



(12) 发明专利申请

(10) 申请公布号 CN 114072515 A

(43) 申请公布日 2022. 02. 18

(21) 申请号 202080049015.0

H·V·杜 R·戈特沙尔

(22) 申请日 2020.04.29

S·图斯克

(30) 优先权数据

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

62/840911 2019.04.30 US

代理人 罗文锋 彭昶

62/913401 2019.10.10 US

(85) PCT国际申请进入国家阶段日

(51) Int.Cl.

2021.12.30

C12N 15/86 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2020/030484 2020.04.29

(87) PCT国际申请的公布数据

W02020/223356 EN 2020.11.05

(71) 申请人 宾夕法尼亚州大学信托人

地址 美国宾夕法尼亚州

申请人 阿米库斯治疗学公司

权利要求书5页 说明书48页

(72) 发明人 J·M·威尔逊 J·奥尔多瓦

序列表74页 附图43页

(54) 发明名称

可用于治疗庞贝病的组合物

(57) 摘要

提供了一种可用于治疗II型糖原贮积病(庞贝)病的重组腺相关病毒(rAAV)。所述rAAV包括AAV衣壳,所述AAV衣壳靶向肌肉、心脏、肾脏和中枢神经系统中的至少一种的细胞并且所述AAV衣壳中包装有载体基因组,所述载体基因组包括在调控序列的控制下编码嵌合融合蛋白的核酸序列,所述嵌合融合蛋白包括信号肽和与人酸性α-葡萄糖苷酶hGAA780I蛋白融合的vIGF2肽,所述调控序列指导所述嵌合融合蛋白的表达。还提供了制造和使用此rAAV的方法。

1. 一种表达盒,其包括在调控序列的控制下编码嵌合融合蛋白的核酸序列,所述嵌合融合蛋白包括信号肽和与人酸性 α -葡萄糖苷酶(hGAA)融合的vIGF2肽,所述hGAA至少包括hGAA780I的活性位点,所述调控序列指导所述嵌合融合蛋白的表达,其中位置780基于SEQ ID NO:3中的氨基酸序列的位置的编号。

2. 根据权利要求1所述的表达盒,其中所述hGAA至少包括SEQ ID NO:3(hGAA780I)的氨基酸204到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

3. 根据权利要求1所述的表达盒,其中所述hGAA至少包括SEQ ID NO:3的氨基酸204到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

4. 根据权利要求1所述的表达盒,其中所述hGAA至少包括SEQ ID NO:3的氨基酸123到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

5. 根据权利要求1所述的表达盒,其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

6. 根据权利要求1所述的表达盒,其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

7. 根据权利要求1到6中任一项所述的表达盒,其中所述hGAA780I由SEQ ID NO:4或与SEQ ID NO:4至少95%相同的序列编码。

8. 根据权利要求1到6中任一项所述的表达盒,其中所述hGAA780I由SEQ ID NO:5或与SEQ ID NO:5至少95%相同的序列编码。

9. 根据任何权利要求1所述的表达盒,其中所述融合蛋白包括SEQ ID NO:6或与SEQ ID NO:6至少95%相同的序列。

10. 根据权利要求9所述的表达盒,其中所述融合蛋白由SEQ ID NO:7或与SEQ ID NO:7至少95%相同的序列编码。

11. 根据权利要求1到10中任一项所述的表达盒,其进一步包括miR靶序列的至少两个串联重复序列,其中所述至少两个串联重复序列包括至少一个第一miRNA靶序列和至少一个第二miRNA靶序列,所述第一miRNA靶序列和所述第二miRNA靶序列可能相同或不同并且在3'处与编码所述融合蛋白的所述序列可操作地连接。

12. 根据权利要求11所述的表达盒,其中所述miR靶序列独立地选自SEQ ID NO:26和SEQ ID NO:27。

13. 根据权利要求11或12所述的表达盒,其中所述miRNA靶序列中的两个或更多个miRNA靶序列由间隔子隔开,并且所述间隔子中的一个或多个间隔子独立地选自(i) GGAT; (ii) CACGTG; 以及 (iii) GCATGC。

14. 根据权利要求1到13中任一项所述的表达盒,其中所述vIGF2肽包括与SEQ ID NO:32至少90%相同并且在选自SEQ ID NO:32的位置6、26、27、43、48、49、50、54、55和65的一个或多个位置处具有至少一个取代的氨基酸序列。

15. 根据权利要求14所述的表达盒,其中所述至少一个取代选自SEQ ID NO:32的E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R。

16. 根据权利要求1或15所述的表达盒,其中所述vIGF2肽在选自SEQ ID NO:32的位置26、27、43、48、49、50、54和55的两个或更多个位置处包括至少两个取代。

17. 根据权利要求16所述的表达盒,其中所述至少两个取代选自SEQ ID NO:32的E6R、

F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R。

18. 根据权利要求1到17中任一项所述的表达盒,其中所述vIGF2肽在SEQ ID NO:32的位置1处包括N末端缺失。

19. 根据权利要求18所述的表达盒,其中所述vIGF2肽在SEQ ID NO:32的位置1到4处包括N末端缺失。

20. 根据权利要求1到19中任一项所述的表达盒,其中与天然IGF2肽相比,所述vIGF2肽对胰岛素受体和IGFR1的亲和力降低或没有亲和力。

21. 根据权利要求1到20中任一项所述的表达盒,其中所述vIGF2肽能够促进将hGAA780I摄取到细胞中的溶酶体中。

22. 根据权利要求1到21中任一项所述的表达盒,其中所述核酸序列进一步包括位于vIGF2核苷酸序列与编码hGAA780I的核酸序列之间的编码接头肽的接头序列。

23. 根据权利要求22中任一项所述的表达盒,其中所述接头肽包括SEQ ID NO:55-60中的任一个。

24. 根据权利要求1到23中任一项所述的表达盒,其中所述信号肽是结合免疫球蛋白(BiP)信号肽或Gaussia信号肽。

25. 根据权利要求24所述的表达盒,其中所述BiP信号肽包括与SEQ ID NO:49-53中的任一个至少90%相同的氨基酸序列。

26. 根据权利要求25所述的表达盒,其中所述BiP信号肽包括SEQ ID NO:49-53中的任一个的氨基酸序列。

27. 根据权利要求24所述的表达盒,其中所述信号肽包括Gaussia信号肽。

28. 根据权利要求27所述的表达盒,其中所述Gaussia信号肽包括与SEQ ID NO:54至少90%相同的氨基酸序列。

29. 根据权利要求28所述的表达盒,其中所述Gaussia信号肽包括SEQ ID NO:54。

30. 根据权利要求1到29中任一项所述的表达盒,其中所述表达盒由选自以下的病毒载体携带:重组细小病毒、重组慢病毒、重组逆转录病毒和重组腺病毒。

31. 根据权利要求30所述的表达盒,其中所述重组细小病毒是进化枝F腺相关病毒。

32. 根据权利要求31所述的表达盒,其中所述进化枝F腺相关病毒是AAVhu68。

33. 根据权利要求1到29中任一项所述的表达盒,其中所述表达盒由选自以下的非病毒载体携带:裸DNA、裸RNA、无机颗粒、脂质颗粒、基于聚合物的载体或基于壳聚糖的调配物。

34. 一种重组腺相关病毒(rAAV),其包括:

(a) AAV衣壳,所述AAV衣壳靶向肌肉、心脏和中枢神经系统中的至少一种的细胞;以及

(b) 载体基因组,所述载体基因组包装在所述AAV衣壳中,所述载体基因组包括在调控序列的控制下编码嵌合融合蛋白的核酸序列,所述嵌合融合蛋白包括信号肽和与hGAA融合的vIGF2肽,所述hGAA至少包括hGAA780I的活性位点,所述调控序列指导所述嵌合融合蛋白的表达,其中位置780基于SEQ ID NO:3中的氨基酸序列的位置的编号。

35. 根据权利要求34所述的rAAV,其中所述hGAA至少包括SEQ ID NO:3(hGAA780I)的氨基酸204到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。

36. 根据权利要求34所述的rAAV,其中所述hGAA至少包括SEQ ID NO:3的氨基酸204到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。

37. 根据权利要求34所述的rAAV, 其中所述hGAA至少包括SEQ ID NO:3的氨基酸123到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

38. 根据权利要求34所述的rAAV, 其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

39. 根据权利要求34所述的rAAV, 其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

40. 根据权利要求34到39中任一项所述的rAAV, 其中所述hGAA780I由SEQ ID NO:4或与SEQ ID NO:4至少95%相同的序列编码。

41. 根据权利要求34到39中任一项所述的rAAV, 其中所述hGAA780I由SEQ ID NO:5或与SEQ ID NO:5至少95%相同的序列编码。

42. 根据权利要求34所述的rAAV, 其中所述融合蛋白包括SEQ ID NO:6或与SEQ ID NO:6至少95%相同的序列。

43. 根据权利要求42所述的rAAV, 其中所述融合蛋白由SEQ ID NO:7或与SEQ ID NO:7至少95%相同的序列编码。

44. 根据权利要求34到43中任一项所述的rAAV, 其中所述载体基因组进一步包括背根神经节 (DRG) 特异性miR-183靶序列的至少两个串联重复序列, 其中所述至少两个串联重复序列包括至少一个第一miRNA靶序列和至少一个第二miRNA靶序列, 所述第一miRNA靶序列和所述第二miRNA靶序列可能相同或不同并且在3'处与编码所述融合蛋白的所述序列可操作地连接。

45. 根据权利要求44所述的rAAV, 其中所述miR-183靶序列是SEQ ID NO:26。

46. 根据权利要求44或45所述的rAAV, 其中所述miRNA靶序列中的两个或更多个miRNA靶序列由间隔子隔开, 并且一个或多个间隔子独立地选自 (i) GGAT; (ii) CACGTG; 以及 (iii) GCATGC。

47. 根据权利要求34到46中任一项所述的rAAV, 其中所述vIGF2肽包括与SEQ ID NO:32至少90%相同并且在选自SEQ ID NO:32的位置6、26、27、43、48、49、50、54、55和65的一个或多个位置处具有至少一个取代的氨基酸序列。

48. 根据权利要求47所述的rAAV, 其中所述至少一个取代选自SEQ ID NO:32的E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R。

49. 根据权利要求47或48所述的rAAV, 其中所述vIGF2肽在选自SEQ ID NO:32的位置26、27、43、48、49、50、54和55的两个或更多个位置处包括至少两个取代。

50. 根据权利要求49所述的rAAV, 其中所述至少两个取代选自SEQ ID NO:32的E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R。

51. 根据权利要求45到48中任一项所述的rAAV, 其中所述vIGF2肽在SEQ ID NO:32的位置1处包括N末端缺失。

52. 根据权利要求51所述的rAAV, 其中所述vIGF2肽在SEQ ID NO:32的位置1到4处包括N末端缺失。

53. 根据权利要求34到52中任一项所述的rAAV, 其中与天然IGF2肽相比, 所述vIGF2肽对胰岛素受体和IGFR1的亲和力降低或没有亲和力。

54. 根据权利要求34到53中任一项所述的rAAV, 其中所述vIGF2肽能够促进将hGAA780I

摄取到细胞中的溶酶体中。

55. 根据权利要求34到54中任一项所述的rAAV, 其中所述核酸序列进一步包括位于vIGF2核苷酸序列与编码hGAA780I的核酸序列之间的编码接头肽的接头序列。

56. 根据权利要求55所述的rAAV, 其中所述接头肽包括SEQ ID NO:55-60中的任一个。

57. 根据权利要求34到56中任一项所述的rAAV, 其中所述信号肽选自结合免疫球蛋白(BiP) 信号肽和Gaussia信号肽。

58. 根据权利要求57所述的rAAV, 其中所述BiP信号肽包括与SEQ ID NO:49-53中的任一个至少90%相同的氨基酸序列。

59. 根据权利要求58所述的rAAV, 其中所述BiP信号肽包括SEQ ID NO:49-53中的任一个的氨基酸序列。

60. 根据权利要求57所述的rAAV, 其中所述信号肽包括Gaussia信号肽。

61. 根据权利要求60所述的rAAV, 其中所述Gaussia信号肽包括与SEQ ID NO:54至少90%相同的氨基酸序列。

62. 根据权利要求61所述的rAAV, 其中所述Gaussia信号肽包括SEQ ID NO:54。

63. 根据权利要求34所述的rAAV, 其中所述载体基因组包括SEQ ID NO:30或与SEQ ID NO:30至少95%相同的序列。

64. 根据权利要求34到63中任一项所述的rAAV, 其中所述衣壳是进化枝F衣壳。

65. 根据权利要求64所述的rAAV, 其中所述衣壳是AAVhu68衣壳。

66. 一种核酸分子, 其包括编码根据权利要求1到33中任一项所述的表达盒的序列。

67. 一种核酸分子, 其包括SEQ ID NO:30或与SEQ ID NO:30至少95%相同的序列。

68. 根据权利要求66或67所述的核酸分子, 其中所述分子是质粒。

69. 一种宿主细胞, 其含有根据权利要求66到68中任一项所述的核酸分子。

70. 一种组合物, 其包括根据权利要求1到33中任一项所述的表达盒和药学上可接受的载体、赋形剂和/或悬浮剂中的至少一种。

71. 一种组合物, 其包括根据权利要求34到65中任一项所述的rAAV和药学上可接受的载体、赋形剂和/或悬浮剂中的至少一种。

72. 根据权利要求70或71所述的组合物, 其是被调配用于静脉内递送的悬浮液。

73. 根据权利要求70或71所述的组合物, 其是被调配用于鞘内、脑池内或脑室内施用的悬浮液。

74. 一种用于治疗患有庞贝病的患者的方法, 所述方法包括向所述患者递送根据权利要求1到33中任一项所述的表达盒或根据权利要求34到65中任一项所述的rAAV。

75. 根据权利要求74所述的方法, 其中所述表达盒和/或所述rAAV通过单独的途径共递送。

76. 根据权利要求74或75所述的方法, 其中通过静脉内和/或鞘内递送向所述患者施用所述表达盒和/或所述rAAV。

77. 一种用于改善患有 α -葡萄糖苷酶(GAA)缺乏症的患者的的心脏、呼吸和/或骨骼肌功能的方法, 所述方法包括向所述患者递送根据权利要求1到33中任一项所述的表达盒或根据权利要求34到65中任一项所述的rAAV。

78. 一种用于治疗患有庞贝病的患者的治疗方案, 所述治疗方案包括使用根据权利要

求1到33中任一项所述的表达盒或根据权利要求34到65中任一项所述的rAAV与免疫调节剂的组合。

79. 根据权利要求78所述的治疗方案,其中所述患者患有晚发性庞贝病。

80. 根据权利要求78所述的治疗方案,其中所述患者患有婴儿型庞贝病。

81. 根据权利要求78到80中任一项所述的治疗方案,其中所述患者接受与支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌力量训练(RMST)、酶替代疗法和/或膈肌起搏疗法的共疗法。

82. 一种根据权利要求1到33中任一项所述的表达盒或根据权利要求34到65中任一项所述的rAAV用于治疗患有庞贝病的患者的用途。

83. 一种根据权利要求1到33中任一项所述的表达盒或根据权利要求34到65中任一项所述的rAAV用于制备药物的用途。

可用于治疗庞贝病的组合物

背景技术

[0001] 如嗜神经AAV血清型(例如, AAV9)等嗜神经病毒已被证明在新生和幼年动物中以高剂量静脉内施用时可转导脊髓 α 运动神经元。这一观察使得最近成功地将AAV9递送应用于治疗患有脊髓性肌萎缩症的婴儿, 所述脊髓性肌萎缩症是一种特征在于下运动神经元的选择性死亡的遗传性运动神经元存活(SMN)蛋白缺陷。在一项涉及另一种嗜神经AAV(AAVhu68)的研究中, 观察到了类似的结果, 其中在全身施用和鞘内(脑脊液)施用后高效转导了脊髓运动神经元和背根神经节的感觉神经元(C.Hinderer等人,《人类基因疗法(Hum Gene Ther.)》2018年3月;29(3):285-298)。然而, DRG神经元的转导伴随着对这些感觉神经元的毒性和脊髓背束中的继发性轴突病变。无论衣壳血清型或转基因如何, 在静脉内和鞘内递送高剂量AAV载体后也遇到了类似的问题(参见J.Hordeaux,《分子疗法:方法与临床发展(Molecular Therapy:Methods&Clinical Development)》第10卷,第79-88页,2018年9月)。

[0002] 庞贝病(Pompe disease), 也称为II型糖原贮积症, 是由酸性 α -葡萄糖苷酶(GAA)基因中的突变引起的导致心脏(心肌病)、肌肉和运动神经元(神经肌肉疾病)中的糖原积累的溶酶体贮积病。在经典的婴儿庞贝病中, 严重的GAA活性丧失会引起多系统和早发性糖原贮积, 尤其是在心脏和肌肉内, 并在最初几年因心肺衰竭而死亡。婴儿庞贝病的特征还在于神经元(尤其是运动神经元)和神经胶质细胞内显著的糖原贮积。目前的护理标准(酶替代疗法(ERT))在矫正肌肉方面的效果欠佳, 并且无法穿过血脑屏障, 从而导致经典婴儿庞贝病长期幸存者的神经系统逐渐恶化。接受ERT的患者由于心脏矫正而活得更久, 显示出具有进行性神经系统表型的新自然病程。另外, 重组人GAA具有高度免疫原性, 并且由于骨骼肌摄取不良而必须大量给药。

[0003] 庞贝病的治疗有若干个未满足的需求, 包含需要矫正所述疾病的CNS组分, 需要改善肌肉矫正, 以及需要更有效、更少免疫原性和/或者更方便的针对当前ERT的替代方案。

发明内容

[0004] 在某些实施例中, 提供了一种表达盒, 其包括在调控序列的控制下编码嵌合融合蛋白的核酸序列, 所述嵌合融合蛋白包括信号肽和与人酸性 α -葡萄糖苷酶(hGAA)融合的vIGF2肽, 所述hGAA至少包括hGAA780I的活性位点, 所述调控序列指导所述嵌合融合蛋白的表达, 其中位置780基于SEQ ID NO:3中的氨基酸序列的位置的编号。在某些实施例中, 所述hGAA至少包括SEQ ID NO:3(hGAA780I)的氨基酸204到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。在某些实施例中, 所述hGAA至少包括SEQ ID NO:3的氨基酸204到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。在某些实施例中, 所述hGAA至少包括SEQ ID NO:3的氨基酸123到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。在某些实施例中, 所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有I1e的序列。在某些实施例中, 所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸890或与SEQ

ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,所述表达盒进一步包括miR靶序列的至少两个串联重复序列,其中所述至少两个串联重复序列包括至少一个第一miRNA靶序列和至少一个第二miRNA靶序列,所述第一miRNA靶序列和所述第二miRNA靶序列可能相同或不同并且在3'处与编码所述融合蛋白的所述序列可操作地连接。

[0005] 在某些实施例中,本文提供的表达盒由选自以下的病毒载体携带:重组细小病毒、重组慢病毒、重组逆转录病毒和重组腺病毒。在某些实施例中,所述重组细小病毒是进化枝F腺相关病毒,任选地AAVhu68。在某些实施例中,本文提供的表达盒由选自以下的非病毒载体携带:裸DNA、裸RNA、无机颗粒、脂质颗粒、基于聚合物的载体或基于壳聚糖的调配物。

[0006] 在某些实施例中,本文提供了一种重组腺相关病毒(rAAV),其包括:(a) AAV衣壳,所述AAV衣壳靶向肌肉、心脏和中枢神经系统中的至少一种的细胞;以及(b)载体基因组,所述载体基因组包装在所述AAV衣壳中,所述载体基因组包括在调控序列的控制下编码嵌合融合蛋白的核酸序列,所述嵌合融合蛋白包括信号肽和与hGAA融合的vIGF2肽,所述hGAA至少包括hGAA780I的活性位点,所述调控序列指导所述嵌合融合蛋白的表达,其中位置780基于SEQ ID NO:3中的氨基酸序列的位置的编号。在某些实施例中,所述hGAA至少包括SEQ ID NO:3(hGAA780I)的氨基酸204到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,所述hGAA至少包括SEQ ID NO:3的氨基酸204到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,所述hGAA至少包括SEQ ID NO:3的氨基酸123到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,其中所述hGAA至少包括SEQ ID NO:3的氨基酸70到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中,所述rAAV载体基因组进一步包括背根神经节(DRG)特异性miR-183靶序列的至少两个串联重复序列,其中所述至少两个串联重复序列包括至少一个第一miRNA靶序列和至少一个第二miRNA靶序列,所述第一miRNA靶序列和所述第二miRNA靶序列可能相同或不同并且在3'处与编码所述融合蛋白的所述序列可操作地连接。

[0007] 在某些实施例中,提供了一种组合物,其包括如本文所述的编码hGAA780I融合蛋白的表达盒和药学上可接受的载体、赋形剂和/或悬浮剂中的至少一种。

[0008] 在某些实施例中,提供了一种组合物,其包含rAAV,所述rAAV包括如本文所述的编码hGAA780I融合蛋白的表达盒和药学上可接受的载体、赋形剂和/或悬浮剂中的至少一种。

[0009] 在某些实施例中,提供了一种用于治疗患有庞贝病的患者和/或用于改善患有 α -葡萄糖苷酶(GAA)缺乏症的患者的心脏、呼吸和/或骨骼肌功能的方法。此方法包括向所述患者递送如本文所述的表达盒、rAAV或组合物。所述表达盒、rAAV或组合物可以静脉内和/或通过鞘内、脑池内或脑室内施用来递送。另外或可替代地,此类基因疗法可以涉及直接递送到心脏(heart)(心脏(cardiac))、递送到肺(鼻内、吸入、气管内)和/或肌肉内注射。其中之一可以是表达盒、载体或组合物的唯一施用途径,或其它递送途径共施用。

[0010] 用于治疗患有庞贝病的患者的治疗方案可以包括向所述患者单独地或与共疗法组合,例如与以下中的一种或多种组合来递送本文所述的表达盒、rAAV或组合物:免疫调节剂、支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌力量训练(RMST)、酶替代疗法和/或膈肌起

搏疗法。

[0011] 在某些实施例中,提供了用于产生本文所述的表达盒和/或rAAV的核酸分子和宿主细胞。

[0012] 在某些实施例中,提供了一种表达盒、rAAV和/或组合物在制备药物中的用途。

[0013] 在某些实施例中,提供了一种适用于治疗患有庞贝病的患者和/或用于改善患有 α -葡萄糖苷酶(GAA)缺乏症的患者的肝脏、呼吸和/或骨骼肌功能的表达盒、rAAV和/或组合物。

[0014] 根据本发明的以下详细描述,本发明的其它方面和优点将显而易见。

附图说明

[0015] 图1A和图1B示出了在静脉内施用具有在CB6启动子(第三列)、CAG启动子(第四列)或UbC启动子(最后一列)的指导下的hGAAV780I的经工程化的编码序列的各种AAVhu68.hGAA后四周,庞贝病(-/-)小鼠的肝脏中的hGAA活性。(图1A)低剂量(1×10^{11} GC)。(图1B)高剂量(1×10^{12})。

[0016] 图2A和图2B示出了在静脉内施用具有在CB6启动子(第三列)、CAG启动子(第四列)或UbC启动子(最后一列)的指导下的hGAAV780I的经工程化的编码序列的各种AAVhu68.hGAA后四周,庞贝病(-/-)小鼠的心脏中的hGAA活性。(图2A)低剂量(1×10^{11} GC)。(图2B)高剂量(1×10^{12})。

[0017] 图3A和图3B示出了在静脉内施用具有在CB6启动子(第三列)、CAG启动子(第四列)或UbC启动子(最后一列)的指导下的hGAAV780I的经工程化的编码序列的各种AAVhu68.hGAA后四周,庞贝病(-/-)小鼠的骨骼肌(四头肌)中的hGAA活性。(图3A)低剂量(1×10^{11} GC)。(图3B)高剂量(1×10^{12})。

[0018] 图4A和图4B示出了在静脉内施用具有在CB6启动子(第三列)、CAG启动子(第四列)或UbC启动子(最后一列)的指导下的hGAAV780I的经工程化的编码序列的各种AAVhu68.hGAA后四周,庞贝病(-/-)小鼠的脑中的hGAA活性。(图4A)低剂量(1×10^{11} GC)。(图4B)高剂量(1×10^{12})。在CB7活性下表达的载体在两种剂量下都具有较低的活性,而在CAG或UbC启动子下表达的载体在较高剂量下具有相当的活性。

[0019] 图5A-图5H示出了在AAVhu68.hGAA递送后四周庞贝病小鼠的心脏的组织学(PAS染色显示糖原贮积)。产生并评估了含有五种不同hGAA表达盒的rAAVhu68载体。媒介对照庞贝病(-/-)(图5D)和野生型(+/-)(图5A)小鼠接受PBS注射。“hGAA”是指由野生型序列编码的具有天然信号肽的参考天然酶(hGAAV780)(图5B)。“BiP-vIGF2.hGAAco”是指参考hGAAV780蛋白的经工程化的编码序列,所述经工程化的编码序列含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体融合的BiP信号肽(图5C)。“hGAAcoV780I”是指由经工程化的序列编码并含有天然信号肽的hGAAV780I变体(图5E)。“BiP-vIGF2.hGAAcoV780I”是指含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体和由经工程化的序列编码的hGAAV780I融合的BiP信号肽的hGAAcoV780I(图5F)。“Sp7. Δ 8.hGAAcoV780I”是指前35个AA缺失的hGAAV780I变体,所述变体由与先前构建体相同但含有编码B2胰凝乳蛋白酶原信号肽而非天然信号肽的序列的经工程化的序列编码(图5G)。(图5H)盲法组织病理学半定量严重性评分。通过职业验证的兽医病理学家以不

知情的方式审查了载玻片,并基于糖原贮积和自噬积累建立严重性评分。

[0020] 图6A-图6H示出了在施用编码各种hGAA的AAVhu68 (2.5×10^{13} GC/kg) 后四周的庞贝病小鼠的四头肌的组织学(PAS染色)的结果。对照庞贝病(-/-) (图6D) 和野生型(+/+) (图6A) 小鼠接受PBS注射。“hGAA”是指由野生型序列编码的具有天然信号肽的参考天然酶(hGAAV780) (图6B)。“hGAAcoV780I”是指由经工程化的序列编码并含有天然信号肽的hGAAV780I变体(图6E)。“Sp7. Δ 8.hGAAcoV780I”是指前35个AA缺失的hGAAV780I变体,所述变体由与先前构建体相同但含有编码B2胰凝乳蛋白酶原信号肽而非天然信号肽的序列的经工程化的序列编码(图6F)。“BiP-vIGF2.hGAAco”是指含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体融合的BiP信号肽且由经工程化的序列编码的参考hGAAV780 (图6C)。“BiP-vIGF2.hGAAcoV780I”是指含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体和由经工程化的序列编码的hGAAV780I融合的BiP信号肽的hGAAV780I (图6G)。(图6H) 盲法组织病理学半定量严重性评分。通过职业验证的兽医病理学家以不知情的方式审查了载玻片,并基于糖原贮积和自噬积累建立严重性评分。评分为0意指无病变;1意指平均少于9%的肌肉纤维受贮积影响;2意指10%到49%;3意指50%到75%,并且4意指76%到100%。

[0021] 图7A-图7H示出了在以 2.5×10^{12} GC/Kg (即,比图6A-图6H中的剂量低10倍的剂量) 施用编码各种hGAA的AAVhu68后四周,来自庞贝病小鼠的四头肌的组织学(过碘酸希夫(Periodic acid-Schiff,PAS) 染色)的结果。对照庞贝病(-/-) (图7D) 和野生型(+/+) (图7A) 小鼠接受PBS注射。“hGAA”是指由野生型序列编码的具有天然信号肽的参考天然酶(hGAAV780) (图7B)。“hGAAcoV780I”是指由经工程化的序列编码并含有天然信号肽的hGAAV780I变体(图7E)。“Sp7. Δ 8.hGAAcoV780I”是指前35个AA缺失的hGAAV780I变体,所述变体由与先前构建体相同但含有编码B2胰凝乳蛋白酶原信号肽而非天然信号肽的序列的经工程化的序列编码(图7F)。“BiP-vIGF2.hGAAco”是指含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体融合的BiP信号肽且由经工程化的序列编码的参考hGAAV780 (图7C)。“BiP-vIGF2.hGAAcoV780I”是指含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体和由经工程化的序列编码的hGAAV780I融合的BiP信号肽的hGAAV780I (图7G)。(图7H) 盲法组织病理学半定量严重性评分。通过职业验证的兽医病理学家以不知情的方式审查了载玻片,并基于糖原贮积和自噬积累建立严重性评分。评分为0意指无病变;1意指平均少于9%的肌肉纤维受贮积影响;2意指10%到49%;3意指50%到75%,并且4意指76%到100%。

[0022] 图8示出了在施用(2.5×10^{12} GC/kg) AAVhu68后四周的庞贝病小鼠的脊髓组织学(PAS和劳克坚牢蓝(luxol fast blue) 染色)的结果,所述AAVhu68具有编码天然hGAA或含有前35个AA的缺失并进一步具有与对胰岛素受体的亲和力较低的IGF2变体和由经工程化的序列编码的hGAAV780I融合的BiP信号肽的hGAAV780I (“BiP-vIGF2.hGAAcoV780I”) 的序列。对脊髓切片进行盲法组织病理学半定量严重性评分。

[0023] 图9A-图9C示出了血浆中的hGAA活性以及与IGF2/CI-MPR的结合。以低剂量(2.5×10^{12} GC) 向庞贝病小鼠施用编码野生型hGAA或BiP-vIGF2.hGAA的载体。(图9A,图9B) 在静脉内施用后四周,在血浆中检测到高水平的野生型和经工程化的hGAA活性。(图9C) 经工程化的hGAA与CI-MPR有效结合。

[0024] 图10示出了在以 2.5×10^{12} GC/Kg (LD) 的剂量施用AAVhu68构建体四周后, 庞贝病小鼠中的糖原清除和自噬积累消退。用DAPI和抗LC3B抗体染色的腓肠肌的石蜡切片。

[0025] 图11示出了BiP-vIGF2.hGAAcoV780I.4XmiR183构建体的示意图。

[0026] 图12示出了在以高剂量(HD: 2.5×10^{13} GC/kg) 或低剂量(LD: 2.5×10^{12} GC/kg) 静脉内施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 (含有drg脱靶序列miR183的四个拷贝) 后四周, 庞贝病小鼠的脑干中的糖原贮积(PAS, 劳克蓝(luxol blue) 染色)。箭头示出神经元内的PAS阳性贮积。

[0027] 图13示出了在以高剂量(HD: 2.5×10^{13} GC/kg) 或低剂量(LD: 2.5×10^{12} GC/kg) 静脉内施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后四周, 庞贝病小鼠的脊髓中的糖原贮积(PAS, 劳克蓝染色)。箭头示出神经元内的PAS阳性贮积。

[0028] 图14示出了在以高剂量(HD: 2.5×10^{13} GC/kg) 或低剂量(LD: 2.5×10^{12} GC/kg) 静脉内施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后四周, 庞贝病小鼠的四头肌中的糖原贮积(PAS染色)。

[0029] 图15示出了在以高剂量(HD: 2.5×10^{13} GC/kg) 或低剂量(LD: 2.5×10^{12} GC/kg) 静脉内施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后四周, 庞贝病小鼠的心脏中的糖原贮积(PAS染色)。

[0030] 图16示出了在以高剂量(HD: 2.5×10^{13} GC/kg) 或低剂量(LD: 2.5×10^{12} GC/kg) 静脉内施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后四周, 庞贝病小鼠的四头肌中自噬空泡标志物LC3b的表达。

[0031] 图17示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I (左) 或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 (右) 后35天恒河猴的颈DRG中的hGAA表达(hGAA的免疫组织化学) 的代表性图像。

[0032] 图18示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I (左) 或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 (右) 后35天恒河猴的腰DRG中的hGAA表达(hGAA的免疫组织化学) 的代表性图像。

[0033] 图19示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I (左) 或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 (右) 后35天恒河猴的脊髓下运动神经元中的hGAA表达(hGAA的免疫组织化学) 的代表性图像。

[0034] 图20示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I (左) 或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 (右) 后35天恒河猴的心脏中的hGAA表达(hGAA的免疫组织化学) 的代表性图像。

[0035] 图21A-图21C示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后35天恒河猴的颈段(图21A)、胸段(图21B) 和腰段(图21C) 的DRG中DRG神经元变性和炎性细胞浸润的组织病理学评分。AAVhu68载体在如前所述的荧光镜引导下在注射到小脑延髓池中的总体积为1mL的无菌人工CSF (媒剂) 中递送(Katz等人,《人类基因治疗方法(Hum Gene Ther.Methods)》, 2018, 29:212-9)。对载体组不知情的通过职业验证的兽医病理学家确立了严重性等级, 定义0为无病变, 1为轻微(<10%), 2为轻度(10-25%), 3为中度(25-50%), 4为显著(50-95%), 并且5为重度(>95%)。每个数据点表示一个DRG。每段和每只动物至少有五个DRG被

评分。

[0036] 图22A-图22C示出了在以 3×10^{13} GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183后,恒河猴的AST水平(图22A)、ALT水平(图22B)和血小板计数(图22C)。

[0037] 图23示出了以 3×10^{13} GC的高剂量被施用(ICM) AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183的NHP在注射后0-35天的血浆hGAA活性水平。

[0038] 图24A-图24G示出了在基线和第35天对被施用(ICM, 3×10^{13} GC) AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183的NHP的神经传导速度测试的结果。

[0039] 图25A和图25B示出了在7个月大时在疾病的晚期阶段进行处理并且在基线时已经有症状的庞贝病小鼠中从载体注射(第0天)到注射后180天的体重纵向随访。所述小鼠通过以下替代施用途径和剂量水平接受AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I:脑室内(ICV),在高剂量(HD) (1×10^{11} GC)或低剂量(LD) (5×10^{10} GC)下;静脉内(IV),在HD (5×10^{13} GC/Kg)或LD (1×10^{13} GC/Kg)下;以及ICV和IV的组合,在低剂量或高剂量下。描绘了平均值和标准偏差。每个时间点的统计分析通过KO PBS对照组与其它组之间的Wilcoxon-Mann-Whitney检验进行。 $*p < 0.05$; $**p < 0.01$

[0040] 图26A和图26B示出了在7个月大时在疾病的晚期阶段进行处理并且在基线时已经有症状的庞贝病小鼠中从载体注射(第0天)到注射后180天相对于体重纵向随访的握力。(图26A)小鼠通过以下替代施用途径和剂量水平接受AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I:脑室内(ICV),在高剂量下(ICV HD: 1×10^{11} GC);静脉内(IV),在高剂量下(IV HD: 5×10^{13} GC/Kg);以及ICV和IV高剂量与ICV和IV低剂量的组合。使用握力计(IITC生命科学(IITC Life Science))在不同时间点测量握力。握力计中的换能器与金属丝网连接,所述金属丝网与阳极化基板连接。通过动物的尾巴来抓住所述动物并且使所述动物轻轻地网格之上移动,直到所述动物用其四只爪子抓住网格。进行了三次握力测量,并且这些读数的平均值表示动物在所述特定时间的握力。(图26B)第180天的结果示出了相对于IV HD, IV+ICV HD的增量益处。值通过动物体重进行归一化。每组N=4个雄性和4个雌性。每个时间点的统计分析通过单因素方差分析(1-way ANOVA)(图26A)或双因素方差分析(2-way ANOVA)(图26B)、与KO PBS对照组相比的事后多重比较检验来确定。 $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

[0041] 图27A和图27B示出了IV、ICV或IV和ICV(双途径)施用了AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I载体的庞贝病小鼠的体积描记法的结果。(图27A) 5%CO₂攻击。(图27B) 7%CO₂攻击。

[0042] 图28示出了在高剂量(HD: 1×10^{11} GC)或低剂量(LD: 5×10^{10} GC) ICV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后,有症状后的庞贝病小鼠的四头肌、心脏和脊髓中的糖原贮积。

[0043] 图29示出了在高剂量(HD: 5×10^{13} GC/Kg)或低剂量(LD: 1×10^{13} GC/Kg) IV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后,有症状后的庞贝病小鼠的四头肌、心脏和脊髓中的糖原贮积。

[0044] 图30A-图30C示出了在第30天(图30A)、第60天(图30B)和第90天(图30C), IV、ICV或IV和ICV(双途径)施用了AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I载体的庞贝病小鼠的血浆中的hGAA活性。

[0045] 图31示出了用于评估NHP中单施用途径(IV或ICM)和双施用途径(IV+ICM)的研究设计。

[0046] 图32A-图32H示出了在IV或ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后,NHP的血浆和CSF中的hGAA和hGAA活性的检测。

[0047] 图33A-图33F示出了在IV(1e13 GC/Kg或5e13 GC/Kg)或ICM(1e13 GC或3e13GC)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后,恒河猴的DRG神经元变性和炎性细胞浸润(图33A-图33C)和脊髓轴突病变(图33D-图33F)的组织病理学评分。对载体组不知情的通过职业验证的兽医病理学家确立了严重性等级,定义0为无病变,1为轻微(<10%),2为轻度(10-25%),3为中度(25-50%),4为显著(50-95%),并且5为重度(>95%)。

[0048] 图34示出了在低剂量(IV-1e13 GC/Kg, ICM-1e13 GC)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后,恒河猴的四头肌、心脏和脊髓中的hGAA表达(hGAA的免疫组织化学)的代表性图像。

具体实施方式

[0049] 提供了用于将融合蛋白递送到患有庞贝病的患者的组合物,所述融合蛋白包括信号肽和与hGAA780I酶的至少活性部分融合的vIGF2肽。本文描述了制备和使用所述组合物的方法,包含用于用这些组合物治疗患者的方案。

[0050] 如本文所使用的,术语“庞贝病”也称为麦芽糖酶缺乏症、II型糖原贮积病(GSDII)或II型糖原贮积症,其旨在指特征在于由编码酸性 α -葡萄糖苷酶的GAA基因中的突变引起的溶酶体酶酸性 α -葡萄糖苷酶(GAA)完全缺失或部分缺乏的遗传性溶酶体贮积病症。所述术语包含但不限于所述疾病的早发型和晚发型形式,包含但不限于婴儿型、青年型和成年型庞贝病。

[0051] 应当理解,希腊字母“alpha”和符号“ α ”在本说明书中可互换使用。类似地,希腊字母“delta”和“ Δ ”在本说明书中可互换使用。

[0052] 如本文中所使用的,术语“酸性 α -葡萄糖苷酶”或“GAA”是指水解糖原、麦芽糖和异麦芽糖的D-葡萄糖单元之间的 α -1,4键的溶酶体酶。替代性名称包含但不限于溶酶体 α -葡萄糖苷酶(EC:3.2.1.20);葡糖淀粉酶;1,4- α -D-葡聚糖葡萄糖水解酶;淀粉葡萄糖苷酶; γ -淀粉酶和外切-1,4- α -葡萄糖苷酶。人酸性 α -葡萄糖苷酶由GAA基因(美国国家生物技术信息中心(National Centre for Biotechnology Information,NCBI)基因ID 2548)编码,所述基因已经被定位到染色体17的长臂(位置17q25.2-q25.3)。在氨基酸残基516-521处的保守六肽WIDMNE对于酸性 α -葡萄糖苷酶蛋白的活性是必需的。术语“hGAA”是指人GAA的编码序列。

[0053] 如本文所使用的,“rAAV.hGAA”是指具有AAV衣壳的rAAV,所述AAV衣壳中包装有载体基因组,所述载体基因组至少含有GAA酶(例如,780I变体、包括信号肽和与hGAA780I酶的至少活性部分融合的vIGF2肽的融合蛋白)的编码序列。rAAVhu68.hGAA或rAAVhu68.hGAA是指其中AAV衣壳是本文定义的AAVhu68衣壳的rAAV。

[0054] 参考全长hGAA的编号,在氨基酸位置1-27处存在信号肽。另外,所述酶已经与多个成熟蛋白,即氨基酸位置70到952处的成熟蛋白、定位于氨基酸位置123到952处的76kD成熟蛋白和氨基酸204到氨基酸952处的70kD成熟蛋白相关联。“活性催化位点”包括六肽WIDMNE (SEQ ID NO:3的氨基酸残基516-521)。在某些实施例中,可以选择更长的片段,例如位置516到616。其它活性位点包含配体结合位点,所述配体结合位点可以定位于位置376、404、405、441、481、516、518、519、600、613、616、649、674中的一个或多个位置处。

[0055] 除非另有说明,否则术语“hGAA780I”或“hGAAV780I”是指具有在SEQ ID NO:3中复制的氨基酸序列的全长前原蛋白。在一些情况下,术语hGAAco780I或hGAAcoV780I用于指代编码hGAA780I的经工程化的序列。与前一段中描述的hGAA参考蛋白相比,hGAA780I在位置780处具有异亮氨酸(Ile或I),参考hGAA在所述位置处含有缬氨酸(Val或V)。出人意料地发现,与在文献中被广泛描述为“参考序列”的在位置780处具有缬氨酸的hGAA序列(hGAAV780)相比,此hGAA780I具有更好的效果和改善的安全性曲线。例如,如在图5A-图5H中可以看到, hGAAV780参考序列诱导毒性(纤维化心肌炎),在相同剂量的hGAA780I下未看到所述毒性。因此,使用hGAA780I可以减少或消除接受hGAA疗法的患者的纤维化心肌炎。hGAA信号肽、成熟蛋白、活性催化位点和结合位点的位置可以基于SEQ ID NO:3中复制的hGAA780I中的类似位置来确定,即氨基酸位置1到27处的信号肽;氨基酸位置70到952处的成熟蛋白;定位于氨基酸位置123到952处的76kD成熟蛋白和氨基酸204到氨基酸952处的70kD成熟蛋白;氨基酸残基516-521处的包括六肽WIDMNE (SEQ ID NO:61)的“活性催化位点”;其它活性位点包含配体结合位点,所述配体结合位点可以定位于位置376、404、405、441、481、516、518、519、600、613、616、649、674中的一个或多个位置处。

[0056] 在某些实施例中,可以选择具有与SEQ ID NO:3的hGAA780I至少95%相同、与所述hGAA780I至少97%相同或与所述hGAA780I至少99%相同的序列的hGAA780I。在某些实施例中,提供了与SEQ ID NO:3的成熟hGAA780I蛋白具有至少95%、至少97%或至少99%同一性的序列。在某些实施例中,与hGAA780I具有至少95%到至少99%同一性的所述序列保留了活性催化位点的序列,而没有任何变化。在某些实施例中,与SEQ ID NO:3的hGAA780I具有至少95%到至少99%同一性的序列的特征在于:当在适当的动物模型中测试时,所述序列与参考hGAAV780相比具有改善的生物学效应和更好的安全性曲线。在某些实施例中,GAA活性测定可以如先前描述的那样(参见例如J. Hordeaux等人,《神经病理学通讯学报 (Acta Neuropathological Communications)》, (2107) 5:66) 或使用其它合适的方法进行。在某些实施例中,hGAA780I酶在hGAA氨基酸序列中的其它位置含有修饰。突变体的实例可以含有例如美国专利9,920,307中描述的那些。在某些实施例中,此类突变体hGAA780I可以至少保留活性催化位点:WIDMNE (SEQ ID NO:61) 和780I区中的氨基酸,如下所述。

[0057] 在某些实施例中,提供了新的hGAA780I融合蛋白,所述融合蛋白包括不同于天然hGAA信号肽的前导肽。在某些实施例中,此外源前导肽优选地是人源的并且可以包含例如IL-2前导肽。在某些实施例中可行的特定外源信号肽包含来自胰凝乳蛋白酶原B2的氨基酸1-20、人 α -1-抗胰蛋白酶的信号肽、来自艾杜糖醛酸-2-硫酸酯酶的氨基酸1-25和来自蛋白酶CI抑制剂的氨基酸1-23。参见例如WO 2018046774。其它信号/前导肽可以天然存在于免疫球蛋白(例如,IgG)、细胞因子(例如,IL-2、IL12、IL18等)、胰岛素、白蛋白、 β -葡萄糖醛酸酶、碱性蛋白酶或纤连蛋白分泌信号肽等中。还参见例如signalpeptide.de/index.php?m

=listspdb_mammalia。

[0058] 此类嵌合hGAA780I可以具有外源前导来代替整个27aa天然信号肽。任选地，hGAA780I酶的N末端截短可能仅缺少信号肽的一部分(例如，约2到约25个氨基酸的缺失或其间的值)、整个信号肽或长于信号肽的片段(例如，基于SEQ ID NO:3的编号，至多到氨基酸70。任选地，此类酶可以含有长度为约5、10、15或20个氨基酸的C末端截短。

[0059] 在某些实施例中，提供了一种新的融合蛋白，其包括与融合配偶体结合的成熟hGAA780I蛋白(aa 70到952)、成熟70kD蛋白(aa 123到aa 952)或成熟76kD蛋白(aa 204到952)。任选地，所述融合蛋白进一步包括对hGAA非天然的信号肽。进一步任选地，这些实施例之一可以进一步含有长度为约5、10、15或20个氨基酸的C末端截短。

[0060] 在某些实施例中，包括hGAA780I蛋白的融合蛋白至少包括SEQ ID NO:3(hGAA780I)的氨基酸204到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白至少包括SEQ ID NO:3的氨基酸204到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白至少包括SEQ ID NO:3的氨基酸123到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中，所述hGAA780I酶至少包括SEQ ID NO:3的氨基酸70到氨基酸952或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。在某些实施例中，所述hGAA780I蛋白至少包括SEQ ID NO:3的氨基酸70到氨基酸890或与SEQ ID NO:3至少95%相同的在位置780处具有Ile的序列。

[0061] 在某些实施例中，融合蛋白包括信号序列和前导序列，并且与SEQ ID NO:7具有至少95%同一性、至少97%同一性或至少99%同一性的hGAA780I序列的活性位点没有变化和/或距活性位点的N端和/或C端氨基酸3到12个氨基酸没有变化。在优选实施例中，经工程化的hGAA表达盒至少编码以下的人hGAA780I片段：T-Val(V)-P-Ile(780I)-Glu(E)-Ala(A)-Leu(L)(SEQ ID NO:62)。在某些实施例中，经工程化的hGAA表达盒编码更长的人hGAA780I片段：Gln(Q)-T-V-P-780I-E-A-L-Gly(G)(SEQ ID NO:63)。在某些实施例中，经工程化的hGAA表达盒编码至少与以下相对应的片段：PLGT-Trp(W)-Tyr(Y)-Asp(D)-LQTVP-780I-EALG-(Ser或S)-L-PPPPAA序列(SEQ ID NO:64)。类似地，在优选实施例中，活性结合位点(SEQ ID NO:3的aa 518到521)没有氨基酸变化。在某些实施例中，位置600、616和/或674处的结合位点保持不变。在某些实施例中，融合蛋白包括信号肽、任选的vIGF+2GS延伸序列、任选的ER蛋白水解肽和hGAA的前35个氨基酸缺失(即，缺乏天然信号肽和氨基酸28到35)的hGAA780I变体。

[0062] 在某些实施例中，提供了分泌的经工程化的GAA，所述分泌的经工程化的GAA包括BiP信号肽、IGF2+2GS延伸序列和hGAA 780I的氨基酸61到952(其中hGAA780I的氨基酸1到60缺失)。在某些实施例中，本文提供包括SEQ ID NO:6或与SEQ ID NO:6至少95%相同的序列的融合蛋白。在某些实施例中，融合蛋白由SEQ ID NO:7或与SEQ ID NO:7至少95%相同的序列编码。在某些实施例中，融合蛋白包括SEQ ID NO:4的序列或与所述序列至少95%相同的序列。在某些实施例中，融合蛋白包括SEQ ID NO:5的序列或与所述序列至少95%相同的序列。

[0063] 下文进一步描述本文提供的融合蛋白的组分。

[0064] 结合CI-MPR的肽

[0065] 本文提供结合CI-MPR的肽(例如,vIGF2肽)。包括此类肽和hGAA780I蛋白的融合蛋白当从基因疗法载体表达时将hGAA780I靶向需要它的细胞,增加此类细胞的细胞摄取并将治疗性蛋白靶向亚细胞位置(例如,溶酶体)。在一些实施例中,所述肽与hGAA780I蛋白的N端融合。在一些实施例中,所述肽与hGAA780I蛋白的C端融合。在一些实施例中,所述肽是vIGF2肽。一些vIGF2肽保持与CI-MPR的高亲和力结合,而其对IGF1受体、胰岛素受体和IGF结合蛋白(IGFBP)的亲和力降低或消除。因此,与野生型IGF2相比,一些变体IGF2肽具有更高的选择性并降低了安全性风险。本文中的vIGF2肽包含具有SEQ ID NO:46的氨基酸序列的那些肽。变体IGF2肽进一步包含与野生型IGF2(SEQ ID NO:34)相比在位置6、26、27、43、48、49、50、54、55或65处具有变体氨基酸的那些肽。在一些实施例中,vIGF2肽的序列具有一个或多个选自由以下组成的组的取代:E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R。在一些实施例中,vIGF2肽的序列具有E6R取代。在一些实施例中,vIGF2肽的序列具有F26S取代。在一些实施例中,vIGF2肽的序列具有Y27L取代。在一些实施例中,vIGF2肽的序列具有V43L取代。在一些实施例中,vIGF2肽的序列具有F48T取代。在一些实施例中,vIGF2肽的序列具有R49S取代。在一些实施例中,vIGF2肽的序列具有S50I取代。在一些实施例中,vIGF2肽的序列具有A54R取代。在一些实施例中,vIGF2肽的序列具有L55R取代。在一些实施例中,vIGF2肽的序列具有K65R取代。在一些实施例中,vIGF2肽的序列具有E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R和L55R取代。在一些实施例中,vIGF2肽具有N末端缺失。在一些实施例中,vIGF2肽具有一个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有两个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有三个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有四个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有四个氨基酸的N末端缺失和E6R、Y27L和K65R取代。在一些实施例中,vIGF2肽具有四个氨基酸的N末端缺失和E6R和Y27L取代。在一些实施例中,vIGF2肽具有五个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有六个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有七个氨基酸的N末端缺失。在一些实施例中,vIGF2肽具有七个氨基酸的N末端缺失和Y27L和K65R取代。

[0066]

IGF2 氨基酸序列 (变体残基加下划线)		
肽	序列	SEQ ID NO:
野生型	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE	32
F26S	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE	33
Y27L	AYRPSETLCGGELVDTLQFVCGDRGFLFSRPASRV	34

[0067]

	SRRSRGIVEECCFRSCDLALLETYCATPAKSE	
V43L	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGILEECCFRSCDLALLETYCATPAKSE	35
F48T	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE	36
R49S	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE	37
S50I	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRICDLALLETYCATPAKSE	38
A54R	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLRLLETYCATPAKSE	39
L55R	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLARLETYCATPAKSE	40
F26S、Y27L、V43L、F48T、 R49S、S50I、A54R、L55R	AYRPSETLCGGELVDTLQFVCGDRGSLFSRPASRV SRRSRGILEECCTSICDLRRLETYCATPAKSE	41
Δ1-6、Y27L、K65R	TLCGGELVDTLQFVCGDRGFLFSRPASRVSRRSRG IVEECCFRSCDLALLETYCATPARSE	42
Δ1-7、Y27L、K65R	LCGGELVDTLQFVCGDRGFLFSRPASRVSRRSRGI VEECCFRSCDLALLETYCATPARSE	43
Δ1-4、E6R、Y27L、K65R	SRTLCCGGELVDTLQFVCGDRGFLFSRPASRVSRRS RGIVEECCFRSCDLALLETYCATPARSE	44
Δ1-4、E6R、Y27L	SRTLCCGGELVDTLQFVCGDRGFLFSRPASRVSRRS RGIVEECCFRSCDLALLETYCATPAKSE	45
E6R	AYRPSRTLCCGGELVDTLQFVCGDRGFYFSRPASRV SRRSRGIVEECCFRSCDLALLETYCATPAKSE	46

[0068]

IGF2 DNA 编码序列		
肽	DNA 序列	SEQ ID NO
成熟 WT IGF2	GCTTACCGCCCCAGTGAGACCCTGTGCGGCGGG GAGCTGGTGGACACCTCCAGTTCGTCTGTGGG GACCGCGGCTTCTACTTCAGCAGGCCCGCAAGC CGTGTGAGCCGTCGCAGCCGTGGCATCGTTGAG GAGTGCTGTTTCCGCAGCTGTGACCTGGCCCTC CTGGAGACGTA CTGTGCTACCCCGCCAAGTCC GAG	47
vIGF2 Δ1-4、E6R、Y27L、 K65R	TCTAGAACACTGTGCGGAGGGGAGCTTGTAGA CACTCTTCAGTTCGTGTGTGGAGATCGCGGGTT CCTCTTCTCTCGCCCCGCTTCCAGAGTTTCACG GAGGTCTAGGGGTATAGTAGAGGAGTGTTGTTT CAGGTCTGTGACTTGCGCTCCTCGAGACCTA TTGCGCGACGCCAGCCAGGTCCGAA	48

[0069] 信号肽

[0070] 在一些实施例中，本文提供的组合物进一步包括信号肽，所述信号肽改善 hGAA780I 从用基因疗法构建体转导的细胞的分泌。在一些实施例中，信号肽改善治疗性蛋白的蛋白加工，并促进新生多肽-核糖体复合物向 ER 的易位并确保适当的共翻译和翻译后修饰。在一些实施例中，信号肽定位于 (i) 信号翻译起始序列的上游位置，(ii) 翻译起始序列与治疗性蛋白之间，或 (iii) 治疗性蛋白质的下游位置。可用于基因疗法构建体的信号肽包含但不限于来自 HSP70 蛋白家族 (例如，HSPA5、热休克蛋白家族 A 成员 5) 和 Gaussia 信号肽及其变体的结合免疫球蛋白 (BiP) 信号肽。这些信号肽对信号识别颗粒具有超高亲和力。下表提供了 BiP 和 Gaussia 氨基酸序列的实例。在一些实施例中，信号肽具有与选自由 SEQ ID No: 49-53 组成的组中的序列至少 90% 相同的氨基酸序列。在一些实施例中，信号肽与选自由 SEQ ID No: 49-53 组成的组的序列相差 5 个或更少、4 个或更少、3 个或更少、2 个或更少、或

1个氨基酸。

[0071]	信号肽序列		
	信号肽	氨基酸序列	SEQ ID NO:
	天然人类 BiP	MKLSLVAAML LLL SAARA	49
	经修饰的 BiP-1	MKLSLVAAML LLL SLVAAML LLL SAARA	50
	经修饰的 BiP-2	MKLSLVAAML LLL WVALL LLL SAARA	51
	经修饰的 BiP-3	MKLSLVAAML LLL SLVAL LLL SAARA	52
	经修饰的 BiP-4	MKLSLVAAML LLL ALVAL LLL SAARA	53
	Gaussia	MGVKVLFALICIAVAEA	54

[0072] Gaussia信号肽源自来自Gaussia princeps的荧光素酶并指导与此信号肽融合的治疗性蛋白的蛋白质合成和分泌增加。在一些实施例中,Gaussia信号肽具有与SEQ ID NO: 54至少90%相同的氨基酸序列。在一些实施例中,信号肽与SEQ ID NO:54相差5个或更少、4个或更少、3个或更少、2个或更少、或1个氨基酸。

[0073] 接头

[0074] 在一些实施例中,本文提供的组合物包括靶向肽与治疗性蛋白之间的接头。在一些实施例中,此类接头维持正确的间距并减轻vIGF2肽与治疗性蛋白之间的空间冲突。在一些实施例中,接头包括重复的甘氨酸残基、重复的甘氨酸-丝氨酸残基和其组合。在一些实施例中,接头由5-20个氨基酸、5-15个氨基酸、5-10个氨基酸、8-12个氨基酸或约5、6、7、8、9、10、11、12或13个氨基酸组成。合适的接头包含但不限于下表中提供的那些:

[0075]	接头序列	
	序列	SEQ ID NO:
	GGGSGGGG	55
	GGGS	56
	GGSGGGGS	57
	GGGSGGGS	58
	GGSGSGTS	59
	GGGSGGGGS	60

[0076] 在整个本说明书中,各种表达盒、载体基因组、载体和组合物被描述为含有hGAA780I编码序列或hGAA780I蛋白或融合蛋白。应当理解,除非另有说明,否则包含N末端截短、C末端截短和融合蛋白如本文所述的那些在内的任何经工程化的hGAA780I蛋白或其编码序列可以类似地工程化到表达盒、载体基因组、载体和组合物中。

[0077] 合适地,提供了包括本文所述的核酸序列的表达盒。

[0078] 表达盒

[0079] 如本文所使用的,“表达盒”是指包括编码功能性基因产物的核酸序列的核酸分子,所述核酸序列可操作地连接到指导所述功能性基因产物在靶细胞中的表达的调控序列(例如,hGAA780I融合蛋白编码序列启动子),并且可以包含用于其的其它调控序列。所需的调控序列以允许hGAA780I融合蛋白编码序列在靶细胞中转录、翻译和/或表达的方式与所述编码序列可操作地连接。

[0080] 在某些实施例中,所述表达盒可以包含非翻译区中的一个或多个miRNA靶序列。如本文所描述的,miRNA靶序列被设计成由存在于细胞中的miRNA特异性地识别,在所述细胞中不期望转基因表达和/或期望降低转基因表达水平。在某些实施例中,所述表达盒包含特

异性地降低背根神经节中hGAA780I融合蛋白表达的miRNA靶序列。在某些实施例中,miRNA靶序列定位于3'UTR、5'UTR和/或3'和5'UTR两者中。在某些实施例中,所述表达盒包括背根神经节(DRG)特异性miRNA靶序列的至少两个串联重复序列,其中所述至少两个串联重复序列包括可能相同或不同的至少一个第一miRNA靶序列和至少一个第二miRNA靶序列。在某些实施例中,所述至少两个drg特异性miRNA串联重复序列中的第一个的起点在距hGAA780I融合蛋白编码序列的3'端20个核苷酸内。在某些实施例中,所述至少两个DRG特异性miRNA串联重复序列中的第一个的起点距hGAA780I融合蛋白编码序列的3'端至少100个核苷酸。在某些实施例中,所述miRNA串联重复序列的长度为200到1200个核苷酸。在某些实施例中,相对于缺少miR靶序列的表达盒或载体基因组,包含miR靶不会改变治疗性转基因在一种或多种靶组织中的表达或功效。

[0081] 在某些实施例中,载体基因组或表达盒含有作为miR-183靶序列的至少一个miRNA靶序列。在某些实施例中,载体基因组或表达盒含有miR-183靶序列,其包含AGTGAATTCTACCACTGCCATA (SEQ ID NO:26),其中与miR-183种子序列互补的序列加下划线。在某些实施例中,载体基因组或表达盒含有与miR-183种子序列100%互补的序列的多于一个拷贝(例如,两个或三个拷贝)。在某些实施例中,miR-183靶序列的长度为约7个核苷酸到约28个核苷酸并且包含与miR-183种子序列至少100%互补的至少一个区。在某些实施例中,miR-183靶序列含有与SEQ ID NO:26具有部分互补性的序列,并且因此当与SEQ ID NO:26比对时,存在一个或多个错配。在某些实施例中,当与SEQ ID NO:26比对时,miR-183靶序列包括具有至少1个、2个、3个、4个、5个、6个、7个、8个、9个或10个错配的序列,其中错配可以不连续。在某些实施例中,miR-183靶序列包含具有100%互补性的区,所述区还包括miR-183靶序列的长度的至少30%。在某些实施例中,具有100%互补性的区包含与miR-183种子序列具有100%互补性的序列。在某些实施例中,miR-183靶序列的剩余部分与miR-183具有至少约80%到约99%的互补性。在某些实施例中,表达盒或载体基因组包含miR-183靶序列,所述miR-183靶序列包括截短的SEQ ID NO:26,即在SEQ ID NO:26的5'或3'端中的任一端或两端处缺少至少1个、2个、3个、4个、5个、6个、7个、8个、9个或10个核苷酸的序列。在某些实施例中,表达盒或载体基因组包括转基因和一个miR-183靶序列。在又其它实施例中,表达盒或载体基因组包括至少两个、三个或四个miR-183靶序列。

[0082] 在某些实施例中,载体基因组或表达盒含有作为miR-182靶序列的至少一个miRNA靶序列。在某些实施例中,载体基因组或表达盒含有miR-182靶序列,其包含AGTGTGAGTTCTACCATGCCCCAA (SEQ ID NO:27)。在某些实施例中,载体基因组或表达盒含有与miR-182种子序列100%互补的序列的多于一个拷贝(例如,两个或三个拷贝)。在某些实施例中,miR-182靶序列的长度为约7个核苷酸到约28个核苷酸并且包含与miR-182种子序列至少100%互补的至少一个区。在某些实施例中,miR-182靶序列含有与SEQ ID NO:27具有部分互补性的序列,并且因此当与SEQ ID NO:27比对时,存在一个或多个错配。在某些实施例中,当与SEQ ID NO:27比对时,miR-182靶序列包括具有至少1个、2个、3个、4个、5个、6个、7个、8个、9个或10个错配的序列,其中错配可以不连续。在某些实施例中,miR-182靶序列包含具有100%互补性的区,所述区还包括miR-182靶序列的长度的至少30%。在某些实施例中,具有100%互补性的区包含与miR-182种子序列具有100%互补性的序列。在某些实施例中,miR-182靶序列的剩余部分与miR-182具有至少约80%到约99%的互补性。在某些

实施例中,表达盒或载体基因组包含miR-182靶序列,所述miR-182靶序列包括截短的SEQ ID NO:27,即在SEQ ID NO:27的5'或3'端中的任一端或两端处缺少至少1个、2个、3个、4个、5个、6个、7个、8个、9个或10个核苷酸的序列。在某些实施例中,表达盒或载体基因组包括转基因和一个miR-182靶序列。在又其它实施例中,表达盒或载体基因组包括至少两个、三个或四个miR-182靶序列。

[0083] 本文所使用的术语“串联重复序列”是指存在两个或更多个连续的miRNA靶序列。这些miRNA靶序列可以是连续的,即一个接一个地直接定位,使得一个靶序列的3'端直接位于下一个靶序列的5'端的上游,没有中间序列,或者反之亦然。在另一个实施例中,miRNA靶序列中的两个或更多个miRNA靶序列由短间隔子序列隔开。

[0084] 如本文所使用的,“间隔子”是任何所选核酸序列,例如,长度为1个、2个、3个、4个、5个、6个、7个、8个、9个或10个核苷酸的定位在两个或更多个连续miRNA靶序列之间的核酸序列。在某些实施例中,间隔子的长度为1个到8个核苷酸、长度为2个到7个核苷酸、长度为3个到6个核苷酸、长度为四个核苷酸、4个到9个核苷酸、3个到7个核苷酸或更大的值。合适地,间隔子是非编码序列。在某些实施例中,间隔子可以具有四(4)个核苷酸。在某些实施例中,间隔子是GGAT。在某些实施例中,间隔子是六(6)个核苷酸。在某些实施例中,间隔子是CACGTG或GCATGC。

[0085] 在某些实施例中,串联重复序列含有相同的miRNA靶序列中的两个、三个、四个或更多个。在某些实施例中,串联重复序列含有至少两个不同的miRNA靶序列、至少三个不同的miRNA靶序列或至少四个不同的miRNA靶序列等。在某些实施例中,串联重复序列可以含有相同的miRNA靶序列中的两个或三个以及不同的第四miRNA靶序列。

[0086] 在某些实施例中,表达盒中可以有至少两组不同的串联重复序列。例如,3'UTR可以含有紧邻转基因下游的串联重复序列、UTR序列和两个或更多个更接近UTR的3'端的串联重复序列。在另一个实例中,5'UTR可以含有一个、两个或更多个miRNA靶序列。在另一个实例中,3'可以含有串联重复序列,并且5'UTR可以含有至少一个miRNA靶序列。

[0087] 在某些实施例中,表达盒含有两个、三个、四个或更多个串联重复序列,所述串联重复序列在转基因的终止密码子的约0个到20个核苷酸内开始。在其它实施例中,表达盒含有距转基因的终止密码子至少100个到约4000个核苷酸的miRNA串联重复序列。

[0088] 参见2019年12月20日提交的PCT/US19/67872,其通过引用并入本文并且要求于2018年12月21日提交的美国临时美国专利申请第62/783,956号的优先权,所述申请通过引用特此并入。

[0089] 如本文所使用的,“BiP-vIGF2.hGAACoV780I.4xmir183”是指表达盒(例如,如图11所描绘),所述表达盒含有用于具有经修饰的BiP-vIGF2信号序列的hGAA780I的经工程化的编码序列和miR183靶序列的四个串联重复序列,所述经工程化的编码序列受遍在的CAG启动子的控制。如本文提供的实例中所示,V780I突变和BiP-vIGF2修饰都有助于提高安全性和功效。在某些实施例中,BiP-vIGF2.hGAACoV780I.4xmir183包含编码SEQ ID NO:3的融合蛋白的序列或与所述序列至少95%相同的序列。在某些实施例中,BiP-vIGF2.hGAACoV780I.4xmir183包含SEQ ID NO:7的核酸序列或与所述核酸序列至少95%到99%相同的序列。在又另一个实施例中,本文提供了载体基因组,其中BiP-vIGF2.hGAACoV780I.4xmir183的侧接有5' ITR和3' ITR。在某些实施例中,载体基因组是SEQ

ID NO:30。在又另外的实施例中,提供了载体基因组,其包含与SEQ ID NO:30至少95%相同的序列并编码SEQ ID NO:6的融合蛋白。

[0090] 如本文所使用的,“可操作地连接的”序列包含与hGAA780I编码序列邻接的表达控制序列和以反式或在远处起作用以控制hGAA780I编码序列的表达控制序列。此类调控序列通常包含例如启动子、增强子、内含子、Kozak序列、聚腺苷酸化序列和TATA信号中的一种或多种。

[0091] 在某些实施例中,调控元件在受庞贝病影响的多种细胞和组织中指导表达,以便允许构建和递送适用于治疗多种靶细胞的单一表达盒。例如,可以选择在肝脏、骨骼肌、心脏和中枢神经系统细胞中的两种或更多种中表达的调控元件(例如,启动子)。例如,可以选择在中枢神经系统(例如,脑)细胞和骨骼肌中表达的调控元件(例如,启动子)。在其它实施例中,调控元件在CNS、骨骼肌和心脏中表达。在其它实施例中,表达盒允许编码的hGAA780I在肝脏、骨骼肌、心脏和中枢神经系统细胞中的全部中表达。在其它实施例中,可以选择调控元件以靶向特定组织并避免在某些细胞或组织中表达(例如,通过使用本文所述的drg-去靶向系统和/或通过选择组织特异性启动子)。在某些实施例中,将本文提供的优先靶向不同组织的不同表达盒施用于患者。

[0092] 调控序列包括启动子。可以选择合适的启动子,包含但不限于将在靶细胞中表达hGAAV780I蛋白的启动子。

[0093] 在某些实施例中,选择组成型启动子或诱导型/调控型启动子。组成型启动子的实例是鸡 β -肌动蛋白启动子。多种鸡 β -肌动蛋白启动子已单独描述,或与各种增强子元件(例如,CB7是具有巨细胞病毒增强子元件的鸡 β -肌动蛋白启动子;CAG启动子,其包含启动子、鸡 β 肌动蛋白的第一外显子和第一内含子以及兔 β 珠蛋白基因的剪接受体;CBh启动子,SJ Gray等人,《人类基因疗法(Hu Gene Ther)》,2011年9月;22(9):1143-1153)组合描述。在某些实施例中,可以选择可调控的启动子。参见例如W0 2011/126808B2,其通过引用并入本文。

[0094] 在某些实施例中,可以选择组织特异性启动子。组织特异性启动子的实例已熟知用于肝(白蛋白,Miyatake等人,(1997)《病毒学杂志(J.Virol.)》,71:5124-32;乙型肝炎病毒核心启动子,Sandig等人,(1996)《基因疗法(Gene Ther.)》,3:1002-9;甲胎蛋白(AFP),Arbuthnot等人,(1996)《人类基因疗法(Hum.Gene Ther.)》,7:1503-14);中枢神经系统,例如神经元(如神经元特异性烯醇酶(NSE)启动子,Andersen等人,(1993)《细胞分子神经生物学(Cell.Mol.Neurobiol.)》,13:503-15;神经丝轻链基因,Piccioli等人,(1991)《美国国家科学院院刊(Proc.Natl.Acad.Sci.USA)》,88:5611-5;以及神经元特异性vgf基因,Piccioli等人,(1995)《神经元(Neuron)》,15:373-84);心肌;骨骼肌;肺和其它组织。在另一个实施例中,合适的启动子可以包含但不限于以下:延伸因子1 α (EF1 α)启动子(参见例如Kim DW等人,使用人延伸因子1 α 启动子作为通用且高效的表达系统(Use of the human elongation factor 1alpha promoter as a versatile and efficient expression system).《基因(Gene)》.1990年7月16日;91(2):217-23);突触蛋白1启动子(参见例如Kügler S等人,人突触蛋白1基因启动子根据转导区域从成年大鼠脑中的腺病毒载体赋予高度神经元特异性的长期转基因表达(Human synapsin I gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in

the adult rat brain depending on the transduced area).《基因疗法(Gene Ther.)》2003年2月;10(4):337-47;神经元特异性烯醇酶(NSE)启动子(参见例如Kim J等人,富含胆固醇的脂质筏参与白介素6诱导的LNCaP前列腺癌细胞的神经内分泌分化(Involvement of cholesterol-rich lipid rafts in interleukin-6-induced neuroendocrine differentiation of LNCaP prostate cancer cells).《内分泌学(Endocrinology)》.2004年2月;145(2):613-9.电子版2003年10月16日);或CB6启动子(参见例如携带人存活运动神经元基因的腺相关病毒载体血清型9的大规模生产(Large-Scale Production of Adeno-Associated Viral Vector Serotype-9 Carrying the Human Survival Motor Neuron Gene),《分子生物技术(Mol Biotechnol.)》2016年1月;58(1):30-6.doi:10.1007/s12033-015-9899-5)。在利用组织特异性启动子的某些实施例中,可以选择涉及具有靶向不同细胞类型的组织特异性启动子的不同表达盒的共疗法。

[0095] 在一个实施例中,调控序列进一步包括增强子。在一个实施例中,调控序列包括一个增强子。在另一个实施例中,调控序列含有两个或更多个表达增强子。这些增强子可以相同或者可以不同。例如,增强子可以包含 α mic/bik增强子或CMV增强子。此增强子可以存在于彼此相邻定位的两个拷贝中。可替代地,增强子的双拷贝可以被一个或多个序列隔开。

[0096] 在一个实施例中,调控序列进一步包括内含子。在另外的实施例中,内含子是鸡 β -肌动蛋白内含子。其它合适的内含子包含本领域已知的内含子,所述内含子可以是人 β -球蛋白内含子和/或可商购获得的 Promega® 内含子以及WO 2011/126808中描述的内含子。

[0097] 在一个实施例中,调控序列进一步包括聚腺苷酸化信号(polyA)。在另外的实施例中,polyA是兔珠蛋白polyA。参见例如WO 2014/151341。可替代地,另一种polyA(例如,人生长激素(hGH)聚腺苷酸化序列、SV40 polyA或合成polyA)可以包含在表达盒中。

[0098] 应当理解,本文所描述的表达盒中的组合物旨在应用于本说明书中描述的其它组合物、方案、方面、实施例和方法。

[0099] 表达盒可以通过任何合适的递送系统递送。合适的非病毒递送系统是本领域已知的(参见例如Ramamoorth和Narvekar.《临床诊断研究(J Clin Diagn Res.)》2015年1月;9(1):GE01-GE06,其通过引用并入本文中)并且可以由本领域技术人员容易地选择并且可以包含例如裸DNA、裸RNA、树状聚合物、PLGA、聚甲基丙烯酸酯、无机颗粒、脂质颗粒(例如,脂质纳米颗粒或LNP)或基于壳聚糖的调配物。

[0100] 在一个实施例中,载体是非病毒质粒,所述非病毒质粒包括其描述的表达盒,例如,“裸DNA”、“裸质粒DNA”、RNA和mRNA;与各种组合物和纳米颗粒偶联,包含例如胶束、脂质体、阳离子脂质-核酸组合物、多聚糖组合物和其它聚合物、基于脂质和/或胆固醇的核酸缀合物以及如本文所描述的其它构建体。参见例如X.Su等人,《分子制药学(Mol.Pharmaceutics)》,2011,8(3),第774-787页;网络出版物:2011年3月21日;WO2013/182683、WO 2010/053572和WO 2012/170930,所有所述文献都通过引用并入本文中。

[0101] 在某些实施例中,本文提供了具有编码如本文所述的hGAA780I变体、融合蛋白或截短的蛋白的序列的核酸分子。在一个期望的实施例中,hGAA780I由SEQ ID NO:4的经工程化的序列或与所述经工程化的序列至少95%相同的编码hGAA780I变体的序列编码。在某些实施例中,SEQ ID NO:4被修饰成使得在位置780I处编码Ile的密码子是ATT或ATC。在某些实施例中,包括SEQ ID NO:4的经工程化的序列或其片段的核酸用于表达融合蛋白或截短

的hGAA780I。尽管不太期望,但在某些实施例中,hGAA780I由SEQ ID NO:5编码。在某些实施例中,核酸编码具有SEQ ID NO:6的氨基酸序列或与所述氨基酸序列至少95%相同的序列的融合蛋白。在某些实施例中,提供具有SEQ ID NO:7的序列或与所述序列至少95%相同的序列的核酸。在某些实施例中,核酸分子是质粒。

[0102] 载体

[0103] 如本文所使用的,“载体”是包括核酸序列的生物或化学部分,所述生物或化学部分可以被引入到合适的靶细胞中以复制或表达所述核酸序列。载体的实例包含但不限于重组病毒、质粒、脂质体、聚合物体、复合物、树状聚合物、细胞穿透肽(CPP)缀合物、磁性颗粒或纳米颗粒。在一个实施例中,载体是具有对功能性基因产物进行编码的外源性或异源性经工程化的核酸的核酸分子,然后将所述核酸分子引入到适当的靶细胞中。此类载体优选地具有一个或多个复制起点以及重组DNA可以插入到的一个或多个位点。载体通常具有装置,通过所述装置,可以从没有载体的细胞中选择具有载体的细胞,例如,所述载体对耐药基因进行编码。常见的载体包含质粒、病毒基因组和“人工染色体”。载体的产生、生产、表征或定量的常规方法对于本领域技术人员是可用的。

[0104] 在某些实施例中,本文所描述的载体是“复制缺陷型病毒”或“病毒载体”,其是指其中含有对功能性hGAA780I融合蛋白进行编码的核酸序列的表达盒包装在病毒衣壳或包膜中的合成或人工病毒颗粒,其中也包装在病毒衣壳或包膜内的任何病毒基因组序列均是复制缺陷型的;即,其不能产生子代病毒粒子,但保留了感染靶细胞的能力。在一个实施例中,病毒载体的基因组不包含对复制所需的酶进行编码的基因(基因组可以被工程化成“无肠的(gutless)”-仅含有核酸序列编码,其侧接扩增和包装人工基因组所需的信号),但是这些基因可以在产生期间供应。因此,这被认为可以安全地用于基因疗法,因为除非存在复制所需的病毒酶,否则不会发生通过子代病毒粒子进行的复制和感染。

[0105] 如本文所使用的,重组病毒载体是靶向期望细胞的任何合适的病毒载体。因此,重组病毒载体优选地靶向受庞贝病影响的细胞和组织中的一种或多种细胞和组织,包含中枢神经系统(例如,脑)、骨骼肌、心脏和/或肝脏。在某些实施例中,病毒载体至少靶向中枢神经系统(例如,脑)细胞、肺、心肌细胞或骨骼肌。在其它实施例中,病毒载体靶向CNS(例如,脑)、骨骼肌和/或心脏。在其它实施例中,病毒载体靶向肝脏、骨骼肌、心脏和中枢神经系统细胞中的全部。实例提供了说明性重组腺相关病毒(rAAV)。然而,其它合适的病毒载体可以包含例如重组腺病毒、重组细小病毒如重组博卡病毒、杂交AAV/博卡病毒、重组单纯疱疹病毒、重组逆转录病毒或重组慢病毒。在优选实施例中,这些重组病毒不能复制。

[0106] 如本文所使用的,术语“宿主细胞”可以指其中产生载体(例如,重组AAV)的包装细胞系。宿主细胞可以是原核或真核细胞(例如,人、昆虫或酵母),所述原核或真核细胞含有通过任何方式(例如,电穿孔、磷酸钙沉淀、显微注射、转化、病毒感染、转染、脂质体递送、膜融合技术、高速DNA包被的团粒、病毒感染和原生质体融合)引入到细胞中的外源性或异源性DNA。宿主细胞的实例可以包含但不限于分离的细胞、细胞培养物、大肠杆菌(*Escherichia coli*)细胞、酵母细胞、人细胞、非人细胞、哺乳动物细胞、非哺乳动物细胞、昆虫细胞、HEK-293细胞、肝细胞、肾细胞、中枢神经系统的细胞、神经元、神经胶质细胞或干细胞。

[0107] 在某些实施例中,宿主细胞含有用于产生hGAA780I的表达盒,使得在体外以足够

的量产生用于分离或纯化的蛋白质。在某些实施例中，宿主细胞含有编码hGAAV780I或其片段的表达盒。如本文所提供的，hGAA780I可以包含在作为治疗剂(即酶替代疗法)施用于受试者的药物组合物中。

[0108] 如本文所使用的，术语“靶细胞”是指其中期望功能性基因产物的表达的任何靶细胞。

[0109] 如本文所使用的，“载体基因组”是指包装在病毒载体内部的核酸序列。在一个实施例中，“载体基因组”从5'到3'至少含有载体特异性序列、核酸序列以及载体特异性序列，所述核酸序列对功能性基因产物(例如，hGAAV780I、融合蛋白hGAAV780I或另一种蛋白质)进行编码，所述核酸序列与指导所述功能性基因产物在靶细胞中的表达的调控序列、载体特异性序列和任选的一个或多个非翻译区中的miRNA靶序列可操作地连接。载体特异性序列可以是特异性地将载体基因组包装到病毒载体衣壳或包膜蛋白中的末端重复序列。例如，AAV反向末端重复序列用于包装到AAV和某些其它细小病毒衣壳中。在期望包装到慢病毒载体中的情况下，可以利用慢病毒长末端重复序列。类似地，可以选择其它末端重复序列(例如，逆转录病毒长末端重复序列)等。

[0110] 应当理解，本文所描述的载体中的组合物旨在应用于本说明书中描述的其它组合物、方案、方面、实施例和方法。

[0111] 腺相关病毒(AAV)

[0112] 在一方面，本文提供了包括AAV衣壳和包装在所述AAV衣壳中的载体基因组的重组AAV(rAAV)，所述载体基因组编码如本文所述的hGAAV780I融合蛋白(酶)。在某些实施例中，所选AAV衣壳靶向肝脏、肌肉、肾脏、心脏和/或中枢神经系统细胞类型中的两种或更多种细胞。在某些实施例中，期望在肝脏、骨骼肌、心脏、肾脏和/或至少一种中枢神经系统细胞类型中的至少两种或更多种中表达hGAA780I融合蛋白。因此，在一个实施例中，所选AAV衣壳靶向心脏组织。在某些实施例中，被选择成靶向心脏组织的AAV衣壳选自AAV1、6、8和9(参见例如Katz等人，《人类基因疗法临床开发(Hum Gene Ther Clin Dev.)》2017年9月1日;28(3):157-164)。在又其它实施例中，所选AAV衣壳靶向肾脏细胞。在一个实施例中，用于靶向肾脏细胞的衣壳选自AAV1、2、6、8、9和Anc80(参见例如Ikeda Y等人，《美国肾脏病学会杂志(J Am Soc Nephrol.)》2018年9月;29(9):2287-2297和Ascio等人《生物化学与生物物理学研究通讯(Biochem Biophys Res Commun.)》2018年2月26日;497(1):19-24)。在某些实施例中，AAV衣壳是天然的或经工程化的进化枝F衣壳。在某些实施例中，衣壳是AAV9衣壳或AAVhu68衣壳。

[0113] 在一个实施例中，载体基因组包括AAV 5'反向末端重复序列(ITR)、如本文所描述的表达盒和AAV 3' ITR。在一个实施例中，载体基因组是指包装在形成rAAV载体的rAAV衣壳内部的核酸序列。此类核酸序列含有侧接表达盒的AAV反向末端重复序列(ITR)。在一个实施例中，用于包装到AAV或博卡病毒衣壳中的“载体基因组”从5'到3'至少含有AAV 5' ITR、编码如本文所述的功能性hGAA780I融合蛋白的核酸序列和AAV3' ITR，所述核酸序列与指导所述融合蛋白在靶细胞中的表达的调控序列可操作地连接。在某些实施例中，ITR来自AAV2，并且衣壳来自不同的AAV。可替代地，可以使用其它ITR。在某些实施例中，载体基因组进一步包括非翻译区中的miRNA靶序列，所述miRNA靶序列被设计成由细胞中的miRNA序列特异性地识别，在所述细胞中不期望转基因表达和/或期望降低转基因表达水平。

[0114] ITR是在载体产生期间负责基因组复制和包装的基因元件,并且是产生rAAV所需的唯一病毒顺式元件。在一个实施例中,ITR来自与供应衣壳的AAV不同的AAV。在优选实施例中,来自AAV2的ITR序列或其缺失版本(Δ ITR)可以为了方便而使用和加速调控批准。然而,可以选择来自其它AAV来源的ITR。在ITR的来源来自AAV2并且AAV衣壳来自另一种AAV来源的情况下,所得载体可以被称为假型。通常,AAV载体基因组包括AAV 5' ITR、hGAA780I编码序列和任何调控序列以及AAV 3' ITR。然而,这些元件的其它构型可以是合适的。已经描述了被称为 Δ ITR的5' ITR的缩短版本,其中缺失了D序列和末端解析位点(trs)。在其它实施例中,使用了全长AAV 5'和3' ITR。

[0115] 如本文所使用的,术语“AAV”是指天然存在的腺相关病毒、本领域技术人员可获得的和/或根据本文所描述的组合物和方法可获得的腺相关病毒以及人工AAV。腺相关病毒(AAV)病毒载体是具有AAV蛋白衣壳的AAV核酸酶(例如,DNase)抗性颗粒,其中包装有侧接AAV反向末端重复序列(ITR)的用于递送到靶细胞的表达盒。核酸酶抗性重组AAV(rAAV)表示AAV衣壳已经完全组装并保护这些包装的载体基因组序列在被设计成去除产生过程中可能存在的污染性核酸的核酸酶温育步骤期间免于降解(消化)。在许多情况下,本文所描述的rAAV是DNase抗性的。

[0116] AAV衣壳由60个衣壳(cap)蛋白亚基VP1、VP2和VP3构成,其以二十面体对称布置,比率为大约1:1:10到1:1:20,具体取决于所选AAV。可以选择各种AAV作为如上文所鉴定的AAV病毒载体的衣壳的来源。参见例如美国公开专利申请第2007-0036760-A1号;美国公开专利申请第2009-0197338-A1号;EP 1310571。还参见WO 2003/042397(AAV7和其它猿猴AAV)、美国专利7790449和美国专利7282199(AAV8)、WO 2005/033321和US 7,906,111(AAV9)和WO 2006/110689以及WO2003/042397(rh.10)。这些文档还描述了可以选择用于产生AAV的其它AAV,并且通过引用并入。在从人或非人灵长类动物(NHP)中分离或工程化以及良好表征的AAV中,人AAV2是第一个被开发为基因转移载体的AAV;其已被广泛用于不同靶组织和动物模型中的高效基因转移实验。除非另有说明,否则本文所描述的AAV衣壳、ITR和其它所选AAV组分可以容易地选自任何AAV,包括但不限于通常鉴定为AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV8bp、AAV7M8和AAVAnc80的AAV。参见例如WO 2005/033321,其通过引用并入本文。在一个实施例中,AAV衣壳是AAV9衣壳或其变体。在某些实施例中,衣壳蛋白由rAAV载体名称中的术语“AAV”之后的数字或数字和字母的组合指定。

[0117] ITR或其它AAV组分可以使用本领域技术人员可用的技术从AAV中容易地分离出来或进行工程化。此类AAV可以从学术、商业或公共来源分离、工程化或获得(例如,维吉尼亚州马纳萨斯的美国典型培养物保藏中心(American Type Culture Collection, Manassas, VA))。可替代地,AAV序列可以通过合成或其它适合的方式通过参考公开的序列(如在文献中或在如GenBank、PubMed等数据库中可获得的公开的序列)进行工程化。AAV病毒可以通过常规的分子生物学技术进行工程化,从而可以优化这些颗粒以用于核酸序列的细胞特异性递送、用于最小化免疫原性、用于调节稳定性和颗粒寿命、用于高效降解、用于准确递送到细胞核等。

[0118] 如本文所使用的,可互换使用的术语“rAAV”和“人工AAV”意指但不限于包括衣壳蛋白和包装在其中的载体基因组的AAV,其中载体基因组包括与AAV异源的核酸。在一个实施例中,衣壳蛋白是非天然存在的衣壳。此类人工衣壳可以通过任何合适的技术使用所选

AAV序列(例如, vp1衣壳蛋白的片段)与异源序列的组合产生,所述异源序列可以从不同的所选AAV、同一AAV的非连续部分、从非AAV病毒来源或从非病毒来源获得。人工AAV可以是但不限于假型AAV、嵌合AAV衣壳、重组AAV衣壳或“人源化”AAV衣壳。假型载体可用于某些实施例中,其中一个AAV的衣壳被异源衣壳蛋白替代。在一个实施例中, AAV2/5和AAV2/8是示例性假型载体。所选基因元件可以通过任何合适的方法递送, 包含转染、电穿孔、脂质体递送、膜融合技术、高速DNA包被的团粒、病毒感染和原生质体融合。用于制备此类构建体的方法对核酸操纵技术人员而言是已知的并且包含基因工程、重组工程以及合成技术。参见例如Green和Sambrook,《分子克隆:实验室手册(Molecular Cloning:A Laboratory Manual)》, 纽约市冷泉港的冷泉港实验室出版社(Cold Spring Harbor Press, Cold Spring Harbor, NY) (2012)。

[0119] 在某些实施例中, AAV衣壳选自天然的和经工程化的进化枝F腺相关病毒。在下面的实例中, 进化枝F腺相关病毒是AAVhu68。参见WO 2018/160582, 其通过引用整体并入本文。然而, 在其它实施例中, AAV衣壳选自不同的进化枝, 例如进化枝A、B、C、D或E, 或选自这些进化枝中的任何进化枝之外的AAV源。

[0120] 如本文所使用的, 与AAV的组有关的术语“进化枝”是指如基于对AAV vp1氨基酸序列进行的比对, 通过(至少1000个复制品的)至少75%的自举值(bootstrap value)和不大于0.05的泊松校正距离测量结果(Poisson correction distance measurement)使用邻接算法(Neighbor-Joining algorithm)确定的一组在系统发育上彼此相关的AAV。在文献中已经描述了邻接算法。参见例如M. Nei和S. Kumar,《分子进化和系统发育学(Molecular Evolution and Phylogenetics)》(牛津大学出版社(Oxford University Press), 纽约(2000))。提供了可以用于实施此算法的可用计算机程序。例如, MEGA v2.1程序实施了经修改的Nei-Gojobori方法。使用这些技术和计算机程序以及AAV vp1衣壳蛋白的序列, 本领域技术人员可以容易地确定所选AAV是含在本文所鉴定的进化枝之一中, 还是在这些进化枝之外的另一个进化枝中。参见例如G. Gao等人,《病毒学杂志(JVirol)》, 2004年6月; 7810: 6381-6388, 所述文献鉴定进化枝A、B、C、D、E和F, 并提供新颖AAV的核酸序列, GenBank登录号AY530553到AY530629。还参见WO2005/033321。

[0121] 如本文所使用的, “AAV9衣壳”是指具有以下的氨基酸序列的AAV9: (a) GenBank登录: AAS99264, 其通过引用并入本文中, 以及AAV vp1衣壳蛋白; 和/或 (b) 由GenBank登录: AY530579.1: (nt 1..2211) 的核苷酸序列编码的氨基酸序列。本发明涵盖来自此经过编码的序列的一些变化, 其可以包含与GenBank登录: AAS99264和US7906111 (也为WO 2005/033321) 中的参考氨基酸序列具有约99%的同一性的序列 (即与参考序列的变化小于约1%)。此类AAV可以包含例如天然分离物 (例如, hu31或hu32) 或具有氨基酸取代、缺失或添加的AAV9的变体, 例如包含但不限于选自从与AAV9衣壳比对的任何其它AAV衣壳中的对应位置“招募”的替代性残基; 例如, 如US 9,102,949、US 8,927,514、US2015/349911、WO 2016/049230A1、US 9,623,120和US 9,585,971中所述。然而, 在其它实施例中, 可以选择与上文所引用的序列具有至少约95%的同一性的AAV9或AAV9衣壳的其它变体。参见例如US 2015/0079038。已经描述了产生衣壳的方法、其编码序列以及产生rAAV病毒载体的方法。参见例如Gao等人,《美国国家科学院院刊(Proc.Natl.Acad.Sci.U.S.A.)》100(10):6081-6086 (2003) 和US2013/0045186A1。

[0122] 在某些实施例中, AAVhu68衣壳如标题为“新型腺相关病毒(AAV)进化枝F载体及其用途(Novel Adeno-associated virus(AAV) Clade F Vector and Uses Therefor)”的W02018/160582中所述, 所述文献在此通过引用并入。在某些实施例中, AAVhu68衣壳蛋白包括: 通过从对SEQ ID NO:2的1到736的预测的氨基酸序列进行编码的核酸序列的表达产生的AAVhu68 vp1蛋白、从SEQ ID NO:2产生的vp1蛋白或从与SEQ ID NO:1至少70%相同的对SEQ ID NO:2的1到736的预测的氨基酸序列进行编码的核酸序列产生的vp1蛋白; 通过从对SEQ ID NO:2的至少约氨基酸138到736的预测的氨基酸序列进行编码的核酸序列的表达产生的AAVhu68 vp2蛋白、从包括SEQ ID NO:1的至少核苷酸412到2211的序列产生的vp2蛋白或从与SEQ ID NO:1的至少核苷酸412到2211至少70%相同的对SEQ ID NO:2的至少约氨基酸138到736的预测的氨基酸序列进行编码的核酸序列产生的vp2蛋白, 和/或通过从对SEQ ID NO:2的至少约氨基酸203到736的预测的氨基酸序列进行编码的核酸序列的表达产生的AAVhu68 vp3蛋白、从包括SEQ ID NO:1的至少核苷酸607到2211的序列产生的vp3蛋白或从与SEQ ID NO:1的至少核苷酸607到2211至少70%相同的对SEQ ID NO:2的至少约氨基酸203到736的预测的氨基酸序列进行编码的核酸序列产生的vp3蛋白。

[0123] AAVhu68 vp1、vp2和vp3蛋白通常表示为由对SEQ ID NO:2的全长vp1氨基酸序列(氨基酸1到736)进行编码的相同核酸序列编码的替代剪接变体。任选地, 单独使用vp1编码序列来表达vp1、vp2和vp3蛋白。可替代地, 此序列可以与以下中的一个或多个共表达: 对SEQ ID NO:2的AAVhu68 vp3氨基酸序列(约aa 203到736)进行编码的核酸序列, 所述AAVhu68 vp3氨基酸序列没有vp1独特区(约aa 1到约aa 137)和/或vp2独特区(约aa 1到约aa 202)或与其互补的链, 即对应的mRNA(SEQ ID NO:1的约nt 607到约nt 2211); 或与SEQ ID NO:1至少70%到至少99%(例如, 至少85%、至少90%、至少95%、至少97%、至少98%或至少99%)相同的对SEQ ID NO:2的aa203到736进行编码的序列。另外地或可替代地, vp1编码和/或vp2编码序列可以与以下共表达: 对SEQ ID NO:2的AAVhu68 vp2氨基酸序列(约aa 138到736)进行编码的核酸序列, 所述AAVhu68 vp2氨基酸序列没有vp1独特区(约aa 1到约aa 137)或与其互补的链, 即对应的mRNA(例如, SEQ ID NO:1的nt 412到2211); 或与SEQ ID NO:1的nt 412到2211至少70%到至少99%(例如, 至少85%、至少90%、至少95%、至少97%、至少98%或至少99%)相同的对SEQ ID NO:2的约aa 138到736进行编码的序列。

[0124] 如本文所描述的, rAAVhu68具有在产生系统中产生的rAAVhu68衣壳, 所述产生系统表达来自以下的衣壳: 对SEQ ID NO:2的vp1氨基酸序列进行编码的AAVhu68核酸以及任选地例如对不含vp1和/或vp2独特区的vp3蛋白进行编码的另外的核酸序列。使用单个核酸序列vp1从生产中产生的rAAVhu68产生vp1蛋白、vp2蛋白和vp3蛋白的异质群体。更具体地, AAVhu68衣壳含有具有来自SEQ ID NO:2中的预测的氨基酸残基的修饰的vp1蛋白内、vp2蛋白内和vp3蛋白内的亚群体。这些亚群体至少包含脱酰胺化的天冬酰胺(N或Asn)残基。例如, 天冬酰胺-甘氨酸对中的天冬酰胺是高度脱酰胺化的。

[0125] 在一个实施例中, AAVhu68 vp1核酸序列具有SEQ ID NO:1的序列或与其互补的链, 例如, 对应的mRNA。在某些实施例中, vp2和/或vp3蛋白可以另外地或可替代地由不同于vp1的核酸序列表达, 例如以改变所选表达系统中的vp蛋白的比率。在某些实施例中, 还提供了对SEQ ID NO:2的AAVhu68 vp3氨基酸序列(约aa 203到736)进行编码的核酸序列, 所述AAVhu68 vp3氨基酸序列没有vp1-独特区(约aa 1到约aa137)和/或vp2独特区(约aa 1到

约aa 202)或与其互补的链,即对应的mRNA(SEQ ID NO:2的约nt 607到约nt 2211)。在某些实施例中,还提供了对SEQ ID NO:2的AAVhu68vp2氨基酸序列(约aa 138到736)进行编码的核酸序列,所述AAVhu68 vp2氨基酸序列没有vp1-独特区(约aa 1到约137)或与其互补的链,即对应的mRNA(SEQ ID NO:1的nt 412到nt 2211)。

[0126] 然而,可以选择对SEQ ID NO:2的氨基酸序列进行编码的其它核酸序列用于产生rAAVhu68衣壳。在某些实施例中,核酸序列具有SEQ ID NO:1的核酸序列或对SEQ ID NO:2进行编码的与SEQ ID NO:1至少70%到99%相同、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%、至少99%相同的序列。在某些实施例中,核酸序列具有SEQ ID NO:1的核酸序列或对SEQ ID NO:2的vp2衣壳蛋白(约aa 138到736)进行编码的与SEQ ID NO:1的约nt 412到约nt 2211至少70%到99%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%或至少99%相同的序列。在某些实施例中,核酸序列具有SEQ ID NO:1的约nt 607到约nt 2211的核酸序列或对SEQ ID NO:1的vp3衣壳蛋白(约aa 203到736)进行编码的与SEQ ID NO:1的nt 412到约nt 2211至少70%到99%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%或至少99%相同的序列。

[0127] 设计对此AAVhu68衣壳进行编码的核酸序列在本领域的技术范围内,包含DNA(基因组或cDNA)或RNA(例如,mRNA)。在某些实施例中,SEQ ID NO:2中提供了对AAVhu68 vp1衣壳蛋白进行编码的核酸序列。在某些实施例中,AAVhu68衣壳是使用SEQ ID NO:1的核酸序列或具有至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%或至少99%的如本文所描述的修饰(例如,脱酰胺化的氨基酸)的对SEQ ID NO:2的vp1氨基酸序列进行编码的序列产生的。在某些实施例中,vp1氨基酸序列在SEQ ID NO:2中重现。

[0128] 在某些实施例中,可以选择具有减少的衣壳脱酰胺的AAV衣壳。参见例如PCT/US19/19804和PCT/US18/19861,两者均于2019年2月27日提交并通过引用整体并入。

[0129] 如本文所使用的,当用于指vp衣壳蛋白时,术语“异质”或其任何语法变型是指由不相同的元件组成的群体,例如具有带有不同的经过修饰的氨基酸序列的vp1、vp2或vp3单体(蛋白质)。SEQ ID NO:2提供了AAVhu68 vp1蛋白的经过编码的氨基酸序列。与vp1、vp2和vp3蛋白(可替代地被称为同种型)结合使用的术语“异质的”是指衣壳内的vp1、vp2和vp3蛋白的氨基酸序列中的差异。AAV衣壳含有具有来自预测的氨基酸残基的修饰的vp1蛋白内、vp2蛋白内和vp3蛋白内的亚群体。这些亚群体至少包含某些脱酰胺化的天冬酰胺(N或Asn)残基。例如,某些亚群体包括天冬酰胺-甘氨酸对中的至少一个、两个、三个或四个高度脱酰胺化的天冬酰胺(N)位置,并且任选地进一步包括其它脱酰胺化的氨基酸,其中脱酰胺化引起氨基酸变化和其它任选的修饰。

[0130] 如本文所使用的,除非另有说明,否则vp蛋白的“亚群体”是指一组vp蛋白,所述一组vp蛋白具有至少一个限定的共同特性,并且由至少一个组成员到少于参考组的所有成员组成。例如,除非另有说明,否则vp1蛋白的“亚群体”是至少一种(1) vp1蛋白,并且少于组装的AAV衣壳中的所有vp1蛋白。除非另有说明,否则vp3蛋白的“亚群体”可以是少于组装的AAV衣壳中的所有vp3蛋白的一种(1) vp3蛋白。例如,vp1蛋白可以是vp蛋白的亚群体;vp2蛋白可以是vp蛋白的单独的亚群体,并且vp3是组装的AAV衣壳中的vp蛋白的又另外的亚群体。在另一个实例中,vp1、vp2和vp3蛋白可以含有具有不同修饰的亚群体,例如,至少一种、两种、三种或四种高度脱酰胺化的天冬酰胺,例如在天冬酰胺-甘氨酸对处。

[0131] 除非另有说明,否则高度脱酰胺化的是指与在参考氨基酸位置处的预测的氨基酸序列相比,在参考的氨基酸位置处被至少45%脱酰胺化的、至少50%脱酰胺化的、至少60%脱酰胺化的、至少65%脱酰胺化的、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%、至少99%或至多约100%脱酰胺化的(例如,在基于SEQ ID NO:2的编号[AAVhu68]的氨基酸57处的天冬酰胺的至少80%可以基于总vp1蛋白脱酰胺化、可以基于总vp1、vp2和vp3蛋白脱酰胺化)。此类百分比可以使用2D凝胶、质谱技术或其它合适的技术来确定。

[0132] 因此,rAAV包含vp1、vp2和/或vp3蛋白的rAAV衣壳内具有脱酰胺化的氨基酸的亚群体,至少包含至少一个包括至少一种高度脱酰胺化的天冬酰胺的亚群体。另外,其它修饰可以包含异构化,特别是在所选天冬氨酸(D或Asp)残基位置处。在仍其它实施例中,修饰可以包含在Asp位置处的酰胺化。

[0133] 在某些实施例中,AAV衣壳含有具有至少4个到至少约25个脱酰胺化的氨基酸残基位置的vp1、vp2和vp3的亚群体,与vp蛋白的经过编码的氨基酸序列相比,所述氨基酸残基位置中的至少1%到10%被脱酰胺化。这些中的大多数可以是N残基。然而,Q残基也可以被脱酰胺化。

[0134] 在某些实施例中,rAAV具有含有vp1、vp2和vp3蛋白的AAV衣壳,所述蛋白具有包括在实例1中提供的表中列出的位置处的两个、三个、四个或更多个脱酰胺化的残基的组合的亚群体,并通过引用并入本文中。rAAV中的脱酰胺化可以使用2D凝胶电泳和/或质谱和/或蛋白质建模技术来确定。在线色谱可以使用Acclaim PepMap柱和与QExactive HF和NanoFlex源(赛默飞世尔科技公司(Thermo Fisher Scientific))耦合的Thermo Ultimate 3000RSLC系统(赛默飞世尔科技公司)执行。MS数据是使用用于QExactive HF的数据依赖性前20种方法获取的,所述方法从调查扫描(200-2000m/z)中动态选择最丰富的尚未测序的前体离子。测序通过高能碰撞解离片段进行,其中通过预测性自动增益控制确定的靶值为1e5离子,并且以4m/z的窗口进行前体分离。在m/z 200下以120,000的分辨率获取调查扫描。HCD光谱的分辨率可以在m/z 200下设置为30,000,其中最大离子注入时间为50毫秒,并且归一化碰撞能量为30。S-透镜RF水平可以设置为50,以使消化肽所占据的m/z区达到最佳传输。可以从片段选择中排除具有单个、未分配或六个和更高电荷状态的前体离子。BioPharma Finder 1.0软件(赛默飞世尔科技公司)可以用于分析所获取的数据。对于肽作图,使用单进入蛋白FASTA数据库进行搜索,其中脛基甲基化设置为固定的修饰;并将氧化、脱酰胺化和磷酸化设置为可变修饰、10ppm质量准确度、高蛋白酶特异性和置信水平为0.8的MS/MS光谱。合适的蛋白酶的实例可以包含例如胰蛋白酶或胰凝乳蛋白酶。脱酰胺化的肽的质谱鉴定相对简单,因为脱酰胺化向完整分子的质量添加了+0.984Da(-OH基团与-NH₂基团之间的质量差)。特定肽的脱酰胺化百分比通过将脱酰胺化的肽的质量面积除以脱酰胺化的和天然的肽的面积之和来确定。考虑到可能的脱酰胺化位点的数量,在不同位点处被脱酰胺化的同量异位物种可以在单个峰中共迁移。因此,源自具有多个潜在的脱酰胺化位点的肽的片段离子可以用于定位或区分多个脱酰胺化位点。在这些情况下,观察到的同位素图案内的相对强度可以用于特异性确定不同的脱酰胺化的肽异构体的相对丰度。此方法假设所有异构物种的片段化效率是相同的,并且在脱酰胺化位点是独立的。本领域的技术人员将理解的是,可以使用这些说明性方法的多种变型。例如,合适的质谱仪可以包含例

如四极杆飞行时间质谱仪 (QTOF), 如Waters Xevo或Agilent 6530, 或Orbitrap仪器, 如Orbitrap Fusion或Orbitrap Velos (赛默飞世尔科技公司)。合适的液相色谱系统包含例如来自沃特世 (Waters) 的Acquity UPLC系统或Agilent系统 (1100或1200系列)。合适的数据分析软件可以包含例如MassLynx (沃特世)、Pinpoint和Petfinder (赛默飞世尔科技公司)、Mascot (矩阵科学公司 (Matrix Science))、Peaks DB (生物信息学解决方案公司 (Bioinformatics Solutions))。可以在例如X.Jin等人, 于2017年6月16日在线公开的《人类基因疗法方法 (Hu Gene Therapy Methods) 》, 第28卷, 第5期, 第255-267页中描述仍其它技术。

[0135] 除脱酰胺化之外, 可能发生不会导致一个氨基酸转化为不同的氨基酸残基的其它修饰。这种修饰可以包含乙酰化残基、异构化、磷酸化或氧化。

[0136] 脱酰胺化的调节: 在某些实施例中, AAV被修饰成改变天冬酰胺-甘氨酸对中的甘氨酸, 以降低脱酰胺化。在其它实施例中, 将天冬酰胺改变为不同的氨基酸, 例如以较慢速率进行脱酰胺化的谷氨酰胺; 或改变为缺乏酰胺基的氨基酸 (例如, 含有酰胺基的谷氨酰胺和天冬酰胺); 和/或改变为缺乏胺基的氨基酸 (例如, 含有胺基的赖氨酸、精氨酸和组氨酸)。如本文所使用的, 缺乏酰胺或胺侧基的氨基酸是指例如甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸、胱氨酸、苯丙氨酸、酪氨酸或色氨酸和/或脯氨酸。如所描述的修饰可以在经过编码的AAV氨基酸序列中存在的的天冬酰胺-甘氨酸对中的一个、两个或三个天冬酰胺-甘氨酸对中。在某些实施例中, 在所有四个天冬酰胺-甘氨酸对中没有进行此类修饰。因此, 一种用于降低具有较低脱酰胺化速率的AAV和/或经工程化的AAV变体的脱酰胺化的方法。另外地或可替代地, 可以将一种或多种其它酰胺氨基酸改变为非酰胺氨基酸以降低AAV的脱酰胺化。在某些实施例中, 本文所描述的突变AAV衣壳含有天冬酰胺-甘氨酸对中的突变, 使得甘氨酸变为丙氨酸或丝氨酸。突变AAV衣壳可以含有一个、两个或三个突变体, 其中参考AAV天然地含有四个NG对。在某些实施例中, AAV衣壳可以含有一个、两个、三个或四个此类突变体, 其中参考AAV天然地含有五个NG对。在某些实施例中, 突变AAV衣壳含有NG对中的仅单个突变。在某些实施例中, 突变AAV衣壳含有两个不同的NG对中的突变。在某些实施例中, 突变AAV衣壳含有定位在AAV衣壳中的结构上分开的位置中的两个不同的NG对中的突变。在某些实施例中, 突变不在VP1-独特区中。在某些实施例中, 突变之一不在VP1-独特区中。任选地, 突变AAV衣壳不含有NG对中的修饰, 但是含有突变以最小化或消除定位在NG对的外部的一个或多个天冬酰胺或谷氨酰胺中的脱酰胺化。在AAVhu68衣壳蛋白中, 4个残基 (N57、N329、N452、N512) 跨不同批次常规地显示出脱酰胺化水平>70%, 并且在大多数情况下脱酰胺化水平>90%。另外的天冬酰胺残基 (N94、N253、N270、N304、N409、N477和Q599) 跨不同批次也显示出至多约20%的脱酰胺化水平。最初使用胰蛋白酶消化物鉴定脱酰胺化水平, 并用胰凝乳蛋白酶消化对其进行验证。

[0137] AAVhu68衣壳含有具有来自SEQ ID NO:2中的预测的氨基酸残基的修饰的vp1蛋白内、vp2蛋白内和vp3蛋白内的亚群体。这些亚群体至少包含某些脱酰胺化的天冬酰胺 (N或Asn) 残基。例如, 某些亚群体包括SEQ ID NO:2中的天冬酰胺-甘氨酸对中的至少一个、两个、三个或四个高度脱酰胺化的天冬酰胺 (N) 位置, 并且任选地进一步包括其它脱酰胺化的氨基酸, 其中脱酰胺化引起氨基酸变化和其它任选的修饰。在本文中描述了这些和其它修饰的各种组合。

[0138] 在某些实施例中,如本文所描述的rAAV是自身互补AAV。“自身互补AAV”是指其中由重组AAV核酸序列所携带的编码区已经被设计成形成分子内双链DNA模板的构建体。感染后,未等待细胞介导的第二条链合成,而是两条互补的半scAAV将缔合以形成易于立即复制和转录的一条双链DNA(dsDNA)。参见例如D M McCarty等人,“自身互补重组腺相关病毒(scAAV)载体独立于DNA合成而促进高效转导(Self-complementary recombinant adeno-associated virus(scAAV) vectors promote efficient transduction independently of DNA synthesis)”,《基因疗法(Gene Therapy)》,(2001年8月),第8卷,第16期,第1248-1254页。自身互补AAV在例如美国专利第6,596,535号;7,125,717;和第7,456,683号中描述,所述美国专利中的每个美国专利通过引用整体并入本文中。

[0139] 可以使用已知的技术产生本文所描述的重组腺相关病毒(AAV)。参见例如W02003/042397;W0 2005/033321、W0 2006/110689;US 7588772 B2。此类方法涉及培养含有以下的宿主细胞:对AAV衣壳进行编码的核酸序列;功能性rep基因;如本文所描述的侧接AAV反向末端重复序列(ITR)的表达盒;以及足够的辅助功能以允许将表达盒包装到AAV衣壳蛋白中。本文还提供了含有以下的宿主细胞:对AAV衣壳进行编码的核酸序列;功能性rep基因;如所描述的载体基因组;以及足够的辅助功能以允许将载体基因组包装到AAV衣壳蛋白中。在一个实施例中,宿主细胞是HEK 293细胞。在W0 2017160360A2中更详细地描述了这些方法,其通过引用并入本文中。

[0140] 可以利用本领域技术人员可用的其它产生rAAV的方法。合适的方法可以包含但不限于杆状病毒表达系统或通过酵母生产。参见例如Robert M.Kotin等人,大规模重组腺相关病毒产生(Large-scale recombinant adeno-associated virus production).《人类分子遗传学(Hum Mol Genet.)》2011年4月15日;20(R1):R2-R6.于2011年4月29日在线公开.doi:10.1093/hmg/ddr141;Aucoin MG等人,使用三重感染在昆虫细胞中产生腺相关病毒载体:杆状病毒浓度比的优化(Production of adeno-associated viral vectors in insect cells using triple infection:optimization of baculovirus concentration ratios).《生物技术与生物工程(Biotechnol Bioeng.)》2006年12月20日;95(6):1081-92;SAMI S.THAKUR,在酵母中产生重组腺相关病毒载体(Production of Recombinant Adeno-associated viral vectors in yeast).论文提交给佛罗里达大学研究生院(Graduate School of the University of Florida),2012;Kondratov O等人,在人类对昆虫细胞中制造的重组腺相关病毒载体的直接头对头评估(Direct Head-to-Head Evaluation of Recombinant Adeno-associated Viral Vectors Manufactured in Human versus Insect Cells),《分子疗法(Mol Ther.)》2017年8月10日.pii:S1525-0016(17)30362-3.doi:10.1016/J.ymthe.2017.08.003.[印刷前的电子出版物];Mietzsch M等人,OneBac 2.0:用于产生使外源DNA的衣壳化最小化的AAV1、AAV2和AAV8载体的Sf9细胞系(OneBac 2.0:Sf9 Cell Lines for Production of AAV1,AAV2,and AAV8 Vectors with Minimal Encapsulation of Foreign DNA).《人类基因疗法方法(Hum Gene Ther Methods)》.2017年2月;28(1):15-22.doi:10.1089/hgtb.2016.164.;Li L等人,新颖的重组腺相关病毒复制型基因组的生产和表征:用于基因转移的真核DNA来源(Production and characterization of novel recombinant adeno-associated virus replicative-form genomes:a eukaryotic source of DNA for gene transfer).《公共科学图书馆·综合

(PLoS One)》.2013年8月1日;8(8):e69879.doi:10.1371/Journal.pone.0069879.于2013年印刷;Galibert L等人,在昆虫细胞中大规模生产腺相关病毒载体以趋向治疗神经肌肉疾病的最新进展(Latest developments in the large-scale production of adeno-associated virus vectors in insect cells toward the treatment of neuromuscular diseases).《无脊椎病理学杂志(J Invertebr Pathol.)》2011年7月;107增刊:S80-93.doi:10.1016/j.jip.2011.05.008;以及Kotin RM,大规模重组腺相关病毒生产(Large-scale recombinant adeno-associated virus production).《人类分子遗传学》2011年4月15日;20(R1):R2-6.doi:10.1093/Hmg/ddr141.电子版2011年4月29日。

[0141] 在高盐浓度下进行两步亲和色谱纯化,然后通过使用阴离子交换树脂色谱来纯化载体药物产物并去除空衣壳。在题为“AAV9的可分级纯化方法(Scalable Purification Method for AAV9)”的WO 2017/160360中更详细地描述了这些方法,其通过引用并入本文中。简而言之,用于从基因组缺陷型AAV9中间体中分离具有包装的基因组序列的rAAV9颗粒的方法涉及使包括重组AAV9病毒颗粒和AAV9衣壳中间体的悬浮液经受高效液相色谱,其中AAV9病毒颗粒和AAV9中间体与在10.2的pH下平衡的强阴离子交换树脂结合并经受盐梯度,同时监测洗脱液在约260和约280下的紫外线吸光度。尽管对于rAAV9不是最佳的,但是pH的范围可以为约10.0到10.4。在此方法中,当A260/A280的比率达到拐点时,从洗脱的级分中收集AAV9完整衣壳。在一个实例中,对于亲和色谱步骤,可以将经渗滤的产物应用于有效捕获AAV2/9血清型的Capture Select™ Poros-AAV2/9亲和树脂(生命科技公司(Life Technologies))。在这些离子条件下,显著百分比的残留的细胞DNA和蛋白质流过柱,而AAV颗粒则被有效捕获。

[0142] 用于表征或定量rAAV的常规方法对于本领域技术人员是可用的。为了计算空颗粒和满颗粒的含量,将所选样品(例如,在本文的实例中经过碘克沙醇(iodixanol)梯度纯化的制剂,其中GC数=颗粒数)的VP3带体积相对于加载的GC颗粒进行作图。所得线性等式($y = mx + c$)用于计算测试制品峰的带状体积中的颗粒的数量。然后将加载的每20 μ L颗粒数量(pt)乘以50,以得到颗粒(pt)/mL。将pt/mL除以GC/mL得到颗粒与基因组拷贝的比率(pt/GC)。pt/mL-GC/mL得到空pt/mL。空pt/mL除以pt/mL并且 $\times 100$ 得到空颗粒的百分比。通常,用于测定具有包装的基因组的空衣壳和AAV载体颗粒的方法是本领域已知的。参见例如Grimm等人,《基因疗法》(1999)6:1322-1330;Sommer等人,《分子疗法(Molec. Ther.)》(2003)7:122-128。为了测试变性的衣壳,所述方法包含使经处理的AAV原液经受SDS-聚丙烯酰胺凝胶电泳(由能够分离三种衣壳蛋白的任何凝胶组成,例如含有含3-8%Tris-乙酸盐的缓冲液的梯度凝胶),然后运行凝胶直到分离出样品材料,并且将凝胶印迹到尼龙或硝酸纤维素膜(优选地尼龙)上。然后,将抗AAV衣壳抗体用作与变性的衣壳蛋白结合的初级抗体,优选地抗AAV衣壳单克隆抗体,最优选地B1抗AAV2单克隆抗体(Wobus等人,《病毒学杂志(J. Viral.)》(2000)74:9281-9293)。然后使用次级抗体,所述次级抗体与初级抗体结合并且含有一种用于检测与初级抗体的结合的装置,更优选地是含有与其共价结合的检测分子的抗IgG抗体,最优选地是与辣根过氧化物酶共价连接的绵羊抗小鼠IgG抗体。一种用于检测结合的方法用于半定量地确定初级抗体与次级抗体之间的结合,优选地是能够检测放射性同位素发射、电磁辐射或比色变化的检测方法,最优选地是化学发光检测试剂盒。例如,对于SDS-PAGE,可以从柱级分中提取样品并在含有还原剂(例如,DTT)的SDS-PAGE上样缓冲

液中加热,并且在预制的梯度聚丙烯酰胺凝胶(例如,Novex)上解析衣壳蛋白。可以根据制造商的说明使用SilverXpress(加利福尼亚州英杰公司(Invitrogen,CA))或其它合适的染色方法(即SYPRO红宝石色或考马斯染色)进行银染色。在一个实施例中,可以通过定量实时PCR(Q-PCR)测量柱级分中的AAV载体基因组(vg)的浓度。将样品稀释并用DNase I(或另一种合适的核酸酶)消化以去除外源DNA。在核酸酶失活后,使用引物和对引物之间的DNA序列具有特异性的TaqManTM荧光探针进一步稀释和扩增样品。在Applied Biosystems Prism 7700序列检测系统上测量每种样品达到定义的荧光水平所需的周期的数量(阈值周期,Ct)。含有与AAV载体中所含序列相同的序列的质粒DNA用于在Q-PCR反应中产生标准曲线。从样品获得的周期阈值(Ct)的值用于通过相对于质粒标准曲线的Ct值对其进行归一化来确定载体基因组效价。也可以使用基于数字PCR的端点测定。

[0143] 在一方面,使用了经过优化的q-PCR方法,所述方法利用了广谱丝氨酸蛋白酶,例如蛋白酶K(如可从凯杰公司(Qiagen)商购获得)。更具体地,经过优化的qPCR基因组效价测定与标准测定类似,不同之处在于在DNase I消化之后,将样品用蛋白酶K缓冲液稀释并用蛋白酶K处理,然后进行热失活。合适地,以等于样品大小的量用蛋白酶K缓冲液稀释样品。蛋白酶K缓冲液可以浓缩2倍或更多倍。通常,蛋白酶K处理为约0.2mg/mL,但是可以在0.1g/mL到约1mg/mL之间变化。处理步骤通常在约55℃下进行持续约15分钟,但是可以在较低温度(例如,约37℃到约50℃)下进行持续较长的时间段(例如,约20分钟到约30分钟),或者在较高的温度(例如,至多约60℃)下进行持续较短的时间段(例如,约5到10分钟)。类似地,热失活通常在约95℃下持续约15分钟,但是温度可以降低(例如,约70℃到约90℃)并且时间延长(例如,约20分钟到约30分钟)。然后将样品稀释(例如,1000倍),并如标准测定中所描述的进行TaqMan分析。

[0144] 另外地或可替代地,可以使用液滴数字PCR(ddPCR)。例如,已经描述了用于通过ddPCR确定单链和自身互补AAV载体基因组效价的方法。参见例如M.Lock等人,《人类基因疗法方法》.2014年4月;25(2):115-25.doi:10.1089/Hgtb.2013.131.电子出版2014年2月14日。

[0145] 用于确定衣壳蛋白的vp1、vp2与vp3之间的比率的方法也是可用的。参见例如Vamseedhar Rayaprolu等人,腺相关病毒衣壳稳定性和动力学的比较分析(Comparative Analysis of Adeno-Associated Virus Capsid Stability and Dynamics),《病毒学杂志(JVirology)》2013年12月;87(24):13150-13160;Buller RM,Rose JA.1978.KB细胞中腺病毒相关病毒诱导多肽的表征(Characterization of adenovirus-associated virus-induced polypeptides in KB cells).《病毒学杂志》25:331-338;以及Rose JA,Maizel JV,Inman JK,Shatkin AJ.1971.腺病毒相关病毒的结构蛋白(Structural proteins of adenovirus-associated viruses).《病毒学杂志》8:766-770。

[0146] 应当理解,本文所描述的rAAV中的组合物旨在应用于本说明书中描述的其它组合物、方案、方面、实施例和方法。

[0147] 药物组合物

[0148] 包括hGAA780I融合蛋白或包括hGAA780I融合蛋白转基因的表达盒的药物组合物可以是液体悬浮液、经冻干或经冷冻的组合物或另一种合适的调配物。在某些实施例中,所述组合物包括hGAA780I融合蛋白或表达盒和形成悬浮液的生理上相容的液体(例如,溶液、

稀释剂、载体)。此类液体优选地是水基的并且可以含有以下中的一种或多种:缓冲剂、表面活性剂、pH调节剂、防腐剂或其它合适的赋形剂。下文更详细地讨论了合适的组分。所述药物组合物包括水性悬浮液和任何所选赋形剂以及hGAA780I融合蛋白或表达盒。

[0149] 在某些实施例中,药物组合物包括包含转基因和非病毒递送系统的表达盒。这可以包含例如裸DNA、裸RNA、无机颗粒、脂质或脂质样颗粒、基于壳聚糖的调配物和本领域已知的并且例如由Ramamoorthi和Narvekar描述的其它调配物,如上所述。在其它实施例中,药物组合物是包括表达盒的悬浮液,所述表达盒包括在病毒载体系统中工程化的转基因。在某些实施例中,药物组合物包括非复制病毒载体。合适的病毒载体可以包含任何合适的递送载体,例如重组腺病毒、重组慢病毒、重组博卡病毒、重组腺相关病毒(AAV)或另一种重组细小病毒。在某些实施例中,病毒载体是用于向有需要的患者递送基因产物的重组AAV。

[0150] 在一个实施例中,药物组合物包括hGAA780I融合蛋白或包括hGAA780I融合蛋白的编码序列的表达盒和适合通过脑室内(ICV)、鞘内(IT)、脑池内或静脉内(IV)注射递送的调配物缓冲液。在一个实施例中,表达盒是重组病毒载体(即,携带融合蛋白的rAAV.hGAA780I)中包装的载体基因组的一部分。

[0151] 在一个实施例中,药物组合物包括用于作为酶替代疗法(ERT)递送给受试者的hGAA780I融合蛋白或其功能片段。此类药物组合物通常静脉内施用,然而在一些情况下皮内、肌肉内或口服施用也是可能的。可以施用组合物来预防性治疗患有庞贝病或有患庞贝病风险的个体。对于治疗应用,将药物组合物以足以降低积累代谢物浓度和/或防止或阻止代谢物进一步积累的量施用于患有已确定疾病的患者。对于有溶酶体酶缺乏症风险的个体,以足以防止或抑制代谢物积累的量预防性地施用药物组合物。本文所述的经修饰的GAA组合物以治疗有效量施用。一般而言,治疗有效量可以根据受试者的医学状况的严重性以及受试者的年龄、一般状况和性别而变化。剂量可以由医生确定,并可以根据需要进行调整以适应所观察到的治疗效果。在一方面,本文提供了一种用于ERT的药物组合物,所述药物组合物被调配成含有单位剂量的hGAA780I融合蛋白或其功能片段。

[0152] 在一个实施例中,组合物包含适合于递送到受试者的最终调配物,所述组合物是例如缓冲到生理上相容的pH和盐浓度的水性液体悬浮液。任选地,调配物中存在一种或多种表面活性剂。在另一个实施例中,可以将组合物作为稀释液以施用于受试者的浓缩物运输。在其它实施例中,可以在施用时将组合物冻干并重构。

[0153] 在一个实施例中,如本文所提供的组合物包括溶解于水性悬浮液中的表面活性剂、防腐剂、赋形剂和/或缓冲液。在一个实施例中,缓冲液是PBS。在另一个实施例中,缓冲液是人工脑脊液(aCSF),例如艾略特调配物缓冲液(Elliott's formulation buffer);或哈佛设备灌注流体(Harvard apparatus perfusion fluid)(具有以下最终离子浓度(以mM为单位)的人工CSF:Na 150;K 3.0;Ca 1.4;Mg 0.8;P 1.0;Cl 155)。各种合适的溶液是已知的,包含那些包含以下中的一种或多种的溶液:缓冲盐水、表面活性剂和生理上相容的盐或盐的混合物,其离子强度被调节到等同于约100mM氯化钠(NaCl)到约250mM氯化钠,或被调节到等离子浓度的生理上相容的盐。

[0154] 合适地,将调配物调节到生理上可接受的pH,例如,在pH 6到8、或pH 6.5到7.5、pH 7.0到7.7或pH 7.2到7.8的范围内。由于脑脊液的pH为约7.28到约7.32,对于鞘内递送,可能期望在此范围内的pH;而对于静脉内递送,可能期望的pH为6.8到约7.2。然而,可以选择

最宽范围和这些子范围内的其它pH用于其它递送途径。

[0155] 可以从无毒的非离子表面活性剂中选择合适的表面活性剂或表面活性剂的组合。在一个实施例中,选择终止于伯羟基的双官能嵌段共聚物表面活性剂,例如**Pluronic®F68** [BASF],也被称为泊洛沙姆(Poloxamer) 188,其具有中性pH,平均分子量为8400。可以选择其它表面活性剂和其它泊洛沙姆,即非离子型三嵌段共聚物,所述非离子型三嵌段共聚物由与聚氧乙烯(聚(环氧乙烷))的两个亲水链侧接的聚氧丙烯(聚(环氧丙烷))的中心疏水链、SOLUTOL HS 15(聚乙二醇-15羟基硬脂酸酯)、LABRASOL(聚氧辛酸甘油酯)、聚氧10油醚、TWEEN(聚氧乙烯山梨聚糖脂肪酸酯)、乙醇和聚乙二醇构成。在一个实施例中,调配物含有泊洛沙姆。这些共聚物通常以字母“P”(对于泊洛沙姆)命名,后跟三个数字:前两位数字 $\times 100$ 给出了聚氧丙烯核的近似分子量,并且最后一位数字 $\times 10$ 给出了聚氧乙烯含量的百分比。在一个实施例中,选择了泊洛沙姆188。表面活性剂可以以悬浮液的至多约0.0005%到约0.001%的量存在。

[0156] 在一个实例中,调配物可以含有例如缓冲盐水溶液,所述缓冲盐水溶液包括水中的氯化钠、碳酸氢钠、葡聚糖、硫酸镁(例如,硫酸镁 $7H_2O$)、氯化钾、氯化钙(例如,氯化钙 $2H_2O$)、磷酸氢二钠及其混合物中的一种或多种。合适地,对于鞘内递送,同渗浓度在与脑脊液相容的范围内(例如,约275到约290);参见例如emedicine.medscape.com/article/2093316-overview。任选地,对于鞘内递送,可以将可商购获得的稀释剂用作悬浮剂,或与另一种悬浮剂和其它任选的赋形剂组合。参见例如Elliotts B®解决方案[Lukare Medical]。

[0157] 在其它实施例中,调配物可以含有一种或多种渗透增强剂。合适的渗透增强剂的实例可以包含例如甘露醇、甘氨酸胆酸钠、牛磺胆酸钠、脱氧胆酸钠、水杨酸钠、辛酸钠、癸酸钠、月桂基硫酸钠、聚氧乙烯-9-月桂基醚或EDTA。

[0158] 另外提供了一种药物组合物,所述药物组合物包括药学上可接受的载体和包括如本文所描述的核酸序列的载体。如本文所使用的,“载体”包含任何和所有溶剂、分散介质、媒剂、涂层、稀释剂、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲液、载体溶液、悬浮液、胶质物等。这些介质和药剂用于药物活性物质的用途在本领域中是熟知的。补充性活性成分也可以并入到组合物中。递送媒剂(如脂质体、纳米胶囊、微颗粒、微球、脂质颗粒、囊泡等)可以用于将本文描述的组合物引入到合适的宿主细胞中。具体地,rAAV载体基可以被调配成用于递送或包封在脂质颗粒、脂质体、囊泡、纳米球或纳米颗粒等中。在一个实施例中,治疗有效量的载体包含在药物组合物中。载体的选择不是对本发明的限制。其它常规的药学上可接受的载体,如防腐剂或化学稳定剂。合适的示例性防腐剂包含氯丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸甲酯、乙基香草醛、甘油、苯酚和对氯苯酚。合适的化学稳定剂包含明胶和白蛋白。

[0159] 短语“药学上可接受的”是指当向宿主施用时不会产生过敏或类似不良反应的分子实体和组合物。

[0160] 如本文所使用的,术语“剂量”或“量”可以指在治疗过程中向受试者递送的总剂量或量或以单一单位(或多单位或分剂量)施用递送的剂量或量。

[0161] 本文所描述的水性悬浮液或药物组合物被设计成用于通过任何合适的途径或不同途径的组合递送到有需要的受试者。在一个实施例中,药物组合物被调配用于通过脑室

内 (ICV)、鞘内 (IT) 或脑池内注射递送。在一个实施例中, 本文所描述的组合物被设计成用于通过静脉内注射递送到有需要的受试者。可替代地, 可以选择其它施用途径 (例如, 口服、吸入、鼻内、气管内、动脉内、眼内、肌内和其它肠胃外途径)。

[0162] 如本文所使用的, 术语“鞘内递送”或“鞘内施用”是指通过注射到椎管中, 更具体地注射到蛛网膜下腔中使得其到达脑脊液 (CSF) 的药物施用途径。鞘内递送可以包含腰椎穿刺、心室内、枕骨下/脑池内和/或C1-2穿刺。例如, 可以通过腰椎穿刺引入材料以在整个蛛网膜下腔扩散。在另一个实例中, 可以向小脑延髓池中注射。脑池内递送可以增加载体扩散和/或减少施用引起的毒性和炎症。参见例如Christian Hinderer等人, 在将AAV9递送到小脑延髓池中之后食蟹猴的中枢神经系统中的广泛基因转移 (Widespread gene transfer in the central nervous system of cynomolgus macaques following delivery of AAV9 into the cisterna magna).《分子疗法方法临床发展 (Mol Ther Methods Clin Dev.)》.2014;1:14051. 于2014年12月10日在线公开.doi:10.1038/mtm.2014.51。

[0163] 如本文所使用的, 术语“脑池内递送”或“脑池内施用”是指药物直接进入脑室或小脑延髓池 (cisterna magna cerebellomedullaris) 的脑脊液中, 更具体地是通过枕骨下穿刺或通过直接注射到小脑延髓池 (cisterna magna) 中或通过永久定位的管的施用途径。

[0164] 在一方面, 本文提供了一种药物组合物, 所述药物组合物包括含如本文所描述的载体的调配物缓冲液。在某些实施例中, 可以以剂量单位调配复制缺陷型病毒组合物, 以使含有的复制缺陷型病毒的量在约 1.0×10^9 GC到约 1.0×10^{16} GC的范围内 (以治疗平均体重为70kg的受试者), 包含所述范围内的所有整数或分数量, 并且对于人类患者而言, 优选地为 1.0×10^{12} GC到 1.0×10^{14} GC。在一个实施例中, 将组合物调配成每剂量含有至少 1×10^9 、 2×10^9 、 3×10^9 、 4×10^9 、 5×10^9 、 6×10^9 、 7×10^9 、 8×10^9 或 9×10^9 GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{10} 、 2×10^{10} 、 3×10^{10} 、 4×10^{10} 、 5×10^{10} 、 6×10^{10} 、 7×10^{10} 、 8×10^{10} 或 9×10^{10} GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{11} 、 2×10^{11} 、 3×10^{11} 、 4×10^{11} 、 5×10^{11} 、 6×10^{11} 、 7×10^{11} 、 8×10^{11} 或 9×10^{11} GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{12} 、 2×10^{12} 、 3×10^{12} 、 4×10^{12} 、 5×10^{12} 、 6×10^{12} 、 7×10^{12} 、 8×10^{12} 或 9×10^{12} GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{13} 、 2×10^{13} 、 3×10^{13} 、 4×10^{13} 、 5×10^{13} 、 6×10^{13} 、 7×10^{13} 、 8×10^{13} 或 9×10^{13} GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{14} 、 2×10^{14} 、 3×10^{14} 、 4×10^{14} 、 5×10^{14} 、 6×10^{14} 、 7×10^{14} 、 8×10^{14} 或 9×10^{14} GC, 包含所述范围内的所有整数或分数量。在另一个实施例中, 将组合物调配成每剂量含有至少 1×10^{15} 、 2×10^{15} 、 3×10^{15} 、 4×10^{15} 、 5×10^{15} 、 6×10^{15} 、 7×10^{15} 、 8×10^{15} 或 9×10^{15} GC, 包含所述范围内的所有整数或分数量。在一个实施例中, 对于人类应用, 剂量的范围可以为每剂量 1×10^{10} 到约 1×10^{12} GC, 包含所述范围内的所有整数或分数量。

[0165] 在一个实施例中, 提供了一种药物组合物, 所述药物组合物包括含如本文所描述的rAAV的调配物缓冲液。在一个实施例中, rAAV以约 1×10^9 基因组拷贝 (GC) /mL到约 1×10^{14} GC/mL调配。在另外的实施例中, rAAV以约 3×10^9 GC/mL到约 3×10^{13} GC/mL调配。在又另外的实施例中, rAAV以约 1×10^9 GC/mL到约 1×10^{13} GC/mL调配。在一个实施例中, rAAV以至少约

1×10^{11} GC/mL 调配。

[0166] 在一个实施例中,包括如本文所描述的rAAV的药物组合物以每克脑质量约 1×10^9 GC到每克脑质量约 1×10^{14} GC的剂量施用。

[0167] 应当理解,本文所描述的药物组合物中的组合物旨在应用于本说明书中描述的其它组合物、方案、方面、实施例和方法。

[0168] 治疗方法

[0169] 提供了用于治疗患有庞贝病的患者的治疗方案,所述治疗方案包括任选地与免疫调节剂组合的如本文所述的表达盒、rAAV和/或hGAA780I融合蛋白。在某些实施例中,所述患者患有晚发性庞贝病。在其它实施例中,所述患者患有儿童期初发型庞贝病。在某些实施例中,共治疗剂与表达盒、rAAV或hGAA780I融合蛋白一起递送,如免疫调节方案。另外地或可替代地,共疗法可以包含支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌力量训练(RMST)、酶替代疗法和/或膈肌起搏疗法中的一种或多种。在某些实施例中,所述患者接受rAAV的单次施用。在某些实施例中,所述患者接受包括如本文所述的表达盒和/或rAAV的组合物的单次施用。在某些实施例中,包括有效量表达盒的组合物的这种单次施用涉及至少一种共治疗剂。在某些实施例中,基本上同时通过两种不同的途径向患者施用表达盒、rAAV和/或hGAA780I融合蛋白或如本文所述。在某些实施例中,两种不同的注射途径是静脉内和鞘内施用。在一个实施例中,所述组合物是悬浮液,通过脑室内、鞘内、脑池内或静脉内递送到受试者。在某些实施例中,向患有 α -葡萄糖苷酶缺乏症的患者施用本文提供的组合物以改善心脏、呼吸和/或骨骼肌功能中的一种或多种。在某些实施例中,由于治疗,心脏、CNS(脑)和/或骨骼肌中的一个或多个中的糖原贮积和/或自噬积累减少。

[0170] 在某些实施例中,表达盒、rAAV、病毒或非病毒载体用于制备药物。在某些实施例中,提供了组合物用于治疗庞贝病的用途。

[0171] 这些组合物可以与其它疗法组合使用,所述其它疗法包含例如免疫疗法、酶替代疗法(例如,Lumizyme,由健赞公司(Genzyme)、赛诺菲公司(Sanofi Corporation)销售,并且在美国以外以Myozyme销售)。庞贝病的另外的治疗是对症和支持性的。例如,可能需要呼吸支持;物理疗法可能有助于增强呼吸肌;一些患者可能需要在夜间和/或白天通过机械通气(即双水平气道正压通气(bipap)或容积通气机)进行呼吸辅助。另外,在呼吸道感染期间可能需要另外的支持。一些患者可能会被推荐使用包含支具在内的矫形装置。如挛缩或脊柱畸形等某些骨科症状可能需要手术。一些婴儿可能需要插入经鼻部、沿食道向下并到胃中的喂食管(鼻胃管)。在一些儿童中,可能需要通过腹壁上的小手术开口将喂食管直接插入胃中。一些患有晚发型庞贝病的个体可能需要软饮食,但很少有个体需要喂食管。

[0172] 如本文所描述的,除非另外指明,否则术语“增加”(例如,在用hGAA780I融合蛋白治疗后增加hGAA水平,如在组织、血液等中所测量的)或“减少”、“降低”、“减轻”、“改善”、“延迟”或其任何语法变体或指示变化的任何类似术语意指与对应的参考(例如,未经治疗的对照或未患有庞贝病的处于正常状况的受试者)相比约5倍、约2倍、约1倍、约90%、约80%、约70%、约60%、约50%、约40%、约30%、约20%、约10%或约5%的变化。

[0173] 在本文中可互换使用的“患者”或“受试者”意指雄性或雌性哺乳动物,包含人、兽医或农场动物、家畜或宠物以及通常用于临床研究的动物。在一个实施例中,这些方法和组合物的受试者是人类患者。在一个实施例中,这些方法和组合物的受试者是男性或女性人

类。

[0174] 在一个实施例中,悬浮液的pH为约7.28到约7.32。

[0175] 可以由本领域的技术人员确定用于递送这些剂量和浓度的合适的体积。例如,可以选择约1 μ L到150mL的体积,其中对于成人而言,选择更大的体积。通常,对于新生婴儿,合适的体积为约0.5mL到约10mL,对于年龄较大的婴儿,可以选择约0.5mL到约15mL。对于幼儿,可以选择约0.5mL到约20mL的体积。对于儿童,可以选择至多约30mL的体积。对于青春期的少年和青少年,可以选择至多约50mL的体积。在仍其它实施例中,患者可以接受鞘内施用的体积选择为约5mL到约15mL或约7.5mL到约10mL。可以确定其它合适的体积和剂量。将调整剂量以平衡治疗益处与任何副作用,并且这种剂量可以根据采用重组载体的治疗应用而变化。

[0176] 在一个实施例中,包括如本文所描述的rAAV的组合物以每克脑质量约 1×10^9 GC到每克脑质量约 1×10^{14} GC的剂量施用。在某些实施例中,rAAV以每千克体重约 1×10^9 GC到每千克体重约 1×10^{13} GC的剂量全身共同施用。

[0177] 在一个实施例中,向受试者递送治疗有效量的本文所述的表达盒、rAAV或hGAA780I融合蛋白。如本文所使用的,“治疗有效量”是指表达盒、rAAV或hGAA780I融合蛋白或其组合的量。因此,在某些实施例中,所述方法包括与施用包括本文提供的hGAA780I融合蛋白酶的组合物组合来向受试者施用用于递送hGAA780I融合蛋白编码核酸序列的rAAV或表达盒。

[0178] 在一个实施例中,表达盒在载体基因组中以每克脑质量约 1×10^9 GC到每克(g)脑质量约 1×10^{13} 个基因组拷贝(GC)的量递送,包含所述范围内的所有整数或分数量和端点。在另一个实施例中,剂量为每克脑质量 1×10^{10} GC到每克脑质量约 1×10^{13} GC。在具体实施例中,向患者施用的载体的剂量为至少约 1.0×10^9 GC/g、约 1.5×10^9 GC/g、约 2.0×10^9 GC/g、约 2.5×10^9 GC/g、约 3.0×10^9 GC/g、约 3.5×10^9 GC/g、约 4.0×10^9 GC/g、约 4.5×10^9 GC/g、约 5.0×10^9 GC/g、约 5.5×10^9 GC/g、约 6.0×10^9 GC/g、约 6.5×10^9 GC/g、约 7.0×10^9 GC/g、约 7.5×10^9 GC/g、约 8.0×10^9 GC/g、约 8.5×10^9 GC/g、约 9.0×10^9 GC/g、约 9.5×10^9 GC/g、约 1.0×10^{10} GC/g、约 1.5×10^{10} GC/g、约 2.0×10^{10} GC/g、约 2.5×10^{10} GC/g、约 3.0×10^{10} GC/g、约 3.5×10^{10} GC/g、约 4.0×10^{10} GC/g、约 4.5×10^{10} GC/g、约 5.0×10^{10} GC/g、约 5.5×10^{10} GC/g、约 6.0×10^{10} GC/g、约 6.5×10^{10} GC/g、约 7.0×10^{10} GC/g、约 7.5×10^{10} GC/g、约 8.0×10^{10} GC/g、约 8.5×10^{10} GC/g、约 9.0×10^{10} GC/g、约 9.5×10^{10} GC/g、约 1.0×10^{11} GC/g、约 1.5×10^{11} GC/g、约 2.0×10^{11} GC/g、约 2.5×10^{11} GC/g、约 3.0×10^{11} GC/g、约 3.5×10^{11} GC/g、约 4.0×10^{11} GC/g、约 4.5×10^{11} GC/g、约 5.0×10^{11} GC/g、约 5.5×10^{11} GC/g、约 6.0×10^{11} GC/g、约 6.5×10^{11} GC/g、约 7.0×10^{11} GC/g、约 7.5×10^{11} GC/g、约 8.0×10^{11} GC/g、约 8.5×10^{11} GC/g、约 9.0×10^{11} GC/g、约 9.5×10^{11} GC/g、约 1.0×10^{12} GC/g、约 1.5×10^{12} GC/g、约 2.0×10^{12} GC/g、约 2.5×10^{12} GC/g、约 3.0×10^{12} GC/g、约 3.5×10^{12} GC/g、约 4.0×10^{12} GC/g、约 4.5×10^{12} GC/g、约 5.0×10^{12} GC/g、约 5.5×10^{12} GC/g、约 6.0×10^{12} GC/g、约 6.5×10^{12} GC/g、约 7.0×10^{12} GC/g、约 7.5×10^{12} GC/g、约 8.0×10^{12} GC/g、约 8.5×10^{12} GC/g、约 9.0×10^{12} GC/g、约 9.5×10^{12} GC/g、约 1.0×10^{13} GC/g、约 1.5×10^{13} GC/g、约 2.0×10^{13} GC/g、约 2.5×10^{13} GC/g、约 3.0×10^{13} GC/g、约 3.5×10^{13} GC/g、约 4.0×10^{13} GC/g、约 4.5×10^{13} GC/g、约 5.0×10^{13} GC/g、约 5.5×10^{13} GC/g、约 6.0×10^{13} GC/g、约 6.5×10^{13} GC/g、约 7.0×10^{13} GC/g、约 7.5×10^{13} GC/g、约 8.0×10^{13} GC/g、约 8.5×10^{13} GC/g、约 $9.0 \times$

10^{13} GC/g、约 9.5×10^{13} GC/g或约 1.0×10^{14} GC/g脑质量。

[0179] 在一个实施例中,治疗方法包括递送hGAA780I融合蛋白作为酶替代疗法。在某些实施例中,hGAA780I融合蛋白作为ERT与基因疗法(包括但不限于本文提供的表达盒或rAAV)组合递送。在某些实施例中,所述方法包括向受试者施用多于一种ERT(例如,包括与另一种治疗性蛋白如Lumizyme组合的hGAA780I融合蛋白的组合物)。可以每1天、2天、3天、4天、5天、6天、7天、8天、9天、10天或更多天向受试者施用包括本文所述的hGAA780I融合蛋白的组合物。施用可以通过静脉内输注给门诊患者,规定每周、每月或每两个月施用。化合物的适当治疗有效剂量由治疗临床医生选择并且包含约 $1\mu\text{g/kg}$ 到约 500mg/kg 、约 10mg/kg 到约 100mg/kg 、约 20mg/kg 到约 100mg/kg 和大约 20mg/kg 到大约 50mg/kg 。在一些实施例中,合适的治疗剂量选自例如0.1、0.25、0.5、0.75、1、5、10、15、20、30、40、50、60、70和 100mg/kg 。

[0180] 在某些实施例中,所述方法包括以每周向患者 10mg/kg 患者体重或更多的剂量向受试者施用hGAA780I融合蛋白。通常每周剂量大于 10mg/kg 。剂量方案的范围可以为每周 10mg/kg 到每周至少 1000mg/kg 。通常,剂量方案是每周 10mg/kg 、每周 15mg/kg 、每周 20mg/kg 、每周 25mg/kg 、每周 30mg/kg 、每周 35mg/kg 、 40mg/kg 周、每周 45mg/kg 、 60mg/kg 周、每周 80mg/kg 和每周 120mg/kg 。在优选方案中, 10mg/kg 、 15mg/kg 、 20mg/kg 、 30mg/kg 或 40mg/kg 每周施用一次、两次或三次。治疗通常持续至少4周,有时持续24周,并且有时持续患者的一生。任选地,在治疗后(例如,在血浆或肌肉中)监测人 α -葡萄糖苷酶的水平,并且当检测到的水平显著低于正常人的值(例如,小于所述值的20%)时施用另外的剂量。在一个实施例中,hGAA780I以初始“高”剂量(即,“负荷剂量”)施用,随后以较低剂量(即,“维持剂量”)施用。负荷剂量的实例是至少约 40mg/kg 患者体重,每周1到3次(例如,持续1、2或3周)。维持剂量的实例是每周至少约5到至少约 10mg/kg 患者体重或更多,如每周 20mg/kg 、每周 30mg/kg 、 40mg/kg 周。在某些实施例中,在剂量时间段期间以增加的速率施用剂量。这可以通过增加静脉内输注的速率或通过使用以恒定速率施用的hGAA780I融合蛋白的浓度增加的梯度来实现。以这种方式施用可以降低免疫原性反应的风险。在某些实施例中,静脉内输注发生在几个小时的时间段内(例如,1-10小时,并且优选地2-8小时,更优选地3-6小时),并且在施用时间段期间每隔一段时间增加输注速率。

[0181] 在一个实施例中,所述方法进一步包括受试者接受免疫抑制共疗法。用于此类共疗法的免疫抑制剂包括但不限于糖皮质激素、类固醇、抗代谢药、T细胞抑制剂、大环内酯类(例如,雷帕霉素或雷帕霉素类似物)以及细胞生长抑制剂,包含烷化剂、抗代谢药、细胞毒性抗生素、抗体或对免疫亲和素具有活性的药剂。免疫抑制剂可以包含氮芥(nitrogen mustard)、亚硝基脲(nitrosourea)、铂化合物、甲氨蝶呤(methotrexate)、硫唑嘌呤(azathioprine)、巯嘌呤(mercaptopurine)、氟尿嘧啶(flourouracil)、更生霉素(dactinomycin)、蒽环霉素(anthracycline)、丝裂霉素C(mitomycin C)、博来霉素(bleomycin)、光神霉素(mithramycin)、IL-2受体或CD3定向抗体、抗IL-2抗体、环孢素(ciclosporin)、他克莫司(tacrolimus)、西罗莫司(sirolimus)、IFN- β 、IFN- γ 、阿片类或TNF- α (肿瘤坏死因子- α)结合剂。在某些实施例中,可以在施用基因疗法之前或之后的第0天、第1天、第2天、第7天或更多天开始免疫抑制疗法。可以在基因疗法施用之后以相同的剂量或经过调整的剂量继续使用这些药物中的一种或多种药物。根据需要,此类疗法可以持续约1周(7天)、约60天或更长时间。

[0182] 在一个实施例中,将包括如本文所描述的表达盒的组合物施用于有需要的受试者一次。在某些实施例中,表达盒通过rAAV递送。应当理解,本文所描述的组合物和方法旨在应用于本说明书中描述的其它组合物、方案、方面、实施例和方法。

[0183] 本文提供的组合物和方法可以用于治疗婴儿型庞贝病或晚发型庞贝病和/或与其相关的症状。在某些实施例中,功效可以通过疾病的一种或多种症状的改善或疾病进展的减缓来确定。婴儿型庞贝病的症状包含但不限于肌张力减退、呼吸(respiratory/breathing)问题、肝肿大、肥厚性心肌病以及心脏、肌肉、CNS(尤其是运动神经元)中的糖原贮积。晚发型庞贝病的症状包含但不限于近端肌肉无力、呼吸问题以及肌肉和运动神经元中的糖原贮积。施用途径可以基于患者的状况和/或诊断确定。在某些实施例中,提供了一种用于治疗被诊断为患有婴儿型庞贝病或晚发型庞贝病的患者的方法,所述方法包含通过IV和ICM途径的组合施用本文所述的用于递送hGAA780I融合蛋白的rAAV。在一些实施例中,向被鉴定为患有晚发型庞贝病的患者施用仅包含全身递送rAAV(例如,仅IV)的治疗。如本文所描述的,包括rAAV的组合物递送可以与酶替代疗法(ERT)组合。在某些实施例中,提供了一种用于治疗被诊断为患有庞贝病的受试者的方法,所述方法包含ICM递送本文所述的rAAV与ERT的组合。在某些实施例中,通过ICM注射向被鉴定为患有婴儿型庞贝病的受试者施用本文所述的rAAV,并且所述受试者还接受ERT以治疗外周疾病的各个方面。

[0184] 如本文所描述的,“核酸”可以是RNA、DNA或其修饰,并且可以是单链或双链的,并且可以选自例如包含以下的组:编码所关注蛋白质的核酸、寡核苷酸、核酸类似物,例如肽-核酸(PNA)、假互补PNA(pc-PNA)、锁核酸(LNA)等。此类核酸序列包含例如但不限于编码例如充当转录阻遏物的蛋白质的核酸序列、反义分子、核酶、小的抑制性核酸序列,例如但不限于RNAi、shRNAi、siRNA、微RNAi(mRNAi)、反义寡核苷酸等。

[0185] 用于“反向翻译”蛋白质、肽或多肽的方法是本领域技术人员已知的。一旦已知蛋白质的序列,就会有基于网络的和可商购获得的计算机程序以及基于服务的公司,所述基于服务的公司将氨基酸序列反向翻译为核酸编码序列。参见例如EMBOSS的backtranseq(可在`ebi.ac.uk/Tools/st`在线获得);Gene Infinity(可在`geneinfinity.org/sms/sms_-backtranslation.html`在线获得);Expasy(可在`expasy.org/tools/`在线获得)。在一个实施例中,RNA和/或cDNA编码序列被设计成在人细胞中最佳表达。

[0186] 在核酸序列的上下文中,术语“百分比(%)同一性”、“序列同一性”、“百分比序列同一性”或“百分比相同的”是指两个序列中的残基在比对以获得对应性时是相同的。期望序列同一性比较的长度可以超过基因组的全长、基因编码序列的全长或至少约500到5000个核苷酸的片段。然而,也可能期望较小片段之间的同一性,例如至少约九个核苷酸,通常至少约20到24个核苷酸、至少约28到32个核苷酸、至少约36个或更多个核苷酸。

[0187] 可以容易地确定蛋白质全长、多肽、约32个氨基酸、约330个氨基酸或其肽片段或对应核酸序列编码序列上的氨基酸序列的百分比同一性。合适的氨基酸片段的长度可以是至少约8个氨基酸并且可以是至多约700个氨基酸。通常,当提及两种不同序列之间的“同一性”、“同源性”或“类似性”时,参考“比对”序列来确定“同一性”、“同源性”或“类似性”。“比对”序列或“比对”是指与参考序列相比,通常含有对丢失的或另外的碱基或氨基酸的校正的多个核酸序列或蛋白质(氨基酸)序列。

[0188] 使用多种公开或可商购获得的多序列比对程序中的任一种进行比对。序列比对程

序可供用于氨基酸序列,所述程序例如“Clustal X”、“Clustal Omega”、“MAP”、“PIMA”、“MSA”、“BLOCKMAKER”、“MEME”以及“Match-Box”程序。通常,以默认设置使用这些程序中的任何程序,但是本领域技术人员可以根据需要改变这些设置。可替代地,本领域技术人员可以利用另一种算法或计算机程序,所述算法或程序提供至少与通过参考算法和程序所提供的一样水平的同一性或比对。参见例如 J. D. Thompson 等人,《核酸研究 (Nucl. Acids. Res.)》, 27 (13): 2682-2690 (1999)。

[0189] 多个序列比对程序也可用于核酸序列。此类程序的实例包含“Clustal W”、“Clustal Omega”、“CAP 序列组装”、“BLAST”、“MAP”和“MEME”,所述程序可通过因特网上的 Web 服务器进行访问。此类程序的其它来源是本领域技术人员已知的。可替代地,也使用了载体 NTI 实用程序。本领域已知的许多算法可以用于测量核苷酸序列同一性,包含上述程序中包含的那些。作为另一个实例,可以使用 GCG 6.1 版本的程序 FastaTM 比较多核苷酸序列。FastaTM 提供了查询序列与搜索序列之间最佳重叠区的比对和序列同一性百分比。例如,核酸序列之间的序列同一性百分比可以是使用如 GCG 6.1 版本中所提供的采用其默认参数 (字号 6 和评分矩阵的 NOPAM 系数) 的 FastaTM 所确定的,所述程序通过引用并入本文中。

[0190] 如本文所使用的,术语“调控序列”或“表达控制序列”是指核酸序列,如起始子序列、增强子序列和启动子序列,所述核酸序列诱导、抑制或以其它方式控制与其可操作地连接的蛋白质编码核酸序列的转录。

[0191] 如用于描述核酸序列或蛋白质的术语“外源性”意指核酸或蛋白质并非天然存在于其在染色体或宿主细胞中的存在位置中。外源性核酸序列还指源自相同宿主细胞或受试者并插入到所述相同宿主细胞或受试者中的序列,但其以非天然状态存在,例如,不同的拷贝数或在不同调控元件的控制下。

[0192] 如用于描述核酸序列或蛋白质的术语“异源性”意指核酸或蛋白质源自与在其中表达所述核酸或蛋白质的宿主细胞或受试者不同的生物体或相同的生物体的不同物种。术语“异源性”当参考质粒、表达盒或载体中的蛋白质或核酸使用时表明蛋白质或核酸与另一序列或子序列一起存在,在自然界中未发现所讨论蛋白质或核酸与所述蛋白质或核酸彼此间的相同关系。

[0193] “包括”是意指包含其它组分或方法步骤的术语。当使用“包括”时,应当理解,相关实施例包含:使用“由……组成”术语的描述,其不包括其它组成或方法步骤;以及使用“基本上由……组成”术语的描述,其不包括基本上改变实施例或本发明的性质的任何组成或方法步骤。应当理解,虽然说明书中的各个实施例是使用“包括”语言来呈现的,但是在各种情况下,也使用“由……组成”或“基本上由……组成”语言来描述相关实施例。

[0194] 如本文所使用的,后跟数字 (nn) 值的术语“e”是指指数并且此术语可与“ $\times 10^{nn}$ ”互换使用。例如,3e13 等效于 3×10^{13} 。

[0195] 应当注意,术语“一个/一种 (a/an)”是指一个或多个/一种或多种,例如,“一种载体”应被理解为表示一种或多种载体。如此,术语“一个” (或“一种”)、“一个或多个”和“至少一个”在本文可互换地使用。

[0196] 如本文所使用的,除非另有说明,否则术语“约”意指相对于给定参考的正或负 10% 的变化性。

[0197] 实例

[0198] 现在将参考以下实例描述本发明。提供这些实例仅为了说明目的,并且本发明绝不应理解为限于这些实例,而是应理解为涵盖由于本文所提供的教导而变得明显的任何和所有变化。

[0199] 实例1:材料与方法

[0200] 载体产生

[0201] 将在780处具有Val的参考GAA序列和具有V780I突变的序列反向翻译,并且将核苷酸序列工程化以产生用于AAV产生的具有在CAG启动子的控制下的表达盒的顺式质粒。另外,将天然hGAA的cDNA序列(参考序列)克隆到相同的AAV-顺式骨架中,以与非经工程化的序列进行比较。如前所述,AAVhu68载体由宾夕法尼亚大学载体核心公司(Penn Vector Core)生产和滴定。(Lock等人2010,《人类基因疗法(Hum Gene Ther)》21(10):1259-1271)。简而言之,HEK293细胞被三重转染,并且采集培养物上清液,浓缩并用碘克沙醇梯度纯化。如先前所述,使用靶向兔 β -珠蛋白polyA序列的引物,通过液滴数字PCR滴定经过纯化的载体(Lock等人(2014).《人类基因疗法方法》25(2):115-125)。

[0202] 动物

[0203] 小鼠

[0204] 创始庞贝病小鼠(Gaa敲除(-/-),C57BL/6/129背景)购自杰克逊实验室(Jackson Labs)(库存编号004154,也称为6neo小鼠)。在基因疗法计划AAALAC认可的屏障小鼠设施中使用杂合子与杂合子交配维持繁殖群,以便在同一窝内产生无效对照和WT对照。Gaa基因敲除小鼠是广泛使用的庞贝病模型。所述基因敲除小鼠表现出在心脏、中枢神经系统、骨骼肌和隔膜中溶酶体糖原的逐渐积累,并伴有活动性降低和进行性肌肉无力。小尺寸、可再现的表型和高效的繁殖允许快速研究,是临床前候选物体内筛选的最佳选择。

[0205] 将动物饲养室维持在64-79°F(18-26°C)的温度范围以及30-70%的湿度范围内。

[0206] 将动物与其父母和同窝出生仔畜一起饲养直到断奶,并且然后在转化研究实验室(Translational Research Laboratories,TRL)GTP动物饲养室中每笼二到五只动物的标准笼子中饲养。所有笼子尺寸和饲养条件均符合《实验室动物护理和使用指南(the Guide for the Care and Use of Laboratory Animals)》。笼子、水瓶和寝具基底被高压灭菌到屏障设施中。

[0207] 维持自动控制的12小时光/暗循环。每个暗周期开始于1900小时(± 30 分钟)。随意提供食物(普瑞纳公司(Purina),LabDiet®,5053,Irradiated,PicoLab®,啮齿动物饲料20、251b)。所有动物都可通过在每个饲养笼子中单独放置的水瓶来随意获取水。在每周更换笼子期间,至少每周更换一次水瓶。供水取自费城市并使用Getinge净水器进行氯化。氯化水平每天由ULAR测试并维持在2-4份每百万份(ppm)。

[0208] 将Nestlets™提供给每个饲养笼子作为富集物。

[0209] 体内研究和组织学

[0210] 通过侧尾静脉(IV)向小鼠施用0.1mL中 5×10^{11} GC(大约 2.5×10^{13} GC/kg)剂量或 5×10^{10} GC(大约 2.5×10^{12} GC/kg)剂量的AAVhu68.CAG.hGAA(各种hGAA构建体),在载体给药后第7天和第21天采血以用于血清分离,并在注射后28天进行终末采血(用于血浆分离)并通过放血执行安乐死。从脑开始,迅速收集组织。

[0211] 器官列表,尸检

[0212]	组织	快速冷冻 (用于蛋白质提取)	福尔马林浸泡 (用于组织学)
	血浆	X	
	左脑		X
	右脑	X	
	颈脊髓		X
	胸+腰脊髓	X	
	心脏	X	X
	肝脏	X	X
	隔膜	X右	X左
	三头肌	X右	X左
	四头肌	X右	X左
	腓肠肌	X右	X左
	胫前肌	X右	X左

[0213] 用于组织学的组织使用标准方法进行福尔马林固定和石蜡包埋。脑和脊髓切片用劳克坚牢蓝 (劳克坚牢蓝染色试剂盒, 艾博抗 (Abcam) ab150675) 染色, 并且外周器官用PAS (过碘酸希夫) 染色, 使用标准方法检测组织中的如糖原等多糖。对福尔马林固定的石蜡包埋样品进行针对hGAA的免疫染色。将切片去石蜡, 在10mM柠檬酸盐缓冲液 (pH 6.0) 中煮沸用于抗原修复, 用含1% 驴血清的PBS+0.2% Triton封闭15分钟, 并且然后依次与初级抗体 (Sigma HPA029126抗hGAA抗体) 和在封闭缓冲液中稀释的生物素化的二级抗体一起温育; 使用基于HRP的比色反应来检测信号。

[0214] 由通过职业验证的兽医病理学家以不知情的方式审查载玻片。建立半定量评分系统以测量如由呈现贮积和/或空泡的细胞的总百分比所确定的肌肉中庞贝病相关组织学病变 (糖原贮积和自噬积累) 的严重性:

[0215] 组织学评分贮积

0 0%

1 1%到 9%

[0216] 2 10%到 49%

3 50%到 74%

4 75%到 100%

[0217] 在适用时还估计了与载体相关的组织病理学病变。

[0218] 非人灵长类动物

[0219] 对于载体施用, 恒河猴用肌肉内右美托咪定 (dexmedetomidine) 和氯胺酮 (ketamine) 镇静并施用单次小脑延髓池内 (ICM) 注射或静脉内注射。如前所述, 使用荧光镜 (OEC9800 C-Arm, GE) 通过脊髓造影术检验用于ICM注射的针放置 (Katz N等人, 《人类基因疗法方法》. 2018年10月; 29 (5) : 212-219)。通过过量巴比妥酸盐对动物实施安乐死。在干冰上立即冷冻所收集的组织, 或将所收集的组织固定在10% 福尔马林中以用于组织学分析。

[0220] hGAA 780I酶体外性能的特征

[0221] GAA活性

[0222] 将匀浆组织的血浆或上清液与5.6mM 4-MU- α -吡喃葡萄糖苷pH 4.0混合,并在37℃下温育三小时。反应用0.4M碳酸钠,pH 11.5终止。相对荧光单位RFU使用Victor3荧光计测量,其中ex为355nm并且发射为460nm。以纳摩尔/毫升/小时为单位的活性是通过从4-MU的标准曲线内插来计算的。将单个组织样品中的活性水平针对匀浆上清液中的总蛋白质含量进行归一化。等体积用于血浆样品。

[0223] 根据LC/MS的GAA特征肽

[0224] 将血浆在100%甲醇中沉淀并离心。弃去上清液。向团粒中掺入稳定同位素标记的对hGAA特有的肽作为内标物并用胰蛋白酶重新悬浮,并在37℃下温育持续一小时。用10%甲酸停止消化。将肽通过C-18反相色谱法分离并通过ESI-质谱法进行鉴定和定量。从特征肽浓度计算血浆中的总GAA浓度。

[0225] 细胞表面受体结合测定

[0226] 将96孔板用受体涂覆、洗涤并用BSA封闭。将含有同等活性的rhGAA或经工程化的GAA的CHO培养条件培养基或血浆连续稀释三倍,以得到一系列九个递减浓度,并与共偶联受体一起温育。在温育后,在37℃下将板洗涤持续一小时以去除任何未结合的GAA和添加的4-MU- α -吡喃葡萄糖苷。用1.0M甘氨酸,pH 10.5终止反应,并通过Spectramax荧光计读取RFU;ex为370,发射为460。每个样品的RFU并通过从4-MU的标准曲线内插转换为纳摩尔/毫升/小时。非线性回归是使用GraphPad Prism完成的。

[0227] 糖原-TFA水解

[0228] 将组织匀浆在100℃下用4N TFA水解四小时,干燥并在水中重建。将经水解的材料注射到CarboPac PA-10 2×250mm柱上,以通过高pH阴离子交换色谱和脉冲安培检测(HPAEC-PAD)测定葡萄糖。每个样品中的游离葡萄糖的浓度是通过从葡萄糖标准曲线内插计算的。最终数据以 μ g糖原/mg蛋白质的形式报告。

[0229] 实例2:在庞贝病小鼠中评估rAAVhu68.hGAA载体

[0230] 将AAV载体在无菌PBS中稀释以用于IV递送到庞贝病小鼠。测试制品包含:AAVhu68.CAG.hGAAco.rBG、AAVhu68.CAG.hGAAcoV780I.rBG、AAVhu68.CAG.BiP-vIGF2.hGAAco.rBG、AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG和AAVhu68.CAG.sp7co. Δ 8.hGAAcoV780I.rBG。研究中包含野生型和媒剂对照。

[0231] 在从经处理的小鼠收集的各种组织中测量hGAA蛋白表达和活性,所述组织包含肝脏(图1A、图1B)、心脏(图2A、图2B)、四头肌(图3A、图3B)、脑(图4A、图4B)、血浆(图9A)。所有启动子在肝脏中在低剂量和高剂量下都表现得同样出色。在UbC启动子下表达的载体的施用使得在两种剂量下在骨骼肌中的活性较低,并且具有CAG启动子的载体具有最佳总体活性。具有UbC启动子的载体在两种剂量下在心脏中的活性也较低。

[0232] 庞贝病小鼠媒剂(PBS)对照(图5D)在心脏中显示出明显的糖原贮积(PAS染色切片上的深色染色)。野生型小鼠和所有经载体处理的小鼠几乎完全清除了贮积。然而,接受编码hGAA参考序列(V780)的载体的两个组表现出中度到显著的纤维化淋巴细胞性心肌炎(图5B和图5C),这在接受hGAA天然转基因的八只动物中有七只出现,并且在接受具有BiP和vIGF2修饰的经工程化的hGAA的八只动物中有三只动物出现。因为接受hGAAcoV780I酶的小鼠都没有心肌炎(图5E、图5F和图5G),所以认为这种病变与载体相关,并且更具体地,是hGAA参考序列特异性的。

[0233] 四头肌组织分析揭示了野生型小鼠和所有用编码具有或不具有另外的修饰的V780I变体的载体处理的小鼠都几乎完全清除了贮积和自噬积累(图6A-图6H)。然而,接受编码hGAAV780的参考序列的载体的两个组显示出最小到中等的糖原贮积剩余以及自噬积累(图10),这共同表现出对庞贝病的两个主要标志的次优矫正。最好的结果是从编码呈天然形式或具有BiP-vIGF2修饰的V780I变体的两种载体的递送中观察到的。sp7-delta8修饰似乎会引起对归因于庞贝病的组织学病变的不一致矫正。编码参考hGAAV780序列的两种构建体在清除糖原贮积和积累方面都是次优的。

[0234] 在高剂量IV施用($5 \times 10^{11} = 2.5 \times 10^{13}$ GC/kg)下,hGAAcoV780I和BiP-vIGF2.hGAAcoV780I在四头肌中表现出接近正常的糖原水平,并且细胞中的hGAA摄取明显更好(图7A-图7H)。对包含胫前肌(TA)和腓肠肌在内的其它骨骼肌的评估显示出类似的结果(V780I的变体并清除了糖原和中央自噬空泡)。所有构建体都减少了心脏中的糖原贮积,其中BiP-vIGF2.hGAAcoV780I施用使得水平最低。虽然四头肌中的糖原水平接近正常,但PAS染色显示出一些差异,其中hGAAcoV780I和BiP-vIGF2.hGAAcoV780I显示出最佳结果。

[0235] 在低剂量IV施用($5 \times 10^{10} = 2.5 \times 10^{12}$ GC/kg)下,与hGAAcoV780I相比,BiP-vIGF2.hGAAcoV780I在心脏和四头肌中表现出更佳的糖原减少。在利用BiP-vIGF2.hGAAcoV780I的情况下,脑和脊髓中的糖原水平接近正常(即使组织水平为约15%),这可能是由于更好的靶向性。在CNS中,观察到经工程化的构建体与V780I变体之间的有效协同作用。仅BiP-vIGF2.hGAAcoV780I清除了CNS糖原。

[0236] 如图8所示,脊髓组织学的评估表明,用AAVhu68.BiP-vIGF2.hGAAcoV780I处理的小鼠几乎完全清除了糖原贮积,而用编码参考hGAAV780酶的载体处理的小鼠具有剩余的糖原贮积。脑切片的染色也揭示了在BiP-vIGF2.hGAAcoV780I的情况下获得矫正,但在天然hGAAV780酶的情况下未获得矫正。结果证明了V780I突变和BiP-vIGF2修饰两者的贡献。

[0237] 实例3:DRG脱靶对庞贝病小鼠中hGAA表达的影响

[0238] 对BiP-vIGF2.hGAAcoV780I进行修饰以包含包装在AAVhu68衣壳中的四个mir183靶位点(BiP-vIGF2.hGAAcoV780I.4xm183,SEQ ID NO:30)(图11)。

[0239] 载体基因组含有以下序列元件:

[0240] 反向末端重复序列(ITR):ITR是源自AAV2(130bp,GenBank:NC_001401)的相同的反向互补序列,所述反向互补序列位于载体基因组的所有组件的两侧。当以反式方式提供AAV和腺病毒辅助功能时,ITR功能既充当载体DNA复制的起点,又充当载体基因组的包装信号。如此,ITR序列表示载体基因组复制和包装所需的唯一顺式序列。

[0241] CAG启动子:由巨细胞病毒(CMV)增强子、鸡 β -肌动蛋白(CB)启动子(282bp,GenBank:X00182.1)和兔 β -珠蛋白内含子组成的杂交构建体。

[0242] 编码序列:编码BiP-vIGF2.hGAAcoV780I(SEQ ID NO:31)的经工程化的cDNA(SEQ ID NO:30的nt 1141到4092)。

[0243] miR靶序列:四个串联miR-183靶序列(SEQ ID NO:26)

[0244] 兔 β -珠蛋白聚腺苷酸化信号(rBG PolyA):rBG PolyA信号(127bp,GenBank:V00882.1)促进顺式转基因mRNA的高效聚腺苷酸化。此元件充当转录终止的信号、新生转录物的3'端处的特异性切割事件以及添加长聚腺苷酸尾的信号。

[0245] 在将AAVhu68 IV递送到庞贝病小鼠后,评估将miR183靶位点引入到BiP-vIGF2-

hGAAcoV780I载体基因组中的效果。正如用BiP-vIGF2.hGAAcoV780I构建体(不具有miR183靶)所观察到的那样,在高剂量静脉内施用包含mir183靶标序列的载体后,CNS中的糖原贮积得到矫正(图12和图13)。在高剂量静脉内施用后,四头肌中的糖原贮积和自噬积累得到完全矫正,而在低剂量施用后观察到糖原贮积的矫正和自噬积累的部分矫正(图14)。在低剂量和高剂量的情况下在心脏中也观察到糖原贮积的矫正(图15)。与施用CAG.BiP-vIGF2.hGAAcoV780I所观察到的相似,自噬积累在高剂量下完全解决并且在低剂量下显著减少(图16)。结果证实,与缺乏miR靶序列的对应载体相比,添加miR183靶没有改变治疗性转基因的功效。

[0246] 实例4:在有症状后的老年庞贝病小鼠中的施用途径和剂量研究

[0247] 在静脉内(IV)和/或通过脑室内(ICV)注射施用了hGAA编码AAVhu68载体(包含例如AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG)的庞贝病小鼠(以及野生型和媒剂对照)中评估施用途径和剂量的影响。使用相同载体的双途径施用方法(静脉内和注射到脑脊液中)应矫正所述疾病的外周和神经系统表现。由于符合基因疗法条件的很大一部分患者已经具有晚期病理,所以选择处理有症状后的庞贝病小鼠(七个月大)并在处理后对所述小鼠进行至少六个月的随访。小鼠使用静脉内(IV)、脑室内(ICV)或双施用途径接受两种剂量水平(低剂量或高剂量)的载体。本研究使用的剂量(1×10^{11} 或 5×10^{10} GC ICV和 1×10^{13} GC/kg或 5×10^{13} GC/kg IV)对应于实例6中描述的李氏研究使用的低剂量和高剂量以及适合施用于人类的剂量(1×10^{13} GC/kg和 5×10^{13} GC/kg)。

[0248] 在研究过程期间,使用转棒仪、钢丝悬挂(wirehang)和握力评估测试小鼠的自发活动,并进行体积描记法。在从经处理的小鼠收集的各种组织中测量hGAA蛋白表达/活性和糖原贮积,所述组织包含血浆、四头肌、腓肠肌、隔膜和脑。进行组织学以评估例如PAS(通过劳克坚牢蓝染色)、hGAA表达和神经炎症(星形细胞增多)。对组织切片进行染色以评估自噬积累或清除(例如,使用标记LC3B的抗体)。

[0249] 下表提供了研究设计。

[0250] 天

	-7 基线	0	30	60	90
	血液收集	载体 给药	血液	血液	血液
	转棒仪		转棒仪	转棒仪	转棒仪
	钢丝悬挂		钢丝悬挂	钢丝悬挂	钢丝悬挂
	握力		握力	握力	握力
	组号	N	基因型	ROA/剂量	
[0251]	1	4M/4F	WT	ICV PBS	
	2	4M/4F	KO	ICV HD (1×10^{11} GC)	
	3	4M/4F	KO	ICV LD (5×10^{10} GC)	
	4	4M/4F	KO	IV LD (1×10^{13} GC/kg)	
	5	4M/4F	KO	IV HD (5×10^{13} GC/kg)	
	6	4M/4F	KO	ICV LD + IV LD	
	7	4M/4F	KO	ICV HD + IV HD	

[0252] 结果表明,通过全身体积描记法评估的接受中枢神经系统定向(ICV)载体的小鼠的呼吸功能因治疗而得到显著改善。庞贝病小鼠(和患者)的呼吸功能障碍被认为与支配呼

吸肌的运动神经元的贮积损伤直接相关。在经高剂量ICV处理的庞贝病小鼠中观察到呼吸功能的改善,但在经IV处理的小鼠中未观察到呼吸功能的改善(图27A和图27B)。

[0253] 对来自经高剂量和低剂量ICV处理的小鼠(图28)以及经高剂量和低剂量IV处理的小鼠(图29)的四头肌、心脏和脊髓样品进行组织学研究。通过ICV途径接受低或高载体剂量的小鼠脊髓中的糖原贮积得到矫正。高剂量IV施用可以有效矫正四头肌、心脏和脊髓中的糖原贮积。

[0254] 在用低剂量和高剂量的ICV和IV载体(双施用途径)的组合处理的雄性中,体重得到显著矫正(图25A)。单途径(单独的IV或单独的ICV)未显著矫正体重。雌性庞贝病小鼠与WT小鼠之间的体重没有差异(图25B)。

[0255] 接受高剂量IV的小鼠的握力显著改善(与基线相比并与PBS对照相比)(图26A)。低剂量的载体ICV和IV施用或双途径施用(ICV LD+IV LC)没有显著益处。然而,早在注射后第30天,高剂量IV和ICV组合的施用就将强度恢复到野生型水平,并且在第180天时所述组合的益处增加(图26B)。

[0256] 研究结果支持双施用途径更适合靶向疾病的所有方面。

[0257] 实例5:向非人灵长类动物施用DRG脱靶基因疗法载体

[0258] 进行NHP灵长类动物研究以评估毒性并评估AAVhu68衣壳中CAG.BiP-IGF2-hGAACoV780I或CAG.BiP-IGF2-hGAACoV780I-4xmir183的ICM递送。以 3×10^{13} GC/kg ICM注射载体,并在第35天处死动物。

[0259] 添加miR183的四个串联重复序列抑制了hGAA转基因在颈DRG的感觉神经元中的表达(图17)。对于miR183载体,在腰DRG的感觉神经元中也观察到hGAA转基因的表达显著降低,但仍有一些表达(图18)。令人惊讶的是,miR183的存在并未改变运动神经元中转基因的表达(图19),这表明载体的施用将有益于减少庞贝病患者的运动神经元中的糖原贮积。另外,在递送含有miR183的构建体后,心脏中的转基因表达没有降低(图20)。事实上,心脏中的表达似乎增加,这表明庞贝病患者的心脏病治疗的功效将得到提高。值得注意的是,miR183的串联重复序列降低了来自颈段和胸段的DRG的感觉神经元中的毒性(图21A和图21B)。在此剂量水平下,腰段中的毒性没有降低(图21C),这可能是由于如图18所描绘的在腰节段处的残留蛋白质表达。

[0260] 实例6:在非人灵长类动物中的施用途径研究

[0261] 进行NHP灵长类动物研究以评估毒性并评估载体施用的替代或组合途径。例如,以 5×10^{13} GC/kg(高剂量)或 1×10^{13} GC/kg(低剂量)IV注射或以 3×10^{13} GC/kg(高剂量)或 1×10^{13} GC(低剂量)ICM注射AAVhu68.CAG.BiP-IGF2-hGAACoV780I或AAVhu68.CAG.BiP-IGF2-hGAACoV780I-4xmir183。例如,通过施用指示的IV高剂量和ICM高剂量或IV低剂量和ICM低剂量来评估双施用途径的可行性和毒性。IV低剂量和ICM低剂量的组合可以揭示协同效应,这将有益于庞贝病患者的治疗。

[0262] 在整个研究中,各种读数用于检测hGAA特征肽(血浆和CSF),评估hGAA酶活性(血清和靶组织),并测量抗hGAA抗体滴度(血液和CSF)。进行组织病理学以评估靶组织的hGAA表达和毒性(例如,CNS、心脏和肌肉的H&E染色)。图31提供了示出施用途径和剂量的研究设计。

[0263] 评估单施用途径的初步研究揭示,低剂量IV注射的动物在四头肌和心脏中表达了

hGAA (图34)。IV注射的动物还表现出比ICM注射的动物更低等级的脊髓轴突病变 (图33D-图33F)。hGAA的表达也通过组织学在低剂量ICM注射的动物的脊髓中观察到 (图34)。ICM注射的动物的DRG变性和脊髓轴突病变不是剂量依赖性的 (图33A-图33F)。另外，一只IV低剂量动物 (RA3607:1e13 GC/Kg) 比IV高剂量注射的动物具有更高的DRG变性、脊髓轴突病变和更高的心脏炎性反应。

[0264] (序列表自由文本)

[0265] 对于在数字标识符<223>下含有自由文本的序列,提供了以下信息。

[0266]	SEQ ID NO: (含有自由文本)	<223>下的自由文本
	3	<223> 合成构建体 <220>

[0267]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
	<221> MISC_FEATURE <222> (1)..(27) <223> 信号肽 <220> <221> MISC_FEATURE <222> (70)..(952) <220> <221> MISC_FEATURE <222> (123)..(952) <223> 具有 V780I 的 76 kD GAA 蛋白 <220> <221> MISC_FEATURE <222> (204)..(952) <223> 具有 V780I 的 70 kD GAA 蛋白
4	<223> 经工程化的 hGAAI 编码序列
6	<223> 包括 hGAA780I 的融合蛋白
7	<223> 编码包括 GAAV780I 的融合蛋白的经工程化的序列 <220> <221> misc_feature <222> (810)..(810) <223> V810I
8	<223> CAG 启动子 <220> <221> misc_feature <222> (1)..(243) <223> CMV 早期增强子元件 <220> <221> misc_feature <222> (244)..(525) <223> 鸡 β 肌动蛋白启动子 <220> <221> misc_feature <222> (526)..(934) <223> 杂交内含子
9	<223> 兔珠蛋白 polyA
12	<223> 经工程化的 hGAAV780I 信号肽 <220> <221> sig_peptide <222> (1)..(81) <220> <221> CDS <222> (1)..(81)
13	<223> 合成构建体
14	<223> 经工程化的 hGAAV780I 成熟蛋白

[0268]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
	<220> <221> CDS <222> (1)..(2649)
15	<223> 合成构建体
16	<223> hGAA780I 123-890 的经工程化的 DNA <220> <221> CDS <222> (1)..(2304)
17	<223> 合成构建体
18	<223> 经工程化的 hGAA 70 kD cDNA <220> <221> CDS <222> (1)..(2247)
19	<223> 合成构建体
20	<223> hGAAV780I 76 kD 蛋白的经工程化的 DNA <220> <221> CDS <222> (1)..(2490)
21	<223> 合成构建体
22	<223> 合成构建体 <220> <221> CDS <222> (1)..(2952) <220> <221> misc_feature <222> (1)..(270) <223> BiP 信号肽+ vIGF2 + 2GS 延伸序列 <220> <221> misc_feature <222> (271)..(2952) <223> hGAA 61 - 952 780I 的经工程化的 DNA <220> <221> misc_feature <222> (2428)..(2430) <223> Ile 密码子
23	<223> 合成构建体
24	<223> 合成构建体 <220> <221> CDS <222> (1)..(2952)

[0269]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
	<220> <221> misc_feature <222> (1)..(270) <223> BiP-vIGF 肽 <220> <221> misc_feature <222> (1)..(270) <223> BiP 信号肽+ vIGF2 + 2GS 延伸序列 <220> <221> misc_feature <222> (271)..(2952) <223> hGAA 61-952 V780 DNA <220> <221> misc_feature <222> (2428)..(2430) <223> hGAA 780 缬氨酸的密码子
25	<223> 合成构建体
26	<223> miRNA 靶序列
27	<223> miRNA 靶序列
28	<223> 合成构建体 <220> <221> misc_feature <222> (1)..(130) <223> 5' ITR <220> <221> 增强子 <222> (195)..(437) <223> CMV IE 增强子 <220> <221> 启动子 <222> (440)..(721) <223> 鸡 β -肌动蛋白启动子 <220> <221> 内含子 <222> (721)..(1128) <223> CAG 中的杂交内含子 <220> <221> CDS <222> (1141)..(4092)

[0270]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
	<223> BiP-vIGF2-hGAAco <220> <221> misc_feature <222> (3568)..(3570) <223> Ile 密码子 <220> <221> polyA_signal <222> (4161)..(4287) <223> 兔 β -珠蛋白 polyA <220> <221> misc_feature <222> (4452)..(4581) <223> 3' ITR
29	<223> 合成构建体
30	<223> 合成构建体 <220> <221> misc_feature <222> (1)..(130) <223> 5' ITR <220> <221> 增强子 <222> (195)..(437) <223> CMV IE 增强子 <220> <221> 启动子 <222> (440)..(721) <223> 鸡 β -肌动蛋白启动子 <220> <221> 内含子 <222> (721)..(1128) <223> CAG 中的杂交内含子 <220> <221> CDS <222> (1141)..(4092) <223> BiP-vIGF2-hGAAco <220> <221> misc_feature <222> (3568)..(3570) <223> Ile 密码子

[0271]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
	<220> <221> misc_feature <222> (4113)..(4134) <223> miR-183 靶 <220> <221> misc_feature <222> (4139)..(4160) <223> miR-183 靶 <220> <221> misc_feature <222> (4167)..(4188) <223> miR-183 靶 <220> <221> misc_feature <222> (4195)..(4216) <223> miR-183 靶 <220> <221> polyA_signal <222> (4267)..(4393) <223> 兔 β -珠蛋白 polyA <220> <221> misc_feature <222> (4558)..(4687) <223> 3' ITR
31	<223> 合成构建体
32	<223> IGF2 F26S
33	<223> IGF2 Y27L
35	<223> V43L
36	<223> IGF2 F48T
37	<223> IGF2 R49S
38	<223> IGF2 S50I
39	<223> IGF2 A54R
40	<223> IGF2 L55R
41	<223> IGF2 F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55
42	<223> IGF2 Δ 1-6、Y27L、K65R
43	<223> IGF2 Δ 1-7、Y27L、K65R
44	<223> IGF2 Δ 1-4、E6R、Y27L、K65R
45	<223> IGF2 Δ 1-4、E6R、Y27L
46	<223> IGF2 E6R
48	<223> vIGF2 Δ 1-4、E6R、Y27L、K65R
20	<223> 经修饰的 BiP-1

[0272]

SEQ ID NO: (含有自由文本)	<223>下的自由文本
51	<223> 经修饰的 BiP-2
52	<223> 经修饰的 BiP-3
53	<223> 经修饰的 BiP-4
55	<223> 接头序列
57	<223> 接头序列
58	<223> 接头序列
59	<223> 接头序列
60	<223> 接头序列

[0273] 本说明书中引用的所有文献均通过引用并入本文。于2019年10月10日提交的美国临时专利申请第62/913,401号以及于2019年4月30日提交的美国临时专利申请第62/840,911号连同其序列列表以全文引用的方式并入。随此提交的名称为“19-8856PCT_ST25.txt”的序列列表及其中的序列和文本通过引用并入。尽管已经参考特定实施例描述了本发明,但是应当理解的是,可以在不脱离本发明的精神的情况下进行修改。此类修改旨在落入所附权利要求的范围内。

序列表

<110>	宾夕法尼亚大学校董会 (THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA) 爱美医疗公司 (AMICUS THERAPEUTICS, INC.)	
<120>	可用于治疗庞贝病的组合物	
<130>	19-8856PCT	
<150>	US 62/913,401	
<151>	2019-10-10	
<150>	US 62/840,911	
<151>	2019-04-30	
<160>	64	
<170>	PatentIn 版本 3.5	
<210>	1	
<211>	2211	
<212>	DNA	
<213>	腺相关病毒 hu.68 (adeno-associated virus hu.68)	
<220>		
<221>	CDS	
<222>	(1)..(2211)	
<400>	1	
[0001]	atg gct gcc gat ggt tat ctt cca gat tgg ctc gag gac aac ctc agt	48
	Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser	
	1 5 10 15	
	gaa ggc att cgc gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc	96
	Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro	
	20 25 30	
	aag gca aat caa caa cat caa gac aac gct cgg ggt ctt gtg ctt ccg	144
	Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro	
	35 40 45	
	ggt tac aaa tac ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg	192
	Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro	
	50 55 60	
	gtc aac gaa gca gac gcg gcg gcc ctc gag cac gac aag gcc tac gac	240
	Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp	
	65 70 75 80	
	cag cag ctc aag gcc gga gac aac ccg tac ctc aag tac aac cac gcc	288
	Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala	
	85 90 95	
	gac gcc gag ttc cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc	336
	Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly	
	100 105 110	
	aac ctc ggg cga gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct	384
	Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro	
	115 120 125	
	ctt ggt ctg gtt gag gaa gcg gct aag acg gct cct gga aag aag agg	432
	Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg	
	130 135 140	
	cct gta gag cag tct cct cag gaa ccg gac tcc tcc gtg ggt att ggc	480
	Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Val Gly Ile Gly	
	145 150 155 160	
	aaa tcg ggt gca cag ccc gct aaa aag aga ctc aat ttc ggt cag act	528

	Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr	
	165 170 175	
	ggc gac aca gag tca gtc ccc gac cct caa cca atc gga gaa cct ccc	576
	Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro	
	180 185 190	
	gca gcc ccc tca ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc	624
	Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly	
	195 200 205	
	gca cca gtg gca gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc	672
	Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser	
	210 215 220	
	tcg gga aat tgg cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc	720
	Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile	
	225 230 235 240	
	acc acc agc acc cga acc tgg gcc ctg ccc acc tac aac aat cac ctc	768
	Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu	
	245 250 255	
	tac aag caa atc tcc aac agc aca tct gga gga tct tca aat gac aac	816
	Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn	
	260 265 270	
	gcc tac ttc ggc tac agc acc ccc tgg ggg tat ttt gac ttc aac aga	864
	Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg	
	275 280 285	
	ttc cac tgc cac ttc tca cca cgt gac tgg caa aga ctc atc aac aac	912
	Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn	
[0002]	290 295 300	
	aac tgg gga ttc cgg cct aag cga ctc aac ttc aag ctc ttc aac att	960
	Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile	
	305 310 315 320	
	cag gtc aaa gag gtt acg gac aac aat gga gtc aag acc atc gct aat	1008
	Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn	
	325 330 335	
	aac ctt acc agc acg gtc cag gtc ttc acg gac tca gac tat cag ctc	1056
	Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu	
	340 345 350	
	ccg tac gtg ctc ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca	1104
	Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro	
	355 360 365	
	gcg gac gtt ttc atg att cct cag tac ggg tat cta acg ctt aat gat	1152
	Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp	
	370 375 380	
	gga agc caa gcc gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc	1200
	Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe	
	385 390 395 400	
	ccg tcg caa atg cta aga acg ggt aac aac ttc cag ttc agc tac gag	1248
	Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu	
	405 410 415	
	ttt gag aac gta cct ttc cat agc agc tat gct cac agc caa agc ctg	1296
	Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu	
	420 425 430	
	gac cga ctc atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	1344

[0003]	Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
	435 440 445	
	aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc agt	1392
	Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser	
	450 455 460	
	gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac ata cct	1440
	Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro	
	465 470 475 480	
	gga ccc agc tac cga caa caa cgt gtc tca acc act gtg act caa aac	1488
	Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn	
	485 490 495	
	aac aac agc gaa ttt gct tgg cct gga gct tct tct tgg gct ctc aat	1536
	Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn	
	500 505 510	
	gga cgt aat agc ttg atg aat cct gga cct gct atg gcc agc cac aaa	1584
	Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys	
	515 520 525	
	gaa gga gag gac cgt ttc ttt cct ttg tct gga tct tta att ttt ggc	1632
	Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly	
	530 535 540	
	aaa caa gga act gga aga gac aac gtg gat gcg gac aaa gtc atg ata	1680
	Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile	
	545 550 555 560	
	acc aac gaa gaa gaa att aaa act acc aac cca gta gca acg gag tcc	1728
	Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser	
	565 570 575	
	tat gga caa gtg gcc aca aac cac cag agt gcc caa gca cag gcg cag	1776
	Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln	
	580 585 590	
	acc ggc tgg gtt caa aac caa gga ata ctt ccg ggt atg gtt tgg cag	1824
	Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln	
	595 600 605	
	gac aga gat gtg tac ctg caa gga ccc att tgg gcc aaa att cct cac	1872
	Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His	
	610 615 620	
	acg gac ggc aac ttt cac cct tct ccg ctg atg gga ggg ttt gga atg	1920
	Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met	
	625 630 635 640	
	aag cac ccg cct cct cag atc ctc atc aaa aac aca cct gta cct gcg	1968
	Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala	
	645 650 655	
	gat cct cca acg gct ttc aac aag gac aag ctg aac tct ttc atc acc	2016
	Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr	
	660 665 670	
	cag tat tct act ggc caa gtc agc gtg gag att gag tgg gag ctg cag	2064
	Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln	
	675 680 685	
	aag gaa aac agc aag cgc tgg aac ccg gag atc cag tac act tcc aac	2112
	Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn	
	690 695 700	
	tat tac aag tct aat aat gtt gaa ttt gct gtt aat act gaa ggt gtt	2160

	Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val	
	705 710 715 720	
	tat tct gaa ccc cgc ccc att ggc acc aga tac ctg act cgt aat ctg	2208
	Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu	
	725 730 735	
	taa	2211
	<210> 2	
	<211> 736	
	<212> PRT	
	<213> 腺相关病毒 hu. 68 (adeno-associated virus hu. 68)	
	<400> 2	
	Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser	
	1 5 10 15	
	Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro	
	20 25 30	
	Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro	
	35 40 45	
	Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro	
	50 55 60	
	Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp	
	65 70 75 80	
	Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala	
	85 90 95	
	Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly	
	100 105 110	
[0004]	Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro	
	115 120 125	
	Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg	
	130 135 140	
	Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Val Gly Ile Gly	
	145 150 155 160	
	Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr	
	165 170 175	
	Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro	
	180 185 190	
	Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly	
	195 200 205	
	Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser	
	210 215 220	
	Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile	
	225 230 235 240	
	Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu	
	245 250 255	
	Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn	
	260 265 270	
	Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg	
	275 280 285	
	Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn	
	290 295 300	
	Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile	
	305 310 315 320	

[0005]

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
 580 585 590
 Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
 595 600 605
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 610 615 620
 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
 625 630 635 640
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
 645 650 655
 Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
 660 665 670
 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
 675 680 685
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
 690 695 700
 Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
 705 710 715 720
 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu

60

[0007]

```

Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe
225                230                235                240
Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr
                245                250                255
Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser
                260                265                270
Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly
                275                280                285
Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly
                290                295                300
Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val
305                310                315                320
Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile
                325                330                335
Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln
                340                345                350
Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly
                355                360                365
Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr
                370                375                380
Arg Gln Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val
385                390                395                400
Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe
                405                410                415
Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His
                420                425                430
Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser
                435                440                445
Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg
                450                455                460
Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val
465                470                475                480
Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu
                485                490                495
Ala Trp Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe
                500                505                510
Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly
                515                520                525
Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val
                530                535                540
Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser
545                550                555                560
Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly
                565                570                575
Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly
                580                585                590
Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg
                595                600                605
Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu
                610                615                620
Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro

```

	625	630	635	640
	Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu			
		645	650	655
	Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg			
		660	665	670
	Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser			
		675	680	685
	Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala			
		690	695	700
	Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly			
	705	710	715	720
	Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser			
		725	730	735
	Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile			
		740	745	750
	Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro			
		755	760	765
	Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly			
		770	775	780
	Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser			
	785	790	795	800
	Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val			
		805	810	815
	His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr			
		820	825	830
[0008]	Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr			
		835	840	845
	Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser			
		850	855	860
	Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala			
	865	870	875	880
	Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly			
		885	890	895
	Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala			
		900	905	910
	Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr			
		915	920	925
	Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly			
		930	935	940
	Glu Gln Phe Leu Val Ser Trp Cys			
	945	950		
	<210> 4			
	<211> 2856			
	<212> DNA			
	<213> 人工序列 (Artificial Sequence)			
	<220>			
	<223> 经工程化的 hGAAI 编码序列			
	<400> 4			
	atgggcgtta gacaccctcc ttgctctcac agactgctgg ccgtgtgtgc tctggtgtct			60
	ctggctacag ctgccctgct gggacatata ctgctgcacg actttctgct ggtgcccaga			120
	gagctgtctg gcagctctcc tgtgctggaa gagacacacc ctgcacatca gcaggcgcc			180

	tctagacctg gacctagaga cgctcaggcc catcctggca gacctagagc tgtgcccaca	240
	cagtgtgacg tgccacctaa cagcagattc gactgcgccc ctgacaaggc catcacacaa	300
	gagcagtgtg aagccagagg ctgctgctac atccctgcca aacaaggact gcaggcgct	360
	cagatgggac agccctggtg cttcttccca ccattctacc ccagctacaa gctggaaaac	420
	ctgagcagca gcgagatggg ctacaccgcc aactgacca gaaccacacc tacattcttc	480
	ccgaaggaca tcttgacact gggctggac gtgatgatgg aaaccgagaa ccggtgcac	540
	ttcaccatca aggacccgc caatcgaga tacgaggtgc cactggaaac ccctcacgtg	600
	cactctagag ccccatctcc actgtacagc gtggaatctt ccgaggaacc cttcgcgctg	660
	atcgtgcgga gacagctgga tggaagagtg ctgctgaaca ccacagtggc ccctctgttc	720
	ttcggcagc agtttctgca gctgagcacc agcctgccta gccagtatat cacaggcctg	780
	gccgagcacc tgtctccact gatgctgagc acatcctgga ccagaatcac cctgtggaac	840
	agagatctgg cccctacacc tggcgccaac ctgtacggt ctcacccctt ttatctggcc	900
	ctggaagacg gcggatctgc ccacggtgtt tttctgtga actccaacgc catggacgtg	960
	gtgtctgacg catctcctgc tctgtcttgg agaagcacag gcggcatcct ggacgtgtac	1020
	atctttctgg gccccgagcc taagagcgtg gtgcagcagt atctggacgt cgtgggctac	1080
	cccttcacgc ctcttattg gggcctgggc ttccacctgt gcaggtgggg atacagcagc	1140
	accgccatca ccagacaggt ggtggaaaac atgaccggg ctcacttccc actggacgtg	1200
	cagtggaacg acctggacta catggacagc agacgggact tcaccttcaa caaggacggc	1260
	ttcagagact tccccccat ggtgcaagaa ctgcaccaag gcggcagacg gtacatgatg	1320
	attgtggacc cagccatcag ctctagcggc cctgccgaa gctacagacc ttacgatgag	1380
	ggcctgagaa gaggcgtgtt catcaccaac gagacagcc agcctctgat cggcaaagtg	1440
	tggcctggca gcacagcctt tccagacttc acaaacccca ccgctctggc ttggtgggaa	1500
	gatatggtgg ccgagttcca cgatcaggtg cccttcgatg gcatgtggat cgacatgaac	1560
	gagcccagca acttcacccg gggcagcgag gacggctgcc ccaacaacga actggaaaat	1620
	ctctcttacg tgcccgcggt gtgcggcgga acacttcagg ccgtacaaat ctgtgccagc	1680
[0009]	agccaccagt tctcagcac ccactacaac ctgcacaacc tgtacggcct gaccgaggcc	1740
	attgcctctc atagagccct ggttaaggcc agaggcacc ggccctttgt gatcagcaga	1800
	agcacattcg ccggccacgg cagatacgcc ggacattgga caggcgacgt gtggtctagt	1860
	tgggagcagc tggctagcag cgtgccagag atcctgcagt tcaatctgct gggcgtgcca	1920
	ctcgtgggag ccgacgtttg tggcttctg ggcaacacca gcgaggaact gtgtgtgct	1980
	tggacacagc tgggcgcctt ctatcccttc atgagaaacc acaacagcct gctgagcctg	2040
	cctcaagagc cctacagctt tagcgagcct gcacagcagg ccatgagaaa ggccctgact	2100
	ctgagatacg ctctgctgcc ccacctgtac accctgttcc accaggtcca cgtggccggg	2160
	gagacagtgg ctagacctct gtctctggaa ttcccaagg acagctccac ctggaccgtg	2220
	gatcatcagc tgtgtgggg agaagccctg ctcatcacac ctgttctgca ggccggaag	2280
	gccgaagtga ccggtatatt tcctctcggc acttggtacg acctgcagac cgtgcctatt	2340
	gaggccctgg gatctcttcc tccacctcct gctgtcctta gagagcctgc catccactct	2400
	gaaggccagt ggggttacct gccgcctcct ctggacacca tcaacgtgca cctgagagct	2460
	ggctacatca tcccactgca aggccttggc ctgaccacaa ccgaatctag acagcagccc	2520
	atggctctgg ccgtggctct taaaaaggc ggagaggcta gaggcgagct gttctgggac	2580
	gacggcgagt ctctggaagt gctggaacgg ggcgcttata cccaagtgat ctctctggcc	2640
	agaaacaaca ccctcgtgaa cgaactcgtg cgcgtgacca gtgaaggtgc tggactgcaa	2700
	ctgcagaaag tgaccgtgct cggagtggcc acagctcctc aacaggtgct gtctaacggc	2760
	gtgcccgtgt ccaacttcac atacagcccc gacaccaagg tcctggacat ctgtgtgtca	2820
	ctgtgatgag gcgagcagtt cctggtgtcc tgggtg	2856
<210>	5	
<211>	2859	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	5	
	atgggagtga ggcacccgcc ctgctccac cggtcctgg ccgtctgcgc cctcgtgtcc	60

	ttggcaaccg	ctgcactcct	ggggcacatc	ctactccatg	atttcctgct	ggttccccga	120
	gagctgagtg	gctcctcccc	agtccctggag	gagactcacc	cagctcacca	gcagggagcc	180
	agtagaccag	ggccccggga	tgcccaggca	caccccggcc	gtcccagagc	agtgcacaca	240
	cagtgcgacg	tcccccccaa	cagccgcttc	gattgcgccc	ctgacaaggc	catcacccag	300
	gaacagtgcg	aggcccgcgg	ctgttgctac	atccctgcaa	agcaggggct	gcagggagcc	360
	cagatggggc	agccctgggt	cttcttccca	cccagctacc	ccagctacaa	gctggagaac	420
	ctgagctcct	ctgaaatggg	ctacacggcc	accctgaccc	gtaccacccc	caccttcttc	480
	ccaagggaca	tcctgaccct	gcggctggac	gtgatgatgg	agactgagaa	ccgcctccac	540
	ttcacgatca	aagatccagc	taacaggcgc	tacgaggtgc	ccttgagagc	cccgcattgc	600
	cacagccggg	caccgtcccc	actctacagc	gtggagtctt	ccgaggagcc	cttcgggggtg	660
	atcgtgcgcc	ggcagctgga	cggccgcgtg	ctgctgaaca	cgacggtggc	gcccctgttc	720
	tttgcggacc	agttccttca	gctgtccacc	tcgctgccct	cgcagtatat	cacaggcctc	780
	gccgagcacc	tcagtccctt	gatgctcagc	accagctgga	ccaggatcac	cctgtggaac	840
	cgggaccttg	cgcccacgcc	cgggtcgaac	ctctacgggt	ctcacccctt	ctacctggcg	900
	ctggaggacg	gcgggtcggc	acacgggggtg	ttcctgctaa	acagcaatgc	catggatgtg	960
	gtctctgcagc	cgagccctgc	ccttagctgg	aggctgcacag	gtgggatcct	ggatgtctac	1020
	atcttcctgg	gccagagacc	caagagcgtg	gtgcagcagt	acctggacgt	tgtgggatac	1080
	ccgttcattg	cgccatactg	gggcctgggc	ttccacctgt	gccgctgggg	ctactcctcc	1140
	accgctatca	cccgccaggt	ggtggagaac	atgaccaggg	cccacttccc	cctggacgtc	1200
	cagtggaaac	acctggacta	catggactcc	cggagggact	tcacgttcaa	caaggatggc	1260
	ttccgggact	tcccggccat	ggtgcaggag	ctgcaccagg	gcggccggcg	ctacatgatg	1320
	atcgtggatc	ctgccatcag	cagctcgggc	cctgccggga	gctacaggcc	ctacgacgag	1380
	ggtctgcgga	gggggggtttt	catcaccaac	gagaccggcc	agccgctgat	tgggaaggtg	1440
	tggcccgggt	ccactgcctt	ccccgacttc	accaacccca	cagccctggc	ctggtgggag	1500
	gacatggttg	ctgagttcca	tgaccagggt	cccttcgacg	gcattgtgat	tgacatgaac	1560
[0010]	gagccttcca	acttcacag	gggtctctgag	gacggctgcc	ccaacaatga	gctggagaac	1620
	ccacctacg	tgcctggggg	ggttgggggg	accctccagg	cgggccaccat	ctgtgcctcc	1680
	agccaccagt	ttctctccac	acactacaac	ctgcacaacc	tctacggcct	gaccgaagcc	1740
	atcgctctcc	acagggcgct	ggtgaaggct	cgggggacac	gcccatttgt	gatctccgc	1800
	tcgacctttg	ctggccacgg	ccgatacgcc	ggccactgga	cgggggacgt	gtggagctcc	1860
	tgggagcagc	tcgcctctc	cgtgccagaa	atcctgcagt	ttaacctgct	gggggtgcct	1920
	ctggtcgggg	ccgacgtctg	cggcttctctg	ggcaaacctt	cagaggagct	gtgtgtgcgc	1980
	tggaccacagc	tgggggcctt	ctacccttcc	atgcggaacc	acaacagcct	gctcagtctg	2040
	ccccaggagc	cgtacagctt	cagcgagccg	gcccagcagg	ccatgaggaa	ggccctcacc	2100
	ctgcgtacg	cactcctccc	ccacctctac	acactgttcc	accaggccca	cgtcgcgggg	2160
	gagaccgtgg	cccggcccc	cttcttgagg	ttccccaagg	actctagcac	ctggactgtg	2220
	gaccaccagc	tcctgtgggg	ggaggccctg	ctcatacccc	cagtgtctca	ggccgggaag	2280
	gccgaagtga	ctggctactt	ccccttgggc	acatggtacg	acctgcagac	ggtgccata	2340
	gaggcccttg	gcagcctccc	acccccacct	gcagctcccc	gtgagccagc	catccacagc	2400
	gaggggcagt	gggtgacgct	gcccggcccc	ctggacacca	tcaacgtcca	cctccgggct	2460
	gggtacatca	tccccctgca	gggccctggc	ctcacaacca	cagagtcccg	ccagcagccc	2520
	atggccctgg	ctgtggccct	gaccaagggt	ggggaggccc	gaggggagct	tttctgggac	2580
	gatggagaga	gcctggaagt	gctggagcga	ggggcctaca	cacaggtcat	cttcttgccc	2640
	aggaataaca	cgatcgtgaa	tgagctggta	cgtgtgacca	gtgaggagc	tggcctgcag	2700
	ctgcagaagg	tgactgtcct	gggcgtggcc	acggcgcccc	agcaggtcct	ctccaacggt	2760
	gtccctgtct	ccaacttcac	ctacagcccc	gacaccaagg	tcctggacat	ctgtgtctcg	2820
	ctgttgatgg	gagagcagtt	tctcgtcagc	tgggtgttag			2859
<210>	6						
<211>	982						
<212>	PRT						
<213>	人工序列 (Artificial Sequence)						

<220>

<223> 包括 hGAA780I 的融合蛋白

<400> 6

Met	Lys	Leu	Ser	Leu	Val	Ala	Ala	Met	Leu	Leu	Leu	Ser	Ala	Ala		
1				5				10				15				
Arg	Ala	Ser	Arg	Thr	Leu	Cys	Gly	Gly	Glu	Leu	Val	Asp	Thr	Leu	Gln	
			20				25					30				
Phe	Val	Cys	Gly	Asp	Arg	Gly	Phe	Leu	Phe	Ser	Arg	Pro	Ala	Ser	Arg	
		35				40					45					
Val	Ser	Arg	Arg	Ser	Arg	Gly	Ile	Val	Glu	Glu	Cys	Cys	Phe	Arg	Ser	
	50				55				60							
Cys	Asp	Leu	Ala	Leu	Leu	Glu	Thr	Tyr	Cys	Ala	Thr	Pro	Ala	Arg	Ser	
65					70				75					80		
Glu	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Arg	Pro	Gly	Pro	Arg	
			85				90						95			
Asp	Ala	Gln	Ala	His	Pro	Gly	Arg	Pro	Arg	Ala	Val	Pro	Thr	Gln	Cys	
		100				105						110				
Asp	Val	Pro	Pro	Asn	Ser	Arg	Phe	Asp	Cys	Ala	Pro	Asp	Lys	Ala	Ile	
	115				120				125							
Thr	Gln	Glu	Gln	Cys	Glu	Ala	Arg	Gly	Cys	Cys	Tyr	Ile	Pro	Ala	Lys	
	130				135				140							
Gln	Gly	Leu	Gln	Gly	Ala	Gln	Met	Gly	Gln	Pro	Trp	Cys	Phe	Phe	Pro	
145					150				155					160		
Pro	Ser	Tyr	Pro	Ser	Tyr	Lys	Leu	Glu	Asn	Leu	Ser	Ser	Ser	Glu	Met	
			165			170						175				
[0011]	Gly	Tyr	Thr	Ala	Thr	Leu	Thr	Arg	Thr	Thr	Pro	Thr	Phe	Phe	Pro	Lys
			180				185					190				
Asp	Ile	Leu	Thr	Leu	Arg	Leu	Asp	Val	Met	Met	Glu	Thr	Glu	Asn	Arg	
	195				200				205							
Leu	His	Phe	Thr	Ile	Lys	Asp	Pro	Ala	Asn	Arg	Arg	Tyr	Glu	Val	Pro	
	210				215				220							
Leu	Glu	Thr	Pro	His	Val	His	Ser	Arg	Ala	Pro	Ser	Pro	Leu	Tyr	Ser	
225				230					235					240		
Val	Glu	Phe	Ser	Glu	Glu	Pro	Phe	Gly	Val	Ile	Val	Arg	Arg	Gln	Leu	
			245			250						255				
Asp	Gly	Arg	Val	Leu	Leu	Asn	Thr	Thr	Val	Ala	Pro	Leu	Phe	Phe	Ala	
		260			265							270				
Asp	Gln	Phe	Leu	Gln	Leu	Ser	Thr	Ser	Leu	Pro	Ser	Gln	Tyr	Ile	Thr	
	275				280							285				
Gly	Leu	Ala	Glu	His	Leu	Ser	Pro	Leu	Met	Leu	Ser	Thr	Ser	Trp	Thr	
	290				295				300							
Arg	Ile	Thr	Leu	Trp	Asn	Arg	Asp	Leu	Ala	Pro	Thr	Pro	Gly	Ala	Asn	
305				310					315					320		
Leu	Tyr	Gly	Ser	His	Pro	Phe	Tyr	Leu	Ala	Leu	Glu	Asp	Gly	Gly	Ser	
			325			330						335				
Ala	His	Gly	Val	Phe	Leu	Leu	Asn	Ser	Asn	Ala	Met	Asp	Val	Val	Leu	
	340				345				350							
Gln	Pro	Ser	Pro	Ala	Leu	Ser	Trp	Arg	Ser	Thr	Gly	Gly	Ile	Leu	Asp	
	355				360				365							
Val	Tyr	Ile	Phe	Leu	Gly	Pro	Glu	Pro	Lys	Ser	Val	Val	Gln	Gln	Tyr	
	370				375				380							

[0012]

```

Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly
385                      390                      395                      400
Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln
                      405                      410                      415
Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp
                      420                      425                      430
Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys
                      435                      440                      445
Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly
                      450                      455                      460
Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly
465                      470                      475                      480
Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val
                      485                      490                      495
Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro
                      500                      505                      510
Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp
                      515                      520                      525
Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly
530                      535                      540
Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu
545                      550                      555                      560
Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly
                      565                      570                      575
Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His
                      580                      585                      590
Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr
595                      600                      605
Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg
610                      615                      620
Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala
625                      630                      635                      640
Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser
                      645                      650                      655
Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val
660                      665                      670
Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys
675                      680                      685
Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His
690                      695                      700
Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro
705                      710                      715                      720
Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu
                      725                      730                      735
Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr
740                      745                      750
Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp
755                      760                      765
Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro
770                      775                      780
Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly

```

```

785          790          795          800
Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu
          805          810          815
Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly
          820          825          830
Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu
          835          840          845
Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr
          850          855          860
Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly
865          870          875          880
Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu
          885          890          895
Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn
          900          905          910
Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly
          915          920          925
Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln
          930          935          940
Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro
945          950          955          960
Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln
          965          970          975
Phe Leu Val Ser Trp Cys
          980

```

[0013]

```

<210> 7
<211> 2952
<212> DNA
<213> 人工序列 (Artificial Sequence)
<220>
<223> 编码包括 GAAV780I 的融合蛋白的经工程化的序列
<220>
<221> misc_feature
<222> (810)..(810)
<223> V810I
<400> 7
atgaagctgt ctctggtggc tgctatgctg ctgctcctgt ctgccgccag agccagcaga      60
acactttgtg gcgagagagct ggtggacacc ctgcagtttg tgtgtggcga cagaggcttc      120
ctgttcagca gacctgccag ccgggtttcc agacggtcta gaggaatcgt ggaagagtgc      180
tgcttcagaa gctgcgatct ggccctgctg gaaacctact gtgccacacc agccagatct      240
gaaggcggcg gaggatctgg cggaggcgga tctagacctg gacctagaga cgcccaggct      300
caccctggta gacctagagc tgtgcctaca cagtgcgacg tgccacctaa cagcagattc      360
gactgcgccc ctgacaaggc catcacacaa gacagtggtg aagccagagg ctgctgctac      420
atccctgcc aacaaggact gcagggcgcc cagatgggac agccttggtg cttcttccca      480
ccatcttacc ccagctacaa gctggaaaac ctgagcagct ccgagatggg ctacaccgcc      540
acactgacca gaaccacacc tacattcttc ccgaaggaca tcctgacact gcggctggac      600
gtgatgatgg aaaccgagaa ccggctgcac ttcacatca aggacccgc caatcggaga      660
tacgaggtgc ccttgaaaac accccacgtg cactctagag caccctctcc actgtacagc      720
gtggaatttt ccgaggaacc cttcggcgtg atcgtgcgga gacagctgga tggcagagtg      780
ctctgaata ccacagtgc ccctctgttc ttgcgcgacc agtttctgca gctgagcacc      840
agcctgccta gccagtatat cacaggcctg gccgagcatc tgagccctct gatgctgagc      900

```

	acatcctgga ccagaatcac cctgtggaac cgcgacctgg ctctacacc tggcgccaat	960
	ctgtacggct ctcacccctt ctacctggca ctggaagacg gtggatctgc ccacggtgtc	1020
	tttctgctga atagcaacgc catggacgtg gtgtgcagc cctctctgc actgtcttgg	1080
	agatctacag gcggcatcct ggacgtgtac atctttctgg gccccgagcc taagagcgtg	1140
	gtgcagcagt atctggacgt cgtgggctac ccttcatgc ctcttattg gggcctgggc	1200
	ttccacctgt gtaggtgggg ctacagcagc accgccatca ccagacaggt ggtggaaaac	1260
	atgacccggg ctacttccc actggacgtg cagtggaaacg acctggacta catggacagc	1320
	agacgggact tcacattcaa caaggacggc ttcagagact tccccgcat ggtgcaagag	1380
	ctgcatcaag gcggacggcg gtacatgatg attgtggacc ctgccatcag cagctctgga	1440
	ccagccggca gctacagacc ttacgatgag ggactgagaa gaggcgtgtt catcaccaac	1500
	gagacaggcc agcctctgat cggcaaagtg tggcctggca gcacagcctt tccagacttc	1560
	acaaacccca ccgctctggc ttggtgggaa gatatgttg ccgagttcca cgatcaggtg	1620
	cccttcgatg gcatgtggat cgacatgaac gagcccagca acttcatccg gggcagcgag	1680
	gatggctgcc ccaacaacga actagaaaat cctccttacg tgcccgccgt tgtcggcgga	1740
	acatttcagg ccgctacaat ctgtgccagc agccatcagt ttctgagcac ccactacaac	1800
	ctgcacaacc tgtacggcct gaccgaggcc attgcctctc atagagccct ggtaaggcc	1860
	agaggcacc ccgcttttgt gatcagcaga agcacattcg ccggccacgg cagatacgca	1920
	ggacattgga caggcgacgt gtggtctagt tgggagcagc tggctagcag cgtgccagag	1980
	atcctgcagt tcaatctgct gggcgtgcca ctctgtggag ccgacgtttg tgcttctctg	2040
	ggcaacacca gcgaggaact gtgtgtgctg tggacacagc tgggcgcctt ctatcccttc	2100
	atgagaaacc acaacagcct gctgagcctg cctcaagagc cctacagctt tagcgagcct	2160
	gcacagcagg ccatgagaaa ggccctgact ctgagatacg ccctgctgcc tcacctgtac	2220
	acctgtttc atcaggccca cgtggcaggc gagacagtgg ctagacctct gttcctggaa	2280
	tttcccaagg acagctccac ctggaccgtg gatcatcagc tgctgtgggg agaagccctg	2340
	ctgattacac cagtgtctga ggccggaaag gccgaagtga caggctattt ccctctcggc	2400
[0014]	acttggtagc acctgcagac cgtgcctatc gaggtctctg gatctcttcc tccacctcct	2460
	gccgtctcta gagagcctgc cattcactct gaaggccagt gggttaccct gcctgtcct	2520
	ctggacacca tcaacgtgca cctgagagcc ggctacatca tccctctgca aggccctggc	2580
	ctgaccacaa ccgaatctag acagcagccc atggcactgg ccgtggctct tacaaaaggc	2640
	ggagaggcta gaggcgagct gttctgggat gatggcgaga gcctagaagt gctggaacgg	2700
	ggcgcttata cccaagtgat cttcctggcc agaaacaaca ccatcgtgaa cgaactcgtg	2760
	cgcgtgacca gtgaaggtgc tggactgcaa ctgcagaaag tgaccgtgct cggagtggcc	2820
	acagctcttc agcaggttct gtctaattgc gtgcccggtt ccaacttcac atacagcccc	2880
	gacaccaagg tcttgacat ctgtgtgtcc ctgcttatgg gcgagcagtt cctggtgtcc	2940
	tggtgctgat aa	2952
	<210> 8	
	<211> 934	
	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> CAG 启动子	
	<220>	
	<221> misc_feature	
	<222> (1)..(243)	
	<223> CMV 早期增强子元件	
	<220>	
	<221> misc_feature	
	<222> (244)..(525)	
	<223> 鸡 β 肌动蛋白启动子	
	<220>	
	<221> misc_feature	

	<222> (526).. (934)	
	<223> CAG 中的杂交内含子 (b-肌动蛋白内含子和兔 β 1-珠蛋白)	
	<220>	
	<221> misc_feature	
	<222> (526).. (934)	
	<223> 杂交内含子	
	<400> 8	
	attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt tccattgacg	60
	tcaatgggtg gagtattttac ggtaaactgc ccacttggca gtacatcaag tgtatcatat	120
	gccaaagtacg ccccttattg acgtcaatga cggtaaattgg ccgcctggc attatgccca	180
	gtacatgacc ttatgggact ttctacttg gcagtacatc tacgtattag tcatcgctat	240
	taccatggtc gaggtgagcc ccacgttctg cttcactctc cccatctccc cccctctccc	300
	accccccaatt ttgtattttat ttatttttta attattttgt gcagcgatgg gggcgggggg	360
	gggggggggg cgcgccgagc gcggggcggg gcggggcgag gggcggggcg gggcgaggcg	420
	gagaggtgctg gcggcagcca atcagagcgg cgcgctccga aagtttcctt ttatggcgag	480
	gcggcgggcg cggcggccct ataaaaagcg aagcgcgcg gcggcgggag tcgctgcgcg	540
	ctgccttcgc cccgtgcccc gctccgcccgc cgctcgcgc cgcccgcccc ggctctgact	600
	gaccgcgtta ctccacagc tgagcgggag ggacggccct tctctccgg gctgtaatta	660
	gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc ttgaggggct	720
	ccgggagggc cctttgtgag gggggagcgg ctcggggctg tccgcggggg gacggctgcc	780
	ttcggggggg acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag	840
	cctctgctaa ccatgttcat gccttcttct ttttctaca gctcctgggc aacgtgctgg	900
	ttattgtgct gtctcatcat ttggcaaaag aatt	934
	<210> 9	
	<211> 127	
[0015]	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 兔珠蛋白 polyA	
	<400> 9	
	gatctttttc cctctgcaa aaattatggg gacatcatga agccccttga gcatctgact	60
	tctggctaataaagaaatt tattttcatt gcaatagtgt gttggaattt ttgtgtctc	120
	tcactcg	127
	<210> 10	
	<211> 188	
	<212> DNA	
	<213> 腺相关病毒 2 (adeno-associated virus 2)	
	<400> 10	
	gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt	60
	tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac	120
	taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacgt agccatgctc	180
	taggaaga	188
	<210> 11	
	<211> 188	
	<212> DNA	
	<213> 腺相关病毒 2 (adeno-associated virus 2)	
	<400> 11	
	tcttctaga gcatggctac gtagataagt agcatggcgg gttaatcatt aactacaagg	60
	aaccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg	120
	ggcgacaaa ggtcgccga cgcccgggct ttgccgggc ggcctcagt agcgagcgag	180
	cgcgcagc	188

	<210>	12	
	<211>	81	
	<212>	DNA	
	<213>	人工序列 (Artificial Sequence)	
	<220>		
	<223>	经工程化的 hGAAV780I 信号肽	
	<220>		
	<221>	sig_peptide	
	<222>	(1).. (81)	
	<220>		
	<221>	CDS	
	<222>	(1).. (81)	
	<400>	12	
		atg ggc gtt aga cac cct cct tgc tct cac aga ctg ctg gcc gtg tgt	48
		Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys	
		1 5 10 15	
		gct ctg gtg tct ctg gct aca gct gcc ctg ctg	81
		Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu	
		20 25	
	<210>	13	
	<211>	27	
	<212>	PRT	
	<213>	人工序列 (Artificial Sequence)	
	<220>		
	<223>	合成构建体	
[0016]	<400>	13	
		Met Gly Val Arg His Pro Pro Cys Ser His Arg Leu Leu Ala Val Cys	
		1 5 10 15	
		Ala Leu Val Ser Leu Ala Thr Ala Ala Leu Leu	
		20 25	
	<210>	14	
	<211>	2649	
	<212>	DNA	
	<213>	人工序列 (Artificial Sequence)	
	<220>		
	<223>	经工程化的 hGAAV780I 成熟蛋白	
	<220>		
	<221>	CDS	
	<222>	(1).. (2649)	
	<400>	14	
		gcc cat cct ggc aga cct aga gct gtg ccc aca cag tgt gac gtg cca	48
		Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro	
		1 5 10 15	
		cct aac agc aga ttc gac tgc gcc cct gac aag gcc atc aca caa gag	96
		Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu	
		20 25 30	
		cag tgt gaa gcc aga ggc tgc tgc tac atc cct gcc aaa caa gga ctg	144
		Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu	
		35 40 45	
		cag ggc gct cag atg gga cag ccc tgg tgc ttc ttc cca cca tct tac	192
		Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr	

	50	55	60	
	ccc agc tac aag ctg gaa aac ctg agc agc agc gag atg ggc tac acc			240
	Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr			
	65	70	75	80
	gcc aca ctg acc aga acc aca cct aca ttc ttc ccg aag gac atc ctg			288
	Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu			
	85	90	95	
	aca ctg cgg ctg gac gtg atg atg gaa acc gag aac cgg ctg cac ttc			336
	Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe			
	100	105	110	
	acc atc aag gac ccc gcc aat cgg aga tac gag gtg cca ctg gaa acc			384
	Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr			
	115	120	125	
	cct cac gtg cac tct aga gcc cca tct cca ctg tac agc gtg gaa ttt			432
	Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe			
	130	135	140	
	tcc gag gaa ccc ttc ggc gtg atc gtg cgg aga cag ctg gat gga aga			480
	Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg			
	145	150	155	160
	gtg ctg ctg aac acc aca gtg gcc cct ctg ttc ttc gcc gac cag ttt			528
	Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe			
	165	170	175	
	ctg cag ctg agc acc agc ctg cct agc cag tat atc aca ggc ctg gcc			576
	Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala			
	180	185	190	
[0017]	gag cac ctg tct cca ctg atg ctg agc aca tcc tgg acc aga atc acc			624
	Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr			
	195	200	205	
	ctg tgg aac aga gat ctg gcc cct aca cct ggc gcc aac ctg tac ggc			672
	Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly			
	210	215	220	
	tct cac ccc ttt tat ctg gcc ctg gaa gac ggc gga tct gcc cac ggt			720
	Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly			
	225	230	235	240
	gtt ttt ctg ctg aac tcc aac gcc atg gac gtg gtg ctg cag cca tct			768
	Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser			
	245	250	255	
	cct gct ctg tct tgg aga agc aca ggc ggc atc ctg gac gtg tac atc			816
	Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile			
	260	265	270	
	ttt ctg ggc ccc gag cct aag agc gtg gtg cag cag tat ctg gac gtc			864
	Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val			
	275	280	285	
	gtg ggc tac ccc ttc atg cct cct tat tgg ggc ctg ggc ttc cac ctg			912
	Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu			
	290	295	300	
	tgc agg tgg gga tac agc agc acc gcc atc acc aga cag gtg gtg gaa			960
	Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu			
	305	310	315	320
	aac atg acc cgg gct cac ttc cca ctg gac gtg cag tgg aac gac ctg			1008
	Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu			

	325	330	335	
	gac tac atg gac agc aga cgg gac ttc acc ttc aac aag gac ggc ttc			1056
	Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe			
	340	345	350	
	aga gac ttc ccc gcc atg gtg caa gaa ctg cac caa ggc ggc aga cgg			1104
	Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg			
	355	360	365	
	tac atg atg att gtg gac cca gcc atc agc tct agc ggc cct gcc gga			1152
	Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly			
	370	375	380	
	agc tac aga cct tac gat gag ggc ctg aga aga ggc gtg ttc atc acc			1200
	Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr			
	385	390	395	400
	aac gag aca ggc cag cct ctg atc ggc aaa gtg tgg cct ggc agc aca			1248
	Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr			
	405	410	415	
	gcc ttt cca gac ttc aca aac ccc acc gct ctg gct tgg tgg gaa gat			1296
	Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp			
	420	425	430	
	atg gtg gcc gag ttc cac gat cag gtg ccc ttc gat ggc atg tgg atc			1344
	Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile			
	435	440	445	
	gac atg aac gag ccc agc aac ttc atc cgg ggc agc gag gac ggc tgc			1392
	Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys			
	450	455	460	
[0018]	ccc aac aac gaa ctg gaa aat cct cct tac gtg ccc ggc gtt gtc ggc			1440
	Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly			
	465	470	475	480
	gga aca ctt cag gcc gct aca atc tgt gcc agc agc cac cag ttc ctc			1488
	Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu			
	485	490	495	
	agc acc cac tac aac ctg cac aac ctg tac ggc ctg acc gag gcc att			1536
	Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile			
	500	505	510	
	gcc tct cat aga gcc ctg gtt aag gcc aga ggc acc cgg cct ttt gtg			1584
	Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val			
	515	520	525	
	atc agc aga agc aca ttc gcc ggc cac ggc aga tac gcc gga cat tgg			1632
	Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp			
	530	535	540	
	aca ggc gac gtg tgg tct agt tgg gag cag ctg gct agc agc gtg cca			1680
	Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro			
	545	550	555	560
	gag atc ctg cag ttc aat ctg ctg ggc gtg cca ctc gtg gga gcc gac			1728
	Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp			
	565	570	575	
	gtt tgt ggc ttc ctg ggc aac acc agc gag gaa ctg tgt gtg cgt tgg			1776
	Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp			
	580	585	590	
	aca cag ctg ggc gcc ttc tat ccc ttc atg aga aac cac aac agc ctg			1824
	Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu			

	595	600	605	
	ctg agc ctg cct caa gag ccc tac agc ttt agc gag cct gca cag cag			1872
	Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln			
	610	615	620	
	gcc atg aga aag gcc ctg act ctg aga tac gct ctg ctg ccc cac ctg			1920
	Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu			
	625	630	635	640
	tac acc ctg ttt cac cag gct cac gtg gcc ggg gag aca gtg gct aga			1968
	Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg			
	645	650	655	
	cct ctg ttc ctg gaa ttt ccc aag gac agc tcc acc tgg acc gtg gat			2016
	Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp			
	660	665	670	
	cat cag ctg ctg tgg gga gaa gcc ctg ctc atc aca cct gtt ctg cag			2064
	His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln			
	675	680	685	
	gcc gga aag gcc gaa gtg acc ggc tat ttt cct ctc ggc act tgg tac			2112
	Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr			
	690	695	700	
	gac ctg cag acc gtg cct att gag gcc ctg gga tct ctt cct cca cct			2160
	Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro			
	705	710	715	720
	cct gct gct cct aga gag cct gcc atc cac tct gaa ggc cag tgg gtt			2208
	Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val			
	725	730	735	
[0019]	aca ctg ccc gct cct ctg gac acc atc aac gtg cac ctg aga gct ggc			2256
	Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly			
	740	745	750	
	tac atc atc cca ctg caa ggc cct ggc ctg acc aca acc gaa tct aga			2304
	Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg			
	755	760	765	
	cag cag ccc atg gct ctg gcc gtg gct ctt aca aaa ggc gga gag gct			2352
	Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala			
	770	775	780	
	aga ggc gag ctg ttc tgg gac gac ggc gag tct ctg gaa gtg ctg gaa			2400
	Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu			
	785	790	795	800
	cgg ggc gct tat acc caa gtg atc ttc ctg gcc aga aac aac acc atc			2448
	Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile			
	805	810	815	
	gtg aac gaa ctc gtg cgc gtg acc agt gaa ggt gct gga ctg caa ctg			2496
	Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu			
	820	825	830	
	cag aaa gtg acc gtg ctc gga gtg gcc aca gct cct caa cag gtg ctg			2544
	Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu			
	835	840	845	
	tct aac ggc gtg ccc gtg tcc aac ttc aca tac agc ccc gac acc aag			2592
	Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys			
	850	855	860	
	gtc ctg gac atc tgt gtg tca ctg ctg atg ggc gag cag ttc ctg gtg			2640
	Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val			

	865	870	875	880	2649
	tcc tgg tgc				
	Ser Trp Cys				
	<210> 15				
	<211> 883				
	<212> PRT				
	<213> 人工序列 (Artificial Sequence)				
	<220>				
	<223> 合成构建体				
	<400> 15				
	Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro				
	1 5 10 15				
	Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu				
	20 25 30				
	Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu				
	35 40 45				
	Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr				
	50 55 60				
	Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr				
	65 70 75 80				
	Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu				
	85 90 95				
	Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe				
	100 105 110				
[0020]	Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr				
	115 120 125				
	Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe				
	130 135 140				
	Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg				
	145 150 155 160				
	Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe				
	165 170 175				
	Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala				
	180 185 190				
	Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr				
	195 200 205				
	Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly				
	210 215 220				
	Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly				
	225 230 235 240				
	Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser				
	245 250 255				
	Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile				
	260 265 270				
	Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val				
	275 280 285				
	Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu				
	290 295 300				
	Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu				
	305 310 315 320				

[0021]

```

Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu
      325                      330                      335
Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe
      340                      345                      350
Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg
      355                      360                      365
Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly
      370                      375                      380
Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr
      385                      390                      395                      400
Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr
      405                      410                      415
Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp
      420                      425                      430
Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile
      435                      440                      445
Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys
      450                      455                      460
Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly
      465                      470                      475                      480
Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu
      485                      490                      495
Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile
      500                      505                      510
Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val
      515                      520                      525
Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp
      530                      535                      540
Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro
      545                      550                      555                      560
Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp
      565                      570                      575
Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp
      580                      585                      590
Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu
      595                      600                      605
Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln
      610                      615                      620
Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu
      625                      630                      635                      640
Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg
      645                      650                      655
Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp
      660                      665                      670
His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln
      675                      680                      685
Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr
      690                      695                      700
Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro
      705                      710                      715                      720
Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val

```

```

              725              730              735
Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly
              740              745              750
Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg
              755              760              765
Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala
              770              775              780
Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu
785              790              795              800
Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile
              805              810              815
Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu
              820              825              830
Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu
              835              840              845
Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys
              850              855              860
Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val
865              870              875              880
Ser Trp Cys

```

<210> 16

<211> 2304

<212> DNA

<213> 人工序列 (Artificial Sequence)

[0022]

<220>

<223> hGAA780I 123-890 的经工程化的 DNA

<220>

<221> CDS

<222> (1)..(2304)

<400> 16

```

gga cag ccc tgg tgc ttc ttc cca cca tct tac ccc agc tac aag ctg      48
Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu
1              5              10              15
gaa aac ctg agc agc agc gag atg ggc tac acc gcc aca ctg acc aga      96
Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg
              20              25              30
acc aca cct aca ttc ttc ccg aag gac atc ctg aca ctg cgg ctg gac      144
Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp
              35              40              45
gtg atg atg gaa acc gag aac cgg ctg cac ttc acc atc aag gac ccc      192
Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile Lys Asp Pro
              50              55              60
gcc aat cgg aga tac gag gtg cca ctg gaa acc cct cac gtg cac tct      240
Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro His Val His Ser
65              70              75              80
aga gcc cca tct cca ctg tac agc gtg gaa ttt tcc gag gaa ccc ttc      288
Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe
              85              90              95
ggc gtg atc gtg cgg aga cag ctg gat gga aga gtg ctg ctg aac acc      336
Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val Leu Leu Asn Thr

```


	100	105	110	
	aca gtg gcc cct ctg ttc ttc gcc gac cag ttt ctg cag ctg agc acc			384
	Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr			
	115	120	125	
	agc ctg cct agc cag tat atc aca ggc ctg gcc gag cac ctg tct cca			432
	Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His Leu Ser Pro			
	130	135	140	
	ctg atg ctg agc aca tcc tgg acc aga atc acc ctg tgg aac aga gat			480
	Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp			
	145	150	155	160
	ctg gcc cct aca cct ggc gcc aac ctg tac ggc tct cac ccc ttt tat			528
	Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr			
	165	170	175	
	ctg gcc ctg gaa gac ggc gga tct gcc cac ggt gtt ttt ctg ctg aac			576
	Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe Leu Leu Asn			
	180	185	190	
	tcc aac gcc atg gac gtg gtg ctg cag cca tct cct gct ctg tct tgg			624
	Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala Leu Ser Trp			
	195	200	205	
	aga agc aca ggc ggc atc ctg gac gtg tac atc ttt ctg ggc ccc gag			672
	Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu			
	210	215	220	
	cct aag agc gtg gtg cag cag tat ctg gac gtc gtg ggc tac ccc ttc			720
	Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe			
	225	230	235	240
[0023]	atg cct cct tat tgg ggc ctg ggc ttc cac ctg tgc agg tgg gga tac			768
	Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg Trp Gly Tyr			
	245	250	255	
	agc agc acc gcc atc acc aga cag gtg gtg gaa aac atg acc cgg gct			816
	Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met Thr Arg Ala			
	260	265	270	
	cac ttc cca ctg gac gtg cag tgg aac gac ctg gac tac atg gac agc			864
	His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser			
	275	280	285	
	aga cgg gac ttc acc ttc aac aag gac ggc ttc aga gac ttc ccc gcc			912
	Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala			
	290	295	300	
	atg gtg caa gaa ctg cac caa ggc ggc aga cgg tac atg atg att gtg			960
	Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met Met Ile Val			
	305	310	315	320
	gac cca gcc atc agc tct agc ggc cct gcc gga agc tac aga cct tac			1008
	Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr			
	325	330	335	
	gat gag ggc ctg aga aga ggc gtg ttc atc acc aac gag aca ggc cag			1056
	Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu Thr Gly Gln			
	340	345	350	
	cct ctg atc ggc aaa gtg tgg cct ggc agc aca gcc ttt cca gac ttc			1104
	Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe			
	355	360	365	
	aca aac ccc acc gct ctg gct tgg tgg gaa gat atg gtg gcc gag ttc			1152
	Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val Ala Glu Phe			

	370	375	380	
	cac gat cag gtg ccc ttc gat ggc atg tgg atc gac atg aac gag ccc			1200
	His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met Asn Glu Pro			
	385	390	395	400
	agc aac ttc atc cgg ggc agc gag gac ggc tgc ccc aac aac gaa ctg			1248
	Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu			
	405	410	415	
	gaa aat cct cct tac gtg ccc ggc gtt gtc ggc gga aca ctt cag gcc			1296
	Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr Leu Gln Ala			
	420	425	430	
	gct aca atc tgt gcc agc agc cac cag ttc ctc agc acc cac tac aac			1344
	Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr His Tyr Asn			
	435	440	445	
	ctg cac aac ctg tac ggc ctg acc gag gcc att gcc tct cat aga gcc			1392
	Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser His Arg Ala			
	450	455	460	
	ctg gtt aag gcc aga ggc acc cgg cct ttt gtg atc agc aga agc aca			1440
	Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser Arg Ser Thr			
	465	470	475	480
	ttc gcc ggc cac ggc aga tac gcc gga cat tgg aca ggc gac gtg tgg			1488
	Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly Asp Val Trp			
	485	490	495	
	tct agt tgg gag cag ctg gct agc agc gtg cca gag atc ctg cag ttc			1536
	Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe			
	500	505	510	
[0024]	aat ctg ctg ggc gtg cca ctc gtg gga gcc gac gtt tgt ggc ttc ctg			1584
	Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu			
	515	520	525	
	ggc aac acc agc gag gaa ctg tgt gtg cgt tgg aca cag ctg ggc gcc			1632
	Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala			
	530	535	540	
	ttc tat ccc ttc atg aga aac cac aac agc ctg ctg agc ctg cct caa			1680
	Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln			
	545	550	555	560
	gag ccc tac agc ttt agc gag cct gca cag cag gcc atg aga aag gcc			1728
	Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala			
	565	570	575	
	ctg act ctg aga tac gct ctg ctg ccc cac ctg tac acc ctg ttt cac			1776
	Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His			
	580	585	590	
	cag gct cac gtg gcc ggg gag aca gtg gct aga cct ctg ttc ctg gaa			1824
	Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu			
	595	600	605	
	ttt ccc aag gac agc tcc acc tgg acc gtg gat cat cag ctg ctg tgg			1872
	Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp			
	610	615	620	
	gga gaa gcc ctg ctc atc aca cct gtt ctg cag gcc gga aag gcc gaa			1920
	Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu			
	625	630	635	640
	gtg acc ggc tat ttt cct ctc ggc act tgg tac gac ctg cag acc gtg			1968
	Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val			

	645	650	655	
	cct att gag gcc ctg gga tct ctt cct cca cct cct gct gct cct aga			2016
	Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Ala Ala Pro Arg			
	660	665	670	
	gag cct gcc atc cac tct gaa ggc cag tgg gtt aca ctg ccc gct cct			2064
	Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro			
	675	680	685	
	ctg gac acc atc aac gtg cac ctg aga gct ggc tac atc atc cca ctg			2112
	Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu			
	690	695	700	
	caa ggc cct ggc ctg acc aca acc gaa tct aga cag cag ccc atg gct			2160
	Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala			
	705	710	715	720
	ctg gcc gtg gct ctt aca aaa ggc gga gag gct aga ggc gag ctg ttc			2208
	Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe			
	725	730	735	
	tgg gac gac ggc gag tct ctg gaa gtg ctg gaa cgg ggc gct tat acc			2256
	Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr			
	740	745	750	
	caa gtg atc ttc ctg gcc aga aac aac acc atc gtg aac gaa ctc gtg			2304
	Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val			
	755	760	765	
<210>	17			
<211>	768			
<212>	PRT			
[0025] <213>	人工序列 (Artificial Sequence)			
<220>				
<223>	合成构建体			
<400>	17			
	Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu			
	1 5 10 15			
	Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg			
	20 25 30			
	Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp			
	35 40 45			
	Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile Lys Asp Pro			
	50 55 60			
	Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro His Val His Ser			
	65 70 75 80			
	Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe			
	85 90 95			
	Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val Leu Leu Asn Thr			
	100 105 110			
	Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr			
	115 120 125			
	Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His Leu Ser Pro			
	130 135 140			
	Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp			
	145 150 155 160			
	Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr			
	165 170 175			

[0026]

```

Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe Leu Leu Asn
    180                      185                      190
Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala Leu Ser Trp
    195                      200                      205
Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu
    210                      215                      220
Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe
    225                      230                      235                      240
Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg Trp Gly Tyr
    245                      250                      255
Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met Thr Arg Ala
    260                      265                      270
His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser
    275                      280                      285
Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala
    290                      295                      300
Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met Met Ile Val
    305                      310                      315                      320
Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr
    325                      330                      335
Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu Thr Gly Gln
    340                      345                      350
Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe
    355                      360                      365
Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val Ala Glu Phe
    370                      375                      380
His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met Asn Glu Pro
    385                      390                      395                      400
Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu
    405                      410                      415
Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr Leu Gln Ala
    420                      425                      430
Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr His Tyr Asn
    435                      440                      445
Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser His Arg Ala
    450                      455                      460
Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser Arg Ser Thr
    465                      470                      475                      480
Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly Asp Val Trp
    485                      490                      495
Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe
    500                      505                      510
Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu
    515                      520                      525
Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala
    530                      535                      540
Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln
    545                      550                      555                      560
Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala
    565                      570                      575
Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His

```

	580	585	590
	Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu		
	595	600	605
	Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp		
	610	615	620
	Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu		
	625	630	635
	Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val		
	645	650	655
	Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg		
	660	665	670
	Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro		
	675	680	685
	Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu		
	690	695	700
	Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala		
	705	710	715
	Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe		
	725	730	735
	Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr		
	740	745	750
	Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val		
	755	760	765
<210>	18		
<211>	2247		
<212>	DNA		
<213>	人工序列 (Artificial Sequence)		
<220>			
<223>	经工程化的 hGAA 70 kD cDNA		
<220>			
<221>	CDS		
<222>	(1).. (2247)		
<400>	18		
	gcc cca tct cca ctg tac agc gtg gaa ttt tcc gag gaa ccc ttc ggc	48	
	Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly		
	1 5 10 15		
	gtg atc gtg cgg aga cag ctg gat gga aga gtg ctg ctg aac acc aca	96	
	Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr		
	20 25 30		
	gtg gcc cct ctg ttc ttc gcc gac cag ttt ctg cag ctg agc acc agc	144	
	Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser		
	35 40 45		
	ctg cct agc cag tat atc aca ggc ctg gcc gag cac ctg tct cca ctg	192	
	Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu		
	50 55 60		
	atg ctg agc aca tcc tgg acc aga atc acc ctg tgg aac aga gat ctg	240	
	Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu		
	65 70 75 80		
	gcc cct aca cct ggc gcc aac ctg tac ggc tct cac ccc ttt tat ctg	288	
	Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu		
	85 90 95		

[0027]

[0028]	gcc ctg gaa gac ggc gga tct gcc cac ggt gtt ttt ctg ctg aac tcc	336
	Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser	
	100 105 110	
	aac gcc atg gac gtg gtg ctg cag cca tct cct gct ctg tct tgg aga	384
	Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg	
	115 120 125	
	agc aca ggc ggc atc ctg gac gtg tac atc ttt ctg ggc ccc gag cct	432
	Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro	
	130 135 140	
	aag agc gtg gtg cag cag tat ctg gac gtc gtg ggc tac ccc ttc atg	480
	Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met	
	145 150 155 160	
	cct cct tat tgg ggc ctg ggc ttc cac ctg tgc agg tgg gga tac agc	528
	Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser	
	165 170 175	
	agc acc gcc atc acc aga cag gtg gtg gaa aac atg acc cgg gct cac	576
	Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met Thr Arg Ala His	
	180 185 190	
	ttc cca ctg gac gtg cag tgg aac gac ctg gac tac atg gac agc aga	624
	Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser Arg	
	195 200 205	
	cgg gac ttc acc ttc aac aag gac ggc ttc aga gac ttc ccc gcc atg	672
	Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala Met	
	210 215 220	
	gtg caa gaa ctg cac caa ggc ggc aga cgg tac atg atg att gtg gac	720
	Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met Met Ile Val Asp	
	225 230 235 240	
	cca gcc atc agc tct agc ggc cct gcc gga agc tac aga cct tac gat	768
	Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr Asp	
	245 250 255	
	gag ggc ctg aga aga ggc gtg ttc atc acc aac gag aca ggc cag cct	816
	Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu Thr Gly Gln Pro	
	260 265 270	
	ctg atc ggc aaa gtg tgg cct ggc agc aca gcc ttt cca gac ttc aca	864
	Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe Thr	
	275 280 285	
	aac ccc acc gct ctg gct tgg tgg gaa gat atg gtg gcc gag ttc cac	912
	Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val Ala Glu Phe His	
	290 295 300	
	gat cag gtg ccc ttc gat ggc atg tgg atc gac atg aac gag ccc agc	960
	Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met Asn Glu Pro Ser	
	305 310 315 320	
	aac ttc atc cgg ggc agc gag gac ggc tgc ccc aac aac gaa ctg gaa	1008
	Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu Glu	
	325 330 335	
	aat cct cct tac gtg ccc ggc gtt gtc ggc gga aca ctt cag gcc gct	1056
	Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr Leu Gln Ala Ala	
	340 345 350	
	aca atc tgt gcc agc agc cac cag ttc ctc agc acc cac tac aac ctg	1104
	Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr His Tyr Asn Leu	
	355 360 365	

	cac aac ctg tac ggc ctg acc gag gcc att gcc tct cat aga gcc ctg His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu 370 375 380	1152
	ggt aag gcc aga ggc acc cgg cct ttt gtg atc agc aga agc aca ttc Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe 385 390 395 400	1200
	gcc ggc cac ggc aga tac gcc gga cat tgg aca ggc gac gtg tgg tct Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser 405 410 415	1248
	agt tgg gag cag ctg gct agc agc gtg cca gag atc ctg cag ttc aat Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn 420 425 430	1296
	ctg ctg ggc gtg cca ctc gtg gga gcc gac gtt tgt ggc ttc ctg ggc Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly 435 440 445	1344
	aac acc agc gag gaa ctg tgt gtg cgt tgg aca cag ctg ggc gcc ttc Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe 450 455 460	1392
	tat ccc ttc atg aga aac cac aac agc ctg ctg agc ctg cct caa gag Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu 465 470 475 480	1440
	ccc tac agc ttt agc gag cct gca cag cag gcc atg aga aag gcc ctg Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu 485 490 495	1488
[0029]	act ctg aga tac gct ctg ctg ccc cac ctg tac acc ctg ttt cac cag Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln 500 505 510	1536
	gct cac gtg gcc ggg gag aca gtg gct aga cct ctg ttc ctg gaa ttt Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe 515 520 525	1584
	ccc aag gac agc tcc acc tgg acc gtg gat cat cag ctg ctg tgg gga Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly 530 535 540	1632
	gaa gcc ctg ctc atc aca cct gtt ctg cag gcc gga aag gcc gaa gtg Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val 545 550 555 560	1680
	acc ggc tat ttt cct ctc ggc act tgg tac gac ctg cag acc gtg cct Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro 565 570 575	1728
	att gag gcc ctg gga tct ctt cct cca cct cct gct gct cct aga gag Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu 580 585 590	1776
	cct gcc atc cac tct gaa ggc cag tgg gtt aca ctg ccc gct cct ctg Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu 595 600 605	1824
	gac acc atc aac gtg cac ctg aga gct ggc tac atc atc cca ctg caa Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln 610 615 620	1872
	ggc cct ggc ctg acc aca acc gaa tct aga cag cag ccc atg gct ctg Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu 625 630 635 640	1920

	gcc gtg gct ctt aca aaa ggc gga gag gct aga ggc gag ctg ttc tgg	1968
	Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp	
	645 650 655	
	gac gac ggc gag tct ctg gaa gtg ctg gaa cgg ggc gct tat acc caa	2016
	Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln	
	660 665 670	
	gtg atc ttc ctg gcc aga aac aac acc atc gtg aac gaa ctc gtg cgc	2064
	Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg	
	675 680 685	
	gtg acc agt gaa ggt gct gga ctg caa ctg cag aaa gtg acc gtg ctc	2112
	Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu	
	690 695 700	
	gga gtg gcc aca gct cct caa cag gtg ctg tct aac ggc gtg ccc gtg	2160
	Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val	
	705 710 715 720	
	tcc aac ttc aca tac agc ccc gac acc aag gtc ctg gac atc tgt gtg	2208
	Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val	
	725 730 735	
	tca ctg ctg atg ggc gag cag ttc ctg gtg tcc tgg tgc	2247
	Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp Cys	
	740 745	
<210>	19	
<211>	749	
<212>	PRT	
<213>	人工序列 (Artificial Sequence)	
<220>		
<223>	合成构建体	
<400>	19	
	Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe Gly	
	1 5 10 15	
	Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val Leu Leu Asn Thr Thr	
	20 25 30	
	Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr Ser	
	35 40 45	
	Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His Leu Ser Pro Leu	
	50 55 60	
	Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp Leu	
	65 70 75 80	
	Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu	
	85 90 95	
	Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe Leu Leu Asn Ser	
	100 105 110	
	Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala Leu Ser Trp Arg	
	115 120 125	
	Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu Pro	
	130 135 140	
	Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe Met	
	145 150 155 160	
	Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg Trp Gly Tyr Ser	
	165 170 175	
	Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met Thr Arg Ala His	

[0030]

				180				185					190			
	Phe	Pro	Leu	Asp	Val	Gln	Trp	Asn	Asp	Leu	Asp	Tyr	Met	Asp	Ser	Arg
				195				200					205			
	Arg	Asp	Phe	Thr	Phe	Asn	Lys	Asp	Gly	Phe	Arg	Asp	Phe	Pro	Ala	Met
				210			215					220				
	Val	Gln	Glu	Leu	His	Gln	Gly	Gly	Arg	Arg	Tyr	Met	Met	Ile	Val	Asp
225						230					235					240
Pro	Ala	Ile	Ser	Ser	Ser	Gly	Pro	Ala	Gly	Ser	Tyr	Arg	Pro	Tyr	Asp	
				245						250					255	
Glu	Gly	Leu	Arg	Arg	Gly	Val	Phe	Ile	Thr	Asn	Glu	Thr	Gly	Gln	Pro	
			260					265					270			
Leu	Ile	Gly	Lys	Val	Trp	Pro	Gly	Ser	Thr	Ala	Phe	Pro	Asp	Phe	Thr	
			275				280					285				
Asn	Pro	Thr	Ala	Leu	Ala	Trp	Trp	Glu	Asp	Met	Val	Ala	Glu	Phe	His	
						295					300					
Asp	Gln	Val	Pro	Phe	Asp	Gly	Met	Trp	Ile	Asp	Met	Asn	Glu	Pro	Ser	
305						310				315						320
Asn	Phe	Ile	Arg	Gly	Ser	Glu	Asp	Gly	Cys	Pro	Asn	Asn	Glu	Leu	Glu	
				325					330						335	
Asn	Pro	Pro	Tyr	Val	Pro	Gly	Val	Val	Gly	Gly	Thr	Leu	Gln	Ala	Ala	
				340				345					350			
Thr	Ile	Cys	Ala	Ser	Ser	His	Gln	Phe	Leu	Ser	Thr	His	Tyr	Asn	Leu	
				355			360					365				
His	Asn	Leu	Tyr	Gly	Leu	Thr	Glu	Ala	Ile	Ala	Ser	His	Arg	Ala	Leu	
				370			375				380					
[0031]	Val	Lys	Ala	Arg	Gly	Thr	Arg	Pro	Phe	Val	Ile	Ser	Arg	Ser	Thr	Phe
385						390					395					400
Ala	Gly	His	Gly	Arg	Tyr	Ala	Gly	His	Trp	Thr	Gly	Asp	Val	Trp	Ser	
				405						410					415	
Ser	Trp	Glu	Gln	Leu	Ala	Ser	Ser	Val	Pro	Glu	Ile	Leu	Gln	Phe	Asn	
				420				425						430		
Leu	Leu	Gly	Val	Pro	Leu	Val	Gly	Ala	Asp	Val	Cys	Gly	Phe	Leu	Gly	
			435				440					445				
Asn	Thr	Ser	Glu	Glu	Leu	Cys	Val	Arg	Trp	Thr	Gln	Leu	Gly	Ala	Phe	
				450			455				460					
Tyr	Pro	Phe	Met	Arg	Asn	His	Asn	Ser	Leu	Leu	Ser	Leu	Pro	Gln	Glu	
465						470					475					480
Pro	Tyr	Ser	Phe	Ser	Glu	Pro	Ala	Gln	Gln	Ala	Met	Arg	Lys	Ala	Leu	
				485						490					495	
Thr	Leu	Arg	Tyr	Ala	Leu	Leu	Pro	His	Leu	Tyr	Thr	Leu	Phe	His	Gln	
				500				505						510		
Ala	His	Val	Ala	Gly	Glu	Thr	Val	Ala	Arg	Pro	Leu	Phe	Leu	Glu	Phe	
				515				520				525				
Pro	Lys	Asp	Ser	Ser	Thr	Trp	Thr	Val	Asp	His	Gln	Leu	Leu	Trp	Gly	
				530			535				540					
Glu	Ala	Leu	Leu	Ile	Thr	Pro	Val	Leu	Gln	Ala	Gly	Lys	Ala	Glu	Val	
545						550					555					560
Thr	Gly	Tyr	Phe	Pro	Leu	Gly	Thr	Trp	Tyr	Asp	Leu	Gln	Thr	Val	Pro	
				565						570					575	
Ile	Glu	Ala	Leu	Gly	Ser	Leu	Pro	Pro	Pro	Pro	Ala	Ala	Pro	Arg	Glu	
				580				585					590			

	Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu	
	595 600 605	
	Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln	
	610 615 620	
	Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu	
	625 630 635 640	
	Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe Trp	
	645 650 655	
	Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr Gln	
	660 665 670	
	Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val Arg	
	675 680 685	
	Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys Val Thr Val Leu	
	690 695 700	
	Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn Gly Val Pro Val	
	705 710 715 720	
	Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu Asp Ile Cys Val	
	725 730 735	
	Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp Cys	
	740 745	
	<210> 20	
	<211> 2490	
	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
[0032]	<223> hGAAV780I 76 kD 蛋白的经工程化的 DNA	
	<220>	
	<221> CDS	
	<222> (1).. (2490)	
	<400> 20	
	gga cag ccc tgg tgc ttc ttc cca cca tct tac ccc agc tac aag ctg	48
	Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu	
	1 5 10 15	
	gaa aac ctg agc agc agc gag atg ggc tac acc gcc aca ctg acc aga	96
	Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg	
	20 25 30	
	acc aca cct aca ttc ttc ccg aag gac atc ctg aca ctg cgg ctg gac	144
	Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp	
	35 40 45	
	gtg atg atg gaa acc gag aac cgg ctg cac ttc acc atc aag gac ccc	192
	Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile Lys Asp Pro	
	50 55 60	
	gcc aat cgg aga tac gag gtg cca ctg gaa acc cct cac gtg cac tct	240
	Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro His Val His Ser	
	65 70 75 80	
	aga gcc cca tct cca ctg tac agc gtg gaa ttt tcc gag gaa ccc ttc	288
	Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe	
	85 90 95	
	ggc gtg atc gtg cgg aga cag ctg gat gga aga gtg ctg ctg aac acc	336
	Gly Val Ile Val Arg Arg Gln Leu Asp Gly Arg Val Leu Leu Asn Thr	
	100 105 110	

[0033]	aca gtg gcc cct ctg ttc ttc gcc gac cag ttt ctg cag ctg agc acc	384
	Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe Leu Gln Leu Ser Thr	
	115 120 125	
	agc ctg cct agc cag tat atc aca ggc ctg gcc gag cac ctg tct cca	432
	Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala Glu His Leu Ser Pro	
	130 135 140	
	ctg atg ctg agc aca tcc tgg acc aga atc acc ctg tgg aac aga gat	480
	Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr Leu Trp Asn Arg Asp	
	145 150 155 160	
	ctg gcc cct aca cct ggc gcc aac ctg tac ggc tct cac ccc ttt tat	528
	Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly Ser His Pro Phe Tyr	
	165 170 175	
	ctg gcc ctg gaa gac ggc gga tct gcc cac ggt gtt ttt ctg ctg aac	576
	Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly Val Phe Leu Leu Asn	
	180 185 190	
	tcc aac gcc atg gac gtg gtg ctg cag cca tct cct gct ctg tct tgg	624
	Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser Pro Ala Leu Ser Trp	
	195 200 205	
	aga agc aca ggc ggc atc ctg gac gtg tac atc ttt ctg ggc ccc gag	672
	Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile Phe Leu Gly Pro Glu	
	210 215 220	
	cct aag agc gtg gtg cag cag tat ctg gac gtc gtg ggc tac ccc ttc	720
	Pro Lys Ser Val Val Gln Gln Tyr Leu Asp Val Val Gly Tyr Pro Phe	
	225 230 235 240	
	atg cct cct tat tgg ggc ctg ggc ttc cac ctg tgc agg tgg gga tac	768
	Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu Cys Arg Trp Gly Tyr	
	245 250 255	
	agc agc acc gcc atc acc aga cag gtg gtg gaa aac atg acc cgg gct	816
	Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu Asn Met Thr Arg Ala	
	260 265 270	
	cac ttc cca ctg gac gtg cag tgg aac gac ctg gac tac atg gac agc	864
	His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu Asp Tyr Met Asp Ser	
	275 280 285	
	aga cgg gac ttc acc ttc aac aag gac ggc ttc aga gac ttc ccc gcc	912
	Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe Arg Asp Phe Pro Ala	
	290 295 300	
	atg gtg caa gaa ctg cac caa ggc ggc aga cgg tac atg atg att gtg	960
	Met Val Gln Glu Leu His Gln Gly Gly Arg Arg Tyr Met Met Ile Val	
	305 310 315 320	
	gac cca gcc atc agc tct agc ggc cct gcc gga agc tac aga cct tac	1008
	Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly Ser Tyr Arg Pro Tyr	
	325 330 335	
	gat gag ggc ctg aga aga ggc gtg ttc atc acc aac gag aca ggc cag	1056
	Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr Asn Glu Thr Gly Gln	
	340 345 350	
	cct ctg atc ggc aaa gtg tgg cct ggc agc aca gcc ttt cca gac ttc	1104
	Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr Ala Phe Pro Asp Phe	
	355 360 365	
	aca aac ccc acc gct ctg gct tgg tgg gaa gat atg gtg gcc gag ttc	1152
	Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp Met Val Ala Glu Phe	
	370 375 380	

	cac gat cag gtg ccc ttc gat ggc atg tgg atc gac atg aac gag ccc	1200	
	His Asp Gln Val Pro Phe Asp Gly Met Trp Ile Asp Met Asn Glu Pro		
385	390	395	400
	agc aac ttc atc cgg ggc agc gag gac ggc tgc ccc aac aac gaa ctg	1248	
	Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys Pro Asn Asn Glu Leu		
	405	410	415
	gaa aat cct cct tac gtg ccc ggc gtt gtc ggc gga aca ctt cag gcc	1296	
	Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly Gly Thr Leu Gln Ala		
	420	425	430
	gct aca atc tgt gcc agc agc cac cag ttc ctc agc acc cac tac aac	1344	
	Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu Ser Thr His Tyr Asn		
	435	440	445
	ctg cac aac ctg tac ggc ctg acc gag gcc att gcc tct cat aga gcc	1392	
	Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser His Arg Ala		
	450	455	460
	ctg gtt aag gcc aga ggc acc cgg cct ttt gtg atc agc aga agc aca	1440	
	Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser Arg Ser Thr		
465	470	475	480
	ttc gcc ggc cac ggc aga tac gcc gga cat tgg aca ggc gac gtg tgg	1488	
	Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly Asp Val Trp		
	485	490	495
	tct agt tgg gag cag ctg gct agc agc gtg cca gag atc ctg cag ttc	1536	
	Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe		
	500	505	510
[0034]	aat ctg ctg ggc gtg cca ctc gtg gga gcc gac gtt tgt ggc ttc ctg	1584	
	Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu		
	515	520	525
	ggc aac acc agc gag gaa ctg tgt gtg cgt tgg aca cag ctg ggc gcc	1632	
	Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala		
	530	535	540
	ttc tat ccc ttc atg aga aac cac aac agc ctg ctg agc ctg cct caa	1680	
	Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln		
545	550	555	560
	gag ccc tac agc ttt agc gag cct gca cag cag gcc atg aga aag gcc	1728	
	Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala		
	565	570	575
	ctg act ctg aga tac gct ctg ctg ccc cac ctg tac acc ctg ttt cac	1776	
	Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His		
	580	585	590
	cag gct cac gtg gcc ggg gag aca gtg gct aga cct ctg ttc ctg gaa	1824	
	Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu		
	595	600	605
	ttt ccc aag gac agc tcc acc tgg acc gtg gat cat cag ctg ctg tgg	1872	
	Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp		
	610	615	620
	gga gaa gcc ctg ctc atc aca cct gtt ctg cag gcc gga aag gcc gaa	1920	
	Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu		
625	630	635	640
	gtg acc ggc tat ttt cct ctc ggc act tgg tac gac ctg cag acc gtg	1968	
	Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val		
	645	650	655

	cct att gag gcc ctg gga tct ctt cct cca cct cct gct gct cct aga	2016
	Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg	
	660 665 670	
	gag cct gcc atc cac tct gaa ggc cag tgg gtt aca ctg ccc gct cct	2064
	Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro	
	675 680 685	
	ctg gac acc atc aac gtg cac ctg aga gct ggc tac atc atc cca ctg	2112
	Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu	
	690 695 700	
	caa ggc cct ggc ctg acc aca acc gaa tct aga cag cag ccc atg gct	2160
	Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala	
	705 710 715 720	
	ctg gcc gtg gct ctt aca aaa ggc gga gag gct aga ggc gag ctg ttc	2208
	Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe	
	725 730 735	
	tgg gac gac ggc gag tct ctg gaa gtg ctg gaa cgg ggc gct tat acc	2256
	Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr	
	740 745 750	
	caa gtg atc ttc ctg gcc aga aac aac acc atc gtg aac gaa ctc gtg	2304
	Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val	
	755 760 765	
	cgc gtg acc agt gaa ggt gct gga ctg caa ctg cag aaa gtg acc gtg	2352
	Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys Val Thr Val	
	770 775 780	
[0035]	ctc gga gtg gcc aca gct cct caa cag gtg ctg tct aac ggc gtg ccc	2400
	Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn Gly Val Pro	
	785 790 795 800	
	gtg tcc aac ttc aca tac agc ccc gac acc aag gtc ctg gac atc tgt	2448
	Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu Asp Ile Cys	
	805 810 815	
	gtg tca ctg ctg atg ggc gag cag ttc ctg gtg tcc tgg tgc	2490
	Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp Cys	
	820 825 830	
	<210> 21	
	<211> 830	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 合成构建体	
	<400> 21	
	Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr Pro Ser Tyr Lys Leu	
	1 5 10 15	
	Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr Ala Thr Leu Thr Arg	
	20 25 30	
	Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu Thr Leu Arg Leu Asp	
	35 40 45	
	Val Met Met Glu Thr Glu Asn Arg Leu His Phe Thr Ile Lys Asp Pro	
	50 55 60	
	Ala Asn Arg Arg Tyr Glu Val Pro Leu Glu Thr Pro His Val His Ser	
	65 70 75 80	
	Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe Ser Glu Glu Pro Phe	

					</											

[0037]

Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe
 500 505 510
 Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu
 515 520 525
 Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala
 530 535 540
 Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln
 545 550 555 560
 Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala
 565 570 575
 Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His
 580 585 590
 Gln Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu
 595 600 605
 Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp
 610 615 620
 Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu
 625 630 635 640
 Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val
 645 650 655
 Pro Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg
 660 665 670
 Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro
 675 680 685
 Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu
 690 695 700
 Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala
 705 710 715 720
 Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala Arg Gly Glu Leu Phe
 725 730 735
 Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu Arg Gly Ala Tyr Thr
 740 745 750
 Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile Val Asn Glu Leu Val
 755 760 765
 Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu Gln Lys Val Thr Val
 770 775 780
 Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu Ser Asn Gly Val Pro
 785 790 795 800
 Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys Val Leu Asp Ile Cys
 805 810 815
 Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val Ser Trp Cys
 820 825 830
 <210> 22
 <211> 2952
 <212> DNA
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> 合成构建体
 <220>
 <221> CDS
 <222> (1).. (2952)

<220>		
<221>	misc_feature	
<222>	(1)..(270)	
<223>	BiP 信号肽+ vIGF2 + 2GS 延伸序列	
<220>		
<221>	misc_feature	
<222>	(271)..(2952)	
<223>	hGAA 61 - 952 780I 的经工程化的 DNA	
<220>		
<221>	misc_feature	
<222>	(2428)..(2430)	
<223>	Ile 密码子	
<400>	22	
	atg aag ctg tct ctg gtg gct gct atg ctg ctg ctc ctg tct gcc gcc	48
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
	aga gcc agc aga aca ctt tgt ggc gga gag ctg gtg gac acc ctg cag	96
	Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln	
	20 25 30	
	ttt gtg tgt ggc gac aga ggc ttc ctg ttc agc aga cct gcc agc cgg	144
	Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg	
	35 40 45	
	gtt tcc aga cgg tct aga gga atc gtg gaa gag tgc tgc ttc aga agc	192
	Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser	
	50 55 60	
[0038]	tgc gat ctg gcc ctg ctg gaa acc tac tgt gcc aca cca gcc aga tct	240
	Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser	
	65 70 75 80	
	gaa ggc ggc gga gga tct ggc gga ggc gga tct aga cct gga cct aga	288
	Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg	
	85 90 95	
	gac gcc cag gct cac cct ggt aga cct aga gct gtg cct aca cag tgc	336
	Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys	
	100 105 110	
	gac gtg cca cct aac agc aga ttc gac tgc gcc cct gac aag gcc atc	384
	Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile	
	115 120 125	
	aca caa gag cag tgt gaa gcc aga ggc tgc tgc tac atc cct gcc aaa	432
	Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys	
	130 135 140	
	caa gga ctg cag ggc gcc cag atg gga cag cct tgg tgc ttc ttc cca	480
	Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro	
	145 150 155 160	
	cca tct tac ccc agc tac aag ctg gaa aac ctg agc agc tcc gag atg	528
	Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	
	165 170 175	
	ggc tac acc gcc aca ctg acc aga acc aca cct aca ttc ttc ccg aag	576
	Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys	
	180 185 190	
	gac atc ctg aca ctg cgg ctg gac gtg atg atg gaa acc gag aac cgg	624
	Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg	

	195	200	205	
	ctg cac ttc acc atc aag gac ccc gcc aat cgg aga tac gag gtg ccc			672
	Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro			
	210	215	220	
	ctg gaa aca ccc cac gtg cac tct aga gca ccc tct cca ctg tac agc			720
	Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser			
	225	230	235	240
	gtg gaa ttt tcc gag gaa ccc ttc ggc gtg atc gtg cgg aga cag ctg			768
	Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu			
	245	250	255	
	gat ggc aga gtg ctc ctg aat acc aca gtg gcc cct ctg ttc ttc gcc			816
	Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala			
	260	265	270	
	gac cag ttt ctg cag ctg agc acc agc ctg cct agc cag tat atc aca			864
	Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr			
	275	280	285	
	ggc ctg gcc gag cat ctg agc cct ctg atg ctg agc aca tcc tgg acc			912
	Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr			
	290	295	300	
	aga atc acc ctg tgg aac cgc gac ctg gct cct aca cct ggc gcc aat			960
	Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn			
	305	310	315	320
	ctg tac ggc tct cac ccc ttc tac ctg gca ctg gaa gac ggt gga tct			1008
	Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser			
	325	330	335	
[0039]	gcc cac ggt gtc ttt ctg ctg aat agc aac gcc atg gac gtg gtg ctg			1056
	Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu			
	340	345	350	
	cag ccc tct cct gca ctg tct tgg aga tct aca ggc ggc atc ctg gac			1104
	Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp			
	355	360	365	
	gtg tac atc ttt ctg ggc ccc gag cct aag agc gtg gtg cag cag tat			1152
	Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr			
	370	375	380	
	ctg gac gtc gtg ggc tac ccc ttc atg cct cct tat tgg ggc ctg ggc			1200
	Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly			
	385	390	395	400
	ttc cac ctg tgt agg tgg ggc tac agc agc acc gcc atc acc aga cag			1248
	Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln			
	405	410	415	
	gtg gtg gaa aac atg acc cgg gct cac ttc cca ctg gac gtg cag tgg			1296
	Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp			
	420	425	430	
	aac gac ctg gac tac atg gac agc aga cgg gac ttc acc ttc aac aag			1344
	Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys			
	435	440	445	
	gac ggc ttc aga gac ttc ccc gcc atg gtg caa gag ctg cat caa ggc			1392
	Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly			
	450	455	460	
	gga cgg cgg tac atg atg att gtg gac cct gcc atc agc agc tct gga			1440
	Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly			

	465	470	475	480	
	cca gcc ggc agc tac aga cct tac gat gag gga ctg aga aga ggc gtg				1488
	Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val				
		485	490	495	
	ttc atc acc aac gag aca ggc cag cct ctg atc ggc aaa gtg tgg cct				1536
	Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro				
		500	505	510	
	ggc agc aca gcc ttt cca gac ttc aca aac ccc acc gct ctg gct tgg				1584
	Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp				
		515	520	525	
	tgg gaa gat atg gtg gcc gag ttc cac gat cag gtg ccc ttc gat ggc				1632
	Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly				
		530	535	540	
	atg tgg atc gac atg aac gag ccc agc aac ttc atc cgg ggc agc gag				1680
	Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu				
		545	550	555	560
	gat ggc tgc ccc aac aac gaa cta gaa aat cct cct tac gtg ccc ggc				1728
	Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly				
		565	570	575	
	gtt gtc ggc gga aca ctt cag gcc gct aca atc tgt gcc agc agc cat				1776
	Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His				
		580	585	590	
	cag ttt ctg agc acc cac tac aac ctg cac aac ctg tac ggc ctg acc				1824
	Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr				
		595	600	605	
[0040]	gag gcc att gcc tct cat aga gcc ctg gtt aag gcc aga ggc acc cgg				1872
	Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg				
		610	615	620	
	cct ttt gtg atc agc aga agc aca ttc gcc ggc cac ggc aga tac gca				1920
	Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala				
		625	630	635	640
	gga cat tgg aca ggc gac gtg tgg tct agt tgg gag cag ctg gct agc				1968
	Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser				
		645	650	655	
	agc gtg cca gag atc ctg cag ttc aat ctg ctg ggc gtg cca ctc gtg				2016
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val				
		660	665	670	
	gga gcc gac gtt tgt ggc ttc ctg ggc aac acc agc gag gaa ctg tgt				2064
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys				
		675	680	685	
	gtg cgt tgg aca cag ctg ggc gcc ttc tat ccc ttc atg aga aac cac				2112
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His				
		690	695	700	
	aac agc ctg ctg agc ctg cct caa gag ccc tac agc ttt agc gag cct				2160
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro				
		705	710	715	720
	gca cag cag gcc atg aga aag gcc ctg act ctg aga tac gcc ctg ctg				2208
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu				
		725	730	735	
	cct cac ctg tac acc ctg ttt cat cag gcc cac gtg gca ggc gag aca				2256
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr				

	740	745	750	
	gtg gct aga cct ctg ttc ctg gaa ttt ccc aag gac agc tcc acc tgg			2304
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp			
	755	760	765	
	acc gtg gat cat cag ctg ctg tgg gga gaa gcc ctg ctg att aca cca			2352
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro			
	770	775	780	
	gtg ctg cag gcc gga aag gcc gaa gtg aca ggc tat ttc cct ctc ggc			2400
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly			
	785	790	795	800
	act tgg tac gac ctg cag acc gtg cct atc gag gct ctg gga tct ctt			2448
	Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu			
	805	810	815	
	cct cca cct cct gcc gct cct aga gag cct gcc att cac tct gaa ggc			2496
	Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly			
	820	825	830	
	cag tgg gtt acc ctg cct gct cct ctg gac acc atc aac gtg cac ctg			2544
	Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu			
	835	840	845	
	aga gcc ggc tac atc atc cct ctg caa ggc cct ggc ctg acc aca acc			2592
	Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr			
	850	855	860	
	gaa tct aga cag cag ccc atg gca ctg gcc gtg gct ctt aca aaa ggc			2640
	Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly			
	865	870	875	880
[0041]	gga gag gct aga ggc gag ctg ttc tgg gat gat ggc gag agc cta gaa			2688
	Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu			
	885	890	895	
	gtg ctg gaa cgg ggc gct tat acc caa gtg atc ttc ctg gcc aga aac			2736
	Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn			
	900	905	910	
	aac acc atc gtg aac gaa ctc gtg cgc gtg acc agt gaa ggt gct gga			2784
	Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly			
	915	920	925	
	ctg caa ctg cag aaa gtg acc gtg ctc gga gtg gcc aca gct cct cag			2832
	Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln			
	930	935	940	
	cag gtt ctg tct aat ggc gtg ccc gtg tcc aac ttc aca tac agc ccc			2880
	Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro			
	945	950	955	960
	gac acc aag gtc ctg gac atc tgt gtg tcc ctg ctt atg ggc gag cag			2928
	Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln			
	965	970	975	
	ttc ctg gtg tcc tgg tgc tga taa			2952
	Phe Leu Val Ser Trp Cys			
	980			
	<210> 23			
	<211> 982			
	<212> PRT			
	<213> 人工序列 (Artificial Sequence)			
	<220>			

<223> 合成构建体
 <400> 23
 Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Ser Ala Ala
 1 5 10 15
 Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln
 20 25 30
 Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg
 35 40 45
 Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser
 50 55 60
 Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser
 65 70 75 80
 Glu Gly Gly Gly Gly Ser Gly Gly Gly Ser Arg Pro Gly Pro Arg
 85 90 95
 Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys
 100 105 110
 Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile
 115 120 125
 Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys
 130 135 140
 Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro
 145 150 155 160
 Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met
 165 170 175
 Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys
 180 185 190
 Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg
 195 200 205
 Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro
 210 215 220
 Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser
 225 230 235 240
 Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu
 245 250 255
 Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala
 260 265 270
 Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr
 275 280 285
 Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr
 290 295 300
 Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn
 305 310 315 320
 Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser
 325 330 335
 Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu
 340 345 350
 Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp
 355 360 365
 Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr
 370 375 380
 Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly

[0042]

	385		390		395		400
	Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln						
		405			410		415
	Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp						
		420			425		430
	Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys						
		435			440		445
	Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly						
		450			455		460
	Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly						
		465			470		480
	Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val						
		485			490		495
	Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro						
		500			505		510
	Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp						
		515			520		525
	Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly						
		530			535		540
	Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu						
		545			550		560
	Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly						
		565			570		575
	Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His						
		580			585		590
[0043]	Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr						
		595			600		605
	Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg						
		610			615		620
	Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala						
		625			630		640
	Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser						
		645			650		655
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val						
		660			665		670
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys						
		675			680		685
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His						
		690			695		700
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro						
		705			710		720
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu						
		725			730		735
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr						
		740			745		750
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp						
		755			760		765
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro						
		770			775		780
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly						
		785			790		800

Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu
 805 810 815
 Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly
 820 825 830
 Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu
 835 840 845
 Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr
 850 855 860
 Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly
 865 870 875 880
 Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu
 885 890 895
 Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn
 900 905 910
 Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly
 915 920 925
 Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln
 930 935 940
 Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro
 945 950 955 960
 Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln
 965 970 975
 Phe Leu Val Ser Trp Cys
 980

[0044]

<210> 24
 <211> 2952
 <212> DNA
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> 合成构建体
 <220>
 <221> CDS
 <222> (1).. (2952)
 <220>
 <221> misc_feature
 <222> (1).. (270)
 <223> BiP-vIGF 肽
 <220>
 <221> misc_feature
 <222> (1).. (270)
 <223> BiP 信号肽+ vIGF2 + 2GS 延伸序列
 <220>
 <221> misc_feature
 <222> (271).. (2952)
 <223> hGAA 61-952 V780 DNA
 <220>
 <221> misc_feature
 <222> (2428).. (2430)
 <223> hGAA 780 缬氨酸的密码子
 <400> 24
 atg aag ctg tct ctg gtg gct gct atg ctg ctg ctc ctg tct gcc gcc

48

[0045]	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
	aga gcc agc aga aca ctt tgt ggc gga gag ctg gtg gac acc ctg cag	96
	Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln	
	20 25 30	
	ttt gtg tgt ggc gac aga ggc ttc ctg ttc agc aga cct gcc agc cgg	144
	Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg	
	35 40 45	
	gtt tcc aga cgg tct aga gga atc gtg gaa gag tgc tgc ttc aga agc	192
	Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser	
	50 55 60	
	tgc gat ctg gcc ctg ctg gaa acc tac tgt gcc aca cca gcc aga tct	240
	Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser	
	65 70 75 80	
	gaa ggc ggc gga gga tct ggc gga ggc gga tct aga cct gga cct aga	288
	Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg	
	85 90 95	
	gac gcc cag gct cac cct ggt aga cct aga gct gtg cct aca cag tgc	336
	Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys	
	100 105 110	
	gac gtg cca cct aac agc aga ttc gac tgc gcc cct gac aag gcc atc	384
	Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile	
	115 120 125	
	aca caa gag cag tgt gaa gcc aga ggc tgc tgc tac atc cct gcc aaa	432
	Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys	
	130 135 140	
	caa gga ctg cag ggc gcc cag atg gga cag cct tgg tgc ttc ttc cca	480
	Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro	
	145 150 155 160	
	cca tct tac ccc agc tac aag ctg gaa aac ctg agc agc tcc gag atg	528
	Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	
	165 170 175	
	ggc tac acc gcc aca ctg acc aga acc aca cct aca ttc ttc ccg aag	576
	Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys	
	180 185 190	
	gac atc ctg aca ctg cgg ctg gac gtg atg atg gaa acc gag aac cgg	624
	Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg	
	195 200 205	
	ctg cac ttc acc atc aag gac ccc gcc aat cgg aga tac gag gtg ccc	672
	Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro	
	210 215 220	
	ctg gaa aca ccc cac gtg cac tct aga gca ccc tct cca ctg tac agc	720
	Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser	
	225 230 235 240	
	gtg gaa ttt tcc gag gaa ccc ttc ggc gtg atc gtg cgg aga cag ctg	768
	Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu	
	245 250 255	
	gat ggc aga gtg ctc ctg aat acc aca gtg gcc cct ctg ttc ttc gcc	816
	Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala	
	260 265 270	
	gac cag ttt ctg cag ctg agc acc agc ctg cct agc cag tat atc aca	864

[0046]	Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr	
	275 280 285	
	ggc ctg gcc gag cat ctg agc cct ctg atg ctg agc aca tcc tgg acc	912
	Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr	
	290 295 300	
	aga atc acc ctg tgg aac cgc gac ctg gct cct aca cct ggc gcc aat	960
	Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn	
	305 310 315 320	
	ctg tac ggc tct cac ccc ttc tac ctg gca ctg gaa gac ggt gga tct	1008
	Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser	
	325 330 335	
	gcc cac ggt gtc ttt ctg ctg aat agc aac gcc atg gac gtg gtg ctg	1056
	Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu	
	340 345 350	
	cag ccc tct cct gca ctg tct tgg aga tct aca ggc ggc atc ctg gac	1104
	Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp	
	355 360 365	
	gtg tac atc ttt ctg ggc ccc gag cct aag agc gtg gtg cag cag tat	1152
	Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr	
	370 375 380	
	ctg gac gtc gtg ggc tac ccc ttc atg cct cct tat tgg ggc ctg ggc	1200
	Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly	
	385 390 395 400	
	ttc cac ctg tgt agg tgg ggc tac agc agc acc gcc atc acc aga cag	1248
	Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln	
	405 410 415	
	gtg gtg gaa aac atg acc cgg gct cac ttc cca ctg gac gtg cag tgg	1296
	Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp	
	420 425 430	
	aac gac ctg gac tac atg gac agc aga cgg gac ttc acc ttc aac aag	1344
	Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys	
	435 440 445	
	gac ggc ttc aga gac ttc ccc gcc atg gtg caa gag ctg cat caa ggc	1392
	Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly	
	450 455 460	
	gga cgg cgg tac atg atg att gtg gac cct gcc atc agc agc tct gga	1440
	Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly	
	465 470 475 480	
	cca gcc ggc agc tac aga cct tac gat gag gga ctg aga aga ggc gtg	1488
	Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val	
	485 490 495	
	ttc atc acc aac gag aca ggc cag cct ctg atc ggc aaa gtg tgg cct	1536
	Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro	
	500 505 510	
	ggc agc aca gcc ttt cca gac ttc aca aac ccc acc gct ctg gct tgg	1584
	Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp	
	515 520 525	
	tgg gaa gat atg gtg gcc gag ttc cac gat cag gtg ccc ttc gat ggc	1632
	Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly	
	530 535 540	
	atg tgg atc gac atg aac gag ccc agc aac ttc atc cgg ggc agc gag	1680

[0047]	Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu	
	545 550 555 560	
	gat ggc tgc ccc aac aac gaa cta gaa aat cct cct tac gtg ccc ggc	1728
	Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly	
	565 570 575	
	gtt gtc ggc gga aca ctt cag gcc gct aca atc tgt gcc agc agc cat	1776
	Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His	
	580 585 590	
	cag ttt ctg agc acc cac tac aac ctg cac aac ctg tac ggc ctg acc	1824
	Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr	
	595 600 605	
	gag gcc att gcc tct cat aga gcc ctg gtt aag gcc aga ggc acc cgg	1872
	Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg	
	610 615 620	
	cct ttt gtg atc agc aga agc aca ttc gcc ggc cac ggc aga tac gca	1920
	Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala	
	625 630 635 640	
	gga cat tgg aca ggc gac gtg tgg tct agt tgg gag cag ctg gct agc	1968
	Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser	
	645 650 655	
	agc gtg cca gag atc ctg cag ttc aat ctg ctg ggc gtg cca ctc gtg	2016
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val	
	660 665 670	
	gga gcc gac gtt tgt ggc ttc ctg ggc aac acc agc gag gaa ctg tgt	2064
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys	
	675 680 685	
	gtg cgt tgg aca cag ctg ggc gcc ttc tat ccc ttc atg aga aac cac	2112
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His	
	690 695 700	
	aac agc ctg ctg agc ctg cct caa gag ccc tac agc ttt agc gag cct	2160
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro	
	705 710 715 720	
	gca cag cag gcc atg aga aag gcc ctg act ctg aga tac gcc ctg ctg	2208
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu	
	725 730 735	
	cct cac ctg tac acc ctg ttt cat cag gcc cac gtg gca ggc gag aca	2256
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr	
	740 745 750	
	gtg gct aga cct ctg ttc ctg gaa ttt ccc aag gac agc tcc acc tgg	2304
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp	
	755 760 765	
	acc gtg gat cat cag ctg ctg tgg gga gaa gcc ctg ctg att aca cca	2352
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro	
	770 775 780	
	gtg ctg cag gcc gga aag gcc gaa gtg aca ggc tat ttc cct ctc ggc	2400
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly	
	785 790 795 800	
	act tgg tac gac ctg cag acc gtg cct atc gag gct ctg gga tct ctt	2448
	Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu	
	805 810 815	
	cct cca cct cct gcc gct cct aga gag cct gcc att cac tct gaa ggc	2496

	Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly	
	820 825 830	
	cag tgg gtt acc ctg cct gct cct ctg gac acc atc aac gtg cac ctg	2544
	Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu	
	835 840 845	
	aga gcc ggc tac atc atc cct ctg caa ggc cct ggc ctg acc aca acc	2592
	Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr	
	850 855 860	
	gaa tct aga cag cag ccc atg gca ctg gcc gtg gct ctt aca aaa ggc	2640
	Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly	
	865 870 875 880	
	gga gag gct aga ggc gag ctg ttc tgg gat gat ggc gag agc cta gaa	2688
	Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu	
	885 890 895	
	gtg ctg gaa cgg ggc gct tat acc caa gtg atc ttc ctg gcc aga aac	2736
	Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn	
	900 905 910	
	aac acc atc gtg aac gaa ctc gtg cgc gtg acc agt gaa ggt gct gga	2784
	Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly	
	915 920 925	
	ctg caa ctg cag aaa gtg acc gtg ctc gga gtg gcc aca gct cct cag	2832
	Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln	
	930 935 940	
	cag gtt ctg tct aat ggc gtg ccc gtg tcc aac ttc aca tac agc ccc	2880
	Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro	
[0048]	945 950 955 960	
	gac acc aag gtc ctg gac atc tgt gtg tcc ctg ctt atg ggc gag cag	2928
	Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln	
	965 970 975	
	ttc ctg gtg tcc tgg tgc tga taa	2952
	Phe Leu Val Ser Trp Cys	
	980	
	<210> 25	
	<211> 982	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 合成构建体	
	<400> 25	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
	Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln	
	20 25 30	
	Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg	
	35 40 45	
	Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser	
	50 55 60	
	Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser	
	65 70 75 80	
	Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg	
	85 90 95	

[0049]

```

Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys
      100                      105                      110
Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile
      115                      120                      125
Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys
      130                      135                      140
Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro
      145                      150                      155                      160
Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met
      165                      170                      175
Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys
      180                      185                      190
Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg
      195                      200                      205
Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro
      210                      215                      220
Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser
      225                      230                      235                      240
Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu
      245                      250                      255
Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala
      260                      265                      270
Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr
      275                      280                      285
Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr
      290                      295                      300
Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn
      305                      310                      315                      320
Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser
      325                      330                      335
Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu
      340                      345                      350
Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp
      355                      360                      365
Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr
      370                      375                      380
Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly
      385                      390                      395                      400
Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln
      405                      410                      415
Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp
      420                      425                      430
Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys
      435                      440                      445
Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly
      450                      455                      460
Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly
      465                      470                      475                      480
Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val
      485                      490                      495
Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro

```

				500						505					510				
	Gly	Ser	Thr	Ala	Phe	Pro	Asp	Phe	Thr	Asn	Pro	Thr	Ala	Leu	Ala	Trp			
				515						520					525				
	Trp	Glu	Asp	Met	Val	Ala	Glu	Phe	His	Asp	Gln	Val	Pro	Phe	Asp	Gly			
				530						535					540				
	Met	Trp	Ile	Asp	Met	Asn	Glu	Pro	Ser	Asn	Phe	Ile	Arg	Gly	Ser	Glu			
	545						550				555					560			
	Asp	Gly	Cys	Pro	Asn	Asn	Glu	Leu	Glu	Asn	Pro	Pro	Tyr	Val	Pro	Gly			
					565						570					575			
	Val	Val	Gly	Gly	Thr	Leu	Gln	Ala	Ala	Thr	Ile	Cys	Ala	Ser	Ser	His			
				580							585					590			
	Gln	Phe	Leu	Ser	Thr	His	Tyr	Asn	Leu	His	Asn	Leu	Tyr	Gly	Leu	Thr			
				595						600						605			
	Glu	Ala	Ile	Ala	Ser	His	Arg	Ala	Leu	Val	Lys	Ala	Arg	Gly	Thr	Arg			
				610						615						620			
	Pro	Phe	Val	Ile	Ser	Arg	Ser	Thr	Phe	Ala	Gly	His	Gly	Arg	Tyr	Ala			
	625						630						635			640			
	Gly	His	Trp	Thr	Gly	Asp	Val	Trp	Ser	Ser	Trp	Glu	Gln	Leu	Ala	Ser			
					645						650					655			
	Ser	Val	Pro	Glu	Ile	Leu	Gln	Phe	Asn	Leu	Leu	Gly	Val	Pro	Leu	Val			
				660							665					670			
	Gly	Ala	Asp	Val	Cys	Gly	Phe	Leu	Gly	Asn	Thr	Ser	Glu	Glu	Leu	Cys			
				675							680					685			
	Val	Arg	Trp	Thr	Gln	Leu	Gly	Ala	Phe	Tyr	Pro	Phe	Met	Arg	Asn	His			
				690							695					700			
[0050]	Asn	Ser	Leu	Leu	Ser	Leu	Pro	Gln	Glu	Pro	Tyr	Ser	Phe	Ser	Glu	Pro			
	705						710				715					720			
	Ala	Gln	Gln	Ala	Met	Arg	Lys	Ala	Leu	Thr	Leu	Arg	Tyr	Ala	Leu	Leu			
					725						730					735			
	Pro	His	Leu	Tyr	Thr	Leu	Phe	His	Gln	Ala	His	Val	Ala	Gly	Glu	Thr			
				740							745					750			
	Val	Ala	Arg	Pro	Leu	Phe	Leu	Glu	Phe	Pro	Lys	Asp	Ser	Ser	Thr	Trp			
				755							760					765			
	Thr	Val	Asp	His	Gln	Leu	Leu	Trp	Gly	Glu	Ala	Leu	Leu	Ile	Thr	Pro			
				770							775					780			
	Val	Leu	Gln	Ala	Gly	Lys	Ala	Glu	Val	Thr	Gly	Tyr	Phe	Pro	Leu	Gly			
	785						790				795					800			
	Thr	Trp	Tyr	Asp	Leu	Gln	Thr	Val	Pro	Ile	Glu	Ala	Leu	Gly	Ser	Leu			
					805						810					815			
	Pro	Pro	Pro	Pro	Ala	Ala	Pro	Arg	Glu	Pro	Ala	Ile	His	Ser	Glu	Gly			
				820							825					830			
	Gln	Trp	Val	Thr	Leu	Pro	Ala	Pro	Leu	Asp	Thr	Ile	Asn	Val	His	Leu			
				835							840					845			
	Arg	Ala	Gly	Tyr	Ile	Ile	Pro	Leu	Gln	Gly	Pro	Gly	Leu	Thr	Thr	Thr			
				850							855					860			
	Glu	Ser	Arg	Gln	Gln	Pro	Met	Ala	Leu	Ala	Val	Ala	Leu	Thr	Lys	Gly			
	865						870									880			
	Gly	Glu	Ala	Arg	Gly	Glu	Leu	Phe	Trp	Asp	Asp	Gly	Glu	Ser	Leu	Glu			
					885						890					895			
	Val	Leu	Glu	Arg	Gly	Ala	Tyr	Thr	Gln	Val	Ile	Phe	Leu	Ala	Arg	Asn			
				900							905					910			

	Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly	
	915 920 925	
	Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln	
	930 935 940	
	Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro	
	945 950 955 960	
	Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln	
	965 970 975	
	Phe Leu Val Ser Trp Cys	
	980	
	<210> 26	
	<211> 22	
	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> miRNA 靶序列	
	<400> 26	
	agtgaattct accagtgcc ta	22
	<210> 27	
	<211> 24	
	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> miRNA 靶序列	
	<400> 27	
[0051]	agtgtgagtt ctaccattgc caaa	24
	<210> 28	
	<211> 4581	
	<212> DNA	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 合成构建体	
	<220>	
	<221> misc_feature	
	<222> (1)..(130)	
	<223> 5' ITR	
	<220>	
	<221> enhancer	
	<222> (195)..(437)	
	<223> CMV IE 增强子	
	<220>	
	<221> promoter	
	<222> (440)..(721)	
	<223> 鸡 β -肌动蛋白启动子	
	<220>	
	<221> Intron	
	<222> (721)..(1128)	
	<223> CAG 中的杂交内含子	
	<220>	
	<221> CDS	
	<222> (1141)..(4092)	

<223>	BiP-vIGF2-hGAAco	
<220>		
<221>	misc_feature	
<222>	(3568)..(3570)	
<223>	Ile 密码子	
<220>		
<221>	polyA_signal	
<222>	(4161)..(4287)	
<223>	兔 β -珠蛋白 poly a	
<220>		
<221>	misc_feature	
<222>	(4452)..(4581)	
<223>	3' ITR	
<400>	28	
	ctgcgcgctc gctcgcctcac tgaggccgcc cgggcaaagc ccgggcgctc ggcgaccttt	60
	ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact	120
	aggggttctt tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct	180
	aggaagatcg taccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg	240
	actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgccactt ggccagtacat	300
	caagtgtatc atatgccaa gacgccccct attgacgtca atgacggtaa atggcccgcc	360
	tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta	420
	ttagtcacg ctattaccat ggtcagagtg agccccacgt tctgcttcac tctccccatc	480
	tccccccct ccccccccc aattttgtat ttattttatt tttaattatt ttgtgcagcg	540
	atgggggcgg gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg	600
	ggcggggcga ggcggagagg tgcggcggca gccaatcaga gcggcgcgct ccgaaagttt	660
[0052]	ccttttatgg cgaggcggcg gcggcggcgg ccctataaaa agcgaagcgc gcggcgggcg	720
	ggagtcgctg cgcgctgcct tcgccccgtg ccccgctccg ccgccgcctc gcgccgccc	780
	ccccggctct gactgaccgc gttactccca cagggtgagcg ggccgggacgg cccttctcct	840
	ccgggctgta attagcgctt ggtttaatga cggttgttt cttttctgtg gctgcgtgaa	900
	agccttgagg ggctccggga gggccctttg tgcgggggga gcggctcggg gctgtccgcg	960
	gggggacggc tgccttcggg ggggacgggg cagggcgggg ttcggcttct ggctgtgac	1020
	cggcggtctc agagcctctg ctaaccatgt tcatgccttc ttctttttcc tacagctcct	1080
	gggcaacgtg ctggttattg tgetgtctca tcatttttgc aaagaattgg atccgccacc	1140
	atg aag ctg tct ctg gtg gct gct atg ctg ctg ctc ctg tct gcc gcc	1188
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
	aga gcc agc aga aca ctt tgt ggc gga gag ctg gtg gac acc ctg cag	1236
	Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln	
	20 25 30	
	ttt gtg tgt ggc gac aga ggc ttc ctg ttc agc aga cct gcc agc cgg	1284
	Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg	
	35 40 45	
	gtt tcc aga cgg tct aga gga atc gtg gaa gag tgc tgc ttc aga agc	1332
	Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser	
	50 55 60	
	tgc gat ctg gcc ctg ctg gaa acc tac tgt gcc aca cca gcc aga tct	1380
	Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser	
	65 70 75 80	
	gaa ggc ggc gga gga tct ggc gga ggc gga tct aga cct gga cct aga	1428
	Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg	
	85 90 95	

[0053]	gac gcc cag gct cac cct ggt aga cct aga gct gtg cct aca cag tgc	1476
	Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys	
	100 105 110	
	gac gtg cca cct aac agc aga ttc gac tgc gcc cct gac aag gcc atc	1524
	Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile	
	115 120 125	
	aca caa gag cag tgt gaa gcc aga ggc tgc tgc tac atc cct gcc aaa	1572
	Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys	
	130 135 140	
	caa gga ctg cag ggc gcc cag atg gga cag cct tgg tgc ttc ttc cca	1620
	Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro	
	145 150 155 160	
	cca tct tac ccc agc tac aag ctg gaa aac ctg agc agc tcc gag atg	1668
	Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	
	165 170 175	
	ggc tac acc gcc aca ctg acc aga acc aca cct aca ttc ttc ccg aag	1716
	Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys	
	180 185 190	
	gac atc ctg aca ctg cgg ctg gac gtg atg atg gaa acc gag aac cgg	1764
	Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg	
	195 200 205	
	ctg cac ttc acc atc aag gac ccc gcc aat cgg aga tac gag gtg ccc	1812
	Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro	
	210 215 220	
	ctg gaa aca ccc cac gtg cac tct aga gca ccc tct cca ctg tac agc	1860
	Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser	
	225 230 235 240	
	gtg gaa ttt tcc gag gaa ccc ttc ggc gtg atc gtg cgg aga cag ctg	1908
	Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu	
	245 250 255	
	gat ggc aga gtg ctc ctg aat acc aca gtg gcc cct ctg ttc ttc gcc	1956
	Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala	
	260 265 270	
	gac cag ttt ctg cag ctg agc acc agc ctg cct agc cag tat atc aca	2004
	Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr	
	275 280 285	
	ggc ctg gcc gag cat ctg agc cct ctg atg ctg agc aca tcc tgg acc	2052
	Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr	
	290 295 300	
	aga atc acc ctg tgg aac cgc gac ctg gct cct aca cct ggc gcc aat	2100
	Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn	
	305 310 315 320	
	ctg tac ggc tct cac ccc ttc tac ctg gca ctg gaa gac ggt gga tct	2148
	Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser	
	325 330 335	
	gcc cac ggt gtc ttt ctg ctg aat agc aac gcc atg gac gtg gtg ctg	2196
	Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu	
	340 345 350	
	cag ccc tct cct gca ctg tct tgg aga tct aca ggc ggc atc ctg gac	2244
	Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp	
	355 360 365	

[0054]	gtg tac atc ttt ctg ggc ccc gag cct aag agc gtg gtg cag cag tat	2292
	Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr	
	370 375 380	
	ctg gac gtc gtg ggc tac ccc ttc atg cct cct tat tgg ggc ctg ggc	2340
	Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly	
	385 390 395 400	
	ttc cac ctg tgt agg tgg ggc tac agc agc acc gcc atc acc aga cag	2388
	Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln	
	405 410 415	
	gtg gtg gaa aac atg acc cgg gct cac ttc cca ctg gac gtg cag tgg	2436
	Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp	
	420 425 430	
	aac gac ctg gac tac atg gac agc aga cgg gac ttc acc ttc aac aag	2484
	Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys	
	435 440 445	
	gac ggc ttc aga gac ttc ccc gcc atg gtg caa gag ctg cat caa ggc	2532
	Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly	
	450 455 460	
	gga cgg cgg tac atg atg att gtg gac cct gcc atc agc agc tct gga	2580
	Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly	
	465 470 475 480	
	cca gcc ggc agc tac aga cct tac gat gag gga ctg aga aga ggc gtg	2628
	Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val	
	485 490 495	
	ttc atc acc aac gag aca ggc cag cct ctg atc ggc aaa gtg tgg cct	2676
	Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro	
	500 505 510	
	ggc agc aca gcc ttt cca gac ttc aca aac ccc acc gct ctg gct tgg	2724
	Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp	
	515 520 525	
	tgg gaa gat atg gtg gcc gag ttc cac gat cag gtg ccc ttc gat ggc	2772
	Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly	
	530 535 540	
	atg tgg atc gac atg aac gag ccc agc aac ttc atc cgg ggc agc gag	2820
	Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu	
	545 550 555 560	
	gat ggc tgc ccc aac aac gaa cta gaa aat cct cct tac gtg ccc ggc	2868
	Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly	
	565 570 575	
	gtt gtc ggc gga aca ctt cag gcc gct aca atc tgt gcc agc agc cat	2916
	Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His	
	580 585 590	
	cag ttt ctg agc acc cac tac aac ctg cac aac ctg tac ggc ctg acc	2964
	Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr	
	595 600 605	
	gag gcc att gcc tct cat aga gcc ctg gtt aag gcc aga ggc acc cgg	3012
	Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg	
	610 615 620	
	cct ttt gtg atc agc aga agc aca ttc gcc ggc cac ggc aga tac gca	3060
	Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala	
	625 630 635 640	

[0055]	gga cat tgg aca ggc gac gtg tgg tct agt tgg gag cag ctg gct agc	3108
	Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser	
	645 650 655	
	agc gtg cca gag atc ctg cag ttc aat ctg ctg ggc gtg cca ctc gtg	3156
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val	
	660 665 670	
	gga gcc gac gtt tgt ggc ttc ctg ggc aac acc agc gag gaa ctg tgt	3204
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys	
	675 680 685	
	gtg cgt tgg aca cag ctg ggc gcc ttc tat ccc ttc atg aga aac cac	3252
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His	
	690 695 700	
	aac agc ctg ctg agc ctg cct caa gag ccc tac agc ttt agc gag cct	3300
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro	
	705 710 715 720	
	gca cag cag gcc atg aga aag gcc ctg act ctg aga tac gcc ctg ctg	3348
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu	
	725 730 735	
	cct cac ctg tac acc ctg ttt cat cag gcc cac gtg gca ggc gag aca	3396
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr	
	740 745 750	
	gtg gct aga cct ctg ttc ctg gaa ttt ccc aag gac agc tcc acc tgg	3444
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp	
	755 760 765	
	acc gtg gat cat cag ctg ctg tgg gga gaa gcc ctg ctg att aca cca	3492
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro	
	770 775 780	
	gtg ctg cag gcc gga aag gcc gaa gtg aca ggc tat ttc cct ctc ggc	3540
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly	
	785 790 795 800	
	act tgg tac gac ctg cag acc gtg cct atc gag gct ctg gga tct ctt	3588
	Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu	
	805 810 815	
	cct cca cct cct gcc gct cct aga gag cct gcc att cac tct gaa ggc	3636
	Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly	
	820 825 830	
	cag tgg gtt acc ctg cct gct cct ctg gac acc atc aac gtg cac ctg	3684
	Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu	
	835 840 845	
	aga gcc ggc tac atc atc cct ctg caa ggc cct ggc ctg acc aca acc	3732
	Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr	
	850 855 860	
	gaa tct aga cag cag ccc atg gca ctg gcc gtg gct ctt aca aaa ggc	3780
	Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly	
	865 870 875 880	
	gga gag gct aga ggc gag ctg ttc tgg gat gat ggc gag agc cta gaa	3828
	Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu	
	885 890 895	
	gtg ctg gaa cgg ggc gct tat acc caa gtg atc ttc ctg gcc aga aac	3876
	Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn	
	900 905 910	

	aac acc atc gtg aac gaa ctc gtg cgc gtg acc agt gaa ggt gct gga Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly 915 920 925	3924
	ctg caa ctg cag aaa gtg acc gtg ctc gga gtg gcc aca gct cct cag Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln 930 935 940	3972
	cag gtt ctg tct aat ggc gtg ccc gtg tcc aac ttc aca tac agc ccc Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro 945 950 955 960	4020
	gac acc aag gtc ctg gac atc tgt gtg tcc ctg ctt atg ggc gag cag Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln 965 970 975	4068
	ttc ctg gtg tcc tgg tgc tga taa gcggcgcac gcgtggtacc tctagagtcg Phe Leu Val Ser Trp Cys 980	4122
	accggggcgg cctcgaggac ggggtgaact acgcctgaga tctttttccc tctgccaaaa attatgggga catcatgaag ccccttgagc atctgacttc tggctaataa aggaaattta ttttcattgc aatagtgtgt tggaattttt tgtgtctctc actcggttaa caacaacaat tgcatcatt ttatgtttca ggttcagggg gagatgtggg aggtttttta aagcaagtaa aacctctaca aatgtggtaa aatcgataag gatcttccta gagcatggct acgtagataa gtagcatggc gggttaatca ttaactacaa ggaacccta gtgatggagt tggccactcc ctctctgcgc gctcgcctgc tcaactgaggc cgggcgacca aaggctgccc gacgcccggg ctttgcccg ggcgcctcag tgagcgagcg agcgcgcgag	4182 4242 4302 4362 4422 4482 4542 4581
	<210> 29	
	<211> 982	
[0056]	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 合成构建体	
	<400> 29	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala 1 5 10 15 Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln 20 25 30 Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg 35 40 45 Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser 50 55 60 Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser 65 70 75 80 Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg 85 90 95 Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys 100 105 110 Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile 115 120 125 Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys 130 135 140 Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro 145 150 155 160 Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	

						165					170					175
	Gly	Tyr	Thr	Ala	Thr	Leu	Thr	Arg	Thr	Thr	Pro	Thr	Phe	Phe	Pro	Lys
						180					185					190
	Asp	Ile	Leu	Thr	Leu	Arg	Leu	Asp	Val	Met	Met	Glu	Thr	Glu	Asn	Arg
						195					200					205
	Leu	His	Phe	Thr	Ile	Lys	Asp	Pro	Ala	Asn	Arg	Arg	Tyr	Glu	Val	Pro
						210					215					220
	Leu	Glu	Thr	Pro	His	Val	His	Ser	Arg	Ala	Pro	Ser	Pro	Leu	Tyr	Ser
	225					230					235					240
	Val	Glu	Phe	Ser	Glu	Glu	Pro	Phe	Gly	Val	Ile	Val	Arg	Arg	Gln	Leu
						245					250					255
	Asp	Gly	Arg	Val	Leu	Leu	Asn	Thr	Thr	Val	Ala	Pro	Leu	Phe	Phe	Ala
						260					265					270
	Asp	Gln	Phe	Leu	Gln	Leu	Ser	Thr	Ser	Leu	Pro	Ser	Gln	Tyr	Ile	Thr
						275					280					285
	Gly	Leu	Ala	Glu	His	Leu	Ser	Pro	Leu	Met	Leu	Ser	Thr	Ser	Trp	Thr
						290					295					300
	Arg	Ile	Thr	Leu	Trp	Asn	Arg	Asp	Leu	Ala	Pro	Thr	Pro	Gly	Ala	Asn
	305					310					315					320
	Leu	Tyr	Gly	Ser	His	Pro	Phe	Tyr	Leu	Ala	Leu	Glu	Asp	Gly	Gly	Ser
						325					330					335
	Ala	His	Gly	Val	Phe	Leu	Leu	Asn	Ser	Asn	Ala	Met	Asp	Val	Val	Leu
						340					345					350
	Gln	Pro	Ser	Pro	Ala	Leu	Ser	Trp	Arg	Ser	Thr	Gly	Gly	Ile	Leu	Asp
						355					360					365
[0057]	Val	Tyr	Ile	Phe	Leu	Gly	Pro	Glu	Pro	Lys	Ser	Val	Val	Gln	Gln	Tyr
						370					375					380
	Leu	Asp	Val	Val	Gly	Tyr	Pro	Phe	Met	Pro	Pro	Tyr	Trp	Gly	Leu	Gly
	385					390					395					400
	Phe	His	Leu	Cys	Arg	Trp	Gly	Tyr	Ser	Ser	Thr	Ala	Ile	Thr	Arg	Gln
						405					410					415
	Val	Val	Glu	Asn	Met	Thr	Arg	Ala	His	Phe	Pro	Leu	Asp	Val	Gln	Trp
						420					425					430
	Asn	Asp	Leu	Asp	Tyr	Met	Asp	Ser	Arg	Arg	Asp	Phe	Thr	Phe	Asn	Lys
						435					440					445
	Asp	Gly	Phe	Arg	Asp	Phe	Pro	Ala	Met	Val	Gln	Glu	Leu	His	Gln	Gly
						450					455					460
	Gly	Arg	Arg	Tyr	Met	Met	Ile	Val	Asp	Pro	Ala	Ile	Ser	Ser	Ser	Gly
	465					470					475					480
	Pro	Ala	Gly	Ser	Tyr	Arg	Pro	Tyr	Asp	Glu	Gly	Leu	Arg	Arg	Gly	Val
						485					490					495
	Phe	Ile	Thr	Asn	Glu	Thr	Gly	Gln	Pro	Leu	Ile	Gly	Lys	Val	Trp	Pro
						500					505					510
	Gly	Ser	Thr	Ala	Phe	Pro	Asp	Phe	Thr	Asn	Pro	Thr	Ala	Leu	Ala	Trp
						515					520					525
	Trp	Glu	Asp	Met	Val	Ala	Glu	Phe	His	Asp	Gln	Val	Pro	Phe	Asp	Gly
						530					535					540
	Met	Trp	Ile	Asp	Met	Asn	Glu	Pro	Ser	Asn	Phe	Ile	Arg	Gly	Ser	Glu
	545					550					555					560
	Asp	Gly	Cys	Pro	Asn	Asn	Glu	Leu	Glu	Asn	Pro	Pro	Tyr	Val	Pro	Gly
						565					570					575

[0058]

Val	Val	Gly	Gly	Thr	Leu	Gln	Ala	Ala	Thr	Ile	Cys	Ala	Ser	Ser	His	580	585	590
Gln	Phe	Leu	Ser	Thr	His	Tyr	Asn	Leu	His	Asn	Leu	Tyr	Gly	Leu	Thr	595	600	605
Glu	Ala	Ile	Ala	Ser	His	Arg	Ala	Leu	Val	Lys	Ala	Arg	Gly	Thr	Arg	610	615	620
Pro	Phe	Val	Ile	Ser	Arg	Ser	Thr	Phe	Ala	Gly	His	Gly	Arg	Tyr	Ala	625	630	635
Gly	His	Trp	Thr	Gly	Asp	Val	Trp	Ser	Ser	Trp	Glu	Gln	Leu	Ala	Ser	645	650	655
Ser	Val	Pro	Glu	Ile	Leu	Gln	Phe	Asn	Leu	Leu	Gly	Val	Pro	Leu	Val	660	665	670
Gly	Ala	Asp	Val	Cys	Gly	Phe	Leu	Gly	Asn	Thr	Ser	Glu	Glu	Leu	Cys	675	680	685
Val	Arg	Trp	Thr	Gln	Leu	Gly	Ala	Phe	Tyr	Pro	Phe	Met	Arg	Asn	His	690	695	700
Asn	Ser	Leu	Leu	Ser	Leu	Pro	Gln	Glu	Pro	Tyr	Ser	Phe	Ser	Glu	Pro	705	710	715
Ala	Gln	Gln	Ala	Met	Arg	Lys	Ala	Leu	Thr	Leu	Arg	Tyr	Ala	Leu	Leu	725	730	735
Pro	His	Leu	Tyr	Thr	Leu	Phe	His	Gln	Ala	His	Val	Ala	Gly	Glu	Thr	740	745	750
Val	Ala	Arg	Pro	Leu	Phe	Leu	Glu	Phe	Pro	Lys	Asp	Ser	Ser	Thr	Trp	755	760	765
Thr	Val	Asp	His	Gln	Leu	Leu	Trp	Gly	Glu	Ala	Leu	Leu	Ile	Thr	Pro	770	775	780
Val	Leu	Gln	Ala	Gly	Lys	Ala	Glu	Val	Thr	Gly	Tyr	Phe	Pro	Leu	Gly	785	790	795
Thr	Trp	Tyr	Asp	Leu	Gln	Thr	Val	Pro	Ile	Glu	Ala	Leu	Gly	Ser	Leu	805	810	815
Pro	Pro	Pro	Pro	Ala	Ala	Pro	Arg	Glu	Pro	Ala	Ile	His	Ser	Glu	Gly	820	825	830
Gln	Trp	Val	Thr	Leu	Pro	Ala	Pro	Leu	Asp	Thr	Ile	Asn	Val	His	Leu	835	840	845
Arg	Ala	Gly	Tyr	Ile	Ile	Pro	Leu	Gln	Gly	Pro	Gly	Leu	Thr	Thr	Thr	850	855	860
Glu	Ser	Arg	Gln	Gln	Pro	Met	Ala	Leu	Ala	Val	Ala	Leu	Thr	Lys	Gly	865	870	875
Gly	Glu	Ala	Arg	Gly	Glu	Leu	Phe	Trp	Asp	Asp	Gly	Glu	Ser	Leu	Glu	885	890	895
Val	Leu	Glu	Arg	Gly	Ala	Tyr	Thr	Gln	Val	Ile	Phe	Leu	Ala	Arg	Asn	900	905	910
Asn	Thr	Ile	Val	Asn	Glu	Leu	Val	Arg	Val	Thr	Ser	Glu	Gly	Ala	Gly	915	920	925
Leu	Gln	Leu	Gln	Lys	Val	Thr	Val	Leu	Gly	Val	Ala	Thr	Ala	Pro	Gln	930	935	940
Gln	Val	Leu	Ser	Asn	Gly	Val	Pro	Val	Ser	Asn	Phe	Thr	Tyr	Ser	Pro	945	950	955
Asp	Thr	Lys	Val	Leu	Asp	Ile	Cys	Val	Ser	Leu	Leu	Met	Gly	Glu	Gln	965	970	975
Phe	Leu	Val	Ser	Trp	Cys													

	980
	<210> 30
	<211> 4687
	<212> DNA
	<213> 人工序列 (Artificial Sequence)
	<220>
	<223> 合成构建体
	<220>
	<221> misc_feature
	<222> (1)..(130)
	<223> 5' ITR
	<220>
	<221> enhancer
	<222> (195)..(437)
	<223> CMV IE 增强子
	<220>
	<221> promoter
	<222> (440)..(721)
	<223> 鸡 β -肌动蛋白启动子
	<220>
	<221> Intron
	<222> (721)..(1128)
	<223> CAG 中的杂交内含子
	<220>
	<221> CDS
[0059]	<222> (1141)..(4092)
	<223> BiP-vIGF2-hGAAco
	<220>
	<221> misc_feature
	<222> (3568)..(3570)
	<223> Ile 密码子
	<220>
	<221> misc_feature
	<222> (4113)..(4134)
	<223> miR-183 靶
	<220>
	<221> misc_feature
	<222> (4139)..(4160)
	<223> miR-183 靶
	<220>
	<221> misc_feature
	<222> (4167)..(4188)
	<223> miR-183 靶
	<220>
	<221> misc_feature
	<222> (4195)..(4216)
	<223> miR-183 靶
	<220>
	<221> polyA_signal
	<222> (4267)..(4393)
	<223> 兔 β -珠蛋白 poly a

<220>		
<221>	misc_feature	
<222>	(4558)..(4687)	
<223>	3' ITR	
<400>	30	
ctgcgcgtc	gctcgtcac	tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg	cctcagttag	cgagcgagcg cgagagagg gagtggccaa ctccatcact 120
aggggttcct	tgtagttaat	gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg	taccattgac	gtcaataatg acgtatgttc ccatagtaac gccaataggg 240
actttccatt	gacgtcaatg	ggtggagtat ttacggtaaa ctgccactt ggcagtacat 300
caagtgtatc	atatgccaag	tacgccccct attgacgtca atgacgtaa atggcccgcc 360
tggcattatg	cccagtacat	gaccttatgg gactttccta cttggcagta catctacgta 420
ttagtcacg	ctattacat	ggtcgaggtg agccccacgt tctgcttcac tctccccatc 480
tccccccct	ccccacccc	aattttgtat ttattttttt ttttaattatt ttgtgcagcg 540
atgggggcgg	gggggggggg	ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg 600
ggcggggcga	ggcggagagg	tgccggcgga gccaatcaga gcggcgcgct ccgaaagttt 660
ccttttatgg	cgaggcgcg	gcggcgcgcg ccctataaaa agcgaagcgc gcggcgggcg 720
ggagtcgtg	cgcgtgcct	tcgccccgtg ccccgctccg ccgcccctc gcggcgcccg 780
ccccggctct	gactgaccgc	gttactccca caggtagcgc gcggggacgg cccttctcct 840
ccgggctgta	attagcgctt	ggtttaatga cggttgttt cttttctgtg gctgcgtgaa 900
agccttgagg	ggctccggga	gggccccttg tcggggggga gcggctcggg gctgtccgcg 960
gggggacggc	tgccttcggg	ggggacgggg cagggcgggg ttgggttctt ggctgtgac 1020
cgcggtctct	agagcctctg	ctaaccatgt tcatgcctc ttctttttcc tacagctcct 1080
gggcaacgtg	ctggttattg	tgtgtctca tcattttggc aaagaattgg atccgccacc 1140
atg aag ctg tct ctg gtg gct gct atg ctg ctg ctc ctg tct gcc gcc		1188
[0060]	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
1	5	10 15
aga gcc agc aga aca ctt tgt ggc gga gag ctg gtg gac acc ctg cag		1236
Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln		
20	25	30
ttt gtg tgt ggc gac aga ggc ttc ctg ttc agc aga cct gcc agc cgg		1284
Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg		
35	40	45
ggt tcc aga cgg tct aga gga atc gtg gaa gag tgc tgc ttc aga agc		1332
Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser		
50	55	60
tgc gat ctg gcc ctg ctg gaa acc tac tgt gcc aca cca gcc aga tct		1380
Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser		
65	70	75 80
gaa ggc ggc gga gga tct ggc gga ggc gga tct aga cct gga cct aga		1428
Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg		
85	90	95
gac gcc cag gct cac cct ggt aga cct aga gct gtg cct aca cag tgc		1476
Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys		
100	105	110
gac gtg cca cct aac agc aga ttc gac tgc gcc cct gac aag gcc atc		1524
Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile		
115	120	125
aca caa gag cag tgt gaa gcc aga ggc tgc tgc tac atc cct gcc aaa		1572
Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys		
130	135	140

	caa gga ctg cag ggc gcc cag atg gga cag cct tgg tgc ttc ttc cca	1620
	Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro	
	145 150 155 160	
	cca tct tac ccc agc tac aag ctg gaa aac ctg agc agc tcc gag atg	1668
	Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	
	165 170 175	
	ggc tac acc gcc aca ctg acc aga acc aca cct aca ttc ttc ccg aag	1716
	Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys	
	180 185 190	
	gac atc ctg aca ctg cgg ctg gac gtg atg atg gaa acc gag aac cgg	1764
	Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg	
	195 200 205	
	ctg cac ttc acc atc aag gac ccc gcc aat cgg aga tac gag gtg ccc	1812
	Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro	
	210 215 220	
	ctg gaa aca ccc cac gtg cac tct aga gca ccc tct cca ctg tac agc	1860
	Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser	
	225 230 235 240	
	gtg gaa ttt tcc gag gaa ccc ttc ggc gtg atc gtg cgg aga cag ctg	1908
	Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu	
	245 250 255	
	gat ggc aga gtg ctc ctg aat acc aca gtg gcc cct ctg ttc ttc gcc	1956
	Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala	
	260 265 270	
[0061]	gac cag ttt ctg cag ctg agc acc agc ctg cct agc cag tat atc aca	2004
	Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr	
	275 280 285	
	ggc ctg gcc gag cat ctg agc cct ctg atg ctg agc aca tcc tgg acc	2052
	Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr	
	290 295 300	
	aga atc acc ctg tgg aac cgc gac ctg gct cct aca cct ggc gcc aat	2100
	Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn	
	305 310 315 320	
	ctg tac ggc tct cac ccc ttc tac ctg gca ctg gaa gac ggt gga tct	2148
	Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser	
	325 330 335	
	gcc cac ggt gtc ttt ctg ctg aat agc aac gcc atg gac gtg gtg ctg	2196
	Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu	
	340 345 350	
	cag ccc tct cct gca ctg tct tgg aga tct aca ggc ggc atc ctg gac	2244
	Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp	
	355 360 365	
	gtg tac atc ttt ctg ggc ccc gag cct aag agc gtg gtg cag cag tat	2292
	Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr	
	370 375 380	
	ctg gac gtc gtg ggc tac ccc ttc atg cct cct tat tgg ggc ctg ggc	2340
	Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly	
	385 390 395 400	
	ttc cac ctg tgt agg tgg ggc tac agc agc acc gcc atc acc aga cag	2388
	Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln	
	405 410 415	

[0062]	gtg gtg gaa aac atg acc cgg gct cac ttc cca ctg gac gtg cag tgg	2436
	Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp	
	420 425 430	
	aac gac ctg gac tac atg gac agc aga cgg gac ttc acc ttc aac aag	2484
	Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys	
	435 440 445	
	gac ggc ttc aga gac ttc ccc gcc atg gtg caa gag ctg cat caa ggc	2532
	Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly	
	450 455 460	
	gga cgg cgg tac atg atg att gtg gac cct gcc atc agc agc tct gga	2580
	Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly	
	465 470 475 480	
	cca gcc ggc agc tac aga cct tac gat gag gga ctg aga aga ggc gtg	2628
	Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val	
	485 490 495	
	ttc atc acc aac gag aca ggc cag cct ctg atc ggc aaa gtg tgg cct	2676
	Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro	
	500 505 510	
	ggc agc aca gcc ttt cca gac ttc aca aac ccc acc gct ctg gct tgg	2724
	Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp	
	515 520 525	
	tgg gaa gat atg gtg gcc gag ttc cac gat cag gtg ccc ttc gat ggc	2772
	Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly	
	530 535 540	
	atg tgg atc gac atg aac gag ccc agc aac ttc atc cgg ggc agc gag	2820
	Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu	
	545 550 555 560	
	gat ggc tgc ccc aac aac gaa cta gaa aat cct cct tac gtg ccc ggc	2868
	Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly	
	565 570 575	
	gtt gtc ggc gga aca ctt cag gcc gct aca atc tgt gcc agc agc cat	2916
	Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His	
	580 585 590	
	cag ttt ctg agc acc cac tac aac ctg cac aac ctg tac ggc ctg acc	2964
	Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr	
	595 600 605	
	gag gcc att gcc tct cat aga gcc ctg gtt aag gcc aga ggc acc cgg	3012
	Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg	
	610 615 620	
	cct ttt gtg atc agc aga agc aca ttc gcc ggc cac ggc aga tac gca	3060
	Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala	
	625 630 635 640	
	gga cat tgg aca ggc gac gtg tgg tct agt tgg gag cag ctg gct agc	3108
	Gly His Trp Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser	
	645 650 655	
	agc gtg cca gag atc ctg cag ttc aat ctg ctg ggc gtg cca ctc gtg	3156
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val	
	660 665 670	
	gga gcc gac gtt tgt ggc ttc ctg ggc aac acc agc gag gaa ctg tgt	3204
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys	
	675 680 685	

	gtg cgt tgg aca cag ctg ggc gcc ttc tat ccc ttc atg aga aac cac	3252
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His	
	690 695 700	
	aac agc ctg ctg agc ctg cct caa gag ccc tac agc ttt agc gag cct	3300
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro	
	705 710 715 720	
	gca cag cag gcc atg aga aag gcc ctg act ctg aga tac gcc ctg ctg	3348
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu	
	725 730 735	
	cct cac ctg tac acc ctg ttt cat cag gcc cac gtg gca ggc gag aca	3396
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr	
	740 745 750	
	gtg gct aga cct ctg ttc ctg gaa ttt ccc aag gac agc tcc acc tgg	3444
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp	
	755 760 765	
	acc gtg gat cat cag ctg ctg tgg gga gaa gcc ctg ctg att aca cca	3492
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro	
	770 775 780	
	gtg ctg cag gcc gga aag gcc gaa gtg aca ggc tat ttc cct ctc ggc	3540
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly	
	785 790 795 800	
	act tgg tac gac ctg cag acc gtg cct atc gag gct ctg gga tct ctt	3588
	Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu	
	805 810 815	
[0063]	cct cca cct cct gcc gct cct aga gag cct gcc att cac tct gaa ggc	3636
	Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly	
	820 825 830	
	cag tgg gtt acc ctg cct gct cct ctg gac acc atc aac gtg cac ctg	3684
	Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu	
	835 840 845	
	aga gcc ggc tac atc atc cct ctg caa ggc cct ggc ctg acc aca acc	3732
	Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr	
	850 855 860	
	gaa tct aga cag cag ccc atg gca ctg gcc gtg gct ctt aca aaa ggc	3780
	Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly	
	865 870 875 880	
	gga gag gct aga ggc gag ctg ttc tgg gat gat ggc gag agc cta gaa	3828
	Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu	
	885 890 895	
	gtg ctg gaa cgg ggc gct tat acc caa gtg atc ttc ctg gcc aga aac	3876
	Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn	
	900 905 910	
	aac acc atc gtg aac gaa ctc gtg cgc gtg acc agt gaa ggt gct gga	3924
	Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly	
	915 920 925	
	ctg caa ctg cag aaa gtg acc gtg ctc gga gtg gcc aca gct cct cag	3972
	Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln	
	930 935 940	
	cag gtt ctg tct aat ggc gtg ccc gtg tcc aac ttc aca tac agc ccc	4020
	Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro	
	945 950 955 960	

	gac acc aag gtc ctg gac atc tgt gtg tcc ctg ctt atg ggc gag cag	4068
	Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln	
	965 970 975	
	ttc ctg gtg tcc tgg tgc tga taa gcggccgcac gcgtggtacc agtgaattct	4122
	Phe Leu Val Ser Trp Cys	
	980	
	accagtgcc taggatagtg aattctacca gtgccataca cgtgagtga ttctaccagt	4182
	gccatagcat gcagtgaatt ctaccagtgc catagcggcc gcctcgaccc gggcggcctc	4242
	gaggacgggg tgaactacgc ctgagatctt ttccctctg ccaaaaatta tggggacatc	4302
	atgaagcccc ttgagcatct gacttctggc taataaagga aatttatttt cattgcaata	4362
	gtgtgttga atttttgtg tctctcactc ggtaacaac aacaattgca ttcattttat	4422
	gtttcaggtt cagggggaga tgtgggaggt tttttaagc aagtaaaacc tctacaaatg	4482
	tggtaaaatc gataaggatc ttcctagagc atggctacgt agataagtag catggcgggt	4542
	taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct ctgcgcgtc	4602
	gctcgtcac tgaggcggg cgaccaaagg tcgcccgcg cccgggcttt gcccgggcgg	4662
	cctcagtgag cgagcgagcg cgcag	4687
	<210> 31	
	<211> 982	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 合成构建体	
	<400> 31	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
[0064]	Arg Ala Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln	
	20 25 30	
	Phe Val Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg	
	35 40 45	
	Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser	
	50 55 60	
	Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser	
	65 70 75 80	
	Glu Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg Pro Gly Pro Arg	
	85 90 95	
	Asp Ala Gln Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys	
	100 105 110	
	Asp Val Pro Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile	
	115 120 125	
	Thr Gln Glu Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys	
	130 135 140	
	Gln Gly Leu Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro	
	145 150 155 160	
	Pro Ser Tyr Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met	
	165 170 175	
	Gly Tyr Thr Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys	
	180 185 190	
	Asp Ile Leu Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg	
	195 200 205	
	Leu His Phe Thr Ile Lys Asp Pro Ala Asn Arg Arg Tyr Glu Val Pro	
	210 215 220	

[0065]

```

Leu Glu Thr Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser
225                230                235                240
Val Glu Phe Ser Glu Glu Pro Phe Gly Val Ile Val Arg Arg Gln Leu
                245                250                255
Asp Gly Arg Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala
                260                265                270
Asp Gln Phe Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr
                275                280                285
Gly Leu Ala Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr
                290                295                300
Arg Ile Thr Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn
305                310                315                320
Leu Tyr Gly Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser
                325                330                335
Ala His Gly Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu
                340                345                350
Gln Pro Ser Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp
                355                360                365
Val Tyr Ile Phe Leu Gly Pro Glu Pro Lys Ser Val Val Gln Gln Tyr
                370                375                380
Leu Asp Val Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly
385                390                395                400
Phe His Leu Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln
                405                410                415
Val Val Glu Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp
                420                425                430
Asn Asp Leu Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys
                435                440                445
Asp Gly Phe Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly
                450                455                460
Gly Arg Arg Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly
465                470                475                480
Pro Ala Gly Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val
                485                490                495
Phe Ile Thr Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro
                500                505                510
Gly Ser Thr Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp
                515                520                525
Trp Glu Asp Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly
                530                535                540
Met Trp Ile Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu
545                550                555                560
Asp Gly Cys Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly
                565                570                575
Val Val Gly Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His
                580                585                590
Gln Phe Leu Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr
                595                600                605
Glu Ala Ile Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg
                610                615                620
Pro Phe Val Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala

```

	625		630		635		640
	Gly His Trp Thr	Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser					
		645		650		655	
	Ser Val Pro Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val						
		660		665		670	
	Gly Ala Asp Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys						
		675		680		685	
	Val Arg Trp Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His						
		690		695		700	
	Asn Ser Leu Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro						
	705		710		715		720
	Ala Gln Gln Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu						
		725		730		735	
	Pro His Leu Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr						
		740		745		750	
	Val Ala Arg Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp						
		755		760		765	
	Thr Val Asp His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro						
		770		775		780	
	Val Leu Gln Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly						
	785		790		795		800
	Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu Gly Ser Leu						
		805		810		815	
	Pro Pro Pro Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly						
		820		825		830	
[0066]	Gln Trp Val Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu						
		835		840		845	
	Arg Ala Gly Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr						
		850		855		860	
	Glu Ser Arg Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly						
	865		870		875		880
	Gly Glu Ala Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu						
		885		890		895	
	Val Leu Glu Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn						
		900		905		910	
	Asn Thr Ile Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly						
		915		920		925	
	Leu Gln Leu Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln						
		930		935		940	
	Gln Val Leu Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro						
	945		950		955		960
	Asp Thr Lys Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln						
		965		970		975	
	Phe Leu Val Ser Trp Cys						
		980					
	<210>	32					
	<211>	67					
	<212>	PRT					
	<213>	智人 (Homo sapiens)					
	<400>	32					
	Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr						

```

1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
65
<210> 33
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 F26S
<400> 33
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
65
<210> 34
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 Y27L
<400> 34
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
65
<210> 35
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> V43L
<400> 35
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr

```

[0067]

```

1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
      20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Leu Glu Glu Cys Cys Phe
      35           40           45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
      50           55           60
Lys Ser Glu
65
<210> 36
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 F48T
<400> 36
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
      20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Thr
      35           40           45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
      50           55           60
Lys Ser Glu
65
<210> 37
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 R49S
<400> 37
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
      20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
      35           40           45
Ser Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
      50           55           60
Lys Ser Glu
65
<210> 38
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 S50I
<400> 38
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr

```

[0068]

```

1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ile Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
65
<210> 39
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 A54R
<400> 39
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ser Cys Asp Leu Arg Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
[0069] 65
<210> 40
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 L55R
<400> 40
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1           5           10           15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala
           20           25           30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
           35           40           45
Arg Ser Cys Asp Leu Ala Arg Leu Glu Thr Tyr Cys Ala Thr Pro Ala
           50           55           60
Lys Ser Glu
65
<210> 41
<211> 67
<212> PRT
<213> 人工序列 (Artificial Sequence)
<220>
<223> IGF2 F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R
<400> 41
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr

```

1 5 10 15
 Leu Gln Phe Val Cys Gly Asp Arg Gly Ser Leu Phe Ser Arg Pro Ala
 20 25 30
 Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Leu Glu Glu Cys Cys Thr
 35 40 45
 Ser Ile Cys Asp Leu Arg Arg Leu Glu Thr Tyr Cys Ala Thr Pro Ala
 50 55 60
 Lys Ser Glu
 65
 <210> 42
 <211> 61
 <212> PRT
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> IGF2 Δ1-6、Y27L、K65R
 <400> 42
 Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly
 1 5 10 15
 Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg
 20 25 30
 Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala
 35 40 45
 Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser Glu
 50 55 60
 <210> 43
 <211> 60
 <212> PRT
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> IGF2 Δ1-7、Y27L、K65R
 <400> 43
 Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly Asp
 1 5 10 15
 Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser
 20 25 30
 Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu
 35 40 45
 Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser Glu
 50 55 60
 <210> 44
 <211> 63
 <212> PRT
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> IGF2 Δ1-4、E6R、Y27L、K65R
 <400> 44
 Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val
 1 5 10 15
 Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg Val Ser
 20 25 30
 Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp

	35	40	45
	Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Arg Ser Glu		
	50	55	60
<210>	45		
<211>	63		
<212>	PRT		
<213>	人工序列 (Artificial Sequence)		
<220>			
<223>	IGF2 Δ 1-4、E6R、Y27L		
<400>	45		
	Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val		
	1 5 10 15		
	Cys Gly Asp Arg Gly Phe Leu Phe Ser Arg Pro Ala Ser Arg Val Ser		
	20 25 30		
	Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp		
	35 40 45		
	Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu		
	50 55 60		
<210>	46		
<211>	67		
<212>	PRT		
<213>	人工序列 (Artificial Sequence)		
<220>			
<223>	IGF2 E6R		
<400>	46		
[0071]	Ala Tyr Arg Pro Ser Arg Thr Leu Cys Gly Gly Glu Leu Val Asp Thr		
	1 5 10 15		
	Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala		
	20 25 30		
	Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe		
	35 40 45		
	Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala		
	50 55 60		
	Lys Ser Glu		
	65		
<210>	47		60
<211>	201		120
<212>	DNA		180
<213>	智人 (Homo sapiens)		201
<400>	47		
	gcttaccgcc ccagtgagac cctgtgcggc ggggagctgg tggacaccct ccagttcgtc		60
	tgtggggacc gcggcttcta cttcagcagg cccgcaagcc gtgtgagccg tcgcagccgt		120
	ggcatcgttg aggagtgtg tttccgcagc tgtgacctgg ccctcctgga gacgtactgt		180
	gctacccccg ccaagtccga g		201
<210>	48		
<211>	189		
<212>	DNA		
<213>	人工序列 (Artificial Sequence)		
<220>			
<223>	vIGF2 Δ 1-4、E6R、Y27L、K65R		
<400>	48		

	tctagaacac tgtgcggagg ggagcttgta gacactcttc agttcgtgtg tggagatcgc	60
	gggttcctct tctctgccc cgcttcaga gtttcacgga ggtctagggg tatagtagag	120
	gagtgttggt tcaggtcctg tgacttggcg ctctcgaga cctattgcgc gacgccagcc	180
	aggtccgaa	189
	<210> 49	
	<211> 18	
	<212> PRT	
	<213> 智人 (Homo sapiens)	
	<400> 49	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Ala Ala	
	1 5 10 15	
	Arg Ala	
	<210> 50	
	<211> 28	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 经修饰的 BiP-1	
	<400> 50	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Leu Val	
	1 5 10 15	
	Ala Ala Met Leu Leu Leu Leu Ser Ala Ala Arg Ala	
	20 25	
	<210> 51	
[0072]	<211> 25	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 经修饰的 BiP-2	
	<400> 51	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Trp Val Ala	
	1 5 10 15	
	Leu Leu Leu Leu Ser Ala Ala Arg Ala	
	20 25	
	<210> 52	
	<211> 26	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	
	<223> 经修饰的 BiP-3	
	<400> 52	
	Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ser Leu Val	
	1 5 10 15	
	Ala Leu Leu Leu Leu Ser Ala Ala Arg Ala	
	20 25	
	<210> 53	
	<211> 26	
	<212> PRT	
	<213> 人工序列 (Artificial Sequence)	
	<220>	

<223> 经修饰的 BiP-4

<400> 53

Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Ala Leu Val
1 5 10 15
Ala Leu Leu Leu Ser Ala Ala Arg Ala
20 25

<210> 54

<211> 17

<212> PRT

<213> Gaussia princeps

<400> 54

Met Gly Val Lys Val Leu Phe Ala Leu Ile Cys Ile Ala Val Ala Glu
1 5 10 15
Ala

<210> 55

<211> 9

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 接头序列

<400> 55

Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5

<210> 56

<211> 5

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 接头序列

<400> 56

Gly Gly Gly Gly Ser
1 5

<210> 57

<211> 9

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 接头序列

<400> 57

Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5

<210> 58

<211> 9

<212> PRT

<213> 人工序列 (Artificial Sequence)

<220>

<223> 接头序列

<400> 58

Gly Gly Gly Gly Ser Gly Gly Gly Ser
1 5

[0073]

<210> 59
 <211> 9
 <212> PRT
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> 接头序列
 <400> 59
 Gly Gly Ser Gly Ser Gly Ser Thr Ser
 1 5
 <210> 60
 <211> 10
 <212> PRT
 <213> 人工序列 (Artificial Sequence)
 <220>
 <223> 接头序列
 <400> 60
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10
 <210> 61
 <211> 6
 <212> PRT
 <213> 智人 (Homo sapiens)
 <400> 61
 [0074] Trp Ile Asp Met Asn Glu
 1 5
 <210> 62
 <211> 7
 <212> PRT
 <213> 智人 (Homo sapiens)
 <400> 62
 Thr Val Pro Ile Glu Ala Leu
 1 5
 <210> 63
 <211> 9
 <212> PRT
 <213> 智人 (Homo sapiens)
 <400> 63
 Gln Thr Val Pro Ile Glu Ala Leu Gly
 1 5
 <210> 64
 <211> 25
 <212> PRT
 <213> 智人 (Homo sapiens)
 <400> 64
 Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro Ile Glu Ala Leu
 1 5 10 15
 Gly Ser Leu Pro Pro Pro Pro Ala Ala
 20 25

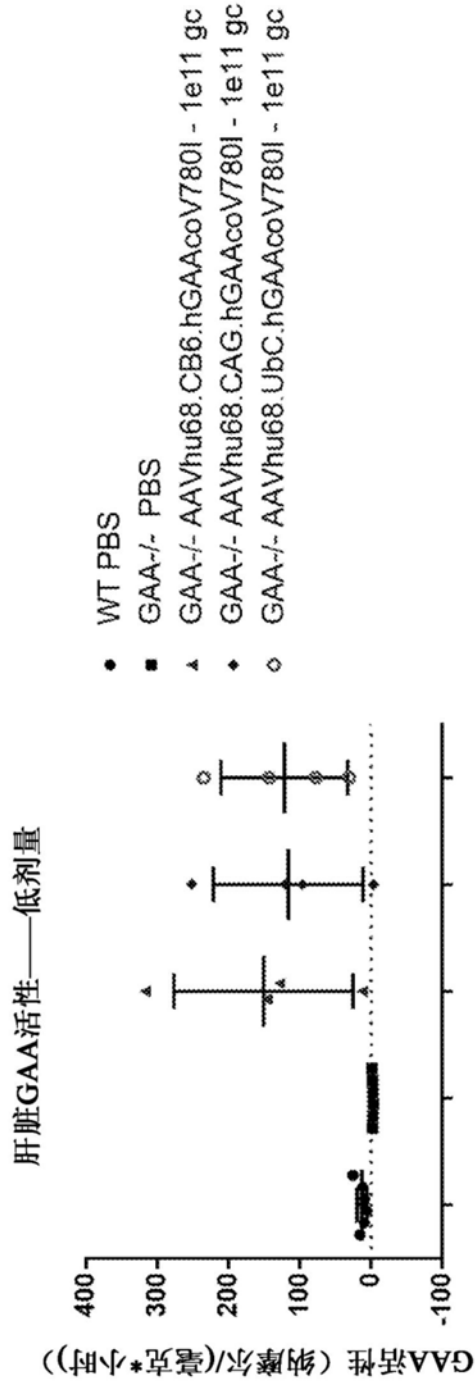


图1A

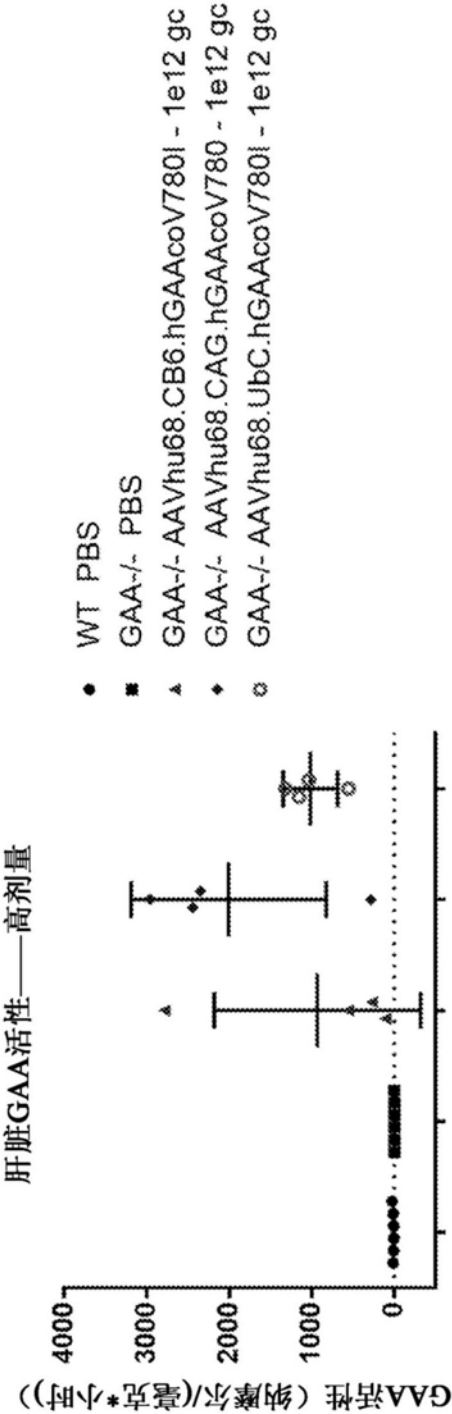


图1B

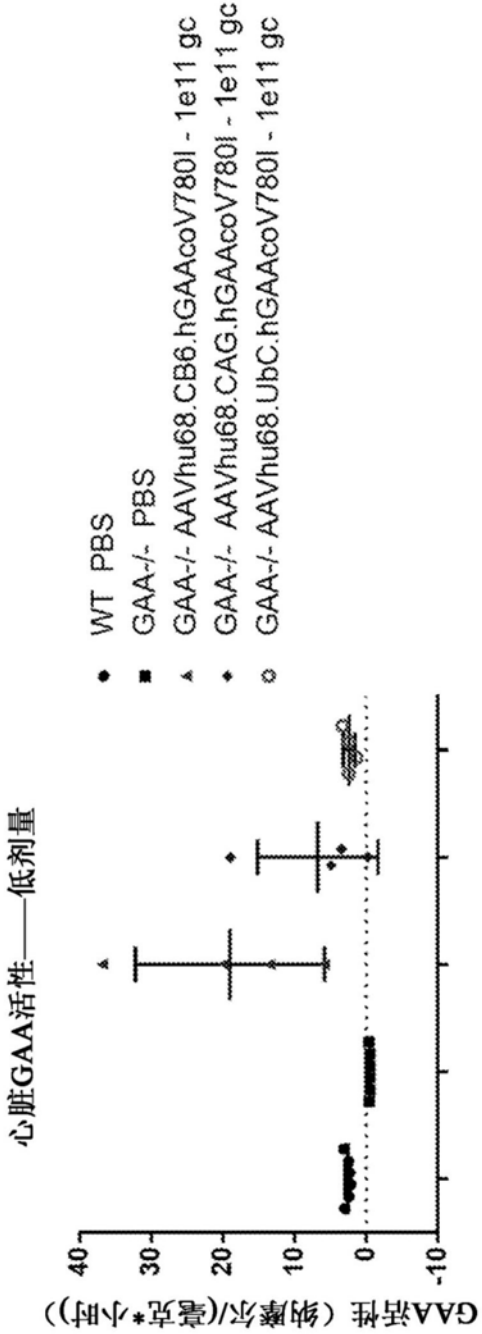


图2A

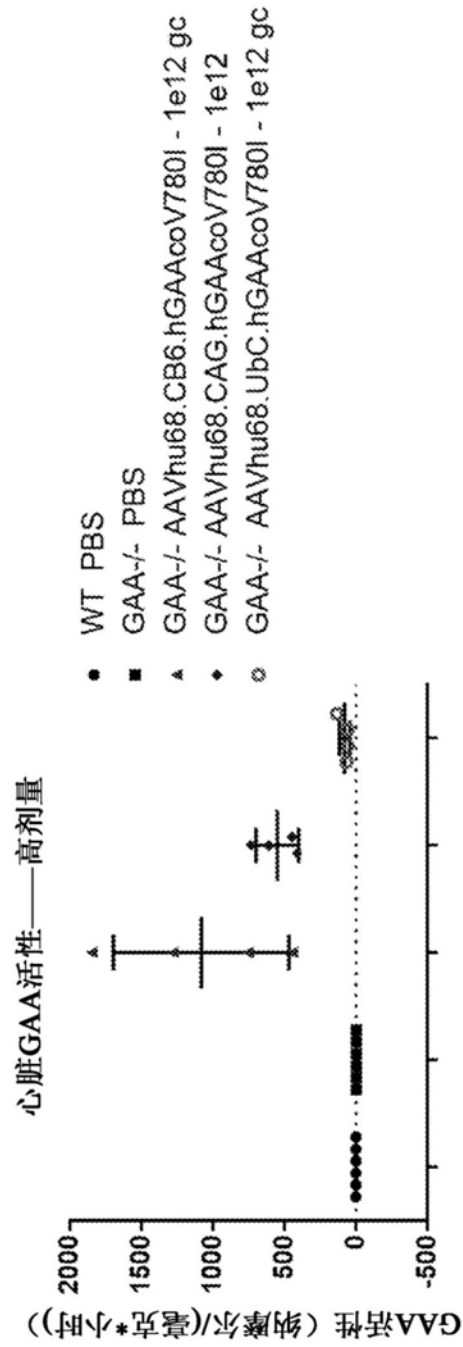


图2B

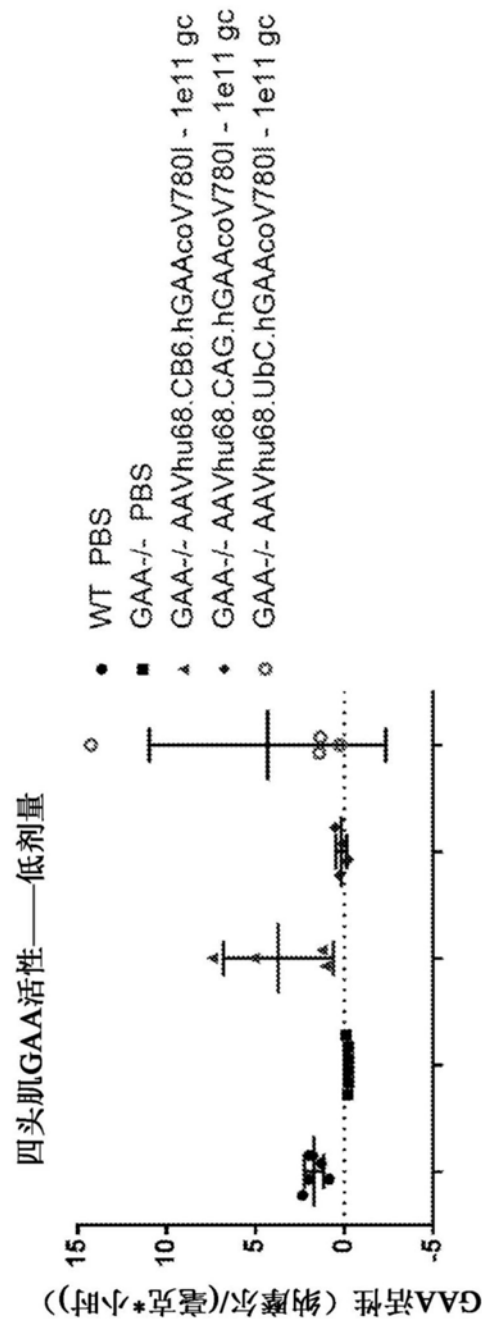


图3A

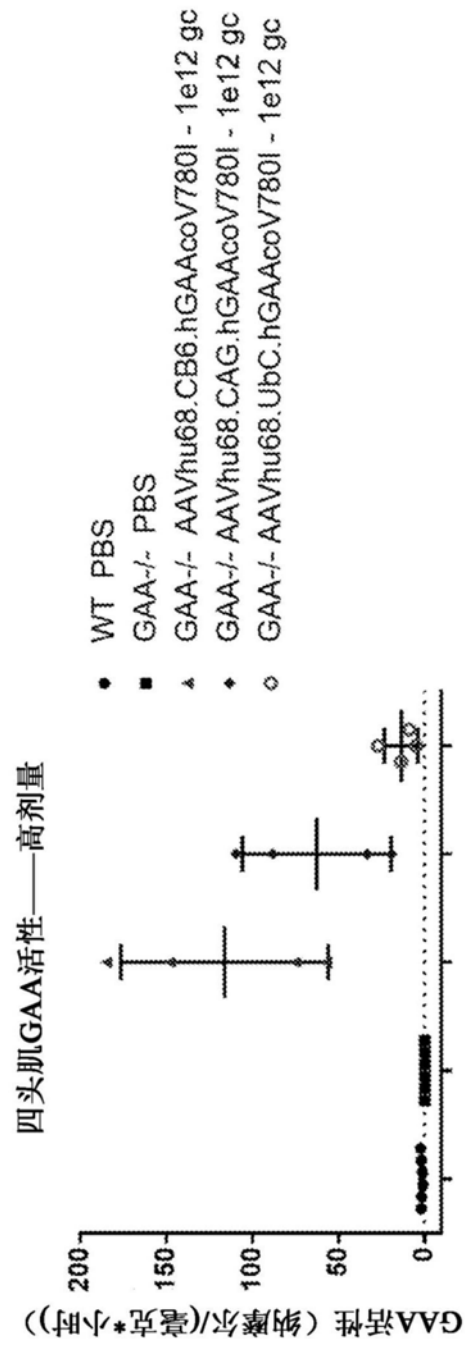


图3B

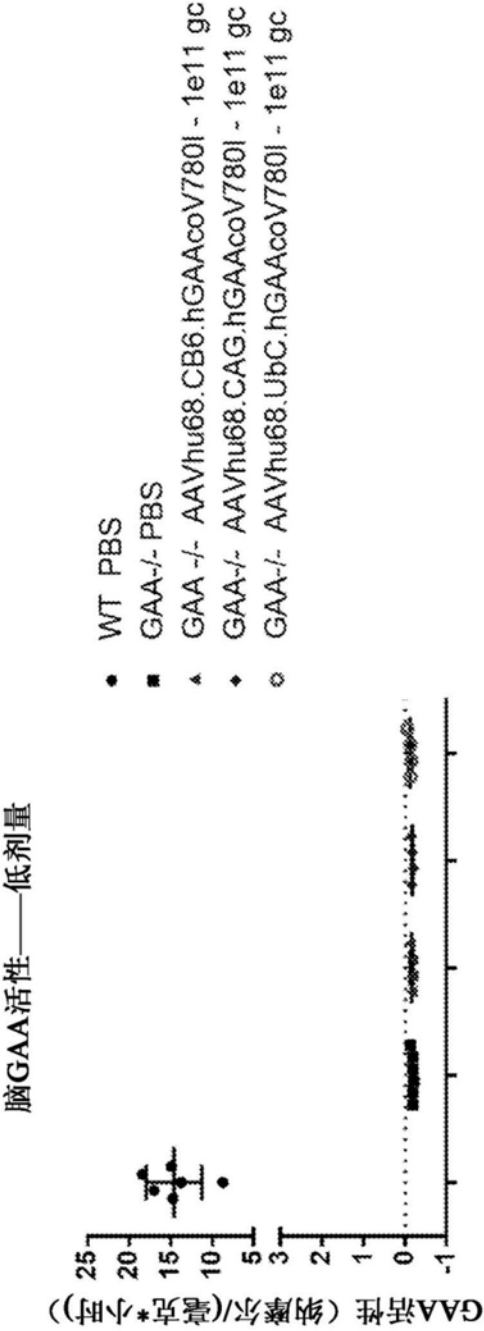


图4A

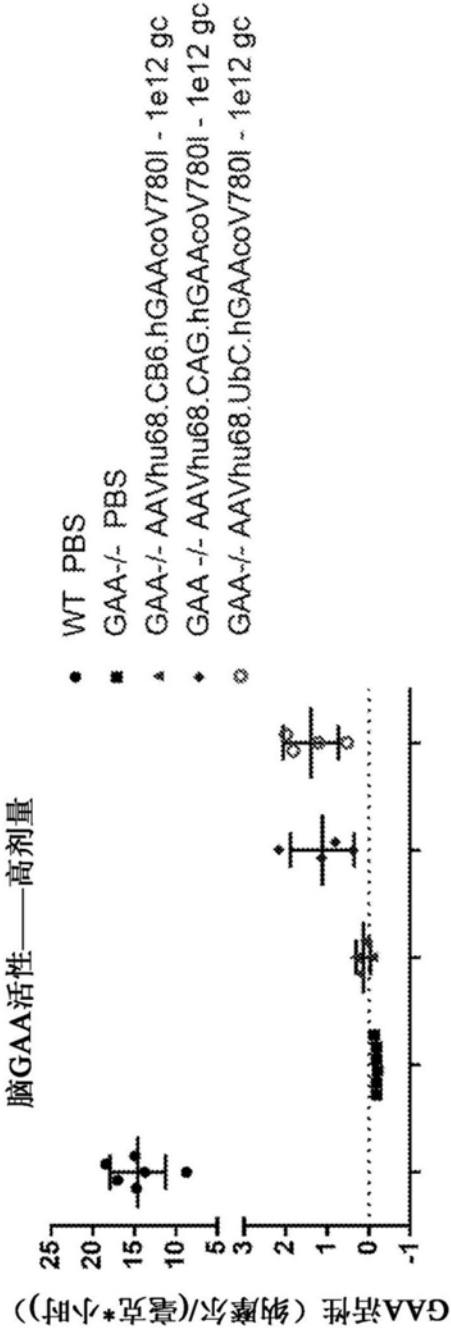


图4B

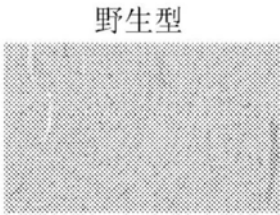


图5A

wt hGAA



图5B

BiP-viGF2-hGAAco

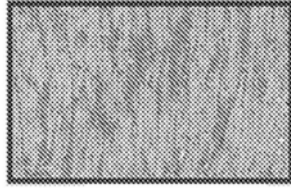


图5C

媒剂

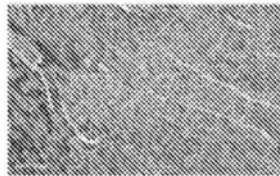


图5D

hGAAco V^{780I}

图5E

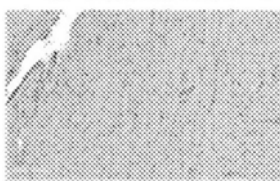
BiP-viGF2-hGAAco V^{780I}

图5F

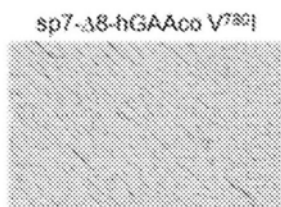


图5G

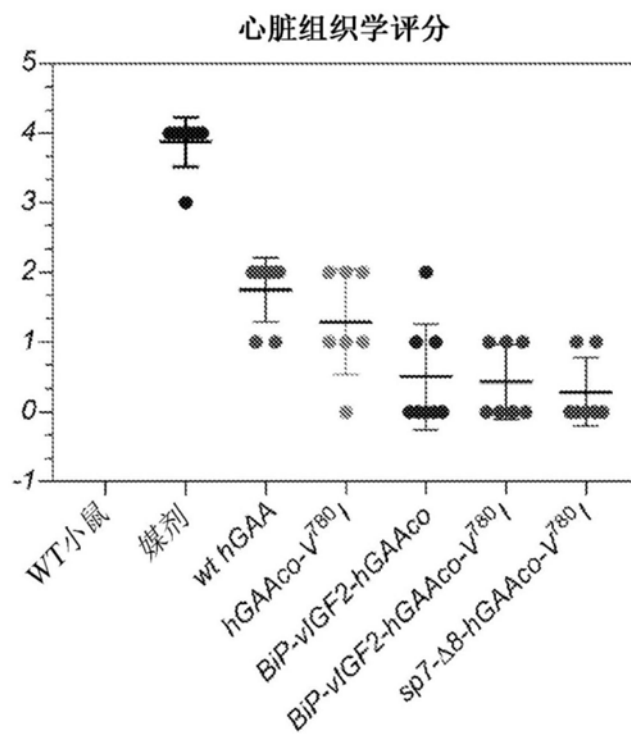


图5H



图6A

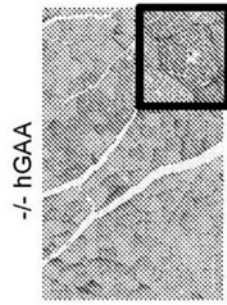


图6B



图6C

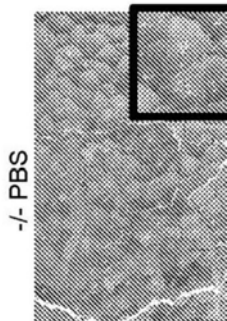
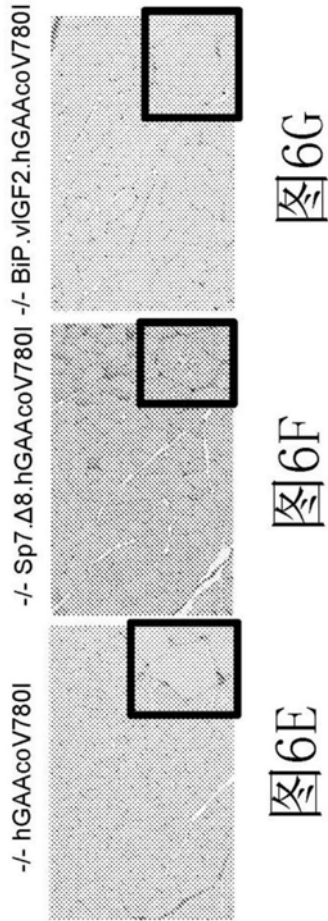


图6D



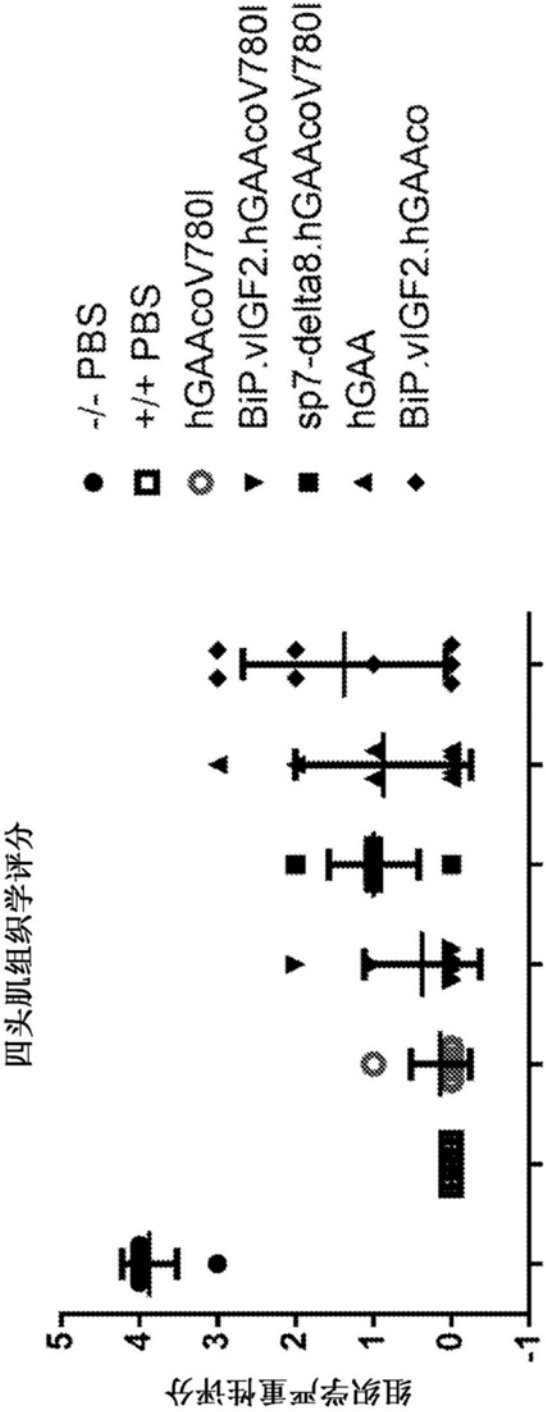


图6H

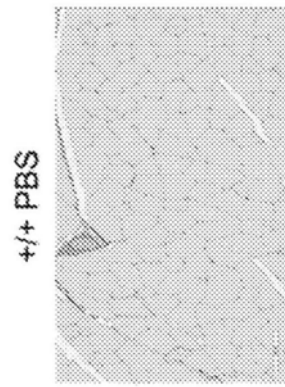


图7A

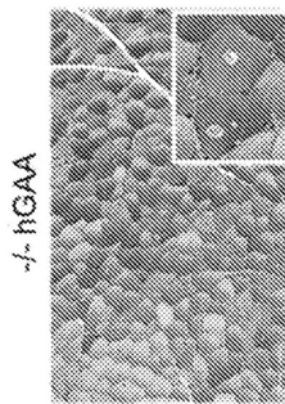


图7B

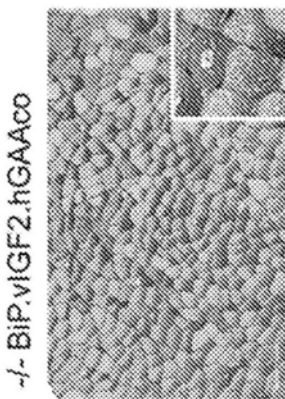


图7C

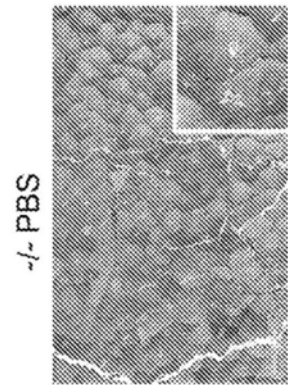


图7D

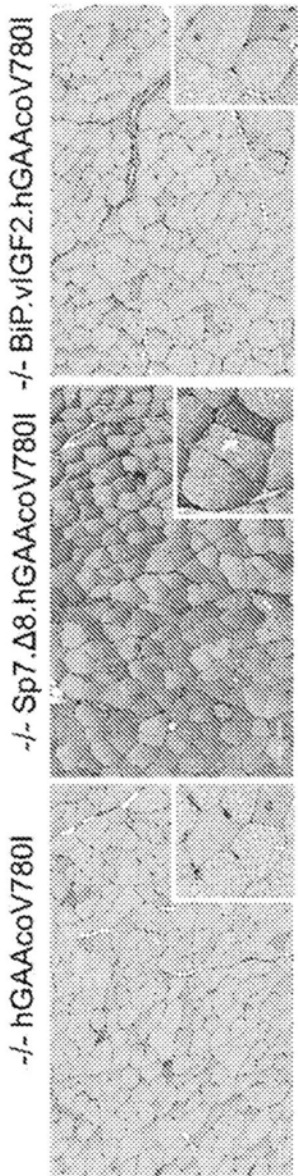


图7E

图7F

图7G

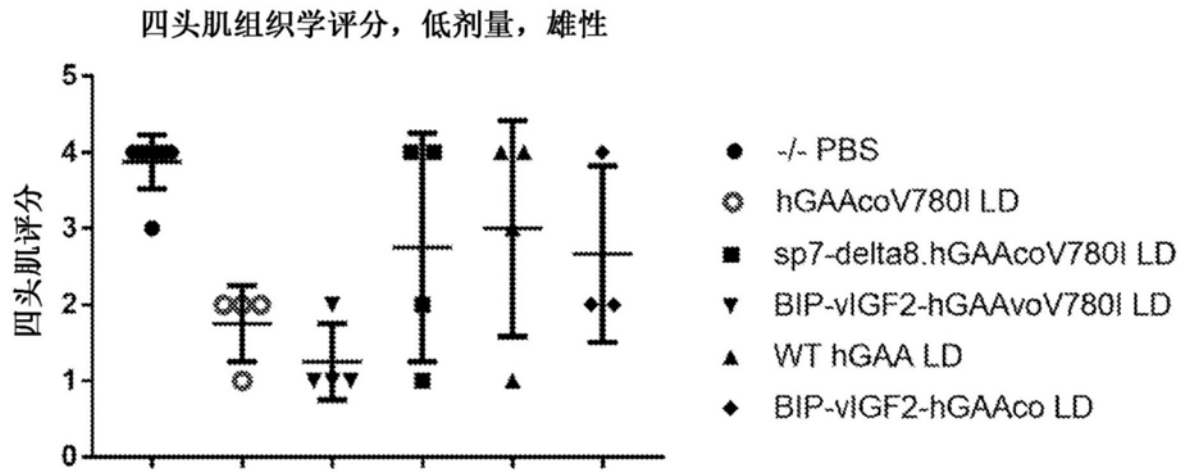


图7H

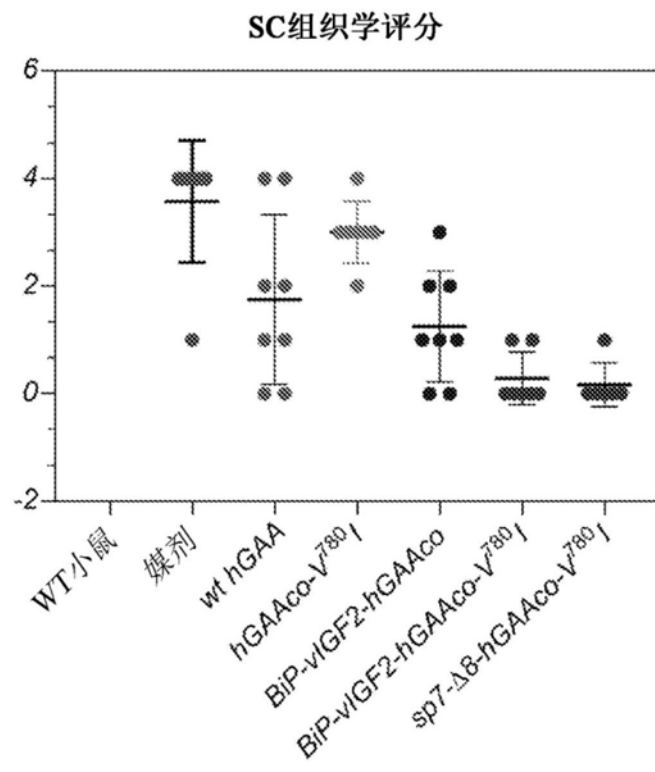


图8

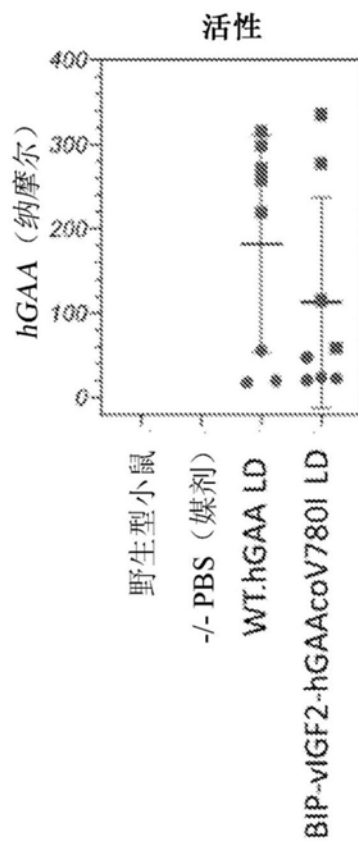


图9A

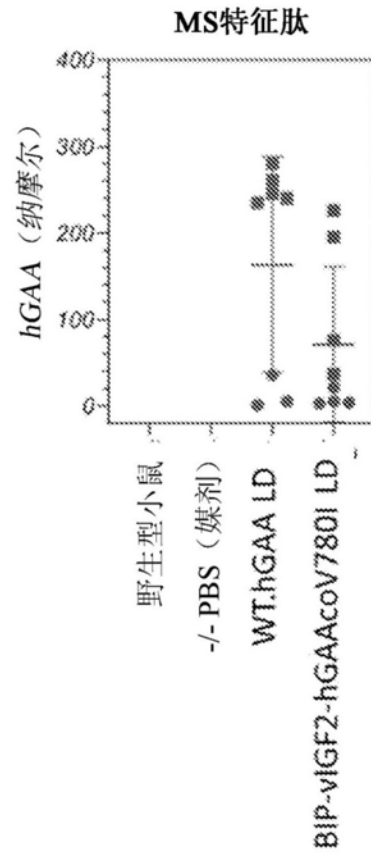


图9B

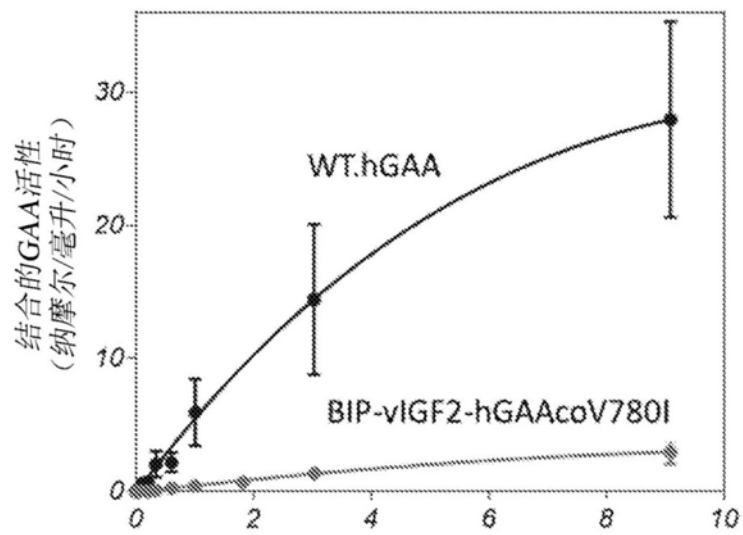


图9C

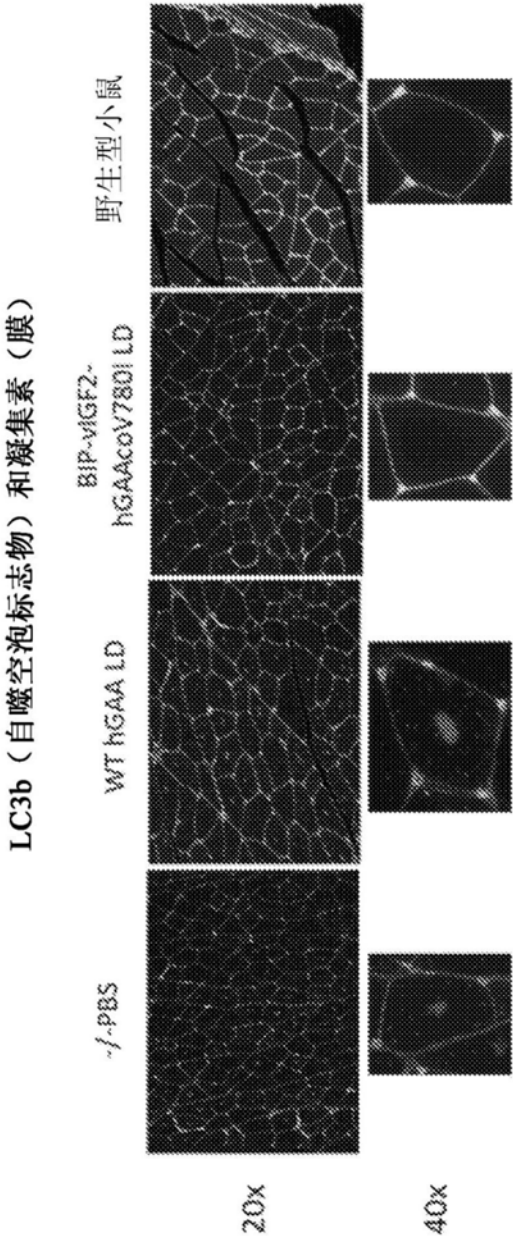


图10

BiP-vIGF2.hGAACO.V780I.4xmiR183

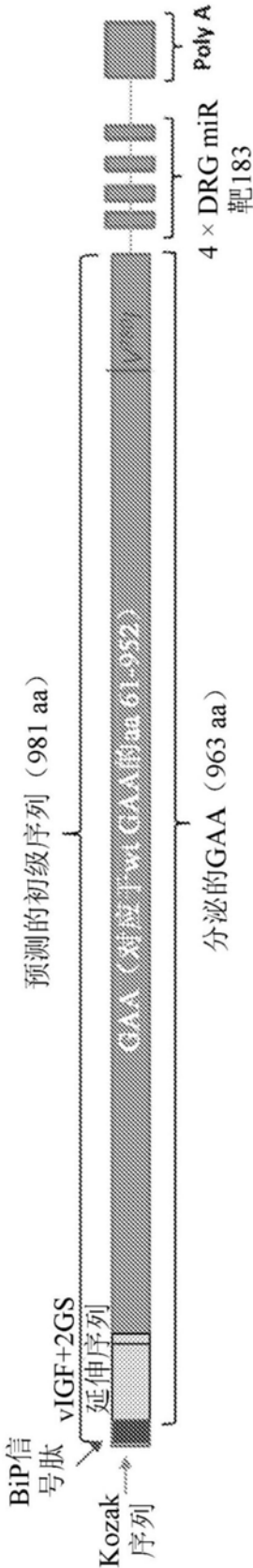


图11

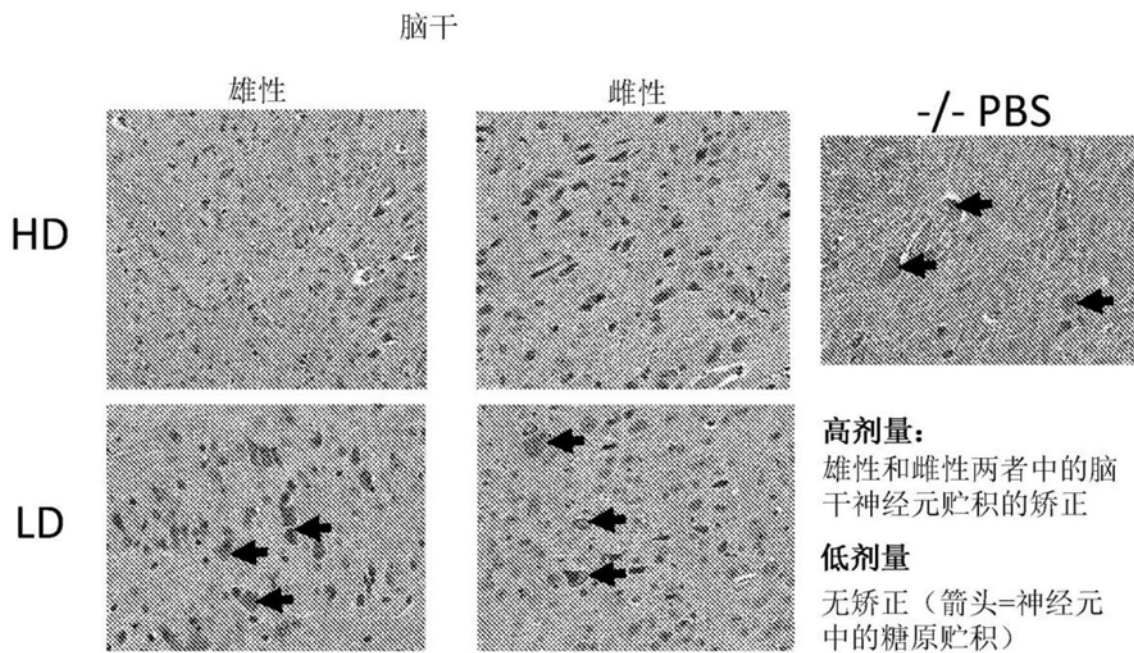


图12

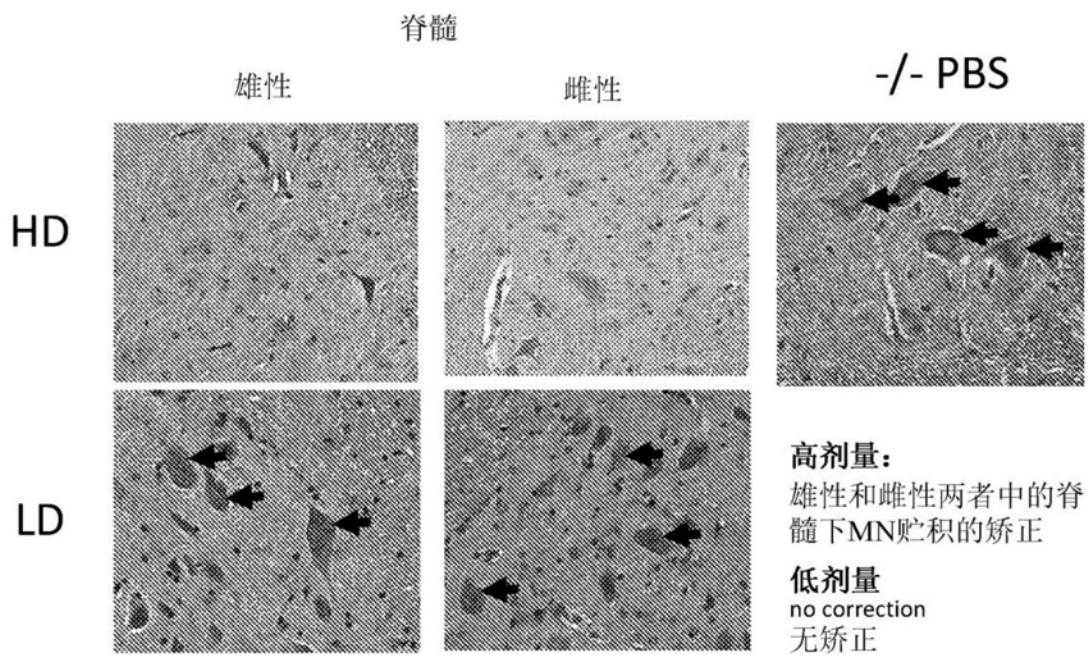


图13

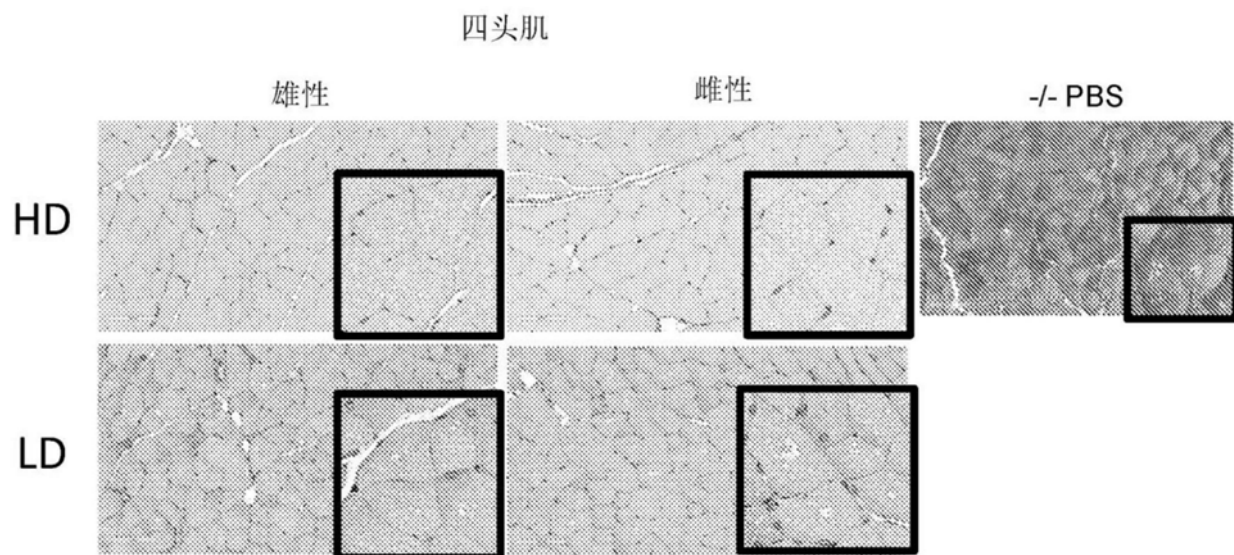


图14

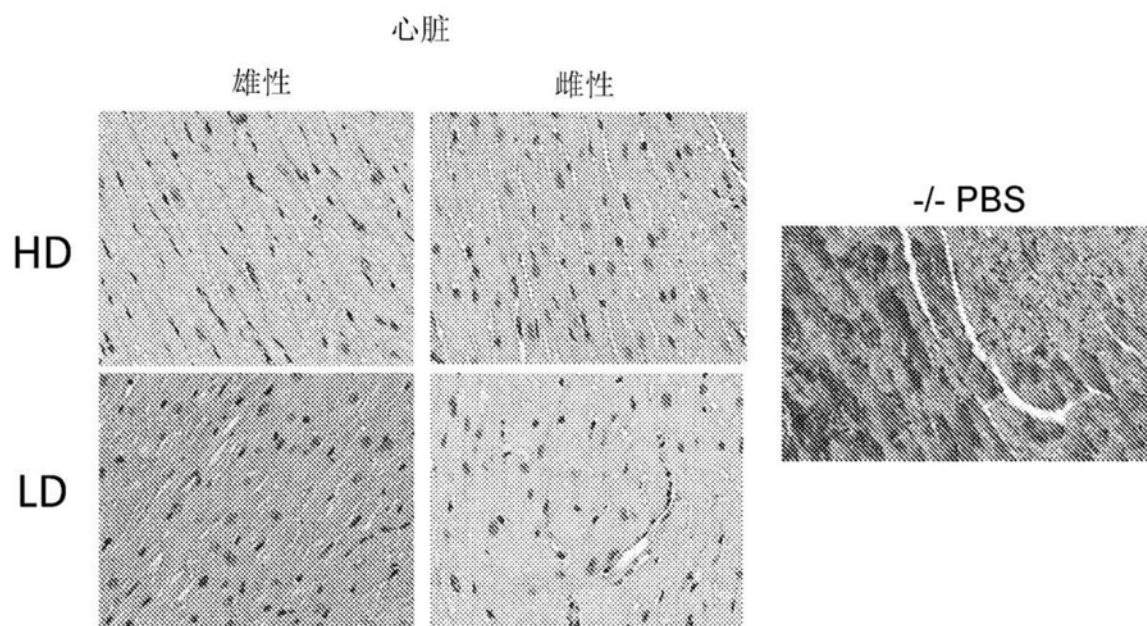


图15

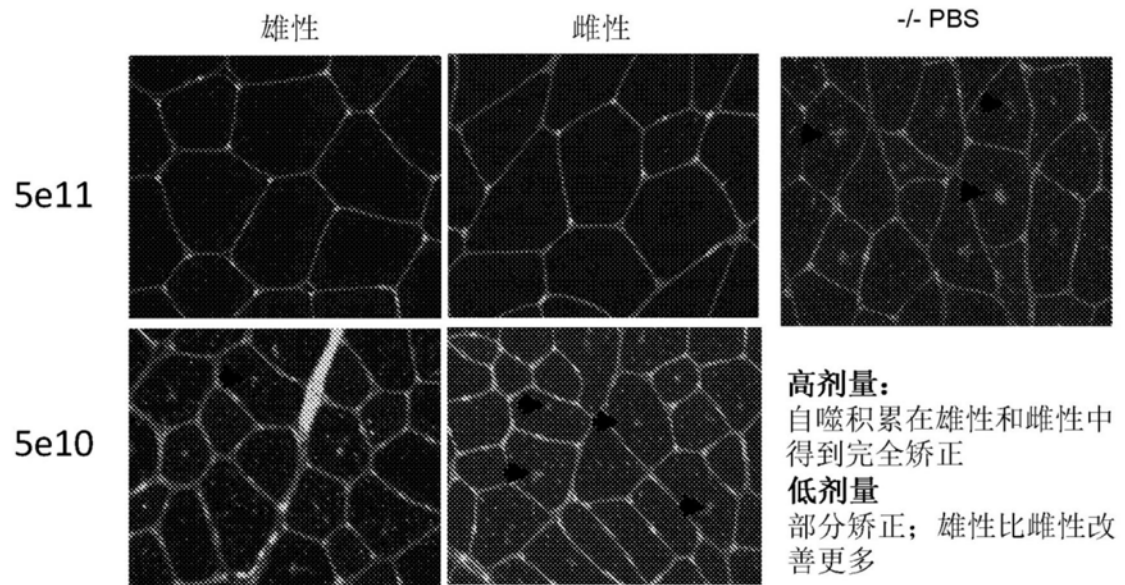


图16

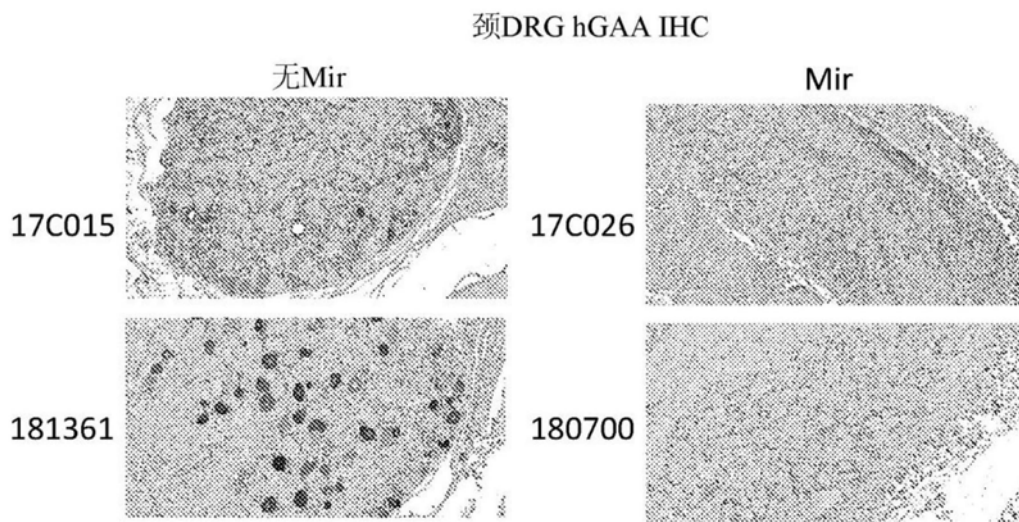


图17

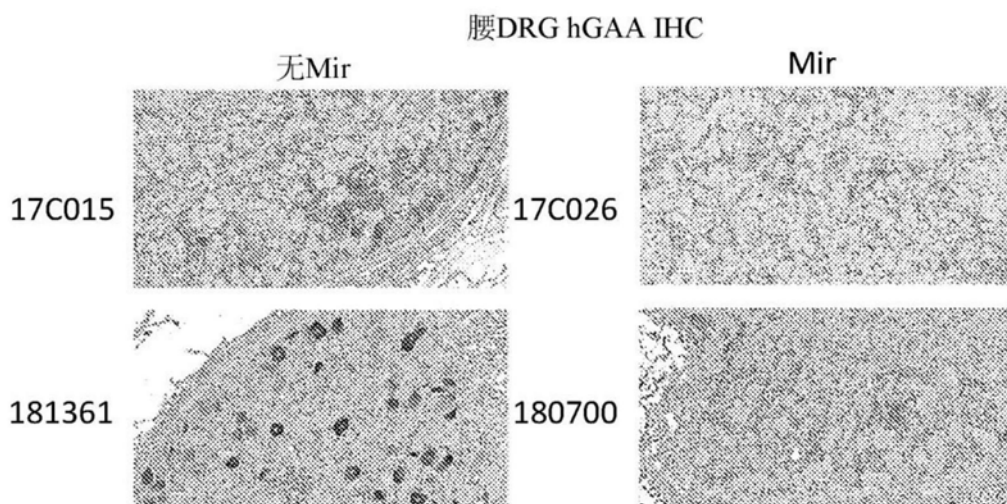


图18

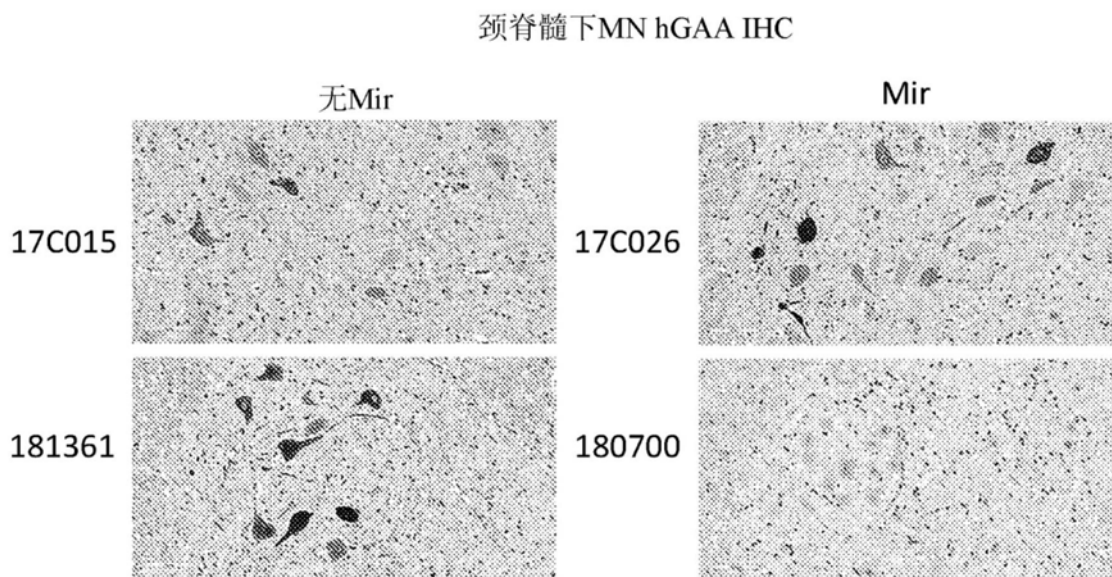


图19

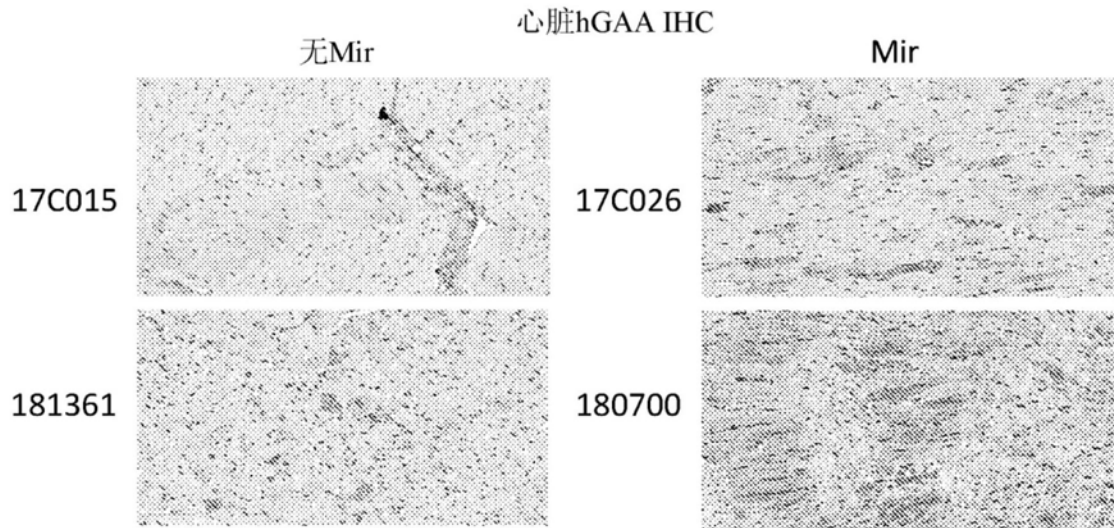


图20

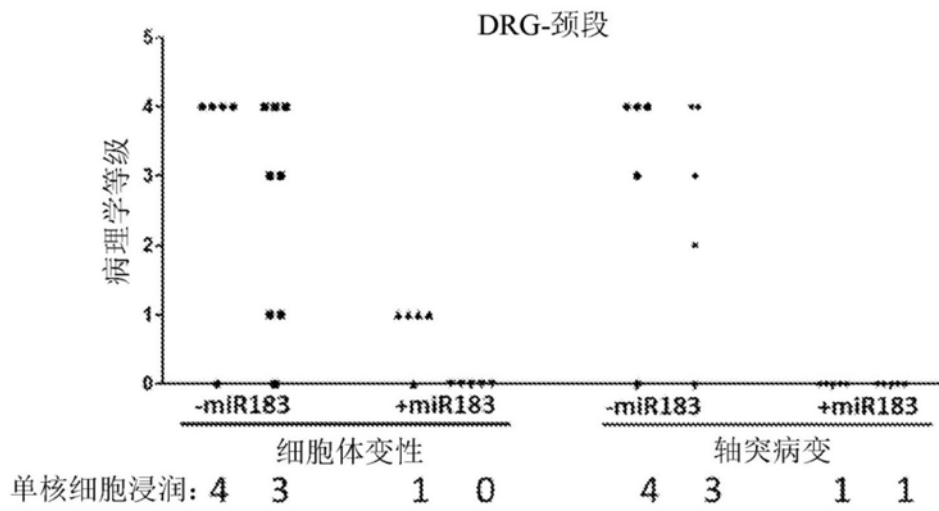


图21A

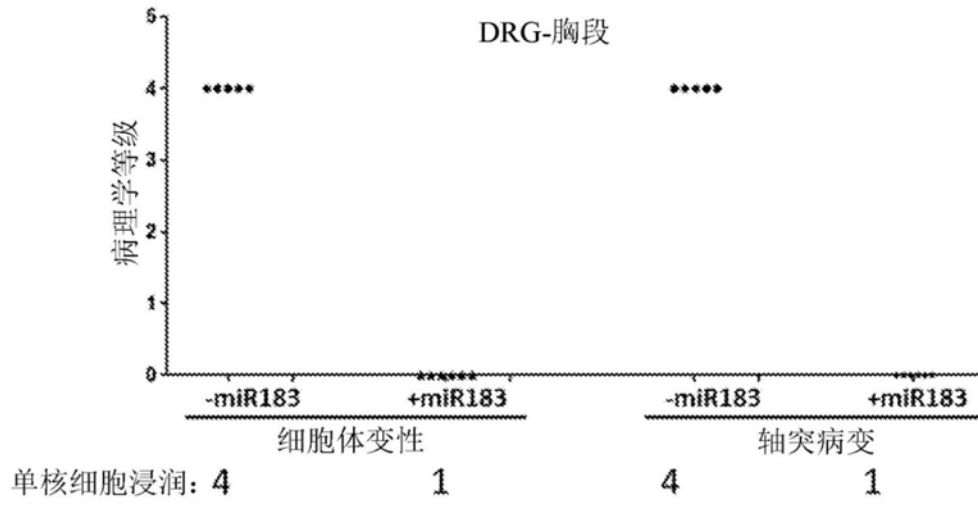


图21B

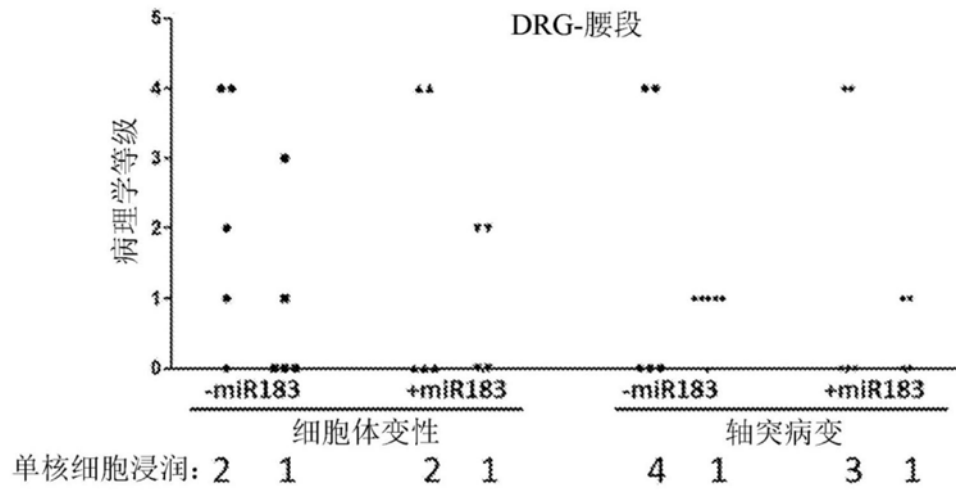


图21C

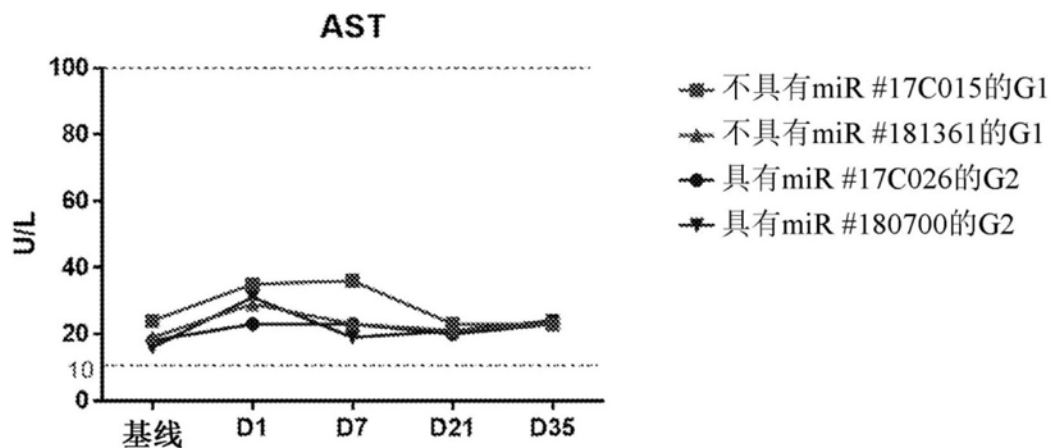


图22A

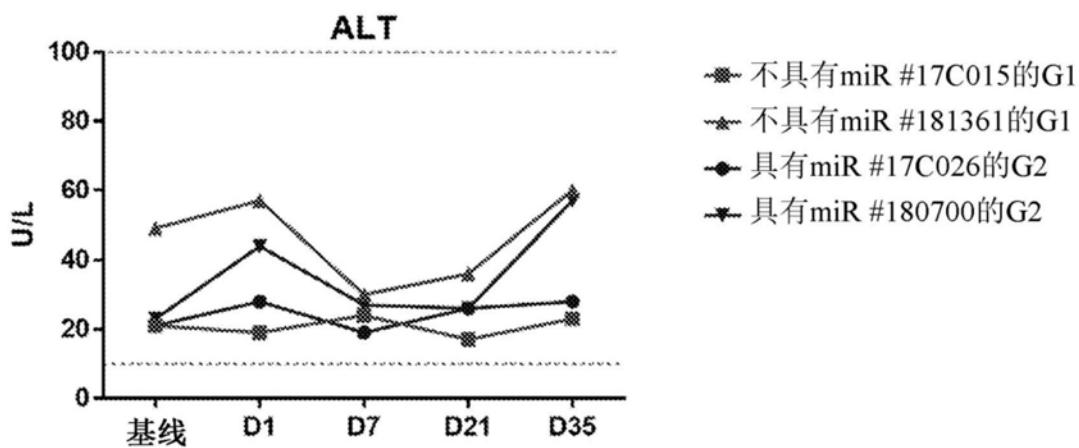


图22B

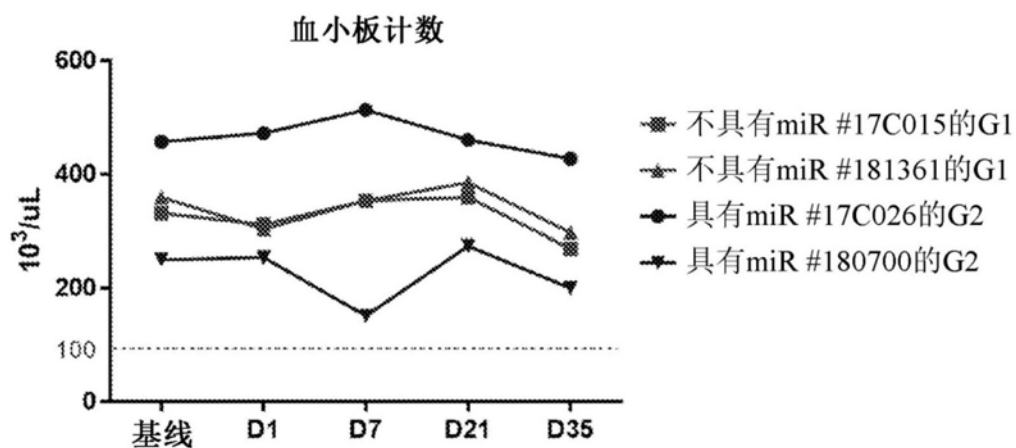


图22C

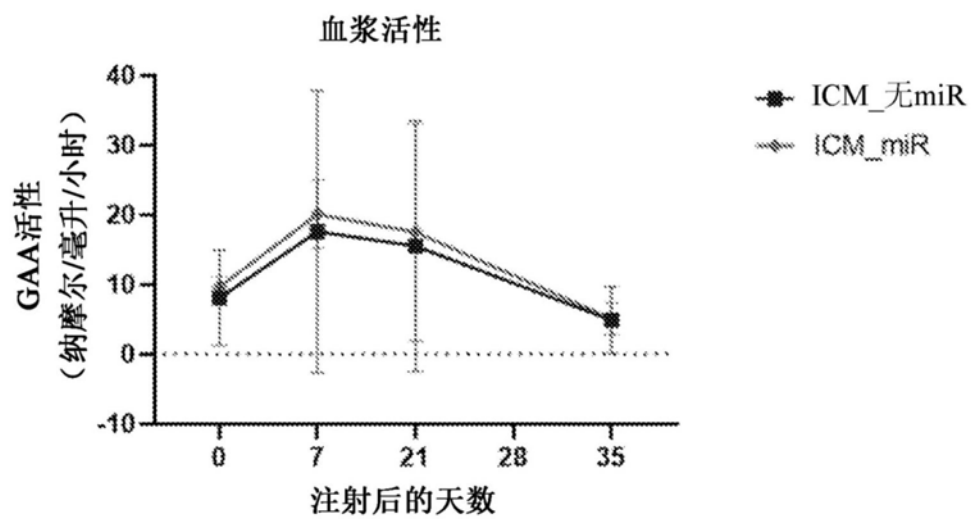
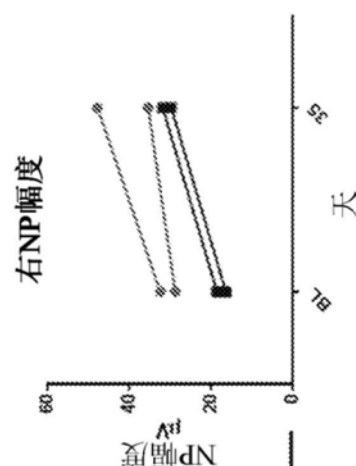
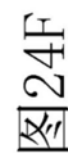
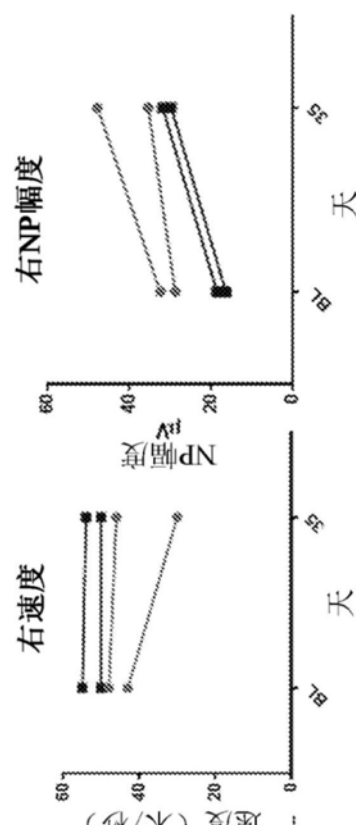
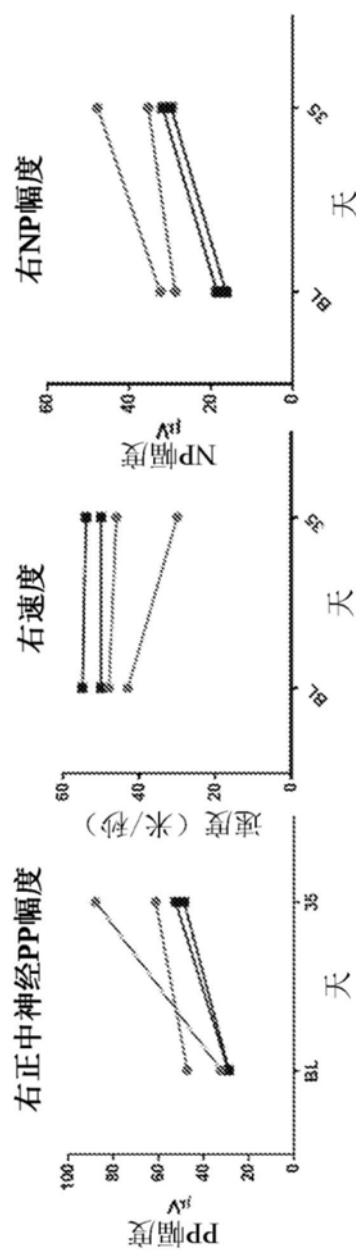
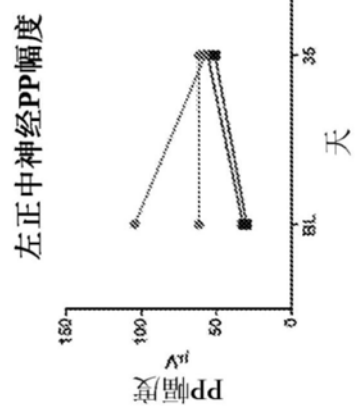
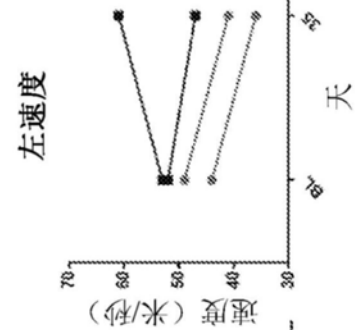
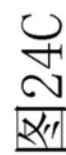
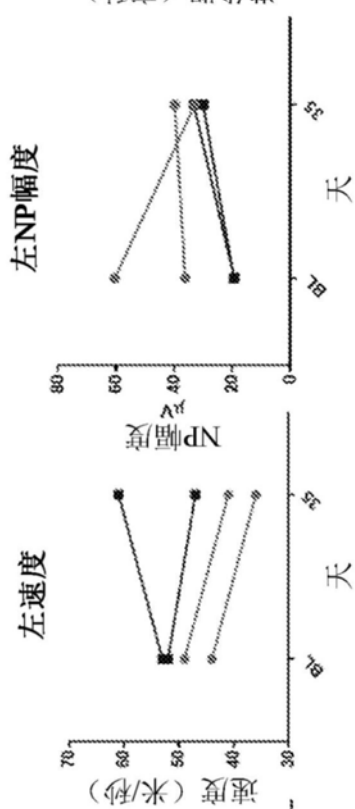
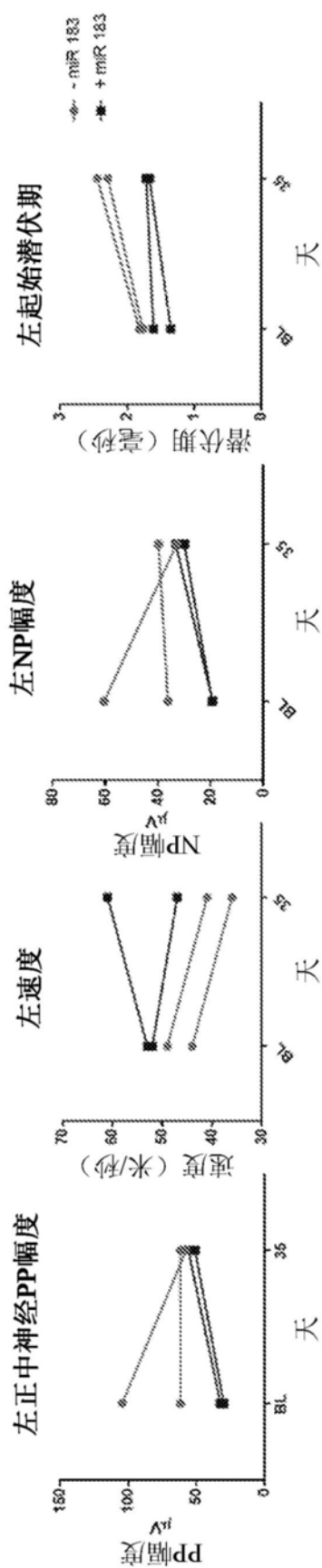


图23



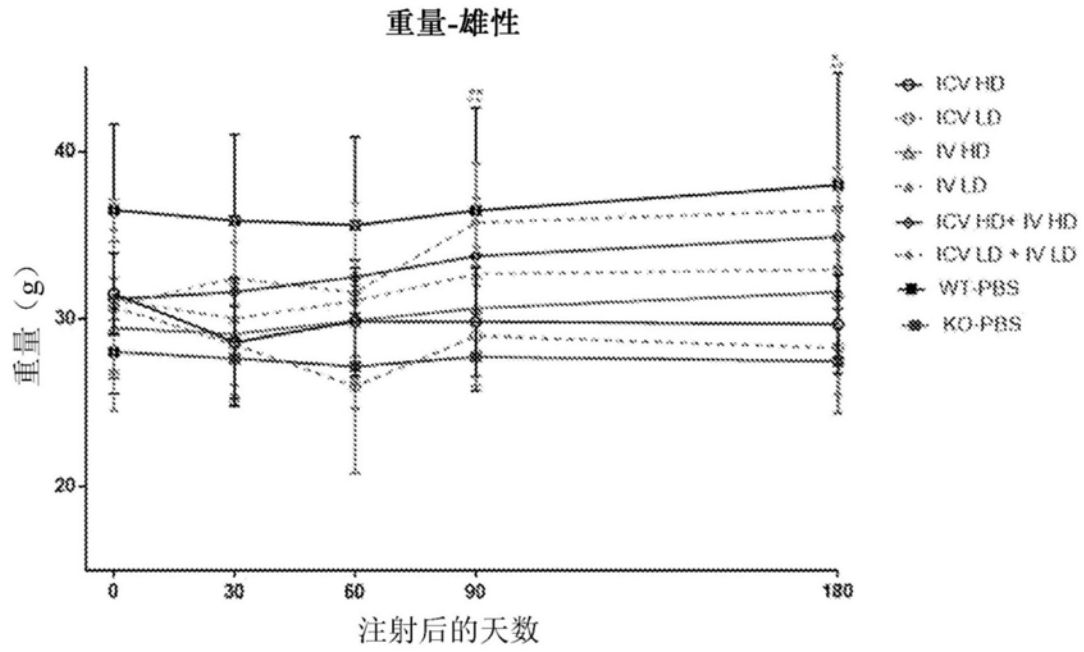


图25A

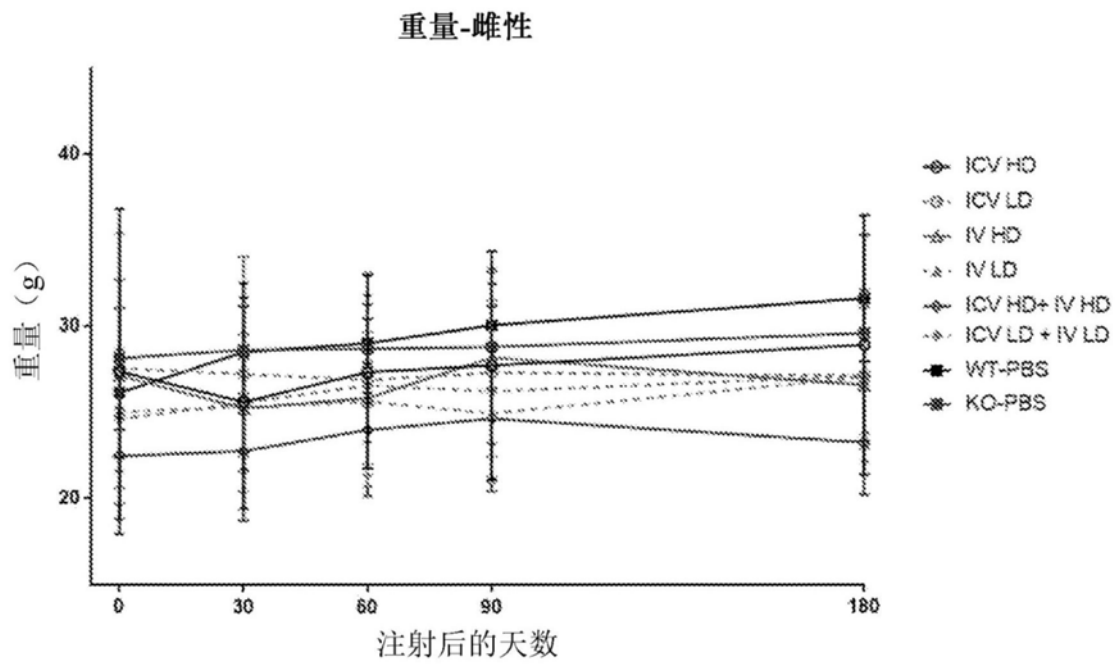


图25B

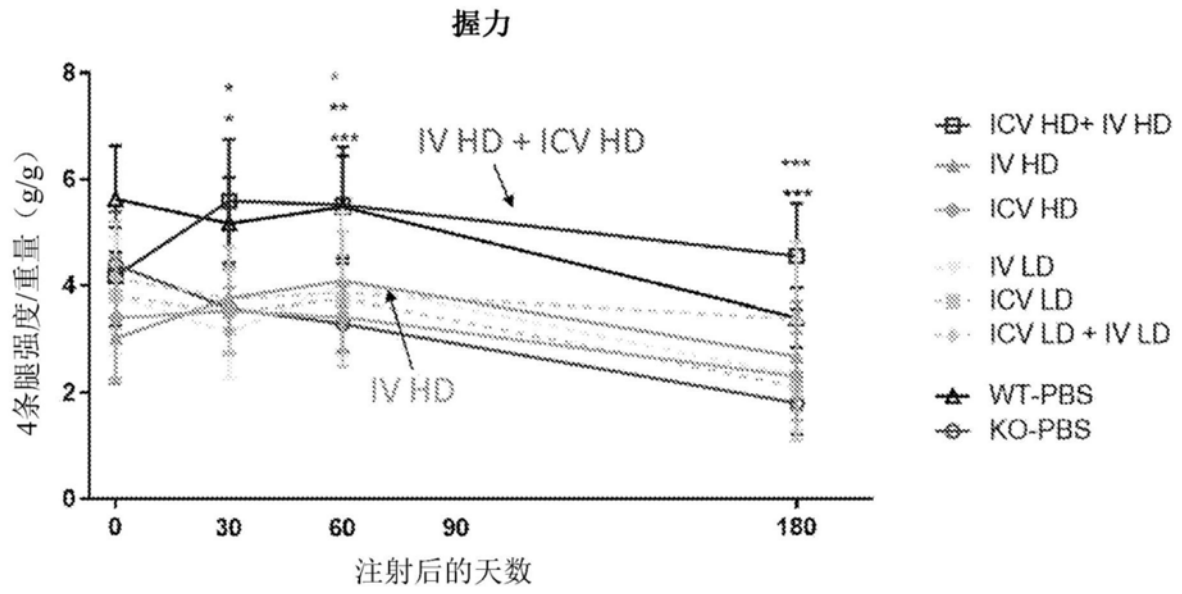


图26A

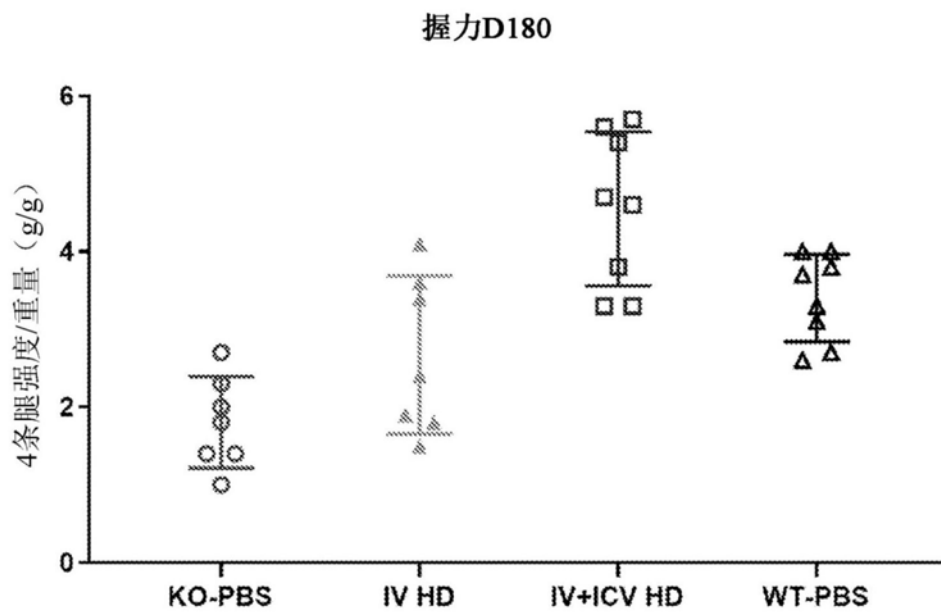


图26B

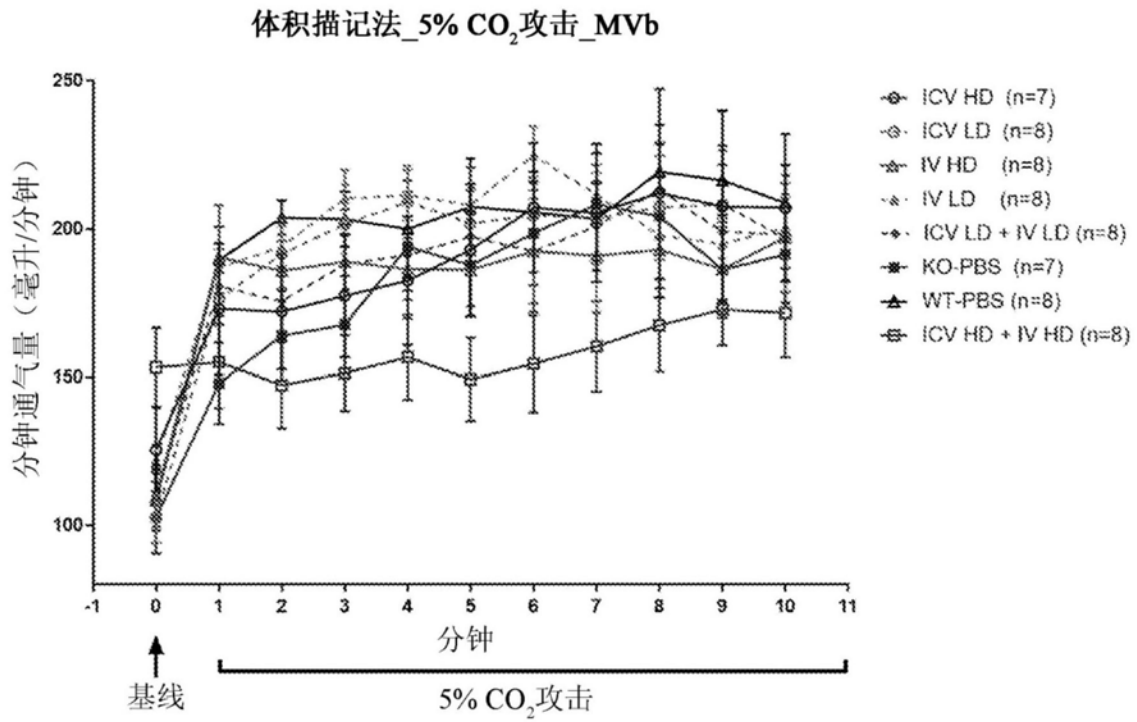


图27A

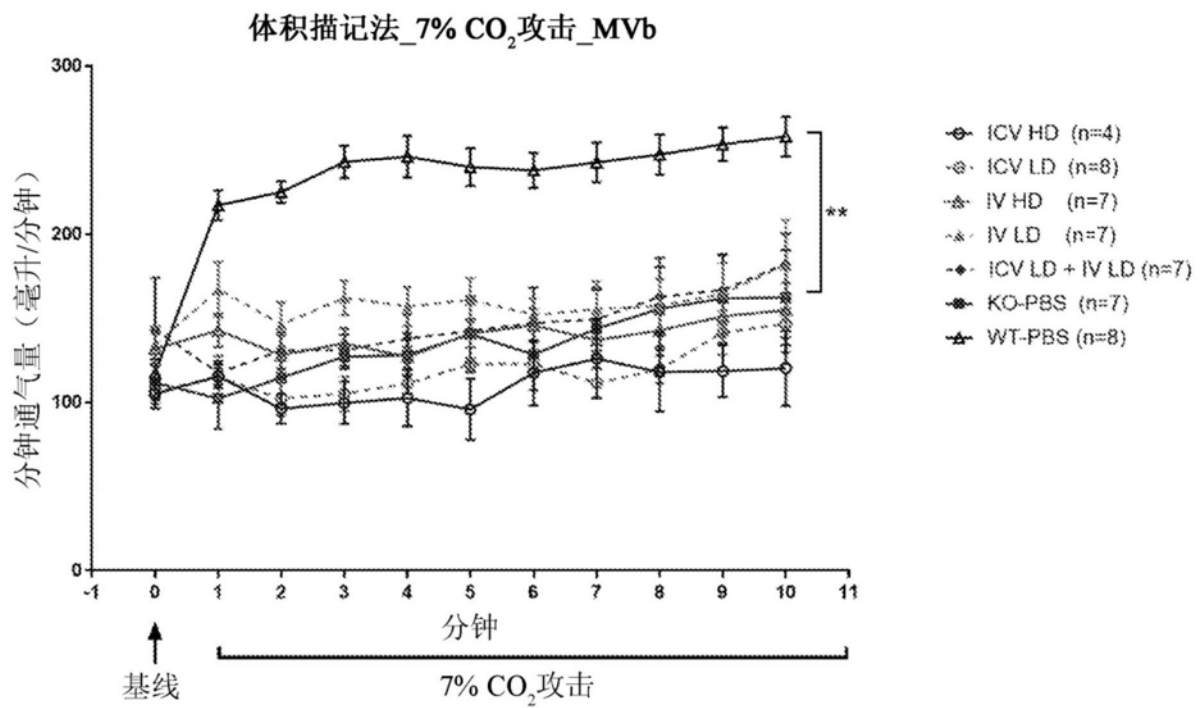


图27B

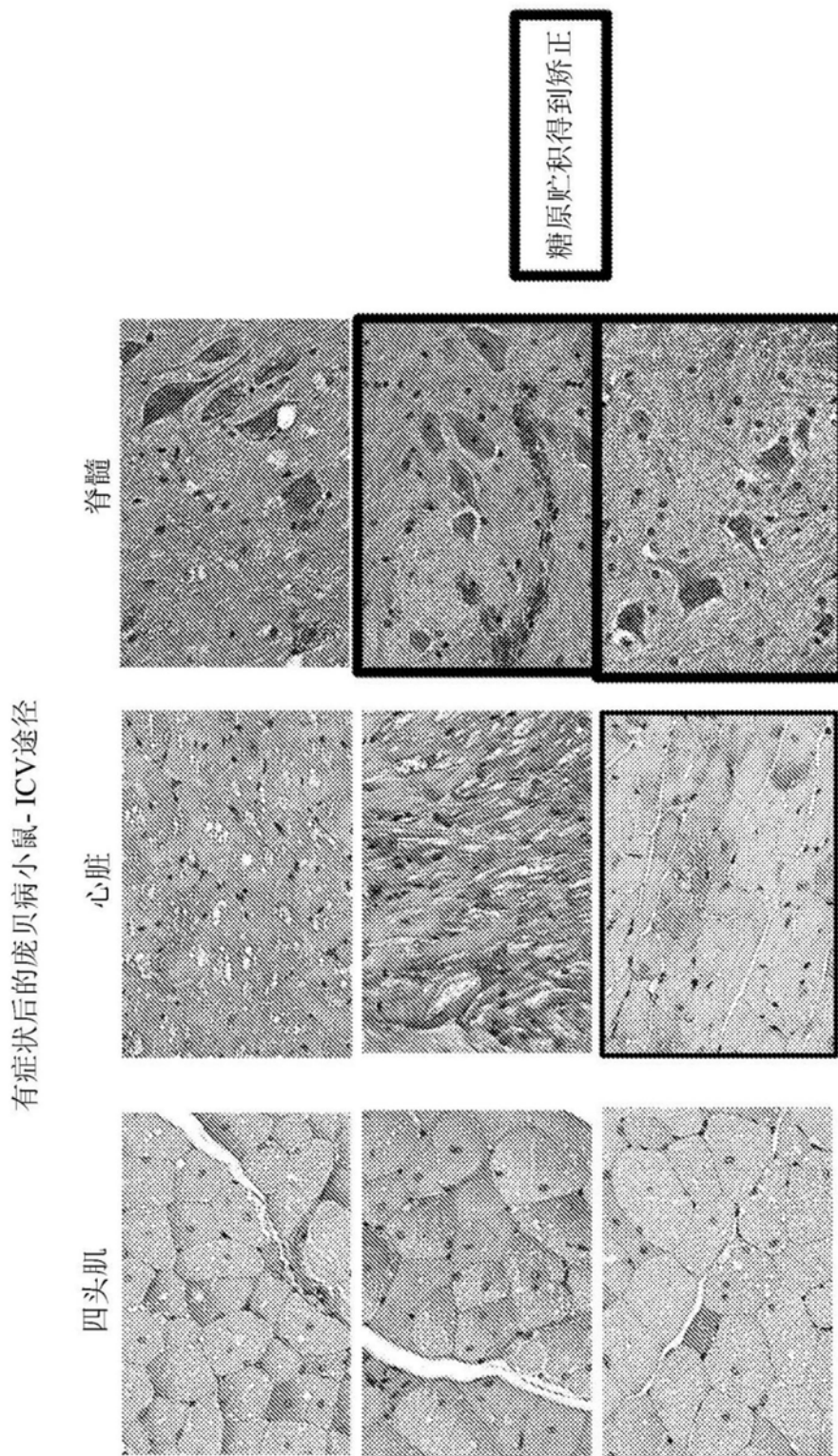


图28

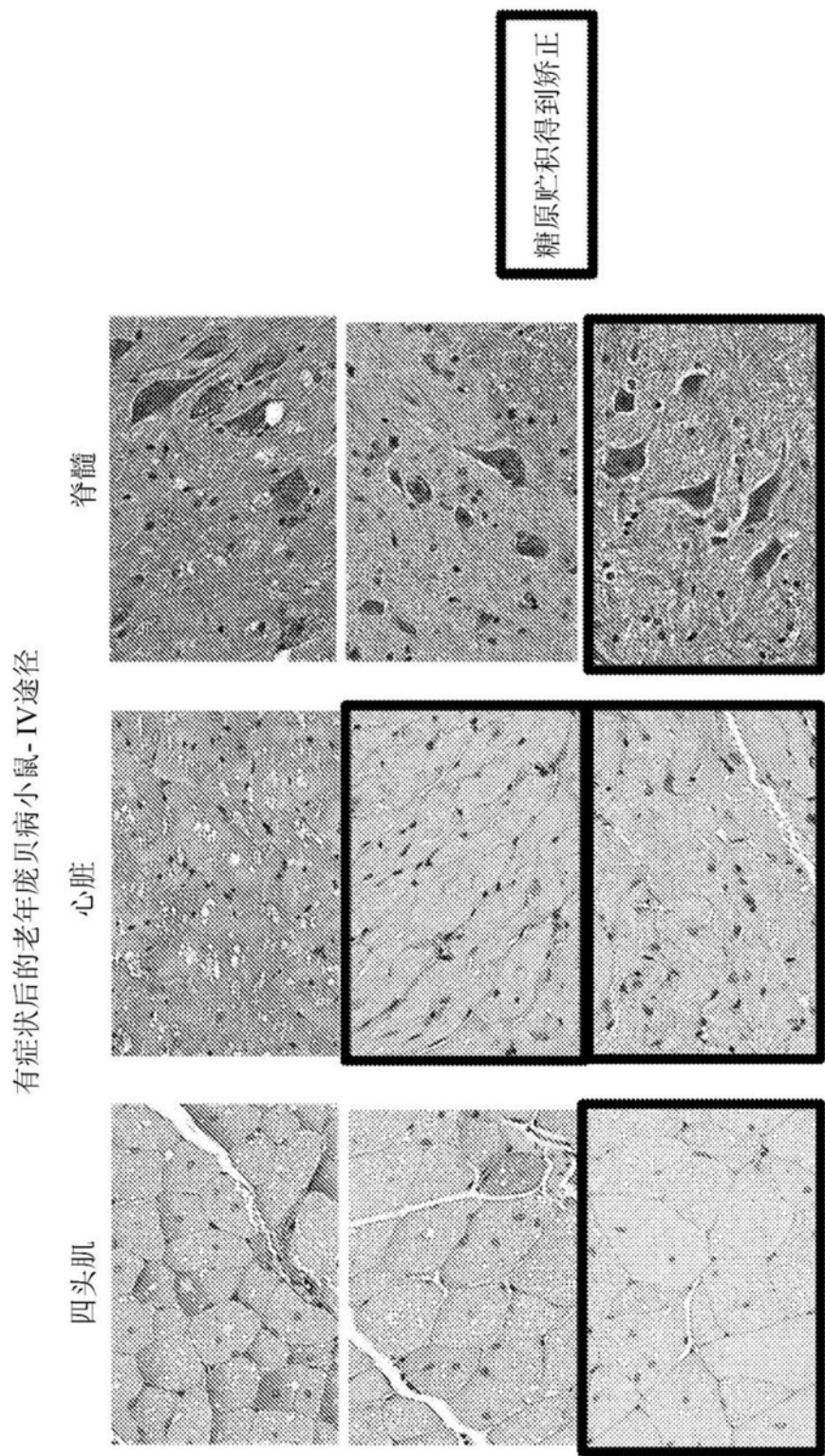


图29

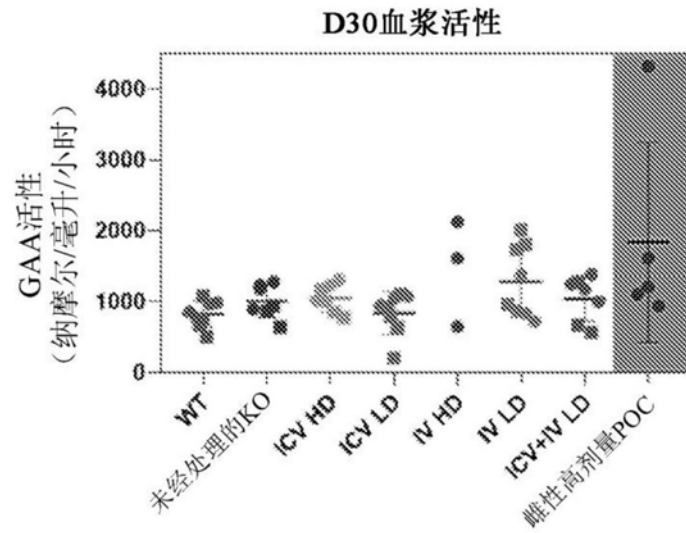


图30A

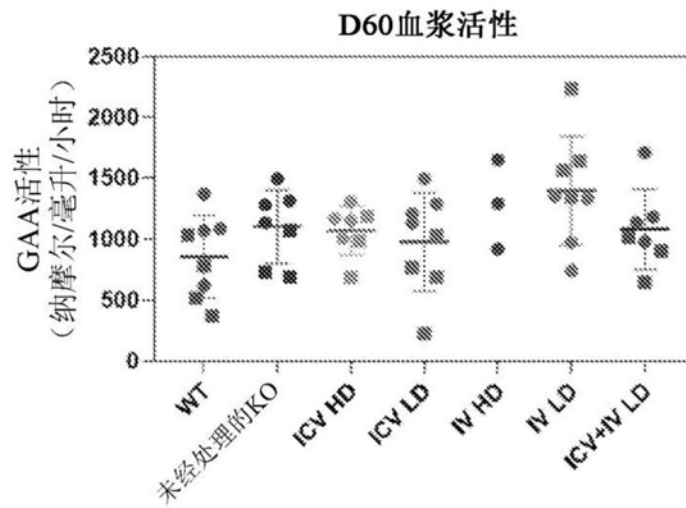


图30B

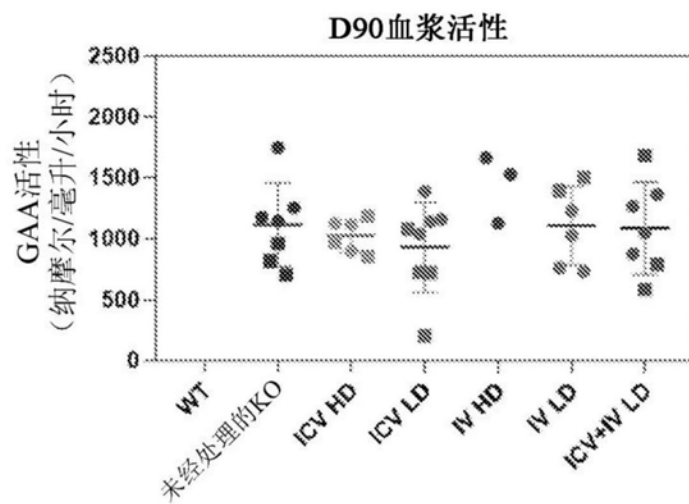


图30C

	-21	-10 (近似)	0	3	7	21	29	35	60	
	基线	强的松	载体给药	临床病理学	临床病理学	临床病理学	强的松逐渐减少	临床病理学	尸检	
	临床病理学		临床病理学	血浆	临床病理学	临床病理学		血浆/CSF	临床病理学	
	NAB		血浆/CSF		血浆/CSF	血浆/CSF		血浆/CSF	血浆/CSF	
	神经系统检查		NAB / PBMC					神经系统检查	NAB / PBMC	
组号	1	2	3	4	5	6*				
N/组	2	2	2	2	2	2				
施用途径	IV LD	IV HD	ICM LD	ICM HD	IV HD ICM HD	IV HHD ICM LD				
载体剂量	1e13 GC/kg	5e13 GC/kg	1e13 GC	3e13 GC	5E13 GC/kg 3e13 GC	1E14 GC/kg 1e13 GC				

图31

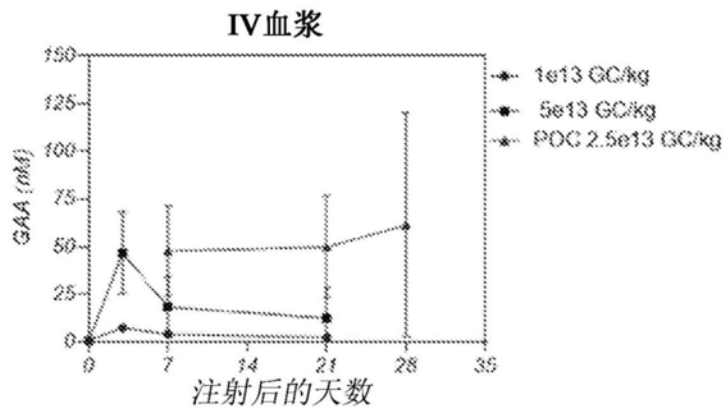


图32A

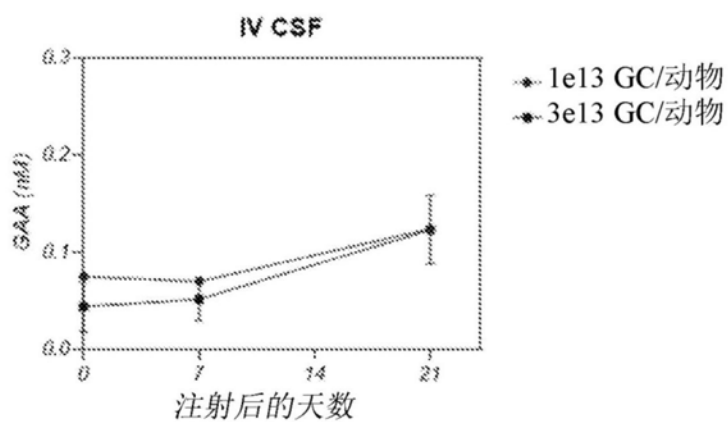


图32B

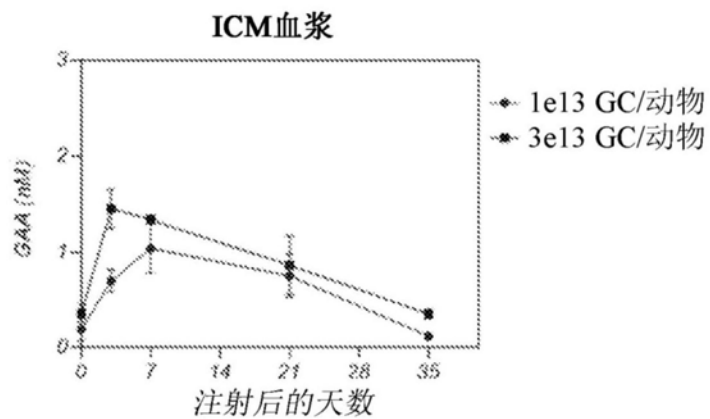


图32C

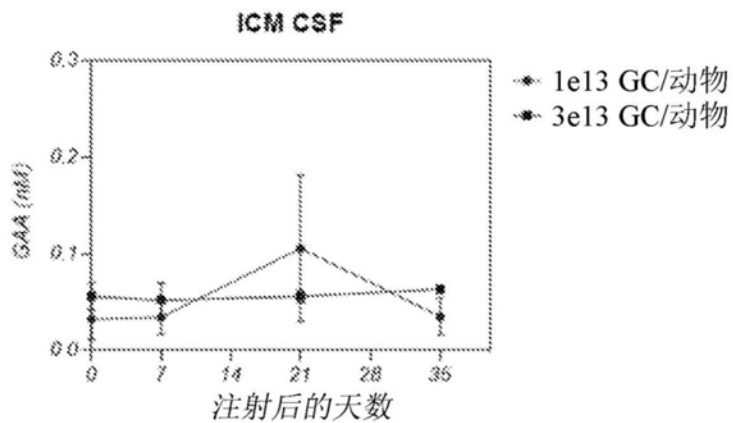


图32D

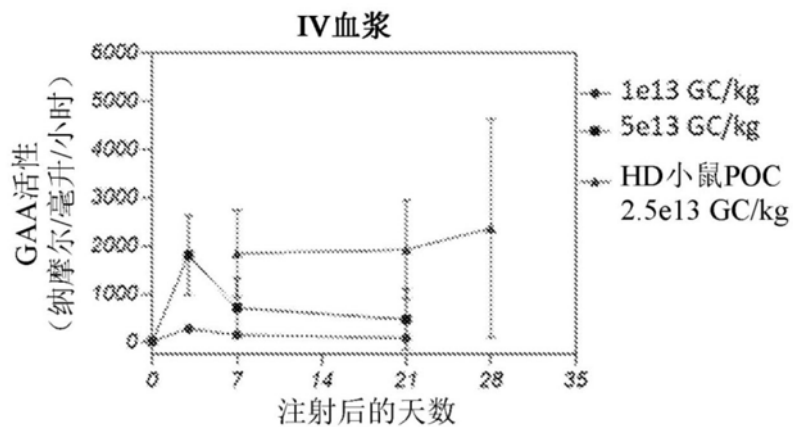


图32E

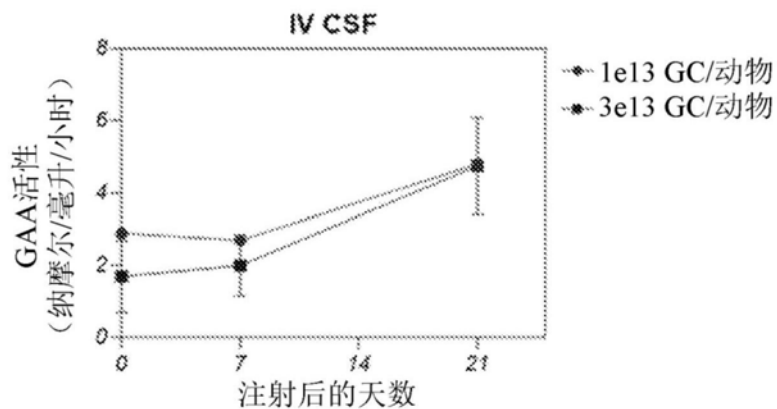


图32F

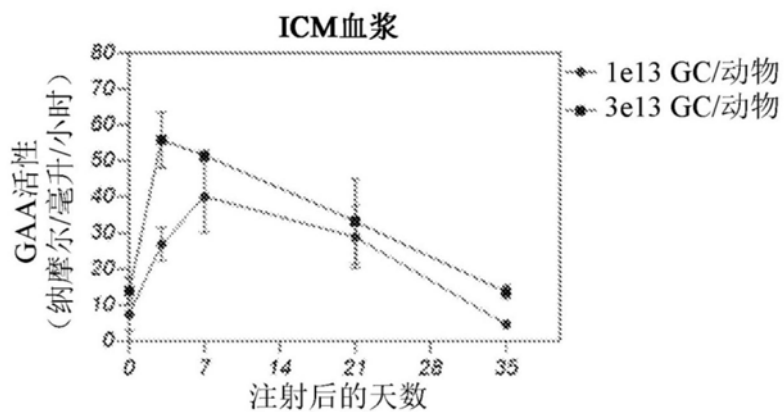


图32G

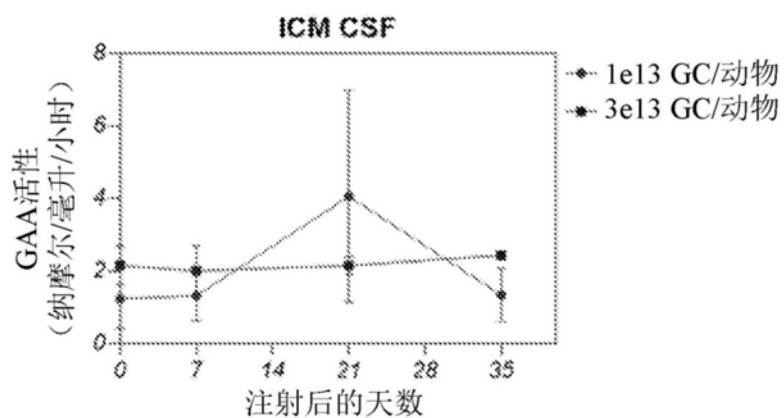


图32H

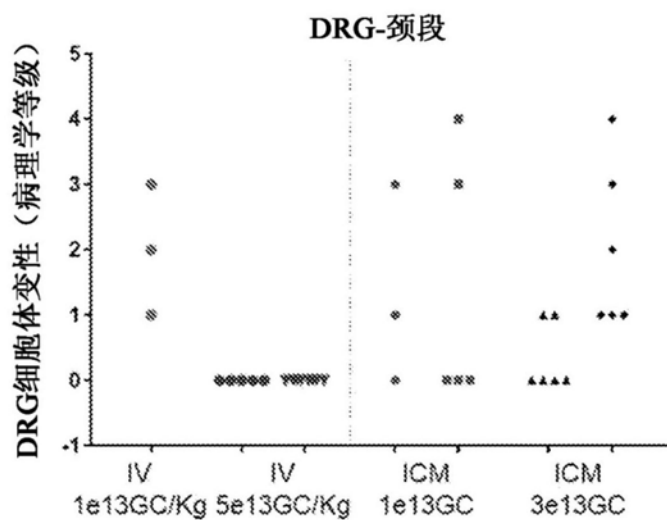


图33A

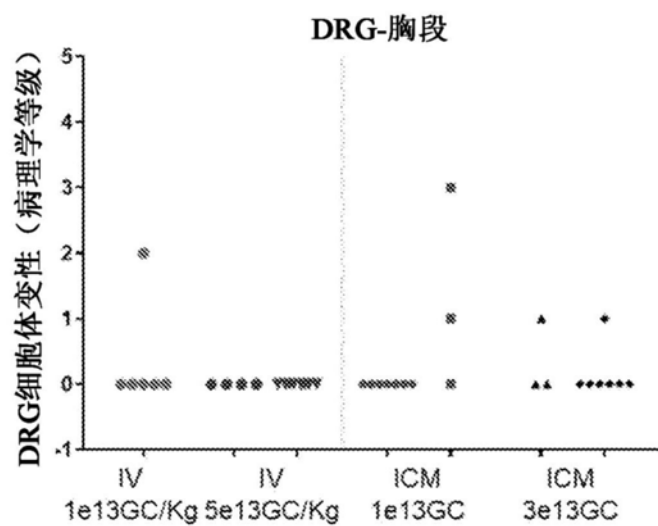


图33B

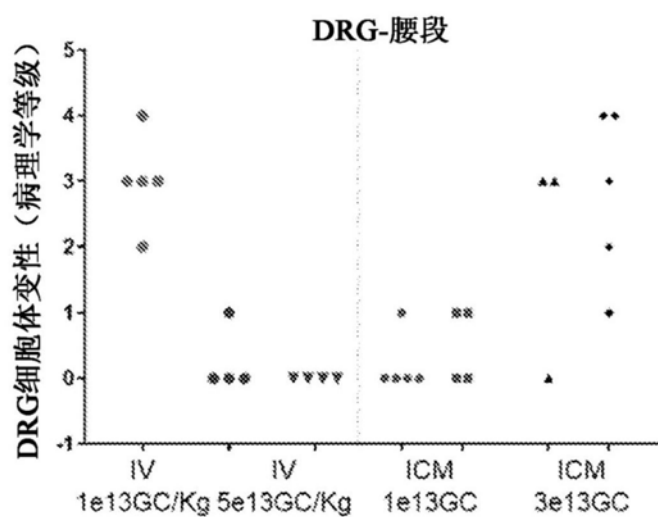


图33C

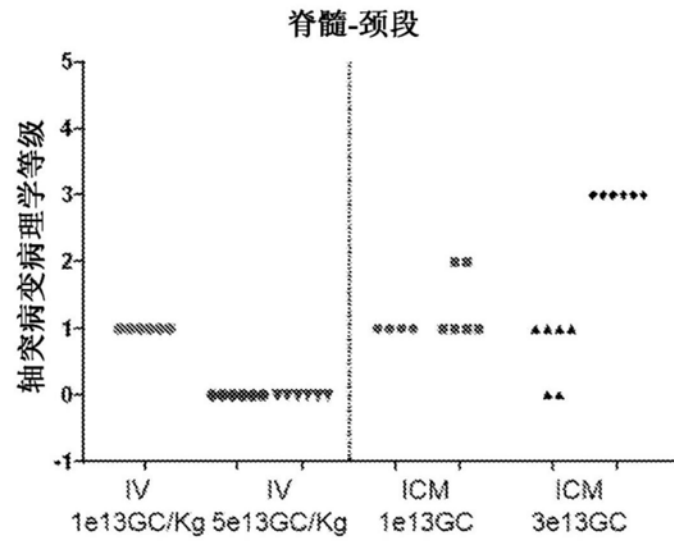


图33D

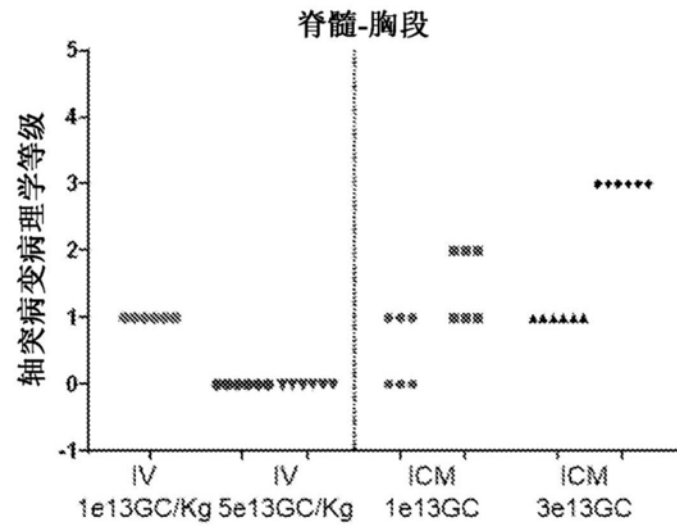


图33E

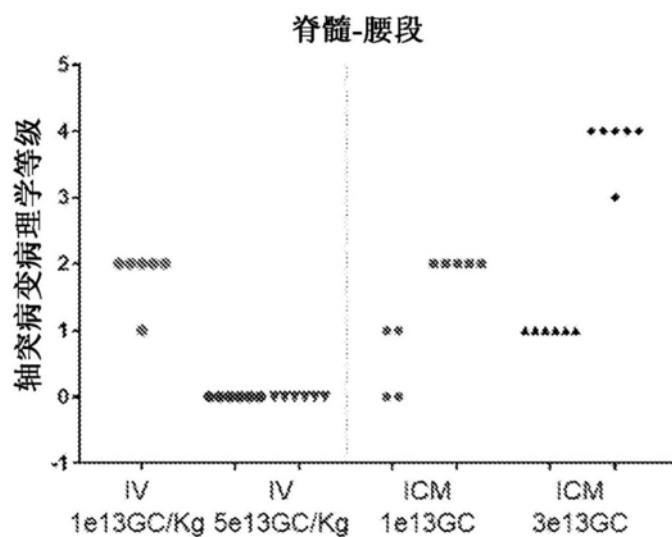


图33F

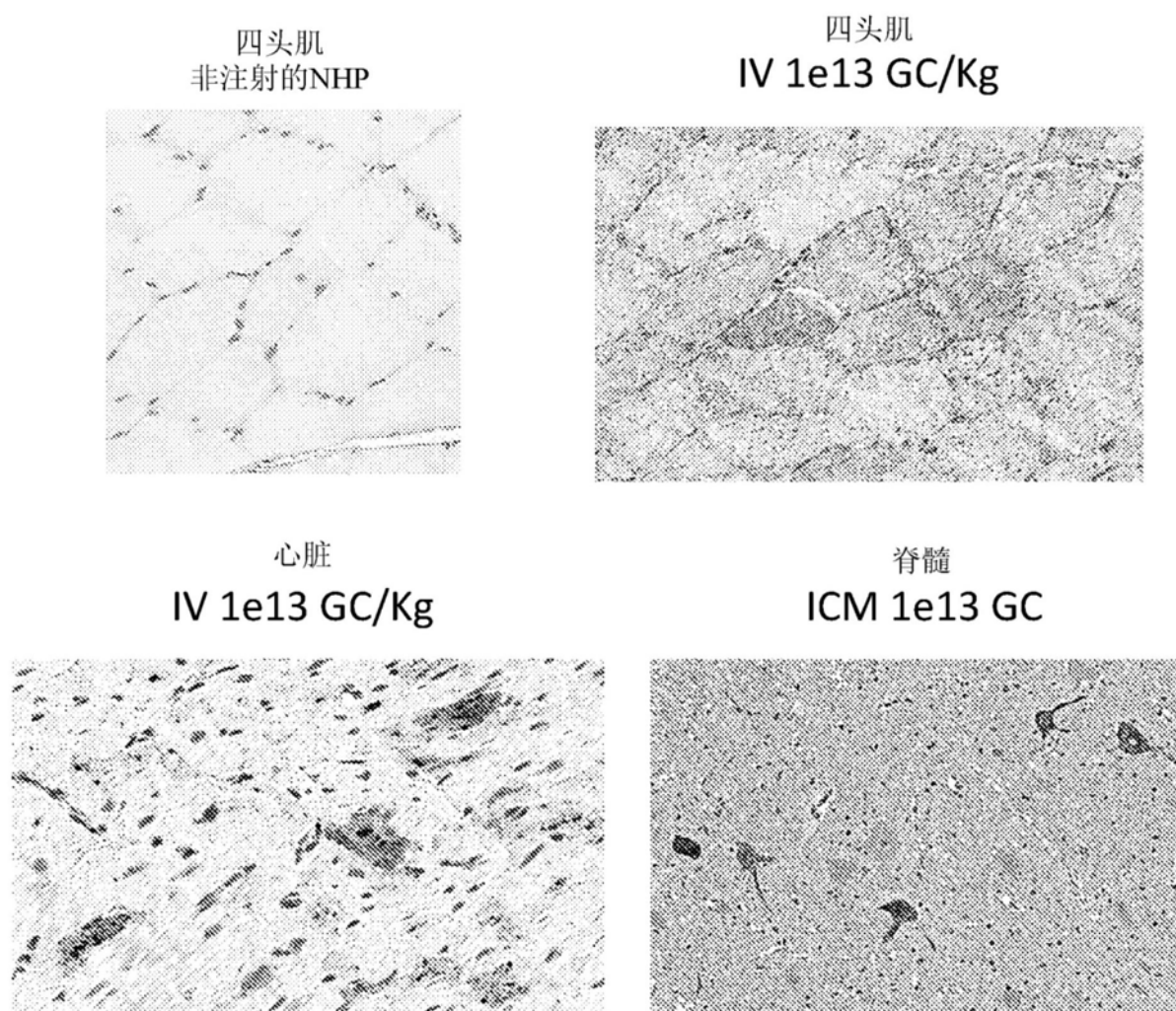


图34