



(12) 发明专利申请

(10) 申请公布号 CN 105579581 A

(43) 申请公布日 2016. 05. 11

(21) 申请号 201480053358. 9

(51) Int. Cl.

(22) 申请日 2014. 08. 05

C12N 15/113(2006. 01)

(30) 优先权数据

61/862, 433 2013. 08. 05 US

(85) PCT国际申请进入国家阶段日

2016. 03. 25

(86) PCT国际申请的申请数据

PCT/US2014/049831 2014. 08. 05

(87) PCT国际申请的公布数据

W02015/021077 EN 2015. 02. 12

(71) 申请人 托卡根公司

地址 美国加利福尼亚州

(72) 发明人 A·H·林 H·E·格鲁伯

C·伊巴涅斯 D·J·约利

(74) 专利代理机构 北京派特恩知识产权代理有

限公司 11270

代理人 姚开丽 迟姗

权利要求书6页 说明书34页

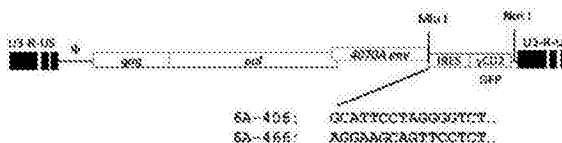
序列表90页 附图8页

(54) 发明名称

具有优化的A凸出物的重组载体

(57) 摘要

本公开描述用于基因疗法和基因递送的可复制型逆转录病毒载体 (RCR)。所述 RCR 包括在分叉区的 A 凸出物中具有 6-7 个 A 的 IRES 序列。本公开提供重组可复制型逆转录病毒, 其包含: 逆转录病毒 GAG 蛋白质; 逆转录病毒 POL 蛋白质; 逆转录病毒包膜; 逆转录病毒多核苷酸, 其包含处于逆转录病毒多核苷酸序列的 3' 端的长末端重复 (LTR) 序列、处于所述逆转录病毒多核苷酸的 5' 端的启动子序列。



1. 一种重组可复制型逆转录病毒,其包含:

逆转录病毒GAG蛋白质;

逆转录病毒POL蛋白质;

逆转录病毒包膜;

逆转录病毒多核苷酸,其包含处于逆转录病毒多核苷酸序列的3'端的长末端重复(LTR)序列;处于所述逆转录病毒多核苷酸的5'端的启动子序列,所述启动子适合于在哺乳动物细胞中表达;gag核酸结构域;pol核酸结构域;和env核酸结构域;

包含内部核糖体进入位点(IRES)的盒,所述IRES的分叉区的A凸出物由6个A组成,其中所述IRES可操作地连接到异源多核苷酸,其中所述盒位于3'LTR的5'和编码所述逆转录病毒包膜的所述env核酸结构域的3';和

在靶细胞中逆转录、包装和整合所必需的顺式作用序列,

其中与包含SEQ ID NO:21的载体(pACE)相比,所述RCR在6次传代后维持更高的复制能力。

2. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述病毒多次感染靶细胞,导致拷贝/二倍体基因组的平均数为5或更大。

3. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述逆转录病毒多核苷酸序列来源于选自由以下组成的组的病毒:鼠白血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV)、狒狒内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、猫源性逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性鼠白血病病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病病毒(GALV)。

4. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述逆转录病毒包膜是双嗜性MLV包膜。

5. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述逆转录病毒是 γ 逆转录病毒。

6. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述靶细胞是具有细胞增殖性病症的细胞。

7. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述靶细胞是赘生性细胞。

8. 根据权利要求6所述的重组可复制型逆转录病毒,其中所述细胞增殖性病选自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿性关节炎或其它自身免疫疾病。

9. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述启动子序列与生长调控基因缔合。

10. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述启动子序列包括组织特异性启动子序列。

11. 根据权利要求10所述的重组可复制型逆转录病毒,其中所述组织特异性启动子序列包含至少一个雄激素应答元件(ARE)。

12. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述启动子包含具有如SEQ ID NO:19、20、22或42中的从核苷酸1到约核苷酸582所示的序列的CMV启动子,并且可包括对一个或多个核酸碱基的修饰,并且所述启动子能够引导和起始转录。

13. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述启动子包含CMV-R-U5结构域多核苷酸。

14. 根据权利要求13所述的重组可复制型逆转录病毒,其中所述CMV-R-U5结构域包含连接到MLV R-U5区的来自人巨细胞病毒的立即早期启动子。

15. 根据权利要求14所述的重组可复制型逆转录病毒,其中所述CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的从约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中所示的序列至少95%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。

16. 根据权利要求1所述的重组可复制型逆转录病毒,其中gag多核苷酸来源于 γ 逆转录病毒。

17. 根据权利要求16所述的重组可复制型逆转录病毒,其中所述gag核酸结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号1203到约核苷酸2819的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。

18. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述多核苷酸的所述pol结构域来源于 γ 逆转录病毒。

19. 根据权利要求18所述的重组可复制型逆转录病毒,其中所述pol结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号2820到约核苷酸6358的序列或与其具有至少95%、98%、99%或99.9%同一性的序列。

20. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述env结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号6359到约核苷酸8323的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。

21. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述IRES由SEQ ID NO:41中所示的序列组成。

22. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述逆转录病毒多核苷酸序列包含(i)SEQ ID NO:42中所示的序列,或(ii)如SEQ ID NO:42中所示的序列,其中T是U。

23. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸包含具有如SEQ ID NO:3、5、11、13、15或17中所示序列的多核苷酸。

24. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸编码包含如SEQ ID NO:4中所示序列的多肽。

25. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸经过人密码子优化并且编码如SEQ ID NO:4中所示的多肽。

26. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸包含如SEQ ID NO:19或22中的从约核苷酸编号8877到约9353所示的序列。

27. 根据权利要求1所述的重组可复制型逆转录病毒,其中所述3'LTR来源于 γ 逆转录病毒。

28. 根据权利要求27所述的重组可复制型逆转录病毒,其中所述3'LTR包含U3-R-U5结构域。

29. 根据权利要求28所述的重组可复制型逆转录病毒,其中所述3'LTR包含如SEQ ID NO:19或22中的从约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%同一

的序列。

30. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸序列编码生物应答调节剂或免疫增强细胞因子。

31. 根据权利要求30所述的重组可复制型逆转录病毒,其中所述免疫增强细胞因子选自自由以下组成的组:白细胞介素1到15、干扰素、肿瘤坏死因子(TNF)和粒细胞-巨噬细胞-集落刺激因子(GM-CSF)。

32. 根据权利要求30所述的重组可复制型逆转录病毒,其中所述免疫增强细胞因子是 γ 干扰素。

33. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸编码将无毒前药转化为毒性药物的多肽。

34. 根据权利要求33所述的重组可复制型逆转录病毒,其中将无毒前药转化为毒性药物的所述多肽是胸苷激酶、嘌呤核苷磷酸化酶(PNP)或胞嘧啶脱氨酶。

35. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸序列编码受体结构域、抗体或抗体片段。

36. 根据权利要求1所述的重组可复制型逆转录病毒,其中异源核酸序列包含抑制性多核苷酸。

37. 根据权利要求36所述的重组可复制型逆转录病毒,其中所述抑制性多核苷酸包含miRNA、RNAi或siRNA序列。

38. 一种重组逆转录病毒多核苷酸基因组,其用于产生根据权利要求1所述的逆转录病毒。

39. 一种治疗细胞增殖性病症的方法,其包括使受试者与根据权利要求34所述的重组可复制型逆转录病毒在使得胞嘧啶脱氨酶多核苷酸被表达的条件下接触,和使所述受试者与5-氟胞嘧啶接触。

40. 根据权利要求39所述的方法,其中所述细胞增殖性病症是多形性胶质母细胞瘤。

41. 根据权利要求39所述的方法,其中所述细胞增殖性病症选自自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌。

42. 一种在哺乳动物细胞中由ECMV IRES表达异源基因的载体,所述ECMV IRES在J-K分叉区的A凸出物中具有6个A。

43. 根据权利要求42所述的载体,其中所述载体是病毒载体。

44. 根据权利要求42和43所述的载体,其中所述载体是逆转录病毒复制载体。

45. 根据权利要求42至44中任一项所述的载体,其中所述载体是来源于 γ -逆转录病毒的逆转录病毒复制载体。

46. 根据权利要求45所述的载体,其中所述 γ -逆转录病毒来源于鼠白血病病毒、狒狒内源性病毒、长臂猿白血病病毒、猫白血病病毒中的一种。

47. 根据权利要求42至46中任一项所述的载体,其中所述异源基因是在哺乳动物中具有治疗活性的基因。

48. 根据权利要求47所述的载体,其中所述治疗活性是抗癌活性。

49. 根据权利要求47所述的载体,其中所述异源基因是前药活化基因。

50. 根据权利要求42至49中任一项所述的载体,其中所述载体可以在不存在蛋白质PTB-1的情况下在哺乳动物细胞中由ECMV IRES表达异源基因。

51. 一种通过施用根据权利要求42至50所述的载体来治疗癌症的方法。

52. 一种重组可复制型逆转录病毒,其包含:

逆转录病毒GAG蛋白质;

逆转录病毒POL蛋白质;

逆转录病毒包膜;

逆转录病毒多核苷酸,其包含处于逆转录病毒多核苷酸序列的3'端的长末端重复(LTR)序列;处于所述逆转录病毒多核苷酸的5'端的启动子序列,所述启动子适合于在哺乳动物细胞中表达;gag核酸结构域;pol核酸结构域;和env核酸结构域;

盒,所述盒包含(i)最小内部核糖体进入位点(IRES),其中所述最小IRES可操作地连接到异源多核苷酸,(ii)连接到miRNA的polIII启动子,或(iii)(ii)之前或之后的可操作地连接到异源多核苷酸的迷你启动子,其中所述盒位于3'LTR的5'和编码所述逆转录病毒包膜的所述env核酸结构域的3';和

在靶细胞中逆转录、包装和整合所必需的顺式作用序列。

53. 根据权利要求52所述的可复制型逆转录病毒,其中所述最小IRES由SEQ ID NO:41的从约碱基123到544的序列组成。

54. 根据权利要求52所述的可复制型逆转录病毒,其中所述最小IRES由SEQ ID NO:41的从约碱基183到544的序列组成。

55. 根据权利要求52至54所述的可复制型逆转录病毒,其中所述IRES在A凸出物中具有6个A。

56. 根据权利要求52所述的可复制型逆转录病毒,其中所述病毒多次感染靶细胞,导致拷贝/二倍体基因组的平均数为5或更大。

57. 根据权利要求52所述的可复制型逆转录病毒,其中所述逆转录病毒多核苷酸序列来源于选自由以下组成的组的病毒:鼠白血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV)、狒狒内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、猫源性逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性鼠白血病病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病病毒(GALV)。

58. 根据权利要求52所述的可复制型逆转录病毒,其中所述逆转录病毒包膜是双嗜性MLV包膜。

59. 根据权利要求52所述的可复制型逆转录病毒,其中所述逆转录病毒是 γ 逆转录病毒。

60. 根据权利要求52所述的可复制型逆转录病毒,其中所述靶细胞是具有细胞增殖性病症的细胞。

61. 根据权利要求52所述的可复制型逆转录病毒,其中所述靶细胞是赘生性细胞。

62. 根据权利要求60所述的可复制型逆转录病毒,其中所述细胞增殖性病选自自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿性关节炎或其它自身免疫疾病。

63. 根据权利要求52所述的可复制型逆转录病毒,其中所述启动子序列与生长调

控基因缔合。

64. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述启动子序列包括组织特异性启动子序列。

65. 根据权利要求64所述的重组可复制型逆转录病毒,其中所述组织特异性启动子序列包含至少一个雄激素应答元件(ARE)。

66. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述启动子包含具有如SEQ ID NO:19、20、22或42中的从核苷酸1到约核苷酸582所示的序列的CMV启动子,并且可包括对一个或多个核酸碱基的修饰,并且所述启动子能够引导和起始转录。

67. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述启动子包含CMV-R-U5结构域多核苷酸。

68. 根据权利要求67所述的重组可复制型逆转录病毒,其中所述CMV-R-U5结构域包含连接到MLV R-U5区的来自人巨细胞病毒的立即早期启动子。

69. 根据权利要求68所述的重组可复制型逆转录病毒,其中所述CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的从约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中所示的序列至少95%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。

70. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述gag多核苷酸来源于 γ 逆转录病毒。

71. 根据权利要求70所述的重组可复制型逆转录病毒,其中所述gag核酸结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号1203到约核苷酸2819的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。

72. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述多核苷酸的pol结构域来源于 γ 逆转录病毒。

73. 根据权利要求72所述的重组可复制型逆转录病毒,其中所述pol结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号2820到约核苷酸6358的序列或与其具有至少95%、98%、99%或99.9%同一性的序列。

74. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述env结构域包含SEQ ID NO:19、20、22或42的从约核苷酸编号6359到约核苷酸8323的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。

75. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸包含具有如SEQ ID NO:3、5、11、13、15或17中所示的序列的多核苷酸。

76. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸编码包含如SEQ ID NO:4中所示的序列的多肽。

77. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸经过人密码子优化并且编码如SEQ ID NO:4中所示的多肽。

78. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸包含如SEQ ID NO:19或22中的从约核苷酸编号8877到约9353所示的序列。

79. 根据权利要求52所述的重组可复制型逆转录病毒,其中所述3'LTR来源于 γ 逆转录病毒。

80. 根据权利要求79所述的重组可复制型逆转录病毒,其中所述3'LTR包含U3-R-U5结构域。

81. 根据权利要求79所述的重组可复制型逆转录病毒,其中所述3'LTR包含如SEQ ID NO:19或22中的从约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%同一的序列。

82. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸序列编码生物应答调节剂或免疫增强细胞因子。

83. 根据权利要求82所述的重组可复制型逆转录病毒,其中所述免疫增强细胞因子选自自由以下组成的组:白细胞介素1到15、干扰素、肿瘤坏死因子(TNF)和粒细胞-巨噬细胞-集落刺激因子(GM-CSF)。

84. 根据权利要求82所述的重组可复制型逆转录病毒,其中所述免疫增强细胞因子是 γ 干扰素。

85. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸编码将无毒前药转化为毒性药物的多肽。

86. 根据权利要求85所述的重组可复制型逆转录病毒,其中将无毒前药转化为毒性药物的所述多肽是胸苷激酶、嘌呤核苷磷酸化酶(PNP)或胞嘧啶脱氨酶。

87. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸序列编码受体结构域、抗体或抗体片段。

88. 根据权利要求52所述的重组可复制型逆转录病毒,其中异源核酸序列包含抑制性多核苷酸。

89. 根据权利要求88所述的重组可复制型逆转录病毒,其中所述抑制性多核苷酸包含miRNA、RNAi或siRNA序列。

90. 一种重组逆转录病毒多核苷酸基因组,其用于产生根据权利要求52所述的的逆转录病毒。

91. 一种治疗细胞增殖性病症的方法,其包括使受试者与根据权利要求86所述的重组可复制型逆转录病毒在使得胞嘧啶脱氨酶多核苷酸被表达的条件下接触,和使所述受试者与5-氟胞嘧啶接触。

92. 根据权利要求91所述的方法,其中所述细胞增殖性病症是多形性胶质母细胞瘤。

93. 根据权利要求91所述的方法,其中所述细胞增殖性病症选自自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌。

具有优化的A凸出物的重组载体

[0001] 相关申请的交叉引用

[0002] 本申请要求2013年8月5日提交的美国临时申请第61/862,433号的优先权,所述美国临时申请的公开内容以引用方式并入本文中。

技术领域

[0003] 本公开涉及优化的内部核糖体进入位点(IRES)和含有包含此类优化的IRES的载体的组合物。更具体地说,本公开涉及用于治疗细胞增殖性病症的可复制型逆转录病毒载体。本公开还涉及此类可复制型逆转录病毒载体用于递送和表达异源核酸的用途。

背景技术

[0004] 将基因和异源核酸递送给细胞和受试者的有效方法已经成为研究人员进行科学研究以及疾病和病症的可能治疗的目标。

[0005] 序列表的并入

[0006] 本申请随电子格式的序列表一起提交。序列表作为题为00014-019W01Sequence_ST25.txt的文档提供,所述文档创建于2014年8月5日,大小为202Kb。电子格式的序列表中的信息以全文引用的方式并入本文中。

发明内容

[0007] 本公开提供重组可复制型逆转录病毒,其包含:逆转录病毒GAG蛋白质;逆转录病毒POL蛋白质;逆转录病毒包膜;逆转录病毒多核苷酸,其包含处于逆转录病毒多核苷酸序列的3'端的长末端重复(LTR)序列、处于逆转录病毒多核苷酸的5'端的启动子序列(所述启动子适合于在哺乳动物细胞中表达)、gag核酸结构域、pol核酸结构域和env核酸结构域;包含内部核糖体进入位点(IRES)的盒,所述IRES的分叉区的A凸出物由6个A组成,其中IRES可操作地连接到异源多核苷酸,其中所述盒位于3'LTR的5'和编码逆转录病毒包膜的env核酸结构域的3';和在靶细胞中逆转录、包装和整合所必需的顺式作用序列,其中与包含SEQ ID NO:21的载体(pACE)相比,RCR在6次传代后维持更高的复制能力。在一个实施方案中,病毒多次感染靶细胞,导致拷贝/二倍体基因组的平均数为5或更大。在任何上文的另一个实施方案中,逆转录病毒多核苷酸序列来源于选自由以下组成的组的病毒:鼠白血病病毒(MLV)、莫洛尼(Mooney)鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV)、狒狒内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、猫源性逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性鼠白血病病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病病毒(GALV)。在任何上文的另一个实施方案中,逆转录病毒包膜是双嗜性MLV包膜。在任何上文的另一个实施方案中,逆转录病毒是 γ 逆转录病毒。在任何上文的另一个实施方案中,靶细胞是具有细胞增殖性病症的细胞。在任何上文的另一个实施方案中,靶细胞是赘生性细胞。在任何上文的另一个实施方案中,细胞增殖性病症选自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿性关

节炎或其它自身免疫疾病。在任何上文的另一个实施方案中,启动子序列与生长调控基因缔合。在任何上文的另一个实施方案中,启动子序列包括组织特异性启动子序列。在任何上文的另一个实施方案中,组织特异性启动子序列包含至少一个雄激素应答元件(ARE)。在任何上文的另一个实施方案中,启动子包含具有如SEQ ID NO:19、20、22或42中的核苷酸1到约核苷酸582所示的序列的CMV启动子,并且可包括对一个或多个核酸碱基的修饰,并且所述启动子能够引导和起始转录。在任何上文的另一个实施方案中,启动子包含CMV-R-U5结构域多核苷酸。在任何上文的另一个实施方案中,CMV-R-U5结构域包含来自人巨细胞病毒的连接到MLV R-U5区的立即早期启动子。在任何上文的另一个实施方案中,CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中所示的序列至少95%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。在任何上文的另一个实施方案中,gag多核苷酸来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,gag核酸结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号1203到约核苷酸2819的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。在任何上文的另一个实施方案中,多核苷酸的pol结构域来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,pol结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号2820到约核苷酸6358的序列或与其具有至少95%、98%、99%或99.9%同一性的序列。在任何上文的另一个实施方案中,env结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号6359到约核苷酸8323的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。在任何上文的另一个实施方案中,IRES由SEQ ID NO:41中所示的序列组成。在任何上文的另一个实施方案中,逆转录病毒多核苷酸序列包含(i)SEQ ID NO:42中所示的序列,或(ii)如SEQ ID NO:42中所示的序列,其中T是U。在任何上文的另一个实施方案中,异源核酸包含具有如SEQ ID NO:3、5、11、13、15或17中所示的序列的多核苷酸。在任何上文的另一个实施方案中,异源核酸编码包含如SEQ ID NO:4中所示的序列的多肽。在任何上文的另一个实施方案中,异源核酸经过人密码子优化并且编码如SEQ ID NO:4中所示的多肽。在任何上文的另一个实施方案中,异源核酸包含如SEQ ID NO:19或22中的约核苷酸编号8877到约9353所示的序列。在任何上文的另一个实施方案中,3'LTR来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,3'LTR包含U3-R-U5结构域。在任何上文的另一个实施方案中,3'LTR包含如SEQ ID NO:19或22中的约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%同一的序列。在任何上文的另一个实施方案中,异源核酸序列编码生物应答调节剂或免疫增强细胞因子。在任何上文的另一个实施方案中,免疫增强细胞因子选自自由以下组成的组:白细胞介素1到15、干扰素、肿瘤坏死因子(TNF)和粒细胞-巨噬细胞-集落刺激因子(GM-CSF)。在任何上文的另一个实施方案中,免疫增强细胞因子是 γ 干扰素。在任何上文的另一个实施方案中,异源核酸编码将无毒前药转化为毒性药物的多肽。在任何上文的另一个实施方案中,将无毒前药转化为毒性药物的多肽是胸苷激酶、嘌呤核苷磷酸化酶(PNP)或胞嘧啶脱氨酶。在任何上文的另一个实施方案中,异源核酸序列编码受体结构域、抗体或抗体片段。在任何上文的另一个实施方案中,异源核酸序列包含抑制性多核苷酸。在任何上文的另一个实施方案中,抑制性多核苷酸包含miRNA、RNAi或siRNA序列。

[0008] 本公开还提供重组逆转录病毒多核苷酸基因组,用于产生如上所述的可复制型逆转录病毒。

[0009] 本公开还提供一种治疗细胞增殖性病症的方法,其包括使受试者与本公开的重组可复制型逆转录病毒在使得胞嘧啶脱氨酶多核苷酸被表达的条件下接触,和使所述受试者与5-氟胞嘧啶接触。在另一个实施方案中,细胞增殖性病症是多形性胶质母细胞瘤。在任何上文的另一个实施方案中,细胞增殖性病症选自以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌。

[0010] 本公开还提供一种载体,其在哺乳动物细胞中由ECMV IRES表达异源基因,所述ECMV IRES在J-K分叉区的A凸出物中具有6个A。在另一个实施方案中,载体是病毒载体。在任何上文的另一个实施方案中,载体是逆转录病毒复制载体。在任何上文的另一个实施方案中,载体是来源于 γ -逆转录病毒的逆转录病毒复制载体。在任何上文的另一个实施方案中, γ -逆转录病毒来源于鼠白血病病毒、狨猴内源性病毒、长臂猿白血病病毒、猫白血病病毒中的一种。在任何上文的另一个实施方案中,异源基因是在哺乳动物中具有治疗活性的基因。在任何上文的另一个实施方案中,治疗活性是抗癌活性。在任何上文的另一个实施方案中,异源基因是前药活化基因。在任何上文的另一个实施方案中,载体可以在不存在蛋白质PTB-1的情况下在哺乳动物细胞中由ECMV IRES表达异源基因。

[0011] 本公开还提供一种通过施用如上所述的载体来治疗癌症的方法。

[0012] 本公开还提供一种重组可复制型逆转录病毒,其包含:逆转录病毒GAG蛋白质;逆转录病毒POL蛋白质;逆转录病毒包膜;逆转录病毒多核苷酸,其包含处于逆转录病毒多核苷酸序列的3'端的长末端重复(LTR)序列、处于逆转录病毒多核苷酸的5'端的启动子序列(所述启动子适合于在哺乳动物细胞中表达)、gag核酸结构域、pol核酸结构域和env核酸结构域;盒,所述盒包含(i)最小内部核糖体进入位点(IRES),其中最小IRES可操作地连接到异源多核苷酸,(ii)连接到miRNA上的polIII启动子,或(iii)(ii)之前或之后的可操作地连接到异源多核苷酸的迷你启动子(mini-promoter),其中所述盒位于3'LTR的5'和编码逆转录病毒包膜的env核酸结构域的3';和在靶细胞中逆转录、包装和整合所必需的顺式作用序列。在一个实施方案中,最小IRES由SEQ ID NO:41的约碱基123到544的序列组成。在任何上文的另一个实施方案中,最小IRES由SEQ ID NO:41的约碱基183到544的序列组成。在任何上文的另一个实施方案中,IRES在A凸出物中具有6个A。在任何上文的另一个实施方案中,病毒多次感染靶细胞,导致拷贝/二倍体基因组的平均数为5或更大。在任何上文的另一个实施方案中,逆转录病毒多核苷酸序列来源于选自以下组成的组的病毒:鼠白血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病毒(FeLV)、狨猴内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、猫源性逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性鼠白血病病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病病毒(GALV)。在任何上文的另一个实施方案中,逆转录病毒包膜是双嗜性MLV包膜。在任何上文的另一个实施方案中,逆转录病毒包膜是 γ 逆转录病毒。在任何上文的另一个实施方案中,靶细胞是具有细胞增殖性病症的细胞。在任何上文的另一个实施方案中,靶细胞是赘生性细胞。在任何上文的另一个实施方案中,细胞增殖性病症选自以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌、类风湿性关节炎或其它自身免疫疾病。在任何上文的另一个实施方案中,启动子序列与生长调控基因缔合。在任何上文的另一个实施方案中,启动子序列包括组织特异性启动子序列。在任何上文的另一个实施方案中,组织特异性启动子序列包含至少一个雄激素应答元件(ARE)。在任

何上文的另一个实施方案中,启动子包含具有如SEQ ID NO:19、20、22或42中的核苷酸1到约核苷酸582所示的序列的CMV启动子,并且可包括对一个或多个核酸碱基的修饰,并且所述启动子能够引导和起始转录。在任何上文的另一个实施方案中,启动子包含CMV-R-U5结构域多核苷酸。在任何上文的另一个实施方案中,CMV-R-U5结构域包含来自人巨细胞病毒的连接到MLV R-U5区的立即早期启动子。在任何上文的另一个实施方案中,CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中所示的序列至少95%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。在任何上文的另一个实施方案中,gag多核苷酸来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,gag核酸结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号1203到约核苷酸2819的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。在任何上文的另一个实施方案中,多核苷酸的pol结构域来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,pol结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号2820到约核苷酸6358的序列或与其具有至少95%、98%、99%或99.9%同一性的序列。在任何上文的另一个实施方案中,env结构域包含SEQ ID NO:19、20、22或42的约核苷酸编号6359到约核苷酸8323的序列或与其具有至少95%、98%、99%或99.8%同一性的序列。在任何上文的另一个实施方案中,异源核酸包含具有如SEQ ID NO:3、5、11、13、15或17中所示的序列的多核苷酸。在任何上文的另一个实施方案中,异源核酸编码包含如SEQ ID NO:4中所示的序列的多肽。在任何上文的另一个实施方案中,异源核酸经过人密码子优化并且编码如SEQ ID NO:4中所示的多肽。在任何上文的另一个实施方案中,异源核酸包含如SEQ ID NO:19或22中的约核苷酸编号8877到约9353所示的序列。在任何上文的另一个实施方案中,3'LTR来源于 γ 逆转录病毒。在任何上文的另一个实施方案中,3'LTR包含U3-R-U5结构域。在任何上文的另一个实施方案中,3'LTR包含如SEQ ID NO:19或22中的约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%同一的序列。在任何上文的另一个实施方案中,异源核酸序列编码生物应答调节剂或免疫增强细胞因子。在任何上文的另一个实施方案中,免疫增强细胞因子选自由以下组成的组:白细胞介素1到15、干扰素、肿瘤坏死因子(TNF)和粒细胞-巨噬细胞-集落刺激因子(GM-CSF)。在任何上文的另一个实施方案中,免疫增强细胞因子是 γ 干扰素。在任何上文的另一个实施方案中,异源核酸编码将无毒前药转化为毒性药物的多肽。在任何上文的另一个实施方案中,将无毒前药转化为毒性药物的多肽是胸苷激酶、嘌呤核苷磷酸化酶(PNP)或胞嘧啶脱氨酶。在任何上文的另一个实施方案中,异源核酸序列编码受体结构域、抗体或抗体片段。在任何上文的另一个实施方案中,异源核酸序列包含抑制性多核苷酸。在任何上文的另一个实施方案中,抑制性多核苷酸包含miRNA、RNAi或siRNA序列。

[0013] 本公开的一个或多个实施方案的细节在附图和以下说明书中加以阐述。其它特征、目标和优点根据说明书和附图以及权利要求书将是显而易见的。

附图说明

[0014] 图1A-C展示含有IRES的复制型逆转录病毒载体和其滴度,所述IRES在A凸出物中具有各种数目的A。(A)EMCV内部核糖体进入位点的预测二级结构。这些序列始于第680位。带圈的大写字母J、K、L和M指示IRES中的限定区。箭头指示J-K区中的分支环。AUG8、AUG9、

AUG10和AUG11带下划线。(B)合并到表达yCD2或GFP的RRV中的EMCV IRES中的J-K分叉区的A凸出物的图。天然ATG8(RNA中的AUG)和ATG9带下划线;被放大且带下划线的序列表示J-K分叉区中的A凸出物;小写字母指示3' IRES中的多聚嘧啶束(polypyrimidine tract)中的5'序列。(C)在A凸出物中含有各种数目的A的RRV的病毒滴度,所述RRV由被感染的HT1080细胞产生。

[0015] 图2A-D展示通过在A凸出物中具有各种数目的A的RRV所获得的细胞病毒RNA和蛋白质表达。(A)细胞病毒RNA同种型的示意图。Env2引物和探针以及yCD2引物和探针分别识别env区和yCD2区中的未剪接和被剪接的病毒RNA,用于通过qRT-PCR测量细胞病毒RNA的水平。实心三角形:env2引物和探针组;空心三角形:yCD2引物和探针组。(B)yCD2和GAPDH蛋白质的免疫印迹。将20微克的细胞裂解物加载到各泳道上,并通过检测遍在标志物GAPDH来控制等同的加载量和印迹效率。PC,阳性对照;NC,阴性对照。图表表示关于yCD2-6A载体的RNA和蛋白质表达水平。(C)关于GFP-6A载体的RNA和GFP表达水平。使用适当的选通法排除GFP阴性细胞,通过流式细胞术确定GFP阳性细胞的百分比。通过使用平均荧光强度对GFP蛋白质表达水平进行定量。(D)通过qPCR测量的被感染的U87-MG细胞(MOI为0.01)的原病毒载体拷贝数。在感染后第14天分离基因组DNA,此时预计具有7个A的载体被最大限度地感染。数据显示被最大限度地感染的U87-MG细胞的载体拷贝数没有显著差异。这和病毒产生数据一致,其中在变体之间没有观察到对于病毒滴度的显著影响。

[0016] 图3展示具有A凸出物的载体序列(SEQ ID NO:22),所述A凸出物含有带下划线和粗体的7个A。

[0017] 图4A-B展示载体稳定性数据。(A)通过终点法PCR测量的在被感染的U87-MG细胞(MOI为0.01)中的载体稳定性。在感染后第14天分离基因组DNA,并且使用跨越env和3' UTR区的3'的引物组扩增IRES-yCD2区(Perez等人,2012)。(B)通过连续感染评定载体稳定性。最初用MOI为0.1的病毒载体感染约 10^5 个原初U87-MG细胞,并使其生长1周以完成单个感染周期。使用来自完全被感染细胞的2ml病毒上清液中的100 μ L来感染原初细胞并重复多达12个周期。通过来自被感染细胞的整合原病毒的PCR扩增来评定IRES-yCD2区的载体稳定性。预计的PCR产物大小为大约1.2kb。小于1.2kb的任何条带的出现都表明IRES-yCD2区的缺失。

[0018] 图5展示被设计成具有最小IRES(SEQ ID NO:41的碱基123-139和183到198)的本公开的构建体的图。

具体实施方式

[0019] 除非上下文另有明确规定,否则如本文和随附权利要求书中所使用的单数形式“一(a/an)”和“所述”包括复数个指示物。因此,举例来说,提及“一细胞”时包括多个此类细胞,并且提及“所述药剂”时包括提及所属领域的技术人员已知的一种或多种药剂,等等。

[0020] 另外,除非另有说明,否则“或”的使用意指“和/或”。类似地,“包含(comprise)”、“包含(comprises)”、“包含(comprising)”、“包括(include)”、“包括(includes)”和“包括(including)”是可互换的并且不旨在是限制性的。

[0021] 应进一步理解,当各种实施方案的描述使用术语“包含”时,所属领域的技术人员将理解,在一些特定情况下,可以使用措辞“基本上由.....组成”或“由.....组成”来替

代性地描述实施方案。

[0022] 除非另有定义,否则本文中使用的所有技术和科学术语的含义都与本公开所属领域的普通技术人员通常所理解的含义相同。虽然与本文中所述的方法和材料类似或等同的那些可以用于实施所公开的方法和组合物,但是本文中描述了示例性的方法、装置和材料。

[0023] 描述本文中可用的分子生物技术(包括载体、启动子的使用和许多其它相关论题)的通用教科书包括Berger和Kimmel,Guide to Molecular Cloning Techniques,Methods in Enzymology,第152卷,(Academic Press,Inc.,San Diego,Calif.)(“Berger”); Sambrook等人,Molecular Cloning--A Laboratory Manual,第2版,第1-3卷,Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y.,1989(“Sambrook”)以及Current Protocols in Molecular Biology,F.M.Ausubel等人编,Current Protocols,Green Publishing Associates,Inc.与John Wiley&Sons,Inc.的合资企业,(1999年增补)(“Ausubel”)。足以指导所属领域的技术人员通过体外扩增方法(包括聚合酶链式反应(PCR)、连接酶链式反应(LCR)、QB-复制酶扩增和其它由RNA聚合酶介导的技术(例如NASBA))例如用于产生本公开的同源核酸的方案示例见于以下文献中:Berger,Sambrook和Ausubel,以及Mullis等人(1987)美国专利第4,683,202号;Innis等人编,(1990)PCR Protocols:A Guide to Methods and Applications(Academic Press Inc.San Diego,Calif.)(“Innis”);Arnheim&Levinson(1990年10月1日)C&EN 36-47;The Journal Of NIH Research (1991)3:81-94;Kwoh等人(1989)Proc.Natl.Acad.Sci.USA 86:1173;Guatelli等人(1990)Proc.Nat'l.Acad.Sci.USA 87:1874;Lomell等人(1989)J.Clin.Chem 35:1826;Landegren等人(1988)Science 241:1077-1080;Van Brunt(1990)Biotechnology 8:291-294;Wu和Wallace(1989)Gene 4:560;Barringer等人(1990)Gene 89:117;以及Sooknanan和Malek (1995)Biotechnology 13:563-564.Wallace等人,美国专利第5,426,039号中描述了用于体外克隆被扩增核酸的改良方法.Cheng等人(1994)Nature 369:684-685和其中引用的参考文献概述了通过PCR扩增核酸的改良方法,其中产生长达40kb的PCR扩增子。所属领域的技术人员将了解,使用逆转录酶和聚合酶,可以将基本上任何RNA转化成适合于限制消化、PCR扩大和测序的双链DNA。参见例如Ausubel、Sambrook和Berger(全部见上文)。

[0024] 本文通篇中所讨论的出版物被提供仅仅因为它们在本申请的提交日期之前公开。本文中的任何内容都不应被解释为承认:发明人因为现有公开而无权占先于该公开。

[0025] 本公开提供可用于将基因或蛋白质递送到细胞或受试者的方法和组合物。此类方法和组合物可用于治疗受试者的各种疾病和病症,包括癌症和其它细胞增殖性疾病和病症。在一方面,本公开提供优化的IRES。此类优化的IRES可用在各种载体中以促进蛋白质表达。在另一方面,本公开提供用于基因递送的可复制型逆转录病毒载体。本公开证明通常使用的在J-K分叉区的A凸出物中含有7个A的IRES不是最佳的,因此本公开提供具有最佳A凸出物序列的IRES,所述IRES与具有更少(3-5个)或更多(7-8个)A的IRES相比具有提高的多肽表达。

[0026] 内部核糖体进入位点(“IRES”)是指在通常位于IRES的3'的编码序列的翻译期间促进核糖体进入或保留的核酸区段。在一些实施方案中,IRES可以包括剪接受体/供体位点,然而,优选的IRES缺少剪接受体/供体位点。通常,核糖体进入信使RNA中通过位于所有真核生物mRNA的5'端的帽发生。然而,这一普遍规律存在例外。在一些病毒mRNA中并不存在

帽,表明存在允许核糖体于这些RNA的内部位点进入的替代结构。迄今为止,已经在没有帽的病毒mRNA的5'非编码区中鉴别到许多这些结构,基于这些结构的功能,它们被命名为IRES,所述没有帽的病毒mRNA例如包括细小核糖核酸病毒如脊髓灰质炎病毒(Pelletier等人,1988,Mol.Cell.Biol.,8,1103-1112)和EMCV病毒(脑心肌炎病毒)(Jang等人,J.Virol.,1988,62,2636-2643)的mRNA。本公开提供在载体且更具体地是可复制型逆转录病毒(RCR)载体的背景下使用优化的IRES。

[0027] 内部核糖体进入位点(IRES)允许病毒RNA以不依赖于帽的方式翻译。来自脑心肌炎病毒(EMCV)的IRES已被全面研究并广泛用于逆转录病毒表达载体和其它哺乳动物表达载体中。IRES的恰当折叠和二级结构决定其功能,并且序列变化可能或可能不对此造成影响。Palmenberg和同事证明了,与5'-IRES区无关,在IRES的3'端的J-K元件在翻译起始中起关键作用(图1A)。可以发现各种载体中的IRES序列在A凸出物中含有不同数目的聚A。举例来说,Logg等人(J.Virol.75:6989-6998,2001)描述了在分叉区的A凸出物中携带有7个腺苷残基(A)而不是6个A的IRES(参见例如Duke等人,J.Virol.66:1924-1932,1992)。如本文其它地方更充分地描述,A凸出物中的A的数目影响可操作地缔合的异源序列的表达。举例来说,本公开确定A凸出物中A的最佳数目在6个A时达到峰值,并且随着在两侧从A的最佳数目偏离,表达发生轻微减少。举例来说,4个A不如6个A有效,8个A也不如6个A有效。

[0028] 如本文中所未使用,“优化的IRES”是指来源于脑心肌炎病毒的在J-K分叉区的A凸出物中具有6个A的IRES。优化的IRES可以是包含待表达的基因或序列(“异源多核苷酸”或“基因”)的盒的一部分。在这些情况下,优化的IRES可操作地连接且位于异源多核苷酸序列的上游,并且用于可操作地引起所连接异源多核苷酸的翻译。与未被优化的IRES(例如在A凸出物中具有3-5个或7-8个A的IRES)相比,优化的IRES盒展示来自所连接异源多核苷酸的增加的蛋白质表达。优化的IRES或IRES盒可被克隆到任何数目的用于表达所连接的异源多核苷酸的载体中。举例来说,可以含有本公开的优化的IRES或IRES盒并与其一起使用的载体包括质粒、表达载体、病毒载体(复制缺陷型和可复制型)等。

[0029] 在一个实施方案中,本公开提供一种优化的IRES,其包含选自以下组成的组的序列:(i)与SEQ ID NO:41具有95%同一性并且在J-K分叉区中具有6个A的序列;(ii)截短的IRES,其包含如SEQ ID NO:41中所示的序列,在分叉区中含有6个A,并且从SEQ ID NO:41的碱基对1到约碱基183之间的任何地方开始并延续到SEQ ID NO:41的544(例如SEQ ID NO:41的约123到544或约183到544),并且与在分叉区中具有7个A的类似IRES相比,具有提高的多肽表达;或(iii)如SEQ ID NO:41中所示的序列;和(iv)其中T可为U的任何前述序列(例如RNA形式)。

[0030] 异源核酸序列可操作地连接到在一个实施方案中A凸出区由6个“A”组成的优化的IRES上。如本文中所未使用,术语“异源”核酸序列或转基因是指(i)在野生型逆转录病毒中通常不存在的序列,(ii)来源于外来物种的序列,(iii)通常不在IRES的下游被发现的序列,或(iv)如果来自同一物种,那么其可以从其原始形式经过显著修饰。或者,通常不在细胞中表达的未改变的核酸序列是异源核酸序列。

[0031] 在一个实施方案中,本公开提供一种包含位于盒中的优化的IRES的载体,所述优化的IRES在J-K分叉区中包含由6个A组成的A凸出物并且可操作地连接到待表达的多核苷酸序列。如下文更详细地描述,与含有3-5个或7-8个A的类似IRES盒相比,由6个A组成的A凸

出物出乎意料地提供更好的蛋白质表达。如将要认识到的,尤其在基因递送中,来自重组载体的蛋白质表达不仅对于体外蛋白质产生很重要,而且对于体内治疗性蛋白质产生也很重要。举例来说,Logg等人(J.Virol.75:6989-6998,2001)描述在分叉区的A凸出物中携带有7个腺苷残基(A)而不是6个A的IRES。

[0032] 优化的IRES盒可以被克隆到任何数目的被所属领域认可的载体中。下文描述了此类载体,但是包括质粒和病毒载体。举例来说,本公开涵盖一种被克隆到表达载体中的本公开的优化的IRES,其中优化的IRES正好位于待表达的异源多核苷酸的上游(例如上游0到约50bp)。使用无需重组受体或辅助细胞即可在哺乳动物组织中感染和传播的可复制型 γ 逆转录病毒载体是尤其受关注的。此类RCR载体包括 γ 逆转录病毒,例如mo-MLV、MLV、GALV、FELV等。典型的 γ 逆转录病毒包含LTR、gag、pol和env基因以及逆转录和整合到宿主基因组中所必需的因子(例如 Φ 因子)。对典型 γ 逆转录病毒载体的修饰已经进行了近20年,包括产生无复制能力的载体、在不同位置携带异源基因的载体和含有IRES盒的载体。举例来说,Kasahara等人在美国专利第6,410,313号中描述了来源于MLV的可复制型逆转录病毒载体的产生,所述可复制型逆转录病毒载体携带有位于env基因下游和3'LTR上游的IRES盒。Gruber等人(美国专利第8,722,867号)描述了进一步优化的载体,其包含正好位于env基因下游和3'LTR上游的IRES盒。在Gruber等人中,IRES盒显示J-K分叉区中的7个A的A凸出物。

[0033] 在一个实施方案中,本公开提供一种可复制型 γ 逆转录病毒载体(RCR),其包含正好位于env基因下游和3'LTR上游的优化的IRES盒,其中优化IRES盒的优化IRES由分叉区中的6个A的A凸出物组成。在另一个实施方案中,与含有具有3-5个或7-8个A的A凸出物的载体相比,RCR具有增加的蛋白质表达。

[0034] 本公开提供在J-K分叉区中具有A凸出物的载体,与现有可复制型逆转录病毒载体(参见例如美国专利公开第2011/0287020-A1和2011/0217267-A1号,其显示在A凸出物中有7个A,所述美国专利公开的公开内容以引用方式并入本文)相比,所述A凸出物由6个A组成。出乎意料地,与7个A的蛋白质产生相比,单个A的改变(即7个A变为6个A)提供增加的蛋白质产生。因此,包含6个A的载体将具有与具有“6A”A凸出物的IRES盒连接的异源基因的提高了的蛋白质表达。

[0035] 术语“载体”、“载体构建体”和“表达载体”意指可借以将DNA或RNA序列(例如外来基因)引入到宿主细胞中以转化宿主并促进所引入序列的表达(例如转录和翻译)的运载体。载体典型地包含其中通过限制酶技术插入编码蛋白质的外来DNA或RNA的可传播物质(transmissible agent)的DNA或RNA。常见类型的载体是“质粒”,其通常是可以容易地接受另外的(外来)DNA并且容易被引入到合适宿主细胞中的双链DNA的自含式分子。已经描述了用于在各种真核宿主和原核宿主中复制和/或表达的大量载体,包括质粒和真菌载体。非限制性示例包括pKK质粒(Clontech)、pUC质粒、pET质粒(Novagen Inc.,Madison,Wis.)、pRSET或pREP质粒(Invitrogen,San Diego,Calif.)或pMAL质粒(New England Biolabs,Beverly,Mass.)和许多适当的宿主细胞,使用了本文中公开或引用的或者相关领域技术人员已知的方法。重组克隆载体通常将包含用于克隆或表达的一种或多种复制系统、用于在宿主中选择的一种或多种标志物(例如抗生素抗性)和一种或多种表达盒。

[0036] 术语“表达(express)”和“表达(expression)”意指允许或引起基因或DNA序列中的信息显现出来,例如通过活化相应基因、RNA或DNA序列的转录和翻译中所涉及的细胞功

能来产生蛋白质。DNA或RNA序列在细胞中表达或被细胞表达以形成“表达产物”，例如蛋白质。表达产物本身，例如得到的蛋白质，也可以称作是被细胞“表达”。多核苷酸或多肽被重组表达，例如当其在外来或天然启动子的控制下在外来宿主细胞中表达或产生时，或者其中天然宿主细胞中的天然基因在外来启动子的控制下表达。

[0037] 本公开提供改良的逆转录病毒载体。改良的逆转录病毒载体可以来源于逆转录病毒科的成员。逆转录病毒科由三个组组成：泡沫病毒(或泡沫状病毒)，例如人泡沫状病毒(HFV)；慢病毒，以及绵羊髓鞘脱落病毒(visna virus)；和致癌病毒(虽然并不是这个组内的所有病毒都具有致癌性)。术语“慢病毒”在其常规意义上用于描述含有逆转录酶的病毒的一个属。慢病毒包括“免疫缺陷病毒”，其包括人免疫缺陷病毒(HIV)1型和2型(HIV-1和HIV-2)和猿免疫缺陷病毒(SIV)。致癌病毒在历史上已经基于如在病毒成熟期间在电子显微镜下所见到的粒子形态被进一步细分为A、B、C和D组。A型粒子代表如在受感染细胞的细胞质中所见到的B型和D型病毒的不成熟粒子。这些粒子不具有感染性。B型粒子通过包裹胞质内的A型粒子而作为成熟病毒粒子从质膜上出芽。在膜上，这些粒子具有75nm的环形核心，从环形核心上伸出长的糖蛋白刺突。在出芽后，B型粒子含有偏心定位的电子致密核心。原型B型病毒是小鼠乳腺肿瘤病毒(MMTV)。在被C型病毒感染的细胞中无法观察到胞质内粒子。而是，成熟粒子通过新月形`C`形凝聚物直接从细胞表面出芽，所述新月形`C`形凝聚物然后自我闭合并且被质膜包围。包膜糖蛋白刺突与均一地电子致密核心一起可以是可见的。出芽可以从表面质膜上发生，或者直接发生到胞内空泡中。C型病毒是被最常研究的，并且包括许多禽白血病病毒和鼠白血病病毒(MLV)。牛白血病病毒(BLV)以及I型和II型人T细胞白血病病毒(HTLV-I/II)因为它们从细胞表面出芽的形态而被类似地归类为C型粒子。然而，它们还具有规则六边形形态和比原型C型病毒如鼠白血病病毒(MLV)更为复杂的基因组结构。D型粒子与B型粒子类似之处在于它们在受感染细胞的细胞质中显示为从细胞表面出芽的环状结构，但是所述病毒粒子合并了短的表面糖蛋白刺突。电子致密核心也偏心地位于粒子内。梅森-菲舍猴病毒(Mason Pfizer monkey virus, MPMV)是原型D型病毒。

[0038] 逆转录病毒已经以多种方式被分类，但是命名法在近十年里已经标准化(参见ICTVdB-The Universal Virus Database, v 4 on the World Wide Web([www](http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/))at [ncbi.nlm.nih.gov/ICTVdb/ICTVdB/](http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/)，和教科书“Retroviruses”Coffin, Hughs和Varmus编，Cold Spring Harbor Press 1997；其公开内容以引用方式并入本文)。在一个实施方案中，可复制型逆转录病毒载体可以包括正逆转录病毒或更典型地 γ 逆转录病毒载体。

[0039] 逆转录病毒是以它们复制其遗传物质的方式被定义的。在复制期间，RNA转化为DNA。在感染细胞之后，通过被称为逆转录的分子过程，从病毒粒子中所携带的两个RNA分子产生一个DNA双链分子。DNA形式作为原病毒共价整合于宿主细胞基因组中，病毒RNA借助于细胞因子和/或病毒因子，从所述宿主细胞基因组表达。表达的病毒RNA被包装成粒子并作为感染性病毒粒子释放。

[0040] 逆转录病毒粒子由两个相同的RNA分子构成。各野生型基因组都具有正义单链RNA分子，其在5'端被加帽并且在3'尾部被多聚腺苷酸化。二倍体病毒粒子含有两条RNA链，这两条RNA链与gag蛋白质、病毒酶(pol基因产物)和gag蛋白质的‘核心’结构内的宿主tRNA分子复合在一起。来源于宿主细胞膜且含有病毒包膜(env)蛋白质的脂质双层包围并保护该衣壳。env蛋白质结合到病毒的细胞受体上，并且所述粒子典型地通过受体介导的胞吞作用

和/或膜融合而进入宿主细胞中。

[0041] 在外层包膜脱落后,病毒RNA通过逆转录被拷贝成DNA。这是由受pol区编码的逆转录酶催化的,并且使用被包装入病毒粒子中的宿主细胞tRNA作为DNA合成的引物。因此,RNA基因组被转化为更为复杂的DNA基因组。

[0042] 通过逆转录产生的双链线性DNA可能必须或者可能不必在核中环化。原病毒此时在任一端都具有两个相同的重复,称为长末端重复(LTR)。这两个LTR序列的末端产生被pol产物(整合酶蛋白质,其催化整合)识别的位点,使得原病毒总是接合到距离LTR两端2个碱基对(bp)的宿主DNA上。在两个LTR的两端都见到细胞序列的复制,这使人想起可转座的遗传元件的整合模式。逆转录病毒可以在宿主DNA中的许多位点处整合其DNA,但是不同的逆转录病毒具有不同的整合位点偏好。HIV-1和猿免疫缺陷病毒DNA优先整合到所表达的基因中,鼠白血病病毒(MLV)DNA优先接近转录起始位点(TSS)整合,并且禽肉瘤造血白细胞组织增生病毒(ASLV)和人T细胞白血病病毒(HTLV)DNA几乎随机地整合,仅显示对基因的轻微偏好(Derse D等人(2007)Human T-cell leukemia virus type 1 integration target sites in the human genome:comparison with those of other retroviruses.J Virol 81: 6731-6741;Lewinski MK等人(2006)Retroviral DNA integration:viral and cellular determinants of target-site selection.PLoS Pathog 2:e601)。

[0043] 宿主细胞蛋白质介导所整合的病毒DNA的转录、RNA剪接和翻译。产生各种剪接的转录物。在人逆转录病毒HIV-1/2和HTLV-I/II的情况下,还使用病毒蛋白质来调控基因表达。细胞因子与病毒因子之间的相互影响是在其中表达病毒基因的病毒潜伏和时间序列的控制中的一个因素。

[0044] 逆转录病毒可以水平地和垂直地传播。逆转录病毒有效的感染性传播需要在特异性地识别病毒包膜蛋白质的受体的靶细胞上表达,虽然病毒可以使用低效率的不依赖于受体的非特异性进入途径。通常,病毒感染在每个细胞中产生单个或几个病毒基因组拷贝,因为存在受体掩蔽或下调,这进而产生对重复感染的抗性(“Retroviruses”,JM Coffin,SH Hughes和HE Varmus 1997Cold Spring Harbor Laboratory Press,Cold Spring Harbor NY中的第3章第104页;Fan等人J.Virol 28:802,1978)。此外,靶细胞类型必须能够在病毒已经结合和穿透之后支持复制周期的所有阶段。当病毒基因组变得整合于宿主的生殖系中时,发生垂直传播。原病毒于是会像细胞基因一样一代一代地传承。因此,形成了内源性原病毒,其经常是潜伏的,但是当宿主暴露于适当因素时,可以变得活化。

[0045] 在在治疗上使用重组可复制型逆转录病毒的许多情形下,有利的是通过重组可复制型逆转录病毒编码的转基因具有高表达水平。举例来说,对于前药活化基因如胞嘧啶脱氨酶基因来说,有利的是细胞中的CD蛋白质具有较高的表达水平,使得前药5-FU更有效地转化为5-FU。类似地,siRNA或shRNA的高表达水平导致靶基因表达被更有效地抑制。同样,对于细胞因子或单链抗体(scAb)来说,表达高水平的细胞因子或scAb通常也是有利的。此外,当在一些载体拷贝中存在使载体或转基因的活性失活或削弱活性的突变时,有利的是在靶细胞中具有多个载体拷贝,因为这提供了完整转基因的有效表达的高可能性。本公开提供重组可复制型逆转录病毒,其能够多次感染靶细胞或靶细胞群体,导致拷贝/二倍体基因组的平均数为5或更大。本公开还提供测试这种性质的方法。还提供了使用重组可复制型逆转录病毒治疗细胞增殖性病症的方法,所述重组可复制型逆转录病毒能够多次感染靶细

胞或靶细胞群体,导致拷贝/二倍体基因组的平均数为5或更大。

[0046] 如上所述,经过整合的DNA中间物被称为原病毒。现有基因疗法或基因递送系统使用这样的方法和逆转录病毒,其需要原病毒在存在适当辅助病毒时或者在含有适当序列的细胞系中转录并装配成感染性病毒,所述适当序列能够实现衣壳化同时不产生污染性的辅助病毒。如下所述,辅助病毒不是产生本公开的重组逆转录病毒所必需的,因为在基因组中提供了用于衣壳化的序列,从而提供了用于基因递送或治疗的可复制型逆转录病毒载体。

[0047] 其它现有的可复制型逆转录病毒载体也往往是不稳定的并且在水平传播或垂直传播到被感染细胞或宿主细胞期间和在复制期间丢失序列。这可能部分地归因于存在包括重复序列或降低聚合酶效率的额外核苷酸序列。

[0048] 本公开的逆转录病毒基因组和原病毒DNA具有至少3种基因:gag、pol和env,这些基因可以侧接有一个或两个长末端(LTR)重复单元,或者在原病毒中,这些基因侧接有两个长末端重复序列(LTR)和含有顺式作用序列如psi的序列。gag基因编码内部结构(基质、衣壳和核衣壳)蛋白质;pol基因编码RNA指导的DNA聚合酶(逆转录酶)、蛋白酶和整合酶;并且env基因编码病毒包膜糖蛋白。5'和/或3'LTR用于促进病毒粒子RNA的转录和多聚腺苷酸化。LTR含有病毒复制所必需的所有其它顺式作用序列。慢病毒具有额外的基因,包括vif、vpr、tat、rev、vpu、nef和vpx(在HIV-1、HIV-2和/或SIV中)。

[0049] 邻近5'LTR的是基因组的逆转录所必需的序列(tRNA引物结合位点)和病毒RNA有效衣壳化成粒子所必需的序列(Psi位点)。如果衣壳化(或逆转录病毒RNA包装成感染性病毒粒子)所必需的序列从病毒基因组上丢失,则会导致顺式缺陷,从而阻止基因组病毒RNA的衣壳化。这种类型的改良载体是典型地已经在现有基因递送系统(即,缺乏病毒粒子的衣壳化所必需的元件的系统)中用作‘辅助’元件的载体,其提供包装非复制性但可包装的RNA基因组的反式病毒蛋白质。

[0050] 本公开提供含有优化的IRES的载体。优化的IRES典型地连接到可以递送给细胞或受试者的异源多核苷酸,所述异源多核苷酸例如编码胞嘧啶脱氨酶或其突变体、胸苷激酶或其突变体、miRNA或siRNA、细胞因子、抗体结合域等。在一个实施方案中,载体是病毒载体。病毒载体可以是腺病毒载体、麻疹载体、疱疹载体、逆转录病毒载体(包括慢病毒载体)、弹状病毒载体如疱疹性口炎病毒载体、呼肠孤病毒载体、塞尼卡谷病毒(Seneca Valley Virus)载体、痘病毒载体(包括来源于动物痘或牛痘的载体)、细小病毒载体(包括AAV载体)、甲病毒载体或所属领域技术人员已知的其它病毒载体(还参见例如Concepts in Genetic Medicine,Boro Dropulic和Barrie Carter编,Wiley,2008,Hoboken,NJ.;The Development of Human Gene Therapy,Theodore Friedmann编,Cold Springs Harbor Laboratory Press,Cold springs Harbor,New York,1999;Gene and Cell Therapy,Nancy Smyth Templeton编,Marcel Dekker Inc.,New York,New York,2000和Gene Therapy:Therapeutic Mechanism and Strategies,Nancy Smyth Templeton和Danilo D Lasic编,Marcel Dekker,Inc.,New York,New York,2000;其公开内容以引用方式并入本文)。

[0051] 在一个实施方案中,本公开的逆转录病毒基因组含有优化的IRES,所述优化的IRES在其下游包含克隆位点,用于插入期望的/异源的多核苷酸。在一个实施方案中,优化的IRES位于逆转录病毒载体中env基因的3',但是位于期望的异源多核苷酸的5'和3'LTR的

5'。在所有前述实施方案中,优化的IRES包含具有6个A的A凸出物。编码期望多肽的异源多核苷酸可以可操作地连接到优化的IRES。

[0052] 在一个实施方案中,病毒载体可以是得自或来源于能够感染复制中的哺乳动物细胞的 γ 逆转录病毒的可复制型逆转录病毒载体。可复制型逆转录病毒载体包含优化的内部核糖体进入位点(IRES),所述IRES包含由6个A组成的A凸出物,所述A凸出物位于例如编码胞嘧啶脱氨酶(SEQ ID NO:3)、胸苷激酶(SEQ ID NO:37)、miRNA、siRNA、细胞因子、受体、抗体等的异源多核苷酸的5'。当异源多核苷酸编码非翻译的RNA如siRNA、miRNA或RNAi时,IRES不是必需的,但是可以被包括以用于另一个翻译的多核苷酸。在一个实施方案中,含有异源多核苷酸的优化的IRES盒位于逆转录病毒载体的ENV多核苷酸的3',但位于3'LTR的5'。在一个实施方案中,病毒载体是能够多次(例如每个二倍体细胞5次或更多次)感染靶细胞的逆转录病毒载体。

[0053] 本公开提供相对于现有逆转录病毒载体具有增加的稳定性且含有优化的IRES的可复制型逆转录病毒载体,所述优化的IRES在A凸出物中具有6个A。在感染和复制期间此类增加的稳定性对于细胞增殖性病症的治疗至关重要。此外,源于优化的A凸出物的增加的蛋白质表达向靶细胞/组织提供了治疗蛋白质的另外递送。可复制型逆转录病毒提供了转导效率、转基因稳定性、转基因表达和靶标选择性的组合。组合物和方法提供插入稳定性并且保持转基因的转录活性和所编码多肽的翻译活力。

[0054] 取决于本公开的载体或逆转录病毒载体的期望用途,可以将任何数目的异源多核苷酸或核酸序列插入到载体或逆转录病毒载体中。举例来说,对于体外研究,可以使用常用的标志基因或报告基因,包括抗生素抗性和荧光分子(例如GFP)。还可以将编码任何期望多肽序列的额外的多核苷酸序列插入到本公开的载体中。当探寻异源核酸序列的体内递送时,可以使用治疗序列和非治疗序列两者。举例来说,异源序列可以编码治疗分子,包括针对与细胞增殖性病症或其它基因相关疾病或病症有关的特定基因的反义分子(miRNA、siRNA)或核酶,异源序列可以是自杀基因(例如HSV-tk或PNP或胞嘧啶脱氨酶;被修饰的或未被修饰的,人源化或非人源化)、生长因子或治疗蛋白质(例如因子IX、IL2等)。在所属领域中容易鉴别适用于本公开的其它治疗蛋白质。

[0055] 在一个实施方案中,载体内的异源多核苷酸包含已经过优化用于在人细胞中表达的胞嘧啶脱氨酶。在另一个实施方案中,胞嘧啶脱氨酶包含已经过人密码子优化的序列,并且包含与野生型胞嘧啶脱氨酶(参见例如SEQ ID NO:4)相比增加胞嘧啶脱氨酶稳定性(例如减少降解或增加热稳定性)的突变。在另一个实施方案中,异源多核苷酸编码融合构建体(参见例如SEQ ID NO:11、13、15和17),所述融合构建体包含与编码具有UPRT或OPRT活性的多肽的多核苷酸可操作地连接的胞嘧啶脱氨酶(经过人密码子优化或非优化的,突变的或非突变的)。此类具有胞嘧啶脱氨酶的多肽和编码此类多肽的多核苷酸的示例可见于国际公开第2010/045002号,其以引用方式并入本文。

[0056] 在另一个实施方案中,载体或可复制型逆转录病毒载体可以包含编码胞嘧啶脱氨酶(如本文中所述)的多肽的异源多核苷酸,并且载体或可复制型逆转录病毒载体可以进一步包含含有miRNA或siRNA分子的多核苷酸,所述miRNA或siRNA分子或者作为来自病毒启动子的初级转录物的一部分,或者连接到可以是细胞型或组织特异性的启动子上。

[0057] 在其它实施方案中,异源多核苷酸可以包含细胞因子,例如白细胞介素、 γ 干扰素

等。可以从本公开的逆转录病毒载体表达的细胞因子包括(但不限于)IL-1 α 、IL-1 β 、IL-2 (SEQ ID NO:40)、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8、IL-9、IL-10、IL-11、IL-12、IL-13、IL-14、IL-15、IL-16、IL-17、IL-18、IL-19、IL-20和IL-21、抗-CD40、CD40L、IFN- γ (人-SEQ ID NO:38;小鼠-SEQ ID NO:39)和TNF- α 、TNF- α 的可溶形式、淋巴毒素- α (LT- α ,也称为TNF- β)、LT- β (在复杂的异质三聚体LT- α 2- β 中发现)、OPGL、FasL、CD27L、CD30L、CD40L、4-1BBL、DcR3、OX40L、TNF- γ (国际公开第W0 96/14328号)、AIM-I(国际公开第W0 97/33899号)、内分泌因子- α (国际公开第W0 98/07880号)、OPG、和neutrokine- α (国际公开第W0 98/18921号)、OX40、和神经生长因子(NGF)、和Fas、CD30、CD27、CD40和4-1BB的可溶形式、TR2(国际公开第W0 96/34095号)、DR3(国际公开第W0 97/33904号)、DR4(国际公开第W0 98/32856号)、TR5(国际公开第W0 98/30693号)、TRANK、TR9(国际公开第W0 98/56892号)、TR10(国际公开第W0 98/54202号)、312C2(国际公开第W0 98/06842号)、和TR12、和CD154、CD70和CD153的可溶形式。在一些实施方案中,血管生成蛋白质可以尤其用于从细胞系产生蛋白质。此类血管生成因子包括(但不限于)神经胶质瘤源性生长因子(GDNF)、血小板源性生长因子-A(PDGF-A)、血小板源性生长因子-B(PDGF-B)、胎盘生长因子(PIGF)、胎盘生长因子-2(PIGF-2)、血管内皮生长因子(VEGF)、血管内皮生长因子-A(VEGF-A)、血管内皮生长因子-2(VEGF-2)、血管内皮生长因子B(VEGF-3)、血管内皮生长因子B-186(VEGF-B186)、血管内皮生长因子-D(VEGF-D)、血管内皮生长因子-D(VEGF-D)和血管内皮生长因子-E(VEGF-E)。成纤维细胞生长因子可以通过本公开的载体递送并且包括(但不限于)FGF-1、FGF-2、FGF-3、FGF-4、FGF-5、FGF-6、FGF-7、FGF-8、FGF-9、FGF-10、FGF-11、FGF-12、FGF-13、FGF-14和FGF-15。造血生长因子可以使用本公开的载体递送,此类生长因子包括(但不限于)粒细胞巨噬细胞集落刺激因子(GM-CSF)(沙格司亭(sargramostim))、粒细胞集落刺激因子(G-CSF)(非格司亭(filgrastim))、巨噬细胞集落刺激因子(M-CSF、CSF-1)、红细胞生成素(阿法依伯汀(epoetin alfa))、干细胞因子(SCF、c-kit配体、青灰因子(steel factor))、巨核细胞集落刺激因子、PIXY321(GMCSF/IL-3)融合蛋白质等。

[0058] 微RNA(miRNA)是非编码小RNA。它们位于编码基因或非编码基因的内含子内、非编码基因的外显子内或基因间区内。miRNA基因由RNA聚合酶II转录,所述RNA聚合酶II产生被称为初级前体miRNA(pri-miRNA)的前体多核苷酸。核糖核酸酶Drosha加工核内的pri-miRNA以产生miRNA前体(pre-miRNA),所述miRNA前体形成短的发夹结构。随后,pre-miRNA通过输出蛋白(Exportin)5运输到细胞质中,并且通过被称为切酶(Dicer)的另一种核糖核酸酶进一步加工以产生活性的成熟miRNA。

[0059] 成熟miRNA的长度为大约21个核苷酸。它通过结合到被靶向基因的mRNA的3'非翻译区并通过阻遏蛋白质翻译或降解mRNA而抑制蛋白质表达来发挥功能。在包括发育、细胞增殖、分化和癌症发展的生物过程中涉及miRNA。miRNA制谱(profiling)研究指示一些miRNA表达是组织特异性的或者在某些组织中富集。举例来说,已经证明miR-142-3p、miR-181和miR-223表达在人和小鼠的造血组织中富集(Baskerville等人,2005RNA 11,241-247;Chen等人,2004Science 303,83-86)。miR-142-3p的靶序列显示在SEQ ID NO:35中。miR-142-3p4X的靶标显示在SEQ ID NO:36中。

[0060] 已经观察到一些miRNA在若干肿瘤中上调(致癌miRNA)或下调(阻遏物)。(Spizzo等人,2009Cell 137,586e1)。举例来说,miR-21在胶质母细胞瘤、乳腺癌、肺癌、前列腺癌、

结肠癌、胃癌、食道癌和子宫颈癌、子宫平滑肌肉瘤、DLBCL、头颈癌中过度表达。相比之下，已经报道了let-7的成员在胶质母细胞瘤、肺癌、乳腺癌、胃癌、卵巢癌、前列腺癌和结肠癌中下调。癌症中的miRNA表达的内稳态的重建是抑制或逆转癌症发展的必要机制。

[0061] 因为miRNA在癌症中调节的重要功能，所以研发潜在治疗方法的焦点已经集中在致癌miRNA的反义介导性抑制(antigomer)。然而，miRNA替换可能代表同样有效的策略。在这种方法中，治疗上最有用的miRNA是在肿瘤中以低水平表达，但在正常组织中以高水平表达且因此可被耐受的miRNA。

[0062] 在癌症中下调的miRNA可用作抗癌剂。示例包括mir-128-1/2(分别是SEQ ID NO: 31和32)、let-7、miR-26、miR-124和miR-137(Esquelea-Kerscher等人,2008Cell Cycle 7, 759-764;Kumar等人,2008Proc Natl Acad Sci USA 105,3903-3908;Kota等人,2009Cell 137,1005-1017;Silber等人,2008BMC Medicine 6:14 1-17)。已经报道了miR-128表达在中枢神经系统中富集，并且已经观察到miR-128在胶质母细胞瘤中下调(Sempere等人,2004Genome Biology 5:R13.5-11;Godlewski等人,2008Cancer Res 68:(22)9125-9130)。miR-128由两种不同的基因miR-128-1和miR-128-2编码。这两种基因被加工成相同的成熟序列。已经报道Bmi-1和E2F3a是miR-128的直接靶标(Godlewski等人,2008Cancer Res 68:(22)9125-9130;Zhang等人,2009J. Mol Med 87:43-51)。此外，已经观察到Bmi-1表达在各种人癌症中上调，所述人癌症包括神经胶质瘤、套细胞淋巴瘤、非小细胞肺癌、B细胞非霍奇金淋巴瘤(B-cell non-Hodgkin's lymphoma)、乳腺癌、结肠直肠癌和前列腺癌。此外，已经证明Bmi-1是来自不同组织的干细胞自我更新所必需的，所述干细胞包括神经元干细胞以及神经胶质瘤中的“干细胞样”细胞群体。

[0063] 虽然由miRNA介导的细胞功能抑制的可能性已经有了许多体外证明，但是仍然难以像其它分子那样以产生体内作用所必需的效率将这些miRNA作为寡聚核苷酸或在病毒载体中递送(例如Li等人,Cell Cycle 5:2103-2109 2006)。

[0064] 复制缺陷型逆转录病毒和慢病毒载体已经被用于通过聚合酶II启动子如CMV或LTR来稳定地表达pri-miRNA，并且证明可产生成熟miRNA。在非复制型逆转录病毒和慢病毒载体中并入III型RNA聚合酶III启动子(例如U6和H1启动子)已被广泛用于表达产生短的发夹结构RNA的功能性小干扰RNA(siRNA)(Bromberg-White等人,2004J Virol 78:9,4914-4916;Sliva等人,2006Virology 351,218-225;Haga等人,2006,Transplant Proc 38(10):3184-8)。用切酶裂解环序列，产生长度为21-22个核苷酸的成熟siRNA。shRNA可以在细胞中稳定表达以下调靶基因表达。SEQ ID NO:33和34包含连接到H1启动子的pre-miR-128。

[0065] 在另一个实施方案中，在A凸出物中包含6个A的优化的IRES可以与核心启动子组合使用，其中优化的IRES可操作地连接到第一异源编码序列，并且核心启动子或迷你启动子连接到第二异源编码序列或siRNA、miRNA或shRNA序列(参见例如WO 2014/066700，其以引用方式并入本文)。

[0066] 如本文中所使用，“核心启动子”是指包含约50-100bp的最小启动子且缺少增强子元件。此类核心启动子包括(但不限于)SCP1、AdML和CMV核心启动子。更具体地说，当存在核心启动子盒时，将会存在第二盒(例如第二迷你启动子盒、p011111启动子盒或IRES盒)。在一些实施方案中，包含具有核心启动子的盒的载体明确排除使用SCP1、AdML和CMV核心启动子，而是利用如此处和下文进一步描述的设计核心启动子。

[0067] 核心启动子包括某些病毒启动子。如本文中所使用的病毒启动子是具有核心序列而且还通常具有一些其它辅助元件的启动子。举例来说,用于SV40的早期启动子含有三种类型的元件:TATA框、起始位点和GC重复序列(Barrera-Saldana等人,EMBO J,4:3839-3849,1985;Yaniv,Virology,384:369-374,2009)。TATA框位于转录起始位点上游大约20个碱基对处。GC重复区是含有6个GC框的21碱基对重复序列并且是决定转录方向的位点。该核心启动子序列为大约100bp。添加额外的72碱基对重复序列从而使其成为“迷你启动子”,可用作转录增强子,其使启动子的功能性增加到约10倍。当SP1蛋白质与21bp重复序列相互作用时,其结合前3个或后3个GC框。前3个GC框的结合起始早期表达,而后3个GC框的结合起始晚期表达。72bp重复序列的功能是提高稳定RNA的量并增加合成速率。这是通过与AP1(活化蛋白质1)结合(二聚)从而得到被3'多聚腺苷酸化和5'加帽的初级转录物而实现的。其它病毒启动子(例如劳氏肉瘤病毒(RSV)、HBV X基因启动子和疱疹胸苷激酶核心启动子)也可以用作选择期望功能的基础。

[0068] 核心启动子典型地涵盖相对于+1转录起始位点的-40到+40(Juven-Gershon和Kadonaga,Dev.Biol.339:225-229,2010),所述+1转录起始位点定义RNA聚合酶II机器起始转录所处的位置。典型地,RNA聚合酶II与结合到启动子中的DNA基序上的许多转录因子相互作用。这些因子通常被称为“一般”或“基础”转录因子并且包括(但不限于)TFIIA(RNA聚合酶IIA的转录因子)、TFIIB、TFIID、TFIIE、TFIIF和TFIIH。这些因子以“一般”方式与所有核心启动子一起作用;因此它们通常被称为“基础”转录因子。

[0069] Juven-Gershon等人(Nat.Methods,3(11):917-922,2006)描述了核心启动子的元件。举例来说,pRC/CMV核心启动子由pRC/CMV框组成并且长度为81bp;CMV核心启动子由TATA框和起始位点组成;而SCP合成核心启动子(SCP1和SCP2)由TATA框、Inr(起始物)、MTE位点(十基序元件)和DPE位点(下游启动子元件)组成并且长度为约81bp。与简单的pRC/CMV核心启动子相比,SCP合成启动子具有提高的表达。

[0070] 如本文中所使用,“迷你启动子”或“小启动子”是指促进可操作地连接的基因或编码核酸序列的转录的调控结构域。顾名思义,迷你启动子包括可操作地连接的编码序列的有效转录和/或翻译所必需的最小量的元件。迷你启动子可以包含“核心启动子”与额外调控元件的组合或“改良的核心启动子”。典型地,迷你启动子或改良的核心启动子的长度将为约100-600bp,而核心启动子典型地小于约100bp(例如约70-80bp)。在其它实施方案中,当存在核心启动子时,盒将典型地包含位于核心启动子序列上游或下游的增强子元件或另一元件,所述增强子元件或另一元件促进可操作地连接的编码序列的表达超过单独核心启动子的表达水平。

[0071] 因此,本公开提供来源于针对“核心启动子”元件所确定的细胞元件的迷你启动子(例如经过修饰的核心启动子)($<100\text{bp}$ 、 $<200\text{bp}$ 、 $<400\text{bp}$ 或 $<600\text{bp}$),所述迷你启动子允许在靶细胞中以显著水平遍在表达,并且一般地说可用于稳定地合并到载体中且具体地说可用于稳定地合并到复制型逆转录病毒载体中,以允许转基因的有效表达。还提供包含核心启动子加最小增强子序列和/或Kozak序列的仍然小于200bp、400bp或600bp的迷你启动子,与缺少此类序列的核心启动子相比,所述迷你启动子允许更好的基因表达。此类迷你启动子包括改良的核心启动子和天然存在的组织特异性启动子,例如弹性蛋白启动子(胰腺腺泡细胞所特有的,204bp;Hammer等人,Mol Cell Biol.,7:2956-2967,1987)以及来自小鼠和

人的细胞周期依赖性ASK基因的启动子(63-380bp;Yamada等人,J.Biol.Chem.,277:27668-27681,2002)。遍在表达的小启动子还包括病毒启动子,例如SV40早期和晚期启动子(约340bp)、RSV LTR启动子(约270bp)和HBV X基因启动子(约180bp)(例如R Anish等人,PLoS One,4:5103,2009),所述HBV X基因启动子没有正则“TATAA框”,而具有5'-CCCCGTTGCCCGG-3'(SEQ ID NO:43)的13bp核心序列。在其它实施方案中,包含至少一个迷你启动子盒的治疗盒的表达水平将超过RRV中存在的IRES盒的表达水平、与RRV中存在的IRES盒的表达水平大致相等、或者是RRV中存在的IRES盒的表达水平的约2/5到1/1。

[0072] 从核心启动子或迷你启动子的转录通过各种元件的相互作用发生。在集中转录中,举例来说,存在单个主要转录起始位点或者在几个核苷酸的狭窄区域内的几个起始位点。集中转录是较简单的生物体中的主要转录模式。在分散转录中,在约50到100个核苷酸的宽阔区域内存在几个弱转录起始位点。分散转录是脊椎动物中最常见的转录模式。举例来说,在约三分之二的人基因中观察到分散转录。在脊椎动物中,集中转录往往与受调控的启动子有关,而分散转录则典型地在CpG岛的组成型启动子中被观察到。

[0073] 表1:可以促成集中核心启动子(几乎总是具有“TATA框和单个转录起始位点(TSS)”)或不含TATA框但通常具有DPE元件的分散启动子的结合位点(参见Dickstein,Trascription,2(5):201-206,2011;Juven-Gershon等人,Nat.Methods,2006,见上文)。用于核苷酸的符号遵循国际惯例(万维网:chem.qmul.ac.uk/iubmb/misc/naseq.html)。

[0074]

转录因子	全名	相对于转录起始位点 (TSS +1) 的结合位点
BREu	TFIIB识别元件, 上游	TATA框的上游, SSRCGCC
TATA box	TATA框	处于-31/-30 TATAWAAR的T, 关键的集中启动子元件
BREd	TFIIB识别元件, 下游	-23到-17 RTDKKKK
XCPE1	HBV X核心启动子元件1	来自 HBV X 基因的 -8 到 +2 DSGYGGRASM
XCPE2	HBV X核心启动子元件2	来自HBV X基因的VCYCRTRTCMY
Inr	起始物	-2到+4 YYANWYY
DCE SI	下游核心元件位点 I	+6到+11 CTTC
DCE SII	下游核心元件位点 II	+16到+21 CTGT
DCE SIII	下游核心元件位点 III	+30到+34 AGC
MTE	十基序元件	+18到+27 CSARCSSAAC, 大部分位于果蝇中
DPE	下游启动子元件	+28到+33 RGWYVT, 常见于果蝇中, 关键的分散启动子元件

[0075] 表2列举可用于构建增强子元件并将其克隆到核心启动子区中的寡核苷酸。如上所述, 本公开的改良/优化的核心启动子可以包括添加有表1的元件的核心序列, 并且可以进一步包括如表2中所列举而被克隆的增强子。在这种情况下, 可以增加迷你启动子的大小。然而, 最终的迷你启动子不应超过600bp并且将典型地为约100bp、200bp、300bp、400bp、500bp和其间的任何整数。

[0076] 表2. 用于构建增强子区段的寡核苷酸

[0077]

编号	寡核苷酸	基序序列	参考文献
1	AP-1	5'-TGTCTCAG-3'	Hallahan 等人 Int. J. Radiat. Oncol. Biol. Phys. 36:355-360 1996.
2	CArG	5'-CCATATAAGG-3'	Datta 等人 Proc. Natl. Acad. Sci. USA 89:10149-10153. 1992
3	NF-κB1	5'-GGAAATCCCC-3'	Ueda 等人 FEBS Lett. 491:40-44 2001
4	NF-κB2	5'-GGAAAGTCCC C-3'	Kanno 等人 EMBO J. 8:4205-4214 1989
5	NF-κB3	5'-GGAGTTCCC-3'	Hong 等人 J. Biol. Chem. 275:18022-18028 2000.
6	NF-Y	5'-CATTGGG-3'	Hu 等人 J. Biol. Chem. 275:2979-2985 2000.

[0078] AP-1, 活化蛋白质-1; NF-κB, 核因子κB。

[0079] 在一个实施方案中, 本公开提供一种能够感染非分裂的宿主细胞、宿主分裂细胞或具有细胞增殖性病症的宿主细胞的重组可复制型逆转录病毒。本公开的重组可复制型逆转录病毒包含编码病毒GAG、病毒POL、病毒ENV的多核苷酸序列、封装在病毒粒子中的优化的内部核糖体进入位点(IRES)之后的异源多核苷酸, 所述IRES在IRES的A凸出物中具有6个A。

[0080] 通常, 本公开的重组载体能够将核酸序列转移到靶细胞中。短语“非分裂”细胞是指未经历有丝分裂的细胞。非分裂细胞可以在细胞周期中的任何点(例如G₀/G₁、G₁/S、G₂/M)被阻断, 只要细胞未在活跃地分裂即可。对于离体感染, 可以通过所属领域的技术人员所使用的标准技术处理分裂细胞以阻断细胞分裂, 所述标准技术包括照射、阿非迪霉素(aphidocolin)处理、血清饥饿和接触抑制。然而, 应理解离体感染通常在不阻断细胞的情况下进行, 因为许多细胞已经被遏制(例如干细胞)。举例来说, 重组慢病毒载体能够感染非分裂细胞。体内预先存在的非分裂细胞的示例包括神经元细胞、肌肉细胞、肝细胞、皮肤细胞、心脏细胞、肺细胞和骨髓细胞以及它们的衍生物。对于分裂细胞, 可以使用致癌逆转录病毒载体。

[0081] “分裂”细胞意指经历活跃的有丝分裂或减数分裂的细胞。此类分裂细胞包括干细胞、皮肤细胞(例如成纤维细胞和角化细胞)、配子细胞和所属领域中已知的其它分裂细胞。术语分裂细胞尤其关注且涵盖具有细胞增殖性病症的细胞, 例如赘生性细胞。术语“细胞增殖性病症”是指特征为细胞数目异常的病况。所述病况可以包括肥大(细胞连续繁殖, 导致组织内细胞群体的过度生长)细胞生长和营养不良型(hypotrophic)(组织内细胞缺乏或不足)细胞生长, 或细胞过度流入或迁移入身体的某个区域中。细胞群体不一定是被转化的细胞、致瘤细胞或恶性细胞, 也可以包括正常细胞。细胞增殖性病症包括与结缔组织的过度生

长有关的病症,例如各种纤维化病症,包括硬皮病、关节炎和肝硬化。细胞增殖性病症包括增生性病症,例如头颈癌。头颈癌将包括例如口癌、食道癌、咽喉癌、喉癌、甲状腺癌、舌癌、唇癌、唾液腺癌、鼻癌、副鼻窦癌、鼻咽癌、上鼻穹窿和鼻窦肿瘤(superior nasal vault and sinus tumors)、成感觉神经细胞瘤、鳞状细胞癌、恶性黑素瘤、鼻腔鼻窦未分化癌(SNUC)、脑瘤(包括胶质母细胞瘤)或血液瘤。还包括局部淋巴结癌,所述局部淋巴结包括颈淋巴结、喉前淋巴结、肺食管旁淋巴结和下颌下淋巴结(Harrison's Principles of Internal Medicine, Isselbacher等人编, McGraw-Hill, Inc., 第13版, 第1850-1853页, 1994)。其它癌症类型包括(但不限于)肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、淋巴瘤、口腔癌、胰腺癌、白血病、黑素瘤、胃癌、皮肤癌和卵巢癌。细胞增殖性疾病还包括类风湿性关节炎(O'Dell NEJM 350:2591 2004)和特征通常为免疫系统细胞不当增殖的其它自身免疫病症(Mackay等人NEJM 345:340 2001)。

[0082] 在其它实施方案中,提供用本公开的可复制型逆转录病毒载体转染的宿主细胞。宿主细胞包括真核细胞,例如酵母细胞、昆虫细胞或动物细胞。宿主细胞还包括原核细胞,例如细菌细胞。在其它实施方案中,宿主细胞已经过修饰或选择以在无血清悬浮液中连续生长(参见例如美国专利公开第2012/0087894-A1号,其以引用方式并入本文)。

[0083] 还提供了用本文中所提供的载体(例如可复制型逆转录病毒载体)转导(转化或转染)的经工程改造的宿主细胞。经工程改造的宿主细胞可以在被适当改良以用于活化启动子、选择转化体或扩增编码多核苷酸的常规营养培养基中培养。例如温度、pH等培养条件是先前被选择用于表达的宿主细胞所用的那些条件,并且对于所属领域的技术人员将是显而易见的且在本文引用的参考文献中,所述参考文献包括例如Sambrook、Ausubel和Berger,以及例如Freshney(1994) Culture of Animal Cells: A Manual of Basic Technique, 第3版(Wiley-Liss, New York)和其中引用的参考文献。

[0084] 适当表达宿主的示例包括哺乳动物细胞,例如CHO、COS、BHK、HEK 293或Bowes黑素瘤等。典型地,将使用人细胞或细胞系;然而,可能期望将本公开的载体和多核苷酸克隆到非人宿主细胞中以用于测序、扩增和克隆目的。

[0085] 在另一个实施方案中,包括靶向多核苷酸序列作为本公开的重组逆转录病毒载体的一部分。靶向多核苷酸序列是靶向配体(例如肽激素,如神经调节蛋白(heregulin)、单链抗体、受体或受体的配体)、组织特异性或细胞类型特异性调控元件(例如组织特异性或细胞类型特异性启动子或增强子)、或靶向配体与组织特异性/细胞类型特异性调控元件的组合。靶向配体可操作地连接到逆转录病毒的env蛋白质上,从而产生嵌合逆转录病毒env蛋白质。病毒GAG、病毒POL和病毒ENV蛋白质可以来源于任何合适的逆转录病毒(例如MLV或慢病毒源性的)。在另一个实施方案中,病毒ENV蛋白质是非逆转录病毒源性的(例如CMV或VSV)。

[0086] 在一个实施方案中,逆转录病毒载体通过结合到在细胞外表面上具有分子的细胞而靶向细胞。这种靶向逆转录病毒的方法利用逆转录病毒包膜上的靶向配体的表达来帮助病毒靶向于细胞或组织,所述细胞或组织具有与逆转录病毒表面上的靶向配体相互作用的受体或结合分子。在病毒感染细胞之后,病毒将其核酸注入细胞中并且逆转录病毒遗传物质可以整合到宿主细胞基因组中。

[0087] 因此,在一个实施方案中,本公开包括一种嵌合env蛋白质,其包含可操作地连接

到靶向多肽的逆转录病毒ENV蛋白质。靶向多肽可以是细胞特异性受体分子、细胞特异性受体的配体、细胞特异性抗原表位的抗体或抗体片段、或所属领域中容易鉴别的能够与靶细胞结合或相互作用的任何其它配体。靶向多肽或分子的示例包括使用生物素-抗生蛋白链菌素作为连接子的二价抗体(Etienne-Julan等人, *J. Of General Virol.*, 73, 3251-3255 (1992); Roux等人, *Proc. Natl. Acad. Sci USA* 86, 9079-9083 (1989)); 在包膜中含有编码针对半抗原的单链抗体可变区的序列的重组病毒(Russell等人, *Nucleic Acids Research*, 21, 1081-1085 (1993)); 将肽激素配体克隆到逆转录病毒包膜中(Kasahara等人, *Science*, 266, 1373-1376 (1994)); 嵌合EPO/env构建体(Kasahara等人, 1994); 针对亲嗜性(ecotropic)MLV包膜中的低密度脂蛋白(LDL)受体的单链抗体, 导致表达LDL受体的HeLa细胞的特异性感染(Somia等人, *Proc. Natl. Acad. Sci USA*, 92, 7570-7574 (1995)); 类似地, 可以通过合并整合素配体来改变ALV的宿主范围, 使得病毒现在能够跨物种以特异性地感染大鼠胶质母细胞瘤细胞(Valsesia-Wittmann等人, *J. Virol.* 68, 4609-4619 (1994)); 并且Dornberg和同事(Chu和Dornburg, *J. Virol* 69, 2659-2663 (1995)); M. Engelstadter等人 *Gene Therapy* 8, 1202-1206 (2001)) 已经报道了使用含有抗肿瘤标志物的单链抗体的包膜实现脾坏死病毒(SNV, 一种禽逆转录病毒)的组织特异性靶向。

[0088] 在一个实施方案中, 本公开的重组逆转录病毒以将病毒靶向于特定细胞类型(例如平滑肌细胞、肝细胞、肾细胞、成纤维细胞、角化细胞、间充质干细胞、骨髓细胞、软骨细胞、上皮细胞、肠细胞、乳腺细胞、赘生性细胞、神经胶质瘤细胞、神经元细胞和所属领域中已知的其它细胞)的方式被遗传修饰, 使得逆转录病毒载体的重组基因组被递送到靶标非分裂细胞、靶标分裂细胞或具有细胞增殖性病症的靶细胞。

[0089] 在另一个实施方案中, 靶向使用细胞特异性或组织特异性调控元件来促进病毒基因组在主动利用调控元件的被靶向细胞中的表达和转录, 如下文更充分地描述。转移的逆转录病毒遗传物质然后在宿主细胞内被转录和翻译成蛋白质。靶向调控元件典型地连接到5'和/或3'LTR上, 从而产生嵌合LTR。

[0090] 在一个实施方案中, 本公开提供一种可复制型逆转录病毒, 其不需要辅助病毒或额外的核酸序列或蛋白质以繁殖并产生病毒粒子。举例来说, 本公开的逆转录病毒的核酸序列分别编码群特异性抗原和逆转录酶(以及成熟和逆转录所必需的整合酶和蛋白酶), 如上文所论述。病毒gag和pol可以来源于慢病毒如HIV或致癌病毒或 γ 逆转录病毒如MoMLV。此外, 本公开的逆转录病毒的核酸基因组包括编码病毒包膜(ENV)蛋白质的序列。env基因可以来源于任何逆转录病毒。env可以是允许人和其它物种的细胞转导的双嗜性包膜蛋白质, 或者可以是只能转导小鼠和大鼠细胞的亲嗜性包膜蛋白质。此外, 可能期望通过将包膜蛋白质与抗体或用于靶向于特定细胞类型的受体的特定配体连接来靶向重组病毒。如上所述, 可以通过插入例如糖脂或蛋白质而逆转录病毒载体具有靶标特异性。通常通过使用抗体将逆转录病毒载体靶向于特定细胞类型(例如在特定组织中发现的细胞类型, 或癌细胞类型)上的抗原来实现靶向。所属领域的技术人员将了解, 或者不经过度实验即可容易地确定实现将逆转录病毒载体递送到特异性靶标的特定方法。在一个实施方案中, env基因来源于非逆转录病毒(例如CMV或VSV)。逆转录病毒源性的env基因的示例包括(但不限于)莫洛尼鼠白血病病毒(MoMuLV)、哈维鼠肉瘤病毒(Harvey murine sarcoma virus)(HaMuSV)、鼠乳腺肿瘤病毒(MuMTV)、长臂猿白血病病毒(GaLV)、人类免疫缺陷病毒(HIV)和劳氏肉瘤病

毒(RSV)。也可以使用其它env基因,例如水泡性口膜炎病毒(VSV)(蛋白质G)、巨细胞病毒包膜(CMV)或流感病毒红细胞凝集素(HA)。

[0091] 在一个实施方案中,逆转录病毒基因组来源于致癌逆转录病毒,且更具体地来源于哺乳动物致癌逆转录病毒。在另一个实施方案中,逆转录病毒基因组来源于 γ 逆转录病毒,且更具体地来源于哺乳动物 γ 逆转录病毒。“来源”意指母体多核苷酸序列是已经通过插入或删除天然存在的序列(例如插入IRES、插入编码目的多肽或抑制性核酸的异源多核苷酸、用来自不同逆转录病毒或病毒的更有效的启动子交换野生型启动子等)而被修饰的野生型致癌病毒。

[0092] 不同于通过所属领域中的标准方法产生的具有缺陷并且需要辅助以产生感染性载体粒子的重组逆转录病毒,本公开提供可复制型逆转录病毒。

[0093] 在另一个实施方案中,本公开提供使用调控序列靶向的逆转录病毒载体。可以利用细胞特异性或组织特异性调控序列(例如启动子)来靶向基因序列在特异性细胞群体中的表达。本文中其它地方描述了用于本公开的合适哺乳动物和病毒启动子。因此,在一个实施方案中,本公开提供一种在逆转录病毒基因组的5'端具有组织特异性启动子元件的逆转录病毒。典型地,组织特异性调控元件/序列位于逆转录病毒基因组的LTR的U3区中,包括例如用于增生性细胞的细胞特异性或组织特异性启动子和增强子(例如肿瘤细胞特异性增强子和启动子),和诱导型启动子(例如四环素)。

[0094] 在一些情况下,可能期望调控表达。举例来说,取决于期望表达水平,可以利用具有不同的活性强度的不同病毒启动子。在哺乳动物细胞中,通常使用CMV立即早期启动子来提供强转录活化。当期望转基因的表达水平降低时,也已经使用了效力较低的CMV启动子的经修改形式。当期望转基因在造血细胞中表达时,可以使用逆转录病毒启动子,例如来自MLV或MMTV的LTR。可以使用的其它病毒启动子包括SV40、RSV LTR、HIV-1和HIV-2LTR、例如来自E1A、E2A或MLP区的腺病毒启动子、AAV LTR、花椰菜花叶病毒、HSV-TK和禽肉瘤病毒。

[0095] 类似地,可使用组织特异性或选择性的启动子来实现特异性组织或细胞中的转录,以降低对于非靶向组织的潜在毒性或不当影响。举例来说,可以使用例如PSA、前列腺碱性蛋白(probacin)、前列腺酸性磷酸酶或前列腺特异性腺体激肽释放酶(hK2)等启动子来靶向前列腺中的基因表达。乳清辅助蛋白质(WAP)可用于乳房组织表达(Andres等人,PNAS 84:1299-1303,1987)。表3列举了可以使用的其它启动子/调控结构域。

[0096] “组织特异性调控元件”是能够驱动基因在一种组织中转录同时在其它组织类型中在很大程度上保持“沉默”的调控元件(例如启动子)。然而,应理解组织特异性启动子可以在其保持沉默的那些组织中具有可检测量的“背景”或“基础”活性。启动子在靶组织中被选择性活化的程度可以表示为选择性比(靶组织中的活性/对照组织中的活性)。就此而言,可用于本公开的实施方案中的组织特异性启动子典型地具有大于约5的选择性比。优选地,选择性比大于约15。

[0097] 在某些适应症中,可能期望在施用本公开的重组可复制型逆转录病毒(RRCR)之后的特定时间活化转录。这可以用激素可调控的或细胞因子可调控的启动子进行。举例来说,在适应症是其中产生或导向特异性类固醇的性腺组织的治疗应用中,使用受雄激素或雌激素调控的启动子可能是有利的。此类激素可调控的启动子包括MMTV、MT-1、蜕皮激素和RuBisco。可以使用其它受激素调控的启动子,例如对甲状腺、垂体和肾上腺激素有应答的

启动子。可以使用的对细胞因子和炎性蛋白质有应答的启动子包括K和T激肽原(Kininogen)(Kageyama等人,1987)、c-fos、TNF- α 、C-反应蛋白质(Arcone等人,1988)、结合珠蛋白(haptoglobin)(Oliviero等人,1987)、血清淀粉样蛋白A2、C/EBP α 、IL-1、IL-6(Poli和Cortese,1989)、补体C3(Wilson等人,1990)、IL-8、 α -1酸性糖蛋白(Prowse和Baumann,1988)、 α -1抗胰蛋白酶、脂蛋白脂酶(Zechner等人,1988)、血管紧张素原(Ron等人,1990)、纤维蛋白原、c-jun(可由佛波醇酯(phorbol ester)、TNF- α 、UV辐射、视黄酸和过氧化氢诱导)、胶原酶(由佛波醇酯和视黄酸诱导)、金属硫蛋白(重金属和糖皮质激素可诱导型)、基质降解酶(Stromelysin)(可由佛波醇酯、白细胞介素-1和EGF诱导)、 α -2巨球蛋白和 α -1抗胰糜蛋白酶。还可以使用例如骨钙蛋白、缺氧应答元件(HRE)、MAGE-4、CEA、 α -甲胎蛋白、GRP78/BiP和酪氨酸酶等肿瘤特异性启动子来调控肿瘤细胞中的基因表达。

[0098] 此外,这个启动子列表不应被解释成穷尽性的或限制性的,所属领域的技术人员将知道可与本文中所公开的启动子和方法结合使用的其它启动子。

[0099] 表3组织特异性启动子

组织	启动子
胰脏	胰島素 弹性蛋白 淀粉酶 pdr-1 pdx-1 葡萄糖激酶
肝脏	白蛋白 PEPCK HBV 增强子 α 甲胎蛋白 载脂蛋白 C α -1 抗胰蛋白酶 卵 黄蛋白原 NF-AB 甲状腺素运载蛋白 (Transthyretin)
骨骼肌	肌球蛋白 H 链 肌肉肌酸激酶 肌营养不良 蛋白 (Dystrophin) 钙蛋白酶 (Calpain) p94 骨骼 α -肌动蛋白 快速肌钙蛋白 1 (fast troponin 1)
皮肤	角蛋白 K6 角蛋白 K1
[0100] 肺	CFTR 人细胞角蛋白 18 (K18) 肺表面活性蛋白 A、B 和 C CC-10 P1
平滑肌	sm22 α SM- α -肌动蛋白
内皮	内皮素-1 E-选择素 温韦伯氏因子 (vonWillebrand factor) TIE (Korhonen 等 人, 1995) KDR/flk-1 黑素细胞酪氨酸酶
脂肪组织	脂蛋白脂肪酶 (Zechner 等人, 1988) 降脂 蛋白 (Spiegelman 等人, 1989) 乙酰辅酶 A 羧化酶 (Pape 和 Kim, 1989) 甘油磷酸脱 氢酶 (Dani 等人, 1989) 脂肪细胞 P2 (Hunt 等人, 1986)
乳房	乳清酸性蛋白 (WAP) (Andres 等人 PNAS 84:1299-1303 1987)
血液	β -球蛋白

[0101] 将进一步理解,某些启动子虽然在活性上并不受限于单一组织类型,但是仍然可能显示选择性,因为它们在一组组织中可能有活性,而在另一组中活性较低或沉默。此类启动子也被称为“组织特异性的”,并且预期用于本公开。举例来说,在各种中枢神经系统(CNS)神经元中有活性的启动子可以在治疗上用于防止因中风所致的损害,中风可以影响脑的许多不同区域中的任一个。因此,本公开中使用的组织特异性调控元件适用于调控异源蛋白质以及适用作本发明的逆转录病毒载体中的靶向多核苷酸序列。

[0102] 在另一个实施方案中,本公开提供包含来源于重组逆转录病毒的构建体的质粒。质粒可以被直接引入靶细胞或细胞培养物中,例如NIH 3T3或其它组织培养细胞。所得细胞将逆转录病毒载体释放到培养基中。

[0103] 鉴于上文和以下实施例,在一个实施方案中,本公开提供一种包含优化的IRES盒的重组可复制型逆转录病毒(RCR)。在一个实施方案中,逆转录病毒多核苷酸序列来源于选自自由以下组成的组的病毒:鼠白血病病毒(MLV)、莫洛尼鼠白血病病毒(MoMLV)、猫白血病病

毒(FeLV)、狨猴内源性逆转录病毒(BEV)、猪内源性病毒(PERV)、猫源性逆转录病毒RD114、松鼠猴逆转录病毒、异嗜性鼠白血病毒相关病毒(XMRV)、禽网状内皮组织增生病毒(REV)或长臂猿白血病毒(GALV)。在另一个实施方案中,RCR包含逆转录病毒GAG蛋白质;逆转录病毒POL蛋白质;逆转录病毒包膜(其可为嵌合型的、亲嗜性的和双嗜性的);逆转录病毒多核苷酸,其包含处于逆转录病毒多核苷酸序列的3'端的长末端重复(LTR)序列、gag、pol和env基因和优化的IRES盒(和/或任选的额外元件,其包括核心启动子、抑制性核酸如miRNA等)和处于逆转录病毒多核苷酸的5'端的LTR内的启动子。在一个实施方案中,3'LTR包含与SEQ ID NO:19、22或42的约核苷酸9405到约9998的序列至少98%同一的序列。在另一个实施方案中,在逆转录病毒多核苷酸的5'端的启动子序列适合于在哺乳动物细胞中表达。在任何上文的另一个实施方案中,启动子、gag、pol和env结构域包含与SEQ ID NO:19、22或42的约1到约8323的序列至少98%同一的序列,其中与SEQ ID NO:21的载体(pACE)相比,逆转录病毒多核苷酸缺少3'LTR下游的MLV序列的70个碱基对。在任何上文的另一个实施方案中,盒包含优化的内部核糖体进入位点(IRES)的,所述优化的内部核糖体进入位点(IRES)包含与SEQ ID NO:19、22或42的约8327到8875的序列至少98%同一的序列,并且在J-K分叉区的A凸出物中由6个A组成。在另一个实施方案中,优化的IRES可操作地连接到异源多核苷酸,其中盒位于3'LTR的5'和编码逆转录病毒包膜的env核酸结构域的3',并且与SEQ ID NO:21的pACE载体(pACE-CD)相比,所述盒在其两侧缺少小重复序列。在任何上文的另一个实施方案中,载体包括在靶细胞中逆转录、包装和整合所必需的顺式作用序列。在另一个实施方案中,与包含SEQ ID NO:21的载体(pACE)相比,RCR在6次传代后维持更高的复制能力,并且其中当表达异源多核苷酸时,与pAC3-yCD2(SEQ ID NO:22)载体相比,其产生至少20%、30%、40%、50%或更多的被表达的异源多肽。在另一个实施方案中,RCR多次感染靶细胞,导致拷贝/二倍体基因组的平均数为5或更大。在另一个实施方案中,逆转录病毒包膜是双嗜性MLV包膜。在一个实施方案中,启动子包括具有如SEQ ID NO:19、20、22或42中的核苷酸1到约核苷酸582所示的序列的CMV启动子,并且可包括对一个或多个核酸碱基的修饰,并且所述启动子能够引导和起始转录。在另一个实施方案中,启动子包含CMV-R-U5结构域多核苷酸。在另一个实施方案中,CMV-R-U5结构域包含来自人巨细胞病毒的连接MLV R-U5区域的立即早期启动子。在另一个实施方案中,CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中所示的序列至少99%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。在另一个实施方案中,gag核酸结构域包含SEQ ID NO:19、22或42的约核苷酸编号1203到约核苷酸2819的序列或与其具有至少99%或99.8%同一性的序列。在另一个实施方案中,多核苷酸的pol结构域来源于 γ 逆转录病毒。在另一个实施方案中,pol结构域包含SEQ ID NO:19、22或42的约核苷酸编号2820到约核苷酸6358的序列或与其具有至少99%或99.9%同一性的序列。在另一个实施方案中,env结构域包含SEQ ID NO:19、22或42的约核苷酸编号6359到约核苷酸8323的序列或与其具有至少99%或99.8%同一性的序列。在另一个实施方案中,IRES包含如SEQ ID NO:41中所示的序列。在另一个实施方案中,异源核酸包含具有如SEQ ID NO:3、5、11、13、15或17中所示的序列的多核苷酸。在另一个实施方案中,异源核酸编码包含如SEQ ID NO:4中所示的序列的多肽。在另一个实施方案中,异源核酸经过人密码子优化并且编码如SEQ ID NO:4中所示的多肽。在另一个实施方案中,异源核酸包

含如SEQ ID NO:19、22或42中的约核苷酸编号8877到约9353所示的序列。在另一个实施方案中,3'LTR包含U3-R-U5结构域。在另一个实施方案中,3'LTR包含如SEQ ID NO:19、22或42中的约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%同一的序列。在一个实施方案中,本公开提供一种包含SEQ ID NO:42的逆转录病毒多核苷酸。在另一个实施方案中,SEQ ID NO:42的逆转录病毒多核苷酸是其中T被U替换的RNA序列。在另一个实施方案中,根据SEQ ID NO:42的其中T是U的逆转录病毒RNA多核苷酸被封装在病毒衣壳中。在任何上文的另一个实施方案中,逆转录病毒多核苷酸可以进一步包含待递送给靶细胞的miRNA、siRNA或shRNA序列。miRNA、siRNA或shRNA可以可操作地连接到polIII启动子。miRNA可以位于优化的IRES盒的上游或下游。在另一个实施方案中,异源多核苷酸可以是任何数目的编码序列,包括细胞因子、免疫增强剂、胸苷激酶、胞嘧啶脱氨酶、嘌呤核苷磷酸化酶、受体、抗体和片段等。

[0104] 本公开还提供一种治疗细胞增殖性病症的方法,其包括使受试者与如本文中所述的逆转录病毒接触。在一个实施方案中,逆转录病毒含有优化的IRES,条件是连接到所述优化的IRES的异源多核苷酸包含胞嘧啶脱氨酶活性,和使受试者与5-氟胞嘧啶接触。在一个实施方案中,逆转录病毒感染细胞,导致包含SEQ ID NO:42的多核苷酸的整合。在另一个实施方案中,细胞增殖性病症是多形性胶质母细胞瘤。在另一个实施方案中,细胞增殖性病症选自自由以下组成的组:肺癌、结肠-直肠癌、乳腺癌、前列腺癌、泌尿道癌、子宫癌、脑癌、头颈癌、胰腺癌、黑素瘤、胃癌和卵巢癌。所述方法可以包括组合疗法,其中使待治疗的受试者与逆转录病毒接触并且进一步与抗癌剂或化学治疗剂接触。举例来说,抗癌剂或化学治疗剂可以选自自由以下组成的组:贝伐单抗(bevacizumab)、培加他尼(pegaptanib)、雷珠单抗(ranibizumab)、索拉非尼(sorafenib)、舒尼替尼(sunitinib)、AE-941、VEGF Trap、帕唑帕尼(pazopanib)、凡德他尼(vandetanib)、瓦他拉尼(vatalanib)、西地尼布(cediranib)、维甲酰酚胺(fenretinide)、角鲨胺、INGN-241、口服四硫钼酸盐、四硫钼酸盐、Panzem NCD、2-甲氧雌二醇、AEE-788、AG-013958、贝伐西尼钠(bevasiranib sodium)、AMG-706、阿昔替尼(axitinib)、BIBF-1120、CDP-791、CP-547632、PI-88、SU-14813、SU-6668、XL-647、XL-999、IMC-1121B、ABT-869、BAY-57-9352、BAY-73-4506、BMS-582664、CEP-7055、CHIR-265、CT-322、CX-3542、E-7080、ENMD-1198、OSI-930、PTC-299、Sirna-027、TKI-258、Veglin、XL-184或ZK-304709。

[0105] 在任何上文的另一个实施方案中,施用约 10^3 到 10^7 TU/g脑重量的逆转录病毒。在另一个实施方案中,施用约 10^4 到 10^6 TU/g脑重量的逆转录病毒。

[0106] 本公开提供一种多核苷酸构建体,其从5'到3'包含:可用于起始转录的启动子或调控区;psi包装信号;gag编码核酸序列;pol编码核酸序列;env编码核酸序列;在A凸出物中包含6个A的内部核糖体进入位点核酸序列;编码标志物、治疗性或诊断性多肽的异源多核苷酸;和LTR核酸序列。如本文中其它地方和下文所述,部分地取决于期望宿主细胞、表达时限或表达量和异源多核苷酸,对本公开的多核苷酸构建体的不同区段(例如重组可复制型逆转录病毒多核苷酸)进行工程改造。本公开的可复制型逆转录病毒构建体可以被分成许多结构域,所述结构域可以通过所属领域的技术人员单独修饰。

[0107] 举例来说,启动子可以包括具有如SEQ ID NO:19、20、22或42的核苷酸1到约核苷酸582中所示的序列的CMV启动子,并且可以包含对一个或多个(例如2-5、5-10、10-20、20-

30、30-50或更多个)核酸碱基的修饰,只要被修饰的启动子能够引导和起始转录即可。在一个实施方案中,启动子或调控区包含CMV-R-U5结构域多核苷酸。CMV-R-U5结构域包含来自人巨细胞病毒的连接到MLV R-U5区域的立即早期启动子。在一个实施方案中,CMV-R-U5结构域多核苷酸包含如SEQ ID NO:19、20、22或42中的约核苷酸1到约核苷酸1202所示的序列或与SEQ ID NO:19、20、22或42中的约核苷酸1到约核苷酸1202所示的序列至少95%同一的序列,其中所述多核苷酸促进与其可操作地连接的核酸分子的转录。多核苷酸的gag结构域可以来源于任何数目的逆转录病毒,但是将典型地来源于致癌逆转录病毒且更具体地来源于哺乳动物致癌逆转录病毒。在一个实施方案中,gag结构域包含来自如SEQ ID NO:19、20、22或42中所示的序列的约核苷酸编号1203到约核苷酸2819的序列或与其具有至少95%、98%、99%或99.8%(四舍五入到小数点后一位)同一性的序列。多核苷酸的pol结构域可以来源于许多逆转录病毒,但是将典型地来源于致癌逆转录病毒且更具体地来源于哺乳动物致癌逆转录病毒。在一个实施方案中,pol结构域包含来自如SEQ ID NO:19、20、22或42中所示的序列的约核苷酸编号2820到约核苷酸6358的序列或与其具有至少95%、98%、99%或99.9%(四舍五入到小数点后一位)同一性的序列。多核苷酸的env结构域可以来源于许多逆转录病毒,但是将典型地来源于致癌逆转录病毒或 γ 逆转录病毒且更具体地来源于哺乳动物致癌逆转录病毒或 γ 逆转录病毒。在一些实施方案中,env编码结构域包含双嗜性env结构域。在一个实施方案中,env结构域包含来自如SEQ ID NO:19、20、22或42中所示的序列的约核苷酸编号6359到约核苷酸8323的序列或与其具有至少95%、98%、99%或99.8%(四舍五入到小数点后一位)同一性的序列。多核苷酸的优化的IRES结构域可以从任何数目的内部核糖体进入位点获得。在一个实施方案中,优化的IRES来源于脑心肌炎病毒。在一个实施方案中,优化的IRES结构域包含如SEQ ID NO:41中所示的序列或与其具有至少95%、98%或99%(四舍五入到小数点后一位)同一性的序列,只要所述结构域允许核糖体的进入并且在A凸出物中包含6个A即可。异源结构域可以包含本公开的胞嘧啶脱氨酶(CD)。在一个实施方案中,CD多核苷酸包含经过人密码子优化的序列。在另一个实施方案中,CD多核苷酸编码具有胞嘧啶脱氨酶的突变多肽,其中突变赋予增加的热稳定性,这使解链温度(T_m)增加10°C,从而允许在更宽广的温度范围内的持续动力学活性和增加的蛋白质累积水平。在另一个实施方案中,本公开包含经过人密码子优化的胸苷激酶。异源结构域之后可以是富含多嘌呤的结构域。3'LTR可以来源于任何数目的逆转录病毒,典型地致癌逆转录病毒和优选地哺乳动物致癌逆转录病毒。在一个实施方案中,3'LTR包含U3-R-U5结构域。在另一个实施方案中,LTR包含如SEQ ID NO:19、20、22或42中的约核苷酸9405到约9998所示的序列或与其至少95%、98%或99.5%(四舍五入到小数点后一位)同一的序列。

[0108] 本公开还提供一种重组逆转录病毒载体,其从5'到3'包含:CMV-R-U5,即来自人巨细胞病毒的立即早期启动子与MLV R-U5区的融合体;PBS,即逆转录酶的引物结合位点;5'剪接位点; ψ 包装信号;gag,即MLV群特异性抗原的ORF;pol,即MLV聚合酶多蛋白质的ORF;3'剪接位点;4070A env,即MLV品系4070A的包膜蛋白质的ORF;优化的IRES,在A凸出物中由6个A组成;经过修饰的胞嘧啶脱氨酶(经过热稳定化和密码子优化)或经过人密码子优化的胸苷激酶;PPT,即多嘌呤束(polypurine tract);以及U3-R-U5,即MLV长末端重复序列。

[0109] 本公开还提供一种用于表达的包含如SEQ ID NO:42(或SEQ ID NO:42,其中T可以是U)中所示的序列的逆转录病毒载体,所述序列包含优化的A凸出物。在一个实施方案中,

IRES的优化的A凸出物由6个A组成。

[0110] 逆转录病毒载体可用于治疗众多种的疾病和病症,包括许多细胞增殖性疾病和病症(参见例如美国专利第4,405,712和4,650,764号;Friedmann,1989,Science,244:1275-1281;Mulligan,1993,Science,260:926-932,R.Crystal,1995,Science 270:404-410,其各自以全文引用的方式并入本文,还参见The Development of Human Gene Therapy, Theodore Friedmann编,Cold Spring Harbor Laboratory Press,Cold Spring Harbor, N.Y.,1999.ISBN 0-87969-528-5,其以全文引用的方式并入本文)。

[0111] 本公开还提供用于治疗细胞增殖性病症的基因疗法。此类疗法通过将适当的治疗性多核苷酸(例如反义分子、核酶、自杀基因、siRNA)引入到患有增殖性病症的受试者的细胞中来实现其治疗效果。多核苷酸构建体的递送可以使用本公开的重组逆转录病毒载体实现,尤其在所述重组逆转录病毒载体是基于能够感染分裂细胞的MLV时。

[0112] 此外,如本文中所述的治疗方法(例如基因疗法或基因递送方法)可以在体内或离体进行。可以优选在基因疗法前去除大部分肿瘤,例如通过外科手术或通过放射。在一些方面,逆转录病毒疗法可以在外科手术、化学疗法或放射疗法之后或之前进行。

[0113] 因此,本公开提供一种能够感染非分裂细胞、分裂细胞或赘生性细胞的重组逆转录病毒,其中重组逆转录病毒包含病毒GAG;病毒POL;病毒ENV;可操作地连接到IRES的异源核酸,所述IRES在A凸出物中由6个A组成;和包装、逆转录和整合所必需的顺式作用核酸序列。重组逆转录病毒可以是慢病毒如HIV,或者可以是致癌病毒。如上文针对产生重组逆转录病毒的方法所描述的,本公开的重组逆转录病毒可以进一步包含VPR、VIF、NEF、VPX、TAT、REV和VPU蛋白质中的至少一种。虽然不希望受特定理论束缚,但是相信这些基因/蛋白质产物中的一种或多种对于增加所产生的重组逆转录病毒(例如NEF)的病毒滴度至关重要,或可能是病毒粒子的感染和包装所必需的。

[0114] 本公开还提供一种将核酸转移到靶细胞以提供特定核酸(例如异源序列)的表达的方法。因此,在另一个实施方案中,本公开提供一种用于在靶细胞中引入和表达异源核酸的方法,其包括用本公开的重组病毒感染靶细胞,和在靶细胞中表达异源核酸。如上所述,靶细胞可以是任何细胞类型,包括分裂细胞、非分裂细胞、赘生性细胞、永生化细胞、被修饰细胞和由所属领域的技术人员知晓的其它细胞类型,只要它们能够被逆转录病毒感染即可。

[0115] 可能期望通过利用本公开的方法引入核酸序列(例如异源核酸序列)来调节细胞中的基因表达,其中核酸序列例如产生反义分子或核酶分子。术语“调节”意指在基因过度表达时抑制基因表达,或者在基因表达不足时增强表达。当细胞增殖性病症与基因表达有关时,可以使用在翻译水平干扰基因表达的核酸序列。这种方法例如利用反义核酸、核酶或三联试剂(triplex agent),通过用反义核酸或三联试剂掩蔽特异性mRNA或通过用核酶裂解特异性mRNA来阻断该特异性mRNA的转录或翻译。

[0116] 可能期望将编码生物应答调节剂(例如细胞因子)的核酸转移到细胞或受试者中。该类别中包括免疫增强剂,包括编码被分类为“白细胞介素”的许多细胞因子的核酸。这些生物应答调节剂包括例如白细胞介素1到15,以及本文中其它地方所描述的其它应答调节剂和因子。虽然不一定根据相同机制工作,但该类别中也包括干扰素且尤其是 γ 干扰素、肿瘤坏死因子(TNF)和粒细胞-巨噬细胞-集落刺激因子(GM-CSF)。其它多肽包括例如血管生

成因子和抗血管生成因子。可能期望将此类核酸递送到骨髓细胞或巨噬细胞以治疗酶缺乏或免疫缺陷。也可以将编码生长因子、毒性肽、配体、受体或其它生理上重要的蛋白质的核酸引入到特定靶细胞中。

[0117] 本公开可用于递送促进药物特异性靶向和作用的异源多核苷酸。举例来说,HER2 (EGF受体家族成员)是药物曲妥珠单抗(trastuzumab)(Herceptin™,Genentech)的结合靶标。曲妥珠单抗是抗体依赖性细胞毒性(ADCC)的介体。活性优先被靶向具有通过免疫组织化学确定的2+和3+过度表达水平的HER2表达细胞,而不是1+和非表达细胞(Herceptin prescribing information,Crommelin 2002)。通过在低HER2肿瘤中引入表达HER2或截短HER2(只表达胞外结构域和跨膜结构域)的载体来增强HER2的表达可以促进ADCC的最佳触发并且克服在临床应用上观察到的对赫赛汀(Herceptin)快速产生抗性。

[0118] 用yCD2(包含SEQ ID NO:19的约8877到9353)代替HER2的胞内结构域允许HER2的细胞表面表达和yCD2的胞质定位。可以通过PCR扩增(Yamamoto等人,Nature 319:230-234,1986;Chen等人,Canc.Res.,58:1965-1971,1998)或化学合成(BioBasic Inc.,Markham,Ontario,Canada)HER2胞外结构域(ECD)和跨膜结构域(TM)(SEQ ID NO:23的约位置175到2200的大约2026bp),并且可以将其插入载体pAC3-yCD2 SEQ ID NO:19中的IRES与yCD2基因之间(例如在SEQ ID NO:19的约核苷酸8876与8877之间)。或者,可以切除yCD基因并用编码HER2多肽或其片段的多核苷酸替换。仅具有ECD和TM结构域的赫赛汀结合结构域IV(来自位置1910到2200的大约290bp)的另一种截短HER2可以如上所述被扩增或化学合成并使用(Landgraf 2007;Garrett等人,J.of Immunol.,178:7120-7131,2007)。具有融合到结构域IV和TM上的天然信号肽(来自位置175-237的大约69bp)的这种截短形式的另一种修饰形式可以如上所述被化学合成并使用。所得病毒可与曲妥珠单抗或曲妥珠单抗和5-FC组合用于治疗受试者的细胞增殖性病征。

[0119] 或者,可以在含有不同ENV基因或其它适当表面蛋白质的单独载体中表达HER2和如上所述的修饰形式。这种载体可以是编码包膜和目的基因的可复制型的(Logg等人J.Mol Biol.369:1214 2007)或非复制型“第一代”逆转录病毒载体(Emi等人J.Virology 65:1202 1991)。在后一种情况下,现有的病毒感染将提供互补性的gag和pol以允许“非复制型”载体从任何先前被感染的细胞中进行感染性传播。替代性的ENV和糖蛋白包括异嗜性和多嗜性(polytropic)ENV和能够感染人细胞的糖蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000,Baum等人,Mol. Therapy,13(6):1050-1063,2006)。举例来说,多核苷酸可以包含以下序列,其中SEQ ID NO:19的GAG和POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV结构域,并且IRES或启动子如RSV可操作地直接连接到HER2、HER2ECDTM、HER2 ECDIVTM或HER2 SECDIVTM。

[0120] VSVG假型病毒和双嗜性逆转录病毒对细胞的混合感染导致产生子代病毒粒子,所述子代病毒粒子携带有被另一种病毒的包膜蛋白质衣壳化的一种病毒的基因组。假型逆转录病毒粒子的其它包膜同样如此。举例来说,上述来源的逆转录病毒所造成的感染导致产生能够在被感染细胞中编码yCD2和HER2(或变体)的子代病毒粒子。所得病毒可与曲妥珠单抗或曲妥珠单抗和5-FC组合用于治疗受试者的细胞增殖性病征。

[0121] 产生对曲妥珠单抗的抗性的另一个方面涉及干扰曲妥珠单抗活性所需的胞内信号传导。抗性细胞显示PTEN的损失和p27kip1的较低表达(Fujita,Brit J.Cancer,94:247,

2006;Lu等人,Journal of the National Cancer Institute,93(24):1852-1857,2001;Kute等人,Cytometry Part A 57A:86-93,2004)。举例来说,可以重组产生或化学合成(BioBasic Inc.,Markham,Canada)编码PTEN的多核苷酸,并且将其直接可操作地插入载体pAC3-yCD2 SEQ ID NO:19或22中的yCD2多核苷酸之后,或者与如上所述的连接子序列一起插入,或者作为yCD2的替代物插入。在另一个示例中,可以如上所述合成编码PTEN的多核苷酸(SEQ ID NO:25)并且将其插入IRES与yCD2序列之间或与如上所述的连接子一起插入。

[0122] 或者,可以在含有不同ENV基因或其它适当表面蛋白质的单独载体中表达PTEN。这种载体可以是可编码包膜和目的基因的复制型的(Logg等人J.Mol Biol.369:1214 2007)或非复制型“第一代”逆转录病毒载体(Emi等人,J.Virol 65:1202 1991)。在后一种情况下,现有的病毒感染将提供互补性的gag和pol以允许“非复制型”载体从任何先前被感染的细胞中进行感染性传播。替代性的ENV和糖蛋白包括异嗜性和多嗜性ENV和能够感染人细胞的糖蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV和其它病毒的糖蛋白(Palu,Rev Med Virol.2000,Baum,Mol.Ther.13(6):1050-1063,2006)。举例来说,多核苷酸可以包含以下序列,其中SEQ ID NO:19的GAG和POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV结构域,并且IRES或启动子如RSV可操作地直接连接到PTEN。

[0123] VSVG假型病毒和双嗜性逆转录病毒对细胞的混合感染导致产生子代病毒粒子,所述子代病毒粒子携带有被另一种病毒的包膜蛋白质衣壳化的一种病毒的基因组[Emi 1991]。假型逆转录病毒粒子的其它包膜同样如此。举例来说,上述来源的逆转录病毒所造成的感染导致产生能够在被感染细胞中编码yCD2和PTEN(或变体)或只编码PTEN的子代病毒粒子。所得病毒可与曲妥珠单抗或曲妥珠单抗和5-FC组合用于治疗受试者的细胞增殖性病征。

[0124] 类似地,可以化学合成(BioBasic Inc.,Markham,Canada)编码p27kip1的多核苷酸(SEQ ID NO:27和28)并且将其直接可操作地插入载体pAC3-yCD2 SEQ ID NO:19或SEQ ID NO:42中的yCD2基因之后,或者与连接子序列一起插入。在另一个示例中,可以如上所述合成编码p27kip1的多核苷酸并且将其插入在A凸出物中由6个A组成的IRES与yCD2序列之间,或者与如上所述的连接子一起插入,或者代替yCD2基因插入。

[0125] 或者,可以在含有不同ENV基因或其它适当表面蛋白质的单独载体中表达p27kip1。这种载体可以是编码包膜和目的基因的可复制型的(Logg等人J.Mol Biol.369:1214 2007)或非复制型“第一代”逆转录病毒载体(Emi等人J.Virol 65:1202 1991)。在后一种情况下,现有的病毒感染将提供互补性的gag和pol以允许“非复制型”载体从任何先前被感染的细胞中进行感染性传播。替代性的ENV和糖蛋白包括异嗜性和多嗜性ENV和能够感染人细胞的糖蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV和其它病毒的糖蛋白(Palu 2000,Baum 2006,见上文)。举例来说,多核苷酸可以包含以下序列,其中SEQ ID NO:19的GAG和POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV结构域,并且在A凸出物中由6个A组成的IRES或启动子如RSV直接可操作地连接到p27kip1。

[0126] VSVG假型病毒和双嗜性逆转录病毒对细胞的混合感染导致产生子代病毒粒子,所述子代病毒粒子携带有被另一种病毒的包膜蛋白质衣壳化的一种病毒的基因组[Emi 1991]。假型逆转录病毒粒子的其它包膜同样如此。举例来说,如上所述来源于SEQ ID NO:19、22和42的逆转录病毒所造成的感染导致产生能够在被感染细胞中编码yCD2和p27kip1

(或变体)的子代病毒粒子。所得病毒可与曲妥珠单抗或曲妥珠单抗和5-FC组合用于治疗受试者的细胞增殖性病征。

[0127] 在另一个示例中,CD20是药物利妥昔单抗(rituximab)(RituxanTM,Genentech)的结合靶标。利妥昔单抗是补体依赖性细胞毒性(CDC)和ADCC的介体。具有通过流式细胞术确定的较高平均荧光强度的细胞显示增强的对利妥昔单抗的敏感性(van Meerten等人,Clin Cancer Res 2006;12(13):4027-4035,2006)。通过在低CD20的B细胞中引入表达CD20的载体来增强CD20表达可以促进ADCC的最佳触发。

[0128] 举例来说,可以化学合成(BioBasic Inc.,Markham,Canada)编码CD20的多核苷酸(SEQ ID NO:29和30),并且将其直接可操作地插入载体pAC3-yCD2(-2)SEQ ID NO:19、22或42中的yCD2基因之后,或者与如上所述的连接子序列一起插入,或者作为yCD2基因的替代物插入。在另一个示例中,可以如上所述合成编码CD20的多核苷酸并且将其插入在A凸出物中由6个A组成的IRES与yCD2序列之间,或者与如上所述的连接子一起插入。作为另一个替代方案,可以在通过Psi1和Not1消化切除CD基因之后将CD20序列插入pAC3-yCD2载体中。

[0129] 在另一个示例中,可以化学合成(BioBasic Inc.,Markham,Canada)编码CD20的多核苷酸(SEQ ID NO:29和30),并且将其插入含有非双嗜性ENV基因或其它适当表面蛋白质的载体中(Tedder等人,PNAS,85:208-212,1988)。替代性的ENV和糖蛋白包括异嗜性和多嗜性ENV和能够感染人细胞的糖蛋白,例如来自MLV的NZB品系的ENV序列和来自MCF、VSV、GALV和其它病毒的糖蛋白[Palu 2000,Baum 2006]。举例来说,多核苷酸可以包含以下序列,其中SEQ ID NO:19的GAG和POL和yCD2基因被缺失,ENV对应于NZB MLV或VSV-g的异嗜性ENV结构域,并且在A凸出物中由6个A组成的IRES或启动子如RSV直接可操作地连接到CD20。

[0130] VSVG假型病毒和双嗜性逆转录病毒对细胞的混合感染导致产生子代病毒粒子,所述子代病毒粒子携带有被另一种病毒的包膜蛋白质衣壳化的一种病毒的基因组(Emi 1991)。假型逆转录病毒粒子的其它包膜同样如此。举例来说,如上所述来源于SEQ ID NO:19、22和42的逆转录病毒所造成的感染导致产生能够在被感染细胞中编码yCD2和CD20的子代病毒粒子。所得病毒可用于与Rituxan和/或5-FC组合治疗受试者的细胞增殖性病征。类似地,用仅编码CD20标志物的载体感染肿瘤使得可以通过使用Rituxan治疗肿瘤。

[0131] 嘧啶合成代谢中所涉及的酶和辅因子的水平可以是限制性的。与敏感细胞系相比,5-FU抗性癌细胞中的OPRT、胸苷激酶(TK)、尿苷单磷酸激酶和嘧啶核苷磷酸化酶的表达很低(Wang等人,Cancer Res.,64:8167-8176,2004)。大群体分析显示酶含量与疾病结果的相关性(Fukui等人,Int'l.J.OF Mol.Med.,22:709-716,2008)。可以利用CD和其它嘧啶合成代谢酶(PAE)的共表达来增加活性并且因此增加氟嘧啶药物的治疗指数。

[0132] 本公开提供用于治疗细胞增殖性病征例如癌症和赘瘤的方法,其包括施用本公开的RCR载体,随后用化学治疗剂或抗癌剂进行治疗。在一方面,在施用化学治疗剂或抗癌剂之前对受试者施用RCR载体并持续一段时间,以允许RCR感染和复制。然后用化学治疗剂或抗癌剂治疗受试者并持续一段时间和一定剂量,以减少增殖或杀死癌细胞。在一方面,如果使用化学治疗剂或抗癌剂的治疗减少但不杀死癌症/肿瘤(例如部分缓解或暂时缓解),那么随后可以用无毒治疗剂(例如5-FC)治疗受试者,所述无毒治疗剂在表达来自RCR的细胞毒性基因(例如胞嘧啶脱氨酶)的细胞中被转化为毒性治疗剂。

[0133] 使用此类方法,在肿瘤细胞的复制过程期间传播本公开的RCR载体,然后可以通过

用抗癌剂或化学治疗剂处理来杀死此类细胞,并且可以使用本文中所描述的RCR处理过程进行进一步的杀死。

[0134] 在本公开的另一个实施方案中,异源基因可以包含靶抗原(例如癌抗原)的编码序列。在这个实施方案中,用包含编码靶抗原的异源多核苷酸的RCR感染包含细胞增殖性病症的细胞以提供靶抗原的表达(例如癌抗原的过度表达)。然后对受试者施用包含与靶抗原特异性地相互作用的靶向同源部分的抗癌剂。靶向同源部分可以可操作地连接到细胞毒性剂或自身可以是抗癌剂。因此,被包含靶向抗原编码序列的RCR感染的癌细胞增加靶标在癌细胞上的表达,从而增加细胞毒性靶向的效率/功效。

[0135] 在另一个实施方案中,本公开的RCR可以包含含有与同源抗原或配体特异性地相互作用的结合结构域(例如抗体、抗体片段、抗体结构域或受体配体)的编码序列。包含结合结构域的编码序列的RCR然后可用于感染患有细胞增殖性病症的受试者的细胞,例如癌细胞或赘生性细胞。被感染细胞然后将表达结合结构域或抗体。然后将可操作地连接到细胞毒性剂的或自身具有细胞毒性的抗原或同源物施用于受试者。细胞毒性同源物然后将选择性地杀死表达结合结构域的被感染细胞。或者,结合结构域自身可以是抗癌剂。

[0136] 如本文中所使用,术语“抗体”是指包含至少一个免疫球蛋白可变域或免疫球蛋白可变域序列的蛋白质。举例来说,抗体可以包含重(H)链可变区(本文中缩写为VH)和轻(L)链可变区(本文中缩写为VL)。在另一个实施例中,抗体包含两个重(H)链可变区和两个轻(L)链可变区。术语“抗体”涵盖抗体的抗原结合片段(例如单链抗体、Fab片段、F(ab')₂、Fd片段、Fv片段和dAb片段)以及完整抗体。

[0137] 本公开提供一种治疗具有细胞增殖性病症的受试者的方法。受试者可以是任何哺乳动物,并且优选地是人。使受试者与本公开的重组可复制型逆转录病毒载体接触。接触可以在体内或离体进行。施用本公开的逆转录病毒载体的方法在所属领域中是已知的,并且包括例如全身施用、局部施用、腹膜内施用、肌肉内施用、颅内施用、脑脊髓施用以及在肿瘤或细胞增殖性病症的部位直接施用。其它施用途径在所属领域中也是已知的。

[0138] 因此,本公开包括可用于治疗细胞增殖性病症的各种药物组合物。根据本公开的药物组合物是通过使用载剂、赋形剂和添加剂或助剂,将根据本公开的含有可用于治疗或调节细胞增殖性病症的异源多核苷酸序列的逆转录病毒载体形成为适合施用于受试者的形式来制备的。常用的载剂或助剂包括碳酸镁、二氧化钛、乳糖、甘露醇和其它糖、滑石、乳蛋白、明胶、淀粉、维生素、纤维素和其衍生物、动物油和植物油、聚乙二醇和溶剂,例如无水、醇、甘油和多元醇。静脉内运载体包括流体和营养补充物。防腐剂包括抗微生物剂、抗氧化剂、螯合剂和惰性气体。其它药学上可接受的载剂包括水溶液、无毒赋形剂(包括盐)、防腐剂、缓冲剂等,例如如Remington's Pharmaceutical Sciences,第15版,Easton:Mack Publishing Co.,1405-1412,1461-1487(1975)和The National Formulary XIV.,第14版,Washington:American Pharmaceutical Association(1975)中所描述的,其内容以引用方式并入本文。根据所属领域中的常规技术调节药物组合物的各种组分的pH和确切浓度。参见Goodman和Gilman的The Pharmacological Basis for Therapeutics(第7版)。

[0139] 举例来说并且不构成限制,可用于治疗细胞增殖性病症的逆转录病毒载体将包含双嗜性ENV蛋白质、GAG和POL蛋白质、U3区逆转录病毒基因组中的启动子序列,和逆转录病毒基因组在靶细胞中复制、包装和整合所必需的所有顺式作用序列。

[0140] 以下实施例旨在说明而不是限制本公开。虽然此类实施例是可能使用的实施例中的典型实施例,但是也可以替代性地利用所属领域的技术人员已知的其它程序。

[0141] 实施例

[0142] 已经证明,当用Toca 511(pAC3-yCD2;SEQ ID NO:22)最大程度地感染细胞时,yCD2的表达水平和5-FC被yCD2转化为5-FU在体外和体内都是有效和稳定的。然而,在长期(大约180天)被感染Ba1b/c小鼠的体内初步研究中,来自一些组织的被整合的原病毒显示在J-K分叉环中携带有扩大或缩小的寡聚A序列。在来自通过分子PCR克隆分析的生物定位研究的4只小鼠的组织中,观察到7A不均匀地扩大为8A、9A、10A、11A和12A,以及7A缩小为6A。这一观察结果和pEMCF中的7个A,而非起初描述的ECMV IRES中的6个A,使得研究通过在A凸出物中具有各种数目的A的IRES所介导的yCD2表达的影响,尤其是对于RRV情形中的蛋白质翻译的影响。因此,产生了特别是分叉区的A凸出物中的一系列缺失和插入突变体。数据显示,A凸出物中缺失或插入寡聚A序列都不影响RRV产生,6个A提供最大的CD和绿色荧光蛋白质(GFP)表达,并且A的数目从6个A发生微小变化具有中度影响,但是更大的变化对IRES介导的mRNA从转基因的翻译的效率具有显著影响。

[0143] 在J-K分叉区的A凸出物中含有各种数目的A的RRV的构建,产生含有EMCV IRES并且编码CD或GFP的RRV,以在J-K分叉区的A凸出物中具有4个、5个、6个、7个、8个、10个或12个A。通过DNA合成(BioBasics Inc.)分别在5'端具有Mlu I和在3'端具有Psi I的完整IRES盒用于直接替换RRV骨架中的等同盒来产生各构建体(图1B)。通过在克隆到RRV骨架中之前和之后的测序分析来确认所有DNA片段。使用转基因的名称、继之以A凸出物中A的数目来命名含有yCD2转基因的RRV构建体(例如yCD2-4A含有yCD2转基因和IRES的A凸出物中的4个A)。

[0144] 在A凸出物中含有各种数目的A的RRV产生类似滴度,通过使用磷酸钙沉淀法瞬时转染293T细胞来产生病毒原液。在转染后大约42小时收集病毒上清液。进行病毒感染以确定滴度。随后使用各载体的病毒上清液来感染HT1080细胞以产生RRV生产细胞。在感染原初U87-MG细胞之前测量所得病毒滴度。图1C显示用含有各种数目的A的RRV感染的HT1080细胞产生类似的病毒水平,表明分叉环中A的数目不影响病毒复制。

[0145] 在J-K分叉区中含有各种数目的A的RRV表达类似水平的转录物但表达不同水平的蛋白质,然后使用来自HT1080细胞的病毒上清液以0.1的感染复数(MOI)感染原初U87-MG细胞。在感染后第10天,当细胞被完全感染时,通过定量实时聚合酶链式反应(qRT-PCR)测量细胞病毒RNA水平,并且通过免疫印迹法检查yCD2的蛋白质表达水平(Perez等人,2012)。使用两组不同的引物测量细胞病毒RNA表达水平,这两组引物分别位于env(5'Env2:5'-ACCCTCAACCTCCCCTACAAGT-3',3'Env2:5'-GTTAAGCGCCTGATAGGCTC-3',探针:5'FAM-AGCCACCCCCAGGAAGTGGAGA TAGA-3'BHQ)和yCD2区(5'yCD2:5'-ATCATCATGTACGGCATCCCTAG-3',3'yCD2:5'-TGAA CTGCTTCATCA GCTTCTTAC-3',探针:5'FAM-TCATCGTCAACAACCACCACCTCGT-3'BHQ)中(图2A)。使用 $2^{-\Delta\Delta(Ct)}$ 方法,相对于含有6个A的载体计算来自各载体的相对RNA水平。细胞病毒RNA水平的比率在0.8到1.1的范围(图2B),表明没有因IRES中的修饰所导致的病毒RNA转录物的显著差异。在通过蛋白质印迹(Western blot)检查这些载体的yCD2蛋白质表达水平时,含有5个A和7个A的载体的yCD2蛋白质表达水平被鉴别为yCD2-6A载体的69%和77%。相比之下,在含有4个、8个、10个和12个A的载体中观察到yCD2蛋白质表达的显著减少。这些载体的CD蛋白质表达水平在yCD2-6A载体的4%

到25%的范围(图2B)。伴有细胞病毒RNA的类似表达水平的yCD2蛋白质表达的显著减少表明,IRES的分叉区中的寡聚A的长度可能对翻译后水平上的基因表达具有很大影响。另外还通过将5FC添加到培养物中并在1小时后测量5-FU来测量相对胞内CD酶活性。与蛋白质印迹数据类似地对活性的差异进行分级,但没有同样显著。这可以归因于基于细胞的测定的局限性和胞内5-FC的低利用度,其低于所用测定中的酶的 K_m 。因此,用已经明确限定蛋白质表达测定的另一种转基因来分析环中A的数目的影响。另外,使用不同的转基因将允许确定yCD2蛋白质表达随A凸出物中A的数目的变化而发生的改变是否具有转基因特异性。

[0146] 产生了编码GFP的RRV的一个等同组。这些载体的病毒滴度彼此也是相当的,并且这一数据看起来与yCD2转基因的非常相似(图1C)。使用流式细胞术,通过选通GFP阳性细胞来测量GFP表达水平。将各载体的平均荧光强度(MFI)归一化为细胞病毒RNA水平,并相对于GFP-6A载体进行计算。来自这组载体的结果(图2C)与对于yCD2载体观察到的结果(图2B)一致,并且含有6个A的载体在两组载体中都从转基因表达最高水平的蛋白质。此外,归因于检测方法的灵敏性,揭示了GFP表达水平的显著差异,显示由含有10个A和12个A的载体所表达的GFP分别减少了大约96%和99%。在这两组载体中,具有7个A的RRV显示蛋白质表达减少了大约30%。与Hoffman等人所报道的发现一致,具有4个A和5个A的RRV分别显示与Hoffman等人所描述的868 Δ 4类似的表型,与具有6个A的RRV相比具有显著降低的蛋白质翻译效率。

[0147] 本公开证明J-K分叉区中A凸出物的长度据推测通过对翻译效率的作用来影响IRES下游的转基因的表达。先前的发现暗示,AUG11周围的环境、位于3' IRES中的多聚嘧啶束与顺反子中的第一个AUG之间的间距以及顺反子在mRNA上的排列在蛋白质翻译的调节中都起作用。数据显示6个A的存在提供了最高水平的转基因蛋白质表达,并且通过缩小或扩大2-4个核苷酸而改变A凸出物中A的数目可显著影响IRES下游的转基因的表达水平。蛋白质表达结果表明最佳的IRES配置通常与分叉环中有6个A相关,而7个A也是可接受的,很可能归因于先前由Kaminiski等人描述的多聚嘧啶束结合蛋白质(PTB)的挽救,显示凸出物A从6个A延长到7个A赋予了依赖于多聚嘧啶束结合蛋白质(PTB)的IRES功能。具有4个、5个、8个、10个和12个A的载体变体可能也需要PBT结合到多聚嘧啶束以实现有效的蛋白质翻译,并且这些载体变体可能显著扭曲IRES的二级和三级结构,从而损害PBT和/或其它反式作用因子与多聚嘧啶束的结合,因此减弱由PBT介导的翻译活性的挽救。除了所制得的用于双顺反子表达载体的EMCV IRES合成构建体以外,在EMCV中尚未描述A凸出物中的腺苷残基数目的变化。在小鼠中,似乎腺苷残基数目的变化不可能由任何种类的选择压力驱动,而是由于RRV易于突变的逆转录酶活性而发生在RRV经过180天的广泛复制期间。总之,在包含ECMV IRES的RRV中,优选使用IRES的6A形式,不仅因为增强的转基因表达,而且还因为寡聚A数目漂变的更频繁方向似乎优先朝向凸出物中更长的寡聚A。因此,如果凸出物始于6个A,那么在转基因表达方面对于单个额外腺苷核苷酸的获取有更高耐受性。

[0148] 所构建的含有具有6个A的最小IRES的RRV产生与只含有6个A的RRV类似水平的滴度、病毒转录物和转基因蛋白质表达。已经显示通过从5'EMCV IRES渐进缺失所产生的突变体具有体外差别翻译效率(Duke等人,J Virol.66:1602-9 1992)。此处,产生含有各种长度的最小IRES的RRV,其被命名为6A-406(例如SEQ ID NO:41的碱基123到544)和6A-466(SEQ ID NO:41的碱基183到544)(参见图5)。可以构建具有其它数目的A和406或466IRES序列的其它类似构建体(命名为7A-406和7A-466(指代含有7个A的最小IRES等)),并且所述其它类

似构建体与具有等同数目的A和全长IRES的构建体表现得大致相称。通过DNA合成(BioBasics Inc.)分别在5'端具有Mlu I和在3'端具有Psi I的完整IRES盒用于直接替换RRV骨架中的等同盒来产生各构建体。通过在克隆到RRV骨架中之前和之后的测序确认所有DNA片段。使用转基因的名称、继之以A凸出物中A的数目来命名含有yCD2转基因的RRV构建体(即yCD2-4A含有yCD2转基因和IRES的A凸出物中的4个A)。数据显示来自瞬时转染的293T细胞和最大程度被感染的HT1080细胞的滴度与凸出物A变体类似。从完全被感染的U87-MG细胞测量yCD2的蛋白质表达。与具有全长IRES的6A变体相比,6A-406变体表达类似水平(在2倍、5倍或10倍内)的yCD2蛋白质。携带5' IRES的又一缺失的6A-466变体显示yCD2的表达。此外,通过连续感染获得的复制动力学和载体稳定性的数据也显示6A-406和6A-466载体稳定长达至少10个感染周期。

[0149] 已经描述了本公开的许多实施方案。然而,应理解可以在不脱离本公开的精神和范围的情况下进行各种修改。因此,其它实施方案属于下列权利要求的范围内。

序列表

<110>	托卡根公司	
<120>	具有优化的A凸出物的重组载体	
<130>	00014-019W01	
<140>	尚未指定	
<141>	2014-08-05	
<150>	US 61/862,433	
<151>	2013-08-05	
<160>	42	
<170>	PatentIn 3.5版	
<210>	1	
<211>	477	
<212>	DNA	
<213>	酿酒酵母 (<i>Saccharomyces cerevisiae</i>)	
<220>		
<221>	CDS	
<222>	(1)..(477)	
[0001] <400>	1	
	atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac	48
	Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
	1 5 10 15	
	att gcc tat gag gag gcg gcc tta ggt tac aaa gag ggt ggt gtt cct	96
	Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro	
	20 25 30	
	att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt	144
	Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
	35 40 45	
	ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag	192
	Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
	50 55 60	
	atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa	240
	Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
	65 70 75 80	
	gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt	288
	Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
	85 90 95	
	gcc atc atc atg tat ggt att cca cgc tgt gtt gtc ggt gag aac gtt	336
	Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val	
	100 105 110	
	aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag	384

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

gtt gtt gtt gtt gac gat gag agg tgt aaa aag atc atg aaa caa ttt 432
 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140

atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag 477
 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 2
 <211> 158
 <212> PRT
 <213> 酿酒酵母

<400> 2

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

[0002] Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210>	3		
<211>	477		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	工程改造的胞嘧啶脱氨酶		
<220>			
<221>	CDS		
<222>	(1)..(477)		
<400>	3		
	atg gtg aca ggg gga atg gca agc aag tgg gat cag aag ggt atg gac		48
	Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp		
	1 5 10 15		
	att gcc tat gag gag gcg tta tta ggt tac aaa gag ggt ggt gtt cct		96
	Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro		
	20 25 30		
	att ggc gga tgt ctt atc aat aac aaa gac gga agt gtt ctc ggt cgt		144
	Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg		
	35 40 45		
	ggt cac aac atg aga ttt caa aag gga tcc gcc aca cta cat ggt gag		192
[0003]	Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu		
	50 55 60		
	atc tcc act ttg gaa aac tgt ggg aga tta gag ggc aaa gtg tac aaa		240
	Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys		
	65 70 75 80		
	gat acc act ttg tat acg acg ctg tct cca tgc gac atg tgt aca ggt		288
	Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly		
	85 90 95		
	gcc atc atc atg tat ggt att cca cgc tgt gtc atc ggt gag aac gtt		336
	Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val		
	100 105 110		
	aat ttc aaa agt aag ggc gag aaa tat tta caa act aga ggt cac gag		384
	Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu		
	115 120 125		
	gtt gtt gtt gtt gac gat gag agg tgt aaa aag tta atg aaa caa ttt		432
	Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe		
	130 135 140		
	atc gat gaa aga cct cag gat tgg ttt gaa gat att ggt gag tag		477
	Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu		
	145 150 155		
<210>	4		
<211>	158		

<212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 4

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

[0004] Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 5
 <211> 480
 <212> DNA
 <213> 人工序列

<220>
 <223> 经过人密码子优化的胞嘧啶脱氨酶

<220>

<221>	CDS	
<222>	(1)..(480)	
<400>	5	
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat		48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp		
1	5	10 15
atc gct tac gag gag gcc gca ctg ggc tac aag gag ggc ggc gtg cct		96
Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro		
	20	25 30
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg		144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg		
	35	40 45
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag		192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu		
	50	55 60
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag		240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys		
65	70	75 80
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc		288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly		
	85	90 95
[0005] gct atc atc atg tac ggc atc cct agg tgt gtg gtc ggc gag aac gtg		336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val		
	100	105 110
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag		384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu		
	115	120 125
gtg gtg gtt gtt gac gat gag agg tgt aag aag atc atg aag cag ttc		432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe		
	130	135 140
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tga taa		480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu		
145	150	155
<210>	6	
<211>	158	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成构建体	
<400>	6	
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp		
1	5	10 15

Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
 100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

[0006] Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu
 145 150 155

<210> 7
 <211> 756
 <212> DNA
 <213> 酿酒酵母

<220>
 <221> CDS
 <222> (1).. (756)

<400> 7
 atg aac ccg tta ttc ttt ttg get tet cca ttc ttg tac ctt aca tat 48
 Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
 1 5 10 15

ctt ata tat tat cca aac aaa ggg tct ttc gtt age aaa cct aga aat 96
 Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
 20 25 30

ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct 144
 Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro

	35	40	45	
	caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat			192
	Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn			
	50	55	60	
	aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga ttg			240
	Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu			
	65	70	75 80	
	ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg			288
	Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val			
		85	90 95	
	gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa			336
	Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys			
		100	105 110	
	atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga			384
	Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly			
		115	120 125	
	tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa			432
	Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln			
		130	135 140	
	agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta cca			480
[0007]	Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro			
	145	150	155 160	
	gag gat ata tet gaa agg tat gtc ttc cta tta gac cca atg ctg gcc			528
	Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala			
		165	170 175	
	acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga ggt			576
	Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly			
		180	185 190	
	gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag gaa			624
	Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu			
		195	200 205	
	ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act			672
	Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr			
		210	215 220	
	ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg			720
	Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly			
		225	230 235 240	
	ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa			756
	Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val			
		245	250	
	<210> 8			
	<211> 251			
	<212> PRT			

<213> 酿酒酵母

<400> 8

Met Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr
1 5 10 15

Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn
20 25 30

Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro
35 40 45

Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn
50 55 60

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

Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val
85 90 95

[0008] Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys
100 105 110

Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly
115 120 125

Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln
130 135 140

Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro
145 150 155 160

Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala
165 170 175

Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly
180 185 190

Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu
195 200 205

Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr
210 215 220

[0009]

Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly
 225 230 235 240

Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
 245 250

<210> 9
 <211> 1443
 <212> DNA
 <213> 智人 (homo sapiens)

<220>
 <221> CDS
 <222> (1)..(1443)

<400> 9

atg gct gtt gct cgt gct gct ctt ggt cct ctt gtt act ggt ctt tat	48
Met Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr	
1 5 10 15	
gat gtt caa get ttt aaa ttt ggt gat ttt gtt ctt aaa tet ggt ctt	96
Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu	
20 25 30	
tct tct cct att tat att gat ctt cgt ggt att gtt tct cgt cct cgt	144
Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg	
35 40 45	
ctt ctt tct caa gtt get gat att ctt ttt caa act gct caa aat get	192
Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala	
50 55 60	
ggt att tct ttt gat act gtt tgt ggt gtt cct tat act gct ctt cct	240
Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro	
65 70 75 80	
ctt gct act gtt att tgt tet act aat caa att cct atg ctt att cgt	288
Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg	
85 90 95	
cgt aaa gaa act aaa gat tat ggt act aaa cgt ctt gtt gaa ggt act	336
Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr	
100 105 110	
att aat cct ggt gaa act tgt ctt att att gaa gat gtt gtt act tct	384
Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser	
115 120 125	
ggt tct tct gtt ctt gaa act gtt gaa gtt ctt caa aaa gaa ggt ctt	432
Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu	
130 135 140	
aaa gtt act gat get att gtt ctt ctt gat cgt gaa caa ggt ggt aaa	480

	Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys	
	145 150 155 160	
	gat aaa ctt caa gct cat ggt att cgt ctt cat tct gtt tgt act ctt	528
	Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu	
	165 170 175	
	tct aaa atg ctt gaa att ctt gaa caa caa aaa aaa gtt gat gct gaa	576
	Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu	
	180 185 190	
	act gtt ggt cgt gtt aaa cgt ttt att caa gaa aat gtt ttt gtt gct	624
	Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala	
	195 200 205	
	gct aat cat aat ggt tct cct ctt tct att aaa gaa gct cct aaa gaa	672
	Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu	
	210 215 220	
	ctt tct ttt ggt gct cgt gct gaa ctt cct cgt att cat cct gtt gct	720
	Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala	
	225 230 235 240	
	tct aaa ctt ctt cgt ctt atg caa aaa aaa gaa act aat ctt tgt ctt	768
	Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu	
	245 250 255	
[0010]	tct gct gat gtt tct ctt gct cgt gaa ctt ctt caa ctt gct gat gct	816
	Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala	
	260 265 270	
	ctt ggt cct tct att tgt atg ctt aaa act cat gtt gat att ctt aat	864
	Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn	
	275 280 285	
	gat ttt act ctt gat gtt atg aaa gaa ctt att act ctt gct aaa tgt	912
	Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys	
	290 295 300	
	cat gaa ttt ctt att ttt gaa gat cgt aaa ttt gct gat att ggt aat	960
	His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn	
	305 310 315 320	
	act gtt aaa aaa caa tat gaa ggt ggt att ttt aaa att gct tct tgg	1008
	Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp	
	325 330 335	
	gct gat ctt gtt aat gct cat gtt gtt cct ggt tct ggt gtt gtt aaa	1056
	Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys	
	340 345 350	
	ggt ctt caa gaa gtt ggt ctt cct ctt cat cgt ggt tgt ctt ctt att	1104
	Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile	
	355 360 365	
	gct gaa atg tct tct act ggt tct ctt gct act ggt gat tat act cgt	1152
	Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg	
	370 375 380	

[0011]

gct gct gtt cgt atg gct gaa gaa cat tct gaa ttt gtt gtt ggt ttt	1200
Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe	
385	390 395 400
att tet ggt tet cgt gtt tet atg aaa cct gaa ttt ctt cat ctt act	1248
Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr	
	405 410 415
cct ggt gtt caa ctt gaa gct ggt ggt gat aat ctt ggt caa caa tat	1296
Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr	
	420 425 430
aat tct cct caa gaa gtt att ggt aaa cgt ggt tct gat att att att	1344
Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile	
	435 440 445
gtt ggt cgt ggt att att tct gct gct gat cgt ctt gaa gct gct gaa	1392
Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu	
	450 455 460
atg tat cgt aaa gct gct tgg gaa gct tat ctt tct cgt ctt ggt gtt	1440
Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val	
	465 470 475 480
taa	1443

<210> 10
 <211> 480
 <212> PRT
 <213> 智人

<400> 10

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

Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr
 100 105 110
 Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser
 115 120 125
 Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu
 130 135 140
 Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys
 145 150 155 160
 Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu
 165 170 175
 Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu
 180 185 190
 Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala
 195 200 205
 [0012] Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu
 210 215 220
 Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala
 225 230 235 240
 Ser Lys Leu Leu Arg Leu Met Gln Lys Lys Glu Thr Asn Leu Cys Leu
 245 250 255
 Ser Ala Asp Val Ser Leu Ala Arg Glu Leu Leu Gln Leu Ala Asp Ala
 260 265 270
 Leu Gly Pro Ser Ile Cys Met Leu Lys Thr His Val Asp Ile Leu Asn
 275 280 285
 Asp Phe Thr Leu Asp Val Met Lys Glu Leu Ile Thr Leu Ala Lys Cys
 290 295 300
 His Glu Phe Leu Ile Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly Asn
 305 310 315 320
 Thr Val Lys Lys Gln Tyr Glu Gly Gly Ile Phe Lys Ile Ala Ser Trp

	325	330	335
Ala Asp Leu Val Asn Ala His Val Val Pro Gly Ser Gly Val Val Lys	340	345	350
Gly Leu Gln Glu Val Gly Leu Pro Leu His Arg Gly Cys Leu Leu Ile	355	360	365
Ala Glu Met Ser Ser Thr Gly Ser Leu Ala Thr Gly Asp Tyr Thr Arg	370	375	380
Ala Ala Val Arg Met Ala Glu Glu His Ser Glu Phe Val Val Gly Phe	385	390	395
Ile Ser Gly Ser Arg Val Ser Met Lys Pro Glu Phe Leu His Leu Thr	405	410	415
Pro Gly Val Gln Leu Glu Ala Gly Gly Asp Asn Leu Gly Gln Gln Tyr	420	425	430
Asn Ser Pro Gln Glu Val Ile Gly Lys Arg Gly Ser Asp Ile Ile Ile	435	440	445
Val Gly Arg Gly Ile Ile Ser Ala Ala Asp Arg Leu Glu Ala Ala Glu	450	455	460
Met Tyr Arg Lys Ala Ala Trp Glu Ala Tyr Leu Ser Arg Leu Gly Val	465	470	475
			480

[0013]

- <210> 11
- <211> 1227
- <212> DNA
- <213> 人工序列

- <220>
- <223> 融合构建体CDopt-UPRT

- <220>
- <221> CDS
- <222> (1).. (1227)

<400> 11	
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	

atc get tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct	96
---	----

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc	432
[0014] Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag aac ccg	480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro	
145 150 155 160	
tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca tat ctt ata tat	528
Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr	
165 170 175	
tat cca aac aaa ggg tct ttc gtt agc aaa cct aga aat ctg caa aaa	576
Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys	
180 185 190	
atg tct tcg gaa cca ttt aag aac gtc tac ttg cta cct caa aca aac	624
Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn	
195 200 205	
caa ttg ctg ggt ttg tac acc atc atc aga aat aag aat aca act aga	672
Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg	
210 215 220	
cct gat ttc att ttc tac tcc gat aga atc atc aga ttg ttg gtt gaa	720
Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu	
225 230 235 240	
gaa ggt ttg aac cat cta cct gtg caa aag caa att gtg gaa act gac	768
Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp	
245 250 255	

acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt aaa atc tgt ggt Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly 260 265 270	816
gtt tcc att gtc aga gct ggt gaa tcg atg gag caa gga tta aga gac Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp 275 280 285	864
tgt tgt agg tct gtg cgt atc ggt aaa att tta att caa agg gac gag Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu 290 295 300	912
gag act gct tta cca aag tta ttc tac gaa aaa tta cca gag gat ata Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile 305 310 315 320	960
tct gaa agg tat gtc ttc cta tta gac cca atg ctg gcc acc ggt ggt Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly 325 330 335	1008
agt gct atc atg gct aca gaa gtc ttg att aag aga ggt gtt aag cca Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro 340 345 350	1056
gag aga att tac ttc tta aac cta atc tgt agt aag gaa ggg att gaa Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu 355 360 365	1104
[0015] aaa tac cat gcc gcc ttc cca gag gtc aga att gtt act ggt gcc etc Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu 370 375 380	1152
gac aga ggt cta gat gaa aac aag tat cta gtt cca ggg ttg ggt gac Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp 385 390 395 400	1200
ttt ggt gac aga tac tac tgt gtt taa Phe Gly Asp Arg Tyr Tyr Cys Val 405	1227
<210> 12	
<211> 408	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 12	
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp 1 5 10 15	
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30	

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
100 105 110

Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
130 135 140

[0016]

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Asn Pro
145 150 155 160

Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr Tyr Leu Ile Tyr
165 170 175

Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg Asn Leu Gln Lys
180 185 190

Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
195 200 205

Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys Asn Thr Thr Arg
210 215 220

Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
225 230 235 240

Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
245 250 255

Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
 260 265 270

Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
 275 280 285

Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
 290 295 300

Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
 305 310 315 320

Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
 325 330 335

Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
 340 345 350

Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
 355 360 365

[0017] Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
 370 375 380

Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
 385 390 395 400

Phe Gly Asp Arg Tyr Tyr Cys Val
 405

<210> 13
 <211> 1287
 <212> DNA
 <213> 人工序列

<220>
 <223> 融合构建体 - CDopt - 连接子 - UPRT

<220>
 <221> CDS
 <222> (1).. (1287)

<400> 13
 atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat
 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

48

	atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro 20 25 30	96
	atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg 35 40 45	144
	ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu 50 55 60	192
	atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys 65 70 75 80	240
	gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly 85 90 95	288
	gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val 100 105 110	336
	aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu 115 120 125	384
[0018]	gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe 130 135 140	432
	atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly 145 150 155 160	480
	ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly 165 170 175	528
	ggc gcc aac ccg tta ttc ttt ttg gct tct cca ttc ttg tac ctt aca Gly Ala Asn Pro Leu Phe Phe Leu Ala Ser Pro Phe Leu Tyr Leu Thr 180 185 190	576
	tat ctt ata tat tat cca aac aaa ggg tct ttc gtt agc aaa cct aga Tyr Leu Ile Tyr Tyr Pro Asn Lys Gly Ser Phe Val Ser Lys Pro Arg 195 200 205	624
	aat ctg caa aaa atg tct tcg gaa cca ttt aag aac gtc tac ttg cta Asn Leu Gln Lys Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu 210 215 220	672
	cct caa aca aac caa ttg ctg ggt ttg tac acc atc atc aga aat aag Pro Gln Thr Asn Gln Leu Leu Gly Leu Tyr Thr Ile Ile Arg Asn Lys 225 230 235 240	720
	aat aca act aga cct gat ttc att ttc tac tcc gat aga atc atc aga Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg 245 250 255 260 265 270	768

	245	250	255	
	ttg ttg gtt gaa gaa ggt ttg aac cat cta cct gtg caa aag caa att Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile 260 265 270			816
	gtg gaa act gac acc aac gaa aac ttc gaa ggt gtc tca ttc atg ggt Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly 275 280 285			864
	aaa atc tgt ggt gtt tcc att gtc aga gct ggt gaa tcg atg gag caa Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln 290 295 300			912
	gga tta aga gac tgt tgt agg tct gtg cgt atc ggt aaa att tta att Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile 305 310 315 320			960
	caa agg gac gag gag act gct tta cca aag tta ttc tac gaa aaa tta Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu 325 330 335			1008
	cca gag gat ata tct gaa agg tat gtc ttc cta tta gac cca atg ctg Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu 340 345 350			1056
[0019]	gcc acc ggt ggt agt gct atc atg gct aca gaa gtc ttg att aag aga Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg 355 360 365			1104
	ggt gtt aag cca gag aga att tac ttc tta aac cta atc tgt agt aag Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys 370 375 380			1152
	gaa ggg att gaa aaa tac cat gcc gcc ttc cca gag gtc aga att gtt Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val 385 390 395 400			1200
	act ggt gcc ctc gac aga ggt cta gat gaa aac aag tat cta gtt cca Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro 405 410 415			1248
	ggg ttg ggt gac ttt ggt gac aga tac tac tgt gtt taa Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val 420 425			1287
	<210> 14 <211> 428 <212> PRT <213> 人工序列			
	<220> <223> 合成构建体			
	<400> 14			
	Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp			

	1	5	10	15
	Ile Ala Tyr	Glu Glu Ala	Leu Leu Gly Tyr Lys	Glu Gly Gly Val Pro
		20	25	30
	Ile Gly Gly	Cys Leu Ile	Asn Asn Lys Asp Gly	Ser Val Leu Gly Arg
		35	40	45
	Gly His Asn	Met Arg Phe	Gln Lys Gly Ser Ala	Thr Leu His Gly Glu
		50	55	60
	Ile Ser Thr	Leu Glu Asn	Cys Gly Arg Leu	Glu Gly Lys Val Tyr Lys
		65	70	75
	Asp Thr Thr	Leu Tyr Thr	Thr Leu Ser Pro	Cys Asp Met Cys Thr Gly
		85	90	95
	Ala Ile Ile	Met Tyr Gly	Ile Pro Arg Cys	Val Ile Gly Glu Asn Val
		100	105	110
[0020]	Asn Phe Lys	Ser Lys Gly	Glu Lys Tyr Leu	Gln Thr Arg Gly His Glu
		115	120	125
	Val Val Val	Val Asp Asp	Glu Arg Cys Lys	Lys Leu Met Lys Gln Phe
		130	135	140
	Ile Asp Glu	Arg Pro Gln	Asp Trp Phe Glu	Asp Ile Gly Glu Ser Gly
		145	150	155
	Gly Gly Ala	Ser Gly Gly	Gly Ala Ser Gly	Gly Gly Ala Ser Gly Gly
		165	170	175
	Gly Ala Asn	Pro Leu Phe	Phe Leu Ala Ser	Pro Phe Leu Tyr Leu Thr
		180	185	190
	Tyr Leu Ile	Tyr Tyr Pro	Asn Lys Gly Ser	Phe Val Ser Lys Pro Arg
		195	200	205
	Asn Leu Gln	Lys Met Ser	Ser Glu Pro Phe	Lys Asn Val Tyr Leu Leu
		210	215	220
	Pro Gln Thr	Asn Gln Leu	Leu Leu Gly Leu	Tyr Thr Ile Ile Arg Asn Lys
		225	230	235
				240

Asn Thr Thr Arg Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg
245 250 255

Leu Leu Val Glu Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile
260 265 270

Val Glu Thr Asp Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly
275 280 285

Lys Ile Cys Gly Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln
290 295 300

Gly Leu Arg Asp Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile
305 310 315 320

Gln Arg Asp Glu Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu
325 330 335

Pro Glu Asp Ile Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu
340 345 350

[0021]

Ala Thr Gly Gly Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg
355 360 365

Gly Val Lys Pro Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys
370 375 380

Glu Gly Ile Glu Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val
385 390 395 400

Thr Gly Ala Leu Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro
405 410 415

Gly Leu Gly Asp Phe Gly Asp Arg Tyr Tyr Cys Val
420 425

<210> 15
<211> 1200
<212> DNA
<213> 人工序列

<220>
<223> 融合构建体 - CDopt3 - OPRT

<220>
 <221> CDS
 <222> (1)..(1200)

<400> 15

atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat	48
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp	
1 5 10 15	
atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct	96
Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	
20 25 30	
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg	144
Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	
35 40 45	
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag	192
Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	
50 55 60	
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag	240
Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	
65 70 75 80	
gac acc acc ctg tac acc acc ctg tcc cct tgt gac atg tgt acc ggc	288
Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	
85 90 95	
[0022] get atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg	336
Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	
100 105 110	
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag	384
Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	
115 120 125	
gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc	432
Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	
130 135 140	
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag gcg gtc	480
Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val	
145 150 155 160	
get cgt gca gct ttg ggg cca ttg gtg acg ggt ctg tac gac gtg cag	528
Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln	
165 170 175	
get ttc aag ttt ggg gac ttc gtg ctg aag agc ggg ctt tcc tcc ccc	576
Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro	
180 185 190	
atc tac atc gat ctg cgg ggc atc gtg tct ega ccg cgt ctt ctg agt	624
Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser	
195 200 205	
cag gtt gca gat att tta ttc caa act gcc caa aat gca ggc atc agt	672

Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser 210 215 220	
ttt gac acc gtg tgt gga gtg cct tat aca gct ttg cca ttg gct aca Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr 225 230 235 240	720
gtt atc tgt tca acc aat caa att cca atg ctt att aga agg aaa gaa Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu 245 250 255	768
aca aag gat tat gga act aag cgt ctt gta gaa gga act att aat cca Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro 260 265 270	816
gga gaa acc tgt tta atc att gaa gat gtt gtc acc agt gga tct agt Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser 275 280 285	864
gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc ttg aag gtc act Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr 290 295 300	912
gat gcc ata gtg ctg ttg gac aga gag cag gga ggc aag gac aag ttg Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu 305 310 315 320	960
[0023] cag gcg cac ggg atc cgc ctc cac tca gtg tgt aca ttg tcc aaa atg Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met 325 330 335	1008
ctg gag att ctc gag cag cag aaa aaa gtt gat gct gag aca gtt ggg Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly 340 345 350	1056
aga gtg aag agg ttt att cag gag aat gtc ttt gtg gca gcg aat cat Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His 355 360 365	1104
aat ggt tet ccc ctt tct ata aag gaa gca ccc aaa gaa ctc agc ttc Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe 370 375 380	1152
ggt gca cgt gca gag ctg ccc agg atc cac cca gtt gca tcg aag taa Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys 385 390 395	1200

<210> 16
 <211> 399
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 16

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

[0024] Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Val
 145 150 155 160

Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu Tyr Asp Val Gln
 165 170 175

Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly Leu Ser Ser Pro
 180 185 190

Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro Arg Leu Leu Ser
 195 200 205

Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn Ala Gly Ile Ser
 210 215 220

Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu Pro Leu Ala Thr
 225 230 235 240

[0025]

Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile Arg Arg Lys Glu
245 250 255

Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly Thr Ile Asn Pro
260 265 270

Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr Ser Gly Ser Ser
275 280 285

Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly Leu Lys Val Thr
290 295 300

Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly Lys Asp Lys Leu
305 310 315 320

Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr Leu Ser Lys Met
325 330 335

Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala Glu Thr Val Gly
340 345 350

Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val Ala Ala Asn His
355 360 365

Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys Glu Leu Ser Phe
370 375 380

Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val Ala Ser Lys
385 390 395

<210> 17
<211> 1260
<212> DNA
<213> 人工序列

<220>
<223> 融合构建体 - CDopt3 - 连接子 - OPRT

<220>
<221> CDS
<222> (1)..(1260)

<400> 17
atg gtg acc ggc ggc atg gcc tcc aag tgg gat caa aag ggc atg gat
Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp

48

1	5	10	15	
atc gct tac gag gag gcc ctg ctg ggc tac aag gag ggc ggc gtg cct Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro	20	25	30	96
atc ggc ggc tgt ctg atc aac aac aag gac ggc agt gtg ctg ggc agg Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg	35	40	45	144
ggc cac aac atg agg ttc cag aag ggc tcc gcc acc ctg cac ggc gag Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu	50	55	60	192
atc tcc acc ctg gag aac tgt ggc agg ctg gag ggc aag gtg tac aag Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys	65	70	75	240
gac acc acc ctg tac ace ace ctg tcc cct tgt gac atg tgt acc ggc Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly	85	90	95	288
gct atc atc atg tac ggc atc cct agg tgt gtg atc ggc gag aac gtg Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val	100	105	110	336
aac ttc aag tcc aag ggc gag aag tac ctg caa acc agg ggc cac gag Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu	115	120	125	384
[0026] gtg gtg gtt gtt gac gat gag agg tgt aag aag ctg atg aag cag ttc Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe	130	135	140	432
atc gac gag agg cct cag gac tgg ttc gag gat atc ggc gag tcc ggc Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly	145	150	155	480
ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc ggc gcc tcc ggc ggc Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly	165	170	175	528
ggc gcc gcg gtc gct cgt gca gct ttg ggg cca ttg gtg acg ggt ctg Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu	180	185	190	576
tac gac gtg cag gct ttc aag ttt ggg gac ttc gtg ctg aag agc ggg Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly	195	200	205	624
ctt tcc tcc ccc atc tac atc gat ctg cgg ggc atc gtg tct cga ccg Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro	210	215	220	672
cgt ctt ctg agt cag gtt gca gat att tta ttc caa act gcc caa aat Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn	225	230	235	720

	gea ggc atc agt ttt gac acc gtg tgt gga gtg cct tat aca get ttg Ala Gly Ile Ser Phe Asp Thr Val Cys Gly Val Pro Tyr Thr Ala Leu 245 250 255	768
	cca ttg gct aca gtt atc tgt tca acc aat caa att cca atg ctt att Pro Leu Ala Thr Val Ile Cys Ser Thr Asn Gln Ile Pro Met Leu Ile 260 265 270	816
	aga agg aaa gaa aca aag gat tat gga act aag cgt ctt gta gaa gga Arg Arg Lys Glu Thr Lys Asp Tyr Gly Thr Lys Arg Leu Val Glu Gly 275 280 285	864
	act att aat cca gga gaa acc tgt tta atc att gaa gat gtt gtc acc Thr Ile Asn Pro Gly Glu Thr Cys Leu Ile Ile Glu Asp Val Val Thr 290 295 300	912
	agt gga tct agt gtt ttg gaa act gtt gag gtt ctt cag aag gag ggc Ser Gly Ser Ser Val Leu Glu Thr Val Glu Val Leu Gln Lys Glu Gly 305 310 315 320	960
	ttg aag gtc act gat gcc ata gtg ctg ttg gac aga gag cag gga ggc Leu Lys Val Thr Asp Ala Ile Val Leu Leu Asp Arg Glu Gln Gly Gly 325 330 335	1008
	aag gac aag ttg cag gcg cac ggg atc cgc etc cac tca gtg tgt aca Lys Asp Lys Leu Gln Ala His Gly Ile Arg Leu His Ser Val Cys Thr 340 345 350	1056
[0027]	ttg tcc aaa atg ctg gag att etc gag cag cag aaa aaa gtt gat gct Leu Ser Lys Met Leu Glu Ile Leu Glu Gln Gln Lys Lys Val Asp Ala 355 360 365	1104
	gag aca gtt ggg aga gtg aag agg ttt att cag gag aat gtc ttt gtg Glu Thr Val Gly Arg Val Lys Arg Phe Ile Gln Glu Asn Val Phe Val 370 375 380	1152
	gea gcg aat cat aat ggt tct ccc ctt tet ata aag gaa gca ccc aaa Ala Ala Asn His Asn Gly Ser Pro Leu Ser Ile Lys Glu Ala Pro Lys 385 390 395 400	1200
	gaa etc agc ttc ggt gca cgt gca gag ctg ccc agg atc cac cca gtt Glu Leu Ser Phe Gly Ala Arg Ala Glu Leu Pro Arg Ile His Pro Val 405 410 415	1248
	gea teg aag taa Ala Ser Lys	1260

<210> 18
 <211> 419
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<400> 18

Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
 1 5 10 15

Ile Ala Tyr Glu Glu Ala Leu Leu Gly Tyr Lys Glu Gly Gly Val Pro
 20 25 30

Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
 35 40 45

Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
 50 55 60

Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
 65 70 75 80

Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
 85 90 95

Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Ile Gly Glu Asn Val
 100 105 110

[0028] Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
 115 120 125

Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Leu Met Lys Gln Phe
 130 135 140

Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ser Gly
 145 150 155 160

Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly Gly Ala Ser Gly Gly
 165 170 175

Gly Ala Ala Val Ala Arg Ala Ala Leu Gly Pro Leu Val Thr Gly Leu
 180 185 190

Tyr Asp Val Gln Ala Phe Lys Phe Gly Asp Phe Val Leu Lys Ser Gly
 195 200 205

Leu Ser Ser Pro Ile Tyr Ile Asp Leu Arg Gly Ile Val Ser Arg Pro
 210 215 220

Arg Leu Leu Ser Gln Val Ala Asp Ile Leu Phe Gln Thr Ala Gln Asn

225	230	235	240
Ala Gly Ile Ser	Phe Asp Thr Val Cys	Gly Val Pro Tyr Thr	Ala Leu
	245	250	255
Pro Leu Ala Thr	Val Ile Cys Ser	Thr Asn Gln Ile Pro	Met Leu Ile
	260	265	270
Arg Arg Lys Glu	Thr Lys Asp Tyr	Gly Thr Lys Arg	Leu Val Glu Gly
	275	280	285
Thr Ile Asn Pro	Gly Glu Thr Cys	Leu Ile Ile Glu	Asp Val Val Thr
	290	295	300
Ser Gly Ser Ser	Val Leu Glu Thr	Val Glu Val Leu	Gln Lys Glu Gly
305	310	315	320
Leu Lys Val Thr	Asp Ala Ile Val	Leu Leu Asp Arg	Glu Gln Gly Gly
	325	330	335
Lys Asp Lys Leu	Gln Ala His Gly	Ile Arg Leu His	Ser Val Cys Thr
	340	345	350
[0029]			
Leu Ser Lys Met	Leu Glu Ile Leu	Glu Gln Gln Lys	Lys Val Asp Ala
	355	360	365
Glu Thr Val Gly	Arg Val Lys Arg	Phe Ile Gln Glu	Asn Val Phe Val
	370	375	380
Ala Ala Asn His	Asn Gly Ser Pro	Leu Ser Ile Lys	Glu Ala Pro Lys
385	390	395	400
Glu Leu Ser Phe	Gly Ala Arg Ala	Glu Leu Pro Arg	Ile His Pro Val
	405	410	415
Ala Ser Lys			
<210>	19		
<211>	11892		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	RCR载体 - pAC3-yCD2		

[0030]

<400> 19	
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggecegee tggetgaccg cccaacgacc ccegccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata ggaetttee attgacgtca	180
atgggtggag tatttaacgt aaactgecca ettggcagta catcaagtgt atcatatgce	240
aagtaegecc cctattgacg tcaatgacgg taaatggccc gectgceatt atgeccagta	300
catgacetta tgggactttc ctacttggca gtacatctac gtattagtca tegctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aatgtcgtg acaactcegc ccattgacg caaatgggag gtaggcgtgt	540
acggtgggag gtctatataa gcagagetgg tttagtgaac cggegccagt cctccgattg	600
actgagtcgc ccgggtaccg gtgtatccaa taaacctct tgcagttgca tccgactgt	660
ggtctcgtg ttctttggga gggctctctc tgagtgattg actaccgctc agcgggggtc	720
tttcatttgg ggctcgtcc gggatcggga gacctctgcc caggaccac cgaccacca	780
ccgggaggtg agetggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctcgtc ggtactagtt agctaactag ctctgtatct ggcggaccgc	900
tggtggaact gacgagttcg gaacaccccg ccgcaacct gggagacgtc ccagggaact	960
cgggggcctg ttttgtggcc cgacctgagt ccaaaaatcc cgategtttt ggactctttg	1020
gtgcaccccc cttagaggag ggataatgtg ttctgttagg agacgagaac ctaaaacagt	1080
tcccgcctcc gtctgaattt ttgctttcgg tttgggaccg aagccgcgcc gcgcgtcttg	1140
tctgtgcag catcgtctg tgttctctct gtctgactgt gtttctgtat ttgtctgaga	1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggteactgg aaagatgtcg	1260
agcggategc tcacaaccag tcggtagatg tcaagaagag acgttgggtt acctctctct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cgcaacctt aaccgagacc	1380
tcataccca ggtaagatc aaggtctttt cacctgccc gcattggacac ccagaccagg	1440
tcccctacat cgtgacctgg gaagccttgg cttttgacct cctcctctgg gtaagcctt	1500
ttgtacccc taagcctcgc cctctctctc ctccatcgc cctctctctc ccccttgaac	1560
ctctctctc gaccccgct cgatcctccc tttatcagc cctcactctt tctctaggcg	1620
ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcactgac ctacttacag	1680

	aagaccccc	gccttatagg	gaccaagac	cacccccctt	cgacaggac	gaaatggtg	1740
	gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatgcatct	cgctacgtg	1800
	ggagacggga	gccccctgtg	gcegactcca	ctacctgca	ggcattcccc	ctccgcgcag	1860
	gaggaacgg	acagcttcaa	tactggcogt	tctctctt	tgaccttac	aactggaaaa	1920
	ataataacc	ttctttttct	gaagatccag	gtaaactgac	agctctgac	gagtctgtt	1980
	tcacaccca	tcagcccc	tgggacgact	gtcagcagct	gttggggact	ctgctgacc	2040
	gagaagaaa	acaacgggtg	ctcttagagg	ctagaaagc	ggtgcggggc	gatgatggc	2100
	gccccactca	actgccccat	gaagtcgatg	ccgcttttcc	cctcgagcgc	ccagactggg	2160
	attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgcccagttg	ctcctagcgg	2220
	gtctccaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280
	ggccccaatga	gtctccctcg	gccttcttag	agagacttaa	ggaagcctat	cgcaggtaca	2340
	ctccttatga	ccctgaggac	ccaggccaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
	agtctgcccc	agacattggg	agaaagttag	agaggttaga	agatttaaaa	aacaagacgc	2460
	ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaaacc	ccggaagaaa	2520
[0031]	gagaggaaag	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
	agcagaaaga	gaaagaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
	ctgtcgttag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccactcg	2700
	atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggactg	ggctaaagat	tgtcccaga	2760
	aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
	gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
	ccgtcacctt	cctggtagat	actggggccc	aacctccgt	gctgacccaa	aatcctggac	2940
	ccctaagtga	taagtctgcc	tgggtccaag	gggctactgg	aggaaagcgg	tatcgtgga	3000
	ccacggatcg	caaagtacat	ctagctaccg	gtaaggtcac	ccactctt	ctccatgtac	3060
	cagactgtcc	ctatctcttg	ttaggaagag	atttgetgac	taaactaaaa	gccccaatcc	3120
	actttgaggg	atcaggagcc	caggttatgg	gaccaatggg	gcagcccctg	caagtgttga	3180
	ccctaaatat	agaagatgag	catcggctac	atgagacctc	aaaagagcca	gatgtttctc	3240
	taggtccac	atggctgtct	gatttctctc	aggcctgggc	gaaaccggg	ggcatgggac	3300
	tggcagttcg	ccaagctcct	ctgatcatac	ctctgaaagc	aacctctacc	ccogtgtcca	3360
	taaaacaata	ccccatgtca	caagaagcca	gactggggat	caagccccac	atacagagac	3420

	tgttggacca	gggaatactg	gtaccctgcc	agtccccctg	gaacacgecc	ctgctacccg	3480
	ttaagaaacc	agggactaat	gattatagge	ctgtccagga	tctgagagaa	gtcaacaage	3540
	gggtggaaga	catccacccc	accgtgcccc	acccttacia	cctcttgagc	gggetecccac	3600
	egtecccacca	gtggtacact	gtgcttgatt	taaaggatgc	ctttttctgc	ctgagactcc	3660
	accccaccag	tcagcctctc	ttegcctttg	agtggagaga	tccagagatg	ggaatctcag	3720
	gacaattgac	ctggaccaga	ctcccacagg	gtttcaaaaa	cagteccacc	ctgtttgatg	3780
	aggcactgca	cagagaceta	gcagacttce	ggatccagca	cccagacttg	atcctgctac	3840
	agtacgtgga	tgacttactg	ctggcccaca	cttctgagct	agactgcca	caaggtactc	3900
	gggcccgtgt	acaaacceta	gggaacctcg	ggatctgggc	ctcggccaag	aaagcccaca	3960
	tttgccagaa	acaggtcaag	tatctggggt	atcttctaaa	agagggtcag	agatggctga	4020
	ctgaggccag	aaaagagact	gtgatgggga	agcctactcc	gaagaccctc	cgacaactaa	4080
	gggagttcct	agggacggca	ggetctgtgc	gcctctggat	ccctgggttt	gcagaaatgg	4140
	cagccccctt	gtaccctctc	accaaaacgg	ggactctgtt	taattggggc	ccagaccaac	4200
	aaaaggceta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagecctg	gggttgccag	4260
[0032]	atttgactaa	gccctttgaa	ctctttgtcg	acgagaagca	gggetacgec	aaaggtgtcc	4320
	taacgcaaaa	actgggacct	tggegtcgge	cgggtggceta	cctgtccaaa	aaagtagacc	4380
	cagtagcagc	tgggtggecc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
	aggatgcagg	caagctaacc	atgggacagc	cactagtcac	tctggcccc	catgcagtag	4500
	aggcactagt	caaacaaccc	cccgaccgct	ggctttccaa	cgcccggatg	actcactate	4560
	aggccttgct	tttgacacag	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaaccggg	4620
	ctacgtgctc	cccactgctc	gaggaaggga	tgcaacacaa	ctgccttgat	atcctggccc	4680
	aagcccacgg	aacccgacc	gaactaacgg	accagccgct	cccagacgce	gaccacacct	4740
	ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcggtga	4800
	ccaccgagac	cgaggtaatc	tgggtctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
	ctgaactgat	agcactcacc	caggeccctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
	atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaagge	4980
	gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
	taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaaagg	5100
	gacacagcgc	cgaggctaga	ggcaaccgga	tggetgacca	agcggcccga	aaggcagcca	5160

	tcacagagac	fccagacacc	tetaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
	aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
	ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
	ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
	tcctagagag	aagccacagt	cectactaca	tgctgaaccg	ggategaaca	ctcaaaaata	5460
	tcaactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
	gaactagggt	ccgcgggcat	eggcccgca	ctcattggga	gategatctc	accgagataa	5580
	agccccgatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
	tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggtcgt	aaccaagaag	ctactagagg	5700
	agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgcctteg	5760
	tctccaaggt	gagtcagaca	gtggccgate	tgttggggat	tgattggaaa	ttacattgtg	5820
	catacagacc	ccaagctca	ggccaggtag	aaagaatgaa	tagaaccate	aaggagactt	5880
	taactaaatt	aacgcttgca	actggetcta	gagactgggt	gtctctactc	cccttagccc	5940
	tgtaccgagc	ccgcaacaeg	ccgggcccc	atggcctcac	cccatatgag	atcttatatg	6000
[0033]	gggcaccccc	gccccttgta	aacttccctg	accctgacat	gacaagagtt	actaacagcc	6060
	ccctcteteca	agctcactta	caggtctctt	acttagtcca	gcacgaagtc	tggagacctc	6120
	tggcggcage	ctaccaagaa	caactggacc	gaccggtggt	acctcaccct	taccgagctg	6180
	ggacacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
	cttacacagt	cctgctgacc	aeccccaccg	ccctcaaagt	agacggcatc	gcagcttgga	6300
	tacacgccgc	ccacgtgaag	getgccgacc	ccgggggttg	accatcctct	agactgacat	6360
	ggcgcgttca	acgctctcaa	aacccccca	agataagatt	aacctgtgga	agcccttaat	6420
	agtcattggga	gtcctgttag	gagtagggat	ggcagagagc	ccccatcagg	tctttaatgt	6480
	aacctggaga	gtaccaacc	tgatgactgg	gcgtaccgcc	aatgacacct	ccctectggg	6540
	aactgtacaa	gatgccttec	caaaattata	ttttgatcta	tgtgatctgg	tggagagga	6600
	gtgggacctt	tcagaccagg	aaccgtatgt	egggtatggc	tgcaagtacc	ccgcagggag	6660
	acagcggacc	cggacttttg	acttttaagt	gtgccctggg	cataccgtaa	agtcgggggtg	6720
	tggggacca	ggagagggtt	actgtggtaa	atgggggtgt	gaaaccaccg	gacaggctta	6780
	ctggaagccc	acatcatcgt	gggacctaat	ctcccttaag	cgggtaaca	ccccctggga	6840
	cacgggatgc	tctaaagttg	cctgtggccc	ctgctaegac	ctctccaaag	tatccaatc	6900

	ettccaaggg gctactcgag ggggcagatg caaccctcta gtectagaat tcaactgatgc	6960
	aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
	aacagatect attaccatgt tctccctgac ceggcaggtc cttaatgtgg gacccccgagt	7080
	ccccataggg cccaaccag tattaccoga ccaaagacte cettccctcac caatagagat	7140
	tgtaccggt ccacagccac ctageccct caataccagt taceccctt ccactaccag	7200
	tacaccctca acctcccta caagtccaag tgtcccacag caeccccag gaactggaga	7260
	tagactacta gctctagtca aaggagccta tcaggcgctt aacctacca atcccgacaa	7320
	gaccaagaa tgttggctgt gettagtgc gggacctct tattacgaag gagtagcgg	7380
	egtgggcact tataccaate atccaccgc tceggccaac tgtaeggcca ctteccaaca	7440
	taagettacc ctatetgaag tgacaggaca gggcctatgc atggggcag tacctaaaac	7500
	teaccaggec ttatgtaaca ccaccaaag cgccggetca ggatectact acctgcagc	7560
	accgcccga acaatgtggg cttgcagcac tggattgact cctgcttgt ccaccacgt	7620
	getcaateta accacagatt attgtgtatt agttgaacte tggeccagag taatttacca	7680
	ctccccgat tatatgtatg gtcagettga acagcgtacc aatataaaa gagagccagt	7740
[0034]	atcattgacc ctggeccctc tactaggagg attaaccatg ggagggattg cagetggaat	7800
	agggaegggg accactgeet taattaaaac ccagcagttt gacagctte atgccgetat	7860
	ccagacagac ctcaacgaag tegaaaagtc aattaccaac ctagaaaagt cactgaccte	7920
	gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgetattcc taaaggaggg	7980
	aggtctctgc gcagecctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040
	gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
	aggccaagga tggtegaag ggctgtttaa tagatecccc tggtttacca cettaatctc	8160
	caccatcatg ggacctetaa tagtactctt actgatctta ctctttggac cttgcattct	8220
	caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
	tcagcaatat caccagctaa aaccataga gtacgagcca tgaacgcgtt actggccgaa	8340
	gcccgttga ataaggccgg tgtgegtgt tetatatgtt atttccacc atattgccgt	8400
	cttttgcaa tgtgagggcc eggaaacctg gccctgtctt cttgacgagc attcctaggg	8460
	gttttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc	8520
	ctctggaagc ttcttgaaga caaacaacgt ctgtagegac ctttgcagg cagcggaaacc	8580
	ccccacctgg cgacaggtgc ctctgccc cc aaaagccagc tgtataagat acacctgcaa	8640

	aggcggcaca	accccagtgc	cacgttgtga	gttggatagt	tgtggaaaga	gtcaaatggc	8700
	tctcctcaag	cgtattcaac	aaggggctga	aggatgccca	gaaggtacce	cattgtatgg	8760
	gatctgatct	ggggectcgg	tgcacatgct	ttacatgtgt	ttagtcgagg	ttaaaaaaac	8820
	gtctaggccc	cccgaaccac	ggggacgtgg	ttttcctttg	aaaaacacga	ttataaatgg	8880
	tgaccggcgg	catggcctcc	aagtgggate	aaaagggcac	ggatatecgt	tacgaggagg	8940
	ccctgctggg	ctacaaggag	ggcggtgtgc	ctatcgccgg	ctgtctgac	aacaacaagg	9000
	acggcagtgt	gctgggcagg	ggccacaaca	tgaggttcca	gaagggtcc	gccaccctgc	9060
	acggcgagat	ctccaccctg	gagaactgtg	gcaggctgga	gggcaagggt	tacaaggaca	9120
	ccaccctgta	caccaccctg	tecccttgtg	acatgtgtac	cggegtatc	atcatgtacg	9180
	gcatccctag	gtgtgtgate	ggegagaacg	tgaacttcaa	gtccaagggc	gagaagtacc	9240
	tgcaaaccag	gggccacgag	gtggtggttg	ttgacgatga	gaggtgtaag	aagctgatga	9300
	agcagttcat	cgacgagagg	cctcaggact	ggttcgagga	tatcggcgag	taagcggccg	9360
	cagataaaat	aaaagatttt	atttagtctc	cagaaaaagg	ggggaatgaa	agaccccacc	9420
	tgtaggtttg	gcaagctagc	ttaagtaacg	ccattttgca	aggeatggaa	aaatacataa	9480
[0035]	ctgagaatag	agaagttcac	ateaaaggta	ggaacagatg	gaacagetga	atatgggccca	9540
	aacaggatat	ctgtggtaag	cagttcctgc	cccggctcag	ggccaagaac	agatggaaca	9600
	gctgaatatg	ggccaaacag	gatatctgtg	gtaagcagtt	cctgccccgg	ctcagggccca	9660
	agaacagatg	gtccccagat	gcggtccagc	cctcagcagt	ttctagagaa	ccatcagatg	9720
	tticcagggt	gccccaaagga	cctgaaatga	ccctgtgcct	tatttgaact	aaccaatcag	9780
	ttcgcttctc	gcttctgttc	gegegttet	gctccccgag	ctcaataaaa	gagcccacaa	9840
	cccctcactc	ggggcgccag	tcctccgatt	gactgagtcg	cccgggtacc	cgtgtatcca	9900
	ataaaccttc	ttgcagttgc	ategacttg	tggctctgct	gttccttggg	agggtctctc	9960
	ctgagtgatt	gactaccctg	cagcgggggt	ctttcattac	atgtgagcaa	aaggccagca	10020
	aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcgttt	ttccatagge	tccgcccccc	10080
	tgacgagcat	cacaaaaatc	gacgtcaag	tcagaggtgg	cgaaacecga	caggactata	10140
	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgegc	tctctgttc	cgaccctgcc	10200
	gcttaccgga	facctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	ctcaatgctc	10260
	acgtgtagg	tatctcagtt	cggtgtaggt	cgttcgtctc	aagetgggct	gtgtgcacga	10320
	acccccggtt	cagcccgacc	gtggtgcctt	atccggtaac	tatcgtcttg	agtccaaccc	10380

gtaagacac gacttatcgc cactggcagc agccactggt aacagatta gcagagcgag	10440
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct aactagaag	10500
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa gagttggtag	10560
ctcttgatcc ggcaaacaaa ccaccctgg tagcggtggt tttttgttt gcaagcagca	10620
gattacgcgc agaaaaaag gatetcaaga agatcctttg atctttteta cggggtctga	10680
cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagttttaa tcaatctaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagttag gcacctatct cagcgatctg	10860
tctatttctg tcatccatag ttgectgaet ccccgctctg tagataacta cgatacggga	10920
gggettacca tctggeccca gtgetgcaat gataccgaga gaccacgct caccgctcc	10980
agattatca gcaataaacc agccagccgg aaggccgag cgcagaagtg gtectgcaac	11040
tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttcgcc	11100
agttaatagt ttgcgcaacg ttgttgcct tgetgcaggc atcgtggtgt cacgctcgtc	11160
gtttggtatg gcttcattca getccggttc ccaacgatca aggegagtta catgatcccc	11220
[0036] catgttctgc aaaaaagcgg ttagctcctt cggctctctg atcgttgtea gaagtaagtt	11280
ggccgcagtg ttatcactca tggttatggt agcactgcat aattctetta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgaactggtga gtactcaacc aagtcattct gagaatagtg	11400
tatgcggcga ccgagttgct ctgcccggc gtcaaacagg gataataccg cccacatag	11460
cagaacttta aaagtgtca tcatggaaa acgttcttcg gggegaaaac tctcaaggat	11520
cttaccgctg ttgagatcca gttgatgta acccaactctg gcaccaact gatcttcagc	11580
atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	11640
aaaggaata agggcgacac ggaaatgtg aatactcata ctcttcttt tcaatatta	11700
ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa	11760
aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga	11820
aaccattatt atcatgacat taacctataa aataggegt atcagagge cctttctct	11880
tcaagaatte at	11892

<210> 20
 <211> 11892
 <212> DNA
 <213> 人工序列

<220>		
<223>	RCR载体 - pAC3-yCD	
<400>	20	
	tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
	cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt	120
	gacgtcaata atgacgtatg tteccatagt aacgccaata ggaatttcc attgacgtca	180
	atgggtggag tattttacggt aaactgccc cttggcagta cateaagtgt atcatatgcc	240
	aagtacgccc cctattgacg teaatgacgg taaatggccc gcctggcatt atgcccagta	300
	catgaccotta tgggactttc ctacttggca gtacatctac gtattagtca tegctattac	360
	catggtgatg cggtttttggc agtacatcaa tgggcgtgga tageggittg actcacgggg	420
	atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
	ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt	540
	acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
	actgagtegc cgggtaccc gtgtatccaa taaccctct tgcagttgca tccgacttgt	660
	ggctctcgtg ttccttggga gggctctctc tgagtgattg actaccgtc agcggggggtc	720
[0037]	tttcatttgg ggctctgtec gggatcgga gaccctgcc cagggaccac cgaaccacca	780
	ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
	tgattttatg cgcctgcgtc ggtaactagt agetaactag ctctgtatct ggegaccgcg	900
	tgggtggaact gacgagttcg gaacaccocg ccgcaaccct gggagacgtc ccagggactt	960
	egggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgategtttt ggactctttg	1020
	gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
	tccgcctcc gtctgaattt ttgcttcgg tttgggaccg aagcgcgcc gcgcgtcttg	1140
	tctgctgcag categttctg tgttctctct gtctgactgt gttctgtat ttgtctgaga	1200
	atatgggcea gactgttacc acteccttaa gtttgacctt aggtcactgg aaagatgtcg	1260
	agcggatcgc tcacaaccag teggtagatg tcaagaagag acgttgggtt acctctctgt	1320
	ctgcagaatg gccaaccttt aacgtcgat ggccgcgaga cggcaccttt aaccgagacc	1380
	tcataccca ggttaagatc aaggctttt cacctggccc gcattggacac ccagaccagg	1440
	tcccctacat cgtgacctgg gaagccttgg cttttgacct cctctcctgg gtaagccct	1500
	ttgtacaccc taagcctccg cctctcttct ctccatccgc cccgtctctc ccccttgaac	1560
	ctctctgttc gaccccgctt cgtctctccc tttatccagc cctcactctt tctctagggc	1620

ccaaacctaa	acctcaagtt	ctttctgaca	gtggggggcc	gctcatcgac	ctacttacag	1680
aagaccccc	gccttatagg	gaccaagac	cacccccttc	cgacaggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatggcatct	cgcttacgtg	1800
ggagacggga	gccccctgtg	gcegaactca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
gaggaaaagg	acagcttcaa	tactggccgt	tctcctcttc	tgacctttac	aactggaaaa	1920
ataataaacc	ttctttttct	gaagatccag	gtaaactgac	agctctgac	gagtctgttc	1980
tcataccca	tcageccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgcccact	gaagtcgatg	ccgcttttcc	cctcgagcgc	ccagactggg	2160
attacaacc	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctcaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280
ggcccaatga	gtctccctcg	gccttcctag	agagacttaa	ggaagcctat	cgcaggtaca	2340
ctccttatga	ccctgaggac	ccagggcaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agctgcccc	agacattggg	agaaagttag	agaggttaga	agatttaaaa	aacaagacgc	2460
[0038] ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaaacc	ccggaagaaa	2520
gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
agcagaaaga	gaaagaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
ctgtcgttag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccaactcg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggcactg	ggctaaagat	tgtcccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctccct	cctgacccta	gatgactagg	2820
gaggtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
ccgtcacctt	cctggtagat	actggggccc	aacactccgt	gctgacccaa	aatcctggac	2940
ccctaagtga	taagtctgcc	tgggtccaag	gggctactgg	aggaaagcgg	tatcgctgga	3000
ccaeggatcg	caaagtacat	ctagctaccg	gtaaggctac	ccactcttct	ctccatgtac	3060
cagactgtcc	ctatectctg	ttaggaagag	atttgetgac	taactaaaa	gccccaaatc	3120
actttgaggg	atcaggagcc	caggttatgg	gaccaatggg	gcagcccctg	caagtgttga	3180
ccctaaatat	agaagatgag	catcggctac	atgagacctc	aaaagagcca	gatgtttctc	3240
tagggtccac	atggctgtct	gattttctct	aggcctgggc	ggaaaccggg	ggcatgggac	3300
tggcagttcg	ccaagctcct	ctgatecata	ctctgaaagc	aacctctacc	cccgtgtcca	3360

	taaaacaata	ccccatgtca	caagaagcca	gactggggat	caagccccac	atacagagac	3420
	tgttgacca	gggaatactg	gtaccctgcc	agteccccctg	gaacaegccc	ctgctacecg	3480
	ttaagaaace	agggactaat	gattataggc	ctgtccagga	tctgagagaa	gtcaacaage	3540
	gggtggaaga	catccacccc	accgtgocca	acccttacia	cctcttgage	gggctccca	3600
	cgccccacca	gtggtacact	gtgettatt	taaaggatgc	ctttttctgc	ctgagactcc	3660
	accccaccag	tcagcctctc	ttegcctttg	agtggagaga	tccagagatg	ggaatctcag	3720
	gacaattgac	ctggaccaga	ctcccacagg	gtttcaaaaa	cagtcccacc	ctgtttgatg	3780
	aggcactgca	cagagaccta	gcagacttcc	ggatccagca	cccagacttg	atcctgctac	3840
	agtacgtgga	tgacttactg	ctggcgcaca	cttctgagct	agaactgcaa	caaggtacte	3900
	gggccctggt	acaaacccta	gggaacctcg	ggatccgggc	ctcggccaag	aaagcccaaa	3960
	tttgccagaa	acaggtcaag	tatctggggt	atcttctaaa	agagggtcag	agatggctga	4020
	ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccctc	cgacaactaa	4080
	gggagtteet	agggacggca	ggcttctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140
	cagccccctt	gtaccctctc	accaaaaagg	ggactctggt	taattggggc	ccagaccaac	4200
[0039]	aaaaggccta	tcaagaaate	aagcaagctc	ttctaactge	cccageccctg	gggttgccag	4260
	atttgactaa	gccctttgaa	ctctttgtcg	acgagaagea	gggctacgcc	aaaggtgtcc	4320
	taacgcaaaa	actgggacct	tggegtcggc	cggtggccta	cctgtccaaa	aagctagacc	4380
	cagtagcagc	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
	aggatgcagg	caagctaacc	atgggacagc	cactagtcac	tctggccccc	catgcagtag	4500
	aggcactagt	caaacaacce	cccgaccget	ggctttccaa	cgcccggatg	actcactate	4560
	agcccttget	tttgacacg	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaacccegg	4620
	ctacgtgct	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccc	4680
	aagcccacgg	aaccggaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
	ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcggtga	4800
	ccaccgagac	cgaggtaate	tgggtctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
	ctgaactgat	agcactcacc	caggccctaa	agatggcaga	aggtagaag	ctaaatgttt	4920
	atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
	gtgggttget	cacatcagaa	ggcaagaga	tcaaaaataa	agaagagatc	ttggccctac	5040
	taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100

	gacacagcgc cgaggctaga ggcaaccgga tggctgacca agcggccccga aaggcagcca	5160
	tcacagagac tccagacacc tetaccctcc tcatagaaaa tteateacce tacacctcag	5220
	aacattttca ttacacagtg actgatataa aggacctaac caagttgggg gccatttatg	5280
	ataaaaacaaa gaagtattgg gtctaccaag gaaaacctgt gatgectgac eagtttactt	5340
	ttgaattatt agacttttctt cateagctga ctcacctcag cttctcaaaa atgaaggctc	5400
	tcctagagag aagccacagt ccctactaca tgcctgaaccg ggategaaca ctcaaaaata	5460
	tactgagac ctgcaaagct tgtgcacaag tcaacgccag caagtctgcc gttaacagg	5520
	gaactagggt ccgcgggcat cggcccggca ctctattggga gatcgatttc accgagataa	5580
	agcccggatt gtatggctat aaatatcttc tagtttttat agataccttt tetggctgga	5640
	tagaagcett cccaaccaag aaagaaaccg ccaaggtcgt aaccaagaag ctactagagg	5700
	agatcttccc caggttcggc atgectcagg tattgggaac tgacaatggg cctgccttcg	5760
	tctccaaggt gagtcagaca gtggccgac tgttggggat tgattggaaa ttacattgtg	5820
	catacagacc ccaaagctca ggccaggtag aaagaatgaa tagaaccatc aaggagactt	5880
	taactaaatt aacgcttgca actggctcta gagactgggt gctcctactc cccttagccc	5940
[0040]	tgtaaccgagc ccgcaacacg ccgggcccc atggcctcac cccatagag atettatag	6000
	gggcaccccc gcccttgta aatttccctg acctgacat gacaagagtt actaacagcc	6060
	ccctctcca agctcactta caggtctct acttagtcca gcaegaagtc tggagacctc	6120
	tggcggcagc ctaccaagaa caactggacc gaccgggtgt acctaccct taccgagtcg	6180
	gcgacacagt gtgggtccgc cgacaccaga ctaagaacct agaacctcgc tggaaaggac	6240
	cttacacagt cctgctgacc accccaccg cctcaaaagt agacggcatc gcagcttggg	6300
	tacacgccgc ccacgtgaag gctgccgacc cgggggttg accatcctct agactgacat	6360
	ggcgcgttca acgctctcaa aaccctca agataagatt aaccctgga agcccttaat	6420
	agtcattgga gtcctgttag gtagggat ggcagagagc cccatcagg tctttaatgt	6480
	aacctggaga gtcaccaacc tgatgactgg gcgtaccgcc aatgccacct ccctcctggg	6540
	aactgtacaa gatgecttcc caaaattata ttttgateta tgtgatctgg tcggagagga	6600
	gtgggacctc tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag	6660
	acagcggacc cggacttttg acttttacgt gtgccctggg cataccttaa agtcgggggtg	6720
	tgggggacca ggagagggct actgtggtaa atgggggtgt gaaaccaccg gacaggctta	6780
	ctggaagccc acatcatcgt gggacctaat ctcccttaag cgcggtaaca ccccctggga	6840

cacgggatgc	tctaaagttg	ctgtggccc	ctgctacgac	ctctccaaag	tatccaatte	6900
cttccaaggg	gctactcgag	ggggcagatg	caacctctta	gtcctagaat	tcaactgatgc	6960
aggaaaaaag	gctaactggg	acgggcccac	atcgtgggga	ctgagactgt	accggacagg	7020
aacagatcct	attaccatgt	tctccctgac	ceggcaggtc	cttaatgtgg	gacccccgagt	7080
ceecataggg	cccaaccag	tattaccga	ccaaagactc	cttctctcac	caatagagat	7140
tgtaccggct	ccacagccac	ctagcccct	caataccagt	taceccccct	ccactaccag	7200
tacacctca	acctcccta	caagtccaag	tgtcccacag	ccacccccag	gaactggaga	7260
tagactacta	gctctagtea	aaggagccta	tcaggcgctt	aacctcacca	atcccgacaa	7320
gacccaagaa	tgttggetgt	getttagtgc	gggacctct	tattaegaag	gagtagcgg	7380
cgtgggcact	tataccaate	atccaccgc	tccggccaac	tgtaeggcca	cttcccaaca	7440
taagettacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgcgggetca	ggatcctact	accttgcagc	7560
accgccegga	acaatgtggg	cttcagcac	tggattgact	cctgcttgt	ccaccacggt	7620
gtccaateta	accacagatt	attgtgtatt	agttgaacte	tggeccagag	taatttacca	7680
[0041] etccccgat	tatatgtatg	gteagcttga	acagcgtacc	aatataaaa	gagagccagt	7740
ateattgacc	ctggcccttc	tactaggagg	attaaccatg	ggagggattg	cagctggaat	7800
agggacgggg	accactgctt	taattaaaac	ccagcagttt	gagcagcttc	atgcccctat	7860
ccagacagac	ctcaacgaag	tcgaaaagtc	aattaccaac	ctagaaaagt	cactgacctc	7920
gttgtctgaa	gtagtcttac	agaaccgcag	aggcctagat	ttgctattcc	taaaggaggg	7980
aggtctctgc	gcagccctaa	aagaagaatg	ttgtttttat	gcagaccaca	cggggctagt	8040
gagagacagc	atggccaaat	taagagaaag	gcttaatcag	agacaaaaac	tatttgagac	8100
aggccaagga	tggttcgaag	ggctgtttaa	tagatcccc	tggtttacca	cttaattctc	8160
caccatcatg	ggacctetaa	tagtactctt	actgatetta	ctctttggac	cttgcattct	8220
caatcgattg	gtccaatttg	ttaaagacag	gatctcagtg	gtccaggctc	tggttttgac	8280
tcagcaatat	caccagctaa	aacceataga	gtacgagcca	tgaacgcgtt	actggccgaa	8340
gccgcttggc	ataaggccgg	tgtgcgtttg	tctatatggt	atcttccacc	atattgcegt	8400
cttttgcaa	tgtgagggcc	cggaaacctg	gccctgtctt	cttgacgagc	attcctaggg	8460
gtctttcccc	tctcgccaaa	ggaatgcaag	gtctgtttgaa	tgctgtgaag	gaagcagttc	8520
ctctggaagc	ttcttgaaga	caaacaacgt	ctgtagegac	ctttgcagg	cagcggaaacc	8580

	ccccacctgg cgacaggfgc ctctgcggcc aaaagccacg tgtataagat acacctgcaa	8640
	aggcggcaca accccagtgc cacgttgtga gttggatagt tgtggaaaga gtcaaatggc	8700
	tctcctcaag cgtattcaac aaggggctga aggatgceca gaaggtaccc cattgtatgg	8760
	gatctgatct ggggcctcgg tgcacatgct ttacatgigt ttagtegagg ttaaaaaaac	8820
	gtctaggecc cccgaaccac ggggacgtgg ttttcctttg aaaaacacga ttataaatgg	8880
	tgacaggggg aatggcaage aagtgggatc agaagggtat ggacattgcc tatgaggagg	8940
	cggccttagg ttacaaagag ggtggtgcte ctattggcgg atgtcttacc aataacaaag	9000
	acggaagtgt tctcggctgt ggtcacaaca tgagatttca aaagggatcc gccacactac	9060
	atggtgagat ctccactttg gaaaactgtg ggagattaga gggcaaagtg tacaaagata	9120
	ccaactttgta tacgacgctg tctccatgag acatgtgtac aggtgccatc atcatgtatg	9180
	gtattccacg ctgtgttctc ggtgagaacg ttaatttcaa aagtaagggc gagaaatatt	9240
	tacaaactag aggtcacgag gttgttgtg ttgacgatga gaggtgtaaa aagatcatga	9300
	aacaatttat cgatgaaaga cctcaggatt ggtttgaaga tattggtgag taggcggccg	9360
	cagataaaat aaaagatttt atttagtctc cagaaaaagg ggggaatgaa agaccccacc	9420
[0042]	tgtaggtttg gcaagctagc ttaagtaacg ccattttgca aggcattgaa aaatacataa	9480
	ctgagaatag agaagtccag atcaaggcca ggaacagatg gaacagctga atatgggcca	9540
	aacaggatat ctgtggttaag cagttcctgc cccggctcag ggccaagaac agatggaaca	9600
	gtgaaatag ggccaaacag gatactctgt gtaagcagtt cctgccccgg ctccaggcca	9660
	agaacagatg gtecccagat ggggtccagc cctcagcagt ttctagagaa ccatcagatg	9720
	ttccagggt gccccaagga cctgaaatga ccctgtgctt tatttgaact aaccaatcag	9780
	ttcgttctc gcttctgttc gcgccttct gctccccgag ctcaataaaa gagcccacaa	9840
	cccctcactc ggggcgccag tctccgatt gactgagtcg cccgggtacc cgtgtatcca	9900
	ataaaccttc ttgcagttgc atccgacttg tggctctgct gttccttggg agggctctct	9960
	ctgagtgatt gactaccctg cagcgggggt ctttcattac atgtgagcaa aaggccagca	10020
	aaaggccagg aaccgtaaaa agcccgctt gctggcgctt ttccatagge tccgcccc	10080
	tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaaccga caggactata	10140
	aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctctgttc cgaccctgcc	10200
	gcttaccgga tacctgtccg cttttctccc ttccgggaagc gtggcgctt ctcaatgctc	10260
	acgtgttagg tatctcagtt cgggttaggt cgttcgctcc aagctgggct gtgtgcacga	10320

acccccggtt cagccccgacc gctgcgcctt atcecgtaac tategtcttg agtccaacec	10380
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagegag	10440
glatgtaggc ggtgctacag agttcttgaa gtggtagcct aactacggct acactagaag	10500
gacagtatit ggtatctgeg ctctgctgaa gecagttacc tteggaaaaa gagttggtag	10560
ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt ttttttgttt geaagcagea	10620
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggctctga	10680
cgetcagtg aacgaaaact cacgttaagg gattttggtc atgagattat caaaaaggat	10740
cttcacctag atccttttaa attaaaaatg aagttttaa tcaatetaaa gtatatatga	10800
gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
tctatttegt tcatecatag ttgcctgact ccccgctgtg tagataacta cgatacggga	10920
ggccttacea tctggeccca gtgctgcaat gataccgoga gacccaecgt caccggctcc	10980
agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg gtcttgcac	11040
tttatccgcc tccatccagt ctattaattg ttgccggaa gctagagtaa gtagttcgcc	11100
agttaatagt ttgcgcaacg ttgttgccat tgctgcagge atcgtggtgt cacgctcgtc	11160
[0043] gtttggtatg gettcattca gctccggtc ccaacgatca aggcgagta catgatcccc	11220
catgttggtc aaaaaagegg ttagctcctt cggctctccg atcgttgtea gaagtaagtt	11280
ggccgcagtg ttateactea tggttatggc agcactgcat aattetetta ctgtcatgcc	11340
atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct gagaatagtg	11400
tatgcggcga ccgagttgct ctgccccggt gtaaacacgg gataataacc cccacatag	11460
cagaacttta aaagtgetca tcattggaaa acgttctctg gggcgaaaac tctcaaggat	11520
cttaccgctg ttgagatcca gttcgatgta acccaactgt gcaccaact gatcttcagc	11580
atcttttact ttaccageg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	11640
aaaggaata agggcgacac ggaaatgttg aatactcata ctcttcctt tcaaatatta	11700
ttgaagcatt taccagggtt attgtctcat gacgggatac atatttgaat gtatttagaa	11760
aaataaaca ataggggtc cgcgcacatt tccccgaaa gtgccacctg acgtctaaga	11820
aaccattatt atcatgacat taacctataa aataggcgt atcacgagc cctttcgtct	11880
tcaagaattc at	11892
<210> 21	
<211> 12007	
<212> DNA	

<213>	人工序列	
<220>		
<223>	RCR载体 - pACE-CD	
<400>	21	
	tagttattaa tagtaatcaa ttacggggtc attagttcat ageccatata tggagttccg	60
	egttacataa cttacggtaa atgcccgcg ttggtgaccg cccaacgacc ccccccatt	120
	gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc attgacgtca	180
	atgggtggag tattttacggt aaactgccc aattggcagta catcaagtgt atcatatgcc	240
	aagtacgecc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
	catgacctta tgggaacttc ctacttgcca gtacatctac gtattagtea tegetattac	360
	catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
	atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
	ggactttcca aaatgtcgt acaactccgc cccattgacg caaatggcg gtaggcgtgt	540
	acggtgggag gtctatataa gcagagctgg tttagtgaac cggcgccagt cctccgattg	600
	actgagtcgc ccgggtaccg gtgtatccaa taaacctct tgcagttgca tccgacttgt	660
[0044]	ggctctcgtg ttctttggga gggctctctc tgagtgattg actaccgtc agcggggggtc	720
	tttcatittg gggctctgccc gggatcggga gacctctgc cagggaccac cgaccacca	780
	ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
	tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcgaccgcg	900
	tgggtggaact gacgagttcg gaacaccgga ccgcaacct gggagacgtc ccagggactt	960
	cggggggcgt ttttgtggcc cgaactgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
	gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080
	tcccgcctcc gtctgaattt ttgcttcgg tttgggaccg aagccgcgcc gcgcgtcttg	1140
	tcctctgcag catcgttctg tgttctctct gtctgactgt gttctgtat ttgtctgaga	1200
	atatgggcca gactgttacc actcccttaa gtttgaactt aggtcaactg aaagatgtcg	1260
	agcggatcgc tcacaaccag teggtagatg tcaagaagag acgttgggtt acctctgct	1320
	ctgcagaatg gccaaccttt aacgtcgat ggccgcgaga cggcaccttt aaccgagacc	1380
	tcataccca ggttaagatc aaggctttt cacctggccc gcatggacac ccagaccagg	1440
	tcccctacat cgtgacctgg gaagccttgg cttttgacct ccttccctgg gtcaagccct	1500
	ttgtacacc taagcctccg cctctcttc ctccatcgc cccgtctct ccccttgaac	1560

	ctcctcgttc gacccccgct cgatcctccc tttatccage cctcactcct tctctaggcg	1620
	ccaaacctaa acctcaagtt ctttctgaca gtggggggcc gctcatcgac ctacttacag	1680
	aagacccccc gccttatagg gacccaagac caccaccttc cgacagggac ggaaatgggtg	1740
	gagaagegac ccctgcggga gaggeaccgg acccctcccc aatggeatct cgcctacgtg	1800
	ggagacggga gccccctgtg gccgactcca ctacctgca ggcattcccc ctccgcgcag	1860
	gaggaaacgg acagcttcaa tactgcccgt tctcctcttc tgaccttac aactggaaaa	1920
	ataataacct ttctttttct gaagatccag gtaaactgac agctctgac gactctgttc	1980
	tcatcaccca tcagcccacc tgggacgact gtcagcagct gttggggact ctgctgaccg	2040
	gagaagaaaa acaacgggtg ctcttagagg ctagaaaggc ggtgeggggc gatgatgggc	2100
	gccccactca actgcccact gaagtcgatg ccgcttttcc cctcgagcgc ccagactggg	2160
	attacaccac ccaggcaggt aggaaccacc tagtccacta tcgccagttg ctectagcgg	2220
	gtctccaaaa cgcgggcaga agccccacca atttggccaa ggtaaaagga ataacacaag	2280
	ggcccaatga gtctccctcg gcttccctag agagacttaa ggaagcctat cgcaggtaca	2340
	ctccttatga ccctgaggac ccagggcaag aaactaatgt gtctatgtct ttcatttggc	2400
[0045]	agctgcccc agacattggg agaaaagttag agaggttaga agatttaaaa aacaagacgc	2460
	ttggagattt ggttagagag gcagaaaaga tctttaataa acgagaaaacc ccggaagaaa	2520
	gagaggaacg tatcaggaga gaaacagagg aaaaagaaga acgccctagg acagaggatg	2580
	agcagaaaaga gaaagaaaaga gatcgtagga gacatagaga gatgagcaag ctattggcca	2640
	ctgtcgttag tggacagaaa caggatagac agggaggaga acgaaggagg tcccaactcg	2700
	atcgcgacca gtgtgcctac tgcaaagaaa aggggcactg ggctaaagat tgtccaaga	2760
	aaccacgagg acctegggga ccaagacccc agacctcct cctgacccta gatgactagg	2820
	gaggtcaggg tcaggagccc ccccctgac ccaggataac cctcaaagtc ggggggcaac	2880
	ccgtcacctt cctggtagat actggggccc aacctccgt gctgacccaa aatcctggac	2940
	ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tatcgctgga	3000
	ccacggatcg caaagtacat ctagnetacc gtaaggteac ccactcttc ctccatgtac	3060
	cagactgtcc ctatctcttg ttaggaagag atttgctgac taaactaaaa gcccaatcc	3120
	actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
	ccctaaatat agaagatgag cctggctac atgagacctc aaaagagcca gatgtttctc	3240
	tagggtccac atggetgtct gatttccctc aggcctgggc ggaaaccggg ggcattggac	3300

tggcagtteg ccaagctcct ctgatacatac ctctgaaagc aacctctacc cccgtgtcca	3360
taaaacaata ccccatgtca caagaagcca gactggggat caagccccac atacagagac	3420
tgttgacca gggaatactg gtacctgcc agtccccctg gaacacgecc ctgctacceg	3480
ttaagaaacc agggactaat gattatagge ctgtccagga tetgagagaa gtcaacaage	3540
gggtggaaga catccacccc aecgtgcca accettacaa cctettgagc gggetccac	3600
cgccccacca gtggtacact gtgettgatt taaaggatgc cttttctgc ctgagactcc	3660
accccaccag tcagectctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
aggcaactgca cagagaacct geagaactcc ggatccagca cccagaactg atcctgctae	3840
agtacgtgga tgacttactg ctggccgcca cttctgagct agaetgcaa caaggtacte	3900
gggccctgtt acaaaccta gggaacctg ggtatcgggc ctggccaag aaagcccaaa	3960
tttgccagaa acaggtcaag tatctggggc atcttctaaa agagggtcag agatggctga	4020
ctgaggccag aaaagagact gtgatgggc agcctactcc gaagaccct cgacaactaa	4080
gggagtctct agggacgcea ggctctgtc gcctctggat cctggggtt gcagaaatgg	4140
[0046] cagccccctt gtacctctc aceaaaaagg ggactctgtt taattggggc ceagaccaac	4200
aaaaggccta tcaagaaatc aageaagctc ttctaactgc cccagccctg gggttgccag	4260
attgactaa gccctttgaa ctctttgtc aegagaagea ggctacgcc aaaggtgtcc	4320
taocgaaaa actgggacct tggcctggc cgggtggcta cctgtccaaa aagctagacc	4380
cagtagcage tgggtggccc cettgcctac ggatggtage agccattgcc gtaactgaaa	4440
aggatgcagg caagetaacc atgggacagc cactagtcac tetggcccc catgcagtag	4500
aggcaactag caaacaacce cccgaccct ggttttcaa cgccggatg actcaactc	4560
aggccttget tttggacacg gaccgggtcc agttcggacc gttggtagcc ctgaaccgg	4620
ctacgctget cccaetgect gaggaagggc tgeaacaaa ctgecttgat atcctggccc	4680
aagcccacgg aaccggacc gacctaacgg accagccct cccagacgc gaccacacct	4740
ggtacacgga tggaagcagt ctcttacaag agggacagcg taaggcggga getgeggtga	4800
ccaccgagac cgaggtaatc tgggctaaag ccctgccagc cgggacatcc getcagcggg	4860
ctgaactgat agcactcacc caggccctaa agatggcaga agtaagaag etaaatgttt	4920
atactgatag ccgttaigtct tttgctactg cccatatcca tggagaaata tacagaaggc	4980
gtgggttget cacatcagaa ggeaaagaga tcaaaaataa agacagatc ttggccctac	5040

taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggtgacca	ageggcccga	aaggcageca	5160
tcacagagac	tccagacacc	tetaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
ataaaaacaaa	gaagtattgg	gtetaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	cateagctga	ctcacctcag	cttetcaaaa	atgaaggctc	5400
tcctagagag	aagccacagt	ccctactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
tactgagac	ctgcaaagct	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggt	ccgcgggcat	cggeccggca	ctcattggga	gategatttc	accgagataa	5580
agcccggatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggctgt	aaccaagaag	ctactagagg	5700
agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgatac	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccate	aaggagactt	5880
[0047] taactaaatt	aacgcttgca	actggetcta	gagactgggt	gctctactc	cccttagccc	5940
tgtaaccgagc	ccgcaacacg	ccgggcccc	atggcctcac	cccataatgag	atcttatatg	6000
gggcaccccc	gccccttgta	aattccctg	accctgacat	gacaagagtt	actaacagcc	6060
ctctctcca	agctcactta	caggetctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcagc	ctaccaagaa	caactggacc	gaccgggtgt	acctaccct	taccgagtcg	6180
ggacacacagt	gtgggtccgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgetgacc	acccccaccg	ccctcaaagt	agacggcctc	gcagcttggga	6300
tacacgccgc	ccacgtgaag	getgccgacc	ccgggggttg	accatctct	agactgacat	6360
ggcgcgttea	acgctctcaa	aacccccca	agataagatt	aacctgtgga	agcccttaat	6420
agtcattggga	gtcctgttag	gagtagggat	ggcagagagc	cccatacagg	tctttaatgt	6480
aacctggaga	gtcaaccaacc	tgatgactgg	gcgtaccgcc	aatgceacct	ccctcctggg	6540
aactgtacaa	gatgecttcc	caaaattata	ttttgateta	tgtgatctgg	tggagagga	6600
gtgggacct	tcagaccagg	aacctatgt	egggtatggc	tgcaagtacc	ccgcagggag	6660
acagcggacc	cggacttttg	acttttacgt	gtgccctggg	cataccgtaa	agtcgggggtg	6720
tgggggacca	ggagagggct	actgtggtaa	atgggggtgt	gaaaccaccg	gacaggctta	6780

	ctggaagecc acatcatcgt gggacctaat ctcccttaag cgcgtaaca ccccttgga	6840
	cacgggatgc tctaaagttg cctgtggccc ctgctacgac ctctccaaag tatecaatte	6900
	cttccaaggg gctactcgag ggggcagatg caacctctta gtctagaat tcaactgatgc	6960
	aggaaaaaag gctaactggg acgggcccac atcgtgggga ctgagactgt accggacagg	7020
	aacagatcct attaccatgt tctccctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
	ccccataggg cccaaccag tattaccga ccaaagactc cttctctcac caatagagat	7140
	tgtaccggct ccacagccac ctgccccct caataccagt taccctctt ccaactaccag	7200
	tacacctca acctccceta caagtccaag tgtcccacag ccaccccag gaactggaga	7260
	tagactacta gctctagtea aaggagccta tcaggcgctt aaectacca atcecgacaa	7320
	gaccaagaa tgttggetgt gettagtgc gggacctctt tattaegaag gagtagcggt	7380
	cgtgggeact tataccaate attccaccgc tccggccaac tgtacggcca ctteccaaca	7440
	taagcttacc ctatetgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
	tcaccaggcc ttatgtaaca ccacccaaag ccggcgctca ggatectact accttgcagc	7560
	accgccgga acaatgtggg cttgcagcac tggattgact cctgtctgt ccaccacggt	7620
[0048]	getcaateta accacagatt attgtgtatt agttgaactc tggeccagag taatttacca	7680
	ctccccgat tatatgtatg gtcagcttga acagegtacc aatataaaa gagagccagt	7740
	atcattgacc ctggcccttc tactaggagg attaaccatg ggagggattg cagetggaat	7800
	agggaacggg accaactgct taattaaaac ccagcagttt gageagcttc atgccgetat	7860
	ccagacagac ctcaacgaag tegaaaagtc aattaccaac ctgaaaaagt cactgaccte	7920
	gttgtctgaa gtagtcctac agaaccgcag aggcctagat ttgetattcc taaaggagg	7980
	aggtctctgc gcagecetaa aagaagaatg ttgtttttat gcagaccaca egggctagt	8040
	gagagacage atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
	aggccaagga tggtttegaag ggtgttttaa tagatcccc tggtttacca cettaatctc	8160
	caccatcatg ggacctctaa tagtaactctt actgatetta ctctttggac cttgcattct	8220
	caatcgatta gtccaatttg ttaaagacag gatatcagtg gtccaggctc tagttttgac	8280
	tcaacaatat caccagctga agcctataga gtacgagcca tgacgtactt tactggccga	8340
	agccgcttgg aataaggccg gtgtgcgttt gtctatatgt tattttccac catattgccg	8400
	tcttttggca atgtgagggc ceggaaacct ggccctgtct tcttgacgag cttctctagg	8460
	ggtcttccc ctctcgccaa aggaatgcaa ggtctgttga atgtctgaa ggaagcagtt	8520

cctctggaag	cttcttgaag	acaaacaacg	tctgtagcga	ccctttgcag	gcagcggaac	8580
ccccacctg	gcgacaggtg	cctctgcggc	caaaagccac	gtgtataaga	tacacctgca	8640
aaggcggeac	aaccccaagt	ccacgttgtg	agttggatag	ttgtggaaag	agtcaaatgg	8700
ctctcctea	gcgtattcaa	caaggggctg	aaggatgccc	agaaggtacc	ccattgtatg	8760
ggatctgac	tggggcctcg	gtgeacatgc	tttacaatgt	tttagtcgag	gttaaaaaaa	8820
egtctaggec	ccccgaacca	cggggacgtg	gttttccttt	gaaaaacacg	ataataccat	8880
ggtgacaggg	ggaatggcaa	gcaagtggga	tcagaagggt	atggacattg	cctatgagga	8940
ggcggcctta	ggttacaaag	agggtggtgt	tcctattggc	ggatgtctta	tcaataacaa	9000
agacggaagt	gttctcggtc	gtggtcacaa	catgagattt	caaaagggat	ccgccacact	9060
acatggtgag	atctccactt	tggaaaactg	tgggagatta	gagggeaaag	tgtacaaaga	9120
taccactttg	tatacgacgc	tgtctccatg	cgacatgtgt	acaggtgcca	tcacatgta	9180
tggattcca	cgctgtgttg	tcggtgagaa	cgtaatttc	aaaagtaagg	gcgagaaata	9240
ttacaaaact	agaggtcacg	aggttgtgt	tgttgacgat	gagaggtgta	aaaagatcat	9300
gaaacaattt	atcgatgaaa	gacctcagga	ttggtttgaa	gatattggtg	agtaggcggc	9360
[0049] cgcgcatag	ataaaataaa	agattttatt	tagtctccag	aaaaaggggg	gaatgaaaga	9420
ccccacctgt	aggtttgcca	agctagctta	agtaacgcca	ttttgcaagg	catggaaaaa	9480
tacataactg	agaatagaga	agttcagate	aaggtcagga	acagatggaa	cagctgaata	9540
tgggccaaac	aggatatctg	tgtaagcag	ttcctgcccc	ggctcagggc	caagaacaga	9600
tggaacagct	gaatatgggc	caaacaggat	atctgtggta	agcagttcct	gccccggcte	9660
agggccaaga	acagatggtc	cccagatgeg	gtccagccct	cagcagtttc	tagagaacca	9720
tcagatgttt	ccagggtgcc	ccaaggacct	gaaatgacce	tgtgcettgt	ttaaactaac	9780
caatcagtte	gcttctcgct	tctgttcgcg	cgtttctgct	ccccgagcte	aataaaagag	9840
cccacaacce	ctcactcggg	gcgccagtc	tccgattgac	tgagtcgccc	gggtaccctg	9900
glatccaata	aacctctttg	cagttgcate	cgacttgtgg	tctcgtgttt	ccttgggagg	9960
gtctctctg	agtgattgac	taccctcag	cgggggtctt	tcatttgggg	gctcgtccgg	10020
gategggaga	cccctgceca	gggaccaccg	accaccacc	gggaggtaa	ctggctgctt	10080
cgcgcgtttc	ggtgatgacg	gtgaaaacct	ctgacatgtg	agcaaaaggc	cagcaaaagg	10140
ccaggaaccg	taaaaaggcc	gcgttgcgtg	cgtttttcca	taggetccgc	ccccctgacg	10200
agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcacagga	ctataaagat	10260

	accaggcggtt	ccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgcccgtta	10320
	ccggataect	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcaa	tgctcacget	10380
	gtaggtatct	cagttcgggtg	taggtcgttc	gctccaagct	ggcctgtgtg	cacgaacccc	10440
	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aaccgcgtaa	10500
	gacacgactt	atgcccaactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	10560
	taggcgggtgc	tacagagttc	ttgaagtgg	ggcctaacta	cgctacact	agaaggacag	10620
	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	10680
	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	10740
	cgcgcaaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgcte	10800
	agtggaaacga	aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	aggatcttca	10860
	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	10920
	cttggctctga	cagttacca	tgettaatca	gtgaggcacc	tatctcagcg	atctgtctat	10980
	ttcgttcate	catagttgcc	tgaactcccc	tcgtgtagat	aactacgata	cgggagggct	11040
	taccatctgg	cccagtgct	gcaatgatac	cgcgagacce	acgctcaccg	gctccagatt	11100
[0050]	tatcagcaat	aaaccagcca	gcgggaaggg	ccgagcgcag	aagtggctct	geaaccttat	11160
	ccgcctccat	ccagttctatt	aattgttgcc	gggaagetag	agtaagtagt	tegccagtta	11220
	atagtttgcg	caacgttggt	gcattgctg	caggcatcgt	ggtgtcacgc	tegtcgtttg	11280
	gtagggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	teccccatgt	11340
	tgtgcaaaaa	agcggttagc	teettcggtc	ctccgatcgt	tgteagaagt	aagttggccg	11400
	cagtgttate	actcatgggt	atggcagcac	tgcataatc	tcttactgtc	atgccatccg	11460
	taagatgett	ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	11520
	ggcgaccgag	ttgctcttgc	ccggcgtaaa	cacgggataa	taccgcgcca	catagcagaa	11580
	ctttaaaagt	gctcatcatt	ggaaaacggt	cttcggggcg	aaaactctca	aggatcttac	11640
	cgtgttgag	atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	11700
	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	11760
	gaataagggc	gacacggaaa	tgttgaatac	tcatactctt	ctttttcaa	tattattgaa	11820
	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	11880
	aaacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	taagaaacca	11940
	ttattateat	gacattaacc	tataaaaaata	ggcgtatcac	gaggeccttt	egtcttcaag	12000

aattcat	12007
<210> 22	
<211> 11893	
<212> DNA	
<213> 人工序列	
<220>	
<223> RCR载体 - pAC3-yCD2	
<400> 22	
tagttattaa tagtaatcaa ttacggggte attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt	120
gacgtcaata atgacgatg tteccatagt aacgccaata ggaactttcc attgacgtca	180
atgggtggag tattttacggt aaactgocca cttggcagta catcaagtgt atcatatgcc	240
aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
attccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
[0051] ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggeg gtaggcgtgt	540
acggtgggag gtctatataa gcagagctgg tttagtgaac cggegccagt cctccgattg	600
actgagtcgc ccgggtaccc gtgtatccaa taaaccctct tgcagttgca tccgacttgt	660
ggtctcgetg ttccttggga gggctcctc tgagtgattg actaccctc agcggggggtc	720
ttcatttgg gggctcgtcc gggatcggga gaccctgcc cagggaccac cgaccacca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgctgcgtc ggtactagtt agctaactag ctctgtatct ggccgaccgc	900
tgggtggaact gacgagttcg gaacaccgg ccgcaaccct gggagacgtc ccagggactt	960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaa ctaaaacagt	1080
tcccgcctcc gtctgaattt ttgcttcgg tttgggaccg aagecgegcc gcgcgtcttg	1140
tctgtgeag catcgttctg tgttctctct gtctgactgt gtttctgtat ttgtctgaaa	1200
atatgggcca gactgttacc actcccttaa gtttgacctt aggtcactgg aaagatgtcg	1260
agcggatcgc tcacaaccag tcggtagatg tcaagaagag acgttgggtt acottctgct	1320
ctgcagaatg gccaaccttt aacgtcggat ggccgcgaga cggcacctt aaccgagacc	1380

	tcataccca	ggttaagafc	aaggtctttt	cacctggccc	gcatggacac	ccagaccagg	1440
	tcccetacat	cgtgacctgg	gaagccttgg	cttttgacce	cceteccctgg	gtcaagccct	1500
	ttgtacacec	taagcctccg	cctcctcttc	ctccatccgc	cccgctctctc	cccccttgaac	1560
	ctcctcgttc	gaccccgctt	cgatcctccc	tttatecagc	cctcaactctt	tctctaggcg	1620
	ccaaacctaa	acctcaagtt	ctttctgaca	gtggggggcc	gctcactgac	ctacttacag	1680
	aagaccccc	gccttatagg	gacccaagac	cacccccttc	cgacagggac	ggaaatggtg	1740
	gagaagcgac	ccctgcggga	gaggeaccgg	acccctcccc	aatggcatct	cgcttacgtg	1800
	ggagacggga	gccccctgtg	gcccactcca	ctacctcgca	ggcattcccc	ctccgcgcag	1860
	gaggaaaacg	acagcttcaa	tactggccgt	tctcctcttc	tgaccttac	aactggaaaa	1920
	ataataacce	ttctttttct	gaagatccag	gtaaaactgac	agctctgac	gagctctgctc	1980
	tcataccca	tcagcccacc	tgggacgact	gtcagcagct	gttggggact	ctgctgaccg	2040
	gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaagc	ggtgcggggc	gatgatgggc	2100
	gccccactca	actgcccact	gaagtcgatg	ccgcttttcc	cctcgagcgc	ccagactggg	2160
	attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgcccagttg	ctcctagcgg	2220
[0052]	gtctccaaaa	cgcgggcaga	ageccccaca	atttggccaa	ggtaaaagga	ataacacaag	2280
	ggcccactga	gtctccctcg	gccttcttag	agagacttaa	ggaagcctat	cgcagggtaca	2340
	ctccttatga	ccctgaggac	ccaggccaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
	agtctgcccc	agacattggg	agaaagttag	agaggttaga	agatttaaaa	aacaagacgc	2460
	ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaacc	ccggaagaaa	2520
	gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
	agcagaaaga	gaaagaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctattggcca	2640
	ctgtcgttag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccactcgc	2700
	atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggcactg	ggctaaagat	tgtcccaga	2760
	aaccacgagg	acctcgggga	ccaagacccc	agacctcctt	cctgacctta	gatgactagg	2820
	gagtcaggg	tcaggagccc	ccccctgaac	ccaggataac	cctcaaagtc	ggggggcaac	2880
	ccgtcacctt	cctggtagat	actggggccc	aacctcctt	gtgacctcaa	aatcctggac	2940
	ccctaagtga	taagtctgcc	tgggtccaag	gggctactgg	aggaaagcgg	tatcgtgga	3000
	ccacggatcg	caaagtacat	ctagetaccg	gtaaggtcac	ccaetcttct	ctccatgtac	3060
	cagaactgtc	ctatectctg	ttaggaagag	atttgcctgac	taaaactaaa	gccccaatcc	3120

actttgagg	atcaggagcc	caggttatgg	gaccaatggg	gcagccctg	caagtgttga	3180
ccctaaatat	agaagatgag	tateggctac	atgagacctc	aaaagagcca	gatgtttctc	3240
tagggctccac	atggctgtct	gattttcctc	aggcctgggc	ggaaaccggg	ggcatgggac	3300
tggcagttcg	ccaagctcct	ctgatcatac	ctctgaaagc	aacctctacc	cccgtgtcca	3360
taaaacaata	ccccatgtca	caagaagcca	gactggggat	caagccccac	atacagagac	3420
tgttggaaca	gggaatactg	gtaccctgcc	agtcctccctg	gaacacgccc	ctgctacccg	3480
ttaagaaacc	agggactaat	gattatagge	ctgtccagga	tctgagagaa	gtcaacaagc	3540
gggtggaaga	catccacccc	accgtgcccc	acccttacia	cctcttgagc	gggctcccc	3600
cgteccacca	gtggtacact	gtgcttgatt	taaaggatgc	ctttttctgc	ctgagactcc	3660
accccaccag	tcagcctctc	ttegectttg	agtggagaga	tccagagatg	ggaatctcag	3720
gacaattgac	ctggaccaga	ctccacacag	gtttcaaaaa	cagteccacc	ctgtttgatg	3780
aggcactgca	cagagaccta	gcagacttcc	ggatccagca	cccagacttg	atcctgctac	3840
agtacgtgga	tgacttactg	ctggccgcca	cttctgagct	agactgcca	caaggtactc	3900
gggcccctgtt	acaaacccta	gggaacctcg	ggtatcgggc	ctcggcccaag	aaagccccaa	3960
[0053] tttgccagaa	acagggtcaag	tatctggggg	atctttctaaa	agagggtcag	agatggctga	4020
ctgaggccag	aaaagagact	gtgatggggc	agcctactcc	gaagaccctc	cgacaactaa	4080
gggagtccct	agggacggca	ggettctgtc	gcctctggat	ccctgggttt	gcagaaatgg	4140
cagccccctt	gtaccctctc	accaaaacgg	ggactctggt	taattggggc	ccagaccaac	4200
aaaaggccta	tcaagaaatc	aagcaagctc	ttctaactgc	cccagccctg	gggttgccag	4260
atttgactaa	gccctttgaa	ctctttgtcg	acgagaagea	gggctacgcc	aaagggtgctc	4320
taacgcaaaa	actgggacct	tggcgtcggc	cggtggccta	cctgtccaaa	aagctagacc	4380
cagtagcagc	tgggtggccc	ccttgccctac	ggatggtagc	agccattgcc	gtactgacaa	4440
aggatgcagg	caagctaacc	atgggacagc	cactagtcac	tctggcccc	catgcagtag	4500
aggcactagt	caaacaaccc	cccgaccgct	ggctttccaa	cgcceggatg	actcactatc	4560
aggccttget	tttgacacg	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaaccocg	4620
ctacgtget	cccactgcct	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccc	4680
aagcccacgg	aaccggacc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcgggtga	4800
ccaccgagac	cgaggtaatc	tgggetaaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860

etgaactgat	agcaetcacc	caggecctaa	agatggcaga	aggtaagaag	ctaaatgttt	4920
atactgatag	ccgttatgct	tttgctactg	cccatatcca	tggagaaata	tacagaaggc	4980
gtgggttget	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
taaaagccct	ctttetgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaagg	5100
gacacagcgc	cgaggctaga	ggcaaccgga	tggetgacca	agcgccccga	aaggcagcca	5160
tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
aacattttca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccattttatg	5280
ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggcte	5400
tcttagagag	aagccacagt	ccctactaca	tgctgaaccg	ggategaaca	ctcaaaaata	5460
tactgagac	ctgcaaaget	tgtgcacaag	tcaacgccag	caagtctgcc	gttaaacagg	5520
gaactagggt	ccgcgggcat	cggccccgca	ctcattggga	gategatttc	accgagataa	5580
agccccgatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
tagaagcett	cccaaccaag	aaagaaaaccg	ccaaggctgt	aaccaagaag	ctactagagg	5700
[0054] agatcttccc	caggttcgge	atgeetcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
tctccaaggt	gagtcagaca	gtggccgate	tgttggggat	tgattggaaa	ttacattgtg	5820
catacagacc	ccaagctca	ggccaggtag	aaagaatgaa	tagaacctc	aaggagactt	5880
taactaaatt	aacgettgea	actggcteta	gagactgggt	getcctactc	cccttagccc	5940
tgtaccgagc	ccgcaacaeg	cegggcccc	atggcctcac	cccataatgag	atcttatatg	6000
gggcaccccc	gccccctgta	aacttccctg	accctgacat	gacaagagtt	actaacagcc	6060
ccctctctca	agctcaetta	caggtctct	acttagtcca	gcacgaagtc	tggagacctc	6120
tggcggcagc	ctaccaagaa	caactggacc	gaccggtggt	acctcaccct	taccgagtctg	6180
gcgacacagt	gtgggtctgc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
cttacacagt	cctgctgacc	acccccaccg	ccctcaagt	agaeggcatc	gcagcttgga	6300
tacacgcgc	ccacgtgaag	getgecgacc	cegggggttg	accatcctct	agaactgacat	6360
ggcgcgttca	acgctctcaa	aacccccctca	agataagatt	aacctgtgga	agcccttaat	6420
agtcattgga	gtcctgttag	gagtagggat	ggcagagagc	ccccatcagg	tctttaatgt	6480
aaactggaga	gtcaccaacc	tgatgactgg	gcgtaccgcc	aatgccacct	ccctctctggg	6540
aactgtacaa	gatgccttcc	caaaattata	ttttgatcta	tgtgatctgg	tccgagagga	6600

gtgggacct	tcagaccagg	aaccgtatgt	cgggtatggc	tgcaagtacc	ccgcagggag	6660
acagcggacc	cggacttttg	acttttacgt	gtgccctggg	cataccgtaa	agtcgggggtg	6720
tgggggacca	ggagagggct	actgtggtaa	atgggggtgt	gaaaccaccg	gacaggctta	6780
ctggaagccc	acatcatcgt	gggacctaat	ctcccttaag	cgcggtaaca	ccccctggga	6840
cacgggatgc	tctaaagttg	cctgtggccc	ctgctacgac	ctctccaaag	tatccaattc	6900
cttccaaggg	gctactcgag	gggceagatg	caaccctcta	gtcctagaat	tactgatgc	6960
aggaaaaaag	gctaactggg	acgggcccac	atcgtgggga	ctgagactgt	accggacagg	7020
aacagatcct	attaccatgt	tctccctgac	ccggcaggte	cttaatgtgg	gacccccgagt	7080
ccccataggg	cccaccccag	tattaccoga	ccaaagactc	ccttctcac	caatagagat	7140
tgtaccggct	ccacagccac	ctageccccct	caataaccagt	tacccccctt	ccactaccag	7200
tacaccctca	acctccccta	caagtccaag	tgteccacag	ccaaccccag	gaactggaga	7260
tagactacta	gtctagtca	aaggagccta	tcaggcgctt	aacctacca	atcccgacaa	7320
gacccaagaa	tgttggctgt	gcttagtgtc	gggacctcct	tattacgaag	gagtagcggc	7380
cgtagggcact	tataccaatc	attccaccgc	tccggccaac	tgtaaggcca	cttcccaca	7440
[0055] taagettacc	ctatctgaag	tgacaggaca	gggcctatgc	atgggggcag	tacctaaaac	7500
tcaccaggcc	ttatgtaaca	ccacccaaag	cgccggctca	ggatctact	accttgcagc	7560
accgccgga	acaatgtggg	cttgcagcac	tggattgact	ccctgcttgt	ccaccacggt	7620
gtcaateta	accacagatt	attgtgtatt	agttgaactc	tggeccagag	taatttacca	7680
ctccccgat	tatatgtatg	gtcagcttga	acagcgtacc	aaatataaaa	gagagccagt	7740
atcattgacc	ctggcccctc	tactaggagg	attaaccatg	ggagggattg	cagctggaat	7800
aggacgggg	accactgctt	taattaaaac	ccagcagttt	gagcagcttc	atgccctat	7860
ccagacagac	ctcaacgaag	tegaaaagtc	aattaccaac	ctagaaaagt	cactgacctc	7920
gttgtctgaa	gtagtcctac	agaaccgcag	aggcctagat	ttgctattcc	taaaggaggg	7980
aggtctctgc	gcagccctaa	aagaagaatg	ttgtttttat	gcagaccaca	cggggctagt	8040
gagagacagc	atggccaaat	taagagaaag	gcttaateag	agacaaaaac	tatttgagac	8100
aggccaagga	tggttcgaag	ggctgtttaa	tagatcccc	tggtttacca	ccttaatctc	8160
caccatcatg	ggacctctaa	tagtactctt	actgatctta	ctctttggac	cttgcattct	8220
caatcgattg	gtccaatttg	ttaaagacag	gatctcagtg	gtccaggctc	tggttttgac	8280
tcagcaatat	caccagctaa	aacctataga	gtacgagcca	tgaacgcgtt	actggccgaa	8340

	gccgcttggg	ataaggccgg	tgtgcgtttg	tctatatgtt	atthccacc	atattgccgt	8400
	cttttgcaa	tgtgagggcc	cggaaacctg	gccctgtctt	cttgacgagc	attcctaggg	8460
	gtctttcccc	tctcgccaaa	ggaatgcaag	gtctgttgaa	tgctgtgaag	gaagcagttc	8520
	ctctggaagc	ttcttgaaga	caaacaacgt	ctgtagegac	cccttgcagg	cagcggaaac	8580
	ccccacctgg	cgacaggtgc	ctctcgggcc	aaaagccacg	tgtataagat	acacctgcaa	8640
	aggcggcaca	acccccagtgc	cacgttgtga	gttgatagat	tgtggaaaga	gtcaaatggc	8700
	tctcctcaag	cgtattcaac	aaggggctga	aggatgcccc	gaaggtacc	cattgtatgg	8760
	gatctgatct	ggggcctcgg	tgacatgct	ttacatgtgt	ttagtcgagg	ttaaaaaac	8820
	gtctaggccc	cccgaaccac	ggggacgtgg	tttctctttg	aaaaacacga	ttataaatgg	8880
	tgaccggcgg	catggcctcc	aagtgggate	aaaagggcat	ggatategct	tacgaggagg	8940
	ccctgctggg	ctacaaggag	ggcggcgtgc	ctatcgccgg	ctgtctgac	aacaacaagg	9000
	acggcagtgt	gttgggcagg	ggccacaaca	tgaggttcca	gaagggtcc	gccacctgc	9060
	acggcgagat	ctccacctg	gagaactgtg	gcaggctgga	gggcaagggt	tacaaggaca	9120
	ccacctgtga	caccacctg	tcccctgtg	acatgtgtac	cggegetate	atcatgtacg	9180
[0056]	geatecctag	gtgtgtgate	ggegagaacg	tgaacttcaa	gtccaagggc	gagaagtacc	9240
	tgcaaaccag	gggccacgag	gtggtggttg	ttgacgatga	gaggtgtaag	aagctgatga	9300
	agcagttcat	cgacgagagg	cctcaggact	ggttcgagga	tateggcgag	taagcggccg	9360
	cagataaaat	aaaagattht	atttagtctc	cagaaaaagg	ggggaatgaa	agacccccacc	9420
	tgtaggtht	gcaagctagc	ttaagtaacg	ccatthtga	aggcatggaa	aaatacataa	9480
	ctgagaatag	agaagttcag	atcaaggtea	ggaacagatg	gaacagctga	atatgggcca	9540
	aacaggataf	ctgtggtaag	cagttcctgc	cccggctcag	ggccaagaac	agatggaaca	9600
	gtgaatatg	ggccaacag	gatatctgtg	gtaagcagtt	cctgccccgg	ctcagggcca	9660
	agaacagatg	gtccccagat	gcggtccagc	cctcagcagt	ttctagagaa	ccatcagatg	9720
	thtccagggt	gccccaaagga	cctgaaatga	ccctgtgcct	tatttgaact	aaccaatcag	9780
	thtgettctc	gcttctgttc	gegegttct	gtccccgag	ctcaataaaa	gagccccaca	9840
	cccccactc	ggggegccag	tctctcgatt	gactgagteg	cccgggtacc	cgtgtatcca	9900
	ataaacctc	ttgcagttgc	atcagacttg	tggtctcgct	gttcttggg	agggtctcct	9960
	ctgagtgatt	gactacctgt	cagcgggggt	ctttcattac	atgtgagcaa	aaggccagca	10020
	aaaggccagg	aaccgtaaaa	aggcgcggtt	gctggcgtht	thtcatagge	thtcccccc	10080

	tgacgagcat cacaaaaafc gacgctcaag tcagaggtgg cgaaaccgga caggactata	10140
	aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc cgaccctgce	10200
	gettaccgga tacctgtccg cttttctccc ttcggaagc gtggcgettt ctcatagctc	10260
	acgetgtagg tatctcagtt egggttaggt cgttcgctcc aagctgggct gtgtgcacga	10320
	acccccgtt cagcccgacc gettgcctt atccggtaac tategtcttg agtccaaccc	10380
	ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag	10440
	gtagtaggc ggtgctacag agttcttgaa gtgggtggcct aactacggct aactagaag	10500
	gacagtatft ggtatctgcg ctctgctgaa gccagttacc ttcggaaaa gagttggtag	10560
	ctcttgatcc ggcaaacaaa ccaccgctgg tagcgggtgt tttttgttt gcaagcagca	10620
	gattacgcgc agaaaaaag gatctcaaga agatcctttg atctttteta cggggctctga	10680
	cgtcagtggt aacgaaaact cacgitaagg gatthtggtc atgagattat caaaaaggat	10740
	cttacctag atccttttaa attaaaaatg aagthttaa tcaatetaaa gtatatatga	10800
	gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct cagcgatctg	10860
	tctatttctg tcatccatag ttgctgact ccccgctgtg tagataacta cgatacggga	10920
[0057]	gggettacea tctggcccc gtgetgcaat gataccgga gaccacgct caccgctcc	10980
	agatttatea gcaataaacc agccagccgg aaggcccgag cgcagaagt gtctgcaac	11040
	ttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa gtagttegcc	11100
	agttaatagt ttgcgcaac ttgttccat tgetgcaggc atcgtggtgt cacgctcgtc	11160
	gthtggtatg gcttcattca gctccggtc ccaacgatca aggegagta catgatcccc	11220
	catgttgtgc aaaaaagcgg ttagctcctt cggctctcag atcgttgtca gaagtaagtt	11280
	ggccgcagtg ttatactca tggttatggc agcactgat aattctctta ctgtcatgcc	11340
	atccgtaaga tgctttctg tgaactgtga gtactcaacc aagtcattct gagaatagtg	11400
	tatgcggcga ccgagttgct ctgcccggc gtcaacacgg gataataccg cgccacatag	11460
	cagaacttta aaagtgetca tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	11520
	cttaccgctg ttgagatcca gttgatgta acccaactgt gaccccaact gatcttcagc	11580
	atcttttaet ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa	11640
	aaaggaata agggcgacac ggaaatgtg aatactcata ctcttcttt ttcaatatta	11700
	ttgaagcatt tatcagggtt attgtctcat gageggatac atatttgaat gtatttagaa	11760
	aaataaacaa ataggggttc cgegeacatt tccccgaaaa gtgccacctg acgtctaaga	11820

	aaccattatt atcatgacat taacctataa aataggggt atcacgagge cetttegtct	11880
	tcaagaattc cat	11893
	<210> 23	
	<211> 4473	
	<212> DNA	
	<213> 智人	
	<220>	
	<221> CDS	
	<222> (175).. (3942)	
	<400> 23	
	aaggggaggt aacctggcc cctttggteg gggcccggg cagccgcgcg ccccttccca	60
	eggggccctt tactgcgcgc cgcgccggc cccaccctt cgcagcacc cgcgcgccgc	120
	gccctccag cgggtccag cggagccat ggggccggag ccgcagtgag cacc atg	177
		Met
		1
	gag ctg gcg gcc ttg tgc cgc tgg ggg etc etc etc gcc etc ttg ecc	225
	Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu Pro	
	5 10 15	
[0058]	ccc gga gcc gcg age acc caa gtg tgc acc ggc aca gac atg aag ctg	273
	Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys Leu	
	20 25 30	
	egg etc cct gcc agt ccc gag acc cac ctg gac atg etc cgc cac etc	321
	Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His Leu	
	35 40 45	
	tac cag gcc tgc cag gtg gtg cag gga aac ctg gaa etc acc tac ctg	369
	Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr Leu	
	50 55 60 65	
	ccc acc aat gcc age ctg tcc ttc ctg cag gat atc cag gag gtg cag	417
	Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val Gln	
	70 75 80	
	ggc tac gtg etc atc gct cac aac caa gtg agg cag gtc cca ctg cag	465
	Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln	
	85 90 95	
	agg ctg cgg att gtg cga ggc acc cag etc ttt gag gac aac tat gcc	513
	Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr Ala	
	100 105 110	
	ctg gcc gtg cta gac aat gga gac ccg ctg aac aat acc acc cct gtc	561
	Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro Val	
	115 120 125	
	aca ggg gcc tcc cca gga ggc ctg cgg gag ctg cag ett cga age etc	609
	Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser Leu	

130	135	140	145	
aca gag atc ttg aaa gga ggg gtc ttg atc cag egg aac ccc cag ctc				657
Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln Leu	150	155	160	
tgc tac cag gac acg att ttg tgg aag gac atc ttc cac aag aac aac				705
Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn Asn	165	170	175	
cag ctg gct ctc aca ctg ata gac acc aac cgc tct cgg gcc tgc cac				753
Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys His	180	185	190	
ccc tgt tet ccg atg tgt aag ggc tcc cgc tgc tgg gga gag agt tet				801
Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser Ser	195	200	205	
gag gat tgt cag agc ctg acg cgc act gtc tgt gcc ggt gcc tgt gcc				849
Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys Ala	210	215	220	225
cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat gag cag tgt gct				897
Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys Ala	230	235	240	
gcc gcc tgc acg gcc ccc aag cac tct gac tgc ctg gcc tgc ctc cac				945
Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu His	245	250	255	
[0059]				
ttc aac cac agt gcc atc tgt gag ctg cac tgc cca gcc ctg gtc acc				993
Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val Thr	260	265	270	
tac aac aca gac acg ttt gag tcc atg ccc aat ccc gag gcc cgg tat				1041
Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg Tyr	275	280	285	
aca ttc gcc gcc agc tgt gtg act gcc tgt ccc tac aac tac ctt tct				1089
Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu Ser	290	295	300	305
acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg cac aac caa gag				1137
Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln Glu	310	315	320	
gtg aca gca gag gat gga aca cag cgg tgt gag aag tgc agc aag ccc				1185
Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys Pro	325	330	335	
tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac ttg cga gag gtg				1233
Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu Val	340	345	350	
agg gca gtt acc agt gcc aat atc cag gag ttt gct gcc tgc aag aag				1281
Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys Lys	355	360	365	

	atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt gat ggg gac cca Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp Pro 370 375 380 385	1329
	gcc tcc aac act gcc ccg etc cag cca gag cag etc caa gtg ttt gag Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe Glu 390 395 400	1377
	act ctg gaa gag atc aca ggt tac cta tac atc tca gca tgg ccg gac Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro Asp 405 410 415	1425
	agc ctg cct gac etc agc gtc ttc cag aac ctg caa gta atc cgg gga Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg Gly 420 425 430	1473
	cga att ctg cac aat ggc gcc tac tgg ctg acc ctg caa ggg ctg ggc Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu Gly 435 440 445	1521
	atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg ggc agt gga ctg Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly Leu 450 455 460 465	1569
	gcc etc atc cac cat aac acc cac etc tgc ttc gtg cac acg gtg ecc Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val Pro 470 475 480	1617
[0060]	tgg gac cag etc ttt cgg aac ccg cac caa gct ctg etc cae act gcc Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr Ala 485 490 495	1665
	aac cgg cca gag gac gag tgt gtg ggc gag ggc ctg gcc tgc cac cag Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His Gln 500 505 510	1713
	ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc acc cag tgt gtc Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys Val 515 520 525	1761
	aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg gag gaa tgc cga Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys Arg 530 535 540 545	1809
	gta ctg cag ggg etc ccc agg gag tat gtg aat gcc agg cac tgt ttg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu 550 555 560	1857
	ccg tgc cac cct gag tgt eag ccc cag aat ggc tca gtg acc tgt ttt Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe 565 570 575	1905
	gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac tat aag gac cct Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro 580 585 590	1953
	ccc ttc tgc gtg gcc cgc tgc ccc age ggt gtg aaa cct gac etc tcc Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser 2001	2001

595	600	605	
tac atg ccc atc tgg aag ttt cca gat gag gag ggc gca tgc cag cct Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro 610 615 620 625			2049
tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg gat gac aag ggc Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly 630 635 640			2097
tgc ccc gcc gag cag aga gcc age cct ctg acg tcc atc atc tct gcg Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser Ala 645 650 655			2145
gtg gtt ggc att ctg ctg gtc gtg gtc ttg ggg gtg gtc ttt ggg atc Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly Ile 660 665 670			2193
ctc atc aag cga cgg cag cag aag atc cgg aag tac acg atg cgg aga Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg 675 680 685			2241
ctg ctg cag gaa acg gag ctg gtg gag ccg ctg aca cct agc gga gcg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala 690 695 700 705			2289
atg ccc aac cag gcg cag atg cgg atc ctg aaa gag acg gag ctg agg Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg [0061] 710 715 720			2337
aag gtg aag gtg ctt gga tct ggc gct ttt ggc aca gtc tac aag ggc Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly 725 730 735			2385
atc tgg atc cct gat ggg gag aat gtg aaa att cca gtg gcc atc aaa Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys 740 745 750			2433
gtg ttg agg gaa aac aca tcc ccc aaa gcc aac aaa gaa atc tta gac Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp 755 760 765			2481
gaa gca tac gtg atg gct ggt gtg ggc tcc cca tat gtc tcc cgc ctt Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu 770 775 780 785			2529
ctg ggc atc tgc ctg aca tcc acg gtg cag ctg gtg aca cag ctt atg Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met 790 795 800			2577
ccc tat ggc tgc etc tta gac cat gtc cgg gaa aac cgc gga cgc ctg Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu 805 810 815			2625
ggc tcc cag gac ctg ctg aac tgg tgt atg cag att gcc aag ggg atg Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met 820 825 830			2673

	agc tac ctg gag gat gtg cgg ctc gta cac agg gac ttg gcc get cgg Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg 835 840 845	2721
	aac gtg ctg gtc aag agt ccc aac cat gtc aaa att aca gac ttc ggg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly 850 855 860 865	2769
	ctg gct cgg ctg ctg gac att gac gag aca gag tac cat gca gat ggg Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly 870 875 880	2817
	ggc aag gtg ccc atc aag tgg atg gcg ctg gag tcc att ctc cgc cgg Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg Arg 885 890 895	2865
	cgg ttc acc cac cag agt gat gtg tgg agt tat ggt gtg act gtg tgg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp 900 905 910	2913
	gag ctg atg act ttt ggg gcc aaa cct tac gat ggg atc cca gcc cgg Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala Arg 915 920 925	2961
	gag atc cct gac ctg ctg gaa aag ggg gag cgg ctg ccc cag ccc ccc Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro 930 935 940 945	3009
[0062]	atc tgc acc att gat gtc tac atg atc atg gtc aaa tgt tgg atg att Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile 950 955 960	3057
	gac tet gaa tgt cgg cca aga ttc cgg gag ttg gtg tet gaa ttc tcc Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe Ser 965 970 975	3105
	cgc atg gcc agg gac ccc cag cgc ttt gtg gtc atc cag aat gag gac Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu Asp 980 985 990	3153
	ttg ggc cca gcc agt ccc ttg gac agc acc ttc tac cgc tca ctg ctg Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu Leu 995 1000 1005	3201
	gag gac gat gac atg ggg gac ctg gtg gat gct gag gag tat ctg Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr Leu 1010 1015 1020	3246
	gta ccc cag cag ggc ttc ttc tgt cca gac cct gcc ccg ggc get Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly Ala 1025 1030 1035	3291
	ggg ggc atg gtc cac cae agg cac cgc age tca tct acc agg agt Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg Ser 1040 1045 1050	3336
	ggc ggt ggg gac ctg aca eta ggg ctg gag ccc tct gaa gag gag Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu Glu 1055 1060 1065	3381

1055	1060	1065	
gcc ccc agg tct cca ctg gca ccc tcc gaa ggg gct ggc tcc gat Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser Asp 1070	1075	1080	3426
gta ttt gat ggt gac ctg gga atg ggg gca gcc aag ggg ctg caa Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu Gln 1085	1090	1095	3471
agc ctc ccc aca cat gac ccc agc cct cta cag cgg tac agt gag Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser Glu 1100	1105	1110	3516
gac ccc aca gta ccc ctg ccc tct gag act gat ggc tac gtt gcc Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val Ala 1115	1120	1125	3561
ccc ctg acc tgc age ccc cag cct gaa tat gtg aac cag cca gat Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro Asp 1130	1135	1140	3606
gtt cgg ccc cag ccc cct tcg ccc cga gag ggc cct ctg cct gct Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro Ala 1145	1150	1155	3651
gcc cga cct gct ggt gcc act ctg gaa agg ccc aag act ctc tcc Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro Lys Thr Leu Ser [0063] 1160	1165	1170	3696
cca ggg aag aat ggg gtc gtc aaa gac gtt ttt gcc ttt ggg ggt Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala Phe Gly Gly 1175	1180	1185	3741
gcc gtg gag aac ccc gag tac ttg aca ccc cag gga gga gct gcc Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly Ala Ala 1190	1195	1200	3786
cct cag ccc cac cct cct cct gcc ttc agc cea gcc ttc gac aac Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp Asn 1205	1210	1215	3831
ctc tat tac tgg gac cag gac cea cea gag cgg ggg gct cea ccc Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro Pro 1220	1225	1230	3876
agc acc ttc aaa ggg aca cct acg gca gag aac cea gag tac ctg Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr Leu 1235	1240	1245	3921
ggt ctg gac gtg cea gtg tga accagaagge caagtccgca gaagccctga Gly Leu Asp Val Pro Val 1250	1255		3972
tgtgtcctca gggagcaggg aaggcctgac ttctgtctggc atcaagaggt gggagggccc			4032
tecgaccaact tccaggggaa cctgccaatgc caggaaacctg tccaaaggaa ccttccttcc			4092

[0064]

tgcttgagtt cccagatggc tggaaggggt ccagcctcgt tggaagagga acagcaactgg 4152
 ggagtctttg tggattctga ggcctgccc aatgagacte tagggtccag tggatgccac 4212
 agcccagett ggcctttcc ttccagatec tgggtactga aagccttagg gaagctggcc 4272
 tgagagggga agcggcccta agggagtgtc taagaacaaa agcgacccat tcagagactg 4332
 tcctgaaac ctagtaactgc ccccatgag gaaggaacag caatggtgtc agtatccagg 4392
 cttgtacag agtgettttc tgtttagttt ttactttttt tgttttgttt ttttaaagat 4452
 gaaataaaga cccaggggga g 4473

<210> 24

<211> 1255

<212> PRT

<213> 智人

<400> 24

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
 1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
 20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
 35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
 50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
 65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
 85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
 100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
 115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
 130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
 145 150 155 160
 Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
 165 170 175
 Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
 180 185 190
 His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
 195 200 205
 Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
 210 215 220
 Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
 225 230 235 240
 Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
 245 250 255
 [0065] His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
 260 265 270
 Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
 275 280 285
 Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
 290 295 300
 Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
 305 310 315 320
 Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
 325 330 335
 Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
 340 345 350
 Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
 355 360 365
 Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
 370 375 380

Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385 390 395 400

Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
405 410 415

Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
420 425 430

Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
435 440 445

Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
450 455 460

Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465 470 475 480

Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
485 490 495

[0066]

Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
500 505 510

Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
515 520 525

Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
530 535 540

Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545 550 555 560

Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
565 570 575

Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
580 585 590

Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
595 600 605

Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
 610 615 620

Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
 625 630 635 640

Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Ile Ser
 645 650 655

Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
 660 665 670

Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
 675 680 685

Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
 690 695 700

Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
 705 710 715 720

[0067] Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
 725 730 735

Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
 740 745 750

Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
 755 760 765

Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
 770 775 780

Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
 785 790 795 800

Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
 805 810 815

Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
 820 825 830

Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
 835 840 845

Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
850 855 860

Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865 870 875 880

Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895

Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910

Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925

Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940

Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960

[0068]

Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975

Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990

Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005

Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr
1010 1015 1020

Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly
1025 1030 1035

Ala Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg
1040 1045 1050

Ser Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu
1055 1060 1065

Glu Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser
1070 1075 1080

Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu
1085 1090 1095

Gln Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser
1100 1105 1110

Glu Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val
1115 1120 1125

Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro
1130 1135 1140

Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro
1145 1150 1155

Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Pro Lys Thr Leu
1160 1165 1170

[0069]

Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala Phe Gly
1175 1180 1185

Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly Ala
1190 1195 1200

Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp
1205 1210 1215

Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro
1220 1225 1230

Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr
1235 1240 1245

Leu Gly Leu Asp Val Pro Val
1250 1255

<210> 25
<211> 1212
<212> DNA
<213> 智人

<220>		
<221>	CDS	
<222>	(1)..(1212)	
<400>	25	
atg aca gcc atc atc aaa gag atc gtt agc aga aac aaa agg aga tat		48
Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr		
1 5 10 15		
caa gag gat gga ttc gac tta gac ttg acc tat att tat cca aac att		96
Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile		
20 25 30		
att gct atg gga ttt cct gca gaa aga ctt gaa ggc gta tac agg aac		144
Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn		
35 40 45		
aat att gat gat gta gta agg ttt ttg gat tca aag cat aaa aac cat		192
Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys His Lys Asn His		
50 55 60		
tac aag ata tac aat ctt tgt gct gaa aga cat tat gac acc gcc aaa		240
Tyr Lys Ile Tyr Asn Leu Cys Ala Glu Arg His Tyr Asp Thr Ala Lys		
65 70 75 80		
ttt aat tgc aga gtt gca caa tat cct ttt gaa gac cat aac cca cca		288
Phe Asn Cys Arg Val Ala Gln Tyr Pro Phe Glu Asp His Asn Pro Pro		
85 90 95		
[0070] cag cta gaa ctt atc aaa ccc ttt tgt gaa gat ctt gac caa tgg cta		336
Gln Leu Glu Leu Ile Lys Pro Phe Cys Glu Asp Leu Asp Gln Trp Leu		
100 105 110		
agt gaa gat gac aat cat gtt gca gca att cac tgt aaa gct gga aag		384
Ser Glu Asp Asp Asn His Val Ala Ala Ile His Cys Lys Ala Gly Lys		
115 120 125		
gga cga act ggt gta atg ata tgt gca tat tta tta cat cgg ggc aaa		432
Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg Gly Lys		
130 135 140		
ttt tta aag gca caa gag gcc cta gat ttc tat ggg gaa gta agg acc		480
Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val Arg Thr		
145 150 155 160		
aga gac aaa aag gga gta act att ccc agt cag agg cgc tat gtg tat		528
Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr Val Tyr		
165 170 175		
tat tat agc tac ctg tta aag aat cat ctg gat tat aga cca gtg gca		576
Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala		
180 185 190		
ctg ttg ttt cac aag atg atg ttt gaa act att cca atg ttc agt ggc		624
Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly		
195 200 205		
gga act tgc aat cct cag ttt gtg gtc tgc cag cta aag gtg aag ata		672

Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile 210 215 220	
tat tcc tcc aat tca gga ccc aca cga cgg gaa gac aag ttc atg tac Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr 225 230 235 240	720
ttt gag ttc cct cag ccg tta cct gtg tgt ggt gat atc aaa gta gag Phe Glu Phe Pro Gln Pro Leu Pro Val Cys Gly Asp Ile Lys Val Glu 245 250 255	768
ttc ttc cac aaa cag aac aag atg cta aaa aag gac aaa atg ttt cac Phe Phe His Lys Gln Asn Lys Met Leu Lys Lys Asp Lys Met Phe His 260 265 270	816
ttt tgg gta aat aca ttc ttc ata cca gga cca gag gaa acc tca gaa Phe Trp Val Asn Thr Phe Phe Ile Pro Gly Pro Glu Glu Thr Ser Glu 275 280 285	864
aaa gta gaa aat gga agt cta tgt gat caa gaa atc gat agc att tgc Lys Val Glu Asn Gly Ser Leu Cys Asp Gln Glu Ile Asp Ser Ile Cys 290 295 300	912
agt ata gag cgt gca gat aat gac aag gaa tat cta gta ctt act tta Ser Ile Glu Arg Ala Asp Asn Asp Lys Glu Tyr Leu Val Leu Thr Leu 305 310 315 320	960
aca aaa aat gat ctt gac aaa gca aat aaa gac aaa gcc aac cga tac [0071] Thr Lys Asn Asp Leu Asp Lys Ala Asn Lys Asp Lys Ala Asn Arg Tyr 325 330 335	1008
ttt tct cca aat ttt aag gtg aag ctg tac ttc aca aaa aca gta gag Phe Ser Pro Asn Phe Lys Val Lys Leu Tyr Phe Thr Lys Thr Val Glu 340 345 350	1056
gag ccg tca aat cca gag gct agc agt tca act tct gta aca cca gat Glu Pro Ser Asn Pro Glu Ala Ser Ser Ser Thr Ser Val Thr Pro Asp 355 360 365	1104
gtt agt gac aat gaa cct gat cat tat aga tat tct gac acc act gac Val Ser Asp Asn Glu Pro Asp His Tyr Arg Tyr Ser Asp Thr Thr Asp 370 375 380	1152
tct gat cca gag aat gaa cct ttt gat gaa gat cag cat aca caa att Ser Asp Pro Glu Asn Glu Pro Phe Asp Glu Asp Gln His Thr Gln Ile 385 390 395 400	1200
aca aaa gtc tga Thr Lys Val	1212

<210> 26
 <211> 403
 <212> PRT
 <213> 智人

 <400> 26

Met Thr Ala Ile Ile Lys Glu Ile Val Ser Arg Asn Lys Arg Arg Tyr
 1 5 10 15
 Gln Glu Asp Gly Phe Asp Leu Asp Leu Thr Tyr Ile Tyr Pro Asn Ile
 20 25 30
 Ile Ala Met Gly Phe Pro Ala Glu Arg Leu Glu Gly Val Tyr Arg Asn
 35 40 45
 Asn Ile Asp Asp Val Val Arg Phe Leu Asp Ser Lys His Lys Asn His
 50 55 60
 Tyr Lys Ile Tyr Asn Leu Cys Ala Glu Arg His Tyr Asp Thr Ala Lys
 65 70 75 80
 Phe Asn Cys Arg Val Ala Gln Tyr Pro Phe Glu Asp His Asn Pro Pro
 85 90 95
 Gln Leu Glu Leu Ile Lys Pro Phe Cys Glu Asp Leu Asp Gln Trp Leu
 100 105 110
 [0072] Ser Glu Asp Asp Asn His Val Ala Ala Ile His Cys Lys Ala Gly Lys
 115 120 125
 Gly Arg Thr Gly Val Met Ile Cys Ala Tyr Leu Leu His Arg Gly Lys
 130 135 140
 Phe Leu Lys Ala Gln Glu Ala Leu Asp Phe Tyr Gly Glu Val Arg Thr
 145 150 155 160
 Arg Asp Lys Lys Gly Val Thr Ile Pro Ser Gln Arg Arg Tyr Val Tyr
 165 170 175
 Tyr Tyr Ser Tyr Leu Leu Lys Asn His Leu Asp Tyr Arg Pro Val Ala
 180 185 190
 Leu Leu Phe His Lys Met Met Phe Glu Thr Ile Pro Met Phe Ser Gly
 195 200 205
 Gly Thr Cys Asn Pro Gln Phe Val Val Cys Gln Leu Lys Val Lys Ile
 210 215 220
 Tyr Ser Ser Asn Ser Gly Pro Thr Arg Arg Glu Asp Lys Phe Met Tyr

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

[0073]

<210> 27
 <211> 597
 <212> DNA
 <213> 智人

<220>
 <221> CDS
 <222> (1)..(597)
 <400> 27

atg tca aac gtg cga gtg tct aac ggg agc cct agc ctg gag cgg atg Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met 1 5 10 15	48
gac gcc agg cag gcg gag cac ccc aag ccc tcg gcc tgc agg aac etc Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu 20 25 30	96
ttc ggc ccg gtg gac cac gaa gag tta acc cgg gac ttg gag aag cac Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His 35 40 45	144
tgc aga gac atg gaa gag gcg agc cag cgc aag tgg aat ttc gat ttt Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe 50 55 60	192
cag aat cac aaa ccc cta gag ggc aag tac gag tgg caa gag gtg gag Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu 65 70 75 80	240
aag ggc agc ttg ccc gag ttc tac tac aga ccc ccg cgg ccc ccc aaa Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys 85 90 95	288
ggt gcc tgc aag gtg ccg gcg cag gag agc cag gat gtc agc ggg agc Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser 100 105 110	336
[0074] cgc ccg gcg gcg cct tta att ggg gct ccg gct aac tet gag gac acg Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr 115 120 125	384
cat ttg gtg gac cca aag act gat ccg tcg gac agc cag acg ggg tta His Leu Val Asp Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu 130 135 140	432
gcg gag caa tgc gca gga ata agg aag cga cct gca acc gac gat tct Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser 145 150 155 160	480
tct act caa aac aaa aga gcc aac aga aca gaa gaa aat gtt tca gac Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp 165 170 175	528
ggt tcc cca aat gcc ggt tct gtg gag cag acg ccc aag aag cct ggc Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly 180 185 190	576
etc aga aga cgt caa acg taa Leu Arg Arg Arg Gln Thr 195	597
<210> 28 <211> 198 <212> PRT <213> 智人	

<400> 28

Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
 1 5 10 15

Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
 20 25 30

Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
 35 40 45

Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
 50 55 60

Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
 65 70 75 80

Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
 85 90 95

Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
 100 105 110

[0075]

Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
 115 120 125

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

Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
 145 150 155 160

Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
 165 170 175

Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
 180 185 190

Leu Arg Arg Arg Gln Thr
 195

<210> 29

<211> 894

<212> DNA

<213> 智人

[0076]

<220>

<221> CDS

<222> (1).. (894)

<400> 29

atg aca aca ccc aga aat tca gta aat ggg act ttc ccg gca gag cca 48
 Met Thr Thr Pro Arg Asn Ser Val Asn Gly Thr Phe Pro Ala Glu Pro
 1 5 10 15

atg aaa ggc cct att gct atg caa tct ggt cca aaa cca ctc ttc agg 96
 Met Lys Gly Pro Ile Ala Met Gln Ser Gly Pro Lys Pro Leu Phe Arg
 20 25 30

agg atg tct tca ctg gtg ggc ccc acg caa agc ttc ttc atg agg gaa 144
 Arg Met Ser Ser Leu Val Gly Pro Thr Gln Ser Phe Phe Met Arg Glu
 35 40 45

tct aag act ttg ggg gct gtc cag att atg aat ggg ctc ttc cac att 192
 Ser Lys Thr Leu Gly Ala Val Gln Ile Met Asn Gly Leu Phe His Ile
 50 55 60

gcc ctg ggg ggt ctt ctg atg atc cca gca ggg atc tat gca ccc atc 240
 Ala Leu Gly Gly Leu Leu Met Ile Pro Ala Gly Ile Tyr Ala Pro Ile
 65 70 75 80

tgt gtg act gtg tgg tac cct ctc tgg gga ggc att atg tat att att 288
 Cys Val Thr Val Trp Tyr Pro Leu Trp Gly Gly Ile Met Tyr Ile Ile
 85 90 95

tcc gga tca ctc ttg gca gca acg gag aaa aac tct agg aag tgt ttg 336
 Ser Gly Ser Leu Leu Ala Ala Thr Glu Lys Asn Ser Arg Lys Cys Leu
 100 105 110

gtc aaa gga aaa atg ata atg aat tca ttg agc ctc ttt gct gcc att 384
 Val Lys Gly Lys Met Ile Met Asn Ser Leu Ser Leu Phe Ala Ala Ile
 115 120 125

tct gga atg att ctt tca atc atg gac ata ctt aat att aaa att tcc 432
 Ser Gly Met Ile Leu Ser Ile Met Asp Ile Leu Asn Ile Lys Ile Ser
 130 135 140

cat ttt tta aaa atg gag agt ctg aat ttt att aga gct cac aca cca 480
 His Phe Leu Lys Met Glu Ser Leu Asn Phe Ile Arg Ala His Thr Pro
 145 150 155 160

tat att aac ata tac aac tgt gaa cca gct aat ccc tct gag aaa aac 528
 Tyr Ile Asn Ile Tyr Asn Cys Glu Pro Ala Asn Pro Ser Glu Lys Asn
 165 170 175

tcc cca tct acc caa tac tgt tac agc ata caa tct ctg ttc ttg gcc 576
 Ser Pro Ser Thr Gln Tyr Cys Tyr Ser Ile Gln Ser Leu Phe Leu Gly
 180 185 190

att ttg tca gtg atg ctg atc ttt gcc ttc ttc cag gaa ctt gta ata 624
 Ile Leu Ser Val Met Leu Ile Phe Ala Phe Phe Gln Glu Leu Val Ile
 195 200 205

gct ggc atc gtt gag aat gaa tgg aaa aga acg tgc tcc aga ccc aaa 672
 Ala Gly Ile Val Glu Asn Glu Trp Lys Arg Thr Cys Ser Arg Pro Lys
 210 215 220

tct aac ata gtt ctc ctg tca gea gaa gaa aaa aaa gaa cag act att 720
 Ser Asn Ile Val Leu Leu Ser Ala Glu Glu Lys Lys Glu Gln Thr Ile
 225 230 235 240

gaa ata aaa gaa gaa gtg gtt ggg cta act gaa aca tct tcc caa cca 768
 Glu Ile Lys Glu Glu Val Val Gly Leu Thr Glu Thr Ser Ser Gln Pro
 245 250 255

aag aat gaa gaa gac att gaa att att cca atc caa gaa gag gaa gaa 816
 Lys Asn Glu Glu Asp Ile Glu Ile Ile Pro Ile Gln Glu Glu Glu Glu
 260 265 270

gaa gaa aca gag acg aac ttt cca gaa cct ccc caa gat cag gaa tcc 864
 Glu Glu Thr Glu Thr Asn Phe Pro Glu Pro Pro Gln Asp Gln Glu Ser
 275 280 285

tca cca ata gaa aat gac agc tet cct taa 894
 Ser Pro Ile Glu Asn Asp Ser Ser Pro
 290 295

<210> 30
 <211> 297
 <212> PRT
 [0077] <213> 智人

<400> 30

Met Thr Thr Pro Arg Asn Ser Val Asn Gly Thr Phe Pro Ala Glu Pro
 1 5 10 15

Met Lys Gly Pro Ile Ala Met Gln Ser Gly Pro Lys Pro Leu Phe Arg
 20 25 30

Arg Met Ser Ser Leu Val Gly Pro Thr Gln Ser Phe Phe Met Arg Glu
 35 40 45

Ser Lys Thr Leu Gly Ala Val Gln Ile Met Asn Gly Leu Phe His Ile
 50 55 60

Ala Leu Gly Gly Leu Leu Met Ile Pro Ala Gly Ile Tyr Ala Pro Ile
 65 70 75 80

Cys Val Thr Val Trp Tyr Pro Leu Trp Gly Gly Ile Met Tyr Ile Ile
 85 90 95

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

	100	105	110
	Val Lys Gly Lys Met Ile Met Asn Ser Leu Ser Leu Phe Ala Ala Ile 115	120	125
	Ser Gly Met Ile Leu Ser Ile Met Asp Ile Leu Asn Ile Lys Ile Ser 130	135	140
	His Phe Leu Lys Met Glu Ser Leu Asn Phe Ile Arg Ala His Thr Pro 145	150	155 160
	Tyr Ile Asn Ile Tyr Asn Cys Glu Pro Ala Asn Pro Ser Glu Lys Asn 165	170	175
	Ser Pro Ser Thr Gln Tyr Cys Tyr Ser Ile Gln Ser Leu Phe Leu Gly 180	185	190
	Ile Leu Ser Val Met Leu Ile Phe Ala Phe Phe Gln Glu Leu Val Ile 195	200	205
[0078]	Ala Gly Ile Val Glu Asn Glu Trp Lys Arg Thr Cys Ser Arg Pro Lys 210	215	220
	Ser Asn Ile Val Leu Leu Ser Ala Glu Glu Lys Lys Glu Gln Thr Ile 225	230	235 240
	Glu Ile Lys Glu Glu Val Val Gly Leu Thr Glu Thr Ser Ser Gln Pro 245	250	255
	Lys Asn Glu Glu Asp Ile Glu Ile Ile Pro Ile Gln Glu Glu Glu Glu 260	265	270
	Glu Glu Thr Glu Thr Asn Phe Pro Glu Pro Pro Gln Asp Gln Glu Ser 275	280	285
	Ser Pro Ile Glu Asn Asp Ser Ser Pro 290	295	
	<210> 31		
	<211> 596		
	<212> DNA		
	<213> 智人		
	<400> 31		
	acgcgtactg gagtcaatga aagcaactat ttcaaaagat cagattactt accagtttca		60

ctaataaaga tttattactt taaaccttta tcataaaatg tafgcttga atactgtgaa	120
gtacactgca tataaggagt gtggtatagt ataaagaaac tttctgcagg tagtaattat	180
agtgaagatt ttaggtttac aaagccctag ctgttttctg tgtagctttt attattctta	240
tgaactctga caagtttgta gcttcaccat atacatttaa tttttgcaa taattggcct	300
tgttcctgag ctgttggatt cggggccgta gcaactgtctg agaggtttac atttctcaca	360
gtgaaccggt ctctttttca getgcttctt ggettctttt tactcaggtt tccactgctt	420
ttttgctttt tttaatgctg tatgaaggctg ttaacatttg tttatattt tcattaattg	480
taataccttt aatcatgca tcatactcag aaatagggat tagaatttaa gtgacatctt	540
tggcctaata taatttacct gttaaaaatt tgtgaaagct attgettage ggccgc	596
<210> 32	
<211> 511	
<212> DNA	
<213> 智人	
<400> 32	
acgcgtccat gtccgtacct ttctagtcca taccttcttt taatttttt tttcttttca	60
atttgaagag agtgettctt ctgttcttaa ggetagggaa ccaaattagg ttgtttcaat	120
[0079] atcgtgctaa aagatactgc ctttagaaga aggetattga caatccagcg tgtctcggtg	180
gaactctgac tccatggttc acttctatga tggccacatg cctctgccc agageccggc	240
agccactgtg cagtgggaag gggggccgat acactgtacg agagtgagta gcaggtctca	300
cagtgaaccg gtctctttcc ctactgtgc acactcctaa tggaatgccg ttatccaaag	360
agcagcacga acccgacagg getgagtggc ttgtgctagg gagaggtttg tgcattctt	420
gctgacaaa ctgcaggaaa aactgctaat tgtcatgctg aagactgcct gacggggaga	480
ctctgccttc tgtaagtagg tcagcgcccg c	511
<210> 33	
<211> 203	
<212> DNA	
<213> 智人	
<400> 33	
acgcgtaatt catatttgca tgctgctatg tgttctggga aatcaccata aacgtgaaat	60
gtctttggat ttgggaatct tataagttct gtatgagacc actcggatga gctgttggat	120
tggggccgt agcactgtct gagaggttta catttctcac agtgaaccgg tctctttttc	180
agetgettct tttttgcggc cgc	203

[0080]

<210>	34	
<211>	205	
<212>	DNA	
<213>	智人	
<400>	34	
	gcggccgcaa ttcataattg caigtcgcta tigtctctgg gaaatcacca taaacgtgaa	60
	atgtctttgg atttgggaat cttataagtt ctgtatgaga ccaactcggat gagctgttgg	120
	attcggggcc gtagcaactgt ctgagaggtt tacatttctc acagtgaacc ggtctctttt	180
	tcagetgett cttttttgcg gccgc	205
<210>	35	
<211>	45	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	MiR-142-3p的DNA靶序列	
<400>	35	
	gcggcccggt cgactccata aagtaggaaa cactacagcg gccgc	45
<210>	36	
<211>	128	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	DNA靶序列四次重复miR-142-3pT4X	
<400>	36	
	gcggcccggt cgactccata aagtaggaaa cactacacga ttccataaag taggaaacac	60
	tacaaccggt tccataaagt aggaaacact acatcaetcc ataaagtagg aaacactaca	120
	gcggccgc	128
<210>	37	
<211>	1131	
<212>	DNA	
<213>	单纯疱疹病毒1	
<400>	37	
	atggettctg accccggcca tcagcacgeg tctgcttgg accaggetgc gcttctctgc	60
	ggccatagea accgacgtac ggcgttgcgc cctcgcgggc agcaagaage cacggaagtc	120
	cgcccggagc agaaaaatgc cacgctactg cgggtttata tagacgtcc ccacgggatg	180
	gggaaaacca ccaccacga actgctggtg gccctgggtt cgcgcgacga tategtctac	240

gtacccgagc cgatgactta ctggcaggtg ctgggggctt ccgagacaat cgcgaacate	300
tacaccacac aacaccgcct cgaccagggt gagatatecg ccggggacgc ggcggtggta	360
atgacaagcg cccagataac aatgggcatg ccttatgccg tgaccgacgc cgtttctggct	420
ctcatatcg ggggggagge tgggagctca catgccccgc ccccggcct caccctcacc	480
ttcgaccgcc atcccategc cgcctcctcg tgttaccggg ccgceggata ccttatgggc	540
agcatgacc cccaggccgt gctggcgctc gtggcctca tcccgcgac cttgcccgcc	600
acaaacatcg tgttgggggc ccttcgggag gacagacaca tcgaccgect ggccaaacgc	660
cagcgcctcg gcgagcggct tgacctggct atgctggccg cgattcgcgc cgtttacgag	720
ctgcttgcca atacggtgcg gtatctgcag ggcggcgggt cgtggcggga ggattgggga	780
cagctttcgg ggaecggcct gccgccccag ggtgccgagc cccagagcaa cgcgggcccc	840
cgaccccata tcggggacac gttatttacc ctgtttcggg cccccagtt gctggcccc	900
aacggcgacc tgtataacgt gttgcctgg gccttggacg tcttgccaa acgcctccgt	960
cccattgcaeg tctttatcct ggattacgac caatcgccg ccggctaccg ggacgcctg	1020
ctgcaactta cctcgggat ggtccagacc cacgtcaaca cccccgctc cataccgacg	1080
[0081] atctgagacc tggegcgcac gtttgcctgg gagatggggg aggctaacta a	1131
<210> 38	
<211> 499	
<212> DNA	
<213> 智人	
<400> 38	
atgaaatata caagttatat cttggctttt cagctctgca tcgttttggg ttctcttggc	60
tgtaactgcc aggaccatat gtaaaagaag cagaaaacct taagaaatat tttaatgcag	120
gtcaatcaga tgtagcggat aatggaacte tttctttagg cattttgaag aattggaaag	180
aggagagtga cagaaaaata atgcagagcc aaattgtctc cttttaette aaacttttta	240
aaaactttaa agatgaccag agcatccaaa agagtgtgga gaccatcaag gaagacatga	300
atgtaagttt ttcaatagca acaaaaagaa acgagatgac ttcgaaaagc tgactaatta	360
ttcggttaact gacttgaatg tccaacgcaa agcaatacat gaactcacc aagtgatggc	420
tgaactgtcg ccagcagcta aaacagggaa gcgaaaaagg agtcagatgc tgtttcgagg	480
tgaagagca tcccagtaa	499
<210> 39	

<211>	468	
<212>	DNA	
<213>	小鼠 (Mus musculus)	
<400>	39	
	atgaacgeta cacactgcat ettggctttg cagctcttcc tcatggetgt ttctggetgt	60
	tactgccacg gcacagtcac tgaaagccta gaaagtctga ataactatct taactcaagt	120
	ggcatagatg tggaagaaaa gagtctcttc ttggatatct ggaggaactg gcaaaaggat	180
	ggtgacatga aaatctgca gagccagatt atctctttct acctcagact ctttgaagtc	240
	ttgaaagaca atcaggccat cagcaacaac ataagegtca ttgaatcaca cctgattact	300
	accttcttca gcaacagcaa ggcgaaaaag gatgcattca tgagtattgc caagtttgag	360
	gtcaacaace cacaggtcca gegccaagea ttcaatgagc tcatecgagt ggteccaccg	420
	ctgttgccgg aatccagcct caggaagcgg aaaaggagtc gctgctga	468
<210>	40	
<211>	462	
<212>	DNA	
<213>	智人	
<400>	40	
[0082]	atgtacagga tgcaactcct gtcttgcatt gaactaagtc ttgcacttgt cacaaacagt	60
	gcacctactt caagttctac aaagaaaaca cagctacaac tggagcattt actgctggat	120
	ttacagatga ttttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc	180
	acatttaagt tttacatgcc caagaaggcc acagaactga aacatcttca gtgtctagaa	240
	gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaa ctctcaetta	300
	agaccaggg acttaatcag caatatcaac gtaatagttc tggaactaaa gggatctgaa	360
	acaacattca tgtgtgaata tgctgatgag acagcaacca ttgtagaatt totgaacaga	420
	tggattaect tttgtcaaag catcatctca acactgactt ga	462
<210>	41	
<211>	544	
<212>	DNA	
<213>	脑心肌炎病毒	
<400>	41	
	cgttactggc cgaagccgct tggaataagg ccggtgtgcg ttgtctata tgttattttc	60
	caccatattg ccgtcttttg gcaatgtgag ggcccggaaa cctgcccctg tcttcttgac	120
	gagcattect aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt	180
	gaaggaagea gttctcttgg aagcttcttg aagacaaca acgtctgtag cgaccctttg	240

caggcagcgg aaccccccaac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata	300
agatacacct gcaaaggcgg cacaaccccc gtgccacggt gtgagttgga tagttgtgga	360
aagagtcaaa tggctctect caagcgtatt caacaagggg ctgaaggatg ccagagaaggt	420
acccattgt atgggatctg atctggggcc tcgggtgcaca tgettfacat gtgtttagtc	480
gaggttaaaa aacgtctagg cccccgaac cacggggacg tggtttctct ttgaaaaaca	540
cgat	544
<210> 42	
<211> 11892	
<212> DNA	
<213> 人工序列	
<220>	
<223> pAC3-yCD2-6A	
<400> 42	
tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccataa gggactttcc attgacgtca	180
[0083] atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc	240
aagtaegecc cctattgacg tcaatgaegg taaatggccc gctggcatt atgcccagta	300
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccaccccc ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt	540
acgggtgggag gtctatataa gcagagctgg tttagtgaac cggcgcagc cctccgattg	600
actgagtcgc ccgggtaacc gtgtatccaa taaacctct tgcagttgca tccgacttgt	660
ggtctcgctg ttccttggga gggctcctc tgagtgattg actaccgctc agcgggggtc	720
tttcatttgg gggetcgtec gggatcgga gaccctgcc caggaccac cgaccacca	780
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgac	840
tgattttatg cgcctgcgtc ggtactagtt agctaactag ctctgtatct ggcggaccg	900
tgggtggaact gacgagttcg gaacaccogg ccgcaaccct gggagacgtc ccagggactt	960
cgggggccgt ttttgtggcc cgacctgagt ccaaaaatcc cgatcgtttt ggactctttg	1020
gtgcaccccc cttagaggag ggatatgtgg ttctggtagg agacgagaac ctaaaacagt	1080

tcccgcctcc	gtctgaattt	ttgctttcgg	tttgggaccg	aagccgcgcc	gcgcgtcttg	1140
tctgctgcag	catcgttctg	tgttgctctt	gtctgactgt	gtttctgtat	ttgtctgaaa	1200
atatgggcca	gactgttacc	actcccctaa	gtttgacctt	aggteactgg	aaagatgtcg	1260
agcggatcgc	tcacaaccag	tcggtagatg	tcaagaagag	acgttggggt	accttctgct	1320
ctgcagaatg	gccaaccttt	aacgtcggat	ggcccgcgaga	cggcaccttt	aaccgagacc	1380
tcatacccca	ggttaagatc	aaggtctttt	cacctggccc	gcataggacac	ccagaccagg	1440
tcccctacat	ctgacactgg	gaagccttgg	cttttgacce	ccctcccctg	gtcaagccct	1500
ttgtacaccc	taagcctccg	ctctctcttc	ctccatccgc	cccgtctctc	ccccttgaac	1560
ctctctgttc	gaccccgcct	cgatcctccc	tttatecage	ctcaactcct	tctctaggcg	1620
ccaaacctaa	acctcaagtt	ctttctgaca	gtggggggcc	gtctatcgac	ctacttacag	1680
aagaccccc	gccttatagg	gacccaagac	cacccccttc	cgacagggac	ggaaatggtg	1740
gagaagcgac	ccctgcggga	gaggcaccgg	acccctcccc	aatggcatct	cgcttacgtg	1800
ggagacggga	gccccctgtg	gcgactcca	ctacctcgca	ggeattcccc	ctcgcgcgag	1860
gaggaaacgg	acagcttcaa	tactggccgt	tctctcttcc	tgacctttac	aactggaaaa	1920
[0084] ataataacce	tttttttctt	gaagatccag	gtaaactgac	agctctgac	gagctctgctc	1980
tcatacccca	tcageccacc	tgggacgaet	gtcagcagct	gttggggact	ctgctgaccg	2040
gagaagaaaa	acaacgggtg	ctcttagagg	ctagaaaggc	ggtgcggggc	gatgatgggc	2100
gccccactca	actgcccact	gaagtcgatg	ccgcttttcc	ctcgcagcgc	ccagactggg	2160
attacaccac	ccaggcaggt	aggaaccacc	tagtccacta	tcgccagttg	ctcctagcgg	2220
gtctccaaaa	cgcgggcaga	agccccacca	atttggccaa	ggtaaaagga	ataacacaag	2280
ggcccaatga	gtctcctctg	gccttctctg	agagacttaa	ggaagcctat	cgcaggtaca	2340
ctccttatga	ccctgaggac	ccaggccaag	aaactaatgt	gtctatgtct	ttcatttggc	2400
agtctgcccc	agacattggg	agaaagttag	agaggttaga	agatttaaaa	aacaagacgc	2460
ttggagattt	ggttagagag	gcagaaaaga	tctttaataa	acgagaaacc	ccggaagaaa	2520
gagaggaacg	tatcaggaga	gaaacagagg	aaaaagaaga	acgccgtagg	acagaggatg	2580
agcagaaaga	gaaagaaaga	gatcgtagga	gacatagaga	gatgagcaag	ctatttgcca	2640
ctgtctgtag	tggacagaaa	caggatagac	agggaggaga	acgaaggagg	tcccactctg	2700
atcgcgacca	gtgtgcctac	tgcaaagaaa	aggggcactg	ggctaaagat	tgteccaaga	2760
aaccacgagg	acctcgggga	ccaagacccc	agacctcctt	ctgacccta	gatgactagg	2820

	gaggtcaggg tcaggagccc cccctgaac ccaggataac cctcaaagtc ggggggcaac	2880
	ccgtcacctt cctggtagat actggggccc aacctccgt getgacccaa aatcctggac	2940
	ccctaagtga taagtctgcc tgggtccaag gggctactgg aggaaagcgg tategctgga	3000
	ccacggateg caaagtacat ctagnetacc gtaaggtcac ccactcttc ctccatgtac	3060
	cagactgtcc ctatctctg ttaggaagag atttgetgac taaactaaaa gcccaaatcc	3120
	actttgaggg atcaggagcc caggttatgg gaccaatggg gcagcccctg caagtgttga	3180
	ccctaaatat agaagatgag tatcggtac atgagacctc aaaagagcca gatgtttctc	3240
	taggtccac atggctgtct gattttctc aggcctgggc gaaaccggg ggcattggac	3300
	tggcagttcg ccaagctct ctgatcatac ctctgaaage aacctctacc cccgtgtcca	3360
	taaaacaata ccccatgtea caagaagcca gactggggat caagcccac atacagagac	3420
	tgttgacca ggaatactg gtacctgcc agtccccctg gaacacgcc ctgctaccg	3480
	ttaagaaacc agggactaat gattatagge ctgtccagga tetgagagaa gtcaacaagc	3540
	gggtggaaga catccacccc accgtgccc acccttaca cctcttgage gggetcccac	3600
	cgteccacca gtggtacact gtgctgatt taaaggatgc cttttctgct ctgagactcc	3660
[0085]	accccaccag teagctctc ttgcctttg agtggagaga tccagagatg ggaatctcag	3720
	gacaattgac ctggaccaga ctcccacagg gtttcaaaaa cagtcccacc ctgtttgatg	3780
	aggcaactgca cagagacctc gcagacttcc ggatccagca ccagacttg atcctgctac	3840
	agtacgtgga tgaettactg ctggccgcca cttctgagct agactgccc caaggtactc	3900
	gggccctggt acaaacctc gggaacctcg ggtatcgggc ctggccaag aaagcccaaa	3960
	tttgccagaa acaggtcaag tatctgggt atcttetaaa agagggtcag agatggctga	4020
	ctgaggccag aaaagagact gtgatgggc agcctactcc gaagacctc cgacaactaa	4080
	gggagttcct agggacggca ggtctctgct gcctctggat cctgggttt gcagaaatgg	4140
	cagccccctt gtacctctc accaaaacgg ggactctggt taattgggc ccagaccaac	4200
	aaaaggccta tcaagaaatc aagcaagctc ttctaactgc ccagccctg gggttgccag	4260
	atttgactaa gcccttgaa ctctttgtcg acgagaagca gggetacgc aaaggtgtcc	4320
	taacgcaaaa actgggacct tggcgtggc cgggtggccta cctgtccaaa aagctagacc	4380
	cagtagcage tgggtggccc ccttgcctac ggatggtagc agccattgcc gtactgacaa	4440
	aggatgcagg caagetaacc atgggacagc cactagtcct tetggcccc catgcagtag	4500
	aggcaactagt caaacaaccc cccgaccgt ggttttccaa cgcccggatg actcaactac	4560

	aggccttgct	tttggacacg	gaccgggtcc	agttcggacc	ggtggtagcc	ctgaacccgg	4620
	ctacgetget	cccactgect	gaggaagggc	tgcaacacaa	ctgccttgat	atcctggccc	4680
	aagccccagg	aacccgaccc	gacctaacgg	accagccgct	cccagacgcc	gaccacacct	4740
	ggtacacgga	tggaagcagt	ctcttacaag	agggacagcg	taaggcggga	gctgcgggtga	4800
	ccaccgagac	cgaggtaate	tgggctaaag	ccctgccagc	cgggacatcc	gctcagcggg	4860
	ctgaactgat	agcactcacc	cagggcctaa	agatggcaga	aggtagaag	ctaaatgttt	4920
	atactgatag	ccgttatgct	tttgctactg	cccataatca	tggagaaata	tacagaaggc	4980
	gtgggttgct	cacatcagaa	ggcaaagaga	tcaaaaataa	agacgagatc	ttggccctac	5040
	taaaagccct	ctttctgccc	aaaagactta	gcataatcca	ttgtccagga	catcaaaaagg	5100
	gacacagcgc	cgaggctaga	ggcaaccgga	tggtcgacca	agcggcccga	aaggcagcca	5160
	tcacagagac	tccagacacc	tctaccctcc	tcatagaaaa	ttcatcacc	tacacctcag	5220
	aacatittca	ttacacagtg	actgatataa	aggacctaac	caagttgggg	gccatttatg	5280
	ataaaacaaa	gaagtattgg	gtctaccaag	gaaaacctgt	gatgcctgac	cagtttactt	5340
	ttgaattatt	agactttctt	catcagctga	ctcacctcag	cttctcaaaa	atgaaggctc	5400
[0086]	tcctagagag	aagccacagt	ccctaactaca	tgctgaaccg	ggatcgaaca	ctcaaaaata	5460
	tcactgagac	ctgcaaaget	tgtgcacaag	tcaacgccag	caagctgccc	gttaaacagg	5520
	gaactagggt	ccgcgggcat	cgccccggca	ctcattggga	gatcgatttc	accgagataa	5580
	agcccgatt	gtatggctat	aaatatcttc	tagtttttat	agataccttt	tctggctgga	5640
	tagaagcctt	cccaaccaag	aaagaaaccg	ccaaggtcgt	aaccaagaag	ctactagagg	5700
	agatcttccc	caggttcggc	atgcctcagg	tattgggaac	tgacaatggg	cctgccttcg	5760
	tctccaaggt	gagtcagaca	gtggccgatac	tgttggggat	tgattggaaa	ttacattgtg	5820
	catacagacc	ccaaagctca	ggccaggtag	aaagaatgaa	tagaaccatc	aaggagactt	5880
	taactaaatt	aacgcttgea	actggetcta	gagactgggt	gctctactc	cccttagccc	5940
	tgtaccgagc	ccgcaacacg	cegggcccc	atggcctcac	cccatatgag	atcttatatg	6000
	gggcaccccc	gccccttgta	aacttcctg	accctgacat	gacaagagtt	actaacagcc	6060
	cctctctcca	agctcactta	caggctctct	acttagtcca	gcacgaagtc	tggagacctc	6120
	tggcggcagc	ctaccaagaa	caactggacc	gaccggtggt	acctcacct	taccgagtcg	6180
	gcgacacagt	gtgggtcctc	cgacaccaga	ctaagaacct	agaacctcgc	tggaaaggac	6240
	cttacacagt	cctgctgacc	acccccaccg	ccctcaaagt	agacggcatc	gcagcttgga	6300

tacacgccgc ccacgtgaag gctgccgacc ccgggggtgg aceatcctct agactgacat	6360
ggcgcgttca acgctctcaa aacccccca agataagatt aaccctgga agcccttaat	6420
agtcattgga gtctgttag gactaggat ggcagagagc cccatcagg tctttaatgt	6480
aacctggaga gtcaccaacc tgatgactgg gctaccgcc aatgccacct ccctctggg	6540
aactgtacaa gatgccttc caaaattata ttttgatcta tgtgatctgg tcggagagga	6600
gtgggacct tcagaccagg aaccgtatgt cgggtatggc tgcaagtacc ccgcagggag	6660
acagcggacc cggacttttg acttttaagt gtgccttggg cataccgtaa agtcgggggtg	6720
tggggacca ggagagggt actgtgtaa atgggggtgt gaaaccaccg gacaggctta	6780
ctggaagccc acatcatcgt gggacctaat ctcccttaag cgeggtaca cccctggga	6840
caaggatgc tctaaagtg cctgtgccc ctgtacgac ctctccaaag tatccaattc	6900
cttccaagg gctactcgag gggcagatg caacctcta gtctagaat tcaactgatc	6960
aggaaaaag gctaactggg acgggcccc atcgtgggga ctgagactgt accggacagg	7020
aacagatcct attaccatgt tctcctgac ccggcaggtc cttaatgtgg gaccccgagt	7080
ccccatagg cccaaccag tattaccga ccaaagactc cttctctac caatagagat	7140
[0087] tgtaccggt ccacagccac ctageccct caataaccagt taccctctt ccaactaccg	7200
tacacctca acctcccta caagtccaag tgtcccacag ccaccccag gaactggaga	7260
tagactacta gctctagtea aaggacctc tcaggcgctt aacctacca atcccgacaa	7320
gacccaagaa tgttggctgt gcttagtgc gggacctct tattacgaag gagtagcgt	7380
cgtgggcact tataccaate atccaccgc tccggccaac tgtaccgcca ctcccaaca	7440
taagcttacc ctatctgaag tgacaggaca gggcctatgc atgggggcag tacctaaaac	7500
tcaccaggcc ttatgtaaca ccaccaaag cgcggctca ggatectact acctgcagc	7560
accgcggga acaatgtggg ctgcagcac tggattgact cctgtttgt ccaccaggt	7620
gctcaatcta accacagatt attgtgtatt agttgaactc tgcccagag taatttacca	7680
ctccccgat tatatgtatg gtcagcttga acagcgtacc aatataaaa gagagccagt	7740
atcattgacc ctggccctc tactaggagg attaacatg ggaggattg cagctggaat	7800
agggaagggg accactgct taattaaaac ccagcagttt gacagcttc atgccctat	7860
ccagacagac ctcaacgaag tcgaaaagtc aattaccaac ctgaaaagt cactgacctc	7920
gttgtctgaa gtagtctac agaaccgcag aggcctagat ttgtattcc taaaggagg	7980
aggtctctgc gcagccctaa aagaagaatg ttgtttttat gcagaccaca cggggctagt	8040

	gagagacagc atggccaaat taagagaaag gcttaatcag agacaaaaac tatttgagac	8100
	aggccaagga tggttcgaag ggcigttaa tagatccccc tggttacca cettaatctc	8160
	caccatcatg ggacctetaa tagtactctt actgatctta ctctttggac ettgcattct	8220
	caatcgattg gtccaatttg ttaaagacag gatctcagtg gtccaggctc tggttttgac	8280
	tcagcaatat caccagctaa aacccataga gtacgagcca tgaacgcgtt actggccgaa	8340
	gccgcttggc ataaggccgg tgtgcgtttg tctatatgtt attttccacc atattgccgt	8400
	cttttggcaa tgtgagggcc cggaaacctg gccctgtctt ctgacgagc attcctaggg	8460
	gtctttcccc tctcgccaaa ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc	8520
	ctctggaagc ttcttgaaga caaacaacgt ctgtagegac cetttgcagg cagcggaaacc	8580
	ccccacctgg cgacaggtgc ctctgcccgc aaaagccacg tgtataagat acacctgcaa	8640
	aggcggcaca accccagtgc cagcttctga gttggatagt tgtgaaaga gtcaaatggc	8700
	tctectcaag cgtattcaac aaggggctga aggatgccca gaaggtacce cattgtatgg	8760
	gatctgatct ggggcctcgg tgcacatgct ttacatgtgt ttagtcgagg ttaaaaaacg	8820
	tctaggcccc ccgaaccacg gggacgtggt tttcctttga aaaacacgat tataaatggt	8880
[0088]	gaccggcggc atggcctcca agtgggatca aaaggccatg gatatcgctt acgaggaggc	8940
	cctgctgggc tacaaggagg gggcgtgcc tateggcggc tgtctgatca acaacaagga	9000
	cggcagtggt ctgggcaggg gccacaacat gaggttccag aagggtccg ccacctgca	9060
	cgcgagatc tccacctgg agaactgtgg caggctggag ggcaagggtg acaaggacac	9120
	cacctgtac accacctgt cccttctga catgtgtacc ggcctatca tcatgtacgg	9180
	catccctagg tgtgtgatc gcgagaacgt gaacttcaag tccaaggcg agaagtacct	9240
	gcaaaccagg ggccacgagg tgggtgtgt tgacgatgag aggtgtaaga agctgatgaa	9300
	gcagttcacc gacgagagge ctccaggactg gttcaggat atcggcgagt aagcggccgc	9360
	agataaaata aaagatttta tttagtctcc agaaaaaggg gggaatgaaa gacccacct	9420
	gtaggtttgg caagctagct taagtaacgc cttttgcaa ggcatggaaa aatacataac	9480
	tgagaataga gaagttcaga tcaaggctag gaacagatgg aacagctgaa tatgggcca	9540
	acaggatata tgtggtaagc agttcctgcc ccggctcagg gccagaaca gatggaacag	9600
	ctgaatatgg gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggcca	9660
	gaacagatgg tccccagatg cgtccagcc ctccagcagtt tetagagaac catcagatgt	9720
	ttccagggtg cccaaggac ctgaaatgac cctgtgcctt atttgaacta accaatcagt	9780

	tcgcttctcg cttctgttcg cgcgcttctg ctccccgagc tcaataaaaag agcccacaac	9840
	ccctcaactcg gggcgccagt cctccgattg actgagtcgc cegggtaccg gtgtatccaa	9900
	taaaccctct tgcagttgea tccgacttgt ggtctcgtcg ttccttggga gggctctctc	9960
	tgagtgattg actaccgctc agcgggggctc tttcattaca tgtgagcaaa aggccageaa	10020
	aaggccagga accgtaaaaa ggcgcgcttg ctggcgcttt tccatagget ccgccccct	10080
	gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgac aggactataa	10140
	agataccagg cgtttcccc tggaagctcc ctctgtcgtc ctctgttcc gacctgccc	10200
	cttaccggat acctgtccgc etttctccct tcgggaagcg tggcgcttc tcatagctca	10260
	cgctgtaggt atctcagtte ggtgtaggct gttcgttcca agctgggctg tgtgcacgaa	10320
	cccccgctc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaaccg	10380
	gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg	10440
	tatgtaggcg gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaagg	10500
	acagtatttg gtatctgcgc tetgtgaag ccagttacct tcggaaaaag agttggtagc	10560
	tcttgatccg gcaaacaaa caccgctggg agcgggtggt tttttgttg caagcagcag	10620
[0089]	attacgcgca gaaaaaaagg atctcaagaa gatccttga tctttctac ggggtctgac	10680
	gctcagtgga acgaaaactc acgtaaggg attttggctc tgagattatc aaaaaggatc	10740
	ttcacctaga tctttttaa ttaaaaatga agtttttaaat caatctaaag tatatatgag	10800
	taaacttggg ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt	10860
	ctatttcggt catccatagt tgctgactc cccgtcgtgt agataactac gatacgggag	10920
	ggcttaccat ctggccccag tgctgcaatg ataccgagc acccaagctc accggctcca	10980
	gatttatcag caataaacca gccagccgga agggccgagc gcagaagtgg tectgcaact	11040
	ttatccgct ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca	11100
	gttaatagtt tgcgcaacgt tgttccatt gctgcaggca tcgtggtgct acgctcgtcg	11160
	tttggtatgg ctccatcag ctccggttcc caacgatcaa ggcgagttac atgatcccc	11220
	atgttgtgca aaaaagcggg tagctccttc ggtctctcca tcgttgcag aagtaagttg	11280
	gccgcagtgt tateactcat ggttatggca gcaactgcata attctcttac tgtcatgcca	11340
	tccgtaagat gcttttctgt gactgggtgag tactcaacca agtcattctg agaatagtgt	11400
	atgcggcgac cgagttgctc ttgcccgcg tcaacacggg ataataaccg gccacatagc	11460
	agaactttaa aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc	11520

[0090]

ttaccgetgt tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca	11580
tcttttactt tcaccagegt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa	11640
aagggataaa gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat	11700
tgaagcattt atcagggtta ttgtctcatg ageggataca tatttgaatg tatttagaaa	11760
aataaacaaa taggggttcc ggcacattt ccccgaaaag tgccacctga cgtetaagaa	11820
accattatta tcatgacatt aacctataaa aataggcgta tcacgaggcc ctttcttett	11880
caagaattcc at	11892

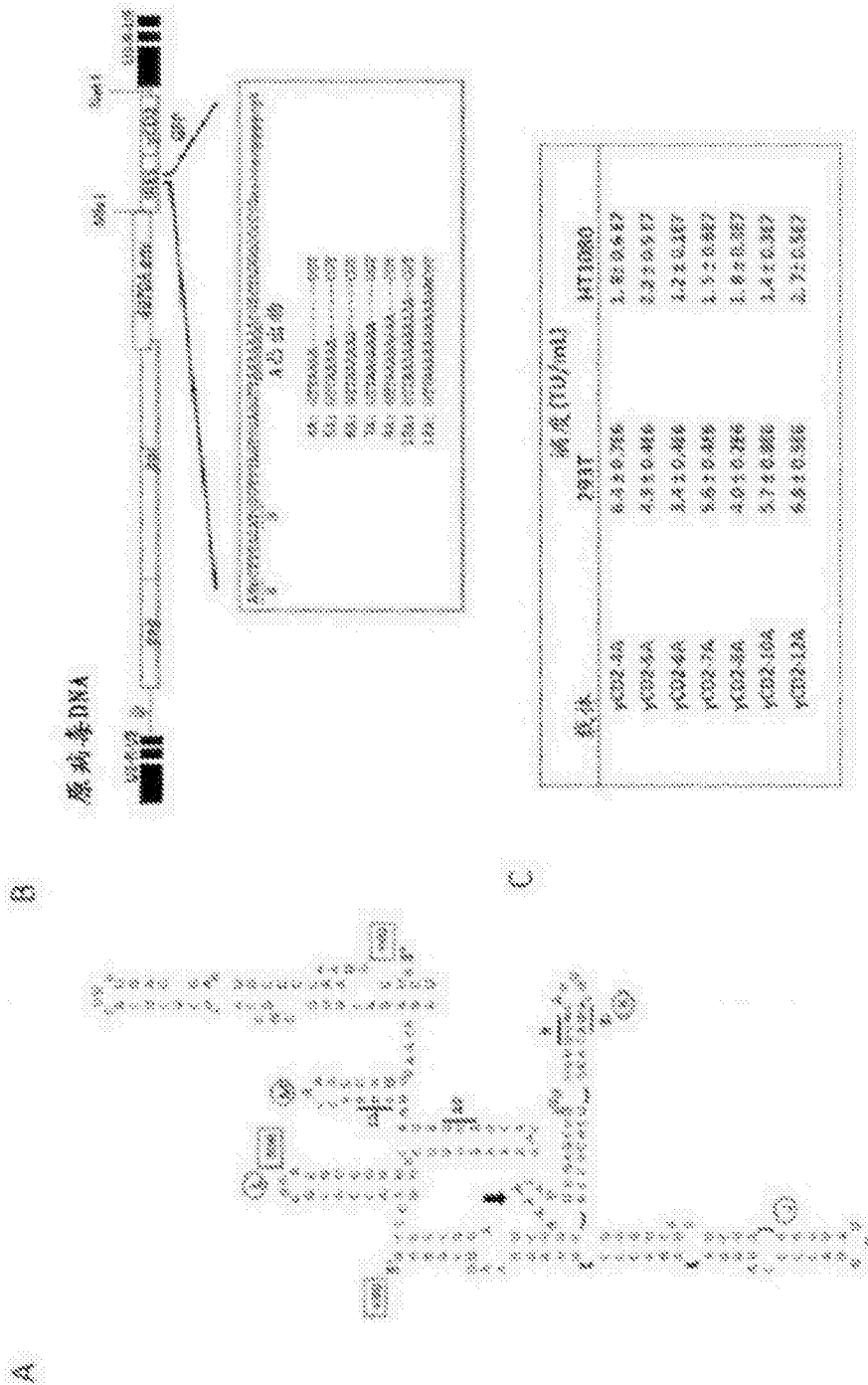


图1A-C

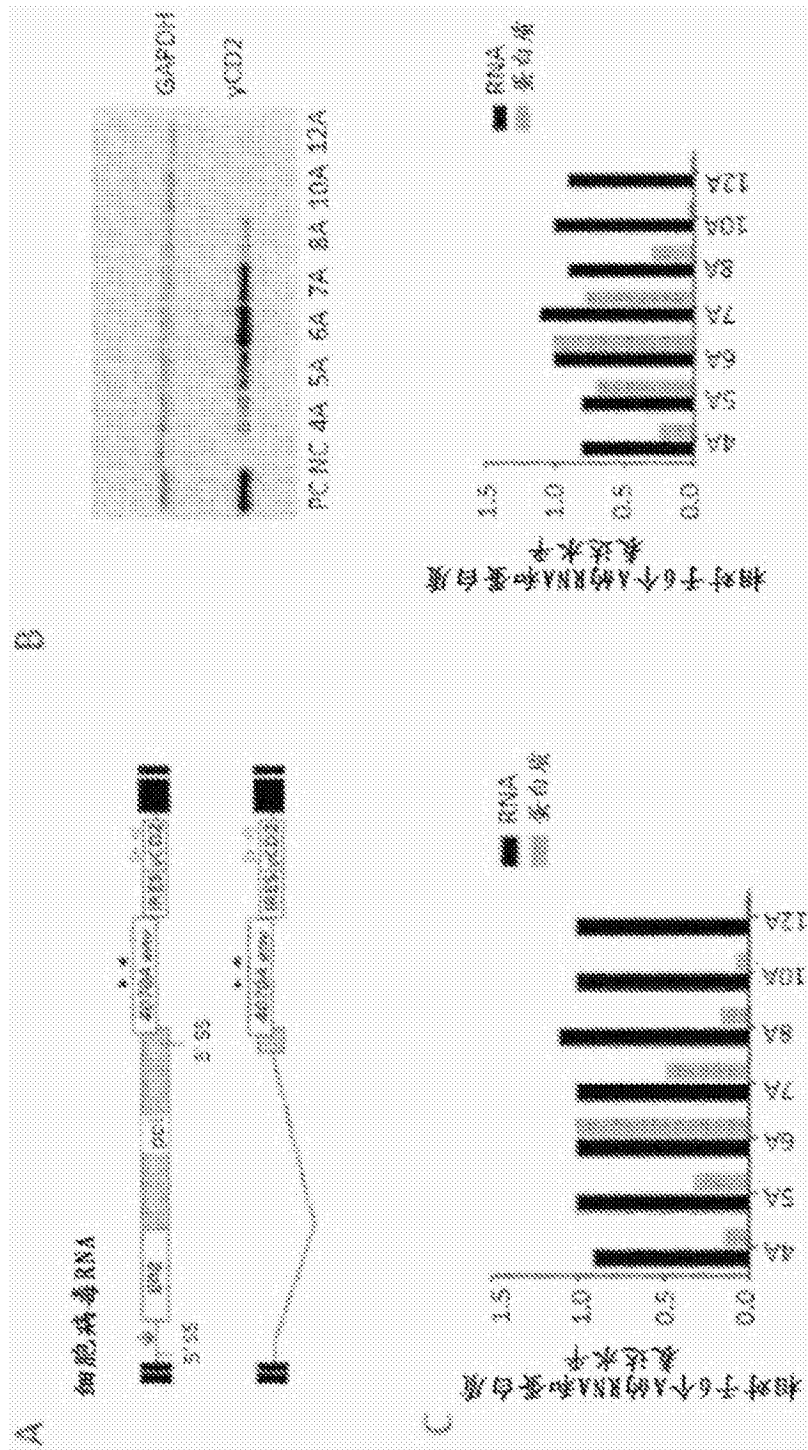


图2A-C

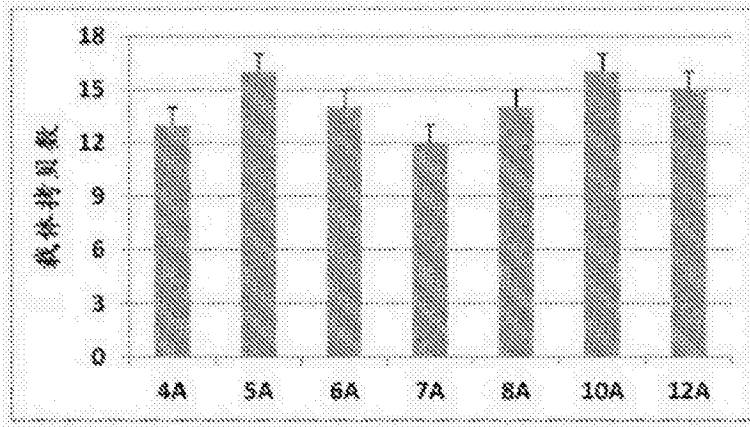


图2D

RCA 载体 - pAC3-pCD2 (SEQ ID NO:121)

tagttattaa	tagtaataa	ttacggggctc	atcagttcat	agcccatata	tggagttcag	80
cgfiasetaa	cttaccgtaa	atggcccgc	tgggtpacag	cccacagacc	cccgcceatt	120
gaagttcaata	atgacgtatg	tccccttagt	aacgccata	gggacttcc	attgacgtca	160
atgggtggag	taattacggt	aaactgcca	cttggcagta	catcaagtyt	atcatatgca	200
aagtaagccc	ccatctgacg	taaatggcgg	taaatggccc	gcctggcaat	atgcccagta	240
caagacccta	tgggacttcc	ctacttggca	ghacactctc	gtattagtca	tccctattac	280
caaggttgatg	gggtcttggc	agtaaccaca	tggggctgga	tggggcttgg	actacagggg	320
atttcaaaat	ctcccccaca	tggacgtcaa	tgggagtttg	ctttggcaac	aaactcaaat	360
agactttcaa	aaatgtctga	acaaactccg	cccattgacg	caaatgggag	gtaggcgtgt	400
ccggtgggag	gtccatataa	gcagagctgg	cttagtgaac	ccgagcaagt	actccagatt	440
actgagctgc	ccgggtaccc	gttatctcaa	taaacctctt	tggagttgca	tccgacttgt	480
ggctctgctg	tcccttggga	gggtctctct	tgggtgattg	actaccctgc	agcggggctc	520
ctttacttgg	gggtctgtcc	gggtctggga	gacccctgpc	ccgggacccc	ggccccca	560
ccggtaggag	agctggccag	caacttatct	gtgtctgtcc	gattgtctag	tgtctatgac	600
tgaatttatg	ggactgctgc	ggtactagtt	agctaacctg	ctctgtatct	ggggacccc	640
tgggtgaaat	gaagcttctg	gaacacccgg	ccgcaacctt	gggagagctc	ccagggactt	680
ccggggccct	ctttggggca	ggactgagtt	ccaaaatccc	cgatcgtttt	ggactctttg	720
gtgacccccc	cttagagggg	ggtatgttgg	tctgtgttag	agacagagac	ctaaacagtt	760
tccggactcc	cttagcaatt	tggcttggg	cttgggcaag	aaccccggcc	gggctcttgg	800
tctgctgag	catgtttctg	tgtttgtctct	gtctgactct	gtttctgtat	ctgtctgaga	840
atactggccc	gactgttacc	actcccttaa	gtttgacctt	agttaccctg	aaactgtctg	880
agggatctgc	tcccacccag	tgggtagatg	tcaagagag	agttggggtt	actctctctg	920
ctgacagatg	gaacaccttt	agcctcaggt	ggcctcagga	ccgcaacctt	aacccagacc	960
tcaatcccac	ggttagatc	aaagtctttt	caactggccc	gcttagaac	cccagccagg	1000
tcccctaaat	ggtgactctg	gaagccttgg	cttttgacc	ccctccctgg	gtcaagccct	1040
tgttccacc	taagctctcc	ccctctctct	ctccctccc	ccgctctctc	ccctctgacc	1080
ctctctgctc	gaacccagct	ggtactccc	ctctctcagc	ctcactccct	ctcttagggg	1120
ccaaacctaa	actccagttt	ctttctgaca	gtggggggcc	gctcctagac	ctacttaccg	1160
agaccccc	gctctctagg	gaacccagac	caaccccttc	ccagagggac	ggactctgtg	1200
gagagcggcc	ccctcgggga	ggggccccc	ccccctccc	actggcctct	ggctctagct	1240
gagagcggga	cccccctgta	ggagactcca	ctcctctcca	ggctctccc	ctctcggcag	1280
ggggaacagg	acagcttcaa	tactggcgtt	tctcctcttc	tgaacttacc	aaactgaaaa	1320
ctataaac	ctctttttct	gaagatccag	gtaactgac	agctctgata	gagctctctc	1360
tcaatcccac	taagccccc	tgggactgct	gtcagagact	gttggggact	ctgctgacag	1400
gagagaaaa	ccaaaggggt	ctctctaggg	ctagaagacc	gttgggggac	gctctctgac	1440
gccccctc	ctggcccaat	gaagctgata	ccgttttccc	ctctagagcc	ccagctctgg	1480
gtctccaaa	ggggggcaga	agcccccca	atttggccaa	gttaaaagga	atacccaag	1520
ggcccaatga	gtctccctcc	gcttctctag	agagacttaa	ggagcctat	ccgagctaca	1560
ctctctctga	ccctcgggac	ccagggccag	aaactaatgt	gtctctgtct	ctctcttggc	1600
agctctgccc	agacatgggg	agaaatllag	agaggtlaga	agatllaaa	aacagagacc	1640
ttggagattt	ggttagagag	gcagaaaaga	tctttataaa	acagaaaacc	ccggaagaaa	1680
gagaggaag	talccgggga	gaacccaggg	aaagagagga	acgggtctgg	ccagagagtg	1720
agcagaaaga	gaaagagaga	gctctctagga	gacatagaga	gatgagagag	ctctctggcc	1760
ctctctctcc	tggagagaaa	ccgctctccc	agggagagga	ccagagagag	ctccactccg	1800
cccccagagg	actctgggga	ccagagcccc	agacctccct	ctctgacctc	gctgactagg	1840
gaggtccagg	tccagagccc	ccccctgacc	ccaggtacc	cttccaaagt	gggggggccc	1880
ccctcaactt	ctctgtgact	actggggccc	aaactctctt	gctgacccc	aatctctgac	1920
ccctcaagtga	taagctctgc	tgggtccaaag	gggtactctg	aggaagaggg	tatctctgga	1960
ccagagctct	caaagtaaat	ctagctccag	gtcaggtacc	ccactcttct	ctccactctac	2000
ccagctctcc	ctctctctct	ttaggaagag	atttctctgac	taacttaaa	gcccaactcc	2040
acttctgggg	ataagagacc	caagttctag	gaccaatggg	ggagcccttg	caagctctga	2080
ccctcaatct	agagctctgc	ctctctctcc	atgagctctc	aaagagagcc	gctgttctct	2120
tgggctctcc	ccagctctct	ctgactctac	ctctgagacc	aaactctaac	ccctgtctcc	2160
taaaacata	ccccctctca	caagagccca	gactgtggat	ccagccctcc	ataccagagc	2200

图3

tgttggacaa	gggaataactg	gtaacctggcc	agtcacccctg	gaaacacgccc	ctgctaacccc	3180
ttaagaaacc	agggactaac	gattataggg	ctgtccagga	tctgagagaa	gttaacaaagc	3240
gggttggaga	cattcaacccc	accgttgccc	acccttaca	cattcttggc	gggttccccc	3300
cgtccacccc	gtggtlaccc	gtgcttgcct	taaggatgc	ctttctctgc	ctgagactcc	3360
acccacccc	ctagcctctc	ttagcctttg	gttggagaa	tccagagatg	ggatctctag	3420
gacactttag	ctgggacaga	ctcccaatag	gtttcaaaaa	cagttccccc	ctggttctag	3480
aggaacttga	ggagacccca	ggagacttcc	ggatccagcc	cccagccctg	atctctgtac	3540
agtcacttga	tgacttacct	ctggcccgcc	cttctgagct	agactgccc	caaggtactc	3600
gggcccgtgt	acaaccccca	gggaacctag	ggtatcgggc	ctggcccaag	aaagccaaa	3660
tttccacaga	acaggtcaag	tatcttgggt	ctctctcaaa	agagggtag	agatgggtga	3720
ctgaggtccg	aaaagagact	ctgalagggc	agctactcc	gagacccct	cgacactta	3780
gggacttcc	agggagggca	ggctctctgc	gctcttggct	ccttgggttt	gcagaaatgg	3840
cagccccc	gtacctctcc	accacaaggg	ggactctgt	taattgggc	ccagaccac	3900
aaaagggcta	tcaagaaatc	agccaagctc	tctaaactgc	cccagccctg	gggttggccg	3960
attttagctaa	ctcttttctg	ctcttttctg	agagagacca	ggcttaccgc	aaaggtctcc	4020
tcaagcctaaa	acttgggacc	tggggtcggc	cgttggctca	ctcttccaaa	aaagtagaac	4080
cagtagctagc	tgggttggcc	cttctgctcc	ggatgttagc	ggatctgccc	gtactgccc	4140
aggttgcagg	caagttaac	atgggacagc	cactagctat	tctggccc	ctgcaagtag	4200
aggaacttag	caaacacccc	cccagccct	ggcttccaa	ggcccgatg	actcaactac	4260
aggaacttgc	tctggacag	gaaaggtcc	agtttggac	gggttaggc	ctgaaacccg	4320
ctagacttgc	cccacttccc	ggggagggc	tgcacacaaa	ctgctttag	actctggccg	4380
agcccccag	aaacttgcac	gccttctag	accagctgct	cccagccctc	gcccacact	4440
ccctccagga	tggagcagc	ctcttctag	agggacggc	taagggggc	gttgggttag	4500
ctgacacttag	agaggttaac	tgggtctaac	cccttccagc	gggacactcc	gtccagggc	4560
atctttagct	ccgttctgct	tcttctctct	agcttggcga	aggttggagc	ctaatgttt	4620
gtaggtttag	ccccttccga	ggcaagcga	tcaaaataa	agacagatc	ttggccctac	4680
taaaagccct	cttcttccc	aaaagctta	gataactta	tcttccagga	catcaaaac	4740
gacccagccc	cgaggttaga	ggcaaccgga	tggcttagca	agcggccga	aaagcagcca	4800
tcacagagac	tccagacccc	tctacccctc	tctatgaaaa	lctatccccc	tacccctcag	4860
aaacttttca	ttacacagtg	actgatata	aggaactaac	caagttgggg	gcaatttag	4920
ataaaacaaa	gaagttattg	gtctaacca	gaaaacctgt	gatgcccgc	cagttactt	4980
ttagatttat	agactttct	cattcagctga	ctcaactcag	cttctcaaaa	atgaaagctc	5040
tccttagagag	agccacagc	ccctactaca	tgttagactc	ggcttagaca	ctcaaaaata	5100
tcacttagac	ctgcaagct	tgttcccaag	tcacagccc	caagttggc	gttaaacccc	5160
gaactagggt	ctggtggccc	gggcccggc	ccacttggc	gacccctcc	accagctta	5220
agcccttagt	cttctgctat	aaaatacttc	tagtttttat	agataccctt	tctggttga	5280
ttagagactt	cccacacag	aaagaaaccc	ccaggtctgt	aaacagagc	ctactagagc	5340
agacttccc	caagttggc	atgcccagc	tatttggaac	tgcacatggc	ctgcccctc	5400
tctccaggt	gagtcagca	ctggccccc	lytttgggt	tgatttggaa	ttcaattgtg	5460
catctagccc	ccaaagctca	ggccagctag	aaagaaatga	tggaaacttc	aaagagactt	5520
taactaaat	aaccttggc	actggctcta	gagactgggt	gctctactc	cccttagccc	5580
tgtacagagc	cngcaacag	ccggcccac	atggccctac	ccctatgag	actttatatg	5640
gggcccacc	gccccttcta	cccttcccct	accctgacat	gcaagagct	actaacccc	5700
ccctcttcca	agcttaacta	caagctctct	acttagtcca	gcaagagctc	tggagacctc	5760
tgggggccc	ctaacagaa	caacttggcc	gacccgtggt	ccctccccc	taacagctc	5820
gggacacagt	gtgggtccgc	cgaacacaga	ctaaagacct	agaaacttgc	tggaaagggc	5880
cttaacacag	cccttggac	cccacacccc	ccctcaagct	agacggcctc	gcaagttgca	5940
tacacagccc	ccacttgaag	gcttcccacc	ccgggggtgg	accctactct	agactgacat	6000
gggcccctca	agctcttcca	aaacccctcc	agataagctt	aaacccctgc	agcccttat	6060
agctctggga	gtcttcttag	caagtagaat	ggcagagagc	cccactcagc	tctttaaagt	6120
aaacttggga	gtctacacac	tgalactctg	ggctacccc	acttcccact	ccctcttggc	6180
aaactttaga	gctgcttccc	caaaatctata	tcttgaatcta	tgtgactctg	tccagagagc	6240
gtgggccc	tcagacacag	aaacgttatg	gggttatggc	tgcagtlacc	cccaggggag	6300
aaagccgacc	ggacttttgc	acttttccgt	gtgcccctgg	caaaccttgc	agtaggggtg	6360
tgggggccc	ggagagggct	actgtgctaa	atgggggtgt	gaaacacccc	gacaggtcta	6420
ctgagagccc	aaactctctg	gggacttaaf	ctcccctaac	ggggttacc	cccctgggc	6480
caagggatgc	tctaaagttg	cccttggccc	ctgctccgac	ctctccaaag	tatccaatct	6540

图3(续)

cttcccaagg	gctactcggg	ggggccagatg	caaccctcta	gctctagcat	tcactgatgc	8880
aggcaaaaag	gctactcggg	aggggcccac	atcgtgggga	ctgagactgt	accggacagg	7920
aacagatcc	attaccatgr	tcfccctgac	agggccggtc	cttcaatgtg	gaccccgagt	7080
ccccataggg	cccaccocag	tattccocga	ccccagatc	ccctccctac	caatagagat	7140
tgtaaccggt	cccacagccc	ctagcccct	ccctaccgt	cccaccctt	ccactagag	7200
taaccctca	accctcccta	caagctcaag	tgctcccag	cccccccag	gacctggaga	7260
tcagctacta	gctctagctc	agggagccta	tcaggacct	cccccccca	atccccaaa	7320
gaccacaaga	tgcttagctg	gcttagctgc	gggaactcct	tattagggag	gagtagcgg	7380
ctagggacat	tataccaatc	attccaccgc	tccggacac	tgtaaggaca	cttcccaaa	7440
taagcttacc	ctatctgaa	tgacaggaac	ggccctctgc	ctgggggag	taactaaa	7500
taaccgggac	ttatgtaaac	cccccccag	ccggggtaa	ggatccctct	cccttgagc	7560
accctggggc	acaaatgtgg	cttgacagac	tgatctgact	ccctgctctg	cccccaagg	7620
gctcaatcta	accacagatc	attgtgtatt	gcttgaactc	tgccccag	taatttacta	7680
ctccccgat	tatactgata	gttagcttga	ccccgctacc	aaatataaa	gagagccag	7740
atcattgacc	ctgggacctc	taactaggg	cttaaccatg	ggagggatg	ccactggat	7800
agggccgggg	accactccct	taattaaa	ccccagctt	gagcagcttc	atggccgat	7860
ccccagccgc	ctccaccag	tgcaaacctc	aatccacc	ctgaaacct	ccctgacct	7920
gctgtctgaa	gtagctccac	agaaaccag	agccctagat	ctgctatctc	taagggagg	7980
aggtctctgc	gcagccctaa	agagagatg	tgcttcttat	ccccccccc	ggggctcgt	8040
gagagaaag	atggcccaat	taagagaaa	gcttaactag	agccaaaa	tatttgagc	8100
aggcccaagg	tggtctgag	ggctgtttaa	tagatcccc	tggttctaca	cccccaatc	8160
ccccatcag	ggacctctaa	tgctactctt	actgatctta	ctctttggac	cttgcaattc	8220
ccctcagctt	ctcccaattg	ctaaagccag	gctctccgtg	ctccaggttc	tggttttag	8280
taagccatct	ccccagctaa	cccccttaca	ctaacagcca	tgacagctt	actggccgaa	8340
gccccttggc	ctaacagccg	tgctgctttg	ctctctcttt	cttttccac	ctcttgccgt	8400
cttttggcca	tgtaggggac	gggaacctg	gcccctcttt	cttgagccc	atctctagg	8460
ctctttccc	tctcccaaaa	ggatctgca	gtctgttga	tgctgtgag	gagccagctc	8520
ctctgggagc	ttcttgaaag	cccccaacct	cttgagccac	cttttgcag	ccccggaaa	8580
ccccccctgg	ccacaggtgc	ctctcgggac	aaaagcccag	tgctataag	acacctgaa	8640
aggggggaca	cccccaagtc	cccccttgta	gctgggact	tggtggaaa	gctcaalgg	8700
tctctccag	cgtaattcacc	aggggctga	ggatctccc	gaggttccc	catcttatg	8760
gctctgatc	ggggccctgg	tgccctctct	ctanaatgt	ctagctggg	tt cccccccc	8820
gcttagggcc	cccccaacc	ggggccctgg	ctttctcttg	ccccccccc	ctctccctg	8880
tgacccggcg	ccctggctcc	agctgggac	aaaagggat	gctatacct	taagagagg	8940
ccctgctggg	ctaacaggg	ggccggctgc	ctctagggg	ctgctgtctc	aaccaacag	9000
agggccagct	gctgggccc	ggcccccca	tgaggtccc	gaggggccc	gccccctgc	9060
ccccctctga	ccccccctg	ccccctctg	acatgtctc	ccccagctc	atctatgac	9120
gcatccctag	gtgtctgata	ggccagagag	tgaccttcaa	gccccaggg	gagagctac	9180
tgcaaccag	gggcaccgg	gctgtgctg	tgagatgca	gaggtgtag	agcctgagc	9240
agcagctcc	ccacagagag	cttccagct	gcttccagga	ctctggccg	taagggccg	9300
ccagataaat	aaaagcttct	ctttagctc	gagaaaagg	ggggaatgaa	agcccccc	9360
tgtaggtttg	gcaagcttag	ctaaagctc	ccattttgca	agccatgga	aaatctata	9420
ctgagctag	agagctccg	ctccagctca	ccccagagc	gacagctga	atctgggca	9480
aacagctat	ctgtgttag	agcttctgc	ccccgctcc	ggccagagc	agatgggaa	9540
gctgataatg	ggcccaacc	gatactctg	gtaagcagct	ctgccccg	ctcagggca	9600
agacagagag	gtccccagat	gggttccag	ctccagagct	ctctagagaa	ccatcagctg	9660
tttccaggg	gccccaagg	cttgaactg	ccccctgct	tatttgact	aaccaatcc	9720
ttcctctct	gcttctcttc	ggccgctct	gctccccag	ctcaataaa	gagccccca	9780
ccccccctc	ggggccccc	ctctccctc	gactgctgc	ccccgctcc	ctgtatcca	9840
ataaacctc	ctgagcttag	ctccagctg	tgctctcct	gctcttggg	aggtctctc	9900
ctgagctat	gactccccc	ccccgggct	cttccctac	ctgtgggca	agcccccca	10020
aaagccagag	aaccttaaaa	agggccgctt	cttccctctt	ctccctagc	cccccccc	10080
tgagagacat	cccccaaatc	gagcttcaag	ctcagagctg	cccccccc	ccccctata	10140
agatccagag	gggttccc	ctggaactc	ctctgctgc	tctctgttc	ccccctgac	10200
gcttaaccga	tcctcttag	ctttctccc	ctggggagc	gctggcttt	ctcaatgctc	10260
acgtcttag	tatctagctc	gggtctagg	cttctccct	agctggctt	gctgcccc	10320
ccccccgctt	ccccccgac	gctgacctt	ctccgctaa	tctctcttg	agcccccc	10380

图3(续)

```

ggtlaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcag 10140
gtatgttagc ggtgtccacg agttcttcaa ggggtgacct aactacgggt acactagag 10150
gacagatatt ggtatctgag ctctgtctaa gccagttacc ttgggaaaa gagttggtag 10160
ctcttgatcc ggcacaacaaa ccacccctgg tagcgggtgt ttttttgtt gcagcagaa 10170
gatctcgggc agaaaaaaag gactcagaag agtctccttg atttttcta cggcgtctg 10180
ccctcagttg aacgcaaacr ccctctcagg gattttgtrc ctgcgattat ccaaaaggt 10190
cttccctcag atccitttaa attcaaatg agqtttttaa tcaectctaa gtatatatg 10200
gtaaaccttg tctgcaagtt aacaaatgctt aalccgttgg gccctcatt ccgcgattg 10210
tctatttctg tcaatcctag ttgactgact ccccgctcgt tagataact cgcatacgg 10220
gggcttaca tctgcaaacr ctgctgcaat gataccgaga gcccaacct ccccgctcc 10230
agattatca gcaatcaaac agccagccgg aagggccggg ccgagaaagt gtcctgaaac 10240
tttatcagcc tcaatcaggt ctattacttg ttgcccggag cctagagtaa gtagttcgc 10250
agttcaatag ttgcccgaag ttgttgcact tctctgagcc atcgtggtgt caccctcgc 10260
gitttgatag gtttatttaa gctccgcttc ccaacgatac agccagatta catgatccc 10270
cctgttctgc aaaaaagcgg ttactcctt ctgctctcng ctgctctca gacttaagt 10280
ggcgcagag ttatcactca ttgttatgpc agcactgact aattctctta ctgctctgc 10290
atcctgaaga tgcctttctg ttactgctga gtaactcaac aagtcattct ggcactagc 10300
tatcggcaga ccagattgct ctgcccggcc gtaaacacgg gataatacng ccgcactag 10310
cagcacttca aaagctctca tcaatgcaaa ccgtctctcg ggcgcaaac tctcaggt 10320
cttccgctg ttgagctcaa gttcagctga accactggt gccccaact gattctcag 10330
ctcttttact ttcccacggg ttctgtgctg agcaaaaca gcaagggcac atgcccacc 10340
aaagggcata agggcagcac cgaactgctg aatactata ctctctctt tcaatata 10350
ttgagcatt taincgggtt attgtctcct gacggctac ctatttgaat gttttaga 10360
aaatacaaa atagcggctc ccgcacatt tcccgcaaa gtcaccactg acgtctaga 10370
aacattatt atnatgcaat taactctaa aataagcgt ctccagacc ctttctct 10380
caagattc at

```

图3(续)

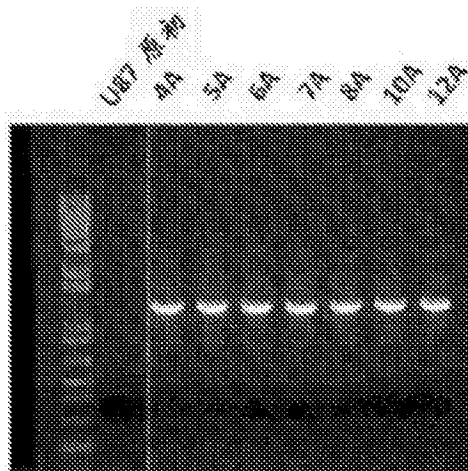


图4A

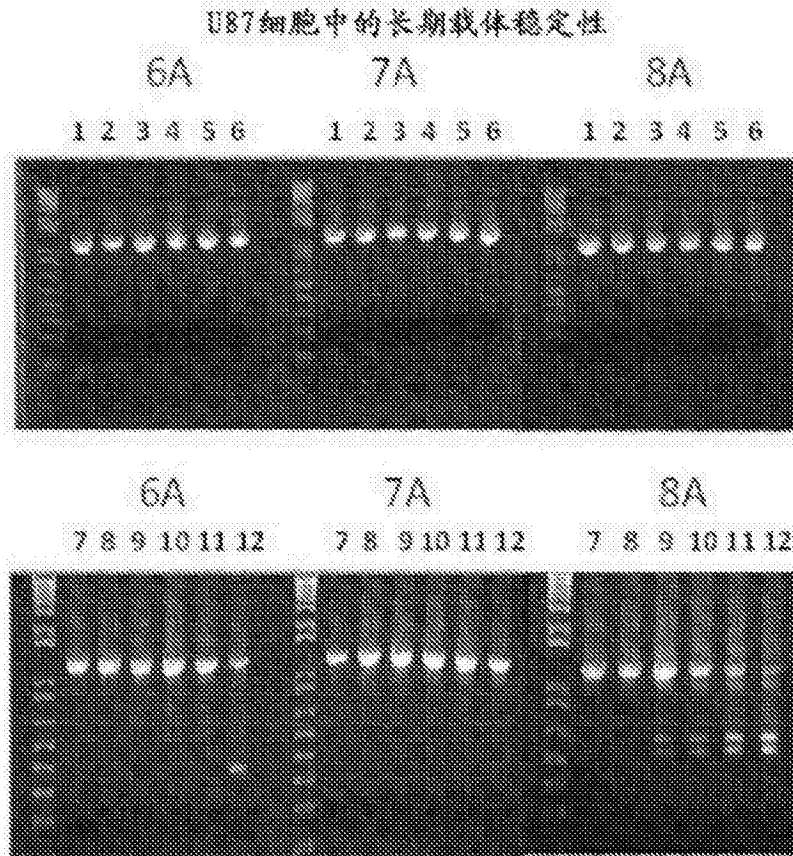


图4B

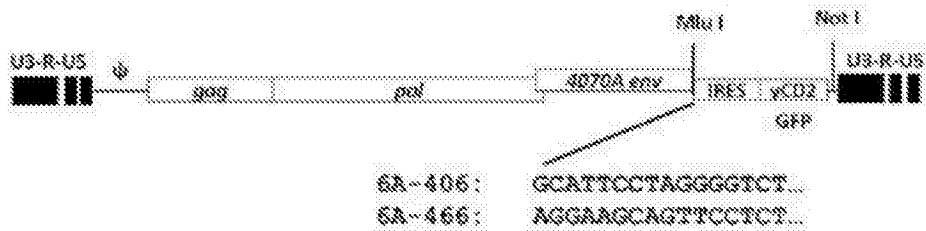


图5