



(12)发明专利申请



(10)申请公布号 CN 110036020 A

(43)申请公布日 2019.07.19

(21)申请号 201780074503.5

(22)申请日 2017.12.08

(30)优先权数据

62/432033 2016.12.09 US

62/465378 2017.03.01 US

(85)PCT国际申请进入国家阶段日
2019.05.31

(86)PCT国际申请的申请数据

PCT/IB2017/057759 2017.12.08

(87)PCT国际申请的公布数据

W02018/104919 EN 2018.06.14

(71)申请人 葛兰素史密丝克莱恩生物有限公司

地址 比利时里克森萨特

(72)发明人 V.阿门多拉 S.科罗卡 A.维特利
B.魏策尔

(74)专利代理机构 中国专利代理(香港)有限公
司 72001

代理人 李唐 周齐宏

(51)Int.Cl.

C07K 14/005(2006.01)

A61K 39/205(2006.01)

C12N 15/86(2006.01)

权利要求书7页 说明书40页
序列表194页 附图22页

(54)发明名称

具有狂犬病病毒抗原的黑猩猩腺病毒构建
体

(57)摘要

本发明提供了包含编码狂犬病病毒抗原的
转基因的腺病毒载体。所述载体可用于生产疫苗
用于预防、改善和治疗由狂犬病病毒疾病引起的
疾病,例如狂犬病。

1. 重组腺病毒,其包含选自以下的至少一种多核苷酸或多肽:

(a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,

(c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸,

(d) 具有根据SEQ ID NO:1的氨基酸序列的多肽;

(e) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(f) 具有根据SEQ ID NO:3的氨基酸序列的多肽;

其中所述腺病毒包含编码狂犬病病毒抗原的核酸序列,其中所述核酸序列与指导所述狂犬病病毒抗原在宿主细胞中表达的一种或多种序列可操作地连接。

2. 组合物,包含根据权利要求1的重组腺病毒和药学上可接受的赋形剂。

3. 根据权利要求1或2的腺病毒或组合物,其包含(a)编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸。

4. 根据权利要求3的腺病毒或组合物,其中所述多核苷酸具有根据SEQ ID NO:2的序列。

5. 根据权利要求1或2的腺病毒或组合物,其包含(b)编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列。

6. 根据权利要求5的腺病毒或组合物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少89.0%同一性,例如至少99.0%同一性,例如至少99.6%同一性的氨基酸序列。

7. 根据权利要求3至6中任一项的腺病毒或组合物,其还包含编码以下的多核苷酸:

(i) 具有根据SEQ ID NO:3的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列,

或

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

8. 根据权利要求1至3中任一项的腺病毒或组合物,其还包含编码以下的多核苷酸:

(i) 具有根据SEQ ID NO:3的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列,

和

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

9. 根据权利要求7或8的腺病毒或组合物,其中具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少70.0%,例如至少90.0%,例如至少99%同一性的氨基酸序列。

10. 根据权利要求7至9中任一项的腺病毒或组合物,其中具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少70.0%,例如至少90.0%,例如至少99%同一性的氨基酸序列。

11. 根据权利要求1或2的腺病毒或组合物,其包含(c)编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸。

12. 根据权利要求11的腺病毒或组合物,其中所述多核苷酸具有根据SEQ ID NO:4的多核苷酸序列。

13. 根据权利要求11或12的腺病毒或组合物,其还包含编码以下的多核苷酸:

(i) 具有根据SEQ ID NO:1的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少50%同一性的氨基酸序列,

或

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

14. 根据权利要求11或12的腺病毒或组合物,其还包含编码以下的多核苷酸:

(i) 具有根据SEQ ID NO:1的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少50%同一性的氨基酸序列,

和

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

15. 根据权利要求13或14的腺病毒或组合物,其中具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少70.0%同一性,例如至少99.0%同一性,例如至少99.4%同一性的氨基酸序列。

16. 根据权利要求13至15中任一项的腺病毒或组合物,其中具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少70.0%同一性,例如至少90.0%同一性,例如至少99%同一性的氨基酸序列。

17. 根据权利要求1或2的腺病毒或组合物,其包含(d)具有根据SEQ ID NO:1的氨基酸序列的多肽。

18. 根据权利要求1或2的腺病毒或组合物,其包含(e)其为具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多肽,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列。

19. 根据权利要求18的腺病毒或组合物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少89.0%同一性,例如至少99.0%同一性,例如至少

99.6%同一性的氨基酸序列。

20. 根据权利要求17至19中任一项的腺病毒或组合物, 其还包含:

(i) 具有根据SEQ ID NO:3的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列,
或

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

21. 根据权利要求17至19中任一项的腺病毒或组合物, 其还包含:

(i) 具有根据SEQ ID NO:3的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列,
和

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

22. 根据权利要求20或21的腺病毒或组合物, 其中具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少70.0%, 例如至少90.0%, 例如至少99%同一性的氨基酸序列。

23. 根据权利要求20至22中任一项的腺病毒或组合物, 其中具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少70.0%, 例如至少90.0%, 例如至少99%同一性的氨基酸序列。

24. 根据权利要求1或2的腺病毒或组合物, 其包含(f) 具有根据SEQ ID NO:3的氨基酸序列的多肽。

25. 根据权利要求24所述的腺病毒或组合物, 其还包含:

(i) 具有根据SEQ ID NO:1的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少50%同一性的氨基酸序列,
或

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

26. 根据权利要求24的腺病毒或组合物, 其还包含:

(i) 具有根据SEQ ID NO:1的氨基酸序列的多肽;或
(ii) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物, 其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少50%同一性的氨基酸序列,
和

(i) 具有根据SEQ ID NO:5的氨基酸序列的多肽;或

(ii) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

27. 根据权利要求25或26的腺病毒或组合物,其中具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少70.0%同一性,例如至少99.0%同一性,例如至少99.4%同一性的氨基酸序列。

28. 根据权利要求25至27中任一项的腺病毒或组合物,其中具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少70.0%,例如至少90.0%,例如至少99%同一性的氨基酸序列。

29. 根据权利要求1-28中任一项的腺病毒或组合物,其中所述多核苷酸包含以下中的至少一种:

- 腺病毒5'-末端,优选腺病毒5'反向末端重复;
- 腺病毒E1A区或其选自E1A_280R和E1A_243R区的片段;
- 腺病毒E1B或IX区,或其选自E1B_19K、E1B_55K或IX区的片段;
- 腺病毒E2b区;或其选自E2B_pTP、E2B_聚合酶和E2B_IVa2区域的片段;
- 腺病毒L1区或其片段,所述片段编码选自L1_13.6k蛋白、L1_52k和L1_IIIa蛋白的腺病毒蛋白;
- 腺病毒L2区或其片段,所述片段编码选自L2_五邻体蛋白、L2_pVII、L2_V和L2_pX蛋白的腺病毒蛋白;
- 腺病毒L3区或其片段,所述片段编码选自L3_pVI蛋白、L3_六邻体蛋白和L3_蛋白酶蛋白的腺病毒蛋白;
- 腺病毒E2A区;
- 腺病毒L4区或其片段,所述片段编码选自L4_100k蛋白、L4_33k蛋白和蛋白L4_VIII的腺病毒蛋白;
- 腺病毒E3区或其选自E3 ORF1、E3 ORF2、E3 ORF3、E3 ORF4、E3 ORF5、E3 ORF6、E3 ORF7、E3 ORF8和E3 ORF9的片段;
- 腺病毒L5区或其片段,所述片段编码L5_纤突纤突蛋白;
- 腺病毒E4区或其选自E4 ORF7、E4 ORF6、E4 ORF4、E4 ORF3、E4 ORF2和E4 ORF1的片段;
- 腺病毒3'-末端,优选腺病毒3'反向末端重复;和/或
- 腺病毒VAI或VAII RNA区,优选来自除ChAd155之外的腺病毒、更优选来自Ad5的腺病毒VAI或VAII RNA区。

30. 根据权利要求29的腺病毒或组合物,其中所述多核苷酸包含以下中的至少一种:

- 腺病毒5'-末端,优选腺病毒5'反向末端重复;
- 腺病毒L1区或其片段,所述片段编码选自L1_13.6k蛋白、L1_52k和L1_IIIa蛋白的腺病毒蛋白;
- 腺病毒L2区或其片段,所述片段编码选自L2_五邻体蛋白、L2_pVII、L2_V和L2_pX蛋白的腺病毒蛋白;
- 腺病毒L3区或其片段,所述片段编码选自L3_pVI蛋白、L3_六邻体六邻体蛋白和L3_蛋白酶的腺病毒蛋白;

● 腺病毒L4区或其片段,所述片段编码选自L4_100k蛋白、L4_33k蛋白和蛋白L4_VIII的腺病毒蛋白;

● 腺病毒L5区或其片段,所述片段编码L5_纤突纤突蛋白;

● 腺病毒3'-末端,优选腺病毒3'反向末端重复。

31. 根据权利要求1-28中任一项的腺病毒或组合物,其中所述多核苷酸包含腺病毒VAI或VAII RNA区。

32. 根据权利要求31的腺病毒或组合物,其中所述VAI或VAII RNA区来自除ChAd155之外的腺病毒。

33. 根据权利要求32的腺病毒或组合物,其中所述VAI或VAII RNA区域来自Ad5。

34. 根据权利要求1-28中任一项的腺病毒或组合物,其中编码狂犬病病毒抗原的核酸序列在其整个长度上与SEQ ID NO:38具有至少98.6%的同一性,例如至少99%的同一性,例如至少99.5%的同一性。

35. 根据权利要求34任一项的腺病毒或组合物,其中编码狂犬病病毒抗原的核酸序列在其整个长度上与SEQ ID NO:38相同。

36. 根据前述权利要求中任一项的腺病毒或组合物,其中所述多核苷酸包含突变或缺失,所述突变或缺失使得选自E1A、E1B、E2A、E2B、E3和E4的基因组区域的至少一个基因无功能。

37. 根据权利要求36的腺病毒或组合物,其中所述多核苷酸缺少选自E1A、E1B、E2A、E2B、E3和/或E4的基因组区域的至少一个基因。

38. 根据权利要求36或37的腺病毒或组合物,其中所述基因组区域是E1A和/或E1B。

39. 根据权利要求1至38中任一项的腺病毒或组合物,其中所述重组腺病毒是复制缺陷型的。

40. 根据权利要求1至39中任一项的腺病毒或组合物,其中所述核酸序列编码来自莫科拉病毒、杜文海格病毒、欧洲蝙蝠狂犬病病毒、欧洲蝙蝠狂犬病病毒 2或澳大利亚蝙蝠狂犬病病毒的免疫原。

41. 根据权利要求40的腺病毒或组合物,其中所述核酸序列编码来自选自CVS11、CVS-N2C、Evelyn Rokitniki Abelseth (ERA)、Flury、Pitman Moore和Wistar毒株的狂犬病病毒的免疫原。

42. 根据权利要求1至41中任一项的腺病毒或组合物,其中所述核酸序列编码源自狂犬病病毒糖蛋白(G)、RNA聚合酶(L)、基质蛋白(M)、核蛋白(N)或磷蛋白(P)的抗原。

43. 根据权利要求1至42中任一项的腺病毒或组合物,其中所述核酸序列编码这样的多肽,所述多肽包含狂犬病病毒糖蛋白(G)、RNA聚合酶(L)、基质蛋白(M)、核蛋白(N)和磷蛋白(P)或包含其片段,例如至少20、至少50、至少100、至少200、至少300、至少400、至少500或至少600个氨基酸的片段。

44. 根据权利要求1至42中任一项的腺病毒或组合物,其中所述核酸序列编码包含来自莫科拉病毒、杜文海格病毒、欧洲蝙蝠狂犬病病毒、欧洲蝙蝠狂犬病病毒 2或澳大利亚蝙蝠狂犬病病毒的糖蛋白的全部或片段的片段的多肽,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

45. 根据权利要求1至42中任一项的腺病毒或组合物,其中所述核酸序列编码包含来自

选自CVS11、CVS-N2C、Evelyn Rokitniki Abelseth (ERA)、Flury、Pitman Moore和Wistar毒株的狂犬病病毒的糖蛋白的全部或片段的多肽,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

46. 根据权利要求1至42中任一项的腺病毒或组合物,其中所述核酸序列编码包含来自多于一种狂犬病病毒的糖蛋白的全部或片段的多肽,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

47. 根据权利要求1至44中任一项的腺病毒或组合物,其中所述核酸序列编码包含来自选自CVS11、CVS-N2C、Evelyn Rokitniki Abelseth (ERA)、Flury、Pitman Moore和Wistar毒株的多于一种狂犬病病毒的糖蛋白的全部或片段的多肽,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

48. 根据权利要求1至42中任一项的腺病毒或组合物,其中所述核酸序列编码包含SEQ ID NO:37的全部或片段的多肽,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

49. 根据权利要求48的腺病毒或组合物,其中所述核酸序列编码SEQ ID NO:37的多肽。

50. 根据权利要求1至49中任一项的腺病毒或组合物,其中所述核酸序列编码狂犬病病毒抗原,所述狂犬病病毒抗原包含位点I、位点IIa、位点IIb、位点III、位点IV和位点a中的至少一个。

51. 根据权利要求50的腺病毒或组合物,其中所述核酸序列编码狂犬病病毒抗原,所述狂犬病病毒抗原包含对应于SEQ ID NO:37的位点I、位点IIa、位点IIb、位点III、位点IV和位点a中的至少一个。

52. 根据权利要求51的腺病毒或组合物,其中所述主体被狂犬病病毒感染。

53. 根据权利要求52的腺病毒或组合物,其中所述狂犬病病毒是选自RABV、ABLV、ARAV、BBLV、DUVV、EBLV-1、EBLV-2、IRKV、KHUV、LBV、MOKV、SHIV、WCBV和IKOV中的一种或多种的狂犬病病毒的毒株。

54. 根据权利要求1至53中任一项的腺病毒或组合物,其中所述核酸序列编码包含这样的序列的狂犬病病毒抗原,所述序列是包含位点I、位点IIa、位点IIb、位点III、位点IV和位点a中的至少一个的免疫原性片段(例如包含一个或多个T细胞表位)。

55. 根据权利要求54的腺病毒或组合物,其中狂犬病病毒抗原包含这样的序列,所述序列是至少20个氨基酸残基的免疫原性片段。

56. 权利要求1至55中任一项的腺病毒或组合物,其中指导所述产物在宿主细胞中表达的一种或多种序列包括选自转录起始、转录终止、启动子和增强子序列的一种或多种的序列。

57. 权利要求56的腺病毒或组合物,其中指导所述产物在宿主细胞中表达的一种或多种序列包括启动子序列。

58. 权利要求57的腺病毒或组合物,其中所述启动子序列选自内部启动子、天然启动子、RSV LTR启动子、CMV启动子、SV40启动子、二氢叶酸还原酶启动子、 β -肌动蛋白启动子、PGK启动子、EF1a启动子和CASI启动子。

59. 根据权利要求1至58中任一项的腺病毒或组合物,其中所述腺病毒在人主体中具有小于10%的血清阳性率,并且优选在人主体中不具有血清阳性率。

60. 根据权利要求1至59中任一项的腺病毒或组合物,其中所述腺病毒能够感染哺乳动物细胞。

61. 根据权利要求2至60中任一项的组合物,其包含佐剂。

62. 根据权利要求1至61中任一项的腺病毒或组合物,其用作药物。

63. 根据权利要求62的腺病毒或组合物,其用于刺激免疫应答。

64. 根据权利要求63的腺病毒或组合物,其用作疫苗。

65. 根据权利要求63的腺病毒或组合物,其用于预防、治疗或改善由狂犬病病毒引起的疾病。

66. 根据权利要求1至61中任一项的腺病毒或组合物在制备用于预防由狂犬病病毒感染引起的疾病的药物中的用途。

67. 根据权利要求1至61中任一项的腺病毒或组合物在制备用于在人中预防由狂犬病病毒感染引起的疾病的药物中的用途。

68. 在主体中诱导免疫应答的方法,其包括向主体施用根据权利要求1至61中任一项的腺病毒或组合物。

69. 根据权利要求68的方法,其中所述主体被狂犬病病毒感染。

70. 根据权利要求68或69的方法,其中所述主体是人。

71. 根据权利要求70的方法,其中所述狂犬病病毒是选自RABV、ABLV、ARAV、BBLV、DUVV、EBLV-1、EBLV-2、IRKV、KHUV、LBV、MOKV、SHIV、WCBV和IKOV的一种或多种的狂犬病病毒的毒株。

72. 非人猿猴腺病毒,其包含SEQ ID NO:3的五邻体、SEQ ID NO:5的六邻体或SEQ ID NO:1的纤突,并且还包含编码狂犬病病毒相关抗原的转基因。

73. 根据权利要求72的非人猿猴腺病毒,其包含来自ChAd155的五邻体(SEQ ID NO:3)、六邻体(SEQ ID NO:5)和纤突(SEQ ID NO:1)蛋白,并且还包含编码狂犬病病毒相关抗原的转基因。

74. 根据权利要求72或73的非人猿猴腺病毒,其中编码的抗原包含与SEQ ID NO. 37具有至少90%同一性的序列。

75. 根据权利要求73-74中任一项的非人猿猴腺病毒,其中所述转基因包含SEQ ID NO.38或由SEQ ID NO.38组成。

76. 根据权利要求70-75中任一项的非人猿猴腺病毒,其是复制缺陷型腺病毒。

77. 根据权利要求75的非人猿猴腺病毒,其中所述腺病毒包含E1基因的功能性失活(例如缺失)。

78. 根据权利要求75或77的非人猿猴腺病毒,其中所述腺病毒包含E4基因的功能性失活(例如缺失)。

79. 根据权利要求72-78中任一项的非人猿猴腺病毒,其中所述腺病毒包含E3基因的功能性失活(例如缺失)。

80. 根据权利要求72-79中任一项的非人猿猴腺病毒,其中所述腺病毒包含Ad5E4orf6基因取代。

具有狂犬病病毒抗原的黑猩猩腺病毒构建体

[0001] 序列表

本申请含有已经以ASCII形式电子提交且在此以其整体通过引用并入的序列表。创建于2017年12月6日的所述ASCII拷贝被命名为VU66242W0_SL.txt并且大小为394,117字节。

发明领域

[0002] 本发明属于改善疾病和治疗并预防病毒感染的领域。特别地,本发明涉及编码狂犬病病毒(*Lyssavirus*)抗原的黑猩猩腺病毒载体。它包括狂犬病病毒抗原用于改善狂犬病病毒的疾病(*Lyssaviral diseases*)以及治疗和预防狂犬病感染的用途。

[0003] 背景

狂犬病病毒是弹状病毒科(*Rhabdoviridae*)中的包膜单链RNA病毒。狂犬病病毒属的成员引起狂犬病并且具有所有已知人类病毒病原体的最高死亡率。狂犬病通过受感染的哺乳动物的唾液传播。它是一种嗜神经病毒,进入其宿主的神经系统,导致几乎总是致命的脑脊髓炎。目前,全世界每年约有60,000例狂犬病死亡,主要是由亚洲和非洲发展中国家的狗咬伤以及北美野生动物和蝙蝠引起的。

[0004] 狂犬病以狂躁或瘫痪的形式呈现。潜伏期在约五天至几年之间变化,但通常在约20至90天之间。临床疾病最通常始于不适、厌食、疲劳、头痛和发烧的前驱症状,然后在暴露部位出现疼痛或感觉异常。在此期间,焦虑、激动或烦躁可能是突出的,随后是极度活跃、方向障碍、癫痫发作、恐水、多涎,并最终导致瘫痪、昏迷和死亡。

[0005] 腺病毒由于它在多种靶组织中实现高效基因转移的能力和大的转基因容量已经被广泛地用于基因转移应用。按照惯例,将腺病毒E1基因缺失并用由所选启动子、目标基因的cDNA序列和聚腺苷酸信号组成的表达盒替换,从而产生复制缺陷型重组病毒。

[0006] 重组腺病毒可用于基因治疗和用作疫苗。基于非人猿猴腺病毒的病毒载体代表了使用人类来源载体开发基因疫苗的替代方案。从非人猿猴分离的某些腺病毒与从人类分离的腺病毒密切相关,如通过它们在人类起源的细胞中的有效繁殖所证明的。

[0007] 需要能够有效递送疫苗抗原的载体。具体而言,狂犬病仍然是全世界重要的病毒性人畜共患病。虽然目前可以进行预防,但在暴露前和暴露后都需要大量剂量,并且依从性(*compliance*)低,从而降低了医疗利益。需要改进的狂犬病疫苗,其具有简化的给药方案,增加的安全性和增强的制造特性。腺病毒制造比现有的人狂犬病疫苗更安全且更便宜,其基于灭活的狂犬病病毒。因此,存在开发用于狂犬病疫苗的腺病毒载体的未满足的要求。

[0008] 发明概述

本发明人提供了可用作免疫原性组合物(所述免疫原性组合物用于在主体中诱导针对狂犬病病毒的疾病和狂犬病病毒感染的免疫应答)组分的构建体,其用于治疗的方法及其制备方法。

[0009] 提供了分离的多核苷酸,其中所述多核苷酸编码选自以下的多肽:

- (a) 具有根据SEQ ID NO:1的氨基酸序列的多肽;
- (b) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物

具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(c) 具有根据SEQ ID NO:3的氨基酸序列的多肽;

其中分离的多核苷酸包含编码狂犬病病毒抗原的核酸序列。

[0010] 还提供了重组多核苷酸,其包含选自以下的多核苷酸:

(a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸;

其中所述重组多核苷酸包含编码狂犬病病毒抗原的核酸序列。

[0011] 还提供了重组载体,其包含选自以下的多核苷酸:

(a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸;

其中所述重组载体包含编码狂犬病病毒抗原的核酸序列。

[0012] 还提供了重组腺病毒,其包含选自以下的至少一种多核苷酸或多肽:

(a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,

(c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸,

(d) 具有根据SEQ ID NO:1的氨基酸序列的多肽;

(e) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(f) 具有根据SEQ ID NO:3的氨基酸序列的多肽;

其中所述重组腺病毒包含编码狂犬病病毒抗原的核酸序列;和其中所述核酸序列与指导所述狂犬病病毒抗原在宿主细胞中表达的一种或多种序列可操作地连接。

[0013] 本发明提供了重组腺病毒,其包含选自以下的一种多核苷酸或多肽:

(a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,

(c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸,

(d) 具有根据SEQ ID NO:1的氨基酸序列的多肽;

(e) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和

(f) 具有根据SEQ ID NO:3的氨基酸序列的多肽;

其中所述腺病毒包含编码狂犬病病毒抗原的核酸序列,其中所述核酸序列与指导所述狂犬病病毒抗原在宿主细胞中表达的一种或多种序列可操作地连接。所述重组腺病毒可包含选自上文所列(a)至(f)的一种或多种其他多核苷酸或多肽。

[0014] 还提供了包含以下中的至少一种的组合物:

- (a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,
- (b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,
- (c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸,
- (d) 具有根据SEQ ID NO:1的氨基酸序列的多肽;
- (e) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,
- (f) 具有根据SEQ ID NO:3的氨基酸序列的多肽;
- (g) 包含如上文(a)、(b)或(c)中所述的多核苷酸的载体,和
- (h) 包含如上文(a)、(b)或(c)中所述的多核苷酸的重组腺病毒和药学上可接受的赋形剂。

[0015] 其中组合物包含编码狂犬病病毒抗原或狂犬病病毒抗原多肽序列的核酸序列;并且,任选地,核酸序列与指导所述狂犬病病毒抗原在宿主细胞中表达的一种或多种序列可操作地连接。

[0016] 还提供了包含以下中的至少一种的细胞:

- (a) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸,
- (b) 编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,
- (c) 编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸,
- (d) 具有根据SEQ ID NO:1的氨基酸序列的多肽;
- (e) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,
- (f) 具有根据SEQ ID NO:3的氨基酸序列的多肽;
- (g) 包含如上文(a)、(b)或(c)中所述的多核苷酸的载体,和
- (h) 包含如上文(a)、(b)或(c)中所述的多核苷酸的重组腺病毒;

其中所述细胞包含腺病毒,所述腺病毒包含编码狂犬病病毒抗原的核酸序列;并且其中核酸序列与指导所述狂犬病病毒抗原在宿主细胞中表达的一种或多种序列可操作地连接。

[0017] 还提供了选自以下的分离的腺病毒多肽:

- (a) 具有根据SEQ ID NO:1的氨基酸序列的多肽;
- (b) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物,其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列,和
- (c) 具有根据SEQ ID NO:3的氨基酸序列的多肽;和

还包含狂犬病病毒抗原多肽序列。

[0018] 还提供了分离的多核苷酸、载体、重组腺病毒、组合物或细胞,其包含根据SEQ ID NO:6的序列并且还包含狂犬病病毒抗原。

[0019] 重组腺病毒和组合物可用作药物,特别是用于刺激针对狂犬病病毒性疾病,例如狂犬病感染的免疫应答。

[0020] 通常,本发明的方法的目的是诱导保护性免疫应答,即针对有关的病原体给主体免疫或接种疫苗。本发明因此可用于预防、治疗或改善由于狂犬病病毒性疾病感染,例如狂犬病病毒感染引起的疾病。

[0021] 在一些方面,本文公开的组合物是免疫原性组合物,其在施用于主体时诱导体液和/或细胞免疫应答,即特异性识别天然存在的狂犬病病毒多肽的免疫应答。例如,免疫原性组合物可以在病毒感染后相对于未治疗的主体诱导记忆T和/或B细胞群,特别是在其中组合物包含含有编码狂犬病病毒抗原的序列的核酸的那些实施方案中。

[0022] 可以提供本发明用于预先暴露预防和暴露后预防由狂犬病病毒性疾病引起的疾病的目的。在一些实施方案中,主体先前已经接种过狂犬病疫苗。例如,本发明的方法可以例如在狂犬病疫苗接种后至少一年,狂犬病疫苗接种后至少两年,狂犬病疫苗接种后至少五年或狂犬病疫苗接种后至少十年用于主体。

[0023] 狂犬病病毒抗原是抗原序列,即来自狂犬病病毒蛋白的序列,其包含至少一个B或T细胞表位。适当地,狂犬病病毒抗原包含至少一个T细胞表位。在本发明的一个实施方案中,腺病毒包含编码狂犬病病毒抗原的核酸序列。在本发明的一个具体实施方案中,腺病毒包含编码衍生自SEQ ID NO:37的多肽的核酸。在本发明的另一个具体实施方案中,腺病毒包含衍生自SEQ ID NO:38的核酸。

[0024] 在本发明的另一个实施方案中,腺病毒可包含编码衍生自SEQ ID NO:39、SEQ ID NO:41、SEQ ID NO:43或SEQ ID NO:45的多肽的核酸。在本发明的其他具体实施方案中,腺病毒可包含衍生自SEQ ID NO:40、SEQ ID NO:42、SEQ ID NO:44或SEQ ID NO:46的核酸。

[0025] 引发的免疫应答可以是抗原特异性B细胞应答,其产生中和抗体。引发的免疫应答可以是抗原特异性T细胞应答,其可以是全身性和/或局部应答。抗原特异性T细胞应答可以包括CD4⁺ T细胞应答,例如涉及表达多种细胞因子,例如IFN γ 、肿瘤坏死因子- α (TNF α) 和/或IL2的CD4⁺ T细胞的应答。可选地或另外,抗原特异性T细胞应答包括CD8⁺ T细胞应答,例如涉及表达多种细胞因子,例如IFN γ 、TNF α 和/或IL2的CD8⁺ T细胞的应答。

附图说明

[0026] 图1:表示主要抗原表位的狂犬病糖蛋白的图。图1以出现顺序分别公开了SEQ ID NO 94、64和78。

图2A-C:来自指示的猿猴腺病毒的纤突蛋白序列的比对。

ChAd3 (SEQ ID NO:27)

PanAd3 (SEQ ID NO:28)

ChAd17 (SEQ ID NO:29)

ChAd19 (SEQ ID NO:30)

ChAd24 (SEQ ID NO:31)

ChAd155 (SEQ ID NO:1)

ChAd11 (SEQ ID NO:32)

ChAd20 (SEQ ID NO:33)

ChAd31 (SEQ ID NO:34)

PanAd1 (SEQ ID NO:35)

PanAd2 (SEQ ID NO:36)

图3:用于产生特异性ChAd155 BAC和质粒载体的流程图

图4:物种C BAC 穿梭物 #1365示意图

图5:pArsChAd155 Ad5E4orf6-2 (#1490)示意图

图6:pChAd155/RSV示意图

图7:BAC ChAd155/RSV示意图

图8:E4 Ad5E4orf6/ Tet0 hCMV RG WPRE (#1509)示意图

图9:表达HIV Gag转基因的ChAd3和ChAd155载体的生产力

图10:表达RSV转基因的ChAd155和PanAd3载体的表达水平 - 蛋白质印迹

图11:表达HIV Gag转基因的ChAd3和ChAd155载体的免疫原性 - IFN- γ ELISpot

图12:表达RSV转基因的PanAd3和ChAd155载体的免疫原性 - IFN- γ ELISpot

图13:在小鼠中表达狂犬病G蛋白转基因的ChAd155载体的免疫原性 - (A)抗体中和和

(B) IFN- γ ELISpot

图14:表达狂犬病G蛋白转基因的ChAd155载体诱导的免疫应答的稳定性

图15:与市售疫苗相比,表达狂犬病G蛋白转基因的ChAd155载体的效力

图16:在小鼠中表达狂犬病G蛋白转基因的ChAd155载体的血清转变和保护率

图17:在小鼠中表达狂犬病G蛋白转基因的ChAd155载体的血清转变和保护率

图18:在非人灵长类动物中表达狂犬病G蛋白转基因的ChAd155载体的中和抗体应答

图19:在非人灵长类动物中对表达狂犬病G蛋白转基因的ChAd155载体的T细胞应答。

[0027] 序列描述

SEQ ID NO:1-ChAd155纤突的多肽序列

SEQ ID NO:2-编码ChAd155纤突的多核苷酸序列

SEQ ID NO:3-ChAd155五邻体的多肽序列

SEQ ID NO:4-编码ChAd155五邻体的多核苷酸序列

SEQ ID NO:5- ChAd155六邻体的多肽序列

SEQ ID NO:6-编码ChAd155六邻体的多核苷酸序列

SEQ ID NO:7-编码ChAd155 # 1434的多核苷酸序列

SEQ ID NO:8-编码ChAd155 # 1390的多核苷酸序列

SEQ ID NO:9-编码ChAd155 # 1375的多核苷酸序列

SEQ ID NO:10-编码野生型ChAd155的多核苷酸序列
SEQ ID NO:11-编码ChAd155/RSV的多核苷酸序列
SEQ ID NO:12-编码CASI启动子的多核苷酸序列
SEQ ID NO:13- Ad5orf6引物1多核苷酸序列
SEQ ID NO:14- Ad5orf6引物2多核苷酸序列
SEQ ID NO:15- BAC/CHAd155 ΔE1_Tet0 hCMV RpsL-Kana引物1多核苷酸序列
SEQ ID NO:16- BAC/CHAd155 ΔE1_Tet0 hCMV RpsL-Kana (#1375) 引物2多核苷酸序列

列

SEQ ID NO:17-1021-FW E4 Del Step1引物多核苷酸序列
SEQ ID NO:18-1022-RW E4 Del Step1引物多核苷酸序列
SEQ ID NO:19-1025-FW E4 Del Step2引物多核苷酸序列
SEQ ID NO:20-1026-RW E4 Del Step2引物多核苷酸序列
SEQ ID NO:21-91-SubMonte FW引物多核苷酸序列
SEQ ID NO:22-90-BghPolyA RW引物多核苷酸序列
SEQ ID NO:23- CMVfor引物多核苷酸序列
SEQ ID NO:24- CMVrev引物多核苷酸序列
SEQ ID NO:25-CMVFAM-TAMRA qPCR探针多核苷酸序列
SEQ ID NO:26-土拨鼠肝炎病毒转录后调节元件(WPRE)多核苷酸序列
SEQ ID NO:27- ChAd3的纤突蛋白的氨基酸序列
SEQ ID NO:28- PanAd3的纤突蛋白的氨基酸序列
SEQ ID NO:29- ChAd17的纤突蛋白的氨基酸序列
SEQ ID NO:30- ChAd19的纤突蛋白的氨基酸序列
SEQ ID NO:31- ChAd24的纤突蛋白的氨基酸序列
SEQ ID NO:32- ChAd11的纤突蛋白的氨基酸序列
SEQ ID NO:33- ChAd20的纤突蛋白的氨基酸序列
SEQ ID NO:34- ChAd31的纤突蛋白的氨基酸序列
SEQ ID NO:35- PanAd1的纤突蛋白的氨基酸序列
SEQ ID NO:36- PanAd2的纤突蛋白的氨基酸序列
SEQ ID NO:37- RG medoid抗原的氨基酸序列
SEQ ID NO:38- RG medoid抗原的核苷酸序列
SEQ ID NO:39- RG AA098的氨基酸序列
SEQ ID NO:40- RG AA098的核酸序列
SEQ ID NO:41- RG AA0093的氨基酸序列
SEQ ID NO:42- RG AA0093的核酸序列
SEQ ID NO:43- RG AA0094的氨基酸序列
SEQ ID NO:44- RG AA0094的核酸序列
SEQ ID NO:45- RG AA0095的氨基酸序列
SEQ ID NO:46- RG AA0095的核酸序列
SEQ ID NO:47- AdC68rab.gp-ERA毒株的氨基酸序列

SEQ ID NO:48- AdC68rab.gp-ERA毒株的核苷酸序列
SEQ ID NO:49-聚组氨酸
SEQ ID NO:50-63-狂犬病G蛋白位点IIb抗原表位
SEQ ID NO:64-77-狂犬病G蛋白位点I抗原表位
SEQ ID NO:78-91-狂犬病G蛋白位点III抗原表位
SEQ ID NO:92- pvjTet0hCMV_WPRE_BghPolyA正向引物
SEQ ID NO:93- pvjTet0hCMV_WPRE_BghPolyA反向引物。

[0028] 发明详述

腺病毒载体

腺病毒由于其经证明的安全性、在多种靶组织中实现高效基因转移的能力和大的转基因容量已经被广泛地用于基因转移应用。在本发明中使用的腺病毒载体可以源自许多哺乳动物宿主。已经分离腺病毒的超过100种不同血清型,其感染多种哺乳动物物种。根据序列同源性和它们凝集红细胞的能力,已经将这些腺病毒血清型分为6个亚类(A-F;B细分为B1和B2)。

[0029] 在一个实施方案中,本发明的腺病毒载体源自非人猿猴腺病毒,也简单地称作猿猴腺病毒。已经从非人猿猴例如黑猩猩、倭黑猩猩、恒河猴和大猩猩中分离出许多腺病毒,并且衍生自这些腺病毒的载体诱导对由这些载体编码的转基因的强免疫应答(Colloca 等人(2012) Sci. Transl. Med. 4:1-9;Roy 等人(2004) Virol.324: 361-372;Roy 等人(2010) J. of Gene Med. 13:17-25)。基于非人猿猴腺病毒的载体的某些优点包括在靶群体中相对缺乏针对这些腺病毒的交叉中和抗体,因此它们的使用克服了对人腺病毒的预先存在的免疫力。例如,与在某些候选人腺病毒载体的情况下的35%相比,某些黑猩猩腺病毒与预先存在的中和抗体应答的交叉反应仅存在于2%的靶群体中。

[0030] 具体地,腺病毒载体可以源自非人腺病毒,例如猿猴腺病毒,且特别是黑猩猩腺病毒,例如ChAd3、ChAd63、ChAd83、ChAd155、Pan 5、Pan 6、Pan 7(也称为C7)或Pan 9并且可以全部或部分包括编码非人腺病毒的纤突、五邻体或六邻体的核苷酸。这样的毒株的实例描述在W003/000283、W02010/086189和GB1510357.5中,并且也可从美国典型培养物保藏中心, 10801 University Boulevard, Manassas, Virginia 20110-2209和其他来源获得。或者,腺病毒载体可以源自从矮黑猩猩分离的非人猿猴腺病毒,例如PanAd1、PanAd2或PanAd3。本文描述的这样的载体的实例可以见于例如W02005/071093和W02010/086189。腺病毒载体也可以源自从大猩猩分离的腺病毒,例如在W02013/52799、W02013/52811和W02013/52832中所述。

[0031] 腺病毒载体结构

腺病毒具有含二十面体衣壳的特征形态,所述二十面体衣壳包含三种主要蛋白:六邻体(II)、五邻体基质(III)和多节纤突(IV)、以及许多其他次要蛋白VI、VIII、IX、IIa和IVa2。所述六邻体占衣壳的结构组分的大部分,所述衣壳由240个三聚的六邻体壳粒和12个五邻体基质组成。所述六邻体具有三个保守双桶,而顶部具有三个塔,每个塔含有来自每个亚基的环,所述亚基形成衣壳的大部分。六邻体的基质在腺病毒血清型之间是高度保守的,而表面环是可变的。五邻体是另一种腺病毒衣壳蛋白,其形成纤突所连接的五聚体基质。三聚的纤突蛋白从在衣壳的12个顶点中的每一个处的五邻体基质伸出且是有节的杆状结构。

纤突蛋白的主要作用是通过结区域(knob region)与细胞受体的相互作用将病毒衣壳拴系至细胞表面,并且纤突的柔性轴以及结区域中的变化是不同血清型的特征。

[0032] 腺病毒基因组已被充分表征。线性双链DNA与高碱性蛋白VII和小肽pX(也称为mu)相关。另一种蛋白V与该DNA-蛋白复合物一起包装,并通过蛋白VI提供与衣壳的结构连接。就类似地定位(例如每种病毒的E1A、E1B、E2A、E2B、E3、E4、L1、L2、L3、L4和L5基因的位置)的特定开放读码框而言,在腺病毒基因组的整体组中存在普遍保守性。腺病毒基因组的每个末端包含被称作反向末端重复(ITR)的序列,其是病毒复制所必需的。腺病毒基因组的5'末端含有包装和复制所必需的5'顺式元件;即,5'ITR序列(其作为复制起点起功能)和天然5'包装增强子结构域(其包含用于包装线性腺病毒基因组所必需的序列和E1启动子的增强子元件)。腺病毒基因组的3'末端包括用于包装和衣壳化所必需的3'顺式元件(包括ITR)。该病毒也包含病毒编码的蛋白酶,其是用于加工产生传染性病毒粒子所需的一些结构蛋白所必需的。

[0033] 基于在宿主细胞转导以后表达病毒基因的次序描述了腺病毒基因组的结构。更具体地,根据在DNA复制开始之前或之后是否发生转录,将所述病毒基因称作早期(E)或晚期(L)基因。在转导的早期阶段,表达腺病毒的E1A、E1B、E2A、E2B、E3和E4基因以制备用于病毒复制的宿主细胞。在感染的晚期阶段,编码病毒颗粒的结构组分的晚期基因L1-L5的表达被活化。

[0034] 腺病毒衣壳蛋白及其编码多核苷酸

如上所述,腺病毒衣壳包含三种主要蛋白质,六邻体、五邻体和纤突。所述六邻体占衣壳的结构组分的大部分,所述衣壳由240个三聚的六邻体壳粒和12个五邻体基质组成。所述六邻体具有三个保守双桶,而顶部具有三个塔,每个塔含有来自每个亚基的环,所述亚基形成衣壳的大部分。六邻体的基质在腺病毒血清型之间是高度保守的,而表面环是可变的。

[0035] 五邻体是另一种腺病毒衣壳蛋白,其形成纤突所连接的五聚体基质。三聚的纤突蛋白从在衣壳的12个顶点中的每一个处的五邻体基质伸出且是有节的杆状结构。腺病毒衣壳的表面相对于大多数其他二十面体病毒的表面的显著差异是细长的纤突蛋白的存在。所述纤突蛋白的主要作用是通过它与细胞受体的相互作用而将病毒衣壳与细胞表面拴系。

[0036] 许多腺病毒血清型的纤突蛋白共享共同的构造:N-端尾巴,由重复序列构成的中心轴,以及C-端球形结结构域(或“头部”)。所述中心轴结构域由可变数目的 β -重复组成。所述 β -重复连接以形成高度刚性且稳定的三个相互缠绕的螺旋链的长形结构。所述轴将N-端尾巴与球形结结构连接,所述球形结结构负责与靶细胞受体的相互作用。腺病毒结结构域的球形性质呈现大表面用于在侧面和在顶点结合所述受体。该构造的作用是使受体结合位点突出远离病毒衣壳,从而使该病毒脱离由相对扁平的衣壳表面呈现的空间约束。

[0037] 尽管许多腺病毒血清型的纤突具有相同的整体构造,但它们具有影响它们的功能以及结构的可变氨基酸序列。例如,在纤突结的表面上的许多暴露区域呈现可容易地适应的受体结合位点。纤突结的球形形状允许受体结合在所述结的侧面处或所述纤突结的顶部上。这些结合位点通常位于表面暴露的环上,所述环连接在人腺病毒之间不太保守的 β -链。这些环上的暴露侧链给予所述结多种表面特征,同时保留三级和四级结构。例如,在结表面处的静电势和电荷分布可以由于纤突结序列中的等电点的宽范围(在对于腺病毒“Ad” 8、Ad 19和Ad 37而言大约9至对于亚组B腺病毒而言大约5的pI范围内变化)而变化。作为在结

构上复杂的病毒配体,所述纤突蛋白允许来自病毒衣壳的呈许多取向和距离(轴)的多种结合表面(结)的呈现。

[0038] 一些血清型之间最明显的变化之一是纤突长度。研究已经证实,纤突轴的长度强烈地影响所述结和所述病毒与其靶受体的相互作用。此外,血清型之间的纤突蛋白还可以在它们的弯曲能力方面存在差异。尽管所述轴中的 β -重复形成高度稳定的且规则的结构,但是电子显微术(EM)研究已经显示所述纤突中的不同铰链。来自几种腺病毒血清型纤突的蛋白序列的分析准确地定位在来自N-端尾巴的第三个 β -重复处所述轴的重复序列的破坏,其与所述轴中的铰链之一强相关,如通过EM看到的。所述纤突中的铰链允许所述结采取相对于该病毒衣壳的多种取向,这可以防止对受体接合的位阻,所述受体接合需要所述结上的受体结合位点的正确呈现。例如,亚组D Ad5的刚性纤突因此需要柔性受体或预先定位用于病毒附着的受体,因为它们不能自身弯曲。

[0039] 通过使用纤突假型化技术已经实现了对不同Ad血清型特异性的细胞受体的鉴别和它们如何促进组织嗜性的知识。尽管一些亚组的Ad使用柯萨奇病毒(Coxsackievirus)和腺病毒受体("CAR")作为第一受体,但越来越清楚的是,许多Ad使用替代性第一受体,从而导致非常不同的体外和体内嗜性。这些血清型的纤突表现出它们的一级结构和三级结构的明显差异,如纤突轴刚度、纤突轴长度,和CAR结合位点和/或推定的HSPG结合基序的缺乏,以及在纤突结内的净电荷的差异。因此,用替代性纤突轴和结将Ad 5颗粒假型化提供除去重要细胞结合结构域的机会,且另外,可以允许与用Ad 5实现的结果相比,更有效的(和潜在地更具细胞选择性的)向确定细胞类型进行转基因递送。如果使用的纤突是来自在人类或实验模型中具有较低血清阳性率的Ad(有利于载体的成功施用的情形),也可以减少纤突假型化的Ad颗粒的中和。此外,全长纤突以及分离的纤突结区域(但不是单独的六邻体或五邻体)能够诱导树突细胞成熟并且与有效的CD8⁺ T细胞应答的诱导相关。总之,腺病毒纤突在腺病毒载体的至少受体结合和免疫原性中起重要作用。

[0040] "低血清阳性率"可以是指具有与人腺病毒5(Ad5)相比降低的预先存在的中和抗体水平。类似地或可选地,"低血清阳性率"可以是指小于约20%血清阳性率、小于约15%血清阳性率、小于约10%血清阳性率、小于约5%血清阳性率、小于约4%血清阳性率、小于约3%血清阳性率、小于约2%血清阳性率、小于约1%血清阳性率或无可检测的血清阳性率。使用在Aste-Amézaga等人, *Hum. Gene Ther.* (2004) 15(3):293-304中描述的方法,可以将血清阳性率测量为具有临床上有关的中和滴度(定义为>200的50%中和滴度)的个体的百分比。

[0041] 图1中提供的比对阐明了C组猿猴腺病毒的纤突蛋白之间的差异。一个突出特征是,这些腺病毒的纤突序列可以宽泛地分组成具有长纤突(如ChAd155)或短纤突(如ChAd3)。该长度差别是由于与长纤突相比在短纤突中大约位置321处的36个氨基酸缺失。另外,存在许多氨基酸取代,它们在短纤突亚组相对于长纤突亚组之间不同,但是在每个亚组内是一致的。尽管尚未阐明这些差异的确切功能,鉴于纤突的功能和免疫原性,它们可能是显著的。已经显示,病毒嗜性的决定簇之一是纤突轴的长度。已经证明,具有较短轴的Ad5载体具有与CAR受体较低的结合效率和较低的感染性。已经推测,这种损害是较短纤突刚性增加的结果,其导致对细胞受体的附着效率较低。这些研究可以解释与以前描述的ChAd3和PanAd3(其携带具有更短轴的纤突)相比携带更长和更柔性纤突的ChAd155的改善的性能。

[0042] 在本发明的一个方面提供了黑猩猩腺病毒ChAd155的分离的纤突、五邻体和六邻

体衣壳多肽,和编码黑猩猩腺病毒ChAd155的纤突、五邻体和六邻体衣壳多肽的分离的多核苷酸。“分离的”多核苷酸是从其原始环境中取出的多核苷酸。例如,如果天然存在的多核苷酸分离自天然系统中共存物质的一些或所有中,则其被分离。如果例如将多核苷酸克隆入不是天然环境的一部分的载体内或如果其包含在cDNA内,则认为将其分离。

[0043] 预期所有三种衣壳蛋白都促成低血清阳性率,且因而可以彼此独立地或组合地用于抑制腺病毒对预先存在的中和抗体的亲和力,例如以制备具有降低的血清阳性率的重组腺病毒。这样的重组腺病毒可以是具有衣壳蛋白的嵌合腺病毒,所述衣壳蛋白来自至少具有来自ChAd155的纤突蛋白的不同血清型。

[0044] 转基因

腺病毒载体可以用于递送期望的RNA或蛋白序列(例如异源序列)用于体内表达。载体可以包括任何遗传元件,包括裸露DNA、噬菌体、转座子、粘粒、附加体、质粒或病毒。这样的载体含有本文中公开的ChAd155的DNA和表达盒。“表达盒”(或“微基因”)是指选择的异源基因(“转基因”)和驱动宿主细胞中的基因产物的翻译、转录和/或表达所必需的其他调节元件的组合。

[0045] 通常,“异源”是指源自与其进行比较的实体的其余部分的基因型不同的实体。异源核酸序列指并非分离自、源自或基于腺病毒载体的天然存在的核酸序列的任何核酸序列。“天然存在的”是指在自然界中发现的且不是合成制备或修饰的序列。当序列分离自一个来源、但是经过适当地修饰(例如,通过缺失、取代(突变)、插入或其他修饰)从而不破坏来源基因的正常功能时,所述序列“源自”所述来源。

[0046] 通常,设计腺病毒载体,使得所述表达盒位于这样的核酸分子中:所述核酸分子在选择腺病毒基因的天然区域中含有其他腺病毒序列。如果需要的话,可以将所述表达盒插入现存的基因区域中以破坏该区域的功能。或者,可以将所述表达盒插入部分或完全缺失的腺病毒基因的位点。例如,所述表达盒可以位于使选自E1A、E1B、E2A、E2B、E3和E4的基因组区域的至少一个基因无功能的突变、插入或缺失的位点。术语“使.....无功能”是指,将足够量的基因区域除去或以其他方式破坏,使得所述基因区域不再能够产生基因表达的功能产物。如果需要的话,可以将整个基因区域除去(和适当地用所述表达盒替换)。适当地,将腺病毒的E1基因缺失并用由所选启动子、目标基因的cDNA序列和聚腺苷酸信号组成的表达盒替换,从而产生复制缺陷型重组病毒。

[0047] 由腺病毒载体编码的转基因是编码可用于生物学和医学的产物(如治疗性或免疫原性蛋白、RNA或酶中的一种或多种)的序列。合乎需要的RNA分子包括tRNA、dsRNA、核糖体RNA、催化RNA、RNA适体和反义RNA。有用的RNA序列的实例是消除治疗的动物中的靶向核酸序列的表达的序列。

[0048] 所述转基因是编码目标蛋白的核酸序列,其与侧接所述转基因的载体序列异源。核酸编码序列以允许宿主细胞中的转基因转录、翻译和/或表达的方式可操作地连接到调节组分。

[0049] 转基因可编码用于疾病治疗、改善或预防的多肽或蛋白,用于诱导免疫应答,和/或用于预防性疫苗目的。如本文中所用,免疫应答的诱导指蛋白(也被称作“抗原”或“免疫原”)诱导针对所述蛋白的T细胞和/或体液免疫应答的能力。

[0050] 由本发明载体表达的免疫原可用于针对其他病原体免疫人或非人动物,或者来自

癌细胞或肿瘤细胞,所述其他病原体包括,例如,感染人和非人脊椎动物的细菌、真菌、寄生微生物或多细胞寄生物。例如,免疫原可以选自多种病毒科。在一个实施方案中,免疫原来自狂犬病病毒,例如莫科拉病毒(Mokola virus)、杜文海格病毒(Duvenhage virus)、欧洲蝙蝠狂犬病病毒、欧洲蝙蝠狂犬病病毒 2和澳大利亚蝙蝠狂犬病病毒。在一个实施方案中,狂犬病病毒免疫原来自狂犬病病毒,例如来自CVS11、CVS-N2C、Evelyn Rokitniki Abelseth (ERA)、Flury、Pitman Moore或Wistar毒株。这些抗原可以源自狂犬病病毒糖蛋白(G)、RNA聚合酶(L)、基质蛋白(M)、核蛋白(N)和磷蛋白(P),例如包含狂犬病病毒糖蛋白(G)、RNA聚合酶(L)、基质蛋白(M)、核蛋白(N)和磷蛋白(P)或包含其片段(合适地是至少20、至少50、至少100、至少200、至少300、至少400、至少500或至少600个氨基酸的片段)。

[0051] 在一个实施方案中,由本发明的载体表达的免疫原包含来自莫科拉病毒、杜文海格病毒、欧洲蝙蝠狂犬病病毒、欧洲蝙蝠狂犬病病毒 2和澳大利亚蝙蝠狂犬病病毒的糖蛋白的全部或片段,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。在一个实施方案中,由本发明的载体表达的免疫原包含来自狂犬病病毒、例如来自CVS11、CVS-N2C、Evelyn Rokitniki Abelseth (ERA)、Flury、Pitman Moore或Wistar毒株的糖蛋白的全部或片段,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。在一个实施方案中,由本发明的载体表达的免疫原包含SEQ ID NO:37的全部或片段,所述片段合适地为至少20、至少50、至少100、至少200、至少300、至少400或至少500个氨基酸的片段。

[0052] 在一个实施方案中,由本发明的载体表达的免疫原包含表1中所示的一种或多种抗原表位。在一个实施方案中,由本发明的载体表达的免疫原包含对应于狂犬病病毒毒株RABV、ABLV、ARAV、BBLV、DUVV、EBLV-1、EBLV-2、IRKV、KHUV、LBV、MOKV、SHIV、WCBV或IKOV的位点I、位点IIa、位点IIb、位点III、位点IV和/或位点a的表位。在一个具体实施方案中,本发明的载体包含对应于SEQ ID NO:37中发现的位点I、位点IIa、位点IIb、位点III、位点IV和/或位点a的表位。

[0053] 在一个实施方案中,可以通过包含抗原的medoid序列来增加疫苗构建体的交叉保护性宽度。“medoid”是指狂犬病病毒序列,其与其他狂犬病病毒序列具有最小的差异性(dissimilarity)。在一个具体实施方案中,本发明的载体包含G糖蛋白的medoid序列。在一个具体实施方案中,本发明的非人灵长类动物载体包含G糖蛋白的medoid序列。在一个具体实施方案中,本发明的ChAd155载体包含G糖蛋白的medoid序列。在一个具体实施方案中,medoid序列源自天然病毒毒株,其在NCBI数据库中注释的所有G蛋白序列中具有最高的氨基酸同一性平均百分比。在一个具体实施方案中,G糖蛋白的medoid序列是NCBI毒株AGN94271。

[0054] 可选地或另外,转基因序列可以包括报告序列(reporter sequence),其在表达后产生可检测信号。这样的报告序列包括但不限于,编码 β -内酰胺酶、 β -半乳糖苷酶(LacZ)、碱性磷酸酶、胸苷激酶、绿色荧光蛋白(GFP)、氯霉素乙酰基转移酶(CAT)、萤光素酶、膜结合蛋白(包括例如,CD2、CD4、CD8、流感血凝素蛋白和本领域众所周知的其他蛋白(存在针对它们的高亲和力抗体或可以通过常规方式生产所述抗体))和包含适当地与抗原标签结构域(尤其来自血凝素或Myc)融合的膜结合蛋白的融合蛋白的DNA序列。这些编码序列当与驱动它们的表达的调节元件结合时会提供可通过常规方式检测的信号,所述常规方式包括酶测

定、放射摄影测定、比色测定、荧光测定或其他光谱测定、荧光活化细胞分选测定和免疫学测定(包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和免疫组织化学)。

[0055] 除了转基因以外,所述表达盒还可以包括常规控制元件,所述控制元件以允许所述转基因在被所述腺病毒载体转染的细胞中转录、翻译和/或表达的方式与所述转基因可操作地连接。如本文所用,“可操作地连接的”序列包括与目标基因邻接的表达控制序列和以反式或在一定距离处起作用以控制目标基因的表达控制序列。

[0056] 表达控制序列包括合适的转录起始、终止、启动子和增强子序列;有效的RNA加工信号,例如剪接和聚腺苷酸化(poly A)信号,包括兔 β -珠蛋白polyA;稳定细胞质mRNA的序列;增强翻译效率的序列(例如,Kozak共有序列);增强蛋白稳定性的序列;并且当需要时,增强编码产物分泌的序列。在其他序列中,可以使用嵌合的内含子。

[0057] “启动子”是允许RNA聚合酶的结合并指导基因的转录的核苷酸序列。通常,启动子位于基因的5'非编码区,在所述基因的转录起始位点的近侧。在转录的起始中起作用的启动子内的序列元件经常特征在于共有核苷酸序列。启动子的实例包括但不限于来自细菌、酵母、植物、病毒和哺乳动物(包括人类)的启动子。大量表达控制序列(包括内部的、天然的、组成型、诱导型和/或组织特异性的启动子)是本领域已知的且可以利用。

[0058] 组成型启动子的实例包括但不限于TBG启动子、逆转录病毒劳氏肉瘤病毒(Rous sarcoma virus)LTR启动子(任选地具有增强子)、巨细胞病毒(CMV)启动子(任选地具有CMV增强子,参见例如Boshart 等人, Cell, 41:521-530 (1985))、CASI启动子(WO2012/115980)、SV40启动子、二氢叶酸还原酶启动子、 β -肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EF1a启动子(Invitrogen)。

[0059] 诱导型启动子允许调节基因表达,并且可以通过外源提供的化合物、环境因素如温度或特定生理学状态(例如急性期)的存在、细胞的特定分化状态或仅在复制中的细胞中被调节。诱导型启动子和诱导型系统可从多种商业来源获得,包括但不限于Invitrogen、Clontech和Ariad。已经描述了许多其他系统,并且本领域技术人员可以容易地选择这些系统。例如,诱导型启动子包括锌诱导型绵羊金属硫蛋白(MT)启动子和地塞米松(Dex)诱导型小鼠乳腺肿瘤病毒(MMTV)启动子。其他诱导型系统包括T7聚合酶启动子系统;蜕皮激素昆虫启动子,四环素可抑制系统和四环素诱导系统。其他系统包括FK506二聚体、VP16或p65(使用甘珀二醇(castradiol)、联苯酚米勒甾酮(diphenol murislerone))、RU486诱导系统和雷帕霉素诱导系统。一些诱导型启动子的有效性随着时间的推移而增加。在这样的情况下,可以通过插入多个串联阻遏物(例如,通过IRES与TetR连接的TetR)增强这样的系统的有效性。

[0060] 在另一个实施方案中,可以使用转基因的天然启动子。当期望转基因的表达应当模拟天然表达时,可以优选所述天然启动子。当转基因的表达必须在时间上或在发育上受到调节或以组织特异性的方式受到调节或响应于特定转录刺激而调节时,可以使用所述天然启动子。在另一个实施方案中,其他天然表达控制元件(例如增强子元件、聚腺苷酸化位点或Kozak共有序列)也可以用于模拟天然表达。

[0061] 转基因可以与组织特异性启动子可操作地连接。例如,如果需要在骨骼肌中表达,则应使用在肌肉中有活性的启动子。这些包括来自编码骨骼 β -肌动蛋白、肌球蛋白轻链2A、肌营养不良蛋白、肌肉肌酸激酶的基因的启动子,以及具有高于天然存在的启动子的活性

的合成肌肉启动子。组织特异性的启动子的实例已知针对肝的；乙型肝炎病毒核心；甲胎蛋白，骨钙素；骨唾液蛋白，淋巴细胞，免疫球蛋白重链；T细胞受体链），针对神经元的例如神经元特异性的烯醇化酶（NSE）启动子，神经丝轻链基因和神经元特异性的vgf基因等。

[0062] 在一些实施方案中，土拨鼠肝炎病毒转录后调节元件（WPRE）（Zuffrey 等人（1999）J. Virol.;73(4):2886-9）可以与转基因可操作地连接。

[0063] 转基因可以用于治疗，例如作为疫苗，用于诱导免疫应答，和/或用于预防性疫苗目的。如本文所用，免疫应答的诱导是指蛋白诱导针对所述蛋白的T细胞和/或体液免疫应答的能力。

[0064] 腺病毒载体构建

通过修饰野生型腺病毒以表达异源基因和/或缺失或失活不希望的腺病毒序列来制备腺病毒载体。腺病毒载体也可能具有改变的复制能力。例如所述载体可以是复制缺陷型或具有有限的复制，使得它在非互补细胞（non-complementing cells）中具有与野生型病毒相比减少的复制能力。这可以如下实现：突变所述病毒，例如通过缺失参与复制的基因，例如缺失E1A、E1B、E3或E4基因。

[0065] 根据本发明的腺病毒载体可包含功能性E1缺失。因此，由于缺乏表达腺病毒E1A和/或E1B的能力，根据本发明的腺病毒载体可能是复制缺陷型的。所述重组腺病毒还可能在其他基因中携带功能性缺失，例如，E3或E4基因中的缺失。所述腺病毒延迟早期基因E3可从腺病毒序列（其构成所述重组病毒的一部分）中被消除。E3的功能对于生产所述重组腺病毒颗粒不是必要的。因此，没有必要替换此基因产物的功能从而包装可用于本发明的重组腺病毒。在一个具体的实施方案中，所述重组腺病毒具有功能上缺失的E1和E3基因。此类载体的构建在Roy 等人，Human Gene Therapy 15:519-530，2004中有述。

[0066] 还可以构建具有E4基因功能缺失的重组腺病毒。在一个具体实施方案中，重组腺病毒具有功能缺失的E1和E4基因，如Colloca 等人（2012）Sci. Transl. Med. 4:1-9；Roy 等人（2004）Virol.324: 361-372中所述。在一些实施方案中，可能需要保留E4 ORF6功能。在一个实施方案中，E4 ORF6区域可以被例如来自人腺病毒5（Ad5）的异源E4 ORF6替换。因此，在一个具体的实施方案中，腺病毒载体可以在E1中功能缺失并且具有来自Ad5的E4 ORF6区域。根据本发明的腺病毒载体还可含有延迟早期基因E2a中的功能缺失。缺失还可发生在腺病毒基因组的晚期基因L1至L5中的任一者上。类似地，中间基因（intermediate genes）IX和IVa中的缺失可能是有用的。

[0067] 可以在其他结构或非结构腺病毒基因中产生其他缺失。以上缺失可单独使用，例如，用于本发明中的腺病毒序列可仅含有E1缺失。或者，可以以任意组合使用有效地破坏它们的生物活性的整个基因或其部分的缺失。例如，在一个示例性载体中，腺病毒序列可具有：E1基因和E4基因缺失，或E1、E2a和E3基因缺失，或E1和E3基因缺失（例如E1a和E1b中的功能性缺失以及至少部分E3的缺失），或E1、E2a和E4基因缺失，具有或没有E3缺失，等等。此类缺失可能是这些基因的部分或完全缺失，并且可能与其他突变（例如温度敏感突变）组合使用以实现所需结果。

[0068] 使用本领域技术人员已知的技术制备这些载体。这样的技术包括常规cDNA克隆技术（如在教科书中描述的那些）、腺病毒基因组的重叠寡核苷酸序列的使用、聚合酶链式反应，和提供期望的核苷酸序列的任何合适方法。特别合适的方法包括标准同源重组方法，例

如Colloca 等人 (2012) Sci. Transl. Med. 4:1-9;Roy 等人 (2004) Virol.324: 361-372;Roy 等人 (2010) J. of Gene Med. 13:17-25;和W02010/085984中提供的那些或 Warming 等人 Nuc. Acids Res. (2005) 33:e36中描述的重组方法。

[0069] 腺病毒载体生产

可以在任何合适的细胞系(所述病毒在其中能够复制)中生产腺病毒载体。具体地,可以使用互补细胞系,其提供了从所述病毒载体缺失的因素,这些缺失的因素导致其受损的复制特性(如E1)。不受限制地,此类细胞系可以是HeLa (ATCC登录号CCL 2)、A549 (ATCC登录号CCL 185)、HEK 293、KB (CCL 17)、Detroit (例如,Detroit 510, CCL 72) 和WI-38 (CCL 75) 细胞等等。这些细胞系均可得自美国典型培养物保藏中心,10801 University Boulevard, Manassas, Virginia 20110-2209, USA。其他合适的亲代细胞系可得自其他来源,例如PGK-E1成视网膜细胞,例如,PER.C6™ 细胞,例如以ECACC号96022940保藏在应用微生物学与研究中心(CAMR, UK)的欧洲动物细胞培养物保藏中心(ECACC)的细胞所示,或Her 96 细胞(Crucell)。

[0070] 在许多情况下,表达病毒的复制和感染性所必需的一种或多种缺失基因(如人E1)的细胞系可以用于反式互补(transcomplement)黑猩猩腺病毒载体。这是特别有利的,因为,由于本发明的黑猩猩腺病毒序列和在目前可得到的包装细胞中发现的人腺病毒序列之间的多样性,当前的含有人E1的细胞的使用会阻止在复制和生产过程中产生复制型的腺病毒。

[0071] 或者,如果需要的话,可以利用本文中提供的序列来产生包装细胞或细胞系,其在选定的亲代细胞系中用于表达的启动子的转录控制下最少表达来自ChAd155的E1基因。诱导型或组成型启动子可以用于此目的。这样的启动子的实例详细描述在该文件的别处。选择亲代细胞用于产生表达任何期望的ChAd155基因的新细胞系。不受限制地,这样的亲代细胞系可以是HeLa [ATCC登录号CCL 2]、A549 [ATCC登录号CCL 185]、HEK 293、KB [CCL 17]、Detroit [例如,Detroit 510、CCL 72]和WI-38 [CCL 75]细胞,等等。这些细胞系都可得自美国典型培养物保藏中心, 10801 University Boulevard, Manassas, Virginia 20110-2209, USA。

[0072] 这样的表达E1的细胞系可用于产生重组腺病毒E1缺失的载体。另外或可选地,可以使用与在重组病毒载体的产生中所用的程序基本上相同的程序构建表达一种或多种腺病毒基因产物(例如,E1A、E1B、E2A、E3和/或E4)的细胞系。这样的细胞系可以用于反式互补(transcomplement)编码那些产物的必需基因有缺失的腺病毒载体,或提供辅助依赖性病毒(例如,腺伴随病毒)的包装所必需的辅助功能。宿主细胞的制备涉及技术例如选择的DNA序列的组装。

[0073] 在实施方案中,必需的腺病毒基因产物由腺病毒载体和/或辅助病毒以反式提供。在这样的实例中,合适的宿主细胞可以选自任何生物,包括原核(例如,细菌)细胞和真核细胞,包括昆虫细胞、酵母细胞和哺乳动物细胞。

[0074] 宿主细胞可以选自任何哺乳动物物种,包括但不限于,细胞如A549、WEHI、3T3、10T1/2、HEK 293细胞或Per.C6 (它们中的后二者表达功能性腺病毒E1)、Saos、C2C12、L细胞、HT1080、HepG2和源自哺乳动物(包括人、猴、小鼠、大鼠、兔和仓鼠)的原代成纤维细胞、肝细胞和成肌细胞。

[0075] 特别合适的互补细胞系是Procell192细胞系。Procell 92细胞系是基于表达腺病毒E1基因(在人磷酸甘油酸激酶-1 (PGK) 启动子的控制下用Tet阻遏物转染)和G418-抗性基因的HEK 293细胞(Vitelli 等人 PLOS One (2013) 8(e55435):1-9)。Procell192.S适合于在悬浮条件中生长,且可用于产生表达毒性蛋白的腺病毒载体(www.okairos.com/e/inner.php?m=00084, 最后一次登录在2015年4月13日)。

[0076] 腺病毒递送方法和剂量

腺病毒载体可以以免疫原性组合物施用。如本文中所述的免疫原性组合物是包含一种或多种重组载体的组合物,所述重组载体能够在递送给哺乳动物(适当地为人)以后诱导针对由所述载体递送的转基因产物的免疫应答,例如体液(例如,抗体)和/或细胞介导的(例如,细胞毒性的T细胞)应答。重组腺病毒可以包含(适当地在它的基因缺失中的任一个中)编码期望的免疫原的基因,且因此可以用在疫苗中。重组腺病毒可用作针对任何病原体的预防性或治疗性疫苗,对于所述病原体,对诱导免疫应答至关重要的抗原能够限制病原体的扩散并且对于所述病原体可获得cDNA。

[0077] 可以在合适的递送媒介物中配制这样的疫苗或其他免疫原性组合物。可以监测选择的基因的免疫水平以确定对强化的需要(如果有的话)。在评估血清中的抗体滴度以后,可以期望任选的强化免疫。

[0078] 任选地,可以将本发明的疫苗或免疫原性组合物配制成含有其他组分,包括例如,佐剂、稳定剂、pH调节剂、防腐剂等。在下面在“佐剂”下提供了合适的佐剂的实例。这样的佐剂可以与编码抗原的引发DNA疫苗一起施用,以与用编码仅抗原的DNA疫苗引发后产生的免疫应答相比增强抗原特异性的免疫应答。或者,这种佐剂可以与多肽抗原一起施用,所述多肽抗原以涉及本发明载体的施用方案施用。

[0079] 可以如下制备用于施用的腺病毒载体:悬浮或溶解于药学上或生理上可接受的载体例如等渗盐水、等渗盐、溶液或本领域技术人员将显而易见的其他制剂。适当的载体是本领域技术人员显而易见的,且很大程度上取决于施用途径。使用可生物降解的生物相容的聚合物,或通过使用胶束、凝胶和脂质体的现场递送,可以在持续释放制剂中将本文描述的组合物施用给哺乳动物。

[0080] 在一些实施方案中,通过肌内注射、静脉内注射、腹膜内注射、皮下注射、表皮施用、皮内施用、透皮施用、阴道内施用、鼻内施用、直肠施用或口腔施用将本发明的重组腺病毒施用于主体。

[0081] 如果治疗方案涉及共同施用一种或多种腺病毒载体和另外的组分,则将每种组分配制在不同的组合物中,它们有利地在相同位点或相同位点附近共位置施用。例如,可以将组分(例如通过选自肌内、透皮、皮内、皮下的施用途径)施用至同一侧或肢体(“同侧”施用)或相对侧或肢体(“对侧”施用)。

[0082] 病毒载体的剂量主要取决于诸如所治疗的病况、所治疗病况的严重程度和患者的年龄、体重和健康等因素,因此可能因患者而异。例如,病毒载体的治疗有效成人剂量通常含有 1×10^5 至 1×10^{15} 个病毒颗粒,例如从 1×10^8 至 1×10^{12} (例如, 1×10^8 、 5×10^8 、 1×10^9 、 5×10^9 、 1×10^{10} 、 2.5×10^{10} 、 5×10^{10} 、 1×10^{11} 、 5×10^{11} 或 1×10^{12} 个颗粒)。或者,病毒载体可以以通常 1×10^5 至 1×10^{10} 噬斑形成单位(PFU),例如 1×10^5 PFU、 5×10^5 PFU、 1×10^6 PFU、 5×10^6 PFU、 1×10^7 PFU、 5×10^7 PFU、 1×10^8 PFU、 5×10^8 PFU、 1×10^9 PFU、 5×10^9 PFU或 1×10^{10} PFU的剂量施用。剂量

将随主体的大小和施用途径而变化。例如,对于单个位点,用于肌肉注射的合适的人剂量(对于约80 kg主体)是在约 1×10^5 至约 5×10^{12} 个颗粒/ml的范围内。任选地,可以使用多个施用位点。在另一个实例中,对于口服制剂,合适的人或兽医学剂量可以是在 1×10^7 至约 1×10^{15} 个颗粒的范围内。

[0083] 可以例如使用基于CMV启动子区域设计的引物和探针,使用含有载体基因组(其含有包括人CMV (hCMV) 启动子的表达盒)的质粒DNA的系列稀释物作为标准曲线,通过定量PCR分析(Q-PCR)定量腺病毒载体。通过平行线分析方法确定测试样品中的拷贝数。用于载体颗粒定量的替代方法包括基于 A_{260nm} 的分析型HPLC或分光光度法。

[0084] 免疫有效量的核酸可适当地在1ng至100mg之间。例如,合适的量可以是1 μ g至100mg。“免疫有效量”是指向主体施用该量对于在主体中诱导针对狂犬病病毒的可测量的免疫应答是有效的。

[0085] 本领域技术人员可以容易地确定特定核酸(例如,载体)的适当量。

[0086] 核酸组分的示例性有效量可以是在1 ng至100 μ g之间,例如在1 ng至1 μ g之间(例如,100 ng-1 μ g),或在1 μ g至100 μ g之间,如10 ng、50 ng、100 ng、150 ng、200 ng、250 ng、500 ng、750 ng或1 μ g。核酸的有效量还可以包括1 μ g至500 μ g,例如1 μ g至200 μ g之间,例如10-100 μ g之间,例如1 μ g、2 μ g、5 μ g、10 μ g、20 μ g、50 μ g、75 μ g、100 μ g、150 μ g或200 μ g。或者,核酸的示例性有效量可以是在100 μ g至1 mg之间,例如100 μ g至500 μ g之间,例如,100 μ g、150 μ g、200 μ g、250 μ g、300 μ g、400 μ g、500 μ g、600 μ g、700 μ g、800 μ g、900 μ g或1 mg。

[0087] 通常,人剂量的体积为0.1ml至2ml,例如0.5ml和2ml。因而,可以将本文描述的组合物配制在例如0.1、0.25、0.5、1.0、1.5或2.0 ml人剂量/单一或组合免疫原性组分的体积中。

[0088] 本领域技术人员可以调节这些剂量,这取决于施用途径和采用重组载体的治疗或疫苗应用。可以监测转基因的表达水平,或对于佐剂而言,监测循环抗体的水平,以确定剂量施用的频率。

[0089] 如果使用一个或多个引发和/或强化步骤,那么该步骤可以包括每小时、每天、每周或每月或每年施用的单剂量。作为实例,哺乳动物可以接受含有在载体中约10 μ g至约50 μ g质粒的一个或两个剂量。期望地基于哺乳动物的身份和状况选择递送的量或位点。

[0090] 可以监测由选定的转基因编码的蛋白的治疗水平或针对所述蛋白的免疫应答的水平,以确定对强化的需要(如果有的话)。在评估血清中的CD8+ T细胞应答或任选的抗体滴度以后,可能期望任选的强化免疫。任选地,可以在单次施用中或在多种组合方案例如与涉及其他活性成分的方案或疗程组合或在引发-强化方案中递送腺病毒载体。

[0091] 重组腺病毒或包含多肽序列的组合物

适当地,本发明的多核苷酸是重组的。重组意味着所述多核苷酸是克隆、限制或连接步骤或产生不同于在自然界中发现的多核苷酸的多核苷酸的其他程序中的至少一者的产物。重组腺病毒是包含重组多核苷酸的腺病毒。重组载体是包含重组多核苷酸的载体。“重组病毒”包括原始重组病毒的后代。“重组载体”包括原始重组载体的重复。“重组多核苷酸”包括原始重组多核苷酸的重复。

[0092] 多肽的“功能衍生物”适当地是指多肽的经修饰形式,例如其中所述多肽的一个或多个氨基酸可以被缺失、插入、修饰和/或取代。例如,在以下情况下,认为未修饰的腺病毒

衣壳蛋白的衍生物是有功能的：

(a) 与包含未修饰的衣壳蛋白的腺病毒相比，在其衣壳内包含衍生衣壳蛋白的腺病毒保持基本上相同或更低的血清阳性率，和/或

(b) 与包含未修饰的衣壳蛋白的腺病毒相比，在其衣壳内包含衍生衣壳蛋白的腺病毒保持基本上相同或更高的宿主细胞感染性，和/或

(c) 与包含未修饰的衣壳蛋白的腺病毒相比，在其衣壳内包含衍生衣壳蛋白的腺病毒保持基本上相同或更高的免疫原性，和/或

(d) 与包含未修饰的衣壳蛋白的腺病毒相比，在其衣壳内包含衍生衣壳蛋白的腺病毒保持基本上相同或更高的转基因生产力水平。

[0093] 本发明的重组腺病毒或组合物适当地包含具有根据SEQ ID NO:1的氨基酸序列的多肽。适当地，本发明的重组腺病毒或组合物包含这样的多肽：其为具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列。适当地，具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性、例如至少85.0%同一性、例如至少90%同一性、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%同一性、例如至少99.2%同一性、例如至少99.4%同一性、例如99.5%同一性、例如至少99.6%同一性、例如至少99.8%同一性、例如99.9%同一性的氨基酸序列。或者，所述功能衍生物相对于SEQ ID NO:1具有不超过130个、更适当地不超过120个、更适当地不超过110个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0094] 根据本发明的重组腺病毒或组合物适当地还包含：

(a) 具有根据SEQ ID NO:3的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列，和/或

(a) 具有根据SEQ ID NO:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

[0095] 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列：其在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少91.0% 同一性、例如至少93.0% 同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5% 同一性、例如至少99.6%、例如至少99.7%同一性、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:3具有不超过300个、更适当地不超过250个、更适当地不超过200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超

过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0096] 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 其在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如99.7%同一性、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:5具有不超过500个、更适当地不超过400个、更适当地不超过450个、更适当地不超过300个、更适当地不超过250个、更适当地不超过200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0097] 本发明的重组腺病毒或组合物适当地包含具有根据SEQ ID NO:3的氨基酸序列的多肽。

[0098] 本发明的重组腺病毒或组合物适当地还包含：

(a) 具有根据SEQ ID NO:1的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列，
和/或

(a) 具有根据SEQ ID NO:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少60%同一性的氨基酸序列。

[0099] 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 其在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少60.0%同一性、例如至少70.0%同一性、例如至少80.0%同一性、例如至少85.0%同一性、例如至少87.0%同一性、例如至少89.0%同一性、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%同一性、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:1具有不超过130个、更适当地不超过120个、更适当地不超过110个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0100] 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基

酸序列： 其在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少95.0%、例如至少97.0%、例如至少99.0%、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:5具有不超过500个、更适当地不超过400个、更适当地不超过450个、更适当地不超过300个、更适当地不超过250个、更适当地不超过200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0101] 本发明的重组腺病毒或组合物适当地包含编码具有根据SEQ ID NO:1的氨基酸序列的多肽的多核苷酸。所述多核苷酸适当地具有根据SEQ ID NO:2的序列。

[0102] 或者，本发明的重组腺病毒或组合物包含编码具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物的多核苷酸，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性的氨基酸序列。具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列：其在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少80%同一性、例如至少85.0%同一性、例如至少90%同一性、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%同一性、例如至少99%同一性、例如至少99.4%同一性、例如至少99.6%同一性或例如至少99.8%同一性。或者，所述功能衍生物相对于SEQ ID NO:1具有不超过130个、更适当地不超过120个、更适当地不超过110个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0103] 本发明的重组腺病毒或组合物适当地还包含这样的多核苷酸，其编码：

(a) 具有根据SEQ ID NO:3的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少50.0%同一性的氨基酸序列，和/或

(a) 具有根据SEQ ID NO:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

[0104] 具有根据SEQ ID NO:3的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 其在其整个长度上与SEQ ID NO:3的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%、例如至少99%、例如至少99.4%、例如至少99.6%、例如至少99.8%同一性。或者，所述功能衍生物相对于SEQ ID NO:3具有不超过300个、更适当地不超过250个、更适当地不超过

200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0105] 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 其在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少95.0%、例如至少97.0%、例如至少98.0%、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如99.7%同一性、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:5具有不超过500个、更适当地不超过400个、更适当地不超过450个、更适当地不超过300个、更适当地不超过250个、更适当地不超过200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0106] 本发明的重组腺病毒或组合物适当地包含编码具有根据SEQ ID NO:3的氨基酸序列的多肽的多核苷酸。所述多核苷酸适当地具有根据SEQ ID NO:4的序列。

[0107] 本发明的重组腺病毒或组合物适当地还包含这样的多核苷酸，其编码：

(a) 具有根据SEQ ID NO:1的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少50%同一性的氨基酸序列，
和/或

(a) 具有根据SEQ ID NO:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物，其中所述功能衍生物具有在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少50%同一性的氨基酸序列。

[0108] 具有根据SEQ ID NO:1的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 其在其整个长度上与SEQ ID NO:1的氨基酸序列具有至少60.0%同一性、例如至少70.0%同一性、例如至少80.0%同一性、例如至少85.0%同一性、例如至少87.0%同一性、例如至少89.0%同一性、例如至少91.0%同一性、例如至少93.0%同一性、例如至少95.0%同一性、例如至少97.0%同一性、例如至少98.0%同一性、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如99.7%同一性、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:1具有不超过130个、更适当地不超过120个、更适当地不超过110个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0109] 具有根据SEQ ID NO:5的氨基酸序列的多肽的功能衍生物适当地具有这样的氨基酸序列： 在其整个长度上与SEQ ID NO:5的氨基酸序列具有至少60.0%、例如至少70.0%、例如至少80.0%、例如至少85.0%、例如至少90.0%、例如至少95.0%、例如至少97.0%、例如至少98.0%、例如至少99.0%、例如至少99.2%、例如至少99.4%、例如99.5%同一性、例如至少99.6%、例如99.7%同一性、例如至少99.8%同一性、例如99.9%同一性。或者，所述功能衍生物相对于SEQ ID NO:5具有不超过500个、更适当地不超过400个、更适当地不超过450个、更适当地不超过300个、更适当地不超过250个、更适当地不超过200个、更适当地不超过150个、更适当地不超过125个、更适当地不超过100个、更适当地不超过90个、更适当地不超过80个、更适当地不超过70个、更适当地不超过60个、更适当地不超过50个、更适当地不超过40个、更适当地不超过30个、更适当地不超过20个、更适当地不超过10个、更适当地不超过5个、更适当地不超过4个、更适当地不超过3个、更适当地不超过2个并且更适当地不超过1个添加、缺失和/或取代。

[0110] *ChAd155*主链

本申请描述了黑猩猩腺病毒ChAd155的分离的多核苷酸序列，其包括野生型未修饰的ChAd155的分离的多核苷酸序列 (SEQ ID NO:10) 和ChAd155的经修饰的主链构建体。这些修饰的主链构建体包括ChAd155#1434 (SEQ ID NO: 7)、ChAd155#1390 (SEQ ID NO: 8) 和ChAd155#1375 (SEQ ID NO: 9)。ChAd155主链可以用于构建复制型的或复制缺陷型的重组腺病毒用于递送转基因。

[0111] 术语“构建体”是指编码本文所述多肽序列的核酸，并且可包含DNA或非天然存在的核酸单体。

[0112] 术语“复制型”腺病毒表示在没有细胞所包含的任何重组辅助蛋白存在下可以在所述宿主细胞中复制的腺病毒。适当地，“复制型”腺病毒包含以下完整或功能性必需早期基因：E1A、E1B、E2A、E2B、E3和E4。从特定动物分离的野生型腺病毒在该动物中将是复制型的。

[0113] 术语“复制缺陷型”或“复制缺陷的”腺病毒表示不能复制的腺病毒，因为它已经被工程改造成至少包含功能缺失(或“功能丧失”突变)，即在不将基因完全除去的情况下损害该基因的功能的缺失或突变，例如人工终止密码子的引入、活性位点或相互作用结构域的缺失或突变、基因的调节序列等的突变或缺失，或编码病毒复制所必需的基因产物的基因的完全除去，例如选自E1A、E1B、E2A、E2B、E3和E4 (例如E3 ORF1、E3 ORF2、E3 ORF3、E3 ORF4、E3 ORF5、E3 ORF6、E3 ORF7、E3 ORF8、E3 ORF9、E4 ORF7、E4 ORF6、E4 ORF4、E4 ORF3、E4 ORF2和/或E4 ORF1) 的腺病毒基因中的一个或多个。特别适当地E1和任选地E3和/或E4是缺失的。如果缺失的话，当确定相对于另一个序列的%同一性时，适当地将在比对中不考虑前述缺失的基因区域。

[0114] 本发明的序列可用作治疗剂和用于构建多种载体系统、重组腺病毒和宿主细胞。适当地，术语“载体”是指相对于野生型序列已经实质上改变(例如，已经缺失和/或失活的基因或功能区域)和/或掺入了异源序列的核酸，即，获得自不同来源(也称为“插入物”)且在引入细胞(例如，宿主细胞)中时复制和/或表达插入的多核苷酸序列的核酸。例如，所述插入物可以是本文描述的ChAd155序列的全部或部分。另外或可选地，ChAd155载体可以是包含病毒基因(例如E1或本文描述的其他病毒基因或功能区域)的一个或多个缺失或失活

的ChAd155腺病毒。这样的ChAd155 (其可以包含或不包含异源序列) 经常被称作“主链”, 且可以原样使用或作为对载体的另外修饰的起始点使用。

[0115] 下面提供了ChAd155野生型序列 (SEQ ID NO:10) 序列的注解。

基因座	ChAd155	37830 bpDNA	线性	2015 年 6 月 10 日
定义	黑猩猩腺病毒 155, 完整基因组。			
解释	根据 ChAd155 与人类腺病毒 2 参考毒株 NC_001405 的比对的注释 在 E3 区域中手工地添加两个推定 ORF			
特征	位置/限定符			
来源	1..37830			
	/生物体="黑猩猩腺病毒 155"			
	/mol_类型="基因组 DNA"			
	/首字母简略词="ChAd155"			
重复_区域	1..101			
	/标准_名称="ITR"			
	/rpt_类型=反向			
基因	466..1622			
	/基因="E1A"			
TATA_信号	466..471			
	/基因="E1A"			
prim_转录物	497..1622			
	/基因="E1A"			
CDS	连接(577..1117,1231..1532)			
	/基因="E1A"			
	/产物="E1A_280R"			
CDS	连接(577..979,1231..1532)			
	/基因="E1A"			
	/产物="E1A_243R"			
polyA_信号	1600..1605			
	/基因="E1A"			
基因	1662..4131			
	/基因="E1B"			
TATA_信号	1662..1667			
	/基因="E1B"			
prim_转录物	1692..4131			
	/基因="E1B"			
CDS	1704..2267			
	/基因="E1B"			
	/产物="E1B_19K"			
CDS	2009..3532			
	/基因="E1B"			
	/产物="E1B_55K"			
基因	3571..4131			

/基因="IX"
TATA_信号 3571..3576
 /基因="IX"
prim_转录物 3601..4131
 /基因="IX"
CDS 3628..4092
 /基因="IX"
 /产物="IX"
polyA_信号 4097..4102
 /注释="E1B, IX"
基因 互补物(4117..27523)
 /基因="E2B"
prim_转录物 互补物(4117..27494)
 /基因="E2B"
基因 互补物(4117..5896)
 /基因="IVa2"
prim_转录物 互补物(4117..5896)
 /基因="IVa2"
CDS 互补物(连接(4151..5487,5766..5778))
 /基因="IVa2"
 /产物="E2B_IVa2"
polyA_信号 互补物(4150..4155)
 /注释="IVa2, E2B"
CDS 互补物(连接(5257..8838,14209..14217))
 /基因="E2B"
 /产物="E2B_聚合酶"
基因 6078..34605
 /基因="L5"
基因 6078..28612
 /基因="L4"
基因 6078..22658
 /基因="L3"
基因 6078..18164
 /基因="L2"
基因 6078..14216
 /基因="L1"
TATA_信号 6078..6083
 /注释="L"
prim_转录物 6109..34605

/基因="L5"
prim_转录物 6109..28612
/基因="L4"
prim_转录物 6109..22658
/基因="L3"
prim_转录物 6109..18164
/基因="L2"
prim_转录物 6109..14216
/基因="L1"
CDS 连接(8038..8457,9722..9742)
/基因="L1"
/产物="L1_13.6K"
CDS 互补物(连接(8637..10640,14209..14217))
/基因="E2B"
/产物="E2B_pTP"
基因 10671..10832
/基因="VAI"
misc_RNA 10671..10832
/基因="VAI"
/产物="VAI"
基因 10902..11072
/基因="VAII"
misc_RNA 10902..11072
/基因="VAII"
/产物="VAII"
CDS 11093..12352
/基因="L1"
/产物="L1_52K"
CDS 12376..14157
/基因="L1"
/产物="L1_pIIIa"
polyA_信号 14197..14202
/基因="L1"
CDS 14254..16035
/基因="L2"
/产物="L2_五邻体"
CDS 16050..16646
/基因="L2"
/产物="L2_pVII"

CDS 16719..17834
 /基因="L2"
 /产物="L2_V"
CDS 17859..18104
 /基因="L2"
 /产物="L2_pX"
polyA_信号 18143..18148
 /基因="L2"
CDS 18196..18951
 /基因="L3"
 /产物="L3_pVI"
CDS 19063..21945
 /基因="L3"
 /产物="L3_六邻体"
CDS 21975..22604
 /基因="L3"
 /产物="L3_蛋白酶"
polyA_信号 22630..22635
 /基因="L3"
基因 互补物(22632..27523)
 /基因="E2A"
prim_转录物 互补物(22632..27494)
 /基因="E2A"
基因 互补物(22632..26357)
 /基因="E2A-L"
prim_转录物 互补物(22632..26328)
 /基因="E2A-L"
polyA_信号 互补物(22649..22654)
 /注释="E2A, E2A-L"
CDS 互补物(22715..24367)
 /基因="E2A"
 /注释="DBP; 常规属 (genus-common); DBP 家族"
 /起始密码子=1
 /产物="E2A"
CDS 24405..26915
 /基因="L4"
 /产物="L4_100k"
TATA_信号 互补物(26352..26357)
 /基因="E2A-L"

CDS 连接(26602..26941,27147..27529)
/基因="L4"
/产物="L4_33K"

CDS 26602..27207
/基因="L4"
/产物="L4_22K"

TATA_信号 互补物(27518..27523)
/注释="E2A, E2B: 标称"

CDS 27604..28287
/基因="L4"
/产物="L4_pVIII"

基因 27969..32686
/基因="E3B"

基因 27969..31611
/基因="E3A"

TATA_信号 27969..27974
/注释="E3A, E3B"

prim_转录物 27998..32686
/基因="E3B"

prim_转录物 27998..31611
/基因="E3A"

CDS 28288..28605
/基因="E3A"
/产物="E3 ORF1"

polyA_信号 28594..28599
/基因="L4"

CDS 29103..29303
/基因="E3A"
/产物="E3 ORF2"

CDS 29300..29797
/基因="E3A"
/产物="E3 ORF3"

CDS 29826..30731
/基因="E3A"
/产物="E3 ORF4"

CDS 30728..31579
/基因="E3A"
/产物="E3 ORF5"

CDS 31283..31579


```

        /基因="E3A"
        /产物="E3 ORF6"
polyA_信号      31578..31584
        /基因="E3A"
CDS      31591..31863
        /基因="E3B"
        /产物="E3 ORF7"
CDS      31866..32264
        /基因="E3B"
        /产物="E3 ORF8"
CDS      32257..32643
        /基因="E3B"
        /产物="E3 ORF9"
polyA_信号      32659..32664
        /基因="E3B"
基因      互补物(<32678..32838)
        /基因="U"
CDS      互补物(<32678..32838)
        /基因="U"
        /注释="编码未鉴定的 C 末端的外显子；
        常规属"
        /产物="蛋白 U"
CDS      32849..34585
        /基因="L5"
        /产物="L5_纤突"
polyA_信号      34581..34586
        /基因="L5"
基因      互补物(34611..37520)
        /基因="E4"
prim_转录物 互补物(34611..37490)
        /基因="E4"
polyA_信号      互补物(34625..34630)
        /基因="E4"
CDS      互补物(连接(34794..35069,35781..35954))
        /基因="E4"
        /产物="E4 ORF7"
CDS      互补物(35070..35954)
        /基因="E4"
        /产物="E4 ORF6"

```

```
CDS      互补物(35875..36219)
          /基因="E4"
          /产物="E4 ORF4"
CDS      互补物(36235..36582)
          /基因="E4"
          /产物="E4 ORF3"
CDS      互补物(36579..36971)
          /基因="E4"
          /产物="E4 ORF2"
CDS      互补物(37029..37415)
          /基因="E4"
          /产物="E4 ORF1"
TATA_信号 互补物(37515..37520)
          /基因="E4"
重复_区域      37740..37830
          /标准_名称="ITR"
          /rpt_类型=反向
```

[0116] 序列同一性

关于序列的同一性在本文中定义为,在比对序列并按需要引入空位以达到最大序列同一性百分比,并且不将任何保守取代视作序列同一性的一部分以后与参照氨基酸序列相同的候选序列中的氨基酸残基的百分比。

[0117] 序列同一性可以通过标准方法确定,所述标准方法通常用于比较两个多肽的氨基酸位置的相似性。使用如BLAST或FASTA的计算机程序,比对两个多肽以使它们各自的氨基酸最佳匹配(沿着一个或两个序列的全长或沿着一个或两个序列的预定部分)。程序提供默认开放罚分和默认空位罚分,以及如PAM 250的评分矩阵(标准评分矩阵可以与计算机程序一起使用。例如,百分比同一性可以计算为相同匹配的总数乘以100,然后除以匹配范围内较长序列的长度和引入较短序列中以便比对两个序列的空位数之和。

[0118] 在本公开通过参考UniProt或Genbank登录码指代序列的情况下,所指的序列是截至本申请的提交日期的当前版本。

[0119] 技术人员会认识到,改变、添加或缺失单个氨基酸或小百分比的氨基酸的对蛋白的个别取代、缺失或添加是这样的“免疫原性衍生物”,其中所述改变导致在功能上类似的氨基酸对氨基酸的取代或者不会实质上影响免疫原性功能的残基的取代/缺失/添加。

[0120] 提供在功能上类似的氨基酸的保守取代表是本领域众所周知的。一般而言,这样的保守取代将落入下文详细说明了氨基酸分组之一中,尽管在一些情况下,其他取代也会是可能的,而不实质上影响所述抗原的免疫原性特性。以下八组各自含有通常彼此保守取代的氨基酸:

- 1) 丙氨酸 (A), 甘氨酸 (G);
- 2) 天冬氨酸 (D), 谷氨酸 (E);

- 3) 天冬酰胺 (N), 谷氨酰胺 (Q);
- 4) 精氨酸 (R), 赖氨酸 (K);
- 5) 异亮氨酸 (I), 亮氨酸 (L), 甲硫氨酸 (M), 缬氨酸 (V);
- 6) 苯丙氨酸 (F), 酪氨酸 (Y), 色氨酸 (W);
- 7) 丝氨酸 (S), 苏氨酸 (T);和
- 8) 半胱氨酸 (C), 甲硫氨酸 (M)。

[0121] 适当地,这样的取代不发生在表位区域,且因此对所述抗原的免疫原性特性不具有显著影响。

[0122] 免疫原性衍生物还可以包括与参照序列相比在其中插入了另外氨基酸的那些。适当地,这样的插入不发生在表位区域,且因此对所述抗原的免疫原性特性不具有显著影响。插入的一个实例包括组氨酸残基的短片段(例如2-6个残基)(SEQ ID NO: 49)以帮助目标抗原的表达和/或纯化。

[0123] 免疫原性衍生物包括与参照序列相比其中已经缺失氨基酸的那些。适当地,这样的缺失不发生在表位区域,且因此对所述抗原的免疫原性特性不具有显著影响。

[0124] 技术人员会认识到,特定的免疫原性衍生物可以包含取代、缺失和添加(或它们的任意组合)。

[0125] 狂犬病毒抗原和疫苗

狂犬病病毒属是弹状病毒科的一个属,是具有单链反义RNA基因组的包膜病毒。RNA以核蛋白(N)、磷蛋白(P)、基质蛋白(M)、糖蛋白(G)和病毒RNA聚合酶(L)的顺序编码五种结构蛋白。P蛋白是核糖核蛋白的结构组分,并且在病毒颗粒的形成和病毒RNA合成中起作用。G蛋白被认为在病毒致病性和保护性免疫中很重要;它是保护性中和抗体的主要靶标。狂犬病病毒是一种嗜神经病毒,其通过中枢神经系统传播,导致大脑和脊髓的严重炎症。

[0126] 狂犬病病毒属包含七种基因型,其中以下六种已经与人狂犬病病例有关:狂犬病病毒(RABV,基因型1)、莫科拉病毒(基因型3)、杜文海格病毒(基因型4)、欧洲蝙蝠狂犬病病毒(基因型5)、欧洲蝙蝠狂犬病病毒 2(基因型6)和澳大利亚蝙蝠狂犬病病毒(基因型7)(Jackson (2016) Curr Infec Dis Rep 18:38)。一旦出现症状,狂犬病几乎百分之百致命。

[0127] 已经在多种狂犬病病毒毒株中鉴定了存在于狂犬病G蛋白上的抗原表位。它们被划分为位点I、位点IIa、位点IIb、位点III、位点IV和位点a并被列于表1。该表以出现顺序分别公开了作为SEQ ID NOS 50-63的'位点 II b'序列,作为SEQ ID NOS 64-77的'位点I'序列,和作为SEQ ID NOS 78-91的'位点III'序列。

[0128] 表1 狂犬病G蛋白抗原表位

病毒	种系组	位点 II b (34-42)	位点 II a (198-200)	位点 I (226-231)	位点 IV (263-264)	位点 III (330-338)	位点 'a' (342-343)
RABV	I	GCTNLSEFS	KRA	KLCGVL	PH	KSVRTWNEI	KG
ABLV	I	GCTSLSGFS	KKA	KLCGIS	FN	KSVRTWDEI	KG
ARAV	I	GCTNLSGFT	KKA	KLCGVM	PH	KSVREWTEV	KG
BBLV	I	GCTTLTVFS	KKA	KLCGVS	PH	KSIRQWTEI	KG
DUVV	I	GCTTLTPFS	KKA	RLCGIS	PH	KSVREWKEI	KG
EBLV-1	I	GCTTLTPFS	KKA	RLCGVP	PH	KSVREWKEV	KG
EBLV-2	I	GCTTLTVFS	KKA	KLCGIS	PH	KSIREWTDV	KG
IRKV	I	GCTTLTAFN	KKA	KLCGMA	DR	KSIREWKEI	KG
KHUV	I	GCTTLSGFT	KRA	KLCGVS	PH	KSIREWSEI	KG
LBV	II	GCSDTATFS	KKS	TLCGKP	NR	LRVDSWNDI	KG
MOKV	II	GCNTESPFT	QKA	TLCGKP	DR	KRVDRWADI	KG
SHIV	II	GCSSSSTFS	KKS	TLCGKP	NR	KRVDRWEEI	KG
WCBV	III	YCTTEQSIT	KLV	SICGRQ	IK	IKVENWSEV	KG
IKOV	?	GCNEGSKVS	ILL	IICGKS	VK	KSVDNWTDI	PI

[0129] 在图1中显示了存在于对应于SEQ ID NO:37的狂犬病G蛋白上的抗原表位。抗原位点I携带构象和线性表位,并且位于氨基酸残基226-231。抗原位点II是残基34-42(IIb)和198-200(IIa)处的不连续构象表位。抗原位点III是残基330-338处的连续构象表位。抗原位点IV位于残基263-264处。抗原位点a位于残基342-343处。

[0130] 狂犬病疫苗目前主要用于暴露后预防,只有一小部分狂犬病疫苗剂量用于暴露前预防。干预时间表由世界卫生组织根据病毒进入的伤口的严重性和类型来定义,并且可以包括用抗狂犬病免疫球蛋白的额外治疗。暴露前预防通常涉及2至3次肌肉剂量的2-3次随访,及根据暴露风险定时加强。暴露后预防通常涉及4-5次肌肉剂量的3-5次随访或4次皮内剂量的4次随访。在一些欠发达国家,仍然通过在受感染动物的大脑中繁殖狂犬病病毒,灭活病毒并且提供每天皮下给予的14-21次注射到腹壁来进行免疫接种。

[0131] 目前有多种狂犬病疫苗可用于人体暴露前和暴露后预防。IMOVAX (Sanofi Pasteur)作为从Wistar Institute获得的毒株PM-1503-3M制备的冷冻干燥的狂犬病病毒而提供。其从感染的人二倍体细胞收获然后灭活。暴露前和暴露后预防由在第0、7和21或28天肌肉施用三次剂量组成。VERORAB (Sanofi Pasteur)作为从Wistar Institute获得的毒株PM/WI 38 1503-3M制备的冷冻干燥的狂犬病病毒提供。从Vero细胞收获然后灭活。暴露前预防由在第0、7和21或28天肌肉施用三次剂量组成。暴露后预防由在第0、3、7、14和28天肌肉施用五次剂量组成。VAXIRAB/ LYSSAVAC (Zydus Cadila/ Novavax)作为由狂犬病病毒的Pitman Moore毒株制备的冷冻干燥的狂犬病病毒提供。它在鸭胚细胞中产生然后灭活。暴露前预防由在第0、7和21或28天肌肉施用三次剂量组成。暴露后预防由在第0、3、7、14和28天肌肉施用五次剂量组成。暴露后预防也可以皮内施用,在第0、3、7和28天在两个部位的每一个处注射。RABIPUR/ RABAVER (GSK)作为从Flury LEP(低胚传代)毒株制备的冷冻

干燥的狂犬病病毒提供。它在鸡成纤维细胞的原代培养物中生长,然后灭活。暴露前预防由在第0、7和21或28天肌肉内施用三次剂量组成。暴露后预防由在第0、3、7、14和28天肌肉内施用五次剂量组成。

[0132] 在文献中已经报道了腺病毒载体的(adeno-vectored)狂犬病疫苗的支持性临床前证据。腺病毒重组病毒载体SAdV24,也称为AdC68或ChAd68,被修饰为复制缺陷并且表达狂犬病的Evelyn Rokitniki Abelseth(ERA)毒株的全长糖蛋白(G),当在狂犬病攻击之前给予时,在食蟹猴中显示出一定程度的免疫原性,但在狂犬病暴露后没有提供可靠的保护(Xiang 等人(2014) *Virol.* 450-451:243-249)。肌肉内给予的表达Evelyn Rokitniki Abelseth(ERA)狂犬病毒株的全长糖蛋白(G)的类似复制缺陷型ChAd68载体针对狂犬病攻击诱导一定程度的保护作用(Zhou 等人(2006) *Mol. Ther.* 14:662-672;部分再现于图16)。

[0133] 佐剂

如本文中使用的“佐剂”是指增强对免疫原的免疫应答的组合物。包含佐剂的根据本发明的组合物可以用作疫苗,例如用于人类主体。与单独施用抗原相比,佐剂加速、延长和/或增强对抗原/免疫原的免疫应答的质量和/或强度,因此,减少任何给定疫苗中必需的抗原/免疫原的量,和/或所必需的注射频率以对目标抗原/免疫原产生足够的免疫应答。

[0134] 可用于本发明组合物背景中的佐剂的实例包括无机佐剂(例如无机金属盐,例如磷酸铝或氢氧化铝)、氢氧化铝(明矾)的凝胶状沉淀物;AlPO₄;铝胶;来自革兰氏阴性细菌外膜的细菌产物,特别是单磷酸脂质A(MPLA)、脂多糖(LPS)、胞壁酰二肽及其衍生物;弗氏不完全佐剂;脂质体,特别是中性脂质体,含有该组合物和任选细胞因子的脂质体;AS01B、AS01E、AS02;非离子嵌段共聚物;ISCOMATRIX佐剂;包含CpG二核苷酸(CpG基序)的未甲基化DNA,特别是具有硫代磷酸酯(PTO)骨架(CpG PTO ODN)或磷酸二酯(PO)骨架(CpG PO ODN)的CpG ODN;合成的脂肽衍生物,特别是Pam₃Cys;脂阿拉伯甘露聚糖;肽聚糖;酵母聚糖;热休克蛋白(HSP),特别是HSP 70;dsRNA及其合成衍生物,特别是Poly I:poly C;聚阳离子肽,特别是聚-L-精氨酸;紫杉醇;纤连蛋白;鞭毛蛋白;咪唑并喹啉;具有佐剂活性的细胞因子,特别是GM-CSF、白细胞介素(IL-1)2、IL-6、IL-7、IL-18、I型和II型干扰素,特别是干扰素- γ (IFN- γ)、TNF- α ;25-二羟基维生素D3(骨化三醇);和合成的寡肽,特别是MHCII呈递的肽。含有聚氧乙烯(POE)和聚氧丙烯(POP)的非离子嵌段聚合物,例如POE-POP-POE嵌段共聚物可用作佐剂。

[0135] 佐剂的其他实例包括无机佐剂(例如无机金属盐,例如磷酸铝或氢氧化铝)、有机佐剂(例如皂苷,例如QS21或角鲨烯)、油基佐剂(例如弗氏完全佐剂和弗氏不完全佐剂)、细胞因子(例如IL-1 β 、IL-2、IL-7、IL-12、IL-18、GM-CSF和INF- γ)颗粒佐剂(例如免疫刺激复合物(ISCOMS)、脂质体、可生物降解的微球、病毒体,细菌佐剂(例如单磷酸脂质A,例如3-脱-O-酰化单磷酸脂质A(3D-MPL)或胞壁酰肽)、合成佐剂(例如单磷酸脂质A(MPL),特别是3-脱-O-酰化单磷酸脂质A(3D-MPL和胞壁酰肽类似物,或合成的脂质A,和合成的多核苷酸佐剂,例如聚精氨酸或聚赖氨酸。

[0136] 皂苷也是合适的佐剂,例如,源自南美树皂树Molina的树皮的皂苷Quil A及其级分。Quil A的纯化级分也称为免疫刺激剂,如角鲨烯、QS21、QS17和QS7、Quil-A的非溶血性级分。QS21和聚山梨醇酯或环糊精的组合也是合适的。

[0137] 佐剂的另一个实例是免疫刺激寡核苷酸,其含有DNA中存在的未甲基化的胞嘧啶-鸟苷二核苷酸基序(“CpG”)。当通过全身和粘膜途径施用, CpG被称为佐剂。当配制成疫苗时,它可以与游离抗原一起在游离溶液中施用或与抗原共价缀合或与载体如氢氧化铝一起配制。

[0138] 特异性受体的活化可以刺激免疫应答。这些受体是本领域技术人员已知的,并且包括,例如,细胞因子受体,特别是I型细胞因子受体、II型细胞因子受体、TNF受体;和作为转录因子的维生素D受体;和Toll样受体1 (TLR1)、TLR-2、TLR3、TLR4、TLR5、TLR-6、TLR7和TLR9。这些受体的激动剂具有佐剂活性,即是免疫刺激性的。其他合适的佐剂包括烷基氨基葡萄糖苷磷酸酯(AGP)或AGP的药学上可接受的盐。一些AGP是TLR4激动剂,而一些是TLR4拮抗剂。本发明组合物的佐剂可以是一种或多种Toll样受体激动剂。在更优选的实施方案中,佐剂是Toll样受体4激动剂。在特别优选的实施方案中,佐剂是Toll样受体9激动剂。

[0139] 佐剂如上述的那些可以与载体如脂质体、水包油乳剂和/或金属盐(包括铝盐如氢氧化铝)一起配制。例如,3D-MPL可以用氢氧化铝或水包油乳剂配制;QS21可以与含胆固醇的脂质体、水包油乳剂或明矾配制;CpG可以与明矾或其他阳离子载体配制。

[0140] 佐剂的组合可用于本发明,特别是单磷酸脂质A和皂苷衍生物的组合,更特别是QS21和3D-MPL的组合或其中QS21在含胆固醇的脂质体中淬灭的组合物(DQ)。或者,CpG加皂苷如QS21的组合是适用于本发明的佐剂,如同在水包油乳剂中涉及QS21、3D-MPL和生育酚的有效佐剂制剂。可以将皂苷佐剂配制在脂质体中,并与免疫刺激的寡核苷酸组合。因此,合适的佐剂系统包括,例如,单磷酸脂质A,优选3D-MPL与铝盐的组合。另一种示例性的佐剂包含QS21和/或MPL和/或CpG。QS21可以在含胆固醇的脂质体中淬灭。

[0141] 如果与腺病毒一起施用,则恒定链与由用于疫苗接种的表达系统所包含的抗原的融合增加了针对所述抗原的免疫应答。因此,在本发明的一个实施方案中,免疫原性的转基因可以与重组ChAd155病毒载体中的恒定链共表达。

[0142] 在另一个实施方案中,本发明提供了通过将ChAd155衣壳递送给主体的ChAd155的衣壳(任选地使用完整或重组病毒颗粒或空衣壳)诱导免疫调节应答或增强或辅助对另一种活性剂的细胞毒性T细胞应答的用途。可以将ChAd155衣壳单独递送,或在组合方案中与活性剂一起递送以增强对其的免疫应答。有利地,在不用腺病毒感染宿主的情况下,可以实现期望的效应。

[0143] 一般原则

除非另有解释,本文使用的所有技术和科学术语具有如本公开内容所属技术领域的普通技术人员通常理解的相同含义。单数术语“一种”、“一个”和“所述/该”包括复数指示对象,除非上下文清楚地另外指示。类似地,词语“或”意在包括“和”,除非上下文清楚地另外指示。术语“复数”是指两个/种或更多。另外,关于物质的浓度或水平(例如溶液组分浓度或其比率)和反应条件(例如温度、压力和循环时间)给出的数值限制旨在是近似的。本文使用的术语“约”意指量 $\pm 10\%$ 。

[0144] 将参考以下非限制性实施例和附图进一步描述本发明。

实施例

[0145] 实施例1:ChAd155的分离

使用如Colloca 等人 (2012) Sci. Transl. Med. 4:1-9和WO 2010086189 (其为描述腺病毒分离和表征技术的目的通过引用并入本文) 中描述的标准程序从 New Iberia Research Center 设施 (New Iberia Research Center, The University of Louisiana at Lafayette) 圈养的健康年幼黑猩猩分离野生型黑猩猩腺病毒155型 (ChAd155)。

[0146] 实施例2:ChAd155-RG载体构建

然后将ChAd155病毒基因组克隆到质粒或BAC载体中,并随后如图3所示进行修饰:

- a) 病毒基因组的E1区 (从bp 449到bp 3529) 的缺失;
- b) 病毒基因组的E4区域 (从bp 34731到bp 37449) 的缺失;
- c) 源自人Ad5的E4orf6的插入; 和
- d) hCMV-RG-WPRE表达盒的插入。

[0147] 2.1: E1区域的缺失: BAC/ChAd155 Δ E1_Tet0 hCMV RpsL-Kana#1375的构建

在用ChAd155病毒DNA和亚组C BAC 穿梭物 (#1365) 共转化的大肠杆菌菌株BJ5183电穿孔感受态细胞 (Stratagene目录号2000154) 中通过同源重组将ChAd155病毒基因组克隆到BAC载体中。如图4的示意图所示,亚组C 穿梭物是源自pBeloBAC11 (GenBank U51113, NEB) 的BAC载体。它致力于克隆属于物种C的黑猩猩腺病毒,并因此含有pIX基因和来自物种C ChAd病毒的右端和左端 (包括右和左ITR) 的DNA片段。

[0148] 物种C BAC 穿梭物还含有插入左端和pIX基因之间的RpsL-Kana盒。另外,侧翼为ISceI限制性位点的Amp-LacZ-SacB选择盒存在于pIX基因和病毒基因组的右端之间。特别地,BAC 穿梭物包括以下特征:左ITR:bp 27至139,hCMV (tet0) RpsL-Kana盒:bp 493至3396,pIX基因:bp 3508至3972,ISceI限制性位点:bp 3990和7481,Amp-LacZ-SacB选择盒:bp 4000至7471,右ITR:bp 7805至7917。

[0149] 用ChAd155纯化的病毒DNA和经ISceI限制酶消化的亚组C BAC 穿梭物载体通过电穿孔共转化BJ5183细胞,然后凝胶纯化。在pIX基因和右ITR序列 (存在于物种C BAC 穿梭物线性化DNA的末端) 之间发生的同源重组和ChAd155病毒DNA中存在的同源序列导致ChAd155病毒基因组DNA插入BAC穿梭载体中。同时,缺失病毒E1区并用RpsL-Kana盒代替,产生BAC/ChAd155 Δ E1/ Tet0 hCMV RpsL-Kana#1375。

[0150] 2.2: 在大肠杆菌BJ5183中通过同源重组构建质粒

2.2.1: E4区的缺失 - pChAd155 Δ E1, E4_Ad5E4orf6/Tet0 hCMV RpsL-Kana (#1434) 的构建

为了提高载体的增殖,通过使用涉及在大肠杆菌中克隆和同源重组的几个步骤的策略用Ad5 E4orf6编码序列替换天然E4区域,在载体主链中引入跨越核苷酸34731-37449 (ChAd155野生型序列) 的E4区域的缺失。E4编码区完全缺失,而E4天然启动子和多腺苷酸化信号被保留。为此,构建穿梭载体以允许通过在大肠杆菌 BJ5183中同源重组替换ChAd155天然E4区来插入Ad5orf6,如下所述。

[0151] pARS 物种C Ad5E4orf6-1的构建

使用Ad5 DNA作为模板,利用寡核苷酸为 5'-ATACGGACTAGTGGAGAAGTACTCGCCTACATG-3' (SEQ ID NO: 13) 和5'-ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC-3' (SEQ ID NO: 14) 通过PCR获得含有Ad5orf6的DNA片段。用BglIII和SpeI消化PCR片段,并克隆到用BglIII和SpeI消化的物种C RLD-EGFP穿梭物中,产生质粒pARS 物种C Ad5orf6-1。有关穿梭物的

详细信息可见于Colloca 等人, Sci. Transl. Med. (2012) 4:115ra2。

[0152] pARS 物种C Ad5E4orf6-2的构建

为了缺失E4区域,使用质粒BAC/ChAd155 Δ E1_TetO hCMV RpsL-Kana (#1375)作为模板,利用以下寡核苷酸5'-ATTCAGTGACAGGCGCGCCAAAGCATGACGCTGTTGATTGATTC-3' (SEQ ID NO: 15)和5'-ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC-3' (SEQ ID NO: 16)通过PCR扩增跨越ChAd155 wt序列(SEQ ID NO:10)的bp 34586至bp 34730的177 bp DNA片段。用BsrGI和SpeI消化PCR片段,并克隆到用BsrGI和SpeI消化的pARS 亚组C Ad5orf6-1中,产生质粒pARS 物种C Ad5orf6-2 (#1490)。在图5中提供该穿梭质粒的示意图。特别地,穿梭质粒包含以下特征:左ITR:bp 1至113,物种C第一个460bp:bp 1至460,ChAd155 wt (SEQ ID NO:10)的bp 34587至bp 34724):bp 516至650, Ad5 orf6:bp 680和1561,物种C最后393 bp:bp 1567至1969,右ITR:bp 1857至1969。

[0153] pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434)的构建

然后使用所得质粒pARS 亚组C Ad5orf6-2用Ad5orf6替换ChAd155主链内的E4区。为此,用PacI/PmeI消化质粒BAC/ChAd155 Δ E1_TetO hCMV RpsL-Kana (#1375),并用消化的质粒pARS 亚组C Ad5orf6-2 BsrGI/AscI共转化到BJ5183细胞中以获得pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) 前腺病毒(pre-adeno)质粒。

[0154] 2.2.2: hCMV-RG表达盒的插入 - pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG#1481的构建

通过利用hCMV启动子和BGH polyA序列之间存在的同源性,通过大肠杆菌中的同源重组将hCMV-RG盒克隆到线性化的前腺病毒受体载体中。如图6所示的质粒pvjTetOhCMV_RG_bghpolyA用SpeI、SphI和AsiSI切割以切除含有具有tetO、RG和BGHpolyA序列的hCMV启动子的2.58Kb片段。通过同源重组将得到的2.58Kb片段克隆到携带在HCMV和BGHpA控制下的RpsL-Kana选择盒的pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) 受体载体中。受体前腺病毒质粒用限制性内切核酸酶SnaBI线性化。得到的构建体是pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG载体(#1481) (图7)。

[0155] 2.2.3: hCMV-RG-WPRE表达盒的插入 - pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG-WPRE#1509的构建

通过利用pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG载体(#1481)的碱基2840-2939和3180-3279之间存在的同源性,通过在大肠杆菌中同源重组将WPRE序列克隆到前腺病毒受体载体中。通过PCR扩增1031bp DNA片段,并含有WPRE,BGHpolyA和对应于#1481 pAdeno载体的碱基2840-2939和3180-3279的重组臂。使用质粒pvjTetOhCMV_WPRE_BghPolyA (#1478)作为模板并利用以下寡核苷酸FW 5'-ggaaggtcagcgtgaccagccagtcgggcaaagtgatttcctcctgggagagctataaaagcggcgagagaccaggctgtgatgagcgccgcgatctgtaatcaacctctggattaca -3' (SEQ ID NO: 92)和RW 5'-ATGGCTCCGGCGGTCTCTGCAACACAAATAAAGAGACCCTAAGACCCCCAACTTATATATTTTCATGACCACCCCAGGCCACGCCCACTACCCACCTCACCATAGAGCCACCGCATCC-3' (SEQ ID NO: 93)进行PCR。通过同源重组将得到的1.03 Kb片段克隆到携带hCMV启动子和BGHpA控制下的RG转基因(SEQ ID NO:38)的pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG载体(#1481) 受体载体中。用限制性内切核酸酶AsiSI消化受体前腺病毒质粒。得到的构建体是pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV-RG-WPRE载体(#

1509),如图8所示。

[0156] 实施例3:ChAd155-RG载体产生

与Procell 92细胞系中的ChAd3和PanAd3相比,评估ChAd155的生产力。

[0157] 3.1:包含HIV Gag转基因的载体的产生

如上所述(ChAd155/GAG)或之前如对于ChAd3/GAG(Colloca等人, Sci. Transl. Med. (2012) 4:115ra2)所述制备表达HIV Gag蛋白的载体。在Procell 92中拯救并扩增ChAd3/GAG和ChAd155/GAG直至第3代(P3);使用P3裂解物感染两个T75烧瓶的用每个载体单层培养的Procell 92细胞。100vp/细胞的感染复数(MOI)用于两个感染实验。当完全细胞病变效应明显(感染后72小时)时收集感染的细胞并进行合并;通过三个冷冻/解冻循环(-70°/ 37°C)从感染细胞释放病毒,然后通过离心澄清裂解物。通过定量PCR(QPCR)分析用与CMV启动子区互补的引物和探针定量澄清的裂解物。寡核苷酸序列如下:CMVfor 5'-CATCTACGTATTAGTCATCGCTATTACCA-3'(SEQ ID NO: 23)、CMVrev 5'-GACTTGGAAATCCCCGTGAGT-3'(SEQ ID NO: 24)、CMVFAM-TAMRA探针5'-ACATCAATGGGCGTGGATAGCGGT-3'(SEQ ID NO: 25)(在ABI Prism 7900 Sequence检测仪-Applied Biosystem上运行QPCR)。在下表2中提供在澄清的裂解物上测量的所得体积滴度(vp/ml)和以病毒颗粒/细胞(vp/细胞)表示的细胞比生产力。

[0158] 表2. 来自P3裂解物的载体生产力

载体	vp/ml	总 vp (20 ml conc.)	vp/ 细胞
ChAd3/GAG	9.82E+09	1.96E+11	6.61E+03
ChAd155/GAG	1.11E+10	2.22E+11	7.46E+03

[0159] 为了证实表达HIV Gag转基因的ChAd155载体的较高生产力,使用经纯化的病毒作为接种物进行了第二个实验。为此目的,将Procell 92细胞接种在T25烧瓶中并在细胞汇合为约80%时,使用MOI=100 vp/细胞,用ChAd3/GAG和ChAd155/GAG进行感染。当完全细胞病变效应明显时收集感染的细胞;通过冷冻/解冻从感染的细胞中释放病毒并通过离心澄清。通过使用与CMV启动子区互补的以下引物和探针通过定量PCR分析定量澄清的裂解物:CMVfor 5'-CATCTACGTATTAGTCATCGCTATTACCA-3'(SEQ ID NO: 23)、CMV rev GACTTGGAAATCCCCGTGAGT (SEQ ID NO: 24)、CMV FAM-TAMRA探针5'-ACATCAATGGGCGTGGATAGCGGT-3'(SEQ ID NO: 25)(样品在ABI Prism 7900序列检测仪-Applied Biosystems上分析)。在表3中提供在澄清的裂解物上测量的所得体积滴度(vp/ml)和以病毒颗粒/细胞(vp/细胞)表示的细胞比生产力。

[0160] 表3 来自纯化的病毒的载体生产力

载体	vp/ml	总vp/T25烧瓶 (5ml裂解物)	vp/细胞
ChAd3/GAG	1.00E+10	5.00E+10	1.67E+04
ChAd155/GAG	1.21E+10	6.05E+10	2.02E+04

[0161] 3.2:包含RSV转基因的载体的产生

进行不同组的实验以评估在悬浮液中培养的Procell 92.S细胞中RSV疫苗载体的生产力。所述实验通过以 5×10^5 个细胞/ml的细胞密度感染Procell 92.S平行比较PanAd3/RSV (在W02012/089833中有所描述) 和ChAd155/RSV。感染后三天收集感染的细胞;通过三个冷冻/解冻循环从感染的细胞中释放病毒,并通过离心澄清裂解物。然后将澄清的裂解物通过如上报告的定量PCR分析定量。在表4中提供在澄清的裂解物上测量的所得体积滴度(vp/ml)和以病毒颗粒/细胞(vp/细胞)表示的细胞比生产力。

[0162] 表4纯化的病毒的载体生产力

病毒	(Vp/ml)	总vp	(vp/细胞)
PanAd3/RSV	5.82E+09	2.91E+11	1.16E+4
ChAd155/RSV	3.16E+10	1.58E+12	6.31E+04

[0163] 实施例4:转基因表达水平

4.1: HIV Gag转基因的表达水平

通过用包含HIV Gag转基因的ChAd3和ChAd155载体感染HeLa细胞,在平行实验中比较表达水平。将HeLa细胞接种在24孔板中,并使用MOI=250 vp/细胞用ChAd3/GAG和ChAd155/GAG纯化的病毒一式两份地感染。在感染后48小时收集HeLa感染的细胞的上清液,并使用市售的ELISA试剂盒(HIV-1 p24 ELISA试剂盒,PerkinElmer Life Science)定量分泌的HIV GAG蛋白的产生。使用HIV-1 p24抗原标准曲线,根据生产商的说明书就行定量。在图9中说明了以pg/ml Gag蛋白表示的结果。

[0164] 4.2: RSV F转基因的表达水平

通过用包含RSV F转基因的上述PanAd3和ChAd155载体感染HeLa细胞,在平行实验中比较表达水平。为此目的,将HeLa细胞接种在6孔板中,并使用MOI=500 vp/细胞用PanAd3/RSV和ChAd155/RSV纯化的病毒一式两份地感染。在感染后48小时收集上清液,并通过ELISA定量分泌的RSV F蛋白的产生。将上清液的5种不同稀释液转移至用市售的小鼠抗-RSV F单克隆抗体包被的微量培养板孔中。使用第二抗-RSV F兔抗血清,随后是生物素缀合的抗-兔IgG,然后加入链霉抗生物素蛋白-AP缀合物(BD Pharmingen目录554065)来揭示抗原捕获。使用RSV F蛋白(Sino Biological目录11049-V08B)标准曲线进行定量。在表5中提供得到的结果(表示为ug/ml的RSV F蛋白)。

[0165] 表5 RSV F转基因的表达水平

样品	μg/ml RSV F 蛋白
ChAd155/RSV	5.9
PanAd3/RSV	4

[0166] 还进行蛋白质印迹分析以证实由ChAd155 RSV载体提供的相对于PanAd3 RSV载体更高的转基因表达水平。使用MOI=250和500 vp/细胞,用PanAd3/RSV和ChAd155/RSV纯化的病毒感染铺板在6孔板中的HeLa细胞。收集HeLa感染细胞的上清液,并通过非还原性SDS凝胶电泳随后通过蛋白质印迹分析来分析分泌的RSV F蛋白的产生。将等量的上清液上样到

非还原性SDS凝胶上;电泳分离后,将蛋白转移到硝酸纤维素膜上,以用可在antibodies-online.com上获得(最后登录为2015年4月13日)的抗RSV F小鼠单克隆抗体(克隆RSV-F-3目录号:ABIN308230)进行探测。与第一抗体温育后,洗涤膜,然后与抗小鼠HRP缀合物第二抗体温育。最后,使用标准技术(ECL检测试剂Pierce目录号W3252282)通过电化学发光(ECL)将测定显色。在图10中显示了蛋白质印迹结果。通过针对F蛋白产生的单克隆抗体mAb 13揭示由箭头指示的约170kD的条带,其对应于三聚的F蛋白的预期分子量。可以看出,ChAd155 RSV载体在250和500 vp/细胞的MOI下产生比PanAd3RSV更深的条带。

[0167] 实施例5:通过小鼠免疫实验评估免疫效力

5.1: 包含HIV Gag转基因的载体的免疫原性

在BALB/c小鼠(5只/组)中,平行地评估ChAd155/GAG载体与ChAd3/GAG载体的免疫原性。通过肌肉注射 10^6 个病毒颗粒进行实验。在免疫接种后3周,使用在BALB/c小鼠中定位的GAG CD8⁺ T细胞表位,通过离体IFN- γ 酶联免疫斑点(ELISpot)测量T-细胞应答。结果显示于图11中,其表示为IFN- γ 点形成细胞/百万个脾细胞(SFC)。每个斑点代表在单一小鼠中的应答,且所述线对应于每个剂量组的平均值。在对两种载体的应答中,五只小鼠中有四只对CD8免疫显性肽呈阳性应答。

[0168] 5.2: 包含RSV转基因的载体的免疫原性

在BALB/c小鼠中评估了PanAd3/RSV和ChAd155/RSV载体的免疫效力。将两种载体以 10^8 、 10^7 和 3×10^6 vp的剂量进行肌肉注射。接种后3周,将免疫的小鼠的脾细胞分离,并使用在BALB/c小鼠中定位的免疫优势肽F和M表位作为抗原,通过IFN- γ -ELISpot进行分析。免疫应答的水平与递减剂量一致地下降(如预期),但是在用ChAd155/RSV载体免疫的小鼠组中的免疫应答明显高于用PanAd3/RSV疫苗免疫的等同小鼠组(图12)。符号显示各个小鼠数据,其表示为IFN- γ 斑点形成细胞(SFC)/百万个脾细胞,计算为对三种优势免疫表位(F₅₁₋₆₆, F₈₅₋₉₃和M2-1282-290)的反应总和并对背景进行校正。水平线代表每个剂量组的IFN- γ SFC/百万个脾细胞的平均数目。

[0169] 总之,上面报告的结果证实,ChAd155是与ChAd3和PanAd3载体相比改善的腺病毒载体。显示了ChAd155是生产力更高的,因此有助于制备过程,并显示能够在体外和在体内表达更高水平的转基因,提供对在动物模型中表达的抗原的更强T-细胞应答。

[0170] 实施例6:ChAd155-RG是免疫原性的并针对狂犬病攻击具有保护性

6.1: ChAd155-AG载体的免疫原性

在CD1小鼠中评估ChAd155-RG载体的免疫效力,并将结果显示在图13中。通过肌肉注射 10^9 vp进行实验。每个点代表单个小鼠的应答。图13证明单次施用编码狂犬病病毒G蛋白抗原的复制缺陷型腺病毒载体诱导了有效的免疫应答。载体诱导中和抗体的保护性水平(图13A)并诱导循环狂犬病特异性T细胞(图13B)。

[0171] 如Cliquet F. 等人, J. Immunol. Methods (1998) 212:79-87所述进行荧光抗体病毒中和测定(FAVN)。图13A证明在单次施用复制缺陷型ChAd155-RG后两周内在血清中检测功能性中和抗体。在施用后第二周检测到中和抗体其量远高于0.5 IU/ml的保护阈值水平(如世界卫生组织指导方针(虚线)中所述),且在第4周没有显示出下降的迹象。图13B证明在注射 10^{-9} pfu/ml ChAd155-RG的CD1小鼠的脾脏中检测到狂犬病特异性T细胞。如上所述对横跨狂犬病G蛋白序列的重叠肽进行的干扰素- γ ELISpot测定证明狂犬病特异性

T细胞的存在。

[0172] 在单次疫苗注射后跟踪抗体动力学至21周；滴度在第8周达到峰值，然后下降，但仍然远高于血清转变阈值（虚线），如图14所示。

[0173] 然后将ChAd155-RG的免疫效力与单剂量方案中的市售狂犬病疫苗进行比较，并将结果显示在图15中。左图显示了用估计的1/500的人剂量的ChAd155-RG (5×10^8 个病毒颗粒)或1/10的犬剂量的兽用狂犬病疫苗NOBIVAC免疫Balb/c小鼠的结果。右图显示了用1/1000的估计的人剂量的ChAd155-RG (10^8 个病毒颗粒)或1/10的人剂量的RABIPUR免疫CD1小鼠的结果。如上所述测量病毒中和抗体滴度，描述为IU/ML，并且在单剂量接种后两个月显示滴度。尽管商业疫苗存在大量过量，但ChAd155-RG载体诱导的免疫力证明优于商业上可获得的兽用和人狂犬病疫苗。

[0174] 6.2: ChAd155-AG载体针对狂犬病攻击进行保护的能力

图16证明单剂量的ChAd155-RG针对狂犬病攻击进行保护。使用ChAd155-RG、ChAd155对照载体或RABIPUR以表6中所示的剂量注射到4至6周龄的远交ICR小鼠的腓肠肌中。组1-6中的每一组由10只小鼠组成。在第0、7和21天给予小鼠三剂量的RABIPUR或以表6中所示的剂量给予单剂量的ChAd155或ChAd155-RG。

[0175] 表6单剂量的ChAd155-RG针对狂犬病攻击进行保护

组	载体	剂量	血清转变率	存活率
1	ChAd155对照	10^8 病毒颗粒	0%	60%
2	在第0、7和21天 RABIPUR	1/10的人剂量x 3	100%	100%
3	ChAd155-RG	10^8 病毒颗粒	100%	100%
4	ChAd155-RG	10^7 病毒颗粒	100%	100%
5	ChAd155-RG	10^6 病毒颗粒	90%	90%
6	ChAd155-RG	10^5 病毒颗粒	20%	60%

[0176] 然后用蝙蝠狂犬病病毒变体的人分离物攻击小鼠并随后进行90天。攻击病毒是从与暴露于狂犬病蝙蝠相关的致命人类病例中分离的街头 RABV变体Ps P4。在先前实验中先前未接种动物中计算的攻击剂量是100%致死的。在这项研究中，相同的剂量是60%致死。通过对狂犬病的快速荧光聚焦抑制试验(RFFIT)进行血清学检测来检测狂犬病特异性中和抗体，如Smith 等人 (1973) Bull. World Health Organ. 48:535-541所进行。此外，使用LIGHT DIAGNOSTICS狂犬病多克隆DFA试剂(Millipore Cat#5199)进行直接免疫荧光以检测脑组织中的病毒抗原。

[0177] 图16显示了每只个体小鼠的狂犬病特异性中和抗体的水平。给予不包含狂犬病抗原的阴性对照载体的组1中的小鼠不发生血清转变，并且组中的60%存活。给予组3-6中的小鼠渐降的ChAd155-RG的病毒颗粒载量。当给予 10^8 或 10^7 个病毒颗粒时，所有小鼠都发生血清转变并存活。注射 10^6 个病毒颗粒的小鼠具有90%的血清转变率并且90%存活。注射 10^5 个病毒颗粒的小鼠具有20%的血清转变率并且60%存活。这证明ChAd155-RG的单次肌肉接种在宽剂量范围内引起中和抗体滴度高于0.5IU/ml的阈值，并赋予针对致死性狂犬病攻击的保护。图16和表6因此证明单次施用重组ChAd155-RG可以至少与常规的、目前使用的灭活病毒疫苗一样有效地针对狂犬病进行保护。

[0178] 实施例7: ChAd155-RG在针对狂犬病攻击进行保护中比AdC68rab.gp更有效

将ChAd155-RG载体针对狂犬病病毒攻击进行保护的效力与如文献中报道的AdC68 rab.gp载体的效力进行比较。如表7中所示,用单剂量的ChAd155-RG肌内免疫Balb/c小鼠。攻击前病毒中和抗体水平是剂量依赖性的,并显示于图17A中。然后如实施例6中所述,用蝙蝠狂犬病病毒变体的人分离物攻击这些小鼠。如表7中所示,给予对照ChAd155载体的60%小鼠存活。用 10^8 或 10^7 vp ChAd155-RG免疫的Balb/c小鼠具有100%的存活率,用 10^6 vp ChAd155-RG免疫的小鼠具有90%的存活率,且用 10^6 vp ChAd155-RG免疫的小鼠具有60%的存活率,并且中和抗体滴度降至接近血清转变阈值。

[0179] 然后将这些结果与Zhou (2006) Mol. Ther. 14:662于670(图17B)公布在结果相关联。Zhou 等人报道用腺病毒重组病毒载体AdC68rab.gp肌内免疫ICR小鼠,然后用CVS-N2C狂犬病病毒鼻内攻击。对照动物未接种并且具有100%的死亡率。用 5×10^5 pfu AdC68rab.gp免疫的小鼠的45%发生血清转变并显示77%的存活率(17B左图),而用 5×10^4 pfu ChAd155-RG(17B右图)免疫的小鼠显示90%血清转变并且具有60%存活率。

[0180] 当使用1/50的剂量(AdC68rab.gp以 5×10^5 pfu 相比于ChAd155-RG以 10^4 pfu)时,在发明人的本研究 and Zhou等人报道的研究中获得类似的血清学和保护效力数据。因此,ChAd155-RG的效力约是AdC68rab.gp的50倍。

[0181] 表7 ChAd155-RG和AdC68rab.gp的效力

组	载体	剂量	血清转变率	存活率
1	ChAd155对照	10^8 个病毒颗粒	0%	60%
2	ChAd155-RG	10^8 个病毒颗粒	100%	100%
3	ChAd155-RG	10^7 个病毒颗粒	100%	100%
4	ChAd155-RG	10^6 个病毒颗粒	90%	90%
5	ChAd155-RG	10^5 个病毒颗粒	20%	60%
6	AdC68rab.gp	5×10^7 个病毒颗粒	44%	77%
7	AdC68rab.gp	5×10^6 个病毒颗粒	10%	60%

[0182] 实施例8:ChAd155-RG为非人灵长类动物提供长期免疫原性

为了评估非人灵长类动物中ChAd155-RG的免疫原性的动力学、广度和寿命,如下处理三组的五只食蟹猴(*Macaca fascicularis*)。用ChAd155-RG 5×10^{10} 病毒颗粒IM免疫第1组,然后在第48周用加强剂量的ChAd155-RG 5×10^{10} 病毒颗粒IM免疫。用ChAd155-RG 5×10^{10} 病毒颗粒IM免疫第2组,随后是在第24周加强剂量的RABIPUR接种,并在第48周IM加强剂量的ChAd155-RG 5×10^{10} 病毒颗粒。第3组接受肌内施用一半人剂量的RABIPUR,并在第7天和第21天接受相同的加强剂量。定期收集血清样品,并收集全血用于外周血单核细胞(PBMC)分析。

[0183] 将单剂量ChAd155-RG诱导的长达48周的免疫原性与RABIPUR的全程进行比较。引入在第24周时使用RABIPUR或是在第48周时使用ChAd155的加强以评估两种疫苗的相容性和加强中期至长期免疫应答的能力。

[0184] 将通过用ChAd155-RG单次免疫诱导的中和抗体滴度与用完全三剂量疗程的RABIPUR诱导的那些进行比较。图18显示了如通过FAVN测定法测定的用重组ChAd155-RG(第1组和第2组)免疫的猴子与用固定的细胞培养病毒疫苗RABIPUR(第3组)免疫长达接种后6个月的那些猴子的中和抗体应答的比较。单剂量的 10^{10} 个病毒颗粒ChAd155-RG诱导与三次剂量RABIPUR相同的免疫应答。

[0185] 这些结果显示单次施用ChAd155-RG能够引发远高于血清转变阈值的中和抗体滴度,其在至少48周内是稳定的并且与三次剂量在RABIPUR相当。由ChAd155-RG诱导的血清转变是快速的。用ChAd155-RG免疫的所有动物在免疫后两周超过阈值,此时用RABIPUR免疫的动物早已经接受第二剂量。

[0186] 在第24周用RABIPUR加强第2组中的动物(图18-正方形)在将病毒中和抗体提高到远高于第0天施用ChAd155-RG后达到的峰值水平方面非常有效。这证明RABIPUR病毒裂解物抗原完全能够加强由ChAd155-RG核酸编码的抗原诱导的免疫力。

[0187] 使用ChAd155-RG加强第1组(图18-三角形)和第2组(图18-倒三角形)中的动物也非常有效。无论它们是否用RABIPUR进行中间加强,用ChAd155-RG免疫的动物在第48周对ChAd155-RG加强产生(mounted)强烈的免疫应答。这证明可以有效地再施用ChAd155-RG核酸编码的抗原。它还证明ChAd155-RG核酸编码的抗原在施用RABIPUR病毒裂解物抗原后有效加强免疫应答。总之,图18证明了编码狂犬病G抗原的猴腺病毒ChAd155与包含病毒裂解物抗原的常规狂犬病疫苗的相容性。

[0188] 实施例9: ChAd155-RG在非人灵长类动物中诱导细胞免疫应答

除了实施例8中证明的体液抗体应答外,ChAd155-RG还诱导强烈的细胞免疫应答。图19显示单剂量的ChAd155-RG在接种动物的外周血中诱导持续水平的狂犬病糖蛋白特异性IFN γ 分泌T细胞,如通过IFN γ -ELIspot测定所检测的。相反,细胞免疫应答低于用RABIPUR接种的动物的检测限。

[0189] 如实施例8中所述,用ChAd155-RG免疫并在第48周用ChAd155-RG加强的第1组中的动物,证明加强再次放大IFN γ 水平。用ChAd155-RG免疫并在第24周用RABIPUR首先加强,然后在第48周用ChAd155-RG加强的第2组中的动物,没有显示出响应于RABIPUR加强的IFN γ 水平的增加,但显示出对ChAd155-RG加强的强烈应答。这证明ChAd155-RG加强可以扩增成熟动物中的记忆T细胞。在整个随访过程中未检测到白介素4应答。

[0190] 实施例10: 人中安全性的剂量递增研究

为了评估ChAd155-RG在人中的安全性,将启动I期研究。主体将是正常健康的成年男性和女性,其没有狂犬病疫苗接种史,未暴露于狂犬病动物或接受基于腺病毒的研究性疫苗。研究规模将足够大以确定主要研究终点安全性的结果。将进行标准统计分析,包括95%置信区间。

[0191] 主体将接受一次或多次肌肉注射ChAd155-RG;RABIPUR将用作比较物。将施用低剂量的ChAd155-RG疫苗,并且在数据审查和批准之后,然后将增加剂量。主体将在施用后跟踪全身性和局部不良事件,包括但不限于发烧、头痛、恶心、呕吐、不适和肌痛;以及注射部位处的疼痛、压痛、硬结、发红或肿胀。将检查血液参数并记录任何其他主动提供的症状。

[0192] 该研究还可以通过评估疫苗相关的免疫应答来评价免疫原性。结果测量可包括但不限于血清中和抗体的水平、循环B细胞分泌的抗体的定量和针对狂犬病病毒抗原的T细胞应答的定量。

序列表

<110> GLAXOSMITHKLINE BIOLOGICALS S. A.

<120> 具有狂犬病病毒抗原的黑猩猩腺病毒构建体

<130> VU66242 W0

<140>

<141>

<150> 62/465, 378

<151> 2017-03-01

<150> 62/432, 033

<151> 2016-12-09

<160> 94

<170> PatentIn version 3.5

<210> 1

<211> 578

<212> PRT

<213> 黑猩猩腺病毒

<400> 1

[0001] Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125

Thr 130	Met	Gln	Ser	Glu	Ala	Pro 135	Leu	Thr	Val	Gln	Asp 140	Ala	Lys	Leu	Thr
Leu 145	Ala	Thr	Lys	Gly	Pro 150	Leu	Thr	Val	Ser	Glu 155	Gly	Lys	Leu	Ala	Leu 160
Gln	Thr	Ser	Ala	Pro 165	Leu	Thr	Ala	Ala	Asp 170	Ser	Ser	Thr	Leu	Thr	Val 175
Ser	Ala	Thr	Pro 180	Pro	Leu	Ser	Thr	Ser	Asn 185	Gly	Ser	Leu	Gly	Ile	Asp 190
Met	Gln	Ala	Pro	Ile	Tyr	Thr	Thr 200	Asn	Gly	Lys	Leu	Gly 205	Leu	Asn	Phe
Gly 210	Ala	Pro	Leu	His	Val	Val 215	Asp	Ser	Leu	Asn	Ala 220	Leu	Thr	Val	Val
Thr 225	Gly	Gln	Gly	Leu	Thr 230	Ile	Asn	Gly	Thr	Ala 235	Leu	Gln	Thr	Arg	Val 240
[0002]															
Ser	Gly	Ala	Leu	Asn 245	Tyr	Asp	Thr	Ser	Gly 250	Asn	Leu	Glu	Leu	Arg 255	Ala
Ala	Gly	Gly	Met 260	Arg	Val	Asp	Ala	Asn 265	Gly	Gln	Leu	Ile	Leu	Asp	Val
Ala	Tyr	Pro 275	Phe	Asp	Ala	Gln	Asn 280	Asn	Leu	Ser	Leu	Arg 285	Leu	Gly	Gln
Gly 290	Pro	Leu	Phe	Val	Asn	Ser 295	Ala	His	Asn	Leu	Asp 300	Val	Asn	Tyr	Asn
Arg 305	Gly	Leu	Tyr	Leu	Phe 310	Thr	Ser	Gly	Asn	Thr 315	Lys	Lys	Leu	Glu	Val 320
Asn	Ile	Lys	Thr	Ala 325	Lys	Gly	Leu	Ile	Tyr 330	Asp	Asp	Thr	Ala	Ile	Ala 335
Ile	Asn	Ala	Gly 340	Asp	Gly	Leu	Gln	Phe 345	Asp	Ser	Gly	Ser	Asp 350	Thr	Asn

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

[0003]

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575

[0004]

Gln Glu

<210> 2

<211> 1734

<212> DNA

<213> 黑猩猩腺病毒

<400> 2

atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtacccta tgacacggaa	60
agcggccctc cctccgtccc ttctctcacc cctcccttcg tgtctccga tggattccaa	120
gaaagtcccc ccggggtcct gtctctgaac ctggccgagc ccctgggtcac ttcccacggc	180
atgctcgccc tgaaaatggg aagtggcctc tccttgagc acgttggtcaa cctcacctct	240
caagatatca ccaccgctag ccctccctc aaaaaacca agaccaacct cagcctagaa	300
acctcatccc ccctaactgt gagcacctca ggccgacctc ccgtagcagc cgccgctccc	360
ctggcggtgg ccggcacctc cctcacctat caatcagagg cccccctgac agtacaggat	420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg	480
caaacatcgg ccccgctgac ggccgctgac agcagcacc tcacagtcag tgccacacca	540
ccccttagca caagcaatgg cagcttggtt attgacatgc aagcccccat ttacaccacc	600
aatggaaaac taggacttaa ctttgccgct cccctgcatg tggtagacag cctaaatgca	660
ctgactgtag ttactggcca aggtcttacg ataaacggaa cagccctaca aactagagtc	720
tcaggtgccc tcaactatga cacatcagga aacctagaat tgagagctgc agggggtatg	780
cgagttgatg caaatggtca acttatcctt gatgtagctt acccatttga tgcacaaaac	840
aatctcagcc ttaggcttgg acagggaccc ctgtttgtta actctgccc caacttgat	900
gttaactaca acagaggcct ctacctgttc acatctggaa ataccaaaaa gctagaagtt	960
aatatcaaaa cagccaaggg tctcatttat gatgacactg ctatagcaat caatgcgggt	1020
gatgggctac agtttgactc aggtcagat acaaatccat taaaaactaa acttgatta	1080
ggactggatt atgactccag cagagccata attgctaaac tgggaactgg cctaagcttt	1140
gacaacacag gtgcatcac agtaggcaac aaaaatgatg acaagcttac cttgtggacc	1200
acaccagacc catcccctaa ctgtagaatc tattcagaga aagatgctaa attcacactt	1260
gttttgacta aatgcggcag tcaggtgttg gccagcgttt ctgttttata tgtaaaaggt	1320
agccttgccg ccatcagtgg cacagtaact agtgctcaga ttgtcctcag atttgatgaa	1380
aatggagttc tactaagcaa ttcttccctt gacctcaat actggaacta cagaaaaggt	1440

[0005]

gaccttacag agggcactgc atataccaac gcagtgggat ttatgcccac cctcacagca 1500
 tacccaaaaa cacagagcca aactgctaaa agcaacattg taagtcaggt ttacttgaat 1560
 ggggacaaat ccaaaccat gaccctcacc attaccctca atggaactaa tgaaacagga 1620
 gatgccacag taagcactta ctccatgtca ttctcatgga actggaatgg aagtaattac 1680
 attaataaaa cgttccaaac caactccttc accttctcct acatcgccca agaa 1734

<210> 3

<211> 593

<212> PRT

<213> 黑猩猩腺病毒

<400> 3

Met Arg Arg Ala Ala Met Tyr Gln Glu Gly Pro Pro Pro Ser Tyr Glu
 1 5 10 15

Ser Val Val Gly Ala Ala Ala Ala Ala Pro Ser Ser Pro Phe Ala Ser
 20 25 30

Gln Leu Leu Glu Pro Pro Tyr Val Pro Pro Arg Tyr Leu Arg Pro Thr
 35 40 45

Gly Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Phe Asp
 50 55 60

Thr Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Val Ala Ser
 65 70 75 80

Leu Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Ile Gln
 85 90 95

Asn Asn Asp Tyr Ser Pro Ser Glu Ala Ser Thr Gln Thr Ile Asn Leu
 100 105 110

Asp Asp Arg Ser His Trp Gly Gly Asp Leu Lys Thr Ile Leu His Thr
 115 120 125

Asn Met Pro Asn Val Asn Glu Phe Met Phe Thr Asn Lys Phe Lys Ala
 130 135 140

Arg Val Met Val Ser Arg Ser His Thr Lys Glu Asp Arg Val Glu Leu
 145 150 155 160

Lys Tyr Glu Trp Val Glu Phe Glu Leu Pro Glu Gly Asn Tyr Ser Glu
165 170 175

Thr Met Thr Ile Asp Leu Met Asn Asn Ala Ile Val Glu His Tyr Leu
180 185 190

Lys Val Gly Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys
195 200 205

Phe Asp Thr Arg Asn Phe Arg Leu Gly Leu Asp Pro Val Thr Gly Leu
210 215 220

Val Met Pro Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Ile
225 230 235 240

Leu Leu Pro Gly Cys Gly Val Asp Phe Thr Tyr Ser Arg Leu Ser Asn
245 250 255

Leu Leu Gly Ile Arg Lys Arg Gln Pro Phe Gln Glu Gly Phe Arg Ile
260 265 270

[0006]

Thr Tyr Glu Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val
275 280 285

Glu Ala Tyr Gln Asp Ser Leu Lys Glu Asn Glu Ala Gly Gln Glu Asp
290 295 300

Thr Ala Pro Ala Ala Ser Ala Ala Ala Glu Gln Gly Glu Asp Ala Ala
305 310 315 320

Asp Thr Ala Ala Ala Asp Gly Ala Glu Ala Asp Pro Ala Met Val Val
325 330 335

Glu Ala Pro Glu Gln Glu Glu Asp Met Asn Asp Ser Ala Val Arg Gly
340 345 350

Asp Thr Phe Val Thr Arg Gly Glu Glu Lys Gln Ala Glu Ala Glu Ala
355 360 365

Ala Ala Glu Glu Lys Gln Leu Ala Ala Ala Ala Ala Ala Ala Ala Leu
370 375 380

[0008]

<211> 1779

<212> DNA

<213> 黑猩猩腺病毒

<400> 4

atgcggcgcg cggcgatgta ccaggaggga cctcctccct cttacgagag cgtggtgggc	60
gcggcgggcg cggcgccctc ttctcccttt gcgtcgcagc tgctggagcc gccgtacgtg	120
cctccgcgct acctgcggcc tacggggggg agaaacagca tccgttactc ggagctggcg	180
cccctgttcg acaccacccg ggtgtacctg gtggacaaca agtcggcgga cgtggcctcc	240
ctgaactacc agaacgacca cagcaatfff ttgaccacgg tcatccagaa caatgactac	300
agcccgagcg aggccagcac ccagaccatc aatctggatg accggtcgca ctggggcggc	360
gacctgaaaa ccatcctgca caccaacatg cccaacgtga acgagttcat gttcaccaat	420
aagttcaagg cgcggtgat ggtgtcgcgc tcgcacacca aggaagaccg ggtggagctg	480
aagtacgagt ggggtggagt cgagctgcca gagggcaact actccgagac catgaccatt	540
gacctgatga acaacgcgat cgtggagcac tatctgaaag tgggcaggca gaacggggtc	600
ctggagagcg acatcggggt caagttcgac accaggaact tccgcctggg gctggacccc	660
gtgaccgggc tggttatgcc cggggtgtac accaacgagg ccttccatcc cgacatcatc	720
ctgctgcccc gctgcgggggt ggacttact tacagccgcc tgagcaacct cctgggcatc	780
cgcaagcggc agcccttcca ggagggttc aggatcacct acgaggacct ggaggggggc	840
aacatccccg cgctcctcga tgtggaggcc taccaggata gcttgaagga aaatgaggcg	900
ggacaggagg ataccgcccc cgccgcctcc gccgccgccc agcagggcga ggatgctgct	960
gacaccgccc ccgcggacgg ggcagaggcc gaccccgcta tgggtggtgga ggctcccag	1020
caggaggagg acatgaatga cagtgcggtg cgcgagaca ccttcgtcac ccggggggag	1080
gaaaagcaag cggaggccga ggccgcggcc gaggaaaagc aactggcggc agcagcggcg	1140
gcggcgcgct tggccgcggc ggaggctgag tctgagggga ccaagcccgc caaggagccc	1200
gtgattaagc ccctgaccga agatagcaag aagcgcagtt acaacctgct caaggacagc	1260
accaacaccg cgtaccgcag ctggtacctg gcctacaact acggcgaccc gtcgacgggg	1320
gtgcgctcct ggaccctgct gtgcacgccg gacgtgacct gcggctcgga gcaggtgtac	1380
tggtegtgc ccgacatgat gcaagacccc gtgaccttcc gctccacgcg gcaggtcagc	1440
aacttcccgg tgggtgggcg cgagctgctg cccgtgcaact ccaagagctt ctacaacgac	1500
caggccgtct actcccagct catccgccag ttcacctctc tgaccacagt gttcaatcgc	1560

[0009]

```

tttcctgaga accagattct ggcgcgcccg cccgccccca ccatcaccac cgtcagtga 1620
aacgttcctg ctctcacaga tcacgggacg ctaccgctgc gcaacagcat cggaggagtc 1680
cagcgagtga cgtttactga cgccagacgc cgcacctgcc cctacgttta caaggccttg 1740
ggcatagtct cgccgcgcgt cctttccage cgcactttt 1779

```

<210> 5

<211> 964

<212> PRT

<213> 黑猩猩腺病毒

<400> 5

```

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
1          5          10          15

```

```

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
          20          25          30

```

```

Arg Ala Thr Asp Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
          35          40          45

```

```

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
          50          55          60

```

```

Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65          70          75          80

```

```

Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
          85          90          95

```

```

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
          100          105          110

```

```

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
          115          120          125

```

```

Ala Pro Asn Ser Cys Glu Trp Glu Gln Glu Glu Thr Gln Thr Ala Glu
          130          135          140

```

```

Glu Ala Gln Asp Glu Glu Glu Asp Glu Ala Glu Ala Glu Glu Glu Met
145          150          155          160

```

```

Pro Gln Glu Glu Gln Ala Pro Val Lys Lys Thr His Val Tyr Ala Gln
          165          170          175

```

Ala Pro Leu Ser Gly Glu Lys Ile Thr Lys Asp Gly Leu Gln Ile Gly
180 185 190

Thr Asp Ala Thr Ala Thr Glu Gln Lys Pro Ile Tyr Ala Asp Pro Thr
195 200 205

Phe Gln Pro Glu Pro Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp
210 215 220

Ala Ser Val Ala Gly Gly Arg Val Leu Lys Lys Thr Thr Pro Met Lys
225 230 235 240

Pro Cys Tyr Gly Ser Tyr Ala Arg Pro Thr Asn Ala Asn Gly Gly Gln
245 250 255

Gly Val Leu Val Glu Lys Asp Gly Gly Lys Met Glu Ser Gln Val Asp
260 265 270

Met Gln Phe Phe Ser Thr Ser Glu Asn Ala Arg Asn Glu Ala Asn Asn
275 280 285

[0010]

Ile Gln Pro Lys Leu Val Leu Tyr Ser Glu Asp Val His Met Glu Thr
290 295 300

Pro Asp Thr His Ile Ser Tyr Lys Pro Ala Lys Ser Asp Asp Asn Ser
305 310 315 320

Lys Val Met Leu Gly Gln Gln Ser Met Pro Asn Arg Pro Asn Tyr Ile
325 330 335

Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly
340 345 350

Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val
355 360 365

Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Leu Leu Asp
370 375 380

Ser Met Gly Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val
385 390 395 400

Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu
405 410 415

Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr
420 425 430

Asp Thr Tyr Gln Ala Ile Lys Thr Asn Gly Asn Gly Asn Gly Gly Gly
435 440 445

Asn Thr Thr Trp Thr Lys Asp Glu Thr Phe Ala Asp Arg Asn Glu Ile
450 455 460

Gly Val Gly Asn Asn Phe Ala Met Glu Ile Asn Leu Ser Ala Asn Leu
465 470 475 480

Trp Arg Asn Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys
485 490 495

Leu Lys Tyr Asn Pro Ser Asn Val Glu Ile Ser Asp Asn Pro Asn Thr
500 505 510

[0011]

Tyr Asp Tyr Met Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys
515 520 525

Tyr Ile Asn Leu Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val
530 535 540

Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met
545 550 555 560

Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln
565 570 575

Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr
580 585 590

Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser
595 600 605

Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser
610 615 620

	Ile	Cys	Leu	Tyr	Ala	Thr	Phe	Phe	Pro	Met	Ala	His	Asn	Thr	Ala	Ser
	625					630					635					640
	Thr	Leu	Glu	Ala	Met	Leu	Arg	Asn	Asp	Thr	Asn	Asp	Gln	Ser	Phe	Asn
					645					650					655	
	Asp	Tyr	Leu	Ser	Ala	Ala	Asn	Met	Leu	Tyr	Pro	Ile	Pro	Ala	Asn	Ala
				660					665					670		
	Thr	Asn	Val	Pro	Ile	Ser	Ile	Pro	Ser	Arg	Asn	Trp	Ala	Ala	Phe	Arg
			675					680					685			
	Gly	Trp	Ala	Phe	Thr	Arg	Leu	Lys	Thr	Lys	Glu	Thr	Pro	Ser	Leu	Gly
	690						695					700				
	Ser	Gly	Phe	Asp	Pro	Tyr	Tyr	Thr	Tyr	Ser	Gly	Ser	Ile	Pro	Tyr	Leu
	705					710					715					720
	Asp	Gly	Thr	Phe	Tyr	Leu	Asn	His	Thr	Phe	Lys	Lys	Val	Ser	Val	Thr
					725					730					735	
[0012]	Phe	Asp	Ser	Ser	Val	Ser	Trp	Pro	Gly	Asn	Asp	Arg	Leu	Leu	Thr	Pro
				740					745					750		
	Asn	Glu	Phe	Glu	Ile	Lys	Arg	Ser	Val	Asp	Gly	Glu	Gly	Tyr	Asn	Val
			755					760					765			
	Ala	Gln	Cys	Asn	Met	Thr	Lys	Asp	Trp	Phe	Leu	Ile	Gln	Met	Leu	Ala
	770						775					780				
	Asn	Tyr	Asn	Ile	Gly	Tyr	Gln	Gly	Phe	Tyr	Ile	Pro	Glu	Ser	Tyr	Lys
	785					790					795					800
	Asp	Arg	Met	Tyr	Ser	Phe	Phe	Arg	Asn	Phe	Gln	Pro	Met	Ser	Arg	Gln
					805					810					815	
	Val	Val	Asp	Glu	Thr	Lys	Tyr	Lys	Asp	Tyr	Gln	Gln	Val	Gly	Ile	Ile
				820					825					830		
	His	Gln	His	Asn	Asn	Ser	Gly	Phe	Val	Gly	Tyr	Leu	Ala	Pro	Thr	Met
			835					840					845			
	Arg	Glu	Gly	Gln	Ala	Tyr	Pro	Ala	Asn	Phe	Pro	Tyr	Pro	Leu	Ile	Gly

[0013]

850	855	860
Lys Thr Ala Val Asp Ser Val Thr Gln Lys Lys Phe Leu Cys Asp Arg		
865	870	875 880
Thr Leu Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly Ala		
	885	890 895
Leu Thr Asp Leu Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala		
	900	905 910
Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu		
	915	920 925
Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Val His Gln Pro His		
	930	935 940
Arg Gly Val Ile Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly		
945	950	955 960
Asn Ala Thr Thr		

<210> 6

<211> 2880

<212> DNA

<213> 黑猩猩腺病毒

<400> 6

atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc	60
tcggagtacc tgagccccgg gctggtgcag ttcgcccgcg ccaccgagag ctacttcagc	120
ctgagtaaca agtttaggaa cccacgggtg gcgccacgc acgatgtgac caccgaccgg	180
tctcagcgcc tgacgtgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac	240
aaggcgcggt tcaccctggc cgtggggcgac aaccgcgtgc tggacatggc ctccacctac	300
tttgacatcc gcgggggtgct ggaccgggggt ccactttca agccctactc tggcaccgcc	360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact	420
caggcagttg aagaagcagc agaagaggaa gaagaagatg ctgacgggtca agctgaggaa	480
gagcaagcag ctacaaaaa gactcatgta tatgctcagg ctcccctttc tggcgaaaaa	540
attagtaaag atggtctgca aataggaacg gacgctacag ctacagaaca aaaacctatt	600

[0014]

tatgcagacc ctacattcca gcccgaaccc caaatcgggg agtcccagtg gaatgaggca	660
gatgctacag tcgccggcgg tagagtgcta aagaaatcta ctcccatgaa accatgctat	720
ggttcctatg caagaccac aaatgctaata ggaggtcagg gtgtactaac ggcaaatgcc	780
caggacagc tagaatctca ggttgaaatg caattctttt caacttctga aaacgcccgt	840
aacgaggcta acaacattca gcccgaattg gtgctgtata gtgaggatgt gcacatggag	900
accccgata cgcacctttc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg	960
ctgggtcagc agtccatgcc caacagacct aattacatcg gcttcagaga caactttatc	1020
ggcctcatgt attacaatag cactggcaac atgggagtcg ttgcaggcca ggcctctcag	1080
ttgaatgcag tgggtgactt gcaagacaga aacacagaac tgtcctacca gctcttgcct	1140
gattccatgg gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagttat	1200
gaccagatg ttagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt	1260
ttccctctgg gtggcatagg ggtaactgac acttaccagg ctgttaaaac caacaatggc	1320
aataacgggg gccaggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata	1380
ggggtgggaa acaatttcgc tatggagatc aacctcagtg ccaacctgtg gagaaacttc	1440
ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaatgtg	1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggg ggccccgggg	1560
ctgggtgact gctacatcaa cctgggcgcg cgctggctcg ttgactacat ggacaacgtc	1620
aacccttca accaccaccg caatgcgggc ctgcgctacc gctccatgct cctgggcaac	1680
gggcgctacg tgcccttcca catccaggtg cccagaagt tctttgccat caagaacctc	1740
ctcctctgc cgggctccta cacctacgag tggaacttca ggaaggatgt caacatggtc	1800
ctccagagct ctctgggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc	1860
atctgcctct acgccacctt ctccccatg gccacaaca cggcctccac gctcgaggcc	1920
atgctcagga acgacaccaa cgaccagtc ttcaatgact acctctccgc cgccaacatg	1980
ctctaccca taccgccaac cgccaccaac gtccccatct ccatccctc gcgcaactgg	2040
gcggccttcc gcggtgggc cttcaccgc ctcaagacca aggagacccc ctccctgggc	2100
tcgggattcg acccctaata cacctactcg ggctccattc cctacctgga cggcaccttc	2160
tacctcaacc aactttcaa gaaggctctg gtcacctcg actcctcggt cagctggccg	2220
ggcaacgacc gtctgtcac cccaacgag ttcgagatca agcgtcgggt cgacggggag	2280
ggctacaacg tggcccagtg caacatgacc aaggactggg tcctgggtcca gatgctggcc	2340

[0015]

aactacaaca tcggctacca gggtttctac atcccagaga gctacaagga caggatgtac	2400
tcctttcttca ggaactttcca gcccattgagc cggcaggtgg tggaccagac caagtacaag	2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggtttcgt gggctacctc	2520
gccccacca tgcgcgaggg acaggcctac cccgccaaact tcccctatcc gctcataggc	2580
aagaccgcgg tcgacagcat caccagaaa aagttcctct gcgaccgcac cctctggcgc	2640
atcccccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg	2700
ctctacgcca actccgcca cgccctcgac atgaccttcg aggtcgaccc catggacgag	2760
cccacccttc tctatgttct gttcgaagtc tttagcgtgg tccgggtcca ccagccgcac	2820
cgcggcgctca tcgagaccgt gtacctgcgt acgcccttct cggccggcaa cgccaccacc	2880

<210> 7

<211> 37912

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 7

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcgggcggg gaggcgggtt tggggcgggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggtt gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttcccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acatttatat tggtcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccac	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac	720
tttccattga cgtcaatggg tggagtattt acggtaaaact gccacttgg cagtacatca	780
agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccccctg	840

[0016]

gcattatgcc cagtacatga ccttatggga ctttctact tggcagtaca tctacgtatt	900
agtcacgcgt attacatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca ccccatigac gtcaatggga gtttgttttg	1020
gcacaaaaat caacgggact ttcaaaaatg tcgtaacaac tccgccccat tgacgcaa	1080
gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctcccta tcagtgatag	1140
agatctccct atcagtgata gagatcgteg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacggtgca ttggaacgcg gattccccgt gccaaagagt agatcttccg	1320
tttatctagg taccgggccc cccctcgagg tcgacggtat cgataagctt cacgctgccg	1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgtagaaag ccagtccgca	1440
gaaacggtgc tgaccccgga tgaatgtcag ctactgggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagcttgcag tgggcttaca tggcgatagc tagactgggc	1560
ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgccctctg gtaaggttgg	1620
gaagccctgc aaagtaaact ggatggcttt cttgccgcca aggatctgat ggcgcagggg	1680
atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac	1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgctgga agcatgcccg caaaaacgtg	1800
gcgtatgtac tcgtgtatat actaccactc ctaaaaaacc gaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggt ggtgaaggtc	1920
acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgttaaa gacctcccgg	1980
gtgttcgtta ccacaccgta cgtggtgcgc ttgactgctc cggcgttaaa gaccgtaagc	2040
aggctcgttc caagtatggc gtgaagcgtc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcgtttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggctatgac tgggcacaa agacaatcgg ctgctctgat	2220
gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg	2280
tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tategtggct ggccacgacg	2340
ggcgttccct gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcattctacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac ctgcccatc	2520

[0017]

gaccaccaag cgaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	2580
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgctgtga cccatggcga tgcctgcttg	2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc	2880
atcgctttct atcgctttct tgacgagttc ttctgagcgg gactctgggg ttcgaaatga	2940
ccgaccaagc gacgccaac ctgccatcac gagatttcga ttccaccgcc gccttctatg	3000
aaaggttggg cttcggaatc gttttccggg acgccggctg gatgacctc cagcgcgggg	3060
atctcatgct ggagtcttc gccaccccg ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgat ctgctgtgcc ttctagtgc cagccatctg ttgtttgcc	3180
ctccccctg ctttcttga ccttgaagg tgccactccc actgtccttt cctaataaaa	3240
tgaggaaatt gcatgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg	3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg	3360
ctctatggcc gatcagcgat cgctgaggtg ggtgagtggg cgtggcctgg ggtggtcatg	3420
aaaatatata agttgggggt cttaggtct ctttatttgt gttgcagaga ccgccggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgcc ttggatggca gcatcgtgag	3540
cccttatttg acgacgcgga tgccccactg ggccgggggtg cgtcagaatg tgatgggctc	3600
cagcatcgac ggccgaccg tctgccccgc aaattccgcc acgctgacct atgcgacct	3660
cgcggggacg ccgttggacg ccaccgccgc cgccgccgcc accgcagccg cctcgccct	3720
gcgcagcctg gccacggact ttgcattcct gggaccactg gcgacagggg ctacttctcg	3780
ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggtctc	3900
ctccctgcaa gctggcgga atgcttctcc cacaaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttattt cataatttc	4020
cgcgcgcgat aggccctaga ccagcgttct cggtcgttga gggcgcggtg tatcttctcc	4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccgggg	4140
tggaggtagc accactgcag agcttcatgc tccggggtg tgtttagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260

[0018]

gggaggccct tgggtgtaagt gtttacaaaa cggttaagtt gggaagggtg cattcgggga	4320
gagatgatgt gcatcttga ctgtatTTTT agattggcga tgtttccgcc cagatccctt	4380
ctgggattca tgttgtcag gaccaccagt acagtgtatc cggtgacctt ggggaatttg	4440
tcatgcagct tagagggaag agcgtggaag aacttggaga cgcctttgtg gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag	4560
atatttctgg ggtcgtgac gtcgtagtgt tgttccagg tgaggtcgtc ataggccatt	4620
tttacaaagc gcgggaggag ggtgcccgc tgggggatga tgggtccctc tggccctggg	4680
gcgtagtgtc cctcgcagat ctgcatttcc caggccttaa tctcggagg gggaaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag	4800
agcaggtttc taagcagctg tgattttcca caaccggtgg gcccataaat aacacctata	4860
accggttgca gctggtagt tagagagctg cagctgccgt cgtcccggag gaggggggcc	4920
acctcgttga gcatgtccct gacgcgatg ttctccccga ccagatccgc cagaaggcgc	4980
tcgccgcca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg	5040
tccgccgtgg gcatgttttt cagggtctgg ctacagcagc ccaggcggtc ccagagctcg	5100
gtgacgtgct ctacggcatc tctatccagc atatctctc gtttcgcggg ttggggcgac	5160
tttcgctgta gggcaccaag cgggtggtgt ccagcggggc cagagtcatt tccttccatg	5220
ggcgcagggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgt ccgggctgag	5280
cgttgccaa ggtgcgcttg aggctggttc tgctggtgct gaagcgtgc cggctcttcgc	5340
cctgcgcgtc ggccaggtag catttgacca tgggtgcata gtccagcccc tccgcggcgt	5400
gtcccttggc gcgcagcttg ccttggagg tggcgccgca cgaggggcag agcaggctct	5460
tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgccgc	5520
agaccccgca cacggtctcg cactccacca gccaggtag ctcggggcgc gccgggtcaa	5580
aaaccagggt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgaggtggt	5640
gtccccgctc ggtgacgaag aggctgtccg tgtctccgta gaccgacttg aggggtcttt	5700
tctccagggg ggtccctcgg tcttctcgt agaggaactc ggaccactct gagacgaagg	5760
cccgcgtcca ggccaggacg aaggaggcta tgtgggagg gtagcggtcg ttgtccacta	5820
gggggtccac cttctccaag gtgtgaagac acatgtcgcc ttctcggcg tccaggaagg	5880
tgattggctt gtaggtgtag gccacgtgac cgggggttcc tgacgggggg gtataaaagg	5940

[0019]

gggtgggggc gcgctcgtcg tcactctctt ccgcatcgct gtctgcgagg gccagctgct	6000
ggggtgagta ttccctctcg aaggcgggca tgacctccgc gctgaggttg tcagtttcca	6060
aaaacgagga ggatttgatg ttcacctgtc ccgaggtgat acctttgagg gtaccgcgct	6120
ccatctggtc agaaaacacg atctttttat tgtccagctt ggtggcgaac gaccgtaga	6180
gggcgttgga gagcagcttg gcgatggagc gcagggtctg gttcttgtcc ctgtcggcgc	6240
gctccttggc cgcgatgttg agctgcacgt actcgcgcgc gacgcagcgc cactcgggga	6300
agacggtggt gcgctcgtcg ggaccaggc gcacgcgcca gccgcggttg tgcagggtga	6360
ccaggtecac gctggtggcg acctcgccgc gcaggcgtc gttggtccag cagagacggc	6420
cgcccttgcg cgagcagaag gggggcaggg ggtcgagctg ggtctcgtcc ggggggtccg	6480
cgtccacggt gaaaaccccg gggcgcaggc gcgcgtcgaa gtagtctatc ttgcaacctt	6540
gcatgtccag cgcctgtgc cagtcgcggg cggcgagcgc gcgctcgtag gggttgagcg	6600
gcgggccccca gggcatgggg tgggtgagtg cggaggcgta catgccgcag atgtcataga	6660
cgtagagggg ctcccgagg accccgatgt aggtggggta gcagcgccg ccgcggatgc	6720
tggcgcgcac gtagtcatac agctcgtgcg agggggcgag gaggtcgggg ccaggttg	6780
tgcgggcggg gcgctccgcg cggaagacga tctgcctgaa gatggcatgc gagttggaag	6840
agatggtggg gcgctggaag acgttgaagc tggcgtcctg caggccgacg gcgtcgcgca	6900
cgaaggaggc gtaggagtcg cgcagcttgt gtaccagtc ggcggtgacc tgcacgtcga	6960
gcgcgcagta gtcgagggtc tcgcggatga tgtcatattt agcctgcccc ttctttttcc	7020
acagctcgcg gttgaggaca aactcttcgc ggtctttcca gtactcttgg atcgggaaac	7080
cgtccggttc cgaacggtaa gagcctagca tgtagaactg gttgacggcc tggtaggcgc	7140
agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg	7200
tcagggcgaa ggtgtccctg accatgactt tgaggtagtg gtgcttgaag tcggagtcgt	7260
cgcagccgcc ccgtcccag agcgagaagt cggcgcgtt cttggagcgg gggttgggca	7320
gagcgaaggt gacatcgttg aagaggattt tgcccgcgcg gggcatgaag ttgcgggtga	7380
tgcggaaggg ccccggcact tcagagcggg tgttgatgac ctgggcggcg agcacgatct	7440
cgtcgaagcc gttgatgttg tggcccacga tgtagagttc caggaaagcg ggccggccct	7500
ttacggtggg cagcttcttt agctcttcgt aggtgagtc ctcgggcgag gcgaggccgt	7560
gctcggccag ggcccagtc gcgaggtgcg ggttgtctct gaggaaggac ttccagaggt	7620
cgcgggccag gaggtctgc aggcggtctc tgaagtcct gaactggcgg ccacggcca	7680

[0020]

ttttttcggg ggtgatgcag tagaaggtga gggggtcttg ctgccagcgg tcccagtcga	7740
gctgcagggc gaggtcgcgc gcggcggtga ccaggcgctc gtcgcccccg aatttcatga	7800
ccagcatgaa gggcacgagc tgctttccga aggcccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaag aactggatct	7920
cccgccacca gttggaggag tggctgttga tgtggtggaa gtagaagtcc cgtcgccggg	7980
ccgaacactc gtgctggctt ttgtaaaagc gagcgcagta ctggcagcgc tgcacgggct	8040
gtacctcatg cacgagatgc acctttcgcc cgcgcacgag gaagccgagg ggaaatctga	8100
gccccccgcc tggctcgcgg catggctggt tctctttctac tttggatgcg tgtccgtctc	8160
cgtctggctc ctcgaggggt gttacggtgg agcggaccac cacgccgcgc gagccgcagg	8220
tccagatata ggccgcgcggc ggtcggagtt tgatgacgac atcgcgcagc tgggagctgt	8280
ccatggtctg gagctcccgc ggccggcgga ggtcagccgg gagttcttgc aggttcacct	8340
cgcagagtcg ggccagggcg cggggcaggt ctaggttgta cctgatctct aggggcgtgt	8400
tgggtggcggc gtcgatggct tgcaggagcc cgcagccccg gggggcgacg acggtgcccc	8460
gcgggggtggt ggtggtggtg gcggtgcagc tcagaagcgg tgccgcgggc gggcccccg	8520
aggtaggggg ggctccggtc ccgcgggcag gggcggcagc ggcacgtcgg cgtggagcgc	8580
gggcaggagt tgggtgctgtg cccggaggtt gctggcgaag gcgacgacgc ggcggttgat	8640
ctcctggatc tggcgcctct gcgtgaagac gacgggcccg gtgagcttga acctgaaaga	8700
gagttcgaca gaatcaatct cggtgtcatt gaccgcggcc tggcgcagga tctcctgcac	8760
gtctcccagag ttgtcttgggt aggcgatctc ggccatgaac tgctcgatct cttcctcctg	8820
gaggtctccg cgtccggcgc gttccacggt ggccgccagg tcgttggaga tgcgccccat	8880
gagctgcgag aaggcgttga gtccgccctc gttccagact cggctgtaga ccacgcccc	8940
ctggatcatc cgggcgcgca tgaccacctg cgcgaggttg agctccacgt gccgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggtggcgg tgtgctcggc	9060
cacgaagaag ttcatgacct agcggcgcaa cgtggattcg ttgatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggcgaagtig aaaaactggg agttgcgcgc	9180
cgacacggtc aactcctcct ccagaagacg gatgagctcg gcgacggtgt cgcgcacctc	9240
gcgctcgaag gctatgggga tctcttctc cgtatgcata accacctcct cctcttctc	9300
ctcttctggc acttccatga tggcttctc ctcttcgggg ggtggcggcg gcggcggttg	9360

[0021]

gggagggggc gctctgcgcc ggcgcgcgcg caccgggagg cggtcacga agcgcgcgat	9420
catctccccg cggcgggcggc gcatggtctc ggtgacggcg cggccgttct cccgggggcg	9480
cagttggaag acgccgccgg acatctggtg ctggggcggg tggccgtgag gcagcgagac	9540
ggcgctgacg atgcatctca acaattgctg cgtaggtacg ccgccgaggg acctgaggga	9600
gtccatatcc accggatccg aaaacctttc gaggaaggcg tctaaccagt cgcagtcgca	9660
aggtaggctg agcaccgtgg cggcgggcgg ggggtggggg gagtgtctgg cggaggtgct	9720
gctgatgatg taattgaagt aggcggactt gacacggcgg atggtcgaca ggagcaccat	9780
gtccttgggt cggcctgct ggatgcggag gcggtcggt atgcccagg cttcgttctg	9840
gcacggcgc aggtccttgt agtagtcttg catgagcctt tccaccggca cctcttctcc	9900
ttcctcttct gcttcttcca tgtctgcttc ggccctgggg cggcgccgcg cccccctgcc	9960
ccccatgcgc gtgaccccg accccctgag cggttggagc agggccaggc cggcgacgac	10020
gcgctcgcc aggatggcct gctgcacctg cgtgagggtg gtttggaagt catccaagtc	10080
cacgaagcgg tggtaggcgc ccgtgttgat ggtgtaggtg cagttggcca tgacggacca	10140
gttgacggtc tggtagcccg gttgcgacat ctcggtgtac ctgagtcgcg agtaggcgcg	10200
ggagtcgaag acgtagtcgt tgcaagtccg caccaggtac tggtagccca ccaggaagtg	10260
cggcgggcgc tggcggtaga ggggccagcg cagggtggcg ggggctccgg gggccaggtc	10320
ttccagcatg aggcggtggt aggcgtagat gtacctggac atccaggatga taccgcggc	10380
ggtggtggag gcgcgcggga agtcgcgcac ccggttcag atgttgcgca ggggcagaaa	10440
gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactctagac	10500
cagggaacac gaaagccggt cagcgggcac tcttccgtgg tctggtgaat agatcgcaag	10560
ggtatcatgg cggagggcct cggttcgagc cccgggtccg ggccggacgg tccgcatga	10620
tccacgcggt taccgcccgc gtgtcgaaac caggtgtgcg acgtcagaca acggtggagt	10680
gttccttttg gcgtttttct ggccgggcgc cggcgccgcg taagagacta agccgcgaaa	10740
gcgaaagcag taagtggctc gctccccgta gccggaggga tccttgctaa gggttgcgtt	10800
gcggcgaacc ccggttcgaa tcccgtactc gggccggccg gaccgcggc taaggtgttg	10860
gattggcctc cccctcgtat aaagaccccg cttgcggatt gactccggac acggggacga	10920
gcccccttta tttttgcttt cccagatgc atccggtgct gcggcagatg cccccccgc	10980
cccagcagca gcaacaacac cagcaagagc ggcagcaaca gcagcgggag tcatgcaggg	11040
ccccctcacc caccctcggc gggccggcca cctcggcgtc cgcggccgtg tctggcgctt	11100

[0022]

gcggcgggcg	cgggggggccg	gctgacgacc	ccgaggagcc	cccgcggcgc	agggccagac	11160
actacctgga	cctggaggag	ggcgagggcc	tggcgcggt	ggggcgccg	tctcccagc	11220
gccacccgcg	ggtgcagctg	aagcgcgact	cgcgcgaggc	gtacgtgcct	cggcagaacc	11280
tgttcaggga	ccgcgcgggc	gaggagcccc	aggagatgcg	ggacaggagg	ttcagcgcag	11340
ggcgggagct	gcggcagggg	ctgaaccgcg	agcggctgct	gcgcgaggag	gactttgagc	11400
ccgacgcgcg	gacggggatc	agccccgcgc	gcgcgcacgt	ggcggccgcc	gacctggtga	11460
cggcgtagca	gcagacggtg	aaccaggaga	tcaacttcca	aaagagtttc	aacaaccacg	11520
tgcgcacgct	ggtggcgcg	gaggaggtga	ccatcggtg	gatgcacctg	tgggactttg	11580
taagcgcgct	ggtgcagaac	cccaacagca	agcctctgac	ggcgcagctg	ttcctgatag	11640
tgcagcacag	cagggacaac	gaggcggtta	gggacgcgct	gctgaacatc	accgagcccc	11700
agggtcggtg	gctgctggac	ctgattaaca	tcctgcagag	catagtgggtg	caggagcgca	11760
gcctgagcct	ggccgacaag	gtggcgggcca	tcaactactc	gatgctgagc	ctgggcaagt	11820
tttacgcgcg	caagatctac	cagacgccgt	acgtgcccac	agacaaggag	gtgaagatcg	11880
acggttttta	catgcgcatg	gcgctgaagg	tgctcacct	gagcgacgac	ctgggcgtgt	11940
accgcaacga	gcgcatccac	aaggccgtga	gcgtgagccg	gcggcgcgag	ctgagcgacc	12000
gcgagctgat	gcacagcctg	cagcgggcgc	tggcgggcgc	cggcagcggc	gacagggagg	12060
cggagtccta	cttcgatgcg	ggggcgggacc	tgcgctgggc	gcccagccgg	cgggccctgg	12120
aggccgcggg	ggtccgcgag	gactatgacg	aggacggcga	ggaggatgag	gagtacgagc	12180
tagaggaggg	cgagtacctg	gactaaaccg	cgggtggtgt	ttccggtaga	tgcaagaccc	12240
gaacgtgggtg	gacccggcgc	tgcgggcggc	tctgcagagc	cagccgtccg	gccttaactc	12300
ctcagacgac	tggcgacagg	tcatggaccg	catcatgtcg	ctgacggcgc	gtaacccgga	12360
cgcgttccgg	cagcagccgc	aggccaacag	gctctccgcc	atcctggagg	cgggtggtgcc	12420
tgcgcgctcg	aacccccacg	acgagaaggt	gctggccata	gtgaacgcgc	tggccgagaa	12480
cagggccatc	cgcgggacg	aggccgggct	ggtgtacgac	gcgctgctgc	agcgcgtggc	12540
ccgctacaac	agcggcaacg	tgcagaccaa	cctggaccgg	ctggtggggg	acgtgcgcga	12600
ggcgggtggcg	cagcgcgagc	gcgcggatcg	gcagggcaac	ctgggctcca	tgggtggcgt	12660
gaatgccttc	ctgagcacgc	agccggccaa	cgtgccgcgg	gggcaggaag	actacaccaa	12720
ctttgtgagc	gcgctgcggc	tgatggtgac	cgagaccccc	cagagcgagg	tgtaccagtc	12780

[0023]

```

gggcccggac tacttcttcc agaccagcag acagggcctg cagacggtga acctgagcca 12840
ggctttcaag aacctgcggg ggctgtgggg cgtgaaggcg cccaccggcg accgggcgac 12900
ggtgtccagc ctgctgacgc ccaactcgcg cctgctgctg ctgctgatcg cgccgttcac 12960
ggacagcggc agcgtgtccc gggacaccta cctggggcac ctgctgaccc tgtaccgca 13020
ggccatcggg caggcgcagg tggacgagca caccttcag gagatcacca gcgtgagccg 13080
cgcgctgggg caggaggaca cgagcagcct ggaggcgact ctgaactacc tgctgaccaa 13140
ccggcggcag aagattccct cgctgcacag cctgacctcc gaggaggagc gcattctgcg 13200
ctacgtgcag cagagcgtga gcctgaacct gatgcgcgac ggggtgacgc ccagcgtggc 13260
gctggacatg accgcgcgca acatggaacc gggcatgtac gccgcgcacc ggccttacat 13320
caaccgcctg atggactacc tgcatcgcgc ggcgccgtg aaccccgagt actttacaa 13380
cgccatcctg aacccgcact ggctcccgcc gccgggttc tacagcgggg gcttcgaggt 13440
cccggagacc aacgatggct tcctgtggga cgacatggac gacagcgtgt tctccccgcg 13500
gccgcaggcg ctggcggaag cgtccctgct gcgtccaag aaggaggagg aggaggaggc 13560
gagtcgccgc cgcggcagca gcggcgtggc ttctctgtcc gagctggggg cggcagccgc 13620
cgcgcgcccc gggtccctgg gcggcagccc cttccgagc ctggtggggg ctctgcacag 13680
cgagcgcacc acccgccctc ggctgctggg cgaggacgag tacctgaata actccctgct 13740
gcagccggtg cgggagaaaa acctgcctcc cgccttcccc aacaacggga tagagagcct 13800
ggtggacaag atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcctgcgct 13860
ccggccgccc acgcggcgcc agcgccacga ccggcagcgg gggctggtgt gggatgacga 13920
ggactccgcg gacgatagca gcgtgctgga cctgggaggg agcggcaacc cgttcgcgca 13980
cctgcgcccc cgctgggga ggatgtttta aaaaaaaaaa aaaaaagcaa gaagcatgat 14040
gcaaaaatta aataaaactc accaaggcca tggcgaccga gcgttggttt cttgtgttcc 14100
cttcagtatg cggcgcgcg cgatgtacca ggaggacct cctccctctt acgagagcgt 14160
ggtgggcgcg gcggcggcgg cgccctcttc tccctttgcg tcgcagctgc tggagccgcc 14220
gtacgtgcct ccgcgctacc tgcggcctac gggggggaga aacagcatcc gttactcgga 14280
gctggcgcgc ctgttcgaca ccaccgggt gtacctggtg gacaacaagt cggcggacgt 14340
ggcctccctg aactaccaga acgaccacag caattttttg accacggtca tccagaacaa 14400
tgactacagc ccgagcgagg ccagcacca gaccatcaat ctggatgacc ggtcgcactg 14460
gggcggcgac ctgaaaacca tcctgcacac caacatgccc aacgtgaacg agttcatgtt 14520

```

[0024]

caccaataag ttcaaggcgc gggatgatggt gtcgcgctcg cacaccaagg aagaccgggt	14580
ggagctgaag tacgagtggg tggagttcga gctgccagag ggcaactact ccgagaccat	14640
gaccattgac ctgatgaaca acgcgatcgt ggagcactat ctgaaagtgg gcaggcagaa	14700
cggggtcctg gagagcgaca tcgggggtcaa gtctgacacc aggaacttcc gcctggggct	14760
ggaccccgctg accgggctgg ttatgcccgg ggtgtacacc aacgaggcct tccatcccga	14820
catcatcctg ctgcccggct gcgggggtgga cttcacttac agccgcctga gcaacctcct	14880
gggcatccgc aagcggcagc ccttccagga gggcttcagg atcacctacg aggacctgga	14940
ggggggcaac atccccgcgc tcctcgatgt ggaggcctac caggatagct tgaaggaaaa	15000
tgaggcggga caggaggata ccgccccgc cgcctccgcc gccgccgagc agggcgagga	15060
tgctgctgac accgcggccg cggacggggc agaggccgac cccgctatgg tgggtggaggc	15120
tcccgagcag gaggaggaca tgaatgacag tgcgggtgcgc ggagacacct tcgtcacccg	15180
gggggaggaa aagcaagcgg aggccgaggc cgcggccgag gaaaagcaac tggcggcagc	15240
agcggcggcg gcggcggttg ccgcggcgga ggctgagtct gaggggacca agcccgccaa	15300
ggagcccgtg attaagcccc tgaccgaaga tagcaagaag cgcagttaca acctgtctaa	15360
ggacagcacc aacaccgcgt accgcagctg gtacctggcc tacaactacg gcgacccgtc	15420
gacgggggtg cgctcctgga ccctgctgtg cacgccggac gtgacctgcg gctcggagca	15480
ggtgtactgg tcgtgccccg acatgatgca agaccccgtg accttccgct ccacgcggca	15540
ggtcagcaac ttcccgggtg tgggcgccga gctgctgccc gtgcactcca agagcttcta	15600
caacgaccag gccgtctact ccagctcat ccgccagttc acctctctga cccacgtgtt	15660
caatcgcttt cctgagaacc agattctggc gcgcccgcgc gccccacca tcaccaccgt	15720
cagtgaaaac gtctctgctc tcacagatca cgggacgcta ccgctgcgca acagcatcgg	15780
aggagtccag cgagtgaccg ttactgacgc cagacgccgc acctgcccct acgtttacaa	15840
ggccttgggc atagtctcgc cgcgcgtcct ttccagccgc actttttgag caacaccacc	15900
atcatgtcca tcctgatctc acccagcaat aactccggct ggggactgct gcgcgcgcgc	15960
agcaagatgt tcggaggggc gaggaagcgt tccgagcagc accccgtgcg cgtgcgcggg	16020
cacttccgcg ccccttgggg agcgcacaaa cgcggccgcg cggggcgcac caccgtggac	16080
gacgccatcg actcggtggt ggagcaggcg cgcaactaca ggcccgcggt ctctaccgtg	16140
gacgcggcca tccagaccgt ggtgcggggc gcgcggcggt acgccaagct gaagagccgc	16200

[0025]

cggaagcgcg	tggccccccg	ccaccgccgc	cgacccgggg	ccgccgcaa	acgcgccgcc	16260
gcggccctgc	ttcgccgggc	caagcgcacg	ggccgccgcg	ccgccatgag	ggccgcgcgc	16320
cgcttgcccg	ccggcatcac	cgcgccacc	atggcccccc	gtacccgaag	acgcgcggcc	16380
gccgccgccg	ccgccgcat	cagtgcacatg	gccagcaggc	gccggggcaa	cgtgtactgg	16440
gtgcgcgact	cggtgaccgg	cacgcgcgtg	cccgtgcgct	tccgcccccc	gcggacttga	16500
gatgatgtga	aaaaacaaca	ctgagtctcc	tgctgtttgtg	tgtatcccag	cggcggcggc	16560
gcgcgcagcg	tcatgtccaa	gcgcaaaatc	aaagaagaga	tgctccaggt	cgtcgcgccg	16620
gagatctatg	ggccccgaa	gaaggaagag	caggattcga	agccccgcaa	gataaagcgg	16680
gtcaaaaaga	aaaagaaaga	tgatgacgat	gccgatgggg	aggtggagtt	cctgcgcgcc	16740
acggcgccca	ggcgcccggt	gcagtgggaag	ggccggcgcg	taaagcgcgt	cctgcgcccc	16800
ggcaccgcgg	tggtcttcac	gcccggcgag	cgctccaccc	ggactttcaa	gcgcgtctat	16860
gacgaggtgt	acggcgacga	agacctgctg	gagcaggcca	acgagcgctt	cggagagttt	16920
gcttacggga	agcgtcagcg	ggcgctgggg	aaggaggacc	tgctggcgct	gccgctggac	16980
cagggcaacc	ccacccccag	tctgaagccc	gtgaccctgc	agcaggtgct	gccgagcagc	17040
gcaccctccg	aggcgaagcg	gggtctgaag	cgcgaggcg	gcgacctggc	gcccaccgtg	17100
cagctcatgg	tgcccaagcg	gcagaggctg	gaggatgtgc	tgagaaaaat	gaaagtagac	17160
cccgtcttgc	agccggacat	cagggtccgc	cccatcaagc	aggtggcgcc	gggcctcggc	17220
gtgcagaccg	tggacgtggt	catccccacc	ggcaactccc	ccgccgccgc	caccactacc	17280
gctgcctcca	cggacatgga	gacacagacc	gatcccgccg	cagccgcagc	cgcagccgcc	17340
gccgcgacct	cctcggcgga	ggtgcagacg	gaccttggc	tgccgccggc	gatgtcagct	17400
ccccgcgcgc	gtcgcgggcg	caggaagtac	ggcgccgcca	acgcgctcct	gcccagtagc	17460
gccttgcatc	cttccatcgc	gcccaccccc	ggctaccgag	gctataccta	ccgcccgcga	17520
agagccaagg	gttccacccg	ccgtccccgc	cgacgcgccg	ccgccaccac	ccgccgccgc	17580
cgccgcagac	gccagcccg	actggctcca	gtctccgtga	ggaaagtggc	gcgcgacgga	17640
cacaccctgg	tgctgcccag	ggcgcgctac	caccccagca	tcgtttaaaa	gcctgtttgtg	17700
gttctttgcag	atatggccct	cacttgccgc	ctccgtttcc	cggtgccggg	ataccgagga	17760
ggaagatcgc	gccgcaggag	gggtctggcc	ggccgcggcc	tgagcggagg	cagccgccgc	17820
gcgcaccggc	ggcgacgcgc	caccagccga	cgcatgcgcg	gcggggtgct	gcccctgtta	17880
atccccctga	tcgccgcggc	gateggcgcc	gtgcccggga	tcgcctccgt	ggccttgcaa	17940

[0026]

gcgtcccaga	ggcattgaca	gacttgcaaa	cttgcaaata	tggaaaaaaa	aaccccaata	18000
aaaaagtcta	gactctcacg	ctcgcttggg	cctgtgacta	ttttgtagaa	tggaagacat	18060
caacttttgcg	tcgctggccc	cgcgtcacgg	ctcgcgcccc	ttcctgggac	actggaacga	18120
tatcggcacc	agcaacatga	gcggtggcgc	cttcagttgg	ggctctctgt	ggagcggcat	18180
taaaagtata	gggtctgccc	ttaaaaatta	cggctccccg	gcctggaaca	gcagcacggg	18240
ccagatgttg	agagacaagt	tgaagagca	gaacttccag	cagaaggtgg	tgaggggcct	18300
ggcctccggc	atcaacgggg	tggtggacct	ggccaaccag	gccgtgcaga	ataagatcaa	18360
cagcagactg	gacccccggc	cgcgggtgga	ggaggtgccc	ccggcgctgg	agacggtgtc	18420
ccccgatggg	cgtggcgaga	agcgccccgc	gcccgatagg	gaagagacca	ctctggtcac	18480
gcagaccgat	gagccgcccc	cgtatgagga	ggccctgaag	caaggtctgc	ccaccacgcg	18540
gcccatecgc	cccatggcca	ccgggggtgt	gggccgccac	acccccgcca	cgctggactt	18600
gcctccgccc	gccgatgtgc	cgcagcagca	gaaggcggca	cagccgggccc	cgcccgcgac	18660
cgcctcccgt	tcctccgccc	gtcctctgcg	ccgcgcggcc	agcggccccc	gcgggggggt	18720
cgcgaggcac	ggcaactggc	agagcacgct	gaacagcatc	gtgggtctgg	gggtgcggtc	18780
cgtgaagcgc	cgcgatgct	actgaatagc	ttagctaacg	tgttgatatgt	gtgtatgcgc	18840
cctatgtcgc	cgccagagga	gctgctgagt	cgcgcgcgtt	cgcgcgcca	ccaccaccgc	18900
cactccgccc	ctcaagatgg	cgaccccatc	gatgatgccg	cagtggctgt	acatgcacat	18960
ctcgggccag	gacgcctcgg	agtacctgag	ccccgggctg	gtgcagttcg	cccgcgccac	19020
cgagagctac	ttcagcctga	gtaacaagtt	taggaacccc	acggtggcgc	ccacgcacga	19080
tgtgaccacc	gaccggtctc	agcgcctgac	gctgcggttc	attcccgtgg	accgcgagga	19140
caccgcgtac	tcgtacaagg	cgcggttcac	cctggccgtg	ggcgacaacc	gcgtgctgga	19200
catggcctcc	acctactttg	acatccgcgg	ggtgctggac	cgggggtccca	ctttcaagcc	19260
ctactctggc	accgcctaca	actccctggc	ccccaggggc	gtccccaact	cctgcgagtg	19320
ggagcaagag	gaaactcagg	cagttgaaga	agcagcagaa	gaggaagaag	aagatgctga	19380
cgggtcaagct	gaggaagagc	aagcagctac	caaaaagact	catgtatatg	ctcaggctcc	19440
cctttctggc	gaaaaaatta	gtaaagatgg	tctgcaaata	ggaacggacg	ctacagctac	19500
agaacaaaaa	cctattttatg	cagaccctac	attccagccc	gaaccccaaa	tcggggagtc	19560
ccagtggaat	gaggcagatg	ctacagtcgc	cggcggtaga	gtgctaaaga	aatctactcc	19620

[0027]

catgaaacca tgctatggtt cctatgcaag acccacaat gctaattggag gtcagggtgt	19680
actaacggca aatgcccagg gacagctaga atctcaggtt gaaatgcaat tcttttcaac	19740
ttctgaaaac gcccgtaacg aggctaacaa cattcagccc aaatttggtgc tgtatagtga	19800
ggatgtgcac atggagaccc cggatacgca cttttcttac aagcccgcaa aaagcgatga	19860
caattcaaaa atcatgctgg gtcagcagtc catgcccac agacctaat acatcggtt	19920
cagagacaac tttatcggcc tcatgtatta caatagcact ggcaacatgg gagtgttgc	19980
aggtcaggcc tctcagttga atgcagtggg ggacttgcaa gacagaaaca cagaactgtc	20040
ctaccagctc ttgcttgatt ccatgggtga cagaaccaga tacttttcca tgtggaatca	20100
ggcagtggac agttatgacc cagatgttag aattattgaa aatcatggaa ctgaagacga	20160
gtcctccaac tattgtttcc ctctgggtgg cataggggta actgacactt accaggctgt	20220
taaaaccaac aatggcaata acgggggcca ggtgacttgg acaaaagatg aaacttttgc	20280
agatcgcaat gaaatagggg tgggaaacaa ttctgctatg gagatcaacc tcagtgccaa	20340
cctgtggaga aacttcctgt actccaacgt ggcgtgtac ctaccagaca agcttaagta	20400
caaccctcc aatgtggaca tctctgacaa cccaacacc tacgattaca tgaacaagcg	20460
agtgggtggc ccggggctgg tggactgcta catcaacctg ggcgcgcgt ggtcgtgga	20520
ctacatggac aacgtcaacc cttcaacca ccaccgcaat gcgggcctgc gctaccgctc	20580
catgctcctg ggcaacgggc gctacgtgcc cttccacatc caggtgccc agaagttctt	20640
tgccatcaag aacctcctcc tcttgcggg ctctacacc tacgagtgga acttcaggaa	20700
ggatgtcaac atggtcctcc agagctctct gggtaacgat ctcagggtgg acggggccag	20760
catcaagttc gagagcatct gcctctacgc caccttcttc cccatggccc acaacacggc	20820
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20880
ctccgcgcc aacatgtctt accccatacc cgccaacgcc accaacgtcc ccatctccat	20940
cccctcgcgc aactgggcgg ccttcgcgg ctgggccttc acccgctca agaccaagga	21000
gacccctcc ctgggctcgg gattcgacct ctactacacc tactcgggct ccattcccta	21060
cctggacggc accttctacc tcaaccacac tttcaagaag gtctcggtca cttcgcactc	21120
ctcggtcagc tggccgggca acgaccgtct gctcaccccc aacgagttcg agatcaagcg	21180
ctcggtcgac ggggagggt acaacgtggc ccagtgaac atgaccaagg actggttctt	21240
ggtccagatg ctggccaact acaacatcgg ctaccagggc ttctacatcc cagagagcta	21300
caaggacagg atgtactcct tcttcaggaa cttccagccc atgagccggc aggtggtgga	21360

[0028]

ccagaccaag tacaaggact accaggaggt gggcatcatc caccagcaca acaactcggg	21420
cttcgtgggc tacctcgccc ccacatgcg cgaggacag gcctaccccg ccaacttccc	21480
ctatccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tcctctgcga	21540
ccgcaccctc tggcgcatcc ctttctccag caacttcatg tccatgggtg cgctctcgga	21600
cctgggccag aacttgctct acgccaactc cgcccacgcc ctcgacatga ccttcgaggt	21660
cgaccccatg gacgagccca cctttctcta tgttctgttc gaagtctttg acgtggtccg	21720
ggtccaccag ccgcaccgcg gcgtcatcga gaccgtgtac ctgcgtacgc ctttctcggc	21780
cggcaacgcc accacctaaa gaagcaagcc gcagtcacgc ccgcctgcat gccgtcgggt	21840
tccaccgagc aagagctcag ggccatcgtc agagacctgg gatgcgggcc ctatTTTTTg	21900
ggcaccttcg acaagcgctt ccctggcttt gtctccccac acaagctggc ctgcgccatc	21960
gtcaacacgg ccggccgcga gaccgggggc gtgcaactggc tggccttcgc ctggaacccg	22020
cgctccaaaa catgcttcct ctttgacccc ttcggtttt cgaccagcg gctcaagcaa	22080
atctacgagt tcgagtacga gggcttgctg cgtcgcagcg ccatgcctc ctgccccgac	22140
cgtgcgtca ccctcgaaaa gtccaccag accgtgcagg ggcccactc ggccgcctgc	22200
ggtctcttct gctgcatgtt tctgcacgcc tttgtgact ggctcagag tccatggac	22260
cgaacccca ccatgaactt gctgacgggg gtgcccact ccatgctca gagccccag	22320
gtcgagccca ccctgcgcg caaccaggag cagctctaca gttcctgga gcgccactcg	22380
ccttacttcc gccgccacag cgcacagatc aggagggcc cctccttctg ccacttgcaa	22440
gagatgcaag aagggttaata acgatgtaca cactttttt ctcaataaat ggcatctttt	22500
tatttataca agctctctgg ggtattcatt tcccaccacc acccgccgtt gtcgccatct	22560
ggctctattt agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcaggac	22620
acgttgcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tcggggaagt tttcgctcca caggctgcgg gtcagcacca gcgcgttcat caggtcgggc	22740
gccgagatct tgaagtcgca gttggggccg ccgccctgcg cgcgcgagtt gcggtacacc	22800
gggttgccagc actggaacac caacagcgcc ggggtgcttca cgttgccag cacgctgcgg	22860
tcggagatca gctcgcgctc caggctctcc gcgttgctca gcgcgaacgg ggtcatcttg	22920
ggcacttgcc gccccaggaa gggcgcgtgc cccggtttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccgtg cccggactcg gcgttgggggt acagcgcgcg catgaaggcc	23040

[0029]

tgcacatctggc ggaaggccat ctgggccttg gcgccctccg agaagaacat gccgcaggac	23100
ttgcccagaga actggtttgc ggggcagctg gcgtcgtgca ggcagcagcg cgcgtcgggtg	23160
ttggcgatct gcaccacgtt gcgccccac cggttcttca cgatcttggc cttggacgat	23220
tgctccttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttcacca tgctgctgcc gtgcagacac ttcagctcgc cctccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtgggctcg aaagacttgt aggtcacctc cgcaaggac	23400
tgcaggtacc cctgcaaaaa gcggcccatc atggtcacga aggtcttggt gctgctgaag	23460
gtcagctgca gcccgcggtg ctctcgttcc agccaggtct tgcacacggc cgccagcgc	23520
tccacctggt cgggcagcat cttgaagtcc accttcagct cattctccac gtggtacttg	23580
tccatcagcg tgccgcgccg ctccatgccc ttctcccagg ccgacaccag cggcaggctc	23640
acggggttct tcaccatcac cgtggccgcc gcctccgccg cgctttcgct ttccgccccg	23700
ctgttctctt cctcttctc ctcttctcgc ccgcgcacca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttg cgcttgccgt tgcgccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcac agcgcggcct cttcttgctc gtcctcgtg	23880
tccagaatga cctccgggga gggggggttg gtcacctca gtaccaggc acgcttctt	23940
ttcttcttg gggcgttcgc cagctccgcg gctgcggccg ctgccagggt cgaaggccga	24000
gggctgggcg tgccgcggcac cagcgcgtcc tgcgagccgt cctcgtctc ctccgactcg	24060
agacggaggc gggcccgtt cttcgggggc gcgcggggcg gcggaggcgg cggcggcgac	24120
ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgcc gcgtccgcgc	24180
tcgggggttg tctcgcgtg gtcctcttcc cgactggcca tctccactg ctcttctcc	24240
tataggcaga aagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgccccct ctgagccctc caccaccgcc gccaccaccg ccaatgccgc cgcggacgac	24360
gcgccaccg agaccaccgc cagtaccacc ctccccagcg acgcaccccc gctcgagaat	24420
gaagtgtga tcgagcagga cccgggtttt gtgagcggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggtcgc cgctcagtg ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata aggatgagac agcagtcggg cgggggaacg gaagccatga tgctgatgac	24600
ggctacctag acgtgggaga cgacgtgctg cttaaagacc tgcaccgcca gtgcgtcatc	24660
gtctgcgacg cgctgcagga gcgtgcgaa gtgcccttg acgtggcgga ggtcagccgc	24720
gcctacgagc ggcacctctt cgcgccgcac gtgccccca agcgcggga gaacggcacc	24780

[0030]

tgcgagccca acccgcgtct caactttctac ccggtcttcg cggtagccga ggtgctggcc	24840
acctaccaca tctttttcca aaactgcaag atccccctct cctgccgcgc caaccgcacc	24900
cgcgccgaca aaacctgac cctgcggcag ggcccccaca tacctgatat cgcctctctg	24960
gaggaagtgc ccaagatctt cgagggtctc ggtcgcgacg agaaacgggc ggccaacgct	25020
ctgcacggag acagcgaaaa cgagagtcac tcgggggtgc tggtaggagct cgagggcgac	25080
aacgcgcgcc tgcccgact caagcgcagc atagaggtea cccactttgc ctaccggcg	25140
ctcaacctgc cccccaaggt catgagtgtg gtcatgggcg agctcatcat gcgccgcgcc	25200
cagccccctgg ccgcggatgc aaacttgcaa gagtctctcg aggaaggcct gcccgcggtc	25260
agcgacgagc agctggcgcg ctggctggag acccgcgacc ccgcgcagct ggaggagcgg	25320
cgcaagctca tgatggccgc ggtgctggtc accgtggagc tcgagtgtct gcagcgcttc	25380
ttcgccgacc ccgagatgca gcgcaagctc gaggagaccc tgcactacac cttccgccag	25440
ggctacgtgc gccaggcctg caagatctcc aacgtggagc tctgcaacct ggtctctac	25500
ctgggcatcc tgcacgagaa ccgcctcggg cagaacgtcc tgcactccac cctcaaagg	25560
gaggcgccgc gcgactacat ccgcgactgc gcctacctct tctctgcta cacctggcag	25620
acggccatgg gggctctggca gcagtgcctg gaggagcgca acctcaagga gctggaaaag	25680
ctcctcaagc gcaccctcag ggacctctgg acgggcttca acgagcgctc ggtggccgcc	25740
gcgctggcgg acatcatctt tcccgagcgc ctgctcaaga ccctgcagca gggcctgccc	25800
gacttcacca gccagagcat gctgcagaac ttcaggactt tcatcctgga gcgctcgggc	25860
atcctgccgg ccacttgctg cgcgctgccc agcgacttcg tgcccatcaa gtacaggag	25920
tgcccgcgcg cgctctgggg ccactgtac ctcttcagc tggccaacta cctcgcctac	25980
cactcggacc tcatggaaga cgtgagcggc gagggcctgc tcgagtgcc a ctgccgtgc	26040
aacctctgca cgcgccaccg ctctctagtc tgcaaccgcg agctgctcag cgagagtcag	26100
attatcggt a ccttcgagct gcagggtccc tcgcctgacg agaagtccgc ggctccagg	26160
ctgaaactca ctccggggct gtggacttcc gcctacctac gcaaatttgt acctgaggac	26220
taccacgccc acgagatcag gttctacgaa gaccaatccc gcccgcccaa ggcggagctc	26280
accgcctgcg tcatcaccca ggggcacatc ctgggccaat tgcaagccat caacaaagcc	26340
cgcgcgagag tcttgctgaa aaagggtcgg ggggtgtacc tggaccccca gtccggcgag	26400
gagctaaacc cgctaccccc gccgcgcgcc cagcagcggg accttgcttc ccaggatggc	26460

[0031]

accagaaag aagcagcagc cgccgccc gccgcagcca tacatgcttc tggaggaaga	26520
ggaggaggac tgggacagtc aggcagagga ggtttcggac gaggagcagg aggagatgat	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgaaca ccatcgccct cggtcgcagc cccctcgccg gggcccctga aatcctccga	26700
accagcacc agcgtataa cctccgctcc tccggcgccg gcgccaccgc cccgcagacc	26760
caaccgtaga tgggacacca caggaaccgg ggtcggtaag tccaagtgcc cgccgcccgc	26820
accgcagcag cagcagcagc agcgccaggg ctaccgctcg tggcgccggc acaagaacgc	26880
catagtcgcc tgcttgcaag actcggggg caacatctct ttgccccgc gtttcttct	26940
attccaccac ggggtcgcct ttccccgcaa tgtcttgcac tactaccgtc atctctacag	27000
cccctactgc agcggcgacc cagaggcggc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatc ctccgcgggc aagacagcgg cagcagcggc caggagacc gcggcagcag	27120
cggcgggagc ggtgggcgca ctgcgcctct cgcccaacga accctctcg acccgggagc	27180
tcagacacag gatcttcccc actttgtatg ccatcttcca acagagcaga ggccaggagc	27240
aggagctgaa aataaaaaac agatctctgc gtcctctcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagcttcgg cgacgctgg aggacgcgga ggcactcttc agcaaatact	27360
gcgcgctcac tcttaaagac tagctccgcg cccttctcga atttaggcgg gagaaaacta	27420
cgtcatcgcc ggccgcccgc cagccccccc agccgagatg agcaaagaga ttcccacgcc	27480
atacatgtgg agctaccagc cgcagatggg actcgcgggc ggagcgggccc aggactactc	27540
caccgcgatg aactacatga gcgcgggacc ccacatgatc tcacaggtca acgggatccg	27600
cgcccagcga aaccaaatac tgctggaaca ggcgggccatc accgccacgc cccgccataa	27660
tctcaacccc cgaaattggc ccgcgcct cgtgtaccag gaaacccct ccgccaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcag gggcgagct	27780
cgcgggcggc tttcgtcacg gggcgcggcc gtcggacca ggtataagac acctgatgat	27840
cagaggccga ggtatccagc tcaacgacga gtcggtgagc tcttcgctcg gtctccgtcc	27900
ggacggaact ttccagctcg ccgatccgg ccgctcttcg ttcacgcccc gccaggcgta	27960
cctgactctg cagacctcgt cctcgagacc ccgctccggc ggcatcgga cctccagtt	28020
cgtggaggag ttcgtgccct cgggtctactt caacccttc tcgggacctc ccggacgcta	28080
ccccgaccag ttcatccga actttgacgc ggtgaaggac tcggcgagc gctacgactg	28140
aatgtcaggt gtcgaggcag agcagcttcg cctgagacac ctcgagcact gccgccgcca	28200

[0032]

caagtgtctt	gccccgcggtt	ctggtgagtt	ctgctacttt	cagctacccg	aggagcatac	28260
cgagggggccg	gcgcacggcg	tccgcctgac	cacccagggc	gaggttacct	gttccctcat	28320
ccgggagttt	accctccgtc	ccctgctagt	ggagcgggag	cggggtcctt	gtgtcctaac	28380
tatgcctgc	aactgcccta	accctggatt	acatcaagat	ctttgctgtc	atctctgtgc	28440
tgagtttaat	aaacgctgag	atcagaatct	actggggctc	ctgtcgccat	cctgtgaacg	28500
ccaccgtctt	cacccacccc	gaccaggccc	aggcgaacct	cacctgcggt	ctgcatcgga	28560
gggccaagaa	gtacctcacc	tggtaactta	acggcacccc	ctttgtggtt	tacaacagct	28620
tcgacgggga	cggagtctcc	ctgaaagacc	agctctccgg	tctcagctac	tccatccaca	28680
agaacaccac	cctccaactc	ttccctccct	acctgccggg	aacctacgag	tgcgtcaccg	28740
gccgtgcac	ccacctcacc	cgcctgatcg	taaaccagag	ctttccggga	acagataact	28800
ccctcttccc	cagaacagga	ggtgagctca	ggaaactccc	cggggaccag	ggcggagacg	28860
taccttcgac	ccttgtgggg	ttaggatttt	ttattaccgg	gttgctggct	cttttaatca	28920
aagtttcctt	gagatttggt	ctttccttct	acgtgtatga	acacctcaac	ctccaataac	28980
tctacccttt	cttcggaatc	aggtgacttc	tctgaaatcg	ggcttggtgt	gctgcttact	29040
ctgttgattt	ttttccttat	catactcagc	cttctgtgcc	tcaggctcgc	cgcctgctgc	29100
gcacacatct	atatctactg	ctggttgctc	aagtgcaggg	gtcgccaccc	aagatgaaca	29160
ggtacatggt	cctatcgatc	ctaggcctgc	tggccctggc	ggcctgcagc	gccgccaaaa	29220
aagagattac	ctttgaggag	cccgttgca	atgtaacttt	caagcccagag	ggtgaccaat	29280
gcaccaccct	cgtcaaatgc	gttaccaatc	atgagaggct	gcgcatcgac	tacaaaaaca	29340
aaactggcca	gtttgcggtc	tatagtgtgt	ttacgcccg	agaccctct	aactactctg	29400
tcaccgtctt	ccagggcgga	cagtctaaga	tattcaatta	cactttccct	ttttatgagt	29460
tatgcgatgc	ggtcatgtac	atgtcaaaac	agtacaacct	gtggcctccc	tctccccagg	29520
cgtgtgtgga	aaatactggg	tcttactgct	gtatggcttt	cgcaatcact	acgctcgtc	29580
taatctgcac	ggtgctatac	ataaaattca	ggcagaggcg	aatctttatc	gatgaaaaga	29640
aaatgccttg	atcgctaaca	ccggctttct	atctgcagaa	tgaatgcaat	cacctcccta	29700
ctaataccca	ccacctcctt	tgcgattgcc	catgggttga	cacgaatcga	agtgccagtg	29760
gggtccaatg	tcaccatggt	gggccccgcc	ggcaattcca	ccctcatgtg	ggaaaaat	29820
gtccgcaatc	aatgggttca	tttctgtctt	aaccgaatca	gtatcaagcc	cagagccatc	29880

[0033]

tgcgatgggc	aaaatctaac	tctgatcaat	gtgcaaatga	tggatgctgg	gtactattac	29940
gggcagcggg	gagaaatcat	taattactgg	cgaccccaca	aggactacat	gctgcatgta	30000
gtcgaggcac	ttcccactac	cacccccact	accacctctc	ccaccaccac	caccactact	30060
actactacta	ctactactac	tactactacc	actaccgctg	cccgccatac	ccgcaaaagc	30120
accatgatta	gcacaaagcc	ccctcgtgct	cactcccacg	ccggcggggc	catcggtgcg	30180
acctcagaaa	ccaccgagct	ttgcttctgc	caatgcacta	acgccagcgc	tcatgaactg	30240
ttcgacctgg	agaatgagga	tgtccagcag	agctccgctt	gcctgaccca	ggaggctgtg	30300
gagcccgttg	ccctgaagca	gatcggtgat	tcaataattg	actcttcttc	ttttgccact	30360
cccgaatacc	ctcccgattc	tactttccac	atcacgggta	ccaaagaccc	taacctctct	30420
ttctacctga	tgctgctgct	ctgtatctct	gtggctctctt	ccgcgctgat	gttactgggg	30480
atgttctgct	gcctgatctg	ccgcagaaag	agaaaagctc	gctctcaggg	ccaaccactg	30540
atgcccttcc	cctaccccc	ggattttgca	gataacaaga	tatgagctcg	ctgctgacac	30600
taaccgcttt	actagcctgc	gctctaacce	ttgtcgcttg	cgactcgaga	ttccacaatg	30660
tcacagctgt	ggcaggagaa	aatgttactt	tcaactccac	ggccgatacc	cagtggctcgt	30720
ggagtggctc	aggtagctac	ttaactatct	gcaatagctc	cacttcccc	ggcatatccc	30780
caaccaagta	ccaatgcaat	gccagcctgt	tcacctcat	caacgcttcc	accctggaca	30840
atggactcta	tgtaggctat	gtaccctttg	gtgggcaagg	aaagaccac	gcttacaacc	30900
tggaagttec	ccagcccaga	accactacce	aagcttctcc	caccaccacc	accaccacca	30960
ccatcaccag	cagcagcagc	agcagcagcc	acagcagcag	cagcagatta	ttgactttgg	31020
ttttggccag	ctcatctgcc	gctaccaggg	ccatctacag	ctctgtgccc	gaaaccactc	31080
agatccaccg	cccagaaacg	accaccgcca	ccacctaca	cacctccagc	gatcagatgc	31140
cgaccaacat	cacccccttg	gctcttcaaa	tgggacttac	aagccccact	ccaaaaccag	31200
tggatgcggc	cgaggtctcc	gccctcgtca	atgactgggc	ggggctggga	atgtggtggt	31260
tcgccatagg	catgatggcg	ctctgcctgc	ttctgctctg	gctcatctgc	tgcctccacc	31320
gcaggcgagc	cagaccccc	atctatagac	ccatcattgt	cctgaacccc	gataatgatg	31380
ggatccatag	attggatggc	ctgaaaaacc	tacttttttc	ttttacagta	tgataaattg	31440
agacatgcct	cgcattttct	tgtacatgtt	ccttctccca	ccttttctgg	ggtgttctac	31500
gctggccgct	gtgtctcacc	tggaggtaga	ctgcctctca	cccttccactg	tctacctgct	31560
ttacggattg	gtcaccctca	ctctcatctg	cagcctaate	acagtaatca	tcgccttcat	31620

[0034]

ccagtgcatt gattacatct gtgtgcgcct cgcatacttc agacaccacc cgcagtaccg	31680
agacaggaac attgccaac ttctaagact gctctaataca tgcataagac tgtgatctgc	31740
cttctgatcc tctgcatcct gcccaccctc acctcctgcc agtacaccac aaaatctccg	31800
cgcaaaagac atgcctcctg ccgcttcacc caactgtgga atatacccaa atgctacaac	31860
gaaaagagcg agctctccga agcttggctg tatgggggtca tctgtgtctt agttttctgc	31920
agcactgtct ttgccctcat aatctacccc tactttgatt tgggatggaa cgcgatcgat	31980
gccatgaatt accccacctt tcccgacccc gagataattc cactgcgaca agttgtaccc	32040
gttgtcgtta atcaacgccc cccatccctt acgcccactg aaatcagcta ctttaacctt	32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatcagta ccgagcagcg	32160
tctcctagag aggcgccaggc aggcggtga gcaagagcgc ctcaatcagg agctccgaga	32220
tctcgttaac ctgcaccagt gcaaaagagg catcttttgt ctggtaaagc aggccaaagt	32280
cacctacgag aagaccggca acagccaccg cctcagttac aaattgcca cccagcgcca	32340
gaagctggtg ctcatggtgg gtgagaatcc catcacgctc acccagcact cggtagagac	32400
cgaggggtgt ctgcactccc cctgtcgggg tccagaagac ctctgcaccc tggtaaagac	32460
cctgtgcggt ctcagagatt tagtcccctt taactaatca aacactggaa tcaataaaaa	32520
gaatcactta cttaaaatca gacagcaggt ctctgtccag ttatttcagc agcacctcct	32580
tcccctcctc ccaactctgg tactccaaac gccttctggc ggcaaacttc ctccacaccc	32640
tgaagggaat gtcagattct tgctcctgtc cctccgcacc cactatcttc atgttgttgc	32700
agatgaagcg caccaaaacg tctgacgaga gtttcaaccc cgtgtacccc tatgacacgg	32760
aaagcggccc tccctccgtc cttttctca cccctccctt cgtgtctccc gatggattcc	32820
aagaaagtcc ccccggggtc ctgtctctga acctggccga gcccctggtc acttcccacg	32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgctggc aacctcacct	32940
ctcaagatat caccaccgct agccctcccc tcaaaaaaac caagaccaac ctcagcctag	33000
aaacctcatc ccccctaact gtgagcacct caggcgccct caccgtagca gccgcgcctc	33060
ccctggcggt ggccggcacc tccctcacca tgcaatcaga ggccccctg acagtacagg	33120
atgcaaaact caccctggcc accaaaggcc cctgaccgt gtctgaaggc aaactggcct	33180
tgcaaacatc ggccccgtg acggccgctg acagcagcac cctcacagtc agtgccacac	33240
cacccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc atttacacca	33300

[0035]

ccaatggaaa	actaggactt	aactttggcg	ctccccctgca	tgtggtagac	agcctaaatg	33360
cactgactgt	agttactggc	caaggtctta	cgataaacgg	aacagcccta	caaactagag	33420
tctcaggtgc	cctcaactat	gacacatcag	gaaacctaga	attgagagct	gcagggggta	33480
tgcgagttga	tgcaaatggt	caacttatcc	ttgatgtagc	ttacccatit	gatgcacaaa	33540
acaatctcag	ccttaggctt	ggacagggac	ccctgtttgt	taactctgcc	cacaacttgg	33600
atgttaacta	caacagaggc	ctctacctgt	tcacatctgg	aaataccaaa	aagctagaag	33660
ttaatatcaa	aacagccaag	ggtctcatit	atgatgacac	tgctatagca	atcaatgcgg	33720
gtgatgggct	acagtttgac	tcaggctcag	atacaaatcc	attaaaaact	aaacttggat	33780
taggactgga	ttatgactcc	agcagagcca	taattgctaa	actgggaact	ggcctaagct	33840
ttgacaacac	aggtgccatc	acagtaggca	acaaaaatga	tgacaagctt	accttgtgga	33900
ccacaccaga	cccatcccct	aactgtagaa	tctattcaga	gaaagatgct	aaattcacac	33960
ttgttttgac	taaatgcggc	agtcaggtgt	tggccagcgt	ttctgtttta	tctgtaaaag	34020
gtagccttgc	gcccacagtc	ggcacagtaa	ctagtgtcga	gattgtcttc	agatttgatg	34080
aaaatggagt	tctactaagc	aattcttccc	ttgacctca	atactggaac	tacagaaaag	34140
gtgaccttac	agagggcact	gcatatacca	acgcagtggg	atttatgccc	aacctcacag	34200
catacccaaa	aacacagagc	caaactgcta	aaagcaacat	tgtaagtcag	gtttacttga	34260
atggggacaa	atccaaaccc	atgacctca	ccattaccct	caatggaact	aatgaaacag	34320
gagatgccac	agtaagcact	tactccatgt	cattctcatg	gaactggaat	ggaagtaatt	34380
acattaatga	aacgttccaa	accaactcct	tcacctctc	ctacatcgcc	caagaataaa	34440
aagcatgacg	ctgttgattt	gattcaatgt	gtttctgttt	tattttcaag	cacaacaaaa	34500
tcattcaagt	cattcttcca	tcttagctta	atagacacag	tagcttaata	gaccagtag	34560
tgcaaagccc	cattctagct	tataactagt	ggagaagtac	tcgcctacat	gggggtagag	34620
tcataatcgt	gcatcaggat	agggcggtgg	tgctgcagca	gcgcgcgaat	aaactgctgc	34680
cgcgcgcgt	ccgtcctgca	ggaatacaac	atggcagtgg	tctcctcagc	gatgattcgc	34740
accgcccgc	gcataaggcg	ccttgctctc	cgggcacagc	agcgcaccct	gatctcactt	34800
aaatcagcac	agtaactgca	gcacagcacc	acaatatgtt	tcaaaatccc	acagtgaag	34860
ggcgtgtatc	caaagctcat	ggcggggacc	acagaaccca	cgtggccatc	ataccacaag	34920
cgcaggtaga	ttaagtggcg	acccctcata	aacacgtgg	acataaacat	tacctttttt	34980
ggcatgttgt	aattcaccac	ctcccgttac	catataaacc	tctgattaaa	catggcgcca	35040

[0036]

tccaccacca	tcctaaacca	gctggccaaa	acctgcccgc	cggtataaca	ctgcaggga	35100
ccgggactgg	aacaatgaca	gtggagagcc	caggactcgt	aaccatggat	catcatgctc	35160
gtcatgatat	caatgttggc	acaacacagg	cacacgtgca	tacacttcct	caggattaca	35220
agctcctccc	gcgttagaac	catatcccag	ggaacaaccc	attcctgaat	cagcgtaa	35280
cccacactgc	agggaagacc	tcgcacgtaa	ctcacgttgt	gcattgtcaa	agtgttacat	35340
tcgggcagca	gcggatgata	ctccagtatg	gtagcgcggg	tttctgtctc	aaaaggaggt	35400
agacgatccc	tactgtacgg	agtgcgccga	gacaaccgag	atcgtgttgg	tcgtagtgtc	35460
atgccaaatg	gaacgccgga	cgtagtcata	tttctgaag	tcttagatct	ctcaacgcag	35520
caccagcacc	aacacttcgc	agtgtaaaag	gccaaagtgc	gagagagtat	atataggaat	35580
aaaaagtgac	gtaaacgggc	aaagtccaaa	aaacgccccag	aaaaaccgca	cgcgaacct	35640
cgccccgaaa	cgaaagccaa	aaaacactag	acactccctt	ccggcgtcaa	cttcgccttt	35700
cccacgttac	gtcacttgcc	ccagtcaaac	aaactacata	tcccgaactt	ccaagtcgcc	35760
acgccccaaa	caccgcctac	acctccccgc	ccgccggccc	gcccccaaac	ccgcctcccg	35820
ccccgcgccc	cgccccgcgc	cgcccatctc	attatcatat	tggcttcaat	ccaaaataag	35880
gtatattatt	gatgatggtt	taaacggatc	caattcttga	agacgaaagg	gcctcgtgat	35940
acgcctat	ttatagggtt	atgtcatgat	aataatggtt	tcttagacgt	caggtggcac	36000
ttttcgggga	aatgtgcgcg	gaaccctat	ttgtttat	ttctaaatac	attcaaata	36060
gtatccgctc	atgagacaat	aaccctgata	aatgcttcaa	taatattgaa	aaaggaagag	36120
tatgagtatt	caacatttcc	gtgtgcacct	tattcccttt	tttgcggcat	tttgccttcc	36180
tgtttttgct	caccagaaa	cgctgggtgaa	agtaaaagat	gctgaagatc	agttgggtgc	36240
acgagtgggt	tacatcgaa	tggatctcaa	cagcggtaag	atccttgaga	gttttcgccc	36300
cgaagaacgt	tttccaatga	tgagcacttt	taaagtcttg	ctatgtggcg	cggatattatc	36360
ccgtgttgac	gccgggcaag	agcaactcgg	tcgccgcata	cactattctc	agaatgactt	36420
ggttgagtac	tcaccagtca	cagaaaagca	tcttacggat	ggcatgacag	taagagaatt	36480
atgcagtgtc	gccataacca	tgagtataaa	cactgcggcc	aacttacttc	tgacaacgat	36540
cggaggaccg	aaggagctaa	ccgttttttt	gcacaacatg	ggggatcatg	taactcgcct	36600
tgatcgttgg	gaaccggagc	tgaatgaagc	cataccaaac	gacgagcgtg	acaccacgat	36660
gcctgtagca	atggcaacaa	cgttgcgcaa	actattaact	ggcgaactac	ttactctagc	36720

[0037]

ttccccgcaa caattaatag actggatgga ggcggataaa gttgcaggac cacttctgcg	36780
ctcggccctt cggctggct ggtttattgc tgataaatct ggagccggtg agcgtgggtc	36840
tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg tagttatcta	36900
cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc	36960
ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac tttagattga	37020
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaata cccttaacgt	37080
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat	37140
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg	37200
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggttaactgg cttcagcaga	37260
gcgcagatac caaatactgt ctttctagt tagccgtagt taggccacca cttcaagaac	37320
tctgtagcac cgcctacata ctcgctctg ctaatcctgt taccagtggc tgctgccagt	37380
ggcgataagt cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgcag	37440
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc	37500
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaaag	37560
gcggacaggt atccggtgag cggcagggtc ggaacaggag agcgcacgag ggagcttcca	37620
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt	37680
cgatttttgt gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc	37740
tttttacggt tcctggcctt ttgctggcct tgaagctgtc cctgatggc gtcacttacc	37800
tgccctggaca gcatggcctg caacgcgggc atcccgatgc cgccggaagc gagaagaatc	37860
ataatgggga aggccatcca gcctcgcgtc gcagatccga attcgtttaa ac	37912

<210> 8

<211> 43428

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 8

catcatcaat aatatacctt attttgatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgccggggcg gaggcgggtt tggggcggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggt gtgacttgct agtgccgggc gcggtaaaag	180

[0038]

tgacgttttc	cgtgcgcgac	aacgcccccg	ggaagtgaca	tttttccgcg	ggtttttacc	240
ggatgttgta	gtgaatttgg	gcgtaaccaa	gtaagatttg	gccattttcg	cgggaaaact	300
gaaacgggga	agtgaatctt	gattaatttt	gcgttagtca	taccgcgtaa	tatttgtcta	360
gggccgaggg	actttggccg	attacgtgga	ggactcgccc	aggtgttttt	tgaggtgaat	420
ttccgcgttc	cgggtcaaag	tctgcgtttt	attattatag	gatatcccat	tgcatacgtt	480
gtatccatat	cataatatgt	acatttatat	tggctcatgt	ccaacattac	cgccatgttg	540
acattgatta	ttgactagtt	attaatagta	atcaattacg	gggtcattag	ttcatagccc	600
atatatggag	ttccgcgtta	cataacttac	ggtaaattgg	ccgcctggct	gaccgcccac	660
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	720
tttccattga	cgtcaatggg	tggagtattd	acggtaaact	gcccacttgg	cagtacatca	780
agtgtatcat	atgccaagta	cgccccctat	tgacgtcaat	gacggtaaat	ggccccctg	840
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	900
agtcategct	attaccatgg	tgatgcggtt	ttggcagtac	atcaatgggc	gtggatagcg	960
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	1020
gcacccaaat	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaat	1080
gggcggtagg	cgtgtacggt	gggaggctta	tataagcaga	gctctcccta	tcagtgatag	1140
agatctccct	atcagtgata	gagatcgctg	acgagctcgt	ttagtgaacc	gtcagatcgc	1200
ctggagacgc	catccacgct	gttttgacct	ccatagaaga	caccgggacc	gatccagcct	1260
ccgcggccgg	gaacggtgca	ttggaacgcg	gattccccgt	gccaagagtg	agatcttccg	1320
tttatctagg	taccgggccc	cccctcgagg	tcgacggtat	cgataagctt	cacgctgccg	1380
caagcactca	gggcgcaagg	gctgctaaag	gaagcggaac	acgtagaaag	ccagtccgca	1440
gaaacggtgc	tgaccccgga	tgaatgtcag	ctactgggct	atctggacaa	gggaaaacgc	1500
aagcgcaaag	agaaagcagg	tagcttgcag	tgggcttaca	tggcgatagc	tagactgggc	1560
ggttttatgg	acagcaagcg	aaccggaatt	gccagctggg	gcgccctctg	gtaaggttgg	1620
gaagccctgc	aaagtaaact	ggatggcttt	cttgccgcca	aggatctgat	ggcgcagggg	1680
atcaagatct	aaccaggagc	tatttaatgg	caacagttaa	ccagctggta	cgcaaaccac	1740
gtgctcgcaa	agttgcgaaa	agcaacgtgc	ctgcgctgga	agcatgcccc	caaaaacgtg	1800
gcgtatgtac	tcgtgtatat	actaccactc	ctaaaaaacc	gaactccgcg	ctgcgtaaag	1860
tatgccgtgt	tcgtctgact	aacggtttct	aagtgacttc	ctacatcggt	ggtgaaggtc	1920

[0039]

acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgttaaa gacctcccgg	1980
gtgttcgtta ccacaccgta cgtgggtgcgc ttgactgctc cggcgttaaa gaccgtaagc	2040
aggctcgttc caagtatggc gtgaagcgtc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcgttttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat	2220
gccgccgtgt tccggctgtc agcgcagggg cgcgcggttc tttttgtcaa gaccgacctg	2280
tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tategtggct ggccacgacg	2340
ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcactcacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcggcggctg catacgcttg atccggctac ctgcccattc	2520
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	2580
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctctcgtga cccatggcga tgcctgcttg	2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc	2880
atgccttct atgccttct tgacgagttc ttctgagcgg gactctgggg ttcgaaatga	2940
ccgaccaagc gacgcccac ctgccatcac gagatttcga ttccaccgcc gccttctatg	3000
aaaggttggg cttcggaatc gttttccggg acgccggctg gatgatcctc cagcgcgggg	3060
atctcatgct ggagtcttc gccaccccc ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggcccgat ctgctgtgcc ttctagtgc cagccatctg ttgtttgcc	3180
ctccccctg ccttccttga ccttgaagg tgccactccc actgtccttt cctaataaaa	3240
tgaggaaatt gcatgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg	3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg	3360
ctctatggcc gatcagcgt cgtgaggtg ggtgagtggg cgtggcctgg ggtggtcatg	3420
aaaatatata agttgggggt cttagggtct ctttatttgt gttgcagaga ccgccggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgc ttggatggca gcatcgtgag	3540
cccttatttg acgacgcgga tgccccactg ggccgggggtg cgtcagaatg tgatgggctc	3600

[0040]

cagcatcgac ggccgacccg tcctgcccgc aaattccgcc acgctgacct atgcgaccgt	3660
cgcggggacg ccgttggacg ccaccgccgc cgccgccgcc accgcagccg cctcggccgt	3720
gcgcagcctg gccacggact ttgcattcct gggaccactg gcgacagggg ctacttctcg	3780
ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggtctc	3900
ctccctgcaa gctggcggga atgcttctcc cacaaatgcc gttaaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttatatt cataattttc	4020
cgcgcgcat aggccctaga ccagegttct cggtcgttga gggcgcggtg tatcttctcc	4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccggggg	4140
tggaggtagc accactgcag agcttcatgc tccggggtgg tgtttagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tgggtgaagt gtttacaaaa cggttaagtt gggaagggtg cattcgggga	4320
gagatgatgt gcatcttga ctgtatTTTT agattggcga tgtttccgcc cagatccctt	4380
ctgggattca tgttgtgcag gaccaccagt acagtgtatc cggcgcactt ggggaatttg	4440
tcatgcagct tagagggaag agcgtggaag aacttggaga cgcctttgtg gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag	4560
atatttctgg ggtcgtgac gtcgtagtgt tgttccagggt tagagtcgtc ataggccatt	4620
tttacaagc gcggggcggag ggtgcccgc tgggggatga tggccccctc tgcccctggg	4680
gcgtagtgtc cctcgcagat ctgcatttcc caggccttaa tctcggagggt gggaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag	4800
agcaggtttc taagcagctg tgattttcca caaccggtgg gcccataaat aacacctata	4860
accggttgca gctggtagtt tagagagctg cagctgccgt cgtcccggag gaggggggcc	4920
acctcgttga gcatgtccct gacgcgcagt ttctccccga ccagatccgc cagaaggcgc	4980
tcgccgccca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg	5040
tccgccgtgg gcatgttttt cagggtctgg ctcagcagct ccaggcggtc ccagagctcg	5100
gtgacgtgct ctacggcatc tctatccagc atatctctc gtttcgcggg ttggggcgac	5160
tttcgctgta gggcaccaag cgggtggtcgt ccagcggggc cagagtcatg tccttccatg	5220
ggcgcagggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgct ccgggctgag	5280
cgcttgccaa ggtgcgcttg aggctggttc tgctggtgct gaagcgtgc cggctcttgc	5340

[0041]

cctgcgcgtc ggccaggtag catttgacca tgggtgtcata gtccagcccc tccgcggcgt	5400
gtcccttggc gcgcagcttg cccttggagg tggcgccgca cgaggggcag agcaggctct	5460
tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgccgc	5520
agacccccga cacggtctcg cactccacca gccaggtgag ctcggggcgc gccgggtcaa	5580
aaaccaggtt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgaggtggt	5640
gtccccgctc ggtgacgaag aggcgtgtccg tgtctccgta gaccgacttg aggggtcttt	5700
tctccagggg ggtccctcgg tcttctctgt agaggaactc ggaccactct gagacgaagg	5760
cccgcgtcca ggccaggacg aaggaggcta tgtgggaggg gtagcggtcg ttgtccacta	5820
gggggtccac cttctccaag gtgtgaagac acatgtcgcc ttctcggcg tccaggaagg	5880
tgattggctt gtaggtgtag gccacgtgac cgggggttcc tgacgggggg gtataaaagg	5940
gggtgggggc gcgctcgtcg tctctctctt ccgcctcgt gtctgcgagg gccagctgct	6000
ggggtgagta ttccctctcg aaggcgggca tgacctcgc gctgaggttg tcagtttcca	6060
aaaacgagga ggatttgatg ttcacctgtc ccgaggtgat acctttgagg gtacccgcgt	6120
ccatctggtc agaaaacacg atctttttat tgtccagctt ggtggcgaac gaccgtaga	6180
gggcgttggg gagcagcttg gcgatggagc gcagggtctg gttcttgtcc ctgtcggcgc	6240
gctccttggc cgcgatgttg agctgcacgt actcgcgcgc gacgcagcgc cactcgggga	6300
agacggtggt gcgctcgtcg ggcaccaggc gcacgcgcca gccgcggttg tgcagggtga	6360
ccaggtccac gctggtggcg acctcgccgc gcaggcgctc gttggtccag cagagacggc	6420
cgcccttgcg cgagcagaag gggggcaggg ggtcgagctg ggtctcgtcc ggggggtccg	6480
cgtccacggt gaaaaccccg gggcgcaggc gcgcgtcgaa gtagtctatc ttgcaacctt	6540
gcatgtccag cgcctgtgc cagtcgcggg cggcgagcgc gcgctcgtag gggttgagcg	6600
gcgggccccca gggcatgggg tgggtgagtg cggaggcgta catgccgcag atgtcataga	6660
cgtagagggg ctcccgcagg accccgatgt aggtggggta gcagcggccg ccgcggatgc	6720
tggcgcgcac gtagtcatac agctcgtgcg agggggcgag gaggtcgggg ccaggttgg	6780
tgcgggcggg gcgctccgcg cggaagacga tctgcctgaa gatggcatgc gagttggaag	6840
agatggtggg gcgctggaag acgttgaagc tggcgtcctg caggccgacg gcgtcgcgca	6900
cgaaggaggc gtaggagtcg cgcagcttgt gtaccagctc ggcggtgacc tgcacgtcga	6960
gcgcgcagta gtcgagggtc tcgcggatga tgtcatatctt agcctgcccc ttctttttcc	7020

[0042]

acagctcgcg gttgaggaca aactcttcgc ggtctttcca gtactcttgg atcgggaaac	7080
cgtccggttc cgaacggtaa gaggctagca tgtagaactg gttgacggcc tggtaggcgc	7140
agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg	7200
tcagggcgaa ggtgtccctg accatgactt tgaggtactg gtgcttgaag tcggagtcgt	7260
cgcagccgcc ccgctcccag agcgagaagt cgggtgcgctt cttggagcgg gggttgggca	7320
gagcgaaggt gacatcgttg aagaggattt tgcccgcgcg gggcatgaag ttgcgggtga	7380
tgcggaaggg ccccgccact tcagagcggg tgttgatgac ctgggcggcg agcacgatct	7440
cgtcgaagcc gttgatgttg tggccacga tgtagagttc caggaaagcg ggccggccct	7500
ttacgggtggg cagcttcttt agctcttcgt aggtgagctc ctcgggcgag gcgaggccgt	7560
gctcgccag ggcccagtc gcgaggtgcg ggttgtctct gaggaaggac ttccagaggt	7620
cgcgggccag gagggtctgc aggcggtctc tgaaggtcct gaactggcgg cccacggcca	7680
ttttttcggg ggtgatgcag tagaaggtga gggggtcttg ctgccagcgg tcccagtcga	7740
gctgcagggc gaggtcgcgc gcggcggtga ccaggcgtc gtcgccccg aatttcatga	7800
ccagcatgaa gggcacgagc tgctttccga aggccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaag aactggatct	7920
cccgccacca gttggaggag tggctgttga tgtggtggaa gtagaagtcc cgtcgccggg	7980
ccgaacactc gtgctggctt ttgtaaaagc gagcgagta ctggcagcgc tgcacgggct	8040
gtacctcatg cacgagatgc acctttcgcc cgcgcagag gaagccgagg ggaaatctga	8100
gcccccgcc tggtcgcgg catggctggt tctcttctac tttggatgcg tgtccgtctc	8160
cgtctggctc ctcgaggggt gttacgggtg agcggaccac cacgccgcgc gagccgcagg	8220
tccagatata ggcgcgcgcc ggtcggagtt tgatgacgac atcgcgagc tgggagctgt	8280
ccatggtctg gagtccccgc ggcgggcgca ggtcagccgg gagttcttgc aggttcacct	8340
cgcagagtcg ggccaggcg cggggcaggt ctaggtggta cctgatctct aggggcgtgt	8400
tggtggcggc gtcgatggct tgcaggagcc cgcagccccg gggggcgacg acggtgcccc	8460
gcggggtggt ggtggtggtg gcggtgcagc tcagaagcgg tgccgcgggc gggcccccg	8520
aggtaggggg ggctccggtc ccgcgggcag gggcggcagc ggcacgtcgg cgtggagcgc	8580
gggcaggagt tgggtctgtg cccggagggt gctggcgaag gcgacgagc ggcggttgat	8640
ctcctggatc tggcgctct gcgtgaagac gacgggcccc gtgagcttga acctgaaaga	8700
gagttcgaca gaatcaatct cgggtgtcatt gaccgcggcc tggcgagga tctcctgcac	8760

[0043]

gtctccccgag ttgtcttggg aggcgatctc ggccatgaac tgctcgatct cttcctcctg	8820
gaggtctccg cgtccggcgc gttccacggg ggccgccagg tcgttgagaga tgcgccccat	8880
gagctgcgag aaggcgttga gtccgccctc gttccagact cggctgtaga ccacgcccc	8940
ctggatcatcg cgggcgcgca tgaccacctg cgcgagggtg agctccacgt gccgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggtggcgg tgtgctcggc	9060
cacgaagaag ttcatgacct agcggcgcaa cgtggattcg ttgatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggccaagtgg aaaaactggg agttgcgcgc	9180
cgcacgggtc aactcctcct ccagaagacg gatgagctcg gcgacggtgt cgcgcacctc	9240
gcgctcgaag gctatgggga tctcttcctc cgctagcctc accacctcct cctcttcctc	9300
ctcttctggc acttccatga tggcttcctc ctcttcgggg ggtggcggcg gcggcggtgg	9360
gggagggggc gctctgcgcc ggccggcgcg caccgggagg cggtcacga agcgcgcgat	9420
catctccccg cggcgcgcg gcattggtctc ggtgacggcg cggccgttct cccggggcg	9480
cagttggaag acgccgccgg acatctggtg ctggggcggg tggccgtgag gcagcgagac	9540
ggcgctgacg atgcatctca acaattgctg cgtaggtacg ccgccgaggg acctgaggga	9600
gtccatatac accggatccg aaaaccttcc gaggaaggcg tctaaccagt cgcagtcgca	9660
aggtaggctg agcacctgg cgggcgcgcg ggggtggggg gagtgtctgg cggaggtgct	9720
gctgatgatg taattgaagt aggcggactt gacacggcgg atggtcgaca ggagcaccat	9780
gtccttgggt ccggcctgct ggatgcggag gcggtcggct atgcccagg cttcgttctg	9840
gcacggcgc aggtccttgt agtagtcttg catgagcctt tccaccggca cctcttctcc	9900
ttcctcttct gcttcttcca tgtctgttcc ggccctgggg cggcgccgcg cccccctgcc	9960
ccccatgcgc gtgacccccga accccctgag cggttggagc agggccagggt cggcgacgac	10020
gcgctcggcc aggatggcct gctgcacctg cgtgagggtg gtttggaagt catccaagtc	10080
cacgaagcgg tggtaggcgc ccgtgttgat ggtgtagggt cagttggcca tgacggacca	10140
gttgacggtc tggtagcccg gttgcacat ctcggtgtac ctgagtcgag agtaggcgcg	10200
ggagtcgaag acgtagtcgt tgcaagtccg caccagggtac tggtagccca ccaggaagtg	10260
cggcgcgcg gcggcgtaga ggggccagcg cagggtggcg ggggtccgg gggccaggtc	10320
ttccagcatg aggcggtggg aggcgtagat gtacctggac atccagggtga taccgcggc	10380
ggtggtggag gcgcgcggga agtcgcgcac ccggttccag atgttgcgca ggggcagaaa	10440

[0044]

gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactctagac	10500
cagggaac gaaagccggt cagcgggcac tcttcctggt tctggtgaat agatcgcaag	10560
ggtatcatgg cggagggcct cggttcgagc cccgggtccg ggccggacgg tccgccatga	10620
tccacgcggt taccgcccgc gtgtcgaacc caggtgtgcg acgtcagaca acggtggagt	10680
gttccttttg gcgtttttct ggccgggcgc cggcgccgcg taagagacta agccgcgaaa	10740
gcgaaagcag taagtggctc gctccccgta gccggaggga tccttgctaa gggttgcgtt	10800
gcggcgaacc ccggttcgaa tcccgtactc gggccggccg gacccgcggc taaggtgttg	10860
gattggcctc cccctcgtat aaagaccccg cttgcggatt gactccggac acggggacga	10920
gcccccttta tttttgcttt ccccagatgc atccggtgct gcggcagatg cccccccgc	10980
cccagcagca gcaacaacac cagcaagagc ggcagcaaca gcagcgggag tcatgcaggg	11040
ccccctcacc caccctcggc gggccggcca cctcggcgtc cgcggccgtg tctggcgcct	11100
gcggcggcgg cggggggccg gctgacgacc ccgaggagcc cccgcggcgc agggccagac	11160
actacctgga cctggaggag ggcgagggcc tggcgcggtt gggggcgccg tctcccagc	11220
gccacccgcg ggtgcagctg aagcgcgact cgcgcgaggc gtacgtgcct cggcagaacc	11280
tgttcaggga ccgcgcgggc gaggagcccg aggagatgcg ggacaggagg ttcagcgcag	11340
ggcgggagct gcggcagggg ctgaaccgcg agcggctgct gcgcgaggag gactttgagc	11400
ccgacgcgcg gacggggatc agccccgcgc gcgcgcacgt ggcggccgcc gacctggtga	11460
cggcgtaaga gcagacggtg aaccaggaga tcaacttcca aaagagtttc aacaaccacg	11520
tgcgcacgct ggtggcgcgc gaggaggtga ccatcgggct gatgcacctg tgggactttg	11580
taagcgcgct ggtgcagaac cccaacagca agcctctgac ggcgcagctg ttcctgatag	11640
tgcagcacag cagggaacac gaggcgttta gggacgcgct gctgaacatc accgagcccc	11700
agggtcggtg gctgctggac ctgattaaca tcctgcagag catagtgggtg caggagcgca	11760
gcctgagcct ggccgacaag gtggcggcca tcaactactc gatgctgagc ctgggcaagt	11820
tttacgcgcg caagatctac cagacgccgt acgtgcccat agacaaggag gtgaagatcg	11880
acggttttta catgcgcatg gcgctgaagg tgctcaccct gagcgacgac ctgggcgtgt	11940
accgcaacga gcgcatccac aaggccgtga gcgtgagccg gcggcgcgag ctgagcgacc	12000
gcgagctgat gcacagcctg cagcgggcgc tggcgggcgc cggcagcggc gacagggagg	12060
cggagtcccta cttcgatgcg ggggcggacc tgcgctgggc gccagccgg cgggccctgg	12120
aggccgcggg ggtccgcgag gactatgacg aggacggcga ggaggatgag gagtacgagc	12180

[0045]

tagaggaggg cgagtacctg gactaaaccg cgggtggtgt ttccggtaga tgcaagaccc	12240
gaacgtggtg gacccggcgc tgcgggcggc tctgcagagc cagccgtccg gccttaactc	12300
ctcagacgac tggcgacagg tcatggaccg catcatgtcg ctgacggcgc gtaacccgga	12360
cgcgttccgg cagcagccgc aggccaacag gctctccgcc atcctggagg cgggtggtgcc	12420
tgcgcgctcg aacccccacgc acgagaaggt gctggccata gtgaacgcgc tggccgagaa	12480
cagggccatc cccccggacg aggcggggct ggtgtacgac gcgctgctgc agcgcgtggc	12540
ccgctacaac agcggcaacg tgcagaccaa cctggaccgg ctggtggggg acgtgcgcga	12600
ggcgggtggcg cagcgcgagc gcgcggatcg gcagggcaac ctgggctcca tgggtggcgct	12660
gaatgccttc ctgagcacgc agccggccaa cgtgccgcgg gggcaggaag actacaccaa	12720
ctttgtgagc gcgctgcggc tgatggtgac cgagaccccc cagagcgagg tgtaccagtc	12780
gggcccggac tactttcttc agaccagcag acagggcctg cagacggtga acctgagcca	12840
ggctttcaag aacctgcggg ggctgtgggg cgtgaaggcg cccaccggcg accgggcgac	12900
ggtgtccagc ctgctgacgc ccaactcgcg cctgctgctg ctgctgatcg cgccgttcac	12960
ggacagcggc agcgtgtccc gggacaccta cctggggcac ctgctgacct tgtaccgca	13020
ggccatcggg caggcgcagg tggacgagca caccttcag gagatcacca gcgtgagccg	13080
cgcgctgggg caggaggaca cgagcagcct ggaggcgact ctgaactacc tgctgaccaa	13140
ccggcggcag aagattccct cgctgcacag cctgacctcc gaggaggagc gcatcttgcg	13200
ctacgtgcag cagagcgtga gcctgaacct gatgcgcgac ggggtgacgc ccagcgtggc	13260
gctggacatg accgcgcgca acatggaacc gggcatgtac gccgcgcacc ggccttacat	13320
caaccgcctg atggactacc tgcatcgcgc ggcggccgtg aaccccgagt actttaccaa	13380
cgccatcctg aacccgcact ggctcccgcc gcccgggttc tacagcgggg gcttcgaggt	13440
cccggagacc aacgatggct tcctgtggga cgacatggac gacagcgtgt tctccccgcg	13500
gccgcaggcg ctggcggaag cgtccctgct gcgtcccaag aaggaggagg aggaggaggc	13560
gagtcgccgc cgcggcagca gcggcgtggc ttctctgtcc gagctggggg cggcagccgc	13620
cgcgcgcccc gggctccctgg gcggcagccc ctttccgagc ctggtgggggt ctctgcacag	13680
cgagcgcacc acccgccctc ggctgctggg cgaggacgag tacctgaata actccctgct	13740
gcagccggtg cgggagaaaa acctgcctcc cgccttcccc aacaacggga tagagagcct	13800
ggtggacaag atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcctgcgct	13860

[0046]

ccggccgccc	acgcggcgcc	agcgccacga	ccggcagcgg	gggctggtgt	gggatgacga	13920
ggactccgcg	gacgatagca	gcgtgctgga	cctgggaggg	agcggcaacc	cgttcgcgca	13980
cctgcgcccc	cgcctgggga	ggatgtttta	aaaaaaaaa	aaaaaagcaa	gaagcatgat	14040
gcaaaaatta	aataaaactc	accaaggcca	tggcgaccga	gcgttggttt	cttgtgttcc	14100
cttcagtatg	cggcgcgcg	cgatgtacca	ggagggacct	cctccctctt	acgagagcgt	14160
ggtgggcgcg	gcggcgcg	cgcctcttc	tccctttg	tcgcagctgc	tggagccgcc	14220
gtacgtgcct	ccgcgctacc	tgcggcctac	gggggggaga	aacagcatcc	gttactcgga	14280
gctggcgccc	ctgttcgaca	ccacccgggt	gtacctggtg	gacaacaagt	cggcggacgt	14340
ggcctccctg	aactaccaga	acgaccacag	caatTTTTT	accacggtca	tccagaacaa	14400
tgactacagc	ccgagcgagg	ccagcaccca	gaccatcaat	ctggatgacc	ggtcgcactg	14460
gggcggcgac	ctgaaaacca	tcctgcacac	caacatgccc	aacgtgaacg	agttcatgtt	14520
caccaataag	ttcaaggcgc	gggtgatggt	gtcgcgctcg	cacaccaagg	aagaccgggt	14580
ggagctgaag	tacgagtggg	tggagttcga	gctgccagag	ggcaactact	ccgagaccat	14640
gaccattgac	ctgatgaaca	acgcgatcgt	ggagcactat	ctgaaagtgg	gcaggcagaa	14700
cggggtcctg	gagagcgaca	tcgggggtcaa	gttcgacacc	aggaacttcc	gcctggggct	14760
ggaccccgtg	accgggctgg	ttatgcccg	ggtgtacacc	aacgaggcct	tccatcccga	14820
catcatectg	ctgcccggct	gcgggggtgga	cttcacttac	agccgcctga	gcaacctcct	14880
gggcatccgc	aagcggcagc	ccttcagga	gggcttcagg	atcacctacg	aggacctgga	14940
ggggggcaac	atccccgcgc	tcctcgatgt	ggaggcctac	caggatagct	tgaaggaaaa	15000
tgaggcgagg	caggaggata	ccgccccgc	cgcctccgcc	gccgccgagc	agggcgagga	15060
tgctgctgac	accgcggccg	cggacggggc	agaggccgac	cccgtatgg	tggtggaggc	15120
tcccagagcag	gaggaggaca	tgaatgacag	tgcggtgcgc	ggagacacct	tcgtcacccg	15180
gggggaggaa	aagcaagcgg	aggccgaggc	cgcggccgag	gaaaagcaac	tggcggcagc	15240
agcggcgggc	gcggcggttg	ccgcggcgga	ggctgagtct	gaggggacca	agcccgccaa	15300
ggagcccgtg	attaagcccc	tgaccgaaga	tagcaagaag	cgcagttaca	acctgtctaa	15360
ggacagcacc	aacaccgcgt	accgcagctg	gtacctggcc	tacaactacg	gcgacccgtc	15420
gacgggggtg	cgtccttgga	ccctgctgtg	cacgccggac	gtgacctgeg	gctcgagaca	15480
ggtgtactgg	tcgtgccccg	acatgatgca	agaccccgtg	accttccgct	ccacgcggca	15540
ggtcagcaac	ttcccgttgg	tgggcgccga	gctgctgccc	gtgcactcca	agagcttcta	15600

[0047]

caacgaccag gccgtctact cccagctcat ccgccagttc acctctctga cccacgtgtt	15660
caatcgcttt cctgagaacc agattctggc gcgcccgcgc gccccacca tcaccaccgt	15720
cagtgaaaac gttcctgctc tcacagatca cgggacgcta ccgctgcgca acagcatcgg	15780
aggagtccag cgagtgaccg ttactgacgc cagacgccgc acctgcccct acgtttacaa	15840
ggccttgggc atagtctcgc cgcgcgtcct ttccagccgc actttttgag caacaccacc	15900
atcatgtcca tcttgatctc acccagcaat aactccggct ggggactgct gcgcgcgcgc	15960
agcaagatgt tcggaggggc gaggaagcgt tccgagcagc accccgtgcg cgtgcgcggg	16020
cacttccgcg ccccttgggg agcgcacaaa cgcggccgcg cggggcgcac caccgtggac	16080
gacgccatcg actcgggtgt ggagcaggcg cgcaactaca ggcccgcggt ctctaccgtg	16140
gacgcggcca tccagaccgt ggtgcggggc gcgcggcggt acgccaagct gaagagccgc	16200
cggaagcgcg tggcccgcgc ccaccgccgc cgaccggggg ccgccgcaa acgcgcgcgc	16260
gcggccctgc ttcgccgggc caagcgcacg ggccgccgcg ccgccatgag ggccgcgcgc	16320
cgcttggccg ccggcatcac cgccgccacc atggcccccc gtaccgaag acgcgcggcc	16380
gccgccgcgc ccgccgcat cagtgcacat gccagcaggc gccggggcaa cgtgtactgg	16440
gtgcgcgact cggtgaccgg cacgcgcgtg cccgtgcgt tccgcccccc gcggacttga	16500
gatgatgtga aaaaacaaca ctgagtctcc tgctgttgtg tgtatcccag cggcggcggc	16560
gcgcgcagcg tcatgtccaa gcgcaaaatc aaagaagaga tgctccaggt cgtcgcgcgc	16620
gagatctatg ggccccgaa gaaggaagag caggattcga agccccgaa gataaagcgg	16680
gtcaaaaaga aaaagaaaga tgatgacgat gccgatgggg aggtggagtt cctgcgcgcc	16740
acggcgccca ggcgcccggg gcagtggaag ggccggcgcg taaagcgcgt cctgcgcccc	16800
ggcaccgcgg tggctttcac gcccggcgag cgctccaccc ggactttcaa gcgcgtctat	16860
gacgaggtgt acggcgacga agacctgctg gagcaggcca acgagcgctt cggagagttt	16920
gcttacggga agcgtcagcg ggcgctgggg aaggaggacc tgctggcgct gccgctggac	16980
cagggaacc ccacccccag tctgaagccc gtgaccctgc agcaggtgct gccgagcagc	17040
gcaccctccg aggcgaagcg gggctctgaag cgcgagggcg gcgacctggc gccaccgtg	17100
cagctcatgg tgcccaagcg gcagaggctg gaggatgtgc tggagaaaat gaaagtagac	17160
cccgtctgc agccggacat cagggtccgc cccatcaagc aggtggcgcc gggcctcggc	17220
gtgcagaccg tggacgtggt cateccccacc ggcaactccc ccgcgcgcgc caccactacc	17280

[0048]

gctgcctcca cggacatgga gacacagacc gatccccccg cagccgcagc cgcagccgcc	17340
gccgcgacct cctcggcgga ggtgcagacg gacccctggc tgccgccggc gatgtcagct	17400
ccccgcgcgc gtcgcgggcg caggaagtac ggcgccgcca acgcgctcct gcccagtac	17460
gccttgcatc cttccatcgc gcccaccccc ggctaccgag gctataccta ccgcccgcga	17520
agagccaagg gttccacccg ccgtccccgc cgacgcgccg ccgccaccac ccgccgccgc	17580
cgccgcagac gccagccccgc actggctcca gtctccgtga ggaaagtggc gcgcgacgga	17640
cacaccctgg tgctgcccag ggcgcgctac caccacagca tcgtttaaaa gcctgttgtg	17700
gttcttgtag atattgccct cacttgccgc ctccgtttcc cggtgccggg ataccgagga	17760
ggaagatcgc gccgcaggag ggggtctggc ggccgcggcc tgagcggagg cagccgccgc	17820
gcgcaccggc ggcgacgcgc caccagccga cgcattgcgc gcggggtgct gcccctgtta	17880
atccccctga tcgccgcggc gatcggcgcc gtgcccggga tcgcctccgt ggccttgcaa	17940
gcgtcccaga ggcatgaca gacttgcaaa cttgcaaata tggaaaaaaa aacccaata	18000
aaaaagtcta gactctcacg ctgccttggt cctgtgacta tttttagtaa tggaagacat	18060
caactttgcg tcgttgccc cgcgtcacgg ctgcgccccg ttcttgggac actggaacga	18120
tatcggcacc agcaacatga gcggtggcgc cttcagttgg ggctctctgt ggagcggcat	18180
taaaagtatc gggctctgcc ttaaaaatta cggctcccgg gccttgaaca gcagcacggg	18240
ccagatgttg agagacaagt tgaaagagca gaattccag cagaaggtagg tggagggcct	18300
ggcctccggc atcaacgggg tggtggacct ggccaaccag gccgtgcaga ataagatcaa	18360
cagcagactg gacccccggc cgccggtgga ggaggtgccg ccggcgctgg agacggtgtc	18420
ccccgatggg cgtggcgaga agcgccccgc gccgatagg gaagagacca ctctggtcac	18480
gcagaccgat gagccgcccc cgtatgagga ggccctgaag caaggtctgc ccaccacgcg	18540
gcccacgcgc cccatggcca ccggggtggt gggccgccac accccgcca cgctggactt	18600
gcctccgccc gccgatgtgc cgcagcagca gaaggcgga cagccgggccc cgccgcgac	18660
cgctcccggt tcctccgccc gtctcttgcg ccgcgcggcc agcggccccc gcgggggggt	18720
cgcgaggcac ggcaactggc agagcacgct gaacagcatc gtgggtcttg ggggtcggtc	18780
cgtgaagcgc cgccgatgct actgaatagc ttagctaacg tgttgatgt gtgtatgcgc	18840
cctatgtcgc cgccagagga gctgctgagt cgccgccgtt cgcgcgcccc ccaccaccgc	18900
cactccgccc ctcaagatgg cgaccccatc gatgatgccg cagtggctgt acatgcacat	18960
ctcgggccag gacgcctcgg agtacctgag ccccgggctg gtgcagttcg cccgcgccac	19020

[0049]

cgagagctac	ttcagcctga	gtaacaagtt	taggaacccc	acggtggcgc	ccacgcacga	19080
tgtgaccacc	gaccggtctc	agcgcctgac	gctgcggttc	attcccgtgg	accgcgagga	19140
caccgcgtac	tcgtacaagg	cgcggttcac	cctggccgtg	ggcgacaacc	gcgtgctgga	19200
catggcctcc	acctactttg	acatccgcgg	ggtgctggac	cgggggtccca	ctttcaagcc	19260
ctactctggc	accgcctaca	actccctggc	ccccaaaggc	gtccccaact	cctgcgagtg	19320
ggagcaagag	gaaactcagg	cagttgaaga	agcagcagaa	gaggaagaag	aagatgctga	19380
cgggtcaagct	gaggaagagc	aagcagctac	caaaaagact	catgtatatg	ctcaggctcc	19440
cctttctggc	gaaaaaatta	gtaaagatgg	tctgcaaata	ggaacggacg	ctacagctac	19500
agaacaaaaa	cctattttatg	cagaccctac	attccagccc	gaaccccaaa	tcggggagtc	19560
ccagtggaat	gaggcagatg	ctacagtcgc	cggcggtaga	gtgctaaaga	aatctactcc	19620
catgaaacca	tgctatggtt	cctatgcaag	accacaaaat	gctaattggag	gtcagggtgt	19680
actaacggca	aatgcccagg	gacagctaga	atctcaggtt	gaaatgcaat	tcttttcaac	19740
ttctgaaaac	gcccgtaacg	aggctaacaa	cattcagccc	aaattggtgc	tgtatagtga	19800
ggatgtgcac	atggagaccc	cggatacgca	cctttcttac	aagcccgcaa	aaagcgatga	19860
caattcaaaa	atcatgctgg	gtcagcagtc	catgcccac	agacctaat	acatcgctt	19920
cagagacaac	tttatcggcc	tcatgtatta	caatagcact	ggcaacatgg	gagtgttgc	19980
aggtcaggcc	tctcagttga	atgcagtgg	ggacttgcaa	gacagaaaca	cagaactgtc	20040
ctaccagctc	ttgcttgatt	ccatgggtga	cagaaccaga	tacttttcca	tgtggaatca	20100
ggcagtggac	agttatgacc	cagatgttag	aattattgaa	aatcatggaa	ctgaagacga	20160
gctccccaac	tattgtttcc	ctctgggtgg	cataggggta	actgacactt	accaggctgt	20220
taaaaccaac	aatggcaata	acggggggcca	ggtgacttgg	acaaaagatg	aaacttttgc	20280
agatcgcaat	gaaatagggg	tgggaaacaa	tttcgctatg	gagatcaacc	tcagtgccaa	20340
cctgtggaga	aacttcctgt	actccaacgt	ggcgtgttac	ctaccagaca	agcttaagta	20400
caaccctcc	aatgtggaca	tctctgacaa	ccccaacacc	tacgattaca	tgaacaagcg	20460
agtgggtggc	ccggggctgg	tggactgcta	catcaacctg	ggcgcgcgct	ggtcgctgga	20520
ctacatggac	aacgtcaacc	ccttcaacca	ccaccgcaat	gcgggcctgc	gtaccgctc	20580
catgctcctg	ggcaacgggc	gtacagtgcc	cttccacatc	caggtgcccc	agaagttctt	20640
tgccatcaag	aacctcctcc	tctgcccggg	ctcctacacc	tacgagtgga	acttcaggaa	20700

[0050]

ggatgtcaac atggtcctcc agagctctct gggtaacgat ctcagggtgg acggggccag	20760
catcaagttc gagagcatct gcctctacgc caccttcttc cccatggccc acaacacggc	20820
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20880
ctccgcccgc aacatgctct accccatacc cgccaacgcc accaacgtcc ccatctccat	20940
cccctcgcgc aactgggcgg ccttcgcgg ctgggccttc acccgctca agaccaagga	21000
gacccctcc ctgggctcgg gattcgaccc ctactacacc tactcgggct ccattcccta	21060
cctggacggc accttctacc tcaaccacac tttcaagaag gtctcggtca ccttcgactc	21120
ctcggtcagc tggccgggca acgaccgtct gtcaccccc aacgagttcg agatcaagcg	21180
ctcggtcgac ggggagggt acaacgtggc ccagtgcac atgaccaagg actggttcct	21240
ggtccagatg ctggccaact acaacatcgg ctaccagggc ttctacatcc cagagagcta	21300
caaggacagg atgtactcct tcttcaggaa cttccagccc atgagccggc aggtggtgga	21360
ccagaccaag tacaaggact accaggaggt gggcatcatc caccagcaca acaactcggg	21420
cttcgtgggc tacctcgccc ccaccatgcg cgagggacag gcctaccccg ccaacttccc	21480
ctatccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tcctctgca	21540
ccgcaccctc tggcgcattc ccttctccag caacttcatt tccatgggtg cgctctcgga	21600
cctgggccag aacttgctct acgccaactc cgcccacgcc ctcgacatga ccttcgaggt	21660
cgaccccatg gacgagccca cccttctcta tgttctgttc gaagtctttg acgtggtccg	21720
ggtccaccag ccgcaccgcg gcgtcatcga gaccgtgtac ctgcgtacgc ccttctcggc	21780
cggcaacgcc accacctaaa gaagcaagcc gcagtcattc ccgcctgcat gccgtcgggt	21840
tccaccgagc aagagctcag ggccatcgtc agagacctgg gatgcgggcc ctatTTTTTg	21900
ggcaccttcg acaagcgtt ccctggcttt gtctccccac acaagctggc ctgcgccatc	21960
gtcaacacgg ccggccgcga gaccgggggc gtgcaactggc tggccttcgc ctggaaccgc	22020
cgtccaaaa catgcttcct ctttgacccc ttcggtttt cggaccagcg gctcaagcaa	22080
atctacgagt tcgagtacga gggcttgctg cgtcgcagcg ccacgcctc ctgcgccgac	22140
cgtgcgtca cctcgaaaa gtccaccag accgtgcagg ggcccgaact ggccgcctgc	22200
ggtctcttct gctgcatgtt tctgcacgcc tttgtgact ggcctcagag tccatggac	22260
cgcaacccca ccatgaactt gctgacgggg gtgcccact ccatgctcca gagccccag	22320
gtcgagccca cctgcgccg caaccaggag cagctctaca gcttctgga gcgccactcg	22380
ccttacttcc gccgccacag cgcacagatc aggagggcc cctccttctg ccacttgcaa	22440

[0051]

gagatgcaag aagggttaata acgatgtaca cacttttttt ctcaataaat ggcatctttt	22500
tatttataca agctctcttg ggtattcatt tcccaccacc acccgccgtt gtcgccatct	22560
ggctctatit agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcagggac	22620
acgttgcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tcggggaagt ttctgctcca caggctgcgg gtcagcacca gcgcgttcat caggtcgggc	22740
gccgagatct tgaagtcgca gttggggccg ccgccctgcg cgcgcgagtt gcggtacacc	22800
gggttgcagc actggaacac caacagcgcc gggtgcttca cgttgccag cacgctgcgg	22860
tcggagatca gctcggcgtc caggctctcc gcgttgctca gcgcgaacgg ggtcatcttg	22920
ggcacttgcc gccccaggaa gggcgcgtgc cccggtttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccgtg cccggactcg gcgttgggggt acagcgcgcg catgaaggcc	23040
tgcacttgcc ggaaggccat ctgggccttg gcgccctccg agaagaacat gccgcaggac	23100
ttgcccagaga actggtttgc ggggcagctg gcgtcgtgca ggcagcagcg cgcgtcggtg	23160
ttggcgatct gcaccacgtt gcgccccac cggttcttca cgatcttggc cttggacgat	23220
tgctccttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttcacca tgctgctgcc gtgcagacac ttcagctcgc cctcgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtgggctcg aaagacttgt aggtcacctc cgcgaaggac	23400
tgcaggtacc cctgcaaaaa gcggcccatc atggtcacga aggtcttggt gctgctgaag	23460
gtcagctgca gcccgcggtg ctctctgttc agccaggtct tgcacacggc cgccagcgcc	23520
tccacctggt cgggcagcat cttgaagttc accttcagct cattctccac gtggtacttg	23580
tccatcagcg tgcgcgccgc ctccatgccc ttctcccagg ccgacaccag cggcaggctc	23640
acggggttct tcaccatcac cgtggccgcc gcctccgccg cgctttcgct ttccgccccg	23700
ctgttctctt cctcttctc ctcttctcg ccgccgccca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttg cgcttgccgt tgcgccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcacc agcgcggcct cttcttgctc gtctcgtg	23880
tccagaatga cctccgggga gggggggttg gtcactctca gtaccgagge acgtttcttt	23940
ttcttctctg gggcgcttcgc cagctccgcg gctgcggccg ctgccaggt cgaaggccga	24000
gggctgggcg tgcgcggcac cagcgcgtcc tgcgagccgt cctcgtctc ctcggactcg	24060
agacggaggc gggcccgtt ctccgggggc gcgcggggcg gcggaggcgg ccggcggcac	24120

[0052]

ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgcc gcgtccgcgc	24180
tcgggggtgg tctcgcgtg gtctcttcc cgactggcca tctcccactg ctctttctcc	24240
tataggcaga aagagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgccccct ctgagccctc caccaccgcc gccaccaccg ccaatgccgc cgcggacgac	24360
gcgcccaccg agaccaccgc cagtaccacc ctccccagcg acgcaccccc gctcgagaat	24420
gaagtgtga tcgagcagga cccgggtttt gtgagcggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggtcgc cgctcagtg ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata aggatgagac agcagtcggg cgggggaacg gaagccatga tgctgatgac	24600
ggctacctag acgtgggaga cgacgtgctg cttaagcacc tgcaccgcca gtgcgtcatc	24660
gtctgcgacg cgctgcagga gcgtgcgaa gtgcccctgg acgtggcgga ggtcagccgc	24720
gcctacgagc ggcacctctt cgcgccgcac gtgccccca agcgccggga gaacggcacc	24780
tgcgagccca acccgctct caacttctac ccggtcttcg cggtaccga ggtgctggcc	24840
acctaccaca tctttttcca aaactgcaag atccccctct cctgccgcgc caaccgcacc	24900
cgcgccgaca aaaccctgac cctgcggcag ggcgccaca tacctgatat cgcctctctg	24960
gaggaagtgc ccaagatctt cgagggtctc ggtcgcgacg agaaacgggc ggcgaacgct	25020
ctgcacggag acagcgaana cgagagtcac tcgggggtgc tggtaggagct cgagggcgac	25080
aacgcgcgcc tggccgtact caagegcagc atagaggatca cccactttgc ctaccggcg	25140
ctcaacctgc ccccaaggt catgagtgtg gtcatgggag agctcatcat gcgccgcgc	25200
cagcccctgg ccgcggatgc aaacttgcaa gattcctccg aggaaggcct gcccgcggtc	25260
agcgacgagc agctggcgcg ctggctggag acccgcgacc ccgcgcagct ggaggagcgg	25320
cgcaagctca tgatggccgc ggtgctggtc accgtggagc tcgagtgtct gcagcgcttc	25380
ttcgcggacc ccgagatgca gcgcaagctc gaggagacc tgcaactac cttccgccag	25440
ggctacgtgc gccaggcctg caagatctcc aacgtggagc tctgcaacct ggtctctctac	25500
ctgggcatcc tgcacgagaa ccgcctcggg cagaacgtcc tgcactccac cctcaaagg	25560
gaggcgcgcc gcgactacat ccgcgactgc gcctacctt tcctctgcta cacctggcag	25620
acggccatgg gggctctggca gcagtgcctg gaggagcgca acctcaagga gctggaaaag	25680
ctctcaagc gcacctcag ggacctctgg acgggcttca acgagcgctc ggtggccgcc	25740
gcgctggcgg acatcatctt tcccagcgc ctgctcaaga ccctgcagca gggcctgccc	25800
gacttcacca gccagagcat gctgcagaac ttcaggactt tcctctgga gcgctcgggc	25860

[0053]

atcctgccgg ccacttgctg cgcgctgccc agcgacttcg tgcccatcaa gtacagggag	25920
tgcccgccgc cgctctgggg ccactgctac ctcttccagc tggccaacta cctcgcctac	25980
cactcggacc tcatggaaga cgtgagcggc gagggcctgc tcgagtgcc a ctgccgctgc	26040
aacctctgca cgccccaccg ctctctagtc tgcaaccgc agctgctcag cgagagtcag	26100
attatcggt a ccttcgagct gcagggtccc tcgctgacg agaagtccgc ggctccaggg	26160
ctgaaactca ctccggggct gtggacttcc gcctacctac gcaaatttgt acctgaggac	26220
taccacgccc acgagatcag gttctacgaa gaccaatccc gcccgcccaa ggcggagctc	26280
accgcctgcg tcatcaccca ggggcacatc ctgggccaat tgcaagccat caacaaagcc	26340
cgccgagagt tcttgctgaa aaagggtcgg ggggtgtacc tggaccccca gtccggcgag	26400
gagctaaacc cgctaccccc gccgccgccc cagcagcggg accttgcttc ccaggatggc	26460
accagaaaag aagcagcagc cgccgccgcc gccgcagcca tacatgcttc tggaggaaga	26520
ggaggaggac tgggacagtc aggcagagga ggtttcggac gaggagcagg aggagatgat	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgcaaca ccatcgccct cggtcgcagc cccctcgccg gggcccctga aatcctccga	26700
accagcacc agcgtataa cctccgctcc tccggcgccg gcgccaccg cccgcagacc	26760
caaccgtaga tgggacacca caggaaccgg ggtcggtaag tccaagtgcc cgccgccgcc	26820
accgcagcag cagcagcagc agcgccaggg ctaccgctcg tggcgcgggc acaagaacgc	26880
catagtcgcc tgcttgcaag actgcggggg caacatctct ttcgccgcc gcttcctgct	26940
attccaccac ggggtcgcct ttccccgaa tgtcctgcat tactaccgtc atctctacag	27000
cccctactgc agcggcgacc cagaggcggc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatc ctccgcgggc aagacagcgg cagcagcggc caggagaccc gcggcagcag	27120
cggcgggagc ggtgggcgca ctgcgcctct cgcccaacga acccctctcg acccgggagc	27180
tcagacacag gatcttcccc actttgtatg ccatcttcca acagagcaga ggccaggagc	27240
aggagctgaa aataaaaaac agatctctgc gctccctcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagcttcgg cgcacgctgg aggacgcgga ggcactcttc agcaaatact	27360
gcgcgctcac tcttaaagac tagctccgcg ccttctctga atttaggcgg gagaaaacta	27420
cgtcategcc ggccgccgcc cagccccccc agccgagatg agcaaagaga ttcccacgcc	27480
atacatgtgg agctaccagc cgcagatggg actcgcggcg ggagcggccc aggactactc	27540

[0054]

cacccgcatg aactacatga gcgcgggacc ccacatgata tcacaggtca acgggatccg	27600
cgcccagcga aaccaaatac tgctggaaca ggcgggccatc accgccacgc cccgccataa	27660
tctcaacccc cgaaattggc ccgcgcgcct cgtgtaccag gaaaccccct ccgccaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcag gggcgcagct	27780
cgcgggcggc tttcgtcacg gggcgcgggc gtcgccacca ggtataagac acctgatgat	27840
cagaggccga ggtatccagc tcaacgacga gtcggtgagc tcttcgctcg gtctccgtcc	27900
ggacggaact ttccagctcg ccggatccgg ccgctcttcg ttcacgcccc gccaggcgta	27960
cctgactctg cagacctgt cctcggagcc ccgctccggc ggcatcgga cctccagtt	28020
cgtggaggag ttcgtgccct cggctctactt caaccccttc tcgggacctc ccggacgcta	28080
ccccgaccag ttcatccga actttgacgc ggtgaaggac tcggcggacg gctacgactg	28140
aatgtcaggt gtcgaggcag agcagcttcg cctgagacac ctcgagcact gccgcccca	28200
caagtgttc gccgcgggt ctggtgagtt ctgctacttt cagctaccgc aggagcatac	28260
cgagggggccg gcgcacggcg tccgcctgac caccagggc gaggttacct gttccctcat	28320
ccgggagttt accctccgtc ccctgctagt ggagcgggag cggggtcctt gtgtcctaac	28380
tatgcctgc aactgcccta accctggatt acatcaagat ctttgctgtc atctctgtgc	28440
tgagtttaaat aaacgtgag atcagaatct actggggctc ctgtcgccat cctgtgaacg	28500
ccaccgtctt caccaccccc gaccagccc aggcgaacct cacctgcggt ctgcatcgga	28560
gggccaagaa gtacctcacc tggctacttc acggcacccc ctttgtggtt tacaacagct	28620
tcgacgggga cggagtctcc ctgaaagacc agctctccgg tctcagctac tccatccaca	28680
agaacaccac cctccaactc ttcctccct acctgccggg aacctacgag tgcgtcaccg	28740
gccgtgcac ccacctcacc cgcctgatcg taaaccagag ctttccggga acagataact	28800
ccctcttccc cagaacagga ggtgagctca ggaaactccc cggggaccag ggcggagacg	28860
taccttcgac ctttgtgggg ttaggatttt ttattaccgg gttgctggct cttttaatca	28920
aagtttcctt gagatttggt ctttcttct acgtgtatga acacctcaac ctccaataac	28980
tctacccttt cttcggaatc aggtgacttc tctgaaatcg ggcttggtgt gctgcttact	29040
ctgttgattt ttttcttat catactcagc cttctgtgcc tcaggtcgc cgcctgctgc	29100
gcacacatct atatctactg ctggttgctc aagtgcaggg gtcgccaccc aagatgaaca	29160
ggtacatggt cctatcgatc ctaggcctgc tggccctggc ggccctgcagc gccgcaaaa	29220
aagagattac ctttgaggag cccgcttgca atgtaacttt caagcccgag ggtgaccaat	29280

[0055]

gcaccaccct cgtcaaagtc gttaccaatc atgagaggct gcgcacgcac tacaaaaaca	29340
aaactggcca gtttgcggtc tatagtgtgt ttacgcccgg agaccctctt aactactctg	29400
tcaccgtctt ccaggcgga cagtctaaga tattcaatta cactttccct ttttatgagt	29460
tatgcgatgc ggtcatgtac atgtcaaac agtacaacct gtggcctccc tctccccagg	29520
cgtgtgtgga aaatactggg tcttactgct gtatggcttt cgcaatcact acgctcgtc	29580
taatctgcac ggtgctatac ataaaattca ggcagaggcg aatctttatc gatgaaaaga	29640
aaatgccttg atcgctaaca ccggctttct atctgcagaa tgaatgcaat caccctcccta	29700
ctaataccca ccaccctcct tgcgattgcc catgggttga cacgaatcga agtgccagtg	29760
gggtccaatg tcaccatggt gggccccgcc ggcaattcca cctcatgtg ggaaaaattt	29820
gtccgcaatc aatgggttca tttctgtctt aaccgaatca gtatcaagcc cagagccatc	29880
tgcgatgggc aaaatctaac tctgatcaat gtgcaaatga tggatgctgg gtactattac	29940
gggcagcggg gagaaatcat taattactgg cgacccaca aggactacat gctgcatgta	30000
gtcgaggcac ttcccactac cccccact accacctctc ccaccaccac caccactact	30060
actactacta ctactactac tactactacc actaccgtg cccgccatac ccgcaaaagc	30120
accatgatta gcacaaagcc cctcgtgct cactcccacg ccggcgggcc catcggtgcg	30180
acctcagaaa ccaccgagct ttgcttctgc caatgcacta acgccagcgc tcatgaactg	30240
ttcgacctgg agaatagagga tgtccagcag agctccgctt gcctgacca ggaggctgtg	30300
gagcccgttg cctgaagca gatcggtgat tcaataattg actcttcttc ttttgccact	30360
cccgaatacc ctccgattc tactttccac atcacgggta ccaaagacc taacctctct	30420
ttctacctga tgctgtgct ctgtatctct gtggtctctt ccgcgtgat gttactgggg	30480
atgttctgct gcctgatctg ccgcagaaag agaaaagctc gctctcaggg ccaaccactg	30540
atgcccttcc cctaccccc ggattttgca gataacaaga tatgagctcg ctgctgacac	30600
taaccgcttt actagcctgc gctctaacc ttgtcgcttg cgactcgaga ttccacaatg	30660
tcacagctgt ggcaggagaa aatgttactt tcaactccac ggccgatacc cagtggctgt	30720
ggagtggctc aggtagctac ttaactatct gcaatagctc cacttcccc ggcatatccc	30780
caaccaagta ccaatgcaat gccagcctgt tcacctcat caacgcttcc accctggaca	30840
atggactcta tgtaggctat gtaccctttg gtgggcaagg aaagaccac gcttacaacc	30900
tggaagtctg ccagcccaga accactaccc aagcttctcc caccaccacc accaccacca	30960

[0056]

ccatcaccag cagcagcagc agcagcagcc acagcagcag cagcagatta ttgacttttg	31020
ttttggccag ctcatctgcc gctacccagg ccatctacag ctctgtgccc gaaaccactc	31080
agateccaccg cccagaaacg accaccgcca ccaccctaca cacctccagc gatcagatgc	31140
cgaccaacat caccctcttg gctcttcaaa tgggacttac aagccccact ccaaaaccag	31200
tggatgcggc cgaggtctcc gccctcgtca atgactgggc ggggctggga atgtggtggt	31260
tcgccatagg catgatggcg ctctgcctgc ttctgctctg gctcatctgc tgcctccacc	31320
gcaggcgagc cagaccccc atctatagac ccatcattgt cctgaacccc gataatgatg	31380
ggatccatag attggatggc ctgaaaaacc tacttttttc ttttacagta tgataaattg	31440
agacatgcct cgcattttct tgtacatgtt ctttctccca ctttttctgg ggtgttctac	31500
gctggccgct gtgtctcacc tggaggtaga ctgcctctca cccttactg tctacctgct	31560
ttacggattg gtcaccctca ctctcatctg cagcctaate acagtaatca tcgccttcat	31620
ccagtgcatt gattacatct gtgtgcgcct cgcatacttc agacaccacc cgcagtaccg	31680
agacaggaac attgcccac ttctaagact gctctaate tgacataagac tgtgatctgc	31740
cttctgatcc tctgcatcct gccaccctc acctcctgcc agtacaccac aaaatctccg	31800
cgcaaaagac atgcctcctg ccgcttcacc caactgtgga atatacccaa atgctacaac	31860
gaaaagagcg agctctccga agcttggctg tatggggctca tctgtgtctt agttttctgc	31920
agcactgtct ttgccctcat aatctacccc tacttttgatt tgggatggaa cgcgatcgat	31980
gccatgaatt accccacctt tcccgcaccc gagataatte cactgcgaca agttgtaccc	32040
gttgtcgtta atcaacgccc cccatcccct acgcccactg aaatcagcta ctttaaccta	32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatcagta ccgagcagcg	32160
tctcctagag aggcgagcgc aggcggctga gcaagagcgc ctcaatcagg agctccgaga	32220
tctcgttaac ctgcaccagt gcaaaagagg catcttttgt ctggtaaagc aggccaaagt	32280
cacctacgag aagaccggca acagccaccg cctcagttac aaattgccc cccagcgcca	32340
gaagctggtg ctcatggtgg gtgagaatcc catcaccgtc acccagcact cggtagagac	32400
cgaggggtgt ctgcactccc cctgtcgggg tccagaagac ctctgcaccc tggtaaagac	32460
cctgtgcggt ctcagagatt tagtcccctt taactaatca aacactggaa tcaataaaaa	32520
gaatcactta cttaaaatca gacagcaggt ctctgtccag tttattcage agcacctcct	32580
tcccctctc ccaactctgg tactccaaac gccttctgga ggcaaacttc ctccacaccc	32640
tgaagggaat gtcagattct tgctcctgtc cctccgcacc cactatcttc atgttgttgc	32700

[0057]

agatgaagcg caccaaaacg tctgacgaga gcttcaaccc cgtgtacccc tatgacacgg	32760
aaagcggccc tccctccgtc cctttcctca cccctccctt cgtgtctccc gatggattcc	32820
aagaaagtcc ccccggggtc ctgtctctga acctggccga gcccctggtc acttcccacg	32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgtggc aacctcacct	32940
ctcaagatat caccaccgct agccctcccc tcaaaaaaac caagaccaac ctacgcctag	33000
aaacctcatc ccccctaact gtgagcacct caggcgccct caccgtagca gccgcgcctc	33060
ccctggcggt ggccggcacc tccctcacca tgcaatcaga ggccccctg acagtacagg	33120
atgcaaaact caccctggcc accaaaggcc ccctgaccgt gtctgaaggc aaactggcct	33180
tgcaaacatc ggccccgtg acggccgtg acagcagcac cctcacagtc agtgccacac	33240
cacccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc atttacacca	33300
ccaatggaaa actaggactt aactttggcg ctccccctgca tgtggtagac agcctaaatg	33360
cactgactgt agttactggc caaggtctta cgataaacgg aacagcccta caaactagag	33420
tctcaggtgc cctcaactat gacacatcag gaaacctaga attgagagct gcaggggta	33480
tgcgagttga tgcaaatggt caacttatcc ttgatgtagc ttaccattt gatgcacaaa	33540
acaatctcag ccttaggctt ggacaggac ccctgtttgt taactctgcc cacaacttgg	33600
atgttaacta caacagaggc ctctacctgt tcacatctgg aaataccaaa aagctagaag	33660
ttaatataca aacagccaag ggtctcattt atgatgacac tgctatagca atcaatgcgg	33720
gtgatgggct acagtttgac tcaggctcag atacaaatcc attaaaaact aaacttggat	33780
taggactgga ttatgactcc agcagagcca taattgctaa actgggaact ggcctaagct	33840
ttgacaacac aggtgccatc acagtaggca acaaaaatga tgacaagctt accttgtgga	33900
ccacaccaga cccatcccct aactgtagaa tctattcaga gaaagatgct aaattcacac	33960
ttgttttgac taaatgcggc agtcaggtgt tggccagcgt ttctgtttta tctgtaaaag	34020
gtagccttgc gcccatcagt ggcacagtaa ctagtgtcga gattgtcctc agatttgatg	34080
aaaatggagt tctactaagc aattcttccc ttgacctca atactggaac tacagaaaag	34140
gtgaccttac agagggcact gcatatacca acgcagtggg atttatgccc aacctcacag	34200
catacccaaa aacacagagc caaactgcta aaagcaacat tgtaagtcag gtttacttga	34260
atggggacaa atccaaaccc atgacctca ccattaccct caatggaact aatgaaacag	34320
gagatgccac agtaagcact tactccatgt cattctcatg gaactggaat ggaagtaatt	34380

[0058]

acattaatga aacgttccaa accaactcct tcaccttctc ctacatcgcc caagaataaa	34440
aagcatgacg ctgttgattt gattcaatgt gtttctgttt tattttcaag cacaacaaaa	34500
tcattcaagt cattcttcca tcttagctta atagacacag tagcttaata gaccagtag	34560
tgcaaagccc cattctagct tataactagt ggagaagtac tcgcctacat gggggtagag	34620
tcataatcgt gcatcaggat agggcggttg tgctgcagca gcgcgcgaat aaactgctgc	34680
cgccgcccgt cgtcctgca ggaatacaac atggcagtgg tctcctcagc gatgattcgc	34740
accgcccga gcataaggcg ccttgctctc cgggcacagc agcgcaccct gatctcactt	34800
aaatcagcac agtaactgca gcacagcacc acaatattgt tcaaaatccc acagtgaag	34860
gcgctgtatc caaagctcat ggccggggacc acagaaccca cgtggccatc ataccacaag	34920
cgcaggtaga ttaagtggcg acccctcata aacacgtgg acataaacat tacctctttt	34980
ggcatgttgt aattcaccac ctcccgttac catataaacc tctgattaaa catggcgcca	35040
tccaccacca tctaaacca gctggccaaa acctgcccgc cggctataca ctgcaggga	35100
ccgggactgg aacaatgaca gtggagagcc caggactcgt aaccatggat catcatgctc	35160
gtcatgatat caatgttggc acaacacagg cacacgtgca tacacttctc caggattaca	35220
agctcctccc gcgttagaac catatcccag ggaacaaccc attcctgaat cagcgtaa	35280
cccacactgc aggaagacc tcgcacgtaa ctacgttgt gcattgtcaa agtgttacat	35340
tcgggcagca gcggatgatc ctccagtatg gtagcgcggg tttctgtctc aaaaggaggt	35400
agacgatccc tactgtacgg agtgcgccga gacaaccgag atcgtgttgg tcgtagtgtc	35460
atgccaaatg gaacgccgga cgtagtcata tttcctgaag tcttagatct ctcaacgcag	35520
caccagcacc aacacttcgc agtgtaaaag gccaaagtgc gagagagtat atataggaat	35580
aaaaagtgc gtaaaccggc aaagtccaaa aaacgcccag aaaaaccgca cgcgaaaccta	35640
cgccccgaaa cgaaagccaa aaaacactag aactcctt ccggcgtcaa cttccgcttt	35700
cccacgtac gtcacttgcc ccagtcaaac aaactacata tcccgaactt ccaagtcgcc	35760
acgccccaaa caccgcctac acctccccgc ccgcccggcc gccccaaac ccgctccc	35820
ccccgcgcc cgccccgcgc cgccatctc attatcatat tggttcaat ccaaaataag	35880
gtatattatt gatgatggtt taaacggatc ctctagagtc gacctgcagg catgcaagct	35940
tgagtattct atagtgtcac ctaaatagtc tggcgtaate atggatcatag ctgtttcctg	36000
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgt	36060
aagcctgggg tgcctaata gtagagctaac tcacattaat tgcgttgccg tcaactgccc	36120

[0059]

ctttccagtc	gggaaacctg	tcgtgccage	tgcattaatg	aatcggccaa	cgcgaacccc	36180
ttgcggccgc	cggggccgtc	gaccaattct	catgtttgac	agcttatcat	cgaatttctg	36240
ccattcatcc	gcttattatc	acttattcag	gcgtagcaac	caggcgttta	agggcaccaa	36300
taactgcctt	aaaaaaatta	cgccccgccc	tgccactcat	cgcagtactg	ttgtaattca	36360
ttaagcattc	tgccgacatg	gaagccatca	caaacggcat	gatgaacctg	aatcgccage	36420
ggcatcagca	ccttgtcgcc	ttgcgtataa	tatttgccca	tggtgaaaac	gggggcgaag	36480
aagtgtgcca	tattggccac	gtttaaatca	aaactgggtga	aactcaccca	gggattggct	36540
gagacgaaaa	acataattctc	aataaacctt	ttagggaaat	aggccagggt	ttcaccgtaa	36600
cacgccacat	cttgcaata	tatgtgtaga	aactgccgga	aatcgtcgtg	gtattcactc	36660
cagagcgatg	aaaacgtttc	agtttgctca	tggaaaacgg	tgtaacaagg	gtgaacacta	36720
tcccatatca	ccagctcacc	gtctttcatt	gccatacgga	attccggatg	agcattcatc	36780
aggcgggcaa	gaatgtgaat	aaaggccgga	taaaacttgt	gcttattttt	ctttacggtc	36840
tttaaaaagg	ccgtaatatc	cagctgaacg	gtctggttat	aggtacattg	agcaactgac	36900
tgaaatgcct	caaaatgttc	tttacgatgc	cattgggata	tatcaacggt	ggtatatcca	36960
gtgatttttt	tctccatttt	agcttcctta	gtcctgaaa	atctcgataa	ctcaaaaaat	37020
acgcccggta	gtgatcttat	ttcattatgg	tgaaagttgg	aacctcttac	gtgccgatca	37080
acgtctcatt	ttcgccaaaa	gttggcccag	ggcttcccgg	tatcaacagg	gacaccagga	37140
tttatttatt	ctgcgaagtg	atcttccgtc	acaggtattt	attcgcgata	agctcatgga	37200
gcggcgtaac	cgctgcacag	gaaggacaga	gaaagcgcg	atctgggaag	tgacggacag	37260
aacggtcagg	acctggattg	gggaggcggt	tgccgccgct	gctgctgacg	gtgtgacggt	37320
ctctgttccg	gtcacaccac	atacgttccg	ccatttcctat	gcgatgcaca	tgctgtatgc	37380
cggatataccg	ctgaaagttc	tgcaaagcct	gatgggacat	aagtccatca	gttcaacgga	37440
agtctacacg	aagggtttttg	cgttggtatg	ggctgcccgg	caccgggtgc	agtttgcgat	37500
gccggagtct	gatgcggttg	cgatgctgaa	acaattatcc	tgagaataaa	tgccttggcc	37560
tttatatgga	aatgtggaac	tgagtggata	tgctgttttt	gtctgtttaa	cagagaagct	37620
ggctgtttatc	cactgagaag	cgaacgaaac	agtcgggaaa	atctcccatt	atcgtagaga	37680
tccgcattat	taatctcagg	agcctgtgta	gcgtttatag	gaagtagtgt	tctgtcatga	37740
tgcctgcaag	cggtaacgaa	aacgatttga	atatgccttc	aggaacaata	gaaatcttcg	37800

[0060]

tgcggtgtta cgttgaagtg gagcggatta tgtcagcaat ggacagaaca acctaataa	37860
cacagaacca tgatgtggtc tgtcctttta cagccagtag tgctcgccgc agtcgagcga	37920
cagggcgaag ccctcgagtg agcgaggaag caccagggaa cagcacttat atattctgct	37980
tacacacgat gcctgaaaaa acttcccttg gggttatcca cttatccacg gggatatttt	38040
tataattatt tttttatag tttttagatc ttctttttta gagcgccttg taggccttta	38100
tccatgctgg ttctagagaa ggtgttgtga caaattgccc tttcagtgtg acaaataacc	38160
ctcaaatgac agtcctgtct gtgacaaatt gcccttaacc ctgtgacaaa ttgccctcag	38220
aagaagctgt tttttcacia agttatccct gcttattgac tcttttttat ttagtgtgac	38280
aatctaaaaa cttgtcacac ttcacatgga tctgtcatgg cggaacacgc ggttatcaat	38340
cacaagaaac gtaaaaatag cccgcgaatc gtccagtcaa acgacctcac tgaggcggca	38400
tatagtctct cccgggatca aaaacgtatg ctgtatctgt tcgttgacca gatcagaaaa	38460
tctgatggca ccctacagga acatgacggg atctgcgaga tccatgttgc taaatatgct	38520
gaaatatctg gattgacctc tgcggaagcc agtaaggata tacggcaggc attgaagagt	38580
ttcgcgggga aggaagtgg tttttatcgc cctgaagagg atgccggcga tgaaaaaggc	38640
tatgaatctt ttccttgggt tatcaaacgt gcgcacagtc catccagagg gctttacagt	38700
gtacatatca acccatatct cattcccttc tttatcgggt tacagaaccg gtttacgcag	38760
tttcggctta gtgaaacaaa agaaataacc aatccgtatg ccatgcgttt atacgaatcc	38820
ctgtgtcagt atcgtaagcc ggatggctca ggcacgtct ctctgaaaat cgactggatc	38880
atagagcggt accagctgcc tcaaagttac cagcgtatgc ctgacttccg ccgccgcttc	38940
ctgcaggtct gtgttaatga gatcaacagc agaactccaa tgcgcctctc atacattgag	39000
aaaaagaaag gccgccagac gactcatatc gtattttcct tccgcgatat cacttccatg	39060
acgacaggat agtctgaggg ttatctgtca cagatttgag ggtggttcgt cacatttggt	39120
ctgacctact gagggttaatt tgtcacagtt ttgctgtttc cttcagcctg catggatttt	39180
ctcatacttt ttgaactgta atttttaagg aagccaaatt tgagggcagt ttgtcacagt	39240
tgatttcctt ctctttccct tcgtcatgtg acctgatatc gggggttagt tcgtcatcat	39300
tgatgagggt tgattatcac agtttattac tctgaattgg ctatccgcgt gtgtacctct	39360
acctggagtt tttccacagg tggatatctc ttcttgctgt gagcgtaaga gctatctgac	39420
agaacagttc ttctttgctt cctcgccagt tcgctcgcta tgctcggtta cacggctgcg	39480
gcgagcgcta gtgataataa gtgactgagg tatgtgctct tcttatctcc ttttgtagtg	39540

[0061]

ttgctcttat tttaaacaac tttgcggttt tttgatgact ttgcgatttt gttgttgctt	39600
tgcagtaaat tgcaagattt aataaaaaaa cgcaaagcaa tgattaaagg atgttcagaa	39660
tgaaactcat ggaaacactt aaccagtgc taaacgctgg tcatgaaatg acgaaggcta	39720
tcgccattgc acagttaa atgatgacagcc cggaagcgag gaaaataacc cggcgctgga	39780
gaataggtga agcagcggat ttagttgggg tttcttctca ggctatcaga gatgccgaga	39840
aagcagggcg actaccgcac ccgatatgg aaattcgagg acgggttgag caacgtgttg	39900
gttatacaat tgaacaaatt aatcatatgc gtgatgtgtt tggtaacgca ttgcgacgtg	39960
ctgaagacgt atttccaccg gtgatcgggg ttgctgcca taaaggtggc gtttacaaaa	40020
cctcagtttc tgttcattct gctcaggatc tggctctgaa ggggctacgt gttttgctcg	40080
tggaaggtaa cgacccccag ggaacagcct caatgtatca cggatgggta ccagatcttc	40140
atattcatgc agaagacact ctctgcctt tctatcttgg ggaaaaggac gatgtcactt	40200
atgcaataaa gccacttgc tggccggggc ttgacattat tccttctgt ctggctctgc	40260
accgtattga aactgagtta atgggcaa attgatgaagg taaactgccc accgatccac	40320
acctgatgct ccgactggcc attgaaactg ttgctcatga ctatgatgtc atagttattg	40380
acagcgcgcc taacctgggt atcggcacga ttaatgtcgt atgtgctgct gatgtgctga	40440
ttgttccac gcctgctgag ttgtttgact acacctccgc actgcagttt ttcgatatgc	40500
ttcgtgatct gctcaagaac gttgatctta aagggttcga gcctgatgta cgtattttgc	40560
ttaccaata cagcaatagt aatggctctc agtccccgtg gatggaggag caaattcggg	40620
atgcctgggg aagcatggtt ctaaaaaatg ttgtacgtga aacggatgaa gtttgtaaag	40680
gtcagatccg gatgagaact gtttttgaac aggccattga tcaacgctct tcaactggtg	40740
cctggagaaa tgctctttct atttgggaac ctgtctgcaa tgaaattttc gatcgtctga	40800
ttaaaccacg ctgggagatt agataatgaa gcgtgcgcct gttattccaa aacatacgct	40860
caatactcaa ccggttgaag atacttcgtt atcgacacca gctgccccga tggatgattc	40920
gttaattgcg cgcgtaggag taatggctcg cggtaatgcc attactttgc ctgtatgtgg	40980
tcgggatgtg aagtttactc ttgaagtgt cgggggtgat agtgttgaga agacctctcg	41040
ggtatggtca ggtaatgaac gtgaccagga gctgcttact gaggacgcac tggatgatct	41100
catcccttct tttctactga ctgggtcaaca gacaccggcg ttcggctgaa gagtatctgg	41160
tgtcatagaa attgccgatg ggagtcgccg tcgtaaagct gctgcactta ccgaaagtga	41220

[0062]

ttatcgtgtt ctggttggcg agctggatga tgagcagatg gctgcattat ccagattggg	41280
taacgattat cgcccaacaa gtgcttatga acgtggtcag cgttatgcaa gccgattgca	41340
gaatgaatth gctggaaata tttctgcgct ggctgatgcg gaaaatattt cacgtaagat	41400
tattaccgcg tgtatcaaca ccgccaatt gcctaaatca gttgttgctc ttttttctca	41460
ccccggtgaa ctatctgccc ggtcaggtga tgcacttcaa aaagccttta cagataaaga	41520
ggaattactt aagcagcagg catctaacct tcatgagcag aaaaaagctg gggtgatatt	41580
tgaagctgaa gaagttatca ctcttttaac ttctgtgctt aaaacgtcat ctgcatcaag	41640
aactagttta agctcacgac atcagtttgc tcctggagcg acagtattgt ataaggcgca	41700
taaaatggtg cttaacctgg acaggtctcg tgttccaact gagtgatatag agaaaattga	41760
ggccattctt aaggaacttg aaaagccagc accctgatgc gaccacgttt tagtctacgt	41820
ttatctgtct ttacttaatg tcctttgtta caggccagaa agcataactg gcctgaatat	41880
tctctctggg cccactgttc cacttgatc gtcggtctga taatcagact gggaccacgg	41940
tcccactcgt atcgtcggtc tgattattag tctgggacca cggcccact cgtatcgtcg	42000
gtctgattat tagtctggga ccacggctcc actcgtatcg tcggtctgat aatcagactg	42060
ggaccacggt cccactcgta tcgtcggctt gattattagt ctgggacat ggtcccactc	42120
gtatcgtcgg tctgattatt agtctgggac caggtccca ctcgtatcgt cggctctgatt	42180
attagtctgg aaccacggtc cactcgtat cgtcggctcg attattagtc tgggaccacg	42240
gtcccactcg tatcgtcggc ctgattatta gtctgggacc acgatccac tcgtgttgctc	42300
ggtctgatta tcggtctggg accacggctc cacttgatt gtcatcaga ctatcagcgt	42360
gagactacga ttccatcaat gcctgtcaag ggcaagtatt gacatgtcgt cgtaacctgt	42420
agaacggagt aacctcggcg tcggttgta tgcctgctgt ggattgctgc tgtgtcctgc	42480
ttatccacaa ctttttgcgc acggttatgt ggacaaaata cctggttacc caggccgtgc	42540
cggcacgtta accgggctgc atccgatgca agtgtgtcgc tgcgacgag ctcgcgagct	42600
cggacatgag gttgccccgt attcagtgtc gctgatttgt attgtctgaa gttgttttta	42660
cgttaagttg atgcagatca attaatacga tacctgcgtc ataattgatt atttgacgtg	42720
gtttgatggc ctccacgcac gttgtgatat gtagatgata atcattatca ctttacgggt	42780
cccttcgggt gatccgacag gttacggggc ggcgacctcg cgggttttcg ctatttatga	42840
aaattttccg gtttaaggcg tttccgttct tcttcgtcat aacttaattgt ttttatttaa	42900
aataccctct gaaaagaaag gaaacgacag gtgctgaaag cgagcttttt ggctctgtc	42960

[0063]

gtttcctttc tctgtttttg tccgtggaat gaacaatgga agtccgagct catcgctaata	43020
aacttcgtat agcatacatt atacgaagtt atattcgatg cggccgcaag gggttcgcgt	43080
cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcatcagag cagattgtac	43140
tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca	43200
tcaggcgcca ttcgccattc aggtcgcga actgttggga agggcgatcg gtgcgggcct	43260
cttcgctatt acgccagctg gcgaaaggg gatgtgctgc aaggcgatta agttgggtaa	43320
cgccagggtt ttccagtcga cgacgttgta aaacgacggc cagtgaattg taatacgact	43380
cactataggg cgaattcgag ctcggtaccc ggggatcctc gtttaaac	43428

<210> 9

<211> 45227

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 9

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tggggcgggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acatttatat tggctcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atatatggag ttccgcgtta cataacttac ggtaaattgg ccgcctggct gaccgccc	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac	720
tttccattga cgtcaatggg tggagtattt acggtaaaact gccacttgg cagtacatca	780
agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggccccctg	840
gcattatgcc cagtacatga ctttatggga ctttcctact tggcagtaca tctacgtatt	900

[0064]

agtcacgcgt attaccatgg tgatgcgggt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg	1020
gcacaaaaat caacgggact ttcaaaaatg tcgtaacaac tccgccccat tgacgcaa	1080
ggcggttagg cgtgtacggg gggagggtcta tataagcaga gctctcccta tcagtgatag	1140
agatctccct atcagtgata gagatcgteg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacgggtgca ttggaacgcg gattccccgt gccaaagagt agatcttccg	1320
tttatctagg taccgggccc cccctcgagg tcgacgggtat cgataagctt cacgctgccg	1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgtagaaag ccagtccgca	1440
gaaacgggtgc tgaccccgga tgaatgtcag ctactgggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagcttgcag tgggcttaca tggcgatagc tagactgggc	1560
ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgccctctg gtaaggttgg	1620
gaagccctgc aaagtaaact ggatggcttt cttgccgcca aggatctgat ggcgcagggg	1680
atcaagatct aaccaggagc tatttaattg caacagttaa ccagctggta cgcaaaccac	1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgctgga agcatgcccg caaaaacgtg	1800
gcgtatgtac tcgtgtatat actaccactc ctaaaaaacc gaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggg ggtgaaggtc	1920
acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgtttaa gacctcccgg	1980
gtgttcgtta ccacaccgta cgtggtgcgc ttgactgctc cggcgtttaa gaccgtaagc	2040
aggctcgttc caagtatggc gtgaagcgtc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcgtttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat	2220
gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg	2280
tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tategtggct ggccacgacg	2340
ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcattctacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcggcggtcg catacgtttg atccggctac ctgccattc	2520
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	2580

[0065]

gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgctgtga cccatggcga tgcctgcttg	2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tcgactgtgg cgggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc	2880
atcgccttct atcgccttct tgacgagttc ttctgagcgg gactctgggg ttcgaaatga	2940
ccgaccaagc gacgcccac ctgccatcac gagatttcga ttccaccgcc gccttctatg	3000
aaaggttggg cttcggaatc gttttccggg acgccggtg gatgatcctc cagcgcgggg	3060
atctcatgct ggagttcttc gccaccccg ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgat ctgctgtgcc ttctagtgc cagccatctg ttgtttgcc	3180
ctccccctg ctttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa	3240
tgaggaaatt gcatcgcat gtctgagtag gtgtcattct attctggggg gtggggtggg	3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg	3360
ctctatggcc gatcagcgat cgctgaggtg ggtgagtggt cgtggcctgg ggtggtcatg	3420
aaaatatata agttgggggt cttagggtct ctttatttgt gttgcagaga ccgccggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgcc ttggatggca gcatcgtgag	3540
cccttatttg acgacgcgga tgccccactg ggccgggggt cgtcagaatg tgatgggctc	3600
cagcatcgac ggccgaccg tccctgccgc aaattccgcc acgctgacct atgcgacct	3660
cgcggggacg ccgttgacg ccaccgccgc cgccgccgcc accgcagccg cctcgccct	3720
gcgcagcctg gccacggact ttgcattcct gggaccactg gcgacagggg ctacttctcg	3780
ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gettactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggtctc	3900
ctccctgcaa gctggcgga atgcttctcc cacaaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttattt cataatttc	4020
cgcgcgcgat aggcctaga ccagcgttct cggtcggtga gggtcggtg tatcttctcc	4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccggggg	4140
tggaggtagc accactgcag agcttcatgc tccggggtg tgttgtagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tgggtgaagt gtttacaaaa cggttaagtt gggaagggtg cattcgggga	4320

[0066]

gagatgatgt gcatcttgga ctgtatTTTT agattggcga tgtttccgcc cagatccctt	4380
ctgggattca tgttgtgcag gaccaccagt acagtgtatc cgggtgcactt ggggaatttg	4440
tcatgcagct tagagggaag agcgtggaag aacttggaga cgcctttgtg gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag	4560
atatttctgg ggtcgtgac gtcgtagtgt tgttccaggg tgaggtcgtc ataggccatt	4620
tttacaagc gcgggcggag ggtgcccgc tgggggatga tgggtccctc tggccctggg	4680
gcgtagtgtc cctcgcagat ctgcatttcc caggccttaa tctcggaggg gggaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag	4800
agcaggtttc taagcagctg tgattttcca caaccggttg gccataaat aacacctata	4860
accggttgca gctggtagtt tagagagctg cagctgccgt cgtcccggag gaggggggcc	4920
acctcgttga gcatgtccct gacgcgatg ttctccccga ccagatccgc cagaaggcgc	4980
tcgccgccca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg	5040
tccgccgtgg gcatgttttt cagggtctgg ctacagcagct ccaggcggtc ccagagctcg	5100
gtgacgtgct ctacggcatc tctatccagc atatctctc gtttcgcggg ttggggcgac	5160
tttcgtgtga gggcaccaag cgggtggtgt ccagcggggc cagagtcatg tccttccatg	5220
ggcgcagggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgt cggggtgag	5280
cgtttgcaa ggtgcgcttg aggctggttc tgctggtgct gaagcgtgc cggctctgc	5340
cctgcgcgtc ggccaggtag catttgacca tgggtgtcata gtccagcccc tccgcggcgt	5400
gtcccttggc gcgcagcttg cccttgaggg tggcgccgca cgaggggcag agcaggctct	5460
tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgccgc	5520
agaccccgca cacggtctcg cactccacca gccagggtgag ctcggggcgc gccgggtcaa	5580
aaaccagggt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgagggtgt	5640
gtccccgctc ggtgacgaag aggctgtccg tgtctccgta gaccgacttg aggggtcttt	5700
tctccagggg ggtccctcgg tcttctcgt agaggaactc ggaccactct gagacgaagg	5760
cccgcgtcca ggccaggacg aaggaggcta tgtgggaggg gtagcggtcg ttgtccacta	5820
gggggtccac cttctccaag gtgtgaagac acatgtcgcc ttctcggcg tccaggaagg	5880
tgattggctt gtaggtgtag gccacgtgac cgggggttcc tgacgggggg gtataaagg	5940
gggtgggggc gcgtcgtcg tcaactctct cgcatecgt gtctgcgagg gccagctgct	6000

[0067]

ggggtgagta	ttccctctcg	aaggcgggca	tgacctccgc	gctgaggttg	tcagtttcca	6060
aaaacgagga	ggatttgatg	ttcacctgtc	ccgaggtgat	acctttgagg	gtacccgcgt	6120
ccatctggtc	agaaaacacg	atctttttat	tgtccagctt	ggtggcgaac	gacccgtaga	6180
gggcgttgga	gagcagcttg	gcgatggagc	gcagggtctg	gttcttgtcc	ctgtcggcgc	6240
gctccttggc	cgcgatgttg	agctgcacgt	actcgcgcgc	gacgcagcgc	cactcgggga	6300
agacggtggt	gcgctcgtcg	ggcaccaggc	gcacgcgcc	gccgcggttg	tgcagggtga	6360
ccagggtccac	gctggtggcg	acctcgccgc	gcaggcgctc	gttggtccag	cagagacggc	6420
cgcccttgcg	cgagcagaag	gggggcaggg	ggtcgagctg	ggtctcgtcc	ggggggtccg	6480
cgtccacggt	gaaaaccccg	gggcgcaggc	gcgcgtcgaa	gtagtctatc	ttgcaacctt	6540
gcatgtccag	cgcctgctgc	cagtcgcggg	cggcgagcgc	gcgctcgtag	gggttgagcg	6600
gcgggcccc	gggcatgggg	tgggtgagtg	cggaggcgta	catgccgcag	atgtcataga	6660
cgtagagggg	ctcccgagg	accccgatgt	aggtggggta	gcagcgcccg	ccgcggatgc	6720
tggcgcgcac	gtagtcatat	agctcgtgcg	agggggcgag	gaggtcgggg	cccaggttgg	6780
tgcgggcggg	gcgctccgcg	cggaaagacga	tctgcctgaa	gatggcatgc	gagttggaag	6840
agatggtggg	gcgctggaag	acgttgaagc	tggcgtcctg	caggccgacg	gcgtcgcgca	6900
cgaaggaggc	gtaggagtcg	cgcagcttgt	gtaccagctc	ggcggtgacc	tgcacgtcga	6960
gcgcgcagta	gtcgagggtc	tcgcggatga	tgtcatat	agcctgcccc	ttctttttcc	7020
acagctcgcg	gttgaggaca	aactcttcgc	ggtctttcca	gtactcttgg	atcgggaaac	7080
cgtccggttc	cgaacggtaa	gagcctagca	tgtagaactg	gttgacggcc	tggtaggcgc	7140
agcagccctt	ctccacgggg	agggcgtagg	cctgcgcggc	cttgcggagc	gaggtgtggg	7200
tcagggcgaa	ggtgtccctg	accatgactt	tgaggtactg	gtgcttgaag	tcggagtcgt	7260
cgcagccgcc	ccgctcccag	agcgagaagt	cggtgcgctt	cttgagcgcg	gggttgggca	7320
gagcgaaggt	gacatcggtg	aagaggat	tgcccgcgcg	gggcatgaag	ttgcgggtga	7380
tgcggaaggg	ccccggcact	tcagagcggt	tgttgatgac	ctgggcggcg	agcacgatct	7440
cgtcgaagcc	gttgatgttg	tggcccacga	tgtagagttc	caggaagcgg	ggccggccct	7500
ttacggtggg	cagcttcttt	agctcttcgt	aggtgagctc	ctcgggcgag	gcgaggccgt	7560
gctcggccag	ggcccagtcc	gcgaggtgcg	ggttgtctct	gaggaaggac	ttccagaggt	7620
cgcgggccag	gagggtctgc	aggcggctct	tgaagtcct	gaactggcgg	cccacggcca	7680
ttttttcggg	ggtgatgcag	tagaaggtga	gggggtcttg	ctgccagcgg	tcccagtcga	7740

[0068]

gctgcagggc gaggtcgcgc gcggcggtga ccaggcgctc gtcgcccccg aatttcatga	7800
ccagcatgaa gggcacgagc tgctttccga aggcccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaag aactggatct	7920
cccgccacca gttggaggag tggctgttga tgtggtggaa gtagaagtcc cgtcgccggg	7980
ccgaacactc gtgctggctt ttgtaaaagc gagcgcagta ctggcagcgc tgcacgggct	8040
gtacctcatg cacgagatgc acctttcgcc cgcgcacgag gaagccgagg ggaaatctga	8100
gccccccgcc tggctcgcgg catggctggt tctctttctac tttaggatgcg tgtccgtctc	8160
cgtctggctc ctcgaggggt gttacggtgg agcggaccac cacgccgcgc gagccgcagg	8220
tccagatata ggccgcgcgc ggtcggagtt tgatgacgac atcgcgcagc tgggagctgt	8280
ccatggtctg gagctccgcg gcggcgcgca ggtcagccgg gagttcttgc aggttcacct	8340
cgcagagtcg ggccagggcg cggggcaggt ctaggtggta cctgatctct aggggcgtgt	8400
tggtggcggc gtcgatggct tgcaggagcc cgcagccccg gggggcgacg acggtgcccc	8460
gcgggggtggt ggtggtggtg gcggtgcagc tcagaagcgg tgccgcgggc gggcccccg	8520
aggtaggggg ggctccggtc ccgcgggcag gggcggcagc ggcacgtcgg cgtggagcgc	8580
gggcaggagt tggtgctgtg cccggaggtt gctggcgaag gcgacgacgc ggcggttgat	8640
ctcctggatc tggcgcctct gcgtgaagac gacgggcccg gtgagcttga acctgaaaga	8700
gagttcgaca gaatcaatct cggtgtcatt gaccgcggcc tggcgcagga tctcctgcac	8760
gtctcccagag ttgtcttggg aggcgatctc ggccatgaac tgctcgatct cttcctcctg	8820
gaggtctccg cgtccggcgc gttccacggt ggccgccagg tcgttgaga tgcgccccat	8880
gagctgcgag aaggcgttga gtccgccctc gttccagact cggctgtaga ccacgcccc	8940
ctggatcatc cgggcgcgca tgaccacctg cgcgaggttg agctccacgt gccgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggtggcgg tgtgctcggc	9060
cacgaagaag ttcattgacc agcggcgcaa cgtggattcg ttgatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggcgaagtig aaaaactggg agttgcgcgc	9180
cgacacggtc aactcctcct ccagaagacg gatgagctcg gcgacggtgt cgcgcacctc	9240
gcgctcgaag gctatgggga tctcttctc cgtatgcata accacctcct cctcttctc	9300
ctcttctggc acttccatga ttgcttctc ctcttcgggg ggtggcggcg gcggcggttg	9360
gggagggggc gctctgcgcc ggccggcgcg caccgggagg cggtccacga agcgcgcgat	9420

[0069]

catctccccg cggcgggcggc gcatggtctc ggtgacggcg cggccgttct cccggggggcg	9480
cagttggaag acgccgccgg acatctggtg ctggggcggg tggccgtgag gcagcgagac	9540
ggcgctgacg atgcatctca acaattgctg cgtaggtacg ccgccgaggg acctgagggg	9600
gtccatatcc accggatccg aaaacctttc gaggaaggcg tctaaccagt cgcagtcgca	9660
aggtaggctg agcaccgtgg cgggcggcgg ggggtggggg gagtgtctgg cggaggtgct	9720
gctgatgatg taattgaagt aggcggactt gacacggcgg atggtcgaca ggagcaccat	9780
gtccttgggt ccggcctgct ggatgcggag gcggtcggct atgccccagg ctctgttctg	9840
gcatcggcgc aggtccttgt agtagtcttg catgagcctt tccaccggca cctcttctcc	9900
ttctctttct gcttcttcca tgtctgtctt ggccctgggg cggcgccgcg cccccctgcc	9960
cccatgcmc gtgaccccg accccctgag cggttggagc agggccaggt cggcgacgac	10020
gcgctcggcc aggatggcct gctgcacctg cgtgagggtg gtttggaagt catccaagtc	10080
cacgaagcgg tggtaggcgc ccgtgttgat ggtgtaggtg cagttggcca tgacggacca	10140
gttgacggtc tggtagcccg gttgcgacat ctcggtgtac ctgagtcgcg agtaggcgcg	10200
ggagtcgaag acgtagtcgt tgcaagtccg caccaggtac tggtagccca ccaggaagtg	10260
cggcgggcggc tggcggtaga ggggccagcg cagggtggcg ggggctccgg gggccaggtc	10320
ttccagcatg aggcgggtgt aggcgtagat gtacctggac atccaggtga taccgcggc	10380
ggtggtggag gcgcgcggga agtcgcgcac ccggttcag atgttgcgca ggggcagaaa	10440
gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactctagac	10500
cagggaaaac gaaagccgt cagcgggcac tcttccgtgg tctggtgaat agatcgcaag	10560
ggtatcatgg cggagggcct cggttcgagc cccgggtccg ggccggacgg tccgcatga	10620
tccacgcggt taccgcccgc gtgtcgaacc cagggtgtcg acgtcagaca acggtggagt	10680
gttctttttg gcgtttttct ggccgggcgc cggcgccgcg taagagacta agccgcgaaa	10740
gcgaaagcag taagtggctc gctccccgta gccggaggga tccttgctaa gggttgcgtt	10800
gcggcgaacc ccggttcgaa tcccgtactc gggccggccg gaccgcggc taagtggtg	10860
gattggcctc cccctcgat aaagaccccg cttgcggatt gactccggac acggggacga	10920
gcccctttta tttttgcttt cccagatgc atccggtgct gcggcagatg cccccccgc	10980
cccagcagca gcaacaacac cagcaagagc ggcagcaaca gcagcgggag tcatgcaggg	11040
ccccctcacc caccctcggc gggccggcca cctcggcgtc cgcggccgtg tctggcgctt	11100
gcggcgggcg cggggggccg gctgacgacc ccgaggagcc cccgcggcgc agggccagac	11160

[0070]

actacctgga cctggaggag ggcgagggcc tggcgcggtt gggggcgccg tctcccagac	11220
gccacccgcg ggtgcagctg aagcgcgact cgcgcgaggc gtacgtgcct cggcagaacc	11280
tgttcaggga ccgcgcgggc gaggagcccc aggagatgcg ggacaggagg ttcagcgcag	11340
ggcgggagct gcggcagggg ctgaaccgcg agcggctgct gcgcgaggag gactttgagc	11400
ccgacgcgcg gacggggatc agccccgcgc gcgcgcacgt ggcgcccgcc gacctggtga	11460
cggcgtagca gcagacggtg aaccaggaga tcaacttcca aaagagtttc aacaaccacg	11520
tgcgcacgct ggtggcgcg caggaggtga ccatcggtt gatgcacctg tgggactttg	11580
taagcgcgct ggtgcagaac cccaacagca agcctctgac ggcgagctg ttcctgatag	11640
tgcagcacag cagggacaac gaggcgttta gggacgcgct gctgaacatc accgagcccc	11700
agggtcgggtg gctgctggac ctgattaaca tcctgcagag catagtgggtg caggagcgca	11760
gcctgagcct ggccgacaag gtggcggcca tcaactactc gatgctgagc ctgggcaagt	11820
tttacgcgcg caagatctac cagacgccgt acgtgcccat agacaaggag gtgaagatcg	11880
acggttttta catgcgcatg gcgctgaagg tgctcaccct gagcgacgac ctgggcgtgt	11940
accgcaacga gcgcatccac aaggccgtga gcgtgagccg gcggcgcgag ctgagcgacc	12000
gcgagctgat gcacagcctg cagcgggcgc tggcgggcgc cggcagcggc gacagggagg	12060
cggagtccta cttegatgcg ggggcggacc tgcgctgggc gccagccgg cgggccctgg	12120
aggccgcggg ggtccgcgag gactatgacg aggacggcga ggaggatgag gagtacgagc	12180
tagaggaggg cgagtacctg gactaaaccg cgggtggtgt ttccggtaga tgcaagaccc	12240
gaacgtggtg gacccggcgc tgcgggcggc tctgcagagc cagccgtccg gccttaactc	12300
ctcagacgac tggcgacagg tcatggaccg catcatgtcg ctgacggcgc gtaaccggga	12360
cgcgttccgg cagcagccgc aggccaacag gctctccgcc atcctggagg cgggtggtgcc	12420
tgcgcgctcg aacccccgc acgagaaggt gctggccata gtgaacgcgc tggccgagaa	12480
cagggccatc cggccggacg aggccgggct ggtgtacgac gcgctgctgc agcgcgtggc	12540
ccgctacaac agcggcaacg tgcagaccaa cctggaccgg ctggtggggg acgtgcgcga	12600
ggcgggtggcg cagcgcgagc gcgcggatcg gcagggcaac ctgggctcca tgggtggcgt	12660
gaatgccttc ctgagcacgc agccggccaa cgtgcccgcg gggcaggaag actacaccaa	12720
ctttgtgagc gcgctgcggc tgatggtgac cgagaccccc cagagcgagg tgtaccagtc	12780
gggcccggac tactttcttc agaccagcag acagggcctg cagacggtga acctgagcca	12840

[0071]

ggctttcaag aacctgcggg ggctgtgggg cgtgaaggcg cccaccggcg accgggcgac	12900
ggtgtccagc ctgctgacgc ccaactcgcg cctgctgctg ctgctgatcg cgccgttcac	12960
ggacagcggc agcgtgtccc gggacaccta cctggggcac ctgctgaccc tgtaccgca	13020
ggccatcggg caggcgcagg tggacgagca caccttcag gagatcacca gcgtgagccg	13080
cgcgctgggg caggaggaca cgagcagcct ggaggcgact ctgaactacc tgctgaccaa	13140
ccggcggcag aagattccct cgctgcacag cctgacctcc gaggaggagc gcatcttgcg	13200
ctacgtgcag cagagcgtga gcctgaacct gatgcgcgac ggggtgacgc ccagcgtggc	13260
gctggacatg accgcgcgca acatggaacc gggcatgtac gccgcgcacc ggccttacat	13320
caaccgcctg atggactacc tgcctgcgcg ggcggccgtg aaccccgagt actttaccaa	13380
cgccatcctg aacccgcact ggctcccgcc gcccggttc tacagcgggg gcttcgaggt	13440
cccggagacc aacgatggct tcctgtggga cgacatggac gacagcgtgt tctccccgcg	13500
gccgcaggcg ctggcggaaag cgtccctgct gcgtcccaag aaggaggagg aggaggaggc	13560
gagtcgccgc cgcggcagca gcggcgtggc ttctctgtcc gagctggggg cggcagccgc	13620
cgcgcgcccc gggtccctgg gcggcagccc ctttccgagc ctggtggggg ctctgcacag	13680
cgagcgcacc acccgccctc ggctgctggg cgaggacgag tacctgaata actccctgct	13740
gcagccggtg cgggagaaaa acctgcctcc cgccttcccc aacaacggga tagagagcct	13800
ggtggacaag atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcctgcgct	13860
ccggccgccc acgcgcgcc agcgccacga ccggcagcgg gggctggtgt gggatgacga	13920
ggactccgcg gacgatagca gcgtgctgga cctgggaggg agcggcaacc cgttcgcgca	13980
cctgcgcccc cgcctgggga ggatgtttta aaaaaaaaaa aaaaaagcaa gaagcatgat	14040
gcaaaaatta aataaaactc accaaggcca tggcgaccga gcgttggttt cttgtgttcc	14100
cttcagtatg cggcgcgcgg cgatgtacca ggagggacct cctccctctt acgagagcgt	14160
ggtgggcgcg gcggcgcgcg cgcctcttcc tccctttgcg tcgcagctgc tggagccgcc	14220
gtacgtgcct ccgcgtacc tgcggcctac gggggggaga aacagcatcc gttactcgga	14280
gctggcgcgc ctgttcgaca ccaccgggt gtacctggtg gacaacaagt cggcggacgt	14340
ggcctccctg aactaccaga acgaccacag caatTTTTTg accacggtca tccagaacaa	14400
tgactacagc ccgagcgagg ccagcaccca gaccatcaat ctggatgacc ggtcgcactg	14460
ggcgcgcgac ctgaaaacca tcctgcacac caacatgccc aacgtgaacg agttcatgtt	14520
caccaataag ttcaaggcgc ggggtgatggt gtcgcgctcg cacaccaagg aagaccgggt	14580

[0072]

ggagctgaag tacgagtggg tggagttcga gctgccagag ggcaactact ccgagaccat	14640
gaccattgac ctgatgaaca acgcgatcgt ggagcactat ctgaaagtgg gcaggcagaa	14700
cggggtcctg gagagcgaca tcgggggtcaa gttcgacacc aggaacttcc gcctggggct	14760
ggacccccgtg accgggctgg ttatgcccgg ggtgtacacc aacgaggcct tccatcccga	14820
catcatcctg ctgcccggct gcgggggtgga cttcacttac agccgcctga gcaacctcct	14880
gggcatccgc aagcggcagc ccttccagga gggcttcagg atcacctacg aggacctgga	14940
ggggggcaac atccccgcgc tctcgtatgt ggaggcctac caggatagct tgaaggaaaa	15000
tgaggcggga caggaggata ccgccccgc cgcctccgcc gccgccgagc agggcgagga	15060
tgctgctgac accgcggccg cggacggggc agaggccgac cccgctatgg tgggtggaggc	15120
tcccagacag gaggaggaca tgaatgacag tgcggtgcgc ggagacacct tcgtcaccgc	15180
gggggaggaa aagcaagcgg aggccgaggc cgcggccgag gaaaagcaac tggcggcagc	15240
agcggcggcg gcggcggttg ccgcggcgga ggctgagtct gaggggacca agcccgccaa	15300
ggagcccgtg attaagcccc tgaccgaaga tagcaagaag cgcagttaca acctgtctaa	15360
ggacagcacc aacaccgcgt accgcagctg gtacctggcc tacaactacg gcgaccgcgc	15420
gacgggggtg cgctcctgga ccctgctgtg cacgccggac gtgacctgcg gctcggagca	15480
ggtgtactgg tcgtgccccg acatgatgca agaccccgtg accttccgct ccacgcggca	15540
ggtcagcaac ttcccgggtg tgggcgccga gctgctgccc gtgcactcca agagcttcta	15600
caacgaccag gccgtctact cccagctcat ccgccagttc acctctctga cccacgtggt	15660
caatcgcttt cctgagaacc agattctggc gcgcccgcgc gccccacca tcaccaccgt	15720
cagtgaaaac gttcctgctc tcacagatca cgggacgcta ccgctgcgca acagcatcgg	15780
aggagtccag cgagtgaccg ttactgacgc cagacgccgc acctgcccct acgtttacaa	15840
ggccttgggc atagtctcgc cgcgcgtcct ttccagccgc actttttgag caacaccacc	15900
atcatgtcca tctgatctc acccagcaat aactccggct ggggactgct gcgcgcgcgc	15960
agcaagatgt tcggaggggc gaggaagcgt tccgagcagc accccgtgcg cgtgcgcggg	16020
cacttccgcg ccccttgggg agcgcacaaa cgcggccgcg cggggcgcac caccgtggac	16080
gacgccatcg actcggtggt ggagcaggcg cgcaactaca ggccccgggt ctctaccgtg	16140
gacgcggcca tccagaccgt ggtgcggggc gcgcggcggt acgccaagct gaagagccgc	16200
cggaagcgcg tggccccgcg ccaccgccgc cgacccgggg ccgccgcca acgcgcgcgc	16260

[0073]

gcggccctgc	ttcgccgggc	caagcgcacg	ggccgccgcg	ccgccatgag	ggccgcgcgc	16320
cgcttggccg	ccggcatcac	cgccgccacc	atggccccc	gtaccgaag	acgcgcggcc	16380
gccgccgcg	ccgccccat	cagtgcacatg	gccagcaggc	gccggggcaa	cgtgtactgg	16440
gtgcgcgact	cggtgaccgg	cacgcgcgtg	cccgtgcgt	tccgcccc	gcggacttga	16500
gatgatgtga	aaaaacaaca	ctgagtctcc	tgctgtttgt	tgtatcccag	cggcggcggc	16560
gcgcgcagcg	tcatgtccaa	gcgcaaaatc	aaagaagaga	tgctccaggt	cgtcgcgccg	16620
gagatctatg	ggccccgaa	gaaggaagag	caggattcga	agccccgcaa	gataaagcgg	16680
gtcaaaaaga	aaaagaaaga	tgatgacgat	gccgatgggg	aggtggagtt	cctgcgcgcc	16740
acggcgccca	ggcgcccggt	gcagtgggaag	ggccggcgcg	taaagcgcgt	cctgcgcccc	16800
ggcaccgcgg	tggtcttcac	gcccggcgag	cgctccaccc	ggactttcaa	gcgcgtctat	16860
gacgaggtgt	acggcgacga	agacctgctg	gagcaggcca	acgagcgcct	cggagagttt	16920
gcttacggga	agcgtcagcg	ggcgtgggg	aaggaggacc	tgctggcgct	gccgctggac	16980
cagggaacc	ccacccccag	tctgaagccc	gtgacctgc	agcaggtgct	gccgagcagc	17040
gcacctccg	aggcgaagcg	gggtctgaag	cgcgaggcg	gcgacctggc	gcccaccgtg	17100
cagctcatgg	tgcccaagcg	gcagaggctg	gaggatgtgc	tggagaaaat	gaaagtagac	17160
cccgtctgc	agccggacat	cagggtccgc	cccatcaagc	aggtggcgcc	gggcctcggc	17220
gtgcagaccg	tggacgtggt	catccccacc	ggcaactccc	ccgccgcgc	caccactacc	17280
gctgcctcca	cggacatgga	gacacagacc	gatcccgccg	cagccgcage	cgcagccgcc	17340
gccgcgacct	cctcggcgga	ggtgcagacg	gaccttggc	tgccgccggc	gatgtcagct	17400
ccccgcgcgc	gtcgcggcg	caggaagtac	ggcgccgcca	acgcgctcct	gcccagtagc	17460
gccttgcatc	cttccatcgc	gcccaccccc	ggctaccgag	gctataccta	ccgcccgcga	17520
agagccaagg	gttccacccg	ccgtccccgc	cgacgcgccg	ccgccaccac	ccgccgcgc	17580
cgcgcagac	gccagccgc	actggctcca	gtctccgtga	ggaaagtggc	gcgcgacgga	17640
cacacctgg	tgctgccag	ggcgcgtac	cacccagca	tcgtttaaaa	gcctgtttgt	17700
gttcttgtag	atatggccct	cacttgccgc	ctccgtttcc	cggtgccggg	ataccgagga	17760
ggaagatcgc	gccgcaggag	gggtctggcc	ggccgcggcc	tgagcggagg	cagccgcgc	17820
gcgcaccggc	ggcgacgcgc	caccagccga	cgcatgcgcg	gcggggtgct	gcccctgtta	17880
atccccctga	tcgcccggc	gateggcgcc	gtgcccggga	tcgcctccgt	ggccttgcaa	17940
gcgtcccaga	ggcattgaca	gacttgcaaa	cttgcaaaata	tggaaaaaaa	aacccaata	18000

[0074]

aaaaagtcta gactctcacg ctcgcttggc cctgtgacta tttttagtaa tggaagacat	18060
caacttttgcg tcgctggccc cgcgtcacgg ctcgcgcccg ttcctgggac actggaacga	18120
tatcggcacc agcaacatga gcggtggcgc cttcagttgg ggctctctgt ggagcggcat	18180
taaaagtatc gggctctgccg ttaaaaatta cggctcccgg gcctggaaca gcagcacggg	18240
ccagatgttg agagacaagt tgaaagagca gaacttccag cagaaggtgg tggagggcct	18300
ggcctccggc atcaacgggg tggtggacct ggccaaccag gccgtgcaga ataagatcaa	18360
cagcagactg gacccccggc cgcgggtgga ggaggtgccg ccggcgctgg agacggtgtc	18420
ccccgatggg cgtggcgaga agcgcgcgcg gcccgatagg gaagagacca ctctggtcac	18480
gcagaccgat gagccgcccc cgtatgagga ggccctgaag caaggtctgc ccaccacgcg	18540
gcccacgcg cccatggcca ccgggggtgtt gggccgccac acccccgcca cgctggactt	18600
gcctccgccc gccgatgtgc cgcagcagca gaaggcggca cagccgggccc cgcccgcgac	18660
cgcctccgtt tcctccgcg gtcctctgcg ccgcgcggcc agcggccccc gcgggggggt	18720
cgcgaggcac ggcaactggc agagcacgct gaacagcatc gtgggtcttg ggggtgcgtc	18780
cgtgaagcgc cgccgatgct actgaatagc ttagctaacg tgttgtatgt gtgtatgcgc	18840
cctatgtgc cgccagagga gctgctgagt cgccgcggtt cgcgcgccca ccaccaccgc	18900
cactccgccc ctcaagatgg cgaccccatc gatgatgccg cagtggctgt acatgcacat	18960
ctcgggccag gacgcctcgg agtacctgag ccccgggctg gtgcagttcg cccgcgccac	19020
cgagagctac ttcagcctga gtaacaagtt taggaacccc acggtggcgc ccacgcacga	19080
tgtgaccacc gaccggtctc agcgcctgac gctgcggttc attcccgtgg accgcgagga	19140
caccgcgtac tcgtacaagg cgcggttcac cctggccgtg ggcgacaacc gcgtgctgga	19200
catggcctcc acctactttg acatccgcgg ggtgctggac cggggtccca ctttcaagcc	19260
ctactctggc accgcctaca actccctggc cccaagggc gctcccaact cctgcgagtg	19320
ggagcaagag gaaactcagg cagttgaaga agcagcagaa gaggaagaag aagatgctga	19380
cggtaagct gaggaagagc aagcagctac caaaaagact catgtatatg ctcaggctcc	19440
cctttctggc gaaaaaatta gtaaagatgg tctgcaaata ggaacggacg ctacagctac	19500
agaacaaaaa cctattttatg cagaccctac attccagccc gaaccccaaa tcggggagtc	19560
ccagtggaat gaggcagatg ctacagtcgc cggcggtaga gtgctaaaga aatctactcc	19620
catgaaacca tgctatggtt cctatgcaag acccacaat gctaattggag gtcagggtgt	19680

[0075]

actaacggca aatgcccagg gacagctaga atctcaggtt gaaatgcaat tcttttcaac	19740
ttctgaaaac gcccgtaacg aggctaacaa cattcagccc aaatttggtgc tgtatagtga	19800
ggatgtgcac atggagaccc cggatacgca cttttcttac aagcccgcaa aaagcgatga	19860
caattcaaaa atcatgctgg gtcagcagtc catgcccac agacctaat acatcggett	19920
cagagacaac tttatcggcc tcatgtatta caatagcact ggcaacatgg gagtgcctgc	19980
aggtcaggcc tctcagttga atgcagtggt ggacttgcaa gacagaaaca cagaactgtc	20040
ctaccagctc ttgcttgatt ccatgggtga cagaaccaga tacttttcca tgtggaatca	20100
ggcagtggac agttatgacc cagatgttag aattattgaa aatcatggaa ctgaagacga	20160
gctccccaac tattgtttcc ctctgggtgg cataggggta actgacactt accaggctgt	20220
taaaaccaac aatggcaata acgggggcca ggtgacttgg acaaaagatg aaacttttgc	20280
agatcgcaat gaaatagggg tgggaaacaa tttcgctatg gagatcaacc tcagtgccaa	20340
cctgtggaga aacttcctgt actccaacgt ggcgctgtac ctaccagaca agcttaagta	20400
caaccctcc aatgtggaca tctctgacaa cccaacacc tacgattaca tgaacaagcg	20460
agtgggtggc cggggctgg tggactgcta catcaacctg ggcgcgcgt ggtcgtgga	20520
ctacatggac aacgtcaacc cttcaacca ccaccgcaat gcgggcctgc gctaccgctc	20580
catgctcctg ggcaacgggc gctacgtgcc cttccacatc caggtgcccc agaagttctt	20640
tgccatcaag aacctcctcc tcctgcggg ctctacacc tacgagtgga acttcaggaa	20700
ggatgtcaac atggtcctcc agagctctct gggtaacgat ctcagggtgg acggggccag	20760
catcaagttc gagagcatct gcctctacgc cactttcttc cccatggccc acaacacggc	20820
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20880
ctccgcccgc aacatgctct accccatacc cgccaacgcc accaacgtcc ccatctccat	20940
cccctcgcgc aactgggcgg cttccgcgg ctgggccttc acccgctca agaccaagga	21000
gacccctcc ctgggctcgg gattcgaccc ctactacacc tactcgggt ccattcccta	21060
cctggacggc accttctacc tcaaccacac tttcaagaag gtctcgggtca cttcgactc	21120
ctcggtcagc tggccgggca acgaccgtct gtcaccccc aacgagttcg agatcaagcg	21180
ctcggtcgac ggggagggt acaacgtggc ccagtgcac atgaccaagg actggttcct	21240
ggtccagatg ctggccaact acaacatcgg ctaccagggc ttctacatcc cagagagcta	21300
caaggacagg atgtactcct tcttcaggaa cttccagccc atgagccggc aggtggtgga	21360
ccagaccaag tacaaggact accaggaggt gggcatcatc caccagcaca acaactcggg	21420

[0076]

cttcgtgggc tacctcgccc ccaccatgcg cgagggacag gcctaccccg ccaacttccc	21480
ctatccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tcctctgcga	21540
ccgcaccctc tggcgcatcc ctttctccag caacttcattg tccatgggtg cgctctcgga	21600
cctggggccag aacttgctct acgccaactc cgcccacgcc ctcgacatga ccttcgaggt	21660
cgaccccatg gacgagccca cctttctcta tgttctgttc gaagtctttg acgtggtccg	21720
ggtccaccag ccgcaccgcg gcgtcatcga gaccgtgtac ctgcgtacgc ctttctcggc	21780
cggcaacgcc accacctaaa gaagcaagcc gcagtcacgc ccgcctgcat gccgtcgggt	21840
tccaccgagc aagagctcag ggccatcgct agagacctgg gatgcggggc ctattttttg	21900
ggcaccttcg acaagcgctt ccttggtttt gtctccccac acaagctggc ctgcgccatc	21960
gtcaacacgg ccggccgcga gaccgggggc gtgcaactggc tggccttcgc ctggaacccg	22020
cgctccaaaa catgcttctt ctttgacccc ttcggttttt cggaccagcg gctcaagcaa	22080
atctacgagt tcgagtacga gggcttgctg cgtcgcagcg ccatgcctc ctgccccgac	22140
cgtgcgtca ccctcgaaaa gtccaccag accgtgcagg ggcccgaact ggccgcctgc	22200
ggtctcttct gctgcatgtt tctgcacgcc tttgtgcaact ggctcagag tcccatggac	22260
cgcaacccca ccatgaactt gctgacgggg gtgcccact ccatgctcca gagccccag	22320
gtcgagccca cctgcgcgc caaccaggag cagctctaca gcttcctgga gcgccactcg	22380
ccttacttcc gccgccacag cgcacagatc aggagggcca cctccttctg ccacttgcaa	22440
gagatgcaag aagggttaata acgatgtaca cacttttttt ctcaataaat ggcatctttt	22500
tatttataca agctctctgg ggtattcatt tcccaccacc acccgccgtt gtcgccatct	22560
ggctctattt agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcagggac	22620
acgttgcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tcggggaagt tttcgctcca caggctgcgg gtcagcacca gcgcgttcat caggtcgggc	22740
gccgagatct tgaagtcgca gttggggccg ccgccctgcg cgcgcgagtt gcggtacacc	22800
gggttgccagc actggaacac caacagcgcc ggggtgcttca cgctggccag cacgctgcgg	22860
tcggagatca gctcggcgct caggctcctc gcgttgctca gcgcgaacgg ggtcatcttg	22920
ggcacttgcc gccccaggaa gggcgcgtgc cccggtttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccgtg cccggactcg gcgttgggggt acagcgcgcg catgaaggcc	23040
tgcacttggc ggaaggccat ctgggccttg gcgccctccg agaagaacat gccgcaggac	23100

[0077]

ttgccccgaga actggtttgc ggggcagctg gcgtcgtgca ggcagcagcg cgcgtcgggtg	23160
ttggcgatct gcaccacgtt gcgccccac cggttcttca cgatcttggc cttggacgat	23220
tgtctcttca gcgcgcgctg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttcacca tgctgctgcc gtgcagacac ttcagctcgc cctccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtgggctcg aaagacttgt aggtcacctc cgcaaggac	23400
tgcaggtacc cctgcaaaaa gcggcccatc atggtcacga aggtcttggt gctgctgaag	23460
gtcagctgca gcccgcgggtg ctctcgttc agccaggtct tgcacacggc cgccagcgcc	23520
tccacctggt cgggcagcat cttgaagtcc accttcagct cattctccac gtggtacttg	23580
tccatcagcg tgcgcgccgc ctccatgccc ttctcccagg ccgacaccag cggcaggtc	23640
acggggttct tcaccatcac cgtggccgcc gcctccgccg cgcttctcgt ttccgccccg	23700
ctgttctctt cctcttcctc ctcttcctcg ccgccgccca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttg cgcttgccgt tgcgcccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcac agcgcggcct cttcttgctc gtctcgtg	23880
tccagaatga cctccgggga ggggggggtg gtcactctca gtaccgaggc acgcttcttt	23940
ttcttcttg gggcgttcgc cagctccgcg gctgcggccg ctgccgaggt cgaaggccga	24000
gggctgggcg tgcgcggcac cagcgcgtcc tgcgagccgt cctcgtcctc ctcgactcg	24060
agacggaggc gggcccgtt ctccgggggc gcgcggggcg gcggaggcgg cggcggcgac	24120
ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgcc gcgtccgcgc	24180
tcgggggtg tctcgcgtg gtctcttcc cgactggcca tctccactg ctcttctcc	24240
tataggcaga aagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgccccct ctgagccctc caccaccgcc gccaccaccg ccaatgccgc cgcggacgac	24360
gcgccaccg agaccaccgc cagtaccacc ctccccagcg acgcaccccc gctcgagaat	24420
gaagtgtga tcgagcagga cccgggtttt gtgagcggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggtcgc cgcctcagtg ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata aggatgagac agcagtcggg cgggggaacg gaagccatga tgctgatgac	24600
ggctacctag acgtgggaga cgacgtgtg ctttaagcacc tgcaccgcca gtgcgtcatc	24660
gtctgcgacg cgtgcagga gcgtgcgaa gtgcccttg acgtggcgga ggtcagccgc	24720
gcctacgagc ggcacctctt cgcgccgcac gtgccccca agcgcgggga gaacggcacc	24780
tgcgagccca acccgcgtct caacttctac ccggtcttcg cggtagccga ggtgctggcc	24840

[0078]

acctaccaca tctttttcca aaactgcaag atccccctct cctgccgcgc caaccgcacc	24900
cgcgcgcgaca aaaccctgac cctgcggcag ggcgcccaca tacctgatat cgcctctctg	24960
gaggaagtgc ccaagatctt cgagggtctc ggtcgcgacg agaaacgggc ggcgaaacgt	25020
ctgcacggag acagcgaaaa cgagagtcac tcgggggtgc tggtaggagct cgagggcgac	25080
aacgcgcgcc tggccgtact caagcgcagc atagaggtea cccactttgc ctaccggcg	25140
ctcaacctgc cccccaaggt catgagtgtg gtcatgggcg agctcatcat gcgcgcgcc	25200
cagccccctgg ccgcggatgc aaacttgcaa gagtcctccg aggaaggcct gcccgcggtc	25260
agcgacgagc agctggcgcg ctggctggag acccgcgacc ccgcgcagct ggaggagcgg	25320
cgcaagctca tgatggccgc ggtgctggtc accgtggagc tcgagtgtct gcagcgcttc	25380
ttcgcggacc ccgagatgca gcgcaagctc gaggagacct tgcactacac cttccgccag	25440
ggctacgtgc gccaggcctg caagatctcc aacgtggagc tctgcaacct ggtctcctac	25500
ctgggcatcc tgcacgagaa ccgcctcggg cagaacgtcc tgcactccac cctcaaagg	25560
gaggcgcgcc gcgactacat ccgcgactgc gcctacctct tctctgcta cacctggcag	25620
acggccatgg gggctctggca gcagtgcctg gaggagcgca acctcaagga gctggaaaag	25680
ctcctcaagc gcacctcag ggacctctgg acgggcttca acgagcgctc ggtggccgcc	25740
gcgctggcgg acatcatctt tcccgagcgc ctgctcaaga ccctgcagca gggcctgccc	25800
gacttcacca gccagagcat gctgcagaac ttcaggactt tcatcctgga gcgctcgggc	25860
atcctgccgg ccacttgctg cgcgctgccc agcgacttcg tgcccatcaa gtacaggag	25920
tgcccgcgcg cgctctgggg ccactgtac ctcttcagc tggccaacta cctcgcctac	25980
cactcggacc tcatggaaga cgtgagcggc gagggcctgc tcgagtgcc a ctgccgtgc	26040
aacctctgca cgccccaccg ctctctagtc tgcaaccgc agctgctcag cgagagtcag	26100
attatcggt a cttcgagct gcagggtccc tcgcctgacg agaagtccgc ggctccagg	26160
ctgaaactca ctccggggct gtggacttcc gcctacctac gcaaatttgt acctgaggac	26220
taccacgccc acgagatcag gttctacgaa gaccaatccc gcccgcccaa ggccgagctc	26280
accgcctgcg tcatcaccca ggggcacatc ctgggccaat tgcaagccat caacaaagcc	26340
cgcgcgagagt tcttgctgaa aaagggtcgg ggggtgtacc tggaccccca gtccggcgag	26400
gagctaaacc cgctaccccc gccgcgcgcc cagcagcggg accttgcttc ccaggatggc	26460
accagaaaag aagcagcagc cgccgcgcgc gccgcagcca tacatgcttc tggaggaaga	26520

[0079]

ggaggaggac tgggacagtc aggcagagga ggtttcggac gaggagcagg aggagatgat	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgcaaca ccatcgccct cggtcgcagc cccctcgccg gggccctga aatcctccga	26700
accagcacc agcgctataa cctccgctcc tccggcgccg gcgccacccg cccgcagacc	26760
caaccgtaga tgggacacca caggaaccgg ggtcggtaag tccaagtgcc cgccgcccgc	26820
accgcagcag cagcagcagc agcgccaggg ctaccgctcg tggcgcgggc acaagaacgc	26880
catagtcgcc tgcttgcaag actgcggggg caacatctct ttcgcccgc gcttcctgct	26940
attccaccac ggggtcgctt ttccccgcaa tgtcctgcat tactaccgtc atctctacag	27000
cccctactgc agcggcgacc cagaggcggc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatc ctccgcgggc aagacagcgg cagcagcggc caggagaccc gcggcagcag	27120
cggcgggagc ggtgggcgca ctgcgcctct cgcccaacga acccctctcg acccgggagc	27180
tcagacacag gatcttcccc actttgtatg ccatcttcca acagagcaga ggccaggagc	27240
aggagctgaa aataaaaaac agatctctgc gtcctctcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagcttcgg cgcacgctgg aggacgcgga ggcactcttc agcaaatact	27360
gcgcgctcac tcttaaagac tagctccgcg cccttctcga atttaggcgg gagaaaacta	27420
cgtcatcgcc ggccgcccgc cagcccgcgc agccgagatg agcaaagaga ttcccacgcc	27480
atacatgtgg agctaccagc cgcagatggg actcgcggcg ggagcggccc aggactactc	27540
cacccgatg aactacatga gcgcgggacc ccacatgatc tcacaggtca acgggatccg	27600
cgcccagcga aaccaaatac tgctggaaca ggcggccatc accgccacgc cccgccataa	27660
tctcaacccc cgaaattggc ccgcgcctt cgtgtaccag gaaacccct ccgccaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcag gggcgagct	27780
cgcggggcgc tttcgtcacg gggcgcgcc gctccgacca ggtataagac acctgatgat	27840
cagaggccga ggtatccagc tcaacgacga gtcggtgagc tcttcgctcg gtctccgtcc	27900
ggacggaact ttccagctcg ccgatccgg ccgctcttcg ttcacgcccc gccaggcgta	27960
cctgactctg cagacctcgt cctcggagcc ccgctccggc ggcacggaa ccctccagtt	28020
cgtggaggag ttcgtgccct cggctctactt caaccccttc tcgggacctc ccggacgcta	28080
ccccgaccag ttcatccga actttgacgc ggtgaaggac tcggcgagc gctacgactg	28140
aatgtcaggt gtcgaggcag agcagcttcg cctgagacac ctcgagcact gccgcccca	28200
caagtgcctt gcccgcggtt ctggtgagtt ctgctacttt cagctacccg aggagcatac	28260

[0080]

cgaggggccc	gcgcacggcg	tccgcctgac	cacccagggc	gaggttacct	gttccctcat	28320
ccgggagttt	accctccgtc	ccctgctagt	ggagcgggag	cggggtcctt	gtgtcctaac	28380
tatcgccctgc	aactgcccta	accctggatt	acatcaagat	ctttgctgtc	atctctgtgc	28440
tgagtttaat	aaacgctgag	atcagaatct	actggggctc	ctgtcgccat	cctgtgaacg	28500
ccaccgtctt	cacccacccc	gaccaggccc	aggcgaacct	cacctgcggt	ctgcatcgga	28560
gggccaagaa	gtacctcacc	tggtaactta	acggcacccc	ctttgtggtt	tacaacagct	28620
tcgacgggga	cggagtctcc	ctgaaagacc	agctctccgg	tctcagctac	tccatccaca	28680
agaacaccac	cctccaactc	tccccctcct	acctgccggg	aacctacgag	tgcgtcaccg	28740
gccgtgcac	ccacctcacc	cgcctgatcg	taaaccagag	ctttccggga	acagataact	28800
ccctcttccc	cagaacagga	ggtgagctca	ggaaactccc	cggggaccag	ggcggagacg	28860
taccttcgac	ccttgtgggg	ttaggatttt	ttattaccgg	gttgcctggc	cttttaatac	28920
aagtttcctt	gagatttggt	ctttccttct	acgtgtatga	acacctcaac	ctccaataac	28980
tctacccttt	cttcggaatc	aggtgacttc	tctgaaatcg	ggcttggtgt	gctgcttact	29040
ctgttgattt	ttttccttat	catactcagc	cttctgtgcc	tcaggctcgc	cgcctgctgc	29100
gcacacatct	atatctactg	ctggttgctc	aagtgcaggg	gtcgccaccc	aagatgaaca	29160
ggtacatggt	cctatcgatc	ctaggcctgc	tggccctggc	ggcctgcagc	gccgccaaaa	29220
aagagattac	ctttgaggag	cccgttgca	atgtaacttt	caagcccagag	ggtgaccaat	29280
gcaccaccct	cgtcaaatgc	gttaccaatc	atgagaggct	gcgcatcgac	tacaaaaaca	29340
aaactggcca	gtttgcggtc	tatagtgtgt	ttacgcccg	agaccctct	aactactctg	29400
tcaccgtctt	ccagggcgga	cagtctaaga	tattcaatta	cactttccct	ttttatgagt	29460
tatgcgatgc	ggtcatgtac	atgtcaaaac	agtacaacct	gtggcctccc	tctccccagg	29520
cgtgtgtgga	aaatactggg	tcttactgct	gtatggcttt	cgcaatcact	acgctcgctc	29580
taatctgcac	ggtgctatac	ataaaattca	ggcagaggcg	aatctttatc	gatgaaaaga	29640
aaatgccttg	atcgctaaca	ccggctttct	atctgcagaa	tgaatgcaat	cacctcccta	29700
ctaataccca	ccacctcctt	tgcgattgcc	catgggttga	cacgaatcga	agtgccagtg	29760
gggtccaatg	tcaccatggt	gggccccgcc	ggcaattcca	ccctcatgtg	ggaaaaat	29820
gtccgcaatc	aatgggttca	tttctgctct	aaccgaatca	gtatcaagcc	cagagccatc	29880
tgcgatgggc	aaaatctaac	tctgatcaat	gtgcaaatga	tggatgctgg	gtactattac	29940

[0081]

gggcagcggg gagaaatcat taattactgg cgaccccaca aggactacat gctgcatgta	30000
gtcgaggcac ttcccactac ccccccaact accacctctc ccaccaccac caccactact	30060
actactacta ctactactac tactactacc actaccgctg cccgccatac ccgcaaaagc	30120
accatgatta gcacaaagcc ccctcgtgct cactcccacg ccggcggggc catcggtgcg	30180
acctcagaaa ccaccgagct ttgcttctgc caatgacta acgccagcgc tcatgaactg	30240
ttcgacctgg agaatgagga tgtccagcag agctccgctt gcctgaccga ggaggctgtg	30300
gagcccgttg ccctgaagca gatcggtgat tcaataattg actcttcttc ttttgccact	30360
cccgaatacc ctcccgattc tactttccac atcacgggta ccaaagacct taacctctct	30420
ttctacctga tgctgctgct ctgtatctct gtggtctctt ccgcgctgat gttactgggg	30480
atgttctgct gcctgatctg ccgcagaaag agaaaagctc gctctcaggg ccaaccactg	30540
atgcccttcc cctaccccc ggattttgca gataacaaga tatgagctcg ctgctgacac	30600
taaccgcttt actagcctgc gctctaacct ttgtcgcttg cgactcgaga ttccacaatg	30660
tcacagctgt ggcaggagaa aatgttactt tcaactccac ggccgatacc cagtggctgt	30720
ggagtggctc aggtagctac ttaactatct gcaatagctc cacttcccc ggcatatccc	30780
caaccaagta ccaatgcaat gccagcctgt tcacctcat caacgcttcc accctggaca	30840
atggactcta tgtaggctat gtaccctttg gtgggcaagg aaagaccac gcttacaacc	30900
tggaagttcg ccagcccaga accactacct aagcttctcc caccaccacc accaccacca	30960
ccatcaccag cagcagcagc agcagcagcc acagcagcag cagcagatta ttgactttgg	31020
ttttggccag ctcatctgcc gctaccagg ccattctacg ctctgtgccc gaaaccactc	31080
agatccaccg ccagaaaacg accaccgcca ccacctaca cacctccagc gatcagatgc	31140
cgaccaacat cacccttg gctcttcaaa tgggacttac aagccccact ccaaaaccag	31200
tggatgcggc cgaggtctcc gccctcgtca atgactgggc ggggctggga atgtggtggt	31260
tcgccatagg catgatggcg ctctgcctgc ttctgctctg gctcatctgc tgcctccacc	31320
gcaggcgagc cagaccccc atctatagac ccatcattgt cctgaacccc gataatgatg	31380
ggatccatag attggatggc ctgaaaaacc tacttttttc ttttacagta tgataaattg	31440
agacatgcct cgcattttct tgtacatgtt ctttctccca cttttcttg ggtgttctac	31500
gttggccgct gtgtctcacc tggaggtaga ctgcctctca cccttactg tctacctgct	31560
ttacggattg gtcacctca ctctcatctg cagcctaate acagtaatca tgccttcat	31620
ccagtgcatt gattacatct gtgtgcgcct cgcatacttc agacaccacc cgcagtaccg	31680

[0082]

agacaggaac attgccaac ttctaagact gctctaataca tgcataagac tgtgatctgc	31740
cttctgatcc tctgcatcct gccaccctc acctcctgcc agtacaccac aaaatctccg	31800
cgcaaaagac atgcctcctg ccgcttcacc caactgtgga atatacccaa atgctacaac	31860
gaaaagagcg agctctccga agcttggctg tatgggggtca tctgtgtctt agttttctgc	31920
agcactgtct ttgccctcat aatctacccc tactttgatt tgggatggaa cgcgatcgat	31980
gccatgaatt accccacctt tcccgacccc gagataattc cactgcgaca agttgtaccc	32040
gttgtcgtaa atcaacgccc cccatccctt acgcccactg aaatcagcta ctttaacctt	32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatcagta ccgagcagcg	32160
tctcctagag aggcgcaggc aggcggctga gcaagagcgc ctcaatcagg agctccgaga	32220
tctcgtaa ac ctgcaccagt gcaaaagagg catcttttgt ctggtaaagc aggccaaagt	32280
cacctacgag aagaccggca acagccaccg cctcagttac aaattgccc cccagcgcca	32340
gaagctggtg ctcatggtgg gtgagaatcc catcacgctc acccagcact cggtagagac	32400
cgaggggtgt ctgcactccc cctgtcgggg tccagaagac ctctgcaccc tggtaaagac	32460
cctgtgcggt ctgagagatt tagtcccctt taactaatca aacactggaa tcaataaaaa	32520
gaatcactta cttaaaatca gacagcaggt ctctgtccag ttatttcagc agcacctcct	32580
tcccctcctc ccaactctgg tactccaaac gccttctggc ggcaaacttc ctccacaccc	32640
tgaaggaat gtcagattct tgctcctgtc cctccgcacc cactatcttc atgttgttgc	32700
agatgaagcg caccaaaacg tctgacgaga gcttcaaccc cgtgtacccc tatgacacgg	32760
aaagcggccc tccctccgtc cctttcctca cccctccctt cgtgtctccc gatggattcc	32820
aagaaagtcc ccccggggtc ctgtctctga acctggccga gcccttggtc acttcccacg	32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgctggc aacctcacct	32940
ctcaagatat caccaccgct agccctcccc tcaaaaaaac caagaccaac ctgagcctag	33000
aaacctcatc ccccctaact gtgagcacct caggcgccct caccgtagca gccgcgctc	33060
ccctggcggt ggccggcacc tccctcacca tgcaatcaga ggccccctg acagtacagg	33120
atgcaaaact caccctggcc accaaaggcc cctgaccgt gtctgaagge aaactggcct	33180
tgcaaacatc ggccccgctg acggccgctg acagcagcac cctcacagtc agtgccacac	33240
cacccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc atttacacca	33300
ccaatggaaa actaggactt aactttggcg ctcccctgca tgtggtagac agcctaaatg	33360

[0083]

cactgactgt agttactggc caaggtctta cgataaacgg aacagcccta caaactagag	33420
tctcaggtgc cctcaactat gacacatcag gaaacctaga attgagagct gcagggggta	33480
tgcgagttga tgcaaatggt caacttatcc ttgatgtagc ttaccatttt gatgcacaaa	33540
acaatctcag ccttaggctt ggacagggac ccctgtttgt taactctgcc cacaacttgg	33600
atgttaacta caacagaggc ctctacctgt tcacatctgg aaataccaaa aagctagaag	33660
ttaatataca aacagccaag ggtctcattt atgatgacac tgctatagca atcaatgcgg	33720
gtgatgggct acagtttgac tcaggctcag atacaaatcc attaaaaact aaacttggat	33780
taggactgga ttatgactcc agcagagcca taattgctaa actgggaact ggcctaagct	33840
ttgacaacac aggtgccatc acagtaggca acaaaaatga tgacaagctt accttgtgga	33900
ccacaccaga cccatcccct aactgtagaa tctattcaga gaaagatgct aaattcacac	33960
ttgttttgac taaatgcggc agtcaggtgt tggccagcgt ttctgtttta tctgtaaaag	34020
gtagccttgc gccatcagt ggcacagtaa ctagtgetca gattgtcctc agatttgatg	34080
aaaatggagt tctactaagc aattcttccc ttgacctca atactggaac tacagaaaag	34140
gtgaccttac agagggcact gcatatacca acgcagtggg atttatgccc aacctcacag	34200
catacccaaa aacacagagc caaactgcta aaagcaacat tgtaagtcag gtttacttga	34260
atggggacaa atccaaaccc atgaccctca ccattaccct caatggaact aatgaaacag	34320
gagatgccac agtaagcaat tactccatgt cattctcatg gaactggaat ggaagtaatt	34380
acattaatga aacgttccaa accaactcct tcaccttctc ctacatcgcc caagaataaa	34440
aagcatgacg ctgttgattt gattcaatgt gtttctgttt tattttcaag cacaacaaaa	34500
tcattcaagt cattcttcca tcttagctta atagacacag tagcttaata gaccagtag	34560
tgcaaagccc cattctagct tatagatcag acagtataa ttaaccacca ccaccacat	34620
accttttgat tcaggaaatc atgatcatca caggatccta gtcttcaggc cgccccctcc	34680
ctcccaagac acagaatata cagtctctc ccccgactg gctttaaata acaccatctg	34740
gttggtcaca gacatgttct taggggttat attccacacg gtctcctgcc gcgccaggcg	34800
ctcgtcggtg atgttgataa actctcccgg cagctcgctc aagttcacgt cgctgtccag	34860
cggctgaacc tccggctgac gcgataactg tgcgaccggc tgctggacga acggaggccg	34920
cgcttacaag ggggtagagt cataatctc ggtcaggata gggcggtgat gcagcagcag	34980
cgagcgaaac atctgtgcc gccgccctc cgtccggcag gaaaacaaca cgccggtggt	35040
ctcctccgcg ataatccgca ccgcccgcag catcagcttc ctcgttctcc gcgcgcagca	35100

[0084]

cctcaccctt atctcgctca aatcggcgca gtaggtacag cacagcacca cgatgttatt	35160
catgatccca cagtgccagg cgctgtatcc aaagctcatg ccgggaacca ccgccccac	35220
gtggccatcg taccacaagc gcacgtaaat caagtgtcga cccctcatga acgcgctgga	35280
cacaaacatt acttccttgg gcatgttgta attcaccacc tcccgggtacc agataaacct	35340
ctggttgaac agggcacctt ccaccacat cctgaaccaa gaggccagaa cctgcccacc	35400
ggctatgcac tgcagggaac ccgggttgga acaatgacaa tgcagactcc aaggctcgta	35460
accgtggatc atccggctgc tgaaggcatc gatgttggca caacacagac acacgtgcat	35520
gcactttctc atgattagca gctcttcct cgtcaggatc atatcccaag gaataacca	35580
ttcttgaatc aacgtaaaac ccacacagca gggaaggcct cgcacataac tcacgttgtg	35640
catggtcagc gtgttgcatt ccggaaacag cggatgatcc tccagtatcg aggcgcgggt	35700
ctccttctca caggaggta aagggtccct gctgtacgga ctgcgccggg acgaccgaga	35760
tcgtgttgag cgtagtgtca tggaaaaggg aacgccggac gtggtcatac ttcttgaagc	35820
agaaccaggt tcgcgcgtgg caggcctcct tgcgtctgcg gtctcgccgt ctagctcgct	35880
ccgtgtgata gttgtagtac agccactccc gcagagcgtc gaggcgcacc ctggcttccg	35940
gatctatgta gactccgtct tgcaccgcgg ccctgataat atccaccacc gtagaataag	36000
caacaccag ccaagcaata cactcgctct gcgagcggca gacaggagga gcgggcagag	36060
atgggagaac catgataaaa aacttttttt aaagaatatt ttccaattct tcgaaagtaa	36120
gatctatcaa gtggcagcgc tcccctccac tggcgcggtc aaactctacg gccaaagcac	36180
agacaacggc atttctaaga tgttccttaa tggcgtccaa aagacacacc gctctcaagt	36240
tgcagtaaac tatgaatgaa aaccatccg gctgattttc caatatagac gcgccggcag	36300
cgtccaccaa accagataa ttttcttctc tccagcgggt tacgatctgt ctaagcaa	36360
cccttatatc aagtccgacc atgccaaaaa tctgtctcaag agcgccctcc accttcatgt	36420
acaagcagcg catcatgatt gcaaaaattc aggttcttca gagacctgta taagattcaa	36480
aatgggaaca ttaacaaaaa ttctctgtc gcgcagatcc cttegcaggg caagctgaac	36540
ataatcagac aggtccgaac ggaccagtga ggccaaatcc ccaccaggaa ccagatccag	36600
agaccctata ctgattatga cgcgcatact cggggctatg ctgaccagcg tagcgccgat	36660
gtaggcgtgc tgcattggcg gcgagataaa atgcaaagtg ctggttaaaa aatcaggcaa	36720
agcctcgcgc aaaaaagcta acacatcata atcatgctca tgcaggtagt tgcaggtaag	36780

[0085]

ctcaggaacc	aaaacggaat	aacacacgat	tttcctctca	aacatgactt	cgcgatact	36840
gcgtaaaaca	aaaaattata	aataaaaaat	taattaaata	acttaaacad	tggaagcctg	36900
tctcacaaca	ggaaaaacca	ctttaatcaa	cataagacgg	gccacgggca	tgccggcata	36960
gccgtaaaaa	aattggtccc	cgtgattaac	aagtaccaca	gacagctccc	cggtcacgtc	37020
gggggtcatc	atgtgagact	ctgtatacac	gtctggattg	tgaacatcag	acaaacaaag	37080
aaatcgagcc	acgtagcccc	gaggtataat	cacccgcagg	cggaggtaca	gcaaaacgac	37140
ccccatagga	ggaatcacaa	aattagtagg	agaaaaaaat	acataaacac	cagaaaaacc	37200
ctgtttgctg	ggcaaaatag	cgcctccccg	atccaaaaca	acataaagcg	cttcacacgg	37260
agcagccata	acaaagaccc	gagtcttacc	agtaaaagaa	aaaagatctc	tcaacgcagc	37320
accagcacca	acacttcgca	gtgtaaaagg	ccaagtgccg	agagagtata	tataggaata	37380
aaaagtgcag	taaacgggca	aagtccaaaa	aacgcccaga	aaaaccgcac	gcgaacctac	37440
gccccgaaac	gaaagccaaa	aaacactaga	cactcccttc	cgcgctcaac	ttccgctttc	37500
ccacgctacg	tcacttgccc	cagtcaaaca	aactacatat	cccgaacttc	caagtcgcca	37560
cgccccaaac	accgcctaca	cctccccgcc	cgccggcccc	cccccaaacc	cgcctccccg	37620
cccgcgcccc	gccccgcgcc	gcccattctc	ttatcatatt	ggcttcaatc	caaaataagg	37680
tatattattg	atgatggttt	aaacggatcc	tctagagtcg	acctgcaggc	atgcaagctt	37740
gagtattcta	tagtgtcacc	taaatagctt	ggcgtaatca	tggtcatagc	tgtttctgt	37800
gtgaaattgt	tatccgtcca	caattccaca	caacatacga	gccggaagca	taaagtgtaa	37860
agcctggggg	gcctaagtga	tgagctaact	cacattaatt	gcgttgccgt	cactgccccg	37920
tttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	atcggccaac	gcgaaccctt	37980
tgcgcccgcc	cgggccgtcg	accaattctc	atgtttgaca	gcttatcatc	gaatttctgc	38040
cattcatccg	cttattatca	cttattcagg	cgtagcaacc	aggcgtttaa	gggcaccaat	38100
aactgcctta	aaaaaattac	gccccgcctt	gccactcatc	gcagtactgt	tgtaattcat	38160
taagcattct	gccgacatgg	aagccatcac	aaacggcatg	atgaacctga	atcgccagcg	38220
gcacagcac	cttgtgcctt	tgcgataaat	atttgcccat	ggtgaaaacg	ggggcgaaga	38280
agttgtccat	attggccacg	tttaaatcaa	aactggtgaa	actcaccag	ggattggctg	38340
agacgaaaaa	catattctca	ataaaccttt	tagggaaata	ggccaggttt	tcaccgtaac	38400
acgccacatc	ttgcgaatat	atgtgtagaa	actgccggaa	atcgctgtgg	tattcactcc	38460
agagcgatga	aaacgtttca	gtttgctcat	ggaaaacggt	gtaacaaggg	tgaacactat	38520

[0086]

cccatatcac	cagctcacgc	tctttcattg	ccatacggaa	ttccggatga	gcattcatca	38580
ggcgggcaag	aatgtgaata	aaggccggat	aaaactttgtg	cttatttttc	tttacggctct	38640
ttaaaaaggc	cgtaatatcc	agctgaacgg	tctggttata	ggtacattga	gcaactgact	38700
gaaatgcctc	aaaatgttct	ttacgatgcc	attgggatat	atcaacgggtg	gtatatccag	38760
tgattttttt	ctccatttta	gcttccttag	ctcctgaaaa	tctcgataac	tcaaaaaata	38820
cgccccgtag	tgatcttatt	tcattatggt	gaaagttgga	acctcttacg	tgccgatcaa	38880
cgtctcattt	tcgcaaaaag	ttggcccagg	gcttcccggg	atcaacaggg	acaccaggat	38940
ttattttatt	tgcaagtga	tcttccgtca	caggatatta	ttcgcgataa	gctcatggag	39000
cggcgtaacc	gtcgcacagg	aaggacagag	aaagcgcgga	tctgggaagt	gacggacaga	39060
acggtcagga	cctggattgg	ggaggcgggt	gccgccgctg	ctgctgacgg	tgtgacgttc	39120
tctgttccgg	tcacaccaca	tacgttccgc	cattcctatg	cgatgcacat	gctgtatgcc	39180
ggtataccgc	tgaaagttct	gcaaagcctg	atgggacata	agtccatcag	ttcaacggaa	39240
gtctacacga	aggtttttgc	gctggatgtg	gctgcccggc	accgggtgca	gtttgcgatg	39300
ccggagtctg	atgcggttgc	gatgctgaaa	caattatcct	gagaataaat	gccttggcct	39360
ttatatggaa	atgtggaact	gagtggatat	gctgtttttg	tctgttaaac	agagaagctg	39420
gctgttatcc	actgagaagc	gaacgaaaca	gtcgggaaaa	tctccatta	tcgtagagat	39480
ccgcattatt	aatctcagga	gcctgtgtag	cgtttatagg	aagtagtggt	ctgtcatgat	39540
gcctgcaagc	ggtaacgaaa	acgatttgaa	tatgccttca	ggaacaatag	aaatcttcgt	39600
gcggtgttac	gttgaagtgg	agcggattat	gtcagcaatg	gacagaacaa	cctaatgaac	39660
acagaacat	gatgtggtct	gtccttttac	agccagtagt	gctcgccgca	gtcgagcgac	39720
agggcgaagc	cctcgagtga	gcgaggaagc	accagggaac	agcacttata	tattctgctt	39780
acacacgatg	cctgaaaaaa	cttcccttgg	ggttatccac	ttatccacgg	ggatattttt	39840
ataattatatt	tttttatagt	tttttagatct	tcttttttag	agcgccttgt	aggcctttat	39900
ccatgctggt	tctagagaag	gtgttgtgac	aaattgccct	ttcagtgtga	caaatacccc	39960
tcaaatgaca	gtcctgtctg	tgacaaattg	cccttaaccc	tgtgacaaat	tgccctcaga	40020
agaagctggt	ttttcacaaa	gttatccctg	cttattgact	cttttttatt	tagtgtgaca	40080
atctaaaaac	ttgtcacact	tcacatggat	ctgtcatggc	ggaaacagcg	gttatcaatc	40140
acaagaaacg	taaaaatagc	ccgcgaatcg	tccagtcaaa	cgacctcact	gaggcggcat	40200

[0087]

atagtctctc ccgggatcaa aaacgtatgc tgtatctgtt cgttgaccag atcagaaaat	40260
ctgatggcac cctacaggaa catgacggta tctgcgagat ccatgttgct aaatatgctg	40320
aaatattegg attgacctct gcggaagcca gtaaggatat acggcaggca ttgaagagtt	40380
tcgcggggaa ggaagtgggt ttttatcgcc ctgaagagga tgccggcgat gaaaaaggct	40440
atgaatcttt tccttgggtt atcaaacgtg cgcacagtcc atccagaggg ctttacagtg	40500
tacatatcaa cccatatctc attcccttct ttatcgggtt acagaaccgg ttacgcagt	40560
ttcggccttag tgaaacaaaa gaaatcacca atccgtatgc catgcgttta tacgaatccc	40620
tgtgtcagta tcgtaagccg gatggctcag gcatcgtctc tctgaaaate gactggatca	40680
tagagcgta ccagctgcct caaagttacc agcgtatgcc tgacttccgc cgccgcttcc	40740
tgcaggtctg tgtaaatgag atcaacagca gaactccaat gcgcctctca tacattgaga	40800
aaaagaaagg ccgccagacg actcatatcg tattttcctt ccgcgatatc acttccatga	40860
cgacaggata gtctgagggt tatctgtcac agatttgagg gtggttcgtc acatttgttc	40920
tgacctactg agggtaatth gtcacagttt tgctgtttcc ttcagcctgc atggattttc	40980
tcatactttt tgaactgtaa tttttaagga agccaaattt gagggcagtt tgtcacagtt	41040
gatttccttc tctttccctt cgtcatgtga cctgatatcg ggggttagtt cgtcatcatt	41100
gatgagggtt gattatcaca gtttattact ctgaattggc tatccgcgtg tgtacctcta	41160
cctggagttt ttcccacggt ggatatttct tcttgcgtg agcgtaagag ctatctgaca	41220
gaacagttct tctttgcttc ctcgccagtt cgctcgtat gctcggttac acggctgcgg	41280
cgagcgctag tgataataag tgactgaggt atgtgctctt cttatctcct tttgtagtgt	41340
tgtctttatt ttaaacaact ttgcggtttt ttgatgactt tgcgattttg ttgttgcttt	41400
gcagtaaatt gcaagattta ataaaaaac gcaaagcaat gattaaagga tgttcagaat	41460
gaaactcatg gaaacactta accagtgcac aaacgtgggt catgaaatga cgaaggctat	41520
cgccattgca cagtttaatg atgacagccc ggaagcgagg aaaataaccc ggcgctggag	41580
aataggtgaa gcagcggatt tagttggggt ttcttctcag gctatcagag atgccgagaa	41640
agcagggcga ctaccgcacc cggatatgga aattcgagga cgggttgagc aacgtgttgg	41700
ttatacaatt gaacaaatta atcatatgcg tgatgtgttt ggtacgcgat tgcgacgtgc	41760
tgaagacgta tttccaccgg tgatcggggt tgctgcccac aaaggtggcg tttaaaaaac	41820
ctcagtttct gttcatcttg ctcaggatct ggctctgaag gggctacgtg ttttgctcgt	41880
ggaaggtaac gacccccagg gaacagcctc aatgtatcac ggatgggtac cagatcttca	41940

[0088]

tattcatgca gaagacactc tcctgccttt ctatcttggg gaaaaggacg atgtcactta	42000
tgcaataaag cccacttgct ggccggggct tgacattatt ccttcctgtc tggctctgca	42060
ccgtattgaa actgagttaa tgggcaaatt tgatgaaggt aaactgcca cccatccaca	42120
cctgatgctc cgactggcca ttgaaactgt tgctcatgac tatgatgtca tagttattga	42180
cagcgcgcct aacctgggta tcggcacgat taatgtcgta tgtgctgctg atgtgctgat	42240
tgttcccacg cctgctgagt tgtttgacta cacctccgca ctgcagtttt tcgatatgct	42300
tcgtgatctg ctcaagaacg ttgatcttaa agggttcgag cctgatgtac gtattttgct	42360
taccaaatac agcaatagta atggctctca gtccccgtgg atggaggagc aaattcggga	42420
tgccctgggga agcatggttc taaaaaatgt tgtacgtgaa acggatgaag ttggtaaagg	42480
tcgatccgg atgagaactg tttttgaaca ggccattgat caacgctctt caactgggtgc	42540
ctggagaaat gctctttcta tttgggaacc tgtctgcaat gaaattttcg atcgtctgat	42600
taaaccacgc tgggagatta gataatgaag cgtgcgcctg ttattccaaa acatacgtc	42660
aatactcaac cggttgaaga tacttcgtta tcgacaccag ctgccccgat ggtggattcg	42720
ttaattgcgc gcgtaggagt aatggctcgc ggtaatgcca ttactttgcc tgtatgtggt	42780
cgggatgtga agtttactct tgaagtgtc cggggtgata gtgttgagaa gacctctcg	42840
gtatggtcag gtaatgaacg tgaccaggag ctgcttactg aggacgcact ggatgatctc	42900
atcccttctt ttctactgac tggtaacag acaccggcgt tcggtcgaag agtatctggt	42960
gtcatagaaa ttgccgatgg gagtcgccgt cgtaaagctg ctgcacttac cgaaagtgat	43020
tatcgtgttc tggttggcga gctggatgat gagcagatgg ctgcattatc cagattgggt	43080
aacgattatc gccaacaag tgcttatgaa cgtggtcagc gttatgcaag ccgattgcag	43140
aatgaatttg ctggaaatat ttctgcgctg gctgatgcgg aaaatatattc acgtaagatt	43200
attaccgcgt gtatcaacac cgccaaattg cctaaatcag ttgttgctct tttttctcac	43260
cccgtgaac tatctgcccg gtcaggatgat gcacttcaaa aagcctttac agataaagag	43320
gaattactta agcagcaggc atctaacctt catgagcaga aaaaagctgg ggtgatattt	43380
gaagctgaag aagttatcac tcttttaact tctgtgctta aaacgtcatc tgcataaga	43440
actagtttaa gctcacgaca tcagtttgct cctggagcga cagtattgta taaggcgat	43500
aaaatgggtgc ttaacctgga caggctctgt gttccaactg agtgtataga gaaaattgag	43560
gccattctta aggaacttga aaagccagca ccctgatgcg accacgtttt agtctacgtt	43620

[0089]

tatctgtctt tacttaatgt cctttgttac aggccagaaa gcataactgg cctgaatatt	43680
ctctctgggc ccaactgttcc acttgtatcg tcggtctgat aatcagactg ggaccacggt	43740
cccactcgta tcgtcggctc gattattagt ctgggaccac ggtcccactc gtatcgtcgg	43800
tctgattatt agtctgggac cacgggtccca ctctgtatcg cggctcgata atcagactgg	43860
gaccacggtc ccaactcgat cgtcggctcg attattagtc tgggaccatg gtcccactcg	43920
tatcgtcggt ctgattatta gtctgggacc acgggtccac tcgtatcgtc ggtctgatta	43980
ttagtctgga accacgggtcc cactcgtatc gtcggtctga ttattagtct gggaccacgg	44040
tcccactcgt atcgtcggtc tgattattag tctgggacca cgatcccact cgtgttgcg	44100
gtctgattat cggctctggga ccacgggtccc acttgtattg tcgatcagac tatcagcgtg	44160
agactacgat tccatcaatg cctgtcaagg gcaagtattg acatgctgc gtaacctgta	44220
gaacggagta acctcgggtg gcggttgtat gcctgctgtg gattgctgct gtgtcctgct	44280
tatccacaac attttgcgca cggttatgtg gacaaaatac ctggttacc aggccgtgcc	44340
ggcacgttaa ccgggctgca tccgatgcaa gtgtgtcgtc gtcgacgagc tcgcgagctc	44400
ggacatgagg ttgccccgta ttcagtgtcg ctgatttgta ttgtctgaag ttgtttttac	44460
gttaagttga tgcagatcaa ttaatacgat acctgcgtca taattgatta tttgacgtgg	44520
tttgatggcc tccacgcacg ttgtgatatg tagatgataa tcattatcac tttacgggtc	44580
ctttccggtg atccgacagg ttacggggcg gcgacctcgc gggttttcgc tatttatgaa	44640
aattttccgg ttttaaggcgt ttccgttctt ctctgtcata acttaatgtt tttatttaaa	44700
ataccctctg aaaagaaagg aaacgacagg tgctgaaagc gagctttttg gcctctgtcg	44760
tttcctttct ctgtttttgt ccgtggaatg aacaatggaa gtccgagctc atcgctaata	44820
acttcgtata gcatacatta tacgaagtta tattcgatgc ggccgcaagg ggttcgcgtc	44880
agcgggtgtt ggcggtgtc ggggctggct taactatgcg gcatcagagc agattgtact	44940
gagagtgcac catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat	45000
caggcgccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc	45060
ttcgtatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac	45120
gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtgaattgt aatacgactc	45180
actatagggc gaattcgagc tcggtaccg gggatcctcg tttaaac	45227

<210> 10

<211> 37830

[0090]

<212> DNA

<213> 黑猩猩腺病毒

<400> 10

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcggga ggcgggtccg ggggcgggcc ggcgggcggg gcggtgtggc ggaagtggac	120
tttgtaagtg tggcggatgt gacttgctag tgccgggcgc ggtaaaagtg acgttttccg	180
tgcgcgacaa cgcccacggg aagtgcatt tttccgcgg tttttaccgg atgtttagt	240
gaatttgggc gtaaccaagt aagatttggc cattttcgc ggaaaactga aacggggaag	300
tgaaatctga ttaatttcgc gttagtcata ccgcgtaata ttgtcgagg gccgaggac	360
tttggccgat tacgtggagg actcggccag gtgttttttg aggtgaattt ccgcgttccg	420
ggtcaaagtc tccgttttat tattatagtc agctgacgcg gagtgtattt atacctctg	480
atctcgtcaa gtggccactc ttgagtcca gcgagtagag tttctcctc tgccgctctc	540
cgtccgctc cgtcggctc tgacaccggg gaaaaaatga gacatttcac ctacgatggc	600
ggtgtgctca ccggccagct ggctgctgaa gtcctggaca ccctgatcga ggaggtattg	660
gccgataatt atcctccctc gactcctttt gagccaccta cacttcacga actctacgat	720
ctggatgtgg tggggcccag cgatccgaac gagcaggcgg tttccagttt ttttccagag	780
tccatgttgt tggccagcca ggagggggtc gaacttgaga cccctcctcc gatcgtggat	840
tccccgatc cgccgcagct gactaggcag cccgagcgt gtgcgggacc tgagactatg	900
ccccagctgc tacctgaggt gatcgatctc acctgtaatg agtctggttt tccaccagc	960
gaggatgagg acgaagagg tgagcagttt gtgttagatt ctgtggaaca acccgggcga	1020
ggatgcaggt cttgtcaata tcaccgaaa aacacaggag actcccagat tatgtgttct	1080
ctgtgttata tgaagatgac ctgtatgttt atttacagta agtttatcat ctgtgggcag	1140
gtgggctata gtgtgggtgg tggcttttgg ggggtttttt aatatatgtc aggggttatg	1200
ctgaagactt ttttattgtg atttttaag gtccagtgtc tgagcccag caagaacctg	1260
aaccggagcc tgagccttct cgcccagga gaaagcctgt aatcttaact agaccagcg	1320
caccggtagc gagaggcctc agcagcgcgg agaccaccga ctccggtgct tcctcatcac	1380
ccccggagat tcacccctg gtgcccctgt gtcccgttaa gcccgttgcc gtgagagtca	1440
gtgggcggcg gtctgctgtg gactgcattg aggacttget ttttgattca caggaacctt	1500
tggacttgag cttgaaacgc ccaggcatt aaacctggtc acctggactg aatgagttga	1560
cgcctatgtt tgcttttgaa tgacttaatg tgtatagata ataaagagtg agataatgtt	1620

[0091]

ttaattgcat	ggtgtgttta	acttgggcgg	agtctgctgg	gtatataagc	ttccctgggc	1680
taaacttgg	tacacttgac	ctcatggagg	cctgggagtg	tttgagaaac	tttgccggag	1740
ttcgtgcctt	gctggacgag	agctctaaca	atacctcttg	gtgggtggagg	tatttgtggg	1800
gctctcccca	gggcaagtta	gtttgtagaa	tcaaggagga	ttacaagtgg	gaatttgaag	1860
agcttttgaa	atcctgtgg	gagctattgg	attctttgaa	tctaggccac	caggctctct	1920
tccaggagaa	ggtcatcagg	actttggatt	tttccacacc	ggggcgcatt	gcagccgcgg	1980
ttgcttttct	agcttttttg	aaggatagat	ggagcgaaga	gaccacttg	agttcgggct	2040
acgtcctgga	ttttctggcc	atgcaactgt	ggagagcatg	gatcagacac	aagaacaggc	2100
tgcaactgtt	gtcttccgtc	cgcccgttgc	tgattccggc	ggaggagcaa	caggccgggt	2160
cagaggaccg	ggcccgtcgg	gatccggagg	agagggcacc	gaggccgggc	gagaggagcg	2220
cgctgaacct	gggaaccggg	ctgagcggcc	atccacatcg	ggagtgaatg	tcgggcaggt	2280
ggtggatctt	tttccagaac	tgcggcggat	tttgactatt	agggaggatg	ggcaatttgt	2340
taagggtctt	aagagggaga	ggggggcttc	tgagcataac	gaggaggcca	gtaatttagc	2400
ttttagcttg	atgaccagac	accgtccaga	gtgcatcact	tttcagcaga	ttaaggacaa	2460
ttgtgccaat	gagttggatc	tgttgggtca	gaagtatagc	atagagcagc	tgaccactta	2520
ctggctgcag	ccgggtgatg	atctggagga	agctattagg	gtgtatgcta	aggtggccct	2580
gcggcccgat	tgcaagtaca	agctcaaggg	gctgggtgaat	atcaggaatt	gttgctacat	2640
ttctggcaac	ggggcggagg	tggagataga	gaccgaagac	agggtggctt	tcagatgcag	2700
catgatgaat	atgtggccgg	gggtgctggg	catggacggg	gtggtgatta	tgaatgtgag	2760
gttcacgggg	cccaacttta	acggcacggt	gtttttgggg	aacaccaacc	tggtcctgca	2820
cggggtgagc	ttctatgggt	ttaacaacac	ctgtgtggag	gcctggaccg	atgtgaaggt	2880
ccgcggttgc	gccttttatg	gatgttggaa	ggccatagtg	agccgcccta	agagcaggag	2940
ttccattaag	aatgcttgt	ttgagagggt	caccttgggg	atcctggccg	agggcaactg	3000
cagggtgcgc	cacaatgtgg	cctccgagtg	cggttgcttc	atgctagtca	agagcgtggc	3060
ggtaatcaag	cataatatgg	tgtgcggcaa	cagcaggagc	aaggcctcac	agatgctgac	3120
ctgcacggat	ggcaactgcc	acttgcgtgaa	gaccatccat	gtaaccagcc	acagccggaa	3180
ggcctggccc	gtgttcgagc	acaacttgct	gacccgctgc	tccttgcatc	tgggcaacag	3240
gcggggggtg	ttcctgccct	atcaatgcaa	ctttagtcac	accaagatct	tgctagagcc	3300

[0092]

cgagagcatg	tccaaggtga	acttgaacgg	ggtgtttgac	atgaccatga	agatctggaa	3360
ggtgctgagg	tacgacgaga	ccaggtcccg	gtgcagaccc	tgcgagtgcg	ggggcaagca	3420
tatgaggaac	cagcccgtga	tgctggatgt	gaccgaggag	ctgaggacag	accacttggt	3480
tctggcctgc	accagggccg	agtttggttc	tagcgatgaa	gacacagatt	gaggtgggtg	3540
agtgggcgtg	gcctgggggtg	gtcatgaaaa	tatataagtt	gggggtctta	gggtctcttt	3600
atttgtgttg	cagagaccgc	cggagccatg	agcgggagca	gcagcagcag	cagtagcagc	3660
agcgccttgg	atggcagcat	cgtgagccct	tatttgacga	cgcggatgcc	ccactgggcc	3720
gggggtgcgtc	agaatgtgat	gggtccagc	atcgacggcc	gaccgctcct	gcccgc aaat	3780
tccgccacgc	tgacctatgc	gaccgtcgcg	gggacgccgt	tggacgccac	cgccgccgcc	3840
gccgccaccg	cagccgcctc	ggccgtgcgc	agcctggcca	cggactttgc	attcctggga	3900
ccactggcga	caggggctac	ttctcgggcc	gctgctgccg	ccgttcgcga	tgacaagctg	3960
accgccctgc	tggcgcagtt	ggatgcgctt	actcgggaac	tgggtgacct	ttctcagcag	4020
gtcatggccc	tgcgccagca	ggtctcctcc	ctgcaagctg	gcgggaatgc	ttctcccaca	4080
aatgccgttt	aagataaata	aaaccagact	ctgtttggat	taaagaaaag	tagcaagtgc	4140
attgctctct	ttatttcata	attttccgcg	cgcgataggc	cctagaccag	cgttctcggt	4200
cgttgagggt	gcggtgtatc	ttctccagga	cgtggtagag	gtggctctgg	acgttgagat	4260
acatgggcat	gagcccgtcc	cgggggtgga	ggtagcacca	ctgcagagct	tcatgctccg	4320
gggtggtggt	gtagatgata	cagtcgtagc	aggagcgtg	ggcatggtgc	ctaaaaatgt	4380
ccttcagcag	caggccgatg	gccaggggga	ggcccttggt	gtaagtgttt	acaaaacggt	4440
taagttggga	agggtgcatt	cggggagaga	tgatgtgcat	cttggactgt	atttttagat	4500
tggcgatggt	tccgcccaga	tcccttctgg	gattcatggt	gtgcaggacc	accagtacag	4560
tgtatccggt	gcacttgggg	aatttgtcat	gcagcttaga	gggaaaagcg	tggaagaact	4620
tggagacgcc	tttgtggcct	cccagatttt	ccatgcattc	gtccatgatg	atggcaatgg	4680
gcccgcggga	ggcagcttgg	gcaaagatat	ttctggggtc	gctgacgtcg	tagttgtgtt	4740
ccagggtgag	gtcgtcatag	gccattttta	caaagcgcgg	gcggagggtg	cccgactggg	4800
ggatgatggt	cccctctggc	cctggggcgt	agttgccctc	gcagatctgc	atttcccagg	4860
ccttaatctc	ggagggggga	atcatatcca	cctgcggggc	gatgaagaaa	acggtttccg	4920
gagccgggga	gattaactgg	gatgagagca	ggttttctaag	cagctgtgat	tttccacaac	4980
cgggtgggccc	ataaataaca	cctataaccg	gttgcagctg	gtagtttaga	gagctgcagc	5040

[0093]

tgccgtcgtc ccggaggagg ggggccacct cgttgagcat gtccctgacg cgcatgttct	5100
ccccgaccag atccgccaga aggcgctcgc cgcccaggga cagcagctct tgcaaggaag	5160
caaagttttt cagcggcttg aggcggtccg ccgtgggcat gtttttcagg gtctggctca	5220
gcagctccag gcggtcccag agctcggtag cgtgctctac ggcatctcta tccagcatat	5280
ctcctcgttt cgcgggttgg ggcgactttc gctgtagggc accaagcggg ggtcgctccag	5340
cggggccaga gtcatgtcct tccatgggag cagggtcctc gtcagggtgg tctgggtcac	5400
ggtgaagggg tgcgctccgg gctgagcgtc tgccaagggt cgcttgaggc tggttctgct	5460
ggtgctgaag cgctgccggg cttegccctg cgcgtcggcc aggtagcatt tgaccatggt	5520
gtcatagtcc agccccctcc cggcgtgtcc cttggcgcgc agcttgccct tggaggtggc	5580
gccgcacgag gggcagagca ggctcttgag cgcgtagagc ttgggggcca ggaagaccga	5640
ttcgggggag taggcgtccg cgccgcagac cccgcacacg gtctcgact ccaccagcca	5700
ggtgagctcg gggcgcgccg ggtcaaaaac caggtttccc ccatgctttt tgatgcgttt	5760
cttacctcgg gtctccatga ggtggtgtcc ccgctcgggt acgaagaggc tgtccgtgtc	5820
tccgtagacc gacttgaggg gtcttttctc caggggggtc cctcggtctt cctcgtagag	5880
gaactcggac cactctgaga cgaaggcccg cgtccaggcc aggacgaagg aggctatgtg	5940
ggaggggtag cggtcgttgt ccactagggg gtccaccttc tccaagggtg gaagacacat	6000
gtcgcccttc tcggcgcca ggaagggtgat tggctttag gtgtaggcca cgtgaccggg	6060
ggttcctgac ggggggggtat aaaagggggg gggggcgcg cgtcgtcac tctcttcgc	6120
atcgctgtct gcgaggcca gctgctgggg taggtattcc ctctcgaagg cgggcatgac	6180
ctccgcgctg aggttgtcag tttccaaaaa cgaggaggat ttgatgttca cctgtcccga	6240
ggtgatacct ttgagggtac ccgcgtccat ctggtcagaa aacacgatct ttttattgtc	6300
cagcttgggt gcgaacgacc cgtagagggc gttggagagc agcttggcga tggagcgcag	6360
ggtctggttc ttgtccctgt cggcgcgctc cttggccgag atgttgagct gcacgtactc	6420
gcgcgcgacg cagcgccact cggggaagac ggtggtgcgc tcgtcgggca ccaggcgcac	6480
gcgccagccg cggttgtgca gggtagaccag gtccacgctg gtggcgacct cgccgcgcag	6540
gcgctcgttg gtccagcaga gacggccgcc cttgcgcgag cagaaggggg gcagggggtc	6600
gagctgggtc tcgtccgggg ggtccgcgtc cacgggtgaaa accccggggc gcaggcgcgc	6660
gtcgaagtag tctatcttgc aaccttgcat gtccagcgcc tgctgccagt cgcgggcggc	6720

[0094]

gagcgcgcgc tcgtaggggt tgagcggcgg gccccagggc atgggggtggg tgagtgcgga	6780
ggcgtacatg ccgcagatgt catagacgta gaggggctcc cgcaggaccc cgatgtaggt	6840
ggggtagcag cggccgccgc ggatgctggc gcgcacgtag tcatacagct cgtgcgaggg	6900
ggcgaggagg tcggggccca ggttggtgcg ggcggggcgc tccgcgcgga agacgatctg	6960
cctgaagatg gcatgcgagt tggaagagat ggtggggcgc tggaagacgt tgaagctggc	7020
gtcctgcagg ccgacggcgt cgcgcacgaa ggaggcgtag gagtcgcgca gcttgtgtac	7080
cagctcggcg gtgacctgca cgtcagcgc gcagtagtcg agggctctcg ggatgatgtc	7140
atattttagcc tgcccccttct ttttccacag ctgcggttg aggacaaact cttcgcggtc	7200
tttccagtac tcttgatcg ggaaaccgtc cggttccgaa cggtaagagc ctagcatgta	7260
gaactggttg acggcctggt aggcgcagca gcccttctcc acggggaggg cgtaggcctg	7320
cgcggccttg cggagcaggg tgtgggtcag ggccaaggtg tccctgacca tgactttgag	7380
gtactggtgc ttgaagtcgg agtcgtcgca gccgccccgc tcccagagcg agaagtcggt	7440
gcgcttcttg gagcgggggt tgggcagagc gaaggtgaca tcgttgaaga ggattttgcc	7500
cgcgcggggc atgaagttgc gggatgatgc gaaggcccc ggcacttcag agcggttggt	7560
gatgacctgg gcggcgagca cgatctcgtc gaagccgttg atgttgtggc ccacgatgta	7620
gagttccagg aagcggggcc ggccctttac ggtgggcagc ttcttttagct cttcgtaggt	7680
gagctcctcg ggcgaggcga ggccgtgctc ggccaggggc cagtccgcga ggtgcgggtt	7740
gtctctgagg aaggacttcc agaggtcgcg ggccaggagg gtctgcaggc ggtctctgaa	7800
ggtcctgaac tggcggccca cggccatttt ttcgggggtg atgcagtaga aggtgagggg	7860
gtcttgctgc cagcgggtccc agtcgagctg cagggcgagg tcgcgcgcgg cggtgaccag	7920
gcgctcgtcg cccccgaatt tcatgaccag catgaagggc acgagctgct ttccgaaggc	7980
ccccatccaa gtgtaggtct ctacatcgta ggtgacaaag aggcgctccg tgcgaggatg	8040
cgagccgatc gggaagaact ggatctcccg ccaccagttg gaggagtggc tgttgatgtg	8100
gtggaagtag aagtcccgtc gccgggccga aactcgtgc ttgcttttgt aaaagcgagc	8160
gcagtactgg cagcgtgca cgggctgtac ctcatgcacg agatgcacct ttcgccccgcg	8220
cacgaggaag ccgaggggaa atctgagccc ccgcctggc tcgcggcatg gctggttctc	8280
ttctactttg gatgcgtgtc cgtctccgtc tggctcctcg aggggtgtta cgggtggagcg	8340
gaccaccacg ccgcgcgagc cgcaggtcca gatatcggcg cgcggcggtc ggagtttgat	8400
gacgacatcg cgcagctggg agctgtccat ggtctggagc tcccgcggcg gcggcaggtc	8460

[0095]

agccgggaggt tcttgcaggt tcacctcgca gagtcggggcc agggcgcggg gcaggtctag	8520
gtggtacctg atctctaggg gcgtgttggg ggcggcgtcg atggcttgca ggagcccgca	8580
gccccggggg gcgacgacgg tgcccccgcg ggtgggtggg gtggtggcgg tgcagctcag	8640
aagcgggtgcc gcgggcgggc ccccgaggt agggggggct ccggtccgc gggcaggggc	8700
ggcagcggca cgtcggcgtg gacgcgggc aggagttggg gctgtgccc gaggttgctg	8760
gcgaaggcga cgacgcggcg gttgatctcc tggatctggc gcctctgcgt gaagacgacg	8820
ggcccgggtga gcttgaacct gaaagagagt tcgacagaat caatctcggt gtcattgacc	8880
gcggcctggc gcaggatctc ctgcacgtct cccgagttgt cttggtaggc gatctcggcc	8940
atgaactgct cgatctcttc ctcttgagg tctccgcgtc cggcgcgttc cacggtggcc	9000
gccaggtcgt tggagatgcg ccccatgagc tgcgagaagg cgttgagtcc gccctcgttc	9060
cagactcggc tgtagaccac gccccctgg tcatcgcggg cgcgcatgac cacctgcgcg	9120
aggttgagct ccacgtgccg cgcgaaagac gcgtagttgc gcagacgtg gaagaggtag	9180
ttgagggtgg tggcggtgtg ctcgccacg aagaagttca tgaccagcg gcgcaacgtg	9240
gattcgttga tgtccccaa ggcctccagc cgttccatgg cctcgtagaa gtccacggcg	9300
aagttgaaaa actgggagtt gcgcgccgac acggtcaact cctcctccag aagacggatg	9360
agctcggcga cgggtgtcgc cacctcgcgc tcgaaggcta tggggatctc ttcctccgct	9420
agcatcacca cctcctctc ttcctctct tctggcactt ccatgatggc ttcctctct	9480
tcgggggggtg gcggcgcgcg cggtggggga gggggcgctc tgcgccggcg gcggcgacc	9540
gggaggcggg ccacgaagcg cgcatcatc tcccccggc ggcgcgcat ggtctcggtg	9600
acggcgcggc cgttctcccg gggcgcgagt tggaagacgc cgccggacat ctggtgctgg	9660
ggcgggtggc cgtgaggcag cgagacggcg ctgacgatgc atctcaaaa ttgctgcgta	9720
ggtacccgc cgagggacct gagggagtc atatccacc gatccgaaaa ctttcgagg	9780
aaggcgtcta accagtcgca gtcgaagg aggtgagca ccgtggcggg cgcgggggg	9840
tggggggagt gtctggcgga ggtgctgct atgatgtaat tgaagtaggc ggacttgaca	9900
cggcggtagg tcgacaggag caccatgtcc ttgggtccgg cctgctggat gcggaggcgg	9960
tcggctatgc ccaggtctt gttctggcat cggcgcaggt cttttagta gtcttgcatg	10020
agcctttcca ccggcacctc ttctcttcc tctctgctt ctccatgtc tgcttcggcc	10080
ctggggcggc gccgcgccc cctgcccc atgcgcgtga cccgaacc cctgagcggg	10140

[0096]

tggagcaggg	ccaggtcggc	gacgacgcgc	tcggccagga	tggcctgctg	cacctgcgtg	10200
agggtggttt	ggaagtcac	caagtccacg	aagcgggtgt	aggcgcccg	gttgatggtg	10260
taggtgcagt	tggccatgac	ggaccagttg	acggtctggt	ggcccggttg	cgacatctcg	10320
gtgtacctga	gtcgcgagta	ggcgcgggag	tcgaagacgt	agtcgttgca	agtccgcacc	10380
aggtactggt	agcccaccag	gaagtgcggc	ggcggctggc	ggtagagggg	ccagcgcagg	10440
gtggcggggg	ctccgggggc	caggtcttcc	agcatgaggc	ggtggtaggc	gtagatgtac	10500
ctggacatcc	aggtgatacc	cgcggcgggtg	gtggaggcgc	gcgggaagtc	gcgcacccgg	10560
ttccagatgt	tgcgcagggg	cagaaagtgc	tccatggtag	gcgtgctctg	tccagtcaga	10620
cgcgcgcagt	cgttgatact	ctagaccagg	gaaaacgaaa	gccggtcagc	gggcactctt	10680
ccgtggtctg	gtgaatagat	cgcaagggtg	tcatggcgga	gggcctcggt	tcgagccccg	10740
ggtccggggc	ggacggtccg	ccatgatcca	cgcggttacc	ccccgcgtgt	cgaaccagg	10800
tgtgcgacgt	cagacaacgg	tggagtgttc	cttttggcgt	ttttctggcc	gggcgcgggc	10860
gccgcgtaag	agactaagcc	gcgaaagcga	aagcagtaag	tggtctgctc	cccgtagccg	10920
gagggatcct	tgctaagggt	tgcgttgccg	cgaaccccg	ttcgaatccc	gtactcgggc	10980
cggccggacc	cgcggctaag	gtgttggatt	ggcctcccc	tcgtataaag	accccgcttg	11040
cggattgact	ccggacacgg	ggacgagccc	cttttatitt	tgctttcccc	agatgcatcc	11100
ggtgctgcgg	cagatgcgcc	ccccgcccc	gcagcagcaa	caacaccagc	aagagcggca	11160
gcaacagcag	cgggagtcac	gcagggcccc	ctcaccacc	ctcggcgggc	cggccacctc	11220
ggcgtccgcg	gccgtgtctg	gcgcctgcgg	cggcggcggg	gggccggctg	acgaccccg	11280
ggagcccccg	cggcgcaggg	ccagacacta	cctggacctg	gaggagggcg	agggcctggc	11340
gcggctgggg	gcgccgtctc	ccgagcgcca	cccgcgggtg	cagctgaagc	gcgactcgcg	11400
cgaggcgtag	gtgcctcggc	agaacctgtt	cagggaccgc	gcggcgagg	agcccgagga	11460
gatgcgggac	aggaggttca	gcgcagggcg	ggagctgcgg	caggggctga	accgcgagcg	11520
gctgctgcgc	gaggaggact	ttgagcccga	cgcgcggacg	gggatcagcc	ccgcgcgcgc	11580
gcacgtggcg	gccgccgacc	tggtgacggc	gtacgagcag	acggtgaacc	aggagatcaa	11640
cttccaaaag	agtttcaaca	accacgtgcg	cacgtggtg	gcgcgcgagg	aggtgacct	11700
cgggctgatg	cacctgtggg	actttgtaag	cgcgtggtg	cagaacccca	acagcaagcc	11760
tctgacggcg	cagctgttcc	tgatagtgc	gcacagcagg	gacaacgagg	cgtttaggga	11820
cgcgctgctg	aacatcaccc	agcccagagg	tcggtggctg	ctggacctga	ttaacatcct	11880

[0097]

gcagagcata	gtggtgcagg	agcgcagcct	gagcctggcc	gacaaggtgg	cggccatcaa	11940
ctactcgatg	ctgagcctgg	gcaagtttta	cgcgcgcaag	atctaccaga	cgccgtacgt	12000
gccccatagac	aaggaggtga	agatcgacgg	tttttacatg	cgcatggcgc	tgaaggtgct	12060
caccctgagc	gacgacctgg	gcgtgtaccg	caacgagcgc	atccacaagg	ccgtgagcgt	12120
gagccggcgg	cgcgagctga	gcgaccgcga	gctgatgcac	agcctgcagc	gggcgctggc	12180
gggcgcgcgg	agcggcgaca	gggaggcgga	gtcctacttc	gatgcggggg	cggacctgcg	12240
ctgggcgccc	agccggcggg	ccctggaggc	cgcgggggtc	cgcgaggact	atgacgagga	12300
cggcgaggag	gatgaggagt	acgagctaga	ggaggcgag	tacctggact	aaaccgcggg	12360
tggtgtttcc	ggtagatgca	agacccgaac	gtggtggacc	cggcgctgcg	ggcggtcttg	12420
cagagccagc	cgtccggcct	taactcctca	gacgactggc	gacaggtcat	ggaccgcctc	12480
atgtcgctga	cggcgcgtaa	cccggacgcg	ttccggcagc	agccgcaggc	caacaggctc	12540
tccgccatcc	tggaggcggt	ggtgcctgcg	cgctcgaacc	ccacgcacga	gaaggtgctg	12600
gccatagtga	acgcgctggc	cgagaacagg	gccatccgcc	cggacgaggc	cgggctggtg	12660
tacgacgcgc	tgctgcagcg	cgtggcccgc	tacaacagcg	gcaacgtgca	gaccaacctg	12720
gaccggctgg	tgggggacgt	gcgcgaggcg	gtggcgagc	gcgagcgcg	ggatcggcag	12780
ggcaacctgg	gctccatggt	ggcgctgaat	gccttcctga	gcacgcagcc	ggccaacctg	12840
ccgcgggggc	aggaagacta	caccaacttt	gtgagcgcg	tgcggctgat	ggtgaccgag	12900
acccccaga	gcgaggtgta	ccagtcgggc	ccggactact	tcttcagac	cagcagacag	12960
ggcctgcaga	cgggtgaacct	gagccaggct	ttcaagaacc	tgcgggggct	gtggggcgctg	13020
aaggcgccca	ccggcgaccg	ggcgacgggtg	tccagcctgc	tgacgcccac	ctcgcgcctg	13080
ctgctgctgc	tgatcgcgcc	gttcacggac	agcggcagcg	tgtcccggga	cacctacctg	13140
gggcacctgc	tgacctgta	ccgcgaggcc	atcgggcagg	cgcagggtga	cgagcacacc	13200
ttccaggaga	tcaccagcgt	gagccgcgcg	ctggggcagg	aggacacgag	cagcctggag	13260
gcgactctga	actacctgct	gaccaaccgg	cggcagaaga	ttccctcgct	gcacagcctg	13320
acctccgagg	aggagcgcat	cttgcgctac	gtgcagcaga	gcgtgagcct	gaacctgatg	13380
cgcgacgggg	tgacgcccag	cgtggcgctg	gacatgaccg	cgcgcaacat	ggaaccgggc	13440
atgtacgcgc	cgcaccggcc	ttacatcaac	cgcctgatgg	actacctgca	tcgcgcggcg	13500
gccgtgaacc	ccgagtactt	taccaacgcc	atcctgaacc	cgcactggct	cccgcgcgcc	13560

[0098]

gggttctaca gcgggggcctt cgaggtcccg gagaccaacg atggcttcct gtgggacgac	13620
atggacgaca gcgtgtttctc cccgcggccg caggcgctgg cggaagcgtc cctgctgcgt	13680
cccaagaagg aggaggagga ggaggcgagt cgccgccgcg gcagcagcgg cgtggcttct	13740
ctgtccgagc tgggggcggc agccgccgcg cgccccgggt ccctgggcgg cagccccctt	13800
ccgagcctgg tggggtctct gcacagcgag cgcaccaccc gccctcggct gctgggcgag	13860
gacgagtacc tgaataactc cctgctgcag ccggtgcggg agaaaaacct gcctcccgcc	13920
ttccccaaca acgggataga gagcctggtg gacaagatga gcagatggaa gacctatgcg	13980
caggagcaca gggacgcgcc tgcgctccgg ccgccacgc ggcgccagcg ccacgaccgg	14040
cagcgggggc tgggtgtgga tgacgaggac tccgcggacg atagcagcgt gctggacctg	14100
ggagggagcg gcaaccggtt cgcgcacctg cgccccgcc tggggaggat gttttaaaaa	14160
aaaaaaaaaa aagcaagaag catgatgcaa aaattaaata aaactcacca aggccatggc	14220
gaccgagcgt tggtttcttg tgttcccttc agtatgcggc gcgcggcgat gtaccaggag	14280
ggacctctc cctcttacga gagcgtggtg ggcgcggcgg cggcggcgcc ctcttctccc	14340
tttgctcgc agctgctgga gccgccgtac gtgcctccgc gctacctgcg gcctacgggg	14400
gggagaaaca gcatccgta ctcggagctg gcgcccctgt tcgacaccac ccgggtgtac	14460
ctggttgaca acaagtcggc ggacgtggcc tcctgaact accagaacga ccacagcaat	14520
tttttgacca cggatcatcca gaacaatgac tacagcccga gcgaggccag caccagacc	14580
atcaatctgg atgaccggtc gcaactgggc ggcgacctga aaaccatcct gcacaccaac	14640
atgccaacg tgaacgagtt catgttcacc aataagttca aggcgcgggt gatggtgtcg	14700
cgctcgcaca ccaaggaaga ccgggtggag ctgaagtacg agtgggtgga gttcgagctg	14760
ccagagggca actactccga gaccatgacc attgacctga tgaacaacgc gatcgtggag	14820
cactatctga aagtgggcag gcagaacggg gtcttgaga gcgacatcgg ggtcaagttc	14880
gacaccagga acttccgcct ggggctggac ccctgaccg ggctggttat gcccgggtg	14940
tacaccaacg aggccttcca tcccacatc atcctgtgc ccgctgcgg ggtggacttc	15000
acttacagcc gcctgagcaa cctcctgggc atccgcaagc ggcagccctt ccaggagggc	15060
ttcaggatca cctacgagga cctggagggg ggcaacatcc ccgcgtcct cgatgtggag	15120
gcctaccagg atagcttgaa ggaaaatgag gcgggacagg aggataccgc ccccgccgcc	15180
tccgcccgcc ccgagcagg cgaggatgct gctgacaccg cgcccgcgga cggggcagag	15240
gccgaccccg ctatggtggt ggaggctccc gagcaggagg aggacatgaa tgacagtgcg	15300

[0099]

gtgcgcggag acaccttcgt caccgggggg gaggaaaagc aagcggaggc cgaggccgcg	15360
gccgaggaaa agcaactggc ggcagcagcg gcggcgggcg cgttggccgc ggcggaggct	15420
gagtcctgagg ggaccaagcc cgccaaggag cccgtgatta agcccctgac cgaagatagc	15480
aagaagcgca gttacaacct gctcaaggac agcaccaaca ccgcgtaccg cagctggtac	15540
ctggcctaca actacggcga cccgtcgacg ggggtgcgct cctggaccct gctgtgcacg	15600
ccggacgtga cctgcggctc ggagcaggtg tactggctgc tgcccacat gatgcaagac	15660
cccgtgacct tccgctccac gcggcaggtc agcaacttcc cgggtggtggg cgccgagctg	15720
ctgcccgtgc actccaagag cttctacaac gaccaggccg tctactccca gctcatccgc	15780
cagttcacct ctctgaccca cgtgttcaat cgctttcctg agaaccagat tctggcgcgc	15840
ccgcccgcgc ccaccatcac caccgtcagt gaaaacgttc ctgctctcac agatcacggg	15900
acgctaccgc tgcgcaacag catcgaggga gtccagcgag tgaccgttac tgacgccaga	15960
cgccgcacct gccctacgt ttacaaggcc ttgggcatag tctcgccgcg cgtcctttcc	16020
agccgcactt tttgagcaac accaccatca tgtccatcct gatctcacc agcaataact	16080
ccggctgggg actgctgcgc gcgcccagca agatgttcgg aggggcgagg aagcgttccg	16140
agcagcacc cgtgcgcgtg cgcgggact tccgcgccc ctggggagcg cacaacgcg	16200
gccgcgcggg gcgcaccacc gtggacgacg ccatcgactc ggtggtggag caggcgcgca	16260
actacaggcc cgcggtctct accgtggacg cggccatcca gaccgtggtg cggggcgcgc	16320
ggcggtagc caagctgaag agccgccgga agcgcgtggc ccgccccac cgccgccgac	16380
ccggggccgc cgccaaacgc gccgccgcgg ccctgttcg ccgggccaag cgcacgggcc	16440
gccgcgccgc catgagggcc gcgcgccgt tggccgccgg catcaccgc gccaccatgg	16500
ccccccgtac ccgaagacgc gcggccgccg ccgccgccgc cgccatcagt gacatggcca	16560
gcaggcgccg gggcaacgtg tactgggtgc gcgactcggt gaccggcacg cgcgtgcccg	16620
tgcgcttccg cccccgcgg acttgagatg atgtgaaaaa acaacactga gtctcctgct	16680
gttgtgtgta tcccagcggc ggcggcgcgc gcagcgtcat gtccaagcgc aaaatcaaag	16740
aagagatgct ccaggctcgt gcgccggaga tctatgggcc cccgaagaag gaagagcagg	16800
attcgaagcc ccgaagata aagcgggtca aaaagaaaaa gaaagatgat gacgatgccg	16860
atggggagggt ggagttcctg cgcgccacgg cgcgccagcg cccggtgcag tggaagggcc	16920
ggcgcgtaaa gcgcgtcctg cgccccggca ccgcggtggt cttcacgccc ggcgagcgt	16980

[0100]

ccacccggac	tttcaagcgc	gtctatgacg	aggtgtacgg	cgacgaagac	ctgctggagc	17040
aggccaacga	gcgcttcgga	gagtttgctt	acgggaagcg	tcagcgggcg	ctggggaagg	17100
aggacctgct	ggcgtgccg	ctggaccagg	gcaacccac	ccccagtctg	aagcccgtga	17160
ccctgcagca	ggtgctgccg	agcagcgcac	cctccgaggc	gaagcggggt	ctgaagcgcg	17220
agggcggcga	cctggcgccc	accgtgcagc	tcatggtgcc	caagcggcag	aggctggagg	17280
atgtgctgga	gaaaatgaaa	gtagaccccg	gtctgcagcc	ggacatcagg	gtccgccccca	17340
tcaagcaggt	ggcgccgggc	ctcggcgtgc	agaccgtgga	cgtggtcatc	cccaccggca	17400
actccccgcg	cgccgccacc	actaccgtg	cctccacgga	catggagaca	cagaccgatc	17460
ccgccgcagc	cgcagccgca	gccgccgccc	cgacctcctc	ggcggagggtg	cagacggacc	17520
cctggctgcc	gccggcgatg	tcagctcccc	gcgcgcgtcg	cgggcgcagg	aagtacggcg	17580
ccgccaacgc	gtcctgccc	gagtacgcct	tgcctccttc	catcgcgccc	acccccggct	17640
accgaggcta	tacctaccgc	ccgcgaagag	ccaagggttc	caccgcgct	ccccgccgac	17700
gcgcgcgcgc	caccacccgc	cgccgccgcc	gcagacgcca	gcccgcactg	gtccagtct	17760
ccgtgaggaa	agtggcgcg	gacggacaca	ccctggtgct	gcccaggcg	cgctaccacc	17820
ccagcatcgt	ttaaaagcct	gttgtggttc	ttgcagatat	ggccctcact	tgccgcctcc	17880
gtttcccgtg	gccgggatac	cgaggaggaa	gatcgcgccg	caggagggtg	ctggccggcc	17940
geggcctgag	cggaggcagc	cgccgcgcgc	accggcggcg	acgcgccacc	agccgacgca	18000
tgcgcggcgg	ggtgctgccc	ctgttaatcc	ccctgatcgc	cgcggcgatc	ggcgccgtgc	18060
ccgggatcgc	ctccgtggcc	ttgcaagcgt	cccagaggca	ttgacagact	tgcaaaacttg	18120
caaatatgga	aaaaaaaaacc	ccaataaaaa	agtctagact	ctcacgctcg	cttggtcctg	18180
tgactatttt	gtagaatgga	agacatcaac	tttgcgtcgc	tggccccgcg	tcacggctcg	18240
cgcgccgttc	tgggacactg	gaacgatata	ggcaccagca	acatgagcgg	tggcgcttc	18300
agttggggct	ctctgtggag	cggcattaaa	agtatcgggt	ctgccgttaa	aaattacggc	18360
tcccgggcct	ggaacagcag	cacgggccag	atgttgagag	acaagttgaa	agagcagaac	18420
ttccagcaga	agtggtgga	gggcctggcc	tccggcatca	acggggtggt	ggacctggcc	18480
aaccaggccg	tgcagaataa	gatcaacagc	agactggacc	cccggccgcc	ggtggaggag	18540
gtgccgcggg	cgttgagac	ggtgtcccc	gatgggcgtg	gcgagaagcg	cccgcggccc	18600
gatagggaag	agaccactct	ggtcacgcag	accgatgagc	cgcggcgta	tgaggaggcc	18660
ctgaagcaag	gtctgcccac	cacgcggccc	atcgcgccca	tggccaccgg	ggtggtgggc	18720

[0101]

cgccacaccc cgcacacgct ggacttgcct ccgcccgcgc atgtgccgca gcagcagaag	18780
gcggcacagc cgggccccgc cgcgaccgcc tcccgttcct ccgcccgtcc tctgcgccgc	18840
gcggccagcg gccccgcgcg gggggtcgcg aggcacggca actggcagag cacgctgaac	18900
agcatcgtgg gtctgggggt gcggtccgtg aagcgccgcc gatgctactg aatagcttag	18960
ctaactgtgt gtatgtgtgt atgcgcccta tgtcgccgcc agaggagctg ctgagtcgcc	19020
gccgttcgcg cccccaccac caccgccact ccgcccctca agatggcgac cccatcgatg	19080
atgccgcagt ggctgtacat gcacatctcg ggccaggacg cctcggagta cctgagcccc	19140
gggctgggtgc agttcgcccc gcgcccgag agctacttca gcctgagtaa caagtttagg	19200
aacccacagg tggcgcccc gcacgatgtg accaccgacc ggtctcagcg cctgacgctg	19260
cggttcattc ccgtggaccg cgaggacacc gcgtactcgt acaaggcgcg gttcaccttg	19320
gccgtgggcg acaaccgcgt gctggacatg gcctccacct actttgacat ccgcggggtg	19380
ctggaccggg gtcccacttt caagccctac tctggcaccg cctacaactc cctggcccc	19440
aaggcgctc ccaactcctg cgagtgggag caagaggaaa ctcaggcagt tgaagaagca	19500
gcagaagagg aagaagaaga tgctgacggt caagctgagg aagagcaagc agctacaaa	19560
aagactcatg tatatgctca ggctcccctt tctggcgaaa aaattagtaa agatggtctg	19620
caaataggaa cggacgctac agctacagaa caaaaaccta tttatgcaga ccctacattc	19680
cagcccgaac cccaaatcgg ggagtcccag tggaatgagg cagatgctac agtcgccggc	19740
ggtagagtgc taaagaaatc tactcccatg aaaccatgct atggttccta tgcaagacct	19800
acaaatgcta atggaggtca ggggtgacta acggcaaatg cccagggaca gctagaatct	19860
caggttgaaa tgcaattctt ttcaacttct gaaaacgccc gtaacgaggc taacaacatt	19920
cagcccaaat tgggtgctgta tagtgaggat gtgcacatgg agaccccgga tacgcacctt	19980
tcttacaagc ccgcaaaaag cgatgacaat tcaaaaatca tgctgggtca gcagtccatg	20040
cccaacagac ctaattacat cggcttcaga gacaacttta tcggcctcat gtattacaat	20100
agcactggca acatgggagt gcttgcaggt caggcctctc agttgaatgc agtgggtggac	20160
ttgcaagaca gaaacacaga actgtcctac cagctcttgc ttgattccat ggggtgacaga	20220
accagatact tttccatgtg gaatcaggca gtggacagtt atgaccaga tgttagaatt	20280
attgaaaatc atggaactga agacgagctc cccaactatt gtttccctct ggggtggcata	20340
ggggtaactg acacttacca ggctgttaaa accaacaatg gcaataacgg gggccagggtg	20400

[0102]

acttggacaa aagatgaaac ttttgcagat cgcaatgaaa taggggtggg aaacaatttc	20460
gctatggaga tcaacctcag tgccaacctg tggagaaact tcctgtactc caacgtggcg	20520
ctgtacctac cagacaagct taagtacaac ccctccaatg tggacatctc tgacaacccc	20580
aacacctacg attacatgaa caagcgagtg gtggccccgg ggctggtgga ctgctacatc	20640
aacctgggcg cgcgctggtc gctggactac atggacaacg tcaaccctt caaccaccac	20700
cgcaatgcgg gcctgcgcta ccgctccatg ctctgggca acgggcgcta cgtgcccttc	20760
cacatccagg tgccccagaa gtcttttgc atcaagaacc tcctcctcct gccgggctcc	20820
tacacctacg agtggaaactt caggaaggat gtcaacatgg tcctccagag ctctctgggt	20880
aacgatctca ggggtggacgg ggccagcatc aagtctgaga gcatctgcct ctacgccacc	20940
ttcttcccca tggcccacaa cacggcctcc acgctcgagg ccatgctcag gaacgacacc	21000
aacgaccagt cttcaatga ctacctctcc gccgccaaca tgctctaccc cataccgcc	21060
aacgccacca acgtcccat ctccatcccc tcgcgcaact gggcggcctt ccgcggtgg	21120
gccttcaccc gcctcaagac caaggagacc ccctccctgg gctcgggatt cgaccctac	21180
tacacctact cgggtccat tccctacctg gacggcacct tctacctcaa ccacacttc	21240
aagaaggtct cggtcacctt cgactcctcg gtcagctggc cgggcaacga ccgtctgctc	21300
acccccaacg agttcgagat caagcgctcg gtcgacgggg agggctacaa cgtggcccag	21360
tgcaacatga ccaaggactg gtctctggtc cagatgctgg ccaactacaa catcggtac	21420
cagggtttct acatcccaga gagctacaag gacaggatgt actccttctt caggaacttc	21480
cagcccatga gccggcagg ggtggaccag accaagtaca aggactacca ggaggtgggc	21540
atcatccacc agcacaacaa ctcgggcttc gtgggctacc tcgccccac catgcgcgag	21600
ggacaggcct accccgcaa cttcccctat ccgtcatag gcaagaccgc ggtcgacagc	21660
atcaccaga aaaagttcct ctgcgaccgc accctctggc gcatccctt ctccagcaac	21720
ttcatgtcca tgggtgcgct ctcggaactg ggccagaact tgctctacgc caactccgcc	21780
cacgcctcg acatgacctt cgaggctgac cccatggacg agcccacct tctctatgtt	21840
ctgttcgaag tctttgacgt ggtccgggtc caccagccgc accgcggcgt catcgagacc	21900
gtgtacctgc gtacgccctt ctcgccggc aacgccacca cctaaagaag caagccgcag	21960
tcategcgc ctgcatgccg tcgggttcca ccgagcaaga gctcagggcc atcgtcagag	22020
acctgggatg cgggccctat tttttgggca ctttcgacaa gcgcttcctt ggctttgtct	22080
ccccacacaa gctggcctgc gccatcgtca acacggccgg ccgcgagacc gggggcgtgc	22140

[0103]

actggctggc cttgcctgg aacccgcgt ccaaaacatg cttcctcttt gaccccttcg	22200
gcttttcgga ccagcggctc aagcaaactc acgagttcga gtacgagggc ttgctgcgtc	22260
gcagcgccat cgcctcctcg cccgaccgt gcgtaccct cgaaaagtcc acccagaccg	22320
tgcaggggcc cgactcggcc gcctgcggtc tcttctgttg catgtttctg cacgcctttg	22380
tgcactggcc tcagagtccc atggaccgca accccacat gaacttgctg acgggggtgc	22440
ccaactccat gctccagagc ccccaggteg agcccacct gcgccgaac caggagcagc	22500
tctacagctt cctggagcgc cactcgcctt acttccgcgc ccacagcgca cagatcagga	22560
gggccacctc cttctgccac ttgcaagaga tgcaagaagg gtaataacga tgtacacact	22620
ttttttctca ataaatggca tctttttatt tatacaagct ctctggggta ttcatttccc	22680
accaccaccc gccgttgctg ccatctggct ctatttagaa atcgaaaggg ttctgccggg	22740
agtcgccgtg cgccacgggc agggacacgt tgcgatactg gtagcgggtg cccacttga	22800
actcgggcac caccaggcga ggcagctcgg ggaagttttc gctccacagg ctgcgggtca	22860
gcaccagcgc gttcatcagg tcgggcgccg agatcttgaa gtcgcagttg gggccgccgc	22920
cctgcgcgcg cgagttgcgg tacaccgggt tgcagcactg gaacaccaac agcgccgggt	22980
gcttcacgtt ggccagcacg ctgcggtcgg agatcagctc ggcgctccagg tcctccgcgt	23040
tgctcagcgc gaacggggtc atcttgggca cttgccgcc caggaagggc gcgtgccccg	23100
gtttcgagtt gcagtcgcag cgcagcggga tcagcaggtg cccgtgcccc gactcggcgt	23160
tggggtacag cgcgcgcatg aaggcctgca tctggcgga ggccatctgg gccttggcgc	23220
cctccgagaa gaacatgccg caggacttgc ccgagaactg gtttgcgggg cagctggcgt	23280
cgtgcaggca gcagcgcgcg tcggtgttgg cgatctgcac cacgttgcgc ccccaccggt	23340
tcttcacgat cttggccttg gacgattgct ccttcagcgc gcgctgcccc ttctcgctgg	23400
tcacatccat ctcgatcaca tgttccttgt tcacatgct gctgccgtgc agacacttca	23460
gctcgccctc cgtctcggtg cagcgggtgt gccacagcgc gcagcccgtg ggctcgaaag	23520
acttgtaggt cacctccgcg aaggactgca ggtacccttg caaaaagcgg cccatcatgg	23580
tcacgaaggt cttgttgctg ctgaaggtea gctgcagccc gcggtgctcc tegttcagcc	23640
aggtcttgca cacggccgcc agcgccctca cctggtcggg cagcatcttg aagttcacct	23700
tcagctcatt ctccacgtgg tacttgctca tcagcgtgcg cgccgcctcc atgcccttct	23760
cccaggccga caccagcggc aggtctacgg ggttcttcac catcaccgtg gccgccgct	23820

[0104]

ccgccgcgct	ttcgctttcc	gccccgctgt	tctcttcctc	ttcctcctct	tcctcgccgc	23880
cgcccactcg	cagcccccg	accacggggt	cgtcttcctg	caggcgctgc	accttgcgct	23940
tgcggttgcg	cccctgcttg	atgcgcacgg	gcgggttgct	gaagcccacc	atcaccagcg	24000
cggcctcttc	ttgctcgcc	tcgctgtcca	gaatgacctc	cggggagggg	gggttggtca	24060
tcctcagtag	cgaggcacgc	ttctttttct	tcctgggggc	gttcgccagc	tccgcggctg	24120
cggccgctgc	cgaggtcgaa	ggccgagggc	tgggcgtgcg	cggcaccagc	gcgtcctgcg	24180
agccgtcctc	gtcctcctcg	gactcgagac	ggaggcgggc	ccgcttcttc	gggggcgcgc	24240
ggggcggcgg	aggcggcggc	ggcgacggag	acggggacga	gacatcgctc	agggtgggtg	24300
gacggcgggc	cgcgccgcgt	ccgcgctcgg	gggtggtctc	gcgctggtcc	tcttcccgac	24360
tggccatctc	ccactgctcc	ttctcctata	ggcagaaaga	gatcatggag	tctctcatgc	24420
gagtcgagaa	ggaggaggac	agcctaaccg	ccccctctga	gccctccacc	accgccgcca	24480
ccaccgccaa	tgccgccgcg	gacgacgcgc	ccaccgagac	caccgccagt	accacctcc	24540
ccagcgacgc	acccccgctc	gagaatgaag	tgctgatcga	gcaggacccg	ggttttgtga	24600
gcggagagga	ggatgaggtg	gatgagaagg	agaaggagga	ggtcgccgcc	tcagtgccaa	24660
aagaggataa	aaagcaagac	caggacgacg	cagataagga	tgagacagca	gtcgggcggg	24720
ggaacggaag	ccatgatgct	gatgacggct	acctagacgt	gggagacgac	gtgctgctta	24780
agcacctgca	ccgccagtgc	gtcatcgtct	gcgacgcgct	gcaggagcgc	tgcaagtgc	24840
ccctggacgt	ggcggaggtc	agccgcgcct	acgagcggca	cctcttcgcg	ccgcacgtgc	24900
cccccaagcg	ccgggagaac	ggcacctgcg	agcccaaccc	gcgtctcaac	ttctaccggg	24960
tcttcgcggt	acccgaggtg	ctggccacct	accacatctt	tttccaaaac	tgcaagatcc	25020
ccctctcctg	ccgcgccaac	cgcacccgcg	ccgacaaaac	cctgaccctg	cggcagggcg	25080
cccacatacc	tgatatcgcc	tctctggagg	aagtgcctaa	gatcttcgag	ggtctcggtc	25140
gcgacgagaa	acgggcggcg	aacgctctgc	acggagacag	cgaaaacgag	agtcactcgg	25200
gggtgctggt	ggagctcgag	ggcgacaacg	cgcgcctggc	cgtactcaag	cgcagcatag	25260
aggtcaccca	ctttgcctac	ccggcgctca	acctgcccc	caaggctcatg	agtgtggtca	25320
tgggcgagct	catcatgcgc	cgcgcccagc	ccctggccgc	ggatgcaaac	ttgcaagagt	25380
cctccgagga	aggcctgccc	gcggtcagcg	acgagcagct	ggcgcgctgg	ctggagaccc	25440
gcgaccccg	gcagctggag	gagcggcgca	agctcatgat	ggccgcggtg	ctggtcaccg	25500
tggagctcga	gtgtctgcag	cgtttcttcg	cggaccccg	gatgcagcgc	aagctcgagg	25560

[0105]

agaccctgca ctacaccttc cgccaggggt acgtgcgcca ggcctgcaag atctccaacg	25620
tggagctctg caacctggtc tcctacctgg gcatcctgca cgagaaccgc ctcgggcaga	25680
acgtcctgca ctccaccctc aaaggggagg cgcgccgcga ctacatccgc gactgcgcct	25740
acctcttcct ctgctacacc tggcagacgg ccatgggggt ctggcagcag tgcctggagg	25800
agcgcaacct caaggagctg gaaaagctcc tcaagcgcac cctcaggga cctctggacgg	25860
gcttcaacga gcgctcggtg gccgccgcgc tggcggacat catctttccc gagcgccctgc	25920
tcaagacctt gcagcagggc ctgcccgaact tcaccagcca gagcatgctg cagaacttca	25980
ggactttcat cctggagcgc tcgggcatcc tgccggccac ttgctgcgcg ctgcccagcg	26040
acttcgtgcc catcaagtac agggagtgcc cgccgccgct ctggggccac tgctacctct	26100
tccagctggc caactacctc gcctaccact cggacctcat ggaagacgtg agcggcgagg	26160
gcctgctcga gtgccactgc cgctgcaacc tctgcacgcc ccaccgctct ctagtctgca	26220
acccgcagct gctcagcgag agtcagatta tcggtacctt cgagctgcag ggtccctcgc	26280
ctgacgagaa gtccgcggct ccagggtga aactcactcc ggggctgtgg acttccgcct	26340
acctacgcaa atttgtacct gaggactacc acgcccacga gatcaggttc tacgaagacc	26400
aatcccgccc gcccaaggcg gagctcaccg cctgcgtcat caccagggg cacatcctgg	26460
gccaattgca agccatcaac aaagcccgcg gagagttctt gctgaaaaag ggtcgggggg	26520
tgtacctgga ccccgagtc gccgaggagc taaacccgct acccccgccg ccgccccagc	26580
agcgggacct tgcttcccag gatggcaccc agaaagaagc agcagccgcc gccgccgccg	26640
cagccataca tgcttctgga ggaagaggag gaggactggg acagtcaggc agaggaggtt	26700
tcggacgagg agcaggagga gatgatggaa gactgggagg aggacagcag cctagacgag	26760
gaagcttcag aggccgaaga ggtggcagac gcaacaccat cgccctcggt cgcagccccc	26820
tcgccggggc ccctgaaatc ctccgaacct agcaccagcg ctataacctc cgctcctccg	26880
gcgccggcgc caccgccccg cagacccaac cgtagatggg acaccacagg aaccggggtc	26940
ggtaagtcca agtgcccgcc gccgccaccg cagcagcagc agcagcagcg ccagggctac	27000
cgctcgtggc gcgggcacaa gaacgccata gtcgcctgct tgcaagactg cgggggcaac	27060
atctcttttcg cccgccgctt cctgctattc caccacgggg tcgcctttcc ccgcaatgtc	27120
ctgcattact accgtcatct ctacagcccc tactgcagcg gcgaccaga ggcggcagcg	27180
gcagccacag cggcgaccac cacctaggaa gatatacctc gcgggcaaga cagcggcagc	27240

[0106]

agcggccagg agacccgcgg cagcagcggc gggagcgggtg ggcgcactgc gcctctcggc	27300
caacgaaccc ctctcgaccc gggagctcag acacaggatc ttccccactt tgtatgccat	27360
cttccaacag agcagaggcc aggagcagga gctgaaaata aaaaacagat ctctgcgctc	27420
cctcaccgc agctgtctgt atcacaaaag cgaagatcag cttcggcgca cgctggagga	27480
cgcggaggca ctcttcagca aatactgcgc gctcactctt aaagactagc tccgcgcctt	27540
tctcgaattt aggcgggaga aaactacgtc atcgccggcc gccgcccagc ccgcccagcc	27600
gagatgagca aagagattcc cacgccatac atgtggagct accagccgca gatgggactc	27660
gcggcgggag cggcccagga ctactccacc cgcatgaact acatgagcgc gggacccccac	27720
atgatctcac aggtcaacgg gatccgcgcc cagcgaaacc aaatactgct ggaacaggcg	27780
gccatcaccg ccacgccccg ccataatctc aacccccgaa attggccccg cgccctcgtg	27840
taccagaaa cccctccgc caccaccgta ctacttccgc gtgacgcca ggccgaagtc	27900
cagatgacta actcaggggc gcagctcgcg ggcggctttc gtcacggggc gcggccgctc	27960
cgaccaggta taagacacct gatgatcaga ggccgaggta tccagctcaa cgacgagtcg	28020
gtgagctctt cgctcggctt ccgtccggac ggaactttcc agctcgccgg atccggccgc	28080
tcttcgttca cgccccgca ggcgtaacctg actctgcaga cctcgctctc ggagccccgc	28140
tccggcggca tcggaaccct ccagtctgtg gaggagtctg tgccctcggg ctacttcaac	28200
cccttctcgg gacctcccg acgtacccc gaccagtcca ttccgaactt tgacgcgggtg	28260
aaggactcgg cggacggcta cgactgaatg tcagggtgtcg aggacagca gcttcgcctg	28320
agacacctcg agcactgccg ccgccacaag tgcttcgccc gcggttctgg tgagttctgc	28380
tactttcagc taccgagga gcataccgag gggccggcgc acggcgctccg cctgaccacc	28440
cagggcgagg ttacctgttc cctcatccgg gaggtttacc tccgtcccct gctagtggag	28500
cgggagcggg gtccctgtgt cctaactatc gcctgcaact gccctaacc tggattacat	28560
caagatcttt gctgtcatct ctgtgtgag tttaataaac gctgagatca gaatctactg	28620
gggctcctgt cgccatcctg tgaacgccac cgtcttcacc caccgccacc aggccaggc	28680
gaacctcacc tgcggtctgc atcggagggc caagaagtac ctcacctggt acttcaacgg	28740
caccccttt gtggtttaca acagcttcga cggggacgga gtctccctga aagaccagct	28800
ctccggtctc agctactcca tccacaagaa caccaccctc caactcttcc ctccctacct	28860
gccgggaacc tacgagtgcg tcaccggccg ctgcacccac ctcaccgcc tgatcgtaaa	28920
ccagagcttt ccgggaacag ataactccct cttccccaga acaggaggtg agctcaggaa	28980

[0107]

actccccggg gaccagggcg gagacgtacc ttcgaccctt gtgggggtag gattttttat	29040
taccgggttg ctggctcttt taatcaaagt ttccttgaga ttgtttcttt ctttctacgt	29100
gtatgaacac ctcaacctcc aataactcta ccttttcttc ggaatcaggt gactttctctg	29160
aaatcgggct tgggtgtgctg cttactctgt tgattttttt ctttatcata ctacgccttc	29220
tgtgcctcag gctcgccgcc tgcctgcgcac acatctatat ctactgctgg ttgctcaagt	29280
gcaggggtcg ccaccaaga tgaacaggta catggctcta tcgacccag gctgctggc	29340
cctggcggcc tgcagcgccg ccaaaaaaga gattaccttt gaggagcccg cttgcaatgt	29400
aactttcaag cccgaggggtg accaatgcac caccctcgtc aaatgcgtta ccaatcatga	29460
gaggctgcgc atcgactaca aaaacaaaac tggccagttt gcggtctata gtgtgtttac	29520
gcccggagac cctctaaact actctgtcac cgtcttccag ggcggacagt ctaagatatt	29580
caattacact ttcccttttt atgagttatg cgatgcggtc atgtacatgt caaacagta	29640
caacctgtgg cctccctctc cccagcgctg tgtggaaaat actgggtctt actgctgtat	29700
ggctttcgca atcactacgc tcgctctaata ctgcacggtg ctatacataa aattcaggca	29760
gaggcgaatc tttatcgatg aaaagaaaat gccttgatcg ctaacaccgg ctttctatct	29820
gcagaatgaa tgcaatcacc tccctactaa tcaccaccac cctccttgcg attgccatg	29880
ggttgacacg aatcgaagtg ccagtggggt ccaatgtcac catggtgggc cccgccggca	29940
attccaccct catgtgggaa aaatttgtcc gcaatcaatg ggttcatttc tgctctaacc	30000
gaatcagtat caagcccaga gccatctgcg atgggcaaaa tctaactctg atcaatgtgc	30060
aaatgatgga tgctgggtac tattacgggc agcggggaga aatcattaat tactggcgac	30120
cccacaagga ctacatgctg catgtagtcg aggcacttcc cactaccacc cccactacca	30180
cctctccac caccaccacc actactacta ctactactac tactactact actaccacta	30240
ccgctgcccc ccataccgcg aaaagcacca tgattagcac aaagccccct cgtgctcact	30300
cccacgccgg cgggcccac ggtgcgacct cagaaaccac cgagctttgc ttctgccaat	30360
gcactaacgc cagcgctcat gaactgttcg acctggagaa tgaggatgtc cagcagagct	30420
ccgcttgccct gaccaggag gctgtggagc ccgttgccct gaagcagatc ggtgattcaa	30480
taattgactc ttctttcttt gccactcccg aataccctcc cgattctact ttccacatca	30540
cgggtaccaa agaccctaac ctctctttct acctgatgct gctgctctgt atctctgtgg	30600
tctcttccgc gctgatgtta ctggggatgt tctgctgcct gatctgccgc agaaagagaa	30660

[0108]

aagctcgctc tcagggccaa ccactgatgc cttcccccta ccccccgat tttgcagata	30720
acaagatatg agctcgctgc tgacactaac cgctttacta gcctgcgctc taacccttgt	30780
cgcttgcgac tcgagattcc acaatgtcac agctgtggca ggagaaaatg ttactttcaa	30840
ctccacggcc gatacccagt ggctgtggag tggtcaggt agctacttaa ctatctgcaa	30900
tagctccact tccccggca tatccccaac caagtaccaa tgcaatgcca gcctgttcac	30960
cctcatcaac gttccaccc tggacaatgg actctatgta ggctatgtac cttttggtgg	31020
gcaaggaaag acccacgctt acaacctgga agttcgccag ccagaaacca ctaccaagc	31080
ttctcccacc accaccacca ccaccacat caccagcagc agcagcagca gcagccacag	31140
cagcagcagc agattattga ctttggtttt ggccagctca tctgccgcta cccaggccat	31200
ctacagctct gtgcccga aa ccactcagat ccaccgcca gaaacgacca ccgccaccac	31260
cctacacacc tccagcgatc agatgccgac caacatcacc cccttggtc ttcaaattggg	31320
acttacaagc cccactccaa aaccagtgga tgcggccgag gtctccgccc tcgtcaatga	31380
ctgggcgggg ctgggaatgt ggtggttcgc cataggcatg atggcgctct gcctgtttct	31440
gctctggctc atctgtgcc tccaccgcag gcgagccaga cccccatct atagacccat	31500
cattgtcctg aaccccgata atgatgggat ccatagattg gatggcctga aaaacctact	31560
tttttctttt acagtatgat aaattgagac atgcctcgca ttttcttgta catgttcctt	31620
ctcccacctt ttctggggtg ttctacgtg gccgctgtgt ctcacctgga ggtagactgc	31680
ctctcaccct tcaactgtcta cctgctttac ggattggtca ccctcactct catctgcagc	31740
ctaatacag taatcatcgc cttcatccag tgcattgatt acatctgtgt gcgcctcgca	31800
tacttcagac accacccgca gtaccgagac aggaacattg cccaacttct aagactgctc	31860
taatcatgca taagactgtg atctgccttc tgatcctctg catcctgccc accctcacct	31920
cctgccagta caccacaaaa tctccgcgca aaagacatgc ctctgcccgc ttcaccaaac	31980
tgtggaatat acccaaatgc tacaacgaaa agagcgagct ctccgaagct tggctgtatg	32040
gggtcatctg tgtcttagtt ttctgcagca ctgtctttgc cctcataatc taccctact	32100
ttgatttggg atggaacgcg atcgatgcca tgaattaccc cactttccc gcacccgaga	32160
taattccact gcgacaagtt gtaccggttg tcgttaatca acgcccccca tcccctacgc	32220
ccactgaaat cagctacttt aacctaacag gcggagatga ctgacgccct agatctagaa	32280
atggacggca tcagtaccga gcagcgtctc ctagagaggc gcaggcaggc ggctgagcaa	32340
gagcgcctca atcaggagct ccgagatctc gttaacctgc accagtgcaa aagaggcatc	32400

[0109]

ttttgtcttg taaagcaggc caaagtcacc tacgagaaga cggcaacag ccaccgcctc	32460
agttacaaat tgcccacca gcgccagaag ctggtgctca tgggtgggtga gaatcccatc	32520
accgtcacc cagcactcgg agagaccgag ggggtgtctgc actccccctg tcgggggtcca	32580
gaagacctct gcaccctgg aaagaccctg tgcgggtctca gagatttagt cccctttaac	32640
taatcaaaca ctggaatcaa taaaaagaat cacttactta aaatcagaca gcaggtctct	32700
gtccagttta ttcagcagca cctccttccc ctctcccaca ctctggtact ccaaagcct	32760
tctggcggca aacttccctc acaccctgaa gggaatgtca gattcttctc cctgtccctc	32820
cgcacccact atcttcatgt tgttgcagat gaagcgcacc aaaacgtctg acgagagctt	32880
caaccccgctg tacccttatg acacggaaag cggccctccc tccgtccctt tctcaccctc	32940
tcccttcgtg tctcccgatg gattccaaga aagtcccccc ggggtcctgt ctctgaacct	33000
ggccgagccc ctggtcactt cccacggcat gctcgcctg aaaatgggaa gtggcctctc	33060
cctggacgac gctggcaacc tcacctctca agatataacc accgctagcc ctccctcaa	33120
aaaaaccaag accaacctca gcctagaaac ctcatcccc ctaactgtga gcacctcagg	33180
cgcctcacc gtagcagccg ccgctccctt ggcgggtggc ggcacctccc tcacatgca	33240
atcagaggcc cccctgacag tacaggatgc aaaactcacc ctggccacca aaggccccct	33300
gaccgtgtct gaaggcaaac tggccttgca aacatcggcc ccgctgacgg ccgctgacag	33360
cagcaccctc acagtcagt cccacaccac ccttagcaca agcaatggca gcttgggtat	33420
tgacatgcaa gccccattt acaccaccaa tggaaaacta ggacttaact ttggcgctcc	33480
cctgcatgtg gtagacagcc taaatgcact gactgtagtt actggccaag gtcttacgat	33540
aaacggaaca gccctacaaa ctagagtctc aggtgcctc aactatgaca catcaggaaa	33600
cctagaattg agagctgcag ggggtatgcg agttgatgca aatggtcaac ttatccttga	33660
tgtagcttac ccatttgatg caaaaaaca tctcagcctt aggcttggac agggaccctt	33720
gtttgttaac tctgcccaca acttggatgt taactacaac agaggcctct acctgttacc	33780
atctggaaat accaaaaagc tagaagttaa tatcaaaaca gccaaagggtc tcatttatga	33840
tgacactgct atagcaatca atgcgggtga tgggctacag tttgactcag gctcagatac	33900
aaatccatta aaaactaaac ttggattagg acttgattat gactccagca gagccataat	33960
tgctaaactg ggaactggcc taagctttga caacacaggt gccatcacag taggcaacaa	34020
aaatgatgac aagcttacct tgtggaccac accagacca tcccctaact gtagaatcta	34080

[0110]

ttcagagaaa gatgctaaat tcacacttgt ttgactaaa tgcggcagtc aggtgttggc	34140
cagcgtttct gttttatctg taaaaggtag ctttgcgccc atcagtggca cagtaactag	34200
tgtcagatt gtcctcagat ttgatgaaaa tggagtctta ctaagcaatt cttcccttga	34260
ccctcaatac tggaactaca gaaaagggtga ctttacagag ggcaactgcat ataccaacgc	34320
agtgggattt atgcccacc tcacagcata cccaaaaaca cagagccaaa ctgctaaaag	34380
caacattgta agtcaggttt acttgaatgg ggacaaatcc aaacccatga ccctcaccat	34440
taccctcaat ggaactaatg aaacaggaga tgccacagta agcacttact ccatgtcatt	34500
ctcatggaac tggaatggaa gtaattacat taatgaaacg ttccaaacca actccttcac	34560
cttctcctac atcgcccaag aataaaaagc atgacgctgt tgatttgatt caatgtgttt	34620
ctgttttatt ttcaagcaca acaaaatcat tcaagtcatt cttccatctt agcttaatag	34680
acacagtagc ttaatagacc cagtagtgca aagccccatt ctagcttata gatcagacag	34740
tgataattaa ccaccaccac caccatacct ttgattcag gaaatcatga tcatcacagg	34800
atcctagtct tcaggccgcc cctccctcc caagacacag aatacacagt cctctcccc	34860
cgactggctt taaataacac catctggttg gtcacagaca tgttcttagg gggttatattc	34920
cacacggtct cctgccgcgc caggcgctcg tcggtgatgt tgataaactc tcccggcagc	34980
tcgctcaagt tcacgtcgt gtccagcggc tgaacctccg gctgacgcga taactgtgcg	35040
accggctgct ggacgaacgg aggcgcgcc tacaaggggg tagagtcata atcctcggtc	35100
aggatagggc ggtgatgcag cagcagcgag cgaaacatct gctgccgcg ccgctccgtc	35160
cggcaggaaa acaacacgcc ggtggtctcc tccgcgataa tccgcaccgc ccgcagcatc	35220
agcttctctg ttctccgcgc gcagcacctc acccttatct cgctcaaate ggcgagtag	35280
gtacagcaca gcaccacgat gttattcatg atcccacagt gcaggcgct gtatccaaag	35340
ctcatgccgg gaaccaccgc cccacgtgg ccatcgtacc acaagcgac gtaaatacaag	35400
tgtcgacccc tcatgaacgc gctggacaca aacattactt ccttgggcat gttgtaattc	35460
accacctccc ggtaccagat aaacctctgg ttgaacaggg caccttcac caccatctg	35520
aaccaagagg ccagaacctg cccaccggt atgcaactga gggaaccgg gttggaacaa	35580
tgacaatgca gactccaagg ctcgtaaccg tggatcatcc ggctgctgaa ggcatcgatg	35640
ttggcacaac acagacacac gtgcatgcac tttctcatga ttagcagctc ttccctcgtc	35700
aggatcatat cccaaggaat aaccattct tgaatcaacg taaaaccac acagcaggga	35760
aggcctcgca cataactcac gttgtgcatg gtcagcgtgt tgcatccgg aaacagcgga	35820

[0111]

tgatcctcca gtatcgaggc gcgggtctcc ttctcacagg gaggtaaagg gtccctgctg	35880
tacggactgc gccgggacga ccgagatcgt gttgagcgta gtgtcatgga aaagggaacg	35940
ccggacgtgg tcatacttct tgaagcagaa ccaggttcgc gcgtggcagg cctccttgcg	36000
tctgcggtct cgccgtctag ctcgctccgt gtgatagttg tagtacagcc actcccgcag	36060
agcgtegagg cgcaccctgg ctcccgatc tatgtagact ccgtcttgca ccgcggccct	36120
gataatatcc accaccgtag aataagcaac acccagccaa gcaatacact cgctctgcga	36180
gcggcagaca ggaggagcgg gcagagatgg gagaacatg ataaaaaact ttttttaaag	36240
aatatatttc aattcttcga aagtaagatc tatcaagtgg cagcgctccc ctccactggc	36300
gcggtcaaac tctacggcca aagcacagac aacggcattt ctaagatgtt ccttaatggc	36360
gtccaaaaga cacaccgctc tcaagttgca gtaaactatg aatgaaaacc catccggctg	36420
attttccaat atagacgcgc cggcagcgctc caccaaacc agataatttt cttctctcca	36480
gcggtttacg atctgtctaa gcaaatccct tatatcaagt ccgaccatgc caaaaatctg	36540
ctcaagagcg ccctccacct tcatgtacaa gcagcgcatc atgattgcaa aaattcaggt	36600
tcttcagaga cctgtataag attcaaatg ggaacattaa caaaaattcc tctgtcgcgc	36660
agatcccttc gcagggaag ctgaacataa tcagacaggt ccgaacggac cagtgaggcc	36720
aaatccccac caggaaccag atccagagac cctatactga ttatgacgcg catactcggg	36780
gctatgctga ccagcgtagc gccgatgtag gcgtgctgca tgggcggcga gataaaatgc	36840
aaagtgtgtg ttaaaaaatc aggcaaagcc tcgcgcaaaa aagctaacac atcataatca	36900
tgctcatgca ggtagttgca ggtaagctca ggaacaaaa cggaataaca cacgattttc	36960
ctctcaaaca tgacttcgcg gatactgcgt aaaacaaaaa attataaata aaaaattaat	37020
taaataactt aaacattgga agcctgtctc acaacaggaa aaaccacttt aatcaacata	37080
agacgggcca cgggcatgcc ggcatagccg taaaaaaatt ggtccccgtg attaacaagt	37140
accacagaca gctccccggt catgtcgggg gtcatcatgt gagactctgt atacacgtct	37200
ggattgtgaa catcagacaa acaaagaaat cgagccacgt agccccgagg tataatcacc	37260
cgcaggcgga ggtacagcaa aacgaccccc ataggaggaa tcacaaaatt agtaggagaa	37320
aaaaatacat aaacaccaga aaaaccctgt tgctgaggca aaatagcgcc ctcccgatcc	37380
aaaacaacat aaagcgcttc cacaggagca gccataacaa agacccgagt cttaccagta	37440
aaagaaaaaa gatctctcaa cgcagcacca gcaccaacac ttgcagtggt aaaaggccaa	37500

[0112]

gtgccgagag agtatatata ggaataaaaa gtgacgtaaa cgggcaaagt ccaaaaaaacg	37560
cccagaaaaa cgcacgcga acctacgccc cgaaacgaaa gccaaaaaac actagacact	37620
cccttcggc gtcaacttcc gttttccac gtacgtcac ttccccggg caaacaact	37680
acatatcccg aacttccaag tcgccacgcc caaaacaccg cctacacct cccgcccgc	37740
ggcccccccc cggaccgcg tcccgcccc cgccgccc atctattatca tatttggttc	37800
aatccaaaat aaggtatatt attgatgatg	37830

<210> 11

<211> 37559

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 11

catcatcaat aatatacctt attttgatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tggggcgggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgta gtgaatttg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acatttatat tggctcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atatatggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct gaccgccc	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caataggac	720
ttccattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca	780
agtgatcat atgccaagta cgccccctat tgacgtcaat gacggtaa atggccgcctg	840
gcattatgcc cagtacatga ctttatggga ctttcctact tggcagtaca tctacgtatt	900
agtcacgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg	1020

[0113]

gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaa	1080
ggcggttagg cgtgtacggt gggagggtcta tataagcaga gctctcccta tcagtgatag	1140
agatctccct atcagtgata gagatcgteg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacggtgca ttggaacgcg gattccccgt gccaaagagt agatcttccg	1320
tttatctagg taccagatat cgccaccatg gaactgctga tcctgaaggc caacgccatc	1380
accaccatcc tgaccgccgt gaccttctgc ttgccagcg gccagaacat caccgaggaa	1440
ttctaccaga gcacctgtag cgccgtgagc aagggtacc tgagcgccct gagaaccggc	1500
tggtacacca gcgtgatcac catcgagctg agcaacatca aagaaaacaa gtgcaacggc	1560
accgacgcca aagtgaagct gatcaagcag gaactggaca agtacaagaa cgccgtgacc	1620
gagctgcagc tgctgatgca gagcaccccc gccaccaaca accgggccag acgggagctg	1680
ccccggttca tgaactacac cctgaacaac gccaaaaaga ccaacgtgac cctgagcaag	1740
aagcggaagc ggcggttcct gggtttctg ctgggcgtgg gcagcgccat tgccagcggc	1800
gtggccgtgt ctaaggtgct gcacctggaa ggcaagtga acaagatcaa gagcgccctg	1860
ctgagcacca acaaggccgt ggtgtccctg agcaacggcg tgagcgtgct gaccagcaag	1920
gtgctggatc tgaagaacta catcgacaag cagctgctgc ccatcgtgaa caagcagagc	1980
tgcagcatca gcaacatcga gacagtgate gagttccagc agaagaacaa ccggctgctg	2040
gaaatcacc cggagttcag cgtgaacgcc ggcgtgacca cccctgtgtc cacctacatg	2100
ctgaccaaca gcgagctgct gagcctgate aacgacatgc ccatcaccaa cgaccagaaa	2160
aagctgatga gcaacaacgt gcagatcgtg cggcagcaga gctactccat catgtccatc	2220
atcaaagaag aggtgctggc ctacgtggtg cagctgcccc tgtacggcgt gatcgacacc	2280
ccctgctgga agctgcacac cageccccctg tgcaccacca acaccaaaga gggcagcaac	2340
atctgcctga cccggaccga cagaggctgg tactgcgaca acgccggcag cgtgtcatte	2400
tttccacagg ccgagacatg caaggtgcag agcaaccggg tgttctgcga caccatgaac	2460
agcctgacct tgccctccga agtgaacctg tgcaacgtgg acatcttcaa cccaagtac	2520
gactgcaaga tcatgacctc caagaccgac gtgtccagct ccgtgatcac ctccctgggc	2580
gccatcgtgt cctgctacgg caagaccaag tgcaccgcca gcaacaagaa ccggggcacc	2640
atcaagacct tcagcaacgg ctgcgactac gtgtccaaca agggggtgga caccgtgtcc	2700
gtgggcaaca ccctgtacta cgtgaacaaa caggaaggca agagcctgta cgtgaagggc	2760

[0114]

gagcccatca tcaacttcta cgacccccctg gtgttcccca gcgacgagtt cgacgccagc	2820
atcagccagg tgaacgagaa gatcaaccag agcctggcct tcatccggaa gtccgacgag	2880
ctgctgcaca atgtgaatgc cggcaagtcc accaccaacc ggaagcggag agccccctgtg	2940
aagcagaccc tgaacttcga cctgctgaag ctggccggcg acgtggagag caatccccggc	3000
cctatggccc tgagcaaagt gaaactgaac gatacactga acaaggacca gctgctgtcc	3060
agcagcaagt acaccatcca gcggagcacc ggcgacagca tcgatacccc caactacgac	3120
gtgcagaagc acatcaacaa gctgtgcggc atgctgtctga tcacagagga cgccaaccac	3180
aagttcaccg gcctgatcgg catgctgtac gccatgagcc ggctggggccg ggaggacacc	3240
atcaagatcc tgcgggacgc cggctaccac gtgaaggcca atggcgtgga cgtgaccaca	3300
caccggcagg acatcaacgg caaagaaatg aagttcgagg tgctgaccct ggccagcctg	3360
accaccgaga tccagatcaa tatcgagatc gagagccgga agtcctacaa gaaaatgctg	3420
aaagaaatgg gcgaggtggc ccccgagtac agacacgaca gccccgactg cggcatgatc	3480
atcctgtgta tcgccgcct ggtgatcaca aagctggccg ctggcgacag atctggcctg	3540
acagccgtga tcagacgggc caacaatgtg ctgaagaacg agatgaagcg gtacaagggc	3600
ctgctgceca aggacattgc caacagcttc tacgaggtgt tcgagaagta cccccacttc	3660
atcgacgtgt tcgtgcactt cggcattgcc cagagcagca ccagaggcgg ctccagagtg	3720
gagggcatct tcgccgcct gttcatgaac gcctacggcg ctggccaggt gatgctgaga	3780
tggggcgtgc tggccaagag cgtgaagaac atcatgctgg gccacgccag cgtgcaggcc	3840
gagatggaac aggtggtgga ggtgtacgag tacgccaga agctgggcgg agaggccggc	3900
ttctaccaca tcctgaacaa ccctaaggcc tccctgctgt cctgaccca gttccccac	3960
ttctccagcg tgggtgctggg aaatgccgcc ggactgggca tcatgggcga gtaccggggc	4020
acccccagaa accaggacct gtacgacgcc gccaaggcct acgccgagca gctgaaagaa	4080
aacggcgtga tcaactacag cgtgctggac ctgaccgtg aggaactgga agccatcaag	4140
caccagctga accccaagga caacgacgtg gagctgggag gcggaggatc tggcggcgga	4200
ggcatgagca gacggaaccc ctgcaagttc gagatccggg gccactgcct gaacggcaag	4260
cggtgccact tcagccacaa ctacttcgag tggeccctc atgctctgct ggtgcggcag	4320
aacttcatgc tgaaccgat cctgaagtcc atggacaaga gcatcgacac cctgagcgag	4380
atcagcggag ccgccgagct ggacagaacc gaggaatatg ccctgggcgt ggtgggagtg	4440

[0115]

ctggaaagct acatcggtc catcaacaac atcacaaagc agagcgctg cgtggccatg	4500
agcaagctgc tgacagagct gaacagcgac gacatcaaga agctgaggga caacgaggaa	4560
ctgaacagcc ccaagatccg ggtgtacaac accgtgatca gctacattga gagcaaccgc	4620
aagaacaaca agcagaccat ccatctgctg aagcggctgc ccgccgacgt gctgaaaaag	4680
accatcaaga acaccctgga catccacaag tccatcacca tcaacaatcc caaagaaagc	4740
accgtgtctg acaccaacga tcacgccaag aacaacgaca ccacctgatg agcggccgcg	4800
atctgctgtg ctttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt	4860
gaccttgaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca	4920
ttgtctgagt aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggga	4980
ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg ccgatcagcg	5040
atcgctgagg tgggtgagtg ggcgtggcct ggggtggtca tgaaaatata taagttgggg	5100
gtcttagggt ctctttattt gtgttgaga gaccgccgga gccatgagcg ggagcagcag	5160
cagcagcagt agcagcagcg ccttgatgg cagcatcgtg agcccttatt tgacgacgcg	5220
gatgccccac tgggccgggg tgcgtcagaa tgtgatgggc tccagcatcg acggccgacc	5280
cgctctgccc gcaaattccg ccacgtgac ctatgcgacc gtcgcgggga cgccgttga	5340
cgccaccgcc gccgccgccc ccaccgcagc cgcctcggcc gtgcgcagcc tggccacgga	5400
ctttgcattc ctgggaccac tggcgacagg ggctacttct cgggccgctg ctgccgccgt	5460
tcgcgatgac aagctgaccg ccctgttggc gcagttggat gcgcttactc gggaactggg	5520
tgacctttct cagcaggtca tggccctgcg ccagcaggtc tcctccctgc aagctggcgg	5580
gaatgcttct cccacaaatg ccgtttaaga taaataaaac cagactctgt ttggattaaa	5640
gaaaagtagc aagtgcattg ctctctttat ttcataattt tccgcgcgcg ataggcccta	5700
gaccagcgtt ctcggtcgtt gaggtgctgg tgtatcttct ccaggacgtg gtagaggtgg	5760
ctctggacgt tgagatacat ggcatgagc ccgtcccggg ggtggaggta gcaccactgc	5820
agagcttcat gtcgggggt ggtgtttag atgatccagt cgtagcagga gcgctgggca	5880
tggtgcctaa aaatgtcctt cagcagcagg ccgatggcca gggggaggcc cttggtgtaa	5940
gtgtttacaa aacggttaag ttgggaaggg tgcatcggg gagagatgat gtgcatcttg	6000
gactgtattt ttagattggc gatgtttccg ccagatccc ttctgggatt catgttgtgc	6060
aggaccacca gtacagtgt tccggtgcac ttggggaatt tgtcatgcag cttagaggga	6120
aaagcgtgga agaacttgga gacgcctttg tgccctccca gatcttccat gcattcgtcc	6180

[0116]

atgatgatgg caatgggccc gcgggaggca gcttgggcaa agatatttct ggggtcgctg	6240
acgtcgtagt tgtgttccag ggtgaggtcg tcataggcca tttttacaaa gcgcgggcgg	6300
agggtgcccc actgggggat gatggtcccc tctggccctg gggcgtagtt gccctcgag	6360
atctgcattt cccaggcctt aatctcggag gggggaatca tatccacctg cggggcgatg	6420
aagaaaacgg tttccggagc cggggagatt aactgggatg agagcaggtt tctaagcagc	6480
tgtgattttc cacaaccggt gggcccataa ataacaccta taaccggttg cagctggtag	6540
tttagagagc tgcagctgcc gtcgtcccgg aggagggggg ccacctcgtt gagcatgtcc	6600
ctgacgcgca tgttctcccc gaccagatcc gccagaaggc gctcgccgcc cagggacagc	6660
agctcttgca aggaagcaaa gtttttcagc ggcttgaggc cgtccgccgt gggcatgttt	6720
ttcagggtct ggctcagcag ctccaggcgg tcccagagct cggtgacgtg ctctacggca	6780
tctctatcca gcatatctcc tcgtttcgcg ggttggggcg actttcgctg tagggcacca	6840
agcggtggtc gtccagcggg gccagagtca tgtccttcca tgggcgcagg gtcctcgtca	6900
gggtggtctg ggtcacggtg aaggggtgcg ctccgggctg agcgcttgcc aaggtgcgct	6960
tgaggctggt tctgctggtg ctgaagcgct gccggtcttc gccctgcgcg tcggccaggt	7020
agcatttgac catggtgtca tagtcagcc cctccgcggc gtgtcccttg gcgcgcagct	7080
tgcccttgga ggtggcgccg cacgaggggc agagcaggct cttgagcgcg tagagcttgg	7140
gggcgaggaa gaccgattcg ggggagtagg cgtccgcgcc gcagaccccg cacacggtct	7200
cgcactccac cagccaggtg agctcggggc gcgccgggtc aaaaaccagg tttccccc	7260
gctttttgat gcgtttctta cctcgggtct ccatgaggtg gtgtccccgc tcggtgacga	7320
agaggctgtc cgtgtctccg tagaccgact tgaggggtct tttctccagg ggggtccctc	7380
ggtcttcctc gtagaggaac tcggaccact ctgagacgaa ggcccgcgtc caggccagga	7440
cgaaggaggc tatgtgggag gggtagcgtt cgttgtccac taggggggtcc accttctcca	7500
aggtgtgaag acacatgtcg ctttcctcgg cgtccaggaa ggtgattggc ttgtaggtgt	7560
aggccacgtg accgggggtt cctgacgggg ggggtataaaa gggggtgggg gcgcgctcgt	7620
cgtcactctc ttccgcacgc ctgtctgcga gggccagctg ctggggtgag tattccctct	7680
cgaaggcggg catgacctcc gcgctgaggt tgtcagtttc caaaaacgag gaggatttga	7740
tgttcacctg tcccaggtg atacctttga gggtagccgc gtccatctgg tcagaaaaca	7800
cgatcttttt attgtccagc ttggtggcga acgaccgta gagggcggtg gagagcagct	7860

[0117]

tggcgatgga ggcgagggtc tggttcttgt ccctgtcggc gcgctccttg gccgcgatgt	7920
tgagctgcac gtactcgcgc gcgacgcagc gccactcggg gaagacgggtg gtgcgctcgt	7980
cgggcaccag gcgcacgcgc cagccgcggt tgtgcagggt gaccagggtcc acgctgggtg	8040
cgacctcgcc gcgcaggcgc tcgttggtcc agcagagacg gccgcccttg cgcgagcaga	8100
aggggggcag ggggtcgagc tgggtctcgt ccgggggggtc cgcgtccacg gtgaaaaccc	8160
cggggcgcag gcgcgcgtcg aagtagtcta tcttgcaacc ttgcatgtcc agcgcctgct	8220
gccagtcgcg ggcggcgagc gcgcgctcgt aggggttagc cggcggggccc cagggcattg	8280
ggtgggttag tgcggaggcg tacatgccgc agatgtcata gacgtagagg ggctcccgca	8340
ggaccccgat gtaggtgggg tagcagcggc cgccgcggat gctggcgcgc acgtagtcac	8400
acagctcgtg cgagggggcg aggaggctcg gggccagggt ggtgcgggcg gggcgctccg	8460
cgcggaagac gatctgcctg aagatggcat gcgagttgga agagatgggtg gggcgctgga	8520
agacgttgaa gctggcgtcc tgcaggccga cggcgtcgcg cacgaaggag gcgtaggagt	8580
cgcgcagctt gtgtaccagc tcggcggtga cctgcacgtc gagcgcgcag tagtcgaggg	8640
tctcgcggat gatgtcatat ttagcctgcc cttctttttt ccacagctcg cggttgagga	8700
caaactcttc gcggtctttc cagtactctt ggatcgggaa accgtccggt tccgaacggt	8760
aagagcctag catgtagaac tggttgacgg cctggtaggc gcagcagccc ttctccacgg	8820
ggagggcgta ggcctgcgcg gccttgcgga gcgaggtgtg ggtcagggcg aaggtgtccc	8880
tgaccatgac tttgaggtac tgggtgctga agtcggagtc gtcgcagccg ccccgctccc	8940
agagcgagaa gtcggtgcgc ttcttgagc gggggttgagg cagagcgaag gtgacatcgt	9000
tgaagaggat tttgcccgcg cggggcatga agttgcgggt gatgcggaag ggccccggca	9060
cttcagagcg gttgttgatg acctgggcgg cgagcacgat ctcgtcgaag ccgttgatgt	9120
tgtggccac gatgtagagt tccaggaagc ggggccggcc ctttacgggtg ggcagcttct	9180
ttagctcttc gtaggtgagc tcctcgggcg aggcgaggcc gtgctcggcc agggcccagt	9240
ccgcgaggtg cgggttgtct ctgaggaagg acttcagag gtcgcgggcc aggagggtct	9300
gcaggcggtc tctgaaggtc ctgaactggc ggcccacggc cattttttcg ggggtgatgc	9360
agtagaaggt gagggggtct tgctgccagc ggtcccagtc gagctgcagg gcgaggtcgc	9420
gcgcggcggt gaccaggcgc tcgtcgcccc cgaatttcat gaccagcatg aagggcacga	9480
gctgctttcc gaaggccccc atccaagtgt aggtctctac atcgtagggtg acaaagaggc	9540
gctccgtgcg aggatgcgag ccgatcggga agaactggat ctcccggcac cagttggagg	9600

[0118]

agtggtctgtt gatgtggtgg aagtagaagt cccgtcgccg ggccgaacac tcgtgctggc	9660
ttttgtaaaa gcgagcgag tactggcagc gctgcacggg ctgtacctca tgcacgagat	9720
gcacctttcg cccgcgcacg aggaagccga ggggaaatct gagccccccg cctggctcgc	9780
ggcatggctg gttctcttct actttggatg cgtgtccgtc tccgtctggc tcctcgaggg	9840
gtgttacggt ggagcggacc accacgccgc gcgagccgca ggtccagata tcggcgcgcg	9900
gcggtcggag ttgatgacg acatcgcgca gctgggagct gtccatggtc tggagctccc	9960
gcggcggcgg caggtcagcc gggagtctct gcaggttcac ctgcagagt cgggccaggg	10020
cgcggggcag gtctaggtgg tactgatct ctaggggctg gttggtggcg gcgtcgatgg	10080
cttgcaggag cccgcagccc cggggggcga cgacggtgcc ccgcggggtg gtggtggtgg	10140
tggcgggtgca gctcagaagc ggtgccgcgg gcgggcccc ggaggtaggg ggggctccgg	10200
tcccgcgggc aggggcggca gcggcacgtc ggctggagc gcgggcagga gttggtgctg	10260
tgcccggagg ttgctggcga aggcgacgac gcggcggttg atctcctgga tctgggcct	10320
ctgcgtgaag acgacgggcc cggtgagctt gaacctgaaa gagagtcca cagaatcaat	10380
ctcggtgtca ttgaccgcgg cctggcgcag gatctcctgc acgtctcccg agttgtcttg	10440
gtaggcgata tcggccatga actgctcgat ctcttctcc tggaggtctc cgcgtccggc	10500
gcgttccacg gtggccgcca ggctggttga gatgcgcccc atgagctgcg agaaggcgtt	10560
gagtccgccc tcgttccaga ctgggtgta gaccacgcc ccctggtcac cgcgggcgcg	10620
catgaccacc tgcgcgaggt tgagctccac gtgccgcgcg aagacggcgt agttgcgcag	10680
acgctggaag aggtagtga ggggtggtggc ggtgtgctcg gccacgaaga agttcatgac	10740
ccagcggcgc aacgtggatt cgttgatgtc cccaaggcc tccagccgtt ccatggcctc	10800
gtagaagtcc acggcgaagt tgaaaaactg ggagttgcgc gccgacacgg tcaactctc	10860
ctccagaaga cggatgagct cggcgacggt gtcgcgcacc tcgcgctcga aggctatggg	10920
gatctcttcc tccgctagca tcaccacctc ctctcttcc tctcttctg gcacttccat	10980
gatggcttcc tctcttccg ggggtggcgg cggcggcggt gggggagggg gcgtctgcg	11040
ccggcggcgg cgcaccggga ggcggtccac gaagcgcgcg atcatctccc cgcggcgcg	11100
gcgcatggtc tcggtgacgg cgcggccgtt ctcccggggg cgcagttgga agacgccgc	11160
ggacatctgg tgctggggcg ggtggccgtg aggcagcgag acggcgctga cgatgcatt	11220
caacaattgc tgcgtaggta cgccgccgag ggacctgagg gagtccatat ccaccgcatc	11280

[0119]

cgaaaacctt	tcgaggaagg	cgtctaacca	gtcgcagtcg	caaggtaggc	tgagcaccgt	11340
ggcgggcggc	ggggggtggg	gggagtgtct	ggcggaggtg	ctgctgatga	tgtaattgaa	11400
gtaggcggac	ttgacacggc	ggatggtcga	caggagcacc	atgtccttgg	gtccggcctg	11460
ctggatgcgg	aggcggtcgg	ctatgcccc	ggcttcgttc	tggcatcggc	gcaggtcctt	11520
gtagtagtct	tgcatgagcc	tttccaccgg	cacctcttct	ccttcctctt	ctgcttcttc	11580
catgtctgct	tcggccctgg	ggcggcgccg	cgccccctg	cccccatgc	gcgtgacccc	11640
gaacccccctg	agcgggttga	gcagggccag	gtcggcgacg	acgcgctcgg	ccaggatggc	11700
ctgctgcacc	tgcgtgaggg	tggtttgaa	gtcatccaag	tccacgaagc	ggtggtaggc	11760
gcccgtgttg	atggtgtagg	tgcagttggc	catgacggac	cagttgacgg	tctggtggcc	11820
cggttgcgac	atctcggtgt	acctgagtcg	cgagtaggcg	cgggagtcga	agacgtagtc	11880
gttgcaagtc	cgcaccaggt	actggtagcc	caccaggaag	tgcggcggcg	gctggcggtg	11940
gaggggcccag	cgcaggggtg	cgggggctcc	gggggcccag	tcttccagca	tgaggcggtg	12000
gtaggcgtag	atgtacctgg	acatccaggt	gatacccgcg	gcggtgggtg	aggcgcgcgg	12060
gaagtcgcgc	acccggttcc	agatgttgcg	caggggcaga	aagtgctcca	tggtaggcgt	12120
gctctgtcca	gtcagacgcg	cgcagtcgtt	gatactctag	accagggaaa	acgaaagccg	12180
gtcagcgggc	actcttccgt	ggtctggtga	atagatcgca	agggtatcat	ggcggagggc	12240
ctcggttcga	gccccgggtc	cgggcccggc	ggtccgcat	gatccacgcg	gttaccgccc	12300
gcgtgtcgaa	cccaggtgtg	cgacgtcaga	caacggtgga	gtgttccttt	tggcgttttt	12360
ctggccgggc	gccggcgccg	cgtaagagac	taagccgcga	aagcgaaagc	agtaagtggc	12420
tcgctccccg	tagccggagg	gaccttgct	aagggttgcg	ttgcggcgaa	ccccggttcg	12480
aatcccgtac	tcgggcccgc	cggacccgcg	gctaaggtgt	tggattggcc	tccccctcgt	12540
ataaagaccc	cgtttgcgga	ttgactccgg	acacggggac	gagccccctt	tatttttget	12600
ttccccagat	gcatccggtg	ctgcggcaga	tgcgcccccc	gccccagcag	cagcaacaac	12660
accagcaaga	gcggcagcaa	cagcagcggg	agtcatgcag	ggccccctca	cccaccctcg	12720
gcgggcccgc	cacctcggcg	tccgcggccg	tgtctggcgc	ctgcggcggc	ggcggggggc	12780
cggctgacga	ccccgaggag	cccccgcggc	gcagggccag	acactacctg	gacctggagg	12840
agggcgaggg	cctggcgcg	ctgggggcgc	cgtctcccga	gcgccacccg	cgggtgcagc	12900
tgaagcgca	ctcgcgcgag	gcgtacgtgc	ctcggcagaa	cctgttcagg	gaccgcgcgg	12960
gcgaggagcc	caggagatg	cgggacagga	ggttcagcgc	agggcgggag	ctgcggcagg	13020

[0120]

ggctgaaccg cgagcggctg ctgcgcgagg aggactttga gcccgcgcg cggacgggga	13080
tcagccccgc gcgcgcgcac gtggcggccg ccgacctggt gacggcgtag gagcagacgg	13140
tgaaccagga gatcaacttc caaaagagtt tcaacaacca cgtgcgcacg ctggtggcgc	13200
gcgaggaggt gaccatcggg ctgatgcacc tgtgggactt tgtaagcgcg ctggtgcaga	13260
accccaacag caagcctctg acggcgcagc tgttctctgat agtgcagcac agcagggaca	13320
acgaggcggt tagggacgcg ctgctgaaca tcaccgagcc cgagggtcgg tggtctctgg	13380
acctgattaa catcctgcag agcatagtgg tgcaggagcg cagcctgagc ctggccgaca	13440
aggtggcggc catcaactac tcgatgctga gcctgggcaa gttttacgcg cgcaagatct	13500
accagacgcc gtacgtgccc atagacaagg aggtgaagat cgacggtttt tacatgcgca	13560
tggcgctgaa ggtgctcacc ctgagcgacg acctgggcgt gtaccgcaac gagcgcattc	13620
acaaggccgt gagcgtgagc cggcggcgcg agctgagcga ccgcgagctg atgcacagcc	13680
tgcagcgggc gctggcgggc gccggcagcg gcgacaggga ggccgagtc tacttcgatg	13740
cgggggcgga cctgcgctgg gcgcccagcc ggccggccct ggaggccgcg ggggtccgcg	13800
aggactatga cgaggacggc gaggaggatg aggagtacga gctagaggag ggcgagtacc	13860
tggactaaac cgcgggtggt gtttccggta gatgcaagac ccgaacgtgg tggaccggc	13920
gctgcgggcg gctctgcaga gccagccgtc cggccttaac tcctcagacg actggcgaca	13980
ggtcatggac cgcattcatgt cgctgacggc gcgtaaccgc gacgcgttc ggacgcagcc	14040
gcaggccaac aggctctccg ccattctgga ggccggtggtg cctgcgcgct cgaacccac	14100
gcacgagaag gtgctggcca tagtgaacgc gctggccgag aacagggcca tccgccgga	14160
cgaggccggg ctggtgtacg acgcgctgct gcagcgcgtg gcccgctaca acagcggcaa	14220
cgtgcagacc aacctggacc ggctgggtgg ggacgtgcgc gaggcggtgg cgcagcgcga	14280
gcgcgcggat cggcagggca acctgggctc catggtggcg ctgaatgcct tcctgagcac	14340
gcagccggcc aacgtgccgc gggggcagga agactacacc aactttgtga gcgcgctgcg	14400
gctgatgggtg accgagaccc ccagagcga ggtgtaccag tcgggcccgg actacttctt	14460
ccagaccagc agacagggcc tgcagacggt gaacctgagc caggctttca agaacctgcg	14520
ggggctgtgg ggcgtgaagg cgcaccagg cgaccgggcg acggtgtcca gcctgctgac	14580
gcccactcg cgctgctgc tgcctgctgat cgcgcgcttc acggacagcg gcagcgtgctc	14640
ccgggacacc tacctggggc acctgctgac cctgtaccgc gaggccatcg ggcaggcgca	14700

[0121]

ggtggacgag cacaccttcc aggagatcac cagcgtgagc cgcgcgctgg ggcaggagga	14760
cacgagcagc ctggaggcga ctctgaacta cctgctgacc aaccggcggc agaagattcc	14820
ctcgtgcac agcctgacct ccgaggagga gcgcattctg cgctacgtgc agcagagcgt	14880
gagcctgaac ctgatgcgcg acgggggtgac gcccagcgtg gcgctggaca tgaccgcgcg	14940
caacatggaa ccgggcatgt acgccgcgca ccggccttac atcaaccgcc tgatggacta	15000
cctgcatcgc gcggcgggcg tgaaccccga gtactttacc aacgccatcc tgaacccgca	15060
ctggctcccc cgccccgggt tctacagcgg gggtctcgag gtccccgaga ccaacgatgg	15120
cttcctgtgg gacgacatgg acgacagcgt gttctcccc cggccgcagg cgctggcgga	15180
agcgtccctg ctgcgtccca agaaggagga ggaggaggag gcgagtcgcc gccgcggcag	15240
cagcggcgtg gcttctctgt ccgagctggg ggccggcagcc gccgcgcgcc ccgggtccct	15300
gggcggcagc ccctttccga gcctggtggg gtctctgcac agcgagcgca ccacccgcc	15360
tcggctgctg ggcgaggacg agtacctgaa taactccctg ctgcagccgg tgcgggagaa	15420
aaacctgcct ccgccttcc ccaacaacgg gatagagagc ctggtggaca agatgagcag	15480
atggaagacc tatgcgcagg agcacaggga cgcgcctgcg ctccggccgc ccacgcggcg	15540
ccagcggcac gaccggcagc gggggctggt gtgggatgac gaggactccg cggacgatag	15600
cagcgtgctg gacctgggag ggagcggcaa ccggttcgcg cacctgcgcc ccgcctggg	15660
gaggatgttt taaaaaaaaa aaaaaaaagc aagaagcatg atgcaaaaat taaataaaac	15720
tcaccaaggc catggcgacc gagcgttggt ttcttgtgtt cccttcagta tgcggcgcg	15780
ggcgatgtac caggagggac ctctccctc ttacgagagc gtggtgggcg cggcggcggc	15840
ggcgccctct tctccctttg cgtcgcagct gctggagccg ccgtacgtgc ctccgcgcta	15900
cctgcggcct acggggggga gaaacagcat ccgttactcg gagctggcg ccctgttcga	15960
caccacccgg gtgtacctgg tggacaacaa gtcggcggac gtggcctccc tgaactacca	16020
gaacgaccac agcaattttt tgaccacggt catccagaac aatgactaca gcccagcga	16080
ggccagcacc cagaccatca atctggatga ccggtcgcac tggggcggcg acctgaaaac	16140
catcctgcac accaacaatgc ccaacgtgaa cgagttcatg ttcaccaata agttcaaggc	16200
gcgggtgatg gtgtcgcgct cgcacaccaa ggaagaccgg gtggagctga agtacgagt	16260
ggtggagttc gagctgccag agggcaacta ctccgagacc atgaccattg acctgatgaa	16320
caacgcgata gtggagcact atctgaaagt gggcaggcag aacgggggtcc tggagagcga	16380
catcggggtc aagttcgaca ccaggaactt ccgcctgggg ctggaccccc tgaccgggct	16440

[0122]

ggttatgccc	ggggtgtaca	ccaacgaggc	cttccatccc	gacatcatcc	tgctgcccgg	16500
ctgcggggtg	gacttcactt	acagccgcct	gagcaacctc	ctgggcatcc	gcaagcggca	16560
gcccttccag	gagggcttca	ggatcaccta	cgaggacctg	gaggggggca	acatccccgc	16620
gctcctcgat	gtggaggcct	accaggatag	cttgaaggaa	aatgaggcgg	gacaggagga	16680
taccgcccc	gccgcctccg	ccgccgccga	gcagggcgag	gatgctgctg	acaccgcggc	16740
cgcggacggg	gcagaggccg	accccgtat	ggtggtggag	gctcccgagc	aggaggagga	16800
catgaatgac	agtgcggtgc	gcggagacac	cttcgtcacc	cggggggagg	aaaagcaagc	16860
ggaggccgag	gccgcggccg	aggaaaagca	actggcggca	gcagcggcgg	cggcggcggtt	16920
ggccgcggcg	gaggctgagt	ctgaggggac	caagcccgcc	aaggagcccg	tgattaagcc	16980
cctgaccgaa	gatagcaaga	agcgcagtta	caacctgctc	aaggacagca	ccaacaccgc	17040
gtaccgcagc	tggtacctgg	cctacaacta	cggcgacccg	tcgacggggg	tgcgctcctg	17100
gaccctgctg	tgcacgccgg	acgtgacctg	cggctcggag	caggtgtact	ggtcgtgcc	17160
cgacatgatg	caagaccccg	tgaccttccg	ctccacgcgg	caggtcagca	acttcccgggt	17220
ggtgggcgcc	gagctgctgc	ccgtgcactc	caagagcttc	tacaacgacc	aggccgtcta	17280
ctcccagctc	atccgccagt	tcacctctct	gacccacgtg	ttcaatcgct	ttcctgagaa	17340
ccagattctg	gcgcgccccg	ccgccccac	catcaccacc	gtcagtgaag	acgttcctgc	17400
tctcacagat	cacgggacgc	taccgctgcg	caacagcatc	ggaggagtcc	agcgagtgc	17460
cgttactgac	gccagacgcc	gcacctgccc	ctacgtttac	aaggccttgg	gcatagtctc	17520
gccgcgcgtc	ctttccagcc	gcactttttg	agcaacacca	ccatcatgtc	catcctgac	17580
tcaccagca	ataactccgg	ctggggactg	ctgcgcgcgc	ccagcaagat	gttcggaggg	17640
gcgaggaagc	gttcgagca	gcaccccggtg	cgcgtgcgcg	ggcacttccg	cgccccctgg	17700
ggagcgcaca	aacgcggccg	cgcggggcgc	accaccgtgg	acgacgccat	cgactcggtg	17760
gtggagcagg	cgcgcaacta	caggcccgcg	gtctctaccg	tggacgcggc	catccagacc	17820
gtggtgcggg	gcgcgcggcg	gtacgccaag	ctgaagagcc	gccggaagcg	cgtggccccg	17880
cgccaccgcc	gccgacccgg	ggccgccgcc	aaacgcgcgc	ccgcggccct	gcttcgcggg	17940
gccaagcgca	cgggcgcgcg	cgcgcctatg	agggccgcgc	gccgcttggc	cgcgcgcac	18000
accgcgcgca	ccatggcccc	ccgtaccgga	agacgcgcgg	ccgcgcgcgc	cgcgcgcgcc	18060
atcagtgaca	tggccagcag	gcgcgggggc	aacgtgtact	gggtgcgcga	ctcggtgacc	18120

[0123]

ggcacgcgcg tgcccggtgcg cttccgcccc ccgcggactt gagatgatgt gaaaaaacia	18180
cactgagtct cctgctgttg tgtgtatccc agcggcgggcg gcgcgcgcag cgtcatgtcc	18240
aagcgcaaaa tcaaagaaga gatgctccag gtcgtcgcgc cggagatcta tgggcccccg	18300
aagaaggaag agcaggattc gaagccccgc aagataaagc gggtaaaaaa gaaaaagaaa	18360
gatgatgacg atgccgatgg ggaggtggag ttcctgcgcg ccacggcgcc caggcgcccc	18420
gtgcagtgga agggccggcg cgtaaagcgc gtcctgcgcc ccggcaccgc ggtggtcttc	18480
acgcccggcg agcgtccac ccggactttc aagcgcgtct atgacgaggt gtacggcgac	18540
gaagacctgc tggagcaggc caacgagcgc ttcggagagt ttgcttacgg gaagcgtcag	18600
cgggcgctgg ggaaggagga cctgctggcg ctgccgctgg accagggcaa cccaccccc	18660
agtctgaagc ccgtgacct gcagcaggtg ctgccgagca gcgcaccctc cgaggcgaag	18720
cggggtctga agcgcgaggc cggcgacctg gcgcccaccg tgcagctcat ggtgccaag	18780
cggcagaggc tggaggatgt gctggagaaa atgaaagtag accccggtct gcagccggac	18840
atcagggtcc gccccatcaa gcagggtggcg ccgggcctcg gcgtgcagac cgtggacgtg	18900
gtcatcccca ccggcaactc ccccgcgcgc gccaccacta ccgtgcctc cacggacatg	18960
gagacacaga ccgatcccg cgcagccgca gccgcagccg ccgccgcgac ctctcggcg	19020
gaggtgcaga cggacccctg gctgccgcgc gcgatgtcag ctccccgcgc gcgtcgcggg	19080
cgcaggaagt acggcgccgc caacgcgctc ctgcccagat acgccttgca tccttcac	19140
gcgcccaccc ccggtaccg aggtataacc taccgcccgc gaagagccaa gggttccacc	19200
cgcgctcccc gccgacgcgc cgcgccacc acccgccgc gccgccgcag acgccagccc	19260
gcactggctc cagtctccgt gaggaaagtg gcgcgcgacg gacacaccct ggtgctgcc	19320
agggcgcgct accaccccag catcgtttaa aagcctgttg tggttcttgc agatatggcc	19380
ctcacttgcc gcctccgttt cccggtgccg ggataccgag gaggaagatc gcgccgcagg	19440
aggggtctgg ccggccgcgg cctgagcggg ggcagccgc gcgcgcaccg gcggcgacgc	19500
gccaccagcc gacgcatgcg cggcgggggtg ctgcccctgt taatccccct gatcgccgcg	19560
gcgatcggcg ccgtgcccgg gatcgccctc gtggccttgc aagcgtccca gaggcattga	19620
cagacttgca aacttgcaaa tatggaaaaa aaaaccccaa taaaaaagtc tagactctca	19680
cgtcgccttg gtctgtgac tatttttag aatggaagac atcaactttg cgtcgtggc	19740
cccgcgtcac ggctcgcgcc cgttcctggg aactggaac gatatcgga ccagcaacat	19800
gagcgggtggc gccttcagtt ggggctctct gtggagcggc attaaaagta tcgggtctgc	19860

[0124]

cgttaaaaat tacggctccc gggcctggaa cagcagcacg ggccagatgt tgagagacaa	19920
gttgaaagag cagaacttcc agcagaaggt ggtggagggc ctggcctccg gcatcaacgg	19980
ggtggtggac ctggccaacc aggccgtgca gaataagatc aacagcagac tggacccccg	20040
gccgccggtg gaggaggtgc cgccggcgct ggagacggtg tccccgatg ggcggtggcga	20100
gaagcgcccc cgccccgata gggaagagac cactctggtc acgcagaccg atgagccgcc	20160
cccgtatgag gaggccctga agcaaggtct gccaccacg cgccccatcg cgcccatggc	20220
caccgggggtg gtgggcccgc acacccccgc cacgctggac ttgcctccgc ccgccgatgt	20280
gccgcagcag cagaaggcgg cacagccggg cccgcccgcg accgcctccc gttcctccgc	20340
cggtcctctg cgccgcgcgg ccagcggccc ccgcgggggg gtcgcgaggc acggcaactg	20400
gcagagcacg ctgaacagca tcgtgggtct ggggggtgcgg tccgtgaagc gccgccgatg	20460
ctactgaata gcttagctaa cgtgttgtat gtgtgtatgc gccctatgtc gccgccagag	20520
gagctgctga gtcgccgcgg ttgcgcgcgc caccaccacc gccactccgc ccctcaagat	20580
ggcgacccca tcgatgatgc cgcagtggtc gtacatgcac atctcgggcc aggacgcctc	20640
ggagtacctg agccccgggc tgggtgcagtt cgcccgcgcc accgagagct acttcagcct	20700
gagtaacaag tttaggaacc ccacggtggc gccacgcac gatgtgacca ccgaccggtc	20760
tcagcgcctg acgctgcggt tcattcccgt ggaccgcgag gacaccgcgt actcgtacaa	20820
ggcgcggttc accctggccg tgggcgacaa ccgcgtgctg gacatggcct ccacctactt	20880
tgacatccgc ggggtgctgg accggggtcc cactttcaag ccctactctg gcaccgccta	20940
caactccctg gcccccaagg gcgctcccaa ctctgcgag tgggagcaag aggaaactca	21000
ggcagttgaa gaagcagcag aagaggaaga agaagatgct gacggtcaag ctgaggaaga	21060
gcaagcagct accaaaaaga ctcatgtata tgctcaggct cccctttctg gcgaaaaaat	21120
tagtaaagat ggtctgcaaa taggaacgga cgctacagct acagaacaaa aacctattta	21180
tgcagaccct acattccagc ccgaacccca aatcggggag tcccagtgga atgaggcaga	21240
tgctacagtc gccggcggtg gagtgctaaa gaaatctact cccatgaaac catgctatgg	21300
ttcctatgca agaccacaaa atgctaattg aggtcagggt gtactaacgg caaatgceca	21360
gggacagcta gaatctcagg ttgaaatgca attcttttca acttctgaaa acgcccgtaa	21420
cgaggctaac aacattcagc ccaaattggt gctgtatagt gaggatgtgc acatggagac	21480
cccggatacg cacctttctt acaagccgcg aaaaagcgat gacaattcaa aaatcatgct	21540

[0125]

gggtcagcag tccatgcccc acagacctaa ttacatcggc ttcagagaca actttatcgg	21600
cctcatgtat tacaatagca ctggcaacat gggagtgttt gcaggtcagg cctctcagtt	21660
gaatgcagtg gtggacttgc aagacagaaa cacagaactg tcctaccage tcttgcttga	21720
ttccatgggt gacagaacca gatacttttc catgtggaat caggcagtgg acagttatga	21780
cccagatgtt agaattattg aaaatcatgg aactgaagac gagctcccca actattgttt	21840
ccctctgggt ggcatagggg taactgacac ttaccaggct gttaaaacca acaatggcaa	21900
taacgggggc caggtgactt ggacaaaaga tgaaactttt gcagatcgca atgaaatagg	21960
ggtgggaaac aatttcgcta tggagatcaa cctcagtgcc aacctgtgga gaaacttcct	22020
gtactccaac gtggcgctgt acctaccaga caagcttaag tacaaccctt ccaatgtgga	22080
catctctgac aacccaaca cctacgatta catgaacaag cgagtgggtg ccccggggct	22140
ggtggactgc tacatcaacc tgggcgcgcg ctggctcgtg gactacatgg acaacgtcaa	22200
ccccttcaac caccaccgca atgcgggcct gcgtaccgc tccatgctcc tgggcaacgg	22260
gcgtacgtg cccctccaca tccaggtgcc ccagaagttc ttgccatca agaactcct	22320
cctcctgccg ggctcctaca cctacgagtg gaacttcagg aaggatgtca acatggtcct	22380
ccagagctct ctgggtaacg atctcagggt ggacggggcc agcatcaagt tcgagagcat	22440
ctgcctctac gccaccttct tccccatggc ccacaacacg gcctccacgc tcgaggccat	22500
gtcaggaac gacaccaacg accagtcctt caatgactac ctctccgcgc ccaacatgct	22560
ctaccccata cccgccaacg ccaccaacgt ccccatctcc atccccctgc gcaactgggc	22620
ggccttccgc ggctgggcct tcacccgcct caagaccaag gagacccctt ccctgggctc	22680
gggattcgac cctactaca cctactcggg ctccattccc tacctggacg gcaccttcta	22740
cctcaaccac actttcaaga aggtctcgggt caccttcgac tcctcgggtca gctggccggg	22800
caacgaccgt ctgctcacc ccaacgagtt cgagatcaag cgctcgggtc acggggaggg	22860
ctacaacgtg gcccagtgc acatgaccaa ggactgggtc ctgggtccaga tgctggccaa	22920
ctacaacatc ggctaccagg gcttctacat ccagagagc tacaaggaca ggatgtactc	22980
cttcttcagg aacttccagc ccatgagccg gcaggtgggt gaccagacca agtacaagga	23040
ctaccaggag gtgggcatca tccaccagca caacaactcg ggcttcgtgg gctacctgc	23100
ccccaccatg cgcgagggac aggcctaccc cgccaacttc ccctatccgc tcataggcaa	23160
gaccgcggtc gacagcatca cccagaaaaa gttcctctgc gaccgcaccc tctggcgcac	23220
ccccttctcc agcaacttca tgtccatggg tgcgtctcgc gacctgggcc agaacttgct	23280

[0126]

ctacgccaac tccgcccacg ccctcgacat gaccttcgag gtcgacccca tggacgagcc	23340
cacccttctc tatgttctgt tcgaagtctt tgacgtggtc cgggtccacc agccgcaccg	23400
cggcgtcatc gagaccgtgt acctgcgtac gcccttctcg gccggcaacg ccaccaccta	23460
aagaagcaag ccgcagtcac cgccgcctgc atgccgtcgg gtccaccga gcaagagctc	23520
agggccatcg tcagagacct gggatgcggg ccctattttt tgggcacctt cgacaagegc	23580
ttccctggct ttgtctcccc acacaagctg gcctgcgcca tcgtcaacac ggccggccgc	23640
gagaccgggg gcgtgcactg gctggccttc gcctggaacc cgcgtccaa aacatgcttc	23700
ctctttgacc ccttcggctt ttccgaccag cggtcgaagc aaatctacga gtccgagtac	23760
gagggttgc tgcgtcgcag cgccatcgcc tcctcgcccg accgtgcgt caccctcgaa	23820
aagtcaccc agaccgtgca ggggcccgcac tcggccgcct gcggtctctt ctgctgcatg	23880
tttctgcacg cctttgtgca ctggcctcag agtcccatgg accgcaaccc caccatgaac	23940
ttgctgacgg ggggtgcccc ctccatgctc cagagccccc aggtcgagcc caccctgcgc	24000
cgcaaccagg agcagctcta cagcttcctg gagcgccact cgccttactt ccgccgccac	24060
agcgcacaga tcaggagggc caccctcttc tgccacttgc aagagatgca agaagggtaa	24120
taacgatgta cacacttttt ttctcaataa atggcatctt tttatttata caagctctct	24180
gggggtattca ttcccacca ccaccgccg ttgtcgccat ctggctctat ttagaaatcg	24240
aaagggttct gccgggagtc gccgtgcgc accgggcagg acacgttgcg atactggtag	24300
cgggtgcccc acttgaactc gggcaccacc aggcgaggca gctcggggaa gttttgcctc	24360
cacaggctgc gggtcagcac cagcgcgttc atcaggtcgg gcgccgagat cttgaagtgc	24420
cagttggggc cgccgccctg cgcgcgcgag ttgcggtaca ccgggttgca gcactggaac	24480
accaacagcg ccgggtgctt cagcttgccc agcacgctgc ggtcggagat cagctcggcg	24540
tccaggctct ccgcttgct cagcgcgaac ggggtcatct tgggcacttg ccgcccagg	24600
aagggcgcgt gccccggttt cgagttgcag tcgcagcgca gcgggatcag caggtgcccg	24660
tgcccgact cggcgttggg gtacagcgcg cgcataagg cctgcatctg gcggaaggcc	24720
atctgggcct tggcgccctc cgagaagaac atgccgcagg acttgcccga gaactggttt	24780
gcggggcagc tggcgctgtg caggcagcag cgcgcgtcgg tgttgcgat ctgcaccacg	24840
ttgcgcccc accggttctt cagcatcttg gccttgacg attgctcctt cagcgcgcgc	24900
tgcccgcttct cgttggtcac atccatctcg atcacatgtt ccttgttcac catgctgctg	24960

[0127]

ccgtgcagac	acttcagctc	gccctccgtc	tcggtgcagc	ggtgctgcc	cagcgcgcag	25020
cccgtgggct	cgaaagactt	gtaggtcacc	tccgcgaagg	actgcaggta	cccctgcaaa	25080
aagcggccca	tcatggtcac	gaaggtcttg	ttgctgctga	aggtcagctg	cagcccgcgg	25140
tgctcctcgt	tcagccaggt	cttgcacacg	gccgccagcg	cctccacctg	gtcgggcagc	25200
atcttgaagt	tcaccttcag	ctcattctcc	acgtgggtact	tgtccatcag	cgtgcgcgcc	25260
gcctccatgc	ccttctccca	ggccgacacc	agcggcaggc	tcacgggggtt	cttcaccatc	25320
accgtggcgc	cgccctccgc	cgcgctttcg	ctttccgccc	cgctgttctc	ttcctcttcc	25380
tcctcttctc	cgccgccgcc	cactcgcagc	ccccgcacca	cggggtcgtc	ttcctgcagg	25440
cgctgcacct	tgcgcttgcc	gttgcgcccc	tgcttgatgc	gcacgggcgg	gttgctgaag	25500
cccaccatca	ccagcgcggc	ctcttcttgc	tcgtcctcgc	tgtccagaat	gacctccggg	25560
gagggggggt	tggtcatcct	cagtaccgag	gcacgcttct	ttttcttctc	gggggcgttc	25620
gccagctccg	cggtgcggc	cgctgccgag	gtcgaaggcc	gagggtggg	cgtgcgcggc	25680
accagcgcgt	cctgcgagcc	gtcctcgtcc	tcctcggact	cgagacggag	gcgggcccgc	25740
ttcttcgggg	gcgcgcgggg	cggcggaggc	ggcggcggcg	acggagacgg	ggacgagaca	25800
tcgtccaggg	tgggtggacg	gcgggcgcgc	ccgcgtccgc	gctcgggggt	ggtctcgcgc	25860
tggtcctctt	cccgactggc	catctcccac	tgctccttct	cctataggca	gaaagagatc	25920
atggagtctc	tcatgcgagt	cgagaaggag	gaggacagcc	taaccgcccc	ctctgagccc	25980
tccaccaccg	cgccaccac	cgccaatgcc	gccgcggacg	acgcgcccac	cgagaccacc	26040
gccagtacca	ccctccccag	cgacgcaccc	ccgctcgaga	atgaagtgct	gatcgagcag	26100
gacctcgggtt	ttgtgagcgg	agaggaggat	gaggtggatg	agaaggagaa	ggaggaggtc	26160
gccgcctcag	tgccaaaaga	ggataaaaag	caagaccagg	acgacgcaga	taaggatgag	26220
acagcagtcg	ggcgggggaa	cggaagccat	gatgctgatg	acggctacct	agacgtggga	26280
gacgacgtgc	tgcttaagca	cctgcaccgc	cagtgcgtca	tcgtctgcga	cgcgctgcag	26340
gagcgtgcg	aagtccccct	ggacgtggcg	gaggtcagcc	gcgcctacga	gcggcacctc	26400
ttcgcgccgc	acgtgcccc	caagcgccgg	gagaacggca	cctgcgagcc	caacctcgcgt	26460
ctcaacttct	acctgggtctt	cgcggtaccc	gaggtgctgg	ccacctacca	catctttttc	26520
caaaactgca	agatccccct	ctcctgcgcg	gccaacgcga	cccgcgccga	caaaacctg	26580
acctgcggc	agggcgccca	catacctgat	atgcctctc	tggaggaagt	gcccagatc	26640
ttcgagggtc	tcggtcgcga	cgagaaacgg	gcggcgaacg	ctctgcacgg	agacagcgaa	26700

[0128]

aacgagagtc actcgggggt gctggtggag ctcgagggcg acaacgcgcg cctggccgta	26760
ctcaagcgca gcatagaggt caccacttt gcctaccgg cgctcaacct gcccccaag	26820
gtcatgagtg tggatcatgg cgagctcatc atgcgcgcg cccagcccct ggccgcggat	26880
gcaaacttgc aagagtcctc cgaggaaggc ctgcccgcgg tcagcgacga gcagctggcg	26940
cgctggctgg agaccgcga cccgcgcgag ctggaggagc ggcgcaagct catgatggcc	27000
gcggtgctgg tcaccgtgga gctcgagtgt ctgcagcgt tcttcgcgga ccccgagatg	27060
cagcgcaagc tcgaggagac cctgcactac accttcgcc agggctacgt gcgccaggcc	27120
tgaagatct ccaacgtgga gctctgcaac ctggtctcct acctgggcat cctgcacgag	27180
aaccgcctcg ggcagaacgt cctgcactcc acctcaaag gggaggcgcg ccgcgactac	27240
atccgcgact gcgcctacct ctctctctgc tacacctggc agacggccat gggggtctgg	27300
cagcagtgcc tggaggagcg caacctcaag gagctggaaa agctcctcaa gcgcaccctc	27360
agggacctct ggacgggctt caacgagcgc tcggtggccg ccgcgctggc ggacatcatc	27420
tttcccgagc gcctgctcaa gacctgcag cagggcctgc ccgacttcac cagccagagc	27480
atgctgcaga acttcaggac ttctatcctg gagcgctcgg gcatacctgcc ggccacttgc	27540
tgcgcgtgc ccagcgactt cgtgcccac aagtacagg agtgcccgc gccgctctgg	27600
ggccactgct acctcttcca gctggccaac tacctgcct accactcgga cctcatggaa	27660
gacgtgagcg gcgaggcct gctcgagtgc cactgccgt gcaacctctg cacgccccac	27720
cgctctctag tctgcaacc gcagctgctc agcgagagtc agattatcgg taccttcgag	27780
ctgcagggtc cctgcctga cgagaagtcc gcggctccag ggctgaaact cactccgggg	27840
ctgtggactt ccgcctacct acgcaaattt gtacctgagg actaccacgc ccacgagatc	27900
aggttctacg aagaccaatc ccgcccgcc aaggcggagc tcaccgcctg cgtcatcacc	27960
caggggcaca tcctgggcca attgcaagcc atcaacaaag cccgccgaga gttcttctg	28020
aaaaagggtc ggggggtgta cctggacccc cagtccggcg aggagctaaa cccgctaccc	28080
ccgccgcgc cccagcagcg ggaccttgc tcccaggatg gcaccagaa agaagcagca	28140
gccgccgcg ccgccgcgc catacatgct tctggaggaa gaggaggagg actgggacag	28200
tcaggcagag gaggtttcgg acgaggagca ggaggagatg atggaagact gggaggagga	28260
cagcagccta gacgaggaag cttcagaggc cgaagaggtg gcagacgcaa caccatcgcc	28320
ctcggtcgca gccccctgc cggggccctt gaaatcctcc gaaccagca ccagcgtat	28380

[0129]

aacctccgct cctccggcgc cggcgccacc cgccccgaga cccaaccgta gatgggacac	28440
cacaggaacc ggggtcggta agtccaagtg cccgccgccg ccaccgcagc agcagcagca	28500
gcagcgccag ggctaccgct cgtggcgcg gcaacaagaac gccatagtcg cctgcttgca	28560
agactgcggg ggcaacatct ctttcgcccc cgcttcctg ctattccacc acggggtcgc	28620
ctttccccgc aatgtcctgc attactaccg tcattctctac agccccctact gcagcggcga	28680
cccagaggcg gcagcggcag ccacagcggc gaccaccacc taggaagata tcctccgcgg	28740
gcaagacagc ggcagcagcg gccaggagac ccgcggcagc agcggcggga gcggtgggcg	28800
cactgcgct ctcgccccac gaacccctct cgacccggga gtcagacac aggatcttcc	28860
ccactttgta tgccatcttc caacagagca gaggccagga gcaggagctg aaaataaaaa	28920
acagatctct gcgtccctc acccgcagct gtctgtatca caaaagcgaa gatcagcttc	28980
ggcgcacgct ggaggacgcg gaggcactct tcagcaaata ctgcgcgctc actcttaaag	29040
actagctccg cgcccttctc gaatttaggc gggagaaaac tacgtcatcg ccggccgccg	29100
cccagccccgc ccagccgaga tgagcaaaga gattcccacg ccatacatgt ggagctacca	29160
gccgcagatg ggactcgcgg cgggagcggc ccaggactac tccacccgca tgaactacat	29220
gagcgcggga ccccatga tctcacaggt caacgggatc cgcgcccgagc gaaaccaa	29280
actgctggaa caggcggcca tcaccgccac gccccgcat aatctcaacc cccgaaattg	29340
gccccgccg ctcgtgtacc aggaacccc ctccgccacc accgtactac ttccgcgtga	29400
cgccccaggcc gaagtccaga tgactaactc aggggcgcag ctgcggggcg gctttcgtca	29460
cggggcgcg cgcgtccgac caggtataag acacctgatg atcagaggcc gaggtatcca	29520
gctcaacgac gagtcggtga gctcttcgct cggtctccgt ccggacggaa ctttcagct	29580
cgccggatcc ggccgctctt cgttcacgcc ccgccaggcg tacctgactc tgcagacctc	29640
gtcctcggag ccccgtccg gcggcatcgg aacctccag ttcgtggagg agttcgtgcc	29700
ctcggctctac ttcaaccct tctcgggacc tcccggacgc taccgccacc agttcattcc	29760
gaactttgac gcggtgaagg actcggcgga cggtacgac tgaatgtcag gtgtcgaggc	29820
agagcagctt cgcctgagac acctcgagca ctgccgccgc cacaagtgt tcgccccggg	29880
ttctggtgag ttctgtact ttcagctacc cgaggagcat accgaggggc cggcgcacgg	29940
cgtcgcctg accaccagg gcgaggttac ctgttcctc atccgggagt ttacctccg	30000
tccctgcta gtggagcggg agcggggctc ctgtgtccta actatcgct gcaactgcc	30060
taacctgga ttacatcaag atctttgctg tcattctctgt gctgagttta ataaacgtg	30120

[0130]

agatcagaat ctactggggc tcctgtcgcc atcctgtgaa cgccaccgtc ttcacccacc	30180
ccgaccaggc ccaggcgaac ctcacctgcg gtctgcatcg gagggccaag aagtacctca	30240
cctgggtactt caacggcacc ccccttgtgg tttaacaacag cttcgacggg gacggagtct	30300
ccctgaaaga ccagctctcc ggtctcagct actccatcca caagaacacc accctccaac	30360
tcttccctcc ctacctgccg ggaacctacg agtgcgtcac cggecgctgc acccacctca	30420
cccgctgat cgtaaaccag agctttccgg gaacagataa ctccctcttc ccagaacag	30480
gaggtgagct caggaaactc cccggggacc agggcggaga cgtaccttcg acccttgtgg	30540
ggttaggatt ttttattacc gggttgctgg ctcttttaaat caaagtttcc ttgagatttg	30600
ttctttcctt ctacgtgtat gaacacctca acctccaata actctaccct ttcttcggaa	30660
tcaggtgact tctctgaaat cgggcttggg gtgctgctta ctctgttgat ttttttcctt	30720
atcatactca gccttctgtg cctcaggctc gccgcctgct gcgcacacat ctatatctac	30780
tgctggttgc tcaagtgcag gggtcgccac ccaagatgaa caggtacatg gtcctatcga	30840
tcctaggcct gctggccctg gcggcctgca gcgccgcaa aaaagagatt acctttgagg	30900
agcccgttg caatgtaact ttcaagcccg agggtgacca atgcaccacc ctctgcaa	30960
gcgttaccaa tcatgagagg ctgcgcatcg actacaaaaa caaaactggc cagtttgccg	31020
tctatagtgt gtttacgcc ggagaccct ctaactactc tgtcaccgtc ttccagggcg	31080
gacagtctaa gatattcaat tacactttcc ctttttatga gttatgcgat gcggtcatgt	31140
acatgtcaaa acagtacaac ctgtggcctc cctctcccca ggcgtgtgtg gaaaatactg	31200
ggctttactg ctgtatggct ttcgcaatca ctacgctgc tctaatctgc acggtgctat	31260
acataaaatt caggcagagg cgaatcttta tcgatgaaaa gaaaatgcct tgatcgctaa	31320
caccggcttt ctatctgcag aatgaatgca atcacctccc tactaatcac caccaccctc	31380
cttgcgattg cccatgggtt gacacgaatc gaagtgccag tggggtccaa tgtcaccatg	31440
gtgggccccg ccggcaattc caccctcatg tgggaaaaat ttgtccgcaa tcaatgggtt	31500
catttctgct ctaaccgaat cagtatcaag ccagagcca tctgcgatgg gcaaaatcta	31560
actctgatca atgtgcaaat gatggatgct ggggtactatt acgggcagcg gggagaaatc	31620
attaattact ggcgacccca caaggactac atgttgcgat tagtcgaggc acttccact	31680
accaccccca ctaccacctc tcccaccacc accaccacta ctactactac tactactact	31740
actactacta ccactaccgc tgcccgccat acccgcaaaa gcaccatgat tagcacaaag	31800

[0131]

ccccctcgtg ctcactccca cgccggcggg cccatcggtg cgacctcaga aaccaccgag	31860
ctttgtcttct gccaatgcac taacgccagc gctcatgaac tgttcgacct ggagaatgag	31920
gatgtccagc agagctccgc ttgcctgacc caggaggctg tggagcccgt tgcctgaag	31980
cagatcggtg attcaataat tgactcttct tcttttgcca ctcccgaata ccctcccgat	32040
tctactttcc acatcacggg taccaaagac cctaacctct ctttctacct gatgctgctg	32100
ctctgtatct ctgtggtctc ttccgcgtg atgttactgg ggatgttctg ctgcctgac	32160
tgccgcagaa agagaaaagc tcgtctcag ggccaaccac tgatgccctt cccctacccc	32220
ccggaatttg cagataacaa gatatgagct cgctgctgac actaacgct ttactagcct	32280
gcgctctaac cettgtcgtc tgcgactcga gattccacaa tgtcacagct gtggcaggag	32340
aaaatgttac tttcaactcc acggccgata cccagtggtc gtggagtggc tcaggtagct	32400
acttaactat ctgcaatagc tccacttccc ccggcatatc cccaaccaag taccaatgca	32460
atgccagcct gttcaccctc atcaacgctt ccacctgga caatggactc tatgtagget	32520
atgtaccctt tgggtgggcaa ggaaagaccc acgcttaciaa cctggaagtt cgccagccca	32580
gaaccactac ccaagcttct cccaccacca ccaccaccac caccatcacc agcagcagca	32640
gcagcagcag ccacagcagc agcagcagat tattgacttt ggttttggcc agctcatctg	32700
ccgctaccca ggccatctac agctctgtgc ccgaaaccac tcagatccac cgcccagaaa	32760
cgaccaccgc caccacccta cacacctcca gcgatcagat gccgaccaac atcacccct	32820
tggtctttca aatgggactt acaagcccca ctccaaaacc agtggatgcg gccgaggtct	32880
ccgccctcgt caatgactgg gcggggctgg gaatgtggtg gttcgccata ggcatgatgg	32940
cgctctgcct gttctgctc ttggtcatct gctgcctcca ccgcaggcga gccagacccc	33000
ccatctatag acccatcatt gtctgaacc ccgataatga tgggatccat agattggatg	33060
gcctgaaaaa cctacttttt tcttttacag tatgataaat tgagacatgc ctgcatttt	33120
cttgtacatg ttccttctcc caccctttct ggggtgttct acgctggccg ctgtgtctca	33180
cctggaggta gactgcctct cacccttcac tgtctacctg ctttacggat tggtcaccct	33240
cactctcatc tgcagcctaa tcacagtaat catcgccttc atccagtga ttgattacat	33300
ctgtgtgcgc ctgcataact tcagacacca ccgcagtag cgagacagga acattgcccc	33360
acttctaaga ctgtctaat catgcataag actgtgatct gccttctgat cctctgcate	33420
ctgcccaccc tcacctcctg ccagtagacc acaaaatctc cgcgcaaaag acatgcctcc	33480
tgcgcgttca cccaactgtg gaatataccc aaatgctaca acgaaaagag cgagctctcc	33540

[0132]

gaagcttggc tgtatggggt catctgtgtc ttagttttct gcagcactgt ctttgcctc	33600
ataatctacc cctactttga tttgggatgg aacgcgatcg atgccatgaa ttacccacc	33660
tttcccgcac ccgagataat tccactgcga caagtgttac ccgttgtcgt taatcaacgc	33720
ccccatccc ctacgccac tgaaatcagc tactttaacc taacaggcgg agatgactga	33780
cgccctagat ctagaaatgg acggcatcag taccgagcag cgtctcctag agaggcgcag	33840
gcaggcggct gagcaagagc gcctcaatca ggagctccga gatctcgta acctgcacca	33900
gtgcaaaaga ggcattttt gtctggtaaa gcaggccaaa gtcacctacg agaagaccgg	33960
caacagccac cgcctcagtt acaaattgcc caccagcgc cagaagctgg tgctcatggt	34020
gggtgagaat cccatcacccg tcaccagca ctcggtagag accgaggggt gtctgcactc	34080
cccctgtcgg ggtccagaag acctctgcac cctggtaaag acctgtgcg gtctcagaga	34140
tttagtcccc tttaactaat caaacactgg aatcaataaa aagaatcact tacttaaat	34200
cagacagcag gtctctgtcc agtttattca gcagcacctc cttccctcc tcccaactct	34260
ggtactccaa acgccttctg gcggcaaact tcctccacac cctgaaggga atgtcagatt	34320
cttgctcctg tccctccgca cccactatct tcatgttggt gcagatgaag cgcacaaaa	34380
cgtctgacga gagcttcaac cccgtgtacc cctatgacac ggaaagcggc cctccctccg	34440
tcccttctct caccctccc ttctgtctc ccgatggatt ccaagaaagt ccccccggg	34500
tcctgtctct gaacctggcc gagcccctgg tcaattccca cggcatgctc gccctgaaaa	34560
tgggaagtgg cctctccctg gacgacgtg gcaacctcac ctctcaagat atcaccaccg	34620
ctagccctcc cctcaaaaa accaagacca acctcagcct agaaacctca tccccctaa	34680
ctgtgagcac ctccaggccc ctccacgtag cagccgccgc tcccctggcg gtggccggca	34740
cctccctcac catgcaatca gagggcccc tgacagtaca ggatgcaaaa ctccacctgg	34800
ccaccaaagg cccctgacc gtgtctgaag gcaaactggc cttgcaaaca tcggccccgc	34860
tgacggccgc tgacagcagc acctcacag tcagtgccac accaccctt agcacaagca	34920
atggcagctt gggtattgac atgcaagccc ccatttacac caccaatgga aaactaggac	34980
ttaacttttg cgtccccctg catgtggtag acagcctaaa tgcactgact gtagttactg	35040
gccaaagtct tacgataaac ggaacagccc tacaactag agtctcaggt gccctcaact	35100
atgacacatc aggaaacct gaattgagag ctgcagggg tatgcagatt gatgcaaatg	35160
gtcaacttat ccttgatgta gcttaccat ttgatgcaca aaacaatctc agccttaggc	35220

[0133]

ttggacaggg	acccctgttt	gttaactctg	cccacaactt	ggatgttaac	tacaacagag	35280
gcctctacct	gttcacatct	ggaaatacca	aaaagctaga	agttaatatc	aaaacagcca	35340
agggtctcat	ttatgatgac	actgctatag	caatcaatgc	gggtgatggg	ctacagtttg	35400
actcaggctc	agatacaaat	ccattaaaaa	ctaaacttgg	attaggactg	gattatgact	35460
ccagcagagc	cataattgct	aaactgggaa	ctggcctaag	ctttgacaac	acaggtgcca	35520
tcacagtagg	caacaaaaat	gatgacaagc	ttaccttggt	gaccacacca	gacccatccc	35580
ctaactgtag	aatctattca	gagaaagatg	ctaaattcac	acttgttttg	actaaatgcg	35640
gcagtcaggt	gttggccagc	gtttctgttt	tatctgtaaa	aggtagcctt	gcgccccatca	35700
gtggcacagt	aactagtgtc	cagattgtcc	tcagatttga	tgaaaatgga	gttctactaa	35760
gcaattcttc	ccttgaccct	caatactgga	actacagaaa	aggtagcctt	acagagggca	35820
ctgcatatac	caacgcagtg	ggatttatgc	ccaacctcac	agcataccca	aaaacacaga	35880
gccaaactgc	taaaagcaac	attgtaagtc	aggtttactt	gaatggggac	aaatccaaac	35940
ccatgaccct	caccattacc	ctcaatggaa	ctaatgaaac	aggagatgcc	acagtaagca	36000
cttactccat	gtcattctca	tggaactgga	atggaagtaa	ttacattaat	gaaacgttcc	36060
aaaccaactc	cttcaccttc	tcctacatcg	ccaagaata	aaaagcatga	cgctgttgat	36120
ttgattcaat	gtgtttctgt	tttattttca	agcacaacaa	aatcattcaa	gtcattcttc	36180
catcttagct	taatagacac	agtagcttaa	tagaccagct	agtgc aaagc	cccattctag	36240
cttataacta	gtggagaagt	actcgcctac	atgggggtag	agtcataatc	gtgcatcagg	36300
atagggcggg	ggtgctgcag	cagcgcgcga	ataaactgct	gccgccgccg	ctccgtcctg	36360
caggaataca	acatggcagct	ggtctcctca	gcgatgattc	gcaccgcccc	cagcataagg	36420
cgccttgtcc	tccgggcaca	gcagcgcacc	ctgatctcac	ttaaatacagc	acagtaactg	36480
cagcacagca	ccacaatatt	gttcaaaaatc	ccacagtgca	aggcgctgta	tccaaagctc	36540
atggcgggga	ccacagaacc	cacgtggcca	tcataccaca	agcgcaggta	gattaagtgg	36600
cgacccctca	taaacacgct	ggacataaac	attacctctt	ttggcatggt	gtaattcacc	36660
acctcccggg	accatataaa	cctctgatta	aacatggcgc	catccaccac	catcctaaac	36720
cagctggcca	aaacctgccc	gccggctata	cactgcaggg	aaccgggact	ggaacaatga	36780
cagtggagag	cccaggactc	gtaaccatgg	atcatcatgc	tcgtcatgat	atcaatgttg	36840
gcacaacaca	ggcacacgtg	catacacttc	ctcaggatta	caagctcctc	ccgcgttaga	36900
accatatccc	agggaacaac	ccattcctga	atcagcgtaa	atcccacact	gcagggaaga	36960

[0134]

cctgcacgt aactcacgtt gtgcattgtc aaagtgttac attcgggcag cagcggatga	37020
tcctccagta tggtagcgcg ggtttctgtc tcaaaaggag gtagacgac cctactgtac	37080
ggagtgcgcc gagacaaccg agatcgtgtt ggtcgtagtg tcatgccaaa tggaacgccg	37140
gacgtagtca tatttcctga agtccttagat ctctcaacgc agcaccagca ccaacacttc	37200
gcagtgtaaa aggccaagtg ccgagagagt atatatagga ataaaaagtg acgtaaacgg	37260
gcaaagtcca aaaaacgccc agaaaaaccg cacgcgaacc tacgccccga aacgaaagcc	37320
aaaaaacact agacactccc ttccggcgtc aacttccgct ttcccacgct acgtcacttg	37380
ccccagtcaa acaaactaca tatcccgaac ttccaagtcg ccacgcccaa aacaccgcct	37440
acacctcccc gcccgccggc ccgcccccaa acccgccctc cgccccgcgc cccgccccgc	37500
gccgcccatc tcattatcat attggcttca atccaaaata aggtatatta ttgatgatg	37559

<210> 12

<211> 1109

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 12

ggagttccgc gttacataac ttacggtaaa tggccccgct ggctgaccgc ccaacgaccc	60
ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca	120
ttgacgtcaa tgggtggagt atttacggtg aactgcccac ttggcagtac atcaagtgtg	180
tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccc cctggcatta	240
tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat	300
cgctattacc atggtcgagg tgagccccac gttctgcttc actctcccca tctccccccc	360
ctccccaccc ccaattttgt atttatattat tttttaatta ttttgtgcag cgatgggggc	420
gggggggggg gggggcgcg gcagggcggg gcggggcggg gcgaggggcg gggcggggcg	480
aggcggagag gtgcggcggc agccaatcag agcggcgcgc tccgaaagtt tccttttatg	540
gcgagggcgc ggcggcggc gccctataaa aagcgaagcg ctccctatca gtgatagaga	600
tctccctatc agtgatagag atcgtcgacg agctcgcggc gggcgggagtg cgctgcgcgc	660
tgccctcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccggcccc gctctgactg	720
accgcgttac taaaacaggt aagtccggcc tccgcgccgg gttttggcgc ctccccgggg	780

[0135]

cgccccctc ctcacggcga gcgctgccac gtcagacgaa gggcgcagcg agcgtcctga	840
tccttccgcc cggacgctca ggacagcggc ccgctgctca taagactcgg ccttagaacc	900
ccagtatcag cagaaggaca ttttaggacg ggacttgggt gactctaggg cactggtttt	960
ctttccagag agcggaacag gcgaggaaa gtagtccctt ctcggcgatt ctgcggaggg	1020
atctccgtgg ggcggtgaac gccgatgatg cctctactaa ccatgttcat gttttctttt	1080
tttttctaca ggtcctgggt gacgaacag	1109

<210> 13

<211> 33

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 13

atacggacta gtggagaagt actcgcctac atg

33

<210> 14

<211> 36

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 14

atacgaaga tctaagactt caggaaatat gactac

36

<210> 15

<211> 46

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 15

attcagtgtgta caggcgcgcc aaagcatgac gctgttgatt tgattc

46

<210> 16

<211> 35

<212> DNA

<213> 人工序列

[0136]

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 16

actaggacta gttataagct agaatggggc ttigc

35

<210> 17

<211> 79

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 17

ttaatagaca cagtagctta atagaccag tagtgcaaag cccattcta gcttataacc

60

cctatttggtt tatttttct

79

<210> 18

<211> 93

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 18

atatatactc tctcggcact tggcctttta cactgcgaag tgttggtgct ggtgctgcgt

60

tgagagatct ttatttggtta actgttaatt gtc

93

<210> 19

<211> 23

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 19

ttaatagaca cagtagctta ata

23

<210> 20

<211> 20

<212> DNA

<213> 人工序列

[0137]

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 20

ggaaggaggt gtctagtgtt

20

<210> 21

<211> 25

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 21

caatgggcgt ggatagcggg ttgac

25

<210> 22

<211> 19

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 22

cagcatgcct gctattgtc

19

<210> 23

<211> 29

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 23

catctacgta ttagtcatcg ctattacca

29

<210> 24

<211> 21

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成引物”

<400> 24

[0138]

gacttggaaa tccccgtgag t 21

<210> 25

<211> 25

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成探针”

<400> 25

acatcaatgg gcgtggatag cggtt 25

<210> 26

<211> 592

<212> DNA

<213> 土拨鼠肝炎病毒

<400> 26

taatcaacct ctggattaca aaatttgtga aagattgact ggtattctta actatgttgc 60

tccttttacg ctatgtggat acgctgcttt aatgcctttg tatcatgcta ttgcttcccg 120

tatggctttc attttctcct ccttgtataa atcctggttg ctgtctcttt atgaggagtt 180

gtggcccggt gtcaggcaac gtggcgtggt gtgcaactgtg tttgctgacg caacccccac 240

tggttggggc attgccacca cctgtcagct cctttccggg actttcgctt tccccctccc 300

tattgccacg gcggaactca tcgccgcctg ccttgcccgc tgettgacag gggctcggct 360

gttgggcact gacaattccg tgggtgtgtc ggggaaatca tcgtcctttc cttggctgct 420

cgccgtgtgt gccacctgga ttctgcgcgg gacgtccttc tgetacgtcc cttcggccct 480

caatccagcg gaccttcctt cccgcggcct gctgccggct ctgcggcctc ttccgcgtct 540

tcgccttcgc cctcagacga gtcggatctc cctttgggcc gcctccccgc ct 592

<210> 27

<211> 543

<212> PRT

<213> 黑猩猩腺病毒

<400> 27

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

	Phe	Val	Ser	Pro	Asp	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser	
			35					40					45				
	Leu	Asn	Leu	Ala	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu	
		50					55					60					
	Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser	
	65					70					75					80	
	Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn	
					85					90					95		
	Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala	
				100					105					110			
	Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	
			115					120					125				
	Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	
		130					135					140					
[0139]	Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	
	145					150					155					160	
	Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	
					165					170					175		
	Ser	Ala	Thr	Pro	Pro	Ile	Asn	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp	
				180					185					190			
	Met	Glu	Asp	Pro	Met	Tyr	Thr	His	Asp	Gly	Lys	Leu	Gly	Ile	Arg	Ile	
			195					200					205				
	Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	
		210					215					220					
	Thr	Gly	Asn	Gly	Leu	Thr	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Arg	Val	
	225					230					235					240	
	Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala	
					245					250					255		
	Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asn	Val	

	260	265	270
	Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln 275 280 285		
	Gly Pro Leu Tyr Ile Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn 290 295 300		
	Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr 305 310 315 320		
	Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile 325 330 335		
	Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val 340 345 350		
	Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro 355 360 365		
[0140]	Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu 370 375 380		
	Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu 385 390 395 400		
	Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val 405 410 415		
	Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser 420 425 430		
	Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn 435 440 445		
	Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala 450 455 460		
	Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln 465 470 475 480		
	Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr 485 490 495		

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 28

<211> 541

<212> PRT

<213> 黑猩猩腺病毒

<400> 28

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

[0141] Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu
100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu
130 135 140

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

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

[0142]

Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu Ala Val
385 390 395 400

Ser Gly Asn Leu Ser Ser Met Thr Gly Thr Val Ser Ser Val Thr Ile
405 410 415

Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser Ser Leu
420 425 430

Asp Lys Glu Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn Ala Thr
435 440 445

Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ser Ala Tyr Pro
450 455 460

Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Glu Val Tyr
465 470 475 480

Leu His Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr Leu Asn
485 490 495

[0143]

Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr Ser Met
500 505 510

Ser Phe Thr Trp Ser Trp Asp Ser Gly Lys Tyr Ala Thr Glu Thr Phe
515 520 525

Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 29

<211> 543

<212> PRT

<213> 黑猩猩腺病毒

<400> 29

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

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

	35		40		45														
	Leu	Asn	Leu	Ala	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu			
	50						55					60							
	Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser			
	65					70					75					80			
	Gln	Asp	Ile	Thr	Ser	Thr	Thr	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn			
					85					90					95				
	Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala			
				100					105					110					
	Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu			
			115					120					125						
	Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Ala	Val	Gln	Asp	Ala	Lys	Leu	Thr			
	130						135					140							
[0144]	Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu			
	145					150					155					160			
	Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val			
					165					170					175				
	Ser	Ser	Thr	Pro	Pro	Ile	Ser	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp			
				180					185					190					
	Met	Glu	Asp	Pro	Met	Tyr	Thr	His	Asp	Gly	Lys	Leu	Gly	Ile	Arg	Ile			
			195					200					205						
	Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val			
			210				215					220							
	Thr	Gly	Asn	Gly	Leu	Thr	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Arg	Val			
	225					230					235					240			
	Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala			
				245						250					255				
	Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val			
				260					265					270					

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320

Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

[0145]

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 30

<211> 543

<212> PRT

<213> 黑猩猩腺病毒

<400> 30

Met Lys Arg Thr Lys Thr Ser Asp Lys Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

[0146]

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110

Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140

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

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175

Ser Ala Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
180 185 190

Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
195 200 205

Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
210 215 220

Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225 230 235 240

Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

[0147]

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320

Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

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

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

[0148]

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 31

<211> 543

<212> PRT

<213> 黑猩猩腺病毒

<400> 31

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110

Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140

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

[0149]

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175

Ser Ala Thr Pro Pro Ile Asn Val Ser Ser Gly Ser Leu Gly Leu Asp
180 185 190

Met Glu Asn Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
195 200 205

Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
210 215 220

Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225 230 235 240

Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320

Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

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

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

[0150]

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 32

<211> 578

<212> PRT

<213> 黑猩猩腺病毒

<400> 32

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

[0151]

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110

Leu Thr Leu Ala Ala Ala Val Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140

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

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Ile
165 170 175

Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
180 185 190

Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
195 200 205

Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
210 215 220

Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
225 230 235 240

Ser Gly Ala Leu Asn Tyr Asp Ser Ser Gly Asn Leu Glu Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Val Asp Ala Asn Gly Lys Leu Ile Leu Asp Val
260 265 270

[0152]

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300

Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320

Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335

Ile Asn Pro Gly Asp Gly Leu Glu Phe Gly Ser Gly Ser Asp Thr Asn
340 345 350

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Ile Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

[0153]

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile
515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575

Gln Glu

<210> 33

<211> 578

<212> PRT

<213> 黑猩猩腺病毒

<400> 33

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80

Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110

[0154]

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
 130 135 140

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

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
 165 170 175

Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
 180 185 190

Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
 195 200 205

Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
 210 215 220

	Thr	Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Arg	Val
	225					230					235					240
	Ser	Gly	Ala	Leu	Asn	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala
					245					250					255	
	Ala	Gly	Gly	Met	Arg	Val	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val
				260					265					270		
	Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln
			275					280					285			
	Gly	Pro	Leu	Phe	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Val	Asn	Tyr	Asn
		290					295					300				
	Arg	Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val
	305					310					315					320
	Asn	Ile	Lys	Thr	Ala	Lys	Gly	Leu	Ile	Tyr	Asp	Asp	Thr	Ala	Ile	Ala
					325					330					335	
[0155]	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Asp	Ser	Gly	Ser	Asp	Thr	Asn
				340					345					350		
	Pro	Leu	Lys	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Asp	Tyr	Asp	Ser	Ser	Arg
			355					360					365			
	Ala	Ile	Ile	Ala	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly
		370					375					380				
	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr
	385					390					395					400
	Thr	Pro	Asp	Pro	Ser	Pro	Asn	Cys	Arg	Ile	Tyr	Ser	Glu	Lys	Asp	Ala
					405					410					415	
	Lys	Phe	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Val	Leu	Ala	Ser
				420					425					430		
	Val	Ser	Val	Leu	Ser	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	Ser	Gly	Thr
			435					440					445			
	Val	Thr	Ser	Ala	Gln	Ile	Val	Leu	Arg	Phe	Asp	Glu	Asn	Gly	Val	Leu

	450	455	460	
	Leu Ser Asn Ser Ser	Leu Asp Pro Gln Tyr	Trp Asn Tyr Arg Lys Gly	465 470 475 480
	Asp Leu Thr Glu Gly Thr Ala Tyr Thr	Asn Ala Val Gly Phe Met Pro		485 490 495
	Asn Leu Thr Ala Tyr Pro Lys Thr	Gln Ser Gln Thr Ala Lys Ser Asn		500 505 510
	Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser	Lys Pro Met Thr		515 520 525
	Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr	Gly Asp Ala Thr Val		530 535 540
	Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr			545 550 555 560
[0156]	Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala			565 570 575
	Gln Glu			
	<210> 34			
	<211> 578			
	<212> PRT			
	<213> 黑猩猩腺病毒			
	<400> 34			
	Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro			1 5 10 15
	Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro			20 25 30
	Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser			35 40 45
	Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu			50 55 60

	Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser	
	65					70					75					80	
	Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn	
					85					90					95		
	Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala	
				100					105					110			
	Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	
			115					120					125				
	Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	
		130					135					140					
	Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	
	145					150					155					160	
	Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	
					165					170					175		
[0157]	Ser	Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Gly	Ile	Asp	
				180					185					190			
	Met	Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Gly	Leu	Asn	Phe	
			195					200					205				
	Gly	Ala	Pro	Leu	His	Val	Val	Asp	Ser	Leu	Asn	Ala	Leu	Thr	Val	Val	
		210					215					220					
	Thr	Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Arg	Val	
	225					230					235					240	
	Ser	Gly	Ala	Leu	Asn	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala	
					245					250					255		
	Ala	Gly	Gly	Met	Arg	Val	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	
				260					265					270			
	Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	
			275					280					285				
	Gly	Pro	Leu	Phe	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Val	Asn	Tyr	Asn	

	290		295		300														
	Arg	Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val			
	305					310					315					320			
	Asn	Ile	Lys	Thr	Ala	Lys	Gly	Leu	Ile	Tyr	Asp	Asp	Thr	Ala	Ile	Ala			
					325					330					335				
	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Asp	Ser	Gly	Ser	Asp	Thr	Asn			
				340					345					350					
	Pro	Leu	Lys	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Asp	Tyr	Asp	Ser	Ser	Arg			
			355					360					365						
	Ala	Ile	Ile	Ala	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly			
	370						375					380							
	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr			
	385					390					395					400			
[0158]	Thr	Pro	Asp	Pro	Ser	Pro	Asn	Cys	Arg	Ile	Tyr	Ser	Glu	Lys	Asp	Ala			
					405					410					415				
	Lys	Phe	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Val	Leu	Ala	Ser			
				420					425					430					
	Val	Ser	Val	Leu	Ser	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	Ser	Gly	Thr			
			435					440					445						
	Val	Thr	Ser	Ala	Gln	Ile	Val	Leu	Arg	Phe	Asp	Glu	Asn	Gly	Val	Leu			
	450					455						460							
	Leu	Ser	Asn	Ser	Ser	Leu	Asp	Pro	Gln	Tyr	Trp	Asn	Tyr	Arg	Lys	Gly			
	465					470					475					480			
	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	Phe	Met	Pro			
					485					490					495				
	Asn	Leu	Thr	Ala	Tyr	Pro	Lys	Thr	Gln	Ser	Gln	Thr	Ala	Lys	Ser	Asn			
				500					505					510					
	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Ser	Lys	Pro	Met	Thr			
			515					520					525						

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575

Gln Glu

<210> 35

<211> 577

<212> PRT

<213> 黑猩猩腺病毒

<400> 35

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro
1 5 10 15

[0159] Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu
100 105 110

Thr Val Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu

	130	135	140	
	Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Thr Leu Gln 145 150 155 160			
	Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val Ser 165 170 175			
	Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Ser Ile Asp Met 180 185 190			
	Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Ala Leu Asn Ile Gly 195 200 205			
	Ala Pro Leu His Val Val Asp Thr Leu Asn Ala Leu Thr Val Val Thr 210 215 220			
	Gly Gln Gly Leu Thr Ile Asn Gly Arg Ala Leu Gln Thr Arg Val Thr 225 230 235 240			
[0160]	Gly Ala Leu Ser Tyr Asp Thr Glu Gly Asn Ile Gln Leu Gln Ala Gly 245 250 255			
	Gly Gly Met Arg Ile Asp Asn Asn Gly Gln Leu Ile Leu Asn Val Ala 260 265 270			
	Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln Gly 275 280 285			
	Pro Leu Ile Val Asn Ser Ala His Asn Leu Asp Leu Asn Leu Asn Arg 290 295 300			
	Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val Asn 305 310 315 320			
	Ile Lys Thr Ala Lys Gly Leu Phe Tyr Asp Gly Thr Ala Ile Ala Ile 325 330 335			
	Asn Ala Gly Asp Gly Leu Gln Phe Gly Ser Gly Ser Asp Thr Asn Pro 340 345 350			
	Leu Gln Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Asn Lys Ala 355 360 365			

Ile Ile Thr Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala
370 375 380

Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr Thr
385 390 395 400

Pro Asp Pro Ser Pro Asn Cys Arg Ile Asn Ser Glu Lys Asp Ala Lys
405 410 415

Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser Val
420 425 430

Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr Val
435 440 445

Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu Leu
450 455 460

Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp
465 470 475 480

[0161]

Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn
485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile
500 505 510

Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu
515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser
530 535 540

Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile
545 550 555 560

Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln
565 570 575

Glu

<210> 36
 <211> 577
 <212> PRT
 <213> 黑猩猩腺病毒

<400> 36
 Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro
 1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80

[0162] Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

Leu Ser Leu Gln Thr Ser Ala Pro Leu Thr Val Ser Ser Gly Ser Leu
 100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu Thr
 115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Gly Leu
 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Thr Leu Gln
 145 150 155 160

Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val Ser
 165 170 175

Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Ser Ile Asp Met
 180 185 190

Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Ala Leu Asn Ile Gly
 195 200 205

Ala Pro Leu His Val Val Asp Thr Leu Asn Ala Leu Thr Val Val Thr
210 215 220

Gly Gln Gly Leu Thr Ile Asn Gly Arg Ala Leu Gln Thr Arg Val Thr
225 230 235 240

Gly Ala Leu Ser Tyr Asp Thr Glu Gly Asn Ile Gln Leu Gln Ala Gly
245 250 255

Gly Gly Met Arg Ile Asp Asn Asn Gly Gln Leu Ile Leu Asn Val Ala
260 265 270

Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln Gly
275 280 285

Pro Leu Ile Val Asn Ser Ala His Asn Leu Asp Leu Asn Leu Asn Arg
290 295 300

Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val Asn
305 310 315 320

[0163]

Ile Lys Thr Ala Lys Gly Leu Phe Tyr Asp Gly Thr Ala Ile Ala Ile
325 330 335

Asn Ala Gly Asp Gly Leu Gln Phe Gly Ser Gly Ser Asp Thr Asn Pro
340 345 350

Leu Gln Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Asn Lys Ala
355 360 365

Ile Ile Thr Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala
370 375 380

Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr Thr
385 390 395 400

Pro Asp Pro Ser Pro Asn Cys Arg Ile Asn Ser Glu Lys Asp Ala Lys
405 410 415

Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser Val
420 425 430

Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr Val
435 440 445

Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu Leu
450 455 460

Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp
465 470 475 480

Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn
485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile
500 505 510

Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu
515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser
530 535 540

[0164]

Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile
545 550 555 560

Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln
565 570 575

Glu

<210> 37

<211> 524

<212> PRT

<213> 狂犬病病毒

<400> 37

Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Ser Leu
1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
20 25 30

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val
35 40 45

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Gly Phe Ser Tyr Met Glu
50 55 60

Leu Lys Val Gly Tyr Ile Ser Ala Ile Lys Val Asn Gly Phe Thr Cys
65 70 75 80

Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
85 90 95

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val
130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Ala Asp
145 150 155 160

[0165]

Leu Asp Pro Tyr Asp Lys Ser Leu His Ser Arg Val Phe Pro Ser Gly
165 170 175

Lys Cys Ser Gly Ile Thr Val Ser Ser Thr Tyr Cys Ser Thr Asn His
180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Leu Gly Thr Ser Cys
195 200 205

Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Ser Lys
210 215 220

Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly
225 230 235 240

Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp
245 250 255

Gly Thr Trp Val Ala Met Gln Thr Ser Asp Glu Thr Lys Trp Cys Pro
260 265 270

Pro Asp Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu
275 280 285

His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp
290 295 300

Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu
305 310 315 320

Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile
325 330 335

Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg
340 345 350

Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly
355 360 365

Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu
370 375 380

[0166]

Gly Pro Asp Gly His Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu
385 390 395 400

Gln Gln His Met Glu Leu Leu Glu Ser Ser Val Ile Pro Leu Met His
405 410 415

Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asp Gly Asp Glu Ala Glu
420 425 430

Asp Phe Val Glu Val His Leu Pro Asp Val His Lys Gln Val Ser Gly
435 440 445

Val Asp Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Leu Ser Ala
450 455 460

Gly Ala Leu Ile Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys
465 470 475 480

Arg Arg Val Asn Arg Pro Glu Ser Thr Gln Arg Ser Leu Gly Gly Thr
485 490 495

[0167]

Gly Arg Lys Val Ser Val Thr Ser Gln Ser Gly Lys Val Ile Ser Ser
 500 505 510

Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Arg Leu
 515 520

<210> 38

<211> 2994

<212> DNA

<213> 狂犬病病毒

<400> 38

```

atggatgccg acaagattgt attcaaagtc aataatcagg tggctctctt gaagcctgag      60
attatcgtgg atcagtatga gtacaagtac cccgctatca aagatttgaa gaagccctgt      120
ataaccctag ggaaagcccc cgacttaaac aaagcataca agtcagtttt atcaggcatg      180
aatgccgcta aacttgatcc tgatgatgta tgttcctact tggcagcagc aatgcagttc      240
tttgagggga cttgtccgga agactggact agctatggaa tcttgattgc acgaaaagga      300
gacaagatca ccccagattc tctggtggaa ataaaacgta ctgatgtaga agggaattgg      360
gctttgacag gaggcattga attgacaagg gacccactg tctctgagca tgcattctta      420
gtcggctctt tcttgagtct gtatagggtg agcaaaatat caggacaaaa caccggtaac      480
tataaaacaa acattgcaga taggatagag cagattttcg agacagcccc ttttgtaaaa      540
atcgttgaac accacactct aatgacaact cacaaaatgt gtgctaattg gagtactata      600
ccgaatttca gatttttggc cggaacctac gacatgtttt tctcccgat tgagcatcta      660
tattcagcaa ttagggtggg cacagttggt actgcttatg aagactgttc agggctggta      720
tcgtttactg ggttcataaa gcagatcaat ctcaccgcaa gagaagcaat ttataacttc      780
ttccacaaga actttgagga agagataaga agaattgttg agccagggca ggaaaccgct      840
gttcctcact cttatttcat tcaattccgc tcattaggct tgagtgggaa gtctccttat      900
tcacgaatg ccgttggta tgtgttcaat ctattcact ttgttggatg ctatatgggt      960
caagtcagat ccctaaatgc aacagttatt gctgcatgtg ctctcatga gatgtctgtt     1020
ctagggggct atctgggaga ggaattcttc gggaaaggaa catttgaaag aagattcttt     1080
agagatgaga aagaacttca agaatacgag gcggtgaac tgacaaagac tgacgtggcg     1140
ttggcagatg atggaacggt taactctgat gatgaggact acttctccgg tgaaaccaga     1200
agtccagaag ccgtttatac tcgaatcatg atgaatggag gtcgactaaa gagatcgcat     1260
atacgaagat atgtctcagt cagttccaat catcaagccc gtccaaactc attcgccgag     1320

```

[0168]

tttctaaaca aaacatactc aagtgactca ggcagagtaa agcagacatt aaactttgat	1380
ttgctgaaac ttgcaggtga tgtagagtca aatccaggtc caatgggtcc tcaggctctt	1440
ttgtttgtgc cccttctggt tttttcattg tgtttcgga aattccctat ttacacgata	1500
ccagacaaac ttggtccctg gagcccgatt gacatccatc acctcagctg tccaaacaat	1560
ttggttgtgg aggacgaagg atgcaccaac ctgtcagggt tctctacat ggaacttaaa	1620
gtcgggtaca tctcagccat aaaagtgaac ggggttcacat gcacgggcgt tgtgacagag	1680
gcagaaacct acactaactt tgttggttat gtcaccacca catttaaaag aaagcatttc	1740
cgcccaacac cagatgcatg tagagccgag tacaattgga agatggccgg tgacccagg	1800
tatgaggagt ctctacacaa ccgtaccca gactatcatt ggcttcgaac cgtgaaaacc	1860
accaaggagt ctctcgttat catatcccca agtgtggcag atttgagccc atatgacaag	1920
tcccttact cgagggtctt ccgagcgga aagtgtcag gaataacggt gtcttctaca	1980
tactgttcga ctaaccacga ttacaccatt tggatgccag agaatccgag actagggaca	2040
tcttgtgaca tctttacca tagtagaggg aagagagcat ccaaaggag taagacctgc	2100
ggctttgtag atgaaagagg cctgtataag tctctaaaag gagcatgcaa actcaagtta	2160
tgtggagttc ttggacttag actcatggat ggaacgtggg ttgcgatgca gacatcgat	2220
gagaccaa at ggtgccctcc cgatcagctg gtgaatttgc acgacttccg ctcggatgaa	2280
attgagcatc tcgtcgtaga ggagctggtc aagaagagag aggagtgtct ggatgcacta	2340
gagtcacatc tgaccaccaa gtcggtgagt ttcagacgtc tcagtcattt aagaaagctt	2400
gtccctgggt ttggaaaagc atacactata ttttaacaaga ctttgatgga ggctgatgct	2460
cattacaa at cagtccgac ttggaatgag atcatcccct caaaggggtg tttgagagta	2520
ggaggagggt gtcacccgca tgtgaacggg gtgtttttca atggtataat attagggcct	2580
gacggccatg ttctaattcc agagatgcaa tcatccctcc tccagcaaca tatggagtta	2640
ttggagtcct cagttatacc cctgatgcac cccctggcag accttcac agttttcaag	2700
gacggtgatg aggctgagga tttgttgaa gttcaccttc ccgatgtgca caaacaggtc	2760
tccgggggtg acctgggtct cccaaactgg gggaaatacg tattactgag tgcaggggt	2820
ctgattgctc tgatgttgat aattttctc atgacctgtt gtagaagagt caatcgacca	2880
gaatctacgc aacgcagtct tggagggaca gggaggaagg tgtcagttac ttcccaaagc	2940
gggaaagtca tatcttcatg ggagtcatat aaaagtgggg gtgagactag actg	2994

<210> 39
 <211> 524
 <212> PRT
 <213> 狂犬病病毒

<400> 39
 Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Ser Leu
 1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
 20 25 30

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val
 35 40 45

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Gly Phe Ser Tyr Met Glu
 50 55 60

Leu Lys Val Gly Tyr Ile Ser Ala Ile Lys Val Asn Gly Phe Thr Cys
 65 70 75 80

[0169]

Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
 85 90 95

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
 100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
 115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val
 130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Ala Asp
 145 150 155 160

Leu Asp Pro Tyr Asp Lys Ser Leu His Ser Arg Val Phe Pro Ser Gly
 165 170 175

Lys Cys Ser Gly Ile Thr Val Ser Ser Thr Tyr Cys Ser Thr Asn His
 180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Leu Gly Thr Ser Cys
 195 200 205

Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Ser Lys
210 215 220

Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly
225 230 235 240

Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp
245 250 255

Gly Thr Trp Val Ala Met Gln Thr Ser Asp Glu Thr Lys Trp Cys Pro
260 265 270

Pro Asp Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu
275 280 285

His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp
290 295 300

Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu
305 310 315 320

[0170]

Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile
325 330 335

Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg
340 345 350

Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly
355 360 365

Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu
370 375 380

Gly Pro Asp Gly His Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu
385 390 395 400

Gln Gln His Met Glu Leu Leu Glu Ser Ser Val Ile Pro Leu Met His
405 410 415

Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asp Gly Asp Glu Ala Glu
420 425 430

[0171]

Asp Phe Val Glu Val His Leu Pro Asp Val His Lys Gln Val Ser Gly
 435 440 445

Val Asp Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Leu Ser Ala
 450 455 460

Gly Ala Leu Ile Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys
 465 470 475 480

Arg Arg Val Asn Arg Pro Glu Ser Thr Gln Arg Ser Leu Gly Gly Thr
 485 490 495

Gly Arg Lys Val Ser Val Thr Ser Gln Ser Gly Lys Val Ile Ser Ser
 500 505 510

Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Arg Leu
 515 520

<210> 40

<211> 1572

<212> DNA

<213> 狂犬病病毒

<400> 40

atggtgcctc aagccctgct gttegtgccc ctgctgggtct tctccctctg ctttggcaag	60
ttccccatct acaccatccc tgacaagctc ggcccctgggt ccccatga catacatcac	120
ctcagctgcc ccaacaacct ggtggtggag gatgagggtc gcacaaacct gagcggcttc	180
tcctatatgg aactcaaggt gggctatata tccgcatca aggtcaatgg attcacatgc	240
accggcgtcg tgacagaggc tgaaacatac accaactttg tgggtacgt caccaccaca	300
ttcaagagga agcattcag gccacccct gacgcttgca gggctgccta caattggaag	360
atggctggcg accccaggta tgaggagtcc ctgcacaatc cctacccga ctaccattgg	420
ctcaggacag tcaagaccac caaggagtcc ctggtcatta tctccctag cgtggccgac	480
ctagaccgt atgacaaaag cctgcactcc aggtcttcc ctageggcaa atgctccggc	540
attacagtga gctccaccta ctgcagcaca aaccacgact acaccatctg gatgcctgag	600
aatcctaggc tcggcacctc ctgtgacata ttacaaata gcaggggcaa gagggcttcc	660
aaaggcagca aaacctgcgg ctttgtcgac gaaagaggcc tgtacaagtc cctcaagggc	720
gcttgtaaac tcaagctgtg cggagtgtg ggactcagac tcatggacgg cacatgggtg	780

[0172]

```

gccatgcaga ccagcgatga gaccaagtgg tgcccccccg atcagctggt gaatctgcac      840
gacttcaggt ccgacgaaat tgagcacctc gtggctcgagg agctggtgaa gaagagagaa      900
gagtgcctgg atgctctgga gtccatcatg accaccaaat ccgtgtcctt cagaaggetg      960
agccacctca ggaagctggt ccccggtttt ggcaaggcct acacaatttt caataagaca     1020
ctgatggagg ccgatgetca ctacaaatcc gtgaggacct ggaacgagat catcccctcc     1080
aaaggctgcc tgagggtggg aggaagatgc cacccccacg tcaacggcgt cttcttcaac     1140
ggcattatcc tcggacccga tggccatgtc ctgatccctg aatgcaaag ctccctgctg     1200
cagcagcaca tggaactcct ggagagctcc gtcatecccc tgatgcaccc tctcgtgac     1260
cccagcaccg tgtttaaaga cggcgtatgag gccgaggact tcgtggaagt gcctctgcct     1320
gatgtgcata agcaagtcag cggcgtcgat ctgggcctgc ctaattgggg caagtatgtc     1380
ctgctctccg ccggagctct gattgccctg atgctgatca tcttctgat gacctgctgc     1440
agaagagtca acagacctga gagcacccaa agatccctcg gcggaaccgg aaggaaggtc     1500
agcgtgacca gccagtccgg caaagtgatt tcctcctggg agagctataa aagcggcgga     1560
gagaccaggc tg                                           1572

```

<210> 41

<211> 524

<212> PRT

<213> 狂犬病病毒

<400> 41

```

Met Val Pro Gln Val Leu Leu Phe Val Pro Leu Leu Gly Phe Ser Leu
1           5           10          15

```

```

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
          20          25          30

```

```

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val
          35          40          45

```

```

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Glu Phe Ser Tyr Met Glu
          50          55          60

```

```

Leu Lys Val Gly Tyr Ile Ser Ala Ile Lys Val Asn Gly Phe Thr Cys
65          70          75          80

```

```

Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
          85          90          95

```

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val
130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Thr Asp
145 150 155 160

Leu Asp Pro Tyr Asp Lys Ser Leu His Ser Arg Val Phe Pro Gly Gly
165 170 175

Asn Cys Ser Gly Ile Thr Val Ser Ser Thr Tyr Cys Ser Thr Asn His
180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Leu Arg Leu Gly Thr Ser Cys
195 200 205

[0173]

Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Gly Lys
210 215 220

Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly
225 230 235 240

Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp
245 250 255

Gly Thr Trp Val Ala Met Gln Thr Ser Asp Glu Thr Lys Trp Cys Pro
260 265 270

Pro Gly Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu
275 280 285

His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp
290 295 300

Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu
305 310 315 320

Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile
325 330 335

Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg
340 345 350

Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly
355 360 365

Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu
370 375 380

Gly Ser Asp Gly His Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu
385 390 395 400

Gln Gln His Met Glu Leu Leu Glu Ser Ser Val Ile Pro Leu Met His
405 410 415

Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asp Gly Asp Glu Val Glu
420 425 430

[0174]

Asp Phe Val Glu Val His Leu Pro Asp Val His Glu Gln Val Ser Gly
435 440 445

Val Glu Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Met Ile Ala
450 455 460

Gly Ala Leu Ile Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys
465 470 475 480

Arg Arg Val Asn Arg Pro Glu Ser Thr Gln Ser Ser Leu Gly Glu Thr
485 490 495

Gly Arg Asn Val Ser Val Thr Ser Gln Ser Gly Lys Val Ile Ser Ser
500 505 510

Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Arg Leu
515 520

<210> 42

<211> 1572

<212> DNA

<213> 狂犬病病毒

[0175]

<400> 42

atggttcctc aggttctttt gtttgtaccc ctcttggggtt tttcattgtg tttcggaag	60
ttccccattt acacgatacc agacaaactt ggtccctgga gccctattga catacaccat	120
ctcagctgtc caaataacct ggttgtggag gacgaaggat gtaccaacct gtccgagttc	180
tcttacatgg aacttaaagt gggatacatc tcagccataa aagtgaacgg gttcacttgc	240
acaggtgttg tgacagaggc agaaacctac accaactttg ttggttatgt cacaaccaca	300
ttcaagagaa agcatttccg cccaccccca gacgcatgta gagccgcgta taactggaag	360
atggccggtg accccagata tgaagagtct ctacacaatc cgtaccccgga ctaccattgg	420
cttcgaactg taaaaaccac caaagagtct ctcgttatca tatccccaag tgtgacagat	480
ttggacccat atgacaaaatc ccttcaactca agggctcttc ctggcggaaa ttgctcagga	540
ataacggtgt cctcgacctc ctgctcaact aatcatgatt acaccatttg gatgcctgag	600
aatctgagac tagggacatc ttgtgacatt tttaccaata gcagaggga gagggcatcc	660
aaaggaggca agacttgagg ctttgtggat gaaagaggcc tgtataagtc tctaaaggga	720
gcatgcaaac tcaagttatg tggagttctc ggacttagac ttatggatgg aacatgggtc	780
gcgatgcaaa catcagatga gaccaaattg tgccctccag gtcagttggt gaatttgcac	840
gactttcgct cagacgagat tgagcatctc gttgtggaag agttagtcaa gaaaagagag	900
gagtgtctgg atgcactaga gtccatcatg accaccaagt cagtgagttt cagacgtctc	960
agtcacttga gaaaacttgt ccctggggtt ggaaaagcat ataccatatt caacaaaacc	1020
ttgatggagg ctgatgctca ctacaagtct gtccggacct ggaatgagat catccctca	1080
aaagggtgtt tgagagttgg ggggaggtgt catcccatg tgaacggggt gtttttcaat	1140
ggtataatat tagggtctga cggccatgtt ctaatccag agatgcagtc atccctcctc	1200
cagcaacata tggagttgtt ggaatcttca gttatcccc tgatgcaccc cttggcagac	1260
ccttctacag ttttcaaaga cggatgatgag gttgaggatt ttgttgaagt tcacctcccc	1320
gatgtgcatg aacaggtctc aggagttgaa ctgggtctcc cgaactgggg gaagtatgta	1380
ttgatgattg caggggcctt gattgccctg atgttgataa ttttctgat gacatgttgc	1440
agaagagtca atcgaccaga atctacgcaa agcagtcttg gagagacagg gagaaatgtg	1500
tcagtcactt cccaaagcgg aaaagtcata tttcatggg agtcatataa gagtggaggc	1560
gagaccagac tg	1572

<210> 43
 <211> 524
 <212> PRT
 <213> 狂犬病病毒

<400> 43
 Met Val Pro Gln Val Leu Leu Phe Val Pro Leu Leu Gly Phe Ser Leu
 1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
 20 25 30

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val
 35 40 45

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Glu Phe Ser Tyr Met Glu
 50 55 60

Leu Lys Val Gly Tyr Ile Ser Ala Ile Lys Val Asn Gly Phe Thr Cys
 65 70 75 80

[0176] Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
 85 90 95

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
 100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
 115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val
 130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Thr Asp
 145 150 155 160

Leu Asp Pro Tyr Asp Lys Ser Leu His Ser Arg Val Phe Pro Gly Gly
 165 170 175

Asn Cys Ser Gly Ile Thr Val Ser Ser Thr Tyr Cys Ser Thr Asn His
 180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Leu Arg Leu Gly Thr Ser Cys
 195 200 205

Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Gly Lys
210 215 220

Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly
225 230 235 240

Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp
245 250 255

Gly Thr Trp Val Ala Met Gln Thr Ser Asp Glu Thr Lys Trp Cys Pro
260 265 270

Pro Gly Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu
275 280 285

His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp
290 295 300

Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu
305 310 315 320

[0177]

Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile
325 330 335

Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg
340 345 350

Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly
355 360 365

Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu
370 375 380

Gly Ser Asp Gly His Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu
385 390 395 400

Gln Gln His Met Glu Leu Leu Glu Ser Ser Val Ile Pro Leu Met His
405 410 415

Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asp Gly Asp Glu Val Glu
420 425 430

[0178]

Asp Phe Val Glu Val His Leu Pro Asp Val His Glu Gln Val Ser Gly
 435 440 445

Val Glu Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Met Ile Ala
 450 455 460

Gly Ala Leu Ile Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys
 465 470 475 480

Arg Arg Val Asn Arg Pro Glu Ser Thr Gln Ser Ser Leu Gly Glu Thr
 485 490 495

Gly Arg Asn Val Ser Val Thr Ser Gln Ser Gly Lys Val Ile Ser Ser
 500 505 510

Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Arg Leu
 515 520

<210> 44

<211> 1572

<212> DNA

<213> 狂犬病病毒

<400> 44

atggtcctc aggtgctcct gtctgctcct ctgctcggt tctccctctg ctttggcaag	60
ttccccatct acacaatccc cgataagctc ggcccttgga gccctattga catccaccac	120
ctctcctgtc ccaacaacct ggtgggtggag gacgaggat gcaccaacct gagcgagtgc	180
tcatacatgg agctgaaggt gggctatatt agcgccatca aggtcaacgg cttcacatgc	240
accggagtgc tgaccgaggc cgagacctac acaaacttcg tcggctacgt cacaacaacc	300
ttcaagagga aacatttcag acccaccct gacgcttgca gggccgctta caattggaag	360
atggctggcg accccagata cgaggagagc ctccacaacc cctaccctga ctaccactgg	420
ctcaggaccg tgaagaccac caaggagtcc ctcgtgatca tctccccag cgtcacagac	480
ctcgaccctt atgataagag cctccactcc aggggtgttc ctggcggcaa ctgttccggc	540
atcaccgtct cctccaccta ctgcagcacc aaccacgact acaccatctg gatgcctgag	600
aacctgaggc tgggcaccag ctgcgacatc ttaccaata gcaggggcaa gagggcctcc	660
aaggaggagaa agacctgcgg atttgtgat gagaggggccc tctacaagtc actgaagggc	720
gcctgcaagc tgaaactctg cggcgtgctg ggactgagc tcatggacgg aacctgggtc	780
gctatgcaaa catccgacga gaccaagtgg tgtccccccg gccagctcgt gaatttcat	840

[0179]

gacttcagga gcgacgaaat cgagcacctc gtggtggagg aactgggtcaa gaagagggag 900
 gagtgcctcg acgctctcga gtccatcatg accaccaaga gcgtgtcatt tagaagactg 960
 agccacctga ggaagctggt ccccggttc ggcaaagcct acaccatctt caacaagacc 1020
 ctgatggagg ccgatgctca ctacaagagc gtcaggacct ggaacgagat catccccagc 1080
 aaaggctgcc tgagagtggg aggaagggtg cccccccacg tgaacggcgt cttcttcaac 1140
 ggcatcattc tcggaagcga cggacacgtc ctgattcccg agatgcagag ctcaactctg 1200
 cagcagcata tggagctcct ggaaagctcc gtcattcctc tgatgcatcc cctcgtgat 1260
 ccctccacag tcttcaaaga tggcgacgag gtggaggact ttgtggaagt gcacctcccc 1320
 gatgttcatg agcaagtctc cggagtggaa ctgggcctcc ccaactgggg caagtacgtc 1380
 ctcatgattg ctggcgtct catcgccctg atgctgatca tcttctgat gacctgctgc 1440
 agaagggtea atagacccga gagcactcag tccagcctcg gcgagaccgg cagaaatgtg 1500
 agcgtgacct cccaatccgg caaagtcac agctcctggg agagctacaa atccggagga 1560
 gaaacaaggc tg 1572

<210> 45

<211> 524

<212> PRT

<213> 狂犬病病毒

<400> 45

Met Val Pro Gln Val Leu Leu Phe Val Pro Leu Leu Gly Phe Ser Leu
 1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
 20 25 30

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Val
 35 40 45

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Glu Phe Ser Tyr Met Glu
 50 55 60

Leu Lys Val Gly Tyr Ile Ser Ala Ile Lys Val Asn Gly Phe Thr Cys
 65 70 75 80

Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
 85 90 95

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr His Trp Leu Arg Thr Val
130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Thr Asp
145 150 155 160

Leu Asp Pro Tyr Asp Lys Ser Leu His Ser Arg Val Phe Pro Gly Gly
165 170 175

Asn Cys Ser Gly Ile Thr Val Ser Ser Thr Tyr Cys Ser Thr Asn His
180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Leu Arg Leu Gly Thr Ser Cys
195 200 205

[0180]

Asp Ile Phe Thr Asn Ser Arg Gly Lys Arg Ala Ser Lys Gly Gly Lys
210 215 220

Thr Cys Gly Phe Val Asp Glu Arg Gly Leu Tyr Lys Ser Leu Lys Gly
225 230 235 240

Ala Cys Lys Leu Lys Leu Cys Gly Val Leu Gly Leu Arg Leu Met Asp
245 250 255

Gly Thr Trp Val Ala Met Gln Thr Ser Asp Glu Thr Lys Trp Cys Pro
260 265 270

Pro Gly Gln Leu Val Asn Leu His Asp Phe Arg Ser Asp Glu Ile Glu
275 280 285

His Leu Val Val Glu Glu Leu Val Lys Lys Arg Glu Glu Cys Leu Asp
290 295 300

Ala Leu Glu Ser Ile Met Thr Thr Lys Ser Val Ser Phe Arg Arg Leu
305 310 315 320

Ser His Leu Arg Lys Leu Val Pro Gly Phe Gly Lys Ala Tyr Thr Ile
325 330 335

Phe Asn Lys Thr Leu Met Glu Ala Asp Ala His Tyr Lys Ser Val Arg
340 345 350

Thr Trp Asn Glu Ile Ile Pro Ser Lys Gly Cys Leu Arg Val Gly Gly
355 360 365

Arg Cys His Pro His Val Asn Gly Val Phe Phe Asn Gly Ile Ile Leu
370 375 380

Gly Ser Asp Gly His Val Leu Ile Pro Glu Met Gln Ser Ser Leu Leu
385 390 395 400

Gln Gln His Met Glu Leu Leu Glu Ser Ser Val Ile Pro Leu Met His
405 410 415

Pro Leu Ala Asp Pro Ser Thr Val Phe Lys Asp Gly Asp Glu Val Glu
420 425 430

[0181]

Asp Phe Val Glu Val His Leu Pro Asp Val His Glu Gln Val Ser Gly
435 440 445

Val Glu Leu Gly Leu Pro Asn Trp Gly Lys Tyr Val Leu Met Ile Ala
450 455 460

Gly Ala Leu Ile Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys
465 470 475 480

Arg Arg Val Asn Arg Pro Glu Ser Thr Gln Ser Ser Leu Gly Glu Thr
485 490 495

Gly Arg Asn Val Ser Val Thr Ser Gln Ser Gly Lys Val Ile Ser Ser
500 505 510

Trp Glu Ser Tyr Lys Ser Gly Gly Glu Thr Arg Leu
515 520

<210> 46

<211> 1572

<212> DNA

<213> 狂犬病病毒

[0182]

<400> 46

atggtgcccc aggtgctgct ctttgtgccc ctgctgggct tcagcctctg cttcggcaag	60
ttecccatct acaccatccc agacaagctg gggccttgga gcccacatga catccaccac	120
ctgagctgcc ccaacaacct ggtggtggaa gatgaaggct gcaccaacct gagcgagtgc	180
tcctacatgg agctgaagggt gggctacatc tctgcatca aggtgaatgg cttcacctgc	240
actggagtgg tcacagaggc cgagacctac accaactttg tgggctatgt gaccaccacc	300
ttcaagagga agcacttccg gccaccccca gatgcctgcc gggccgccta caactggaag	360
atggctgggg acccccgcta tgaggagagc ctgcacaacc cctaccaga ctaccactgg	420
ctgaggacag tgaagaccac caaggagagc ctggtgatca tcagccccag cgtgacagac	480
ctggaccctt atgacaagag cctgcacagc cgggtgttcc ctggcggcaa ctgcagcggc	540
atcacctga gcagcaccta ctgcagcacc aaccacgact acaccatctg gatgccagaa	600
aacctgcggc tgggcaccag ctgtgacatc ttcaccaaca gccggggcaa gagggccagc	660
aaggcgggca agacctgtgg ctttgtggat gagcggggcc tctacaagag cctgaagggg	720
gcctgcaagc tgaagctctg tggggtgctg ggcctgcggc tgatggatgg cacctgggtg	780
gccatgcaga cctcagatga gaccaagtgg tgccccccag gccagctggt gaacctgcat	840
gacttccgca gcgacgagat tgagcacctg gtggtggagg agctggtgaa gaagaggag	900
gagtgcctgg atgccctga gagcatcatg accaccaaga gcgtgtcctt ccgccgcctg	960
agccacctgc ggaagctggt gcctggcttt ggaaaggcct acaccatctt caacaagacc	1020
ctgatggagg cagatgcccc ctacaagagt gtgcggacct ggaatgagat catccccagc	1080
aagggtgcc tgcgggtggg cggccggtgc cccccacg tgaatggagt gttcttcaat	1140
ggcatcatcc tgggcagcga cggccacgtg ctgatecctg agatgcagag cagcctgctg	1200
cagcagcaca tggagctgct ggagagctct gtcatcccc tgatgcaccc cctcgctgac	1260
cccagcaccg tgttcaagga tggagatgaa gtggaggact tcgtggagggt gcacctgcca	1320
gatgtgcacg agcaggtgtc tggcgtggag ctgggcctgc ccaactgggg caagtacgtg	1380
ctgatgattg ctggcgccct gatcgccctg atgctgatca tcttcctgat gacctgctgc	1440
cggcgggtga acagacctga gagcaccag agcagcctgg gagagactgg aagaaatgtg	1500
tctgtacca gccagagcgg caaggtgatc agcagctggg agagctacaa gagtggcggc	1560
gagaccaggc tg	1572

<210> 47

<211> 524
 <212> PRT
 <213> 狂犬病病毒

<400> 47

Met Val Pro Gln Ala Leu Leu Phe Val Pro Leu Leu Val Phe Pro Leu
 1 5 10 15

Cys Phe Gly Lys Phe Pro Ile Tyr Thr Ile Pro Asp Lys Leu Gly Pro
 20 25 30

Trp Ser Pro Ile Asp Ile His His Leu Ser Cys Pro Asn Asn Leu Ile
 35 40 45

Val Glu Asp Glu Gly Cys Thr Asn Leu Ser Gly Phe Ser Tyr Met Glu
 50 55 60

Leu Lys Val Gly Tyr Ile Leu Ala Ile Lys Met Asn Gly Phe Thr Cys
 65 70 75 80

Thr Gly Val Val Thr Glu Ala Glu Thr Tyr Thr Asn Phe Val Gly Tyr
 85 90 95

[0183]

Val Thr Thr Thr Phe Lys Arg Lys His Phe Arg Pro Thr Pro Asp Ala
 100 105 110

Cys Arg Ala Ala Tyr Asn Trp Lys Met Ala Gly Asp Pro Arg Tyr Glu
 115 120 125

Glu Ser Leu His Asn Pro Tyr Pro Asp Tyr Arg Trp Leu Arg Thr Val
 130 135 140

Lys Thr Thr Lys Glu Ser Leu Val Ile Ile Ser Pro Ser Val Ala Asp
 145 150 155 160

Leu Asp Pro Tyr Asp Arg Ser Leu His Ser Arg Val Phe Pro Ser Gly
 165 170 175

Lys Cys Ser Gly Val Ala Val Ser Ser Thr Tyr Cys Ser Thr Asn His
 180 185 190

Asp Tyr Thr Ile Trp Met Pro Glu Asn Pro Arg Leu Gly Met Ser Cys
 195 200 205

	Asp	Ile	Phe	Thr	Asn	Ser	Arg	Gly	Lys	Arg	Val	Ser	Lys	Gly	Ser	Glu
	210						215					220				
	Thr	Cys	Gly	Phe	Val	Asp	Glu	Arg	Gly	Leu	Tyr	Lys	Ser	Leu	Lys	Gly
	225					230					235					240
	Ala	Cys	Lys	Leu	Lys	Leu	Cys	Gly	Val	Leu	Gly	Leu	Arg	Leu	Met	Asp
					245					250					255	
	Gly	Thr	Trp	Val	Ala	Met	Gln	Thr	Ser	Asn	Glu	Thr	Lys	Trp	Cys	Pro
				260					265					270		
	Pro	Asp	Gln	Leu	Val	Asn	Leu	His	Asp	Phe	Arg	Ser	Asp	Glu	Ile	Glu
			275					280					285			
	His	Leu	Val	Val	Glu	Glu	Leu	Val	Arg	Lys	Arg	Glu	Glu	Cys	Leu	Asp
	290						295					300				
	Ala	Leu	Glu	Ser	Ile	Met	Thr	Thr	Lys	Ser	Val	Ser	Phe	Arg	Arg	Leu
	305					310					315					320
[0184]	Ser	His	Leu	Arg	Lys	Leu	Val	Pro	Gly	Phe	Gly	Lys	Ala	Tyr	Thr	Ile
					325					330					335	
	Phe	Asn	Lys	Thr	Leu	Met	Glu	Ala	Asp	Ala	His	Tyr	Lys	Ser	Val	Arg
				340					345					350		
	Thr	Trp	Asn	Glu	Ile	Leu	Pro	Ser	Lys	Gly	Cys	Leu	Arg	Val	Gly	Gly
			355					360					365			
	Arg	Cys	His	Pro	His	Val	Asn	Gly	Val	Phe	Phe	Asn	Gly	Ile	Ile	Leu
	370						375					380				
	Gly	Pro	Asp	Gly	Asn	Val	Leu	Ile	Pro	Glu	Met	Gln	Ser	Ser	Leu	Leu
	385					390					395					400
	Gln	Gln	His	Met	Glu	Leu	Leu	Glu	Ser	Ser	Val	Ile	Pro	Leu	Val	His
					405					410				415		
	Pro	Leu	Ala	Asp	Pro	Ser	Thr	Val	Phe	Lys	Asp	Gly	Asp	Glu	Ala	Glu
				420					425					430		
	Asp	Phe	Val	Glu	Val	His	Leu	Pro	Asp	Val	His	Asn	Gln	Val	Ser	Gly

[0185]

435	440	445
Val Asp Leu Gly Leu Pro Asn Trp Gly Lys Tyr	Val Leu Leu Ser Ala	
450	455	460
Gly Ala Leu Thr Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys		
465	470	475 480
Arg Arg Val Asn Arg Ser Glu Pro Thr Gln Leu Asn Leu Arg Gly Thr		
485	490	495
Gly Arg Glu Val Ser Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser		
500	505	510
Trp Glu Ser His Lys Ser Gly Gly Glu Thr Arg Leu		
515	520	

<210> 48

<211> 1572

<212> DNA

<213> 狂犬病病毒

<400> 48

atggttcctc aggcctctcct gtttgtaccc cttctggttt ttccattgtg ttttgggaaa	60
ttccctatatt acacgatacc agacaagctt ggtccctgga gcccgattga catacatcac	120
ctcagctgcc caaacaattt gatagtggag gacgaaggat gcaccaacct gtcagggttc	180
tcctacatgg aacttaaagt tggatacatc ttagccataa aaatgaacgg gttcacttgc	240
acaggcggtg tgacggaggc tgaaacctac actaacttcg ttggttatgt cacaaccacg	300
tcaaaagaaa gcatttcgcg ccaacaccag atgcatgtag agccgcgtac aactggaaga	360
tggccggtga cccagatat gaagagtctc tacacaatcc gtaccctgac taccgctggc	420
ttcgaactgt aaaaaccacc aaggagtctc tcgttatcat atctccaagt gtggcagatt	480
tggaccata tgacagatcc cttcactcga gggctctccc tagcgggaag tgctcaggag	540
tagcgggtgc ttctacctac tgctccacta accacgatta caccatttgg atgcccgaga	600
atccgagact agggatgtct tgtgacattt ttaccaatag tagagggaag agagtatcca	660
aaggagtgga gacttgcggc tttgtagatg aaagaggcct atataagtct ttaaaaggag	720
catgcaaact caagttatgt ggagttctag gacttagact tatggatgga acatgggtcg	780
cgatgcaaac atcaaatgaa accaaatggt gccctcccga tcagttggtg aacctgcacg	840

[0186]

```

actttcgctc agacgaaatt gagcaccttg ttgtagagga gttggtcagg aagagagagg      900
agtgtctgga tgcactagag tccatcatat aaccaagtca gtgagtttca gacgtctcag      960
tcatttaaga aaacttgtcc ctgggttttg aaaagcatat accatattca acaagacctt     1020
gatggaagcc gatgctcact acaagtcagt cagaacttgg aatgagatcc tcccttcaaa     1080
agggtgttta agagttaggg ggaggtgtca tcctcatgtg aacggggtgt ttttcaatgg     1140
tataatatta ggacctgacg gcaatgtctt aatcccagag atgcaatcat cctcctcca     1200
gcaacatatg gagttgttgg aatcctcggg tatccccctt gtgcaccccc tggcagaccc     1260
gtctaccgtt ttcaaggacg gtgacgaggc tgaggatttt gttgaagttc accttcccga     1320
tgtgcacaat caggtctcag gagttgactt ggggtctccc aactggggga agtatgtatt     1380
actgagtgca ggggccctga ctgccttgat gttgataatt ttctgatga catgtttag      1440
aagagtcaat cgatcagaac ctacgcaact caatctcaga gggacaggga gggaggtgtc     1500
agcactcccc aaagcgggaa gatcatatct tcatgggaat cacacaagag tgggggtgag     1560
accagactgt ga                      1572

```

<210> 49

<211> 6

<212> PRT

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成肽”

<220>

<221> MISC_FEATURE

<222> (1)..(6)

<223> /注释= “该序列可以包含2-6个残基”

<400> 49

His His His His His His

1

5

<210> 50

<211> 9

<212> PRT

<213> 狂犬病病毒

<400> 50

Gly Cys Thr Asn Leu Ser Glu Phe Ser

1

5

<210> 51
<211> 9
<212> PRT
<213> 澳大利亚蝙蝠狂犬病病毒

<400> 51
Gly Cys Thr Ser Leu Ser Gly Phe Ser
1 5

<210> 52
<211> 9
<212> PRT
<213> Aravan狂犬病病毒

<400> 52
Gly Cys Thr Asn Leu Ser Gly Phe Thr
1 5

<210> 53
<211> 9
<212> PRT
<213> Bokeloh蝙蝠狂犬病病毒

<400> 53
Gly Cys Thr Thr Leu Thr Val Phe Ser
1 5

[0187]

<210> 54
<211> 9
<212> PRT
<213> 杜文海格狂犬病病毒

<400> 54
Gly Cys Thr Thr Leu Thr Pro Phe Ser
1 5

<210> 55
<211> 9
<212> PRT
<213> 欧洲蝙蝠狂犬病病毒1

<400> 55
Gly Cys Thr Thr Leu Thr Pro Phe Ser
1 5

<210> 56
<211> 9
<212> PRT
<213> 欧洲蝙蝠狂犬病病毒2

<400> 56
Gly Cys Thr Thr Leu Thr Val Phe Ser

	1	5
	<p><210> 57 <211> 9 <212> PRT <213> Irkut狂犬病病毒</p>	
	<p><400> 57 Gly Cys Thr Thr Leu Thr Ala Phe Asn 1 5</p>	
	<p><210> 58 <211> 9 <212> PRT <213> Khujand狂犬病病毒</p>	
	<p><400> 58 Gly Cys Thr Thr Leu Ser Gly Phe Thr 1 5</p>	
	<p><210> 59 <211> 9 <212> PRT <213> Lagos蝙蝠狂犬病病毒</p>	
[0188]	<p><400> 59 Gly Cys Ser Asp Thr Ala Thr Phe Ser 1 5</p>	
	<p><210> 60 <211> 9 <212> PRT <213> 莫科拉狂犬病病毒</p>	
	<p><400> 60 Gly Cys Asn Thr Glu Ser Pro Phe Thr 1 5</p>	
	<p><210> 61 <211> 9 <212> PRT <213> Shimoni蝙蝠狂犬病病毒</p>	
	<p><400> 61 Gly Cys Ser Ser Ser Ser Thr Phe Ser 1 5</p>	
	<p><210> 62 <211> 9 <212> PRT <213> 西高加索蝙蝠狂犬病病毒</p>	

<400> 62
Tyr Cys Thr Thr Glu Gln Ser Ile Thr
1 5

<210> 63
<211> 9
<212> PRT
<213> Ikoma狂犬病病毒

<400> 63
Gly Cys Asn Glu Gly Ser Lys Val Ser
1 5

<210> 64
<211> 6
<212> PRT
<213> 狂犬病病毒

<400> 64
Lys Leu Cys Gly Val Leu
1 5

[0189] <210> 65
<211> 6
<212> PRT
<213> 澳大利亚蝙蝠狂犬病病毒

<400> 65
Lys Leu Cys Gly Ile Ser
1 5

<210> 66
<211> 6
<212> PRT
<213> Aravan狂犬病病毒

<400> 66
Lys Leu Cys Gly Val Met
1 5

<210> 67
<211> 6
<212> PRT
<213> Bokeloh蝙蝠狂犬病病毒

<400> 67
Lys Leu Cys Gly Val Ser
1 5

<210> 68

<211> 6
<212> PRT
<213> 杜文海格狂犬病病毒

<400> 68
Arg Leu Cys Gly Ile Ser
1 5

<210> 69
<211> 6
<212> PRT
<213> 欧洲蝙蝠狂犬病病毒 1

<400> 69
Arg Leu Cys Gly Val Pro
1 5

<210> 70
<211> 6
<212> PRT
<213> 欧洲蝙蝠狂犬病病毒 2

<400> 70
Lys Leu Cys Gly Ile Ser
1 5

[0190]

<210> 71
<211> 6
<212> PRT
<213> Irkut狂犬病病毒

<400> 71
Lys Leu Cys Gly Met Ala
1 5

<210> 72
<211> 6
<212> PRT
<213> Khujand狂犬病病毒

<400> 72
Lys Leu Cys Gly Val Ser
1 5

<210> 73
<211> 6
<212> PRT
<213> Lagos蝙蝠狂犬病病毒

<400> 73
Thr Leu Cys Gly Lys Pro
1 5

- <210> 74
<211> 6
<212> PRT
<213> 莫科拉狂犬病病毒
- <400> 74
Thr Leu Cys Gly Lys Pro
1 5
- <210> 75
<211> 6
<212> PRT
<213> Shimoni蝙蝠狂犬病病毒
- <400> 75
Thr Leu Cys Gly Lys Pro
1 5
- <210> 76
<211> 6
<212> PRT
<213> 西高加索蝙蝠狂犬病病毒
- [0191] <400> 76
Ser Ile Cys Gly Arg Gln
1 5
- <210> 77
<211> 6
<212> PRT
<213> Ikoma狂犬病病毒
- <400> 77
Ile Ile Cys Gly Lys Ser
1 5
- <210> 78
<211> 9
<212> PRT
<213> 狂犬病病毒
- <400> 78
Lys Ser Val Arg Thr Trp Asn Glu Ile
1 5
- <210> 79
<211> 9
<212> PRT
<213> 澳大利亚蝙蝠狂犬病病毒

	<400> 79 Lys Ser Val Arg Thr Trp Asp Glu Ile 1 5
	<210> 80 <211> 9 <212> PRT <213> Aravan狂犬病病毒
	<400> 80 Lys Ser Val Arg Glu Trp Thr Glu Val 1 5
	<210> 81 <211> 9 <212> PRT <213> Bokeloh蝙蝠狂犬病病毒
	<400> 81 Lys Ser Ile Arg Gln Trp Thr Glu Ile 1 5
[0192]	<210> 82 <211> 9 <212> PRT <213> 欧洲狂犬病病毒
	<400> 82 Lys Ser Val Arg Glu Trp Lys Glu Ile 1 5
	<210> 83 <211> 9 <212> PRT <213> 欧洲蝙蝠狂犬病病毒 1
	<400> 83 Lys Ser Val Arg Glu Trp Lys Glu Val 1 5
	<210> 84 <211> 9 <212> PRT <213> 欧洲蝙蝠狂犬病病毒 2
	<400> 84 Lys Ser Ile Arg Glu Trp Thr Asp Val 1 5
	<210> 85 <211> 9

<212> PRT
<213> Irkut狂犬病病毒

<400> 85
Lys Ser Ile Arg Glu Trp Lys Glu Ile
1 5

<210> 86
<211> 9
<212> PRT
<213> Khujand狂犬病病毒

<400> 86
Lys Ser Ile Arg Glu Trp Ser Glu Ile
1 5

<210> 87
<211> 9
<212> PRT
<213> Lagos蝙蝠狂犬病病毒

<400> 87
Leu Arg Val Asp Ser Trp Asn Asp Ile
1 5

[0193]

<210> 88
<211> 9
<212> PRT
<213> 莫科拉狂犬病病毒

<400> 88
Lys Arg Val Asp Arg Trp Ala Asp Ile
1 5

<210> 89
<211> 9
<212> PRT
<213> Shimoni蝙蝠狂犬病病毒

<400> 89
Lys Arg Val Asp Arg Trp Glu Glu Ile
1 5

<210> 90
<211> 9
<212> PRT
<213> 西高加索蝙蝠狂犬病病毒

<400> 90
Ile Lys Val Glu Asn Trp Ser Glu Val
1 5

[0194]

<210> 91

<211> 9

<212> PRT

<213> Ikoma狂犬病病毒

<400> 91

Lys Ser Val Asp Asn Trp Thr Asp Ile

1

5

<210> 92

<211> 120

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 92

ggaaggtcag cgtgaccagc cagtccggca aagtgatttc ctcctgggag agctataaaa 60

gcggcggaga gaccaggctg tgatgagcgg ccgcgatctg taatcaacct ctggattaca 120

<210> 93

<211> 120

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释= “人工序列的描述：合成多核苷酸”

<400> 93

atggctccgg cggctctctgc aacacaaata aagagaccct aagaccccca acttatatat 60

tttcatgacc accccaggcc acgcccactc acccacctca ccatagagcc caccgcatcc 120

<210> 94

<211> 9

<212> PRT

<213> 狂犬病病毒

<400> 94

Gly Cys Thr Asn Leu Ser Gly Phe Ser

1

5

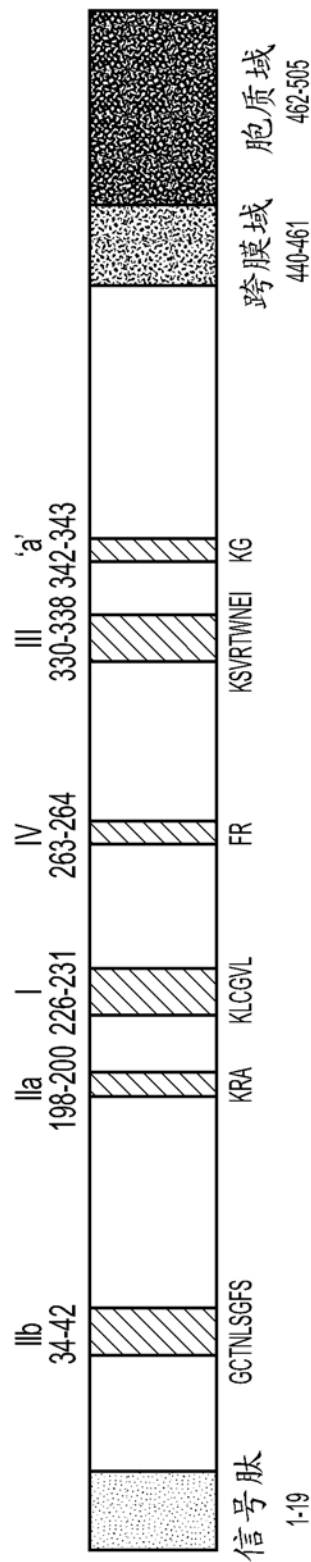


图 1

medoid G蛋白的主要抗原位点及其氨基酸位置。抗原位点I携带构象和线性表位并且位于位置226-231。抗原位点II是残基34-42 (IIB) 和198-200 (IIA) 处的不连续构象表位。抗原位点III是残基330-338处的连续构象表位。抗原位点IV位于位置263-264。次要位点 'A' 位于位置342-343。

ChAd3	(1)	MKR TK TSDE SE FNVPVYDTE SG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D IT TA SP PLK KT KNLS LE
PanAd3	(1)	MKR AK TSDE TE FNVPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS RL SEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TV TP PLK KT KNLS LQ
ChAd17	(1)	MKR TK TSDE SE FNVPVYDTE SG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D ITS TT PP PLK KT KNLS LE
ChAd19	(1)	MKR TK TS DK SE FN VPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TT TP PLK KT KNLS LE
ChAd24	(1)	MKR TK TSDE SE FNVPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TT TP PLK KT KNLS LE
ChAd155	(1)	MKR TK TSDE SE FNVPVYDTE SG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D IT TA SP PLK KT KNLS LE
ChAd11	(1)	MKR TK TSDE SE FNVPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TT TP PLK KT KNLS LE
ChAd20	(1)	MKR TK TSDE SE FNVPVYDTE SG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D IT TA SP PLK KT KNLS LE
ChAd31	(1)	MKR TK TSDE SE FNVPVYDTE SG PPSVPELT PP FVSPDGFQ ES PPGVLS LN AEPLVTS HG MLAK MG SLDDAGNLT SQ D IT TA SP PLK KT KNLS LE
PanAd1	(1)	MKR AK TSDE TE FNVPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS RL SEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TV TP PLK KT KNLS LQ
PanAd2	(1)	MKR AK TSDE TE FNVPVYDTE NG PPSVPELT PP FVSPDGFQ ES PPGVLS RL SEPLVTS HG MLAK MG SLDDAGNLT SQ D VT TV TP PLK KT KNLS LQ
ChAd3	(101)	TSSPLTV ST SCAL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PI N VS SG SL G LD ME DP MY TH
PanAd3	(101)	TS AP LT VS -SG SL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL GL AT Q GP LT V SE GKL TL Q TS AP LT AA DS ST LT VG TT PP IS VS SG SL G LD ME DP MY TH
ChAd17	(101)	TSSPLTV ST SCAL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT VS ST PP IS VS SG SL G LD ME DP MY TH
ChAd19	(101)	TS AP LT VS TS CA LT IA AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PI IS VS SG SL G LD ME DP MY TH
ChAd24	(101)	TS AP LT VS TS CA LT IA AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PI N VS SG SL G LD ME DP MY TH
ChAd155	(101)	TSSPLTV ST SCAL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PL ST SN GS LS G ID M Q AP I Y TT
ChAd11	(101)	TS AP LT VS TS CA LT IA AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT IS AT PP PL ST SN GS LS G ID M Q AP I Y TT
ChAd20	(101)	TSSPLTV ST SCAL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PL ST SN GS LS G ID M Q AP I Y TT
ChAd31	(101)	TSSPLTV ST SCAL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL TL AT KG PLTV SE GKL AL Q TS AP LT AA DS ST LT TV SA TP PL ST SN GS LS G ID M Q AP I Y TT
PanAd1	(101)	TS AP LT VS -SG SL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL GL AT Q GP LT V SE GKL TL Q TS AP LT AA DS ST LT VS AT PP PL ST SN GS LS ID M Q AP I Y TT
PanAd2	(101)	TS AP LT VS -SG SL TV AA AP LA VG TS LT MQ SE AP LT VQ DA KL GL AT Q GP LT V SE GKL TL Q TS AP LT AA DS ST LT VS AT PP PL ST SN GS LS ID M Q AP I Y TT

图 2A

ChAd3	(201)	DKLGIRIGGP ^{LR} VVDS ^{LHT} TVVTGNG ^{LTV} DNNA ^{LQ} TRVTGALGYDTSNG ^{LQ} RAAGMRIDANG ^Q LILNWAYPFDQAQNNLSRLGGP ^{LY} INT ^D HNLD	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{TV} GN ^{TG} DDK ^{LT} LWT
PanAd3	(200)	DKLGIRIGGP ^{LQ} VVDS ^{LHT} TVVTGNG ^{LTV} ANNA ^{LQ} TKVAGALGYDSSGN ^{LE} RAAGMRINTG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LY} VNT ^I HNLD	-----KIDSGLDYN ^{AN} GAVI ^{IA} KL ^{GT} GLF ^{ED} NTGAL ^{IT} GN ^{TG} DDK ^{LT} LWT
ChAd17	(201)	DKLGIRIGGP ^{LR} VVDS ^{LHT} TVVTGNG ^{LTV} DNNA ^{LQ} TRVTGALGYDTSNG ^{LQ} RAAGMRIDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LY} VNT ^I HNLD	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{TV} GN ^{TG} DDK ^{LT} LWT
ChAd19	(201)	DKLGIRIGGP ^{LR} VVDS ^{LHT} TVVTGNG ^{LTV} DNNA ^{LQ} TRVTGALGYDTSNG ^{LQ} RAAGMRIDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LY} VNT ^I HNLD	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} SVF ^{ED} STGALS ^{VGN} TGDDK ^{LT} LWT
ChAd24	(201)	DKLGIRIGGP ^{LR} VVDS ^{LHT} TVVTGNG ^{LTV} DNNA ^{LQ} TRVTGALGYDTSNG ^{LQ} RAAGMRIDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LY} VNT ^I HNLD	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} SVF ^{ED} STGALS ^{VGN} TGDDK ^{LT} LWT
ChAd155	(201)	NGKLG ^{LN} FGAP ^{LH} VVDS ^{INA} LT ^{VVT} GQGL ^{TNG} TALQ ^{TR} VSGAL ^{NY} DTSGN ^{LE} RAAGMRVDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LE} VNSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd11	(201)	NGKLG ^{LN} FGAP ^{LH} VVDS ^{INA} LT ^{VVT} GQGL ^{TNG} TALQ ^{TR} VSGAL ^{NY} DTSGN ^{LE} RAAGMRVDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LE} VNSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd20	(201)	NGKLG ^{LN} FGAP ^{LH} VVDS ^{INA} LT ^{VVT} GQGL ^{TNG} TALQ ^{TR} VSGAL ^{NY} DTSGN ^{LE} RAAGMRVDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LE} VNSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd31	(201)	NGKLG ^{LN} FGAP ^{LH} VVDS ^{INA} LT ^{VVT} GQGL ^{TNG} TALQ ^{TR} VSGAL ^{NY} DTSGN ^{LE} RAAGMRVDANG ^Q LILDVAYPFDQAQNNLSRLGGP ^{LE} VNSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
PanAd1	(200)	NGKLA ^{LN} IGAP ^{LH} VVD ^{TINA} LT ^{VVT} GQGL ^{TNG} RALQ ^{TR} VTGALSYDTG ^{NI} Q ^{LQ} AGGGMRIDNNG ^Q LILNWAYPFDQAQNNLSRLGGP ^{LIV} NSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
PanAd2	(200)	NGKLA ^{LN} IGAP ^{LH} VVD ^{TINA} LT ^{VVT} GQGL ^{TNG} RALQ ^{TR} VTGALSYDTG ^{NI} Q ^{LQ} AGGGMRIDNNG ^Q LILNWAYPFDQAQNNLSRLGGP ^{LIV} NSA ^{HN} LD	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd3	(301)	LN ^{CNR} GL ^{TTTT} INT ^{NT} KKLE ^T -----	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{TV} GN ^{TG} DDK ^{LT} LWT
PanAd3	(300)	LN ^{CNR} GL ^{TTTT} ISS ^{NT} TKLE ^T -----	-----KIDSGLDYN ^{AN} GAVI ^{IA} KL ^{GT} GLF ^{ED} NTGAL ^{IT} GN ^{TG} DDK ^{LT} LWT
ChAd17	(301)	LN ^{CNR} GL ^{TTTT} INT ^{NT} KKLE ^T -----	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{TV} GN ^{TG} DDK ^{LT} LWT
ChAd19	(301)	LN ^{CNR} GL ^{TTTT} INT ^{NT} KKLE ^T -----	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} SVF ^{ED} STGALS ^{VGN} TGDDK ^{LT} LWT
ChAd24	(301)	LN ^{CNR} GL ^{TTTT} INT ^{NT} KKLE ^T -----	-----KISSGLDYD ^{TN} GAVI ^{IK} LG ^{TG} SVF ^{ED} STGALS ^{VGN} TGDDK ^{LT} LWT
ChAd155	(301)	VN ^{YNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{KT} KLGL ^{GLD} YDS ^{SRA} IIAKLG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd11	(301)	VN ^{YNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{KT} KLGL ^{GLD} YDS ^{SRA} IIAKLG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd20	(301)	VN ^{YNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{KT} KLGL ^{GLD} YDS ^{SRA} IIAKLG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
ChAd31	(301)	VN ^{YNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{KT} KLGL ^{GLD} YDS ^{SRA} IIAKLG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
PanAd1	(300)	LN ^{ZNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{QT} KLGL ^{GLD} YDS ^{NKA} II ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT
PanAd2	(300)	LN ^{ZNR} GLY ^{LFT} SGNT ^{KK} LEVNI ^{KT} AKGL ^{LYD} DTAIAINAGDGLQ ^{FD} SGSD ^{TN} PL ^{QT} KLGL ^{GLD} YDS ^{NKA} II ^{IK} LG ^{TG} LSF ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT	-----KISSGLDYD ^{TN} GAVI ^{IA} KL ^{GT} GLS ^{ED} NTGAL ^{IT} GN ^{TG} KNDDK ^{LT} LWT

图 2B

ChAd3	(365)	TPDPSPNCRIHS ^{DKD} CKFTLLVLT ^{TKGSQ} LLASV ^{AA} LA ^{AV} SGNLASTIGTVASVT ^{FL} RF RD QNGVIM ^{EN} SSLD ^{RQ} YWN ^{FR} NCNS ^{TNA} APY ^{TNA} VGFM ^{PN} LA ^{AA}
PanAd3	(364)	TPDPSPNCRIHAD ^{KD} CKFTLLVLT ^{TKGSQ} LLASV ^{AA} LA ^{AV} SGNLSSITGTVSSVT ^{FL} RF RD QNGVIM ^{EN} SSLD ^{KE} YWN ^{FR} NCNS ^{TNA} TPY ^{TNA} VGFM ^{PN} LSA
ChAd17	(365)	TPDPSPNCRIHS ^{DKD} CKFTLLVLT ^{TKGSQ} LLASV ^{AA} LA ^{AV} SGNLASTIGTVASVT ^{FL} RF RD QNGVIM ^{EN} SSLD ^{KQ} YWN ^{FR} NCNS ^{TNA} APY ^{TNA} VGFM ^{PN} LA ^{AA}
ChAd19	(365)	TPDPSPNCRIHS ^{DKD} CKFTLLVLT ^{TKGSQ} LLASV ^{AA} LA ^{AV} SGNLASTIGTVSSVT ^{FL} RF RD QNGVIM ^{EN} SSLD ^{KQ} YWN ^{FR} NCNS ^{TNA} TPY ^{TNA} VGFM ^{PN} LA ^{AA}
ChAd24	(365)	TPDPSPNCRIHS ^{DKD} CKFTLLVLT ^{TKGSQ} LLASV ^{AA} LA ^{AV} SGNLASTIGTVSSVT ^{FL} RF RD QNGVIM ^{EN} SSLD ^{KQ} YWN ^{FR} NCNS ^{TNA} TPY ^{TNA} VGFM ^{PN} LA ^{AA}
ChAd155	(401)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
ChAd11	(401)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
ChAd20	(401)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
ChAd31	(401)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
PanAd1	(400)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
PanAd2	(400)	TPDPSPNCRIYSEK ^{DA} KFTLLVLT ^{TKGSQ} VLASV ^{SV} LSVKGSLAPISGTVTSAQ ^{IV} LR ^{FD} DENGVL ^{LSN} SSLD ^{PQ} YWN ^{YR} KGD ^{LE} GTAY ^{TNA} VGFM ^{PN} LA ^{TA}
ChAd3	(465)	YPKTQSQTAKNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} SET ^{SQ} VS ^{HY} SMST ^{WAW} ESGQY ^{ATE} TF ^{AT} NSFT ^{FSY} IAE ^Q
PanAd3	(464)	YPKTQSQTAKNIVSEVYL ^{HGDK} SKPMT ^{LT} ITILNGTNE ^{SS} SET ^{SQ} VS ^{HY} SMST ^{WAW} ESGQY ^{ATE} TF ^{AT} NSFT ^{FSY} IAE ^Q
ChAd17	(465)	YPKTQSQTAKNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} SET ^{SQ} VS ^{HY} SMST ^{WAW} ESGQY ^{ATE} TF ^{AT} NSFT ^{FSY} IAE ^Q
ChAd19	(465)	YPKTQSQTAKNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} SET ^{SQ} VS ^{HY} SMST ^{WAW} ESGQY ^{ATE} TF ^{AT} NSFT ^{FSY} IAE ^Q
ChAd24	(465)	YPKTQSQTAKNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} SET ^{SQ} VS ^{HY} SMST ^{WAW} ESGQY ^{ATE} TF ^{AT} NSFT ^{FSY} IAE ^Q
ChAd155	(501)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}
ChAd11	(501)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}
ChAd20	(501)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}
ChAd31	(501)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}
PanAd1	(500)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}
PanAd2	(500)	YPKTQSQTAKSNIVSQVYL ^{NGDK} SKPMT ^{LT} ITILNGTNE ^{SS} TG-DAT ^{VS} YSMSF ^{SNW} NGSN ^{YIN} ET ^{FF} QTNSFT ^{FSY} IAQ ^{EF}

图 2C

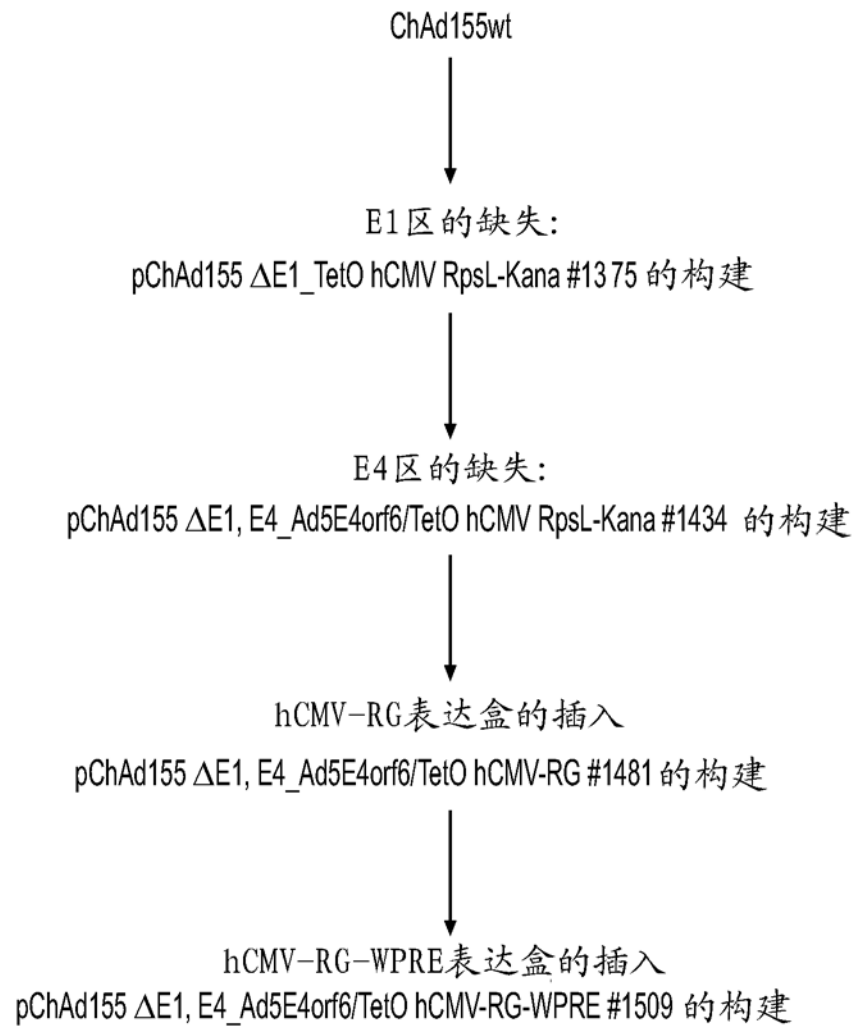


图 3

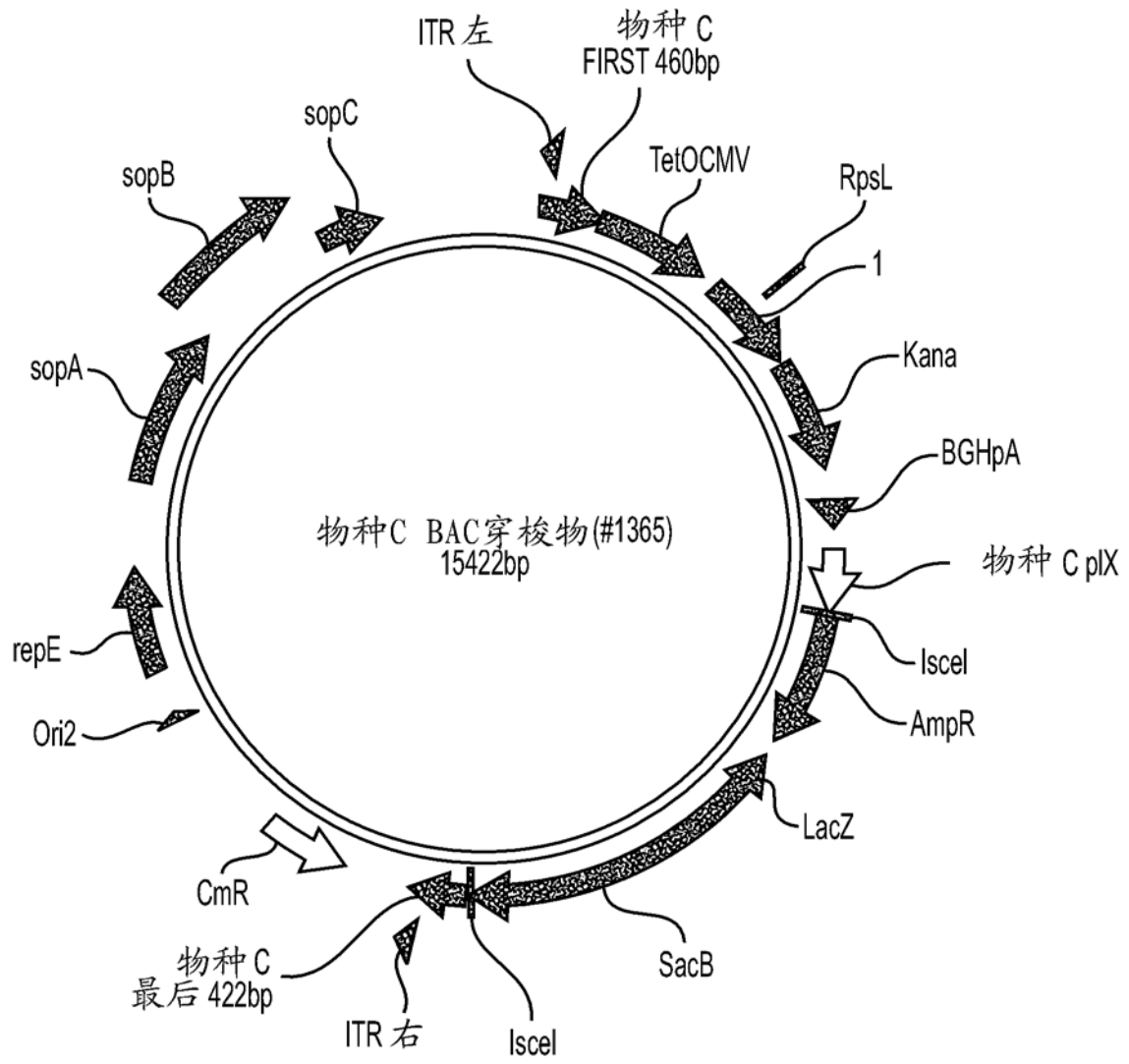


图 4

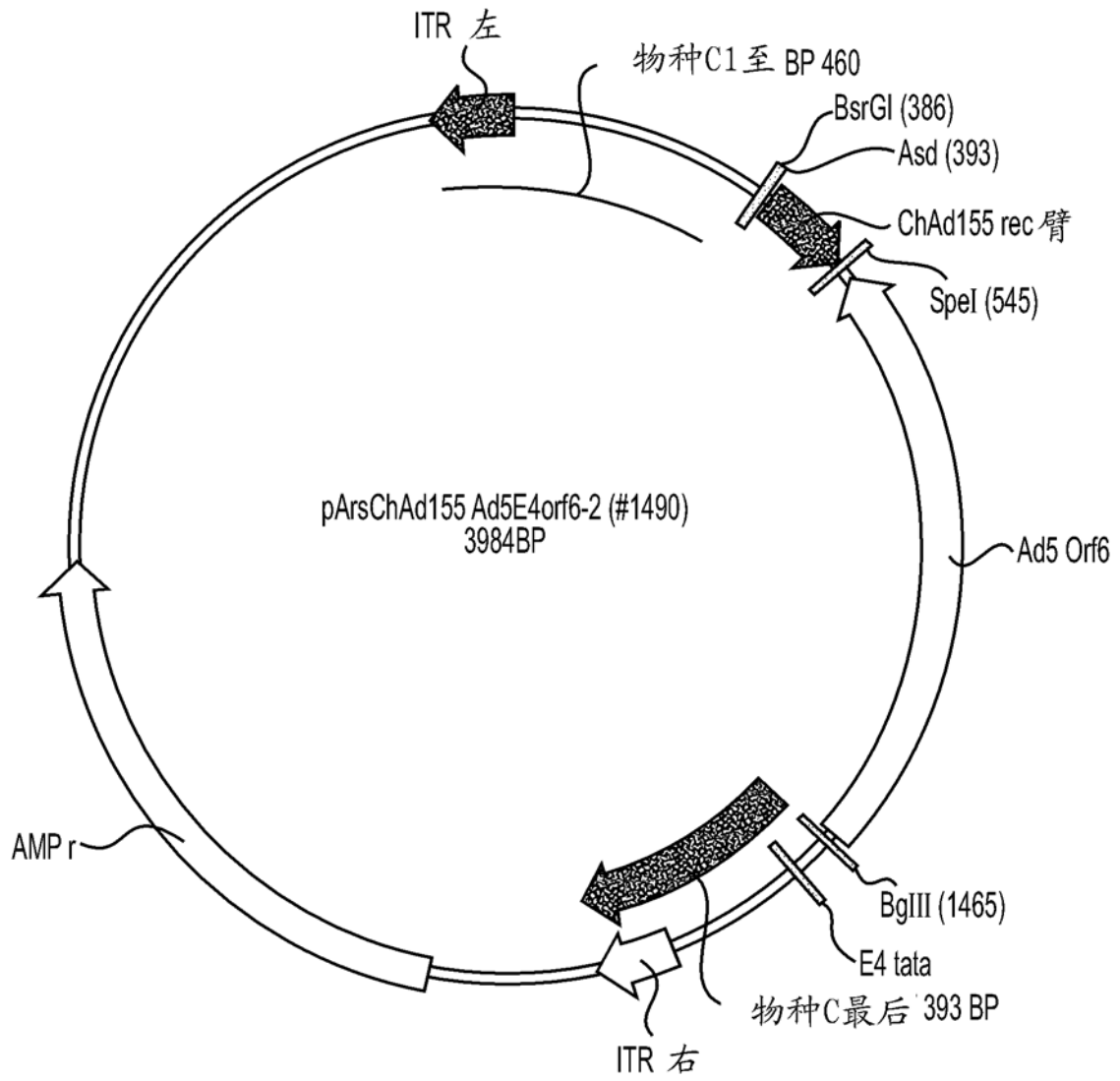


图 5

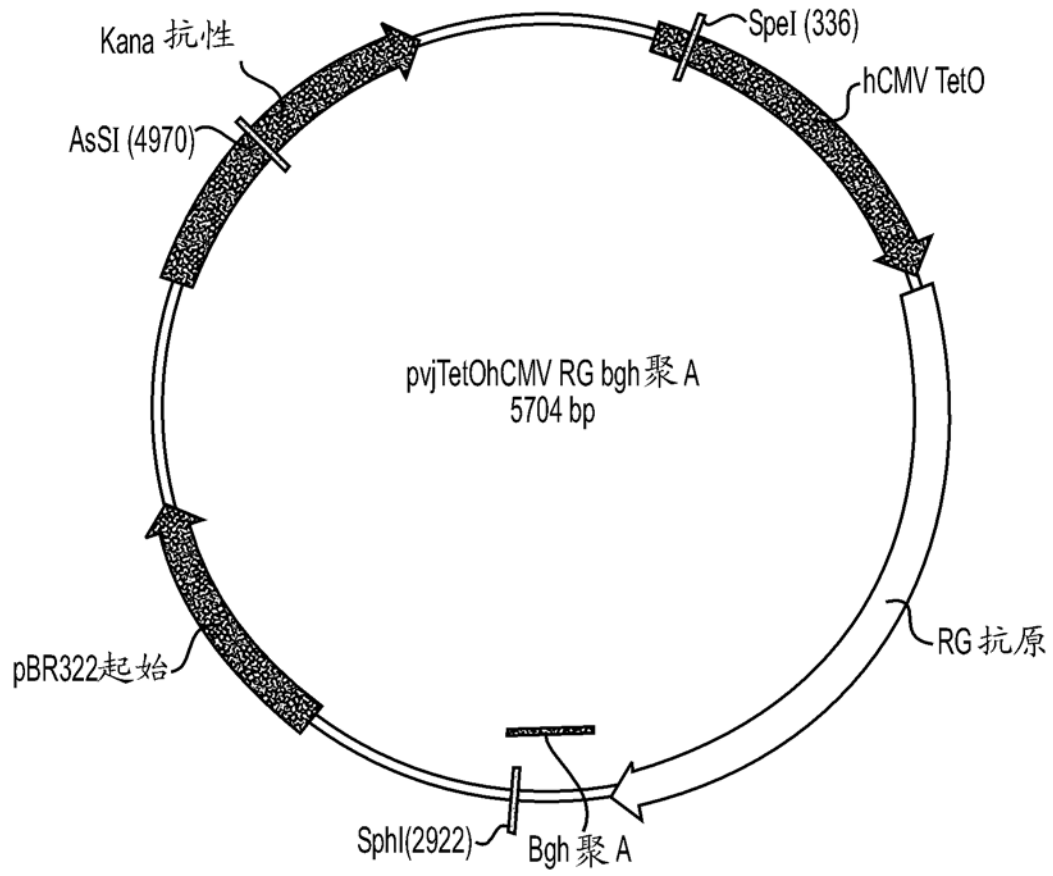


图 6

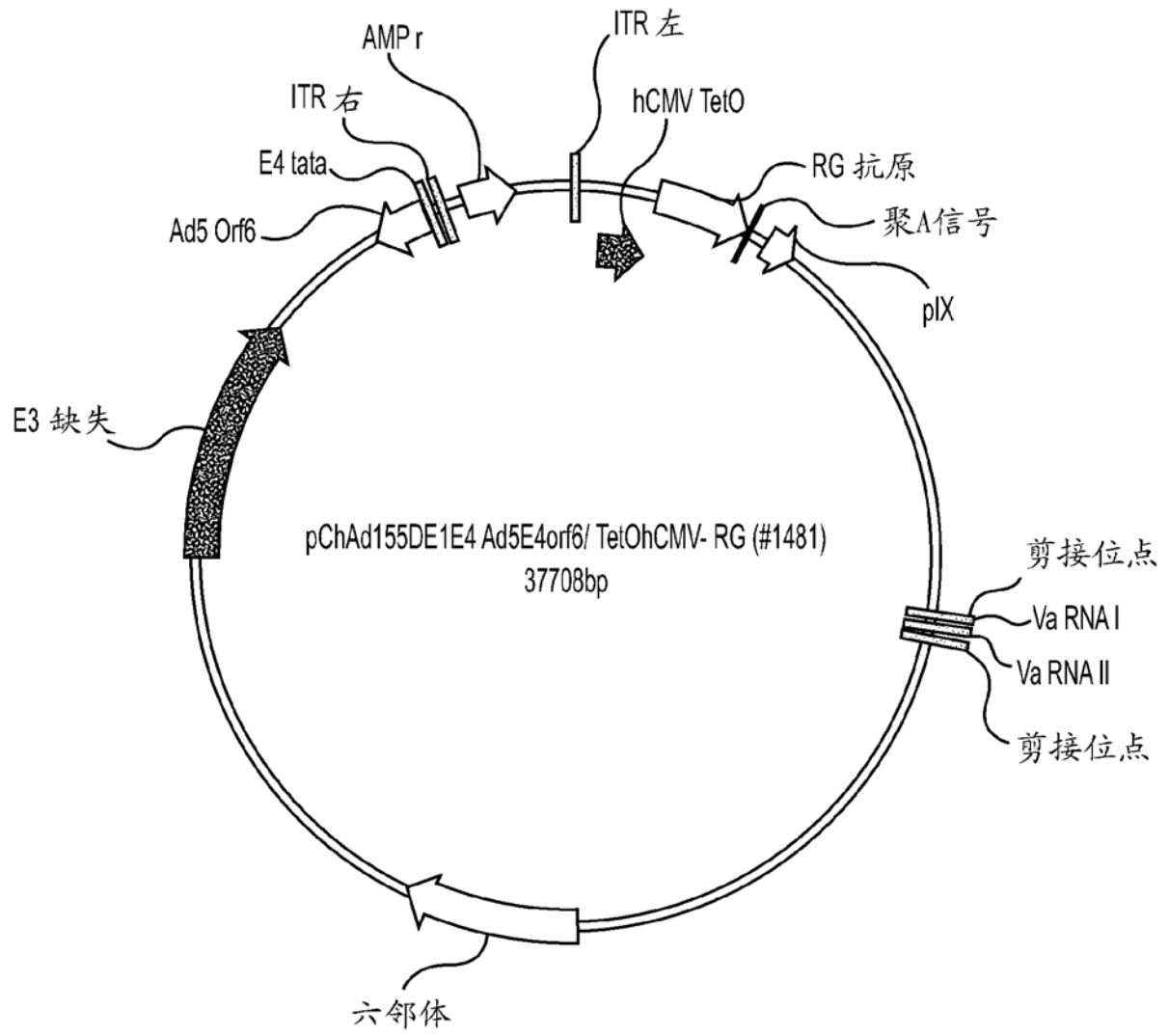


图 7

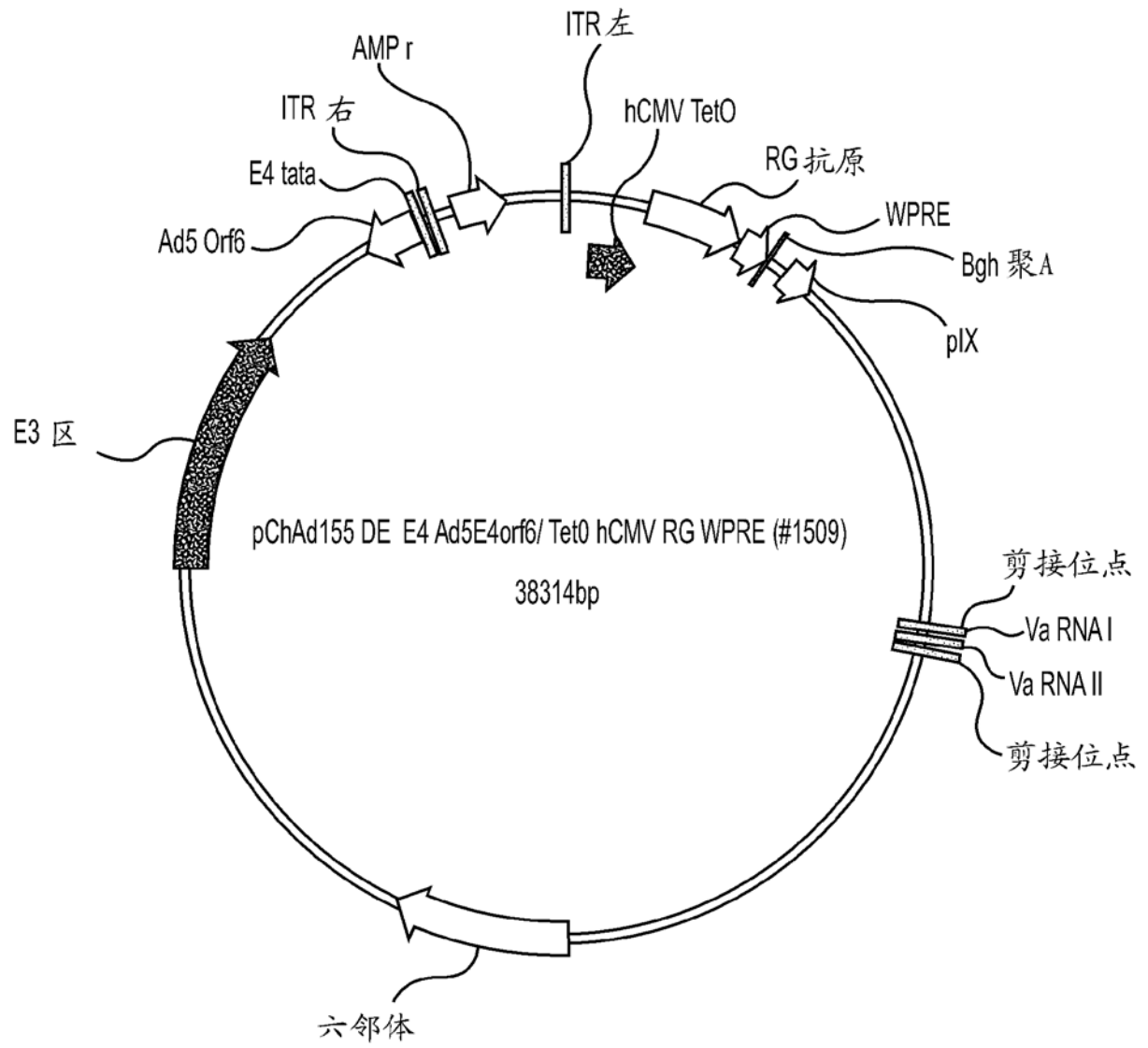


图 8

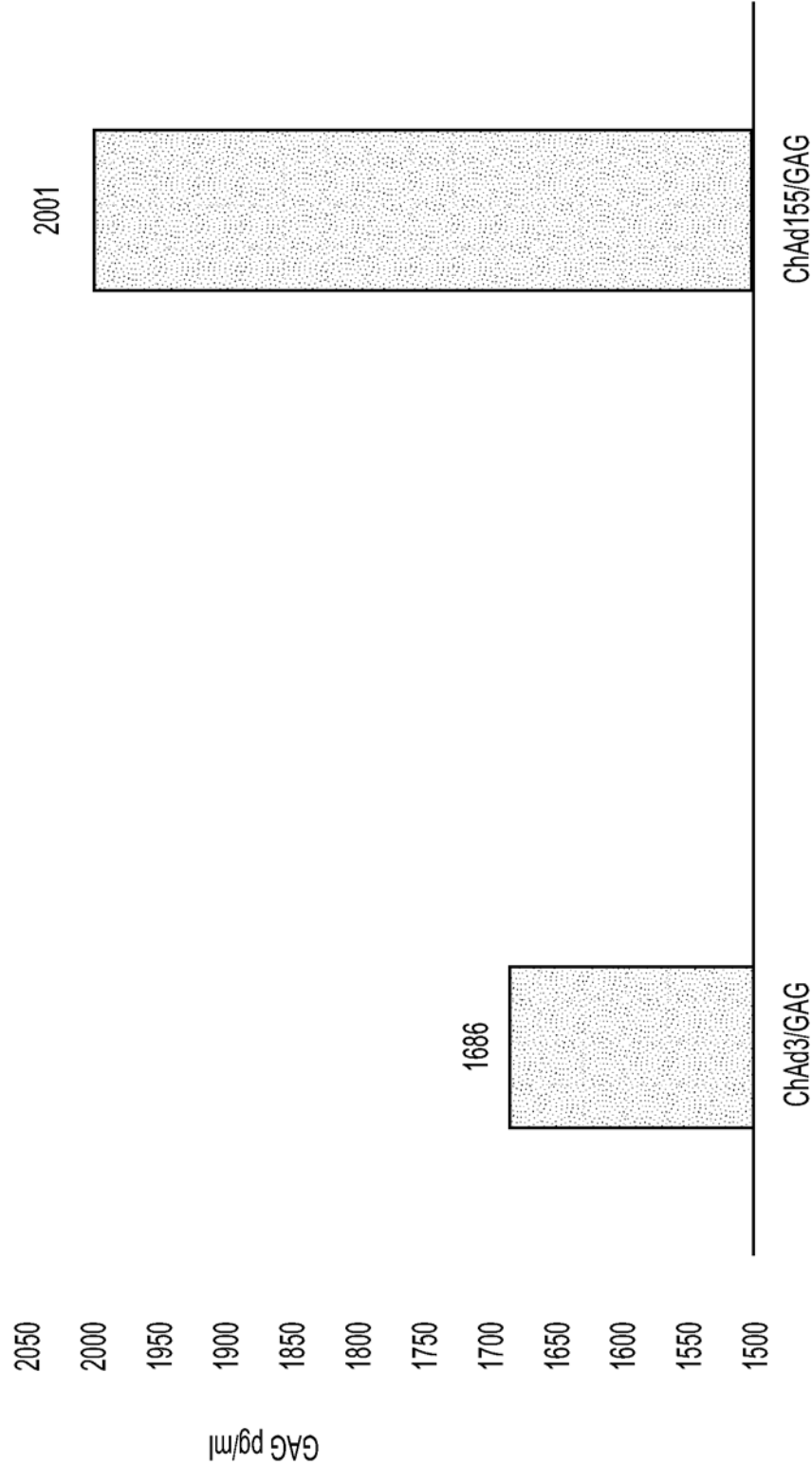


图 9

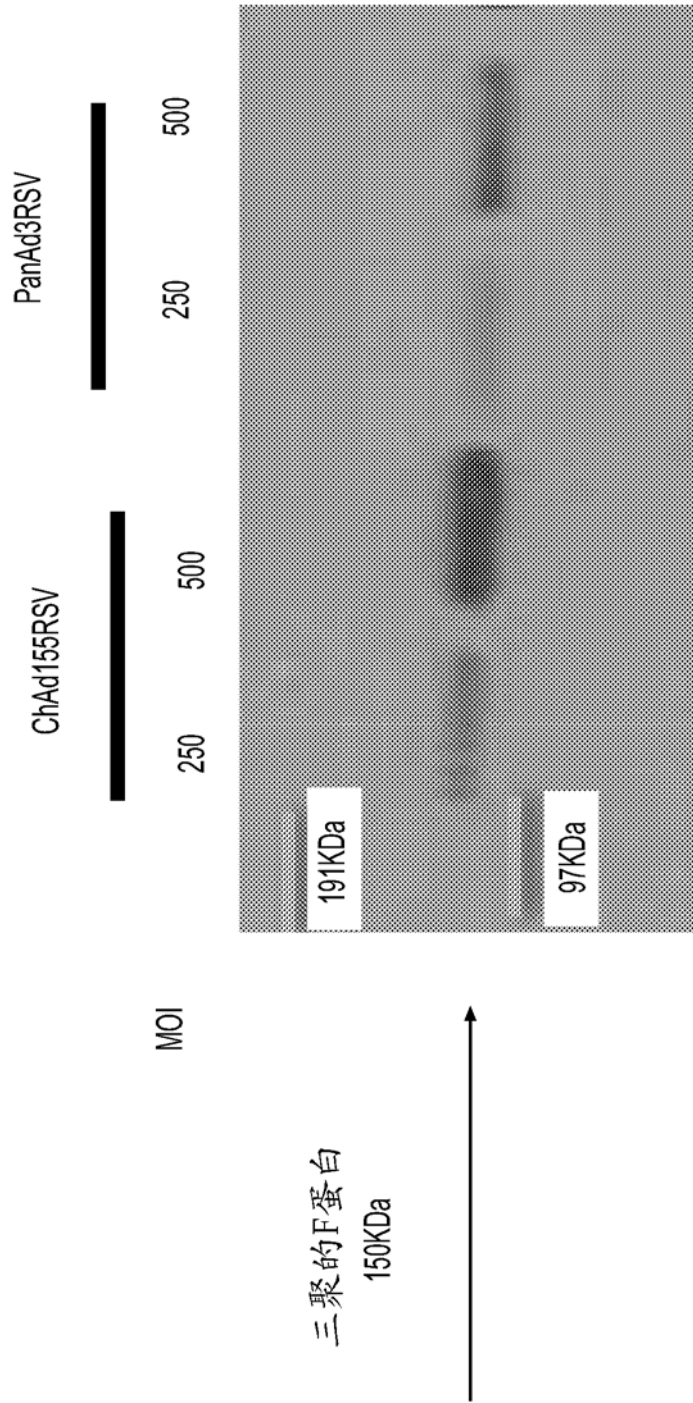


图 10

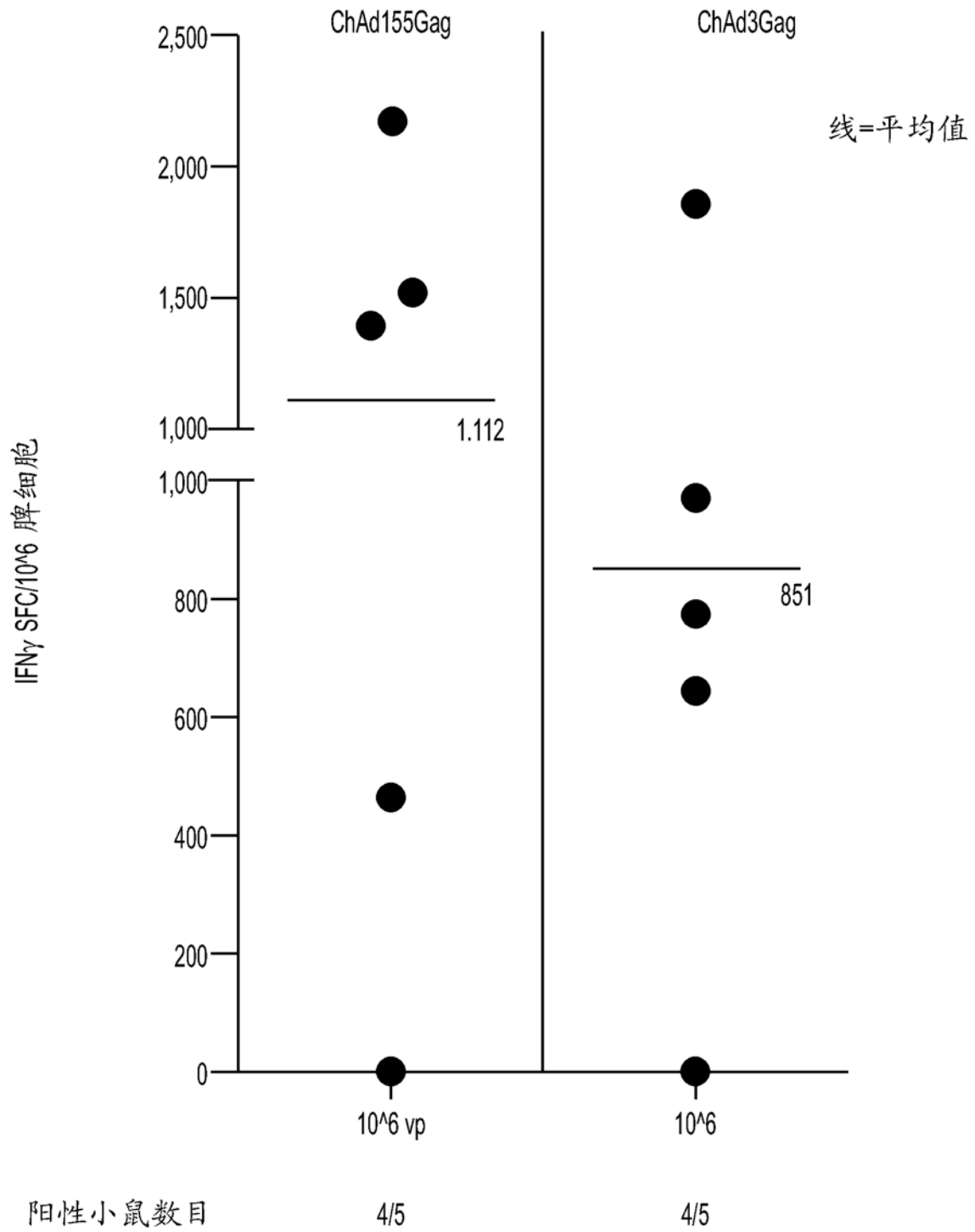


图 11

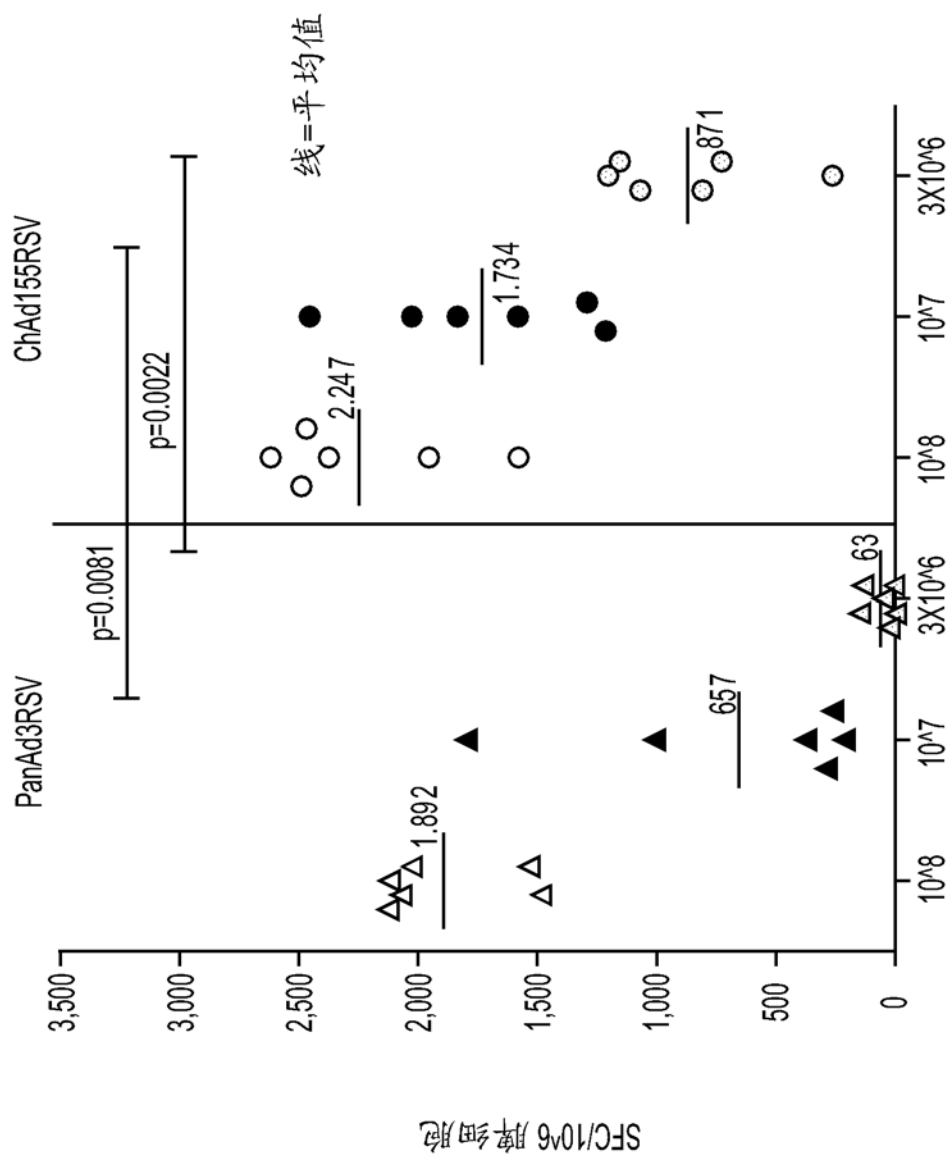
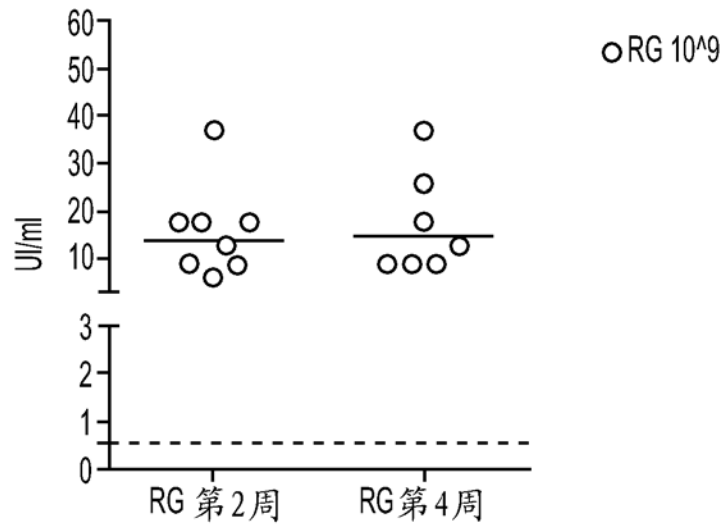


图 12

CD1小鼠中单次IM施用 10^9 VP的ChAd155-RG

功能性抗体



抗狂犬病FAVN: 荧光抗体病毒中和

图 13A

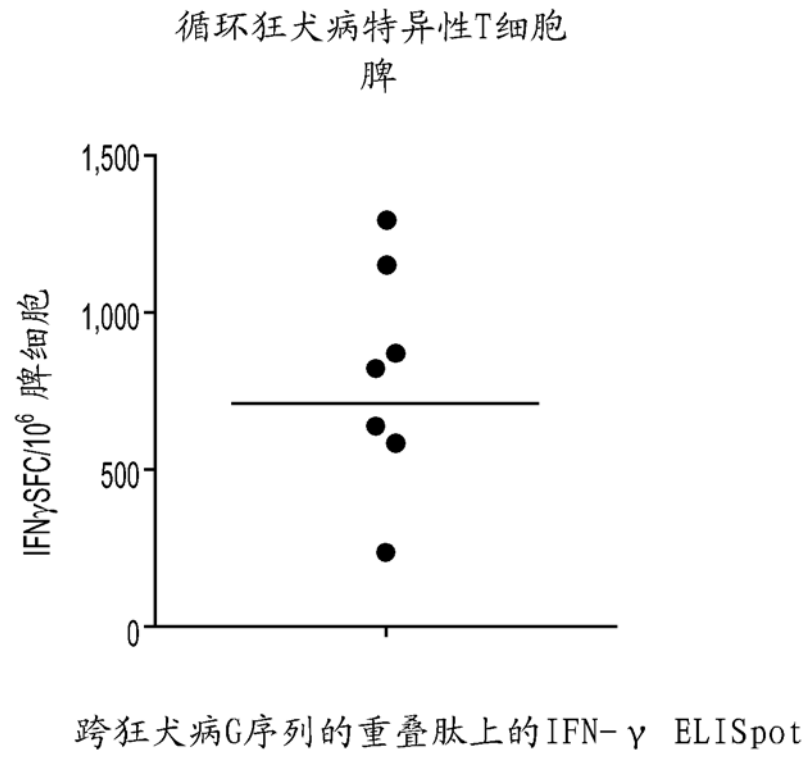


图 13B

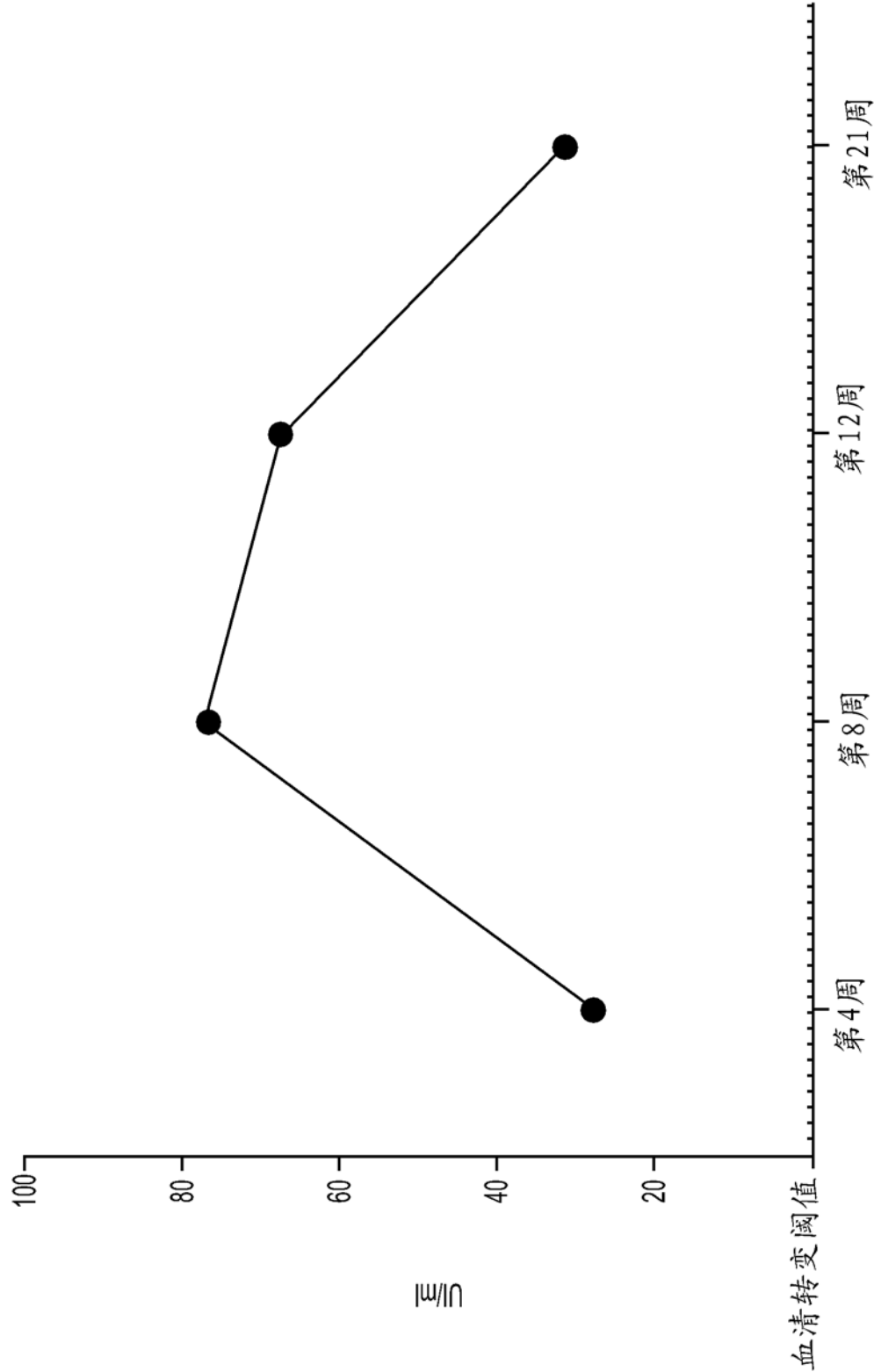


图 14

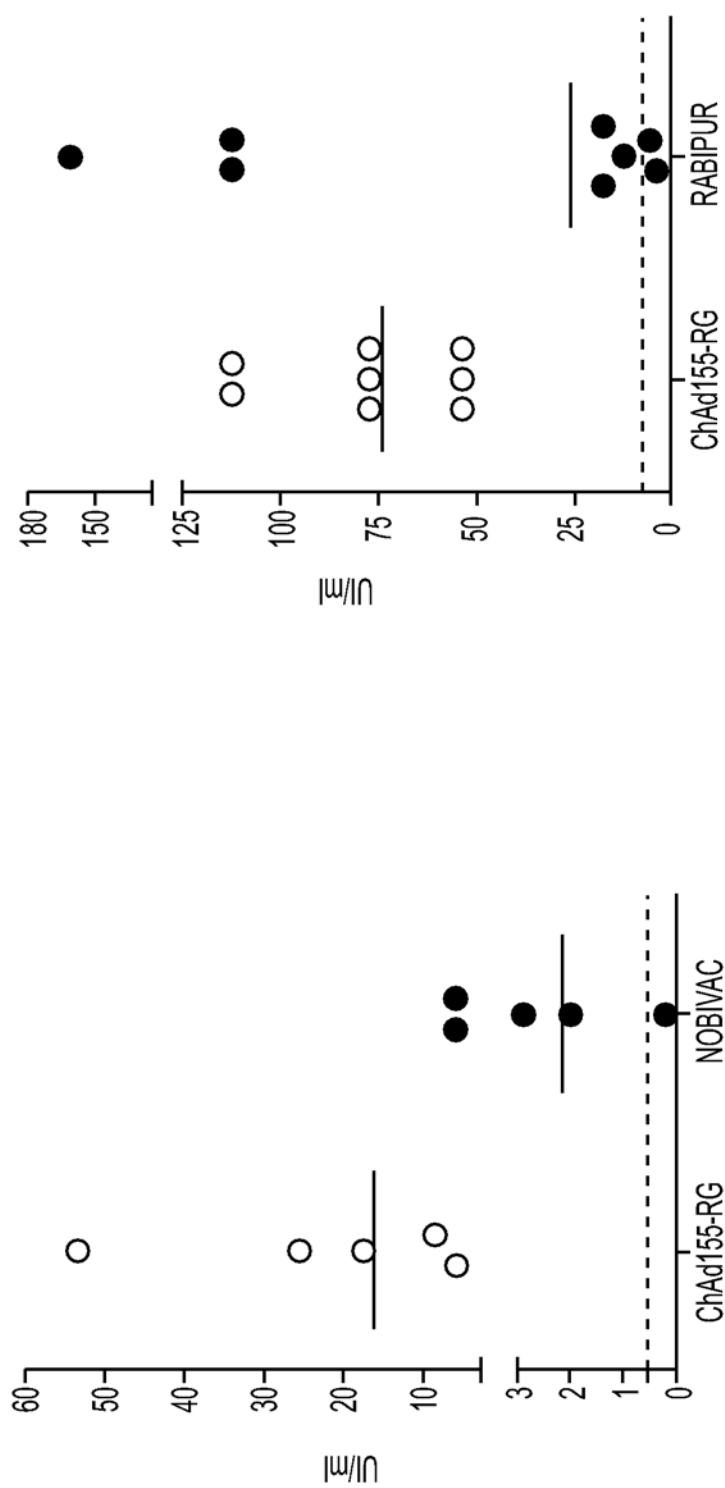


图 15

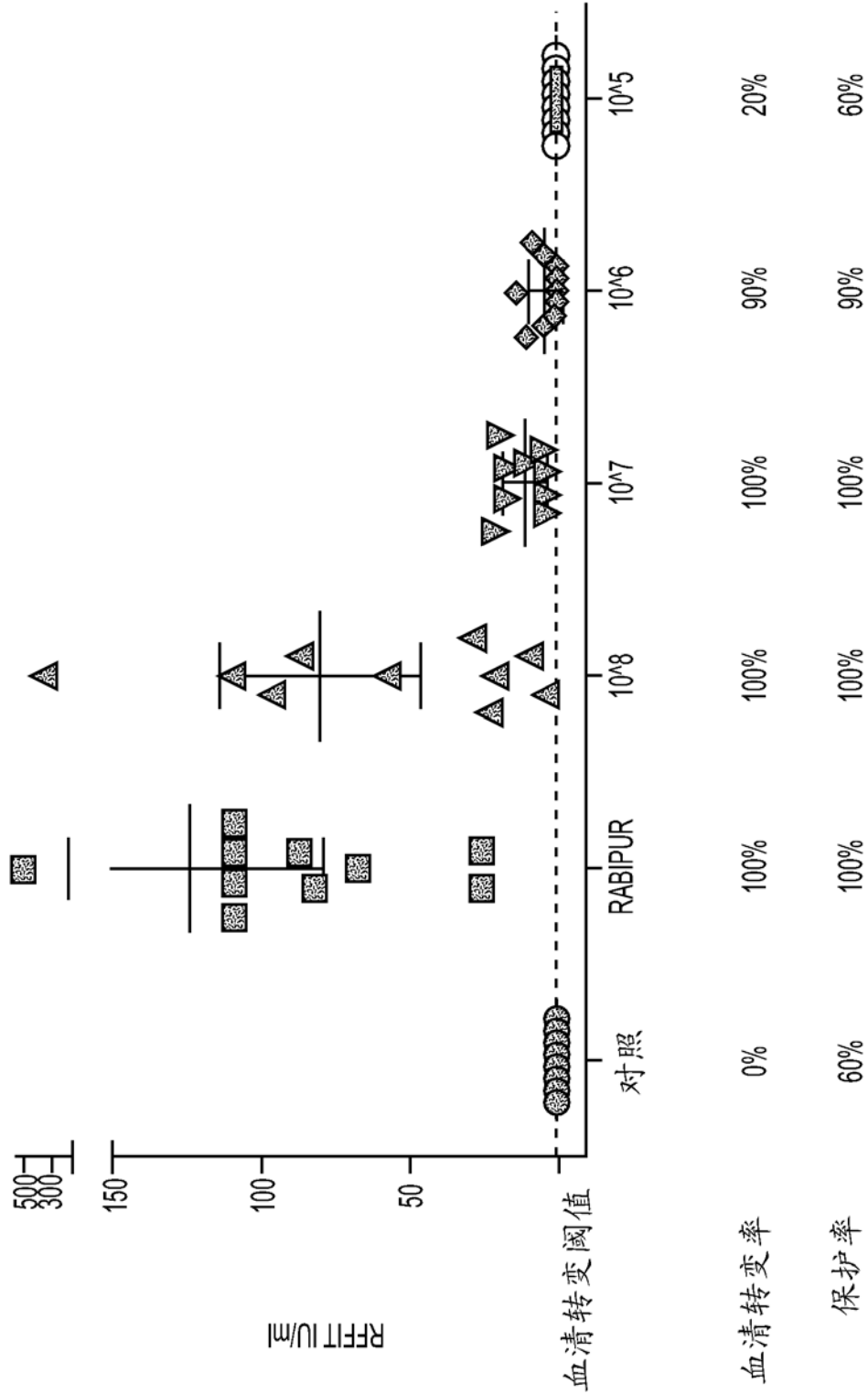


图 16

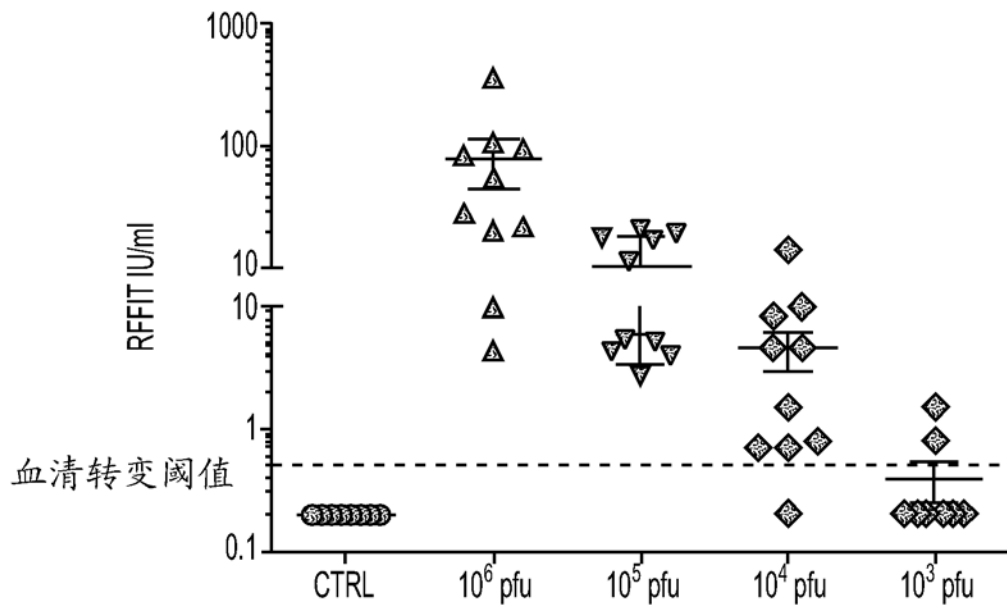


图 17A

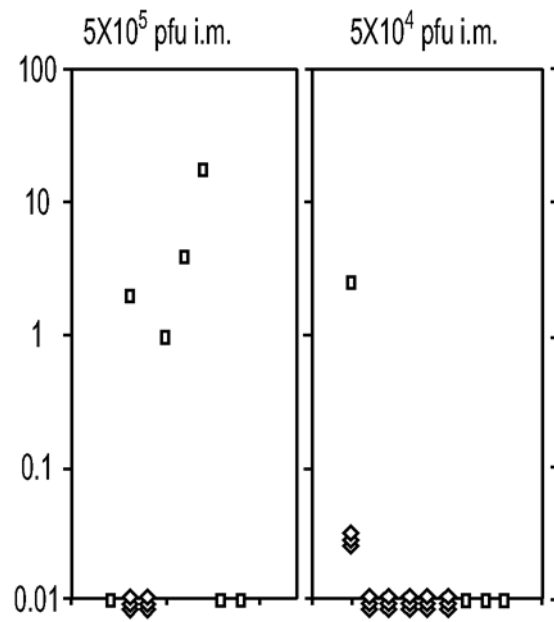


图 17B

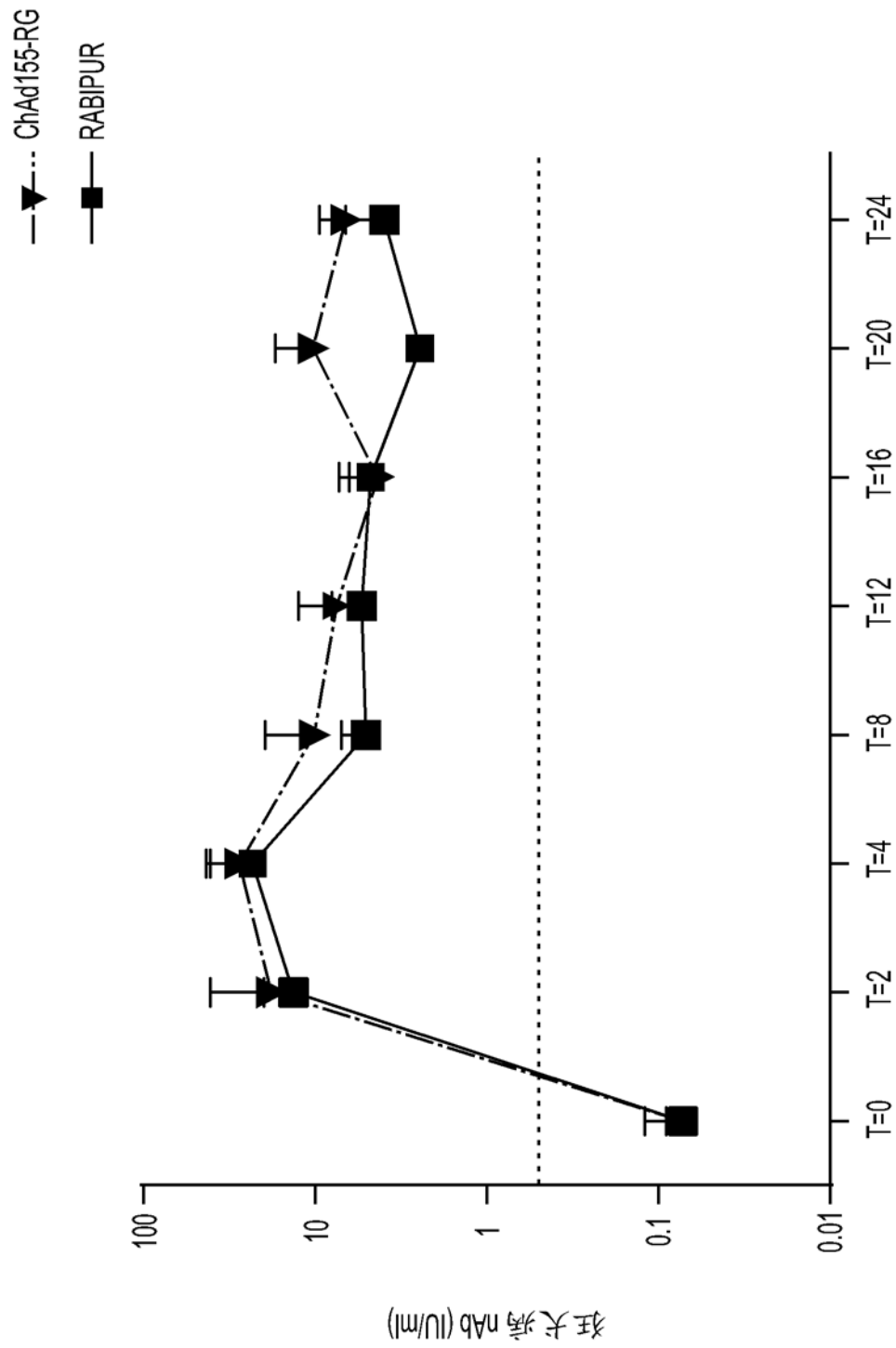


图 18

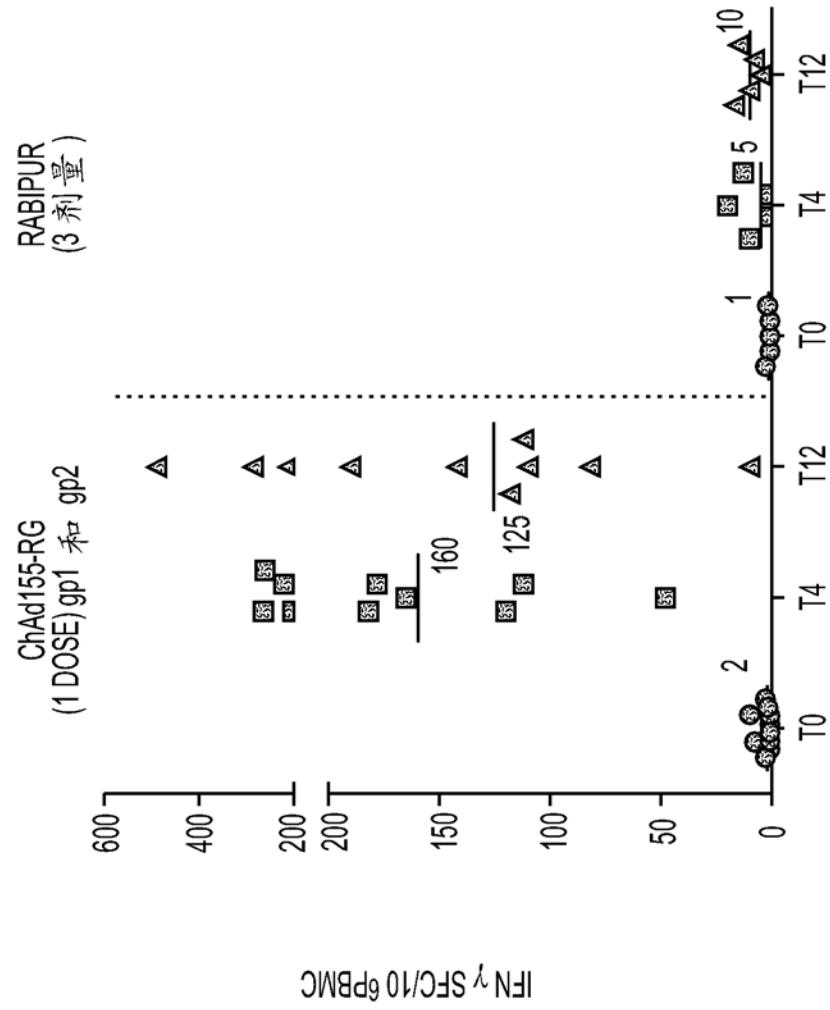


图 19