



(12) 发明专利申请

(10) 申请公布号 CN 115768437 A

(43) 申请公布日 2023. 03. 07

(21) 申请号 202180044091.7

S-J·洪 A·阿库派

(22) 申请日 2021.04.21

(74) 专利代理机构 北京市金杜律师事务所

11256

(30) 优先权数据

专利代理师 陈文平 袁泉

63/013,456 2020.04.21 US

63/020,473 2020.05.05 US

(85) PCT国际申请进入国家阶段日

(51) Int.Cl.

A61K 31/7088 (2006.01)

A61K 31/7115 (2006.01)

A61K 39/00 (2006.01)

G12N 15/86 (2006.01)

2022.12.20

(86) PCT国际申请的申请数据

PCT/US2021/028486 2021.04.21

(87) PCT国际申请的公布数据

W02021/216776 EN 2021.10.28

(71) 申请人 磨石生物公司

地址 美国加利福尼亚州

(72) 发明人 K·朱斯 A·R·拉帕波特

权利要求书30页 说明书74页

C·D·斯卡伦 L·吉特林

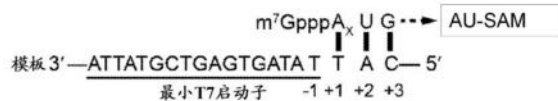
序列表221页 附图7页

(54) 发明名称

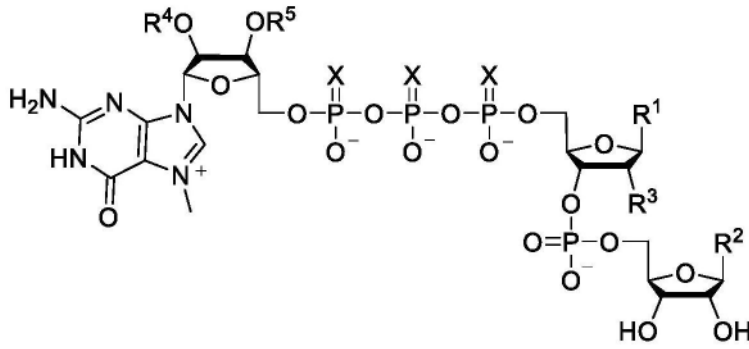
加帽化合物、组合物及其使用方法

(57) 摘要

本公开尤其包括可用作RNA核苷酸的5'帽的非天然核苷酸。本公开尤其还包括使用包括非天然核苷酸作为5'帽的递送和疫苗RNA核苷酸组合物的组合物和方法。



1. 一种式 (I) 化合物



(I),

或其药学上可接受的盐,

其中

R¹是核苷;

R²是核苷;

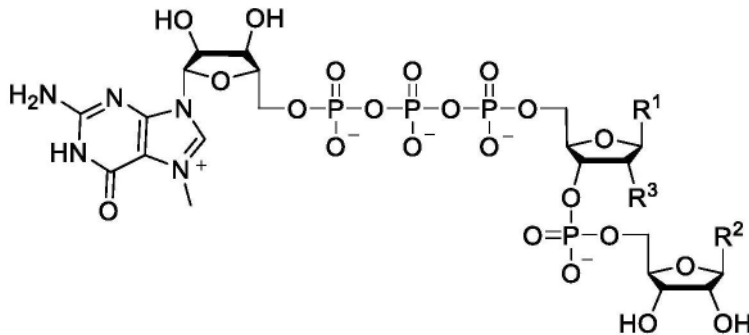
R³是卤素、任选地经取代的C₁-C₃烷基或经取代的C₁-C₃烷氧基;

R⁴是氢或任选地经取代的C₁-C₃脂族;

R⁵是氢或任选地经取代的C₁-C₃脂族;并且

每个X独立地是O或S,并且

任选地,其中所述化合物具有式(I-1):



(I-1),

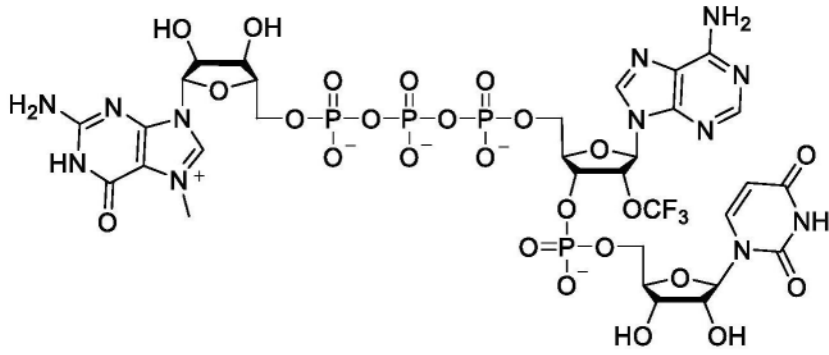
或其药学上可接受的盐。

2. 如权利要求1所述的化合物,其中R¹是腺嘌呤,任选地其中R¹是N6-甲基化腺嘌呤。

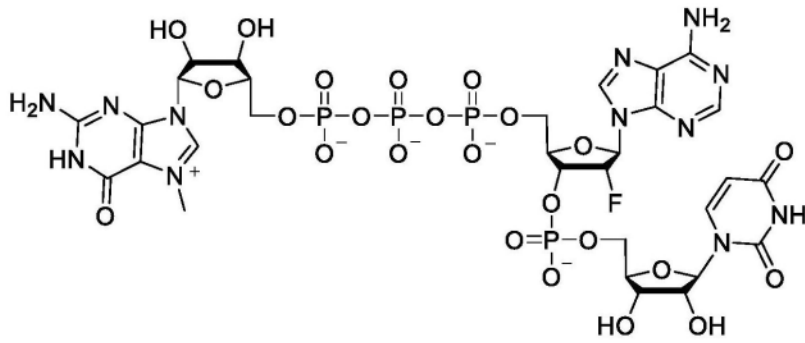
3. 如权利要求1或2所述的化合物,其中R²是尿嘧啶。

4. 如权利要求1-3中任一项所述的化合物,其中R³选自由氟、-CF₃、-OCF₃和-OCH₂CH₂OCH₃组成的组。

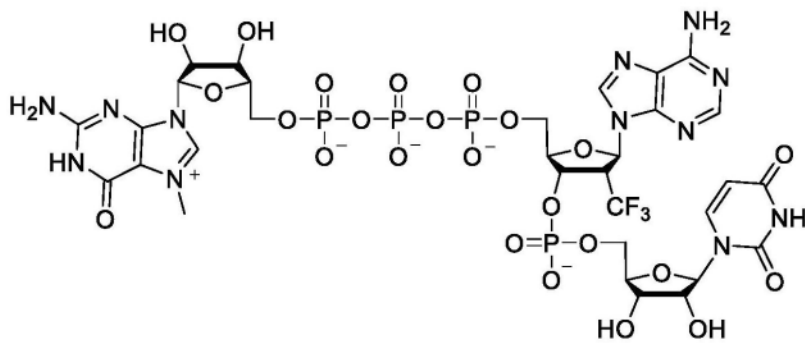
5. 如权利要求1-4中任一项所述的化合物,其中所述化合物选自由以下组成的组:



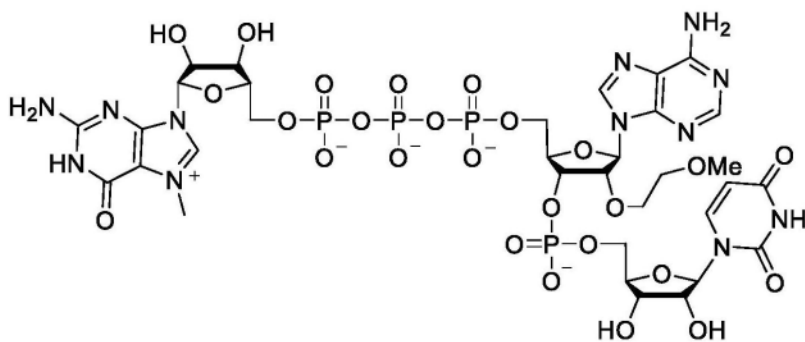
1、



2、



3、



4、

及其药学上可接受的盐。

6. 一种刺激免疫应答的方法, 任选地其中所述免疫应答治疗癌症, 所述方法包括向有需要的患者施用RNA寡核苷酸, 其中所述RNA寡核苷酸包含如权利要求1-5中任一项所述的化合物。

7. 如权利要求6所述的方法, 其中所述癌症选自由以下组成的组: 肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性髓细胞白血病、成人急性淋巴母细胞白血病、慢性髓细胞白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

8. 如权利要求6所述的方法, 其中所述癌症是实体瘤。

9. 如权利要求6或8所述的方法, 其中所述癌症选自由以下组成的组: MSS-CRC、NSCLC和PDA。

10. 如权利要求6或8所述的方法, 其中所述癌症选自由以下组成的组: 微卫星稳定结直肠癌(MSS-CRC)、非小细胞肺癌(NSCLC)、胰腺导管腺癌(PDA)和胃食管腺癌(GEA)。

11. 一种免疫或治疗感染的方法, 所述方法包括向有需要的患者施用RNA寡核苷酸, 其中所述RNA寡核苷酸包含如权利要求1-5中任一项所述的化合物。

12. 如权利要求11所述的方法, 其中所述感染是细菌感染。

13. 如权利要求11所述的方法, 其中所述感染是真菌感染。

14. 如权利要求11所述的方法, 其中所述感染是病毒感染。

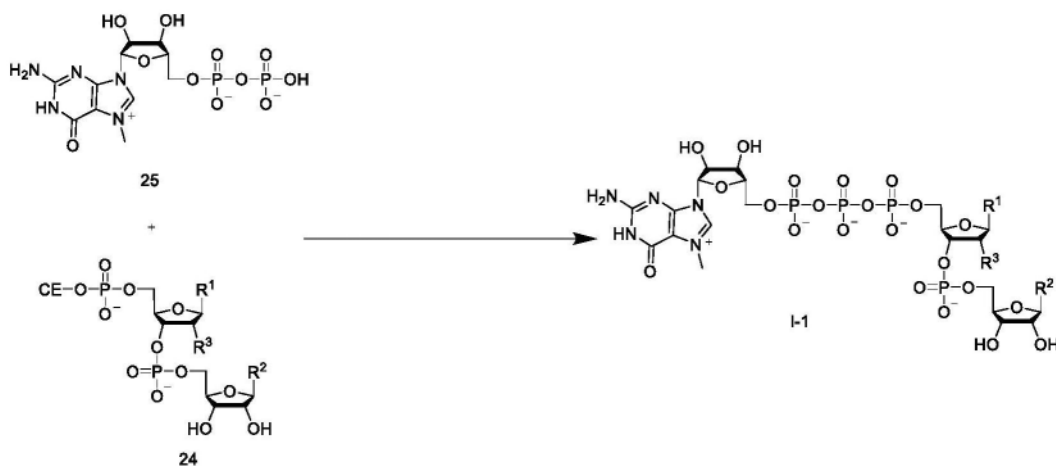
15. 如权利要求14所述的方法, 其中所述病毒感染是HIV感染。

16. 一种包含起始加帽寡核苷酸引物和DNA模板的复合物, 其中所述起始加帽寡核苷酸引物包含如权利要求1-5中任一项所述的化合物,

其中所述DNA模板包含包括转录起始位点的启动子区, 所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸; 并且

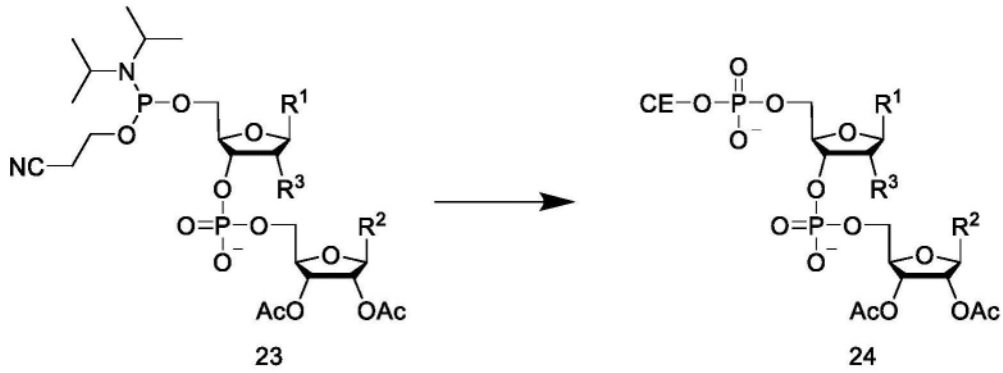
其中所述起始加帽寡核苷酸引物至少在核苷酸位置+1和+2处与所述DNA模板杂交。

17. 一种用于制备如权利要求1-5中任一项所述的化合物的方法, 所述方法包括以下步骤:



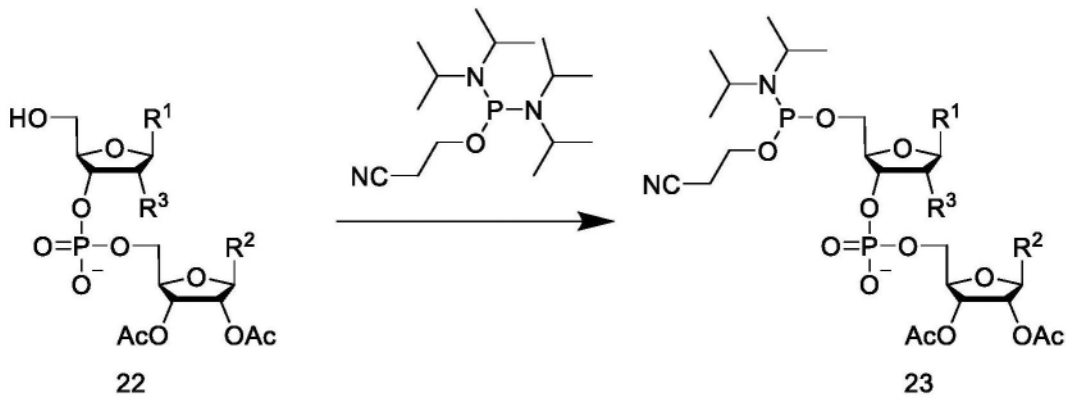
18. 如权利要求17所述的方法, 其中所述步骤包括用三氟乙酸酐和三乙胺处理24和25, 并且还包括用N-甲基咪唑处理所得混合物。

19. 如权利要求17所述的方法, 所述方法还包括以下步骤:

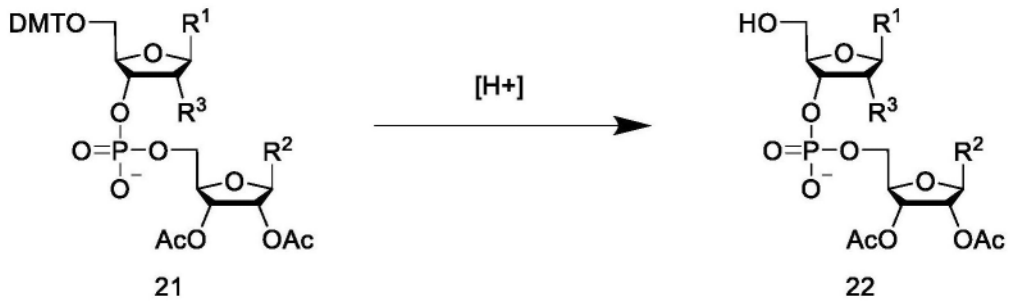


20. 如权利要求19所述的方法,其中所述步骤包括用 I_2 和水处理化合物23。

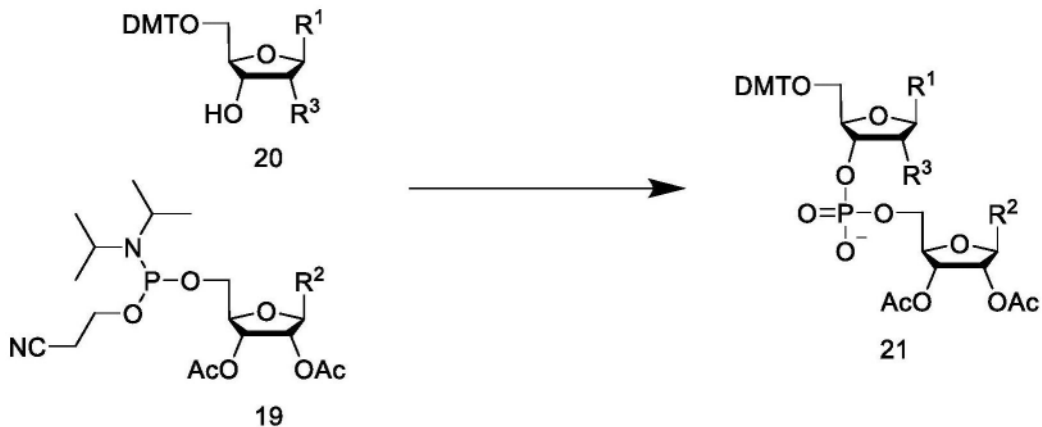
21. 如权利要求19所述的方法,所述方法还包括以下步骤:



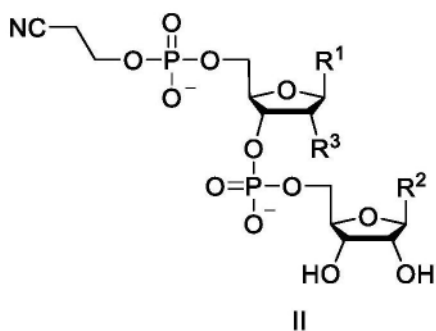
22. 如权利要求21所述的方法,所述方法还包括以下步骤:



23. 如权利要求22所述的方法,所述方法还包括以下步骤:



24. 一种式II化合物:



或其药学上可接受的盐，

其中

R¹是核苷；

R²是核苷；并且

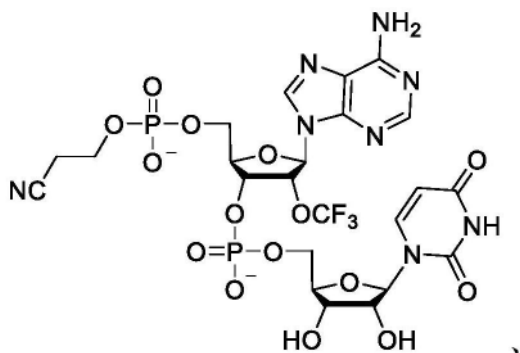
R³是卤素、任选地经取代的C₁-C₃烷基或经取代的C₁-C₃烷氧基。

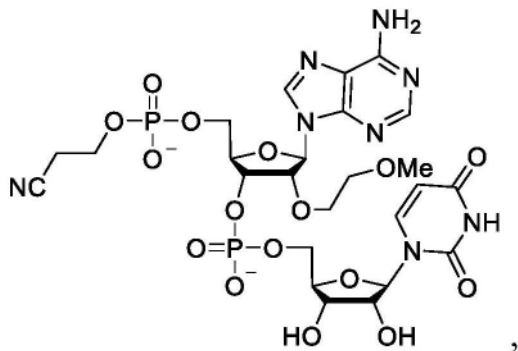
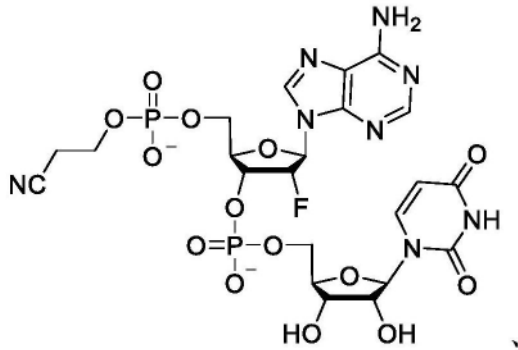
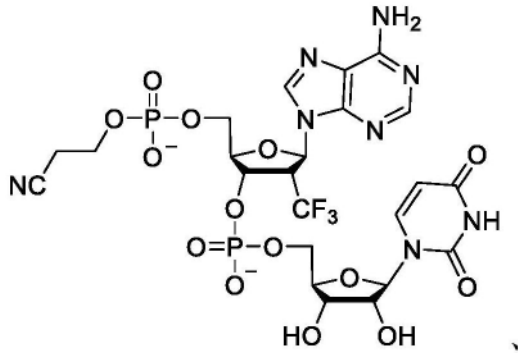
25. 如权利要求24所述的化合物，其中R¹是腺嘌呤，任选地其中R¹是N6-甲基化腺嘌呤。

26. 如权利要求24或25所述的化合物，其中R²是尿嘧啶。

27. 如权利要求24-26中任一项所述的化合物，其中R³选自由氟、-CF₃、-OCF₃和-OCH₂CH₂OCH₃组成的组。

28. 如权利要求24-26中任一项所述的化合物，其中所述化合物选自由以下组成的组：





及其药学上可接受的盐。

29. 一种自扩增表达系统，

其中所述自扩增表达系统包含自扩增骨架，其中所述自扩增骨架包含自我复制的RNA病毒的一种或多种多核苷酸序列；并且

其中所述自扩增表达系统包含核酸序列，其中每个元件从5'至3'由下式所述连接：

$m^7G\text{-ppp-N}_1\text{-N}_2\text{-N}_v$ ，其中

m^7G 是7-甲基鸟苷酸 (m^7G) 帽，

ppp是三磷酸根桥，

N_1 是所述自扩增骨架的第一个核苷酸，对应于所述自我复制的RNA病毒的第一个内源5'核苷酸，

N_2 是所述自扩增骨架的第二个核苷酸，对应于所述自我复制的RNA病毒的第二个内源5'核苷酸，并且

N_v 包含 (1) 所述自扩增骨架的一种或多种额外核酸序列，和 (2) 包含至少一种用于递送的外源核酸序列的盒，任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列，任

选地其中所述多肽编码核酸序列是抗原编码核酸序列,并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中。

30. 如权利要求29所述的组合物;

其中用于递送所述自扩增表达系统的所述组合物包含:

(A) 所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA (SAM) 载体,其中所述一种或多种SAM载体包含:

(a) 所述自扩增骨架,其中所述自扩增骨架包含:

(i) 至少一种启动子核苷酸序列,

(ii) 至少一种聚腺苷酸化(聚(A))序列,和

(b) 所述盒,任选地其中所述盒包含以下中的一者或多者:

(i) 至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:

a. 表位编码核酸序列,任选地包含:

(1) 使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2) 编码感染性疾病生物体肽的核酸序列,所述肽选自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,

b. 任选的5'接头序列,和

c. 任选的3'接头序列;

(ii) 可操作地连接到所述至少一种抗原编码核酸序列的第二启动子核苷酸序列;或

(iii) 任选地,至少一种第二聚(A)序列,其中所述第二聚(A)序列是所述自我复制的RNA病毒的原生聚(A)序列或外源聚(A)序列;和

(B) 任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

31. 如权利要求29所述的组合物,

其中用于递送所述自扩增表达系统的所述组合物包含:

(A) 所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA (SAM) 载体,其中所述一种或多种SAM载体包含:

(a) 所述自扩增骨架,其中所述自扩增骨架包含以SEQ ID NO:6所示的核酸序列,其中所述自扩增骨架序列包含亚基因组启动子核苷酸序列和聚(A)序列,其中所述亚基因组启动子序列对于所述自我复制的RNA病毒是内源的,其中所述聚(A)序列对于所述自我复制的RNA病毒骨架是内源的;和

(b) 整合在所述亚基因组启动子核苷酸序列与所述聚(A)序列之间的盒,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列,并且任选地其中所述盒包含至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:

a. 表位编码核酸序列,任选地包含:(1) 使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2) 编码感染性疾病生物体肽的核酸序列,所述肽选自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,

b. 任选的5'接头序列,和

c. 任选的3'接头序列;和

(B) 任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

32. 如权利要求29-31中任一项所述的组合物,其中N₁是修饰的核苷酸,任选地其中所述

修饰的核苷酸包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

33. 如权利要求29-31中任一项所述的组合物,其中 N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

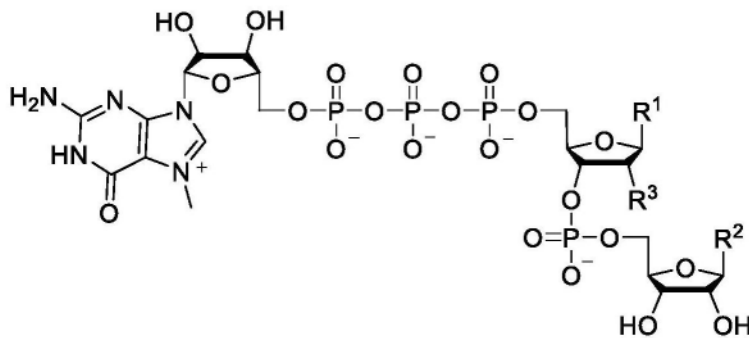
34. 如权利要求29-31中任一项所述的组合物,其中 N_1 和 N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸各自独立地包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

35. 如权利要求29-34中任一项所述的组合物,其中 N_1 是腺苷或修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

36. 如权利要求29-35中任一项所述的组合物,其中 N_2 是尿苷或修饰的尿苷,任选地其中所述修饰的尿苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

37. 如权利要求29-36中任一项所述的组合物,其中 N_1 是修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖,并且 N_2 是尿苷。

38. 如权利要求29-37中任一项所述的组合物,其中 $m^7G\text{-ppp-}N_1\text{-}N_2$ 由式(I-1)表示:



(I-1),

或其药学上可接受的盐,

其中

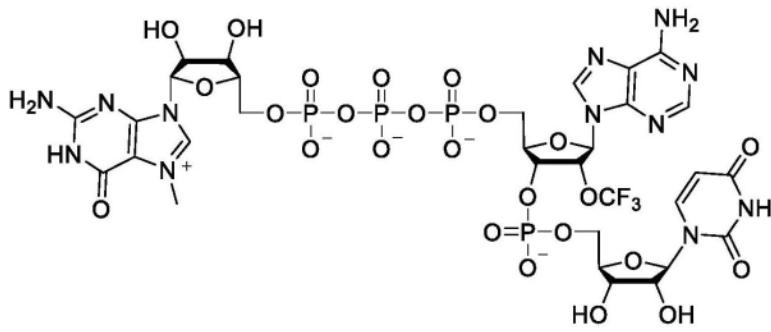
其中 R^1 是核苷,任选地其中 R^1 是腺嘌呤,任选地其中 R^1 是N6-甲基化腺嘌呤;

R^2 是核苷,任选地其中 R^2 是尿嘧啶;并且

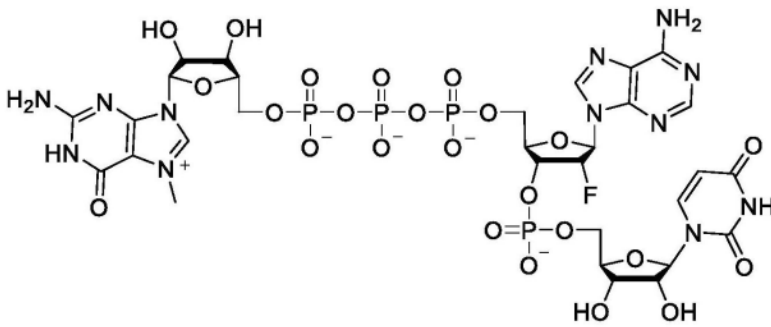
R^3 是卤素、任选地经取代的 $C_1\text{-}C_3$ 烷基或经取代的 $C_1\text{-}C_3$ 烷氧基。

39. 如权利要求38所述的组合物,其中 R^3 选自由氟、 $-\text{CF}_3$ 、 $-\text{OCF}_3$ 和 $-\text{OCH}_2\text{CH}_2\text{OCH}_3$ 组成的组。

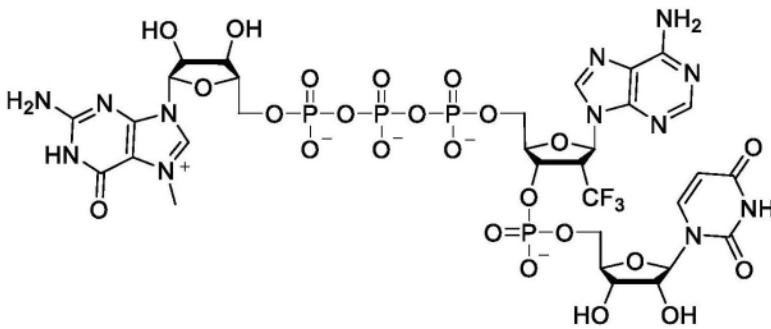
40. 如权利要求38或39所述的组合物,其中 $m^7G\text{-ppp-}N_1\text{-}N_2$ 由选自由以下组成的组的式表示:



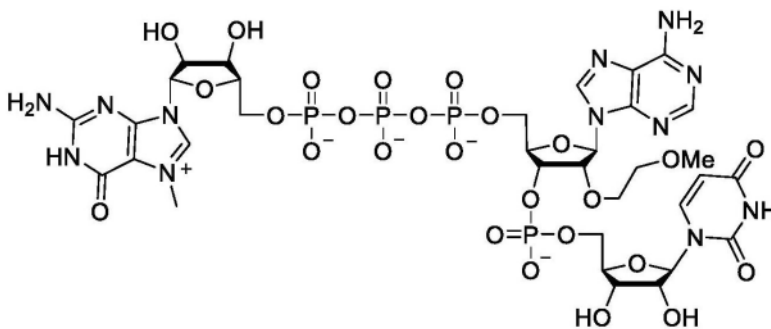
1、



2、



3、



4、

及其药学上可接受的盐。

41. 如权利要求29-40中任一项所述的组合物,其中所述自扩增表达系统是通过体外转

录产生的。

42. 如权利要求41所述的组合物,其中所述体外转录过程包括使用包含如权利要求29-40中任一项所述的 $m^7G\text{-ppp-N}_1\text{-N}_2$ 的起始加帽寡核苷酸。

43. 一种包含起始加帽寡核苷酸引物和DNA模板的复合物,其中所述起始加帽寡核苷酸引物包含如上述权利要求中任一项所述的 $m^7G\text{-ppp-N}_1\text{-N}_2$,

其中所述DNA模板从5'至3'包含:

(A) 包含转录起始位点的RNA转录启动子区,所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸,和

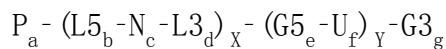
(B) 可操作地连接到所述RNA转录启动子区的包含如上述权利要求中任一项所述的 $N_1\text{-N}_2\text{-N}_v$ 的序列。

44. 如权利要求43所述的复合物,其中所述RNA转录启动子区包含T7启动子序列,任选地其中所述T7启动子序列是核苷酸序列TAATACGACTCACTATA或TAATACGACTCACTATT;SP6启动子序列,任选地其中所述SP6启动子序列是核苷酸序列ATTTAGGTGACTATA;或K11 RNAP启动子序列,任选地其中所述K11 RNAP启动子序列是核苷酸序列AATTAGGGCAGACTATA。

45. 如权利要求43或44所述的复合物,其中所述DNA模板包含以SEQ ID NO:57所示的序列,并且

其中所述盒被插入在如SEQ ID NO:6的序列中所示的位置7544处以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

46. 如权利要求29-45中任一项所述的组合物,其中用于递送所述自扩增表达系统的所述组合物中的所述盒的每个元件的有序序列描述于下式,从5'至3'包含:



其中P包含第二启动子核苷酸序列,其中a=0或1,

N包含表位编码核酸序列之一,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,其中c=1,

L5包含5'接头序列,其中b=0或1,

L3包含3'接头序列,其中d=0或1,

G5包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中e=0或1,

G3包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中g=0或1,

U包含至少一种MHC II类表位编码核酸序列之一,其中f=1,

X=1至400,其中对于每个X,相应的 N_c 是MHC I类表位编码核酸序列,并且

Y=0、1或2,其中对于每个Y,相应的 U_f 是MHC II类表位编码核酸序列。

47. 如权利要求46所述的组合物,其中对于每个X,相应的 N_c 是不同的MHC I类表位编码核酸序列。

48. 如权利要求46或47所述的组合物,其中对于每个Y,相应的 U_f 是不同的MHC II类表位编码核酸序列。

49. 如权利要求46-48中任一项所述的组合物,其中

a=0, b=1, d=1, e=1, g=1, h=1, X=10, Y=2,

所述至少一种启动子核苷酸序列是由所述自扩增骨架提供的单个亚基因组启动子核苷酸序列,

所述至少一种聚腺苷酸化聚(A)序列是由所述自扩增骨架提供的至少80个连续A核苷酸的聚(A)序列,

所述盒整合在所述亚基因组启动子核苷酸序列和所述聚(A)序列之间,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列和所述聚(A)序列,

每个N编码长度为7-15个氨基酸的MHC I类表位,

L5是编码所述MHC I类表位的原生N末端氨基酸序列的原生5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,

L3是编码所述MHC I类表位的原生C末端氨基酸序列的原生3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,

U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一者,

所述自扩增骨架是以SEQ ID NO:6所示的序列,并且

所述MHC I类表位编码核酸序列中的每一者编码长度为13至25个氨基酸的多肽。

50. 如权利要求29-45中任一项所述的组合物,其中用于递送的所述至少一种外源核酸序列包含多肽编码核酸序列。

51. 如权利要求50所述的组合物,其中所述多肽编码核酸序列编码所述抗原编码核酸序列。

52. 如权利要求51所述的组合物,其中所述抗原编码核酸序列包含MHC I类表位、MHC II类表位、能够刺激B细胞应答的表位或其组合。

53. 如权利要求51或52所述的组合物,其中所述抗原编码核酸序列包含编码全长蛋白质、蛋白质亚基、蛋白质结构域或其组合的序列。

54. 如权利要求50所述的组合物,其中所述多肽编码核酸序列编码全长蛋白质或其功能部分。

55. 如权利要求54所述的组合物,其中所述全长蛋白质或其功能部分选自由以下组成的组:抗体、细胞因子、嵌合抗原受体(CAR)、T细胞受体和基因组编辑系统核酸酶。

56. 如权利要求29-45中任一项所述的组合物,其中用于递送的所述至少一种外源核酸序列包含至少一种包含非编码核酸序列的核酸序列。

57. 如权利要求56所述的组合物,其中所述非编码核酸序列是RNA干扰(RNAi)多核苷酸或基因组编辑系统多核苷酸。

58. 如上述权利要求中任一项所述的组合物,其中所述LNP包含选自由以下组成的组的脂质:可电离氨基脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。

59. 如上述权利要求中任一项所述的组合物,其中所述LNP包含可电离氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。

60. 如权利要求58或59所述的组合物,其中所述可电离氨基脂质包含MC3样(二亚油醇甲基-4-二甲基氨基丁酸酯)分子。

61. 如上述权利要求中任一项所述的组合物,其中所述LNP包封的表达系统具有60-140nm之间的直径。

62. 如上述权利要求中任一项所述的组合物,其中用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用。

63. 如上述权利要求中任一项所述的组合物,其中用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)施用。

64. 如权利要求29、30、32-48或58-63中任一项所述的组合物,其中所述盒整合在所述至少一种启动子核苷酸序列和所述至少一种聚(A)序列之间。

65. 如权利要求29、30、32-48或58-64中任一项所述的组合物,其中所述至少一种启动子核苷酸序列可操作地连接到所述盒。

66. 如权利要求29、30、32-48或58-65中任一项所述的组合物,其中所述一种或多种SAM载体包含一种或多种正链RNA载体。

67. 如权利要求29、30、32-48或58-65中任一项所述的组合物,其中所述一种或多种SAM载体包含一种或多种负链RNA载体。

68. 如权利要求67所述的组合物,其中所述一种或多种负链RNA载体包含麻疹病毒或弹状病毒的至少一种多核苷酸序列。

69. 如权利要求29、30、32-48或58-68中任一项所述的组合物,其中所述一种或多种SAM载体在哺乳动物细胞内自扩增。

70. 如权利要求29、30、32-48或58-69中任一项所述的组合物,其中所述自我复制的RNA病毒选自自由以下组成的组:甲病毒;黄病毒、麻疹病毒和弹状病毒。

71. 如权利要求29、30、32-48或58-69中任一项所述的组合物,其中所述自扩增骨架包含甲病毒的至少一种多核苷酸序列,任选地其中所述甲病毒选自自由以下组成的组:奥拉病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、塞姆利基森林病毒、辛德毕斯病毒和马雅罗病毒。

72. 如权利要求29、30、32-48或58-69中任一项所述的组合物,其中所述自扩增骨架包含委内瑞拉马脑炎病毒的至少一种核苷酸序列。

73. 如权利要求71或72所述的组合物,其中所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因。

74. 如权利要求71或72所述的组合物,其中所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列和聚(A)序列。

75. 如权利要求73或74所述的组合物,其中用于非结构蛋白介导的扩增的序列选自自由以下组成的组:甲病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、甲病毒3'UTR或其组合。

76. 如权利要求73-75中任一项所述的组合物,其中所述自扩增骨架不编码结构性病毒粒子蛋白衣壳E2和E1,任选地其中E1是全长E1,或不编码结构性病毒粒子蛋白衣壳E3、E2、6K。

77. 如权利要求76所述的组合物,其中所述盒被插入以代替所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯

病毒或所述马雅罗病毒的多核苷酸序列内的结构性病毒粒子蛋白。

78. 如权利要求71或72所述的组合物,其中所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列。

79. 如权利要求71或72所述的组合物,其中所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列,其还包含碱基对7544和11175之间的缺失。

80. 如权利要求79所述的组合物,其中所述自扩增骨架包含SEQ ID NO:6或SEQ ID NO:7所示的序列。

81. 如权利要求79或80所述的组合物,其中所述盒被插入在位置7544处以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

82. 如权利要求77-81所述的组合物,其中所述盒的插入提供了包含nsP1-4基因和至少一种核酸序列的多顺反子RNA的转录,其中所述nsP1-4基因和所述至少一种核酸序列在单独的开放阅读框中。

83. 如权利要求29、30、32-48或58-82中任一项所述的组合物,其中所述至少一种启动子核苷酸序列是由所述自我复制的RNA病毒编码的原生启动子核苷酸序列,任选地其中所述原生启动子核苷酸序列是亚基因组启动子核苷酸序列。

84. 如权利要求29、30、32-48或58-82中任一项所述的组合物,其中所述至少一种启动子核苷酸序列是外源RNA启动子。

85. 如权利要求29、30、32-48或58-84中任一项所述的组合物,其中所述第二启动子核苷酸序列是亚基因组启动子核苷酸序列。

86. 如权利要求29、30、32-48或58-84中任一项所述的组合物,其中所述第二启动子核苷酸序列包含多个亚基因组启动子核苷酸序列,其中每个亚基因组启动子核苷酸序列提供所述单独的开放阅读框中的一者或多者的转录。

87. 如上述权利要求中任一项所述的组合物,其中所述一种或多种SAM载体的大小各自为至少300nt。

88. 如上述权利要求中任一项所述的组合物,其中所述一种或多种SAM载体的大小各自为至少1kb。

89. 如上述权利要求中任一项所述的组合物,其中所述一种或多种SAM载体的大小各自为2kb。

90. 如上述权利要求中任一项所述的组合物,其中所述SAM载体的大小各自小于5kb。

91. 如权利要求29-48或58-90中任一项所述的组合物,其中所述至少一种抗原编码核酸序列包含两种或更多种抗原编码核酸序列。

92. 如权利要求91所述的组合物,其中每种抗原编码核酸序列彼此直接连接。

93. 如权利要求29-48或58-92中任一项所述的组合物,其中每种抗原编码核酸序列用编码接头的核酸序列连接到不同的抗原编码核酸序列。

94. 如权利要求93所述的组合物,其中所述接头连接两个MHC I类表位编码核酸序列或一个MHC I类表位编码核酸序列至一个MHC II类表位编码核酸序列。

95. 如权利要求94所述的组合物,其中所述接头选自由以下组成的组:(1)连续甘氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(2)连续丙氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)由哺

乳动物蛋白酶体有效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列；和(6)侧接源自同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或2-20个氨基酸残基的一种或多种原生序列。

96. 如权利要求93所述的组合物，其中所述接头连接两个MHC II类表位编码核酸序列或一个MHCII类序列至一个MHCI类表位编码核酸序列。

97. 如权利要求96所述的组合物，其中所述接头包含序列GPGPG。

98. 如权利要求29-48或58-97中任一项所述的组合物，其中所述抗原编码核酸序列可操作地或直接地连接到增强所述表位编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性的单独或连续的序列。

99. 如权利要求98所述的组合物，其中所述单独或连续的序列包含以下中的至少一者：泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如，所述泛素序列在位置76处含有Gly至Ala取代)、免疫球蛋白信号序列(例如，IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列；任选地其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

100. 如权利要求29-48或58-99中任一项所述的组合物，其中所述至少一种抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10种抗原编码核酸序列，任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。

101. 如权利要求29-48或58-99中任一项所述的组合物，其中所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列，任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。

102. 如权利要求29-48或58-99中任一项所述的组合物，其中所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列。

103. 如权利要求29-48或58-99中任一项所述的组合物，其中所述至少一种抗原编码核酸序列包含至少2-400种抗原编码核酸序列，并且其中所述抗原编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。

104. 如权利要求29-48或58-99中任一项所述的组合物，其中每种抗原编码核酸序列独立地包含至少2-10、2、3、4、5、6、7、8、9或10种表位编码核酸序列，任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。

105. 如权利要求29-48或58-99中任一项所述的组合物，其中每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列，任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。

106. 如权利要求29-48或58-99中任一项所述的组合物，其中每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列。

107. 如权利要求29-48或58-99中任一项所述的组合物，其中每种抗原编码核酸序列独立地包含至少2-400种表位编码核酸序列，并且其中所述表位编码核酸序列中的至少两者

编码由MHC I类呈递在细胞表面上的表位序列或其部分。

108. 如权利要求49所述的组合物,其中所述MHC I类表位中的至少两者由MHC I类呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上。

109. 如权利要求29-48或58-108中任一项所述的组合物,其中所述表位编码核酸序列包含至少一种MHC I类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为8至35个氨基酸,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸的多肽序列。

110. 如权利要求46-48或58-109中任一项所述的组合物,其中存在所述至少一种MHC II类表位编码核酸序列。

111. 如权利要求46-48或58-109中任一项所述的组合物,其中所述至少一种MHC II类表位编码核酸序列存在并包含至少一种MHC II类表位编码核酸序列,所述MHC II类表位编码核酸序列包含使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变。

112. 如权利要求29-48或58-111中任一项所述的组合物,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。

113. 如权利要求29-48或58-112中任一项所述的组合物,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在至少一种MHC II类表位编码核酸序列,并且其中所述至少一种MHC II类表位编码核酸序列包含至少一种通用MHC II类表位编码核酸序列,任选地其中所述至少一种通用序列包含破伤风类毒素和PADRE中的至少一者。

114. 如权利要求29、30、32-48或58-113中任一项所述的组合物,其中所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是诱导型的。

115. 如权利要求29、30、32-48或58-113中任一项所述的组合物,其中所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是非诱导型的。

116. 如权利要求29、30、32-48或58-115中任一项所述的组合物,其中所述至少一种聚(A)序列包含所述自我复制病毒原生的聚(A)序列。

117. 如权利要求29、30、32-48或58-115中任一项所述的组合物,其中所述至少一种聚(A)序列包含所述自我复制病毒外源的聚(A)序列。

118. 如权利要求29、30、32-48或58-117中任一项所述的组合物,其中所述至少一种聚(A)序列可操作地连接到所述至少一种核酸序列中的至少一者。

119. 如权利要求29、30、32-48或58-118中任一项所述的组合物,其中所述至少一种聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少110个或至少120个连续A核苷酸。

120. 如权利要求29、30、32-48或58-118中任一项所述的组合物,其中所述至少一种聚(A)序列是至少80个连续A核苷酸。

121. 如权利要求29-48或58-120中任一项所述的组合物,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,并且其中所述MHC I类表位编码核酸序列是通过执行以下步骤来选择的:

(a) 从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸

测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;

(b) 将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及

(c) 基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生所述MHC I类表位编码核酸序列。

122. 如权利要求49所述的组合物,其中所述MHCI类表位编码核酸序列中的每一者是通过执行以下步骤来选择:

(a) 从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;

(b) 将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及

(c) 基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生至少20种MHCI类表位编码核酸序列。

123. 如权利要求121所述的组合物,其中所述经选择表位集合的数目为2-20。

124. 如权利要求121-123中任一项所述的组合物,其中所述呈递模型表示以下两者之间的依赖性:

(a) 所述MHC等位基因中的特定一者和在肽序列的特定位置处的特定氨基酸的一对的存在;和

(b) 由所述对的所述MHC等位基因中的所述特定一者将在所述特定位置处包含所述特定氨基酸的所述肽序列呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性。

125. 如权利要求121-124中任一项所述的组合物,其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位被呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性增加的表位。

126. 如权利要求121-125中任一项所述的组合物,其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激肿瘤特异性或感染性疾病生物体特异性免疫应答的可能性增加的表位。

127. 如权利要求121-126中任一项所述的组合物,其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够由专职抗原呈递细胞(APC)呈递给初始T细胞的可能性增加的表位,任选地其中所述APC是树突状细胞(DC)。

128. 如权利要求121-127中任一项所述的组合物,其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位经由中心或外周耐受性受抑制的可能性降低的表位。

129. 如权利要求121-128中任一项所述的组合物,其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激对正常组织的自身免疫

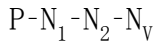
应答的可能性降低的表位。

130. 如权利要求121-129中任一项所述的组合物,其中外显子组或转录组核苷酸测序数据是通过肿瘤细胞或组织、受感染细胞或感染性疾病生物体进行测序而获得的。

131. 如权利要求130所述的组合物,其中所述测序是下一代测序(NGS)或任何大规模并行测序方法。

132. 一种产生自扩增表达系统的方法,其中所述方法包括以下步骤:

a) 提供DNA模板,其中每个元件从5'至3'由下式所述连接:



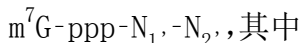
其中,P包含含有转录起始位点的RNA转录启动子区,所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸,

N_1 是自扩增骨架的第一个核苷酸,对应于自我复制的RNA病毒的第一个内源5'核苷酸,

N_2 是所述自扩增骨架的第二个核苷酸,对应于所述自我复制的RNA病毒的第二个内源5'核苷酸,并且

N_V 包含(1)所述自扩增骨架的一种或多种额外核酸序列,和(2)包含至少一种用于递送的外源核酸序列的盒,任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列,任选地其中所述多肽编码核酸序列是抗原编码核酸序列,并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中;

b) 提供起始加帽寡核苷酸引物,其中所述起始加帽寡核苷酸引物包含核酸序列,其中每个元件从5'至3'由下式所述连接:



m^7G 是7-甲基鸟苷酸(m^7G)帽,

ppp是三磷酸根桥,

N_1 是对应于所述DNA模板的 N_1 的核苷酸,并且

N_2 是对应于所述DNA模板的 N_2 的核苷酸,以及

c) 提供能够从所述RNA转录启动子区起始转录的RNA聚合酶

d) 在足以产生包含核酸序列的所述自扩增表达系统的条件下使所述DNA模板、所述起始加帽寡核苷酸引物和所述RNA聚合酶接触,其中每个元件从5'至3'由式 $m^7G-ppp-N_1-N_2-N_V$ 所述连接。

133. 如权利要求132所述的方法,其中所述RNA转录启动子区包含T7启动子序列,任选地其中所述T7启动子序列是核苷酸序列TAATACGACTCACTATA或TAATACGACTCACTATT;SP6启动子序列,任选地其中所述SP6启动子序列是核苷酸序列ATTTAGGTGACACTATA;或K11 RNAP启动子序列,任选地其中所述K11 RNAP启动子序列是核苷酸序列AATTAGGGCACACTATA。

134. 如权利要求132或133所述的方法,其中所述DNA模板包含以SEQ ID NO:57所示的序列,并且

其中所述盒被插入在如SEQ ID NO:6的序列中所示的位置7544处以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

135. 如权利要求132-134中任一项所述的方法,其中 N_1 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

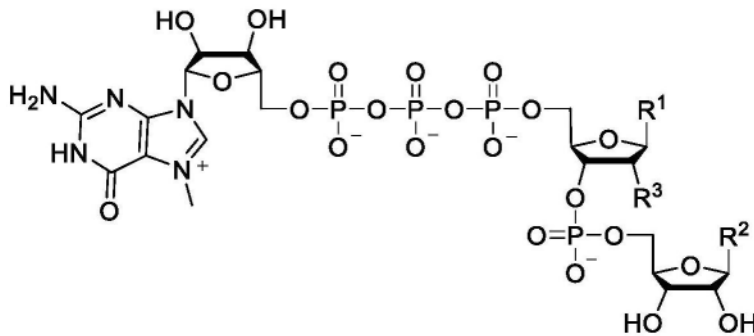
136. 如权利要求132-135中任一项所述的方法,其中 N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

137. 如权利要求132-136中任一项所述的方法,其中 N_1 是腺苷或修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

138. 如权利要求132-137中任一项所述的方法,其中 N_2 是尿苷或修饰的尿苷,任选地其中所述修饰的尿苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

139. 如权利要求132-138中任一项所述的方法,其中 N_1 是修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖,并且 N_2 是尿苷。

140. 如权利要求132-139中任一项所述的方法,其中所述起始加帽寡核苷酸引物由式(I-1)表示:



(I-1),

或其药学上可接受的盐,

其中

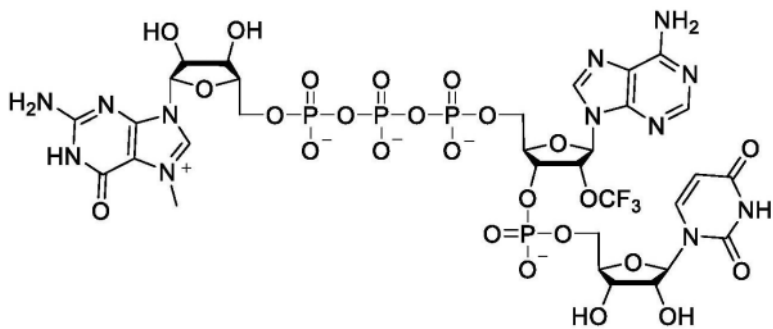
R^1 是核苷,任选地其中 R^1 是腺嘌呤,任选地其中 R^1 是N6-甲基化腺嘌呤;

R^2 是核苷,任选地其中 R^2 是尿嘧啶;并且

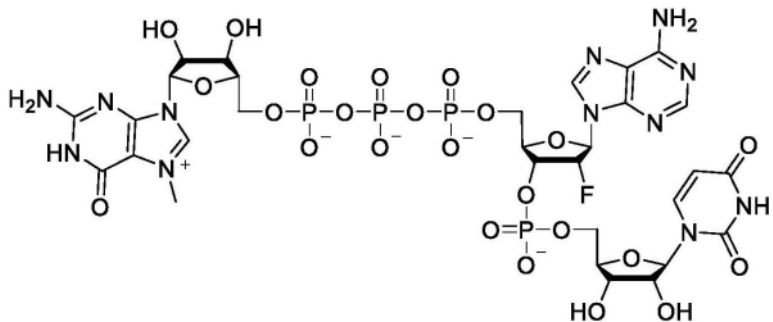
R^3 是卤素、任选地经取代的 C_1-C_3 烷基或经取代的 C_1-C_3 烷氧基。

141. 如权利要求140所述的方法,其中 R^3 选自由氟、 $-CF_3$ 、 $-OCF_3$ 和 $-OCH_2CH_2OCH_3$ 组成的组。

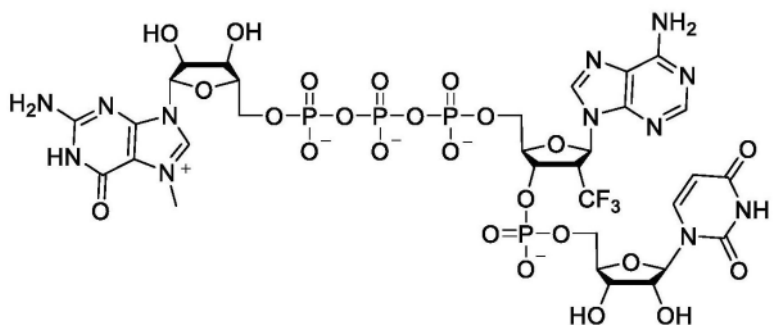
142. 如权利要求140或141所述的方法,其中所述起始加帽寡核苷酸引物由选自由以下组成的组的式表示:



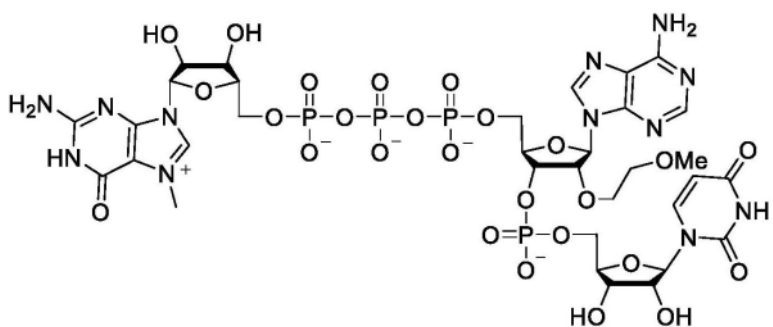
1、



2、



3、



4、

及其药学上可接受的盐。

143. 一种刺激受试者中的免疫应答的方法,所述方法包括向所述受试者施用用于递送

自扩增表达系统的组合物，

其中所述自扩增表达系统包含自扩增骨架，其中所述自扩增骨架包含自我复制的RNA病毒的一种或多种多核苷酸序列；并且

其中所述自扩增表达系统包含核酸序列，其中每个元件从5'至3'由下式所述连接：

$m^7G\text{-ppp-N}_1\text{-N}_2\text{-N}_v$ ，其中

m^7G 是7-甲基鸟苷酸 (m^7G) 帽，

ppp是三磷酸根桥，

N_1 是所述自扩增骨架的第一个核苷酸，对应于所述自我复制的RNA病毒的第一个内源5'核苷酸，

N_2 是所述自扩增骨架的第二个核苷酸，对应于所述自我复制的RNA病毒的第二个内源5'核苷酸，并且

N_v 包含 (1) 所述自扩增骨架的一种或多种额外核酸序列，和 (2) 包含至少一种用于递送的外源核酸序列的盒，任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列，任选地其中所述多肽编码核酸序列是抗原编码核酸序列，并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中。

144. 如权利要求143所述的方法；

其中用于递送所述自扩增表达系统的所述组合物包含：

(A) 所述自扩增表达系统，其中所述自扩增表达系统包含一种或多种自扩增mRNA (SAM) 载体，其中所述一种或多种SAM载体包含：

(a) 所述自扩增骨架，其中所述自扩增骨架包含：

(i) 至少一种启动子核苷酸序列，

(ii) 至少一种聚腺苷酸化 (聚 (A)) 序列，和

(b) 所述盒，任选地其中所述盒包含以下中的一者或多者：

(i) 至少一种抗原编码核酸序列，所述抗原编码核酸序列包含：

a. 表位编码核酸序列，任选地包含：(1) 使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变，或 (2) 编码感染性疾病生物体肽的核酸序列，所述肽选自以下组成的组：病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽，

b. 任选的5'接头序列，和

c. 任选的3'接头序列；

(ii) 可操作地连接到所述至少一种抗原编码核酸序列的第二启动子核苷酸序列；或

(iii) 任选地，至少一种第二聚 (A) 序列，其中所述第二聚 (A) 序列是所述自我复制的RNA病毒的原生聚 (A) 序列或外源聚 (A) 序列；和

(B) 任选地，脂质纳米粒子 (LNP)，其中所述LNP包封所述自扩增表达系统。

145. 如权利要求143所述的方法，

其中用于递送所述自扩增表达系统的所述组合物包含：

(A) 所述自扩增表达系统，其中所述自扩增表达系统包含一种或多种自扩增mRNA (SAM) 载体，其中所述一种或多种SAM载体包含：

(a) 所述自扩增骨架，其中所述自扩增骨架包含以SEQ ID NO:6所示的核酸序列，其中所述自扩增骨架序列包含亚基因组启动子核苷酸序列和聚 (A) 序列，其中所述亚基因组启

动子序列对于所述自我复制的RNA病毒是内源的,其中所述聚(A)序列对于所述自扩增骨架是内源的;和

(b) 整合在所述亚基因组启动子核苷酸序列与所述聚(A)序列之间的盒,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列,并且任选地其中所述盒包含至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:

a. 表位编码核酸序列,任选地包含:(1)使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2)编码感染性疾病生物体肽的核酸序列,所述肽选自自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,

b. 任选的5'接头序列,和

c. 任选的3'接头序列;和

(B) 任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

146. 如权利要求143-145中任一项所述的方法,其中 N_1 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

147. 如权利要求143-145中任一项所述的方法,其中 N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

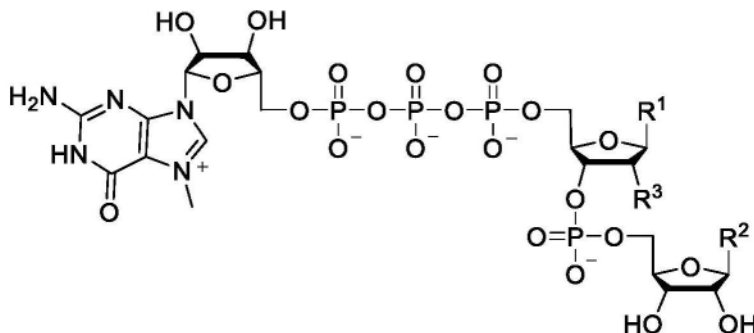
148. 如权利要求143-145中任一项所述的方法,其中 N_1 和 N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸各自独立地包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

149. 如权利要求143-148中任一项所述的方法,其中 N_1 是腺苷或修饰的腺苷,任选地其中所述修饰的腺苷包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

150. 如权利要求143-149中任一项所述的方法,其中 N_2 是尿苷或修饰的尿苷,任选地其中所述修饰的尿苷包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。

151. 如权利要求143-150中任一项所述的方法,其中 N_1 是修饰的腺苷,任选地其中所述修饰的腺苷包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖,并且 N_2 是尿苷。

152. 如权利要求143-151中任一项所述的方法,其中 $m^7G\text{-ppp-}N_1\text{-}N_2$ 由式(I-1)表示:



(I-1),

或其药学上可接受的盐，

其中

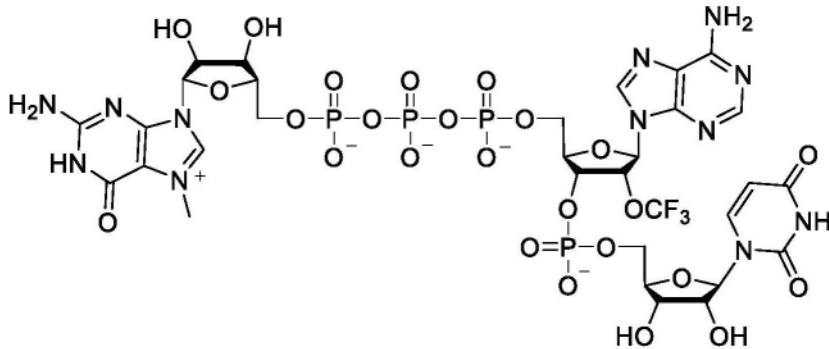
R^1 是核苷，任选地其中 R^1 是腺嘌呤，任选地其中 R^1 是N6-甲基化腺嘌呤；

R^2 是核苷，任选地其中 R^2 是尿嘧啶；并且

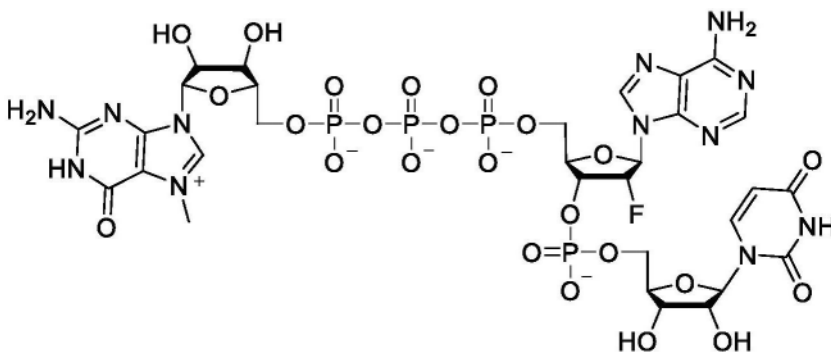
R^3 是卤素、任选地经取代的 C_1-C_3 烷基或经取代的 C_1-C_3 烷氧基。

153. 如权利要求152所述的方法，其中 R^3 选自由氟、 $-CF_3$ 、 $-OCF_3$ 和 $-OCH_2CH_2OCH_3$ 组成的组。

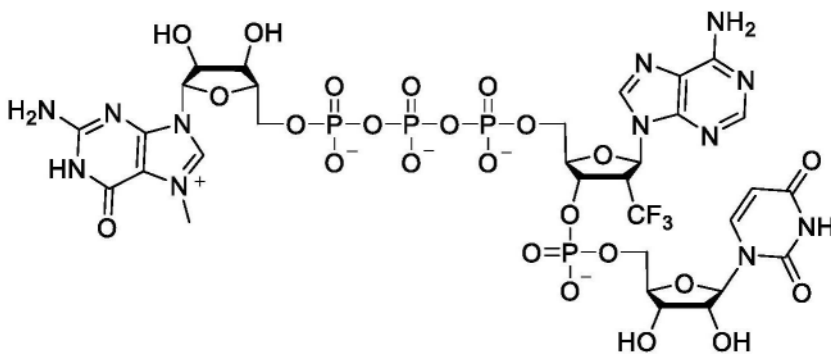
154. 如权利要求152或153所述的方法，其中 $m^7G\text{-ppp-N}_1\text{-N}_2$ 由选自由以下组成的组的式表示：



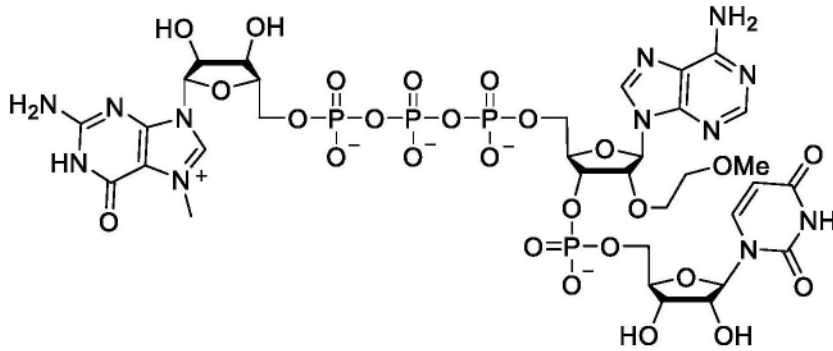
1、



2、



3、



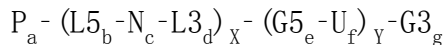
4,

及其药学上可接受的盐。

155. 如权利要求143-154中任一项所述的方法,其中所述自扩增表达系统是通过体外转录产生的。

156. 如权利要求155所述的方法,其中所述体外转录过程包括使用包含如权利要求143-154中任一项所述的 $m^7G\text{-ppp-}N_1\text{-}N_2$ 的起始加帽寡核苷酸。

157. 如权利要求143-156中任一项所述的方法,其中用于递送所述自扩增表达系统的所述组合物中的所述盒的每个元件的有序序列以下式描述,从5'至3'包含



其中P包含第二启动子核苷酸序列,其中 $a=0$ 或 1 ,

N包含表位编码核酸序列之一,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,其中 $c=1$,

L5包含5'接头序列,其中 $b=0$ 或 1 ,

L3包含3'接头序列,其中 $d=0$ 或 1 ,

G5包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中 $e=0$ 或 1 ,

G3包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中 $g=0$ 或 1 ,

U包含至少一种MHC II类表位编码核酸序列之一,其中 $f=1$,

$X=1$ 至 400 ,其中对于每个 X ,相应的 N_c 是MHC I类表位编码核酸序列,并且

$Y=0, 1$ 或 2 ,其中对于每个 Y ,相应的 U_f 是MHC II类表位编码核酸序列。

158. 如权利要求157所述的方法,其中对于每个 X ,相应的 N_c 是不同的MHC I类表位编码核酸序列。

159. 如权利要求157或158所述的方法,其中对于每个 Y ,相应的 U_f 是不同的MHC II类表位编码核酸序列。

160. 如权利要求157-159中任一项所述的方法,其中

$a=0, b=1, d=1, e=1, g=1, h=1, X=10, Y=2$,

所述至少一种启动子核苷酸序列是由所述自扩增骨架提供的单个亚基因组启动子核苷酸序列,

所述至少一种聚腺苷酸化聚(A)序列是由所述自扩增骨架提供的至少80个连续A核苷酸的聚(A)序列,

所述盒整合在所述亚基因组启动子核苷酸序列和所述聚(A)序列之间,其中所述盒可

操作地连接到所述亚基因组启动子核苷酸序列和所述聚 (A) 序列,

每个N编码长度为7-15个氨基酸的MHC I类表位,

L5是编码所述MHCI表位的原生N末端氨基酸序列的原生5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,

L3是编码所述MHCI表位的原生C末端氨基酸序列的原生3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,

U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一者,

所述自扩增骨架是以SEQ ID NO:6所示的序列,并且

所述MHC I类表位编码核酸序列中的每一者编码长度为13至25个氨基酸的多肽。

161. 如权利要求143-156中任一项所述的方法,其中用于递送的至少一种外源核酸序列包含所述多肽编码核酸序列。

162. 如权利要求161所述的方法,其中所述多肽编码核酸序列编码所述抗原编码核酸序列。

163. 如权利要求162所述的方法,其中所述抗原编码核酸序列包含MHCI类表位、MHCII类表位、能够刺激B细胞应答的表位或其组合。

164. 如权利要求162或163所述的方法,其中所述抗原编码核酸序列包含编码全长蛋白质、蛋白质亚基、蛋白质结构域或其组合的序列。

165. 如权利要求161所述的方法,其中所述多肽编码核酸序列编码全长蛋白质或其功能部分。

166. 如权利要求165所述的方法,其中所述全长蛋白质或其功能部分选自由以下组成的组:抗体、细胞因子、嵌合抗原受体 (CAR)、T细胞受体和基因组编辑系统核酸酶。

167. 如权利要求143-156中任一项所述的方法,其中用于递送的至少一种外源核酸序列包含至少一种包含非编码核酸序列的核酸序列。

168. 如权利要求167所述的方法,其中所述非编码核酸序列是RNA干扰 (RNAi) 多核苷酸或基因组编辑系统多核苷酸。

169. 如上述权利要求中任一项所述的方法,其中所述LNP包含选自由以下组成的组的脂质:可电离氨基脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。

170. 如上述权利要求中任一项所述的方法,其中所述LNP包含可电离氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。

171. 如权利要求169或170所述的方法,其中所述可电离氨基脂质包含MC3样(二亚油醇甲基-4-二甲基氨基丁酸酯)分子。

172. 如上述权利要求中任一项所述的方法,其中所述LNP包封的表达系统具有约60-140nm之间的直径。

173. 如上述权利要求中任一项所述的方法,其中用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用。

174. 如上述权利要求中任一项所述的方法,其中用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)施用。

175. 如权利要求143、144、146-159或169-174中任一项所述的方法,其中所述盒整合在

所述至少一种启动子核苷酸序列和所述至少一种聚(A)序列之间。

176. 如权利要求143、144、146-159或169-175中任一项所述的方法,其中所述至少一种启动子核苷酸序列可操作地连接到所述盒。

177. 如权利要求143、144、146-159或169-176中任一项所述的方法,其中所述一种或多种SAM载体包含一种或多种正链RNA载体。

178. 如权利要求143、144、146-159或169-177中任一项所述的方法,其中所述一种或多种SAM载体包含一种或多种负链RNA载体。

179. 如权利要求178所述的方法,其中所述一种或多种负链RNA载体包含麻疹病毒或弹状病毒的至少一种多核苷酸序列。

180. 如权利要求143、144、146-159或169-179中任一项所述的方法,其中所述一种或多种SAM载体在哺乳动物细胞内自扩增。

181. 如权利要求142、144、146-159或169-180中任一项所述的方法,其中所述自扩增骨架包含自我复制的RNA病毒的至少一种多核苷酸序列,所述自我复制的RNA病毒选自自由以下组成的组:甲病毒;黄病毒、麻疹病毒和弹状病毒。

182. 如权利要求142、144、146-159或169-180中任一项所述的方法,其中所述自扩增骨架包含甲病毒的至少一种多核苷酸序列,任选地其中所述甲病毒选自自由以下组成的组:奥拉病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、塞姆利基森林病毒、辛德毕斯病毒和马雅罗病毒。

183. 如权利要求143、144、146-159或169-180中任一项所述的方法,其中所述自扩增骨架包含委内瑞拉马脑炎病毒的至少一种核苷酸序列。

184. 如权利要求182或183所述的方法,其中所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因。

185. 如权利要求182或183所述的方法,其中所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列和聚(A)序列。

186. 如权利要求184或185所述的方法,其中用于非结构蛋白介导的扩增的序列选自自由以下组成的组:甲病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、甲病毒3'UTR或其组合。

187. 如权利要求184-186中任一项所述的方法,其中所述自扩增骨架不编码结构性病毒粒子蛋白衣壳E2和E1,任选地其中E1是全长E1,或不编码结构性病毒粒子蛋白衣壳E3、E2、6K。

188. 如权利要求187所述的方法,其中所述盒被插入以代替所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的多核苷酸序列内的结构性病毒粒子蛋白。

189. 如权利要求182或183所述的方法,其中所述委内瑞拉马脑炎病毒包含SEQ ID NO:

3或SEQ ID NO:5的序列。

190. 如权利要求182或183所述的方法,其中所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列,其还包含碱基对7544和11175之间的缺失。

191. 如权利要求190所述的方法,其中所述自扩增骨架包含以SEQ ID NO:6或SEQ ID NO:7所示的序列。

192. 如权利要求190或191所述的方法,其中所述盒被插入在位置7544处以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

193. 如权利要求188-192所述的方法,其中所述盒的插入提供了包含nsP1-4基因和至少一种核酸序列的多顺反子RNA的转录,其中所述nsP1-4基因和所述至少一种核酸序列在单独的开放阅读框中。

194. 如权利要求143、144、146-159或169-193中任一项所述的方法,其中所述至少一种启动子核苷酸序列是由所述自扩增骨架编码的原生启动子核苷酸序列,任选地其中所述原生启动子核苷酸序列是亚基因组启动子核苷酸序列。

195. 如权利要求143、144、146-159或169-193中任一项所述的方法,其中所述至少一种启动子核苷酸序列是外源RNA启动子。

196. 如权利要求143、144、146-159或169-195中任一项所述的方法,其中所述第二启动子核苷酸序列是亚基因组启动子核苷酸序列。

197. 如权利要求143、144、146-159或169-195中任一项所述的方法,其中所述第二启动子核苷酸序列包含多个亚基因组启动子核苷酸序列,其中每个亚基因组启动子核苷酸序列提供所述单独的开放阅读框中的一者或多者的转录。

198. 如上述权利要求中任一项所述的方法,其中所述一种或多种SAM载体的大小各自为至少300nt。

199. 如上述权利要求中任一项所述的方法,其中所述一种或多种SAM载体的大小各自为至少1kb。

200. 如上述权利要求中任一项所述的方法,其中所述一种或多种SAM载体的大小各自为2kb。

201. 如上述权利要求中任一项所述的方法,其中所述一种或多种SAM载体的大小各自小于5kb。

202. 如权利要求143-159或169-201中任一项所述的方法,其中所述至少一种抗原编码核酸序列包含两种或更多种抗原编码核酸序列。

203. 如权利要求202所述的方法,其中每种抗原编码核酸序列彼此直接连接。

204. 如权利要求143-159或169-203中任一项所述的方法,其中每种抗原编码核酸序列用编码接头的核酸序列连接到不同的抗原编码核酸序列。

205. 如权利要求204所述的方法,其中所述接头连接两个MHC I类表位编码核酸序列或一个MHC I类表位编码核酸序列至一个MHC II类表位编码核酸序列。

206. 如权利要求205所述的方法,其中所述接头选自由以下组成的组:(1)连续甘氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(2)连续丙氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)由哺乳动物蛋白酶体有效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;

和(6)侧接源自同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或2-20个氨基酸残基的一种或多种原生序列。

207.如权利要求204所述的方法,其中所述接头连接两个MHC II类表位编码核酸序列或一个MHCII类序列至一个MHCI类表位编码核酸序列。

208.如权利要求207所述的方法,其中所述接头包含序列GPGPG。

209.如权利要求143-159或169-208中任一项所述的方法,其中所述抗原编码核酸序列可操作地或直接地连接到增强所述表位编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性的单独或连续的序列。

210.如权利要求209所述的方法,其中所述单独或连续的序列包含以下中的至少一者:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,所述泛素序列在位置76处含有Gly至Ala取代)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

211.如权利要求143-159或169-210中任一项所述的方法,其中所述至少一种抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。

212.如权利要求143-159或169-210中任一项所述的方法,其中所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。

213.如权利要求143-159或169-210中任一项所述的方法,其中所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列。

214.如权利要求143-159或169-210中任一项所述的方法,其中所述至少一种抗原编码核酸序列包含至少2-400种抗原编码核酸序列,并且其中所述抗原编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。

215.如权利要求143-159或169-210中任一项所述的方法,其中每种抗原编码核酸序列独立地包含至少2-10、2、3、4、5、6、7、8、9或10种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。

216.如权利要求143-159或169-210中任一项所述的方法,其中每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。

217.如权利要求143-159或169-210中任一项所述的方法,其中每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列。

218.如权利要求143-159或169-210中任一项所述的方法,其中每种抗原编码核酸序列独立地包含至少2-400种表位编码核酸序列,并且其中所述表位编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。

219. 如160所述的方法,其中所述MHC I类表位中的至少两者由MHC I类呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上。

220. 如权利要求143-159或169-219中任一项所述的方法,其中所述表位编码核酸序列包含至少一种MHC I类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为8至35个氨基酸,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸的多肽序列。

221. 如权利要求157-159或169-220中任一项所述的方法,其中存在所述至少一种MHC II类表位编码核酸序列。

222. 如权利要求157-159或169-220中任一项所述的方法,其中所述至少一种MHC II类表位编码核酸序列存在并包含至少一种MHC II类表位编码核酸序列,所述MHC II类表位编码核酸序列包含使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变。

223. 如权利要求143-159或169-222中任一项所述的方法,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。

224. 如权利要求143-159或169-223中任一项所述的方法,其中所述表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在至少一种MHC II类表位编码核酸序列,并且其中所述至少一种MHC II类表位编码核酸序列包含至少一种通用MHC II类表位编码核酸序列,任选地其中所述至少一种通用序列包含破伤风类毒素和PADRE中的至少一者。

225. 如权利要求143、144、146-159或169-224中任一项所述的方法,其中所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是诱导型的。

226. 如权利要求143、144、146-159或169-224中任一项所述的方法,其中所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是非诱导型的。

227. 如权利要求143、144、146-159或169-226中任一项所述的方法,其中所述至少一种聚(A)序列包含所述自我复制RNA原生的聚(A)序列。

228. 如权利要求143、144、146-159或169-226中任一项所述的方法,其中所述至少一种聚(A)序列包含所述自我复制RNA外源的聚(A)序列。

229. 如权利要求143、144、146-159或169-228中任一项所述的方法,其中所述至少一种聚(A)序列可操作地连接到所述至少一种核酸序列中的至少一者。

230. 如权利要求143、144、146-159或169-229中任一项所述的方法,其中所述至少一种聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少110个或至少120个连续A核苷酸。

231. 如权利要求143、144、146-159或169-229中任一项所述的方法,其中所述至少一种聚(A)序列是至少80个连续A核苷酸。

232. 如权利要求143-159或169-231中任一项所述的方法,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,并且其中所述MHC I类表位编码核酸序列是通过执行以下步骤来选择:

(a) 从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽

序列的数据；

(b) 将每个表位的肽序列输入到呈递模型中，以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合，所述数值可能性集合已至少基于所接收的质谱数据进行鉴定；以及

(c) 基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合，所述经选择表位用于产生所述MHC I类表位编码核酸序列。

233. 如权利要求160所述的方法，其中所述MHC I类表位编码核酸序列中的每一者是通过执行以下步骤来选择：

(a) 从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者，其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据；

(b) 将每个表位的肽序列输入到呈递模型中，以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合，所述数值可能性集合已至少基于所接收的质谱数据进行鉴定；以及

(c) 基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合，所述经选择表位用于产生至少20种MHC I类表位编码核酸序列。

234. 如权利要求232所述的方法，其中所述经选择表位集合的数目为2-20。

235. 如权利要求232-234中任一项所述的方法，其中所述呈递模型表示以下两者之间的依赖性：

(a) 所述MHC等位基因中的特定一者和在肽序列的特定位置处的特定氨基酸的一对的存在；和

(b) 由所述对的所述MHC等位基因中的所述特定一者将在所述特定位置处包含所述特定氨基酸的所述肽序列呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性。

236. 如权利要求232-235中任一项所述的方法，其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位被呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性增加的表位。

237. 如权利要求232-236中任一项所述的方法，其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激肿瘤特异性或感染性疾病生物体特异性免疫应答的可能性增加的表位。

238. 如权利要求232-237中任一项所述的方法，其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够由专职抗原呈递细胞 (APC) 呈递给初始T细胞的可能性增加的表位，任选地其中所述APC是树突状细胞 (DC)。

239. 如权利要求232-238中任一项所述的方法，其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位经由中心或外周耐受性受抑制的可能性降低的表位。

240. 如权利要求232-239中任一项所述的方法，其中选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激对正常组织的自身免疫应答的可能性降低的表位。

241. 如权利要求232-240中任一项所述的方法,其中外显子组或转录组核苷酸测序数据是通过肿瘤细胞或组织、受感染细胞或感染性疾病生物体进行测序而获得的。

242. 如权利要求241所述的方法,其中所述测序是下一代测序(NGS)或任何大规模并行测序方法。

243. 如权利要求143-242中任一项所述的方法,其中用于递送所述自扩增表达系统的所述组合物作为初免疫苗施用。

244. 如权利要求143-243中任一项所述的方法,其中所述方法还包括施用第二组合物,任选地其中所述第二组合物是疫苗组合物。

245. 如权利要求244所述的方法,其中所述第二组合物是在用于递送所述自扩增表达系统的所述组合物之前施用的。

246. 如权利要求244所述的方法,其中所述第二组合物是在施用用于递送所述自扩增表达系统的所述组合物之后施用的。

247. 如权利要求244-246中任一项所述的方法,其中所述第二组合物与用于递送所述自扩增表达系统的所述组合物相同。

248. 如权利要求244-246中任一项所述的方法,其中所述第二组合物与用于递送所述自扩增表达系统的所述组合物不同。

249. 如权利要求248所述的方法,其中所述第二组合物包含所述自扩增表达系统的所述盒,任选地其中所述第二组合物包含编码所述自扩增表达系统的所述盒的黑猩猩腺病毒载体。

250. 如权利要求244-249中任一项所述的方法,其中施用两种或更多种第二组合物,任选地其中用于递送所述自扩增表达系统的所述组合物作为初免疫苗施用。

251. 如权利要求143-250中任一项所述的方法,其中用于递送所述自扩增表达系统的所述组合物是肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用的。

252. 如权利要求143-251中任一项所述的方法,其中所述方法还包括施用免疫调节剂,任选地其中所述免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段、抗4-1BB抗体或其抗原结合片段、抗OX-40抗体或其抗原结合片段、或细胞因子,任选地其中所述细胞因子是IL-2、IL-7、IL-12、IL-15或IL-21或其变体中的至少一者。

253. 如权利要求143-252中任一项所述的方法,其中所述方法还包括施用佐剂。

加帽化合物、组合物及其使用方法

[0001] 相关申请的交叉引用

[0002] 本申请要求于2020年4月21日提交的美国临时申请第63/013,456号和2020年5月5日提交的美国临时申请第63/020,473号的权益,所述临时申请各自出于所有目的特此通过引用整体并入。

[0003] 序列表

[0004] 本申请含有序列表,所述序列表已以ASCII格式以电子方式提交并特此通过引用整体并入。所述ASCII副本创建于2021年4月21日,命名为GS0_088W0_sequencelisting.txt并且大小为422,240字节。

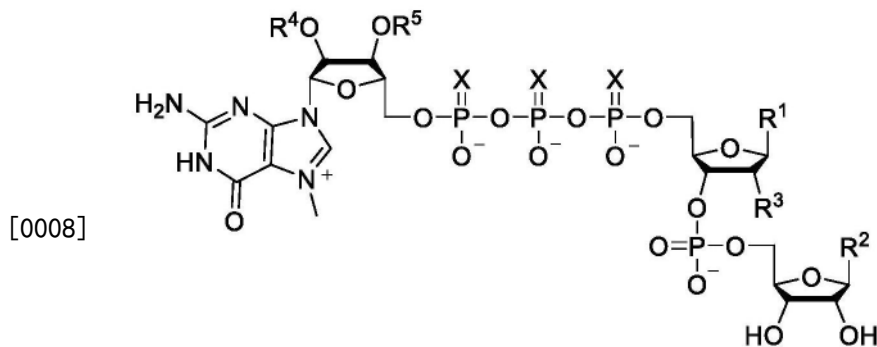
背景技术

[0005] 编码用于治疗应用的生理上重要的蛋白质的信使RNA (mRNA),与基于DNA的质粒和病毒载体相比,在递送遗传物质方面显示出显著优势。存在于每个活性mRNA分子中的若干结构元件用于有效地翻译所编码的蛋白质。这些元件之一是mRNA的5'-端的帽结构,它存在于所有真核生物(和一些病毒)中。天然存在的帽结构包含在鸟嘌呤碱基的N7位处甲基化的核糖-鸟苷残基。该7-甲基鸟苷(^{7m}G)经由5'-至5'-三磷酸根链在mRNA分子的5'端连接。5'端上^{7m}Gppp片段的存在对于mRNA成熟至关重要,它保护mRNA免受外切核酸酶的降解,促进mRNA从细胞核到细胞质的转运,并在翻译起始复合物的组装中起关键作用。

[0006] 工业中需要允许大规模合成mRNA的组合物和方法,它们(a)比常规方法省力,(b)在转录期间消除或减少双向起始,(c)导致更高的mRNA产量,(d)与当前方法相比成本降低,(e)减少了具有不同5'-序列的异质产物的产生,并且(f)不需要额外的酶促反应以将帽1和帽2结构掺入到合成的mRNA中。还需要合成含有经修饰和/或非天然核苷的各种mRNA,其在分子的5'端处或附近携带特定修饰和/或亲和标签如荧光染料、放射性同位素、质量标签和/或分子结合对的一个配偶体如生物素。

发明内容

[0007] 本公开尤其包括一种式(I)化合物:

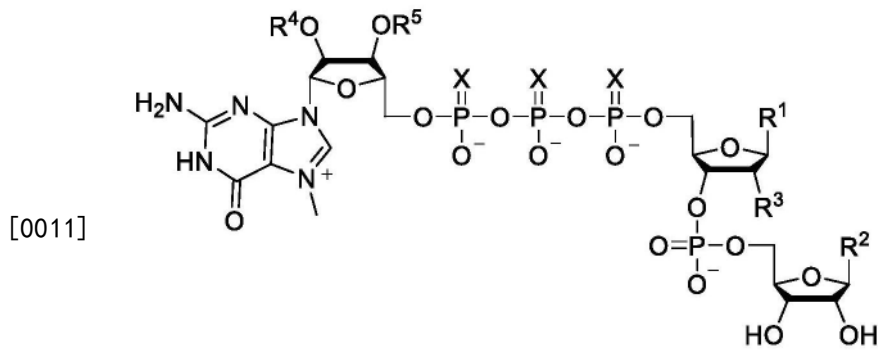


(I),

[0009] 或其药学上可接受的盐。此外,本公开尤其包括药物组合物、使用方法和制备式

(I) 化合物的方法。

[0010] 本文提供了一种式 (I) 化合物



(I),

[0012] 或其药学上可接受的盐,

[0013] 其中

[0014] R^1 是核苷;

[0015] R^2 是核苷;

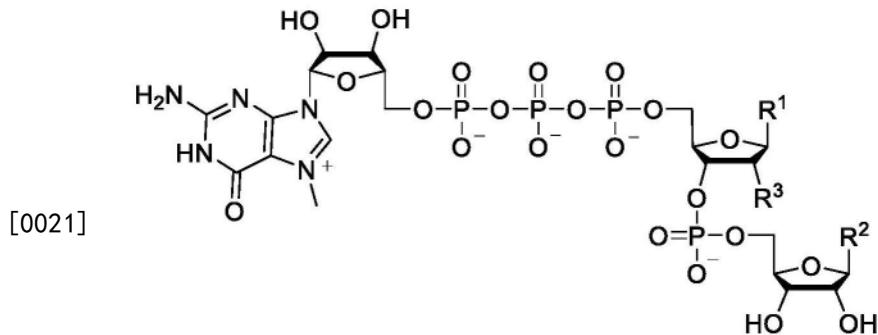
[0016] R^3 是卤素、任选地经取代的 C_1-C_3 烷基或经取代的 C_1-C_3 烷氧基;

[0017] R^4 是氢或任选地经取代的 C_1-C_3 脂族;

[0018] R^5 是氢或任选地经取代的 C_1-C_3 脂族; 并且

[0019] 每个 X 独立地是 O 或 S, 并且

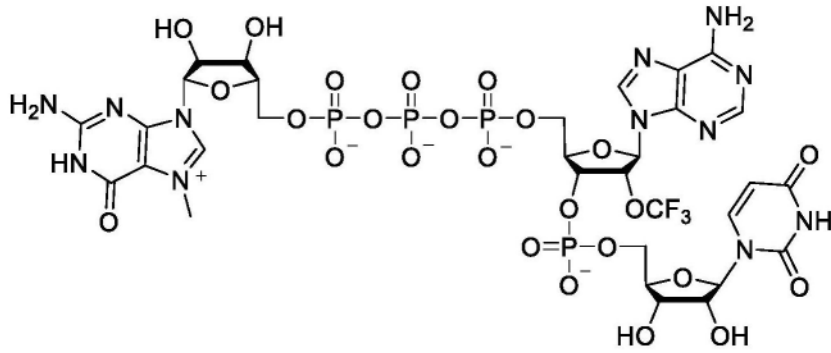
[0020] 任选地, 其中所述化合物具有式 (I-1):



(I-1),

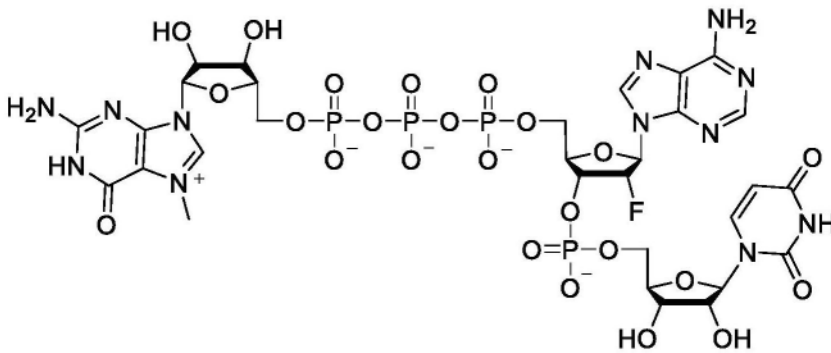
[0022] 或其药学上可接受的盐。

[0023] 在一些方面, R^1 是腺嘌呤。在一些方面, R^1 是 N6-甲基化腺嘌呤。在一些方面, R^2 是尿嘧啶。在一些方面, R^3 选自由氟、 $-CF_3$ 、 $-OCF_3$ 和 $-OCH_2CH_2OCH_3$ 组成的组。在一些方面, 所述化合物选自由以下组成的组:

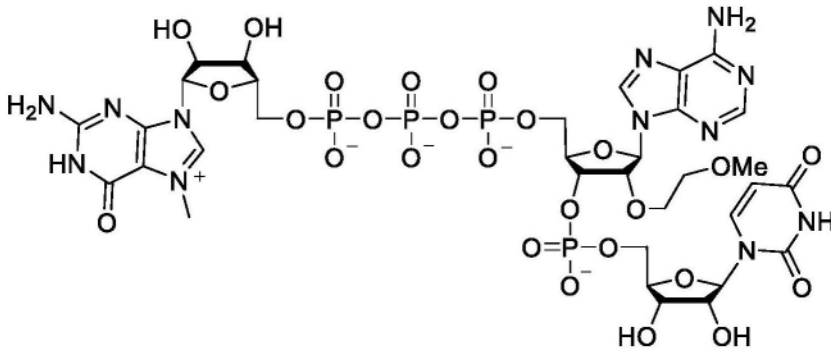


1、

[0024]



2、



3、

[0025] 及其药学上可接受的盐。

[0026] 本文还提供了一种刺激免疫应答的方法，任选地其中所述免疫应答治疗癌症，所述方法包括向有需要的患者施用RNA寡核苷酸，其中所述RNA寡核苷酸包含本文所述化合物中的任一者。在一些方面，所述癌症选自由以下组成的组：肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性髓细胞白血病、成人急性淋巴母细胞白血病、慢性髓细胞白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。在一些方面，所述癌症是实体瘤。在一些方面，所述癌症选自由以下组成的组：MSS-CRC、NSCLC和PDA。在一些方面，所述癌症选自由以下组成的组：微卫星稳定结直肠癌 (MSS-CRC)、非小细胞肺癌 (NSCLC)、胰腺导管腺癌 (PDA)

和胃食管腺癌(GEA)。

[0027] 本文还提供了一种免疫或治疗感染的方法,所述方法包括向有需要的患者施用RNA寡核苷酸,其中所述RNA寡核苷酸包含本文所述化合物中的任一者。在一些方面,所述感染是真菌感染。在一些方面,所述感染是病毒感染。在一些方面,所述病毒感染是HIV感染。

[0028] 本文还提供了一种包含起始加帽寡核苷酸引物和DNA模板的复合物,其中所述起始加帽寡核苷酸引物包含本文所述化合物中的任一者,其中所述DNA模板包含包括转录起始位点的启动子区,所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸;并且其中所述起始加帽寡核苷酸引物至少在核苷酸位置+1和+2处与所述DNA模板杂交。

[0029] 本文还提供了一种自扩增表达系统,

[0030] 其中所述自扩增表达系统包含自扩增骨架,其中所述自扩增骨架包含自我复制的RNA病毒的一种或多种多核苷酸序列;并且

[0031] 其中所述自扩增表达系统包含核酸序列,其中每个元件从5'至3'由下式所述连接:

[0032] $m^7G\text{-ppp-N}_1\text{-N}_2\text{-N}_v$, 其中

[0033] m^7G 是7-甲基鸟苷酸(m^7G)帽,

[0034] ppp是三磷酸根桥,

[0035] N_1 是所述自扩增骨架的第一个核苷酸,对应于所述自我复制的RNA病毒的第一个内源5'核苷酸,

[0036] N_2 是所述自扩增骨架的第二个核苷酸,对应于所述自我复制的RNA病毒的第二个内源5'核苷酸,并且

[0037] N_v 包含(1)所述自扩增骨架的一种或多种额外核酸序列,和(2)包含至少一种用于递送的外源核酸序列的盒,任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列,任选地其中所述多肽编码核酸序列是抗原编码核酸序列,并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中。

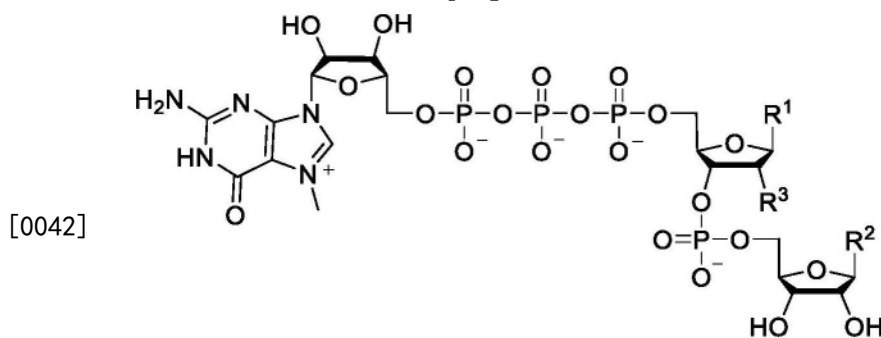
[0038] 在一些方面,用于递送所述自扩增表达系统的所述组合物包含:(A)所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA(SAM)载体,其中所述一种或多种SAM载体包含:(a)所述自扩增骨架,其中所述自扩增骨架包含:(i)至少一种启动子核苷酸序列,(ii)至少一种聚腺苷酸化(聚(A))序列,和(b)所述盒,任选地其中所述盒包含以下中的一者或多者:(i)至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:a.表位编码核酸序列,任选地包含:(1)使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2)编码感染性疾病生物体肽的核酸序列,所述肽选自自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,b.任选的5'接头序列,和c.任选的3'接头序列;(ii)可操作地连接到所述至少一种抗原编码核酸序列的第二启动子核苷酸序列;或(iii)任选地,至少一种第二聚(A)序列,其中所述第二聚(A)序列是所述自我复制的RNA病毒的原生聚(A)序列或外源聚(A)序列;和(B)任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

[0039] 在一些方面,用于递送所述自扩增表达系统的所述组合物包含:(A)所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA(SAM)载体,其中所述一种或

多种SAM载体包含：(a) 所述自扩增骨架，其中所述自扩增骨架包含以SEQ ID NO:6所示的核酸序列，其中所述自扩增骨架序列包含亚基因组启动子核苷酸序列和聚(A)序列，其中所述亚基因组启动子序列对于所述自我复制的RNA病毒是内源的，其中所述聚(A)序列对于所述自我复制的RNA病毒骨架是内源的；和(b) 整合在所述亚基因组启动子核苷酸序列与所述聚(A)序列之间的盒，其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列，并且任选地其中所述盒包含至少一种抗原编码核酸序列，所述抗原编码核酸序列包含：a. 表位编码核酸序列，任选地包含：(1) 使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变，或(2) 编码感染性疾病生物体肽的核酸序列，所述肽选自由以下组成的组：病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽，b. 任选的5'接头序列，和c. 任选的3'接头序列；和(B) 任选地，脂质纳米粒子(LNP)，其中所述LNP包封所述自扩增表达系统。

[0040] 在一些方面， N_1 是修饰的核苷酸，任选地其中所述修饰的核苷酸包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖。在一些方面， N_1 是修饰的腺苷。在一些方面， N_1 是2'-OH-甲基化N6-甲基腺苷。在一些方面， N_2 是修饰的核苷酸，任选地其中所述修饰的核苷酸包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖。在一些方面， N_1 和 N_2 是修饰的核苷酸，任选地其中所述修饰的核苷酸各自独立地包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖。在一些方面， N_1 是腺苷或修饰的腺苷，任选地其中所述修饰的腺苷包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖。在一些方面， N_2 是尿苷或修饰的尿苷，任选地其中所述修饰的尿苷包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖。在一些方面， N_1 是修饰的腺苷，任选地其中所述修饰的腺苷包含选自以下组成的组的修饰：修饰的糖、修饰的核苷、核苷类似物或其组合，任选地其中所述修饰的糖是修饰的核糖，并且 N_2 是尿苷。

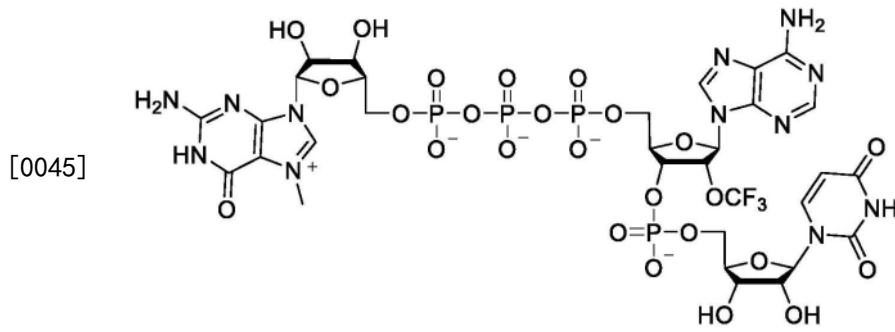
[0041] 在一些方面， $m^7G\text{-ppp-}N_1\text{-}N_2$ 由式(I-1)表示：



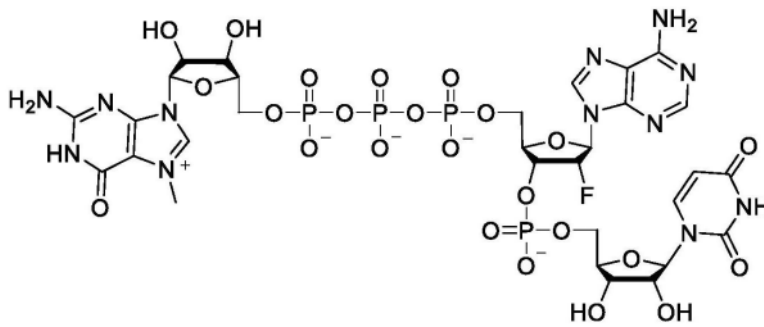
(I-1),

[0043] 或其药学上可接受的盐，其中 R^1 是核苷，任选地其中 R^1 是腺嘌呤，任选地其中 R^1 是N6-甲基化腺嘌呤； R^2 是核苷，任选地其中 R^2 是尿嘧啶；并且 R^3 是卤素、任选地经取代的 $C_1\text{-}C_3$ 烷基或经取代的 $C_1\text{-}C_3$ 烷氧基。在一些方面， R^3 选自由氟、 $-\text{CF}_3$ 、 $-\text{OCF}_3$ 和 $-\text{OCH}_2\text{CH}_2\text{OCH}_3$ 组成的组。

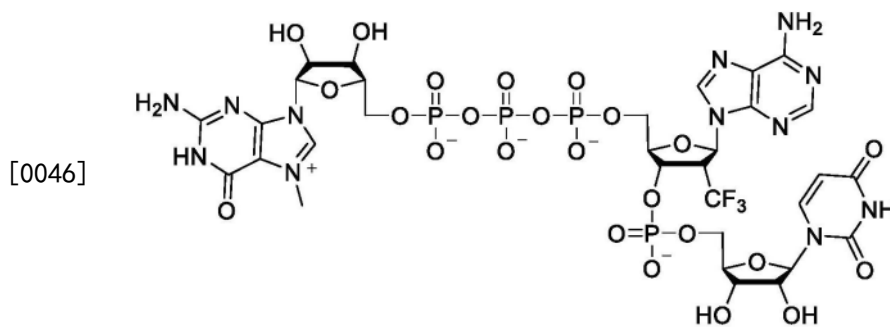
[0044] 在一些方面, $m^7\text{G}-\text{ppp}-\text{N}_1-\text{N}_2$ 由选自以下组成的组的式表示:



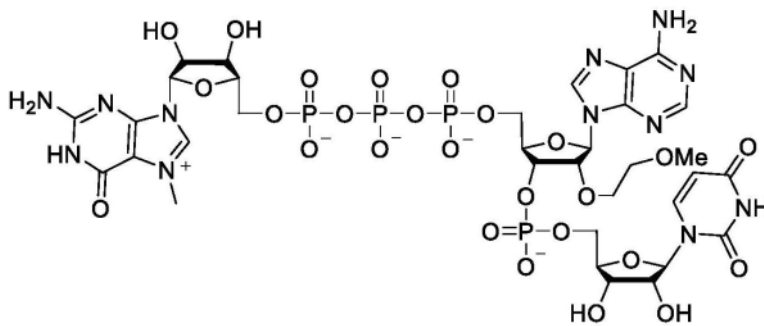
1、



2、



3、



4、

[0047] 及其药学上可接受的盐。

[0048] 在一些方面,所述自扩增表达系统是通过体外转录产生的。在一些方面,所述体外转录过程包括使用包含本文所述 $m^7G\text{-ppp-N}_1\text{-N}_2$ 中的任一者的起始加帽寡核苷酸。

[0049] 本文还提供了一种包含起始加帽寡核苷酸引物和DNA模板的复合物,其中所述起始加帽寡核苷酸引物包含本文所述的具有式 $m^7G\text{-ppp-N}_1\text{-N}_2$ 的任何化合物,其中所述DNA模板从5'至3'包含:(A)包含转录起始位点的RNA转录启动子区,所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸,和(B)可操作地连接到所述RNA转录启动子区的包含本文所述的具有式 $N_1\text{-N}_2\text{-N}_v$ 的任何序列的序列。

[0050] 在一些方面,所述RNA转录启动子区包含T7启动子序列,任选地其中所述T7启动子序列是核苷酸序列TAATACGACTCACTATA或TAATACGACTCACTATT;SP6启动子序列,任选地其中所述SP6启动子序列是核苷酸序列ATTTAGGTGACACTATA;或K11 RNAP启动子序列,任选地其中所述K11 RNAP启动子序列是核苷酸序列AATTAGGGCACACTATA。在一些方面,所述DNA模板包含以SEQ ID NO:57所示的序列,并且其中所述盒被插入在如SEQ ID NO:6的序列中所示的位置7544以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

[0051] 在一些方面,用于递送所述自扩增表达系统的所述组合物中的所述盒的每个元件的有序序列描述于下式,从5'至3'包含:

[0052] $P_a - (L5_b - N_c - L3_d)_X - (G5_e - U_f)_Y - G3_g$

[0053] 其中P包含第二启动子核苷酸序列,其中a=0或1,N包含表位编码核酸序列之一,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,其中c=1,L5包含5'接头序列,其中b=0或1,L3包含3'接头序列,其中d=0或1,G5包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中e=0或1,G3包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中g=0或1,U包含至少一种MHC II类表位编码核酸序列之一,其中f=1,X=1至400,其中对于每个X,相应的 N_c 是MHC I类表位编码核酸序列,并且Y=0、1或2,其中对于每个Y,相应的 U_f 是MHC II类表位编码核酸序列。

[0054] 在一些方面,对于每个X,相应的 N_c 是不同的MHC I类表位编码核酸序列。在一些方面,对于每个Y,相应的 U_f 是不同的MHC II类表位编码核酸序列。在一些方面,a=0,b=1,d=1,e=1,g=1,h=1,X=10,Y=2,所述至少一种启动子核苷酸序列是由所述自扩增骨架提供的单个亚基因组启动子核苷酸序列,所述至少一种聚腺苷酸化聚(A)序列是由所述自扩增骨架提供的至少80个连续A核苷酸的聚(A)序列,所述盒整合在所述亚基因组启动子核苷酸序列和所述聚(A)序列之间,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列和所述聚(A)序列,每个N编码长度为7-15个氨基酸的MHC I类表位,L5是编码所述MHC I表位的原生N末端氨基酸序列的原生5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,L3是编码所述MHC I表位的原生C末端氨基酸序列的原生3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一者,所述自扩增骨架是以SEQ ID NO:6所示的序列,并且所述MHC I类表位编码核酸序列中的每一者编码长度为13至25个氨基酸的多肽。

[0055] 在一些方面,用于递送的至少一种外源核酸序列包含所述多肽编码核酸序列。在一些方面,所述多肽编码核酸序列编码抗原编码核酸序列。在一些方面,所述抗原编码核酸序列包含MHC I类表位、MHC II类表位、能够刺激B细胞应答的表位或其组合。在一些方面,

所述抗原编码核酸序列包含编码全长蛋白质、蛋白质亚基、蛋白质结构域或其组合的序列。在一些方面,所述多肽编码核酸序列编码全长蛋白质或其功能部分。在一些方面,所述全长蛋白质或其功能部分选自由以下组成的组:抗体、细胞因子、嵌合抗原受体(CAR)、T细胞受体和基因组编辑系统核酸酶。

[0056] 在一些方面,用于递送的至少一种外源核酸序列包含至少一种包含非编码核酸序列的核酸序列。在一些方面,所述非编码核酸序列是RNA干扰(RNAi)多核苷酸或基因组编辑系统多核苷酸。

[0057] 在一些方面,所述LNP包含选自由以下组成的组的脂质:可电离氨基脂质、阳离子脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。在一些方面,所述LNP包含可电离氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。在一些方面,所述可电离氨基脂质包含MC3样(二亚油醇甲基-4-二甲基氨基丁酸酯)分子。在一些方面,所述LNP包封的表达系统具有约100nm的直径。在一些方面,所述LNP包封的表达系统具有60-140nm之间的直径。

[0058] 在一些方面,用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用。在一些方面,用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)施用。

[0059] 在一些方面,所述盒整合在所述至少一种启动子核苷酸序列和所述至少一种聚(A)序列之间。在一些方面,所述至少一种启动子核苷酸序列可操作地连接到所述盒。

[0060] 在一些方面,所述一种或多种SAM载体包含一种或多种正链RNA载体。在一些方面,所述一种或多种SAM载体包含一种或多种负链RNA载体。在一些方面,所述一种或多种负链RNA载体包含麻疹病毒或弹状病毒的至少一种多核苷酸序列。

[0061] 在一些方面,所述一种或多种SAM载体在哺乳动物细胞内自扩增。在一些方面,所述自我复制的RNA病毒选自由以下组成的组:甲病毒;黄病毒、麻疹病毒和弹状病毒。

[0062] 在一些方面,所述自扩增骨架包含甲病毒的至少一种多核苷酸序列,任选地其中所述甲病毒选自由以下组成的组:奥拉病毒(Aura virus)、摩根堡病毒(Fort Morgan virus)、委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus)、罗斯河病毒(Ross River virus)、塞姆利基森林病毒(Semiliki Forest virus)、辛德毕斯病毒(Sindbis virus)和马雅罗病毒(Mayaro virus)。在一些方面,所述自扩增骨架包含委内瑞拉马脑炎病毒的至少一种核苷酸序列。在一些方面,所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因。在一些方面,所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列和聚(A)序列。在一些方面,所述用于非结构蛋白介导的扩增的序列选自由以下组成的组:甲病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、甲病毒3'UTR或其组合。在一些方面,所述自扩增骨架不编码结构性病毒粒子蛋白衣壳E2和E1,任选地其中E1是全长E1,或不编码结构性病毒粒子蛋白衣壳E3、E2、6K。在一些方面,所述盒被插入以代替所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆

利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的多核苷酸序列内的结构性病毒粒子蛋白。在一些方面,所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列。在一些方面,所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列,其还包含碱基对7544和11175之间的缺失。在一些方面,所述自扩增骨架包含以SEQ ID NO:6或SEQ ID NO:7所示的序列。在一些方面,所述盒被插入在位置7544以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。在一些方面,所述盒的插入提供了包含nsP1-4基因和至少一种核酸序列的多顺反子RNA的转录,其中所述nsP1-4基因和所述至少一种核酸序列在单独的开放阅读框中。

[0063] 在一些方面,所述至少一种启动子核苷酸序列是由所述自我复制的RNA病毒编码的原生(也称为“内源”)启动子核苷酸序列,任选地其中所述原生启动子核苷酸序列是亚基因组启动子核苷酸序列。在一些方面,所述至少一种启动子核苷酸序列是外源RNA启动子。在一些方面,所述第二启动子核苷酸序列是亚基因组启动子核苷酸序列。在一些方面,所述第二启动子核苷酸序列包含多个亚基因组启动子核苷酸序列,其中每个亚基因组启动子核苷酸序列提供所述单独的开放阅读框中的一者或多者的转录。

[0064] 在一些方面,所述一种或多种SAM载体的大小各自为至少300nt。在一些方面,所述一种或多种SAM载体的大小各自为至少1kb。在一些方面,所述一种或多种SAM载体的大小各自为2kb。在一些方面,所述SAM载体的大小各自小于5kb。

[0065] 在一些方面,所述至少一种抗原编码核酸序列包含两种或更多种抗原编码核酸序列。在一些方面,每种抗原编码核酸序列彼此直接连接。

[0066] 在一些方面,每种抗原编码核酸序列用编码接头的核酸序列连接到不同的抗原编码核酸序列。在一些方面,所述接头连接两个MHC I类表位编码核酸序列或一个MHC I类表位编码核酸序列至一个MHC II类表位编码核酸序列。在一些方面,所述接头选自自由以下组成的组:(1)连续甘氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(2)连续丙氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)由哺乳动物蛋白酶体有效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;和(6)侧接源自同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或2-20个氨基酸残基的一种或多种原生序列。在一些方面,所述接头连接两个MHC II类表位编码核酸序列或一个MHC II类序列至一个MHC I类表位编码核酸序列。在一些方面,所述接头包含序列GPGPG。

[0067] 在一些方面,所述抗原编码核酸序列可操作地或直接地连接到增强所述表位编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性的单独或连续的序列。在一些方面,所述单独或连续的序列包含以下中的至少一者:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,所述泛素序列在位置76处含有Gly至Ala取代)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

[0068] 在一些方面,所述至少一种抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、

11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少2-400种抗原编码核酸序列,并且其中所述抗原编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。在一些方面,每种抗原编码核酸序列独立地包含至少2-10、2、3、4、5、6、7、8、9或10种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少2-400种表位编码核酸序列,并且其中所述表位编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。

[0069] 在一些方面,所述MHC I类表位中的至少两者由MHC I类呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上。

[0070] 在一些方面,所述表位编码核酸序列包含至少一种MHC I类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为8至35个氨基酸,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸的多肽序列。

[0071] 在一些方面,存在所述至少一种MHC II类表位编码核酸序列。在一些方面,所述至少一种MHC II类表位编码核酸序列存在并包含至少一种MHC II类表位编码核酸序列,所述MHC II类表位编码核酸序列包含使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变。

[0072] 在一些方面,所述表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。在一些方面,所述表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在至少一种MHC II类表位编码核酸序列,并且其中所述至少一种MHC II类表位编码核酸序列包含至少一种通用MHC II类表位编码核酸序列,任选地其中所述至少一种通用序列包含破伤风类毒素和PADRE中的至少一者。

[0073] 在一些方面,所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是诱导型的。在一些方面,所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是非诱导型的。在一些方面,所述至少一种聚(A)序列包含所述自我复制病毒原生的聚(A)序列。在一些方面,所述至少一种聚(A)序列包含所述自我复制病毒外源的聚(A)序列。在一些方面,所述至少一种聚(A)序列可操作地连接到所述至少一种核酸序列中的至少一者。在一些方面,所述至少一种聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少110个或至少120个连续A核苷酸。在一些方面,所述至少一种聚(A)序列是至少80个连续A核苷酸。

[0074] 在一些方面,所述表位编码核酸序列包含MHC I类表位编码核酸序列,并且其中所述MHC I类表位编码核酸序列是通过执行以下步骤来选择:(a)从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;(b)将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及(c)基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生所述MHC I类表位编码核酸序列。

[0075] 在一些方面,所述MHC I类表位编码核酸序列中的每一者是通过执行以下步骤来选择:(a)从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;(b)将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及(c)基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生至少20种MHC I类表位编码核酸序列。在一些方面,所述经选择表位集合的数目为2-20。

[0076] 在一些方面,所述呈递模型表示以下两者之间的依赖性:(a)所述MHC等位基因中的特定一者和在肽序列的特定位置处的特定氨基酸的一对的存在;和(b)由所述对的所述MHC等位基因中的所述特定一者将在所述特定位置处包含所述特定氨基酸的所述肽序列呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位被呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性增加的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激肿瘤特异性或感染性疾病生物体特异性免疫应答的可能性增加的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够由专职抗原呈递细胞(APC)呈递给初始T细胞的可能性增加的表位,任选地其中所述APC是树突状细胞(DC)。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位经由中心或外周耐受性受抑制的可能性降低的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激对正常组织的自身免疫应答的可能性降低的表位。在一些方面,外显子组或转录组核苷酸测序数据是通过对肿瘤细胞或组织、受感染细胞或感染性疾病生物体进行测序而获得的。在一些方面,所述测序是下一代测序(NGS)或任何大规模并行测序方法。

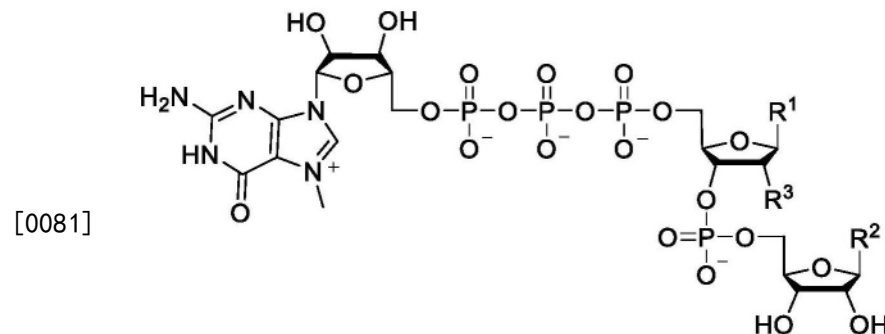
[0077] 本文还提供了一种产生自扩增表达系统的方法,其中所述方法包括以下步骤:a)提供DNA模板,其中每个元件从5'至3'由下式所述连接: $P-N_1-N_2-N_V$,其中,P包含含有转录起始位点的RNA转录启动子区,所述转录起始位点具有在核苷酸位置+1处的第一个核苷酸和在核苷酸位置+2处的第二个核苷酸, N_1 是自扩增骨架的第一个核苷酸,对应于自我复制的RNA病毒的第一个内源5'核苷酸, N_2 是所述自扩增骨架的第二个核苷酸,对应于所述自我复

制的RNA病毒的第二个内源5'核苷酸,并且 N_V 包含(1)所述自扩增骨架的一种或多种额外核酸序列,和(2)包含至少一种用于递送的外源核酸序列的盒,任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列,任选地其中所述多肽编码核酸序列是抗原编码核酸序列,并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中;
b) 提供起始加帽寡核苷酸引物,其中所述起始加帽寡核苷酸引物包含核酸序列,其中每个元件从5'至3'由下式所述连接: $m^7G\text{-ppp-}N_1\text{-}N_2$,其中 m^7G 是7-甲基鸟苷酸(m^7G)帽,ppp是三磷酸根桥, N_1 是对应于所述DNA模板的 N_1 的核苷酸,并且 N_2 是对应于所述DNA模板的 N_2 的核苷酸,以及c) 提供能够从所述RNA转录启动子区起始转录的RNA聚合酶,d) 在足以产生包含核酸序列的所述自扩增表达系统的条件下使所述DNA模板、所述起始加帽寡核苷酸引物和所述RNA聚合酶接触,其中每个元件从5'至3'由式 $m^7G\text{-ppp-}N_1\text{-}N_2\text{-}N_V$ 所述连接。

[0078] 在一些方面,所述RNA转录启动子区包含T7启动子序列,任选地其中所述T7启动子序列是核苷酸序列TAATACGACTCACTATA或TAATACGACTCACTATT;SP6启动子序列,任选地其中所述SP6启动子序列是核苷酸序列ATTTAGGTGACACTATA;或K11 RNAP启动子序列,任选地其中所述K11 RNAP启动子序列是核苷酸序列AATTAGGGCACACTATA。在一些方面,所述DNA模板包含以SEQ ID NO:57所示的序列,并且其中所述盒被插入在如SEQ ID NO:6的序列中所示的位置7544以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。

[0079] 在一些方面, N_1 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_1 是腺苷或修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_2 是尿苷或修饰的尿苷,任选地其中所述修饰的尿苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_1 是修饰的腺苷,任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖,并且 N_2 是尿苷。

[0080] 在一些方面,所述起始加帽寡核苷酸引物由式(I-1)表示:

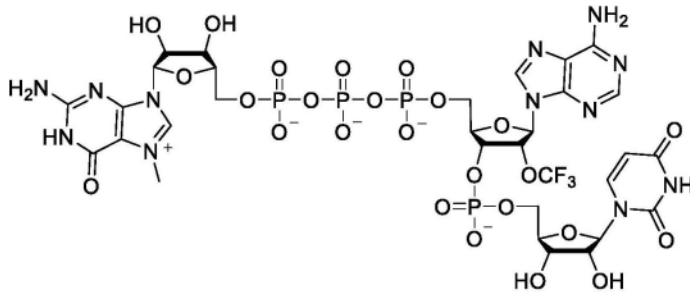


(I-1),

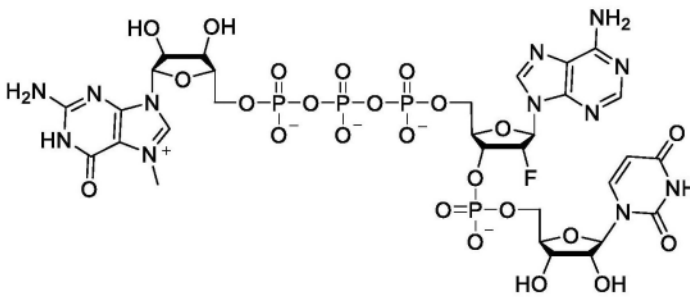
[0082] 或其药学上可接受的盐,其中 R^1 是核苷,任选地其中 R^1 是腺嘌呤,任选地其中 R^1 是

N6-甲基化腺嘌呤; R^2 是核苷, 任选地其中 R^2 是尿嘧啶; 并且 R^3 是卤素、任选地经取代的 C_1-C_3 烷基或经取代的 C_1-C_3 烷氧基。

[0083] 在一些方面, R^3 选自由氟、 $-CF_3$ 、 $-OCF_3$ 和 $-OCH_2CH_2OCH_3$ 组成的组。在一些方面, 所述起始加帽寡核苷酸引物由选自由以下组成的组的式表示:

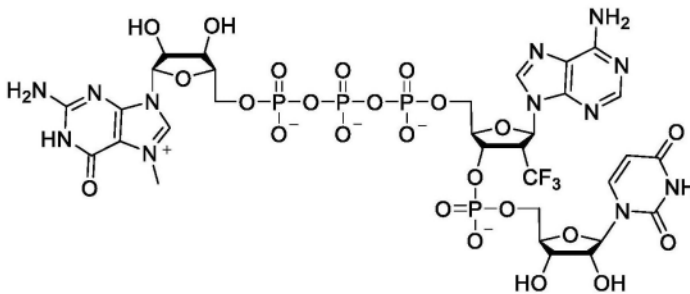


1、

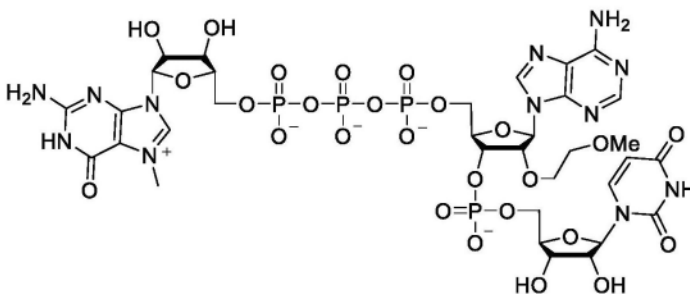


2、

[0084]



3、



4、

[0085]

[0086] 及其药学上可接受的盐。

[0087] 本文还提供了一种刺激受试者中的免疫应答的方法, 所述方法包括向所述受试者

施用用于递送自扩增表达系统的组合物,其中所述自扩增表达系统包含自扩增骨架,其中所述自扩增骨架包含自我复制的RNA病毒的一种或多种多核苷酸序列;并且其中所述自扩增表达系统包含核酸序列,其中每个元件从5'至3'由下式所述连接: $m^7G\text{-ppp-N}_1\text{-N}_2\text{-N}_v$,其中 m^7G 是7-甲基鸟苷酸(m^7G)帽,ppp是三磷酸根桥, N_1 是所述自扩增骨架的第一个核苷酸,对应于所述自我复制的RNA病毒的第一个内源5'核苷酸, N_2 是所述自扩增骨架的第二个核苷酸,对应于所述自我复制的RNA病毒的第二个内源5'核苷酸,并且 N_v 包含(1)所述自扩增骨架的一种或多种额外核酸序列,和(2)包含至少一种用于递送的外源核酸序列的盒,任选地其中所述至少一种外源核酸序列包含多肽编码核酸序列,任选地其中所述多肽编码核酸序列是抗原编码核酸序列,并且其中所述盒可操作地连接到所述自扩增骨架或可操作地插入所述自扩增骨架中。

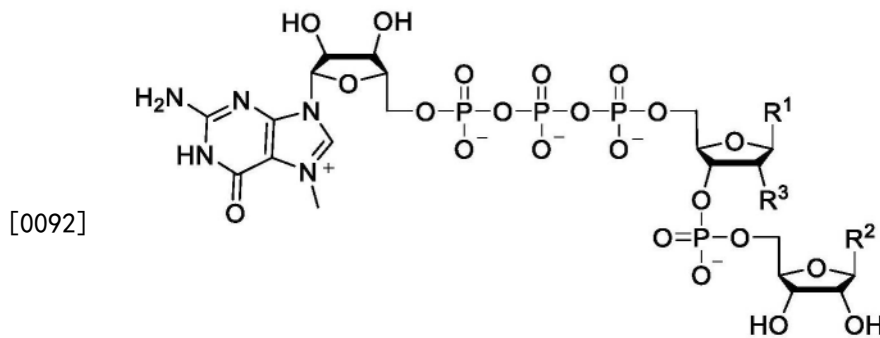
[0088] 在一些方面,用于递送所述自扩增表达系统的所述组合物包含:(A)所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA(SAM)载体,其中所述一种或多种SAM载体包含:(a)所述自扩增骨架,其中所述自扩增骨架包含:(i)至少一种启动子核苷酸序列,(ii)至少一种聚腺苷酸化(聚(A))序列,和(b)所述盒,任选地其中所述盒包含以下中的一者或多者:(i)至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:a.表位编码核酸序列,任选地包含:(1)使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2)编码感染性疾病生物体肽的核酸序列,所述肽选自自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,b.任选的5'接头序列,和c.任选的3'接头序列;(ii)可操作地连接到所述至少一种抗原编码核酸序列的第二启动子核苷酸序列;或(iii)任选地,至少一种第二聚(A)序列,其中所述第二聚(A)序列是所述自我复制的RNA病毒的原生聚(A)序列或外源聚(A)序列;和(B)任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

[0089] 在一些方面,用于递送所述自扩增表达系统的所述组合物包含:(A)所述自扩增表达系统,其中所述自扩增表达系统包含一种或多种自扩增mRNA(SAM)载体,其中所述一种或多种SAM载体包含:(a)所述自扩增骨架,其中所述自扩增骨架包含以SEQ ID NO:6所示的核酸序列,其中所述自扩增骨架序列包含亚基因组启动子核苷酸序列和聚(A)序列,其中所述亚基因组启动子序列对于所述自我复制的RNA病毒是内源的,其中所述聚(A)序列对于所述自扩增骨架是内源的;和(b)整合在所述亚基因组启动子核苷酸序列与所述聚(A)序列之间的盒,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列,并且任选地其中所述盒包含至少一种抗原编码核酸序列,所述抗原编码核酸序列包含:a.表位编码核酸序列,任选地包含:(1)使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变,或(2)编码感染性疾病生物体肽的核酸序列,所述肽选自自由以下组成的组:病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和寄生虫衍生肽,b.任选的5'接头序列,和c.任选的3'接头序列;和(B)任选地,脂质纳米粒子(LNP),其中所述LNP包封所述自扩增表达系统。

[0090] 在一些方面, N_1 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_2 是修饰的核苷酸,任选地其中所述修饰的核苷酸包含选自自由以下组成的组的修饰:修饰的糖、修饰的核苷、核苷类似物或其组合,任选地其中所述修饰

的糖是修饰的核糖。在一些方面, N_1 和 N_2 是修饰的核苷酸, 任选地其中所述修饰的核苷酸各自独立地包含选自由以下组成的组的修饰: 修饰的糖、修饰的核苷、核苷类似物或其组合, 任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_1 是腺苷或修饰的腺苷, 任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰: 修饰的糖、修饰的核苷、核苷类似物或其组合, 任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_2 是尿苷或修饰的尿苷, 任选地其中所述修饰的尿苷包含选自由以下组成的组的修饰: 修饰的糖、修饰的核苷、核苷类似物或其组合, 任选地其中所述修饰的糖是修饰的核糖。在一些方面, N_1 是修饰的腺苷, 任选地其中所述修饰的腺苷包含选自由以下组成的组的修饰: 修饰的糖、修饰的核苷、核苷类似物或其组合, 任选地其中所述修饰的糖是修饰的核糖, 并且 N_2 是尿苷。

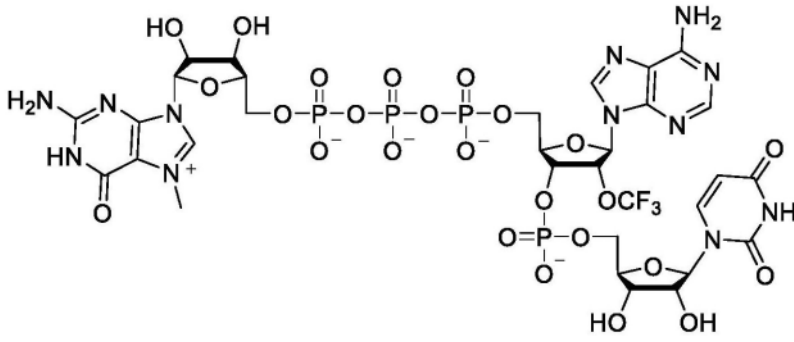
[0091] 在一些方面, $m^7G\text{-ppp-}N_1\text{-}N_2$ 由式 (I-1) 表示:



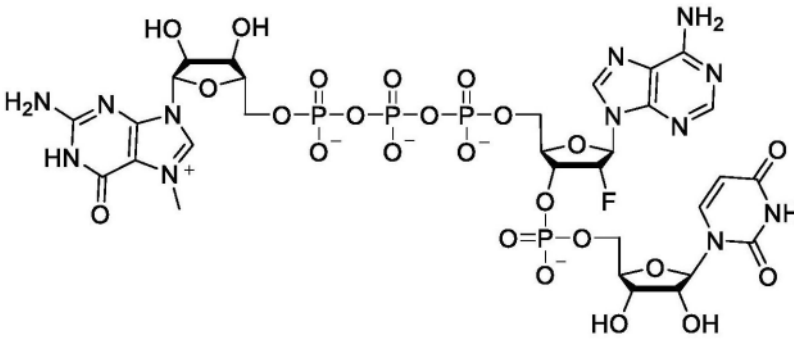
(I-1),

[0093] 或其药学上可接受的盐, 其中 R^1 是核苷, 任选地其中 R^1 是腺嘌呤, 任选地其中 R^1 是 N6-甲基化腺嘌呤; R^2 是核苷, 任选地其中 R^2 是尿嘧啶; 并且 R^3 是卤素或经取代的 $C_1\text{-}C_3$ 烷氧基。

[0094] 在一些方面, R^3 选自由氟、 $-\text{CF}_3$ 、 $-\text{OCF}_3$ 和 $-\text{OCH}_2\text{CH}_2\text{OCH}_3$ 组成的组。在一些方面, $m^7G\text{-ppp-}N_1\text{-}N_2$ 由选自由以下组成的组的式表示:

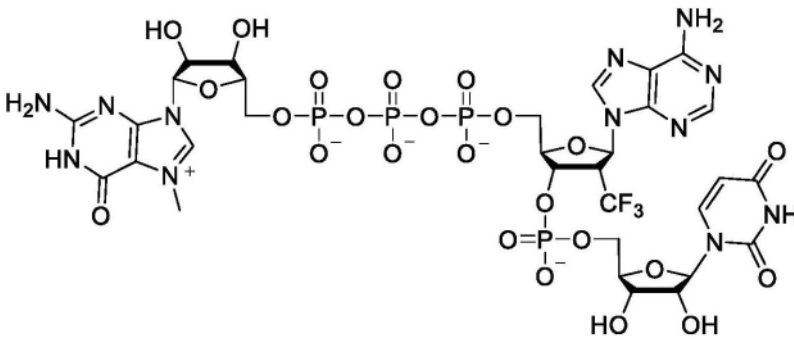


1、

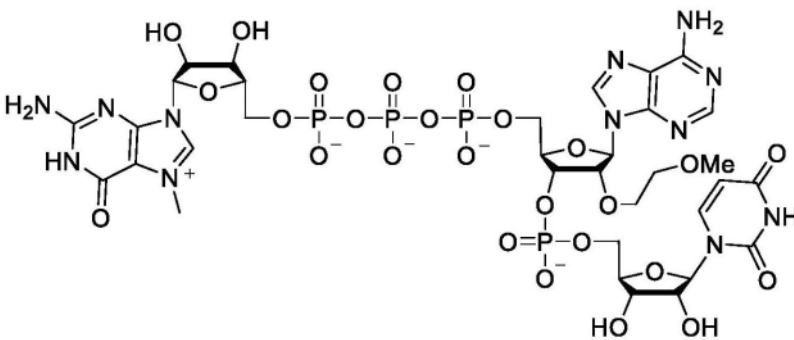


[0095]

2、



3、



[0096]

4、

[0097] 及其药学上可接受的盐。

[0098] 在一些方面,所述自扩增表达系统是通过体外转录产生的。在一些方面,所述体外转录过程包括使用包含本文所述的 $m^7G\text{-ppp-N}_1\text{-N}_2$ 组合中任一者的起始加帽寡核苷酸。在一些方面,用于递送所述自扩增表达系统的所述组合物中的所述盒的每个元件的有序序列以下式描述,从5'至3'包含 $P_a\text{-(L5}_b\text{-N}_c\text{-L3}_d)_X\text{-(G5}_e\text{-U}_f)_Y\text{-G3}_g$,其中P包含第二启动子核苷酸序列,其中 $a=0$ 或 1 ,N包含表位编码核酸序列之一,其中所述表位编码核酸序列包含MHC I类表位编码核酸序列,其中 $c=1$,L5包含5'接头序列,其中 $b=0$ 或 1 ,L3包含3'接头序列,其中 $d=0$ 或 1 ,G5包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中 $e=0$ 或 1 ,G3包含至少一种编码GPGPG氨基酸接头的核酸序列之一,其中 $g=0$ 或 1 ,U包含至少一种MHC II类表位编码核酸序列之一,其中 $f=1$, $X=1$ 至 400 ,其中对于每个X,相应的 N_c 是MHC I类表位编码核酸序列,并且 $Y=0,1$ 或 2 ,其中对于每个Y,相应的 U_f 是MHC II类表位编码核酸序列。

[0099] 在一些方面,对于每个X,相应的 N_c 是不同的MHC I类表位编码核酸序列。在一些方面,对于每个Y,相应的 U_f 是不同的MHC II类表位编码核酸序列。在一些方面, $a=0, b=1, d=1, e=1, g=1, h=1, X=10, Y=2$,所述至少一种启动子核苷酸序列是由所述自扩增骨架提供的单个亚基因组启动子核苷酸序列,所述至少一种聚腺苷酸化聚(A)序列是由所述自扩增骨架提供的至少80个连续A核苷酸的聚(A)序列,所述盒整合在所述亚基因组启动子核苷酸序列和所述聚(A)序列之间,其中所述盒可操作地连接到所述亚基因组启动子核苷酸序列和所述聚(A)序列,每个N编码长度为7-15个氨基酸的MHC I类表位,L5是编码所述MHC I表位的原生N末端氨基酸序列的原生5'接头序列,并且其中所述5'接头序列编码长度为至少3个氨基酸的肽,L3是编码所述MHC I表位的原生C末端氨基酸序列的原生3'接头序列,并且其中所述3'接头序列编码长度为至少3个氨基酸的肽,U是PADRE II类序列和破伤风类毒素MHC II类序列中的每一者,所述自扩增骨架是以SEQ ID NO:6所示的序列,并且所述MHC I类表位编码核酸序列中的每一者编码长度为13至25个氨基酸的多肽。

[0100] 在一些方面,用于递送的至少一种外源核酸序列包含所述多肽编码核酸序列。在一些方面,所述多肽编码核酸序列编码抗原编码核酸序列。在一些方面,所述抗原编码核酸序列包含MHC I类表位、MHC II类表位、能够刺激B细胞应答的表位或其组合。在一些方面,所述抗原编码核酸序列包含编码全长蛋白质、蛋白质亚基、蛋白质结构域或其组合的序列。在一些方面,所述多肽编码核酸序列编码全长蛋白质或其功能部分。在一些方面,所述全长蛋白质或其功能部分选自由以下组成的组:抗体、细胞因子、嵌合抗原受体(CAR)、T细胞受体和基因组编辑系统核酸酶。

[0101] 在一些方面,用于递送的至少一种外源核酸序列包含至少一种包含非编码核酸序列的核酸序列。在一些方面,所述非编码核酸序列是RNA干扰(RNAi)多核苷酸或基因组编辑系统多核苷酸。

[0102] 在一些方面,所述LNP包含选自由以下组成的组的脂质:可电离氨基脂质、磷脂酰胆碱、胆固醇、基于PEG的外壳脂质或其组合。在一些方面,所述LNP包含可电离氨基脂质、磷脂酰胆碱、胆固醇和基于PEG的外壳脂质。在一些方面,所述可电离氨基脂质包含MC3样(二亚油醇甲基-4-二甲氨基丁酸酯)分子。在一些方面,所述LNP包封的表达系统具有约100nm的直径。在一些方面,所述LNP包封的表达系统具有60-140nm之间的直径。

[0103] 在一些方面,用于递送所述自扩增表达系统的所述组合物被配制用于肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用。在一些方面,用于递送

所述自扩增表达系统的所述组合物被配制用于肌肉内 (IM) 施用。

[0104] 在一些方面,盒整合在所述至少一种启动子核苷酸序列和所述至少一种聚(A)序列之间。在一些方面,所述至少一种启动子核苷酸序列可操作地连接到所述盒。

[0105] 在一些方面,所述一种或多种SAM载体包含一种或多种正链RNA载体。在一些方面,所述一种或多种SAM载体包含一种或多种负链RNA载体。在一些方面,所述一种或多种负链RNA载体包含麻疹病毒或弹状病毒的至少一种多核苷酸序列。

[0106] 在一些方面,所述一种或多种SAM载体在哺乳动物细胞内自扩增。在一些方面,所述自扩增骨架包含自我复制的RNA病毒的至少一种多核苷酸序列,所述自我复制的RNA病毒选自自由以下组成的组:甲病毒;黄病毒、麻疹病毒和弹状病毒。

[0107] 在一些方面,所述自扩增骨架包含甲病毒的至少一种多核苷酸序列,任选地其中所述甲病毒选自自由以下组成的组:奥拉病毒、摩根堡病毒、委内瑞拉马脑炎病毒、罗斯河病毒、塞姆利基森林病毒、辛德毕斯病毒和马雅罗病毒。在一些方面,所述自扩增骨架包含委内瑞拉马脑炎病毒的至少一种核苷酸序列。在一些方面,所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列、聚(A)序列、非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因和nsP4基因。在一些方面,所述自扩增骨架至少包含由所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的核苷酸序列编码的用于非结构蛋白介导的扩增的序列、亚基因组启动子序列和聚(A)序列。在一些方面,所述用于非结构蛋白介导的扩增的序列选自自由以下组成的组:甲病毒5'UTR、51-nt CSE、24-nt CSE、26S亚基因组启动子序列、19-nt CSE、甲病毒3'UTR或其组合。在一些方面,所述自扩增骨架不编码结构性病毒粒子蛋白衣壳E2和E1,任选地其中E1是全长E1,或不编码结构性病毒粒子蛋白衣壳E3、E2、6K。在一些方面,所述盒被插入以代替所述奥拉病毒、所述摩根堡病毒、所述委内瑞拉马脑炎病毒、所述罗斯河病毒、所述塞姆利基森林病毒、所述辛德毕斯病毒或所述马雅罗病毒的多核苷酸序列内的结构性病毒粒子蛋白。在一些方面,所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列。在一些方面,所述委内瑞拉马脑炎病毒包含SEQ ID NO:3或SEQ ID NO:5的序列,其还包含碱基对7544和11175之间的缺失。在一些方面,所述自扩增骨架包含以SEQ ID NO:6或SEQ ID NO:7所示的序列。在一些方面,所述盒被插入在位置7544以代替如SEQ ID NO:3或SEQ ID NO:5的序列中所示的碱基对7544和11175之间的缺失。在一些方面,所述盒的插入提供了包含nsP1-4基因和至少一种核酸序列的多顺反子RNA的转录,其中所述nsP1-4基因和所述至少一种核酸序列在单独的开放阅读框中。

[0108] 在一些方面,所述至少一种启动子核苷酸序列是由所述自扩增骨架编码的原生启动子核苷酸序列,任选地其中所述原生启动子核苷酸序列是亚基因组启动子核苷酸序列。在一些方面,所述至少一种启动子核苷酸序列是外源RNA启动子。在一些方面,所述第二启动子核苷酸序列是亚基因组启动子核苷酸序列。在一些方面,所述第二启动子核苷酸序列包含多个亚基因组启动子核苷酸序列,其中每个亚基因组启动子核苷酸序列提供所述单独的开放阅读框中的一者或多者的转录。

[0109] 在一些方面,所述一种或多种SAM载体大小各自为至少300nt。在一些方面,所述

一种或多种SAM载体的大小各自为至少1kb。在一些方面,所述一种或多种SAM载体的大小各自为2kb。在一些方面,所述一种或多种SAM载体的大小各自小于5kb。

[0110] 在一些方面,所述至少一种抗原编码核酸序列包含两种或更多种抗原编码核酸序列。在一些方面,每种抗原编码核酸序列彼此直接连接。在一些方面,每种抗原编码核酸序列用编码接头的核酸序列连接到不同的抗原编码核酸序列。在一些方面,所述接头连接两个MHC I类表位编码核酸序列或一个MHC I类表位编码核酸序列至一个MHC II类表位编码核酸序列。在一些方面,所述接头选自由以下组成的组:(1)连续甘氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(2)连续丙氨酸残基,长度为至少2、3、4、5、6、7、8、9或10个残基;(3)两个精氨酸残基(RR);(4)丙氨酸、丙氨酸、酪氨酸(AAY);(5)由哺乳动物蛋白酶体有效加工的长度为至少2、3、4、5、6、7、8、9或10个氨基酸残基的共有序列;和(6)侧接源自同源蛋白的抗原并且长度为至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或2-20个氨基酸残基的一种或多种原生序列。在一些方面,所述接头连接两个MHC II类表位编码核酸序列或一个MHC II类序列至一个MHC I类表位编码核酸序列。在一些方面,所述接头包含序列GPGPG。

[0111] 在一些方面,所述抗原编码核酸序列可操作地或直接地连接到增强所述表位编码核酸序列的表达、稳定性、细胞运输、加工和呈递和/或免疫原性的单独或连续的序列。在一些方面,所述单独或连续的序列包含以下中的至少一者:泛素序列、经修饰以增加蛋白酶体靶向的泛素序列(例如,所述泛素序列在位置76处含有Gly至Ala取代)、免疫球蛋白信号序列(例如,IgK)、主要组织相容性I类序列、溶酶体相关膜蛋白(LAMP)-1、人树突状细胞溶酶体相关膜蛋白和主要组织相容性II类序列;任选地其中所述经修饰以增加蛋白酶体靶向的泛素序列是A76。

[0112] 在一些方面,所述至少一种抗原编码核酸序列包含至少2-10、2、3、4、5、6、7、8、9或10种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列,任选地其中每种抗原编码核酸序列编码不同的抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种抗原编码核酸序列。在一些方面,所述至少一种抗原编码核酸序列包含至少2-400种抗原编码核酸序列,并且其中所述抗原编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。在一些方面,每种抗原编码核酸序列独立地包含至少2-10、2、3、4、5、6、7、8、9或10种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列,任选地其中每种表位编码核酸序列编码不同的表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少11-20、15-20、11-100、11-200、11-300、11-400、11、12、13、14、15、16、17、18、19、20或多达400种表位编码核酸序列。在一些方面,每种抗原编码核酸序列独立地包含至少2-400种表位编码核酸序列,并且其中所述表位编码核酸序列中的至少两者编码由MHC I类呈递在细胞表面上的表位序列或其部分。在一些方面,所述MHC I类表位中的至少两者由MHC I类呈递在细胞表面、任选地肿瘤细

胞表面或受感染细胞表面上。

[0113] 在一些方面,所述表位编码核酸序列包含至少一种MHC I类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为8至35个氨基酸,任选地长度为9-17、9-25、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34或35个氨基酸的多肽序列。

[0114] 在一些方面,存在所述至少一种MHC II类表位编码核酸序列。在一些方面,所述至少一种MHC II类表位编码核酸序列存在并包含至少一种MHC II类表位编码核酸序列,所述MHC II类表位编码核酸序列包含使所编码的表位序列与由野生型核酸序列编码的相应肽序列不同的至少一种改变。

[0115] 在一些方面,所述表位编码核酸序列包含MHC II类表位编码核酸序列,并且其中每种抗原编码核酸序列编码长度为12-20、12、13、14、15、16、17、18、19、20或20-40个氨基酸的多肽序列。在一些方面,所述表位编码核酸序列包含MHC II类表位编码核酸序列,其中存在至少一种MHC II类表位编码核酸序列,并且其中所述至少一种MHC II类表位编码核酸序列包含至少一种通用MHC II类表位编码核酸序列,任选地其中所述至少一种通用序列包含破伤风类毒素和PADRE中的至少一者。

[0116] 在一些方面,所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是诱导型的。在一些方面,所述至少一种启动子核苷酸序列或所述第二启动子核苷酸序列是非诱导型的。

[0117] 在一些方面,所述至少一种聚(A)序列包含所述自我复制RNA原生的聚(A)序列。在一些方面,所述至少一种聚(A)序列包含所述自我复制RNA外源的聚(A)序列。在一些方面,所述至少一种聚(A)序列可操作地连接到所述至少一种核酸序列中的至少一者。在一些方面,所述至少一种聚(A)序列是至少20个、至少30个、至少40个、至少50个、至少60个、至少70个、至少80个、至少90个、至少100个、至少110个或至少120个连续A核苷酸。在一些方面,所述至少一种聚(A)序列是至少80个连续A核苷酸。在一些方面,所述至少一种聚(A)序列是至少100个连续A核苷酸。

[0118] 在一些方面,所述表位编码核酸序列包含MHC I类表位编码核酸序列,并且其中所述MHC I类表位编码核酸序列是通过执行以下步骤来选择:(a)从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;(b)将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及(c)基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生所述MHC I类表位编码核酸序列。

[0119] 在一些方面,所述MHC I类表位编码核酸序列中的每一者是通过执行以下步骤来选择:(a)从肿瘤、受感染细胞或感染性疾病生物体获得外显子组、转录组或全基因组核苷酸测序数据中的至少一者,其中所述核苷酸测序数据用于获得代表一组表位中的每一者的肽序列的数据;(b)将每个表位的肽序列输入到呈递模型中,以产生所述表位中的每一者由MHC等位基因中的一者或多者在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上呈递

的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及(c)基于所述数值可能性集合选择所述表位集合的子集以产生经选择表位集合,所述经选择表位用于产生至少20种MHC I类表位编码核酸序列。在一些方面,所述经选择表位集合的数目为2-20。在一些方面,所述呈递模型表示以下两者之间的依赖性:(a)所述MHC等位基因中的特定一者和在肽序列的特定位置处的特定氨基酸的一对的存在;和(b)由所述对的所述MHC等位基因中的所述特定一者将在所述特定位置处包含所述特定氨基酸的所述肽序列呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位被呈递在细胞表面、任选地肿瘤细胞表面或受感染细胞表面上的可能性增加的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激肿瘤特异性或感染性疾病生物体特异性免疫应答的可能性增加的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够由专职抗原呈递细胞(APC)呈递给初始T细胞的可能性增加的表位,任选地其中所述APC是树突状细胞(DC)。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位经由中心或外周耐受性受抑制的可能性降低的表位。在一些方面,选择所述经选择表位集合包括基于所述呈递模型选择相对于未选择的表位能够在受试者中刺激对正常组织的自身免疫应答的可能性降低的表位。在一些方面,外显子组或转录组核苷酸测序数据是通过肿瘤细胞或组织、受感染细胞或感染性疾病生物体进行测序而获得的。在一些方面,所述测序是下一代测序(NGS)或任何大规模并行测序方法。

[0120] 在一些方面,用于递送所述自扩增表达系统的所述组合物作为初免疫苗施用。在一些方面,所述方法还包括施用第二组合物,任选地其中所述第二组合物是疫苗组合物。在一些方面,所述第二组合物是在用于递送所述自扩增表达系统的所述组合物之前施用的。在一些方面,所述第二组合物是在施用用于递送所述自扩增表达系统的所述组合物之后施用的。在一些方面,所述第二组合物与用于递送所述自扩增表达系统的所述组合物相同。在一些方面,所述第二组合物与用于递送所述自扩增表达系统的所述组合物不同。在一些方面,所述第二组合物包含所述自扩增表达系统的盒,任选地其中所述第二组合物包含编码所述自扩增表达系统的盒的黑猩猩腺病毒载体。在一些方面,施用两种或更多种第二组合物,任选地其中用于递送所述自扩增表达系统的所述组合物作为初免疫苗施用。

[0121] 在一些方面,用于递送所述自扩增表达系统的所述组合物是肌肉内(IM)、皮内(ID)、皮下(SC)、玻璃体内(IVT)、鞘内或静脉内(IV)施用的。在一些方面,所述方法还包括施用免疫调节剂,任选地其中所述免疫调节剂是抗CTLA4抗体或其抗原结合片段、抗PD-1抗体或其抗原结合片段、抗PD-L1抗体或其抗原结合片段、抗4-1BB抗体或其抗原结合片段、抗OX-40抗体或其抗原结合片段、或细胞因子,任选地其中所述细胞因子是IL-2、IL-7、IL-12、IL-15或IL-21或其变体中的至少一者。在一些方面,所述方法还包括施用佐剂。

附图说明

[0122] 关于以下描述和附图将更好地理解本发明的这些和其他特征、方面和优势,在所述附图中:

[0123] 图(FIG.) 1图示了使用典型T7启动子或修饰的(“最小”)T7启动子进行的SAM载体

转录。

[0124] 图2提供了代表性的AU-SAM载体的示意图。

[0125] 图3显示了通过IVT使用三核苷酸 m^7G -ppp-A-U帽类似物或二核苷酸 m^7G -ppp-A帽类似物产生的加帽AU-SAM RNA产量。

[0126] 图4显示了用10ug指定SAM-LNP免疫的Ba1b/c小鼠(每组n=8)和免疫后12天分离的脾细胞。在用AH1-A5抗原(PSYAYHQF)刺激6小时后,通过IFN γ 的细胞内细胞因子染色来测量抗原特异性T细胞的数量。数据呈现为IFN γ +细胞占CD8+细胞的百分比,减去阴性对照肽的背景信号。条形表示中位数。

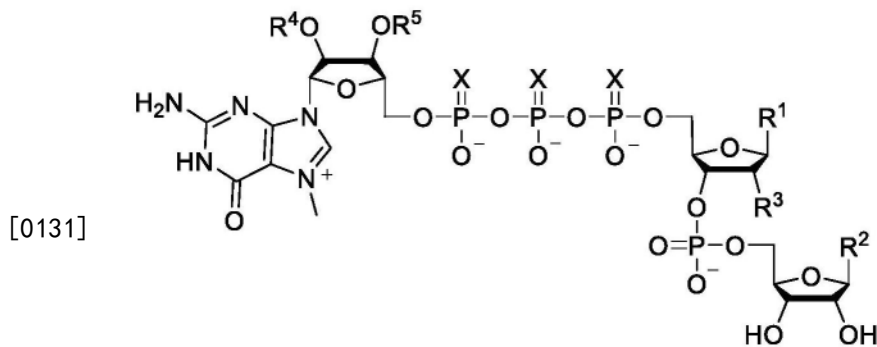
[0127] 图5说明了AU-SAM研究组细节(上图)和使用的模型抗原(下图)。

[0128] 图6显示了在用AU-SAM免疫(初免/加强)后六种Mamu-A*01中的每一者的抗原特异性免疫应答的时间进程。

[0129] 图7显示了在用AU-SAM免疫(初免/加强)后六种Mamu-A*01中的每一者的抗原特异性免疫应答的时间进程。

具体实施方式

[0130] 在一些实施方案中,本公开包括一种式(I)化合物:



(I),

[0132] 或其药学上可接受的盐,

[0133] 其中

[0134] R^1 是核苷;

[0135] R^2 是核苷;

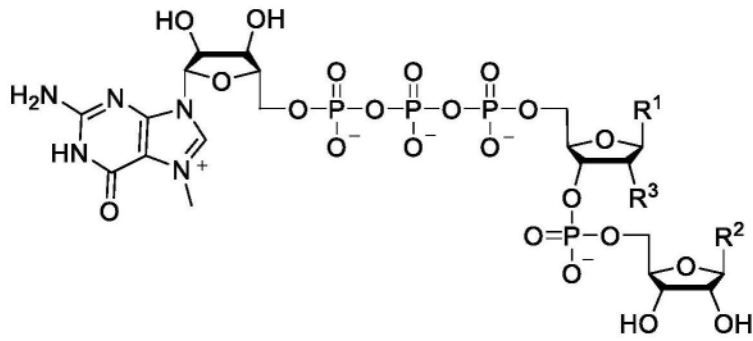
[0136] R^3 是卤素、任选地经取代的 C_1 - C_3 烷基或经取代的 C_1 - C_3 烷氧基; R^4 是氢或任选地经取代的 C_1 - C_3 脂族;

[0137] R^5 是氢或任选地经取代的 C_1 - C_3 脂族;并且

[0138] 每个X独立地是O或S。

[0139] 在一些实施方案中,本公开包括一种式(I-1)化合物:

[0140]

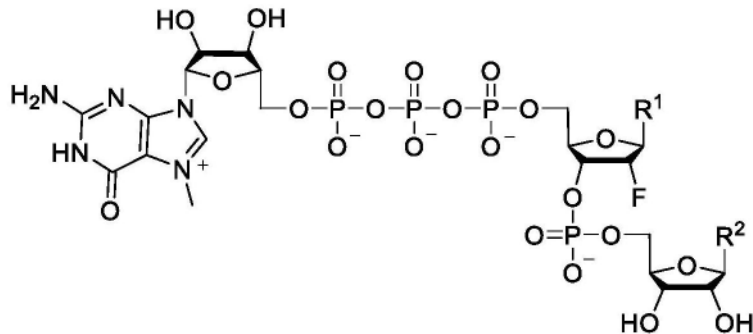


(I-1),

[0141] 或其药学上可接受的盐,其中 R^1 、 R^2 和 R^3 如上文所定义并在本文的类别和亚类中描述。

[0142] 在一些实施方案中,本公开包括一种式 (I-2) 化合物:

[0143]

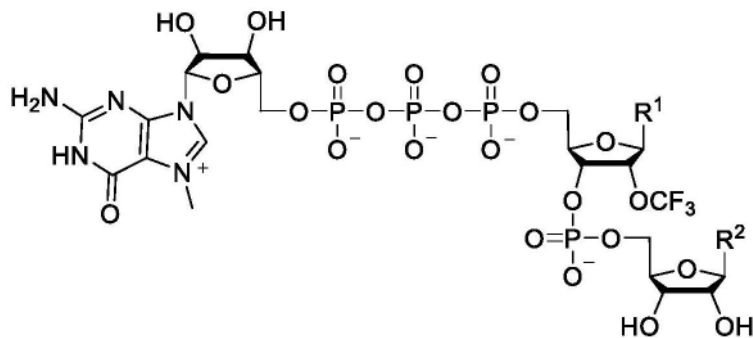


(I-2),

[0144] 或其药学上可接受的盐,其中 R^1 和 R^2 如上文所定义并在本文的类别和亚类中描述。

[0145] 在一些实施方案中,本公开包括一种式 (I-3) 化合物:

[0146]

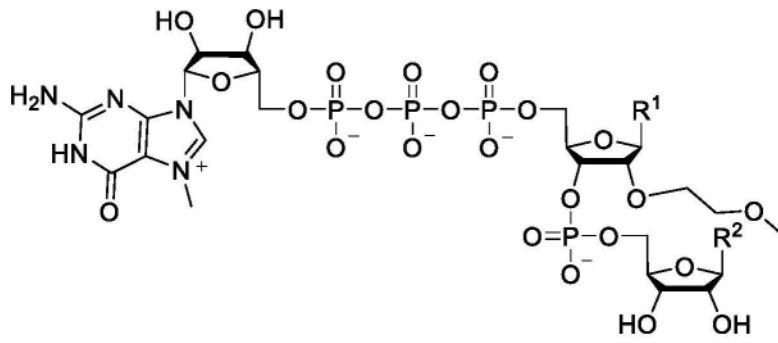


(I-3),

[0147] 或其药学上可接受的盐,其中 R^1 和 R^2 如上文所定义并在本文的类别和亚类中描述。

[0148] 在一些实施方案中,本公开包括一种式 (I-4) 化合物:

[0149]

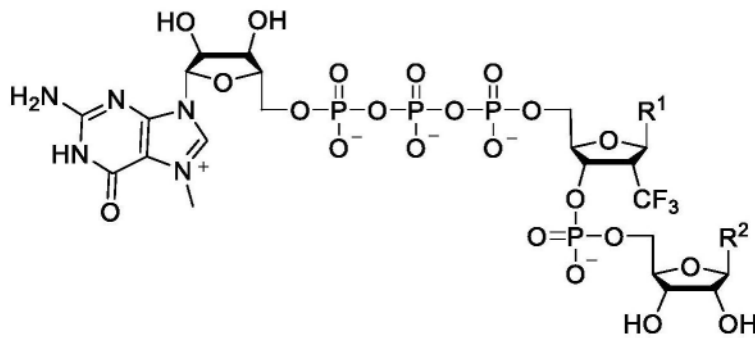


(I-4),

[0150] 或其药学上可接受的盐,其中R¹和R²如上文所定义并在本文的类别和亚类中描述。

[0151] 在一些实施方案中,本公开包括一种式 (I-5) 化合物:

[0152]

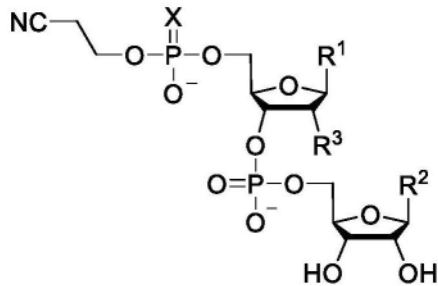


(I-5),

[0153] 或其药学上可接受的盐,其中R¹和R²如上文所定义并在本文的类别和亚类中描述。

[0154] 在一些实施方案中,本公开包括一种式 (II) 化合物:

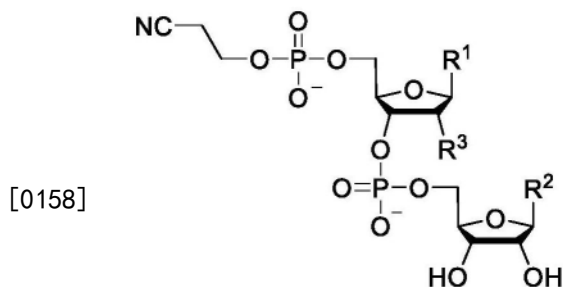
[0155]



(II),

[0156] 或其药学上可接受的盐,其中R¹、R²、R³和X如上文所定义并在本文的类别和亚类中描述。

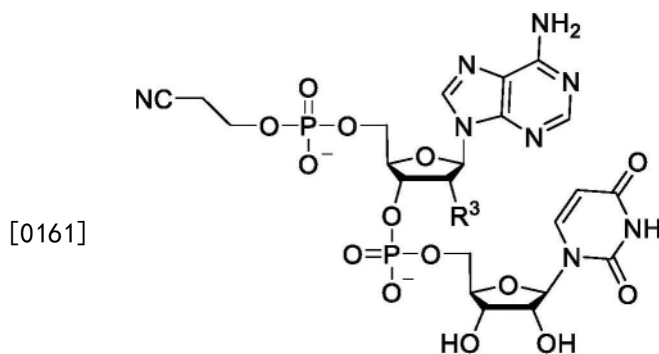
[0157] 在一些实施方案中,本公开包括一种式 (II-1) 化合物:



(II-1),

[0159] 或其药学上可接受的盐,其中 R^1 、 R^2 和 R^3 如上文所定义并在本文的类别和亚类中描述。

[0160] 在一些实施方案中,本公开包括一种式(II-2)化合物:



(II-2),

[0162] 或其药学上可接受的盐,其中 R^3 如上文所定义并在本文的类别和亚类中描述。

[0163] 在一些实施方案中, R^1 选自由腺嘌呤、尿嘧啶、鸟嘌呤和胞嘧啶组成的组。在一些实施方案中, R^1 是腺嘌呤。在一些实施方案中, R^1 是N6-甲基化腺嘌呤。在一些实施方案中, R^1 是尿嘧啶。在一些实施方案中, R^1 是鸟嘌呤。在一些实施方案中, R^1 是胞嘧啶。在一些实施方案中, R^1 是胸腺嘧啶。

[0164] 在一些实施方案中, R^2 选自由腺嘌呤、尿嘧啶、鸟嘌呤和胞嘧啶组成的组。在一些实施方案中, R^2 是腺嘌呤。在一些实施方案中, R^2 是尿嘧啶。在一些实施方案中, R^2 是鸟嘌呤。在一些实施方案中, R^2 是胞嘧啶。在一些实施方案中, R^2 是胸腺嘧啶。

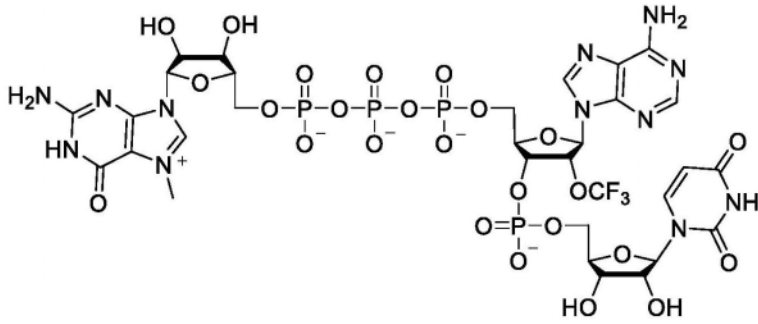
[0165] 在一些实施方案中, R^3 是卤素、任选地经取代的 C_1 - C_3 烷基或经取代的 C_1 - C_3 烷氧基。在一些实施方案中, R^3 是卤素。在一些实施方案中, R^3 是F。在一些实施方案中, R^3 是任选地经取代的 C_1 - C_3 烷基。在一些实施方案中, R^3 是 $-CF_3$ 。在一些实施方案中, R^3 是经取代的 C_1 - C_3 烷氧基。在一些实施方案中, R^3 是 C_1 - C_3 卤代烷氧基。在一些实施方案中, R^3 是 $-OCF_3$ 。在一些实施方案中, R^3 是被 C_1 - C_3 烷氧基取代的 C_1 - C_3 烷氧基。在一些实施方案中, R^3 是 $-OCH_2CH_2OCH_3$ 。

[0166] 在一些实施方案中, R^4 是氢或任选地经取代的 C_1 - C_3 脂族。在一些实施方案中, R^4 是氢。在一些实施方案中, R^4 是任选地经取代的 C_1 - C_3 脂族。在一些实施方案中, R^4 是氢或任选地经取代的甲基。在一些实施方案中, R^4 是甲基。

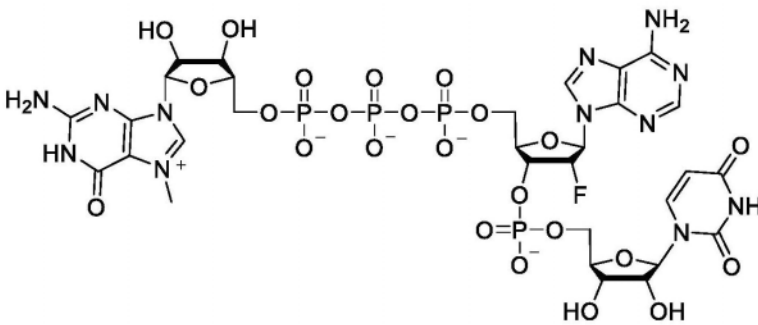
[0167] 在一些实施方案中, R^5 是氢或任选地经取代的 C_1 - C_3 脂族。在一些实施方案中, R^5 是氢。在一些实施方案中, R^5 是任选地经取代的 C_1 - C_3 脂族。在一些实施方案中, R^5 是氢或任选

地经取代的甲基。在一些实施方案中，R⁵是甲基。

[0168] 在一些实施方案中，本公开包括一种选自由以下组成的组的化合物：

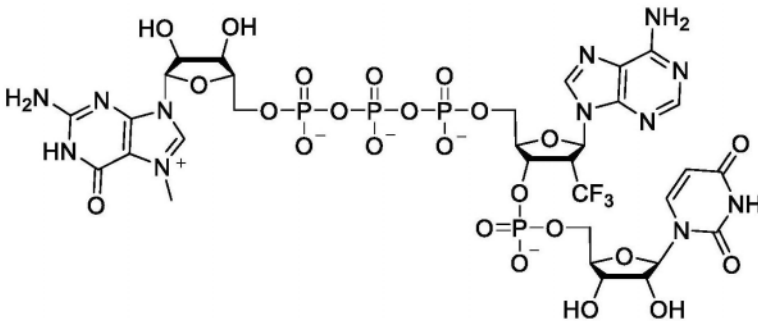


1、

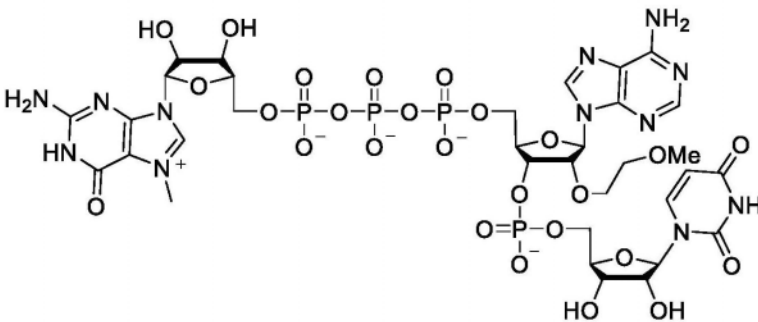


[0169]

2、



3、和

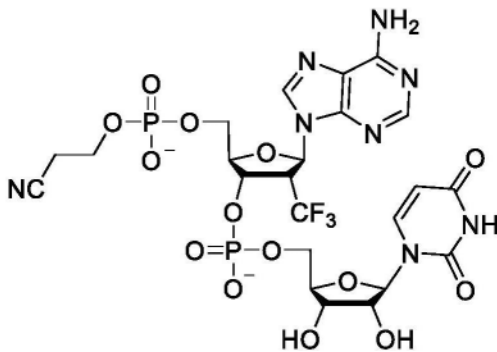
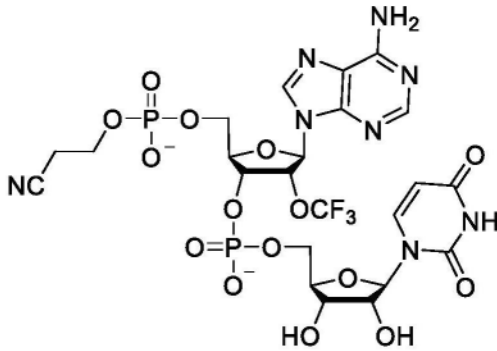


[0170]

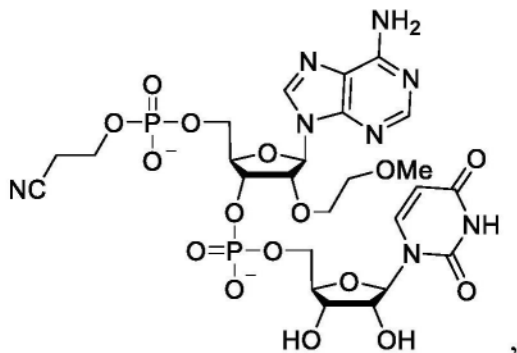
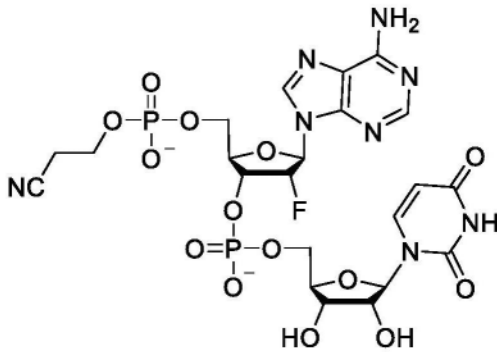
4、

[0171] 或其药学上可接受的盐。

[0172] 在一些实施方案中,本公开包括一种化合物,其包括:



[0173]



[0174] 或其药学上可接受的盐。

[0175] 定义

[0176] 如本文所用,术语“脂族”或“脂族团”是指完全饱和或含有一个或多个不饱和单元的直接(即,无支链)或支链、被取代或未经取代的烃链,或完全饱和或含有一个或多个不饱

和单元但不是芳族的单环烃或双环烃(本文也称为“碳环”、“脂环族”或“环烷基”),其具有与分子其余部分的单一连接点。除非另有说明,否则脂族团含有1-6个脂族碳原子。在一些实施方案中,脂族团含有1-5个脂族碳原子。在其他实施方案中,脂族团含有1-4个脂族碳原子。在其他实施方案中,脂族团含有1-3个脂族碳原子,并且在其他实施方案中,脂族团含有1-2个脂族碳原子。在一些实施方案中,“脂环族”(或“碳环”或“环烷基”)是指完全饱和或含有一个或多个不饱和单元但不是芳族的单环 C_3 - C_6 烃,其具有与分子其余部分的单一连接点。合适的脂族团包括但不限于直链或支链的被取代或未经取代的烷基、烯基、炔基及其杂化物,例如(环烷基)烷基、(环烯基)烷基或(环烷基)烯基。

[0177] 术语“卤代脂族”是指被一个或多个卤素原子取代的脂族团。

[0178] 术语“烷基”是指直链或支链烷基。示例性烷基是甲基、乙基、丙基、异丙基、丁基、异丁基和叔丁基。

[0179] 术语“卤代烷基”是指被一个或多个卤素原子取代的直链或支链烷基。

[0180] 术语“卤素”是指F、Cl、Br或I。

[0181] 单独使用或如在“芳烷基”、“芳烷氧基”或“芳氧基烷基”中作为较大部分的一部分使用的术语“芳基”是指具有总共五至十四个环成员的单环和双环环系,其中体系中的至少一个环是芳族的并且其中体系中的每个环含有三至七个环成员。术语“芳基”可与术语“芳基环”互换使用。在本公开的某些实施方案中,“芳基”是指芳族环系,其包括但不限于苯基、联苯基、萘基、蒽基等,其可以带有一个或多个取代基。如本文所用,术语“芳基”的范围内还包括芳环稠合至一个或多个非芳环的基团,例如茚满基、邻苯二甲酰亚胺基、萘酰亚胺基、菲啶基或四氢萘基等。

[0182] 如本文所用,术语“部分不饱和”是指包括至少一个双键或三键的环部分。术语“部分不饱和”旨在涵盖具有多个不饱和位点的环,但不旨在包括如本文所定义的芳基或杂芳基部分。

[0183] 如本文所述,本公开的化合物可含有“任选地经取代的”部分。通常,术语“经取代的”,无论前面是否有术语“任选地”,是指指定部分的一个或多个氢被合适的取代基代替。除非另有说明,否则“任选地经取代的”基团可在所述基团的每个可取代位置具有合适的取代基,并且当任何给定结构中的多于一个位置可被多于一个选自指定基团的取代基取代时,所述取代基可在每个位置相同或不同。本公开所设想的取代基的组合优选是导致形成稳定的或化学上可行的化合物的那些。如本文所用,术语“稳定的”是指当经受允许其生产、检测以及在某些实施方案中其回收、纯化和用于本文公开的一个或多个目的的条件时基本上没有改变的化合物。

[0184] “任选地经取代的”基团的可取代碳原子上合适的单价取代基独立地是卤素; $-(CH_2)_{0-4}R^{\circ}$; $-(CH_2)_{0-4}OR^{\circ}$; $-O(CH_2)_{0-4}R^{\circ}$; $-O-(CH_2)_{0-4}C(O)OR^{\circ}$; $-(CH_2)_{0-4}CH(OR^{\circ})_2$; $-(CH_2)_{0-4}SR^{\circ}$; $-(CH_2)_{0-4}Ph$,其可被 R° 取代; $-(CH_2)_{0-4}O(CH_2)_{0-1}Ph$,其可被 R° 取代; $-CH=CHPh$,其可被 R° 取代; $-(CH_2)_{0-4}O(CH_2)_{0-1}$ -吡啶基,其可被 R° 取代; $-NO_2$; $-CN$; $-N_3$; $-(CH_2)_{0-4}N(R^{\circ})_2$; $-(CH_2)_{0-4}N(R^{\circ})C(O)R^{\circ}$; $-N(R^{\circ})C(S)R^{\circ}$; $-(CH_2)_{0-4}N(R^{\circ})C(O)NR^{\circ}_2$; $-N(R^{\circ})C(S)NR^{\circ}_2$; $-(CH_2)_{0-4}N(R^{\circ})C(O)OR^{\circ}$; $-N(R^{\circ})N(R^{\circ})C(O)R^{\circ}$; $-N(R^{\circ})N(R^{\circ})C(O)NR^{\circ}_2$; $-N(R^{\circ})N(R^{\circ})C(O)OR^{\circ}$; $-(CH_2)_{0-4}C(O)R^{\circ}$; $-C(S)R^{\circ}$; $-(CH_2)_{0-4}C(O)OR^{\circ}$; $-(CH_2)_{0-4}C(O)SR^{\circ}$; $-(CH_2)_{0-4}C(O)OSiR^{\circ}_3$; $-(CH_2)_{0-4}OC(O)R^{\circ}$; $-OC(O)(CH_2)_{0-4}SR^{\circ}$; $SC(S)SR^{\circ}$; $-(CH_2)_{0-4}SC(O)R^{\circ}$; $-(CH_2)_{0-4}C(O)NR^{\circ}_2$; $-C(S)NR^{\circ}_2$; $-C(S)SR^{\circ}$; $-SC(S)$

SR^o; -(CH₂)₀₋₄OC(O)NR^o₂; -C(O)N(OR^o)R^o; -C(O)C(O)R^o; -C(O)CH₂C(O)R^o; -C(NOR^o)R^o; -(CH₂)₀₋₄SSR^o; -(CH₂)₀₋₄S(O)₂R^o; -(CH₂)₀₋₄S(O)₂OR^o; -(CH₂)₀₋₄OS(O)₂R^o; -S(O)₂NR^o₂; -(CH₂)₀₋₄S(O)R^o; -N(R^o)S(O)₂NR^o₂; -N(R^o)S(O)₂R^o; -N(OR^o)R^o; -C(NH)NR^o₂; -P(O)₂R^o; -P(O)R^o₂; -OP(O)R^o₂; -OP(O)(OR^o)₂; SiR^o₃; -(C₁₋₄直链或支链亚烷基)O-N(R^o)₂; 或-(C₁₋₄直链或支链亚烷基)C(O)O-N(R^o)₂, 其中每个R^o可如下文所定义被取代并且独立地是氢、C₁₋₆脂族、-CH₂Ph、-O(CH₂)₀₋₁Ph、-CH₂- (5-6元杂芳基环) 或具有0-4个独立地选自氮、氧或硫的杂原子的5-6元饱和、部分不饱和或芳基环, 或尽管有上述定义, 两个独立出现的R^o与其间插原子一起形成具有0-4个独立地选自氮、氧或硫的杂原子的3-12元饱和、部分不饱和或芳基单环或双环, 其可以如下文所定义被取代。

[0185] R^o (或由两次独立出现的R^o与其间插原子一起形成的环) 上合适的单价取代基独立地是卤素、-(CH₂)₀₋₂R[•]、-(卤基R[•])、-(CH₂)₀₋₂OH、-(CH₂)₀₋₂OR[•]、-(CH₂)₀₋₂CH(OR[•])₂; -O(卤基R[•])、-CN、-N₃、-(CH₂)₀₋₂C(O)R[•]、-(CH₂)₀₋₂C(O)OH、-(CH₂)₀₋₂C(O)OR[•]、-(CH₂)₀₋₂SR[•]、-(CH₂)₀₋₂SH、-(CH₂)₀₋₂NH₂、-(CH₂)₀₋₂NHR[•]、-(CH₂)₀₋₂NR[•]₂、-NO₂、-SiR[•]₃、-OSiR[•]₃、-C(O)SR[•]、-(C₁₋₄直链或支链亚烷基)C(O)OR[•]或-SSR[•], 其中每个R[•]未被取代或在前面有“卤基”时仅被一个或多个卤素取代, 并且独立地选自C₁₋₄脂族、-CH₂Ph、-O(CH₂)₀₋₁Ph或具有0-4个独立地选自氮、氧或硫的杂原子的5-6元饱和、部分不饱和或芳基环。R^o的饱和碳原子上的合适的二价取代基包括 -O和 -S。

[0186] “任选地经取代的”基团的饱和碳原子上的合适二价取代基包括以下: -O、-S、-NNR*₂、-NNHC(O)R*、-NNHC(O)OR*、-NNHS(O)₂R*、-NR*、-NOR*、-O(C(R*₂))₂₋₃O-或-S(C(R*₂))₂₋₃S-, 其中每个独立出现的R*选自氢、C₁₋₆脂族 (其可以如下文所定义被取代), 或具有0-4个独立地选自氮、氧或硫的杂原子的未经取代的5-6元饱和、部分不饱和或芳基环。与“任选地经取代的”基团的邻位可取代碳结合的合适二价取代基包括: -O(CR*₂)₂₋₃O-, 其中每个独立出现的R*选自氢、C₁₋₆脂族 (其可以如下文所定义被取代), 或具有0-4个独立地选自氮、氧或硫的杂原子的未经取代的5-6元饱和、部分不饱和或芳基环。

[0187] R*的脂族团上合适的取代基包括卤素、-R[•]、-(卤基R[•])、-OH、-OR[•]、-O(卤基R[•])、-CN、-C(O)OH、-C(O)OR[•]、-NH₂、-NHR[•]、-NR[•]₂或-NO₂, 其中每个R[•]未被取代或在前面有“卤基”时仅被一个或多个卤素取代, 并且独立地是C₁₋₄脂族、-CH₂Ph、-O(CH₂)₀₋₁Ph, 或具有0-4个独立地选自氮、氧或硫的杂原子的5-6元饱和、部分不饱和或芳基环。

[0188] “任选地经取代的”基团的可取代氮上的合适取代基包括 -R[†]、-NR[†]₂、-C(O)R[†]、-C(O)OR[†]、-C(O)C(O)R[†]、-C(O)CH₂C(O)R[†]、-S(O)₂R[†]、-S(O)₂NR[†]₂、-C(S)NR[†]₂、-C(NH)NR[†]₂或-N(R[†])S(O)₂R[†]; 其中每个R[†]独立地是氢、C₁₋₆脂族 (其可以如下文所定义被取代)、未经取代的-OPh, 或具有0-4个独立地选自氮、氧或硫的杂原子的未经取代的5-6元饱和、部分不饱和或芳基环, 或者, 尽管有上述定义, 两个独立出现的R[†]连同其间插原子一起形成具有0-4个独立地选自氮、氧或硫的杂原子的未经取代的3-12元饱和、部分不饱和或芳基单环或双环。

[0189] R[†]的脂族团上的合适取代基独立地是卤素、-R[•]、-(卤基R[•])、-OH、-OR[•]、-O(卤基R[•])、-CN、-C(O)OH、-C(O)OR[•]、-NH₂、-NHR[•]、-NR[•]₂或-NO₂, 其中每个R[•]未被取代或在前面有

“卤基”时仅被一个或多个卤素取代,并且独立地是 C_{1-4} 脂族、 $-CH_2Ph$ 、 $-O(CH_2)_{0-1}Ph$,或具有0-4个独立地选自氮、氧或硫的杂原子的5-6元饱和、部分不饱和或芳基环。

[0190] 如本文所用,术语“药学上可接受的盐”是指在合理的医学判断范围内,适用于与人类和低等动物的组织接触而没有过度毒性、刺激性、过敏反应等并且与合理的益处/风险比相称的那些盐。药学上可接受的盐是本领域所熟知的。例如,S.M.Berge等人在J.Pharmaceutical Sciences,1977,66,1-19中详细描述了药学上可接受的盐,所述文献通过引用并入本文。本公开的化合物的药学上可接受的盐包括衍生自合适的无机和有机酸和碱的那些。药学上可接受的无毒酸加成盐的实例是氨基与无机酸如盐酸、氢溴酸、磷酸、硫酸和高氯酸或与有机酸如乙酸、草酸、马来酸、酒石酸、柠檬酸、琥珀酸或丙二酸或通过使用本领域中使用的其他方法如离子交换形成的盐。其他药学上可接受的盐包括己二酸盐、藻酸盐、抗坏血酸盐、天冬氨酸盐、苯磺酸盐、苯甲酸盐、硫酸氢盐、硼酸盐、丁酸盐、樟脑酸盐、樟脑磺酸盐、柠檬酸盐、环戊烷丙酸盐、二葡萄糖酸盐、十二烷基硫酸盐、乙磺酸盐、甲酸盐、富马酸盐、葡庚糖酸盐、甘油磷酸盐、葡糖酸盐、半硫酸盐、庚酸盐、己酸盐、氢碘酸盐、2-羟基-乙磺酸盐、乳糖酸盐、乳酸盐、月桂酸盐、月桂基硫酸盐、苹果酸盐、马来酸盐、丙二酸盐、甲磺酸盐、2-萘磺酸盐、烟酸盐、硝酸盐、油酸盐、草酸盐、棕榈酸盐、双羟萘酸盐、果胶酸盐、过硫酸盐、3-苯丙酸盐、磷酸盐、新戊酸盐、丙酸盐、硬脂酸盐、琥珀酸盐、硫酸盐、酒石酸盐、硫氰酸盐、对甲苯磺酸盐、十一烷酸盐、戊酸盐等。

[0191] 衍生自适当碱的盐包括碱金属、碱土金属、铵和N(C1-4烷基)4盐。代表性的碱金属或碱土金属盐包括钠、锂、钾、钙、镁等。其他药学上可接受的盐在适当时包括无毒的铵盐、季铵盐和使用抗衡离子如卤离子、氢氧根离子、羧酸根离子、硫酸根离子、磷酸根离子、硝酸根离子、低级烷基磺酸根离子和芳基磺酸根离子形成的胺阳离子。

[0192] 在本文的任何变量定义中叙述的化学基团列表包括将该变量定义为任何单个基团或所列基团的组合。此处对变量的一个实施方案的叙述包括作为任何单个实施方案或与任何其他实施方案或其部分组合的实施方案。

[0193] 如本文所用,术语“生物样品”包括但不限于细胞培养物或其提取物;获自哺乳动物或其提取物的活检材料;以及血液、唾液、尿液、粪便、精液、眼泪或其他体液或其提取物。

[0194] 如本文所用,“治疗有效量”是指刺激所需生物反应的物质(例如,治疗剂、组合物和/或制剂)的量。在一些实施方案中,物质的治疗有效量是当作为给药方案的一部分施用于患有或易患疾病、病症和/或疾患的受试者时足以治疗、诊断、预防和/或延迟疾病、病症和/或疾患的发作的量。如本领域普通技术人员将理解的,物质的有效量可以根据例如期望的生物学终点、待递送的物质、目标细胞或组织等因素而变化。例如,用于治疗疾病、病症和/或疾患的制剂中提供的化合物的有效量是减轻、改善、缓解、抑制、预防、延迟疾病、病症和/或疾患的一种或多种症状或特征的发作,降低其严重程度和/或降低其发生率的量。在一些实施方案中,“治疗有效量”是所提供的化合物或含有所提供的化合物的组合物的至少最小量,其足以治疗疾病或病症的一种或多种症状。

[0195] 疾病、病症和疾患在本文中可互换使用。

[0196] 如本文所用,术语“治疗(treatment/treat/treating)”是指如本文所述的病症或疾患或者所述病症或疾患的一种或多种症状的部分或完全减轻、抑制、延迟发作、预防、改善和/或缓解。在一些实施方案中,可在一种或多种症状出现后施用治疗。在一些实施方案

中,术语“治疗”包括预防或停止疾病或病症的进展。在其他实施方案中,可在没有症状的情况下施用治疗。例如,可在症状发作之前对易感个体施用治疗(例如,根据症状史和/或根据遗传或其他易感性因素)。在症状消退后也可继续治疗,例如预防或延迟其再发生。因此,在一些实施方案中,术语“治疗”包括预防疾病或病症的复发或再发生。

[0197] 预期施用的“受试者”包括但不限于人类(即任何年龄组的男性或女性,例如儿科受试者(例如婴儿、儿童、青少年)或成人受试者(例如,年轻人、中年人或老年人))和/或非人类动物,例如哺乳动物如灵长类动物(例如食蟹猴、恒河猴)、牛、猪、马、绵羊、山羊、啮齿动物、猫和/或狗。在某些实施方案中,受试者是人类。在某些实施方案中,受试者是非人类动物。术语“患者”和“受试者”在本文中可互换使用。

[0198] 术语“药学上可接受的载剂、佐剂或媒介物”是指不破坏与其配制的化合物的药理学活性的无毒载体、佐剂或媒介物。可用于本文公开的化合物的组合物中的药学上可接受的载剂、佐剂或媒介物包括但不限于离子交换剂、氧化铝、硬脂酸铝、卵磷脂、血清蛋白如人血清白蛋白、缓冲物质如磷酸盐、甘氨酸、山梨酸、山梨酸钾、饱和植物脂肪酸的部分甘油酯混合物、水、盐或电解质,例如硫酸鱼精蛋白、磷酸氢二钠、磷酸氢钾、氯化钠、锌盐、胶体二氧化硅、三硅酸镁、聚乙烯吡咯烷酮、纤维素基物质、聚乙二醇、羧甲基纤维素钠、聚丙烯酸酯、蜡、聚乙烯-聚氧丙烯嵌段聚合物、聚乙二醇和羊毛脂。

[0199] 替代实施方案

[0200] 在一个替代实施方案中,本文所述的化合物还可包含一个或多个同位素取代。例如,氢可以是 ^2H (D或氘)或 ^3H (T或氚);碳可以是例如 ^{13}C 或 ^{14}C ;氧可以是例如 ^{18}O ;氮可以是例如 ^{15}N 等。在其他实施方案中,特定同位素(例如, ^3H 、 ^{13}C 、 ^{14}C 、 ^{18}O 或 ^{15}N)可代表占据化合物特定位点的元素总同位素丰度的至少1%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少99%或至少99.9%。

[0201] 药物组合物

[0202] 在一些实施方案中,本公开提供了包含式(I)化合物和药学上可接受的载剂、佐剂或媒介物的组合物。为了便于施用和剂量均匀性,本公开的化合物优选配制成剂量单位形式。

[0203] 使用本公开化合物的方法-RNA寡核苷酸的合成

[0204] 在一些实施方案中,式(I)化合物可用于制备5'加帽的RNA。本文考虑的用于制备5'加帽的RNA的方法和组合物包括但不限于mRNA、小核RNA(snRNA)、小核仁RNA(snoRNA)、小卡哈尔体(cajal body)特异性RNA(scaRNA)。在一些实施方案中,一种方法涉及使用含有寡核苷酸引物、核苷5'-三磷酸(NTP)和RNA聚合酶的帽用于DNA模板化和启动子控制的RNA合成。在某些方面,一种方法使用在RNA合成、特别是加帽mRNA的合成中提供效用的起始加帽寡核苷酸引物。

[0205] 在一些实施方案中,式(I)化合物可用于制备RNA的方法中,所述RNA包括但不限于mRNA、snRNA、snoRNA、scaRNA、转移RNA(tRNA)、核糖体RNA(rRNA)和转移-信使RNA(tmRNA),其在分子的5'端处或附近携带修饰。在一些实施方案中,一种方法涉及使用带有或不带有帽、核苷5'-三磷酸(NTP)和RNA聚合酶的起始寡核苷酸引物用于DNA模板化和启动子控制的RNA合成。在某些方面,一种方法使用修饰的起始寡核苷酸引物,其携带在RNA合成、特别是

5'-修饰的RNA合成中提供效用的结构修饰。

[0206] 起始加帽寡核苷酸引物具有开放的3'-OH基团,其通过在引物的3'端添加核苷酸单元而允许在DNA模板上起始RNA聚合酶介导的RNA合成。起始加帽寡核苷酸引物在转录起始位点(即,起始位点位于更接近启动子序列的3'末端并且可与启动子序列重叠)与模板DNA序列基本互补,在某些实施方案中,起始加帽寡核苷酸引物主要从引物的3'端开始在一个方向(“正向”)指导RNA的合成。在某些方面和实施方案中,起始加帽寡核苷酸引物在起始RNA合成方面胜过任何核苷5'-三磷酸,从而使以起始加帽寡核苷酸引物开始的RNA产生最大化并且使以5'-三磷酸核苷(通常是GTP)开始的RNA产生最小化。

[0207] 本公开的起始加帽寡核苷酸引物具有杂交序列,该杂交序列可与起始位点处的DNA模板上的序列互补。杂交序列的存在迫使起始加帽寡核苷酸引物主要与起始位点处的DNA模板的互补序列仅在所需方向(即“正向”方向)上对齐。在正向方向上,RNA转录物以倒置的鸟苷残基(即^{7m}G(5')ppp(5')N.....)开始。通过杂交复合物的热力学维持了DNA模板上引物对齐的正向方向相对于不正确的“反向”方向的优势。后者由起始加帽寡核苷酸引物的杂交序列的长度和参与与DNA模板杂交的碱基的身份确定。所需正向方向的杂交也可能取决于在体外转录期间DNA模板和起始加帽寡核苷酸引物杂交或使用的温度和反应条件。

[0208] 与使用标准GTP、ATP、CTP或UTP的起始功效相比,本公开的起始加帽寡核苷酸引物增强了转录起始功效。在一些实施方案中,当RNA的合成主要从起始加帽寡核苷酸引物而不是从转录混合物中的任何NTP开始时,转录的起始被认为是增强的。转录起始效率的增强导致RNA转录物的产量更高。在无起始加帽引物的情况下,相比于使用常规方法合成RNA,增强的转录起始效率可增加到约10%、约20%、约40%、约60%、约80%、约90%、约100%、约150%、约200%或约500%。在某些实施方案中,“起始加帽寡核苷酸引物”在起始转录方面胜过任何NTP(包括GTP)。本领域普通技术人员能够容易地确定底物活性水平和起始加帽寡核苷酸引物的功效。确定底物功效的方法的一个实例在实施例13)中进行了说明。在某些实施方案中,起始从加帽寡核苷酸引物而不是NTP发生,这导致转录的mRNA的更高水平的加帽。

[0209] 在一些方面,提供了其中利用具有取代或修饰的起始加帽寡核苷酸引物合成RNA的方法。在一些方面,起始加帽寡核苷酸引物的取代和修饰基本上不损害RNA的合成。可预先形成常规测试合成以确定使用修饰的起始加帽寡核苷酸引物是否可获得期望的合成结果。本领域技术人员可进行这样的常规实验以确定是否可获得期望的结果。起始加帽寡核苷酸引物的取代或修饰包括例如一种或多种修饰的核苷碱基、一种或多种修饰的糖、一种或多种修饰的核苷酸间键和/或一种或多种修饰的三磷酸根桥。

[0210] 可包括本文提供的方法和组合物的一个或多个修饰基团的修饰的起始加帽寡核苷酸引物可通过将NTP并入到开放的3'-OH基团上而在DNA模板上由RNA聚合酶延长。起始加帽寡核苷酸引物可包括天然RNA和DNA核苷、修饰的核苷或核苷类似物。起始加帽寡核苷酸引物可含有天然核苷酸间磷酸二酯键或其修饰,或其组合。

[0211] 使用本公开化合物的方法-治疗方法

[0212] 在一些实施方案中,本公开提供了一种用于治疗患者癌症或减轻癌症严重程度的方法,所述方法包括向所述患者施用RNA寡核苷酸的步骤,其中所述RNA寡核苷酸包含式(I)化合物。

[0213] 在一些实施方案中,根据本公开的方法,化合物和组合物可使用有效治疗癌症或减轻癌症严重程度的任何量和任何施用途径进行施用。在一些实施方案中,癌症选自由以下组成的组:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性髓细胞白血病、成人急性淋巴母细胞白血病、慢性髓细胞白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

[0214] 在一些实施方案中,癌症是实体瘤。在一些实施方案中,癌症选自由以下组成的组:微卫星稳定结直肠癌(MSS-CRC)、非小细胞肺癌(NSCLC)、胰腺导管腺癌(PDA)和胃食管腺癌(GEA)。在一些实施方案中,癌症选自由以下组成的组:MSS-CRC、NSCLC和PDA。

[0215] 在一些实施方案中,将包含本公开的式(I)化合物的RNA寡核苷酸施用于患有癌症的患者,所述癌症选自由以下组成的组:肺癌、黑色素瘤、乳腺癌、卵巢癌、前列腺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、膀胱癌、脑癌、B细胞淋巴瘤、急性髓细胞白血病、成人急性淋巴母细胞白血病、慢性髓细胞白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病、非小细胞肺癌和小细胞肺癌。

[0216] 在一些实施方案中,将包含式(I)化合物的RNA寡核苷酸施用于患有感染的患者。在一些实施方案中,感染是病毒感染、真菌感染或细菌感染。在一些实施方案中,感染是病毒感染。在一些实施方案中,病毒感染是被病毒感染,其中所述病毒是HIV。在一些实施方案中,将包含式(I)化合物的RNA寡核苷酸施用于患有AIDS的患者。在一些实施方案中,病毒感染是被病毒感染,其中所述病毒是冠状病毒。在一些实施方案中,将包含式(I)化合物的RNA寡核苷酸施用于患有COVID-19的患者。

[0217] 在一些实施方案中,本公开涉及一种使生物样品与包含式(I)化合物的RNA寡核苷酸接触的方法。

[0218] 在一些实施方案中,一种或多种额外治疗剂也可与包含式(I)化合物的RNA寡核苷酸组合施用。在一些实施方案中,包含式(I)化合物和一种或多种额外治疗剂的RNA寡核苷酸可作为多剂量方案的一部分施用。在一些实施方案中,包含式(I)化合物和一种或多种额外治疗剂的RNA寡核苷酸可同时、依次或在一段时间内施用。在一些实施方案中,包含式(I)化合物和一种或多种额外治疗剂的RNA寡核苷酸可在彼此间隔五小时内施用。在一些实施方案中,包含式(I)化合物和一种或多种额外治疗剂的RNA寡核苷酸可在彼此间隔24小时内施用。在一些实施方案中,包含式(I)化合物和一种或多种额外治疗剂的RNA寡核苷酸可在彼此间隔一周内施用。

[0219] 自扩增mRNA载体

[0220] 一般而言,所有自扩增RNA(SAM)载体都含有源自自我复制病毒的自扩增骨架。术语“自扩增骨架”是指允许病毒基因组自我复制的自我复制病毒的最小序列。例如,允许甲病毒自我复制的最小序列可包括用于非结构蛋白介导的扩增的保守序列(例如,非结构蛋白1(nsP1)基因、nsP2基因、nsP3基因、nsP4基因和/或聚A序列)。自扩增骨架还可包括用于表达亚基因组病毒RNA的序列(例如,甲病毒的亚基因组启动子如26S启动子元件)。SAM载体可以是正义RNA多核苷酸或负义RNA多核苷酸,例如具有源自正义或负义自我复制病毒的骨架的载体。自我复制病毒包括但不限于甲病毒、黄病毒(例如昆津病毒(Kunjin virus))、麻疹病毒和弹状病毒(例如狂犬病病毒和水疱性口炎病毒)。源自自我复制病毒的SAM载体系

统的实例更详细地描述于Lundstrom (Molecules. 2018年12月13日; 23(12). pii: E3310. doi:10.3390/molecules23123310), 其出于所有目的通过引用并入本文。

[0221] 体外自扩增产生

[0222] 本领域中熟知的用于RNA产生的便利技术是体外转录 (IVT)。在这种技术中, 首先通过本领域技术人员熟知的技术产生所需载体的DNA模板, 包括标准分子生物学技术, 例如克隆、限制性消化、连接、基因合成 (例如化学和/或酶促合成) 和聚合酶链反应 (PCR)。

[0223] DNA模板在需要转录为RNA的序列的5'端含有一个RNA聚合酶启动子 (例如SAM)。启动子包括但不限于噬菌体聚合酶启动子, 例如T3、T7、SP6或K11。取决于所选的特定RNA聚合酶启动子序列, 除了所需序列外, 还可转录额外的5'核苷酸。例如, 典型的T7启动子可用序列TAATACGACTCACTATAGG提及, 其中使用DNA模板TAATACGACTCACTATAGG_{N_v}用于产生所需序列N的IVT反应将产生mRNA序列GG-N_v。一般而言, 并且不希望受理论束缚, T7聚合酶更有效地转录以鸟苷开始的RNA转录物。然而, 额外的5'核苷酸可能不需要和/或可能是有害的。因此, 包含在DNA模板中的RNA聚合酶启动子可以是导致转录物仅含有所需序列的5'核苷酸的序列, 例如, 具有SAM载体所来源的自我复制病毒的内源性 (也称为“原生”或“基因组”) 5'序列的SAM, 是指自我复制病毒的原生基因组序列 (例如, 具有内源5'VEEV核苷酸AU, 也称为“AU-SAM”)。例如, 最小T7启动子可由序列TAATACGACTCACTATA (定向5'-3'; ϕ 6.5 T7启动子) 提及, 其中使用DNA模板TAATACGACTCACTATAN₁N₂N_v用于产生所需序列N的IVT反应将产生mRNA序列N₁N₂N_v。另一种最小T7启动子可由序列TAATACGACTCACTATT (定向5'-3'; ϕ 2.5 T7启动子) 提及。同样, 由序列ATTTAGGTGACACTATA提及的最小SP6启动子可用于在不存在额外5'核苷酸下产生转录物。同样, 由序列AATTAGGGCACACTATA提及的最小K11启动子可用于在不存在额外5'核苷酸下产生转录物。在典型的IVT反应中, DNA模板与适当的RNA聚合酶、缓冲剂和核苷酸 (NTP) 一起孵育。

[0224] 所得的RNA多核苷酸可任选地进一步经修饰, 包括但不限于添加5'帽结构, 例如7-甲基鸟苷或相关结构, 以及任选地修饰3'端以包括聚腺苷酸 (聚A) 尾。在改进的IVT反应中, 通过在IVT期间添加帽类似物, RNA以5'帽结构共转录加帽。帽类似物可包括二核苷酸 (m⁷G-ppp-N) 帽类似物或三核苷酸 (m⁷G-ppp-N₁-N₂) 帽类似物, 其中N表示核苷酸或修饰的核苷酸 (例如, 核糖核苷, 包括但不限于腺苷、鸟苷、胞苷和尿嘧啶)。修饰的核苷酸可包括修饰的腺苷, 例如2'-OH-甲基化N₆-甲基腺苷。在包括三核苷酸 (m⁷G-ppp-N₁-N₂) 帽类似物的说明性非限制性实例中, N₁可以是2'-OH-甲基化N₆-甲基腺苷。帽类似物可包括本文所述的任何结构或式。示例性的帽类似物及其在IVT反应中的用途也更详细地描述于美国专利第10,519,189号中, 其出于所有目的通过引用并入本文。如所讨论的, T7聚合酶更有效地转录以鸟苷开始的RNA转录物。为了提高不以鸟苷开始的模板中的转录效率, 可使用三核苷酸帽类似物 (m⁷G-ppp-N-N)。三核苷酸帽类似物可相对于使用二核苷酸帽类似物 (m⁷G-ppp-N) 的IVT反应提高转录效率2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20倍或更多倍。

[0225] 转录后也可添加5'帽结构, 例如使用含有mRNA 2'-O-甲基转移酶和S-腺苷甲硫氨酸的牛痘加帽系统 (例如, NEB目录号M2080)。

[0226] RNA然后可使用本领域熟知的技术如苯酚-氯仿提取或柱纯化 (例如, 基于色谱的纯化) 进行纯化。

[0227] 甲病毒生物学

[0228] 甲病毒是披膜病毒科(Togaviridae)的成员,并且是正义单链RNA病毒。成员通常分类为旧世界,例如辛德毕斯、罗斯河、马雅罗、基孔肯雅(Chikungunya)和塞姆利基森林病毒,或新世界,例如东部马脑炎、奥拉、摩根堡或委内瑞拉马脑炎及其衍生菌株TC-83 (Strauss Microbial Review 1994)。天然甲病毒基因组通常约12kb长,其中前三分之二含有编码非结构蛋白(nsP)的基因,所述非结构蛋白形成用于病毒基因组自我复制的RNA复制复合物,并且最后三分之一含有编码用于病毒粒子产生的结构蛋白的亚基因组表达盒(Frolov RNA2001)。

[0229] 甲病毒的模型生命周期涉及数个不同步骤(Strauss Microbial Review 1994, Jose Future Microbiol 2009)。在病毒附着于宿主细胞后,病毒粒子与内饮区室内的膜融合,导致基因组RNA最终释放至细胞溶质中。以正链定向并包含5'甲基鸟苷酸帽和3'聚A尾部的基因组RNA被翻译以产生形成复制复合物的非结构蛋白nsP1-4。在感染早期,正链然后由复合物复制到负链模板中。在当前模型中,复制复合物随着感染进展被进一步加工,使得所得经加工的复合物转换成将负链转录成全长正链基因组RNA以及含有结构基因的26S亚基因组正链RNA。甲病毒的数个保守序列元件(CSE)已鉴定为可能在各种RNA复制步骤中起作用,包括:从负链模板复制正链RNA中的5'UTR的互补序列、从基因组模板复制负链合成中的51-nt CSE、从负链转录亚基因组RNA中的在nsP与26S RNA之间的接合区中的24-nt CSE、和从正链模板的负链合成中的3'19-nt CSE。

[0230] 在各种RNA物种复制后,病毒颗粒然后通常在病毒的天然生命周期中组装。26S RNA被翻译并且所得蛋白质经进一步加工以产生结构蛋白,其包括衣壳蛋白、糖蛋白E1和E2以及两个小多肽E3和6K (Strauss 1994)。发生病毒RNA的衣壳化,衣壳蛋白通常仅特异性针对所包装的基因组RNA,接着病毒粒子组装并在膜表面出芽。

[0231] 甲病毒递送载体

[0232] 甲病毒(包括甲病毒序列、特征和其他元件)可用于产生基于甲病毒的递送载体(也称为甲病毒载体、甲病毒病毒载体、甲病毒疫苗载体、自我复制RNA(srRNA)载体或自扩增mRNA(SAM)载体)。甲病毒先前已被工程化以用作表达载体系统(Pushko 1997, Rheme 2004)。甲病毒提供数种优势,特别是在可能需要异源抗原表达的疫苗环境中。由于在宿主细胞溶质中自我复制的能力,因此甲病毒载体通常能够在细胞内产生高拷贝数的表达盒,从而导致高水平的异源抗原产生。另外,载体通常是瞬时的,从而使得生物安全性得以提高以及减少对载体的免疫耐受性的诱导。与其他标准病毒载体(例如人类腺病毒)相比,公众通常还缺乏对甲病毒载体预先存在的免疫性。基于甲病毒的载体也通常导致对受感染细胞的细胞毒性应答。在一定程度上,细胞毒性在疫苗环境中对于适当刺激对所表达的异源抗原的免疫应答可为重要的。然而,所需细胞毒性的程度可为平衡作用,并且因此已开发数种减毒甲病毒,包括VEEV的TC-83菌株。因此,本文所述的抗原表达载体的实例可利用甲病毒骨架,其允许高水平的抗原表达、刺激对抗原的稳固免疫应答、不刺激对载体本身的免疫应答,并且可以安全方式使用。此外,抗原表达盒可经设计以经由优化载体使用的甲病毒序列(包括但不限于源自VEEV或其减毒衍生物TC-83的序列)而刺激不同水平的免疫应答。

[0233] 已使用甲病毒序列对数种表达载体设计策略进行工程化(Pushko1997)。在一种策略中,甲病毒载体设计包括在结构蛋白基因下游插入26S启动子序列元件的第二拷贝,接着是异源基因(Frolov 1993)。因此,除天然非结构蛋白和结构蛋白之外,还产生表达异源蛋

白的额外亚基因组RNA。在这种系统中,存在用于产生感染性病毒粒子的所有元件,并且因此可能发生在非感染细胞中表达载体的反复轮感染。

[0234] 另一种表达载体设计利用辅助病毒系统(Pushko 1997)。在这种策略中,结构蛋白由异源基因代替。因此,在由仍完整的非结构基因介导病毒RNA的自我复制之后,26S亚基因组RNA提供异源蛋白的表达。传统上,表达结构蛋白的额外载体然后例如通过细胞系的共转染以反式供应,以产生感染性病毒。系统详细描述于USPN 8,093,021中,其出于所有目的通过引用整体并入本文。辅助载体系统提供限制形成感染性颗粒的可能性的益处,因此提高生物安全性。另外,辅助载体系统减小总载体长度,潜在提高复制和表达效率。因此,本文所述的抗原表达载体的实例可利用结构蛋白由抗原盒代替的甲病毒骨架,所得载体降低生物安全问题,同时由于整体表达载体尺寸减小而促进有效表达。

[0235] 经由脂质纳米粒子(LNP)递送

[0236] 在疫苗载体设计中考虑的一个重要方面是针对载体本身的免疫性(Riley 2017)。这可呈对载体本身(例如某些人类腺病毒系统)预先存在的免疫性形式,或呈在疫苗施用后对载体产生免疫性的形式。如果进行相同疫苗的多次施用(例如分开的初免和增强剂量),或者如果使用相同疫苗载体系统递送不同抗原盒,则后者是重要的考虑因素。

[0237] 在甲病毒载体的情况下,标准的递送方法是前面讨论的辅助病毒系统,其以反式提供衣壳、E1和E2蛋白以产生感染性病毒颗粒。然而,重要的是要注意E1和E2蛋白通常是中和抗体的主要靶标(Strauss1994)。因此,如果感染性颗粒被中和抗体靶向,则使用甲病毒载体将目标抗原递送至目标细胞的功效可能会降低。

[0238] 病毒颗粒介导的基因递送的替代方案是使用纳米材料递送表达载体(Riley 2017)。重要的是,纳米材料媒介物可由非免疫原性材料制成并且通常避免引发对递送载体本身的免疫性。这些材料可包括但不限于脂质、无机纳米材料和其他聚合物。脂质可为阳离子、阴离子或中性的。材料可为合成或天然来源的,并且在一些情况下是可生物降解的。脂质可包括脂肪、胆固醇、磷脂、脂质缀合物,包括但不限于聚乙二醇(PEG)缀合物(聚乙二醇化脂质)、蜡、油、甘油酯和脂溶性维生素。

[0239] 脂质纳米粒子(LNP)是一种有吸引力的递送系统,因为脂质的两亲性使得能够形成膜和囊泡状结构(Riley 2017)。一般而言,这些囊泡通过吸收至目标细胞的膜中并且将核酸释放至细胞溶质中来递送表达载体。另外,LNP可被进一步修饰或官能化以有助于靶向特定细胞类型。作为说明性实例,LNP的选择性和靶向递送可通过以下实现:1)将脂质缀合配体(例如,甘露糖)结合到细胞类型特异性受体至LNP中,和/或2)将与靶向抗体相互作用的膜束缚脂蛋白(锚)结合到LNP中。所述锚可以是蛋白质A/G和任何结构形式的抗体,包括scFv、Fab和VHH单域抗体或具有在其N末端或在其C末端编码的外在脂化信号(例如,棕榈酰化、异戊二烯化和肉豆蔻酰化)的纳米抗体。LNP设计中的另一个考虑因素是靶向效率与细胞毒性之间的平衡。脂质组合物通常包括阳离子、中性、阴离子和两性脂质的限定混合物。在一些情况下,包括特定脂质以防止LNP聚集、防止脂质氧化、或提供有助于额外部分附着的功能性化学基团。脂质组合物可影响整体LNP大小和稳定性。在一个实例中,脂质组合物包含二亚油醇甲基-4-二甲基氨基丁酸酯(MC3)和MC3样分子。MC3和MC3样脂质组合物可经配制以包括一种或多种其他脂质,例如PEG或PEG缀合脂质、磷酸胆碱、磷酸乙醇胺、固醇或中性脂质。

[0240] 直接暴露于血清的核酸载体(例如表达载体)可具有数种不期望的结果,包括核酸由血清核酸酶降解或游离核酸对免疫系统的脱靶刺激。因此,包封甲病毒载体可用于避免降解,同时还避免潜在的脱靶效应。在某些实例中,甲病毒载体完全包封在递送媒介物内,例如在LNP的含水内部。甲病毒载体包封在LNP内可通过本领域技术人员熟知的技术来进行,例如在微流液体滴生成装置上进行的微流体混合和液滴生成。此类装置包括但不限于标准T形接头装置或流动聚焦装置。在一个实例中,所需脂质制剂(例如含有MC3或MC3样的组合物)与甲病毒递送载体和其他所需药剂并行提供至液滴生成装置,使得递送载体和所需药剂完全包封在基于MC3或MC3样的LNP内部。在一个实例中,液滴生成装置可控制所产生的LNP的尺寸范围和尺寸分布。例如,LNP的尺寸可在1至1000纳米直径范围内,例如1、10、50、100、500或1000纳米。在液滴生成后,包封表达载体的递送媒介物可经进一步处理或修饰以使其准备用于施用。

[0241] 其他载体

[0242] 本文所述的基于自扩增mRNA(SAM)的组合物可与具有不同(例如,非SAM)载体骨架的其他组合物一起使用。例如,SAM组合物可用作疫苗策略的一部分,所述疫苗策略还使用黑猩猩来源的载体骨架来编码抗原盒。黑猩猩C68腺病毒(本文也称为ChAdV68)的核苷酸序列可用于抗原递送的疫苗组合物中(参见SEQ ID NO:1)。源自C68腺病毒的载体的使用进一步描述于USPN 6,083,716、美国申请公开第US20200197500A1号和国际专利申请公开WO2020/243719中,所述文献各自出于所有目的通过引用整体并入本文。

[0243] 抗原

[0244] 抗原可包括核苷酸或多肽。例如,抗原可为编码多肽序列的RNA序列。可用于疫苗中的抗原可因此包括核苷酸序列或多肽序列。

[0245] 本文公开了包含通过本文公开的方法鉴定的肿瘤特异性突变的分离肽、包含已知肿瘤特异性突变的肽和通过本文公开的方法鉴定的突变多肽或其片段。新抗原肽可描述于其编码序列的上下文中,其中新抗原包括编码相关多肽序列的核苷酸序列(例如DNA或RNA)。

[0246] 本文还公开了这样的肽,其源自已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中具有改变表达的任何多肽,例如已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中异常表达的任何多肽。可例如在COSMIC数据库中发现可获得抗原肽的合适的多肽。COSMIC整理了关于人类癌症体细胞突变的全面信息。肽可含有肿瘤特异性突变。肿瘤抗原(例如,共有肿瘤抗原和肿瘤新抗原)可包括但不限于美国申请第17/058,128号中描述的那些,所述文献出于所有目的通过引用并入本文。

[0247] 本文还公开了衍生自与感染性疾病生物体、受试者中的感染或受试者的受感染细胞相关的任何多肽的肽。抗原可来源于感染性疾病生物体的核苷酸序列或多肽序列。感染性疾病生物体的多肽序列包括但不限于病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和/或寄生虫衍生肽。感染性疾病生物体包括但不限于严重急性呼吸综合征相关冠状病毒(SARS)、严重急性呼吸综合征冠状病毒2(SARS-CoV-2)、埃博拉、HIV、乙型肝炎病毒(HBV)、流感、丙型肝炎病毒(HCV)、人乳头瘤病毒(HPV)、巨细胞病毒(CMV)、基孔肯雅病毒、呼吸道合胞病毒(RSV)、登革热病毒、正粘病毒科病毒和肺结核。

[0248] 本文公开了包含通过本文公开的方法鉴定的感染性疾病生物体特异性抗原或表

位的分离肽、包含已知感染性疾病生物体特异性抗原或表位的肽、以及通过本文公开的方法鉴定的突变多肽或其片段。抗原肽可以在其编码序列的背景下进行描述,其中抗原包括编码相关多肽序列的核苷酸序列(例如,DNA或RNA)。

[0249] 本文所述的载体和相关组合物可用于递送来自任何生物体的抗原,包括它们的毒素或其他副产物,以预防和/或治疗与生物体或其副产物相关的感染或其他不良反应。

[0250] 可并入疫苗(例如,编码在盒中)的抗原包括可用于使人类或非人类动物免疫以抵抗病毒,例如感染人类和非人类脊椎动物的病原病毒的免疫原。抗原可选自多种病毒家族。需要免疫应答的理想病毒家族的实例包括小核糖核酸病毒家族,其包括引起约50%普通感冒病例的鼻病毒属;肠道病毒属,其包括脊髓灰质炎病毒、柯萨奇病毒(coxsackievirus)、埃可病毒(echovirus)和人类肠道病毒如甲型肝炎病毒;以及口蹄疫病毒属,其主要在非人类动物中引起口蹄疫。在小核糖核酸病毒家族中,靶抗原包括VP1、VP2、VP3、VP4和VPG。另一个病毒家族包括杯状病毒家族,其涵盖诺沃克(Norwalk)病毒群,它们是流行性胃肠炎的重要病原体。另一个可用于靶向抗原以在人类和非人类动物中刺激免疫应答的理想病毒家族是披膜病毒家族,其包括甲病毒属,其包括辛德毕斯病毒、罗斯河病毒和委内瑞拉、东方和西方马脑炎,以及风疹病毒属(rubivirus),包括风疹病毒。黄病毒科包括登革热、黄热病、日本脑炎、圣路易斯脑炎和蜱传脑炎病毒。其他靶抗原可能产生自丙型肝炎或冠状病毒家族,其包括多种非人类病毒,例如传染性支气管炎病毒(家禽)、猪传染性胃肠病毒(猪)、猪血凝性脑脊髓炎病毒(猪)、猫传染性腹膜炎病毒(猫)、猫肠道冠状病毒(猫)、犬冠状病毒(狗)和人类呼吸道冠状病毒,其可导致普通感冒和/或非甲、乙或丙型肝炎。在冠状病毒家族中,靶抗原包括E1(也称为M或基质蛋白)、E2(也称为S或刺突蛋白)、E3(也称为HE或血凝素-埃尔特糖(elterose))糖蛋白(并非在所有冠状病毒中都存在)或N(核衣壳)。还有其他抗原可靶向弹状病毒家族,其包括水疱病毒属(例如,水疱性口炎病毒)和一般的狂犬病病毒(例如,狂犬病)。在弹状病毒家族中,合适的抗原可来源于G蛋白或N蛋白。丝状病毒科(其包括出血热病毒如马尔堡病毒和埃博拉病毒)可能是合适的抗原来源。副粘病毒家族包括1型副流感病毒、3型副流感病毒、3型牛副流感病毒、风疹病毒(腮腺炎病毒)、2型副流感病毒、4型副流感病毒、新城疫病毒(鸡)、牛瘟、麻疹病毒,其包括麻疹和犬瘟热,以及肺炎病毒,其包括呼吸道合胞病毒(例如,糖(G)蛋白和融合(F)蛋白,其序列可从GenBank获得)。流感病毒属于正粘病毒科并且可以是合适的抗原来源(例如,HA蛋白、N1蛋白)。布尼亚病毒家族包括布尼亚病毒属(加利福尼亚脑炎、拉克罗斯(La Crosse))、白蛉病毒(裂谷热)、汉坦病毒(puremala是一种血红热病毒)、内罗病毒(内罗毕绵羊病)和各种未指定的布尼亚病毒(bungavirus)。沙粒病毒家族提供了抗LCM和拉沙热病毒的抗原来源。呼肠孤病毒家族包括呼肠孤病毒属、轮状病毒属(其导致儿童急性胃肠炎)、环状病毒和培养病毒(科罗拉多婢热、利庞博(Lebombo)(人类)、马脑病、蓝舌病)。逆转录病毒家族包括涵盖人类和兽医学疾病如猫白血病病毒、HTLVI和HTLVII、慢病毒(其包括人类免疫缺陷病毒(HIV)、猿免疫缺陷病毒(SIV)、猫免疫缺陷病毒(FIV)、马传染性贫血病毒和泡沫病毒(spumavirinal))的致癌病毒亚科。在慢病毒中,已经描述了许多合适的抗原并且可容易地选择。合适的HIV和SIV抗原的实例包括但不限于gag、pol、Vif、Vpx、VPR、Env、Tat、Nef和Rev蛋白,以及它们的各种片段。例如,Env蛋白的合适片段可包括其任何亚基,例如gp120、gp160、gp41,或其更小的片段,例如,长度至少约8个氨基酸。类似地,可选择tat蛋白的片段。[参见美国专利第5,891,

994号和美国专利第6,193,981号。]还参见D.H.Barouch等人, *J.Virol.*, 75(5):2462-2467 (2001年3月)和R.R.Amara等人, *Science*, 292:69-74 (2001年4月6日)中描述的HIV和SIV蛋白。在另一个实例中, HIV和/或SIV免疫原性蛋白或肽可用于形成融合蛋白或其他免疫原性分子。参见例如2001年8月2日公开的WO 01/54719和1999年4月8日公开的WO 99/16884中描述的HIV-1Tat和/或Nef融合蛋白和免疫方案。本发明不限于本文所述的HIV和/或SIV免疫原性蛋白质或肽。此外, 对这些蛋白质的多种修饰已被描述或可由本领域技术人员容易地进行。参见例如美国专利第5,972,596号中描述的经修饰的gag蛋白。此外, 任何所需的HIV和/或SIV免疫原可单独或组合递送。此类组合可包括来自单个载体或来自多个载体的表达。乳多空病毒家族包括多瘤病毒亚科(BKU和JCU病毒)和乳头瘤病毒亚科(与癌症或乳头的恶性进展相关)。腺病毒家族包括引起呼吸道疾病和/或肠炎的病毒(EX、AD7、ARD、O.B.)。细小病毒家族猫细小病毒(猫肠炎)、猫泛白细胞减少病毒、犬细小病毒和猪细小病毒。疱疹病毒家族包括 α 疱疹病毒亚科, 其涵盖单纯病毒属(HSVI、HSVII)、水痘病毒(伪狂犬病、水痘带状疱疹), 和 β 疱疹病毒亚科, 其包括巨细胞病毒属(人CMV、鼠巨细胞病毒), 和 γ 疱疹病毒亚科, 其包括淋巴潜隐病毒属、EBV(伯基特淋巴瘤)、传染性鼻气管炎、马立克氏病(Marek's disease)病毒和细长病毒属(rhadinovirus)。痘病毒家族包括脊索动物痘病毒亚科(chordopoxvirinae), 其涵盖正痘病毒属(天花(Variola/Smallpox)和牛痘(Vaccinia/Cowpox)、副痘病毒、禽痘病毒、山羊痘病毒、兔痘病毒、猪痘病毒和昆虫痘病毒亚科。肝炎病毒病毒包括乙型肝炎病毒。一种可能是合适的抗原来源的未分类病毒是丁型肝炎病毒。其他病毒来源可包括禽传染性法氏囊病病毒以及猪呼吸和生殖综合征病毒。甲病毒家族包括马动脉炎病毒和各种脑炎病毒。

[0251] 可并入疫苗中(例如, 编码在盒中)的抗原还包括可用于免疫人类或非人类动物以抵抗病原体的免疫原, 所述病原体包括感染人类和非人类脊椎动物的细菌、真菌、寄生微生物或多细胞寄生虫。细菌病原体的实例包括致病性革兰氏阳性球菌, 包括肺炎球菌; 葡萄球菌; 和链球菌。致病性革兰氏阴性球菌包括脑膜炎球菌; 淋球菌。致病性肠道革兰氏阴性杆菌包括肠杆菌科(enterobacteriaceae); 假单胞菌属(pseudomonas)、不动杆菌属(acinetobacteria)和艾肯氏菌属(eikenella); 类鼻疽; 沙门氏菌属(salmonella); 志贺氏菌属(shigella); 嗜血杆菌属(haemophilus)(流感嗜血杆菌(Haemophilus influenzae)、睡眠嗜血杆菌(Haemophilus somnus)); 莫拉菌属(moraxella); 杜克雷嗜血杆菌(H. ducreyi)(其导致软下疳); 布鲁氏菌属(brucella); 土拉热弗朗西斯菌(Franisella tularensis)(其导致土拉菌病); 耶尔森氏菌属(yersinia)(巴氏杆菌属(pasteurella)); 串珠状链杆菌(streptobacillus moniliformis)和螺旋菌属(spirillum)。革兰氏阳性杆菌包括单核细胞增生李斯特菌(listeria monocytogenes); 红斑丹毒丝菌(erysipelothrix rhusiopathiae); 白喉棒状杆菌(Corynebacterium diphtheria)(白喉); 霍乱; 炭疽芽胞杆菌(B. anthracis)(炭疽病); 多诺瓦斯病(donovanosis)(腹股沟肉芽肿); 和巴尔通体病(bartonellosis)。由致病性厌氧菌引起的疾病包括破伤风; 肉毒杆菌中毒; 其他梭菌; 肺结核; 麻风; 和其他分枝杆菌。具体细菌种类的实例是但不限于肺炎链球菌(Streptococcus pneumoniae)、化脓性链球菌(Streptococcus pyogenes)、无乳链球菌(Streptococcus agalactiae)、粪链球菌(Streptococcus faecalis)、卡他莫拉菌(Moraxella catarrhalis)、幽门螺杆菌(Helicobacter pylori)、脑膜炎奈瑟菌

(*Neisseria meningitidis*)、淋病奈瑟菌(*Neisseria gonorrhoeae*)、沙眼衣原体(*Chlamydia trachomatis*)、肺炎衣原体(*Chlamydia pneumoniae*)、鹦鹉热衣原体(*Chlamydia psittaci*)、百日咳杆菌(*Bordetella pertussis*)、伤寒杆菌(*Salmonella typhi*)、鼠伤寒沙门氏菌(*Salmonella typhimurium*)、霍乱沙门氏菌(*Salmonella choleraesuis*)、大肠杆菌(*Escherichia coli*)、志贺氏菌(*Shigella*)、霍乱弧菌(*Vibrio cholerae*)、白喉棒杆菌(*Corynebacterium diphtheriae*)、结核分枝杆菌(*Mycobacterium tuberculosis*)、鸟分枝杆菌(*Mycobacterium avium*)、胞内分枝杆菌(*Mycobacterium intracellulare*)复合体、奇异变形杆菌(*Proteus mirabilis*)、普通变形杆菌(*Proteus vulgaris*)、金黄色葡萄球菌(*Staphylococcus aureus*)、破伤风梭菌(*Clostridium tetani*)、问号钩端螺旋体(*Leptospira interrogans*)、伯氏疏螺旋体(*Borrelia burgdorferi*)、溶血巴斯德菌(*Pasteurella haemolytica*)、多杀巴斯德菌(*Pasteurella multocida*)、胸膜肺炎放线杆菌(*Actinobacillus pleuropneumoniae*)和鸡毒支原体(*Mycoplasma gallisepticum*)。致病性螺旋体疾病包括梅毒；密螺旋体病：雅司病(yaws)、品他病(pinta)和地方性梅毒；和钩端螺旋体病。由高等病原体细菌和病原真菌引起的其他感染包括放线菌病；诺卡氏菌病；隐球菌病(隐球菌)、芽生菌病(芽生菌)、组织胞浆菌病(组织胞浆菌)和球孢子菌病(球孢子菌)；念珠菌病(念珠菌)、曲霉病(曲霉)和毛霉菌病；孢子丝菌病；副球孢子菌病、皮氏菌病、圆球菌病、足菌肿和色霉菌病；和皮肤癣菌病。立克次体感染包括斑疹伤寒、落基山斑疹热、Q热和立克次体痘。支原体和衣原体感染的实例包括：肺炎支原体(*Mycoplasma pneumoniae*)；性病性淋巴肉芽肿；鹦鹉热；和围产期衣原体感染。致病性真核生物涵盖致病性原生动物和蠕虫，并且由此产生的感染包括：阿米巴病；疟疾；利什曼病(例如，由硕大利什曼原虫(*Leishmania major*)引起)；锥虫病；弓形虫病(例如，由刚地弓形虫(*Toxoplasma gondii*)引起)；卡氏肺孢子虫(*Pneumocystis carinii*)；Trichans；刚地弓形虫；巴贝虫病；贾第鞭毛虫病(例如，由贾第鞭毛虫(*Giardia*)引起)；旋毛虫病(例如，由滴虫(*Trichomonas*)引起)；丝虫病；血吸虫病(例如，由血吸虫(*Schistosoma*)引起)；线虫；吸虫(trematodes/fluke)；和绦虫(cestode/tapeworm)感染。其他寄生虫感染可能由蛔虫(*Ascaris*)、鞭虫(*Trichuris*)、隐孢子虫(*Cryptosporidium*)和卡氏肺孢子虫等引起。

[0252] 本文还公开了衍生自与感染性疾病生物体、受试者中的感染或受试者的受感染细胞相关的任何多肽的肽。抗原可来源于感染性疾病生物体的核酸序列或多肽序列。感染性疾病生物体的多肽序列包括但不限于病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和/或寄生虫衍生肽。感染性疾病生物体包括但不限于严重急性呼吸综合征相关冠状病毒(SARS)、严重急性呼吸综合征冠状病毒2(SARS-CoV-2)、埃博拉、HIV、乙型肝炎病毒(HBV)、流感、丙型肝炎病毒(HCV)、人乳头瘤病毒(HPV)、巨细胞病毒(CMV)、基孔肯雅病毒、呼吸道合胞病毒(RSV)、登革热病毒、正粘病毒科家族病毒和肺结核。

[0253] 可选择预计呈递在细胞例如肿瘤细胞、受感染细胞或免疫细胞(包括专职抗原呈递细胞如树突状细胞)的细胞表面上的抗原。可选择预计具有免疫原性的抗原。

[0254] 由抗原核苷酸序列编码的一种或多种多肽可包含以下中的至少一者：IC50值小于1000nM的与MHC的结合亲和力，对于MHC I类肽，长度为8-15、8、9、10、11、12、13、14或15个氨基酸，在肽内或肽附近存在促进蛋白酶体切割的序列基序，以及存在促进TAP转运的序列基序。对于长度为6-30、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、

27、28、29或30个氨基酸的MHC II类肽,在肽内或附近存在促进通过胞外或溶酶体蛋白酶(例如组织蛋白酶)切割或HLA-DM催化的HLA结合的序列基序。

[0255] 一种或多种抗原可呈递在肿瘤的表面。一种或多种抗原可呈递在受感染细胞的表面上。

[0256] 一种或多种抗原在患有肿瘤的受试者中可以是免疫原性的,例如,能够在受试者中刺激T细胞应答和/或B细胞应答。一种或多种抗原在患有或怀疑患有感染的受试者中可以是免疫原性的,例如,能够在受试者中刺激T细胞应答和/或B细胞应答。一种或多种抗原在处于感染风险的受试者中可以是免疫原性的,例如,能够在受试者中刺激T细胞应答和/或B细胞应答,从而提供针对感染的免疫保护(即免疫性),例如,刺激记忆T细胞、记忆B细胞和/或感染特异性抗体的产生。

[0257] 一种或多种抗原能够刺激B细胞应答,例如产生识别所述一种或多种抗原的抗体(例如识别感染性疾病抗原的抗体)。抗体可识别线性多肽序列或识别二级和三级结构。因此,B细胞抗原可包括线性多肽序列或具有二级和三级结构的多肽,包括但不限于全长蛋白质、蛋白质亚基、蛋白质结构域或已知或预测具有二级和三级结构的任何多肽序列。能够刺激B细胞对感染的应答的抗原可为在感染性疾病生物体表面发现的抗原。能够引发B细胞对感染的应答的抗原可为在感染性疾病生物体中表达的细胞内抗原。

[0258] 一种或多种抗原可包括能够刺激T细胞应答的抗原(例如,包括预测的T细胞表位序列的肽)和能够刺激B细胞应答的不同抗原(例如,全长蛋白质、蛋白质亚基、蛋白质结构域)的组合。

[0259] 在针对受试者产生疫苗的情况下,可排除在受试者中刺激自身免疫应答的一种或多种抗原。

[0260] 至少一种抗原肽分子(例如,表位序列)的大小可包含但不限于约5、约6、约7、约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30、约31、约32、约33、约34、约35、约36、约37、约38、约39、约40、约41、约42、约43、约44、约45、约46、约47、约48、约49、约50、约60、约70、约80、约90、约100、约110、约120或更多个氨基酸残基,和可源自其中的任何范围。在具体实施方案中,抗原肽分子等于或小于50个氨基酸。

[0261] 抗原肽和多肽可为:对于MHC I类,15个残基或更少的长度并且通常由约8个至约11个残基、特别是9或10个残基组成;对于MHC II类,6-30个残基(包括端值)。

[0262] 如果需要,可以数种方式设计较长的肽。在一种情况下,当HLA等位基因上肽的呈递可能性经预测或已知时,较长的肽可由以下任一者组成:(1)个别呈递的具有朝向每个相应基因产物的N末端和C末端延伸2-5个氨基酸的肽;(2)所呈递的肽中的一些或全部与各自的延伸序列的串接。在另一种情况下,当测序揭示肿瘤中存在长(>10个残基)新表位序列(例如归因于产生新颖肽序列的移码、通读或内含子包含)时,较长的肽将由以下组成:(3)新型肿瘤特异性或感染性疾病特异性氨基酸的整个区段,从而绕过了对最强HLA呈递的较短肽的基于计算或体外测试的选择的需要。在这两种情况下,使用较长的肽允许患者细胞进行内源性加工,并且可引起更有效的抗原呈递和刺激T细胞应答。较长的肽还可包括肽的全长蛋白质、蛋白质亚基、蛋白质结构域及其组合,例如在感染性疾病生物体中表达的那些。可包括较长的肽(例如,全长蛋白质、蛋白质亚基、蛋白质结构域)及其组合以刺激B细胞

应答。

[0263] 抗原肽和多肽可呈递在HLA蛋白上。在一些方面,抗原肽和多肽以比野生型肽更大的亲和力呈递在HLA蛋白上。在一些方面,抗原肽或多肽的IC₅₀可至少小于5000nM、至少小于1000nM、至少小于500nM、至少小于250nM、至少小于200nM、至少小于150nM、至少小于100nM、至少小于50nM或更小。

[0264] 在一些方面,抗原肽和多肽在施用至受试者时不刺激自身免疫应答和/或引起免疫耐受性。

[0265] 还提供了包含至少两种或更多种抗原肽的组合物。在一些实施方案中,所述组合物含有至少两种不同的肽。至少两种不同的肽可衍生自相同的多肽。不同的多肽是指肽因长度、氨基酸序列或两者而异。肿瘤特异性肽可衍生自己知或已发现含有肿瘤特异性突变的任何多肽,或衍生自己知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中具有改变表达的任何多肽的肽,例如已知或已发现与正常细胞或组织相比在肿瘤细胞或癌组织中异常表达的任何多肽。肽可衍生自己知或怀疑与感染性疾病生物体相关的任何多肽,或衍生自己知或已发现与正常细胞或组织相比在受感染细胞中具有改变表达的任何多肽(例如,感染性疾病多核苷酸或多肽,包括在宿主细胞中表达受限的感染性疾病多核苷酸或多肽)的肽。可衍生抗原肽的合适多肽可见于例如COSMIC数据库或AACR基因组学证据瘤形成信息交换(Genomics Evidence Neoplasia Information Exchange, GENIE)数据库。COSMIC整理了关于人类癌症体细胞突变的全面信息。AACR GENIE汇总临床级癌症基因组数据并将其与数万名癌症患者的临床结果联系起来。肽可包括肿瘤特异性突变。在一些方面,所述肿瘤特异性突变是特定癌症类型的驱动突变。

[0266] 具有所需活性或特性的抗原肽和多肽可经修饰以提供某些所需属性,例如改善的药理学特征,同时增加或至少保留基本上所有未经修饰的肽与所需MHC分子结合和激活适当T细胞的生物活性。例如,抗原肽和多肽可进行各种变化,例如保守或非保守取代,其中此类变化可在其使用中提供某些优势,例如改进的MHC结合、稳定性或呈递。保守取代意指用生物和/或化学类似的另一氨基酸残基置换氨基酸残基(例如,一个疏水性残基置换另一个氨基酸残基,或一个极性残基置换另一氨基酸残基)。取代包括以下组合,例如Gly、Ala; Val、Ile、Leu、Met; Asp、Glu; Asn、Gln; Ser、Thr; Lys、Arg; 和Phe、Tyr。单氨基酸取代的效应也可使用D-氨基酸探测。此类修饰可使用熟知的肽合成程序进行,如例如Merrifield, Science 232:341-347 (1986), Barany和Merrifield, The Peptides, Gross和Meienhofer编(N.Y., Academic Press), 第1-284页(1979); 以及Stewart和Young, Solid Phase Peptide Synthesis, (Rockford, Ill., Pierce), 第2版(1984)中所述。

[0267] 肽和多肽用各种氨基酸模拟物或非天然氨基酸修饰可在提高肽和多肽的体内稳定性方面特别有用。稳定性可以多种方式加以测定。例如,肽酶和各种生物介质(例如人类血浆和血清)已用于测试稳定性。参见例如Verhoef等人, Eur. J. Drug Metab Pharmacokin. 11:291-302 (1986)。肽的半衰期可使用25%人类血清(v/v)测定方便地确定。方案一般如下。汇集的人类血清(AB型,非热灭活)在使用之前通过离心去脂。血清然后用RPMI组织培养基稀释至25%并用于测试肽稳定性。在预定时间间隔下,移出少量反应溶液并添加至6%三氯乙酸或乙醇水溶液中。将混浊的反应样品冷却(4°C)15分钟,并且然后旋转集结沉淀的血清蛋白质。然后使用稳定性特定的色谱条件通过反相HPLC来判定肽的存

在。

[0268] 肽和多肽可经修饰以提供除改进的血清半衰期以外的所需属性。例如，肽刺激CTL活性的能力可通过与含有至少一个能够刺激T辅助细胞应答的表位的序列连接来增强。免疫原性肽/T辅助细胞缀合物可通过间隔分子连接。间隔子通常由相对较小的中性分子(例如氨基酸或氨基酸模拟物)构成，其在生理条件下基本上不带电。间隔子通常选自例如Ala、Gly或非极性氨基酸或中性极性氨基酸的其他中性间隔子。应理解，任选存在的间隔子无需由相同残基组成，并且因此可为杂寡聚物或均寡聚物。当存在时，间隔子将通常是至少一个或两个残基，更通常是三至六个残基。或者，肽可在无间隔子的情况下连接到T辅助肽。

[0269] 抗原肽可直接或经由在肽的氨基或羧基端处的间隔子连接到T辅助肽。抗原肽或T辅助肽的氨基端可被酰化。示例性T辅助肽包括破伤风类毒素830-843、流感307-319、疟疾孢子382-398和378-389。

[0270] 蛋白质或肽可通过本领域技术人员已知的任何技术制造，包括经由标准分子生物学技术表达蛋白质、多肽或肽；从天然来源分离蛋白质或肽；或化学合成蛋白质或肽。先前已公开对应于各种基因的核苷酸和蛋白质、多肽和肽序列，并且可见于本领域普通技术人员已知的计算机化数据库中。一种此类数据库是位于美国国家卫生研究院(National Institutes of Health)网站的美国国家生物技术信息中心(National Center for Biotechnology Information)的Genbank和GenPept数据库。已知基因的编码区可使用本文所公开或本领域普通技术人员应知晓的技术扩增和/或表达。或者，蛋白质、多肽和肽的各种市售制剂是本领域技术人员已知的。

[0271] 在另一方面，抗原包括编码抗原肽或其部分的核酸(例如多核苷酸)。多核苷酸可为例如DNA、cDNA、PNA、CNA、RNA(例如mRNA)、单链和/或双链、或原生或稳定化形式的多核苷酸，例如具有硫代磷酸骨架的多核苷酸，或其组合，并且其可含有或可不含内含子。可对编码抗原的多核苷酸序列进行序列优化以改进表达，例如通过改进转录、翻译、转录后加工和/或RNA稳定性。例如，可对编码抗原的多核苷酸序列进行密码子优化。本文中的“密码子优化”是指就给定生物体的密码子偏倚而言，用经常使用的同义密码子替换不经常使用的密码子。可对多核苷酸序列进行优化以改进转录后加工，例如优化以减少意外剪接，例如通过去除剪接基序(例如，典型和/或隐秘/非典型剪接供体、分支和/或受体序列)和/或引入外源剪接基序(例如剪接供体、分支和/或受体序列)以偏向偏好的剪接事件。外源内含子序列包括但不限于源自SV40(例如，SV40微型内含子)和源自免疫球蛋白(例如，人 β -珠蛋白基因)的那些。外源内含子序列可并入在启动子/增强子序列和抗原序列之间。Callendret等人(Virology.2007年7月5日;363(2):288-302)更详细地描述了用于表达载体的外源内含子序列，所述文献出于所有目的通过引用并入本文。可对多核苷酸序列进行优化以改进转录物稳定性，例如通过去除RNA不稳定性基序(例如，富含AU的元件和3'UTR基序)和/或重复核苷酸序列。可对多核苷酸序列进行优化以改进准确转录，例如通过去除隐秘的转录起始子和/或终止子。可对多核苷酸序列进行优化以改进翻译和翻译准确性，例如通过去除隐秘的AUG起始密码子、过早的聚A序列和/或二级结构基序。可对多核苷酸序列进行优化以改进转录物的核输出，例如通过添加组成型转运元件(CTE)、RNA转运元件(RTE)或土拨鼠转录后调控元件(WPRE)。Callendret等人(Virology.2007年7月5日;363(2):288-302)更详细地描述了用于表达载体的核输出信号，所述文献出于所有目的通过引用并入本文。多核苷酸序

列可关于GC含量进行优化,例如以反映给定生物体的平均GC含量。序列优化可平衡一种或多种序列特性,例如转录、翻译、转录后加工和/或RNA稳定性。序列优化可产生平衡了转录、翻译、转录后加工和RNA稳定性中的每一者的最佳序列。序列优化算法是本领域技术人员已知的,例如GeneArt (Thermo Fisher)、Codon Optimization Tool (IDT)、Cool Tool (新加坡大学)、SGI-DNA (La Jolla California)。可分别对抗原编码蛋白的一个或多个区域进行序列优化。另一方面提供了一种能够表达多肽或其部分的表达载体。不同细胞类型的表达载体在本领域中是熟知的并且无需过度实验便可选择。一般而言,DNA以适当定向插入至表达载体(例如质粒)中并以正确阅读框进行表达。如果需要,DNA可连接到由所需宿主识别的适当转录和翻译调节控制核苷酸序列,但此类控制件通常可用于表达载体中。载体然后经由标准技术引入至宿主中。指导可见于例如Sambrook等人(1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 中。

[0272] 疫苗组合物

[0273] 本文还公开了一种免疫原性组合物,例如疫苗组合物,其能够引起特异性免疫应答,例如肿瘤特异性免疫应答或感染性疾病生物体特异性免疫应答。疫苗组合物通常包含一种或多种抗原,例如,使用本文描述的方法选择或选自病原体衍生肽、病毒衍生肽、细菌衍生肽、真菌衍生肽和/或寄生虫衍生肽。疫苗组合物也可以称为疫苗。

[0274] 疫苗可含有1至30种肽;2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30种不同的肽;6、7、8、9、10、11、12、13或14种不同的肽;或12、13或14种不同的肽。肽可包括翻译后修饰。疫苗可含有1至100或更多种核苷酸序列;2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多种不同的核苷酸序列;6、7、8、9、10、11、12、13或14种不同的核苷酸序列;或12、13或14种不同的核苷酸序列。疫苗可含有1至30种抗原序列;2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多种不同的抗原序列;6、7、8、9、10、11、12、13或14种不同的抗原序列;或12、13或14种不同的抗原序列。

[0275] 疫苗可含有1至30种抗原编码核酸序列;2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多种不同的抗原编码核酸序列;6、7、8、9、10、11、12、13或14种不同的抗原编码核酸序列;或12、13或14种不同的抗原编码核酸序列。抗原编码核酸序列可指抗原“盒”的抗原编码部分。本文更详细地描述了抗原盒的特征。抗原编码核酸序列可含有一种或多种表位编码核酸序列(例如,编码串接的T细胞表位的抗原编码核酸序列)。

[0276] 疫苗可含有1至30种不同的表位编码核酸序列;2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多种不同的表位编码核酸序列;6、7、8、9、10、11、12、13或14种不同的表位编码核酸序列;或12、13或14种不同的表位编码核酸序列。表位编码核酸序列可指单个表位序列的序列,例如编码串接的T细胞表位的抗原编码核酸序列中的T细胞表位中的每一者。

[0277] 疫苗可含有表位编码核酸序列的至少两个重复序列。如本文所用,“重复序列”是指抗原编码核酸序列内相同的核酸表位编码核酸序列(包括本文所述的任选的5'接头序列和/或任选的3'接头序列)的两次或更多次迭代。在一个实例中,盒的抗原编码核酸序列部分编码表位编码核酸序列的至少两个重复序列。在进一步的非限制性实例中,盒的抗原编码核酸序列部分编码超过一个不同的表位,并且至少一个不同的表位由编码所述不同表位的核酸序列的至少两个重复序列编码(即,至少两个不同的表位编码核酸序列)。在说明性的非限制性实例中,抗原编码核酸序列编码由表位编码核酸序列表位编码序列A(E_A)、表位编码序列B(E_B)和表位编码序列C(E_C)编码的表位A、B和C,并且具有至少一个不同表位的重复序列的示例性抗原编码核酸序列由但不限于下式说明:

[0278] -一个不同表位的重复序列(表位A的重复序列):

[0279] $E_A-E_B-E_C-E_A$;或

[0280] $E_A-E_A-E_B-E_C$

[0281] -多个不同表位的重复序列(表位A、B和C的重复序列):

[0282] $E_A-E_B-E_C-E_A-E_B-E_C$;或

[0283] $E_A-E_A-E_B-E_B-E_C-E_C$

[0284] -多个不同表位的多次重复序列(表位A、B和C的重复序列):

[0285] $E_A-E_B-E_C-E_A-E_B-E_C-E_A-E_B-E_C$;或

[0286] $E_A-E_A-E_A-E_B-E_B-E_B-E_C-E_C-E_C$

[0287] 上述实例不是限制性的,并且具有至少一个不同表位的重复序列的抗原编码核酸序列可以任何顺序或频率编码每个不同表位。例如,顺序和频率可以是不同表位的随机排列,例如,在具有表位A、B和C的实例中,通过式 $E_A-E_B-E_C-E_C-E_A-E_B-E_A-E_C-E_A-E_C-E_C-E_B$ 。

[0288] 本文还提供了一种抗原编码盒,所述抗原编码盒具有至少一个由下式从5'至3'描述的抗原编码核酸序列:

[0289] $(E_x - (E_n^N)_y)_z$

[0290] 其中E表示包含至少一种不同的表位编码核酸序列中的至少一者的核苷酸序列,

[0291] n表示单独的不同的表位编码核酸序列的数目并且是包括0的任何整数,

[0292] E^N 表示包含对于每个相应n的单独不同的表位编码核酸序列的核苷酸序列,

[0293] 对于z的每次迭代: $x=0$ 或 1 ,对于每个n, $y=0$ 或 1 ,和x或y=1中的至少一者,并且

[0294] $z=2$ 或更大,其中所述抗原编码核酸序列包含E、给定 E^N 或其组合的至少两次迭代。

[0295] 每个E或 E^N 可独立地包含本文所述的任何表位编码核酸序列(例如编码感染性疾

病T细胞表位和/或新抗原表位的肽)。例如,每个E或E^N可独立地包含由式(L5_b-N_c-L3_d)从5'至3'描述的核苷酸序列,其中N包含与每个E或E^N相关的不同表位编码核酸序列,其中c=1,L5包含5'接头序列,其中b=0或1,并且L3包含3'接头序列,其中d=0或1。本文进一步描述了可使用的表位和接头。

[0296] 表位编码核酸序列的重复序列(包括任选的5'接头序列和/或任选的3'接头序列)可彼此直接线性连接(例如,如上所示的E_A-E_A-.....)。表位编码核酸序列的重复序列可被一个或多个额外的核苷酸序列分开。一般而言,表位编码核酸序列的重复序列可被适用于本文所述组合物的任何大小的核苷酸序列分开。在一个实例中,表位编码核酸序列的重复序列可由单独的不同的表位编码核酸序列分开(例如,如上所示的E_A-E_B-E_C-E_A.....)。在重复序列被单个单独的不同表位编码核酸序列分开并且每个表位编码核酸序列(包括任选的5'接头序列和/或任选的3'接头序列)编码长度为25个氨基酸的肽的实例中,重复序列可被75个核苷酸分开,例如在以E_A-E_B-E_A.....表示的抗原编码核酸中,E_A被75个核苷酸分开。在一个说明性实例中,具有编码25聚体抗原Trp1(VTNTMFVTAPDNLGYMYEVQWPGQ)和Trp2(TQPQIANCSVYDFVWLHYYSVRDT)的重复序列的序列VTNTMFVTAPDNLGYMYEVQWPGQTQPQIANCSVYDFVWLHYYSVRDTVTNTMFVTAPDNLGYMYEVQWPGQTQPQIANCSVYDFVWLHYYSVRDT的抗原编码核酸序列,Trp1的重复序列由25聚体Trp2分开并且因此Trp1表位编码核酸序列的重复序列由75个核苷酸的Trp2表位编码核酸序列分开。在重复序列由2、3、4、5、6、7、8或9个单独的不同表位编码核酸序列分开并且每个表位编码核酸序列(包括任选的5'接头序列和/或任选的3'接头序列)编码长度为25个氨基酸的肽的实例中,所述重复序列可分别由150、225、300、375、450、525、600或675个核苷酸分开。

[0297] 在一个实施方案中,选择不同肽和/或多肽或编码其的核苷酸序列,使得肽和/或多肽能够与不同MHC分子(例如不同MHC I类分子和/或不同MHC II类分子)结合。在一些方面,一种疫苗组合物包含能够与最常出现的MHC I类分子和/或不同MHC II类分子缔合的肽和/或多肽的编码序列。因此,疫苗组合物可包含能够与至少2个优选的、至少3个优选的或至少4个优选的MHC I类分子和/或不同MHC II类分子缔合的不同片段。

[0298] 疫苗组合物能够刺激特异性细胞毒性T细胞应答和/或特异性辅助T细胞应答。疫苗组合物能够刺激特异性细胞毒性T细胞应答和特异性辅助T细胞应答。

[0299] 疫苗组合物能够刺激特异性B细胞应答(例如,抗体应答)。

[0300] 疫苗组合物能够刺激特异性细胞毒性T细胞应答、特异性辅助T细胞应答和/或特异性B细胞应答。疫苗组合物能够刺激特异性细胞毒性T细胞应答和特异性B细胞应答。疫苗组合物能够刺激特异性辅助T细胞应答和特异性B细胞应答。疫苗组合物能够刺激特异性细胞毒性T细胞应答、特异性辅助T细胞应答和特异性B细胞应答。

[0301] 疫苗组合物还可包含佐剂和/或载剂。有用的佐剂和载剂的实例在下文中给出。组合物可与载剂缔合,例如蛋白质或抗原呈递细胞,例如能够将肽呈递至T细胞的树突状细胞(DC)。

[0302] 佐剂是混合至疫苗组合物中增加或以其他方式修饰对抗原的免疫应答的任何物质。载剂可为支架结构,例如能够与抗原缔合的多肽或多糖。任选地,佐剂是共价或非共价缀合的。

[0303] 佐剂提高对抗原的免疫应答的能力通常显现为免疫介导性反应的显著或实质性

增加或疾病症状的减少。例如,体液免疫的增强典型地显现为针对抗原所产生的抗体滴度的显著增大,并且T细胞活性的增强典型地显现为细胞增殖或细胞性细胞毒性或细胞因子分泌的增强。佐剂还可改变免疫应答,例如通过将主要体液或Th应答变为主要细胞或Th应答。

[0304] 合适的佐剂包括但不限于1018ISS、明矾、铝盐、Amplivax、AS15、BCG、CP-870,893、CpG7909、CyaA、dSLIM、GM-CSF、IC30、IC31、咪喹莫特(Imiquimod)、ImuFact IMP321、IS Patch、ISS、ISCOMATRIX、JuvImmune、LipoVac、MF59、单磷酸基脂质A、Montanide IMS 1312、Montanide ISA 206、Montanide ISA 50V、Montanide ISA-51、OK-432、OM-174、OM-197-MP-EC、ONTAK、PepTel载体系统、PLG微粒、雷西莫特(resiquimod)、SRL172、病毒颗粒和其他病毒样颗粒、YF-17D、VEGF捕获剂、R848、 β -葡聚糖、Pam3Cys、Aquila的源自皂素的QS21刺激子(Aquila Biotech, Worcester, Mass., USA)、分支杆菌提取物和合成细菌细胞壁模拟物,以及其他专用佐剂,例如Ribi的Detox、Quil或Superfos。例如弗氏(Freund's)不完全或GM-CSF的佐剂是有用的。先前已描述对树突状细胞具有特异性的数种免疫佐剂(例如MF59)及其制备(Dupuis M等人, Cell Immunol. 1998; 186(1): 18-27; Allison A C; Dev Biol Stand. 1998; 92: 3-11)。还可使用细胞因子。数种细胞因子已直接关联于:影响树突状细胞迁移至淋巴组织(例如TNF- α)、加速树突状细胞成熟变为T-淋巴细胞的有效抗原呈递细胞(例如GM-CSF、IL-1和IL-4)(美国专利第5,849,589号,其通过引用整体明确并入本文)以及充当免疫佐剂(例如IL-12)(Gabrilovich D I等人, J Immunother Emphasis Tumor Immunol. 1996(6): 414-418)。

[0305] 还报道CpG免疫刺激性寡核苷酸增强佐剂在疫苗环境中的效应。还可使用其他TLR结合分子,例如结合RNA的TLR 7、TLR 8和/或TLR 9。

[0306] 有用佐剂的其他实例包括但不限于经化学修饰的CpG(例如CpR、Idera)、聚(I:C)(例如聚i:CI2U)、非CpG细菌DNA或RNA以及免疫活性小分子和抗体,例如环磷酰胺、舒尼替尼(sunitinib)、贝伐单抗(bevacizumab)、西乐葆(celebrex)、NCX-4016、西地那非(sildenafil)、他达拉非(tadalafil)、伐地那非(vardenafil)、索拉菲尼(sorafenib)、XL-999、CP-547632、帕佐洋尼(pazopanib)、ZD2171、AZD2171、伊匹单抗、曲美单抗(tremelimumab)和SC58175,其可起治疗作用和/或充当佐剂。佐剂和添加剂的量和浓度可容易由本领域技术人员确定而无需过度实验。额外佐剂包括集落刺激因子,例如粒细胞巨噬细胞集落刺激因子(GM-CSF,沙格司亭(sargramostim))。

[0307] 疫苗组合物可包含超过一种的不同佐剂。此外,治疗性组合物可包含任何佐剂物质,包括以上各者中的任一者或其组合。还预期,疫苗和佐剂可一起或以任何适当的顺序分开施用。

[0308] 载剂(或赋形剂)可独立于佐剂存在。载剂的功能可例如是增加特定突变体的分子量以提高活性或免疫原性、赋予稳定性、增加生物活性或增加血清半衰期。此外,载剂可辅助呈递肽至T细胞。载剂可为本领域技术人员已知的任何合适的载剂,例如蛋白质或抗原呈递细胞。载剂蛋白质可为但不限于匙孔螺血氰蛋白、血清蛋白质(例如转铁蛋白)、牛血清白蛋白、人类血清白蛋白、甲状腺球蛋白或卵白蛋白、免疫球蛋白或激素,例如胰岛素或棕榈酸。对于人类免疫,载剂通常是生理学上可接受的载剂,其为人类可接受的并且安全的。然而,破伤风类毒素和/或白喉类毒素是合适的载剂。或者,载剂可为葡聚糖,例如琼脂糖。

[0309] 细胞毒性T细胞(CTL)识别呈结合至MHC分子的肽形式而非完整外来抗原自身的抗原。MHC分子本身位于抗原呈递细胞的细胞表面上。因此,如果存在肽抗原、MHC分子和APC的三聚体复合物,则可能激活CTL。相应地,如果不仅将肽用于激活CTL,而且如果另外添加具有相应MHC分子的APC,则可加强免疫应答。因此,在一些实施方案中,疫苗组合物另外含有至少一种抗原呈递细胞。

[0310] 抗原还可包括于基于病毒载体的疫苗平台中,例如牛痘、禽痘、自我复制甲病毒、马拉巴病毒(marabavirus)、腺病毒(参见例如Tatsis等人,Adenoviruses, Molecular Therapy (2004) 10,616-629)或慢病毒,包括但不限于第二、第三或杂交第二/第三代慢病毒和任一一代的重组慢病毒,其设计成靶向特定细胞类型或受体(参见例如Hu等人, Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases, Immunol Rev. (2011) 239 (1):45-61; Sakuma等人, Lentiviral vectors: basic to translational, Biochem J. (2012) 443 (3):603-18; Cooper等人, Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter, Nucl. Acids Res. (2015) 43 (1):682-690; Zufferey等人, Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery, J. Virol. (1998) 72 (12):9873-9880)。取决于上述基于病毒载体的疫苗平台的包装能力,这种方法可递送编码一种或多种抗原肽的一种或多种核苷酸序列。序列可侧接非突变序列,可由接头分开或者可在前面有一种或多种靶向亚细胞区室的序列(参见例如Gros等人, Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients, Nat Med. (2016) 22 (4):433-8; Stronen等人, Targeting of cancer neoantigens with donor-derived T cell receptor repertoires, Science. (2016) 352 (6291):1337-41; Lu等人, Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions, Clin Cancer Res. (2014) 20 (13):3401-10)。在引入宿主中后,受感染细胞表达抗原,从而刺激针对一种或多种肽的宿主免疫(例如CTL)应答。可用于免疫方案中的牛痘载体和方法描述于例如美国专利第4,722,848号中。另一种载体是卡介苗(Bacille Calmet te Guerin, BCG)。BCG载体描述于Stover等人(Nature 351:456-460 (1991))中。根据本文描述,可用于抗原的治疗性施用或免疫的各种其他疫苗载体,例如伤寒沙门氏菌载体等,对于本领域技术人员将是显而易见的。

[0311] 抗原盒

[0312] 鉴于本文提供的教导,用于选择一种或多种抗原、“抗原盒”的克隆和构建以及将其插入病毒载体中的方法在本领域技术范围内。“抗原盒”是指经选择的抗原或多种抗原(例如,抗原编码核酸序列)与转录抗原和表达转录产物所必需的其他调控元件的组合。经选择的抗原或多种抗原可指不同的表位序列,例如,盒中的抗原编码核酸序列可编码表位编码核酸序列(或多个表位编码核酸序列),使得表位被转录和表达。一种抗原或多种抗原可以允许转录的方式可操作地连接到调控组件。此类组件包括可驱动一种或多种抗原在用病毒载体转染的细胞中表达的常规调控元件。因此,抗原盒还可含有经选择的启动子,所述启动子与一种或多种抗原连接并且与其他任选的调控元件一起位于重组载体的经选择的病毒序列内。盒可包括一种或多种抗原,例如一种或多种病原体衍生肽、病毒衍生肽、细菌

衍生肽、真菌衍生肽、寄生虫衍生肽和/或肿瘤衍生肽。盒可具有一种或多种抗原编码核酸序列,例如含有多个抗原编码核酸序列的盒,每个抗原编码核酸序列独立地可操作地连接到单独的启动子和/或使用其他多顺反子系统例如2A核糖体跳跃序列元件(例如,E2A、P2A、F2A或T2A序列)或内部核糖体进入位点(IRES)序列元件连接在一起。接头也可具有切割位点,例如TEV或弗林蛋白酶切割位点。具有切割位点的接头可与其他元件组合使用,例如多顺反子系统的那些。在一个非限制性说明性实例中,弗林蛋白酶切割位点可与2A核糖体跳跃序列元件结合使用,使得弗林蛋白酶切割位点被配置为促进翻译后2A序列的去除。在含有超过一个抗原编码核酸序列的盒中,每个抗原编码核酸序列可含有一种或多种表位编码核酸序列(例如,编码串接的T细胞表位的抗原编码核酸序列)。

[0313] 有用的启动子可为组成型启动子或经调控(诱导型)启动子,其将能够控制有待表达的抗原的量。例如,合乎需要的启动子是巨细胞病毒立即早期启动子/增强子的启动子[参见例如Boshart等人,Cell,41:521-530(1985)]。另一种合乎需要的启动子包括劳斯肉瘤(Rous sarcoma)病毒LTR启动子/增强子。另一种启动子/增强子序列是鸡细胞质β-肌动蛋白启动子[T.A.Kost等人,Nucl.Acids Res.,11(23):8287(1983)]。其他适合或合乎需要的启动子可由本领域技术人员选择。

[0314] 抗原盒还可包括对病毒载体序列异源的核酸序列,包括提供转录物的有效聚腺苷酸化(聚(A)、聚-A或pA)的信号的序列和具有功能性剪接供体和受体位点的内含子。本发明的示例性载体中采用的普通聚-A序列是源自乳多泡病毒SV-40的聚-A序列。聚-A序列通常可在基于抗原的序列之后和在病毒载体序列之前插入盒中。普通内含子序列也可源自SV-40,并且被称为SV-40T内含子序列。抗原盒还可含有此类内含子,位于启动子/增强子序列与抗原之间。这些和其他普通载体元件的选择是常规的[参见例如Sambrook等人,“Molecular Cloning.A Laboratory Manual.”,第2版,Cold Spring Harbor Laboratory,New York(1989)和其中引用的参考文献]并且许多此类序列可从商业和工业来源以及Genbank获得。

[0315] 抗原盒可具有一种或多种抗原。例如,给定盒可包括1-10、1-20、1-30、10-20、15-25、15-20、1-2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或更多种抗原。抗原可直接彼此连接。抗原也可利用接头彼此连接。抗原相对于彼此可处于任何方向,包括N至C或C至N。

[0316] 如本文别处所述,抗原盒可定位于病毒载体中的任何经选择缺失的位点中,例如VEEV骨架的缺失结构蛋白或基于ChAd的载体的E1基因区缺失或E3基因区缺失的位点,以及可选择的其他位点。

[0317] 抗原盒可使用下式描述以描述每个元件的有序序列,从5'至3':

[0318] $(P_a - (L5_b - N_c - L3_d)_x)_z - (P2_h - (G5_e - U_f)_y)_w - G3_g$

[0319] 其中P和P2包含启动子核苷酸序列,N包含MHC I类表位编码核酸序列,L5包含5'接头序列,L3包含3'接头序列,G5包含编码氨基酸接头的核酸序列,G3包含编码氨基酸接头的至少一个核酸序列中的一者,U包含MHC II类抗原编码核酸序列,其中对于每个X,相应Nc是表位编码核酸序列,其中对于每个Y,相应Uf是通用MHC II类表位编码核酸序列。通用序列可包含破伤风类毒素和PADRE中的至少一者。通用序列可包含破伤风类毒素肽。通用序列可包含PADRE肽。通用序列可包含破伤风类毒素和PADRE肽。所述组合物和有序序列可进一步

通过选择存在元件的数目来定义,例如其中 $a=0$ 或 1 ,其中 $b=0$ 或 1 ,其中 $c=1$,其中 $d=0$ 或 1 ,其中 $e=0$ 或 1 ,其中 $f=1$,其中 $g=0$ 或 1 ,其中 $h=0$ 或 1 , $X=1$ 至 400 , $Y=0,1,2,3,4$ 或 5 , $Z=1$ 至 400 ,并且 $W=0,1,2,3,4$ 或 5 。

[0320] 在一个实例中,存在的元件包括其中 $a=0, b=1, d=1, e=1, g=1, h=0, X=10, Y=2, Z=1$,和 $W=1$,其描述其中无额外启动子存在(例如,仅存在由载体骨架如RNA甲病毒骨架提供的启动子核苷酸序列),存在10个MHC I类表位,每个N存在5'接头,每个N存在3'接头,存在2个MHC II类表位,存在连接两个MHC II类表位的接头,存在将两个MHC II类表位的5'端连接到最终MHC I类表位的3'接头的接头,并且存在将两个MHC II类表位的3'端连接到载体骨架(例如,RNA甲病毒骨架)的接头。将抗原盒的3'端连接到载体骨架(例如,RNA甲病毒骨架)的实例包括直接连接到由载体骨架提供的3'UTR元件(例如3'19-nt CSE)。将抗原盒的5'端连接到载体骨架(例如,RNA甲病毒骨架)的实例包括直接连接到载体骨架的启动子或5'UTR元件,例如亚基因组启动子序列(例如,26S亚基因组启动子序列)、甲病毒5'UTR、51-nt CSE或24-nt CSE。

[0321] 其他实例包括:其中 $a=1$,描述其中存在除载体骨架(例如,RNA甲病毒骨架)所提供的启动子核苷酸序列以外的启动子;其中 $a=1$ 并且 Z 大于 1 ,其中存在除载体骨架所提供的启动子核苷酸序列以外的多个启动子,其各自驱动1个或更多个不同MHC I类表位编码核酸序列的表达;其中 $h=1$,描述其中存在单独启动子以驱动MHC II类表位编码核酸序列的表达;以及其中 $g=0$,描述MHC II类表位编码核酸序列(如果存在)直接连接到载体骨架(例如,RNA甲病毒骨架)。

[0322] 其他实例包括其中所存在的每个MHC I类表位可具有5'接头、3'接头,两者都不具有,或具有两者。在其中同一抗原盒中存在超过一种MHC I类表位的实例中,一些MHC I类表位可具有5'接头和3'接头两者,而其他MHC I类表位可具有5'接头、3'接头或两者均不具有。在其中同一抗原盒中存在超过一种MHC I类表位的其他实例中,一些MHC I类表位可具有5'接头或3'接头,而其他MHC I类表位可具有5'接头、3'接头或两者均不具有。

[0323] 在其中同一抗原盒中存在超过一种MHC II类表位的实例中,一些MHC II类表位可具有5'接头和3'接头两者,而其他MHC II类表位可具有5'接头、3'接头或两者均不具有。在其中同一抗原盒中存在超过一种MHC II类表位的其他实例中,一些MHC II类表位可具有5'接头或3'接头,而其他MHC II类表位可具有5'接头、3'接头或两者均不具有。

[0324] 其他实例包括其中所存在的每个抗原可具有5'接头、3'接头,两者都不具有,或具有两者。在其中同一抗原盒中存在超过一种抗原的实例中,一些抗原可具有5'接头和3'接头两者,而其他抗原可具有5'接头、3'接头或两者均不具有。在其中同一抗原盒中存在超过一种抗原的其他实例中,一些抗原可具有5'接头或3'接头,而其他抗原可具有5'接头、3'接头或两者均不具有。

[0325] 启动子核苷酸序列P和/或P2可与载体骨架如RNA甲病毒骨架所提供的启动子核苷酸序列相同。例如,由载体骨架提供的启动子序列P_n和P2可各自包含亚基因组启动子序列(例如,26S亚基因组启动子序列)或CMV启动子。启动子核苷酸序列P和/或P2可不同于载体骨架(例如,RNA甲病毒骨架)所提供的启动子核苷酸序列,以及可彼此不同。

[0326] 5'接头L5可为原生序列或非天然序列。非天然序列包括但不限于AAY、RR和DPP。3'接头L3也可为原生序列或非天然序列。另外,L5和L3可两者均为原生序列,两者均为非天然

序列,或者一者可为原生的并且另一者可为非天然的。对于每个X,氨基酸接头的长度可为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。对于每个X,氨基酸接头的长度也可为至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29、或至少30个氨基酸。

[0327] 对于每个Y,氨基酸接头G5的长度可为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。对于每个Y,氨基酸接头的长度也可为至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29、或至少30个氨基酸。

[0328] 氨基酸接头G3的长度可为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100或更多个氨基酸。G3的长度也可为至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29、或至少30个氨基酸。

[0329] 对于每个X,每个N可编码MHC I类表位、MHC II类表位、能够刺激B细胞应答的表位/抗原,或其组合。对于每个X,每个N可编码MHC I类表位、MHC II类表位和能够刺激B细胞应答的表位/抗原的组合。对于每个X,每个N可编码MHC I类表位和MHC II类表位的组合。对于每个X,每个N可编码MHC I类表位和能够刺激B细胞应答的表位/抗原的组合。对于每个X,每个N可编码MHC II类表位和能够刺激B细胞应答的表位/抗原的组合。对于每个X,每个N可编码MHC II类表位。对于每个X,每个N可编码能够刺激B细胞应答的表位/抗原。对于每个X,每个N可编码长度为7-15个氨基酸的MHC I类表位。对于每个X,每个N也可编码长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个氨基酸的MHC I类表位。对于每个X,每个N也可编码长度为至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少25、至少26、至少27、至少28、至少29或至少30个氨基酸的MHC I类表位。

[0330] 编码一种或多种抗原的盒可以是700个核苷酸或更短。编码一种或多种抗原的盒可以是700个核苷酸或更短并编码2个不同的表位编码核酸序列(例如,编码2个不同的感染

性疾病或肿瘤衍生的编码免疫原性多肽的核酸序列)。编码一种或多种抗原的盒可以是700个核苷酸或更短并编码至少2个不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更短并编码3个不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更短并编码至少3个不同的表位编码核酸序列。编码一种或多种抗原的盒可以是700个核苷酸或更短并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0331] 编码一种或多种抗原的盒的长度可在375-700个核苷酸之间。编码一种或多种抗原的盒的长度可在375-700个核苷酸之间并编码2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可在375-700个核苷酸之间并编码至少2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可在375-700个核苷酸之间并编码3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度在375-700个核苷酸之间并编码至少3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可在375-700个核苷酸之间并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0332] 编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短并编码2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短并编码至少2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短并编码3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短并编码至少3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可以是600、500、400、300、200或100个核苷酸或更短并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0333] 编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸。编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸并编码2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸并编码至少2个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸并编码3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸并编码至少3个不同的表位编码核酸序列。编码一种或多种抗原的盒的长度可为375-600、375-500或375-400个核苷酸并且包括1-10、1-5、1、2、3、4、5、6、7、8、9、10或更多种抗原。

[0334] 免疫调节剂

[0335] 本文所述的载体,例如本文所述的C68载体或本文所述的甲病毒载体,可包含编码至少一种抗原的核酸,并且相同或单独的载体可包含编码至少一种免疫调节剂的核酸。免疫调节剂可包括结合于并阻断免疫检查点分子的活性的结合分子(例如,抗体如scFv)。免疫调节剂可包括细胞因子,例如IL-2、IL-7、IL-12(包括IL-12p35、p40、p70和/或p70-融合构建体)、IL-15或IL-21。免疫调节剂可包括修饰的细胞因子(例如,pegIL-2)。载体可包含抗原盒和一种或多种编码免疫调节剂的核酸分子。

[0336] 可靶向用于阻断或抑制的说明性免疫检查点分子包括但不限于CTLA-4、4-1BB(CD137)、4-1BBL(CD137L)、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、TIM3、B7H3、B7H4、VISTA、KIR、2B4(属于CD2分子家族并且在所有NK、 γ δ 和记忆CD8+ ($\alpha\beta$) T细

胞上表达)、CD160(也称为BY55)和CGEN-15049。免疫检查点抑制剂包括抗体或其抗原结合片段或其他结合蛋白,其结合于并阻断或抑制CTLA-4、PDL1、PDL2、PD1、B7-H3、B7-H4、BTLA、HVEM、TIM3、GAL9、LAG3、TIM3、B7H3、B7H4、VISTA、KIR、2B4、CD160和CGEN-15049中的一者或多者的活性。说明性免疫检查点抑制剂包括曲美单抗(CTLA-4阻断抗体)、抗OX40、PD-L1单克隆抗体(抗B7-H1;MEDI4736)、伊匹单抗、MK-3475(PD-1阻断剂)、纳武单抗(Nivolumab)(抗PD1抗体)、CT-011(抗PD1抗体)、BY55单克隆抗体、AMP224(抗PDL1抗体)、BMS-936559(抗PDL1抗体)、MPLDL3280A(抗PDL1抗体)、MSB0010718C(抗PDL1抗体)和Yervoy/伊匹单抗(抗CTLA-4检查点抑制剂)。可使用本领域普通技术将抗体编码序列工程化到例如C68的载体中。一种示例性方法描述于Fang等人,Stable antibody expression at therapeutic levels using the2A peptide.Nat Biotechnol.2005年5月;23(5):584-90.电子版2005年4月17日;出于所有目的通过引用并入本文。

[0337] 有效负载编码SAM组合物

[0338] 本文还公开了具有SAM载体所来源的自我复制病毒的内源5'序列的SAM载体(例如,具有内源5'VEEV核苷酸AU,也称为“AU-SAM”),其例如在盒中编码一种或多种有效负载核酸序列。“盒”是指所选择的多核苷酸(例如,抗原编码核酸序列)和转录所述多核苷酸并且通常在编码序列的情况下表达所转录的产物所需的其他调控元件的组合。本文还公开了能够递送一种或多种有效负载核酸序列的SAM载体递送组合物。有效负载核酸序列可以是希望被递送至目标细胞的任何核酸序列。一般而言,有效负载是与启动子或任何翻译工具(例如IRES、任何2A自切割肽序列如P2A、E2A、F2A和T2A)连接以驱动核酸序列表达的核酸序列。有效负载核酸序列可编码多肽(即,能够被转录和翻译成蛋白质的核酸序列)。通常,编码肽的有效负载核酸序列可编码任何希望在细胞中表达的蛋白质。蛋白质的实例包括但不限于抗原(例如,MHCI类表位、MHC II类表位或能够刺激B细胞应答的表位)、抗体、细胞因子、嵌合抗原受体(CAR)、T细胞受体或基因组编辑系统组件(例如,基因组编辑系统中使用的核酸酶)。基因组编辑系统包括但不限于CRISPR系统、锌指系统、大范围核酸酶系统或TALEN系统。有效负载核酸序列可以是非编码的(即,能够转录但不翻译成蛋白质的核酸序列)。一般而言,非编码有效负载核酸序列可以是希望在细胞中表达的任何非编码多核苷酸。非编码多核苷酸的实例包括但不限于RNA干扰(RNAi)多核苷酸(例如,反义寡核苷酸、shRNA、siRNA、miRNA等)或基因组编辑系统多核苷酸(例如,具有各种/不同长度的指导RNA [gRNA]、单指导RNA [sgRNA]、反式激活CRISPR [tracrRNA]和/或CRISPR RNA [crRNA])。有效负载核酸序列可编码两个或更多个(例如,2、3、4、5或更多个)不同的多肽(例如,两个或更多个连接在一起的不同表位序列)或含有两个或更多个不同的非编码核酸序列(例如,两个或更多个不同的RNAi多核苷酸)。有效负载核酸序列可具有多肽编码核酸序列和非编码核酸序列的组合。

[0339] 抗原鉴定

[0340] 肿瘤和正常外显子组和转录组的NGS分析研究方法已被描述并应用于抗原鉴定领域。^{6,14,15}可考虑某些优化以提高临床环境中抗原鉴定的灵敏度和特异性。这些优化可分为两个领域,与实验室过程相关的领域和与NGS数据分析相关的领域。所描述的研究方法也可应用于鉴定其他环境中的抗原,例如鉴定来自感染性疾病生物体、受试者中的感染或受试者的受感染细胞的抗原的鉴定。优化的实例对于本领域技术人员来说是已知的,例如在美

国专利第10,055,540号、美国申请公开第US20200010849A1号以及国际专利申请公开WO/2018/195357和WO/2018/208856中更详细描述的方法,所述文献各自出于所有目的通过引用整体并入本文。

[0341] 用于鉴定抗原(例如,源自肿瘤或感染性疾病生物体的抗原)的方法包括鉴定可能呈递在细胞表面上(例如,由MHC呈递在肿瘤细胞、受感染细胞或免疫细胞(包括专职抗原呈递细胞如树突状细胞)上)和/或可能具免疫原性的抗原。举例来说,一种这样的方法可包括以下步骤:从肿瘤、受感染细胞或感染性疾病生物体中获得外显子组、转录组或全基因组核苷酸测序和/或表达数据中的至少一者,其中所述核苷酸测序数据和/或表达数据用于获得代表一组抗原(例如,源自肿瘤或感染性疾病生物体的抗原)中的每一者的肽序列的数据;将每种抗原的肽序列输入到一个或多个呈递模型中,以产生所述抗原中的每一者由一种或多种MHC等位基因在受试者的细胞表面、例如肿瘤细胞或受感染细胞上呈递的数值可能性集合,所述数值可能性集合已至少基于所接收的质谱数据进行鉴定;以及基于所述数值可能性集合选择所述抗原集合的子集以产生经选择抗原集合。

[0342] 躯干肽,意指由所有或大多数亚克隆呈递的肽,可优先排序以纳入疫苗中。任选地,如果没有预测呈递且以高概率具有免疫原性的躯干肽,或者如果预测呈递且以高概率具有免疫原性的躯干肽数目足够小,使得可在疫苗中包括额外的非躯干肽,则可通过估计亚克隆的数目和身份并选择肽来优先排序更多的肽,以使疫苗覆盖的亚克隆数目最大化。

[0343] 在应用了所有上述抗原过滤器之后,可能仍有比疫苗技术可支持的更多候选抗原可用于疫苗包涵。另外,可保留关于抗原分析的各个方面的不确定性,并且候选疫苗抗原的不同特性之间可存在折衷。因此,可考虑整合式多维模型代替选择过程的每个步骤中的预定过滤器,将候选抗原置于具有至少以下轴的空间中并使用整合方法优化选择。

[0344] 1.自身免疫或耐受性的风险(生殖系的风险)(自身免疫的风险较低通常是优选的)

[0345] 2.测序伪影的概率(伪影的概率较低通常是优选的)

[0346] 3.免疫原性的概率(免疫原性的概率较高通常是优选的)

[0347] 4.呈递的概率(呈递的概率较高通常是优选的)

[0348] 5.基因表达(较高表达通常是优选的)

[0349] 6.HLA基因的覆盖率(参与抗原集合呈递的HLA分子数目较大可降低肿瘤、感染性疾病和/或受感染细胞经由HLA分子的下调或突变逃避免疫攻击的概率)

[0350] 7.HLA类别的覆盖率(覆盖HLA-I和HLA-II两者可增加治疗应答的概率并降低肿瘤或感染性疾病逃避的概率)

[0351] 另外,任选地,如果预测抗原将由在患者的全部或部分肿瘤或受感染细胞中丢失或失活的HLA等位基因呈递,则可从疫苗接种去除所述抗原的优先排序(例如排除所述抗原)。HLA等位基因损失可通过体细胞突变、杂合性缺失或基因座的同型接合缺失发生。用于检测HLA等位基因体细胞突变的方法是本领域中所熟知,例如(Shukla等人,2015)。用于检测体细胞LOH和同型接合缺失(包括HLA基因座)的方法同样被充分描述。(Carter等人,2012;McGranahan等人,2017;Van Loo等人,2010)。如果质谱数据表明所预测的抗原未被预测的HLA等位基因呈递,则还可去除抗原的优先排序。

[0352] 治疗和制造方法

[0353] 还提供了一种通过向受试者施用一种或多种抗原(例如使用本文公开的方法鉴定的多种抗原)而在受试者中刺激肿瘤特异性免疫应答、针对肿瘤进行疫苗接种、治疗和/或减轻受试者中的癌症症状的方法。

[0354] 还提供了一种通过向受试者施用一种或多种抗原(例如使用本文公开的方法鉴定的多种抗原)而在受试者中刺激感染性疾病生物体特异性免疫应答、针对感染性疾病生物体进行疫苗接种、治疗和/或减轻与感染性疾病生物体相关的感染症状的方法。

[0355] 在一些方面,受试者已诊断患有癌症或具有患癌症的风险。受试者可为人、狗、猫、马或需要肿瘤特异性免疫应答的任何动物。肿瘤可为任何实体肿瘤,例如乳房肿瘤、卵巢肿瘤、前列腺肿瘤、肺肿瘤、肾脏肿瘤、胃肿瘤、结肠肿瘤、睾丸肿瘤、头颈部肿瘤、胰腺肿瘤、脑肿瘤、黑色素瘤和其他组织器官肿瘤,以及血液肿瘤,例如淋巴瘤和白血病,包括急性髓细胞白血病、慢性髓细胞白血病、慢性淋巴细胞白血病、T细胞淋巴细胞白血病和B细胞淋巴瘤。

[0356] 在一些方面,受试者已被诊断患有感染或处于感染风险中,例如年龄、地理/旅行和/或工作相关的感染风险或易感性增加,或处于季节性和/或新型疾病感染风险中。

[0357] 可以足以刺激CTL应答的量施用抗原。可以足以刺激T细胞应答的量施用抗原。可以足以刺激B细胞应答的量施用抗原。

[0358] 抗原可以单独施用或与其他治疗剂组合施用。治疗剂可包括靶向感染性疾病生物体的那些,例如抗病毒剂或抗生素剂。

[0359] 此外,可进一步向受试者施用抗免疫抑制/免疫刺激剂,例如检查点抑制剂。例如,可进一步向受试者施用抗CTLA抗体或抗PD-1或抗PD-L1。抗体对CTLA-4或PD-L1的阻断可增强患者对癌细胞的免疫应答。特别是,在遵循疫苗接种方案时,CTLA-4阻断已被证明是有效的。

[0360] 可确定疫苗组合物中每种抗原的最佳量和最佳给药方案。例如,抗原或其变体可制备用于静脉内(i.v.)注射、皮下(s.c.)注射、皮内(i.d.)注射、腹膜内(i.p.)注射、肌肉内(i.m.)注射。注射方法包括s.c.、i.d.、i.p.、i.m.和i.v.。DNA或RNA注射方法包括i.d.、i.m.、s.c.、i.p.和i.v.。其他施用疫苗组合物的方法是本领域技术人员已知的。

[0361] 可对疫苗进行编译,使得组合物中存在的抗原的选择、数目和/或量具组织、癌症、感染性疾病和/或患者特异性。例如,肽的精确选择可由给定组织中亲本蛋白的表达模式指导,或者由患者的突变或疾病状态指导。选择可取决于特定的癌症类型、特定的感染性疾病(例如,受试者感染或有感染风险的特定感染性疾病分离株/菌株)、疾病状态、疫苗接种的目标(例如,预防性或针对正在进行的疾病)、早期治疗方案、患者的免疫状态,当然还有患者的HLA-单倍型。此外,根据特定患者的个人需要,疫苗可含有个性化组分。实施例包括根据特定患者中抗原的表达来改变抗原的选择或在第一轮或治疗方案之后调整二次治疗。

[0362] 通过使用各种诊断方法,例如下文进一步描述的患者选择方法,可鉴定患者是否施用抗原疫苗。患者选择可涉及鉴定一种或多种基因的突变或表达模式。患者选择可涉及鉴定正在进行的感染的感染性疾病。患者选择可涉及鉴定感染性疾病感染的风险。在一些情况下,患者选择涉及鉴定患者的单倍型。可并行执行各种患者选择方法,例如,测序诊断可鉴定患者的突变和单倍型。各种患者选择方法可按顺序进行,例如,一种诊断测试鉴定突变并且另一种诊断测试鉴定患者的单倍型,并且其中每个测试可以相同(例如,高通量测

序)或不同(例如,一种高通量测序和另一种Sanger测序)诊断方法。

[0363] 对于用作癌症或感染性疾病疫苗的组合物,在本文所述的组合物中可避免或以低量存在在正常组织中大量表达的具有类似正常自身肽的抗原。另一方面,如果已知患者的肿瘤或受感染细胞表达大量的某种抗原,则用于治疗这种癌症或感染的相应药物组合物可以大量存在和/或可包括多于一种对这种特定抗原或这种抗原途径具特异性的抗原。

[0364] 可将包含抗原的组合物施用于已经患有癌症或感染的个体。在治疗性应用中,将组合物以足以刺激对肿瘤抗原或感染性疾病生物体抗原的有效CTL应答并治愈或至少部分阻遏症状和/或并发症的量施用于患者。足以实现这一点的量被定义为“治疗有效剂量”。对该用途有效的量将取决于例如组合物、施用方式、所治疗疾病的阶段和严重程度、患者的体重和一般健康状况以及处方医生的判断。应记住,组合物通常可用于严重疾病状态,即危及生命或可能危及生命的情况,尤其是当癌症已转移或感染性疾病生物体已诱发器官损伤和/或其他免疫病理学时。在这些情况下,鉴于外来物质的最小化和抗原的相对无毒性质,主治医师可以并且认为需要施用实质过量的这些组合物。

[0365] 对于治疗用途,可在检测或手术切除肿瘤时或者在检测或治疗感染时开始施用。这之后可为加强剂量直到至少症状基本上减轻并且持续此后一段时间,或认为提供了免疫性(例如,记忆B细胞或T细胞群,或产生抗原特异性B细胞或抗体)。

[0366] 用于治疗性治疗的药物组合物(例如疫苗组合物)旨在用于肠胃外、局部、经鼻、口服或局部施用。药物组合物可经肠胃外施用,例如静脉内、皮下、皮内或肌肉内施用。组合物可施用在手术切除部位以刺激针对肿瘤的局部免疫应答。可施用组合物以靶向受试者的特定感染组织和/或细胞。本文公开了用于肠胃外施用的组合物,其包含抗原溶液,并且将疫苗组合物溶解或悬浮在可接受的载剂例如水性载剂中。可使用多种水性载剂,例如水、缓冲水、0.9%盐水、0.3%甘氨酸、透明质酸等。这些组合物可通过常规的熟知灭菌技术灭菌,或者可经无菌过滤。所得水溶液可以原样包装使用,或冻干,冻干制剂在施用之前与无菌溶液组合。组合物可根据需要含有药学上可接受的辅助物质以接近生理条件,例如pH调节剂和缓冲剂、张力调节剂、润湿剂等,例如乙酸钠、乳酸钠、氯化钠、氯化钾、氯化钙、脱水山梨糖醇单月桂酸酯、三乙醇胺油酸酯等。

[0367] 抗原也可经由脂质体施用,脂质体将它们靶向特定的细胞组织,例如淋巴组织。脂质体也可用于增加半衰期。脂质体包括乳液、发泡体、胶束、不溶性单层、液晶、磷脂分散体、层状层等。在这些制剂中,待递送的抗原作为脂质体的一部分被并入,单独或联合与结合于例如淋巴样细胞中普遍存在的受体的分子,例如结合于CD45抗原的单克隆抗体,或联合其他治疗性或免疫原性组合物。因此,可将填充有所需抗原的脂质体导向淋巴样细胞的部位,然后所述脂质体在该部位递送选定的治疗性/免疫原性组合物。脂质体可由标准的囊泡形成脂质形成,其通常包括中性和带负电荷的磷脂和固醇,例如胆固醇。脂质的选择通常通过考虑例如脂质体大小、酸不稳定性和脂质体在血流中的稳定性来指导。多种方法可用于制备脂质体,如例如Szoka等人,Ann.Rev.Biophys.Bioeng.9;467(1980),美国专利第4,235,871号、第4,501,728号、第4,501,728号、第4,837,028号和第5,019,369号中所述。

[0368] 为了靶向免疫细胞,要并入脂质体中的配体可包括例如对所需免疫系统细胞的细胞表面决定簇特异的抗体或其片段。脂质体悬浮液可以静脉内、局部、表面等施用,其剂量尤其根据施用方式、所递送的肽和所治疗疾病的阶段而变化。

[0369] 出于治疗或免疫目的,编码肽和任选地一种或多种本文所述肽的核酸也可施用于患者。许多方法可方便地用于将核酸递送至患者。例如,核酸可作为“裸DNA”直接递送。例如在Wolff等人,Science 247:1465-1468(1990)以及美国专利第5,580,859号和第5,589,466号中描述了这种方法。核酸也可使用弹道递送来施用,如例如美国专利第5,204,253号中所述。可施用仅由DNA组成的粒子。或者,可将DNA粘附到粒子如金粒子上。用于递送核酸序列的方法可包括带有或不带有电穿孔的病毒载体、mRNA载体和DNA载体。

[0370] 还可将核酸与阳离子化合物例如阳离子脂质复合递送。脂质介导的基因递送方法描述于例如9618372WOAWO 96/18372;9324640WOAWO 93/24640;Mannino和Gould-Fogerite,BioTechniques6(7):682-691(1988);美国专利第5,279,833号Rose美国专利第5,279,833号;9106309WOAWO 91/06309;和Felgner等人,Proc.Natl.Acad.Sci.USA 84:7413-7414(1987)。

[0371] 抗原还可包括在基于病毒载体的疫苗平台中,例如牛痘、禽痘、自我复制甲病毒、马拉巴病毒、腺病毒(参见例如Tatsis等人,Adenoviruses,Molecular Therapy(2004)10,616-629)或慢病毒,包括但不限于第二、第三或杂交第二/第三代慢病毒和任一一代的重组慢病毒,其设计成靶向特定细胞类型或受体(参见例如Hu等人,Immunization Delivered by Lentiviral Vectors for Cancer and Infectious Diseases,Immunol Rev.(2011)239(1):45-61;Sakuma等人,Lentiviral vectors:basic to translational,Biochem J.(2012)443(3):603-18;Cooper等人,Rescue of splicing-mediated intron loss maximizes expression in lentiviral vectors containing the human ubiquitin C promoter,Nucl.Acids Res.(2015)43(1):682-690;Zufferey等人,Self-Inactivating Lentivirus Vector for Safe and Efficient In Vivo Gene Delivery,J.Virol.(1998)72(12):9873-9880)。取决于上述基于病毒载体的疫苗平台的包装能力,这种方法可递送编码一种或多种抗原肽的一种或多种核苷酸序列。序列可侧接非突变序列,可由接头分开或者可在前面有一种或多种靶向亚细胞区室的序列(参见例如Gros等人,Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients,Nat Med.(2016)22(4):433-8;Stronen等人,Targeting of cancer neoantigens with donor-derived T cell receptor repertoires,Science.(2016)352(6291):1337-41;Lu等人,Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions,Clin Cancer Res.(2014)20(13):3401-10)。在引入宿主中后,受感染细胞表达抗原,从而刺激针对一种或多种肽的宿主免疫(例如CTL)应答。可用于免疫方案中的牛痘载体和方法描述于例如美国专利第4,722,848号中。另一种载体是卡介苗(Bacille Calmet te Guerin,BCG)。BCG载体描述于Stover等人(Nature 351:456-460(1991))中。根据本文描述,可用于抗原的治疗性施用或免疫的各种其他疫苗载体,例如伤寒沙门氏菌载体等,对于本领域技术人员将是显而易见的。

[0372] 一种施用核酸的方法使用编码一个或多个表位的微型基因构建体。为了创建编码所选CTL表位(微型基因)以在人类细胞中表达的DNA序列,将表位的氨基酸序列进行反向翻译。人类密码子使用表用于指导每个氨基酸的密码子选择。这些表位编码DNA序列直接相连,形成一个连续的多肽序列。为了优化表达和/或免疫原性,可在微型基因设计中并入额

外的元件。可反向翻译并包括在微型基因序列中的氨基酸序列的实例包括：辅助T淋巴细胞、表位、前导（信号）序列和内质网保留信号。此外，可通过包括与CTL表位相邻的合成（例如聚丙氨酸）或天然存在的侧翼序列来改进CTL表位的MHC呈递。通过组装编码微型基因正链和负链的寡核苷酸，将微型基因序列转化为DNA。使用熟知技术在适当的条件下合成、磷酸化、纯化和退火重叠的寡核苷酸（30-100个碱基长）。使用T4 DNA连接酶连接寡核苷酸的末端。然后可将这种编码CTL表位多肽的合成微型基因克隆到所需的表达载体中。

[0373] 纯化的质粒DNA可使用多种配方制备用于注射。其中最简单的是在无菌磷酸盐缓冲液（PBS）中重建冻干DNA。已经描述了多种方法，并且可使用新技术。如上所述，核酸方便地与阳离子脂质一起配制。此外，糖脂、融合脂质体、肽和统称为保护性、相互作用、非缩合（PINC）的化合物也可与纯化的质粒DNA复合，以影响例如稳定性、肌肉内分散或运输到特定器官或细胞类型的变量。

[0374] 还公开了一种制造疫苗的方法，所述方法包括执行本文公开的方法的步骤；以及生产包含多种抗原或多种抗原子集的疫苗。

[0375] 本文公开的抗原可使用本领域已知的方法制造。例如，本文公开的产生抗原或载体（例如，包括至少一种编码一种或多种抗原的序列的载体）的方法可包括在适合表达抗原或载体的条件下培养宿主细胞，其中所述宿主细胞包含至少一种编码抗原或载体的多核苷酸，并纯化所述抗原或载体。标准纯化方法包括色谱技术、电泳、免疫学、沉淀、透析、过滤、浓缩和色谱聚焦技术。

[0376] 宿主细胞可包括中国仓鼠卵巢（CHO）细胞、NS0细胞、酵母或HEK293细胞。宿主细胞可用一种或多种多核苷酸转化，所述多核苷酸包含至少一种编码本文公开的抗原或载体的核酸序列，任选地其中经分离的多核苷酸还包含可操作地连接到至少一种编码所述抗原或载体的核酸序列的启动子序列。在某些实施方案中，经分离的多核苷酸可为cDNA。

[0377] 抗原使用和施用

[0378] 疫苗接种方案可用于向受试者给与一种或多种抗原。可使用初免疫苗和加强疫苗对受试者进行给药。

[0379] 初免疫苗可基于本文所述的SAM疫苗组合物，其中SAM具有SAM载体所来源的自我复制病毒的内源5'序列（例如，内源5'VEEV核苷酸AU，也称为“AU-SAM”）。

[0380] 加强疫苗（包括两次或更多次加强施用）可基于本文所述的SAM疫苗组合物，其中SAM具有SAM载体所来源的自我复制病毒的内源5'序列（例如，内源5'VEEV核苷酸AU，也称为“AU-SAM”）。

[0381] 疫苗接种方案可包括初免疫苗和加强疫苗两者，各自基于本文所述的SAM疫苗组合物，其中SAM具有SAM载体所来源的自我复制病毒的内源5'序列（例如，内源5'VEEV核苷酸AU，也称为“AU-SAM”）。

[0382] 初免疫苗（包括与具有内源5'序列的SAM组合使用）也可基于C68（例如，SEQ ID NO:1或2所示的序列）或SAM（例如，SEQ ID NO:3或4所示的序列）。加强疫苗（包括与具有内源5'序列的SAM组合使用）也可基于C68（例如，SEQ ID NO:1或2所示的序列）或SAM（例如，SEQ ID NO:3或4所示的序列）。

[0383] 初免/加强策略中的每个载体通常包括包含抗原的盒。盒可包括约1-50种抗原，由间隔子例如通常围绕每种抗原的天然序列或其他非天然间隔子序列如AAY分开。盒还可包

括MHCII抗原,例如破伤风类毒素抗原和PADRE抗原,它们可被认为是通用II类抗原。盒还可包括靶向序列,例如泛素靶向序列。此外,每个疫苗剂量可与免疫调节剂一起(例如,同时、之前或之后)施用于受试者。每个疫苗剂量可与检查点抑制剂(CPI)一起(例如,同时、之前或之后)施用于受试者。CPI可包括抑制CTLA4、PD1和/或PDL1的那些,例如抗体或其抗原结合部分。此类抗体可包括曲美单抗或德瓦鲁单抗(durvalumab)。每个疫苗剂量可与细胞因子如IL-2、IL-7、IL-12(包括IL-12p35、p40、p70和/或p70融合构建体)、IL-15或IL-21一起(例如,同时、之前或之后)施用于受试者。每个疫苗剂量可与修饰的细胞因子(例如pegIL-2)一起(例如,同时、之前或之后)施用于受试者。

[0384] 初免疫苗可注射(例如肌肉内)于受试者。可使用每剂量双侧注射。例如,可使用一次或多次ChAdV68(C68)注射(例如总剂量 1×10^{12} 个病毒颗粒);可使用选自0.001至1ug RNA、特别是0.1或1ug范围的低疫苗剂量的一次或多次SAM载体注射;或者可使用选自1至100ug RNA、特别是10或100ug范围的高疫苗剂量的一次或多次SAM载体注射。

[0385] 可在初免疫苗接种之后注射(例如肌肉内)疫苗加强剂(加强疫苗)。加强疫苗可在初免后约每1、2、3、4、5、6、7、8、9或10周,例如每4周和/或8周施用。可使用每剂量双侧注射。例如,可使用一次或多次ChAdV68(C68)注射(例如总剂量 1×10^{12} 个病毒颗粒);可使用选自0.001至1ug RNA、特别是0.1或1ug范围的低疫苗剂量的一次或多次SAM载体注射;或者可使用选自1至100ug RNA、特别是10或100ug范围的高疫苗剂量的一次或多次SAM载体注射。

[0386] 也可向受试者施用抗CTLA-4(例如,曲美单抗)。例如,抗CTLA4可在肌肉内疫苗注射(ChAdV68初免或SAM低剂量)部位附近皮下施用,以确保引流到同一淋巴结。曲美单抗是CTLA-4的选择性人类IgG2 mAb抑制剂。靶向抗CTLA-4(曲美单抗)皮下剂量通常为70-75mg(特别是75mg),剂量范围例如1-100mg或5-420mg。

[0387] 在某些情况下,可使用抗PD-L1抗体,例如德瓦鲁单抗(MEDI 4736)。德瓦鲁单抗是一种选择性、高亲和力的人类IgG1 mAb,其阻断PD-L1与PD-1和CD80的结合。德瓦鲁单抗通常每4周以20mg/kg i.v.施用。

[0388] 免疫监测可在疫苗施用之前、期间和/或之后进行。这种监测可告知安全性和功效以及其他参数。

[0389] 为了进行免疫监测,通常使用PBMC。PBMC可在初免疫苗接种之前和在初免疫苗接种之后(例如4周和8周)分离。PBMC可在即将加强疫苗接种之前和在每次加强疫苗接种之后(例如4周和8周)收集。

[0390] 免疫应答,例如T细胞应答和B细胞应答,可作为免疫监测方案的一部分进行评估。例如,可监测和/或评估本文所述的疫苗组合物刺激免疫应答的能力。如本文所用,“刺激免疫应答”是指免疫应答的任何增加,例如引发免疫应答(例如,在未接受过治疗的受试者中刺激引发免疫应答的初免疫苗)或增强免疫应答(例如,在对抗原具有预先存在的免疫应答,例如由初免疫苗引发的预先存在的免疫应答的受试者中刺激免疫应答增强的加强疫苗)。可使用本领域中已知的一种或多种方法测量T细胞应答,例如ELISpot、细胞内细胞因子染色、细胞因子分泌和细胞表面捕捉、T细胞增殖、MHC多聚体染色或通过细胞毒性测定。针对疫苗中编码的表位的T细胞应答可通过使用ELISpot测定来测量细胞因子(例如IFN- γ)的诱导而从PBMC监测。针对疫苗中编码的表位的特异性CD4或CD8 T细胞应答可通过使用流式细胞术测量胞内或胞外捕捉的细胞因子(例如IFN- γ)的诱导而从PBMC监测。针对疫

苗中编码的表位的特异性CD4或CD8 T细胞应答可通过使用MHC多聚体染色测量表达特异性针对表位/MHC I类复合物的T细胞受体的T细胞群体而从PBMC监测。针对疫苗中编码的表位的特异性CD4或CD8 T细胞应答可通过在3H-胸苷、溴脱氧尿苷和羧基荧光素-二乙酸酯-琥珀酰亚胺酯(CFSE)并入后测量T细胞群的离体扩增而从PBMC监测。特异性针对疫苗中编码的表位的源自PBMC的T细胞的抗原识别能力和溶解活性可通过铬释放测定或替代性比色细胞毒性测定来功能性评估。

[0391] B细胞应答可使用本领域已知的一种或多种方法测量,例如用于确定B细胞分化(例如,分化为浆细胞)、B细胞或浆细胞增殖、B细胞或浆细胞活化(例如,共刺激标志物如CD80或CD86的上调)、抗体类别转换和/或抗体产生(例如,ELISA)的测定法。也可评估抗体的功能,例如评估中和能力。

[0392] 范例

[0393] 为了可以更充分地理解本文所述的本公开,提出以下实施例。本申请中描述的合成和生物学实施例被提供用于说明本文提供的化合物、药物组合和方法,而不应以任何方式解释为限制它们的范围。

[0394] 材料和方法

[0395] 本文提供的化合物可使用以下一般方法和程序从容易获得的起始材料制备。应理解,在给出典型或优选的工艺条件(即,反应温度、时间、反应物的摩尔比、溶剂、压力等)的情况下,除非另有说明,否则也可使用其他工艺条件。最佳反应条件可随所用的特定反应物或溶剂而变化,但此类条件可由本领域技术人员通过常规优化来确定。

[0396] 此外,如对本领域技术人员显而易见的,常规保护基团可能是必要的以防止某些官能团发生不希望的反应。针对特定官能团选择合适的保护基团以及保护和脱保护的合适条件是本领域所熟知的。例如,T. W. Greene和P. G. M. Wuts, *Protecting Groups in Organic Synthesis*, 第二版, Wiley, New York, 1991和其中引用的参考文献中描述了许多保护基团及其引入和去除。

[0397] 本文提供的化合物可通过已知的标准程序分离和纯化。此类程序包括(但不限于)湿磨、柱色谱、HPLC或超临界流体色谱法(SFC)。以下方案详细呈现了本文列出的代表性氧甾醇的制备。本文提供的化合物可由有机合成领域的技术人员从已知或可商购获得的起始材料和试剂制备。可用于本文提供的对映异构体/非对映异构体的分离/纯化的示例性手性柱包括但不限于CHIRALPAK® AD-10、CHIRALCEL® 0B、CHIRALCEL® 0B-H、CHIRALCEL® 0D、CHIRALCEL® 0D-H、CHIRALCEL® 0F、CHIRALCEL® 0G、CHIRALCEL® 0J和CHIRALCEL® 0K。

[0398] 缩写:

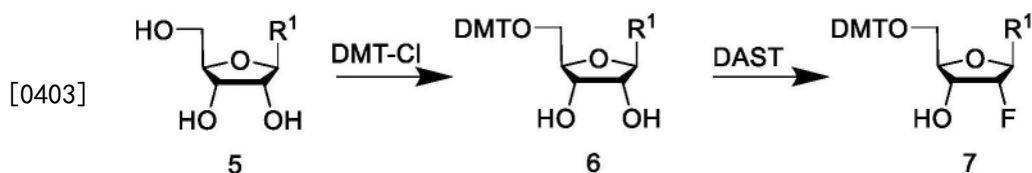
[0399] PE:石油醚;EtOAc:乙酸乙酯;THF:四氢呋喃;PCC:氯铬酸吡啶鎓;TLC:薄层色谱法;PCC:氯铬酸吡啶鎓;t-BuOK:叔丁醇钾;9-BBN:9-硼杂双环[3.3.1]壬烷;Pd(t-Bu₃P)₂:双(三叔丁基膦)钯(0);AcCl:乙酰氯;i-PrMgCl:异丙基氯化镁;TBSCl:叔丁基(氯)二甲基硅烷;(i-PrO)₄Ti:四异丙醇钛;BHT:2,6-二叔丁基-4-甲基苯酚;Me:甲基;i-Pr:异丙基;t-Bu:叔丁基;Ph:苯基;Et:乙基;Bz:苯甲酰基;BzCl:苯甲酰氯;CsF:氟化铯;DAST:二乙氨基三氟化硫;DCC:二环己基碳化二亚胺;DCM:二氯甲烷;DMAP:4-二甲氨基吡啶;DMP:戴斯-马丁过碘烷;EtMgBr:乙基溴化镁;EtOAc:乙酸乙酯;TEA:三乙胺;AlaOH:丙氨酸;Boc:叔丁氧

羰基;Py:吡啶;TBAF:四正丁基氟化铵;THF:四氢呋喃;TBS:叔丁基二甲基甲硅烷基;TMS:三甲基甲硅烷基;TMSCF₃: (三氟甲基) 三甲基硅烷;Ts:对甲苯磺酰基;Bu:丁基;Ti(OiPr)₄:四异丙氧基钛;LAH:氢化铝锂;LDA:二异丙基氨基锂;LiOH·H₂O:氢氧化锂水合物;MAD:甲基铝双(2,6-二叔丁基-4-甲基苯醚);MeCN:乙腈;NBS:N-溴代琥珀酰亚胺;Na₂SO₄:硫酸钠;Na₂S₂O₃:硫代硫酸钠;PE:石油醚;MeCN:乙腈;MeOH:甲醇;Boc:叔丁氧羰基;DMT:4,4'-二甲氧基三苯甲基;MTBE:甲基叔丁基醚;K-selectride:三(仲丁基)硼氢化钾。

[0400] 实施例1.2'-氟核苷酸7的合成

[0401] 本领域普通技术人员将理解,2'-氟核苷酸6可经由一般方案I中概述的合成步骤等来制备。

[0402] 一般合成方案I

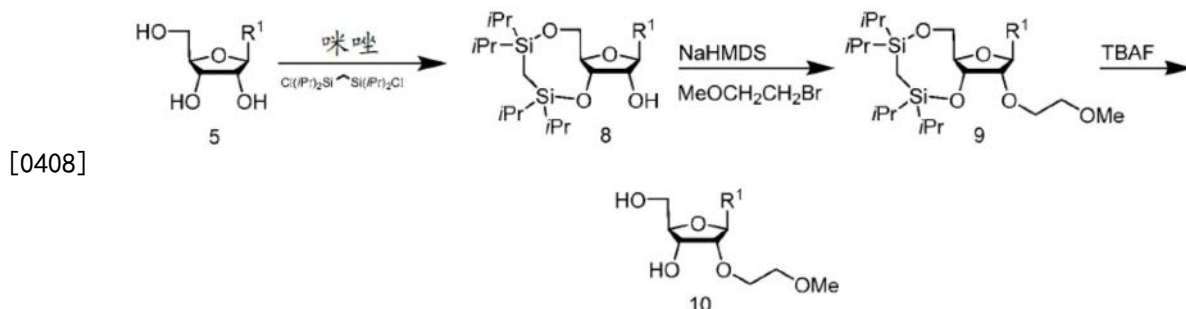


[0404] 可使用DMT-Cl对核苷酸5的伯醇进行选择性的DMT保护,以提供4,4'-二甲氧基三苯甲基保护的核苷酸6。随后暴露于DAST可产生2'-氟核苷酸7。

[0405] 实施例2.2'-甲氧基乙基-核苷酸10的合成

[0406] 本领域普通技术人员将理解,2'-甲氧基乙基-核苷酸10可经由一般方案II中概述的合成步骤等来制备。

[0407] 一般合成方案II

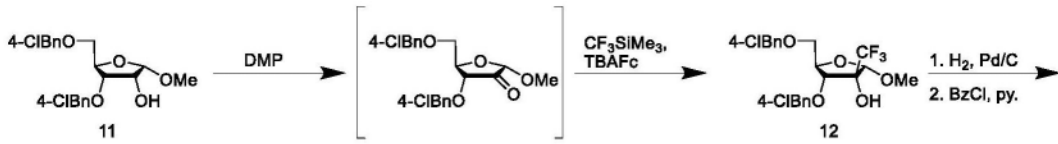


[0409] 核苷酸5可与咪唑和1,1双(二异丙基)氯甲硅烷基)甲烷反应形成脱保护的核苷酸8。8暴露于NaHMDS和MeOCH₂CH₂Br将产生受保护的2'-甲氧基乙基-核苷酸8。使用TBAF对核苷酸9进行脱保护可得到2'-甲氧基乙基-核苷酸10。

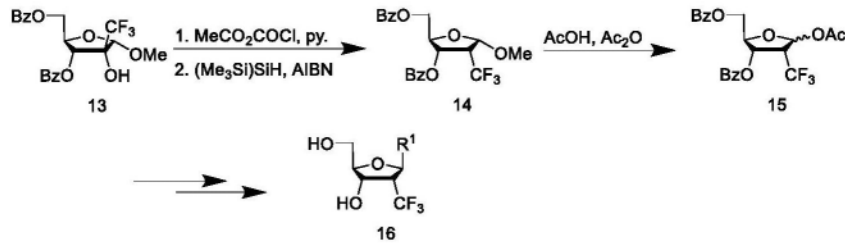
[0410] 实施例3.2'-三氟甲基-核苷酸16的合成

[0411] 本领域普通技术人员将理解2'-三氟甲基核苷酸16可经由一般方案III中概述的合成步骤等来制备。例如,可使用Jeannot,F.等人,“Synthesis and antiviral evaluation of 2'-deoxy-2'-C-trifluoromethyl-β-D-ribonucleoside analogues bearing the five naturally occurring nucleic bases”Org.Biomol.Chem.,2003,1,2096-2102中概述的步骤来复制该核苷酸的合成。

[0412] 一般合成方案III



[0413]

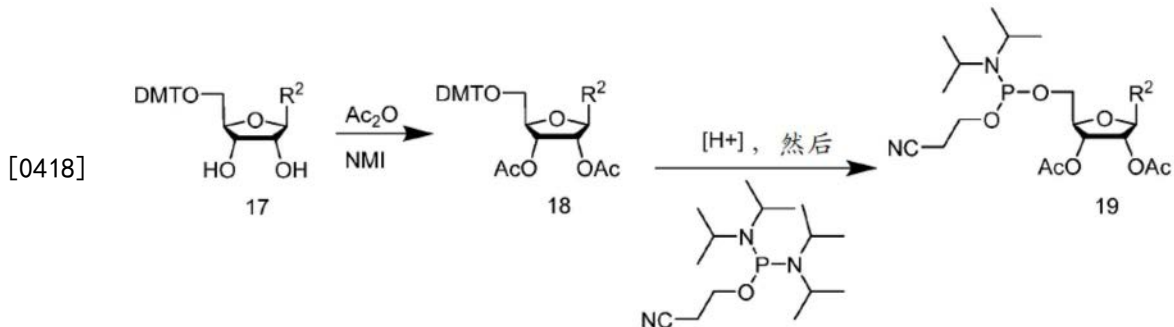


[0414] 使用DMP氧化4-Cl-苄基保护的核苷酸11并随后用 CF_3SiMe_3 处理可提供3-三氟甲基核苷酸12。还原性脱保护,接着使用BzCl重新保护可产生苯甲酰基保护的核苷酸13。自由基介导的脱氧作用可产生苯甲酰基保护的脱氧核苷酸14。甲氧基部分的置换可通过暴露于乙酸和乙酸酐来完成。可使用Jeannot等人描述的条件将1'-乙酸核苷酸15转化为各种核苷酸类似物(16)。

[0415] 实施例4.2',3'二乙酸核苷酸19的合成

[0416] 本领域普通技术人员将理解,2',3'二乙酸核苷酸19可经由一般方案III中概述的合成步骤等来制备。

[0417] 一般合成方案IV

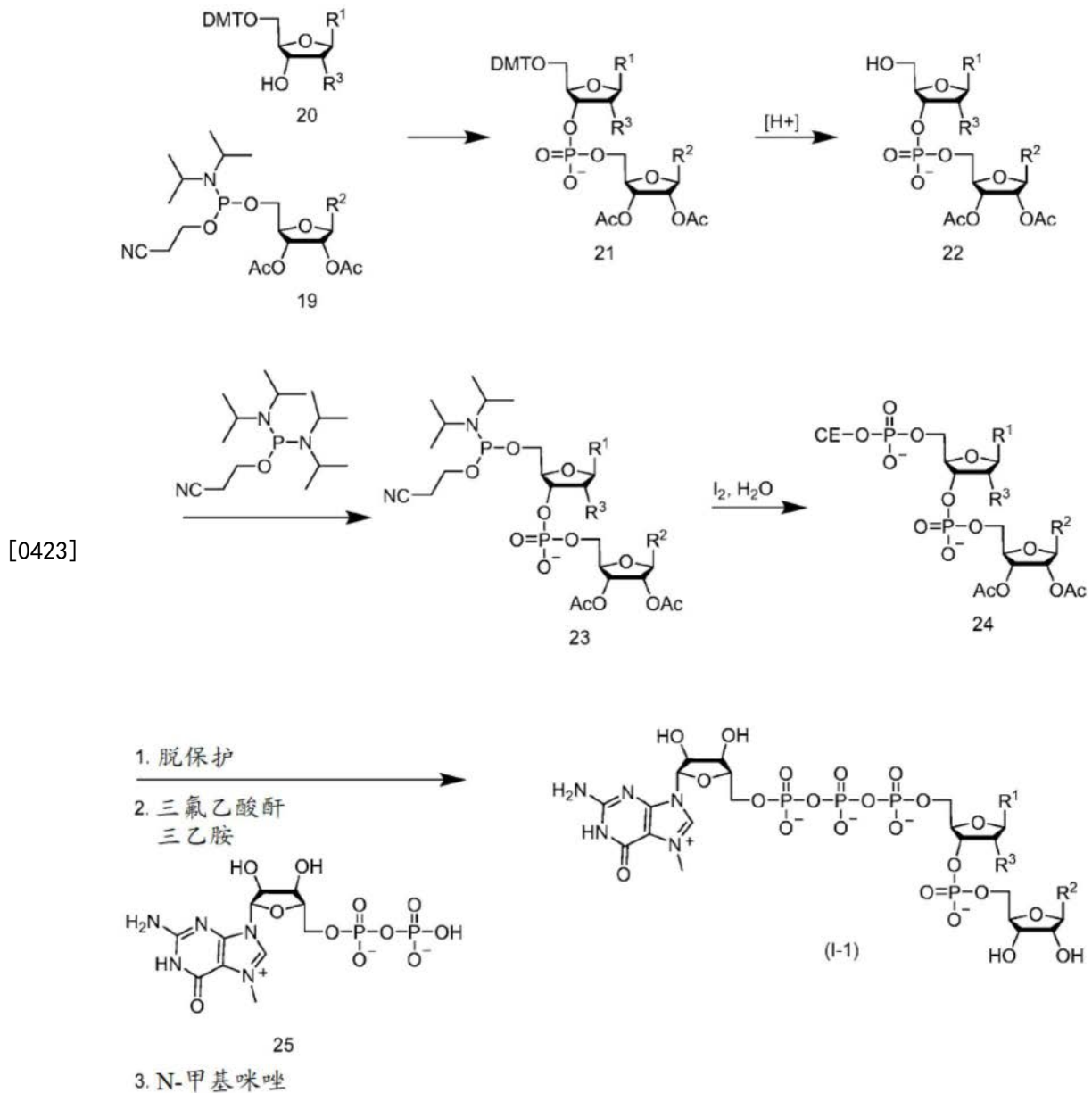


[0419] 用乙酸酐(Ac_2O)和N-甲基咪唑(NMI)处理DMT保护的核苷酸17(参见实施例1)可产生二乙酸酯18。用酸将二乙酸酯18脱保护,随后与N,N,N',N'-四异丙基亚磷酰二胺2-氰乙酯反应可提供2',3'二乙酸核苷酸19。

[0420] 实施例5.式(I-1)化合物的合成

[0421] 本领域普通技术人员将理解,式(I-1)化合物可经由一般方案V中概述的合成步骤等来制备。

[0422] 一般合成方案V



[0424] 具体而言,化合物亚磷酰胺19可在合适的条件下与受保护的核苷酸20反应,得到二核苷酸21。4,4'-二甲氧基三苯甲基二核苷酸21的脱保护可通过暴露于质子酸来实现,产生二核苷酸22。用N,N,N',N'-四异丙基亚磷酰二胺2-氰乙酯处理羟基二核苷酸22可得到亚磷酰二胺2-氰乙酯23。在合适的条件(例如, I₂、H₂O)下氧化亚磷酰二胺2-氰基乙酯23可得到磷酸2-氰基乙酯24。磷酸2-氰基乙酯24可在合适的条件下脱保护。所得二核苷酸可与^{m7}G二磷酸酯25偶联以完成式(I-1)化合物的合成。

[0425] ^{m7}G二磷酸酯25可使用本领域已知的方法制备。例如,参见Kore,A.R.等人,“An Industrial Process for Selective Synthesis of 7-methyl Guanosine 5'-Diphosphate:Versatile Synthons of Synthesis of mRNA Cap Analogues”Nucleosides, Nucleotides,and Nucleic Acids 25:337-340,2006,DOI:10.1080/15257770500544552。

[0426] 实施例6.自扩增表达系统

[0427] A.自我复制RNA病毒骨架和SAM生成

[0428] 在本发明的一种实现方式中,用于抗原表达系统的RNA甲病毒骨架由自我复制的

委内瑞拉马脑炎病毒 (“VEEV”; Kinney, 1986, *Virology* 152:400-413) 通过缺失位于26S亚基因组启动子3'的VEEV的结构蛋白(除了E1的最后50个氨基酸)而生成(VEEV序列7544至11,175缺失;编号基于Kinney等人,1986;SEQ ID NO:6)。为了生成自扩增mRNA (“SAM”)载体,将缺失序列替换为抗原序列。含有20种模型抗原的代表性SAM载体是“VEE-MAG25聚体”(SEQ ID NO:4)。将缺少典型3'二核苷酸GG的修饰T7 RNA聚合酶启动子(TAATACGACTCACTATA)添加到SAM载体的5'端以生成体外转录模板DNA(SEQ ID NO:57;7544至11,175缺失,无插入的抗原盒)。产生额外的模板生产载体,添加PCR正向引物序列和3'限制性位点(SEQ ID NO:58;7544至11,175缺失,无插入的抗原盒)。

[0429] 使用上述模板产生的RNA含有与内源5'VEEV核苷酸序列直接连接的m⁷G帽,即,在m⁷G帽与内源5'VEEV核苷酸序列之间不存在额外的中间核苷酸,例如当使用典型的T7 RNA聚合酶时通常存在的二核苷酸GG。具有以内源核苷酸AUG开始并使用典型的或修饰的(“最小的”)T7启动子的骨架的SAM载体的RNA生产在图1中说明。不具有位于m⁷G帽与内源5'AU核苷酸之间的额外中间核苷酸的SAM载体在本文中称为“AU-SAM”载体。代表性的AU-SAM载体的示意图示于图2中。

[0430] 使用以下步骤共转录产生含有编码代表性抗原的盒的加帽AU-SAM RNA:

[0431] -通过将目标抗原盒克隆到体外转录模板DNA(SEQ ID NO:57)中产生DNA模板

[0432] -加帽RNA是通过体外转录(IVT)产生的,概述如下:

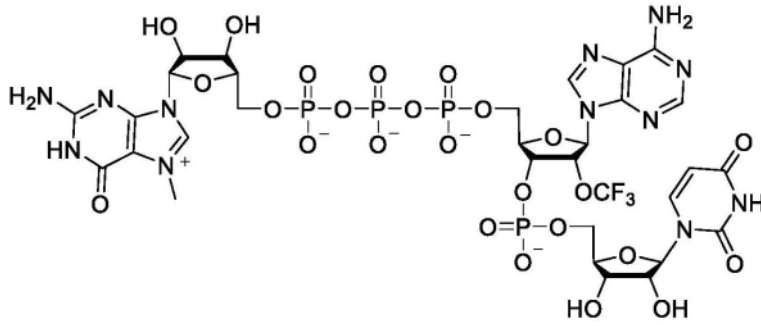
[0433] ○反应含有:使用最终浓度为1x T7 RNA聚合酶混合物(E2040S)的1x转录缓冲液(40mM Tris、10mM二硫苏糖醇、2mM亚精胺、0.002% Triton X-100和27mM氯化镁);0.025mg/mL DNA转录模板(通过限制性消化线性化或PCR扩增);8mM三核苷酸m⁷G-ppp-A-U帽类似物(CleanCap试剂AU;目录号N-7114)和各10mM的ATP、三磷酸胞苷(CTP)、GTP和三磷酸尿苷(UTP)[CleanCap试剂AU取代二核苷酸m⁷G-ppp-A帽类似物(NEB),如下所示]

[0434] ○IVT反应条件:转录反应在37°C孵育2小时并在37°C用DNA酶I缓冲液中的最终2U DNA酶I(AM2239)/0.001mg DNA转录模板处理1小时

[0435] ○加帽的AU-SAM通过RNeasy Maxi(QIAGEN,75162)或液相色谱法纯化

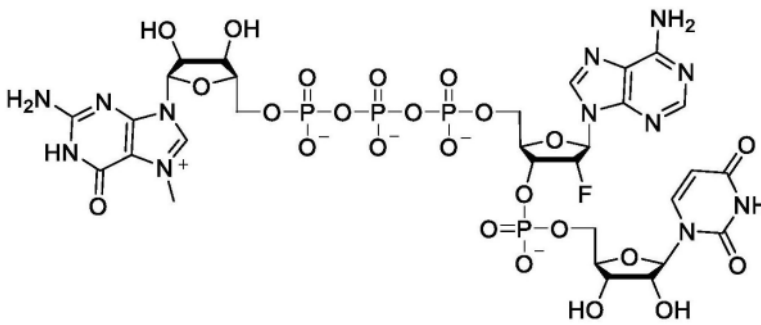
[0436] 将模型抗原盒(“MAG25聚体”;核苷酸SEQ ID NO:34和肽SEQ ID NO:35)插入到VEEV骨架的缺失区域中。如上所述,使用三核苷酸m⁷G-ppp-A-U帽类似物或二核苷酸m⁷G-ppp-A帽类似物产生加帽的AU-SAM RNA。如图3所示,含有三核苷酸m⁷G-ppp-A-U帽的反应产生的RNA可比二核苷酸m⁷G-ppp-A帽类似物多20倍以上。

[0437] 加帽的AU-SAM RNA也在IVT反应中使用本文(诸如下文)所述的三核苷酸m⁷G-ppp-A-U帽类似物,以大于使用二核苷酸m⁷G-ppp-A帽类似物的量产生。

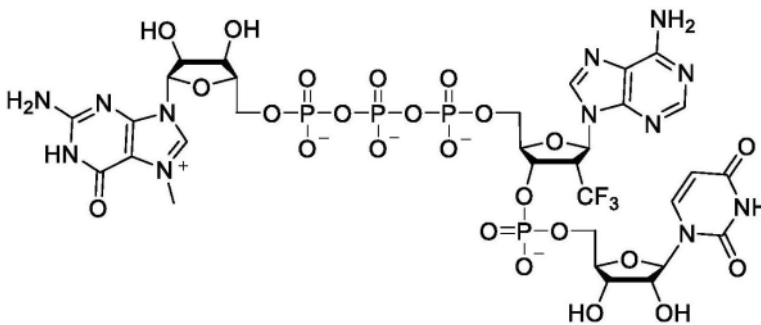


1、

[0438]

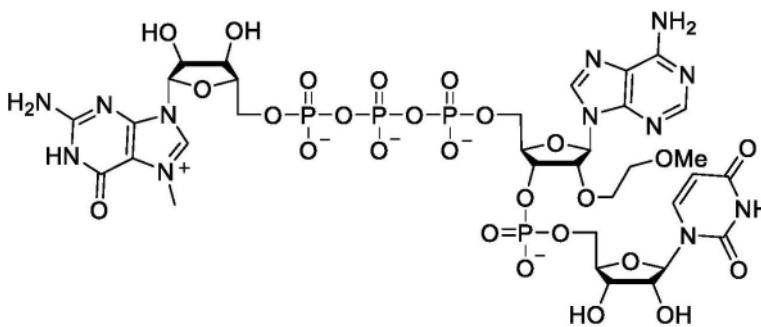


2、



3、和

[0439]



4。

[0440] B. 小鼠中自扩增mRNA病毒载体评价

[0441] 免疫

[0442] 用10ug SAM-LNP使Ba1b/c小鼠(每组n=8)免疫。SAM是AU-SAM(如上所述产生),或使用含有典型T7启动子(SEQ ID NO:8)的DNA模板产生的GG-SAM,其中所产生的RNA的特征在于在m⁷G帽与内源5'VEEV核苷酸序列之间的GG二核苷酸。

[0443] 离体细胞内细胞因子染色(ICS)和流式细胞术分析

[0444] 对于研究中的每只小鼠,通过测量细胞因子诸如IFN- γ 的诱导,在脾细胞中监测T细胞对疫苗中编码的AH1-A5抗原I类表位(PSYAYHQF)的应答。将密度为2-5x10⁶个细胞/毫升的新鲜分离的淋巴细胞与10uM所示肽一起孵育2小时。两小时后,加入布雷菲尔德菌素A(brefeldin A)至5ug/ml的浓度,并将细胞与刺激剂一起再孵育4小时。刺激后,用可固定的活力染料eFluor780,根据制造商的方案标记活细胞,并用以1:400稀释的抗CD8 APC(克隆53-6.7, BioLegend)染色。对于细胞内染色,以1:100使用抗IFN γ PE(克隆XMG1.2, BioLegend)。将样品收集在Cytotflex LX(Beckman Coulter)上。使用FlowJo标绘流式细胞术数据并分析。为了评估抗原特异性应答的程度,计算应答于每个肽刺激剂的CD8+细胞的IFN γ +百分比。

[0445] 小鼠中的免疫原性结果

[0446] 本研究旨在评价和比较使用SAM载体的小鼠中的免疫,所述SAM载体含有直接连接到内源5'VEEV核苷酸序列的m⁷G帽(AU-SAM)或在m⁷G帽与内源5'VEEV核苷酸序列之间的GG二核苷酸(GG-SAM)。插入自扩增骨架中的MAG25聚体模型抗原盒的特征是AH1-A5抗原I类表位SPSYAYHQF作为模型非自身抗原。

[0447] 如上所述,对小鼠进行免疫,并在初始免疫后第12天收集脾细胞并评估抗原特异性免疫应答。如图4和表1所示,相对于GG-SAM,用AU-SAM进行疫苗接种产生的IFN γ +CD8细胞百分比增加约2倍,表明用AU-SAM进行疫苗接种导致抗原特异性免疫应答相对于在RNA的5'末端具有非内源性核苷酸的SAM载体有所增加。

[0448] 表1-小鼠中使用AU-SAM或GG-SAM的IFN γ +CD8细胞

	GG-SAM	AU-SAM
	20.2	32.3
	17.3	22.8
	12.9	31.5
	11.3	18.1
	18.9	35.1
	16.1	31.4
	14.0	22.1
	10.6	30.9
中位数	15.1	31.1
平均值	15.2	28.0
SD	3.5	6.1

[0450] C. 非人类灵长类动物中的自扩增mRNA病毒载体同源初免/增强免疫原性评价

[0451] 免疫

[0452] 用含有MAG25聚体抗原盒(如上所述由IVT共转录产生)并配制在LNP中的AU-SAM递

送组合物使Mamu-A*01印度恒河猴(N=5)免疫。在免疫当天,将SAM-LNP在室温下解冻,用PBS稀释至所需浓度,并使用蠕动泵(Masterflex)和滤筒(Sartorius Sartopore 2过滤囊尺寸4,150cm²,0.2μm)过滤。动物没有接受过任何先前免疫调节抗体治疗或针对SIV的疫苗接种,并且先前没有暴露于SIV。SAM作为双侧肌肉内注射施用到股四头肌中。在初免疫苗接种后第4周、第8周和第20周肌肉内施用AU-SAM的同源加强。每只动物的所有4次剂量总计为1mg。对于前3剂(第0周、第4周、第8周),施用2mL SAM(每条腿1mL)。对于第4剂(第20周),注射体积减少至1mL(每条腿0.5mL)。

[0453] 免疫监测

[0454] 对于免疫监测,将10-20mL血液收集到含有肝素的真空采血管中,并保持在室温下直至分离。使用淋巴细胞分离培养基(LSM)和Leucosep分离管通过密度梯度离心分离PBMC。PBMC用碘化丙锭染色,并使用Cytotflex LX(Beckman Coulter)对活细胞计数。然后在RPMI完全(10% FBS)中以4x 10⁶个细胞/毫升重悬样品。

[0455] IFN γ ELISPOT测定是使用预包装的96孔板(MAbtech,猴IFN γ ELISPOT PLUS,ALP(试剂盒批号36,板批号19))按照制造商的方案进行的。对于每个样品和刺激物,每孔2.5x10⁴和1x10⁵个PBMC一式三份地与10ug/mL肽刺激物(GenScript)一起涂铺,并在完全RPMI中孵育过夜。盒中不包含的人HBV S-抗原肽(WLSLLVPPFV, Genscript)用作每个样品的阴性对照。用PBS洗涤板并与抗猴IFN γ MAb生物素(MAbtech)一起孵育两小时,接着再次洗涤并与链霉亲和素-ALP(MAbtech)一起孵育一小时。最后一次洗涤后,将板与BCIP/NBT(MAbtech)一起孵育十分钟以形成免疫斑点并在37°C干燥过夜。使用AID读取器(Autoimmun Diagnostika)对斑点进行成像和计数。

[0456] 重复孔变异性(变异性=方差/[中位数+1])大于10并且中位数大于10的样品被排除在外。斑点值基于孔饱和度根据下式进行调整:调整斑点=原始斑点+2*(原始斑点*饱和度/[100-饱和度])。孔饱和度大于33%的孔被认为是“数量众多而难以计数”(TNTC)并被排除在外。通过减去阴性对照肽孔的平均值对每个样品进行背景校正。通过将校正的斑点数乘以1x10⁶/涂铺细胞数,将数据相对于每1x10⁶个PBMC的斑点形成集落(SFC)进行归一化。对于总体总结分析,利用通过以1x10⁵个细胞/孔涂铺细胞而产生的计算值,除非当样品是TNTC时,在这种情况下,由以2.5x10⁴个细胞涂铺细胞产生的值用于该特定样品/刺激物/时间点。使用R编程语言执行数据处理。

[0457] 恒河猴中的免疫原性结果

[0458] 本研究旨在评价基于SAM(特别是AU-SAM)的同源初免/加强免疫策略在恒河猴中的免疫原性和初步安全性,恒河猴是一种高度预测人类疫苗效力的模型。对于AU-SAM研究组,如上所述,对恒河猴进行免疫,并在免疫前和初始免疫后的第1、2、3、4、5、6、8、9、10和14周收集PBMC用于免疫监测(AU-SAM研究组的详细信息在图5上图中说明)。MAG25聚体模型抗原盒的特征是六个Mamu-A*01限制性I类限制性病毒抗原作为模型非自身抗原(图5下图中所示的模型抗原)。

[0459] 评估了六种Mamu-A*01抗原中的每一者的抗原特异性免疫应答。如图6所示,在免疫后评估的所有时间点观察直至研究第6周PBMC中的抗原特异性免疫应答。在初免剂量(第2周和第3周)后,观察到Mamu-A*01抗原的每10⁶个PBMC中SFC的初始增加,接着是收缩(第4周)。值得注意的是,早在加强剂量后1周(第5周和第6周)就观察到每10⁶个PBMC的SFC增加

超过初始初免峰值应答。

[0460] 将抗原特异性免疫应答评估为对六种Mamu-A*01抗原的总应答。如图7所示,在免疫后评估的所有时间点观察直至研究第22周PBMC中的抗原特异性免疫应答。在初免剂量(第2周和第3周)后,对于六种Mamu-A*01抗原的总应答,观察到每 10^6 个PBMC的SFC初始增加,接着是收缩(第4周)。早在第4周施用的第一次加强剂量后的1周(第5周和第6周),就观察到每 10^6 个PBMC的SFC增加超过初始初免峰值应答,接着是收缩(第8周)。在第8周施用的第二次加强剂量后1周(第9周)再次观察到每 10^6 个PBMC的SFC增加,接着是收缩,其中每 10^6 个PBMC的SFC恢复稳定持续10周(第10-20周)。值得注意的是,早在前一次加强剂量后12周(第20周)施用的第三次加强剂量后的1周(第21周和第22周),就再次观察到每 10^6 个PBMC的SFC增加。

[0461] 因此,数据表明基于同源初免/加强AU-SAM的免疫策略导致恒河猴对非自身抗原产生有效、快速和稳定的抗原特异性免疫原性应答。

[0462] 序列

[0463] 本文提及的载体、盒和抗体在下文描述并由SEQ ID NO提及。

[0464]

曲美单抗 VL (SEQ ID NO:16)
曲美单抗 VH (SEQ ID NO:17)
曲美单抗 VH CDR1 (SEQ ID NO:18)
曲美单抗 VH CDR2 (SEQ ID NO:19)
曲美单抗 VH CDR3 (SEQ ID NO:20)
曲美单抗 VL CDR1 (SEQ ID NO:21)
曲美单抗 VL CDR2 (SEQ ID NO:22)
曲美单抗 VL CDR3 (SEQ ID NO:23)
德瓦鲁单抗(MEDI4736) VL (SEQ ID NO:24)
MEDI4736 VH (SEQ ID NO:25)
MEDI4736 VH CDR1 (SEQ ID NO:26)
MEDI4736 VH CDR2 (SEQ ID NO:27)
MEDI4736 VH CDR3 (SEQ ID NO:28)
MEDI4736 VL CDR1 (SEQ ID NO:29)
MEDI4736 VL CDR2 (SEQ ID NO:30)
MEDI4736 VL CDR3 (SEQ ID NO:31)
UbA76-25 聚体 PDTT 核苷酸(SEQ ID NO:32)
UbA76-25 聚体 PDTT 多肽(SEQ ID NO:33)
MAG-25 聚体 PDTT 核苷酸(SEQ ID NO:34)
MAG-25 聚体 PDTT 多肽(SEQ ID NO:35)
Ub7625 聚体 PDTT_NoSFL 核苷酸(SEQ ID NO:36)
Ub7625 聚体 PDTT_NoSFL 多肽(SEQ ID NO:37)

[0465]

ChAdV68.5WTnt.MAG25 聚体(SEQ ID NO:2); 具有 E1 (nt 577 至 3403)和 E3 (nt 27,125- 31,825)序列缺失的 AC_000011.1;在五个位置处被取代的相应 ATCC VR-594 核苷酸; 被插入以代替缺失 E1 的 CMV 启动子/增强子控制下的模型新抗原盒; 盒的 SV40 聚 A 3'
委内瑞拉马脑炎病毒[VEEV] (SEQ ID NO:3) GenBank: L01442.2
VEE-MAG25 聚体(SEQ ID NO:4); 含有 MAG-25 聚体 PDDT 核苷酸(碱基 30-1755)
委内瑞拉马脑炎病毒株 TC-83 [TC-83](SEQ ID NO:5) GenBank: L01443.1
VEEV 递送载体(SEQ ID NO:6); 具有核苷酸 7544-11175 缺失的 VEEV 基因组 [甲病毒结构蛋白被去除, 除了 E1 的最后 50 个氨基酸]
TC-83 递送载体(SEQ ID NO:7); 具有核苷酸 7544-11175 缺失的 TC-83 基因组 [甲病毒结构蛋白被去除]
VEEV 生产载体(SEQ ID NO:8); 具有核苷酸 7544-11175 缺失的 VEEV 基因组, 加 5' T7-启动子, 加 3'限制性位点
TC-83 生产载体(SEQ ID NO:9); 具有核苷酸 7544-11175 缺失的 TC-83 基因组, 加 5' T7-启动子, 加 3'限制性位点
VEE-UbAAY (SEQ ID NO:14); 具有 MHC I 类小鼠肿瘤表位 SIINFEKL 和 AH1-A5 插入的 VEEV 递送载体
VEE-荧光素酶(SEQ ID NO:15); 具有插入在 7545 处的荧光素酶基因的 VEEV 递送载体
泛素(SEQ ID NO:38)>UbG76 0-228
泛素 A76 (SEQ ID NO:39)>UbA76 0-228
HLA-A2 (MHC I 类)信号肽(SEQ ID NO:40)>MHC 信号肽 0-78
HLA-A2 (MHC I 类)跨膜结构域(SEQ ID NO:41)>HLA A2 TM 结构域 0-201
IgK 前导序列(SEQ ID NO:42)>IgK 前导序列 0-60
人类 DC-Lamp (SEQ ID NO:43)>人类 DCLAMP 0-3178
小鼠 LAMP1 (SEQ ID NO:44)>小鼠 Lamp1 0-1858
人类 Lamp1 cDNA (SEQ ID NO:45)>人类 Lamp1 0-2339
破伤风类毒素核酸序列(SEQ ID NO:46)
破伤风类毒素氨基酸序列(SEQ ID NO:47)
PADRE 核苷酸序列(SEQ ID NO:48)
PADRE 氨基酸序列(SEQ ID NO:49)
WPRE (SEQ ID NO:50)>WPRE 0-593
IRES (SEQ ID NO:51)>eGFP_IRES_SEAP_插入 1746-2335
GFP (SEQ ID NO:52)
SEAP (SEQ ID NO:53)
萤火虫荧光素酶(SEQ ID NO:54)
FMDV 2A (SEQ ID NO:55)
GPGPG 接头(SEQ ID NO:56)
SAM 体外转录模板 DNA (SEQ ID NO:57); 具有核苷酸 7544-11175 缺失的 VEEV 基因组, 加最小 5' T7-启动子(粗斜体)
<i>TAATACGACTCACTATAAT</i> GggcggcgcgatgagagaagcccagaccaattacctacccaaaATGG agaaagttcacgttgacatcgaggaagacagccattcctcagagctttgcagcggagcttcccgcagtttgaggtaga agccaagcaggtcactgataatgaccatgctaagccagagcgttttcgcactctggcttcaaaactgatcgaacggag

[0466]

```

gtggacccatccgacacgatccttgacattggaagtgcgccgccgcagaatgtattctaagcacaagtatcattgtat
ctgtccgatgagatgtgcggaagatccggacagattgtataagatgcaactaagctgaagaaaaactgaaggaata
actgataaggaattggacaagaaaaatgaaggagctcggccgctcatgagcgaccctgacctggaactgagactat
gtgctccacgacgacgagtcgtgctgctacgaagggaagtcgctgtttaccaggatgtatacgggtgacggacc
gacaagtctctataccaagccaataaggaggttagagtcgctactggataggcttgacaccacccttttatgtttaa
gaacttgctggagcatalccatcactctaccaactggccgacgaaaccgtgtaacggctcgtaacataggcctat
gcagctctgacgttatggagcggtcacgtagaggatgtcactcttagaaaagatattgaaaccatccaacaatgttc
tattctctgttgctcgcacctaccacgagaagaggactactgaggagctggcacctgccgtctgtattcacttac
gtggcaagcaaaattacacatgtcgggtgtgagactatagttagttgacgggtacgtcgttaaaagaatgctatcagt
ccaggcctgtatgggaagccttcaggctatgctgctacgatgcaccgcgagggtattctgtgctgcaaaagtacagaca
cattgaaacggggagagggtctctttcccgctgtgcacgtatgtccagctacattgtgtgaccaaactgactggcactgg
caacagatgtcagtgccggacgacgcgcaaaaactgctggttgggctcaaccagcgtatagtcgtcaacggctgcacc
cagagaaacccaataacatgaaaaattacctttgccgtagtggccaggcatttgctagggtggcaaaaggaataa
aggaagatcaagaagatgaaaggccactaggactacgagatagacagttatgcatgggggtgttgggcttttgaag
gcacaagataacatctattataagcgcccggatcccaaacatcatcaaaagtgaacagcgattccactcattcgtgct
gcccaggataggcagtaaacacattggagatcgggctgagaacaagaatcaggaaaaattgtagaggagcacaaggag
ccgtcacctctcattaccgcccaggacgtacaagaagctaaagtgcgacgccgatgaggctaaggaggtgctggaagc
cgaggagttgcgagctctaccactttggcagctgatgttgaggagccactctggaaagccgatgctgactgatg
ttacaagggtggggccggctcagtgagacacctcgtggcttgataaagggtaccagctacgctggcgaggacaa
gatcggctcttacgctgtgctttcccgaggctgtactcaagagtgaaaaattatctgcatccacctctcgtgaacaa
gtcatagtataacacactctggccgaaaaggcggttatgccgtggaaccataccatggtaaaagtgtggtgccagag
ggacatgcaataccgctccaggacttcaagctctgagtgaaggtgccaccattgtgtacaacgaactgagttcgtaaa
caggtagctgcaccatattgccacacatggaggagcgtgaacactgatgaagaatattacaaaactgtcaagcccag
cgagcacgacggcgaatacctgtacgacatcgcaggaacagtgctcaagaaagaactagctactgggctaggg
ctcacaggcgagctggtggatcctccttccatgaattgcctacgagagctctgagaacacgaccagccgctcctacc
aagtaccaacataggggtgtatggcggtccaggatcaggcaagctctggcatcattaaaagcgcagtcacaaaaaag
atctagtggtagcgcgaagaaagaaaactgtgcagaaaattataagggacgtcaagaaaatgaaagggctggacgtc
aatgccagaactgtggactcagtgctcttgaatggatgcaaacaccccgtagagaccctgtatattgacgaagctttgct
tgtcatgcaggtaactctcagagcgtcatagccattataagacctaaaaaggcagtgctctgcccgggatcccaaacagt
gcggttttttaacatgatgtgcctgaaagtgcatttaaccacgagattgacacacaagtctccacaaaagcatctctgc
cgtgactaaatctgtgactcggctgtctcaacctgttttacgacaaaaaatgagaacgacgaatccgaaagagact
aagattgtgattgacactaccggcagtaacaaactaagcaggacgatctcattctactgtttcagagggtgggtgaa
gcagttgcaaatagattacaaggcaacgaataatgacggcagctgcctctcaagggtgaccgtaaaaggtgtgtat
gccgtcgggtacaaggtgaatgaaaatcctctgtacgacccacctcagaacatgtgaacgtcctactgaccgcacgg
aggaccgcatcgtgtgaaaacactagccggcgacccatggataaaaactgactgccaagtaccctgggaatttca
ctgccacgatagaggagtggcaagcagagcatgatgccatcatgaggcacatctggagagaccggaccctaccgac
gtctccagaataaggcaaacgtgtgtggccaaggcttttagtgcgggtgctgaagaccgctggcatagacatgacca
ctgaacaatggaacactgtggatttttgaacggacaaagctcactcagcagagatagattgaaccaactatgctg
aggttctttggactcagctggactccggctatcttctgacccactgttccgttatccattaggaataatcactgggataac
tccccgtcgcctaacatgtacgggctgaataaagaagtgtccgtcagctctctgcaggatcccacaactgcctcggg
cagttgccactggaagatgctatgacatgaacactggtacactgcgcaattatgatccgcgataaacctagtacgtga
aacagaagactgcctcatgcttagtctccaccataatgaacaccacagagtacttttcttctcgtcagcaaatga
agggcagaactgtcctggtggtcggggaaaagttgtccgtccaggcaaaatggtgactgggtgacagaccggcctg
aggctacctcagagctcggctggatttagcatccagggtgatgtgccaaatatgacataatattgttaatgtgagga
ccccataaaataccatcactatcagcagtgtagaacatgccattaaactgacatgttgaccaagaaagctgtctgc
atctgaatcccggcggaacctgtgtcagcataggttatggttacgctgacagggccagcgaagcatcattggtgctata
gcgcgggcagttcaagtttccgggtatgcaaacgaaatcctcactgaagagacggaagttctgtttgtattcattgggt
acgatgcaagggccgtacgcacaatccttacaagcttcatcaaccttgaccaacattatacaggttccagactccagc
aagccggatgtgcacctcatatcatgtggtcggaggggatattgccacggccaccgaaggagtattataaatgctgc
taacagcaaaaggacaacctggcggaggggtgtcggagcgtgtataagaaattcccgaaagcttgatttacggc

```

[0467]

```

gatcgaagtaggaaaagcgcgactggtcaaaggcagctaaacatatcattcatgccgtaggaccaaaactcaaca
agtttcggagggtgaaggtagacaacagttggcagaggctatgagtcctcgcctaagattgtcaacgataacaattaca
agtcagtagcgttccactgtgtccaccggcatctttccgggaacaaagatcgactaacccaatcattgaaccatttgc
gacagctttagacaccactgatgcagatgtagccatatactgcagggacaagaaatgggaaatgactctcaaggaagc
agtggctaggagagaagcagtggaaggagatgcatatccgacgactctcagtgacagaacctgatgcagagctggt
gagggtgcatccgaagattcttggctggaaggaaagggtcagcacaaagcagtggaacaaacttctcatatttggaa
gggaccaagtaccaggcggccaaggatagcagaaatfaatgcatgtggcccgttgaacggaggccaatga
gcaggtatgcatgtatctcgggagaagcatgagcagtaggtcgaatgccccgtcgaagagtcggaagcctc
cacaccactagcagctgccttctgtgcatccatgccatgactccagaagagtagcgcctaaagcctcacgt
ccagaacaaactgtgtgctccttccattgccgaagtatagaatcactggtgtgcagaagatccaatgtccccag
cctatattgttctaccgaaagtgcctgcgtatattcatccaaggaaatctcgtggaacaccaccggtagacgagact
ccggagccatcggcagagaaccaatccacagaggggacacctgaacaaccaccactataaccgaggatgagacca
ggactagaacgcctgagccgatcatcagagaggaagaggatagcataagtttgcctgcatagggcccacc
caccaggtgctgcaagtcgagcgacatcacgggcccctctgtatctagctcatctggtccattctcatgcatc
cgactttgatgtggacagttatccatacttgacacctggagggagctagcgtgaccagcggggcaacgtcagccga
gactaaccttacttcgcaagagatggagttctggcgcgaccgggtgctgcgcctcgaacagtattcaggaaccctc
cacatcccgtccgcgcacaagaacaccgtcacttgacccagcaggggcctgctcgagaaccagcctagttccacc
cgccaggcgtgaatagggtgactagagaggagctcagggcgttaccctgcacgcactcctagcaggtcggctc
tcgagaaccagcctggtctcaaccgccaggcgtaaatagggtgattacaagagaggagttagggcgtcgtagca
caacaacaatgacgggttgatcgggtgcatacatcttctccgacaccggtcaagggcatttacaacaaaaatcagt
aaggcaaacggtgctatccgaagtgttggagaggaccgaattggagattcgtatgccccgcgcctcgaacaaga
aaaagaagaattactacgcaagaaattacagttaaatcccacacctgtaacagaagcagataccagtcaggaaggt
ggagaacatgaaagccataacagctagacgtattctgcaaggcctagggcattatttgaaggcagaaggaaaagtggg
gtgctaccgaacctgcatcctgttcttcttattcatctagtgtgaaccgtgcctttcaagccccaggcgcagtgga
gctgtaacgccatgttgaagagaacttccgactgtggcttctactgtattattccagagtacgatgcctatttggacat
ggtgacggagcttcatgctcttagacactgccagttttgccctgcaagctgcgcagcttccaagaaacactccta
tttgaaccacaatacagtcggcagtccttcagcgtaccagaacacgctccagaacgctctggcagctgccacaaa
aagaaatgcaatgtcacgcaaatgagagaattgccgtattggattcggcggccttfaatgtggaatgctcaagaata
tgcgtgtaataatgaatattgggaaacgtttaaagaaaacccatcaggcttactgaagaaaacgtggtaaattacattac
caaataaaaaggacaaaagctgctgctcttttgcgaagacacataatttgaatattgtgcaggacataccaatggacag
gtttgtaatggactaaagagagacgtgaaagtgactccaggaaacaaacatactgaagaacggcccaaggtagcaggt
gatccaggctgccgatccgctagcaacagcgtatctgtcggaaatccaccgagagctggttaggagattaaatgcggt
cctgctccgaacattcatacactgtttgatalgtcggctgaagactttgacgctattatagccgagcactccagcctggg
gattgtgtctggaactgacatcgcgtcgttgataaaagtgaggacgacgccatggctctgaccgcgttaatgattctg
gaagacttaggtgtggacgcagagctgtgacgctgattgaggcggcttccggcgaatattcatcaatacatttgccacc
aaaactaaatttaattcggagccatgatgaaatctggaatgttctctcacactgtttgtgaacacagtcattaacattgta
cgcaagcagagtggtgagagaacggcctaaccggatccatgtgcagcattcattggagatgacaatcgtgaaagg
agtcaaatcggacaaatfaatggcagacaggtgcgccacctggtgaaatggaaagtcagattatagatgctgtggtg
ggcgagaaagcgccttattctgtggagggttatttctgtgactccgtgaccggcacagcgtgccgtgtggcagacc
cctaaaaaggctgtttaaagcttggcaaaccttggcagcagacgatgaacatgatgacaggagaaggcattgcat
gaagagtcaacacgctggaaccagtggttattcttcagagctgtgcaaggcagtagaatcaaggatgaaaccgta
ggaactccatcatagttatggccatgactactctagctagcagtgtaaatcattcagctacctgagagggggcccctata
actctctacggcTAAcctgaatggactacgactTatcacgcccaaacatttacagccgcgggtgtcaaaaaccgcgtg
gacgtggttaacatccctgctgggaggatcagccgaattattataattggcttggctgctggctactattgtggccatgtac
gtgctgaccaaccagaaacataattgaatacagcagcaattggcaagctgctfacatagaactcgcggcgattggcatg
ccgcttaaaattttattttttctttcttccgaatcggttttttaaatattcAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

```

SAM 模板生产载体(SEQ ID NO:58); 具有核苷酸 7544-11175 缺失的 VEEV 基

因组，加 5' T7-启动子(粗斜体)和正向引物结合位点，加 3'限制性位点

gggtatgtggacgcggccgc**TAATACGACTCACTATAAT**Gggcggcgcgatgagagaagcccagacc
 aattacctacccaaaATGGagaaagttcacgttgacatcgaggaagacagccattcctcagagctttgcagcggag
 ctccccgagtttgaggtagaagccaagcaggtcactgataatgacctgctaagccagagcgtttcgcacttggttc
 aaaactgatcgaaacggaggtggaccatccgacacgatccttgacattggaagtgcgcccggcgagaatgtattc
 taagcacaagtatcattglatctgctcgatgagatgtgcggaagatccggacagattgataagatgcaactaagctgaa
 gaaaaactgtaaggaaataactgataaggaaatggacaagaaaatgaaggagctcggccgctcatgagcgacctg
 acctggaactgagactatgtgctccacgacgacgagtcgtgctgctacgaagggaagtcgctgttaccaggatgt
 atacgcggtgacggaccgacaagtctctatcaccaagccaataagggagttagagtcgcctactggataggcttgac
 accacccttttatgttaagaactggctggagcatatccatcactctaccaactgggcccagcgaaccggtgtaacg
 gctcgtacaatagccctatgcagctcgtgactgtaggagcggcgtacgtagaggatgtccallcttagaaagaatgtt
 aaaccatccaacaatgttctatctctgttggctgaccatctaccacgagaagagggacttactgaggagctggcacct
 gccgtctgtattcacttacgtggcaagcaaaattacacatgctggtgtgagactatagttagtgacgagggtacgtct
 aaaaagaatagctatcagtcaggcctgtatgggaagccttcaggctatgctgctacgatgaccgaggggattctgtg
 ctgcaaaagtacagacacattgaacggggagaggggtctctttccctgtgacgatggtgcccagctacattgtgtgacc
 aaatgactggcactggcaacagatgtcagtgccgacgacgcgcaaaaactgctggtgggctcaaccagcgtatag
 tcgtcaacggctgcaaccagagaacaccaataacatgaaaaattacctttgccgtagtggcccaggcattgtctagg
 tgggcaaaaggaataaggaagatcaagaagatgaaaggccactaggactacgagatagacagttatgcatggggtg
 ttgttgggcttttagaaggcacaagataacatctattataagcggccgatacccaaacatcatcaaaagtgaacagcg
 atttccactcattcgtgctgccaggataggcagtaaacattggagatcgggctgagaacaagaatcaggaaaaatgta
 gaggagcacaaggagccgtcactctcattaccgccgaggacgtacaagaagctaaagtgcgcagccgatgaggcta
 aggaggtgctgaagccgaggaggtgcgcgacgtctaccacctttggcagctgatgttgaggagccactctggaa
 gccgatgtcactgtatgtacaagaggctgggcccggctcagtgagacacctcgtggctgataaaggttaccagct
 acgctggcaggacaagatcggctcttacgctgtctttctcgcaggctgactcaagagtgaaaaattacttgcaccc
 acctctcgtgaacaagtcatagtgataacacacttggccgaaaaaggcgttatgccgtggaaaccataacctggtaa
 agtagtggtgccagaggacatgcaataccgtccaggacttcaagctctgagtgaaagtgccaccattgtgtacaac
 gaactgtagttcgtaaacaggtacctgcaccatattgccacacatggaggagcgtgaacactgatgaaatattaca
 aaactgtcaagcccagcagcagcagcggcgaatacctgtacgacatcgacaggaacaagtgcgtcaagaaagaact
 agtcactgggctagggtcacaggcagctggtggatcctccctccatgaattgcctacgagagctgagaacacg
 accagccgtccttaccaaagtaccaacataggggtgtatggcgtgccaggatcaggcaagtctggcatcattaaag
 cgcagtcacaaaaaagatctagtgtgagcgaagaagaagaactgtgcagaaattataaggagctcaagaaaa
 tgaagggtggagctcaatgccagaactgtggactcagctcttgaatggatgcaaacaccccgtagagaccctgta
 tattgacgaagctttgctgtcatgaggtactctcagagcgtcatagccattataagacctaaaaaggcagtgctctgc
 ggggatcccaacaagtgcggtttttaacatgatgtgcctgaaagtgcattttaaccacgagatttgacacaaagtctcc
 acaaaagcatctctcggcttgcactaaatctgtgactcggctgctcaacctgttttacgacaaaaaatgagaacga
 cgaatccgaaagagactaagattgtgattgacactaccggcagtaacaaaactaaagcaggacgatctcattctactgt
 ttcagagggtgggtgaagcagttgcaaatagattacaaggaacgaaataatgacggcagctgctctcaagggctg
 acccgtaaaagggtgtatgccgtcgggtacaaggtgaatgaaatcctctgtacgacccacctcagaacatgtaacgt
 cctactgaccgcacggaggaccgcatcgtgtggaaaactagccggcgaccatggataaaaactgactgcca
 agtaccctgggaatttactgccacgatagaggagtggaagcagagcatgatgccatcatgaggcacatcttgaga
 gaccggaccctaccgactcttcagaataaggcaaacgtgtgtgggccaaggcttagtgccggtgctgaagaccg
 ctggcatagacatgaccactgaacaatggaacactgtggatttttgaacggacaagctcactcagcagagatag
 attgaaccaactatgcgtgaggtctttggactcgtctggactccggtctatttctgacccactgttccgttatccattg
 gaataactcactgggataactcccgtcgcctaacatgtacgggctgaataaagaagtggtccgtcagctctctcaggt
 acccacaactgcctcgggaggtgacctggaagagctatgacatgaacactggtacactgcgcaattatgatccgcg
 cataaacctagtacctgtaaacagaagactgcctcatgctttagctccaccataatgaacacccacagagtgactttct
 tcatcgtcagcaaatgaaaggcagaactgtcctggtgctggggaaaaagtgtccgtccaggcaaaatggtgact
 ggtgtcagaccgctgaggctacctcagagctcggctggatttagcatccagggtgatgtgccaaatgatgacata
 atattgtaatgtgaggaccatataaaatccatcactatcagcagtgtaagaccatgccattaaagcttagcatgtga
 ccaagaaagctgtctgcatcgaatcccggcgaacctgtgfcagcataggttatggttacgctgacaggggcagcga

[0468]

[0469]

```

aagcatcattggtgctatagcgcggcagttcaagtttccgggtatgcaaaccgaatcctcacttgaagagacggaa
gttctgttgtatcattgggtacgacgcgaaggcccgtacgcacaatccttacaagcttcatcaaccttgaccaacattat
acaggttccagactccacgaagccggatgtgcaccctcatatcatgtggtgaggggatattgccacggccaccgaa
ggagtgtattataaatgctgctaacagcaaaggacaacctggcggagggtgtgcggagcgtgtataagaaatcccg
gaaagcttcgatttacagccgatcgaagtaggaaaagcgcgactggtcaaagggtgcagctaaacatacattcatgccg
taggaccaaactcaacaagttcggagggtgaaggtgacaacagttggcagaggcttatgagtcacgctaagatt
gtcaacgataacaattacaagtcagtagcgaatccactgtgtccaccggcatctttccgggaacaagaatcgactaac
caatcattgaaccatttctgacagctttagacaccactgatgcagatgtagccatatactgcagggacaagaatggga
aatgactctcaaggaagcagtggtgctagggagagaagcagtgaggagatagcatalccgacgactctcagtgacaga
acctgatgcagagctggtgaggggtgcattccgaaggttcttggctggaaggaagggtacagcacaagcgatggca
aaactttcatalattggaagggaccaagttcaccaggcggccaaggatagcagaaattaatgccatgtggcccgtt
gcaacggaggccaatgagcaggtatgcatgtatctcctggagaaagcatgagcagattaggtcgaatgccccgtc
gaagagtcggaaagcctccacaccactagcacgctgcttgccttgcattgcattccatgccatgactccagaaagagtacag
cgctaaaagcctcacgtccagaacaattactgtgtcctatcttccattgccgaagatagaatcactggtgtgcag
aagatccaatgctcccagcctatattgttctcaccgaaagtgcctgcgtatattcatccaaggaagtatctcgtggaaaca
ccaccggtagacgagactccggagccatcggcagagaaccaatccacagaggggacacctgaacaaccaccactta
taaccgaggatgagaccaggactagaacgctgagccgatcatatcgaagaggaagaaggatagcataagtttg
ctgtcagatggcccaccaccaggtgctgaagtcgaggcagacattcacggccgccctctgtatctagctcatcct
ggtcattcctcatgcatccgactttgatgtggacagttatccatacttgacaccctggagggaagctagcgtgaccagcg
gggcaacgtcagccgagactaactcttacttcgaaagagatggaagttctggcgcgaccggtgctgcgctcga
cagtattcaggaaccctccacatcccgtccgcgcacaagaacaccgtcacttgacccagcagggcctgctcgaga
accagcctagttccaccccggcaggcgtgaatagggtgactactagagaggagctcgaagcgttaccctgcacg
cactcctagcaggtcggctcgcagaaccagcctggtctcaaccgccaggcgtaaatagggtgattacaagagagg
agtttgaggcgtcgtgacacaacaacaatgacggttgatgcgggtgcatacatctttcctccgacaccggtcaaggg
catttaacaacaaaatcagtaaggcaaacgggtgctatccgaagtggtgtggagaggaccgaattggagattcgtatgc
cccgcctcgcaccaagaaaaagaaatctacgcaagaaatcagttaaatccacacctgtaacagaagcag
ataccagtcagggaaggtggagaacatgaaagccataacagctagacgtatctgcaaggcctagggcattattgaa
gcagaaggaaggtggagtgtaccgaacctgcatcctgttcccttattcatctagtgtgaacctgctctttcaagcc
ccaaggtcgcagtggaaagcctgtaacgccatgttgaagagaacttccgactgtggcttctactgtattatccagagt
acgatgcctatttggacatggttgacggagcttcatgctgcttagacactgccagttttgccctgcaagctcgcagctt
tccaaagaacactcctatttggaaccacaatacgtcggcagtgcttcagcgatccagaacacgctccagaacgtc
ctggcagctgccacaaaagaaatgcaatgtcacgcaaatgagagaattgccgtaattggattcggcggccttfaatgt
ggaatgctcaagaaatgctgtgtaataatgaatattgggaaacgtttaaagaaaacccatcaggcttactgaagaaa
acgtggtaaatfacattaccaaatataaggacaaaagctgctgctcttttgcgaagacacataattgaaatgttgc
ggacataccaatggacaggttgaatggacttaagagagacgtgaaagtgactccaggaacaaaacatactgaaga
acggccaaggtacaggtgatccaggctgccgatccgtagcaacagcgtatctgtcggaaatccaccgagagctgg
ttaggagattaaatcggctcctgctccgaactcatacactgtttgatatgtcggctgaagactttgacgctattatagcc
gagcactccagcctggggattgttctggaactgacatcgcgtcgttgataaaagtgaggacgacgccatggctct
gaccgcgttaatgattctggaaagacttaggtgtggacgcagagctgttgacgctgattgagcggccttccggcgaattt
catcaatacatttcccactaaaactaaatataatcggagccatgatgaaatcggaaatgttccctcacactgtttgtgaa
acagtcattaacattgtaatcgaagcagagtgtgagagaacggtaaccggatcaccatgtgcagcattcattggaga
tgacaatcgtgaaaggagtcacaatcggacaaatataggcagacaggtgcgccacctgggtgaaatggaagtcaag
attatagatgctgtggtggcgagaaagcgccttattctgtggagggtttattttgtgactccgtgaccggcacagcgt
gccgtgtggcagaccccctaaaaggctgttaagctggcaaacctctggcagcagacgatgaacatgatgatgaca
ggagaagggcattgcatgaagagtcacacgctggaaccgagtggtatfcttccagagctgtgcaaggcagtagaat
caaggtatgaaaccgtaggaactccatcatagttatggccatgactactctagctagcaggtgtaaatcattcagctacct
gagaggggcccctataactctctacggcTAActgaaatggactacgactTatcacgcccacaaatccacagccgcg
gtgtcaaaaaccgcgtggacgtggttaacatccctgctgggaggatcagccgtaattattataattggcttggctggct
actattgtggccatgtacgtgctgaccaaccagaacataattgaatacagcagcaattggcaagctgttacatagaac
tcggcggcattggcatgccgcttaaaattttattttttcttttccgaatcgattttgatttaattttcAAA

```

[0470]

AA
 AA
 AAatcgtagttaaac

[0471] 等效内容和范围

[0472] 在权利要求书中,除非指示相反情形或从上下文中另外明显可见,否则例如“一个”、“一种”和“所述”的冠词可意指一个或多个。除非指示相反情形或从上下文中另外明显可见,否则如果一个、多于一个或所有组成员存在于、用于给定产物或工艺或以其他方式与给定产物或工艺相关,则认为其符合在组的一个或多个成员之间包括“或”的权利要求书或说明书。本公开包括其中恰好一个组成员存在于、用于给定产物或工艺或以其他方式与给定产物或工艺相关的实施方案。本公开包括其中多于一个或所有的组成员存在于、用于给定产物或工艺或以其他方式与给定产物或工艺相关的实施方案。

[0473] 此外,本公开涵盖其中将来自一项或多项所列示权利要求的一种或多种限制、要素、条款和说明性术语引入到另一项权利要求中的所有变化形式、组合和排列。例如,从属于另一项权利要求的任何权利要求可修改为包括一种或多种在从属于同一基本权利要求的任何其他权利要求中所发现的限制。在要素以列表形式(例如以马库什组(Markush group)格式)呈现的情况下,还公开了所述要素的每个亚组,并且可从所述组中移除任何要素。应理解,一般而言,在本公开或本公开的方面被称为包括特定要素和/或特征的情况下,本公开的某些实施方案或本公开的方面由此类要素和/或特征组成或基本上由其组成。出于简单性目的,那些实施方案在本文中未以所述语言明确陈述。还应注意,术语“包含”和“含有”旨在为开放性的并允许包括其他要素或步骤。在给出范围的情况下,端点包括在内。此外,除非另有指示或从上下文和本领域技术人员的理解另外显而易见,否则表示为范围的值在本公开的不同实施方案中可采用所陈述范围内的任何具体值或子范围,除非上下文另外明确地指明,否则精确到所述范围下限单位的十分之一。

[0474] 本申请涉及各种已颁布专利、公开专利申请、期刊文章和其他出版物,所有这些都通过引用并入本文。如果所并入的任何参考文献与本说明书之间存在冲突,则应以本说明书为准。另外,可从任何一项或多项权利要求中明确地排除属于现有技术的本公开的任何特定实施方案。因为这样的实施方案被认为是本领域普通技术人员已知的,所以即使在本文中未明确阐述排除,它们也可被排除。出于任何原因,无论是否与现有技术的存在相关,本公开的任何特定实施方案都可从任何权利要求中排除。

[0475] 本领域技术人员将认识到或能够仅使用常规实验来确定本文所阐述的具体实施方案的许多等效形式。本文所阐述的本发明实施方案的范围不旨在限于上述说明书,而是如所附权利要求中所陈述的。本领域普通技术人员将理解,在不脱离如以下权利要求所限定的本公开的精神或范围的情况下,可对本说明书进行各种改变和修改。

序列表

<110> 磨石生物公司 (GRITSTONE ONCOLOGY, INC.)

<120> 加帽化合物、组合物及其使用方法

<130> GS0-088W0

<140>

<141>

<150> US63/013, 456

<151> 2020-04-21

<150> US63/020, 473

<151> 2020-05-05

<160> 56

<170> PatentIn version 3.5

<210> 1

<211> 36519

<212> DNA

<213> 人工序列 (Artificial Sequence)

[0001]

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 1

```

ccatcttcaa taatatacct caaacttttt gtgcgcgta atatgcaaat gaggcgtttg      60
aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgttagg ggcggggcga      120
gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag      180
tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttccgc gctctctgac      240
aggaaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact      300
gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg caggaggag tatttgccga      360
gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa      420
tttccgcgta cgggtgcaaa gtccggtgtt tttacgtagg tgtcagctga tcgccagggt      480
atntaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcgag aagagttttc      540
tcctccgcgc cgcgagtcag atctacactt tgaaagatga ggcacctgag agacctgccc      600
gatgagaaaa tcatcatcgc ttccgggaac gagattctgg aactggtggt aaatgcatg      660

```

	atgggcgacg accctccgga gccccccacc ccatttgaga caccttcgct gcacgatttg	720
	tatgatctgg aggtggatgt gcccaggac gatcccaatg aggaggcggg aaatgatttt	780
	tttagcgatg ccgcgctgct agctgccgag gaggcttcga gctctagctc agacagcgac	840
	tcttcaactgc ataccctag acccggcaga ggtgagaaaa agatccccga gcttaaaggg	900
	gaagagatgg acttgcgctg ctatgaggaa tgcttgcccc cgagcgatga tgaggacgag	960
	caggcgatcc agaacgcagc gagccagga gtgcaagccg ccagcgagag ctttgcgctg	1020
	gactgccccg ctctgccccg acacggctgt aagtcttctg aatttcatcg catgaatact	1080
	ggagataaag ctgtgttctg tgcactttgc tatatgagag cttacaacca ttgtgtttac	1140
	agtaagtgtg attaagttga acttttagagg gaggcagaga gcagggtgac tgggcgatga	1200
	ctggtttatt tatgtatata tgttctttat ataggtcccg tctctgacgc agatgatgag	1260
	acccccacta caaagtccac ttcgtcacc ccagaaattg gcacatctcc acctgagaat	1320
	attgttagac cagttcctgt tagagccact gggaggagag cagctgtgga atgtttggat	1380
[0002]	gacttgctac aggggtgggg tgaacctttg gacttgtgta cccgaaacg ccccaggcac	1440
	taagtccac acatgtgtgt ttacttgagg tgatgtcagt atttataggg tgtggagtgc	1500
	aataaaaaat gtgttgactt taagtgcgtg gtttatgact caggggtggg gactgtgagt	1560
	atataagcag gtgcagacct gtgtggttag ctcagagcgg catggagatt tggacggtct	1620
	tggaagactt tcacaagact agacagctgc tagagaacgc ctcgaacgga gtctcttacc	1680
	tgtggagatt ctgcttcggt ggcgacctag ctaggctagt ctacagggcc aaacaggatt	1740
	atagtgaaca atttgagggt attttgagag agtgttctgg tctttttgac gctcttaact	1800
	tgggccatca gtctcacttt aaccagagga ttctgagagc cettgatttt actactctg	1860
	gcagaaccac tgcagcagta gccttttttg cttttattct tgacaaatgg agtcaagaaa	1920
	cccatttcag cagggattac cagctggatt tcttagcagt agctttgtgg agaacatgga	1980
	agtgccagcg cctgaatgca atctccggct acttgccggt acagccgcta gacactctga	2040
	ggatcctgaa tctccaggag agtcccaggg cacccaacg tcgccagcag cagcagcagg	2100
	aggaggatca agaagagaac ccgagagccg gcctggacct tccggcggag gaggaggagt	2160

	agctgacctg tttcctgaac tgcgccgggt gctgactagg tcttcgagtg gtcgggagag	2220
	ggggattaag cgggagaggc atgatgagac taatcacaga actgaactga ctgtgggtct	2280
	gatgagtcgc aagcgcaccag aaacagtgtg gtggcatgag gtgcagtcga ctggcacaga	2340
	tgagggtgctg gtgatgcatg agaggttttc tctagaacaa gtcaagactt gttgggttaga	2400
	gcctgaggat gattgggagg tagccatcag gaattatgcc aagctggctc tgaggccaga	2460
	caagaagtac aagattacta agctgataaa tatcagaaat gcctgctaca tctcagggaa	2520
	tggggctgaa gtggagatct gtctccagga aagggtggct ttcagatgct gcatgatgaa	2580
	tatgtaccgg ggagtggtgg gcatggatgg ggttaccttt atgaacatga ggttcagggg	2640
	agatgggtat aatggcacgg tctttatggc caataccaag ctgacagtcc atggctgctc	2700
	cttctttggg ttttaataaca cctgcatcga ggccctgggt caggctcgtg tgaggggctg	2760
	cagtttttca gccaaactgga tgggggtcgt gggcaggacc aagagtatgc tgtccgtgaa	2820
	gaaatgcttg tttgagaggt gccacctggg ggtgatgagc gagggcgaag ccagaatccg	2880
	ccactgcgcc tctaccgaga cgggctgctt tgtgctgtgc aaggccaatg ctaagatcaa	2940
[0003]	gcataaatag atctgtggag cctcggacga gcgcggctac cagatgctga cctgcgccgg	3000
	cgggaacagc catatgctgg ccaccgtaca tgtgcttcc catgctcga agccctggcc	3060
	cgagttcgag cacaatgtca tgaccaggtg caatatgcat ctgggggtccc gccgaggcat	3120
	gttcatgccc taccagtga acctgaatta tgtgaagggt ctgctggagc ccgatgcat	3180
	gtccagagtg agcctgacgg ggggtgttga catgaatgtg gaggtgtgga agattctgag	3240
	atatgatgaa tccaagacca ggtgccgagc ctgcgagtc ggagggaagc atgccaggtt	3300
	ccagcccgtg tgtgtggatg tgaccgagga cctgcgacce gatcatttgg tgttgcctg	3360
	caccgggacg gagttcggtt ccagcgggga agaactgac tagagtgagt agtgttctgg	3420
	ggcgggggag gacctgcatg agggccagaa taactgaaat ctgtgctttt ctgtgtgttg	3480
	cagcagcatg agcggaaagc gctcctttga gggaggggta ttcagccctt atctgacggg	3540
	gcgtctccc tcttggcggg gagtgcgtca gaatgtgatg ggatccacgg tggacggccg	3600
	gcccgtgcag cccgcgaact cttcaaccct gacctatgca accctgagct cttcgtcgtt	3660
	ggacgcagct gccgccgag ctgctgcata tgcccagc gccgtgcgcg gaatggccat	3720

	gggcgccggc tactacggca ctctggtggc caactcgagt tccaccaata atccccccag	3780
	cctgaacgag gagaagctgt tgctgctgat ggcccagctc gaggccttga cccagcgcct	3840
	gggcgagctg acccagcagg tggtcagct gcaggagcag acgcgggccc cggttgccac	3900
	ggtgaaatcc aaataaaaaa tgaatcaata aataaacgga gacggttgtt gattttaaca	3960
	cagagtctga atctttattt gatttttcgc gcgcggtagg ccctggacca ccggtctcga	4020
	tcattgagca cccggtgat cttttccagg acccggtaga ggtgggcttg gatgttgagg	4080
	tacatgggca tgagcccgtc ccgggggtgg aggtagctcc attgcagggc ctctgctcgc	4140
	ggggtggtgt tgtaaatcac ccagtcatac caggggcgca gggcatggtg ttgcacaata	4200
	tctttgagga ggagactgat ggccacgggc agccctttgg ttaggtgtt tacaaatctg	4260
	ttgagctggg agggatgcat gcggggggag atgaggtgca tcttggcctg gatcttgaga	4320
	ttggcgatgt taccgccag atcccgcctg gggttcatgt tgtgcaggac caccagcacg	4380
	gtglatccgg tgcacttggg gaatttatca tgcaacttgg aagggaaggc gtgaaagaat	4440
[0004]	ttggcgacgc ctttgtgcc gccaggttt tccatgact catccatgat gatggcgatg	4500
	ggcccgtggg cggcggcctg ggcaaagac tttcgggggt cggacacatc atagtgttg	4560
	tcttgggtga ggtcatcata ggccatttta atgaatttgg ggcggagggt gccggactgg	4620
	gggacaaagg taccctcgat cccgggggcg tagttccct cacagatctg catctcccag	4680
	gctttgagct cggagggggg gatcatgtcc acctgcgggg cgataaagaa cacggttcc	4740
	ggggcggggg agatgagctg ggccgaaagc aagtccgga gcagctggga cttgccgcag	4800
	ccggtggggc cgtagatgac cccgatgacc ggctgcaggt ggtagttgag ggagagacag	4860
	ctgccgtect cccggaggag gggggccacc tcgttcatca tctcgcgac gtcatgttc	4920
	tcgvcacca gttccgccag gaggcgtct cccccaggg ataggagctc ctggagcgag	4980
	gcgaagtttt tcagcggctt gagtccgtcg gccatgggca ttttggagag ggtttgttgc	5040
	aagagtcca gccggtcca gagctcgggt atgtgctcta cgcatctcg atccagcaga	5100
	cctcctcgtt tcgcggttg ggacggctgc gggagtaggg caccagacga tggcgctcca	5160
	gcgcagccag ggtccggtcc ttccagggtc gcagcgtccg cgtcagggtg gtctccgtca	5220

	cggtgaaggg gtgcgcgccg ggctgggcgc ttgcgagggt gcgcttcagg ctcacccggc	5280
	tggtcgaaaa ccgctcccga tcggcgcctt gcgcgtcggc caggtagcaa ttgaccatga	5340
	gttcgtagtt gagcgcctcg gccgcgtggc ctttggcgcg gagcttacct ttggaagtct	5400
	gcccgcaggc gggacagagg agggacttga gggcgtagag cttgggggcg aggaagacgg	5460
	actcgggggc gtaggcgtcc gcgccgcagt gggcgcagac ggtctcgcac tccacgagcc	5520
	aggtgaggtc gggctggtcg gggtcaaaaa ccagtttccc gccgttcttt ttgatgcgtt	5580
	tcttaccttt ggtctccatg agctcgtgtc cccgctgggt gacaaagagg ctgtccgtgt	5640
	ccccgtagac cgactttatg ggccggctct cgagcgggtg gccgcggtcc tctcgtaga	5700
	ggaacccccg cactccgag acgaaagccc gggtcaggc cagcacgaag gaggccactg	5760
	gggacgggta gcggtcgtt tccaccagcg ggtccacctt ttccagggta tgcaaacaca	5820
	tgtccccctc gtccacatcc aggaaggatga ttggcttgta agtgtaggcc acgtgaccgg	5880
	gggtcccggc cgggggggta taaaagggtg cgggtccctg ctogtctca ctgtcttccg	5940
	gatcgtgtc caggagcgc agctgttggg gtaggtatc cctctcgaag gcgggcatga	6000
[0005]	cctcggcact caggttgta gtttctagaa acgaggagga ttgatattg acggtgccgg	6060
	cggagatgcc tttcaagagc ccctcgtcca tctggtcaga aaagacgac tttttgttgt	6120
	cgagcttggg gcgcaaggag ccgtagaggc cgttgagag gagcttggcg atggagcgca	6180
	tggtctggtt ttttcttg tcggcgcgct ccttggcggc gatgttgagc tgcacgtact	6240
	cgcgccac gcacttccat tcggggaaga cggtggtcag ctctcgggc acgattctga	6300
	cctgccagcc ccgattatgc agggatgatga ggtccacact ggtggccacc tcgccgcgca	6360
	ggggctcatt agtccagcag aggcgtccgc ccttgcgca gcagaagggg ggcagggggg	6420
	ccagcatgac ctctcgggg gggtcggcat cgatggtgaa gatgccgggc aggaggtcgg	6480
	ggtcaaagta gctgatgaa gtggccagat cgtccagggc agcttgccat tcgpcacgg	6540
	ccagcgcgcg ctcttaggga ctgaggggcg tgccccaggc catgggatgg gtaagcgcgg	6600
	aggcgtacat gccgcagat tcgtagacgt agaggggctc ctgaggatg ccgatgtagg	6660
	tgggtagca gcgcccccg cggatgctgg cgcgcacgta gtcatacagc tcgtgcgagg	6720
	gggcgaggag cccccggccc aggttgggtc gactgggctt ttcggcgcgg tagacgatct	6780

	ggcggaaaat ggcattgcgag ttggaggaga tggtagggcct ttggaagatg ttgaagtggg	6840
	cgtggggcag tccgaccgag tcgcggatga agtgggcgta ggagtcttgc agcttggcga	6900
	cgagctcggc ggtgactagg acgtccagag cgcagtagtc gagggctctcc tggatgatgt	6960
	catacttgag ctgtcccttt tgtttccaca gctcgcggtt gagaaggaac tcttcgcggt	7020
	ccttccagta ctcttcgagg gggaacctgt cctgatctgc acgtaagag cctagcatgt	7080
	agaactgggt gacggccttg taggcgcagc agcccttctc cacggggagg gcgtaggcct	7140
	gggcggcctt gcgcaggag gtgtgcgtga gggcgaaagt gtccctgacc atgaccttga	7200
	ggaactgggt cttgaagtgc atatcgtgc agccccctg ctcccagagc tggaaagtccg	7260
	tgcgcttctt gtaggcgggg ttgggcaaag cgaagtaac atcgttgaag aggatcttgc	7320
	ccgcgcgggg cataaagttg cgagtgatgc gaaaggttg gggcacctcg gcccggttgt	7380
	tgatgacctg ggcggcgagc acgatctcgt cgaagccgtt gatgttggg cccacgatgt	7440
	agagtccac gaatcgcgga cggcccttga cgtggggcag tttcttgagc tcctcgtagg	7500
[0006]	tgagctcgtc ggggtcgtc agccccgtct gctcgcgcgc ccagtcggcg agatgggggt	7560
	tggcgcggag gaaggaagtc cagagatcca cggccagggc ggtttgcaga cggctccggt	7620
	actgacggaa ctgctgcccc acggccattt tttcgggggt gacgcagtag aaggtgcggg	7680
	ggtccccgtg ccagcgatcc catttgagct ggagggcgag atcgaggcg agctcgacga	7740
	gccgctcgtc cccggagagt ttcatgacca gcatgaagg gacgagctgc ttgccgaagg	7800
	acccatcca ggtgtaggtt tccacatcgt aggtgaggaa gagccttctg gtgcgaggat	7860
	gcgagccgat ggggaagaac tggatctcct gccaccaatt ggaggaatgg ctgttgatgt	7920
	gatggaagta gaaatgccga cggcgcgccg aacctcgtg cttgtgttta tacaagcggc	7980
	cacagtctc gcaacgctgc acgggatgca cgtgctgcac gagctgtacc tgagttcctt	8040
	tgacgaggaa tttcagtggg aagtggagtc gtggcgcctg catctcgtgc tgtactacgt	8100
	cgtggtggtc ggccctggccc tcttctgcct cgatggtggt catgctgacg agcccgcgcg	8160
	ggaggcaggt ccagacctc gcgcgagcgg gtcggagagc gaggacgagg gcgcgcagc	8220
	cggagctgct cagggtcctg agacgctgc gagtcaggtc agtgggcagc ggcggcgcgc	8280

	ggttgacttg caggagtttt tccagggcgc gcgggaggtc cagatggtac ttgatctcca	8340
	ccgcgccatt ggtggcgacg tcgatggctt gcaggtccc gtgccctgg ggtgtgacca	8400
	ccgtcccccg tttcttcttg ggcggtctgg gcgacggggg cgggtcctct tccatggtta	8460
	gaagcggcgg cgaggacgcg cgccgggcgg caggggcggc tcggggcccg gaggcagggg	8520
	cggcaggggc acgtcggcgc cgcgcgcggg taggttcttg tactgcgcc ggagaagact	8580
	ggcgtgagcg acgacgcgac ggttgacgtc ctggatctga cgcctctggg tgaaggccac	8640
	gggaccctg agtttgaacc tgaaagagag ttcgacagaa tcaatctcgg tatcgttgac	8700
	ggcggcctgc cgcaggatct cttgcacgtc gcccagttg tcttggtagg cgatctcgg	8760
	catgaactgc tcgatctcct cctcttgaag gtctccgcg cgggcgct ccacggtggc	8820
	cgcgaggtcg ttggagatgc ggcccatgag ctgcgagaag gcgttcatgc ccgcctcgtt	8880
	ccagacgagg ctgtagacca cgacgccctc gggatcgcgg gcgcgcatga ccacctgggc	8940
	gaggttgagc tccacgtggc gcgtgaagac cgcgtagttg cagaggcgt ggtagagta	9000
[0007]	gttgagcgtg gtggcagatg gctcggtagc gaagaaatac atgatccagc ggcggagcgg	9060
	catctcgtg acgtcgccea gcgcctccaa acgttccatg gcctcgtaaa agtccacggc	9120
	gaagttgaaa aactgggagt tgcgcgccga gacggtcaac tcctcctcca gaagacggat	9180
	gagctcggcg atggtggcgc gcacctcgc ctcgaaggcc ccgggagtt cctccactc	9240
	ctcttcttc tctccacta acatctcttc tacttctcctc tcaggcggca gtggtggcgg	9300
	gggagggggc ctgcgtgcc ggcgcgcac gggcagacgg tcgatgaagc gctcagatgt	9360
	ctcggcgc cggcgtcga tggctcggg gacggcgcgc ccgtcctcgc ggggccgag	9420
	cgtgaagacg ccgccgcga tctccagtg gccggggggg tcccgttg gcaggagag	9480
	ggcgtgacg atgcatctta tcaattgcc cgtagggact ccgcgcaagg acctgagcgt	9540
	ctcagatcc acgggatctg aaaaccgtg aacgaaggct tcgagccagt cgcagtcgca	9600
	aggtaggctg agcacggttt cttctggcgg gtcattgttg ttgggagcgg ggcggcgat	9660
	gctgctggtg atgaagtga aataggcgg tctgagacgg cggatggtgg cgaggagcac	9720
	caggtctttg ggccccgctt gctggatcgc cagacggtcg gccatgcccc aggcgtggtc	9780
	ctgacacctg gccaggtcct tgtagtagtc ctgcatgagc cgctccacgg gcacctctc	9840

	ctcgcccgcg cggccgtgca tgcgcgtgag cccgaagccg cgctggggct ggacgagcgc	9900
	caggtcggcg acgacgcgct cggcgaggat ggcttgctgg atctgggtga ggggtgtctg	9960
	gaagtcacga aagtcgacga agcgggtgta ggctccggtg ttgatggtgt aggagcagtt	10020
	ggccatgacg gaccagttga cggctcggtg gcccgacgc acgagctcgt ggtacttgag	10080
	gcgcgagtag gcgcgcgtgt cgaagatgta gtcgttcgag gtgcgcacca ggtactgta	10140
	gccgatgagg aagtgcggcg gcggctggcg gtagagcggc catcgctcgg tggcgggggc	10200
	gccggcgcg aggtcctcga gcatggtgcg gtggtagccg tagatgtacc tggacatcca	10260
	ggtgatgccg gcggcgggtg tggaggcgcg cgggaactcg cggacgcggt tccagatgtt	10320
	gcgcagcggc aggaagtagt tcatggtggg cacggtctgg cccgtgaggc gcgcgcagtc	10380
	gtggatgctc tatacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag	10440
	gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag	10500
	ccgcagctaa cgtggtattg gcactcccgt ctcgacccaa gcctgcacca accctccagg	10560
[0008]	atacggaggc gggctgtttt gcaacttttt tttggaggcc ggatgagact agtaagcgcg	10620
	gaaagcggcc gaccgcgatg gctcgtgccc gtagtctgga gaagaatcgc cagggttgcg	10680
	ttgcggtgtg ccccggttcg aggccggccc gattccgcgg ctaacgaggc cgtggctgcc	10740
	cegtcgtttc caagacccca tagccagccc acttctccag ttacggagcg agcccctctt	10800
	ttgtttgttt tgtttttgcc agatgcatcc cgtactgcgg cagatgcgcc cccaccacce	10860
	tccaccgcaa caacagcccc ctccacagcc ggcgcttctg ccccccccc agcagcaact	10920
	tccagccacg accgccgcgg ccgccgtgag cggggctgga cagagtatg atcaccagct	10980
	ggccttgaa gagggcgagg ggctggcgcg cctggggggcg tcgtcggcgg agcggcaccc	11040
	gcgcgtgcag atgaaaagg acgctcgcga ggcctacgtg cccaagcaga acctgttcag	11100
	agacaggagc ggcgaggagc ccgaggagat gcgcgcggcc cggttccacg cggggcggga	11160
	gctgcggcgc ggcctggacc gaaagagggt gctgagggac gaggatttcg aggcggacga	11220
	gctgacgggg atcagccccg cgcgcgcgca cgtggccgcg gccaacctgg tcacggcgta	11280
	cgagcagacc gtgaaggagg agagcaactt ccaaaaatcc ttcaacaacc acgtgcgcac	11340

	cctgatcgcg cgcgaggagg tgaccctggg cctgatgcac ctgtgggacc tgctggagge	11400
	catcgtgcag aaccccacca gcaagccgct gacggcgag ctgttcctgg tgggtgcagca	11460
	tagtcgggac aacgaagcgt tcagggagge gctgctgaat atcaccgagc ccgagggccg	11520
	ctggctcctg gacctgggta acattctgca gagcatcgtg gtgcaggagc gcgggctgcc	11580
	gctgtccgag aagctggcgg ccatcaactt ctcgggtctg agtttgggca agtactacgc	11640
	taggaagatc tacaagacce cgtacgtgcc catagacaag gaggtgaaga tcgacgggtt	11700
	ttacatgcgc atgacctga aagtgtgac cctgagcgac gatctggggg tgtaccgcaa	11760
	cgacaggatg caccgtgcgg tgagcggcag caggcggcgc gagctgagcg accaggagct	11820
	gatgcatagt ctgcagcggg ccctgaccgg ggcccggacc gagggggaga gctactttga	11880
	catgggcgcg gacctgcaact ggcagcccag ccgccgggccc ttggaggcgg cggcaggacc	11940
	ctacgtagaa gaggtggacg atgaggtgga cgaggagggc gagtacctgg aagactgatg	12000
	gcgcgaccgt atttttgcta gatgcaacaa caacagccac ctctgatcc cgcgatgcgg	12060
	gcggcgctgc agagccagcc gtccggcatt aactcctcg acgattggac ccaggccatg	12120
[0009]	caacgcatca tggcgtgac gacctgcaac ccggaagcct ttagacagca gccccaggcc	12180
	aaccggctct cggccatcct ggaggccgtg gtgccctcgc gctccaacce cacgcacgag	12240
	aaggctctgg ccatcgtgaa cgcgctgggtg gagaacaagg ccatccgagg cgacgaggcc	12300
	ggcctggtgt acaacgcgct gctggagcgc gtggcccgt acaacagcac caacgtgcag	12360
	accaacctgg accgcatggt gaccgacgtg cgcgaggccg tggcccagcg cgagcggttc	12420
	caccgcgagt ccaacctggg atccatggtg gcgctgaacg ccttcctcag caccagccc	12480
	gccaacgtgc cccggggcca ggaggactac accaacttca tcagcgcctt gcgcctgatg	12540
	gtgaccgagg tgccccagag cgaggtgtac cagtccgggc cggactactt cttccagacc	12600
	agtcgccagg gcttgcagac cgtgaacctg agccaggctt tcaagaactt gcagggcctg	12660
	tggggcgtgc aggccccggt cggggaccgc gcgacgggtg cgagcctgct gacgccgaac	12720
	tcgcccctgc tgctgctgct ggtggcccc ttcacggaca gcggcagcat caaccgcaac	12780
	tcgtacctgg gctacctgat taacctgtac cgcgaggcca tcggccagge gcacgtggac	12840
	gagcagacct accaggagat caccacgtg agccgcgccc tgggccagga cgaccgggc	12900

	aacctggaag ccaccctgaa ctttttgctg accaaccggt cgcagaagat cccgccccag	12960
	tacgcgctca gcaccgagga ggagcgcate ctgcgttacg tgcagcagag cgtgggcctg	13020
	ttcctgatgc aggagggggc cacccccagc gccgcgctcg acatgaccgc gcgcaacatg	13080
	gagcccagca tgtacgccag caaccgcccg ttcataata aactgatgga ctacttgcac	13140
	cgggcggccg ccatgaactc tgactatrtc accaacgcca tcctgaatcc ccaactggctc	13200
	ccgccgccgg ggttctacac gggcgagtac gacatgcccg accccaatga cgggttcctg	13260
	tgggacgatg tggacagcag cgtgttctcc ccccgaccgg gtgctaacga gcgcccttg	13320
	tggaagaagg aaggcagcga ccgaccccc tcctcggcgc tgtccggccg cgagggtgct	13380
	gccgcggcgg tgcccagagc ccccagtcct ttcccagact tgcccttctc gctgaacagt	13440
	atccgcagca gcgagctggg caggatcacg cccccgcgt tgctgggcga agaggagtac	13500
	ttgaatgact cgctgttgag acccgagcgg gagaagaact tcccataa cgggatagaa	13560
	agcctggtgg acaagatgag ccgctggaag acgtatgccc aggagcacag ggacgatccc	13620
[0010]	cgggcgtcgc agggggccac gagccggggc agcgcgccc gtaaaccgcg gtggcacgac	13680
	aggcagcggg gacagatgtg ggacgatgag gactccgccg acgacagcag cgtgttgac	13740
	ttgggtggga gtggtaacc gttcgtcac ctgcgcccc gtatcgggcg catgatgtaa	13800
	gagaaaaccg aaataaatga tactcaccaa ggccatggcg accagcgtgc gttcgtttct	13860
	tctctgttgt tgttgtatct agtatgatga ggcgtgcgta cccggagggt cctcctcct	13920
	cgtagcagag cgtgatgcag caggcagatg cggcggcggc gatgcagccc ccgctggagg	13980
	ctccttacgt gccccgcgg tacctggcgc ctacggaggg gcggaacagc attcgttact	14040
	cggagctggc acccttgtag gataccacc gttgtacct ggtggacaac aagtcggcgg	14100
	acatgcctc gctgaactac cagaacgacc acagcaact cctgaccacc gtggtgcaga	14160
	acaatgactt cacccccag gagccagca cccagaccat caactttgac gacgctcgc	14220
	ggtggggcgg ccagctgaaa accatcatgc acaccaacat gcccaacgtg aacgagttca	14280
	tgtacagcaa caagttcaag gcgcgggtga tggctctccg caagaccccc aatgggggtga	14340
	cagtacaga ggattatgat ggtagtcagg atgagctgaa gtatgaatgg gtggaatttg	14400

	agctgcccga aggcaacttc tcggtgacca tgaccatoga cctgatgaac aacgccatca	14460
	tcgacaatta cttggcgggtg gggcggcaga acggggtgct ggagagcgac atcggcgtga	14520
	agttcgacac taggaacttc aggctgggct gggaccccgt gaccgagctg gtcgatgcccg	14580
	gggtgtacac caacgaggct ttccatcccg atattgtctt gctgcccggc tgcgggggtgg	14640
	acttcaccga gagccgctc agcaacctgc tgggcattcg caagaggcag cccttcagg	14700
	aaggettcca gatcatgtac gaggatctgg aggggggcaa catccccgcg ctctggatg	14760
	tcgacgccta tgagaaaagc aaggaggatg cagcagctga agcaactgca gccgtagcta	14820
	ccgcctctac cgaggtcagg ggcgataatt ttgcaagcgc cgcagcagtg gcagcggccg	14880
	aggcggctga aaccgaaagt aagatagtca ttcagccggt ggagaaggat agcaagaaca	14940
	ggagctacaa cgtactaccg gacaagataa acaccgccta ccgcagctgg tacctagcct	15000
	acaactatgg cgaccccag aagggcgtgc gctcctggac getgctcacc acctcggacg	15060
	tcacctgceg cgtggagcaa gtctactggt cgctgcccga catgatgcaa gacccggtca	15120
	ccttcgctc cacgcgtcaa gttagcaact acccgggtgt gggcggcag ctctgcccg	15180
[0011]	tctactcaa gagcttcttc aacgagcagg ccgtctactc gcagcagctg cgcgcttca	15240
	cctcgcttac gcacgtcttc aaccgcttc ccgagaacca gatcctcgtc cgcccgcgg	15300
	cgcccacat taccaccgtc agtgaaaacg ttctgtctt cacagatcac gggaccctgc	15360
	cgctgcgcag cagtatccgg ggagtccagc gcgtgaccgt tactgacgcc agacgccga	15420
	cctgccccta cgtctacaag gccctgggca tagtegcgcc gcgctcctc tcgagccga	15480
	ccttctaaat gtccattctc atctcggcca gtaataacac cggttggggc ctgcgcgdc	15540
	ccagcaagat gtacggaggc gctcgccaac gctccacgca acaccccgtg cgcgtgcgcg	15600
	ggcacttccg cgctccctgg ggcgccctca agggccgct gcggtcgcgc accaccgtc	15660
	acgacgtgat cgaccagggt gtggccgacg cgcgcaacta ccccccgcc gccgcgccc	15720
	tctccaccgt ggacgccgtc atcgacagcg tgggtggcca cgcgcgcccg tacgcccgcg	15780
	ccaagagccg gcggcggcgc atcgcccggc ggcaccggag ccccccgcc atgcgcgcg	15840
	cgcgagcctt gctgcgcagg gccaggcgca cgggacgcag ggccatgctc agggcggcca	15900
	gacgcgcccg ttcaggcgcc agcggccgca ggacccggag acgcgcggcc acggcggcgg	15960

	cagcggccat cgccagcatg tcccggccgc ggcgagggaa cgtgtactgg gtgcgcgacg	16020
	ccgccaccgg tgtgcgcgtg cccgtgcgca cccgcccccc tgcacttga agatgttcac	16080
	ttcgcgatgt tgatgtgtcc cagcggcgag gaggatgtcc aagcgcaaat tcaaggaaga	16140
	gatgtccag gtcacgcgc ctgagatcta cgccctgcg gtggtgaagg aggaaagaaa	16200
	gccccgcaaa atcaagcggg tcaaaaagga caaaaaggaa gaagaaagtg atgtggacgg	16260
	attggtggag tttgtgcgag agttcggccc ccggcggcgc gtgcagtggc gcgggcggaa	16320
	ggtgcaaccg gtgctgagac ccggcaccac cgtggtcttc acgcccggcg agcgtccgg	16380
	caccgttcc aagcgtcct acgacgaggt gtacgggat gatgatattc tggagcaggc	16440
	ggccgagcgc ctggcgagt ttgcttacgg caagcgcagc cgttccgac cgaaggaaga	16500
	ggcgggtgcc atcccgtgg accacggcaa cccacgccg agcctcaagc ccgtgacctt	16560
	gcagcaggtg ctgccaccg cggcggcgc ccgggggttc aagcgcgagg gcgaggatct	16620
	gtaccccacc atgcagctga tgggtcccaa gcgccagaag ctggaagacg tgctggagac	16680
[0012]	catgaaggtg gaccggcag tgcagcccga ggtcaaggtg cggcccatca agcaggtggc	16740
	cccggcctg ggcgtcaga ccgtggacat caagattccc acggagccca tggaaacgca	16800
	gaccgagccc atgatcaagc ccagcaccag caccatggag gtgcagacgg atccctggat	16860
	gccatcgct cctagtcgaa gacccggcg caagtacggc gcggccagcc tgcctgatcc	16920
	caactacgcg ctgcatcctt ccatcatccc cagccgggc taccggca cgcgcttcta	16980
	ccgggtcat accagcagcc gccggcga gaccaccaact cggccgcc gtcgccgac	17040
	cgccgtgca accaccctg ccgccctggt gcggagagt taccggcg gccgcgacc	17100
	tctgacctg ccgcgcgc gctaccacc gagcatgcc atttaaactt tcgctgctt	17160
	tgcagatcaa tgccctcac atgccctt cgcgttccca ttacgggcta ccgaggaaga	17220
	aaaccgccc gtagaaggct ggccgggaa gggatgcgtc gccaccacca ccggcggcg	17280
	cgcccatca gcaagcggtt ggggggaggc ttctgccc cgctgatccc catcatgcc	17340
	gcggcgatcg gggcgatccc cggcattgct tccgtggcgg tgcaggctc tcagcggcac	17400
	tgagacacac ttgaaacat cttgtaataa accaatggac tctgacgctc ctggtcctgt	17460

	gatgtgtttt cgtagacaga tggaagacat caatTTTTcg tccttgctc cgcgacacgg	17520
	cacgcggccg ttcatgggca cctggagcga catcggcacc agccaactga acgggggccc	17580
	cttcaattgg agcagtctct ggagcgggct taagaatttc gggTccacgc ttaaaaccta	17640
	tggcagcaag gcgtggaaca gcaccacagg gcaggcctg agggataagc tgaagagca	17700
	gaacttccag cagaaggTgg tcgatgggct cgcctcgggc atcaacgggg TggTggacct	17760
	ggccaaccag gccTgcagc ggCagatcaa cagccgctg gaccCGgtgc cgcCCcgg	17820
	ctccgtggag atgccgagg Tggaggagga gctgcctccc ctggacaagc gggcgagaa	17880
	gcgacccccc cccgatcgg aggagacgct gctgacgac acggacgagc cgcCCcgtA	17940
	cgaggaggcg gtgaaactgg gtctgcccac cacgcggccc atcgcgccc Tggccaccg	18000
	ggtgctgaaa cccgaaaagc ccgcgacct ggactTgct cctccccagc ctccccccc	18060
	ctctacagtg gctaagcccc Tgccgccggt ggccgtggcc cgcgcgcgac cgggggac	18120
	cgcCCcct catgcgaact ggCagagcac tctgaacagc atcgtgggTc TgggagtGca	18180
[0013]	gagtgtgaag cgcCCcct gctattaaac ctaccgtagc gcttaactg cttgtctgtg	18240
	Tgtgtatgta ttatgtgcc gccgCCctg tccaccagaa ggaggagtga agaggcgcT	18300
	cgcCCgattg caagatggcc accccatcga Tgctgcccc gtggcgTac atgcacatc	18360
	ccggacagga cgctcggag tacctgagtc cgggtctggt gcagttTgc cgcCCcag	18420
	acacctactt cagtctggg aacaagtTta ggaacccac gTggcgccc acgcacgat	18480
	Tgaccaccga ccgcagccag cggctgacgc Tgcgcttctg gcccgTggac cgcgaggaca	18540
	acacctactc gtacaaagt cgctacacgc Tggccgtggg cgacaaccgc gtctggaca	18600
	Tggccagcac ctactTgac atccgCGcg Tgctggatc gggccctagc ttcaaacct	18660
	actcggcac cgcctacaac agtctggccc ccaaggagc acccaact TgtcagtGga	18720
	catataaagc cgatggTgaa actgccacag aaaaaaccta tacatatgga aatgcacccg	18780
	Tgcaggcat taacatcaca aaagatgTta ttcaactTg aactgacacc gatgatcagc	18840
	caatctacgc agataaaacc tatcagcctg aacctcaagt gggTgatgct gaatggcat	18900
	acatcactgg tactgatgaa aagtatggag gcagagctct taagcctgat accaaaatga	18960
	agcctTgTta Tggttctttt gccaaGccta ctaataaaga aggaggtcag gcaaatgtga	19020

	aaacaggaac aggcactact aaagaatatg acatagacat ggctttcttt gacaacagaa	19080
	gtgcggctgc tgctggccta gctccagaaa ttgttttgta tactgaaaat gtggatttgg	19140
	aaactccaga taccatatt gtatacaaag caggcacaga tgacagcagc tcttctatta	19200
	at ttgggtca gcaagccatg cccaacagac ctaactacat tggtttcaga gacaacttta	19260
	tcgggctcat gtactacaac agcactggca atatgggggt gctggccggt caggcttctc	19320
	agctgaatgc tgtggttgac ttgcaagaca gaaacaccga gctgtcctac cagctcttgc	19380
	ttgactctct ggggtacaga acccggtatt tcagtatgtg gaatcaggcg gtggacagct	19440
	atgatcctga tgtgcgcatt attgaaaatc atgggtgtgga ggatgaactt cccaactatt	19500
	gtttccctct ggatgctgtt ggcagaacag atacttatca gggaattaag gctaatggaa	19560
	ctgatcaaac cacatggacc aaagatgaca gtgtcaatga tgctaataag ataggcaagg	19620
	gtaatccatt cgccatggaa atcaacatcc aagccaacct gtggaggaaac ttcctctacg	19680
	ccaacgtggc cctgtacctg cccgactctt acaagtacac gccggccaat gtaccctgc	19740
[0014]	ccaccaacac caacacctac gattacatga acggccgggt ggtggcgcgc tcgctggtgg	19800
	actcctacat caacatcggg gcgcgctggt cgctggatcc catggacaac gtgaaccctc	19860
	tcaaccacca ccgcaatgcg gggctgcgct accgctccat gctcctgggc aacgggcgct	19920
	acgtgccctt ccacatccag gtgccccaga aatTTTTgc catcaagagc ctctgctcc	19980
	tgccccggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga	20040
	gctcctcgg caacgacctg cgcacggacg gggcctccat ctcttcacc agcatcaacc	20100
	tctacgccac cttcttcccc atggcgcaaca acacggcctc cacgctcgag gccatgctgc	20160
	gcaacgacac caacgaccag tcttcaacg actacctctc ggcggccaac atgctctacc	20220
	ccatcccggc caacgccacc aacgtgcccc tctccatccc ctgcgcaac tgggccgct	20280
	tccgcgctg gtccttcacg cgtctcaaga ccaaggagac gcctcctg ggctccgggt	20340
	tcgaccccta cttegtctac tcgggctcca tcccctacct cgacggcacc ttctacctca	20400
	accacacctt caagaaggtc tccatcacct tcgactcctc cgtcagctgg cccggcaacg	20460
	accgctcct gacgccccaac gagttcgaaa tcaagegcac cgtcgcaggc gagggctaca	20520

	acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgctg gccactaca	20580
	acatcggcta ccagggttc tacgtgcccg agggctacaa ggaccgcatg tactccttct	20640
	tccgcaactt ccageccatg agccgccagg tggtaggacga ggtcaactac aaggactacc	20700
	aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctgcgcccc	20760
	ccatgcgcca gggccagccc taccgcgcca actacccta cccgctcacc ggcaagagcg	20820
	ccgtcaccag cgtcaccag aaaaagtcc tctgcgacag ggtcatgtgg cgcacccct	20880
	tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgtctatg	20940
	ccaactccgc ccacgcgcta gacatgaatt tcgaagtca ccccatggat gagtccacc	21000
	ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg	21060
	tcacgagge cgtctacctg cgcacccct tctcggccgg taacgccacc acctaagtc	21120
	ttgttcttg caagccatgg ccgcgggctc cggcgagcag gagctcagg ccatcatccg	21180
	cgacctgggc tgcgggacct acttctggg caccttcgat aagcgttcc cgggattcat	21240
[0015]	ggccccgcac aagctggcct gcgccatcgt caacacggcc ggccgcgaga ccggggcgca	21300
	gcactggctg gccttcgect ggaacccgag ctcgaaacacc tgetacctct tcgacctt	21360
	cgggttctcg gacgagcgc tcaagcagat ctaccagttc gactacgagg gcctgctgcg	21420
	ccgcagcgc ctggccaccg aggaccgctg cgtcacctg gaaaagtcca ccagaccgt	21480
	gcagggtccg cgtcggccg cctcgggct cttctgctgc atgttctgc acgcttctg	21540
	gcactggccc gaccgcccc tggacaagaa cccaccatg aacttgctga cgggggtgcc	21600
	caacggcatg ctccagtcgc cccaggtgga acccacctg ccccgcaacc aggagcgcgt	21660
	ctaccgcttc ctcaactccc actccgcta cttctgctcc caccgcgcgc gcatcgagaa	21720
	ggccaccgccc ttcgaccgca tgaatcaaga catgtaaacc gtgtgtgtat gttaaatgtc	21780
	tttaataaac agcacttca tgttacacat gcatctgaga tgatttattt agaaatcgaa	21840
	agggttctgc cgggtctcgg catggcccgc gggcaggac acgttgcgga actggtactt	21900
	ggccagccac ttgaactcgg ggatcagcag tttgggcagc ggggtgtcgg ggaaggagtc	21960
	ggtccacagc ttcccgctca gttgcagggc gccagcagg tcgggcgcgg agatcttgaa	22020
	atcgcagttg ggaccgcgt tctgcgcgcg ggagttcgg tacacggggt tgcagcactg	22080

	gaacaccatc agggccgggt gcttcacgct cgccagcacc gtcgcgctcg tgatgctctc	22140
	cacgtcgagg tcctcggcgt tggccatccc gaagggggtc atcttgagg tctgccttcc	22200
	catggtgggc acgcacccgg gcttgtggtt gcaatcgcag tgcaggggga tcagcatcat	22260
	ctgggccttg tcggcgttca tccccggta catggccttc atgaaagcct ccaattgcct	22320
	gaacgcctgc tgggccttgg ctccctcggt gaagaagacc ccgcaggact tgctagagaa	22380
	ctggttggtg gcgcacccgg cgtcgtgcac gcagcagcgc gcgtcgttgt tggccagctg	22440
	caccacgctg cccccacg ggttctgggt gatcttggcc cggtcggggt tctccttcag	22500
	cgcgcgctgc ccgttctcgc tcgccacatc catctcgatc atgtgctcct tctggatcat	22560
	ggtggtcccg tgcaggcacc gcagcttgcc ctccggcctc gtgcaccctg gcagccacag	22620
	cgcgaccccg gtgcactccc agttcttgtg ggcgatctgg gaatgcgctg gcacgaagcc	22680
	ctgcaggaag cggcccatca tgggtggtcag ggtcttgttg ctagtgaagg tcagcggaat	22740
	gcccggtgct tcctcgttga tgtacaggtg gcagatgcgg cggtagacct cgccctgctc	22800
[0016]	gggcatcagc tggaagttgg ctttcaggtc ggtctccacg cggtagcggg ccatcagcat	22860
	agtcatgatt tccataccct tctcccaggc cgagacgatg ggcaggctca tagggttctt	22920
	caccatcadc ttagcgctag cagcccgggc cagggggctc ctctcgtcca ggtctcaaa	22980
	gctccgcttg ccgtccttct cggatgatcc caccgggggg tagctgaagc ccacggccgc	23040
	cagctcctcc tcggcctgct tttcgtctc gctgtcctgg ctgacgtcct gcaggaccac	23100
	atgcttggtc ttgcggggtt tcttcttggg cggcagcggc ggccgagatg ttggagatgg	23160
	cgagggggag cgcgagttct cgctcaccac tactatctct tcctcttctt ggtccgaggc	23220
	cacgcggcgg taggtatgct tcttcggggg cagaggcgga ggcgacgggc tctcgcgcc	23280
	gcgacttggc ggatggctgg cagagcccct tccgcgttcg ggggtgcgct cccggcggcg	23340
	ctctgactga cttcctccgc ggccggccat tgtgttctcc tagggaggaa caacaagcat	23400
	ggagactcag ccatcgccaa cctcgccatc tgccccacc gccgacgaga agcagcagca	23460
	gcagaatgaa agcttaaccg ccccggcgc cagccccgc acctccgacg cggccgtccc	23520
	agacatgcaa gagatggagg aatccatcga gattgacctg ggctatgtga cgccccgga	23580

	gcacgaggag gagctggcag tgcgcttttc acaagaagag atacaccaag aacagccaga	23640
	gcaggaagca gagaatgagc agagtcaggc tgggctcgag catgacggcg actacctcca	23700
	cctgagcggg ggggaggacg cgctcatcaa gcctctggcc cggcaggcca ccatcgtcaa	23760
	ggatgcgctg ctcgaccgca ccgaggtgcc cctcagcgtg gaggagctca gccgcgccta	23820
	cgagttgaac ctcttctcgc cgcgcgtgcc cccaagcgc cagcccaatg gcacctgca	23880
	gccaaccg cgctcaact tctaccggc cttcgcggcg cccgaggccc tgccacctta	23940
	ccacatcttt ttcaagaacc aaaagatccc cgtctctcgc cgcgccaacc gcacccgcgc	24000
	cgacgcctt ttcaacctgg gtcccggcgc ccgcctacct gatatgcct ccttgaaga	24060
	ggttccaag atcttcgagg gtctgggcag cgacgagact cgggccgca acgctctgca	24120
	aggagaagga ggagagcatg agcaccacag cgccctggc gagttggaag gcgacaacgc	24180
	gcggctggcg gtgtcaaac gcacggcga gctgacctt ttcgcctacc cggctctgaa	24240
	cctgcccc aaagtcata gcgcggcct ggaccaggtg ctcatcaagc gcgcgtcgcc	24300
	catctccgag gacgaggca tgcaagactc cgaggaggc aagcccgtg tcagcgacga	24360
[0017]	gcagctggcc cgggtggtg gtctaatgc tagtcccag agtttgaag agcggcgcaa	24420
	actcatgat gccgtggtc tggtgaccgt ggagctggag tgcctgcgcc gcttctcgc	24480
	cgacgcggag accctgcga aggtcgagga gaacctgcac tacctctca ggcacgggtt	24540
	cgtgcgccag gcctgcaaga tctccaacgt ggagctgacc aacctggtct cctacatggg	24600
	catcttgca gagaaccgcc tggggcagaa cgtgctgca accacctgc gcggggaggc	24660
	ccggcgcgac tacatccgag actgcgtcta cctctacct tgccacacct ggcagacggg	24720
	catggcgctg tggcagcagt gtctggagga gcagaacctg aaagagctct gcaagctcct	24780
	gcagaagaac ctcaagggtc tgtggaccg gttcgacgag cgcaccaccg cctcggacct	24840
	ggccgacctc attttcccc agcgcctcag gctgacgctg cgcaacggcc tgcccactt	24900
	tatgagccaa agcatgttg aaaactttcg ctctttcacc ctogaacgct ccggaatcct	24960
	gcccgccacc tgctccgagc tgccctcgga cttcgtgccg ctgacctcc gcgagtgcc	25020
	cccgcgctg tggagccact gctacctgct gcgcctggcc aactacctg cctaccactc	25080
	ggacgtgatc gaggacgca gcggcgaggg cctgctcgag tgccactgcc gctgcaacct	25140

	ctgcacgccg caccgctccc tggcctgcaa cccccagctg ctgagcgaga cccagatcat	25200
	cggcaccttc gagttgcaag ggcccagcga aggcgagggt tcagccgcca aggggggtct	25260
	gaaactcacc ccggggctgt ggacctcggc ctacttgcgc aagttcgtgc ccgaggacta	25320
	ccatcccttc gagatcaggt tctacgagga ccaateccat ccgccaagg ccgagctgtc	25380
	ggcctgcgtc atcaccagcagg gggcgatcct ggccaattg caagccatcc agaaatccc	25440
	ccaagaattc ttgctgaaaa agggccgagg ggtctacct gacccccaga ccggtgagga	25500
	gctcaacccc ggcttcccc aggatgcccc gaggaacaaa gaagctgaaa gtggagctgc	25560
	cgcccgtaga ggatttggag gaagactggg agaacagcag tcaggcagag gaggaggaga	25620
	tggaggaaga ctgggacagc actcaggcag aggaggacag cctgcaagac agtctggagg	25680
	aagacgagga ggaggcagag gaggaggtgg aagaagcagc cgccgccaga ccgtcgtcct	25740
	cggcggggga gaaagcaagc agcacggata ccatctccgc tccgggtcgg ggtcccgtc	25800
	gaccacacag tagatgggac gagaccggac gattcccga cccaccacc cagaccgta	25860
[0018]	agaaggagcg gcagggatac aagtctggc gggggcaca aaacgccatc gtctctgct	25920
	tgcaggcctg cgggggcaac atctcctca cccggcgcta cctgctcttc caccgctggg	25980
	tgaactttcc ccgcaacatc ttgcattact accgtcacct ccacagcccc tactacttcc	26040
	aagaagaggc agcagcagca gaaaaagacc agcagaaaac cagcagctag aaaatccaca	26100
	gcggcggcag caggtggact gaggatcgcg gcgaacgagc cggcgcaaac ccgggagctg	26160
	aggaaccgga tctttccac cctctatgcc atcttccagc agagtcgggg gcaggagcag	26220
	gaactgaaag tcaagaaccg ttctctgcgc tcgctcacc gcagttgtct gtatcacaag	26280
	agcgaagacc aacttcagcg cactctcgag gacgcccagg ctctcttcaa caagtactgc	26340
	gcgctcactc ttaaagagta gcccgcgcc gccagtcgc agaaaaagc gggaattacg	26400
	tcacctgtgc ccttcgccct agccgcctcc acccatcacc atgagcaaag agattcccac	26460
	gccttacatg tggagctacc agccccagat gggcctggcc gccggtgccg cccaggacta	26520
	ctccaccgc atgaattggc tcagcgcgg gcccgcatg atctcacggg tgaatgacat	26580
	ccgcgccac cgaaaccaga tactcctaga acagtcagc ctcaccgcca cgccccgaa	26640

	tcacctcaat cgcgtaatt ggccccccgc cctgggtgtac caggaaatc cccagcccac	26700
	gaccgtacta cttccgcgag acgcccaggc cgaagtccag ctgactaact caggtgtcca	26760
	gctggcgggc ggcgccaccc tgtgtcgtca ccgccccgct cagggtataa agcggctgg	26820
	gatccggggc agaggcacac agctcaacga cgagggtgtg agctcttcgc tgggtctgcg	26880
	acctgacgga gtcttccaac tcgccgatc ggggagatct tccttcacgc ctgctcaggc	26940
	cgtcctgact ttggagagtt cgtcctcgca gccccgctcg ggtggcatcg gcaacttcca	27000
	gttcgtggag gagttcactc cctcggctca cttcaacccc ttctccggct cccccggcca	27060
	ctaccgggac gagttcatcc cgaacttcga cgccatcagc gagtcgggtg acggctacga	27120
	ttgaatgtcc catggtggcg cagctgacct agctcggctt cgacacctgg accactgccg	27180
	ccgcttcgcg tgcctcgcgc gggatctcgc cgagtttgc tactttgagc tgcccaggga	27240
	gcaccctcag ggccccggcc acggagtgcg gatcgtcgtc gaagggggcc tcgactcca	27300
	cctgcttcgg atcttcagcc agcgtccgat cctggctcag cgcgagcaag gacagacct	27360
	tctgactctg tactgcatct gcaaccaccc cggcctgcat gaaagtctt gttgtctgct	27420
[0019]	gtgtactgag tataataaaa gctgagatca gcgactactc cggacttccg tgtgttctg	27480
	aatccatcaa ccagcttttg ttcttcaccg ggaacgagac cgagctccag ctccagtgt	27540
	agccccacaa gaagtacctc acctggctgt tccagggtc cccgatcgc gttgtcaacc	27600
	actgcgacaa cgacggagtc ctgctgagcg gccctgccaa ccttactttt tccaccgca	27660
	gaagcaagct ccagetcttc caacccttc tccccgggac ctatcagtgc gtctcgggac	27720
	cctgccatca caccttcac ctgatcccga ataccacagc gtcgctcccc gctactaaca	27780
	accaaactaa cctccacaa cgccaccgtc gcgaccttc tgaatctaat actaccacc	27840
	acaccggagg tgagctccga ggtcaacaa cctctgggat ttactacggc ccctgggagg	27900
	tggttgggtt aatagcgcta ggcctagtig cgggtgggct tttggttctc tgetacctat	27960
	acctcccttg ctgttcgtac ttagtggtgc tgtgttgctg gtttaagaaa tggggaagat	28020
	caccctagtg agctgcggtg cgctgggtgc ggtgttgctt tcgattgtgg gactgggcgg	28080
	tgcggctgta gtgaaggaga aggccgatcc ctgcttgcac ttcaatccca acaaatgcca	28140
	gctgagtttt cagccccgatg gcaatcggtg cgcggtactg atcaagtgcg gatgggaatg	28200

	cgagaacgtg agaatcgagt acaataacaa gactcggaac aatactctcg cgtccgtgtg	28260
	gcagcccggg gaccccgagt ggtacaccgt ctctgtcccc ggtgctgacg gctccccgcg	28320
	caccgtgaat aatactttca tttttgcgca catgtgacgac acggatcatgt ggatgagcaa	28380
	gcagtaacgat atgtggcccc ccacgaagga gaacatcgtg gtcttctcca tcgcttacag	28440
	cctgtgcacg gcgctaataca ccgctatcgt gtgcctgagc attcacatgc tcatcgctat	28500
	tcgccccaga aataatgccg aaaaagaaaa acagccataa cgTTTTTTTT cacacctttt	28560
	tcagaccatg gcctctgtta aatttttgc tttatttgc agtctcattg ccgtcattca	28620
	tggaatgagt aatgagaaaa ttactattta cactggcact aatcacacat tgaaaggtcc	28680
	agaaaaagcc acagaagttt catggattg ttattttaat gaatcagatg tatctactga	28740
	actctgtgga aacaataca aaaaaatga gagcattact ctcatcaagt ttcaatgtgg	28800
	atctgactta accctaatta acatcactag agactatgta ggtatgtatt atggaactac	28860
	agcaggcatt tcggacatgg aattttatca agtttctgtg tctgaacca ccacgcctag	28920
[0020]	aatgaccaca accacaaaa ctacacctgt taccactatg cagctcacta ccaataacat	28980
	ttttgccatg cgtcaaatgg tcaacaatag cactcaacce accccacca gtgaggaaat	29040
	tcccaaatcc atgattggca ttattgttgc tgtagtggtg tgcattgtga tcatgcctt	29100
	gtgcatggtg tactatgcct tctgtacag aaagcacaga ctgaacgaca agctggaaca	29160
	cttactaagt gttgaatttt aattttttag aaccatgaag atcctaggcc ttttaatttt	29220
	ttctatcatt acctctgtc tatgcaattc tgacaatgag gacgttactg tcgttgtegg	29280
	atcaaattat aactgaaag gtccagcgaa gggatgctt tcgtgggtatt gctatttttg	29340
	atctgacact acagaaactg aattatgcaa tcttaagaat ggcaaaattc aaaattctaa	29400
	aattaacaat tatatatgca atggfactga tctgatactc ctcaatatca cgaaatcata	29460
	tgctggcagt tacacctgcc ctggagatga tctgacagc atgatttttt acaaagtaac	29520
	tgttgtgat cccactactc cacciccacc caccacaact actcacacca cacacacaga	29580
	tcaaaccgca gcagaggagg cagcaaagtt agccttgacg gtccaagaca gttcatttgt	29640
	tggcattacc cctacacctg atcagcggtg tccggggctg ctagtcagcg gcattgtcgg	29700

	tgtgctttcg ggattagcag tcataatcat ctgcatgttc atttttgctt gctgctatag	29760
	aaggctttac cgacaaaaat cagaccact gctgaacctc tatgtttaat tttttcaga	29820
	gtcatgaagg cagttagegc tctagttttt tgttctttga ttggcattgt tttttgcaat	29880
	cctattccta aagttagctt tattaagat gtgaatgta ctgagggggg caatgtgaca	29940
	ctggtaggtg tagagggtgc tgaaaacacc acctggacaa aataccact caatgggtg	30000
	aaagatattt gcaattggag tgtattagtt tatacatgtg aggaggttaa tcttaccatt	30060
	gtcaatgcca cctcagctca aaatggtaga attcaaggac aaagtgtcag tgtatcta	30120
	gggtatttta cccaacatac ttttatctat gacgttaaag tcataaccact gcctacgct	30180
	agcccaccta gcactaccac acagacaacc cacactacac agacaaccac atacagtaca	30240
	ttaaatcagc ctaccaccac tacagcagca gaggttgcca gctcgtctgg ggtccgagtg	30300
	gcatttttga tgtgggcccc atctagcagt cccactgcta gtaccaatga gcagactact	30360
	gaatttttgt ccactgtcga gagccacacc acagctacct ccagtgcctt ctctagcacc	30420
	gccaatctct cctcgtttc ctctacacca atcagtcctg ctactactcc tagccccgt	30480
[0021]	cctcttcca ctcccctgaa gcaaacagac ggcgcatgc aatggcagat caccctgctc	30540
	atttgtatcg ggttggctcat cctggccgtg ttgctctact acatcttctg ccgccgatt	30600
	cccaacgcgc accgcaagcc ggtctacaag cccatcattg tcgggcagcc ggagccgctt	30660
	cagtggaag ggggtctaag gaatcttctc ttctctttta cagtatggtg attgaactat	30720
	gattcctaga caattcttga tcaactattct tatctgcctc ctccaagtct gtgccacct	30780
	cgctctggtg gccaacgcca gtccagactg tattgggccc ttgcctcctt acgtgctctt	30840
	tgccttcacc acctgcatct gctgctgtag catagtctgc ctgcttatca ccttcttcca	30900
	gttcattgac tggatctttg tgcgcatgc ctacctgcgc caccacccc agtaccgca	30960
	ccagcgagtg gcgcggctgc tcaggetcct ctgataagca tgcgggctct getacttctc	31020
	gcgcttctgc tgttagtgct cccccgtccc gtcgaccccc ggtccccac ccagtcccc	31080
	gaggaggtcc gcaaatgcaa attccaagaa ccctggaaat tctcaaatg ctaccgcaa	31140
	aaatcagaca tgcacccag ctggatcatg atcattggga tcgtgaacat tctggcctgc	31200
	accctcatct cctttgtgat ttaccctgc tttgactttg gttggaactc gccagaggcg	31260

	ctctatctcc cgctgaacc tgacacacca ccacagcaac ctcaggcaca cgactacca	31320
	ccactacagc ctaggccaca atacatgccc atattagact atgaggccga gccacagcga	31380
	cccatgctcc ccgctattag ttacttcaat ctaaccggcg gagatgactg acccactggc	31440
	caacaacaac gtcaacgacc ttctcctgga catggacggc cgcgcctcgg agcagcgact	31500
	cgcccaactt cgcattcgcc agcagcagga gagagccgtc aaggagctgc aggatgcggt	31560
	ggccatccac cagtgaaga gaggcatctt ctgcctggtg aaacaggcca agatctccta	31620
	cgaggtcact ccaaacgacc atgcctctc ctacgagctc ctgcagcagc gccagaagtt	31680
	cacctgcctg gtcggagtca accccatcgt catcaccag cagtctggcg ataccaagg	31740
	gtgcatccac tgctcctgcg actccccga ctgcgtccac actctgatca agaccctctg	31800
	cggcctccgc gacctctcc ccatgaacta atcaccctc tatccagtga aataaagatc	31860
	atattgatga tgattttaca gaaataaaaa ataatcattt gatttgaat aaagatacaa	31920
	tcatattgat gatttgagtt taacaaaaaa ataaagaatc acttacttga aatctgatac	31980
[0022]	caggctctg tccatgtttt ctgccaacac cacttcactc cctcttccc agctctggta	32040
	ctgcaggccc cggcgggctg caaacttctc ccacacgctg aaggggatgt caaattcctc	32100
	ctgtccctca atcttcattt tatcttctat cagatgtcca aaaagcgcgt ccgggtgat	32160
	gatgactteg acccgtcta ccctacgat gcagacaacg caccgaccgt gcccttcatc	32220
	aacccccct tcgtctcttc agatggatc caagagaagc ccctgggggt gttgtccctg	32280
	cgaetggccg acccgtcac caccaagaac ggggaaatca ccctcaagct gggagagggg	32340
	gtggacctcg attcctcggg aaaactcatc tccaacacgg ccaccaaggc cgccgccct	32400
	ctcagttttt ccaacaacac catttccctt aacatggate accccttta cactaaagat	32460
	ggaaaattat ccttacaagt ttctccacca ttaaataac tgagaacaag cattctaac	32520
	acactagctt taggttttgg atcaggttta ggactccgtg gctctgcctt ggcagtacag	32580
	ttagctctc cacttacatt tgatactgat gaaacataa agcttacctt agacagaggt	32640
	ttgcatgtta caacaggaga tgcaattgaa agcaacataa gctgggctaa aggtttaaaa	32700
	tttgaagatg gagccatagc aaccaacatt ggaaatgggt tagagtttgg aagcagtagt	32760

	acagaaacag gtgttgatga tgcttaccca atccaagtta aacttggatc tggccttagc	32820
	tttgacagta caggagccat aatggctggt aacaaagaag acgataaact cactttgttg	32880
	acaacacctg atccatcacc aaactgtcaa atactcgcag aaaatgatgc aaaactaaca	32940
	ctttgcttga ctaaagtgg tagtcaaata ctggccactg tgtcagtctt agttgtagga	33000
	agtggaacc taaaccccat tactggcacc gtaagcagtg ctgaggtgtt tctacgtttt	33060
	gatgcaaacg gtgttctttt aacagaacat tctacactaa aaaaatactg ggggtatagg	33120
	caggagata gcatagatgg cactccatat accaatgctg taggattcat gcccaattta	33180
	aaagcttata caaagtcaca aagttctact actaaaaata atatagtagg gcaagtatac	33240
	atgaatggag atgtttcaaa acctatgctt ctactataa ccctcaatgg tactgatgac	33300
	agcaacagta catattcaat gtcattttca tacacctgga ctaatggaag ctatgttga	33360
	gcaacatttg gggctaactc ttataccttc tcatacatcg cccaagaatg aacctgtat	33420
	cccacctgc atgccaacce ttcccacccc actctgtgga acaaactctg aaacacaaaa	33480
	taaaataaag ttcaagtgtt ttattgattc aacagtttta caggattcga gcagttattt	33540
[0023]	ttctccacc ctcccaggac atggaataca ccacctctc cccccgaca gccttgaaca	33600
	tctgaatgcc attggtgatg gacatgcttt tggctctccac gtccacaca gtttcagagc	33660
	gagccagtct cgggtcggtc agggagatga aaccctcgg gcactcccgc atctgcacct	33720
	cacagctcaa cagctgagga ttgtcctcgg tggctgggat cacggttata tggagaagc	33780
	agaagagcgg cgggtgggaat catagtccgc gaacgggac gcccgggtgt gtcgcatcag	33840
	gccccgcagc agtcgctgcc gccgccgctc cgtaagctg ctgctcagg ggtccgggtc	33900
	cagggactcc ctgagcatga tgcccacggc cctcagcacc agtcgtctgg tgcggcgggc	33960
	gcagcagcgc atgcggatct cgctcaggtc gctgcagtac gtgcaacaca gaaccaccag	34020
	gttgttcaac agtccatagt tcaacacgct ccagccgaaa ctcatcgcgg gaaggatgct	34080
	accacagtgg ccgtcgtacc agatcctcag gtaaatacaag tggtgcccc tccagaacac	34140
	gctgcccacg tacatgatct ccttgggcat gtggcggttc accacctccc ggtaccacat	34200
	cacctctggt ttgaacatgc agccccggat gatcctcggg aaccacaggg ccagcaccgc	34260
	ccccccgcc atgcagcgaa gagacccccg gtccccgcaa tggcaatgga ggaccaccgc	34320

	ctcgtacccg tggatcatct gggagctgaa caagtctatg ttggcacagc acaggcatat	34380
	gctcatgcat ctcttcagea ctctcaactc ctctgggggtc aaaaccatat cccagggcac	34440
	ggggaactct tgcaggacag cgaacccccgc agaacagggc aatcctcgca cagaacttac	34500
	attgtgcatg gacagggtat cgcaatcagg cagcaccggg tgatcctcca ccagagaage	34560
	gcgggtctcg gtctctcac agcgtggtaa gggggccggc cgatacgggt gatggcggga	34620
	cgcggtgat cgtgttcgag accgtgtcat gatgcagttg ctttcggaca ttttcgtact	34680
	tgctgtagca gaacctggtc cgggcgctgc acaccgatcg ccggcggcgg tctcggcget	34740
	tggaacgctc ggtgttgaag ttgtaaaaca gccactctct cagaccgtgc agcagatcta	34800
	gggcctcagg agtgatgaag atcccatcat gcctgatggc tctgatcaca tcgaccaccg	34860
	tggaatgggc cagaccagc cagatgatgc aattttggtg ggtttcggtg acggcggggg	34920
	agggaagaac aggaagaacc atgattaact tttaatccaa acggtctcgg agtacttcaa	34980
	aatgaagatc gcggagatgg cacctctcgc ccccgctgtg ttggtggaag ataacagcca	35040
[0024]	ggtcaaaggt gatacggttc tcgagatggt ccacgggtgc ttccagcaaa gcctccacgc	35100
	gcacatccag aaacaagaca atagcgaagc cgggagggtt ctctaattcc tcaatcatca	35160
	tgttacactc ctgcaccatc cccagataat tttcattttt ccagccttga atgattcgaa	35220
	ctagttcctg aggtaaatcc aagccagcca tgataaagag ctgcgcaga gcgccctcca	35280
	ccggcattct taagcacacc ctcataatc caagatattc tgctcctggt tcacctgcag	35340
	cagattgaca agcgggaatat caaaatctct gccgcgatcc ctgagctcct cctcagcaa	35400
	taactgtaag tactctttca tatectctcc gaaattttta gccataggac caccaggaat	35460
	aagattaggg caagccacag tacagataaa ccgaagtcct cccagtgag cattgccaaa	35520
	tgcaagactg ctataagcat gctggctaga cccggtgata tcttccagat aactggacag	35580
	aaaatcggcc aggcaathtt taagaaaatc aacaaaagaa aaatcctcca ggtggacggt	35640
	tagagcctcg ggaacaacga tgaagtaaat gcaagcgggt cgttccagca tggtagtta	35700
	gctgatctgt agaaaaaaca aaaatgaaca ttaaaccatg ctagcctggc gaacaggtgg	35760
	gtaaactggt ctctccagca ccaggcagge cacgggtctt ccggcgcgac cctcgtaaaa	35820

attgtcgcta tgattgaaaa ccatcacaga gagacgttcc cggtagcccg cgtgaatgat	35880
tcgacaagat gaatacaccc ccggaacatt ggcgtccgag agtgaaaaaa agcgcccag	35940
gaagcaataa ggcaactaaa tgetcagtct caagtccagc aaagcgatgc catgcggatg	36000
aagcacaanaa ttctcaggtg cgtacanaat gtaattactc ccctcctgca caggcagcaa	36060
agccccgat ccctccaggt acacatacaa agcctcagcg tccatagctt accgagcagc	36120
agcacacaac aggcgcaaga gtcagagaaa ggctgagctc taacctgtcc acccgctctc	36180
tgetcaatat atagcccaga tctacactga cgtaaaggcc aaagtctaaa aatacccgc	36240
aaataatcac acagcccag cacacgcca gaaaccggtg acacactcaa aaaaatacgc	36300
gcacttctc aaacgcccnaa aactgcccgc atttccgggt tcccacgcta cgtcatcaaa	36360
acacgacttt caaatccgt cgaccgttaa aaacgtcacc cggcccgcct ctaacggtcg	36420
cccgctctc agccaatcag cggcccgat ccccaaatc aaacacctca tttgcatatt	36480
aacgcgcaca aaaagtttga ggtatattat tgatgatgg	36519

[0025]

<210> 2
 <211> 31588
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 2

ccatcttcaa taatatacct caaacctttt gtgcgcgtta atatgcaaat gaggcgtttg	60
aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgttagg ggcggggcga	120
gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag	180
tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttcccgc gctctctgac	240
aggaaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact	300
gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg caggaggag tatttgccga	360
gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa	420
tttccgcgta cgggtgcaaa gtccggtgtt tttacgtagg tgcagctga tcgccagggt	480
atttaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcagc aagagttttc	540

	tcctccgcgc cgcgagtcag atctacactt tgaagtagg gataacaggg taatgacatt	600
	gattattgac tagttgttaa tagtaatcaa ttacggggtc attagttcat agcccatata	660
	tggagtccg cgttacataa cttacggtaa atggcccgc tggctgaccg cccaacgacc	720
	cccgccatt gacgtcaata atgacgtatg ttcccatagt aacccaata gggactttcc	780
	attgacgtca atgggtggag tatttacggt aaactgceca cttggcagta catcaagtgt	840
	atcatatgcc aagtccgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt	900
	atgccagta catgacctta cgggactttc ctacttggca gtacatctac gtattagtca	960
	tcgctattac catggtgatg cggttttggc agtacaccaa tgggcgtgga tagcggtttg	1020
	actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc	1080
	aaaatcaacg ggactttcca aatgtcgtataaataccccgc cccgttgacg caaatgggcg	1140
	gtagcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg	1200
	cctggaacgc catccacgct gttttgacct ccatagaaga cagcgatcgc gccaccatgg	1260
[0026]	ccgggatggt ccaggcactg tccgaaggct gcacacceta tgatattaac cagatgctga	1320
	atgtcctggg agaccaccag gtctctggcc tggagcagct ggagagcadc atcaacttcg	1380
	agaagctgac cgagtggaca agctccaatg tgatgcctat cctgtcccca ctgaccaagg	1440
	geatcctggg cttegtgttt accctgacag tgccttctga gcggggcctg tcttgcata	1500
	gcgaggcaga cgcaaccaca ccagagtccg ccaatctggg cgaggagadc ctgtctcagc	1560
	tgtacctgtg gccccgggtg acatatact ccccttctta cgcctatcac cagtctgagc	1620
	ggagagccaa gtacaagaga cacttcccag gctttggcca gtctctgctg ttcggctacc	1680
	ccgtgtacgt gtteggcgat tgcgtgcagg gcgactggga tgccatccgg tttagatact	1740
	gcgcaccacc tggatatgca ctgctgaggt gtaacgacac caattattcc gcctgctgg	1800
	cagtgggcgc cctggagggc cctcgaate aggattggct gggcgtgcca aggcagctgg	1860
	tgacacgcat gcagccatc cagaacgcag gcctgtgcac cctggtggca atgctggagg	1920
	agacaatctt ctggctgcag gccttctga tggcctgac cgacagcggc cccaagacaa	1980
	acatcatcgt ggattcccag tacgtgatgg gcacttccaa gccttcttc caggagtttg	2040

	tggactggga gaacgtgagc ccagagctga attccaccga tcagccattc tggcaggcag	2100
	gaatcctggc aaggaacctg gtcctatgg tggccacagt gcagggccag aatctgaagt	2160
	accagggcca gagcctggtc atcagcgcct ccatcatcgt gtttaacctg ctggagctgg	2220
	agggcgacta tcgggacgat ggcaacgtgt ggggtgcacac cccactgagc cccagaacac	2280
	tgaacgcctg ggtgaaggcc gtggaggaga agaaggcat cccagtgcac ctggagctgg	2340
	cctccatgac caatatggag ctgatgtcta gcacgtgca ccagcagggtg aggacatacg	2400
	gaccctgttt catgtgcctg ggaggcctgc tgaccatggt ggcaggagcc gtgtggctga	2460
	cagtgcgggt gctggagctg ttcagagccg cccagctggc caacgatgtg gtgtgcaga	2520
	tcattggagct gtgcggagca gcctttgcc aggtgtgcca caccacagtg ccatggccca	2580
	atgcctcct gacccccaa tggaacaatg agacaacaca gcctcagatc gccaaactgta	2640
	gcgtgtacga cttcttcgtg tggctgcact actatagcgt gagggatacc ctgtggcccc	2700
	gcgtgacata ccacatgaat aagtacgcct atcacatgct ggagaggcgc gccaaagtata	2760
[0027]	agagaggccc tggcccaggc gcaaagttt tggcagcatg gaccctgaag gccgccgccg	2820
	gccccggccc cggccagtat atcaaggcta acagtaagtt cattggaatc acagagctgg	2880
	gaccgggacc tggataatga gtttaaacctc ccatttaaat gtgagggtta atgcttcgag	2940
	cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa	3000
	aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	3060
	ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt	3120
	gggaggtttt ttaaagcaag taaaacctct acaaatgtgg taaaataact ataacggtcc	3180
	taaggtagcg agtgagtagt gttctggggc gggggaggac ctgcatgagg gccagaataa	3240
	ctgaaatctg tgcttttctg tgtgttcag cagcatgagc ggaagcggct cctttgaggg	3300
	agggtattc agcccttate tgacggggcg tctccccctc tggcgaggag tgcgtcagaa	3360
	tgtgatggga tccacggtgg acggccggcc cgtgcagccc gcgaactctt caaccctgac	3420
	ctatgaacc ctgagctctt cgtcgttga cgcagctgcc gccgcagctg ctgcatctgc	3480
	cgccagcgcc gtgcgcggaa tggccatggg cgccggtac tacggcactc tgggtggccaa	3540
	ctcagattcc accaataatc ccgccagcct gaacaggag aagctgttgc tgctgatggc	3600

	ccagctcgag gccttgacc agcgctggg cgagctgacc cagcaggtgg ctgagctgca	3660
	ggagcagacg cgggcccggg ttgccacggt gaaatccaaa taaaaaatga atcaataaat	3720
	aaacggagac ggttgttgat tttaacacag agtctgaatc tttatttgat ttttcgcgcg	3780
	cggtaggccc tggaccaccg gtctcgatca ttgagcacc ggtggatctt ttccaggacc	3840
	cggtagaggt gggcttgat gttgaggtac atgggcatga gcccgccccg ggggtggagg	3900
	tagctccatt gcagggcctc gtgctcgggg gtggtgttgt aatcaccca gtcatacgag	3960
	gggcgcaggg catggtgttg cacaatatct ttgaggagga gactgatggc cacgggcagc	4020
	cctttggtgt aggtgtttac aaatctgttg agctgggagg gatgcatgcg gggggagatg	4080
	aggtgcatct tggcctggat cttgagattg gcgatgttac cgcccagatc ccgcctgggg	4140
	ttcatgttgt gcaggaccac cagcacggtg tatccggtgc acttggggaa ttatcatgc	4200
	aacttgaag ggaaggcgtg aaagaatttg gcgacgcctt tgtgcccgcc caggttttcc	4260
	atgcaactcat ccatgatgat ggcgatgggc ccgtgggcgg cggcctgggc aaagacgttt	4320
[0028]	cgggggtcgg acacatcata gttgtgttcc tgggtgaggt catcataggc cattttaatg	4380
	aatttggggc ggagggtgcc ggactggggg acaaaggtag cctcgatccc gggggcgtag	4440
	ttcccctcac agatctgcat ctcccaggct ttgagctcgg agggggggat catgtccacc	4500
	tgcggggcga taaagaacac ggtttccggg gcgggggaga tgagctgggc cgaaagcaag	4560
	ttccggagca gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc	4620
	tgcaggtggt agttgaggga gagacagctg ccgtcctccc ggaggagggg ggccacctcg	4680
	ttcatcatct cgcgcacgtg catgttctcg cgcaccagtt ccgccaggag gcgctctccc	4740
	cccagggata ggagctctcg gagcgaggcg aagtttttca gcggcttgag tccgtcggcc	4800
	atgggcattt tggagagggt ttgttcaag agttccaggc ggtcccagag ctcggtgatg	4860
	tgctctacgg catctcgatc cagcagacct cctcgtttcg cgggttggga cgctgcggg	4920
	agtagggcac cagacgatgg gcgtccagcg cagccagggt ccggtccttc cagggtcgca	4980
	gcgtccgcgt cagggtggtc tccgtcacgg tgaaggggtg cgcgccgggc tgggcgcttg	5040
	cgagggtgcg cttcaggctc atccggctgg tcgaaaaccg ctcccgatcg gcgccctgcg	5100

	cgctggccag gtagcaattg accatgagtt cgtagttgag cgccctggcc gcgtggcctt	5160
	tggcgcggag cttacctttg gaagtctgcc cgcaggcggg acagaggagg gacttgaggg	5220
	cgtagagctt gggggcgagg aagacggact cggggcgta ggcgtccgcg ccgcagtggg	5280
	cgcagacggt ctcgcactcc acgagccagg tgaggctggg ctggctgggg tcaaaaacca	5340
	gtttcccgcc gttctttttg atgcgtttct tacctttggt ctccatgagc tcgtgtcccc	5400
	gctgggtgac aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cgtcctcga	5460
	gcggtgtgcc gcggtcctcc tcgtagagga accccgcccc ctccgagacg aaagcccggg	5520
	tccaggccag cacgaaggag gccacgtggg acgggtagcg gtcgttgtcc accagcgggt	5580
	ccacctttc cagggtatgc aaacacatgt cccctcgtc cacatccagg aaggtgattg	5640
	gcttgaagt gtagccacg tgaccggggg tcccggccgg gggggtataa aagggtgcgg	5700
	gtccctgctc gtccctactg tcttccgat cgctgtccag gagcgcacg tgttgggta	5760
	ggtattcct ctcgaaggcg ggcatgacct cggcactcag gttgtcagtt tctagaaacg	5820
	aggaggattt gatattgacg gtgccggcgg agatgccttt caagagcccc tcgtccatct	5880
[0029]	ggtcagaaaa gacgatcttt ttgttgcga gcttgggtgc gaaggagccg tagaggcgt	5940
	tggagaggag cttggcgatg gagcgcattg tctggttttt ttccttgtcg gcgcgtcct	6000
	tggcggcgat gttgagctgc acgtactcgc gcgccacgca cttccattcg gggaagacgg	6060
	tggtcagctc gtcgggcacg attctgacct gccagccccg attatgcagg gtgatgaggt	6120
	ccacactggt ggccacctcg ccgcgcaggg gctcattagt ccagcagagg cgtccgcct	6180
	tgcgcgagca gaaggggggc aggggtcca gcatgacctc gtcggggggg tcggcatcga	6240
	tggtgaagat gccgggcagg aggtcggggt caaagtagct gatggaagtg gccagatcgt	6300
	ccagggcagc ttgccattcg cgcacggcca gcgcgcctc gtagggactg agggcgtgc	6360
	cccaggcat gggatgggta agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga	6420
	ggggctctc gaggatgccg atgtaggtgg ggtagcagcg cccccgcgg atgctggcgc	6480
	gcacgtagtc atacagctcg tgcgaggggg cgaggagccc cgggccagg ttggtgcgac	6540
	tggcctttc ggccggtag acgatctggc ggaaaatggc atgcgagttg gaggagatgg	6600
	tggcctttg gaagatggtg aagtggcgt ggggcagtcc gaccgagtcg cggatgaagt	6660

	gggcgtagga gtcttcgagc ttggcgacga gctcggcggg gactaggacg tccagagcgc	6720
	agtagtcgag ggtctctggt atgatgtcat acttgagctg tcccttttgt ttccacagct	6780
	cgcggttgag aaggaactct tcgcggtcct tccagtactc ttcgaggggg aaccctgcct	6840
	gatctgcacg gtaagagcct agcatgtaga actggttgac ggcctttag ggcgagcagc	6900
	ccttctccac ggggagggcg taggcctggg cggccttgcg cagggagggtg tgcgtgaggg	6960
	cgaaagtgtc cctgaccatg accttgagga actggtgctt gaagtcgata tcgtcgcagc	7020
	ccccctgctc ccagagctgg aagtccgtgc gcttcttcta ggcggggttg ggcaaagcga	7080
	aagtaacatc gttgaagagg atcttgcccg cgcggggcat aaagttgcga gtgatgcgga	7140
	aaggttgggg cacctcggcc cggttgttga tgacctgggc ggcgagcacg atctcgtcga	7200
	agccgttgat gttgtggccc acgatgtaga gtccacgaa tcgcgacgg cccttgacgt	7260
	ggggcagttt cttgagctcc tcgtaggtga gctcgtcggg gtcgctgagc ccgtgctgct	7320
	cgagcgccca gtcgcgaga tgggggttg cgcggaggaa ggaagtccag agatccacgg	7380
[0030]	ccagggcggt ttgcagacgg tcccgtact gacggaactg ctgcccagc gccatttttt	7440
	cgggggtgac gcagtagaag gtgcgggggt ccccggtcca gcgatcccat ttgagctgga	7500
	gggcgagatc gagggcgagc tcgacgagcc ggtcgtcccc ggagagtttc atgaccagca	7560
	tgaaggggac gagctgcttg ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg	7620
	tgaggaagag cctttcggtg cgaggatgag agccgatggg gaagaactgg atctcctgcc	7680
	accaattgga ggaatggctg ttgatgtgat ggaagtagaa atgccgacgg cgcgccgaac	7740
	actcgtgctt gtgtttatac aagcgccac agtgctcgca acgctgcacg ggatgcacgt	7800
	gctgcacgag ctgtacctga gttcctttga cgaggaattt cagtgggaag tggagtcgtg	7860
	gcgcctgcat ctcgtgctgt actacgtcgt ggtggtcggc ctggccctct tctgcctcga	7920
	tggtggtcat gctgacgagc ccgcgcgga ggcaggtcca gacctcggcg cgagcgggtc	7980
	ggagagcgag gacgagggcg cgcagggcgg agctgtccag ggtcctgaga cgctcggag	8040
	tcaggtcagt gggcagcggc ggccgcgggt tgacttgacg gagtttttcc agggcgcgcg	8100
	ggaggtccag atggtacttg atctccaccg cgccattggt ggcgacgtcg atggcttgca	8160

	gggtcccgtg cccctggggt gtgaccaccg tccccgttt cttcttgggc ggctggggcg	8220
	acgggggcgg tgcctcttcc atggttagaa gcggcggcga ggacgcgcgc cgggcggcag	8280
	gggcggctcg gggcccggag gcaggggcgg caggggcacg tcggcgccgc gcgcgggtag	8340
	gttctggtac tgcgcccga gaagactggc gtgagcgacg acgcgacggt tgacgtcctg	8400
	gatctgacgc ctctgggtga aggccacggg acccgtgagt ttgaacctga aagagagttc	8460
	gacagaatca atctcggtat cgttgacggc ggccctgccgc aggatctctt gcacgtcgcc	8520
	cgagttgtcc tggtaggcga tctcggtcat gaactgctcg atctctctct ctgaaaggtc	8580
	tccgcggccg gcgcgctcca cggtagccgc gaggtcgttg gagatgcggc ccatgagctg	8640
	cgagaaggcg ttcattgccc cctcgttcca gacgcggctg tagaccacga cgcctcggg	8700
	atcgcgggcg cgcattgacca cctgggcgag gttgagctcc acgtggcgcg tgaagaccgc	8760
	gtagttgcag aggcgctggt agaggtagtt gagcgtggtg gcgatgtgct cggtagcгаа	8820
	gaaatacatg atccagcggc ggagcggcat ctcgctgacg tcgcccagcg cctccaaacg	8880
[0031]	ttccatggcc tcgtaaaagt ccacggcgaa gttgaaaaac tgggagttgc gcgccgagac	8940
	ggtcaactcc tectccagaa gacggatgag ctcgcgatg gtggcgcgca cctcgcgctc	9000
	gaaggccccc gggagttcct ccacttcctc ttcttctctc tccactaaca tctcttctac	9060
	ttctctctca ggccgacgtg gtggcggggg agggggcctg cgtcgccggc ggccgacggg	9120
	cagacggtcg atgaagcgtc cgatggtctc gccgcgccgg cgtcgcatgg tctcggtagc	9180
	ggcgcgcccg tctcgcggg gccgcagcgt gaagacccc cgcgcacatc ccaggtggcc	9240
	gggggggtcc ccgttgggca gggagagggc gctgacgatg catcttatca attgccccgt	9300
	agggactccg cgcaaggacc tgagcgtctc gagatccacg ggatctgaaa accgctgaac	9360
	gaaggcttcg agccagtcgc agtcgcaagg taggctgagc acggtttctt ctggcgggtc	9420
	atgttggttg ggagcggggc gggcgatgct gctggtgatg aagttgaaat aggcggttct	9480
	gagacggcgg atggtggcga ggagcaccag gtctttgggc ccggcttgct ggatgcgcag	9540
	acggtcggcc atgccccagg cgtggtcctg acacctggcc aggtccttgt agtagtcctg	9600
	catgagccgc tccacgggca cctctctctc gcccgccggc ccgtgcatgc gcgtgagccc	9660
	gaagcccgcg tggggctgga cgagcggcag gtcggcgacg acgcgctcgg cgaggatggc	9720

	ttgctggatc tgggtgaggg tggctctgga gtcacaaag tcgacgaagc ggtggtaggc	9780
	tccggtgttg atggtgtagg agcagttggc catgacggac cagttgacgg tctggtggcc	9840
	cggacgcacg agctcgtggt acttgaggcg cgagtaggcg cgcgtgtcga agatgtagtc	9900
	gttgcaggtg cgcaccaggt actggtagcc gatgaggaag tgcggcggcg gctggcggta	9960
	gagcggccat cgctcgggtg cggggcgccc gggcgcgagg tcctcagca tgggtcgggtg	10020
	gtagccgtag atgtacctg acatccaggt gatgccggcg gcggtggtgg aggcgcgcgg	10080
	gaactcgcgg acgcggttcc agatggtgcg cagcggcagg aagtagttca tgggtggcac	10140
	ggtctggccc gtgaggcgcg cgcagtcgtg gatgctctat acgggcaaaa acgaaagcgg	10200
	tcagcggctc gactccgtgg cctggaggct aagcgaacgg gttgggctgc gcgtgtacc	10260
	cggttcgaat ctcgaatcag gctggagccg cagctaactg ggtattggca ctcccgtctc	10320
	gaccaagcc tgcaccaacc ctccaggata cggaggcggg tcgttttgca actttttttt	10380
	ggaggccgga tgagactagt aagcgcggaa agcggcccac cgcgatggct cgtgcctga	10440
[0032]	gtctggagaa gaatgccag ggttgcgttg cgggtgcccc cggttcgagg ccggccgat	10500
	tccgcggcta acgagggcgt ggctgccccg tcgtttccaa gaccccatag ccagccgact	10560
	tctccagtta cggagcgagc cctcttttg tttgtttgt ttttgccaga tgcattccgt	10620
	actcggcag atgcgcccc accaccctcc accgcaaaa cagccccctc cacagccggc	10680
	gcttctgcc ccgccccagc agcaacttcc agccacgacc gccgcggccg ccgtgagcgg	10740
	ggctggacag agttatgatc accagctggc cttggaagag ggcgaggggc tggcgcgct	10800
	ggggcgctcg tcgccggagc ggcaccgcg cgtgcagatg aaaagggacg ctgcgaggc	10860
	ctacgtgccc aagcagaacc tgttcagaga caggagcggc gaggagcccg aggagatgcg	10920
	cgcggccccg ttccacgcgg ggcgggagct gcggcgcggc ctggaccgaa agagggtgct	10980
	gaggacgag gatttcgagg cggacgagct gacggggatc agccccgcgc gcgcgcacgt	11040
	ggcccggccc aacctggtca cggcgtacga gcagaccgtg aaggaggaga gcaacttcca	11100
	aaaatccttc aacaaccacg tgcgcacct gatcgcgcgc gaggaggtga ccctggcct	11160
	gatgcacctg tgggacctgc tggaggccat cgtgcagaac cccaccagca agccgctgac	11220

	ggcgcagctg ttcttggtgg tgcagcatag tcgggacaac gaagcgttca gggaggcgtc	11280
	gctgaatata accgagcccc agggccgctg gctcctggac ctggtgaaca ttctgcagag	11340
	catcgtggtg caggagcgcg ggctgccgct gtccgagaag ctggcggcca tcaacttctc	11400
	ggtgctgagt ttgggcaagt actacgctag gaagatctac aagaccccgt acgtgcccat	11460
	agacaaggag gtgaagatcg acgggtttta catgcgcatg accctgaaag tgctgaccct	11520
	gagcgacgat ctgggggtgt accgcaacga caggatgcac cgtgcggtga gcgccagcag	11580
	gcggcgcgag ctgagcgacc aggagctgat gcatagtctg cagcgggccc tgaccggggc	11640
	cgggaccgag ggggagagct actttgacat gggcgcggac ctgcaactggc agcccagccg	11700
	ccgggccttg gaggcggcgg caggacccta cgtagaagag gtggacgatg aggtggacga	11760
	ggagggcgag tacctggaag actgatggcg cgaccgtatt ttgctagat gcaacaacaa	11820
	cagccacctc ctgatcccc gatgcggcgg gcgctgcaga gccagccgtc cggcattaac	11880
	tcctcggacg attggacce ggccatgcaa cgcacatgg cgctgacgac ccgcaacccc	11940
	gaagccttta gacagcagcc ccaggccaac cggctctcgg ccatcctgga ggcctggtg	12000
[0033]	ccctcgcgct ccaaccccc gcacgagaag gtccctggcca tcgtgaacgc gctggtggag	12060
	aacaaggcca tccgcggcga cgaggccggc ctggtgtaca acgcgctgct ggagcgcgtg	12120
	gcccgtaca acagcaccaa cgtgcagacc aacctggacc gcatggtgac cgacgtgcgc	12180
	gaggccgtgg cccagcgcga gcggttccac cgcgagtcca acctgggac catggtggcg	12240
	ctgaacgect tcctcagcac ccagcccgcc aacgtgcccc ggggccagga ggactacacc	12300
	aacttcatca gcgccctgcg cctgatggtg accgaggtgc cccagagcga ggtgtaccag	12360
	tccgggccgg actacttctt ccagaccagt cgccagggt tgcagaccgt gaacctgagc	12420
	caggctttca agaacttgca gggcctgtgg ggcgtgcagg ccccggtcgg ggaccgcgcg	12480
	acggtgtcga gcctgctgac gccgaactcg cgcctgctgc tgctgctggt ggcccccttc	12540
	acggacagcg gcagcatcaa ccgcaactcg tacctgggct acctgattaa cctgtaccgc	12600
	gaggccatcg gccaggcgca cgtggacgag cagacctacc aggagatcac ccacgtgagc	12660
	cgcccccttg gccaggacga cccgggcaac ctggaagcca cctgaactt ttgctgacc	12720
	aaccggtcgc agaagatccc gccccagtac gcgctcagca ccgaggagga gcgcatcctg	12780

	cgttacgtgc agcagagcgt gggcctgttc ctgatgcagg agggggccac ccccagcgcc	12840
	gcgctcgaca tgaccgcgcg caacatggag cccagcatgt acgccagcaa ccgcccgttc	12900
	atcaataaac tgatggacta cttgcatcgg gcggccgcca tgaactctga ctatttcacc	12960
	aacgccatcc tgaatcccca ctggtccccg ccgcccgggt tctacacggg cgagtacgac	13020
	atgccccgacc ccaatgacgg gttcctgtgg gacgatgtgg acagcagcgt gttctcccc	13080
	cgaccgggtg ctaacgagcg ccccttgtgg aagaaggaag gcagcgaccg acgcccgtcc	13140
	tggcgctgt ccggccgca ggggtctgcc gcggcgggtc ccgaggccgc cagtcctttc	13200
	ccgagcttgc cttctcgtc gaacagtatc cgcagcagcg agctgggcag gatcacgcgc	13260
	ccgcgcttgc tggcggaaga ggagtacttg aatgactcgc tgttgagacc cgagcgggag	13320
	aagaacttcc ccaataacgg gatagaaagc ctggtggaca agatgagccg ctggaagacg	13380
	tatgcgcagg agcacaggga cgatccccgg gcgtcgcagg gggccacgag ccggggcagc	13440
	gccgcccgtg aacgcccgtg gcacgacagg cagcggggac agatgtggga cgatgaggac	13500
[0034]	tccgccgacg acagcagcgt gttggacttg ggtgggagtg gtaaccggtt cgctcacctg	13560
	cgccccgta tcggcgcat gatgtaagag aaaccgaaaa taaatgatac tcaccaaggc	13620
	catggcgacc agcgtgcgtt cgtttcttct ctgttgttgt tgtatctagt atgatgagge	13680
	gtgcgtacce ggagggtcct cctccctcgt acgagagcgt gatgcagcag gcgatggcgg	13740
	cgggcgcat gcagccccg ctggaggctc cttacgtgcc cccgcggtac ctggcgcta	13800
	cggagggcg gaacagcatt cgttactcgg agctggcacc cttgtacgat accaccggt	13860
	tgtacctggt ggacaacaag tcggcggaca tcgcctcgt gaactaccag aacgaccaca	13920
	gcaacttct gaccaccgtg gtgcagaaca atgacttca cccacggag gccagcacc	13980
	agaccatcaa ctttgacgag cgctcgcgtt gggcgccca gctgaaaacc atcatgcaca	14040
	ccaacatgcc caacgtgaac gagttcatgt acagcaaca gttcaaggcg cgggtgatgg	14100
	tctcccga gacccccaat ggggtgacag tgacagagga ttatgatggt agtcaggatg	14160
	agctgaagta tgaatgggtg gaatttgagc tgcccgaagg caacttctc gtgaccatga	14220
	ccatcgacct gatgaacaac gccatcatcg acaattactt ggcggtgggg cggcagaacg	14280

	gggtgctgga gagcgacatc ggcgtgaagt tcgacactag gaacttcagg ctgggctggg	14340
	accccgtagc cgagctggtc atgcccgggg tgtacaccaa cgaggctttc catcccgata	14400
	ttgtcttget gcccggctgc ggggtggact tcaccgagag ccgcctcagc aacctgctgg	14460
	gcattcgcaa gaggcagccc ttccaggaag gcttccagat catgtacgag gatctggagg	14520
	ggggcaacat ccccgcgctc ctggatgtcg acgcctatga gaaaagcaag gaggatgcag	14580
	cagctgaage aactgcagcc gtagctaccg cctctaccga ggtcaggggc gataattttg	14640
	caagcgccgc agcagtgga gggccgagg cggctgaaac cgaaagtaag atagtcattc	14700
	agccggtgga gaaggatagc aagaacagga gctacaacgt actaccggac aagataaaca	14760
	ccgcctaccg cagctggtag cttagcctaca actatggcga ccccgagaag ggcgtgcgct	14820
	cctggacgct gctcaccacc tcggacgtca cctgcggcgt ggagcaagtc tactggtcgc	14880
	tgcccagat gatgcaagac ccggtcacct tccgctccac gcgtcaagtt agcaactacc	14940
	cgttggtggg cgccgagctc ctgcccgtct actccaagag cttcttcaac gagcaggccg	15000
	tctactcgca gcagctgcgc gccttcacct cgcttacgca cgtcttcaac cgcttccccg	15060
[0035]	agaaccagat cctcgtccgc ccgcccgcgc ccaccattac caccgtcagt gaaaacgttc	15120
	ctgctctcac agatcacggg accctgccgc tgcgcagcag tatccgggga gtccagcgcg	15180
	tgaccgttac tgaccgaca cgccgcacct gccctacgt ctacaaggcc ctgggcatag	15240
	tcgcccgcgc cgtcctctcg agccgcacct tctaaatgct cattctcacc tcgcccagta	15300
	ataacaccgg ttggggcctg cgcgccccca gcaagatgta cggaggcgt cgccaacgt	15360
	ccacgcaaca ccccgtagc gtgcccgggc acttccgcgc tccctggggc gcctcaagg	15420
	gcccgtagc gtcgcccacc accgtcagc acgtgatcga ccaggtggtg gccgacgcgc	15480
	gcaactacac ccccgcgccc gcgcccgtct ccaccgtgga cgccgtcacc gacagcgtgg	15540
	tgcccagcgc gcgcccgtac gcccgcccga agagccggcg gcgcccacc gcccgccgc	15600
	accggagcac ccccgccatg cgcgccggcg gacccttctt gcgcccggcc aggcgcacgg	15660
	gacgagggc catgctcagg gcggccagac gcgcccgttc aggcgccagc gccggcagga	15720
	cccggagacg cgcccaccg gcggccggcg cgccatcgc cagcatgtcc cgcccgcgccc	15780
	gagggaacgt gtactgggtg cgcgaccgccc ccaccggtgt gcgcccgtccc gtgcccacc	15840

	gccccctcg cacttgaaga tgttcacttc gcgatgttga tgtgtcccag cggcgaggag	15900
	gatgtccaag cgcaaattca aggaagagat gctccaggtc atcgcgctg agatctacgg	15960
	ccctgcggtg gtgaaggagg aaagaaagcc ccgcaaaatc aagcgggtca aaaaggacaa	16020
	aaaggaagaa gaaagtgatg tggacggatt ggtggagttt gtgcgcgagt tcgcccccg	16080
	gcggcgcgtg cagtggcgcg ggcggaaggt gcaaccggtg ctgagaccg gcaccaccgt	16140
	ggtcttcacg cccggcgagc gctccggcac cgcttccaag cgctcctacg acgaggtgta	16200
	cggggatgat gatattctgg agcaggcggc cgagcgctg ggcgagttg cttacggcaa	16260
	gcgcagccgt tccgcaccga aggaagaggc ggtgtccatc ccgctggacc acggcaacce	16320
	cacgccgagc ctcaagcccg tgaccttga gcaggtgctg ccgaccgcg cgccgcgccg	16380
	ggggttcaag cgcgaggcg aggatctgta cccaccatg cagctgatgg tgcccaagcg	16440
	ccagaagctg gaagacgtgc tggagaccat gaaggtggac ccggacgtgc agcccaggt	16500
	caaggtgctg cccatcaagc aggtggcccc gggcctgggc gtgcagaccg tggacatcaa	16560
[0036]	gattcccacg gagcccatgg aaacgcagac cgagcccatg atcaagccca gcaccagcac	16620
	catggaggtg cagacggatc cctggatgcc atcggctcct agtcgaagac cccggcgcaa	16680
	gtacggcgcg gccagcctgc tgatgccc aa ctacgcgtg catcctcca tcatccccac	16740
	gccgggctac cgcggcacgc gttctaccg cggtcatacc agcagccgcc gccgcaagac	16800
	caccactcgc ccccgcctc gccgcaccgc cgtgcaacc acccctgcc ccttgggtgc	16860
	gagagtgtac cgcgcggcc gcgcacctc gacctgccg cgcgcgcgt accaccgag	16920
	catgccatt taaactttcg cctgctttgc agatcaatgg ccctcatatg ccgcttgc	16980
	gttcccatta cgggctaccg aggaagaaaa ccgcgccgta gaaggtggc ggggaacggg	17040
	atgcgtgcc accaccaccg gcggcggcgc gccatcagca agcggttggg gggaggcttc	17100
	ctgcccgcgc tgatcccat catcgccgcg gcgategggg cgatccccg cattgettcc	17160
	gtggcgggtc aggcctctca gcgccactga gacacacttg gaaacatctt gtaataaacc	17220
	aatggactct gacgtcctg gtctgtgat gtgttttctg agacagatgg aagacatcaa	17280
	ttttctgcc ctggctccgc gacacggcac gcggccgttc atgggcacct ggagcgacat	17340

	cggcaccagc caactgaacg ggggcgcctt caattggagc agtctctgga gcgggcttaa	17400
	gaatttcggg tccacgctta aaacctatgg cagcaaggcg tggaacagca ccacagggca	17460
	ggcgctgagg gataagctga aagagcagaa cttccagcag aaggtggtcg atgggctcgc	17520
	ctcgggcatac aacgggggtg tggacctggc caaccaggcc gtgcagcggc agatcaacag	17580
	ccgcctggac ccggtgccgc ccgccggctc cgtggagatg ccgcaggtgg aggaggagct	17640
	gcctcccctg gacaagcggg gcgagaagcg accccgcccc gatgcggagg agacgctgct	17700
	gacgcacacg gacgagccgc ccccgtagca ggaggcggtg aaactgggtc tgcccaccac	17760
	gcggcccatc gcgccccctg ccaccggggt gctgaaacct gaaaagcccc cgaccctgga	17820
	cttgccctct cccagcctt ccgccccctc tacagtggct aagcccctgc cgccggtggc	17880
	cgtggccccgc gcgcgacccg ggggcaccgc ccgcctcat gcgaaactggc agagcactct	17940
	gaacagcadc gtgggtctgg gactgcagag tgtgaagcgc cgccgctgct attaaacct	18000
	ccgtagcget taacttgctt gtctgtgtgt gtatgtatta tctcgccgcc gccgctgtcc	18060
[0037]	accagaagga ggagtgaaga ggcgctcgc cgagttgcaa gatggccacc ccatcgatgc	18120
	tgccccagtg ggcgtacatg cacatgcgcg gacaggacgc ttcggagtac ctgagtccgg	18180
	gtctggtgca gtttgccccg gccacagaca cctacttcag tctggggaac aagtttagga	18240
	acccacaggt ggcgccccac cacgatgtga ccaccgaccg cagccagcgg ctgacgctgc	18300
	gcttcgtgcc cgtggaccgc gaggacaaca cctactcgta caaagtgcgc tacacgctgg	18360
	ccgtgggcga caaccgctg ctggacatgg ccagcaccta ctttgacatc cgcggcgtgc	18420
	tggatcgggg ccctagcttc aaacctact ccggcaccgc ctacaacagt ctggcccca	18480
	agggagcacc caacacttgt cagtggacat ataaagccga tggtgaaact gccacagaaa	18540
	aaacctatac atatggaaat gcacccgtgc agggcattaa catcacaaaa gatggtatc	18600
	aacttgaac tgacaccgat gatcagccaa tctacgcaga taaaacctat cagcctgaac	18660
	ctcaagtggg tgatgctgaa tggcatgaca tcaactgtac tgatgaaaag tatggaggca	18720
	gagctcttaa gcctgatacc aaaatgaagc cttgttatgg ttcttttgcc aagcctacta	18780
	ataaagaagg aggtcaggca aatgtgaaaa caggaacagg cactactaaa gaatatgaca	18840
	tagacatggc tttcttgac aacagaagtg cggctgctgc tggcctagct ccagaaattg	18900

	ttttgtatac tgaaaatgtg gatttggaaa ctccagatac ccatattgta tacaagcag	18960
	gcacagatga cagcagctct tetattaatt tgggtcagca agccatgccc aacagaccta	19020
	actacattgg tttcagagac aactttatcg ggctcatgta ctacaacagc actggcaata	19080
	tgggggtgct ggccggtcag gcttctcagc tgaatgctgt ggttgacttg caagacagaa	19140
	acaccgagct gtccaccag ctcttgcttg actctctggg tgacagaacc cggatattca	19200
	gtatgtggaa tcaggcgtg gacagctatg atcctgatgt gcgcattatt gaaaatcatg	19260
	gtgtggagga tgaacttccc aactattgtt tccctctgga tgctgttggc agaacagata	19320
	cttatcaggg aattaaggct aatggaactg atcaaaccac atggaccaa gatgacagt	19380
	tcaatgatgc taatgagata ggcaaggga atccattcgc catggaaatc aacatccaag	19440
	ccaacctgtg gaggaacttc ctctacgcca acgtggccct gtacctgccc gactcttaca	19500
	agtacacgcc ggccaatgtt accctgccc ccaacaccaa cacctacgat tacatgaacg	19560
	gccgggtggt ggcgccctcg ctggiggact cctacatcaa catcggggcg cgctggtcgc	19620
[0038]	tgatcccat ggacaactg aacccttca accaccacg caatgcgggg ctgcgtacc	19680
	gctccatgct cctgggcaac gggcgctacg tgccccttca catccaggtg ccccagaaat	19740
	ttttgccat caagagcctc ctgctcctgc ccgggtccta cacctacgag tggaaacttc	19800
	gcaaggacgt caacatgatc ctgcagagct cctcggcaa cgacctgccc acggacgggg	19860
	cctccatctc cttcaccagc atcaacctct acgccacctt cttcccatg gcgcacaaca	19920
	cggcctccac gctcgaggcc atgctgcgca acgacaccaa cgaccagtcc ttcaacgact	19980
	acctctcggc ggccaacatg ctctacceca tcccggcaa cgccaccaac gtgccatct	20040
	ccatcccctc gcgcaactgg gccgccttcc gcggttggtc cttcacgct ctcaagacca	20100
	aggagacgcc ctgctgggc tccgggttcg acccctactt cgtctactcg ggtccatcc	20160
	cctacctega cggcaacttc tactcaacc acaccttcaa gaaggtctcc atcaccttcg	20220
	actctccgt cagctggccc ggcaacgacc ggctcctgac gcccaacgag ttcgaaatca	20280
	agcgcaccgt cgacggcgag ggctacaacg tggeccagt caacatgacc aaggactggt	20340
	tcttggtcca gatgctggcc cactacaaca tcggctacca gggcttctac gtgcccagg	20400

	gctacaagga cgcgatgtac tctttcttcc gcaacttcca gcccatgagc cgccaggtgg	20460
	tggacgaggt caactacaag gactaccagg ccgtcaccct ggctaccag cacaacaact	20520
	egggettctg eggtacctc ggcgccacca tgcgccagg ccagccctac cccgccaact	20580
	accctaccc gctcatcggc aagagcgccg tcaccagcgt caccagaaa aagttcctct	20640
	gcgacagggc catgtggcgc atccccttct ccagcaactt catgtccatg ggcgcgctca	20700
	ccgacctcgg ccagaacatg ctctatgcca actccgcca cgcgctagac atgaatttcg	20760
	aagtcgacc catggatgag tccacccttc tctatgttgt cttcgaagtc ttcgacgtcg	20820
	tccgagtgea ccagccccac cgcggcgtca tcgaggcgt ctacctgcgc acccccttct	20880
	cggccggtaa cgccaccacc taagctcttg cttcttgcaa gccatggccg cgggctccgg	20940
	cgagcaggag ctcagggcca tcatccgca cctgggctgc gggccctact tcctgggcac	21000
	cttcgataag cgcttcccgg gattcatggc cccgcacaag ctggcctgcg ccatcgtcaa	21060
	cacggccggc cgcgagaccg ggggcgagca ctggctggcc ttcgcctgga acccgcctc	21120
	gaacacctgc tacctcttcg accccttcgg gttctcggac gagcgcctca agcagatcta	21180
[0039]	ccagttcgag tacgagggcc tgctgcgccc cagcgcctg gccaccgagg accgctgcgt	21240
	caccctgga aagtccacc agacctgca gggctccgccc tcggccgccc gcgggctctt	21300
	ctgtgcatg ttctgcacg cttctgtgca ctggcccgc cccccatgg acaagaacc	21360
	caccatgaac ttgtgacgg gggtgcccaa cggcatgctc cagtcgccc agtggaacc	21420
	caccctgcgc cgcaaccagg aggcgctcta ccgcttctc aactcccact ccgctactt	21480
	tcgtccccc cgcgcgcgca tcgagaaggc caccgcttc gaccgcatga atcaagacat	21540
	gtaaaccgtg tgtgtatggt aaatgtcttt aataaacagc actttcatgt tacacatgca	21600
	tctgagatga tttattttaga aatcgaaagg gttctgccc gtctcggcat ggcccgcggg	21660
	caggacacg ttgcggaact ggtacttggc cagccaattg aactcgggga tcagcagttt	21720
	gggcagcggg gtgtcgggga aggagtcggt ccacagcttc cgcgtcagtt gcagggcgcc	21780
	cagcaggtcg ggcgaggaga tcttgaaatc gcagttggga cccgcttct gcgcgaggga	21840
	gttcggttac acggggttgc agcactggaa caccatcagg gccgggtgct tcacgctcgc	21900
	cagcaccgtc gcgtcggta tgctctccac gtcgaggctc tcggcgttg ccatcccga	21960

	gggggtcatc ttgcaggtct gccttcccat ggtgggcacg caccgggct tgtggttgca	22020
	atcgcagtgc aggggatca gcatcatctg ggcctggtcg gcgttcatcc cgggtacat	22080
	ggccttcatg aaagcctcca attgcctgaa cgctgctgg gccttgctc cctcggtgaa	22140
	gaagaccccg caggacttgc tagagaactg gttggtggcg caccggcgt cgtgcacgca	22200
	gcagcgcgcg tcgttgttgg ccagctgcac cacgctgcgc cccagcggg tctgggtgat	22260
	cttgccccgg tcggggttct ccttcagcgc gcgctgccc ttctcgtcg ccacatccat	22320
	ctcgatcatg tgctccttct ggatcatggt ggtcccgtgc aggcaccgca gettgcctc	22380
	ggcctcggtg caccgtgca gccacagcgc gcaccgggtg cactcccagt tcttgtggc	22440
	gatctgggaa tgcgcgtgca cgaagccctg caggaagcgg cccatcatgg tggtcagggt	22500
	cttgttgcta gtgaaggta gcggaatgcc gcggtgctcc tcgttgatgt acagggtgca	22560
	gatcggcggg tacacctcgc cctgctcggg catcagctgg aagttggctt tcaggtegg	22620
	ctccacgcgg tagcgggtcca tcagcatagt catgatttcc atacccttct cccaggccga	22680
[0040]	gacgatgggc aggctcatag ggttcttcac catcatctta gcgctagcag ccgcgccag	22740
	ggggtcgtc tcgtccaggg tctcaaagct ccgcttgccg tccttctcgg tgatccgcac	22800
	cgggggtag ctgaagccca cggccgccag ctctctctcg gcctgtcttt cgtcctcgt	22860
	gtcctggctg acgtcctgca ggaccacatg cttggctttg cggggtttct tcttggcgg	22920
	cagcggcggc ggagatggtg gagatggcga gggggagcgc gagttctcgc tcaccactac	22980
	tatctcttcc tcttcttggg ccgaggccac gcggcggtag gtatgtctct tcggggcag	23040
	aggcggaggc gacgggctct cgccgccgag acttgccgga tggctggcag agccccttc	23100
	gcgttcgggg gtgcetccc ggcggcgtc tgactgactt cctccgcgce cgccattgt	23160
	gttctcctag ggaggaacaa caagcatgga gactcagcca tcgccaacct cgccatctgc	23220
	ccccaccgcc gacgagaagc agcagcagca gaatgaaage ttaaccgcc cgccgccag	23280
	ccccccacc tccgacgcgg ccgtcccaga catgcaagag atggaggaat ccatcgagat	23340
	tgacctgggc tatgtgacgc ccgaggagca cgaggaggag ctggcagtgc gcttttaca	23400
	agaagagata caccaagaac agccagagca ggaagcagag aatgagcaga gtcaggctgg	23460

	gctcgagcat gacggcgact acctccacct gagcgggggg gaggacgcgc tcatcaagca	23520
	tctgccccgg caggccacca tcgtcaagga tgcgctgctc gaccgcaccg aggtgccct	23580
	cagcgtggag gagctcagcc ggcctacga gttgaacctc ttctcgccgc gcgtgcccc	23640
	caagcgccag cccaatggca cctgcgagcc caaccgcgc ctcaacttct acccgttctt	23700
	cgcggtgcc caggccctgg ccacctacca catcttttc aagaaccaa agatccccgt	23760
	ctcctgccgc gccaacgcga cccgcgccga cgccttttc aacctgggtc ccggcgccc	23820
	cctacctgat atgcctcct tggaagaggt tcccaagatc ttcgagggtc tgggcagcga	23880
	cgagactcgg gccgcgaacg ctctgcaagg agaaggagga gagcatgagc accacagcgc	23940
	cctggtcgag ttggaaggcg acaacgcgcg gctggcgggtg ctcaaagca cggtcgagct	24000
	gaccatttc gctaccgct ctctgaacct gcccccaaa gtcagagcg cggtcatgga	24060
	ccaggtgctc atcaagcgcg cgtcgcccat ctccgaggac gagggcatgc aagactccga	24120
	ggagggcaag cccgtggtca ggcagagca gctggccccg tggtgggtc ctaatgctag	24180
	tcccagagt ttggaagagc ggcgcaaact catgatggcc gtggtcctgg tgaccgtgga	24240
[0041]	gctggagtgc ctgcgccgt tcttcgccga cgcggagacc ctgcgcaagg tcgaggagaa	24300
	cctgcactac ctcttcagcg acgggttcgt ggcgagccc tgcaagatct ccaacgtgga	24360
	gctgaccaac ctggtctcct acatgggcat cttgcacgag aaccgcctgg ggcagaact	24420
	gctgcacacc accctgcgcg gggaggcccc gcgcgactac atccgcgact gcgtctacct	24480
	ctacctctgc cacacctggc agacgggcat gggcgtgtgg cagcagtgc tggaggagca	24540
	gaacctgaaa gagctctgca agctcctgca gaagaacctc aagggtctgt ggaccgggtt	24600
	cgacgagcgc accaccgct cggacctggc cgacctcatt ttccccgagc gcctcaggct	24660
	gacgtgcgc aacggcctgc ccgactttat gagccaaagc atgttgcaaa actttcgtc	24720
	tttcatctc gaacgctccg gaatcctgcc cgccacctgc tccgcgtgc cctcggactt	24780
	cgtgccgctg acctcccg agtgcccc gccgctgtgg agccactgct acctgctgcg	24840
	cctggccaac tacctggcct accactcgga cgtgatcgag gacgtcagcg gcgaggcct	24900
	gctcgagtgc cactgccgt gcaacctctg cacgccgac cgtccctgg cctgcaacc	24960
	ccagctgctg agcgagacc agatcatcgg caccttcgag ttgcaaggc ccagcgaagg	25020

	cgagggttca gccgccaagg ggggtctgaa actcaccg gggctgtgga cctcggccta	25080
	cttgcgcaag ttcgtgcccc aggactacca tcccttcgag atcaggttct acgaggacca	25140
	atcccatccg cccaaggccg agctgtcggc ctgcgtcatc acccaggggg cgatcctggc	25200
	ccaattgcaa gccatccaga aatcccgcca agaattcttg ctgaaaaagg gccgcggggg	25260
	ctacctcgac ccccagaccg gtgaggagct caacccccggc ttccccagc atgccccgag	25320
	gaaacaagaa gctgaaagtg gagctgccgc ccgtggagga tttggaggaa gactgggaga	25380
	acagcagtca ggcagaggag gaggagatgg aggaagactg ggacagcact caggcagagg	25440
	aggacagcct gcaagacagt ctggaggaag acgaggagga ggcagaggag gaggtggaag	25500
	aagcagccgc cgccagaccg tcgtcctcgg cgggggagaa agcaagcagc acggatacca	25560
	tctccgctcc gggctcggggg cccgctcgac cacacagtag atgggacgag accggacgat	25620
	tcccgaacce caccaccag accgtaaga aggagcggca gggatacaag tcctggcggg	25680
	ggcacaaaaa cgccatcgtc tcctgcttgc aggcctgcgg gggcaacatc tccttcacce	25740
[0042]	ggcgtacct gctcttcac cgcgggggta actttcccc caacatcttg cattaactacc	25800
	gtcacctcca cagcccctac tacttccaag aagaggcagc agcagcagaa aaagaccagc	25860
	agaaaaccag cagctagaaa atccacagcg gcggcagcag gtggactgag gatcgcggcg	25920
	aacgagccgg cgcaaaccg ggagctgagg aaccggatct tcccaccct ctatgceate	25980
	ttcagcaga gtcgggggca ggagcaggaa ctgaaagtca agaaccgttc tctgcgctcg	26040
	ctcaccgca gtgtctgta tcacaagagc gaagaccaac ttcagcgcac tctcaggagc	26100
	gccgaggctc tcttcaaaa gtactgcgcg ctactctta aagagtagcc cgcgccgcc	26160
	cagtcgcaga aaaaggcggg aattacgtea cctgtgcct tcgccctagc cgcctccacc	26220
	catcatcatg agcaaagaga ttcccacgcc ttacatgtgg agctaccagc cccagatggg	26280
	cctggccgcc ggtgccgcc aggactactc cacccgatg aattggctca gcgccggcc	26340
	cgcgatgata tcacgggtga atgacatccg cgcccaccga aaccagatac tcctagaaca	26400
	gtcagcgtc accgccagc cccgcaatca cctcaatccg cgtaattggc ccgccccct	26460
	ggtgtaccag gaaattcccc agcccacgac cgtactactt ccgcgagacg cccagccga	26520

	agtccagctg actaactcag gtgtccagct ggcgggcggc gccaccctgt gtcgtcaccg	26580
	ccccgctcag ggtataaagc ggctggatgat ccggggcaga ggcacacagc tcaacgacga	26640
	ggtggatgagc tcttcgctgg gtctgcgacc tgacggagtc ttccaactcg ccggatcggg	26700
	gagatcttcc ttcacgcctc gtcaggccgt cctgactttg gagagtctgt cctcgcagcc	26760
	ccgctcgggt ggcatcggca ctctccagtt cgtggaggag ttcactccct cggctctactt	26820
	caacccttc tccgctccc ccggcacta cccggacgag ttcattcccga acttcgacgc	26880
	catcagcgag tcggtggacg gctacgattg aaactaatca ccccttatac cagtgaata	26940
	aagatcatat tgatgatgat tttacagaaa taaaaataa tcatttgatt tgaataaag	27000
	atacaatcat attgatgatt tgagttaac aaaaaataa agaactcactt acttgaaatc	27060
	tgataccagg tctctgtcca tgttttctgc caacaccact tcactccctt cttcccagct	27120
	ctggtactgc aggccccggc gggctgcaaa ctctctccac acgctgaagg ggatgtcaaa	27180
	ttctctctgt cctcaatct tcattttatc ttctatcaga tgcctcaaaa gcgcgtccgg	27240
	gtggatgatg acttcgacce cgtctacccc tacgatgcag acaacgcacc gaccgtgcc	27300
[0043]	ttcatcaacc ccccttctgt ctcttcagat ggattccaag agaagccctt gggggtgttg	27360
	tccttgcgac tggccgacce cgtcaccacc aagaacgggg aaatcacctt caagctggga	27420
	gagggggtgg acctcgatc ctcgggaaaa ctcatctcca acacggccac caagccgccc	27480
	gcccctctca gtttttccaa caacaccatt tcccttaaca tggatcacc cttttact	27540
	aaagatgaa aattatcctt acaagtttct ccaccattaa atatactgag aacaagcatt	27600
	ctaaacacac tagctttagg ttttgatca ggtttaggac tccgtggctc tgccttggca	27660
	gtacagttag tctctccact tacatttgat actgatggaa acataaagct taccttagac	27720
	agaggtttgc atgttacaac aggagatgca attgaaagca acataagctg ggctaaaggt	27780
	ttaaaatttg aagatggagc catagcaacc aacattggaa atgggttaga gtttgaagc	27840
	agtagtacag aaacagggtg tgatgatgct tacccaatcc aagttaaact tggatctggc	27900
	cttagctttg acagtacagg agccataatg gctggttaaca aagaagacga taaactcact	27960
	ttgtggacaa cacctgatcc atcaccaaac tgtcaaatac tcgcagaaaa tgatgcaaaa	28020
	ctaacacttt gcttgactaa atgtggtagt caaatactgg ccactgtgtc agtcttagtt	28080

	gtaggaagtg gaaacctaaa cccattact ggcaccgtaa gcagtgetca ggtgtttcta	28140
	cgttttgatg caaacggtgt tcttttaaca gaacattcta cactaaaaaa atactggggg	28200
	tataggcagg gagatagcat agatggcact ccatatacca atgctgtagg attcatgccc	28260
	aatttaaaag cttatccaaa gtcacaaagt tctactacta aaaataatat agtagggcaa	28320
	gtatacatga atggagatgt ttcaaaacct atgctttctca ctataaccct caatggctact	28380
	gatgacagca acagtacata ttcaatgtca ttttcataca cctggactaa tggaagctat	28440
	gttgagcaa catttggggc taactcttat accttctcat acatcggcca agaatgaaca	28500
	ctgtatccca cctgcatgc caacccttc caccctcctc tgtggaacaa actctgaaac	28560
	acaaaataaa ataaagttca agtgttttat tgattcaaca gttttacagg attcgagcag	28620
	ttattttcc tccaccttc caggacatgg aatacaccac cctctcccc cgcacagcct	28680
	tgaacatctg aatgccattg gtgatggaca tgcttttggc ctccacgttc cacacagttt	28740
	cagagcgagc cagtctcggg tcggtcaggg agatgaaacc ctccgggcac tcccgcattc	28800
[0044]	gcacctcaca gctcaacagc tgaggattgt cctcgggtgt cgggatcacg gttatctgga	28860
	agaagcagaa gagcggcggg gggaatcata gtccgcgaac gggatcggcc ggtggtgtcg	28920
	catcaggccc cgcagcagtc gctgccgccg ccgctccgtc aagctgctgc tcagggggtc	28980
	cgggtccagg gactcctca gcatgatgcc cacggcctc agcatcagtc gtctggtgctg	29040
	gcgggcgcag cagcgcagtc ggatctcgtc caggctcgtc cagtacgtgc aacacagaac	29100
	caccaggttg ttcaacagtc catagttcaa cacgctccag ccgaaactca tcgcggaag	29160
	gatgctacc acgtggccgt cgtaccagat cctcaggtaa atcaagtgtg gccccctca	29220
	gaacacgctg cccacgtaca tgatctcctt gggcatgtgg cggttcacca cctcccggta	29280
	ccacatcacc ctctggttga acatgcagcc ccggatgatc ctgcggaacc acagggccag	29340
	caccgccccg cccgcatgc agcgaagaga ccccggttc cggcaatggc aatggaggac	29400
	ccaccgctcg taccgtgga tcatctggga gctgaacaag tctatgttg cacagcacag	29460
	gcatatgctc atgcatctct tcagactct caactctctg ggggtcaaaa ccatatccca	29520
	gggcacgggg aactcttgca ggacagcgaa ccccgagaa cagggaatc ctgcacaga	29580

	acttacattg tgcatggaca gggtatcgca atcaggcagc accgggtgat cctccaccag	29640
	agaagcgcgg gtctcggctt cctcacagcg tggtaagggg gccggccgat acgggtgatg	29700
	gcgggacgcg gctgatcgtg ttcgcgaccg tgcgatgatg cagttgcttt cggacatttt	29760
	cgtacttgct gtagcagaac ctgggccggg cgctgcacac cgatcgccgg cggcggcttc	29820
	ggcgccttga acgctcgggtg ttgaaattgt aaaacagcca ctctctcaga ccgtgcagca	29880
	gatctagggc ctcaggagtg atgaagatcc catcatgcct gatggctctg atcacatcga	29940
	ccaccgtgga atgggccaga cccagccaga tgatgcaatt ttgttgggtt tcggtgacgg	30000
	cgggggaggg aagaacagga agaaccatga ttaactttta atccaaacgg tctcggagta	30060
	cttcaaaatg aagatcgcgg agatggcacc tctcgcccc gctgtgttgg tggaaaataa	30120
	cagccaggtc aaagtgata cggttctcga gatgttccac ggtggcttcc agcaaagcct	30180
	ccacgcgcac atccagaaac aagacaatag cgaaagcggg agggttctct aattcctcaa	30240
	tcatcatggt acactcctgc accatcccca gataattttc atttttccag ccttgaatga	30300
	ttcgaactag ttctgaggt aaatccaagc cagccatgat aaagagctcg cgcagagcgc	30360
[0045]	cctccaccgg cattcttaag cacaccctca taattccaag atattctgct cctggttca	30420
	ctgcagcaga ttgacaagcg gaatatcaaa atctctgccg cgatccctga gctcctcct	30480
	cagcaataac tgtaagtact ctttcatatc ctctccgaaa tttttagcca taggaccacc	30540
	aggaataaga ttagggcaag ccacagtaca gataaacga agtcctcccc agtgagcatt	30600
	gccaaatgca agactgctat aagcatgctg gctagaccgg gtgatatctt ccagataact	30660
	ggacagaaaa tcgccaggc aatttttaag aaaatcaaca aaagaaaaat cctccaggtg	30720
	gacgtttaga gcctcgggaa caacgatgaa gtaaatacaa gcggtgcgtt ccagcatggt	30780
	tagttagctg atctgtagaa aaaacaaaaa tgaacattaa accatgctag cctggcgaac	30840
	aggtgggtaa atcgttctct ccagcaccag gcaggccacg gggctctcgg cgcgaccctc	30900
	gtaaaaattg tcgctatgat tgaaaacat cacagagaga cgttcccggg ggccggcgtg	30960
	aatgattcga caagatgaat acacccccgg aacattggcg tccgcgagtg aaaaaagcg	31020
	cccaggaag caataaggca ctacaatgct cagtctcaag tccagcaaag cgatgcatg	31080
	cggatgaag acaaaattct caggtgcgta caaatgtaa ttactccct cctgcacagg	31140

	cagcaaagcc cccgatccct ccaggtacac atacaaagcc tcagcgtcca tagcttaccg	31200
	agcagcagca cacaacaggc gcaagagtca gagaaaggct gagctctaac ctgtccaccc	31260
	gctctctgct caatatatag cccagatcta cactgacgta aaggccaaag tctaaaaata	31320
	cccgccaaat aatcacacac gcccagcaca cgcccagaaa ccggtgacac actcaaaaaa	31380
	atagcgcac ttcctcaaac gcccaaaact gccgtcattt cggggttccc acgtacgtc	31440
	atcaaaacac gactttcaaa ttccgtcgac cgttaaaaac gtcaccgcc ccgccctaa	31500
	cggtgcgccg tctctcagcc aatcagcgcc ccgcatcccc aaattcaaac acctcatttg	31560
	catattaacg cgcacaaaaa gtttgagg	31588
	<210> 3	
	<211> 11447	
	<212> DNA	
	<213> 委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus)	
	<400> 3	
	atgggcgcg catgagagaa gccagacca attacctacc caaatggag aaagttcacg	60
[0046]	ttgacatcga ggaagacagc ccattcctca gagctttgca gggagcttc ccgagtttg	120
	aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc	180
	tggcttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa	240
	gtgcgcccgc ccgcagaatg tattctaage acaagtatca ttgtatctgt ccgatgagat	300
	gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
	aaataactga taaggaattg gacaagaaaa tgaaggagct cgccgccgtc atgagcgacc	420
	ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
	aagtcgctgt ttaccaggat gtatacgcg ttgacggacc gacaagtctc taccaccaag	540
	ccaataaggg agttagagtc gctactgga taggctttga caccaccct tttatgttta	600
	agaacttgge tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa	660
	cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttcaact	840

	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagetatc agtccaggcc tgtatgggaa gccttcagge tatgctgcta	960
	cgatgcaccg cgagggattc ttgtgctgca aagtgcacaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaatg actggcatac	1080
	tggcaacaga tgcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140
	tagtgcgtaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccc	1200
	tagtggccca ggcatttget aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacgagat agacagttag tcatgggggtg ttgttgggct tttagaaggc	1320
	acaagataac atctatttat aagcggccgg ataccctaac catcatcaaa gtgaacagcg	1380
	atttcactc attcgtgctg cccaggatag gcagtaacac attggagatc gggtgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
[0047]	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gcccactctg gaagccgatg	1620
	tcgacttgat gtacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
	aggttaccag ctacgctggc gaggacaaga tcggctctta cgctgtgctt tctccgagg	1740
	ctgtactcaa gagtgaaaaa ttatcttgca tccacctctc cgctgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtgaa agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctggg gatcctccct tccatgaatt cgcctacgag agtctgagaa	2160
	cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaa agcgcagtca caaaaaaga tctagtgtg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataaggagc tcaagaaat gaaagggtg gacgtcaatg	2340

	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
	ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
	ctaaaaaggc agtgcctctgc ggggatccca aacagtgcgg tttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactaa atctgtgact tcggctctct caaccttgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagatttga ttgacactac cggcagtacc aaacctaagc	2700
	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggtc gaccctgaaa ggtgtgtatg	2820
	ccgttcggta caagtgtaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tctactgac cgcacggag gaccgcatcg tgtggaaaac actagccggc gaccatgga	2940
	taaaaacact gactgccaa gaccctggga atttcaactgc cacgatagag gagtggcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtcttcc	3060
	agaataagc aaacgtgtgt tggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
[0048]	tagacatgac cactgaacaa tggaacactg tggattattt tgaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactcga ctggactccg	3240
	gtctattttc tgcacccact gtccggttat ccattaggaa taatcaactgg gataactccc	3300
	cgtcgcctaa catgtacggg ctgaataaag aagtgtccg tcagctctct cgcaggtacc	3360
	cacaactgcc tcgggcagtt gccactggaa gactctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgata aacctagtag ctgtaaacag aagactgcct catgcttag	3480
	tctccacca taatgaacac ccacagagtg acttttcttc attcgctcagc aaattgaagg	3540
	gcagaactgt cctgtgtgct ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
	tgtcagaccg gcctgaggct accttcagag ctcggtgga tttaggcac ccaggtgatg	3660
	tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcac	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggtta cgctgacagg gccagcga	3840
	gcacattgg tgctatagcg cggcagttca agttttccc ggtatgcaaa ccgaaatcct	3900

	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
	acaatcctta caagetttca tcaaccttga ccaacattta tacaggttcc agactccacg	4020
	aagccggatg tgcaccctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tggtaaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgct aagattgtca	4320
	acgataacaa ttacaagtca gtagcgattc cactgttgtc caccggcacc tttccggga	4380
	acaaagatcg actaacccaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
	cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgagggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
[0049]	caagcgatgg caaaactttc tcatatattg aagggaccaa gtttcaccag gcggccaag	4680
	atatagcaga aattaatgcc atgtggcccc ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatcct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
	aagcctccac accacctagc acgtgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcaact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccgtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcacc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggccccgac ccaccaggtg ctgcaagtcg	5220
	aggcagacat tcacgggccg ccctctgtat ctagctcacc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacaccct ggaggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400

	gaccggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgt cgcgcacaaa	5460
	gaacaccgtc acttgcaccc agcagggcct gctcgagaac cagcctagtt tccacccgc	5520
	caggcgtgaa tagggtgatc actagagagg agctcgagge gettaccocg tcacgcactc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaaccc gccaggcgta aatagggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
	catacatctt ttcctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
	cgggtctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatccc acacctgcta	5880
	acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacce	6000
	tgcatcctgt tcctttgtat tcatctagtg tgaacctgc cttttcaagc cccaaggctc	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
[0050]	ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
	ccacaatagc atcggcagtg ccttcagcga tccagaacac gctccagaac gtcttggcag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgta ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaactg	6420
	ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggacaaa agctgctgct ctttttgcga agacacataa tttgaatatg ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccagge tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
	tccagcctgg ggattgtggt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatgac tetgaccgag ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgctgat tgaggcggct ttcggcgaaa tttcatcaat acatttgcce actaaaacta	6960

	aatttaaatt cggagccatg atgaaatctg gaatgttctt cactctgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acagtgccgc cacctggttg aatatggaag tcaagattat agatgctgtg gtggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga cccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
	gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaccgta ggaacttcca	7440
	tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgcaa	7560
	gatgttcccg ttccagccaa tgtatccgat gcagccaatg ccctatcga acccgttcgc	7620
	ggccccgcgc aggccctggt tccccagaac cgaccctttt ctggcgtatgc aggtgcagga	7680
[0051]	attaaccgcg tcgatggcta acctgacgtt caagcaacgc cgggacgcgc cacctgaggg	7740
	gccatccgct aagaaaccga agaaggaggc ctcgcaaaaa cagaaagggg gaggccaagg	7800
	gaagaagaag aagaaccaag ggaagaagaa ggctaagaca gggccgccta atccgaaggc	7860
	acagaatgga aacaagaaga agaccaacaa gaaaccaggc aagagacagc gcatggtcat	7920
	gaaattggaa tctgacaaga cgttccaat catgttggaa gggaagataa acggctacgc	7980
	ttgtgtggtc ggagggaagt tattcaggcc gatgcatgtg gaaggcaaga tcgacaacga	8040
	cgttctggcc gcgcttaaga cgaagaaagc atccaaatac gatcttgagt atgcagatgt	8100
	gccacagaac atgcccggcg atacattcaa atacacccat gagaaacccc aaggctatta	8160
	cagctggcat catggagcag tccaatatga aatgggcgt ttcacggtgc cgaaaggagt	8220
	tggggccaag ggagacagcg gacgacccat tctggataac cagggacggg tggctcctat	8280
	tgtgctggga ggtgtgaatg aaggatctag gacagccctt tcagtcgtca tgtggaacga	8340
	gaaggagatt accgtgaagt atactccgga gaactgcgag caatggtcac tagtgaccac	8400
	catgtgtctg ctgcctaatg tgacgttccc atgtgctcaa ccaccaattt gctacgacag	8460

	aaaaccagca gagactttgg ccatgctcag cgtaacggt gacaaccgg gctacgatga	8520
	gctgctggaa gcagctgtta agtgccccg aaggaagagg agatccaccg aggagctgtt	8580
	taaggagtat aagctaacgc gcccttacat ggccagatgc atcagatgtg cagttgggag	8640
	ctgccatagt ccaatagcaa tcgaggcagt aaagagcgac gggcacgacg gttatgttag	8700
	acttcagact tcctcgcagt atggcctgga ttcctccggc aacttaaagg gcaggacat	8760
	gcggtatgac atgcacggga ccattaaaga gataccata catcaagtgt cactccatac	8820
	atctcggccg tgtcacattg tggatgggca cggttatttc ctgcttgcca ggtgcccggc	8880
	aggggactcc atcaccatgg aatttaagaa agattccgtc acacactcct gctcggtgcc	8940
	gtatgaagtg aaatttaatc ctgtaggcag agaactctat actcatcccc cagaacacgg	9000
	agtagagcaa gcgtgccaa tctacgcaca tgatgcacag aacagaggag cttatgtcga	9060
	gatgcacctc cgggctcag aagtggacag cagtttggtt tccttgagcg gcagttcagt	9120
	caccgtgaca cctcctgttg ggactagcgc cctggtggaa tgcgagtggt gcggcacaaa	9180
	gatctccgag accatcaaca agacaaaaca gttcagccag tgcacaaaaga aggagcagtg	9240
[0052]	cagagcatat cggctgcaga acgataagtg ggtgtataat tctgacaaac tgcccaaagc	9300
	agcgggagcc acctaaaag gaaaactgca tgtccattc ttgctggcag acggcaaatg	9360
	caccgtgcct ctagcaccag aacctatgat aaccttgggt ttcagatcag tgcactgaa	9420
	actgcaccct aagaatccca catatctaac caccgcca cttgctgatg agcctcacta	9480
	cacgcacgag ctcatatctg aaccagctgt taggaatttt accgtcaccg aaaaagggtg	9540
	ggagtttgta tgggaaacc acccgccgaa aaggttttgg gcacaggaaa cagcaccgg	9600
	aatccacat gggctaccgc acgagtgat aactcattat taccacagat acctatgtc	9660
	caccatctg ggtttgtcaa tttgtgccgc cattgcaacc gtttccgttg cagcgtctac	9720
	ctggctgttt tgcagatcta gagttgcgtg cctaactcct taccggctaa cacctaacgc	9780
	taggatacca tttgtctg ctgtgctttg ctgcgccgc actgcccggg ccgagaccac	9840
	ctgggagtcc ttggatcacc tatggaacaa taaccaacag atgttctgga ttcaattgct	9900
	gatccctctg gccgccttga tcgtagtac tcgctgctc aggtgcgtgt gctgtgtcgt	9960
	gcctttttta gtcatggccg gcgccgagg cgccggcgc tacgagcag cgaccacgat	10020

	gccgagccaa gcggaatct cgtataacac tatagtcaac agagcaggt acgcaccact	10080
	ccctatcagc ataacaccaa caaagatcaa gctgatacct acagtgaact tggagtacgt	10140
	cacctgccac tacaaaaacag gaatggattc accagccatc aaatgctgcg gatctcagga	10200
	atgcaactca acttacaggc ctgatgaaca gtgcaaagtc ttcacagggg tttaccggt	10260
	catgtggggt ggtgcatatt gcttttgcga cactgagaac acccaagtca gcaaggccta	10320
	cgtaatgaaa tctgacgact gccttgcgga tcatgctgaa gcatataaag cgcacacagc	10380
	ctcagtgcag gcgttcctca acatcacagt gggagaacac tctattgtga ctaccgtgta	10440
	tgtgaatgga gaaactcctg tgaatttcaa tggggtcaaa ttaactgcag gtccgctttc	10500
	cacagcttgg acacccttg atcgcaaaat cgtgcagtat gccggggaga tctataatta	10560
	tgattttcct gagtatgggg caggacaacc aggagcattt ggagatatac aatccagaac	10620
	agtetcaage tcagatctgt atgccaatac caacctagtg ctgcagagac ccaaagcagg	10680
	agcgatccac gtgccataca ctcaggcacc ttcgggtttt gagcaatgga agaaagataa	10740
[0053]	agctccatca ttgaaattta cgccccctt cggatgcgaa atatatacaa accccattcg	10800
	cgccgaaaac tgtgctgtag ggtcaattcc attagcctt gacattccc acgcttgtt	10860
	caccagggtg tcagaaacac cgacacttcc agcgccgaa tgcactctta acgagtgcgt	10920
	gtattcttcc gactttggtg ggatgccac ggtcaagtac tcggccagca agtcaggcaa	10980
	gtgctcagtc catgtgccat cagggactgc taccctaaaa gaagcagcag tcgagctaac	11040
	cgagcaaggg tcggcgacta tccatttctc gaccgcaaat atccaccgg agtccaggct	11100
	ccaaatatgc acatcatatg ttactgcaa aggtgattgt cacccccga aagaccatat	11160
	tgtgacacac cctcagtatc acgcccacac atttacagcc gcggtgtcaa aaaccgctg	11220
	gacgtggtta acatccctgc tgggaggatc agccgtaatt attataattg gcttgggtgt	11280
	ggctactatt gtggccatgt acgtgctgac caaccagaaa cataattgaa tacagcagca	11340
	attgcaagc tgcttacata gaactcgcgg cgattggcat gccgcctaa aatttttatt	11400
	ttatTTTTTc ttttcttttc cgaatcggat tttgttttta atatttc	11447

<210> 4

<211> 9577
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多核苷酸

<400> 4

```

atgggcggcg catgagagaa gccagacca attacctacc caaatggag aaagttcacg      60
ttgacatcga ggaagacagc ccattcctca gagctttgca gggagcttc ccgagtttg      120
aggtagaagc caagcaggtc actgataatg accatgetaa tgccagagcg ttttcgcatc      180
tggttcaaaa actgatcgaa acggagggtg acccatccga cacgatcctt gacattggaa      240
gtgcgccccg ccgcagaatg tattctaagc acaaglatca ttgtatctgt ccgatgagat      300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg      360
aaataactga taaggaattg gacaagaaaa tgaaggagct ccccgccgtc atgagcgacc      420
ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc      480
aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc taccaccaag      540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta      600
agaacttggc tggagcatat ccatacact ctaccaactg gcccagcga accgtgttaa      660
cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt      720
ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga      780
ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact      840
tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg      900
tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta      960
cgatgcaccg cgagggatc ttgtgctgca aagtacaga cacattgaac ggggagaggg      1020
tctettttcc cgtgtgcacg tatgtgccag ctacattgtg tgacaaatg actggcatac      1080
tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta      1140
tagtcgtaa cggtcgacc cagagaaaca ccaataccat gaaaaattac cttttgcccg      1200
tagtggccca ggcatttgct aggtgggcaa agaatataa ggaagatcaa gaagatgaaa      1260
ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc      1320
  
```

[0054]

	acaagataac atctatttat aagcgcccgg ataccctaac catcatcaaa gtgaacagcg	1380
	atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgtc acctctcatt accgcccagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgctgtaa gccgaggagt	1560
	tgcgcgagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgacttgat gttacaagag gctggggccg gctcagtga gacacctctg ggcttgataa	1680
	aggttaccag ctacgtggc gaggacaaga tcggctctta cgctgtgctt tctccgagc	1740
	ctgtactcaa gagtgaaaaa ttatcttga tccacctct cgctgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggegttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtga agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcctgaa cactgatgaa gaatattaca aaactgtcaa gccagcagc cagcagcg	2040
[0055]	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctggg gatcctccct tccatgaatt cgcctacgag agtctgagaa	2160
	cagcaccagc cgctccttac caagtacaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctagtggg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataaggagc tcaagaaat gaaagggtg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca cccgtagag accctgtata	2400
	ttgacgaagc ttttcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
	ctaaaaagc agtgcctgc ggggatccca aacagtgcgg ttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttga cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactaa atctgtgact tcggtcgtct caaccttgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagatttga ttgacactac cggcagtacc aaacctaagc	2700
	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgaa atagattaca	2760
	aaggcaacga aataatgac gcagctgct ctcaagggt gaccgtaaa ggtgtgtatg	2820

	ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tcctactgac cgcacggag gaccgcatcg tgtggaaaac actagccggc gaccatgga	2940
	taaaaaact gactgccaag taccctggga atttcaactgc cacgatagag gagtggcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtcttcc	3060
	agaataaggc aaacgtgtgt tggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattattt tgaaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactgat ctggactccg	3240
	gtctattttc tgcacccact gttccgttat ccattaggaa taactactgg gataactccc	3300
	cgctgcctaa catgtacggg ctgaataaag aagtgggccg tcagctctct cgcaggtacc	3360
	cacaactgcc tgggacgatt gccactggaa gactctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
	tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcage aaattgaagg	3540
	gcagaactgt cctgggtggtc ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
[0056]	tgtcagaccg gctgaggct accttcagag ctcggtgga tttaggcatc ccaggtgatg	3660
	tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcac	3780
	tgaatcccgg cggaaacctgt gtcagcatag gttatggta cgtgacagg gccagcghaa	3840
	gcatcattgg tgctatagcg cggcagttca agttttccc ggtatgcaa cggaaatcct	3900
	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgctacg	3960
	acaatcctta caagctttca tcaacctga ccaacattta tacaggttcc agactccacg	4020
	aagccgatg tgcacctca tatcatgttg tgcgagggga tattgccacg gccaccgaag	4080
	gagtattat aatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccghaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tgttcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc acaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgt aagattgtca	4320
	acgataacaa ttacaagtca gtagegattc cactgttgtc caccggcatc ttttccghaa	4380

	acaaagatcg actaacccaa tcattgaacc atttgctgac agctttagac accactgatg	4440
	cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgagggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaactttc tcatatattg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatacct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
	aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatcaa ggaagtatct cgtggaaca ccaccggtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
[0057]	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcacc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggcccagc ccaccaggtg ctgcaagtcg	5220
	aggcagacat tcacgggccc cctctgtat ctagctcacc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacacct ggagggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgaaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgtc ccgcgcacaa	5460
	gaacaccgtc acttgcacc agcaggcct gctcgagaac cagcctagtt tccaccccgc	5520
	caggcgtgaa tagggtgacc actagagagg agctcgagge gcttaccctg tcacgcactc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaacc gccaggcgtg aatagggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
	catacatctt ttcctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
	cggtgctatc cgaagtgtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaattccc acacctgcta	5880

	acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacce	6000
	tgcatcctgt tectttgtat tcatctagtg tgaaccgtgc cttttcaage cccaaggtcg	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
	ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcttggcag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgta ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt	6420
	ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggaccaa agctgctgct ctttttgcga agacacataa ttgaaatag ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660
[0058]	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
	tccagcctgg ggattgtgtt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatgac tctgaccgag ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgtgat tgaggcggct ttcggcgaaa tttcatcaat acatttgecc actaaaacta	6960
	aatttaaatt cggagccatg atgaaatctg gaatgttctt cacactgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga cccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
	gtattcttcc agagctgtgc aaggcagtag aatcaaggta tgaaccgta ggaacttcca	7440

	tcatagttat ggccatgact actctagcta gcagtgtaa atcattcagc tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gactctagaa tagtctttaa	7560
	ttaagccacc atggcaggca tgtttcaggc gctgagcгаа ggctgcaccc cgtatgatat	7620
	taaccagatg ctgaacgtgc tgggcgatca tcaggctca ggccttgagc agcttgagag	7680
	tataatcaac tttgaaaaac tgactgaatg gaccagttct aatgttatgc ctatcctgtc	7740
	tcctctgaca aagggcaccc tgggcttcgt gttaccctg accgtgcctt ctgagagagg	7800
	acttagctgc attagcgaag cggatgcgac caccccgaa agcgcgaacc tgggcgaaga	7860
	aattctgagc cagctgtatc tttggccaag ggtgacctac cattccccta gttatgctta	7920
	ccaccaatth gaaagacgag ccaaataaa aagacacttc cccgctttg gccagagcct	7980
	gctgtttggc taccctgtgt acgtgttcgg cgattgcgtg cagggcgatt gggatgcgat	8040
	tcgctttcgc tattgcgcgc cgccgggcta tgcctgctg cgctgcaacg ataccaacta	8100
	tagcctctg ctggctgtgg gggccctaga aggaccagc aatcaggact ggcttggtgt	8160
[0059]	cccaagacaa cttgtaactc ggatgcaggc tattcagaat gccggcctgt gtaccctggt	8220
	ggccatgctg gaagagacaa tcttctggct gcaagcgth ctgatggcgc tgaccgatag	8280
	cggcccgaaa accaacatta ttgtggatag ccagtatgtg atgggcatta gcaaaccgag	8340
	ctttcaggaa tttgtggatt gggaaaacgt gagcccggaa ctgaacagca ccgatcagcc	8400
	gttttgcaa gccggaatcc tggccagaaa tctggtgcct atggtggcca cagtgcaggg	8460
	ccagaacctg aagtaccagg gtcagtcact agtcatctct gcttctatca ttgtcttcaa	8520
	cctgctggaa ctggaaggctg attatcgaga tgatggcaac gtgtgggtgc ataccctgct	8580
	gagcccgcgc accctgaacg cgtgggtgaa agcggtgгаа gaaaaaaaaag gtattccagt	8640
	tcacctagag ctggccagta tgaccaacat ggagctcatg agcagtattg tgcatcagca	8700
	ggtcagaaca tacggccccg tgttcatgtg tctcggcгаа ctgcttacia tgggtgctgg	8760
	tgctgtgtgg ctgacagtgc gagtgctcga gctgttccgg gccgcgcagc tggccaacga	8820
	cgtggtctc cagatcatgg agctttgtgg tgcagcgtht cgccaggtgt gccataccac	8880
	cgtgccgtgg ccgaacgcga gcctgacccc gaaatggaac aacgaaacca cccagcccca	8940

gatcgccaac tgcagcgtgt atgacttttt tgtgtggctc cattattatt ctgttcgaga 9000
cacactttgg ccaaggtga cctaccatat gaacaaatat gcgtatcata tgctggaaag 9060
acgagccaaa tataaaagag gaccaggacc tggcgctaaa tttgtggccg cctggacact 9120
gaaagccgct gctggctctg gacctggcca gtacatcaag gccaacagca agttcatcgg 9180
catcaccgaa ctcggacccg gaccaggctg atgattcgaa cggccgtatc acgccc aaac 9240
atttacagcc gcggtgtcaa aaaccgctg gacgtggtta acatccctgc tgggaggatc 9300
agccgtaatt attataattg gcttgggtct ggctactatt gtggccatgt acgtgctgac 9360
caaccagaaa cataattgaa tacagcagca attggcaagc tgcttacata gaactcgcgg 9420
cgattggcat gccgccttaa aatTTTTATT ttatTTTTTc ttttcttttc cgaatcggat 9480
tttgttttta atatttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 9540
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 9577

<210> 5

<211> 11447

<212> DNA

<213> 委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus)

<400> 5

atggcgcgcg catgagagaa gccagacca attacctacc caaatggag aaagttcacg 60
ttgacatcga ggaagacagc ccattctca gagctttgca gggagcttc ccgagtttg 120
aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc 180
tggttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa 240
gtgcgcccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat 300
gtcggaaga tccgacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg 360
aaataactga taaggaattg gacaagaaaa tgaaggagct cggcccgctc atgagcgacc 420
ctgacctgga aactgagact atgtgectcc acgacgacga gtcgtgtcgc tacgaagggc 480
aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc taccaccaag 540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta 600
agaacttggc tggagcatat ccatcact ctaccaactg ggccgacgaa accgtgttaa 660

[0060]

	cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttcaact	840
	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcagge tatgctgcta	960
	cgatgcaccg cgagggatte ttgtgctgca aagtgcaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
	tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgtg	1140
	tagtctgtaa cggctgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
	tagtggccca ggcatttgcg aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
	acaagataac atctatttat aagcgcgccg atacccaaac catcatcaa gtgaacagcg	1380
[0061]	atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc ggctgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgct acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctctg ggcttgataa	1680
	aggttaccag ctacgctggc gaggacaaga tcggctctta cgtctgtgctt tctccgcagg	1740
	ctgtactcaa gagtgaaaaa ttatcttgca tccacctctc cgctgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtgaa agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctggtg gatcctccct tccatgaatt cgctacgag agtctgagaa	2160
	cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220

	gcaagtctgg catcattaa agcgcagtca ccaaaaaaga tctagtgggtg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataagggacg tcaagaaaat gaaagggctg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
	ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
	ctaaaaaggc agtgctctgc ggggatccca aacagtgcgg tttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgcaactaa atctgtgact tcggtcgtct caaccttgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtacc aaacctaagc	2700
	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggtc gaccctgaaa ggtgtgtatg	2820
	ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tcctactgac ccgcacggag gaccgcatcg tgtggaaaac actagccggc gaccctatgga	2940
[0062]	taaaaaact gactgccaag taccctggga atttactgc cacgatagag gactggcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggacctacc gacgtcttcc	3060
	agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattattt tgaaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240
	gtctatitct tgcacceact gtccggttat ccattaggaa taatcactgg gataactccc	3300
	cgtcgcctaa catgtacggg ctgaataaag aagtggctcc tcagctctct cgcaggtacc	3360
	cacaactgcc tcgggcagtt gccactggaa gactctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgcata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
	tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
	gcagaactgt cctgggtggtc ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
	tgtcagaccg gcctgaggct accttcagag ctccgctgga tttaggcatc ccaggtgatg	3660
	tgcccaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720

	agcagtgatga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcatc	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggta cgctgacagg gccagcga	3840
	gcatcattgg tgctatagcg cggcagttca agttttcccg ggtatgcaaa ccgaaatcct	3900
	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
	acaatectta caagctttca tcaacctga ccaacattta tacaggttcc agactccacg	4020
	aagccggatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcgagcgc	4140
	tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tgttcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc acaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgt aagattgtca	4320
	acgataacaa ttacaagtca gtagcgattc cactgttgc caccggcatc tttccggga	4380
	acaagatcg actaaccaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
	cagatgtagc cataactgc agggacaaga aatgggaaat gactctcaag gaagcagtg	4500
[0063]	ctaggagaga agcagtgag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgaggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaacttcc tcataatttg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatacct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
	aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggcccagc ccaccaggtg ctgcaagtgc	5220
	aggcagacat tcacgggccc ccctctgtat ctagctcatc ctggtccatt cctcatgcat	5280

	ccgactttga tgtggacagt ttatccatac ttgacaccct ggagggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaaccctcc acatcccgt cgcgcacaaa	5460
	gaacaccgtc acttgcaccc agcagggcct gctcgagaac cagcctagtt tccaccccgc	5520
	caggcgtgaa tagggtgatc actagagagg agctcgagge gcttaccgct tcacgcactc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaacc gccaggcgtg aatagggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
	catacatctt ttcctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
	cgggtctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gcccgcgcc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatecc acacctgcta	5880
	acagaagcag ataccagtcc aggaagggtg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacc	6000
[0064]	tgcacctgt tcctttgtat tcactagtg tgaaccgtgc cttttcaagc cccaaggtcg	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaa	6240
	ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcttgccag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgtg ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaactg	6420
	ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggacaaa agctgctgct ctttttgcga agacacataa tttgaatatg ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtagcag tgatccaggc tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780

	tccagcctgg ggattgtggt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgtgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
	aatttaaatt cggagccatg atgaaatctg gaatgttctt cacactgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcacatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga ccccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
	gtattcttcc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
	tcatagttat ggccatgact actctagcta gcagtgttaa atcattcage tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgcaa	7560
[0065]	gatgttcccg ttccagccaa tgtatccgat gcagccaatg ccctatcgca acccgttcgc	7620
	ggccccgcgc aggccctggt tccccagaac cgaccctttt ctggcgatgc aggtgcagga	7680
	attaaccgcg tcgatggcta acctgacggt caagcaacgc cgggacgcgc cacctgaggg	7740
	gccatccgct aagaaaccga agaaggaggc ctcgcaaaaa cagaaagggg gaggccaagg	7800
	gaagaagaag aagaaccaag ggaagaagaa ggctaagaca gggccgccta atccgaagc	7860
	acagaatgga aacaagaaga agaccaacaa gaaaccaggc aagagacagc gcatggtcat	7920
	gaaattgaa tctgacaaga cgttccaat catgttgaa gggaagataa acggctacgc	7980
	ttgtgtggtc ggaggaagt tattcaggcc gatgcatgtg gaaggcaaga tcgacaacga	8040
	cgttctggcc gcgcttaaga cgaagaaagc atccaaatac gatcttgagt atgcagatgt	8100
	gccacagaac atgcgggccg atacattcaa atacacccat gagaaacccc aaggctatta	8160
	cagctggcat catggagcag tccaatatga aatgggcgt ttcacggtgc cgaaaggagt	8220
	tggggccaag ggagacagcg gacgacccat tctggataac cagggacggg tggtcgctat	8280
	tgtgctggga ggtgtgaatg aaggatctag gacagccctt tcagtcgtca tgtggaacga	8340

	gaagggagtt accgtgaagt atactccgga gaactgcgag caatggtcac tagtgaccac	8400
	catgtgtctg ctgcceaatg tgacgttccc atgtgetcaa ccaccaattt gctacgacag	8460
	aaaaccagca gagactttgg ccatgctcag cgtaaactgt gacaacccgg gctacgatga	8520
	gctgctggaa gcagctgtta agtgccccgg aaggaaaagg agatccaccg aggagctgtt	8580
	taaggagtat aagctaacgc gcccttacat gcccagatgc atcagatgtg cagttgggag	8640
	ctgccatagt ccaatagcaa tcgaggcagt aaagagcgac gggcacgacg gttatgttag	8700
	acttcagact tcctcgcagt atggcctgga ttctcggc aacttaaagg gcaggacat	8760
	gcggtatgac atgcacggga ccattaaaga gataccacta catcaagtgt cactccatac	8820
	atctcggccg tgtcacattg tggatgggca cggttatttc ctgcttgcca ggtgccccgg	8880
	aggggactcc atcacatgg aatttaagaa agattccgtc acacactcct gctcggtgcc	8940
	gtatgaagtg aaatttaatc ctgtaggcag agaactctat actcatcccc cagaacacgg	9000
	agtagagcaa gcgtgccaag tctacgcaca tgatgcacag aacagaggag cttatgtcga	9060
[0066]	gatgcacctc ccgggctcag aagtggacag cagtttggtt tccttgagcg gcagttcagt	9120
	caccgtgaca cctcctgttg ggactagcgc cctggtggaa tgcgagtgtg gcggcacaaa	9180
	gatctccgag accatcaaca agacaaaaca gttcagccag tgcacaaaga aggagcagtg	9240
	cagagcatat cggtgcaga acgataagtg ggtgtataat tctgacaaac tgcccaaage	9300
	agcgggagcc accttaaaag gaaaactgca tgtcccattc ttgctggcag acggcaaatg	9360
	caccgtgcct ctagcaccag aacctatgat aacctttggt ttcagatcag tgtcactgaa	9420
	actgcacctt aagaatccca catatctaac caccgccaa cttgctgatg agcctcacta	9480
	cacgcacgag ctcatatctg aaccagctgt taggaatttt accgtcaccg aaaaagggtg	9540
	ggagtttgta tggggaacc acccgccgaa aaggttttgg gcacaggaaa cagcaccgg	9600
	aatccacat gggtaccgc acgaggtgat aactcattat taccacagat accctatgtc	9660
	caccatcctg ggtttgtcaa tttgtccgc cattgcaacc gtttccgttg cagcgtctac	9720
	ctggtgtttt tgcagatcta gagttgcgtg cctaactcct taccgctaa cacctaacgc	9780
	taggatacca ttttgtctgg ctgtgctttg ctgcgccccg actgccccgg ccgagaccac	9840

	ctgggagtcc ttggatcacc tatggaacaa taaccaacag atgttctgga ttcaattgct	9900
	gatccctctg gccgccttga tcgtagtgac tcgcctgctc aggtgcgtgt gctgtgtcgt	9960
	gcctttttta gtcatggccg gcgccgcagg cgccggcgcc tacgagcacg cgaccacgat	10020
	gccgagccaa gcgggaatct cgtataacac tatagtcaac agagcaggct acgcaccact	10080
	ccctatcagc ataacaccaa caaagatcaa gctgatacct acagtgaact tggagtacgt	10140
	cacctgccac taaaaacag gaatggattc accagccatc aaatgctgcg gatctcagga	10200
	atgcactcca acttacaggc ctgatgaaca gtgcaaagtc ttcacagggg tttaccggt	10260
	catgtggggt ggtgcatatt gcttttgca cactgagaac acccaagtca gcaaggccta	10320
	cgtaatgaaa tctgacgact gccttgcgga tcatgctgaa gcatataaag cgcacacagc	10380
	ctcagtgcag gcgttctca acatcacagt gggagaacac tctattgtga ctaccgtgta	10440
	tgtgaatgga gaaactcctg tgaatttcaa tggggtcaaa ttaactgcag gtccgctttc	10500
	cacagcttgg acacccttg atcgaaaat cgtgcagtat gccggggaga tctataatta	10560
[0067]	tgattttcct gagtatgggg caggacaacc aggagcattt ggagatatac aatccagaac	10620
	agtctcaage tcagatctgt atgccaatac caacctagtg ctgcagagac ccaaagcagg	10680
	agcgatccac gtgccataca ctcaggcacc ttcgggtttt gagcaatgga agaaagataa	10740
	agctccatca ttgaaattta ccgcccttt cggatgcaaa atatatacaa accccattcg	10800
	cgccgaaaac tgtgctgtag ggtcaattcc attagccttt gacattcccg acgccttgtt	10860
	caccaggtg tcagaaacac cgacaattc agcgccgaa tgcactctta acgagtgcgt	10920
	gtattcttcc gactttggtg ggatgccac ggtcaagtac tcggccagca agtcaggcaa	10980
	gtgcgcagtc catgtccat cagggactgc taccctaaa gaagcagcag tcgagctaac	11040
	cgagcaaggg tcggcgacta tccatttctc gaccgcaaat atccaccgg agttcaggct	11100
	ccaaatatgc acatcatatg ttacgtgcaa aggtgattgt caccctccga aagaccatat	11160
	tgtgacacac cctcagtatc acgcccacac atttacagcc gcggtgtcaa aaaccgctg	11220
	gacgtggtta acatccctgc tgggaggatc agccgtaatt attataattg gcttgggtgct	11280
	ggctactatt gtggccatgt acgtgctgac caaccagaaa cataattgaa tacagcagca	11340
	attggcaage tgcttacata gaactcgagg cgattggcat gccgccttaa aatttttatt	11400

	ttatTTTTtTc ttttctTTTTc cgaatcggat ttgtTTTTa atatttc	11447
	<210> 6	
	<211> 7894	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 6	
	atgggcggcg catgagagaa gcccagacca attacctacc caaaatggag aaagttcacg	60
	ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg	120
	aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcate	180
	tggettcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa	240
	gtgcgcccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
	gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
[0068]	aaataactga taaggaattg gacaagaaaa tgaaggagct cgccgccgtc atgagcgacc	420
	ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
	aagtcgctgt ttaccaggat gtatagcggg ttgacggacc gacaagtctc taccaccaag	540
	ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta	600
	agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa	660
	cggetcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttcaact	840
	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta	960
	cgatgcaccg cgagggattc ttgtgctgca aagtgcacaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
	tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140

	tagtcgtcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccc	1200
	tagtggccca ggcatattgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacgagat agacagttag tcatgggggtg ttgttgggct tttagaaggc	1320
	acaagataac atctatattat aagcgccccg atacccaaac catcatcaaa gtgaacagcg	1380
	atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
	caagaatcag gaaaatgtta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgacttgat gttacaagag gctggggccg gctcagtga gacacctcg ggcttgataa	1680
	aggttaccag ctacgctggc gaggacaaga tcggctctta cgctgtgctt tctccgagc	1740
	ctgtactcaa gagtgaaaaa ttatcttgca tccacctct cgctgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
[0069]	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtga agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gcccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctggtg gatcctcct tccatgaatt cgctacgag agtctgagaa	2160
	cacgaccage cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctagtgggtg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataaggacg tcaagaaat gaaagggtg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaacaa ccccgtagag acctgtata	2400
	ttgacgaage ttttgcttgt catgcaggta ctctcagagc getcatagcc attataagac	2460
	ctaaaaaggc agtgctctgc ggggatccca aacagtgcgg ttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactaa atctgtgact tcggctgtct caacctgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagatttga ttgacactac cggcagtacc aaacctaagc	2700

	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggtc gaccctaaa ggtgtgtatg	2820
	ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tctactgac ccgcacggag gaccgcatcg tgtggaaaac actagccggc gaccatgga	2940
	taaaaaact gactgccaag taccctggga atttactgc cacgatagag gagggtgcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gactcttcc	3060
	agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattatct tgaaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgag tgaggttctt tggactcgat ctggactccg	3240
	gtctatcttc tgcacccact gttccgttat ccattaggaa taactactgg gataactccc	3300
	cgctgcctaa catgtacggg ctgaataaag aagtggctcc tcagctctct cgcaggtacc	3360
	cacaactgcc tcgggcagtt gccactggaa gactctatga catgaacact ggtacactgc	3420
[0070]	gcaattatga tccgcgcata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
	tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
	gcagaactgt cctggtggtc ggggaaaagt tgcctgtccc aggcaaatg gttgactggt	3600
	tgtcagaccg gcctgagget accttcagag ctccgctgga tttaggcatc ccaggtgatg	3660
	tgcccaaata tgacataata tttgtaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcatc	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggta cgctgacagg gccagcgaag	3840
	gcatcattgg tgctatagcg cggcagttca agttttccc ggtatgcaaa ccgaaatcct	3900
	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgaag gccctgacgc	3960
	acaatcctta caagcttca tcaaccttga ccaacattta tacagttcc agactccacg	4020
	aagccggatg tgcaccctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200

	tggtcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgct aagattgtca	4320
	acgataacaa ttacaagtca gtagcgattc cactgttgte caccggcattc ttttccggga	4380
	acaaagatcg actaacccaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
	cagatgtagc catafactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgaggggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaaactttc tcatatttgg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatacct cggagaaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
	aagcctccac accacctage acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
[0071]	caccgaaagt gctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcattc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggcccgac ccaccaggtg ctgcaagtcg	5220
	aggcagacat tcacgggccc ccctctgtat ctagctcattc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacaccct ggaggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaaccctcc acatcccgtc ccgcgcacaa	5460
	gaacaccgtc acttgcaccc agcagggcct gctcgagaac cagcctagtt tccacccgc	5520
	caggcgtgaa tagggtgatc actagagagg agctcgaggc gcttaccgctc tcacgcactc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaacct gccaggcgta aataggggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcgggtg	5700
	catacatctt ttctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760

	cggtgctatc cgaagtgggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaateccc acacctgcta	5880
	acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacce	6000
	tgcatcctgt tcctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggtcg	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
	ccacaatacg atcggcagtg ctttcagcga tccagaacac gctccagaac gtcctggcag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aatgagaga attgcccgta ttggattcgg	6360
	cggccttta tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt	6420
	ttaaagaaaa ccccatcagg cttactgaag aaaacgtggg aaattacatt accaaattaa	6480
[0072]	aaggacaaa agctgctgct ctttttgcga agacacataa tttgaatatg ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccagge tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
	tccagcctgg ggattgtgtt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatggc tctgaccgcg ttaatgattc tgaagactt aggtgtggac gcagagctgt	6900
	tgacgtgat tgaggcgget ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
	aatttaaatt cggagccatg atgaaatctg gaatgttctt cactgttt gtgaacacag	7020
	tcattaacat tgtaatcgea agcagagtgt tgagagaac gctaaccgga tcacatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260

gtgtggcaga ccccctaaaa aggctgttta agcttggcaa acctctggca gcagacgatg 7320
aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg 7380
gtattctttc agagetgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca 7440
tcatagtatt ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag 7500
gggcccctat aactctctac ggctaacctg aatggactac gactatcacg cccaaacatt 7560
tacagccgcg gtgtcaaaaa ccgcgtggac gtggttaaca tccttctgg gaggatcagc 7620
cgtaattatt ataattggct tgggtctggc tactattgtg gccatgtacg tgctgaccaa 7680
ccagaaacat aattgaatac agcagcaatt ggcaagctgc ttacatagaa ctgcgcgca 7740
ttggcatgcc gccttaaaat ttttatttta tttttcttt tcttttccga atcggatitt 7800
gtttttaata tttcaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 7860
aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaa 7894

<210> 7

<211> 7893

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 7

ataggcggcg catgagagaa gcccagacca attacctacc caaaatggag aaagtccacg 60
ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg 120
aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc 180
tggcttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa 240
gtgcgcccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat 300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg 360
aaataactga taaggaattg gacaagaaaa tgaaggagct cgccgccgtc atgagcgacc 420
ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc 480
aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc taccaccaag 540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccct tttatgttta 600

[0073]

	agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa	660
	cggtctgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact	840
	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta	960
	cgatgcaccg cgagggattc ttgtgctgca aagtgacaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
	tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct gggttgggctc aaccagcgta	1140
	tagtctgcaa cggctgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
	tagtggccca ggcatttgcg aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
[0074]	acaagataac atctatttat aagcggccgg ataccctaac catcatcaaa gtgaacagcg	1380
	atttccactc attcgtgctg cccaggatag gcagtaaac attggagatc gggtgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgaactgat gtacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
	aggttaccag ctacgatggc gaggacaaga tcggctctta cgctgtgctt tctccgcagg	1740
	ctgtactcaa gagtgaaaaa ttatcttgea tccacctctc cgctgaacaa gtcatagtga	1800
	taacacactc tgccgaaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaage tctgagtgaa agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100

	ggctcacagg cgagctggtg gatcctcctt tccatgaatt cgcttacgag agtctgagaa	2160
	cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaa agcgcagtca ccaaaaaaga tctagtgggtg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataagggacg tcaagaaaat gaaagggtg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
	ttgacgaagc ttttgettgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
	ctaaaaaggc agtgcctctgc ggggatccca aacagtgcgg ttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttga cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactaa atctgtgact tcggctgtct caacctgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagatttga ttgacactac cggcagtacc aaacctaagc	2700
	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggt gacctgaaa ggtgtgtatg	2820
	ccgttcggta caagtgat gaaaatctc tgtacgcacc cacctcagaa catgtgaacg	2880
[0075]	tctactgac cgcacggag gaccgatcg tgtgaaaac actagccggc gaccatgga	2940
	taaaaaact gactgccaag taccctggga atttactgc cacgatagag gagtggcaag	3000
	cagagcatga tgccatcatg aggacatct tggagagacc ggaccctacc gacgtcttcc	3060
	agaataagc aaacgtgtgt tggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattattt tgaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240
	gtctattttc tgcaccact gttccgttat ccattaggaa taactactgg gataactccc	3300
	cgtcgcctaa catgtacggg ctgaataaag aagtgtccg tcagctctct cgcaggtacc	3360
	cacaactgcc tcgggcagtt gccactgga gagtctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgata aacctagtac ctgtaaacag aagactgcct catgctttag	3480
	tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
	gcagaactgt cctggtggtc ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
	tgtcagaccg gcctgaggct accttcagag ctcggtgga tttaggcac ccaggtgatg	3660

	tgcccaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcatc	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggtta cgctgacagg gccagcgaaa	3840
	gcatcattgg tgctatagcg cggcagttca agttttccc ggtatgcaaa ccgaaatcct	3900
	cacttgaaga gacggaagt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
	acaatcctta caagctttca tcaaccttga ccaacattta tacagttcc agactccacg	4020
	aagccggatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tggtaaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgt aagattgtca	4320
	acgataacaa ttacaagtca gtagcgattc cactgtttgc caccggcatc tttccggga	4380
[0076]	acaaagatcg actaacccaa tcattgaacc atttgetgac agctttagac accactgatg	4440
	cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagt acagaacctg	4560
	atgcagagct ggtgagggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaactttc tcatatttgg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatcct cggagaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtctgg	4800
	aagcctccac accacctagc acgtgcctt gettgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgetca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160

	aagaagagga tagcataagt ttgctgtcag atggcccagc ccaccaggtg ctgcaagtcg	5220
	aggcagacat tcacgggccg ccctctgtat ctagctcadc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacaccct ggagggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaaccctcc acatcccgtc ccgcgcacaa	5460
	gaacaccgtc acttgacccc agcaggcctt gctcgagaac cagcctagtt tccaccccgc	5520
	caggcgtgaa tagggtgadc actagagagg agctcagagg gcttaccctg tcacgcactc	5580
	ctagcaggtc ggtctcagaa accagcctgg tctccaaccc gccaggcgtg aatagggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcggtg	5700
	catacatctt ttctccgac accggtcaag ggcatattaca acaaaaatca gtaaggcaaa	5760
	cgggtctadc cgaagtgggt ttggagagga ccgaattgga gatttcgtat gccccgcgcc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatccc acacctgcta	5880
[0077]	acagaagcag ataccagtcc aggaagggtg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacce	6000
	tgcatcctgt tcctttgtat tcatctagtg tgaaccgtgc cttttcaagc cccaaggctg	6060
	cagtggaaag ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgccctgca aagctgcgca gctttccaaa gaaacactcc tatttggaac	6240
	ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtcctggcag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aatgagaga attgccgta ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaactg	6420
	ttaaagaaaa ccccatcagg ctactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggacaaa agctgctgct ctttttgca agacacataa ttgaaatag ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccagcc tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aatgcggtc ctgcttccga	6720

	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
	tccagcctgg ggatttggtt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgtgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
	aatttaaatt cggagccatg atgaaatctg gaatgttctt cacactgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtgggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga cccctaaaa agctgtttaa agcttgcaa acctctggca gcagacgatg	7320
	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgettg aaccgagtgg	7380
	gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaaccgta ggaacttcca	7440
[0078]	tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gactatcacg cccaaacatt	7560
	tacagccgcg gtgtcaaaaa ccgcgtggac gtggtaaca tcctgtctgg gaggatcagc	7620
	cgtaattatt ataattggct tgggtctggc tactattgtg gccatgtacg tgetgaccaa	7680
	ccagaaacat aattgaatac agcagcaatt ggcaagctgc ttacatagaa ctgcgcgca	7740
	ttggcatgcc gccttaaaat ttttatttta tttttctttt cttttccgaa tcggattttg	7800
	tttttaatat ttcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	7860
	aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa	7893
	<210> 8	
	<211> 7927	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 8	

	taatacgact cactatagga tgggcggcgc atgagagaag cccagaccaa ttacctacce	60
	aaaatggaga aagttcacgt tgacatcgag gaagacagcc cattcctcag agctttgcag	120
	eggagcttcc cgcagtttga ggtagaagcc aagcaggtca ctgataatga ccatgctaat	180
	gccagagcgt tttcgcactt ggcttcaaaa ctgatcgaaa cggagggtga cccatccgac	240
	acgatecctg acattggaag tgcgcccgcc cgcagaatgt attctaagca caagtatcat	300
	tgtatctgtc cgatgagatg tgcggaagat ccggacagat tgtataagta tgcaactaag	360
	ctgaagaaaa actgtaagga aataactgat aaggaattgg acaagaaaaat gaaggagctc	420
	gccgccgtca tgagcgacct tgacctggaa actgagacta tgtgcctcca cgacgacgag	480
	tcgtgtcgtc acgaagggca agtcgctgtt taccaggatg tatacgcggt tgacggaccg	540
	acaagtctct atcaccaage caataaggga gttagagtcg cctactggat aggcctttgac	600
	accacccctt ttatgtttaa gaacttggtt ggagcatatc catcatactc taccaactgg	660
	gccgacgaaa ccgtgttaac ggctcgtaac atagccctat gcagctctga cgttatggag	720
	cggtcacgta gagggatgtc cattcttaga aagaagtatt tgaaacctac caacaatgtt	780
[0079]	ctattctctg ttgctcgac catctaccac gagaagaggg acttactgag gagctggcac	840
	ctgccgtctg tatttcactt acgtggcaag caaaattaca catgtcggtg tgagactata	900
	gttagttgcg acgggtacgt cgttaaaaga atagctatca gtccaggcct gtatgggaag	960
	ccttcaggct atgtctctac gatgcaccgc gagggattct tgtgctgcaa agtgacagac	1020
	acattgaacg gggagagggt ctcttttccc gtgtgcacgt atgtgccagc tacattgtgt	1080
	gaccaaatac ctggcatact ggcaacagat gtcagtgcgg acgacgcgca aaaactgctg	1140
	gttgggctca accagcgtat agtcgtcaac ggtcgcacc cagagaaacac caataccatg	1200
	aaaaattacc ttttgccgt agtggcccag gcatttgcta ggtgggcaaa ggaatataag	1260
	gaagatcaag aagatgaaag gccactagga ctacgagata gacagttagt catgggggtg	1320
	tgttgggctt ttagaaggca caagataaca tctatttata agcggcccga tacccaaacc	1380
	atcatcaaag tgaacagcga tttcactca ttcgtgctgc ccaggatagg cagtaacaca	1440
	ttggagatcg ggctgagaac aagaatcagg aaaatgttag aggagcaca ggagccgtca	1500
	cctctcatta ccgccgagga cgtacaagaa gctaagtgcg cagccgatga ggctaaggag	1560

	gtgcgtgaag ccgaggagtt ggcgcagct ctaccacctt tggcagctga tgttgaggag	1620
	cccactctgg aagccgatgt cgacttgatg ttacaagagg ctggggccgg ctcaagtggag	1680
	acacctcgtg gcttgataaa ggttaccagc tacgctggcg aggacaagat cggctcttac	1740
	gctgtgcttt ctccgcaggc tgtactcaag agtgaaaaat tatcttgcac ccaccctctc	1800
	gctgaacaag tcatagtgat aacacactct gcccgaaaag ggcgttatgc cgtggaacca	1860
	taccatggta aagtagtggg gccagaggga catgcaatac ccgtccagga ctttcaagct	1920
	ctgagtghaa gtgccacat tgtgtacaac gaacgtgagt tcgtaaacag gtacctgcac	1980
	catattgcca cacatggagg agcgtgaac actgatgaag aatattacaa aactgtcaag	2040
	cccagcgagc acgacggcga atacctgtac gacatcgaca ggaaacagtg cgtcaagaaa	2100
	gaactagtca ctgggctagg gctcacaggc gagctggtgg atcctccctt ccatgaattc	2160
	gcctacgaga gtctgagaac acgaccagcc gctccttacc aagtaccaac cataggggtg	2220
	tatggcgtgc caggatcagg caagtctggc atcattaaaa ggcagtcac caaaaaagat	2280
[0080]	ctagtggatga gcgccaagaa agaaaactgt gcagaaatta taaggacgt caagaaaatg	2340
	aaagggtctg acgtcaatgc cagaactgtg gactcagtc tcttgaatgg atgcaaacac	2400
	cccgtagaga cctgtatat tgacgaagct tttgcttgtc atgcaggtac tctcagagcg	2460
	ctcatagcca ttataagacc taaaaaggca gtgctctgcg gggatcccaa acagtgcggt	2520
	tttttaaca tgatgtgcct gaaagtgcac ttaaccacg agatttgcac acaagtcttc	2580
	cacaaaagca tctctgccc ttgcaactaa tctgtgactt cggctcctc aacctgttt	2640
	tacgacaaaa aatgagaac gacgaatccg aaagagacta agattgtgat tgacactacc	2700
	ggcagtacca aacctagca ggacgatctc attctcactt gtttcagagg gtgggtgaag	2760
	cagttgcaaa tagattacaa aggcaacgaa ataatgacgg cagctgcctc tcaagggtg	2820
	accgtaaag gtgtgtatgc cgttcggtac aaggatgatg aaaatcctct gtacgcacc	2880
	acctcagaac atgtgaacgt cctactgacc cgcacggagg accgcatcgt gtggaaaaca	2940
	ctagccggcg acctatgat aaaaactg actgccaagt acctgggaa tttactgccc	3000
	acgatagagg agtggcaagc agagcatgat gccatcatga ggcacatctt ggagagaccg	3060

	gaccctaccg acgtcttcca gaataaggca aacgtgtggtt gggccaagge tttagtgccg	3120
	gtgctgaaga cgcctggcat agacatgacc actgaacaat ggaacactgt ggattatitt	3180
	gaaacggaca aagctcactc agcagagata gtattgaacc aactatgcgt gaggttcttt	3240
	ggactcgatc tggactccgg tctatittct gcacccactg ttccgttate cattaggaat	3300
	aatcactggg ataactcccc gtcgcctaac atgtacgggc tgaataaaga agtggtcctg	3360
	cagctctctc gcaggtacce acaactgcct cgggcagttg ccaactggaag agtctatgac	3420
	atgaacactg gtacactgcg caattatgat ccgcgcataa acctagtacc tgtaaacaga	3480
	agactgcctc atgctttagt cctccacat aatgaacacc cacagagtga cttttcttca	3540
	ttcgtcagca aattgaaggg cagaactgtc ctggtggtcg gggaaaagt gtccgtccca	3600
	ggcaaatgg ttgactgggt gtcagaccgg cctgaggcta ccttcagagc tcggctggat	3660
	ttaggcatcc caggatgat gcccataat gacataatat ttgttaatgt gaggaccca	3720
	tataaatacc atcactatca gcagtgtgaa gaccatgcca ttaagcttag catgttgacc	3780
[0081]	aagaaagctt gtctgcatct gaatccggc ggaacctgtg tcagcatagg ttatggttac	3840
	gctgacaggg ccagcgaag catcattggt gctatagcgc ggcagttcaa gttttccgg	3900
	gtatgcaaac cgaaatcctc acttgaagag acggaagttc tgtttgtatt cattgggtac	3960
	gatcgcaagg cccgtacgca caatccttac aagctttcat caacctgac caacattat	4020
	acaggttcca gactccacga agccggatgt gcacctcat atcatgtggt gcgaggggat	4080
	attgccacgg ccaccgaagg agtgattata aatgctgcta acagcaaagg acaacctggc	4140
	ggaggggtgt gcggagcgt gtataagaaa ttcccgaaa gcttcgattt acagccgatc	4200
	gaagtaggaa aagcgcgact ggtcaaaggt gcagctaac atatcattca tgccgtagga	4260
	ccaaacttca acaaagtttc ggaggtttaa ggtgacaaac agttggcaga ggcttatgag	4320
	tccatcgcta agattgtcaa cgataacaat tacaagtcag tagcgattcc actgttgctc	4380
	accggcatct tttccggaa caaagatcga ctaaccaat cattgaacca ttgctgaca	4440
	gcttagaca ccaactgatc agatgtagcc atatactgca gggacaagaa atgggaaatg	4500
	actctcaagg aagcagtggc taggagagaa gcagtggagg agatatgcat atccgacgac	4560
	tcttcagtga cagaacctga tgcagagctg gtgaggggtc atccgaagag ttctttggct	4620

	ggaaggaagg gctacagcac aagcgatggc aaaactttct catatttga agggaccaag	4680
	tttcaccagg cggccaagga tatagcagaa attaatgcca tgtggcccgt tgcaacggag	4740
	gccaatgagc aggtatgcat gtatatectc ggagaaagca tgagcagtat taggtcghaa	4800
	tgccccgtcg aagagtcgga agcctccaca ccacctagca cgctgccttg cttgtgcatc	4860
	catgccatga ctccagaaaag agtacagcgc ctaaaagcct cacgtccaga acaaattact	4920
	gtgtgctcat cctttccatt gccgaagtat agaatcactg gtgtgcagaa gatccaatgc	4980
	tcccagccta tattgttttc accgaaagtg cctgcgtata ttcaccaag gaagtatctc	5040
	gtggaacac caccggtaga cgagactccg gagccatcgg cagagaacca atccacagag	5100
	gggacacctg aacaaccacc acttataacc gaggatgaga ccaggactag aacgcctgag	5160
	ccgatcatca tcgaagagga agaagaggat agcataagtt tgctgtcaga tgccccgacc	5220
	caccaggtgc tgcaagtcga ggcagacatt cacgggccgc cctctgtatc tagctcatcc	5280
	tggtccattc ctcatgcatc cgactttgat tgggacagtt tatccatact tgacaccctg	5340
[0082]	gagggagcta gcgtgaccag cggggcaacg tcagccgaga ctaactctta cttcgcaaag	5400
	agtatggagt ttctggcgcg accggtgcct gcgcctcga cagtattcag gaaccctcca	5460
	catcccgtc cgcgcacaag aacaccgtca cttgcacca gcagggcctg ctcgagaacc	5520
	agcctagttt ccaccccgcc aggcgtgaat agggatgatca ctagagagga gctcgaggcg	5580
	cttaccctgt cacgcactcc tagcaggtcg gtctcgagaa ccagcctggt ctccaaccg	5640
	ccagcgtaa atagggatgat tacaagagag gagtttgagg cgttcgtagc acaacaacaa	5700
	tgacggtttg atgcgggtgc atacatcttt tctccgaca ccggtcaagg gcatttacia	5760
	caaaaatcag taaggcaaac ggtgctatcc gaagtgggtg tggagaggac cgaattggag	5820
	atttcgtatg ccccgcgcct cgaccaagaa aaagaagaat tactacgcaa gaaattacag	5880
	ttaaatecca cacctgctaa cagaagcaga taccagtcca ggaaggtgga gaacatgaaa	5940
	gccataacag ctagacgiat tctgcaagc ctagggcatt atttgaagc agaaggaaaa	6000
	gtggagtgt accgaaccct gcatcctgtt cctttgtatt catctagtgt gaaccgtgcc	6060
	ttttcaagcc ccaaggtcgc agtggaaagc tgtaacgcca tgttgaaaga gaactttccg	6120

	actgtggcct cttactgtat tattccagag tacgatgcct atttggacat gtttgacgga	6180
	gcttcatgct gcttagacac tgccagtttt tgcctgcaa agctgcgcag ctttccaaag	6240
	aaacactcct atttgaacc cacaatacga tggcagtgcc cttcagcgat ccagaacacg	6300
	ctccagaacg tcctggcagc tgccacaaaa agaaattgca atgtcacgca aatgagagaa	6360
	ttgcccgtat tggattcgcc gccctttaat gtggaatgct tcaagaaata tgcgtgtaat	6420
	aatgaatatt gggaaacggt taaagaaaac cccatcagcc ttactgaaga aaacgtggt	6480
	aattacatta ccaaattaaa aggaccacaaa gctgctgctc tttttgcgaa gacacataat	6540
	ttgaatatgt tgcaggacat accaatggac aggtttgtaa tggacttaaa gagagacgtg	6600
	aaagtgactc caggaacaaa acatactgaa gaacggccca aggtacaggt gatccaggct	6660
	gccgatccgc tagcaacagc gtatctgtgc ggaatccacc gagagctggt taggagatta	6720
	aatgcggctc tgcttccgaa cattcataca ctgtttgata tgcggtgga agactttgac	6780
	gctattatag ccgagcactt ccagcctggg gattgtgttc tggaaactga catcgcgtcg	6840
[0083]	tttgataaaa gtgaggacga cgccatggct ctgaccgcgt taatgattct ggaagactta	6900
	ggtgtggacg cagagctggt gacgctgatt gaggcggctt tcggcgaaat ttcacata	6960
	catttgccca ctaaaactaa atttaaattc ggagccatga tgaatctgg aatgttctc	7020
	acactgtttg tgaacacagt cattaacatt gtaatcga gacagctgtt gagagaacgg	7080
	ctaaccgat caccatgtgc agcattcatt ggagatgaca atatcgtgaa aggagtcaaa	7140
	tcggacaaat taatggcaga caggtgcgcc acctggttga atatggaagt caagattata	7200
	gatgctgtgg tggcgagaa agcgccttat ttctgtggag ggtttatttt gtgtgactcc	7260
	gtgaccggca cagcgtgccg tgtggcagac cccctaaaa ggctgtttaa gcttgcaaa	7320
	cctctggcag cagacgatga acatgatgat gacaggagaa gggcattgca tgaagagtca	7380
	acacgtgga accgagtggt tattctttca gagctgtgca aggcagtaga atcaaggtat	7440
	gaaaccgtag gaacttccat catagttatg gccatgacta ctctagctag cagtgttaaa	7500
	tcattcagct acctgagagg ggcccctata actctctacg gctaacctga atggactacg	7560
	actatcacgc ccaaacattt acagccgagg tgcacaaaac cgcgtggacg tggtaacat	7620
	ccctgctggg aggatcagcc gtaattatta taattggctt ggtgctggct actattgtgg	7680

	ccatgtacgt gctgaccaac cagaacata attgaatata gcagcaattg gcaagctgct	7740
	tacatagaac tcgcgcgat tggcatgccg ccttaaaatt tttatTTTTat tttttctttt	7800
	cttttccgaa tcggattttg tttttaatat ttcaaaaaaa aaaaaaaaaa aaaaaaaaaa	7860
	aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaatacgtag	7920
	tttaaac	7927
	<210> 9	
	<211> 7926	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 9	
	taatacgact cactatagga taggcggcgc atgagagaag cccagaccaa ttacctacc	60
	aaaatggaga aagttcacgt tgacatcgag gaagacagcc cattcctcag agctttgcag	120
[0084]	cggagcttcc cgcagtttga ggtagaagcc aagcaggtca ctgataatga ccattgcta	180
	gccagagcgt tttcgcatct ggcttcaaaa ctgatcgaac cggaggtgga cccatccgac	240
	acgatccttg acattggaag tgcgcccgcc cgcagaatgt attctaagca caagtatcat	300
	tgtatctgtc cgatgagatg tgcggaagat ccggacagat tgtataagta tgcaactaag	360
	ctgaagaaaa actgtaagga aataactgat aaggaattgg acaagaaaat gaaggagctc	420
	gccgccgtca tgagcgacc tgacctggaa actgagacta tgtgcctcca cgacgacgag	480
	tcgtgtcgct acgaaggca agtcgctggt taccaggatg tatacgcggt tgacggaccg	540
	acaagtctct atcaccaagc caataaggga gttagagtcg cctactggat aggctttgac	600
	accacccctt ttatgtttaa gaacttgctt ggagcatatc catcactc taccaactgg	660
	gccgacgaaa ccgtgttaac ggctcgtaac ataggcctat gcagctctga cgttatggag	720
	cggtcacgta gagggatgct cattcttaga aagaagtatt tgaaaccatc caacaatggt	780
	ctattctctg ttggctcgac catctaccac gagaagaggg acttactgag gagctggcac	840
	ctgccgtctg tatttactt acgtggcaag caaaattaca catgtcgggt tgagactata	900

	gttagttgcg acgggtacgt cgtaaaga atagctatca gtccaggcct gtaggggaag	960
	ccttcaggct atgctgctac gatgcaccgc gagggattct tgtgctgcaa agtgacagac	1020
	acattgaacg gggagagggt ctcttttccc gtgtgcacgt atgtgccagc tacattgtgt	1080
	gaccaaata ctggcatact ggcaacagat gtcagtgcgg acgacgcgca aaaactgctg	1140
	gttgggctca accagcgtat agtcgtcaac ggtcgcaccc agagaaacac caatacatg	1200
	aaaaattacc ttttgcccgt agtggcccag gcatttgcta ggtgggcaaa ggaatataag	1260
	gaagatcaag aagatgaaag gccactagga ctacgagata gacagttagt catgggggtg	1320
	tgttgggctt ttagaaggca caagataaca tctatttata agcggcccga tacccaaacc	1380
	atcatcaaag tgaacagcga tttccactca ttcgtgctgc ccaggatagg cagtaacaca	1440
	ttggagatcg ggctgagaac aagaatcagg aaaatgttag aggagcaca ggagccgtca	1500
	cctctcatta ccgccgagga cgtacaagaa gctaagtgcg cagccgatga ggctaaggag	1560
	gtgcgtgaag ccgaggagtt gcgcgcagct ctaccacctt tggcagctga tgttgaggag	1620
	cccactctgg aagccgatgt cgacttgatg ttacaagagg ctggggccgg ctcagtggag	1680
[0085]	acacctctgt gcttgataaa ggttaccagc tacgatggcg aggacaagat cgctcttac	1740
	gctgtgcttt ctccgcagc tgtactcaag agtgaaaaat tatcttgcac ccacctctc	1800
	gctgaacaag tcatagtgat aacacactct ggccgaaaag ggcgttatgc cgtggaacca	1860
	taccatggta aagtagtggg gccagaggga catgcaatac ccgtccagga ctttcaagct	1920
	ctgagtgaag gtgccacat tgtgtacaac gaacgtgagt tcgtaaacag gtacctgcac	1980
	catattgcca cacatggagg agcgcgtgac actgatgaag aatattaca aactgtcaag	2040
	cccagcgagc acgacggcga atacctgtac gacatcgaca gaaacagtg cgtcaagaaa	2100
	gaactagtca ctgggctagg gctcacagc gagctggtgg atcctccctt ccatgaattc	2160
	gcctacgaga gtctgagaac acgaccagcc gctccttacc aagtaccaac cataggggtg	2220
	tatggcgtgc caggatcagg caagtctgac atcattaaaa gcgcagtcac caaaaaagat	2280
	ctagtggatg gcgccaagaa agaaaactgt gcagaaatta taaggacgt caagaaaatg	2340
	aaagggtg acgtcaatgc cagaactgtg gactcagtc tcttgaatgg atgcaaacac	2400
	cccgtagaga ccctgtatat tgacgaagct tttgcttgc atgcaggtac tctcagagcg	2460

	ctcatagcca ttataagacc taaaaaggca gtgctctgcg gggatcccaa acagtgcggt	2520
	ttttttaaca tgatgtgcct gaaagtgcac ttttaaccacg agatttgcaac acaagtcttc	2580
	cacaaaagca tctctcgccg ttgcactaaa tctgtgactt cggtcgtctc aaccttgttt	2640
	tacgacaaaa aaatgagaac gacgaatccg aaagagacta agattgtgat tgacactacc	2700
	ggcagtacca aacctaaagca ggacgatctc attctcactt gtttcagagg gtgggtgaag	2760
	cagttgcaaa tagattacaa aggcaacgaa ataatgacgg cagctgcctc tcaagggtg	2820
	acctgtaaag gtgtgtatgc cgttcggtac aaggtgaatg aaaatcctct gtacgcacce	2880
	acctcagaac atgtgaacgt cctactgacc cgcacggagg accgcatcgt gtggaaaaca	2940
	ctagccggcg acctatggat aaaaacactg actgccaaagt acctgggaa tttcactgcc	3000
	acgatagagg agtggcaagc agagcatgat gccatcatga ggcacatctt ggagagaccg	3060
	gacctaccg acgtcttcca gaataaggca aacgtgtgtt gggccaaggc tttagtgcg	3120
	gtgtgaaga ccgctggcat agacatgacc actgaacaat ggaacactgt ggattatfff	3180
[0086]	gaaacggaca aagctcactc agcagagata gtattgaacc aactatgcgt gaggttcttt	3240
	ggactcgatc tggactccgg tctatfffct gcacccactg ttccgttacc cattaggaat	3300
	aatcactggg ataactcccc gtcgcctaac atgtacgggc tgaataaaga agtggtcctg	3360
	cagctctctc gcaggtacc acaactgcct cgggcagttg cacttgaag agtctatgac	3420
	atgaacactg gtacactgcg caattatgat ccgcgcataa acctagtacc tgtaaacaga	3480
	agactgcctc atgctttagt cctccaccat aatgaacacc cacagagtga cttttcttca	3540
	ttcgtcagca aattgaaggg cagaactgtc ctgggtggtc gggaaaagtt gtccgtccca	3600
	ggcaaaatgg ttgactggtt gtcagaccgg cctgaggcta cttcagagc tcggctggat	3660
	ttaggcatcc caggtgatgt gcccaaatat gacataatat ttgttaatgt gaggaccca	3720
	tataaatacc atcaactatca gcaggtgaa gaccatgcca ttaagcttag catgttgacc	3780
	aagaaagctt gtctgcatct gaatcccggc ggaacctgtg tcagcatagg ttatggttac	3840
	gctgacaggg ccagcgaag catcattggt gctatagcgc ggcagttcaa gttttcccgg	3900
	gtatgcaaac cgaaatcctc acttgaagag acggaagtgc tgtttgtatt cattgggtac	3960

	gatcgcaagg cccgtacgca caatccttac aagctttcat caaccttgac caacatttat	4020
	acaggttcca gactccacga agccggatgt gcacctcat atcatgtggt gcgaggggat	4080
	attgccacgg ccaccgaagg agtgattata aatgctgcta acagcaaagg acaacctggc	4140
	ggaggggtgt gcggagcgct gtataagaaa ttcccgaaa gcttcgattt acagccgatc	4200
	gaagtaggaa aagcgcgact ggtcaaaggt gcagctaac atatcattca tgccgtagga	4260
	ccaaacttca acaaagtttc ggaggttga ggtgacaaac agttggcaga ggcttatgag	4320
	tccatcgcta agattgtcaa cgataacaat tacaagtcag tagcgattcc actgttgtcc	4380
	accggcatct tttccgggaa caaagatcga ctaaccaat cattgaacca ttgtctgaca	4440
	gcttagaca ccaactgatgc agatgtagcc atatactgca gggacaagaa atgggaaatg	4500
	actctcaagg aagcagtggc taggagagaa gcagtggagg agatatgcat atccgacgac	4560
	tcttcagtga cagaacctga tgcagagctg gtgaggggtgc atccgaagag ttctttgget	4620
	ggaaggaagg gctacagcac aagcgatggc aaaactttct catatttga agggaccaag	4680
	ttcaccagg cgccaagga tatagcagaa attaatgcca tgtggcccgt tgcaacggag	4740
[0087]	gccaatgagc aggtatgcat gtatatctc ggagaaagca tgagcagtat taggtcgaaa	4800
	tgccccgtcg aagagtcgga agcctccaca ccacctgca cgctgccttg cttgtgcatc	4860
	catgccatga ctccagaaag agtacagcgc ctaaaagcct cacgtccaga acaaattact	4920
	gtgtgctcat ctttccatt gccgaagtat agaatactg gtgtgcagaa gatccaatgc	4980
	tcccagccta tattgttctc accgaaagtg cctgcgtata ttcaccaag gaagtatctc	5040
	gtggaaacac caccggtaga cgagactccg gagccatcgg cagagaacca atccacagag	5100
	gggacacctg aacaaccacc acttataacc gaggatgaga ccaggactag aacgcctgag	5160
	ccgatcatca tgaagagga agaagaggat agcataagtt tgctgtcaga tgccccgacc	5220
	caccaggtgc tgcaagtcca ggcagacatt cacgggcccgc cctctgtatc tagctcatcc	5280
	tggccattc ctcatgcatc cgactttgat gtggacagtt tatccatact tgacacctg	5340
	gagggagcta gcgtgaccag cggggcaacg tcagccgaga ctaactctta cttcgcaaag	5400
	agtatggagt ttctggcgcg accggtgcct gcgcctcgaa cagtattcag gaacctcca	5460
	cateccgctc cgcgcaaacg aacaccgtca cttgcacca gcagggcctg ctcgagaacc	5520

	agcctagttt ccaccccgcc aggcgtgaat agggatgatca ctagagagga gctcagggcg	5580
	cttaccocgt cacgcactcc tagcaggteg gtctcgagaa ccagcctggt ctccaacccg	5640
	ccaggcgtaa atagggtgat tacaagagag gagtttgagg cgttcgtagc acaacaacaa	5700
	tgacggtttg atgcccgtgc atacatcttt tcctccgaca cgggtcaagg gcatttacia	5760
	caaaaatcag taaggcaaac ggtgctatcc gaagtgggtg tggagaggac cgaattggag	5820
	atttcgtatg ccccgcgcct cgaccaagaa aaagaagaat tactacgcaa gaaattacag	5880
	ttaaatccca cacctgctaa cagaagcaga taccagtcca ggaaggtgga gaacatgaaa	5940
	gccataacag ctagacgtat tctgcaaggc ctagggcatt atttgaaggc agaaggaaaa	6000
	gtggagtgt accgaacct gcacccgtgt cctttgtatt catctagtgt gaaccgtgcc	6060
	tttcaagcc ccaaggcgc agtggaagcc tgtaacgcca tgttgaaaga gaactttccg	6120
	actgtggett ctactgtat tattccagag tacgatgcct atttggacat ggttgacgga	6180
	gcttcatgct gcttagacac tgccagtttt tgccctgcaa agctgcgcag ctttccaaag	6240
[0088]	aaacactcct atttgaacc cacaatacga tcggcagtgc cttcagcgat ccagaacacg	6300
	ctccagaacg tcctggcagc tgccacaaaa agaaattgca atgtcacgca aatgagagaa	6360
	ttgcccgtat tggattcggc ggccittaat gtggaatgct tcaagaaata tgcgtgtaat	6420
	aatgaatatt gggaaacgtt taaagaaaac cccatcaggc ttactgaaga aaacgtggta	6480
	aattacatta ccaaattaa aggaccacaaa gctgctgctc tttttcgaa gacacataat	6540
	ttgaatatgt tgcaggacat accaatggac aggtttgtaa tggacttaaa gagagacgtg	6600
	aaagtgactc caggaacaaa acatactgaa gaacggccca aggtacaggt gatccaggct	6660
	gccgatccgc tagcaacagc gtatctgtgc ggaatccacc gagagctggt taggagatta	6720
	aatgcggctc tgcttccgaa cattcataca ctgtttgata tgtcggctga agactttgac	6780
	gctattatag ccgagcactt ccagcctggg gattgtgttc tggaaactga catcgcgtcg	6840
	tttgataaaa gtgaggacga cgccatggct ctgaccgcgt taatgattct ggaagactta	6900
	ggtgtggacg cagagctgtt gacgctgatt gaggcggctt tcggcgaat ttcataata	6960
	catttgccca ctaaaactaa atttaaatc ggagccatga tgaaatctgg aatgttctc	7020

	acactgtttg tgaacacagt cattaacatt gtaatcгаа gcagagtgtt gagagaacgg	7080
	ctaaccggat caccatgtgc agcattcatt ggagatgaca atatcgtgaa aggagtcaaa	7140
	teggacaaat taatggcaga caggtgcgcc acctggttga atatggaagt caagattata	7200
	gatgctgtgg tggcgagaa agcgccttat ttctgtggag ggtttatttt gtgtgactcc	7260
	gtgaccggca cagcgtgccg tgtggcagac cccctaaaaa ggctgtttaa gcttggcaaa	7320
	cctctggcag cagacgatga acatgatgat gacaggagaa gggcattgca tgaagagtca	7380
	acacgctgga accgagtggg tattctttca gagctgtgca aggcagtaga atcaaggtat	7440
	gaaaccgtag gaacttccat catagttagt gccatgacta ctctagctag cagtgttaaa	7500
	tcattcagct acctgagagg ggcccctata actctctacg gctaacctga atggactacg	7560
	actatcacgc ccaaacattt acagccgcgg tgtcaaaaac cgcgtggacg tggttaacat	7620
	ccctgctggg aggatcagcc gtaattatta taattggctt ggtgctggct actattgtgg	7680
	ccatgtacgt gctgaccaac cagaaacata attgaataca gcagcaattg gcaagctgct	7740
	tacatagaac tcgcggcgat tggcatgccg ccttaaaatt tttattttat ttttcttttc	7800
[0089]	ttttccgaat cggattttgt ttttaattt tcaaaaaaaaa aaaaaaaaa aaaaaaaaa	7860
	aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aatacgtagt	7920
	ttaaac	7926
	<210> 10	
	<211> 36519	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 10	
	ccatcttcaa taatatacct caaacTTTT gtgcgcgтта atatgcaaat gaggcgtttg	60
	aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgttagg ggcggggcga	120
	gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag	180
	tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttcccgc gctctctgac	240
	aggaaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact	300

	gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg cagggaggag tatttgccga	360
	gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa	420
	tttccgcgta cgggtgtcaaa gtccgggtgtt ttacgttagg tgtcagctga tcgccagggt	480
	atthaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcgag aagagttttc	540
	tcctccgcgc cgcgagtcag atctacactt tgaagatga ggcacctgag agacctgccc	600
	gatgagaaaa tcatcatcgc ttccgggaac gagattctgg aactggtggt aaatgccatg	660
	atgggcgacg accctccgga gccccacc ccatttgaga caccttcgct gcacgatttg	720
	tatgatctgg aggtggatgt gcccaggac gatcccaatg aggaggcggg aaatgatttt	780
	tttagcgatg ccgcgctgct agctgccgag gaggttcga gctctagctc agacagcgac	840
	tcttactgc ataccctag acccggcaga ggtgagaaaa agatccccga gcttaaaggg	900
	gaagagatgg acttgcgctg ctatgaggaa tgcttgcccc cgagcgatga tgaggacgag	960
	caggcgatcc agaacgcagc gagccaggga gtgcaagccg ccagcgagag ctttgcgctg	1020
[0090]	gactgcccgc ctctgcccgg acacggctgt aagtcttggt aatttcacg catgaatact	1080
	ggagataaag ctgtgttggtg tgcactttgc tatatgagag cttacaacca ttgtgtttac	1140
	agtaagtgtg attaagttga actttagagg gaggcagaga gcagggtgac tgggcgatga	1200
	ctggtttatt tatgtatata tgttctttat ataggtcccg tctctgacgc agatgatgag	1260
	acccccacta caaagtccac ttcgtcacc ccagaaattg gcacatctcc acctgagaat	1320
	attgttagac cagttcctgt tagagccact gggaggagag cagctgtgga atgtttgat	1380
	gacttgctac aggggtgggt tgaaccttg gacttgtgta cccgaaacg ccccaggcac	1440
	taagtgccac acatgtgtgt ttacttgagg tgatgtcagt atttataggg tgtggagtgc	1500
	aataaaaaat gtgttgactt taagtgcgtg gtttatgact caggggtggg gactgtgagt	1560
	atataagcag gtgcagacct gtgtggttag ctcagagcgg catggagatt tggacggtct	1620
	tggaagactt tcacaagact agacagctgc tagagaacgc ctcgaacgga gtctcttacc	1680
	tgtggagatt ctgcttcggt ggcgacctag ctaggctagt ctacagggcc aaacaggatt	1740
	atagtgaaca atttgagggt attttgagag agtgttctgg tctttttgac gtcttaact	1800

	tgggccatca gtctcacttt aaccagagga tttcgagagc ctttgatttt actactcctg	1860
	gcagaaccac tgcagcagta gccttttttg cttttattct tgacaaatgg agtcaagaaa	1920
	cccatttcag cagggattac cagctggatt tcttagcagt agctttgtgg agaacatgga	1980
	agtgccagcg cctgaatgca atctccggct acttgccggc acagccgcta gacactctga	2040
	ggatcctgaa tctccaggag agtcccaggc cacgccaacg tcgccagcag cagcagcagg	2100
	aggaggatca agaagagaac ccgagagccg gcctggacce tccggcggag gaggaggagt	2160
	agctgacctg tttcctgaac tgcgccgggt gctgactagg tcttcgagtg gtcgggagag	2220
	ggggattaag cgggagaggc atgatgagac taatcacaga actgaactga ctgtgggtct	2280
	gatgagtcgc aagcggccag aaacagtgtg gtggcatgag gtgcagtcca ctggcacaga	2340
	tgagggtgctg gtgatgatg agaggttttc tctagaacaa gtcaagactt gttggttaga	2400
	gcctgaggat gattgggagg tagccatcag gaattatgcc aagctggctc tgaggccaga	2460
	caagaagtac aagattacta agctgataaa tatcagaaat gcctgtaca tctcagggaa	2520
	tggggctgaa gtggagatct gtctccagga aagggtggct ttcagatgct gcatgatgaa	2580
[0091]	tatgtaccgg ggagtgggtg gcatggatgg ggttaccttt atgaacatga ggttcagggg	2640
	agatgggtat aatggcacgg tctttatggc caataccaag ctgacagtcc atggctgctc	2700
	cttctttggg ttttaataaca cctgcatcga ggcctggggc caggctcggc tgaggggctg	2760
	cagttttca gccaaactgga tgggggtcgt gggcaggacc aagagtatgc tgtccgtgaa	2820
	gaaatgcttg tttgagaggt gccacctggg ggtgatgagc gagggcgaag ccagaatccg	2880
	ccactgcgcc tctaccgaga cgggctgctt tgtgctgtgc aaggcfaatg ctaagatcaa	2940
	gcataaatat atctgtggag cctcggacga gcgcggctac cagatgctga cctgcgcccg	3000
	cgggaacagc catatgctgg ccaccgtaca tgtggettcc catgctcga agccctggcc	3060
	cgagtccgag cacaatgtca tgaccagggt caaatatgcat ctggggctccc gccgaggcat	3120
	gttcatgccc taccagtcca acctgaatta tgtgaagggt ctgctggagc ccgatgccat	3180
	gtccagagtg agcctgacgg ggggtgttga catgaatgtg gaggtgtgga agattctgag	3240
	atatgatgaa tccaagacca ggtgccgagc ctgcgagtgc ggagggaagc atgccaggtt	3300
	ccagcccgtg tgtgtggatg tgaccggagga cctgcgacce gatcatttgg tgttgcctg	3360

	caccgggacg gagttcgggt ccagcgggga agaatctgac tagagtgagt agtgttctgg	3420
	ggcgggggag gacctgcatg agggccagaa taactgaaat ctgtgctttt ctgtgtgttg	3480
	cagcagcatg agcgggaagcg gctcctttga gggaggggta ttcagccctt atctgacggg	3540
	gcgtctcccc tcctgggcgg gagtgcgta gaatgtgatg ggatccacgg tggacggccg	3600
	gccccgtcag cccgcgaact cttcaacctt gacctatgca accctgagct cttcgtcgtt	3660
	ggacgcagct gccgccgag ctgctgcatc tgccgccagc gccgtgcgcg gaatggccat	3720
	gggcgccggc tactacggca ctctgggtgc caactcgagt tccaccaata atccccag	3780
	cctgaacgag gagaagctgt tgctgctgat ggcccagctc gaggccttga cccagcgcct	3840
	ggcgagctg acccagcagg tggtcagct gcaggagcag acgcgggccc cggttgccac	3900
	ggtgaaatcc aaataaaaaa tgaatcaata aataaacgga gacggttgtt gattttaaca	3960
	cagagtctga atctttattt gatttttcgc gcgcggtagg ccctggacca ccggtctcga	4020
	tcattgagca cccggtggat cttttccagg acccggtaga ggtgggcttg gatgttgagg	4080
[0092]	tacatgggca tgagcccgtc ccgggggtgg aggtagctcc attgcagggc ctctgctcg	4140
	gggtgggtgt tgtaaatcac ccagtcatag caggggcgca gggcatggtg ttgcacaata	4200
	tctttgagga ggagactgat ggccacgggc agccctttgg tgtagtggtt taaaaatctg	4260
	ttgagctggg agggatgcat gcggggggag atgaggtgca tcttgccctg gatcttgaga	4320
	ttggcgatgt taccgccag atcccctg ggttcatgt tgtgcaggac caccagcacg	4380
	gtgtatccgg tgcacttggg gaatttatca tgcaacttgg aagggaaggc gtgaaagaat	4440
	ttggcgacgc ctttgtgcc gcccaggttt tccatgact catccatgat gatggcgatg	4500
	ggccccgtgg cgggcgccct ggcaaagacg tttcgggggt cggacacatc atagttgttg	4560
	tcctgggtga ggtcatcata ggccatttta atgaatttgg ggcggagggt gccggactgg	4620
	gggacaaagg taccctgat cccggggcg tagttccct cacagatctg catctcccag	4680
	gctttgagct cggagggggg gatcatgtcc acctgcgggg cgataaagaa cacggtttcc	4740
	ggggcggggg agatgagctg ggccgaaagc aagttccgga gcagctggga cttgccgag	4800
	ccggtggggc cgtagatgac cccgatgacc ggctgcaggt ggtagttgag ggagagacag	4860

	ctgcgcctct cccggaggag gggggccacc tcgttcatca tctcgcgcac gtgcatgttc	4920
	tcgcgcacca gttccgccag gaggcgctct cccccaggg ataggagctc ctggagcgag	4980
	gcgaagtttt tcagcggctt gagtccgtcg gccatgggca ttttgagag gttttgttc	5040
	aagagttcca ggcgggtcca gagctcggtg atgtgtctta cggcatctcg atccagcaga	5100
	cctcctcgtt tcgcgggttg ggacggctgc gggagtaggg caccagacga tggcgtcca	5160
	gcgcagccag ggtccggtcc ttccagggtc gcagcgtccg cgtcagggtg gtctccgtca	5220
	cggtaaggg gtgcgcgcc ggctgggccc ttgcgagggt gcgcttcagg ctcatccggc	5280
	tggtcgaaaa ccgctcccga tcggcgcctt gcgcgtcggc caggtagcaa ttgacctga	5340
	gttcgtagtt gagcgcctcg gccgcgtggc ctttggcgcg gagcttacct ttggaagtct	5400
	gcccgcaggc gggacagagg agggacttga gggcgtagag cttggggcgc aggaagacgg	5460
	actcgggggc gtaggcgtcc gcgccgcagt gggcgcagac ggtctcgcac tccacgagcc	5520
	aggtgaggtc gggctggctc gggtaaaaa ccagtttccc gcgcttcttt ttgatgcgtt	5580
[0093]	tcttaccttt ggtctccatg agctcgtgtc cccgctgggt gacaaagagg ctgtccgtgt	5640
	ccccgtagac cgactttatg ggccggtcct cgagcgggtg gccgcggtcc tcctcgtaga	5700
	ggaaccccgc cactccgag acgaaagccc ggggtccaggc cagcacgaag gaggccacgt	5760
	gggacgggta gcggtcgttg tccaccagcg ggtccacctt ttccagggta tgcaaacaca	5820
	tgtccccctc gtccacatcc aggaaggtga ttggcttgta agtgtaggcc acgtgaccgg	5880
	gggtcccggc cgggggggta taaaagggtg cgggtccctg ctcgctctca ctgtcttccg	5940
	gatcgtgtc caggagcgc agctgttggg gtaggtattc cctctcgaag gcgggcatga	6000
	cctcggcact caggttgtca gtttctagaa acgaggagga ttgatattg acggtgccgg	6060
	cggagatgcc tttcaagagc ccctcgtcca tctggtcaga aaagacgac tttttgttgt	6120
	cgagcttggg gcgcaaggag ccgtagaggg cgttggagag gagcttggcg atggagcgca	6180
	tggtctggtt ttttctctg tcggcgcgct ccttggcggc gatgttgagc tgcacgtact	6240
	cgcgccccac gcacttccat tcggggaaga cgggtgtcag ctcgtcgggc acgattctga	6300
	cctgccagcc ccgattatgc agggatgatga ggtccacact ggtggccacc tcgccgcga	6360
	ggggctcatt agtccagcag aggcgtccgc ccttgcgcga gcagaagggg ggcagggggt	6420

	ccagcatgac ctcgtcgggg gggtcggcat cgatggtgaa gatgccgggc aggaggtcgg	6480
	ggtcaaagta gctgatggaa gtggccagat cgtccagggc agcttgccat tcgcgcacgg	6540
	ccagcgcgcg ctcgtaggga ctgagggcg tccccaggg catgggatgg gtaagcgcgg	6600
	aggcgtacat gccgcagatg tcgtagacgt agaggggctc ctcgaggatg ccgatgtagg	6660
	tgggtagca gcgcccccg cggatgctgg cgcgcacgta gtcatacagc tcgtgcgagg	6720
	gggcgaggag ccccgggcc aggttgggtgc gactgggctt ttcggcgcgg tagacgatct	6780
	ggcgaaaaat ggcatgcgag ttggaggaga tggtaggctt ttggaagatg ttgaagtggg	6840
	cgtggggcag tccgaccgag tcgcgatga agtgggcgta ggagtcttgc agcttggcga	6900
	cgagctcggc ggtgactagg acgtccagag cgcagtagtc gagggtctcc tggatgatgt	6960
	catacttgag ctgtcccttt tgtttccaca gctcgcggtt gagaaggaac tcttcgcggt	7020
	ccttcagta ctcttcgagg gggaaccctt cctgatctgc acgtaagag cctagcatgt	7080
	agaactgggt gacggccttg taggcgcagc agcccttctc cacggggagg gcgtaggcct	7140
[0094]	gggcggcctt gcgcaggag gtgtcgtga gggcgaaagt gtcctgacc atgaccttga	7200
	ggaactggtg cttgaagtcg atatcgtcg agccccctg ctcccagagc tggaaagtcg	7260
	tgcgcttctt gtaggcgggg ttggcmaaag cgaaagtaac atcgttgaag aggatcttgc	7320
	ccgcgcgggg cataaagttg cgagtgatgc ggaaaggttg gggcacctcg gcccgttgt	7380
	tgatgacctg ggcggcgagc acgatctcgt cgaagccgtt gatgttgggg cccacgatgt	7440
	agagttccac gaatcgcgga cggcccttga cgtggggcag tttcttgagc tcctcgtagg	7500
	tgagctcgtc ggggtcgtg agcccgtgct gctcagcgc ccagtcggcg agatgggggt	7560
	tggcgcggag gaaggaagtc cagagatcca cggccagggc ggtttgcaga cggtcccgtt	7620
	actgacgaa ctgctgcccg acggccattt tttcgggggt gacgcagtag aaggtgcggg	7680
	ggtcccctg ccagcgatcc catttgagct ggagggcgag atcgaggcg agctcgacga	7740
	gccgctcgtc cccggagagt ttcattacca gcatgaaggg gacgagctgc ttgccgaagg	7800
	accccatcca ggtgtaggtt tccacatcgt aggtgaggaa gagcctttcg gtgcgaggat	7860
	gcgagccgat ggggaagaac tggatctcct gccaccaatt ggaggaatgg ctgttgatgt	7920

	gatggaagta gaaatgccga cggcgcgccg aacactcgtg cttgtgttta tacaagcggc	7980
	cacagtgtc gcaacgtgc acgggatgca cgtgctgcac gagctgtacc tgagttcctt	8040
	tgacaggaa tttcagtggg aagtggagtc gtggcgccg catctcgtgc tgtactacgt	8100
	cgtggtggc ggccctggccc tcttctgcct cgatggtggt catgctgacg agcccgcgcg	8160
	ggaggcaggt ccagacctc gcgcgagcgg gtcggagagc gaggacgagg gcgcgcaggc	8220
	cggagctgtc cagggtcctg agacgctgcg gagtcaggtc agtgggcagc ggccgcgcgc	8280
	ggttgacttg caggagtttt tccagggcgc gcgggaggtc cagatggtac ttgatctcca	8340
	ccgcgccatt ggtggcgacg tcgatggctt gcagggtccc gtgcccctgg ggtgtgacca	8400
	ccgtcccccg tttcttcttg ggccgctggg gcgacggggg cggtcctct tccatggtta	8460
	gaagcggcgg cgaggacgcg cgcggggcgg caggggcggc tcggggcccg gaggcagggg	8520
	cggcaggggc acgtcggcgc cgcgcgcggg taggttctgg tactgcgccc ggagaagact	8580
	ggcgtgagcg acgacgcgac ggttgacgtc ctggatctga cgctctggg tgaaggccac	8640
[0095]	gggaccctg agtttgaacc tgaagagag ttcgacagaa tcaatctcg tatcgttgac	8700
	ggcggcctgc cgcaggatct cttgcacgtc gcccagttg tcttggtagg cgatctcgtt	8760
	catgaactgc tcgatctcct cctcttgaag gtctccgcgg ccggcgcgct ccacggtggc	8820
	cgcgaggtcg ttggagatgc ggcccatgag ctgcgagaag gcgttcatgc ccgcctcgtt	8880
	ccagacgcgg ctgtagacca cgacgccctc gggatcgcgg gcgcgcatga ccacctggc	8940
	gaggttgagc tccacgtggc gcgtgaagac cgcgtagttg cagaggcgtt ggtagaggta	9000
	gttgagcgtg gtggcgtatg gctcggtgac gaagaaatac atgatccagc ggcggagcgg	9060
	catctcgtg acgtcgccea gcgcctcaa acgttccatg gcctcgtaaa agtccacggc	9120
	gaagttgaaa aactgggagt tgcgcgccga gacggtcaac tcctctcca gaagacggat	9180
	gagctcggcg atggtggcgc gcacctcgcg ctcgaaggcc cccgggagtt cctccactc	9240
	ctcttcttc tcctccacta acatctcttc tacttctcc tcaggcggca gtggtggcgg	9300
	gggagggggc ctgcctgcc ggccggcgcac gggcagacgg tcgatgaagc gctcgtggt	9360
	ctcggcgcgc cggcgtgca tggctcggg gacggcgcgc ccgtcctcgc ggggccgcag	9420
	cgtgaagacg ccgccgcga tctccaggtg gccggggggg tcccgttg gcaggagag	9480

	ggcgctgacg atgcatctta tcaattgccc cgtagggact ccgcgcaagg acctgagcgt	9540
	ctcgagatcc acgggatctg aaaaccgctg aacgaaggct tcgagccagt cgcagtcgca	9600
	aggtaggctg agcacggttt cttctggcgg gtcattgttg ttgggagcgg ggcgggcat	9660
	gctgctggtg atgaagttga aataggcggg tctgagacgg cggatggtgg cgaggagcac	9720
	caggctcttg ggcccggctt gctggatgcg cagacggtcg gccatgcccc aggcgtggtc	9780
	ctgacacctg gccaggctct ttagtagtc ctgcatgagc cgctccacgg gcacctctc	9840
	ctcggccgcg cggccgtgca tgcgcgtgag cccgaagccg cgctggggct ggacgagcgc	9900
	caggctggcg acgacgcgct cggcgaggat ggcttctggt atctgggtga ggggtgctc	9960
	gaagtcacga aagtcgacga agcgggtgga ggctccgggt ttgatggtgt aggagcagtt	10020
	ggccatgacg gaccagtga cggctctggt gcccgacgc acgagctcgt ggtacttgag	10080
	gcgcgagtag gcgcgctgt cgaagatgta gtcgttgacg gtgcgcacca ggtactggtg	10140
	gccgatgagg aagtgcggcg gcggctggcg gtagagcggc catcgcctcg tggcggggc	10200
[0096]	gccggcgcg aggtcctcga gcatggtgcg gtggtagccg tagatgtacc tggacatcca	10260
	ggtgatgccg gcggcgggtg tggaggcgcg cgggaactcg cggacgcggt tccagatggt	10320
	gcgcagcggc aggaagtagt tcatggtggg cacggtctgg cccgtgaggc gcgcgcagtc	10380
	gtggatgctc tatacgggca aaaacgaaag cggtcagcgg ctcgactccg tggcctggag	10440
	gctaagcgaa cgggttgggc tgcgcgtgta ccccggttcg aatctcgaat caggctggag	10500
	ccgcagctaa cgtggtattg gcaactccgt ctcgacccaa gcctgcacca accctccagg	10560
	atacggaggc gggtcgtttt gcaacttttt ttggaggcc ggatgagact agtaagcgcg	10620
	gaaagcggcc gaccgcgatg gctcgtgcc gtagtctgga gaagaatgc cagggttgcg	10680
	ttgcggtgtg ccccggttcg aggccggccg gattccgcgg ctaacgaggc cgtggctgcc	10740
	ccgtcgttcc caagaccca tagccagccg acttctccag ttacggagcg agcccctctt	10800
	ttgtttgtt tgtttttgcc agatgcatcc cgtactgcgg cagatgcgcc cccaccacc	10860
	tccaccgcaa caacagcccc ctccacagcc ggcgttctg ccccccccc agcagcaact	10920
	tccagccacg accgccgcgg ccgccgtgag cggggctgga cagagttatg atcaccagct	10980

	ggccttgaa gaggcgagg ggctggcgcg cctgggggcg tcgtcgccgg agcggcaccc	11040
	gcgctgcag atgaaaagg acgctcgcga ggcctacgtg cccaagcaga acctgttcag	11100
	agacaggagc ggcgaggagc ccgaggagat gcgcgcggcc cggttccacg cggggcggga	11160
	gctcggcgc ggcctggacc gaaagagggt gctgaggac gaggatttcg aggcggacga	11220
	gctgacgggg atcagccccg cgcgcgcga cgtggccgcg gccaacctgg tcacggcgta	11280
	cgagcagacc gtgaaggagg agagcaactt ccaaaaatcc ttcaacaacc acgtgcgac	11340
	cctgatcgcg cgcgaggagg tgacctggg cctgatgcac ctgtgggacc tgctggaggc	11400
	catcgtgcag aaccccacca gcaagccgct gacggcgcag ctgttcctgg tggcgcagca	11460
	tagtcgggac aacgaagcgt tcaggaggc gctgctgaat atcaccgagc ccgagggccg	11520
	ctggctcctg gacctggtga acattctgca gagcatcgtg gtgcaggagc gcgggctgcc	11580
	gctgtccgag aagctggcgg ccatcaactt ctcggtgctg agtttgggca agtactacgc	11640
	taggaagatc tacaagacc cgtacgtgcc catagacaag gaggtgaaga tcgacgggtt	11700
	ttacatgcgc atgacctga aagtgtgac cctgagcgc gatctggggg tgtaccgcaa	11760
[0097]	cgacaggatg caccgtgcgg tgagcgcag caggcggcgc gagctgagcg accaggagct	11820
	gatgcatagt ctgcagcgg ccctgaccgg ggccgggacc gagggggaga gctactttga	11880
	catggcgcgc gacctgcact ggcagcccag ccgccgggcc ttggaggcgg cggcaggacc	11940
	ctacgtagaa gaggtggac atgaggtgga cgaggaggc gagtacctgg aagactgatg	12000
	gcgcgaccgt attttgcta gatgcaacaa caacagccac ctctgatcc cgcgatgcgg	12060
	gcggcgtgc agagccagcc gtccggcatt aactcctcg acgattggac ccagccatg	12120
	caacgcatca tggcctgac gaccgcaac ccggaagcct ttagacagca gcccaggcc	12180
	aaccgctct cggccatct ggaggccgtg gtgccctgc gctccaacc cacgcagag	12240
	aagtcctgg ccatcgtgaa cgcctggtg gagaacaagg ccatccgcgg cgacgaggcc	12300
	ggcctggtgt acaacgcct gctggagcgc gtggcccgt acaacagcac caacgtgcag	12360
	accaacctgg accgcatggt gaccgacgt cgcgaggccg tggcccagcg cgagcggttc	12420
	caccgcgagt ccaacctggg atccatggtg gcgctgaac ccttctcag caccagccc	12480
	gccaacgtgc cccggggcca ggaggactac accaactca tcagegcct gcgcctgatg	12540

	gtgaccgagg tgccccagag cgagggtgtac cagtccgggc cggactactt cttccagacc	12600
	agtcgccagg gcttgcagac cgtgaacctg agccaggctt tcaagaactt gcagggcctg	12660
	tggggcgtgc aggccccggt cggggaccgc gcgacggtgt cgagcctgct gacgccgaac	12720
	tcgvcctgc tgctgctgct ggtggcccc ttcacggaca gcggcagcat caaccgcaac	12780
	tcgtacctgg gctacctgat taacctgtac cgcgaggcca tcggccaggc gcacgtggac	12840
	gagcagacct accaggagat caccacgtg agccgcgcc tgggccagga cgaccgggc	12900
	aacctggaag ccacctgaa ctttttctg accaaccggt cgcagaagat cccgccccag	12960
	tacgcctca gcaccgagga ggagcgcac ctgcgttac tgcagcagag cgtgggcctg	13020
	ttctgatgc aggagggggc cccccagc gccgcgctcg acatgaccgc gcgcaacatg	13080
	gagcccagca tgtaccag caaccgccg ttcataata aactgatgga ctacttgcac	13140
	cgggcggccg ccatgaactc tgactatttc accaacgcca tectgaatcc ccaactggctc	13200
	ccgccccgg ggttctacac gggcgagtac gacatgccc accccaatga cgggttctg	13260
[0098]	tgggacgatg tggacagcag cgtgttctcc ccccgaccgg gtgctaacga gcgccccttg	13320
	tggaagaagg aaggcagcga ccgaccccc tctcggcgc tgtccggccg cgagggtgct	13380
	gccgcggcgg tgcccagagc cggcagctct tcccagact tgccctctc gctgaacagt	13440
	atccgcagca gcgagctggg caggatcacg cccccgct tgctgggcga agaggagtac	13500
	ttgaatgact cgctgtttag acccgagcgg gagaagaact tccccataa cgggatagaa	13560
	agcctggtgg acaagatgag ccgctggaag acgtatgcgc aggagcacag ggacgatccc	13620
	cgggcgtgc agggggccac gagccggggc agcgcgcc gtaaaccggt gtggcacgac	13680
	aggcagcggg gacagatgtg ggacgatgag gactccgcc acgacagcag cgtgttgac	13740
	ttgggtgga gtggtaacc gttcgtcac ctgcgcccc gtatcggcg catgatgtaa	13800
	gagaaaccga aaataaatga tactaccaa ggccatggcg accagcgtgc gttcgttct	13860
	tctctgttgt tgttgtatct agtatgatga ggcgtgcgta cccggagggt cctcctcct	13920
	cgtacgagag cgtgatgcag caggcagtg cggcggcggc gatgcagccc ccgctggagg	13980
	ctccttacgt gccccgcgg tacctggcgc ctacggaggg gcggaacagc attcgttact	14040

	cggagctggc acccttgtac gataccaccc ggttgtacct ggtggacaac aagtcggcgg	14100
	acatcgctc gctgaactac cagaacgacc acagcaactt cctgaccacc gtggtgcaga	14160
	acaatgactt cacccccacg gaggccagca cccagaccat caactttgac gagcgcctgc	14220
	ggtggggcgg ccagctgaaa accatcatgc acaccaacat gcccaacgtg aacgagttca	14280
	tgtacagcaa caagttcaag gcgcgggtga tggctctccg caagaccccc aatgggggtga	14340
	cagtgcacaga ggattatgat ggtagtcagg atgagctgaa gtatgaatgg gtggaatttg	14400
	agctgcccga aggcaacttc tcggtgacca tgaccatcga cctgatgaac aacgccatca	14460
	tcgacaatta cttggcggtg gggcggcaga acgggggtgct ggagagcgac atcggcgtga	14520
	agttcgacac taggaacttc aggctgggct gggaccccgt gaccgagctg gtcatgcccc	14580
	gggtgtacac caacgaggct ttccatcccg atattgtctt gctgccccgc tgcgggggtg	14640
	acttcaccga gagccgcctc agcaacctgc tgggcattcg caagaggcag cccttcagg	14700
	aaggettcca gatcatgtac gaggatctgg aggggggcaa catccccgcg ctcttgatg	14760
	tcgacgccta tgagaaaagc aaggaggatg cagcagctga agcaactgca gccgtagcta	14820
[0099]	ccgcctctac cgaggtcagg ggcgataatt ttgcaagcgc cgcagcagtg gcagcggccg	14880
	aggcggctga aaccgaaagt aagatagtca ttcagccggt ggagaaggat agcaagaaca	14940
	ggagctacaa cgtactaccg gacaagataa acaccgccta ccgcagctgg tacctagcct	15000
	acaactatgg cgaccccag aagggcgtgc gctcctggac gctgctcacc acctcgacg	15060
	tcacctgegg cgtggagcaa gtctactggt cgctgcccga catgatgcaa gaccgggtca	15120
	cttccgctc cacgcgtcaa gttagcaact acccgggtgtt gggcggccgag ctctgcccc	15180
	tctactcaa gagcttcttc aacgagcagg ccgtctactc gcagcagctg cgcgccttca	15240
	cctcgcttac gcacgtcttc aaccgcttcc ccgagaacca gatcctcgtc cgcccggccg	15300
	cgcccacat taccaccgtc agtgaanaac ttctgtctt cacagatcac gggaccctgc	15360
	cgctgcgcag cagtatccgg ggagtccagc gcgtgaccgt tactgacgcc agacggcgca	15420
	cctgccccta cgtctacaag gccctgggca tagtcgcgcc gcgcgtctc tcgagccgca	15480
	ccttctaaat gtccattctc atctcgccca gtaataaac cggttggggc ctgcgcgcgc	15540
	ccagcaagat gtacggaggc gctcgccaac gctccacgca acaccccgtg cgcgtgcgcg	15600

	ggcacttccg cgctccctgg ggcgcctca agggcccgct gcggtcgcgc accaccgtcg	15660
	acgacgtgat cgaccaggtg gtggccgacg cgcgcaacta ccccccgcc gccgcgccc	15720
	tctccaccgt ggacgccgtc atcgacagcg tgggtggccga cgcgcgccgg tacgcccgcg	15780
	ccaagagccg gcggcggcgc atcgcccggc ggcaccggag ccccccgcc atgcgcgcgg	15840
	cgcgagcctt gctgcgcagg gccagcgca cgggacgcag ggccatgctc agggcggcca	15900
	gacgcgcggc ttcaggcgcc agcgcggca ggaccggag acgcgcggcc acggcggcgg	15960
	cagcggccat cgccagcatg tcccggcgc ggcgagggaa cgtgtactgg gtgcgcgacg	16020
	ccgccaccgg tgtgcgcgtg cccgtgcga cccgcccc tcgacttga agatgttac	16080
	ttcgcgatgt tgatgtgtcc cagcggcgag gaggatgtcc aagcgcaaat tcaaggaaga	16140
	gatgctccag gtcacgcgc ctgagatcta cggccctgcg gtggtgaagg aggaaagaaa	16200
	gccccgaaa atcaagcggg tcaaaaagga caaaaaggaa gaagaaagtg atgtggacgg	16260
	attggtggag tttgtgcgc agttcggc cggcggcgc gtgcagtggc gcggcgga	16320
[0100]	ggtgcaaccg gtgctgagac ccggcaccac cgtggtcttc acgcccggcg agcgtcccg	16380
	caccgttcc aagcgtcct acgacaggt gtacgggat gatgatattc tggagcaggc	16440
	ggccgagcgc ctgggcgagt ttgcttacgg caagcgcagc cgttccgac cgaaggaaga	16500
	ggcgggtgctc atcccgtgg accacggcaa cccacgccg agcctcaagc ccgtgacct	16560
	gcagcaggtg ctgccgaccg cggcggcgc cggggggttc aagcgcgagg gcgaggatct	16620
	gtaccccacc atgcagctga tgggtcccaa gcgccagaag ctggaagacg tgctggagac	16680
	catgaaggtg gaccggacg tcagcccga ggtcaaggtg cggccatca agcaggtggc	16740
	ccgggctg ggcgtgcaga ccgtggacat caagattccc acggagccca tggaaacgca	16800
	gaccgagccc atgatcaagc ccagcaccag caccatggag gtgcagacgg atccctgat	16860
	gccatcggt cctagtcgaa gacccggcg caagtacggc gcggccagcc tgctgatgcc	16920
	caactacgcg ctgcatcctt ccatcatccc cacgcccggc taccgcgga cgcgcttcta	16980
	ccgcggtcat accagcagcc gccgccgaa gaccaccact cgccggccc gtcgccgac	17040
	cgccgctgca accaccctg ccgcctggt gcggagagt taccggcg gcgcgcacc	17100

	tctgaccctg ccgcgcgcgc gctaccaccc gagcatcgcc atttaaactt tcgctgctt	17160
	tgcagatcaa tggccctcac atgccgcctt cgcgttccca ttacgggcta ccgaggaaga	17220
	aaaccgcgcc gtagaaggct ggccggggaac gggatgcgtc gccaccacca ccggcggcgg	17280
	cgcccatca gcaagcggtt ggggggagc ttcctgcccg cgctgatccc catcatcgcc	17340
	ggcgatcg gggcgatccc cggcattgct tccgtggcgg tgcaggcctc tcagcgccac	17400
	tgagacacac ttgaaacat cttgtaataa accaatggac tctgacgctc ctggtcctgt	17460
	gatgtgtttt cgtagacaga tggaagacat caatttttcg tccctggtc cgcgacacgg	17520
	cacgcggccg ttcattggca cctggagcga catcggcacc agccaactga acggggcgc	17580
	ctcaattgg agcagtctct ggagcgggct taagaatttc ggtccacgc ttaaaccta	17640
	tggcagcaag gcgtggaaca gcaccacagg gcaggcctg agggataagc taaaagagca	17700
	gaacttcag cagaaggtag tcgatgggct cgcctcgggc atcaacgggg tggtagacct	17760
	ggccaaccag gccgtgcagc ggcagatcaa cagccgctg gaccgggtgc cgcggccgg	17820
	ctccgtggag atgccgcagg tggaggagga gctgcctccc ctggacaagc gggcgagaa	17880
[0101]	gcgacccgc cccgatgcgg aggagacgct gctgacgac acggacgagc cgccccgta	17940
	cgaggaggcg gtgaaactgg gtctgcccac cacgcggccc atcgcgcccc tggccaccgg	18000
	ggtgctgaaa cccgaaaagc ccgcgacct ggacttgct cctccccagc ctccccgcc	18060
	ctctacagtg gctaagcccc tgccgcccgt ggccgtggcc cgcgcgcgac ccgggggac	18120
	cgccccct catgcgaact ggcagagcac tctgaacagc atcgtgggtc tgggagtga	18180
	gagtgtgaag cgcccgcgt gctattaaac ctaccgtagc gcttaactg cttgtctgtg	18240
	tgtgtatgta ttatgtcgcc gccgcccgt tccaccagaa ggaggagtga agaggcgcgt	18300
	cgccgagttg caagatggcc accccatcga tgctgcccc gtgggcgtac atgcacatc	18360
	ccggacagga cgcttcggag tacctgagtc cgggtctggt gcagtttgcc cgcgccacag	18420
	acacctactt cagtctgggg aacaagtta ggaacccac ggtggcggcc acgcacgatg	18480
	tgaccaccga ccgcagccag cggctgacgc tgcgcttctg gccgctggac cgcgaggaca	18540
	acacctactc gtacaaagtg cgctacacgc tggccgtggg cgacaaccgc gtgctggaca	18600
	tggccagcac ctacttgac atccgcggcg tgctggatcg gggccctagc ttcaaacct	18660

	actccggcac cgctacaac agtctggccc ccaagggagc acccaacact tgtcagtgga	18720
	catataaagc cgatggtgaa actgccacag aaaaaaccta tacatatgga aatgcacccg	18780
	tgcagggcat taacatcaca aaagatggta ttcaacttgg aactgacacc gatgatcagc	18840
	caatctacgc agataaaacc tatcagcctg aacctcaagt gggatgatgct gaatggcatg	18900
	acatcactgg tactgatgaa aagtatggag gcagagctct taagcctgat accaaaatga	18960
	agccttgta tggttctttt gccaaacct ctaataaaga aggaggtcag gcaaagtga	19020
	aaacaggaac aggcactact aaagaatatg acatagacat ggctttcttt gacaacagaa	19080
	gtgcggctgc tgctggccta gctccagaaa ttgttttga tactgaaaat gtggatttgg	19140
	aaactcaga taccatatt gtatacaaag caggcacaga tgacagcagc tcttctatta	19200
	atttgggtca gcaagccatg cccaacagac ctaactacat tggtttcaga gacaacttta	19260
	tccggctcat gtactacaac agcactggca atatgggggt gctggccggt caggcttctc	19320
	agctgaatgc tgtggtgac ttgcaagaca gaaacaccga gctgtcctac cagctcttgc	19380
[0102]	ttgactctct gggtagacaga acccgtatt tcagtatgtg gaatcaggcg gtggacagct	19440
	atgatcctga tgtgcgcatt attgaaaatc atggtgtgga ggatgaactt cccaactatt	19500
	gtttccctct ggatgctgtt ggcagaacag atacttatca gggaattaag gctaattgaa	19560
	ctgatcaaac cacatggacc aaagatgaca gtgtcaatga tgctaatgag ataggcaagg	19620
	gtaatccatt cgccatgaa atcaacatcc aagccaacct gtggaggaac ttctctacg	19680
	ccaacgtggc cctgtacctg cccgactctt acaagtacac gccggccaat gttaccctgc	19740
	ccaccaacac caacacctac gattacatga acggccgggt ggtggcggcc tcgctggtgg	19800
	actcctacat caacatcggg gcgcgctggt cgctggatcc catggacaac gtgaaccct	19860
	tcaaccacca ccgcaatgcg gggctgogct accgctccat gctcctgggc aacgggcgct	19920
	acgtgccctt ccacatccag gtgccccaga aatttttcgc catcaagagc ctctgtctcc	19980
	tgcccgggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga	20040
	gctccctcgg caacgacctg cgcacggacg gggcctccat ctcttcacc agcatcaacc	20100
	tctacgccac cttcttccc atggcgcaca acacggcctc cacgctcgag gccatgctgc	20160

	gcaacgacac caacgaccag tecttcaacg actacctctc ggcggccaac atgctctacc	20220
	ccatcccggc caacgccacc aacgtgccca tctccatccc ctgcgcgaac tgggccgcct	20280
	tccgcggtg gtccttcacg cgtctcaaga ccaaggagac gcctctgctg ggetccgggt	20340
	tegaccctta cttcgtctac tcgggtccca tcccctacct cgacggcacc ttctacctca	20400
	accacacctt caagaaggtc tccatcaact tcgactcctc cgtcagctgg cccggcaacg	20460
	accggtcctt gacgcccac gagttcgaat tcaagcgac cgtcgacggc gagggtaca	20520
	acgtggccca gtgcaacatg accaaggact ggttcttggc ccagatgctg gccactaca	20580
	acatcggtta ccagggttc tacgtgcccg agggctacaa ggaccgcatg tactccttct	20640
	tccgcaactt ccagcccatt agccgccagg tgggtgacga ggtcaactac aaggactacc	20700
	aggccgtcac cctggcctac cagcacaaca actcgggctt cgtcggctac ctgcgccc	20760
	ccatgcgcca gggccagccc taccgcgcca actaccctta cccgctcctc ggcaagagcg	20820
	ccgtcaccag cgtcaccag aaaaagtcc tctgcgacag ggtcatgttg cgcctccct	20880
	tctccagcaa cttcatgtcc atgggcgcgc tcaccgacct cggccagaac atgctctatg	20940
[0103]	ccaactcgc ccacgcgta gacatgaatt tcgaagtca ccccatggat gaggccacc	21000
	ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg	21060
	tcctcgagge cgtctacctg cgcacccctt tctcggccgg taacgccacc acctaagctc	21120
	ttgcttcttg caagccatgg ccgcgggctc cggcgagcag gagctcaggc ccatcatccg	21180
	cgacctgggc tgcgggacct acttctggg caccttcgat aagcgttcc cgggattcat	21240
	ggccccgcac aagctggcct gcgccatcgt caacacggcc ggccgcgaga ccgggggcga	21300
	gcaactggctg gccttcgctt ggaaccgctc ctcgaaacc tgctacctct tcgaccctt	21360
	cgggttctcg gacgagcgc tcaagcagat ctaccagttc gactacgagg gcctgctgcg	21420
	ccgcagcgc ctggccaccg aggaccgctg cgtcacctg gaaaagtcca cccagaccgt	21480
	gcagggtccg cgctcggccg cctgcgggct cttctgctgc atgttctgc acgccttctg	21540
	gcaactggccc gaccgcccc tggacaagaa cccaccatg aacttgctga cgggggtgcc	21600
	caacggcatg ctccagtcgc cccaggtgga accaccctg cggcgcaacc aggaggcgt	21660
	ctaccgcttc ctcaactccc actccgcta cttctgctcc caccgcgcgc gcatcgagaa	21720

	ggccaccgcc ttcgaccgca tgaatcaaga catgtaaacc gtgtgtgtat gttaaatgtc	21780
	tttaataaac agcactttca tgttacacat gcatctgaga tgatttattt agaaatcgaa	21840
	agggttctgc cgggtctcgg catggcccgc gggcaggac acgttgcgga actggtactt	21900
	ggccagccac ttgaactcgg ggatcagcag tttgggcage ggggtgtcgg ggaaggagtc	21960
	ggtccacagc ttcccgctca gttgcagggc gccagcagg tcgggcgcgg agatcttgaa	22020
	atcgcagttg ggaccgcgt tctgcgcgcg ggagttgcgg tacacggggt tgcagcactg	22080
	gaacaccatc agggccgggt gcttcacgct cgcagcacc gtcgcgtcgg tgatgctctc	22140
	cacgtcgagg tcctcggcgt tggccatccc gaagggggtc atcttgcagg tctgccttcc	22200
	catggtgggc acgcaccgg gcttgtggtt gcaatgcag tgcaggggga tcagcateat	22260
	ctgggcctgg tcggcgttca tccccggta catggccttc atgaaagcct ccaattgcct	22320
	gaacgcctgc tgggccttgg ctcctcgggt gaagaagacc ccgcaggact tgctagagaa	22380
	ctggttggtg gcgcaccgg cgtcgtgcac gcagcagcgc gcgtcgttgt tggccagctg	22440
[0104]	caccacgctg cgccccagc ggttctgggt gatcttggcc cggtcggggt tctcctttag	22500
	cgcgcgctgc ccgttctcgc tcgccacatc catctcgatc atgtgctcct tctggatcat	22560
	ggtggtcccg tgcaggcacc gcagcttgcc ctccgctcgc gtgcaccgt gcagccacag	22620
	cgcgaccgg gtgcaactcc agttcttctg ggcgatctgg gaatgcgcgt gcacgaagcc	22680
	ctgcaggaag cggcccatca tgggtgctag ggtcttgttg ctagtgaagg tcagcggaat	22740
	gccgcggtgc tcctcgttga tgtacaggtg gcagatgcgg cggtagacct cgcctgctc	22800
	gggcatcagc tggaaagtgg ctttcaggtc ggtctccac cggtagcggg ccatcagcat	22860
	agtcatgatt tccataacct tctcccagc cgagacgatg ggcaggctca tagggttctt	22920
	caccatcadc ttagcgctag cagccgcggc cagggggtcg ctctcgtcca ggtctcaaa	22980
	gctccgcttg ccgtccttct cggatgatcc caccgggggg tagctgaagc ccacggccgc	23040
	cagtcctcc tcggcctgic tttcgtctc gctgtcctgg ctgacgtcct gcaggaccac	23100
	atgcttggtc ttgcggggt tcttcttggg cggcagcggc ggcggagatg ttggagatgg	23160
	cgagggggag cgcgagttct cgctcaccac tactatctct tctcttctt ggtccgaggc	23220

	cacgcggcgg taggtatgtc tcttcggggg cagaggcggg ggcgacgggc tctcgccgcc	23280
	gcgacttggc ggatggctgg cagagcccct tccgcgttcg ggggtgcgct cccggcggcg	23340
	ctctgactga cttcctccgc ggccggccat tgtgttctcc tagggaggaa caacaagcat	23400
	ggagactcag ccatcgccaa cctcgccatc tgccccacc gccgacgaga agcagcagca	23460
	gcagaatgaa agcttaaccg ccccgccgcc cagccccgcc acctccgacg cggccgtccc	23520
	agacatgcaa gagatggagg aatccatcga gattgacctg ggctatgtga cgcccgcgga	23580
	gcacgaggag gagctggcag tgcgcttttc acaagaagag atacaccaag aacagccaga	23640
	gcaggaagca gagaatgagc agagtcaggc tgggctcgag catgacggcg actacctca	23700
	cctgagcggg ggggaggacg cgctcatcaa gcatctggcc cggcaggcca ccatcgtcaa	23760
	ggatgcgctg ctcgaccgca ccgaggtgcc cctcagcgtg gaggagctca gccgcgcta	23820
	cgagtgaac ctcttctcgc cgcgcgtgcc cccaagcgc cagcccaatg gcacctgca	23880
	gccaacccg cgcctcaact tctaccgggt cttcgcggtg cccgaggccc tggccaccta	23940
	ccacatcttt ttcaagaacc aaaagatccc cgtctctcgc cgcgccaacc gcaccgcgc	24000
[0105]	cgacgcctt ttcaacctgg gtcccggcgc ccgctacct gatategect ccttgaaga	24060
	ggttccaag atcttcgagg gtctgggcag cgacgagact cgggccgca acgctctgca	24120
	aggagaagga ggagagcatg agcaccacag cgccttggtc gaggttggaag gcgacaacgc	24180
	gcggtggcg gtgtcaaac gcacggtcga gctgacctat ttgcctacc cggtctgaa	24240
	cctgcccccc aaagtcatga gcgcggtcat ggaccagggt ctcataaagc gcgcgtgcc	24300
	catctccgag gacgagggca tgcaagactc cgaggagggc aagcccgtgg tcagcgacga	24360
	gcagctggcc cgggtggctgg gtcctaatgc tagtcccag agtttggaag agcggcgcaa	24420
	actcatgatg gccgtggtcc tggtgacctg ggagctggag tgcttgcgcc gcttctcgc	24480
	cgacgcggag acctgcgca aggtcgagga gaacctgcac tacctcttca ggcacgggtt	24540
	cgtgcgccag gcctgcaaga tctccaacgt ggagctgacc aacctggtct cctacatggg	24600
	catcttgca gagaaccgcc tggggcagaa cgtgctgcac accaccctgc gcggggaggc	24660
	ccggcgcgac tacatccgag actgcgtcta cctctaccte tgccacacct ggcagacggg	24720
	catgggcgtg tggcagcagt gtctggagga gcagaacctg aaagagctct gcaagctcct	24780

	gcagaagaac ctcaagggtc tgtggaccgg gttcgacgag cgcaccaccg cctcggacct	24840
	ggccgacctc attttccccg agegcctcag gctgacgctg cgcaacggcc tgccccgactt	24900
	tatgagccaa agcatgttgc aaaactttcg ctctttcatc ctggaacgct ccggaatcct	24960
	gccccccacc tgctccgcgc tgccctcgga ctctcgtgcc ctgaccttcc gcgagtgcc	25020
	ccccccgctg tggagccact gctacctgct gcgcctggcc aactacctgg cctaccactc	25080
	ggacgtgacg gaggacgtca gcggcgaggg cctgctcagag tgccactgcc gctgcaacct	25140
	ctgcacgccc caccgctccc tggcctgcaa cccccagctg ctgagcgaga cccagatcat	25200
	cggcaccttc gagttgcaag ggcccagcga aggcgagggc tcagcccca aggggggtct	25260
	gaaactcacc ccggggctgt ggacctcggc ctacttgcgc aagttcgtgc ccgaggacta	25320
	ccatcccttc gagatcaggt tctacgagga ccaatccat ccgccaagc ccgagctgtc	25380
	ggcctgcgtc atcaccagcgg gggcgatcct ggcccattg caagccatcc agaaatccc	25440
	ccaagaattc ttgctgaaaa agggccgcgg ggtctacctc gacccccaga ccggtgagga	25500
[0106]	gctcaacccc ggcttcccc aggatgcccc gaggaaacaa gaagctgaaa gtggagctgc	25560
	cgcccgtgga ggatttggag gaagactggg agaacagcag tcaggcagag gaggaggaga	25620
	tggaggaaga ctgggacagc actcaggcag aggaggacag cctgcaagac agtctggagg	25680
	aagacgagga ggaggcagag gaggaggtgg aagaagcagc cgccgccaga ccgtcgtcct	25740
	cggcggggga gaaagcaagc agcacggata ccatctccgc tccgggtcgg ggtcccgtc	25800
	gaccacacag tagatgggac gagaccggac gattcccga cccaccacc cagaccgta	25860
	agaaggagcg gcagggatac aagtcctggc gggggcaca aaacccatc gtctcctgct	25920
	tgcaggcctg cgggggcaac atctcctca cccggcgcta cctgctcttc caccgcgggg	25980
	tgaactttcc ccgcaacatc ttgcattact accgtcacct ccacagcccc tactacttcc	26040
	aagaagaggc agcagcagca gaaaaagacc agcagaaaac cagcagctag aaaatccaca	26100
	gcggcggcag caggtggact gaggatcgcg gcgaacgagc cggcgcaaac ccgggagctg	26160
	aggaaccgga tctttccac cctctatgcc atcttccagc agagtcgggg gcaggagcag	26220
	gaactgaaag tcaagaaccg ttctctgcgc tcgctcacc gcagttgtct gtatcacaag	26280

	agcgaagacc aacttcagcg cactctcgag gacgccgagg ctctcttcaa caagtactgc	26340
	gcgctcactc ttaaagagta gcccgcgccc gcccgatcgc agaaaaaggc gggaattacg	26400
	tcacctgtgc ccttcgcctt agccgcctcc acccatcacc atgagcaaag agattcccac	26460
	gccttacatg tggagctacc agccccagat gggcctggcc gccggtgccg cccaggacta	26520
	ctccacccgc atgaattggc tcagcgccgg gcccgcatg atctcacggg tgaatgacat	26580
	ccgcgcccac cgaaaccaga tactcctaga acagtcagcg ctcaccgcca cgccccgaa	26640
	tcacctcaat ccgcgtaatt ggccccgcgc cctgggtgac caggaaatc cccagcccac	26700
	gaccgtacta cttccgcgag acgcccaggc cgaagtccag ctgactaact cagggtgcca	26760
	gctggcgggc ggcgccacc tgtgtcgtca ccgccccgt cagggtataa agcggctggt	26820
	gatccggggc agaggcacac agctcaacga cgaggtggtg agctcttcgc tgggtctgcg	26880
	acctgacgga gtcttccaac tcgccgcatc ggggagatct tccttcacgc ctgctcaggc	26940
	cgctctgact ttggagagtt cgtcctcgca gccccgctcg ggtggcatcg gcaactcca	27000
[0107]	gttcgtggag gagttcactc cctcggctca cttcaacccc ttctccggt cccccgcca	27060
	ctaccggac gagttcacc cgaacttga cgccatcagc gagtcggtgg acggctacga	27120
	ttgaatgtcc catggtggcg cagctgacct agctcggctt cgacacctgg accactgccg	27180
	ccgcttcgc tgcttcgctc gggatctcgc cgagtttgcc tactttgagc tgcccagga	27240
	gcaccctcag ggccccgccc acggagtgcg gatcgtcgtc gaagggggcc tcgactcca	27300
	cctgcttcgg atcttcagcc agcgtccgat cctggctcag cgcgagcaag gacagacct	27360
	tctgactctg tactgcatct gcaaccaccc cggcctgcat gaaagtcttt gttgtctgct	27420
	gtgtactgag tataataaaa gctgagatca gcgactactc cggacttccg tgtgttctg	27480
	aatccatcaa ccagtctttg ttcttcaccg ggaacgagac cgagctccag ctccagtgt	27540
	agccccacaa gaagtacctc acctggctgt tccagggctc cccgatcgc gttgtcaacc	27600
	actgcgacaa cgacggagtc ctgctgagcg gccctgcaaa cttactttt tccaccgca	27660
	gaagcaagct ccagctcttc caacccttc tccccggac ctatcagtgc gtctcggac	27720
	cctgccatca caccttcac ctgatcccga ataccacagc gtcgctcccc gtactaaca	27780
	accaaactaa cctccacaa cgccaccgtc gcgaccttc tgaatctaat actaccacc	27840

	acaccggagg tgagctccga ggtcaaccaa cctctgggat ttactacggc ccctgggagg	27900
	tggttgggtt aatagegcta ggcctagttg cgggtgggct tttggttctc tgctacctat	27960
	acctcccttg ctgttcgtac ttagtgggtc tgtgttgctg gtttaagaaa tggggaagat	28020
	caccctagtg agctgcggtg cgctgggtgc ggtgttgctt tcgattgtgg gactgggcgg	28080
	tgcggtgta gtgaaggaga aggccgatcc ctgcttgcat ttcaatccca acaaatgcca	28140
	gctgagtttt cagcccgatg gcaatcgggtg cgcggtactg atcaagtgcg gatgggaatg	28200
	cgagaacgtg agaatcgagt acaataacaa gactcggaac aatactctcg cgtccgtgtg	28260
	gcagcccggg gaccccgatg ggtacaccgt ctctgtcccc ggtgctgacg gctccccgcg	28320
	caccgtgaat aatactttca tttttgcga catgtgcgac acggtcatgt ggatgagcaa	28380
	gcagtacgat atgtggcccc ccacgaagga gaacatcgtg gtcttctcca tcgettacag	28440
	cctgtgcacg gcgctaatca ccgctatcgt gtgcctgagc attcacatgc tcatcgctat	28500
	tcgccccaga aataatgccg aaaaagaaaa acagccataa cgtttttttt cacacctttt	28560
[0108]	tcagaccatg gcctctgtta aatttttgct tttatttgce agtctcattg ccgtcattca	28620
	tggaatgagt aatgagaaaa ttactattta cactggcact aatcacacat tgaaggtcc	28680
	agaaaaagcc acagaagttt catggtattg ttattttaat gaatcagatg tatctactga	28740
	actctgtgga aacaataaca aaaaaaatga gagcattact ctcatcaagt ttcaatgtgg	28800
	atctgactta accctaatta acatcactag agactatgta ggtatgtatt atggaactac	28860
	agcaggcatt tcggacatgg aattttatca agtttctgtg tctgaacca ccacgcctag	28920
	aatgaccaca accacaaaa ctacacctgt taccactatg cagctcacta ccaataacat	28980
	ttttgccatg cgtcaaatgg tcaacaatag cactcaacce accccacca gtgaggaaat	29040
	tcccaaatcc atgattggca ttattgttgc tgtagtggig tgcatttga tcatgcctt	29100
	gtgcatggtg tactatgcct tctgctacag aaagcacaga ctgaacgaca agctggaaca	29160
	cttactaagt gttgaatfff aatfffittag aaccatgaag atcctaggcc ttttaatfff	29220
	ttctatcatt acctctgctc tatgcaattc tgacaatgag gacgttactg tcgttgctcg	29280
	atcaaattat aactgaaag gtccagcgaa ggtatgctt tcgtggtatt gctatfffgg	29340

	atctgacact acagaaactg aattatgcaa tcttaagaat ggcaaaatc aaaattctaa	29400
	aattaacaat tataatgca atggtactga tctgatactc ctcaatatca cgaaatcata	29460
	tgctggcagt tacacctgcc ctggagatga tgctgacagt atgatttttt acaaagtaac	29520
	tgttgttgat cccactactc cacctccacc caccacaact actcacacca cacacacaga	29580
	tcaaaccgca gcagaggagg cagcaaagtt agccttgacg gtccaagaca gttcatttgt	29640
	tggcattacc cctacacctg atcagcgggtg tccggggctg ctagtcagcg gcattgtcgg	29700
	tgtgctttcg ggattagcag tcataatcat ctgcatgttc atttttgctt gctgctatag	29760
	aaggctttac cgacaaaaat cagaccctact gctgaacctc tatgtttaat tttttccaga	29820
	gtcatgaagg cagttagcgc tctagttttt tgttctttga ttggcattgt tttttgcaat	29880
	cctattccta aagttagctt tattaagat gtgaatgtta ctgagggggg caatgtgaca	29940
	ctggtaggtg tagagggtgc tgaaaacacc acctggacaa aataccacct caatgggtg	30000
	aaagatattt gcaattggag tgtattagtt tatacatgtg agggagttaa tcttaccatt	30060
[0109]	gtcaatgcca cctcagctca aaatggtaga attcaaggac aaagtgtcag tgtatctaat	30120
	gggtatttta cccaacatac ttttatctat gacgttaaag tcataccact gcctacgcct	30180
	agcccaccta gcactaccac acagacaacc cacactacac agacaaccac atacagtaca	30240
	ttaaatcagc ctaccaccac tacagcagca gaggttgcca gctcgtctgg ggtccgagtg	30300
	gcatttttga tgttgcccc atctagcagt cccactgcta gtaccaatga gcagactact	30360
	gaatttttgt ccaactgtga gagccacacc acagctacct ccagtgcctt ctetagcacc	30420
	gccaatctct cctcgtttc ctctacacca atcagtcceg ctactactcc tagccccgct	30480
	cctcttcca ctcccctgaa gcaaacagac ggcggcatgc aatggcagat caccctgctc	30540
	attgtgatcg ggttggtcat cctggccgtg ttgctctact acatcttctg ccgccgatt	30600
	cccaacgcgc accgcaagcc ggtctacaag cccatcattg tcgggcagcc ggagccgctt	30660
	caggtggaag ggggtctaag gaatcttctc ttctctttta cagtatgggtg attgaactat	30720
	gattcctaga caattcttga tcactattct tatctgcctc ctccaagtct gtgccacct	30780
	cgctctggtg gccaacgcca gtccagactg tattggcccc ttgcctcct acgtgctctt	30840
	tgccttcacc acctgcatct gctgctgtag catagtctgc ctgcttatca ccttcttcca	30900

	gttcattgac tggatctttg tgcgcatcgc ctacctgcgc caccaccccc agtaccgga	30960
	ccagegagtg gcgcggtgc tcaggctcct ctgataagca tgcgggctct getacttctc	31020
	gcgcttctgc tgttagtgc cccccgtccc gtcgaccccc ggtccccac ccagtcccc	31080
	gaggaggtcc gcaaatgcaa attccaagaa ccctggaaat tectcaaatg ctaccgcaa	31140
	aatcagaca tgcattcccag ctggatcatg atcattggga tcgtgaacat tctggcctgc	31200
	acctcatct cctttgtgat ttaccctgc ttgactttg gttggaactc gccagaggcg	31260
	ctctatctcc cgctgaacc tgacacacca ccacagcaac ctcaggcaca cgcactacca	31320
	ccactacagc ctaggccaca atacatgccc atattagact atgaggccga gccacagcga	31380
	cccatgctcc ccgctattag ttacttcaat ctaaccggcg gagatgactg acccactggc	31440
	caacaacaac gtcaacgacc ttctcctgga catggacggc cgcgcctcgg agcagcgact	31500
	cgcccaactt cgcattcgc agcagcagga gagagccgc aaggagctgc aggatgcggt	31560
	ggccatccac cagtgaaga gaggcatctt ctgcctggtg aaacaggcca agatctccta	31620
[0110]	cgaggtcaact ccaaacgacc atgcctctc ctacgagctc ctgcagcagc gccagaagt	31680
	cacctgcctg gtcggagtca accccatcgt catcaccag cagtctggcg ataccaagg	31740
	gtgcatccac tgctctcgc actccccga ctgcgtccac actctgatca agaccctctg	31800
	cggcctccgc gacctctcc ccatgaacta atcacccct tatccagtga aataaagatc	31860
	atattgatga tgattttaca gaaataaaaa ataatcattt gatttgaat aaagatacaa	31920
	tcatatgat gatttgagt taacaaaaa ataaagaatc acttacttga aatctgatac	31980
	caggtctctg tccatgtttt ctgccaacac cacttactc ccctcttccc agctctggta	32040
	ctgcaggccc cggcgggctg caaacttct ccacacgctg aaggggatgt caaattctc	32100
	ctgtccctca atcttcattt tatctctat cagatgtcca aaaagcgcgt ccgggtggat	32160
	gatgactteg acccgtcta cccctacgat gcagacaac caccgaccgt gcccttcatc	32220
	aacccccct tcgtctctc agatggattc caagagaagc ccctgggggt gttgtccctg	32280
	cgactggccg acccgtcac caccaagaac ggggaaatca ccctcaagct gggagagggg	32340
	gtggacctcg attcctcggg aaaactcatc tccaacacgg ccaccaagc ccccgcct	32400

	ctcagttttt ccaacaacac catttccctt aacatggatc acccctttta cactaaagat	32460
	ggaaaattat ccttacaagt ttctccacca ttaaataac tgagaacaag cattctaac	32520
	acactagctt taggttttgg atcaggttta ggactccgtg gctctgcctt ggcagtacag	32580
	ttagtctctc cacttacatt tgatactgat ggaaacataa agcttacctt agacagaggt	32640
	ttgcatgtta caacaggaga tgcaattgaa agcaacataa gctgggctaa aggtttaaaa	32700
	tttgaagatg gagccatagc aaccaacatt ggaaatgggt tagagtttgg aagcagtagt	32760
	acagaaacag gtgttgatga tgcttaccca atccaagtta aacttggatc tggccttagc	32820
	tttgacagta caggagccat aatggctggt aacaaagaag acgataaact cactttgtgg	32880
	acaacacctg atccatcacc aaactgtcaa atactcgcag aaaatgatgc aaaactaaca	32940
	ctttgcttga ctaaagtgg tagtcaaata ctggccactg tgtcagtctt agttgtagga	33000
	agtggaaacc taaaccccat tactggcacc gtaagcagtg ctcaggtgtt tctacgtttt	33060
	gatgcaaacg gtgttctttt aacagaacat tctacactaa aaaaactctg ggggtatagg	33120
	caggagata gcatagatgg cactccatat accaatgctg taggattcat gcccaattta	33180
[0111]	aaagcttate caaagtcaca aagttctact actaaaaata atatagtagg gcaagtatac	33240
	atgaatggag atgtttcaaa acctatgctt ctactataa ccctcaatgg tactgatgac	33300
	agcaacagta catattcaat gtcattttca tacacctgga ctaatggaag ctatgttggg	33360
	gcaacatttg gggctaacte ttataccttc tcatacatcg cccaagaatg aacctgtat	33420
	cccaccctgc atgccaacce ttcccacccc actctgtgga acaaactctg aacacaaaa	33480
	taaaataaag ttcaagtgtt ttattgattc aacagtttta caggattcga gcagttattt	33540
	ttctccacc ctcccaggac atggaataca ccacctctc cccccgaca gccttgaaca	33600
	tctgaatgcc attggtgatg gacatgcttt tggctctcac gttccacaca gtttcagagc	33660
	gagccagtct cgggtcggtc agggagatga aaccctccgg gcactcccgc atctgcacct	33720
	cacagctcaa cagctgagga ttgtcctcgg tggctgggat cacggttatac tggaagaagc	33780
	agaagagcgg cgggtgggaat catagtccgc gaacgggatc ggccggtggt gtcgcatcag	33840
	gccccgcagc agtcgctgcc gccccgcctc cgtcaagctg ctgctcaggg ggtccgggtc	33900
	cagggactcc ctcagcatga tgcccacggc cctcagcatc agtcgtcttg tgcggcgggc	33960

	gcagcagcgc atgcggatct cgctcaggtc gctgcagtac gtgcaacaca gaaccaccag	34020
	gttgttcaac agtccatagt tcaacacgct ccagccgaaa ctcatcgcg gaaggatget	34080
	accacagtgg ccgctgtacc agatcctcag gtaaatacaag tgggtccccc tccagaacac	34140
	gctgcccacg tacatgatct ccttgggcat gtggcggttc accacctccc ggtaccacat	34200
	cacctctggt ttgaacatgc agccccggat gatcctgcgg aaccacaggg ccagcaccgc	34260
	cccgccgcc atgcagcga gagaccccgg gtcccggcaa tggcaatgga ggaccaccg	34320
	ctcgtaccgg tggatcatct gggagctgaa caagtctatg ttggcacagc acaggcatat	34380
	gctcatgcat ctcttcagca ctctcaactc ctcgggggtc aaaaccatat ccagggcac	34440
	ggggaactct tgcaggacag cgaaccccgc agaacagggc aatcctcgca cagaacttac	34500
	attgtgcatg gacagggtat cgcaatcagg cagcaccggg tgatcctcca ccagagaagc	34560
	gcgggtctcg gtctctcac agcgtggtaa gggggccggc cgatacgggt gatggcggga	34620
	cgcggtgat cgtgttcgcg accgtgtcat gatgcagttg ctttcggaca ttttcgtact	34680
[0112]	tgctgtagca gaacctggtc cgggcgctgc acaccgatcg ccggcggcgg tctcggcgct	34740
	tggaacgctc ggtgttgaaa ttgtaaaaca gccactctct cagaccgtgc agcagatcta	34800
	gggcctcagg agtgatgaag atcccatcat gctgatggtc tctgatcaca tcgaccaccg	34860
	tggaatgggc cagaccagc cagatgatgc aattttgttg ggtttcggtg acggcggggg	34920
	agggaagaac aggaagaacc atgattaact tttaatcaa acggtctcgg agtacttcaa	34980
	aatgaagatc gcggagatgg cacctctcgc ccccgctgtg ttggtggaaa ataacagcca	35040
	ggtcaaaggt gatacggttc tcgagatggt ccacgggtgc ttccagcaaa gcctccacgc	35100
	gcacatccag aaacaagaca atagcgaag cgggagggtt ctctaattcc tcaatcatca	35160
	tgttacactc ctgcaccatc ccagataat ttctatttt ccagccttga atgattcgaa	35220
	ctagtctctg aggtaaatcc aagccagcca tgataaagag ctcgccgaga gcgccctcca	35280
	ccggcattct taagcacacc ctcataattc caagatattc tgctcctggt tcacctgcag	35340
	cagattgaca agcggaatat caaaatctct gcccgatcc ctgagctcct ccctcagcaa	35400
	taactgtaag tactctttca tctctctcc gaaattttta gccataggac caccaggaat	35460

	aagattaggg caagccacag tacagataaa ccgaagtccct cccagtgag cattgccaaa	35520
	tgcaagactg ctataagcat gctggctaga cccggtgata tcttcagat aactggacag	35580
	aaaatgccc aggcaatfff taagaaaatc aacaaaagaa aaatcctcca ggtggacgtt	35640
	tagagcctcg ggaacaacga tgaagtaaat gcaagcggg cgttcagca tggtagtta	35700
	gctgatctgt agaaaaaca aaaatgaaca ttaaaccatg ctagcctggc gaacaggtgg	35760
	gtaaatcggt ctctccagca ccaggcaggc cacgggtct cggcgcgac cctcgtaaaa	35820
	attgtcgcta tgattgaaaa ccatcacaga gagacgttcc cggtgccgg cgtgaatgat	35880
	tcgacaagat gaatacacc ccggaacatt ggcgtccgag agtgaaaaa agcggccgag	35940
	gaagcaataa ggcaactaca tgctcagtct caagtccagc aaagcagatc catgcggatg	36000
	aagcacaaaa ttctcaggtg cgtacaaaat gtaattactc ccctctgca caggcagcaa	36060
	agccccgat ccctccaggt acacatacaa agcctcagcg tccatagett accgagcagc	36120
	agcacacaac aggcgcaaga gtcagagaaa ggctgagctc taacctgtcc acccgctctc	36180
	tgctcaatat atagcccaga tctacactga cgtaaaggcc aaagtctaaa aataccgcc	36240
[0113]	aaataatcac acacgccag cacacgcca gaaaccggtg acacactcaa aaaaatacgc	36300
	gcacttctc aaacgccc aaactgccc atttccgggt tcccacgcta cgtcatcaaa	36360
	acacgacttt caaatccgt cgaccgttaa aaacgtcacc cgccccgccc ctaacggtcg	36420
	cccgtctctc agccaatcag cgccccgat ccccaaatc aaacacctca tttgcatatt	36480
	aacgcgcaca aaaagtttga ggtatattat tgatgatgg	36519
	<210> 11	
	<211> 31867	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 11	
	ccatcttcaa taatatacct caaactffff gtgcgcgtta atatgcaaat gaggcgtttg	60
	aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgttagg ggcggggcga	120
	gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag	180

	tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttcccgc gctctctgac	240
	agghaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact	300
	gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg cagggaggag tatttgccga	360
	gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa	420
	tttccgctga cgggtgtcaaa gtccgggtgtt tttacgtagg tgtcagctga tcgccagggt	480
	atttaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcgag aagagttttc	540
	tectccgcgc cgcgagtcag atctacactt tgaagtagg gataacaggg taatgacatt	600
	gattattgac tagttgttaa tagtaatcaa ttacggggtc attagtcat agcccatata	660
	tggagtccg cgttacataa cttacggtaa atggccccgc tggctgaccg cccaacgacc	720
	cccgccatt gacgtcaata atgacgatg ttcccatagt aacgccaata gggactttcc	780
	attgacgtca atgggtggag tatttacggg aaactgcccc cttggcagta catcaagtgt	840
	atcatatgcc aagtccgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt	900
[0114]	atgcccagta catgacctta cgggactttc ctacttgcca gtacatctac gtattagtca	960
	tcgctattac catggtgatg cggttttggc agtacaccaa tgggcgtgga tagcggtttg	1020
	actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc	1080
	aaaatcaacg ggactttcca aaatgtcgta ataacccccgc cccgttgacg caaatgggcg	1140
	gtaggcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg	1200
	cctggaacgc catccacget gttttgacct ccatagaaga cagcgatcgc gccaccatgg	1260
	tgagcaaggg cgaggagctg ttcaccgggg tgggtgccat cctggtcgag ctggacggcg	1320
	acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc acctacggca	1380
	agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg cccaccctcg	1440
	tgaccacct gacctacggc gtgcagtget tcagccgcta ccccaccac atgaagcagc	1500
	acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc atcttcttca	1560
	aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac acctgtgtga	1620
	accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg gggcacaagc	1680

	tggagtacaa ctacaacagc cacaacgtct atatcatggc cgacaagcag aagaacggca	1740
	tcaaggtgaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag ctgccgacc	1800
	actaccagca gaacaccccc atcggcgacg gccccgtgct gctgccccgac aaccactacc	1860
	tgagcaccca gtccgccctg agcaaagacc ccaacgagaa gcgcatcac atggtcctgc	1920
	tggagttcgt gaccgccgc gggatcactc tcggcatgga cgagctttac aagtagtgag	1980
	tttaactcc catttaaatg tgagggttaa tgcttcgagc agacatgata agatacattg	2040
	atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt	2100
	gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca	2160
	attgcattca ttttatgttt caggttcagg gggagatgtg ggaggttttt taaagcaagt	2220
	aaaacctcta caaatgtggt aaaataacta taacggtcct aaggtagcga gtgagtagtg	2280
	ttctggggcg ggggaggacc tgcatgaggg ccagaataac taaaatctgt gcttttctgt	2340
	gtgttcgagc agcatgagcg gaagcggctc ctttgagga ggggtattca gcccttatct	2400
	gacggggcgt ctcccctct gggcgggagt gcgtcagaat gtgatgggat ccacggtgga	2460
[0115]	cggccggccc gtgcagccc cgaactcttc aacctgacc tatgcaacce tgagctcttc	2520
	gtcgttgac gcagctgcc ccgcagctgc tgcatctgcc gccagcggc tgcgcggaat	2580
	ggccatgggc gccggctact acggcactct ggtggccaac tcgagttcca ccaataatcc	2640
	cgccagcctg aacgaggaga agctgttgct gctgatggcc cagctcgagg ccttgacca	2700
	gcgcctgggc gagctgacce agcaggtggc tcagctgcag gagcagacgc gggccgcggt	2760
	tgccacggtg aaatccaaat aaaaaatgaa tcaataaata aacggagacg gttgttgatt	2820
	ttaacacaga gtctgaatct ttatttgatt tttcgcgcg gtaggcct ggaccaccgg	2880
	tctgatcat tgagcaccg gtgatcttt tccaggacce gtagagggtg ggcttgatg	2940
	ttgaggtaca tgggcatgag cccgtcccgg gggtagggt agctccattg caggcctcg	3000
	tgctcggggg tgggttgta aatcaccag tcatagcagg ggcgcagggc atggtgttgc	3060
	acaatatctt tgaggaggag actgatggcc acgggcagcc ctttggtgta ggtgtttaca	3120
	aatctgttga gctgggagg atgcatcgg ggggagatga ggtgcatctt ggccctggatc	3180
	ttgagattgg cgatgttacc gccagatcc cgcctggggt tcatgttgtg caggaccacc	3240

	agcacggtgt atccggtgca cttggggaat ttatcatgca acttgggaagg gaaggcgtga	3300
	aagaatttgg cgacgccttt gtgcccgcc aggttttcca tgcactcacc catgatgatg	3360
	gcgatgggcc cgtgggcggc ggcctgggca aagacgtttc gggggtcggga cacatcatag	3420
	ttgtggtcct ggggtgagtc atcataggcc attttaatga atttggggcg gaggggtgcc	3480
	gactggggga caaaggtacc ctcgatcccc ggggcgtagt tcccctcaca gatctgcatc	3540
	tcccaggctt tgagctcggga gggggggatc atgtccacct gcggggcgat aaagaacacg	3600
	gtttccgggg cgggggagat gagctgggcc gaaagcaagt tccggagcag ctgggacttg	3660
	ccgcagccgg tggggccgta gatgaccccc atgaccggct gcaggtggta gttgaggag	3720
	agacagctgc cgtcctccc gaggaggggg gccacctcgt tcatcatctc gcgcacgtgc	3780
	atgttctcgc gcaccagttc cgccaggagg cgctctcccc ccagggatag gagctcctgg	3840
	agcgaggcga agttttcag cggcttgagt ccgtcggcca tgggcatttt ggagagggtt	3900
	tgttcaaga gttccaggcg gtcccagagc tcggtgatgt gctctacggc atctcgatcc	3960
[0116]	agcagacctc ctcgtttcgc gggttgggac ggctgcggga gtagggcacc agacgatggg	4020
	cgtccagcgc agccagggtc cggtccttcc agggtcgcag cgtccgcgtc aggggtggtct	4080
	ccgtcacggt gaaggggtgc gcgccgggct gggcgcttgc gaggtgccc ttcaggctca	4140
	tccgctggt cgaaaaccgc tcccgatcgg cgccctgcgc gtcggccagg tagcaattga	4200
	ccatgagttc gtagttgagc gcctcggccg cgtggccttt ggcgcggagc ttacctttgg	4260
	aagtctgccc gcaggcggga cagaggagg acttgagggc gtagagcttg ggggcgagga	4320
	agacggactc gggggcgtag gcgtccgcgc cgcagtgggc gcagacggtc tcgactcca	4380
	cgagccaggt gaggtcgggc tggtcggggt caaaaaccag tttcccgcg ttctttttga	4440
	tgcgtttctt acctttggtc tccatgagct cgtgtcccc ctgggtgaca aagaggctgt	4500
	ccgtgtcccc gtagaccgac tttatgggcc ggtcctcgag cgggtgtccc cggtcctcct	4560
	cgtagaggaa ccccgccac tccgagacga aagcccgggt ccaggccagc acgaaggagg	4620
	ccacgtggga cgggtagcgg tcgtgtcca ccagcgggtc caccttttcc aggttatgca	4680
	aacacatgtc cccctcgtcc acatccagga aggtgattgg cttgtaagt taggccacgt	4740

	gaccgggggt cccggccggg ggggtataaa aggggtgcggg tccttgctcg tcctcactgt	4800
	cttccggatc gctgtccagg agcgcagct gttgggtag gtattccctc tcgaaggcgg	4860
	gcatgacctc ggcaactcagg ttgtcagttt ctagaaacga ggaggatttg atattgacgg	4920
	tgccggcgga gatgcctttc aagagcccct cgtccatctg gtcagaaaag acgatctttt	4980
	tgttgtcgag cttggtggcg aaggagccgt agagggcgtt ggagaggagc ttggcgatgg	5040
	agcgcattgt ctggtttttt tccttgctcg cgcgctcctt ggccggcagc ttgagctgca	5100
	cgtaactcgc cgccacgcac ttccattcgg ggaagacggt ggtcagctcg tcgggcacga	5160
	ttctgacctg ccagccccga ttatgcaggg tgatgaggtc cacactgggtg gccacctcgc	5220
	cgcgaggggg ctcatagtc cagcagaggc gtccgccctt gcgcgagcag aaggggggca	5280
	gggggtccag catgacctcg tcgggggggt cggcatcgat ggtgaagatg ccgggcagga	5340
	ggtcggggtc aaagtagctg atggaagtgg ccagatcgtc cagggcagct tgccattcgc	5400
	gcacggccag cgcgctctcg tagggactga ggggcgtgcc ccagggcagc ggatgggtaa	5460
[0117]	gcgcggaggc gtacatgccg cagatgtcgt agacgtagag gggctcctcg aggatgccga	5520
	tgtaggtggg gtagcagcgc ccccccgga tgctggcgcg cacgtagtca tacagctcgt	5580
	gcgagggggc gaggagcccc gggcccaggt tgggtgcgact gggcttttcg gcgcggtaga	5640
	cgatctggcg gaaaatggca tgcgagttgg aggagatggt gggcctttgg aagatgttga	5700
	agtgggcgtg gggcagtcgc accgagtcgc ggatgaagtg ggcgtaggag tcttgcagct	5760
	tggcgacgag ctcgccggtg actaggacgt ccagagcgca gtagtcgagg gtctcctgga	5820
	tgatgtcata cttgagctgt ccctttgtt tccacagctc gcggttgaga aggaactctt	5880
	cgcggtcctt ccagtactct tcgaggggga acccgtcctg atctgcacgg taagagccta	5940
	gcatgtagaa ctggttgacg gccttgtagg cgcagcagcc cttctccacg gggagggcgt	6000
	aggcctgggc ggccctgcgc agggagggtg gcgtgagggc gaaagtgtcc ctgacctga	6060
	ccctgaggaa ctggtgcttg aagtcgatat cgtcgcagcc ccctgctcc cagagctgga	6120
	agtccgtgcg cttctttagt gcggggttgg gcaaagcgaa agtaacatcg ttgaagagga	6180
	tcttccccgc gcggggcata aagttgcgag tgatgcggaa aggttggggc acctcgcccc	6240
	ggttgtgat gacctggcg gcgagcacga tctcgtcgaa gccgttgatg ttgtggccca	6300

	cgatgtagag ttccacgaat cgcggacggc ccttgacgtg gggcagtttc ttgagctcct	6360
	cgtaggtgag ctctegggg tegctgagcc cgtgctgctc gagcgcccag tcggcgagat	6420
	gggggttggc gcggaggaag gaagtccaga gatccacggc cagggcggtt tgcagacggt	6480
	cccgtactg acggaactgc tgcccagcgg ccattttttc gggggtgacg cagtagaagg	6540
	tgcgggggtc cccgtgccag cgatcccatt tgagctggag ggcgagatcg agggcgagct	6600
	cgacgagccg gtctgccccg gagagtttca tgaccagcat gaaggggacg agctgcttgc	6660
	cgaaggacc catccaggtg taggtttcca catcgtaggt gaggaagagc ctttcggtgc	6720
	gaggatgcga gccgatgggg aagaactgga tctcctgcc ccaattggag gaatggctgt	6780
	tgatgtgatg gaagtagaaa tgccgacggc gcgccgaaca ctctgtcttg tgtttataca	6840
	agcggccaca gtgctcgcaa cgctgcacgg gatgcacgtg ctgcacgagc tgtacctgag	6900
	ttcctttgac gaggaatttc agtgggaagt ggagtcgtgg cgcctgcac tcgtgctgta	6960
	ctacgtctg gtggtcggcc tggccctctt ctgcctcgat ggtggtcatg ctgacgagcc	7020
[0118]	cgcggggag gcaggtccag acctcggcgc gagcgggtcg gagagcgagg acgagggcgc	7080
	gcaggccgga gctgtccagg gtctgagac gctgcggagt caggtcagtg ggcagcggcg	7140
	gcgcgcggtt gacttgcagg agtttttcca gggcgcgcgg gaggtccaga tggacttga	7200
	tctccaccgc gccattggtg gcgacgtcga tggcttgacg ggtcccgtgc ccttggggtg	7260
	tgaccaccgt cccccgttcc ttcttgggcg gctggggcga cggggcggt gcctcttcca	7320
	tggttagaag cggcggcgag gacgcgcgc gggcggcagg ggcggctcgg ggcccggagg	7380
	cagggcggc aggggcacgt cggcggcgcg cgcgggtagg ttctggtact gccccggag	7440
	aagactggcg tgagcgacga cgcgacggtt gacgtcctgg atctgacgcc tctgggtgaa	7500
	ggccacggga cccgtgagtt tgaacctgaa agagagttcg acagaatcaa tctcgtatc	7560
	gttgacggcg gcctgccgca ggatctcttg cacgtcgcg gagttgtcct ggtaggcgat	7620
	ctcggctatg aactgctcga tctcctctc ttgaaggctt ccgcgcccg cgcgctccac	7680
	ggtggcccg aggtcgttgg agatgcggcc catgagctgc gagaaggcgt tcatgcccgc	7740
	ctcgttccag acgcggctgt agaccacgac gccctcggga tcgcggcgcc gcatgaccac	7800

	ctgggcgagg ttgagctcca cgtggcgcgt gaagaccgcg tagttgcaga ggcgctggtg	7860
	gaggtagttg agcgtggtgg cgatgtgctc ggtgacgaag aaatacatga tccagcggcg	7920
	gagcggcadc tegctgacgt cgcccagcgc ctccaaacgt tccatggcct cgtaaaagtc	7980
	cacggcgaag ttgaaaaact gggagttgcg cgccgagacg gtcaactcct cctccagaag	8040
	acggatgagc tcggcgatgg tggcgcgcac ctccgctcg aaggccccg ggagtctctc	8100
	cacttctct tcttctct ccaactaacat ctcttctact tctctctcag gcggcagtgg	8160
	tggcggggga gggggcctgc gtcgccggcg gcgcacgggc agacggtcga tgaagcctc	8220
	gatggtctcg ccgcgccggc gtcgcatggt ctccggtgac gcgcgcccg cctcgcgggg	8280
	ccgcagcgtg aagacgccg cgcgcatctc caggtggccg ggggggtccc cgttgggcag	8340
	ggagagggcg ctgacgatgc atcttatcaa ttgccccgta gggactccgc gcaaggacct	8400
	gagcgtctcg agatccacgg gatctgaaaa ccgctgaacg aaggcttcga gccagtcgca	8460
	gtcgaaggt agcctgagca cggtttcttc tggcgggtca tgttggttg gagcggggcg	8520
[0119]	ggcgatgctg ctggtgatga agttgaaata ggcggttctg agacggcggg tggtagcgag	8580
	gagcaccagg tctttgggcc cggcttctg gatgcgcaga cggtcggcca tgccccagc	8640
	gtggtctga cacctggcca ggtccttga gtagtctgc atgagccgct ccacgggcac	8700
	ctctctctc cccgcgcggc cgtgcatgcg cgtgagccc aagccgcgct ggggctggac	8760
	gagcggcagg tcggcgacga cgcgctcggc gaggatggct tgctggatct gggtagggg	8820
	ggtctggaag tcatcaaagt cgacgaagcg gtggtaggct ccggtgttga tggtagtagg	8880
	gcagttggcc atgacggacc agttgacggt ctggtggccc ggacgcacga gctcgtggtg	8940
	cttagggcg gagtaggcgc gcgtgtcga gatgtagtcg ttgcaggtgc gcaccaggta	9000
	ctgtagccg atgaggaagt gcggcggcgg ctggcggtag agcggccatc gctcgtggtc	9060
	gggggcgccg ggcgcgaggt cctcgagcat ggtgcggtgg tagccgtaga tgtacctgga	9120
	catccagggt atgccggcgg cggtaggtga gcgcgcggg aactcgcgga cgcggttcca	9180
	gatgttgcg agcggcagga agtagttcat ggtgggcacg gcttgcccg tagggcgcgc	9240
	gcagtcgtgg atgctctata cgggcaaaaa cgaaagcgg cagcggctcg actccgtggc	9300
	ctggaggcta agcgaacggg ttgggctgcg cgtgtacccc ggttcgaatc tcgaatcagg	9360

	ctggagccgc agctaactg gtattggcac tcccgtctcg acccaagcct gcaccaacce	9420
	tccaggatac ggaggegggt cgttttgcaa ctttttttg gaggccgat gagactagta	9480
	agcgcggaaa gcggccgacc gcgatggctc gctgccgtag tctggagaag aatcgccagg	9540
	gttgcttgc ggtgtgcccc ggttcgaggc cgcccgatt ccgcgctaa cgagggcgtg	9600
	gctgccccgt cgtttccaag accccatagc cagccgactt ctccagttac ggagcgagcc	9660
	cctcttttgt tttgtttgt tttgccagat gcatcccgtc ctgcggcaga tgcgccccca	9720
	ccaccctcca ccgcaacaac agccccctcc acagccggcg cttctgcccc cgccccagca	9780
	gcaacttcca gccacgaccg ccgcgccgc cgtgagcggg gctggacaga gttatgatca	9840
	ccagctggcc ttggaagagg gcgaggggct ggcgcgcctg ggggcgtcgt cgccggagcg	9900
	gcaccgcgc gtgcagatga aaagggacgc tcgcgaggcc tacgtgcca agcagaacct	9960
	gttcagagac aggagcggcg aggagcccga ggagatgcgc gcggcccgt tccacgcggg	10020
	gcgggagctg cggcgcggcc tggaccgaaa gagggtgctg agggacgagg atttcgaggc	10080
[0120]	ggacgagctg acgggatca gccccgcgc gcgcacgtg gccgcggcca acctggtcac	10140
	ggcgtacgag cagaccgtga aggaggagag caacttcaa aaatcctca acaaccagct	10200
	gcgcaccctg atcgcgcgc aggaggtgac cctgggcctg atgcacctgt gggacctgct	10260
	ggagccatc gtgcagaacc ccaccagcaa gccgctgacg gcgcagctgt tcttgggtgt	10320
	gcagcatagt cgggacaacg aagcgttcag ggaggcgtc ctgaatatca ccgagcccga	10380
	gggccgctgg ctctggacc tggatgaacat tctgcagagc atcgtggtgc aggagcgcgg	10440
	gctgccgtg tccgagaagc tggcggccat caacttctc gtgctgagtt tgggcaagta	10500
	ctacgctagg aagatctaca agaccccgtc cgtgccata gacaaggagg tgaagatcga	10560
	cgggttttac atgcgatga ccctgaaagt gctgaccctg agcgacgac tgggggtgta	10620
	ccgcaacgac aggatgcacc gtgcggtgag cgccagcagg cggcgcgagc tgagcgacca	10680
	ggagctgatg catagtctgc agcgggccct gaccggggcc gggaccgagg gggagagcta	10740
	ctttgacatg ggcgcggacc tgcaactgca gccagccgc cgggccttgg aggcggcggc	10800
	aggaccctac gtagaagagg tggacgatga ggtggacgag gagggcgagt acctggaaga	10860

	ctgatggcgc gaccgtatth ttgctagatg caacaacaac agccacctcc tgatccccgcg	10920
	atgcgggicgg cgctgcagag ccagccgtcc ggcattaact cctcggacga ttggaccag	10980
	gccatgcaac gcatcatggc gctgacgacc cgcaaccccc aagcctttag acagcagccc	11040
	caggccaacc ggctctcggc catcctggag gccgtggcgc cctcgcgctc caacccccag	11100
	cacgagaagg tcctggccat cgtgaacgcg ctggtggaga acaaggccat ccgcggcgac	11160
	gagcccggcc tgggttaciaa cgcgctgctg gagcgcgtgg cccgctaciaa cagcacciaac	11220
	gtgcagacca acctggaccg catggtgacc gacgtcgcgc aggcctggc ccagcgcgag	11280
	cggttccacc gcgagtccaa cctgggatcc atggtggcgc tgaacgcctt cctcagcacc	11340
	cagcccccca acgtgccccg gggccaggag gactacacca acttcatcag cgcctcgcgc	11400
	ctgatggtga ccgaggtgcc ccagagcgag gtgtaccagt ccgggccgga ctacttcttc	11460
	cagaccagtc gccagggctt gcagaccgtg aacctgagcc aggccttcaa gaacttgacg	11520
	ggcctgtggg gcgtgcaggc cccggtcggg gaccgcgcga cgggtgcgag cctgctgacg	11580
	ccgaactcgc gcctgctgct gctgctggtg gccccctca cggacagcgg cagcatcaac	11640
[0121]	cgcaactcgt acctgggcta cctgattaac ctgtaccgcg aggcctcgg ccagcgcac	11700
	gtggacgagc agacctacca ggagatcacc cacgtgagcc gcgccctggg ccagcagcgc	11760
	ccgggcaacc tggaagccac cctgaacttt ttgctgacca accggtcgcga gaagatcccg	11820
	ccccagtacg cgctcagcac cgaggaggag cgcctcctgc gttacgtgca gcagagcgtg	11880
	ggcctgttcc tgatgcagga gggggccacc cccagcggc cgctcgacat gaccgcgcgc	11940
	aacatggagc ccagcatgta cgccagcaac cgcccgttca tcaataaact gatggactac	12000
	ttgcatcggg cggccgccat gaactctgac tatttaccac acgccatcct gaatccccac	12060
	tggctcccgc cgccggggtt ctacacgggc gactacgaca tgcccgacct caatgacggg	12120
	ttctgtggg acgatgtgga cagcagcgtg ttctcccccc gaccgggtgc taacgagcgc	12180
	cccttggtga agaaggaagg cagcagccga cgcccgtcct cggcgcctgc cggccgcgag	12240
	ggtgctgccg cggcgggtgcc cgaggccgcc agtcctttcc cgagcttgcc cttctcgtg	12300
	aacagtatcc gcagcagcga gctgggcagg atcacgcgcc cgcgcttgct gggcgaagag	12360
	gagtacttga atgactcgtt gttgagacct gagcgggaga agaacttccc caataacggg	12420

	atagaaagcc tgggtggacaa gatgagccgc tggaaagacgt atgcgcagga gcacagggac	12480
	gatccccggg cgctgcaggg ggccacgagc cggggcagcg ccgcccgtaa acgccggtgg	12540
	cacgacaggc agcggggaca gatgtgggac gatgaggact ccgccgacga cagcagcgtg	12600
	ttggacttgg gtgggagtgg taacccttc gctcacctgc gccccgtat cgggcgcatg	12660
	atgtaagaga aaccgaaaat aatgatact caccaaggcc atggcgacca gcgtgcgttc	12720
	gtttcttctc tgttgttgtt gtatctagta tgatgaggcg tgcgtaccg gagggtcctc	12780
	ctccctcgta cgagagcgtg atgcagcagg cgatggcggc ggccggcgtg cagccccgc	12840
	tggaggctcc ttacgtgcc ccgcggtacc tggcgcctac ggagggcgg aacagcattc	12900
	gttactcgga gctggcacc ttgtacgata ccaccggtt gtacctggtg gacaacaagt	12960
	cggcggacat cgctcgtg aactaccaga acgaccacag caacttctg accaccgtg	13020
	tgcagaacaa tgacttcacc cccacggagg ccagcacca gaccatcaac tttgacgagc	13080
	gctcgcggtg gggcggccag ctgaaaacca tcatgcacac caacatgcc aacgtgaac	13140
[0122]	agttcatgta cagcaacaag ttcaaggcgc gggatgatgt ctcccgaag acccccaatg	13200
	gggtgacagt gacagaggat tatgatgta gtcaggatga gctgaagtat gaatgggtgg	13260
	aatttgagct gccgaaggc aacttctcgg tgacctgac catcgacctg atgaacaacg	13320
	ccatcatcga caattacttg gcggtggggc ggcagaacgg ggtgctggag agcgacatcg	13380
	gcgtgaagtt cgacactagg aacttcaggc tgggctggga cccctgacc gagctggtca	13440
	tgccggggt gtacaccaac gaggettcc atcccgatat tgtcttctg cccggtcgcg	13500
	gggtggactt caccgagagc cgctcagca acctgctggg cattcgaag aggcagcct	13560
	tccaggaagg ctccagatc atgtacgagg atctggagg gggcaacatc cccgcgtcc	13620
	tggatgtcga cgctatgag aaaagcaagg aggatgcagc agctgaagca actgcagccg	13680
	tagctaccgc ctctaccgag gtcagggcg ataattttgc aagcggcga gcagtggcag	13740
	cggccgaggc ggctgaaacc gaaagtaaga tagtcattca gccggtggag aaggatagca	13800
	agaacaggag ctacaacgta ctaccgaca agataaacac cgctaccgc agctgtgacc	13860
	tagctataca ctatggcgac cccgagaagg gcgtgcgctc ctggacgctg ctaccacct	13920

	cggacgtcac ctgcgcgctg gagcaagtct actggtcgct gcccgacatg atgcaagacc	13980
	cggtcacctt ccgctccacg cgtcaagtta gcaactaccc ggtggtgggc gccgagctcc	14040
	tgcccgteta ctccaagagc ttettcaacg agcaggccgt ctactcgcag cagctgcgcg	14100
	ccttcacctc gcttacgcac gtcttcaacc gcttccccga gaaccagatc ctgctccgcc	14160
	cgcccgcgcc caccattacc accgtcagtg aaaacgttcc tgctctcaca gatcacggga	14220
	ccctgccgct ggcgagcagt atccggggag tccagcgcgt gaccgttact gacgccagac	14280
	gccgcacctg cccctacgtc tacaaggccc tgggcatagt cgcgcccgcg gtctctcga	14340
	gccgcacctt ctaaatgtcc attctcatct cgcccagtaa taacaccggt tggggcctgc	14400
	gcgcgcccag caagatgtac ggaggcgcct gccaacgctc cagcaaacac cccgtgcgcg	14460
	tgcgcgggca cttccgcgct ccctggggcg cctcaaggg ccgcgtgcgg tcgcgacca	14520
	ccgtcgacga cgtgatgcac caggtggtgg ccgacgcgcg caactacacc cccgccgccg	14580
	cgcccgtctc caccgtggac gccgtcatcg acagcgtggt ggccgacgcg cgccggtacg	14640
	cccgcgcaa gagccggcgg cggcgcacgc cccggcggca ccggagcacc cccgccatgc	14700
[0123]	gcgcggcgcg agccttgctg cgcagggcca ggcgcacggg acgcagggcc atgctcaggg	14760
	cggccagacg cgcggcttca ggccagcgcg ccggcaggac ccggagacgc gcggccacgg	14820
	cggcggcagc ggccatgcgc agcatgtccc gcccgcggcg agggaacgtg tactgggtgc	14880
	gcgacgccgc caccggtgtg cgcgtgcccg tgcgcacccg cccccctgc acttgaagat	14940
	gttcaacttc cgatgttgat gtgtcccagc ggcgaggagg atgtccaagc gcaaattcaa	15000
	ggaagagatg ctccaggtca tcgcgctga gatctacggc cctgcggtgg tgaaggagga	15060
	aagaaagccc cgcaaatca agcgggtcaa aaaggacaaa aaggaagaag aaagtgatgt	15120
	ggacggattg gtggagtttg tgcgcgagtt cgcccccg cggcgcgtgc agtggcgcg	15180
	gcggaagggt caaccggtgc tgagaccgag caccaccgtg gtcttcacgc ccggcgagcg	15240
	ctccggcacc gcttccaagc gctcctacga cgaggtgtac ggggatgatg atattctgga	15300
	gcaggcggcc gagcgcctgg gcgagtttgc ttacggcaag cgcagccgtt ccgcaccgaa	15360
	ggaagaggcg gtgtccatcc cgetggacca cggcaacccc acgccagacc tcaagcccgt	15420
	gaccttgcag caggtgctgc cgaccgcgac gccgcgcccg gggttcaagc gcgaggcga	15480

	ggatctgtac cccacatgc agctgatggt gcccaagcgc cagaagctgg aagacgtgct	15540
	ggagaccatg aaggtggacc cggacgtgca gcccgaggtc aaggtgcggc ccatcaagca	15600
	ggtggccccg ggcctgggcg tgcagaccgt ggacatcaag attcccacgg agcccatgga	15660
	aacgcagacc gagcccatga tcaagcccag caccagcacc atggaggtgc agacggatcc	15720
	ctggatgcca tcggtccta gtcgaagacc ccggcgcaag tacggcgcgg ccagcctgct	15780
	gatgccaac tacgcgctgc atccttccat catccccacg ccgggctacc gcggcacgcg	15840
	cttctaccgc ggtcatacca gcagccgccc ccgcaagacc accactcgcc gccgcgctcg	15900
	ccgcaccgcc gctgcaacca cccctgccgc cctggtgcgg agagtgtacc gccgcggccc	15960
	cgcacctctg accctgccgc gcgcgcgcta ccaccgagc atcgccattt aaactttcgc	16020
	ctgctttgca gatcaatggc cctcacatgc cgcttcgcg ttccattac gggetaccga	16080
	ggaagaaaac cgcgccgtag aaggctggcg ggaacggga tgcgtcgcca ccaccaccgg	16140
	ccggcggcgc ccatcagcaa gcggttgggg ggaggcttcc tgcccgcgct gatcccatc	16200
[0124]	atcgccgcgg cgateggggc gatccccgc attgcttccg tggcgggtgca ggcctctcag	16260
	cgccactgag acacacttgg aaacatcttg taataaacca atggactctg acgctcctgg	16320
	tctgtgatg tgttttcgta gacagatgga agacatcaat ttttcgtccc tggtccgcg	16380
	acacggcacg cggccgttca tgggcacctg gagcgacatc ggcaccagcc aactgaacgg	16440
	gggcgccttc aattggagca gtctctggag cgggcttaag aatttcgggt ccacgcttaa	16500
	aacctatggc agcaaggcgt ggaacagcac cacaggcgag gcgctgaggg ataagctgaa	16560
	agagcagaac ttccagcaga aggtggtcga tgggctcgcc tcgggcatca acggggtggt	16620
	ggacctggcc aaccaggccc tgcagcggca gatcaacage cgcttgacc cggtgccgcc	16680
	cgccggctcc gtggagatgc cgcaggtgga ggaggagctg cctcccctgg acaagcgggg	16740
	cgagaagcga ccccgccccg atgcggagga gacgctgctg acgcacacgg acgagccgcc	16800
	cccgtacgag gaggcggtga aactgggtct gccaccacg cggcccatcg cgcccctggc	16860
	caccggggtg ctgaaacctg aaaagcccgc gacctggac ttgcctctc cccagccttc	16920
	ccgcccctct acagtggcta agcccctgcc gccggtggcc gtggcccgcg cgcgaccggc	16980

	gggcaccgcc cgccctcatg cgaactggca gagcactctg aacagcatcg tgggtctggg	17040
	agtcagaggt gtgaagcgcc gccgctgcta ttaaacctac cgtagcgctt aacttgcttg	17100
	tctgtgtgtg tatgtattat gtcgccgccc ccgctgtcca ccagaaggag gagtgaagag	17160
	gcgcgctgcc gagttgcaag atggccaccc catcgatgct gccccagtgg gcgtacatgc	17220
	acatcgccgg acaggacgct tcggagtacc tgagtccggg tctgggtgcag tttgccccgc	17280
	ccacagacac ctacttcagt ctggggaaca agtttaggaa cccacgggtg gcgccccgc	17340
	acgatgtgac caccgaccgc agccagcggc tgacgctgct cttcgtgccc gtggaccgcg	17400
	aggacaacac ctactcgtac aaagtgcgct acacgctggc cgtgggcgac aaccgcgtgc	17460
	tggacatggc cagcacctac tttgacatcc gcggcgtgct ggatcggggc cctagcttca	17520
	aaccctactc cgccaccgcc tacaacagtc tggccccaa gggagcacc aacacttgtc	17580
	agtggacata taaagccgat ggtgaaactg ccacagaaaa aacctataca tatggaaatg	17640
	caccgctgca gggcattaac atcacaaaag atggatttca acttggaact gacaccgatg	17700
	atcagccaat ctacgcagat aaaacctatc agcctgaacc tcaagtgggt gatgctgaat	17760
[0125]	ggcatgacat cactggctact gatgaaaagt atggaggcag agctcttaag cctgatacca	17820
	aatgaagcc ttgttatggt tcttttgcca agcctactaa taaagaagga ggtcaggcaa	17880
	atgtgaaaac aggaacaggc actactaaag aatatgacat agacatggct ttctttgaca	17940
	acagaagtgc ggctgctgct ggcctagctc cagaaattgt tttgtatact gaaaatgtgg	18000
	atttgaaaac tccagatacc catattgtat acaaagcagg cacagatgac agcagctctt	18060
	ctattaattt gggtcagcaa gccatgccca acagacctaa ctacattggt ttcagagaca	18120
	actttatcgg gctcatgtac tacaacagca ctggcaatat gggggtgctg gccggtcagg	18180
	cttctcagct gaatgctgtg gttgacttgc aagacagaaa caccgagctg tcctaccage	18240
	tcttgcttga ctctctgggt gacagaaccc ggtatttcag tatgtggaat caggcgggtg	18300
	acagctatga tcctgatgtg cgcattattg aaaatcatgg tgtggaggat gaacttccca	18360
	actattgttt cctctggat gctgttggca gaacagatac ttatcaggga attaaggcta	18420
	atggaactga tcaaaccaca tggaccaaag atgacagtgt caatgatgct aatgagatag	18480
	gcaagggtaa tccattcgcc atggaatca acatccaagc caacctgtgg aggaacttcc	18540

	tctacgcaa cgtggccctg tacctgcccg actcttaca gtacacgccg gccaatgtta	18600
	ccctgcccac caacaccaac acctacgatt acatgaacgg ccgggtggtg gcgccctcgc	18660
	tggtggactc ctacatcaac atcggggcgc gctggtcgtt ggatcccatg gacaacgtga	18720
	acccttcaa ccaccaccgc aatgcggggc tgcgctaccg ctccatgctc ctgggcaacg	18780
	ggcgtacgt gcccttccac atccaggtgc cccagaaatt tttcgccatc aagagcctcc	18840
	tgctcctgcc cgggtcctac acctacgagt ggaacttccg caaggacgtc aacatgatcc	18900
	tgcagagctc cctcggcaac gacctgcga cggacggggc ctccatctcc ttcaccagca	18960
	tcaaccteta cgccacctc ttccccatgg cgcacaacac ggctccacg ctcgaggcca	19020
	tgctgcgcaa cgacaccaac gaccagtctt tcaacgacta cctctcggcg gccaacatgc	19080
	tctaccccat ccggccaac gccaccaac tgcccatctc catcccctcg cgcaactggg	19140
	ccgccttcg cggttggtcc ttcacgcgtc tcaagaccaa ggagacgcc tcgctgggct	19200
	ccgggttcga cccctacttc gtctactcgg gctccatccc ctacctcgac ggcaccttct	19260
[0126]	acctcaacca caccttcaag aaggtctcca tcaccttcca ctctccgctc agctggcccc	19320
	gcaacgaccg gctcctgacg cccaacgagt tcgaaatcaa gcgcaccgtc gacggcgagg	19380
	gctacaacgt ggcccagtgc aacatgacca aggactggtt cctggtccag atgctggccc	19440
	actacaacat cggtaccag ggcttctacg tgcccagagg ctacaaggac cgcatgtact	19500
	ccttcttcg caacttcag cccatgagcc gccaggtggt ggacgaggtc aactacaagg	19560
	actaccaggc cgtcacctg gctaccagc acaacaactc gggcttcgct ggctacctg	19620
	cgcccacat gcgccaggc cagccctacc ccgccaacta cccctaccg ctcatcgga	19680
	agagcgccgt caccagcgtc acccagaaaa agttcctctg cgacagggtc atgtggcgca	19740
	tccccttctc cagcaacttc atgtccatgg gcgcgctcac cgacctcggc cagaacatgc	19800
	tctatgcaa ctccgccac gcgctagaca tgaatttcca agtcgacccc atggatgagt	19860
	ccacccttct ctatgttgc ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc	19920
	gcggcgtcat cgaggccgct tacctgcgca ccccttctc ggccggtaac gccaccact	19980
	aagctcttgc ttcttgcaag ccatggccgc gggctccggc gagcaggagc tcagggccat	20040

	catccgcgac ctgggctgcg ggccctactt cctgggcacc ttcgataage gcttccccggg	20100
	attcatggcc cgcacaagc tggcctgcbc catcgtcaac acggccggcc gcgagaccgg	20160
	gggcgagcac tggctggcct tcgcctggaa cccgcgctcg aacacctgct acctcttcga	20220
	ccccttcggg ttctcggacg agcgcctcaa gcagatctac cagtctgagt acgagggcct	20280
	gctgcgccgc agcgccttg ccaccgagga ccgctgcgtc accctggaaa agtccacca	20340
	gaccgtgcag ggtccgcgct cggccgcctg cgggctcttc tgetgcatgt tcctgcacgc	20400
	cttcgtgcac tggccccgacc gccccatgga caagaacccc accatgaact tgetgacggg	20460
	ggtgcccaac ggcatgctcc agtcgcccc ggtggaaccc accctgcgcc gcaaccagga	20520
	ggcgtctac cgcttctca actcccactc cgcctacttt cgtcccacc gcgcgcgat	20580
	cgagaaggcc accgccttcg accgcatgaa tcaagacatg taaaccgtgt gtgtatgta	20640
	aatgtcttta ataaacagca ctttcatggt acacatgcat ctgagatgat ttatttagaa	20700
	atcгааaggg ttctgccggg tctcggcatg gcccgcgggc agggacacgt tgcggaactg	20760
[0127]	gtacttggcc agccacttga actcggggat cagcagttg ggcagcgggg tgcggggaa	20820
	ggagtgggtc cacagcttcc gcgtcagttg caggcgcccc agcaggtcgg gcgcgggat	20880
	cttгааatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca cggggttga	20940
	gcaactggaac accatcaggg ccgggtgctt cacgctcgcc agcacctcg cgtcgggtgat	21000
	gctctccacg tcgaggtcct cggcgttggc catcccgaag ggggtcatct tgcaggtctg	21060
	ccttcccatg gtgggcacgc acccgggctt gtggttgcga tcgagtgca ggggatcag	21120
	catcatctgg gcctggctcg cgttcatccc cgggtacatg gccttcatga aagcctcaa	21180
	ttgcctgaac gcctgctggg ctttggctcc ctcggtgaag aagaccccgc aggacttgct	21240
	agagaactgg ttggtggcgc acccggcgtc gtgcacgcag cagcgcgct cgttgttggc	21300
	cagetgcacc acgtgcgcc ccacgcggtt ctgggtgate ttggcccggg cggggttctc	21360
	cttcagcgcg cgctgcccgt tctcgtcgc cacatccatc tcgatcatgt gctccttctg	21420
	gatcatggtg gtcccgtgca ggcaccgcag cttgccctcg gcctcgggtc acccgtgcag	21480
	ccacagcgcg caccgggtgc actcccagtt cttgtgggcg atctgggaat gcgcgtgcac	21540
	gaagccctgc aggaagcggc ccatcatggt ggtcagggtc ttgttgctag tgaaggtcag	21600

	cggaatgccg cgggtgctcct cgttgatgta caggtggcag atgcggcggg acacctcgcc	21660
	ctgctcgggc atcagctgga agttggcttt caggteggtc tccacgcggg agcgggccat	21720
	cagcatagtc atgatttcca tacccttctc ccaggccgag acgatgggca ggctcatagg	21780
	gttcttcacc atcatcttag cgctagcage cgcggccagg gggtcgctct cgtccagggt	21840
	ctcaaagctc cgcttgccgt ccttctcggg gatccgcacc ggggggtagc tgaagcccac	21900
	ggccgccagc tcctctcggg cctgtcttcc gtctctcgtg tcctggctga cgtctgcag	21960
	gaccacatgc ttggtcttgc ggggtttctt cttgggcggc agcggcggcg gagatgttgg	22020
	agatggcgag ggggagcgcg agttctcgtc caccactact atctcttctt cttcttggtc	22080
	cgaggccacg cggcggtagg tatgtctctt cgggggcaga ggccggaggcg acgggctctc	22140
	gccgccgca cttggcggat ggctggcaga gccccttccg cgttcggggg tgcgctccc	22200
	gcggcgctct gactgacttc ctccgcggcc ggccattgtg ttctcttagg gaggaacaac	22260
	aagcatggag actcagccat cgccaacctc gccatctgcc cccaccgccg acgagaagca	22320
[0128]	gcagcagcag aatgaaagct taaccgcccc gccgccagc cccgccacct ccgacgcggc	22380
	cgtcccagac atgcaagaga tggaggaatc catcgagatt gacctgggct atgtgacgcc	22440
	cgcgagcac gaggaggagc tggcagtgcg cttttcaca gaagagatac accaagaaca	22500
	gccagagcag gaagcagaga atgagcagag tcaggctggg ctcgagcatg acggcgacta	22560
	cctccacctg agcggggggg aggacgcgt catcaagcat ctggcccggc aggccacct	22620
	cgtcaaggat gcgctgctcg accgcaccga ggtgccctc agcgtggagg agctcagccg	22680
	cgctacgag ttgaacctct tctcgcgcg cgtgcccc aagcggcagc ccaatggcac	22740
	ctgcgagccc aacctgcgcc tcaactteta cccggtcttc gcggtgcccg aggccctggc	22800
	cacctaccac atctttttca agaaccaaaa gatccccgtc tcctgccgcg ccaaccgcac	22860
	ccgcgccgac gcccttttca acctgggtcc cggcgcccgc ctacctgata tcgctctctt	22920
	ggaagagggt cccaagatct tcgagggtct gggcagcgac gagactcggg ccgcgaacgc	22980
	tctgcaagga gaaggaggag agcatgagca ccacagcgcc ctggctcagt tggaaggcga	23040
	caacgcgcgg ctggcgggtc tcaaacgcac ggtcgagctg acccatttct cctaccggc	23100

	tctgaacctg ccccccaag tcatgagcgc ggtcatggac caggtgctca tcaagcgcgc	23160
	gtcgcctatc tccgaggacg agggcatgca agactccgag gagggaagc cctgtgtcag	23220
	cgacgagcag ctggcccgtt ggctgggtcc taatgctagt ccccagagtt tggaagagcg	23280
	gcgcaaacctc atgatggcgg tggctctggt gaccgtggag ctggagtgcc tgcgccgctt	23340
	cttcgccgac gcggagacct tgcgcaaggt cgaggagaac ctgcactacc tcttcaggca	23400
	cggttctgtg cgccaggcct gcaagatctc caacgtggag ctgaccaacc tggctctcta	23460
	catgggcctc ttgcaagaga accgctggg gcagaacgtg ctgcacacca ccttgcgcgg	23520
	ggaggccccg cgcgactaca tccgcgactg cgtctacctc tacctctgcc acacctggca	23580
	gacgggcatg ggcgtgtggc agcagtgtct ggaggagcag aacctgaaag agctctgcaa	23640
	gctcctgcag aagaacctca agggctctgtg gaccgggttc gacgagcgc caaccgctc	23700
	ggacctggcc gacctcattt tccccgagcg cctcaggctg acgctgcgca acggcctgcc	23760
	cgactttatg agccaaagca tgttgcaaaa ctttcgctct ttcatectcg aacgctccgg	23820
	aatcctgccc gccacctgct ccgctgtgcc ctcggacttc gtgccgctga ccttccgca	23880
[0129]	gtgcccccg ccgctgtgga gccactgcta cctgctgcgc ctggccaact acctggccta	23940
	ccactcggac gtgatcgagg acgtcagcgg cgagggcctg ctcgagtgcc actgccgctg	24000
	caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga gcgagacca	24060
	gatcatcggc accttcgagt tgcaagggcc cagcgaaggc gagggttcag ccgccaagg	24120
	gggtctgaaa ctcaccccgg ggctgtggac ctcggcctac ttgcgcaagt tcgtgcccga	24180
	ggactacat cccttcgaga tcaggttcta cgaggaccaa tcccatccgc ccaaggccga	24240
	gctgtcggcc tgcgtcatca cccagggggc gatcctggcc caattgcaag ccatccagaa	24300
	atcccgccaa gaattcttgc tgaaaaaggg ccgcggggtc tacctcgacc cccagaccgg	24360
	tgaggagctc aaccccggct tccccagga tgccccgagg aaacaagaag ctgaaagtgg	24420
	agctgccgcc cgtggaggat ttggaggaag actgggagaa cagcagtcag gcagaggagg	24480
	aggagatgga ggaagactgg gacagcactc aggcagagga ggacagcctg caagacagtc	24540
	tggaggaaga cgaggaggag gcagaggagg aggtggaaga agcagccgcc gccagaccgt	24600
	cgctcctggc gggggagaaa gcaagcagca cggataccat ctccgctccg ggtcggggtc	24660

	ccgctcgacc acacagtaga tgggacgaga ccggacgatt cccgaacccc accaccaga	24720
	ccgtaagaa ggagcggcag ggatacaagt cctggcgggg gcacaaaaac gccatcgtct	24780
	cctgcttgca ggccctgcggg ggcaacatct ccttcacccg gcgctacctg ctcttcacc	24840
	gcggggtgaa ctttccccgc aacatcttgc attactaccg tcacctccac agcccctact	24900
	acttccaaga agaggcagca gcagcagaaa aagaccagca gaaaaccagc agctagaaaa	24960
	tccacagcgg cggcagcagg tggactgagg atcgcggcga acgagccggc gcaaaccggg	25020
	gagctgagga accggatctt tcccaccctc tatgccatct tccagcagag tcgggggcag	25080
	gagcaggaac tgaagtcaa gaaccgttct ctgcgctcgc tcaccgcag ttgtctgtat	25140
	cacaagagcg aagaccaact tcagcgcact ctcgaggacg ccgaggtctt cttcaacaag	25200
	tactgcgcgc tcaactctta agagtagccc gcgcccggcc agtcgcagaa aaaggcggga	25260
	attacgtcac ctgtgccctt cgccttagcc gcctccacc atcatcatga gcaaagagat	25320
	tcccacgcct tacatgtgga gctaccagcc ccagatgggc ctggccggcg gtgccgcca	25380
[0130]	ggactactcc acccgcata attggctcag cgccgggccc gcgatgatct cacgggtgaa	25440
	tgacatccgc gccaccgaa accagatact cctagaacag tcagcgtca ccgccacgcc	25500
	ccgcaatcac ctcaatccgc gtaattggcc cgccgccctg gtgtaccagg aaattcccca	25560
	gccacgacc gtactacttc cgcgagacgc ccaggccgaa gtccagctga ctaactcagg	25620
	tgtccagctg gcgggcgcg ccaccctgtg tcgtcaccgc cccgctcagg gtataaagcg	25680
	gctggtgatc cggggcagag gcacacagct caacgacgag gtggtgagct cttegtggg	25740
	tctgcacct gacggagtct tccaactcgc cggatcgggg agatcttct tcacgcctcg	25800
	tcaggccgtc ctgactttgg agagtctgtc ctgcagccc cgctcgggtg geatcggcac	25860
	tctccagttc gtggaggagt tcactccctc ggtctacttc aacccttct ccgctcccc	25920
	cgccactac ccggacgagt tcatccgaa ctctgacgcc atcagcgagt cggtgacgg	25980
	ctacgattga atgtccatg gtggcgcagc tgacctagct cggcttcgac acctggacca	26040
	ctgccgccgc ttccgtgct tegetcggga tctcggcag tttgcctact ttgagctgcc	26100
	cgaggagcac cctcagggcc cggcccacgg agtgcggatc gtcgtcgaag gggcctcga	26160

	ctcccacctg cttcggatct tcagccagcg tccgatcctg gtcgagcgcg agcaaggaca	26220
	gacccttctg actctgtact gcatctgcaa ccaccccggc ctgcatgaaa gtctttgttg	26280
	tctgctgtgt actgagtata ataaaagctg agatcagcga ctactccgga cttccgtgtg	26340
	ttcctgaate catcaaccag tctttgttct tcaccgggaa cgagaccgag ctccagctcc	26400
	agtgtaaacc ccacaagaag tacctcaact ggctgttcca gggctccccg atcgccgttg	26460
	tcaaccactg cgacaacgac ggagtcctgc tgagcggccc tgccaacctt actttttcca	26520
	cccgcagaag caagetccag ctcttccaac ctttctccc cgggacctat cagtgcgtct	26580
	cgggacctg ccatcacacc ttccacctga tcccgaatac cacagcgtcg ctccccgcta	26640
	ctaacaacca aactaacctc caccaacgcc accgtcgcga cggccacaat acatgccat	26700
	attagactat gaggccgagc cacagcagcc catgctcccc gctattagtt acttcaatct	26760
	aaccggcgga gatgactgac ccaactggcca acaacaacgt caacgacctt ctctggaca	26820
	tggacggccg cgcctcggag cagcagctcg cccaacttcg cattcgcag cagcaggaga	26880
	gagccgtcaa ggagctgcag gatgcggtgg ccatccacca gtgcaagaga ggcatttct	26940
[0131]	gcctggtgaa acagccaag atctcctacg aggtcactcc aaacgacct cgcctctcct	27000
	acgagctcct gcagcagcgc cagaagtcca cctgcctggt cggagtcaac cccatcgtca	27060
	tcaccagca gtctggcgat accaaggggt gcatccactg ctctcgcgac tccccgact	27120
	gcgtccacac tctgatcaag accctctgcg gcctcgcga cctcctccc atgaactaat	27180
	caccccctta tccagtgaaa taaagatcat attgatgatg attttacaga aataaaaaat	27240
	aatcatttga tttgaataa agatacaatc atattgatga tttgagtta acaaaaaaat	27300
	aaagaatcac ttacttgaat tctgatacca ggtctctgtc catgttttct gccaacacca	27360
	cttcactccc ctcttcccag ctctggtact gcaggccccg gcgggctgca aacttctcc	27420
	acacgtgaa ggggatgta aattctcct gtcctcaat ctccatttta tctctatca	27480
	gatgtccaaa aagcgcgtcc ggggtgatga tgacttcgac cccgtctacc cctacgatgc	27540
	agacaacgca ccgaccgtgc cttcatcaa ccccccttc gtctcttcag atggattcca	27600
	agagaagccc ctgggggtgt tgtccctgcg actggccgac cccgtcacca ccaagaacgg	27660
	ggaaatcacc ctcaagctgg gagaggggt ggacctgat tcctcgggaa aactcatctc	27720

	caacacggcc accaaggccg cgcacctct cagtttttcc aacaacacca ttcccttaa	27780
	catggatcac cccttttaca ctaaagatgg aaaattatcc ttacaagttt ctccaccatt	27840
	aaatatactg agaacaagca ttctaaacac actagcttta ggTTTTGGAT caggTTtagg	27900
	actccgtggc tctgccttgg cagtacagtt agtctctcca cttacatttg atactgatgg	27960
	aaacataaag cttaccttag acagaggttt gcatgttaca acaggagatg caattgaaag	28020
	caacataagc tgggctaaag gtttaaaatt tgaagatgga gccatagcaa ccaacattgg	28080
	aatgggtta gagtttggaa gcagtagtac agaaacaggt gttgatgatg cttaccaat	28140
	ccaagttaa cttggatctg gccttagctt tgacagtaca ggagccataa tggctggtaa	28200
	caaagaagac gataaactca ctttgggac aacacctgat ccatcaccaa actgtcaaat	28260
	actgcagaa aatgatgcaa aactaacact ttgcttgact aaatgtggtg gtcaaaact	28320
	ggccactgtg tcagtcttag ttgtaggaag tggaaaccta aaccaccata ctggcaccgt	28380
	aagcagtgct caggTgttTc tacgttttga tgcaaacggT gttcttttaa cagaacattc	28440
[0132]	tactactaaa aaatactggg ggtataggca gggagatagc atagatggca ctccatatac	28500
	caatgctgta ggattcatgc ccaatttaa agcttatcca aagtcacaaa gttctactac	28560
	taaaaataat atagtagggc aagtatacat gaatggagat gtttcaaac ctatgcttct	28620
	cactataacc ctcaatggta ctgatgacag caacagtaca tattcaatgt cattttcata	28680
	cacctggact aatggaagct atgttggagc aacatttggg gctaactctt ataccttctc	28740
	atacatgcc caagaatgaa cactgtatcc caccctgcat gccaacctt cccacccac	28800
	tctgtggaac aaactctgaa acacaaaata aaataaagtt caagtgtttt attgattcaa	28860
	cagttttaca ggattcgagc agttatTTTT cctccacct cccaggacat ggaatacacc	28920
	acctctccc cccgcacagc cttgaacatc tgaatgcat tggatgatgga catgcttttg	28980
	gtctccacgt tccacacagt ttcagagcga gccagtctcg ggtcggtcag ggagatgaaa	29040
	ccctccgggc actcccgcct ctgcacctca cagctcaaca gctgaggatt gtctcgggtg	29100
	gtcgggatca cggttatctg gaagaagcag aagagcggcg gtgggaatca tagtccgcga	29160
	acgggatcgg ccggtggtgt cgcctcagc cccgcagcag tcgctgccgc cgcctcgcg	29220

	tcaagctgct gctcaggggg tccgggtcca gggactccct cagcatgatg cccacggccc	29280
	tcagcatcag tcgtctggtg cggcgggcgc agcagcgcag gcggatctcg ctcaggtcgc	29340
	tgcagtagct gcaacacaga accaccaggt tgttcaacag tccatagtcc aacacgctcc	29400
	agccgaaact catcgcggga aggatgctac ccacgtggcc gtcgtaccag atcctcaggt	29460
	aatcaagtg gtgccccctc cagaacacgc tgcccacgta catgatctcc ttgggcatgt	29520
	ggcggttcac cacctcccgg taccacatca ccctctggtt gaacatgcag ccccggatga	29580
	tctcgggaa ccacagggcc agcaccgcc cgcccgccat gcagcgaaga gaccccgggt	29640
	cccggcaatg gcaatggagg acccaccgct cgtaccctg gatcatctgg gagctgaaca	29700
	agtctatggt ggcacagcac aggcataatgc tcatgcatct cttcagcact ctcaactcct	29760
	cgggggtcaa aaccatatcc cagggcacgg ggaactcttg caggacagcg aaccccgcag	29820
	aacagggcaa tcctcgcaca gaacttacat tgtgcatgga cagggtatcg caatcaggca	29880
	gcaccgggtg atcctccacc agagaagcgc gggctctcgtt ctctcacag cgtggtaagg	29940
	gggccggccg atacgggtga tggcgggacg cggctgatcg tttcgcgac cgtgtcatga	30000
[0133]	tgcagttgct ttcggacatt ttcgtacttg ctgtagcaga acctggtccg ggcgctgcac	30060
	accgatgcc ggcggcggtc tcggcgcttg gaacgctcgg tgttgaaatt gtaaacagc	30120
	cactctctca gaccgtgcag cagatctagg gcctcaggag tgatgaagat cccatcatgc	30180
	ctgatggctc tgatcacatc gaccaccgtg gaatgggcca gaccagcca gatgatgcaa	30240
	ttttgttggg tttcggtgac ggcgggggag ggaagaacag gaagaacat gattaacttt	30300
	taatccaaac ggtctcggag tacttcaaaa tgaagatcgc ggagatggca cctctcggcc	30360
	ccgctgtggt ggtggaaaat aacagccagg tcaaaggta tacggttctc gagatgttcc	30420
	acggtggctt ccagcaaagc ctccacgcgc acatccagaa acaagacaat agcgaagcg	30480
	ggagggttct ctaattctc aatcatcatg ttacactcct gcaccatccc cagataattt	30540
	tcatttttcc agccttgaat gattcgaact agttcgtgag gtaaatcaa gccagccatg	30600
	ataaagagct cgcgcagagc gccctccacc ggcattctta agcacaccct cataattcca	30660
	agatattctg ctcttggttc acctgcagca gattgacaag cggaatatca aaatctctgc	30720
	cgcgatccct gagctcctcc ctacagcaata actgtaagta ctctttcata tcctctccga	30780

	aatttttagc cataggacca ccaggaataa gattagggca agccacagta cagataaacc	30840
	gaagtcctcc ccagtgagea ttgccaaatg caagactgct ataagcatgc tggctagacc	30900
	cggtgatatc ttccagataa ctggacagaa aatcgcccag gcaattttta agaaaatcaa	30960
	caaaagaaaa atcctccagg tggacgttta gagcctcggg aacaacgatg aagtaaattgc	31020
	aagcggtagc ttccagcatg gttagtttagc tgatctgtag aaaaaacaaa aatgaacatt	31080
	aaaccatgct agcctggcga acaggtgggt aaatcgttct ctccagcacc aggcaggcca	31140
	cggggtctcc ggcgcgacct tcgtaaaaat tgcgctatg attgaaaacc atcacagaga	31200
	gacgttcccg gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg	31260
	cgtccgcgag tgaaaaaaag cgcccagga agcaataagg cactacaatg ctcaagtctca	31320
	agtccagcaa agcgatgcca tgcggatgaa gcacaaaatt ctcaagtgcg tacaaaatgt	31380
	aattactccc ctctgcaca ggcagcaaag cccccgatcc ctccaggtac acatacaaag	31440
	cctcagcgtc catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg	31500
[0134]	ctgagcteta acctgtccac ccgctctctg ctcaatatat agcccagatc tacactgacg	31560
	taaaggccaa agtctaaaaa taccgccaa ataatcacac acgccagca cacgccaga	31620
	aaccggtgac aactcaaaa aaatacgcgc acttctcaa acgcccacaa ctgccgtcat	31680
	ttccgggttc ccacgetacg tcataaaaac acgactttca aattccgtcg accgttaaaa	31740
	acgtcaccgg ccccgcccct aacggtcgcc cgtctctcag ccaatcagcg ccccgcatcc	31800
	ccaaattcaa acacctcatt tgcataataa cgcgcacaaa aagtttgagg tatattattg	31860
	atgatgg	31867
	<210> 12	
	<211> 32788	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 12	
	ccatcttcaa taatatacct caaacttttt gtgcgcgtta atatgcaaat gaggcgtttg	60

aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgtagg ggcggggcga	120
gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag	180
tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttcccgc gctctctgac	240
aggaaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact	300
gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg cagggaggag tatttgccga	360
gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa	420
tttccgcgta cgggtgtcaaa gtccggtgtt tttacgtagg tgtcagctga tcgccagggt	480
atntaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcgag aagagttttc	540
tcctccgcgc cgcgagtcag atctacactt tgaaagtagg gataacaggg taatgacatt	600
gattattgac tagttgttaa tagtaatcaa ttacggggtc attagttcat agcccatata	660
tggagtccg cgttacataa cttacggtaa atggcccgc tggtgaccg cccaacgacc	720
cccgccatt gacgtcaata atgacgtatg tteccatagt aacgccaata gggactttcc	780
attgacgtca atgggtggag tatttacggt aaactgccca ctggcagta catcaagtgt	840
[0135] atcatatgcc aagtcgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	900
atgccagta catgacctta cgggactttc ctacttggca gtacatctac gtattagtca	960
tcgctattac catggtgatg cggttttggc agtacaccaa tgggcgtgga tagcggtttg	1020
actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc	1080
aaaatcaacg ggactttcca aaatgtcgta ataaccgcc cccgttgacg caaatgggcg	1140
gtaggcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg	1200
cctggaacgc catccacgct gttttgacct ccatagaaga cagcagatcg gccaccatgg	1260
ccgggatgtt ccaggcactg tccgaaggct gcacacccta tgatattaac cagatgctga	1320
atgtcctggg agaccaccag gtctctggcc tggagcagct ggagagcatc atcaacttcg	1380
agaagctgac cgagtgaca agctccaatg tgatgcctat cctgtcccca ctgaccaagg	1440
gcatcctggg cttcgtgttt accctgacag tgccttctga gcggggcctg tcttgcata	1500
gcgaggcaga cgcaaccaca ccagagtccg ccaatctggg cgaggagatc ctgtctcagc	1560
tgtacctgtg gccccgggtg acatatcact ccccttctta cgcctatcac cagttcgagc	1620

	ggagagccaa gtacaagaga cacttcccag gctttggcca gtctctgctg ttcggctacc	1680
	ccgtgtacgt gttcggcgat tgcgtgcagg gcgactggga tgccatccgg tttagatact	1740
	gcgcaccacc tggatatgca ctgctgaggt gtaacgacac caattattcc gccttgctgg	1800
	cagtgggcgc cctggagggc cctcgcaate aggattggct gggcgtgcca aggcagctgg	1860
	tgacacgcat gcaggccatc cagaacgcag gcctgtgcac cctggtggca atgctggagg	1920
	agacaatctt ctggctgcag gcctttctga tggcctgac cgacagcggc cccaagacaa	1980
	acatcatcgt ggattcccag tacgtgatgg gcatctccaa gccttcttcc caggagtttg	2040
	tggactggga gaacgtgagc ccagagctga attccaccga tcagccattc tggcaggcag	2100
	gaatcctggc aaggaacctg gtgcctatgg tggccacagt gcagggccag aatctgaagt	2160
	accagggcca gagcctggtc atcagcgcct ccatcatcgt gtttaacctg ctggagctgg	2220
	aggcgcacta tcgggacgat ggcaacgtgt gggcgcacac cccactgagc cccagaacac	2280
	tgaacgcctg ggtgaaggcc gtggaggaga agaaggcat cccagtgcac ctggagctgg	2340
[0136]	cctccatgac caatatggag ctgatgtcta gcacgtgca ccagcagtg aggacatacg	2400
	gaccctgttt catgtgcctg ggaggcctgc tgaccatggt ggaggagcc gtgtggctga	2460
	cagtgcgggt gctggagctg ttcagagccg cccagctggc caacgatgtg gtgctgcaga	2520
	tcattggagct gtgcggagca gcctttcgcc aggtgtgcca caccacagtg ccatggccca	2580
	atgcctcctt gacccccaaag tggaacaatg agacaacaca gcctcagatc gccaaactgta	2640
	gcgtgtacga cttcttcgtg ttgctgcaact actatagcgt gagggatacc ctgtggcccc	2700
	gcgtgacata ccacatgaat aagtagcct atcacatgct ggaggagcgc gccaaagtata	2760
	agagaggccc tggcccaggc gcaaagtttg tggcagcatg gacctgaag gccgcccgcg	2820
	gccccggccc cggccagtat atcaaggcta acagtaagtt cattggaatc acagagctgg	2880
	gaccggacc tggataatga gtttaaactc ccatttaaat gtgagggtta atgcttcgag	2940
	cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa	3000
	aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca	3060
	ataaacaagt taacaacaac aattgcattc attttatggt tcaggttcag ggggagatgt	3120

	gggaggtttt ttaaagcaag taaaacctct acaaattggtg taaaataact ataacggctc	3180
	taaggtagcg agtgagtagt gttctggggc gggggaggac ctgcatgagg gccagaataa	3240
	ctgaaatctg tgcttttctg tgtgttcag cagcatgagc ggaagcggct cctttgagg	3300
	agggtattc agcccttacc tgacggggcg tctcccctcc tggcggggag tgcgtcagaa	3360
	tgtgatggga tccacggtgg acggccggcc cgtgcagccc gcgaactctt caaccctgac	3420
	ctatgcaacc ctgagctctt cgtcgttga cgcagctgcc gccgcagctg ctgcatctgc	3480
	cgccagcgcc gtgcgcggaa tggccatggg cgccggetac tacggcactc tgggtggccaa	3540
	ctcgagtcc accaataatc ccgccagcct gaacgaggag aagctgttgc tgctgatggc	3600
	ccagctcgag gccttgacc agcgcctggg cgagctgacc cagcaggtgg ctgagctgca	3660
	ggagcagacg cgggccggg ttgccacggt gaaatccaaa taaaaatga atcaataat	3720
	aaacggagac ggttgtgat tttacacag agtctgaatc tttattgat ttttcgcgcg	3780
	cggtaggccc tggaccaccg gtctcagca ttgagacccc ggtggatctt ttccaggacc	3840
	cggtagaggt gggcttggat gttgaggtac atgggcatga gcccgctccg ggggtggagg	3900
[0137]	tagctccatt gcagggcctc gtgctcgggg gtggtgttgt aaatcaccca gtcatagcag	3960
	gggcgcaggg catggtgttg cacaatatct ttgaggagga gactgatggc cacgggcagc	4020
	cctttggtgt aggtgtttac aaatctgttg agctgggagg gatgcatgcg gggggagatg	4080
	aggtgcatct tggcctggat cttgagattg gcgatgttac cgcccagatc ccgctgggg	4140
	ttcatgttgt gcaggaccac cagcacggtg tatccggtgc acttggggaa tttatcatgc	4200
	aacttgaag ggaaggcgtg aaagaatttg gcgacgcctt tgtgcccc caggttttcc	4260
	atgcaactcat ccatgatgat ggcgatgggc ccgtgggcgg cggcctgggc aaagacgttt	4320
	cgggggtcgg acacatcata gttgtgttcc tgggtgaggt catcataggc cattttaatg	4380
	aatttggggc ggaggggtgc ggactggggg acaaaggtac cctcgatccc gggggcgtag	4440
	ttcccctcac agatctgcat ctcccaggct ttgagctcgg agggggggat catgtccacc	4500
	tgcggggcga taaagaacac ggtttccggg gcgggggaga tgagctgggc cgaagcaag	4560
	ttccggagca gctgggactt gccgcagccg gtggggccgt agatgacccc gatgaccggc	4620
	tgcaggtggt agttgaggga gagacagctg ccgtcctccc ggaggagggg ggccacctcg	4680

	ttcatcatct cgcgcacgtg catgttctcg cgcaccagtt ccgccaggag gcgctctccc	4740
	cccaggata ggagctcctg gagcgaggcg aagtttttca gcggcttgag tccgtcggcc	4800
	atgggcattt tggagagggt ttgttcaag agttccaggc ggtcccagag ctcggtgatg	4860
	tgctctacgg catctcgatc cagcagacct cctcgtttcg cgggttggga cggtcggg	4920
	agtagggcac cagacgatgg gcgtccagcg cagccagggt ccggtccttc cagggtcgca	4980
	gcgtccgcgt cagggtggtc tccgtcacgg tgaaggggtg cgcgccgggc tgggcgcttg	5040
	cgagggtgcg cttcaggctc atccggctgg tcgaaaaccg ctcccgatcg gcgccctgcg	5100
	cgtcggccag gtagcaattg accatgagtt cgtagttgag cgcctcggcc gcgtggcctt	5160
	tggcgcggag cttacctttg gaagtctgcc cgcaggcggg acagaggagg gacttgaggg	5220
	cgtagagctt gggggcgagg aagacggact cggggcgta ggcgtccgcg ccgcagtggg	5280
	cgcagacggt ctgcactcc acgagccagg tgaggtcggg ctggtcgggg tcaaaaacca	5340
	gtttcccgcc gttctttttg atgcgtttct tacctttggt ctccatgagc tcgtgtcccc	5400
[0138]	gctgggtgac aaagaggctg tccgtgtccc cgtagaccga ctttatgggc cggtcctcga	5460
	gcggtgtgcc gcggtcctcc tcgtagagga accccgccca ctccgagacg aaagcccggg	5520
	tccaggccag cacgaaggag gccacgtggg acgggtagcg gtcgttgtcc accagcgggt	5580
	ccaccttttc cagggtatgc aaacacatgt cccctcgtc cacatccagg aaggtgattg	5640
	gcttgaagt gtaggccacg tgaccggggg tcccggccgg ggggtataa aagggtgcgg	5700
	gtcctgctc gtcctcactg tcttccggat cgtgtccag gagcgcacg tgttgggta	5760
	ggtattccct ctcaaggcg ggcatgacct cggcactcag gttgtcagtt tctagaaacg	5820
	aggaggattt gatattgacg gtgccggcgg agatgccttt caagagcccc tcgtccatct	5880
	ggtcagaaaa gacgatcttt ttgttgcga gcttgggtgc gaaggagccg tagaggcgt	5940
	tggagaggag cttggcgatg gagcgcacatg tctggttttt ttcttctgct gcgcctcct	6000
	tggcggcgat gttgagctgc acgtactcgc gcgccacgca cttccattcg gggaagacgg	6060
	tggtcagctc gtcgggcacg attctgacct gccagccccg attatgcagg gtgatgaggt	6120
	ccacactggt ggccacctcg ccgcgcaggg gctcattagt ccagcagagg cgtccgcct	6180

	tgcgcgagca gaaggggggc aggggggtcca gcatgacctc gtcggggggg tcggcatcga	6240
	tggtgaagat gccgggcagg aggtcggggt caaagtagct gatggaagt gccagatcgt	6300
	ccagggcagc ttgccattcg cgcacggcca gcgcgtctc gtagggactg aggggcgtgc	6360
	cccagggcat gggatgggta agcgcggagg cgtacatgcc gcagatgtcg tagacgtaga	6420
	ggggctctc gaggatgccg atgtagggtg ggtagcagcg ccccccgcg atgtggcgc	6480
	gcacgtagtc atacagctcg tgcgagggg cgaggagccc cgggccagg ttggtgcgac	6540
	tgggcttttc ggcgcggtag acgatctggc ggaaaatggc atgcgagttg gaggagatgg	6600
	tgggcctttg gaagatggtt aagtgggctg ggggcagtc gaccgagtc cgatgaagt	6660
	gggcgtagga gtcttcagc ttggcgacga gctcggcgt gactaggacg tccagagcg	6720
	agtagtcgag ggtctcttg atgatgtcat acttgagctg tccctttgt ttccacagct	6780
	cgcggttag aaggaactct tcgcggtcct tccagtactc ttcgagggg aaccgctct	6840
	gatctgcag gtaagagcct agcatgtaga actggttgac ggcctttag gcgcagcagc	6900
	ccttctccac ggggagggcg taggcctggg cggccttgcg cagggagtg tgcgtgagg	6960
[0139]	cgaaagtgtc cctgaccatg accttgagga actggtgctt gaagtcgata tcgtcgcagc	7020
	ccccctgctc ccagagctgg aagtccgtgc gcttcttgta ggcggggtg ggcaaagcga	7080
	aagtaacatc gttgaagagg atcttgcccg cgcggggcat aaagttgcga gtgatgcgga	7140
	aagttgggg cacctcgcc cggttggtga tgacctgggc ggcgagcac atctcgtcga	7200
	agccgttgat gttgtggccc acgatgtaga gttccacgaa tcgcgacgg cccttgacgt	7260
	ggggcagttt cttgagctc tcgtaggtga gctcgtcggg gtcgctgagc ccgtgctgct	7320
	cgagcgcca gtcggcgaga tgggggttg cgcggaggaa ggaagtccag agatccacgg	7380
	ccagggcggg ttgcagacgg tcccgtact gacggaactg ctgcccagc gccattttt	7440
	cgggggtgac gcagtagaag gtgcgggggt ccccggtcca gcgatccat ttgagctgga	7500
	gggcgagatc gagggcgagc tcgacgagcc ggtcgtcccc ggagagtttc atgaccagca	7560
	tgaaggggac gagctgctt ccgaaggacc ccatccaggt gtaggtttcc acatcgtagg	7620
	tgaggaagag cctttcgggt cgaggatgag agccgatggg gaagaactgg atctcctgcc	7680
	accaattgga ggaatggctg ttgatgtgat ggaagtagaa atgccgacg cgcgccgaac	7740

	actcgtgctt gtgtttatac aagcggccac agtgctcgca acgctgcacg ggatgcacgt	7800
	gctgcacgag ctgtacctga gttcctttga cgaggaattt cagtgggaag tggagtcgtg	7860
	gcgcctgcat ctcgtgctgt actacgtcgt ggtggtcggc ctggccctct tctgcctcga	7920
	tggtggtcac gctgacgagc ccgcgcggga ggcaggtcca gacctcggcg cgagcgggtc	7980
	ggagagcgag gacgagggcg cgcagggcgg agctgtccag ggtcctgaga cgctgcggag	8040
	tcaggtcagt gggcagcggc ggccgcgggt tgacttgacg gagttttcc agggcgcgcg	8100
	ggaggtccag atggtacttg atctccaccg cgccattggt ggcgacgtcg atggcttgca	8160
	gggtcccgtg cccttgggtg gtgaccaccg tccccgttt cttcttgggc ggctggggcg	8220
	acggggggcg tgcctcttcc atggttagaa gcggcggcga ggacgcgcgc cgggcggcag	8280
	gggcggctcg gggcccggag gcagggcgcg caggggcacg tcggcggcgc gcgcgggtag	8340
	gttctggtac tgcgcccgga gaagactggc gtgagcgacg acgcgacggt tgacgtcctg	8400
	gatctgacgc ctctgggtga aggccacggg acccgtgagt ttgaacctga aagagagttc	8460
[0140]	gacagaatca atctcggtat cgttgacggc ggccctgccg aggatctctt gcacgtgcc	8520
	cgagttgtcc tggtaggcga tctcggcat gaactgctcg atctcctct cttgaaggtc	8580
	tccgcggcgg gcgcgctcca cggcggcgc gaggtcgttg gagatgcggc ccatgagctg	8640
	cgagaaggcg ttcattgccc cctcgttcca gacgcggctg tagaccacga cgccctcggg	8700
	atcgcggcgg cgcatgacca cctgggcgag gttgagctcc acgtggcgcg tgaagaccgc	8760
	gtagttgcag aggcgctggt agaggtagtt gagcgtggtg gcgatgtgct cggtagcгаа	8820
	gaaatacatg atccagcggc ggagcggcat ctgctgacg tcgcccagcg cctccaaacg	8880
	ttccatggcc tcgtaaaagt ccacggcgaa gttgaaaaac tgggagttgc gcgccgagac	8940
	ggtcaactcc tcctccagaa gacggatgag ctggcgatg gtggcgcgca cctcgcgctc	9000
	gaagccccc gggagttcct ccaactctc ttcttctc tccactaaca tctcttctac	9060
	ttctcctca ggcggcagtg gtggcggggg agggggcctg cgtcgcggc ggccgacggg	9120
	cagacggtcg atgaagcgtc cgatggtctc gccgcggcgg cgtcgcagtg tctcggtagc	9180
	ggcgcggccc tcctcgggg gccgcagcgt gaagacgcc cgcgcacatc ccaggtggcc	9240

	gggggggtcc cgttgggca gggagagggc gctgacgatg catcttatca attgccccgt	9300
	agggactccg cgcaaggacc tgagcgtctc gagatccacg ggatctgaaa accgctgaac	9360
	gaaggcttcg agccagtcgc agtcgcaagg taggctgagc acggtttctt ctggcgggtc	9420
	atgttggttg ggagcggggc gggcgatgct gctggtgatg aagttgaaat aggcggttct	9480
	gagacggcgg atggtggcga ggagcaccag gtctttgggc ccggcttgct ggatgcbgag	9540
	acggtcggcc atgccccagg cgtggtcctg acacctggcc aggtccttgt agtagtctctg	9600
	catgagccgc tccacgggca cctcctcctc gcccgcgcgg ccgtgcatgc gcgtgagccc	9660
	gaagccgcgc tggggctgga cgagcggcag gtcggcgacg acgcgctcgg cgaggatggc	9720
	ttgctggatc tgggtgaggg tggctgga gtcacaaag tcgacgaagc ggtggtaggc	9780
	tccggtgttg atggtgtagg agcagttggc catgacggac cagttgacgg tctggtggcc	9840
	cggacgcacg agctcgtggt acttgaggcg cgagtaggcg cgcgtgtcga agatgtagtc	9900
	gttcaggtg cgcaccaggt actggtagcc gatgaggaag tgcggcggcg gctggcggta	9960
	gagcggccat cgctcgttgg cgggggcgcc gggcgcgagg tcctcgagca tggtagcgtg	10020
[0141]	gtagccgtag atgtacctgg acatccaggt gatgccggcg gcggtggtgg aggcgcgcgg	10080
	gaactcgcgg acgcggttcc agatgttgcg cagcggcagg aagtagttca tggtaggcac	10140
	ggtctggccc gtgagggcgc cgacagtcgt gatgctctat acgggcaaaa acgaaagcgg	10200
	tcagcggctc gactccgtgg cctggaggct aagcgaacgg gttgggctgc gcgtgtacce	10260
	cggttcgaat ctgaaatcag gctggagccg cagctaactg ggtattggca ctcccgtctc	10320
	gaccaagcc tgcaccaacc ctccaggata cggaggcggg tcgttttgca actttttttt	10380
	ggaggccgga tgagactagt aagcgcgga agcggccgac cgcgatggct cgctgccgta	10440
	gtctggagaa gaatgccag ggttgcgttg cgggtgtccc cggttcgagg ccggccggat	10500
	tccgcggcta acgagggcgt ggctgccccg tcgtttccaa gaccccatag ccagccgact	10560
	tctccagtta cggagcgcgc cctctttttg ttttgtttgt ttttgccaga tgcateccgt	10620
	actcggcag atgccccc accaccctcc accgcaacaa cagccccctc cacagccggc	10680
	gcttctgccc ccgccccage agcaacttcc agccacgacc gcccgggccg ccgtgagcgg	10740
	ggctggacag agttatgatc accagctggc cttggaagag ggcgaggggc tggcgcgcct	10800

	gggggcgtcg tcgccggagc ggcacccgcg cgtgcagatg aaaaggagc ctcgcgaggc	10860
	ctacgtgccc aagcagaacc tgttcagaga caggagcggc gaggagcccg aggagatgcg	10920
	cgcgccccgg ttccacgcgg ggcgggagct gcggcgcggc ctggaccgaa agagggtgct	10980
	gaggacgag gatttcgagg cggacgagct gacggggatc agccccgcgc gcgcgcacgt	11040
	ggccgcggcc aacctggtca cggcgtacga gcagaccgtg aaggaggaga gcaacttcca	11100
	aaaatccttc aacaaccacg tgcgcaccct gatcgcgcgc gaggaggtga ccctgggcct	11160
	gatgcacctg tgggacctgc tggaggccat cgtgcagaac cccaccagca agccgctgac	11220
	ggcgcagctg ttcttggtgg tgcagcatag tcgggacaac gaagcgttca gggaggcgt	11280
	gctgaatata accgagcccg agggccgctg gctcctggac ctggtgaaca ttctgcagag	11340
	catcgtggtg caggagcgcg ggctgccgct gtccgagaag ctggcggcca tcaacttctc	11400
	ggtgetgagt ttggcaagt actacgctag gaagatctac aagaccccgt acgtgccc	11460
	agacaaggag gtgaagatcg acgggtttta catgcgatg accctgaaag tgctgaccct	11520
[0142]	gagcgacgat ctgggggtgt accgcaacga caggatgcac cgtgcggtga gcgccagcag	11580
	gcggcgcgag ctgagcgacc aggagctgat gcatagtctg cagcgggccc tgaccggggc	11640
	cgggaccgag ggggagagct actttgacat gggcgcggac ctgcactggc agcccagccg	11700
	ccgggccttg gagcggcgg caggacccta ctagaagag gtggacgatg aggtggacga	11760
	ggagggcgag tacctggaag actgatggcg cgaccgtatt ttgctagat gcaacaacaa	11820
	cagccacctc ctgatcccgc gatgcgggcg gcgctgcaga gccagccgtc cggcattaac	11880
	tcctcggacg attggacca ggccatgcaa cgcacatg cgctgacgac ccgcaacccc	11940
	gaagccttta gacagcagcc ccagccaac cggtctcgg ccatcctgga ggccgtggtg	12000
	ccctcgcgct ccaaccccac gcacgagaag gtccctggcca tcgtgaacgc gctggtggag	12060
	aacaaggcca tccgcggcga cgaggccggc ctggtgtaca acgcgctgct ggagcgcgtg	12120
	gcccctaca acagcaccaa cgtgcagacc aacctggacc gcatggtgac cgacgtgcgc	12180
	gaggccgtgg cccagcgcga gcggttccac cgcgagtcca acctgggatc catggtggcg	12240
	ctgaacgcct tcctcagcac ccagcccgc aacgtgcccc ggggccagga ggactacacc	12300

	aacttcatca ggcacctgcg cctgatggtg accgaggtgc cccagagcga ggtgtaccag	12360
	tccgggccgg actacttctt ccagaccagt cgccagggtt tgcagaccgt gaacctgagc	12420
	caggctttca agaacttgea gggcctgtgg ggcgctcagg ccccggctcg ggaccgcgcg	12480
	acggtgtcga gcctgctgac gccgaactcg cgcctgctgc tgctgctggt ggcccccttc	12540
	acggacagcg gcagcatcaa ccgcaactcg tacctgggtt acctgattaa cctgtaccgc	12600
	gaggccatcg gccaggcgca cgtggacgag cagacctacc aggagatcac ccacgtgagc	12660
	cgcgccctgg gccaggacga cccgggcaac ctggaagcca ccctgaactt tttgctgacc	12720
	aaccggtcgc agaagatccc gccccagtac gcgctcagca ccgaggagga gcgcatcctg	12780
	cgttacgtgc agcagagcgt gggcctgttc ctgatgcagg agggggccac cccagcgc	12840
	gcgctcgaca tgaccgcgcg caacatggag cccagcatgt acgccagcaa ccgcccgttc	12900
	atcaataaac tgatggacta cttgcatcgg gcggcccca tgaactctga ctatttcacc	12960
	aacccatcc tgaatcccca ctggtccccg ccgcccgggt tctacacggg cgagtacgac	13020
	atgcccgacc ccaatgacgg gttcctgtgg gacgatgtgg acagcagcgt gttctcccc	13080
[0143]	cgaccgggtg ctaacgagcg ccccttgtgg aagaaggaag gcagcgaccg acgcccgtcc	13140
	tcggcgctgt ccggccgca gggctgtgcc gcggcggtgc ccgaggccgc cagtcccttc	13200
	ccgagcttgc ccttctcgct gaacagtatc cgcagcagcg agctgggcag gatcacgcgc	13260
	ccgcgcttgc tgggcgaaga ggagtacttg aatgactcgc tgttgagacc cgagcgggag	13320
	aagaacttcc ccaataacgg gatagaaagc ctggtggaca agatgagccg ctggaagacg	13380
	tatgcgcagg agcacaggga cgatccccg gcgctgcagg gggccacgag ccggggcagc	13440
	gcccccgtg aacgcccgtg gcacgacagg cagcggggac agatgtggga cgatgaggac	13500
	tcccccagc acagcagcgt gttggacttg ggtgggagtg gtaaccggtt cgctcacctg	13560
	cgccccgtg tcgggcgcat gatgtaagag aaaccgaaaa taaatgatac tcaccaaggc	13620
	catggcgacc agcgtgcggt cgtttcttct ctgtttgtgt tgtatctagt atgatgagc	13680
	gtgctgacct ggagggtcct cctccctcgt acgagagcgt gatgcagcag gcgatggcgg	13740
	cggcggcgat gcagccccg ctggaggctc cttacgtgcc cccgcggtac ctggcgcta	13800
	cggagggcg gaacagcatt cgttactcgg agctggcacc cttgtacgat accaccggt	13860

	tgtacctggt ggacaacaag tcggcggaca tcgcctcgct gaactaccag aacgaccaca	13920
	gcaacttctt gaccaccgtg gtgcagaaca atgacttcaac cccacaggag gccagcacc	13980
	agaccatcaa ctttgacgag cgctcgcggt ggggcggcca gctgaaaacc atcatgcaca	14040
	ccaacatgcc caacgtgaac gagttcatgt acagcaaca gttcaaggcg cgggtgatgg	14100
	tctcccga gacccccaat ggggtgacag tgacagagga ttatgatggt agtcaggatg	14160
	agctgaagta tgaatgggtg gaatttgagc tgcccgaagg caacttctcg gtgaccatga	14220
	ccatcgacct gatgaacaac gccatcatcg acaattactt ggcggtgggg cggcagaacg	14280
	gggtgctgga gagcgacatc ggcgtaagt tcgacactag gaattcagg ctgggctggg	14340
	accccgtagc cgagctggtc atgcccgggg tgtacaccaa cgaggetttc catcccata	14400
	ttgtcttgct gcccgctgc ggggtggact tcaccgagag ccgcctcagc aacctgctgg	14460
	gcattcgca gaggcagccc ttccaggaag gcttccagat catgtacgag gatctggagg	14520
	ggggcaacat ccccgcgctc ctggatgtcg acgcctatga gaaaagcaag gaggatgcag	14580
[0144]	cagctgaagc aactgcagcc gtagctaccg cctctaccga ggtaggggc gataatcttg	14640
	caagcgccgc agcagtggca gggccgagg cggctgaaac cgaaagtaag atagtcatc	14700
	agccggtgga gaaggatagc aagaacagga gctacaacgt actaccggac aagataaaca	14760
	ccgcctaccg cagctggtac ctagcctaca actatggcga ccccgagaag ggcgtgcgct	14820
	cctggacgct gctcaccacc tcggacgtca cctgcggcgt ggagcaagtc tactggtcgc	14880
	tgcccacat gatgcaagac ccggtcacct tccgctccac gcgtcaagtt agcaactacc	14940
	cgggtgggg cgccgagctc ctgcccgtct actccaagag cttcttcaac gagcaggccg	15000
	tctactcgca gcagctgcgc gccttcacct cgtttacgca cgtcttcaac cgtttccccg	15060
	agaaccagat cctcgtccgc ccgcccgcgc ccaccattac caccgtcagt gaaaacgttc	15120
	ctgctctcac agatcacggg accctgccgc tgcgcagcag tatccgggga gtccagcgcg	15180
	tgaccgttac tgacgccaga cgccgcacct gccctacgt ctacaaggcc ctgggcatag	15240
	tcgcgccgc cgtcctctcg agccgcacct tctaaatgct cattctcctc tcgccagta	15300
	ataacaccgg ttggggcctg cgcgccccca gcaagatgta cggaggcgt cgccaacgt	15360

	ccacgcaaca ccccgtgcgc gtgcgcgggc acttccgcgc tccttggggc gcctcaagg	15420
	gccgcgtgcg gtcgcgcacc accgtcgacg acgtgatcga ccaggtggtg gccgacgcgc	15480
	gcaactacac ccccgccgcc gcgcccgtet ccaccgtgga cgccgtcadc gacagcgtgg	15540
	tggccgacgc gcgccggtac gcccgcgcca agagccggcg gcggcgcatc gcccgcgggc	15600
	accggagcac ccccgccatg cgcgcggcgc gacccttctt gcgcagggcc aggcgcacgg	15660
	gacgcagggc catgctcagg gcggccagac gcgcggcttc aggcgccagc gccggcagga	15720
	cccgagacg cgcggccacg gcggcggcag cgcccatcgc cagcatgtcc cgcggcgggc	15780
	gagggaacgt gtactgggtg cgcgacgccg ccaccggtgt gcgcgtgccc gtgcgcaccc	15840
	gccccctcg cacttgaaga tgttcacttc gcgatgttga tgtgtcccag cggcgaggag	15900
	gatgtccaag cgcaaattca aggaagagat gctccaggtc atcgcgcctg agatctacgg	15960
	ccctgcggtg gtgaaggagg aaagaaagcc ccgcaaaatc aagcgggtca aaaaggacaa	16020
	aaaggaagaa gaaagtgatg tggacggatt ggtggagttt gtgcgcgagt tcgcccccg	16080
	gcggcgcgtg cagtggcgcg ggcggaaggt gcaaccggtg ctgagaccgc gcaccaccgt	16140
[0145]	ggtcttcacg cccggcgagc gctccggcac cgcttccaag cgctcctacg acgaggtgta	16200
	cggggatgat gatattctgg agcaggcggc cgagcgcctg ggcgagtttg cttacggcaa	16260
	gcgcagccgt tccgaccga aggaagaggc ggtgtccatc ccgctggacc acggcaacce	16320
	caccccgagc ctcaagcccc tgacctgca gcaggtgctg ccgaccgcgg cgcgcgccg	16380
	ggggttcaag cgcgagggcg aggatctgta cccaccatg cagctgatgg tgcccaagcg	16440
	ccagaagctg gaagacgtgc tggagaccat gaaggtggac ccggacgtgc agcccgaggt	16500
	caaggtgcgg cccatcaagc aggtggcccc gggcctgggc gtgcagaccg tggacatcaa	16560
	gattcccacg gagcccatgg aaacgcagac cgagcccatg atcaagccca gcaccagcac	16620
	catggagggtg cagacggate cctggatgcc atcggtcctt agtcgaagac cccggcgcaa	16680
	gtacggcgcg gccagcctgc tgatgcccga ctacgcgctg catccttcca tcatccccac	16740
	gccgggctac cgcggcacgc gcttctaccg cggtcatacc agcagccgcc gccgcaagac	16800
	caccactcgc cgcggcgtc gccgcaccgc cgctgcaacc acccctgccg ccctggtgcg	16860
	gagagtgtac cgcggcgccc gcgcacctct gaccctgccg cgcgcgcgct accaccgag	16920

	catgccatt taaactttcg cctgctttgc agatcaatgg ccctcacatg ccgccttcgc	16980
	gttcccatta cgggctaccg aggaagaaaa ccgcgccgta gaaggtggc ggggaacggg	17040
	atgcgtcgcc accaccaccg gcggcggcgc gccatcagca agcggttggg gggaggcttc	17100
	ctgcccgcgc tgatecccat catcgccgeg gcgategggg cgatccccgg cattgettcc	17160
	gtggcgggtgc aggcctctca gcgccactga gacacacttg gaaacatctt gtaataaacc	17220
	aatggactct gacgctcctg gtctctgat gtgttttcgt agacagatgg aagacatcaa	17280
	tttttcgtcc ctggctccgc gacacggcac gcggccgttc atgggcacct ggagcgacat	17340
	cggcaccagc caactgaacg ggggcgcctt caattggagc agtctctgga gcgggcttaa	17400
	gaatttcggg tccacgctta aaacetatgg cagcaaggcg tggaacagca ccacagggca	17460
	ggcgtgagg gataagctga aagagcagaa cttccagcag aagggtgtcg atgggctcgc	17520
	ctcgggcatc aacggggtgg tggacctggc caaccaggcc gtgcagcggc agatcaacag	17580
	ccgcctggac ccggtgccgc ccgccggctc cgtggagatg ccgcaggtgg aggaggagct	17640
[0146]	gcctcccctg gacaagcggg gcgagaagcg accccgcccc gatgcggagg agacgtgct	17700
	gacgcacacg gacgagccgc ccccgtagca ggaggcggtg aaactgggtc tgcccaccac	17760
	gcggcccatc gcgcccctgg ccaccggggt gctgaaacc gaaaagcccg cgaccctgga	17820
	cttgctcct ccccagcctt cccgcccctc tacagtggct aagcccctgc cgccggtggc	17880
	cgtggcccgc gcgcgaccg ggggcaccgc ccgccctcat gcgaactggc agagcactct	17940
	gaacagcatc gtgggtctgg gagtgcagag tgtgaagcgc cgccgtgct attaaaccta	18000
	ccgtagcgct taacttgctt gtctgtgtgt gtatgtatta tgtcgcgcc gcgctgtcc	18060
	accagaagga ggagtgaaga ggcgcgtcgc cgagttgcaa gatggccacc ccatcgatgc	18120
	tgccccagtg ggcgtacatg cacatcgccg gacaggacgc ttcggagtac ctgagtccgg	18180
	gtctggtgca gtttgcccgc gccacagaca cctacttcag tctggggaac aagtttagga	18240
	accccacggt ggcgcccacg cacgatgtga ccaccgaccg cagccagcgg ctgacgctgc	18300
	gcttcgtgcc cgtggaccgc gaggacaaca cctactcgtg caaagtgcgc tacacgctgg	18360
	ccgtgggcga caaccgcgtg ctggacatgg ccagcaccta ctttgacatc cgcggcgtgc	18420

	tggatcgggg ccctagcttc aaacctact cggcaccgc ctacaacagt ctggcccca	18480
	agggagcacc caacacttgt cagtggacat ataaagccga tggtgaaact gccacagaaa	18540
	aaacctatac atatggaaat gcaccctgc agggcattaa catcacaaaa gatggtatc	18600
	aacttggaa tgacaccgat gatcagccaa tctacgcaga taaaacctat cagcctgaac	18660
	ctcaagtggg tgatgctgaa tggcatgaca tcaactgtac tgatgaaaag tatggaggca	18720
	gagctcttaa gcctgatacc aaaatgaagc cttgttatgg ttcttttgcc aagcctacta	18780
	ataaagaagg aggtcaggca aatgtgaaaa caggaacagg cactactaaa gaatatgaca	18840
	tagacatggc tttctttgac aacagaagtg cggctgctgc tggcctagct ccagaaattg	18900
	ttttgtatac tgaanaatgtg gatttggaaa ctccagatac ccatattgta tacaagcag	18960
	gcacagatga cagcagctct tctattaatt tgggtcagca agccatgcc aacagaccta	19020
	actacattgg tttcagagac aactttatcg ggctcatgta ctacaacagc actggcaata	19080
	tgggggtgct ggccggtcag gcttctcagc tgaatgctgt gtttgacttg caagacagaa	19140
	acaccgagct gtcctaccag ctcttgcttg actctctggg tgacagaacc cggtatctca	19200
[0147]	gtatgtggaa tcaggcgggtg gacagctatg atcctgatgt gcgcattatt gaaaatcatg	19260
	gtgtggagga tgaacttccc aactattggt tccctctgga tgctgttggc agaacagata	19320
	cttatcaggg aattaaggct aatggaactg atcaaaccac atggacaaa gatgacagtg	19380
	tcaatgatgc taatgagata ggcaagggtg atccattgc catggaaatc aacatccaag	19440
	ccaacctgtg gaggaacttc ctctacgcca acgtggccct gtacctgcc gactcttaca	19500
	agtacacgcc ggccaatggt accctgccca ccaacaccaa cacctacgat tacatgaacg	19560
	gccgggtggt ggccgctcg ctggtggact cctacatcaa catcggggcg cgctggtcgc	19620
	tggatcccat ggacaacgtg aacccttca accaccaccg caatgcgggg ctgcgtacc	19680
	gctccatgct cctgggcaac gggcgetacg tgccttcca catccaggtg cccagaaat	19740
	ttttgccat caagacctc ctgctcctgc cgggtccta cacctacgag tggaacttcc	19800
	gcaaggacgt caacatgate ctgcagagct ccctcgcaa cgacctgctc acggacgggg	19860
	cctccatctc cttcaccage atcaacctct acgccacctt cttccccatg gcgcacaaca	19920
	cggcctccac gctcaggcc atgctgcgca acgacaccaa cgaccagtcc ttcaacgact	19980

	acctctcggc ggccaacatg ctctacceca tcccggccaa cgccaccaac gtgcccattc	20040
	ccatcccctc gcgcaactgg gccgccttcc gcggetggtc cttcacgcgt ctcaagacca	20100
	aggagacgcc ctogctgggc tccgggttcg acccctactt cgtctactcg ggctccatcc	20160
	cctacctcga cggcaccttc tacctcaacc acaccttcaa gaaggtctcc atcaccttcg	20220
	actcctccgt cagctggccc ggcaacgacc ggctcctgac gcccaacgag ttcgaaatca	20280
	agcgcaccgt cgacggcgag ggctacaacg tggcccagtg caacatgacc aaggactggc	20340
	tectggcca gatgctggcc cactacaaca tcggtacca gggcttctac gtgcccagg	20400
	gctacaagga ccgcatgtac tccttcttcc gcaacttcca gcccatgagc cggcaggcgg	20460
	tggacgaggt caactacaag gactaccagg ccgtcacctt ggctaccag cacaacaact	20520
	cgggcttcgt cggtacctc gcgcccacca tgcgccaggg ccagccctac cccgccaact	20580
	accctaccg gctcatcggc aagagcgccg tcaccagcgt caccagaaa aagtctctct	20640
	gcgacagggt catgtggcgc atccccttct ccagcaactt catgtccatg ggcgcgctca	20700
[0148]	ccgacctcgg ccagaacatg ctctatgcca actccgccc cgcgctagac atgaatttcg	20760
	aagtcgacc catggatgag tccacccttc tctatgttgt cttcgaagtc ttcgacgtcg	20820
	tccgagtgea ccagcccac cgcggcgtca tcgaggccgt ctacctgcgc accccttct	20880
	cggccggtaa cggcaccacc taagetcttg cttcttgcaa gccatggccg cgggctccgg	20940
	cgagcaggag ctcaggcca tcatccgca cctgggctgc gggccctact tctgggcac	21000
	cttcgataag cgcttcccgg gattcatggc cccgcacaag ctggcctgcg ccatcgtcaa	21060
	cacggccggc cgcgagaccg ggggcgagca ctggctggcc ttcgcctgga acccgcgctc	21120
	gaacacctgc tacctcttcg accccttcgg gttctcggac gagcgcctca agcagateta	21180
	ccagttcgag tacgagggcc tgctgcgcc cagcgccttg gccaccgagg accgctgcgt	21240
	caccctgga aagtcaccc agaccgtgca gggctcgcgc teggccctt cgggctctt	21300
	ctgtgcatg ttctgcacg cttctgtgca ctggcccac cggcccatgg acaagaacc	21360
	caccatgaac ttgctgacgg gggtgcccaa cggcatgctc cagtcgccc aggtggaacc	21420
	caccctgcgc cgcaaccagg aggcgtctta ccgcttctc aactcccact ccgctactt	21480

	tcgctccac cgcgcgcgca tcgagaaggc caccgccttc gaccgcatga atcaagacat	21540
	gtaaaccgtg tgtgtatggt aaatgtcttt aataaacagc actttcatgt tacacatgca	21600
	tctgagatga tttatrtaga aatcgaaagg gttctgccgg gtctcggcat ggcccgcggg	21660
	caggacacg ttgcggaact ggtacttggc cagccacttg aactcgggga tcagcagttt	21720
	gggcagcggg gtgtcgggga aggagtcggt ccacagcttc cgcgtcagtt gcagggcgcc	21780
	cagcaggtcg ggcgcggaga tcttgaatc gcagttggga cccgcgttct gcgcgcggga	21840
	gttgcggtac acggggttgc agcactggaa caccatcagg gccgggtgct tcacgctcgc	21900
	cagcaccgtc gcgtcgggta tgctctccac gtcgaggctc tcggcgttgg ccatcccga	21960
	gggggtcatc ttgcaggtct gccttcccat ggtgggcacg caccgggct tgtggttga	22020
	atgcagtg cagggggatca gcatcatctg ggcttggctg gcgttcatcc ccgggtacat	22080
	ggccttcatg aaagcctcca attgcctgaa cgcctgetgg gccttggctc cctcggtgaa	22140
	gaagaccccg caggacttgc tagagaactg gttggtggcg caccggcgt cgtgcacgca	22200
	gcagcgcgcg tcgttgttgg ccagctgcac cacgctgcgc cccagcggg tctgggtgat	22260
[0149]	cttggcccgg teggggttct ccttcagcgc gcgctgccc ttctcctcgc ccacatccat	22320
	ctcgatcatg tgctccttct ggatcatggt ggtcccgtgc aggaccgca gcttgcctc	22380
	ggcctcggtg caccctgca gccacagcgc gcaccgggtg cactcccagt tcttgtggc	22440
	gatctgggaa tgcgcgtgca cgaagccctg caggaagcgg cccatcatgg tggtcagggt	22500
	cttgttgcta gtgaaggtea gcggaatgcc gcggtgctcc tcgttgatgt acaggtggca	22560
	gatcgcggcg tacacctgc cctgctcggg catcagctgg aagttggctt tcaggtcgg	22620
	ctccacgcgg tagcgggtcca tcagcatagt catgatttc atacccttct cccaggccga	22680
	gacgatgggc aggtcatag ggttcttcac catcatctta gcgctagcag ccgcccag	22740
	gggtcgtc tcgtccagg tctcaaagct ccgcttccc tcttctcgg tgatccgac	22800
	cgggggtag ctgaagccca cggcccccag ctctctcgc gcctgtctt cgtcctcgt	22860
	gtcctggctg acgtcctgca ggaccacatg cttggtcttg cggggtttct tcttggcgg	22920
	cagcggcggc ggagatggtg gagatggcga gggggagcgc gagttctcgc tcaccactac	22980
	tatctcttcc tcttcttggg ccgaggccac gcggcggtag gtatgtctct tcggggcag	23040

	aggcggaggc gacgggctct cgcccgcg acttggcgga tggctggcag agccccttcc	23100
	gcgttcgggg gtgcgtccc ggcggcgtc tgactgactt cctccgcggc cggccattgt	23160
	gttctcctag ggaggaacaa caagcatgga gactcagcca tcgccaacct cgccatctgc	23220
	ccccaccgcc gacgagaagc agcagcagca gaatgaaage ttaaccgccc cgccgccag	23280
	ccccccacc tccgacgcgg ccgtcccaga catgcaagag atggaggaat ccatcgagat	23340
	tgacctgggc tatgtgacgc ccgcggagca cgaggaggag ctggcagtgc gcttttcaca	23400
	agaagagata caccaagaac agccagagca ggaagcagag aatgagcaga gtcaggctgg	23460
	gctcgagcat gacggcgact acctccacct gagcgggggg gaggacgcgc tcatcaagca	23520
	tctggcccgg caggccacca tegtcaagga tgcgtgctc gaccgcaccg aggtgccct	23580
	cagcgtggag gagctcagcc ggcctacga gttgaacctc ttctcgcgc gcgtgcccc	23640
	caagcgcag cccaatggca cctgcgagcc caaccgcgc ctcaacttct acccgttctt	23700
	cgcggtgccc gaggccctgg ccacctacca catcttttcc aagaacaaa agatccccgt	23760
[0150]	ctctgcccgc gccaaccgca ccgcgcccga cgcccttttc aacctgggtc ccggcgcggc	23820
	cctacctgat atgcctcct tggaagaggt tcccaagatc ttcgagggtc tgggcagcga	23880
	cgagactcgg gccgcgaacg ctctgcaagg agaaggagga gagcatgagc accacagcgc	23940
	cctggtcgag ttggaaggcg acaacgcgcg gctggcgggtg ctcaaacgca cggtcgagct	24000
	gaccatttc gcctaccggt ctctgaacct gcccccaaa gtcatgagcg cggtcatgga	24060
	ccaggtgctc atcaagcgcg cgtcgcccat ctccgaggac gagggeatgc aagactccga	24120
	ggagggcaag cccgttgtca ggcagagca gctggcccgg tggctgggtc ctaatgctag	24180
	tccccagagt ttggaagagc ggcgcaaaact catgatggcc gtggtcctgg tgaccgtgga	24240
	gctggagtgc ctgcgccgt tcttcgccga cgcggagacc ctgcaagagg tcgaggagaa	24300
	cctgcactac ctcttcagc acgggttcgt gcgccaggcc tgcaagatct ccaacgtgga	24360
	gctgaccaac ctggtctcct acatgggcat cttgcacgag aaccgcctgg ggcagaacgt	24420
	gctgcacacc accctgcgcg gggaggcccg gcgcgactac atccgcgact gcgtctacct	24480
	ctacctctgc cacacctggc agacgggcat gggcgtgtgg cagcagtgtc tggaggagca	24540

	gaacctgaaa gagctctgca agctcctgca gaagaacctc aagggtctgt ggaccgggtt	24600
	cgacgagcgc accaccgct cggacctggc cgacctcatt ttccccgagc gcctcaggct	24660
	gacgctgcgc aacggcctgc ccgactttat gagccaaagc atgttgcaaa actttegctc	24720
	tttcatcctc gaacgctccg gaatcctgcc cgccacctgc tccgcgctgc cctcggactt	24780
	cgtgccgctg accttccgag agtgcccccc gccgctgtgg agccactgct acctgctgcg	24840
	cctggccaac tacctggcct accactcgga cgtgatcgag gacgtcagcg gcgagggcct	24900
	gctcgagtgc cactgccgct gcaacctctg cacgccgcac cgctccctgg cctgcaacct	24960
	ccagctgctg agcgagacct agatcatcgg caccttcgag ttgcaagggc ccagcgaagg	25020
	cgagggttca gccgccaagg ggggtctgaa actcaccctg gggctgtgga cctcggccta	25080
	cttgcgcaag ttcgtgcccc aggactacca tcccttcgag atcaggttct acgaggacca	25140
	atcccatccg cccaaggccg agctgtcggc ctgcgtcacc acccaggggg cgatcctggc	25200
	ccaattgcaa gccatccaga aatcccgcca agaattcttg ctgaaaaagg gccgcggggt	25260
	ctacctgac ccccagaccg gtgaggagct caaccccggc ttccccagg atgccccgag	25320
[0151]	gaaacaagaa gctgaaagtg gagctgccgc ccgtggagga ttggaggaa gactgggaga	25380
	acagcagtca ggcagaggag gaggagatgg aggaagactg ggacagcact caggcagagg	25440
	aggacagcct gcaagacagt ctggaggaag acgaggagga ggcagaggag gaggtggaag	25500
	aagcagccgc cgccagaccg tcgtcctcgg cgggggagaa agcaagcagc acggatacca	25560
	tctccgctcc gggtcggggt cccgctcagc cacacagtag atgggacgag accggacgat	25620
	tccgaacct caccaccag accggtaaga aggagcggca gggatacaag tcttggcggg	25680
	ggcacaacaa cgccatcgtc tctgcttgc aggcctcgg gggcaacatc tcctcacc	25740
	ggcgtacct gctcttccac cgcggggtga actttcccc caacatcttg cttactacc	25800
	gtcacctcca cagcccctac tacttccaag aagaggcagc agcagcagaa aaagaccagc	25860
	agaaaaccag cagctagaaa atccacagcg gcggcagcag gtggactgag gatcgcggcg	25920
	aacgagccgg cgcaaacccg ggagctgagg aaccgatct tccccacct ctatgccatc	25980
	ttccagcaga gtcgggggca ggagcaggaa ctgaaagtca agaaccgttc tctgcgctcg	26040
	ctcaccgca gttgtctgta tcacaagagc gaagaccaac ttcagcgcac tctcaggac	26100

	gccgaggctc tcttcaaca gtactgcgcg ctactctta aagagtagcc cgcgcccgcc	26160
	cagtcgcaga aaaaggcggg aattacgtea cctgtgcct tegccctage cgctccacc	26220
	catcatcatg agcaaagaga ttcccacgcc ttacatgtgg agctaccagc cccagatggg	26280
	cctggccgcc ggtgccgcc aggactactc cacccgatg aattggctca gcgccgggcc	26340
	cgcgatgata tcacgggtga atgacatccg cccccaccga aaccagatac tcctagaaca	26400
	gtcagcgctc accgccacgc cccgcaatca cctcaatccg cgtaattggc ccgccgcct	26460
	ggtgtaccag gaaattcccc agcccacgac cgtactactt ccgcgagacg cccaggccga	26520
	agtccagctg actaactcag gtgtccagct ggccggcggc gccaccctgt gtcgtcaccg	26580
	ccccgctcag ggtataaagc ggctggatgat ccggggcaga ggcacacagc tcaacgacga	26640
	ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc ttccaactcg ccgatcggg	26700
	gagatcttcc ttcacgctc gtcaggccgt cctgactttg gagagtctgt cctcgcagcc	26760
	ccgctcgggt ggcatcggca ctctccagtt cgtggaggag ttactccct cggctactt	26820
[0152]	caacccttc tccgctccc ccggcacta cccggacgag ttcattccga acttcgacgc	26880
	catcagcgag tcggtggacg gctacgattg aatgtcccat ggtggcgcag ctgacctagc	26940
	tcgcttcga cacctggacc actgccccg ctccgctgc ttcgctcggg atctcgcga	27000
	gtttgcctac tttgagctgc ccgaggagca cctcagggc ccggcccacg gagtgcggat	27060
	cgctcgcgaa gggggcctcg actcccact gcttcggatc ttcagccagc gtccgatcct	27120
	ggtcgagcgc gagcaaggac agacccttct gactctgtac tgcatctgca accaccccg	27180
	cctgcatgaa agtctttgtt gtctgctgtg tactgagtat aataaaagct gagatcagcg	27240
	actactccgg acttcctgtt gttcctgaat ccatcaacca gtctttgttc ttcaccggga	27300
	acgagaccga gctccagctc cagtgtgaagc cccacaagaa gtacctcacc tggetgttcc	27360
	aggctcccc gatcgcctt gtaaccact gcgacaacga cggagtctg ctgagcggcc	27420
	ctgccaacct tactttttcc acccgagaa gcaagctcca gctcttcaa cccttctcc	27480
	ccgggacctc tcagtgcgtc tcgggacct gccatcacac cttccacctg atcccgaata	27540
	ccacagcgtc gctccccgt actaacaacc aaactaacct ccaccaacgc caccgtcgcg	27600

	acggccacaa tacatgceca tattagacta tgaggccgag ccacagcgac ccatgctccc	27660
	cgctattagt tacttcaatc taaccggcgg agatgactga cccactggcc aacaacaacg	27720
	tcaacgacct tetctggac atggacggcc gcgcctcgga gcagcgactc gcccaacttc	27780
	gcattcgcca gcagcaggag agagccgtca aggagctgca ggatgcgggtg gccatccacc	27840
	agtgcaagag aggcaccttc tgcctggtga aacaggccaa gatctcctac gaggtcactc	27900
	caaacgacca tgcctctcc tacgagctcc tgcagcagcg ccagaagttc acctgcctgg	27960
	tggagtcaa ccccatgctc atcaccagc agtctggcga taccaagggg tgcacccact	28020
	gctcctgca ctccccgac tgcgtccaca ctctgatcaa gaccctctgc ggctccgcg	28080
	acctcctccc catgaaacta tccccctt atccagtga ataaagatca tattgatgat	28140
	gattttacag aaataaaaaa taatcatttg atttgaata aagatacaat catattgatg	28200
	atttgagttt aacaaaaaaa taaagaatca cttacttgaa atctgatacc aggtctctgt	28260
	ccatgttttc tgccaacacc acttcactcc cctcttccca gctctggtac tgcaggcccc	28320
	ggcgggctgc aaacttctc cacacgtga aggggatgac aaattcctcc tgtccctcaa	28380
[0153]	tcttcatttt atctctatc agatgtccaa aaagcgcgtc cgggtggatg atgacttcca	28440
	ccccgtctac cctacgatg cagacaacgc accgaccgtg cccttcatca accccccctt	28500
	cgtctcttca gatggattcc aagagaagcc cctgggggtg ttgtccctgc gactggccga	28560
	ccccgtcacc accaagaacg gggaaatcac cctcaagctg ggagaggggg tggacctcga	28620
	ttctcggga aaactcatct ccaacacggc caccaaggcc gccgccctc tcagtttttc	28680
	caacaacacc atttccctta acatggatca ccccttttac actaaagatg gaaaattatc	28740
	cttacaagtt tctccacat taaatatact gagaacaagc attctaaaca cactagcttt	28800
	aggttttgga tcagttttag gactccgtgg ctctgccttg gcagtacagt tagtctctcc	28860
	acttacattt gatactgatg gaaacataaa gcttacetta gacagaggtt tgcattgtac	28920
	aacaggagat gcaattgaaa gcaacataag ctgggctaaa ggtttaaaat ttgaagatgg	28980
	agccatagca accaacattg gaaatgggtt agagtttggga agcagtagta cagaaacagg	29040
	tgttgatgat gcttacccea tccaagttaa acttggatct ggcccttagct ttgacagtac	29100
	aggagccata atggctggta acaagaaga cgataaactc actttgtgga caacacctga	29160

	tccatcacca aactgtcaaa tactcgcaga aatgatgca aaactaacac ttgtcttgac	29220
	taaatgtggt agtcaaatac tggccactgt gtcagtctta gttgtaggaa gtggaaacct	29280
	aaacccatt actggcaccg taagcagtgc tcaggtgttt ctacgttttg atgcaaacgg	29340
	tgttctttta acagaacatt ctacactaaa aaaatactgg gggatataggc agggagatag	29400
	catagatggc actccatata ccaatgctgt aggattcatg cccaatttaa aagcttatcc	29460
	aaagtcacaa agttctacta ctaaaaataa tatagtaggg caagtataca tgaatggaga	29520
	tgtttcaaaa cctatgcttc tcaactataac cctcaatggt actgatgaca gcaacagtac	29580
	atattcaatg tcattttcat acacctggac taatggaagc tatgttggag caacatttgg	29640
	ggctaactct tataccttct catacatgc ccaagaatga aactgtatc ccacctgca	29700
	tgccaacct tcccaccca ctctgtggaa caactctga aacacaaaat aaaataaagt	29760
	tcaagtgttt tattgattca acagttttac aggattcgag cagttatfff tcctccacce	29820
	tcccaggaca tggaatacac caccctctcc ccccgcacag ccttgaacat ctgaatgcca	29880
[0154]	ttggtgatgg acatgctfff ggtctccacg ttccacacag ttccagagcg agccagtctc	29940
	gggtcggtea gggagatgaa accctccggg cactcccga tctgcacctc acagctcaac	30000
	agctgaggat tgcctcggg ggtcgggatc acggttatct ggaagaagca gaagagcggc	30060
	ggtgggaatc atagtcgcg aacgggatcg gccggtggtg tcgcatcagg ccccgtagca	30120
	gtcgtgccc cgcgctcc gtcaagctgc tgcctcaggg gtccgggtcc agggactccc	30180
	tcagcatgat gcccacggcc ctacatca gtcgtctggt gcggcgggcg cagcagcgca	30240
	tgcggatctc gctcaggtcg ctgcagtac tgcaacacag aaccaccagg ttgttcaaca	30300
	gtccatagtt caaacgctc cagccgaaac tcatecgagg aaggatgcta cccacgtgge	30360
	cgctgtacca gatcctcagg taaatcaagt ggtgccccct ccagaacacg ctgcccacgt	30420
	acatgatctc cttgggcatg tggcggttca ccacctccc gtaccacatc accctctggt	30480
	tgaacatgca gccccggatg atcctgcgga accacagggc cagcaccgcc ccgcccga	30540
	tgcagcgaag agacccggg tcccggcaat ggcaatggag gaccaccgc tcgtaccctg	30600
	ggatcatctg ggagctgaac aagtctatgt tggcacagca caggcatatg ctcatgcatc	30660

	tcttcagcac tctcaactcc tcgggggtca aaacatatac ccagggcacg gggaactctt	30720
	gcaggacagc gaacccccga gaacagggca atcctcgcac agaacttaca ttgtgcatgg	30780
	acagggtatc gcaatcagge agcaccgggt gatcctccac cagagaagcg cgggtctcgg	30840
	tctcctcaca gcgtggtaag ggggccggcc gatacgggtg atggcgggac gcggctgac	30900
	gtgttcgca ccgtgcatg atgcagttgc tttcggacat tttcgtactt gctgtagcag	30960
	aacctggtcc gggcgtgca caccgatcgc cggcggcggc ctcggcgctt ggaacgctcg	31020
	gtgttgaaat tgtaaaacag ccactctctc agaccgtgca gcagatctag ggctcagga	31080
	gtgatgaaga tcccatcatg cctgatggct ctgatcacat cgaccaccgt ggaatgggcc	31140
	agaccagcc agatgatgca atttgttgg gtttcgggta cggcggggga gggaagaaca	31200
	ggaagaacca tgattaactt ttaatcaaaa cggctctgga gtacttcaaa atgaagatcg	31260
	cggagatgge acctctgcc cccgctgtgt tgggtgaaaa taacagccag gtcaaagggtg	31320
	atacggttct cgagatgttc cacggtggct tccagcaaaag cctccacgag cacatccaga	31380
	aacaagacia tagcgaaagc gggagggttc tctaattcct caatcatcat gttactctc	31440
[0155]	tgaccatcc ccagataatt ttcatttttc cagcctttaa tgattcgaac tagttcgtga	31500
	ggtaaatcca agccagccat gataaagagc tcgcgagag cgccctccac cggcattctt	31560
	aagcacacc tcataattcc aagatattct gctcctggtt cacctgcagc agattgacaa	31620
	gcggaatata aaaatctctg ccgcatccc tgagctctc cctcagcaat aactgtaagt	31680
	actctttcat atctctccg aaatttttag ccataggacc accaggaata agattagggc	31740
	aagccacagt acagataaac cgaagtcctc cccagtgagc attgccaaat gcaagactgc	31800
	tataagcatg ctggctagac ccggtgatat cttccagata actggacaga aaatcgccca	31860
	ggcaattttt aagaaaatca acaaaagaaa aatcctccag gtggacgttt agagcctcgg	31920
	gaacaacgat gaagtaaatg caagcgggtc gttccagcat ggttagttag ctgatctgta	31980
	gaaaaaacia aatgaacat taaaccatgc tagcctggcg aacaggtggg taaatcgttc	32040
	tctccagcac caggcaggcc acgggtctc cggcgcgacc ctcgtaaaaa ttgtcgctat	32100
	gattgaaaac catcacagag agacgttccc ggtggccggc gtgaatgatt cgacaagatg	32160
	aatacaccce cggaacattg gcgtccgca gtgaaaaaaaa gcgcccagg aagcaataag	32220

gcactacaat gctcagtctc aagtccagca aagcgatgcc atgcggatga agcacaaaaat 32280
 tctcaggtgc gtacaaaatg taattactcc cctcctgcac aggcagcaaa gcccccgate 32340
 cctccaggta cacatacaaa gcctcagcgt ccatagctta ccgagcagca gcacacaaca 32400
 ggcgcaagag tcagagaaag gctgagctct aacctgtcca cccgctctct getcaatata 32460
 tagcccagat ctacactgac gtaaaggcca aagtctaaaa ataccgcca aataatcaca 32520
 cacgcccagc acacgcccag aaaccggtga cacactcaaa aaaatacgcg cacttctca 32580
 aacgcccana actgcegtca tttccgggtt cccacgctac gtcacaaaa cacgactttc 32640
 aaattccgtc gaccgttaaa aacgtcacc gccccgccc taacggtcgc ccgtctctca 32700
 gccaatcagc gccccgcatc cccaaattca aacacctcat ttgcatatta acgcgcaaaa 32760
 aaagtttgag gtatattatt gatgatgg 32788

<210> 13

<211> 30684

<212> DNA

[0156] <213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 13

ccatcttcaa taatatacct caaacttttt gtgcgcgtta atatgcaaat gaggcgtttg 60
 aatttgggga ggaagggcgg tgattggtcg agggatgagc gaccgttagg ggcggggcga 120
 gtgacgtttt gatgacgtgg ttgcgaggag gagccagttt gcaagttctc gtgggaaaag 180
 tgacgtcaaa cgaggtgtgg tttgaacacg gaaatactca attttcccgc gctctctgac 240
 aggaaatgag gtgtttctgg gcggatgcaa gtgaaaacgg gccattttcg cgcgaaaact 300
 gaatgaggaa gtgaaaatct gagtaatttc gcgtttatgg caggaggag tatttgccga 360
 gggccgagta gactttgacc gattacgtgg gggtttcgat taccgtgttt ttcacctaaa 420
 tttccgcgta cgggtgtcaaa gtccggtgtt tttacgtagg tgtcagctga tcgccagggt 480
 atttaaacct gcgctctcca gtcaagaggc cactcttgag tgccagcagc aagagttttc 540
 tctccgcgc cgcgagtcag atctacactt tgaagtagg gataacaggg taatgacatt 600

	gattattgac tagttgtaa tagtaatcaa ttacggggtc attagttcat agcccatata	660
	tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc	720
	cccgccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc	780
	attgacgtca atgggtggag tatttacggt aaactgccc a ttggcagta catcaagtgt	840
	atcatatgcc aagtcgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt	900
	atgcccagta catgacctta cgggactttc ctacttgga gtacatctac gtattagtca	960
	tcgetattac catggtgatg cggttttggc agtacaccaa tgggcgtgga tagcggtttg	1020
	actcacgggg atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc	1080
	aaaatcaacg ggactttcca aaatgtcgtataaaccgc cccgttgacg caaatgggcg	1140
	gtaggcgtgt acggtgggag gtctatataa gcagagctcg tttagtgaac cgtcagatcg	1200
	cctggaacgc catccacgct gttttgacct ccatagaaga cagcagatcg gccaccatgg	1260
	tgagcaaggg cgaggagctg ttcaccgggg tgggtgccat cctggtcag ctggacggcg	1320
	acgtaaacgg ccacaagttc agcgtgtccg gcgaggcgga gggcgatgcc acctacggca	1380
[0157]	agctgacct gaagtccatc tgcaccaccg gcaagctgcc cgtgccctgg cccacctcg	1440
	tgaccacct gacctacggc gtgcagtgtc tcagccgcta ccccgaccac atgaagcagc	1500
	acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc atcttcttca	1560
	aggacgacgg caactacaag acccgcgcc aggtgaagtt cgaggcgac acctggtga	1620
	accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg gggcacaagc	1680
	tggagtacaa ctacaacagc cacaacgtct atatcatggc cgacaagcag aagaacggca	1740
	tcaagtgtaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag ctgccgacc	1800
	actaccagca gaacacccc atcggcgacg gccccgtgt gctgcccac aaccactacc	1860
	tgagcaccca gtccgccctg agcaaagacc ccaacgagaa gcgcgatcac atggtcctgc	1920
	tggagttcgt gaccgccgc gggatcactc tcggcatgga cgagctttac aagtagtgag	1980
	tttaaactcc catttaaatg tgagggttaa tgcttcgagc agacatgata agatacattg	2040
	atgagtttgg acaaaccaca actagaatgc agtgaaaaa atgctttatt tgtgaaattt	2100
	gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca	2160

	attgcattca ttttatgttt caggttcagg gggagatgtg ggaggttttt taaagcaagt	2220
	aaaacctcta caaatgtggt aaaataacta taacggctct aaggtagcga gtgagtagtg	2280
	ttctggggcg ggggaggacc tgcatgaggg ccagaataac tgaaatctgt gcttttctgt	2340
	gtgttcagc agcatgagcg gaagcggctc ctttgagga ggggtattca gcccttatct	2400
	gacggggcgt ctccccctct gggcgggagt gcgtcagaat gtgatgggat ccacggtgga	2460
	cggccggccc gtgcagcccg cgaactcttc aaccctgacc tatgcaacc tgagctcttc	2520
	gtcgttgac gcagctgceg ccgcagctgc tgcatctgcc gccagcggcg tgcgcggaat	2580
	ggccatgggc gccggtact acggcactct ggtggccaac tcgagtcca ccaataatcc	2640
	cgccagcctg aacgaggaga agctgttget gctgatggcc cagctcgagg ccttgacca	2700
	gcgcctgggc gagctgacc agcaggtggc tcagctgcag gagcagacgc ggcccgcggt	2760
	tgccacggtg aaatccaaat aaaaaatgaa tcaataaata aacggagacg gttgttgatt	2820
	ttaacacaga gtctgaatct ttatttgatt ttctgcgcgc ggtaggccct ggaccaccgg	2880
[0158]	tctgatcat tgagcaccg gtggatcttt tccaggacce ggtagaggtg ggcttgatg	2940
	ttgaggtaca tgggcatgag cccgtcccgg gggtaggagt agctccattg cagggcctcg	3000
	tgctcggggg tgggtttgta aatcaccag tcatagcagg ggcgcagggc atggtgttgc	3060
	acaatatctt tgaggaggag actgatggcc acgggcagcc ctttggtgta ggtgtttaca	3120
	aatctgttga gctgggagg atgcatcgcg ggggagatga ggtgcatctt ggcttgatc	3180
	ttgagattgg cgatgttacc gccagatcc gcctgggggt tcatgttggt caggaccacc	3240
	agcacggtgt atccggtgca cttggggaat ttatcatgca acttgaagga gaaggcgtga	3300
	aagaatttgg cgacgccttt gtgcccgcc aggttttcca tgcaactcacc catgatgatg	3360
	gcgatgggcc cgtgggcggc ggccctggca aagacgtttc gggggtcgga cacatcatag	3420
	ttgtggtcct gggtaggtc atcataggcc attttaatga atttggggcg gaggggtgcc	3480
	gactggggga caaaggtacc ctcatcccg ggggcgtagt tcccctcaca gatctgcatc	3540
	tcccaggctt tgagctcgga gggggggatc atgtccacct gcggggcgat aaagaacacg	3600
	gtttccgggg cgggggagat gagctgggcc gaaagcaagt tccggagcag ctgggacttg	3660

	ccgcagccgg tggggccgta gatgaccccg atgaccggct gcaggtggta gttgagggag	3720
	agacagctgc cgtcctcccg gaggaggggg gccacctcgt tcatcatctc gcgcacgtgc	3780
	atgttctcgc gcaccagttc cgccaggagg cgctctcccc ccagggatag gagctcctgg	3840
	agcgaggcga agtttttcag cggcttgagt ccgtcggcca tgggcatttt ggagagggtt	3900
	tgttgcaaga gttccaggcg gtcccagagc tcggtgatgt gctctacggc atctcgatcc	3960
	agcagacctc ctcgtttcgc gggttgggac ggctgcggga gtagggcacc agacgatggg	4020
	cgtccagcgc agccagggtc cggtccttcc agggctgcag cgtccgcgtc aggggtgtct	4080
	ccgtcacggt gaaggggtgc gcgccgggct gggcgttgc gaggtgcgc ttcaggctca	4140
	tccggctggt cgaaaaccgc tcccgatcgg cgccctgcgc gtcggccagg tagcaattga	4200
	ccatgagttc gtagttgagc gcctcggccg cgtggccttt ggcgcggagc ttacctttgg	4260
	aagtctgcc gcaggcggga cagaggaggg acttgagggc gtagagcttg gggcgagga	4320
	agacggactc gggggcgtag gcgtccgcgc cgcagtgggc gcagacggtc tcgcactcca	4380
	cgagccaggt gaggtcgggc tggtcggggt caaaaaccag tttcccgcg ttccttttga	4440
[0159]	tgcgtttctt acctttggtc tccatgaget cgtgtccccg ctgggtgaca aagaggctgt	4500
	ccgtgtcccc gtagaccgac tttatgggcc ggtcctcgag cgggtgtccg cggctcctct	4560
	cgtagaggaa ccccgccac tccgagacga aagcccgggt ccaggccagc acgaaggagg	4620
	ccacgtggga cgggtagcgg tcgttgcca ccagcgggtc caccttttcc agggtatgca	4680
	aacacatgtc cccctcgtcc acatccagga aggtgattgg cttgtaagt taggccacgt	4740
	gaccgggggt cccggccggg ggggtataaa aggggtcggg tccctgctcg tcctcactgt	4800
	cttccgcatc gctgtccagg agcgcagct gttgggtag gtattccctc tcgaaggcgg	4860
	gcatgacctc ggcactcagg ttgtcagttt ctagaaacga ggaggatttg atattgacgg	4920
	tgccggcgga gatgcctttc aagagcccct cgtccatctg gtcagaaaag acgatctttt	4980
	tgttgtcgag cttggtggcg aaggagccgt agaggcggtt ggagaggagc ttggcgatgg	5040
	agcgcatggt ctggtttttt tccttctcgg cgcgctcctt ggcggcgatg ttgagctgca	5100
	cgtactcgcg cgccacgcac ttccattcgg ggaagacggt ggtcagctcg tcgggcacga	5160
	ttctgacctg ccagccccga ttatgcaggg tgatgaggtc cacactggtg gccacctcgc	5220

	cgcgagggg ctcattagtc cagcagaggc gtccgccctt gcgcgagcag aaggggggca	5280
	gggggtccag catgacctcg tcgggggggt cggcatcgat ggtgaagatg ccgggcagga	5340
	ggtcggggtc aaagtagctg atggaagtgg ccagatcgtc cagggcagct tgccattcgc	5400
	gcacggccag cgcgcgctcg tagggactga ggggcgtgcc ccagggcatg ggatgggtaa	5460
	gcgcggaggc gtacatgccg cagatgtcgt agacgtagag gggctcctcg aggatgccga	5520
	tgtaggtggg gtagcagcgc cccccgga tgctggcgcg cacgtagtca tacagctcgt	5580
	gcgagggggc gaggagcccc gggcccaggt tggcgcgact gggcttttcg gcgcggtaga	5640
	cgatctggcg gaaaatggca tgcgagttgg aggagatggt gggcctttgg aagatgttga	5700
	agtgggcgtg gggcagtcag accgagtcgc ggatgaagtg ggcgtaggag tcttgcagct	5760
	tggcgacgag ctcggcggtg actaggacgt ccagagcgca gtagtcgagg gtctcctgga	5820
	tgatgtcata cttgagctgt cctttttgtt tccacagctc gcggttgaga aggaactctt	5880
	cgcggtcctt ccagtactct tcgaggggga acccgtcctg atctgcacgg taagagccta	5940
[0160]	gcatgtagaa ctggttgacg gcctttagg cgcagcagcc cttctccacg gggagggcgt	6000
	aggcctgggc ggccttgccg agggaggtgt gcgtgagggc gaaagtgtcc ctgaccatga	6060
	ccttgaggaa ctggtgcttg aagtcgatat cgtcgcagcc cccctgctcc cagagctgga	6120
	agtccgtgcg cttctttag gcggggttgg gcaaagcgaa agtaacatcg ttgaagagga	6180
	tcttccccgc gcggggcata aagttagcag tgatgcggaa aggttggggc acctcgccc	6240
	ggttgtgat gacctgggcg gcgagcacga tctcgtcgaa gccgttgatg ttgtggccca	6300
	cgatgtagag ttccacgaat cgcggacggc ccttgacgtg gggcagtttc ttgagctcct	6360
	cgtaggtgag ctcgtcgggg tcgctgagcc cgtgctgctc gagcggccag tcggcgagat	6420
	gggggttggc gcggaggaag gaagtccaga gatccacggc cagggcggtt tgcagacggt	6480
	cccgtactg acggaactgc tgcccagcgg ccattttttc gggggtgacg cagtagaagg	6540
	tgcgggggtc cccgtgccag cgatcccatt tgagctggag ggcgagatcg agggcgagct	6600
	cgacgagccg gtcgtccccg gagagtcca tgaccagcat gaaggggacg agctgcttgc	6660
	cgaaggaccc catccaggtg taggtttcca catcgtaggt gaggaagagc ctttcggtgc	6720

	gaggatgcga gccgatgggg aagaactgga tctcctgcc acaattggag gaatggctgt	6780
	tgatgtgatg gaagtagaaa tgccgacggc gcgccgaaca ctctgtcttg tgttataca	6840
	agcggccaca gtgctcgsaa cgctgcacgg gatgcacgtg ctgcacgagc tgtacctgag	6900
	ttcctttgac gaggaatttc agtgggaagt ggagtcgtgg cgcctgcatc tcgtgctgta	6960
	ctacgtctg gtggtcggcc tggccctctt ctgcctcgat ggtggtcatt ctgacgagcc	7020
	cgcgcgggag gcaggtccag acctcggcgc gagcgggtcg gagagcgagg acgagggcgc	7080
	gcagggcggg gctgtccagg gtctctgagac gctgcggagt caggtcagtg ggcagcggcg	7140
	gcgcgcggtt gacttgcagg agtttttcca gggcgcgcgg gaggtccaga tggacttga	7200
	tctccaccgc gccattggtg gcgacgtcga tggcttgtag ggtcccgtgc ccttggggtg	7260
	tgaccaccgt cccccgttct ttcttgggcg gctggggcga cgggggctgg gcctcttcca	7320
	tggttagaag cggcggcgag gacgcgcgcc gggcggcagg ggcggctcgg ggcccggagg	7380
	cagggcgggc aggggcacgt cggcgcgcgc cgcgggtagg ttctggtact gcgccggag	7440
	aagactggcg tgagcgacga cgcgacggtt gacgtcctgg atctgacgc tctgggtgaa	7500
[0161]	ggccacggga cccgtgagtt tgaacctgaa agagagtctg acagaatcaa tctcggtatc	7560
	gttgacggcg gcctgccgca ggatctcttg cacgtcgccc gagttgtcct ggtagcgat	7620
	ctcggtcatt aactgctcga tctctctctc ttgaaggctt ccgcggccgg cgcgctccac	7680
	ggtggccgcg aggtcgttgg agatgcggcc catgagctgc gagaaggcgt tcatgccgc	7740
	ctcgttccag acgcggctgt agaccacgac gccctcggga tcgcgggcgc gcatgaccac	7800
	ctgggcgagg ttgagctcca cgtggcgcgt gaagaccgcg tagttgcaga ggcgctgta	7860
	gaggtagttg agcgtggtgg cgatgtgctc ggtgacgaag aaatacatga tccagcggcg	7920
	gagcggcacc tcgctgacgt cgcaccgcgc ctccaaactg tccatggcct cgtaaaagtc	7980
	cacggcgaag ttgaaaaact gggagttgcg cgcgagacg gtcaactcct cctccagaag	8040
	acggatgagc tcggcgatgg tggcgcgcac ctgcgctcg aaggcccccg ggagttcctc	8100
	cacttctct tcttctctct ccaactaacat ctcttctact tctctctcag gcggcagtgg	8160
	tggcggggga gggggcctgc gtcgccggcg gcgcacgggc agacggtcga tgaagcctc	8220
	gatggtctcg ccgcgccggc gtcgcatggt ctcggtgacg gcgcgccctg cctcgcgggg	8280

	ccgcagcgtg aagacgccgc cgcgcatctc caggtggccg ggggggtccc cgttgggcag	8340
	ggagagggcg ctgacgatgc atcttatcaa ttgccccgta gggactccgc gcaaggacct	8400
	gagcgtctcg agatccacgg gatctgaaaa ccgctgaacg aaggcttcga gccagtcgca	8460
	gtcgaaggt aggctgagca cggtttcttc tggcgggtca tgttggttg gagcggggcg	8520
	ggcgatgctg ctggtgatga agttgaaata ggcggttctg agacggcgga tggtagcgag	8580
	gagcaccagg tctttgggcc cggcttctg gatgcgcaga cggtcggcca tgccccaggc	8640
	gtggtcctga cacctggcca ggtcettgta gtagtctgc atgagccgct ccacgggcac	8700
	ctctcctcg cccgcgcggc cgtgcatgcg cgtgagccc aagcccgct ggggctggac	8760
	gagcaccagg tcggcgacga cgcgctcggc gaggatggt tgcctgatct gggtaggggt	8820
	ggtctggaag tcatcaaagt cgacgaagcg gtggtaggct ccggtgtga tggtagga	8880
	gcagttggcc atgacggacc agttgacggt ctggtggccc ggacgcacga gctcgtgta	8940
	cttagggcg gagtaggcgc gcgtgtcgaa gatgtagtgc ttgcaggtgc gcaccaggta	9000
[0162]	ctgtagccg atgaggaagt gcggcggcgg ctggcggtag agcggccatc gctcgggtgc	9060
	ggggcgccg ggcgcgaggt cctcagcat ggtgcggtg tagccgtaga tgtacctgga	9120
	catccaggtg atgccggcgg cggtagtga ggccgcggg aactcgcga cgcggtcca	9180
	gatgttgcgc agcggcagga agtagtcat ggtgggcac gtctggccc tgaggcgcgc	9240
	gcagtcgtg atgctctata cgggcaaaaa cgaaagcgg cagcggctc actccgtggc	9300
	ctggaggcta agcgaacggg ttggctgcg cgtgtacccc ggttcgaatc tcgaatcagg	9360
	ctggagccgc agctaactg gtattggcac tcccgtctc acccaagcct gcaccaacc	9420
	tccaggatac ggaggcgggt cgttttcaa ctttttttg gaggccgat gagactagta	9480
	agcgcgaaa gcggccgacc gcgatggctc gctgccgtag tctggagaag aatcgccagg	9540
	gttgcttgc ggtgtgcccc ggttcgagc cgcccgatt ccgcgctaa cgagggcgtg	9600
	gctgccccgt cgtttccaag accccatagc cagccgactt ctccagttac ggagcgagcc	9660
	cctctttgt tttgtttgt tttgccagat gcatcccgta ctgcggcaga tgcccccca	9720
	ccaccctcca ccgcaacaac agccccctc acagccggcg cttctgcccc cgccccagca	9780

	gcaacttcca gccacgaccg ccgcgccgc cgtgagcggg gctggacaga gttatgatca	9840
	ccagctggcc ttggaagagg gcgaggggct ggcgcgctg gggcgctcgt cgccggagcg	9900
	gcacccgcgc gtgcagatga aaagggacgc tcgagaggcc tacgtgcccga agcagaacct	9960
	gttcagagac aggagcggcg aggagcccga ggagatgcgc gcggcccggg tccacgcggg	10020
	gcgggagctg cggcgcggcc tggaccgaaa gaggggtgctg agggacgagg atttcgaggc	10080
	ggacgagctg acggggatca gccccgcgc cgcgcacgtg gccgcggcca acctggtcac	10140
	ggcgtacgag cagaccgtga aggaggagag caacttccaa aaatccttca acaaccacgt	10200
	gcgcaccctg atcgcgcgcg aggaggtgac cctgggctctg atgcacctgt gggacctgct	10260
	ggaggccatc gtgcagaacc ccaccagcaa gccgctgacg gcgcagctgt tcctggtggt	10320
	gcagcatagt cgggacaacg aagcgttcag ggagcgcctg ctgaatatca ccgagcccga	10380
	gggccgctgg ctcttgacc tggatgaacat tctgcagagc atcgtggtgc aggagcgcgg	10440
	gctgccgctg tccgagaagc tggcgccat caacttctcg gtgctgagtt tgggcaagta	10500
	ctacgctagg aagatctaca agacccgta cgtgcccata gacaaggagg tgaagatcga	10560
[0163]	cgggttttac atgcgcatga ccctgaaagt gctgaccctg agcgacgacg tgggggtgta	10620
	ccgcaacgac aggatgcacc gtgcggtgag cgccagcagg cggcgcgagc tgagcgacca	10680
	ggagctgatg catagtctgc agcgggccc gaccggggcc gggaccgagg gggagagcta	10740
	ctttgacatg ggcgcgacc tgcactggca gccagccgc cgggccttgg aggcggcggc	10800
	aggaccctac gtagaagagg tggacgatga ggtggacgag gagggcgagt acctggaaga	10860
	ctgatggcgc gaccgtattt ttgctagatg caacaacaac agccacctcc tgatcccgcg	10920
	atcgggcgg cgctgcagag ccagccgtcc ggcattaact cctcggacga ttgaccag	10980
	gccatgcaac gcatcatggc gctgacgacc cgcaaccccg aagcctttag acagcagccc	11040
	caggccaacc ggctctcggc catcctggag gccgtggtgc cctcgcgctc caaccccacg	11100
	cacgagaagg tcctggccat cgtgaacgcg ctggtggaga acaaggccat ccgcgcgac	11160
	gaggccggcc tgggtgataa cgcgctgctg gagcgcgtgg cccgctataa cagcaccaac	11220
	gtgcagacca acctggaccg catggtgacc gacgtgcgcg aggcctggc ccagcgcgag	11280
	cggttcacc gcgagtccaa cctgggatcc atggtggcgc tgaacgcctt cctcagcacc	11340

	cagcccccca acgtgccccg gggccaggag gactacacca acttcatcag cgccctgcgc	11400
	ctgatggtga ccgaggtgcc ccagagcgag gtgtaccagt ccgggccgga ctacttette	11460
	cagaccagtc gccagggcctt gcagaccgtg aacctgagcc aggctttcaa gaacttgacg	11520
	ggcctgtggg gcgtgcaggc cccggtcggg gaccgcgca cggtgtcag cctgctgacg	11580
	ccgaactcgc gcctgctgct gctgctggtg gcccccttca cggacagcgg cagcatcaac	11640
	cgcaactcgt acctgggcta cctgattaac ctgtaccgag aggccatcgg ccaggcgcac	11700
	gtggacgagc agacctacca ggagatcacc cacgtgagcc gcgccctggg ccaggacgac	11760
	ccgggcaacc tggagccac cctgaacttt ttgctgacca accggtcga gaagatcccg	11820
	ccccagtacg cgctcagcac cgaggaggag cgcacctcgc gttacgtgca gcagagcgtg	11880
	ggcctgttcc tgatgcagga gggggccacc cccagcgccg cgctcgacat gaccgcgagc	11940
	aacatggagc ccagcatgta cgccagcaac cgcccgttca tcaataaact gatggactac	12000
	ttgcatcggg cggccgcat gaactctgac tatttcacca acgccatcct gaatccccac	12060
[0164]	tggtccccgc cgccggggtt ctacacgggc gactacgaca tgcccgaccc caatgacggg	12120
	ttctgtggg acgatgtgga cagcagcgtg ttctccccc gaccgggtgc taacgagcgc	12180
	cccttgtgga agaaggaagg cagcagccga cgcccgtcct cggcgctgtc cggccgcgag	12240
	ggtgctgccc cggcggtgcc cgaggccgcc agtcctttcc cgagcttgcc cttctcgtg	12300
	aacagtatcc gcagcagcga gctgggcagg atcacgcgcc cgcgcttgct gggcgaagag	12360
	gagtacttga atgactcget gttgagacc gagcgggaga agaacttccc caataacggg	12420
	atagaaagcc tgggtgacaa gatgagccgc tggaagacgt atgcgagga gcacaggac	12480
	gatccccggg cgctgcaggg ggccacgagc cggggcagcg ccgcccgtaa acgccggtgg	12540
	cacgacagc agcggggaca gatgtgggac gatgaggact ccgccacga cagcagcgtg	12600
	ttgacttgg gtgggagtgg taaccgcttc gctcacctgc gccccgtat cgggcgcatg	12660
	atgiaagaga aaccgaaaat aatgatact caccaaggcc atggcgacca gcgtgcgttc	12720
	gtttcttctc tgttgttgtt gtatctagta tgatgagcgc tgcgtaccg gagggctctc	12780
	ctccctcgta cgagagcgtg atgcagcagg cgatggcggc ggccggcagc cagccccgc	12840

	tggaggctcc ttacgtgcc cgcggtacc tggcgcctac ggaggggCGG aacagcattc	12900
	gttactcgga gctggcacc ttgtacgata ccaccgggt gtacctggtg gacaacaagt	12960
	cgccggacat cgcctcgtg aactaccaga acgaccacag caacttcctg accaccgtgg	13020
	tgcagaacaa tgacttcacc cccacggagg ccagcaccca gaccatcaac tttgacgagc	13080
	gctcgcggtg gggcggccag ctgaaaacca tcatgcacac caacatgccc aacgtgaacg	13140
	agttcatgta cagcaacaag ttcaaggcgc gggatgatgt ctcccgaag acccccaatg	13200
	gggtgacagt gacagaggat tatgatggtg gtcaggatga gctgaagtat gaatgggtgg	13260
	aatttgagct gcccgaaggc aacttctcgg tgaccatgac catcgacctg atgaacaacg	13320
	ccatcatcga caattacttg gcggtggggc ggcagaacgg ggtgctggag agcgacatcg	13380
	gcgtgaagtt cgacactagg aacttcaggc tgggctggga ccccgtagc gagctggtca	13440
	tgcccggggt gtacaccaac gaggetttcc atcccgatat tgtcttctg cccgctgcg	13500
	gggtggactt caccgagagc cgcctcagca acctgctggg cattcgcaag aggcagccct	13560
	tccaggaagg cttccagatc atgtacgagg atctggagg gggcaacatc cccgctc	13620
[0165]	tggatgtcga cgcctatgag aaaagcaagg aggatgcagc agctgaagca actgcagccg	13680
	tagctaccgc ctctaccgag gtcagggcg ataattttgc aagcggcga gcagtggcag	13740
	cggccgaggc ggctgaaacc gaaagtaaga tagtcattca gccggtggag aaggatagca	13800
	agaacaggag ctacaacgta ctaccggaca agataaacac cgcctaccgc agctgttacc	13860
	tagcctacaa ctatggcgac cccgagaagg gcgtgcgctc ctggacgtg ctcaccact	13920
	cggacgtcac ctgcggcgtg gagcaagtct actggtcgt gcccgacatg atgcaagacc	13980
	cggtcacctt ccgctccacg cgtcaagtta gcaactacc ggtggtggc gccgagctc	14040
	tgcccgtcta ctccaagagc ttcttcaacg agcaggcgt ctactcagc cagctgcg	14100
	ctttcacctc gcttacgca gtcttcaacc gcttccccga gaaccagatc ctctccgcc	14160
	cgcccgccc caccattacc accgtcagtg aaaacgttcc tgctctcaca gatcacggga	14220
	ccctgccgt gcgcagcagt atccggggag tccagcgtg gaccgttact gacccagac	14280
	gccgcacctg cccctacgtc tacaaggccc tgggcatagt cgcgccgcg gtctctcga	14340
	gccgcacct ctaaagtcc attctcatct cgcccagtaa taacaccggt tggggcctgc	14400

	gcgcgcccag caagatgtac ggaggcgctc gccaacgctc cagcaaacac cccgtgcgcg	14460
	tgcgcgggca cttccgcgct cectggggcg cctcaaggg ccgcgtgcgg tcgcgcacca	14520
	ccgtcgacga cgtgatcgac caggtagtg gacgcgcg caactacacc cccgccgccc	14580
	cgcccgtctc caccgtggac gccgtcatcg acagcgtggt ggccgacgcg cgccggtacg	14640
	cccgcgcaa gagccggcgg cggcgcatcg cccggcggca ccggagcacc cccgccatgc	14700
	gcgcggcgcg agccttgctg cgcagggcca ggcgcacggg acgcagggcc atgctcaggg	14760
	cgccagacg cgcggcttca ggcgccagcg ccggcaggac ccggagacgc gcggccacgg	14820
	cggcggcagc ggccatcgcc agcatgtccc gcccgcgcg agggaacctg tactgggtgc	14880
	gcgacgccgc caccggtgtg cgcgtgcccg tgcgcacccg cccccctgc acttgaagat	14940
	gttcaacttc cgatgttgat gtgtcccagc ggcgaggagg atgtccaagc gcaaattcaa	15000
	ggaagagatg ctccaggtca tcgcgcctga gatctacggc cctgcggtgg tgaaggagga	15060
	aagaaagccc cgcaaatca agcgggtcaa aaaggacaaa aaggaagaag aaagtgatgt	15120
[0166]	ggacggattg gtggagttg tgcgcgagtt cgcccccg cggcgcgtgc agtggcgcgg	15180
	gcggaaggtg caaccggtgc tgagaccg caccaccgtg gtcttcacgc ccggcgagcg	15240
	ctccggcacc gttccaagc gctcctacga cgaggtgtac ggggatgatg atattctgga	15300
	gcagcggcc gagegcctgg gcgagttgc ttacggcaag cgcagccgtt ccgcaccgaa	15360
	ggaagaggcg gtgtccatcc cgctggacca cggcaacccc acgccgagcc tcaagcccgt	15420
	gacctgcag caggtgctgc cgaccgcgg gccgcgccgg gggttcaagc gcgaggcga	15480
	ggatctgtac cccaccatgc agctgatggt gcccaagcgc cagaagctgg aagacgtgct	15540
	ggagaccatg aaggtggacc cggacgtgca gcccgaggtc aaggtgcggc ccatcaagca	15600
	ggtggccccg ggctgggcg tcagaccgt ggacatcaag attcccacgg agcccatgga	15660
	aacgcagacc gagcccatga tcaagcccag caccagcacc atggaggtgc agacggatcc	15720
	ctggatgcca tcggctccta gtcgaagacc ccggcgcaag tacggcgcg ccagcctgct	15780
	gatgccaac tacgcgctgc atccttccat cateccccag ccgggctacc gcggcacgcg	15840
	cttctaccgc ggtcatacca gcagccgccc ccgcaagacc accactcgcc gccgccgtcg	15900

	ccgcaccgcc gctgcaacca cccctgccgc cctgggtcgg agagtgtacc gccgcggccg	15960
	cgcacctctg accctgccgc gcgcgcgcta ccacccgagc atcgccattt aaactttcgc	16020
	ctgetttgca gatcaatggc cctcacatgc cgccttcgcg ttcccattac gggctaccga	16080
	ggaagaaaac cgcgccgtag aaggctggcg gggaacggga tgcgtcgcca ccaccaccgg	16140
	cggcggcgcg ccatcagcaa gcggttgggg ggaggttcc tccccgcgt gatccccatc	16200
	atcgccgcgg ccatcggggc gatccccggc attgcttccg tggcgggtgca ggcctctcag	16260
	cgccactgag acacacttgg aaacatcttg taataaacca atggactctg acgtccttgg	16320
	tcctgtgatg tgttttcgta gacagatgga agacatcaat ttttcgtccc tggctccgcg	16380
	acacggcacg cggccgttca tgggcacctg gagcgacatc ggcaccagcc aactgaacgg	16440
	gggcgccttc aattggagca gtctctggag cgggcttaag aatttcgggt ccacgcttaa	16500
	aacctatggc agcaaggcgt ggaacagcac cacagggcag gcgctgaggg ataagctgaa	16560
	agagcagaac ttccagcaga aggtggtcga tgggctcgcc tcgggcatca acggggtggt	16620
	ggacctggcc aaccaggccg tgcagcggca gatcaacagc cgcttgacc cggtgccgcc	16680
[0167]	cgccggctcc gtggagatgc cgcaggtgga ggaggagctg cctcccctgg acaagcgggg	16740
	cgagaagcga ccccccccg atgcggagga gacgctgctg acgcacacgg acgagccgcc	16800
	cccgtacgag gaggcggtga aactgggtct gccaccacg cggcccatcg cgcccctggc	16860
	caccggggtg ctgaaaccg aaaagcccgc gaccctggac ttgcctctc cccagccttc	16920
	cgcccctct acagtggcta agcccctgcc gccggtggcc gtggcccgcg cgcgaccgg	16980
	gggcaccgcc cgccctcatg cgaactggca gagcactctg aacagcatcg tgggtctggg	17040
	agtgcagagt gtgaagcgc gccgctgcta ttaaacctac cgtagcgtt aacttgctt	17100
	tctgtgtgtg tatgtattat gtcgcccg cgctgtcca ccagaaggag gagtgaagag	17160
	gcgctcgcc gagttgcaag atggccaccc catcgatgct gcccagtg gctacatgc	17220
	acatcgccgg acaggacgct tcggagtacc tgagtccggg tctggtgcag ttgcccgcg	17280
	ccacagacac ctactcagt ctggggaaca agtttaggaa cccacggtg gcgcccacg	17340
	acgatgtgac caccgaccgc agccagcggc tgacgtgcg cttcgtgccc gtggaccgcg	17400
	aggacaacac ctactcgtac aaagtgcgct acacgtggc cgtgggcgac aaccgcgtgc	17460

	tggacatggc cagcacctac tttgacatcc gcggcgtgct ggatcggggc ctagcttca	17520
	aaccctactc eggcaccgcc tacaacagtc tggccccaa gggagcacc aacacttgtc	17580
	agtggacata taaagccgat ggtgaaactg ccacagaaaa aacctataca tatggaaatg	17640
	caccctgca gggcattaac atcacaaaag atggtattca acttggaaact gacaccgatg	17700
	atcagccaat ctacgcagat aaaacctatc agcctgaacc tcaagtgggt gatgctgaat	17760
	ggcatgacat cactggtact gatgaaaagt atggaggcag agctcttaag cctgatacca	17820
	aatgaagcc ttgttatggt tcttttgcca agcctactaa taaagaagga ggtcaggcaa	17880
	atgtgaaaac aggaacaggc actactaaag aatatgacat agacatggct ttctttgaca	17940
	acagaagtgc ggctgctgct ggcctagctc cagaaattgt tttgtatact gaaaatgtgg	18000
	atttggaaac tccagatacc catattgtat acaaagcagg cacagatgac agcagctctt	18060
	ctattaattht gggtcagcaa gccatgccc acagacctaa ctacattggt ttcagagaca	18120
	actttatcgg gctcatgtac tacaacagca ctggcaatat gggggtgctg gccggtcagg	18180
[0168]	cttctcagct gaatgctgtg gttgacttgc aagacagaaa caccgagctg tctaccage	18240
	tcttgcttga ctctctgggt gacagaacct ggtatttcag tatgtggaat caggcgggtg	18300
	acagctatga tctgatgtg cgcattattg aaaatcatgg tgtggaggat gaacttcca	18360
	actattgttt ccctctggat gctgttgca gaacagatac ttatcaggga attaaggcta	18420
	atggaactga tcaaaccaca tggaccaaag atgacagtgt caatgatgct aatgagatag	18480
	gcaagggtaa tccattcgcc atggaatca acatccaage caacctgtgg aggaacttcc	18540
	tctacgcaa cgtggccctg tacctgccg actcttaca gtacacgcc gccaatgtta	18600
	ccctgcccac caacaccaac acctacgatt acatgaacgg ccgggtggtg gcgccctcgc	18660
	tgggtgactc ctacatcaac atcggggcgc gctggtcgtt ggatcccatg gacaacgtga	18720
	acccttcaa ccaccaccgc aatgcggggc tgcgtaccg ctccatgctc ctgggcaacg	18780
	ggcgtacgt gcccttccac atccaggtgc ccagaaatt tttgccatc aagagcctcc	18840
	tgtctctgcc cgggtctac acctacgagt ggaacttccg caaggacgct aacatgatcc	18900
	tgcagagctc cctcggcaac gacctgcga cggacggggc ctccatctcc ttcaccagca	18960

tcaacctcta cgccacette ttccccatgg cgcacaacac ggcctccacg ctcgaggcca	19020
tgctgcgcaa cgacaccaac gaccagtctt tcaacgacta cctctcggcg gccaacatgc	19080
tctaccccat cccggccaac gccaccaacg tgcccatctc catccctcgc cgcaactggg	19140
ccgccttccg cggtctgttc ttcacgcgtc tcaagaccaa ggagacgccc tcgctgggct	19200
ccgggttcga cccctacttc gtctactcgg gctccatccc ctacctgac ggcaccttct	19260
acctcaacca caccttcaag aaggtctcca tcaccttoga ctctctcgtc agctggcccc	19320
gcaacgaccg gctcctgacg cccaacgagt tcgaaatcaa gcgcaccgtc gacggcgagg	19380
gctacaacgt ggcccagtcg aacatgacca aggactggtt cctggtccag atgctggccc	19440
actacaacat cggctaccag ggcttctacg tgcccagggg ctacaaggac cgcatgtact	19500
ccttcttccg caacttccag cccatgagcc gccaggtggt ggacgaggtc aactacaagg	19560
actaccagge cgtcacctg gcctaccagc acaacaacte gggcttcgtc ggctacctgc	19620
cgcccacat gcgccagggc cagccctacc ccgccaacta ccctaccgc ctcatcggca	19680
agagcgccgt caccagcgtc acccagaaaa agttctctg cgacagggtc atgtggcgca	19740
[0169] tccccttctc cagcaacttc atgtccatgg gcgcgtcac cgacctcggc cagaacatgc	19800
tctatgcaa ctccgccac gcgctagaca tgaatttoga agtcgacccc atggatgagt	19860
ccacccttct ctatgttgtc ttcgaagtct tcgacgtcgt ccgagtgcac cagccccacc	19920
gcggcgtcat cgaggccgtc tacttgcgca ccccttctc ggccggtaac gccaccacct	19980
aagctcttgc ttcttgcaag ccatggccgc gggctccggc gagcaggagc tcagggccat	20040
catccgcgac ctgggctcgc ggccctactt cctgggcacc ttcgataagc gcttcccggg	20100
attcatggcc ccgcacaagc tggcctgcgc catcgtcaac acggccggcc gcgagaccgg	20160
gggcgagcac tggttggect tcgcttgaa cccgcgtctg aacacctgct acctcttga	20220
ccccctcggg ttctcggacg agcgcctcaa gcagatctac cagttcgagt acgaggcct	20280
gctgcgccgc agcgccttg ccaccagga ccgctgcgtc accctggaaa agtccacca	20340
gaccgtgcag ggtccgcgt cggccgcctg cgggctcttc tgctgcatgt tctgcacgc	20400
cttctgacac tggccccgacc gccccatgga caagaacccc accatgaact tgctgacggg	20460
ggtgcccac ggcctgctcc agtcgcccc ggtggaaccc accctgcgcc gcaaccagga	20520

	ggcgctctac cgcttctca actcccactc cgctacttt cgctcccacc gcgcgcgat	20580
	cgagaaggcc accgccttcg accgcatgaa tcaagacatg taaaccgtgt gtgtatgtta	20640
	aatgtcttta ataaacagca ctttcatggt acacatgcat ctgagatgat ttatttagaa	20700
	atcgaaaggg ttctgccggg tctcggcatg gcccgcgggc agggacacgt tgcggaactg	20760
	gtacttggcc agccacttga actcggggat cagcagtttg ggcagcgggg tgcggggaa	20820
	ggagtcggtc cacagcttcc gcgtcagttg cagggcgccc agcaggtcgg gcgcggagat	20880
	cttgaatcg cagttgggac ccgcgttctg cgcgcgggag ttgcggtaca cggggttga	20940
	gcactggaac accatcaggg ccgggtgctt cacgctcgcc agcacctcg cgtcggatg	21000
	gctctccacg tcgaggtcct cggegttggc catcccgaag ggggtcatct tgcaggtctg	21060
	ccttcccatg gtgggcacgc acccgggctt gtggttgcaa tcgtagtca ggggatcag	21120
	catcatctgg gcctggtcgg cgttcatccc cgggtacatg gccttcatga aagcctcaa	21180
	ttgcctgaac gcctgctggg ccttggctcc ctcggtgaag aagacccgc aggacttgc	21240
[0170]	agagaactgg ttggtggcg acccggcgtc gtgcacgcag cagcgcgct cgttgttggc	21300
	cagctgcacc acgctgcgc cccagcgggt ctgggtgatc ttggcccggc cggggttctc	21360
	cttcagcgcg cgctgccggt tctcgtcgc cacatccatc tcgatcatgt gtccttctg	21420
	gatcatggtg gtcccgtgca ggcaccgcag cttgccctcg gcctcgggtc acccgtgcag	21480
	ccacagcgcg caccgggtgc actcccagtt cttgtggcg atctgggaat gcgcgtgcac	21540
	gaagccctgc aggaagcggc ccatcatggt ggtcagggtc ttgttctag tgaaggtcag	21600
	cggaatgccg cgggtctcct cgttgatgta caggtggcag atcggcggt acacctgcc	21660
	ctgctcgggc atcagctgga agttggcttt caggtcggtc tccacgcggt agcggctcat	21720
	cagcatagtc atgatttcca tacccttctc ccaggccgag acgatgggca ggctcatagg	21780
	gttcttcacc atcatcttag cgctagcagc cgcggccagg gggctcgtct cgtccagggt	21840
	ctcaaagctc cgcttgccgt ccttctcggg gatccgcacc ggggggtagc tgaagcccac	21900
	ggccgccagc tcctctcgg cctgtcttcc gtctcgtcgc tcctggctga cgtcctgcag	21960
	gaccacatgc ttggtcttgc ggggtttctt cttggcggc agcggcgcg gagatgttgg	22020

	agatggcgag ggggagcgcg agttctcgct caccactact atctcttctt cttcttggtc	22080
	cgaggccacg cggcggtagg tatgtctctt cgggggcaga ggcggaggcg acgggctctc	22140
	gccgccgca cttggcggat ggctggcaga gcccttccg cgttcggggg tgcgctcccg	22200
	gcggcgctct gactgacttc ctccgcggcc ggccattgtg ttctcctagg gaggaacaac	22260
	aagcatggag actcagccat cgccaacctc gccatctgcc cccaccgccc acgagaagca	22320
	gcagcagcag aatgaaagct taaccgcccc gccgcccagc cccgccacct ccgacgcggc	22380
	cgtcccagac atgcaagaga tggaggaatc catcgagatt gacctgggct atgtgacgcc	22440
	cgcgagcac gaggaggagc tggcagtgcg cttttcacia gaagagatac accaagaaca	22500
	gccagagcag gaagcagaga atgagcagag tcaggctggg ctcgagcatg acggcgacta	22560
	cctccacctg agcggggggg aggacgcgct catcaagcat ctggcccggc aggccacct	22620
	cgtaaggat gcgctgctcg accgcaccga ggtgcccctc agcgtggagg agctcagccg	22680
	cgctacgag ttgaacctct tctcgcgcg cgtgcccc aagcgcagc ccaatggcac	22740
	ctgcgagccc aaccgcgccc tcaacttcta cccggtcttc gcggtgcccg aggccctggc	22800
[0171]	cacctaccac atctttttca agaaccaaaa gatccccgtc tctgcccgcg ccaaccgcac	22860
	ccgcgccgac gcccttttca acctgggtcc cggcgcccgc ctacctgata tgcctctctt	22920
	ggaagagggt cccaagatct tcgagggtct gggcagcgac gagactcggg ccgcgaacgc	22980
	tctgcaagga gaaggaggag agcatgagca ccacagcgcc ctggctcagt tggaaggcga	23040
	caacgcgagg ctggcgggtc tcaaacgcac ggtcagctg acccatttcg cctaccggc	23100
	tctgaacctg cccccaaag tcatgagcgc ggtcatggac caggtgetca tcaagcgcgc	23160
	gtcggccatc tccgaggacg agggcatgca agactccgag gagggcaagc ccgtggtcag	23220
	cgacgagcag ctggcccggg ggctgggtcc taatgetagt cccagagtt tggaagagcg	23280
	gcgcaaacctc atgatggcgg tggctctggt gaccgtggag ctggagtgcc tgcgccgctt	23340
	cttcgccgac gcggagacct tgcgcaaggt cgaggagaac ctgcactacc tcttcaggca	23400
	cgggttcgtg cggcaggcct gcaagatctc caacgtggag ctgaccaacc tggctctcta	23460
	catgggcate ttgcacgaga accgcctggg gcagaacgtg ctgcacacca ccttgcgagg	23520
	ggaggcccgg cgcgactaca tccgcgactg cgtctacctc tacctctgcc acacctggca	23580

	gacgggcatg ggcgtgtggc agcagtgtct ggaggagcag aacctgaaag agctctgcaa	23640
	gctcctgcag aagaacctca aggtctgtg gaccgggttc gacgagcgca ccaccgcctc	23700
	ggacctggcc gacctcattt tccccgagcg cctcaggctg acgctgcgca acggcctgcc	23760
	cgactttatg agccaaagca tgttgcaaaa ctctcgtctt ttcctcctcg aacgctccgg	23820
	aatcctgccc gccacctgct ccgcctgccc ctccgacttc gtgccctga cttcccgca	23880
	gtgccccccg ccgctgtgga gccactgcta cctgctgcgc ctggccaact acctggccta	23940
	ccactcggac gtgatcgagg acgtcagcgg cgagggcctg ctcgagtgcc actgccgctg	24000
	caacctctgc acgccgcacc gctccctggc ctgcaacccc cagctgctga gcgagacca	24060
	gatcatcggc accttcgagt tgcaagggcc cagcgaagge gagggttcag ccgccaaggg	24120
	gggtctgaaa ctacccccg ggctgtggac ctccgcctac ttgcgcaagt tcgtgcccg	24180
	ggactacat cccttcgaga tcaggttcta cgaggaccaa tccatccgc ccaaggccga	24240
	gctgtcggcc tgcgtcatca cccagggggc gatcctggcc caattgcaag ccatccagaa	24300
[0172]	atcccgccaa gaattcttgc tgaaaaagg ccgcggggtc tacctcgacc cccagaccgg	24360
	tgaggagctc aacccccgct tccccagga tgccccgagg aaacaagaag ctgaaagtgg	24420
	agctgccgcc cgtggaggat ttggaggaag actgggagaa cagcagtcag gcagaggagg	24480
	aggagatgga ggaagactgg gacagcactc aggcagagga ggacagcctg caagacagtc	24540
	tggaggaaga cgaggaggag gcagaggagg aggtggaaga agcagccgcc gccagaccgt	24600
	cgctctcgcc gggggagaaa gcaagcagca cggataccat ctccgctccg ggtcggggtc	24660
	ccgctcgacc acacagtaga tgggacgaga ccggacgatt cccgaacccc accaccaga	24720
	ccgtaagaa ggagcggcag ggatacaagt cctggcgggg gcacaaaaac gccatcgtct	24780
	cctgcttgca ggctgcggg ggcaacatct ccttcacccg gcgctacctg ctcttcacc	24840
	gcggggtgaa ctttccccgc aacatcttgc attactaccg tcacctccac agcccctact	24900
	acttccaaga agaggcagca gcagcagaaa aagaccagca gaaaaccagc agctagaaaa	24960
	tccacagcgg cggcagcagg tggactgagg atcgcggcga acgagccggc gcaaaccgg	25020
	gagctgagga accggatctt tcccaccctc tatgcatctt tccagcagag tcgggggcag	25080

	gagcaggaac tgaagtgca gaaccgttct ctgcgctcgc tcaccgcag ttgtctgtat	25140
	cacaagagcg aagaccaact tcagcgcaact ctcgaggacg ccgaggctct cttcaacaag	25200
	tactgcgcgc teactcttaa agagtagccc gcgccccccc agtcgcagaa aaaggcggga	25260
	attacgtcac ctgtgccctt cgccctagcc gcctccaccc atcatcatga gcaaagagat	25320
	tcccacgcct tacatgtgga gctaccagcc ccagatgggc ctggccgccc gtgccgcca	25380
	ggactactcc acccgcata attggctcag cgccgggccc gcgatgatct cacgggtgaa	25440
	tgacatccgc gcccaccgaa accagatact cctagaacag tcagcgctca cggccacgcc	25500
	ccgcaatcac ctcaatccgc gtaattggcc cgccgcctg gtgtaccagg aaattcccca	25560
	gcccacgacc gtactacttc cgcgagacgc ccaggccgaa gtccagctga ctaactcagg	25620
	tgtccagctg gggggcggcg ccaccctgtg tcgtaccgc cccgctcagg gtataaagcg	25680
	gctggtgatc cggggcagag gcacacagct caacgacgag gtggtgagct cttcgctggg	25740
	tctgcgacct gacggagtct tccaactcgc cggatcgggg agatcttct tcacgcctcg	25800
	tcaggccgtc ctgactttgg agagtctgtc ctcgcagccc cgctcgggtg gcacccggac	25860
[0173]	tctccagttc gtggaggagt tcaactccctc ggtctacttc aacccttct cggctccccc	25920
	cggccactac ccggacgagt tcatcccga cttcgacgcc atcagcgagt cgggtgacgg	25980
	ctacgattga aactaatcac ccccttatcc agtgaaataa agatcatatt gatgatgatt	26040
	ttacagaaat aaaaaataat catttgattt gaaataaaga tacaatcata ttgatgattt	26100
	gagttaaca aaaaaataaa gaatcactta cttgaaatct gataccaggc ctctgtccat	26160
	gttttctgcc aacaccactt cactcccctc ttcccagctc tggctactgca ggccccggcg	26220
	ggctgcaaac ttctccaca cgctgaaggg gatgtcaaat tcctcctgtc cctcaatctt	26280
	cattttatct tctatcagat gtccaaaaag cgcgtccggg tggatgatga cttcgacccc	26340
	gtctaccctt acgatgcaga caacgcaccg accgtgccct tcatcaacce ccccttcgtc	26400
	tcttcagatg gattccaaga gaagcccctg ggggtgttgt ccctgcgact ggccgacccc	26460
	gtcaccacca agaacgggga aatcacctc aagctgggag agggggtgga cctcgattcc	26520
	tcgggaaaac teatctcaa cacggccacc aaggccccc ccctctcag tttttcaac	26580
	aacaccattt cccttaacat ggatcacccc ttttaccta aagatggaaa attatcctta	26640

	caagtttctc caccattaata tatactgaga acaagcattc taaacacact agcttttaggt	26700
	tttgatcag gtttaggact ccgtggctct gccttggcag tacagttagt ctctccactt	26760
	acatttgata ctgatggaaa cataaagctt accttagaca gaggtttgca tgttacaaca	26820
	ggagatgcaa ttgaaagcaa cataagctgg gctaaagggt taaaatttga agatggagcc	26880
	atagcaacca acattggaaa tgggttagag tttggaagca gtagtacaga aacaggtgtt	26940
	gatgatgctt acccaatcca agttaaactt ggatctggcc ttagctttga cagtacagga	27000
	gccataatgg ctggtaacaa agaagacgat aaactcactt tgtggacaac acctgatcca	27060
	tcaccaaact gtcaataact cgcagaaaat gatgcaaac taacactttg cttgactaaa	27120
	tgtggtagtc aaatactggc cactgtgtca gtcttagttg taggaagtgg aaacctaac	27180
	cccattactg gcaccgtaag cagtgtctcag gtgtttctac gttttgatgc aaacggtgtt	27240
	cttttaacag aacattctac actaaaaaaaa tactgggggt ataggcaggg agatagcata	27300
	gatggcactc catataccaa tgctgtagga ttcattgcca atttaaaagc ttatccaaag	27360
[0174]	tcacaaagtt ctactactaa aaataatata gtagggcaag tatacatgaa tggagatgtt	27420
	tcaaaaccta tgcttctcac tataaccctc aatggtactg atgacagcaa cagtacatat	27480
	tcaatgtcat tttcatacac ctggactaat ggaagctatg ttggagcaac atttggggct	27540
	aactettata ccttctcata catcgcccaa gaatgaacac tgtatccac cctgcatgcc	27600
	aaccttccc acccactct gtggaacaaa ctctgaaaca caaaataaaa taaagttcaa	27660
	gtgttttatt gattcaacag ttttacagga ttcgagcagt tatttttctt ccacctccc	27720
	aggacatgga atacaccacc ctctccccc gcacagcctt gaacatctga atgccattgg	27780
	tgatggacat gcttttggtc tccacgttcc acacagtttc agagcgagcc agtctcgggt	27840
	cggtcaggga gatgaaacce tccgggcaact cccgcatctg cacctcacag ctcaacagct	27900
	gaggattgtc ctcggtggtc gggatcacgg ttatctggaa gaagcagaag agcggcgggtg	27960
	ggaatcatag tccgcgaacg ggatcggccg gtggtgtcgc atcaggcccc gcagcagtcg	28020
	ctgccgccgc cgctcgtca agctgctgct caggggttcc ggggtccaggg actccctcag	28080
	catgatgccc acggccctca gcatcagtcg tctggtgcgg cgggcgcagc agcgcagtcg	28140

	gatctcgtc aggtcgtgc agtacgtgca acacagaacc accaggttgt tcaacagtcc	28200
	atagttcaac acgctccagc cgaaactcat cgcgggaagg atgctacca cgtggccgtc	28260
	gtaccagatc ctcaggtaaa tcaagtggg cccctccag aacacgtgc ccacgtacat	28320
	gatctccttg ggcatgtggc ggttcaccac ctcccgttac cacatcacc tctggttgaa	28380
	catgcagccc cggatgatcc tgcggaacca cagggccagc accgccccgc ccgcatgca	28440
	gcgaagagac cccgggtccc ggcaatggca atggaggacc caccgctcgt acccgtgat	28500
	catctgggag ctgaacaagt ctatgttggc acagcacagg catatgctca tgcattctt	28560
	cagcactctc aactcctcgg gggtaaaaac catatcccag ggcacgggga actcttgacg	28620
	gacagcgaac cccgcagaac agggcaatcc tcgcacagaa cttacattgt gcatggacag	28680
	ggtatcgcaa tcaggcagca ccgggtgatc ctccaccaga gaagcgcggg tctcgtctc	28740
	ctcacagcgt ggtaaggggg ccggccgata cgggtgatgg cgggacgcgg ctgatcgtg	28800
	tcgcgaccgt gtcatgatgc agttgcttcc ggacatttcc gtacttgctg tagcagaacc	28860
	tggtccgggc gctgcacacc gatcgcgggc ggcggtctcg gcgcttgaa cgctcgggtg	28920
[0175]	tgaaattgta aaacagccac tctctcagac cgtgcagcag atctagggcc tcaggagtga	28980
	tgaagatccc atcatgcctg atggctctga tcacatcgac caccgtggaa tgggccagac	29040
	ccagccagat gatgcaattt tgttgggttt cgggtgacggc gggggaggga agaacaggaa	29100
	gaaccatgat taacttttaa tccaaacggt ctcggagtac ttcaaatga agatcgcgga	29160
	gatggcacct ctcgccccg ctgtgttggg ggaaaataac agccaggtca aaggtgatac	29220
	ggttctcgag atgttccacg gtggcttcca gcaaagcctc cacgcgcaca tccagaaaca	29280
	agacaatagc gaaagcggga gggttctcta attcctcaat catcatgtta cactcctgca	29340
	ccatccccag ataattttca ttttccagc cttgaatgat tcgaactagt tcctgaggtg	29400
	aatccaagcc agccatgata aagagctcgc gcagagcgc ctcaccggc attcttaagc	29460
	acaccctcat aattccaaga tattctgctc ctggttcacc tgcagcagat tgacaagcgg	29520
	aatatcaaaa tctctgccgc gatccctgag ctctccctc agcaataact gtaagtactc	29580
	tttcatatcc tctccgaaat ttttagccat aggaccacca ggaataagat tagggcaagc	29640
	cacagtacag ataaaccgaa gtcctcccca gtgagcattg ccaaatgcaa gactgctata	29700

	agcatgctgg ctagaccggt tgatatcttc cagataactg gacagaaaat cgcccaggca	29760
	atthttaaga aaatcaacaa aagaaaaate ctccagggtg acgttttagag cctcgggaac	29820
	aacgatgaag taaatgcaag cgggtgcgttc cagcatggtt agtttagctga tctgtagaaa	29880
	aaacaaaaat gaacattaaa ccatgctagc ctggcgaaca ggtgggtaaa tcgttctctc	29940
	cagcaccagg caggccacgg ggtctccggc gcgaccctcg taaaaattgt cgctatgatt	30000
	gaaaaccatc acagagagac gttcccgggtg gccggcgtga atgattcgac aagatgaata	30060
	cacccccgga acattggcgt ccgcgagtga aaaaaagcgc ccgaggaagc aataaggcac	30120
	tacaatgctc agtctcaagt ccagcaaagc gatgccatgc ggatgaagca caaaattctc	30180
	aggtgcgtac aaaatgtaat tactcccctc ctgcacagge agcaaagccc ccgatacctc	30240
	caggtacaca tacaagcct cagcgtccat agcttaccga gcagcagcac acaacaggcg	30300
	caagagtcag agaaaggctg agctctaacc tgtccaccg ctctctgctc aatatatagc	30360
	ccagatctac actgacgtaa aggccaaagt ctaaaaatac ccgccaata atcacacacg	30420
[0176]	cccagcacac gccagaaac cgggtacaca ctcaaaaaaa tacgcgcaact tcctcaaacg	30480
	ccccaaactg ccgtcatttc cgggttccca cgctacgtca tcaaaacacg actttcaaat	30540
	tccgtcgacc gttaaaaacg tcaccgccc cgcccetaac ggtcgcccgt ctctcagcca	30600
	atcagcgecc cgcatccca aattcaaaca cctcatttgc atattaacgc gcacaaaaag	30660
	tttgaggtat attattgatg atgg	30684
	<210> 14	
	<211> 8602	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 14	
	atggcgcgcg catgagagaa gccagacca attacctacc caaaatggag aaagttcacg	60
	ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg	120
	aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc	180

	tggcttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa	240
	gtgcgccccg cgcgagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
	gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
	aaataactga taaggaattg gacaagaaaa tgaaggagct cggcccgctc atgagcgacc	420
	ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
	aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc taccaccaag	540
	ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta	600
	agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa	660
	cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttcaact	840
	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta	960
[0177]	cgatgcaccg cgagggatc ttgtgctgca aagtgcaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
	tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140
	tagtcgtcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgcccg	1200
	tagtggccca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacgagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
	acaagataac atctatttat aagcggcccg atacccaaac catcatcaa gtgaacagcg	1380
	atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc ggctgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgct acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctctg ggcttgataa	1680
	aggttaccag ctacgctggc gaggacaaga tcggctctta cgctgtgctt tctccgcagg	1740

	ctgtactcaa gaggtaaaaa ttatcttgca tccaccctct cgctgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggegttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtgaa agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaaaca ggtacctgca ccatattgcc acacatggag	1980
	gagcgtgaa cactgatgaa gaatattaca aaactgtcaa gcccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctggtg gatcctccct tccatgaatt cgcctacgag agtctgagaa	2160
	cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctagtggtg agcgccaaga	2280
	aagaaaactg tgcagaaatt ataagggacg tcaagaaat gaaagggtg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
	ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
[0178]	ctaaaaaggc agtgetctgc ggggatccca aacagtgcgg ttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttgca cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactaa atctgtgact tgggtcgtct caaccttgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtacc aaacctaagc	2700
	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggt gaccgtaaa ggtgtgtatg	2820
	ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tcctactgac ccgcacggag gaccgatcg tgtggaaaac actagccggc gaccatgga	2940
	taaaaacact gactgccaag taccctggga atttactgc cacgatagag gagtggcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtettcc	3060
	agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattatct tgaaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240

	gtctattttc tgcacccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
	cgtcgcctaa catgtacggg ctgaataaag aagtgggccg tcagctctct cgcaggtacc	3360
	cacaactgcc tgggcagtt gccactggaa gagtctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgcata aacctagtag ctgtaaacag aagactgcct catgctttag	3480
	tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaattgaagg	3540
	gcagaactgt cctggtggtc ggggaaaagt tgtccgtccc aggcaaatg gttgactggt	3600
	tgtcagaccg gcctgaggct accttcagag ctcggctgga tttaggcata ccaggtgatg	3660
	tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcata	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggta cgctgacagg gccagcghaa	3840
	gcacattgg tgctatagcg cggcagttca agttttccc ggtatgcaaa ccgaaatcct	3900
	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
	acaatcctta caagctttca tcaacctga ccaacattta tacaggttcc agactccacg	4020
[0179]	aagccggatg tgcaccctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccghaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tggtaaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc acaaaagttt	4260
	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgt aagattgtca	4320
	acgataaaa ttacaagtca gtagcgattc cactgttgc caccggcata tttccggga	4380
	acaaagatcg actaaccaa tcattgaacc atttctgac agcttttagac accactgatg	4440
	cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgagggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaactttc tcatatttg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatcct cggagaaaagc atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800

	aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040
	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcatc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggcccagc ccaccaggtg ctgcaagtcg	5220
	aggcagacat tcacgggccg cctctgtat ctagctcatc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacacct ggagggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgt ccgcgcacaa	5460
	gaacaccgtc acttgcacc agcaggcct gctcgagaac cagcctagtt tccaccccgc	5520
[0180]	caggcgtgaa tagggtgatc actagagagg agctcgagge gcttaccceg tcacgactc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaacc gccaggcgta aatagggtag	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcggtg	5700
	catacatctt ttctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
	cggtgctatc cgaagtgtg ttggagagga ccgaattgga gatttcgtat gccccgcgc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatccc acacctgta	5880
	acagaagcag ataccagtcc aggaaggtgg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacc	6000
	tgcatcctgt tcctttgtat tcacttagtg tgaaccgtgc cttttcaagc cccaaggtcg	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtgget tcttactgta	6120
	ttattccaga gtacgatgcc tatttggaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgcctgca aagctgcgca gctttccaaa gaaacactcc tatttgaac	6240
	ccacaatacg atcggcagtg ccttcagcga tccagaacac gctccagaac gtccctggcag	6300

	ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgta ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt	6420
	ttaaagaaaa ccccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggacaaaa agctgctgct ctttttgcga agacacataa ttggaatatg ttgcaggaca	6540
	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttcoga	6720
	acattcatac actgtttgat atgtcggctg aagactttga cgctattata gccgagcact	6780
	tccagcctgg ggattgtgtt ctggaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatggc tctgaccgcg ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgctgat tgaggcggct ttcggcgaaa tttcatcaat acatttgccc actaaaacta	6960
	aatttaaatt cggagccatg atgaaatctg gaatgttctt cacactgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
[0181]	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcmc cacctggttg aatatggaag tcaagattat agatgctgtg gtggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga cccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
	gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaccgta ggaacttcca	7440
	tcatagtatt ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gactctagaa tagtctttaa	7560
	ttaaagtccg ccatatgagg ccaccatgca gatcttcgtg aagacctga ccggcaagac	7620
	catcaccta gaggtggagc ccagtgacac catcgagaac gtgaaggcca agatccagga	7680
	taaagagggc atccccctg accagcagag gctgatcttt gccggcaagc agctggaaga	7740
	tggccgcacc ctctctgatt acaacatcca gaaggagtca acctgcacc tggctcttcg	7800
	cctgagaggt ggcgctgctt acagtataat caactttgaa aaactggctg cttacggcat	7860

	cctgggcttt gtgtttacac tggctgccta cctgctgttt ggctatcctg tgtacgtggc	7920
	cgcttatgga ctgtgtaccc tgggtggccat gctggctgct tacaatctgg tgcctatggt	7980
	ggccacagtg gccgcctatt gtcttggcgg actgctgaca atgggtggcag cctacagccc	8040
	gagctatgcg tatcatcagt ttgcagccta cggeccagga ccaggcgcta aatttgtggc	8100
	tgcctggaca ctgaaagccg ccgctggacc aggtcctgga cagtacatca aggccaacag	8160
	caagttcatc ggcatcaccg aactcggccc aggaccaggc tatccctacg atgtgcctga	8220
	ttacgcctga tagtgatgat tcgaacggcc gtatcacgcc caaacattta cagcccggt	8280
	gtcaaaaacc gcgtggacgt ggtaacatc cctgctggga ggatcagccg taattattat	8340
	aattggcttg gtgctggcta ctattgtggc catgtacgtg ctgaccaacc agaaacataa	8400
	ttgaatacag cagcaattgg caagctgctt acatagaact cgcggcgatt ggcatgccgc	8460
	cttaaaatth ttatthttatt thttctthtc thttccgaat cggatthttgt thttaatatt	8520
	tcaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	8580
[0182]	aaaaaaaaaa aaaaaaaaaa aa	8602

<210> 15

<211> 9595

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 15

atgggctggc	catgagagaa gccagacca attacctacc caaaatggag aaagttcacg	60
ttgacatcga	ggaagacagc ccattcctca gagctttgca gcgagcttc ccgcagtttg	120
aggtagaagc	caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc	180
tgcttcaaaa	actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa	240
gtgcgcccgc	ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
gtgcggaaga	tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
aaataactga	taaggaattg gacaagaaaa tgaaggagct cgccgcccgc atgagcgacc	420

	ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
	aagtcgctgt ttaccaggat gtatacgcggt ttgacggacc gacaagtctc taccaccaag	540
	ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta	600
	agaacttggc tggagcatat ccatcactat ctaccaactg ggccgacgaa accgtgttaa	660
	cggtcgttaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
	ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
	ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttctact	840
	tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
	tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta	960
	cgatgcaccg cgagggatc ttgtgctgca aagtgcaga cacattgaac ggggagaggg	1020
	tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
	tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgta	1140
	tagtcgtcaa cggtcgacc cagagaaaca ccaataccat gaaaaattac cttttgccc	1200
[0183]	tagtggccca ggcatttgcg aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
	ggccactagg actacagat agacagttag tcatggggtg ttgttgggct tttagaaggc	1320
	acaagataac atctatttat aagcggcccg ataccctaac catcatcaa gtgaacagcg	1380
	attccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
	caagaatcag gaaaatgta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
	acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
	tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
	tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa	1680
	aggttaccag ctacgtggc gaggacaaga tcggctctta cgtgtgctt tctccgcagg	1740
	ctgtactcaa gaggtaaaaa ttatcttgca tccacctct cgtgaacaa gtcatagtga	1800
	taacacactc tggccgaaaa gggcgttatg ccgtggaacc ataccatggt aaagtagtgg	1860
	tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtgaa agtgccacca	1920
	ttgtgtacaa cgaacgtgag ttcgtaacaa ggtacctgca ccatattgcc acacatggag	1980

	gagcgctgaa cactgatgaa gaatattaca aaactgtcaa gcccagcgag cacgacggcg	2040
	aatacctgta cgacatcgac aggaaacagt gcgtcaagaa agaactagtc actgggctag	2100
	ggctcacagg cgagctgggt gatcctccct tccatgaatt cgcctacgag agtctgagaa	2160
	cacgaccagc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccaggatcag	2220
	gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctagtgggt agcgccaaga	2280
	aagaaaactg tgcagaaatt ataagggacg tcaagaaat gaaagggtg gacgtcaatg	2340
	ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata	2400
	ttgacgaagc ttttgcttgt catgcaggta ctctcagagc gctcatagcc attataagac	2460
	ctaaaaaggc agtgctctgc ggggatccca aacagtgcgg ttttttaac atgatgtgcc	2520
	tgaaagtgca ttttaaccac gagatttga cacaagtctt ccacaaaagc atctctcgcc	2580
	gttgactactaa atctgtgact tgggtcgtct caaccttgtt ttacgacaaa aaaatgagaa	2640
	cgacgaatcc gaaagagact aagattgtga ttgacactac cggcagtacc aaacctaagc	2700
[0184]	aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca	2760
	aaggcaacga aataatgacg gcagctgcct ctcaagggt gaccctaaa ggtgtgtatg	2820
	ccgttcggta caaggtgaat gaaaatcctc tgtacgcacc cacctcagaa catgtgaacg	2880
	tectactgac ccgcacggag gaccgcatcg tgtggaaaac actagccggc gaccatgga	2940
	taaaaacact gactgccaag taccctggga atttactgc cacgatagag gagtggcaag	3000
	cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgtcttcc	3060
	agaataaggc aaacgtgtgt tgggccaagg ctttagtgcc ggtgctgaag accgctggca	3120
	tagacatgac cactgaacaa tggaacactg tggattattt tgaaacggac aaagctcact	3180
	cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgat ctggactccg	3240
	gtctattttc tgcaccact gttccgttat ccattaggaa taatcactgg gataactccc	3300
	cgtcgcctaa catgtacggg ctgaataaag aagtgggtccg tcagctctct cgcaggtacc	3360
	cacaactgcc tcgggcagtt gccactggaa gagtctatga catgaacact ggtacactgc	3420
	gcaattatga tccgcgcata aacctagtac ctgtaaacag aagactgcct catgctttag	3480

	tcctccacca taatgaacac ccacagagtg actttttcttc attcgtcage aaattgaagg	3540
	gcagaactgt cctggtggtc ggggaaaagt tgtccgtccc aggcaaaatg gttgactggt	3600
	tgtcagaccg gcctgaggct accttcagag ctcggtgga tttaggcatc ccaggtgatg	3660
	tgcccaaata tgacataata tttgttaatg tgaggacccc atataaatac catcactatc	3720
	agcagtgtga agaccatgcc attaagctta gcatgttgac caagaaagct tgtctgcatc	3780
	tgaatcccgg cggaacctgt gtcagcatag gttatggta cgtgacagg gccagcгаа	3840
	gcatcattgg tgctatagcg cggcagttca agttttcccg ggtatgcaaа ccgaaatcct	3900
	cacttgaaga gacggaagtt ctgtttgtat tcattgggta cgatcgcaag gcccgtagc	3960
	acaatcctta caagctttca tcaacctga ccaacattta tacaggttcc agactccacg	4020
	aagccggatg tgcacctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag	4080
	gagtgattat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc	4140
	tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac	4200
	tgtcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc acaaagttt	4260
[0185]	cggaggttga aggtgacaaa cagttggcag aggcttatga gtccatcgtc aagattgtca	4320
	acgataacaa ttacaagtca gtagcgattc cactgtttgc caccggcatc tttccggga	4380
	acaaagatcg actaaccaa tcattgaacc atttgctgac agcttttagac accactgatg	4440
	cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg	4500
	ctaggagaga agcagtggag gagatatgca tatccgacga ctcttcagtg acagaacctg	4560
	atgcagagct ggtgaggggtg catccgaaga gttctttggc tggaaggaag ggctacagca	4620
	caagcgatgg caaaactttc tcatatttgg aagggaccaa gtttcaccag gcggccaagg	4680
	atatagcaga aattaatgcc atgtggcccg ttgcaacgga ggccaatgag caggtatgca	4740
	tgtatatacct cggagaaage atgagcagta ttaggtcgaa atgccccgtc gaagagtcgg	4800
	aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgccatg actccagaaa	4860
	gagtacagcg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat	4920
	tgccgaagta tagaatcact ggtgtgcaga agatccaatg ctcccagcct atattgttct	4980
	caccgaaagt gcctgcgtat attcatccaa ggaagtatct cgtggaaaca ccaccggtag	5040

	acgagactcc ggagccatcg gcagagaacc aatccacaga ggggacacct gaacaaccac	5100
	cacttataac cgaggatgag accaggacta gaacgcctga gccgatcadc atcgaagagg	5160
	aagaagagga tagcataagt ttgctgtcag atggcccagc ccaccagggtg ctgcaagtcg	5220
	aggcagacat tcacgggccc cctctgtat ctagctcadc ctggtccatt cctcatgcat	5280
	ccgactttga tgtggacagt ttatccatac ttgacacct ggagggagct agcgtgacca	5340
	gcggggcaac gtcagccgag actaactctt acttcgcaaa gagtatggag tttctggcgc	5400
	gaccggtgcc tgcgcctcga acagtattca ggaacctcc acatcccgt cgcgcacaaa	5460
	gaacaccgtc acttgcacc agcaggcct gctcgagaac cagcctagtt tccaccccgc	5520
	caggcgtgaa tagggtgadc actagagagg agctcgaggc gcttaccocg tcacgcadc	5580
	ctagcaggtc ggtctcgaga accagcctgg tctccaacc gccaggcgt aatagggtga	5640
	ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gatgcccgtg	5700
	catacatctt ttctccgac accggtcaag ggcatttaca acaaaaatca gtaaggcaaa	5760
[0186]	cggtgctadc cgaagtgtg ttggagagga ccgaattgga gatttcgtat gccccgcgc	5820
	tcgaccaaga aaaagaagaa ttactacgca agaaattaca gttaaatccc acacctgcta	5880
	acagaagcag ataccagtc aggaagggtg agaacatgaa agccataaca gctagacgta	5940
	ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc taccgaacc	6000
	tgcacctgt tcctttgtat tcactagtg tgaaccgtgc ctttcaagc cccaaggtcg	6060
	cagtggaagc ctgtaacgcc atgttgaaag agaactttcc gactgtggct tcttactgta	6120
	ttattccaga gtacgatgcc tatttgaca tggttgacgg agcttcatgc tgcttagaca	6180
	ctgccagttt ttgcctgca aagctgcgca gttttcaaaa gaaacactcc tatttggaa	6240
	ccacaatacg atcggcagtg cttcagcga tccagaacac gctccagaac gtccctggcag	6300
	ctgccacaaa aagaaattgc aatgtcacgc aaatgagaga attgcccgtt ttggattcgg	6360
	cggcctttaa tgtggaatgc ttcaagaaat atgcgtgtaa taatgaatat tgggaaacgt	6420
	ttaaagaaaa cccatcagg cttactgaag aaaacgtggt aaattacatt accaaattaa	6480
	aaggacaaaa agctgctgct cttttgca agacacataa tttgaatatg ttgcaggaca	6540

	taccaatgga caggtttgta atggacttaa agagagacgt gaaagtgact ccaggaacaa	6600
	aacatactga agaacggccc aaggtacagg tgatccaggc tgccgatccg ctagcaacag	6660
	cgtatctgtg cggaatccac cgagagctgg ttaggagatt aaatgcggtc ctgcttccga	6720
	acattcatac actgtttgat atgtcggctg aagacttga cgctattata gccgagcact	6780
	tccagcctgg ggattgtgtt ctggaaactg acatcgcgtc gtttgataaa agtgaggacg	6840
	acgccatgac tctgaccgag ttaatgattc tggaagactt aggtgtggac gcagagctgt	6900
	tgacgtgat tgaggcggct ttcggcgaaa tttcatcaat acatttggcc actaaaacta	6960
	aatttaaatt cggagccatg atgaaactcg gaatgttctt cacactgttt gtgaacacag	7020
	tcattaacat tgtaatcgca agcagagtgt tgagagaacg gctaaccgga tcaccatgtg	7080
	cagcattcat tggagatgac aatatcgtga aaggagtcaa atcggacaaa ttaatggcag	7140
	acaggtgcgc cacctggttg aatatggaag tcaagattat agatgctgtg gtggcgaga	7200
	aagcgcctta tttctgtgga gggtttattt tgtgtgactc cgtgaccggc acagcgtgcc	7260
	gtgtggcaga cccctaaaa aggctgttta agcttgcaa acctctggca gcagacgatg	7320
[0187]	aacatgatga tgacaggaga agggcattgc atgaagagtc aacacgctgg aaccgagtgg	7380
	gtattctttc agagctgtgc aaggcagtag aatcaaggta tgaaccgta ggaacttcca	7440
	tcatagtatt ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag	7500
	gggcccctat aactctctac ggctaacctg aatggactac gactctagaa tagtctttaa	7560
	ttaaagtccg ccatatgaga tggaagatgc caaaaacatt aagaagggcc cagcgcatt	7620
	ctaccactc gaagacggga ccgccggcga gcagctgcac aaagccatga agcgcctacg	7680
	cctggtgccc ggcaccatcg cctttaccga cgcacatata gaggtggaca ttacctacg	7740
	cgagtacttc gagatgagcg ttcggctggc agaagctatg aagcgtatg ggctgaatac	7800
	aaaccatcgg atcgtggtgt gcagcgagaa tagcttgacg ttcttcatgc ccgtgttggg	7860
	tgccctgttc atcgggtgtg ctgtggcccc agctaacgac atctacaacg agcgcgagct	7920
	gctgaacagc atgggcatca gccagcccac cgtcgtattc gtgagcaaga aagggtgca	7980
	aaagatctc aacgtgcaaa agaagctacc gatcatacaa aagatcatca tcatggatag	8040
	caagaccgac taccagggt tccaaagcat gtacacctc gtgacttccc atttgcacc	8100

	cggettcaac gaggtagact tegtgtccga gagcttcgac cgggacaaaa ccatcgcctt	8160
	gatcatgaac agtagtggca gtaccggatt gcccaagggc gtagccctac cgcaccgcac	8220
	cgcttgtgtc cgattcagtc atgcccgcga ccccatcttc ggcaaccaga tcatccccga	8280
	caccgctatc ctcagcgtgg tgccatttca ccacgcttc ggcattgtca ccacgctggg	8340
	ctacttgatc tgcggctttc gggtcgtgct catgtaccgc ttcgaggagg agctattctt	8400
	gcgagccttg caagactata agattcaatc tgccctgctg gtgcccacac tatttagctt	8460
	cttcgctaag agcactctca tcgacaagta cgacctaaag aacttgcacg agatcgccag	8520
	cggcggggcg ccgctcagca aggaggtagg tgaggccgtg gccaaacgt tccacctacc	8580
	aggcatccgc cagggtctac gcctgacaga aacaaccage gccattctga tcacccccga	8640
	aggggacgac aagcctggcg cagtaggcaa ggtggtgccc ttcttcgagg ctaaggtggt	8700
	ggacttggac accggtgaaga cactgggtgt gaaccagcgc ggcgagctgt gcgtccgtgg	8760
	ccccatgac atgagcggct acgttaacaa ccccgaggt acaaacgctc tcacgacaa	8820
[0188]	ggacggctgg ctgcacagcg gcgacatcgc ctactgggac gaggacgagc acttcttcat	8880
	cgtggaccgg ctgaagagcc tgatcaaata caagggtac caggtagccc cagccgaact	8940
	ggagagcadc ctgctgcaac accccaacat cttcgacgce ggggtcgcg gcctgcccga	9000
	cgacgatgcc ggcgagctgc ccgccagct cgtcgtgctg gaacacggta aaacctgac	9060
	cgagaaggag atcgtggact atgtggccag ccaggttaca accgccaaga agctgcgcgg	9120
	tggtgtgtg ttcgtggacg aggtgcctaa aggactgacc ggcaagttgg acgcccga	9180
	gatccgcgag attctcatta aggccaagaa gggcggcaag atcgcctgt aattcgaacg	9240
	gccgtatcac gcccaaacat ttacagccgc ggtgtcaaaa accgcgtgga cgtggttaac	9300
	atccctgctg ggaggatcag ccgtaattat tataattggc ttggtgctgg ctactattgt	9360
	ggccatgtac gtgctgacca accagaaaca taattgaata cagcagcaat tgcaagctg	9420
	cttacataga actcgcggcg attggcatgc gccttaaaa tttttatfff atttttctt	9480
	ttcttttccg aatcggattt tgtttttaat atttcaaaaa aaaaaaaaaa aaaaaaaaaa	9540
	aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa	9595

<210> 16
 <211> 139
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多肽

<400> 16
 Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
 1 5 10 15

 Arg Ala Ser Gln Ser Ile Asn Ser Tyr Leu Asp Trp Tyr Gln Gln Lys
 20 25 30

 Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Ala Ala Ser Ser Leu Gln
 35 40 45

 Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
 50 55 60

 [0189]
 Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr
 65 70 75 80

 Cys Gln Gln Tyr Tyr Ser Thr Pro Phe Thr Phe Gly Pro Gly Thr Lys
 85 90 95

 Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro
 100 105 110

 Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
 115 120 125

 Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val
 130 135

 <210> 17
 <211> 167
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多肽

<400> 17

Gly Val Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser
 1 5 10 15

Gly Phe Thr Phe Ser Ser Tyr Gly Met His Trp Val Arg Gln Ala Pro
 20 25 30

Gly Lys Gly Leu Glu Trp Val Ala Val Ile Trp Tyr Asp Gly Ser Asn
 35 40 45

Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp
 50 55 60

Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu
 65 70 75 80

[0190]

Asp Thr Ala Val Tyr Tyr Cys Ala Arg Asp Pro Arg Gly Ala Thr Leu
 85 90 95

Tyr Tyr Tyr Tyr Tyr Gly Met Asp Val Trp Gly Gln Gly Thr Thr Val
 100 105 110

Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala
 115 120 125

Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu
 130 135 140

Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly
 145 150 155 160

Ala Leu Thr Ser Gly Val His
 165

<210> 18

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成肽

<400> 18

Gly Phe Thr Phe Ser Ser Tyr Gly Met His
1 5 10

<210> 19

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成肽

<400> 19

Val Ile Trp Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser Val
1 5 10 15

<210> 20

<211> 16

[0191]

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成肽

<400> 20

Asp Pro Arg Gly Ala Thr Leu Tyr Tyr Tyr Tyr Tyr Gly Met Asp Val
1 5 10 15

<210> 21

<211> 11

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成肽

<400> 21

Arg Ala Ser Gln Ser Ile Asn Ser Tyr Leu Asp
1 5 10

<210> 22

<211> 7

<212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 22
 Ala Ala Ser Ser Leu Gln Ser
 1 5

<210> 23
 <211> 9
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 23
 Gln Gln Tyr Tyr Ser Thr Pro Phe Thr
 1 5

[0192]

<210> 24
 <211> 108
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多肽

<400> 24
 Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
 1 5 10 15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Arg Val Ser Ser Ser
 20 25 30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu
 35 40 45

Ile Tyr Asp Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser
 50 55 60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu
 65 70 75 80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Ser Leu Pro
 85 90 95

Trp Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
 100 105

<210> 25

<211> 121

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多肽

<400> 25

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 Phe Thr Phe Ser Arg Tyr
 20 25 30

[0193]

Trp Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ala Asn Ile Lys Gln Asp Gly Ser Glu Lys Tyr Tyr Val Asp Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu 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 Arg Glu Gly Gly Trp Phe Gly Glu Leu Ala Phe Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

<210> 26
 <211> 5
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 26
 Arg Tyr Trp Met Ser
 1 5

<210> 27
 <211> 17
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 27
 Asn Ile Lys Gln Asp Gly Ser Glu Lys Tyr Tyr Val Asp Ser Val Lys
 1 5 10 15

[0194]

Gly

<210> 28
 <211> 12
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 28
 Glu Gly Gly Trp Phe Gly Glu Leu Ala Phe Asp Tyr
 1 5 10

<210> 29
 <211> 12
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 29
 Arg Ala Ser Gln Arg Val Ser Ser Ser Tyr Leu Ala
 1 5 10

<210> 30
 <211> 7
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

<400> 30
 Asp Ala Ser Ser Arg Ala Thr
 1 5

<210> 31
 <211> 9
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成肽

[0195]

<400> 31
 Gln Gln Tyr Gly Ser Leu Pro Trp Thr
 1 5

<210> 32
 <211> 2019
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多核苷酸

<400> 32
 gcccgggcat ttaaattgca tgcgatgat tacgactcta gaatagtcta gtccgcagge 60
 caccatgcag atcttcgtga agacctgac cggcaagacc atcacctag aggtggagcc 120
 cagtgcacc atcgagaacg tgaaggccaa gatccaggat aaagaggca tccccctga 180
 ccagcagagg ctgatctttg cggcaagca gctggaagat ggccgcaccc tctctgatta 240
 caacatccag aaggagtcaa cctgcacct ggtccttcgc ctgagaggtg ccatgtttca 300
 ggcgctgagc gaaggetgca ccccgatga tattaaccag atgctgaacg tgctgggcga 360

	tcatacaggtc tcaggccttg agcagcttga gagtataatc aactttgaaa aactgactga	420
	atggaccagt tctaattgta tgcctatcct gtctcctctg acaaaggga tccctgggctt	480
	cgtgtttacc ctgaccgtgc cttctgagag aggacttagc tgcattagcg aagcggatgc	540
	gaccaccccg gaaagcgcga acctgggcga agaaattctg agccagctgt atctttggcc	600
	aagggtgacc taccattccc ctagtattgc ttaccaccaa tttgaaagac gagccaaata	660
	taaaagacac ttccccggct ttggccagag cctgctgttt ggctaccctg tgtacgtgtt	720
	cggcgattgc gtgcaggcgc attgggatgc gattcgcttt cgctattgcg cgccgccggg	780
	ctatgcgctg ctgcgctgca acgataccea ctatagcgt ctgctggctg tgggggcct	840
	agaaggacce aggaatcagg actggcttgg tgtccaaga caactttaa ctcggatgca	900
	ggctattcag aatgccggcc tgtgtaccct ggtggccatg ctggaagaga caatcttctg	960
	gctgcaagcg tttctgatgg cgctgaccga tagcggcccg aaaaccaaca ttattgtgga	1020
	tagccagtat gtgatgggca ttagcaaac gagctttcag gaatttggg attgggaaaa	1080
[0196]	cgtgagcccg gaactgaaca gcaccgatca gccgttttgg caagccgga tccctggccag	1140
	aaatctggtg cctatggtgg ccacagtgca gggccagaac ctgaagtacc agggtcagtc	1200
	actagtcate tctgcttcta tcattgtctt caacctgctg gaactggaag gtgattatcg	1260
	agatgatggc aacgtgtggg tgcatacccc gctgagcccg cgcaccctga acgctgggt	1320
	gaaagcggtg gaagaaaaa aaggtattcc agttcaccta gagctggcca gtatgaccaa	1380
	catggagctc atgagcagta ttgtgcatca gcaggtcaga acatacggcc ccgtgttcat	1440
	gtgtctcggc ggactgctta caatggtggc tgggtgctgtg tggctgacag tgcgagtgt	1500
	cgagetgttc cgggcgcgc agctggccaa cgacgtggte ctccagatca tggagctttg	1560
	tgggtcagcg tttccagcagg tgtgccatac caccgtgccg tggccgaacg cgagcctgac	1620
	cccgaatgg aacaacgaaa ccaccagcc ccagatgcc aactgcagcg tgtatgactt	1680
	ttttgtgtgg ctccattatt attctgttcg agacacactt tggccaaggg tgacctacca	1740
	tatgaacaaa tatgcgtatc atatgctgga aagacgagcc aaatataaaa gaggaccagg	1800
	acctggcgct aaatttgtgg ccgcctggac actgaaagcc gctgctggtc ctggacctgg	1860

ccagtacatc aaggccaaca gcaagttcat cggcatcacc gaactcggac cggaccagg 1920
 ctgatgattt cgaaatttaa ataagcttgc ggccgctagg gataacaggg taattatcac 1980
 gcccaaacat ttacagccgc ggtgtcaaaa accgcgtgg 2019

<210> 33
 <211> 619
 <212> PRT
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多肽

<400> 33
 Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
 1 5 10 15

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
 20 25 30

[0197] Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
 50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Ala Met Phe Gln Ala
 65 70 75 80

Leu Ser Glu Gly Cys Thr Pro Tyr Asp Ile Asn Gln Met Leu Asn Val
 85 90 95

Leu Gly Asp His Gln Val Ser Gly Leu Glu Gln Leu Glu Ser Ile Ile
 100 105 110

Asn Phe Glu Lys Leu Thr Glu Trp Thr Ser Ser Asn Val Met Pro Ile
 115 120 125

Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe Thr Leu Thr
 130 135 140

Val Pro Ser Glu Arg Gly Leu Ser Cys Ile Ser Glu Ala Asp Ala Thr
 145 150 155 160

Thr Pro Glu Ser Ala Asn Leu Gly Glu Glu Ile Leu Ser Gln Leu Tyr
 165 170 175

Leu Trp Pro Arg Val Thr Tyr His Ser Pro Ser Tyr Ala Tyr His Gln
 180 185 190

Phe Glu Arg Arg Ala Lys Tyr Lys Arg His Phe Pro Gly Phe Gly Gln
 195 200 205

Ser Leu Leu Phe Gly Tyr Pro Val Tyr Val Phe Gly Asp Cys Val Gln
 210 215 220

Gly Asp Trp Asp Ala Ile Arg Phe Arg Tyr Cys Ala Pro Pro Gly Tyr
 225 230 235 240

[0198] Ala Leu Leu Arg Cys Asn Asp Thr Asn Tyr Ser Ala Leu Leu Ala Val
 245 250 255

Gly Ala Leu Glu Gly Pro Arg Asn Gln Asp Trp Leu Gly Val Pro Arg
 260 265 270

Gln Leu Val Thr Arg Met Gln Ala Ile Gln Asn Ala Gly Leu Cys Thr
 275 280 285

Leu Val Ala Met Leu Glu Glu Thr Ile Phe Trp Leu Gln Ala Phe Leu
 290 295 300

Met Ala Leu Thr Asp Ser Gly Pro Lys Thr Asn Ile Ile Val Asp Ser
 305 310 315 320

Gln Tyr Val Met Gly Ile Ser Lys Pro Ser Phe Gln Glu Phe Val Asp
 325 330 335

Trp Glu Asn Val Ser Pro Glu Leu Asn Ser Thr Asp Gln Pro Phe Trp
 340 345 350

Gln Ala Gly Ile Leu Ala Arg Asn Leu Val Pro Met Val Ala Thr Val
 355 360 365

Gln Gly Gln Asn Leu Lys Tyr Gln Gly Gln Ser Leu Val Ile Ser Ala
 370 375 380

Ser Ile Ile Val Phe Asn Leu Leu Glu Leu Glu Gly Asp Tyr Arg Asp
 385 390 395 400

Asp Gly Asn Val Trp Val His Thr Pro Leu Ser Pro Arg Thr Leu Asn
 405 410 415

Ala Trp Val Lys Ala Val Glu Glu Lys Lys Gly Ile Pro Val His Leu
 420 425 430

Glu Leu Ala Ser Met Thr Asn Met Glu Leu Met Ser Ser Ile Val His
 435 440 445

[0199] Gln Gln Val Arg Thr Tyr Gly Pro Val Phe Met Cys Leu Gly Gly Leu
 450 455 460

Leu Thr Met Val Ala Gly Ala Val Trp Leu Thr Val Arg Val Leu Glu
 465 470 475 480

Leu Phe Arg Ala Ala Gln Leu Ala Asn Asp Val Val Leu Gln Ile Met
 485 490 495

Glu Leu Cys Gly Ala Ala Phe Arg Gln Val Cys His Thr Thr Val Pro
 500 505 510

Trp Pro Asn Ala Ser Leu Thr Pro Lys Trp Asn Asn Glu Thr Thr Gln
 515 520 525

Pro Gln Ile Ala Asn Cys Ser Val Tyr Asp Phe Phe Val Trp Leu His
 530 535 540

Tyr Tyr Ser Val Arg Asp Thr Leu Trp Pro Arg Val Thr Tyr His Met

acaacatca tcgtggattc ccagtagctg atgggcatct ccaagccttc tttccaggag	780
tttgtggact gggagaacgt gagcccagag ctgaattcca ccgatcagcc attctggcag	840
gcaggaatcc tggcaaggaa cctggtgcct atggtggcca cagtgcaggg ccagaatctg	900
aagtaccagg gccagagcct ggtcatcagc gcctccatca tcgtgtttaa cctgctggag	960
ctggagggcg actatcggga cgatggcaac gtgtgggtgc acacccact gagcccaga	1020
acactgaacg cctgggtgaa ggccgtggag gagaagaagg gcatcccagt gcacctggag	1080
ctggcctcca tgaccaatat ggagctgatg tctagcatcg tgcaccagca ggtgaggaca	1140
tacggacccg tgttcatgtg cctgggaggc ctgctgacca tgggtgcagg agccgtgtgg	1200
ctgacagtgc ggggtctgga gctgttcaga gccgcccagc tggccaacga tgtggtgctg	1260
cagatcatgg agctgtgcgg agcagccttt cgccagggtg gccacaccac agtgccatgg	1320
cccaatgcct cctgacccc caagtggaac aatgagaaa cacagcctca gatcgccaac	1380
tgtagcgtgt acgacttctt cgtgtggctg cactactata gcgtgagga taccctgtgg	1440
[0201] ccccgctga cataccacat gaataagtac gcctatcaca tgctggagag gcgcgccaag	1500
tataagagag gccctggccc aggcgcaaag tttgtggcag catggaccct gaaggccgcc	1560
gccggccccg gccccggcca gtatatcaag gctaacagta agttcattgg aatcacagag	1620
ctgggacccg gacctgga	1638

<210> 35

<211> 546

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多肽

<400> 35

Met Ala Gly Met Phe Gln Ala Leu Ser Glu Gly Cys Thr Pro Tyr Asp
 1 5 10 15

Ile Asn Gln Met Leu Asn Val Leu Gly Asp His Gln Val Ser Gly Leu
 20 25 30

Glu Gln Leu Glu Ser Ile Ile Asn Phe Glu Lys Leu Thr Glu Trp Thr
 35 40 45

Ser Ser Asn Val Met Pro Ile Leu Ser Pro Leu Thr Lys Gly Ile Leu
 50 55 60

Gly Phe Val Phe Thr Leu Thr Val Pro Ser Glu Arg Gly Leu Ser Cys
 65 70 75 80

Ile Ser Glu Ala Asp Ala Thr Thr Pro Glu Ser Ala Asn Leu Gly Glu
 85 90 95

Glu Ile Leu Ser Gln Leu Tyr Leu Trp Pro Arg Val Thr Tyr His Ser
 100 105 110

Pro Ser Tyr Ala Tyr His Gln Phe Glu Arg Arg Ala Lys Tyr Lys Arg
 115 120 125

[0202] His Phe Pro Gly Phe Gly Gln Ser Leu Leu Phe Gly Tyr Pro Val Tyr
 130 135 140

Val Phe Gly Asp Cys Val Gln Gly Asp Trp Asp Ala Ile Arg Phe Arg
 145 150 155 160

Tyr Cys Ala Pro Pro Gly Tyr Ala Leu Leu Arg Cys Asn Asp Thr Asn
 165 170 175

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

Asp Trp Leu Gly Val Pro Arg Gln Leu Val Thr Arg Met Gln Ala Ile
 195 200 205

Gln Asn Ala Gly Leu Cys Thr Leu Val Ala Met Leu Glu Glu Thr Ile
 210 215 220

Phe Trp Leu Gln Ala Phe Leu Met Ala Leu Thr Asp Ser Gly Pro Lys
 225 230 235 240

Thr Asn Ile Ile Val Asp Ser Gln Tyr Val Met Gly Ile Ser Lys Pro
 245 250 255

Ser Phe Gln Glu Phe Val Asp Trp Glu Asn Val Ser Pro Glu Leu Asn
 260 265 270

Ser Thr Asp Gln Pro Phe Trp Gln Ala Gly Ile Leu Ala Arg Asn Leu
 275 280 285

Val Pro Met Val Ala Thr Val Gln Gly Gln Asn Leu Lys Tyr Gln Gly
 290 295 300

Gln Ser Leu Val Ile Ser Ala Ser Ile Ile Val Phe Asn Leu Leu Glu
 305 310 315 320

Leu Glu Gly Asp Tyr Arg Asp Asp Gly Asn Val Trp Val His Thr Pro
 325 330 335

[0203] Leu Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Ala Val Glu Glu Lys
 340 345 350

Lys Gly Ile Pro Val His Leu Glu Leu Ala Ser Met Thr Asn Met Glu
 355 360 365

Leu Met Ser Ser Ile Val His Gln Gln Val Arg Thr Tyr Gly Pro Val
 370 375 380

Phe Met Cys Leu Gly Gly Leu Leu Thr Met Val Ala Gly Ala Val Trp
 385 390 395 400

Leu Thr Val Arg Val Leu Glu Leu Phe Arg Ala Ala Gln Leu Ala Asn
 405 410 415

Asp Val Val Leu Gln Ile Met Glu Leu Cys Gly Ala Ala Phe Arg Gln
 420 425 430

Val Cys His Thr Thr Val Pro Trp Pro Asn Ala Ser Leu Thr Pro Lys
 435 440 445

Trp Asn Asn Glu Thr Thr Gln Pro Gln Ile Ala Asn Cys Ser Val Tyr
 450 455 460

Asp Phe Phe Val Trp Leu His Tyr Tyr Ser Val Arg Asp Thr Leu Trp
 465 470 475 480

Pro Arg Val Thr Tyr His Met Asn Lys Tyr Ala Tyr His Met Leu Glu
 485 490 495

Arg Arg Ala Lys Tyr Lys Arg Gly Pro Gly Pro Gly Ala Lys Phe Val
 500 505 510

Ala Ala Trp Thr Leu Lys Ala Ala Ala Gly Pro Gly Pro Gly Gln Tyr
 515 520 525

Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Leu Gly Pro Gly
 530 535 540

[0204]

Pro Gly
 545

<210> 36

<211> 2019

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 36

gccccgggcat ttaaattgcga tgcacatgat tacgactcta gaatagtcta gtccgcagge 60

caccatgcag atcttcgtga agaccctgac cggcaagacc atcaccctag aggtggagcc 120

cagtgcacacc atcgagaacg tgaaggccaa gatccaggat aaagagggca tccccctga 180

ccagcagagg ctgatctttg ccggcaagca gctggaagat ggccgcaccc tctctgatta 240

caacatccag aaggagtcaa cctgcacct ggtccttcgc ctgagaggtg ccatgtttca 300

ggcgetgagc gaaggctgca ccccgtatga tattaaccag atgetgaacg tgetgggcga 360

	tcatacagttt aagcacatca aagcctttga cggacattht gctaacaacc caggtcccat	420
	ggttggtgttt gccacacctg ggcctatcct gtctcctctg acaaaggcca tcctgggctt	480
	cgtgtttacc ctgacctgce cttctgagag aggaacttagc tgcattagcg aagcggatgc	540
	gaccaccccc gaaagcgcga acctgggcga agaaattctg agccagctgt atctttggcc	600
	aagggtgacc taccattccc ctagtattgc ttaccaccaa ttgaaagac gagccaaata	660
	taaaagacac ttccccggct ttggccagag cctgctgttt ggctaccctg tgtacgtgtt	720
	cggcgattgc gtgcaggcg attgggatgc gattcctttt cgtatttgcg cgcgcgccgg	780
	ctatgcgctg ctgcgctgca acgataccaa ctatagcgt ctgctggctg tgggggcct	840
	agaaggacc aggaatcagg actggcttgg tgtccaaga caacttgha ctcggatgca	900
	ggctattcag aatgccggcc tgtgtaccct ggtggccatg ctggaagaga caatcttctg	960
	gctgcaagcg tttctgatgg cgctgaccga tagcggcccc aaaaccaaca ttattgtgga	1020
	tagccagtat gtgatggcca ttagcaaac gagctttcag gaatttgtgg attgggaaaa	1080
	cgtagcccc gaactgaaca gcaccgatca gccgttttgg caagccgga tcctggccag	1140
[0205]	aaatctggtg cctatggtgg ccacagtgca gggccagaac ctgaagtacc aggtcagtc	1200
	actagtcatc tctgcttcta tcattgtctt caacctgctg gaactggaag gtgattatcg	1260
	agatgatggc aacgtgtggg tgcatacccc gctgagcccc cgcaccctga acgctgtggg	1320
	gaaagcggtg gaagaaaaaa aaggtattcc agttcaccta gagctggcca gtatgaccaa	1380
	catggagctc atgagcagta ttgtgcatca gcaggtcaga acatacggcc ccgtgttcat	1440
	gtgtctcggc ggactgctta caatggtggc tgggtctgtg tggctgacag tgcgagtgt	1500
	cgagctgttc cgggccgcgc agctggccaa cgacgtggtc ctccagatca tggagctttg	1560
	tgggtcagcg tttcgcagg tgtgccatac caccgtgccg tggccgaac cgagcctgac	1620
	cccgaatgg aacaacgaaa ccaccagcc ccagatgcc aactgcagcg tgtatgactt	1680
	ttttgtgtgg ctccattatt attctgttcg agacacactt tggccaaggg tgacctacca	1740
	tatgaacaaa tatgcgtatc atatgctgga aagacgagcc aatataaaa gaggaccagg	1800
	acctggcgt aaatttgtgg ccgcctggac actgaaagcc gctgctggtc ctggacctgg	1860
	ccagtacatc aaggccaaca gcaagttcat cggcatcacc gaactcggac ccggaccagg	1920

ctgatgattt cgaatttaa ataagcttgc ggccgctagg gataacaggg taattatcac 1980

gccccaaacat ttacagccgc ggtgtcaaaa accgcgtgg 2019

<210> 37

<211> 619

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> 人工序列的描述: 合成多肽

<400> 37

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
 20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 35 40 45

[0206]

Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Ala Met Phe Gln Ala
65 70 75 80

Leu Ser Glu Gly Cys Thr Pro Tyr Asp Ile Asn Gln Met Leu Asn Val
 85 90 95

Leu Gly Asp His Gln Phe Lys His Ile Lys Ala Phe Asp Arg Thr Phe
 100 105 110

Ala Asn Asn Pro Gly Pro Met Val Val Phe Ala Thr Pro Gly Pro Ile
 115 120 125

Leu Ser Pro Leu Thr Lys Gly Ile Leu Gly Phe Val Phe Thr Leu Thr
130 135 140

Val Pro Ser Glu Arg Gly Leu Ser Cys Ile Ser Glu Ala Asp Ala Thr
 145 150 155 160

Thr Pro Glu Ser Ala Asn Leu Gly Glu Glu Ile Leu Ser Gln Leu Tyr
 165 170 175

Leu Trp Pro Arg Val Thr Tyr His Ser Pro Ser Tyr Ala Tyr His Gln
 180 185 190

Phe Glu Arg Arg Ala Lys Tyr Lys Arg His Phe Pro Gly Phe Gly Gln
 195 200 205

Ser Leu Leu Phe Gly Tyr Pro Val Tyr Val Phe Gly Asp Cys Val Gln
 210 215 220

Gly Asp Trp Asp Ala Ile Arg Phe Arg Tyr Cys Ala Pro Pro Gly Tyr
 225 230 235 240

[0207] Ala Leu Leu Arg Cys Asn Asp Thr Asn Tyr Ser Ala Leu Leu Ala Val
 245 250 255

Gly Ala Leu Glu Gly Pro Arg Asn Gln Asp Trp Leu Gly Val Pro Arg
 260 265 270

Gln Leu Val Thr Arg Met Gln Ala Ile Gln Asn Ala Gly Leu Cys Thr
 275 280 285

Leu Val Ala Met Leu Glu Glu Thr Ile Phe Trp Leu Gln Ala Phe Leu
 290 295 300

Met Ala Leu Thr Asp Ser Gly Pro Lys Thr Asn Ile Ile Val Asp Ser
 305 310 315 320

Gln Tyr Val Met Gly Ile Ser Lys Pro Ser Phe Gln Glu Phe Val Asp
 325 330 335

Trp Glu Asn Val Ser Pro Glu Leu Asn Ser Thr Asp Gln Pro Phe Trp
 340 345 350

Gln Ala Gly Ile Leu Ala Arg Asn Leu Val Pro Met Val Ala Thr Val
 355 360 365

Gln Gly Gln Asn Leu Lys Tyr Gln Gly Gln Ser Leu Val Ile Ser Ala
 370 375 380

Ser Ile Ile Val Phe Asn Leu Leu Glu Leu Glu Gly Asp Tyr Arg Asp
 385 390 395 400

Asp Gly Asn Val Trp Val His Thr Pro Leu Ser Pro Arg Thr Leu Asn
 405 410 415

Ala Trp Val Lys Ala Val Glu Glu Lys Lys Gly Ile Pro Val His Leu
 420 425 430

Glu Leu Ala Ser Met Thr Asn Met Glu Leu Met Ser Ser Ile Val His
 435 440 445

[0208]

Gln Gln Val Arg Thr Tyr Gly Pro Val Phe Met Cys Leu Gly Gly Leu
 450 455 460

Leu Thr Met Val Ala Gly Ala Val Trp Leu Thr Val Arg Val Leu Glu
 465 470 475 480

Leu Phe Arg Ala Ala Gln Leu Ala Asn Asp Val Val Leu Gln Ile Met
 485 490 495

Glu Leu Cys Gly Ala Ala Phe Arg Gln Val Cys His Thr Thr Val Pro
 500 505 510

Trp Pro Asn Ala Ser Leu Thr Pro Lys Trp Asn Asn Glu Thr Thr Gln
 515 520 525

Pro Gln Ile Ala Asn Cys Ser Val Tyr Asp Phe Phe Val Trp Leu His
 530 535 540

Tyr Tyr Ser Val Arg Asp Thr Leu Trp Pro Arg Val Thr Tyr His Met
 545 550 555 560

Asn Lys Tyr Ala Tyr His Met Leu Glu Arg Arg Ala Lys Tyr Lys Arg
 565 570 575

Gly Pro Gly Pro Gly Ala Lys Phe Val Ala Ala Trp Thr Leu Lys Ala
 580 585 590

Ala Ala Gly Pro Gly Pro Gly Gln Tyr Ile Lys Ala Asn Ser Lys Phe
 595 600 605

Ile Gly Ile Thr Glu Leu Gly Pro Gly Pro Gly
 610 615

<210> 38
 <211> 228
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多核苷酸

[0209]

<400> 38
 atgcagatct tcgtgaagac cctgaccggc aagaccatca ccctagaggt ggagcccagt 60
 gacaccatcg agaacgtgaa ggccaagatc caggataaag agggcatccc ccctgaccag 120
 cagaggctga tctttgccgg caagcagctg gaagatggcc gcaccctctc tgattacaac 180
 atccagaagg agtcaaccct gcacctggtc cttegcctga gaggtggc 228

<210> 39
 <211> 228
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 人工序列的描述: 合成多核苷酸

<400> 39
 atgcagatct tcgtgaagac cctgaccggc aagaccatca ccctagaggt ggagcccagt 60
 gacaccatcg agaacgtgaa ggccaagatc caggataaag agggcatccc ccctgaccag 120
 cagaggctga tctttgccgg caagcagctg gaagatggcc gcaccctctc tgattacaac 180

	atccagaagg agtcaaccct gcacctggtc cttcgctga gaggtgcc	228
	<210> 40	
	<211> 78	
	<212> DNA	
	<213> 智人(Homo sapiens)	
	<400> 40	
	atggccgtca tggcgccccg aaccctegtc ctgetactct cgggggetct ggcctgacc	60
	cagacctggg cgggctct	78
	<210> 41	
	<211> 201	
	<212> DNA	
	<213> 智人(Homo sapiens)	
	<400> 41	
	ccgtcttccc agcccacat ccccatcgtg ggcacattg ctggcctgt tctctttgga	60
	gctgtgatca ctggagctgt ggtcgctgct gtgatgtgga ggaggaagag ctcatataga	120
[0210]	aaaggaggga gctactctca ggctgcaagc agtgacagtg cccagggtc tgatgtgtct	180
	ctcacagctt gtaaagtgtg a	201
	<210> 42	
	<211> 60	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成寡核苷酸	
	<400> 42	
	atggagaccg atacactgct gctgtgggtg ctgctcctgt ggggtccagg aagcacagcc	60
	<210> 43	
	<211> 3178	
	<212> DNA	
	<213> 智人(Homo sapiens)	
	<400> 43	
	ggcaccgatt cggggcctgc cggacttcg ccgcacgctg cagaacctcg cccagcgcgc	60
	accatgcccc ggcagctcag cgcggcggcc gcgctcttcg cgteccctggc cgtaattttg	120

	cacgatggca gtcaaatgag agcaaaagca tttccagaaa ccagagatta ttctcaacct	180
	actgcagcag caacagtaca ggacataaaa aaacctgtcc agcaaccagc taagcaagca	240
	cctcaccaaa ctttagcagc aagattcatg gatggtcata tcacctttca aacageggcc	300
	acagtaaaaa ttccaacaac taccacagca actacaaaaa aactgcaac caccagccca	360
	attacctaca ccctgggtcac aaccaggcc acaccaaca actcacacac agctcctcca	420
	gttactgaag ttacagtgg ccctagetta gccccttatt cactgccacc caccatcacc	480
	ccaccagctc atacagctgg aaccagtcca tcaaccgtca gccacacaac tgggaacacc	540
	actcaacca gtaaccagac caccctcca gcaactttat cgatagcact gcacaaaagc	600
	acaaccggtc agaagcctga tcaaccacc catgccccag gaacaacggc agctgccac	660
	aataccacc gcacagctgc acctgcctcc acggttcctg ggcccacct tgcacctcag	720
	ccatcgtcag tcaagactgg aatttatcag gttctaaacg gaagcagact ctgtataaaa	780
	gcagagatgg ggatacagct gattgttcaa gacaaggagt cggtttttc acctcggaga	840
	tacttcaaca tgaccccaa cgcaacgcaa gcctctggga actgtggcac ccgaaaatcc	900
[0211]	aaccttctgt tgaattttca gggcggattt gtgaatctca catttaccaa ggatgaagaa	960
	tcatattata tcagtgaagt gggagcctat ttgacctct cagatccaga gacagtttac	1020
	caaggaatca aacatgcggt ggtgatgttc cagacagcag tcgggcattc cttcaagtgc	1080
	gtgagtgaac agagcctcca gttgtcagcc cacctgcagg tgaaaacaac cgatgtccaa	1140
	cttcaagcct ttgattttga agatgaccac tttggaaatg tggatgagtg ctcgtctgac	1200
	tacacaattg tgcttctgt gattggggcc atcgtggttg gtctctgcct tatgggtatg	1260
	ggtgtctata aaatccgctt aagggtcaa tcatctggat accagagaat ctaattgttg	1320
	ccccggggga atgaaaataa tggaatttag agaactctt catccctcc aggatggatg	1380
	ttgggaaatt ccctcagagt gtgggtcctt caaacaatgt aaaccacat cttctattca	1440
	aatgaagtga gtcatgtgtg atttaagtcc aggcagcaca tcaatttcta aatactttt	1500
	gtttatttta tgaaagatat agtgagctgt ttattttcta gtttcctta gaatatttta	1560
	gccactcaaa gtcaacattt gagatatgtt gaattaacat aatatatgta aagtagaata	1620
	agccttcaaa ttataaacca agggccaatt gtaactaata ctactgtgtg tgcattgaag	1680

	atattatattt acccttgatc ttaacaaagc ctttgctttg ttatcaaatg gactttcagt	1740
	gcttttacta tctgtgtttt atggtttcat gtaacataca tattcctggt gtagcactta	1800
	actccttttc cactttaaat ttgtttttgt tttttgagac ggagtttcac tcttgtcacc	1860
	caggctggag tacagtggca cgatctggc ttatggcaac ctccgcctcc cgggttcaag	1920
	tgattctcct gcttcagctt cccgagtagc tgggattaca ggcacacact accacgcctg	1980
	gctaattttt gtatttttat tatagacggg tttcaccatg ttggccagac tggctttgaa	2040
	ctcttgacct caggtgatcc acccacctca gcctcccaaa gtgctgggat tacaggeatg	2100
	agccattgcg cccggcctta aatgtttttt ttaatcatca aaaagaaca catatctcag	2160
	gttgcttaag tgtttttatg taaaaccaac aaaaagaaca aatcagctta tattttttat	2220
	cttgatgact cctgctccag aattgctaga ctaagaatta ggtggctaca gatggtagaa	2280
	ctaacaata agcaagagac aataataatg gcccttaatt attaacaag tgccagagtc	2340
	taggctaagc actttatcta tatctcattt cattctcaca acttataagt gaatgagtaa	2400
[0212]	actgagactt aagggaactg aatcacttaa atgtcacctg gctaactgat ggcagagcca	2460
	gagcttgaat tcatgtttgt ctgacatcaa ggtcttttgg cttctcccta caccaagtta	2520
	cctacaagaa caatgacacc aactctgcc tgaaggctca cacctcatac cagcatagc	2580
	tcacettaca gggaaatggg tttatccagg atcatgagac attaggtag atgaaaggag	2640
	agctttgcag ataacaaaat agcctatcct taataaatcc tccactctct ggaaggagac	2700
	tgaggggctt tgtaaaacat tagtcagttg ctcatTTTTA tgggattgct tagctggget	2760
	gtaaagatga aggcacaaa taaactcaaa gtatttttaa attttttga taatagagaa	2820
	acttegetaa ccaactgttc tttcttgagt gtatagcccc atcttgggt aacttgctgc	2880
	ttctgcactt catatccata tttctattg ttcactttat tctgtagagc agcctgccaa	2940
	gaattttatt tctgctgttt tttttgctgc taaagaaagg aactaagtca ggatgttaac	3000
	agaaaagtcc acataaccct agaattctta gtcaaggaat aattcaagtc agcctagaga	3060
	ccatgttgac tttctcatg tgtttcctta tgactcagta agttggcaag gtctgactt	3120
	tagtcttaat aaaacattga attgtagtaa aggtttttgc aataaaaact tactttgg	3178

<210> 44	
<211> 1858	
<212> DNA	
<213> 小鼠属某种 (Mus sp.)	
<400> 44	
attccggagg tgaaaaacaa tggcacaacg tgtataatgg ccagcttctc tgctccttt	60
ctgaccacct acgagactgc gaatggttct cagatcgtga acatttcctt gccagcctct	120
gcagaagtac tgaaaaatgg cagtcttctg ggtaaagaaa atgtttctga ccccagcctc	180
acaattactt ttggaagagg atatttactg acaactcaact tcacaaaaaa tacaacacgt	240
tacagtgtcc agcatatgta ttttacatat aacttgtcag atacagaaca ttttcccaat	300
gccatcagca aagagatcta caccatggat tccacaactg acatcaaggc agacatcaac	360
aaagcatacc ggtgtgtcag tgatatccgg gtctacatga agaattgtac cgttgtgctc	420
cgggatgcca ctatccagc ctacctgtcg agtggcaact tcagcaagga agagacacac	480
tgcacacagg atggacctc cccaaccact gggccacca gccctcacc accacttgtg	540
[0213] cccacaaacc ccaactgtatc caagtacaat gttacttgta acaacggaac ctgcctgctg	600
gcctctatgg cactgcaact gaatatcacc tacctgaaaa aggacaacaa gacggtgacc	660
agagcgttca acatcagccc aatgacaca tctagtggga gttgcggtat caacttgggtg	720
accctgaaag tggagaacaa gaacagagcc ctggaattgc agtttgggat gaatgccagc	780
tctagcctgt ttttcttgca aggagtgcgc ttgaatatga ctcttctga tgccttagtg	840
cccacattca gcatctcaa ccattcactg aaagctcttc aggccactgt gggaaactca	900
tacaagtgca aactgagga acacatcttt gtcagcaaga tgctctccct caatgtcttc	960
agtgtgcagg tccaggcttt caaggtggac agtgacaggt ttgggtctgt ggaagagtgt	1020
gttcaggatg gtaacaacat gttgatcccc attgctgtgg gcggtgccct ggcagggctg	1080
atctccteg tcctcattgc ctacctcatt ggcaggaaga ggagtcacgc cgctatcag	1140
accatctagc ctggtgggca ggtgcaccag agatgcacag gggcctgttc tcacatcccc	1200
aagcttagat aggtgtggaa gggaggcaca ctttctgca aactgtttta aaatctgctt	1260
tatcaaatgt gaagttcatc ttgcaacatt tactatgcac aaaggaataa ctattgaaat	1320

	gacgggtgta attttgctaa ctgggttaaa tattgatgag aaggctccac tgatttgact	1380
	tttaagactt ggtgttttgg tcttcattct tttactcaga tttaagccta tcaaagggat	1440
	actctgggcc agaccttggc ctggcaaggg tggtgatgg ttaggctgca cacacttaag	1500
	aagcaacggg agcaggaag gcttgcacac aggcacgcac aggtcaacc tctggacact	1560
	tggttgggc tacctggcct tgggggggct gaactctggc atctggctgg gtacacacce	1620
	ccccaatttc tgtgctctgc caccctgag ctgccacttt cctaaataga aaatggcatt	1680
	atthttattt actthtttgt aaagtgatt ccagcttctg gttggcgttc aggttgccc	1740
	tgtctctgca ctgtgtacaa taatagattc aactctgta cgtgtcttgc agcgtaggtg	1800
	ggttgtagac tgggcatcag ctacacgtaa gcattgcctg taacgatgct aataaaaa	1858
	<210> 45	
	<211> 2339	
	<212> DNA	
	<213> 智人(Homo sapiens)	
[0214]	<400> 45	
	ggcccaaccg ccgcccgcgc ccccgtcttc cgcaccgtac ccggcccct cgcgccatgg	60
	cgccccccgg cagcgcggcg cgacccttc tgctgtact gctgttctg ctgctcggcc	120
	tcatgcattg tgcgtcagca gcaatgttta tggtagaaaa tggcaacggg accgcgtgca	180
	taatggccaa cttctctgct gccttctcag tgaactacga caccaagagt ggcctaaga	240
	acatgacctt tgacctgcca tcagatgcca cagtggtgct caaccgcagc tctgtggaa	300
	aagagaacac ttctgacccc agtctcgtga ttgcttttgg aagaggacat aactcactc	360
	tcaatttcac gagaaatgca acacgttaca ggtccagct catgagtttt gttataact	420
	tgtagacac acacctttc ccaatgca gctccaaaga aatcaagact gtggaatcta	480
	taactgacat cagggcagat atagataaaa aatacagatg tgttagtggc acccaggtcc	540
	acatgaacaa cgtgaccgta acgctccatg atgccacat ccaggcgtac cttccaaca	600
	gcagcttcag caggggagag acacgctgtg aacaagacag gccttccca accacagcgc	660
	cccctgcgcc acccagcccc tcgccctcac ccgtgcccga gagcccctct gtggacaagt	720
	acaacgtgag cggcaccaac gggacctgcc tgctggccag catggggctg cagctgaacc	780

	tcacctatga gaggaaggac aacacgacgg tgacaaggct tctcaacatc aacccaaca	840
	agacctcggc cagcgggagc tgcggcggcc acctggtgac tctggagctg cacagcgagg	900
	gcaccaccgt cctgetcttc cagtccggga tgaatgcaag ttctagccgg ttttctctac	960
	aaggaatcca gttgaataca attcttctctg acgccagaga ccctgccttt aaagctgcca	1020
	acggctccct gcgagcgctg caggccacag tcggcaattc ctacaagtgc aacgcggagg	1080
	agcacgtccg tgtcacgaag gcgttttcag tcaatatatt caaagtgtgg gtccaggctt	1140
	tcaagggtgga aggtggccag tttggtctctg tggaggagtg tctgctggac gagaacagca	1200
	tgctgatecc catcgtctgt ggtgggtgcc tggcggggct ggtcctcctc gtcctcctcg	1260
	cctacctcgt cggcaggaag aggagtcacg caggctacca gactatctag cctggtgcac	1320
	gcaggcacag cagctgcagg ggccctctgt cctttctctg ggcttagggt cctgtcgaag	1380
	gggaggcaca ctttctggca aacgtttctc aaatctgctt catccaatgt gaagttcctc	1440
	ttgcagcatt tactatgcac aacagagtaa ctatcgaat gacgggtgta attttgctaa	1500
	ctgggttaaa tattttgcta actgggtaaa cattaatatt taccaaagta ggattttgag	1560
[0215]	ggtgggggtg ctctctctga ggggggtggg gtgccgctgt ctctgagggg tgggggtgcc	1620
	gctgtctctg aggggtggg gtgccgctct ctctgagggg gtgggggtgc cgtttctct	1680
	gaggggtgg ggggtcccgt ctctctgagg ggggtgggggt gctgtctctc ccgaggggtg	1740
	gaatcccgt gtctctgagg ggtgggggtg ccgctctaaa ttggtccat atcatttgag	1800
	tttagggttc tgggttttg tttcttcatt ctttactgca ctcagattta agccttaca	1860
	agggaaagcc tctggccgtc acacgtagga cgcagtaagg tcaactcgtg tgaggctgac	1920
	atgctcacac attacaacag tagagagga aaatcctaag acagaggaac tccagagatg	1980
	agtgtctgga gcgcttcagt tcagctttaa aggccaggac gggccacacg tgcttgccg	2040
	cctcgttcca gtggcggcac gtccttgggc gtctctaatg tctgcagctc aagggtggc	2100
	acttttttaa atataaaaat ggggtgttatt tttatTTTTT ttgttaaagt gatttttgg	2160
	cttctgttga cattcggggt gatcctgttc tgcgctgtgt acaatgtgag atcgggtcgt	2220
	tctcctgatg ttttcccgtg gcttggggat tgtacacggg accagctcac gtaatgcatt	2280
	gcctgtaaca atgtaataaa aagcctcttt cttttaaaa aaaaaaaaa aaaaaaaaa	2339

<210> 46
 <211> 45
 <212> DNA
 <213> 人工序列(Artificial Sequence)

 <220>
 <223> 人工序列的描述: 合成寡核苷酸

 <400> 46
 cagtacatca aggccaacag caagttcatc ggcatcaccg aactc

45

<210> 47
 <211> 15
 <212> PRT
 <213> 人工序列(Artificial Sequence)

 <220>
 <223> 人工序列的描述: 合成肽

 <400> 47
 Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Leu
 1 5 10 15

[0216]

<210> 48
 <211> 39
 <212> DNA
 <213> 人工序列(Artificial Sequence)

 <220>
 <223> 人工序列的描述: 合成寡核苷酸

 <400> 48
 gctaaatttg tggctgcctg gacactgaaa gccgcccgt

39

<210> 49
 <211> 13
 <212> PRT
 <213> 人工序列(Artificial Sequence)

 <220>
 <223> 人工序列的描述: 合成肽

 <400> 49
 Ala Lys Phe Val Ala Ala Trp Thr Leu Lys Ala Ala Ala
 1 5 10

<210> 50	
<211> 593	
<212> DNA	
<213> 土拨鼠肝炎病毒(Woodchuck hepatitis virus)	
<400> 50	
aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttget	60
ccttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat tgcttcccg	120
atggctttca ttttctctc cttgtataaa tcttggttgc tgtctcttta tgaggagt	180
tggcccgttg tcagcaacg tggcgtggtg tgcactgtgt ttgctgacgc aacccccact	240
ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgcttt cccctcct	300
attgccacgg cggaactcat cgccgctgc cttgccgct gctggacagg ggctcggtg	360
ttggcactg acaattccgt ggtgttgcg gggaagctga cgtccttcc atggctgctc	420
gcctgtgttg ccacctggat tctgcgctg acgtccttct gctacgtccc ttcggcctc	480
aatccagcgg accttcctc ccgcgccctg ctgccgctc tgcggcctct tccgctctt	540
[0217] cgcttcgcc ctcagacgag tggatctcc ctttgggccc cctccccgcc tgt	593
<210> 51	
<211> 589	
<212> DNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 51	
tctcccccc cccctctcc ctccccccc ctaacgtta ctggccgaag ccgcttgaa	60
taagccggt gtgcgtttgt ctatatgta tttccacca tattgccgtc tttggcaat	120
gtgagggccc ggaaacctgg ccctgtcttc ttgacgagca ttctagggg tctttccct	180
ctcgccaaag gaatgcaagg tetgttgaat gtcgtgaagg aagcagttcc tctggaagct	240
tcttgaagac aaacaacgtc tgtagcgacc ctttgcaggc agcggaacct cccacctggc	300
gacaggtgcc tctgcggcca aaagccact gtataagata cacctgcaaa ggcggcacia	360
ccccagtgcc acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctctcaagc	420

	gtattcaaca aggggctgaa ggatgccag aaggtacccc attgtatggg atctgatctg	480
	gggcctcggg gcacatgctt tacatgtgtt tagtcgaggt taaaaaacg tctaggcccc	540
	ccgaaccacg gggacgtggt tttccttga aaaacacgat gataatatg	589
	<210> 52	
	<211> 720	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 52	
	atggtgagca agggcgagga gctgttcacc ggggttggtc ccatcctggt cgagctggac	60
	ggcgacgtaa acggccacaa gttcagcgtg tccggcgagg gcgagggcga tgccacctac	120
	ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggcccacc	180
	ctcgtgacca ccctgacctc cggcgtgcag tgettccagc gctaccccga ccacatgaag	240
	cagcagcact tcttcaagtc cgccatgccc gaaggctacg tccaggagcg caccatcttc	300
[0218]	ttcaaggacg acggcaacta caagaccgc gccgaggtga agttcgagg gcacacctg	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 tgetgctgcc cgacaaccac	600
	tacctgagca cccagtccgc cctgagcaaa gacccaacg agaagcgcga tcacatggtc	660
	ctgctggagt tcgtgaccgc cgccgggatc actctcgcca tggacgagct gtacaagtag	720
	<210> 53	
	<211> 1563	
	<212> DNA	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 人工序列的描述: 合成多核苷酸	
	<400> 53	
	atgctgctgc tgctgctgct gctggcctg aggctacagc tctcctggg catcatccca	60

	gttgaggagg agaaccgga cttctggaac cgcgaggcag ccgaggccct gggtgccgcc	120
	aagaagctgc agcctgcaca gacagccgcc aagaacctca tcattctcct gggcgatggg	180
	atgggggtgt ctacggtgac agctgccagg atcctaaaag ggcagaagaa ggacaaactg	240
	ggcctgaga taccctggc catggaccgc ttcccatatg tggctctgtc caagacatac	300
	aatgtagaca aacatgtgcc agacagtga gccacagcca cgccctacct gtgcggggtc	360
	aagggaact tccagacat tggcttgagt gcagccgcc gctttaacca gtgcaacacg	420
	acacgcggca acgaggtcat ctccgtgatg aatcgggcca agaaagcagg gaagtcagt	480
	ggagtggtaa ccaccacacg agtgcagcac gcctgccag ccggcaccta cggccacacg	540
	gtgaaccgca actggtactc ggacccgac gtgcctgcct cggcccgcca ggaggggtgc	600
	caggacatcg ctacgcagct catctccaac atggacattg acgtgatcct aggtggaggc	660
	cgaaagtaca tgtttcgcat gggaaccca gacctgagt acccagatga ctacagccaa	720
	ggtgggacca ggctggacgg gaagaatctg gtgcaggaat ggctggcgaa gcgccagggt	780
[0219]	gcccgtatg tgtggaaccg cactgagctc atgcaggctt ccctggacce gtctgtgacc	840
	catctcatgg gtctctttga gcctggagac atgaaatacg agatccaccg agactccaca	900
	ctggaccct ccctgatgga gatgacagag gctgcctgc gcctgctgag caggaacccc	960
	cgcgcttct tcctcttctg ggaggggtgt cgcacgacc atggtcatca tgaagcagg	1020
	gcttaccggg cactgactga gacgatcatg ttcgacgac ccattgagag ggcgggccag	1080
	ctcaccagcg aggaggacac gctgagcctc gtcactgcc accactccca cgtcttctcc	1140
	ttcggaggct acccctgcg agggagctcc atcttcgggc tggcccctgg caaggcccgg	1200
	gacaggaagg cctacacggt cctctatac ggaaacggte caggctatgt gctcaaggac	1260
	ggcggccggc cggatgttac cgagagcgag agcgggagcc ccgagtatcg gcagcagtca	1320
	gcagtgcccc tggacgaaga gaccacgca ggcgaggacg tggcgggtgt cgcgcgcggc	1380
	ccgcaggcgc acctggttca cggcgtgcag gagcagacct tcatagcgca cgtcatggcc	1440
	ttcggcct gctggagcc ctacaccgcc tgcacctgg cggccccgc cggcaccacc	1500
	gacggcgc acccgggtta ctctagagtc gggcgggccg gccgcttca gcagacatga	1560

taa	1563
<210> 54	
<211> 1653	
<212> DNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 54	
atggaagatg ccaaaaaacat taagaagggc ccagcgcctat tctaccact cgaagacggg	60
accgccggcg agcagctgca caaagccatg aagcgcctac ccctgggtgcc cggcaccatc	120
gcctttaccg acgcacatat cgaggtggac attacctacg ccgagtactt cgagatgagc	180
gttcggctgg cagaagctat gaagcgcctat gggtgaata caaacctcg gatcgtggtg	240
tgcagcgaga atagcttgca gttcttcatg cccgtgttgg gtgccctgtt catcggtgtg	300
gctgtggccc cagctaacga catctacaac gagcgcgagc tgetgaacag catgggcac	360
agccagccca ccgtcgtatt cgtgagcaag aaagggtgc aaaagatcct caacgtgcaa	420
[0220] aagaagctac cgatcataca aaagatcacc atcatggata gcaagaccga ctaccagggc	480
ttcaaagca tgtacacctt cgtgacttcc catttggcac ccggttcaa cgagtacgac	540
ttcgtgcccg agagcttcca ccgggacaaa accatcggcc tgatcatgaa cagtagtggc	600
agtaccgat tgccaaggg cgtagcccta ccgcaccgca ccgcttgtgt ccgattcagt	660
catgcccgcg accccatctt cggcaaccag atcatccccg acaccgctat cctcagcgtg	720
gtgccatttc accacggctt cggcatgttc accacgctgg gctacttgat ctgcggcttt	780
cggtcgtgc tcatgtaccg cttcaggag gagctattct tgcgcagctt gcaagactat	840
aagattcaat ctgccctgct ggtgccca caatttagct tcttcgctaa gagcactctc	900
atcgacaagt acgacctaac caacttgac gagatgccca gcggcggggc gccgctcagc	960
aaggaggtag gtgaggccgt ggccaaacgc ttccacctac caggcatccg ccagggtac	1020
ggcctgacag aaacaaccag cgccattctg atcacccccg aaggggacga caagcctggc	1080
gcagtaggca aggtggtgcc cttcttcgag gctaaggtgg tggacttga caccggtgtaag	1140
acactgggtg tgaaccagcg cggcgagctg tgcgtccgtg gccccatgat catgagcggc	1200

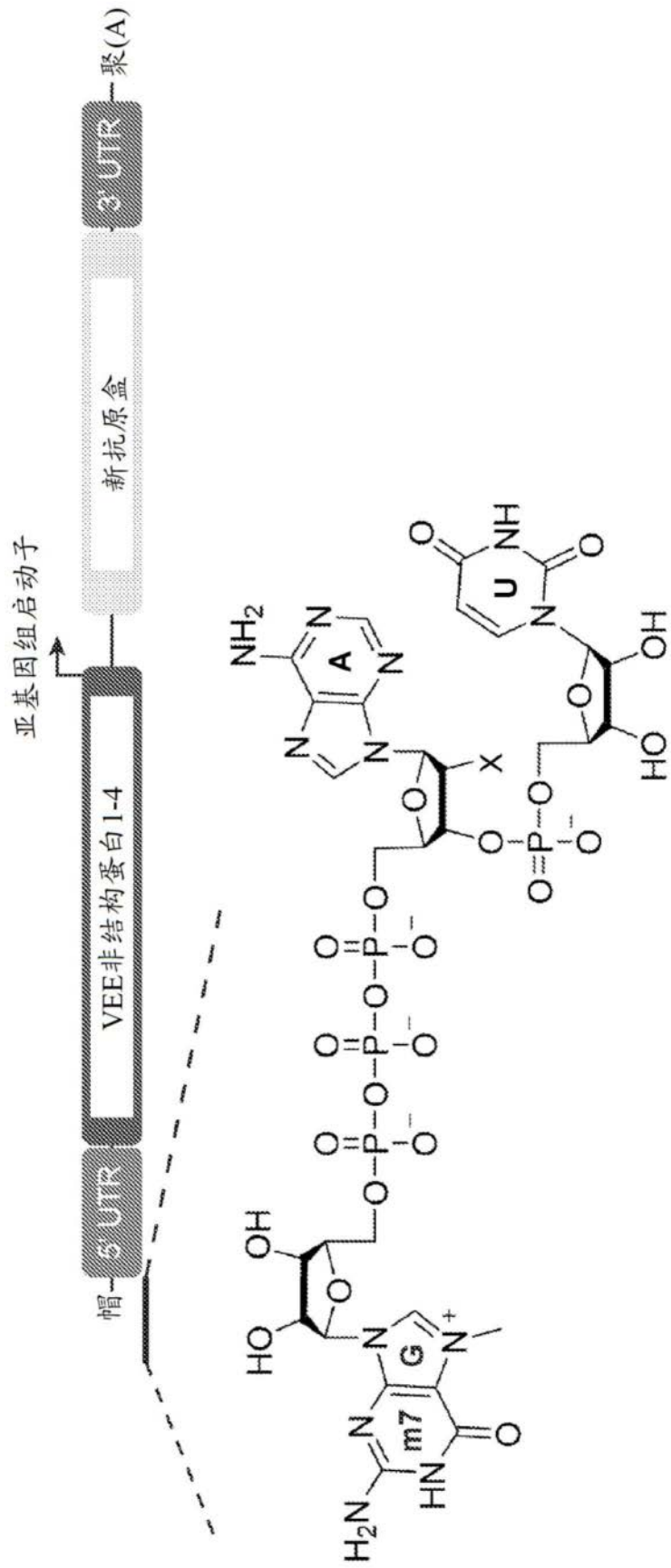


图2

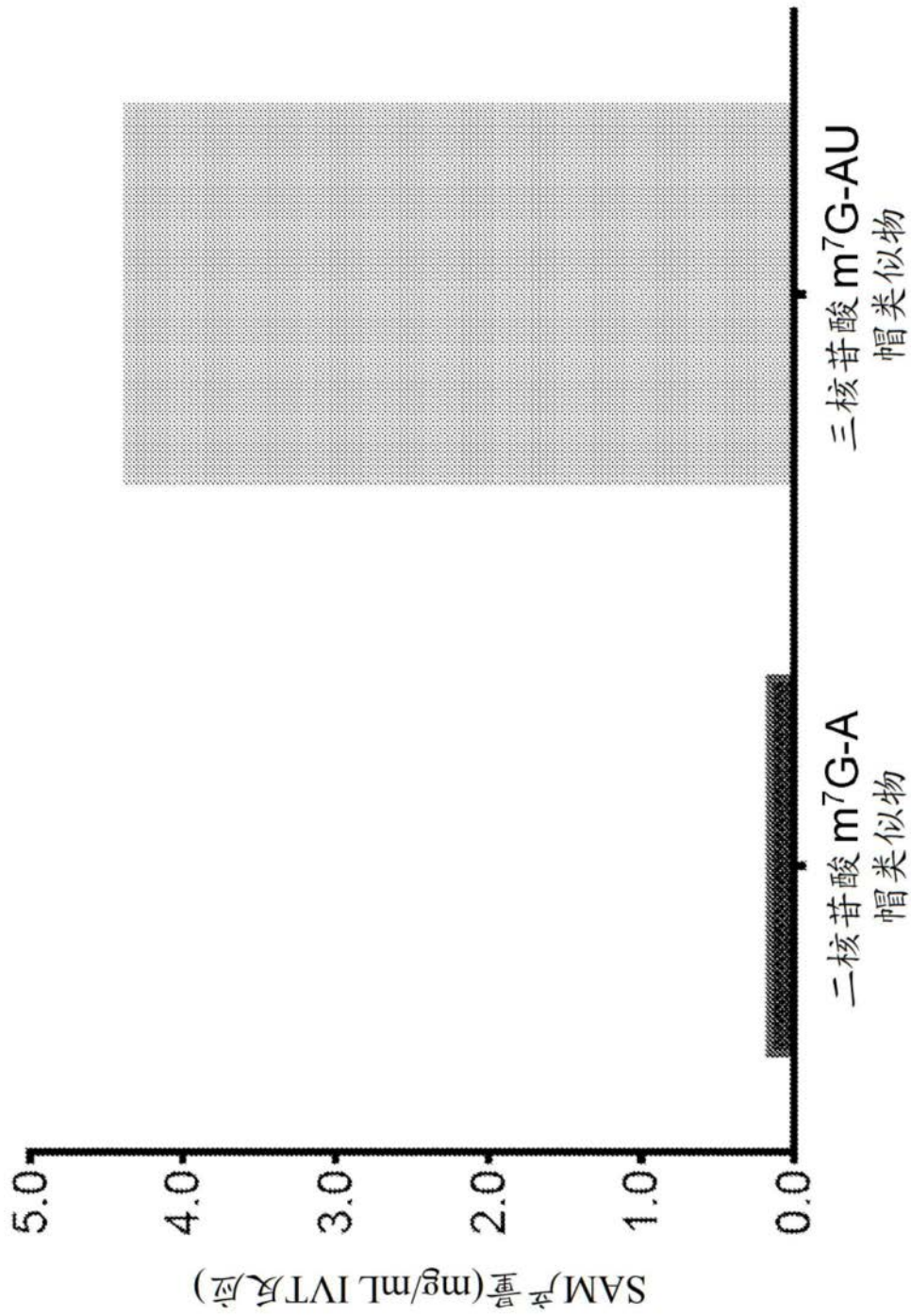


图3

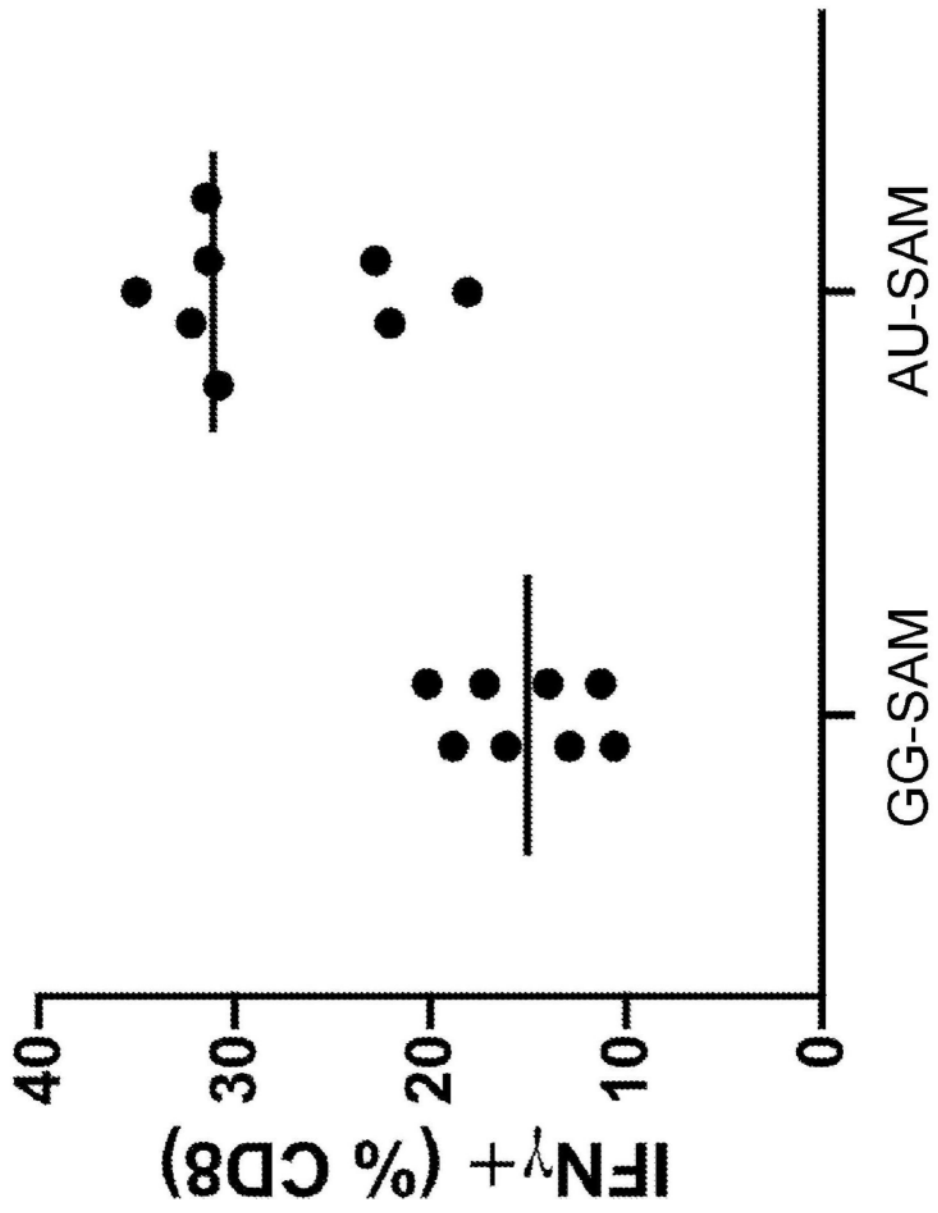


图4

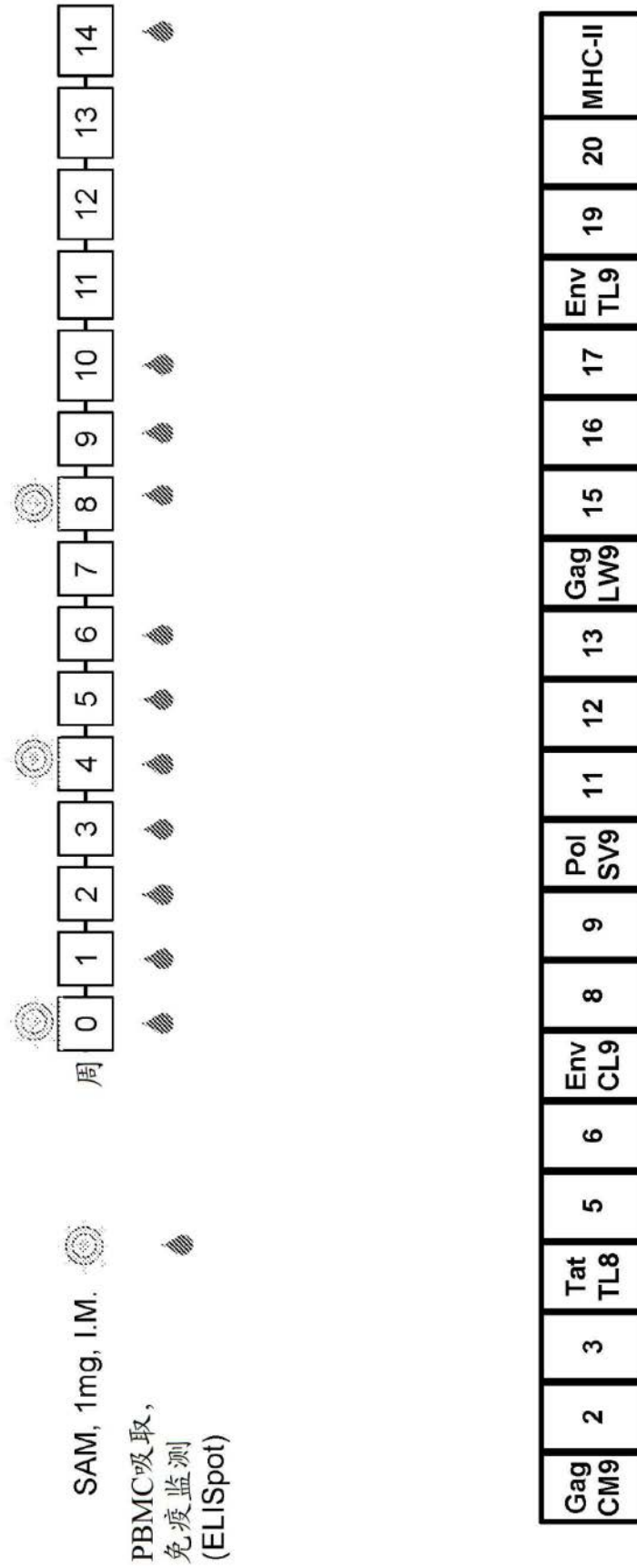


图5

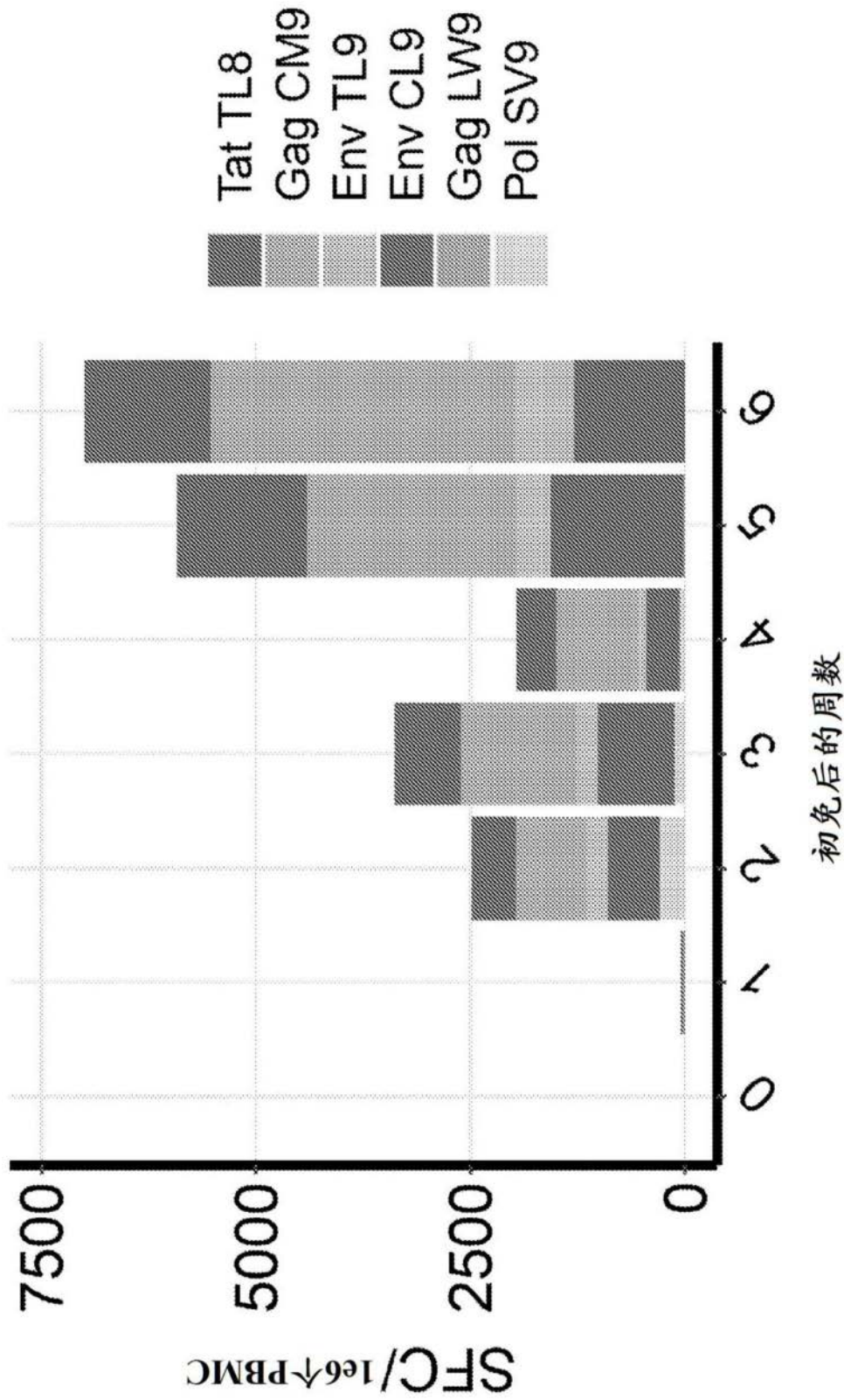


图6

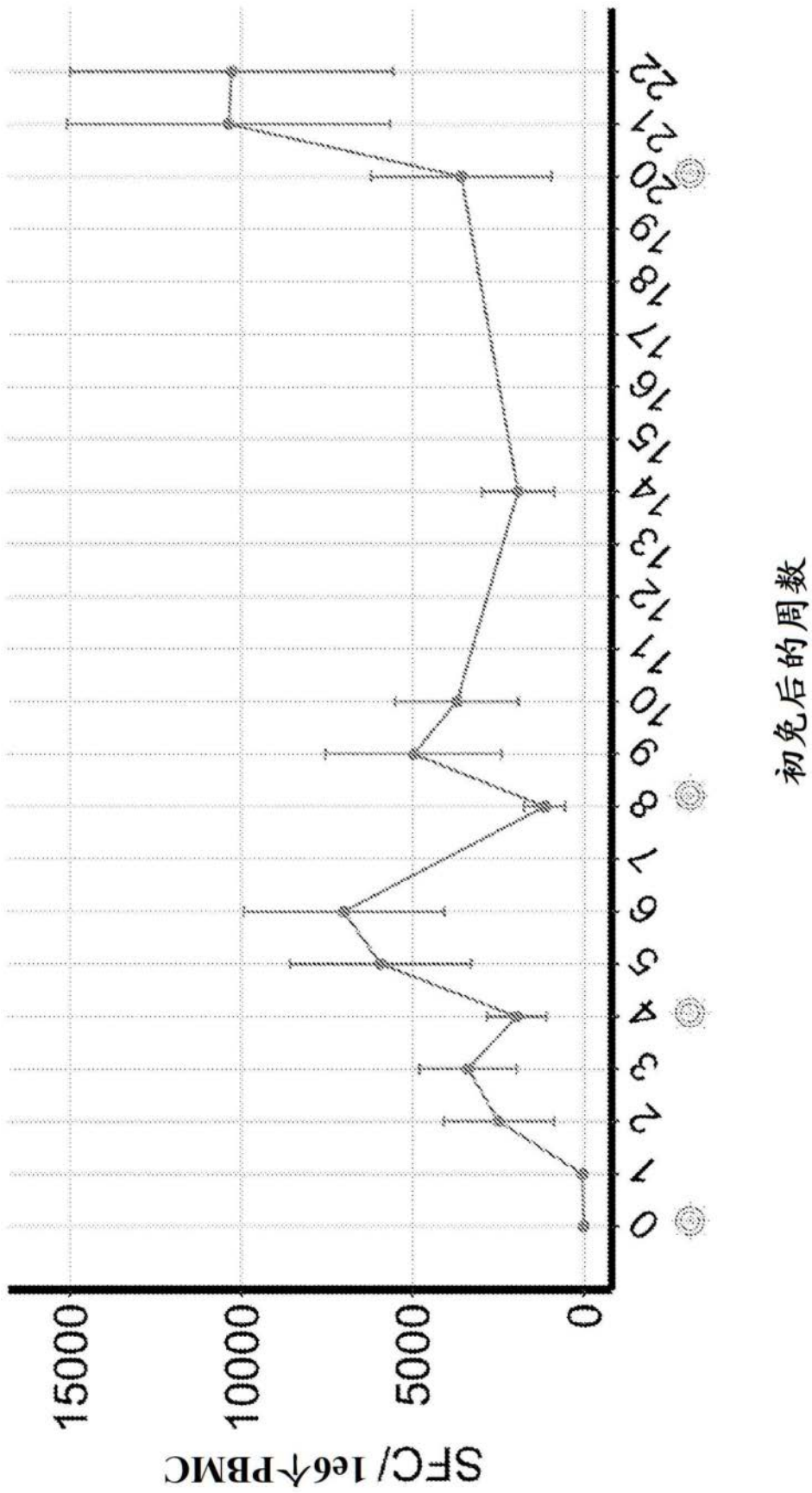


图7