTCCACCCTT
TGATGGCTGCCCTGCCGGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCGTTATCTATGTCTTTATGAACCG
GCAGTTTCGAAACTGCATCTTGCAGCTTTTCGGGAAGAAGGTTGACGATGGCTCTGAACTCTCCAGCGCCTCCAAA
ACGGAGGTC(SEQ ID NO:84)

[0250] S尾:TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72)(C端12个S视蛋白氨基酸的密码)

[0251] GfP:

[0252] ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGT
AAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC
TGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCC
GCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATC
GAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA
ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGG
CAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGA
CCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGTC(SEQ ID NO:85)

[0253] 原生L视蛋白尾(最后10个氨基酸的密码)

[0254] TCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID NO:73)

[0255] WPRE(短,246bp)

[0256] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTAC
GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGC
TGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0257] 来自L视蛋白的3'UTR(这是我们载体中的polyA信号的唯一来源)

[0258] GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATC
CCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAACAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACC
TGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACT

TTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC(SEQ ID NO:86)

[0259] 3’间隔子：

[0260] AAGCTTATCGATAAGGATCTTCCTAGAGCATGGCTA(SEQ ID NO:87)

[0261] 3’ITR

[0262] Cgtagataagtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactcctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgca(SEQ ID NO:88)

[0263] 序列#92.表达盒L视蛋白的组分(包含ITR的总长度＝3633(全长序列为SEQ ID NO:92;组分在下面分别说明)。

[0264] 5’ITR

[0265] tgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgtcgggcgacctttggtcgcccggcctcagtgagcgagcgagcgcgcagagagggagtggccaactccatcactaggggttccttgtagttaatgattaacccgccatgctacttatctacg(SEQ ID NO:79)

[0266] 5’间隔子

[0267] TAGCCATGCTCTAGGAAGATCT(SEQ ID NO:81)

[0268] 1.2kb LCR(与现有技术中的1.6kb LCR相比,缺失了5’325bp)

[0269] GGAGGCTGAGGGGTGGGGAAAGGGCATGGGTGTTTCATGAGGACAGAGCTTCCGTTTCATGCAATGAAAAGAGTTTGGAGACGGATGGTGGTGACTGGACTATACACTTACACACGGTAGCGATGGTACACTTTGTATTATGTATATTTTACCACGATCTTTTTAAAGTGTCAAAGGCAAATGGCCAAATGGTTCCTTGTCCTATAGCTGTAGCAGCCATCGGCTGTTAGTGACAAAGCCCCTGAGTCAAGATGACAGCAGCCCCCATAACTCCTAATCGGCTCTCCCGCGTGGAGTCATTTAGGAGTAGTCGCATTAGAGACAAGTCCAACATCTAATCTTCCACCCTGGCCAGGGCCCCAGCTGGCAGCGAGGGTGGGAGACTCCGGGCAGAGCAGAGGGCGCTGACATTGGGGCCCGGCCTGGCTTGGGTCCCTCTGGCCTTTCCCCAGGGGCCCTCTTTCCTTGGGGCTTTCTTGGGCCGCCACTGCTCCCGCTCCTCTCCCCCCATCCCACCCCCTCACCCCCTCGTTCTTCATATCCTTCTCTAGTGCTCCCTCCACTTTCATCCACCCTTCTGCAAGAGTGTGGGACCACAAATGAGTTTTCACCTGGCCTGGGGACACACGTGCCCCCACAGGTGCTGAGTGACTTTCTAGGACAGTAATCTGCTTTAGGCTAAAATGGGACTTGATCTTCTGTTAGCCCTAATCATCAATTAGCAGAGCCGGTGAAGGTGCAGAACCTACCGCCTTTCCAGGCCTCCTCCCACCTCTGCCACCTCCACTCTCCTTCCTGGGATGTGGGGGCTGGCACACGTGTGGCCCAGGGCATTGGTGGGATTGCACTGAGCTGGGTCATTAGCGTAATCCTGGACAAGGGCAGACAGGGCGAGCGGAGGGCCAGCTCCGGGGCTCAGGCAAGGCTGGGGGCTTCCCCCAGACACCCCACTCCTCCTCTGCTGGACCCCCACTTCATAGGGCACTTCGTGTTCTCAAAGGGCTTCCAAATAGCATGGTGGCCTTGGATGCCCAGGGAAGCCTCAGAGTTGCTTATCTCCCTCTAGACAGAAGGGGAATCTCGGTCAAGAGGGAGAGGTCGCCCTGTTCAAGGCCACCCAGCCAGCTCATGGCGGTAATGGGACAAGGCTGGCCAGCCATCCCACCCTCAGAAGGGACCCGGTGGGGCAGGTGATCTCAGAGGAGGCTCACTTCTGGGTCTCACATTCTT(SEQ ID NO:2)

[0270] 171+9个经过优化的L启动子,T>C,Kpn,置换了ATG(缺失了-190到-130区域,所述区域同时存在于抑制转录的495L启动子和M启动子两者中)

[0271] GGGGAGGAGGAGGTCTAAGTCCCAGGCCCAATTAAGAGATCAGGTAGTGTAGGG TTTGGGAGCTTTTAAGGTGAAGAGGCCCGGGCTGATCCCACAGGCCAGTATAAAGCG CCGTGACCCTCAGGTGA**C**GCGCCAGGGCCGGCTGCCGTCGGGGACAGGGCTTTCCAT AGCC**GGTACC**ATG（SEQ ID NO: 82）

[0272] 突出显示的说明：**GGTACC**被插入KpnI位点，以取代ATG并破坏诱饵Kozak

[0273] **C**=单核苷酸取代增加了表达（通常这是 T 残基）

[0274] SV40微小内含子

[0275] CTAGAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTT TAGTCTTTTTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAA CTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTG CTCTAAAAGCTGCGGAATTGTACCC**GCCGCCACC**（SEQ ID NO: 83）

[0276] 突出显示的说明：**GCCGCCACC**=经过优化的 Kozak（ATG 除外）

[0277] L视蛋白(任何视蛋白)

[0278] ATGGCCCAGCAGTGGAGCCTCCAAAGGCTCGCAGGCCGCCATCCGCAGGACAGCTATGAGGACAGCAC CCAGTCCAGCATCTTCACCTACACCAACAGCAACTCCACCAGAGGCCCCTTCGAAGGCCCGAATTACCACATCGCT CCCAGATGGGTGTACCACCTCACCAGTGTCTGGATGATCTTTGTGGTCACTGCATCCGTCTTCACAAATGGGCTTG TGCTGGCGGCCACCATGAAGTTCAAGAAGCTGCGCCACCCGCTGAACTGGATCCTGGTGAACCTGGCGGTCGCTGA CCTAGCAGAGACCGTCATCGCCAGCACTATCAGCATTGTGAACCAGGTCTCTGGCTACTTCGTGCTGGGCCACCCT ATGTGTGTCCTGGAGGGCTACACCGTCTCCCTGTGTGGGATCACAGGTCTCTGGTCTCTGGCCATCATTTCCTGGG AGAGGTGGCTGGTGGTCTGCAAGCCCTTTGGCAATGTGAGATTTGATGCCAAGCTGGCCATCGTGGGCATTGCCTT CTCCTGGATCTGGTCTGCTGTGTGGACAGCCCCGCCCATCTTTGGTTGGAGCAGGTACTGGCCCCACGGCCTGAAG ACTTCATGCGGCCCAGACGTGTTCAGCGGCAGCTCGTACCCCGGGGTGCAGTCTTACATGATTGTCCTCATGGTCA CCTGCTGCATCATCCCACTCGCTATCATCATGCTCTGCTACCTCCAAGTGTGGCTGGCCATCCGAGCGGTGGCAAA GCAGCAGAAAGAGTCTGAATCCACCCAGAAGGCAGAGAAGGAAGTGACGCGCATGGTGGTGGTGATGATCTTTGCG TACTGCGTCTGCTGGGGACCCTACACCTTCTTCGCATGCTTTGCTGCTGCCAACCCTGGTTACGCCTTCCACCCTT TGATGGCTGCCCTGCCGGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCGTTATCTATGTCTTTATGAACCG GCAGTTTCGAAACTGCATCTTGCAGCTTTTCGGGAAGAAGGTTGACGATGGCTCTGAACTCTCCAGCGCCTCCAAA ACGGAGGTCTCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID NO:84)

[0279] WPRE(短)

[0280] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTAC GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTG TATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGC TGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0281] 3'UTR(这是经过优化的载体中的polyA信号的唯一来源)

[0282] GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATC CCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAACAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACC

TGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACT TTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC(SEQ ID NO:86)

[0283] 3'间隔子：

[0284] AAGCTTATCGATAAGGATCTTCCTAGAGCATGGCTA(SEQ ID NO:87)

[0285] 3'ITR

[0286] Cgtagataagtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactc cctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggc ctcagtgagcgagcgagcgcgca(SEQ ID NO:88)

[0287] 序列#93L/M视锥表达的带有柠檬黄标签的VEGF trap(全长序列为SEQ ID NO:93；组分在下面分别说明)。

[0288] AAV2 5'ITR

[0289] TGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCG GCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTA ACCCGCCATGCTACTTATCTACG(SEQ ID NO:79--)

[0290] 间隔子:TAGCCATGCTCTAGGAAGATCT(SEQ ID NO:81)

[0291] LCR 1.2Kb：

[0292] GGAGGCTGAGGGGTGGGGAAAGGGCATGGGTGTTTCATGAGGACAGAGCTTCCGTTTCATGCAATGAA AAGAGTTTGGAGACGGATGGTGGTGACTGGACTATACACTTACACACGGTAGCGATGGTACACTTTGTATTATGTA TATTTTACCACGATCTTTTTAAAGTGTCAAAGGCAAATGGCCAAATGGTTCCTTGTCCTATAGCTGTAGCAGCCAT CGGCTGTTAGTGACAAAGCCCCTGAGTCAAGATGACAGCAGCCCCCATAACTCCTAATCGGCTCTCCCGCGTGGAG TCATTTAGGAGTAGTCGCATTAGAGACAAGTCCAACATCTAATCTTCCACCCTGGCCAGGGCCCCAGCTGGCAGCG AGGGTGGGAGACTCCGGGCAGAGCAGAGGGCGCTGACATTGGGGCCCGGCCTGGCTTGGGTCCCTCTGGCCTTTCC CCAGGGGCCCTCTTTCCTTGGGGCTTTCTTGGGCCGCCACTGCTCCCGCTCCTCTCCCCCCATCCCACCCCCTCAC CCCCTCGTTCTTCATATCCTTCTCTAGTGCTCCCTCCACTTTCATCCACCCTTCTGCAAGAGTGTGGGACCACAAA TGAGTTTTCACCTGGCCTGGGGACACACGTGCCCCCACAGGTGCTGAGTGACTTTCTAGGACAGTAATCTGCTTTA GGCTAAAATGGGACTTGATCTTCTGTTAGCCCTAATCATCAATTAGCAGAGCCGGTGAAGGTGCAGAACCTACCGC CTTTCCAGGCCTCCTCCCACCTCTGCCACCTCCACTCTCCTTCCTGGGATGTGGGGGCTGGCACACGTGTGGCCCA GGGCATTGGTGGGATTGCACTGAGCTGGGTCATTAGCGTAATCCTGGACAAGGGCAGACAGGGCGAGCGGAGGGCC AGCTCCGGGGCTCAGGCAAGGCTGGGGGCTTCCCCCAGACACCCCACTCCTCCTCTGCTGGACCCCCACTTCATAG GGCACTTCGTGTTCTCAAAGGGCTTCCAAATAGCATGGTGGCCTTGGATGCCCAGGGAAGCCTCAGAGTTGCTTAT CTCCCTCTAGACAGAAGGGGAATCTCGGTCAAGAGGGAGAGGTCGCCCTGTTCAAGGCCACCCAGCCAGCTCATGG CGGTAATGGGACAAGGCTGGCCAGCCATCCCACCCTCAGAAGGGACCCGGTGGGGCAGGTGATCTCAGAGGAGGCT CACTTCTGGGTCTCACATTCTTG(SEQ ID NO:2)

[0293] 经过优化的171bp L启动子

[0294] GGGAGGAGGAGGTCTAAGTCCCAGGCCCAATTAAGAGATCAGGTAGTGTAGGGTTTGGGAGCTTTTAA GGTGAAGAGGCCCGGGCTGATCCCACAGGCCAGTATAAAGCGCCGTGACCCTCAGGTGACGCGCCAGGGCCGGCTG CCGTCGGGGACAGGGCTTTCCATAGCC(SEQ ID NO:89)

[0295] SV40微小内含子

[0296] GGTACCATGCTAGAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTC TTTTTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTT TACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGCc(SEQ ID NO:90)

[0297] 经过优化的Kozak(部分)

[0298] GCCACC

[0299] Vegf trap分为4部分(sflt信号序列、sflt结构域2、VEGFR2结构域3、IgG1fc)：

[0300] Sflt信号序列：

[0301] ATGGTCAGCTACTGGGACACCGGGGTCCTGCTGTGCGCGCTGCTCAGCTGTCTGCTTCTCACAGGATC TAGTTCCGGA(SEQ ID NO:78)

[0302] Sflt结构域2

[0303] AGTGATACCGGTAGACCTTTCGTAGAGATGTACAGTGAAATCCCCGAAATTATACACATGACTGAAGG AAGGGAGCTCGTCATTCCCTGCCGGGTTACGTCACCTAACATCACTGTTACTTTAAAAAAGTTTCCACTTGACACT TTGATCCCTGATGGAAAACGCATAATCTGGGACAGTAGAAAGGGCTTCATCATATCAAATGCAACGTACAAAGAAA TAGGGCTTCTGACCTGTGAAGCAACAGTCAATGGGCATTTGTATAAGACAAACTATCTCACACATCGACAAACCAA TACAATCATAGATGTG(SEQ ID NO:17)

[0304] VEGFR2结构域3

[0305] GTTCTGAGTCCGTCTCATGGAATTGAACTATCTGTTGGAGAAAAGCTTGTCTTAAATTGTACAGCAAG AACTGAACTAAATGTGGGGATTGACTTCAACTGGGAATACCCTTCTTCGAAGCATCAGCATAAGAAACTTGTAAAC CGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAGAAATTTTTGAGCACCTTAACTATAGATGGTGTAACCCGGA GTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCTGATGACCAAGAAGAACAGCACATTTGTCAGGGTCCA TGAAAAG(SEQ ID NO:18)

[0306] IgG1 fc

[0307] GACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTT CCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCAC GAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGG AGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTA CAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGA GAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCA AAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCC TCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGG AACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG GTAAA(SEQ ID NO:19)

[0308] VEGF Trap与柠檬黄之间的连接子

[0309] Ggaggtggaggttctggtggaggaggttcc(SEQ ID NO:76)

[0310] 柠檬黄

[0311] ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGT AAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC TGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGATGTGCTTCGCCC GCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT

CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATC
GAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACA
ACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGG
CAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC
TACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGA
CCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAG(SEQ ID NO:74)

[0312] WPRE短

[0313] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTAC
GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGC
TGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0314] 3'UTR OPN1LW

[0315] GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATC
CCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAACAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACC
TGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACT
TTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC(SEQ ID
NO:86)

[0316] 间隔子:aagcttatcgataaggatcttcctagagcatggcta(SEQ ID NO:87)

[0317] AAV2 3'ITR

[0318] Cgtagataagtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactc
cctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggc
ctcagtgagcgagcgagcgcgca(SEQ ID NO:88)

[0319] 序列#94L/M视锥表达的VEGF trap(无柠檬黄)(全长序列为SEQ ID NO:94;组分在下面分别说明)。

[0320] AAV2 5'ITR

[0321] TGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCG
GCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTA
ACCCGCCATGCTACTTATCTACG(SEQ ID NO:79)

[0322] 间隔子:TAGCCATGCTCTAGGAAGATCT(SEQ ID NO:81)

[0323] LCR 1.2Kb:

[0324] GGAGGCTGAGGGGTGGGGAAAGGGCATGGGTGTTTCATGAGGACAGAGCTTCCGTTTCATGCAATGAA
AAGAGTTTGGAGACGGATGGTGGTGACTGGACTATACACTTACACACGGTAGCGATGGTACACTTTGTATTATGTA
TATTTTACCACGATCTTTTTAAAGTGTCAAAGGCAAATGGCCAAATGGTTCCTTGTCCTATAGCTGTAGCAGCCAT
CGGCTGTTAGTGACAAAGCCCCTGAGTCAAGATGACAGCAGCCCCCATAACTCCTAATCGGCTCTCCCGCGTGGAG
TCATTTAGGAGTAGTCGCATTAGAGACAAGTCCAACATCTAATCTTCCACCCTGGCCAGGGCCCCAGCTGGCAGCG
AGGGTGGGAGACTCCGGGCAGAGCAGAGGGCGCTGACATTGGGGCCCGGCCTGGCTTGGGTCCCTCTGGCCTTTCC
CCAGGGGCCCTCTTTCCTTGGGGCTTTCTTGGGCCGCCACTGCTCCCGCTCCTCTCCCCCCATCCCACCCCCTCAC
CCCCTCGTTCTTCATATCCTTCTCTAGTGCTCCCTCCACTTTCATCCACCCTTCTGCAAGAGTGTGGGACCACAAA
TGAGTTTTCACCTGGCCTGGGGACACACGTGCCCCCACAGGTGCTGAGTGACTTTCTAGGACAGTAATCTGCTTTA

GGCTAAAATGGGACTTGATCTTCTGTTAGCCCTAATCATCAATTAGCAGAGCCGGTGAAGGTGCAGAACCTACCGC
CTTTCCAGGCCTCCTCCCACCTCTGCCACCTCCACTCTCCTTCCTGGGATGTGGGGGCTGGCACACGTGTGGCCCA
GGGCATTGGTGGGATTGCACTGAGCTGGGTCATTAGCGTAATCCTGGACAAGGGCAGACAGGGCGAGCGGAGGGCC
AGCTCCGGGGCTCAGGCAAGGCTGGGGGCTTCCCCCAGACACCCCACTCCTCCTCTGCTGGACCCCCACTTCATAG
GGCACTTCGTGTTCTCAAAGGGCTTCCAAATAGCATGGTGGCCTTGGATGCCCAGGGAAGCCTCAGAGTTGCTTAT
CTCCCTCTAGACAGAAGGGGAATCTCGGTCAAGAGGGAGAGGTCGCCCTGTTCAAGGCCACCCAGCCAGCTCATGG
CGGTAATGGGACAAGGCTGGCCAGCCATCCCACCCTCAGAAGGGACCCGGTGGGGCAGGTGATCTCAGAGGAGGCT
CACTTCTGGGTCTCACATTCTTG(SEQ ID NO:2)

[0325] 经过优化的171bp L启动子

[0326] GGGAGGAGGAGGTCTAAGTCCCAGGCCCAATTAAGAGATCAGGTAGTGTAGGGTTTGGGAGCTTTTAA
GGTGAAGAGGCCCGGGCTGATCCCACAGGCCAGTATAAAGCGCCGTGACCCTCAGGTGACGCGCCAGGGCCGGCTG
CCGTCGGGGACAGGGCTTTCCATAGCC(SEQ ID NO:89)

[0327] SV40微小内含子

[0328] GGTACCATGCTAGAGGATCCGGTACTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTAAGTTTAGTC
TTTTTGTCTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGATGTTGCCTT
TACTTCTAGGCCTGTACGGAAGTGTTACTTCTGCTCTAAAAGCTGCGGAATTGTACCCGCc(SEQ ID NO:90)

[0329] 经过优化的Kozak(部分)

[0330] GCCACC

[0331] Vegf trap分为4部分(sflt信号序列、sflt结构域2、VEGFR2结构域3、IgG1fc)：

[0332] Sflt信号序列：

[0333] ATGGTCAGCTACTGGGACACCGGGGTCCTGCTGTGCGCGCTGCTCAGCTGTCTG
CTTCTCACAGGATCTAGTTCCGGA(SEQ ID NO:78)

[0334] Sflt结构域2

[0335] AGTGATACCGGTAGACCTTTCGTAGAGATGTACAGTGAAATCCCCGAAATTATACACATGACTGAAGG
AAGGGAGCTCGTCATTCCCTGCCGGGTTACGTCACCTAACATCACTGTTACTTTAAAAAAGTTTCCACTTGACACT
TTGATCCCTGATGGAAAACGCATAATCTGGGACAGTAGAAAGGGCTTCATCATATCAAATGCAACGTACAAAGAAA
TAGGGCTTCTGACCTGTGAAGCAACAGTCAATGGGCATTTGTATAAGACAAACTATCTCACACATCGACAAACCAA
TACAATCATAGATGTG(SEQ ID NO:17)

[0336] VEGFR2结构域3

[0337] GTTCTGAGTCCGTCTCATGGAATTGAACTATCTGTTGGAGAAAAGCTTGTCTTAAATTGTACAGCAAG
AACTGAACTAAATGTGGGGATTGACTTCAACTGGGAATACCCTTCTTCGAAGCATCAGCATAAGAAACTTGTAAAC
CGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAGAAATTTTTGAGCACCTTAACTATAGATGGTGTAACCCGGA
GTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCTGATGACCAAGAAGAACAGCACATTTGTCAGGGTCCA
TGAAAAG(SEQ ID NO:18)

[0338] IgG1 fc

[0339] GACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTT
CCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCAC
GAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGG
AGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTA

CAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGA
GAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCA
AAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCC
TCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGG
AACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
GTAAATAG(SEQ ID NO:19)

[0340] WPRE短

[0341] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTAC
GCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTG
TATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGC
TGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0342] 3'UTR OPN1LW

[0343] GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATC
CCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAACAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACC
TGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACT
TTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC(SEQ ID
NO:86)

[0344] 间隔子:aagcttatcgataaggatcttcctagagcatggcta(SEQ ID NO:87)

[0345] AAV2 3'ITR

[0346] Cgtagataagtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactc
cctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggc
ctcagtgagcgagcgagcgcgca(SEQ ID NO:88)

[0347] 序列#95L/M视锥表达盒的具有经过修饰的人β球蛋白内含子的版本(全长序列为SEQ ID NO:95;组分在下面分别说明)。

[0348] AAV2 5'ITR:

[0349] TGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCG
GCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCTTGTAGTTAATGATTA
ACCCGCCATGCTACTTATCTACG(SEQ ID NO:79)

[0350] 5'间隔子:TAGCCATGCTCTAGGAAGATCT(SEQ ID NO:81)

[0351] 1.2kb短LCR

[0352] GGAGGCTGAGGGGTGGGGAAAGGGCATGGGTGTTTCATGAGGACAGAGCTTCCGTTTCATGCAATGAA
AAGAGTTTGGAGACGGATGGTGGTGACTGGACTATACACTTACACACGGTAGCGATGGTACACTTTGTATTATGTA
TATTTTACCACGATCTTTTTAAAGTGTCAAAGGCAAATGGCCAAATGGTTCCTTGTCCTATAGCTGTAGCAGCCAT
CGGCTGTTAGTGACAAAGCCCCTGAGTCAAGATGACAGCAGCCCCCATAACTCCTAATCGGCTCTCCCGCGTGGAG
TCATTTAGGAGTAGTCGCATTAGAGACAAGTCCAACATCTAATCTTCCACCCTGGCCAGGGCCCCAGCTGGCAGCG
AGGGTGGGAGACTCCGGGCAGAGCAGAGGGCGCTGACATTGGGGCCCGGCCTGGCTTGGGTCCCTCTGGCCTTTCC
CCAGGGGCCCTCTTTCCTTGGGGCTTTCTTGGGCCGCCACTGCTCCCGCTCCTCTCCCCCCATCCCACCCCCTCAC
CCCCTCGTTCTTCATATCCTTCTCTAGTGCTCCCTCCACTTTCATCCACCCTTCTGCAAGAGTGTGGGACCACAAA
TGAGTTTTCACCTGGCCTGGGGACACACGTGCCCCCACAGGTGCTGAGTGACTTTCTAGGACAGTAATCTGCTTTA

GGCTAAAATGGGACTTGATCTTCTGTTAGCCCTAATCATCAATTAGCAGAGCCGGTGAAGGTGCAGAACCTACCGC
CTTTCCAGGCCTCCTCCCACCTCTGCCACCTCCACTCTCCTTCCTGGGATGTGGGGGCTGGCACACGTGTGGCCCA
GGGCATTGGTGGGATTGCACTGAGCTGGGTCATTAGCGTAATCCTGGACAAGGGCAGACAGGGCGAGCGGAGGGCC
AGCTCCGGGGCTCAGGCAAGGCTGGGGGCTTCCCCCAGACACCCCACTCCTCCTCTGCTGGACCCCCACTTCATAG
GGCACTTCGTGTTCTCAAAGGGCTTCCAAATAGCATGGTGGCCTTGGATGCCCAGGGAAGCCTCAGAGTTGCTTAT
CTCCCTCTAGACAGAAGGGGAATCTCGGTCAAGAGGGAGAGGTCGCCCTGTTCAAGGCCACCCAGCCAGCTCATGG
CGGTAATGGGACAAGGCTGGCCAGCCATCCCACCCTCAGAAGGGACCCGGTGGGGCAGGTGATCTCAGAGGAGGCT
CACTTCTGGGTCTCACATTCTT(SEQ ID NO:2)

[0353] 具有T>C的经过优化的短L启动子和经过优化的Kozak

[0354] GGGGAGGAGGAGGTCTAAGTCCCAGGCCCAATTAAGAGATCAGGTAGTGTAGGGTTTGGGAGCTTTTA
AGGTGAAGAGGCCCGGGCTGATCCCACAGGCCAGTATAAAGCGCCGTGACCCTCAGGTGACGCGCCAGGGCCGGCT
GCCGTCGGGGACAGGGCTTTCCATAGCCGCCACC(SEQ ID NO:96)

[0355] (经过优化的Kozak突出显示)

[0356] OPN1LW外显子1和2

[0357] ATGGCCCAGCAGTGGAGCCTCCAAAGGCTCGCAGGCCGCCATCCGCAGGACAGCTATGAGGACAGCAC
CCAGTCCAGCATCTTCACCTACACCAACAGCAACTCCACCAGAGGCCCCTTCGAAGGCCCGAATTACCACATCGCT
CCCAGATGGGTGTACCACCTCACCAGTGTCTGGATGATCTTTGTGGTCACTGCATCCGTCTTCACAAATGGGCTTG
TGCTGGCGGCCACCATGAAGTTCAAGAAGCTGCGCCACCCGCTGAACTGGATCCTGGTGAACCTGGCGGTCGCTGA
CCTAGCAGAGACCGTCATCGCCAGCACTATCAGCATTGTGAACCAGGTCTCTGGCTACTTCGTGCTGGGCCACCCT
ATGTGTGTCCTGGAGGGCTACACCGTCTCCCTGTGTG(SEQ ID NO:97)

[0358] 内含子1(源自pFLARE载体的经过修饰的人β球蛋白内含子)

[0359] GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATG TGGAGACAGAGAA
GACTCACGCGTTTCTGAATTCACTGACTCTCTCTGCCTATTGGT CTATTTTCTCACCCTTAG(SEQ ID NO:7)

[0360] OPN1LW外显子3

[0361] GGATCACAGGTCTCTGGTCTCTGGCCATCATTTCCTGGGAGAGATGGATGGTGGTCTGCAAGCCCTTT
GGCAATGTGAGATTTGATGCCAAGCTGGCCATCGTGGGCATTGCCTTCTCCTGGATCTGGTCTGCTGTGTGGACAG
CCCCGCCCATCTTTGGTTGGAGCAG(SEQ ID NO:98)

[0362] 内含子2(源自pFLARE载体的经过修饰的人β球蛋白内含子)

[0363] GTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATGGATCCATAGTCGACCACCATGGTGGCT
TAGATCCGGGCATGTGGAGACAGAGAAGACTGTTGAGTTTGTGATAAGCACTGACTCTCTCTGCCTATTGGTCTAT
TTTCCCTCCCTCAG(SEQ ID NO:8)

[0364] OPN1LW外显子4、5和6

[0365] GTACTGGCCCCACGGCCTGAAGACTTCATGCGGCCCAGACGTGTTCAGCGGCAGCTCGTACCCCGGGG
TGCAGTCTTACATGATTGTCCTCATGGTCACCTGCTGCATCATCCCACTCGCTATCATCATGCTCTGCTACCTCCA
AGTGTGGCTGGCCATCCGAGCGGTGGCAAAGCAGCAGAAAGAGTCTGAATCCACCCAGAAGGCAGAGAAGGAAGTG
ACGCGCATGGTGGTGGTGATGATCTTTGCGTACTGCGTCTGCTGGGGACCCTACACCTTCTTCGCATGCTTTGCTG
CTGCCAACCCTGGTTACGCCTTCCACCCTTTGATGGCTGCCCTGCCGGCCTACTTTGCCAAAAGTGCCACTATCTA
CAACCCCGTTATCTATGTCTTTATGAACCGGCAGTTTCGAAACTGCATCTTGCAGCTTTTCGGGAAGAAGGTTGAC
GATGGCTCTGAACTCTCCAGCGCCTCCAAAACGGAGGTCTCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID

NO:99)

[0366] OPN1LW 3'UTR

[0367] GGTCTGCCTCCTACCCATCCCGCCCACCGGGGCTTTGGCCACCTCTCCTTTCCCCCTCCTTCTCCATCCCTGTAAAATAAATGTAATTTATCTTTGCCAAAACCAACAAAGTCACAGAGGCTTTCACTGCAGTGTGGGACCACCTGAGCCTCTGCGTGTGCAGGCACTGGGTCTCGAGAGGGTGCAAGGGGGATAAAGAGGAGAGAGCGCTTCATAGACTTTAAGTTTTCCCGAGCCTCATGTCTACCGATGGCGTGAAAGGATCCTGGCAAAACAGAAGTGTGAGGC(SEQ ID NO:86)

[0368] 短WPRE

[0369] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTG(SEQ ID NO:15)

[0370] 3'间隔子

[0371] AAGCTTATCGATAAGGATCTTCCTAGAGCATGGCTA(SEQ ID NO:87)

[0372] AAV2 3'ITR

[0373] 3'ITR

[0374] Cgtagataagtagcatggcgggttaatcattaactacaaggaacccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactgaggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcgagcgagcgcgca(SEQ ID NO:88)

[0375] 在本文的任何实施例或实施例组合中的一个实施例中,核酸表达盒包括核酸表达载体。如本文所使用的,"表达载体"涵盖载体(例如,如上文所讨论的或如本领域中已知的质粒、微环、病毒载体、脂质体等)并且用于在预期的靶细胞中实现基因产物的表达,所述载体包括编码所关注的基因产物的核酸表达盒。

[0376] 在另一个方面,提供了重组宿主细胞,所述重组宿主细胞包括根据本公开的任何实施例或实施例组合所述的核酸表达盒。细胞可以具有可以用本文所公开的表达载体转染或转导的任何类型。应当理解,术语"宿主细胞"是指原始转导、感染、转染或转化的细胞及其子代。在表达载体是rAAV载体的一个实施例中,所述细胞包括用本公开的无复制能力的rAAV表达载体转导的生产细胞,如上文所描述的和本领域众所周知的,通过引入AAV辅助构建体可以从所述生产细胞中获得病毒颗粒。

[0377] 在另一个方面,提供了重组腺相关病毒(rAAV)颗粒,其包括

[0378] (a)AAV衣壳蛋白;以及

[0379] (b)根据本公开的任何实施例或实施例组合所述的核酸表达盒或表达载体。

[0380] 术语"AAV"是腺相关病毒的缩写,并且可以用于指病毒本身或其衍生物。除非另有要求,否则所述术语涵盖所有亚型以及天然存在的形式和重组形式两者。术语"AAV"包含1型AAV(AAV-1)、2型AAV(AAV-2)、3型AAV(AAV-3)、4型AAV(AAV-4)、5型AAV(AAV-5)、6型AAV(AAV-6)、7型AAV(AAV-7)、8型AAV(AAV-8)、禽类AAV、牛类AAV、犬类AAV、马类AAV、灵长类动物AAV、非灵长类动物AAV和绵羊类AAV。"灵长类动物AAV"是指感染灵长类动物的AAV,"非灵长类动物AAV"是指感染非灵长类哺乳动物的AAV,"牛类AAV"是指感染牛类哺乳动物的AAV,等等。

[0381] “AAV病毒”或“AAV病毒颗粒”或“rAAV载体颗粒”是指由至少一种AAV衣壳蛋白(通常由野生型AAV的所有衣壳蛋白)和衣壳化的多核苷酸rAAV表达盒或载体构成的病毒颗粒。如果颗粒包括异源多核苷酸(即,除野生型AAV基因组以外的多核苷酸,如待递送至哺乳动物细胞的治疗性基因),则所述颗粒通常被称为“rAAV载体颗粒”或简称为“rAAV载体”。因此,rAAV颗粒的产生必然包含rAAV载体的产生,如此rAAV载体颗粒内含有载体。

[0382] 如本文相对于本公开的AAV病毒载体所使用的,术语“复制缺陷”意指AAV载体无法独立地复制和包装其基因组。例如,当受试者的细胞感染了rAAV病毒粒子时,异源基因在受感染细胞中表达,然而,由于受感染细胞缺乏AAV rep和cap基因以及辅助功能基因,因此rAAV不能进一步复制。

[0383] 缩写“rAAV”是指重组腺相关病毒,也被称为重组AAV载体(或“rAAV载体”)。如本文所使用的,“rAAV载体”是指包括不属于AAV来源的多核苷酸序列(即,与AAV异源的多核苷酸)的AAV载体,所述多核苷酸序列通常是对细胞进行基因转化时所关注的序列。通常,异源多核苷酸侧接有至少一个(并且通常有两个)AAV反向末端重复序列(ITR)。术语rAAV载体涵盖rAAV载体颗粒和rAAV载体质粒两者。

[0384] 在另一个方面,提供了药物组合物,所述药物组合物包括

[0385] (a)根据本公开的任何实施例或实施例组合所述的核酸表达盒、表达载体、重组宿主细胞或rAAV颗粒;以及

[0386] (b)药学上可接受的载剂。

[0387] 在另一个方面,提供了配制物,所述配制物包括经过包装的病毒颗粒(如rAAV颗粒),所述经过包装的病毒颗粒包括根据本公开的任何实施例或实施例组合所述的核酸表达盒或表达载体。在一个实施例中,病毒颗粒的存在浓度为每mL至少10^{10}个含颗粒的载体基因组;在各个其它实施例中,病毒颗粒的存在浓度为每mL至少7.5×10^{10}个;10^{11}个;5×10^{11}个;10^{12}个;5×10^{12}个;10^{13}个;1.5×10^{13}个;3×10^{13}个;5×10^{13}个;$7.5\times9\times10^{13}$个;或9×10^{13}个含颗粒的载体基因组。配制物可以进一步包括药学上可接受的载剂、稀释剂和试剂。配制物可以呈液体溶液、糊剂、水凝胶的形式,或者可以包埋在底物(包含但不限于泡沫基质)内或支撑在容器中。

[0388] 对于视锥细胞在体内接触的情况,可以适当地处理核酸表达盒、表达载体、重组宿主细胞或rAAV颗粒,以递送至眼睛。病毒原液可以与可用于制备通常是安全、无毒且期望的并且包含可接受以供灵长类动物使用的赋形剂的配制物的药学上可接受的载剂、稀释剂和试剂组合。此类赋形剂可以是固体、液体、半固体或气体(在气溶胶组合物的情况下)。此类载剂或稀释剂的实例包含但不限于水、盐水、林格氏溶液、右旋糖溶液和5%人血清白蛋白。还可以将补充性活性化合物掺入配制物中。用于配制物的溶液或悬浮液可以包含:无菌稀释剂,如注射用水、盐水溶液、不挥发性油、聚乙二醇、甘油、丙二醇或其它合成溶剂;抗菌化合物,如苯甲醇或对羟基苯甲酸甲酯;抗氧化剂,如抗坏血酸或亚硫酸氢钠;螯合化合物,如乙二胺四乙酸(EDTA);缓冲剂,如乙酸盐、柠檬酸盐或磷酸盐;洗涤剂,如用于防止聚合的吐温(Tween)20;以及用于调节张力的化合物,如氯化钠或右旋糖。可以用酸或碱(如盐酸或氢氧化钠)调节pH。

[0389] 适合于在本公开中内部使用的药物组合物进一步包含无菌水溶液或分散液以及用于临时制备无菌可注射溶液或分散液的无菌粉末。在所有情况下,组合物是无菌的,并且

应当具有达到易于注射的程度的流动性。载剂可以是溶剂或分散介质，所述溶剂或分散介质含有例如水、乙醇、多元醇(例如，甘油、丙二醇和液体聚乙二醇等)及其合适的混合物。可以例如通过使用如卵磷脂等包衣、通过在分散情况下保持所需粒径以及通过使用表面活性剂来保持适当的流动性。可以通过各种抗细菌剂和抗真菌剂(例如，对羟苯甲酸酯、三氯叔丁醇、苯酚、抗坏血酸、硫柳汞等)来实现防止微生物的作用。在许多情况下，优选在组合物中纳入等渗剂，例如糖、多元醇(如甘露醇、山梨醇)、氯化钠。可以通过在组合物中纳入延迟吸收的试剂(例如，单硬脂酸铝和明胶)来实现内服组合物的延长吸收。

[0390] 在一个实施例中，将活性化合物与将保护这些化合物免于从体内快速消除的载剂一起制备，如控释配制物，包含植入物和微囊化的递送系统。可以使用可生物降解、生物相容的聚合物，如乙烯乙酸乙烯酯、聚酸酐、聚乙醇酸、胶原、聚原酸酯和聚乳酸。用于制备此类配制物的方法对本领域技术人员而言将是显而易见的。材料也可以商购获得。脂质体悬浮液(包含靶向受感染细胞的脂质体，其中单克隆抗体针对病毒抗原)也可以用作药学上可接受的载剂。这些可以根据本领域技术人员已知的方法来制备，例如，如在美国专利第4,522,811号中所描述的方法。

[0391] 尤其有利的是，将组合物调配成剂量单位形式，以便于施用以及剂量的均匀性。如本文所使用的，剂量单位形式是指适合作为针对有待治疗的受试者的单一剂量的物理上离散单位；每单位含有预定量的活性化合物，所述预定量的活性化合物经计算可与所需的药物载剂缔合产生所期望的治疗效果。本公开的剂量单位形式的规格由活性化合物的独特特性和待实现的特定治疗效果以及本领域中在混配这种活性化合物以用于治疗个体时固有的局限性决定并且直接依赖于此。

[0392] 药物组合物可以与施用说明书一起包含在容器、包装或分配器(例如注射器，例如载药注射器)中。本公开的药物组合物涵盖任何药学上可接受的盐、酯或这种酯的盐、或在施用于包括人的动物时能够(直接或间接)提供生物活性代谢物或其残留物的任何其它化合物。因此，例如，本公开还涉及本公开化合物的前药和药学上可接受的盐、此类前药的药学上可接受的盐以及其它生物等效物。术语“前药”表示以非活性形式制备的治疗剂，其通过内源酶或其它化学物质和/或条件的作用在体内或其细胞内转化为活性形式(即，药物)。术语“药学上可接受的盐”是指本公开化合物的生理学上可接受的盐和药学上可接受的盐：即，保留了亲本化合物的期望生物活性并且不对亲本化合物产生不期望的毒理效果的盐。

[0393] 用如碱金属和碱土金属或有机胺等金属或胺形成药学上可接受的碱加成盐。用作阳离子的金属包括钠、钾、镁、钙等。胺包括N-N'-二苄基乙二胺、氯普鲁卡因、胆碱、二乙醇胺、二环己胺、乙二胺、N-甲基葡糖胺和普鲁卡因(参见例如Berge等人，“药用盐(Pharmaceutical Salts)”，《药学杂志(J.Pharma Sci)》，1977,66,119)。酸性化合物的碱加成盐是通过使游离酸形式与足量的期望的碱接触以常规方式产生盐来制备的。游离酸形式可以通过使盐形式与酸接触并以常规方式分离游离酸来再生。游离酸形式在某些物理性质(如在极性溶剂中的溶解度)方面与其各自的盐形式略有不同，但是出于本公开的目的，盐以其它方式等同于其各自的游离酸。

[0394] 在另一个方面，本公开提供了用于在视锥细胞中表达如蛋白质等基因产物的方法，所述方法包括使一个或多个视锥细胞与有效量的根据本公开的任何实施例或实施例组合所述的核酸表达盒、重组宿主细胞、rAAV、药物组合物或配制物接触，其中在所述一个或

多个视锥细胞中以可检测的水平表达由所述编码区编码的所述基因产物，如蛋白质。

[0395] 在另一个方面，本公开提供了用于在需要治疗或预防视锥细胞病症的哺乳动物中治疗或预防视锥细胞病症的方法，所述方法包括向所述哺乳动物的眼睛施用有效量的根据本公开的任何实施例或实施例组合所述的核酸表达盒、重组宿主细胞、rAAV、药物组合物或配制物，其中所述编码区包括编码治疗性基因产物的核酸序列。

[0396] 术语“治疗(treatment、treating)”等在本文中通常用于意指获得期望的药理学和/或生理学效果。效果在完全或部分地预防疾病或其症状方面可以是预防性的，例如降低受试者体内发生疾病或其症状的可能性，和/或在部分或完全治愈疾病和/或由疾病引起的不良反应方面可以是治疗性的。如本文所使用的，“治疗”覆盖了对哺乳动物的疾病的任何治疗，并且包含：(a)防止疾病在可能倾向于患有疾病但尚未被诊断为患有疾病的受试者身上发生；(b)抑制疾病，即，阻止其发展；或者(c)缓解疾病，即，使疾病消退。可以在疾病或损伤发作之前、期间或之后施用治疗剂。特别关注了对发展中的疾病的治疗，其中所述治疗稳定或减少了患者的不期望的临床症状。理想的是，在受影响组织的功能完全丧失之前执行这种治疗。理想的是，将在疾病的症状阶段期间并且在一些情况下在疾病的症状阶段之后施用主题疗法。

[0397] 术语“个体”、“宿主”、“受试者”和“患者”在本文中可互换使用，并且是指哺乳动物，例如啮齿动物(例如小鼠、大鼠、沙鼠)、兔子、猫、犬、山羊、绵羊、猪、马、牛或灵长类动物。在某些实施例中，受试者是狭鼻小目(Parvorder Catarrhini)的灵长类动物。如本领域中已知的，狭鼻小目是高等灵长类动物的两个细分之一(另一个是新世界猴)，并且包含旧世界猴和猿，猿又被进一步划分成较小的猿或长臂猿和巨猿，所述猿由猩猩、大猩猩、黑猩猩、倭黑猩猩和人组成。在另外的优选实施例中，灵长类动物是人。

[0398] 在各个实施例中，所述视锥细胞病症选择由以下组成的组：黄斑营养不良、色觉障碍或中央黄斑视力障碍。在其它实施例中，所述视锥细胞病症选自由以下组成的组：全色盲、蓝视锥单色视、红绿色盲、红色觉缺陷、绿色觉缺陷、蓝色觉缺陷、黄斑营养不良，如斯特格氏黄斑营养不良、视锥营养不良、视锥-视杆营养不良、X连锁视锥营养不良、7型脊髓小脑性共济失调和巴尔得-别德尔氏综合征1，年龄相关性黄斑变性、黄斑毛细管扩张症、视网膜色素变性、糖尿病性视网膜病变、视网膜静脉阻塞、青光眼、索斯比氏眼底营养不良、成人型卵黄样黄斑营养不良、贝斯特氏病、视杆-视锥营养不良、莱伯氏先天性黑蒙症和X连锁视网膜劈裂症、博恩霍姆氏眼病和伴有近视的X连锁视锥功能障碍综合征。在另外的实施例中，所述视锥细胞病症选自由以下组成的组：红绿色盲、蓝视锥单色视、博恩霍姆氏眼病、伴有近视的X连锁视锥功能障碍综合征和X连锁视锥营养不良、X连锁视网膜劈裂症、常染色体隐性遗传性视网膜色素变性和年龄相关性黄斑变性(AMD;湿性或干性)。

[0399] 为了促进基因产物的表达，根据本文的教导，可以在使用者认为适当时施用组合物。可以将组合物提供给受试者一次或多次，例如一次、两次、三次或三次以上。通常，提供有效量的组合物以在细胞中产生基因产物的表达。有效量可以容易地凭经验确定，例如通过检测基因产物基因产物的存在或水平、通过检测对视锥细胞的生存力或功能的影响等。通常，与本领域已知的相同量的多核苷酸盒(例如pR2.1(SEQ ID NO:50的核苷酸1-2274)、pR1.7、pR1.5、pR1.1或IRBP/GNAT2盒)相比，有效量的组合物将促进视锥细胞中基因产物的更大表达。通常，相对于来自本领域已知的参照或对照多核苷酸盒的表达，表达将被增强2

倍或更多倍，例如3倍、4倍或5倍或更多倍，在一些情况下，10倍、20倍或50倍或更多倍，例如100倍。

[0400] 可以通过任何合适的方法将组合物施用于受试者的视网膜。例如，可以通过玻璃体内注射或视网膜下注射眼内施用组合物。用于通过玻璃体内注射或通过视网膜下注射递送载体的一般性方法可以通过以下简要概述来说明。这些实例仅旨在说明方法的某些特征，而绝不意味着是限制性的。

[0401] 对于视网膜下施用，可以在使用手术显微镜的直接观察下以视网膜下注射的悬浮液形式递送组合物。此过程可能涉及玻璃体切除术，然后使用细套管通过一个或多个小视网膜切开术将组合物悬浮液注入视网膜下腔。简而言之，可以在整个手术过程中将输注套管缝合在适当的位置，以通过输注(例如输注盐水)来维持正常的球体容积。使用适当孔径(例如20到27号规格)的套管进行玻璃体切除术，其中通过从输注套管中输注盐水或其它等渗溶液来替换被去除的玻璃体凝胶的体积。进行玻璃体切割术是有利的，因为(1)去除其皮层(后透明膜)便于通过套管穿透视网膜；(2)将其去除并用液体(例如生理盐水)替代会产生用于容纳眼内注射载体的空间，并且(3)对其进行有控制的去除会降低视网膜撕裂和计划外的视网膜脱离的可能性。

[0402] 对于玻璃体内施用，可以呈悬浮液的形式递送组合物。最初，将局部麻醉剂施加到眼睛表面，然后再施加局部防腐剂溶液。在有或没有仪器的情况下，使眼睛保持张开，并且在直接观察下将组合物用短而窄的(例如30号)针头通过巩膜注射到受试者的眼玻璃体腔中。玻璃体内施用通常被很好地耐受。手术结束时，注射位点处有时会存在轻微发红。偶尔会有压痛，但大多数患者没有报告任何疼痛。此手术后，无需眼贴或眼罩，并且活动不受限制。有时，开了几天的抗生素滴眼剂有助于预防感染。

[0403] 本公开的方法和组合物可用于治疗可以至少部分地通过视锥光感受器细胞的基因疗法解决的任何病状。因此，本公开的组合物和方法可用于治疗需要视锥细胞治疗的个体。需要视锥细胞治疗的人是指患有视锥细胞病症或具有患视锥细胞病症风险的个体。“视锥细胞病症”是指影响视网膜视锥细胞的任何病症，包含但不限于与视锥细胞内的缺陷(即视锥细胞内在的缺陷)相关的眼睛视力障碍，例如黄斑营养不良，如斯特格氏黄斑营养不良、视锥营养不良、视锥-视杆营养不良、7型脊髓小脑性共济失调和巴尔得-别德尔氏综合征1；以及色觉障碍，包含全色盲、不完全性全色盲、蓝视锥单色视以及红色觉缺陷、绿色觉缺陷和蓝色觉缺陷；以及可以通过靶向视锥细胞进行治疗的中央黄斑(灵长类动物)视力障碍，例如年龄相关性黄斑变性、黄斑毛细管扩张症、视网膜色素变性、糖尿病性视网膜病变、视网膜静脉阻塞、青光眼、索斯比氏眼底营养不良、成人型卵黄样黄斑营养不良、贝斯特氏病、视杆-视锥营养不良、莱伯氏先天性黑蒙症和X连锁视网膜劈裂症。

[0404] 斯特格氏黄斑营养不良。斯特格氏黄斑营养不良也被称为斯特格氏病和眼底黄色斑点症，是通常导致进行性视力丧失，甚至达到法定盲点的青少年型黄斑变性的一种遗传形式。症状的发作通常出现在六至三十岁的年龄之间(平均约16-18岁)。包含ABCA4、CNGB3、ELOVL4、PROM1的几种基因的突变与所述病症相关。症状通常在二十岁时出现，并且包含波浪形视觉、盲点、模糊、色觉受损和难以适应昏暗的照明。斯特格氏病的主要症状是视觉敏锐度下降，其范围为20/50到20/200。另外，患有斯特格氏病的人对眩光敏感；阴天提供了一些缓解。当黄斑受损时，视觉最明显受损，这可以通过眼底检查观察到。

[0405] 视锥营养不良。视锥营养不良(COD)是一种遗传性眼病,其特征在于视锥细胞的丢失。视锥营养不良最常见的症状是视力觉丧失(发病年龄范围为十几岁到六十岁)、对强光敏感和色觉差。视觉敏锐度通常会逐渐退化,但其可以迅速退化到20/200;后来,在更严重的情况下,其下降到"数指"视觉。使用颜色测试板(HRR系列)进行的色觉测试揭示了红绿色板和蓝黄色板两者上的许多错误。据信营养不良是原发性的,因为在观察到镜检变化之前发现了视锥功能的主观和客观异常。然而,视网膜色素上皮(RPE)被迅速波及,从而导致视网膜营养不良,其主要涉及黄斑。在视锥营养不良的早期,通过检眼镜进行的眼底检查基本上是正常的,并且明确的黄斑改变通常在视觉丧失后很容易发生。在检眼镜检查期间看到的最常见的黄斑病变类型具有牛眼外观,并且由围绕中央较暗区域的萎缩性色素上皮甜甜圈样区域(doughnut-like zone)组成。在另一种较不常见的视锥营养不良形式中,后极发生弥漫性萎缩,其中黄斑区有斑点色素团块。在早期患者中很少见脉络膜毛细血管萎缩和较大的脉络膜血管萎缩。荧光素血管造影术(FA)在疑似患有视锥营养不良的人的诊断检查中是有用的辅助手段,因为其可以检测出很难用检眼镜看到的视网膜的早期变化。由于眼底变化范围广且早期难以诊断,因此视网膜电描记法(ERG)仍然是用于进行诊断的最佳测试。在光线充足的房间内进行测试时(适应光的ERG),通过降低的单闪和闪烁响应指示ERG上的视锥功能异常。包含GUCA1A、PDE6C、PDE6H和RPGR的几种基因的突变与所述病症相关。

[0406] 视锥-视杆营养不良。视锥-视杆营养不良(CRD,或CORD)是一种属于色素性视网膜病变的组的遗传性视网膜营养不良。CRD的特征在于眼底检查可见视网膜色素沉积、主要定位于黄斑区域以及视锥细胞和视杆细胞均丢失。与视杆光感受器的原发性丢失和随后的视锥光感受器的继发性丢失所导致的视杆-视锥营养不良(RCD)相比,CRD反映了相反的事件顺序:原发性视锥参与,或者有时通过视锥和视杆两者同时丧失。症状包含视觉敏锐度下降、色觉缺陷、光厌恶和中央视野的敏感性降低,随后是周围视觉的进行性丧失和夜盲症。包含ADAM9、PCDH21、CRX、GUCY2D、PITPNM3、PROM1、PRPH2、RAX2、RIMS1、RPGR和RPGRIP1的几种基因的突变与所述病症相关。

[0407] 7型脊髓小脑性共济失调。脊髓小脑性共济失调是一种进行性、退化性、遗传性疾病,其特征在于步态缓慢进行性不协调,并且通常与手、言语和眼球运动协调性差相关。SCA有多种类型,其中7型脊髓小脑性共济失调(SCA-7)与大多数其它SCA有所不同,不同之处在于除了协调差之外,还会出现视觉问题。SCA-7与ATXN7/SCA7基因中的常染色体显性突变相关。当疾病在40岁之前表现出来时,所述疾病的最早征兆通常是视觉问题而不是协调性差。早期症状包含难以区分颜色和中央视觉下降。另外,共济失调的症状(不协调、眼球移动缓慢以及感觉或反射的轻度变化)可能是可检测的。随着疾病的进展,失去运动控制、言语不清和吞咽困难变得更加突出。

[0408] 巴尔得-别德尔氏综合征1。巴尔得-别德尔氏综合征1(BBS-1)是一种具有在家庭内部和家庭之间均观察到的可变的表达性和广泛的临床变异性的多效性病症。主要临床特征是视杆-视锥营养不良,伴有儿童期发作的视觉丧失和之前的夜盲症;轴后多指;婴幼儿时期显现的躯干肥胖并且到成年后仍然存在问题;在一些而非全部个体中存在特定的学习困难;男性生殖腺发育不全和复杂的女性泌尿生殖系统畸形;以及作为发病率和死亡率的主要原因的肾功能障碍。视力丧失是巴尔得-别德尔氏综合征的主要特征之一。夜间视觉的问题在儿童中期就变得很明显,随后在周边视觉中出现盲点。随着时间的流逝,这些盲点会

扩大并合并,从而产生隧道视觉。大多数患有巴尔得-别德尔氏综合征的人也会发展出中央视觉模糊(视觉敏锐度差),并在青春期或成年早期法定失明。巴尔得-别德尔氏综合征可能是由于至少14种不同的基因(通常被称为BBS基因)发生突变而引起的,这些基因已知或怀疑在纤毛功能中起关键作用,其中最常见的是BBS1和BBS10的突变。

[0409] 全色盲。全色盲或视杆单色视是一种这样的病症:受试者完全缺乏对颜色的感知,使得受试者只能看到黑色、白色和灰色阴影。其它症状包含视觉敏锐度下降、畏光、眼球震颤、小的中央巩膜瘤和偏心注视。通常在六个月左右大的孩子中通过其畏光活动和/或其眼球震颤而首先注意到所述病症。在生命的前6-7年期间,视觉敏锐度和眼球运动的稳定性通常会提高(但仍保持在20/200附近)。CNGB3、CNGA3、GNAT2、PDE6C和PDE6HI的突变与所述病症相关。

[0410] 不完全性全色盲。不完全性全色盲类似于全色盲,但渗透率较低。在不完全性全色盲中,除了形式减少以外,其症状与完全性全色盲的症状类似。患有不完全性全色盲的个体视觉敏锐度下降,伴有或不伴有眼球震颤或畏光。此外,这些个体仅示出视锥细胞功能的部分损伤,但仍保留了视杆细胞功能。

[0411] 蓝视锥单色视。蓝视锥(S视锥)单色视(BCM)是一种罕见的X连锁先天性固定视锥功能障碍综合征,大约每100,000人中有1人受影响。患有BCM的受影响男性在视网膜中没有功能性长波长敏感性(L)视锥或中等波长敏感性(M)视锥,这是由于L和M-视蛋白基因的基因座处的突变。出生起颜色辨别力被严重损害,并且视觉源自保留下来的其余S视锥和视杆光感受器。BCM通常表现为视觉敏锐度下降(6/24到6/60)、摆动性眼球震颤、畏光,并且患者经常患有近视。视杆特异性和最大视网膜电图(ERG)通常显示出明确异常,而30Hz的视锥ERG无法被检测到。单闪光明视ERG经常是可记录的,尽管很小而且很晚,并且S视锥的ERG保存完好。

[0412] 颜色视觉缺陷。颜色视觉缺陷(CVD)或色盲是在正常照明条件下无法看见颜色或感知色差的能力或看见颜色或感知色差的能力下降。可以使用多种颜色视觉测试中的任何一种来鉴别患有色盲的个体,例如,颜色ERG(cERG)、伪等色板(石原板(Ishihara plates),哈迪-兰特-里特多色板(Hardy-Rand-Ritter polychromatic plates))、法斯沃斯-蒙赛尔100色相测试(Farnsworth-Munsell 100hue test)、法恩斯沃思氏板(Farnsworth's panel)D-15、城市大学测试、科尔纳氏法则(Kollner's rule)等。颜色视觉缺陷的实例包含红色觉缺陷、绿色觉缺陷和蓝色觉缺陷。红色觉缺陷包含红色盲(对红光不敏感)和红色弱(对红光的敏感性降低),并且与L-视蛋白基因(OPN1LW)的突变相关。绿色觉缺陷包含绿色盲(对绿光不敏感)和绿色弱(对绿光的敏感性降低),并且与M-视蛋白基因(OPN1MW))的突变相关。蓝色觉缺陷包含蓝色盲(对蓝光不敏感)和蓝色弱(对蓝光的敏感性降低),并且与S-视蛋白基因(OPN1SW)的突变相关。

[0413] 年龄相关性黄斑变性。年龄相关性黄斑变性(AMD)是50岁年龄以上人群视力丧失的主要原因之一。AMD主要影响中央视觉,所述中央视觉是如阅读、驾驶和识别面部等详细任务所需的。在这种病状下的视力丧失是由黄斑中的光感受器逐渐退化引起的。侧面(周边)视觉和夜晚视觉通常不会受到影响。

[0414] 研究人员已描述了年龄相关性黄斑变性的两种主要类型,被称为干性或“非渗出性”形式和湿性或“渗出性”或“新生血管性”形式,这两种形式都可以通过在主题多核苷酸

盒的背景下递送基因产物来进行治疗。

[0415] 干性AMD的特征在于视网膜色素上皮与黄斑下方脉络膜之间被称为玻璃疣的黄色沉积物堆积，这可以通过眼底照相术观察到。这导致视觉的缓慢进行性丧失。所述病状通常会影响两只眼睛的视觉，但视力丧失通常在一只眼中先于另一只眼发生。其它变化可以包含色素变化和RPE萎缩。例如，在某些情况(被称为中央地图样萎缩或“GA”)下，观察到视网膜色素上皮萎缩并且随后在眼中部失去光感受器。干性AMD与CD59和补体级联中的基因的突变相关。

[0416] 湿性AMD是干性AMD的进展状态，在约10％的干性AMD患者中发生。病理变化包含视网膜色素上皮细胞(RPE)功能障碍、在RPE下的液体收集以及黄斑区中的脉络膜新血管形成(CNV)。在严重情况下，可能会发生流体渗漏、RPE或神经视网膜脱离以及血管破裂出血。湿性AMD的症状可以包含视觉失真，如直线出现波浪形或弯曲，门口或路牌看起来偏斜，物体看起来比实际更小或更远；中央视觉下降；颜色的强度或亮度降低；以及视野中非常明确的模糊点或盲点。可能突然发作并迅速恶化。诊断可以包含：使用阿姆斯勒方格表(Amsler grid)测试受试者的中央视觉的缺陷(黄斑变性可能导致方格中的直线出现褪色、折断或扭曲)；使用荧光素血管造影术来观察血管或视网膜异常；以及使用光学相干断层扫描来检测视网膜肿胀或血管渗漏。CNV的产生涉及多个细胞因子，其中包括血管内皮生长因子(VEGF)、血小板衍生生长因子(PDGF)、色素上皮衍生因子(PEDF)、缺氧诱导因子(HIF)、血管生成素(Ang)和其它细胞因子、促分裂原活化蛋白激酶(MAPK)等。

[0417] 黄斑毛细血管扩张症。黄斑毛细血管扩张症(MacTel)是在黄斑中央凹区进行病理性扩张的血管(毛细血管扩张症)的形式。由于充液囊肿的发展，组织恶化，并且视网膜结构变得瘢痕性，从而损害了光感受器的营养并永久破坏了视觉。MacTel有两种类型，1型和2型。2型黄斑性毛细血管扩张症是一种双侧疾病，最近示出其患病率在40岁及40岁以上的人群中高达0.1％。生物显微镜检查可能示出视网膜透明度降低、晶体沉积、轻度扩张的毛细血管、微弱的小静脉、视网膜色素斑、中央凹萎缩和新生血管复合物。荧光素血管造影术显示，毛细血管扩张性毛细血管在早期主要暂时位于中央凹，而在后期存在弥漫性高荧光。高分辨率光学相干断层扫描(OCT)可能显示光感受器内部区段-外部区段边界的破坏、内部或外部视网膜水平处的反射不良腔以及后期视网膜的萎缩。在1型黄斑毛细血管扩张症中，所述疾病几乎总是在一只眼睛中发生，这使其与2型区别开来。虽然MacTel通常不会导致完全失明，但其通常会导致阅读和驾驶视觉所需的中央视力的丧失10-20年。

[0418] 视网膜色素变性。视网膜色素变性(RP)是一组遗传性病症，其特征在于进行性周边视力丧失和可能导致中央视力丧失的夜视困难(夜盲)。RP的表现体征和症状各不相同，但经典的表现体征和症状包含夜盲(夜盲症，最常见的RP的最早症状)；视觉丧失(通常为周边视觉丧失，但在晚期病例中，为中央视觉丧失)；以及闪光幻视(看见闪光)。由于RP是许多遗传疾病的集合，因此物理发现存在显著的变异性。眼部检查涉及对视觉敏锐度和瞳孔反应的评估，以及眼前节、视网膜和眼底镜评估。在一些情况下，RP是以下综合症的一个方面：例如也与听力丧失相关的综合征(乌谢尔综合征、瓦登伯格综合征、奥尔波特综合征、雷夫叙姆氏病)；卡恩斯-塞尔氏综合征(外部眼肌麻痹、眼睑下垂、心脏传导阻滞和色素性视网膜病)；无β脂蛋白血症(脂肪吸收不良、脂溶性维生素缺乏症、脊髓小脑变性和色素性视网膜变性)；粘多糖酶(例如，贺勒氏综合征、施艾氏综合征、圣菲利柏氏症)；巴尔得-别德尔氏

综合征(多发性、截断性肥胖、肾功能障碍、身材矮小和色素性视网膜病);以及神经元类脂褐藻病(痴呆、癫痫和色素性视网膜病;婴儿型被称为詹-比二氏病(Jansky-Bielschowsky disease),青少年型为沃格特-施皮尔迈尔-巴顿氏病(Vogt-Spielmeyer-Batten disease),并且成人型为库夫斯氏综合征)。视网膜色素变性最常见与RHO、RP2、RPGR、RPGRIP1、PDE6A、PDE6B、MERTK、PRPH2、CNGB1、USH2A、ABCA4、BBS基因的突变相关。

[0419] 糖尿病性视网膜病变。糖尿病性视网膜病变(DR)是由糖尿病并发症引起的视网膜损伤,其最终可能导致失明。不希望受理论的束缚,据信高血糖症引起的壁内周细胞死亡和基底膜增厚会导致血管壁功能不全。这些损害改变了血-视网膜屏障的形成,并且还使视网膜血管变得更加可渗透。

[0420] 糖尿病性视网膜病变分为两个阶段:非增生性糖尿病性视网膜病变(NPDR)和增生性糖尿病性视网膜病变(PDR)。非增生性糖尿病性视网膜病变是糖尿病性视网膜病变的第一阶段,并通过眼底镜检查和并存的糖尿病进行诊断。在视觉下降的情况下,可以进行荧光素血管造影术以可视化眼后部的血管以及可能存在的任何视网膜缺血。所有糖尿病患者都有患NPDR的风险,并且因此,将成为使用主题载体进行预防性治疗的候选人。增生性糖尿病性视网膜病变是糖尿病性视网膜病变的第二阶段,其特征在于视网膜的新血管形成、玻璃体出血和视力模糊。在一些情况下,纤维血管增生导致牵引性视网膜脱离。在一些情况下,血管也可以长到眼前房的角中,并导致新生血管性青光眼。患有NPDR的个体患PDR的风险增加,并且因此,将成为使用主题载体进行预防性治疗的候选人。

[0421] 糖尿病性黄斑水肿。糖尿病性黄斑水肿(DME)是糖尿病性视网膜病变的一种晚期视觉受限并发症,其影响了将近30%患有糖尿病至少20年的患者,并且是由DR导致的大部分视力丧失的原因。糖尿病性黄斑水肿是由视网膜微血管变化所致,所述视网膜微血管变化损害了血-视网膜屏障,从而导致血浆成分泄漏到周围的视网膜中,进而导致视网膜水肿。不希望受理论的束缚,据信高血糖症、细胞信号传导途径的持续改变以及伴有白细胞介导的损伤的慢性微血管发炎会导致慢性视网膜微血管损伤,这触发了眼内VEGF水平的升高,其反过来又增加脉管系统的渗透性。

[0422] 具有患DME风险的患者包含患有糖尿病的时间较长的患者以及患有严重的高血压病(高血压)、液体潴留、白蛋白血症或高脂血症中的一种或多种的患者。DME的常见症状是视觉模糊、漂浮物、双重视觉和最终失明(如果不进行治疗而任疾病进展)。通过眼底镜检查将DME诊断为黄斑中心2个椎间盘直径内的视网膜增厚。可以采用的其它方法包含:用于检测视网膜肿胀、囊样水肿和浆液性视网膜脱离的光学相干断层扫描(OCT);荧光素血管造影术,其区分和定位局灶性和弥散性渗漏区域,从而在使用激光光凝术治疗水肿时指导激光光凝术的施加;以及彩色立体眼底照片,其可以用于评估视网膜的长期变化。还可以测量视觉敏锐度,特别是用于跟踪黄斑水肿的进展并在施用主题药物组合物后观察其治疗。

[0423] 视网膜静脉阻塞。视网膜静脉阻塞(RVO)是排出血液视网膜的循环部分的阻塞。阻塞可以导致毛细血管中的后备压力,所述后备压力可能导致出血以及流体和其它血液成分的泄漏。

[0424] 青光眼。青光眼是描述一组导致视神经损伤的眼(眼睛)病症的术语,其通常与眼中的流体压力(眼内压)(IOP)增加相关。所述病症可以大致划分成两大类,“开角型”青光眼和“闭角型(closed-angle或angle closure)”青光眼。开角型青光眼占美国青光眼病例的

90％。开角型青光眼无痛且没有急性发作。唯一的迹象是逐渐进行性视野丧失和视神经改变(眼底镜检查时杯盘比增加)。在美国,闭角型青光眼占青光眼病例的不到10％,但在其它国家(尤其是亚洲国家)中,闭角型青光眼占青光眼病例的一半。约有10％的闭角患者出现急性闭角危机,其特征在于突然的眼痛、见光晕、红眼、眼压过高(>30mmHg)、恶心和呕吐、视力突然下降以及固定的轻微散瞳。在一些情况下,闭角型青光眼还与椭圆形的瞳孔相关。调节由DLK、NMDA、INOS、CASP-3、Bcl-2或Bcl-xl编码的蛋白质的活性可以治疗所述病状。

[0425] 索斯比氏眼底营养不良。索斯比氏眼底营养不良是一种与TIMP3基因的突变相关的常染色体显性遗传性视网膜疾病。临床上,发现了早期外周中部玻璃疣和色觉缺陷。一些患者抱怨夜盲症。最常见地,主要症状是由于无法治疗的黄斑下新生血管所导致的突然的敏锐度下降,在生命的第三个十年到第四个十年中表现出来。从组织学上讲,在布鲁赫膜的水平处存在含有30μm厚材料的融合脂质的聚集体。

[0426] 卵黄样黄斑营养不良。卵黄样黄斑营养不良是一种可以导致进行性视力丧失的遗传性眼部病症。卵黄样黄斑营养不良与黄斑下面的细胞中脂肪黄色素(脂褐素)的积累相关。随着时间的流逝,此物质的异常累积会损害对于清晰的中央视觉至关重要的细胞。因此,患有此病症的人常常失去中央视觉,并且其视力可能变得模糊或扭曲。卵黄样黄斑营养不良通常不影响侧面(周边)视觉或夜间视力。研究人员已描述了具有类似特征的卵黄样黄斑营养不良的两种形式。早发形式(被称为贝斯特氏病)通常出现在儿童时期;症状的发作和视力丧失的严重程度差异很大。其与VMD2/BEST1基因的突变相关。成人发作形式(成人型卵黄样黄斑营养不良)开始较晚,通常在成年中期开始,并倾向于导致随着时间的流逝慢慢恶化的视力丧失。其与PRPH2基因的突变相关。卵黄样黄斑营养不良的两种形式均具有可以在眼睛检查期间检测到的黄斑的特征性变化。

[0427] 视杆-视锥营养不良。视杆-视锥营养不良是一类进行性疾病,其中导致夜盲症和周边视野扩大丧失的视杆功能障碍是普遍存在的问题,或者至少与视锥功能障碍一样严重。可以在视网膜的视网膜赤道部看见扇贝状的腔隙性萎缩。尽管在所有病例中均可见视网膜中央变薄,但通过临床检查,黄斑仅轻度波及。色觉障碍早期较轻,并且通常会变得更加严重。视野是中等收缩到严重收缩的,尽管在较年轻的个体中存在典型的环型暗点。周围视网膜含有“白点”,并且通常类似于在白点状视网膜炎中看到的视网膜变化。视网膜色素变性是此定义下包含的疾病的主要分组,总体上估计每3,500人中就有大约一人受到影响。根据所使用的分类标准,所有视网膜色素变性患者中约有60-80％的患者具有清晰的视网膜疾病的视杆-视锥营养不良模式,并且一旦考虑到其它症状,则所有视网膜色素变性中约有50-60％的视网膜色素变性落入视杆-视锥营养不良的非综合征型类别中。

[0428] 莱伯氏先天性黑蒙症。莱伯氏先天性黑蒙症(LCA)是一种严重的视网膜营养不良,其通常在生命的第一年变得明显。视觉功能通常较差,并且经常伴有眼球震颤、迟钝或几乎没有瞳孔反应、畏光、高度远视和圆锥形角膜。视觉敏锐度很少超过20/400。一个特征性发现是弗兰切斯凯蒂氏眼动数字体征(Franceschetti’s oculo-digital sign),其包括戳眼、按压和摩擦。眼底的外观变化很大。虽然视网膜最初可能看起来是正常的,但在童年后期经常观察到色素性视网膜病,其使人联想起视网膜色素变性。视网膜电图(ERG)是典型地“非可检测的”或严重低于正常的。已知17个基因的突变会导致LCA:GUCY2D(基因座名称:LCA1)、RPE65(LCA2)、SPATA7(LCA3)、AIPL1(LCA4)、LCA5(LCA5)、RPGRIP1(LCA6)、CRX

(LCA7)、CRB1(LCA8)、NMNAT1(LCA9)、CEP290(LCA10)、IMPDH1(LCA11)、RD3(LCA12)、RDH12(LCA13)、LRAT(LCA14)、TULP1(LCA15)、KCNJ13(LCA16)和IQCB1。总共地,这些基因的突变估计占所有LCA诊断的一半以上。已经报道了LCA的至少一种其它疾病基因座,但所述基因是未知的。

[0429] X连锁视网膜劈裂症。X连锁视网膜劈裂症(XLRS)的特征在于对称性双侧黄斑参与,其发作于生命的第一个十年,在一些情况下,甚至早在三个月大时发作。眼底检查示出了黄斑中的裂隙区域(视网膜神经纤维层分裂),有时给出辐轮图案的印象。个体中大约50%的个体出现神经周边视网膜裂痕(主要在颞下)。受影响的男性通常视力为20/60到20/120。视觉敏锐度通常在生命的第一个十年和第二个十年期间恶化,但是直到第五个十年或第六个十年一直保持相对稳定。X连锁青少年型视网膜劈裂症的诊断基于眼底发现、电生理测试结果和分子遗传学测试。RS1是唯一已知与X连锁青少年型视网膜劈裂症相关的基因。

[0430] 可以使用本领域已知的用于检测病症症状的技术来容易地鉴别患有视锥细胞病症或具有患视锥细胞病症风险的个体,所述技术包含但不限于:眼底照相术;光学相干断层扫描(OCT);自适应光学(AO)检眼镜/眼底镜检查;视网膜电描记法,例如ERG、颜色ERG(cERG);颜色视觉测试,如伪等色板(石原板、哈迪-兰特-里特多色板)、法斯沃斯-蒙赛尔100色相测试、法恩斯沃思氏板D-15、城市大学测试、科尔纳氏法则等;以及视觉敏锐度测试,如ETDRS字母测试、斯内伦(Snellen)视觉敏锐度测试等;如普通技术人员将知道的。另外或可替代地,可以使用本领域已知的用于检测与视锥细胞病症相关的基因突变的技术来容易地鉴别受视锥细胞病症影响或具有患视锥细胞病症风险的个体,所述技术包含但不限于PCR、DNA序列分析、限制性消化、Southern印迹杂交、质谱分析法等。在一些实施例中,所述方法包括鉴别需要视锥细胞治疗的个体的步骤。在这种情况下,可以利用例如通过检测本文所描述或本领域已知的症状、通过检测如本文所述或本领域已知的基因的突变等确定个体是否患有视锥细胞病症的一种或多种症状或具有患视锥细胞病症风险的任何便利的方法来鉴别需要视锥细胞治疗的个体。

[0431] 在实践受试者方法中,通常以有效导致视锥细胞中基因产物表达的量将组合物递送至受试者的视网膜。在一些实施例中,所述方法包括检测视锥细胞中的基因产物表达的步骤。在某些实施例中,所述方法包括检测视锥细胞中的基因产物表达的步骤。检测基因产物表达的方法有多种,其中任何一种都可以用于主题实施例中。例如,可以直接检测表达,即通过测量基因产物的量,例如:在RNA水平下测量基因产物的量,例如通过RT-PCR、Northern印迹、RNA酶保护;或在蛋白质水平下测量基因产物的量,例如通过Western印迹、ELISA、免疫组织化学等。作为另一个实例,可以间接检测表达,即通过检测基因产物对受试者的视锥光感受器的生存力或功能的影响。例如,如果由基因产物编码的基因产物改善了视锥细胞的生存力,则可以通过检测视锥细胞的生存力的改善来检测基因产物的表达,例如通过眼底照相术、光学相干断层扫描(OCT)、自适应光学(AO)检眼镜/眼底镜检查等。如果由基因产物编码的基因产物改变了视锥细胞的活性,则可以通过检测视锥细胞的活性的变化来检测所述基因产物的表达,例如通过视网膜电图(ERG)和颜色ERG(cERG);颜色视觉测试,如伪等色板(石原板、哈迪-兰特-里特多色板)、法斯沃斯-蒙赛尔100色相测试、法恩斯沃思氏板D-15、城市大学测试、科尔纳氏法则等;以及视觉敏锐度测试,如ETDRS字母测试,斯内伦(Snellen)视觉敏锐度测试等,作为检测所递送的多核苷酸的存在的一种方式。在一

些情况下,可以检测生存力的改善和视锥细胞功能的修饰。

[0432] 在一些实施例中,所述方法产生了治疗益处,例如预防病症的发展、停止病症的进展、逆转病症的进展等。在一些实施例中,主题方法包括检测到已经实现了治疗益处的步骤。普通技术人员应了解,治疗功效的这种量度将适用于被修饰的特定疾病,并且应认识到用于测量治疗功效的适当的检测方法。例如,可以将治疗黄斑变性的治疗功效观察为黄斑变性的速率降低或黄斑变性的进展停止,通过比较施用主题组合物之后的测试结果与施用主题组合物之前的测试结果,其效果可以通过例如眼底照相术、OCT或AO观察到。作为另一个实例,可以将治疗进行性视锥功能障碍的治疗功效观察为视锥功能障碍进展速度的降低、视锥功能障碍进展的停止或视锥功能的改善,其效果可以通过例如ERG和/或cERG;颜色视觉测试;和/或视觉敏锐度测试观察到,例如,通过比较施用主题组合物之后的测试结果与施用主题组合物之前的测试结果以及检测视锥细胞生存力和/或功能的变化。作为第三个实例,可以将治疗色觉缺陷的治疗功效观察为个体对颜色的感知改变,例如对红色波长的感知改变、对绿色波长的感知改变、对蓝色波长的感知改变,其效果可以通过例如cERG和颜色视觉测试观察到,例如通过比较施用主题组合物之后的测试结果与施用主题组合物之前的测试结果以及检测视锥细胞生存力和/或功能的变化。

[0433] 在一些情况下,在施用主题组合物后2周或更长时间(例如3周、4周、5周或6周或更长时间,例如2个月或更长时间,例如4个月、6个月、8个月或10个月或更长时间,在一些情况下为1年或更长时间,例如2年、3年、4年或5年,在某些情况下为5年以上)将观察到表达的变化。在一些情况下,在施用主题rAAV后2周或更长时间(例如3周、4周、5周或6周或更长时间,例如2个月或更长时间,例如4个月、6个月、8个月或10个月或更长时间,在一些情况下为1年或更长时间,例如2年、3年、4年或5年,在某些情况下为5年以上)将观察到治疗效果。

[0434] 通常,如果主题组合物为包括本公开的主题多核苷酸盒的rAAV,则用于实现变化的有效量将为主题rAAV的约1×10^{4}个载体基因组或更多个,例如1×10^{5}个载体基因组或更多个,例如1×10^{6}个、1×10^{7}个、1×10^{8}个载体基因组或更多个,在一些情况下,1×10^{9}个、1×10^{10}个、1×10^{11}个、1×10^{12}或1×10^{13}个载体基因组或更多个,在某些情况下,1×10^{14}个载体基因组或更多个,并且通常不多于1×10^{15}个载体基因组。在一些情况下,所递送的载体基因组的量至多为约1×10^{15}个载体基因组,例如1×10^{14}个载体基因组或更少,例如1×10^{13}个、1×10^{12}个、1×10^{11}个、1×10^{10}个或1×10^{9}个载体基因组或更少,在某些情况下,1×10^{8}个、1×10^{7}个、1×10^{6}个或1×10^{5}个载体基因组或更少,并且通常不少于1×10^{4}个载体基因组。在一些情况下,所递送的载体基因组的量为1×10^{10}个到1×10^{11}个载体基因组。在一些情况下,所递送的载体基因组的量为1×10^{10}个到3×10^{12}个载体基因组。在一些情况下,所递送的载体基因组的量为1×10^{9}个到3×10^{13}个载体基因组。在一些情况下,所递送的载体基因组的量为1×10^{8}个到3×10^{14}个载体基因组。在一些情况下,所递送的载体基因组的量为1×10^{6}个到1×10^{15}个载体基因组。

[0435] 在一些情况下,可以使用感染复数(MOI)测量待施用的药物组合物的量。在一些情况下,MOI可以指载体或病毒基因组与核酸可以递送至的细胞的比率或倍数。在一些情况下,MOI可以为1×10^{6}个。在一些情况下,MOI可以为1×10^{5}个到1×10^{7}个。在一些情况下,MOI可以为1×10^{4}个到1×10^{8}个。在一些情况下,本公开的重组病毒至少为约1×10^{1}个、1×10^{2}个、1×10^{3}个、1×10^{4}个、1×10^{5}个、1×10^{6}个、1×10^{7}个、1×10^{8}个、1×10^{9}个、1×10^{10}个、$1\times$

10^{11}个、1×10^{12}个、1×10^{13}个、1×10^{14}个、1×10^{15}个、1×10^{16}个、1×10^{17}和1×10^{18}个MOI。在一些情况下，本公开的重组病毒为1×10^{8}个到3×10^{14}个MOI。在一些情况下，本公开的重组病毒至多为约1×10^{1}个、1×10^{2}个、1×10^{3}个、1×10^{4}个、1×10^{5}个、1×10^{6}个、1×10^{7}个、1×10^{8}个、1×10^{9}个、1×10^{10}个、1×10^{11}个、1×10^{12}个、1×10^{13}个、1×10^{14}个、1×10^{15}个、1×10^{16}个、1×10^{17}和1×10^{18}个MOI。

[0436] 在一些方面，药物组合物的量包括约1×10^{6}个到约1×10^{15}个重组病毒颗粒、约1×10^{7}个到约1×10^{14}个重组病毒颗粒、约1×10^{8}个到约1×10^{13}个重组病毒颗粒、约1×10^{9}个到约3×10^{12}个重组病毒颗粒或约1×10^{10}个到约3×10^{12}个重组病毒颗粒。

[0437] 单独剂量通常不小于对受试者产生可测量效果所需的量，并且可以基于主题组合物或其副产物的吸收、分布、代谢和排泄(“ADME”)的药代动力学和药理学来确定，并且因此基于受试者内的组合物的分布来确定。这包含考虑施用途径和剂量，可以针对视网膜下应用(直接应用于主要需要局部作用的部位)、玻璃体内应用(针对全视网膜作用而应用于玻璃体)或肠胃外应用(通过全身途径应用，例如静脉内、肌肉内等)进行调整。根据临床前测定、安全性和递增以及剂量范围试验、个体临床医师-患者关系等，可以凭经验容易地确定有效剂量和/或剂量方案。

[0438] 在另一个方面，本公开提供了分离的核酸序列，其包括选自由以下组成的组的序列：

[0439] (i)序列10(SEQ ID NO:10)；以及

[0440] (ii)序列12(SEQ ID NO:12)。

[0441] 这些序列是上面讨论的突变的视蛋白内含子，其可以用于多种构建体中以改善所关注的基因(如视蛋白基因)的表达。

[0442] 序列10:OPN1LW/MW微小内含子2(SEQ ID NO:10)

[0443] Gtaagccagtcggggcccaggctcggcggaaaccactcattcaccctgcaagctcctccagccacctcatgatgatcggggcccagctgctcctgtaggcctgtctccctccccatctgcgcctcacatccatatactgaagggttctggggtggaaagaaagatgtcgtttttttccacctcagtccgtggagccctgaattctgtgtgcagacgtttgggtctaagcaggac

[0444] 序列12:OPN1LW/MW微小内含子3(SEQ ID NO:12)

[0445] gtaagggtgcgaggacgcaagatggagtgggcagggtcagactctgtgaccttaaggcaaatcacttcctttctctgggcccctctgagcgtgcaatgtctatcaatgtatgaatgtggctgcaacataggaaaggctctgtggtccccgttgatcacttcaaattgggtaatctcatgcaacatgaatttcacctcaatttaaaaaaacaaaccccaccgagttagcaccgtgcctgggccgggggtcctgggtcaccccaccctgcatcaggactggctgccggcccttctctccag

[0446] 在另一个方面，本公开提供了分离的核酸，其包括通式A-B的核酸序列，其中A编码包括氨基酸序列MSRKIEGFLLLLLFGYEATLGLSS(SEQ ID NO:21)、基本上由所述氨基酸序列组成或由所述氨基酸序列组成的多肽，并且其中B编码用于治疗视锥细胞病症的多肽。这些核酸编码针对视锥细胞病症的治疗性治疗特别有用的基因产物。

[0447] 在一个实施例中，A结构域由核酸序列ATGTCACGCAAGATAGAAGGCTTTTTGTTATTACTTCTCTTTGGCTATGAAGCCACATTGGGATTATCGTCT(SEQ ID NO:100)编码。

[0448] 用于治疗视锥细胞病症的多肽可以是本文所公开的多肽中的任何多肽，如

OPN1LW/MW、阿柏西普或经过修饰的阿柏西普或其功能片段、衍生物或变体。在一个实施例中,B包括以下序列或由以下序列组成:

[0449] AGTGATACCGGTAGACCTTTCGTAGAGATGTACAGTGAAATCCCCGAAATTATACACATGACTGAAGG
AAGGGAGCTCGTCATTCCCTGCCGGGTTACGTCACCTAACATCACTGTTACTTTAAAAAAGTTTCCACTTGACACT
TTGATCCCTGATGGAAAACGCATAATCTGGGACAGTAGAAAGGGCTTCATCATATCAAATGCAACGTACAAAGAAA
TAGGGCTTCTGACCTGTGAAGCAACAGTCAATGGGCATTTGTATAAGACAAACTATCTCACACATCGACAAACCAA
TACAATCATAGATGTGGTTCTGAGTCCGTCTCATGGAATTGAACTATCTGTTGGAGAAAAGCTTGTCTTAAATTGT
ACAGCAAGAACTGAACTAAATGTGGGGATTGACTTCAACTGGGAATACCCTTCTTCGAAGCATCAGCATAAGAAAC
TTGTAAACCGAGACCTAAAAACCCAGTCTGGGAGTGAGATGAAGAAATTTTTGAGCACCTTAACTATAGATGGTGT
AACCCGGAGTGACCAAGGATTGTACACCTGTGCAGCATCCAGTGGGCTGATGACCAAGAAGAACAGCACATTTGTC
AGGGTCCATGAAAAGGACAAAACTCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCT
TCCTCTTCCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGT
GAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCG
CGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCA
AGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCA
GCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGC
CTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGA
CCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCA
GCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTG
TCTCCGGGTAAA(TAG)(SEQ ID NO 101)。

[0450] 其中括号中的残基可以被修饰成不同的终止密码子,或者可以不存在(如当多肽被表达为融合蛋白时)。

[0451] 在另一个实施例中,核酸编码经过修饰的阿柏西普多肽或其功能片段、衍生物或变体,所述经过修饰的阿柏西普多肽包括以下氨基酸序列、基本上由以下氨基酸序列组成或由以下氨基酸序列组成:

[0452] MSRKIEGFLLLLLFGYEATLGLSS

[0453] DTGRPFVEMYSEIPEIIHMTEGRELVIPCRVTSPNITVTLKKFPLDTLIPDGKRIIWDSRKGFIISNA
TYKEIGLLTCEATVNGHLYKTNYLTHRQTNTIIDVVLSPSHGIELSVGEKLVLNCTARTELNVGIDFNWEYPSSKH
QHKKLVNRDLKTQSGSEMKKFLSTLTIDGVTRSDQGLYTCAASSGLMTKKNSTFVRVHEKDKTHTCPPCPAPELLG
GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQD
WLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPE
NNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK(SEQ ID NO:102)。

[0454] 在另一个方面,提供了一种经过修饰的阿柏西普多肽或其功能片段、衍生物或变体,所述经过修饰的阿柏西普多肽包括以下氨基酸序列、基本上由以下氨基酸序列组成或由以下氨基酸序列组成:

[0455] MSRKIEGFLLLLLFGYEATLGLSSDTGRPFVEMYSEIPEIIHMTEGRELVIPCRVTSPNITVTLKKFP
LDTLIPDGKRIIWDSRKGFIISNATYKEIGLLTCEATVNGHLYKTNYLTHRQTNTIIDVVLSPSHGIELSVGEKLV
LNCTARTELNVGIDFNWEYPSSKHQHKKLVNRDLKTQSGSEMKKFLSTLTIDGVTRSDQGLYTCAASSGLMTKKNS
TFVRVHEKDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK

TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK(SEQ ID NO:103)。

[0456] 实例：

[0457] 我们已经系统地特异性修饰了表达载体中的各种元件,以找到可以协同提供治疗性膜结合和分泌蛋白的大大增强的功能表达的组合。在此过程中操纵的元件包含经过改善的增强子和经过改善的近端启动子,所述经过改善的增强子和经过改善的近端启动子可以用于例如驱动长(L)和中等(M)波长视锥光感受器的特异性表达,并且包含改善转录起始信号和翻译信号(Kozak序列)、3'和5'非翻译区(包含原生视蛋白聚腺苷酸化信号)、两个不同的经过改善的内含子、转录起始片段和土拨鼠肝炎病毒转录后调节元件(WPRE)。

[0458] 这里我们给出两个具体的实例,一个是膜结合蛋白,另一个是人L视锥视蛋白中的示例蛋白。第二个实例针对分泌蛋白,并且所述实例是血管内皮生长因子-Trap(VEGF Trap)。在使用这些试剂进行各项研究时,功能蛋白具有不会明显干扰表达或功能的荧光标签的版本会是有用的。因此,除了治疗性版本外,我们还包含了这样的试剂:其中在视蛋白的情况下使用GFP标记治疗性蛋白质,而在VEGF-trap的情况下使用柠檬黄标记治疗性蛋白质。最后,有效的mRNA剪接是有效的基因表达的关键,表达盒中包含的内含子提供了明显的益处。还存在一种担忧,即包含在盒中的病毒序列可能被靶向基因沉默,并且不利于转基因的长期表达。因此,我们包了含两种不同版本的试剂,一种带有病毒内含子,而第二选项采用真核内含子。

[0459] 本文所公开的试剂首次允许从注射入眼玻璃体开始在灵长类动物黄斑的视锥光感受器中稳定表达治疗性蛋白质。这些试剂用于治疗红绿色盲,并且用于拯救由X染色体基因OPN1LW和OPN1MW的突变引起的严重视觉障碍。通过介导分泌蛋白的表达,所述试剂已应用于通过基因疗法治疗人的黄斑疾病,这有望消除对视网膜下注射的需要,并且可以避免相关的风险。

[0460] 用于制备表达盒以在视锥光感受器中表达治疗性基因的示例性方法。

[0461] 用于产生本文所描述的表达盒的常用部件清单包含：

[0462] A.在L/M视锥光感受器中有选择地、稳健地实现表达所需的调节元件。

[0463] a.将基因座控制区(LCR)截短以产生用于表达盒的短LCR(序列2;(SEQ ID NO:2))。

[0464] b.修饰OPN1LW基因的启动子,以产生两个经过优化的L启动子：

[0465] i.经过修饰的L启动子v1.0,当内含子紧随启动子(序列4)(SEQ ID NO:4)时使用。

[0466] ii.经过修饰的L启动子v2.0,当治疗性基因编码区紧随启动子(序列5)(SEQ ID NO:5)时使用。

[0467] B.一个或两个内含子。内含子可用于实现高水平的转录,但是其位置和序列也可能对转录产生负面影响。在我们的新表达盒中使用经过修饰的内含子。

[0468] a.SV40微小内含子(序列6;(SEQ ID NO:6))。

[0469] b.pFLARE内含子1和2(源自人β球蛋白基因)(序列7(SEQ ID NO:7)和8(SEQ ID NO:8))。

[0470] c.经过修饰的OPN1LW/OPN1MW微小内含子2和微小内含子3(序列10(SEQ ID NO:

10)和12(SEQ ID NO:12))。

[0471] C.对有效转录和翻译有用的调节元件(通用)

[0472] a.用于翻译起始的Kozak共有序列。

[0473] b.土拨鼠肝炎病毒转录后调节元件(WPRE),以及具体地说,已证明是有效的并能增强转录的完整的三方WPRE的子组分(序列15;(SEQ ID NO:15))。

[0474] c.从OPN1LW基因延伸的3'非翻译区,包含原生视蛋白基因多腺苷酸化信号和在其处将poly A尾添加至mRNA的位点(序列16;(SEQ ID NO:16))。

[0475] 特定于两个治疗性基因类别(L/M视蛋白或VEGF-TRAP)的部件清单包含:

[0476] a.OPN1LW/MW基因和cDNA,包含3'非翻译区以及3'UTR的3'侧上另外的侧接DNA。(作为OPN1LW的实例的序列13在没有5'非翻译区(UTR)的情况下使用)。

[0477] b.信号肽的遗传密码,其负责从光感受器中分泌视网膜劈裂症(RS1)(在序列14中突出显示)

[0478] c.阿柏西普的部分序列,包含VEGFR1结构域2、VEGFR2结构域3和IgG1FC(序列14)

[0479] 用于在视锥光感受器中表达治疗性基因的表达盒的示例性程序:

[0480] A.使用SV40微小内含子的表达盒

[0481] a.上游和下游调节模块如下进行组装:

[0482] 上游调节模块:将短LCR、经过优化的L启动子v1.0和SV40微小内含子以及部分Kozak序列(GCCGCCACC)按5'到3'的顺序连接。

[0483] 下游调节模块:将视蛋白的短WPRE和延伸的3'UTR按照从5'到3'的顺序连接。

[0484] b.通过将上游调节模块与5'侧的治疗性基因连接,并将下游调节模块与3'侧的治疗性基因连接,完整的表达盒得以连接。最终构建体5'到3'中的元件顺序为:LCR、启动子、内含子、Kozak、治疗性基因、WPRE和延伸的3'UTR。治疗性基因可以是例如OPN1LW(外显子1-6)、OPN1MW(外显子1-6)或经过修饰的阿柏西普或其功能片段、衍生物或变体。

[0485] B.用于治疗性视蛋白基因的表达盒,所述表达盒使用插入视蛋白基因中的内含子。

[0486] a.通过按以下指定顺序连接序列来产生上游调节模块:短LCR和经过优化的L启动子v2.0。

[0487] b.以与使用SV40微小内含子的表达盒相同的组装方法组装下游调节模块。

[0488] c.将内含子插入到治疗性视蛋白基因中:

[0489] i.在pFlare内含子的情况下,将内含子1插入外显子3的上游,并且将内含子2插入外显子3的下游。

[0490] ii.在OPN1LW/MW微小内含子的情况下,将微小内含子2插入外显子3的上游,并且将微小内含子3插入外显子3的下游。

[0491] d.通过将上游调节模块与5'侧上的含有内含子的治疗性视蛋白基因连接,并将下游调节模块与含有内含子的治疗性基因的3'端连接,完整的表达盒得以组装:

[0492] i.使用pFLARE内含子的最终表达盒的从5'到3'的元件呈以下顺序:LCR、启动子(通过连接启动子和视蛋白基因产生Kozak)、视蛋白的外显子1和2、pFLARE内含子1、视蛋白的外显子3、pFLARE内含子2、视蛋白的外显子4、5和6、WPRE、延伸的视蛋白3'UTR。

[0493] ii.使用OPN1LW/MW内含子的最终表达盒的从5'到3'的元件呈以下顺序:LCR、启动

子(通过连接启动子和视蛋白产生Kozak)、视蛋白的外显子1和2、OPN1LW/MW微小内含子2、视蛋白的外显子3、OPN1LW/MW微小内含子3、视蛋白的外显子4、5和6、WPRE、延伸的视蛋白3’UTR。

[0494] C.用于实验目的的荧光蛋白的基因插入：

[0495] a.视蛋白/GFP融合物。紧邻视蛋白基因翻译终止密码子的上游插入GFP编码区。注意：视蛋白/GFP融合物与视蛋白微小内含子并非优选的，因为视蛋白微小内含子可活化GFP中的隐蔽剪接位点。在一个选项中，当用编码S视蛋白的最后12个氨基酸的TCAACTGTGTCCTCGACCCAGGTAGGGCCTAAC(SEQ ID NO:72)替换OPN1LW/MW编码区的最后27个核苷酸(TCATCTGTGTCCTCGGTATCGCCTGCA(SEQ OD NO:71))时，可以增强融合蛋白在光感受器膜中的表达。可以将指定了OPN1LW/MW C端的最后10个氨基酸的序列TCATCTGTGTCCTCGGTATCGCCTGCATAG(SEQ ID NO:73)插入GFP的最后一个氨基酸编码密码子之间。插入物包含翻译终止密码子。

[0496] i.最终盒中的元件顺序为：上游调节模块、具有S视蛋白C端尾的经过修饰的OPN1LW/MW基因、具有L/M视蛋白C端尾的经过修饰的GFP、翻译终止密码子、下游调节模块。

[0497] ii.可以使用经过修饰的视蛋白C端的其它组合。例如，视蛋白可以保持其原生的C端，并且可以将视蛋白C端的12个氨基酸添加到GFP。关键要求是包含用于将视蛋白定位于GFP的C端处的膜的信号。

[0498] b.经过修饰的阿柏西普/荧光基因(柠檬黄)需要在IgG1 FC与柠檬黄编码区之间的连接子，以保持融合蛋白两个部分的功能。使用的连接子是：5’ggaggtggaggttctggtggaggaggttcc(SEQ ID NO:76)

[0499] 上文详细描述了构建体的组分，并且SEQ ID NO:91-95是用于促进视锥细胞中的基因表达的示例性全长表达载体。

[0500] 结果

[0501] 玻璃体内注射后，在来自基因治疗载体的视锥光感受器特异性启动子控制下的功能蛋白的表达以前从未被证实。本文显示的结果证明，在玻璃体内注射到啮齿动物和灵长类动物的眼睛后，本文所公开的新表达盒可产生功能性治疗性蛋白质的稳健持久的表达。实验还证明，与经过修饰的阿柏西普基因偶联的构建体可产生治疗性分子的稳健表达和分泌，并且本文所公开的pFLARE内含子携带非常强的剪接信号，所述剪接信号克服了由L和M视蛋白中的外显子3变异而引起的剪接缺陷，并且因此产生用于视蛋白基因疗法的理想内含子。

[0502] 在第一项研究中，将SEQ ID NO:91的构建体包装到HEK293细胞中的AAV2_7m8衣壳中，并通过碘克沙醇密度梯度离心法进行纯化。将含有约5×10^{11}个病毒基因组的三微升病毒溶液注入到小鼠眼睛的玻璃体中。使用纳升注射泵进行玻璃体内注射，同时在显微镜下观察注射。注射后一个月，用过剂量的戊巴比妥钠处死小鼠，并收获眼睛用于组织学，并且获得平整固定的小鼠视网膜的共聚焦显微镜荧光图像。图1示出了来自L视蛋白上的GFP标签的荧光。小鼠M视蛋白表达表现出了特征性的空间表达模式，其中表达集中在上侧视网膜中，而在下侧视网膜种很少或没有表达。(图1A)携带SEQ ID NO:91的AAV2_7m8的空间表达模式视蛋白-GFP遵循小鼠M视蛋白的原生表达模式，这表明SEQ ID NO:91不驱动小鼠S视锥中的表达，所述视锥在下侧小鼠视网膜中占主导地位。(图1B)示出了由白框勾勒出的A中的

区域的放大图,并表明了来自SEQ ID NO:91的视蛋白基因正确定位在被称为外部区段的细胞的专门区室中。

[0503] 在第二项研究中,将SEQ ID NO:92的构建体包装到HEK293细胞中的AAV2_7m8衣壳中,并通过碘克沙醇密度梯度超速离心法进行纯化。使用注射泵将含有约5×10^{11}个病毒基因组的三微升病毒溶液注入到玻璃体中,并用显微镜进行观察。用氯胺酮/甲苯噻嗪麻醉小鼠,并获得视网膜电图(ERG)结果。对于ERG,响应于一系列交替的634nm(红色)和535nm(绿色)发光二极管(LED)光脉冲,从放置在眼睛角膜上的电极记录ERG电位。数据在图2中示出。图的Y轴表示与记录的对红色LED响应的振幅相匹配所需的绿色LED的光强度。野生型小鼠不具有长波长敏感性(L)感光色素,并且其内源性M视锥介导了对634nm光的所有敏感性。因此,未经过处理的小鼠的基线值平均为5的相对较低强度值。如此处所示,平均而言,经过处理的动物需要约6倍的绿光(值接近30)才能匹配由红光引起的响应。此处显示的对红光敏感性的急剧增加表明,在通过玻璃体内注射携带SEQ ID NO:92的AAV2-7m8处理一个月后,人L-视蛋白的功能性表达水平很高。使用这项基于ERG的测定,先前使用表达盒pR2.1和pMNTC进行玻璃体内处理不会产生任何显著的功能性L-视蛋白表达。

[0504] 玻璃体内注射携带SEQ ID NO:91的AAV2_7m8后1年,我们接下来获得在体内拍摄的灵长类动物视网膜黄斑的retcam图像。使用0.33ml的结核菌素注射器和25号针头,将含有约3×10^{12}个病毒基因组的三十微升病毒溶液注射到用氯胺酮/甲苯噻嗪麻醉的灵长类动物的玻璃体中。将SEQ ID NO:91包装到HEK293细胞中的AAV2_7m8衣壳中,并通过碘克沙醇密度梯度超速离心法进行纯化。结果在图3中示出。如(图3A)所示,经过处理的眼底白光下的图像示出了黄斑区的解剖学标志。左上方可见视神经和视网膜血管的小新月形。右中下部的较暗区域是黄斑区,中央凹处在正中心。图3B示出了在蓝光下拍摄的,A中所示的相同眼底区域,以允许检测GFP荧光。中央凹中心的亮白点(图像右中下部)证明,注入到灵长类动物眼玻璃体中的AAV2_7m8中的SEQ ID NO:91产生持久、稳健的表达,所述表达遵循原生L视蛋白的空间表达,后者处于集中于中央凹的视锥光感受器中。

[0505] 图4示出了拯救灵长类动物视网膜疾病模型的视锥中的L-视蛋白表达。在玻璃体内注射携带SEQ ID NO:91的AAV2_7m8一年后,图3的灵长类动物眼睛中的GFP表达的共聚焦显微镜图像。穿过眼中央凹的横截面证明,单次玻璃体内注射导致GFP在中央黄斑区在近100%的L/M视锥中稳健表达,然而,所述单次玻璃体内注射不会转导S视锥或视杆光感受器,所述处理也不转导任何其它细胞类型,如视网膜中的神经节细胞。图像在中央凹上从右到左居中。白色标记是来自在视锥光感受器中特异性表达的视蛋白-GFP融合转基因的原生GFP荧光。灰色标记是用于使视网膜中所有细胞类型的细胞核视觉化的DAPI。结果表示对天然存在的灵长类动物视网膜疾病模型的成功的基因疗法处理。所述疾病模型是从大群体的基因筛查中鉴别出的松鼠猴中的红绿色盲。就像人类红绿色盲的一种形式一样,这些动物缺乏正常三色颜色视觉所需要的L-视锥(红色视锥)中正常表达的L-视蛋白。所述治疗在视锥光感受器中产生了人L-视蛋白转基因的表达,从而拯救此灵长类动物疾病模型中的缺陷。此处所测试的红色/绿色色盲“疾病模型”显示了人类蓝视锥单色视的重要特性。蓝视锥单色视的症状包含视觉敏锐度差(介于20/50与20/200之间)、畏光和全色盲。红绿色盲与蓝视锥单色视之间的共同特征包含:两者都是出生时就存在的固定性疾病;由OPN1MW/LW基因的突变引起;并且视锥在两种病症中均保持生存力,这使其成为本文所展示的基因替代疗

法的良好靶标。

[0506] 然后用携带与柠檬黄融合的VEGF-TRAP的合成基因的质粒转染HEK293T细胞，分泌VEGF-TRAP-柠檬黄融合蛋白。用X轴上指示的三个质粒之一转染HEK293T细胞。一种质粒含有在人突触蛋白1启动子控制下的柠檬黄基因(AAV2柠檬黄)，另一种质粒含有图7B所示的构建体(VEGF-Trap)，并且第三种质粒含有图7B所示的构建体，不同之处在于用RS1信号序列替换“sflt信号序列”以产生RS1-VEGF-Trap。VEGF-Trap构建体是SEQ ID NO:94，而VEGF-Trap柠檬黄融合物是SEQ ID NO:93。转染两天后，收集培养基并通过离心法除去细胞碎片。数据在图5中示出；在标记为“VEGF-Trap”的数据使用SEQ ID NO:93中的构建体，而标记为“RS1-VEGF-Trap”的数据使用SEQ ID NO:93中的构建体，不同之处在于标记为sflt信号序列的区域被来自另一个分泌分子(RS1)的信号序列替换。使用BioRad Glomax™发光计测量培养基的荧光强度。每个质粒转染六份重复培养物。测量六个重复的“无转染”对照，以给出背景荧光水平。不应分泌AAV2-Citrine，并且因此培养基不应示出高表达水平。如图5中所观察的，预计一些荧光来自裂解的细胞。已经用VEGF-Trap或RS1-VEGF-Trap转染的HEK293T细胞的荧光强度最高，如经过转染的细胞将VEGF-柠檬黄融合蛋白分泌到培养基中所预期的那样。为了使VEGF-Trap对眼部疾病具有治疗效果，必须从表达VEGF-Trap的细胞中将其分泌出来，并且此实验证明VEGF-Trap构建体(SEQ ID NO:93)是从细胞中分泌的。

[0507] 经过修饰的β-球蛋白内含子被称为pFLARE内含子，当插入OPN1LW cDNA的特定位置处时，所述内含子会提供强大的剪接信号，并且可以通过玻璃体内注射携带视蛋白基因的AAV来增加视蛋白基因的表达。OPN1LW外显子3的两个变体(被称为LIAVA和LVAVA)表现出完整的或接近完整的剪接缺陷，其中最终信使RNA(mRNA)不包含外显子3。如图8所展示的，我们将pFLARE内含子插入了外显子3的两侧，并通过转染携带图8的构建体区段的质粒进行了剪接测定，所述区段从OPN1LW x1/x2穿过短的WPRE延伸到HEK293细胞中。转染后24小时到48小时，从细胞中分离mRNA，并通过凝胶电泳和对凝胶上观察到的条带的直接测序进行检查。数据在图6中示出。凝胶上的每对通道均示出重复转染的结果。通道1a和1b示出了来自对照构建体的mRNA，所述对照构建体携带外显子3的未表现出剪接缺陷的版本。pFLARE内含子对此构建体中的剪接没有影响，因为观察到的唯一条带对应于包含外显子3的全长mRNA。通道2a和2b显示，pFLARE内含子抑制了OPN1LW外显子3的LIAVA版本的剪接缺陷，所述版本通常表现出完整的剪接缺陷，所述完整的剪接缺陷仅产生缺少外显子3的mRNA。同样，通道3a和3b显示，pFLARE内含子完全抑制了通常针对外显子3的LVAVA版本观察到的剪接缺陷。通道4是示出了未经过剪接的构建体的对照，并且mw是分子量标志物。*表示对应于全长mRNA的条带，并且**表示缺少外显子3的mRNA的条带，所述条带比全长mRNA短169bp。这些观察结果与存在原生内含子时，针对LIAVA和LVAVA OPN1LW外显子变体的剪接观察到的结果相反。这些数据表明，pFLARE内含子携带非常强的剪接信号，并且因此成为用于视蛋白基因疗法的理想内含子。

[0508] 从接受玻璃体内注射包装在AAV2_7m8衣壳中的SEQ ID NO:93(图7B)的小鼠的视锥光感受器中分泌VEGF-Trap-柠檬黄。使用HEK293细胞产生病毒，并通过碘克沙醇梯度超速离心法将所述病毒纯化。经过麻醉的小鼠接受玻璃体内注射约8×10^{11}个病毒基因组，体积为3微升。大约7周后，通过过剂量的戊巴比妥处死小鼠，并对眼睛进行加工以用于组织学。数据在图9中示出。整个固定的视网膜(图9A)示出了来自SEQ ID NO:93的柠檬黄荧光的

从上至下的图案,其遵循M-视蛋白表达的正常梯度。柠檬黄荧光可见于通过视锥内部区段的光学部分(图9B)。11周后,处死小鼠,并对眼睛进行加工以用于组织学。注射后11周,通过视网膜的冰冻切片(c,d)示出了VEGF-Trap-柠檬黄表达。在上侧视网膜看见最高浓度。视锥鞘用花生凝集素(d)染色,并且细胞核用DAPI(a,b,d)染色。

序列表

<110> 华盛顿大学

M•尼茨

J•尼茨

<120> 用于增强治疗性基因在光感受器中的功能性表达的组合物和方法

<130> 17-1292-PCT

<150> 62/594811

<151> 2017-12-05

<160> 115

<170> PatentIn版本3.5

<210> 1

<211> 166

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> 尚未归类的特征

<222> (166)..(166)

<223> 可选地缺失

<400> 1

gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60

gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120

caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccn 166

<210> 2

<211> 1230

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 2

ggaggctgag gggtggggaa agggcatggg tgtttcatga ggacagagct tccgtttcat 60

gcaatgaaaa gagtttggag acggatggtg gtgactggac tatacactta cacacggtag 120

cgatggtaca ctttgtatta tgtatatttt accacgatct ttttaaagtg tcaaaggcaa 180

atggccaaat ggttccttgt cctatagctg tagcagccat cggctgttag tgacaaagcc 240

cctgagtcaa gatgacagca gcccccataa ctcctaatcg gctctcccgc gtggagtcat 300

ttaggagtag tcgcattaga gacaagtcca acatctaatc ttccaccctg gccagggccc 360

cagctggcag cgagggtggg agactccggg cagagcagag ggcgctgaca ttggggcccg 420

gcctggcttg ggtccctctg gcctttcccc aggggccctc tttccttggg gctttcttgg 480
gccgccactg ctcccgctcc tctcccccca tcccaccccc tcacccccctc gttcttcata 540
tccttctcta gtgctccctc cactttcatc cacccttctg caagagtgtg ggaccacaaa 600
tgagttttca cctggcctgg ggacacacgt gcccccacag gtgctgagtg actttctagg 660
acagtaatct gctttaggct aaaatgggac ttgatcttct gttagcccta atcatcaatt 720
agcagagccg gtgaaggtgc agaacctacc gcctttccag gcctcctccc acctctgcca 780
cctccactct ccttcctggg atgtgggggc tggcacacgt gtggcccagg gcattggtgg 840
gattgcactg agctgggtca ttagcgtaat cctggacaag ggcagacagg gcgagcggag 900
ggccagctcc ggggctcagg caaggctggg ggcttccccc agacacccca ctcctcctct 960
gctggacccc cacttcatag ggcacttcgt gttctcaaag ggcttccaaa tagcatggtg 1020
gccttggatg cccagggaag cctcagagtt gcttatctcc ctctagacag aaggggaatc 1080
tcggtcaaga gggagaggtc gccctgttca aggccaccca gccagctcat ggcggtaatg 1140
ggacaaggct ggccagccat cccaccctca gaagggaccc ggtggggcag gtgatctcag 1200
aggaggctca cttctgggtc tcacattctt 1230

<210> 3
<211> 180
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 3
gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60
gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120
caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccatagc cggtaccatg 180

<210> 4
<211> 177
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 4
gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60
gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120
caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccatagc cggtacc 177

<210> 5
<211> 174
<212> DNA
<213> 人工序列
<220>

<223> 合成的

<400> 5

gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60

gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120

caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccgccgc cacc 174

<210> 6

<211> 193

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 6

ctagaggatc cggtactcga ggaactgaaa aaccagaaag ttaactggta agtttagtct 60

ttttgtcttt tatttcaggt cccggatccg gtggtggtgc aaatcaaaga actgctcctc 120

agtggatgtt gcctttactt ctaggcctgt acggaagtgt tacttctgct ctaaaagctg 180

cggaattgta ccc 193

<210> 7

<211> 129

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 7

gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg catgtggaga 60

cagagaagac tcacgcgttt ctgaattcac tgactctctc tgcctattgg tctattttct 120

cacccttag 129

<210> 8

<211> 158

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 8

gttggtatca aggttacaag acaggtttaa ggagaccaat ggatccatag tcgaccacca 60

tggtggctta gatccgggca tgtggagaca gagaagactg ttgagtttgt gataagcact 120

gactctctct gcctattggt ctattttccc tccctcag 158

<210> 9

<211> 409

<212> DNA

<213> 人工序列
<220>
<223> 合成的
<400> 9
atggcccagc agtggagcct ccaaaggctc gcaggccgcc atccgcagga cagctatgag 60
gacagcaccc agtccagcat cttcacctac accaacagca actccaccag aggccccttc 120
gaaggcccga attaccacat cgctcccaga tgggtgtacc acctcaccag tgtctggatg 180
atctttgtgg tcactgcatc cgtcttcaca aatgggcttg tgctggcggc caccatgaag 240
ttcaagaagc tgcgccaccc gctgaactgg atcctggtga acctggcggt cgctgaccta 300
gcagagaccg tcatcgccag cactatcagc attgtgaacc aggtctctgg ctacttcgtg 360
ctgggccacc ctatgtgtgt cctggagggc tacaccgtct ccctgtgtg 409
<210> 10
<211> 233
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 10
taagccagtc ggggcccagg ctcggcggaa accactcatt caccctgcaa gctcctccag 60
ccacctcatg atgatcgggg cccagctgct cctgtaggcc tgtctccctc cccatctgcg 120
cctcacatcc atatactgaa gggttctggg gtggaaagaa agatgtcgtt ttttccacct 180
cagtccgtgg agccctgaat tctgtgtgca gacgtttggg gtctaagcag gac 233
<210> 11
<211> 169
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 11
ggatcacagg tctctggtct ctggccatca tttcctggga gaggtggctg gtggtgtgca 60
agcccttttgg caatgtgaga tttgatgcca agctggccat cgtgggcatt gccttctcct 120
ggatctggtc tgctgtgtgg acagccccgc ccatctttgg ttggagcag 169
<210> 12
<211> 300
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 12

gtaagggtgc gaggacgcaa gatggagtgg gcagggtcag actctgtgac cttaaggcaa 60
atcacttcct ttctctgggc ccctctgagc gtgcaatgtc tatcaatgta tgaatgtggc 120
tgcaacatag gaaaggctct gtggtccccg ttgatcactt caaattgggt aatctcatgc 180
aacatgaatt tcacctcaat ttaaaaaaac aaaccccacc cgagttagca ccgtgcctgg 240
gccgggggtc ctgggtcacc ccaccctgca tcaggactgg ctgccggccc ttctctccag 300

<210> 13
<211> 1095
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (1093)..(1095)
<223> 可选地缺失
<400> 13

atggcccagc agtggagcct ccaaaggctc gcaggccgcc atccgcagga cagctatgag 60
gacagcaccc agtccagcat cttcacctac accaacagca actccaccag aggccccttc 120
gaaggcccga attaccacat cgctcccaga tgggtgtacc acctcaccag tgtctggatg 180
atctttgtgg tcactgcatc cgtcttcaca aatgggcttg tgctggcggc caccatgaag 240
ttcaagaagc tgcgccaccc gctgaactgg atcctggtga acctggcggt cgctgaccta 300
gcagagaccg tcatcgccag cactatcagc attgtgaacc aggtctctgg ctacttcgtg 360
ctgggccacc ctatgtgtgt cctggagggc tacaccgtct ccctgtgtgg gatcacaggt 420
ctctggtctc tggccatcat ttcctgggag aggtggctgg tggtgtgcaa gccctttggc 480
aatgtgagat ttgatgccaa gctggccatc gtgggcattg ccttctcctg gatctggtct 540
gctgtgtgga cagccccgcc catctttggt tggagcaggt actggcccca cggcctgaag 600
acttcatgcg gcccagacgt gttcagcggc agctcgtacc ccggggtgca gtcttacatg 660
attgtcctca tggtcacctg ctgcatcatc ccactcgcta tcatcatgct ctgctacctc 720
caagtgtggc tggccatccg agcggtggca aagcagcaga aagagtctga atccacccag 780
aaggcagaga aggaagtgac gcgcatggtg gtggtgatga tctttgcgta ctgcgtctgc 840
tggggaccct acaccttctt cgcatgcttt gctgctgcca accctggtta cgccttccac 900
cctttgatgg ctgccctgcc ggcctacttt gccaaaagtg ccactatcta caaccccgtt 960
atctatgtct ttatgaaccg gcagtttcga aactgcatct tgcagctttt cgggaagaag 1020
gttgacgatg gctctgaact ctccagcgcc tccaaaacgg aggtctcatc tgtgtcctcg 1080
gtatcgcctg catga 1095

<210> 14
<211> 1371
<212> DNA
<213> 人工序列

<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (1369)..(1371)
<223> 可选地缺失
<400> 14
atgtcacgca agatagaagg ctttttgtta ttacttctct ttggctatga agccacattg 60
ggattatcgt ctagtgatac cggtagacct ttcgtagaga tgtacagtga aatccccgaa 120
attatacaca tgactgaagg aagggagctc gtcattccct gccgggttac gtcacctaac 180
atcactgtta ctttaaaaaa gtttccactt gacactttga tccctgatgg aaaacgcata 240
atctgggaca gtagaaaggg cttcatcata tcaaatgcaa cgtacaaaga aatagggctt 300
ctgacctgtg aagcaacagt caatgggcat ttgtataaga caaactatct cacacatcga 360
caaaccaata caatcataga tgtggttctg agtccgtctc atggaattga actatctgtt 420
ggagaaaagc ttgtcttaaa ttgtacagca agaactgaac taaatgtggg gattgacttc 480
aactgggaat acccttcttc gaagcatcag cataagaaac ttgtaaaccg agacctaaaa 540
acccagtctg ggagtgagat gaagaaattt ttgagcacct taactataga tggtgtaacc 600
cggagtgacc aaggattgta cacctgtgca gcatccagtg ggctgatgac caagaagaac 660
agcacatttg tcagggtcca tgaaaaggac aaaactcaca catgcccacc gtgcccagca 720
cctgaactcc tgggggacc gtcagtcttc ctcttccccc caaaacccaa ggacaccctc 780
atgatctccc ggaccctga ggtcacatgc gtggtggtgg acgtgagcca cgaagaccct 840
gaggtcaagt tcaactggta cgtggacggc gtggaggtgc ataatgccaa gacaaagccg 900
cgggaggagc agtacaacag cacgtaccgt gtggtcagcg tcctcaccgt cctgcaccag 960
gactggctga atggcaagga gtacaagtgc aaggtctcca acaaagccct cccagccccc 1020
atcgagaaaa ccatctccaa agccaaaggg cagccccgag aaccacaggt gtacaccctg 1080
cccccatccc gggatgagct gaccaagaac caggtcagcc tgacctgcct ggtcaaaggc 1140
ttctatccca gcgacatcgc cgtggagtgg gagagcaatg ggcagccgga gaacaactac 1200
aagaccacgc ctcccgtgct ggactccgac ggctccttct tcctctacag caagctcacc 1260
gtggacaaga gcaggtggca gcaggggaac gtcttctcat gctccgtgat gcatgaggct 1320
ctgcacaacc actacacgca gaagagcctc tccctgtctc cgggtaaata g 1371
<210> 15
<211> 246
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 15
aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct 60
ccttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat tgcttcccgt 120

atggctttca ttttctcctc cttgtataaa tcctggttag ttcttgccac ggcggaactc 180
atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac tgacaattcc 240
gtggtg 246
<210> 16
<211> 288
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (107)..(288)
<223> 可选地缺失
<400> 16
ggtctgcctc ctacccatcc cgcccaccgg ggctttggcc acctctcctt tccccctcct 60
tctccatccc tgtaaaataa atgtaattta tctttgccaa aaccaacaaa gtcacagagg 120
ctttcactgc agtgtgggac cacctgagcc tctgcgtgtg caggcactgg gtctcgagag 180
ggtgcaaggg ggataaagag gagagagcgc ttcatagact ttaagttttc ccgagcctca 240
tgtctaccga tggcgtgaaa ggatcctggc aaaacagaag tgtgaggc 288
<210> 17
<211> 312
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 17
agtgataccg gtagaccttt cgtagagatg tacagtgaaa tccccgaaat tatacacatg 60
actgaaggaa gggagctcgt cattccctgc cgggttacgt cacctaacat cactgttact 120
ttaaaaaagt ttccacttga cactttgatc cctgatggaa aacgcataat ctgggacagt 180
agaaagggct tcatcatatc aaatgcaacg tacaaagaaa tagggcttct gacctgtgaa 240
gcaacagtca atgggcattt gtataagaca aactatctca cacatcgaca aaccaataca 300
atcatagatg tg 312
<210> 18
<211> 303
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 18

gttctgagtc cgtctcatgg aattgaacta tctgttggag aaaagcttgt cttaaattgt 60
acagcaagaa ctgaactaaa tgtggggatt gacttcaact gggaataccc ttcttcgaag 120
catcagcata agaaacttgt aaaccgagac ctaaaaaccc agtctgggag tgagatgaag 180
aaatttttga gcaccttaac tatagatggt gtaacccgga gtgaccaagg attgtacacc 240
tgtgcagcat ccagtgggct gatgaccaag aagaacagca catttgtcag ggtccatgaa 300
aag 303

<210> 19
<211> 681
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 19

gacaaaactc acacatgccc accgtgccca gcacctgaac tcctgggggg accgtcagtc 60
ttcctcttcc ccccaaaacc caaggacacc ctcatgatct cccggacccc tgaggtcaca 120
tgcgtggtgg tggacgtgag ccacgaagac cctgaggtca agttcaactg gtacgtggac 180
ggcgtggagg tgcataatgc caagacaaag ccgcgggagg agcagtacaa cagcacgtac 240
cgtgtggtca gcgtcctcac cgtcctgcac caggactggc tgaatggcaa ggagtacaag 300
tgcaaggtct ccaacaaagc cctcccagcc cccatcgaga aaaccatctc caaagccaaa 360
gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggatga gctgaccaag 420
aaccaggtca gcctgacctg cctggtcaaa ggcttctatc ccagcgacat cgccgtggag 480
tgggagagca atgggcagcc ggagaacaac tacaagacca cgcctcccgt gctggactcc 540
gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg 600
aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc 660
ctctccctgt ctccgggtaa a 681

<210> 20
<211> 456
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 20

Met Ser Arg Lys Ile Glu Gly Phe Leu Leu Leu Leu Leu Phe Gly Tyr
1 5 10 15
Glu Ala Thr Leu Gly Leu Ser Ser Ser Asp Thr Gly Arg Pro Phe Val
20 25 30
Glu Met Tyr Ser Glu Ile Pro Glu Ile Ile His Met Thr Glu Gly Arg
35 40 45
Glu Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile Thr Val Thr

50 55 60

Leu Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly Lys Arg Ile
65 70 75 80

Ile Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys
85 90 95

Glu Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr
100 105 110

Lys Thr Asn Tyr Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val
115 120 125

Val Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu Lys Leu
130 135 140

Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile Asp Phe
145 150 155 160

Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu Val Asn
165 170 175

Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe Leu Ser
180 185 190

Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr
195 200 205

Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val
210 215 220

Arg Val His Glu Lys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
225 230 235 240

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
245 250 255

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
260 265 270

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
275 280 285

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
290 295 300

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
305 310 315 320

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
325 330 335

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
340 345 350

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr
355 360 365

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
370 375 380
Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
385 390 395 400
Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
405 410 415
Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
420 425 430
Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
435 440 445
Ser Leu Ser Leu Ser Pro Gly Lys
450 455

<210> 21
<211> 24
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 21
Met Ser Arg Lys Ile Glu Gly Phe Leu Leu Leu Leu Leu Phe Gly Tyr
1 5 10 15
Glu Ala Thr Leu Gly Leu Ser Ser
20

<210> 22
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 22
Ser Asp Thr Gly Arg
1 5

<210> 23
<211> 99
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 23

Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu Ile Ile His Met Thr
1 5 10 15
Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile
20 25 30
Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly
35 40 45
Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala
50 55 60
Thr Tyr Lys Glu Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly
65 70 75 80
His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg Gln Thr Asn Thr Ile
85 90 95
Ile Asp Val

<210> 24
<211> 101
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 24

Val Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu Lys Leu
1 5 10 15
Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile Asp Phe
20 25 30
Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu Val Asn
35 40 45
Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe Leu Ser
50 55 60
Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr
65 70 75 80
Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val
85 90 95
Arg Val His Glu Lys
100

<210> 25
<211> 227
<212> PRT
<213> 人工序列
<220>

<223> 合成的
<400> 25

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
1 5 10 15
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
20 25 30
Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
35 40 45
Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
50 55 60
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
65 70 75 80
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
85 90 95
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
100 105 110
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
115 120 125
Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
130 135 140
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
145 150 155 160
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
165 170 175
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
180 185 190
Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
195 200 205
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
210 215 220
Pro Gly Lys
225

<210> 26
<211> 224
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 26

Met Ser Arg Lys Ile Glu Gly Phe Leu Leu Leu Leu Leu Phe Gly Tyr
1 5 10 15
Glu Ala Thr Leu Gly Leu Ser Ser Thr Glu Asp Glu Gly Glu Asp Pro
20 25 30
Trp Tyr Gln Lys Ala Cys Lys Cys Asp Cys Gln Gly Gly Pro Asn Ala
35 40 45
Leu Trp Ser Ala Gly Ala Thr Ser Leu Asp Cys Ile Pro Glu Cys Pro
50 55 60
Tyr His Lys Pro Leu Gly Phe Glu Ser Gly Glu Val Thr Pro Asp Gln
65 70 75 80
Ile Thr Cys Ser Asn Pro Glu Gln Tyr Val Gly Trp Tyr Ser Ser Trp
85 90 95
Thr Ala Asn Lys Ala Arg Leu Asn Ser Gln Gly Phe Gly Cys Ala Trp
100 105 110
Leu Ser Lys Phe Gln Asp Ser Ser Gln Trp Leu Gln Ile Asp Leu Lys
115 120 125
Glu Ile Lys Val Ile Ser Gly Ile Leu Thr Gln Gly Arg Cys Asp Ile
130 135 140
Asp Glu Trp Met Thr Lys Tyr Ser Val Gln Tyr Arg Thr Asp Glu Arg
145 150 155 160
Leu Asn Trp Ile Tyr Tyr Lys Asp Gln Thr Gly Asn Asn Arg Val Phe
165 170 175
Tyr Gly Asn Ser Asp Arg Thr Ser Thr Val Gln Asn Leu Leu Arg Pro
180 185 190
Pro Ile Ile Ser Arg Phe Ile Arg Leu Ile Pro Leu Gly Trp His Val
195 200 205
Arg Ile Ala Ile Arg Met Glu Leu Leu Glu Cys Val Ser Lys Cys Ala
210 215 220

<210> 27
<211> 214
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 27

Asp Ile Gln Leu Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15
Asp Arg Val Thr Ile Thr Cys Ser Ala Ser Gln Asp Ile Ser Asn Tyr
20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Val Leu Ile
35 40 45
Tyr Phe Thr Ser Ser Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80
Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Thr Val Pro Trp
85 90 95
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140
Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160
Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175
Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205
Phe Asn Arg Gly Glu Cys
210

<210> 28
<211> 231
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 28

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Asp Phe Thr His Tyr
20 25 30
Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45
Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Ala Asp Phe
50 55 60

Lys Arg Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr
65 70 75 80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Lys Tyr Pro Tyr Tyr Tyr Gly Thr Ser His Trp Tyr Phe Asp Val
100 105 110
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly
115 120 125
Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly
130 135 140
Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val
145 150 155 160
Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe
165 170 175
Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val
180 185 190
Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val
195 200 205
Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
210 215 220
Ser Cys Asp Lys Thr His Leu
225 230

<210> 29
<211> 214
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 29

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15
Asp Arg Val Thr Ile Thr Cys Ser Ala Ser Gln Asp Ile Ser Asn Tyr
20 25 30
Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Val Leu Ile
35 40 45
Tyr Phe Thr Ser Ser Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly
50 55 60
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Thr Val Pro Trp
85 90 95
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140
Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160
Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175
Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205
Phe Asn Arg Gly Glu Cys
210

<210> 30
<211> 453
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 30

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Thr Phe Thr Asn Tyr
20 25 30
Gly Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45
Gly Trp Ile Asn Thr Tyr Thr Gly Glu Pro Thr Tyr Ala Ala Asp Phe
50 55 60
Lys Arg Arg Phe Thr Phe Ser Leu Asp Thr Ser Lys Ser Thr Ala Tyr
65 70 75 80
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Lys Tyr Pro His Tyr Tyr Gly Ser Ser His Trp Tyr Phe Asp Val
100 105 110

Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly
115 120 125
Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly
130 135 140
Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val
145 150 155 160
Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe
165 170 175
Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val
180 185 190
Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val
195 200 205
Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
210 215 220
Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu
225 230 235 240
Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
245 250 255
Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
260 265 270
Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val
275 280 285
Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser
290 295 300
Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
305 310 315 320
Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala
325 330 335
Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
340 345 350
Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln
355 360 365
Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
370 375 380
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
385 390 395 400
Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu
405 410 415
Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser

420 425 430

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser

435 440 445

Leu Ser Pro Gly Lys

450

<210> 31

<211> 418

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 31

Met Gln Ala Leu Val Leu Leu Leu Cys Ile Gly Ala Leu Leu Gly His

1 5 10 15

Ser Ser Cys Gln Asn Pro Ala Ser Pro Pro Glu Glu Gly Ser Pro Asp

20 25 30

Pro Asp Ser Thr Gly Ala Leu Val Glu Glu Glu Asp Pro Phe Phe Lys

35 40 45

Val Pro Val Asn Lys Leu Ala Ala Ala Val Ser Asn Phe Gly Tyr Asp

50 55 60

Leu Tyr Arg Val Arg Ser Ser Thr Ser Pro Thr Thr Asn Val Leu Leu

65 70 75 80

Ser Pro Leu Ser Val Ala Thr Ala Leu Ser Ala Leu Ser Leu Gly Ala

85 90 95

Glu Gln Arg Thr Glu Ser Ile Ile His Arg Ala Leu Tyr Tyr Asp Leu

100 105 110

Ile Ser Ser Pro Asp Ile His Gly Thr Tyr Lys Glu Leu Leu Asp Thr

115 120 125

Val Thr Ala Pro Gln Lys Asn Leu Lys Ser Ala Ser Arg Ile Val Phe

130 135 140

Glu Lys Lys Leu Arg Ile Lys Ser Ser Phe Val Ala Pro Leu Glu Lys

145 150 155 160

Ser Tyr Gly Thr Arg Pro Arg Val Leu Thr Gly Asn Pro Arg Leu Asp

165 170 175

Leu Gln Glu Ile Asn Asn Trp Val Gln Ala Gln Met Lys Gly Lys Leu

180 185 190

Ala Arg Ser Thr Lys Glu Ile Pro Asp Glu Ile Ser Ile Leu Leu Leu

195 200 205

Gly Val Ala His Phe Lys Gly Gln Trp Val Thr Lys Phe Asp Ser Arg

210 215 220
Lys Thr Ser Leu Glu Asp Phe Tyr Leu Asp Glu Glu Arg Thr Val Arg
225 230 235 240
Val Pro Met Met Ser Asp Pro Lys Ala Val Leu Arg Tyr Gly Leu Asp
245 250 255
Ser Asp Leu Ser Cys Lys Ile Ala Gln Leu Pro Leu Thr Gly Ser Met
260 265 270
Ser Ile Ile Phe Phe Leu Pro Leu Lys Val Thr Gln Asn Leu Thr Leu
275 280 285
Ile Glu Glu Ser Leu Thr Ser Glu Phe Ile His Asp Ile Asp Arg Glu
290 295 300
Leu Lys Thr Val Gln Ala Val Leu Thr Val Pro Lys Leu Lys Leu Ser
305 310 315 320
Tyr Glu Gly Glu Val Thr Lys Ser Leu Gln Glu Met Lys Leu Gln Ser
325 330 335
Leu Phe Asp Ser Pro Asp Phe Ser Lys Ile Thr Gly Lys Pro Ile Lys
340 345 350
Leu Thr Gln Val Glu His Arg Ala Gly Phe Glu Trp Asn Glu Asp Gly
355 360 365
Ala Gly Thr Thr Pro Ser Pro Gly Leu Gln Pro Ala His Leu Thr Phe
370 375 380
Pro Leu Asp Tyr His Leu Asn Gln Pro Phe Ile Phe Val Leu Arg Asp
385 390 395 400
Thr Asp Thr Gly Ala Leu Leu Phe Ile Gly Lys Ile Leu Asp Pro Arg
405 410 415
Gly Pro

<210> 32
<211> 687
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 32
Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser
1 5 10 15
Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly Ser Lys Leu Lys Asp Pro
20 25 30
Glu Leu Ser Leu Lys Gly Thr Gln His Ile Met Gln Ala Gly Gln Thr
35 40 45

Leu His Leu Gln Cys Arg Gly Glu Ala Ala His Lys Trp Ser Leu Pro
50 55 60
Glu Met Val Ser Lys Glu Ser Glu Arg Leu Ser Ile Thr Lys Ser Ala
65 70 75 80
Cys Gly Arg Asn Gly Lys Gln Phe Cys Ser Thr Leu Thr Leu Asn Thr
85 90 95
Ala Gln Ala Asn His Thr Gly Phe Tyr Ser Cys Lys Tyr Leu Ala Val
100 105 110
Pro Thr Ser Lys Lys Lys Glu Thr Glu Ser Ala Ile Tyr Ile Phe Ile
115 120 125
Ser Asp Thr Gly Arg Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu
130 135 140
Ile Ile His Met Thr Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val
145 150 155 160
Thr Ser Pro Asn Ile Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr
165 170 175
Leu Ile Pro Asp Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe
180 185 190
Ile Ile Ser Asn Ala Thr Tyr Lys Glu Ile Gly Leu Leu Thr Cys Glu
195 200 205
Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg
210 215 220
Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val
225 230 235 240
Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr
245 250 255
Pro Leu Asn Thr Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Lys
260 265 270
Asn Lys Arg Ala Ser Val Arg Arg Arg Ile Asp Gln Ser Asn Ser His
275 280 285
Ala Asn Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys Met Gln Asn Lys
290 295 300
Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser Gly Pro Ser Phe Lys
305 310 315 320
Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Phe Ile Thr Val
325 330 335
Lys His Arg Lys Gln Gln Val Leu Glu Thr Val Ala Gly Lys Arg Ser
340 345 350
Tyr Arg Leu Ser Met Lys Val Lys Ala Phe Pro Ser Pro Glu Val Val

355 360 365

Trp Leu Lys Asp Gly Leu Pro Ala Thr Glu Lys Ser Ala Arg Tyr Leu

370 375 380

Thr Arg Gly Tyr Ser Leu Ile Ile Lys Asp Val Thr Glu Glu Asp Ala

385 390 395 400

Gly Asn Tyr Thr Ile Leu Leu Ser Ile Lys Gln Ser Asn Val Phe Lys

405 410 415

Asn Leu Thr Ala Thr Leu Ile Val Asn Val Lys Pro Gln Ile Tyr Glu

420 425 430

Lys Ala Val Ser Ser Phe Pro Asp Pro Ala Leu Tyr Pro Leu Gly Ser

435 440 445

Arg Gln Ile Leu Thr Cys Thr Ala Tyr Gly Ile Pro Gln Pro Thr Ile

450 455 460

Lys Trp Phe Trp His Pro Cys Asn His Asn His Ser Glu Ala Arg Cys

465 470 475 480

Asp Phe Cys Ser Asn Asn Glu Glu Ser Phe Ile Leu Asp Ala Asp Ser

485 490 495

Asn Met Gly Asn Arg Ile Glu Ser Ile Thr Gln Arg Met Ala Ile Ile

500 505 510

Glu Gly Lys Asn Lys Met Ala Ser Thr Leu Val Val Ala Asp Ser Arg

515 520 525

Ile Ser Gly Ile Tyr Ile Cys Ile Ala Ser Asn Lys Val Gly Thr Val

530 535 540

Gly Arg Asn Ile Ser Phe Tyr Ile Thr Asp Val Pro Asn Gly Phe His

545 550 555 560

Val Asn Leu Glu Lys Met Pro Thr Glu Gly Glu Asp Leu Lys Leu Ser

565 570 575

Cys Thr Val Asn Lys Phe Leu Tyr Arg Asp Val Thr Trp Ile Leu Leu

580 585 590

Arg Thr Val Asn Asn Arg Thr Met His Tyr Ser Ile Ser Lys Gln Lys

595 600 605

Met Ala Ile Thr Lys Glu His Ser Ile Thr Leu Asn Leu Thr Ile Met

610 615 620

Asn Val Ser Leu Gln Asp Ser Gly Thr Tyr Ala Cys Arg Ala Arg Asn

625 630 635 640

Val Tyr Thr Gly Glu Glu Ile Leu Gln Lys Lys Glu Ile Thr Ile Arg

645 650 655

Gly Glu His Cys Asn Lys Lys Ala Val Phe Ser Arg Ile Ser Lys Phe

660 665 670

Lys Ser Thr Arg Asn Asp Cys Thr Thr Gln Ser Asn Val Lys His
675 680 685

<210> 33
<211> 128
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 33

Met Gly Ile Gln Gly Gly Ser Val Leu Phe Gly Leu Leu Leu Val Leu
1 5 10 15
Ala Val Phe Cys His Ser Gly His Ser Leu Gln Cys Tyr Asn Cys Pro
20 25 30
Asn Pro Thr Ala Asp Cys Lys Thr Ala Val Asn Cys Ser Ser Asp Phe
35 40 45
Asp Ala Cys Leu Ile Thr Lys Ala Gly Leu Gln Val Tyr Asn Lys Cys
50 55 60
Trp Lys Phe Glu His Cys Asn Phe Asn Asp Val Thr Thr Arg Leu Arg
65 70 75 80
Glu Asn Glu Leu Thr Tyr Tyr Cys Cys Lys Lys Asp Leu Cys Asn Phe
85 90 95
Asn Glu Gln Leu Glu Asn Gly Gly Thr Ser Leu Ser Glu Lys Thr Val
100 105 110
Leu Leu Leu Val Thr Pro Phe Leu Ala Ala Ala Trp Ser Leu His Pro
115 120 125

<210> 34
<211> 698
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 34

Met Glu Thr Arg Gly Leu Ala Asp Ser Gly Gln Gly Ser Phe Thr Gly
1 5 10 15
Gln Gly Ile Ala Arg Phe Gly Arg Ile Gln Lys Lys Ser Gln Pro Glu
20 25 30
Lys Val Val Arg Ala Ala Ser Arg Gly Arg Pro Leu Ile Gly Trp Thr
35 40 45
Gln Trp Cys Ala Glu Asp Gly Gly Asp Glu Ser Glu Met Ala Leu Ala

50 55 60

Gly Ser Pro Gly Cys Ser Ser Gly Pro Gln Gly Arg Leu Ser Arg Leu
65 70 75 80

Ile Phe Leu Leu Arg Arg Trp Ala Ala Arg His Val His His Gln Asp
85 90 95

Gln Gly Pro Asp Ser Phe Pro Asp Arg Phe Arg Gly Ala Glu Leu Lys
100 105 110

Glu Val Ser Ser Gln Glu Ser Asn Ala Gln Ala Asn Val Gly Ser Gln
115 120 125

Glu Pro Ala Asp Arg Gly Arg Ser Ala Trp Pro Leu Ala Lys Cys Asn
130 135 140

Thr Asn Thr Ser Asn Asn Thr Glu Glu Glu Lys Lys Thr Lys Lys Lys
145 150 155 160

Asp Ala Ile Val Val Asp Pro Ser Ser Asn Leu Tyr Tyr Arg Trp Leu
165 170 175

Thr Ala Ile Ala Leu Pro Val Phe Tyr Asn Trp Tyr Leu Leu Ile Cys
180 185 190

Arg Ala Cys Phe Asp Glu Leu Gln Ser Glu Tyr Leu Met Leu Trp Leu
195 200 205

Val Leu Asp Tyr Ser Ala Asp Val Leu Tyr Val Leu Asp Val Leu Val
210 215 220

Arg Ala Arg Thr Gly Phe Leu Glu Gln Gly Leu Met Val Ser Asp Thr
225 230 235 240

Asn Arg Leu Trp Gln His Tyr Lys Thr Thr Thr Gln Phe Lys Leu Asp
245 250 255

Val Leu Ser Leu Val Pro Thr Asp Leu Ala Tyr Leu Lys Val Gly Thr
260 265 270

Asn Tyr Pro Glu Val Arg Phe Asn Arg Leu Leu Lys Phe Ser Arg Leu
275 280 285

Phe Glu Phe Phe Asp Arg Thr Glu Thr Arg Thr Asn Tyr Pro Asn Met
290 295 300

Phe Arg Ile Gly Asn Leu Val Leu Tyr Ile Leu Ile Ile Ile His Trp
305 310 315 320

Asn Ala Cys Ile Tyr Phe Ala Ile Ser Lys Phe Ile Gly Phe Gly Thr
325 330 335

Asp Ser Trp Val Tyr Pro Asn Ile Ser Ile Pro Glu His Gly Arg Leu
340 345 350

Ser Arg Lys Tyr Ile Tyr Ser Leu Tyr Trp Ser Thr Leu Thr Leu Thr
355 360 365

Thr Ile Gly Glu Thr Pro Pro Pro Val Lys Asp Glu Glu Tyr Leu Phe
370 375 380
Val Val Val Asp Phe Leu Val Gly Val Leu Ile Phe Ala Thr Ile Val
385 390 395 400
Gly Asn Val Gly Ser Met Ile Ser Asn Met Asn Ala Ser Arg Ala Glu
405 410 415
Phe Gln Ala Lys Ile Asp Ser Ile Lys Gln Tyr Met Gln Phe Arg Lys
420 425 430
Val Thr Lys Asp Leu Glu Thr Arg Val Ile Arg Trp Phe Asp Tyr Leu
435 440 445
Trp Ala Asn Lys Lys Thr Val Asp Glu Lys Glu Val Leu Lys Ser Leu
450 455 460
Pro Asp Lys Leu Lys Ala Glu Ile Ala Ile Asn Val His Leu Asp Thr
465 470 475 480
Leu Lys Lys Val Arg Ile Phe Gln Asp Cys Glu Ala Gly Leu Leu Val
485 490 495
Glu Leu Val Leu Lys Leu Arg Pro Thr Val Phe Ser Pro Gly Asp Tyr
500 505 510
Ile Cys Lys Lys Gly Asp Ile Gly Lys Glu Met Tyr Ile Ile Asn Glu
515 520 525
Gly Lys Leu Ala Val Val Ala Asp Asp Gly Val Thr Gln Phe Val Val
530 535 540
Leu Ser Asp Gly Ser Tyr Phe Gly Glu Ile Ser Ile Leu Asn Ile Lys
545 550 555 560
Gly Ser Lys Ser Gly Asn Arg Arg Thr Ala Asn Ile Arg Ser Ile Gly
565 570 575
Tyr Ser Asp Leu Phe Cys Leu Ser Lys Asp Asp Leu Met Glu Ala Leu
580 585 590
Thr Glu Tyr Pro Glu Ala Lys Lys Ala Leu Glu Glu Lys Gly Arg Gln
595 600 605
Ile Leu Met Lys Asp Asn Leu Ile Asp Glu Glu Leu Ala Arg Ala Gly
610 615 620
Ala Asp Pro Lys Asp Leu Glu Glu Lys Val Glu Gln Leu Gly Ser Ser
625 630 635 640
Leu Asp Thr Leu Gln Thr Arg Phe Ala Arg Leu Leu Ala Glu Tyr Asn
645 650 655
Ala Thr Gln Met Lys Met Lys Gln Arg Leu Ser Gln Leu Glu Ser Gln
660 665 670
Val Lys Gly Gly Gly Asp Lys Pro Leu Ala Asp Gly Glu Val Pro Gly

675 680 685

Asp Ala Thr Lys Thr Glu Asp Lys Gln Gln

690 695

<210> 35

<211> 694

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 35

Met Ala Lys Ile Asn Thr Gln Tyr Ser His Pro Ser Arg Thr His Leu
1 5 10 15
Lys Val Lys Thr Ser Asp Arg Asp Leu Asn Arg Ala Glu Asn Gly Leu
20 25 30
Ser Arg Ala His Ser Ser Ser Glu Glu Thr Ser Ser Val Leu Gln Pro
35 40 45
Gly Ile Ala Met Glu Thr Arg Gly Leu Ala Asp Ser Gly Gln Gly Ser
50 55 60
Phe Thr Gly Gln Gly Ile Ala Arg Leu Ser Arg Leu Ile Phe Leu Leu
65 70 75 80
Arg Arg Trp Ala Ala Arg His Val His His Gln Asp Gln Gly Pro Asp
85 90 95
Ser Phe Pro Asp Arg Phe Arg Gly Ala Glu Leu Lys Glu Val Ser Ser
100 105 110
Gln Glu Ser Asn Ala Gln Ala Asn Val Gly Ser Gln Glu Pro Ala Asp
115 120 125
Arg Gly Arg Ser Ala Trp Pro Leu Ala Lys Cys Asn Thr Asn Thr Ser
130 135 140
Asn Asn Thr Glu Glu Glu Lys Lys Thr Lys Lys Lys Asp Ala Ile Val
145 150 155 160
Val Asp Pro Ser Ser Asn Leu Tyr Tyr Arg Trp Leu Thr Ala Ile Ala
165 170 175
Leu Pro Val Phe Tyr Asn Trp Tyr Leu Leu Ile Cys Arg Ala Cys Phe
180 185 190
Asp Glu Leu Gln Ser Glu Tyr Leu Met Leu Trp Leu Val Leu Asp Tyr
195 200 205
Ser Ala Asp Val Leu Tyr Val Leu Asp Val Leu Val Arg Ala Arg Thr
210 215 220
Gly Phe Leu Glu Gln Gly Leu Met Val Ser Asp Thr Asn Arg Leu Trp

225 230 235 240
Gln His Tyr Lys Thr Thr Thr Gln Phe Lys Leu Asp Val Leu Ser Leu
245 250 255
Val Pro Thr Asp Leu Ala Tyr Leu Lys Val Gly Thr Asn Tyr Pro Glu
260 265 270
Val Arg Phe Asn Arg Leu Leu Lys Phe Ser Arg Leu Phe Glu Phe Phe
275 280 285
Asp Arg Thr Glu Thr Arg Thr Asn Tyr Pro Asn Met Phe Arg Ile Gly
290 295 300
Asn Leu Val Leu Tyr Ile Leu Ile Ile Ile His Trp Asn Ala Cys Ile
305 310 315 320
Tyr Phe Ala Ile Ser Lys Phe Ile Gly Phe Gly Thr Asp Ser Trp Val
325 330 335
Tyr Pro Asn Ile Ser Ile Pro Glu His Gly Arg Leu Ser Arg Lys Tyr
340 345 350
Ile Tyr Ser Leu Tyr Trp Ser Thr Leu Thr Leu Thr Thr Ile Gly Glu
355 360 365
Thr Pro Pro Pro Val Lys Asp Glu Glu Tyr Leu Phe Val Val Val Asp
370 375 380
Phe Leu Val Gly Val Leu Ile Phe Ala Thr Ile Val Gly Asn Val Gly
385 390 395 400
Ser Met Ile Ser Asn Met Asn Ala Ser Arg Ala Glu Phe Gln Ala Lys
405 410 415
Ile Asp Ser Ile Lys Gln Tyr Met Gln Phe Arg Lys Val Thr Lys Asp
420 425 430
Leu Glu Thr Arg Val Ile Arg Trp Phe Asp Tyr Leu Trp Ala Asn Lys
435 440 445
Lys Thr Val Asp Glu Lys Glu Val Leu Lys Ser Leu Pro Asp Lys Leu
450 455 460
Lys Ala Glu Ile Ala Ile Asn Val His Leu Asp Thr Leu Lys Lys Val
465 470 475 480
Arg Ile Phe Gln Asp Cys Glu Ala Gly Leu Leu Val Glu Leu Val Leu
485 490 495
Lys Leu Arg Pro Thr Val Phe Ser Pro Gly Asp Tyr Ile Cys Lys Lys
500 505 510
Gly Asp Ile Gly Lys Glu Met Tyr Ile Ile Asn Glu Gly Lys Leu Ala
515 520 525
Val Val Ala Asp Asp Gly Val Thr Gln Phe Val Val Leu Ser Asp Gly
530 535 540

Ser Tyr Phe Gly Glu Ile Ser Ile Leu Asn Ile Lys Gly Ser Lys Ser
545 550 555 560
Gly Asn Arg Arg Thr Ala Asn Ile Arg Ser Ile Gly Tyr Ser Asp Leu
565 570 575
Phe Cys Leu Ser Lys Asp Asp Leu Met Glu Ala Leu Thr Glu Tyr Pro
580 585 590
Glu Ala Lys Lys Ala Leu Glu Glu Lys Gly Arg Gln Ile Leu Met Lys
595 600 605
Asp Asn Leu Ile Asp Glu Glu Leu Ala Arg Ala Gly Ala Asp Pro Lys
610 615 620
Asp Leu Glu Glu Lys Val Glu Gln Leu Gly Ser Ser Leu Asp Thr Leu
625 630 635 640
Gln Thr Arg Phe Ala Arg Leu Leu Ala Glu Tyr Asn Ala Thr Gln Met
645 650 655
Lys Met Lys Gln Arg Leu Ser Gln Leu Glu Ser Gln Val Lys Gly Gly
660 665 670
Gly Asp Lys Pro Leu Ala Asp Gly Glu Val Pro Gly Asp Ala Thr Lys
675 680 685
Thr Glu Asp Lys Gln Gln
690

<210> 36
<211> 348
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 36

Met Arg Lys Met Ser Glu Glu Glu Phe Tyr Leu Phe Lys Asn Ile Ser
1 5 10 15
Ser Val Gly Pro Trp Asp Gly Pro Gln Tyr His Ile Ala Pro Val Trp
20 25 30
Ala Phe Tyr Leu Gln Ala Ala Phe Met Gly Thr Val Phe Leu Ile Gly
35 40 45
Phe Pro Leu Asn Ala Met Val Leu Val Ala Thr Leu Arg Tyr Lys Lys
50 55 60
Leu Arg Gln Pro Leu Asn Tyr Ile Leu Val Asn Val Ser Phe Gly Gly
65 70 75 80
Phe Leu Leu Cys Ile Phe Ser Val Phe Pro Val Phe Val Ala Ser Cys
85 90 95

Asn Gly Tyr Phe Val Phe Gly Arg His Val Cys Ala Leu Glu Gly Phe
100 105 110
Leu Gly Thr Val Ala Gly Leu Val Thr Gly Trp Ser Leu Ala Phe Leu
115 120 125
Ala Phe Glu Arg Tyr Ile Val Ile Cys Lys Pro Phe Gly Asn Phe Arg
130 135 140
Phe Ser Ser Lys His Ala Leu Thr Val Val Leu Ala Thr Trp Thr Ile
145 150 155 160
Gly Ile Gly Val Ser Ile Pro Pro Phe Phe Gly Trp Ser Arg Phe Ile
165 170 175
Pro Glu Gly Leu Gln Cys Ser Cys Gly Pro Asp Trp Tyr Thr Val Gly
180 185 190
Thr Lys Tyr Arg Ser Glu Ser Tyr Thr Trp Phe Leu Phe Ile Phe Cys
195 200 205
Phe Ile Val Pro Leu Ser Leu Ile Cys Phe Ser Tyr Thr Gln Leu Leu
210 215 220
Arg Ala Leu Lys Ala Val Ala Ala Gln Gln Gln Glu Ser Ala Thr Thr
225 230 235 240
Gln Lys Ala Glu Arg Glu Val Ser Arg Met Val Val Val Met Val Gly
245 250 255
Ser Phe Cys Val Cys Tyr Val Pro Tyr Ala Ala Phe Ala Met Tyr Met
260 265 270
Val Asn Asn Arg Asn His Gly Leu Asp Leu Arg Leu Val Thr Ile Pro
275 280 285
Ser Phe Phe Ser Lys Ser Ala Cys Ile Tyr Asn Pro Ile Ile Tyr Cys
290 295 300
Phe Met Asn Lys Gln Phe Gln Ala Cys Ile Met Lys Met Val Cys Gly
305 310 315 320
Lys Ala Met Thr Asp Glu Ser Asp Thr Cys Ser Ser Gln Lys Thr Glu
325 330 335
Val Ser Thr Val Ser Ser Thr Gln Val Gly Pro Asn
340 345

<210> 37
<211> 364
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 37

Met Ala Gln Gln Trp Ser Leu Gln Arg Leu Ala Gly Arg His Pro Gln
1 5 10 15
Asp Ser Tyr Glu Asp Ser Thr Gln Ser Ser Ile Phe Thr Tyr Thr Asn
20 25 30
Ser Asn Ser Thr Arg Gly Pro Phe Glu Gly Pro Asn Tyr His Ile Ala
35 40 45
Pro Arg Trp Val Tyr His Leu Thr Ser Val Trp Met Ile Phe Val Val
50 55 60
Ile Ala Ser Val Phe Thr Asn Gly Leu Val Leu Ala Ala Thr Met Lys
65 70 75 80
Phe Lys Lys Leu Arg His Pro Leu Asn Trp Ile Leu Val Asn Leu Ala
85 90 95
Val Ala Asp Leu Ala Glu Thr Val Ile Ala Ser Thr Ile Ser Val Val
100 105 110
Asn Gln Val Tyr Gly Tyr Phe Val Leu Gly His Pro Met Cys Val Leu
115 120 125
Glu Gly Tyr Thr Val Ser Leu Cys Gly Ile Thr Gly Leu Trp Ser Leu
130 135 140
Ala Ile Ile Ser Trp Glu Arg Trp Met Val Val Cys Lys Pro Phe Gly
145 150 155 160
Asn Val Arg Phe Asp Ala Lys Leu Ala Ile Val Gly Ile Ala Phe Ser
165 170 175
Trp Ile Trp Ala Ala Val Trp Thr Ala Pro Pro Ile Phe Gly Trp Ser
180 185 190
Arg Tyr Trp Pro His Gly Leu Lys Thr Ser Cys Gly Pro Asp Val Phe
195 200 205
Ser Gly Ser Ser Tyr Pro Gly Val Gln Ser Tyr Met Ile Val Leu Met
210 215 220
Val Thr Cys Cys Ile Thr Pro Leu Ser Ile Ile Val Leu Cys Tyr Leu
225 230 235 240
Gln Val Trp Leu Ala Ile Arg Ala Val Ala Lys Gln Gln Lys Glu Ser
245 250 255
Glu Ser Thr Gln Lys Ala Glu Lys Glu Val Thr Arg Met Val Val Val
260 265 270
Met Val Leu Ala Phe Cys Phe Cys Trp Gly Pro Tyr Ala Phe Phe Ala
275 280 285
Cys Phe Ala Ala Ala Asn Pro Gly Tyr Pro Phe His Pro Leu Met Ala
290 295 300
Ala Leu Pro Ala Phe Phe Ala Lys Ser Ala Thr Ile Tyr Asn Pro Val

305 310 315 320
Ile Tyr Val Phe Met Asn Arg Gln Phe Arg Asn Cys Ile Leu Gln Leu
325 330 335
Phe Gly Lys Lys Val Asp Asp Gly Ser Glu Leu Ser Ser Ala Ser Lys
340 345 350
Thr Glu Val Ser Ser Val Ser Ser Val Ser Pro Ala
355 360

<210> 38
<211> 364
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 38

Met Ala Gln Gln Trp Ser Leu Gln Arg Leu Ala Gly Arg His Pro Gln
1 5 10 15
Asp Ser Tyr Glu Asp Ser Thr Gln Ser Ser Ile Phe Thr Tyr Thr Asn
20 25 30
Ser Asn Ser Thr Arg Gly Pro Phe Glu Gly Pro Asn Tyr His Ile Ala
35 40 45
Pro Arg Trp Val Tyr His Leu Thr Ser Val Trp Met Ile Phe Val Val
50 55 60
Thr Ala Ser Val Phe Thr Asn Gly Leu Val Leu Ala Ala Thr Met Lys
65 70 75 80
Phe Lys Lys Leu Arg His Pro Leu Asn Trp Ile Leu Val Asn Leu Ala
85 90 95
Val Ala Asp Leu Ala Glu Thr Val Ile Ala Ser Thr Ile Ser Ile Val
100 105 110
Asn Gln Val Ser Gly Tyr Phe Val Leu Gly His Pro Met Cys Val Leu
115 120 125
Glu Gly Tyr Thr Val Ser Leu Cys Gly Ile Thr Gly Leu Trp Ser Leu
130 135 140
Ala Ile Ile Ser Trp Glu Arg Trp Leu Val Val Cys Lys Pro Phe Gly
145 150 155 160
Asn Val Arg Phe Asp Ala Lys Leu Ala Ile Val Gly Ile Ala Phe Ser
165 170 175
Trp Ile Trp Ser Ala Val Trp Thr Ala Pro Pro Ile Phe Gly Trp Ser
180 185 190
Arg Tyr Trp Pro His Gly Leu Lys Thr Ser Cys Gly Pro Asp Val Phe

195 200 205

Ser Gly Ser Ser Tyr Pro Gly Val Gln Ser Tyr Met Ile Val Leu Met
210 215 220

Val Thr Cys Cys Ile Ile Pro Leu Ala Ile Ile Met Leu Cys Tyr Leu
225 230 235 240

Gln Val Trp Leu Ala Ile Arg Ala Val Ala Lys Gln Gln Lys Glu Ser
245 250 255

Glu Ser Thr Gln Lys Ala Glu Lys Glu Val Thr Arg Met Val Val Val
260 265 270

Met Ile Phe Ala Tyr Cys Val Cys Trp Gly Pro Tyr Thr Phe Phe Ala
275 280 285

Cys Phe Ala Ala Ala Asn Pro Gly Tyr Ala Phe His Pro Leu Met Ala
290 295 300

Ala Leu Pro Ala Tyr Phe Ala Lys Ser Ala Thr Ile Tyr Asn Pro Val
305 310 315 320

Ile Tyr Val Phe Met Asn Arg Gln Phe Arg Asn Cys Ile Leu Gln Leu
325 330 335

Phe Gly Lys Lys Val Asp Asp Gly Ser Glu Leu Ser Ser Ala Ser Lys
340 345 350

Thr Glu Val Ser Ser Val Ser Ser Val Ser Pro Ala
355 360

<210> 39

<211> 2273

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 39

Met Gly Phe Val Arg Gln Ile Gln Leu Leu Leu Trp Lys Asn Trp Thr
1 5 10 15

Leu Arg Lys Arg Gln Lys Ile Arg Phe Val Val Glu Leu Val Trp Pro
20 25 30

Leu Ser Leu Phe Leu Val Leu Ile Trp Leu Arg Asn Ala Asn Pro Leu
35 40 45

Tyr Ser His His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
50 55 60

Gly Met Leu Pro Trp Leu Gln Gly Ile Phe Cys Asn Val Asn Asn Pro
65 70 75 80

Cys Phe Gln Ser Pro Thr Pro Gly Glu Ser Pro Gly Ile Val Ser Asn

85 90 95

Tyr Asn Asn Ser Ile Leu Ala Arg Val Tyr Arg Asp Phe Gln Glu Leu
100 105 110

Leu Met Asn Ala Pro Glu Ser Gln His Leu Gly Arg Ile Trp Thr Glu
115 120 125

Leu His Ile Leu Ser Gln Phe Met Asp Thr Leu Arg Thr His Pro Glu
130 135 140

Arg Ile Ala Gly Arg Gly Ile Arg Ile Arg Asp Ile Leu Lys Asp Glu
145 150 155 160

Glu Thr Leu Thr Leu Phe Leu Ile Lys Asn Ile Gly Leu Ser Asp Ser
165 170 175

Val Val Tyr Leu Leu Ile Asn Ser Gln Val Arg Pro Glu Gln Phe Ala
180 185 190

His Gly Val Pro Asp Leu Ala Leu Lys Asp Ile Ala Cys Ser Glu Ala
195 200 205

Leu Leu Glu Arg Phe Ile Ile Phe Ser Gln Arg Arg Gly Ala Lys Thr
210 215 220

Val Arg Tyr Ala Leu Cys Ser Leu Ser Gln Gly Thr Leu Gln Trp Ile
225 230 235 240

Glu Asp Thr Leu Tyr Ala Asn Val Asp Phe Phe Lys Leu Phe Arg Val
245 250 255

Leu Pro Thr Leu Leu Asp Ser Arg Ser Gln Gly Ile Asn Leu Arg Ser
260 265 270

Trp Gly Gly Ile Leu Ser Asp Met Ser Pro Arg Ile Gln Glu Phe Ile
275 280 285

His Arg Pro Ser Met Gln Asp Leu Leu Trp Val Thr Arg Pro Leu Met
290 295 300

Gln Asn Gly Gly Pro Glu Thr Phe Thr Lys Leu Met Gly Ile Leu Ser
305 310 315 320

Asp Leu Leu Cys Gly Tyr Pro Glu Gly Gly Gly Ser Arg Val Leu Ser
325 330 335

Phe Asn Trp Tyr Glu Asp Asn Asn Tyr Lys Ala Phe Leu Gly Ile Asp
340 345 350

Ser Thr Arg Lys Asp Pro Ile Tyr Ser Tyr Asp Arg Arg Thr Thr Ser
355 360 365

Phe Cys Asn Ala Leu Ile Gln Ser Leu Glu Ser Asn Pro Leu Thr Lys
370 375 380

Ile Ala Trp Arg Ala Ala Lys Pro Leu Leu Met Gly Lys Ile Leu Tyr
385 390 395 400

Thr Pro Asp Ser Pro Ala Ala Arg Arg Ile Leu Lys Asn Ala Asn Ser
405 410 415
Thr Phe Glu Glu Leu Glu His Val Arg Lys Leu Val Lys Ala Trp Glu
420 425 430
Glu Val Gly Pro Gln Ile Trp Tyr Phe Phe Asp Asn Ser Thr Gln Met
435 440 445
Asn Met Ile Arg Asp Thr Leu Gly Asn Pro Thr Val Lys Asp Phe Leu
450 455 460
Asn Arg Gln Leu Gly Glu Glu Gly Ile Thr Ala Glu Ala Ile Leu Asn
465 470 475 480
Phe Leu Tyr Lys Gly Pro Arg Glu Ser Gln Ala Asp Asp Met Ala Asn
485 490 495
Phe Asp Trp Arg Asp Ile Phe Asn Ile Thr Asp Arg Thr Leu Arg Leu
500 505 510
Val Asn Gln Tyr Leu Glu Cys Leu Val Leu Asp Lys Phe Glu Ser Tyr
515 520 525
Asn Asp Glu Thr Gln Leu Thr Gln Arg Ala Leu Ser Leu Leu Glu Glu
530 535 540
Asn Met Phe Trp Ala Gly Val Val Phe Pro Asp Met Tyr Pro Trp Thr
545 550 555 560
Ser Ser Leu Pro Pro His Val Lys Tyr Lys Ile Arg Met Asp Ile Asp
565 570 575
Val Val Glu Lys Thr Asn Lys Ile Lys Asp Arg Tyr Trp Asp Ser Gly
580 585 590
Pro Arg Ala Asp Pro Val Glu Asp Phe Arg Tyr Ile Trp Gly Gly Phe
595 600 605
Ala Tyr Leu Gln Asp Met Val Glu Gln Gly Ile Thr Arg Ser Gln Val
610 615 620
Gln Ala Glu Ala Pro Val Gly Ile Tyr Leu Gln Gln Met Pro Tyr Pro
625 630 635 640
Cys Phe Val Asp Asp Ser Phe Met Ile Ile Leu Asn Arg Cys Phe Pro
645 650 655
Ile Phe Met Val Leu Ala Trp Ile Tyr Ser Val Ser Met Thr Val Lys
660 665 670
Ser Ile Val Leu Glu Lys Glu Leu Arg Leu Lys Glu Thr Leu Lys Asn
675 680 685
Gln Gly Val Ser Asn Ala Val Ile Trp Cys Thr Trp Phe Leu Asp Ser
690 695 700
Phe Ser Ile Met Ser Met Ser Ile Phe Leu Leu Thr Ile Phe Ile Met

705 710 715 720
His Gly Arg Ile Leu His Tyr Ser Asp Pro Phe Ile Leu Phe Leu Phe
725 730 735
Leu Leu Ala Phe Ser Thr Ala Thr Ile Met Leu Cys Phe Leu Leu Ser
740 745 750
Thr Phe Phe Ser Lys Ala Ser Leu Ala Ala Ala Cys Ser Gly Val Ile
755 760 765
Tyr Phe Thr Leu Tyr Leu Pro His Ile Leu Cys Phe Ala Trp Gln Asp
770 775 780
Arg Met Thr Ala Glu Leu Lys Lys Ala Val Ser Leu Leu Ser Pro Val
785 790 795 800
Ala Phe Gly Phe Gly Thr Glu Tyr Leu Val Arg Phe Glu Glu Gln Gly
805 810 815
Leu Gly Leu Gln Trp Ser Asn Ile Gly Asn Ser Pro Thr Glu Gly Asp
820 825 830
Glu Phe Ser Phe Leu Leu Ser Met Gln Met Met Leu Leu Asp Ala Ala
835 840 845
Val Tyr Gly Leu Leu Ala Trp Tyr Leu Asp Gln Val Phe Pro Gly Asp
850 855 860
Tyr Gly Thr Pro Leu Pro Trp Tyr Phe Leu Leu Gln Glu Ser Tyr Trp
865 870 875 880
Leu Gly Gly Glu Gly Cys Ser Thr Arg Glu Glu Arg Ala Leu Glu Lys
885 890 895
Thr Glu Pro Leu Thr Glu Glu Thr Glu Asp Pro Glu His Pro Glu Gly
900 905 910
Ile His Asp Ser Phe Phe Glu Arg Glu His Pro Gly Trp Val Pro Gly
915 920 925
Val Cys Val Lys Asn Leu Val Lys Ile Phe Glu Pro Cys Gly Arg Pro
930 935 940
Ala Val Asp Arg Leu Asn Ile Thr Phe Tyr Glu Asn Gln Ile Thr Ala
945 950 955 960
Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu Ser Ile Leu
965 970 975
Thr Gly Leu Leu Pro Pro Thr Ser Gly Thr Val Leu Val Gly Gly Arg
980 985 990
Asp Ile Glu Thr Ser Leu Asp Ala Val Arg Gln Ser Leu Gly Met Cys
995 1000 1005
Pro Gln His Asn Ile Leu Phe His His Leu Thr Val Ala Glu His
1010 1015 1020

Met Leu Phe Tyr Ala Gln Leu Lys Gly Lys Ser Gln Glu Glu Ala
1025 1030 1035
Gln Leu Glu Met Glu Ala Met Leu Glu Asp Thr Gly Leu His His
1040 1045 1050
Lys Arg Asn Glu Glu Ala Gln Asp Leu Ser Gly Gly Met Gln Arg
1055 1060 1065
Lys Leu Ser Val Ala Ile Ala Phe Val Gly Asp Ala Lys Val Val
1070 1075 1080
Ile Leu Asp Glu Pro Thr Ser Gly Val Asp Pro Tyr Ser Arg Arg
1085 1090 1095
Ser Ile Trp Asp Leu Leu Leu Lys Tyr Arg Ser Gly Arg Thr Ile
1100 1105 1110
Ile Met Ser Thr His His Met Asp Glu Ala Asp Leu Leu Gly Asp
1115 1120 1125
Arg Ile Ala Ile Ile Ala Gln Gly Arg Leu Tyr Cys Ser Gly Thr
1130 1135 1140
Pro Leu Phe Leu Lys Asn Cys Phe Gly Thr Gly Leu Tyr Leu Thr
1145 1150 1155
Leu Val Arg Lys Met Lys Asn Ile Gln Ser Gln Arg Lys Gly Ser
1160 1165 1170
Glu Gly Thr Cys Ser Cys Ser Ser Lys Gly Phe Ser Thr Thr Cys
1175 1180 1185
Pro Ala His Val Asp Asp Leu Thr Pro Glu Gln Val Leu Asp Gly
1190 1195 1200
Asp Val Asn Glu Leu Met Asp Val Val Leu His His Val Pro Glu
1205 1210 1215
Ala Lys Leu Val Glu Cys Ile Gly Gln Glu Leu Ile Phe Leu Leu
1220 1225 1230
Pro Asn Lys Asn Phe Lys His Arg Ala Tyr Ala Ser Leu Phe Arg
1235 1240 1245
Glu Leu Glu Glu Thr Leu Ala Asp Leu Gly Leu Ser Ser Phe Gly
1250 1255 1260
Ile Ser Asp Thr Pro Leu Glu Glu Ile Phe Leu Lys Val Thr Glu
1265 1270 1275
Asp Ser Asp Ser Gly Pro Leu Phe Ala Gly Gly Ala Gln Gln Lys
1280 1285 1290
Arg Glu Asn Val Asn Pro Arg His Pro Cys Leu Gly Pro Arg Glu
1295 1300 1305
Lys Ala Gly Gln Thr Pro Gln Asp Ser Asn Val Cys Ser Pro Gly

1310 1315 1320

Ala Pro Ala Ala His Pro Glu Gly Gln Pro Pro Pro Glu Pro Glu

1325 1330 1335

Cys Pro Gly Pro Gln Leu Asn Thr Gly Thr Gln Leu Val Leu Gln

1340 1345 1350

His Val Gln Ala Leu Leu Val Lys Arg Phe Gln His Thr Ile Arg

1355 1360 1365

Ser His Lys Asp Phe Leu Ala Gln Ile Val Leu Pro Ala Thr Phe

1370 1375 1380

Val Phe Leu Ala Leu Met Leu Ser Ile Val Ile Pro Pro Phe Gly

1385 1390 1395

Glu Tyr Pro Ala Leu Thr Leu His Pro Trp Ile Tyr Gly Gln Gln

1400 1405 1410

Tyr Thr Phe Phe Ser Met Asp Glu Pro Gly Ser Glu Gln Phe Thr

1415 1420 1425

Val Leu Ala Asp Val Leu Leu Asn Lys Pro Gly Phe Gly Asn Arg

1430 1435 1440

Cys Leu Lys Glu Gly Trp Leu Pro Glu Tyr Pro Cys Gly Asn Ser

1445 1450 1455

Thr Pro Trp Lys Thr Pro Ser Val Ser Pro Asn Ile Thr Gln Leu

1460 1465 1470

Phe Gln Lys Gln Lys Trp Thr Gln Val Asn Pro Ser Pro Ser Cys

1475 1480 1485

Arg Cys Ser Thr Arg Glu Lys Leu Thr Met Leu Pro Glu Cys Pro

1490 1495 1500

Glu Gly Ala Gly Gly Leu Pro Pro Pro Gln Arg Thr Gln Arg Ser

1505 1510 1515

Thr Glu Ile Leu Gln Asp Leu Thr Asp Arg Asn Ile Ser Asp Phe

1520 1525 1530

Leu Val Lys Thr Tyr Pro Ala Leu Ile Arg Ser Ser Leu Lys Ser

1535 1540 1545

Lys Phe Trp Val Asn Glu Gln Arg Tyr Gly Gly Ile Ser Ile Gly

1550 1555 1560

Gly Lys Leu Pro Val Val Pro Ile Thr Gly Glu Ala Leu Val Gly

1565 1570 1575

Phe Leu Ser Asp Leu Gly Arg Ile Met Asn Val Ser Gly Gly Pro

1580 1585 1590

Ile Thr Arg Glu Ala Ser Lys Glu Ile Pro Asp Phe Leu Lys His

1595 1600 1605

Leu Glu Thr Glu Asp Asn Ile Lys Val Trp Phe Asn Asn Lys Gly
1610 1615 1620
Trp His Ala Leu Val Ser Phe Leu Asn Val Ala His Asn Ala Ile
1625 1630 1635
Leu Arg Ala Ser Leu Pro Lys Asp Arg Ser Pro Glu Glu Tyr Gly
1640 1645 1650
Ile Thr Val Ile Ser Gln Pro Leu Asn Leu Thr Lys Glu Gln Leu
1655 1660 1665
Ser Glu Ile Thr Val Leu Thr Thr Ser Val Asp Ala Val Val Ala
1670 1675 1680
Ile Cys Val Ile Phe Ser Met Ser Phe Val Pro Ala Ser Phe Val
1685 1690 1695
Leu Tyr Leu Ile Gln Glu Arg Val Asn Lys Ser Lys His Leu Gln
1700 1705 1710
Phe Ile Ser Gly Val Ser Pro Thr Thr Tyr Trp Val Thr Asn Phe
1715 1720 1725
Leu Trp Asp Ile Met Asn Tyr Ser Val Ser Ala Gly Leu Val Val
1730 1735 1740
Gly Ile Phe Ile Gly Phe Gln Lys Lys Ala Tyr Thr Ser Pro Glu
1745 1750 1755
Asn Leu Pro Ala Leu Val Ala Leu Leu Leu Leu Tyr Gly Trp Ala
1760 1765 1770
Val Ile Pro Met Met Tyr Pro Ala Ser Phe Leu Phe Asp Val Pro
1775 1780 1785
Ser Thr Ala Tyr Val Ala Leu Ser Cys Ala Asn Leu Phe Ile Gly
1790 1795 1800
Ile Asn Ser Ser Ala Ile Thr Phe Ile Leu Glu Leu Phe Glu Asn
1805 1810 1815
Asn Arg Thr Leu Leu Arg Phe Asn Ala Val Leu Arg Lys Leu Leu
1820 1825 1830
Ile Val Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Leu
1835 1840 1845
Ala Leu Ser Gln Ala Val Thr Asp Val Tyr Ala Arg Phe Gly Glu
1850 1855 1860
Glu His Ser Ala Asn Pro Phe His Trp Asp Leu Ile Gly Lys Asn
1865 1870 1875
Leu Phe Ala Met Val Val Glu Gly Val Val Tyr Phe Leu Leu Thr
1880 1885 1890
Leu Leu Val Gln Arg His Phe Phe Leu Ser Gln Trp Ile Ala Glu

1895 1900 1905

Pro Thr Lys Glu Pro Ile Val Asp Glu Asp Asp Asp Val Ala Glu

1910 1915 1920

Glu Arg Gln Arg Ile Ile Thr Gly Gly Asn Lys Thr Asp Ile Leu

1925 1930 1935

Arg Leu His Glu Leu Thr Lys Ile Tyr Pro Gly Thr Ser Ser Pro

1940 1945 1950

Ala Val Asp Arg Leu Cys Val Gly Val Arg Pro Gly Glu Cys Phe

1955 1960 1965

Gly Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Thr Thr Phe Lys

1970 1975 1980

Met Leu Thr Gly Asp Thr Thr Val Thr Ser Gly Asp Ala Thr Val

1985 1990 1995

Ala Gly Lys Ser Ile Leu Thr Asn Ile Ser Glu Val His Gln Asn

2000 2005 2010

Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Asp Glu Leu Leu Thr

2015 2020 2025

Gly Arg Glu His Leu Tyr Leu Tyr Ala Arg Leu Arg Gly Val Pro

2030 2035 2040

Ala Glu Glu Ile Glu Lys Val Ala Asn Trp Ser Ile Lys Ser Leu

2045 2050 2055

Gly Leu Thr Val Tyr Ala Asp Cys Leu Ala Gly Thr Tyr Ser Gly

2060 2065 2070

Gly Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Cys

2075 2080 2085

Pro Pro Leu Val Leu Leu Asp Glu Pro Thr Thr Gly Met Asp Pro

2090 2095 2100

Gln Ala Arg Arg Met Leu Trp Asn Val Ile Val Ser Ile Ile Arg

2105 2110 2115

Glu Gly Arg Ala Val Val Leu Thr Ser His Ser Met Glu Glu Cys

2120 2125 2130

Glu Ala Leu Cys Thr Arg Leu Ala Ile Met Val Lys Gly Ala Phe

2135 2140 2145

Arg Cys Met Gly Thr Ile Gln His Leu Lys Ser Lys Phe Gly Asp

2150 2155 2160

Gly Tyr Ile Val Thr Met Lys Ile Lys Ser Pro Lys Asp Asp Leu

2165 2170 2175

Leu Pro Asp Leu Asn Pro Val Glu Gln Phe Phe Gln Gly Asn Phe

2180 2185 2190

Pro Gly Ser Val Gln Arg Glu Arg His Tyr Asn Met Leu Gln Phe
2195 2200 2205
Gln Val Ser Ser Ser Ser Leu Ala Arg Ile Phe Gln Leu Leu Leu
2210 2215 2220
Ser His Lys Asp Ser Leu Leu Ile Glu Glu Tyr Ser Val Thr Gln
2225 2230 2235
Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Gln Gln Thr
2240 2245 2250
Glu Ser His Asp Leu Pro Leu His Pro Arg Ala Ala Gly Ala Ser
2255 2260 2265
Arg Gln Ala Gln Asp
2270

<210> 40
<211> 533
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 40

Met Ser Ile Gln Val Glu His Pro Ala Gly Gly Tyr Lys Lys Leu Phe
1 5 10 15
Glu Thr Val Glu Glu Leu Ser Ser Pro Leu Thr Ala His Val Thr Gly
20 25 30
Arg Ile Pro Leu Trp Leu Thr Gly Ser Leu Leu Arg Cys Gly Pro Gly
35 40 45
Leu Phe Glu Val Gly Ser Glu Pro Phe Tyr His Leu Phe Asp Gly Gln
50 55 60
Ala Leu Leu His Lys Phe Asp Phe Lys Glu Gly His Val Thr Tyr His
65 70 75 80
Arg Arg Phe Ile Arg Thr Asp Ala Tyr Val Arg Ala Met Thr Glu Lys
85 90 95
Arg Ile Val Ile Thr Glu Phe Gly Thr Cys Ala Phe Pro Asp Pro Cys
100 105 110
Lys Asn Ile Phe Ser Arg Phe Phe Ser Tyr Phe Arg Gly Val Glu Val
115 120 125
Thr Asp Asn Ala Leu Val Asn Val Tyr Pro Val Gly Glu Asp Tyr Tyr
130 135 140
Ala Cys Thr Glu Thr Asn Phe Ile Thr Lys Ile Asn Pro Glu Thr Leu
145 150 155 160

Glu Thr Ile Lys Gln Val Asp Leu Cys Asn Tyr Val Ser Val Asn Gly
165 170 175
Ala Thr Ala His Pro His Ile Glu Asn Asp Gly Thr Val Tyr Asn Ile
180 185 190
Gly Asn Cys Phe Gly Lys Asn Phe Ser Ile Ala Tyr Asn Ile Val Lys
195 200 205
Ile Pro Pro Leu Gln Ala Asp Lys Glu Asp Pro Ile Ser Lys Ser Glu
210 215 220
Ile Val Val Gln Phe Pro Cys Ser Asp Arg Phe Lys Pro Ser Tyr Val
225 230 235 240
His Ser Phe Gly Leu Thr Pro Asn Tyr Ile Val Phe Val Glu Thr Pro
245 250 255
Val Lys Ile Asn Leu Phe Lys Phe Leu Ser Ser Trp Ser Leu Trp Gly
260 265 270
Ala Asn Tyr Met Asp Cys Phe Glu Ser Asn Glu Thr Met Gly Val Trp
275 280 285
Leu His Ile Ala Asp Lys Lys Arg Lys Lys Tyr Leu Asn Asn Lys Tyr
290 295 300
Arg Thr Ser Pro Phe Asn Leu Phe His His Ile Asn Thr Tyr Glu Asp
305 310 315 320
Asn Gly Phe Leu Ile Val Asp Leu Cys Cys Trp Lys Gly Phe Glu Phe
325 330 335
Val Tyr Asn Tyr Leu Tyr Leu Ala Asn Leu Arg Glu Asn Trp Glu Glu
340 345 350
Val Lys Lys Asn Ala Arg Lys Ala Pro Gln Pro Glu Val Arg Arg Tyr
355 360 365
Val Leu Pro Leu Asn Ile Asp Lys Ala Asp Thr Gly Lys Asn Leu Val
370 375 380
Thr Leu Pro Asn Thr Thr Ala Thr Ala Ile Leu Cys Ser Asp Glu Thr
385 390 395 400
Ile Trp Leu Glu Pro Glu Val Leu Phe Ser Gly Pro Arg Gln Ala Phe
405 410 415
Glu Phe Pro Gln Ile Asn Tyr Gln Lys Tyr Cys Gly Lys Pro Tyr Thr
420 425 430
Tyr Ala Tyr Gly Leu Gly Leu Asn His Phe Val Pro Asp Arg Leu Cys
435 440 445
Lys Leu Asn Val Lys Thr Lys Glu Thr Trp Val Trp Gln Glu Pro Asp
450 455 460
Ser Tyr Pro Ser Glu Pro Ile Phe Val Ser His Pro Asp Ala Leu Glu

465 470 475 480
Glu Asp Asp Gly Val Val Leu Ser Val Val Val Ser Pro Gly Ala Gly
485 490 495
Gln Lys Pro Ala Tyr Leu Leu Ile Leu Asn Ala Lys Asp Leu Ser Glu
500 505 510
Val Ala Arg Ala Glu Val Glu Ile Asn Ile Pro Val Thr Phe His Gly
515 520 525
Leu Phe Lys Lys Ser
530

<210> 41
<211> 317
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 41

Met Ser Glu Gly Val Gly Thr Phe Arg Met Val Pro Glu Glu Glu Gln
1 5 10 15
Glu Leu Arg Ala Gln Leu Glu Gln Leu Thr Thr Lys Asp His Gly Pro
20 25 30
Val Phe Gly Pro Cys Ser Gln Leu Pro Arg His Thr Leu Gln Lys Ala
35 40 45
Lys Asp Glu Leu Asn Glu Arg Glu Glu Thr Arg Glu Glu Ala Val Arg
50 55 60
Glu Leu Gln Glu Met Val Gln Ala Gln Ala Ala Ser Gly Glu Glu Leu
65 70 75 80
Ala Val Ala Val Ala Glu Arg Val Gln Glu Lys Asp Ser Gly Phe Phe
85 90 95
Leu Arg Phe Ile Arg Ala Arg Lys Phe Asn Val Gly Arg Ala Tyr Glu
100 105 110
Leu Leu Arg Gly Tyr Val Asn Phe Arg Leu Gln Tyr Pro Glu Leu Phe
115 120 125
Asp Ser Leu Ser Pro Glu Ala Val Arg Cys Thr Ile Glu Ala Gly Tyr
130 135 140
Pro Gly Val Leu Ser Ser Arg Asp Lys Tyr Gly Arg Val Val Met Leu
145 150 155 160
Phe Asn Ile Glu Asn Trp Gln Ser Gln Glu Ile Thr Phe Asp Glu Ile
165 170 175
Leu Gln Ala Tyr Cys Phe Ile Leu Glu Lys Leu Leu Glu Asn Glu Glu

180 185 190
Thr Gln Ile Asn Gly Phe Cys Ile Ile Glu Asn Phe Lys Gly Phe Thr
195 200 205
Met Gln Gln Ala Ala Ser Leu Arg Thr Ser Asp Leu Arg Lys Met Val
210 215 220
Asp Met Leu Gln Asp Ser Phe Pro Ala Arg Phe Lys Ala Ile His Phe
225 230 235 240
Ile His Gln Pro Trp Tyr Phe Thr Thr Thr Tyr Asn Val Val Lys Pro
245 250 255
Phe Leu Lys Ser Lys Leu Leu Glu Arg Val Phe Val His Gly Asp Asp
260 265 270
Leu Ser Gly Phe Tyr Gln Glu Ile Asp Glu Asn Ile Leu Pro Ser Asp
275 280 285
Phe Gly Gly Thr Leu Pro Lys Tyr Asp Gly Lys Ala Val Ala Glu Gln
290 295 300
Leu Phe Gly Pro Gln Ala Gln Ala Glu Asn Thr Ala Phe
305 310 315
<210> 42
<211> 346
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 42
Met Ala Leu Leu Lys Val Lys Phe Asp Gln Lys Lys Arg Val Lys Leu
1 5 10 15
Ala Gln Gly Leu Trp Leu Met Asn Trp Phe Ser Val Leu Ala Gly Ile
20 25 30
Ile Ile Phe Ser Leu Gly Leu Phe Leu Lys Ile Glu Leu Arg Lys Arg
35 40 45
Ser Asp Val Met Asn Asn Ser Glu Ser His Phe Val Pro Asn Ser Leu
50 55 60
Ile Gly Met Gly Val Leu Ser Cys Val Phe Asn Ser Leu Ala Gly Lys
65 70 75 80
Ile Cys Tyr Asp Ala Leu Asp Pro Ala Lys Tyr Ala Arg Trp Lys Pro
85 90 95
Trp Leu Lys Pro Tyr Leu Ala Ile Cys Val Leu Phe Asn Ile Ile Leu
100 105 110
Phe Leu Val Ala Leu Cys Cys Phe Leu Leu Arg Gly Ser Leu Glu Asn

115 120 125
Thr Leu Gly Gln Gly Leu Lys Asn Gly Met Lys Tyr Tyr Arg Asp Thr
130 135 140
Asp Thr Pro Gly Arg Cys Phe Met Lys Lys Thr Ile Asp Met Leu Gln
145 150 155 160
Ile Glu Phe Lys Cys Cys Gly Asn Asn Gly Phe Arg Asp Trp Phe Glu
165 170 175
Ile Gln Trp Ile Ser Asn Arg Tyr Leu Asp Phe Ser Ser Lys Glu Val
180 185 190
Lys Asp Arg Ile Lys Ser Asn Val Asp Gly Arg Tyr Leu Val Asp Gly
195 200 205
Val Pro Phe Ser Cys Cys Asn Pro Ser Ser Pro Arg Pro Cys Ile Gln
210 215 220
Tyr Gln Ile Thr Asn Asn Ser Ala His Tyr Ser Tyr Asp His Gln Thr
225 230 235 240
Glu Glu Leu Asn Leu Trp Val Arg Gly Cys Arg Ala Ala Leu Leu Ser
245 250 255
Tyr Tyr Ser Ser Leu Met Asn Ser Met Gly Val Val Thr Leu Leu Ile
260 265 270
Trp Leu Phe Glu Val Thr Ile Thr Ile Gly Leu Arg Tyr Leu Gln Thr
275 280 285
Ser Leu Asp Gly Val Ser Asn Pro Glu Glu Ser Glu Ser Glu Ser Gln
290 295 300
Gly Trp Leu Leu Glu Arg Ser Val Pro Glu Thr Trp Lys Ala Phe Leu
305 310 315 320
Glu Ser Val Lys Lys Leu Gly Lys Gly Asn Gln Val Glu Ala Glu Gly
325 330 335
Ala Asp Ala Gly Gln Ala Pro Glu Ala Gly
340 345

<210> 43
<211> 405
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 43
Met Ala Ala Ser Gly Lys Thr Ser Lys Ser Glu Pro Asn His Val Ile
1 5 10 15
Phe Lys Lys Ile Ser Arg Asp Lys Ser Val Thr Ile Tyr Leu Gly Asn

20 25 30

Arg Asp Tyr Ile Asp His Val Ser Gln Val Gln Pro Val Asp Gly Val

35 40 45

Val Leu Val Asp Pro Asp Leu Val Lys Gly Lys Lys Val Tyr Val Thr

50 55 60

Leu Thr Cys Ala Phe Arg Tyr Gly Gln Glu Asp Ile Asp Val Ile Gly

65 70 75 80

Leu Thr Phe Arg Arg Asp Leu Tyr Phe Ser Arg Val Gln Val Tyr Pro

85 90 95

Pro Val Gly Ala Ala Ser Thr Pro Thr Lys Leu Gln Glu Ser Leu Leu

100 105 110

Lys Lys Leu Gly Ser Asn Thr Tyr Pro Phe Leu Leu Thr Phe Pro Asp

115 120 125

Tyr Leu Pro Cys Ser Val Met Leu Gln Pro Ala Pro Gln Asp Ser Gly

130 135 140

Lys Ser Cys Gly Val Asp Phe Glu Val Lys Ala Phe Ala Thr Asp Ser

145 150 155 160

Thr Asp Ala Glu Glu Asp Lys Ile Pro Lys Lys Ser Ser Val Arg Leu

165 170 175

Leu Ile Arg Lys Val Gln His Ala Pro Leu Glu Met Gly Pro Gln Pro

180 185 190

Arg Ala Glu Ala Ala Trp Gln Phe Phe Met Ser Asp Lys Pro Leu His

195 200 205

Leu Ala Val Ser Leu Asn Lys Glu Ile Tyr Phe His Gly Glu Pro Ile

210 215 220

Pro Val Thr Val Thr Val Thr Asn Asn Thr Glu Lys Thr Val Lys Lys

225 230 235 240

Ile Lys Ala Phe Val Glu Gln Val Ala Asn Val Val Leu Tyr Ser Ser

245 250 255

Asp Tyr Tyr Val Lys Pro Val Ala Met Glu Glu Ala Gln Glu Lys Val

260 265 270

Pro Pro Asn Ser Thr Leu Thr Lys Thr Leu Thr Leu Leu Pro Leu Leu

275 280 285

Ala Asn Asn Arg Glu Arg Arg Gly Ile Ala Leu Asp Gly Lys Ile Lys

290 295 300

His Glu Asp Thr Asn Leu Ala Ser Ser Thr Ile Ile Lys Glu Gly Ile

305 310 315 320

Asp Arg Thr Val Leu Gly Ile Leu Val Ser Tyr Gln Ile Lys Val Lys

325 330 335

Leu Thr Val Ser Gly Phe Leu Gly Glu Leu Thr Ser Ser Glu Val Ala
340 345 350
Thr Glu Val Pro Phe Arg Leu Met His Pro Gln Pro Glu Asp Pro Ala
355 360 365
Lys Glu Ser Tyr Gln Asp Ala Asn Leu Val Phe Glu Glu Phe Ala Arg
370 375 380
His Asn Leu Lys Asp Ala Gly Glu Ala Glu Glu Gly Lys Arg Asp Lys
385 390 395 400
Asn Asp Val Asp Glu
405

<210> 44
<211> 350
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 44

Met Gly Ala Gly Ala Ser Ala Glu Glu Lys His Ser Arg Glu Leu Glu
1 5 10 15
Lys Lys Leu Lys Glu Asp Ala Glu Lys Asp Ala Arg Thr Val Lys Leu
20 25 30
Leu Leu Leu Gly Ala Gly Glu Ser Gly Lys Ser Thr Ile Val Lys Gln
35 40 45
Met Lys Ile Ile His Gln Asp Gly Tyr Ser Leu Glu Glu Cys Leu Glu
50 55 60
Phe Ile Ala Ile Ile Tyr Gly Asn Thr Leu Gln Ser Ile Leu Ala Ile
65 70 75 80
Val Arg Ala Met Thr Thr Leu Asn Ile Gln Tyr Gly Asp Ser Ala Arg
85 90 95
Gln Asp Asp Ala Arg Lys Leu Met His Met Ala Asp Thr Ile Glu Glu
100 105 110
Gly Thr Met Pro Lys Glu Met Ser Asp Ile Ile Gln Arg Leu Trp Lys
115 120 125
Asp Ser Gly Ile Gln Ala Cys Phe Glu Arg Ala Ser Glu Tyr Gln Leu
130 135 140
Asn Asp Ser Ala Gly Tyr Tyr Leu Ser Asp Leu Glu Arg Leu Val Thr
145 150 155 160
Pro Gly Tyr Val Pro Thr Glu Gln Asp Val Leu Arg Ser Arg Val Lys
165 170 175

Thr Thr Gly Ile Ile Glu Thr Gln Phe Ser Phe Lys Asp Leu Asn Phe
180 185 190
Arg Met Phe Asp Val Gly Gly Gln Arg Ser Glu Arg Lys Lys Trp Ile
195 200 205
His Cys Phe Glu Gly Val Thr Cys Ile Ile Phe Ile Ala Ala Leu Ser
210 215 220
Ala Tyr Asp Met Val Leu Val Glu Asp Asp Glu Val Asn Arg Met His
225 230 235 240
Glu Ser Leu His Leu Phe Asn Ser Ile Cys Asn His Arg Tyr Phe Ala
245 250 255
Thr Thr Ser Ile Val Leu Phe Leu Asn Lys Lys Asp Val Phe Phe Glu
260 265 270
Lys Ile Lys Lys Ala His Leu Ser Ile Cys Phe Pro Asp Tyr Asp Gly
275 280 285
Pro Asn Thr Tyr Glu Asp Ala Gly Asn Tyr Ile Lys Val Gln Phe Leu
290 295 300
Glu Leu Asn Met Arg Arg Asp Val Lys Glu Ile Tyr Ser His Met Thr
305 310 315 320
Cys Ala Thr Asp Thr Gln Asn Val Lys Phe Val Phe Asp Ala Val Thr
325 330 335
Asp Ile Ile Ile Lys Glu Asn Leu Lys Asp Cys Gly Leu Phe
340 345 350

<210> 45
<211> 201
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 45

Met Gly Asn Val Met Glu Gly Lys Ser Val Glu Glu Leu Ser Ser Thr
1 5 10 15
Glu Cys His Gln Trp Tyr Lys Lys Phe Met Thr Glu Cys Pro Ser Gly
20 25 30
Gln Leu Thr Leu Tyr Glu Phe Arg Gln Phe Phe Gly Leu Lys Asn Leu
35 40 45
Ser Pro Ser Ala Ser Gln Tyr Val Glu Gln Met Phe Glu Thr Phe Asp
50 55 60
Phe Asn Lys Asp Gly Tyr Ile Asp Phe Met Glu Tyr Val Ala Ala Leu
65 70 75 80

Ser Leu Val Leu Lys Gly Lys Val Glu Gln Lys Leu Arg Trp Tyr Phe
85 90 95
Lys Leu Tyr Asp Val Asp Gly Asn Gly Cys Ile Asp Arg Asp Glu Leu
100 105 110
Leu Thr Ile Ile Gln Ala Ile Arg Ala Ile Asn Pro Cys Ser Asp Thr
115 120 125
Thr Met Thr Ala Glu Glu Phe Thr Asp Thr Val Phe Ser Lys Ile Asp
130 135 140
Val Asn Gly Asp Gly Glu Leu Ser Leu Glu Glu Phe Ile Glu Gly Val
145 150 155 160
Gln Lys Asp Gln Met Leu Leu Asp Thr Leu Thr Arg Ser Leu Asp Leu
165 170 175
Thr Arg Ile Val Arg Arg Leu Gln Asn Gly Glu Gln Asp Glu Glu Gly
180 185 190
Ala Asp Glu Ala Ala Glu Ala Ala Gly
195 200

<210> 46
<211> 1103
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 46

Met Thr Ala Cys Ala Arg Arg Ala Gly Gly Leu Pro Asp Pro Gly Leu
1 5 10 15
Cys Gly Pro Ala Trp Trp Ala Pro Ser Leu Pro Arg Leu Pro Arg Ala
20 25 30
Leu Pro Arg Leu Pro Leu Leu Leu Leu Leu Leu Leu Leu Gln Pro Pro
35 40 45
Ala Leu Ser Ala Val Phe Thr Val Gly Val Leu Gly Pro Trp Ala Cys
50 55 60
Asp Pro Ile Phe Ser Arg Ala Arg Pro Asp Leu Ala Ala Arg Leu Ala
65 70 75 80
Ala Ala Arg Leu Asn Arg Asp Pro Gly Leu Ala Gly Gly Pro Arg Phe
85 90 95
Glu Val Ala Leu Leu Pro Glu Pro Cys Arg Thr Pro Gly Ser Leu Gly
100 105 110
Ala Val Ser Ser Ala Leu Ala Arg Val Ser Gly Leu Val Gly Pro Val
115 120 125

Asn Pro Ala Ala Cys Arg Pro Ala Glu Leu Leu Ala Glu Glu Ala Gly
130 135 140
Ile Ala Leu Val Pro Trp Gly Cys Pro Trp Thr Gln Ala Glu Gly Thr
145 150 155 160
Thr Ala Pro Ala Val Thr Pro Ala Ala Asp Ala Leu Tyr Ala Leu Leu
165 170 175
Arg Ala Phe Gly Trp Ala Arg Val Ala Leu Val Thr Ala Pro Gln Asp
180 185 190
Leu Trp Val Glu Ala Gly Arg Ser Leu Ser Thr Ala Leu Arg Ala Arg
195 200 205
Gly Leu Pro Val Ala Ser Val Thr Ser Met Glu Pro Leu Asp Leu Ser
210 215 220
Gly Ala Arg Glu Ala Leu Arg Lys Val Arg Asp Gly Pro Arg Val Thr
225 230 235 240
Ala Val Ile Met Val Met His Ser Val Leu Leu Gly Gly Glu Glu Gln
245 250 255
Arg Tyr Leu Leu Glu Ala Ala Glu Glu Leu Gly Leu Thr Asp Gly Ser
260 265 270
Leu Val Phe Leu Pro Phe Asp Thr Ile His Tyr Ala Leu Ser Pro Gly
275 280 285
Pro Glu Ala Leu Ala Ala Leu Ala Asn Ser Ser Gln Leu Arg Arg Ala
290 295 300
His Asp Ala Val Leu Thr Leu Thr Arg His Cys Pro Ser Glu Gly Ser
305 310 315 320
Val Leu Asp Ser Leu Arg Arg Ala Gln Glu Arg Arg Glu Leu Pro Ser
325 330 335
Asp Leu Asn Leu Gln Gln Val Ser Pro Leu Phe Gly Thr Ile Tyr Asp
340 345 350
Ala Val Phe Leu Leu Ala Arg Gly Val Ala Glu Ala Arg Ala Ala Ala
355 360 365
Gly Gly Arg Trp Val Ser Gly Ala Ala Val Ala Arg His Ile Arg Asp
370 375 380
Ala Gln Val Pro Gly Phe Cys Gly Asp Leu Gly Gly Asp Glu Glu Pro
385 390 395 400
Pro Phe Val Leu Leu Asp Thr Asp Ala Ala Gly Asp Arg Leu Phe Ala
405 410 415
Thr Tyr Met Leu Asp Pro Ala Arg Gly Ser Phe Leu Ser Ala Gly Thr
420 425 430
Arg Met His Phe Pro Arg Gly Gly Ser Ala Pro Gly Pro Asp Pro Ser

435 440 445

Cys Trp Phe Asp Pro Asn Asn Ile Cys Gly Gly Gly Leu Glu Pro Gly
450 455 460

Leu Val Phe Leu Gly Phe Leu Leu Val Val Gly Met Gly Leu Ala Gly
465 470 475 480

Ala Phe Leu Ala His Tyr Val Arg His Arg Leu Leu His Met Gln Met
485 490 495

Val Ser Gly Pro Asn Lys Ile Ile Leu Thr Val Asp Asp Ile Thr Phe
500 505 510

Leu His Pro His Gly Gly Thr Ser Arg Lys Val Ala Gln Gly Ser Arg
515 520 525

Ser Ser Leu Gly Ala Arg Ser Met Ser Asp Ile Arg Ser Gly Pro Ser
530 535 540

Gln His Leu Asp Ser Pro Asn Ile Gly Val Tyr Glu Gly Asp Arg Val
545 550 555 560

Trp Leu Lys Lys Phe Pro Gly Asp Gln His Ile Ala Ile Arg Pro Ala
565 570 575

Thr Lys Thr Ala Phe Ser Lys Leu Gln Glu Leu Arg His Glu Asn Val
580 585 590

Ala Leu Tyr Leu Gly Leu Phe Leu Ala Arg Gly Ala Glu Gly Pro Ala
595 600 605

Ala Leu Trp Glu Gly Asn Leu Ala Val Val Ser Glu His Cys Thr Arg
610 615 620

Gly Ser Leu Gln Asp Leu Leu Ala Gln Arg Glu Ile Lys Leu Asp Trp
625 630 635 640

Met Phe Lys Ser Ser Leu Leu Leu Asp Leu Ile Lys Gly Ile Arg Tyr
645 650 655

Leu His His Arg Gly Val Ala His Gly Arg Leu Lys Ser Arg Asn Cys
660 665 670

Ile Val Asp Gly Arg Phe Val Leu Lys Ile Thr Asp His Gly His Gly
675 680 685

Arg Leu Leu Glu Ala Gln Lys Val Leu Pro Glu Pro Pro Arg Ala Glu
690 695 700

Asp Gln Leu Trp Thr Ala Pro Glu Leu Leu Arg Asp Pro Ala Leu Glu
705 710 715 720

Arg Arg Gly Thr Leu Ala Gly Asp Val Phe Ser Leu Ala Ile Ile Met
725 730 735

Gln Glu Val Val Cys Arg Ser Ala Pro Tyr Ala Met Leu Glu Leu Thr
740 745 750

Pro Glu Glu Val Val Gln Arg Val Arg Ser Pro Pro Pro Leu Cys Arg
755 760 765
Pro Leu Val Ser Met Asp Gln Ala Pro Val Glu Cys Ile Leu Leu Met
770 775 780
Lys Gln Cys Trp Ala Glu Gln Pro Glu Leu Arg Pro Ser Met Asp His
785 790 795 800
Thr Phe Asp Leu Phe Lys Asn Ile Asn Lys Gly Arg Lys Thr Asn Ile
805 810 815
Ile Asp Ser Met Leu Arg Met Leu Glu Gln Tyr Ser Ser Asn Leu Glu
820 825 830
Asp Leu Ile Arg Glu Arg Thr Glu Glu Leu Glu Leu Glu Lys Gln Lys
835 840 845
Thr Asp Arg Leu Leu Thr Gln Met Leu Pro Pro Ser Val Ala Glu Ala
850 855 860
Leu Lys Thr Gly Thr Pro Val Glu Pro Glu Tyr Phe Glu Gln Val Thr
865 870 875 880
Leu Tyr Phe Ser Asp Ile Val Gly Phe Thr Thr Ile Ser Ala Met Ser
885 890 895
Glu Pro Ile Glu Val Val Asp Leu Leu Asn Asp Leu Tyr Thr Leu Phe
900 905 910
Asp Ala Ile Ile Gly Ser His Asp Val Tyr Lys Val Glu Thr Ile Gly
915 920 925
Asp Ala Tyr Met Val Ala Ser Gly Leu Pro Gln Arg Asn Gly Gln Arg
930 935 940
His Ala Ala Glu Ile Ala Asn Met Ser Leu Asp Ile Leu Ser Ala Val
945 950 955 960
Gly Thr Phe Arg Met Arg His Met Pro Glu Val Pro Val Arg Ile Arg
965 970 975
Ile Gly Leu His Ser Gly Pro Cys Val Ala Gly Val Val Gly Leu Thr
980 985 990
Met Pro Arg Tyr Cys Leu Phe Gly Asp Thr Val Asn Thr Ala Ser Arg
995 1000 1005
Met Glu Ser Thr Gly Leu Pro Tyr Arg Ile His Val Asn Leu Ser
1010 1015 1020
Thr Val Gly Ile Leu Arg Ala Leu Asp Ser Gly Tyr Gln Val Glu
1025 1030 1035
Leu Arg Gly Arg Thr Glu Leu Lys Gly Lys Gly Ala Glu Asp Thr
1040 1045 1050
Phe Trp Leu Val Gly Arg Arg Gly Phe Asn Lys Pro Ile Pro Lys

1055 1060 1065

Pro Pro Asp Leu Gln Pro Gly Ser Ser Asn His Gly Ile Ser Leu

1070 1075 1080

Gln Glu Ile Pro Pro Glu Arg Arg Arg Lys Leu Glu Lys Ala Arg

1085 1090 1095

Pro Gly Gln Phe Ser

1100

<210> 47

<211> 676

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 47

Met Ala Lys Ile Asn Thr Gln Tyr Ser His Pro Ser Arg Thr His Leu

1 5 10 15

Lys Val Lys Thr Ser Asp Arg Asp Leu Asn Arg Ala Glu Asn Gly Leu

20 25 30

Ser Arg Ala His Ser Ser Ser Glu Glu Thr Ser Ser Val Leu Gln Pro

35 40 45

Gly Ile Ala Met Glu Thr Arg Gly Leu Ala Asp Ser Gly Gln Gly Ser

50 55 60

Phe Thr Gly Gln Gly Ile Ala Arg Leu Ser Arg Leu Ile Phe Leu Leu

65 70 75 80

Arg Arg Trp Ala Ala Arg His Val His His Gln Asp Gln Gly Pro Asp

85 90 95

Ser Phe Pro Asp Arg Phe Arg Gly Ala Glu Leu Lys Glu Val Ser Ser

100 105 110

Gln Glu Ser Asn Ala Gln Ala Asn Val Gly Ser Gln Glu Pro Ala Asp

115 120 125

Arg Gly Arg Arg Lys Lys Thr Lys Lys Lys Asp Ala Ile Val Val Asp

130 135 140

Pro Ser Ser Asn Leu Tyr Tyr Arg Trp Leu Thr Ala Ile Ala Leu Pro

145 150 155 160

Val Phe Tyr Asn Trp Tyr Leu Leu Ile Cys Arg Ala Cys Phe Asp Glu

165 170 175

Leu Gln Ser Glu Tyr Leu Met Leu Trp Leu Val Leu Asp Tyr Ser Ala

180 185 190

Asp Val Leu Tyr Val Leu Asp Val Leu Val Arg Ala Arg Thr Gly Phe

195 200 205

Leu Glu Gln Gly Leu Met Val Ser Asp Thr Asn Arg Leu Trp Gln His

210 215 220

Tyr Lys Thr Thr Thr Gln Phe Lys Leu Asp Val Leu Ser Leu Val Pro

225 230 235 240

Thr Asp Leu Ala Tyr Leu Lys Val Gly Thr Asn Tyr Pro Glu Val Arg

245 250 255

Phe Asn Arg Leu Leu Lys Phe Ser Arg Leu Phe Glu Phe Phe Asp Arg

260 265 270

Thr Glu Thr Arg Thr Asn Tyr Pro Asn Met Phe Arg Ile Gly Asn Leu

275 280 285

Val Leu Tyr Ile Leu Ile Ile Ile His Trp Asn Ala Cys Ile Tyr Phe

290 295 300

Ala Ile Ser Lys Phe Ile Gly Phe Gly Thr Asp Ser Trp Val Tyr Pro

305 310 315 320

Asn Ile Ser Ile Pro Glu His Gly Arg Leu Ser Arg Lys Tyr Ile Tyr

325 330 335

Ser Leu Tyr Trp Ser Thr Leu Thr Leu Thr Thr Ile Gly Glu Thr Pro

340 345 350

Pro Pro Val Lys Asp Glu Glu Tyr Leu Phe Val Val Val Asp Phe Leu

355 360 365

Val Gly Val Leu Ile Phe Ala Thr Ile Val Gly Asn Val Gly Ser Met

370 375 380

Ile Ser Asn Met Asn Ala Ser Arg Ala Glu Phe Gln Ala Lys Ile Asp

385 390 395 400

Ser Ile Lys Gln Tyr Met Gln Phe Arg Lys Val Thr Lys Asp Leu Glu

405 410 415

Thr Arg Val Ile Arg Trp Phe Asp Tyr Leu Trp Ala Asn Lys Lys Thr

420 425 430

Val Asp Glu Lys Glu Val Leu Lys Ser Leu Pro Asp Lys Leu Lys Ala

435 440 445

Glu Ile Ala Ile Asn Val His Leu Asp Thr Leu Lys Lys Val Arg Ile

450 455 460

Phe Gln Asp Cys Glu Ala Gly Leu Leu Val Glu Leu Val Leu Lys Leu

465 470 475 480

Arg Pro Thr Val Phe Ser Pro Gly Asp Tyr Ile Cys Lys Lys Gly Asp

485 490 495

Ile Gly Lys Glu Met Tyr Ile Ile Asn Glu Gly Lys Leu Ala Val Val

500 505 510

Ala Asp Asp Gly Val Thr Gln Phe Val Val Leu Ser Asp Gly Ser Tyr
515 520 525
Phe Gly Glu Ile Ser Ile Leu Asn Ile Lys Gly Ser Lys Ser Gly Asn
530 535 540
Arg Arg Thr Ala Asn Ile Arg Ser Ile Gly Tyr Ser Asp Leu Phe Cys
545 550 555 560
Leu Ser Lys Asp Asp Leu Met Glu Ala Leu Thr Glu Tyr Pro Glu Ala
565 570 575
Lys Lys Ala Leu Glu Glu Lys Gly Arg Gln Ile Leu Met Lys Asp Asn
580 585 590
Leu Ile Asp Glu Glu Leu Ala Arg Ala Gly Ala Asp Pro Lys Asp Leu
595 600 605
Glu Glu Lys Val Glu Gln Leu Gly Ser Ser Leu Asp Thr Leu Gln Thr
610 615 620
Arg Phe Ala Arg Leu Leu Ala Glu Tyr Asn Ala Thr Gln Met Lys Met
625 630 635 640
Lys Gln Arg Leu Ser Gln Leu Glu Ser Gln Val Lys Gly Gly Gly Asp
645 650 655
Lys Pro Leu Ala Asp Gly Glu Val Pro Gly Asp Ala Thr Lys Thr Glu
660 665 670
Asp Lys Gln Gln
675

<210> 48
<211> 694
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 48

Met Ala Lys Ile Asn Thr Gln Tyr Ser His Pro Ser Arg Thr His Leu
1 5 10 15
Lys Val Lys Thr Ser Asp Arg Asp Leu Asn Arg Ala Glu Asn Gly Leu
20 25 30
Ser Arg Ala His Ser Ser Ser Glu Glu Thr Ser Ser Val Leu Gln Pro
35 40 45
Gly Ile Ala Met Glu Thr Arg Gly Leu Ala Asp Ser Gly Gln Gly Ser
50 55 60
Phe Thr Gly Gln Gly Ile Ala Arg Leu Ser Arg Leu Ile Phe Leu Leu
65 70 75 80

Arg Arg Trp Ala Ala Arg His Val His His Gln Asp Gln Gly Pro Asp
85 90 95
Ser Phe Pro Asp Arg Phe Arg Gly Ala Glu Leu Lys Glu Val Ser Ser
100 105 110
Gln Glu Ser Asn Ala Gln Ala Asn Val Gly Ser Gln Glu Pro Ala Asp
115 120 125
Arg Gly Arg Ser Ala Trp Pro Leu Ala Lys Cys Asn Thr Asn Thr Ser
130 135 140
Asn Asn Thr Glu Glu Glu Lys Lys Thr Lys Lys Lys Asp Ala Ile Val
145 150 155 160
Val Asp Pro Ser Ser Asn Leu Tyr Tyr Arg Trp Leu Thr Ala Ile Ala
165 170 175
Leu Pro Val Phe Tyr Asn Trp Tyr Leu Leu Ile Cys Arg Ala Cys Phe
180 185 190
Asp Glu Leu Gln Ser Glu Tyr Leu Met Leu Trp Leu Val Leu Asp Tyr
195 200 205
Ser Ala Asp Val Leu Tyr Val Leu Asp Val Leu Val Arg Ala Arg Thr
210 215 220
Gly Phe Leu Glu Gln Gly Leu Met Val Ser Asp Thr Asn Arg Leu Trp
225 230 235 240
Gln His Tyr Lys Thr Thr Thr Gln Phe Lys Leu Asp Val Leu Ser Leu
245 250 255
Val Pro Thr Asp Leu Ala Tyr Leu Lys Val Gly Thr Asn Tyr Pro Glu
260 265 270
Val Arg Phe Asn Arg Leu Leu Lys Phe Ser Arg Leu Phe Glu Phe Phe
275 280 285
Asp Arg Thr Glu Thr Arg Thr Asn Tyr Pro Asn Met Phe Arg Ile Gly
290 295 300
Asn Leu Val Leu Tyr Ile Leu Ile Ile Ile His Trp Asn Ala Cys Ile
305 310 315 320
Tyr Phe Ala Ile Ser Lys Phe Ile Gly Phe Gly Thr Asp Ser Trp Val
325 330 335
Tyr Pro Asn Ile Ser Ile Pro Glu His Gly Arg Leu Ser Arg Lys Tyr
340 345 350
Ile Tyr Ser Leu Tyr Trp Ser Thr Leu Thr Leu Thr Thr Ile Gly Glu
355 360 365
Thr Pro Pro Pro Val Lys Asp Glu Glu Tyr Leu Phe Val Val Val Asp
370 375 380
Phe Leu Val Gly Val Leu Ile Phe Ala Thr Ile Val Gly Asn Val Gly

385 390 395 400
Ser Met Ile Ser Asn Met Asn Ala Ser Arg Ala Glu Phe Gln Ala Lys
405 410 415
Ile Asp Ser Ile Lys Gln Tyr Met Gln Phe Arg Lys Val Thr Lys Asp
420 425 430
Leu Glu Thr Arg Val Ile Arg Trp Phe Asp Tyr Leu Trp Ala Asn Lys
435 440 445
Lys Thr Val Asp Glu Lys Glu Val Leu Lys Ser Leu Pro Asp Lys Leu
450 455 460
Lys Ala Glu Ile Ala Ile Asn Val His Leu Asp Thr Leu Lys Lys Val
465 470 475 480
Arg Ile Phe Gln Asp Cys Glu Ala Gly Leu Leu Val Glu Leu Val Leu
485 490 495
Lys Leu Arg Pro Thr Val Phe Ser Pro Gly Asp Tyr Ile Cys Lys Lys
500 505 510
Gly Asp Ile Gly Lys Glu Met Tyr Ile Ile Asn Glu Gly Lys Leu Ala
515 520 525
Val Val Ala Asp Asp Gly Val Thr Gln Phe Val Val Leu Ser Asp Gly
530 535 540
Ser Tyr Phe Gly Glu Ile Ser Ile Leu Asn Ile Lys Gly Ser Lys Ser
545 550 555 560
Gly Asn Arg Arg Thr Ala Asn Ile Arg Ser Ile Gly Tyr Ser Asp Leu
565 570 575
Phe Cys Leu Ser Lys Asp Asp Leu Met Glu Ala Leu Thr Glu Tyr Pro
580 585 590
Glu Ala Lys Lys Ala Leu Glu Glu Lys Gly Arg Gln Ile Leu Met Lys
595 600 605
Asp Asn Leu Ile Asp Glu Glu Leu Ala Arg Ala Gly Ala Asp Pro Lys
610 615 620
Asp Leu Glu Glu Lys Val Glu Gln Leu Gly Ser Ser Leu Asp Thr Leu
625 630 635 640
Gln Thr Arg Phe Ala Arg Leu Leu Ala Glu Tyr Asn Ala Thr Gln Met
645 650 655
Lys Met Lys Gln Arg Leu Ser Gln Leu Glu Ser Gln Val Lys Gly Gly
660 665 670
Gly Asp Lys Pro Leu Ala Asp Gly Glu Val Pro Gly Asp Ala Thr Lys
675 680 685
Thr Glu Asp Lys Gln Gln
690

<210> 49
<211> 354
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 49

Met Gly Ser Gly Ala Ser Ala Glu Asp Lys Glu Leu Ala Lys Arg Ser
1 5 10 15
Lys Glu Leu Glu Lys Lys Leu Gln Glu Asp Ala Asp Lys Glu Ala Lys
20 25 30
Thr Val Lys Leu Leu Leu Leu Gly Ala Gly Glu Ser Gly Lys Ser Thr
35 40 45
Ile Val Lys Gln Met Lys Ile Ile His Gln Asp Gly Tyr Ser Pro Glu
50 55 60
Glu Cys Leu Glu Phe Lys Ala Ile Ile Tyr Gly Asn Val Leu Gln Ser
65 70 75 80
Ile Leu Ala Ile Ile Arg Ala Met Thr Thr Leu Gly Ile Asp Tyr Ala
85 90 95
Glu Pro Ser Cys Ala Asp Asp Gly Arg Gln Leu Asn Asn Leu Ala Asp
100 105 110
Ser Ile Glu Glu Gly Thr Met Pro Pro Glu Leu Val Glu Val Ile Arg
115 120 125
Arg Leu Trp Lys Asp Gly Gly Val Gln Ala Cys Phe Glu Arg Ala Ala
130 135 140
Glu Tyr Gln Leu Asn Asp Ser Ala Ser Tyr Tyr Leu Asn Gln Leu Glu
145 150 155 160
Arg Ile Thr Asp Pro Glu Tyr Leu Pro Ser Glu Gln Asp Val Leu Arg
165 170 175
Ser Arg Val Lys Thr Thr Gly Ile Ile Glu Thr Lys Phe Ser Val Lys
180 185 190
Asp Leu Asn Phe Arg Met Phe Asp Val Gly Gly Gln Arg Ser Glu Arg
195 200 205
Lys Lys Trp Ile His Cys Phe Glu Gly Val Thr Cys Ile Ile Phe Cys
210 215 220
Ala Ala Leu Ser Ala Tyr Asp Met Val Leu Val Glu Asp Asp Glu Val
225 230 235 240
Asn Arg Met His Glu Ser Leu His Leu Phe Asn Ser Ile Cys Asn His
245 250 255

Lys Phe Phe Ala Ala Thr Ser Ile Val Leu Phe Leu Asn Lys Lys Asp
260 265 270
Leu Phe Glu Glu Lys Ile Lys Lys Val His Leu Ser Ile Cys Phe Pro
275 280 285
Glu Tyr Asp Gly Asn Asn Ser Tyr Asp Asp Ala Gly Asn Tyr Ile Lys
290 295 300
Ser Gln Phe Leu Asp Leu Asn Met Arg Lys Asp Val Lys Glu Ile Tyr
305 310 315 320
Ser His Met Thr Cys Ala Thr Asp Thr Gln Asn Val Lys Phe Val Phe
325 330 335
Asp Ala Val Thr Asp Ile Ile Ile Lys Glu Asn Leu Lys Asp Cys Gly
340 345 350
Leu Phe

<210> 50
<211> 858
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 50

Met Gly Glu Ile Asn Gln Val Ala Val Glu Lys Tyr Leu Glu Glu Asn
1 5 10 15
Pro Gln Phe Ala Lys Glu Tyr Phe Asp Arg Lys Leu Arg Val Glu Val
20 25 30
Leu Gly Glu Ile Phe Lys Asn Ser Gln Val Pro Val Gln Ser Ser Met
35 40 45
Ser Phe Ser Glu Leu Thr Gln Val Glu Glu Ser Ala Leu Cys Leu Glu
50 55 60
Leu Leu Trp Thr Val Gln Glu Glu Gly Gly Thr Pro Glu Gln Gly Val
65 70 75 80
His Arg Ala Leu Gln Arg Leu Ala His Leu Leu Gln Ala Asp Arg Cys
85 90 95
Ser Met Phe Leu Cys Arg Ser Arg Asn Gly Ile Pro Glu Val Ala Ser
100 105 110
Arg Leu Leu Asp Val Thr Pro Thr Ser Lys Phe Glu Asp Asn Leu Val
115 120 125
Gly Pro Asp Lys Glu Val Val Phe Pro Leu Asp Ile Gly Ile Val Gly
130 135 140
Trp Ala Ala His Thr Lys Lys Thr His Asn Val Pro Asp Val Lys Lys

145 150 155 160
Asn Ser His Phe Ser Asp Phe Met Asp Lys Gln Thr Gly Tyr Val Thr
165 170 175
Lys Asn Leu Leu Ala Thr Pro Ile Val Val Gly Lys Glu Val Leu Ala
180 185 190
Val Ile Met Ala Val Asn Lys Val Asn Ala Ser Glu Phe Ser Lys Gln
195 200 205
Asp Glu Glu Val Phe Ser Lys Tyr Leu Asn Phe Val Ser Ile Ile Leu
210 215 220
Arg Leu His His Thr Ser Tyr Met Tyr Asn Ile Glu Ser Arg Arg Ser
225 230 235 240
Gln Ile Leu Met Trp Ser Ala Asn Lys Val Phe Glu Glu Leu Thr Asp
245 250 255
Val Glu Arg Gln Phe His Lys Ala Leu Tyr Thr Val Arg Ser Tyr Leu
260 265 270
Asn Cys Glu Arg Tyr Ser Ile Gly Leu Leu Asp Met Thr Lys Glu Lys
275 280 285
Glu Phe Tyr Asp Glu Trp Pro Ile Lys Leu Gly Glu Val Glu Pro Tyr
290 295 300
Lys Gly Pro Lys Thr Pro Asp Gly Arg Glu Val Asn Phe Tyr Lys Ile
305 310 315 320
Ile Asp Tyr Ile Leu His Gly Lys Glu Glu Ile Lys Val Ile Pro Thr
325 330 335
Pro Pro Ala Asp His Trp Thr Leu Ile Ser Gly Leu Pro Thr Tyr Val
340 345 350
Ala Glu Asn Gly Phe Ile Cys Asn Met Met Asn Ala Pro Ala Asp Glu
355 360 365
Tyr Phe Thr Phe Gln Lys Gly Pro Val Asp Glu Thr Gly Trp Val Ile
370 375 380
Lys Asn Val Leu Ser Leu Pro Ile Val Asn Lys Lys Glu Asp Ile Val
385 390 395 400
Gly Val Ala Thr Phe Tyr Asn Arg Lys Asp Gly Lys Pro Phe Asp Glu
405 410 415
His Asp Glu Tyr Ile Thr Glu Thr Leu Thr Gln Phe Leu Gly Trp Ser
420 425 430
Leu Leu Asn Thr Asp Thr Tyr Asp Lys Met Asn Lys Leu Glu Asn Arg
435 440 445
Lys Asp Ile Ala Gln Glu Met Leu Met Asn Gln Thr Lys Ala Thr Pro
450 455 460

Glu Glu Ile Lys Ser Ile Leu Lys Phe Gln Glu Lys Leu Asn Val Asp
465 470 475 480
Val Ile Asp Asp Cys Glu Glu Lys Gln Leu Val Ala Ile Leu Lys Glu
485 490 495
Asp Leu Pro Asp Pro Arg Ser Ala Glu Leu Tyr Glu Phe Arg Phe Ser
500 505 510
Asp Phe Pro Leu Thr Glu His Gly Leu Ile Lys Cys Gly Ile Arg Leu
515 520 525
Phe Phe Glu Ile Asn Val Val Glu Lys Phe Lys Val Pro Val Glu Val
530 535 540
Leu Thr Arg Trp Met Tyr Thr Val Arg Lys Gly Tyr Arg Ala Val Thr
545 550 555 560
Tyr His Asn Trp Arg His Gly Phe Asn Val Gly Gln Thr Met Phe Thr
565 570 575
Leu Leu Met Thr Gly Arg Leu Lys Lys Tyr Tyr Thr Asp Leu Glu Ala
580 585 590
Phe Ala Met Leu Ala Ala Ala Phe Cys His Asp Ile Asp His Arg Gly
595 600 605
Thr Asn Asn Leu Tyr Gln Met Lys Ser Thr Ser Pro Leu Ala Arg Leu
610 615 620
His Gly Ser Ser Ile Leu Glu Arg His His Leu Glu Tyr Ser Lys Thr
625 630 635 640
Leu Leu Gln Asp Glu Ser Leu Asn Ile Phe Gln Asn Leu Asn Lys Arg
645 650 655
Gln Phe Glu Thr Val Ile His Leu Phe Glu Val Ala Ile Ile Ala Thr
660 665 670
Asp Leu Ala Leu Tyr Phe Lys Lys Arg Thr Met Phe Gln Lys Ile Val
675 680 685
Asp Ala Cys Glu Gln Met Gln Thr Glu Glu Glu Ala Ile Lys Tyr Val
690 695 700
Thr Val Asp Pro Thr Lys Lys Glu Ile Ile Met Ala Met Met Met Thr
705 710 715 720
Ala Cys Asp Leu Ser Ala Ile Thr Lys Pro Trp Glu Val Gln Ser Gln
725 730 735
Val Ala Leu Met Val Ala Asn Glu Phe Trp Glu Gln Gly Asp Leu Glu
740 745 750
Arg Thr Val Leu Gln Gln Gln Pro Ile Pro Met Met Asp Arg Asn Lys
755 760 765
Arg Asp Glu Leu Pro Lys Leu Gln Val Gly Phe Ile Asp Phe Val Cys

770 775 780
Thr Phe Val Tyr Lys Glu Phe Ser Arg Phe His Lys Glu Ile Thr Pro
785 790 795 800
Met Leu Ser Gly Leu Gln Asn Asn Arg Val Glu Trp Lys Ser Leu Ala
805 810 815
Asp Glu Tyr Asp Ala Lys Met Lys Val Ile Glu Glu Glu Ala Lys Lys
820 825 830
Gln Glu Gly Gly Ala Glu Lys Ala Ala Glu Asp Ser Gly Gly Gly Asp
835 840 845
Asp Lys Lys Ser Lys Thr Cys Leu Met Leu
850 855

<210> 51
<211> 83
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 51
Met Ser Asp Asn Thr Thr Leu Pro Ala Pro Ala Ser Asn Gln Gly Pro
1 5 10 15
Thr Thr Pro Arg Lys Gly Pro Pro Lys Phe Lys Gln Arg Gln Thr Arg
20 25 30
Gln Phe Lys Ser Lys Pro Pro Lys Lys Gly Val Lys Gly Phe Gly Asp
35 40 45
Asp Ile Pro Gly Met Glu Gly Leu Gly Thr Asp Ile Thr Val Ile Cys
50 55 60
Pro Trp Glu Ala Phe Ser His Leu Glu Leu His Glu Leu Ala Gln Phe
65 70 75 80
Gly Ile Ile

<210> 52
<211> 299
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 52
Met Met Ala Tyr Met Asn Pro Gly Pro His Tyr Ser Val Asn Ala Leu
1 5 10 15
Ala Leu Ser Gly Pro Ser Val Asp Leu Met His Gln Ala Val Pro Tyr

20 25 30

Pro Ser Ala Pro Arg Lys Gln Arg Arg Glu Arg Thr Thr Phe Thr Arg
35 40 45

Ser Gln Leu Glu Glu Leu Glu Ala Leu Phe Ala Lys Thr Gln Tyr Pro
50 55 60

Asp Val Tyr Ala Arg Glu Glu Val Ala Leu Lys Ile Asn Leu Pro Glu
65 70 75 80

Ser Arg Val Gln Val Trp Phe Lys Asn Arg Arg Ala Lys Cys Arg Gln
85 90 95

Gln Arg Gln Gln Gln Lys Gln Gln Gln Gln Pro Pro Gly Gly Gln Ala
100 105 110

Lys Ala Arg Pro Ala Lys Arg Lys Ala Gly Thr Ser Pro Arg Pro Ser
115 120 125

Thr Asp Val Cys Pro Asp Pro Leu Gly Ile Ser Asp Ser Tyr Ser Pro
130 135 140

Pro Leu Pro Gly Pro Ser Gly Ser Pro Thr Thr Ala Val Ala Thr Val
145 150 155 160

Ser Ile Trp Ser Pro Ala Ser Glu Ser Pro Leu Pro Glu Ala Gln Arg
165 170 175

Ala Gly Leu Val Ala Ser Gly Pro Ser Leu Thr Ser Ala Pro Tyr Ala
180 185 190

Met Thr Tyr Ala Pro Ala Ser Ala Phe Cys Ser Ser Pro Ser Ala Tyr
195 200 205

Gly Ser Pro Ser Ser Tyr Phe Ser Gly Leu Asp Pro Tyr Leu Ser Pro
210 215 220

Met Val Pro Gln Leu Gly Gly Pro Ala Leu Ser Pro Leu Ser Gly Pro
225 230 235 240

Ser Val Gly Pro Ser Leu Ala Gln Ser Pro Thr Ser Leu Ser Gly Gln
245 250 255

Ser Tyr Gly Ala Tyr Ser Pro Val Asp Ser Leu Glu Phe Lys Asp Pro
260 265 270

Thr Gly Thr Trp Lys Phe Thr Tyr Asn Pro Met Asp Pro Leu Asp Tyr
275 280 285

Lys Asp Gln Ser Ala Trp Lys Phe Gln Ile Leu
290 295

<210> 53

<211> 809

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 53

Met Phe Lys Ser Leu Thr Lys Val Asn Lys Val Lys Pro Ile Gly Glu
1 5 10 15
Asn Asn Glu Asn Glu Gln Ser Ser Arg Arg Asn Glu Glu Gly Ser His
20 25 30
Pro Ser Asn Gln Ser Gln Gln Thr Thr Ala Gln Glu Glu Asn Lys Gly
35 40 45
Glu Glu Lys Ser Leu Lys Thr Lys Ser Thr Pro Val Thr Ser Glu Glu
50 55 60
Pro His Thr Asn Ile Gln Asp Lys Leu Ser Lys Lys Asn Ser Ser Gly
65 70 75 80
Asp Leu Thr Thr Asn Pro Asp Pro Gln Asn Ala Ala Glu Pro Thr Gly
85 90 95
Thr Val Pro Glu Gln Lys Glu Met Asp Pro Gly Lys Glu Gly Pro Asn
100 105 110
Ser Pro Gln Asn Lys Pro Pro Ala Ala Pro Val Ile Asn Glu Tyr Ala
115 120 125
Asp Ala Gln Leu His Asn Leu Val Lys Arg Met Arg Gln Arg Thr Ala
130 135 140
Leu Tyr Lys Lys Lys Leu Val Glu Gly Asp Leu Ser Ser Pro Glu Ala
145 150 155 160
Ser Pro Gln Thr Ala Lys Pro Thr Ala Val Pro Pro Val Lys Glu Ser
165 170 175
Asp Asp Lys Pro Thr Glu His Tyr Tyr Arg Leu Leu Trp Phe Lys Val
180 185 190
Lys Lys Met Pro Leu Thr Glu Tyr Leu Lys Arg Ile Lys Leu Pro Asn
195 200 205
Ser Ile Asp Ser Tyr Thr Asp Arg Leu Tyr Leu Leu Trp Leu Leu Leu
210 215 220
Val Thr Leu Ala Tyr Asn Trp Asn Cys Trp Phe Ile Pro Leu Arg Leu
225 230 235 240
Val Phe Pro Tyr Gln Thr Ala Asp Asn Ile His Tyr Trp Leu Ile Ala
245 250 255
Asp Ile Ile Cys Asp Ile Ile Tyr Leu Tyr Asp Met Leu Phe Ile Gln
260 265 270
Pro Arg Leu Gln Phe Val Arg Gly Gly Asp Ile Ile Val Asp Ser Asn
275 280 285

Glu Leu Arg Lys His Tyr Arg Thr Ser Thr Lys Phe Gln Leu Asp Val
290 295 300
Ala Ser Ile Ile Pro Phe Asp Ile Cys Tyr Leu Phe Phe Gly Phe Asn
305 310 315 320
Pro Met Phe Arg Ala Asn Arg Met Leu Lys Tyr Thr Ser Phe Phe Glu
325 330 335
Phe Asn His His Leu Glu Ser Ile Met Asp Lys Ala Tyr Ile Tyr Arg
340 345 350
Val Ile Arg Thr Thr Gly Tyr Leu Leu Phe Ile Leu His Ile Asn Ala
355 360 365
Cys Val Tyr Tyr Trp Ala Ser Asn Tyr Glu Gly Ile Gly Thr Thr Arg
370 375 380
Trp Val Tyr Asp Gly Glu Gly Asn Glu Tyr Leu Arg Cys Tyr Tyr Trp
385 390 395 400
Ala Val Arg Thr Leu Ile Thr Ile Gly Gly Leu Pro Glu Pro Gln Thr
405 410 415
Leu Phe Glu Ile Val Phe Gln Leu Leu Asn Phe Phe Ser Gly Val Phe
420 425 430
Val Phe Ser Ser Leu Ile Gly Gln Met Arg Asp Val Ile Gly Ala Ala
435 440 445
Thr Ala Asn Gln Asn Tyr Phe Arg Ala Cys Met Asp Asp Thr Ile Ala
450 455 460
Tyr Met Asn Asn Tyr Ser Ile Pro Lys Leu Val Gln Lys Arg Val Arg
465 470 475 480
Thr Trp Tyr Glu Tyr Thr Trp Asp Ser Gln Arg Met Leu Asp Glu Ser
485 490 495
Asp Leu Leu Lys Thr Leu Pro Thr Thr Val Gln Leu Ala Leu Ala Ile
500 505 510
Asp Val Asn Phe Ser Ile Ile Ser Lys Val Asp Leu Phe Lys Gly Cys
515 520 525
Asp Thr Gln Met Ile Tyr Asp Met Leu Leu Arg Leu Lys Ser Val Leu
530 535 540
Tyr Leu Pro Gly Asp Phe Val Cys Lys Lys Gly Glu Ile Gly Lys Glu
545 550 555 560
Met Tyr Ile Ile Lys His Gly Glu Val Gln Val Leu Gly Gly Pro Asp
565 570 575
Gly Thr Lys Val Leu Val Thr Leu Lys Ala Gly Ser Val Phe Gly Glu
580 585 590
Ile Ser Leu Leu Ala Ala Gly Gly Gly Asn Arg Arg Thr Ala Asn Val

595 600 605
Val Ala His Gly Phe Ala Asn Leu Leu Thr Leu Asp Lys Lys Thr Leu
610 615 620
Gln Glu Ile Leu Val His Tyr Pro Asp Ser Glu Arg Ile Leu Met Lys
625 630 635 640
Lys Ala Arg Val Leu Leu Lys Gln Lys Ala Lys Thr Ala Glu Ala Thr
645 650 655
Pro Pro Arg Lys Asp Leu Ala Leu Leu Phe Pro Pro Lys Glu Glu Thr
660 665 670
Pro Lys Leu Phe Lys Thr Leu Leu Gly Gly Thr Gly Lys Ala Ser Leu
675 680 685
Ala Arg Leu Leu Lys Leu Lys Arg Glu Gln Ala Ala Gln Lys Lys Glu
690 695 700
Asn Ser Glu Gly Gly Glu Glu Glu Gly Lys Glu Asn Glu Asp Lys Gln
705 710 715 720
Lys Glu Asn Glu Asp Lys Gln Lys Glu Asn Glu Asp Lys Gly Lys Glu
725 730 735
Asn Glu Asp Lys Asp Lys Gly Arg Glu Pro Glu Glu Lys Pro Leu Asp
740 745 750
Arg Pro Glu Cys Thr Ala Ser Pro Ile Ala Val Glu Glu Glu Pro His
755 760 765
Ser Val Arg Arg Thr Val Leu Pro Arg Gly Thr Ser Arg Gln Ser Leu
770 775 780
Ile Ile Ser Met Ala Pro Ser Ala Glu Gly Gly Glu Glu Val Leu Thr
785 790 795 800
Ile Glu Val Lys Glu Lys Ala Lys Gln
805

<210> 54
<211> 610
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 54
Met Pro Leu Thr Glu Tyr Leu Lys Arg Ile Lys Leu Pro Asn Ser Ile
1 5 10 15
Asp Ser Tyr Thr Asp Arg Leu Tyr Leu Leu Trp Leu Leu Leu Val Thr
20 25 30
Leu Ala Tyr Asn Trp Asn Cys Cys Phe Ile Pro Leu Arg Leu Val Phe

35 40 45

Pro Tyr Gln Thr Ala Asp Asn Ile His Tyr Trp Leu Ile Ala Asp Ile
50 55 60

Ile Cys Asp Ile Ile Tyr Leu Tyr Asp Met Leu Phe Ile Gln Pro Arg
65 70 75 80

Leu Gln Phe Val Arg Gly Gly Asp Ile Ile Val Asp Ser Asn Glu Leu
85 90 95

Arg Lys His Tyr Arg Thr Ser Thr Lys Phe Gln Leu Asp Val Ala Ser
100 105 110

Ile Ile Pro Phe Asp Ile Cys Tyr Leu Phe Phe Gly Phe Asn Pro Met
115 120 125

Phe Arg Ala Asn Arg Met Leu Lys Tyr Thr Ser Phe Phe Glu Phe Asn
130 135 140

His His Leu Glu Ser Ile Met Asp Lys Ala Tyr Ile Tyr Arg Val Ile
145 150 155 160

Arg Thr Thr Gly Tyr Leu Leu Phe Ile Leu His Ile Asn Ala Cys Val
165 170 175

Tyr Tyr Trp Ala Ser Asn Tyr Glu Gly Ile Gly Thr Thr Arg Trp Val
180 185 190

Tyr Asp Gly Glu Gly Asn Glu Tyr Leu Arg Cys Tyr Tyr Trp Ala Val
195 200 205

Arg Thr Leu Ile Thr Ile Gly Gly Leu Pro Glu Pro Gln Thr Leu Phe
210 215 220

Glu Ile Val Phe Gln Leu Leu Asn Phe Phe Ser Gly Val Phe Val Phe
225 230 235 240

Ser Ser Leu Ile Gly Gln Met Arg Asp Val Ile Gly Ala Ala Thr Ala
245 250 255

Asn Gln Asn Tyr Phe Arg Ala Cys Met Asp Asp Thr Ile Ala Tyr Met
260 265 270

Asn Asn Tyr Ser Ile Pro Lys Leu Val Gln Lys Arg Val Arg Thr Trp
275 280 285

Tyr Glu Tyr Thr Trp Asp Ser Gln Arg Met Leu Asp Glu Ser Asp Leu
290 295 300

Leu Lys Thr Leu Pro Thr Thr Val Gln Leu Ala Leu Ala Ile Asp Val
305 310 315 320

Asn Phe Ser Ile Ile Ser Lys Val Asp Leu Phe Lys Gly Cys Asp Thr
325 330 335

Gln Met Ile Tyr Asp Met Leu Leu Arg Leu Lys Ser Val Leu Tyr Leu
340 345 350

Pro Gly Asp Phe Val Cys Lys Lys Gly Glu Ile Gly Lys Glu Met Tyr
355 360 365
Ile Ile Lys His Gly Glu Val Gln Val Leu Gly Gly Pro Asp Gly Thr
370 375 380
Lys Val Leu Val Thr Leu Lys Ala Gly Ser Val Leu Leu Ala Ala Gly
385 390 395 400
Gly Gly Asn Arg Arg Thr Ala Asn Val Val Ala His Gly Phe Ala Asn
405 410 415
Leu Leu Thr Leu Asp Lys Lys Thr Leu Gln Glu Ile Leu Val His Tyr
420 425 430
Pro Asp Ser Glu Arg Ile Leu Met Lys Lys Ala Arg Val Leu Leu Lys
435 440 445
Gln Lys Ala Lys Thr Ala Glu Ala Thr Pro Pro Arg Lys Asp Leu Ala
450 455 460
Leu Leu Phe Pro Pro Lys Glu Glu Thr Pro Lys Leu Phe Lys Thr Leu
465 470 475 480
Leu Gly Gly Thr Gly Lys Ala Ser Leu Ala Arg Leu Leu Lys Leu Lys
485 490 495
Arg Glu Gln Ala Ala Gln Lys Lys Glu Asn Ser Glu Gly Gly Glu Glu
500 505 510
Glu Gly Lys Glu Asn Glu Asp Lys Gln Lys Glu Asn Glu Asp Lys Gln
515 520 525
Lys Glu Asn Glu Asp Lys Gly Lys Glu Asn Glu Asp Lys Asp Lys Gly
530 535 540
Arg Glu Pro Glu Glu Lys Pro Leu Asp Arg Pro Glu Cys Thr Ala Ser
545 550 555 560
Pro Ile Ala Val Glu Glu Glu Pro His Ser Val Arg Arg Thr Val Leu
565 570 575
Pro Arg Gly Thr Ser Arg Gln Ser Leu Ile Ile Ser Met Ala Pro Ser
580 585 590
Ala Glu Gly Gly Glu Glu Val Leu Thr Ile Glu Val Lys Glu Lys Ala
595 600 605
Lys Gln
610

<210> 55
<211> 2156
<212> PRT
<213> 人工序列
<220>

<223> 合成的

<400> 55

Met Ser Asp Thr Pro Ser Thr Gly Phe Ser Ile Ile His Pro Thr Ser
1 5 10 15
Ser Glu Gly Gln Val Pro Pro Pro Arg His Leu Ser Leu Thr His Pro
20 25 30
Val Val Ala Lys Arg Ile Ser Phe Tyr Lys Ser Gly Asp Pro Gln Phe
35 40 45
Gly Gly Val Arg Val Val Val Asn Pro Arg Ser Phe Lys Ser Phe Asp
50 55 60
Ala Leu Leu Asp Asn Leu Ser Arg Lys Val Pro Leu Pro Phe Gly Val
65 70 75 80
Arg Asn Ile Ser Thr Pro Arg Gly Arg His Ser Ile Thr Arg Leu Glu
85 90 95
Glu Leu Glu Asp Gly Glu Ser Tyr Leu Cys Ser His Gly Arg Lys Val
100 105 110
Gln Pro Val Asp Leu Asp Lys Ala Arg Arg Arg Pro Arg Pro Trp Leu
115 120 125
Ser Ser Arg Ala Ile Ser Ala His Ser Pro Pro His Pro Val Ala Val
130 135 140
Ala Ala Pro Gly Met Pro Arg Pro Pro Arg Ser Leu Val Val Phe Arg
145 150 155 160
Asn Gly Asp Pro Lys Thr Arg Arg Ala Val Leu Leu Ser Arg Arg Val
165 170 175
Thr Gln Ser Phe Glu Ala Phe Leu Gln His Leu Thr Glu Val Met Gln
180 185 190
Arg Pro Val Val Lys Leu Tyr Ala Thr Asp Gly Arg Arg Val Pro Ser
195 200 205
Leu Gln Ala Val Ile Leu Ser Ser Gly Ala Val Val Ala Ala Gly Arg
210 215 220
Glu Pro Phe Lys Pro Gly Asn Tyr Asp Ile Gln Lys Tyr Leu Leu Pro
225 230 235 240
Ala Arg Leu Pro Gly Ile Ser Gln Arg Val Tyr Pro Lys Gly Asn Ala
245 250 255
Lys Ser Glu Ser Arg Lys Ile Ser Thr His Met Ser Ser Ser Ser Arg
260 265 270
Ser Gln Ile Tyr Ser Val Ser Ser Glu Lys Thr His Asn Asn Asp Cys
275 280 285
Tyr Leu Asp Tyr Ser Phe Val Pro Glu Lys Tyr Leu Ala Leu Glu Lys

290 295 300

Asn Asp Ser Gln Asn Leu Pro Ile Tyr Pro Ser Glu Asp Asp Ile Glu
305 310 315 320

Lys Ser Ile Ile Phe Asn Gln Asp Gly Thr Met Thr Val Glu Met Lys
325 330 335

Val Arg Phe Arg Ile Lys Glu Glu Glu Thr Ile Lys Trp Thr Thr Thr
340 345 350

Val Ser Lys Thr Gly Pro Ser Asn Asn Asp Glu Lys Ser Glu Met Ser
355 360 365

Phe Pro Gly Arg Thr Glu Ser Arg Ser Ser Gly Leu Lys Leu Ala Ala
370 375 380

Cys Ser Phe Ser Ala Asp Val Ser Pro Met Glu Arg Ser Ser Asn Gln
385 390 395 400

Glu Gly Ser Leu Ala Glu Glu Ile Asn Ile Gln Met Thr Asp Gln Val
405 410 415

Ala Glu Thr Cys Ser Ser Ala Ser Trp Glu Asn Ala Thr Val Asp Thr
420 425 430

Asp Ile Ile Gln Gly Thr Gln Asp Gln Ala Lys His Arg Phe Tyr Arg
435 440 445

Pro Pro Thr Pro Gly Leu Arg Arg Val Arg Gln Lys Lys Ser Val Ile
450 455 460

Gly Ser Val Thr Leu Val Ser Glu Thr Glu Val Gln Glu Lys Met Ile
465 470 475 480

Gly Gln Phe Ser Tyr Ser Glu Glu Arg Glu Ser Gly Glu Asn Lys Ser
485 490 495

Glu Tyr His Met Phe Thr His Ser Cys Ser Lys Met Ser Ser Val Ser
500 505 510

Asn Lys Pro Val Leu Val Gln Ile Asn Asn Asn Asp Gln Met Glu Glu
515 520 525

Ser Ser Leu Glu Arg Lys Lys Glu Asn Ser Leu Leu Lys Ser Ser Ala
530 535 540

Ile Ser Ala Gly Val Ile Glu Ile Thr Ser Gln Lys Met Leu Glu Met
545 550 555 560

Ser His Asn Asn Gly Leu Pro Ser Thr Ile Ser Asn Asn Ser Ile Val
565 570 575

Glu Glu Asp Val Val Asp Cys Val Val Leu Asp Asn Lys Thr Gly Ile
580 585 590

Lys Asn Phe Lys Thr Tyr Gly Asn Thr Asn Asp Arg Phe Ser Pro Ile
595 600 605

Ser Ala Asp Ala Thr His Phe Ser Ser Asn Asn Ser Gly Thr Asp Lys
610 615 620
Asn Ile Ser Glu Ala Pro Ala Ser Glu Ala Ser Ser Thr Val Thr Ala
625 630 635 640
Arg Ile Asp Arg Leu Ile Asn Glu Phe Ala Gln Cys Gly Leu Thr Lys
645 650 655
Leu Pro Lys Asn Glu Lys Lys Ile Leu Ser Ser Val Ala Ser Lys Lys
660 665 670
Lys Lys Lys Ser Arg Gln Gln Ala Ile Asn Ser Arg Tyr Gln Asp Gly
675 680 685
Gln Leu Ala Thr Lys Gly Ile Leu Asn Lys Asn Glu Arg Ile Asn Thr
690 695 700
Lys Gly Arg Ile Thr Lys Glu Met Ile Val Gln Asp Ser Asp Ser Pro
705 710 715 720
Leu Lys Gly Gly Ile Leu Cys Glu Glu Asp Leu Gln Lys Ser Asp Thr
725 730 735
Val Ile Glu Ser Asn Thr Phe Cys Ser Lys Ser Asn Leu Asn Ser Thr
740 745 750
Ile Ser Lys Asn Phe His Arg Asn Lys Leu Asn Thr Thr Gln Asn Ser
755 760 765
Lys Val Gln Gly Leu Leu Thr Lys Arg Lys Ser Arg Ser Leu Asn Lys
770 775 780
Ile Ser Leu Gly Ala Pro Lys Lys Arg Glu Ile Gly Gln Arg Asp Lys
785 790 795 800
Val Phe Pro His Asn Glu Ser Lys Tyr Cys Lys Ser Thr Phe Glu Asn
805 810 815
Lys Ser Leu Phe His Val Phe Asn Ile Leu Glu Gln Lys Pro Lys Asp
820 825 830
Phe Tyr Ala Pro Gln Ser Gln Ala Glu Val Ala Ser Gly Tyr Leu Arg
835 840 845
Gly Met Ala Lys Lys Ser Leu Val Ser Lys Val Thr Asp Ser His Ile
850 855 860
Thr Leu Lys Ser Gln Lys Lys Arg Lys Gly Asp Lys Val Lys Ala Ser
865 870 875 880
Ala Ile Leu Ser Lys Gln His Ala Thr Thr Arg Ala Asn Ser Leu Ala
885 890 895
Ser Leu Lys Lys Pro Asp Phe Pro Glu Ala Ile Ala His His Ser Ile
900 905 910
Gln Asn Tyr Ile Gln Ser Trp Leu Gln Asn Ile Asn Pro Tyr Pro Thr

915 920 925

Leu Lys Pro Ile Lys Ser Ala Pro Val Cys Arg Asn Glu Thr Ser Val

930 935 940

Val Asn Cys Ser Asn Asn Ser Phe Ser Gly Asn Asp Pro His Thr Asn

945 950 955 960

Ser Gly Lys Ile Ser Asn Phe Val Met Glu Ser Asn Lys His Ile Thr

965 970 975

Lys Ile Ala Gly Leu Thr Gly Asp Asn Leu Cys Lys Glu Gly Asp Lys

980 985 990

Ser Phe Ile Ala Asn Asp Thr Gly Glu Glu Asp Leu His Glu Thr Gln

995 1000 1005

Val Gly Ser Leu Asn Asp Ala Tyr Leu Val Pro Leu His Glu His

1010 1015 1020

Cys Thr Leu Ser Gln Ser Ala Ile Asn Asp His Asn Thr Lys Ser

1025 1030 1035

His Ile Ala Ala Glu Lys Ser Gly Pro Glu Lys Lys Leu Val Tyr

1040 1045 1050

Gln Glu Ile Asn Leu Ala Arg Lys Arg Gln Ser Val Glu Ala Ala

1055 1060 1065

Ile Gln Val Asp Pro Ile Glu Glu Glu Thr Pro Lys Asp Leu Leu

1070 1075 1080

Pro Val Leu Met Leu His Gln Leu Gln Ala Ser Val Pro Gly Ile

1085 1090 1095

His Lys Thr Gln Asn Gly Val Val Gln Met Pro Gly Ser Leu Ala

1100 1105 1110

Gly Val Pro Phe His Ser Ala Ile Cys Asn Ser Ser Thr Asn Leu

1115 1120 1125

Leu Leu Ala Trp Leu Leu Val Leu Asn Leu Lys Gly Ser Met Asn

1130 1135 1140

Ser Phe Cys Gln Val Asp Ala His Lys Ala Thr Asn Lys Ser Ser

1145 1150 1155

Glu Thr Leu Ala Leu Leu Glu Ile Leu Lys His Ile Ala Ile Thr

1160 1165 1170

Glu Glu Ala Asp Asp Leu Lys Ala Ala Val Ala Asn Leu Val Glu

1175 1180 1185

Ser Thr Thr Ser His Phe Gly Leu Ser Glu Lys Glu Gln Asp Met

1190 1195 1200

Val Pro Ile Asp Leu Ser Ala Asn Cys Ser Thr Val Asn Ile Gln

1205 1210 1215

Ser Val Pro Lys Cys Ser Glu Asn Glu Arg Thr Gln Gly Ile Ser
1220 1225 1230
Ser Leu Asp Gly Gly Cys Ser Ala Ser Glu Ala Cys Ala Pro Glu
1235 1240 1245
Val Cys Val Leu Glu Val Thr Cys Ser Pro Cys Glu Met Cys Thr
1250 1255 1260
Val Asn Lys Ala Tyr Ser Pro Lys Glu Thr Cys Asn Pro Ser Asp
1265 1270 1275
Thr Phe Phe Pro Ser Asp Gly Tyr Gly Val Asp Gln Thr Ser Met
1280 1285 1290
Asn Lys Ala Cys Phe Leu Gly Glu Val Cys Ser Leu Thr Asp Thr
1295 1300 1305
Val Phe Ser Asp Lys Ala Cys Ala Gln Lys Glu Asn His Thr Tyr
1310 1315 1320
Glu Gly Ala Cys Pro Ile Asp Glu Thr Tyr Val Pro Val Asn Val
1325 1330 1335
Cys Asn Thr Ile Asp Phe Leu Asn Ser Lys Glu Asn Thr Tyr Thr
1340 1345 1350
Asp Asn Leu Asp Ser Thr Glu Glu Leu Glu Arg Gly Asp Asp Ile
1355 1360 1365
Gln Lys Asp Leu Asn Ile Leu Thr Asp Pro Glu Tyr Lys Asn Gly
1370 1375 1380
Phe Asn Thr Leu Val Ser His Gln Asn Val Ser Asn Leu Ser Ser
1385 1390 1395
Cys Gly Leu Cys Leu Ser Glu Lys Glu Ala Glu Leu Asp Lys Lys
1400 1405 1410
His Ser Ser Leu Asp Asp Phe Glu Asn Cys Ser Leu Arg Lys Phe
1415 1420 1425
Gln Asp Glu Asn Ala Tyr Thr Ser Phe Asp Met Glu Glu Pro Arg
1430 1435 1440
Thr Ser Glu Glu Pro Gly Ser Ile Thr Asn Ser Met Thr Ser Ser
1445 1450 1455
Glu Arg Asn Ile Ser Glu Leu Glu Ser Phe Glu Glu Leu Glu Asn
1460 1465 1470
His Asp Thr Asp Ile Phe Asn Thr Val Val Asn Gly Gly Glu Gln
1475 1480 1485
Ala Thr Glu Glu Leu Ile Gln Glu Glu Val Glu Ala Ser Lys Thr
1490 1495 1500
Leu Glu Leu Ile Asp Ile Ser Ser Lys Asn Ile Met Glu Glu Lys

1505 1510 1515

Arg Met Asn Gly Ile Ile Tyr Glu Ile Ile Ser Lys Arg Leu Ala
1520 1525 1530

Thr Pro Pro Ser Leu Asp Phe Cys Tyr Asp Ser Lys Gln Asn Ser
1535 1540 1545

Glu Lys Glu Thr Asn Glu Gly Glu Thr Lys Met Val Lys Met Met
1550 1555 1560

Val Lys Thr Met Glu Thr Gly Ser Tyr Ser Glu Ser Ser Pro Asp
1565 1570 1575

Leu Lys Lys Cys Ile Lys Ser Pro Val Thr Ser Asp Trp Ser Asp
1580 1585 1590

Tyr Arg Pro Asp Ser Asp Ser Glu Gln Pro Tyr Lys Thr Ser Ser
1595 1600 1605

Asp Asp Pro Asn Asp Ser Gly Glu Leu Thr Gln Glu Lys Glu Tyr
1610 1615 1620

Asn Ile Gly Phe Val Lys Arg Ala Ile Glu Lys Leu Tyr Gly Lys
1625 1630 1635

Ala Asp Ile Ile Lys Pro Ser Phe Phe Pro Gly Ser Thr Arg Lys
1640 1645 1650

Ser Gln Val Cys Pro Tyr Asn Ser Val Glu Phe Gln Cys Ser Arg
1655 1660 1665

Lys Ala Ser Leu Tyr Asp Ser Glu Gly Gln Ser Phe Gly Ser Ser
1670 1675 1680

Glu Gln Val Ser Ser Ser Ser Ser Met Leu Gln Glu Phe Gln Glu
1685 1690 1695

Glu Arg Gln Asp Lys Cys Asp Val Ser Ala Val Arg Asp Asn Tyr
1700 1705 1710

Cys Arg Gly Asp Ile Val Glu Pro Gly Thr Lys Gln Asn Asp Asp
1715 1720 1725

Ser Arg Ile Leu Thr Asp Ile Glu Glu Gly Val Leu Ile Asp Lys
1730 1735 1740

Gly Lys Trp Leu Leu Lys Glu Asn His Leu Leu Arg Met Ser Ser
1745 1750 1755

Glu Asn Pro Gly Met Cys Gly Asn Ala Asp Thr Thr Ser Val Asp
1760 1765 1770

Thr Leu Leu Asp Asn Asn Ser Ser Glu Val Pro Tyr Ser His Phe
1775 1780 1785

Gly Asn Leu Ala Pro Gly Pro Thr Met Asp Glu Leu Ser Ser Ser
1790 1795 1800

Glu Leu Glu Glu Leu Thr Gln Pro Leu Glu Leu Lys Cys Asn Tyr
1805 1810 1815
Phe Asn Met Pro His Gly Ser Asp Ser Glu Pro Phe His Glu Asp
1820 1825 1830
Leu Leu Asp Val Arg Asn Glu Thr Cys Ala Lys Glu Arg Ile Ala
1835 1840 1845
Asn His His Thr Glu Glu Lys Gly Ser His Gln Ser Glu Arg Val
1850 1855 1860
Cys Thr Ser Val Thr His Ser Phe Ile Ser Ala Gly Asn Lys Val
1865 1870 1875
Tyr Pro Val Ser Asp Asp Ala Ile Lys Asn Gln Pro Leu Pro Gly
1880 1885 1890
Ser Asn Met Ile His Gly Thr Leu Gln Glu Ala Asp Ser Leu Asp
1895 1900 1905
Lys Leu Tyr Ala Leu Cys Gly Gln His Cys Pro Ile Leu Thr Val
1910 1915 1920
Ile Ile Gln Pro Met Asn Glu Glu Asp Arg Gly Phe Ala Tyr Arg
1925 1930 1935
Lys Glu Ser Asp Ile Glu Asn Phe Leu Gly Phe Tyr Leu Trp Met
1940 1945 1950
Lys Ile His Pro Tyr Leu Leu Gln Thr Asp Lys Asn Val Phe Arg
1955 1960 1965
Glu Glu Asn Asn Lys Ala Ser Met Arg Gln Asn Leu Ile Asp Asn
1970 1975 1980
Ala Ile Gly Asp Ile Phe Asp Gln Phe Tyr Phe Ser Asn Thr Phe
1985 1990 1995
Asp Leu Met Gly Lys Arg Arg Lys Gln Lys Arg Ile Asn Phe Leu
2000 2005 2010
Gly Leu Glu Glu Glu Gly Asn Leu Lys Lys Phe Gln Pro Asp Leu
2015 2020 2025
Lys Glu Arg Phe Cys Met Asn Phe Leu His Thr Ser Leu Leu Val
2030 2035 2040
Val Gly Asn Val Asp Ser Asn Thr Gln Asp Leu Ser Gly Gln Thr
2045 2050 2055
Asn Glu Ile Phe Lys Ala Val Asp Glu Asn Asn Asn Leu Leu Asn
2060 2065 2070
Asn Arg Phe Gln Gly Ser Arg Thr Asn Leu Asn Gln Val Val Arg
2075 2080 2085
Glu Asn Ile Asn Cys His Tyr Phe Phe Glu Met Leu Gly Gln Ala

2090 2095 2100

Cys Leu Leu Asp Ile Cys Gln Val Glu Thr Ser Leu Asn Ile Ser

2105 2110 2115

Asn Arg Asn Ile Leu Glu Leu Cys Met Phe Glu Gly Glu Asn Leu

2120 2125 2130

Phe Ile Trp Glu Glu Glu Asp Ile Leu Asn Leu Thr Asp Leu Glu

2135 2140 2145

Ser Ser Arg Glu Gln Glu Asp Leu

2150 2155

<210> 56

<211> 7100

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 56

gacatactga gaataaatcc aaagacatta gtttctttgc acgaaatgag gttacatatc 60

cagtgacatt tatttgagct atttaaacaa cttaaacatc tttttctttt cttaataagg 120

gacgtttcaa gttgtggtct cagccaaaat gagtgatacc ccttctactg gtttttccat 180

cattcatcct acgtcttctg aaggtcaagt tccacccccc cgccatttga gcctcactca 240

tcctgttgtg gccaagcgaa tcagtttcta caagagcgga gacccccaat tcggcggggt 300

cagggtggtg gtcaaccctc gctcctttaa gtcctttgat gctctgctgg ataacttgtc 360

caggaaggtg cccctccctt ttggagtgag gaacatcagc acccctcggg gcaggcacag 420

catcacgcgc ctggaggagc tggaggacgg cgagtcctac ctatgttccc acggcaggaa 480

ggtgcagcct gtagacctgg acaaagcccg tcggcgcccg cggccctggc tcagcagccg 540

ggccattagc gcgcactcac cgccccaccc cgtagccgtc gctgctcccg gcatgccccg 600

ccccccacgg agcctagtgg tcttcaggaa tggcgacccg aagacgaggc gtgcggttct 660

tctgagcagg agggtcaccc agagcttcga ggcatttcta cagcacctga cagaggtcat 720

gcagcgccct gtggtcaagc tgtacgctac ggacggaagg agggttccca gcctccaggc 780

agtgatcctg agctctggag ctgtggtggc ggcaggaagg gagccattta aaccaggaaa 840

ttatgacatc caaaaatact tgcttcctgc tagattacca gggatctctc agcgtgtgta 900

ccccaaggga aatgcaaagt cagaaagcag aaagataagc acacatatgt cttcaagctc 960

aaggtcccag atttattctg tttcttctga gaaaacacat aataatgatt gctacttaga 1020

ctattctttt gttcctgaaa agtacttggc cttagaaaag aatgattctc agaatttacc 1080

aatatatcct tctgaagatg atattgagaa atcaattatt tttaatcaag acggcactat 1140

gacagttgag atgaaagttc gattcagaat aaaagaggaa gaaaccataa aatggacaac 1200

tactgtcagt aaaactggtc cttctaataa tgatgaaaag agtgagatga gttttccagg 1260

aagaacagaa agtcgatcat ctggtttaaa gcttgcagca tgttcattct ctgcagatgt 1320

gtcacctatg gagcgaagca gtaatcaaga gggcagtttg gcagaggaga taaacattca 1380

aatgacagat caagtggctg aaacttgcag ttctgctagt tgggagaatg ctactgtgga 1440
cacagatatc atccagggaa ctcaagacca agcaaagcat cgtttttata ggccccctac 1500
acctggacta agaagagtga gacaaaagaa atctgtgatt ggcagtgtga ccttagtatc 1560
tgaaactgag gttcaagaga aaatgattgg acagttttca tatagtgaag aaagggaaag 1620
tggggaaaac aagtctgagt atcacatgtt tacacattct tgcagtaaaa tgtcatcagt 1680
atctaacaaa ccagtacttg ttcagatcaa taacaatgat caaatggagg agtcatcatt 1740
agaaagaaaa aaggaaaaca gtctgcttaa gtcaagtgca ataagtgctg gtgttataga 1800
aattacaagt cagaagatgt tagagatgtc acataataat ggtttgccat caactatatc 1860
aaataactca attgtggagg aagatgtagt tgattgtgtg gtattggaca acaaaactgg 1920
tatcaagaac ttcaaaactt atggtaacac caatgatagg ttcagtccta tttcagcaga 1980
tgcaacccat ttttcaagta ataactctgg aactgacaaa aatatttctg aggctccagc 2040
ttcagaagca tcctctactg tcactgcaag aattgacaga ctaattaatg aatttgctca 2100
gtgtggttta acaaaacttc caaaaaatga aaagaagatt ttgtcatctg ttgccagcaa 2160
aaagaagaaa aaatctcgac agcaagcaat aaattccagg tatcaagatg gacagcttgc 2220
aaccaaagga attcttaata agaatgagag aataaacaca aaaggtagaa ttacaaagga 2280
aatgatagtg caagattcag atagtcccct taaaggaggg atactttgtg aggaagacct 2340
ccagaaaagt gatactgtaa ttgaatcaaa tacttttttgt tccaaaagta atctcaattc 2400
cacgatttcc aagaatttcc atagaaataa attaaatact actcaaaatt ccaaggttca 2460
aggactttta accaaaagaa aatctagatc actaaataaa ataagcttag gagcacctaa 2520
aaaaagagaa atcggtcaaa gagataaagt gtttcctcac aatgaatcta aatattgcaa 2580
aagtactttt gaaaacaaaa gtttatttca tgtatttaac atccttgagc aaaaacccaa 2640
agatttttat gcaccgcaat ctcaagcaga agtggcatct gggtatttga gaggaatggc 2700
aaagaagagt ttagtttcaa aagttactga ttcacacata actttaaaaa gccagaaaaa 2760
acgtaaaggg gataaagtga aagcaagtgc tattttaagt aaacaacatg ctacaaccag 2820
ggcaaattct ttagcttctt tgaaaaaacc tgattttcct gaggctattg ctcatcattc 2880
aattcaaaat tatatacaga gttggttgca gaacataaat ccatatccaa ctttaaagcc 2940
tataaaatca gctccagtat gtagaaatga aacgagtgtg gtaaattgta gcaataatag 3000
tttttcaggg aatgatcccc atacaaattc tggaaaaata agtaattttg ttatggaaag 3060
taataagcac ataactaaaa ttgccggttt gacaggagat aatctatgta aagagggaga 3120
taagtctttt attgccaatg acactggtga agaagatctc catgagacac aggttggatc 3180
tctgaatgat gcttatttgg ttcccctgca tgaacactgt actttgtcac agtcagctat 3240
taatgatcat aatactaaaa gtcatatagc tgctgaaaaa tcaggaccag agaaaaaact 3300
tgtttaccag gaaataaacc tagctagaaa aaggcaaagt gtagaggctg ccattcaagt 3360
agatcctata gaagaggaaa ctccaaaaga cctcttacca gtcctgatgc ttcaccaatt 3420
gcaagcttca gttcctggta ttcacaagac tcagaatgga gttgttcaaa tgccaggttc 3480
acttgcaggt gttccctttc attctgcaat atgtaattca tccactaatc tccttctagc 3540
ttggctcttg gtgctaaacc taaagggaag tatgaatagc ttctgtcaag ttgatgctca 3600
caaggctacc aacaaatctt cagaaacact tgcattgttg gagattctaa agcacatagc 3660
tatcacagag gaagctgatg acttgaaagc tgctgttgcc aatttagtgg agtcaactac 3720

aagccacttt ggactcagtg agaaagaaca agacatggtt ccaatagatc tttctgcaaa 3780
ttgttccacg gtcaacattc agagtgttcc taagtgcagt gaaaatgaaa gaacacaagg 3840
aatctcctct ttggatggag gttgctctgc cagtgaggca tgtgcccctg aagtctgtgt 3900
tttggaagtg acttgctctc catgtgagat gtgcactgta aataaggctt attctccaaa 3960
agagacatgt aaccccagtg acacttttt tcctagtgat ggttatggtg tggatcagac 4020
ttctatgaat aaggcttgtt tcctaggaga ggtctgttca cttactgata ctgtgttttc 4080
tgataaggct tgtgctcaaa aggagaacca tacctatgag ggagcttgcc caattgatga 4140
gacctacgtt cctgtcaatg tctgcaatac cattgacttt ttaaactcca aagaaaacac 4200
atatactgat aacttggatt caactgaaga gttagaaaga ggtgatgaca ttcagaaaga 4260
tctaaatatt ttgacagacc ctgaatataa aaatggattt aatacattgg tgtcacatca 4320
aaatgtcagt aatttaagct cctgtggcct ttgcctaagt gaaaaagaag cagaacttga 4380
taagaaacat agttctctag atgattttga aaattgttca ctaaggaagt ttcaggatga 4440
aaatgcatat acttcctttg atatggaaga accacggact tctgaagaac caggctcaat 4500
aaccaacagc atgacatcaa gtgaaagaaa catttcagaa ttggaatctt ttgaagaatt 4560
agaaaaccat gacactgata tctttaatac agtggtaaat ggaggagagc aagccactga 4620
agaattaatc caagaagagg tagaggctag taaaacttta gaattgatag acatctctag 4680
taagaatatt atggaagaaa aaagaatgaa cggtataatt tatgaaataa tcagtaagag 4740
gctggcaaca ccaccatctt tagatttttg ctatgattct aagcaaaata gtgaaaagga 4800
gaccaatgaa ggagaaacta agatggtaaa aatgatggtg aaaactatgg aaactggaag 4860
ttattcagag tcctctcctg atttaaaaaa atgcatcaaa agtccagtga cttctgattg 4920
gtcagactat cggcctgaca gtgacagtga gcagccatat aaaacatcca gtgatgatcc 4980
caatgacagt ggcgaactta cccaagagaa agaatataac ataggatttg ttaaaagggc 5040
aatagaaaaa ctgtacggta aagcagatat tatcaaacca tctttttttc ctgggtctac 5100
ccgcaaatct caggtttgtc cttataattc tgtggaattt cagtgttcca ggaaagcaag 5160
tctttatgat tctgaagggc agtcatttgg ctcttctgaa caggtatcta gtagttcatc 5220
tatgttgcag gaattccagg aggaaagaca agataagtgt gatgttagtg ctgtgaggga 5280
caattattgt aggggtgaca ttgtagaacc tggtacaaaa caaaatgatg atagcagaat 5340
cctcacagac atagaggaag gagtactgat tgacaaaggc aaatggcttc tgaaagaaaa 5400
tcatttgcta aggatgtcat ctgaaaatcc tggcatgtgt ggcaatgcag acaccacatc 5460
agtggacacc ctacttgata ataacagcag tgaggtacca tattcacatt ttggtaattt 5520
ggccccaggc ccaacgatgg atgaactctc ctcttcagaa ctcgaggaac tgactcaacc 5580
ccttgaacta aaatgcaatt actttaacat gcctcatggt agtgactcag aacctttca 5640
tgaggacttg ctggatgttc gcaatgaaac ctgtgccaag gaaagaatag caaatcatca 5700
tacagaggag aagggtagtc atcagtcaga aagagtatgc acatctgtca ctcattcctt 5760
tatttctgct ggtaacaaag tctaccctgt ctctgatgat gctattaaaa accaaccatt 5820
gcctggcagt aatatgattc atggtacact tcaggaagct gactctttgg ataaactgta 5880
tgctctttgt ggtcaacatt gcccaatact aactgttatt atccaaccca tgaatgagga 5940
agaccgagga tttgcatatc gcaaagaatc tgatattgaa aatttcttgg gtttttattt 6000
atggatgaaa atacacccat atttacttca gacagacaaa aatgtgttca gggaagagaa 6060

caataaagca agtatgagac aaaatcttat tgataatgcc attggtgata tatttgatca 6120
gttttatttc agtaacacat ttgacttgat gggtaaaaga agaaaacaaa aaagaattaa 6180
cttcttgggg ttagaggaag aaggtaattt aaagaaattt caaccagatt tgaaggaaag 6240
gttttgtatg aatttcttgc acacatcatt gttagttgtg ggtaatgtgg attcaaatac 6300
acaagacctc agcggtcaga caaatgaaat ctttaaagca gtcgatgaga ataacaactt 6360
attaaataac agattccagg gctcaagaac aaatctcaac caagtagtaa gagaaaatat 6420
caactgtcat tacttctttg aaatgcttgg tcaagcttgc ctcttagata tttgccaagt 6480
tgagacctcc ttaaatatta gcaacagaaa tattttagaa ctttgtatgt ttgagggtga 6540
aaatcttttc atttgggaag aggaagacat attaaattta actgatcttg aaagcagtag 6600
agaacaagaa gatttataat ttcaatatca gcacactcat tctttgtcaa ttcatttttt 6660
cccatgagat gaagcacatg tgacgaatac ggactagata acctctaaga attttccact 6720
tcttcaaaat gaacttactc tagaaagctt acccttggat aaccagtttg actttcataa 6780
tgtctctgtt ttttgttttt ccaacaatta cagactcagg ttctcttatt ttggaagttt 6840
ctatctggtt ttgttctgaa cttacatttt ttttttttt ggtatctatg attttttttg 6900
ctcagggcat caaaatgtgc taaggacaag aattatatcc tttttaaaaa atgttgttag 6960
cttggtgtaa aatgtatatt gactgtattg gtgaataaat tgaatagaca taacctcaaa 7020
gtacttcact tattcttttt aactactgat ttgataaaaa gtatgattat aagatatcca 7080
cgacaatctc atagtttctt 7100

<210> 57
<211> 1286
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 57

Met Ser His Leu Val Asp Pro Thr Ser Gly Asp Leu Pro Val Arg Asp
1 5 10 15
Ile Asp Ala Ile Pro Leu Val Leu Pro Ala Ser Lys Gly Lys Asn Met
20 25 30
Lys Thr Gln Pro Pro Leu Ser Arg Met Asn Arg Glu Glu Leu Glu Asp
35 40 45
Ser Phe Phe Arg Leu Arg Glu Asp His Met Leu Val Lys Glu Leu Ser
50 55 60
Trp Lys Gln Gln Asp Glu Ile Lys Arg Leu Arg Thr Thr Leu Leu Arg
65 70 75 80
Leu Thr Ala Ala Gly Arg Asp Leu Arg Val Ala Glu Glu Ala Ala Pro
85 90 95
Leu Ser Glu Thr Ala Arg Arg Gly Gln Lys Ala Gly Trp Arg Gln Arg
100 105 110

Leu Ser Met His Gln Arg Pro Gln Met His Arg Leu Gln Gly His Phe
115 120 125
His Cys Val Gly Pro Ala Ser Pro Arg Arg Ala Gln Pro Arg Val Gln
130 135 140
Val Gly His Arg Gln Leu His Thr Ala Gly Ala Pro Val Pro Glu Lys
145 150 155 160
Pro Lys Arg Gly Pro Arg Asp Arg Leu Ser Tyr Thr Ala Pro Pro Ser
165 170 175
Phe Lys Glu His Ala Thr Asn Glu Asn Arg Gly Glu Val Ala Ser Lys
180 185 190
Pro Ser Glu Leu Val Ser Gly Ser Asn Ser Ile Ile Ser Phe Ser Ser
195 200 205
Val Ile Ser Met Ala Lys Pro Ile Gly Leu Cys Met Pro Asn Ser Ala
210 215 220
His Ile Met Ala Ser Asn Thr Met Gln Val Glu Glu Pro Pro Lys Ser
225 230 235 240
Pro Glu Lys Met Trp Pro Lys Asp Glu Asn Phe Glu Gln Arg Ser Ser
245 250 255
Leu Glu Cys Ala Gln Lys Ala Ala Glu Leu Arg Ala Ser Ile Lys Glu
260 265 270
Lys Val Glu Leu Ile Arg Leu Lys Lys Leu Leu His Glu Arg Asn Ala
275 280 285
Ser Leu Val Met Thr Lys Ala Gln Leu Thr Glu Val Gln Glu Ala Tyr
290 295 300
Glu Thr Leu Leu Gln Lys Asn Gln Gly Ile Leu Ser Ala Ala His Glu
305 310 315 320
Ala Leu Leu Lys Gln Val Asn Glu Leu Arg Ala Glu Leu Lys Glu Glu
325 330 335
Ser Lys Lys Ala Val Ser Leu Lys Ser Gln Leu Glu Asp Val Ser Ile
340 345 350
Leu Gln Met Thr Leu Lys Glu Phe Gln Glu Arg Val Glu Asp Leu Glu
355 360 365
Lys Glu Arg Lys Leu Leu Asn Asp Asn Tyr Asp Lys Leu Leu Glu Ser
370 375 380
Met Leu Asp Ser Ser Asp Ser Ser Ser Gln Pro His Trp Ser Asn Glu
385 390 395 400
Leu Ile Ala Glu Gln Leu Gln Gln Gln Val Ser Gln Leu Gln Asp Gln
405 410 415
Leu Asp Ala Glu Leu Glu Asp Lys Arg Lys Val Leu Leu Glu Leu Ser

420 425 430

Arg Glu Lys Ala Gln Asn Glu Asp Leu Lys Leu Glu Val Thr Asn Ile

435 440 445

Leu Gln Lys His Lys Gln Glu Val Glu Leu Leu Gln Asn Ala Ala Thr

450 455 460

Ile Ser Gln Pro Pro Asp Arg Gln Ser Glu Pro Ala Thr His Pro Ala

465 470 475 480

Val Leu Gln Glu Asn Thr Gln Ile Glu Pro Ser Glu Pro Lys Asn Gln

485 490 495

Glu Glu Lys Lys Leu Ser Gln Val Leu Asn Glu Leu Gln Val Ser His

500 505 510

Ala Glu Thr Thr Leu Glu Leu Glu Lys Thr Arg Asp Met Leu Ile Leu

515 520 525

Gln Arg Lys Ile Asn Val Cys Tyr Gln Glu Glu Leu Glu Ala Met Met

530 535 540

Thr Lys Ala Asp Asn Asp Asn Arg Asp His Lys Glu Lys Leu Glu Arg

545 550 555 560

Leu Thr Arg Leu Leu Asp Leu Lys Asn Asn Arg Ile Lys Gln Leu Glu

565 570 575

Gly Ile Leu Arg Ser His Asp Leu Pro Thr Ser Glu Gln Leu Lys Asp

580 585 590

Val Ala Tyr Gly Thr Arg Pro Leu Ser Leu Cys Leu Glu Thr Leu Pro

595 600 605

Ala His Gly Asp Glu Asp Lys Val Asp Ile Ser Leu Leu His Gln Gly

610 615 620

Glu Asn Leu Phe Glu Leu His Ile His Gln Ala Phe Leu Thr Ser Ala

625 630 635 640

Ala Leu Ala Gln Ala Gly Asp Thr Gln Pro Thr Thr Phe Cys Thr Tyr

645 650 655

Ser Phe Tyr Asp Phe Glu Thr His Cys Thr Pro Leu Ser Val Gly Pro

660 665 670

Gln Pro Leu Tyr Asp Phe Thr Ser Gln Tyr Val Met Glu Thr Asp Ser

675 680 685

Leu Phe Leu His Tyr Leu Gln Glu Ala Ser Ala Arg Leu Asp Ile His

690 695 700

Gln Ala Met Ala Ser Glu His Ser Thr Leu Ala Ala Gly Trp Ile Cys

705 710 715 720

Phe Asp Arg Val Leu Glu Thr Val Glu Lys Val His Gly Leu Ala Thr

725 730 735

Leu Ile Gly Ala Gly Gly Glu Glu Phe Gly Val Leu Glu Tyr Trp Met
740 745 750
Arg Leu Arg Phe Pro Ile Lys Pro Ser Leu Gln Ala Cys Asn Lys Arg
755 760 765
Lys Lys Ala Gln Val Tyr Leu Ser Thr Asp Val Leu Gly Gly Arg Lys
770 775 780
Ala Gln Glu Glu Glu Phe Arg Ser Glu Ser Trp Glu Pro Gln Asn Glu
785 790 795 800
Leu Trp Ile Glu Ile Thr Lys Cys Cys Gly Leu Arg Ser Arg Trp Leu
805 810 815
Gly Thr Gln Pro Ser Pro Tyr Ala Val Tyr Arg Phe Phe Thr Phe Ser
820 825 830
Asp His Asp Thr Ala Ile Ile Pro Ala Ser Asn Asn Pro Tyr Phe Arg
835 840 845
Asp Gln Ala Arg Phe Pro Val Leu Val Thr Ser Asp Leu Asp His Tyr
850 855 860
Leu Arg Arg Glu Ala Leu Ser Ile His Val Phe Asp Asp Glu Asp Leu
865 870 875 880
Glu Pro Gly Ser Tyr Leu Gly Arg Ala Arg Val Pro Leu Leu Pro Leu
885 890 895
Ala Lys Asn Glu Ser Ile Lys Gly Asp Phe Asn Leu Thr Asp Pro Ala
900 905 910
Glu Lys Pro Asn Gly Ser Ile Gln Val Gln Leu Asp Trp Lys Phe Pro
915 920 925
Tyr Ile Pro Pro Glu Ser Phe Leu Lys Pro Glu Ala Gln Thr Lys Gly
930 935 940
Lys Asp Thr Lys Asp Ser Ser Lys Ile Ser Ser Glu Glu Glu Lys Ala
945 950 955 960
Ser Phe Pro Ser Gln Asp Gln Met Ala Ser Pro Glu Val Pro Ile Glu
965 970 975
Ala Gly Gln Tyr Arg Ser Lys Arg Lys Pro Pro His Gly Gly Glu Arg
980 985 990
Lys Glu Lys Glu His Gln Val Val Ser Tyr Ser Arg Arg Lys His Gly
995 1000 1005
Lys Arg Ile Gly Val Gln Gly Lys Asn Arg Met Glu Tyr Leu Ser
1010 1015 1020
Leu Asn Ile Leu Asn Gly Asn Thr Pro Glu Gln Val Asn Tyr Thr
1025 1030 1035
Glu Trp Lys Phe Ser Glu Thr Asn Ser Phe Ile Gly Asp Gly Phe

1040 1045 1050

Lys Asn Gln His Glu Glu Glu Glu Met Thr Leu Ser His Ser Ala

1055 1060 1065

Leu Lys Gln Lys Glu Pro Leu His Pro Val Asn Asp Lys Glu Ser

1070 1075 1080

Ser Glu Gln Gly Ser Glu Val Ser Glu Ala Gln Thr Thr Asp Ser

1085 1090 1095

Asp Asp Val Ile Val Pro Pro Met Ser Gln Lys Tyr Pro Lys Ala

1100 1105 1110

Asp Ser Glu Lys Met Cys Ile Glu Ile Val Ser Leu Ala Phe Tyr

1115 1120 1125

Pro Glu Ala Glu Val Met Ser Asp Glu Asn Ile Lys Gln Val Tyr

1130 1135 1140

Val Glu Tyr Lys Phe Tyr Asp Leu Pro Leu Ser Glu Thr Glu Thr

1145 1150 1155

Pro Val Ser Leu Arg Lys Pro Arg Ala Gly Glu Glu Ile His Phe

1160 1165 1170

His Phe Ser Lys Val Ile Asp Leu Asp Pro Gln Glu Gln Gln Gly

1175 1180 1185

Arg Arg Arg Phe Leu Phe Asp Met Leu Asn Gly Gln Asp Pro Asp

1190 1195 1200

Gln Gly His Leu Lys Phe Thr Val Val Ser Asp Pro Leu Asp Glu

1205 1210 1215

Glu Lys Lys Glu Cys Glu Glu Val Gly Tyr Ala Tyr Leu Gln Leu

1220 1225 1230

Trp Gln Ile Leu Glu Ser Gly Arg Asp Ile Leu Glu Gln Glu Leu

1235 1240 1245

Asp Ile Val Ser Pro Glu Asp Leu Ala Thr Pro Ile Gly Arg Leu

1250 1255 1260

Lys Val Ser Leu Gln Ala Ala Ala Val Leu His Ala Ile Tyr Lys

1265 1270 1275

Glu Met Thr Glu Asp Leu Phe Ser

1280 1285

<210> 58

<211> 3950

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 58

atgtcacatc tggtggaccc tacatcagga gacttgccag ttagagacat agatgctata 60
cctctggtgc taccagcctc aaaaggtaag aatatgaaaa ctcaaccacc cttgagcagg 120
atgaaccggg aggaattgga ggacagtttc tttcgacttc gcgaagatca catgttggtg 180
aaggagcttt cttggaagca acaggatgag atcaaaaggc tgaggaccac cttgctgcgg 240
ttgaccgctg ctggccggga cctgcgggtc gcggaggagg cggcgccgct ctcggagacc 300
gcaaggcgcg ggcagaaggc gggatggcgg cagcgcctct ccatgcacca gcgcccccag 360
atgcaccgac tgcaagggca tttccactgc gtcggccctg ccagcccccg ccgcgcccag 420
cctcgcgtcc aagtgggaca cagacagctc cacacagccg gtgcaccggt gccggagaaa 480
cccaagaggg ggccaaggga caggctgagc tacacagccc ctccatcgtt taaggagcat 540
gcgacaaatg aaaacagagg tgaagtagcc agtaaaccca gtgaacttgt ttctggttct 600
aacagcataa tttctttcag cagtgtcata agtatggcta aacccattgg tctatgcatg 660
cctaacagtg cccacatcat ggccagcaat accatgcaag tggaagagcc acccaagtct 720
cctgagaaaa tgtggcctaa agatgaaaat tttgaacaga gaagctcatt ggagtgtgct 780
cagaaggctg cagagcttcg agcttccatt aaagagaagg tagagctgat tcgacttaag 840
aagctcttac atgaaagaaa tgcttcattg gttatgacaa aagcacaatt aacagaagtt 900
caagaggcat acgaaacctt gctccagaag aatcagggaa tcctgagtgc agcccatgag 960
gccctcctca agcaagtgaa tgagctcagg gcagagctga aggaagaaag caagaaggct 1020
gtgagcttga agagccaact ggaagatgtg tctatcttgc agatgactct gaaggagttt 1080
caggagagag ttgaagattt ggaaaaagaa cgaaaattgc tgaatgacaa ttatgacaaa 1140
ctcttagaaa gcatgctgga cagcagtgac agctccagtc agccccactg gagcaacgag 1200
ctcatagcgg aacagctaca gcagcaagtc tctcagctgc aggatcagct ggatgctgag 1260
ctggaggaca agagaaaagt tttacttgag ctgtccaggg agaaagccca aaatgaggat 1320
ctgaagcttg aagtcaccaa catacttcag aagcataaac aggaagtaga gctcctccaa 1380
aatgcagcca caatttccca acctcctgac aggcaatctg aaccagccac tcacccagct 1440
gtattgcaag agaacactca gatcgagcca agtgaaccca aaaaccaaga agaaaagaaa 1500
ctgtcccagg tgctaaatga gttgcaagta tcacacgcag agaccacatt ggaactagaa 1560
aagaccaggg acatgcttat tctgcagcgc aaaatcaacg tgtgttatca ggaggaactg 1620
gaggcaatga tgacaaaagc tgacaatgat aatagagatc acaaagaaaa gctggagagg 1680
ttgactcgac tactagacct caagaataac cgtatcaagc agctggaagg tattttaaga 1740
agccatgacc ttccaacatc tgaacagctc aaagatgttg cttatggcac ccgaccgttg 1800
tcgttatgtt tggaaacact gccagcccat ggagatgagg ataaagtgga tatttctctg 1860
ctgcatcagg gtgagaatct ttttgaactg cacatccacc aggccttcct gacatctgcc 1920
gccctagctc aggctggaga tacccaacct accactttct gcacctattc cttctatgac 1980
tttgaaaccc actgtacccc attatctgtg gggccacagc ccctctatga cttcacctcc 2040
cagtatgtga tggagacaga ttcgcttttc ttacactacc ttcaagaggc ttcagcccgg 2100
cttgacatac accaggccat ggccagtgaa cacagcactc ttgctgcagg atggatttgc 2160
tttgacaggg tgctagagac tgtggagaaa gtccatggct tggccacact gattggagct 2220
ggtggagaag agttcggggt tctagagtac tggatgaggc tgcgtttccc cataaaaccc 2280

agcctacagg cgtgcaataa acgaaagaaa gcccaggtct acctgtcaac cgatgtgctt 2340
ggaggccgga aggcccagga agaggagttc agatcggagt cttgggaacc tcagaacgag 2400
ctgtggattg aaatcaccaa gtgctgtggc ctccggagtc gatggctggg aactcaaccc 2460
agtccatatg ctgtgtaccg cttcttcacc ttttctgacc atgacactgc catcattcca 2520
gccagtaaca acccctactt tagagaccag gctcgattcc cagtgcttgt gacctctgac 2580
ctggaccatt atctgagacg ggaggccttg tctatacatg tttttgatga tgaagactta 2640
gagcctggct cgtatcttgg ccgagcccga gtgcctttac tgcctcttgc aaaaaatgaa 2700
tctatcaaag gtgattttaa cctcactgac cctgcagaga aacccaacgg atctattcaa 2760
gtgcaactgg attggaagtt tccctacata ccccctgaga gcttcctgaa accagaagct 2820
cagactaagg ggaaggatac caaggacagt tcaaagatct catctgaaga ggaaaaggct 2880
tcatttcctt cccaggatca gatggcatct cctgaggttc ccattgaagc tggccagtat 2940
cgatctaaga gaaaacctcc tcatggggga gaaagaaagg agaaggagca ccaggttgtg 3000
agctactcaa gaagaaaaca tggcaaaaga ataggtgttc aaggaaagaa tagaatggag 3060
tatcttagcc ttaacatctt aaatggaaat acaccagagc aggtgaatta cactgagtgg 3120
aagttctcag agactaacag cttcataggt gatggcttta aaaatcagca cgaggaagag 3180
gaaatgacat tatcccattc agcactgaaa cagaaggaac ctctacatcc tgtaaatgac 3240
aaagaatcct ctgaacaagg ttctgaagtc agtgaagcac aaactaccga cagtgatgat 3300
gtcatagtgc cacccatgtc tcagaaatat cctaaggcag attcagagaa gatgtgcatt 3360
gaaattgtct ccctggcctt ctacccagag gcagaagtga tgtctgatga gaacataaaa 3420
caggtgtatg tggagtacaa attctacgac ctacccttgt cggagacaga gactccagtg 3480
tccctaagga agcctagggc aggagaagaa atccactttc actttagcaa ggtaatagac 3540
ctggacccac aggagcagca aggccgaagg cggtttctgt tcgacatgct gaatggacaa 3600
gatcctgatc aaggacattt aaagtttaca gtggtaagtg atcctctgga tgaagaaaag 3660
aaagaatgtg aagaagtggg atatgcatat cttcaactgt ggcagatcct ggagtcagga 3720
agagatattc tagagcaaga gctagacatt gttagccctg aagatctggc taccccaata 3780
ggaaggctga aggtttccct tcaagcagct gctgtcctcc atgctattta caaggagatg 3840
actgaagatt tgttttcatg aaggaacaag tgctattcca atctaaaagt ctctgaggga 3900
accatagtaa aaagtctctt ataaagttag cttgctataa catgaaaaaa 3950

<210> 59
<211> 2335
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 59

Met Ala Gly Val Phe Pro Tyr Arg Gly Pro Gly Asn Pro Val Pro Gly
1 5 10 15
Pro Leu Ala Pro Leu Pro Asp Tyr Met Ser Glu Glu Lys Leu Gln Glu
20 25 30

Lys Ala Arg Lys Trp Gln Gln Leu Gln Ala Lys Arg Tyr Ala Glu Lys
35 40 45
Arg Lys Phe Gly Phe Val Asp Ala Gln Lys Glu Asp Met Pro Pro Glu
50 55 60
His Val Arg Glu Ile Ile Arg Asp His Gly Asp Met Thr Asn Arg Lys
65 70 75 80
Phe Arg His Asp Lys Arg Val Tyr Leu Gly Ala Leu Lys Tyr Met Pro
85 90 95
His Ala Val Leu Lys Leu Leu Glu Asn Met Pro Met Pro Trp Glu Gln
100 105 110
Ile Arg Asp Val Pro Val Leu Tyr His Ile Thr Gly Ala Ile Ser Phe
115 120 125
Val Asn Glu Ile Pro Trp Val Ile Glu Pro Val Tyr Ile Ser Gln Trp
130 135 140
Gly Ser Met Trp Ile Met Met Arg Arg Glu Lys Arg Asp Arg Arg His
145 150 155 160
Phe Lys Arg Met Arg Phe Pro Pro Phe Asp Asp Glu Glu Pro Pro Leu
165 170 175
Asp Tyr Ala Asp Asn Ile Leu Asn Val Glu Pro Leu Glu Ala Ile Gln
180 185 190
Leu Glu Leu Asp Pro Glu Glu Asp Ala Pro Val Leu Asp Trp Phe Tyr
195 200 205
Asp His Gln Pro Leu Arg Asp Ser Arg Lys Tyr Val Asn Gly Ser Thr
210 215 220
Tyr Gln Arg Trp Gln Phe Thr Leu Pro Met Met Ser Thr Leu Tyr Arg
225 230 235 240
Leu Ala Asn Gln Leu Leu Thr Asp Leu Val Asp Asp Asn Tyr Phe Tyr
245 250 255
Leu Phe Asp Leu Lys Ala Phe Phe Thr Ser Lys Ala Leu Asn Met Ala
260 265 270
Ile Pro Gly Gly Pro Lys Phe Glu Pro Leu Val Arg Asp Ile Asn Leu
275 280 285
Gln Asp Glu Asp Trp Asn Glu Phe Asn Asp Ile Asn Lys Ile Ile Ile
290 295 300
Arg Gln Pro Ile Arg Thr Glu Tyr Lys Ile Ala Phe Pro Tyr Leu Tyr
305 310 315 320
Asn Asn Leu Pro His His Val His Leu Thr Trp Tyr His Thr Pro Asn
325 330 335
Val Val Phe Ile Lys Thr Glu Asp Pro Asp Leu Pro Ala Phe Tyr Phe

340 345 350

Asp Pro Leu Ile Asn Pro Ile Ser His Arg His Ser Val Lys Ser Gln
355 360 365

Glu Pro Leu Pro Asp Asp Asp Glu Glu Phe Glu Leu Pro Glu Phe Val
370 375 380

Glu Pro Phe Leu Lys Asp Thr Pro Leu Tyr Thr Asp Asn Thr Ala Asn
385 390 395 400

Gly Ile Ala Leu Leu Trp Ala Pro Arg Pro Phe Asn Leu Arg Ser Gly
405 410 415

Arg Thr Arg Arg Ala Leu Asp Ile Pro Leu Val Lys Asn Trp Tyr Arg
420 425 430

Glu His Cys Pro Ala Gly Gln Pro Val Lys Val Arg Val Ser Tyr Gln
435 440 445

Lys Leu Leu Lys Tyr Tyr Val Leu Asn Ala Leu Lys His Arg Pro Pro
450 455 460

Lys Ala Gln Lys Lys Arg Tyr Leu Phe Arg Ser Phe Lys Ala Thr Lys
465 470 475 480

Phe Phe Gln Ser Thr Lys Leu Asp Trp Val Glu Gly Trp Leu Gln Val
485 490 495

Cys Arg Gln Gly Tyr Asn Met Leu Asn Leu Leu Ile His Arg Lys Asn
500 505 510

Leu Asn Tyr Leu His Leu Asp Tyr Asn Phe Asn Leu Lys Pro Val Lys
515 520 525

Thr Leu Thr Thr Lys Glu Arg Lys Lys Ser Arg Phe Gly Asn Ala Phe
530 535 540

His Leu Cys Arg Glu Val Leu Arg Leu Thr Lys Leu Val Val Asp Ser
545 550 555 560

His Val Gln Tyr Arg Leu Gly Asn Val Asp Ala Phe Gln Leu Ala Asp
565 570 575

Gly Leu Gln Tyr Ile Phe Ala His Val Gly Gln Leu Thr Gly Met Tyr
580 585 590

Arg Tyr Lys Tyr Lys Leu Met Arg Gln Ile Arg Val Cys Lys Asp Leu
595 600 605

Lys His Leu Ile Tyr Tyr Arg Phe Asn Thr Gly Pro Val Gly Lys Gly
610 615 620

Pro Gly Cys Gly Phe Trp Ala Ala Gly Trp Arg Val Trp Leu Phe Phe
625 630 635 640

Met Arg Gly Ile Thr Pro Leu Leu Glu Arg Trp Leu Gly Asn Leu Leu
645 650 655

Ala Arg Gln Phe Glu Gly Arg His Ser Lys Gly Val Ala Lys Thr Val
660 665 670
Thr Lys Gln Arg Val Glu Ser His Phe Asp Leu Glu Leu Arg Ala Ala
675 680 685
Val Met His Asp Ile Leu Asp Met Met Pro Glu Gly Ile Lys Gln Asn
690 695 700
Lys Ala Arg Thr Ile Leu Gln His Leu Ser Glu Ala Trp Arg Cys Trp
705 710 715 720
Lys Ala Asn Ile Pro Trp Lys Val Pro Gly Leu Pro Thr Pro Ile Glu
725 730 735
Asn Met Ile Leu Arg Tyr Val Lys Ala Lys Ala Asp Trp Trp Thr Asn
740 745 750
Thr Ala His Tyr Asn Arg Glu Arg Ile Arg Arg Gly Ala Thr Val Asp
755 760 765
Lys Thr Val Cys Lys Lys Asn Leu Gly Arg Leu Thr Arg Leu Tyr Leu
770 775 780
Lys Ala Glu Gln Glu Arg Gln His Asn Tyr Leu Lys Asp Gly Pro Tyr
785 790 795 800
Ile Thr Ala Glu Glu Thr Val Ala Val Tyr Thr Thr Thr Val His Trp
805 810 815
Leu Glu Ser Arg Arg Phe Ser Pro Ile Pro Phe Pro Pro Leu Ser Tyr
820 825 830
Lys His Asp Thr Lys Leu Leu Ile Leu Ala Leu Glu Arg Leu Lys Glu
835 840 845
Ala Tyr Ser Val Lys Ser Arg Leu Asn Gln Ser Gln Arg Glu Glu Leu
850 855 860
Gly Leu Ile Glu Gln Ala Tyr Asp Asn Leu His Glu Ala Leu Ser Arg
865 870 875 880
Ile Lys Arg His Leu Leu Thr Gln Arg Ala Phe Lys Glu Val Gly Ile
885 890 895
Glu Phe Met Asp Leu Tyr Ser His Leu Val Pro Val Tyr Asp Val Glu
900 905 910
Pro Leu Glu Lys Ile Thr Asp Ala Tyr Leu Asp Gln Tyr Leu Trp Tyr
915 920 925
Glu Ala Asp Lys Arg Arg Leu Phe Pro Pro Trp Ile Lys Pro Ala Asp
930 935 940
Thr Glu Pro Pro Pro Leu Leu Val Tyr Lys Trp Cys Gln Gly Ile Asn
945 950 955 960
Asn Leu Gln Asp Val Trp Glu Thr Ser Glu Gly Glu Cys Asn Val Met

965 970 975

Leu Glu Ser Arg Phe Glu Lys Met Tyr Glu Lys Ile Asp Leu Thr Leu

980 985 990

Leu Asn Arg Leu Val Arg Leu Ile Val Asp His Asn Ile Ala Asp Tyr

995 1000 1005

Met Thr Ala Lys Asn Asn Val Val Ile Asn Tyr Lys Asp Met Asn

1010 1015 1020

His Thr Asn Ser Tyr Gly Ile Ile Arg Gly Leu Gln Phe Ala Ser

1025 1030 1035

Phe Ile Val Gln Tyr Tyr Gly Leu Val Met Asp Leu Leu Val Leu

1040 1045 1050

Gly Leu His Arg Ala Ser Glu Met Ala Gly Pro Pro Gln Met Pro

1055 1060 1065

Asn Asp Phe Leu Ser Phe Gln Asp Ile Ala Thr Glu Ala Ala His

1070 1075 1080

Pro Ile Arg Leu Phe Cys Arg Tyr Ile Asp Arg Ile His Ile Phe

1085 1090 1095

Phe Arg Phe Thr Ala Asp Glu Ala Arg Asp Leu Ile Gln Arg Tyr

1100 1105 1110

Leu Thr Glu His Pro Asp Pro Asn Asn Glu Asn Ile Val Gly Tyr

1115 1120 1125

Asn Asn Lys Lys Cys Trp Pro Arg Asp Ala Arg Met Arg Leu Met

1130 1135 1140

Lys His Asp Val Asn Leu Gly Arg Ala Val Phe Trp Asp Ile Lys

1145 1150 1155

Asn Arg Leu Pro Arg Ser Val Thr Thr Val Gln Trp Glu Asn Ser

1160 1165 1170

Phe Val Ser Val Tyr Ser Lys Asp Asn Pro Asn Leu Leu Phe Asn

1175 1180 1185

Met Cys Gly Phe Glu Cys Arg Ile Leu Pro Lys Cys Arg Thr Ser

1190 1195 1200

Tyr Glu Glu Phe Thr His Lys Asp Gly Val Trp Asn Leu Gln Asn

1205 1210 1215

Glu Val Thr Lys Glu Arg Thr Ala Gln Cys Phe Leu Arg Val Asp

1220 1225 1230

Asp Glu Ser Met Gln Arg Phe His Asn Arg Val Arg Gln Ile Leu

1235 1240 1245

Met Ala Ser Gly Ser Thr Thr Phe Thr Lys Ile Val Asn Lys Trp

1250 1255 1260

Asn Thr Ala Leu Ile Gly Leu Met Thr Tyr Phe Arg Glu Ala Val
1265 1270 1275
Val Asn Thr Gln Glu Leu Leu Asp Leu Leu Val Lys Cys Glu His
1280 1285 1290
Lys Ile Gln Thr Arg Ile Lys Ile Gly Leu Asn Ser Lys Met Pro
1295 1300 1305
Ser Arg Phe Pro Pro Val Val Phe Tyr Thr Pro Lys Glu Leu Gly
1310 1315 1320
Gly Leu Gly Met Leu Ser Met Gly His Val Leu Ile Pro Gln Ser
1325 1330 1335
Asp Leu Arg Trp Ser Lys Gln Thr Asp Val Gly Ile Thr His Phe
1340 1345 1350
Arg Ser Gly Met Ser His Glu Glu Asp Gln Leu Ile Pro Asn Leu
1355 1360 1365
Tyr Arg Tyr Ile Gln Pro Trp Glu Ser Glu Phe Ile Asp Ser Gln
1370 1375 1380
Arg Val Trp Ala Glu Tyr Ser Leu Lys Arg Gln Glu Ala Ile Ala
1385 1390 1395
Gln Asn Arg Arg Leu Thr Leu Glu Asp Leu Glu Asp Ser Trp Asp
1400 1405 1410
Arg Gly Ile Pro Arg Ile Asn Thr Leu Phe Gln Lys Asp Arg His
1415 1420 1425
Thr Leu Ala Tyr Asp Lys Gly Trp Arg Val Arg Thr Asp Phe Lys
1430 1435 1440
Gln Tyr Gln Val Leu Lys Gln Asn Pro Phe Trp Trp Thr His Gln
1445 1450 1455
Arg His Asp Gly Lys Leu Trp Asn Leu Asn Asn Tyr Arg Thr Asp
1460 1465 1470
Met Ile Gln Ala Leu Gly Gly Val Glu Gly Ile Leu Glu His Thr
1475 1480 1485
Leu Phe Lys Gly Thr Tyr Phe Pro Thr Trp Glu Gly Leu Phe Trp
1490 1495 1500
Glu Lys Ala Ser Gly Phe Glu Glu Ser Met Lys Trp Lys Lys Leu
1505 1510 1515
Thr Asn Ala Gln Arg Ser Gly Leu Asn Gln Ile Pro Asn Arg Arg
1520 1525 1530
Phe Thr Leu Trp Trp Ser Pro Thr Ile Asn Arg Ala Asn Val Tyr
1535 1540 1545
Val Gly Phe Gln Val Gln Leu Asp Leu Thr Gly Ile Phe Met His

1550 1555 1560

Gly Lys Ile Pro Thr Leu Lys Ile Ser Leu Ile Gln Ile Phe Arg
1565 1570 1575

Ala His Leu Trp Gln Lys Ile His Glu Ser Ile Val Met Asp Leu
1580 1585 1590

Cys Gln Val Phe Asp Gln Glu Leu Asp Ala Leu Glu Ile Glu Thr
1595 1600 1605

Val Gln Lys Glu Thr Ile His Pro Arg Lys Ser Tyr Lys Met Asn
1610 1615 1620

Ser Ser Cys Ala Asp Ile Leu Leu Phe Ala Ser Tyr Lys Trp Asn
1625 1630 1635

Val Ser Arg Pro Ser Leu Leu Ala Asp Ser Lys Asp Val Met Asp
1640 1645 1650

Ser Thr Thr Thr Gln Lys Tyr Trp Ile Asp Ile Gln Leu Arg Trp
1655 1660 1665

Gly Asp Tyr Asp Ser His Asp Ile Glu Arg Tyr Ala Arg Ala Lys
1670 1675 1680

Phe Leu Asp Tyr Thr Thr Asp Asn Met Ser Ile Tyr Pro Ser Pro
1685 1690 1695

Thr Gly Val Leu Ile Ala Ile Asp Leu Ala Tyr Asn Leu His Ser
1700 1705 1710

Ala Tyr Gly Asn Trp Phe Pro Gly Ser Lys Pro Leu Ile Gln Gln
1715 1720 1725

Ala Met Ala Lys Ile Met Lys Ala Asn Pro Ala Leu Tyr Val Leu
1730 1735 1740

Arg Glu Arg Ile Arg Lys Gly Leu Gln Leu Tyr Ser Ser Glu Pro
1745 1750 1755

Thr Glu Pro Tyr Leu Ser Ser Gln Asn Tyr Gly Glu Leu Phe Ser
1760 1765 1770

Asn Gln Ile Ile Trp Phe Val Asp Asp Thr Asn Val Tyr Arg Val
1775 1780 1785

Thr Ile His Lys Thr Phe Glu Gly Asn Leu Thr Thr Lys Pro Ile
1790 1795 1800

Asn Gly Ala Ile Phe Ile Phe Asn Pro Arg Thr Gly Gln Leu Phe
1805 1810 1815

Leu Lys Ile Ile His Thr Ser Val Trp Ala Gly Gln Lys Arg Leu
1820 1825 1830

Gly Gln Leu Ala Lys Trp Lys Thr Ala Glu Glu Val Ala Ala Leu
1835 1840 1845

Ile Arg Ser Leu Pro Val Glu Glu Gln Pro Lys Gln Ile Ile Val
1850 1855 1860
Thr Arg Lys Asp Met Leu Asp Pro Leu Glu Val His Leu Leu Asp
1865 1870 1875
Phe Pro Asn Ile Val Ile Lys Gly Ser Glu Leu Gln Leu Pro Phe
1880 1885 1890
Gln Ala Cys Leu Lys Val Glu Lys Phe Gly Asp Leu Ile Leu Lys
1895 1900 1905
Ala Thr Glu Pro Gln Met Val Leu Phe Asn Leu Tyr Asp Asp Trp
1910 1915 1920
Leu Lys Thr Ile Ser Ser Tyr Thr Ala Phe Ser Arg Leu Ile Leu
1925 1930 1935
Ile Leu Arg Ala Leu His Val Asn Asn Asp Arg Ala Lys Val Ile
1940 1945 1950
Leu Lys Pro Asp Lys Thr Thr Ile Thr Glu Pro His His Ile Trp
1955 1960 1965
Pro Thr Leu Thr Asp Glu Glu Trp Ile Lys Val Glu Val Gln Leu
1970 1975 1980
Lys Asp Leu Ile Leu Ala Asp Tyr Gly Lys Lys Asn Asn Val Asn
1985 1990 1995
Val Ala Ser Leu Thr Gln Ser Glu Ile Arg Asp Ile Ile Leu Gly
2000 2005 2010
Met Glu Ile Ser Ala Pro Ser Gln Gln Arg Gln Gln Ile Ala Glu
2015 2020 2025
Ile Glu Lys Gln Thr Lys Glu Gln Ser Gln Leu Thr Ala Thr Gln
2030 2035 2040
Thr Arg Thr Val Asn Lys His Gly Asp Glu Ile Ile Thr Ser Thr
2045 2050 2055
Thr Ser Asn Tyr Glu Thr Gln Thr Phe Ser Ser Lys Thr Glu Trp
2060 2065 2070
Arg Val Arg Ala Ile Ser Ala Ala Asn Leu His Leu Arg Thr Asn
2075 2080 2085
His Ile Tyr Val Ser Ser Asp Asp Ile Lys Glu Thr Gly Tyr Thr
2090 2095 2100
Tyr Ile Leu Pro Lys Asn Val Leu Lys Lys Phe Ile Cys Ile Ser
2105 2110 2115
Asp Leu Arg Ala Gln Ile Ala Gly Tyr Leu Tyr Gly Val Ser Pro
2120 2125 2130
Pro Asp Asn Pro Gln Val Lys Glu Ile Arg Cys Ile Val Met Val

2135 2140 2145
Pro Gln Trp Gly Thr His Gln Thr Val His Leu Pro Gly Gln Leu
2150 2155 2160
Pro Gln His Glu Tyr Leu Lys Glu Met Glu Pro Leu Gly Trp Ile
2165 2170 2175
His Thr Gln Pro Asn Glu Ser Pro Gln Leu Ser Pro Gln Asp Val
2180 2185 2190
Thr Thr His Ala Lys Ile Met Ala Asp Asn Pro Ser Trp Asp Gly
2195 2200 2205
Glu Lys Thr Ile Ile Ile Thr Cys Ser Phe Thr Pro Gly Ser Cys
2210 2215 2220
Thr Leu Thr Ala Tyr Lys Leu Thr Pro Ser Gly Tyr Glu Trp Gly
2225 2230 2235
Arg Gln Asn Thr Asp Lys Gly Asn Asn Pro Lys Gly Tyr Leu Pro
2240 2245 2250
Ser His Tyr Glu Arg Val Gln Met Leu Leu Ser Asp Arg Phe Leu
2255 2260 2265
Gly Phe Phe Met Val Pro Ala Gln Ser Ser Trp Asn Tyr Asn Phe
2270 2275 2280
Met Gly Val Arg His Asp Pro Asn Met Lys Tyr Glu Leu Gln Leu
2285 2290 2295
Ala Asn Pro Lys Glu Phe Tyr His Glu Val His Arg Pro Ser His
2300 2305 2310
Phe Leu Asn Phe Ala Leu Leu Gln Glu Gly Glu Val Tyr Ser Ala
2315 2320 2325
Asp Arg Glu Asp Leu Tyr Ala
2330 2335

<210> 60
<211> 7221
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 60
cgggcggcct cttgtgtgag ggcctgtggg attctccgga tatggccgga gtgtttcctt 60
atcgagggcc gggtaacccg gtgcctggcc ctctagcccc gctaccggac tacatgtcgg 120
aggagaagct gcaggagaaa gctcgaaaat ggcagcaatt gcaggccaag cgctatgcag 180
aaaagcggaa gtttgggttt gtggatgccc agaaggaaga catgccccca gaacatgtca 240
gggagatcat tcgagaccat ggagacatga ccaacaggaa gttccgccat gacaaaaggg 300

tttacttggg tgccctaaag tacatgcccc acgcagtcct caaactcctg gagaacatgc 360
ctatgccttg ggagcagatt cgggatgtgc ccgtgctgta ccacatcact ggagccattt 420
ccttcgtcaa tgagattccc tgggtcattg aacctgtcta catctcccag tgggggtcaa 480
tgtggattat gatgcgccga gaaaaaagag ataggaggca tttcaagaga atgcgttttc 540
ccccttttga tgatgaggag ccgcccttgg actatgctga caacatccta aatgttgagc 600
cactggaggc cattcagcta gagctggacc ctgaggagga cgcccctgtg ttggactggt 660
tctatgacca ccagccgttg agggacagca ggaagtatgt aaatggctcc acttaccagc 720
gctggcagtt cacactacct atgatgtcaa ctctctaccg cctggctaat cagctcctga 780
cagacttggt ggatgacaac tacttctacc tgtttgattt gaaggccttc tttacgtcca 840
aggcactcaa tatggccatt cctggaggcc ccaaatttga acctcttgtt cgagacatca 900
acctacagga tgaagactgg aatgaattca atgatattaa caagattatc atccggcagc 960
ctatccggac tgagtacaag attgcttttc cttacttgta caacaatctt ccacaccatg 1020
tccacctcac ctggtaccat actcccaatg ttgtattcat caaaactgaa gatcctgact 1080
tgccagcttt ctactttgac cctttgatca acccaatctc ccataggcac tcagtcaaga 1140
gccaggaacc attgccggat gatgatgagg aatttgagct cccggagttt gtggagccct 1200
tcctgaagga cacacccctc tatacagaca atacagccaa tggcattgcc ctgctctggg 1260
ccccgcggcc cttcaaccta cgctctggtc gcacccgtcg ggccctggac atacccctg 1320
tcaagaactg gtatcgggag cattgtcctg ccgggcagcc tgtgaaagtg agggtctcct 1380
accagaagct gcttaagtac tatgtgctga atgccctgaa gcatcggccc cctaaggctc 1440
aaaagaagag gtatttgttc cgctccttca aagccaccaa attctttcag tccacaaagc 1500
tggactgggt ggagggttgg ctccaggttt gccgccaggg ctacaacatg ctcaaccttc 1560
tcattcaccg caaaaacctc aactacctgc acctggacta caacttcaac ctcaagcctg 1620
tgaaaacgct caccaccaag gaaagaaaga aatctcgttt tgggaatgct ttccacctgt 1680
gtcgggaagt tctgcgtttg actaagctgg tggtggatag tcacgtgcag tatcggctgg 1740
gcaatgtgga tgccttccag ctggcagatg gattgcagta tatatttgcc catgttgggc 1800
agttgacggg catgtatcga tacaaataca agctgatgcg acagattcgc gtgtgcaagg 1860
acctgaagca tctcatctat tatcgtttca acacaggccc tgtagggaag ggtcctggct 1920
gtggcttctg ggctgccggt tggcgagtct ggctcttttt catgcgtggc attacccctt 1980
tattagagcg atggcttggc aacctcctgg cccggcagtt tgaaggtcga cactcaaagg 2040
gggtggcaaa gacagtaaca aagcagcgag tggagtcaca ttttgacctt gagctgcggg 2100
cagctgtgat gcatgatatt ctggacatga tgcctgaggg gatcaagcag aacaaggccc 2160
ggacaatcct gcagcacctc agtgaagcct ggcgctgctg gaaagccaac attccctgga 2220
aggtccctgg gctgccgacg cccatagaga atatgatcct tcgatacgtg aaggccaagg 2280
ctgactggtg gaccaacact gcccactaca accgagaacg gatccgccga ggggccactg 2340
tggacaagac tgtttgtaaa aagaatctgg gccgcctcac ccggctctat ctgaaggcag 2400
aacaggagcg gcagcacaac tacctgaagg acgggcctta catcacagcg gaggaaacag 2460
tggcagtata taccaccaca gtgcattggt tggaaagccg caggttttca cccatcccat 2520
tccccccact ctcctataag catgacacca agttgctcat cttggcattg gagcggctca 2580
aggaagctta tagtgtgaag tctcggttga accagtctca gagggaggag ctaggtctga 2640

tcgagcaggc ctacgataac ctccacgagg cgctgtcccg cataaagcgt cacctcctca 2700
cacagagagc cttcaaagag gtgggcattg agttcatgga tctgtatagc cacctcgttc 2760
cagtatatga tgttgagccc ctggagaaga taactgatgc ttacctggac cagtacctgt 2820
ggtatgaagc cgacaagcgc cgcctgttcc caccctggat taagcctgca gacacagaac 2880
cacctccact gcttgtttac aagtggtgtc aaggcatcaa taacctgcag gacgtgtggg 2940
agacgagtga aggcgagtgc aatgtcatgc tggaatcccg ctttgagaag atgtatgaga 3000
agatcgactt gactctgctc aacaggctcg tgcgcctcat cgtggaccac aacatagccg 3060
actacatgac agccaagaac aacgtcgtca tcaactataa ggacatgaac catacgaatt 3120
catatgggat catcagaggc ctgcagtttg cctcattcat agtgcagtat tatggcctgg 3180
tgatggattt gcttgtattg ggattgcacc gggccagtga gatggctggg ccccctcaga 3240
tgccaaatga ctttctcagt ttccaggaca tagccactga ggctgcccac cccatccgtc 3300
tcttctgcag atacattgat cgcatccata ttttttttcag gttcacagca gatgaggctc 3360
gggacctgat tcaacgttac ctgacagagc accctgaccc caataatgaa aacatcgttg 3420
gctataataa caagaagtgc tggccccgag atgcccgcat gcgcctcatg aaacatgatg 3480
ttaacttagg ccgggcggta ttctgggaca tcaagaaccg cttgccacgg tcagtgacta 3540
cagttcagtg ggagaacagc ttcgtgtctg tgtacagtaa ggacaacccc aacctgctgt 3600
tcaacatgtg tggcttcgag tgccgcatcc tgcctaagtg ccgcaccagc tatgaggagt 3660
tcacccacaa ggacggggtc tggaacctgc agaatgaggt tactaaggag cgcacagctc 3720
agtgtttcct gcgtgtggac gatgagtcaa tgcagcgctt ccacaaccgc gtgcgtcaga 3780
ttctcatggc ctctgggtcc accaccttca ccaagattgt gaataagtgg aatacagctc 3840
tcattggcct tatgacatac tttcgggagg ctgtggtgaa cacccaagag ctcttggact 3900
tactggtgaa gtgtgagcac aaaatccaga cacgtatcaa gattggactc aactccaaga 3960
tgccaagtcg gttcccccg gttgtgttct acacccctaa ggagttgggt ggactcggca 4020
tgctctcaat gggccatgtg ctcatccccc aatccgacct caggtggtcc aaacagacag 4080
atgtaggtat cacacacttt cgttcaggaa tgagccatga agaagaccag ctcattccca 4140
acttgtaccg ctacatacag ccatgggaga gcgagttcat tgattctcag cgggtctggg 4200
ctgagtactc actcaagaga caagaggcca ttgctcagaa cagacgcctg actttagaag 4260
acctagaaga ttcatgggat cgtggcattc ctcgaatcaa taccctcttc cagaaggacc 4320
ggcacacact ggcttatgat aagggctggc gtgtcagaac tgactttaag cagtatcagg 4380
ttttgaagca gaatccgttc tggtggacac accagcggca tgatgggaag ctctggaacc 4440
tgaacaacta ccgtacagac atgatccagg ccctgggcgg tgtggaaggc attctggaac 4500
acacactctt taagggcact tacttcccta cctgggaggg gcttttctgg gagaaggcca 4560
gtggctttga ggaatctatg aagtggaaga agctaactaa tgctcagcga tcaggactga 4620
accagattcc caatcgtaga ttcaccctct ggtggtcccc gaccattaat cgagccaatg 4680
tatatgtagg ctttcaggtg cagctagacc tgacgggtat cttcatgcac ggcaagatcc 4740
ccacgctgaa gatctctctc atccagatct tccgagctca cttgtggcag aagatccatg 4800
agagcattgt tatggactta tgtcaggtgt ttgaccagga acttgatgca ctggaaattg 4860
agacagtaca aaaggagaca atccatcccc gaaagtcata taagatgaac tcttcctgtg 4920
cagatatcct gctctttgcc tcctataagt ggaatgtctc ccggccctca ttgctggctg 4980

actccaagga tgtgatggac agcaccacca cccagaaata ctggattgac atccagttgc 5040
gctgggggga ctatgattcc cacgacattg agcgctacgc ccgggccaag ttcctggact 5100
acaccaccga caacatgagt atctaccctt cgcccacagg tgtactcatc gccattgacc 5160
tggcctataa cttgcacagt gcctatggaa actggttccc aggcagcaag cctctcatac 5220
aacaggccat ggccaagatc atgaaggcaa accctgccct gtatgtgtta cgtgaacgga 5280
tccgcaaggg gctacagctc tattcatctg aacccactga gccttatttg tcttctcaga 5340
actatggtga gctcttctcc aaccagatta tctggtttgt ggatgacacc aacgtctaca 5400
gagtgactat tcacaagacc tttgaaggga acttgacaac caagcccatc aacggagcca 5460
tcttcatctt caacccacgc acagggcagc tgttcctcaa gataatccac acgtccgtgt 5520
gggcgggaca gaagcgtttg gggcagttgg ctaagtggaa gacagctgag gaggtggccg 5580
ccctgatccg atctctgcct gtggaggagc agcccaagca gatcattgtc accaggaagg 5640
acatgctgga cccactggag gtgcacttac tggacttccc caatattgtc atcaaaggat 5700
cggagctcca actccctttc caggcgtgtc tcaaggtgga aaaattcggg gatctcatcc 5760
ttaaagccac tgagccccag atggttctct tcaacctcta tgacgactgg ctcaagacta 5820
tttcatctta cacggccttc tcccgtctca tcctgattct gcgtgcccta catgtgaaca 5880
acgatcgggc aaaagtgatc ctgaagccag acaagactac tattacagaa ccacaccaca 5940
tctggcccac tctgactgac gaagaatgga tcaaggtcga ggtgcagctc aaggatctga 6000
tcttggctga ctacggcaag aaaaacaatg tgaacgtggc atcactgaca caatcagaaa 6060
ttcgagacat catcctgggt atggagatct cggcaccgtc acagcagcgg cagcagatcg 6120
ctgagatcga gaagcagacc aaggaacaat cgcagctgac ggcaacacag actcgcactg 6180
tcaacaagca tggcgatgag atcatcacct ccaccaccag caactatgag acccagactt 6240
tctcatccaa gactgagtgg agggtcaggg ccatctctgc tgccaacctg cacctaagga 6300
ccaatcacat ctatgtttca tctgacgaca tcaaggagac tggctacacc tacatccttc 6360
ccaagaatgt gcttaagaag ttcatctgca tatctgacct tcgggcccaa attgcaggat 6420
acctatatgg ggtgagccca ccagataacc cccaggtgaa ggagatccgc tgcattgtga 6480
tggtgccgca gtggggcact caccagaccg tgcacctgcc tggccagctg ccccagcatg 6540
agtacctcaa ggagatggaa cccttaggtt ggatccacac tcagcccaat gagtccccgc 6600
agttatcacc ccaggatgtc accacccatg ccaagatcat ggctgacaac ccatcttggg 6660
atggcgagaa gaccattatc atcacatgca gcttcacgcc aggctcctgt acactgacgg 6720
cctacaagct gacccctagt ggctacgaat ggggccgcca gaacacagac aagggcaaca 6780
accccaaggg ctacctgcct tcacactatg agagggtgca gatgctgctg tcggaccgtt 6840
tccttggctt cttcatggtc cctgcccagt cctcgtggaa ctacaacttc atgggtgttc 6900
ggcatgaccc caacatgaaa tatgagctac agctggcgaa ccccaaagag ttctaccacg 6960
aggtgcacag gccctctcac ttcctcaact ttgctctcct gcaggagggg gaggtttact 7020
ctgcggatcg ggaggacctg tatgcctgac cgtttccctg cctcctgctt cagcctcccg 7080
aggccgaagc ctcagcccct ccagacaggc cgctgacatt cagcagtttg gcctctttcc 7140
ctctgtctgt gcttgtgttg ttgacctcct gatggcttgt catcctgaat aaaatataat 7200
aataaatttt gtataaatag g 7221

<210> 61

<211> 2479
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 61

Met Pro Pro Asn Ile Asn Trp Lys Glu Ile Met Lys Val Asp Pro Asp
1 5 10 15
Asp Leu Pro Arg Gln Glu Glu Leu Ala Asp Asn Leu Leu Ile Ser Leu
20 25 30
Ser Lys Val Glu Val Asn Glu Leu Lys Ser Glu Lys Gln Glu Asn Val
35 40 45
Ile His Leu Phe Arg Ile Thr Gln Ser Leu Met Lys Met Lys Ala Gln
50 55 60
Glu Val Glu Leu Ala Leu Glu Glu Val Glu Lys Ala Gly Glu Glu Gln
65 70 75 80
Ala Lys Phe Glu Asn Gln Leu Lys Thr Lys Val Met Lys Leu Glu Asn
85 90 95
Glu Leu Glu Met Ala Gln Gln Ser Ala Gly Gly Arg Asp Thr Arg Phe
100 105 110
Leu Arg Asn Glu Ile Cys Gln Leu Glu Lys Gln Leu Glu Gln Lys Asp
115 120 125
Arg Glu Leu Glu Asp Met Glu Lys Glu Leu Glu Lys Glu Lys Lys Val
130 135 140
Asn Glu Gln Leu Ala Leu Arg Asn Glu Glu Ala Glu Asn Glu Asn Ser
145 150 155 160
Lys Leu Arg Arg Glu Asn Lys Arg Leu Lys Lys Lys Asn Glu Gln Leu
165 170 175
Cys Gln Asp Ile Ile Asp Tyr Gln Lys Gln Ile Asp Ser Gln Lys Glu
180 185 190
Thr Leu Leu Ser Arg Arg Gly Glu Asp Ser Asp Tyr Arg Ser Gln Leu
195 200 205
Ser Lys Lys Asn Tyr Glu Leu Ile Gln Tyr Leu Asp Glu Ile Gln Thr
210 215 220
Leu Thr Glu Ala Asn Glu Lys Ile Glu Val Gln Asn Gln Glu Met Arg
225 230 235 240
Lys Asn Leu Glu Glu Ser Val Gln Glu Met Glu Lys Met Thr Asp Glu
245 250 255
Tyr Asn Arg Met Lys Ala Ile Val His Gln Thr Asp Asn Val Ile Asp

260 265 270

Gln Leu Lys Lys Glu Asn Asp His Tyr Gln Leu Gln Val Gln Glu Leu

275 280 285

Thr Asp Leu Leu Lys Ser Lys Asn Glu Glu Asp Asp Pro Ile Met Val

290 295 300

Ala Val Asn Ala Lys Val Glu Glu Trp Lys Leu Ile Leu Ser Ser Lys

305 310 315 320

Asp Asp Glu Ile Ile Glu Tyr Gln Gln Met Leu His Asn Leu Arg Glu

325 330 335

Lys Leu Lys Asn Ala Gln Leu Asp Ala Asp Lys Ser Asn Val Met Ala

340 345 350

Leu Gln Gln Gly Ile Gln Glu Arg Asp Ser Gln Ile Lys Met Leu Thr

355 360 365

Glu Gln Val Glu Gln Tyr Thr Lys Glu Met Glu Lys Asn Thr Cys Ile

370 375 380

Ile Glu Asp Leu Lys Asn Glu Leu Gln Arg Asn Lys Gly Ala Ser Thr

385 390 395 400

Leu Ser Gln Gln Thr His Met Lys Ile Gln Ser Thr Leu Asp Ile Leu

405 410 415

Lys Glu Lys Thr Lys Glu Ala Glu Arg Thr Ala Glu Leu Ala Glu Ala

420 425 430

Asp Ala Arg Glu Lys Asp Lys Glu Leu Val Glu Ala Leu Lys Arg Leu

435 440 445

Lys Asp Tyr Glu Ser Gly Val Tyr Gly Leu Glu Asp Ala Val Val Glu

450 455 460

Ile Lys Asn Cys Lys Asn Gln Ile Lys Ile Arg Asp Arg Glu Ile Glu

465 470 475 480

Ile Leu Thr Lys Glu Ile Asn Lys Leu Glu Leu Lys Ile Ser Asp Phe

485 490 495

Leu Asp Glu Asn Glu Ala Leu Arg Glu Arg Val Gly Leu Glu Pro Lys

500 505 510

Thr Met Ile Asp Leu Thr Glu Phe Arg Asn Ser Lys His Leu Lys Gln

515 520 525

Gln Gln Tyr Arg Ala Glu Asn Gln Ile Leu Leu Lys Glu Ile Glu Ser

530 535 540

Leu Glu Glu Glu Arg Leu Asp Leu Lys Lys Lys Ile Arg Gln Met Ala

545 550 555 560

Gln Glu Arg Gly Lys Arg Ser Ala Thr Ser Gly Leu Thr Thr Glu Asp

565 570 575

Leu Asn Leu Thr Glu Asn Ile Ser Gln Gly Asp Arg Ile Ser Glu Arg
580 585 590
Lys Leu Asp Leu Leu Ser Leu Lys Asn Met Ser Glu Ala Gln Ser Lys
595 600 605
Asn Glu Phe Leu Ser Arg Glu Leu Ile Glu Lys Glu Arg Asp Leu Glu
610 615 620
Arg Ser Arg Thr Val Ile Ala Lys Phe Gln Asn Lys Leu Lys Glu Leu
625 630 635 640
Val Glu Glu Asn Lys Gln Leu Glu Glu Gly Met Lys Glu Ile Leu Gln
645 650 655
Ala Ile Lys Glu Met Gln Lys Asp Pro Asp Val Lys Gly Gly Glu Thr
660 665 670
Ser Leu Ile Ile Pro Ser Leu Glu Arg Leu Val Asn Ala Ile Glu Ser
675 680 685
Lys Asn Ala Glu Gly Ile Phe Asp Ala Ser Leu His Leu Lys Ala Gln
690 695 700
Val Asp Gln Leu Thr Gly Arg Asn Glu Glu Leu Arg Gln Glu Leu Arg
705 710 715 720
Glu Ser Arg Lys Glu Ala Ile Asn Tyr Ser Gln Gln Leu Ala Lys Ala
725 730 735
Asn Leu Lys Ile Asp His Leu Glu Lys Glu Thr Ser Leu Leu Arg Gln
740 745 750
Ser Glu Gly Ser Asn Val Val Phe Lys Gly Ile Asp Leu Pro Asp Gly
755 760 765
Ile Ala Pro Ser Ser Ala Ser Ile Ile Asn Ser Gln Asn Glu Tyr Leu
770 775 780
Ile His Leu Leu Gln Glu Leu Glu Asn Lys Glu Lys Lys Leu Lys Asn
785 790 795 800
Leu Glu Asp Ser Leu Glu Asp Tyr Asn Arg Lys Phe Ala Val Ile Arg
805 810 815
His Gln Gln Ser Leu Leu Tyr Lys Glu Tyr Leu Ser Glu Lys Glu Thr
820 825 830
Trp Lys Thr Glu Ser Lys Thr Ile Lys Glu Glu Lys Arg Lys Leu Glu
835 840 845
Asp Gln Val Gln Gln Asp Ala Ile Lys Val Lys Glu Tyr Asn Asn Leu
850 855 860
Leu Asn Ala Leu Gln Met Asp Ser Asp Glu Met Lys Lys Ile Leu Ala
865 870 875 880
Glu Asn Ser Arg Lys Ile Thr Val Leu Gln Val Asn Glu Lys Ser Leu

885 890 895
Ile Arg Gln Tyr Thr Thr Leu Val Glu Leu Glu Arg Gln Leu Arg Lys
900 905 910
Glu Asn Glu Lys Gln Lys Asn Glu Leu Leu Ser Met Glu Ala Glu Val
915 920 925
Cys Glu Lys Ile Gly Cys Leu Gln Arg Phe Lys Glu Met Ala Ile Phe
930 935 940
Lys Ile Ala Ala Leu Gln Lys Val Val Asp Asn Ser Val Ser Leu Ser
945 950 955 960
Glu Leu Glu Leu Ala Asn Lys Gln Tyr Asn Glu Leu Thr Ala Lys Tyr
965 970 975
Arg Asp Ile Leu Gln Lys Asp Asn Met Leu Val Gln Arg Thr Ser Asn
980 985 990
Leu Glu His Leu Glu Cys Glu Asn Ile Ser Leu Lys Glu Gln Val Glu
995 1000 1005
Ser Ile Asn Lys Glu Leu Glu Ile Thr Lys Glu Lys Leu His Thr
1010 1015 1020
Ile Glu Gln Ala Trp Glu Gln Glu Thr Lys Leu Gly Asn Glu Ser
1025 1030 1035
Ser Met Asp Lys Ala Lys Lys Ser Ile Thr Asn Ser Asp Ile Val
1040 1045 1050
Ser Ile Ser Lys Lys Ile Thr Met Leu Glu Met Lys Glu Leu Asn
1055 1060 1065
Glu Arg Gln Arg Ala Glu His Cys Gln Lys Met Tyr Glu His Leu
1070 1075 1080
Arg Thr Ser Leu Lys Gln Met Glu Glu Arg Asn Phe Glu Leu Glu
1085 1090 1095
Thr Lys Phe Ala Glu Leu Thr Lys Ile Asn Leu Asp Ala Gln Lys
1100 1105 1110
Val Glu Gln Met Leu Arg Asp Glu Leu Ala Asp Ser Val Ser Lys
1115 1120 1125
Ala Val Ser Asp Ala Asp Arg Gln Arg Ile Leu Glu Leu Glu Lys
1130 1135 1140
Asn Glu Met Glu Leu Lys Val Glu Val Ser Lys Leu Arg Glu Ile
1145 1150 1155
Ser Asp Ile Ala Arg Arg Gln Val Glu Ile Leu Asn Ala Gln Gln
1160 1165 1170
Gln Ser Arg Asp Lys Glu Val Glu Ser Leu Arg Met Gln Leu Leu
1175 1180 1185

Asp Tyr Gln Ala Gln Ser Asp Glu Lys Ser Leu Ile Ala Lys Leu
1190 1195 1200
His Gln His Asn Val Ser Leu Gln Leu Ser Glu Ala Thr Ala Leu
1205 1210 1215
Gly Lys Leu Glu Ser Ile Thr Ser Lys Leu Gln Lys Met Glu Ala
1220 1225 1230
Tyr Asn Leu Arg Leu Glu Gln Lys Leu Asp Glu Lys Glu Gln Ala
1235 1240 1245
Leu Tyr Tyr Ala Arg Leu Glu Gly Arg Asn Arg Ala Lys His Leu
1250 1255 1260
Arg Gln Thr Ile Gln Ser Leu Arg Arg Gln Phe Ser Gly Ala Leu
1265 1270 1275
Pro Leu Ala Gln Gln Glu Lys Phe Ser Lys Thr Met Ile Gln Leu
1280 1285 1290
Gln Asn Asp Lys Leu Lys Ile Met Gln Glu Met Lys Asn Ser Gln
1295 1300 1305
Gln Glu His Arg Asn Met Glu Asn Lys Thr Leu Glu Met Glu Leu
1310 1315 1320
Lys Leu Lys Gly Leu Glu Glu Leu Ile Ser Thr Leu Lys Asp Thr
1325 1330 1335
Lys Gly Ala Gln Lys Val Ile Asn Trp His Met Lys Ile Glu Glu
1340 1345 1350
Leu Arg Leu Gln Glu Leu Lys Leu Asn Arg Glu Leu Val Lys Asp
1355 1360 1365
Lys Glu Glu Ile Lys Tyr Leu Asn Asn Ile Ile Ser Glu Tyr Glu
1370 1375 1380
Arg Thr Ile Ser Ser Leu Glu Glu Glu Ile Val Gln Gln Asn Lys
1385 1390 1395
Phe His Glu Glu Arg Gln Met Ala Trp Asp Gln Arg Glu Val Asp
1400 1405 1410
Leu Glu Arg Gln Leu Asp Ile Phe Asp Arg Gln Gln Asn Glu Ile
1415 1420 1425
Leu Asn Ala Ala Gln Lys Phe Glu Glu Ala Thr Gly Ser Ile Pro
1430 1435 1440
Asp Pro Ser Leu Pro Leu Pro Asn Gln Leu Glu Ile Ala Leu Arg
1445 1450 1455
Lys Ile Lys Glu Asn Ile Arg Ile Ile Leu Glu Thr Arg Ala Thr
1460 1465 1470
Cys Lys Ser Leu Glu Glu Lys Leu Lys Glu Lys Glu Ser Ala Leu

1475 1480 1485

Arg Leu Ala Glu Gln Asn Ile Leu Ser Arg Asp Lys Val Ile Asn

1490 1495 1500

Glu Leu Arg Leu Arg Leu Pro Ala Thr Ala Glu Arg Glu Lys Leu

1505 1510 1515

Ile Ala Glu Leu Gly Arg Lys Glu Met Glu Pro Lys Ser His His

1520 1525 1530

Thr Leu Lys Ile Ala His Gln Thr Ile Ala Asn Met Gln Ala Arg

1535 1540 1545

Leu Asn Gln Lys Glu Glu Val Leu Lys Lys Tyr Gln Arg Leu Leu

1550 1555 1560

Glu Lys Ala Arg Glu Glu Gln Arg Glu Ile Val Lys Lys His Glu

1565 1570 1575

Glu Asp Leu His Ile Leu His His Arg Leu Glu Leu Gln Ala Asp

1580 1585 1590

Ser Ser Leu Asn Lys Phe Lys Gln Thr Ala Trp Asp Leu Met Lys

1595 1600 1605

Gln Ser Pro Thr Pro Val Pro Thr Asn Lys His Phe Ile Arg Leu

1610 1615 1620

Ala Glu Met Glu Gln Thr Val Ala Glu Gln Asp Asp Ser Leu Ser

1625 1630 1635

Ser Leu Leu Val Lys Leu Lys Lys Val Ser Gln Asp Leu Glu Arg

1640 1645 1650

Gln Arg Glu Ile Thr Glu Leu Lys Val Lys Glu Phe Glu Asn Ile

1655 1660 1665

Lys Leu Gln Leu Gln Glu Asn His Glu Asp Glu Val Lys Lys Val

1670 1675 1680

Lys Ala Glu Val Glu Asp Leu Lys Tyr Leu Leu Asp Gln Ser Gln

1685 1690 1695

Lys Glu Ser Gln Cys Leu Lys Ser Glu Leu Gln Ala Gln Lys Glu

1700 1705 1710

Ala Asn Ser Arg Ala Pro Thr Thr Thr Met Arg Asn Leu Val Glu

1715 1720 1725

Arg Leu Lys Ser Gln Leu Ala Leu Lys Glu Lys Gln Gln Lys Ala

1730 1735 1740

Leu Ser Arg Ala Leu Leu Glu Leu Arg Ala Glu Met Thr Ala Ala

1745 1750 1755

Ala Glu Glu Arg Ile Ile Ser Ala Thr Ser Gln Lys Glu Ala His

1760 1765 1770

Leu Asn Val Gln Gln Ile Val Asp Arg His Thr Arg Glu Leu Lys
1775 1780 1785
Thr Gln Val Glu Asp Leu Asn Glu Asn Leu Leu Lys Leu Lys Glu
1790 1795 1800
Ala Leu Lys Thr Ser Lys Asn Arg Glu Asn Ser Leu Thr Asp Asn
1805 1810 1815
Leu Asn Asp Leu Asn Asn Glu Leu Gln Lys Lys Gln Lys Ala Tyr
1820 1825 1830
Asn Lys Ile Leu Arg Glu Lys Glu Glu Ile Asp Gln Glu Asn Asp
1835 1840 1845
Glu Leu Lys Arg Gln Ile Lys Arg Leu Thr Ser Gly Leu Gln Gly
1850 1855 1860
Lys Pro Leu Thr Asp Asn Lys Gln Ser Leu Ile Glu Glu Leu Gln
1865 1870 1875
Arg Lys Val Lys Lys Leu Glu Asn Gln Leu Glu Gly Lys Val Glu
1880 1885 1890
Glu Val Asp Leu Lys Pro Met Lys Glu Lys Asn Ala Lys Glu Glu
1895 1900 1905
Leu Ile Arg Trp Glu Glu Gly Lys Lys Trp Gln Ala Lys Ile Glu
1910 1915 1920
Gly Ile Arg Asn Lys Leu Lys Glu Lys Glu Gly Glu Val Phe Thr
1925 1930 1935
Leu Thr Lys Gln Leu Asn Thr Leu Lys Asp Leu Phe Ala Lys Ala
1940 1945 1950
Asp Lys Glu Lys Leu Thr Leu Gln Arg Lys Leu Lys Thr Thr Gly
1955 1960 1965
Met Thr Val Asp Gln Val Leu Gly Ile Arg Ala Leu Glu Ser Glu
1970 1975 1980
Lys Glu Leu Glu Glu Leu Lys Lys Arg Asn Leu Asp Leu Glu Asn
1985 1990 1995
Asp Ile Leu Tyr Met Arg Ala His Gln Ala Leu Pro Arg Asp Ser
2000 2005 2010
Val Val Glu Asp Leu His Leu Gln Asn Arg Tyr Leu Gln Glu Lys
2015 2020 2025
Leu His Ala Leu Glu Lys Gln Phe Ser Lys Asp Thr Tyr Ser Lys
2030 2035 2040
Pro Ser Ile Ser Gly Ile Glu Ser Asp Asp His Cys Gln Arg Glu
2045 2050 2055
Gln Glu Leu Gln Lys Glu Asn Leu Lys Leu Ser Ser Glu Asn Ile

2060 2065 2070

Glu Leu Lys Phe Gln Leu Glu Gln Ala Asn Lys Asp Leu Pro Arg
2075 2080 2085

Leu Lys Asn Gln Val Arg Asp Leu Lys Glu Met Cys Glu Phe Leu
2090 2095 2100

Lys Lys Glu Lys Ala Glu Val Gln Arg Lys Leu Gly His Val Arg
2105 2110 2115

Gly Ser Gly Arg Ser Gly Lys Thr Ile Pro Glu Leu Glu Lys Thr
2120 2125 2130

Ile Gly Leu Met Lys Lys Val Val Glu Lys Val Gln Arg Glu Asn
2135 2140 2145

Glu Gln Leu Lys Lys Ala Ser Gly Ile Leu Thr Ser Glu Lys Met
2150 2155 2160

Ala Asn Ile Glu Gln Glu Asn Glu Lys Leu Lys Ala Glu Leu Glu
2165 2170 2175

Lys Leu Lys Ala His Leu Gly His Gln Leu Ser Met His Tyr Glu
2180 2185 2190

Ser Lys Thr Lys Gly Thr Glu Lys Ile Ile Ala Glu Asn Glu Arg
2195 2200 2205

Leu Arg Lys Glu Leu Lys Lys Glu Thr Asp Ala Ala Glu Lys Leu
2210 2215 2220

Arg Ile Ala Lys Asn Asn Leu Glu Ile Leu Asn Glu Lys Met Thr
2225 2230 2235

Val Gln Leu Glu Glu Thr Gly Lys Arg Leu Gln Phe Ala Glu Ser
2240 2245 2250

Arg Gly Pro Gln Leu Glu Gly Ala Asp Ser Lys Ser Trp Lys Ser
2255 2260 2265

Ile Val Val Thr Arg Met Tyr Glu Thr Lys Leu Lys Glu Leu Glu
2270 2275 2280

Thr Asp Ile Ala Lys Lys Asn Gln Ser Ile Thr Asp Leu Lys Gln
2285 2290 2295

Leu Val Lys Glu Ala Thr Glu Arg Glu Gln Lys Val Asn Lys Tyr
2300 2305 2310

Asn Glu Asp Leu Glu Gln Gln Ile Lys Ile Leu Lys His Val Pro
2315 2320 2325

Glu Gly Ala Glu Thr Glu Gln Gly Leu Lys Arg Glu Leu Gln Val
2330 2335 2340

Leu Arg Leu Ala Asn His Gln Leu Asp Lys Glu Lys Ala Glu Leu
2345 2350 2355

Ile His Gln Ile Glu Ala Asn Lys Asp Gln Ser Gly Ala Glu Ser
2360 2365 2370
Thr Ile Pro Asp Ala Asp Gln Leu Lys Glu Lys Ile Lys Asp Leu
2375 2380 2385
Glu Thr Gln Leu Lys Met Ser Asp Leu Glu Lys Gln His Leu Lys
2390 2395 2400
Glu Glu Ile Lys Lys Leu Lys Lys Glu Leu Glu Asn Phe Asp Pro
2405 2410 2415
Ser Phe Phe Glu Glu Ile Glu Asp Leu Lys Tyr Asn Tyr Lys Glu
2420 2425 2430
Glu Val Lys Lys Asn Ile Leu Leu Glu Glu Lys Val Lys Lys Leu
2435 2440 2445
Ser Glu Gln Leu Gly Val Glu Leu Thr Ser Pro Val Ala Ala Ser
2450 2455 2460
Glu Glu Phe Glu Asp Glu Glu Glu Ser Pro Val Asn Phe Pro Ile
2465 2470 2475
Tyr

<210> 62
<211> 7972
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 62

atttgaagtc ctcgttccac gccttctcat catcctgaac accgagctct gggactccgg 60
cggagaatct aaacgtaaag catcacccac ggtcgtgaac tgtaggctct cctggcatcc 120
gggatcttat tctggccttg gcggagttgg ggatggtgtc gcctagcagc cgctgccgct 180
ttggcttgct cgggaccatt tggctggacc cagagtccgc gtggaaccgc gatagggatc 240
tgtcagggcc cgcggccggg tccagcttgg tggttgcggt agtgagaggc ctccgctggt 300
tgccaggctt ggtctagagg tggagcacag tgaaagaatt caagatgcca cctaatataa 360
actggaaaga aataatgaaa gttgacccag atgacctgcc ccgtcaagaa gaactggcag 420
ataatttatt gatttcctta tccaaggtgg aagtaaatga gctaaaaagt gaaaagcaag 480
aaaatgtgat acacctttc agaattactc agtcactaat gaagatgaaa gctcaagaag 540
tggagctggc tttggaagaa gtagaaaaag ctggagaaga acaagcaaaa tttgaaaatc 600
aattaaaaac taaagtaatg aaactggaaa atgaactgga gatggctcag cagtctgcag 660
gtggacgaga tactcggttt ttacgtaatg aaatttgcca acttgaaaaa caattagaac 720
aaaaagatag agaattggag gacatggaaa aggagttgga gaaagagaag aaagttaatg 780
agcaattggc tcttcgaaat gaggaggcag aaaatgaaaa cagcaaatta agaagagaga 840
acaaacgtct aaagaaaaag aatgaacaac tttgtcagga tattattgac taccagaaac 900

aaatagattc acagaaagaa acacttttat caagaagagg ggaagacagt gactaccgat 960
cacagttgtc taaaaaaaac tatgagctta tccaatatct tgatgaaatt cagactttaa 1020
cagaagctaa tgagaaaatt gaagttcaga atcaagaaat gagaaaaaat ttagaagagt 1080
ctgtacagga aatggagaag atgactgatg aatataatag aatgaaagct attgtgcatc 1140
agacagataa tgtaatagat cagttaaaaa aagaaaacga tcattatcaa cttcaagtgc 1200
aggagcttac agatcttctg aaatcaaaaa atgaagaaga tgatccaatt atggtagctg 1260
tcaatgcaaa agtagaagaa tggaagctaa ttttgtcttc taaagatgat gaaattattg 1320
agtatcagca aatgttacat aacctaaggg agaaacttaa gaatgctcag cttgatgctg 1380
ataaaagtaa tgttatggct ctacagcagg gtatacagga acgagacagt caaattaaga 1440
tgctcaccga acaagtagaa caatatacaa aagaaatgga aaagaatact tgtattattg 1500
aagatttgaa aaatgagctc caaagaaaca aaggtgcttc aacccttttct caacagactc 1560
atatgaaaat tcagtcaacg ttagacattt taaaagagaa aactaaagag gctgagagaa 1620
cagctgaact ggctgaggct gatgctaggg aaaaggataa agaattagtt gaggctctga 1680
agaggttaaa agattatgaa tcgggagtat atggtttaga agatgctgtc gttgaaataa 1740
agaattgtaa aaaccaaatt aaaataagag atcgagagat tgaaatatta acaaaggaaa 1800
tcaataaact tgaattgaag atcagtgatt tccttgatga aaatgaggca cttagagagc 1860
gtgtgggcct tgaaccaaag acaatgattg atttaactga atttagaaat agcaaacact 1920
taaaacagca gcagtacaga gctgaaaacc agattctttt gaaagagatt gaaagtctag 1980
aggaagaacg acttgatctg aaaaaaaaaa ttcgtcaaat ggctcaagaa agaggaaaaa 2040
gaagtgcaac ttcaggatta accactgagg acctgaacct aactgaaaac atttctcaag 2100
gagatagaat aagtgaaaga aaattggatt tattgagcct caaaaatatg agtgaagcac 2160
aatcaaagaa tgaatttctt tcaagagaac taattgaaaa agaaagagat ttagaaagga 2220
gtaggacagt gatagccaaa tttcagaata aattaaaaga attagttgaa gaaaataagc 2280
aacttgaaga aggtatgaaa gaaatattgc aagcaattaa ggaaatgcag aaagatcctg 2340
atgttaaagg aggagaaaca tctctaatta tccctagcct tgaaagacta gttaatgcta 2400
tagaatcaaa gaatgcagaa ggaatctttg atgcgagtct gcatttgaaa gcccaagttg 2460
atcagcttac cggaagaaat gaagaattaa gacaggagct cagggaatct cggaaagagg 2520
ctataaatta ttcacagcag ttggcaaaag ctaatttaaa gatagaccat cttgaaaaag 2580
aaactagtct tttacgacaa tcagaaggat cgaatgttgt ttttaaagga attgacttac 2640
ctgatgggat agcaccatct agtgccagta tcattaattc tcagaatgaa tatttaatac 2700
atttgttaca ggaactagaa aataaagaaa aaaagttaaa gaatttagaa gattctcttg 2760
aagattacaa cagaaaattt gctgtaattc gtcatcaaca aagtttgttg tataaagaat 2820
acctaagtga aaaggagacc tggaaaacag aatctaaaac aataaaagag gaaaagagaa 2880
aacttgagga tcaagtccaa caagatgcta taaaagtaaa agaatataat aatttgctca 2940
atgctcttca gatggattcg gatgaaatga aaaaaatact tgcagaaaat agtaggaaaa 3000
ttactgtttt gcaagtgaat gaaaaatcac ttataaggca atatacaacc ttagtagaat 3060
tggagcgaca acttagaaaa gaaaatgaga agcaaaagaa tgaattgttg tcaatggagg 3120
ctgaagtttg tgaaaaaatt gggtgtttgc aaagatttaa ggaaatggcc attttcaaga 3180
ttgcagctct ccaaaaagtt gtagataata gtgtttcttt gtctgaacta gaactggcta 3240

ataaacagta caatgaactg actgctaagt acagggacat cttgcaaaaa gataatatgc 3300
ttgttcaaag aacaagtaac ttggaacacc tggagtgtga aaacatctcc ttaaaagaac 3360
aagtggagtc tataaataaa gaactggaga ttaccaagga aaaacttcac actattgaac 3420
aagcctggga acaggaaact aaattaggta atgaatctag catggataag gcaaagaaat 3480
caataaccaa cagtgacatt gtttccattt caaaaaaaat aactatgctg gaaatgaagg 3540
aattaaatga aaggcagcgg gctgaacatt gtcaaaaaat gtatgaacac ttacggactt 3600
cgttaaagca aatggaggaa cgtaattttg aattggaaac caaatttgct gagcttacca 3660
aaatcaattt ggatgcacag aaggtggaac agatgttaag agatgaatta gctgatagtg 3720
tgagcaaggc agtaagtgat gctgataggc aacggattct agaattagag aagaatgaaa 3780
tggaactaaa agttgaagtg tcaaaactga gagagatttc tgatattgcc agaagacaag 3840
ttgaaatttt gaatgcacaa caacaatcta gggacaagga agtagagtcc ctcagaatgc 3900
aactgctaga ctatcaggca cagtctgatg aaaagtcgct cattgccaag ttgcaccaac 3960
ataatgtctc tcttcaactg agtgaggcta ctgctcttgg taagttggag tcaattacat 4020
ctaaactgca gaagatggag gcctacaact tgcgcttaga gcagaaactt gatgaaaaag 4080
aacaggctct ctattatgct cgtttggagg gaagaaacag agcaaaacat ctgcgccaaa 4140
caattcagtc tctacgacga cagtttagtg gagctttacc cttggcacaa caggaaaagt 4200
tctccaaaac aatgattcaa ctacaaaatg acaaacttaa gataatgcaa gaaatgaaaa 4260
attctcaaca agaacataga aatatggaga acaaaacatt ggagatggaa ttaaaattaa 4320
agggcctgga agagttaata agcactttaa aggataccaa aggagcccaa aaggtaatca 4380
actggcatat gaaaatagaa gaacttcgtc ttcaagaact taaactaaat cgggaattag 4440
tcaaggataa agaagaaata aaatatttga ataacataat ttctgaatat gaacgtacaa 4500
tcagcagtct tgaagaagaa attgtgcaac agaacaagtt tcatgaagaa agacaaatgg 4560
cctgggatca aagagaagtt gacctggaac gccaactaga cattttgac cgtcagcaaa 4620
atgaaatact aaatgcggca caaaagtttg aagaagctac aggatcaatc cctgacccta 4680
gtttgcccct tccaaatcaa cttgagatcg ctctaaggaa aattaaggag aacattcgaa 4740
taattctaga aacacgggca acttgcaaat cactagaaga gaaactaaaa gagaaagaat 4800
ctgctttaag gttagcagaa caaaatatac tgtcaagaga caaagtaatc aatgaactga 4860
ggcttcgatt gcctgccact gcagaaagag aaaagctcat agctgagcta ggcagaaaag 4920
agatggaacc aaaatctcac cacacattga aaattgctca tcaaaccatt gcaaacatgc 4980
aagcaaggtt aaatcaaaaa gaagaagtat taaagaagta tcaacgtctt ctagaaaaag 5040
ccagagagga gcaaagagaa attgtgaaga aacatgagga agaccttcat attcttcatc 5100
acagattaga actacaggct gatagttcac taaataaatt caaacaaacg gcttgggatt 5160
taatgaaaca gtctcccact ccagttccta ccaacaagca ttttattcgt ctggctgaga 5220
tggaacagac agtagcagaa caagatgact ctctttcctc actcttggtc aaactaaaga 5280
aagtatcaca agatttggag agacaaagag aaatcactga attaaaagta aaagaatttg 5340
aaaatatcaa attacagctt caagaaaacc atgaagatga agtgaaaaaa gtaaaagcgg 5400
aagtagagga tttaaagtat cttctggacc agtcacaaaa ggagtcacag tgtttaaaat 5460
ctgaacttca ggctcaaaaa gaagcaaatt caagagctcc aacaactaca atgagaaatc 5520
tagtagaacg gctaaagagc caattagcct tgaaggagaa acaacagaaa gcacttagtc 5580

gggcactttt agaactccgg gcagaaatga cagcagctgc tgaagaacgt attatttctg 5640
caacttctca aaaagaggcc catctcaatg ttcaacaaat cgttgatcga catactagag 5700
agctaaagac acaagttgaa gatttaaatg aaaatctttt aaaattgaaa gaagcactta 5760
aaacaagtaa aaacagagaa aactcactaa ctgataattt gaatgactta aataatgaac 5820
tgcaaaagaa acaaaaagcc tataataaaa tacttagaga gaaagaggaa attgatcaag 5880
agaatgatga actgaaaagg caaattaaaa gactaaccag tggattacag ggcaaacccc 5940
tgacagataa taaacaaagt ctaattgaag aactccaaag gaaagttaaa aaactagaga 6000
accaattaga gggaaaggtg gaggaagtag acctaaaacc tatgaaagaa aagaatgcta 6060
aagaagaatt aattaggtgg gaagaaggta aaaagtggca agccaaaata gaaggaattc 6120
gaaacaagtt aaaagagaaa gaggggggaag tctttacttt aacaaagcag ttgaatactt 6180
tgaaggatct ttttgccaaa gccgataaag agaaacttac tttgcagagg aaactaaaaa 6240
caactggcat gactgttgat caggttttgg gaatacgagc tttggagtca gaaaaagaat 6300
tggaagaatt aaaaaagaga aatcttgact tagaaaatga tatattgtat atgagggccc 6360
accaagctct tcctcgagat tctgttgtag aagatttaca tttacaaaat agatacctcc 6420
aagaaaaact tcatgcttta gaaaaacagt tttcaaagga tacatattct aagccttcaa 6480
tttcaggaat agagtcagat gatcattgtc agagagaaca ggagcttcag aaggaaaact 6540
tgaagttgtc atctgaaaat attgaactga aatttcagct tgaacaagca aataaagatt 6600
tgccaagatt aaagaatcaa gtcagagatt tgaaggaaat gtgtgaattt cttaagaaag 6660
aaaaagcaga agttcagcgg aaacttggcc atgttagagg gtctggtaga agtggaaaga 6720
caatcccaga actggaaaaa accattggtt taatgaaaaa agtagttgaa aaagtccaga 6780
gagaaaatga acagttgaaa aaagcatcag gaatattgac tagtgaaaaa atggctaata 6840
ttgagcagga aaatgaaaaa ttgaaggctg aattagaaaa acttaaagct catcttgggc 6900
atcagttgag catgcactat gaatccaaga ccaaaggcac agaaaaaatt attgctgaaa 6960
atgaaaggct tcgtaaagaa cttaaaaaag aaactgatgc tgcagagaaa ttacggatag 7020
caaagaataa tttagagata ttaaatgaga agatgacagt tcaactagaa gagactggta 7080
agagattgca gtttgcagaa agcagaggtc cacagcttga aggtgctgac agtaagagct 7140
ggaaatccat tgtggttaca agaatgtatg aaaccaagtt aaaagaattg gaaactgata 7200
ttgccaaaaa aaatcaaagc attactgacc ttaaacagct tgtaaaagaa gcaacagaga 7260
gagaacaaaa agttaacaaa tacaatgaag accttgaaca acagattaag attcttaaac 7320
atgttcctga aggtgctgag acagagcaag gccttaaacg ggagcttcaa gttcttagat 7380
tagctaatca tcagctggat aaagagaaag cagaattaat ccatcagata gaagctaaca 7440
aggaccaaag tggagctgaa agcaccatac ctgatgctga tcaactaaag gaaaaaataa 7500
aagatctaga gacacagctc aaaatgtcag atctagaaaa gcagcatttg aaggaggaaa 7560
taaagaagct gaaaaaagaa ctggaaaatt ttgatccttc atttttgaa gaaattgaag 7620
atcttaagta taattacaag gaagaagtga agaagaatat tctcttagaa gagaaggtaa 7680
aaaaactttc agaacaattg ggagttgaat taactagccc tgttgctgct tctgaagagt 7740
ttgaagatga agaagaaagt cctgttaatt tccccattta ctaaaggtca cctataaact 7800
ttgtttcatt taactattta ttaactttat aagttaaata tacttggaaa taagcagttc 7860
tccgaactgt agtatttcct tctcactacc ttgtaccttt atacttagat tggaattctt 7920

aataaataaa attatatgaa attttcaact tattaaaaaa aaaaaaaaaa aa 7972

<210> 63

<211> 599

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 63

Met Glu Gly Pro Leu Thr Pro Pro Pro Leu Gln Gly Gly Gly Ala Ala
1 5 10 15
Ala Val Pro Glu Pro Gly Ala Arg Gln His Pro Gly His Glu Thr Ala
20 25 30
Ala Gln Arg Tyr Ser Ala Arg Leu Leu Gln Ala Gly Tyr Glu Pro Glu
35 40 45
Ser Pro Arg Leu Asp Leu Ala Thr His Pro Thr Thr Pro Arg Ser Glu
50 55 60
Leu Ser Ser Val Val Leu Leu Ala Gly Val Gly Val Gln Met Asp Arg
65 70 75 80
Leu Arg Arg Ala Ser Met Ala Asp Tyr Leu Ile Ser Gly Gly Thr Gly
85 90 95
Tyr Val Pro Glu Asp Gly Leu Thr Ala Gln Gln Leu Phe Ala Ser Ala
100 105 110
Asp Gly Leu Thr Tyr Asn Asp Phe Leu Ile Leu Pro Gly Phe Ile Asp
115 120 125
Phe Ile Ala Asp Glu Val Asp Leu Thr Ser Ala Leu Thr Arg Lys Ile
130 135 140
Thr Leu Lys Thr Pro Leu Ile Ser Ser Pro Met Asp Thr Val Thr Glu
145 150 155 160
Ala Asp Met Ala Ile Ala Met Ala Leu Met Gly Gly Ile Gly Phe Ile
165 170 175
His His Asn Cys Thr Pro Glu Phe Gln Ala Asn Glu Val Arg Lys Val
180 185 190
Lys Lys Phe Glu Gln Gly Phe Ile Thr Asp Pro Val Val Leu Ser Pro
195 200 205
Ser His Thr Val Gly Asp Val Leu Glu Ala Lys Met Arg His Gly Phe
210 215 220
Ser Gly Ile Pro Ile Thr Glu Thr Gly Thr Met Gly Ser Lys Leu Val
225 230 235 240
Gly Ile Val Thr Ser Arg Asp Ile Asp Phe Leu Ala Glu Lys Asp His

245 250 255
Thr Thr Leu Leu Ser Glu Val Met Thr Pro Arg Ile Glu Leu Val Val
260 265 270
Ala Pro Ala Gly Val Thr Leu Lys Glu Ala Asn Glu Ile Leu Gln Arg
275 280 285
Ser Lys Lys Gly Lys Leu Pro Ile Val Asn Asp Cys Asp Glu Leu Val
290 295 300
Ala Ile Ile Ala Arg Thr Asp Leu Lys Lys Asn Arg Asp Tyr Pro Leu
305 310 315 320
Ala Ser Lys Asp Ser Gln Lys Gln Leu Leu Cys Gly Ala Ala Val Gly
325 330 335
Thr Arg Glu Asp Asp Lys Tyr Arg Leu Asp Leu Leu Thr Gln Ala Gly
340 345 350
Val Asp Val Ile Val Leu Asp Ser Ser Gln Gly Asn Ser Val Tyr Gln
355 360 365
Ile Ala Met Val His Tyr Ile Lys Gln Lys Tyr Pro His Leu Gln Val
370 375 380
Ile Gly Gly Asn Val Val Thr Ala Ala Gln Ala Lys Asn Leu Ile Asp
385 390 395 400
Ala Gly Val Asp Gly Leu Arg Val Gly Met Gly Cys Gly Ser Ile Cys
405 410 415
Ile Thr Gln Glu Val Met Ala Cys Gly Arg Pro Gln Gly Thr Ala Val
420 425 430
Tyr Lys Val Ala Glu Tyr Ala Arg Arg Phe Gly Val Pro Ile Ile Ala
435 440 445
Asp Gly Gly Ile Gln Thr Val Gly His Val Val Lys Ala Leu Ala Leu
450 455 460
Gly Ala Ser Thr Val Met Met Gly Ser Leu Leu Ala Ala Thr Thr Glu
465 470 475 480
Ala Pro Gly Glu Tyr Phe Phe Ser Asp Gly Val Arg Leu Lys Lys Tyr
485 490 495
Arg Gly Met Gly Ser Leu Asp Ala Met Glu Lys Ser Ser Ser Ser Gln
500 505 510
Lys Arg Tyr Phe Ser Glu Gly Asp Lys Val Lys Ile Ala Gln Gly Val
515 520 525
Ser Gly Ser Ile Gln Asp Lys Gly Ser Ile Gln Lys Phe Val Pro Tyr
530 535 540
Leu Ile Ala Gly Ile Gln His Gly Cys Gln Asp Ile Gly Ala Arg Ser
545 550 555 560

Leu Ser Val Leu Arg Ser Met Met Tyr Ser Gly Glu Leu Lys Phe Glu
565 570 575
Lys Arg Thr Met Ser Ala Gln Ile Glu Gly Gly Val His Gly Leu His
580 585 590
Ser Tyr Glu Lys Arg Leu Tyr
595

<210> 64
<211> 2625
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 64

actcgaccgg gctgcgctca ctgcccagcc ggggccccgg gagcctccag gctcccgccc 60
gccctgagct gcggcctccg catggagggg ccactcactc caccaccgct gcagggaggc 120
ggagccgccg ctgttccgga gcccggagcc cggcaacacc cgggacacga gacggcggcg 180
cagcggtaca gcgcccgact gctgcaggcc ggctacgagc ccgagagccc tagattggac 240
ctcgctacac acccgacgac accccgttca gaactatctt cagtggtctt actggcaggt 300
gttggtgtcc agatggatcg ccttcgcagg gctagcatgg cggactacct gatcagcggc 360
ggcaccggct acgtgcccga ggatgggctc accgcgcagc agctcttcgc cagcgccgac 420
ggcctcacct acaacgactt cctgattctc ccaggattca tagacttcat agctgatgag 480
gtggacctga cctcagccct gacccggaag atcacgctga agacgccact gatctcctcc 540
cccatggaca ctgtgacaga ggctgacatg gccattgcca tggctctgat gggaggtatt 600
ggtttcattc accacaactg caccccagag ttccaggcca acgaggtgcg gaaggtcaag 660
aagtttgaac agggcttcat cacggaccct gtggtgctga gcccctcgca cactgtgggc 720
gatgtgctgg aggccaagat gcggcatggc ttctctggca tccccatcac tgagacgggc 780
accatgggca gcaagctggt gggcatcgtc acctcccgag acatcgactt tcttgctgag 840
aaggaccaca ccaccctcct cagtgaggtg atgacgccaa ggattgaact ggtggtggct 900
ccagcaggtg tgacgttgaa agaggcaaat gagatcctgc agcgtagcaa gaaagggaag 960
ctgcctatcg tcaatgattg cgatgagctg gtggccatca tcgcccgcac cgacctgaag 1020
aagaaccgag actaccctct ggcctccaag gattcccaga agcagctgct ctgtggggca 1080
gctgtgggca cccgtgagga tgacaaatac cgtctggacc tgctcaccca ggcgggcgtc 1140
gacgtcatag tcttggactc gtcccaaggg aattcggtgt atcagatcgc catggtgcat 1200
tacatcaaac agaagtaccc ccacctccag gtgattgggg ggaacgtggt gacagcagcc 1260
caggccaaga acctgattga tgctggtgtg gacgggctgc gcgtgggcat gggctgcggc 1320
tccatctgca tcacccagga agtgatggcc tgtggtcggc cccagggcac tgctgtgtac 1380
aaggtggctg agtatgcccg gcgctttggt gtgcccatca tagccgatgg cggcatccag 1440
accgtgggac acgtggtcaa ggccctggcc cttggagcct ccacagtgat gatgggctcc 1500
ctgctggccg ccactacgga ggcccctggc gagtacttct tctcagacgg ggtgcggctc 1560

aagaagtacc ggggcatggg ctcactggat gccatggaga agagcagcag cagccagaaa 1620
cgatacttca gcgaggggga taaagtgaag atcgcgcagg gtgtctcggg ctccatccag 1680
gacaaaggat ccattcagaa gttcgtgccc tacctcatag caggcatcca acacggctgc 1740
caggatatcg gggcccgcag cctgtctgtc cttcggtcca tgatgtactc aggagagctc 1800
aagtttgaga agcggaccat gtcggcccag attgagggtg gtgtccatgg cctgcactct 1860
tacgaaaagc ggctgtactg aggacagcgg tggaggccga ggtggtggag gggatgcacc 1920
ccagtgtcca cttttgggca cagcctccct ccataactga gtggtccaca gatttgcact 1980
acgggttctc cagctccttt ccaggcagag aggaggggag gtcctgaggg gactgctgcc 2040
cctcactcgg catcccctgc agagtcagga ctgctcccgg ggccaggctg ccctgggagc 2100
ccccctccga gcccagccag ccaggctctc aggccctgcg cctgcctcag gtctttcttg 2160
ctgcagcctg ctccagcctg gccccccaccc caggggcagg cggcccctcc tggcttctcc 2220
tgtagggcac ctccctgccc ctagcctccc aggaaatggt gctctcctgg ccctgcctct 2280
ggcccttccc gggccgctgc ccctcagcca tgtggcactt ctgagctcct gacctaggcc 2340
aaggggaggt ctctgccccc ttccccggcc ctgggctacc cttgggtcct gctcctcagg 2400
ccgctcccct gtccctggcc atgggtagga gactgccctg gtcatggccg cctgcctgtc 2460
attcctgact caccaccgtc cccaggtgaa ccattcctcc cttctcctca gctgcagtcg 2520
aaggctttaa ctttgcacac ttgggatcac agttgcgtca ttgtgtatta aatacttgga 2580
ataaatcaag caggtctcaa cgcctccact aaaaaaaaaa aaaaa 2625

<210> 65
<211> 384
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 65

Met Asp Ala Ala Leu Leu Leu Asn Val Glu Gly Val Lys Lys Thr Ile
1 5 10 15
Leu His Gly Gly Thr Gly Glu Leu Pro Asn Phe Ile Thr Gly Ser Arg
20 25 30
Val Ile Phe His Phe Arg Thr Met Lys Cys Asp Glu Glu Arg Thr Val
35 40 45
Ile Asp Asp Ser Arg Gln Val Gly Gln Pro Met His Ile Ile Ile Gly
50 55 60
Asn Met Phe Lys Leu Glu Val Trp Glu Ile Leu Leu Thr Ser Met Arg
65 70 75 80
Val His Glu Val Ala Glu Phe Trp Cys Asp Thr Ile His Thr Gly Val
85 90 95
Tyr Pro Ile Leu Ser Arg Ser Leu Arg Gln Met Ala Gln Gly Lys Asp
100 105 110

Pro Thr Glu Trp His Val His Thr Cys Gly Leu Ala Asn Met Phe Ala
115 120 125
Tyr His Thr Leu Gly Tyr Glu Asp Leu Asp Glu Leu Gln Lys Glu Pro
130 135 140
Gln Pro Leu Val Phe Val Ile Glu Leu Leu Gln Val Asp Ala Pro Ser
145 150 155 160
Asp Tyr Gln Arg Glu Thr Trp Asn Leu Ser Asn His Glu Lys Met Lys
165 170 175
Ala Val Pro Val Leu His Gly Glu Gly Asn Arg Leu Phe Lys Leu Gly
180 185 190
Arg Tyr Glu Glu Ala Ser Ser Lys Tyr Gln Glu Ala Ile Ile Cys Leu
195 200 205
Arg Asn Leu Gln Thr Lys Glu Lys Pro Trp Glu Val Gln Trp Leu Lys
210 215 220
Leu Glu Lys Met Ile Asn Thr Leu Ile Leu Asn Tyr Cys Gln Cys Leu
225 230 235 240
Leu Lys Lys Glu Glu Tyr Tyr Glu Val Leu Glu His Thr Ser Asp Ile
245 250 255
Leu Arg His His Pro Gly Ile Val Lys Ala Tyr Tyr Val Arg Ala Arg
260 265 270
Ala His Ala Glu Val Trp Asn Glu Ala Glu Ala Lys Ala Asp Leu Gln
275 280 285
Lys Val Leu Glu Leu Glu Pro Ser Met Gln Lys Ala Val Arg Arg Glu
290 295 300
Leu Arg Leu Leu Glu Asn Arg Met Ala Glu Lys Gln Glu Glu Glu Arg
305 310 315 320
Leu Arg Cys Arg Asn Met Leu Ser Gln Gly Ala Thr Gln Pro Pro Ala
325 330 335
Glu Pro Pro Thr Glu Pro Pro Ala Gln Ser Ser Thr Glu Pro Pro Ala
340 345 350
Glu Pro Pro Thr Ala Pro Ser Ala Glu Leu Ser Ala Gly Pro Pro Ala
355 360 365
Glu Pro Ala Thr Glu Pro Pro Pro Ser Pro Gly His Ser Leu Gln His
370 375 380

<210> 66
<211> 316
<212> PRT
<213> 人工序列
<220>

<223> 合成的

<400> 66

Met Leu Val Thr Leu Gly Leu Leu Thr Ser Phe Phe Ser Phe Leu Tyr
1 5 10 15
Met Val Ala Pro Ser Ile Arg Lys Phe Phe Ala Gly Gly Val Cys Arg
20 25 30
Thr Asn Val Gln Leu Pro Gly Lys Val Val Val Ile Thr Gly Ala Asn
35 40 45
Thr Gly Ile Gly Lys Glu Thr Ala Arg Glu Leu Ala Ser Arg Gly Ala
50 55 60
Arg Val Tyr Ile Ala Cys Arg Asp Val Leu Lys Gly Glu Ser Ala Ala
65 70 75 80
Ser Glu Ile Arg Val Asp Thr Lys Asn Ser Gln Val Leu Val Arg Lys
85 90 95
Leu Asp Leu Ser Asp Thr Lys Ser Ile Arg Ala Phe Ala Glu Gly Phe
100 105 110
Leu Ala Glu Glu Lys Gln Leu His Ile Leu Ile Asn Asn Ala Gly Val
115 120 125
Met Met Cys Pro Tyr Ser Lys Thr Ala Asp Gly Phe Glu Thr His Leu
130 135 140
Gly Val Asn His Leu Gly His Phe Leu Leu Thr Tyr Leu Leu Leu Glu
145 150 155 160
Arg Leu Lys Val Ser Ala Pro Ala Arg Val Val Asn Val Ser Ser Val
165 170 175
Ala His His Ile Gly Lys Ile Pro Phe His Asp Leu Gln Ser Glu Lys
180 185 190
Arg Tyr Ser Arg Gly Phe Ala Tyr Cys His Ser Lys Leu Ala Asn Val
195 200 205
Leu Phe Thr Arg Glu Leu Ala Lys Arg Leu Gln Gly Thr Gly Val Thr
210 215 220
Thr Tyr Ala Val His Pro Gly Val Val Arg Ser Glu Leu Val Arg His
225 230 235 240
Ser Ser Leu Leu Cys Leu Leu Trp Arg Leu Phe Ser Pro Phe Val Lys
245 250 255
Thr Ala Arg Glu Gly Ala Gln Thr Ser Leu His Cys Ala Leu Ala Glu
260 265 270
Gly Leu Glu Pro Leu Ser Gly Lys Tyr Phe Ser Asp Cys Lys Arg Thr
275 280 285
Trp Val Ser Pro Arg Ala Arg Asn Asn Lys Thr Ala Glu Arg Leu Trp

290 295 300
Asn Val Ser Cys Glu Leu Leu Gly Ile Arg Trp Glu
305 310 315
<210> 67
<211> 697
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 67
Met Gly Glu Arg Ala Gly Ser Pro Gly Thr Asp Gln Glu Arg Lys Ala
1 5 10 15
Gly Lys His His Tyr Ser Tyr Leu Ser Asp Phe Glu Thr Pro Gln Ser
20 25 30
Ser Gly Arg Ser Ser Leu Val Ser Ser Ser Pro Ala Ser Val Arg Arg
35 40 45
Lys Asn Pro Lys Arg Gln Thr Ser Asp Gly Gln Val His His Gln Ala
50 55 60
Pro Arg Lys Pro Ser Pro Lys Gly Leu Pro Asn Arg Lys Gly Val Arg
65 70 75 80
Val Gly Phe Arg Ser Gln Ser Leu Asn Arg Glu Pro Leu Arg Lys Asp
85 90 95
Thr Asp Leu Val Thr Lys Arg Ile Leu Ser Ala Arg Leu Leu Lys Ile
100 105 110
Asn Glu Leu Gln Asn Glu Val Ser Glu Leu Gln Val Lys Leu Ala Glu
115 120 125
Leu Leu Lys Glu Asn Lys Ser Leu Lys Arg Leu Gln Tyr Arg Gln Glu
130 135 140
Lys Ala Leu Asn Lys Phe Glu Asp Ala Glu Asn Glu Ile Ser Gln Leu
145 150 155 160
Ile Phe Arg His Asn Asn Glu Ile Thr Ala Leu Lys Glu Arg Leu Arg
165 170 175
Lys Ser Gln Glu Lys Glu Arg Ala Thr Glu Lys Arg Val Lys Asp Thr
180 185 190
Glu Ser Glu Leu Phe Arg Thr Lys Phe Ser Leu Gln Lys Leu Lys Glu
195 200 205
Ile Ser Glu Ala Arg His Leu Pro Glu Arg Asp Asp Leu Ala Lys Lys
210 215 220
Leu Val Ser Ala Glu Leu Lys Leu Asp Asp Thr Glu Arg Arg Ile Lys

225 230 235 240
Glu Leu Ser Lys Asn Leu Glu Leu Ser Thr Asn Ser Phe Gln Arg Gln
245 250 255
Leu Leu Ala Glu Arg Lys Arg Ala Tyr Glu Ala His Asp Glu Asn Lys
260 265 270
Val Leu Gln Lys Glu Val Gln Arg Leu Tyr His Lys Leu Lys Glu Lys
275 280 285
Glu Arg Glu Leu Asp Ile Lys Asn Ile Tyr Ser Asn Arg Leu Pro Lys
290 295 300
Ser Ser Pro Asn Lys Glu Lys Glu Leu Ala Leu Arg Lys Asn Ala Ala
305 310 315 320
Cys Gln Ser Asp Phe Ala Asp Leu Cys Thr Lys Gly Val Gln Thr Met
325 330 335
Glu Asp Phe Lys Pro Glu Glu Tyr Pro Leu Thr Pro Glu Thr Ile Met
340 345 350
Cys Tyr Glu Asn Lys Trp Glu Glu Pro Gly His Leu Thr Leu Asp Leu
355 360 365
Gln Ser Gln Lys Gln Asp Arg His Gly Glu Ala Gly Ile Leu Asn Pro
370 375 380
Ile Met Glu Arg Glu Glu Lys Phe Val Thr Asp Glu Glu Leu His Val
385 390 395 400
Val Lys Gln Glu Val Glu Lys Leu Glu Asp Glu Trp Glu Arg Glu Glu
405 410 415
Leu Asp Lys Lys Gln Lys Glu Lys Ala Ser Leu Leu Glu Arg Glu Glu
420 425 430
Lys Pro Glu Trp Glu Thr Gly Arg Tyr Gln Leu Gly Met Tyr Pro Ile
435 440 445
Gln Asn Met Asp Lys Leu Gln Gly Glu Glu Glu Glu Arg Leu Lys Arg
450 455 460
Glu Met Leu Leu Ala Lys Leu Asn Glu Ile Asp Arg Glu Leu Gln Asp
465 470 475 480
Ser Arg Asn Leu Lys Tyr Pro Val Leu Pro Leu Leu Pro Asp Phe Glu
485 490 495
Ser Lys Leu His Ser Pro Glu Arg Ser Pro Lys Thr Tyr Arg Phe Ser
500 505 510
Glu Ser Ser Glu Arg Leu Phe Asn Gly His His Leu Gln Asp Ile Ser
515 520 525
Phe Ser Thr Pro Lys Gly Glu Gly Gln Asn Ser Gly Asn Val Arg Ser
530 535 540

Pro Ala Ser Pro Asn Glu Phe Ala Phe Gly Ser Tyr Val Pro Ser Phe
545 550 555 560
Ala Lys Thr Ser Glu Arg Ser Asn Pro Phe Ser Gln Lys Ser Ser Phe
565 570 575
Leu Asp Phe Gln Arg Asn Ser Met Glu Lys Leu Ser Lys Asp Gly Val
580 585 590
Asp Leu Ile Thr Arg Lys Glu Lys Lys Ala Asn Leu Met Glu Gln Leu
595 600 605
Phe Gly Ala Ser Gly Ser Ser Thr Ile Ser Ser Lys Ser Ser Asp Pro
610 615 620
Asn Ser Val Ala Ser Ser Lys Gly Asp Ile Asp Pro Leu Asn Phe Leu
625 630 635 640
Pro Gly Asn Lys Gly Ser Arg Asp Gln Glu His Asp Glu Asp Glu Gly
645 650 655
Phe Phe Leu Ser Glu Gly Arg Ser Phe Asn Pro Asn Arg His Arg Leu
660 665 670
Lys His Ala Asp Asp Lys Pro Ala Val Lys Ala Ala Asp Ser Val Glu
675 680 685
Asp Glu Ile Glu Glu Val Ala Leu Arg
690 695

<210> 68
<211> 364
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (65)..(65)
<223> X为T或I
<220>
<221> 尚未归类的特征
<222> (111)..(111)
<223> X为V或I
<220>
<221> 尚未归类的特征
<222> (116)..(116)
<223> X为S或Y
<220>

<221> 尚未归类的特征
<222> (153)..(153)
<223> X为L或M
<220>
<221> 尚未归类的特征
<222> (171)..(171)
<223> X为V或I
<220>
<221> 尚未归类的特征
<222> (174)..(174)
<223> X为V或A
<220>
<221> 尚未归类的特征
<222> (178)..(178)
<223> X为V或I
<220>
<221> 尚未归类的特征
<222> (180)..(180)
<223> X为A或S
<220>
<221> 尚未归类的特征
<222> (230)..(230)
<223> X为I或T
<220>
<221> 尚未归类的特征
<222> (233)..(233)
<223> X为A或S
<220>
<221> 尚未归类的特征
<222> (236)..(236)
<223> X为M或V
<220>
<221> 尚未归类的特征
<222> (274)..(274)
<223> X为I或V
<220>
<221> 尚未归类的特征
<222> (275)..(275)
<223> X为F或L

<220>
<221> 尚未归类的特征
<222> (277)..(277)
<223> X为F或Y
<220>
<221> 尚未归类的特征
<222> (279)..(279)
<223> X为F或V
<220>
<221> 尚未归类的特征
<222> (285)..(285)
<223> X为T或A
<220>
<221> 尚未归类的特征
<222> (298)..(298)
<223> X为P或A
<400> 68

```
Met Ala Gln Gln Trp Ser Leu Gln Arg Leu Ala Gly Arg His Pro Gln
1               5                  10                  15
Asp Ser Tyr Glu Asp Ser Thr Gln Ser Ser Ile Phe Thr Tyr Thr Asn
            20                  25                  30
Ser Asn Ser Thr Arg Gly Pro Phe Glu Gly Pro Asn Tyr His Ile Ala
        35                  40                  45
Pro Arg Trp Val Tyr His Leu Thr Ser Val Trp Met Ile Phe Val Val
    50                  55                  60
Xaa Ala Ser Val Phe Thr Asn Gly Leu Val Leu Ala Ala Thr Met Lys
65                  70                  75                  80
Phe Lys Lys Leu Arg His Pro Leu Asn Trp Ile Leu Val Asn Leu Ala
                85                  90                  95
Val Ala Asp Leu Ala Glu Thr Val Ile Ala Ser Thr Ile Ser Xaa Val
            100                 105                 110
Asn Gln Val Xaa Gly Tyr Phe Val Leu Gly His Pro Met Cys Val Leu
        115                 120                 125
Glu Gly Tyr Thr Val Ser Leu Cys Gly Ile Thr Gly Leu Trp Ser Leu
    130                 135                 140
Ala Ile Ile Ser Trp Glu Arg Trp Xaa Val Val Cys Lys Pro Phe Gly
145                 150                 155                 160
Asn Val Arg Phe Asp Ala Lys Leu Ala Ile Xaa Gly Ile Xaa Phe Ser
                165                 170                 175
```

Trp Xaa Trp Xaa Ala Val Trp Thr Ala Pro Pro Ile Phe Gly Trp Ser
180 185 190
Arg Tyr Trp Pro His Gly Leu Lys Thr Ser Cys Gly Pro Asp Val Phe
195 200 205
Ser Gly Ser Ser Tyr Pro Gly Val Gln Ser Tyr Met Ile Val Leu Met
210 215 220
Val Thr Cys Cys Ile Xaa Pro Leu Xaa Ile Ile Xaa Leu Cys Tyr Leu
225 230 235 240
Gln Val Trp Leu Ala Ile Arg Ala Val Ala Lys Gln Gln Lys Glu Ser
245 250 255
Glu Ser Thr Gln Lys Ala Glu Lys Glu Val Thr Arg Met Val Val Val
260 265 270
Met Xaa Xaa Ala Xaa Cys Xaa Cys Trp Gly Pro Tyr Xaa Phe Phe Ala
275 280 285
Cys Phe Ala Ala Ala Asn Pro Gly Tyr Xaa Phe His Pro Leu Met Ala
290 295 300
Ala Leu Pro Ala Tyr Phe Ala Lys Ser Ala Thr Ile Tyr Asn Pro Val
305 310 315 320
Ile Tyr Val Phe Met Asn Arg Gln Phe Arg Asn Cys Ile Leu Gln Leu
325 330 335
Phe Gly Lys Lys Val Asp Asp Gly Ser Glu Leu Ser Ser Ala Ser Lys
340 345 350
Thr Glu Val Ser Ser Val Ser Ser Val Ser Pro Ala
355 360

<210> 69
<211> 174
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 69
gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60
gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120
caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccgccgc cacc 174
<210> 70
<211> 322
<212> DNA
<213> 人工序列
<220>

<223> 合成的
<400> 70
gatccggttc caggcctcgg ccctaaatag tctccctggg ctttcaagag aaccacatga 60
gaaaggagga ttcgggctct gagcagtttc accacccacc ccccagtctg caaatcctga 120
cccgtgggtc cacctgcccc aaaggcggac gcaggacagt agaagggaac agagaacaca 180
taaacacaga gagggccaca gcggctccca cagtcaccgc caccttcctg gcggggatgg 240
gtggggcgtc tgagtttggt tcccagcaaa tccctctgag ccgcccttgc gggctcgcct 300
caggagcagg ggagcaagag gt 322
<210> 71
<211> 27
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 71
tcatctgtgt cctcggtatc gcctgca 27
<210> 72
<211> 33
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 72
tcaactgtgt cctcgaccca ggtagggcct aac 33
<210> 73
<211> 30
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 73
tcatctgtgt cctcggtatc gcctgcatag 30
<210> 74
<211> 720
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<220>

<221> 尚未归类的特征
<222> (718)..(720)
<223> 可选地缺失
<400> 74
atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180
ctcgtgacca ccttcggcta cggcctgatg tgcttcgccc gctaccccga ccacatgaag 240
cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagct accagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaagtag 720
<210> 75
<211> 239
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 75
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1 5 10 15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
20 25 30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
35 40 45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
50 55 60
Phe Gly Tyr Gly Leu Met Cys Phe Ala Arg Tyr Pro Asp His Met Lys
65 70 75 80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
85 90 95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
100 105 110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
130 135 140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145 150 155 160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165 170 175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180 185 190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
195 200 205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
210 215 220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 76
<211> 30
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 76
ggaggtggag gttctggtgg aggaggttcc 30
<210> 77
<211> 517
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (515)..(517)
<223> 可选地缺失
<400> 77
gtactggccc cacggcctga agacttcatg cggcccagac gtgttcagcg gcagctcgta 60
ccccggggtg cagtcttaca tgattgtcct catggtcacc tgctgcatca tcccactcgc 120
tatcatcatg ctctgctacc tccaagtgtg gctggccatc cgagcggtgg caaagcagca 180
gaaagagtct gaatccaccc agaaggcaga gaaggaagtg acgcgcatgg tggtggtgat 240
gatctttgcg tactgcgtct gctggggacc ctacaccttc ttcgcatgct ttgctgctgc 300
caaccctggt tacgccttcc acccttttgat ggctgccctg ccggcctact ttgccaaaag 360

tgccactatc tacaaccccg ttatctatgt ctttatgaac cggcagtttc gaaactgcat 420
cttgcagctt ttcgggaaga aggttgacga tggctctgaa ctctccagcg cctccaaaac 480
ggaggtctca tctgtgtcct cggtatcgcc tgcatga 517
<210> 78
<211> 78
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 78
atggtcagct actgggacac cggggtcctg ctgtgcgcgc tgctcagctg tctgcttctc 60
acaggatcta gttccgga 78
<210> 79
<211> 167
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 79
tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacg 167
<210> 80
<211> 145
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 80
gttaatcatt aactacaagg aacccctagt gatggagttg gccactccct ctctgcgcgc 60
tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc 120
ggcctcagtg agcgagcgag cgcgc 145
<210> 81
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 81

tagccatgct ctaggaagat ct 22

<210> 82

<211> 181

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 82

ggggaggagg aggtctaagt cccaggccca attaagagat caggtagtgt agggtttggg 60

agcttttaag gtgaagaggc ccgggctgat cccacaggcc agtataaagc gccgtgaccc 120

tcaggtgacg cgccagggcc ggctgccgtc ggggacaggg ctttccatag ccggtaccat 180

g 181

<210> 83

<211> 202

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 83

ctagaggatc cggtactcga ggaactgaaa aaccagaaag ttaactggta agtttagtct 60

ttttgtcttt tatttcaggt cccggatccg gtggtggtgc aaatcaaaga actgctcctc 120

agtggatgtt gcctttactt ctaggcctgt acggaagtgt tacttctgct ctaaaagctg 180

cggaattgta cccgccgcca cc 202

<210> 84

<211> 1065

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 84

atggcccagc agtggagcct ccaaaggctc gcaggccgcc atccgcagga cagctatgag 60

gacagcaccc agtccagcat cttcacctac accaacagca actccaccag aggccccttc 120

gaaggcccga attaccacat cgctcccaga tgggtgtacc acctcaccag tgtctggatg 180

atctttgtgg tcactgcatc cgtcttcaca aatgggcttg tgctggcggc caccatgaag 240

ttcaagaagc tgcgccaccc gctgaactgg atcctggtga acctggcggt cgctgaccta 300

gcagagaccg tcatcgccag cactatcagc attgtgaacc aggtctctgg ctacttcgtg 360

ctgggccacc ctatgtgtgt cctggagggc tacaccgtct ccctgtgtgg gatcacaggt 420

ctctggtctc tggccatcat ttcctgggag aggtggctgg tggtgtgcaa gccctttggc 480

aatgtgagat ttgatgccaa gctggccatc gtgggcattg ccttctcctg gatctggtct 540

gctgtgtgga cagccccgcc catctttggt tggagcaggt actggcccca cggcctgaag 600
acttcatgcg gcccagacgt gttcagcggc agctcgtacc ccggggtgca gtcttacatg 660
attgtcctca tggtcacctg ctgcatcatc ccactcgcta tcatcatgct ctgctacctc 720
caagtgtggc tggccatccg agcggtggca aagcagcaga aagagtctga atccacccag 780
aaggcagaga aggaagtgac gcgcatggtg gtggtgatga tctttgcgta ctgcgtctgc 840
tggggaccct acaccttctt cgcatgcttt gctgctgcca accctggtta cgccttccac 900
cctttgatgg ctgccctgcc ggcctacttt gccaaaagtg ccactatcta caaccccgtt 960
atctatgtct ttatgaaccg gcagtttcga aactgcatct tgcagctttt cgggaagaag 1020
gttgacgatg gctctgaact ctccagcgcc tccaaaacgg aggtc 1065

<210> 85
<211> 720
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 85

atggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac 120
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc 180
ctcgtgacca ccctgaccta cggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240
cagcacgact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc 300
ttcaaggacg acggcaacta caagacccgc gccgaggtga agttcgaggg cgacaccctg 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac 480
ggcatcaagg tgaacttcaa gatccgccac aacatcgagg acggcagcgt gcagctcgcc 540
gaccactacc agcagaacac ccccatcggc gacggccccg tgctgctgcc cgacaaccac 600
tacctgagca cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatggtc 660
ctgctggagt tcgtgaccgc cgccgggatc actctcggca tggacgagct gtacaaggtc 720

<210> 86
<211> 288
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 86

ggtctgcctc ctacccatcc cgcccaccgg ggctttggcc acctctcctt tccccctcct 60
tctccatccc tgtaaaataa atgtaattta tctttgccaa aaccaacaaa gtcacagagg 120
ctttcactgc agtgtgggac cacctgagcc tctgcgtgtg caggcactgg gtctcgagag 180
ggtgcaaggg ggataaagag gagagagcgc ttcatagact ttaagttttc ccgagcctca 240

tgtctaccga tggcgtgaaa ggatcctggc aaaacagaag tgtgaggc 288

<210> 87

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 87

aagcttatcg ataaggatct tcctagagca tggcta 36

<210> 88

<211> 167

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 88

cgtagataag tagcatggcg ggttaatcat taactacaag gaacccctag tgatggagtt 60

ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa aggtcgcccg 120

acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgca 167

<210> 89

<211> 171

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 89

gggaggagga ggtctaagtc ccaggcccaa ttaagagatc aggtagtgta gggtttggga 60

gcttttaagg tgaagaggcc cgggctgatc ccacaggcca gtataaagcg ccgtgaccct 120

caggtgacgc gccagggccg gctgccgtcg gggacagggc tttccatagc c 171

<210> 90

<211> 205

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 90

ggtaccatgc tagaggatcc ggtactcgag gaactgaaaa accagaaagt taactggtaa 60

gtttagtctt tttgtctttt atttcaggtc ccggatccgg tggtggtgca aatcaaagaa 120

ctgctcctca gtggatgttg cctttacttc taggcctgta cggaagtgtt acttctgctc 180

taaaagctgc ggaattgtac ccgcc 205

<210> 91

<211> 4386

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 91

tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacgtag ccatgctcta 180
ggaagatctg gaggctgagg ggtggggaaa gggcatgggt gtttcatgag gacagagctt 240
ccgtttcatg caatgaaaag agtttggaga cggatggtgg tgactggact atacacttac 300
acacggtagc gatggtacac tttgtattat gtatatttta ccacgatctt tttaaagtgt 360
caaaggcaaa tggccaaatg gttccttgtc ctatagctgt agcagccatc ggctgttagt 420
gacaaagccc ctgagtcaag atgacagcag cccccataac tcctaatcgg ctctcccgcg 480
tggagtcatt taggagtagt cgcattagag acaagtccaa catctaatct tccaccctgg 540
ccagggcccc agctggcagc gagggtggga gactccgggc agagcagagg gcgctgacat 600
tggggcccgg cctggcttgg gtccctctgg cctttcccca ggggccctct ttccttgggg 660
ctttcttggg ccgccactgc tcccgctcct ctccccccat cccacccccc caccccctcg 720
ttcttcatat ccttctctag tgctccctcc actttcatcc acccttctgc aagagtgtgg 780
gaccacaaat gagttttcac ctggcctggg gacacacgtg cccccacagg tgctgagtga 840
ctttctagga cagtaatctg ctttaggcta aaatgggact tgatcttctg ttagccctaa 900
tcatcaatta gcagagccgg tgaaggtgca gaacctaccg cctttccagg cctcctccca 960
cctctgccac ctccactctc cttcctggga tgtgggggct ggcacacgtg tggcccaggg 1020
cattggtggg attgcactga gctgggtcat tagcgtaatc ctggacaagg gcagacaggg 1080
cgagcggagg gccagctccg gggctcaggc aaggctgggg gcttccccca gacaccccac 1140
tcctcctctg ctggaccccc acttcatagg gcacttcgtg ttctcaaagg gcttccaaat 1200
agcatggtgg ccttggatgc ccagggaagc ctcagagttg cttatctccc tctagacaga 1260
aggggaatct cggtcaagag ggagaggtcg ccctgttcaa ggccacccag ccagctcatg 1320
gcggtaatgg gacaaggctg gccagccatc ccaccctcag aagggacccg gtggggcagg 1380
tgatctcaga ggaggctcac ttctgggtct cacattcttg gggaggagga ggtctaagtc 1440
ccaggcccaa ttaagagatc aggtagtgta gggtttggga gcttttaagg tgaagaggcc 1500
cgggctgatc ccacaggcca gtataaagcg ccgtgaccct caggtgacgc gccagggccg 1560
gctgccgtcg gggacagggc tttccatagc cggtaccatg ctagaggatc cggtactcga 1620
ggaactgaaa aaccagaaag ttaactggta agtttagtct ttttgtcttt tatttcaggt 1680
cccggatccg gtggtggtgc aaatcaaaga actgctcctc agtggatgtt gcctttactt 1740
ctaggcctgt acggaagtgt tacttctgct ctaaaagctg cggaattgta cccgccgcca 1800
ccatggccca gcagtggagc ctccaaaggc tcgcaggccg ccatccgcag gacagctatg 1860

aggacagcac ccagtccagc atcttcacct acaccaacag caactccacc agaggcccct 1920
tcgaaggccc gaattaccac atcgctccca gatgggtgta ccacctcacc agtgtctgga 1980
tgatctttgt ggtcactgca tccgtcttca caaatgggct tgtgctggcg gccaccatga 2040
agttcaagaa gctgcgccac ccgctgaact ggatcctggt gaacctggcg gtcgctgacc 2100
tagcagagac cgtcatcgcc agcactatca gcattgtgaa ccaggtctct ggctacttcg 2160
tgctgggcca ccctatgtgt gtcctggagg gctacaccgt ctccctgtgt gggatcacag 2220
gtctctggtc tctggccatc atttcctggg agaggtggct ggtggtgtgc aagccctttg 2280
gcaatgtgag atttgatgcc aagctggcca tcgtgggcat tgccttctcc tggatctggt 2340
ctgctgtgtg gacagccccg cccatctttg gttggagcag gtactggccc cacggcctga 2400
agacttcatg cggcccagac gtgttcagcg gcagctcgta ccccggggtg cagtcttaca 2460
tgattgtcct catggtcacc tgctgcatca tcccactcgc tatcatcatg ctctgctacc 2520
tccaagtgtg gctggccatc cgagcggtgg caaagcagca gaaagagtct gaatccaccc 2580
agaaggcaga gaaggaagtg acgcgcatgg tggtggtgat gatctttgcg tactgcgtct 2640
gctggggacc ctacaccttc ttcgcatgct ttgctgctgc caaccctggt tacgccttcc 2700
acccctttgat ggctgccctg ccggcctact ttgccaaaag tgccactatc tacaaccccg 2760
ttatctatgt ctttatgaac cggcagtttc gaaactgcat cttgcagctt ttcgggaaga 2820
aggttgacga tggctctgaa ctctccagcg cctccaaaac ggaggtccaa ctgtgtcctc 2880
gacccaggta gggcctaaca tggtgagcaa gggcgaggag ctgttcaccg gggtggtgcc 2940
catcctggtc gagctggacg gcgacgtaaa cggccacaag ttcagcgtgt ccggcgaggg 3000
cgagggcgat gccacctacg gcaagctgac cctgaagttc atctgcacca ccggcaagct 3060
gcccgtgccc tggcccaccc tcgtgaccac cctgacctac ggcgtgcagt gcttcagccg 3120
ctaccccgac cacatgaagc agcacgactt cttcaagtcc gccatgcccg aaggctacgt 3180
ccaggagcgc accatcttct tcaaggacga cggcaactac aagacccgcg ccgaggtgaa 3240
gttcgagggc gacaccctgg tgaaccgcat cgagctgaag ggcatcgact tcaaggagga 3300
cggcaacatc ctggggcaca agctggagta caactacaac agccacaacg tctatatcat 3360
ggccgacaag cagaagaacg gcatcaaggt gaacttcaag atccgccaca acatcgagga 3420
cggcagcgtg cagctcgccg accactacca gcagaacacc cccatcggcg acggccccgt 3480
gctgctgccc gacaaccact acctgagcac ccagtccgcc ctgagcaaag accccaacga 3540
gaagcgcgat cacatggtcc tgctggagtt cgtgaccgcc gccgggatca ctctcggcat 3600
ggacgagctg tacaaggtct catctgtgtc ctcggtatcg cctgcataga atcaacctct 3660
ggattacaaa atttgtgaaa gattgactgg tattcttaac tatgttgctc cttttacgct 3720
atgtggatac gctgctttaa tgcctttgta tcatgctatt gcttcccgta tggctttcat 3780
tttctcctcc ttgtataaat cctggttagt tcttgccacg gcggaactca tcgccgcctg 3840
ccttgcccgc tgctggacag gggctcggct gttgggcact gacaattccg tggtgggtct 3900
gcctcctacc catcccgccc accggggctt tggccacctc tcctttcccc ctccttctcc 3960
atccctgtaa aataaatgta atttatcttt gccaaaacca acaaagtcac agaggctttc 4020
actgcagtgt gggaccacct gagcctctgc gtgtgcaggc actgggtctc gagagggtgc 4080
aaggggata aagaggagag agcgcttcat agactttaag ttttcccgag cctcatgtct 4140
accgatggcg tgaaaggatc ctggcaaaac agaagtgtga ggcaagctta tcgataagga 4200

tcttcctaga gcatggctac gtagataagt agcatggcgg gttaatcatt aactacaagg 4260
aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 4320
ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag 4380
cgcgca 4386

<210> 92
<211> 3634
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 92

tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacgtag ccatgctcta 180
ggaagatctg gaggctgagg ggtggggaaa gggcatgggt gtttcatgag gacagagctt 240
ccgtttcatg caatgaaaag agtttggaga cggatggtgg tgactggact atacacttac 300
acacggtagc gatggtacac tttgtattat gtatatttta ccacgatctt tttaaagtgt 360
caaaggcaaa tggccaaatg gttccttgtc ctatagctgt agcagccatc ggctgttagt 420
gacaaagccc ctgagtcaag atgacagcag cccccataac tcctaatcgg ctctcccgcg 480
tggagtcatt taggagtagt cgcattagag acaagtccaa catctaatct tccaccctgg 540
ccagggcccc agctggcagc gagggtggga gactccgggc agagcagagg gcgctgacat 600
tggggcccgg cctggcttgg gtccctctgg cctttcccca ggggccctct ttccttgggg 660
ctttcttggg ccgccactgc tcccgctcct ctccccccat cccacccccc caccccctcg 720
ttcttcatat ccttctctag tgctccctcc actttcatcc acccttctgc aagagtgtgg 780
gaccacaaat gagttttcac ctggcctggg gacacacgtg cccccacagg tgctgagtga 840
ctttctagga cagtaatctg ctttaggcta aaatgggact tgatcttctg ttagccctaa 900
tcatcaatta gcagagccgg tgaaggtgca gaacctaccg cctttccagg cctcctccca 960
cctctgccac ctccactctc cttcctggga tgtgggggct ggcacacgtg tggcccaggg 1020
cattggtggg attgcactga gctgggtcat tagcgtaatc ctggacaagg gcagacaggg 1080
cgagcggagg gccagctccg gggctcaggc aaggctgggg gcttccccca gacaccccac 1140
tcctcctctg ctggaccccc acttcatagg gcacttcgtg ttctcaaagg gcttccaaat 1200
agcatggtgg ccttggatgc ccagggaagc ctcagagttg cttatctccc tctagacaga 1260
aggggaatct cggtcaagag ggagaggtcg ccctgttcaa ggccacccag ccagctcatg 1320
gcggtaatgg gacaaggctg gccagccatc ccaccctcag aagggacccg gtggggcagg 1380
tgatctcaga ggaggctcac ttctgggtct cacattcttg gggaggagga ggtctaagtc 1440
ccaggcccaa ttaagagatc aggtagtgta gggtttggga gcttttaagg tgaagaggcc 1500
cgggctgatc ccacaggcca gtataaagcg ccgtgaccct caggtgacgc gccagggccg 1560
gctgccgtcg gggacagggc tttccatagc cggtaccatg ctagaggatc cggtactcga 1620
ggaactgaaa aaccagaaag ttaactggta agtttagtct ttttgtcttt tatttcaggt 1680

cccggatccg gtggtggtgc aaatcaaaga actgctcctc agtggatgtt gcctttactt 1740
ctaggcctgt acggaagtgt tacttctgct ctaaaagctg cggaattgta cccgccgcca 1800
ccatggccca gcagtggagc ctccaaaggc tcgcaggccg ccatccgcag gacagctatg 1860
aggacagcac ccagtccagc atcttcacct acaccaacag caactccacc agaggcccct 1920
tcgaaggccc gaattaccac atcgctccca gatgggtgta ccacctcacc agtgtctgga 1980
tgatctttgt ggtcactgca tccgtcttca caaatgggct tgtgctggcg gccaccatga 2040
agttcaagaa gctgcgccac ccgctgaact ggatcctggt gaacctggcg gtcgctgacc 2100
tagcagagac cgtcatcgcc agcactatca gcattgtgaa ccaggtctct ggctacttcg 2160
tgctgggcca ccctatgtgt gtcctggagg gctacaccgt ctccctgtgt gggatcacag 2220
gtctctggtc tctggccatc atttcctggg agaggtggct ggtggtctgc aagccctttg 2280
gcaatgtgag atttgatgcc aagctggcca tcgtgggcat tgccttctcc tggatctggt 2340
ctgctgtgtg gacagccccg cccatctttg gttggagcag gtactggccc cacggcctga 2400
agacttcatg cggcccagac gtgttcagcg gcagctcgta ccccggggtg cagtcttaca 2460
tgattgtcct catggtcacc tgctgcatca tcccactcgc tatcatcatg ctctgctacc 2520
tccaagtgtg gctggccatc cgagcggtgg caaagcagca gaaagagtct gaatccaccc 2580
agaaggcaga gaaggaagtg acgcgcatgg tggtggtgat gatctttgcg tactgcgtct 2640
gctggggacc ctacaccttc ttcgcatgct ttgctgctgc caaccctggt tacgccttcc 2700
acccctttgat ggctgccctg ccggcctact ttgccaaaag tgccactatc tacaaccccg 2760
ttatctatgt ctttatgaac cggcagtttc gaaactgcat cttgcagctt ttcgggaaga 2820
aggttgacga tggctctgaa ctctccagcg cctccaaaac ggaggtctca tctgtgtcct 2880
cggtatcgcc tgcatagaat caacctctgg attacaaaat ttgtgaaaga ttgactggta 2940
ttcttaacta tgttgctcct tttacgctat gtggatacgc tgctttaatg cctttgtatc 3000
atgctattgc ttcccgtatg gctttcattt tctcctcctt gtataaatcc tggttagttc 3060
ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt 3120
tgggcactga caattccgtg gtgggtctgc ctcctaccca tcccgcccac cggggctttg 3180
gccacctctc ctttcccccc ccttctccat ccctgtaaaa taaatgtaat ttatctttgc 3240
caaaaccaac aaagtcacag aggctttcac tgcagtgtgg gaccacctga gcctctgcgt 3300
gtgcaggcac tgggtctcga gagggtgcaa gggggataaa gaggagagag cgcttcatag 3360
actttaagtt ttcccgagcc tcatgtctac cgatggcgtg aaaggatcct ggcaaaacag 3420
aagtgtgagg caagcttatc gataaggatc ttcctagagc atggctacgt agataagtag 3480
catggcgggt taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct 3540
ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt 3600
gcccgggcgg cctcagtgag cgagcgagcg cgca 3634

<210> 93
<211> 4663
<212> DNA
<213> 人工序列
<220>
<223> 合成的

<400> 93

tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacgtag ccatgctcta 180
ggaagatctg gaggctgagg ggtggggaaa gggcatgggt gtttcatgag gacagagctt 240
ccgtttcatg caatgaaaag agtttggaga cggatggtgg tgactggact atacacttac 300
acacggtagc gatggtacac tttgtattat gtatatttta ccacgatctt tttaaagtgt 360
caaaggcaaa tggccaaatg gttccttgtc ctatagctgt agcagccatc ggctgttagt 420
gacaaagccc ctgagtcaag atgacagcag cccccataac tcctaatcgg ctctcccgcg 480
tggagtcatt taggagtagt cgcattagag acaagtccaa catctaatct tccaccctgg 540
ccagggcccc agctggcagc gagggtggga gactccgggc agagcagagg gcgctgacat 600
tggggcccgg cctggcttgg gtccctctgg cctttcccca ggggccctct ttccttgggg 660
ctttcttggg ccgccactgc tcccgctcct ctccccccat cccacccccc caccccctcg 720
ttcttcatat ccttctctag tgctccctcc actttcatcc acccttctgc aagagtgtgg 780
gaccacaaat gagttttcac ctggcctggg gacacacgtg cccccacagg tgctgagtga 840
ctttctagga cagtaatctg ctttaggcta aaatgggact tgatcttctg ttagccctaa 900
tcatcaatta gcagagccgg tgaaggtgca gaacctaccg cctttccagg cctcctccca 960
cctctgccac ctccactctc cttcctggga tgtgggggct ggcacacgtg tggcccaggg 1020
cattggtggg attgcactga gctgggtcat tagcgtaatc ctggacaagg gcagacaggg 1080
cgagcggagg gccagctccg gggctcaggc aaggctgggg gcttccccca gacaccccac 1140
tcctcctctg ctggaccccc acttcatagg gcacttcgtg ttctcaaagg gcttccaaat 1200
agcatggtgg ccttggatgc ccagggaagc ctcagagttg cttatctccc tctagacaga 1260
aggggaatct cggtcaagag ggagaggtcg ccctgttcaa ggccacccag ccagctcatg 1320
gcggtaatgg gacaaggctg gccagccatc ccaccctcag aagggacccg gtggggcagg 1380
tgatctcaga ggaggctcac ttctgggtct cacattcttg gggaggagga ggtctaagtc 1440
ccaggcccaa ttaagagatc aggtagtgta gggtttggga gcttttaagg tgaagaggcc 1500
cgggctgatc ccacaggcca gtataaagcg ccgtgaccct caggtgacgc gccagggccg 1560
gctgccgtcg gggacagggc tttccatagc cggtaccatg ctagaggatc cggtactcga 1620
ggaactgaaa aaccagaaag ttaactggta agtttagtct ttttgtcttt tatttcaggt 1680
cccggatccg gtggtggtgc aaatcaaaga actgctcctc agtggatgtt gcctttactt 1740
ctaggcctgt acggaagtgt tacttctgct ctaaaagctg cggaattgta cccgccgcca 1800
ccatggtcag ctactgggac accggggtcc tgctgtgcgc gctgctcagc tgtctgcttc 1860
tcacaggatc tagttccgga agtgataccg gtagaccttt cgtagagatg tacagtgaaa 1920
tccccgaaat tatacacatg actgaaggaa gggagctcgt cattccctgc cgggttacgt 1980
cacctaacat cactgttact ttaaaaaagt ttccacttga cactttgatc cctgatggaa 2040
aacgcataat ctgggacagt agaaagggct tcatcatatc aaatgcaacg tacaaagaaa 2100
tagggcttct gacctgtgaa gcaacagtca atgggcattt gtataagaca aactatctca 2160
cacatcgaca aaccaataca atcatagatg tggttctgag tccgtctcat ggaattgaac 2220
tatctgttgg agaaaagctt gtcttaaatt gtacagcaag aactgaacta aatgtgggga 2280

ttgacttcaa ctgggaatac ccttcttcga agcatcagca taagaaactt gtaaaccgag 2340
acctaaaaac ccagtctggg agtgagatga agaaattttt gagcacctta actatagatg 2400
gtgtaacccg gagtgaccaa ggattgtaca cctgtgcagc atccagtggg ctgatgacca 2460
agaagaacag cacatttgtc agggtccatg aaaaggacaa aactcacaca tgcccaccgt 2520
gcccagcacc tgaactcctg gggggaccgt cagtcttcct cttcccccca aaacccaagg 2580
acaccctcat gatctcccgg acccctgagg tcacatgcgt ggtggtggac gtgagccacg 2640
aagaccctga ggtcaagttc aactggtacg tggacggcgt ggaggtgcat aatgccaaga 2700
caaagccgcg ggaggagcag tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc 2760
tgcaccagga ctggctgaat ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc 2820
cagccccat cgagaaaacc atctccaaag ccaaagggca gccccgagaa ccacaggtgt 2880
acaccctgcc cccatcccgg gatgagctga ccaagaacca ggtcagcctg acctgcctgg 2940
tcaaaggctt ctatcccagc gacatcgccg tggagtggga gagcaatggg cagccggaga 3000
acaactacaa gaccacgcct cccgtgctgg actccgacgg ctccttcttc ctctacagca 3060
agctcaccgt ggacaagagc aggtggcagc aggggaacgt cttctcatgc tccgtgatgc 3120
atgaggctct gcacaaccac tacacgcaga agagcctctc cctgtctccg ggtaaaggag 3180
gtggaggttc tggtggagga ggttccatgg tgagcaaggg cgaggagctg ttcaccgggg 3240
tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 3300
gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 3360
gcaagctgcc cgtgccctgg cccaccctcg tgaccacctt cggctacggc ctgatgtgct 3420
tcgcccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 3480
gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 3540
aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 3600
aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 3660
atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 3720
tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 3780
gccccgtgct gctgcccgac aaccactacc tgagctacca gtccgccctg agcaaagacc 3840
ccaacgagaa gcgcgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 3900
tcggcatgga cgagctgtac aagtagaatc aacctctgga ttacaaaatt tgtgaaagat 3960
tgactggtat tcttaactat gttgctcctt ttacgctatg tggatacgct gctttaatgc 4020
ctttgtatca tgctattgct tcccgtatgg ctttcatttt ctcctccttg tataaatcct 4080
ggttagttct tgccacggcg gaactcatcg ccgcctgcct tgcccgctgc tggacagggg 4140
ctcggctgtt gggcactgac aattccgtgg tgggtctgcc tcctacccat cccgcccacc 4200
ggggctttgg ccacctctcc tttccccctc cttctccatc cctgtaaaat aaatgtaatt 4260
tatctttgcc aaaaccaaca aagtcacaga ggctttcact gcagtgtggg accacctgag 4320
cctctgcgtg tgcaggcact gggtctcgag agggtgcaag ggggataaag aggagagagc 4380
gcttcataga ctttaagttt tcccgagcct catgtctacc gatggcgtga aaggatcctg 4440
gcaaaacaga agtgtgaggc aagcttatcg ataaggatct tcctagagca tggctacgta 4500
gataagtagc atggcgggtt aatcattaac tacaaggaac ccctagtgat ggagttggcc 4560
actccctctc tgcgcgctcg ctcgctcact gaggccgggc gaccaaaggt cgcccgacgc 4620

ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc gca 4663

<210> 94

<211> 3916

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 94

tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacgtag ccatgctcta 180
ggaagatctg gaggctgagg ggtggggaaa gggcatgggt gtttcatgag gacagagctt 240
ccgtttcatg caatgaaaag agtttggaga cggatggtgg tgactggact atacacttac 300
acacggtagc gatggtacac tttgtattat gtatatttta ccacgatctt tttaaagtgt 360
caaaggcaaa tggccaaatg gttccttgtc ctatagctgt agcagccatc ggctgttagt 420
gacaaagccc ctgagtcaag atgacagcag cccccataac tcctaatcgg ctctcccgcg 480
tggagtcatt taggagtagt cgcattagag acaagtccaa catctaatct tccaccctgg 540
ccagggcccc agctggcagc gagggtggga gactccgggc agagcagagg gcgctgacat 600
tggggcccgg cctggcttgg gtccctctgg cctttcccca ggggccctct ttccttgggg 660
ctttcttggg ccgccactgc tcccgctcct ctccccccat cccacccccc caccccctcg 720
ttcttcatat ccttctctag tgctccctcc actttcatcc acccttctgc aagagtgtgg 780
gaccacaaat gagttttcac ctggcctggg gacacacgtg cccccacagg tgctgagtga 840
ctttctagga cagtaatctg ctttaggcta aaatgggact tgatcttctg ttagccctaa 900
tcatcaatta gcagagccgg tgaaggtgca gaacctaccg cctttccagg cctcctccca 960
cctctgccac ctccactctc cttcctggga tgtgggggct ggcacacgtg tggcccaggg 1020
cattggtggg attgcactga gctgggtcat tagcgtaatc ctggacaagg gcagacaggg 1080
cgagcggagg gccagctccg gggctcaggc aaggctgggg gcttccccca gacaccccac 1140
tcctcctctg ctggaccccc acttcatagg gcacttcgtg ttctcaaagg gcttccaaat 1200
agcatggtgg ccttggatgc ccagggaagc ctcagagttg cttatctccc tctagacaga 1260
aggggaatct cggtcaagag ggagaggtcg ccctgttcaa ggccacccag ccagctcatg 1320
gcggtaatgg gacaaggctg gccagccatc ccaccctcag aagggacccg gtggggcagg 1380
tgatctcaga ggaggctcac ttctgggtct cacattcttg gggaggagga ggtctaagtc 1440
ccaggcccaa ttaagagatc aggtagtgta gggtttggga gcttttaagg tgaagaggcc 1500
cgggctgatc ccacaggcca gtataaagcg ccgtgaccct caggtgacgc gccagggccg 1560
gctgccgtcg gggacagggc tttccatagc cggtaccatg ctagaggatc cggtactcga 1620
ggaactgaaa aaccagaaag ttaactggta agtttagtct ttttgtcttt tatttcaggt 1680
cccggatccg gtggtggtgc aaatcaaaga actgctcctc agtggatgtt gcctttactt 1740
ctaggcctgt acggaagtgt tacttctgct ctaaaagctg cggaattgta cccgccgcca 1800
ccatggtcag ctactgggac accggggtcc tgctgtgcgc gctgctcagc tgtctgcttc 1860

tcacaggatc tagttccgga agtgataccg gtagaccttt cgtagagatg tacagtgaaa 1920
tccccgaaat tatacacatg actgaaggaa gggagctcgt cattccctgc cgggttacgt 1980
cacctaacat cactgttact ttaaaaaagt ttccacttga cactttgatc cctgatggaa 2040
aacgcataat ctgggacagt agaaagggct tcatcatatc aaatgcaacg tacaaagaaa 2100
tagggcttct gacctgtgaa gcaacagtca atgggcattt gtataagaca aactatctca 2160
cacatcgaca aaccaataca atcatagatg tggttctgag tccgtctcat ggaattgaac 2220
tatctgttgg agaaaagctt gtcttaaatt gtacagcaag aactgaacta aatgtgggga 2280
ttgacttcaa ctgggaatac ccttcttcga agcatcagca taagaaactt gtaaaccgag 2340
acctaaaaac ccagtctggg agtgagatga agaaattttt gagcacctta actatagatg 2400
gtgtaacccg gagtgaccaa ggattgtaca cctgtgcagc atccagtggg ctgatgacca 2460
agaagaacag cacatttgtc agggtccatg aaaaggacaa aactcacaca tgcccaccgt 2520
gcccagcacc tgaactcctg gggggaccgt cagtcttcct cttcccccca aaacccaagg 2580
acaccctcat gatctcccgg acccctgagg tcacatgcgt ggtggtggac gtgagccacg 2640
aagaccctga ggtcaagttc aactggtacg tggacggcgt ggaggtgcat aatgccaaga 2700
caaagccgcg ggaggagcag tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc 2760
tgcaccagga ctggctgaat ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc 2820
cagcccccat cgagaaaacc atctccaaag ccaaagggca gccccgagaa ccacaggtgt 2880
acaccctgcc cccatcccgg gatgagctga ccaagaacca ggtcagcctg acctgcctgg 2940
tcaaaggctt ctatcccagc gacatcgccg tggagtggga gagcaatggg cagccggaga 3000
acaactacaa gaccacgcct cccgtgctgg actccgacgg ctccttcttc ctctacagca 3060
agctcaccgt ggacaagagc aggtggcagc aggggaacgt cttctcatgc tccgtgatgc 3120
atgaggctct gcacaaccac tacacgcaga agagcctctc cctgtctccg ggtaaataga 3180
atcaacctct ggattacaaa atttgtgaaa gattgactgg tattcttaac tatgttgctc 3240
cttttacgct atgtggatac gctgctttaa tgcctttgta tcatgctatt gcttcccgta 3300
tggctttcat tttctcctcc ttgtataaat cctggttagt tcttgccacg gcggaactca 3360
tcgccgcctg ccttgcccgc tgctggacag gggctcggct gttgggcact gacaattccg 3420
tggtgggtct gcctcctacc catcccgccc accggggctt tggccacctc tcctttcccc 3480
ctccttctcc atccctgtaa aataaatgta atttatcttt gccaaaacca acaaagtcac 3540
agaggctttc actgcagtgt gggaccacct gagcctctgc gtgtgcaggc actgggtctc 3600
gagagggtgc aaggggata aagaggagag agcgcttcat agactttaag ttttcccgag 3660
cctcatgtct accgatggcg tgaaaggatc ctggcaaaac agaagtgtga ggcaagctta 3720
tcgataagga tcttcctaga gcatggctac gtagataagt agcatggcgg gttaatcatt 3780
aactacaagg aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc 3840
actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg 3900
agcgagcgag cgcgca 3916

<210> 95
<211> 3715
<212> DNA
<213> 人工序列

<220>

<223> 合成的

<400> 95

tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc cgggcgtcgg gcgacctttg 60
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120
ggggttcctt gtagttaatg attaacccgc catgctactt atctacgtag ccatgctcta 180
ggaagatctg gaggctgagg ggtggggaaa gggcatgggt gtttcatgag gacagagctt 240
ccgtttcatg caatgaaaag agtttggaga cggatggtgg tgactggact atacacttac 300
acacggtagc gatggtacac tttgtattat gtatatttta ccacgatctt tttaaagtgt 360
caaaggcaaa tggccaaatg gttccttgtc ctatagctgt agcagccatc ggctgttagt 420
gacaaagccc ctgagtcaag atgacagcag cccccataac tcctaatcgg ctctcccgcg 480
tggagtcatt taggagtagt cgcattagag acaagtccaa catctaatct tccaccctgg 540
ccagggcccc agctggcagc gagggtggga gactccgggc agagcagagg gcgctgacat 600
tggggcccgg cctggcttgg gtccctctgg cctttcccca ggggccctct ttccttgggg 660
ctttcttggg ccgccactgc tcccgctcct ctccccccat cccaccccct caccccctcg 720
ttcttcatat ccttctctag tgctccctcc actttcatcc acccttctgc aagagtgtgg 780
gaccacaaat gagttttcac ctggcctggg gacacacgtg cccccacagg tgctgagtga 840
ctttctagga cagtaatctg ctttaggcta aaatgggact tgatcttctg ttagccctaa 900
tcatcaatta gcagagccgg tgaaggtgca gaacctaccg cctttccagg cctcctccca 960
cctctgccac ctccactctc cttcctggga tgtgggggct ggcacacgtg tggcccaggg 1020
cattggtggg attgcactga gctgggtcat tagcgtaatc ctggacaagg gcagacaggg 1080
cgagcggagg gccagctccg gggctcaggc aaggctgggg gcttccccca gacaccccac 1140
tcctcctctg ctggaccccc acttcatagg gcacttcgtg ttctcaaagg gcttccaaat 1200
agcatggtgg ccttggatgc ccagggaagc ctcagagttg cttatctccc tctagacaga 1260
aggggaatct cggtcaagag ggagaggtcg ccctgttcaa ggccacccag ccagctcatg 1320
gcggtaatgg gacaaggctg gccagccatc ccaccctcag aagggacccg gtggggcagg 1380
tgatctcaga ggaggctcac ttctgggtct cacattcttg gggaggagga ggtctaagtc 1440
ccaggcccaa ttaagagatc aggtagtgta gggtttggga gcttttaagg tgaagaggcc 1500
cgggctgatc ccacaggcca gtataaagcg ccgtgaccct caggtgacgc gccagggccg 1560
gctgccgtcg gggacagggc tttccatagc cgccaccatg gcccagcagt ggagcctcca 1620
aaggctcgca ggccgccatc cgcaggacag ctatgaggac agcacccagt ccagcatctt 1680
cacctacacc aacagcaact ccaccagagg ccccttcgaa ggcccgaatt accacatcgc 1740
tcccagatgg gtgtaccacc tcaccagtgt ctggatgatc tttgtggtca ctgcatccgt 1800
cttcacaaat gggcttgtgc tggcggccac catgaagttc aagaagctgc gccacccgct 1860
gaactggatc ctggtgaacc tggcggtcgc tgacctagca gagaccgtca tcgccagcac 1920
tatcagcatt gtgaaccagg tctctggcta cttcgtgctg ggccacccta tgtgtgtcct 1980
ggagggctac accgtctccc tgtgtggtaa gtatcaaggt tacaagacag gtttaaggag 2040
accaatagaa actgggcatg tggagacaga gaagactcac gcgtttctga attcactgac 2100
tctctctgcc tattggtcta ttttctcacc cttagggatc acaggtctct ggtctctggc 2160

catcatttcc tgggagagat ggatggtggt ctgcaagccc tttggcaatg tgagatttga 2220
tgccaagctg gccatcgtgg gcattgcctt ctcctggatc tggtctgctg tgtggacagc 2280
cccgcccatc tttggttgga gcaggttggt atcaaggtta caagacaggt ttaaggagac 2340
caatggatcc atagtcgacc accatggtgg cttagatccg ggcatgtgga gacagagaag 2400
actgttgagt ttgtgataag cactgactct ctctgcctat tggtctattt tccctccctc 2460
aggtactggc cccacggcct gaagacttca tgcggcccag acgtgttcag cggcagctcg 2520
taccccgggg tgcagtctta catgattgtc ctcatggtca cctgctgcat catcccactc 2580
gctatcatca tgctctgcta cctccaagtg tggctggcca tccgagcggt ggcaaagcag 2640
cagaaagagt ctgaatccac ccagaaggca gagaaggaag tgacgcgcat ggtggtggtg 2700
atgatctttg cgtactgcgt ctgctgggga ccctacacct tcttcgcatg ctttgctgct 2760
gccaaccctg gttacgcctt ccacccttg atggctgccc tgccggccta ctttgccaaa 2820
agtgccacta tctacaaccc cgttatctat gtctttatga accggcagtt tcgaaactgc 2880
atcttgcagc ttttcgggaa gaaggttgac gatggctctg aactctccag cgcctccaaa 2940
acggaggtct catctgtgtc ctcggtatcg cctgcatagg gtctgcctcc tacccatccc 3000
gcccaccggg gctttggcca cctctccttt ccccctcctt ctccatccct gtaaaataaa 3060
tgtaatttat ctttgccaaa accaacaaag tcacagaggc tttcactgca gtgtgggacc 3120
acctgagcct ctgcgtgtgc aggcactggg tctcgagagg gtgcaagggg gataaagagg 3180
agagagcgct tcatagactt taagttttcc cgagcctcat gtctaccgat ggcgtgaaag 3240
gatcctggca aaacagaagt gtgaggcaat caacctctgg attacaaaat ttgtgaaaga 3300
ttgactggta ttcttaacta tgttgctcct tttacgctat gtggatacgc tgctttaatg 3360
cctttgtatc atgctattgc ttcccgtatg gctttcattt tctcctcctt gtataaatcc 3420
tggttagttc ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg 3480
gctcggctgt tgggcactga caattccgtg gtgaagctta tcgataagga tcttcctaga 3540
gcatggctag tagataagta gcatggcggg ttaatcatta actacaagga acccctagtg 3600
atggagttgg ccactccctc tctgcgcgct cgctcgctca ctgaggccgg gcgaccaaag 3660
gtcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc gcgca 3715

<210> 96
<211> 178
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 96
ggggaggagg aggtctaagt cccaggccca attaagagat caggtagtgt agggtttggg 60
agcttttaag gtgaagaggc ccgggctgat cccacaggcc agtataaagc gccgtgaccc 120
tcaggtgacg cgccagggcc ggctgccgtc ggggacaggg ctttccatag ccgccacc 178

<210> 97
<211> 409
<212> DNA

<213> 人工序列
<220>
<223> 合成的
<400> 97
atggcccagc agtggagcct ccaaaggctc gcaggccgcc atccgcagga cagctatgag 60
gacagcaccc agtccagcat cttcacctac accaacagca actccaccag aggccccttc 120
gaaggcccga attaccacat cgctcccaga tgggtgtacc acctcaccag tgtctggatg 180
atctttgtgg tcactgcatc cgtcttcaca aatgggcttg tgctggcggc caccatgaag 240
ttcaagaagc tgcgccaccc gctgaactgg atcctggtga acctggcggt cgctgaccta 300
gcagagaccg tcatcgccag cactatcagc attgtgaacc aggtctctgg ctacttcgtg 360
ctgggccacc ctatgtgtgt cctggagggc tacaccgtct ccctgtgtg 409
<210> 98
<211> 169
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 98
ggatcacagg tctctggtct ctggccatca tttcctggga gagatggatg gtggtctgca 60
agcccctttgg caatgtgaga tttgatgcca agctggccat cgtgggcatt gccttctcct 120
ggatctggtc tgctgtgtgg acagccccgc ccatctttgg ttggagcag 169
<210> 99
<211> 517
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 99
gtactggccc cacggcctga agacttcatg cggcccagac gtgttcagcg gcagctcgta 60
ccccggggtg cagtcttaca tgattgtcct catggtcacc tgctgcatca tcccactcgc 120
tatcatcatg ctctgctacc tccaagtgtg gctggccatc cgagcggtgg caaagcagca 180
gaaagagtct gaatccaccc agaaggcaga gaaggaagtg acgcgcatgg tggtggtgat 240
gatctttgcg tactgcgtct gctggggacc ctacaccttc ttcgcatgct ttgctgctgc 300
caaccctggt tacgccttcc acccctttgat ggctgccctg ccggcctact ttgccaaaag 360
tgccactatc tacaaccccg ttatctatgt ctttatgaac cggcagtttc gaaactgcat 420
cttgcagctt ttcgggaaga aggttgacga tggctctgaa ctctccagcg cctccaaaac 480
ggaggtctca tctgtgtcct cggtatcgcc tgcatag 517
<210> 100
<211> 72

<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 100
atgtcacgca agatagaagg ctttttgtta ttacttctct ttggctatga agccacattg 60
ggattatcgt ct 72
<210> 101
<211> 1299
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> 尚未归类的特征
<222> (1297)..(1299)
<223> 可选地缺失
<400> 101
agtgataccg gtagaccttt cgtagagatg tacagtgaaa tccccgaaat tatacacatg 60
actgaaggaa gggagctcgt cattccctgc cgggttacgt cacctaacat cactgttact 120
ttaaaaaagt ttccacttga cactttgatc cctgatggaa aacgcataat ctgggacagt 180
agaaagggct tcatcatatc aaatgcaacg tacaaagaaa tagggcttct gacctgtgaa 240
gcaacagtca atgggcattt gtataagaca aactatctca cacatcgaca aaccaataca 300
atcatagatg tggttctgag tccgtctcat ggaattgaac tatctgttgg agaaaagctt 360
gtcttaaatt gtacagcaag aactgaacta aatgtgggga ttgacttcaa ctgggaatac 420
ccttcttcga agcatcagca taagaaactt gtaaaccgag acctaaaaac ccagtctggg 480
agtgagatga agaaattttt gagcacctta actatagatg gtgtaacccg gagtgaccaa 540
ggattgtaca cctgtgcagc atccagtggg ctgatgacca agaagaacag cacatttgtc 600
agggtccatg aaaaggacaa aactcacaca tgcccaccgt gcccagcacc tgaactcctg 660
gggggaccgt cagtcttcct cttcccccca aaacccaagg acaccctcat gatctcccgg 720
acccctgagg tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagttc 780
aactggtacg tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag 840
tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 900
ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagccccat cgagaaaacc 960
atctccaaag ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg 1020
gatgagctga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc 1080
gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct 1140
cccgtgctgg actccgacgg ctccttcttc ctctacagca agctcaccgt ggacaagagc 1200
aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac 1260

tacacgcaga agagcctctc cctgtctccg ggtaaatag 1299

<210> 102

<211> 455

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 102

Met Ser Arg Lys Ile Glu Gly Phe Leu Leu Leu Leu Leu Phe Gly Tyr
1 5 10 15
Glu Ala Thr Leu Gly Leu Ser Ser Asp Thr Gly Arg Pro Phe Val Glu
20 25 30
Met Tyr Ser Glu Ile Pro Glu Ile Ile His Met Thr Glu Gly Arg Glu
35 40 45
Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile Thr Val Thr Leu
50 55 60
Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly Lys Arg Ile Ile
65 70 75 80
Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys Glu
85 90 95
Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys
100 105 110
Thr Asn Tyr Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val Val
115 120 125
Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu Lys Leu Val
130 135 140
Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile Asp Phe Asn
145 150 155 160
Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu Val Asn Arg
165 170 175
Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe Leu Ser Thr
180 185 190
Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr Cys
195 200 205
Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg
210 215 220
Val His Glu Lys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
225 230 235 240
Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys

245 250 255
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
260 265 270
Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp
275 280 285
Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr
290 295 300
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
305 310 315 320
Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu
325 330 335
Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
340 345 350
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys
355 360 365
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
370 375 380
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
385 390 395 400
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
405 410 415
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser
420 425 430
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
435 440 445
Leu Ser Leu Ser Pro Gly Lys
450 455

<210> 103
<211> 455
<212> PRT
<213> 人工序列
<220>
<223> 合成的
<400> 103
Met Ser Arg Lys Ile Glu Gly Phe Leu Leu Leu Leu Leu Phe Gly Tyr
1 5 10 15
Glu Ala Thr Leu Gly Leu Ser Ser Asp Thr Gly Arg Pro Phe Val Glu
20 25 30
Met Tyr Ser Glu Ile Pro Glu Ile Ile His Met Thr Glu Gly Arg Glu

35 40 45

Leu Val Ile Pro Cys Arg Val Thr Ser Pro Asn Ile Thr Val Thr Leu
50 55 60

Lys Lys Phe Pro Leu Asp Thr Leu Ile Pro Asp Gly Lys Arg Ile Ile
65 70 75 80

Trp Asp Ser Arg Lys Gly Phe Ile Ile Ser Asn Ala Thr Tyr Lys Glu
85 90 95

Ile Gly Leu Leu Thr Cys Glu Ala Thr Val Asn Gly His Leu Tyr Lys
100 105 110

Thr Asn Tyr Leu Thr His Arg Gln Thr Asn Thr Ile Ile Asp Val Val
115 120 125

Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu Lys Leu Val
130 135 140

Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile Asp Phe Asn
145 150 155 160

Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu Val Asn Arg
165 170 175

Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe Leu Ser Thr
180 185 190

Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu Tyr Thr Cys
195 200 205

Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr Phe Val Arg
210 215 220

Val His Glu Lys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
225 230 235 240

Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
245 250 255

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
260 265 270

Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp
275 280 285

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr
290 295 300

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
305 310 315 320

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu
325 330 335

Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
340 345 350

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys
355 360 365
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
370 375 380
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
385 390 395 400
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
405 410 415
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser
420 425 430
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
435 440 445
Leu Ser Leu Ser Pro Gly Lys
450 455

<210> 104
<211> 15
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 104
atagccggta ccatg 15
<210> 105
<211> 1114
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 105
atgagaaaaa tgtcggagga agagttttat ctgttcaaaa atatctcttc agtggggccg 60
tgggatgggc ctcagtacca cattgcccct gtctgggcct tctacctcca ggcagctttc 120
atgggcactg tcttccttat agggttccca ctcaatgcca tggtgctggt ggccacactg 180
cgctacaaaa agttgcggca gcccctcaac tacattctgg tcaacgtgtc cttcggaggc 240
ttcctcctct gcatcttctc tgtcttccct gtcttcgtcg ccagctgtaa cggatacttc 300
gtcttcggtc gccatgtttg tgctttggag ggcttcctgg gcactgtagc aggtctggtt 360
acaggatggt cactggcctt cctggccttt gagcgctaca ttgtcatctg taagcccttc 420
ggcaacttcc gcttcagctc caagcatgca ctgacggtgg tcctggctac ctggaccatt 480
ggtattggcg tctccatccc acccttcttt ggctggagcc ggttcatccc tgagggcctg 540
cagtgttcct gtggccctga ctggtacacc gtgggcacca aataccgcag cgagtcctat 600

acgtggttcc tcttcatctt ctgcttcatt gtgcctctct ccctcatctg cttctcctac 660
actcagctgc tgagggccct gaaagctgtt gcagctcagc agcaggagtc agctacgacc 720
cagaaggctg aacgggaggt gagccgcatg gtggttgtga tggtaggatc cttctgtgtc 780
tgctacgtgc cctacgcggc cttcgccatg tacatggtca acaaccgtaa ccatgggctg 840
gacttacggc ttgtcaccat tccttcattc ttctccaaga gtgcttgcat ctacaatccc 900
atcatctact gcttcatgaa taagcagttc caagcttgca tcatgaagat ggtgtgtggg 960
aaggccatga cagatgaatc cgacacatgc agctcccaga aaacagaagt ttctactgtc 1020
tcgtctaccc aagttggccc caactgagga cccaatattg gcctgtttgc aacagctaga 1080
attaaatttt acttttaaaa aaaaaaaaaa aaaa 1114

<210> 106
<211> 1998
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 106

cccactggcc ggtataaagc accgtgaccc tcaggtgacg caccagggcc ggctgccgtc 60
ggggacaggg ctttccatag ccatggccca gcagtggagc ctccaaaggc tcgcaggccg 120
ccatccgcag gacagctatg aggacagcac ccagtccagc atcttcacct acaccaacag 180
caactccacc agaggcccct tcgaaggccc gaattaccac atcgctccca gatgggtgta 240
ccacctcacc agtgtctgga tgatctttgt ggtcattgca tccgtcttca caaatgggct 300
tgtgctggcg gccaccatga agttcaagaa gctgcgccac ccgctgaact ggatcctggt 360
gaacctggcg gtcgctgacc tggcagagac cgtcatcgcc agcactatca gcgttgtgaa 420
ccaggtctat ggctacttcg tgctgggcca ccctatgtgt gtcctggagg gctacaccgt 480
ctccctgtgt gggatcacag gtctctggtc tctggccatc atttcctggg agagatggat 540
ggtggtctgc aagccctttg gcaatgtgag atttgatgcc aagctggcca tcgtgggcat 600
tgccttctcc tggatctggg ctgctgtgtg gacagccccg cccatctttg gttggagcag 660
gtactggccc cacggcctga agacttcatg cggcccagac gtgttcagcg gcagctcgta 720
ccccggggtg cagtcttaca tgattgtcct catggtcacc tgctgcatca ccccactcag 780
catcatcgtg ctctgctacc tccaagtgtg gctggccatc cgagcggtgg caaagcagca 840
gaaagagtct gaatccaccc agaaggcaga gaaggaagtg acgcgcatgg tggtggtgat 900
ggtcctggca ttctgcttct gctggggacc atacgccttc ttcgcatgct ttgctgctgc 960
caaccctggc taccccttcc acccctttgat ggctgccctg ccggccttct ttgccaaaag 1020
tgccactatc tacaaccccg ttatctatgt ctttatgaac cggcagtttc gaaactgcat 1080
cttgcagctt ttcgggaaga aggttgacga tggctctgaa ctctccagcg cctccaaaac 1140
ggaggtctca tctgtgtcct cggtatcgcc tgcatgaggt ctgcctccta cccatcccgc 1200
ccaccggggc tttggccacc tctcctttcc ccctccttct ccatccctgt aaaataaatg 1260
taatttatct ttgccaaaac caacaaagtc acagaggctt tcactgcagt gtgggaccac 1320
ctgagcctct gcgtgtgcag gcactgggtc tcgagagggt gcaaggggga taaagaggag 1380

agagcgcttc atagacttta agttttcccg agcctcatgt ctaccgatgg cgtgaaagga 1440
tcctggcaaa acagaagtgt gaggcaggtg ggcgtctata tccatttcac caggctggtg 1500
gttacataat cggcaagcaa gagctgtgga ggggcttgct ggatgccctc agcacccagg 1560
aggagggagg gagctagcaa gctaaggcag gtggccctcc tggcccctta aggtccatct 1620
gctggaggcc cagagtcctt ggagtacagt ctacacctgg aggggaccca ttcctgccag 1680
tctgtggcag ggatggcgcg ccacctctgc caggccagga ccccaagccc gatcagcatc 1740
agcatggtgc aggtgcacag gcgtgagctg atcagtgacg aggggcaggc acacaaggtg 1800
gagacaaaga ccaagaggac ggttgccagt gagaggcgcg gactcaggaa cttgaacaac 1860
atctgcgggg gacggctttg gaggtgctcc gctgcctcca gttgggtgac ttgctgtagc 1920
atctccagct tggatattcg gctcttgaag gtctccgtga tctcctgcag gagacgaaaa 1980
tgcacgcacc agaagtca 1998

<210> 107
<211> 1234
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 107

cggctgccgt cggggacagg gctttccata gccatggccc agcagtggag cctccaaagg 60
ctcgcaggcc gccatccgca ggacagctat gaggacagca cccagtccag catcttcacc 120
tacaccaaca gcaactccac cagaggcccc ttcgaaggcc cgaattacca catcgctccc 180
agatgggtgt accacctcac cagtgtctgg atgatctttg tggtcactgc atccgtcttc 240
acaaatgggc ttgtgctggc ggccaccatg aagttcaaga agctgcgcca cccgctgaac 300
tggatcctgg tgaacctggc ggtcgctgac ctagcagaga ccgtcatcgc cagcactatc 360
agcattgtga accaggtctc tggctacttc gtgctgggcc accctatgtg tgtcctggag 420
ggctacaccg tctccctgtg tgggatcaca ggtctctggt ctctggccat catttcctgg 480
gagaggtggc tggtggtgtg caagcccttt ggcaatgtga gatttgatgc caagctggcc 540
atcgtgggca ttgccttctc ctggatctgg tctgctgtgt ggacagcccc gcccatcttt 600
ggttggagca ggtactggcc ccacggcctg aagacttcat gcggcccaga cgtgttcagc 660
ggcagctcgt accccggggt gcagtcttac atgattgtcc tcatggtcac ctgctgcatc 720
atcccactcg ctatcatcat gctctgctac ctccaagtgt ggctggccat ccgagcggtg 780
gcaaagcagc agaaagagtc tgaatccacc cagaaggcag agaaggaagt gacgcgcatg 840
gtggtggtga tgatctttgc gtactgcgtc tgctggggac cctacacctt cttcgcatgc 900
tttgctgctg ccaaccctgg ttacgccttc acccctttga tggctgccct gccggcctac 960
tttgccaaaa gtgccactat ctacaacccc gttatctatg tctttatgaa ccggcagttt 1020
cgaaactgca tcttgcagct tttcgggaag aaggttgacg atggctctga actctccagc 1080
gcctccaaaa cggaggtctc atctgtgtcc tcggtatcgc ctgcatgagg tctgcctcct 1140
acccatcccg cccaccgggg ctttggccac ctctcctttc cccctccttc tccatccctg 1200
taaaataaat gtaatttatc tttgccaaaa ccaa 1234

<210> 108
<211> 7326
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 108

aggacacagc gtccggagcc agaggcgctc ttaacggcgt ttatgtcctt tgctgtctga 60
ggggcctcag ctctgaccaa tctggtcttc gtgtggtcat tagcatgggc ttcgtgagac 120
agatacagct tttgctctgg aagaactgga ccctgcggaa aaggcaaaag attcgctttg 180
tggtggaact cgtgtggcct ttatctttat ttctggtctt gatctggtta aggaatgcca 240
acccgctcta cagccatcat gaatgccatt tccccaacaa ggcgatgccc tcagcaggaa 300
tgctgccgtg gctccagggg atcttctgca atgtgaacaa tccctgtttt caaagcccca 360
ccccaggaga atctcctgga attgtgtcaa actataacaa ctccatcttg gcaagggtat 420
atcgagattt tcaagaactc ctcatgaatg caccagagag ccagcacctt ggccgtattt 480
ggacagagct acacatcttg tcccaattca tggacaccct ccggactcac ccggagagaa 540
ttgcaggaag aggaatacga ataagggata tcttgaaaga tgaagaaaca ctgacactat 600
ttctcattaa aaacatcggc ctgtctgact cagtggtcta ccttctgatc aactctcaag 660
tccgtccaga gcagttcgct catggagtcc cggacctggc gctgaaggac atcgcctgca 720
gcgaggccct cctggagcgc ttcatcatct tcagccagag acgcggggca aagacggtgc 780
gctatgccct gtgctccctc tcccagggca ccctacagtg gatagaagac actctgtatg 840
ccaacgtgga cttcttcaag ctcttccgtg tgcttcccac actcctagac agccgttctc 900
aaggtatcaa tctgagatct tggggaggaa tattatctga tatgtcacca agaattcaag 960
agtttatcca tcggccgagt atgcaggact tgctgtgggt gaccaggccc ctcatgcaga 1020
atggtggtcc agagaccttt acaaagctga tgggcatcct gtctgacctc ctgtgtggct 1080
accccgaggg aggtggctct cgggtgctct ccttcaactg gtatgaagac aataactata 1140
aggcctttct ggggattgac tccacaagga aggatcctat ctattcttat gacagaagaa 1200
caacatcctt ttgtaatgca ttgatccaga gcctggagtc aaatccttta accaaaatcg 1260
cttggagggc ggcaaagcct ttgctgatgg gaaaaatcct gtacactcct gattcacctg 1320
cagcacgaag gatactgaag aatgccaact caacttttga agaactggaa cacgttagga 1380
agttggtcaa agcctgggaa gaagtagggc cccagatctg gtacttcttt gacaacagca 1440
cacagatgaa catgatcaga gataccctgg ggaacccaac agtaaaagac tttttgaata 1500
ggcagcttgg tgaagaaggt attactgctg aagccatcct aaacttcctc tacaagggcc 1560
ctcgggaaag ccaggctgac gacatggcca acttcgactg gagggacata tttaacatca 1620
ctgatcgcac cctccgcctg gtcaatcaat acctggagtg cttggtcctg gataagtttg 1680
aaagctacaa tgatgaaact cagctcaccc aacgtgccct ctctctactg gaggaaaaca 1740
tgttctgggc cggagtggta ttccctgaca tgtatccctg gaccagctct ctaccacccc 1800
acgtgaagta taagatccga atggacatag acgtggtgga gaaaaccaat aagattaaag 1860
acaggtattg ggattctggt cccagagctg atcccgtgga agatttccgg tacatctggg 1920

gcgggtttgc ctatctgcag gacatggttg aacaggggat cacaaggagc caggtgcagg 1980
cggaggctcc agttggaatc tacctccagc agatgcccta cccctgcttc gtggacgatt 2040
ctttcatgat catcctgaac cgctgtttcc ctatcttcat ggtgctggca tggatctact 2100
ctgtctccat gactgtgaag agcatcgtct tggagaagga gttgcgactg aaggagacct 2160
tgaaaaatca gggtgtctcc aatgcagtga tttggtgtac ctggttcctg gacagcttct 2220
ccatcatgtc gatgagcatc ttcctcctga cgatattcat catgcatgga agaatcctac 2280
attacagcga cccattcatc ctcttcctgt tcttgttggc tttctccact gccaccatca 2340
tgctgtgctt tctgctcagc accttcttct ccaaggccag tctggcagca gcctgtagtg 2400
gtgtcatcta tttcaccctc tacctgccac acatcctgtg cttcgcctgg caggaccgca 2460
tgaccgctga gctgaagaag gctgtgagct tactgtctcc ggtggcattt ggatttggca 2520
ctgagtacct ggttcgcttt gaagagcaag gcctggggct gcagtggagc aacatcggga 2580
acagtcccac ggaaggggac gaattcagct tcctgctgtc catgcagatg atgctccttg 2640
atgctgctgt ctatggctta ctcgcttggt accttgatca ggtgtttcca ggagactatg 2700
gaaccccact tccttggtac tttcttctac aagagtcgta ttggcttggc ggtgaagggt 2760
gttcaaccag agaagaaaga gccctggaaa agaccgagcc cctaacagag gaaacggagg 2820
atccagagca cccagaagga atacacgact ccttctttga acgtgagcat ccagggtggg 2880
ttcctggggt atgcgtgaag aatctggtaa agatttttga gccctgtggc cggccagctg 2940
tggaccgtct gaacatcacc ttctacgaga accagatcac cgcattcctg ggccacaatg 3000
gagctgggaa aaccaccacc ttgtccatcc tgacgggtct gttgccacca acctctggga 3060
ctgtgctcgt tgggggaagg gacattgaaa ccagcctgga tgcagtccgg cagagccttg 3120
gcatgtgtcc acagcacaac atcctgttcc accacctcac ggtggctgag cacatgctgt 3180
tctatgccca gctgaaagga aagtcccagg aggaggccca gctggagatg gaagccatgt 3240
tggaggacac aggcctccac cacaagcgga atgaagaggc tcaggaccta tcaggtggca 3300
tgcagagaaa gctgtcggtt gccattgcct ttgtgggaga tgccaaggtg gtgattctgg 3360
acgaacccac ctctggggtg gacccttact cgagacgctc aatctgggat ctgctcctga 3420
agtatcgctc aggcagaacc atcatcatgt ccactcacca catggacgag gccgacctcc 3480
ttggggaccg cattgccatc attgcccagg gaaggctcta ctgctcaggc accccactct 3540
tcctgaagaa ctgctttggc acaggcttgt acttaacctt ggtgcgcaag atgaaaaaca 3600
tccagagcca aaggaaaggc agtgagggga cctgcagctg ctcgtctaag ggtttctcca 3660
ccacgtgtcc agcccacgtc gatgacctaa ctccagaaca agtcctggat ggggatgtaa 3720
atgagctgat ggatgtagtt ctccaccatg ttccagaggc aaagctggtg gagtgcattg 3780
gtcaagaact tatcttcctt cttccaaata agaacttcaa gcacagagca tatgccagcc 3840
ttttcagaga gctggaggag acgctggctg accttggtct cagcagtttt ggaatttctg 3900
acactcccct ggaagagatt tttctgaagg tcacggagga ttctgattca ggacctctgt 3960
ttgcgggtgg cgctcagcag aaaagagaaa acgtcaaccc ccgacacccc tgcttgggtc 4020
ccagagagaa ggctggacag acaccccagg actccaatgt ctgctcccca ggggcgccgg 4080
ctgctcaccc agagggccag cctcccccag agccagagtg cccaggcccg cagctcaaca 4140
cggggacaca gctggtcctc cagcatgtgc aggcgctgct ggtcaagaga ttccaacaca 4200
ccatccgcag ccacaaggac ttcctggcgc agatcgtgct cccggctacc tttgtgtttt 4260

tggctctgat gctttctatt gttatccctc cttttggcga ataccccgct ttgacccttc 4320
acccctggat atatgggcag cagtacacct tcttcagcat ggatgaacca ggcagtgagc 4380
agttcacggt acttgcagac gtcctcctga ataagccagg ctttggcaac cgctgcctga 4440
aggaagggtg gcttccggag taccctgtg gcaactcaac accctggaag actccttctg 4500
tgtccccaaa catcacccag ctgttccaga agcagaaatg gacacaggtc aacccttcac 4560
catcctgcag gtgcagcacc agggagaagc tcaccatgct gccagagtgc cccgagggtg 4620
ccgggggcct cccgcccccc cagagaacac agcgcagcac ggaaattcta caagacctga 4680
cggacaggaa catctccgac ttcttggtaa aaacgtatcc tgctcttata agaagcagct 4740
taaagagcaa attctgggtc aatgaacaga ggtatggagg aatttccatt ggaggaaagc 4800
tcccagtcgt ccccatcacg ggggaagcac ttgttgggtt tttaagcgac cttggccgga 4860
tcatgaatgt gagcgggggc cctatcacta gagaggcctc taaagaaata cctgatttcc 4920
ttaaacatct agaaactgaa gacaacatta aggtgtggtt taataacaaa ggctggcatg 4980
ccctggtcag ctttctcaat gtggcccaca acgccatctt acgggccagc ctgcctaagg 5040
acaggagccc cgaggagtat ggaatcaccg tcattagcca acccctgaac ctgaccaagg 5100
agcagctctc agagattaca gtgctgacca cttcagtgga tgctgtggtt gccatctgcg 5160
tgattttctc catgtccttc gtcccagcca gctttgtcct ttatttgatc caggagcggg 5220
tgaacaaatc caagcacctc cagtttatca gtggagtgag ccccaccacc tactgggtga 5280
ccaacttcct ctgggacatc atgaattatt ccgtgagtgc tgggctggtg gtgggcatct 5340
tcatcgggtt tcagaagaaa gcctacactt ctccagaaaa ccttcctgcc cttgtggcac 5400
tgctcctgct gtatggatgg gcggtcattc ccatgatgta cccagcatcc ttcctgtttg 5460
atgtccccag cacagcctat gtggctttat cttgtgctaa tctgttcatc ggcatcaaca 5520
gcagtgctat taccttcatc ttggaattat ttgagaataa ccggacgctg ctcaggttca 5580
acgccgtgct gaggaagctg ctcattgtct tcccccactt ctgcctgggc cggggcctca 5640
ttgaccttgc actgagccag gctgtgacag atgtctatgc ccggtttggt gaggagcact 5700
ctgcaaatcc gttccactgg gacctgattg ggaagaacct gtttgccatg gtggtggaag 5760
gggtggtgta cttcctcctg accctgctgg tccagcgcca cttcttcctc tcccaatgga 5820
ttgccgagcc cactaaggag cccattgttg atgaagatga tgatgtggct gaagaaagac 5880
aaagaattat tactggtgga aataaaactg acatcttaag gctacatgaa ctaaccaaga 5940
tttatccagg cacctccagc ccagcagtgg acaggctgtg tgtcggagtt cgccctggag 6000
agtgctttgg cctcctggga gtgaatggtg ccggcaaaac aaccacattc aagatgctca 6060
ctggggacac cacagtgacc tcaggggatg ccaccgtagc aggcaagagt attttaacca 6120
atatttctga agtccatcaa aatatgggct actgtcctca gtttgatgca attgatgagc 6180
tgctcacagg acgagaacat ctttaccttt atgcccggct tcgaggtgta ccagcagaag 6240
aaatcgaaaa ggttgcaaac tggagtatta agagcctggg cctgactgtc tacgccgact 6300
gcctggctgg cacgtacagt gggggcaaca agcggaaact ctccacagcc atcgcactca 6360
ttggctgccc accgctggtg ctgctggatg agcccaccac agggatggac ccccaggcac 6420
gccgcatgct gtggaacgtc atcgtgagca tcatcagaga agggagggct gtggtcctca 6480
catcccacag catggaagaa tgtgaggcac tgtgtacccg gctggccatc atggtaaagg 6540
gcgcctttcg atgtatgggc accattcagc atctcaagtc caaatttgga gatggctata 6600

tcgtcacaat gaagatcaaa tccccgaagg acgacctgct tcctgacctg aaccctgtgg 6660
agcagttctt ccaggggaac ttcccaggca gtgtgcagag ggagaggcac tacaacatgc 6720
tccagttcca ggtctcctcc tcctccctgg cgaggatctt ccagctcctc ctctcccaca 6780
aggacagcct gctcatcgag gagtactcag tcacacagac cacactggac caggtgtttg 6840
taaattttgc taaacagcag actgaaagtc atgacctccc tctgcaccct cgagctgctg 6900
gagccagtcg acaagcccag gactgatctt tcacaccgct cgttcctgca gccagaaagg 6960
aactctgggc agctggaggc gcaggagcct gtgcccatat ggtcatccaa atggactggc 7020
cagcgtaaat gaccccactg cagcagaaaa caaacacacg aggagcatgc agcgaattca 7080
gaaagaggtc tttcagaagg aaaccgaaac tgacttgctc acctggaaca cctgatggtg 7140
aaaccaaaca aatacaaaat ccttctccag accccagaac tagaaacccc gggccatccc 7200
actagcagct ttggcctcca tattgctctc atttcaagca gatctgcttt tctgcatgtt 7260
tgtctgtgtg tctgcgttgt gtgtgatttt catggaaaaa taaaatgcaa atgcactcat 7320
cacaaa 7326

<210> 109
<211> 2608
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 109

tccttcttca ttctgcagtt ggtgccagaa ctctggatcc tgaactggaa gaaaatgtct 60
atccaggttg agcatcctgc tggtggttac aagaaactgt ttgaaactgt ggaggaactg 120
tcctcgccgc tcacagctca tgtaacaggc aggatccccc tctggctcac cggcagtctc 180
cttcgatgtg ggccaggact ctttgaagtt ggatctgagc cattttacca cctgtttgat 240
gggcaagccc tcctgcacaa gtttgacttt aaagaaggac atgtcacata ccacagaagg 300
ttcatccgca ctgatgctta cgtacgggca atgactgaga aaaggatcgt cataacagaa 360
tttggcacct gtgctttccc agatccctgc aagaatatat tttccaggtt tttttcttac 420
tttcgaggag tagaggttac tgacaatgcc cttgttaatg tctacccagt gggggaagat 480
tactacgctt gcacagagac caactttatt acaaagatta atccagagac cttggagaca 540
attaagcagg ttgatctttg caactatgtc tctgtcaatg gggccactgc tcacccccac 600
attgaaaatg atggaaccgt ttacaatatt ggtaattgct ttggaaaaaa tttttcaatt 660
gcctacaaca ttgtaaagat cccaccactg caagcagaca aggaagatcc aataagcaag 720
tcagagatcg ttgtacaatt cccctgcagt gaccgattca agccatctta cgttcatagt 780
tttggtctga ctcccaacta tatcgttttt gtggagacac cagtcaaaat taacctgttc 840
aagttccttt cttcatggag tctttgggga gccaactaca tggattgttt tgagtccaat 900
gaaaccatgg gggtttggct tcatattgct gacaaaaaaa ggaaaaagta cctcaataat 960
aaatacagaa cttctccttt caacctcttc catcacatca acacctatga agacaatggg 1020
tttctgattg tggatctctg ctgctggaaa ggatttgagt ttgtttataa ttacttatat 1080
ttagccaatt tacgtgagaa ctgggaagag gtgaaaaaaa atgccagaaa ggctccccaa 1140

cctgaagtta ggagatatgt acttcctttg aatattgaca aggctgacac aggcaagaat 1200
ttagtcacgc tccccaatac aactgccact gcaattctgt gcagtgacga gactatctgg 1260
ctggagcctg aagttctctt ttcagggcct cgtcaagcat ttgagtttcc tcaaatcaat 1320
taccagaagt attgtgggaa accttacaca tatgcgtatg gacttggctt gaatcacttt 1380
gttccagata ggctctgtaa gctgaatgtc aaaactaaag aaacttgggt ttggcaagag 1440
cctgattcat acccatcaga acccatcttt gtttctcacc cagatgcctt ggaagaagat 1500
gatggtgtag ttctgagtgt ggtggtgagc ccaggagcag gacaaaagcc tgcttatctc 1560
ctgattctga atgccaagga cttaagtgaa gttgcccggg ctgaagtgga gattaacatc 1620
cctgtcacct ttcatggact gttcaaaaaa tcttgagcat actccagcaa gatatgtttt 1680
tggtagcaaa actgagaaaa tcagcttcag gtctgcaatc aaattctgtt caattttagc 1740
ctgctatatg tcatggtttt aacttgcaga tgcgcacaat tttgcaatgt tttacagaaa 1800
gcactgagtt gagcaagcaa ttcctttatt taaaaaaaaa agtacgtatt tagataatca 1860
tacttcctct gtgagacagg ccataactga aaaactctta aatatttagc aatcaaatag 1920
gaaatgaatg tggacttact aaatggcttt taattcctat tataagagca tattttaggt 1980
acctatctgc tccaattata tttttaacat ttaaaaacca aagtcctcta cacttgattt 2040
atattatatg tggctttgct gagtcaagga agtatcatgc aataaggctt aattactaaa 2100
tgtcaaacca aacttttttct caaaccaggg actatcatct aagattaatt acagtaatta 2160
ttttgcgtat acgtaactgc tcaaagatta tgaatcttat gaatgttaac ctttccgttt 2220
attacaagca agtactatta tttctgattt tataataaga aaatctgtgt ttaatcaact 2280
gaggcctctc aaccaaataa catctcagag attaagttat atattaaaag cttatgtaac 2340
ataaaagcaa gtacatatag tagtgactat atttaaaaaa acagcataaa atgcttaaaa 2400
atgtaatatt tactaaaatc agattatggg ataatgttgc aggattatac tttattgcat 2460
ctttttgtt taattgtatt taagcattgt gcaatcactt gggaaaaata ttaaattatt 2520
aacattgagg tattaataca ttttaagcct tttgttttta aatttctttt cttccagaga 2580
ttgtttaaaa ataaatattg acaaaaat 2608

<210> 110
<211> 1767
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 110

gtcagctggg tgactcatac atctgtgaga ccctccgtat ggagaccggg gcgggggtg 60
ttgcttaatg cttagaactg ctgaaactac cagttcactg tctcagcact aatccaactc 120
tgctccttaa gtgggatttg gcttttagac attgagacct ggatgctggg catcctcgct 180
agatccccta caaattcccc acatacgtag gccaggagcc tcagcggtgc cccttcaggc 240
tcatctggca agacggtacc agcttgctca gaacaggggc tggctattca tcatctcaga 300
gcatagagac cctctccttg ccacccggcc cttcccacct ggttggtgac aaatcacaag 360
gtggtagaag ttgccaggga cagataacat ggcagccagc gggaagacca gcaagtccga 420

accgaaccat gttatcttca agaagatctc ccgggacaaa tcggtgacca tctacctggg 480
gaacagagac tacatagacc atgtcagcca agtccagcct gtggatggtg tcgtgttggt 540
tgatcctgat cttgtgaagg gaaagaaagt gtatgtcact ctgacctgcg ccttccgcta 600
tggccaagag gacattgacg tgatcggctt gaccttccgc agggacctgt acttctcccg 660
ggtccaggtg tatcctcctg tgggggccgc gagcaccccc acaaaactgc aagagagcct 720
gcttaaaaag ctggggagca acacgtaccc ctttctcctg acgtttcctg actacttgcc 780
ctgttcagtg atgttgcagc cagctccaca agattcaggg aagtcctgtg gggttgactt 840
tgaggtcaaa gcattcgcca cagacagcac cgatgccgaa gaggacaaaa tccccaagaa 900
gagctccgtg cgattactga tccgcaaagt acagcatgcc ccacttgaga tgggtcccca 960
gccccgagct gaggcggcct ggcagttctt catgtctgac aagcccctgc accttgcggt 1020
ctctctcaac aaagagatct atttccatgg ggagcccatc cctgtgaccg tgactgtcac 1080
caataacaca gagaagaccg tgaagaagat taaagcattc gtggaacagg tggccaatgt 1140
ggttctctac tcgagtgatt attacgtcaa gcccgtggct atggaggaag cgcaagaaaa 1200
agtgccacca aacagcactt tgaccaagac gctgacgctg ctgcccttgc tggctaacaa 1260
tcgagaaagg agaggcattg ccctggatgg gaaaatcaag cacgaggaca caaaccttgc 1320
ctccagcacc atcattaagg agggcataga ccggaccgtc ctgggaatcc tggtgtctta 1380
ccagatcaag gtgaagctca cagtgtcagg ctttctggga gagctcacct ccagtgaagt 1440
cgccactgag gtcccattcc gcctcatgca ccctcagcct gaggacccag ctaaggaaag 1500
ttatcaggat gcaaatttag tttttgagga gtttgctcgc cataatctga aagatgcagg 1560
agaagctgag gaggggaaga gagacaagaa tgacgttgat gagtgaagat gtcggctcag 1620
gatgccggaa aatgacctgt agttaccagt gcaacgagca aagccccaca gtttagtcct 1680
ttggagttat gctgcgtatg aaaggatgag tcttcttccg agaaataaag cttgtttgtt 1740
ctcccctgga aaaaaaaaaa aaaaaaa 1767

<210> 111
<211> 2483
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 111

agttgattgc aggtcctcct ggggccagaa gggtgcctgg gaggccaggt tctggggatc 60
ccctccatcc agaagaacca cctgctcact ctgtcccttc gcctgctgct gggaccatgg 120
gggctggggc cagtgctgag gagaagcact ccagggagct ggaaaagaag ctgaaagagg 180
acgctgagaa ggatgctcga accgtgaagc tgctgcttct gggtgccggt gagtccggga 240
agagcaccat cgtcaagcag atgaagatta tccaccagga cgggtactcg ctggaagagt 300
gcctcgagtt tatcgccatc atctacggca acacgttgca gtccatcctg gccatcgtac 360
gcgccatgac cacactcaac atccagtacg gagactctgc acgccaggac gacgcccgga 420
agctgatgca catggcagac actatcgagg agggcacgat gcccaaggag atgtcggaca 480
tcatccagcg gctgtggaag gactccggta tccaggcctg ttttgagcgc gcctcggagt 540

accagctcaa cgactcggcg ggctactacc tctccgacct ggagcgcctg gtaaccccgg 600
gctacgtgcc caccgagcag gacgtgctgc gctcgcgagt caagaccact ggcatcatcg 660
agacgcagtt ctccttcaag gatctcaact tccggatgtt cgatgtgggc gggcagcgct 720
cggagcgcaa gaagtggatc cactgcttcg agggcgtgac ctgcatcatc ttcatcgcgg 780
cgctgagcgc ctacgacatg gtgctagtgg aggacgacga agtgaaccgc atgcacgaga 840
gcctgcacct gttcaacagc atctgcaacc accgctactt cgccacgacg tccatcgtgc 900
tcttccttaa caagaaggac gtcttcttcg agaagatcaa gaaggcgcac ctcagcatct 960
gtttcccgga ctacgatgga cccaacacct acgaggacgc cggcaactac atcaaggtgc 1020
agttcctcga gctcaacatg cggcgcgacg tgaaggagat ctattcccac atgacgtgcg 1080
ccaccgacac gcagaacgtc aaatttgtct tcgacgctgt caccgacatc atcatcaagg 1140
agaacctcaa agactgtggc ctcttctgag gccagggcct gtgctgcagt cggggacaag 1200
gagcttccgt ctggcaaggc cggggcacaa tttgcactcc cctcagctag acgcacagac 1260
tcagcaataa acctttgcat caggctccag ctgtcctttc ttggtggagg acttaattat 1320
cacaagtcat gggcatttat taagtgccca gtgctgggtt gggcatgaag tgggaagatg 1380
gcccctccca ggaagaagta cctggcctga caaggtgggg cactcttggg ggtatgggac 1440
caactcatgg cttttcacgg gagttgagga gagaggagct gtggaaaata ttcactggga 1500
cagtcttgga tcaagaggga gttttgaggt ggaggctcat tctggcaggg accgtagtgt 1560
ctaccagccc cagaaacatg ggcttatggc cacaggagtt cagtggagca agagcagggg 1620
aggagagacg tggacaggtg cccaaagcca gtcggagggc ctgggctttc tcagaaggtg 1680
atggagagtc ttggaagccc tcgaggcagg aacataattg cagggctggg attagggtga 1740
gggaagtgag gcacactcac cttgggtgca acatttaagg cgatgccaaa aaatttagta 1800
accaaggtaa ataatattag gataatattt ttaaaaatca aatgaatgca aaaccccaca 1860
atgaatgaaa tatcaaaatc caacagagga tcaaacagag gcatgctaag atatattggg 1920
gcttgaagca aagggaaaac tatttgttgc tatatgtttg tagggatttt ttgccagttt 1980
taaaaataca tgtatcataa agtttactat ctcagccact tgccggtgta tagtttggtg 2040
gtgttaagta cattcataat gttgtacaac caccgcaact gttcatctcc agaactcctt 2100
tcctcttgta aaactgtaac tctgtaccca tgaaaaaata acccccccatt cctgccttcc 2160
cccggctcct ggcatccacc attctacttt ccatctctat gaatgtgact gctctaagtg 2220
cctcagatgt gtgggtccat gaagtctttg tctttttgca actggcttat ttcacttagc 2280
atcatgtctt caaggtttat tcatgtgtag catatggcag aatctccttc ctttttaagg 2340
ttgaataata ttccattgta tatattccac actttgttta tttattcatc tattgatgaa 2400
tggttacatc tgccttttgg ctattgtgaa taatgctgct atgaacatgg gtgtacaaat 2460
ctctcaaaaa aaaaaaaaaa aaa 2483

<210> 112
<211> 2981
<212> DNA
<213> 人工序列
<220>
<223> 合成的

<400> 112

gctgggtaaa tcccagagtc tcagccgcct aagtgtcttc cccggaggtg agattatctc 60
cgcctgtgct ggacacctcc ctttctcctg cagccatgga tgccgctctg ctcctgaacg 120
tggaaggggt caagaaaacc attctgcacg ggggcacggg cgagctccca aacttcatca 180
ccggatcccg agtgatcttt catttccgca ccatgaaatg tgatgaggag cggacagtca 240
ttgacgacag tcggcaggtg ggccagccca tgcacatcat catcggaaac atgttcaagc 300
tcgaggtctg ggagatcctg cttacctcca tgcgggtgca cgaggtggcc gagttctggt 360
gcgacaccat ccacacgggg gtctacccca tcctatcccg gagcctgagg cagatggccc 420
agggcaagga ccccacagag tggcacgtgc acacgtgcgg gctggccaac atgttcgcct 480
accacacgct gggctacgag gacctggacg agctgcagaa ggagcctcag cctctggtct 540
ttgtgatcga gctgctgcag gttgatgccc cgagtgatta ccagagggag acctggaacc 600
tgagcaatca tgagaagatg aaggcggtgc ccgtcctcca cggagaggga aatcggctct 660
tcaagctggg ccgctacgag gaggcctctt ccaagtacca ggaggccatc atctgcctaa 720
ggaacctgca gaccaaggag aagccatggg aggtgcagtg gctgaagctg gagaagatga 780
tcaatactct gatcctcaac tactgccagt gcctgctgaa gaaggaggag tactatgagg 840
tgctggagca caccagtgat attctccggc accacccagg catcgtgaag gcctactacg 900
tgcgtgcccg ggctcacgca gaggtgtgga atgaggccga ggccaaggcg gacctccaga 960
aagtgctgga gctggagccg tccatgcaga aggcggtgcg cagggagctg aggctgctgg 1020
agaaccgcat ggcggagaag caggaggagg agcggctgcg ctgccggaac atgctgagcc 1080
agggtgccac gcagcctccc gcagagccac ccacagagcc acccgcacag tcatccacag 1140
agccacctgc agagccaccc acagcaccat ctgcagagct gtccgcaggg ccccctgcag 1200
agccagccac agagccaccc ccgtccccag ggcactcgct gcagcactga gccccctgag 1260
gcccacagcc acccaggcag ggagcaagtg gcctggtcac ttctggttcg attgaccagg 1320
atcgtggtgt cacttttaa aatttaaaat taatttttga aatcaaagtc agacacaccc 1380
atggtaaaaa aaaaaaaaaa aacaatccca agggtacaga agagcttatg aataaaagta 1440
gttttctcct ctacccctct cattccttcc gtgccatggt tttaattgac cctgttttta 1500
attcttctgg tagttttctc tatttccaag taatctgttt aaatcagttt ctagatttta 1560
ccccatgtca atgacaaatg aggatttgat gctctgatcc tttctcatgc ctgatacccc 1620
tccctgtctc cccattttgg atagttacat ttgggggtca tctcggtgat ttttgtaact 1680
ttacgcagga cacttagagc tctctagaat cccactgact ttagtgggtc ttgatgtagg 1740
gtgggcaagc cccgacactg gagcttagcc tgagagggtt cttggcctcc cccaggaaag 1800
atttcaaagg caagcgccag tggtagggta gaagaaaaca gctgtggtcg ggcacggtgg 1860
ctcacgccta taatcccagc actttgggag gccgaggcgg gtggatcacc tgaggtcagg 1920
agttccagac cagcctggcc aacatggtga aacctcatct ctactaaaaa tacaaaaaaa 1980
ctagctgggc gtggtggcgg gcgcctataa tcccagctac ttgggaggct gaggcaggag 2040
aattgattga acctgggagg cggaggttgc agtgagccaa gatcacgtca ttgcactcca 2100
gcctggtcaa caagagtaaa acttcatctc aaaaaacaaa acaaaacaaa aacaacaaca 2160
aaaaacaaaa gaaaaacaaa caaaaccaaa accaaaacag ctgtattgaa gctgcagtgt 2220
tgcagctctg tgactgccct gcagagcagg gctaccccat aggcagcgag cagcagctca 2280

gggcagttct gcagtcagat ttatacccac ttttaattac atgtagatta aggagctgca 2340
tatacaaaga tttctaggga aggagtagta acttctgggt cctggggtct ttgccacgga 2400
acagggcagt atgccagggt gttgccacgg caatggtaaa ctgacatggc accctggggg 2460
tcatgcctta gggaaagccg cttccactcg cccctgtttt agctcatctt caagttagtc 2520
tggtgtccaa gctccaccgc ctgcctcagt ctggtgacct ccttctgtgt ctgatgagca 2580
tggcagcgtt gggaccttcc ccttccaact ctctccctcc tcttcgtcct ccctaaagga 2640
cgggtacgag gaggggctat cacgccagcg acatcctcta gcaccaccca ggtgtgtggg 2700
gtggggcagg ggggcgacga agtatccagc ccagggccac gtagtcaact gccaagggct 2760
tcctgggctt ctcttctgcc ccagagcttg tctccaccca gcaggggttc ccccagcgct 2820
aactgtatcc ctaaagttct gatgtacttt acttttccat cttccttgtt gttaacatct 2880
accttctgct ctgtaagcaa aactaaatct tctgtgcttt gtccataggt tgattctaca 2940
atctgaaaat caataaacag catttgcatg aaaaaaaaaa a 2981

<210> 113
<211> 1934
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 113

gggcacaagc aatcctccct tctcagcttc ctgagtggcc aggactacag aggactgtat 60
gctgttctta aggactctct gcttcctgga caagctcaag ctaaggacta catctcccag 120
caggctgtgc tctgacagct cttggattta aataggattc tgggctctgc tcagagtcag 180
gctgctgctc agcacccagg acggagagga gcagagaagc agcagaagca gccaagagct 240
ggagccagac caggaacctg agccagagct ggggttgaag ctggagcagc agcaaaagca 300
acagcagcta cagaagttgg aacgatgctg gtcaccttgg gactgctcac ctccttcttc 360
tcgttcctgt atatggtagc tccatccatc aggaagttct ttgctggtgg agtgtgtaga 420
acaaatgtgc agcttcctgg caaggtagtg gtgatcactg gcgccaacac gggcattggc 480
aaggagacgg ccagagagct cgctagccga ggagcccgag tctatattgc ctgcagagat 540
gtactgaagg gggagtctgc tgccagtgaa atccgagtgg atacaaagaa ctcccaggtg 600
ctggtgcgga aattggacct atccgacacc aaatctatcc gagcctttgc tgagggcttt 660
ctggcagagg aaaagcagct ccatattctg atcaacaatg cgggagtaat gatgtgtcca 720
tattccaaga cagctgatgg ctttgaaacc cacctgggag tcaaccacct gggccacttc 780
ctcctcacct acctgctcct ggagcggcta aaggtgtctg cccctgcacg ggtggttaat 840
gtgtcctcgg tggctcacca cattggcaag attcccttcc acgacctcca gagcgagaag 900
cgctacagca ggggttttgc ctattgccac agcaagctgg ccaatgtgct ttttactcgt 960
gagctggcca agaggctcca aggcaccggg gtcaccacct acgcagtgca cccaggcgtc 1020
gtccgctctg agctggtccg gcactcctcc ctgctctgcc tgctctggcg gctcttctcc 1080
ccctttgtca agacggcacg ggaggggggcg cagaccagcc tgcactgcgc cctggctgag 1140
ggcctggagc ccctgagtgg caagtacttc agtgactgca agaggacctg ggtgtctcca 1200

agggcccgaa ataacaaaac agctgagcgc ctatggaatg tcagctgtga gcttctagga 1260
atccggtggg agtagctggt ggaagagctg cagctttatc aggcccaatc catgccataa 1320
tgaacaggga ccaaggagaa ggccaaccct aaaggattgt cctcttggcc agctggtgct 1380
gcgaatcctg cctgctctga tcctcttgac ccttctggga atgtttgcac acctgacact 1440
cttgtgagac tggcttatgg catgagttgt ggacacctat agagtgttct tctctaagac 1500
ctggaaagtc agcaaccctc tgggggcagc aggactgggc agatcccagg ctgggcatgg 1560
gggtggcaga agagcccgag aaattgggtc agttccctca tcagcaccag aggctcagct 1620
gaggcaagaa gagcaccatc actgcctatt tctaggggct atacactcca actcttggtt 1680
gatctctttc tttttaaaaa tatttgccac caccctggag tctagaccaa cacacaaaga 1740
tcctggctaa ccctggccta tttagattcc ttcctctcac ctggaccttc ccatttcaat 1800
catgcagatg gtttcttttt gtaaagagtt ccgtttgcct ttcaattttt agagaaaata 1860
aagactgcat tcatctcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1920
aaaaaaaaaa aaaa 1934

<210> 114
<211> 4729
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 114

agtctaggcc tccgcctccg ttaccctgga gcccaggtta ccgccgctgc cacccaggag 60
ccccgatcct cgcctctgtc ccatccttgt gttcaaacct cccgcatctc ggcaacctcg 120
gcacccgccc ggcagcctcc gcaggaacca ggcacccgct ctttggcggt cagacgccga 180
ggccccagct gggagtttgg tcctaagagg gaaggcaagg aggcgggacg ccgcatcggc 240
tctgctgaag agcctgcggg ttggaggtgg gctttgaagt gggcgtggag acggcggggg 300
aggtggaggt gcgagtataa atgatcaacc agaaattatc ttcaaaggaa taaaaccaga 360
agtatgtaaa taaaaagccc aagataaaga aacagaaaag ctgacactac atgaagcaga 420
gggcaaaaaa gtttatcttc tggatgccaa tgtgaattgt ggtctacaaa tacattgtgg 480
agaaaataga ttgcacagaa atgaatatta tcaggatctg aagactgtga aaatgttttt 540
cagtattgtc atagtctcct ctggagaaaa taatctgtga aattatgtga atagagacca 600
tttttcaaaa caatggggga aagagcagga agtccaggta ctgatcaaga aagaaaggca 660
ggcaaacacc attattctta cttatctgat tttgaaacgc cacagtcttc tggccgatca 720
tcgctggtca gttcttcacc tgcaagtgtt aggagaaaaa atcctaaaag acaaacttca 780
gatggccaag tacatcacca agcccctcgg aaaccaagcc ctaagggtct accaaacaga 840
aagggagtcc gagtgggatt tcgctcccag agcctcaata gagagccact tcggaaagat 900
actgatcttg ttacaaaacg gattctgtct gcaagactgc taaaaatcaa tgagttgcag 960
aatgaagtat ctgaactcca ggtcaagtta gctgagctgc taaaagaaaa taaatctttg 1020
aaaaggcttc agtacagaca ggagaaagcc ctgaataagt ttgaagatgc cgaaaatgaa 1080
atctcacaac ttatatttcg tcataacaat gagattacag cactcaaaga acgcttaaga 1140

aaatctcaag agaaagaacg ggcaactgag aaaagggtaa aagatacaga aagtgaacta 1200
tttaggacaa aattttcctt acagaaactg aaagagatct ctgaagctag acacctacct 1260
gaacgagatg atttggcaaa gaaactagtt tcagcagagt taaagttaga tgacaccgag 1320
agaagaatta aggagctatc gaaaaacctt gaactgagta ctaacagttt ccaacgacag 1380
ttgcttgctg aaaggaaaag ggcatatgag gctcatgatg aaaataaagt tcttcaaaag 1440
gaggtacagc gactatatca caaattaaag gaaaaggaga gagaactgga tataaaaaat 1500
atatattcta atcgtctgcc aaagtcctct ccaaataaag agaaagaact tgcattaaga 1560
aaaaatgctg catgccagag tgattttgca gacctgtgta caaaaggagt acaaaccatg 1620
gaagacttca agccagaaga atatccttta actccagaaa caattatgtg ttacgaaaac 1680
aaatgggaag aaccaggaca tcttactttg gacttgcaat ctcaaaagca agacaggcat 1740
ggagaagcag ggattctaaa cccaattatg gaaagagaag aaaaatttgt tacagatgaa 1800
gaactccatg tcgtaaaaca ggaggttgaa aagctggagg atgaatggga aagagaagaa 1860
cttgataaaa agcaaaaaga aaaggcatct ttactggaaa gagaagaaaa gccagagtgg 1920
gaaactggaa ggtaccaact aggaatgtat ccaattcaga atatggataa attgcaagga 1980
gaggaagaag aaagactgaa gagagaaatg ctacttgcta aactgaatga aattgacaga 2040
gaactccaag attctcgaaa tctaaaatac cctgttttgc cattgttacc tgattttgaa 2100
tcaaaactac actccccaga gagaagcccc aaaacataca ggttctctga atcctcagag 2160
agattattta atgggcatca tttgcaagac atcagtttct caactccaaa aggagaaggt 2220
cagaattcag gaaatgttag aagtccagcc tcccctaatg agttcgcatt tggtagctac 2280
gtgccttcgt ttgcaaaaac atcagagagg tcaaatccat ttagtcaaaa aagtagtttt 2340
ttggatttcc aaagaaacag tatggaaaaa cttagtaaag atggtgtaga tttaattaca 2400
agaaaagaga aaaaagctaa tttgatggaa cagttatttg gtgccagtgg tagcagcacc 2460
atttcctcca aaagcagtga cccaaattct gtggcttcca gtaaaggaga cattgaccct 2520
ctaaattttc tccctgggaa taaaggcagc agagatcaag aacatgatga agatgaaggc 2580
tttttcctca gtgaaggaag aagttttaat ccaaataggc accgattaaa acatgcagac 2640
gataaaccag cagtaaaagc agctgattct gtagaagatg aaattgaaga agtagcactg 2700
agatgactga ctagagtata catttttct aattgtaaat attgaaatat tttaatacag 2760
tatttattat aaacatttag acttttaat gctaaaatgt ccttattaag gaatgatttt 2820
taatagctaa atacaatgca gttaaaaaga agtagatcat acatacacca catagatagt 2880
ttgccaagaa tgaaggaggc ttttttgaat gaaaccaaaa ataaaaatat gtcttgaaaa 2940
atgaaataga ttttaactct tcatccagtg ctatggcaag tttaactgca ctggaggtgg 3000
tattccttt ctacttttat tcctatttat gtatttattt ttaatcatat tctcactgtg 3060
ctaaatacag tcttcccact aattgttgaa tgaaaattaa gggtgaaagt ctgtgttggg 3120
aagtgtagct ccggatggtg gagaaccagt gcttcttagg agcccttccc tttatggata 3180
gggccagggg tccctatctt acgtggcagt cctaagctac tcttgagtga tagcagtcac 3240
aacattggga tttcccattc ctctgaaagt acccacagct aaacactact catttcccaa 3300
tttcagtatt tactgaaatc acttacctac aattctgtta gaatattatg ttgagttcgg 3360
aagacatttc tctgtttgca aggacagtta gttgctctct gtcatagccc gcagaagctt 3420
ggctacagtg taattcctct ccttgttctc ctacactctt tgtatctcag tgactgggtg 3480

taagatttca tgccaaatgc aaggaatagt ggaaatacat accaactgca aagaatagaa 3540
aaactttatc ccaatatatt atagaaagat cttcagttca attatgtcca gagtaatata 3600
atttggcaat gttaatgctt ttaagatttg aacttgtcct caaaccaagg agacaacaat 3660
agtgtaatac tattggactt acaggattag ttttaaagca actttgaata gaagtgtttc 3720
agacatagga cttctgcctt gttgactaga gtggatagtt ttctgtttaa atgcattggt 3780
cttttggctg tttgtaattc acttcagctt atagagaagt tgatgacctt ggatcatgct 3840
gggtatgatt ggtctttaaa aagcagtaat ataccaccag cccaaggaga aaacattgtt 3900
aaaatgaaga gtggtggaaa tagtgtgttg ggaagaataa aaaatttaga ttggcactta 3960
ttctaaagtg agctgttttt ttcaagaatt tactagcatt tgctcgtagt atgatttctg 4020
acgccagtca tatatactga tgaggggaag gagttactgt gttattctga gttcttataa 4080
atgcatatta ttgaatttca ttgcatgact atttgtcaag acttacctgg ttaggtctcc 4140
aattaaaaga tgtaccacct gtcatcttct aaattgtgtt cctttcattt attgggaaca 4200
gaatctctca aaagtgctaa actatattaa aaagttttct taatagattt gtatgtctga 4260
tatgtataac catttactat aatatgttgc aaatcattct aatatatgta aaacaatatt 4320
atttgtaatg cagttattgg taaaatatta tgtaatgtta attttgttcc tagctgctaa 4380
tttttctgcc aaaagtattc taattttcag gttgttttaa aggcttttaa aacttttag 4440
ttaagttttt tattcacgtc atttaaatta gtttttgct ttttttctac ctgataatct 4500
ctttaacaag atgcaacaag acagaattat taaaattaaa cctaaagtta agttacagat 4560
ttaaaggcat tatgatatta agcattaaaa ttggagatta aaatgtacaa aacagggttt 4620
tccttatgaa caaaccttac agggtaaatt gttttttttt ctaaatgtca ttaaatttat 4680
ttgtactcag aactgttact aataaaaatt aaaaatgcaa aaaaaaaaa 4729

<210> 115
<211> 1374
<212> DNA
<213> 人工序列
<220>
<223> 合成的
<400> 115

atggtcagct actgggacac cggggtcctg ctgtgcgcgc tgctcagctg tctgcttctc 60
acaggatcta gttccggaag tgataccggt agacctttcg tagagatgta cagtgaaatc 120
cccgaaatta tacacatgac tgaaggaagg gagctcgtca ttccctgccg ggttacgtca 180
cctaacatca ctgttacttt aaaaaagttt ccacttgaca ctttgatccc tgatggaaaa 240
cgcataatct gggacagtag aaagggcttc atcatatcaa atgcaacgta caaagaaata 300
gggcttctga cctgtgaagc aacagtcaat gggcatttgt ataagacaaa ctatctcaca 360
catcgacaaa ccaatacaat catagatgtg gttctgagtc cgtctcatgg aattgaacta 420
tctgttggag aaaagcttgt cttaaattgt acagcaagaa ctgaactaaa tgtggggatt 480
gacttcaact gggaataccc ttcttcgaag catcagcata agaaacttgt aaaccgagac 540
ctaaaaaccc agtctgggag tgagatgaag aaatttttga gcaccttaac tatagatggt 600
gtaacccgga gtgaccaagg attgtacacc tgtgcagcat ccagtgggct gatgaccaag 660

aagaacagca catttgtcag ggtccatgaa aaggacaaaa ctcacacatg cccaccgtgc 720
ccagcacctg aactcctggg gggaccgtca gtcttcctct tccccccaaa acccaaggac 780
accctcatga tctcccggac ccctgaggtc acatgcgtgg tggtggacgt gagccacgaa 840
gaccctgagg tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca 900
aagccgcggg aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg 960
caccaggact ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca 1020
gcccccatcg agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac 1080
accctgcccc catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc 1140
aaaggcttct atcccagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac 1200
aactacaaga ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag 1260
ctcaccgtgg acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat 1320
gaggctctgc acaaccacta cacgcagaag agcctctccc tgtctccggg taaa 1374

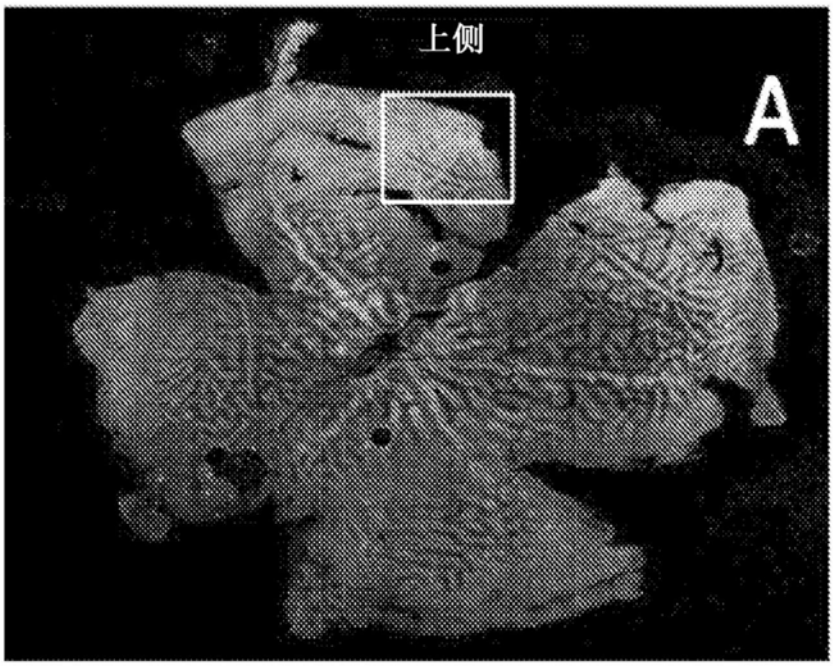

图1A

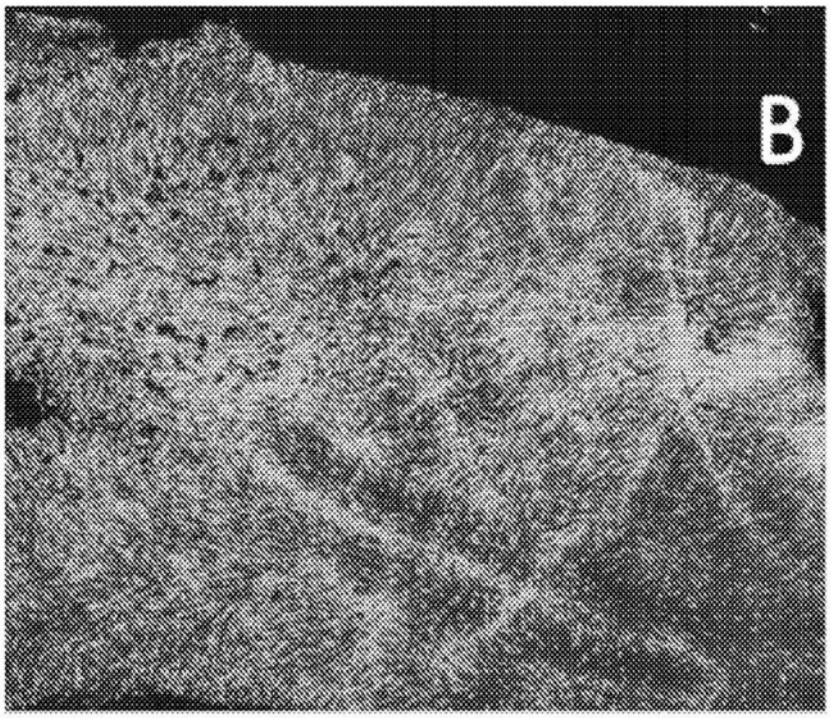

图1B

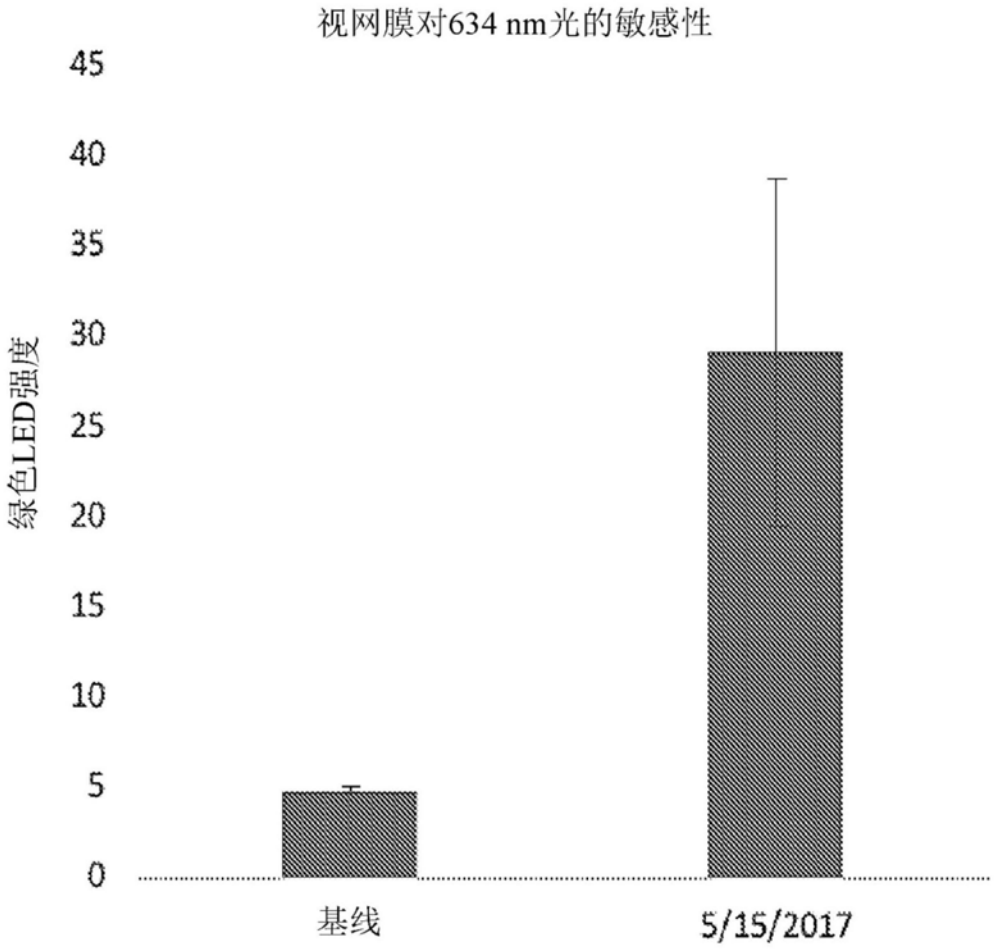

图2

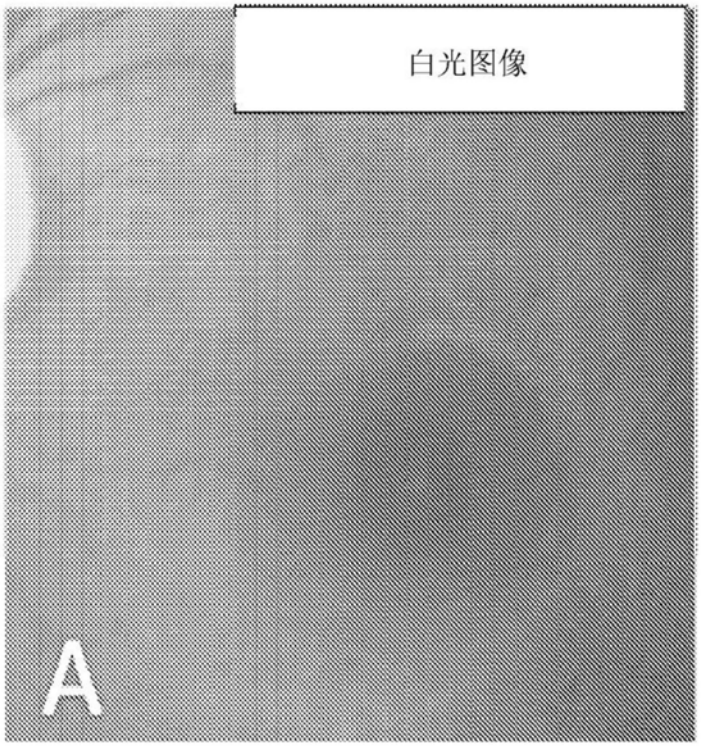

图3A

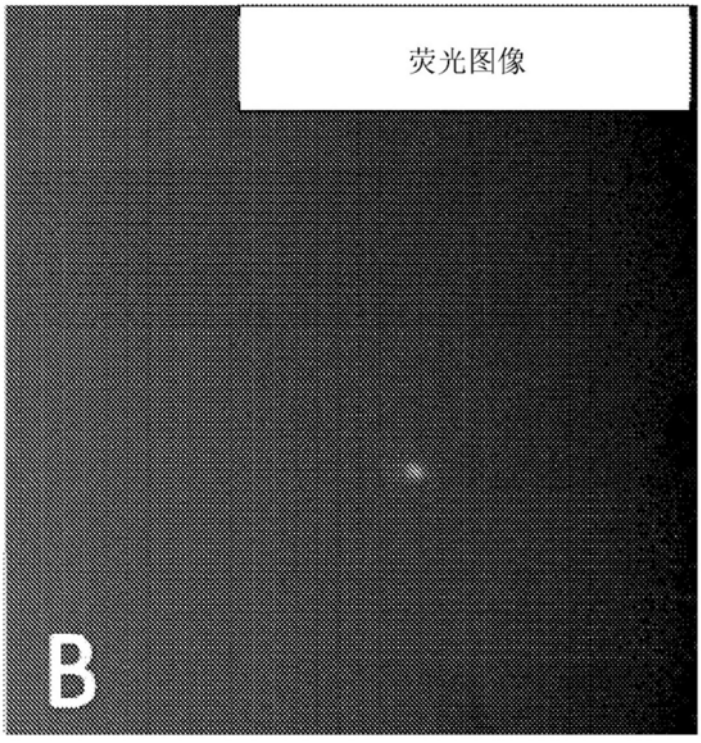

图3B

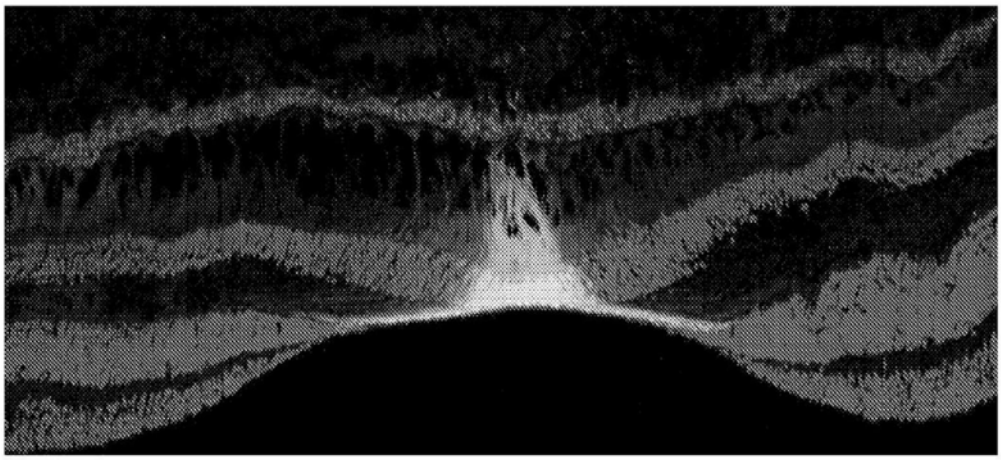

图4

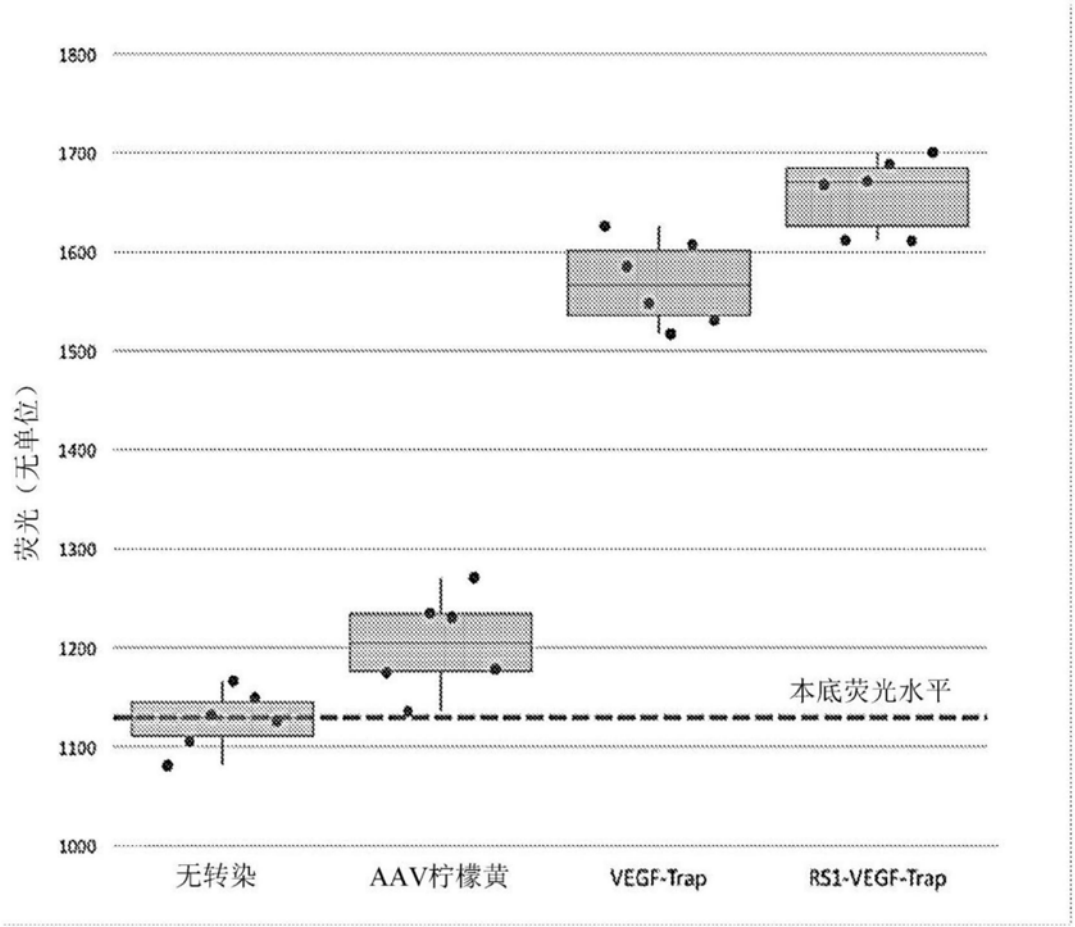
1800
1700
1600
1500
1400
1300
1200
1100
1000
荧光（无单位）
本底荧光水平
无转染
AAV柠檬黄
VEGF-Trap
RS1-VEGF-Trap

图5

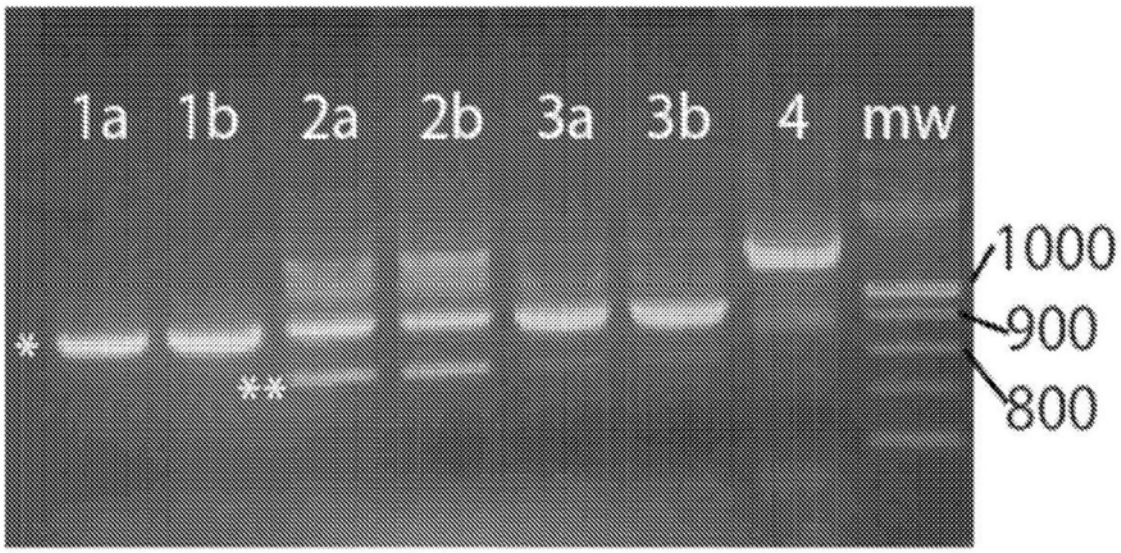

图6

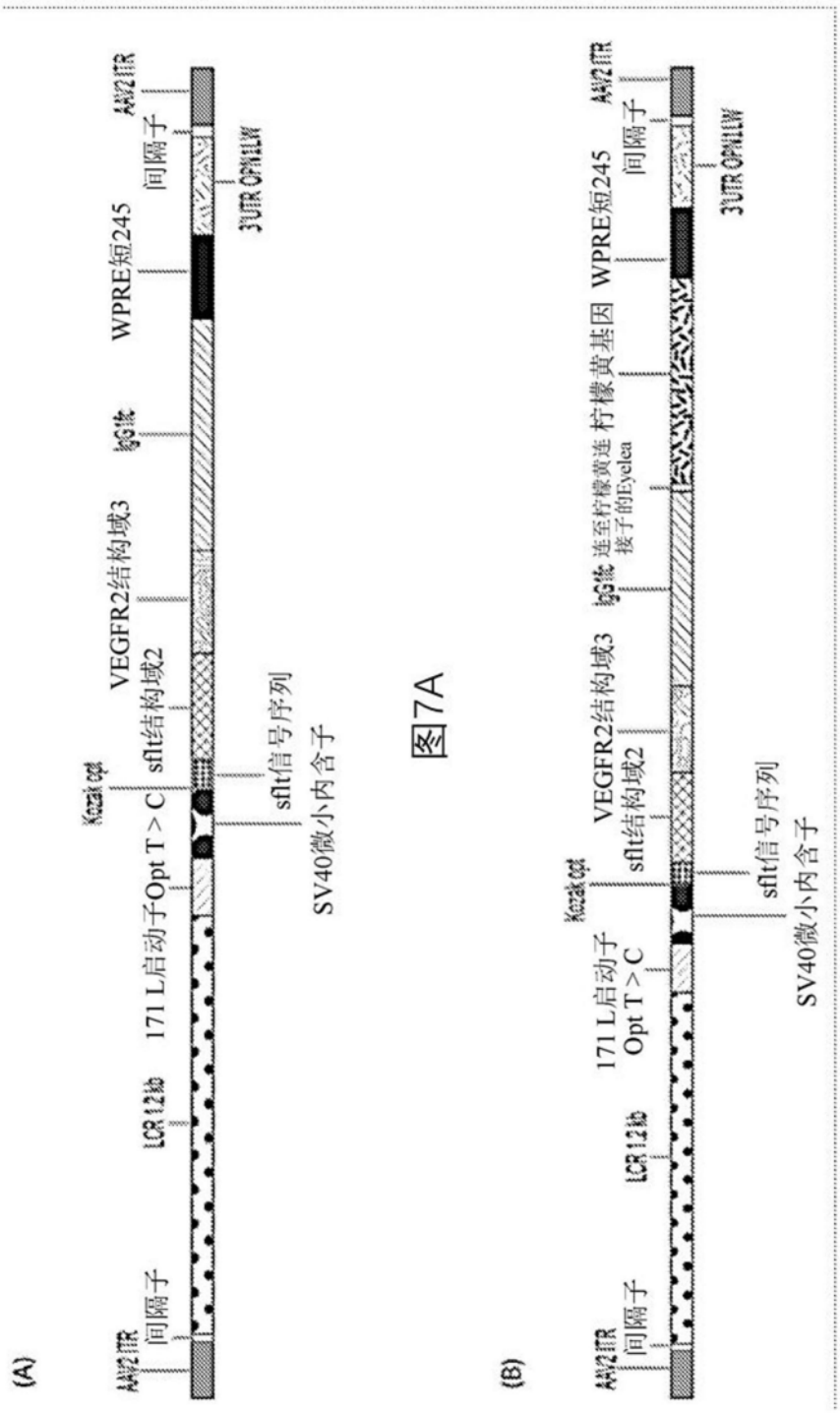
(A)
AAV2 ITR
间隔子
LCR 12kb
171 L启动子Opt T > C
Kozak opt
sflt结构域2
VEGFR2结构域3
IgG1κ
WPRE短245
AAV2 ITR
间隔子
3'UTR SPA/LUW
sflt信号序列
SV40微小内含子
(B)
AAV2 ITR
间隔子
LCR 12kb
171 L启动子
Opt T > C
Kozak opt
VEGFR2结构域2
sflt结构域2
VEGFR2结构域3
IgG1κ
连至柠檬黄连接子的Eyelea
柠檬黄基因
WPRE短245
AAV2 ITR
间隔子
3'UTR SPA/LUW
sflt信号序列
SV40微小内含子

图7A

图7B

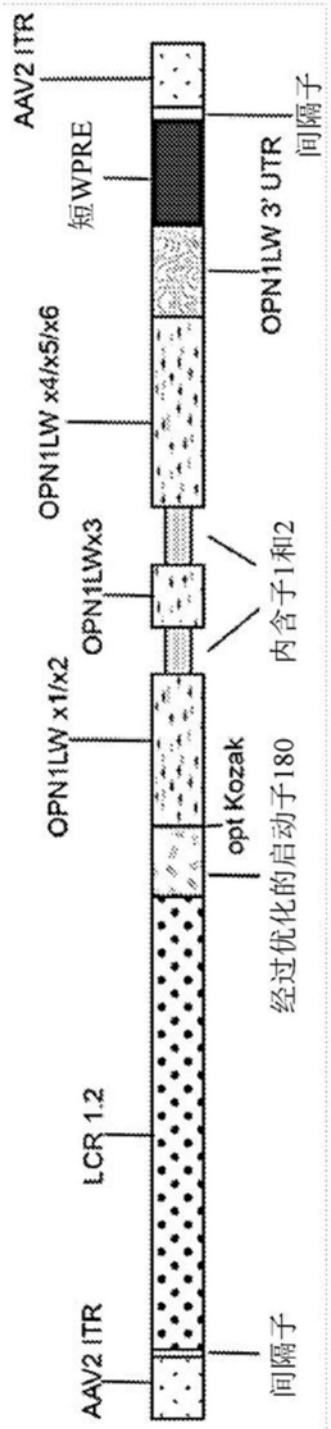
AAV2 ITR
短WPRE
间隔子
OPN1LW 3' UTR
OPN1LW x4/x5/x6
OPN1LWx3
内含子1和2
OPN1LW x1/x2
opt Kozak
经过优化的启动子180
LCR 1.2
AAV2 ITR
间隔子

图8

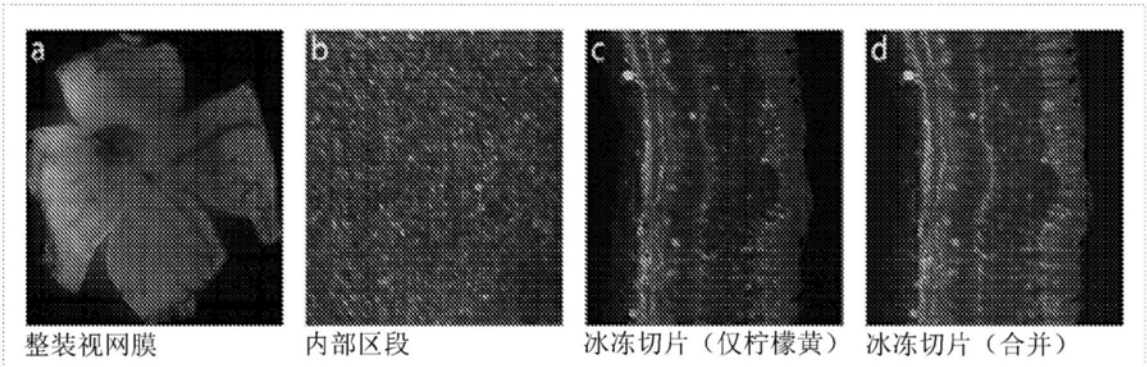

图9