



(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

(74)专利代理机构 中国专利代理(香港)有限公司

(22)申请日 2017.12.08

72001

(30)优先权数据

代理人 李唐 周齐宏

62/432033 2016.12.09 US

(51)Int.Cl.

62/465378 2017.03.01 US

C07K 14/005(2006.01)

(85)PCT国际申请进入国家阶段日

A61K 39/205(2006.01)

2019.05.31

C12N 15/86(2006.01)

(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

权利要求书7页 说明书40页

地址 比利时里克森萨特

序列表194页 附图22页

(72)发明人 V.阿门多拉 S.科罗卡 A.维特利

B.魏策尔

(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 N0:5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 N0:5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 N0:1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 N0:1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 N0:5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 N0:5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、 $\beta$ -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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3的氨基酸序列的多肽的多核苷酸,

(d) 具有根据SEQ ID N0:1的氨基酸序列的多肽;

(e) 具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,

(b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3的氨基酸序列的多肽的多核苷酸,

(d) 具有根据SEQ ID N0:1的氨基酸序列的多肽;

(e) 具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,
- (b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3的氨基酸序列的多肽的多核苷酸,
- (d) 具有根据SEQ ID N0:1的氨基酸序列的多肽;
- (e) 具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽的多核苷酸,
- (b) 编码具有根据SEQ ID N0:1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 N0:1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 N0:3的氨基酸序列的多肽的多核苷酸,
- (d) 具有根据SEQ ID N0:1的氨基酸序列的多肽;
- (e) 具有根据SEQ ID N0:1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 N0:1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 N0:3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 N0:1的氨基酸序列的多肽;
- (b) 具有根据SEQ ID N0:1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 N0:1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 N0:3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的疾病,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的疾病的感染,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的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  $\gamma$ 、肿瘤坏死因子- $\alpha$  (TNF $\alpha$ ) 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  $\gamma$ 、TNF $\alpha$  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/ TetO 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- $\gamma$  ELISpot

图12:表达RSV转基因的PanAd3和ChAd155载体的免疫原性 - IFN- $\gamma$  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- $\gamma$  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 Δ E1\_TetO hCMV RpsL-Kana引物1多核苷酸序列  
SEQ ID NO:16- BAC/CHAd155 Δ E1\_TetO hCMV RpsL-Kana (#1375) 引物2多核苷酸序列  
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- pjTetO(hCMV\_WPRE\_BghPolyA正向引物  
SEQ ID NO:93- pjTetO(hCMV\_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、I1Ia和I1Va2。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构中存在普遍保守性。腺病毒基因组的每个末端包含被称作反向末端重复(LTR)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'LTR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LTR)。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 $\beta$ -重复组成。所述 $\beta$ -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 $\beta$ -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 $\beta$ -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 $\beta$ -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 Ads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幊上有关的中和滴度(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的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、ABL V、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Procell92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)。Procell92.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/inners.php?m=00084](http://www.okairos.com/e/inners.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 $1 \times 10^5$ 至 $1 \times 10^{15}$ 个病毒颗粒,例如从 $1 \times 10^8$ 至 $1 \times 10^{12}$ (例如, $1 \times 10^8$ 、 $5 \times 10^8$ 、 $1 \times 10^9$ 、 $5 \times 10^9$ 、 $1 \times 10^{10}$ 、 $2.5 \times 10^{10}$ 、 $5 \times 10^{10}$ 、 $1 \times 10^{11}$ 、 $5 \times 10^{11}$ 或 $1 \times 10^{12}$ 个颗粒)。或者,病毒载体可以以通常 $1 \times 10^5$ 至 $1 \times 10^{10}$ 噬斑形成单位(PFU),例如 $1 \times 10^5$  PFU、 $5 \times 10^5$  PFU、 $1 \times 10^6$  PFU、 $5 \times 10^6$  PFU、 $1 \times 10^7$  PFU、 $5 \times 10^7$  PFU、 $1 \times 10^8$  PFU、 $5 \times 10^8$  PFU、 $1 \times 10^9$  PFU、 $5 \times 10^9$  PFU或 $1 \times 10^{10}$  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,用于肌内注射的合适的人剂量(对于约80 kg主体)是在约 $1\times10^5$ 至约 $5\times10^{12}$ 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 $1\times10^7$ 至约 $1\times10^{15}$ 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 $A_{260\text{nm}}$ 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 $\mu\text{g}$ 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 $\mu\text{g}$ 之间,例如在1 ng至1  $\mu\text{g}$ 之间(例如,100 ng-1  $\mu\text{g}$ ),或在1 $\mu\text{g}$ 至100 $\mu\text{g}$ 之间,如10 ng、50 ng、100 ng、150 ng、200 ng、250 ng、500 ng、750 ng或1 $\mu\text{g}$ 。核酸的有效量还可以包括1  $\mu\text{g}$ 至500 $\mu\text{g}$ ,例如1 $\mu\text{g}$ 至200 $\mu\text{g}$ 之间,例如10-100 $\mu\text{g}$ 之间,例如1 $\mu\text{g}$ 、2 $\mu\text{g}$ 、5 $\mu\text{g}$ 、10 $\mu\text{g}$ 、20 $\mu\text{g}$ 、50 $\mu\text{g}$ 、75 $\mu\text{g}$ 、100 $\mu\text{g}$ 、150 $\mu\text{g}$ 或200 $\mu\text{g}$ 。或者,核酸的示例性有效量可以是在100 $\mu\text{g}$ 至1 mg之间,例如100 $\mu\text{g}$ 至500 $\mu\text{g}$ 之间,例如,100  $\mu\text{g}$ 、150  $\mu\text{g}$ 、200  $\mu\text{g}$ 、250  $\mu\text{g}$ 、300  $\mu\text{g}$ 、400  $\mu\text{g}$ 、500  $\mu\text{g}$ 、600  $\mu\text{g}$ 、700  $\mu\text{g}$ 、800  $\mu\text{g}$ 、900  $\mu\text{g}$ 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  $\mu\text{g}$ 至约50  $\mu\text{g}$ 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。

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 N0:5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 N0:5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 N0:5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 N0:3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 N0:1的氨基酸序列的多肽；或

(b) 具有根据SEQ ID N0:1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 N0:1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 N0:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID N0:5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 N0:5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 N0:1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 N0:1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 N0:1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 N0:5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 N0:5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 N0:5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 N0:1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 N0:2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 N0:1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 N0:1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 N0:1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 N0:1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 N0:1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 N0:3的氨基酸序列的多肽；或
- (b) 具有根据SEQ ID N0:3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 N0:3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 N0:5的氨基酸序列的多肽；或
  - (b) 具有根据SEQ ID N0:5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 N0:5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 N0:3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 N0:3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 N0:3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 N0:5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 N0:5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 N0:5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 N0:3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 N0:4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 N0:1的氨基酸序列的多肽；或

(b) 具有根据SEQ ID N0:1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 N0:1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 N0:5的氨基酸序列的多肽；或

(b) 具有根据SEQ ID N0:5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 N0:5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 N0:1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 N0:1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 N0:1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：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。

[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的氨基酸分组之一中,尽管在一些情况下,其他取代也是可能的,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	<b>GCTNLSEFS</b>	KRA	KLCGVL	FH	<b>KSVRTWNEI</b>	KG
ABLV	I	<b>GCTSLSGFS</b>	KKA	KLCGIS	FN	<b>KSVRTWDEI</b>	KG
ARAV	I	<b>GCTNLSGFT</b>	KKA	KLCGVN	FH	<b>KSVREWTEV</b>	KG
BBLV	I	<b>GCTTLTVFS</b>	KKA	KLCGVN	FH	<b>KSIROWTEI</b>	KG
DUVV	I	<b>GCTTLTPFS</b>	KKA	KLCGIS	FH	<b>KSVREWKEI</b>	KG
EBLV-1	I	<b>GCTTLTPFS</b>	KKA	KLCGVP	FH	<b>KSVREWKEV</b>	KG
EBLV-2	I	<b>GCTTLTVFS</b>	KKA	KLCGIS	FH	<b>KSIREWTDV</b>	KG
IRKV	I	<b>GCTTLTAFN</b>	KKA	KLCGMA	DR	<b>KSIREWKEI</b>	KG
KHUV	I	<b>GCTTLSGFT</b>	KRA	KLCGVN	FH	<b>KSIREWSEI</b>	KG
LBV	II	<b>GCSDTATFS</b>	KKS	TLCGKP	NR	<b>LRVDSWNDI</b>	KG
MOKV	II	<b>GCNTESPFT</b>	QKA	TLCGKP	DR	<b>KRVDRWADI</b>	KG
SHIV	II	<b>GCSSSSTFS</b>	KKS	TLCGKP	NR	<b>KRVDRWEEI</b>	KG
WCBV	III	<b>YCTTEQSIT</b>	KLV	SICGRQ	IK	<b>IKVENWSEV</b>	KG
IKOV	?	<b>GCNEGSKVS</b>	ILL	IICGKS	VK	<b>KSVDNWTDI</b>	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次肌内剂量的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次肌内剂量的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] 目前有几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天肌内施用三次剂量组成。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天肌内施用五次剂量组成。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天肌内施用五次剂量组成。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/ RABAVERT (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;AlPO<sub>4</sub>;铝胶;来自革兰氏阴性细菌外膜的细菌产物,特别是单磷酰脂质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Pam<sub>3</sub>Cys;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(例如皂昔,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-1β、IL-2、IL-7、IL-12、IL-18、GM-CFS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(例如单磷酰脂质A,例如3-脱-0-酰化单磷酰脂质A(3D-MPL)或胞壁酰肽)、合成佐剂(例如单磷酰脂质A(MPL),特别是3-脱-0-酰化单磷酰脂质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] 皂昔也是合适的佐剂,例如,源自南美树皂树Molina的树皮的皂昔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时,CpG被称为佐剂。当配制成为疫苗时,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,特别是单磷酰脂质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,单磷酰脂质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±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\_TetO 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 (tetO) 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/ TetO 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/TetO 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BglII和SpeI消化PCR片段,并克隆到用BglII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  $\Delta$  E1\_Tet0 hCMV RpsL-Kana (#1375) 作为模板,利用以下寡核苷酸5' -ATTCAGTGTACAGGCGCGCAAAGCATGACGCTGTTGATTGATTC-3' (SEQ ID NO: 15) 和5' -ACTAGGACTAGTTATAAGCTAGAATGGGGCTTGCG-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  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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  $\Delta$  E1\_Tet0 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  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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pvjTet0hCMV\_RG\_bghpolyA用SpeI、SphI和AsiSI切割以切除含有具有tet0、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BGHpolyA控制下的RpsL-Kana选择盒的pChAd155  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 hCMV-RG载体 (#1481) (图7)。

[0155] 2.2.3: *hCMV-RG-WPRE表达盒的插入 – pChAd155  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 hCMV-RG-WPRE#1509的构建*

通过利用pChAd155  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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pvjTet0hCMV\_WPRE\_BghPolyA (#1478) 作为模板并利用以下寡核苷酸FW 5' -ggaaggtcagcgtgaccagccagtcgg caaagtgattcctcggagagctataaaaagcggcggagagaccaggctgtgatgagcggccgcgtctgtat caacctctggattaca -3' (SEQ ID NO: 92) 和RW 5' - ATGGCTCCGGCGGTCTCTGCAACACAAATAA AGAGACCCTAAGACCCCCAACTTATATATTTCATGACCACCCCAGGCCACGCCACTCACCCACCTCACCATAGA GCCCACCGCATCC-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BGHpolyA控制下的RG转基因(SEQ ID NO:38)的pChAd155  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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  $\Delta$  E1, E4\_Ad5E4orf6/Tet0 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' - ACATCAATGGCGTGGATAGCGGTT-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' - ACATCAATGGCGTGGATAGCGGTT-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 $5 \times 10^5$ 个细胞/ml的细胞密度感染Procell 92.S平行比较PanAd3/RSV (在WO2012/089833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	ug/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。通过肌内注射 $10^6$ 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- $\gamma$ 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- $\gamma$ 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 $3 \times 10^6$  vp的剂量进行肌内注射。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- $\gamma$ -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- $\gamma$ 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（F<sub>51-66</sub>，F<sub>85-93</sub>和M2-1<sub>282-290</sub>）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- $\gamma$  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。通过肌内注射 $10^9$  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（如在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- $\gamma$  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( $5 \times 10^8$ 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的单次肌内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 $5 \times 10^5$  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 $5 \times 10^4$  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 $5 \times 10^5$  pfu 相比于ChAd155-RG以 $10^4$  pfu)时,在发明人的本研究和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)	$5 \times 10^7$ 个病毒颗粒	44%	77%
7	AdC68rab(gp)	$5 \times 10^6$ 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  $5 \times 10^{10}$ 病毒颗粒IM免疫第1组,然后在第48周用加强剂量的ChAd155-RG  $5 \times 10^{10}$ 病毒颗粒IM免疫。用ChAd155-RG  $5 \times 10^{10}$ 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,和在第48周IM加强剂量的ChAd155-RG  $5 \times 10^{10}$ 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抗原的猿猴腺病毒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  $\gamma$  分泌T细胞,如通过IFN  $\gamma$  -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  $\gamma$  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  $\gamma$  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内注射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 WO

<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 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

[0002]

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

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 tgtaccccta tgacacggaa	60
agcggccctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa	120
gaaagtcccc ccggggtctt gtctctgaac ctggccgagc ccctggcac ttcccacggc	180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgctggcaa ctcacccct	240
caagatatac ccaccgctag ccctccctc aaaaaaacca agaccaaccc cagcctagaa	300
acctcatecc ccctaactgt gagcaccta ggcgcctca ccgtagcagc cgccgctccc	360
ctggcggtgg ccggcacctc ctcaccatg caatcagagg ccccccgtac agtacaggat	420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg	480
caaacatcgg ccccgctgac ggccgctgac agcagcaccc tcacagtcag tgccacacca	540
cccttagca caagcaatgg cagcttgggt attgacatgc aagccccat ttacaccacc	600
aatggaaaac taggacttaa ctttggcgct cccctgcatg tggtagacag cctaaatgca	660
ctgactgttag ttactggcca aggtcttacg ataaacggaa cagccctaca aactagagtc	720
tcaggtgccc tcaactatga cacatcagga aacctagaat tgagagctgc agggggat	780
cgagttgatg caaatggtca acttacccctt gatgttagctt acccatttga tgcacaaaac	840
aatctcagcc ttaggcttgg acagggaccc ctgtttgtta actctgccc caacttggat	900
gttaactaca acagaggcct ctacctgttc acatctggaa atacaaaaaa gctagaagtt	960
aatatcaaaa cagccaaggg tctcattttat gatgacactg ctatagcaat caatgcgggt	1020
gatggctac agtttactc aggtctagat acaaattccat taaaaactaa acttggatta	1080
ggactggatt atgactccag cagagccata attgctaaac tggaaactgg cctaagctt	1140
gacaacacag gtgccatcac agtaggcaac aaaaatgtg acaagcttac cttgtggacc	1200
acaccagacc catccctaa ctgtagaatc tattcagaga aagatgctaa attcacactt	1260
gttttacta aatgcggcag tcaggtgttg gccagcgttt ctgttttac tggaaaaggt	1320
agccttgcgc ccatcagtgg cacagtaact agtgcctaga ttgtcctcag atttgatgaa	1380
aatggagttc tactaagcaa ttctccctt gaccctcaat actggaaacta cagaaaaggt	1440

[0005]

gaccttacag	agggcactgc	atataccaaac	gcagtggat	ttatgccaa	cctcacagca	1500
tacccaaaaa	cacagagcca	aactgctaaa	agcaacattg	taagtcaggt	ttacttgaat	1560
ggggacaaat	ccaaacccat	gaccctcacc	attaccctca	atggaactaa	tgaaacagga	1620
gatgccacag	taagcactta	ctccatgtca	ttctcatgga	actggaatgg	aagtaattac	1680
attaatgaaa	cgttccaaac	caactccttc	acttctcct	acatgccca	agaa	1734

&lt;210&gt; 3

&lt;211&gt; 593

&lt;212&gt; PRT

&lt;213&gt; 黑猩猩腺病毒

&lt;400&gt; 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	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 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 Leu  
370 375 380

Ala Ala Ala Glu Ala Glu Ser Glu Gly Thr Lys Pro Ala Lys Glu Pro  
385 390 395 400

Val Ile Lys Pro Leu Thr Glu Asp Ser Lys Lys Arg Ser Tyr Asn Leu  
405 410 415

Leu Lys Asp Ser Thr Asn Thr Ala Tyr Arg Ser Trp Tyr Leu Ala Tyr  
420 425 430

Asn Tyr Gly Asp Pro Ser Thr Gly Val Arg Ser Trp Thr Leu Leu Cys  
435 440 445

Thr Pro Asp Val Thr Cys Gly Ser Glu Gln Val Tyr Trp Ser Leu Pro  
450 455 460

Asp Met Met Gln Asp Pro Val Thr Phe Arg Ser Thr Arg Gln Val Ser  
465 470 475 480

Asn Phe Pro Val Val Gly Ala Glu Leu Leu Pro Val His Ser Lys Ser  
485 490 495

[0007] Phe Tyr Asn Asp Gln Ala Val Tyr Ser Gln Leu Ile Arg Gln Phe Thr  
500 505 510

Ser Leu Thr His Val Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Ala  
515 520 525

Arg Pro Pro Ala Pro Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala  
530 535 540

Leu Thr Asp His Gly Thr Leu Pro Leu Arg Asn Ser Ile Gly Gly Val  
545 550 555 560

Gln Arg Val Thr Val Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val  
565 570 575

Tyr Lys Ala Leu Gly Ile Val Ser Pro Arg Val Leu Ser Ser Arg Thr  
580 585 590

Phe

<210> 4

[0008]

<211> 1779  
<212> DNA  
<213> 黑猩猩腺病毒

<400> 4  
atgcggcgcg cggcgatgta ccaggaggga cctcctccct cttacgagag cgtggtggc 60  
gcgcgccgg cgccgcctc ttctccctt gcgtcgacg tgctggagcc gccgtacgt 120  
cctccgcgt acctgccc tacgggggg agaaacagca tccgttactc ggagctggc 180  
ccctgttcg acaccacccg ggtgtacctg gtggacaaca agtcggcggc cgtggcctcc 240  
ctgaactacc agaacgacca cagcaatttt ttgaccacgg tcatccagaa caatgactac 300  
agcccgagcg aggcacgac ccagaccatc aatctggatg accggtcgca ctggggcggc 360  
gacctgaaaa ccattctgca caccaacatg cccaacgtga acgagttcat gttcaccat 420  
aagttcaagg cgccgggtgat ggtgtcgcc tcgcacacca aggaagacccg ggtggagctg 480  
aagtacgagt gggtggagtt cgagctgcca gagggcaact actccgagac catgaccatt 540  
gacctgatga acaacgcgt cgtggagcac tatctgaaag tggcaggca gaacggggtc 600  
ctggagagcg acatcggggt caagttcgac accaggaact tccgcctgg gctggacccc 660  
gtgaccgggc tggttatgcc cgggggtgtac accaacgagg cttccatcc cgacatcatc 720  
ctgctgccc gctcggggt ggacttcact tacagccgcc tgagcaacct cctggcatac 780  
cgcaagcggc agcccttcca ggaggccttc aggatcacct acgaggaccc ggagggggc 840  
aacatccccg cgctctcgta tgtggaggcc taccaggata gcttgaagga aaatgaggcg 900  
ggacaggagg ataccgcccc cgccgcctcc gcccgcggc agcagggcga ggatgctgct 960  
gacaccgcgg ccgcggacgg ggcagaggcc gacccgcta tgggtggta ggctcccgag 1020  
caggaggagg acatgaatga cagtgcggtg cgccggagaca cttcgaccc ccggggggag 1080  
gaaaagcaag cggaggccga gcccggcc gaggaaaagc aactggcggc agcagggcgc 1140  
gcggcggcgt tggccggc ggaggctgag tctgagggga ccaagcccgc caaggagccc 1200  
gtgattaagc ccctgaccga agatagcaag aagcgcagtt acaacactgct caaggacagc 1260  
accaacaccc cgtaccgcag ctggtaacct gcctacaact acggcggacc gtcgacgggg 1320  
gtgcgcctt ggaccctgct gtgcacgcgg gacgtgaccc gcccgcggc gcagggttac 1380  
tggtcgtgc ccgacatgtat gcaagacccc gtgacccccc gctccacgcg gcagggtcagc 1440  
aacttcccg tggtgccgc cggactgctg cccgtcaact ccaagagctt ctacaacgac 1500  
caggccgtct actcccgact catccgcag ttcacccctc tgacccacgt gttcaatcgc 1560

[0009]

tttcctgaga accagattct ggcgcggccg cccgccccca ccatcaccac cgtcagtcaa	1620
aacgttctcg ctctcacaga tcacgggacg ctaccgctgc gcaacagcat cggaggagtc	1680
cagcgagtga ccgttactga cgccagacgc cgcacctgcc cctacgtta caaggccttg	1740
ggcatagtct cgccgcgcgt ccttccagc cgcaacttt	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 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 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 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	catcgatgtat	gccgcagtgg	tcgtacatgc	acatctcgaa	ccaggacgcc	60
tcggagttacc	tgagccccgg	gctgggtcag	ttcgcccccg	ccaccgagag	ctacttcagc	120
ctgagtaaca	agtttaggaa	ccccacggtg	gcccacgc	acgatgtgac	caccgaccgg	180
tctcagcgcc	tgacgctgct	gttcattccc	gtggaccgcg	aggacaccgc	gtactcgtac	240
aaggcgcgg	tcaccctggc	cgtggcgcac	aaccgcgtgc	tggacatggc	ctccacctac	300
tttgacatcc	gcggggtgct	ggaccgggt	cccacttca	agccctactc	tggcaccgcc	360
tacaactccc	tggcccccaa	gggcgcctcc	aactcctgcg	agtgggagca	agagggaaact	420
caggcagttg	aagaagcagc	agaagaggaa	gaagaagatg	ctgacggta	agctgaggaa	480
gagcaagcag	ctacaaaaaa	gactcatgta	tatgctcagg	ctcccttcc	tggcgaaaaaa	540
attagtaaag	atggtctgca	aataggaacg	gacgctacag	ctacagaaca	aaaacctatt	600

[0014]

tatgcagacc ctacattcca gcccgaaccc caaatcgaaa agtcccagtg gaatgaggca	660
gatgctacag tcgcccggcgg tagagtgcta aagaaatcta ctcccatgaa accatgctat	720
ggttcctatg caagacccac aaatgcta at ggaggtcagg gtgtactaac ggcaaattgc	780
cagggacagc tagaatctca gttgaaaatg caattcttt caacttctga aaacgcccgt	840
aacgaggcta acaacattca gcccaaattg gtgctgtata gtgaggatgt gcacatggag	900
accccgata cgcacccccc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg	960
ctgggtcagc agtccatgcc caacagaccc aattacatcg gcttcagaga caactttatc	1020
ggcctcatgt attacaatag cactggcaac atgggagtgc ttgcaggta ggcctctcag	1080
ttgaatgcag tgggtggactt gcaagacaga aacacagaac tgtcctacca gctttgctt	1140
gattccatgg gtgacagaac cagatacttt tccatgtgaa atcaggcagt ggacagttat	1200
gacccagatg tttagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt	1260
ttccctctgg gtggcatagg ggttaactgac acttaccagg ctgttaaaac caacaatggc	1320
aataacgggg gccaggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata	1380
gggttggaa acaatttcgc tatggagatc aacctcagtg ccaacctgtg gagaaacttc	1440
ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaaatgt	1500
gacatctctg acaacccaa cacctacgat tacatgaaca agcgagtggt ggccccgggg	1560
ctgggtggact gctacatcaa cctggcgcg cgctggtcgc tggactacat ggacaacgtc	1620
aacccttca accaccacccg caatgcgggc ctgcgctacc gctccatgct cctggcaac	1680
ggcgctacg tgcccttcca catccaggtg ccccagaagt tctttgccat caagaacctc	1740
ctcctctgc cggcctcta cacctacgag tggaaactca ggaaggatgt caacatggtc	1800
ctccagagct ctctggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc	1860
atctgcctct acgccaccc ttccccatg gcccacaaca cggcctccac gctcgaggcc	1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctctccgc cgccaaacatg	1980
ctctacccca taccggccaa cgccaccaac gtccccatct ccatccccctc ggcacactgg	2040
gcggccttcc gcggctggc cttcacccgc ctcaagacca aggagacccc ctccctggc	2100
tcgggattcg acccctacta cacctactcg ggctccattc cctacctgga cggcaccc	2160
tacctcaacc acactttcaa gaaggctcg gtcacccctcg actcctcggt cagctggccg	2220
ggcaacgacc gtctgctcac ccccaacgag ttgcagatca agcgctcggt cgacggggag	2280
ggctacaacg tggcccagtg caacatgacc aaggactggt tcctggtcca gatgctggcc	2340

[0015]

aactacaaca tcggcttacca gggcttctac atcccaagaga gctacaagga caggatgtac	2400
tccttcttca ggaacttcca gcccatgagc cggcagggtgg tggaccagac caagtacaag	2460
gactaccagg aggtggcat catccaccag cacaacaact cgggcttcgt gggctacctc	2520
gccccccacca tgcgcgaggg acaggcctac cccgccaact tcccctatcc gctcataggc	2580
aagaccgcgg tcgacagcat cacccagaaa aagttcctct gcgaccgcac cctctggcgc	2640
atccccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg	2700
ctctacgcca actccggcca cgcctcgac atgaccttcg aggtcgaccc catggacgag	2760
cccaccccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac	2820
cgcggcgtca tcgagaccgt gtacctgcgt acgccttct 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 attttgattt gaagccaata tgataatgag atggcgccg  
cgggcgcccc cgccggcg gaggcggtt tggggcgcc cgccggcg gggcggttg  
gcggaaagtgg actttgttaag tgtggcgat gtgacttgct agtgcgggc gcgtaaaag  
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca ttttccgc ggttttacc  
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttgc cattttcg cggaaaaact  
gaaacgggga agtgaatct gattaatttt gcgttagtca taccgcgtaa tatttgctta  
ggccgaggg actttggccg attacgtgga ggactcgccc aggttttt tgaggtgaat  
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacg  
gtatccatataataatgt acatttatat tggctcatgt ccaacattac cgccatgtt  
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc  
atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccc  
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac  
tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagttacatca  
agtgtatcat atgccaagta cgcccccatt tgacgtcaat gacggtaaat ggccgcctg  
840

[0016]

gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt	900
agtcatcgct attaccatgg tcatgcgggtt ttggcagtac atcaatggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga gtttgggg	1020
gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaa	1080
ggcggtagg cgtgtacggt gggaggtcta tataaggcaga gctctcccta tcagtgatag	1140
agatctccct atcgtgata gagatcgctg acgagctcg ttagtgaacc gtcagatcg	1200
ctggagacgc catccacgct gtttgaccc ccatagaaga caccgggacc gatccagc	1260
cccgccggcga aacggtgca ttggAACGCG gattccccgt gccaagagtg agatttccg	1320
tttatctagg taccgggccc cccctcgagg tcgacggat cgataagctt cacgctgcc	1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgttagaaag ccagtccgca	1440
gaaacggtgc tgaccccgga tgaatgtcag ctactggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagcttgcag tggcgttaca tggcgatagc tagactggc	1560
ggttttatgg acagcaagcg aaccggattt gccagctggg gcccctctg gtaaggttgg	1620
gaagccctgc aaagtaaact ggatggcttt cttgcccaca aggatctgat ggccgagg	1680
atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac	1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgttgcg agcatgccc caaaaacgt	1800
gcgtatgtac tcgtgtatct actaccactc ctaaaaaacc gaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggt ggtgaagg	1920
acaacctgca ggagcactcc gtgatctga tccgtggcgg tcgtgtaaa gaccccccgg	1980
gtgttcgtta ccacaccgta cgtggtgcgc ttgactgctc cggcgtaaaa gaccgta	2040
aggctcggtc caagtatggc gtgaagcgtc ctaaggctt atggtagatc tgatcaagag	2100
acaggatgac ggtcggttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggtatgac tggcacaac agacaatcg ctgctctgat	2220
gccggcgtgt tccggctgtc agcgcaggccc cgccgggttc ttttgcgtt gaccgac	2280
tccgggtccc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct ggccacgac	2340
ggcggttcctt ggcgcgtgt gctcgacgtt gtcactgaag cgggaaggaa ctggctgct	2400
ttggcgttgc ggcgcgtgt gctcgacgtt gtcactgaag cgggaaggaa ctggctgct	2460
tccatcatgg ctgatgcaat gggcggtc catacgctt atccggctac ctgcccattc	2520

[0017]

gaccaccaag cgaaacatcg catcgagcga gcacgtactc gnatggaagc cggtcttgtc	2580
gatcaggatg atctggacga agagcatcgag gggctcgcc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgtcgta cccatggcga tgcctgcttg	2700
ccgaatatca tggtgaaaaa tggccgctt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccgtg atattgtga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcggtt tacggatcg ccgctcccga ttgcagcgc	2880
atgccttct atgccttct tgacgagttc ttctgagcgg gactctgggg ttgcggaaatga	2940
ccgaccaagc gacgcccac ctgcacatcac gagatttgcg ttccaccggc gccttctatg	3000
aaagggttggg cttcggaaatc gtttccggg acgcccggctg gatgatcctc cagcgcgggg	3060
atctcatgct ggagttttc gcccaccccg ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgat ctgtgtgcc ttctagttgc cagccatctg ttgtttgcc	3180
ctccccgtg cttccttga ccctggagg tgccactccc actgtcctt cctaataaaaa	3240
tgaggaaatt gcatcgatt gtctgagtag gtgtcattct attctgggg gtgggggtgg	3300
gcaggacagc aagggggagg attgggaaaga caatagcagg catgctgggg atgcgggtgg	3360
ctctatggcc gatcagcgat cgctgagggtg ggtgagttgg cgtggcctgg ggtggcgt	3420
aaaatatata agttgggggt cttagggctt ctttatttgt gttgcagaga ccgcccggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgc ttggatggca gcatcgtag	3540
cccttatttgc acgacgcgga tgccccactg ggccgggtg cgtcagaatg tgcggctc	3600
cagcatcgac ggccgaccccg tcctgcccgc aaattccgc acgctgaccc atgcgaccgt	3660
cgcggggacg ccgttggacg ccaccggcgc cgccgccc accgcagccg cctcgccgt	3720
gcccgcctg gccacggact ttgcattctt gggaccactg ggcacagggg ctacttctg	3780
ggccgctgt gccgcgttc gcatgacaa gctgaccgc ctgtggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcgcc agcaggctc	3900
ctccctgcaa gctggggga atgcttctcc cacaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgt ctctttattt cataatttc	4020
cgcgcgcgat aggcctaga ccagcggtt cggcggtga gggcggtg tatcttctcc	4080
aggacgtgggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccggggg	4140
tggaggttagc accactgcag agcttcatgc tccgggtgg tggtagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260

[0018]

gggaggccct	tggtgttaagt	gtttacaaaa	cggtaagtt	gggaagggtg	cattcgggga	4320
gagatgatgt	gcatcttgg	ctgtatTTT	agattggcga	tgttccg	cagatcc	4380
ctgggattca	tgttgcag	gaccacc	acagtgtatc	cgggcactt	ggggaa	4440
tcatgcagct	tagaggaaa	agcgtggaag	aactggaga	cgcctt	gcctccc	4500
ttttccatgc	attcg	ccat	gatgtggca	atggcccgc	gggaggc	4560
atatttctgg	ggtcg	ctgac	gtcgtat	tg	ttggc	4620
tttacaaagc	gcgggcggag	ggtgccc	gac	tggggat	ga	4680
gcgtat	ttgc	cctcg	cagat	cttcc	tgc	4740
tccac	ctg	cg	gat	gg	atc	4800
agcagg	ttc	ta	agc	at	tttca	4860
accgg	ttg	ca	ggc	tttcc	cc	4920
ac	cgtt	gat	gtt	cc	cc	4980
tcg	cc	gg	cc	cc	cc	5040
tcc	cc	gtt	gg	cc	cc	5100
gt	ac	gtt	ttt	cc	cc	5160
ttt	cg	ct	ttt	cc	cc	5220
ggc	cac	ca	gg	ttt	cc	5280
cc	cc	cc	gg	cc	cc	5340
cct	cg	cg	cc	cc	cc	5400
gt	cc	cc	cc	cc	cc	5460
tc	cc	cc	cc	cc	cc	5520
ag	cc	cc	cc	cc	cc	5580
aa	acc	cc	cc	cc	cc	5640
gt	ccc	cc	cc	cc	cc	5700
tc	cc	cc	cc	cc	cc	5760
cc	cc	cc	cc	cc	cc	5820
gg	gg	gg	gg	gg	gg	5880
tg	at	gg	cc	cc	cc	5940

[0019]

gggtggggc gcgctcgctcg tcactctttt ccgcacatcgct gtctgcgagg gccagctgct	6000
gggtgagta ttccctctcg aaggcggca tgacccgc gctgagggttgc tagttcca	6060
aaaacgagga ggatttgatg ttcacctgtc ccgaggtgat acctttgagg gtacccgcgt	6120
ccatctggtc agaaaacacg atcttttat tgtccagctt ggtggcgaac gaccgcgt	6180
ggcggttggaa gaggcgttgc gcgatggagc gcagggtctg gttttgtcc ctgtcgccgc	6240
gctccttggc cgcgatgttgc agctgcacgt actcgccgc gacgcagcgc cactcgccgt	6300
agacggtggt gcgctcgctcg ggcaccaggc gcacgcgc gccgcgggttgc tgcagggt	6360
ccaggtccac gctggtggcg acctcgccgc gcaggcgctc gttggtccag cagagacggc	6420
cgccttgcg cgagcagaag ggggcaggc ggcgtcgaa gtagtctatc ttgcaacctt	6480
gcatgtccag cgcctgctgc cagtcgcggc cggcgagcgc gcgtcgtag ggttggcg	6540
gcgggccccca gggcatgggg tgggtgagtg cggaggcgta catgcccgc atgtcataga	6600
cgtagagggg ctcccgcagg accccgatgt aggtgggttgc agcagcggccg ccgcggatgc	6660
tggcgccac gtagtcatac agctcgctcg agggggcgag gaggtcgccccc cccagggttgg	6720
tgcggccggc ggcgtccgcg cggaaagacga tctgcctgaa gatggcatgc gagttggaaag	6780
agatggtggg ggcgtggaaag acgttgaagc tggcgctctg caggccgacg gcgtcgccca	6840
cgaaggaggc gtaggagtcg cgcagttgt gtaccagctc ggcggtgacc tgcacgtcga	6900
gcgcgcagta gtcgagggtc tcgcggatga tgtcatatcc agcctgcgc ttcttttcc	6960
acagctcgcg gttgaggaca aactttcgcc ggtctttcca gtacttttgg atcggaaac	7020
cgtccgggttc cgaacggtaa gagcctagca tgtagaactg gttgacggcc tggtaggcgc	7080
agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg	7140
tcagggcgaa ggtgtccctg accatgactt tgaggtactg gtgcttgaag tcggagtcgt	7200
cgcagccgc cgcgtcccgag agcgagaagt cgggtcgctt cttggagcgg gggttggca	7260
gagcgaaggt gacatcggtt aagaggattt tgccgcgc gggcatgaag ttgcgggtga	7320
tgcggaaaggc ccccgact tcagagcggt tggtgatgac ctggcgccgc agcacgatct	7380
cgtcgaagcc gttgatgttgc tggccacga tgttagagttc caggaagcgg ggcggccct	7440
ttacgggtggc cagtttttgc agctttcgat aggtgagctc ctcggcgag ggcggccgt	7500
gctcgccag gcccagtc gcgaggcg ggttgcgtt gaggaggac ttccagagggt	7560
cgcggccag gagggtctgc aggcggtctc tgaaggctt gaaactggcgg cccacggcca	7620
cgcggccag gagggtctgc aggcggtctc tgaaggctt gaaactggcgg cccacggcca	7680

[0020]

tttttcggg ggtgatgcag tagaaggtga ggggtcttgc tgccagcgg tcccagtcga	7740
gctgcaggc gaggtcgcgc gcggcggtga ccaggcgctc gtcgcccccg aatttcatga	7800
ccagcatgaa gggcacgagc tgcttccga aggccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtgcgag gatgcgagcc gatcggaaag aactggatct	7920
cccgccacca gttggaggag tggctgttga tgtggtgaa gtagaagtcc cgtcgccggg	7980
ccgaacactc gtgctggctt ttgtaaaagc gagcgcagta ctggcagcgc tgcacggct	8040
gtacctcatg cacgagatgc accttcgccc cgccacgag gaagccgagg gaaatctga	8100
gcccccggcc tggctcgccg catggctggg tctttctac tttggatgctg tgtccgtctc	8160
cgtctggctc ctcgaggggt gttacgggtt agcggaccac cacggccgcg gagccgcagg	8220
tccagatatac ggccgcggc ggtcgagtt ttagtgcac atcgcgcagc tggagctgt	8280
ccatggctg gagctcccgc ggcggcggca ggtcagccgg gagttcttc aggttcacct	8340
cgcagagtcg ggccagggcg cggggcaggt ctaggtggta cctgatctt agggcgtgt	8400
tggggcggc gtcgatggct tgcaggagcc cgcagcccg gggggcgacg acgggtcccc	8460
gcgggggtgg ggtgggtggc gcggtgcagc tcagaagcgg tgccgcggc gggcccccgg	8520
aggtaggggg ggctccggc cccggggcag gggcgccagc ggcacgtcgg cgtggagcgc	8580
ggcaggagt tggctgttg cccggagggtt gctggcgaag ggcacgcacgc ggcgggttat	8640
ctcctggatc tggcgcctt cgtgaagac gacggcccg gtgagcttga acctgaaaga	8700
gagttcgaca gaatcaatct cgggttcatt gaccggcc tggcgcagga ttcctgcac	8760
gtctcccgag ttgtcttggt aggcgatctc ggcattaaac tgctcgatct cttccctctg	8820
gaggtctccg cgtccggcgc gttccacggt ggcggccagg tcgttgaga tgcgccccat	8880
gagctgcgag aaggcgttga gtccgcctc gttccagact cggctgtaga ccaccccc	8940
ctggcatcg cggcgcgca tgaccacctg cgcgagggtt agctccacgt ggcgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaaagag gtagttgagg gtggggcgg tgtgctcgcc	9060
cacaaagaag ttcatgaccc agcggcgcaa cgtggattcg ttgtatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggcgaagttt aaaaactggg agttgcgcgc	9180
cgacacggtc aactccctt ccagaagacg gatgagctcg ggcacgggtt cgcgcaccc	9240
cgctcgaaag gctatggga tctttccctc cgctagcatc accacccctt cctttccctc	9300
ctcttctggc acttccatga tggcttcctc ctcttcgggg ggtggcgccg gcggcggtgg	9360

[0021]

gggagggggc gctctgcgcc ggcggcgccg caccgggagg cggtccacga agcgcgcgat 9420  
catctccccg cggcgccggc gcatggtctc ggtgacggcg cggccgttct cccggggcg 9480  
cagtttgaag acggccggc acatctggtg ctggggcg ggccgtgag gcagcgagac 9540  
ggcgctgacg atgcatctca acaattgctg cgtaggtacg cggccgaggg acctgaggga 9600  
gtccatatacc accggatccg aaaaccttc gaggaaggcg tctaaccagt cgcaatcgca 9660  
aggtaggctg agcaccgtgg cggcgccgg ggggtgggg gagtgtctgg cggagggtct 9720  
gctgatgatg taattgaagt aggccgactt gacacggcg atggtcgaca ggagcaccat 9780  
gtccttgggt ccggcctgct ggatgcggag gcggcggct atgcccagg ctgcgttctg 9840  
gcatcgccgc aggtccttgt agtagtcttgc catgacccctt tccaccggca cctttctcc 9900  
ttccttttct gcttcttcca tgtctgttcc ggcctgggg cggcgcccgccccctgccc 9960  
ccccatgcgc gtgaccccgaa cccctgag cgggtggagc agggccaggt cggcgacgac 10020  
gctcgccgc agatggcct gctgcacctg cgtgagggtg gtttggaaat catccaagtc 10080  
cacgaagcgg tggtaggcgc ccgtgttgc ggttaggtg cagttggcca tgacggacca 10140  
gttgcggc tggtagggccg gttgcacat ctcgggttac ctgagtcgcg agtaggcgcg 10200  
ggagtcgaag acgtagtcgt tgcaagtccg caccaggtac tggtagccca ccaggaagtg 10260  
cgccggcgcc tggcggtaga gggccagcg cagggtggcg gggctccgg gggccaggc 10320  
ttccagcatg aggccgttgtt aggcgttagat gtacctggac atccaggtga taccggcggc 10380  
ggtgtggag gctcgccgcg agtcgcac ccgttccag atgttgcgcgaa gggcagaaa 10440  
gtgctccatg gtaggcgtgc tctgtccagt cagacgcgcg cagtcgttga tactcttagac 10500  
cagggaaaaac gaaagccggt cagcggcac tctccgtgg tctggtaat agatgcgaag 10560  
ggtatcatgg cggagggcct cggtcgagc cccgggtccg ggccggacgg tccggcatga 10620  
tccacgcggc taccggccgc gtgtcgaacc caggtgtgcg acgtcagaca acggtgagg 10680  
gttccttttgc gcgttttctt ggccggcgcc cggcgccgcg taagagacta agccgcgaaa 10740  
gcgaaagcag taagtggctc gctccccgtt gccggaggga tccttgctaa gggttgcgtt 10800  
gcggcgaacc cgggttcgaa tcccgactc gggccggccg gaccggcggc taaggtgttgc 10860  
gattggcctc cccctcgat aaagaccccg cttgcggatt gactccggac acggggacga 10920  
gcccccttttta ttttgctt cccagatgc atccgggtct gcggcagatg cgccccccgc 10980  
cccagcagca gcaacaacac cagcaagagc ggcagcaaca gcagcggag tcatgcaggg 11040  
ccccctcacc caccctcgcc gggccggcca ctcggcgctc cggccgcgtg tctggcgct 11100

[0022]

gcggcggcgg cggggggccg gctgacgacc ccgaggagcc cccgcggcgc agggccagac	11160
actaacctgga cctggaggag ggcgaggggcc tggcgcggct gggggcgccg tctcccgagc	11220
gccaccccgcg ggtcagactg aagcgcgact cgcgcgaggc gtacgtgcct cggcagaacc	11280
tgttcaggga ccgcgcgggc gaggagcccg aggagatgcg ggacaggagg ttcagcgcag	11340
ggcgggagct gcggcagggg ctgaaccgcg agcggctgct gcgcgaggag gactttgagc	11400
ccgacgcgcg gacggggatc agccccgcgc gcgcgcacgt ggcggccgcc gacctggta	11460
cggcgtacga gcagacggtg aaccaggaga tcaacttcca aaagagttc aacaaccacg	11520
tgcgcacgct ggtggcgcgc gaggaggtga ccatggct gatgcacctg tggactttg	11580
taagcgcgct ggtcagaac cccaacagca agcctctgac gcgcagctg ttccgtatag	11640
tgcagcacag cagggacaac gaggcgtaa gggacgcgct gctgaacatc accgagcccg	11700
agggtcggtg gctgctggac ctgattaaca tcctgcagag catagtggtg caggagcgc	11760
gcctgagcct ggccgacaag gtggcggcca tcaactactc gatgctgagc ctggcaagt	11820
tttacgcgcg caagatctac cagacgcgt acgtccccat agacaaggag gtgaagatcg	11880
acggtttta catgcgcattt gcgcgtgaagg tgctcaccct gagcgacgc acggcgtgt	11940
accgcaacga gcgcattccac aaggccgtga gcgtgagccg gcggcgcgag ctgagcgc	12000
gcgagctgat gcacagcctg cagcggcgc tggcggcgc cggcagcggc gacagggagg	12060
cggagtccta cttcgatgcg gggcggacc tgcgcggc gcccagccgg cggccctgg	12120
aggccgcggg ggtccgcgag gactatgacg aggacggcga ggaggatgag gagtacgagc	12180
tagaggaggg cgagtacctg gactaaaccg cgggtgggt ttccgtatgc tgcaagaccc	12240
gaacgtggtg gaccggcgc tgcggcggc tctcagagc cagccgtccg gccttaactc	12300
ctcagacgac tggcgcacagg tcatggaccg catcatgtcg ctgacggcgc gtaacccgga	12360
cgcgttccgg cagcagccgc aggccaacag gctctccgcc atcctggagg cgggtggcgc	12420
tgcgcgcctg aaccccacgc acgagaaggt gctggccata gtgaacgcgc tggccgagaa	12480
cagggccatc cgccggacg aggccgggt ggtgtacgac gcgcgtgc acgcgtggc	12540
ccgctacaac agcggcaacg tgcagaccaa cctggaccgg ctggggggg acgtgcgcga	12600
ggcggtggcg cagcgcgagc gcgcggatcg gcaggcaac ctgggctcca tgggtggcg	12660
aatgccttc ctgagcgcgc agccggccaa cgtgcgcgg gggcaggaag actacaccaa	12720
ctttgtgagc gcgcgtgcgc tgatggtgac cgagacccc cagagcggagg tgtaccagtc	12780

[0023]

ggccccggac tacttcttcc agaccagcag acagggcctg cagacggta acctgagcca	12840
ggcttcaag aacctgcggg ggctgtggg cgtgaaggcg cccaccggcg accgggcac	12900
ggtgtccagc ctgctgacgc ccaactcgcg cctgctgctg ctgctgatcg cgccgttcac	12960
ggacagcggc agcgtgtccc gggacaccta cctggggcac ctgctgaccc tgtaccgcga	13020
ggccatcggg caggcgcagg tggacgagca cacttccag gagatcacca gcgtgagccg	13080
cgcgctgggg caggaggaca cgagcgcct 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 tgcatcgccg ggcggccgtg aaccccgagt actttaccaa	13380
cgcgcctg aacccgcact ggctccgc gcccgggtt tacagcgggg gcttcgaggt	13440
cccgagacc aacgatggct tcctgtggg cgacatggac gacagcgtgt tctcccgcg	13500
gccgcaggcg ctggcggaaag cgccctgtc gcgtcccaag aaggaggagg aggaggaggc	13560
gagtcgcgc cgccgcagca gcggcgtggc ttctctgtcc gagctgggg cgccagccgc	13620
cgcgcgcacc acccgccctc ggctgctggg cgaggacgag tacctaata actccctgtc	13680
cgagccggtg cgggagaaaa acctgcctcc cgcctcccc aacaacggga tagagagcct	13740
ggtgacaag atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcctgcgt	13800
ccggccgccc acgcggcgcc aggccacga cggcagcgg gggctgggt gggatgacga	13860
ggactccgcg gacgatacgca gcgtgctgga cctgggaggg agcggcaacc cggtcgcc	13920
cctgcgcacc cgcctgggaa ggatgtttt aaaaaaaaaaaa aaaaaagcaa gaagcatgat	13980
gcaaaaatta aataaaactc accaaggcca tggcgaccga gcgttggttt cttgtttcc	14040
cttcagtagt cggcgccgg cgatgtacca ggagggaccc cctccctttt acgagagcgt	14100
ggtgccgcgcg gcggcggcg cgcctcttc tcccttgcg tcgcagctgc tggagccgc	14160
gtacgtgcct cgcgcctacc tgccgcctac ggggggaga aacagcatcc gttactcgga	14220
gctggcgccc ctgttcgaca ccacccgggt gtacctggtg gacaacaagt cggcggacgt	14280
ggcctccctg aactaccaga acgaccacag caattttg accacggta tccagaacaa	14340
tgactacagc ccgagcggagg ccagcaccca gaccatcaat ctggatgacc ggtcgactg	14400
ggcgccgac ctgaaaacca tcctgcacac caacatgccc aacgtgaacg agttcatgtt	14460
	14520

[0024]

caccaataag ttcaaggcgc gggtgatggt gtcgcgctcg cacaccaagg aagaccgggt	14580
ggagctgaag tacgagtggg tggagttcga gctgccagag ggcaactact ccgagaccat	14640
gaccattgac ctgatgaaca acgcgatcgt ggagcactat ctgaaagtgg gcaggcagaa	14700
cgggtcctg gagagcgaca tcgggtcaa gttcgacacc aggaacttcc gcctgggct	14760
ggaccccgta accgggctgg ttatgcccgg ggtgtacacc aacgaggcct tccatcccga	14820
catcatcctg ctgcccggct gcgggggtgga cttcaattac agccgcctga gcaacctcct	14880
ggcattccgc aagcggcagc cttccagga gggcttcagg atcacctacg aggacctgga	14940
ggggggcaac atccccgcgc tcctcgatgt ggaggcctac caggatagct tgaaggaaaa	15000
tgaggcggga caggaggata ccgcggccgc cgcctccgccc gccgcccggc agggcgagga	15060
tgctgctgac accgcggccg cggacggggc agaggccgac cccgctatgg tggtggaggc	15120
tcccgagcag gaggaggaca tgaatgacag tgcggtgccg ggagacacct tcgtcaccgg	15180
gggggaggaa aagcaagcgg aggccgaggc cgccggcggag gaaaagcaac tggccggcagc	15240
agccggcggcgc gcggcggtgg ccgcggcgga ggctgagtct gaggggacca agccgcggaa	15300
ggagcccgta attaagcccc tgaccgaaga tagcaagaag cgcatgtaca acctgctcaa	15360
ggacagcacc aacaccgcgt accgcagctg gtacctggcc tacaactacg ggcgcggcgtc	15420
gacgggggtg cgctcctgga ccctgctgtg cacgcggac gtgacctgctg gctggagca	15480
ggtgtactgg tcgctgccc acatgatgca agaccggta accttccgct ccacgcggca	15540
ggtcagcaac ttcccggtgg tggcgccga gctgctgccc gtgcactcca agagttcta	15600
caacgaccag gccgtctact cccagctcat ccgcggatcc acctctctga cccacgtgtt	15660
caatcgctt cctgagaacc agattctggc gcgcggccc gccccacca tcaccaccgt	15720
cagtaaaaac gttcctgctc tcacagatca cgggacgcta ccgcgtgcga acagcatcggt	15780
aggagtccag cgagtgaccg ttactgacgc cagacgcgc acctgccc acgtttacaa	15840
ggccttggc atagtctcgc cgccgtcct ttccagccgc actttttgag caacaccacc	15900
atcatgtcca tcctgatctc acccagcaat aactccggct gggactgct gcgcgcgc	15960
agcaagatgt tcggagggc gaggaagcgt tccgagcagc accccgtcg cgtgcgcggg	16020
cacttccgcg cccccctgggg agccgcacaaa cgcggccgcg cggggcgcac caccgtggac	16080
gacgccatcg actcggtggt ggagcaggcg cgcaactaca gccccgcgtt ctctaccgt	16140
gacgcggcca tccagaccgt ggtgcggggc gcgcggcggt acgccaagct gaagagccgc	16200

[0025]

cggaagcgcg tggcccggccg ccaccgcccgc cgaccgggg ccgcccggccaa acgcgcggcc	16260
gcggccctgc ttgcggggc caagcgacg ggccgcccgc cgccatgag ggccgcgcgc	16320
cgcttggccg ccggcatcac cgccgcccacc atggcccccgtacccgaag acgcgcggcc	16380
gccggccggccg ccggcccat cagtacatg gccagcaggc gccggggcaa cgtgtactgg	16440
gtgcgcact cggtgaccgg cacgcgcgtg cccgtgcgt tccggccccc gcggacttga	16500
gatgatgtga aaaaacaaca ctgagtcctcc tgctgttgtg tgtatccag cggcggccgc	16560
gcgcgacgcg tcatgtccaa gcgcggaaatc aaagaagaga tgctccaggt cgtcgcggcc	16620
gagatctatg ggcccccggaa gaaggaagag caggattcga agcccccggaa gataaagcgg	16680
gtcaaaaaga aaaagaaaaga ttagtgcgtat ggcgtatgggg aggtggagtt cctgcgcgc	16740
acggcgccca ggcccgggt gcagtggaaag ggccggcgcg taaagcgcgt cctgcgcgc	16800
ggcacccggg tggcttcac gcccggcgag cgctccaccc ggacttcaa gcgcgtctat	16860
gacgagggtgt acggcgacga agacctgctg gagcaggcca acgagcgtt cggagagttt	16920
gcttacggga agcgtcagcg ggcgtgggg aaggaggacc tgctggcgct gccgctggac	16980
cagggcaacc ccaccccgat tctgaagccc gtgaccctgc agcaggtgct gccgagcagc	17040
gcaccctccg aggcgaagcg gggctgaag cgcgaggcg gcgacctggc gcccaccgtg	17100
cagctcatgg tgcccaagcg gcagaggctg gaggatgtgc tggagaaaat gaaagttagac	17160
cccggtctgc agccggacat cagggtccgc cccatcaagc agtggcgcc gggctcgcc	17220
gtgcagaccc tggacgtggt catccccacc ggcaactccc ccggccggcc caccactacc	17280
gctgcctcca cggacatgga gacacagacc gatcccggc cagccgcagc cgcagccgc	17340
gccgcgaccc ctcggcgga ggtgcagacg gaccctggc tgccggccgc gatgtcagct	17400
cccccggcgc gtcggggcg caggaagtac ggccggccca acgcgcctcct gcccggatc	17460
gccttgcata cttccatcgc gcccacccccc ggctaccgag gctataccta ccgcggcgaa	17520
agagccaagg gttccaccccg ccgtccccgc cgacgcgcgc ccggccaccac ccggccggc	17580
cgccgcagac gccagccgc actggctcca gtctccgtga ggaaagtggc gcgcgacgga	17640
cacaccctgg tgctggccag ggcgcgtac cacccagca tcgtttaaaa gcctgttgt	17700
gttcttgcag atatggccct cacttgcgc ctcgtttcc cggtggccggg ataccgagga	17760
ggaagatcgc gccgcaggag gggctggcc ggccggccgc tgagcggagg cagccggcgc	17820
gcccacccggc ggcgacgcgc caccagccga cgcatgcgcgc gcgggggtgt gcccgttta	17880
atccccctga tcggccggc gatcgccgcgtt gtcggccggaa tcgcctccgt ggccttgaa	17940

[0026]

gcgtcccaga ggcattgaca gacttgcaaa cttgcaaata tggaaaaaaaaaacccaaata	18000
aaaaagtcta gactctcacg ctcgcttggc cctgtgacta tttttagaa tggaagacat	18060
caactttgcg tcgctggccc cgcgtcacgg ctcgcgcgg ttcctggac actggaacga	18120
tatcggcacc agcaacatga gcggtggcgc cttcagttgg ggctctctgt ggagcggcat	18180
taaaagtatac gggctgccc ttaaaaattt cggctccgg gcctggaaca gcagcacggg	18240
ccagatgtt agagacaagt taaaagagca gaacttccag cagaagggtgg tggagggcct	18300
ggcctccggc atcaacgggg tggggaccc gccaaccag ggcgtgcaga ataagatcaa	18360
cagcagactg gaccccccggc cgccgggtgga ggaggtgccc ccggcgctgg agacgggtgc	18420
ccccgatggg cgtggcgaga agcgcccgcg gcccataagg gaagagacca ctctggtcac	18480
cgagaccat gagccgcccc cgtatgagga ggcctgaag caaggtctgc ccaccacgcg	18540
gcccatcgcg cccatggcca cccgggtgg tggccggcac accccggcca cgctggactt	18600
gcctccgccc gccgatgtgc cgccggcagca gaaggcggca cagccgggccc cgcccgac	18660
cgcctccgt tcctccgccc gtcctctgca ccgcgggccc agcggcccccc ggggggggt	18720
cgcgaggcac ggcaactggc agagcacgct gaacagcatc gtgggtctgg gggtgcggc	18780
cgtgaagcgc cgccgatgct actgaatagc ttagctaact tttgtatgt gtgtatgcgc	18840
cctatgtcgc cgccagagga gctgctgagt cgcggccgtt cgccggccca ccaccaccgc	18900
cactccgccc ctcaagatgg cgacccatc gatgatgccc cagtggctgt acatgcacat	18960
ctcggggccag gacgcctcgg agtacctgag cccgggctg gtgcagttcg cccggccac	19020
cgagagctac ttccgcctga gtaacaagtt taggaacccc acggtggcgc ccacgcacga	19080
tgtgaccacc gaccggtctc agcgccctgac gctgcgggatc attccgtgg accgcgagga	19140
caccgcgtac tcgtacaagg cgccgttcac cctggccgtg ggcgacaacc gcgtgctgga	19200
catggcctcc acctactttg acatccgcgg ggtgctggac cgggggtccca cttcaagcc	19260
ctactctggc accgcctaca actccctggc ccccaaggac gctcccaact cctgcgagtg	19320
ggagcaagag gaaactcagg cagttgaaga agcagcagaa gaggaagaag aagatgtga	19380
cggtaagct gaggaagagc aagcagctac caaaaagact catgtatatg ctcaggctcc	19440
cctttctggc gaaaaattt gtaaaagatgg tctgcaaata ggaacggacg ctacagctac	19500
agaacaaaaaa cctattttatg cagaccctac attccagccc gaacccaaa tcggggagtc	19560
ccagtggaat gaggcagatg ctacagtcgc cggcggtaga gtgctaaaga aatctactcc	19620

[0027]

catgaaacca	tgctatggtt	cctatgcaag	acccacaaaat	gctaatggag	gtcagggtgt	19680
actaacggca	aatgcccagg	gacagctaga	atctcagggtt	gaaatgcaat	tctttcaac	19740
ttctgaaaac	gcccgttaacg	aggctaaca	cattcagccc	aaattggtgc	tgtatagtga	19800
ggatgtgcac	atggagaccc	cggatacgca	cctttcttac	aagccgc当地	aaagcgatga	19860
caattcaaaa	atcatgctgg	gtcagcagtc	catgccc当地	agacctaatt	acatcggtt	19920
cagagacaac	tttatcggcc	tcatgttata	caatagcact	ggcaacatgg	gagtgc当地	19980
aggcaggcc	tctcagttga	atgcagtggt	ggacttgcaa	gacagaaaca	cagaactgtc	20040
ctaccagctc	ttgcttgatt	ccatgggtga	cagaaccaga	tactttcca	tgtggaatca	20100
ggcagtggac	agttatgacc	cagatgttag	aattattgaa	aatcatggaa	ctgaagacga	20160
gctccccaaac	tattgttcc	ctctgggtgg	catagggta	actgacactt	accaggctgt	20220
taaaaccaac	aatggcaata	acgggggcca	ggtgacttgg	acaaaagatg	aaactttgc	20280
agatcgcaat	gaaatagggg	tggaaaccaa	tttcgctatg	gagatcaacc	tcagtgccaa	20340
cctgtggaga	aacttcctgt	actccaacgt	ggcgctgtac	ctaccagaca	agcttaagta	20400
caaccctcc	aatgtggaca	tctctgacaa	ccccaacacc	tacgattaca	tgaacaagcg	20460
agtggtggcc	ccggggctgg	tggactgcta	catcaacctg	ggcgcgcgct	ggtcgctgga	20520
ctacatggac	aacgtcaacc	cctcaacca	ccaccgcaat	gcgggcetgc	gctaccgctc	20580
catgtctcg	ggcaacgggc	gtacgtgcc	cttccacatc	caggtgc当地	agaagttt	20640
tgccatcaag	aacccctcc	tcctgccggg	ctcctacacc	tacgagtggaa	acttcaggaa	20700
ggatgtcaac	atggcctcc	agagctctct	ggtaacgat	ctcagggtgg	acggggccag	20760
catcaagttc	gagagcatct	gcctctacgc	caccccttc	ccatggccc	acaacacggc	20820
ctccacgctc	gaggccatgc	ttaggaacga	caccaacgac	cagtcctca	atgactaccc	20880
ctccggcc	aacatgtct	acccatacc	cggcaacgccc	accaacgtcc	ccatctccat	20940
ccctctcg	aactggcg	cctcccg	ctggccttc	acccgctca	agaccaagga	21000
gaccctcc	ctggcctcg	gattcgaccc	ctactacacc	tactcgggct	ccatcccta	21060
cctggacggc	accccttacc	tcaaccacac	tttcaagaag	gtctcggtca	ccttcgactc	21120
ctcggtcagc	tggccggca	acgaccgtct	gtcaccccc	aacgagttcg	agatcaagcg	21180
ctcggtcagc	ggggagggct	acaacgtggc	ccagtgc当地	atgaccaagg	actggttcct	21240
gttccagatg	ctggccaact	acaacatcg	ctaccaggc	ttctacatcc	cagagagcta	21300
caaggacagg	atgtactct	tcttcaggaa	cttccagccc	atgagccggc	aggtggtgaa	21360

[0028]

ccagaccaag	tacaaggact	accaggaggt	gggcatcatc	caccagcaca	acaactcggt	21420
cttcgtggc	tacctcgccc	ccaccatgctg	cgagggacag	gcctaccccg	ccaacttccc	21480
ctatccgctc	ataggcaaga	ccgcggtcga	cagcatcacc	cagaaaaagt	tcctctgcga	21540
ccgcaccctc	tggcgcatcc	ccttctccag	caacttcatg	tccatgggtg	cgctctcgga	21600
cctgggccag	aacttgctct	acgccaactc	cgcacacgcc	ctcgacatga	cttcgaggt	21660
cgacccccatg	gacgagccca	cccttctcta	tgttctgttc	gaagtcttg	acgtggtccg	21720
ggtccaccag	ccgcaccgcg	gcgtcatcga	gaccgtgtac	ctgcgtacgc	cttctcgcc	21780
cggcaacgcc	accacctaataa	gaagcaagcc	gcagtcatcg	ccgcctgcat	gccgtcggt	21840
tccaccgagc	aagagctcag	ggccatcgctc	agagacctgg	gatgcgggcc	ctatttttg	21900
ggcaccttcg	acaagcgctt	ccctggcttt	gtctccccac	acaagctggc	ctgcgccatc	21960
gtcaacacgg	ccggccgcga	gaccgggggc	gtgcactggc	tggccttcgc	ctgaaacccg	22020
cgctccaaaa	catgcttcct	cttgaccccc	ttcggctttt	cggaccagcg	gctcaagcaa	22080
atctacgagt	tcgagttacga	gggcttgctg	cgtcgacagcg	ccatcgccctc	ctcgcccgac	22140
cgctgcgtca	ccctcgaaaa	gtccacccag	accgtgcagg	ggcccgactc	ggccgcctgc	22200
ggtctttct	gctgcatttt	tctgcacgcc	tttgcact	ggcctcagag	tcccatggac	22260
cgcacccca	ccatgaactt	gctgacgggg	gtgccaact	ccatgtccca	gagccccag	22320
gtcgagccca	ccctgcgcgg	caaccaggag	cagctctaca	gcttcctgga	gcccactcg	22380
ccttacttcc	gccgccacag	cgcacagatc	aggagggcca	cctccttctg	ccacttgcaa	22440
gagatgcaag	aagggttaata	acgatgtaca	cactttttt	ctcaataaaat	ggcatctttt	22500
tatttataca	agctctctgg	ggtattcatt	tcccaccacc	acccggcgtt	gtgccatct	22560
ggctctattt	agaaatcgaa	agggttctgc	cgggagtcgc	cgtgcgccac	ggcagggac	22620
acgttgcgtat	actggtagcg	ggtgccccac	ttgaactcgg	gcaccaccag	gcgaggcagc	22680
tcgggaaat	tttcgctcca	caggctgcgg	gtcagcacca	gcmcgttcat	caggtcggc	22740
gccgagatct	tgaagtcgca	gttggggccg	ccgcctgctg	cgcgcgagtt	gcgg tacacc	22800
gggttgcagc	actggaacac	caacagcgcc	gggtgcttca	cgctggccag	cacgctgcgg	22860
tcggagatca	gctcggcg	caggtcctcc	gcgttgctca	gcmcgaacgg	ggtcatcttgc	22920
ggcacttgcc	gccccaggaa	gggcgcgtgc	cccggtttcg	agttgcagtc	gcagcgcagc	22980
gggatcagca	ggtgcggcgtg	cccgactcg	gcgttgggt	acagcgcgcg	catgaaggcc	23040

[0029]

tgcatctggc ggaaggccat ctgggccttg ggcgcctccg agaagaacat gccgcaggac	23100
ttgcccgaga actggtttc gggcagctg gcgtcgatca ggcagcagcg cgcgtcggtg	23160
ttggcgatct gcaccacgtt gcgcacccac cggttctca cgatcttgc cttggacgat	23220
tgctccttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttcacca tgctgtgcc gtgcagacac ttcaagtcgc cctccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtggcgtcg aaagacttgt aggtcacctc cgcaaggac	23400
tgcaggtacc cctgcaaaaaa gcggcccatc atggtcacga aggtcttgtt gctgctgaag	23460
gtcagctgca gcccgccgtg ctcctcggtc agccaggctt tgcacacggc cgccagcgcc	23520
tccacctggc cggcagcat cttgaagttc accttcagct cattctccac gtggacttg	23580
tccatcagcg tgcgcgcgc ctcatgcggc ttctccagg ccgacaccag cggcaggctc	23640
acggggttct tcaccatcac cgtggccgac gcctccgcg cgcttcgtt ttccgcggcg	23700
ctgttcttt ccttttcctc ctttcctcg ccggccggcc ctgcagccc ccgcaccacg	23760
gggtcgttctt cctgcaggcg ctgcacccgt cgcttgcgt tgcgcgcgtt cttgatgcgc	23820
acggcggt tgctgaagcc caccatcacc agcgccgcctt cttttgcgc tgcctcgctg	23880
tccagaatga cctccgggaa gggggggttt gtcatcctca gtaccgaggc acgcttctt	23940
ttcttcctgg gggcggtcg cagctccgcg gctgcggccg ctggcagggt cgaaggccga	24000
gggctggcgc tgcgcggcac cagcgcgtcc tgcgagccgt ctcgtcctc ctggactcg	24060
agacggaggc gggccgcgtt ctgcggggc gcgcggggcg gggaggcgg cggcggcgac	24120
ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgc gcgtccgcgc	24180
tcgggggtgg tctcgcgctg gtcctttcc cgactggcca tctccactg ctcttcctcc	24240
tataggcaga aagagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accgcgcctt ctgagccctc caccacgcgc gcacccacgg ccaatgcgcgc cgccggac	24360
gcgcgcaccc agaccacccgc cagtaccacc ctccccagcg acgcacccccc gtcgagaat	24420
gaagtgcgtga tcgagcagga cccgggtttt gtgagcggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggctgc cgcctcgtt ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata aggtgagac agcagtcggc cgggggaacg gaagccatga tgctgatgac	24600
ggctacccat acgtgggaga cgacgtcgctg cttaaggacc tgcaccgcgc gtgcgtcatc	24660
gtctcgacg cgctgcagga gcgcgtcgaa gtgcgcgtt acgtggcgga ggtcagccgc	24720
gcctacgagc ggcacccctt cgcgcgcac gtgcgcgcac aggcggggaa gaacggcacc	24780

[0030]

tgcgagcca acccgctct caacttctac ccgttctcg cggtacccga ggtgctggcc	24840
acctaccaca tcttttcca aaactgcaag atccccctct cctgccgcgca aaccgcacc	24900
cgcgccgaca aaaccctgac cctgcggcag ggcgcaca tacctgatat cgcctctg	24960
gaggaagtgc ccaagatctt cgagggtctc ggtcgacg agaaacggc ggcgaacgct	25020
ctgcacggag acagcgaaaa cgagagtac tcgggggtgc tggggagct cgagggcgac	25080
aacgcgcgc tggccgtact caagcgcagc atagaggtca cccacttgc ctaccggcg	25140
ctcaacctgc ccccaaggt catgagtgtg gtcattggcg agctcatcat ggcgcgcgc	25200
cagccctgg ccggatgc aaacttgcaa gagtcctccg aggaaggcct gcccggc	25260
agcgacgagc agctggcgcg ctggctggag acccgaccc cgcgcagct ggaggagcgg	25320
cgcaagctca tggccgc ggtgctggc accgtggagc tcgagtgtct gcagcgcttc	25380
ttcgcggacc ccgagatgca ggcgaagctc gaggagaccc tgcactacac cttccgcag	25440
ggctacgtgc gccaggcctg caagatctcc aacgtggagc tctgcaacct ggtctctac	25500
ctggcatcc tgcacgagaa ccgcctcgaa cagaacgtcc tgcactccac cctcaaaggg	25560
gaggcgcgc gcgactacat ccgcgactgc gcctaccttc tcctctgcta cacctggcag	25620
acggccatgg gggtctggca gcagtgcctg gaggagcga acctaagga gctggaaaag	25680
ctcctaagc gcaccctcag ggacctctgg acgggcttca acgagcgctc ggtggccgc	25740
gcgcgtggc acatcatctt tcccgagcgc ctgctcaaga ccctgcagca gggcctgccc	25800
gacttcacca gccagagcat gctgcagaac ttcaggactt tcatcctgga ggcgcggc	25860
atcctgcgg ccacttgctg cgcgtgccc agcgacttcg tgcccatcaa gtacagggag	25920
tgccgcgcg cgtctgggg ccactgtac ctttccagc tggccaacta cctgcctac	25980
cactcgacc tcatggaaga cgtgagcggc gagggctgc tcgagtgcac ctgcgcgtc	26040
aacctctgca cgccccaccg ctctctagtc tgcaacccgc agctgctcag cgagagtcag	26100
attatcggtt cttcgagct gcagggtccc tgcctgacg agaagtccgc ggctccagg	26160
ctgaaactca ctccgggct gtggacttcc gcctacccac gcaaatttgc acctgaggac	26220
taccacgccc acgagatcag gttctacgaa gaccaatccc gcccgcacaa ggcggagctc	26280
accgcctgcg tcatcacccca gggcacatc ctggccaat tgcaagccat caacaaagcc	26340
cgcgcgagat tttgctgaa aaagggtcgg ggggtgtacc tggacccca gtccggcag	26400
gagctaaacc cgctacccccc gccgcgcgc cagcagcggg accttgcttc ccaggatggc	26460

[0031]

acccagaaaag aagcagcagc cgccgcccgc gcccgcagcca tacatgcgtt tggaggaaga	26520
ggaggaggac tggacagtc aggcagagga gtttggac gaggagcagg aggagatgt	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgcaaca ccatgcctt cggcgcagc cccctgcgg gggccctga aatcctccga	26700
acccagcacc agcgctataa cctccgcctt tccggcgccg gcgcacccg cccgcagacc	26760
caaccgtaga tggacacca caggaaccgg gtcggtaag tccaagtgcc cgccgcccgc	26820
accgcagcag cagcagcagc agcgcaggc ctaccgcgt tggcgccggc acaagaacgc	26880
catagtcgcc tgcttgcaag actgcggggg caacatctt ttgcggccgc gttcctgt	26940
attccaccac ggggtcgcc ttcccgcaaa tgtcctgcat tactaccgtc atctctacag	27000
ccctactgc agcggcgacc cagaggcggc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatac ctccgcggc aagacagcgg cagcagcggc caggagaccc gcggcagcag	27120
cggcgggagc ggtggcgca ctgcgcctt cgcccaacga acccctctcg acccgggagc	27180
tcagacacag gatttcccc actttgtatg ccatcttcca acagagcaga gcccaggagc	27240
aggagctgaa aataaaaaac agatctctgc gtccttcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagttcgg cgcacgctgg aggacgcggg ggcactttc agcaaatact	27360
gcgcgctcac tcttaaagac tagtcccgcc cccttctcg atttaggcgg gagaaaacta	27420
cgtcatcgcc ggccgcgcgc cagccgcgc agccgagatg agcaaagaga ttcccacgcc	27480
atacatgtgg agtaccaggc cgcaaatggg actcgccggc ggagcggccc aggactactc	27540
cacccgcattt aactacatga gcgcgggacc ccacatgtac tcacaggta acggatccg	27600
cgcccaagcga aaccaaatac tgcttggaca ggcggccatc accgcacgc cccgcataaa	27660
tctcaacccc cgaaattggc cgcgcgcctt cgtgtaccag gaaacccctt cgcgcaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcg gggcgcagct	27780
cgcggcggc ttctcgacg gggcgccgc gtcggacca ggtataagac acctgtatgt	27840
cagaggccga ggtatccagc tcaacgcga gtcgggtggc ttctcgctcg gtctccgtcc	27900
ggacggaaact ttccagctcg ccggatccgg ccgccttcg ttacgcggcc gccaggcgta	27960
cctgactctg cagacctcg cctcgaggcc ccgcgcggc ggcacgcggaa ccctccagg	28020
cgtggaggag ttctcgccct cggctactt caaccccttc tcggacccctc ccggacgcata	28080
cccccggccag ttcatccga actttgacgc ggtgaaggac tcggcgacgc gctacgactg	28140
aatgtcaggcgt gtcgaggcag agcagctcg cctgagacac ctcgagcact gcccgcgc	28200

[0032]

caagtgcctc	gcccgcggtt	ctggtgagtt	ctgctacttt	cagctaccgg	aggagcatac	28260
cgagggggccg	gcmcacggcg	tccgcctgac	cacccagggc	gaggttacct	gttccctcat	28320
ccgggagttt	accctccgtc	ccctgctagt	ggagcgggag	cggggccct	gtgtcctaacc	28380
tatcgccctgc	aactgcccta	accctggatt	acatcaagat	ctttgctgtc	atctctgtgc	28440
ttagtttaat	aaacgctgag	atcagaatct	actggggctc	ctgtcgccat	cctgtgaacg	28500
ccaccgtctt	cacccacccc	gaccaggccc	aggcgaacct	cacctgcggt	ctgcacatcgga	28560
gggccaagaa	gtacctcacc	tggtaattca	acggcacccc	ctttgtggtt	tacaacagct	28620
tcgacgggga	cggagtcctcc	ctgaaagacc	agcttcgg	tctcagctac	tccatccaca	28680
agaacaccac	cctccaaactc	ttccctccct	acctgccggg	aacctacgag	tgcgtcaccg	28740
gccgctgcac	ccacccatcacc	cgcctgatcg	taaaccagag	ctttccggg	acagataact	28800
ccctcttccc	cagaacagga	ggtgagctca	ggaaactccc	cggggaccag	ggcggagacg	28860
tacccctcgac	ccttgcgggg	ttaggatttt	ttattaccgg	gttgcggct	cttttaatca	28920
aagtttcctt	gagattgtt	cttccttct	acgtgtatga	acacctcaac	ctccaataaac	28980
tctacccttt	cttcggaaatc	aggtgacttc	tctgaaatcg	ggcttgggt	gctgcttact	29040
ctgttgattt	ttttccttat	catactcagc	cttctgtgcc	tcaggctcgc	cgcctgctgc	29100
gcacacatct	atatctactg	ctggttgctc	aagtgcaggg	gtcgccaccc	aagatgaaca	29160
ggtacatgg	cctatcgatc	ctaggcctgc	tggccctggc	ggcctgcagc	gccccaaaa	29220
aagagattac	ctttgaggag	cccgcttgca	atgtaacttt	caagcccgg	ggtgaccaat	29280
gcaccaccct	cgtcaaatgc	gttaccaatc	atgagaggct	gcmcacatcgac	tacaaaaaca	29340
aaactggcca	gtttgcggc	tatagtgtgt	ttaccccccgg	agaccctct	aactactctg	29400
tcaccgtctt	ccagggcgga	cagtctaaga	tattcaatta	cacttccct	ttttatgagt	29460
tatgcgatgc	ggtcatgtac	atgtcaaaac	agtacaacct	gtggcctccc	tctccccagg	29520
cgtgtgtgga	aaatactggg	tcttactgct	gtatggctt	cgcaatcact	acgctcgctc	29580
taatctgcac	ggtgctatac	ataaaattca	ggcagaggcg	aatctttatc	gataaaaaga	29640
aaatgccttg	atcgctaaca	ccggcttct	atctgcagaa	tgaatgcaat	cacccctcta	29700
ctaataccacca	ccaccctcct	tgcgattgcc	catgggttga	cacgaatcga	agtgccagtg	29760
gggtccaaatg	tcaccatgg	gggccccggc	ggcaattcca	ccctcatgtg	ggaaaaattt	29820
gtccgcaatc	aatgggttca	tttctgtct	aaccgaatca	gtatcaagcc	cagagccatc	29880

[0033]

tgcatggc aaaatcta ac tctgatcaat gtgcaa atga tggatgctgg gtactattac	29940
ggcagcgg gagaatcat taattactgg cgacccaca aggactacat gctgcatgt	30000
gtcaggcac ttcccactac caccactacc accaccttc ccaccaccac caccactact	30060
actactacta ctactactac tactactacc actaccgctg cccgcatac ccgcaaaagc	30120
accatgatta gcacaaagcc ccctcggtc cactcccacg ccggcgggccc catcggtcg	30180
acccagaaa ccaccgagct ttgcttctgc caatgcacta acgccagcgc tcatgaactg	30240
ttcgacctgg agaatgagga tgtccagcag agctccgctt gcctgaccca ggaggctgt	30300
gagcccgttg ccctgaagca gatecggtat tcaataattg acttttttc ttttgcact	30360
cccgaaatacc ctcccgattc tactttccac atcacggta ccaaagaccc taacctct	30420
tttacactga tgctgctgct ctgtatctct gtggctctt ccgcgtatgt gttactgggg	30480
atgttctgct gcctgatctg ccgcagaaag agaaaagctc gctctcaggc ccaaccactg	30540
atgccttcc cttacccccc ggattttgca gataacaaga tatgagctcg ctgctgacac	30600
taaccgctt actagcctgc gctctaacc ttgtcgctt cgaactcgaga ttccacaatg	30660
tcacagctgt ggcaggagaa aatgttactt tcaactccac ggccgatacc cagtggctgt	30720
ggagtggctc aggtagctac ttaactatct gcaatagctc cacttccccc gcatatccc	30780
caaccaagta ccaatgcaat gccagcctgt tcaccctcat caacgcttcc accctggaca	30840
atggactcta tgtaggctat gtaccctttt gttggcaagg aaagacccac gcttacaacc	30900
tggaaagttcg ccagcccaga accactaccc aagcttctcc caccaccacc accaccacca	30960
ccatcaccag cagcagcagc agcagcagcc acagcagcag cagcagatta ttgactttgg	31020
ttttggccag ctcatctgcc gctacccagg ccatctacag ctctgtgcc gaaaccactc	31080
agatccaccc cccagaaacg accaccccca ccaccctaca cacccctcagc gatcagatgc	31140
cgaccaacat caccccttg gctttcaaa tggacttac aagcccccact ccaaaccag	31200
tggatgcccgc cgaggtctcc gcccctgtca atgactggc gggctgggaa atgtgggt	31260
tcgccatagg catgatggcg ctctgcctgc ttctgctctg gctcatctgc tgcctccacc	31320
cgaggcggc cagacccccc atctatagac ccatcattgt cctgaacccc gataatgatg	31380
ggatccatag attggatggc ctgaaaaacc tactttttc ttttacagta tgataaattg	31440
agacatgcct cgcatttct tgtacatgtt cttctccca cttttctgg ggtgttctac	31500
gctggccgct gtgtctcacc tggaggtaga ctgcctctca cccttcactg tctacctgct	31560
ttacggattt gtcaccctca ctctcatctg cagcctaattc acagtaatca tcgccttcat	31620

[0034]

ccagtgcatt gattacatct gtgtgcgcct cgcatactc agacaccacc cgca	taccg 31680
agacaggaac attgccaac ttctaagact gctctaatac tgcataagac tgtgat	ctgc 31740
cttctgatcc tctgcattcct gcccacccctc acctcctgcc agtacaccac aaaat	tcctcg 31800
cgcaaaaagac atgcctcctg ccgcttcacc caactgtgga atataccaa atgct	acaac 31860
gaaaagagcg agctctccga agcttggctg tatgggtca tctgtgtctt agttt	ctgc 31920
agcaactgtct ttgcctcat aatctacccc tacttgatt tggatggaa cgcgat	cgat 31980
gccatgaatt accccacctt tcccgaccc gagataattc cactgcgaca agttgtac	ccc 32040
gttgcgtta atcaacgccc cccatcccc acgcccactg aaatcagcta cttaac	cta 32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatcagta ccgagc	agc 32160
tctcctagag aggccgcaggc aggccgcgtga gcaagagcgc ctcaatcagg agct	ccgc 32220
tccgttaac ctgcaccagt gcaaaagagg catctttgt ctggtaaagc aggcca	aaagt 32280
cacctacgag aagaccggca acagccaccc cctcagttac aaattgcca cccagc	gcgc 32340
gaagctggc ctcatggcgt gtgagaatcc catcaccgtc acccagcact cggt	agagac 32400
cgaggggtgt ctgcactccc cctgtcgcccc tccagaagac ctctgcaccc tgg	taaagac 32460
cctgtcggt ctcagagatt tagtcccatt taactaatca aacactggaa tcaata	aaaaaa 32520
aatcactta cttaaaatca gacagcagg ctctgtccag tttattcagc agcac	ctct 32580
tccctcctc ccaactctgg tactccaaac gccttctggc ggcaaaacttc ctcc	acacaccc 32640
tgaagggaat gtcagattct tgctcctgtc cctccgcacc cactatctt atgtt	gttgc 32700
agatgaagcg caccaaaacg tctgacgaga gcttcaaccc cgtgtacccc tat	gacacagg 32760
aaagcggccc tccctccgtc ccttcctca cccctccctt cgtgtctccc gatgg	attcc 32820
aagaaaagtcc ccccggggtc ctgtctctga acctggccga gcccctggc actt	ccacacg 32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgctggc aac	ctcacct 32940
ctcaagatat caccaccgt agccctcccc tcaaaaaaac caagaccaac ctc	agcct 33000
aaacctcatc ccccctaact gtgagcacct caggccct caccgtagca gccc	ccgcgtc 33060
ccctggcggt ggcggcacc tccctcacca tgcaatcaga ggcggccctg acag	tacagg 33120
atgcaaaaact caccctggcc accaaaggcc ccctgaccgt gtctgaaggc aa	actggcct 33180
tgcaaacatc ggcggcgtg acggccgctg acagcagcac ctcacagtc agt	gccacac 33240
cacccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc attt	acacca 33300

[0035]

ccaatggaaa actaggactt aactttggcg ctcccctgca tgtggtagac agcctaaatg	33360
cactgactgt agttactggc caaggtctta cgataaacgg aacagcccta caaactagag	33420
tctcagggtgc cctcaactat gacacatcag gaaacctaga attgagagct gcaggggta	33480
tgcgagttga tgcaaatttgt caacttatcc ttgatgtagc ttacccattt gatgcacaaa	33540
acaatctcg ccttaggctt ggacagggac ccctgttgt taactctgcc cacaacttgg	33600
atgttaacta caacagaggc ctctacctgt tcacatctgg aaataccaaa aagctagaag	33660
ttaatatcaa aacagccaag ggtctcattt atgatgacac tgctatagca atcaatgcgg	33720
gtgatgggct acagtttgac tcaggctcag atacaatcc attaaaaact aaacttggat	33780
taggactgga ttatgactcc agcagagcca taattgctaa actggaaact ggcctaagct	33840
ttgacaacac aggtgccatc acagtaggca acaaaaatga tgacaagctt accttggaa	33900
ccacaccaga cccatcccct aactgttagaa tctattcaga gaaagatgct aaattcacac	33960
ttgtttgac taaatgcggc agtcaggtgt tggccagcgt ttctgtttta tctgtaaaag	34020
gtgccttgc gcccattcagt ggcacagtaa ctagtgcctca gattgtcctc agatttgatg	34080
aaaatggagt tctactaagc aattttccc ttgaccctca atactggAAC tacagaaaag	34140
gtgaccttac agagggact gcatatacca acgcagtggg atttatgccc aacctcacag	34200
cataccaaa aacacagagc caaactgcta aaagcaacat tgtaagtcag gtttacttga	34260
atggggacaa atccaaaccc atgaccctca ccattaccct caatggact aatgaaacag	34320
gagatgccac agtaagcact tactccatgt cattctcatg gaactggaaat ggaagtaatt	34380
acattaatga aacgttccaa accaactcct tcaccttctc ctacatcgcc caagaataaa	34440
aagcatgacg ctgttgattt gattcaatgt gtttctgttt tattttcaag cacaacaaaa	34500
tcattcaagt cattttcca tcttagctt atagacacag tagcttaata gacccagtag	34560
tgcaaagccc cattcttagct tataactagt ggagaagtac tcgcctacat gggggtagag	34620
tcataatgt gcatcaggat agggcggtgg tgctgcagca ggcgcgcaat aaactgctgc	34680
cggcccgct ccgtcctgca ggaataacaac atggcagtgg tctcctcagc gatgattcgc	34740
accgccccca gcataaggcg cttgtcctc cgggcacagc agcgcacccct gatctcactt	34800
aaatcagcac agtaactgca gcacagcacc acaatattgt tcaaaatccc acagtgcag	34860
gcgcgttatac caaagctcat ggccgggacc acagaaccca cgtggccatc ataccacaag	34920
cgcaggtaga ttaagtggcg acccctcata aacacgctgg acataaaacat taccttttt	34980
ggcatgttgtt aattcaccac ctcccggtac catataaacc tctgattaaa catggcgcca	35040

[0036]

tccaccacca tcctaaacca gctggccaaa acctgcccgc	35100
cgggactgg aacaatgaca gtggagagcc caggactcgt	35160
aaccatggat catcatgctc gtcatgat	
aat caatgttgc acaacacagg cacacgtca tacacttctt	35220
caggattaca agctcctccc gcgttagaac catatcccag	35280
ggaacaaccc attcctgaat cagcgtaaat cccacactgc	
agggaagacc tcgcacgtaa ctcacgttgc gcattgtcaa	35340
agtgttacat tcggcagca gcggatgatc ctccagttatg	35400
gtagcgcggg tttctgtctc aaaaggaggt agacgatccc	
tactgtacgg agtgcgcga gacaaccgag atcgtttgg	35460
tcgttagtgc atgccaatg gaacgcggga cgtagtcata	35520
tttccctgaag tcttagatct ctcaacgcag caccagcacc	
aacacttcgc agtgtaaaag gccaagtgcc gagagagtat	35580
atataggaat aaaaagtgac gtaaacgggc aaagtccaaa	35640
aaacgcggc aaaaaccgca cgcgaaccta cgccccgaaa	
cgaaagccaa aaaacactag acactccctt cggcgtcaa	35700
cttccgctt cccacgctac gtcacttgcc ccagtcaa	35760
aaactacata tcccgaactt ccaagtcgc acgcccggaaa	
caccgcctac acctccccgc cgccggccc gcccccaaac	35820
ccgcctcccg ccccgcgccc cgccccgcgc cgcccatctc	
attatcatat tggcttcaat ccaaaataag gtatattatt	35880
gatgatggtt taaacggatc caattttga agacgaaagg	35940
gcctcgat acgcctat ttataggtta atgtcatgat	
aataatggtt tcttagacgt caggtggcac tttcgggaa	36000
aatgtgcgcg gaaccctat ttgttatttt ttctaaatac	
attcaaataat gtatccgctc atgagacaat aaccctgata	36060
aatgcttcaa taatattgaa aaaggaagag tatgagtatt	
caacatttcc gtgtcgccct tattccctt tttcgccat	36120
tttgccttcc tggtttgct cacccagaaa cgctggtgaa	
agtaaaagat gctgaagatc agttgggtgc acgagtgggt	36180
tacatcgaac tggatctcaa cagcgtaag atccttgaga	
gtttcgccc cgaagaacgt ttccaatga tgagcacttt	36240
taaagttctg ctatgtggcg cggtattatc ccgtgttgac	
gccggcaag agcaactcgg tcgcccata cactattctc	36300
agaatgactt ggtttagtac tcaccagtca cagaaaagca	
tcttacggat ggcatgacag taagagaatt atgcagtgtc	36360
gccataacca tgagtgataa cactgcggcc aacttacttc	
tgacaacgtat cggaggaccg aaggagctaa ccgctttttt	36420
gcacaacatg ggggatcatg taactcgct tgatcggtgg	
gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	36480
acaccacgat gcctgttagca atggcaacaa cggtgcgcaa	
actattaact ggcgaactac ttactctagc	36540
36600	
tgatcggtgg gaaccggagc tgaatgaagc cataccaaac	
gacgagcgtg acaccacgat gcctgttagca atggcaacaa	36660
cggtgcgcaa actattaact ggcgaactac ttactctagc	
36720	

[0037]

ttccggcaa caattaatag actggatgga ggcggataaa gttgcaggac cacttctgcg	36780
ctcgccctt ccggctggct ggtttattgc tgataaatct ggagccggtg agcgtggtc	36840
tcgcgtatc attgcagcac tggggccaga tggtaagccc tcccgatcg tagttatcta	36900
cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg agataggtgc	36960
ctcaactgatt aagcattggt aactgtcaga ccaagttac tcataatatac tttagattga	37020
tttaaaagga tctaggtgaa gatcctttt gataatctca tgacccaaat cccttaacgt	37080
gagtttcgt tccactgagc gtcagacccc gtagaaaaaga tcaaaggatc ttcttgagat	37140
ccttttttc tgcgcgtaat ctgctgctt gaaacaaaaa aaccaccgct accagcggtg	37200
gtttgttgc cggtcaaga gctaccaact cttttccga aggttaactgg cttcagcaga	37260
gcccagatac caaatactgt cttctagtg tagccgtatg taggccacca cttcaagaac	37320
tctgttagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt	37380
ggcgataagt cgtgtttac cgggttggac tcaagacgt agttaccgga taaggcgcag	37440
cggtcgggct gaacgggggg ttcgtgcaca cagccagct tggagcgaac gacctacacc	37500
gaactgagat acctacagcg ttagctatga gaaagcgcca cgttcccgaa agggagaaag	37560
gcccacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca	37620
ggggaaacg cctggtatct ttatagtcct gtcgggttc gccacctctg acttgagcgt	37680
cgattttgt gatgctcgtc agggggcgg agcctatgga aaaacgccag caacgcggcc	37740
ttttacggt tcctggcctt ttgctggcct tgaagctgtc cctgatggtc gtcatctacc	37800
tgctggaca gcatggctg caacgcggc atcccgatgc cgccggaaagc gagaagaatc	37860
ataatgggaa aggccatcca gcctcgctc gcagatccga attcgtaaa ac	37912

&lt;210&gt; 8

&lt;211&gt; 43428

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 8

catcatcaat aatatacctt atttggatt gaagccaata tgataatgag atggcggcgt	60
cggggcgggg cgccggcgg gaggcgggtt tggggcggg ccggcggcgt gggcgggtgt	120
gcggaaagtgg actttgttaag tgtggcgat gtgacttgct agtgccggc gcggtaaaag	180

[0038]

tgacgttttc cgtgcgcgac aacgcggccg ggaagtgaca ttttccgcg gtttttacc	240
ggatgttgta gtgaatttgg gcgtAACAA gtaagatttgc cccattttcg cgggaaaact	300
gaaacgggaa agtgaatct gattaatttt gcgttagtca taccgcgtaa tatttgccta	360
ggcccgaggc actttggccg attacgtgga ggactcgccc aggtgtttt tgaggtaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gataccat tgcatacg	480
gtatccatatacataatatgt acatTTTAT tggctcatgt ccaacattac cgccatgtt	540
acattgatta ttgacttagtt attaatagta atcaattacg gggcatttag ttcataccc	600
atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgc	660
cgaccccccgc ccattgacgt caataatgac gtatgtccc atagtaacgc caataggac	720
tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtagatca	780
agtgtatcat atgccaagta cgccccat tgacgtcaat gacggtaat gccccgc	840
gcattatgcc cagtagatca ctttatggc ctttcctact tggcagtaca tctacgtatt	900
agtcatcgat attaccatgg tcatgcgtt ttggcagtac atcaatggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccatttgcgtcaatggg gtttgggg	1020
gcaccaaaat caacggact ttccaaaatg tcgtaacaac tccgcggcat tgacgcaa	1080
ggcggtagg cgtgtacggt gggaggtcta tataaggcaga gctctcccta tcagtatag	1140
agatctccct atcagtgata gagatcgatcg acgagctgt ttagtgaacc gtcagatgc	1200
ctggagacgc catccacgct gtttgcactt ccatagaaga caccggacc gatccagc	1260
ccgcggccgg gaacggtgca ttggAACGCG gattcccggt gccaagagt gatctccg	1320
tttatctagg taccgggccc cccctcgagg tcgacggtat cgataagctt cacgctgc	1380
caagcactca gggcgcaagg gctgctaaag gaagcgaaac acgttagaaag ccagtccg	1440
gaaacgggtgc tgacccggta tgaatgtcag ctactggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagttgcag tggcgttaca tggcgtatgc tagactggc	1560
ggtttatgg acagcaagcg aaccggattt gccagctggg gcgcctctg gtaagggtgg	1620
gaagccctgc aaagtaaact ggatggctt cttggccca aggtatgtat ggcgcagg	1680
atcaagatct aaccaggagc tattttatgg caacagttaa ccagctggta cgcaaaaccac	1740
gtgctcgaa agttgcgaaa agcaacgtgc ctgcgttgc agcatgcggc caaaaacgt	1800
cgctatgtac tcgtgtatatacacttccctt ctaaaaaacc gaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtgacttc ctacatcggt ggtgaagg	1920

[0039]

acaacactgca ggagcactcc gtgatcctga tccgtggcg tcgtgttaaa gaccccccgg	1980
gtgttcgtta ccacaccgta cgtggtgcgc ttgactgctc cggcgtaaaa gaccgtaagc	2040
aggctcggttc caagtatggc gtgaagcgctc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcggttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggttatgac tgggcacaac agacaatcg ctgctctgat	2220
gccggccgtgt tccggctgtc agcgcagggg cgcccggttc ttttgcgttcaaa gaccgacactg	2280
tccgggtccc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct ggcacgcacg	2340
ggcggttcctt ggcgcgtgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcccggctg catacgcttg atccggctac ctgcccattc	2520
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctttgtc	2580
gatcaggatg atctggacga agagcatcag gggctcgccgc cagccgaact gttgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgtcgtga cccatggcga tgcctgcttg	2700
ccgaatatca tgggtggaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccgtg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctccgaa ttgcagcgc	2880
atcgccttct atcgccttct tgacgagttc ttctgagcgg gactctgggg ttgcggatga	2940
ccgaccaagc gacgccaac ctgcccac gagattcga ttccaccgccc gccttctatg	3000
aaagggttggg cttcggaatc gtttccggg acgcccgtg gatgatcctc cagcgcgggg	3060
atctcatgct ggagttcttc gcccaccccg ggctcgatcc ctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgtat ctgctgtgcc ttctagttgc cagccatctg ttgtttgccc	3180
ctccccgtg cttcccttga ccctggagg tgccactccc actgtccttt ctaataaaaa	3240
tgaggaaatt gcatcgatt gtctgagtag gtgtcattct attctggggg gtgggtggg	3300
gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcgggtggg	3360
ctctatggcc gatcagcgat cgctgaggtg ggtgagtgccc cgtggcctgg ggtggcgtatg	3420
aaaatatata agttgggggt cttagggctt cttaattgt gttgcagaga ccgcggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcggcc ttggatggca gcatcgatg	3540
cccttatttg acgacgcggaa tgccccactg ggccgggtg cgtcagaatg tgcgtggctc	3600

[0040]

cagcatcgac ggccgacccg tcctgcccgc aaattccgcc acgctgacct atgcgaccgt	3660
cgcggggacg ccgttggacg ccacccgcgc cgccgcccgc accgcagccg cctcgccgt	3720
gcccagcctg gccacggact ttgcattcct gggaccactg ggcacagggg ctacttctcg	3780
ggccgctgct gccgcccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggtctc	3900
ctccctgcaa gctggcggga atgcttctcc cacaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctcttattt cataatttc	4020
cgcgcgcgat aggcctaga ccagcgttct cggcgttga gggtcgggtg tatcttctcc	4080
aggacgtggc agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccgggg	4140
tggaggttagc accactgcag agcttcatgc tccgggttgg tttgttagat gatccagtcg	4200
tagcaggagc gctgggcattt gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tggtaagt gttacaaaa cggtaagtt gggagggtg cattcggga	4320
gagatgatgt gcatcttggc ctgtatTTT agattggcga tttccgc cagatccctt	4380
ctgggattca tgggtgcag gaccaccagt acagtgtatc cggcactt gggaaatttgc	4440
tcatgcagct tagagggaaa agcgtggaaag aacttggaga cgcctttgtg gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atggccgcgc gggaggcagc ttggcaaaag	4560
atatttctgg ggtcgctgac gtcgtatTTT tggccagg tgaggtcgatc ataggcatt	4620
tttacaaagc gcggcggag ggtccccgac tggggatga tggcccctc tggccctgg	4680
gcgtagttgc cctcgcagat ctgcatttcc caggcctaa tctggagggg gggaaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctggatgag	4800
agcaggTTTC taagcagctg tgatTTTca acccggtgg gcccataaat aacacccata	4860
accggTTGca gctggtagtt tagagagctg cagctccgt cgtccggag gggggggcc	4920
acctcgTTGA gcatgtccct gacgcgcattt ttctccccga ccagatccgc cagaaggcgc	4980
tcggcccccga gggacagcag ctcttgcagaa gaagcaaagt tttcagcgg cttgaggccg	5040
tccggcgtgg gcatgtttt cagggtctgg ctcagcagct ccaggcggc ccagagctcg	5100
gtgacgtgct ctacggcattc tctatccagc atatctccctc gttcgcggg ttggggcgac	5160
tttcgctgtt gggcaccaag cgggtgtcgat ccagcggggc cagagtcatg tccttccatg	5220
ggcgcagggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgt ccgggctgag	5280
cgcttgccaa ggtgcgcattt aggctggttc tgctggatgc gaagcgctgc cggcttcgc	5340

[0041]

cctgcgcgtc	ggccaggtag	catttacca	tggtgtcata	gtccagcccc	tccgcggcgt	5400
gtcccttggc	gcmcagcttg	cccttggagg	tggccgcga	cgagggcag	agcaggctct	5460
tgagcgcgt	gagcttgggg	gcmaggaaga	ccgattcggg	ggagtaggcg	tccgcgcgc	5520
agaccccgca	cacggctctcg	cactccacca	gccaggtgag	ctcggggcgc	gccgggtcaa	5580
aaaccagggtt	tccccatgc	tttttgcgt	gtttcttacc	tcgggtctcc	atgaggtgg	5640
gtccccgctc	ggtgacgaag	aggctgtccg	tgtctccgt	gaccgacttg	aggggtcttt	5700
tctccagggg	gttccctcgg	tcttcctcg	agaggaactc	ggaccactct	gagacgaagg	5760
cccgcttcca	ggccaggacg	aaggaggcta	tgtgggaggg	gtacgggtcg	ttgtccacta	5820
gggggtccac	cttctccaag	gtgtgaagac	acatgtcgcc	ttcctcggcg	tccaggaagg	5880
tgattggctt	gtaggtgtag	gccacgtgac	cgggggttcc	tgacgggggg	gtataaaagg	5940
gggtggggc	gcmctcgctcg	tcactcttct	ccgcatcgct	gtctgcgagg	gccagctgct	6000
ggggtgagta	ttccctctcg	aaggcgggca	tgacctccgc	gctgagggttg	tcagttcca	6060
aaaacgagga	ggatttgatg	ttcacctgtc	ccgaggtgat	acctttgagg	gtacccgcgt	6120
ccatctggtc	agaaaacacg	atcttttat	tgtccagctt	ggtggcgaac	gaccgtaga	6180
ggcggttgg	gagcagcttg	gcmatggagc	gcagggctcg	gttcttgc	ctgtcggcgc	6240
gctcccttggc	cgcgatgttg	agctgcacgt	actcgcgcgc	gacgcagcgc	cactcgggga	6300
agacggtggt	gcmctcgctcg	ggcaccaggc	gcacgcgc	gccgcgggttg	tgcaggggtga	6360
ccaggtccac	gctgggtggcg	acctcgccgc	gcaggcgctc	gttggtccag	cagagacggc	6420
cgcccttgcg	cgagcagaag	gggggcaggg	ggtcgagctg	ggtctcg	gggggtcccg	6480
cgtccacgg	gaaaaccccg	gggcgcaggc	gcmctcgaa	gtatctatc	ttgcaacctt	6540
gcatgtccag	cgcctgctgc	cagtcgcggg	cggcgagcgc	gcmctcgtag	gggttgagcg	6600
gccccccca	ggcatgggg	tgggtgagtg	cggaggcgta	catgcgcag	atgtcataga	6660
cgtagagggg	ctcccgagg	accccgatgt	aggtgggtta	gcagcggccg	ccgcccgtgc	6720
tggcgccac	gtatctatac	agctcggtcg	agggggcgag	gaggtcg	cccagggttg	6780
tgcgggcggg	gcmctccgcg	cgaaagacga	tctgcctgaa	gatggcatgc	gagttggaag	6840
agatggtggg	gcmctggaag	acgttgaagc	tggcgctcg	caggccgacg	gcmctcg	6900
cgaaggaggc	gtaggagtcg	cgcagttgt	gtaccagctc	ggcggtgacc	tgcacgtcga	6960
gcmctcgagta	gtcgagggtc	tcgcggatga	tgtcatattt	agcctgc	cccc ttctttcc	7020

[0042]

acagctcgcg gttgaggaca aactttcgc ggtcttcca gtactttgg atcggaaac	7080
cgtccggttc cgaacggtaa gagcctagca tgtagaactg gttgacggcc tggtaggcgc	7140
agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg	7200
tcagggcgaa ggtgtccctg accatgactt tgaggtactg gtgcttgaag tcggagtcgt	7260
cgcagccgcc ccgctcccg agcgagaagt cgggtcgctt cttggagcgg gggttggca	7320
gagcgaaggt gacatcgtt aagaggattt tgcccgccg gggcatgaag ttgcgggtga	7380
tgccgaaggg ccccgccact tcagagcggt tggatgac ctggcgccg agcacatct	7440
cgtcgaagcc gttgatgttg tggccacga tgtagatcc caggaagcgg gcccggccct	7500
ttacggtggg cagttcttt agctttcgta aggtgagctc ctcggcgag gcgaggccgt	7560
gctcgccag gcccagtcg cggaggtcgc gttgtctct gaggaaggac ttccagaggt	7620
cgcggccag gagggtctgc aggccgtctc tgaaggctt gaactggcgg cccacggcca	7680
tttttcggg ggtgatgcag tagaaggta ggggttttgc tggccagcgg tccagtcga	7740
gctgcaggc gaggtcgccg cggcggtga ccaggcgctc gtcggcccg aatttcatga	7800
ccagcatgaa gggcacgagc tgcttcga aggccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtcgag gatgcgagcc gatcgggaag aactggatct	7920
ccgcacactc gtgctggctt ttgtaaaagc gagcgcagta ctggcagcgc tgcacggct	8040
gtacctcatg cacgagatgc accttcgccc cgccacgag gaagccgagg gaaatctga	8100
ccccccgccc tggctcgccg catggctggt tctttctac tttggatgac tggccgtctc	8160
cgtctggctc ctcgaggggt gttacgggtt agcggaccac cacggcgccg gagccgcagg	8220
tccagatatac ggccgcggc ggtcgagtt tggatgcac atcgccgcgc tggagctgt	8280
ccatggctcg gagctccgc ggcggccggca ggtcagccgg gagttttgcg agttcacct	8340
cgcagagtgc ggccaggcgc cggggcaggt ctaggtggta cctgatctt agggcgtgt	8400
tggtggccgc gtcgatggct tgcaggagcc cgcagcccg gggggcgcacg acgggtcccc	8460
cgccccgtgtt ggtgggtgtt gcggtgcagc tcagaagcgg tgccgcggc gggcccccgg	8520
aggttaggggg ggctccggc cgcggccag gggcggcagc ggcacgtcgg cgtggagcgc	8580
ggccaggagt tggtgctgtt cccggagggtt gctggcgaag ggcacgcgc ggcgggttgc	8640
ctcctggatc tggcgccctc gcgtgaagac gacggcccg gtgagttga acctgaaaga	8700
gagttcgaca gaatcaatct cgggtcatt gaccgcggcc tggcgcagga tctcctgcac	8760

[0043]

gtctcccgag ttgtcttggc	aggcgatctc gccatgaac tgctcgatct cttccctc	8820
gaggtctccg cgtccggcgc	gttccacggc ggccgccagg tcgttgaga tgcgc	8880
gagctgcgag aaggcgttga	gtccgcctc gttccagact cggctgtaga ccacg	8940
ctggtcatcg cggcgcgc	tgaccacctg cgcgagggtg agctccacgt gccgc	9000
gacggcgtag ttgcgcagac	gctggaagag gtagttgagg gtggcggcgtg tgcgc	9060
cacgaagaag ttcatgaccc	agcggcgca cgtggattcg ttgatgtccc ccaagg	9120
cagccgttcc atggcctcgt	agaagtccac ggcgaagttg aaaaactggg agttgcgc	9180
cgacacggtc aactccct	ccagaagacg gatgagctcg gcgacggtgt cgccac	9240
gcgctcgaag gctatggga	tctccctc cgctagcatc accacctcct cctcc	9300
ctcttcgtgc acttccatga	tgcgttccctc ctctcgggg ggtggcggcgc gggcgg	9360
gggagggggc gctctgcgc	ggcggcggcgc caccgggagg cggccacga agcgc	9420
catctccccg cggcggcggc	gcatggtctc ggtgacggcg cggcgttct cccgggg	9480
cagttggaag acgcccgg	acatctggtg ctggggcggg tggccgtgag gcagc	9540
ggcgctgacg atgcata	acaattgctc cgtaggtacg ccggcgaggg acctgagg	9600
gtccatatcc accggatcc	aaaacccccc gaggaaggcg tctaaccagt cgca	9660
aggtaggctg agcaccgtgg	cggcggcggcgg ggggtgggg gagtgtctgg cggagg	9720
gctgatgatg taattgaat	aggcggactt gacacggcg atggtcaca ggagcaccat	9780
gtccttgggt ccggcctgct	ggatgcggag gcggcggct atgcggcagg ctgc	9840
gcatcggcgc aggtcctt	gttgcggctt catgagcctt tccacccggca cctcc	9900
ttccctttct gcttcttcc	tgtctgttc ggcctgggg cggcggcgc cccccc	9960
ccccatgcgc gtgaccc	cccccgtgag cgggtggagc agggccaggt cggc	10020
gcgctcggcc aggtggc	ctgcacctg cgtgagggtg gtttggaaat catcca	10080
cacgaagcgg tggtaggc	ccgtgttgc ggtgttaggtg cagttggcca tgacgg	10140
gttgcggc tggtggcc	gttgcgcacat ctgggtgtac ctgagtcgc agtaggc	10200
ggagtcgaag acgtagtc	tgcaagtccg caccaggtac tggtagccca ccaggaa	10260
cgccggcggc tggcggt	aga gggccagcg cagggtggcg gggctccgg gggc	10320
ttccagcatg aggccgtgg	gtacccatggac atccaggtga taccgcggc	10380
ggtggtggag gcgcgcgg	gac agtgcgcac ccgggtccag atgttgcgc aaaa	10440

[0044]

gtgctccatg	gtaggcgtgc	tctgtccagt	cagacgcgcg	cagtcgttga	tactctagac	10500
cagggaaaac	gaaagccggt	cagcgggcac	tcttccgtgg	tctggtaat	agatcgaag	10560
ggtatcatgg	cggagggcct	cgggtcgagc	cccggtccg	ggccggacgg	tccgccatga	10620
tccacgcgtt	taccgcgc	gtgtcgaacc	caggtgtcgc	acgtcagaca	acggtgagg	10680
gttccttttgcgttttct	ggccgggcgc	cggcgcgcgc	taagagacta	agccgcgaaa		10740
gcgaaagcag	taagtggctc	gctcccgta	gccggaggga	tccttgctaa	gggttgcgtt	10800
gcggcgaacc	ccggttcgaa	tccgtactc	gggcccggccg	gaccgcggc	taaggttttgcgttttct	10860
gattggcctc	cccctcgat	aaagaccccg	cttgcggatt	gactccggac	acggggacga	10920
gccctttta	tttttgcgtt	ccccagatgc	atccggtgct	gccccccgc	ccccccgc	10980
cccagcagca	gcaacaacac	cagcaagagc	ggcagcaaca	gcagcgggag	tcatgcaggg	11040
ccccctcacc	caccctcgcc	gggcccggca	cctcggcg	cgccggcg	tctggcg	11100
gcggcggcgg	cggggggccg	gctgacgacc	ccgaggagcc	cccgcggcgc	agggccagac	11160
actacctgga	cctggaggag	ggcgagggcc	tggcgccg	ggggcgccg	tctcccg	11220
gccacccgcg	ggtgcagctg	aagcgcact	cgcgcgaggc	gtacgtgcct	cgccagaacc	11280
tgttcaggga	ccgcgcggc	gaggagcccg	aggagatgcg	ggacaggagg	ttcagcgcag	11340
ggcggagct	gcggcagggg	ctgaaccgcg	agcggctgct	gcgcgaggag	gactttgagc	11400
ccgacgcgcg	gacggggatc	agccccgcgc	gcgcgcacgt	ggcggccg	gacctggta	11460
cggcgtacga	gcagacgg	aaccaggaga	tcaacttcca	aaagatttc	aacaaccacg	11520
tgccacgct	ggtggcg	gaggaggta	ccatcg	gatgcac	tggactttg	11580
taagcgcgt	ggtgcaga	ccaaacagca	agcctctgac	ggcgcag	ttcctgatag	11640
tgcagcacag	cagggaca	gaggcg	ggacgcg	gctgaacatc	accgaggccg	11700
agggtcggt	gctgctgg	ctgattaaca	tcctgc	catagtgg	caggaggc	11760
gcctgagc	ggccgaca	gtggcg	tcaactactc	gatg	ctggcaag	11820
tttacgcgcg	caagatctac	cagacgcgt	acgt	gccc	agacaaggag	11880
acg	tttta	catgcgc	atg	gac	gtgaagatcg	11940
accgcaacga	gcgc	atccac	aggcg	gtgt	gacgc	12000
gcgagctgat	gcac	aggc	ccgt	gaggcc	ctgagc	12060
cggagtccta	cttc	cgatgcg	ggggcg	tgcg	ctggcc	12120
aggccgcggg	gg	tcg	ggc	gtgt	ggatg	12180
gtccgcg	actatgac	aggacggc	ggaggat	gagta	cgac	

[0045]

tagaggaggg	cgagtacctg	gactaaaccg	cgggtggtgt	ttccggtaga	tgcaagaccc	12240
gaacgtggtg	gaccggcg	tcggggcg	tctgcagagc	cagccgtccg	gccttaactc	12300
ctcagacgac	tggcgacagg	tcatggaccg	catcatgtcg	ctgacggcg	gtaaccgg	12360
cgcgttccgg	cagcagccgc	aggccaacag	gctctccg	atcctggagg	cggtggtgc	12420
tgcgcgctcg	aaccccacgc	acgagaaggt	gctggccata	gtgaacgcgc	tggccgagaa	12480
cagggccatc	cgcgggacg	aggccgggct	ggtgtacgac	gcgctgctgc	agcgcgtggc	12540
ccgctacaac	agcggcaacg	tgcagaccaa	cctggaccgg	ctggtgggg	acgtgcgcga	12600
ggcggtggcg	cagcgcgagc	gcgcggatcg	gcagggcaac	ctgggctcca	tggtggcgct	12660
gaatgccttc	ctgagcacgc	agccggccaa	cgtccgcgg	ggcaggaag	actacaccaa	12720
ctttgtgagc	gcgcgtgcggc	tgtatggtgac	cgagaccccc	cagagcgagg	tgtaccagtc	12780
ggcccccggac	tacttcttcc	agaccagcag	acagggcctg	cagacggtga	acctgagcca	12840
ggctttcaag	aacctgcggg	ggctgtgggg	cgtgaaggcg	cccaccggcg	accgggcgac	12900
gggtgtccagc	ctgctgacgc	ccaactcg	cctgctgctg	ctgctgatcg	cgccgttcac	12960
ggacagcggc	agcgtgtccc	gggacaccta	cctggggcac	ctgctgaccc	tgtaccgcga	13020
ggccatcggg	caggcgcagg	tggacgagca	cacccccc	gagatcacca	gcgtgagccg	13080
cgcgctgggg	caggaggaca	cgagcagcct	ggaggcgact	ctgaactacc	tgctgaccaa	13140
ccggcggcag	aagattccct	cgctgcacag	cctgaccc	gaggaggagc	gcatttgcg	13200
ctacgtgcag	cagagcgtga	gcctgaacct	gatgcgcgac	gggtgtacgc	ccagcgtggc	13260
gctggacatg	accgcgcgca	acatggacc	ggcatgtac	gccgcgcacc	ggccttacat	13320
caaccgcctg	atggactacc	tgcacgcgc	ggcggccgt	aaccccgagt	actttaccaa	13380
cgccatcctg	aacccgcact	ggctcccgc	gcccgggttc	tacagcgggg	gcttcgaggt	13440
cccgagacc	aacgatggct	tcctgtggga	cgacatggac	gacagcgtgt	tctcccgcg	13500
gccgcaggcg	ctggcgaaag	cgtccctgct	gcgtcccaag	aaggaggagg	aggaggaggc	13560
gagtcgcccgc	cgcggcagca	gccccgtggc	ttctctgtcc	gagctgggg	ccgcagccgc	13620
cgcgcgcccc	gggtccctgg	gccccggccc	ctttccgagc	ctggtgggg	ctctgcacag	13680
cgagcgcacc	acccgcctc	ggctgctggg	cgaggacgag	tacctgaata	actccctgct	13740
gcagccggtg	cgggagaaaa	acctgcctcc	cgcctcccc	aacaacggga	tagagacgc	13800
ggtgacaaag	atgagcagat	ggaagaccta	tgcgcaggag	cacagggacg	cgccctgcgc	13860

[0046]

ccggccgccc acgcggcgcc aggccacga ccggcagcgg gggctggtgt gggatgacga	13920
ggactcccg gacgatagca gcgtgctgga cctgggaggg agcggcaacc cgttcgcgca	13980
cctgcgcccc cgccctggga ggatgtttt aaaaaaaaaaaa aaaaaaagcaa gaagcatgat	14040
gcaaaaatta aataaaactc accaaggcca tggcgaccga gcgttggttt cttgtttcc	14100
cttcagtatg cggcgccg ggatgtacca ggagggaccc cttccctttt acgagagcgt	14160
ggtgggcgca gggcgccgg cgccctttc tcccttgcg tcgcagctgc tggagccgc	14220
gtacgtgcct ccgcgctacc tgcggctac ggggggaga aacagcatcc gttactcgga	14280
gctggcgccc ctgttcgaca ccacccgggt gtacctgggt gacaacaagt cggcggacgt	14340
ggcctccctg aactaccaga acgaccacag caatttttt accacggtca tccagaacaa	14400
tgactacagg ccgagcgagg ccagcacca gaccatcaat ctggatgacc ggtcgactg	14460
ggcgccgac ctgaaaacca tcctgcacac caacatgccc aacgtgaacg agttcatgtt	14520
caccaataag ttcaaggcgc gggtgatggt gtcgcgtcg cacaccaagg aagaccgggt	14580
ggagctgaag tacgagtggg tggagttcga gctgccagag ggcaactact ccgagaccat	14640
gaccattgac ctgatgaaca acgcgatcgt ggagcactat ctgaaagtgg gcaggcagaa	14700
cgggtcctg gagagcgaca tcgggtcaa gttcgacacc aggaacttcc gcctgggct	14760
ggaccccggt accgggctgg ttatgcccgg ggtgtacacc aacgaggcct tccatcccga	14820
catcatctg ctgcccggct gcggggtgga cttcacttac agccgcctga gcaacccct	14880
ggccatccgc aagcggcagc cttccagga gggcttcagg atcacctacg aggacctgga	14940
ggggggcaac atcccccgc tcctcgatgt ggaggctac caggatagct tgaaggaaaa	15000
tgaggcggga caggaggata ccgcggccgc cgcctccgccc gcccggagc agggcgagga	15060
tgctgctgac accgcggccg cggacggggc agaggccgac cccgctatgg tggggggc	15120
tcccgagcag gaggaggaca tgaatgacag tgcggtgccg ggagacacct tcgtcacccg	15180
gggggaggaa aagcaagcgg aggccgaggg cggccggag gaaaagcaac tggcgccagc	15240
agccgcggcg gcggcggtgg ccgcggcgga ggctgagtct gaggggacca agccgcggca	15300
ggagcccggt attaagcccc tgaccgaaga tagcaagaag cgcaattaca acctgctcaa	15360
ggacagcacc aacaccgcgt accgcagctg gtacctggcc tacaactacg gcaaccgc	15420
gacgggggtg cgctcctgga ccctgctgtg cacgcggac gtgacctgca gctcgagca	15480
ggtgtactgg tcgctgccc acatgatgca agacccgtg accttccgct ccacgcggca	15540
ggtcagcaac ttcccggtgg tggcgccga gctgctgccc gtgcactcca agagcttcta	15600

[0047]

caacgaccag	gccgtctact	cccagctcat	ccgccagttc	acctctctga	cccacgtgtt	15660
caatcgcttt	cctgagaacc	agattctggc	gcccggccc	gcccccacca	tcaccaccgt	15720
cagtgaaaac	gttcctgctc	tcacagatca	cgggacgcta	ccgctgcgca	acagcatcg	15780
aggagtccag	cgagtgaccg	ttactgacgc	cagacgccgc	acctgcccct	acgtttacaa	15840
ggccttggc	atagtctcgc	cgcgcgtcct	ttccagccgc	acttttttag	caacaccacc	15900
atcatgtcca	tcctgatctc	acccagcaat	aactccggct	ggggactgt	gcccgcgc	15960
agcaagatgt	tcggagggc	gaggaagcgt	tccgagcagc	accccgtgcg	cgtgcgcggg	16020
cacttccg	ccccctgggg	agcgcacaaa	cgcggccgcg	cggggcgcac	caccgtggac	16080
gacgccatcg	actcggtggt	ggagcaggcg	cgcaactaca	gccccgcggt	ctctaccgt	16140
gacgcggcca	tccagaccgt	ggtgcggggc	gcccggcggt	acgccaagct	gaagagccgc	16200
cggaaagcgcg	tggccgccc	ccaccgcccgc	cgaccgggg	ccgcccgc	acgcgcgc	16260
cgggccctgc	ttcgccggc	caagcgcacg	ggccggccgc	ccgccatgag	ggccgcgcgc	16320
cgcttggccg	ccggcatcac	cggccacc	atggccccc	gtacccgaag	acgcgcggcc	16380
gccggccgcg	ccgcccgc	cagtgcacatg	gccagcaggc	gccggggcaa	cgtgtactgg	16440
gtgcgcact	cggtgaccgg	cacgcgcgt	cccggtcgct	tccgc	cgggacttga	16500
gatgatgtga	aaaaacaaca	ctgagtctcc	tgctttgt	tgtatcccag	cggccggcggc	16560
gcgcgcagcg	tcatgtccaa	gcgcacaaatc	aaagaagaga	tgctccaggt	cgtgcgcgc	16620
gagatctatg	ggccccc	gaaggaagag	caggattcga	agccccc	gataaagcgg	16680
gtcaaaaaga	aaaagaaaga	tgtacgtat	gccgatgggg	aggtggagtt	cctgcgcgc	16740
acggcgccca	ggcgc	gcagtggaa	ggccggcg	taaagcgcgt	cctgcgc	16800
ggcacccgcgg	tggtcttcac	gcccg	cgctccaccc	ggactttcaa	gcgcgtctat	16860
gacgagggt	acggc	gacga	agac	gtcg	cgagagtt	16920
gcttacgg	agcgt	cagcg	ggcgt	gggg	tgctggcg	16980
cagggcaacc	ccac	cccc	tct	gtgacc	ctgc	17040
gcaccctccg	aggc	gaagcg	gggt	ctgaag	cg	17100
cagctcatgg	tgccc	aaagcg	gcag	aggct	gtgc	17160
cccggtctgc	agcc	ggacat	cagg	gtccgc	ccatcaagc	17220
gtcagaccg	tgg	acgtggt	cat	cccc	ggcaactccc	17280
					caccactacc	

[0048]

gctgcctcca	cgacatgga	gacacagacc	gatcccggcg	cagccgcagc	cgcagccgcc	17340
gccgcgaccc	cctcggcgga	ggtgcagacg	gaccctggc	tgccgcccggc	gatgtcagct	17400
ccccgcgcgc	gtcgccccggc	caggaagttac	ggcgccggca	acgcgtctct	gcccggatc	17460
gccttgcata	cttccatcg	gcccaccccc	ggctaccgag	gctataccta	ccgccccgca	17520
agagccaagg	gttccaccccg	cggtccccgc	cgacgcggcg	ccggccaccac	ccggccggcgc	17580
cgccgcagac	gccagccgc	actggctcca	gtctccgtga	ggaaagtggc	gchgacgg	17640
cacaccctgg	tgctgcccag	ggcgcgtac	cacccagca	tcgtttaaaa	gcctgttgt	17700
gttcttgcag	atatggccct	cacttgcgc	ctccgttcc	cggtgccggg	ataccgagga	17760
ggaagatcgc	gccgcaggag	gggtctggcc	ggccgcggcc	tgagcggagg	cagccggcgc	17820
gcccgcggc	ggcgacgcgc	caccagccga	cgcatgcgc	gcgggggtgct	gcccctgtta	17880
atccccctga	tcgcccggc	gatggcgcc	gtgcccggga	tcgcctccgt	ggccttgcaa	17940
gcgtcccaga	ggcattgaca	gacttgcaaa	cttgcaaata	tggaaaaaaaaa	aaccccaata	18000
aaaaagtcta	gactctcact	ctcgcttggt	cctgtgacta	ttttgttagaa	tggaagacat	18060
caactttgcg	tcgctggccc	cgcgtaacgg	ctcgccccc	ttcctgggac	actggaaacga	18120
tatcgccacc	agcaacatga	gccccggcgc	cttcagttgg	ggctctctgt	ggagcggcat	18180
taaaaagtatc	gggtctgccc	ttaaaaattt	cggtcccg	gcctggaaaca	gcagcacggg	18240
ccagatgtt	agagacaagt	tgaaagagca	gaacttccag	cagaagggtgg	tggaggcct	18300
ggcctccggc	atcaacgggg	tggggaccc	ggccaaccag	gccgtgcaga	ataagatcaa	18360
cagcagactg	gaccccccggc	cgccgggtgg	ggaggtgccg	ccggcgctgg	agacgggtgc	18420
ccccgatggg	cgtggcgaga	agcgcggcgc	gcccgtatgg	gaagagacca	ctctggtcac	18480
gcagaccat	gagccgcggc	cgtatgagga	ggccctgaag	caaggtctgc	ccaccacgc	18540
gcccatcgcg	cccatggcca	ccgggggtgg	ggccgcggcc	accccccggca	cgctggactt	18600
gcctccggcc	gccgatgtc	cgcagcagca	gaaggcggca	cagccggggcc	cgcccgccgc	18660
cgccctccgt	tcctccggcg	gtccctctgc	ccgcggccggc	agcggccccc	gcgggggggt	18720
cgcgaggcac	ggcaactggc	agagcacgt	gaacagcatc	gtgggtctgg	gggtgcggc	18780
cgtgaagcgc	cggcgatgt	actgaatagc	ttagctaact	tgttgtatgt	gtgtatgcgc	18840
cctatgtcgc	cggcgagagga	gctgctgagt	cgccggcg	cgccgcggcca	ccaccacccgc	18900
cactccggcc	ctcaagatgg	cgacccatc	gatgtgccc	cagtggcgt	acatgcacat	18960
ctcgggccag	gacgcctcgg	agtacctgag	ccccgggctg	gtcagttcg	cccgccac	19020

[0049]

cgagagctac ttccagcctga gtaacaagg taggaacccc acggtggcgc ccacgcacga	19080
tgtgaccacc gaccggtctc agcgccctgac gctgcggttc attcccggtt accgcgagga	19140
caccgcgtac tcgtacaagg cgcggttcac cctggccgtg ggcgacaacc gcgtgctgga	19200
catggcctcc acctactttg acatccgcgg ggtgctggac cggggtccca ctttaagcc	19260
ctactctggc accgcctaca actccctggc ccccaaggc 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 cctatttatg cagaccctac attccagccc gaaccccaa tcggggagtc	19560
ccagtggaat gaggcagatg ctacagtcgc cggcggtaga gtgctaaaga aatctactcc	19620
catgaaacca tgctatggtt cctatgcaag acccacaaat gctaattggag gtcagggtgt	19680
actaacggca aatgcccagg gacagctaga atctcagggtt gaaatgcaat tctttcaac	19740
ttctgaaaac gcccgttaacg aggctaacaa cattcagccc aaattgggtc tgtatagtga	19800
ggatgtgcac atggagaccc cgatgcga cttttttac aagccgcacaa aaagcgatga	19860
caattcaaaa atcatgctgg gtcagcagtc catgcccac agacctaatt acatggctt	19920
cagagacaac ttatcggcc tcatgttata caatagcact ggcaacatgg gagtgcttgc	19980
aggcaggcc tctcaggta atgcagtgg ggacttgca gacagaaaca cagaactgtc	20040
ctaccagctc ttgcttgatt ccatgggtga cagaaccaga tactttcca tgtgaatca	20100
ggcagtggac agttatgacc cagatgttag aattattgaa aatcatggaa ctgaagacga	20160
gctcccaac tattgtttcc ctctgggtgg catagggta actgacactt accaggctgt	20220
taaaaccaac aatggcaata acggggccca ggtgacttgg acaaaagatg aaactttgc	20280
agatcgcaat gaaatagggg tggaaacaa ttgcgtatg gagatcaacc tcagtgcacaa	20340
cctgtggaga aactccctgt actccaacgt ggcgcgtac ctaccagaca agcttaagta	20400
caacccctcc aatgtggaca tctctgacaa ccccaacacc tacgattaca tgaacaagcg	20460
agtgggtggcc cggggctgg tggactgcta catcaacctg ggcgcgcgt ggtcgctgga	20520
ctacatggac aacgtcaacc cttcaacca ccaccgaat gcccgcctgc gctaccgctc	20580
catgctcctg ggcaacgggc gctacgtgcc cttccacatc caggtgcccc agaagttctt	20640
tgccatcaag aacccctcc tcctggggg ctcctacacc tacgagtggaa acttcaggaa	20700

[0050]

ggatgtcaac atggcctcc agagctctc gggtaacgat ctcagggtgg acggggccag	20760
catcaagttc gagagcatct gcctctacgc caccttcttc cccatggccc acaacacggc	20820
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20880
ctccggccgc aacatgctc accccatacc cgccaacgcc accaacgtcc ccatctccat	20940
cccctcgcc aactggcgcc cttccgcgg ctgggccttc acccgctca agaccaagga	21000
gaccggctcc ctgggctcgg gattcgaccc ctactacacc tactcgggct ccattcccta	21060
cctggacggc accttctacc tcaaccacac tttcaagaag gtctcggtca cttcgactc	21120
ctcggtcagc tggccggca acgaccgtct gtcacccccc aacgagttcg agatcaagcg	21180
ctcggtcgac ggggaggggct acaacgtggc ccagtgcac atgaccaagg actggttcct	21240
ggtccagatg ctggccaact acaacatcgg ctaccaggc ttctacatcc cagagagcta	21300
caaggacagg atgtactct tcttcaggaa cttccagccc atgagccggc aggtggtgaa	21360
ccagaccaag tacaaggact accaggaggt gggcatcatc caccagcaca acaactcggg	21420
cttcgtggc tacctcgccc ccaccatgca cgagggacag gcctaccccg ccaacttccc	21480
ctatccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tctctgcga	21540
ccgcaccctc tggcgcatcc cttctccag caacttcatg tccatgggt cgctctcgga	21600
cctggccag aacttgctct acgccaactc cgcccacgcc ctcgacatga cttcgaggt	21660
cgaccccatg gacgagccca cccttctcta tttctgttc gaagtcttg acgtggccg	21720
ggtccaccag ccgcaccgcg gcgtcatcga gaccgtgtac ctgcgtacgc cttctcgcc	21780
cgccaacgcc accacctaaa gaagcaagcc gcagtcatcg ccgcctgcat gccgtgggt	21840
tccaccgagc aagagctcag gccatcgta agagacctgg gatgcgggcc ctatttttg	21900
ggcacctcg acaagcgctt ccctggcttt gtctccac acaagctggc ctgcgcacatc	21960
gtcaacacgg ccggccgcga gaccggggc gtgcactggc tggccttcgc ctggaacccg	22020
cgctccaaaa catgttctt cttgacccc ttggctttt cgaccagcg getcaagcaa	22080
atctacgagt tcgagttacga gggcttgctg cgtcgacgcg ccatcgctc ctgcggcgc	22140
cgctcgta ccctcgaaaa gtccacccag accgtgcagg ggcccactc ggccgcctgc	22200
ggtcttttct gctgcatttt tctgcacgcc tttgtgcact ggcctcagag tcccatggac	22260
cgcaacccca ccatgaactt gctgacgggg gtgcccact ccatgctcca gagccccag	22320
gtcgagccca ccctcgcccg caaccaggag cagcttacaa gttcttgaa gcccactcg	22380
ccttacttcc gccgccacag cgacacagatc aggaggccca cttcttctg ccacttgaa	22440

[0051]

gagatgcaag aaggtaata acgatgtaca cactttttt ctaataaat ggcattttt	22500
tatttataca agctctctgg ggtattcatt tcccaccacc acccgccgtt gtcgccatct	22560
ggctctattt agaaatcgaa agggttctgc cgggagtcgc cgtgcgccac gggcagggac	22620
acgttgcgat actggtagcg ggtgccccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tcgggaaagt tttcgctcca caggctcgg gtcagcacca gcgcgttcat caggtcggc	22740
gccgagatct tgaagtcgca gttggggccg ccgcctcgc cgcgcgagtt gcggtacacc	22800
gggttgcagc actggaacac caacagcgcc gggtgcttca cgctggccag cacgctcgg	22860
tcggagatca gtcggcgtc caggtectcc gcgttgctca gcgcgaacgg ggtcatcttgc	22920
ggcacttgcc gccccagaa gggcgcgtgc cccggttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccgtg cccggactcg gcgttgggt acagcgcgcg catgaaggcc	23040
tgcatttggc ggaaggccat ctgggccttg gcgcctccg agaagaacat gccgcaggac	23100
ttgcccgaga actggtttc gggcagctg gcgtcgtcga ggcagcagcg cgctcggtg	23160
ttggcgatct gcaccacgtt gcgcacccac cggttcttca cgatcttggc cttggacgat	23220
tgctcatttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttacca tgctgtgcc gtgcagacac ttcaagtcgc cctccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtggcgtcg aaagacttgt aggtcacctc cgcaaggac	23400
tgcaggtacc cctgaaaaaa gggcccatc atggcacga aggtttttt gtcgtgaag	23460
gtcagctgca gcccgcgtg ctccctgttc agccaggct tgcacacggc cgccagcgc	23520
tccacctggt cggcagcat cttaagttt acetttagt catttccac gtggacttg	23580
tccatcagcg tgcgcgcgc ctccatgccc ttctccagg ccgacaccag cggcaggctc	23640
acgggttct tcaccatcac cgtggccgc gcctccgcg cgcttcgtt ttccgcggc	23700
ctgttctttt cctttccctc ctttctcg ccgcgcctca ctgcagccc ccgcaccacg	23760
gggtcgtctt cctgcaggcg ctgcaccttgc gcttgcgt tgcccccctg cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcacc agcgccgcct ttcttgctc gtcctcgctg	23880
tccagaatga cctccggga ggggggttg gtcattctca gtaccgaggc acgcttcttt	23940
ttcttcctgg gggcggtcgc cagctccgcg gtcggccgc ctgcccggat cgaaggccga	24000
gggctggcgc tgcgcggcac cagcgcgtcc tgcgagccgt cctcgcttc ctcggactcg	24060
agacggaggc gggcccgctt ctgcggggc gcgcggggcgcg gcgaggcgg cggcggcgac	24120

[0052]

ggagacgggg acgagacatc gtccagggtg ggtggacggc gggccgcgc gcgtccgc	24180
tcgggggtgg tctcgcgctg gtcctttcc cgactggca tctccactg ctccttctcc	24240
tataggcaga aagagatcat ggagtctctc atgcgagtcg agaaggagga ggacagccta	24300
accggccctt ctgagccctc caccaccgcc gccaccaccc ccaatgccgc cgccggac	24360
gcccacccg agaccaccgc cagtaccacc ctccccagcg acgcacccccc gctcgagaat	24420
gaagtgctga tcgagcagga cccgggtttt gtgagcggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggtcgc cgcctcagtg ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata agatgagac agcagtcggg cggggaaacg gaagccatga tgctgatgac	24600
ggctacctag acgtgggaga cgacgtgctg cttaagcacc tgcaccgcc gtgcgtcatc	24660
gtctgcgacg cgctgcagga gcgcgtcgaa gtgccttgg acgtggcgga gtcagccgc	24720
gcctacgagc ggcaccttt cgccgcac gtgccttca agcgcggga gaacggcacc	24780
tgcgagccca acccgctct caacttctac cccgttcccg cgttacccga ggtgctggcc	24840
acctaccaca tcttttcca aaactgcaag atccccctct cctgccgcgc caaccgcacc	24900
cgcgcgaca aaaccctgac cctgcggcag ggcccaca tacatgatc cgcctcttg	24960
gaggaagtgc ccaagatctt cgagggtctc ggtcgacgc agaaacggc ggcgaacgct	25020
ctgcacggag acagcgaaaa cgagagtcac tcgggggtgc tggggagct cgagggcgac	25080
aacgcgcgcc tggcgact caagcgcagc atagaggta cccacttgc ctacccggcg	25140
ctcaacctgc ccccaaggt catagtggtg gtcatggcgc agtcatcat ggcgcgcgc	25200
cagccctgg ccgcggatgc aaacttgcaa gagtcctccg aggaaggcct gcccggcgc	25260
agcgacgagc agctggcgc ctggctggag acccgcgacc cgcgcagct ggaggagcgg	25320
cgcaagctca tggatggcgc ggtgctggc accgtggagc tcgagtgctc gcagcgcttc	25380
ttcgcggacc ccgagatgca ggcgaagctc gaggagaccc tgcactacac cttccggcag	25440
ggctacgtgc gccaggcctg caagatctcc aacgtggagc tctgcaacct ggtctctac	25500
ctggcatcc tgcacgagaa ccgcctcgcc cagaacgtcc tgcactccac cctcaaagg	25560
gaggcgcgc ggcgactacat ccgcgactgc gcctaccttc tcctctgcta cacctggcag	25620
acggccatgg gggtctggca gcagtgcctg gaggagcga acctcaagga gctggaaaag	25680
ctcctcaagc gcaccctcag ggacctctgg acgggcttca acgagcgctc ggtggccgc	25740
ggcgtggcgg acatcatctt tcccgagcgc ctgctcaaga ccctgcagca gggcctgccc	25800
gacttcacca gccagagcat gctgcagaac ttcaggactt tcatcctgga ggcgtcgcc	25860

[0053]

atccctgccgg ccacttgctg cgcgctgcc agcgacttcg tgcccatcaa gtacagggag	25920
tgcggccgc cgctctgggg ccactgctac ctcttccagc tggccaacta cctgcctac	25980
cactcgacc tcatggaaga cgtgagcggc gagggcctgc tcgagtgcca ctgcccgtgc	26040
aacctctgca cgccccaccg ctctctagtc tgcaacccgc agctgctcag cgagagtcag	26100
attatcggtt cttcgagct gcagggtccc tcgcctgacg agaagtccgc ggctccaggg	26160
ctgaaaactca ctccggggct gtggacttcc gcctacctac gcaaatttgc acctgaggac	26220
taccacgccc acgagatcag gttctacgaa gaccaatccc gcccggccaa ggccggagctc	26280
accgcctgctg tcatcaccca ggggcacatc ctgggcaat tgcaagccat caacaaagcc	26340
cgccgagagt tcttgctgaa aaagggtcgg ggggtgtacc tggaccccca gtccggcgag	26400
gagctaaacc cgctacccccc gccggccgccc cagcagcggg accttgcttc ccaggatggc	26460
accaggaaag aagcagcagc cgccggccgccc gccgcagcca tacatgcttc tggaggaaga	26520
ggaggaggac tgggacagtc aggcagagga ggtttcggac gaggagcagg aggagatgt	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgcaaca ccatcgccct cggtcgcagc cccctcgccg gggccctga aatcctccga	26700
accaggacc agcgctataa cctccgctcc tccggcgccg ggcgcacccg cccgcagacc	26760
caaccgtaga tgggacacca caggaaccgg ggtcggtaaag tccaagtgcc cgccggcc	26820
accgcagcag cagcagcagc agcgccaggg ctaccgctcg tggcgccggc acaagaacgc	26880
catagtcgcc tgcttgcaag actgcgggggg caacatctt ttcgcccggc gttcctgtct	26940
attccaccac ggggtcgccct ttccccgcaa tgtcctgcat tactaccgtc atctctacag	27000
cccctactgc agcggcgacc cagaggcgcc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatac ctccgcccc aagacagcgg cagcagcggc caggagaccc gggcagcag	27120
cgccgggagc ggtggcgca ctgcgcctct cggccaaacga acccctctcg accccgggagc	27180
tcagacacag gatctcccc actttgtatg ccatttcca acagagcaga ggccaggagc	27240
aggagctgaa aataaaaaac agatctctgc gtccttcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagcttcgg cgcacgctgg aggacgcgg ggcactcttc agcaaatact	27360
gcgcgctcac tcttaaagac tagctccgctgg cccttctcgat ttttaggcgg gagaaaacta	27420
cgtcatcgcc ggccggccccc cagccccccc agccgagatg agcaaagaga ttcccacgccc	27480
atacatgtgg agctaccagc cgcagatggg actcgccggc ggagcggccc aggactactc	27540

[0054]

cacccgcatg aactacatga gcgcgggacc ccacatgatc tcacaggtca acggatccg	27600
cgcggcggc aaccaaatac tgctggaaca ggcggccatc accgccacgc cccgccataa	27660
tctcaacccc cgaaattggc ccggcccccct cgtgtaccag gaaacccctt ccggccaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcag gggcgcagct	27780
cgcggcggc tttcggtacg gggcgccggc gctccgacca ggtataagac acctgtatgat	27840
cagaggccga ggtatccagc tcaacgacga gtcggtgagc tcttcgctcg gtctccgtcc	27900
ggacggaact ttccagctcg ccggatccgg ccgctttcg ttacgcccc gccaggcgta	27960
cctgactctg cagacctcg tctcgagcc ccgctccggc ggcacccggaa ccctccagtt	28020
cgtggaggag ttctgtccct cggctactt caaccccttc tcgggaccc tcggacgcta	28080
cccccaccag ttcatccga actttgacgc ggtgaaggac tcggcggacg gctacgactg	28140
aatgtcaggt gtcgaggcag agcagctcg cctgagacac ctcgagcact gccgcggcca	28200
caagtgtttc gcccgcgggtt ctggtgagtt ctgctacttt cagctacccg aggagcatac	28260
cgagggcccg ggcacccggc tccgcctgac cacccaggc gaggttaccc gttccctcat	28320
ccgggagttt accctccgtc ccctgtactt ggagcgggag cggggccctt gtgtcttaac	28380
tatcgctgc aactgccta accctggatt acatcaagat ctttgctgtc atctctgtgc	28440
ttagtttaat aaacgcttag atcagaatct actggggctc ctgtcgccat cctgtgaacg	28500
ccaccgtttt cacccacccccc gaccaggccc aggccaaacct cacctgcggt ctgcacccg	28560
ggcccaagaa gtacctcacc tggtaattca acggcaccccc ctttgggtt tacaacagct	28620
tcgacgggaa cggagtcctcc ctgaaagacc agctctccgg tctcagctac tccatccaca	28680
agaacaccac cctccaactc ttccctccct acctgcggg aacctacgag tgcgtcaccg	28740
gccgctgcac ccacccatcc cgcctgatcg taaaccagag cttccggaa acagataact	28800
ccctcttccc cagaacagga ggtgagctca ggaaactccc cggggaccag ggcggagacg	28860
tacccgtac ctttgtgggg ttaggatttt ttattaccgg gttgtggct ctttaatca	28920
aagtttcctt gagatttggt ctttccttct acgtgtatga acacccaaac ctccaaataac	28980
tctacccttt ctccgaatc aggtgacttc tctgaaatcg ggcttgggt gctgttact	29040
ctgttgattt ttttccttat catactcagc cttctgtgcc tcaggctgc cgcctgtgc	29100
gcacacatct atatctactg ctgggtgctc aagtgcagg gtcggccaccc aagatgaaca	29160
ggtacatggt cctatcgatc cttaggcctgc tggccctggc ggcctgcagc gccggccaaaa	29220
aagagattac ctttgaggag cccgcttgca atgtaacttt caagcccgag ggtgaccaat	29280

[0055]

gcaccaccct cgtcaaatgc gttaccaatc atgagaggct ggcgcac tacaaaaaca	29340
aaactggcca gtttgcggtc tatagtgtgt ttacgcccgg agacccctct aactactctg	29400
tcaccgtctt ccagggcggc 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 cacccctta	29700
ctaattacca ccaccctcct tgcgattgcc catgggttga cacgaatcga agtgcagtg	29760
gggtccaaatg tcaccatggt gggcccccgc ggcaattcca ccctcatgtg ggaaaaattt	29820
gtccgcaatc aatgggttca tttctgctct aaccgaatca gtatcaagcc cagagccatc	29880
tgcgatggc aaaatctaac tctgatcaat gtgcaatga tggatgctgg gtactattac	29940
ggcagcggg gagaaatcat taattactgg cgaccacaca aggactacat gctgcattgt	30000
gtcgaggcac ttcccactac caccccaact accacctctc ccaccaccac caccactact	30060
actactacta ctactactac tactactacc actaccgctg cccgccatac ccgcaaaagc	30120
accatgatta gcacaaagcc ccctcggtct cactcccacg cggcgcccc catcggtgcg	30180
acctcagaaa ccaccgagct ttgcttctgc caatgcacta acgccagcgc tcatgaactg	30240
ttcgacctgg agaatgagga tgtccagcag agctccgctt gcctgaccca ggaggctgt	30300
gagcccggtt ccctgaagca gatcggtgat tcaataattt actcttcttc tttgccact	30360
cccgaaatacc ctcccgattt tactttccac atcacggta ccaaagaccc taacctctt	30420
ttctacctga tgctgctgct ctgtatctct gtggctctt ccgcgtgtat gttactgggg	30480
atgttctgct gcctgatctg ccgcagaaaag agaaaagctc gctctcaggc ccaaccactg	30540
atgcccttcc ctttccccc ggattttgca gataacaaga tatgagctcg ctgctgacac	30600
taaccgcttt actagcctgc gctctaaccct ttgtcgcttg cgactcgaga ttccacaatg	30660
tcacagctgt ggcaggagaa aatgttactt tcaactccac ggccgataacc cagtggtcg	30720
ggagtggctc aggtagctac ttaactatct gcaatagctc cacttcccccc ggcataatccc	30780
caaccaagta ccaatgcaat gccaggctgt tcaccctcat caacgcttcc accctggaca	30840
atggactcta tgttaggctat gtaccctttg gtggcaagg aaagacccac gtttacaacc	30900
tggaaagttcg ccagcccaga accactaccc aagcttctcc caccaccacc accaccacca	30960

[0056]

ccatcaccag cagcagcagc agcagcagcc acagcagcag cagcagatta ttgactttgg	31020
ttttggccag ctcatctgcc gctacccagg ccatctacag ctctgtgcc gaaaccactc	31080
agatccaccg cccagaaaacg accaccgcca ccaccctaca caccctcagc gatcagatgc	31140
cgaccaacat cacccccattg gctttcaaa tgggacttac aagccccact ccaaaaccag	31200
tggatgcggc cgaggtctcc gccctcgta atgactggc ggggctggga atgtggtggt	31260
tcgccatagg catgatggcg ctctgcctgc ttctgctctg gctcatctgc tgcctccacc	31320
gcagggcggc cagacccccc atctatagac ccatcattgt cctgaacccc gataatgatg	31380
ggatccatag attggatggc ctgaaaaacc tacttttgc ttttacagta tgataaattt	31440
agacatgcct cgcatttct tgtacatgtt cttctccca cttttctgg ggtgttctac	31500
gctggccgct gtgtctcacc tggaggtaga ctgcctctca cccttcaactg tctacctgct	31560
ttacggattt gtcaccctca ctctcatctg cagcctaatac acagtaatca tgccttcatt	31620
ccagtgcatt gattacatct gtgtgcgcct cgcatacttc agacaccacc cgcaagtaccg	31680
agacaggaac attgccaac ttctaagact gctctaatac tgcataagac tgtgatctgc	31740
cttctgatcc tctgcatect gcccacccctc acctctgcc agtacaccac aaaatctccg	31800
cgcaaaagac atgcctcctg ccgcttcacc caactgtgga atataccaa atgctacaac	31860
gaaaagagcg agctctccga agcttggctg tatgggtca tctgtgtctt agtttctgc	31920
agcaactgtct ttgcctcat aatctacccc tactttgatt tggatggaa cgcgatcgat	31980
gccatgaatt accccaccc tcccgacacc gagataattc cactgcgaca agttgtaccc	32040
gttgcgtta atcaacgccc cccatccc acgcccactg aaatcagcta cttaaccta	32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatcagta ccgagcagcg	32160
tctcctagag aggcgcaggc aggcggctga gcaagagcgc ctcaatcagg agtccgaga	32220
tctcgtaac ctgcaccagt gcaaaagagg catctttgt ctggtaaagc aggccaaagt	32280
cacctacgag aagaccggca acagccaccc cctcagttac aaattgccc cccagcgcca	32340
gaagctggtg ctcatggtg gtgagaatcc catcaccgtc acccagcact cggtagagac	32400
cgaggggtgt ctgcactccc cctgtcgcccc tccagaagac ctctgcaccc tggtaaagac	32460
cctgtcggt ctcagagatt tagtcccatt taactaatca aacactggaa tcaataaaaa	32520
gaatcactta cttaaatca gacagcaggt ctctgtccag tttattcagc agcacctcct	32580
tccctcctc ccaactctgg tactccaaac gccttctggc ggcaaacttc ctccacaccc	32640
tgaaggaaat gtcagattct tgctcctgtc cctccgcacc cactatctt atgttggc	32700

[0057]

agatgaagcg caccaaaacg tctgacgaga gcttcaaccc cgtgtacccc tatgacacgg	32760
aaagcggccc tccctccgtc cctttcctca cccctccctt cgtgtctccc gatggattcc	32820
aagaaaagtcc ccccggggtc ctgtctctga acctggccga gcccctggtc acttcccacg	32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgctggc aacctcacct	32940
ctcaagatat caccaccgct agccctcccc tcaaaaaaac caagaccaac ctcagcctag	33000
aaacctcatc ccccctaact gtgagcacct caggccctt caccgttagca gcccggcgtc	33060
ccctggcggt ggccggcacc tccctcacca tgcaatcaga ggcggccctg acagtacagg	33120
atgcaaaact caccctggcc accaaaggcc ccctgaccgt gtctgaaggc aaactggcct	33180
tgcaaacatc ggccccgctg acggccgctg acagcagcac ctcacagtc agtgccacac	33240
caccccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc atttacacca	33300
ccaatggaaa actaggactt aactttggcg ctcccctgca tgtggtagac agcctaaatg	33360
cactgactgt agttactggc caaggtctta cgataaacgg aacagcccta caaactagag	33420
tctcaggtgc cctcaactat gacacatcag gaaacctaga attgagagct gcaggggta	33480
tgcgagttga tgcaaatggt caacttatcc ttgatgttagc ttaccattt gatgcacaaa	33540
acaatctcag ccttaggctt ggacagggac ccctgtttgt taactctgcc cacaacttgg	33600
atgttaacta caacagaggc ctctacctgt tcacatctgg aaataccaaa aagctagaag	33660
ttaatatcaa 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 acttgtgga	33900
ccacaccaga cccatcccct aactgtagaa tctattcaga gaaagatgct aaattcacac	33960
ttgttttgc taaatgcggc agtcaggtgt tggccagcgt ttctgtttta tctgtaaaag	34020
gtagccttgc gcccattcagt ggcacagtaa ctagtgctca gattgtcctc agatttgatg	34080
aaaatggagt tctactaaggc aattcttccc ttgaccctca atactggAAC tacagaaaag	34140
gtgaccttac agagggcact gcatataccca acgcagtggg atttatgccc aacctcacag	34200
cataccaaa aacacagagc caaactgcta aaagcaacat tgtaagttagt gtttacttga	34260
atggggacaa atccaaaccc atgaccctca ccattaccct caatggact aatgaaacag	34320
gagatgccac agtaagcact tactccatgt cattctcatg gaactggaaat ggaagtaatt	34380

[0058]

acattaatga aacgttccaa accaactcct tcaccttctc ctacatcgcc caagaataaa	34440
aagcatgacg ctgttgattt gattcaatgt gtttctgttt tattttcaag cacaacaaaa	34500
tcattcaagt cattcttcca tcttagctt atagacacag tagcttaata gacccagtag	34560
tgcaaagccc cattcttagct tataactagt ggagaagtac tcgcctacat ggggtagag	34620
tcataatcgt gcatcaggat agggcggtgg tgctgcagca gcgcgcgaat aaactgctgc	34680
cggcccgct ccgtcctgca ggaataacaac atggcagtgg tctcctcagc gatgattcgc	34740
accgcccgcga gcataaggcg cttgtcctc cgggcacagc agcgcaccct gatctca	34800
aaatcagcac agtaactgca gcacagcacc acaatattgt tcaaaatccc acagtgc	34860
gcgcgttgc taaagctcat ggcgggacc acagaaccca cgtggccatc ataccacaag	34920
cgcaggtaga ttaagtggcg acccctcata aacacgctgg acataaacat taccttttt	34980
ggcatgttgt aattcaccac ctcccggtac catataaacc tctgattaaa catggcgcca	35040
tccaccacca tcctaaacca gctggccaaa acctgcccgc cggtataca ctgcaggaa	35100
ccgggactgg aacaatgaca gtggagagcc caggactgt aaccatggat catcatgctc	35160
gtcatgatcat caatgttggc acaacacagg cacacgtca tacacttcct caggattaca	35220
agctcctccc gcgttagaac catatcccag ggaacaaccc attcctgaat cagcgtaat	35280
cccacactgc agggaaagacc tcgcacgtaa ctcacgtgt gcattgtcaa agtgttacat	35340
tcgggcagca gcggatgatc ctccagttatg gtagcgcggg tttctgtctc aaaaggaggt	35400
agacgateccc tactgtacgg agtgcgcga gacaaccgag atcgtgttgg tcgttagtgc	35460
atgccaaatg gaacgcccga cgtagtcata tttcctgaag tcttagatct ctcaacgcag	35520
caccagcacc aacacttcgc agtgtaaaag gccaagtgcg gagagagtt atataggaat	35580
aaaaagtgac gtaaacgggc aaagtccaaa aaacgcccag aaaaacgcgca cgcaaccta	35640
cgcggcggaaa cgaaagccaa aaaacactag acactccctt cggcgtaa cttccgttt	35700
cccacgtac gtcacttgcc ccagtcaaac aaactacata tcccgaaactt ccaagtgcgc	35760
acgccccaaa caccgcctac acctccccgc cggccggccc gcccccaaac cgcctcccg	35820
cccccgcccc cggcccgccc cgcccatctc attatcatat tgggtcaat ccaaataag	35880
gtatattatt gatgatggtt taaacggatc ctcttagagtc gacctgcagg catgcaagct	35940
ttagtattct atagtgtcac ctaaatagct tggcgtaatc atggcatag ctgtttcctg	36000
tgtgaaattt ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgt	36060
aagcctgggg tgcctaataatgt gtagactaac tcacattaat tgcgttgcgc tcactgc	36120

[0059]

cttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcgccaa cgcaacccc	36180
ttgcggccgc ccgggcccgtc gaccaattct catgttgac agcttatcat cgaatttctg	36240
ccattcatcc gcttatttac acttattttag gcgttagcaac caggcggtta agggcaccaa	36300
taactgcctt aaaaaaatta cgccccgccc tgccactcat cgcatgtactg ttgttaattca	36360
ttaagcattc tgccgacatg gaagccatca caaacggcat gatgaacctg aatgccagc	36420
ggcatcagca cttgtcgcc ttgcgtataa tatttgccta tggtaaaaac gggggcgaag	36480
aagttgtcca tattggccac gtttaaatca aaactggta aactcacca gggattggct	36540
gagacgaaaa acatatttctc aataaaccct ttagggaaat aggccaggtt ttcaccgtaa	36600
cacgccacat cttgcgataa tatgtgtaga aactgccga aatgcgtcg gtattcactc	36660
cagagcgatg aaaacgtttc agtttgctca tggaaaacgg tgtaacaagg gtgaacacta	36720
tcccatatca ccagctcacc gtcttcatt gccatacgga attccggatg agcattcatc	36780
aggcgggcaa gaatgtgaat aaaggccgga taaaacttgt gcttattttt cttaacggtc	36840
tttaaaaagg ccgtaatatc cagctgaacg gtctggttat aggtacattg agcaactgac	36900
tgaaatgcct caaaatgttc tttacgatgc cattggata tatcaacgg ggtatatacca	36960
gtgattttt tctccatttt agtttcctta gctcctgaaa atctcgataa ctcaaaaaat	37020
acgcccggta gtgatctt ttcattatgg taaaagttgg aacctttac gtgcgcata	37080
acgtctcatt ttgcacaaaa gttggcccg ggcttcccg tatcaacagg gacaccagg	37140
tttattttt ctgcgaagtg atcttcgtc acaggattt attcgcata agctcatgga	37200
gcggcgtaac cgtcgacag gaaggacaga gaaagcgcgg atctggaaag tgacggacag	37260
aacggtcagg acctggattt gggaggcggt tgccggct gctgctgacg gtgtgacgtt	37320
ctctgttccg gtcacaccac atacgttccg ccattctat gcgtgcaca tgctgtatgc	37380
cggtataccg ctgaaagttc tgcaaagcct gatggacat aagtccatca gttcaacgg	37440
agtctacacg aaggttttt cgctggatgt ggcttcccg caccgggtgc agtttgcgt	37500
gccggagtct gatgcgggtt cgatgctgaa acaattatcc tgagaataaa tgccttggcc	37560
tttatatgga aatgtggaac tgagtggata tgctttttt gtctgttaaa cagagaagct	37620
ggctgttatac cactgagaag cgaacgaaac agtcggaaa atctcccatt atcgttagaga	37680
tccgcattat taatctcagg agcctgtgta gcgtttatag gaagtagtgt tctgtcatga	37740
tgcctgcaag cgtaacgaa aacgatttga atatgccttc aggaacaata gaaatcttcg	37800

[0060]

tgccgtgtta cgttgaagt gaggcggatta tgtcagcaat ggacagaaca acctaata 37860  
cacagaacca tggatgtggtc tgtcctttt cagccagtag tgctcgccgc agtcgagcga 37920  
cagggcgaag ccctcgagtg agcgaggaag caccaggaa cagcacttat atattctgct 37980  
tacacacgt gcctgaaaaa acttcccttg gggttatcca cttatccacg gggatatttt 38040  
tataattatt tttttatag ttttttagatc ttctttttt gagcgccttg taggccttt 38100  
tccatgctgg ttcttagagaa ggtgttgtga caaattgccc tttcagtgatc acaaattcacc 38160  
ctcaaattgac agtcctgtct gtgacaaatt gcccttaacc ctgtgacaaa ttgcctcag 38220  
aagaagctgt ttttcacaa agttatccct gcttattgac tctttttt ttagtgtac 38280  
aatctaaaaa cttgtcacac ttcacatgga tctgtcatgg cgaaacacgc gtttatcaat 38340  
cacaagaaac gtaaaaatag cccgcgaatc gtccagtcac acgacctcac tgaggcggca 38400  
tatagtctct cccggatca aaaacgtatg ctgtatctgt tcgttgacca gatcagaaaa 38460  
tctgatggca ccctacagga acatgacggt atctgcgaga tccatgttgc taaatatgct 38520  
gaaatattcg gattgacctc tgcggaaagcc agtaaggata tacggcaggc attgaagagt 38580  
ttcgcgggaa aggaagtgg ttttatcgc cctgaagagg atgccggcga tgaaaaaggc 38640  
tatgaatctt ttccctgggt tatcaaacgt ggcacagtc catccagagg gctttacagt 38700  
gtacatatca acccatatct cattcccttc tttatcggt tacagaaccg gtttacgcag 38760  
tttcggctta gtgaaacaaa agaaatcacc aatccgtatg ccatgcgtt atacgaatcc 38820  
ctgtgtcagt atcgtaagcc ggatggctca ggcacgtct ctctgaaaat cgactggatc 38880  
atagagcgtt accagctgcc tcaaagtac cagcgtatgc ctgacttccg ccggcccttc 38940  
ctgcaggctt gtgttaatga gatcaacacgc agaactccaa tgcgcctctc atacatttag 39000  
aaaaagaaaag gccgccagac gactcatatc gtatttcct tccgcgatata cacttccatg 39060  
acgacaggat agtctgaggg ttatctgtca cagatttgag ggtggttcgt cacattgtt 39120  
ctgacactt gaggtaatt tgtcacagtt ttgctgttcc ttccagcctg catggatttt 39180  
ctcatacttt ttgaactgtatc attttaagg aagccaaatt tgagggcagt ttgtcacagt 39240  
tgatttccctt ctcttccct tcgtcatgtc acctgatatc gggggtagt tcgtcatcat 39300  
tgatgagggt tgattatcac agtttattac tctgaattgg ctatccgcgt gtgtacctct 39360  
acctggagtt ttcccacgg tggatatttc ttcttgcgt gacgcgtaa gctatctgac 39420  
agaacagttc ttcttgctt cctcgccagt tcgctcgcta tgctcggtta cacggctcgc 39480  
gcgagcgtta gtgataataa gtgactgagg tatgtgtct tcttatctcc tttttagtgc 39540

[0061]

ttgctttat tttaaacaaac ttgcggttt tttatgact ttgcatttt gttttgctt	39600
tgcagtaaat tgcaagattt aataaaaaaa cgcacaaagcaa tgattaaagg atgttcagaa	39660
tgaaactcat ggaaacactt aaccagtgc taaacgctgg tcatgaaatg acgaaggcta	39720
tcgccattgc acagtttaat gatgacagcc cggaagcgg gaaaataacc cggcgctgg	39780
gaataggtga agcagcggat ttagttgggg tttcttctca ggctatcaga gatgccgaga	39840
aagcagggcg actaccgcac ccggatatgg aaattcgagg acgggtttag caacgtttg	39900
gttatacat tgaacaaatt aatcatatgc gtatgtgtt tggtaacgcga ttgcacgtg	39960
ctgaagacgt atttccacccg gtatcgcccc taaaggtggc gtttacaaaa	40020
cctcagtttc ttttcatctt gtcaggatc tggctctgaa gggctacgt gtttgcctg	40080
tggaaaggtaa cgaccccccag ggaacacgcct caatgtatca cggatggta ccagatctc	40140
atattcatgc agaagacact ctccctgcctt tctatcttgg gaaaaaggac gatgtcactt	40200
atgcaataaa gcccacttgc tggccggggc ttgacattat tccttcctgt ctggctctgc	40260
accgtattga aactgagttt atggcaaat ttgatgaagg taaactgccc accgatccac	40320
acctgatgct ccgactggcc attgaaactg ttgctcatga ctatgatgatc atagttattt	40380
acagcgcgcc taacctgggt atcggcacga ttaatgtcgatgt gatgtgctga	40440
ttgttccac gcctgctgag ttgtttgact acaccccgcc actgcagttt ttgcataatgc	40500
ttcgtgatct gctcaagaac gttatctta aagggtcga gcctgatgta cgtatggc	40560
ttaccaaata cagcaatagt aatggctctc agtccccgtg gatggaggag caaattcggg	40620
atgcctgggg aagcatgggtt ctaaaaaatg ttgtacgtga aacggatgaa gttggtaaag	40680
gtcagatccg gatgagaact gttttgaac aggccattga tcaacgctct tcaactgggt	40740
cctggagaaa tgcttttctt atttggaaac ctgtctgaa tggaaattttc gatgtctga	40800
ttaaaccacg ctgggagatt agataatgaa gcgtgcgcct gttattccaa aacatacgct	40860
caataactcaa ccgggtgaag atacttcgtt atcgacacca gctgccccga tgggtggattc	40920
gttaattgctc cgcttaggat taatggctcg cggtaatgcc attactttgc ctgtatgtgg	40980
tcggatgtg aagtttactc ttgaagtgcgat cgggggtgat agtggatgaga agacactctcg	41040
ggatgttca ggtatgaac gtgaccagga gctgcgtact gaggacgcac tggatgtatctcg	41100
catcccttctt ttttacttgc ctggtaacaaca gacaccggcg ttcggatcgaa gagtatctgg	41160
tgtcatagaa attgcccattt ggagtcgccc tcgtaaagct gctgcactta ccgaaagtga	41220

[0062]

ttatcggtt ctgggtggcg agctggatga tgagcagatg gctgcattat ccagatggg	41280
taacgattat cgcccaacaa gtgcttatga acgtggtcag cgttatgcaa gccgattgca	41340
gaatgaattt gctggaaata tttctgcgct ggctgatgcg gaaaatattt cacgtaagat	41400
tattacccgc tgtatcaaca ccgccaaatt gcctaaatca gttgttgctc tttttctca	41460
ccccgggtgaa ctatctgccc ggtcaggtga tgcacttcaa aaagcctta cagataaaga	41520
ggaattactt aagcagcagg catctaacct tcatgagcag aaaaaagctg gggtgatatt	41580
tgaagctgaa gaagttatca ctcttttaac ttctgtgctt aaaacgtcat ctgcatcaag	41640
aactagtttta agctcacgac atcagttgc tcctggagcg acagtattgt ataagggcga	41700
taaaaatggtg cttaacctgg acaggtctcg tggccaact gagtgtatag agaaaattga	41760
ggccattctt aaggaacttgg aaaagccagc accctgatgc gaccacgtt tagtctacgt	41820
ttatctgtct ttacttaatg tcctttgtta cagggccagaa agcataactg gcctgaatat	41880
tctctctggg cccactgttc cacttgtatc gtcggctga taatcagact gggaccacgg	41940
tcccactcgt atcgtcggtc tgattattag tctgggacca cggtcccact cgtatcg	42000
gtctgattat tagtctggg ccacggtccc actcgtatcg tcggctgtat aatcagactg	42060
ggaccacggc cccactcgta tcgtcggtct gattattagt ctgggaccat ggtcccactc	42120
gtatcgtcggt tctgattatt agtctgggac cacggtccc ctcgtatcgt cggtctgatt	42180
attagtctgg aaccacggtc ccactcgat cgtcggtctg attattagtc tgggaccacg	42240
gtcccactcg tatcgtcggt ctgattatta gtctgggacc acgatcccac tcgtgttgc	42300
ggtctgatta tcggtctggg accacggtcc cacttgtatt gtcgatcaga ctatcagcgt	42360
gagactacga ttccatcaat gcctgtcaag ggcaagtatt gacatgtcgt cgtaacctgt	42420
agaacggagt aacctcggtg tgccgttgta tgcctgtgt ggattgtgc tgcgtctgc	42480
ttatccacaa cattttgcgc acggttatgt ggacaaaata cctggttacc cagggcgtgc	42540
cggcacgtta accgggctgc atccgatgca agtgtgtcgc tgcgacgag ctcgcgagct	42600
cggacatgag gttccccgt attcagtgtc gctgatttgc attgtctgaa gttgtttta	42660
cgttaagttt atgcagatca attaatacga tacctcggtc ataattgatt atttgacgt	42720
gtttgatggc ctccacgcac gttgtgatgt gtagatgata atcattatca cttaacgggt	42780
cctttccggat gatccgacag gttacggggc ggcgaccccg cgggtttcg ctatattga	42840
aaattttccg gtttaaggcg tttccgttct tcttcgtcat aacttaatgt ttttattaa	42900
aataccctct gaaaagaaaag gaaacgacag gtgctgaaag cgagctttt ggcctctgc	42960

[0063]

gttccttc tctgttttgc cgtggaaat gaacaatggaa agtccgagct catcgctaat	43020
aacttcgtat agcatacatt atacgaagtt atattcgatg cggccgcaag gggttcgcgt	43080
cagcgggtgt tggcgggtgt cggggctggc ttaactatgc ggcacagag cagattgtac	43140
tgagagtgcacccatgcgg tgtgaaatac cgacacagatg cgttaaggaga aaataccgca	43200
tcaggcgcca ttgcgcattc aggctgcgcactgtggaa agggcgatcg gtgcggcct	43260
cttcgctatt acgccagctg gcgaaaggaa gatgtgctgc aaggcgattaa agttggtaa	43320
cgcagggtt ttcccagtca cgacgttgta aaacgacggc cagtgaatttga taatacgact	43380
cactataggg cgaatttcgag ctgcgttaccc gggatcctc gtttaaac	43428

&lt;210&gt; 9

&lt;211&gt; 45227

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 9

catcatcaat aatatacctt atttggattt gaagccaata tgataatgag atggcgccg	60
cggggcgccccggcgg gaggcgggtt tggggcgccc cccggcgccggcggcgggtgt	120
gcggaaagtgg actttgttaag tggcgccat gtgacttgct agtgcgggc gcggtaaaag	180
tgacgttttc cgtgcgcac aacgcggccggc ggaagtgaca ttttccgc ggttttacc	240
ggatgttgta gtgaatttgg gcttaaccaa gtaagatttgc cccatggc cgggaaaact	300
gaaacgggaa agtgaatct gattaattttt gcgttagtca taccgcgtaa tattttctta	360
ggcccgaggg actttggccg attacgtggaa ggactcgccc aggtgttttt tgaggtaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgat	480
gtatccatataataatgt acatggat tggctcatgt ccaacattac cgcgttgc	540
acattgatta ttgacttagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atatatggag ttccgcgtta cataacttac ggttaatggc cccgcgttgc gaccggccaa	660
cgcggccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caataggac	720
tttccattga cgtcaatggg tggagtattt acggtaact gcccacttgg cagtacatca	780
agtgtatcat atgccaagta cgcggccat tgcgtcaat gacggtaat ggcggccctg	840
gcattatgcc cagtacatga ctttatggga ctttcctact tggcagtaca tctacgtatt	900

[0064]

agtcatcgct attaccatgg ttagtgcgtt ttggcagttac atcaatggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca ccccattgac gtcaatgggat gtttgttttgc	1020
gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccggcccat tgacgcaaataat	1080
ggcggttagg cgtgtacggt gggaggtcta tataaggcaga gctctcccta tcagtgtatag	1140
agatctccct atcagtgata gagatcgctg acgagctcg ttagtgaacc gtcagatcg	1200
ctggagacgc catccacgct gtttgacccct ccatagaaga caccgggacc gatccagc	1260
ccgcggccgg gaacggtgca ttggaacgctg gattccccgt gccaagagtg agatcttccg	1320
tttatctagg taccgggccc cccctcgagg tcgacggtat cgataagctt cacgctgccc	1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgttagaaag ccagtccgca	1440
gaaacggtgc tgaccggga tgaatgtcag ctactggct atctggacaa gggaaaacgc	1500
aagcgcaaag agaaagcagg tagcttgac tggcgttaca tggcgatagc tagactggc	1560
ggtttatgg acagcaagcg aaccggaaatt gccagctggg gcgcctctg gtaaggttgg	1620
gaagccctgc aaagtaaact ggatggctt cttgcgcacca aggatctgat ggcgcagggg	1680
atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac	1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgttgcg agcatgccc caaaaacgtg	1800
gcgtatgtac tcgtgtatct actaccactc ctaaaaaacc gaaactccgcg ctgcgtaaag	1860
tatgccgtgt tcgtctgact aacggtttcg aagtacttc ctacatcggt ggtgaaggc	1920
acaacctgca ggagcactcc gtgatcctga tccgtggcgg tcgtgttaaa gaccccccgg	1980
gtgttcgtta ccacaccgtta cgtggtgcgc ttgactgctc cggcgtaaaa gaccgtaagc	2040
aggctcggtc caagtatggc gtgaagcgac ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcggttcg catgcttgcgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggttatgac tggcacaac agacaatcggtt ctgcgttgc	2220
gccgcgtgt tccggctgtc agcgccagggg cgcccggttc ttttgc gaccgac	2280
tccgggtccc tgaatgaact gcaggacgag gcagcgccgc tattcggtt ggcacgcacg	2340
ggcggttccctt ggcgcgtgt gctcgacgtt gtcactgaaat cggaaaggaa ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcccggctg catacgcttgc atccggctac ctgcccattc	2520
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggtatggaaatc cggttgc	2580

[0065]

gatcaggatg atctggacga agagcatcag gggctcgcbc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgtcgtga cccatggcga tgcctgcttg	2700
ccgaatatca tggtgtggaaaa tggccgcctt tctggattca tcgactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgtt gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgcctccga ttgcagcgc	2880
atgccttct atgccttct tgacgagttc ttctgagcgg gactctgggg ttgaaatga	2940
ccgaccaagc gacgcccac ctgccatcac gagatttcga ttccaccgcc gccttctatg	3000
aaagggttggg cttcggaatc gtttccggg acgcccgtg gatgatectc cagcgcgggg	3060
atctcatgct ggagttcttc gcccaccccg ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgtat ctgctgtgcc ttctagttgc cagccatctg ttgtttgcc	3180
ctccccgtg cttccttga ccctggaagg tgccactccc actgtcctt cctaataaaa	3240
tgaggaaatt gcatcgatt gtctgagtag gtgtcattct attctgggg gtgggggtggg	3300
gcaggacagc aagggggagg attggaaaga caatagcagg catgctgggg atgcgggtgg	3360
ctctatggcc gatcagcgat cgctgagggt ggtgagttgg cgtggctgg ggtggcgt	3420
aaaatatata agttgggggt cttagggtct ctttatttgc gttgcagaga ccgcgggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgc ttggatggca gcatcgtag	3540
cccttatttgc acgacgcggta tgccccactg ggccgggtg cgtcagaatg tgatggctc	3600
cagcatcgac ggccgaccccg tcctgcccgc aaattccgc acgctgaccc atgcgaccgt	3660
cgcggggacg ccgttggacg ccaccgcccgc cgccgcccgc accgcagccg cctcggccgt	3720
gcccggctg gccacggact ttgcattctt gggaccactg ggcacagggg ctacttctcg	3780
ggccgctgt gccgcgttc gcgatgacaa gctgaccgc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgc accaggtctc	3900
ctccctgcaa gctggggga atgcttctcc cacaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgt ctcttattt cataatttc	4020
cgcgcgcgtat aggcctaga ccagcgttct cggcgttga gggcgggtg tatcttctcc	4080
aggacgtgggt agaggtggct ctggacgtt agatacatgg gcatgagccc gtcccggggg	4140
tggaggttagc accactgcag agcttcatgc tccgggggtgg tggtagat gatccagtc	4200
tagcaggagc gctggccatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tggtaagt gtttacaaaa cggtaagtt gggagggtg cattcgggaa	4320

[0066]

gagatgatgt gcatcttgg a	4380
ctgggattca tgggtgcag gaccaccagt	4440
acagtgtatc cggtgactt gggaaatttgc	
tcatgcagct tagagggaaa agcgtgaaag aacttgaga cgccttgc	4500
gcctcccaga tttccatgc attcgccat gatgatggca atggccgc gggaggcagc	4560
ttggc aaag atatttctgg ggtcgctgac gtcgtatgg tggccagg tgaggtcgat	4620
ataggccatt tttacaaagc gcgggcggag ggtgcccac tggggatga tggcccctc	4680
tggccctgg gcgtagttgc cctcgccat ctgcattcc caggcattaa tctcgagg ggaaatcata	4740
tccacactgca gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctggatgag	4800
agcaggttc taagcagctg tgatttcca caaccgtgg gcccataaat aacacctata	4860
accggttgca gctggtagtt tagagagctg cagctgccgt cgtcccgag gagggggcc	4920
acctcggtga gcatgtccct gacgcgcatt ttctccccga ccagatccgc cagaaggcgc	4980
tcgcccggca gggacagcag ctcttgcaag gaagcaaagt tttcagcgg cttgaggccg	5040
tccggcgtgg gcatgtttt cagggtctgg ctcagcagct ccaggcggc ccagagctcg	5100
gtgacgtgct ctacggcatt tctatccagc atatctcctc gttcgcgg ttggggcgc	5160
tttcgctgta gggcaccaag cggtggtcg ccagggggc cagagtcatg tcctccatg	5220
ggcgcagggt cctcgccagg gtggctctgg tcacggtaa ggggtgcgct ccggctcg	5280
cgcttgccaa ggtgcgttg aggctggttc tgctggtgcgta gaagcgctgc cggtttcgc	5340
cctcgccgtc ggcaggtag catttgcatt tggtgtcata gtccagcccc tccggcgt	5400
gtcccttggc gcgcaggctt cccttggagg tggccggca cgaggggcag agcaggctct	5460
tgagcgcgtt gagcttggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgcgc	5520
agacccgcac cacggctcg cactccacca gccaggtag ctcggggcgc gccgggtcaa	5580
aaaccaggaa tccccatgc ttttgatgc gtttcttacc tcgggtctcc atgaggtgg	5640
gtccccgtc ggtgacgaag aggctgtccg tgtctccgtt gaccgactt aggggtctt	5700
tctccagggg ggtccctcg tcttcctcg agaggaactc ggaccactt gagacgaagg	5760
cccgccgtcca gccaggacg aaggaggcta tgtggaggg ttagcggtcg ttgtccacta	5820
gggggtccac cttctccaag gtgtgaagac acatgtgcc ttccctggc tccaggaagg	5880
tgattggctt gtaggtgttag gccacgtgac cgggggttcc tgacgggggg gtataaaagg	5940
gggtggggc ggcgtcgctg tcactcttt ccgcattcgct gtctgcgagg gccagctgc	6000

[0067]

gggtgagta ttccctctcg aaggcggca tgacctccgc gctgagggtt tcagttcca	6060
aaaacgagga ggatttgatg ttcacctgtc ccgaggtgat acctttggg gtacccgcgt	6120
ccatctggtc agaaaacacg atcttttat tgtccagctt ggtggcgaac gaccggtaga	6180
ggcggttggaa gaggcgttgc gcgatggagc gcagggtctg gttcttgcc ctgtcgccgc	6240
gctccttggc cgcgatgttgc agctgcacgt actcgccgc gacgcagcgc cactcgaaa	6300
agacggtggt ggcgtcgcc ggcaccaggc gcacgcgcca gccgcgggttgc tgcagggtga	6360
ccaggtccac gctggtggcg acctcgccgc gcaggcgctc gttggtccag cagagacggc	6420
cgcccttgcg cgagcagaag gggggcaggg ggtcgagctg ggtctcgcc ggggggtccg	6480
cgtccacggt gaaaaccccg gggcgccaggc ggcgtcgaa gtagtctatc ttgcaacctt	6540
gcatgtccag cgcctgtcg cagtcgcggg cggcgagcgc ggcgtcgtag ggttggcg	6600
gcggggccca gggcatgggg tgggtgagtg cggaggcgta catgcgcag atgtcataga	6660
cgtagagggg ctcccgccagg accccgatgt aggtgggttgcagcggcccg ccgcggatgc	6720
tggcgccac gtagtcatac agtcgtcg agggggcgag gaggtcgccccc cccaggttgg	6780
tgccggccgg ggcgtcccg cgaaagacga tctgcctgaa gatggcatgc gagttggaaag	6840
agatggtggg ggcgtggaaag acgttgaagc tggcgctctg caggccgacg ggcgtcgca	6900
cgaaggaggc gtaggagtcg cgcagcttgcgtt accatgactc ggcgttgcacc tgcacgtcgaa	6960
gcccggcggta gtcgagggtc tcgcggatga tgtcatatcc accctgcggcc ttcttttcc	7020
acagctcgcg gttgaggaca aactcttcgc ggtctttcca gtactttgg atcggaaac	7080
cgtccgggttc cgaacggtaa gagcctagca tgtagaactg gttgacggcc tggtaggcgc	7140
agcagccctt ctccacgggg agggcgtagg cctgcgcggc cttgcggagc gaggtgtggg	7200
tcagggcgaa ggtgtccctg accatgactt tgaggactg gtgcttgcgg tcggagtcgt	7260
cgcagccggcc cgcgtcccg agcgagaagt cggcgccgtt cttggagcgg ggggtggca	7320
gagcgaaggt gacatgttgc aagaggattt tgcccgccgg gggcatgaag ttgcgggtga	7380
tgcggaaaggc ccccgccact tcagagcggt tggtgatgac ctggcgccgg agcacgatct	7440
cgtcgaagcc gttgatgttgc tgcccgccacgt tgtagagttc caggaagcgg ggccggccct	7500
ttacgggtggg cagcttctt agctttcgat aggtgagctc ctcggcgag ggcggccgt	7560
gctcgccag ggcccagtc gcgaggtgcg ggttgcgttctt gaggaggac ttccagaggt	7620
cgcggccag gagggtctgc aggcggcttc tgaaggcttgc gaaactggcgg cccacggcca	7680
tttttcggg ggtgatgcag tagaagggtga ggggtcttgc tggccagcgg tcccaatcgaa	7740

[0068]

gctgcagggc gaggtcgcgc gcggcggtga ccaggcgctc gtcgcccccg aattcatga	7800
ccagcatgaa gggcacgagc tgcttccga aggccccat ccaagtgtag gtctctacat	7860
cgtaggtgac aaagaggcgc tccgtcgag gatgcgagcc gatcgaaag aactggatct	7920
ccggccacca gttggaggag tggctgtta tgtggtgaa gtagaagtcc cgtcgccggg	7980
ccgaacactc gtgctggctt ttgtaaaagc gagcgcagta ctggcagcgc tgcacggct	8040
gtacctcatg cacgagatgc acctttcgcc cgccacgag gaagccgagg ggaaatctga	8100
ccccccgccc tggctcgccg catggctggg tctcttctac tttggatgctg tgtccgtctc	8160
cgtctggctc ctcgaggggt gttacgggtgg agcggaccac cacggcgccg gagccgcagg	8220
tccagatatac ggccgcggc ggtcgagtt ttagtgcac atcgccgcgc tggagctgt	8280
ccatggctcg gagctccgcg ggcggcggca ggtcagccgg gagttcttcg aggttcaccc	8340
cgcagatcg ggccaggcg cgccccaggt ctaggtggta cctgatctt agggcgtgt	8400
tggtggccgc gtcgatggct tgcaggagcc cgccggccgg gggggcgacg acgggtcccc	8460
gcgggggtgg ggtgggttg gcggtgcagc tcagaagcgg tgccgcggc gggccccgg	8520
aggttaggggg ggctccggc cccggggcag gggcggcagc ggcacgtcgg cgtggagcgc	8580
ggccaggagt tggtgctgtg cccggaggtt gctggcgaag ggcacgcgc ggcgttgat	8640
ctcctggatc tggcgccctc gcgtgaagac gacggcccg gtgagttga acctgaaaga	8700
gagttcgaca gaatcaatct cgggtcatt gaccgcggcc tggcgcagga tctcctgcac	8760
gtctcccgag ttgtcttggt aggcgatctc ggcgcataac tgctcgatct cttccctctg	8820
gaggtctccg cgtccggcgc gttccacggt ggcgcagg tcgttgaga tgccggccat	8880
gagctgcgag aaggcggtga gtccgcctc gttccagact cggctgtaga ccaccccc	8940
ctggtcatcg cggcgccca tgaccacctg cgcgagggtt agctccacgt ggcgcgcgaa	9000
gacggcgtag ttgcgcagac gctggaagag gtagttgagg gtggcggcgg tgtgcgcgc	9060
cacgaagaag ttcatgaccc agcggcgcaa cgtggattcg ttgtatgtccc ccaaggcctc	9120
cagccgttcc atggcctcgt agaagtccac ggcgaagttt aaaaactggg agttgcgcgc	9180
cgcacacggtc aactcctcct ccagaagacg gatgagctcg ggcacgggtgt cgccaccc	9240
gcgctcgaaag gctatggga tctcttcctc cgctagcatc accacccctt cctttccctc	9300
ctcttctggc acttccatga tggcttcctc ctcttcgggg ggtggcggcgc gggcggtgg	9360
gggagggggc gctctgcgc ggcggcggcgc caccgggagg cggtccacga agcgcgcgat	9420

[0069]

[0070]

actacacctgga	cctggaggag	ggcgagggcc	tggcgccgt	gggggcgcgc	tctcccgagc	11220
gccaccccgcg	ggtgcagctg	aagcgcgact	cgcgcgaggc	gtacgtgcct	cggcagaacc	11280
tgttcaggga	ccgcgcgggc	gaggagcccg	aggagatgcg	ggacaggagg	ttcagcgcag	11340
ggcgggagct	gcggcagggg	ctgaaccgcg	agcggctgct	gcfgcaggag	gactttgagc	11400
ccgacgcgcg	gacggggatc	agccccgcgc	gcfgcacf	ggcggccgc	gacctggta	11460
cggcgtacga	gcagacggtg	aaccaggaga	tcaacttcca	aaagagtttc	aacaaccacg	11520
tgcgcacgct	ggtggcgcgc	gaggaggtga	ccatggct	gatgcacctg	tggactttg	11580
taagcgcgct	ggtgcagaac	cccaacagca	agcctctgac	ggcgcagctg	ttcctgatag	11640
tgcagcacag	cagggacaac	gaggcgttt	gggacgcgct	gctgaacatc	accgagcccg	11700
agggtcggtg	gctgctggac	ctgattaaca	tcctcagag	catagtggt	caggagcgc	11760
gcctgagcct	ggccgacaag	gtggcggcca	tcaactactc	gatgctgagc	ctggcaagt	11820
tttacgcgcg	caagatctac	cagacgcgt	acgtgccat	agacaaggag	gtgaagatcg	11880
acggtttta	catgcgcac	gcgcgtgaagg	tgctcaccct	gagcgacgac	ctggcgtgt	11940
accgcaacga	gcgcatccac	aaggccgtga	gcgtgagccg	gcggcgcgag	ctgagcgacc	12000
gcgagctgat	gcacagcctg	cagcgggcgc	tggcggcgc	ccgcagcggc	gacagggagg	12060
cggagtccta	cttcgatgct	ggggcggacc	tgcgctggc	gcccagccgg	cggccctgg	12120
aggccgcggg	ggtccgcgag	gactatgacg	aggacggcga	ggaggatgag	gagtacgagc	12180
tagaggaggg	cgagtacctg	gactaaaccg	cgggtggtgt	ttccggtaga	tgcaagaccc	12240
gaacgtggtg	gaccggcgc	tgcggcggc	tctcagagc	cagccgtccg	gccttaactc	12300
ctcagacgac	tggcgcacagg	tcatggaccg	catcatgtcg	ctgacggcgc	gtaacccgga	12360
cgcgttccgg	cagcagccgc	aggccaacag	gctctccgcc	atcctggagg	cgggtgggcc	12420
tgcgcgctcg	aaccccacgc	acgagaaggt	gctggccata	gtgaacgcgc	tggccgagaa	12480
cagggccatc	cgcggcgc	aggccggcgt	ggtgtacgac	gcgcgtctgc	agcgcgtggc	12540
ccgctacaac	agcggcaacg	tgcagaccaa	cctggaccgg	ctggtgcccc	acgtgcgcga	12600
ggcgggtggcg	cagcgcgagc	gcfgcggatcg	gcaggcaac	ctgggctcca	tggtggcgct	12660
gaatgccttc	ctgagcacgc	agccggccaa	cgtccgcgg	ggcaggaaag	actacaccaa	12720
ctttgtgagc	gcfgcggcgc	tgtggtgac	cgagaccccc	cagacgcagg	tgtaccagtc	12780
ggcccccggac	tacttcttcc	agaccagcag	acagggcctg	cagacggta	acctgagcca	12840

[0071]

ggcttcaag aacctgcggg ggctgtgggg cgtgaaggcg cccaccggcg accggggcac	12900
ggtgtccagc ctgctgacgc ccaactcgcg cctgctgctg ctgctgatcg cgccgttcac	12960
ggacagcggc agcgtgtccc gggacaccta cctggggcac ctgctgaccc tgtaccgcga	13020
ggccatcggg caggcgcagg tggacgagca cacttccag gagatcacca gcgtgagccg	13080
cgcgctgggg caggaggaca cgagcagcct ggaggcgact ctgaactacc tgctgaccaa	13140
ccggcggcag aagattccct cgctgcacag cctgacctcc gaggaggagc gcatttgcg	13200
ctacgtgcag cagagcgtga gcctgaacct gatgcgcgac ggggtgacgc ccagcgtggc	13260
gctggacatg accgcgcgca acatggaacc gggcatgtac gccgcgcacc ggccttacat	13320
caaccgcctg atggactacc tgcattgcgc ggccggcgtg aaccccgagt actttaccaa	13380
cgccatcctg aacccgcact ggctcccgcc gcccgggttc tacagcgggg gcttcgaggt	13440
cccgagacc aacgatggct tcctgtggga cgacatggac gacagcgtgt tctcccgcg	13500
gccgcaggcg ctggcggaaag cgtccctgct gcgtcccaag aaggaggagg aggaggaggc	13560
gagtcgcccgc cgccgcagca gcggcgtggc ttctctgtcc gagctgggg cgccagccgc	13620
cgcgcccc gggtccctgg gcggcagccc cttccgagc ctggtgggt ctctgcacag	13680
cgagcgcacc acccgccctc ggctgctggg cgaggacgag tacctaata actccctgct	13740
cgagccggtg cgggagaaaa acctgcctcc cgcctcccc aacaacggga tagagagcct	13800
ggtggacaag atgagcagat ggaagaccta tgcgcaggag cacagggacg cgcctgcgt	13860
ccggccgccc acgccccgcc agcgcacga cggcagcgg gggctgggt gggatgacga	13920
ggactcccgcg gacgatacgca gcgtgctgga cctgggaggg agcggcaacc cggtcgccca	13980
cctgcgcccc cgcctggga ggtgtttta aaaaaaaaaa aaaaaagcaa gaagcatgat	14040
gcaaaaatta aataaaactc accaaggcca tggcgaccga gcgttggttt cttgtgttcc	14100
ttcagttatcg cggcgccgg cgatgtacca ggagggaccc cttccctttt acgagagcgt	14160
ggtggcgcgc gcggcggcgg cgcccttttcc tccctttcg tcgcagctgc tggagccgc	14220
gtacgtgcct cgcgcctacc tgccgcctac gggggggaga aacagcatcc gttactcgga	14280
gctggcgccc ctgttcgaca ccacccgggt gtacctgggt gacaacaagt cggcggacgt	14340
ggcctccctg aactaccaga acgaccacag caatttttt accacggtca tccagaacaa	14400
tgactacagg ccgagcggagg ccagcaccca gaccatcaat ctggatgacc ggtcgactg	14460
ggcggcgcac ctgaaaacca tcctgcacac caacatgccc aacgtgaacg agttcatgtt	14520
caccaataag ttcaaggcgc gggatgggt gtcgcctcg cacaccaagg aagaccgggt	14580

[0072]

ggagctgaag tacgagtggg tggagttcga gctgccagag ggcaactact ccgagaccat	14640
gaccattgac ctgatgaaca acgcgatcgt ggagcactat ctgaaagtgg gcagggagaa	14700
cgggtcctg gagagcgaca tcgggtcaa gttcgacacc aggaacttcc gcctgggct	14760
ggacccctg accgggctgg ttatgcccgg ggtgtacacc aacgaggcct tccatcccga	14820
catcatcctg ctgcccggct gcggggtgga cttcaattac agccgcctga gcaacccct	14880
ggcattccgc aagcggcagc cttccagga gggcttcagg atcacctacg aggacctgga	14940
ggggggcaac atccccgcgc tcctcgatgt ggaggcctac caggatagct tgaaggaaaa	15000
tgaggcggga caggaggata ccgcggccgc cgcctccgccc gcccggcagc agggcgagga	15060
tgctgctgac accgcggccg cggacggggc agaggccgac cccgctatgg tggggaggc	15120
tcccgagcag gaggaggaca tgaatgacag tgcggtgcgc ggagacacct tcgtcacccg	15180
gggggaggaa aagcaagcgg aggccgaggc cgcggccgag gaaaagcaac tggggcagc	15240
agcggcggcg gcggcggtgg ccgcggcgga ggctgagtct gaggggacca agccgc当地	15300
ggagccctg attaagcccc tgaccgaaga tagcaagaag cgcagttaca acctgctcaa	15360
ggacagcacc aacaccgcgt accgcagctg gtacctggcc tacaactacg ggcgc当地	15420
gacgggggtg cgctcctgga ccctgctgtg cacgcggac gtgacctgcg gctggagca	15480
ggtgtactgg tcgctgccc acatgatgca agacccctg accttccgct ccacgc当地	15540
ggtcagcaac ttcccggtgg tggcgccga gctgctgccc gtgcactcca agagttcta	15600
caacgaccag gccgtctact cccagctcat ccgc当地gttcc acctctctga cccacgtgtt	15660
caatcgctt cctgagaacc agattctggc gc当地ccgcccc gccccacca tcaccaccgt	15720
cagtgaaaac gttcctgctc tcacagatca cgggacgcta cc当地gtgc当地 acagcatcgg	15780
aggagtccag cgagtgaccg ttactgacgc cagacccgc acctgccc acgttacaa	15840
ggccttggc atagtctcgc cgccgctcct ttccagccgc acttttttag caacaccacc	15900
atcatgtcca tcctgatctc acccagcaat aactccgct gggactgct gc当地ccgcccc	15960
agcaagatgt tcggagggc gaggaagcgt tccgagcagc acccctgctg cgtgc当地gg	16020
cacttcccgcc cccctgggg agcgc当地aaa cgc当地ccgccg cggggccgac caccgtggac	16080
gacgc当地atcg actcggtgg ggagcaggcg cgcaactaca gc当地ccgctg ctctaccgt	16140
gacgc当地ccca tccagaccgt ggtgc当地ggc gc当地ccgctg acgccaagct gaagagccgc	16200
cggaagcgc当地 tggccc当地ccg ccaccgc当地ccg cgc当地ccg cccaa acgc当地ccgcc	16260

[0073]

gcggccctgc ttgcgcgggc caagcgacg ggccgcccgc cgccatgag ggccgcgc	16320
cgcttggccg cggcatcac cgccgccacc atggccccc gtacccgaag acgcgcggcc	16380
gccgccgcgc cgccgcaccat cagtacatg gccagcaggc gccggggcaa cgttactgg	16440
gtgcgcact cggtgaccgg cacgcgcgtg cccgtgcgt tccgcgcgc gcggacttga	16500
gatgtatgtga aaaaacaaca ctgagtcctcc tgctgttgtg tgtatccag cggcgccgc	16560
gcgcgacgc tcatgtccaa gcgcacaaatc aaagaagaga tgctccaggt cgtcgccgc	16620
gagatctatg ggcccccga gaaggaagag caggattcga agcccccga gataaagcgg	16680
gtcaaaaaga aaaagaaaaga ttagtgcgtt gccgtgggg aggtggagtt cctgcgcgc	16740
acggcgccca ggcccgggt gcagtggaaag ggccggcg taaaagcgcgt cctgcgcgc	16800
ggcaccgcgg tggcttcac gcccggcgag cgctccaccc ggactttcaa gcgcgttat	16860
gacgagggtgt acggcgacga agacctgctg gagcaggcca acgagcgcgtt cggagagtt	16920
gcttacggga agcgtcagcg ggctgggg aaggaggacc tgctggcg tccgctggac	16980
cagggcaacc ccaccccaag tctgaagccc gtgaccctgc agcaggtgct gccgagcagc	17040
gcacccctcg aggcaagcg ggtctgaag cgcgaggcg gcgcacctggc gccaccgtg	17100
cagctcatgg tgcccaagcg gcagaggctg gaggatgtgc tggagaaaat gaaagttagac	17160
cccggtctgc agccggacat cagggctccgc cccatcaagc aggtggcgcc gggctcgcc	17220
gtgcagacgc tggacgtggt catccccacc ggcaactccc ccgcgcgc caccactacc	17280
gctgcctcca cggacatgga gacacagacc gatccgcgc cagccgcagc cgcagccgc	17340
gccgcgaccc ctcggcgga ggtgcagacg gacccctggc tgccgcggc gatgtcagct	17400
ccccgcgcgc gtcgcggcg caggaagtac ggccgcgcca acgcgcctt gcccgagtac	17460
gccttgcata cttccatcgcc gcccacccccc ggctaccgag gctataccta ccgcgcgcga	17520
agagccaagg gttccaccccg ccgtccccgc cgacgcgcgc ccgcaccac ccgcgcgc	17580
cgccgcagac gccagccgc actggctcca gtctccgtga ggaaagtggc ggcgcacgg	17640
cacaccctgg tgctgcccag ggccgcgtac cacccagca tcgtttaaaa gcctgttgt	17700
tttcttgcag atatggccct cacttgcgc ctccgtttcc cggtgccggg ataccgagga	17760
ggaagatcgc gccgcaggag ggtctggcc ggccgcggc tgagcggagg cagccgcgc	17820
gcccgcgcgc ggcgcacgcgc caccagccga cgcatgcgcgc gcgggggtgtt gcccctgtt	17880
atccccctga tcgcccggc gatcgccgcgt tccgcgcgtt ggccttgcaaa	17940
gcgtcccaga ggcattgaca gacttgcaaa cttgcaata tgaaaaaaa aaccccaata	18000

[0074]

aaaaagtcta gactctcacg ctcgcttggc cctgtacta tttttagaa tggaagacat	18060
caactttgcg tcgctggccc cgccgtcacgg ctcgcgcgg ttcctggac actggaacga	18120
tatcggcacc agcaacatga gcgggtggcgc cttcagttgg ggctctctgt ggagccgcat	18180
taaaaagtatc gggtctgccc taaaaaaatta cggctcccg gcctggaaca gcagcacggg	18240
ccagatgtt agagacaagt tgaaagagca gaacttccag cagaagggtgg tggagggcct	18300
ggcctccggc atcaacgggg tggtgaccc gccaaccag gccgtgcaga ataagatcaa	18360
cagcagactg gaccccccgc cgccgggtgg ggaggtgccc ccggcgctgg agacgggtgc	18420
ccccgatggg cgtggcgaga agcgcccgcg gcccgtatgg gaagagacca ctctggtcac	18480
gcagaccat gagccgcccc cgtatgagga gcccgtatgg caaggtctgc ccaccacgcg	18540
gccccatcgcg cccatggcca cccgggtggg gggccgccc accccccgcg cgctggactt	18600
gcctccgccc gccgatgtgc cgccagcagca gaaggcggca cagccggccc cgcccgac	18660
cgcctccgt tcctccgccc gtcctctgcg ccgcgcggcc agcggccccc gcgggggggt	18720
cgcgaggcac ggcaactggc agagcacgct gaacagcatc gtgggtctgg gggtgcggc	18780
cgtgaagcgc cgccgatgct actgaatagc ttagctaact tggttatgt gtgtatgcgc	18840
cctatgtcgc cgccagagga gctgctgagt cgcgcgcgtt cgccgcgcac ccaccaccgc	18900
cactccgccc ctcaagatgg cgacccatc gatgatgcgc cagtggcgt acatgcacat	18960
ctcggccag gacgcctcgg agtacctgag cccggctg gtgcagttcg cccgcgcac	19020
cggagactac ttcaagctga gtaacaagtt taggaacccc acggtgccgc ccacgcacga	19080
tgtgaccacc gaccggcttc agcgcctgac gctgcggttc attccgtgg accgcgagga	19140
caccgcgtac tcgtacaagg cgccgttcac cctggccgtg ggcgacaacc gcgtgctgg	19200
catggcctcc acctactttg acatccgcgg ggtgctggac cgggtccca cttcaagcc	19260
ctactctggc accgcctaca actccctggc ccccaaggcc gctcccaact cctgcgagtg	19320
ggagcaagag gaaactcagg cagttgaaga agcagcagaa gaggaagaag aagatgctga	19380
cggtaagct gaggaagagc aagcagactac caaaaagact catgtatatg ctcaggctcc	19440
cctttctggc gaaaaaaatta gtaaaagatgg tctgcaaata ggaacggacg ctacagctac	19500
agaacaaaaa cctattttatc cagaccctac attccagccc gaacccaaa tcggggagtc	19560
ccagtggaat gaggcagatg ctacagtcgc cggcggtaga gtgctaaaga aatctactcc	19620
catgaaacca tgctatggtt cctatgcaag acccacaaat gctaattggag gtcagggtgt	19680

[0075]

actaacggca aatgccagg gacagctaga atctcaggtt gaaatgcaat tctttcaac	19740
ttctgaaaac gcccgtaacg aggctaaca cattcagccc aaattggtgc tgtatagtga	19800
ggatgtgcac atggagaccc cggatacgca cctttttac aagccgcaa aaagcgatga	19860
caattcaaaa atcatgctgg gtcagcagtc catgccaac agacctaatt acatcgctt	19920
cagagacaac tttatcgcc tcatgttata caatagcact ggcaacatgg gagtgcttgc	19980
aggcaggcc tctcagttga atgcagtggt ggacttgcaa gacagaaaca cagaactgtc	20040
ctaccagctc ttgcttgatt ccatgggtga cagaaccaga tactttcca tgtggatca	20100
ggcagtggac agttatgacc cagatgttag aattattgaa aatcatggaa ctgaagacga	20160
gctcccaac tattgttcc ctctgggtgg catagggta actgacactt accaggctgt	20220
taaaaccaac aatggcaata acggggcca ggtgacttgg acaaaagatg aaactttgc	20280
agatcgcaat gaaatagggg tggaaacaa tttcgctatg gagatcaacc tcagtccaa	20340
cctgtggaga aacttcctgt actccaacgt ggctgtac ctaccagaca agcttaagta	20400
caaccctcc aatgtggaca tctctgacaa ccccaacacc tacgattaca tgaacaagcg	20460
agtggtggcc cggggctgg tggactgcta catcaacctg ggccgcgcgt ggtcgcttgg	20520
ctacatggac aacgtcaacc cttcaacca ccaccgcaat gccccctgc gctaccgctc	20580
catgctctg ggcaacggc gctacgtgcc cttccacatc caggtgccttcc agaagttctt	20640
tgccatcaag aacctctcc tcctgccggg ctcctacacc tacgagtggaa acttcaggaa	20700
ggatgtcaac atggcctcc agagctctt gggtaacgat ctcagggtgg acggggccag	20760
catcaagttc gagagcatct gcctctacgc caccttcttc cccatggccc acaacacggc	20820
ctccacgctc gaggccatgc tcaggaacga caccaacgac cagtccttca atgactacct	20880
ctccgcgc aacatgtct accccatacc cgccaaacgcc accaacgtcc ccatctccat	20940
ccctcgcc aactggcgcc cttccgcgg ctgggccttc acccgctca agaccaagga	21000
gacccttcc ctggctcgg gatcgaccc ctactacacc tactcggct ccatcccta	21060
cctggacggc accttctacc tcaaccacac tttcaagaag gtctcggtca cttcgactc	21120
ctcggtcagc tggccggca acgaccgtct gtcacccccc aacgagttcg agatcaagcg	21180
ctcggtcgac ggggagggt acaacgtggc ccagtgcac atgaccaagg actggttcct	21240
ggtccagatg ctggccaact acaacatcggtt acccgatcc ttctacatcc cagagagcta	21300
caaggacagg atgtacttct tttcaggaa cttccagccc atgagccggc aggtggtgaa	21360
ccagaccaag tacaaggact accaggaggt gggcatcatc caccagcaca acaactcggtt	21420

[0076]

cttcgtggc tacctcgccc ccaccatgcg cgagggacag gcctacccg ccaacttccc	21480
ctatccgctc ataggcaaga ccgcggtcga cagcatcacc cagaaaaagt tcctctgcga	21540
ccgcaccctc tggcgcatcc cttctccag caacttcatg tccatgggtg cgctctcgga	21600
cctgggccag aacttgctct acgccaactc cgcccacgcc ctcgacatga cttcgaggt	21660
cgaccccatg gacgagccca cccttctcta tttctgttc gaagtcttg acgtggtccg	21720
ggtccaccag ccgcaccgcg gcgtcatcga gaccgtgtac ctgcgtacgc cttctcgcc	21780
cgcaacgcc accacctaataa gaagcaagcc gcagtcatcg ccgcctgcat ggcgtcggt	21840
tccaccgagc aagagctcag ggcacatgctc agagacctgg gatgcgggccc tatttttt	21900
ggcaccttcg acaagcgctt ccctggcttt gtctccccac acaagctggc ctgcgccatc	21960
gtcaacacgg ccggccgcga gaccgggggc gtgcactggc tggccttcgc ctggAACCCG	22020
cgctccaaaa catgcttcct ctttgaccctt ttcggctttt cggaccagcg gctcaagcaa	22080
atctacgagt tcgagttacga gggcttgctg cgtcgacgcg ccatgcctc ctcggccgac	22140
cgcgtcgta ccctcgaaaa gtcacccag accgtgcagg gccccgactc ggcgcctgc	22200
ggtcttttct gctgcattttt tctgcacgccc tttgtgcact ggcctcagag tcccatggac	22260
cgcaacccca ccatgaactt gtcgacgggg gtgcacaact ccatgctcca gagccccag	22320
gtcgagccca ccctgcgcgg caaccaggag cagctctaca gttcctggc ggcactcg	22380
ccttacttcc gccccacag cgcacagatc aggagggcca cttcttcgc ccacttgcaa	22440
gagatgcaag aagggttaata acgtatgtaca cactttttt ctaataaat ggcattttt	22500
tatttataca agctctctgg ggtattcatt tccaccacc acccgccgtt gtcgcacatct	22560
ggctctatTTT agaaatcgaa agggttctgc cggagtcgc cgtgcgcac gggcaggac	22620
acgttgcgtat actggtagcg ggtgcacccac ttgaactcgg gcaccaccag gcgaggcagc	22680
tcgggaaagt ttctgcctca caggctgcgg gtcagcacca ggcgttcat caggtcggc	22740
gccgagatct tgaagtcgca gttggggccg ccgcctgcg cgccgcgagg ttgcgtacacc	22800
gggttgcgc actggAACAC caacagcgcc gggtgctca cgctggccag cacgtgcgg	22860
tcggagatca gctcgccgtc caggtcctcc gcgttgctca ggcgaacgg gtcatcttgc	22920
ggcacttgcc gccccaggaa gggcgctgc cccggttcg agttgcagtc gcagcgcagc	22980
gggatcagca ggtgcccgtg cccggactcg gcgttgggtt acagcgcgcg catgaaggcc	23040
tgcatttgcg ggaaggccat ctggccttg ggcgcctccg agaagaacat ggcgcaggac	23100

[0077]

ttgcccggaga actggtttgc gggcagctg gcgtcggtca ggcagcagcg cgcgctgg	23160
ttggcgatct gcaccacgtt gcgcacccac cggttcttca cgatcttggc cttggacgtat	23220
tgctccttca gcgcgcgtg cccgttctcg ctggtcacat ccatctcgat cacatgttcc	23280
ttgttacca tgctgctgcc gtgcagacac tttagtgcgc cttccgtctc ggtgcagcgg	23340
tgctgccaca gcgcgcagcc cgtgggtcg aaagacttgtt aggtcacctc cgcaaggac	23400
tgcaggtaacc cctgcaaaaaa gcggcccatc atggtcacga aggtcttggt gctgctgaag	23460
gtcagctgca gcccgcgtg ctccctgttcc agccagggtct tgacacacggc cgccagcgcc	23520
tccacctggc cgggcagcat cttgaagttc accttcagct catttcacac gtggtaacttg	23580
tccatcagcg tgcgcgcgc ctccatgccc ttctccagg ccgacaccag cgccaggctc	23640
acggggttct tcaccatcac cgtggccgc gcctccgcg cgcttcgtt ttccgcggc	23700
ctgttctttt cctttcttc cttttcgtt ccgcgcgcac ctcgcagccc ccgcaccacg	23760
gggtcgtt cctgcaggcg ctgcacccgtt cgcttgcgt tgcccccgtt cttgatgcgc	23820
acgggcgggt tgctgaagcc caccatcacc agcgcggcctt cttttgtctc gtcctcgctg	23880
tccagaatga cctccgggaa ggggggttg gtcatctca gtaccgggc acgcttctt	23940
ttcttcctgg gggcggtcgc cagctccgcg gctgcggccg ctggcgaggt cgaaggccga	24000
gggctggcg tgcgcggcac cagcgcgtcc tgcgaggccgt cctcgcttc ctcggactcg	24060
agacggaggc gggccgcgtt ctccggggc gcgcggggcg gcggaggccgg cgccggcgac	24120
ggagacgggg acgagacatc gtccagggtt ggtggacggc gggccgcgc gcgtccgc	24180
tcgggggtgg tctcgctgt gtcctttcc cgaactggcca tctccactg ctcccttc	24240
tataggcaga aagagatcat ggagtcttc atgcgagtcg agaaggagga ggacagccta	24300
accgcgcctt ctgagccctc caccacggcc gccaccaccc ccaatgcgc cgccggac	24360
gcgcgcaccc agaccaccgc cagtaccacc ctccccagcg acgcacccccc gctcgagaat	24420
gaagtgcgtga tcgagcagga cccgggtttt gtgagggag aggaggatga ggtggatgag	24480
aaggagaagg aggaggctgc cgcctcgtt ccaaaagagg ataaaaagca agaccaggac	24540
gacgcagata agatgagac agcagtcggg cggggaaacg gaagccatga tgctgatgac	24600
ggctacccat acgtgggaga cgacgtgtcg cttaagcacc tgccacccca gtgcgtcatc	24660
gtctgcgacg cgctgcagga gcgcgtcgaa gtgcgcgtt acgtggcgaa ggtcagccgc	24720
gcctacgagc ggcacccctt cgcgcgcac gtgcgcgcac agcgcggga gaacggcacc	24780
tgcgagccca acccgcggtct caacttctac ccggtcttcg cggtacccga ggtgctggcc	24840

[0078]

acctaccaca	tcttttcca	aaactgcaag	atccccctct	cctgccgcgc	caaccgcacc	24900
cgccgcgaca	aaaccctgac	cctgcggcag	ggcgcccaca	tacctgatat	cgcctctctg	24960
gaggaagtgc	ccaagatctt	cgagggtctc	ggtcgcgacg	agaaaacgggc	ggcgaacgct	25020
ctgcacggag	acagcgaaaa	cgagagtac	tcgggggtgc	tggtggagct	cgagggcgac	25080
aacgcgcgcc	tggccgtact	caagcgcagc	atagaggtca	cccactttgc	ctacccggcg	25140
ctcaacctgc	cccccaaggt	catgagtgtg	gtcatggcgc	agctcatcat	gcccgcgcgc	25200
cagccctgg	ccgcggatgc	aaacttgcaa	gagtccctcg	aggaaggcct	gcccgcggtc	25260
agcgcgagc	agctggcgcg	ctggctggag	acccgcgacc	ccgcgcagct	ggaggagcgg	25320
cgcaagctca	tgatggccgc	ggtgctggtc	accgtggagc	tcgagtgtct	gcagcgcttc	25380
ttcgcggacc	ccgagatgca	gchgcaagctc	gaggagaccc	tgcactacac	cttccgcgcag	25440
ggctacgtgc	gccaggcctg	caagatctcc	aacgtggagc	tctgcaacct	ggtctcctac	25500
ctgggcattcc	tgcacgagaa	ccgcctcggg	cagaacgtcc	tgcactccac	cctcaaaggg	25560
gaggcgcgcc	gchgactacat	ccgcgactgc	gcctacctct	tcctctgcta	cacctggcag	25620
acggccatgg	gggtctggca	gcagtgcctg	gaggagcgc	acctaaggaa	gctggaaaag	25680
ctcctcaagc	gcaccctcag	ggacctctgg	acgggcttca	acgagcgctc	ggtggccgc	25740
gchgctggcg	acatcatctt	tcccgagcgc	ctgctcaaga	ccctgcagca	gggcctgccc	25800
gacttcacca	gccagagcat	gctgcagaac	ttcaggactt	tcatcctgga	gchgctgggc	25860
atcctgcgg	ccacttgctg	cgcgctgccc	agcgacttcg	tgcccatcaa	gtacagggag	25920
tgcccgcgc	cgctctgggg	ccactgtctac	ctcttcagc	tggccaacta	cctgcctac	25980
cactcgacc	tcatggaaga	cgtgagcgcc	gagggctgc	tcgagtgcac	ctgcccgtgc	26040
aacctctgca	cgcaccaccg	ctctctagtc	tgcaacccgc	agctgctcag	cgagagtac	26100
attatcggt	ccttcgagct	gcagggtccc	tcgcctgacg	agaagtccgc	ggctccaggg	26160
ctgaaactca	ctccggggct	gtggacttcc	gcctacctac	gcaaatttgc	acctgaggac	26220
taccacgccc	acgagatcag	gttctacgaa	gaccaatccc	gcccgcggaa	ggcggagctc	26280
accgcctgcg	tcatcacca	ggggcacatc	ctggccaaat	tgcaagccat	caacaaagcc	26340
cgcgcagagt	tcttgctgaa	aaagggtcgg	gggggtgtacc	tggaccccca	gtccggcgag	26400
gagctaaacc	cgctacccccc	gccgcgcgcc	cagcagcgcc	accttgcttc	ccaggatggc	26460
acccagaaag	aagcagcagc	cgccgcgcgc	gccgcagcca	tacatgcttc	tggaggaaga	26520

[0079]

ggaggaggac tggcacagtc aggcagagga gtttcggac gaggagcagg aggagatgat	26580
ggaagactgg gaggaggaca gcagcctaga cgaggaagct tcagaggccg aagaggtggc	26640
agacgcaaca ccatcgccct cggtcgccgc cccctcgccg gggccctga aatcctccga	26700
acccagcacc agcgctataa cctccgctcc tccggcgccg gcgccacccg cccgcagacc	26760
caaccgtaga tgggacacca caggaaccgg ggtcgtaag tccaagtgcc cgccgcgc	26820
accgcagcag cagcagcagc agcgccaggg ctaccgctcg tggcgccggc acaagaacgc	26880
catagtcgcc tgcttgcaag actgcggggg caacatctt ttcgcccgc gcttcgtct	26940
attccaccac ggggtcgctt tccccgcaa tgtcctgcat tactaccgtc atctctacag	27000
ccctactgc agcggcgacc cagaggcgcc agcggcagcc acagcggcga ccaccaccta	27060
ggaagatatac ctccgcccc aagacagcgg cagcagcggc caggagaccc gcggcagcag	27120
cggcgggagc ggtggcgca ctgcgcctct cgcccaacga acccctctcg acccgggagc	27180
tcagacacag gatcttcccc actttgtatg ccatttcca acagagcaga ggccaggagc	27240
aggagctgaa aataaaaaaac agatctctgc gtccttcac ccgcagctgt ctgtatcaca	27300
aaagcgaaga tcagcttcgg cgacgctgg aggacgcgg ggcactttc agcaaatact	27360
gcgcgctcac tcttaaagac tagctccgca cccttcgatg atttaggcgg gagaaaacta	27420
cgtcatcgcc ggccgccc cagccccc agccgagatg agcaaagaga ttcccacgcc	27480
atacatgtgg agctaccagc cgcaaatggg actcgccggc ggagcggccc aggactactc	27540
cacccgcattt aactacatga gcgcgggacc ccacatgatc tcacaggtca acggatccg	27600
cgcgcgca aaccaaatac tgcttggaca ggcggccatc accgcacgc cccgcataaa	27660
tctcaacccc cgaaatttggc ccgcgcctt cgttaccag gaaacccctt ccgcaccac	27720
cgtactactt ccgcgtgacg cccaggccga agtccagatg actaactcag gggcgacgt	27780
cgcggcgcc tttcgatcag gggcgccggc gctccgatca ggtataagac acctgtatgt	27840
cagaggccga ggtatccagc tcaacgacga gtccgtggc tcttcgtctg gtctccgtcc	27900
ggacggact ttccagctcg ccggatccgg ccgccttcg ttacgcggcc gccaggcgta	27960
cctgactctg cagacctcg tctccggagcc ccgcctccgc ggcacatggaa ccctccagg	28020
cgtggaggag ttccgtccctt cggctactt caaccccttc tcgggaccc cccgcacgt	28080
cccccggccatc ttcatccga actttgacgc ggtgaaggac tcggcgacgc gctacgactg	28140
aatgtcagggt gtcgaggcag agcagctcg cctgagacac ctcgagcact gccgcgc	28200
caagtgcattt gcccgcgggtt ctggtgagtt ctgctacttt cagctacccg aggagcatac	28260

[0080]

cgaggggccg	gcmcacggcg	tccgcctgac	caccaggc	gaggttacct	gttccctcat	28320
ccggagttt	accctccgtc	ccctgctagt	ggagcgggag	cggggtccct	gtgtcctaac	28380
tatgcctgc	aactgcccta	accctggatt	acatcaagat	cttgctgtc	atctctgtgc	28440
ttagttaat	aaacgctgag	atcagaatct	actggggctc	ctgtcgccat	cctgtgaacg	28500
ccaccgtctt	caccacccc	gaccaggccc	aggcgaacct	cacctgcggt	ctgcatcgga	28560
ggccaagaa	gtacctcacc	tggtaacttca	acggcacccc	ctttgtggtt	tacaacagct	28620
tcgacggga	cggagtcctcc	ctgaaagacc	agctctccgg	tctcagctac	tccatccaca	28680
agaacaccac	cctccaactc	ttccctccct	acctgccggg	aacctacgag	tgcgtcaccg	28740
gccgctgcac	ccacctcacc	cgcctgatcg	taaaccagag	cttccggga	acagataact	28800
ccctcttccc	cagaacagga	ggtgagctca	ggaaactccc	cggggaccag	ggcggagacg	28860
tacttcgac	ccttgggg	ttaggatttt	ttattaccgg	gttgcggct	cttttaatca	28920
aagtttcctt	gagatttgg	cttccttct	acgtgtatga	acacctaacc	ctccaataac	28980
tctacccttt	tttcggaatc	aggtgacttc	tctgaaatcg	ggcttgggt	gctgcttact	29040
ctgttgattt	tttccttat	catactcagc	cttctgtgcc	tcaggctcgc	cgcctgctgc	29100
gcacacatct	atatctactg	ctgggtgctc	aagtgcaggg	gtcgccaccc	aagatgaaca	29160
ggtacatgg	cctatcgatc	ctaggcctgc	tggccctggc	ggcctgcagc	gccgcca	29220
aagagattac	cttggaggag	cccgcttgca	atgtaacttt	caagcccgg	ggtgaccaat	29280
gcaccaccct	cgtcaaatgc	gttaccaatc	atgagaggct	gcmcatacgac	tacaaaaca	29340
aaactggcca	gttgcggtc	tatagtgtgt	ttacccgg	agacccctct	aactactctg	29400
tcaccgtctt	ccagggcgga	cagtctaaga	tattcaatta	cactttccct	ttttatgagt	29460
tatgcgatgc	ggtcatgtac	atgtcaaaac	agtacaacct	gtggcctccc	tctccccagg	29520
cgtgtgtgga	aaatactggg	tcttactgct	gtatggcttt	cgcaatca	acgctcgctc	29580
taatctgcac	ggtgctatac	ataaaattca	ggcagaggcg	aatctttatc	gatgaaaaga	29640
aaatgccttg	atcgctaaca	ccggctttct	atctgcagaa	tgaatgcaat	cacccctta	29700
ctaataccacca	ccaccctcct	tgcgattgcc	catgggttga	cacgaatcga	agtgccagtg	29760
gggtccaaatg	tcaccatgg	gggccccgccc	ggcaattcca	ccctcatgtg	ggaaaaattt	29820
gtccgcaatc	aatgggttca	tttctgctct	aaccgaatca	gtatcaagcc	cagagccatc	29880
tgcgatggc	aaaatcta	tctgatcaat	gtgcaatga	tggatgctgg	gtactattac	29940

[0081]

ggcagcggg gagaatcat taattactgg cgacccaca aggactacat gctgcgtgt	30000
gtcgaggcac ttcccactac cacccccact accaccttc ccaccaccac caccactact	30060
actactacta ctactactac tactactacc actaccgctg cccgcatac cgcacaaagc	30120
accatgatta gcacaaagcc ccctcgtgct cactcccacg cggcgggccc catcggtgcg	30180
acccatcagaaa ccaccgagct ttgttttgc caatgcacta acgcgcgc tcatgaactg	30240
ttcgacctgg agaatgagga tgtccagcag agctccgctt gcctgaccca ggaggctgtg	30300
gagcccggtt ccctgaagca gatcggtgat tcaataattt acttttcttc ttttgcact	30360
cccgaaatacc ctcccgattt tactttccac atcacggta ccaaagaccc taacctctt	30420
ttcttacactga tgctgctgct ctgtatctt gtggtcttt ccgcgcgtat gttactgggg	30480
atgttctgct gcctgatctg ccgcagaaag agaaaagctc gctctcaggg ccaaccactg	30540
atgccttcc cttttttttt ggatttttgc aataacaaga tatgagctcg ctgctgacac	30600
taaccgcctt actagcctgc gctctaacc ttgtcggtt cgactcgaga ttccacaatg	30660
tcacagctgt ggcaggagaa aatgttactt tcaactccac ggccgatacc cagtggcgt	30720
ggagtggttc aggtagctac ttaactatct gcaatagctc cactttttt ggcataatccc	30780
caaccaagta ccaatgcaat gccagcgtt tcaccctcat caacgttcc accctggaca	30840
atggactcta tgtaggctat gtaccctttt gtgggcaagg aaagacccac gcttacaacc	30900
tggaaagttcg ccagccaga accactaccc aagcttctcc caccaccacc accaccacca	30960
ccatcaccag cagcagcagc agcagcagcc acagcagcag cagcagatta ttgactttgg	31020
ttttggccag ctcatctgcc gctaccagg ccatctacag ctctgtgtccc gaaaccactc	31080
agatccaccg cccagaaacg accaccgcca ccaccctaca cacccctcagc gatcagatgc	31140
cgaccaacat caccccttgc gctttcaaa tggacttac aagccccact ccaaaaccag	31200
tggatgcggc cgaggtctcc gccctcgatca atgactggc gggctggg atgtgggt	31260
tcgcctagg catgatggcg ctctgcctgc ttctgtctg gtcctatgc tgcctccacc	31320
gcaggcggc cagacccccc atctatagac ccatcattgt cctgaacccc gataatgatg	31380
ggatccatag attggatggc ctgaaaaacc tactttttt ttttacagta tgataaattt	31440
agacatgcct cgcattttct tgtacatgtt cttctccca cttttctgg ggtgttctac	31500
gctggccgct gtgtctcacc tggaggtaga ctgcctctca cccttcactg tctacctgt	31560
ttacggattt gtcaccctca ctctcatctg cagcctaatac acagtaatca tgccttcat	31620
ccagtgcatt gattacatct gtgtgcgcct cgcataactc agacaccacc cgcaagtaccg	31680

[0082]

agacaggaac attgccaaac ttcttaagact gctctaataca tgcataagac tgtgatctgc	31740
cttctgatcc tctgcataatccc gcccaccctc acctcctgcc agtacaccac aaaatctccg	31800
cgcaaaaagac atgcctcctg ccgcattcacc caactgtgga atataccaa atgctacaac	31860
gaaaagagcg agctctccga agcttggctg tatgggtca tctgtgtctt agtttctgc	31920
agcactgtct ttgcctcat aatctacccc tactttgatt tggatggaa cgcatcgat	31980
gccatgaatt accccacatt tccgcaccc gagataattc cactgcgaca agttgtaccc	32040
gttgcgtta atcaacgccc cccatcccct acgcccactg aaatcagcta cttaaccta	32100
acaggcggag atgactgacg ccctagatct agaaatggac ggcatacgta ccgagcagcg	32160
tctcctagag aggccgcaggc aggccgctga gcaagagcgc ctaatcagg agctccgaga	32220
tctcgtaac ctgcaccagt gcaaaaaggag catctttgt ctggtaaagc aggccaaagt	32280
cacctacgag aagaccggca acagccaccc cctcagttac aaattgcca cccagcgc	32340
gaagctggtg ctcatggtgg gtgagaatcc catcaccgtc acccagcact cggtagagac	32400
cgaggggtgt ctgcactccc cctgtcgaaa tccagaagac ctctgcaccc tggtaaagac	32460
cctgtgcggt ctcagagatt tagtcccatt taactaatca aacactggaa tcaataaaaa	32520
gaatcactta cttaaaatca gacagcaggc ctctgtccag tttattcagc agcacctcct	32580
tccctcctc ccaactctgg tactccaaac gccttctggc ggcaaaacttc ctccacaccc	32640
tgaaggaaat gtcagattct tgctctgtc cctccgcacc cactatctt atgtgttgc	32700
agatgaagcg cacaaaaacg tctgacgaga gttcaaccc cgtgtaccc tatgacacgg	32760
aaagcggccc tccctccgtc ctttctca cccctccctt cgtgtctccc gatggattcc	32820
aagaaagtcc cccggggtc ctgtctctga acctggccga gcccctggc acttcccacg	32880
gcatgctcgc cctgaaaatg ggaagtggcc tctccctgga cgacgctggc aacctcacct	32940
ctcaagatata caccaccgt agccctcccc tcaaaaaaac caagaccaac ctcagcctag	33000
aaacctcatc ccccctaact gtgagcaccc caggccctt caccgttagca gccggcgctc	33060
ccctggcggt ggccggcacc tccctcacca tgcaatcaga ggcccccctg acgtacagg	33120
atgcaaaaact caccctggcc accaaaggcc ccctgaccgt gtctgaaggc aaactggct	33180
tgcaaacatc ggccccgctg acggccgctg acagcagcac cctcacagtc agtgcacac	33240
cacccttag cacaagcaat ggcagcttgg gtattgacat gcaagcccc attacacca	33300
ccaatggaaa actaggactt aactttggcg ctcccctgca tgtggtagac agcctaaatg	33360

[0083]

cactgactgt agttactggc caaggtctta cgataaacgg aacagcccta caaactagag	33420
tctcaggtgc cctcaactat gacacatcgaa gaaacctaga attgagagct gcaggggta	33480
tgcgagttga tgcaaattgtt caacttatcc ttgatgttagc ttaccattt gatgcacaaa	33540
acaatctcgat ccttaggctt ggacagggac ccctgttgtt taactctgcc cacaacttgg	33600
atgttaacta caacagaggc ctctacatgt tcacatctgg aaataccaaa aagctagaag	33660
ttaatatcaa aacagccaag ggtctcattt atgatgacac tgctatagca atcaatgcgg	33720
gtgatggct acagtttgac tcaggctcgat atacaatcc attaaaaact aaacttggat	33780
taggactgga ttatgactcc agcagagcca taattgctaa actggaaact ggcctaagct	33840
ttgacaacac aggtgccatc acagtaggca acaaaaatga tgacaagctt accttggtt	33900
ccacaccaga cccatccccctt aactgttagaa tctattcaga gaaagatgct aaattcacac	33960
ttgtttgac taaatgcggc agtcagggtgt tggccagcgt ttctgttttta tctgtaaaag	34020
gtgccttgc gcccattcagt ggcacagtaa ctagtgccta gattgtcctc agatggat	34080
aaaatggagt tctactaagc aattttccc ttgaccctca atactggAAC tacagaaaag	34140
gtgaccccttac agagggact gcatatacca acgcagtggg atttatgccc aacccatcag	34200
cataccaaa aacacagagc caaaactgcta aaagcaacat tgtaagttagt gtttacttga	34260
atggggacaa atccaaaccc atgaccctca ccattaccctt caatggact aatgaaacag	34320
gagatgccac agtaaggact tactccatgt catttcatg gaactggaaat ggaagtaatt	34380
acattaatga aacgttccaa accaactcct tcaccccttc ctacatcgcc caagaataaa	34440
aagcatgacg ctgttgattt gattcaatgt gtttctgtttt tattttcaag cacaacaaaa	34500
tcattcaagt cattttcca tcttagcttta atagacacag tagtttaata gacccagtag	34560
tgcaaaggccc cattcttagct tatacatcgat acagtgtttaa ttaaccacca ccaccat	34620
accttttgcat tcaggaaatc atgatcatca caggatccta gtcttcaggc cgccccctcc	34680
ctcccaagac acagaataca cagtcctctc ccccccggactg gctttaaata acaccatctg	34740
gttggtcaca gacatgttct taggggttat attccacacg gtctcctgcc ggcggcaggcg	34800
ctcgctcggtg atgttataa actctccgg cagctcgctc aagttcacgt cgctgtccag	34860
cggtgtacc tccggctgac gcgataactg tgcgaccggc tgctggacga acggaggccg	34920
cgccatcataag gggtagatgttataatcctc ggtcaggata gggcggtgtt gcagcagcag	34980
cgagcgaaac atctgctgcc gccggcgctc cgtccggcag gaaaacaaca cggccgttgtt	35040
ctccctcccgca ataatccgca cggcccgccag catcagcttc ctcgttctcc ggcgcagca	35100

[0084]

cctcaccctt atctcgctca aatcgccgca gtaggtacag cacagcacca cgatgttatt	35160
catgateccca cagtgcaggg cgctgtatcc aaagctcatg ccggaaacca ccgcacccac	35220
gtggccatcg taccacaaggc gcacgtaaat caagtgtcga cccctcatga acgcgctgga	35280
cacaaacatt acttccttgg gcatgttgta attcaccacc tcccggtacc agataaacct	35340
ctggttgaac agggcacctt ccaccacca cctgaaccaa gaggccagaa cctgcccacc	35400
ggctatgcac tgcagggAAC ccgggttgga acaatgacaa tgcagactcc aaggctcgta	35460
accgtggatc atccggctgc tgaaggcatc gatgttgca caacacagac acacgtgcat	35520
gcactttctc atgattagca gctttccct cgtcaggatc atatcccaag gaataaccca	35580
ttcttgaatc aacgtaaaac ccacacagca ggaaaggcct cgcacataac tcacgttgc	35640
catggtcagc gtgttcatt ccggaaacag cggatgatcc tccagtatcg aggccgggt	35700
ctccttctca cagggaggta aagggtccct gctgtacgga ctgcgccggg acgaccgaga	35760
tcgtgtttag cgtagtgtca tggaaaagg aacgcggac gtggcatac ttcttgaagc	35820
agaaccaggt tcgcgcgtgg caggcctcct tgcgtctgcg gtctgccgt ctagctcgct	35880
ccgtgtgata gttgttagtac agccactccc gcagagcgac gaggcgcacc ctggcttccg	35940
gatctatgta gactccgtct tgcaccgcgg ccctgataat atccaccacc gtagaataag	36000
caacacccag ccaagcaata cactcgctc gcgagcggca gacaggagga gcggcagag	36060
atgggagaac catgataaaa aactttttt aaagaatatt ttccaattct tcgaaagtaa	36120
gatctatcaa gtggcagcgc tcccctccac tggcgcggc aaactctacg gccaaggcac	36180
agacaacggc atttctaaga tttccctaa tggcgtccaa aagacacacc gctctcaagt	36240
tgcagtaaac tatgaatgaa aaccatccg gctgatttc caatatacgc gcgcggcag	36300
cgtccaccaa acccagataa tttcttctc tccagcggt tacgatctgt ctaagcaa	36360
cccttatatac aagtccgacc atgcaaaaaa tctgctcaag agcgccctcc accttcatgt	36420
acaaggcagcg catcatgatt gcaaaaattc agttttca gagacctgta taagattcaa	36480
aatgggaaca ttaacaaaaa ttccctgtc ggcgcgatcc ctgcgcaggaa caagctgaac	36540
ataatcagac aggtccgaac ggaccagtga ggcacaaatcc ccaccaggaa ccagatccag	36600
agaccctata ctgattatga cgcgcatact cggggctatg ctgaccagcg taggcgcgt	36660
gtaggcgtgc tgcattggcg gcgagataaa atgcaagtg ctggtaaaa aatcaggcaa	36720
agcctcgccgca aaaaaagcta acacatcata atcatgctca tgcaggttagt tgcaggtaag	36780

[0085]

ctcaggaacc	aaaacggaat	aacacacgat	tttcctctca	aacatgactt	cgcgatact	36840
gcgtaaaaca	aaaaattata	aataaaaaat	taattaaata	acttaaacat	tggaaggctg	36900
tctcacaaca	ggaaaaacca	cttaatcaa	cataagacgg	gccacgggca	tgccggcata	36960
gccgtaaaaa	aattggtccc	cgtgattaac	aagtaccaca	gacagctccc	cggtcatgtc	37020
gggggtcatc	atgtgagact	ctgtatacac	gtctggattg	tgaacatcag	acaaacaaag	37080
aaatcgagcc	acgtagcccg	gaggtataat	cacccgcagg	cgaggtaca	gcaaaacgac	37140
ccccatagga	ggaatcacaa	aatttagtagg	agaaaaaaat	acataaacac	cagaaaaacc	37200
ctgttgctga	ggcaaaatag	cgcctcccg	atccaaaaca	acataaagcg	cttccacagg	37260
agcagccata	acaaagaccc	gagtcttacc	agtaaaagaa	aaaagatctc	tcaacgcagc	37320
accagcacca	acacttcgca	gtgtaaaagg	ccaagtgccg	agagagtata	tataggaata	37380
aaaagtgacg	taaacgggca	aagtccaaaa	aacgcccaga	aaaacccgac	gcaaacctac	37440
cccccgaaac	gaaagccaaa	aaacactaga	cactcccttc	cgcgtaac	ttccgcttcc	37500
ccacgctacg	tcacttgccc	cagtcaaaca	aactacatat	cccgaacttc	caagtcgcca	37560
cgcccaaacc	accgcctaca	cctccccgcc	cgccggcccg	cccccaaacc	cgcctccgc	37620
cccgccggcc	gccccgccc	gccccatctca	ttatcatatt	ggcttcaatc	caaaataagg	37680
tatattattt	atgatggttt	aaacggatcc	tctagagtgc	acctgcaggc	atgcaagctt	37740
gagtattcta	tagtgtcacc	taaatagctt	ggcgtaatca	tggcatagc	tgtttcctgt	37800
gtgaaattgt	tatccgctca	caattccaca	caacatacga	gccggaagca	taaagtgtaa	37860
agcctgggt	gcctaattgag	tgagctaact	cacattaatt	gcgttgcgct	cactgcccgc	37920
tttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	atcggccaac	gcaaccct	37980
tgccggccgc	cggcccgctcg	accaattctc	atgtttgaca	gcttatcatc	gaatttctgc	38040
cattcatccg	cttattatca	cttattcagg	cgtagcaacc	aggcgttaa	gggcaccaat	38100
aactgcctta	aaaaaattac	gccccgcct	gccactcatc	gcagtactgt	tgtattcat	38160
taagcattct	gccgacatgg	aagccatcac	aaacggcatg	atgaacctga	atgccagcg	38220
gcatcagcac	cttgtcgct	tgcgtataat	atttgcccatt	ggtgaaaacg	ggggcgaaga	38280
agttgtccat	attggccacg	tttaaatcaa	aactggtgaa	actcacccag	ggattggctg	38340
agacgaaaaaa	catattctca	ataaacccctt	tagggaaata	ggccaggttt	tcaccgtaac	38400
acgccacatc	ttgcgaatat	atgtgtagaa	actgccggaa	atcgctgtgg	tattcactcc	38460
agagcgatga	aaacgttca	gtttgctcat	ggaaaacggt	gtaacaaggg	tgaacactat	38520

[0086]

cccatatcac cagctcacccg tcttcattt ccatacgaa ttccggatga gcattcatca	38580
ggcgggcaag aatgtgaata aaggccggat aaaacttgtt cttattttc tttacggtct	38640
ttaaaaaggc cgtaatatcc agctgaacgg tctggttata ggtacattga gcaactgact	38700
gaaatgcctc aaaatgttct ttacgatgcc attggatata atcaacggtg gtatatccag	38760
tgatTTTTT ctccattta gcttccttag ctcctgaaaa tctcgataac tcaaaaaata	38820
cgcggtag tgatcttatt tcattatggt gaaagttgga acctcttacg tgccgatcaa	38880
cgtctcattt tcgccaaaag ttggcccagg gcttcccgtt atcaacaggg acaccaggat	38940
ttatTTTTC tgcaagtga tcttccgtca caggtattta ttgcgcataa gctcatggag	39000
cggcgtaacc gtcgcacagg aaggacagag aaagcgcgga tctggaaat gacggacaga	39060
acggtcagga cctggattgg ggaggcggtt gccgcgcgt ctgctgacgg tgtgacgttc	39120
tctgttccgg tcacaccaca tacgttccgc cattctatg cgatgcacat gctgtatgcc	39180
ggtataccgc tgaaagttct gcaaagcctg atggacata agtccatcag ttcaacggaa	39240
gtctacacga aggttttgc gctggatgtg gctgccggc accgggtgca gtttgcgt	39300
ccggagtctg atgcgggtgc gatgctgaaa caattatcct gagaataat gccttggcct	39360
ttatATGAA atgtgaaact gagtggatata gctgttttg tctgttaaac agagaagctg	39420
gctgttatcc actgagaagc gaacgaaaca gtcggaaaaa tctccattt tcgttagagat	39480
ccgcattatt aatctcagga gcctgtgtag cgttatagg aagtagtgtt ctgtcatgat	39540
gcctgcaagc ggtacgaaa acgatttggaa tatgccttca ggaacaatag aaatttcg	39600
gcgggtttac gttgaagtgg agcggattat gtcagcaatg gacagaacaa cctaataac	39660
acagaaccat gatgtggtct gtcctttac agccagtagt gctgccgca gtcgagcgc	39720
agggcgaagc cctcgagtga gcgaggaagc accaggaaac agcacttata tattctgct	39780
acacacgatg cctgaaaaaa ctcccttgg gtttatccac ttatccacgg ggtatTTTT	39840
ataattattt ttttatagt ttttagatct tcttttttag agcgccttgtt aggcctttat	39900
ccatgctggt tctagagaag gtgttgtgac aaattgcct ttcagtgtga caaatcaccc	39960
tcaaatacaca gtcctgtctg tgacaaattt cccttaaccc tgtgacaaat tgccctcaga	40020
agaagctgtt tttcacaaa gttatccctg cttattgact ctTTTTTATT tagtgtgaca	40080
atctaaaaac ttgtcacact tcacatggat ctgtcatggc ggaaacagcg gttatcaatc	40140
acaagaaacg taaaaatagc ccgcgaatcg tccagtcataa cgacctcact gagggggcat	40200

[0087]

atagtctctc ccgggatcaa aaacgtatgc tgtatctgtt cggtgaccag atcagaaaaat	40260
ctgatggcac cctacaggaa catgacggta tctgcgagat ccatgttgct aaatatgctg	40320
aaatattcgg attgacctct gcggaaagcca gtaaggatat acggcaggca ttgaagagtt	40380
tcgcgggaa ggaagtgggt ttttatcgcc ctgaagagga tgccggcgat gaaaaaggct	40440
atgaatctt tccttggttt atcaaacgtg cgcacagtcc atccagaggg ctttacagtg	40500
tacatatcaa cccatatctc attcccttct ttatcgggtt acagaaccgg tttacgcagt	40560
ttcggcttag tgaaacaaaaa gaaatcacca atccgtatgc catgcgtta tacgaatccc	40620
tgtgtcagta tcgtaagccg gatggctcag gcatcgctc tctgaaaatc gactggatca	40680
tagagcgta ccagctgcct caaagttacc agcgtatgcc tgacttccgc cgccgcttcc	40740
tgcaggtctg tgttaatgag atcaacagca gaactccaat gcgcctctca tacattgaga	40800
aaaagaaaagg ccgccagacg actcatatcg tatttcctt ccgcgatatc acttccatga	40860
cgacaggata gtctgaggggt tatctgtcac agatttgagg gtggttcgac acatttggc	40920
tgacctactg agggtaattt gtcacagttt tgctgttcc ttcaagctgc atggattttc	40980
tcatacttt tgaactgtaa ttttaagga agccaaattt gagggcagtt tgtcacagtt	41040
gatttccttc tcttcctt cgtcatgtga cctgatatcg gggtagtt cgtcatcatt	41100
gatgagggtt gattatcaca gtttattact ctgaattggc tatccgcgtg tgtacctcta	41160
cctggagttt ttcccacggt ggatatttct tcttgcgtg agcgtaaagag ctatctgaca	41220
gaacagttct tcttgcttc ctgcgcagtt cgctcgctat gctcggttac acggctgcgg	41280
cgagcgctag tgataataag tgactgaggt atgtgcctt cttatctctt tttgttagtgt	41340
tgctcttatt ttaaacaact ttgcggtttt ttgatgactt tgcgattttt ttgttgcttt	41400
gcagtaaatt gcaagattta ataaaaaaac gcaaagcaat gattaaagga tgttcagaat	41460
gaaactcatg gaaacactta accagtgcattt aaacgctggc catgaaatga cgaaggctat	41520
cgcatttgcattt cagtttaatg atgacagccc ggaagcgagg aaaataaccc ggctggag	41580
aataggtgaa gcagcggatt tagttgggt ttcttctcag gctatcagag atgccgagaa	41640
agcagggcga ctaccgcacc cggatatgga aattcgagga cgggttgcac aacgtgttgg	41700
ttatacaatt gaacaaatta atcatatgcg tgatgtgtt ggtacgcgtat tgacgtgc	41760
tgaagacgta ttccacccgg tgatcgggt tgctgcccattt aaaggtggcg tttacaaaac	41820
ctcagttctt gttcatcttgc ctcaggatct ggctctgaag gggctacgtg tttgctgt	41880
ggaaggtaac gaccccccagg gaacagcctc aatgtatcac ggatgggtac cagatctca	41940

[0088]

tattcatgca gaagacactc tcctgcctt ctagttggg gaaaaggacg atgtcactta	42000
tgcaataaag cccacttgct ggccggggct tgacattatt cttcctgct tggctctgca	42060
ccgtattgaa actgagttaa tggcaaaatt tggatgggt aaactgccc cccatccaca	42120
cctgatgctc cgactggcca ttgaaaactgt tgctcatgac tatgatgtca tagttattga	42180
cagcgcgcct aacctggta tcggcacat taatgtcgta tgtgctgctg atgtgctgat	42240
tgttcccacg cctgctgagt tggatggacta cacctccgca ctgcagttt tcgatatgct	42300
tcgtgatctg ctcagaacg ttgatcttaa agggttcgag cctgatgtac gtatggct	42360
taccaaatac agcaatagta atggctctca gtcccggtt atggaggagc aaattcggg	42420
tgcctggga agcatggttc taaaaaatgt tggatgtgaa acggatgaag ttggtaaagg	42480
tcagatccgg atgagaactg ttttgaaca ggcattgtat caacgcttt caactggc	42540
ctggagaaat gctttctta tttgggacc tggatgtcaat gaaatttcg atcgatgtat	42600
taaaccacgc tgggagatta gataatgaag cgtgcgcctg ttattccaaa acatacgctc	42660
aatactcaac cggttgaaga tacttcgtt tcgacaccag ctgccccat ggtggattcg	42720
ttaattgcgc gcgttaggat aatggctcgc ggtatgcca ttacttgcg tggatgtggt	42780
cggatgtga agtttactt tggatgtcgc cgggtgata gtggatggaa gacatctcg	42840
gtatggtcag gtaatgaacg tgaccaggag ctgcttactg aggacgcact ggtatgttc	42900
atcccttctt ttctactgac tggtaacag acaccggcgt tcggatcgaa agtattgtgt	42960
gtcatagaaa ttggcgatgg gagtcgcccgt cgtaaagctg ctgcacttac cgaaagtgtat	43020
tatcgatgttc tggatggcgat gctggatgtat gagcagatgg ctgcatttac cagattgggt	43080
aacgatttac gcccaacaag tgcttatgaa cgtggcagc gttatgcaag ccgattgcag	43140
aatgaatttg ctggaaatat ttctgcgt gctgatgcgg aaaatatttc acgtaaagatt	43200
attacccgtt gttatcaacac cgccaaatttgc cttaaatcag ttgttgcctt ttttctcac	43260
cccggtgaac tatctgcccgt gtcaggtgat gcacttacaaa aacccatttac agataaagag	43320
gaattactta agcagcaggc atctaacctt catgagcaga aaaaagctgg ggtatgtt	43380
gaagctgaag aagttatcac tcttttaact tctgtgcctt aaacgtcatc tgcataaga	43440
actagttaa gctcacgaca tcagttgcgt cctggagcga cagtattgtt taagggcgat	43500
aaaatggtc ttaacctgga caggtctcgat gttccaaactg agtgtataga gaaaatttgag	43560
gccattctta aggaacttga aaagccagca ccctgatgcg accacgtttt agtctacgtt	43620

[0089]

tatctgtctt tacttaatgt cctttgttac aggccagaaa gcataactgg cctgaatatt	43680
ctctctggc ccactgttcc acttgtatcg tcggctctgat aatcagactg ggaccacggt	43740
cccaactcgta tcgtcggtct gattattagt ctgggaccac ggtcccactc gtatcgctgg	43800
tctgattatt agtctgggac cacggtccca ctcgtatcgta cggtctgata atcagactgg	43860
gaccacggtc ccactcgat cgtcggtctg attattagtc tgggaccatg gtcccaactcg	43920
tatcgctggc ctgattatta gtctgggacc acggtcccac tcgtatcgta ggtctgatta	43980
ttagtctgga accacggtcc cactcgatc gtcggctga ttattagtct gggaccacgg	44040
tcccaactcgat atcgctggc tgattattag tctgggacca cgatcccact cgtgttgc	44100
gtctgattat cggtctgggaa ccacggtccc acttgtattt tcgatcagac tatcagcgtg	44160
agactacgat tccatcaatg cctgtcaagg gcaagtattt acatgtcgta gtaacctgta	44220
gaacggagta acctcggtgt gcgggtgtat gcctgctgtg gattgctgct gtgtcctgct	44280
tatccacaac attttgcgca cggttatgtt gacaaaatac ctggttaccc aggccgtgcc	44340
ggcacgttaa ccgggctgca tccgatgcaa gtgtgtcgct gtgcacgagc tgcgagctc	44400
ggacatgagg ttgccccgta ttcaagtgtcg ctgatttgc ttgtctgaag ttgttttac	44460
gttaagttga tgcagatcaa ttaatacgat acctgcgtca taattgatta tttgacgtgg	44520
tttgcgttgc tccacgcacg ttgtatgtg tagatgataa tcattatcac tttacgggtc	44580
ctttccggcgt atccgacagg ttacggggcg gcgcacctgc gggtttcgc tatttatgaa	44640
aattttccgg tttaaggcgt ttccgttctt ctgcgtcata acttaatgtt tttatttaaa	44700
ataccctctg aaaagaaaagg aaacgacagg tgctgaaagc gagcttttg gcctctgtcg	44760
tttcctttct ctgttttgcgtt ccgtggaatg aacaatggaa gtccgagctc atcgctaata	44820
acttcgtata gcatacatta tacgaagttt tattcgatgc ggcgcgcaagg ggttcgcgtc	44880
agcgggtgtt ggcgggtgtc gggctggct taactatgcg gcatcagagc agattgtact	44940
gagagtgcac catatgcgtt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat	45000
caggcgccat tcgcccattca ggctgcgca ctgttggaa gggcgatcgg tgcggccctc	45060
ttcgcttattt cggccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac	45120
gccagggttt tcccaagtac gacgttgc aacgcacggcc agtgaattgt aatacgactc	45180
actatagggc gaattcgagc tcggtaaccgg gggatcctcg tttaaac	45227

&lt;210&gt; 10

&lt;211&gt; 37830

[0090]

&lt;212&gt; DNA

&lt;213&gt; 黑猩猩腺病毒

<400> 10		
catcatcaat aatacacatt atttggatt gaagccaata tgataatgag atggcggcg	60	
cggggcggga ggcgggtccg gggcgggccc ggcgggggg gcgggtgtggc ggaagtggac	120	
ttttaagtg tggcgatgt gacttgctag tgccggcgc ggtaaaagtg acgtttccg	180	
tgcgacaa cgcccacggg aagtgacatt tttccgcgg ttttaccgg atgtttagt	240	
gaatttggc gtaaccaagt aagatttggc catttcgcg ggaaaactga aacgggaag	300	
tgaatctga ttaatttgcg gttagtcata ccgcgtaata tttgtcgagg gccgagggac	360	
tttggccgat tacgtggagg actcgccag gtgttttg aggtgaattt ccgcgttccg	420	
ggtcaaagtc tccgtttat tattatagtc agctgacgag gagtgtat ataccctctg	480	
atctcgtaa gtggccactc ttgagtgcca gcgagtagag tttctctc tgccgctctc	540	
cgctccgctc cgctcgctc tgacaccggg gaaaaaatga gacatttcac ctacgatggc	600	
ggtgtgctca cggccagct ggctgctgaa gtcctggaca ccctgatcga ggaggtattt	660	
gccgataatt atcctccctc gactccttt gagccaccta cacttcacga actctacgat	720	
ctggatgtgg tggggccag cgatccgaac gagcaggcgg tttccagttt tttccagag	780	
tccatgttgt tggccagcca ggaggggtc gaacttgaga cccctctcc gatcgtggat	840	
tcccccgatc cgccgcagct gactaggcag cccgagcgct gtgcgggacc tgagactatg	900	
ccccagctgc tacctgaggt gatcgatctc acctgtaatg agtctggttt tccacccagc	960	
gaggatgagg acgaagaggg tgagcagttt gtgttagatt ctgtggaca acccggcga	1020	
ggatgcaggt ctgtcaata tcacggaaa aacacaggag actcccagat tatgtttct	1080	
ctgtgtata tgaagatgac ctgtatgtt atttacagta agtttatcat ctgtggcag	1140	
gtggctata gtgtgggtgg tggctttgg ggggtttttt aatatatgtc aggggttatg	1200	
ctgaagactt ttttattgtt attttaaag gtccagtgac tgagcccgag caagaacctg	1260	
aaccggagcc tgagccttct cgccccagga gaaagcctgt aatcttaact agacccagcg	1320	
caccggtagc gagaggcctc agcagcgcgg agaccaccga ctccgggtct tcctcatcac	1380	
ccccggagat tcacccctg gtccccctgt gtcccgttaa gcccggtgcc gtgagagtca	1440	
gtggcggcgc gtctgctgtg gagtcatttgg aggacttgct ttttgcattca caggaacctt	1500	
tggacttgag cttgaaacgc cccaggcatt aaacctggtc acctggactg aatgagttga	1560	
cgcctatgtt tgctttgaa tgacttaatg tgtatagata ataaagagtg agataatgtt	1620	

[0091]

ttaattgcat ggtgtgtta acttggcggt	1680
taaacttgt tacacttgac ctcatggagg cctggagtg tttggagaac tttgccggag	1740
ttcgtgcctt gctggacgag agctctaaca ataccttgc gtggtgagg tatttgtgg	1800
gctctccca gggcaagttt gttttagaa tcaaggagga ttacaagtgg gaatttgaag	1860
agctttgaa atcctgttgtt gagctattgg attcttgaa tctaggccac caggctctt	1920
tccaggagaa ggtcatcagg actttggatt tttccacacc ggggcgcatt gcagccgcgg	1980
ttgctttctt agctttttt aaggatagat ggagcgaaga gacccacttg agttcggct	2040
acgtcctgga tttctggcc atgcaactgt ggagagcatg gatcagacac aagaacaggc	2100
tgcaactgtt gtcttcgtc cggccgttgc tgattccggc ggaggagcaa cagggcggt	2160
cagaggaccg ggcccgtcgg gatccggagg agagggcacc gaggccggc gagaggagcg	2220
cgctgaacctt gggAACCGGG ctgagcggcc atccacatcg ggagtgaatg tcggcaggt	2280
ggtgatctt tttccagaac tgcggcgat tttgactatt agggaggatg ggcaatttgc	2340
taagggtctt aagaggaga gggggcttc tgagcataac gaggaggcca gtaatttgc	2400
tttagctt atgaccagac accgtccaga gtgcatact tttcagcaga ttaaggacaa	2460
tttgccaat gagttggatc tgggggtca gaagtatagc atagagcagc tgaccactta	2520
ctggctgcag ccgggtgatg atctggagga agctattagg gtgtatgcta aggtggccct	2580
cgccccgat tgcaagtaca agctcaaggg gctggtaat atcaggaatt gttgtacat	2640
ttctggcaac gggcggagg tggagataga gaccgaagac agggtggctt tcagatgcag	2700
catgtaat atgtggccgg ggggtctggg catggacggg gtggtgatc tgaatgtgag	2760
gttcacgggg cccaaacttta acggcacggt gttttgggg aacaccaacc tggcctgca	2820
cggggtgagc ttctatgggt ttaacaacac ctgtgtggag gcctggaccc atgtgaaggt	2880
ccgcgggtgc gcctttatg gatgttgaa gccatagtg agccgcctt agagcaggag	2940
ttccattaag aaatgcttgt ttgagaggtt cacctgggg atcctggccg agggcaactg	3000
cagggtgcgc cacaatgtgg cctccgagtg cggttgcctt atgctagtc agagcgtggc	3060
ggtaatcaag cataatatgg tgtcggcaa cagcgaggac aaggcctcac agatgctgac	3120
ctgcacggat ggcaactgcc acttgctgaa gaccatccat gtaaccagcc acagccggaa	3180
ggcctggccc gtgttcgagc acaacttgct gacccgtgc tccttgcatc tggcaacag	3240
gcgggggggtg ttccctggccct atcaatgcaa cttagtcac accaagatct tgctagagcc	3300

[0092]

cgagagcatg tccaagggtga acttgaacgg ggtgtttgac atgaccatga agatctggaa	3360
ggtgctgagg tacgacgaga ccaggtcccg gtgcagaccc tgcgagtgcg gggcaagca	3420
tatgaggaac cagcccgta tgctggatgt gaccgaggag ctgaggacag accacttggt	3480
tctggcctgc accagggccg agtttggttc tagcgatgaa gacacagatt gaggtgggtg	3540
agtgggcgtg gcctgggtg gtcatgaaaa tatataagtt ggggtctta gggctcttt	3600
atttgtttg cagagaccgc cggagccatg agcgggagca gcagcagcag cagtagcagc	3660
agcgccttgg atggcagcat cgtgagccct tatttgacga cgcggatgcc ccactggcc	3720
ggggtgcgtc agaatgtgat gggctccagc atcgacggcc gacccgtcct gcccgc当地	3780
tccgccacgc tgacctatgc gaccgtcgcg gggacggcgt tggacgccac cgccgc当地	3840
ccactggc当地 cagccgc当地 ggcgtcgcg agcctggcca cggacttgc attcctggg	3900
ccactggc当地 cagggctac ttctcggcc gctgctgc当地 ccgttgc当地 tgacaagctg	3960
accgcctgc tggcgc当地 ggatgc当地 actcgggaaac tgggtgacct ttctcagcag	4020
gtcatggccc tgc当地 cagca ggtctccctcc ctgcaagctg gcggaaatgc ttctccc当地	4080
aatgccgtt aagataaata aaaccagact ctgtttggat taaaagaaaag tagcaagtgc	4140
attgctctct ttatttc当地 atttccgc当地 cgc当地 tagaccag cgttctcggt	4200
cgttggggt gcgggtgtatc ttctccagga cgtggtagag gtggctctgg acgttggagat	4260
acatgggcat gagccgc当地 cgggggtgga gtagcacca ctgc当地 gagct tcatgc当地	4320
gggtgggtt gtagatgatc cagtc当地 tagc aggagcgctg ggc当地 ggtgc当地 ctaaaaatgt	4380
cttcagcag caggccgatg gccaggggga ggc当地 ttgggtt gtaagtgtt acaaaaacgg	4440
taagttggga agggtgc当地 cggggagaga ttagtgc当地 cttggactgt atttttagat	4500
tggc当地 atgtt tccccc当地 aatttgc当地 gatgc当地 gtgc当地 aggacc accagtagc	4560
tgtatccggt gcacttgggg aatttgc当地 gca gtttgc当地 gggaaaagcg tggaagaact	4620
tggagacgcc tttgtggc当地 cccagat当地 ccatgc当地 ttccatgc当地 atggcaatgg	4680
gcccgccggga ggc当地 ggtt gcaaagat当地 ttctggggtc gctgacgtcg tagttgtt	4740
ccagggttag gtc当地 catag gcaat当地 tta caaagcgccg gcggagggtg cccgactgg	4800
ggatgatggt cccctctggc cctggggcgt agttgc当地 ctc gc当地 atgttccagg	4860
ccttaatctc ggagggggga atcatatcca cctgc当地 ggtt gatgaagaaa acggttccg	4920
gagccggggga gattaactgg gatgagagca gtttctaaag cagctgtgat ttccacaac	4980
cggtggccc ataaataaca cctataaccg gttgc当地 gtagtttaga gagctgc当地	5040

[0093]

tgccgtcgctc	ccggaggagg	ggggccacct	cgttgagcat	gtccctgacg	cgcatgttct	5100
ccccgaccag	atccgccaga	aggcgctcgc	cgcgcaggga	cagcagctct	tgcaaggaag	5160
caaagttttt	cagcggcttg	aggccgtccg	ccgtggcat	gttttcagg	gtctggctca	5220
gcagctccag	gcggtcccag	agctcggtga	cgtgcctac	ggcatctcta	tccagcatat	5280
ctcctcgttt	cgcgggttgg	ggcgacttcc	gctgttagggc	accaagcggt	ggtcgccag	5340
cggggccaga	gtcatgtcct	tccatggcgc	cagggcctc	gtcagggtgg	tctgggtcac	5400
ggtaagggg	tgcgctccgg	gctgagcgct	tgccaaagggt	cgctttaggc	tggttctgct	5460
ggtgctgaag	cgctgcccgt	cttcgcctg	cgcgtccgc	aggtagcatt	tgaccatgg	5520
gtcatagtcc	agcccccctccg	cggcgtgtcc	cttggcgcgc	agcttgcct	tggaggtggc	5580
gccgcacgag	gggcagagca	ggctctttag	cgcgtagagc	ttggggcga	ggaagaccga	5640
ttcgggggag	taggcgtccg	cggcgcagac	ccgcacacg	gtctcgact	ccaccagcca	5700
ggtgagctcg	gggcgcgcgc	ggtcaaaaaac	caggttccc	ccatgctttt	tgtcgcttt	5760
cttacctcg	gtctccatga	ggtggtgtcc	ccgctcggt	acgaagaggc	tgtccgtgtc	5820
tccgtagacc	gacttgaggg	gtctttctc	caggggggtc	cctcggtt	cctcgtagag	5880
gaactcgac	cactctgaga	cgaaggcccg	cgtccaggcc	aggacgaagg	aggctatgt	5940
ggaggggtag	cggtcgttgt	ccactagggg	gtccaccttc	tccaagggtgt	gaagacacat	6000
gtcgccttcc	tcggcgtcca	ggaagggtat	tggctgttag	gtgttaggcca	cgtgaccggg	6060
ggttcctgac	gggggggtat	aaaaggggg	ggggcgcgc	tcgtcgac	tctttccgc	6120
atcgctgtct	gcgagggcca	gctgctgggg	ttagtattcc	ctctcgaagg	cggcatgac	6180
ctccgcgtg	aggttgtcag	tttccaaaaa	cgaggaggat	ttgatgttca	cctgtcccga	6240
ggtgataacct	ttgagggtac	ccgcgtccat	ctggtcagaa	aacacgatct	ttttattgtc	6300
cagcttggtg	gcgaacgacc	cgtagagggc	gttggagagc	agcttggcga	tggagcgcag	6360
ggtctggttc	ttgtccctgt	cggcgcgtc	cttggccgc	atgtttagct	gcacgtact	6420
gcgcgcgacg	cagcgccact	cgggaaagac	ggtggtgcc	tcgtcggca	ccagcgcac	6480
gcgccagccg	cgttggtca	gggtgaccag	gtccacgct	gtggcgacct	cgccgcgcag	6540
gcgcgttg	gtccagcaga	gacggccgccc	cttgcgcgag	cagaaggggg	gcaggggtc	6600
gagctgggtc	tgtccgggg	ggtccgcgtc	cacggtaaaa	accccggggc	gcaggcgcgc	6660
gtcgaagtag	tctatcttgc	aaccttgc	gtccagcgc	tgctgccagt	cgcggcggc	6720

[0094]

gagcgcgcgc tcgttaggggt tgagcggcgg gccccagggc atgggtggg tgagtgcgg	6780
ggcgtacatg ccgcagatgt catagacgt aaggggctcc cgcaaggaccc cgatgttaggt	6840
ggggtagcag cggccggccgc ggatgctggc ggcacgtac tcatacagct cgtgcgaggg	6900
ggcgaggagg tcggggccca gtttgtcgc ggcggggcgc tccgcgcgg aacgatctg	6960
cctgaagatg gcatgcgagt tggaaagat ggtggggcgc tggaaagacgt tgaagctggc	7020
gtcctgcagg ccgacggcgt cgccacgaa ggaggcgtag gagtcgcgc gcttgtgtac	7080
cagctcggcg gtgacctgca cgtcgagcgc gcagtagtcg agggtctcgc ggatgtatgtc	7140
atatttagcc tgccccttct tttccacag ctcgcggttt aggacaaact cttcgccgtc	7200
tttccagttac tcttggatcg gaaaccgtc cggttccgaa cgtaagagc ctagcatgt	7260
gaactggttt acggccttgtt aggcgcagca gccccttcc acggggaggg cgtaggcctg	7320
cgcggccttg cggagcggagg tgtgggtcag ggcgaagggtg tccctgacca tgactttgag	7380
gtactggtgc ttgaagtctgg agtcgtcgc gcccggccgc tccagagcg agaagtcggt	7440
gcgcttcttgc ggggggtt tggcagagc gaaggtgaca tcgttgaaga ggattttgcc	7500
cgcgcggggc atgaagttgc gggatgcg gaaggcccc ggcacttcag agcggttgtt	7560
gatgacctgg gcccgcgacgca cgatctcgta gaagccgttgc atgttgtggc ccacgatgt	7620
gagttccagg aagcggggcc gcccatttac ggtgggcagc ttcttagct cttcgtaggt	7680
gagctcctcg ggcgaggcga ggcgtgctc ggcaggccc cagtcgcga ggtgcgggtt	7740
gtctctgagg aaggacttcc agaggtcgcg ggcaggagg gtctgcaggc ggtctctgaa	7800
ggtcctgaac tggcggccca cggccatttt ttcgggggtt atgcagtaga aggtgagggg	7860
gtcttgcgtc cagcggtccc agtcgagctg cagggcgagg tcgcgcgcgg cggtgaccag	7920
gcgctcgctg ccccgaaatt tcatgaccag catgaaggc acgagctgtt ttccgaaggc	7980
cccccatttca gtgttaggtct ctacatcgta ggtacaaag aggcgcctcg tgcgaggatg	8040
cgagccgatc gggaaact ggatctcccg ccaccagttg gaggagtggc tggatgtg	8100
gtgaaagtag aagtccgtc gcccggccga acactcggtc tggctttgt aaaagcgagc	8160
gcagttactgg cagcgctgca cggcgtgtac ctcgtcgcg agatgcaccc ttgcggccgc	8220
cacgaggaag ccgaggggaa atctgagccc cccgcgttgc tcgcggcatg gctggtttc	8280
ttctactttt gatgcgtgtc cgtctccgtc tggctcctcg aggggtgtta cggtgagcg	8340
gaccaccacg ccgcgcgagc cgcagggtcca gatatcgccg cgcggcggtc ggagttgtat	8400
gacgacatcg cgcagctggg agctgtccat ggtctggagc tcccgccggc gcggcaggc	8460

[0095]

agccggagt tcttcaggt tcacctcgca gagtcgggcc agggcgcgg gcaggtctag	8520
gtggtacctg atctctaggg gcgtgttggg ggcggcgtcg atggcttgca ggagcccgca	8580
gccccggggg gcgacgacgg tgccccgcgg ggtgggtggt gtggtggcgg tgcagctcag	8640
aagcggtgcc gcgggcgggc ccccgaggt agggggggct ccggtcccgc gggcagggc	8700
ggcagcggca cgtcggcgtg gagcgcggc aggagttggt gctgtgcccg gaggttgctg	8760
gcgaaggcga cgacgcggcg gttgatctcc tggatctggc gcctctgcgt gaagacgacg	8820
ggcccggtga gcttgaacct gaaagaggt tcgacagaat caatctcgt gtcattgacc	8880
gcggcctggc gcaggatctc ctgcacgtct cccgagttgt cttggtaggc gatctggcc	8940
atgaactgct ccatctttc ctcctggagg tctccgcgtc cggcgcgttc cacgtggcc	9000
gccaggtcgt tggagatgcg ccccatgagc tgcgagaagg cgttgagtcc gccctcggtc	9060
cagactcggc tgttagaccac gccccctgg tcatcgcggc cgccatgac cacgtgcgc	9120
aggtttagt ccacgtgccg cgccaaagacg gcgttagttgc gcagacgctg gaagaggttag	9180
ttgagggtgg tggcggtgtc ctcggccacg aagaagttca tgacccagcg ggcacacgt	9240
gattcgttga tgtccccaa ggcctccagc cgttccatgg cctcgtagaa gtccacggcg	9300
aagttgaaaa actggagtt ggcgcgcgac acggtaact ctcctccag aagacggatg	9360
agctcggcga cgggtgcgcg cacctcgcgc tcgaaggcta tgggatctc ttccctcgct	9420
agcatcacca ctcctccctc ttccctctct tctggactt ccatgatggc ttccctctct	9480
tcgggggtg gggcgccgg cggtggggga gggggcgctc tgcgcggcg gggcgccacc	9540
gggaggcggt ccacgaagcg cgccatcatc tccccgcggc ggcggcgcat ggtctcggt	9600
acggcgcggc cgttctcccg gggcgcaagt tggaagacgc cgccggacat ctggtgcgt	9660
ggcgggtggc cgtgaggcag cgagacggcg ctgacgatgc atctcaacaa ttgctgcgt	9720
ggtacgcgcg cggggacact gagggagtcc atatccaccc gatccgaaaa ctttcgagg	9780
aaggcgtcta accagtcga gtcgcaaggt aggctgagca ccgtggcggc ggggggggg	9840
tggggggagt gtctggcga ggtgctgctg atgatgtaat tgaagtaggc ggacttgaca	9900
cggcggatgg tcgacaggag caccatgtcc ttgggtccgg cctgctggat gcggaggcgg	9960
tcggctatgc cccaggcttc gttctggcat cggcgcaggc cttgttagta gtctgcatt	10020
agcctttcca cggcacctc ttctccttcc tttctgctt cttccatgtc tgcttcggcc	10080
ctggggcgcc gcccgcggcc cctgcccccc atgcgcgtga ccccgaaaccc cctgagcggt	10140

[0096]

tggagcaggg ccaggtcggc gacgacgcgc tcggccagga tggccctgctg cacctgcgtg	10200
agggtggttt ggaagtcatc caagtccacg aagcggtggt aggcggccgt gttgatggtg	10260
taggtgcagt tggccatgac ggaccagttg acggctgtgt ggcccggttgcgacatctcg	10320
gtgtaccta gtcgcgagta ggcgcgggag tcgaagacgt agtcgttgca agtccgcacc	10380
aggtaactggt agcccaccag gaagtgcggc ggcggctggc ggtagagggg ccagcgcagg	10440
gtggcggggg ctccggggc caggtcttcc agcatgagggc ggtggtaggc gtagatgtac	10500
ctggacatcc aggtgatacc cgccgcggtg gtggaggcgc gcgggaagtc ggcacccgg	10560
ttccagatgt tgcgcagggg cagaaagtgc tccatggtag gcgtgctctg tccagtcaga	10620
cgcgcgcagt cggtgatact ctagaccagg gaaaacgaaa gccggcgtcgc gggactctt	10680
ccgtggctgt gtgaatagat cgcaaggta tcatggcga gggctcggt tcgagccccg	10740
ggtccgggcc ggacggtccg ccatgatcca cgcggttacc gcccgctgtt cgaaccagg	10800
tgtgcgacgt cagacaacgg tggagtgttc ctttggcgt tttctggcc gggcgccggc	10860
gccgcgttaag agactaagcc gcgaaagcga aagcagtaag tggctcgctc cccgtagccg	10920
gagggatect tgctaagggt tgcgttgcgg cgaacccgg ttcaatccc gtactcggtc	10980
cggccggacc cgccgctaag gtgttggatt ggcctcccc tcgtataaag accccgcttgc	11040
cggattgact cggacacagg ggacgagccc cttttatccc tgcttcccc agatgcattcc	11100
gggtgctgcgg cagatgcgcc ccccgccccca gcagcagcaa caacaccagg aagagcggca	11160
gcaacagcag cgggagtcatttgcagggcccc ctcacccacc ctcggcgggc cggccaccc	11220
ggcgtccgcg gccgtgtctg gcgcctgcgg cggcggcggg gggccggctg acgacccga	11280
ggagcccccg cggcgcaggcc ctagacacta cctggacactg gaggaggcgc agggcctggc	11340
ggggctgggg gcgcgtctc cggagcgcacccgcggcgggtgcagctgaagc gcgactcg	11400
cgaggcgtac gtgcctcgcc agaacctgtt cagggaccgc gcggcggagg agcccgagga	11460
gatgcgggac aggaggttca gcgcaggcgc ggagctgcgg cagggctgtga accgcgagc	11520
gctgctgcgc gaggaggact ttgagcccga cgcgcggacg gggatcagcc cgcgcgcgc	11580
gcacgtggcg gccgcgcacc tggtaacggc gtacgagcag acggtaacc aggagatcaa	11640
cttccaaaag agtttcaaca accacgtgcg cacgctggtg gcgcgcgagg aggtgaccat	11700
cgggctgatg cacctgtggg actttgttaag cgcgtgggtc cagaacccca acagcaagcc	11760
tctgacggcg cagctgttcc ttagatgtgc gacacaggcagg gacaacggcagg cggttaggaa	11820
cgcgcgtctg aacatcacccg agcccgaggc tcgggtggctg ctggacactga ttaacatcct	11880

[0097]

gcagagcata	gtggcgcagg	agcgcagcct	gagcctggcc	gacaagggtgg	cggccatcaa	11940
ctactcgatg	ctgagcctgg	gcaagttta	cgcgcgaag	atctaccaga	cggcgatcg	12000
gcccatagac	aaggagggtga	agatcgacgg	ttttacatg	cgcacggcgc	tgaagggtgt	12060
caccctgagc	gacgacctgg	gcgtgtaccg	caacgagcgc	atccacaagg	ccgtgagcgt	12120
gagccggcgg	cgcgagctga	gcgacccgca	gctgatgcac	agcctgcagc	gggcgcgtggc	12180
gggcgcggc	agcggcgaca	gggaggcggg	gtcctacttc	gatgcggggg	cggacctgcg	12240
ctggcgccc	agccggcggg	ccctggaggc	cgcgggggtc	cgcgaggact	atgacgagga	12300
cggcgaggag	gatgaggagt	acgagctaga	ggagggcgag	tacctggact	aaaccgcggg	12360
tgggtttcc	ggtagatgca	agacccgaac	gtggtgacc	cggcgctgcg	ggcgctctg	12420
cagagccagc	cgtccggcct	taactcctca	gacgactggc	gacaggtcat	ggaccgcata	12480
atgtcgctga	cggcgctaa	cccgacgcg	ttccggcagc	agccgcaggc	caacaggctc	12540
tccgccatcc	tggaggcggt	ggtgcctgcg	cgctgaacc	ccacgcacga	gaaggtgctg	12600
gccatagtga	acgcgctggc	cgagaacagg	gccatccgcc	cgacgcaggc	cgggctgggt	12660
tacgacgcgc	tgctgcagcg	cgtggcccgc	tacaacagcg	gcaacgtgca	gaccaacctg	12720
gaccggctgg	tggggacgt	gkgcgaggcg	gtggcgcagc	gcgagcgcgc	ggatcggcag	12780
ggcaacctgg	gctccatgg	ggcgctgaat	gccttcctga	gcacgcagcc	ggccaacgtg	12840
cccgccccgc	aggaagacta	caccaacttt	gtgagcgcgc	tgcggctgtat	ggtgaccgag	12900
accccccaga	gcgagggtgt	ccagtcggc	ccggactact	tcttccagac	cagcagacag	12960
ggcctgcaga	cggtaaacct	gagccaggct	ttcaagaacc	tgcggggct	gtggggcgtg	13020
aaggcgccca	ccggcgaccg	ggcgacggtg	tccagcctgc	tgacgccccaa	ctcgccctg	13080
ctgctgtgc	tgatcgcgc	gttcacggac	agcggcagcg	tgtcccggga	cacctacctg	13140
ggcacctgc	tgaccctgt	ccgcgaggcc	atcggcagg	cgcagggtgga	cgagcacacc	13200
ttccaggaga	tcaccagcg	gagccgcgcg	ctggggcagg	aggacacgag	cagcctggag	13260
gcgactctga	actacctgt	gaccaaccgg	cggcagaaga	ttccctcgct	gcacagcctg	13320
acctccgagg	aggagcgcac	cttgcgtac	gtgcagcaga	gcgtgagcct	gaacctgatg	13380
cgcgacgggg	tgacgcccag	cgtggcgctg	gacatgaccg	cgcgcaacat	gaaaccgggc	13440
atgtacgccc	cgcaccggcc	ttacatcaac	cgcctgtatgg	actacctgca	tgcgcggcg	13500
gccgtgaacc	ccgagtactt	taccaacgccc	atcctgaacc	cgcactggct	cccgccgccc	13560

[0098]

gggttctaca gcggggcgtt cgaggtcccg gagaccaacg atggcttcct gtggacac	13620
atggacgaca gcgtgttctc cccgcggccg caggcgctgg cgaaagcgctc cctgctgcgt	13680
cccaagaagg aggaggagga ggaggcgagt cgccgcccgcg gcagcagcgg cgtggcttct	13740
ctgtccgagc tggggccggc agccgcccgcg cgccccgggt ccctggcgg cagccccctt	13800
ccgagcctgg tgggtctct gcacagcgag cgacaccaccc gccctcggt gctggcgag	13860
gacgagtacc tgaataactc cctgctgcag ccggtgccgg agaaaaacct gcctccgccc	13920
ttccccaaaca acggataga gagcctgggt gacaagatga gcagatggaa gacatatgcg	13980
caggagcaca gggacgcgc tgcgctccgg ccgcacgc ggcgcagcg ccacgaccgg	14040
cagcgggggc tgggtggga tgacgaggac tccgcggacg atagcagcgt gctggacctg	14100
ggagggagcg gcaaccctt cgccacactg cgcccccgc tggggaggat gttttaaaaaa	14160
aaaaaaaaaaa aagcaagaag catgatgcaa aaattaaata aaactcacca aggccatggc	14220
gaccgagcgt tggtttttgc tggcccttc agtatgcggc gcgcggcgat gtaccaggag	14280
ggacctcctc cctcttacga gagcgtgggt ggcgcggccg cggcggccct ctcttctccc	14340
tttgcgtcgc agctgctgga gccgcgtac gtgcctccgc gctacctgcg gcctacgggg	14400
gggagaaaca gcatccgtta ctggagctg gcgcctgt tcgacaccac cgggtgtac	14460
ctggtgaca acaagtcggc ggacgtggcc tccctgaact accagaacga ccacagcaat	14520
ttttgacca cggcatcca gaacaatgac tacagccga gcgaggccag cacccagacc	14580
atcaatctgg atgaccggc gcactgggc ggcacactga aaaccatcct gcacaccaac	14640
atgccaacg tgaacgagtt catgttacc aataagtca aggcgcgggt gatgggtcgc	14700
cgctcgacca ccaaggaaga cgggtggag ctgaagtacg agtgggtgga gttcgagctg	14760
ccagagggca actactccga gaccatgacc attgacactga tgaacaacgc gatcgtggag	14820
cactatctga aagtggcag gcagaacggg gtcctggaga ggcacatcgg ggtcaagttc	14880
gacaccagga acttccgcct gggctggac cccgtgaccg ggctggttat gcccgggtg	14940
tacaccaacg aggccttcca tcccgacatc atcctgctgc cggctgcgg ggtggacttc	15000
acttacagcc gcctgagcaa ctcctggc atccgcaagc ggcagccctt ccaggaggc	15060
ttcaggatca ctcacgagga cttggagggg ggcaacatcc cgcgcctcct cgatgtggag	15120
gccttaccagg atagcttcaa ggaaatgag gcccggacagg aggataccgc ccccgccccc	15180
tccggcccg ccgagcaggc cgaggatgct gtcgacaccg cggccgcggc cggggcagag	15240
gccgaccccg ctatgggtt ggaggctccc gacaggaggagg aggacatgaa tgacagtgcg	15300

[0099]

gtgcgcccgg	acacacccgtt	cacccggggg	gaggaaaagc	aagcggaggc	cgaggccgcg	15360
gccgaggaaa	agcaactggc	ggcagcagcg	gcggccggcgg	cgttggccgc	ggcggaggct	15420
gagtctgagg	ggaccaagcc	cgc当地	cccgtgatta	agcccctgac	cgaagatagc	15480
aagaagcgca	gttacaacct	gctcaaggac	agcaccaaca	ccgc当地	cagctggta	15540
ctggcctaca	actacggcga	cccgtcgacg	ggggtgcgt	cctggaccct	gctgtgcacg	15600
ccggacgtga	cctgc当地	ggagcaggta	tactggtcgc	tgccc当地	gatgcaagac	15660
cccggtgacct	tccgctccac	gccc当地	agcaacttcc	cgg当地	cgccgagctg	15720
ctgcccgtgc	actccaagag	cttctacaac	gaccaggccg	tctactccca	gctcatccgc	15780
cagttcacct	ctctgaccca	cgtgttcaat	cgcttcc	agaaccagat	tctggcgcgc	15840
ccgccc当地	ccaccatcac	caccgtcagt	gaaaacgttc	ctgctctcac	agatcacggg	15900
acgctaccgc	tgc当地	catcgagga	gtccagc当地	tgaccgttac	tgacgccaga	15960
cgccgc当地	gccc当地	ttacaaggcc	ttgggcatag	tctgc当地	cgtc当地	16020
agccgcactt	tttgagcaac	accaccatca	tgtccatcct	gatctcaccc	agcaataact	16080
ccggctgggg	actgctgcgc	gccc当地	agatgttccg	aggggc当地	aagcgttccg	16140
agcagcaccc	cgtgc当地	cgc当地	tccgc当地	ctgggg当地	cacaaacgc当地	16200
gccgc当地	gccc当地	gtggacgacg	ccatgc当地	ggtgg当地	caggc当地	16260
actacaggcc	cgc当地	accgtggacg	cggccatcca	gaccgtgg	cgccgc当地	16320
ggc当地	caagctgaag	agccgc当地	agc当地	ccgccc当地	cgc当地	16380
ccggg当地	cgc当地	aaacgc当地	gccgc当地	ccctg当地	ccggg当地	16440
gccgc当地	catgaggcc	gccc当地	tggccgc当地	catcaccgc当地	gccaccatgg	16500
ccccccgtac	ccgaagacgc当地	gccc当地	ccgccc当地	cgccatc当地	gacatggcc当地	16560
gcaggc当地	ggcaacgtg	tactgggtgc当地	gcaactcggt当地	gaccggc当地	cgc当地	16620
tgc当地	ccccccgc当地	acttgagatg	atgtgaaaaa	acaactgta	gtctc当地	16680
tttgtgtta	tccc当地	ggc当地	gc当地	gtccaagc当地	aaaatcaaag	16740
aagagatgct	ccaggtcg	gccc当地	tctatggcc当地	ccc当地	gaagagcagg当地	16800
attcgaagcc	ccgcaagata	aagc当地	aaaagaaaaa	gaaagatgat	gacgatgc当地	16860
atggggaggt	ggagttc当地	cgc当地	cgc当地	cccg当地	tggaaggcc当地	16920
ggc当地	gtaaa	gc当地	cgccc当地	ccgc当地	ggc当地	16980

[0100]

ccacccggac tttcaagcgc gtctatgacg aggtgtacgg cgacgaagac ctgctggagc	17040
aggccaaacga ggcgcttcgga gagtttgctt acgggaagcg tcagcggcgc ctggggaaagg	17100
aggacctgtt ggcgctgccg ctggaccagg gcaacccac ccccagtctg aagccgtga	17160
ccctgcagca ggtgctgccg agcagcgcac cctccgaggc gaagcggggt ctgaagcgcg	17220
aggcgccgca cctggcgccc accgtgcagc tcatggtgc caagcggcag aggctggagg	17280
atgtgctgga gaaaatgaaa gtagaccccg gtctgcagcc ggacatcagg gtccgcggca	17340
tcaaggcagg ggcgccggc ctcggcgtgc agaccgtga cgtggtcattc cccaccggca	17400
actccccgcg cgcgcgcacc actaccgctg cctccacggc catggagaca cagaccgatc	17460
ccggccgcagc cgcaagccgca gccggccggc cgacccttc ggcggaggtg cagacggacc	17520
cctggctgcc gccggcgatg tcagctcccc ggcgcgtcg cggcgcagg aagtacggcg	17580
ccgccaacgc gctctcgccc gagtaacgcct tgcatttc catcgccccc accccggct	17640
accgaggcta tacctaccgc ccgcgaagag ccaagggttc caccggccgt cccggccgac	17700
gcggccgcgc caccacccgc cgccggccgc gcagacgcca gcccgcactg gctccagtct	17760
ccgtgaggaa agtggcgccgac gacggacaca ccctggtgct gcccaggcgc cgctaccacc	17820
ccagcatcg ttaaaagcct gttgtggttc ttgcagatat ggcctcaact tgccgcctcc	17880
gtttcccggt gccggatac cgaggaggaa gatcgccgcg caggagggtt ctggccggcc	17940
gcggcctgag cggaggcagc cgccgcgcgc accggcggcgc acgcgcacc accgcacgca	18000
tgcgcggcgg ggtgctgccct ctgttaatcc ccctgatcgc cgccgcgatc ggcgcgtgc	18060
ccgggatcgc ctccgtggcc ttgcaagcgt cccagaggca ttgacagact tgcaaacttg	18120
caaataatgga aaaaaaaaaacc ccaataaaaaa agtctagact ctcacgctcg cttggctctg	18180
tgactatttt gtagaatgga agacatcaac tttgcgtcgc tggcccgcg tcacggctcg	18240
cgcccggttcc tgggacactg gaacgatatac ggcaccagca acatgagcgg tggcgccttc	18300
agtggggct ctctgtggag cggcattaaa agtacgggt ctggcgtaa aaattacggc	18360
tcccgccct ggaacagcag cacggccag atgttgagag acaagttgaa agagcagaac	18420
ttccagcaga aggtggtgga gggctggcc tccggcatca acgggggtgg ggacctggcc	18480
aaccaggccg tgcagaataa gatcaacagc agactggacc cccggccgc ggtggaggag	18540
gtgccgcgg cgctggagac ggtgtcccc gatggcgtg gcgagaagcg cccgcggccc	18600
gatagggaaag agaccactt ggtcacgcag accgatgagc cgcggccgt a tgaggaggcc	18660
ctgaagcaag gtctgcccac cacgcggccc atcgcgcaca tggccaccgg ggtggtgccc	18720

[0101]

cgccacacccc ccgcccacgct ggacttgcct ccgccccgccc atgtgccgca gcagcagaag 18780  
gcggcacagc cggggcccgcc cgccgaccgccc tcccgttccc ccgcgggtcc tctgcgcgc 18840  
gcggccagcg gcccccgccgg gggggtcgccc aggcacggca actggcagag cacgctgaac 18900  
agcatcgtgg gtctgggggt gcgggtccgtg aagcgccgccc gatgctactg aatagcttag 18960  
ctaacgtgtt gtatgtgtgt atgcgcccta tgcgcgcgc agaggagctg ctgagtcgc 19020  
gcgcgttcgcg cgccccaccac cacccgcact ccgcgcctca agatggcgac cccatcgatg 19080  
atgcccgcagt ggtcgtaat gcacatctcg gccaggacgc cctcgagta cctgagcccc 19140  
gggctgggtgc agttcgcccg cgccaccgag agctacttca gcctgagtaa caagtttagg 19200  
aaccggccacgg tggcgcccac gcacgatgtg accaccgacc ggtctcagcg cctgacgc 19260  
cggttcattc ccgtggaccgcg cgaggacacc gcgtactcgt acaaggcgcg gttcacccctg 19320  
gccgtggcg acaaccgcgt gctggacatg gcctccaccc actttgacat ccgcgggtgc 19380  
ctggaccggg gtcccacttt caagccctac tctggcaccgc cctacaactc cctggccccc 19440  
aaggggcgtc ccaactcctg cgagtggag caagaggaaa ctcaggcgt tgaagaagca 19500  
gcagaagagg aagaagaaga tgctgacggt caagctgagg aagagcaagc agctaccaaa 19560  
aagactcatg tatatgctca ggctccctt tctggcgaaa aaattagtaa agatggtctg 19620  
caaataggaa cggacgctac agtacagaa caaaaaccta tttatgcaga ccctacattc 19680  
cagccccgaac cccaaatcg ggagtcccg tggaatgagg cagatgctac agtcgcggc 19740  
ggtagagtgc taaagaaatc tactccatg aaaccatgct atggcccta tgcaagaccc 19800  
acaaaatgcta atggaggtca ggggttacta acggcaaattt cccaggacgc gctagaatct 19860  
cagggtgaaa tgcaattttt ttcaacttctt gaaaacgcgc gtaacgaggc taacaacatt 19920  
cagccccaaat tggtgctgtt tagtgaggat gtgcacatgg agacccggta tacgcaccc 19980  
tcttacaaggcc cggccaaaaag cgatgacaat tcaaaaatca tgctgggtca gcagtc 20040  
cccaacagac ctaattacat cggcttcaga gacaacttta tcggcctcat gtattacaat 20100  
agcactggca acatgggagt gcttgcaggt caggccttc agttgaatgc agtgggtggac 20160  
ttgcaagacca gaaacacaga actgtcctac cagctttgc ttgattccat ggggtgacaga 20220  
accagatact tttccatgtt gaatcaggca gtggacagtt atgacccaga tgtagaaattt 20280  
attgaaaatc atggaactga agacgagctc cccaaactatt gttccctct ggggtggcata 20340  
ggggtaactg acacttacca ggctgttaaa accaacaatg gcaataacgg gggccagggtg 20400

[0102]

acttggacaa aagatgaaac tttgcagat cgcaatgaaa taggggtggg aaacaatttc	20460
gctatggaga tcaacacctag tgccaacctg tggagaaaact tcctgtactc caacgtggcg	20520
ctgtacccatc cagacaagct taagtacaac ccctccaatg tggacatctc tgacaacccc	20580
aacacccatcg attacatgaa caagcgagtg gtggcccccgg ggctggtgga ctgctacatc	20640
aacctggcg cgcgctggtc gctggactac atggacaacg tcaacccctt caaccaccac	20700
cgcaatgcgg gcctgcgcta ccgctccatg ctccctggca acgggcgcta cgtcccttc	20760
cacatccagg tgccccagaa gttctttgcc atcaagaacc tcctcctcct gccgggctcc	20820
tacacccatcg agtggaaactt caggaaggat gtcaacatgg tcctccagag ctctctgggt	20880
aacgatctca gggtggacgg gcccagcatc aagttcgaga gcatctgcct ctacgcccacc	20940
tttttccccca tggcccacaa cacggcctcc acgctcgagg ccatgctcag gaacgacacc	21000
aacgaccagt cttcaatga ctacctctcc gccgccaaca tgctctaccc cataccggcc	21060
aacgccacca acgtccccat ctccatcccc tcgcgcact gggggcctt ccggggctgg	21120
gccttcaccc gcctcaagac caaggagacc ccctccctgg gctcgggatt cgaccctac	21180
tacacccatcg cgggctccat tccctacctg gacggcacct tctacctcaa ccacacttc	21240
aagaaggctt cggtcacctt cgactcctcg gtcagctggc cggcaacga ccgtctgctc	21300
acccccaacg agttcgagat caagcgctcg gtcgacgggg agggctacaa cgtggcccg	21360
tgcaacatga ccaaggactg gttcttggtc cagatgtgg ccaactacaa catcggtac	21420
cagggcttct acatcccaga gagctacaag gacaggatgt actccttctt caggaacttc	21480
cagcccatga gccggcaggt ggtggaccag accaagtaca aggactacca ggaggtggc	21540
atcatccacc agcacaacaa ctcggccttc gtgggttacc tcggcccttccatgcgcgag	21600
ggacaggcct accccgccaa ctccccatcg cgcgtcatag gcaagaccgc ggtcgacagc	21660
atcacccaga aaaagttctt ctgcgaccgc accctctggc gcatccctt ctccagcaac	21720
ttcatgtcca tgggtgcgt ctggacactg ggccagaact tgctctacgc caactccgccc	21780
cacggccctcg acatgacctt cgaggtcgac cccatggacg agcccaccct tctctatgtt	21840
ctgttgcag tctttgacgt ggtccgggtc caccagccgc accgcggcgt catcgagacc	21900
gtgtacctgc gtacgcctt ctggccggc aacgccacca cctaaagaag caagccgcag	21960
tcatcgccgc ctgcgtgcg tcgggttcca ccgagcaaga gctcaggccatc atcgatcg	22020
acctgggatg cggccctat ttttggca cttcgacaa ggcgttccct ggctttgtct	22080
ccccacacaa gctggcctgc gccatgtca acacggccgg ccgcgagacc gggggcgtgc	22140

[0103]

actggctggc	cttcgcctgg	aacccgcgct	ccaaaacatg	cttcctctt	gacccttcg	22200
gctttcgga	ccagcggtc	aagcaaattct	acgagttcga	gtacgaggc	ttgctgcgtc	22260
gcagcgccat	cgcctcctcg	cccgaccgct	gcgtcaccct	cgaaaagtcc	acccagaccg	22320
tgcaggggcc	cgactcggtc	gcctgcggc	tcttcgtctg	catgtttctg	cacgcctttg	22380
tgcactggcc	tcagagtccc	atggaccgca	accccaccat	gaacttgctg	acgggggtgc	22440
ccaaactccat	gctccagagc	ccccaggtcg	agcccaccct	gcccgcgcaac	caggagcagc	22500
tctacagctt	cctggagcgc	cactgcctt	acttccgcgg	ccacagcgc	cagatcagga	22560
gggccacctc	cttctgccac	ttgcaagaga	tgcaagaagg	gtaataacga	tgtacacact	22620
tttttctca	ataaatggca	tcttttatt	tatacaagct	ctctgggtt	ttcatttccc	22680
accaccaccc	gccgttgcg	ccatctggct	ctatttagaa	atcgaaagg	ttctgcggg	22740
agtcggcgtg	cggcacggc	agggacacgt	tgcgatactg	gtagcgggtg	ccccacttga	22800
actcgggcac	caccaggcga	ggcagctcg	ggaagtttc	gctccacagg	ctgcgggtca	22860
gcaccagcgc	gttcatcagg	tcggcgccg	agatcttcaa	gtcgagttt	ggccgcgc	22920
cctgcgcgcg	cgagttgcgg	tacaccgggt	tgcagcactg	gaacaccaac	agcggcgggt	22980
gcttcacgct	ggccagcacg	ctgcggtcgg	agatcagctc	ggcgtccagg	tcctccgcgt	23040
tgctcagcgc	gaacgggttc	atcttggca	cttgcgcgc	caggaaggc	gcgtccccg	23100
gttgcagtt	gcagtgcag	cgcagcggg	tcaagcgttg	cccgtgcgg	gactcggcgt	23160
tgggtacag	cgcgcgcgt	aaggcctgca	tctggcgaa	ggccatctgg	gccttggcgc	23220
cctccgagaa	gaacatgccg	caggacttgc	ccgagaactg	gtttgcgggg	cagctggcgt	23280
cgtgcaggca	gcagcgcg	tcgggtttgg	cgtatgcac	cacgttgcgc	ccccaccgg	23340
tcttcacgat	cttggcctt	gacgattgt	cttcagcgc	gctgcgtcc	ttctcgctgg	23400
tcacatccat	ctcgatcaca	tgttcctt	tcaccatgt	gctgcgtgc	agacacttca	23460
gctgcgcctc	cgtctcggt	cagcggtgt	gccacagcgc	gcagccgtg	ggctcgaaag	23520
acttgttaggt	cacctccgcg	aaggactgca	ggtacccctg	caaaaagcgg	cccatcatgg	23580
tcacgaaggt	cttgcgttg	ctgaaggtca	gctgcagccc	gcggtgctcc	tcgttcagcc	23640
aggcttgca	cacggccgccc	agcgcctcca	cctggtcggg	cagcatctt	aagttcacct	23700
tcagctcatt	ctccacgtgg	tacttgtcca	tcagcgtgc	cgccgcctcc	atgccttct	23760
cccaggccga	caccaggcgc	aggctcacgg	ggttcttcac	catcaccgtg	gccggccct	23820

[0104]

ccggccgcgtt	ttcgctttcc	gccccgcgtt	tcttttcctc	ttccttcctt	tcctcgccgc	23880
cggccactcg	cagccccgc	accacgggg	cgtttctg	caggcgctgc	accttgcgt	23940
tgccgttgcg	cccctgctt	atgcgcacgg	gcgggttgct	gaagcccacc	atcaccagcg	24000
cggccttcc	ttgctcg	tcgctgtcca	aatgacacc	cggggagggg	gggttggtca	24060
tcctcagttac	cgaggcacgc	ttcttttct	tcctggggc	tttgcgc	tccgcggctg	24120
cggccgtgc	cgaggtcgaa	ggccgagg	tggcgtgcg	cggcacc	gcgtcctgcg	24180
agccgtcctc	gtcctcctcg	gactcgagac	ggaggcgg	ccgcttctt	ggggcgcgc	24240
ggggcggcgg	aggcggcggc	ggcgacgg	acggggacga	gacatgtcc	agggtgggt	24300
gacggcggc	cgcgcgcgt	ccgcgc	gggtggtctc	gcgtgg	tcttccgc	24360
tggccatctc	ccactgtcc	tttcctata	ggcagaaaga	gatcatggag	tctctcatgc	24420
gagtcgagaa	ggaggaggac	agcctaacc	ccccctctga	gccctccacc	accgcgc	24480
ccaccgc	tgccgcgc	gacgacgc	ccaccgagac	caccgc	accacc	24540
ccagcgacgc	accccgctc	gagaatgaag	tgctgatcga	gcaggaccc	gttttgta	24600
gcggagagga	ggatgagg	gatgagaagg	agaaggagga	ggtcgcgc	tcagtgc	24660
aagaggataa	aaagcaagac	caggacgac	cagataagga	tgagacagca	gtcggcgg	24720
ggaacggaag	ccatgatg	gatgacgg	acctagacgt	ggagacgac	gtgctg	24780
agcacctgca	ccgc	gtcatgtct	gcgcgcgt	gcaggagc	tgcgaagt	24840
ccctggacgt	ggcgagg	agccgc	acgagcgg	ccttcgc	ccgcacgt	24900
cccccaagcg	ccggagaac	ggcac	agcccaac	gcgtctca	tttaccc	24960
tcttcgcgt	acccgagg	ctggccac	accacat	tttccaa	tgcaagat	25020
cccttcctg	ccgc	cgaccc	ccgacaa	cctgacc	ccgcagg	25080
cccacatacc	tgatatgc	tcttgagg	aagtgc	gatttcg	ggtctcg	25140
gcgacgagaa	acggcgg	aacgtct	acggagac	cgaaaac	agtca	25200
gggtgctggt	ggagctcg	ggcgaca	cgccctgg	cgtact	cgcag	25260
aggcaccca	cttgcctac	ccggcgtca	acgtcccc	caagg	catg	25320
tggcgagct	catcatgc	cgccccc	ccctggcc	ggatgc	aaac	25380
cctccgagga	aggcctg	gcggc	acgagc	ggcgc	ctgg	25440
gcgaccc	gcagctgg	gagcgg	agctcat	ggccgc	ctgg	25500
tggagctcg	gtgtctcg	cgcttctc	cgga	gatgc	agctcg	25560

[0105]

agaccctgca ctacacccttc cgccaggcgt acgtgcgcca ggcctgcaag atctccaacg	25620
tggagctctg caacctggtc tcctacctgg gcacccctgca cgagaaccgc ctcggcaga	25680
acgtcctgca ctccaccctc aaaggggagg cgccgcgcga ctacatccgc gactgcgcct	25740
acctcttcct ctgctacacc tggcagacgg ccatgggggt ctggcagcag tgcctggagg	25800
agcgcaacct caaggagctg gaaaagctcc tcaagcgcac cctcagggac ctctggacgg	25860
gcttcaacga gcgctcggtg gccgcccgcgc tggcggacat catcttcggc gagcgcctgc	25920
tcaagaccct gcagcagggc ctgcccact tcaccagcca gagcatgctg cagaacttca	25980
ggactttcat cctggagcgc tcggcatcc tgccggccac ttgctgcgcg ctgcccagcg	26040
acttcgtgcc catcaagtac agggagtgcc cgccggcgt ctggggccac tgctacctct	26100
tccagctggc caactacctc gcctaccact cggacccat ggaagacgtg agcggcgagg	26160
gcctgctcga gtgccactgc cgctgcaacc tctgcacgcc ccaccgctct ctagtctgca	26220
accccgagct gctcagcgag agtcagatta tcggtacctt cgagctgcag ggtccctcgc	26280
ctgacgagaa gtccggcgt ccagggctga aactcaactcc ggggctgtgg acttccgcct	26340
acctacgcaa atttgtaccc gaggactacc acgcccacga gatcagggttc tacgaagacc	26400
aatccggccc gcccaaggcg gagctcaccc cctgcgtcat cacccagggg cacatcctgg	26460
gccaattgca agccatcaac aaagccgcgc gagagttctt gctaaaaaag ggtcgaaaaa	26520
tgtacctgga cccccagtc ggcgaggagc taaaccgcgt acccccgcgc ccccccagc	26580
agcgggaccc tgcctccag gatggcaccc agaaagaagc agcagccgc gccggccgc	26640
cagccataca tgcttctgga ggaagaggag gaggactggg acagtcaggc agaggaggtt	26700
tcggacgagg agcaggagga gatgatggaa gactggagg aggacagcag cctagacgag	26760
gaagcttcag aggccgaaga ggtggcagac gcaacaccat cgccctcggt cgccggcc	26820
tcggccgggc ccctgaaatc ctccgaaccc agcaccagcg ctataacctc cgctcctccg	26880
gcgcggcgc caccggcccg cagaccaac cgtagatggg acaccacagg aaccggggtc	26940
ggttaagtcca agtgcgcgc gccgcccaccc cagcagcagc agcagcagcg ccaggctac	27000
cgctcgtggc gcgggcacaa gaacgccata gtcgcctgct tgcaagactg cggggcaac	27060
atctcttcg cccggcgtt cctgctattc caccacgggg tcgccttcc ccgcaatgtc	27120
ctgcattact accgtcatct ctacagcccc tactgcagcg gcgacccaga ggcggcagcg	27180
gcagccacag cggcgaccac cacctaggaa gatattcctcc gcgggcaaga cagcggcagc	27240

[0106]

agcggccagg agacccgcgg cagcagcggc gggagcggtg ggcgcactgc gcctctcgcc	27300
caacgaaccc ctctcgaccc gggagcttag acacaggatc ttccccactt tgtatgccat	27360
cttccaacag agcagaggcc aggaggcagga gctgaaaata aaaaacagat ctctgcgctc	27420
cctcacccgc agctgtctgt atcacaaaag cgaagatcag cttcggcgca cgctggagga	27480
cgccgaggca ctcttcagca aatactgcgc gctcacttt aaagacttagc tccgcgcct	27540
tctcgaattt aggccggaga aaactacgtc atcgccggcc gccgcccagc ccgcccagcc	27600
gagatgagca aagagattcc cacgccatac atgtggagct accagccgca gatgggactc	27660
gcggcgggag cggcccgagga ctactccacc cgcatgaact acatgagcgc gggaccccac	27720
atgatctcac aggtcaacgg gatccgcgc cagcgaaacc aaatactgct ggaacaggcg	27780
gccatcacccg ccacgccccg ccataatctc aaccccgaa attggccgc cgccctcg	27840
taccaggaaa cccctccgc caccaccgtt ctactccgc gtgacgcccga ggccgaagtc	27900
cagatgacta actcaggggc gcagctcgcg ggcggcttc gtcacggggc gcggccgctc	27960
cgaccaggtt taagacacact gatgatcaga ggccgaggtt tccagctaa cgacgagtcg	28020
gtgagcttt cgctcggtt ccgtccggac ggaactttcc agctcgccgg atccggccgc	28080
tcttcgttca cgcccccgcga gggttacctt actctgcaga cctcgtcctc ggagcccccgc	28140
tccggcggca tcggaaccctt ccagttcggtt gaggagttcg tgccctcggtt ctacttcaac	28200
cccttctcggtt gaccccccgg acgctacccc gaccaggta ttccgaactt tgacgagggtt	28260
aaggactcgg cggacggcta cgactgaatg tcaggtgtcg aggagagca gttcgccctg	28320
agacacctcg agcaactgcgc ccggcacaag tgcttcgcgc gcggttctgg tgagttctgc	28380
tactttcagc taccggagga gcataccggag gggccggcgc acggcgccg cctgaccacc	28440
cagggcgagg ttacctgttc cctcataccgg gagtttaccc tccgtccctt gctagtggag	28500
cgggagcggg gtccctgtgt cctaactatc gcctgcaact gccctaacc tggattacat	28560
caagatcttt gctgtcatct ctgtgttagt tttataaaac gctgagatca gaatctactg	28620
gggctcctgtt cgccatcctg tgaacgccac cgtcttcacc caccggacc agggccaggc	28680
gaacctcacc tgcggcttc atcgaggggc caagaagtac ctcacctggt acttcaacgg	28740
cacccctttt gtggtttaca acagttcga cggggacgga gtctccctga aagaccagct	28800
ctccggcttc agctacttca tccacaagaa caccaccctc caactttcc ctccctaccc	28860
gccgggaacc tacgagtgcg tcaccggccg ctgcacccac ctcacccgcc tgatcgtaaa	28920
ccagagcttt ccgggaacag ataactccctt cttcccaaga acaggaggtg agctcaggaa	28980

[0107]

actccccggg gaccaggcg gagacgtacc ttgcaccctt gtggggtag gatTTTTat	29040
taccgggttg ctggctttt taatcaaagt ttccctgaga ttgttctt cttctacgt	29100
gtatgaacac ctcacacctcc aataactcta cccttcttc ggaatcagg gacttctcg	29160
aaatcgggct tggtgtgctg cttaactctgt tgatTTTT ctttatcata ctcagccttc	29220
tgtgcctcag gctgcccgc tgctgcgcac acatctatat ctactgctgg ttgctcaagt	29280
gcaggggtcg ccacccaaga tgaacaggta catggccta tcgatcctag gcctgctggc	29340
cctggcggcc tgcagcgccg ccaaaaaaaga gattacctt gaggagcccg cttgaatgt	29400
aacttcaag cccgagggtg accaatgcac caccctcgac aaatgcgtt ccaatcatga	29460
gaggctgcgc atcgactaca aaaacaaaac tggccagttt gcggctata gtgttttac	29520
gcccgagac ccctctaact actctgtcac cgtttccag ggcggacagt ctaagatatt	29580
caattacact ttccctttt atgagttatg cgatgcgtc atgtacatgt caaaacagta	29640
caacctgtgg cctccctctc cccaggcgtg tgtggaaat actgggtctt actgctgtat	29700
ggcttcgca atcactacgc tcgctctaattt ctgcacggc ctatacataa aattcaggca	29760
gaggcgaatc ttatcgatg aaaagaaaaat gccttgatcg ctaacaccgg ctttctatct	29820
gcagaatgaa tgcaatcacc tccctactaa tcaccaccac ctccttgcg attggccatg	29880
ggttgcacacg aatcgaagtg ccagtgggtt ccaatgtcac catggtggc cccggcggca	29940
attccaccct catgtggaa aaatttgc gcaatcaatg gtttcatttc tgctctaacc	30000
gaatcagtat caagcccaga gccatctgcg atggcaaaa tctaactctg atcaatgtgc	30060
aaatgatgga tgctgggtac tattacggc agcggggaga aatcattaaat tactggcgcac	30120
cccacaagga ctacatgctg catgttagtcg aggcaattcc cactaccacc cccactacca	30180
cctctccac caccaccacc actactacta ctactactac tactactact actaccacta	30240
ccgctgcccgc ccataccgc aaaagcacca tgattagcac aaagccccct cgtgctact	30300
cccacgcggc cggcccatc ggtgcgaccc cagaaaccac cgagcttgc ttctgccaat	30360
gcactaacgc cagcgctcat gaactgttcg acctggagaa tgaggatgtc cagcagagct	30420
ccgcttgcct gacccaggag gctgtggagc ccgttgcct gaagcagatc ggtgattcaa	30480
taattgactc ttcttctttt gccactcccg aataccctcc cgattctact ttccacatca	30540
cgggtaccaa agaccctaac ctcttttctt acctgatgtc gctgctgt atctctgtgg	30600
tctctccgc gctgatgtta ctggggatgt tctgctgcct gatctgccgc agaaagagaa	30660

[0108]

aagctcgctc tcagggccaa ccactgatgc cttccctta ccccccggat tttgcagata	30720
acaagatatg agctcgctgc tgacactaac cgctttacta gcctgcgctc taacccttgt	30780
cgcttgcgac tcgagattcc acaatgtcac agctgtggca ggagaaaaatg ttactttcaa	30840
ctccacggcc gatacccaagt ggctgtggag tggctcaggt agctacttaa ctatctgcaa	30900
tagctccact tccccggca tatccccaac caagtaccaa tgcaatgcca gcctgttcac	30960
cctcatcaac gttccaccc tggacaatgg actctatgta ggctatgtac ctttggtgg	31020
gcaaggaaag acccacgctt acaacctgga agttcggcag cccagaacca ctacccaagc	31080
ttctcccacc accaccacca ccaccacat caccagcagc agcagcagca gcagccacag	31140
cagcagcagc agattattga ctttggtttt ggccagctca tctggcgcta cccaggccat	31200
ctacagctct gtggccgaaa ccactcagat ccaccggcca gaaacgacca cggccaccac	31260
cctacacacc tccagcgtac agatgccgac caacatcacc cccttggctc ttcaaattgg	31320
acttacaagg cccactccaa aaccagtggta tgccggcgag gtctccgccc tcgtcaatga	31380
ctggcgggg ctggaatgt ggtgggtcgc cataggcatg atggcgctct gcctgcttct	31440
gctctggctc atctgctgcc tccaccgcag gcgagccaga ccccccattct atagaccat	31500
cattgtcctg aaccccgata atgatggat ccatagattt gatggcctga aaaacctact	31560
tttttctttt acagttatgtt aaattgagac atgcctcgca ttttcttgcata gatgttcctt	31620
ctcccaccc ttctgggtg ttctacgctg gccgctgtgt ctcacctgga ggttagactgc	31680
ctctcacccct tcactgtcta cctgcttac ggattggtca ccctcactct catctgcagc	31740
ctaattcacag taatcatcgc cttcatccag tgcattgatt acatctgtgt gcgcctcgca	31800
tacttcagac accacccgca gtaccgagac aggaacattt cccaaacttct aagactgctc	31860
taatcatgca taagactgtg atctgccttc tgatcctctg catcctgccc accctcacct	31920
cctgccagta caccacaaaaa tctccgcgca aaagacatgc ctccctggc ttcacccaaac	31980
tgtggatat acccaaattgc tacaacgaaa agagcgagct ctccgaagct tggctgtatg	32040
gggtcatctg tgtcttagtt ttctgcagca ctgtcttgc cctcataatc taccctact	32100
ttgatttggg atggaacgca atcgatgcca tgaattaccc caccttccc gcacccgaga	32160
taattccact gcgacaagtt gtacccgttg tcgttaatca acgccccca tcccctacgc	32220
ccactgaaat cagctacttt aacctaacag gcggagatga ctgacgcctt agatctagaa	32280
atggacggca tcagtaccga gcagcgtctc ctagagaggc gcaggcaggc ggctgagcaa	32340
gagcgcctca atcaggagct ccgagatctc gttaacctgc accagtgcaa aagaggcatt	32400

[0109]

tttgtctgg taaagcaggc caaagtcacc tacgagaaga ccggcaacag ccaccgcctc	32460
agttacaat tgcccaccca gcgccagaag ctgggtctca tgggtgggtga gaatcccatc	32520
accgtcacc accgtcacc agcactcggt agagaccgag ggggtctgc actccccctg tcgggtccca	32580
gaagacctct gcaccctggt aaagaccctg tgccgtctca gagatttagt ccccttaac	32640
taatcaaaca ctggaatcaa taaaaagaat cacttactta aaatcagaca gcaggtctct	32700
gtccagttt ttcagcagca ctccttccc ctccctccaa ctctggtaact ccaaacgcct	32760
tctggcggca aacttcctcc acaccctgaa gggatgtca gattcttgcct cctgtccctc	32820
cgacccact atttcatgt tggcagat gaagcgcacc aaaacgtctg acgagagctt	32880
caacccctg taccctatg acacggaaag cggccctccc tccgtccctt tcctcacc	32940
tcccttcgtg tctcccgatg gattccaaga aagtccccc ggggtctgt ctctgaacct	33000
ggccgagccc ctggtaactt cccacggcat gctccctg aaaatggaa gtggctctc	33060
cctggacgac gctggcaacc tcacctctca agatatcacc accgctagcc ctccctcaa	33120
aaaaaccaag accaacctca gcctagaaac ctcatccccc ctaactgtga gcacccctagg	33180
cgcctcacc gtagcagccg ccgctccctt ggcggggcc ggcacccccc tcaccatgca	33240
atcagaggcc cccctgacag tacaggatgc aaaactcacc ctggccacca aaggccccct	33300
gaccgtgtct gaaggcaaac tggcatttgc aacatggcc ccgctgacgg ccgctgacag	33360
cagcaccctc acagtcagt ccacaccacc ctttagcaca agcaatggca gttgggtat	33420
tgacatgcaa gccccattt acaccacaa tggaaaacta ggacttaact ttggcgctcc	33480
cctgcattgt gtagacagcc taaatgcact gactgttagt actggccaag gtcttacgat	33540
aaacggaaaca gcccataaaa ctagatctc aggtccctc aactatgaca catcaggaaa	33600
cctagaattt agagctgcag ggggtatgcg agttgatgca aatggtaac ttatcatttga	33660
tgttagttac ccatttgcac cacaacaa tctcagcattt aggcttggac agggaccctt	33720
gtttgttaac tctgcacaca acttggatgt taactacaac agaggcctt acctgttac	33780
atctggaaat accaaaaaggc tagaagttaa tatcaaaca gccaagggtc tcattttat	33840
tgacactgct atagcaatca atgcgggtga tggctacag tttgactcag gctcagatac	33900
aaatccattt aaaaactaaac ttggattagg actggattat gactccagca gagccataat	33960
tgctaaactg ggaactggcc taagcttgc caacacaggt gccatcacag taggaacaa	34020
aaatgtatgac aagcttacact tggaccac accagaccca tcccctaact gttagaatcta	34080

[0110]

ttcagagaaa gatgctaaat tcacacttgt tttgactaaa tgcggcagtc aggtgtggc	34140
cagcgtttct gtttatctg taaaaggtag cttgcgccc atcagtggca cagtaactag	34200
tgctcagatt gtcctcagat ttgatgaaaa tggagttcta ctaagcaatt cttcccttga	34260
ccctcaatac tggaactaca gaaaaggta cttacagag ggcactgcat ataccaacgc	34320
agtgggattt atgccaacc tcacagcata cccaaaaaca cagagccaa ctgctaaaag	34380
caacattgta agtcaggaaa acttgaatgg ggacaaatcc aaacccatga ccctcaccat	34440
taccctcaat ggaactaatg aaacaggaga tgccacagta agcacttact ccatgtcatt	34500
ctcatggaac tggaaatggaa gtaattacat taatgaaacg ttccaaacca actccttcac	34560
cttctcctac atcgcccaag aataaaaagc atgacgctgt tgatttgatt caatgtgttt	34620
ctgttttatt ttcaagcaca acaaaatcat tcaagtcatt cttccatctt agcttaatag	34680
acacagtagc ttaatagacc cagtagtgca aagccccatt ctagttata gatcagacag	34740
tgataattaa ccaccaccac caccatacct tttgatttag gaaatcatga tcacacagg	34800
atcctagtc tcaggccgcc ccctccctcc caagacacag aatcacagt cctctcccc	34860
cgactggctt taaataacac catctggtt gtcacagaca tggcttttagg gtttatattc	34920
cacacggctt cctgcccgc caggcgctcg tcgggtatgt tgataaactc tccggcagc	34980
tcgctcaagt tcacgtcgct gtccagcggc tgaacctccg gctgacgcga taactgtcg	35040
accggctgct ggacgaacgg aggcgcgc tacaaggggg tagagtcata atcctcggtc	35100
aggatagggc ggtgatgcag cagcagcgg cgaaacatct gctgccgcg ccgctccgtc	35160
cgccaggaaaa acaacacgcc ggtggctcc tccgcataa tccgcaccgc ccgcagcatc	35220
agcttcctcg ttctccgcgc gcagcacctc acccttatct cgctcaaattc ggcgcagtag	35280
gtacagcaca gcaccacgat gttattcatg atcccacagt gcagggcgct gtatccaaag	35340
ctcatgccgg gaaccaccgc cccacgtgg ccacgttacc acaagcgcac gtaaatcaag	35400
tgtcgacccc tcatgaacgc gctggacaca aacattactt cttggcat gttgtattc	35460
accacctccc ggtaccagat aaacctctgg ttgaacaggg caccttccac caccatctg	35520
aaccaagagg ccagaacctg cccacggct atgcactgca gggaaacccgg gttggacaa	35580
tgacaatgca gactccaagg ctcgttaaccg tggatcatcc ggctgctgaa ggcacatcgat	35640
ttggcacaac acagacacac gtgcacatgcac tttctcatga ttagcagctc ttccctcg	35700
aggatcatat cccaaaggaaat aacccattct tgaatcaacg taaaacccac acagcaggaa	35760
aggcctcgca cataactcac gttgtgcatt gtcagcgtgt tgcattccgg aaacagcgg	35820

[0111]

tgatcctcca	gtatcgaggc	gcgggtctcc	ttctcacagg	gaggtaaagg	gtccctgctg	35880
tacggactgc	gccgggacga	ccgagatcgt	gttgagcgta	gtgtcatgga	aaaggaaacg	35940
ccggacgtgg	tcataacttct	tgaagcagaaa	ccaggttcgc	gcgtggcagg	cctccttgcg	36000
tctgcggtct	cggcgtctag	ctcgctccgt	gtgatagttg	tagtacagcc	actcccgcag	36060
agcgtcgagg	cgcaccctgg	cttccggatc	tatgttagact	ccgtcttgca	ccgccccct	36120
gataatatcc	accaccgtag	aataagcaac	acccagccaa	gcaatacact	cgctctgcga	36180
gcggcagaca	ggaggagcgg	gcagagatgg	gagaaccatg	ataaaaaaact	ttttttaaag	36240
aatattttcc	aattcttcga	aagtaagatc	tatcaagtgg	cagcgctccc	ctccactggc	36300
gcggtcaaac	tctacggcca	aagcacagac	aacggcattt	ctaagatgtt	ccttaatggc	36360
gtccaaaaga	cacaccgctc	tcaagttgca	gtaaactatg	aatgaaaacc	catccggctg	36420
atttccaat	atagacgcgc	cggcagcgtc	caccaaaccc	agataatttt	cttctctcca	36480
gcggtttacg	atctgtctaa	gcaaatccct	tatataagt	ccgaccatgc	caaaaatctg	36540
ctcaagagcg	ccctccac	tcatgtacaa	gcagcgcac	atgattgcaa	aaattcaggt	36600
tcttcagaga	cctgtataag	attcaaaatg	ggaacattaa	caaaaattcc	tctgtcgcc	36660
agatcccttc	gcagggcaag	ctgaacataa	tcagacaggt	cgaaacggac	cagtgaggcc	36720
aaatccccac	caggaaccag	atccagagac	cctatactga	ttatgacgcg	catactcggg	36780
gctatgtga	ccagcgtagc	gccgatgtag	gcgtgtcga	tggcggcga	gataaaatgc	36840
aaagtgtgg	ttaaaaaatc	aggcaaagcc	tcgcgaaaaa	aagctaacac	atcataatca	36900
tgctcatgca	ggtagttgca	ggttaagctca	ggaacaaaaa	cggaataaca	cacgattttc	36960
ctctcaaaca	tgacttcgca	gatactgcgt	aaaacaaaaa	attataaata	aaaaattaat	37020
taaataactt	aaacattgga	agcctgtctc	acaacaggaa	aaaccacttt	aatcaacata	37080
agacgggccca	cgggcatgcc	ggcatagccg	taaaaaaatt	ggtccccgtg	attaaaaagt	37140
accacagaca	gctcccggt	catgtcgaaa	gtcatcatgt	gagactctgt	atacacgtct	37200
ggattgtgaa	catcagacaa	acaaagaaat	cgagccacgt	agccggagg	tataatcacc	37260
cgcaggcgga	ggtacagcaa	aacgacccccc	ataggaggaa	tcacaaaatt	agtaggagaa	37320
aaaaatacat	aaacaccaga	aaaaccctgt	tgctgaggca	aaatagcgcc	ctcccgatcc	37380
aaaacaacat	aaagcgcttc	cacaggagca	gccataacaa	agacccgagt	cttaccagta	37440
aaagaaaaaa	gatctctcaa	cgcagcacca	gcaccaacac	ttcgcagtgt	aaaaggccaa	37500

[0112]

gtgccgagag agtatataaaa ggaataaaaaa gtgacgtaaa cgggcaaagt ccaaaaaacg	37560
cccagaaaaa ccgcacgcga acctacgccc cgaaacgaaa gccaaaaaac actagacact	37620
cccttccggc gtcaacttcc gcttccac gctacgtcac ttccccggt caaacaact	37680
acatatccg aacttccaag tcgccacgccc caaaacaccg cctacaccctc cccgcccggc	37740
ggcccgcccc cggaccggcc tccgccccgg cggccggccat ctcattatca tattggcttc	37800
aatccaaaat aaggtatatt attgatgatg	37830

&lt;210&gt; 11

&lt;211&gt; 37559

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 11

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atggcgccg	60
cggggcgccccggg cgccggggcg gaggcggtt tggggcgccc cggccggcg gggcggtgt	120
gcggaaagtgg actttgttaag tgtggcgat gtgacttgct agtgccggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgccccgg ggaagtgaca tttttccgc gtttttacc	240
ggatgttgta gtgaatttgg gcgttaaccaa gtaagatttgc cccatttcg cgggaaaact	300
gaaacgggaa agtgaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
ggcccgaggg actttggccg attacgtgga ggactcgccc aggtgtttt tgaggtgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacg	480
gtatccatat cataatatgt acatttatgt tggctcatgt ccaacattac cgcgttgc	540
acattgatta ttgactagtt attaatagta atcaattacg gggcatttag ttcatacgccc	600
atatatggag ttccgcgtta cataacttac ggtaaatggc cccctggct gaccggccaa	660
cgcacccgcg ccattgacgt caataatgac gtatgtccc atagtaacgc caataggac	720
tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtagtca	780
agtgtatcat atgccaagta cgccttat tgacgtcaat gacggtaaat ggccgcctg	840
gcattatgcc cagtagtca ctttatggc ctttcctact tggcagtaca tctacgtatt	900
agtcatcgct attaccatgg tggatgcgtt ttggcagtac atcaatggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatgggatgttttg	1020

[0113]

gcaccaaaaat caacggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat	1080
ggcggttagg cgtgtacggt gggaggtcta tataaggcaga gctctcccta tcagtgatag	1140
agatctccct atcagtgata gagatcgtcg acgagctcg ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gtttgacct ccatagaaga caccgggacc gatccagcct	1260
cccgccggcgg gaacggtgca ttggaacgcg gattccccgt gccaagagtg agatcttccg	1320
tttatctagg taccagatat cgccaccatg gaactgctga tcctgaaggc caacgccatc	1380
accaccatcc tgaccggcgt gacccctgc ttcgccagcg gccagaacat caccgaggaa	1440
ttctaccaga gcacctgttag cgccgtgagc aagggttacc tgagcgcct gagaaccggc	1500
tggcacca gcgtgatcac catcgagctg agcaacatca aagaaaacaa gtgcaacggc	1560
accgacgcca aagtgaagct gatcaaggcag gaactggaca agtacaagaa cgccgtgacc	1620
gagctgcagc tgctgatgca gagcaccccc gccaccaaca accggggccag acgggagctg	1680
ccccggttca tgaactacac cctgaacaac gccaaaaaaga ccaacgtgac cctgagcaag	1740
aagcggaaagc ggcggggttctt gggcttctg ctggcgttgg gcagcgccat tgccagcggc	1800
gtggccgtgt ctaaggtgct gcacctggaa ggcgaagtga acaagatcaa gagcgcctg	1860
ctgagcacca acaaggccgt ggtgtccctg agcaacggcg tgagcgtgct gaccagcaag	1920
gtgctggatc tgaagaacta catcgacaag cagctgctgc ccatcgtaa caagcagagc	1980
tgcagcatca gcaacatcga gacagtgate gagttccagc agaagaacaa ccggctgctg	2040
gaaatcaccc gggagtttag cgtgaacgcc ggcgtgacca cccctgtgtc cacctacatg	2100
ctgaccaaca gcgagctgct gagcctgatc aacgacatgc ccatcaccaa cgaccagaaa	2160
aagctgatga gcaacaacgt gcagatcgta cggcagcaga gctactccat catgtccatc	2220
atcaaagaag aggtgctggc ctacgtggc cagctgcccc tgtacggcgt gatcgacacc	2280
ccctgctgga agctgcacac cagccccctg tgcaccacca acaccaaaga gggcagcaac	2340
atctgcctga cccggaccga cagaggctgg tactgacgaca acgcccggcag cgtgtcattc	2400
tttccacagg ccgagacatg caaggtgcag agcaaccggg tggctgcga caccatgaac	2460
agcctgaccc tgccctccga agtgaacctg tgcaacgtgg acatcttcaa ccccaagtac	2520
gactgcaaga tcatgaccc tcagaccgac gtgtccagct ccgtgatcac ctccctggc	2580
gccatcggtgt cctgctacgg caagaccaag tgcaccgcc gcaacaagaa ccggggcattc	2640
atcaagacct tcagcaacgg ctgcgactac gtgtccaaca aggggggtgga caccgtgtcc	2700
gtggcaaca ccctgtacta cgtgaacaaa caggaaggca agagcctgta cgtgaaggc	2760

[0114]

gagcccatca tcaacttcta cgacccctg gtttccca ggcacgagg cgacccagc	2820
atcagccagg tgaacgagaa gatcaaccag agcctggct tcatccggaa gtccgacgag	2880
ctgctgcaca atgtaatgc cggcaagtcc accaccaacc ggaagcggag agccctgtg	2940
aagcagaccc tgaacttcga cctgctgaag ctggccggcg acgtggagag caatccggc	3000
cctatggccc tgagcaaagt gaaactgaac gatacactga acaaggacca gctgctgtcc	3060
agcagcaagt acaccatcca gcggagcacc ggcgacagca tcgataccca caactacgac	3120
gtgcagaagc acatcaacaa gctgtcgcc atgctgctga tcacagagga cgccaaaccac	3180
aagttcaccg gcctgatcgg catgctgtac gccatgagcc ggctggccg ggaggacacc	3240
atcaagatcc tgcggacgc 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 tcgcccctt ggtgatcaca aagctggccg ctggcgacag atctggcctg	3540
acagccgtga tcagacggc caacaatgtg ctgaagaacg agatgaagcg gtacaagggc	3600
ctgctgcca aggacattgc caacagcttc tacgaggtgt tcgagaagta cccccacttc	3660
atcgacgtgt tcgtgcactt cggcattgcc cagacgacca ccagaggcgg ctccagagt	3720
gagggcatct tcgcccctt gttcatgaac gcctacggcg ctggccaggt gatgctgaga	3780
tggggcgtgc tggccaagag cgtgaagaac atcatgctgg gccacgcccag cgtcaggcc	3840
gagatggAAC aggtggtgga ggtgtacgag tacgcccaga agctggccgg agaggccggc	3900
ttctaccaca tcctgaacaa ccctaaggcc tccctgctgt ccctgaccca gttccccac	3960
ttctccagcg tggcgtggg aaatgccgccc ggactggca tcatggcga gtaccgggac	4020
accccccagaa accaggacccgtacgacgccc gccaaggcct acgcccggac gctgaaagaa	4080
aacggcgtga tcaactacag cgtgctggac ctgaccgctg aggaactggaa agccatcaag	4140
caccagctga accccaagga caacgacgtg gagctggag gcccggac gacggcaag	4200
ggcatgagca gacggAACcc ctgcaagtcc gagatccggg gccactgcct gaacggcaag	4260
cgggccact tcagccacaa ctacttcgag tggcccccctc atgctctgt ggtggccag	4320
aacttcatgc tgaaccggat cctgaagtcc atggacaaga gcatcgacac cctgagcggag	4380
atcagcggag ccggcggagct ggacagaacc gaggaatatg ccctggcgt ggtggagtg	4440

[0115]

ctgaaagct acatcggtc catcaacaac atcacaaagc agagcgctg cgtggccatg	4500
agcaagctgc tgacagagct gaacagcgac gacatcaaga agctgaggga caacgaggaa	4560
ctgaacagcc ccaagatccg ggtgtacaac accgtgatca gctacattga gagcaaccgc	4620
aagaacaaca agcagaccat ccatctgctg aagcggctgc ccggcgtacgt gctgaaaaag	4680
accatcaaga acaccctgga catccacaag tccatcacca tcaacaatcc caaagaaagc	4740
accgtgtctg acaccaacga tcacgccaag aacaacgaca ccacctgatg agcggccg	4800
atctgctgtc cttcttagtt gccagccatc tttttttgc ccctccccg tgccttcctt	4860
gaccctggaa ggtgccactc ccactgtcct ttccataataa aatgaggaaa ttgcacatcg	4920
ttgtctgagt aggtgtcatt ctattctggg ggggtgggtg gggcaggaca gcaaggggaa	4980
ggattggaa gacaatagca ggcattgtgg ggtgggtg ggctctatgg ccgtatcg	5040
atcgctgagg tgggtgagtg ggcgtggcct ggggtggta tgaaaatata taagttggg	5100
gtcttaggt ctcttattt gtgtgcaga gaccgcggaa gccatgagcg ggagcagcag	5160
cagcagcagt agcagcagcg cttggatgg cagcatcg agcccttatt tgacgacg	5220
gatccccac tggccgggg tgcgtcagaa tgtatgggc tccagcatcg acggccgacc	5280
cgtcctgccc gcaaattccg ccacgctgac ctatgcgacc gtcgcgggaa cgccgttgg	5340
cgccaccgccc gccgcgcgcg ccacgcgcgc cgcctggcc gtgcgcagcc tggccacgg	5400
ctttgcattc ctgggaccac tggcgacagg ggctacttct cggccgcgt ctggcccg	5460
tcgcgtac aagctgaccg ccctgctggc gcagttggat gcgttactc gggactgg	5520
tgacctttct cagcaggta tggccctgctc ccagcaggta tcctccctgc aagctggcgg	5580
gaatgcttct cccacaaatg ccgttaaga taaataaaac cagactctgt ttggattaa	5640
gaaaagtagc aagtgcattt ctcttttat ttcataattt tccgcgcgc ataggcccta	5700
gaccagcggtt ctcggcggtt gagggtgcgg tgtatttct ccaggacgtg gttaggg	5760
ctctggacgt tgagatacat gggcatgagc ccgtccggg ggtggaggta gcaccactgc	5820
agagcttcat gctccgggtt ggtgtttagt atgatccagt cgtacgagga ggcgtggca	5880
tggcctaa aatgtcctt cagcaggcagg ccgtggcca gggggaggcc cttgggtaa	5940
gtgtttacaa aacggtaag ttgggaagg tgcattcggg gagagatgtat gtgcatttg	6000
gactgtattt ttagattggc gatgtttccg cccagatccc ttctggatt catgttgtc	6060
aggaccacca gtacagtgtt tccggcgcac ttggggatt tgcacatgcag ctttagggaa	6120
aaagcgtgga agaacttggaa gacgcctttg tggcctccca gattttccat gcattcg	6180

[0116]

atgatgatgg caatgggccc gcgggaggca gcttggcaa agatatttct ggggtcgctg	6240
acgtcgtagt tgtgttccag ggtgaggtcg tcataggcca tttttacaaa ggcggggcgg	6300
agggtgcccgg actggggat gatggtcccc tctggccctg gggcgtagtt gcccctcgac	6360
atctgcattt cccaggcctt aatctcgag gggggaatca tatccacctg cggggcgatg	6420
aagaaaaacgg tttccggagc cggggagatt aactgggatg agagcaggtt tctaagcagc	6480
tgtgattttc cacaaccggt gggcccataa ataacaccta taaccggttt cagctggtag	6540
tttagagagc tgcagctgcc gtcgtcccg aggagggggg ccacctcggtt gagcatgtcc	6600
ctgacgcgca tggtctcccc gaccagatcc gccagaaggc gctgcccggc cagggacagc	6660
agctcttgca aggaagcaaa gttttcagc ggcttgaggc cgtccggcgtt gggcatgttt	6720
ttcagggtct ggctcagcag ctccaggcgg tcccagagct cggtgacgtg ctctacggca	6780
tctctatcca gcatatctcc tcggttcgag gggtggggcg actttcgctg tagggcacca	6840
agcggtggtc gtccagcggg gccagagtca tgtccttcca tgggcgcagg gtcctcgta	6900
gggtggctg ggtcacggtg aaggggtcg ctcgggctg agcgcttgc aaggtgcgt	6960
tgaggctggt tctgctggtg ctgaagcgct gccggcttc gcccgtcg tcggccaggt	7020
agcatttgac catgggttca tagtccagcc ctcggcggc gtgtcccttg ggcgcagct	7080
tgcccttgaa ggtggcgccg cacgaggggc agagcaggctt cttgagcgcg tagagcttgg	7140
ggcgaggaa gaccgattcg ggggagtagg cgtccgcgc gcagaccccg cacacggct	7200
cgcactccac cagccaggtg agctcggggc gcgcgggtc aaaaaccagg tttccccat	7260
gcttttgcgtt gctttctta cctcgggtct ccatgagggtt gtgtccccgc tcgggtacga	7320
agaggctgtc cgtgtctccg tagaccgact tgagggtct tttctccagg ggggtccctc	7380
ggtcttcctc gtagaggaac tcggaccact ctgagacgaa ggccgcgtc cagggcagga	7440
cgaaggaggc tatgtgggag gggtagcggt cgttgcac taggggttcc accttctcca	7500
agggtgtaa acacatgtcg cttcctcg cgtccaggaa ggtgattggc ttgttaggtgt	7560
aggccacgtg accgggggtt cctgacgggg gggtaaaaaa ggggtgggg gcgcgtcg	7620
cgtcactctc ttccgcattcg ctgtctgcga gggccagctg ctgggggtgag tattccctt	7680
cgaaggcggg catgacctcc gcgcgtgagggt tgtagtttc caaaaacgag gaggatttga	7740
tgttcacctg tcccgggtt ataccttga gggtacccgc gtccatctgg tcagaaaaca	7800
cgtatttttt attgtccagc ttgggtggcga acgaccgtt gaggcgtt gagagcagct	7860

[0117]

tggcgatgga	gcmcagggtc	tggttcttgt	ccctgtcggc	gcgcctccttgcgcatgt	7920
tgagctgcac	gtactcgccgc	gcgacgcagc	gccactcggtt	gaagacgggtgtgcgc	7980
cgggcaccag	gcmcacgcgc	cagccgcggt	tgtgcagggtt	gaccagggtcc	8040
cgacactcgcc	gcmcaggcgc	tcgttggtcc	agcagagacg	gccgccttgcgc	8100
agggggcag	ggggtcgagc	tgggtctcg	ccgggggtc	cgcgtccacg	8160
cggggcgcag	gcmcgcgtcg	aagtatcta	tcttgcacc	ttgcgtgtcc	8220
gccagtcgcg	ggcggcgagc	gcmcgtcg	aggggtttag	cggcgggccc	8280
ggtgtgttag	tgcggaggcg	tacatgcgc	agatgtata	gacgttagagg	8340
ggaccccgat	gttagtgggg	tagcagcggc	cggcgggtt	gctggcgcgc	8400
acagctcg	cgagggggcg	aggaggtcgg	ggcccagggtt	ggtgcggcgc	8460
cgcgaagac	gatctgcctg	aagatggcat	gcgagttgga	agagatgggt	8520
agacgttgaa	gctggcgtcc	tgcaggccg	cggcgtcg	cacgaaggag	8580
cgcgcagctt	gtgtaccagc	tcggcggtga	cctgcacgtc	gagcgcgcag	8640
tctcgccgtt	gatgtcatat	ttagcctgcc	ccttctttt	ccacagctcg	8700
caaactcttc	gcggctttc	cagtagtctt	ggatcggaa	accgtccgg	8760
aagagcctag	catgtagaac	tggttgacgg	cctggtaggc	gcagcagccc	8820
ggagggcgta	ggcctgcgcg	gccttgcgg	gcgagggtgt	ggtcaggcgc	8880
tgaccatgac	tttgaggta	tggtgcttga	agtcggagtc	gtcgcagccg	8940
agagcgagaa	gtcggcgcgc	ttcttggagc	gggggttggg	cagagcgaag	9000
tgaagaggat	tttgcggcgc	cgggcatga	agttgcgggt	gatgcggaa	9060
cttcagagcg	gttggat	acgtggcgg	cgagcacgt	ctcgctcgaag	9120
tgtggccac	gatgttagagt	tccaggaagc	ggggccggcc	ctttacgggt	9180
ttagctttc	gtaggtgagc	tcctcgccgc	aggcgaggcc	gtgctcgcc	9240
ccgcgaggtg	cggttgtct	ctgaggaagg	acttccagag	gtcgcggcc	9300
cgaggcggc	tctgaaggc	ctgaactggc	ggcccacggc	catttttgc	9360
agtagaagg	gaggggtct	tgctgccagc	ggtcccagtc	gagctgcagg	9420
gcgcggcggt	gaccaggcgc	tcgtcgcccc	cgaatttcat	gaccagcatg	9480
gctgcttcc	gaaggcccc	atccaagtgt	aggtctctac	atcgtaggtg	9540
gctccgtcgc	aggatgcgag	ccgatcggtt	agaactggat	ctccgcac	9600

[0118]

agtggctgtt	gatgtggtgg	aagtagaagt	cccgctcgccg	ggccgaacac	tcgtgctggc	9660
ttttgtaaaa	gcgagcgcag	tactggcagc	gctgcacggg	ctgtacacctca	tgcacgagat	9720
gcacctttcg	cccgcgacg	aggaagccga	ggggaaatct	gagccccccg	cctggctcgc	9780
ggcatggctg	gttctttct	actttggatg	cgtgtccgtc	tccgtctggc	tcctcgaggg	9840
gtgttacggt	ggagcggacc	accacgccgc	gcgagccgca	ggtccagata	tcggcgcg	9900
gcggtcggag	ttttagatgacg	acatcgcgca	gctgggagct	gtccatggc	tggagctccc	9960
gcggcggcgg	caggtcagcc	gggagttctt	gcaggttcac	ctcgcagagt	cgggccaggg	10020
cgcgccccag	gtctaggtgg	tacctgatct	ctagggcgt	gttggtggcg	gcgtcgatgg	10080
cttgcaggag	cccgcagccc	cggggggcga	cgacggtgcc	ccgcggggtg	gtgggtgg	10140
tggcggtgca	gctcagaagc	ggtgccgcgg	gcggggccccc	ggaggtaggg	ggggctccgg	10200
tcccgccggc	aggggcccga	gcggcacgtc	ggcgtggagc	gcgggcagga	gttggtgctg	10260
tgcccggagg	ttgctggcga	aggcgacgac	gcggcggtt	atctcctgga	tctggcgct	10320
ctgcgtgaag	acgacggcc	cggtgagctt	gaacctgaaa	gagagttcga	cagaatcaat	10380
ctcgggtgtca	ttgaccgcgg	cctggcgcag	gatctcctgc	acgtctcccg	agttgtctt	10440
gttaggcgatc	tcggccatga	actgctcgat	ctcttcctcc	tggaggtctc	cgcgtccggc	10500
gcgttccacg	gtggccgcca	ggtcggttga	gatgcgc	atgagctgcg	agaaggcg	10560
gagtccgccc	tcgttccaga	ctcggctgt	gaccacgccc	ccctggtcat	cgcggcg	10620
catgaccacc	tgcgcgaggt	ttagctccac	gtgccgcgc	aagacggcgt	agttgcgcag	10680
acgcttggaa	aggttagttga	gggtggtggc	ggtgtctcg	gccacgaaga	agttcatgac	10740
ccagcggcgc	aacgtggatt	cgttgatgtc	ccccaaaggcc	tccagccgtt	ccatggcctc	10800
gtagaagtcc	acggcgaagt	tgaaaaactg	ggagttgcgc	gccgacacgg	tcaactcctc	10860
ctccagaaga	cggatgagct	cggcgacggt	gtcgccgacc	tcgcgtcg	aggctatgg	10920
gatctttcc	tccgctagca	tcaccacctc	ctcctttcc	tcctttctg	gcacttccat	10980
gatggcttcc	tcctttcgg	gggggtggcgg	cggcgccgt	gggggagggg	gcgctctgcg	11040
ccggcggcgg	cgcaccggga	ggcggtccac	gaagcgcgc	atcatctcc	cgcggcggc	11100
gcgcatggtc	tgcgtgacgg	cgcggccgtt	ctccgggggg	cgcagttgga	agacgccg	11160
ggacatctgg	tgctggggcg	ggtggccgt	aggcagcgag	acggcgctga	cgtgcacat	11220
caacaattgc	tgcgttagt	cgccgcccag	ggacctgagg	gagtccat	ccaccggatc	11280

[0119]

cgaaaacctt	tcgaggaagg	cgtctaacca	gtcgcagtgc	caaggtaggc	tgagcaccgt	11340
ggcgggcggc	gggggggtggg	gggaggtgtct	ggcggaggtg	ctgctgatga	tgttaattgaa	11400
gtaggcggac	ttgacacggc	ggatggtcga	caggagcacc	atgtccttgg	gtccggcctg	11460
ctggatgcgg	aggcggtcgg	ctatgccccca	ggcttcgttc	tggcatcggc	gcaggtcctt	11520
gtagtagtct	tgcacatgagcc	tttccacccgg	cacctttcttct	ccttcttctt	ctgctttttc	11580
catgtctgtct	tcggccctgg	ggcggcgccg	cgccccctg	ccccccatgc	gcgtgacccc	11640
gaacccctg	agcggttgaa	gcagggccag	gtcggcgacg	acgcgctcgg	ccaggatggc	11700
ctgctgcacc	tgcgtgaggg	tggtttgaa	gtcatccaag	tccacgaagc	ggtggtaggc	11760
gcccgtgttgc	atggtgttagg	tgcagttggc	catgacggac	cagttgacgg	tctggtgcc	11820
cggttgcac	atctcggtgt	acctgagtcg	cgagtagggc	cgggagtcga	agacgtagtc	11880
tttgcaggatc	cgcaccaggat	actggtagcc	caccaggaag	tgcggcgccg	gctggcggtt	11940
gaggggcccag	cgcaggggtgg	cggggctcc	gggggcccagg	tcttccagca	tgaggcggtg	12000
gtaggcgttag	atgtacctgg	acatccaggat	gatacccgcg	gcgggtggtgg	aggcgccgg	12060
gaagtcgcgc	acccggttcc	agatgttgcg	cagggcaga	aagtgcctca	tggtaggcgt	12120
gctctgttcca	gtcagacgcg	cgcagtcgtt	gataactctag	accaggaa	acgaaagccg	12180
gtcagcgggc	actcttccgt	ggtctggta	atagatcgca	aggatcat	ggcggagggc	12240
ctcggttca	gccccgggtc	cggccggac	ggtccggcat	gatccacgcg	gttaccgccc	12300
gcgtgtcgaa	cccaggtgtg	cgacgtcaga	caacggtgaa	gtttccctt	tggcgtttt	12360
ctggccgggc	gccggcgccg	cgtaaagagac	taagccgcga	aagcgaaagc	agtaagtggc	12420
tcgctccccg	tagccggagg	gatecttgct	aagggttgcg	ttcggcgaa	ccccgggtcg	12480
aatcccgta	tcggccggc	cggacccgcg	gctaagggtgt	tggattggcc	tccccctcg	12540
ataaaagaccc	cgcttgcgga	ttgactccgg	acacggggac	gagcccttt	tatttttgt	12600
ttccccagat	gcatccggtg	ctgcggcaga	tgcgcgggg	gccccagcag	cagcaacaac	12660
accagcaaga	gcggcagcaa	cagcagcggg	agtcatgcag	ggccccccta	cccaccctcg	12720
gcggccggc	cacctcggcg	tccgcggccg	tgtctggcgc	ctgcggcgcc	ggcggggggc	12780
cggctgacga	ccccgaggag	cccccgccgc	gcagggccag	acactacctg	gacctggagg	12840
aggcgcagg	cctggcgccgg	ctggggcgc	cgtctcccg	gcccaccccg	cgggtgcagc	12900
tgaagcgca	ctcgcgcgag	gcgtacgtgc	ctcggcagaa	cctgttcagg	gaccgcgcgg	12960
gcgaggagcc	cgaggagatg	cggacagga	gttcagcgc	agggcgggag	ctgcggcagg	13020

[0120]

ggctgaaccg cgagcggctg ctgcgcgagg aggacttga gcccgacgcg cggacgggga	13080
tcaagccccgc gcgcgacac gtggcggccg ccgacacttgt gacggcgtac gagcagacgg	13140
tgaaccagga gatcaacttc caaaagagtt tcaacaacca cgtgcgcacg ctggtggcgc	13200
gcgaggaggt gaccatcggg ctgatgcacc tgtggactt tgtaagcgcg ctggtgcaga	13260
accccaacag caagcctctg acggcgcagc tggcctgat agtgcagcac agcagggaca	13320
acgaggcggt tagggacgcg ctgctgaaca tcaccgagcc cgagggtcgg tggctgctgg	13380
acctgattaa catcctgcag agcatagtgg tgcaggagcg cagcctgagc ctggccgaca	13440
aggtgtggc catcaactac tcgatgctga gcctggcaa gtttacgcg cgcaagatct	13500
accagacgcc gtacgtgcc atagacaagg aggtgaagat cgacggttt tacatgcgca	13560
tggcgctgaa ggtgctcacc ctgagcgcacg acctggcgt gtaccgcaac gagcgcattcc	13620
acaaggccgt gagcgtgagc cggcggcgcg agctgagcga ccgcgagctg atgcacagcc	13680
tgcagcgggc gctggcgggc gccggcagcg gcgcacaggga ggcggagtcc tactcgatg	13740
cggggcggga cctgcgctgg gcgcgcagcc ggcggccct ggaggccgcg ggggtccgcg	13800
aggactatga cgaggacggc gaggaggatg aggactacga gctagaggag ggcgagtacc	13860
tggactaac acgccccgtt gttccggta gatgaagac ccgaacgtgg tggacccggc	13920
gctgcggcgc gctctgcaga gccagccgtc cggcctaac tcctcagacg actggcgcaca	13980
ggtcatggac cgcacatgt cgctgacggc gcgttaaccg gacgcgttcc ggcagcagcc	14040
gcaggccaac aggctctccg ccatcctgga ggcggtggtg cctgcgcgt cgaaccccac	14100
gcacgagaag gtgctggcca tagtgaacgc gctggccgag aacagggcca tccgcggga	14160
cgaggccggg ctgggtacg acgcgtgt gcagcgcgtg gcccgtaca acagggcaa	14220
cgtgcagacc aacctggacc ggctggtggtt ggacgtgcgc gaggcggtgg cgcagcgcga	14280
gcgcgcggat cggcaggca acctggcgtc catggcggcg ctgaatgcct tcctgagcac	14340
gcagccggcc aacgtgccgc gggggcagga agactacacc aactttgtga gcgcgtgcg	14400
gctgatggtg accgagaccc cccagagcga ggtgtaccag tcggcccccgg actacttctt	14460
ccagaccagc agacagggcc tgcagacggt gaacctgagc caggcttca agaacctgcg	14520
ggggctgtgg ggcgtgaagg cggccaccgg cgaccggcgc acgggttcca gcctgctgac	14580
gcccaactcg cgcctgctgc tgctgctgat cgcgcgttc acggacagcg gcagcgtgtc	14640
ccgggacacc tacctggggc acctgctgac cctgtaccgc gaggccatcg ggcaggcgca	14700

[0121]

ggtggacgag cacaccttcc aggagatcac cagcgtgagc cgccgcgtgg ggcaggagga	14760
cacgagcagc ctggaggcga ctctgaacta cctgctgacc aaccggcggc agaagattcc	14820
ctcgctgcac agcctgaccc ctgaggagga gcgcatttg cgctacgtgc agcagagcgt	14880
gaggctgaac ctgatgcgac acgggggtgac gcccagcgtg gcgctggaca tgaccgcgc	14940
caacatggaa cgggcatgt acgcccgcga cggcattac atcaaccgcc tgatggacta	15000
cctgcatcgc gcggcggccg tgaacccga gtactttacc aacgccatcc tgaacccgc	15060
ctggctcccg cggccgggt tctacagcgg gggcttcgag gtcccgaga ccaacgatgg	15120
cttcctgtgg gacgacatgg acgacagcgt gttctcccg cggccgcagg cgctggcgga	15180
agcgtccctg ctgcgtccca agaaggagga ggaggaggag gcgagtcgcc gccgcggcag	15240
cagcggcgtg gttctctgt ccgagctggg ggcggcagcc gccgcgcgc ccgggtccct	15300
ggcggcagc cccttccga gcctggtggg gtctctgcac agcagcgcga ccacccgc	15360
tcggctgtg ggcaaggacg agtacctgaa taactccctg ctgcagccgg tgccggagaa	15420
aaacctgcct cccgccttcc ccaacaacgg gatagagagc ctggtgaca agatgagcag	15480
atgaaagacc tatgcgcagg agcacaggga cgcgcctgcg ctccggccgc ccacgcggcg	15540
ccagcgccac gaccggcagc ggggctgggt gtggatgac gaggactccg cggacgatag	15600
cagcgtgctg gacctggag ggagcggcaa cccgttcgcg cacctgcgcc cccgcctgg	15660
gaggatgttt taaaaaaaaaaa aaaaaaaaaagc aagaagcatg atgaaaaat taaataaaac	15720
tcaccaaggc catggcgtacc gagcgttggt ttcttgtt ccctcagta tgccgcgc	15780
ggcgatgtac caggagggac ctccctccctc ttacgagagc gtggtgccgc cggcggccgc	15840
ggcccccctct tctccctttg cgtcgagct gctggagccg ccgtacgtgc ctccgcgcta	15900
cctgccccct acggggggga gaaacagcat cggtaactcg gagctggcgc ccctgttgc	15960
caccacccgg gtgtacctgg tggacaacaa gtcggcggac gtggctccc tgaactacca	16020
gaacgaccac agcaatttt tgaccacggt catccagaac aatgactaca gcccgagcga	16080
ggccagcacc cagaccatca atctggatga cggcgcac tggggcggcgc acctgaaaac	16140
catcctgcac accaacatgc ccaacgtaa cgagttcatg ttaccaata agttcaaggc	16200
gcgggtgatg gtgtcgct cgcacaccaa ggaagaccgg gtggagctga agtacgagtg	16260
ggggagttc gagctgccag agggcaacta ctccgagacc atgaccattg acctgatgaa	16320
caacgcgatc gtggagcact atctgaaagt gggcaggcag aacggggtcc tggagagcga	16380
catcggggtc aagttcgaca ccaggaactt ccgcctgggg ctggaccccg tgaccggc	16440

[0122]

ggtatgccc	gggtgtaca	ccaacgaggc	cttccatccc	gacatcatcc	tgctgcccgg	16500
ctcggggtg	gacttcactt	acagccgcct	gagcaacctc	ctgggcattc	gcaagcggca	16560
gcccttcag	gagggcttca	ggatcaccta	cgaggacctg	gagggggca	acatccccgc	16620
gctcctcgat	gtggaggcct	accaggatag	cttgaaggaa	aatgaggcgg	gacaggagga	16680
taccgcccc	gccgcctccg	ccgcccggca	gcagggcggag	gatgctgctg	acaccgcggc	16740
cgccggacggg	gcagaggccg	accccgctat	ggtggggag	gctcccggagc	aggaggagga	16800
catgaatgac	agtgcggtgc	gccccggacac	cttcgtcacc	cggggggagg	aaaagcaagc	16860
ggaggccgag	gccgcggccg	aggaaaagca	actggcggca	gcagcggcgg	cggccggcgtt	16920
ggccgcggcg	gaggctgagt	ctgaggggac	caagccggcc	aaggagcccg	tgattaagcc	16980
cctgaccgaa	gatagcaaga	agcgcagtta	caacctgctc	aaggacagca	ccaacaccgc	17040
gtaccgcagc	tggtacctgg	cctacaacta	cgccgaccccg	tcgacggggg	tgcgctcctg	17100
gaccctgctg	tgcacgcgg	acgtgacctg	cggctggag	caggtgtact	ggtcgctgcc	17160
cgcacatgatg	caagaccccg	tgaccttccg	ctccacgcgg	caggtcagca	acttcccggt	17220
ggtgggcgcc	gagctgctgc	ccgtgcactc	caagagcttc	tacaacgacc	aggccgtcta	17280
ctcccgctc	atccgcagt	tcacctctct	gaccacgtg	ttcaatcgct	ttcctgagaa	17340
ccagattctg	gcgcgcggc	ccgcggccac	catcaccacc	gtcagtgaaa	acgttcctgc	17400
tctcacagat	cacggacgc	taccgctgctg	caacagcatc	ggaggagttcc	agcgagtgac	17460
cgttactgac	gccagacgcc	gcacctgccc	ctacgtttac	aaggccttgg	gcatagtctc	17520
gccgcgcgtc	ctttccagcc	gcactttttg	agcaacacca	ccatcatgtc	catcctgate	17580
tcacccagca	ataactccgg	ctggggactg	ctgcgcgcgc	ccagcaagat	gttcggaggg	17640
gcgaggaagc	gttccgagca	gcacccctgt	cgcgtgcgcg	ggcacttccg	cgccccctgg	17700
ggagcgcaca	aacgcggccg	cgcggggcgc	accaccgtgg	acgacgccat	cgactcgggt	17760
gtggagcagg	cgcgcacta	caggcccgcg	gtctctaccg	tggacgcggc	catccagacc	17820
gtgggtgcggg	gcgcgcggcg	gtacgccaag	ctgaagagcc	gccggaagcg	cgtggccgc	17880
cgcaccgcgc	gccgacccgg	ggccgcgcgc	aaacgcgcgc	ccgcggccct	gcttcgcgcgg	17940
gccaaggcgca	cggccgcgcg	cgccgcgcgt	aggccgcgc	gccgcttgc	cgccggcatc	18000
accgcgcaca	ccatggcccc	ccgtacccga	agacgcgcgg	ccgcgcgcgc	cgccgcgcgc	18060
atcagtgaca	tggccagcag	gcgcgcgggc	aacgtgtact	gggtgcgcga	ctcggtgacc	18120

[0123]

ggcacgcg	tgcccgtg	cttccgcccc	ccgcggactt	gagatgtatgt	aaaaaaacaa	18180
cactgagtct	cctgctgtt	tgtgtatccc	agcggcggcg	gcgcgcgcag	cgtcatgtcc	18240
aagcgcaaaa	tcaaagaaga	gatgctccag	gtcgtcg	cgagatcta	tggcccccg	18300
aagaaggaag	agcaggattc	gaagccccgc	aagataaagc	ggtcaaaaaa	aaaaaagaaa	18360
gatgatgacg	atgccgatgg	ggaggtggag	ttcctgcg	ccacggcgcc	caggcggcc	18420
gtcagtg	agggccggcg	cgtaaagcgc	gtcctgcg	ccggcaccgc	gtgg	18480
acgcccggcg	agcgctccac	ccggactt	aagcgct	atgacgaggt	gtacggcgac	18540
gaagac	tggagcaggc	caacgagcgc	ttcggagagt	ttgttacgg	gaagcg	18600
cggcg	ggaaggagga	cctgctggcg	ctgccc	accaggc	ccccac	18660
agtctga	ccgtgacc	gcagcagg	ctgccc	gcgcacc	cgaggc	18720
cgggtctg	agcgcgagg	cggc	gcgc	tgca	gtgccc	18780
cggcagaggc	tggaggatgt	gctggagaaa	atgaa	acc	gcagccgg	18840
atcagg	gccccatcaa	gcaggtggcg	ccgg	gcgt	cgac	18900
gtcat	ccggcaactc	ccccggcc	gccacc	ccg	ctg	18960
gagac	ccgatcccgc	cgcagccgca	gccgc	ccg	ccgc	19020
gaggt	cggacc	cgtggcc	gcgat	ctcc	gcgg	19080
cgc	acggccgc	caac	ctgc	acgc	tc	19140
ccgc	ccggctacc	agg	catt	acc	ttcc	19200
gcac	gcgc	ccgc	ccgc	ccgc	ccgc	19260
gcac	ctccgt	gagg	aaagt	gcgc	gac	19320
agg	acc	cc	atc	cgtt	ttgc	19380
ctca	tcg	ccgtt	cc	ggat	acc	19440
agg	ccgg	ccgg	c	gcgc	acc	19500
ccac	ac	ccat	cc	gcgc	ccgt	19560
gcgat	ccgt	cccg	gat	cc	tt	19620
cac	tt	cc	tc	gt	cc	19680
cgctcg	gtc	ctgt	at	ttgt	at	19740
cccg	gtc	cg	at	ttt	tt	19800
gac	ttc	at	ttt	gtt	ttt	19860

[0124]

cgtaaaaat tacggctccc gggcctggaa cagcagcacg ggccagatgt tgagagacaa	19920
gttcaaagag cagaacttcc agcagaaggt ggtggagggc ctggcctccg gcatcaacgg	19980
ggtgtggac ctggccaacc aggccgtgca gaataagatc aacagcagac tggacccccc	20040
gccgccggtg gaggaggtgc cgccggcgct ggagacggtg tccccgatg ggcgtggcga	20100
gaagcgcccg cggcccgata gggaaagagac cactctggc acgcagaccc atgagccgcc	20160
cccgtatgag gaggccctga agcaagggtct gccaccacg cggcccatcg cgcccatggc	20220
caccgggtg gtgggcccgc acaccccccgc cacgctggac ttgcctccgc cgcgcgtgt	20280
gccgcagcag cagaaggcgg cacagccggg cccgcccgcg accgcctccc gttcctccgc	20340
cggcctctg cgccgcgcgg ccagcggccc ccgcgggggg gtcgcgaggc acggcaactg	20400
gcagagcacg ctgaacagca tcgtgggtct ggggtgcgg tccgtgaagc gccggcgatg	20460
ctactgaata gcttagctaa cgtgttgtat gtgttatgc gccctatgtc gccggcagag	20520
gagctgtga gtcgcgcgcg ttgcgcgcgc caccaccacc gccactccgc ccctaagat	20580
ggcgacccca tcgatgatgc cgcaagtggc gtacatgcac atctcgggcc aggacgcctc	20640
ggagttacctg agccccggc tggtcagtt cgcccgcc accgagagct acttcagcct	20700
gagtaacaag ttttaggaacc ccacggtggc gcccacgcac gatgtgacca ccgaccggc	20760
tcagcgcctg acgctgcggt tcattccgt ggaccgcgag gacaccgcgt actcgtacaa	20820
ggcgcgggttc accctggccg tggcgacaa ccgcgtctg gacatggcct ccacctactt	20880
tgacatccgc ggggtgctgg accggggtcc cactttcaag ccctactctg gcaccgccta	20940
caactccctg gcccccaagg ggcgtcccaa ctcctgcgag tggagcaag agaaactca	21000
ggcagttgaa gaagcagcag aagaggaaga agaagatgct gacggtaag ctgaggaaga	21060
gcaaggagct accaaaaaga ctcatgtata tgctcaggct cccctttctg gcgaaaaaat	21120
tagtaaagat ggtctgcaaa taggaacgga cgctacagct acagaacaaa aacattttta	21180
tgcagaccct acattccagc ccgaacccca aatcggggag tcccagtggaa atgaggcaga	21240
tgctacagtc gcccgggta gagtgtctaa gaaatctact cccatgaaac catgtatgg	21300
ttcctatgca agacccacaa atgctaattgg aggtcagggt gtactaacgg caaatgccc	21360
gggacagcta gaatctcagg ttgaaatgca attctttca acttctgaaa acgcccgtaa	21420
cgaggctaac aacattcagc ccaaattggt gctgtatagt gaggatgtgc acatggagac	21480
cccgatacg caccttctt acaagcccgc aaaaagcgat gacaattcaa aaatcatgct	21540

[0125]

gggtcagcag tccatgccca acagaccta ttacatcgcc ttcaagagaca actttatcg	21600
cctcatgtat tacaatagca ctggcaacat gggagtgc ttcaaggtcagg cctctcagg	21660
aatgcagtg gtggacttgc aagacagaaa cacagaactg tcctaccagg tcggcttga	21720
ttccatgggt gacagaacca gatactttc catgtggaa caggcagtgg acagttatga	21780
cccagatgtt agaattattt aaaaatcatgg aactgaagac gagctccccca actattttt	21840
ccctctgggt ggcatagggg taactgacac ttaccaggct gttaaaacca acaatggcaa	21900
taacggggc caggtgactt ggacaaaaga tgaaactttt gcagatcgca atgaaatagg	21960
ggtggaaac aatttcgcta tggagatcaa cctcagtgc aacctgtggaa gaaacttct	22020
gtactccaac gtggcgctgt acctaccaga caagcttaag tacaaccctt ccaatgtgg	22080
catctctgac aaccccaaca cctacgatta catgaacaag cgagtgggtt ccccgggct	22140
ggtgactgc tacatcaacc tggcgcgcg ctggcgctg gactacatgg acaacgtcaa	22200
cccttcaac caccaccgca atgcggcct gcgcgtaccgc tccatgtcc tggcaacgg	22260
gcgcgtacgt cccttccaca tccaggtgcc ccagaagttc tttccatca agaacctct	22320
cctcctgcgc ggctcctaca cctacgagt gaaactcagg aaggatgtca acatggctt	22380
ccagagctt ctggtaacg atctcagggt ggacggggcc agcatcaagt tcgagagcat	22440
ctgcctctac gccacccctt tccccatggc ccacaacacg gcctccacgc tcgaggccat	22500
gctcaggaaac gacaccaacg accagtcctt caatgactac ctctccgcgc ccaacatgt	22560
ctacccata cccgccaacg ccaccaacgt cccatctcc atcccctcgc gcaactggc	22620
ggccttccgc ggctggcct tcacccgc caagaccaag gagacccctt ccctggc	22680
gggattcgac ccctactaca cctactcggt ctccattccc tacctggacg gcacccctta	22740
cctcaaccac actttcaaga aggtctcggt caccttcgac tcctcggtca gctggccgg	22800
caacgaccgt ctgctcaccc ccaacgagtt cgagatcaag cgctcggtcg acggggaggg	22860
ctacaacgtg gcccagtgc acatgaccaa ggactggttc ctggtccaga tgctggccaa	22920
ctacaacatc ggctaccagg gcttctacat cccagagagc tacaaggaca ggatgtactc	22980
cttcttcagg aacttccagc ccatgagccg gcaggtgggt gaccagacca agtacaagga	23040
ctaccaggag gtggcatca tccaccagca caacaactcg ggcttcgtgg gctacccgc	23100
ccccaccatg cgcgaggac aggctaccc cgccaaacttc ccctatccgc tcataggcaa	23160
gaccgcggc gacagcatca cccagaaaaa gttcctctgc gaccgcaccc tctggcgc	23220
cccttctcc agcaacttca tgtccatggg tgcgcctcg gacctggcc agaacttgc	23280

[0126]

ctacgccaac	tccgcccacg	ccctcgacat	gacttcgag	gtcgacccca	tggacgagcc	23340
cacccttctc	tatgttctgt	tcgaagtctt	tgacgtggc	cgggtccacc	agccgcacccg	23400
cggcgcatc	gagaccgtgt	acctgcgtac	gcccttctcg	gccggcaacg	ccaccaccta	23460
aagaagcaag	ccgcagtcat	cggcgctgc	atgccgtcg	gttccaccga	gcaagagctc	23520
agggccatcg	tcagagacct	gggatgcggg	ccctatTTT	tggcacctt	cgacaagcgc	23580
ttccctggct	ttgtctcccc	acacaagctg	gcctgcgcca	tcgtcaacac	ggccggccgc	23640
gagaccgggg	gcgtgcactg	gctggccttc	gcctggaacc	cgcgctccaa	aacatgcttc	23700
ctctttgacc	cttcggctt	ttcggaccag	cggctcaagc	aatctacga	gttcgagttac	23760
gagggcttgc	tgcgtcgac	cggcatcgcc	tcctcgcccg	accgctgcgt	caccctcgaa	23820
aagtccaccc	agaccgtgca	ggggcccgac	tcggccgcct	gcggctcttt	ctgctgcatg	23880
tttctgcacg	ccttttgca	ctggcctcag	agtcccatgg	accgcaaccc	caccatgaac	23940
ttgctgacgg	gggtgcccaa	ctccatgctc	cagagccccc	aggtcgagcc	caccctgcgc	24000
cgcaaccagg	agcagctcta	cagttcctg	gagcggcact	cgccttactt	ccgcccacac	24060
agcgcacaga	tcaaggaggc	caccccttc	tgccacttgc	aagagatgca	agaaggtaa	24120
taacgatgta	cacactttt	ttctcaataa	atggcatctt	tttatttata	caagctctct	24180
gggttattca	tttcccacca	ccacccgccc	ttgtcgccat	ctggctctat	ttagaaatcg	24240
aaagggttct	gccgggagtc	gccgtgcgc	acgggcaggg	acacgttgc	atactggtag	24300
cgggtgcccc	acttgaactc	gggcaccacc	aggcgaggca	gctcggggaa	gtttcgctc	24360
cacaggctgc	gggtcagcac	cagcgcgttc	atcaggtcgg	gcgcgcagat	cttgaagtcg	24420
cagttgggc	cggccccc	cgcgcgcgag	ttgcgttaca	ccgggttgca	gcactggaac	24480
accaacagcg	ccgggtgctt	cacgctggcc	agcacgtgc	gtcggagat	cagctcggcg	24540
tccaggtcct	ccgcgttgct	cagcgcgaac	ggggcatct	tggcacttg	ccgcggcagg	24600
aagggcgcgt	ccccgggtt	cgagttgcag	tcgcagcgca	gcgggatcag	caggtgcccc	24660
tgcccgact	cgcggtggg	gtacagcgcg	cgcataagg	cctgcacatc	gcggaaaggcc	24720
atctgggcct	tggccccc	cgagaagaac	atgccgcagg	acttgcggca	gaactggttt	24780
gcggggcagc	tggcgctgt	cagggcagcag	cgcgcgtcg	tgttggcgat	ctgcaccacg	24840
ttgcgcggcc	accggttctt	cacgatcttgc	gccttggacg	attgctcctt	cagcgcgcgc	24900
tgcccggttct	cgctggtcac	atccatctcg	atcacatgtt	ccttggcac	catgctgctg	24960

[0127]

ccgtgcagac acttcagctc gccctccgtc tcgggtgcagc ggtgtgcaca cagcgccag	25020
cccggtggct cgaaagactt gtaggtcacc tccgcgaagg actgcaggta cccctgaaa	25080
aaggccccca tcatggtcac gaaggcttg ttgctgctga aggtcagctg cagccgcgg	25140
tgctccttgt tcagccaggt cttgcacacg gccgcacagc cttccacactg gtcggcagc	25200
atcttgaagt tcacccatcg ctcatttc acgtggtaact tgccatcag cgtgcgcgc	25260
gcctccatgc cttctccca ggccgacacc agcggcaggc tcacgggtt cttcaccatc	25320
accgtggccg ccgcctccgc cgcgcttgc cttccgccc cgctgttctc ttccctttcc	25380
tcctcttctt cggccggcc cactcgacgc cccgcacca cggggctgctc ttccctgagg	25440
cgctgcacct tgcgcttgcc gttgcgcggc tgcttgatgc gcacggcgg gttgtgaag	25500
ccaccatca ccagcgcggc cttttcttgc tcgtcctcgc tgccagaat gacctccgg	25560
gaggggggggt tggtcatctt cagtaaccggc gcacgcttct ttttcttctt gggggcggttc	25620
gccagctccg cggctgcggc cgctgcccgg gtcgaaggcc gagggctggg cgtgcgcggc	25680
accagcgcgt cctgcgagcc gtccctcgcc tcctcgact cgagacggag gcggcccccgc	25740
ttcttcgggg gcgcgcgggg cggcgaggc ggcggccggc acggagacgg ggacgagaca	25800
tcgtccaggg tgggtggacg gcggccgcg ccgcgtccgc gtcgggggtt ggtctcgcc	25860
tggtcctttt cccgactggc catctccac tgctccttct cctataggca gaaagagatc	25920
atggagtctc tcatgcgagt cgagaaggag gaggacagcc taaccgcggc ctctgagccc	25980
tccaccaccc cggccaccac cgccaatgcc gccgcggacg acgcgcaccc acggccacc	26040
gccagtagcca ccctccccag cgacgcaccc cgcgtcgaga atgaagtgt gatcgaggc	26100
gaccgggtt ttgtgagcgg agaggaggat gaggtggatg agaaggagaa ggaggaggc	26160
gccgcctcag tgccaaaaga ggataaaaag caagaccagg acgacgcaga taaggatgag	26220
acagcagtcg ggcggggaa cggaaaggcat gatgctgatg acggctaccc agacgtggaa	26280
gacgacgtgc tgcttaagca cctgcaccgc cagtgcgtca tcgtctgcga cgctgcag	26340
gagcgctgca aagtccccct ggacgtggcg gaggtcagcc gcgcctacga gcggcaccc	26400
ttcgcgcgc acgtcccccc caagcgccgg gagaacggca cctgcgagcc caaccgcgt	26460
ctcaacttct acccggttcc cgccgtaccc gaggtgctgg ccacctacca catttttc	26520
caaaaactgca agatccccct ctcctgcgcg gcacaaccgc cccgcgcgca caaaaaccctg	26580
accctgcggc agggcgccca catacctgat atgcctctc tggaggaagt gcccaagatc	26640
ttcgagggtc tcgggtcgca cgagaaacgg gcggcgaacg ctctgcacgg agacagcga	26700

[0128]

aacgagagtc actcggggt gctgggtggag ctcgagggcg acaacgcgcg cctggccgta	26760
ctcaagcgca gcatagaggt cacccacttt gcctaccgg cgctcaacct gcccccaag	26820
gtcatgagtg tggcatggg cgagctcatc atgcgccgcg cccagccct ggccgcggat	26880
gcaaacttgc aagagtctc cgaggaaggc ctgcccgg tcagcgacga gcagctggcg	26940
cgctggctgg agacccgcga ccccgccag ctggaggagc ggcgcaagct catgatggcc	27000
gcggtgctgg tcaccgtgga gctcgagtgt ctgcagcgct tcttcgcgga ccccgagatg	27060
cagcgcaagc tcgaggagac cctgcactac accttccgc agggctacgt ggcgcaggcc	27120
tgcaagatct ccaacgtgga gctctgcaac ctggctctt acctggcat cctgcacgag	27180
aaccgcctcg ggcagaacgt cctgcactcc accctcaaag gggaggccgcg ccgcgactac	27240
atccgcact ggccttaccc ttccctctgc tacacctggc agacggccat ggggtctgg	27300
cagcagtgcc tggaggagcg caacctaag gagctggaaa agctcctcaa ggcacccctc	27360
agggacctct ggacgggctt caacgagcgc tcgggtggccg ccgcgtggc ggacatcatc	27420
tttcccggc gcctgctcaa gaccctgcag cagggcttcg ccgacttcac cagccagagc	27480
atgctgcaga acttcaggac tttcatcctg gagcgctcg gcattctgcc ggccacttgc	27540
tgccgcgtgc ccagcgactt cgtccccatc aagtacaggg agtgcggcc ggcgtctgg	27600
ggccactgct acctttcca gctggccaac tacctcgctt accactcgga cctcatggaa	27660
gacgtgagcg gcgagggct gctcgagtgc cactggcgct gcaacctctg cacccccac	27720
cgctctctag tctgcaaccc gcagctgctc agcgagagtc agattatcg taccttcgag	27780
ctgcagggtc cctgcctga cgagaagtcc gggctccag ggctgaaact cactccgggg	27840
ctgtggactt ccgccttaccc acgcaaattt gtacctgagg actaccacgc ccacgagatc	27900
aggttctacg aagaccaatc ccgcggccccc aaggcggagc tcaccgcctg cgtcatcacc	27960
caggggcaca tcctgggcca attgcaagcc atcaacaaag cccgcccaga gttcttgctg	28020
aaaaagggtc ggggggtgta cctggacccc cagtcggcg aggagctaaa cccgctaccc	28080
ccggccgcgc cccagcagcg ggaccttgc tccaggatg gcacccagaa agaagcagca	28140
gccggccgcg ccgcccgcagc catacatgct tctggaggaa gaggaggagg actggacag	28200
tcaggcagag gaggttcgg acgaggagca ggaggagatg atggaagact gggaggagga	28260
cagcagccta gacgaggaag cttcagaggc cgaagaggtg gcagacgcaa caccatcgcc	28320
ctcggtcgca gccccctcgc cggggccctt gaaatcctcc gaacccagca ccagcgctat	28380

[0129]

aacctccgtc	cctccggcgc	cggcgccacc	cgcgcgaga	ccaaaccgta	gatggacac	28440
cacaggaacc	gggttcggta	agtccaagt	ccgcgcgcg	ccacccgac	agcagcagca	28500
gcagcgccag	ggctaccgct	cgtggcgcgg	gcacaagaac	gccatagtcg	cctgcttgc	28560
agactgcggg	ggcaacatct	cttcgcggc	ccgcttctg	ctattccacc	acggggtcgc	28620
cttccccgc	aatgtctgc	attactaccg	tcatctctac	agccctact	gcagcggcga	28680
cccagaggcg	gcagcggcag	ccacagcggc	gaccaccacc	taggaagata	tcctccgcgg	28740
gcaagacagc	ggcagcagcg	gccaggagac	ccgcggcagc	agcggcgga	gcggtggcgc	28800
cactgcgcct	ctgcggcaac	gaaccctct	cgaccggga	gctcagacac	aggatcttc	28860
ccactttgta	tgcgcatttc	caacagagca	gaggccagga	gcaggagctg	aaaataaaaa	28920
acagatctct	gcgcgcctc	acccgcagct	gtctgtatca	caaaagcgaa	gatcagcttc	28980
ggcgcacgct	ggaggacgcg	gaggcactct	ttagcaata	ctgcgcgc	actcttaaag	29040
actagctccg	cgccttctc	gaatttaggc	gggagaaaac	tacgtcatcg	ccggccgcgc	29100
cccagccgc	ccagccgaga	ttagcaaaaga	gattccacg	ccatacatgt	ggagctacca	29160
gccgcagatg	ggactcgccg	cggagcggc	ccaggactac	tccacccgca	tgaactacat	29220
gagcgcggga	ccccacatga	tctcacaggt	caacggatc	cgcccccagc	gaaaccaaatt	29280
actgctggaa	caggcggcca	tcaccgcac	gccccccat	aatctcaacc	ccgaaatttgc	29340
gccccccgc	ctcggttacc	agaaaacccc	ctccgcacc	accgtactac	ttccgcgtga	29400
cggccaggcc	gaagtccaga	tgactaactc	aggggcgcag	ctcgccggcg	gcttcgtca	29460
cggggcgcgg	ccgctccgac	caggtataag	acacctgatg	atcagaggcc	gaggtatcca	29520
gctcaacgac	gagtcggta	gcttcgtct	cgggtccgt	ccggacggaa	ctttccagct	29580
cgcggatcc	ggccgcttt	cgttcacgccc	ccgcccaggcg	tacctgactc	tgcagacctc	29640
gtcctcgag	ccccgctccg	gcggcatcg	aaccctccag	ttcgtggagg	agttcggtcc	29700
ctcggtctac	ttcaacccct	tctcgacc	tcccgacgc	taccccgacc	agttcattcc	29760
gaactttgac	gcggtaagg	actcgccgga	cggctacgac	tgaatgtcag	gtgtcgaggc	29820
agagcagctt	cgcctgagac	acctcgagca	ctgcccgc	cacaagtgc	tgcggccgg	29880
ttctggtgag	ttctgtact	ttcagctacc	cgaggagcat	accgaggggc	cgccgcacgg	29940
cgtccgcctg	accacccagg	gcgaggttac	ctgtccctc	atccggaggt	ttaccctccg	30000
tccctgcta	gtggagcggg	agcggggtcc	ctgtgtctta	actatgcct	gcaactgccc	30060
taaccctgga	ttacatcaag	atcttgctg	tcatctctgt	gctgagttt	ataaacgctg	30120

[0130]

agatcagaat ctactgggc tcctgtcgcc atcctgtcaa cgccaccgtc ttcacccacc	30180
ccgaccaggc ccaggcgaac ctcacctgcg gtctgcacatg gagggccaag aagtaccta	30240
cctggtaactt caacggcacc ccctttgtgg tttacaacag cttcgacggg gacggagtct	30300
ccctgaaaga ccagctctcc ggtctcagct actccatcca caagaacacc accctccaac	30360
tcttccctcc ctacctgccc ggaacctacg agtgcgtcac cggccgctgc acccaccta	30420
cccgccctgat cgtaaaccag agctttccgg gaacagataa ctccctcttc cccagaacag	30480
gaggtgagct cagaaactc cccggggacc agggcggaga cgtaccttcg acccttgcgg	30540
ggtaggatt ttttattacc gggttgctgg ctctttaat caaagttcc ttgagatttg	30600
ttcttcctt ctacgtgtat gaacacctca acctccaata actctaccct ttcttcggaa	30660
tcaggtgact tctctgaaat cgggcttggt gtgctgctta ctctgttgc tttttcctt	30720
atcataactca gccttctgtg cctcaggctc gccgcctgct gcgcacacat ctatatctac	30780
tgctggttgc tcaagtgcag gggtcgccc ccaagatgaa caggtacatg gtcctatcga	30840
tcctaggcct gctggccctg gcccctgca gcgcgccaa aaaagagatt accttgagg	30900
agcccgcttgaatgttaact ttcaagcccg agggtgacca atgcaccacc ctcgtcaa	30960
gcgttaccaa tcatgagagg ctgcgcacatg actacaaaaaa caaaaactggc cagttgcgg	31020
tctatagtgt gtttacgccc ggagaccct ctaactactc tgtcaccgtc ttccaggcg	31080
gacagtctaa gatattcaat tacacttcc ctttttatga gttatgcgt gcggcatgt	31140
acatgtcaaa acagtacaac ctgtggcctc cctctccca ggcgtgtgt gaaaatactg	31200
ggcttactg ctgtatggct ttgcataatca ctacgctcgc tctaactgc acgggtctat	31260
acataaaaatt caggcagagg cgaattttgcataatgaaa gaaaatgcct tgatgcata	31320
caccggctt ctatctgcag aatgaatgca atcacccc tactaatcac caccaccctc	31380
cttgcgattt cccatgggtt gacacgaatc gaagtgcac tgggtccaa tgtcaccatg	31440
gtggggcccg ccggcaattt caccctcatg tggaaaaat ttgtccgcaa tcaatgggtt	31500
catttctgtct ctaaccgaat cagttatcaag cccagagcca tctgcgtatgg gcaaaatcta	31560
actctgtatca atgtgcataat gatggatgct ggttactatt acgggcagcg gggagaaatc	31620
atataattact ggcgacccca caaggactac atgcgtcatg tagtcgaggc acttccact	31680
accacccca ctaccaccc tccaccacc accaccacta ctactactac tactactact	31740
actactacta ccactaccgc tgccgcatt acccgaaaaa gcaccatgt tagcacaaag	31800

[0131]

ccccctcgta	ctca	ccccccca	cgccggcggg	cccatcggt	cgac	ctcaga	aacc	accgag	31860	
ctttgcttct	gcca	aatgcac	taac	gccc	agc	gtc	atgaa	tgag	31920	
gatgtccagc	agag	ctccgc	ttgc	cctgacc	cagg	aggct	tgg	agccgt	31980	
cagatcggt	attc	aataat	tgact	tttct	tctt	ttgcca	ctcc	gaata	32040	
tctactttcc	acat	caacggg	tacca	aaagac	ccta	accct	ctt	tac	32100	
ctctgtatct	ctgt	gggtctc	ttcc	cgctg	atgt	tactgg	ggat	gttctg	32160	
tgccgcagaa	agaga	aaaagc	tcg	ctctc	ag	gcc	aaaccac	tgt	gccc	32220
ccggat	tttgc	agataacaa	gat	atgag	ct	cg	tgt	acta	accg	32280
g	ct	ct	ct	act	ac	cc	at	ac	cc	32340
aaaatgttac	ttt	caactcc	acgg	ccgata	ccc	agtg	gg	tt	gg	32400
acttaactat	ctg	caat	atgc	tcc	act	ccc	ccg	atata	ccaa	32460
atgccagcct	gtt	cacc	tc	atca	acg	ctt	cc	acc	tgg	32520
atgtaccctt	tggt	ggg	caa	gaa	agac	cc	ac	gttac	aa	32580
gaaccactac	cca	agctt	t	cc	acc	cca	cc	acc	ac	32640
gcagcagcag	cc	ac	agc	agc	at	tgt	ttt	gg	cc	32700
ccgctaccca	gg	cc	at	ct	ta	ct	cc	ac	cc	32760
cgaccaccgc	cac	acc	c	c	ta	cc	g	ac	cc	32820
tggctttca	aat	ggg	actt	aca	ggcc	aa	acc	at	cc	32880
ccgccc	ct	gt	ct	ca	tt	gg	cc	at	gt	32940
ccatctatag	acc	cat	catt	gt	c	ct	cc	cg	cc	33000
gcctgaaaaa	cct	act	tttt	tct	ttt	ac	at	gt	at	33060
cttgcata	ttc	c	ttt	ttt	tt	tt	tt	tt	tt	33120
cactctcatc	tc	g	cc	ct	cc	tt	cc	tt	cc	33180
ctgtgtgcgc	ct	gc	ca	ct	cc	tt	cc	tt	cc	33240
acttctaaga	ct	g	ct	ta	at	ca	tc	at	ca	33300
ctggcc	cc	ac	cc	c	cc	tt	cc	tt	cc	33360
tcac	cc	cc	cc	cc	cc	cc	cc	cc	cc	33420
tgccg	cc	cc	cc	cc	cc	cc	cc	cc	cc	33480
ttca	cc	cc	cc	cc	cc	cc	cc	cc	cc	33540

[0132]

gaagcttggc tgtatgggt catctgtgc tttagtttct gcagcactgt ctggccctc	33600
ataatctacc cctactttga tttgggatgg aacgcgatcg atgccatgaa ttaccccacc	33660
tttccgcac ccgagataat tccactgcga caagttgtac ccgttgcgt taatcaacgc	33720
cccccatccc ctacgcccac taaaatcagc tacttaacc 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 cgctcagtt acaaattgcc cacccagcgc cagaagctgg tgctcatgg	34020
gggtgagaat cccatcaccc tcacccagca ctggtagag accgaggggt gtctgcactc	34080
ccccgtcgg ggtccagaag acctctgcac cctggtaaag accctgtcgc gtctcagaga	34140
tttagtcccc tttaactaat caaacactgg aatcaataaa aagaatcact tactaaaat	34200
cagacagcag gtctctgtcc agtttattca gcagcacctc cttccctcc tcccaactct	34260
ggtactccaa acgccttctg gggcaaaact tcctccacac cctgaaggaa atgtcagatt	34320
cttgctcctg tccctccgca cccactatct tcatgttgc ttgcagatgaag cgccacaaaa	34380
cgctcgtcga gagcttcaac cccgtgtacc cctatgacac ggaaagcggc cctccctccg	34440
tcccttcct caccctccc ttcgtgtc tcgatggatt ccaagaaagt ccccccgggg	34500
tcctgtctc tAACCTGGCC gagccccctgg tcacttcca cggcatgctc gcccgtaaaa	34560
tggaaagtgg cctctccctg gacgacgctg gcaacacctac ctctcaagat atcaccaccc	34620
ctagccctcc cctcaaaaaa accaagacca acctcagcct agaaacctca tccccctaa	34680
ctgtgagcac ctcaggcgcc ctcaccgtag cagccgcgc tccctggcg gtggccggca	34740
cctccctcac catgcaatca gaggcccccc tgacagtaca ggatgcaaaa ctcaccctgg	34800
ccaccaaagg cccctgacc gtgtctgaag gcaaaactggc cttgcaaaaca tcggccccgc	34860
tgacggccgc tgacagcagc accctcacag tcaagtccac accacccctt agcacaagca	34920
atggcagctt gggatttgac atgcaagccc ccatttacac caccaatggaa aactaggac	34980
ttaactttgg cgctccctg catgtggtag acagcctaaa tgcactgact gtagttactg	35040
gccaagggtct tacgataaac ggaacagccc tacaaactag agtctcagggt gcccctcaact	35100
atgacacatc aggaaaccta gaattgagag ctgcaggggg tatgcgaggat gatgcaatg	35160
gtcaacttat cttgtatgtta gcttacccat ttgtatgcaca aaacaatctc agccttaggc	35220

[0133]

ttggacaggg acccctgttt gttaactctg cccacaactt ggatgttaac tacaacagag	35280
gcctctacct gttcacatct ggaaatacca aaaagctaga agttaatatc aaaacagcca	35340
agggtctcat ttatgtatgac actgctatag caatcaatgc gggtgatggg ctacagtttgc	35400
actcaggctc agatacaaataat ccattaaaaa ctaaacttgg attaggactg gattatgact	35460
ccagcagagc cataattgct aaactggaa ctggcctaag ctttgacaac acaggtgccat	35520
tcacagtagg caacaaaaat gatgacaagc ttaccttgc gaccacacca gaccatccc	35580
ctaaactgttag aatctattca gagaaagatg ctaaattcac acttgttttgc actaaatgcg	35640
gcagtcaggt gttggccagc gtttctgttt tatctgtaaa aggtgcctt gcccgcata	35700
gtggcacagt aactagtgc cagattgtcc tcagatttgc tgaaaatggc gttctactaa	35760
gcaattcttc cttgaccct caatactggc actacagaaa aggtgacctt acagagggca	35820
ctgcatatac caacgcagtggc ggatttatgc ccaacccatc acgataccca aaaacacaga	35880
ccaaactgc taaaagcaac attgttaatgc aggtttactt gaatggggac aaatccaaac	35940
ccatgaccct caccattacc ctcaatggc ctaatggaa aggagatgcc acagtaagca	36000
cttactccat gtcattctca tggaactggc atggaaatggc ttacattaaat gaaacgttcc	36060
aaaccaactc cttcacccct tcctacatgc cccagaataaaa aggtcatgc cgctgttgc	36120
ttgattcaat gtgtttctgt tttatccatc agcacaacaa aatcattcaat gtcattcttc	36180
catcttagct taatagacac agtagcttaa tagaccaggc agtgcaagc cccattctgc	36240
cttataacta gtggagaagt actcgctac atggggtagt agtcataatc gtgcattcagg	36300
atagggcggt ggtgctgcag cagcgccgca ataaactgc gcccggccg ctccgtcctg	36360
caggaataca acatggcagt ggtctctca gcgatgattc gcaccggccg cagcataagg	36420
cgccctgtcc tccggcaca gcagcgccacc ctgatctac ttaaaatcgc acagtaactg	36480
cagcacagca ccacaatatt gttcaaaatc ccacagtgc aggcgcgttgc tccaaagctc	36540
atggcggtt ccacagaacc cacgtggccatc tcataccaca agcgcaggta gattaatgg	36600
cgaccctca taaacacgcgtt ggacataaac attaccttt ttggcatgtt gtaattcacc	36660
accccccgtt accatataaa cctctgatta aacatggcgc catccaccac catctaaac	36720
cagctggccatc aaacctgccc gccggctata cactgcaggaa accggggact ggaacaatgc	36780
cagttggagag cccaggactc gtaaccatgg atcatcatgc tcgtcatgtt atcaatgttgc	36840
gcacaacaca ggcacacgtt catacacttc ctcaggatca caagctcctc ccgcgttgc	36900
accatatccc agggaaacaac ccattcctga atcagcgtaa atcccacact gcagggaaaga	36960

[0134]

cctcgacgt aactcacgtt gtgcattgtc aaagtttac attcggcag cagcgatga	37020
tcctccagta tggtagcgcg ggtttctgtc tcaaaaggag gtagacgatc cctactgtac	37080
ggagtgcgcc gagacaaccc agatcgtgtt ggtcgtagtg tcatgccaaa tggAACGCCG	37140
gacgttagtca tatttcctga agtcttagat ctctcaacgc agcaccagca ccaacacttc	37200
gcagtgtaaa aggccaagtg ccgagagagt atatataggg ataaaaaagtg acgtaaacgg	37260
gcaaagtcca aaaaacgccc agaaaaaccg cacgcgaacc tacgccccga aacgaaagcc	37320
aaaaaaacact agacactccc ttccggcgte aacttccgct ttcccacgct acgtcacttg	37380
ccccagtc aaacaaactaca tatcccgaac ttccaagtcg ccacgccccaa aacaccgcct	37440
acacctcccc gcccgcggc ccgcccccaa acccgctcc cgccccgcgc cccgccccgc	37500
gccgccccatc tcattatcat attggcttca atccaaaata aggtatatta ttgatgatg	37559

&lt;210&gt; 12

&lt;211&gt; 1109

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 12

ggagttccgc gttacataac ttacggtaaa tggccgcct ggctgaccgc ccaacgaccc	60
ccgccccattt acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca	120
ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagttac atcaagtgtt	180
tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcattt	240
tgcccagttac atgaccttat gggactttcc tacttggcag tacatctacg tattagttcat	300
cgcttattacc atggtcgagg tgagccccac gttctgttttc actctccccca tctcccccc	360
ctccccaccc ccaattttgt atttattttat ttttaattt ttttgtgcag cgatggggc	420
gggggggggg gggggcgccgc gccaggcggtt gcgccccgg gcgaggggcg gggcgccccg	480
aggcggagag gtgcggcgcc agccaatcag agcggcgccgc tccgaaagtt tcctttatgt	540
gcgaggcgcc ggcggcgccgc gcccataaaa aagcgaagcg ctccctatca gtgatagaga	600
tctccctatc agtgatagag atcgtcgacg agctcgcggc gggcgaggat cgctgcgcgc	660
tgccttcgccc ccgtcccccg ctccgcgc gctcgccgc gcccgcggc gctctgactg	720
accgcgttac taaaacaggt aagtccggcc tccgcggc gttttggcgc ctccgcggg	780

[0135]

cgccccccctc	ctcacggcga	gcgcgtgccac	gtcagacgaa	gggcgcagcg	agcgtcctga	840
tccttccgcc	cggaacgctca	ggacagcggc	ccgctgctca	taagactcgg	ccttagaacc	900
ccagtatcag	cagaaggaca	ttttaggacg	ggacttgggt	gactctaggg	cactggttt	960
ctttccagag	agcggAACAG	gCGAGGAAAA	gtagccctt	ctcggcgatt	ctgcggaggg	1020
atctccgtgg	ggcggtgaac	gccgatgatg	cctctactaa	ccatgttcat	gttttctttt	1080
tttttctaca	ggtcctgggt	gacgaacag				1109

&lt;210&gt; 13

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 13

atacggacta

gtggagaagt

actcgctac

atg

33

&lt;210&gt; 14

&lt;211&gt; 36

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 14

atacggaga

tctaagactt

caggaaatat

36

&lt;210&gt; 15

&lt;211&gt; 46

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 15

attcagtgt

a caggcgcgc

c

46

&lt;210&gt; 16

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; 人工序列

[0136]

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 16

actaggacta gttataagct agaatggggc tttgc

35

&lt;210&gt; 17

&lt;211&gt; 79

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 17

ttaatagaca cagtagctta atagacccag tagtgcaag ccccattctta gcttataacc

60

cctatttggtt tatttttct

79

&lt;210&gt; 18

&lt;211&gt; 93

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 18

atataatactc tctcggcaact tggcctttta cactgcgaag tgggggtgct ggtgctgcgt

60

tgagagatct ttattttgtta actgttaatt gtc

93

&lt;210&gt; 19

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 19

ttaatagaca cagtagctta ata

23

&lt;210&gt; 20

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; 人工序列

[0137]

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 20

ggaaggaggt gtctagtg

20

&lt;210&gt; 21

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 21

caatggcggt ggatagcggt ttgac

25

&lt;210&gt; 22

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 22

cagcatgcct gctattg

19

&lt;210&gt; 23

&lt;211&gt; 29

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 23

catctacgtt ttagtcatcg ctattacca

29

&lt;210&gt; 24

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成引物”

&lt;400&gt; 24

[0138]

gacttggaaa tccccgttag t 21

&lt;210&gt; 25

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成探针”

&lt;400&gt; 25

acatcaatgg gcgtggatag cggtt 25

&lt;210&gt; 26

&lt;211&gt; 592

&lt;212&gt; DNA

&lt;213&gt; 土拨鼠肝炎病毒

&lt;400&gt; 26

taatcaacct ctggattaca aaatttgc aagattgact ggtattctta actatgttgc 60

tcctttacg ctatgtggat acgctgcttt aatgccttg tatcatgcta ttgcttccg 120

tatggctttc attttctctt ccttgtataa atcctggttt ctgtctctt atgaggagtt 180

gtggcccggtt gtcaggcaac gtggcgtggt gtgcactgtt tttgctgacg caaccccac 240

tgggtggggc attgccacca cctgtcagct ccttccggg actttcgctt tccccctccc 300

tattgccacg gcggactca tcgcccctg ccttgcgcgc tgctggacag gggctcggct 360

gttggcact gacaattccg tgggtttgtc gggaaatca tcgtccttc cttggctgct 420

cgccctgtgtt gccacactgga ttctgcgcgg gacgtccttc tgctacgtcc cttcgccct 480

caatccagcg gaccttcctt cccgcggcct gctgccggct ctgcggcctc ttccgcgtct 540

tcgccttcgc cctcagacga gtggatctc ctttggcc gcctccccgc ct 592

&lt;210&gt; 27

&lt;211&gt; 543

&lt;212&gt; PRT

&lt;213&gt; 黑猩猩腺病毒

&lt;400&gt; 27

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro  
1 5 10 15Tyr 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 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			
[0140] 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

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser  
[0141] 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  
 225 230 235 240

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

[0142] 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 Ser Ser Asn Thr Thr Lys Leu Glu Thr Lys  
 305 310 315 320

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

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 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 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 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 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 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 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 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  
 [0151] 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 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 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
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
Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala		
[0158]	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
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

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro  
[0159] 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 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		
[0160] 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

&lt;210&gt; 36

&lt;211&gt; 577

&lt;212&gt; PRT

&lt;213&gt; 黑猩猩腺病毒

&lt;400&gt; 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
								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			

[0162]

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			

&lt;210&gt; 38

&lt;211&gt; 2994

&lt;212&gt; DNA

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 38

atggatgccg	acaagattgt	attcaaagtc	aataatcagg	tggctcttt	gaagccttag	60
attatcgtag	atcagttatga	gtacaagttac	cccgctatca	aagattttaga	gaagccctgt	120
ataaccctag	ggaaagcccc	cgacttaaac	aaagcataca	agtcagtttt	atcaggcatg	180
aatgccgcta	aacttgatcc	tgtatgtat	tgttcctact	tggcagcagc	aatgcagttc	240
tttgaggggg	cttgcgggaa	agactggact	agctatggaa	tcctgattgc	acgaaaagga	300
gacaagatca	ccccagattc	tctggggaa	ataaaacgta	ctgtatgtaga	aggaaattgg	360
gcttgacag	gaggcatgg	attgacaagg	gacccactg	tctctgagca	tgcatctta	420
gtcggtcttc	tcttgagtct	gtataggttg	agcaaaatat	caggacaaaa	caccggtaac	480
tataaaacaa	acattgcaga	taggatagag	cagatttcg	agacagcccc	ttttgttaaa	540
atcggttggac	accacactct	aatgacaact	cacaaaatgt	gtgctaattt	gagttactata	600
ccgaatttca	gattttggc	cggaacctac	gacatgttt	tctccggat	tgagcatcta	660
tattcagcaa	ttagggtgg	cacagttgtt	actgctttag	aagactgttc	agggctggta	720
tcgttactg	ggttcataaa	gcagatcaat	ctcaccgcaa	gagaagcaat	tttatacttc	780
ttccacaaga	actttgagga	agagataaga	agaatgttt	agccagggca	ggaaaccgct	840
gttcctact	cttatttcat	tcacttccgc	tcatttaggt	tgagtggaa	gttccttat	900
tcatcgaatg	ccgttggtca	tgtgttcaat	ctcatttact	ttgttggatg	ctatatgggt	960
caagtcagat	ccctaaatgc	aacagttatt	gctgcatttg	ctcctcatga	gatgtctgtt	1020
ctaggggct	atctgggaga	ggaatttttc	gggaaaggaa	cattgaaag	aagattcttt	1080
agagatgaga	aagaacttca	agaatacgg	gccccctgaac	tgacaaagac	tgacgtggcg	1140
ttggcagatg	atgaaacgg	taactctgtat	gatgaggact	acttctccgg	tgaaaccaga	1200
agtccagaag	ccgtttatac	tcaaatcatg	atgaatggag	gtcgactaaa	gagatcgat	1260
atacgaagat	atgtctcagt	cagttccaaat	catcaagccc	gtccaaactc	attcgccgag	1320

[0168]

tttctaaaca aaacataactc aagtgactca gcgcagtaa agcagacatt aaactttgat	1380
ttgctgaaac ttgcaggtga tgttaggtca aatccaggc caatggttcc tcaggcttt	1440
ttgttgtgc ccctctgg ttttcattt tgttcggga aattccctat ttacacgata	1500
ccagacaaac ttggccctg gagcccgatt gacatccatc acctcagctg tccaaacaat	1560
ttggttgtgg aggacgaagg atgcaccaac ctgtcagggt tctctacat ggaacttaaa	1620
gtcgggtaca tctcagccat aaaagtgaac gggttcacat gcacggcggt tgtgacagag	1680
gcagaaacct acactaactt ttttggttat gtcaccacca catttaaaag aaagcatttc	1740
cgcccaacac cagatgcatg tagagccgctg tacaatttggaa agatggccgg tgaccccagg	1800
tatgaggagt ctctacacaa cccgtaccca gactatcatt ggcttcgaac cgtaaaacc	1860
accaaggagt ctctcggttat catatccccaa agtgtggcag atttggaccc atatgacaag	1920
tcccttcact cgagggctt cccgagcgga aagtgtcag gaataacggt gtcttctaca	1980
tactgttcga ctaaccacga ttacaccatt tggatgccag agaatccgag actaggacaa	2040
tcttgtgaca tctttaccaa tagtagaggg aagagagcat ccaaagggag taagacctgc	2100
ggctttgttag atgaaagagg cctgtataag tctctaaaag gagcatgcaa actcaagtt	2160
tgtggagttc ttggacttag actcatggat ggaacgtggg ttgcgtatgca gacatcggt	2220
gagaccaaat ggtgccctcc cgatcagctg gtgaatttgc acgacttccg ctcggatgaa	2280
attgagcatc tcgtcgtaga ggagctggc aagaagagag aggagtgtct ggatgcacta	2340
gagtccatca tgaccaccaa gtcggtgagt ttcagacgtc tcagtcattt aagaaagctt	2400
gtccctgggt ttggaaaagc atacactata tttaacaaga cttgtatgga ggctgtatgct	2460
cattacaaat cagtccggac ttggaatgag atcatcccct caaaggggtg tttgagagta	2520
ggagggaggt gtcatccgca tgtgaacggg gtgttttca atggtataat attagggcct	2580
gacggccatg ttctaatccc agagatgcaa tcatccctcc tccagcaaca tatggagtt	2640
ttggagtctt cagttatacc cctgatgcac cccctggcag acccggtccac agtttcaag	2700
gacggtgatg aggctgagga ttttggaa gttcaccttc ccgatgtgca caaacaggc	2760
tccggggttt acctgggtct cccaaactgg gggaaatacg tattactgag tgcagggct	2820
ctgattgctc tcatgttgc aattttcctc atgaccctgtt gtatgttgc aatcgacca	2880
aatctacgc aacgcagtct tggagggaca gggaggaagg tgtcagttac ttccaaagc	2940
gggaaagtca tatcttcatg ggagtcatat aaaagtgggg gtgagactag actg	2994

&lt;210&gt; 39

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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			

Thr	Gly	Val	Val	Thr	Glu	Ala	Glu	Thr	Tyr	Thr	Asn	Phe	Val	Gly	Tyr
				85				90			95				

[0169]

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

&lt;210&gt; 40

&lt;211&gt; 1572

&lt;212&gt; DNA

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 40

atgggcctc aagccctgct gttcgtgccc ctgctggctct tctccctctg ctggcaag	60
ttccccatct acaccatccc tgacaagctc ggccctgtgt ccccccattga catacatcac	120
ctcagctgcc ccaacaacct ggtggtgag gatgagggct gcacaaacct gagcggcttc	180
tcctatatgg aactcaaggt gggctatatac tccgccatca aggtcaatgg attcacatgc	240
accggcgtcg tgacagaggc taaaacatac accaactttg tggctacgt caccaccaca	300
ttcaagagga agcacttcag gcccacccct gacgcttgca gggctgccta caattggaag	360
atggctggcg accccaggta tgaggagtcc ctgcacaatc cttaccccgat cttaccattgg	420
ctcaggacag tcaagaccac caaggagtcc ctggtcatta tctcccttag cgtggccgac	480
ctagaccgt atgacaaaag cctgcactcc agggcttcc ctagcggcaa atgctccggc	540
attacagtga gctccaccta ctgcagcaca aaccacgact acaccatctg gatgcctgag	600
aatcctaggc tcggcacctc ctgtgacata ttacaataa gcagggccaa gagggcttcc	660
aaaggcagca aaacctgcgg ctttgcgac gaaagaggcc tgtacaagtc cctcaaggc	720
gcttgtaaac tcaagctgtg cggagtgctg ggactcagac tcatggacgg cacatgggtg	780

[0172]

gccatgcaga ccagcgatga gaccaagtgg tgccccccg atcagctggt gaatctgcac	840
gacttcaggt ccgacgaaat tgagcacctc gtggtcgagg agctggtgaa gaagagagaa	900
gagtgccctgg atgctctgga gtccatcatg accaccaaatt ccgtgtcctt cagaaggctg	960
agccacaccta ggaagctggt cccggcttt ggcaaggcct acacaatttt caataagaca	1020
ctgatggagg ccgatgctca ctacaaatcc gtgaggacct ggaacgagat catcccctcc	1080
aaaggctgcc tgagggtggg aggaagatgc caccccccacg tcaacggcgt cttttcaac	1140
ggcattatcc tcggaccga tggccatgtc ctgatccctg aaatgcaaag ctccctgctg	1200
cagcagcaca tggaactcct ggagagctcc gtcatcccc tcatgcaccc tctcgctgac	1260
cccagcaccg tggtaaaga cggcgatgag gccgaggact tcgtggaagt gcatctgcct	1320
gatgtgcata agcaagtcag cggcgatcgat ctgggcctgc ctaattgggg caagtatgtc	1380
ctgctctccg ccggagctct gattgccctg atgctgatca tttcctgat gacctgctgc	1440
agaagagtca acagacctga gagcacccaa agatccctcg gcggAACCGG aaggaaggc	1500
agcgtgacca gccagtcgg caaagtgatt tcctcctggg agagctataa aagcggcgg	1560
gagaccaggc tg	1572

&lt;210&gt; 41

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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 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]

&lt;400&gt; 42

atggttcctc aggttcttt gtttgtaccc ctccctgggtt tttcattgtg tttcgggaag	60
ttccccattt acacgatacc agacaaactt ggtccctgga gccctattga catacaccat	120
ctcagctgtc caaataacct ggttgtggag gacgaaggat gtaccaacct gtccgagttc	180
tcttacatgg aacttaagt gggatacatc tcagccataa aagtgaacgg gttcacttgc	240
acaggtgtt gtagcagaggc agaaacctac accaactttg ttggttatgt cacaaccaca	300
ttcaagagaa agcatttccg ccccacccca gacgcattgt aagccgcgtt taactggaag	360
atggccgggtg accccagata tgaagagtct ctacacaatc cgtaccccgat ctaccattgg	420
cttcgaactg taaaaaccac caaagagtct ctcgttatca tatcccaag tgtgacagat	480
ttggaccat atgacaaatc ctttcactca agggtcttcc ctggcggaaa ttgctcagga	540
ataacgggtt cctcgaccta ctgctcaact aatcatgatt acaccattt gatgcctgag	600
aatctgagac tagggacatc ttgtgacatt tttaccaata gcagagggaa gagggcatcc	660
aaaggaggca agacttgcgg ctttggat gaaagaggcc tgtataagtc tctaaaggga	720
gcatgcaaacc tcaagttatg tggagttctc ggacttagac ttatggatgg aacatgggtc	780
gcgatgcaaa catcagatga gaccaatgg tgccctccag gtcagttggt gaatttgcac	840
gactttcgct cagacgagat tgagcatctc gttgtggaag agttagtcaa gaaaagagag	900
gagtgtctgg atgcactaga gtcccatcatg accaccaagt cagttagttt cagacgtctc	960
agtcaatttga gaaaacttgc ccctgggtt ggaaaagcat ataccatatt caacaaaacc	1020
ttgatggagg ctgatgctca ctacaagtct gtccggaccc ggaatgagat catccctca	1080
aaagggtgtt tgagagttgg ggggaggtgt catccccatg tgaacgggtt gttttcaat	1140
ggtataatat tagggcttga cggccatgtt ctaatccag agatgcagtc atccctcctc	1200
cagcaacata tggagttgtt ggaatcttca gttatcccc tcatgcaccc cttggcagac	1260
ccttctacag ttttcaaaga cgggtatgag gttgaggatt ttgttgaagt tcacccccc	1320
gatgtgcattt aacaggtctc aggagttgaa ctgggtctcc cgaactgggg gaagtatgt	1380
ttgatgattt cagggcctt gattgccctg atgttataa tttcctgtat gacatgttgc	1440
agaagagtca atcgaccaga atctacgcaa agcagtctt gagagacagg gagaatgtg	1500
tcagtcactt cccaaagcgg aaaagtctata tcttcattgg agtcatataa gagtggaggc	1560
gagaccagac tg	1572

&lt;210&gt; 43

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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				

Thr	Gly	Val	Val	Thr	Glu	Ala	Glu	Thr	Tyr	Thr	Asn	Phe	Val	Gly	Tyr
				85				90			95				

[0176]

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

&lt;210&gt; 44

&lt;211&gt; 1572

&lt;212&gt; DNA

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 44

atggtcctc	aggtgctcct	gttcgtccct	ctgctcggt	tctccctctg	ctttggcaag	60
ttccccatct	acacaatccc	cgataagctc	ggcccttgga	gccctattga	catccaccac	120
ctctcctgtc	ccaacaacct	ggtggtgag	gacgagggat	gcaccaacct	gagcgagttc	180
tcatacatgg	agctgaaggt	gggctatatt	agcgccatca	aggtaaacgg	cttcacatgc	240
accggagtcg	tgaccgaggc	cgagacctac	acaaacttcg	tcggctacgt	cacaacaacc	300
ttcaagagga	aacatttcag	accacccct	gacgcttgca	ggcccgctta	caattggaag	360
atggctggcg	accccagata	cgaggagagc	ctccacaacc	cctaccctga	ctaccactgg	420
ctcaggaccg	tgaagaccac	caaggagtc	ctcggtatca	tctccccag	cgtcacagac	480
ctcgaccctt	atgataagag	cctccactcc	agggtgttcc	ctggcggcaa	ctgttccggc	540
atcaccgtct	cctccaccta	ctgcagcacc	aaccacgact	acaccatctg	gatgccttag	600
aacctgaggc	tggcaccag	ctgcgacatc	ttcaccata	gcaggggcaa	gagggcctcc	660
aaggaggaa	agacctgcgg	atttgtggat	gagaggggcc	tctacaagtc	actgaaggc	720
gcctgcaagc	tgaaaactctg	cggcgtgctg	ggactgaggc	tcatggacgg	aacctgggtc	780
gctatgcaaa	catccgacga	gaccaagtgg	tgtccccccg	gccagctcgt	aatcttcat	840

[0179]

gacttcagga	gcgacgaaat	cgagcacctc	gtggtgagg	aactggtaa	gaagagggag	900
gagtgcctcg	acgctctcga	gtccatcatg	accaccaaga	gcgtgtcatt	tagaagactg	960
agccaccta	ggaagctgg	ccccggcttc	ggcaaaggct	acaccatctt	caacaagacc	1020
ctgatggagg	ccgatgctca	ctacaagagc	gtcaggacct	ggaacgagat	catccccagc	1080
aaaggctgcc	tgagagtgg	aggaaggtgt	caccccaacg	tgaacggcgt	cttcttcaac	1140
ggcatcattc	tcggaagcga	cggacacg	ctgattcccg	agatgcagag	ctcactc	1200
cagcagcata	tggagctcct	gaaagctcc	gtcattcctc	tgatgcattc	cctcgctgat	1260
ccctccacag	tcttcaaaga	tggcgacgag	gtggaggact	ttgtggaagt	gcaccccc	1320
gatgttcatg	agcaagtctc	cgagtgaa	ctggcctcc	ccaactgggg	caagtacg	1380
ctcatgattt	ctggcgctc	catgc	atgctgatca	tctcctgat	gac	1440
agaagggtca	atagacccga	gagcactc	tccagc	g	caag	1500
agcgtgac	ccaaatccgg	caaagtcatc	agtc	c	atccggag	1560
gaaacaaggc	tg					1572

&lt;210&gt; 45

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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 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]

&lt;400&gt; 46

atggtgcccc aggtgctgct ctttgtcccc ctgctggct tcagcctctg cttcgccaag	60
ttccccatct acaccatccc agacaagctg gggccttggaa gccccatcga catccaccac	120
ctgagctgcc ccaacaacctt ggtgggtggaa gatgaaggct gcaccaacctt gagcgagttc	180
tccttacatgg agctgaaggt gggctacatc tctgccatca aggtgaatgg cttcacctgc	240
actggagtggt tcacagaggc cgagacctac accaactttg tggctatgt gaccaccacc	300
ttcaagagga agcacttccg gcccacccca gatgcctgcc gggccgccta caactggaag	360
atggctgggg acccccgcta tgaggagagc ctgcacaacc cctaccaggatctaccactgg	420
ctgaggacag tgaagaccac caaggagagc ctgggtatca tcagccccag cgtgacagac	480
ctggaccctt atgacaagag cctgcacagc cgggtgttcc ctggcggcaa ctgcagcggc	540
atcaccgtga gcagcaccta ctgcagcacc aaccacgact acaccatctg gatgccagaa	600
aacctgcggc tgggcaccag ctgtgacatc ttaccaaca gccggggcaa gagggccagc	660
aagggcggca agacctgtgg ctttgtggat gagcggggcc tctacaagag cctgaagggg	720
gcctgcaagc tgaagctctg tgggtgtctg ggcctgcggc tcatggatgg cacctgggtg	780
gccatgcaga cctcagatga gaccaagtgg tgcccccag gccagctggt gaacctgcat	840
gacttccgca gcgacgagat tgagcacctg gtggtgagg agctggtaa gaagagggag	900
gagtgccctgg atgcctcga gacatcatg accaccaaga gcgtgtcctt ccgcgcctg	960
agccacactgc ggaagctggt gcctggctt ggaaaggcct acaccatctt caacaagacc	1020
ctgatggagg cagatccccatcataagagt gtgcggacct ggaatgagat catccccagc	1080
aagggctgcc tgcgggtggg cggccgggtgc caccccccacg tgaatggagt gttttcaat	1140
ggcatcatcc tggcagcga cggccacgtg ctgatccctg agatgcagag cagcctgctg	1200
cagcagcaca tggagctgct ggagagctct gtcatccccatcataagagt gtgcggacct ggaatgagat catccccagc	1260
cccagcacccg tggtaagga tggagatgaa gtggaggact tcgtggaggt gcacctgcca	1320
gatgtgcacg agcaggtgtc tggcgtggag ctggcctgc ccaactgggg caagtacgt	1380
ctgatgattt ctggccctt gatgcctgc atgctgatca tcttcctgat gacctgctgc	1440
cggcgggtga acagacccatcataagagt gtgcggacct ggaatgagat catccccagc	1500
tctgtcacca gccagagcgg caaggtgatc agcagctggg agagctacaa gagtggcggc	1560
gagaccaggc tg	1572

&lt;210&gt; 47

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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 460Gly Ala Leu Thr Ala Leu Met Leu Ile Ile Phe Leu Met Thr Cys Cys  
465 470 475 480Arg Arg Val Asn Arg Ser Glu Pro Thr Gln Leu Asn Leu Arg Gly Thr  
485 490 495Gly Arg Glu Val Ser Val Thr Pro Gln Ser Gly Lys Ile Ile Ser Ser  
500 505 510Trp Glu Ser His Lys Ser Gly Gly Glu Thr Arg Leu  
515 520

&lt;210&gt; 48

&lt;211&gt; 1572

&lt;212&gt; DNA

&lt;213&gt; 狂犬病病毒

<400> 48	60
atggttcctc aggctctcct gtttgtaccc cttctggttt ttccattgtg ttttggaaa	60
ttccctattt acacgatacc agacaagctt ggtccctgga gcccgattga catacatcac	120
ctcagctgcc caaacaattt gatagtggag gacgaaggat gcaccaacct gtcagggttc	180
tcctacatgg aacttaaagt tggatacatc tttagccataa aaatgaacgg gttcacttgc	240
acaggcgttg tgacggaggc tgaaacctac actaacttcg ttggttatgt cacaaccacg	300
tcaaaaagaaa gcatttccgc ccaacaccag atgcatgtag agccgcgtac aactggaaga	360
tggccggta ccccagatat gaagagtctc tacacaatcc gtaccctgac taccgctggc	420
ttcgaactgt aaaaaccacc aaggagtctc tcgttatcat atctccaagt gtggcagatt	480
tggaccata tgacagatcc cttcactcga gggcttccc tagcgggaag tgctcaggag	540
tagcggtgtc ttctacctac tgctccacta accacgatta caccatttg atgcccgaga	600
atccgagact agggatgtct tgtgacattt ttaccaatag tagagggaag agagtatcca	660
aagggagtga gacttgcggc ttttagatg aaagaggcct atataagtct ttaaaaggag	720
catgcaaact caagttatgt ggagttctag gacttagact tatggatgga acatgggtcg	780
cgtatgcaaac atcaaatgaa accaaatggt gccctcccgaa tcagttggtg aacctgcacg	840

[0186]

actttcgctc agacgaaatt gagcaccttg ttgttagagga gttggtcagg aagagagagg	900
agtgtctgga tgcactagag tccatcatac aaccaagtca gtgagttca gacgtctcag	960
tcatttaaga aaacttgtcc ctgggttgg aaaagcatat accatattca acaagacctt	1020
gatggaagcc gatgctcaact acaagtcagt cagaacttgg aatgagatcc tcccttcaaa	1080
agggtgttta agagttgggg ggaggtgtca tcctcatgtg aacgggtgt ttttcaatgg	1140
tataatatta ggacctgacg gcaatgtctt aatcccagag atgcaatcat ccctcctcca	1200
gcaacatatg gagttgttgg aatcctcggt tatccccctt gtgcacccccc tggcagaccc	1260
gtctaccgtt ttcaaggacg gtgacgaggc tgaggatttt gttgaagttc accttcccga	1320
tgtgcacaat caggtctcag gagttgactt gggtctcccg aactggggga agtatgtatt	1380
actgagtgca gggccctga ctgccttgat gttgataatt ttcctgatga catgtttag	1440
aagagtcaat cgatcagaac ctacgcaact caatctcaga gggacaggga gggaggtgtc	1500
agcactcccc aaagcgggaa gatcatatct tcatggaat cacacaagag tgggggtgag	1560
accagactgt ga	1572

&lt;210&gt; 49

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成肽”

&lt;220&gt;

&lt;221&gt; MISC\_FEATURE

&lt;222&gt; (1)..(6)

&lt;223&gt; /注释=“该序列可以包含2-6个残基”

&lt;400&gt; 49

His His His His His His

1 5

&lt;210&gt; 50

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 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

<210> 57  
<211> 9  
<212> PRT  
<213> Irkut狂犬病病毒

<400> 57  
Gly Cys Thr Thr Leu Thr Ala Phe Asn  
1 5

<210> 58  
<211> 9  
<212> PRT  
<213> Khu.jand狂犬病病毒

<400> 58  
Gly Cys Thr Thr Leu Ser Gly Phe Thr  
1 5

<210> 59  
<211> 9  
<212> PRT  
<213> Lagos蝙蝠狂犬病病毒

[0188] <400> 59  
Gly Cys Ser Asp Thr Ala Thr Phe Ser  
1 5

<210> 60  
<211> 9  
<212> PRT  
<213> 莫科拉狂犬病病毒

<400> 60  
Gly Cys Asn Thr Glu Ser Pro Phe Thr  
1 5

<210> 61  
<211> 9  
<212> PRT  
<213> Shimoni蝙蝠狂犬病病毒

<400> 61  
Gly Cys Ser Ser Ser Ser Thr Phe Ser  
1 5

<210> 62  
<211> 9  
<212> PRT  
<213> 西高加索蝙蝠狂犬病病毒

<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]

&lt;210&gt; 91

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Ikoma狂犬病病毒

&lt;400&gt; 91

Lys Ser Val Asp Asn Trp Thr Asp Ile  
1 5

&lt;210&gt; 92

&lt;211&gt; 120

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 92

ggaaggtcag cgtgaccaggc cagtccggca aagtgatttc ctcctgggag agctataaaa 60

gcggcggaga gaccaggctg tcatgagcgg ccgcgatctg taatcaacct ctggattaca 120

&lt;210&gt; 93

&lt;211&gt; 120

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;221&gt; 来源

&lt;223&gt; /注释=“人工序列的描述：合成多核苷酸”

&lt;400&gt; 93

atggctccgg cggctctgc aacacaata aagagaccct aagacccca acttatatat 60

tttcatgacc accccaggcc acgcccactc acccacctca ccatagagcc caccgcatcc 120

&lt;210&gt; 94

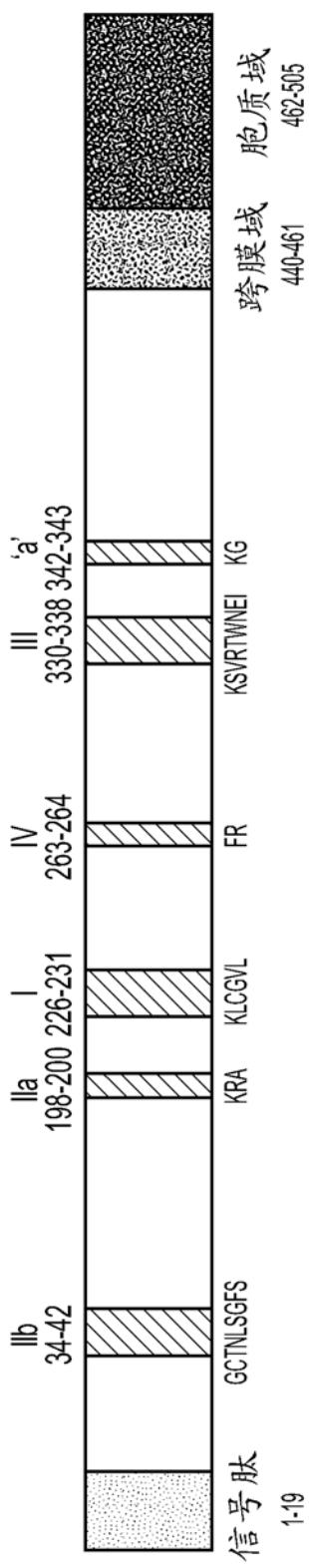
&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; 狂犬病病毒

&lt;400&gt; 94

Gly Cys Thr Asn Leu Ser Gly Phe Ser  
1 5



merozoite surface protein 1 (MSP1) domain I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 (IIA) 和198-200 (IIB) 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 <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDIT <sub>T</sub> <u>T</u> ASPP <sub>L</sub> KKTKTNLISLQ
PanAd3	(1)	MKR <sub>A</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLR <sub>E</sub> SEPLVSHGMLALKMGNGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd17	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDIT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd19	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd24	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDIT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd155	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDIT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd11	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd20	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGSGLSLDAGNLTSDQDIT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd31	(1)	MKR <sub>T</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLAEP <sub>P</sub> VSHGMLALKMGNGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
PanAd1	(1)	MKR <sub>A</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLR <sub>E</sub> SEPLVSHGMLALKMGNGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
PanAd2	(1)	MKR <sub>A</sub> KTSDE <sub>E</sub> ENPVVPPYDTE SGPPSVFVSPFLTPPFVSPDGFQESP <sub>P</sub> GVLSLNLR <sub>E</sub> SEPLVSHGMLALKMGNGLSLDAGNLTSDQDVT <sub>T</sub> <u>T</u> TP <sub>L</sub> KKTKTNLISLQ
ChAd3	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LTIVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
PanAd3	(101)	TSA <sub>P</sub> LTIVS-SGS <sub>L</sub> TVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> GLAT <sub>Q</sub> GPLTVSEGKLTQTSAPLTAADSSTLT <sub>V</sub> G <sub>T</sub> TP <sub>P</sub> ISVS <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd17	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LTIVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd19	(101)	TSA <sub>P</sub> LTIVSTSG <sub>A</sub> LT <sub>L</sub> AAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd24	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LT <sub>L</sub> AAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd155	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LTIVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd11	(101)	TSA <sub>P</sub> LTIVSTSG <sub>A</sub> LT <sub>L</sub> AAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> GLAT <sub>Q</sub> GPLTVSEGKLTQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd20	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LTIVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
ChAd31	(101)	TSS <sub>E</sub> PLIVSTSG <sub>A</sub> LTIVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> LAT <sub>K</sub> GPLTVSEGKLAQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
PanAd1	(101)	TSA <sub>P</sub> LTIVS-SGS <sub>L</sub> TVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> GLAT <sub>Q</sub> GPLTVSEGKLTQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH
PanAd2	(101)	TSA <sub>P</sub> LTIVS-SGS <sub>L</sub> TVAAAAPLAVAGTSLTMQSEAP <sub>P</sub> LTQDAKLT <sub>E</sub> GLAT <sub>Q</sub> GPLTVSEGKLTQTSAPLTAADSSTLT <sub>V</sub> SAT <sub>P</sub> PL <sub>N</sub> V <sub>S</sub> SGS <sub>L</sub> GLDMD <sub>E</sub> DPWYTH

图 2A

图 2B

ChAd3	(365)	TPDPSPNCR <sub>1</sub> <u>HS</u> DKDCK <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> AVSGN <sub>1</sub> LASI <sub>1</sub> GT <sub>1</sub> VASV <sub>1</sub> FTLRF <sub>1</sub> QGV <sub>1</sub> MEN <sub>1</sub> SSLD <sub>1</sub> DRQ <sub>1</sub> YNNER <sub>1</sub> NG <sub>1</sub> MS <sub>1</sub> TA <sub>1</sub> AP <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
PanAd3	(364)	TPDPSPNCR <sub>1</sub> <u>HA</u> DKDCK <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> AVSGN <sub>1</sub> SSMT <sub>1</sub> GT <sub>1</sub> VSSV <sub>1</sub> FTLRF <sub>1</sub> QGV <sub>1</sub> MEN <sub>1</sub> SSLD <sub>1</sub> KEY <sub>1</sub> YNNER <sub>1</sub> NG <sub>1</sub> MS <sub>1</sub> TA <sub>1</sub> AP <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> TP <sub>1</sub> PN <sub>1</sub> NA <sub>1</sub>
ChAd17	(365)	TPDPSPNCR <sub>1</sub> <u>HS</u> DKDCK <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> AVSGN <sub>1</sub> LASI <sub>1</sub> GT <sub>1</sub> VASV <sub>1</sub> FTLRF <sub>1</sub> QGV <sub>1</sub> MEN <sub>1</sub> SSLD <sub>1</sub> KQ <sub>1</sub> YNNER <sub>1</sub> NG <sub>1</sub> MS <sub>1</sub> TA <sub>1</sub> AP <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd19	(365)	TPDPSPNCR <sub>1</sub> <u>HS</u> DKDCK <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> AVSGN <sub>1</sub> LASI <sub>1</sub> GT <sub>1</sub> VSSV <sub>1</sub> FTLRF <sub>1</sub> QGV <sub>1</sub> MEN <sub>1</sub> SSLD <sub>1</sub> KQ <sub>1</sub> YNNER <sub>1</sub> NG <sub>1</sub> MS <sub>1</sub> TA <sub>1</sub> AP <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd24	(365)	TPDPSPNCR <sub>1</sub> <u>HS</u> DKDCK <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> AVSGN <sub>1</sub> LASI <sub>1</sub> GT <sub>1</sub> VSSV <sub>1</sub> FTLRF <sub>1</sub> QGV <sub>1</sub> MEN <sub>1</sub> SSLD <sub>1</sub> KQ <sub>1</sub> YNNER <sub>1</sub> NG <sub>1</sub> MS <sub>1</sub> TA <sub>1</sub> AP <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd155	(401)	TPDPSPNCR <sub>1</sub> <u>SE</u> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd11	(401)	TPDPSPNCR <sub>1</sub> <u>SE</u> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd20	(401)	TPDPSPNCR <sub>1</sub> <u>SE</u> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd31	(401)	TPDPSPNCR <sub>1</sub> <u>SE</u> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
PanAd1	(400)	TPDPSPNCR <sub>1</sub> <u>IN</u> SE <sub>1</sub> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
PanAd2	(400)	TPDPSPNCR <sub>1</sub> <u>IN</u> SE <sub>1</sub> DKA <sub>1</sub> <u>FT</u> LVLTKCGSQ <sub>1</sub> LASV <sub>1</sub> SV <sub>1</sub> SV <sub>1</sub> VKG <sub>1</sub> SLA <sub>1</sub> PIS <sub>1</sub> GT <sub>1</sub> V <sub>1</sub> TS <sub>1</sub> AQ <sub>1</sub> IV <sub>1</sub> LR <sub>1</sub> DENG <sub>1</sub> V <sub>1</sub> U <sub>1</sub> SSLD <sub>1</sub> DP <sub>1</sub> Q <sub>1</sub> Y <sub>1</sub> NR <sub>1</sub> KG <sub>1</sub> D <sub>1</sub> LT <sub>1</sub> E <sub>1</sub> G <sub>1</sub> T <sub>1</sub> A <sub>1</sub> Y <sub>1</sub> TA <sub>1</sub> VG <sub>1</sub> MP <sub>1</sub> NA <sub>1</sub>
ChAd3	(465)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
PanAd3	(464)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> D <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd17	(465)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd19	(465)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd24	(465)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd155	(501)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd11	(501)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> D <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd20	(501)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
ChAd31	(501)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> D <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
PanAd1	(500)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> A <sub>1</sub> E <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ
PanAd2	(500)	YPKIQSQ <sub>1</sub> TA <sub>1</sub> K <sub>1</sub> IV <sub>1</sub> SV <sub>1</sub> Y <sub>1</sub> NG <sub>1</sub> DK <sub>1</sub> SK <sub>1</sub> PM <sub>1</sub> LT <sub>1</sub> LN <sub>1</sub> GT <sub>1</sub> NE <sub>1</sub> SET <sub>1</sub> Q <sub>1</sub> SH <sub>1</sub> YS <sub>1</sub> MS <sub>1</sub> SET <sub>1</sub> W <sub>1</sub> D <sub>1</sub> SG <sub>1</sub> Y <sub>1</sub> ATE <sub>1</sub> AT <sub>1</sub> NS <sub>1</sub> FT <sub>1</sub> SY <sub>1</sub> IA <sub>1</sub> EQ

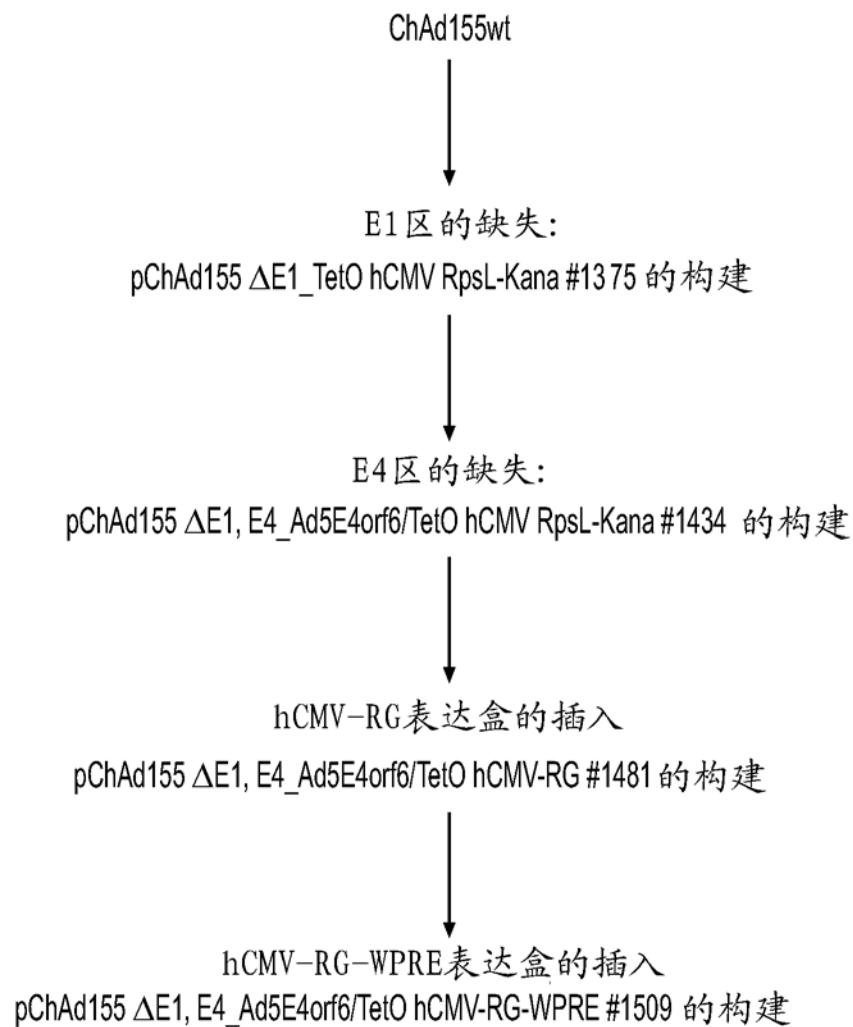


图 3

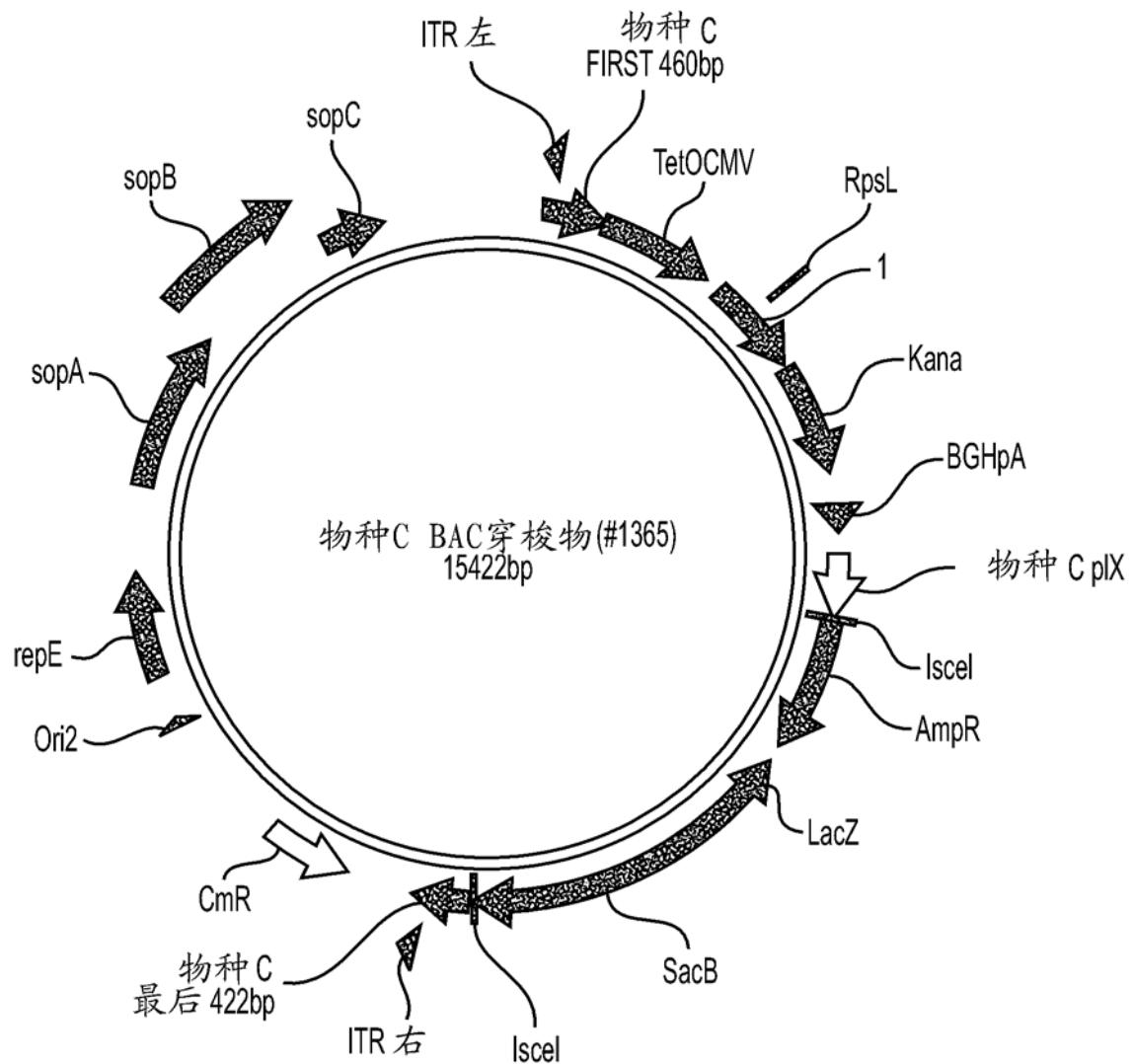


图 4

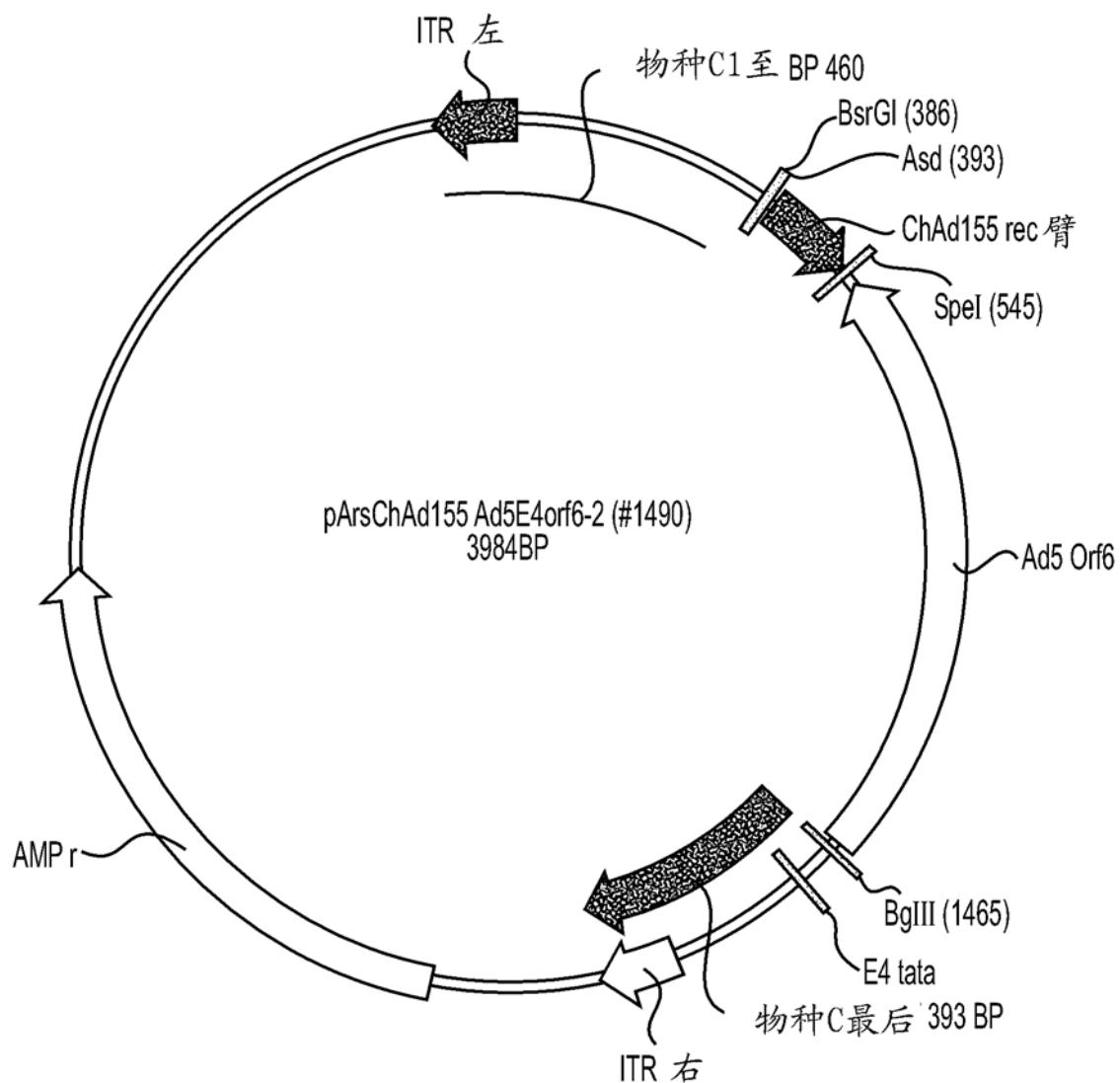
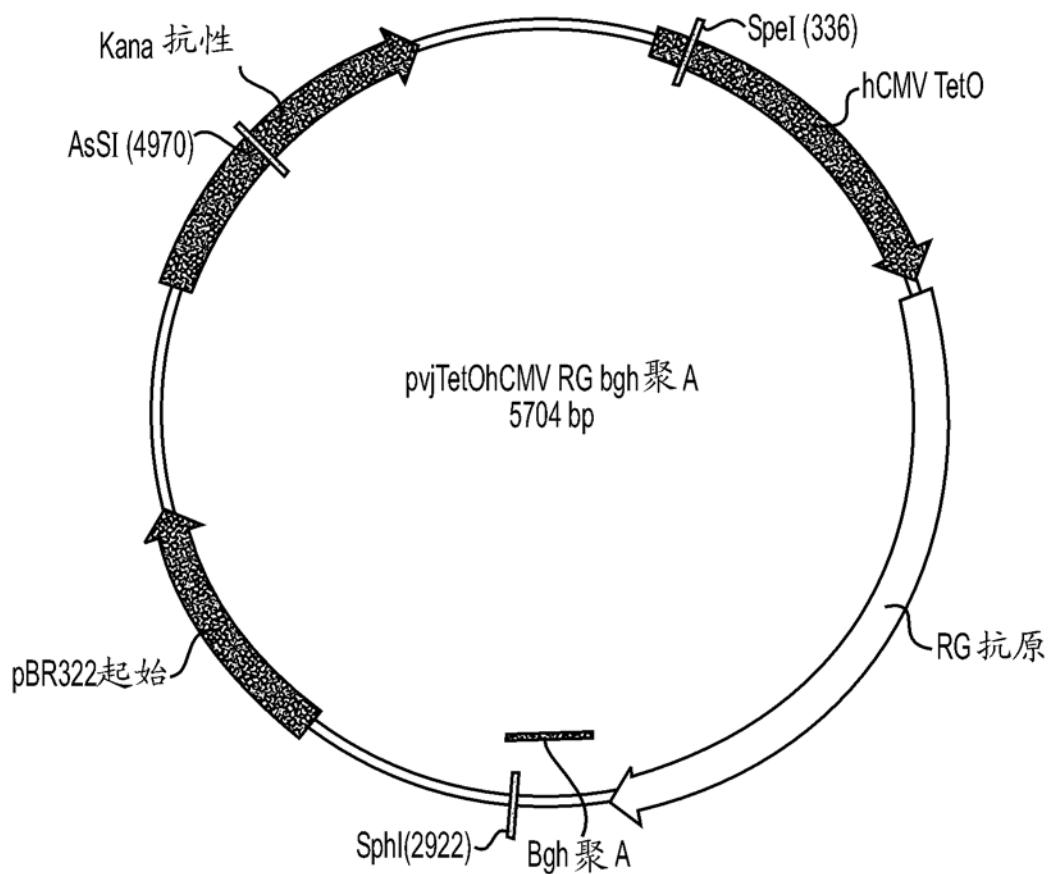


图 5



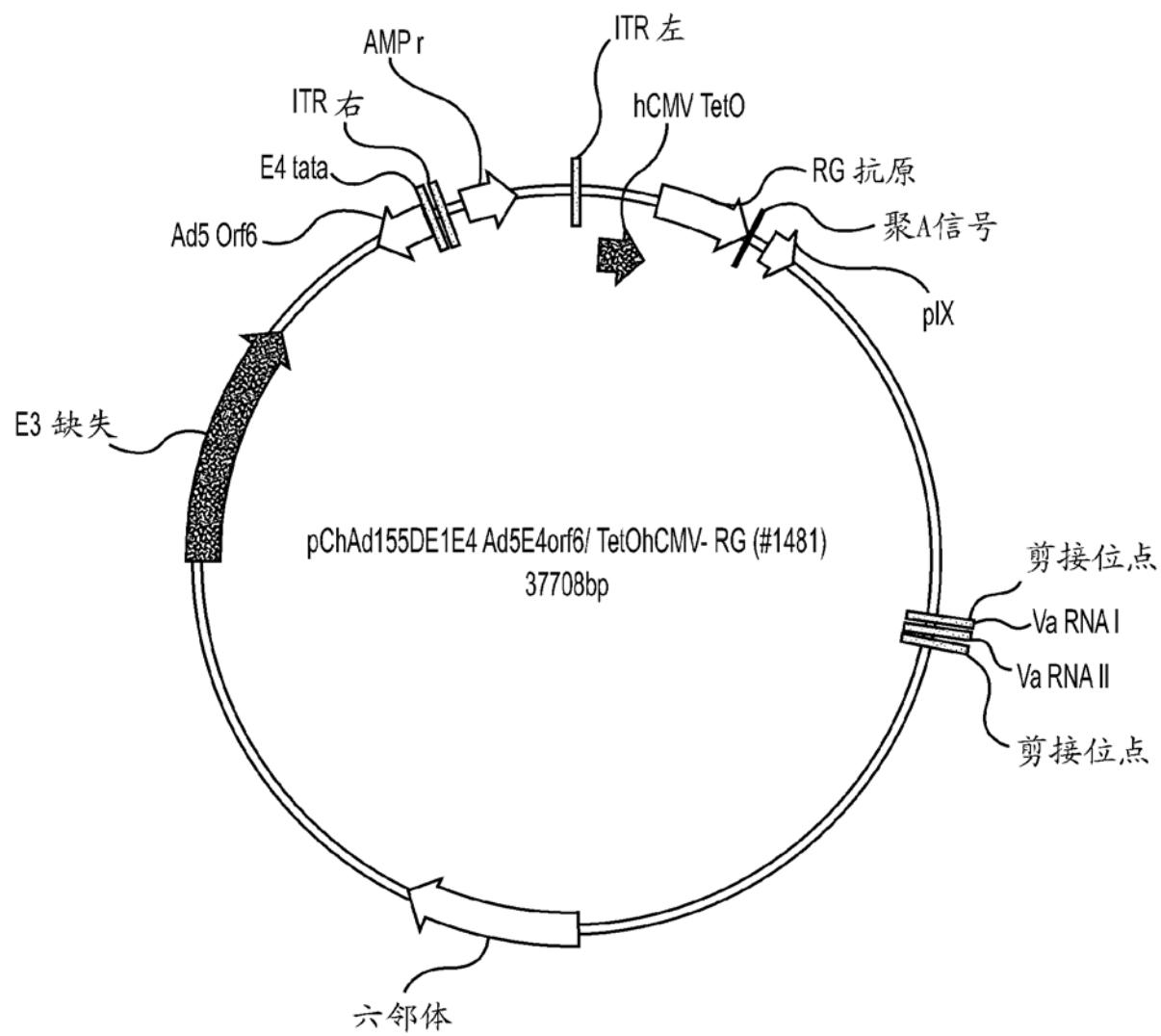


图 7

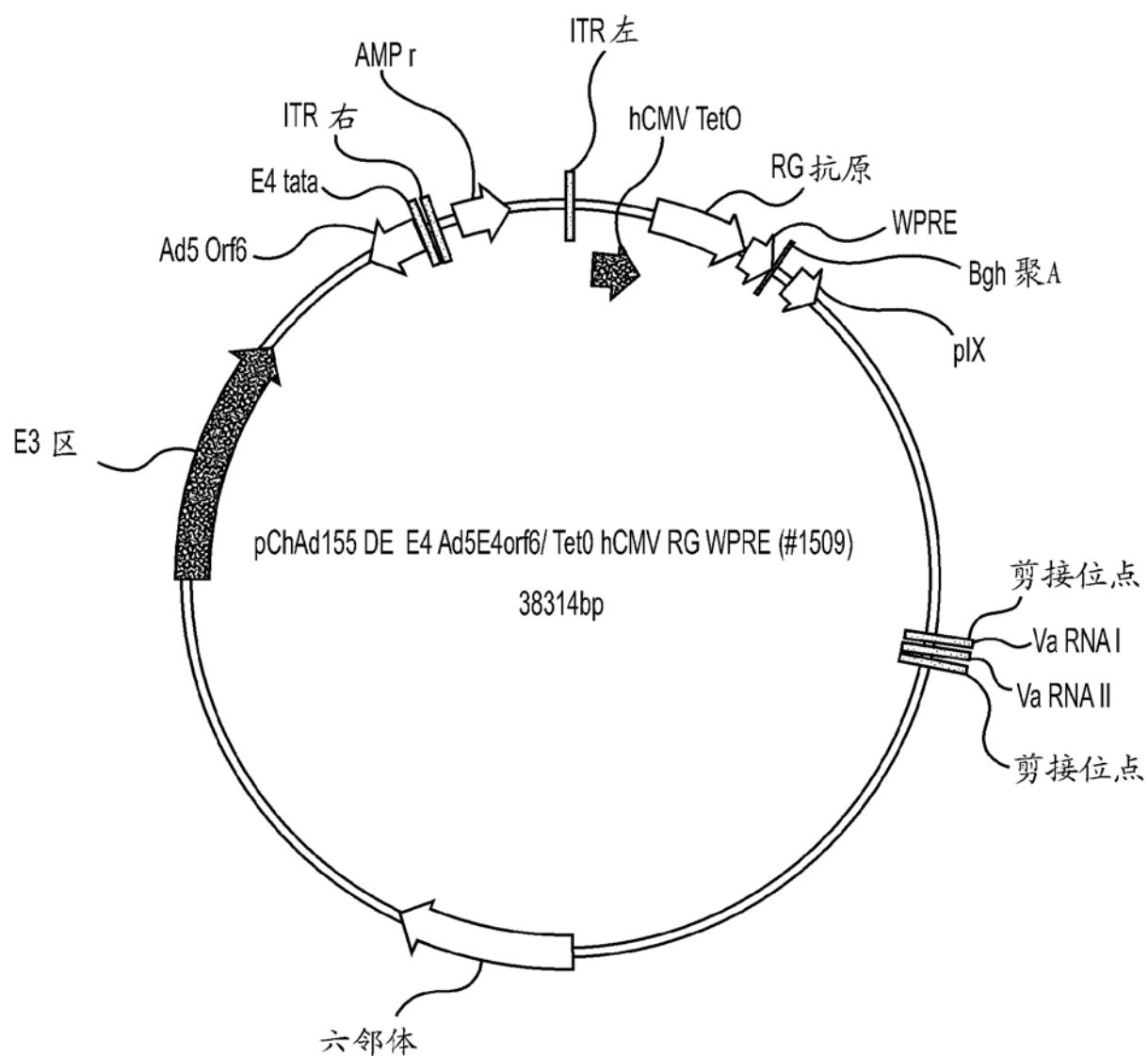


图 8

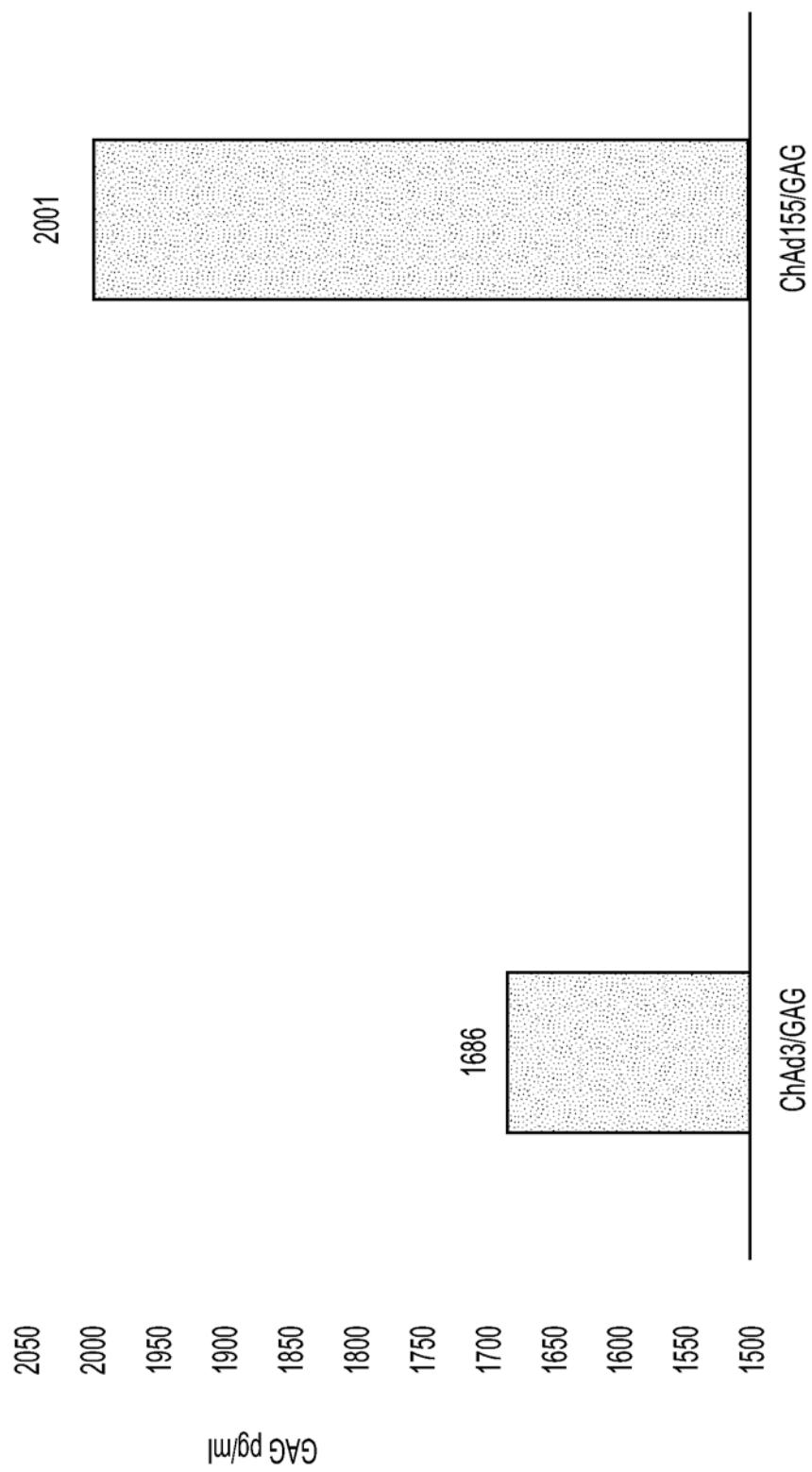


图 9

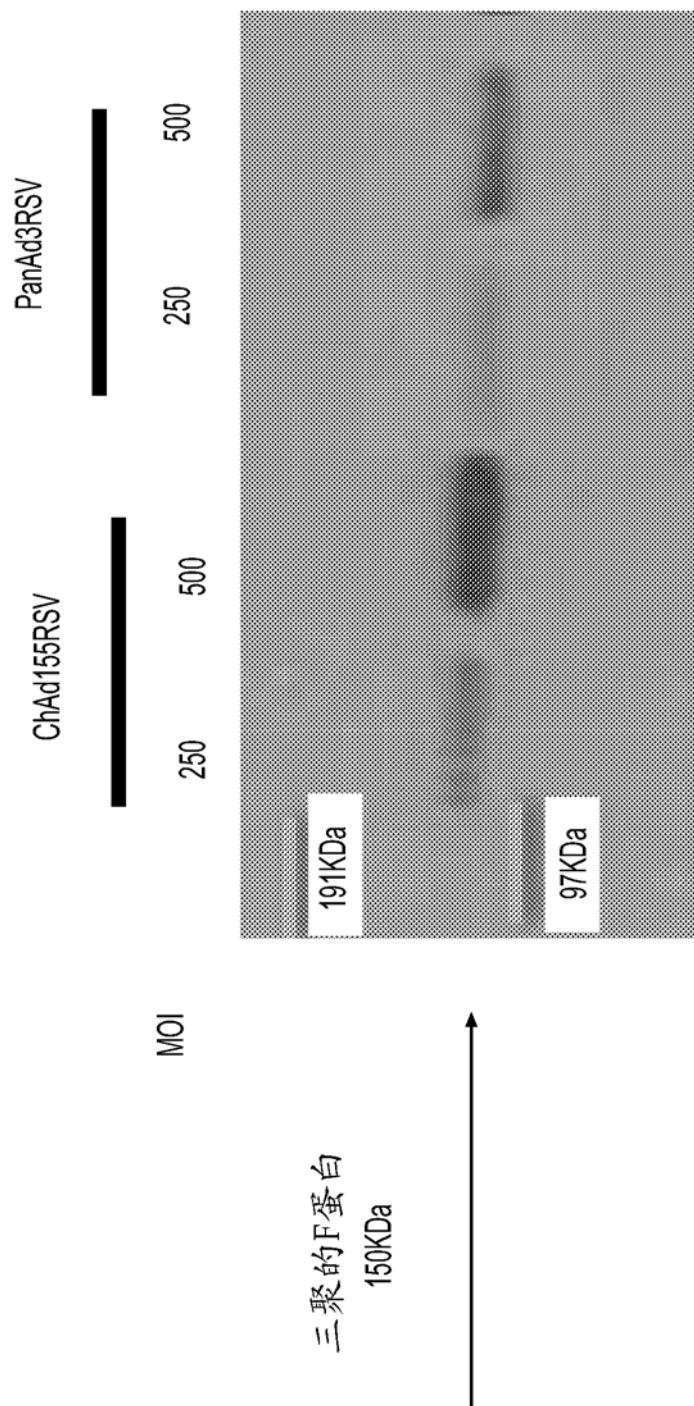


图 10

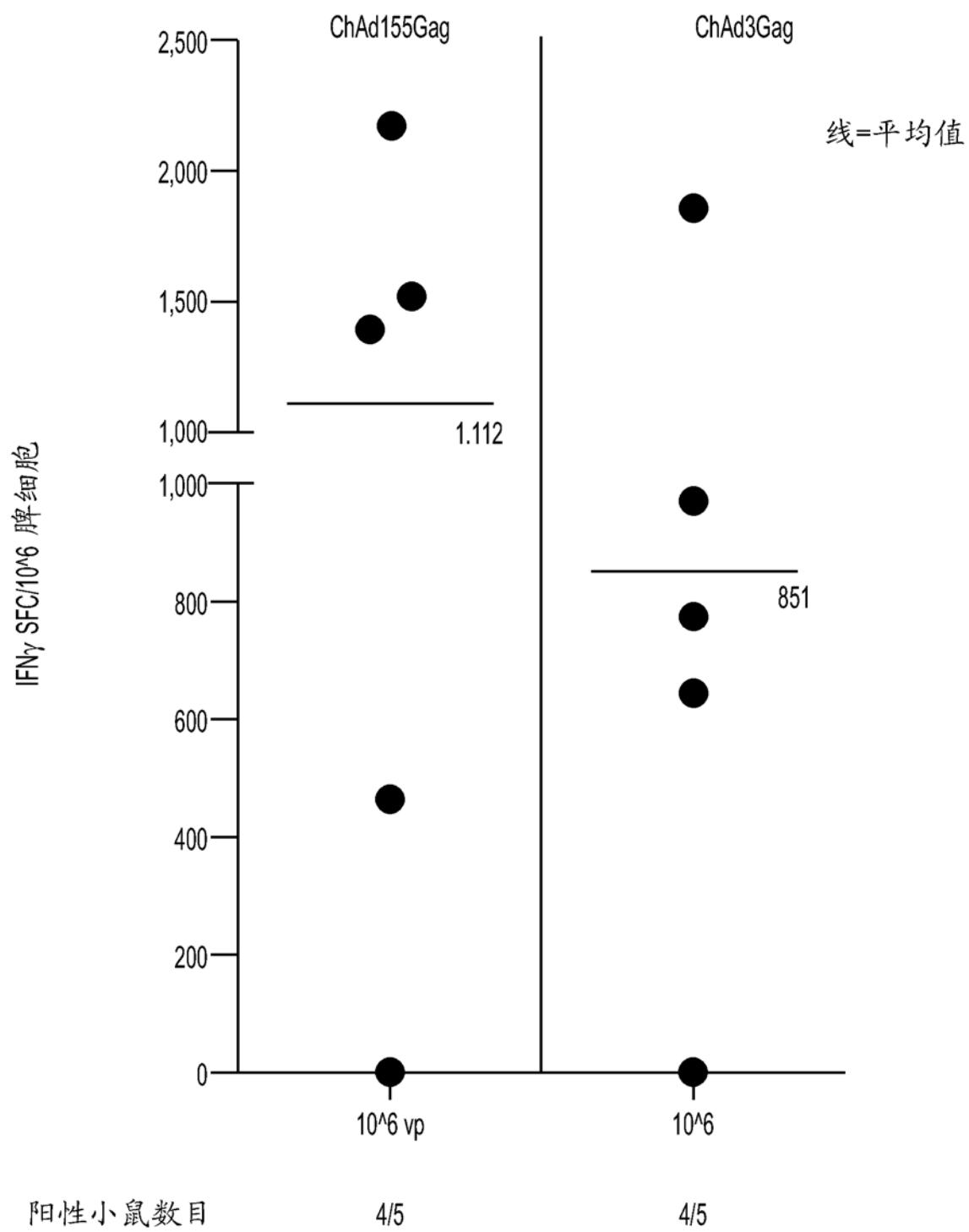


图 11

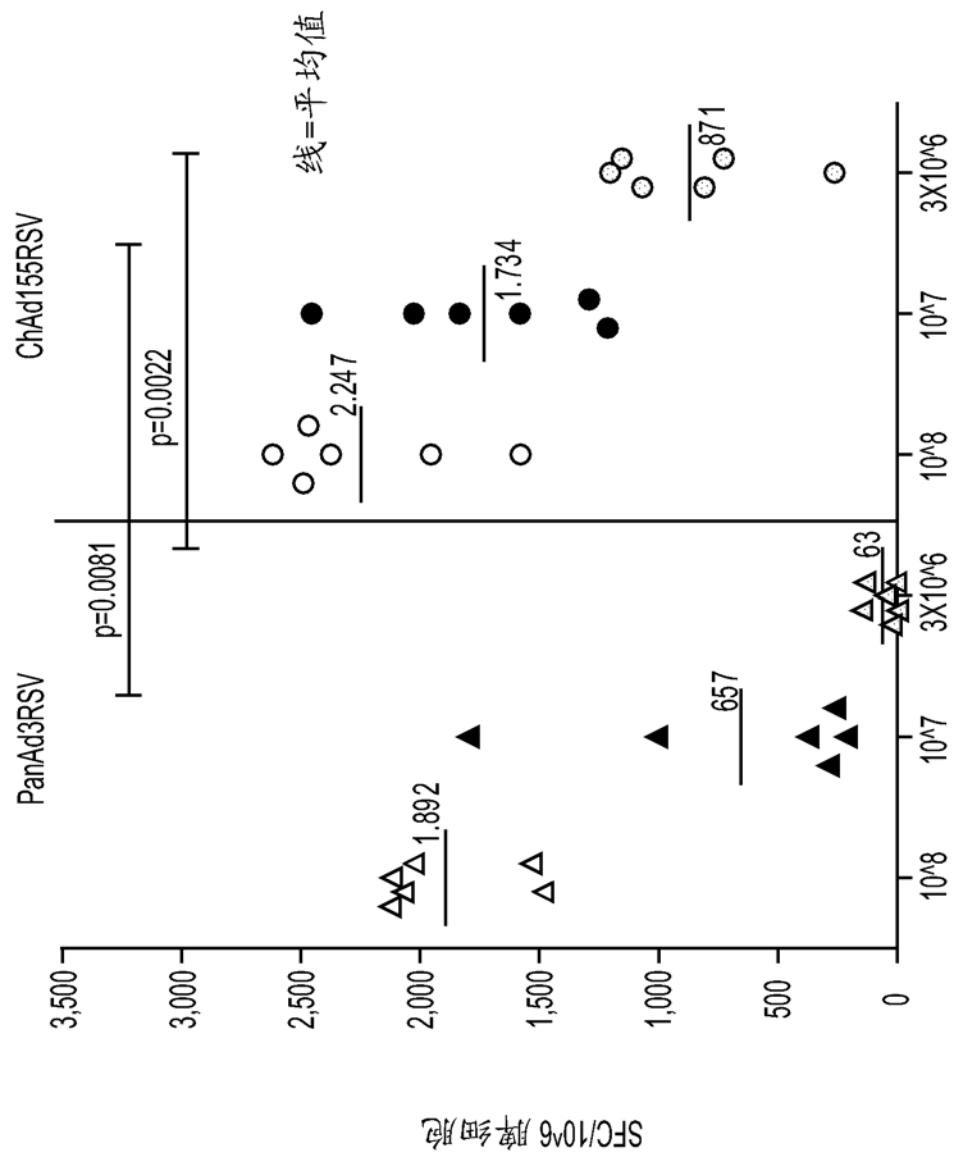
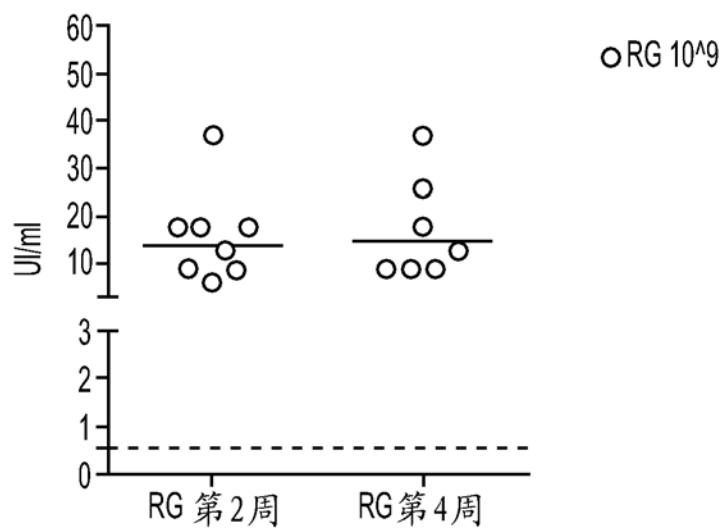


图 12

CD1小鼠中单次IM施用10<sup>9</sup>VP的ChAd155-RG

功能性抗体



抗狂犬病FAVN: 荧光抗体病毒中和

图 13A

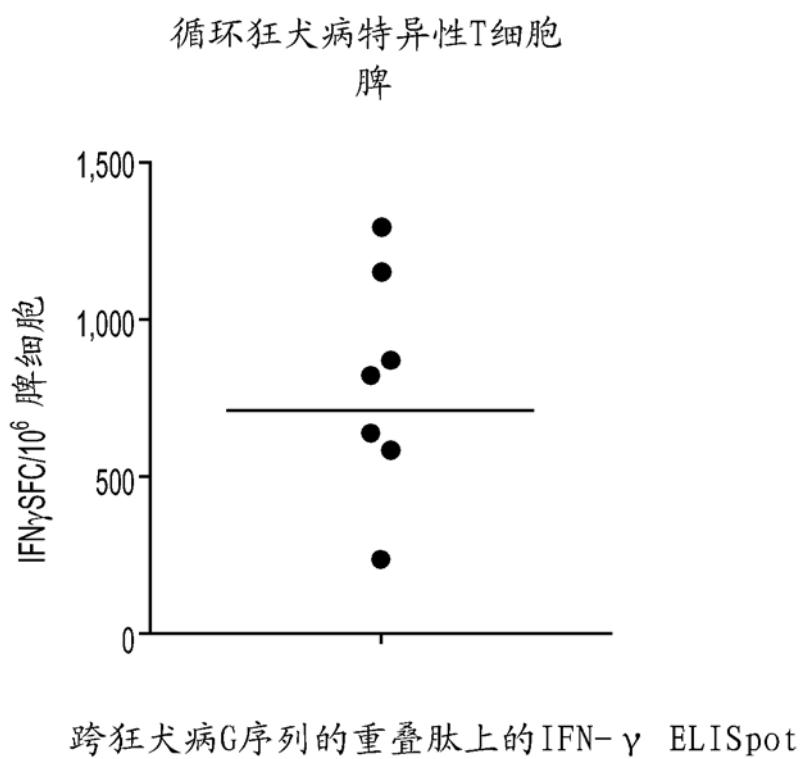


图 13B

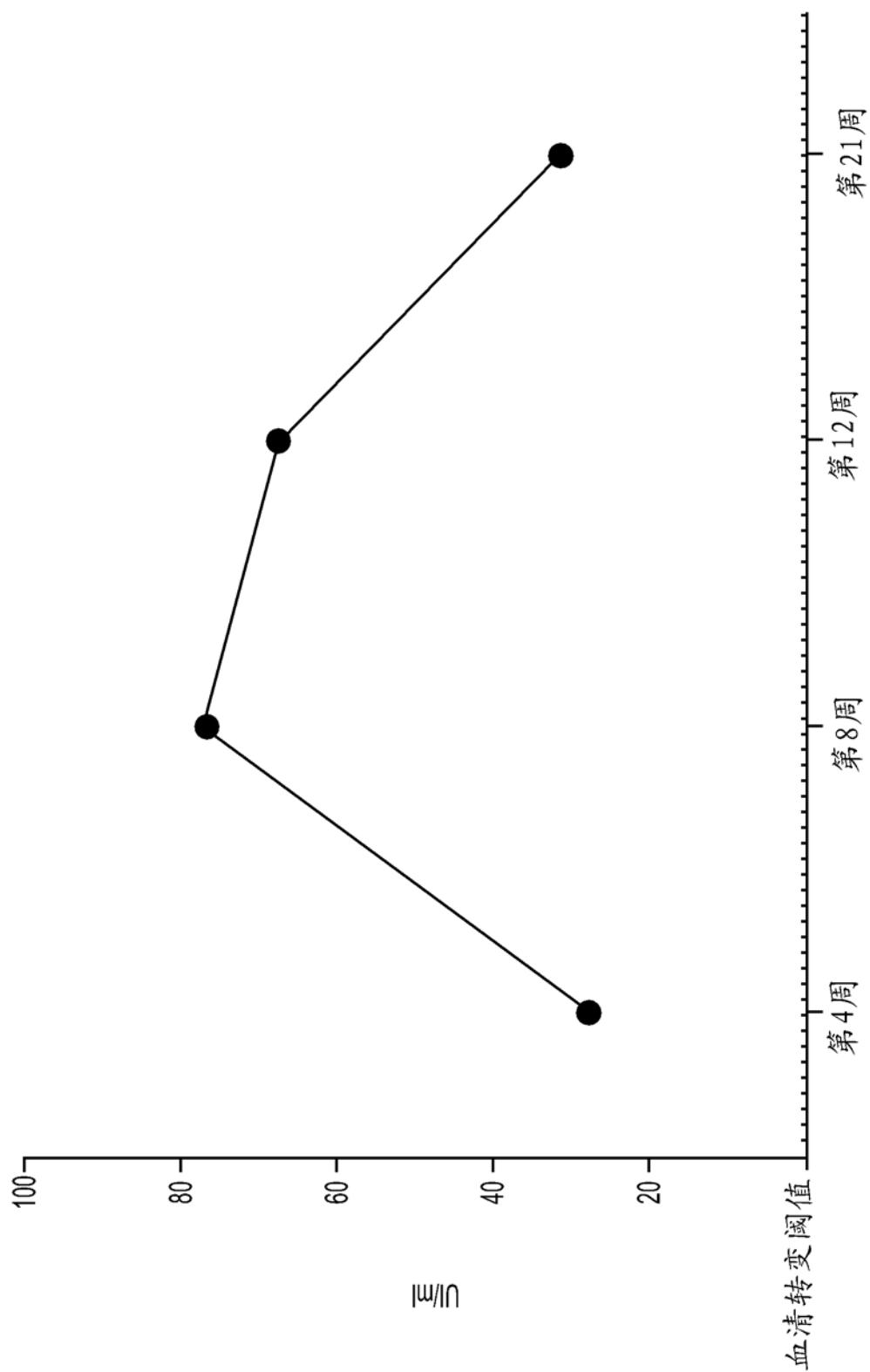
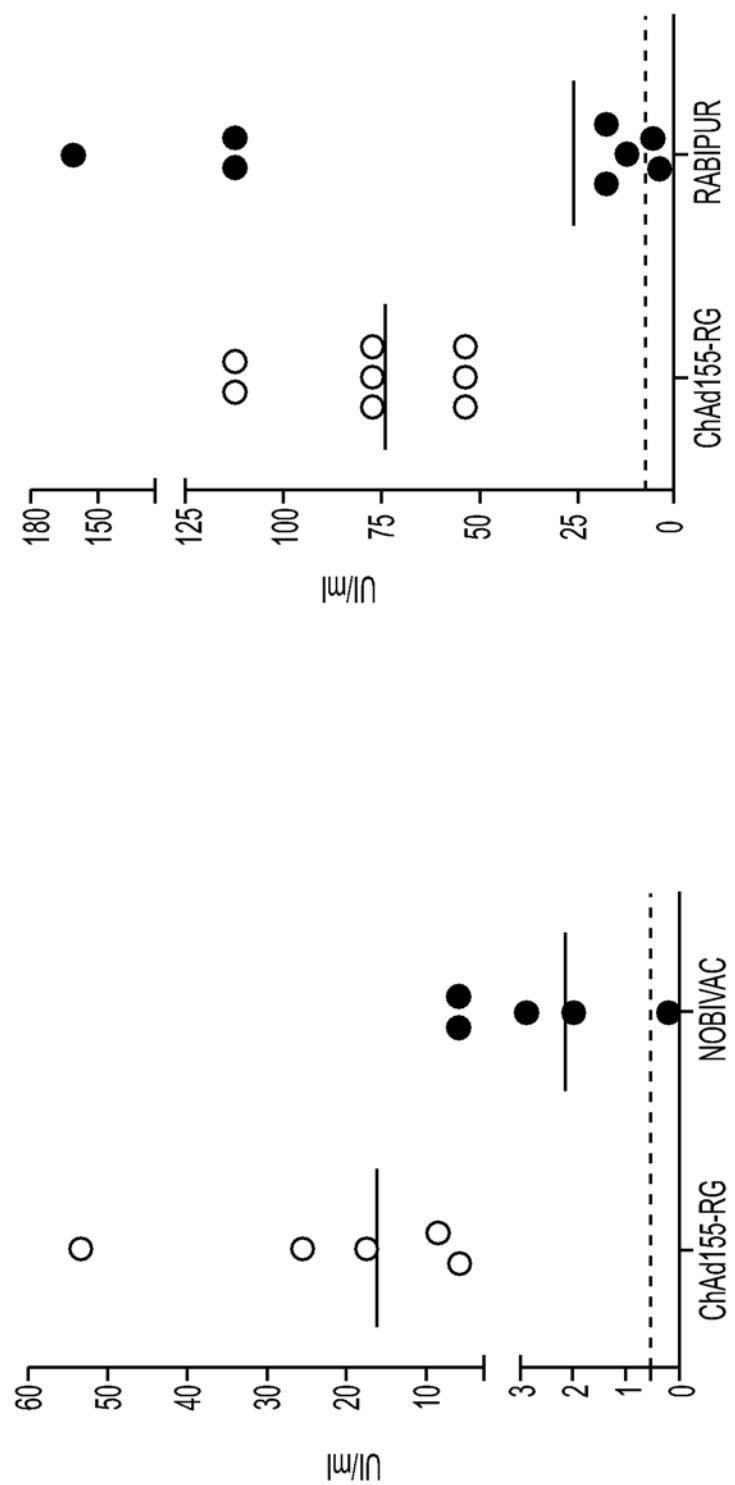


图 14



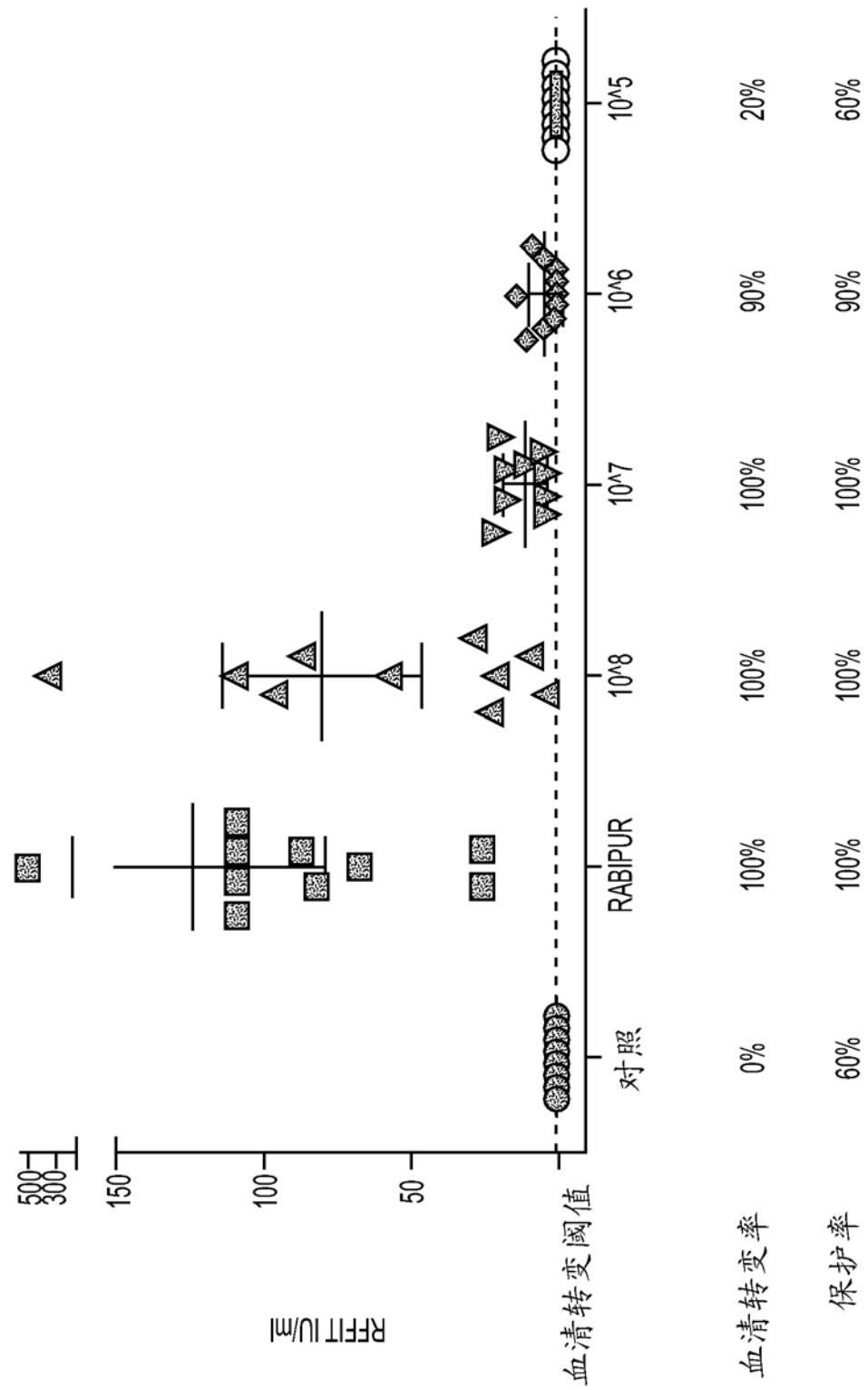


图 16

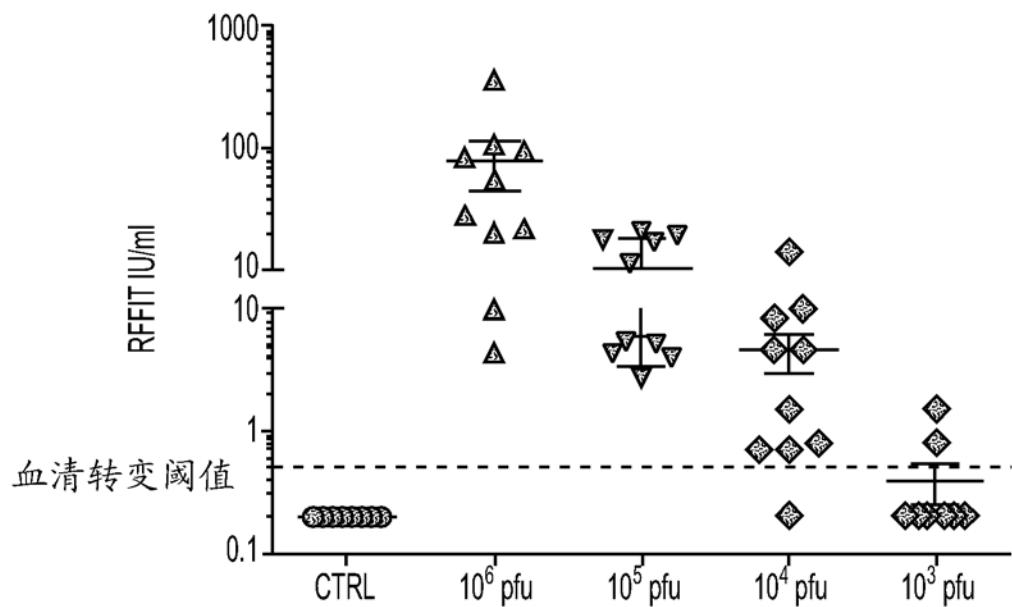


图 17A

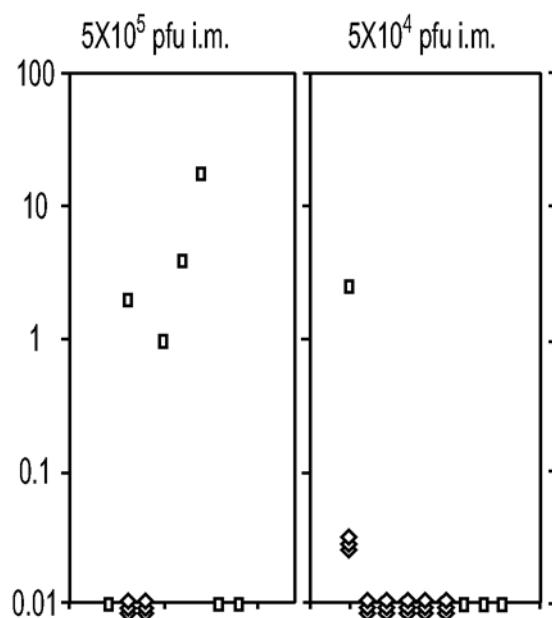


图 17B

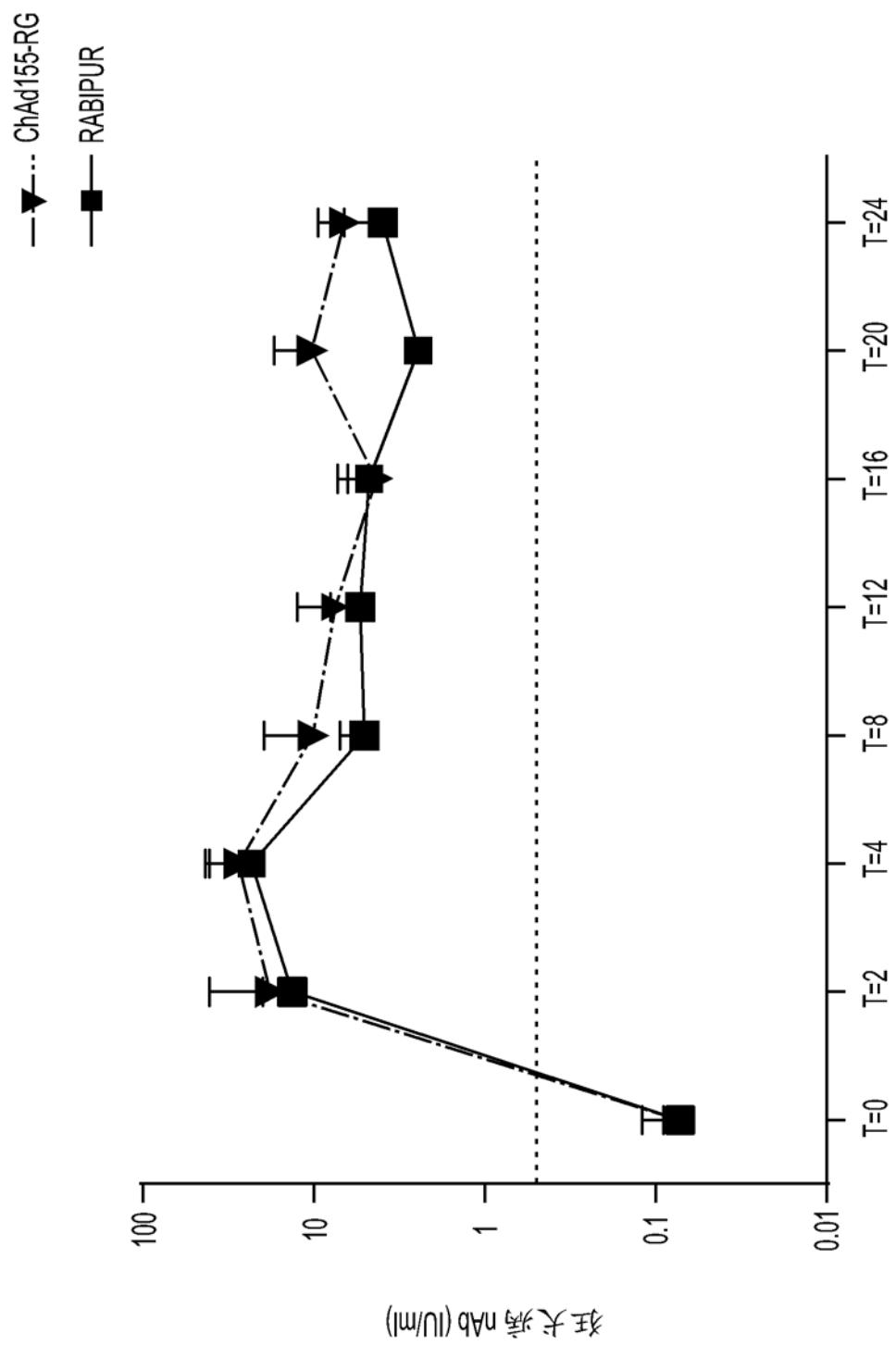


图 18

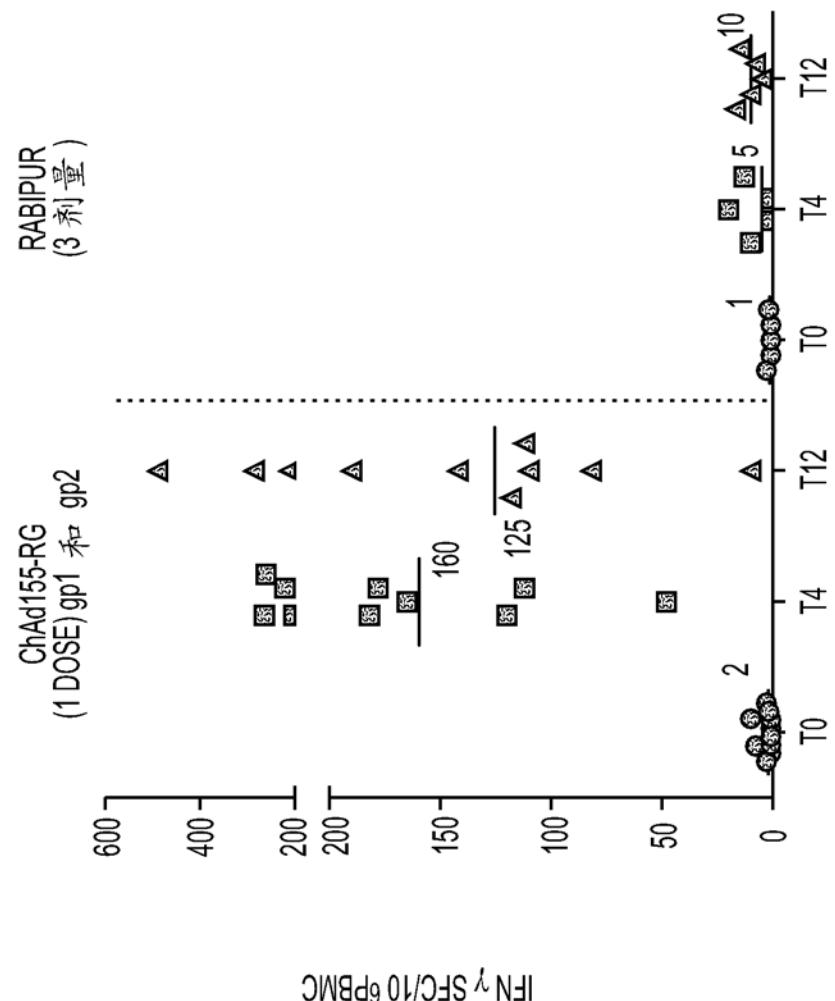


图 19