(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115916334 A
(43)申请公布日 2023.04.04

(21)申请号 202180049054.5

(22)申请日 2021.05.14

(30)优先权数据
63/024941 2020.05.14 US
63/109677 2020.11.04 US
63/180379 2021.04.27 US

(85)PCT国际申请进入国家阶段日
2023.01.09

(86)PCT国际申请的申请数据
PCT/US2021/032451 2021.05.14

(87)PCT国际申请的公布数据
WO2021/231863 EN 2021.11.18

(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

(72)发明人 J·M·威尔逊 J·豪多 H·V·董 R·戈夏尔 S·图斯克

(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 罗文锋 梅黎

(51)Int.Cl.
A61P 3/00(2006.01)

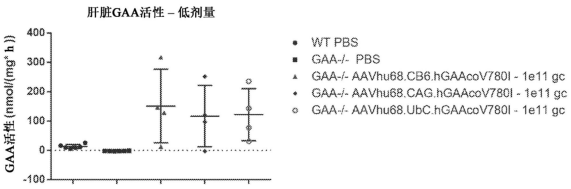
权利要求书2页 说明书49页
序列表97页 附图63页

(54)发明名称

可用于治疗庞贝病的组合物

(57)摘要

本文提供了一种用于在患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理的方法，其中所述患者已经被诊断患有庞贝病或被怀疑患有庞贝病。所述方法包括向所述患者施用具有AAV衣壳和包装在其中的载体基因组的重组AAV(rAAV)，其中所述载体基因组包括：(a)5'反向末端重复序列(ITR)；(b)启动子；(c)编码嵌合融合蛋白的核苷酸序列，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，(d)polyA；和(e)3'ITR。还提供了用于治疗患有或怀疑患有庞贝病的患者的包括本文所述的rAAV的药物组合物。



CN 115916334 A

1.一种用于在患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理的方法,其中所述患者已经被诊断患有庞贝病或被怀疑患有庞贝病,所述方法包括向所述患者施用具有AAV衣壳和包装在其中的载体基因组的重组AAV(rAAV),其中所述载体基因组包括:

(a)5'反向末端重复序列(ITR);

(b)启动子;

(c)编码嵌合融合蛋白的核苷酸序列,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,其中编码所述嵌合融合蛋白的所述序列与指导其表达的调节序列可操作地连接,并且包括SEQ ID NO:7或与其至少95%相同的编码SEQ ID NO:6的氨基酸1至氨基酸982的序列;

(d)polyA;和

(e)3'ITR。

2.根据权利要求1所述的方法,其中,所述启动子是组成型启动子,任选地CAG启动子或CB7启动子。

3.根据权利要求1所述的方法,其中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均,iv)自噬积聚,v)空泡化,和vi)无力。

4.根据权利要求1所述的方法,其中,所述患者患有晚发性庞贝病。

5.根据权利要求1所述的方法,其中,所述患者患有婴儿发病的庞贝病。

6.根据权利要求1所述的方法,其中,所述载体基因组进一步包括至少四个、至少五个、至少六个、至少七个或至少八个miR靶序列,任选地,其中每个所述miR靶序列对miR-183是特异性的。

7.根据权利要求1所述的方法,其中,所述AAV衣壳是AAVhu68衣壳。

8.根据权利要求1所述的方法,其中,所述rAAV通过静脉内和/或鞘内施用。

9.根据权利要求1所述的方法,其中,所述rAAV经由双重施用途径被施用于所述患者,任选地,其中所述双重途径是静脉内施用和枕大池内(ICM)施用。

10.根据权利要求1所述的方法,其中,所述rAAV以约1x 10^{11}基因组拷贝(GC)/kg至约5x 10^{13}GC/kg的剂量被静脉内施用。

11.根据权利要求1所述的方法,其中,所述rAAV以约1x 10^{12}GC至约5x 10^{13}GC的剂量经由ICM施用。

12.根据权利要求1所述的方法,其中,所述rAAV以约1x 10^{11}GC/kg至约5x 10^{13}GC/kg的剂量被IV施用并且以约1x 10^{12}GC至约5x 10^{13}GC的剂量经由ICM施用。

13.根据权利要求1所述的方法,其进一步包括所述患者接受施用联合疗法,任选地,其中所述联合疗法是支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌强度训练(RMST)、酶替代疗法和/或膈肌起搏疗法。

14.一种药物组合物,所述药物组合物包括具有AAV衣壳和包装在其中的载体基因组的重组AAV(rAAV),其中所述载体基因组包括:

(a)5'反向末端重复序列(ITR);

(b)启动子;

(c)编码嵌合融合蛋白的核苷酸序列,所述嵌合融合蛋白包括信号肽和与人酸性-α-葡

葡糖苷酶(hGAA)融合的vIGF2肽,其中编码所述嵌合融合蛋白的所述序列与指导其表达的调节序列可操作地连接,并且包括SEQ ID NO:7或与其至少95%相同的编码SEQ ID NO:6的氨基酸1至氨基酸982的序列;

(d)polyA;和

(e)3'ITR。

15.根据权利要求14所述的药物组合物,其中,所述启动子是组成型启动子,任选地CAG启动子或CB7启动子。

16.根据权利要求14或15所述的药物组合物,其中,所述载体基因组进一步包括至少四个、至少五个、至少六个、至少七个或至少八个miR靶序列,任选地,其中每个所述miR靶序列对miR-183是特异性的。

17.根据权利要求14至16中任一项所述的药物组合物,其中,所述AAV衣壳是AAVhu68衣壳。

18.根据权利要求14至17中任一项所述的药物组合物,其中,所述组合物被配制用于静脉内和/或鞘内递送。

19.根据权利要求14至18中任一项所述的药物组合物,其用于治疗患有庞贝病的患者,其中所述治疗在所述患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理。

20.根据权利要求19所述的药物组合物,其中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均,iv)自噬积聚,v)空泡化,和vi)无力。

21.根据权利要求19或20所述的药物组合物,其中,所述患者患有晚发性庞贝病。

22.根据权利要求19或20所述的药物组合物,其中,所述患者患有婴儿发病的庞贝病。

23.根据权利要求19至22中任一项所述的药物组合物,其中,所述rAAV经由双重施用途径被施用于所述患者,任选地,其中所述双重途径是静脉内施用和枕大池内(ICM)施用。

24.根据权利要求14至18中任一项所述的药物组合物,其适合于施用于已经被诊断患有庞贝病的症状后患者。

25.根据权利要求14至18中任一项所述的药物组合物,其适合于逆转患有庞贝病的症状后患者的异常肌肉病理。

26.根据权利要求25所述的药物组合物,其中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均,iv)自噬积聚,v)空泡化,和vi)无力。

27.根据权利要求14至18或24至26中任一项所述的药物组合物,其适用于联合疗法,任选地,其特征在于所述患者进一步接受使用支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌强度训练(RMST)、酶替代疗法和/或膈肌起搏疗法的治疗。

28.一种根据权利要求14至18中任一项所述的药物组合物用于在有需要的患者中治疗庞贝病的用途,其中,所述治疗在所述患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理。

可用于治疗庞贝病的组合物

背景技术

[0001] 庞贝病(Pompe disease)(也被称为II型糖原病)是一种由酸-α-葡萄糖苷酶(GAA)基因的突变引起的溶酶体贮积病,导致心脏(心肌病)、肌肉和运动神经元(神经肌肉疾病)中的糖原累积。在典型的婴儿庞贝病中,严重的GAA活性丧失导致多系统和早发性糖原贮积,尤其是在心脏和肌肉内,并且在头几年期间因心肺衰竭而死亡。婴儿庞贝病的特征还在于神经元(特别是运动神经元)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)对矫正肌肉的效率次优,并且不能穿过血脑屏障,导致典型婴儿庞贝病的长期存活者的进行性神经功能恶化。接受ERT的患者(其由于心脏矫正而活得更长)揭示了具有进行性神经表型的新自然病史。此外,重组人类GAA是高度免疫原性的,并且由于骨骼肌摄取不良,必须以非常大的量给药。

[0002] 对庞贝病的治疗存在几个未满足的需求,包含需要纠正该疾病的CNS组分,需要改善肌肉纠正,以及需要一种对当前ERT更有效、更低免疫原性和/或更方便的替代方案。

发明内容

[0003] 在一个方面中,本文提供了一种用于在患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理的方法,其中所述患者已经被诊断患有庞贝病或被怀疑患有庞贝病,所述方法包括向所述患者施用具有AAV衣壳和包装在其中的载体基因组的重组AAV(rAAV),其中所述载体基因组包括:(a)5'反向末端重复序列(ITR);(b)启动子;(c)编码嵌合融合蛋白的核苷酸序列,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,其中编码嵌合融合蛋白的序列与指导其表达的调节序列可操作地连接,并且包括SEQ ID NO:7或与其至少95%相同的编码SEQ ID NO:6的氨基酸1至氨基酸982的序列;(d)polyA;和(e)3'ITR。在某些实施例中,所述启动子是组成型启动子,任选地,CAG启动子或CB7启动子。在某些实施例中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均(anisocytosis),iv)自噬积聚,v)空泡化,和vi)无力。在某些实施例中,所述患者患有晚发性庞贝病。在某些实施例中,所述患者患有婴儿发病的庞贝病。在某些实施例中,所述载体基因组进一步包括至少四个、至少五个、至少六个、至少七个或至少八个miR靶序列,任选地,其中每个所述miR靶序列对miR-183是特异性的。在某些实施例中,所述AAV衣壳是AAVhu68衣壳。在某些实施例中,所述rAAV通过静脉内和/或鞘内施用。在某些实施例中,所述rAAV经由双重施用途径被施用于患者,任选地,其中双重途径是静脉内施用和枕大池内(ICM)施用。

[0004] 在一个方面中,本文提供了一种药物组合物,所述组合物包括具有AAV衣壳和包装在其中的载体基因组的重组AAV(rAAV),其中所述载体基因组包括:(a)5'反向末端重复序列(ITR);(b)启动子;(c)编码嵌合融合蛋白的核苷酸序列,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,其中所述编码嵌合融合蛋白的序列与指导其表达的调节序列可操作地连接,并且包括SEQ ID NO:7,或与其至少95%相同的编码

SEQ ID NO:6的氨基酸1至氨基酸982的序列;(d)polyA;和(e)3'ITR。在某些实施例中,所述启动子是组成型启动子,任选地,CAG启动子或CB7启动子。在某些实施例中,所述载体基因组包括至少四个、至少五个、至少六个、至少七个或至少八个miR靶序列,任选地,其中每个所述miR靶序列对miR-183是特异性的。在某些实施例中,所述AAV衣壳是AAVhu68衣壳。在某些实施例中,所述组合物被配制用于静脉内和/或鞘内递送。

[0005] 在一个方面中,本文提供了一种用于治疗患有庞贝病的患者的药物组合物,其中所述治疗在患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理。在某些实施例中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均,iv)自噬积聚,v)空泡化,和vi)无力。在某些实施例中,所述患者患有晚发性庞贝病(late-onset Pompe disease)。在某些实施例中,所述患者患有婴儿发病的庞贝病。在某些实施例中,所述rAAV经由双重施用途径被施用于患者,任选地,其中双重途径是静脉内施用和枕大池内(ICM)施用。

[0006] 在一个方面中,本文提供的药物组合物适合于施用于已经被诊断患有庞贝病的症状后患者。在某些实施例中,所述组合物适合于逆转患有庞贝病的症状后患者的异常肌肉病理。在某些实施例中,所述异常肌肉病理的特征在于以下中的一种或多种:i)具有中央核的肌肉细胞的百分比升高;ii)肌纤维萎缩,iii)肌细胞纤维中的红细胞大小不均,iv)自噬积聚,v)空泡化,和vi)无力。在某些实施例中,所述药物组合物适用于联合疗法,任选地,其特征在于所述患者进一步接受使用支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌强度训练(RMST)、酶替代疗法和/或膈肌起搏疗法的治疗。

[0007] 在一个方面,提供了药物组合物的用途,其包括施用本文所述的rAAV以治疗本文所提供的有需要的患者的庞贝病,其中所述治疗在患者中减少异常肌肉病理的进展和/或逆转患者的异常肌肉病理。

[0008] 根据本发明的以下详细描述,本发明的其他方面和优点将变得显而易见。

附图说明

[0009] 图1A和图1B示出了在CB6(第三列)、CAG(第四列)或UbC启动子(最后一列)的指导下,静脉内施用具有hGAAV780I的工程改造的编码序列的各种AAVhu68.hGAA四周后Pompe(-/-)小鼠肝脏中的hGAA活性。(图1A)低剂量($1x10^{11}$GC)。(图1B)高剂量($1x10^{12}$)。

[0010] 图2A和图2B示出了在CB6(第三列)、CAG(第四列)或UbC启动子(最后一列)的指导下,静脉内施用具有hGAAV780I的工程改造的编码序列的各种AAVhu68.hGAA四周后Pompe(-/-)小鼠心脏中的hGAA活性。(图2A)低剂量($1x10^{11}$GC)。(图2B)高剂量($1x10^{12}$)。

[0011] 图3A和图3B示出了在CB6(第三列)、CAG(第四列)或UbC启动子(最后一列)的指导下,静脉内施用具有hGAAV780I的工程改造的编码序列的各种AAVhu68.hGAA四周后Pompe(-/-)小鼠的骨骼肌(四头肌)中的hGAA活性。(图3A)低剂量($1x10^{11}$GC)。(图3B)高剂量($1x10^{12}$)。

[0012] 图4A和图4B示出了在CB6(第三列)、CAG(第四列)或UbC启动子(最后一列)的指导下,静脉内施用具有hGAAV780I的工程改造的编码序列的各种AAVhu68.hGAA四周后Pompe(-/-)小鼠的大脑中的hGAA活性。(图4A)低剂量($1x10^{11}$GC)。(图4B)高剂量($1x10^{12}$)。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。

[0013] 图5A至图5H示出了递送AAVhu68.hGAA四周后Pompe小鼠的心脏组织学(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。媒介物对照Pompe(-/-)(图5D)和野生型(+/+)(图5A)小鼠接受PBS注射。“hGAA”是指由具有天然信号肽的野生型序列编码的参考天然酶(hGAAV780)(图5B)。“BiP-vIGF2.hGAAco”是指参考hGAAV780蛋白的工程改造的编码序列,其含有前35AA的缺失,并进一步具有BiP信号肽,与对胰岛素受体具有低亲和力的IGF2变体融合(图5C)。(图5D)来自媒介物处理的对照的图像。“hGAAcoV780I”是指由工程改造的序列编码并含有天然信号肽的hGAAV780I变体(图5E)。“BiP-vIGF2.hGAAcoV780I”是指hGAAcoV780I,其含有前35AA的缺失,并且进一步具有与对胰岛素受体具有低亲和力的IGF2变体融合的BiP信号肽和由工程改造的序列编码的hGAAV780I(图5F)。“Sp7.Δ8.hGAAcoV780I”是指hGAAV780I变体,其缺失了由与先前构建体相同的工程改造的序列编码的前35AA,但含有编码B2胰凝乳蛋白酶原信号肽的序列来代替天然信号肽(图5G)。(图5H)盲法组织病理学半定量严重程度评分。经委员会认证的兽医病理家以盲法方式审查了载玻片,并且基于糖原贮积和自噬积聚建立了严重程度评分。

[0014] 图6A至图6H示出了施用编码各种hGAA($2.5x10^{13}$GC/kg)的AAVhu68四周后Pompe小鼠的四头肌肌肉的组织学(PAS染色)的结果。对照Pompe(-/-)(图6D)和野生型(+/+)(图6A)小鼠接受PBS注射。“hGAA”是指由具有天然信号肽的野生型序列编码的参考天然酶(hGAAV780)(图6B)。“hGAAcoV780I”是指由工程改造的序列编码并含有天然信号肽的hGAAV780I变体(图6E)。“Sp7.Δ8.hGAAcoV780I”是指hGAAV780I变体,其缺失了由与先前构建体相同的工程改造的序列编码的前35AA,但含有编码B2胰凝乳蛋白酶原信号肽的序列来代替天然信号肽(图6F)。“BiP-vIGF2.hGAAco”是指参考hGAAV780,其含有前35AA的缺失,并且进一步具有BiP信号肽,与对胰岛素受体具有低亲和力的IGF2变体融合,并且由工程改造的序列编码(图6C)。“BiP-vIGF2.hGAAcoV780I”是指hGAAV780I,其含有前35AA的缺失,并且进一步具有与对胰岛素受体具有低亲和力的IGF2变体融合的BiP信号肽和由工程改造的序列编码的hGAAV780I(图6G)。(图6H)盲法组织病理学半定量严重程度评分。经委员会认证的兽医病理家以盲法方式审查了载玻片,并且基于糖原贮积和自噬积聚建立了严重程度评分。0的评分意指没有病变;1意指平均不到9%的肌纤维受到贮积的影响;2意指10%至49%的肌纤维受到贮积的影响;3意指50%至75%的肌纤维受到贮积的影响,并且4意指76%至100%的肌纤维受到贮积的影响。

[0015] 图7A至图7H示出了在以$2.5x10^{12}$GC/Kg下(即,比图6A至图6H中的剂量低10倍)施用编码各种hGAA的AAVhu68四周后来自Pompe小鼠的四头肌肌肉的组织学结果(高碘酸-希夫(PAS)染色)。对照Pompe(-/-)(图7D)和野生型(+/+)(图7A)小鼠接受PBS注射。“hGAA”是指由具有天然信号肽的野生型序列编码的参考天然酶(hGAAV780)(图7B)。“hGAAcoV780I”是指由工程改造的序列编码并含有天然信号肽的hGAAV780I变体(图7E)。“Sp7.Δ8.hGAAcoV780I”是指hGAAV780I变体,其缺失了由与先前构建体相同的工程改造的序列编码的前35AA,但含有编码B2胰凝乳蛋白酶原信号肽的序列来代替天然信号肽(图7F)。“BiP-vIGF2.hGAAco”是指参考hGAAV780,其含有前35AA的缺失,并且进一步具有BiP信号肽,与对胰岛素受体具有低亲和力的IGF2变体融合,并且由工程改造的序列编码(图7C)。“BiP-vIGF2.hGAAcoV780I”是指hGAAV780I,其含有前35AA的缺失,并且进一步具有与对胰岛素受

体具有低亲和力的IGF2变体融合的BiP信号肽和由工程改造的序列编码的hGAAV780I(图7G)。(图7H)盲法组织病理学半定量严重程度评分。经委员会认证的兽医病理家以盲法方式审查了载玻片,并且基于糖原贮积和自噬积聚建立了严重程度评分。0的评分意指没有病变;1意指平均不到9%的肌纤维受到贮积的影响;2意指10%至49%的肌纤维受到贮积的影响;3意指50%至75%的肌纤维受到贮积的影响,并且4意指76%至100%的肌纤维受到贮积的影响。

[0016] 图8示出了来自在施用($2.5x10^{12}$GC/kg)具有编码天然hGAA或含有前35AA的缺失的hGAAV780I的AAVhu68四周后Pompe小鼠的脊髓的组织学(PAS和luxol fast blue染色)的结果,并且进一步具有与对胰岛素受体具有低亲和力的IGF2变体融合的BiP信号肽和由工程改造的序列编码的hGAAV780I(“BiP-vIGF2.hGAAcoV780I”)。对脊髓切片进行盲法组织病理学半定量严重程度评分。

[0017] 图9A至图9C示出了血浆中hGAA活性以及与IGF2/CI-MPR的结合。Pompe小鼠被施用低剂量($2.5x10^{12}$GC)的编码野生型hGAA或BiP-vIGF2.hGAA的载体。(图9A,图9B)静脉内施用四周后,在血浆中检测到高水平的野生型和工程改造的hGAA活性。(图9C)工程改造的hGAA有效地结合至CI-MPR。

[0018] 图10示出了以$2.5x10^{12}$GC/Kg(LD)的剂量施用AAVhu68构建体四周后Pompe小鼠中的糖原清除率和自噬积聚的消除。腓肠肌的石蜡切片用DAPI和抗LC3B抗体染色。

[0019] 图11示出了BiP-vIGF2.hGAAcoV780I.4xmiR183构建体的示意图。

[0020] 图12示出了在静脉内施用高剂量(HD:$2.5x10^{13}$GC/kg)或低剂量(LD:$2.5x10^{12}$GC/kg)的AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183(含有drg去靶向序列的四个拷贝,miR183)四周后Pompe小鼠的脑干中的糖原贮积(PAS,luxol蓝染色)。箭头示出了神经元内的PAS阳性贮积。

[0021] 图13示出了在静脉内施用高剂量(HD:$2.5x10^{13}$GC/kg)或低剂量(LD:$2.5x10^{12}$GC/kg)的AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183四周后Pompe小鼠的脊髓中的糖原贮积(PAS,luxol蓝染色)。箭头示出了神经元内的PAS阳性贮积。

[0022] 图14示出了在静脉内施用高剂量(HD:$2.5x10^{13}$GC/kg)或低剂量(LD:$2.5x10^{12}$GC/kg)的AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183四周后Pompe小鼠的四头肌肌肉中的糖原贮积(PAS染色)。

[0023] 图15示出了在静脉内施用高剂量(HD:$2.5x10^{13}$GC/kg)或低剂量(LD:$2.5x10^{12}$GC/kg)的AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183四周后Pompe小鼠的心脏中的糖原贮积(PAS染色)。

[0024] 图16示出了在静脉内施用高剂量(HD:$2.5x10^{13}$GC/kg)或低剂量(LD:$2.5x10^{12}$GC/kg)的AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183四周后Pompe小鼠的四头肌肌肉中的自噬空泡标记物LC3b的表达。

[0025] 图17示出了在以$3x10^{13}$GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I(左)或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183(右)35天后恒河猴颈部DRG中hGAA表达的代表性图像(hGAA的免疫组织化学)。

[0026] 图18示出了在以$3x10^{13}$GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I(左)或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183(右)35天后恒河猴腰椎DRG中hGAA

表达的代表性图像(hGAA的免疫组织化学)。

[0027] 图19示出了在以3x10^{13}GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I(左)或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183(右)35天后恒河猴脊髓下运动神经元中hGAA表达的代表性图像(hGAA的免疫组织化学)。

[0028] 图20示出了在以3x10^{13}GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I(左)或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183(右)35天后恒河猴心脏中hGAA表达的代表性图像(hGAA的免疫组织化学)。

[0029] 图21A至图21C示出了在以3x10^{13}GC的高剂量ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183 35天后恒河猴颈段(图21A)、胸段(图21B)和腰段(图21C)的DRG中DRG神经元变性和炎症细胞浸润的组织病理评分。如前所述(Katz等人,《人类基因治疗方法(Hum Gene Ther.Methods),2018,29:212-9),在荧光镜的指导下,将AAVhu68载体以总体积为1mL的无菌人工CSF(媒介物)注射到枕大池中对载体组不知情的经委员会认证的兽医病理家建立了严重程度等级,其定义为0为无病变,1为最小(<10%),2为轻度(10%至25%),3为中度(25%至50%),4为显著(50%至95%),并且5为严重(>95%)。每个数据点代表一个DRG。对每个片段和每个动物最少五个DRG进行评分。

[0030] 图22A至图22C示出了在3x10^{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)。

[0031] 图23示出了在注射后第0天至第35天,以3x10^{13}GC的高剂量(ICM)NHP施用的(ICM)AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183的血浆hGAA活性水平。

[0032] 图24A至图24G示出了在基线和NHP施用(ICM、3x10^{13}GC)AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I或AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.4XmiR183的第35天的来自神经传导速度测试的结果。

[0033] 图25A和图25B示出了在Pompe小鼠中从载体注射(第0天)至注射后180天的体重纵向随访,该小鼠在7个月大时在疾病晚期被治疗并且在基线时已经有症状。他们经由替代的施用途径和剂量水平接受AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I:在高剂量(HD)(1x10^{11}GC)或低剂量(LD)(5x10^{10}GC)下的侧脑室(ICV)、在HD(5x10^{13}GC/Kg)或LD(1x10^{13}GC/Kg)下的静脉内(IV),以及在低剂量或高剂量下的ICV和IV的组合。描绘了平均值和标准偏差。在KO PBS对照组和其他组之间,通过Wilcoxon-Mann-Whitney检验进行每个时间点的统计分析。*p<0.05;**p<0.01

[0034] 图26和图27示出了在Pompe小鼠中从载体注射(第0天)至注射后180天相对于体重的握力的纵向随访,该小鼠在7个月大时在疾病晚期被治疗并且在基线时已经有症状。(图26)小鼠经由替代的施用途径和剂量水平接受AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I:在高剂量(ICV HD:1x10^{11}GC)下的侧脑室(ICV)、在高剂量(IV HD:5x10^{13}GC/Kg)下的静脉内(IV),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))在不同的时间点测量握力。抓力度计中的传感器(transducer)被连接到与阳极化基板连接的丝网栅格。动物被它的尾巴抓住,并且轻轻地从栅格经过,直到它用其四只爪

子抓住栅格。进行三次握力测量，并且这些读数的平均值代表动物在该特定时间的握力。(图27)来自第180天的结果示出了IV+ICV HD相对于IV HD的增量益处。数值通过动物体重被归一化。N＝每组4只雄性和4只雌性。通过单向ANOVA(图26)或双向ANOVA(图27)确定每个时间点的统计分析，与KO PBS对照组比较的事后多重比较测试。*p<0.05、**p<0.01、***p<0.001

[0035] 图28示出了在高剂量(HD:$1x10^{11}$GC)或低剂量(LD:$5x10^{10}$GC)ICV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后，症状后Pompe小鼠的四头肌、心脏和脊髓中的糖原贮积。

[0036] 图29示出了在高剂量(HD:$5x10^{13}$GC/Kg)或低剂量(LD:$1x10^{13}$GC/Kg)IV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后，症状后Pompe小鼠的四头肌、心脏和脊髓中的糖原贮积。

[0037] 图30A至图30C示出了在第30天(图30A)、第60天(图30B)和第90天(图30C)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I载体IV、ICV或IV和ICV(双重途径)的Pompe小鼠的血浆中的hGAA活性。

[0038] 图31示出了用于评估NHP中单一(IV或ICM)和双重施用途径(IV+ICM)的研究设计。

[0039] 图32A至图32H示出了在IV或ICM施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I，hGAA的检测(图32A和图32C-血浆；图32B和图32D-CSF)和hGAA活性(图32E和图32G-血浆；图32F和图32H-CSF)。

[0040] 图33A至图33F示出了在IV($1x10^{13}$GC/Kg或$5x10^{13}$GC/Kg)、ICM($1x10^{13}$GC或$3x10^{13}$GC)、或IV+ICM($5x10^{13}$GC/Kg IV+$3x10^{13}$GC ICM或$1x10^{14}$GC/Kg IV+$1x10^{13}$GC ICM)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后，恒河猴的颈段(图33A)、胸段(图33B)和腰段(图33C)中的DRG神经元变性和炎症细胞浸润以及颈段(图33D)、胸段(图33E)和腰段(图33F)中的脊髓轴突病变的组织病理评分。对载体组不知情的经委员会认证的兽医病理家建立了严重程度等级，其定义为0为无病变，1为最小(<10％)，2为轻度(10％至25％)，3为中度(25％至50％)，4为显著(50％至95％)，并且5为严重(>95％)。

[0041] 图34示出了在IV($1x10^{13}$GC/Kg或$5x10^{13}$GC/Kg)、ICM($1x10^{13}$GC或$3x10^{13}$GC)、或IV+ICM($5x10^{13}$GC/Kg IV+$3x10^{13}$GC ICM)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后，恒河猴血浆中的GAA活性。

[0042] 图35A至图35D示出了在IV($1x10^{13}$GC/Kg或$5x10^{13}$GC/Kg)、ICM($1x10^{13}$GC或$3x10^{13}$GC)、或IV+ICM($5x10^{13}$GC/Kg IV+$3x10^{13}$GC ICM)施用AAVhu68.CAG.BiP-vIGF2.hGAA.coV780I后，来自恒河猴的肝脏和心脏(图35A)以及膈肌、三头肌和胫骨前肌(图35B)中GAA活性的测量。

[0043] 图36示出了在IV($1x10^{13}$GC/Kg或$5x10^{13}$GC/Kg)、ICM($1x10^{13}$GC或$3x10^{13}$GC)、或IV+ICM($5x10^{13}$GC/Kg IV+$3x10^{13}$GC ICM)施用AAVhu68.CAG.BiP-vIGF2.hGAA.coV780I后，来自恒河猴的血清中的抗GAA抗体滴度。

[0044] 图37示出了在低剂量(IV-$1x10^{13}$GC/Kg、ICM-$1x10^{13}$GC)施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I后，恒河猴的四头肌、心脏和脊髓中hGAA表达的代表性图像(hGAA的免疫组织化学)。

[0045] 图38示出了在以高剂量(HD:$2.5x10^{12}$GC/kg)静脉内施用BiP-

vIGF2.hGAAcoV780I.4XmiR183(AAV.hGAAeng)或由具有天然信号肽的野生型序列(AAV.hGAAnat)编码的hGAAV780I四周后,Pompe小鼠四头肌肌肉和心脏GAA活性和糖原贮积的分析(PAS染色)。

[0046] 图39示出了在以高剂量(HD:2.5x10^{13}GC/kg)静脉内施用BiP-vIGF2.hGAAcoV780I.4XmiR183(AAV.hGAAeng)或由具有天然信号肽的野生型序列(AAV.hGAAnat)编码的hGAAV780I四周后,Pompe小鼠CNS GAA活性和糖原贮积的分析(PAS染色)。

[0047] 图40示出了在静脉内施用BiP-vIGF2.hGAAcoV780I.4XmiR183(IV 2.5x10^{13}GC/kg)四周后,Pompe小鼠的四头肌、心脏和脊髓中hGAA表达的代表性图像(hGAA的免疫组织化学)。

[0048] 图41示出了关于评估症状前(年轻)Pompe小鼠的治疗的研究的研究概况,其包含IV施用不同剂量的AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183。

[0049] 图42示出了来自年轻WT、对照PBS处理的GAA-/-和AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183处理的GAA-/-小鼠的四头肌肌肉切片的代表性免疫荧光图像。WGA(细胞膜)、DAPI(细胞核)和LC3b抗体(自噬体)。

[0050] 图43示出了来自施用AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183的年轻Pompe小鼠的四头肌肌肉中央核的量化。

[0051] 图44示出了来自施用AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183的年轻Pompe小鼠的四头肌肌肉中LC3b染色后自噬积聚的量化。

[0052] 图45示出了来自施用AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183的年轻Pompe小鼠的四头肌肌肉中肌纤维直径的量化。

[0053] 图46A至图46F示出了在比目鱼肌(图46A)、膈肌(图46B)、四头肌(图46C)、三头肌(图46D)、腓肠肌(图46E)和胫骨前肌(图46F)中IV施用不同剂量的AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183后,肌肉溶酶体贮积病理的半定量评分(空泡化的严重程度)。正方形代表雄性小鼠,圆形代表雌性小鼠。统计:单向ANOVA(Kruskall Wallis检验),然后进行事后Dunn多重比较检验。与GAA KO Pompe PBS组相比,星号示出显著性(第二列)。评分:具有空泡的肌纤维的比例-0:0%,1:1%至9%,2:10%至24%,3:25%至49%,4:50%至74%,和5:>5%.

[0054] 图47提供了在Pompe GAA敲除小鼠中评估IV、ICV和IV+ICV施用途径的研究设计。

[0055] 图48示出了来自6-7个月大的WT、对照PBS处理的GAA-/-和AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183处理的GAA-/-小鼠的四头肌肌肉切片的代表性免疫荧光图像。WGA(细胞膜)、DAPI(细胞核)和LC3b抗体(自噬体)。

[0056] 图49示出了来自施用给予AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183的症状后Pompe小鼠的四头肌肌肉中的中央核的量化。

[0057] 图50示出了在IV、ICV和IV+ICV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG后,对来自GAA KO小鼠的四头肌肌肉组织进行LC3b染色后,自噬积聚的量化。

[0058] 图51示出了在IV、ICV和IV+ICV施用AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG后,Pompe GAA敲除小鼠中的肌纤维大小分布。纤维直径被分为S=小(<30μm)、M=中(30-50μm)和L=大(>50μm)的类别。

[0059] 图52示出了对非人灵长类IV、ICM或IV+ICM施用AAVhu68.CAG.BiP-IGF2-hGAAcoV780I载体后脊髓段中表达运动神经元的hGAA的定量。

具体实施方式

[0060] 提供了用于向患有庞贝病的患者递送融合蛋白的组合物，该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组合物的方法，包含用这些组合物治疗患者的方案。

[0061] 如本文所用，术语“庞贝病”(也被称为麦芽糖酶缺乏症、糖原贮积病II型(GSDII)或糖原病II型)旨在是指遗传性溶酶体贮积障碍，其特征在于由编码酸性α-葡萄糖苷酶的GAA基因的突变引起的溶酶体酶酸a-葡萄糖苷酶(GAA)的完全缺失或部分缺失。该术语包含但不限于疾病的早发性和晚发性，包含但不限于婴儿、青少年和成人发病的庞贝病。

[0062] 应当理解，希腊字母“α”和符号“α”在整个本说明书中可互换使用。类似地，希腊字母“Δ”和“Δ”在整个本说明书中可互换使用。

[0063] 如本文所用，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。替代的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(国家生物技术信息中心(National Centre for Biotechnology Information)(NCBI)基因ID 2548)编码，该基因已经被映射到17号染色体的长臂(位置17q25.2-q25.3)。氨基酸残基516-521处的保守六肽WIDMNE是酸性α-葡萄糖苷酶蛋白的活性所必需的。术语“hGAA”是指人类GAA的编码序列。

[0064] 如本文所用，“rAAV.hGAA”是指具有AAV衣壳的rAAV，在该衣壳中包装有载体基因组，该基因组至少含有GAA酶的编码序列(例如，780I变体，包括信号肽和至少与hGAA780I酶的活性部分融合的vIGF2肽的融合蛋白)。rAAVhu68.hGAA或rAAVhu68.hGAA是指其中AAV衣壳是AAVhu68衣壳的rAAV，如本文所定义。

[0065] 参照全长hGAA的编号，在氨基酸位置1至氨基酸位置27处存在信号肽。此外，该酶与多种成熟蛋白相关，即氨基酸位置70至氨基酸位置952处的成熟蛋白、位于氨基酸位置123至氨基酸位置952处的76kD成熟蛋白和氨基酸204至氨基酸位置952处的70kD成熟蛋白。“活性催化位点”包括六肽WIDMNE(SEQ ID NO:3的氨基酸残基516至氨基酸残基521处)。在某些实施例中，可以选择更长的片段，例如位置516至位置616。其他活性位点包括配体结合位点，其可以位于位置376、404、405、441、481、516、518、519、600、613、616、649、674中的一个或多个处。

[0066] 除非另有说明，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(pre-pro-protein)。在一些情况下，术语hGAAco780I或hGAAcoV780I用于指编码hGAA780I的工程改造的序列。如与前段中描述的hGAA参考蛋白相比，hGAA780I在位置780处具有异亮氨酸(Ile或I)，其中参考hGAA含有缬氨酸(Val或V)。出人意料地发现这种hGAA780I比在位置780处具有缬氨酸的hGAA序列(hGAAV780)具有更好的效果和改善的安全性概况，该序列在文献中已经被广泛地描述为“参考序列”。例如，如从图5A至图5H中可以看出，hGAAV780参考序列诱导毒性(纤维化心肌炎)，这在hGAA780I酶的相同剂量中没有观察到。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在SEQ ID NO:3中复制的hGAA780I中的类似位置来确定，即在氨基酸位置1至氨基酸位置27的信号肽；在氨基酸位置70至氨基酸位置952的成熟蛋白质；位于氨基酸位置123至氨基酸位置952的76kD成熟蛋白，和在氨基酸204至氨基酸952的70kD成熟蛋白；在氨基酸残基516至氨基酸残基521处包括六肽WIDMNE(SEQ ID NO:61)的“活性催化位点”；其他活性位点包含配体结合位点，其可以位于位置376、404、405、441、481、516、518、519、600、613、616、649、674中的一个或多个处。

[0067] 在某些实施例中，可以选择hGAA780I，其具有与hGAA780I至少95％相同、与hGAA780I至少97％相同、或与SEQ ID NO:3的hGAA780I至少99％相同的序列。在某些实施例中，提供了与SEQ ID NO:3的成熟hGAA780I蛋白至少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％至至少99％同一性的序列保留了活性催化位点的序列而没有任何改变。在某些实施例中，与SEQ ID NO:3的hGAA780I具有至少95％至至少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，与参考hGAAV780相比，具有改进的生物学效应和更好的安全性概况。在某些实施例中，可以如先前所述(参见，例如，J.Hordeaux等人，《神经病理通讯学报(Acta Neuropathological Communications)》，(2107)5:66)或使用其他合适的方法进行GAA活性测定。在某些实施例中，hGAA780I酶在hGAA氨基酸序列的其他位置中含有修饰。在某些实施例中，此类突变体hGAA780I可以最少保留780I的区域中的活性催化位点:WIDMNE(SEQ ID NO:61)和氨基酸，如下所述。

[0068] 在某些实施例中，提供了新的hGAA780I融合蛋白，其包括除天然hGAA信号肽以外的前导肽。在某些实施例中，此类外源性前导肽优选地来源于人类，并且可以包含例如IL-2前导肽。在某些实施例中可用的特定外源性信号肽包含来自胰凝乳蛋白酶原B2的氨基酸1至氨基酸20、人类α-1-抗胰蛋白酶的信号肽、来自伊杜糖醛酸-2-硫酸酯酶的氨基酸1至氨基酸25和来自蛋白酶CI抑制剂的氨基酸1至氨基酸23。参见，例如WO2018046774。其他信号/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。

[0069] 此类嵌合hGAA780I可以具有外源性前导序列以代替整个27aa天然信号肽。任选地，hGAA780I酶的N-末端截短可能仅缺失信号肽的一部分(例如，缺失约2个氨基酸至约25个氨基酸，或其间的值)、完整的信号肽、或比信号肽更长的片段(例如，基于SEQ ID NO:3的编号，多达70个氨基酸)。任选地，这种酶可以含有长度约为5、10、15或20个氨基酸的C-末端截短。

[0070] 在某些实施例中，提供了新的融合蛋白，其包括与融合伴侣结合的成熟hGAA780I蛋白(aa 70至aa 952)、成熟70kD蛋白(aa 123至aa 952)或成熟76kD蛋白(aa 204至aa 952)。任选地，融合蛋白进一步包括对于hGAA非天然的信号肽。进一步任选地，这些实施例中的一个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。

[0071] 在某些实施例中，包括hGAA780I蛋白的融合蛋白包括SEQ ID NO:3的至少氨基酸204至氨基酸890(hGAA780I)，或与其至少95％相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白包括SEQ ID NO:3的至少氨基酸204至氨基酸952或与其至少95％相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白包括SEQ ID NO:3的至

少氨基酸123至氨基酸890或与其至少95％相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白包括SEQ ID NO:3的至少氨基酸70至氨基酸952或与其至少95％相同的在位置780处具有Ile的序列。在某些实施例中，hGAA780I蛋白包括SEQ ID NO:3的至少氨基酸70至氨基酸890或与其至少95％相同的在位置780处具有Ile的序列。

[0072] 在某些实施例中，融合蛋白包括信号序列和前导序列以及hGAA780I序列，hGAA780I序列与SEQ ID NO:7具有至少95％的同一性、至少97％的同一性或至少99％的同一性，融合蛋白的活性位点没有变化和/或活性位点的氨基酸3至氨基酸12的N-末端和/或C-末端没有变化。在优选的实施例中，工程改造的hGAA表达盒至少编码以下的人类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，工程改造的hGAA表达盒编码较长的人类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，工程改造的hGAA表达盒编码至少对应于以下的片段：PLGT-Trp(W)-Tyr(Y)-Asp(D)-LQTVP-780I-EALG-(Ser或S)-L-PPPPAA序列(SEQ ID NO:64)。类似地，在优选的实施例中，在活性结合位点(SEQ ID NO:3的aa 518至aa 521)中没有氨基酸变化。在某些实施例中，在位置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延伸段、任选的ER蛋白水解肽和缺失hGAA的前35个氨基酸(即，缺少天然信号肽和氨基酸28至氨基酸35)的hGAA780I变体。

[0073] 在某些实施例中，提供了分泌型工程改造的GAA，其包括BiP信号肽、IGF2+2GS延伸段和hGAA 780I的氨基酸61至氨基酸952(缺失hGAA780I的氨基酸1至氨基酸60)。在某些实施例中，本文提供了一种包括SEQ ID NO:6或与其至少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或与其至少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，或与其至少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，或与其至少95％相同的序列。

[0074] 下文进一步描述了本文提供的融合蛋白的组分。

[0075] 结合CI-MPR的肽

[0076] 本文提供了结合CI-MPR的肽(例如，vIGF2肽)。当从基因疗法载体表达时，包括此类肽和hGAA780I蛋白的融合蛋白将hGAA780I靶向到需要它的细胞，增加通过此类细胞的细胞摄取，并将治疗性蛋白靶向到亚细胞位置(例如溶酶体)。在一些实施例中，肽与hGAA780I蛋白的N-末端融合。在一些实施例中，肽与hGAA780I蛋白的C-末端融合。在一些实施例中，肽是vIGF2肽。一些vIGF2肽保持与CI-MPR的高亲和力结合，而它们对IGF1受体、胰岛素受体和IGF结合蛋白(IGFBP)的亲和力被降低或消除。因此，与野生型IGF2相比，一些变体IGF2肽实质上更具选择性，并且具有降低的安全风险。本文中的vIGF2肽包含具有SEQ ID NO:46的氨基酸序列的肽。变体IGF2肽进一步包含与野生型IGF2(SEQ ID NO:34)相比在6、26、27、43、48、49、50、54、55或65的位置处具有变体氨基酸的那些肽。在一些实施例中，vIGF2肽具有由E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R和K65R组成的组的一个或多个取代的序列。在一些实施例中，vIGF2肽具有E6R取代的序列。在一些实施例中，vIGF2肽具有F26S取代的序列。在一些实施例中，vIGF2肽具有Y27L取代的序列。在一些实施例中，vIGF2肽具有V43L取代的序列。在一些实施例中，vIGF2肽具有F48T取代的序列。在一些实施例中，vIGF2肽具有R49S取代的序列在一些实施例中，vIGF2肽具有S50I取代的序列。在一些实施例中，vIGF2肽具有A54R取代的序列。在一些实施例中，vIGF2肽具有L55R取代的序列。在一

些实施例中，vIGF2肽具有K65R取代的序列。在一些实施例中，vIGF2肽具有E6R、F26S、Y27L、V43L、F48T、R49S、S50I、A54R和L55R取代的序列。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-末端缺失以及E6R、Y27L和K65R的取代。在一些实施例中，vIGF2肽具有四个氨基酸的N-末端缺失以及E6R和Y27L的取代。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-末端缺失以及Y27L和K65R的取代。

[0077]

IGF2 氨基酸序列（变体残基用下划线标出）		
肽	序列	SEQ ID NO:
野生型	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRSCDLALLETYCATPAKSE	32
F26S	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRSCDLALLETYCATPAKSE	33
Y27L	AYRPSETLCGGELVDTLQFVCGDRGFLFSRPASRVS RRSRGIVEECCFRSCDLALLETYCATPAKSE	34
V43L	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGILEECCFRSCDLALLETYCATPAKSE	35
F48T	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCTRSCDLALLETYCATPAKSE	36

[0078]

R49S	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFSSCDLALLETYCATPAKSE	37
S50I	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRICDLALLETYCATPAKSE	38
A54R	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRSCDLRLLETYCATPAKSE	39
L55R	AYRPSETLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRSCDLARLETYCATPAKSE	40
F26S、Y27L、V43L、F48T、R49S、S50I、A54R、L55R	AYRPSETLCGGELVDTLQFVCGDRGSLFSRPASRVS RRSRGILEECCTSICDLRRLETYCATPAKSE	41
Δ1-6、Y27L、K65R	TLCGGELVDTLQFVCGDRGFLFSRPASRVSRRSRGI VEECCFRSCDLALLETYCATPARSE	42
Δ1-7、Y27L、K65R	LCGGELVDTLQFVCGDRGFLFSRPASRVSRRSRGIV EECCFRSCDLALLETYCATPARSE	43
Δ1-4、E6R、Y27L、K65R	SRTLCGGELVDTLQFVCGDRGFLFSRPASRVSRRSR GIVEECCFRSCDLALLETYCATPARSE	44
Δ1-4、E6R、Y27L	SRTLCGGELVDTLQFVCGDRGFLFSRPASRVSRRSR GIVEECCFRSCDLALLETYCATPAKSE	45
E6R	AYRPSRTLCGGELVDTLQFVCGDRGFYFSRPASRVS RRSRGIVEECCFRSCDLALLETYCATPAKSE	46

[0079]

IGF2 DNA 编码序列		
肽	DNA 序列	SEQ ID NO
成熟 WT IGF2	GCTTACCGCCCCAGTGAGACCCTGTGCGGCGGGG AGCTGGTGGACACCCTCCAGTTCGTCTGTGGGGA CCGCGGCTTCTACTTCAGCAGGCCCGCAAGCCGT GTGAGCCGTCGCAGCCGTGGCATCGTTGAGGAGT GCTGTTTCCGCAGCTGTGACCTGGCCCTCCTGGA GACGTACTGTGCTACCCCCGCCAAGTCCGAG	47
vIGF2 Δ1-4、E6R、Y27L、K65R	TCTAGAACACTGTGCGGAGGGGAGCTTGTAGACA CTCTTCAGTTCGTGTGTGGAGATCGCGGGTTCCTC TTCTCTCGCCCCGCTTCCAGAGTTTCACGGAGGT CTAGGGGTATAGTAGAGGAGTGTTGTTTCAGGTCC TGTGACTTGGCGCTCCTCGAGACCTATTGCGCGA CGCCAGCCAGGTCCGAA	48

[0080] 信号肽

[0081] 在一些实施例中，本文提供的组合物进一步包括信号肽，其改善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质的蛋白质加工，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，信号肽位于(i)信号翻译起始序列的上游位置，(ii)翻译起始序列和治疗性蛋白之间，或(iii)治疗性蛋白的下游位置。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)的结合免疫球蛋白蛋白(BiP)信号肽和Gaussia信号肽及其变体。这些信号肽对信号识别颗粒具有超高的亲和力。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至SEQ ID No:53组成的组的序列至少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至SEQ ID No: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。

[0082]

信号肽序列		
信号肽	氨基酸序列	SEQ ID NO:
天然人类 BiP	MKLSLVAAMLLLLSAARA	49
修饰的 BiP-1	MKLSLVAAMLLLLSLVAAMLLLLSAARA	50
修饰的 BiP-2	MKLSLVAAMLLLLWVALLLLSAARA	51
修饰的 BiP-3	MKLSLVAAMLLLLSLVALLLLSAARA	52
修饰的 BiP-4	MKLSLVAAMLLLLALVALLLLSAARA	53
Gaussia	MGVKVLFALICIAVAEA	54

[0083] Gaussia信号肽来源于来自Gaussia princeps的荧光素酶，并且指导增加的蛋白质合成和与该信号肽融合的治疗性蛋白质的分泌。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。

[0084] 接头

[0085] 在一些实施例中，本文提供的组合物包括靶向肽和治疗性蛋白质之间的接头。在一些实施例中，此类接头维持正确的间距并减轻vIGF2肽与治疗性蛋白质之间的空间冲突。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-丝氨酸残基及其组合。在一些实施例中，接头由5个氨基酸至20个氨基酸、5个氨基酸至15个氨基酸、5个氨基酸至10个氨基酸、8个氨基酸至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接头：

[0086]

接头序列	
序列	SEQ ID NO:
GGGGSGGGG	55
GGGGS	56
GGGSGGGGS	57
GGGGSGGGS	58
GGSGSGSTS	59
GGGGSGGGGS	60

[0087] 在整个说明书中，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，否则任何工程改造的hGAA780I蛋白，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)，或其编码序列，可以类似地被工程改造成表达盒、载体基因组、载体和组合物。

[0088] 合适地，提供了包括本文所述的核酸序列的表达盒。

[0089] 表达盒

[0090] 如本文所用，“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，功能性基因产物与指导其在靶细胞中表达的调节序列(例如hGAA780I融合蛋白编码序列)启动子可操作地连接，并且可以包含其的其他调节序列。必要的调节序列以允许其在靶细胞中转录、翻译和/或表达的方式与hGAA780I融合蛋白编码序列可操作地连接。

[0091] 在某些实施例中，表达盒可以在非翻译区中包含一个或多个miRNA靶序列。miRNA靶序列被设计为被细胞中存在的miRNA特异性识别，在细胞中不期望转基因表达和/或期望降低转基因表达的水平。在某些实施例中，表达盒包含特异性降低背根神经节中hGAA780I融合蛋白表达的miRNA靶序列。在某些实施例中，miRNA靶序列位于3’UTR、5’UTR中和/或位于3’UTR和5’UTR两者中。在某些实施例中，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，其中至少两个串联重复序列包括可以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，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，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，miRNA串联重复序列的长度为200个核苷酸至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靶不改变治疗性转基因在一个或多个靶组织中的表达或功效。

[0092] 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AGTGAATTCTACCAGTGCCATA(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个核苷酸至约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部分互补的序列，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、2、3、4、5、6、7、8、9或10个错配的序列，其中错配可以是不连续的。在某些实施例

中，miR-183靶序列包含具有100％互补性的区，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％互补性的区域包含与miR-183种子序列具有100％互补性的序列。在某些实施例中，miR-183靶序列的其余部分与miR-183具有至少约80％至约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，靶序列包括截短的SEQ ID NO:26，即在SEQ ID NO:26的5'端或3'端中的任一端或两端处缺乏至少1、2、3、4、5、6、7、8、9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。在又其他实施例中，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。在某些实施例中，表达盒或载体基因组包括至少四个、至少五个、至少六个、至少七个或至少八个miR-183靶序列。在某些实施例中，在表达盒或载体基因组中包含至少两个、三个或四个miR-183靶序列导致靶组织如心脏中转基因表达水平的增加。

[0093] 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AGTGTGAGTTCTACCATTGCCAAA(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个核苷酸至约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部分互补的序列，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-183靶序列包括具有至少1、2、3、4、5、6、7、8、9或10个错配的序列，其中错配可以是不连续的。在某些实施例中，miR-182靶序列包含具有100％互补性的区，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靶序列的其余部分与miR-182具有至少约80％至约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，所靶序列包括截短的SEQ ID NO:27，即在SEQ ID NO:27的5'端或3'端中的任一端或两端处缺乏至少1、2、3、4、5、6、7、8、9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。在其他实施例中，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。

[0094] 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由短间隔序列隔开。

[0095] 如本文所用，“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个核苷酸至8个核苷酸、长度为2个核苷酸至7个核苷酸、长度为3个核苷酸至6个核苷酸、长度为四个核苷酸、4个核苷酸至9个核苷酸、3个核苷酸至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。

[0096] 在某些实施例中，串联重复序列含有相同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miRNA靶序列、至少三个不同miRNA

靶序列或至少四个不同miRNA靶序列等。在某些实施例中，串联重复序列可含有相同miRNA靶序列中的两个或三个和不同的第四miRNA靶序列。在某些实施例中，表达盒或载体基因组包含至少一个、至少两个、至少三个或至少四个miR183靶序列和至少一个、至少两个、至少三个或至少四个miR182靶序列的组合。

[0097] 在某些实施例中，表达盒中可存在至少两组不同的串联重复序列。例如，3'UTR可含有紧接在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。在另一实例中，5'UTR可含有一个、两个或更多个miRNA靶序列。在另一实例中，3'可含有串联重复序列，并且5'UTR可含有至少一个miRNA靶序列。

[0098] 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，串联重复序列在转基因的终止密码子的约0个核苷酸至20个核苷酸内开始。在其他实施例中，表达盒含有距转基因的终止密码子至少100个核苷酸至约4000个核苷酸的miRNA串联重复序列。

[0099] 参见于2019年12月20日提交的并且现在已公布为WO 2020/132455的国际专利申请第PCT/US19/67872号，其要求于2018年12月21日提交的美国临时专利申请第62/783,956号的优先权，其据此通过引用并入。还参见2020年5月12日提交的美国临时专利申请第63/023,593号、2020年6月12日提交的美国临时专利申请第63/038,488号、2020年6月24日提交的美国临时专利申请第63/043,562,号、2020年9月16日提交的美国临时专利申请第63/079,299号、2021年2月22日提交的美国临时专利申请第63/152,042号，以及2021年5月12日提交的国际专利申请第PCT/US21/32003号，所有这些专利申请据此通过引用并入。

[0100] 如本文所用，“BiP-vIGF2.hGAAcoV780I.4xmir183”是指含有hGAA780I的工程改造的编码序列的表达盒(例如，如图11所示)，hGAA780I在普遍存在的CAG启动子的控制下具有修饰的BiP-vIGF2信号序列，以及miR183靶序列的四个串联重复序列。如本文提供的实例所示，V780I突变和BiP-vIGF2修饰两者都有助于提高安全性和有效性。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，或与其至少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，或与其至少95％至99％相同的序列。在又另一实施例中，本文提供了一种载体基因组，其中BiP-vIGF2.hGAAcoV780I.4xmir183的侧翼是5'ITR和3'ITR。在某些实施例中，所述载体基因组是SEQ ID NO:30。在又一个实施例中，提供了一种载体基因组，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。

[0101] 如本文所使用，“可操作连接的”序列包含与hGAA780I编码序列相邻的表达控制序列和以反式或以一定距离作用以控制hGAA780I编码序列的表达控制序列。此类调节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。

[0102] 在某些实施例中，调节元件指导受庞贝病影响的多种细胞和组织中的表达，以便允许构建和递送适合用于治疗多种靶细胞的单个表达盒。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节元件(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)中表达的调节元件(例如启动子)。在其他实施例中，调节元件在CNS、骨骼肌和心脏中表达。在其他实施例中，表达盒允许编码的hGAA780I在所有肝脏、骨骼肌、心脏和中枢神经系统细胞中表达。在其他实施例中，可以选择调节元件以用于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,将本文提供的不同表达盒施用于患者,其优先靶向不同组织。

[0103] 调节序列包括启动子。可以选择合适的启动子,包含但不限于将在靶向的细胞中表达hGAAV780I蛋白的启动子。

[0104] 在某些实施例中,选择组成型启动子或诱导型/调节型启动子。组成型启动子的实例是鸡β-肌动蛋白启动子。已经描述了多种单独的或与多种增强子元件组合的鸡β-肌动蛋白启动子(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,其包含鸡β肌动蛋白的启动子、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)。在某些实施例中,可以选择可调控的启动子。参见例如WO2011/126808B2,其通过引用并入本文。

[0105] 在某些实施例中,可以选择组织特异性启动子。组织特异性的启动子的实例对于以下是熟知的:肝脏(白蛋白,Miyatake等人,(1997),《病毒学杂志(J.Virol)》,71:512432;乙型肝炎病毒核心启动子,Sandig等人,(1996),《基因疗法(Gene Ther)》,3:10029;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化酶(NSE)启动子,Andersen等人,(1993),《细胞和分子神经生物学(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;和神经元特异性vgf基因,Piccioli等人,(1995),《神经元(Neuron)》,15:373 84),心肌、骨骼肌、肺和其他组织。在另一个实施例中,合适的启动子可以包含但不限于延伸因子1α(EF1α)启动子(参见,例如,Kim DW等人,使用人延伸因子1α启动子作为通用和有效的表达系统)。《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基因启动子赋予成年大鼠脑中来自腺病毒载体的高度神经元特异性长期转基因表达,这取决于被转导的区域。《基因疗法(Gene Ther)》,2003年2月;10(4):337-47),神经元特异性烯醇化酶(NSE)启动子(参见,例如,Kim J等,富含胆固醇的脂筏参与白细胞介素-6诱导的LNCaP前列腺癌细胞的神经内分泌分化)。《内分泌学(Endocrinology)》,2004年2月;145(2):613-9,电子公布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,《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。

[0106] 在一个实施例中,调节序列进一步包括增强子。在一个实施例中,调节序列包括一个增强子。在另一个实施例中,调节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,增强子可以包括αmic/bik增强子或CMV增强子。这种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分开。

[0107] 在一个实施例中,调节序列进一步包括内含子。在另一个实施例中,内含子是鸡β-肌动蛋白内含子。其他合适的内含子包含本领域已知的可以由人β-球蛋白获得的内含子、和/或可商购的Promega®内含子,以及在WO 2011/126808中描述的内含子。

[0108] 在一个实施例中,调节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poly A。参见,例如,WO 2014/151341。可替代地,在表达盒中可以包含另一种polyA,例如人生长激素(hGH)聚腺苷酸化序列、SV40 polyA或合成polyA。

[0109] 应当理解,本文描述的表达盒中的组合物旨在应用于说明书中描述的其他组合物、方案、方面、实施例和方法。

[0110] 表达盒可以经由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中已知的(参见例如Ramamoorth和Narvekar,《临床与诊断研究杂志(J Clin Diagn Res)》,2015年1月;9(1):GE01-GE06,其通过引用并入本文)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、树枝状聚合物、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)或基于壳聚糖的制剂。

[0111] 在一个实施例中,载体是非病毒质粒,其包括其所述的表达盒,例如,“裸DNA”、“裸质粒DNA”、RNA和mRNA;与各种组合物和纳米颗粒偶联,所述组合物和纳米颗粒包含例如胶束、脂质体、阳离子脂质-核酸组合物、多聚糖蛋白组合物和其他聚合物、脂质和/或胆固醇基-核酸缀合物以及如本文所述的其他构建体。参见,例如X.Su等人,《分子药剂学(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,其所有均通过引用并入本文。

[0112] 在某些实施例中,本文提供了具有编码hGAA780I变体、融合蛋白或截短的蛋白的序列的核酸分子,如本文所述。在一个理想的实施例中,hGAA780I由SEQ ID NO:4的工程改造的序列或与其至少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,使得在位置780I处编码Ile的密码子为ATT或ATC。在某些实施例中,包括SEQ ID NO:4的工程改造的序列或其片段的核酸用于表达融合蛋白或截短的hGAA780I。尽管不太理想,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的氨基酸序列或与其至少95%相同的序列的融合蛋白。在某些实施例中,提供了一种具有SEQ ID NO:7的序列或与其至少95%相同的序列的核酸。在某些实施例中,核酸分子是质粒。

[0113] 载体

[0114] 如本文所用的“载体”是包括核酸序列的生物或化学部分,核酸序列可以被引入到合适的靶细胞中以用于复制或表达核酸序列。载体的实例包含但不限于重组病毒、质粒、脂质体(Lipoplexes)、多聚体(Polymersome)、多复合物(Polyplexes)、树枝状聚合物、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,载体是具有编码功能性基因产物的外源或异源性工程改造的核酸的核酸分子,其然后可以被引入到合适的靶细胞中。此类载体优选地具有一个或多个复制起点以及重组DNA可以插入到其中的一个或多个位点。载体通常具有可以从没有载体的细胞中选择具有载体的细胞的装置,例如,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”。本领域技术人员可以获得产生、生产、表征或定量载体的常规方法。

[0115] 在一些实施例中,本文描述的载体是“复制缺陷型病毒”或“病毒载体”,“复制缺陷型病毒”或“病毒载体”是指其中含有编码功能性hGAA780I融合蛋白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,它们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

可以被工程改造成“无肠的(gutless)”-仅含有编码侧翼为扩增和包装人工基因组所需的信号的核酸序列),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。

[0116] 如本文所用,重组病毒载体是靶向所需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、心脏细胞或骨骼肌。在其他实施例中,病毒载体靶向CNS(例如,脑)、骨骼肌和/或心脏。在其他实施例中,病毒载体靶向所有肝脏、骨骼肌、心脏和中枢神经系统细胞。实例提供了说明性的重组腺相关病毒(rAAV)。然而,其他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。在优选的实施例中,这些重组病毒是复制缺陷型的。

[0117] 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(人类、昆虫或酵母),原核或真核细胞含有通过任何手段(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包被颗粒、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、中枢神经系统细胞、神经元、神经胶质细胞或干细胞。

[0118] 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生产蛋白质用于分离或纯化。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蛋白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。

[0119] 如本文所用,术语“靶细胞”是指需要表达功能基因产物的任何靶细胞。

[0120] 如本文所用,“载体基因组”是指包装在病毒载体内的核酸序列。在一个实例中,“载体基因组”从5’至3’最少含有载体特异性序列、编码功能基因产物(例如hGAAV780I、融合蛋白hGAAV780I或另一种蛋白)的核酸序列(其与指导其在靶细胞中表达的调节控制序列可操作地连接)、载体特异性序列和任选地非翻译区中的miRNA靶序列以及载体特异性序列。载体特异性序列可以是末端重复序列,其将载体基因组特异性包装到病毒载体衣壳或包膜蛋白中。例如,AAV反向末端重复序列被用于包装到AAV和某些其他细小病毒衣壳中。在需要包装到慢病毒载体中的情况下,可以使用慢病毒长末端重复序列。类似地,可以选择其他末端重复序列(例如,逆转录病毒长末端重复序列)等。

[0121] 应当理解,本文描述的载体中的组合物旨在应用于说明书中描述的其他组合物、方案、方面、实施例和方法。

[0122] 腺相关病毒(AAV)

[0123] 在一个方面中,本文提供了一种重组AAV(rAAV),其包括AAV衣壳和包装在其中的载体基因组,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,所选择的AAV衣壳靶向肝脏、肌肉、肾、心脏和/或中枢神经系统细胞类型中的两种或更多种的细胞。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，所选择的AAV衣壳靶向心脏组织。在某些实施例中，被选择靶向心脏组织的AAV衣壳选自AAV 1、6、8和9(参见，例如，Katz等人，《人类基因疗法临床发展(Hum Gene Ther Clin Dev)》，2017年9月1日；28(3):157-164)。在又其他实施例中，所选择的AAV衣壳靶向肾脏的细胞。在一个实施例中，用于靶向肾细胞的衣壳选自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，《生物化学和生物物理研究通讯(Biochem Biophys Res Commun)》，2018年2月26日；497(1):19-24)。在某些实施例中，AAV衣壳是天然的或工程改造的进化枝F衣壳。在某些实施例中，衣壳是AAV9衣壳或AAVhu68衣壳。

[0124] 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rAAV衣壳内形成rAAV载体的核酸序列。此类核酸序列含有位于表达盒侧翼的AAV反向末端重复序列(ITR)。在一个实例中，用于包装到AAV或博卡病毒衣壳中的“载体基因组”从5'至3'最少包含AAV 5'ITR，编码如本文所述的功能性hGAA780I融合蛋白的核酸序列(其可操作地连接到指导其在靶细胞中表达的调节控制序列)，以及AAV 3'ITR。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，可以使用其他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，其被设计为在不希望转基因表达和/或希望转基因表达水平降低的细胞中被miRNA序列特异性识别。在某些实施例中，载体基因组中的miR183靶序列导致转基因在心脏中的表达增加。

[0125] ITR是在载体生产期间负责基因组的复制和包装的遗传元件，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。在优选的实施例中，来自AAV2的ITR序列，或其缺失的版本(ΔITR)，其可以用于方便和加速监管批准。然而，可以选择来自其他AAV来源的ITR。在ITR的来源来自AAV2并且AAV衣壳来自另一个AAV来源的情况下，所得载体可以被称为假型的。典型地，AAV载体基因组包括AAV 5'ITR、hGAA780I编码序列和任何调节序列，以及AAV 3'ITR。然而，这些元件的其他构型可以是合适的。已经描述了被称为ΔITR的5'ITR的缩短版本，其中缺失了D序列和末端解析位点(trs)。在其他实施例中，使用全长的AAV 5'ITR和AAV 3'ITR。在某些实施例中，载体基因组包含130个碱基对的缩短的5'和/或3'AAV2ITR，其中缺失外部“a”元件。在使用内部A元件作为模板的载体DNA扩增期间，缩短的ITR被恢复到145个碱基对的野生型长度。

[0126] 如本文所用的术语“AAV”是指天然存在的腺相关病毒、本领域技术人员和/或根据本文所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，在该衣壳中包装有表达盒，表达盒的侧翼是AAV反向末端重复序列(ITRs)，用于递送至靶细胞。核酸酶抗性重组AAV(rAAV)表明AAV衣壳已经被完全组装，并且在设计以去除生产过程中可能存在的污染核酸的核酸酶孵育步骤期间保护这些包装的载体基因组序列免于降解(消化)。在许多情况下，本文描述的rAAV是DNase抗性的。

[0127] 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至1:1:20，具体取决于所选AAV。可以选择各种AAV作为如上文所识别的AAV病毒载体的衣壳的来源。参见，例如美国公开的专利申请第2007-0036760-A1号、美国公

开的专利申请第2009-0197338-A1号、EP 1310571。还参见WO 2003/042397(AAV7和其他猿猴AAV)、美国专利7790449和美国专利7282199(AAV8)、WO 2005/033321和US 7,906,111(AAV9)、以及WO 2006/110689和WO 2003/042397(rh.10)。这些文献还描述了可以被选择用于产生AAV的其他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,否则本文所述的AAV衣壳、ITR和其他选定的AAV组分可以容易地选自任何AAV,包含但不限于通常鉴定为AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV6.2、AAV7、AAV8、AAV9、AAV8bp、AAVrh10、AAVhu37、AAV7M8和AAVAnc80、AAVrh90(PCT/US20/30273,2020年4月28日提交,其通过引用并入本文)、AAVrh91(PCT/US20/30266,2020年4月28日提交,其通过引用并入本文)、以及AAVrh92、rh93和rh91.93(PCT/US20/30281,2020年4月28日提交,其通过引用并入本文)的AAV以及任何已知或提及的AAV或有待发现的AAV的变体或其变体或混合物。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载体名称中术语"AAV"后的数字或数字和字母的组合来命名。

[0128] ITR或其他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改造。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他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。

[0129] 如本文所用,术语"rAAV"和"人工AAV"可互换使用,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,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,其中一种AAV的衣壳被异源衣壳蛋白取代。在一个实施例中,AAV2/5和AAV2/8是示例性的假型载体。所选基因元件可以通过任何合适的方法递送,方法包含转染、电穿孔、脂质体递送、膜融合技术、高速DNA包被颗粒、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)。

[0130] 在某些实施例中,AAV衣壳选自天然和工程改造的进化枝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,其通过引用以其整体并入本文。然而,在其他实施例中,AAV衣壳选自不同的进化枝,例如进化枝A、B、C、D或E,或者选自这些进化枝之外的AAV来源。

[0131] 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的泊松纠正距离测量结果(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程序实现了改进的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(J.Virol)》,2004年6月;7810:6381-6388,文献鉴定进化枝A、B、C、D、E和F,并且提供了新的AAV的核酸序列,GenBank登录号AY530553至AY530629。还参见WO 2005/033321。

[0132] 如本文所用,“AAV9衣壳”是指具有(a)GenBank登录号:AAS99264(其通过引用并入本文)的氨基酸序列和AAV vp1衣壳蛋白和/或(b)由GenBank登录号AY530579.1:(nt 1..2211)的核苷酸序列编码的氨基酸序列的AAV9。本发明涵盖了来自该编码的序列的一些变异,其可以包含与GenBank登录号:AAS99264和US7906111(也是WO 2005/033321)中的参考氨基酸序列具有约99%同一性的序列(即,与参考序列的变异小于约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衣壳对齐的任何其他AAV衣壳中的相应位置“招募”的替代残基的氨基酸取代;例如,如在US 9,102,949、US 8,927,514、US2015/349911、WO 2016/049230A1、US 9,623,120和US 9,585,971中描述的。然而,在其他实施例中,可以选择与上述序列具有至少约95%同一性的AAV9或AAV9衣壳的其他变体。参见,例如US2015/0079038。因此已经描述了产生衣壳的方法、编码序列以及用于产生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)》,100(10),6081-6086(2003)和US 2013/0045186A1。

[0133] 在某些实施例中,AAVhu68衣壳如在标题为“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WO 2018/160582中所述,其据此通过引用并入。在某些实施例中,AAVhu68衣壳包括:通过从对SEQ ID NO:2的1至736的预测的氨基酸序列进行编码的核酸序列的表达的AAVhu68 vp1蛋白、由SEQ ID NO:1产生的vp1蛋白、或由与SEQ ID NO:1具有至少70%相同的对SEQ ID NO:2的1至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至氨基酸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、由包括SEQ ID NO:1的至少核苷酸412至核苷酸2211的序列产生的vp2蛋白、由与SEQ ID NO:1的至少核苷酸412至核苷酸2211至少70%相同的对SEQ ID NO:2的至少约氨基酸138至氨基酸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/或蛋白通过从对SEQ ID NO:2的至少约氨基酸203至氨基酸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、由包括SEQ ID NO:1的至少核苷酸607至核苷酸2211的序列产生的vp3蛋白、或由与SEQ ID NO:1的至少核苷酸607至核苷酸2211至少70%相同的对SEQ ID NO:2的至少约氨基酸203至氨基酸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。

[0134] 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至氨基酸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至aa 736)进行编码的核酸序列,AAVhu68 vp3氨基酸序列没有vp1独特区(约aa 1至约aa 137)和/或vp2独特区(约aa 1至约aa 202)或与其互补的链,即对应的mRNA或tRNA(SEQ ID NO:1的约nt 607至约nt 2211);或与SEQ ID NO:1至少70%至至少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aa 203至aa 736进行编码的序列。任选地或可替代地,vp1编码和/或vp2编码序列可以与以下中的一个或多个共表达:对SEQ ID NO:2(约aa 138至aa 736)的AAVhu68 vp2氨基酸序列进行编码的核酸序列,AAVhu68 vp2氨基酸序列没有vp1独特区(约aa 1至约aa 137)或与其互补的链,即对应的mRNA或tRNA(SEQ ID NO:1的nt 412至nt 2211);或与SEQ ID NO:1至少70%至至少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至aa 736进行编码的序列。

[0135] 如本文所述,rAAVhu68具有在表达来自AAVhu68核酸的衣壳的生产系统中产生的rAAVhu68衣壳,AAVhu68核酸编码SEQ ID NO:2的vp1氨基酸序列,和任选地额外的核酸序列,例如编码不含vp1和/或vp2独特区的vp3蛋白。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。

[0136] 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(约aa 203至aa 736)的AAVhu68 vp3氨基酸序列进行编码的核酸序列,AAVhu68 vp3氨基酸序列没有vp1独特区(约aa 1至约aa 137)和/或vp2独特区(约aa 1至约aa 202)或与其互补的链,即对应的mRNA(SEQ ID NO:2的约nt 607至约nt 2211)。在某些实施例中,还提供了对SEQ ID NO:2(约aa 138至aa736)的AAVhu68 vp2氨基酸序列进行编码的核酸序列,AAVhu68 vp2氨基酸序没有vp1独特区(约aa 1至约aa 137)或与其互补的链,即对应的mRNA(SEQ ID NO:1的nt 412至nt 2211)。

[0137] 然而,可以选择对SEQ ID NO:2的氨基酸序列进行编码的其他核酸序列用于产生rAAVhu68衣壳。在某些实施例中,核酸序列具有SEQ ID NO:1的核酸序列或与SEQ ID NO:1至少70%至99%相同、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%、至少99%相同的对SEQ ID NO:2进行编码的序列。在某些实施例中,核酸序列具有SEQ ID NO:1的核酸序列或与SEQ ID NO:1的约nt 412至约nt 2211至少70%至99%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%、或至少99%相同的对SEQ ID NO:2的vp2衣壳蛋白(约aa 138至736)进行编码的序列。在某些实施例中,核酸序列具有SEQ ID NO:1的约nt 607至约nt 2211的核酸序列,或与SEQ ID NO:1的nt 412至约nt 2211至少70%至

99％、至少75％、至少80％、至少85％、至少90％、至少95％、至少97％、或至少99％相同的对SEQ ID NO:1的vp1衣壳蛋白(约aa203至aa 736)进行编码的序列。

[0138] 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，在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，或通过如本文所描述的修饰(例如，脱酰胺化的氨基酸)对SEQ ID NO:2的vp1氨基酸序列进行编码的与其至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少97％、至少99％相同的序列生产的。在某些实施例中，vp1氨基酸序列在SEQ ID NO:2中被复制。

[0139] 在某些实施例中，可以选择具有降低的衣壳脱酰胺作用的AAV衣壳。参见，例如，PCT/US19/19804和PCT/US18/19861，两者均于2019年2月27日提交并通过引用以其整体并入。

[0140] 如本文所用，当用于指vp衣壳蛋白时，术语“异质”或其任何语法变异是指由不相同的元件组成的群体，例如，具有带有不同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蛋白的经编码的氨基酸序列。如与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”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。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含其他脱酰胺的氨基酸，其中脱酰胺引起氨基酸变化和其他任选的修饰。

[0141] 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，该组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至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蛋白的单独亚群体，并且vp3是组装的AAV衣壳中的vp蛋白的又一亚群体。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。

[0142] 除非另有说明，否则与在参考氨基酸位置处的预测的氨基酸序列相比，高度脱酰胺化是指在参考的氨基酸位置处被至少45％脱酰胺化、至少50％脱酰胺化、至少60％脱酰胺化、至少65％脱酰胺化、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％、至少97％、至少99％、或高达约100％脱酰胺化(例如，基于SEQ ID NO:2[AAVhu68]的编号，氨基酸57处的天冬酰胺中至少80％可以基于总vp1蛋白脱酰胺，可以基于总vp1、vp2和vp3蛋白脱酰胺)。此类百分比可以使用2D凝胶、质谱技术或其他合适的技术来确定。

[0143] 因此，rAAV包括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，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。另外，其他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。在仍其他实施例中，修饰可以包含在Asp位置处的酰胺化。

[0144] 在某些实施例中,AAVhu衣壳含有具有至少4个至至少约25个脱酰胺化的氨基酸残基位置的vp1、vp2和vp3的亚群体,其中与vp蛋白的编码的氨基酸序列相比,至少1%至10%被脱酰胺。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。

[0145] 在某些实施例中,rAAV具有含有vp1、vp2和vp3蛋白的AAV衣壳,vp1、vp2和vp3蛋白具有包含在实例1提供的并以引用的方式并入本文的表中所阐述的位置处的两个、三个、四个或更多个脱酰胺化的残基的组合的亚群。rAAV中的脱酰胺作用可以使用2D凝胶电泳、和/或质谱、和/或蛋白质建模技术来确定。在线色谱法可以用Acclaim PepMap柱和与带有NanoFlex源的Q Exactive HF(Thermo Fisher Scientific)偶联的Thermo UltiMate 3000RSLC系统(Thermo Fisher Scientific)进行。MS数据是使用用于Q Exactive HF的数据依赖性前20种方法获取的,方法从调查扫描(200m/z至2000m/z)中动态选择最丰富的尚未测序的前体离子。经由较高能量的碰撞解离碎裂进行测序,其中用预测自动增益控制来确定1e5离子的目标值,并且用4m/z的窗口进行前体的分离。在m/z 200下以120,000的分辨率获得了调查扫描。在m/z200下,HCD光谱的分辨率可以被设置为30,000,其中最大离子注入时间为50ms,并且归一化的碰撞能量为30。S-lens RF水平可以被设置为50,以给出由来自消化物的肽占据的m/z区域的最佳传输。可以从碎裂选择中排除具有单个、未分配或六个及更高电荷状态的前体离子。BioPharma Finder 1.0软件(Thermo Fischer Scientific)可以用于分析获得的数据。对于肽图谱,使用单条目蛋白质FASTA数据库进行搜索,其中将氨基甲基化设置为固定的修饰;并且将氧化、脱酰胺和磷酸化设置为可变修饰,10ppm的质量准确度,高蛋白酶特异性,并且MS/MS谱的置信水平为0.8。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-NH_2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,来源于具有多个潜在脱酰胺位点的肽的片段离子可以用于定位或区分多个脱酰胺位点。在这些情况下,观察到的同位素模式中的相对强度可以用于特异性地确定不同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上是独立的。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(Thermo Fisher Scientific))。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(Waters))、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(Hum Gene Ther Methods)》,第28卷,第5期,第255-267页中描述仍其他技术。

[0146] 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他修饰。这种修饰可包括乙酰化残基、异构化、磷酸化或氧化。

[0147] 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。在其他实施例中,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氨基酸序列中发现的天冬酰胺-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/或工程化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他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,如本文所述的突变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衣壳中的结构上分离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,突变中的一个突变不在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衣壳蛋白的情况下,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高达约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。

[0148] 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他脱酰胺化的氨基酸,其中脱酰胺化导致氨基酸改变和其他任选的修饰。在本文中描述了这些和其他修饰的各种组合。

[0149] 在某些实施例中,如本文所述的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至1254页。自身互补AAV在例如美国专利第6,596,535号;第7,125,717号;和第7,456,683号中描述,这些美国专利中的每个美国专利通过引用整体并入本文中。

[0150] 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WO 2003/042397;WO 2005/033321、WO 2006/110689;US 7588772 B2。此类方法涉及培养含有

对AAV衣壳进行编码的核酸序列的宿主细胞;功能性rep基因;侧翼是AAV反向末端重复序列(ITR)的如本文所述的表达盒;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。本文还提供了含有对AAV衣壳进行编码的核酸序列的宿主细胞;功能性rep基因;如所述的载体基因组;和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。这些方法在WO2017160360 A2中更详细地描述,其通过引用并入本文。

[0151] 可以利用本领域技术人员可获得的其他生产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经由酵母生产。参见,例如,Robert M.Kotin,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(Hum Mol Genet)》,2011年4月15日;20(R1):R2-R6.2011年4月29日在线发布,doi:10.1093/hmg/ddr141;Aucoin MG等人,使用三重感染在昆虫细胞中生产腺相关病毒载体:杆状病毒浓度比率的优化(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、在酵母中生产重组腺相关病毒载体(Production of Recombinant Adeno-associated viral vectors in yeast),向佛罗里达大学研究生院(Graduate School of the University of Florida)提交的论文,2012;Kondratov O等人,“在人类相对于昆虫细胞中制造的重组腺相关病毒载体的直接面对面评估(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,[在印刷之前电子出版];Mietzsch M等人,OneBac 2.0:“用于生产具有最小外源DNA包封的AAV1、AAV2和AAV8载体的Sf9细胞系(Sf9 Cell Lines for Production of AAV1,AAV2,and AAV8 Vectors with Minimal Encapsidation 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,“新型重组腺相关病毒复制型基因组的生产和鉴定: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.印刷,2013;Galibert L等人,“在昆虫细胞中大规模生产腺相关病毒载体以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)”,《无脊椎动物病理杂志(J Invertebr Pathol)》,2011年7月;107补刊:S80-93.doi:10.1016/j.jip.2011.05.008;和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(Hum Mol Genet)》,2011年4月15日;20(R1):R2-6.doi:10.1093/hmg/ddr141.电子公布2011年4月29日。

[0152] 在高盐浓度下进行两步亲和色谱法纯化,然后使用阴离子交换树脂色谱法来纯化载体药物产物并去除空衣壳。这些方法在题为“AAV9的可扩展纯化方法(Scalable

Purification Method for AAV9)”的WO 2017/160360中更详细地描述,其通过引用并入本文。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法,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,并且经受盐梯度,同时监测洗脱液在约260nm和约280nm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至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,对于亲和色谱法步骤,可以将经过渗滤的产物应用于有效捕获AAV2/9血清型的Capture SelectTM Poros-AAV2/9亲和树脂(生命科技公司(Life Technologies))上。在这些离子条件下,显著比例的残留的细胞DNA和蛋白质流过柱,而AAV颗粒则被有效捕获。

[0153] 本领域技术人员可获得用于表征或定量rAAV的常规方法。为了计算空颗粒和完整颗粒的含量,将所选样本(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#＝颗粒#)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)用于计算测试品峰的带状体积中的颗粒的数量。然后将每20μL装载的颗粒数(pt)乘以50,以得出颗粒数(pt)/mL。Pt/mL除以GC/mL得到颗粒与基因组拷贝数的比率(pt/GC)。Pt/mL-GC/mL给出空pt/mL。空pt/mL除以pt/mL并且×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(Gene Therapy)》(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,该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,例如在缓冲液中含有3%-8%三乙酸盐的梯度凝胶),然后运行凝胶直到分离出样本材料,并且将凝胶印迹到尼龙或硝酸纤维素膜(优选地是尼龙)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AAV-2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,该次级抗体与初级抗体结合并含有用于检测与初级抗体结合的工具,更优选地含有与其共价结合的检测分子的抗IgG抗体,最优选地共价连接至辣根过氧化物酶的绵羊抗小鼠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本并在含有还原剂(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)或其他合适的染色方法(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。将样本稀释并用DNase I(或另一种合适的核酸酶)消化以除去外源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TaqMan™荧光探针进一步稀释和扩增样本。在Applied Biosystems Prism 7700序列检测系统上测量每个样本达到定义荧光水平所需的循环次数(阈值循环,Ct)。使用含有与AAV载体中含有的序列相同序列的质粒DNA以在Q-PCR反应中生成标准曲线。从样本获得的周期阈值(Ct)的值用于通过对其进行归一化为质粒标准曲线的Ct值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。

[0154] 一方面，使用了经过优化的q-PCR方法，该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，将样本用蛋白酶K缓冲液稀释并用蛋白酶K处理，然后进行热失活。合适地，以等于样本大小的量用蛋白酶K缓冲液稀释样本。蛋白酶K缓冲液可以被浓缩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至约1mg/mL之间变化。处理步骤通常在约55℃下进行约15分钟，但可以在较低的温度(例如，约37℃至约50℃)下进行较长的时间段(例如，约20分钟至约30分钟)，或在较高的温度(例如，高达约60℃)下进行较短的时间段(例如，约5分钟至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℃至约90℃)并且时间延长(例如，约20分钟至约30分钟)。然后将样本稀释(例如，1000倍)，并如标准测定中所描述的进行TaqMan分析。

[0155] 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(Hum Gene Ther Methods)》，2014年4月;25(2):115-25.doi:10.1089/hgtb.2013.131.电子公布2014年2月14日。

[0156] 用于确定衣壳蛋白的vp1、vp2和vp3之间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(J.Virol)》，2013年12月;87(24):13150-13160;Buller RM、Rose JA.1978.“在KB细胞中腺病毒相关病毒诱导的多肽的特性(Characterization of adenovirus-associated virus-induced polypeptides in KB cells)”，《病毒学杂志(J.Virol)》，25:331-338;和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(J.Virol)》，8:766-770。

[0157] 应当理解，本文所述的rAAV中的组合物旨在应用于本说明书中所述的其他组合物、方案、方面、实施例和方法。

[0158] 药物组合物

[0159] 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、冻干的或冷冻的组合物或另一种合适的制剂。在某些实施例中，组合物包括hGAA780I融合蛋白或表达盒和形成悬浮液的生理相容性液体(例如，溶液、稀释剂、载剂)。此类液体优选地是水基的，并且可以含有以下的一种或多种:缓冲剂、表面活性剂、pH调节剂、防腐剂或其他合适的赋形剂。合适的组分将在下面更详细地讨论。药物组合物包括水性悬浮液体和任何选择的赋形剂，以及hGAA780I融合蛋白或表达盒。

[0160] 在某些实施例中，药物组合物包括有包括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、壳聚糖基制剂和本领域已知的其他制剂，以及例如如上引用的Ramamoorth和Narvekar所述的制剂。在其他实施例中，药物组合物是包括表达盒的悬浮液，该表达盒包括在病毒载体系统中工程改造的转基因。在某些实施例中，药物组合物包括非复制型病毒载体。合适的病毒载体可以包含任何合适的递送载体，如例如重组腺病毒、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。

[0161] 在一个实施例中，药物组合物包括hGAA780I融合蛋白或包括hGAA780If融合蛋白的编码序列的表达盒以及适于经由侧脑室(ICV)、鞘内(IT)、脑池内或静脉内(IV)注射递送的制剂缓冲液。在一个实施例中，表达盒是包装在重组病毒载体(即携带融合蛋白的rAAV.hGAA780I)中的载体基因组的一部分。

[0162] 在一个实施例中，药物组合物包括hGAA780I融合蛋白或其功能片段，用于作为酶替代疗法(ERT)递送至受试者。此类药物组合物通常静脉内施用，然而在一些情况下皮内、肌内或口服施用也是可能的。组合物可以被施用以用于预防性治疗患有庞贝病或处于患庞贝病风险的个体。对于治疗应用，以足以降低累积代谢物的浓度和/或防止或阻止代谢物的进一步累积的量将药物组合物施用于患有已确定疾病的患者。对于处于溶酶体酶缺乏症风险中的个体，药物组合物以足以防止或抑制代谢物积累的量被预防性施用。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程度以及受试者的年龄、一般状况和性别而变化。剂量可以由医生确定，并且可以根据需要进行调整，以适应所观察的治疗的效果。在一个方面中，本文提供了一种用于ERT的药物组合物，其被配制成含有单位剂量的hGAA780I融合蛋白或其功能片段。

[0163] 在一个实施例中，组合物包含适合于递送到受试者的最终制剂，该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，制剂中存在一种或多种表面活性剂。在另一个实施例中，组合物可以作为浓缩物运输，该浓缩物被稀释用于施用于受试者。在其他实施例中，可以在施用时将组合物冻干并重构。

[0164] 在一个实施例中，如本文提供的组合物包括溶解在水性悬浮液体中的表面活性剂、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，Eliott的制剂缓冲液；或Harvard仪器灌注液(一种具有最终离子浓度的人工CSF(以mM计):Na 150;K 3.0;Ca 1.4;Mg 0.8;P 1.0;Cl 155的人工脑脊液)。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、调节至离子强度相当于约100mM氯化钠(NaCl)至约250mM氯化钠的生理相容性盐或盐的混合物，或调节至等效离子浓度的生理相容性盐。

[0165] 合适地，将制剂调整到生理上可接受的pH，例如，在pH 6至8、或pH 6.5至7.5、pH 7.0至7.7或pH 7.2至7.8的范围内。由于脑脊液的pH为约7.28至约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6.8至约7.2的pH。然而，可以选择最宽范围和这些子范围内的其他pH用于其他递送途径。

[0166] 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[巴斯夫(BASF)]，也被称为泊洛沙姆188，其具有中性pH，平均分子量为8400。可以选择其他表面活性剂和其他泊洛沙姆，即非离子型三嵌段共聚物，该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)、乙醇和聚乙二醇组成。在一个实施例中，制剂含有泊洛沙姆。这些共聚物通常以字母“P”(泊洛沙姆)后跟三个数字命名：前两个数字x 100给出

聚氧丙烯核的近似分子量，并且最后一个数字x 10给出聚氧乙烯含量的百分比。在一个实施例中，选择泊洛沙姆188。表面活性剂可以以悬浮液的高达约0.0005％至约0.001％的量存在。

[0167] 在一个实例中，制剂可以含有例如缓冲盐水溶液，该缓冲盐水溶液包括水中的氯化钠、碳酸氢钠、葡萄糖、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，渗透压在与脑脊液相容的范围内(例如，约275至约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他任选的赋形剂组合。参见，例如，ElliottsB®溶液[Lukare Medical]。

[0168] 在其他实施例中，制剂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。

[0169] 另外提供了一种药物组合物，其包括药学上可接受的载剂和包括如本文所述的核酸序列的载体。如本文所使用的，“载剂”包含任何和所有溶剂、分散介质、媒剂、涂层、稀释剂、抗细菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲剂、载剂溶液、悬浮液、胶质物等。将此类培养基和药剂用于药物活性物质在本领域中是众所周知的。还可以将补充性活性成分并入到组合物中。递送媒剂(如脂质体、纳米胶囊、微颗粒、微球、脂质颗粒、囊泡等)可以用于将本文所述的组合物引入到合适的宿主细胞中。具体地，rAAV载体可以被配制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，在药物组合物中包含治疗有效量的载体。载剂的选择不是本发明的限制。其他常规药学上可接受的载剂，如防腐剂或化学稳定剂。适合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、对羟基苯甲酸酯、乙基香兰素、甘油、苯酚以及对氯苯酚。适合的化学稳定剂包含明胶和白蛋白。

[0170] 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。

[0171] 如本文所用，术语“剂量”或“量”可以指在治疗过程中递送至受试者的总剂量或量，或以单个单位(或多个单位或分次剂量)施用递送的剂量或量。

[0172] 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此需要的受试者。在一个实施例中，药物组合物被配制成经由侧脑室(ICV)、鞘内(IT)或脑池内注射来递送。在一个实施例中，本文所述的组合物被设计成用于通过静脉内注射递送至有此需要的受试者。或者，可以选择其他施用途径(例如，口服、吸入、鼻内、气管内、动脉内、眼内、肌内和其他胃肠外途径)。

[0173] 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，“在将AAV9递送到小脑延髓池后，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-方法与临床发展(Mol Ther Methods Clin Dev)》,2014;1:14051.于2014年12月10日在线发布,doi:10.1038/mtm.2014.51。

[0174] 如本文所使用的,术语“脑池内递送”或“脑池内施用”是指药物直接进入到脑室的脑脊液中或小脑延髓池(cisterna magna cerebellomedularis)中,更具体地是经由枕骨下穿刺或通过直接注射到小脑延髓池(cisterna magna)中或经由永久定位的管的施用途径。

[0175] 在一个方面中,本文提供了一种药物组合物,其在制剂缓冲液中包括如本文所述的载体。在一些实施例中,可以以剂量单位配制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至约1.0×10^{16}GC的范围内(以治疗平均体重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至1.0×10^{14}GC。在一个实施例中,将组合物配制成每剂量含有至少$1x10^{9}$、$2x10^{9}$、$3x10^{9}$、$4x10^{9}$、$5x10^{9}$、$6x10^{9}$、$7x10^{9}$、$8x10^{9}$或$9x10^{9}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{10}$、$2x10^{10}$、$3x10^{10}$、$4x10^{10}$、$5x10^{10}$、$6x10^{10}$、$7x10^{10}$、$8x10^{10}$或$9x10^{10}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{11}$、$2x10^{11}$、$3x10^{11}$、$4x10^{11}$、$5x10^{11}$、$6x10^{11}$、$7x10^{11}$、$8x10^{11}$或$9x10^{11}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{12}$、$2x10^{12}$、$3x10^{12}$、$4x10^{12}$、$5x10^{12}$、$6x10^{12}$、$7x10^{12}$、$8x10^{12}$或$9x10^{12}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{13}$、$2x10^{13}$、$3x10^{13}$、$4x10^{13}$、$5x10^{13}$、$6x10^{13}$、$7x10^{13}$、$8x10^{13}$或$9x10^{13}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{14}$、$2x10^{14}$、$3x10^{14}$、$4x10^{14}$、$5x10^{14}$、$6x10^{14}$、$7x10^{14}$、$8x10^{14}$或$9x10^{14}$GC,包含该范围内的所有整数或分数量。在另一个实施例中,将组合物配制成每剂量含有至少$1x10^{15}$、$2x10^{15}$、$3x10^{15}$、$4x10^{15}$、$5x10^{15}$、$6x10^{15}$、$7x10^{15}$、$8x10^{15}$或$9x10^{15}$GC,包含该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$1x10^{10}$至约$1x10^{12}$GC,包含该范围内的所有整数或分数量。

[0176] 在一个实施例中,提供了一种在制剂缓冲液中包括如本文所述的rAAV的药物组合物。在一个实施例中,rAAV以约$1x10^{9}$基因组拷贝(GC)/mL至约$1x10^{14}$GC/mL配制。在一个实施例中,rAAV以约$3x10^{9}$GC/mL至约$3x10^{13}$GC/mL配制。在一个实施例中,rAAV以约$1x10^{9}$GC/mL至约$1x10^{13}$GC/mL配制。在一个实施例中,rAAV以至少约$1x10^{11}$GC/mL配制。

[0177] 在一个实施例中,包括如本文所述rAAV的药物组合物可以约1×10^{9}GC/克脑质量至约$1x10^{14}$GC/克脑质量的剂量施用。

[0178] 应当理解,本文所述的药物组合物中的组合物旨在应用于说明书中描述的其他组合物、方案、方面、实施例和方法。

[0179] 治疗方法

[0180] 提供了一种用于治疗患有庞贝病的患者的治疗方案,其包括如本文所述的表达盒、rAAV和/或hGAA780I融合蛋白,任选地与免疫调节剂组合。在某些实施例中,患者患有晚发性庞贝病。在其他实施例中,患者患有儿童期发病的庞贝病。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。此外或可替代地,联合疗法可以包含支气管扩张剂、乙酰胆碱酯酶抑制剂、呼吸肌肉力量训练(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,患者接受rAAV的单次施用。在某些

实施例中，患者接受单次施用包括如本文所述的表达盒和/或rAAV的组合物。在某些实施例中，这种包括有效量的表达盒的组合物的单次施用涉及至少一种共治疗剂。在某些实施例中，如本文所述基本上同时经由两种不同的途径向患者施用表达盒、rAAV和/或hGAA780I融合蛋白。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，组合物是悬浮液，其经侧脑室、鞘内、脑池内或静脉内被递送至受试者。在某些实施例中，向患有α-葡糖苷酶缺乏症的患者施用如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，作为治疗的结果，在心脏、CNS(脑)和/或骨骼肌中的一个或多个中存在减少的糖原贮积和/或自噬积聚。

[0181] 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。

[0182] 这些组合物可以与其他疗法联合使用，包含例如免疫疗法、酶替代疗法(例如，Lumizyme，由赛诺菲公司(Sanofi Corporation)以Genzyme销售，并且在美国境外作为Myozyme销售)。庞贝病的额外治疗是有症状的和支持性的。例如，可能需要呼吸支持；物理疗法可能有助于加强呼吸肌；一些患者可能需要在夜间和/或白天期间通过机械通气(即bipap或容积通气)的呼吸辅助。此外，在呼吸道感染期间，可能需要额外的支持。对于一些患者，可能会推荐使用包含支架(brace)的矫形装置。对于某些矫形症状，如挛缩或脊柱畸形，可能需要手术。一些婴儿可能需要插入饲管，该饲管穿过鼻子，沿着食道进入胃(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饲管直接插入到胃中。一些患有晚发性庞贝病的个体可能需要清淡饮食，但很少有人需要饲管。

[0183] 尽管ERT显著改善了患有典型婴儿庞贝病的患者的生存率，但其无法完全逆转骨骼肌病理，这部分地是由于抑制酶到达溶酶体的自噬积聚。我们已经表明，本文提供的组合物可有效治疗和逆转肌肉病理。例如，在治疗时在患有预先存在的病理的老年Pompe小鼠中，自噬体累积完全消退。这些发现还表明，用本文提供的载体治疗可以显著增加骨骼肌中大肌纤维的百分比，并降低小肌纤维的百分比。因此，典型的治疗抗性病理(如肌纤维大小和自噬积聚)对治疗有应答。

[0184] 在某些实施例中，本文提供了用于在诊断为庞贝病或怀疑患有庞贝病的患者中减少异常肌肉病理的进展和/或逆转异常肌肉病理的方法。在某些实施例中，患者为症状前的。在其他实施例中，患者是症状后的，包含具有更晚期疾病阶段的老年患者，并且治疗包含改善(或逆转)庞贝病的症状。在某些实施例中，异常肌肉病理的特征在于以下中的一种或多种：i)具有中央核的肌肉细胞百分比升高；ii)肌纤维萎缩，iii)肌细胞纤维中的红细胞大小不均，iv)自噬积聚，v)空泡化，和vi)无力。在某些实施例中，该方法改善了患者的呼吸和/或运动。

[0185] 如本文所述，术语“增加”(例如，在用hGAA780I融合蛋白处理后hGAA水平的增加，如在组织、血液等中测量的)或“减少”、“降低”、“改进”、“改善”、“延迟”或其任何语法变化，或指示变化的任何类似术语，是指与相应的参考(例如，未治疗的对照或没有庞贝的正常条件下的受试者)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，除非另有说明。

[0186] 如本文可互换使用的，“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

物的受试者是人类受试者。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。

[0187] 在一个实施例中，悬浮液具有约7.28至约7.32的pH。

[0188] 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1mL至150mL的体积，其中对于成人，选择更大的体积。典型地，对于新生婴儿，合适的体积是约0.5mL至约10mL，对于较大的婴儿，可以选择约0.5mL至约15mL。对于幼儿，可以选择约0.5mL至约20mL的体积。对于儿童，可以选择高达约30mL的体积。对于青春期前和青少年，可以选择高达约50mL的体积。在又其他实施例中，患者可以接受以选择的约5mL至约15mL，或约7.5mL至约10mL的体积的鞘内施用。可以确定其他合适的体积和剂量。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。

[0189] 在一个实施例中，包括如本文所述的rAAV的组合物可以以约1×10^{9}GC/克脑质量至约1×10^{14}GC/克脑质量的剂量施用。在某些实施例中，rAAV以约1×10^{9}GC/kg体重至约1×10^{13}GC/kg体重的剂量被全身性共同施用。在某些实施例中，rAAV以约1×10^{11}GC/kg体重至约$5x10^{13}$GC/kg体重的剂量被施用或全身性共同施用。

[0190] 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。如本文所用，“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，方法包括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，并联合施用包括本文提供的hGAA780I融合蛋白酶的组合物。

[0191] 在一个实施例中，表达盒在载体基因组中，以约1×10^{9}GC/克脑质量至约1×10^{13}基因组拷贝(GC)/克(g)脑质量的量递送，包含范围和端点内的所有整数或分数量。在另一个实施例中，剂量为$1x10^{10}$GC/克脑质量至约$1x10^{13}$GC/克脑质量。在具体的实施例中，施用于患者的载体的剂量为至少约$1.0x10^{9}$GC/g、约$1.5x10^{9}$GC/g、约$2.0x10^{9}$GC/g、约$2.5x10^{9}$GC/g、约$3.0x10^{9}$GC/g、约$3.5x10^{9}$GC/g、约$4.0x10^{9}$GC/g、约$4.5x10^{9}$GC/g、约$5.0x10^{9}$GC/g、约$5.5x10^{9}$GC/g、约$6.0x10^{9}$GC/g、约$6.5x10^{9}$GC/g、约$7.0x10^{9}$GC/g、约$7.5x10^{9}$GC/g、约$8.0x10^{9}$GC/g、约$8.5x10^{9}$GC/g、约$9.0x10^{9}$GC/g、约$9.5x10^{9}$GC/g、约$1.0x10^{10}$GC/g、约$1.5x10^{10}$GC/g、约$2.0x10^{10}$GC/g、约$2.5x10^{10}$GC/g、约$3.0x10^{10}$GC/g、约$3.5x10^{10}$GC/g、约$4.0x10^{10}$GC/g、约$4.5x10^{10}$GC/g、约$5.0x10^{10}$GC/g、约$5.5x10^{10}$GC/g、约$6.0x10^{10}$GC/g、约$6.5x10^{10}$GC/g、约$7.0x10^{10}$GC/g、约$7.5x10^{10}$GC/g、约$8.0x10^{10}$GC/g、约$8.5x10^{10}$GC/g、约$9.0x10^{10}$GC/g、约$9.5x10^{10}$GC/g、约$1.0x10^{11}$GC/g、约$1.5x10^{11}$GC/g、约$2.0x10^{11}$GC/g、约$2.5x10^{11}$GC/g、约$3.0x10^{11}$GC/g、约$3.5x10^{11}$GC/g、约$4.0x10^{11}$GC/g、约$4.5x10_{11}$GC/g、约$5.0x10^{11}$GC/g、约$5.5x10^{11}$GC/g、约$6.0x10^{11}$GC/g、约$6.5x10^{11}$GC/g、约$7.0x10^{11}$GC/g、约$7.5x10^{11}$GC/g、约$8.0x10^{11}$GC/g、约$8.5x10^{11}$GC/g、约$9.0x10^{11}$GC/g、约$9.5x10^{11}$GC/g、约$1.0x10^{12}$GC/g、约$1.5x10^{12}$GC/g、约$2.0x10^{12}$GC/g、约$2.5x10^{12}$GC/g、约$3.0x10^{12}$GC/g、约$3.5x10^{12}$GC/g、约$4.0x10^{12}$GC/g、约$4.5x10^{12}$GC/g、约$5.0x10^{12}$GC/g、约$5.5x10^{12}$GC/g、约$6.0x10^{12}$GC/g、约$6.5x10^{12}$GC/g、约$7.0x10^{12}$GC/g、约$7.5x10^{12}$GC/g、约$8.0x10^{12}$GC/g、约$8.5x10^{12}$GC/g、约$9.0x10^{12}$GC/g、约$9.5x10^{12}$GC/g、约$1.0x10^{13}$GC/g、约$1.5x10^{13}$GC/g、约$2.0x10^{13}$GC/g、约$2.5x10^{13}$GC/g、约$3.0x10^{13}$GC/g、约$3.5x10^{13}$GC/g、约$4.0x10^{13}$GC/g、约$4.5x10^{13}$GC/g、约$5.0x10^{13}$GC/g、约$5.5x10^{13}$GC/g、约$6.0x10^{13}$GC/g、约$6.5x10^{13}$GC/g、约$7.0x10^{13}$GC/g、约$7.5x10^{13}$GC/g、约$8.0x10^{13}$GC/g、约$8.5x10^{13}$GC/g、约$9.0x10^{13}$GC/g、约

$9.5x10^{13}$GC/g或约$1.0x10^{14}$GC/g脑质量。

[0192] 在某些实施例中，包括如本文所述的rAAV的组合物以约$1x10^{11}$GC/kg体重至约$5x10^{13}$GC/kg体重的剂量被全身施用。在某些实施例中，rAAV以约$1x10^{12}$GC至约$5x10^{13}$GC的剂量经由ICM施用。在其他实施例中，rAAV经由静脉内和ICM途径被共同施用，其中向患者施用约$1x10^{11}$GC/kg体重至约$5x10^{13}$GC/kg体重(IV)的剂量和约$1x10^{12}$GC至约$5x10^{13}$GC(ICM)的剂量。

[0193] 在一个实施例中，治疗的方法包括递送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(包含但不限于如本文提供的表达盒或rAAV)联合递送。在某些实施例中，该方法包括向受试者施用多于一种的ERT(例如，包括hGAA780I融合蛋白与另一种治疗性蛋白如Lumizyme组合的组合物)。包括本文所述的hGAA780I融合蛋白的组合物可以每1、2、3、4、5、6、7、8、9、10或更多天施用于受试者。可以通过静脉输注向门诊患者施用，按处方每周、每月或每两个月施用一次。由治疗临床医师选择化合物的适当治疗有效剂量，并且包含约1μg/kg至约500mg/kg、约10mg/kg至约100mg/kg、约20mg/kg至约100mg/kg和约20mg/kg至约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。

[0194] 在某些实施例中，该方法包括以对患者每周10mg/kg患者体重或更多的剂量向受试者施用hGAA780I融合蛋白。剂量通常大于10mg/kg/周。剂量方案可以在10mg/kg/周至至少1000mg/kg/周的范围内。典型的剂量方案是10mg/kg/周、15mg/kg/周、20mg/kg/周、25mg/kg/周、30mg/kg/周、35mg/kg/周、40mg/kg/周、45mg/kg/周、60mg/kg/周、80mg/kg/周和120mg/kg/周。在优选的方案中，每周施用10mg/kg、15mg/kg、20mg/kg、30mg/kg或40mg/kg一次、两次或三次。治疗通常持续至少4周，有时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(例如，低于20％)正常人的值时，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次至3次(例如，持续1、2或3周)。维持剂量的实例是至少约5mg/kg至至少约10mg/kg患者体重/周，或更多，例如20mg/kg/周、30mg/kg/周、40mg/kg/周。在某些实施例中，在给药期间以增加的速率施用剂量。这可以通过增加静脉输注流速或通过使用以恒定速率施用的hGAA780I融合蛋白的增加的浓度梯度来实现。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，静脉输注发生在几小时的时间段内(例如，1小时至10小时，并且优选地2小时至8小时，更优选地3小时至6小时)，并且输注的速率在施用的时间段期间以一定的间隔增加。

[0195] 在一个实施例中，方法进一步包括使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含但不限于糖皮质激素、类固醇、抗代谢物、T细胞抑制剂、大环内酯(例如雷帕霉素或雷帕霉素类似物)和细胞抑制剂，细胞抑制剂包含烷化剂、抗代谢物、细胞毒性抗生素、抗体或对亲免疫蛋白有活性的药剂。免疫抑制剂可以包含氮芥、亚硝脲、铂化合物、甲氨蝶呤、硫唑嘌呤、巯基嘌呤、氟尿嘧啶、更生霉素、蒽环类抗生素、丝裂霉素C、博来霉素、光辉霉素、IL-2受体导向抗体或CD3导向抗体、抗IL-2抗体、环孢素、他克莫司、西罗莫司、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，免疫抑制疗法可以在基因疗法施用前0、1、2、7或更多天开始。这些药物中的一种或多种可以在基因疗

法施用后以相同剂量或调整的剂量继续施用。根据需要，这种疗法可以持续约1周(7天)、约60天或更长时间。

[0196] 在一个实施例中，将包括如本文所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，表达盒经由rAAV递送。应当理解，本文描述的组合物和方法旨在应用于说明书中描述的其他组合物、方案、方面、实施例和方法。

[0197] 本文提供的组合物和方法可以用于治疗婴儿发病的庞贝病或晚发性庞贝病和/或与其相关的症状。在某些实施例中，疗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。婴儿发病的庞贝病的症状包含但不限于张力减退、呼吸/呼吸问题、肝肿大、肥厚性心肌病以及心脏、肌肉和CNS(尤其是运动神经元)中的糖原贮积。晚发性庞贝病的症状包含但不限于近端肌无力、呼吸/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/或诊断来确定。在某些实施例中，提供了一种用于治疗被诊断为婴儿发病的庞贝病或晚发性庞贝病的患者的方法，该方法包含施用本文所述的rAAV，用于经由IV和ICM途径的组合递送hGAA780I融合蛋白。在一些实施例中，对被鉴定为患有晚发性庞贝病的患者施用仅包含rAAV的全身递送(例如，仅IV)的治疗。如本文所述，包括rAAV的组合物的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诊断为庞贝病的受试者的方法，其包含ICM递送本文所述的rAAV与ERT的组合。在某些实施例中，被鉴定为患有婴儿发病的庞贝病的受试者经由ICM注射被施用本文所述的rAAV，并且还接受ERT以用于治疗外周疾病的方面。

[0198] 如本文所述的“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：编码感兴趣的蛋白的核酸、寡核苷酸、核酸类似物，例如肽-核酸(PNA)、假补体PNA(pc-PNA)、锁定核酸(LNA)等。此类核酸序列包含，例如但不限于，编码蛋白质的核酸序列，例如充当转录阻遏物的蛋白质，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。

[0199] 用于“反向翻译”蛋白质、肽或多肽的方法是本领域的技术人员已知的。一旦已知蛋白质的序列，就存在基于网络的和商业上可获得的计算机程序，以及基于服务的公司，其将氨基酸序列反向翻译成核酸编码序列。参见，例如，通过EMBOSS(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)的反向翻译。在一个实施例中，RNA和/或cDNA编码序列被设计成用于在人类细胞中的最佳表达。

[0200] 在核酸序列的情况下，术语“百分比(%)一致性”、“序列一致性”、“百分比序列一致性”或“百分比一致”是指当比对对应关系时相同的两个序列中的残基。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个核苷酸至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个核苷酸至24个核苷酸、至少约28个核苷酸至32个核苷酸、至少约36个核苷酸或更多个核苷酸。

[0201] 可以容易地确定蛋白质、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的全长上的氨基酸序列的百分比一致性。合适的氨基酸片段可以是至少约8个氨基酸长度，并且可以是高达约700个氨基酸。通常，当提及两个不同序列之间的“一致性”、“同源性”或“类似性”时，参考“比对”序列来确定“一致性”、“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纠正的多个核酸序列或蛋白质(氨基酸)序列。

[0202] 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。序列比对程序可用于氨基酸序列，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，尽管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，该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.Thomson等人，《核酸研究(Nucl.Acids.Res)》，27(13):2682-2690(1999)。

[0203] 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 Sequence Assembly”、“BLAST”、“MAP”和“MEME”，该程序可通过因特网上的Web服务器进行访问。此类程序的其他来源是本领域技术人员已知的。可替代地，也可以使用载体NTI工具。在本领域还存在许多已知的算法(其可以用于测量核苷酸序列同一性)，包含在上述程序中含有的算法。作为另一实例，可以使用Fasta™(GCG版本6.1中的程序)比较多核苷酸序列。Fasta™提供了查询序列和搜索序列之间最佳重叠区域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Fasta™所确定的，该程序通过引用并入本文中。

[0204] 如本文所使用的，术语“调节序列”或“表达控制序列”是指核酸序列，如起始子序列、增强子序列或启动子序列，该核酸序列诱导、抑制或以其他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。

[0205] 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存在的位置。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该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。

[0206] 用于描述核酸序列或蛋白质的术语“异源的”是指该核酸或蛋白质源自与其表达的宿主细胞或受试者不同的生物体或同一生物体的不同物种。术语“异源的”当用于指质粒、表达盒或载体中的蛋白质或核酸时，指示该蛋白质或核酸与另一序列或亚序列一起存在，该另一序列或亚序列与所讨论的蛋白质或核酸在自然界中彼此没有相同的关系。

[0207] “包括”是包含其他组分或方法步骤的术语。当使用“包括”时，应当理解，相关实施例包括使用“由…组成”术语，其不包括其他组分或方法步骤以及“基本上由…组成”术语，其不包括实质上改变实施例或发明的性质的任何组分或方法步骤的描述。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”的语言来描述相关的实施例。

[0208] 如本文所用，术语“e”后跟数字(nn)值是指指数，并且该术语可与“x 10nn”互换使用。例如，3e13相当于$3x10^{13}$。

[0209] 应当注意，术语“一个(种)(a)”或“一个(种)(an)”是指一个(种)或多个(种)，例如“一个载体”应理解为表示一个或多个载体。因此，术语“一个(a)”(或“an”)、“一个或多个”和“至少一个”在本文中可互换使用。

[0210] 如本文所使用的,除非另有说明,否则术语“约”意指与给定参考的正或负10%的变化性。

[0211] 实例

[0212] 现在参考以下实例来描述本发明。提供以下实例仅仅是为了说明的目的,并且本发明决不应该被解释为局限于这些实例,而是应该被解释为涵盖由于本文提供的教导而变得明显的任何和所有变化。

[0213] 实例1:材料和方法

[0214] 载体产生

[0215] 在780处具有Val的参考GAA序列和具有V780I突变的序列被反向翻译,并且核苷酸序列被工程改造以产生用于AAV生产的顺式质粒,其中表达盒在CAG启动子下。此外,将天然hGAA的cDNA序列(参考序列)克隆到相同的AAV-顺式主链中,以用于与非工程改造的序列进行比较。如前所述生产和滴定AAVhu68载体。(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,并且获得培养物上清液,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滴定纯化的载体(Lock等人,(2014)。《人类基因疗法方法(Hum Gene Ther Methods)》,25(2):115-125)。

[0216] 动物

[0217] 小鼠

[0218] Pompe小鼠(Gaa敲除(-/-),C57BL/6/129背景)的创建者购自杰克逊实验室(Jackson Labs)(库存#004154,也被称为6neo小鼠)。在基因疗法计划AAALAC认证的屏障小鼠设施中维持育种群,使用杂合子与杂合子交配,以便在相同窝中产生无效对照和WT对照。Gaa敲除小鼠是广泛使用的庞贝病模型。它们表现为溶酶体糖原在心脏、中枢神经系统、骨骼肌和膈肌中的进行性积聚,活动能力降低和进行性肌无力。小尺寸、可再现的表型和高效的育种允许进行快速研究,这对于临床前候选体内筛选是最佳的。

[0219] 动物饲养室保持在64°F至79°F(18℃至26℃)范围内的温度,其中湿度范围为30%至70%。

[0220] 将动物与其父母和同窝动物关在一起,直至断奶,然后关在转化研究实验室(TRL)GTP生态箱中的标准笼舍中,两至五只动物/笼舍。所有笼舍尺寸和居住条件均符合《实验动物护理和使用指南(Guide for the Care and Use of Laboratory Animals)》。笼舍、水瓶和垫料基底被高压灭菌放入屏障设施中。

[0221] 维持自动控制的12小时光/暗循环。每个黑暗期开始于19:00时(±30分钟)。随意提供食物(Purina,LabDiet®,5053,被照射过,PicoLab®,啮齿动物饮食20,25lb)。所有动物都可以经由单独放置在每个笼舍中的水瓶随意获取水。在每周更换笼子期间,每周至少更换一次水瓶。水供应来自费城(City of Philadelphia),并使用Getinge净水器进行氯化处理。通过ULAR每天检测氯化水平,并保持在百万分之二至百万分之四(ppm)。Nestlets™作为装饰物(enrichment)被提供给每个笼子。

[0222] 体内研究和组织学

[0223] 小鼠经由侧尾静脉(IV)被施用$5x10^{11}$GC(约$2.5x10^{13}$GC/kg)的剂量或$5x10^{10}$GC(约$2.5x10^{12}$GC/kg)的剂量的0.1mL中的AAVhu68.CAG.hGAA(各种hGAA构建体),在载体给药后第

7天和第21天放血用于血清分离，并且在注射后28天最后放血(用于血浆分离)并通过放血实施安乐死。迅速收集组织，从大脑开始。

[0224] 使用标准方法将用于组织学的组织福尔马林固定并石蜡包埋。脑和脊髓切片用luxol fast blue(luxol fast blue染色试剂盒，Abcam ab150675)染色，并且外周器官用PAS(高碘酸-Schiff)染色，使用标准方法来检测组织中的多糖，如糖原。对福尔马林固定的石蜡包埋的样本进行hGAA免疫染色。将切片脱蜡，在10mM柠檬酸盐缓冲液(pH 6.0)中煮沸以用于抗原修复，用PBS+0.2% Triton中的1％驴血清封闭15分钟，然后依次与在封闭缓冲液中稀释的初级抗体(Sigma HPA029126抗hGAA抗体)和生物素化次级抗体一起温育；使用基于HRP的比色反应来检测信号。

[0225] 由委员会认证的兽医病理家以盲法方式审查载玻片。建立了半定量评分系统以测量肌肉中庞贝相关组织学病变的严重程度(糖原贮积和自噬积聚)，以通过呈现贮积和/或空泡的细胞的总百分比所确定：

[0226]

组织评分	贮积
0	0％
1	1%至9％
2	10%至49%
3	50%至74%
4	75%至100%

[0227] 在适用的情况下，还对载体相关的组织病理病变进行了评估。

[0228] 非人灵长类动物

[0229] 对于载体施用，用肌肉注射右美托咪定和氯胺酮使恒河猴镇静，并施用单次枕大池内(ICM)注射或静脉内注射。如先前所述，使用荧光镜(OEC9800 C-Arm，GE)通过脊髓造影核实ICM注射的针放置(Katz N等人，《人类基因疗法方法(Hum Gene Ther Methods)》，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。

[0230] hGAA780I酶的体外性能的表征

[0231] GAA活性

[0232] 将匀浆组织的血浆或上清液与pH 4.0的5.6mM 4-MU-α-吡喃葡萄糖苷混合，并且在37℃下温育三小时。用pH 11.5的0.4M碳酸钠终止反应。使用Victor3荧光计测量相对荧光单位RFU，在355nm处激发和在460nm处发射。以nmol/mL/h为单位的活性通过从4-MU的标准曲线插值计算。针对匀浆上清液中的总蛋白质含量，对单独的组织样本中的活性水平进行归一化。等体积用于血浆样本。

[0233] 通过LC/MS的GAA特征肽

[0234] 血浆在100％甲醇中沉淀并离心。丢弃上清液。颗粒被掺入hGAA特有的稳定同位素标记的肽作为内标，并用胰蛋白酶重悬，并在37℃下温育一小时。用10％甲酸终止消化。肽通过C-18反相色谱法分离，并通过ESI-质谱法鉴定和定量。血浆中的总GAA浓度由特征肽浓度计算。

[0235] 细胞表面受体结合测定

[0236] 用受体包被96孔板，洗涤并用BSA封闭。将含有相同活性的rhGAA或工程改造的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，洗涤板以除去任何未结合的GAA，并在37℃下加入4-MU-α-吡喃葡萄糖苷一小时。用1.0M甘氨酸(pH 10.5)终止反应，并通过Spectramax荧光计读取RFU；在370nm处激发，在460nm处发射。每个样本的RFU通过从4-MU的标准曲线插值转换成nmol/mL/h。使用GraphPad Prism进行非线性回归。

[0237] 糖原-TFA水解

[0238] 组织匀浆用4N TFA在100℃下水解四小时，干燥并在水中重构。将水解的材料注射到CarboPac PA-10 2x250 mm柱上，用于通过高pH阴离子交换色谱法和脉冲电流计检测(HPAEC-PAD)进行葡萄糖确定。每个样本中游离葡萄糖的浓度通过从葡萄糖标准曲线插值计算。以μg糖原/mg蛋白报告最终数据。

[0239] 实例2:rAAVhu68.hGAA载体在Pompe小鼠中的评价

[0240] 将AAV载体稀释在无菌PBS中，用于IV递送至Pompe小鼠。测试物品包含：AAVhu68.CAG.hGAAco.rBG、AAVhu68.CAG.hGAAcoV780I.rBG、AAVhu68.CAG.BiP-vIGF2.hGAAco.rBG、AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG和AAVhu68.CAG.sp7co.Δ8.hGAAcoV780I.rBG。在研究中包含野生型对照和媒介物对照。

[0241] 在从治疗的小鼠收集的各种组织中测量hGAA蛋白表达和活性，包含肝(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)。在低剂量和高剂量两者下，所有启动子在肝脏中表现相同。施用在UbC启动子下表达的载体在两种剂量下均导致骨骼肌中的活性较低，并且具有CAG启动子的载体具有最好的总体活性。具有UbC启动子的载体在两种剂量下心脏中也具有较低的活性。

[0242] Pompe小鼠载体(PBS)对照(图5D)在心脏中显示出显著的糖原贮积(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具有接近完全至完全的贮积清除率。然而，接受编码hGAA参考序列(V780)的载体的两组显示中度至显著的纤维化淋巴细胞性心肌炎(图5B和图5C)，这存在于接受hGAA天然转基因的八只动物中的七只中和接受具有BiP和vIGF2修饰的工程改造的hGAA的八只动物中的三只中。由于接受hGAAcoV780I酶的小鼠都没有心肌炎(图5E、图5F和图5G)，因此认为该病变是载体相关的，并且更具体而言是hGAA参考序列特异性的。

[0243] 四头肌组织的分析显示，野生型小鼠和用编码V780I变体的载体(经进一步修饰或未经进一步修饰)处理的所有小鼠，具有接近完全至完全的贮积和自噬积聚的清除(图6A至图6H)。然而，接受编码hGAAV780的参考序列的载体的两组显示出极小至中度的糖原贮积剩余以及自噬积聚(图10)，共同证明了庞贝病的两个主要特征的次优纠正。从递送编码V780I变体的两个载体中观察到最佳结果，无论是其天然形式还是具有BiP-vIGF2修饰。sp7-Δ8修饰似乎导致庞贝病引起的组织学病变的不一致纠正。两种编码参考hGAAV780序列的构建体在清除糖原贮积和积聚方面均是次优的。

[0244] 在高剂量IV施用($5x10^{11}=2.5x10^{13}$GC/kg)下，hGAAcoV780I和BiP-vIGF2.hGAAcoV780I在四头肌肌肉中表现出接近正常的糖原水平，并且在细胞中摄取hGA有明显更好(图7A至图7H和图42)。对其他骨骼肌(包含胫骨前肌(TA)和腓肠肌)的评估显示了

类似的结果(具有V780I的变体,并且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施用导致最低水平。虽然四头肌肌肉中的糖原水平接近正常,但PAS染色显示了一些差异,其中hGAAcoV780I和BiP-vIGF2.hGAAcoV780I显示了最好的结果。免疫组织化学证实了hGAA在接受的小鼠的骨骼肌、心脏和脊髓中的表达(图43)。

[0245] 在低剂量IV施用时($5x10^{10}=2.5x10^{12}$GC/kg),BiP-vIGF2.hGAAcoV780I在心脏和四头肌肌肉中显示出比天然hGAAV780I更好的糖原减少(图41)。在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%,推测是由于更好的靶向性。在CNS中,观察到工程改造的构建体和V780I变体之间的有效协同效应。只有BiP-vIGF2.hGAAcoV780I清除了CNS糖原。

[0246] 如图8所示,脊髓组织学的评价显示,用AAVhu68.BiP-vIGF2.hGAAcoV780I处理的小鼠具有接近完全到完全的糖原贮积的清除,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。脑和脊髓切片的染色也揭示了用BiP-vIGF2.hGAAcoV780I进行纠正,但不是用天然hGAAV780酶进行纠正(图42)。结果证明了V780I突变和BiP-vIGF2修饰两者的贡献。

[0247] 实例3:DRG去靶向对Pompe小鼠中的hGAA表达的影响

[0248] BiP-vIGF2.hGAAcoV780I被修饰以包含四个mir183靶位点(BiP-vIGF2.hGAAcoV780I.4xmir183、SEQ ID NO:30)(图11),并被包装在AAVhu68衣壳中。

[0249] 载体基因组含有以下序列元件:

[0250] 反向末端重复序列(ITR):ITR是源自侧接载体基因组的所有组分的AAV2(130bp,GenBank:NC_001401)的相同、相反互补序列。当反式提供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。

[0251] 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。

[0252] 编码序列:编码BiP-vIGF2.hGAAcoV780I(SEQ ID NO:31)的工程改造的cDNA(SEQ ID NO:30的nt 1141至nt 4092)。

[0253] miR靶序列:四个串联的miR-183靶序列(SEQ ID NO:26)。

[0254] 兔β珠蛋白聚腺苷酸化信号(rBG PolyA):rBG PolyA信号(127bp)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′端处特异性切割事件以及添加长聚腺苷酸尾的信号。

[0255] 将miR183靶位点引入到BiP-vIGF2-hGAAcoV780I载体基因组中的效果在将AAVHu68 IV递送至Pompe小鼠后进行评估。如使用BiP-vIGF2.hGAAcoV780I构建体(没有mir183靶标)所观察到的,在包含miR183靶序列的载体的高剂量静脉内施用后,在CNS中纠正了糖原贮积(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聚得到完全纠正,而在低剂量施用后,观察到糖原贮积纠正和自噬积聚的部分纠正(图14)。在低剂量和高剂量的情况下也观察到心脏中糖原贮积的纠正(图15)。类似于在施用CAG.BiP-vIGF2.hGAAcoV780I的情况下所观察到的,自噬积聚在高剂量时完全消除,并且在低剂量时显著减少(图16)。结果证实,与不含miR靶序列的相应载体相比,添加miR183靶标不改变治

疗性转基因的功效。

[0256] 进行了剂量范围研究，以确定IV施用后BiP-vIGF2.hGAAcoV780I.4xmir183构建体的疗效概况和MED。在下表中提供了研究设计。

[0257]

组名称	1	2	3	4	5	6	7
N /组	4M/4F	4M/4F	4M/4F	4M/4F	4M/4F	4M/4F	4M/4F (WT)
施用途径	IV	IV	IV	IV	IV	IV	IV
载体剂量	5e12 GC/kg	1e13 GC/kg	2.5e13 GC/kg	5e13 GC/kg	1e14 GC/kg	PBS	PBS
持续时间				60 天			

[0258] 在IV施用AAVhu68.BiP-vIGF2.hGAAcoV780I.4xmir183载体60天后，对PAS染色的肌肉切片进行肌肉病理评分。通过使用用WGA(细胞膜;以允许测量肌纤维直径)，DAPI(细胞核;以量化中央核的存在)和LC3b抗体(自噬体;以量化自噬积聚)染色的免疫组织化学方法分析四头肌肌肉切片。对切片进行扫描并自动数字化，然后使用Visiopharm软件进行分析。

[0259] 中央核量化

[0260] 健康的肌纤维很少含有中央核(CN)(在WT肌肉中低于3％)，并且CN的存在表明肌肉再生。在WT对照和KO PBS对照之间，具有CN的纤维的百分比显著不同(图43)。与KO-PBS对照小鼠相比，在用$2x10^{11}$GC($1x10^{13}$GC/Kg)和更高载体剂量处理的组中，具有CN的纤维的百分比显著降低。在年轻小鼠中，增加CN的退化/再生循环在研究开始前没有发生，并通过导致Pompe小鼠表型纠正的治疗进行了预防。

[0261] LC3b量化

[0262] 在正常的生产性自噬条件下，自噬体被溶酶体快速降解，并且几乎无法检测到LC3阳性结构。在GAA KO小鼠中，受损/功能障碍的溶酶体可能触发自噬的增加;溶酶体不能与自噬体融合并降解自噬体的内容物，导致自噬积聚。

[0263] 在从$5x10^{11}$GC($2.5x10^{13}$GC/Kg)开始的所有剂量下，自噬积聚(LC3b+细胞％)均被防止(图44)。在三个月大的PBS对照组中观察到显著的自噬体积聚(纤维>20％)。

[0264] 股四头肌肌纤维的更小直径量化

[0265] 为了量化当与野生型对照动物相比时纤维大小的差异，将纤维直径分为小(<30μm)、中(30μm至50μm)和大(>50μm)的类别。KO PBS对照组在3个月大时显示出显著的萎缩。与KO PBS组相比，大股四头肌肌纤维的百分比显著增加，并且小股四头肌肌纤维的百分比减少(图45)。在用$2x10^{11}$GC($1x10^{13}$GC/kg)剂量和更高剂量治疗的GAA-/-小鼠中，小纤维(S)的比例显著降低，表明肌肉萎缩的预防。

[0266] 空泡化的严重程度

[0267] 结果显示在所有评估的肌肉中溶酶体贮积的剂量依赖性纠正(图46A至图46F)。在比目鱼肌和膈肌中在测试的最低剂量($5x10^{12}$GC/kg、$1x10^{11}$GC)下实现了纠正，而与在$2x10^{13}$GC/Kg($5x10^{11}$GC)的中间剂量下GAA KO PBS对照小鼠相比，大多数其他测试的肌肉得到了显著改善。肌肉病理完全不存在，并且切片显示与最高两个剂量-$5x10^{13}$GC/Kg($1x10^{12}$GC)和$1x10^{14}$GC/Kg($2x10^{12}$GC)的WT小鼠肌肉相似。

[0268] 实例4:在症状后的老年Pompe小鼠中的施用途径和剂量研究

[0269] 在静脉内(IV)和/或经由侧脑室(ICV)注射施用了编码hGAA的AAVhu68载体(包含

例如AAVhu68.CAG.BiP-vIGF2.hGAAcoV780I.rBG)的Pompe小鼠(以及野生型对照和媒介物对照)中评估了施用途径和剂量的效果。使用相同载体的双重途径施用方法(静脉内和注射入脑脊液)应该纠正该疾病的外周和神经表现。由于将有资格接受基因疗法的患者中的很大一部分已经患有晚期病理,因此我们选择对症状后的Pompe小鼠(六至七个月大)进行治疗,并在治疗后对其进行至少六个月的随访。

[0270] 小鼠使用静脉内(IV)、侧脑室内(ICV)或双重施用途径接受两种剂量水平(低剂量或高剂量)的载体。本研究中使用的剂量($5x10^{10}$或$1x10^{11}$GC ICV和$1x10^{13}$GC/kg或$5x10^{13}$GC/kg IV)对应于实例6中描述的NHP研究中使用的低剂量和高剂量以及适合于施用于人类的剂量($1x10^{13}$GC/kg和$5x10^{13}$GC/kg)。注射后约210天,在13个月至14个月龄时处死小鼠,以收集组织用于分析。在图47中提供了研究设计。

[0271] 在研究过程期间,使用旋转棒、钢丝悬吊和握力评估来测试小鼠的运动活动,并进行体积描记术。在从处理的小鼠收集的各种组织(包含血浆、四头肌、腓肠肌、膈肌和脑)中测量hGAA蛋白表达/活性和糖原贮积。进行组织学以评估,例如,PAS(经由Luxol fast blue染色)、hGAA表达和神经炎症(星形细胞病)。将组织切片染色以评估自噬积聚或清除(例如,使用标记LC3B的抗体)。

[0272] 对来自高剂量和低剂量ICV处理的(图28)以及高剂量和低剂量IV处理的(图29)小鼠的四头肌肌肉、心脏和脊髓样本进行组织学研究。在经由ICV途径接受低或高载体剂量的小鼠的脊髓中糖原贮积被纠正。高剂量IV施用可有效纠正四头肌肌肉、心脏和脊髓中的糖原贮积。

[0273] 在用低剂量和高剂量的ICV和IV载体的组合(双重施用途径)治疗的雄性中,体重得到显著纠正(图25A)。单一途径(单独的IV或单独的ICV)不能显著纠正体重。雌性Pompe和WT小鼠之间的体重没有差异(图25B)。

[0274] 对于接受高剂量IV(与基线相比并且与PBS对照组相比)的小鼠,握力显著改善(图26)。对于低剂量的ICV和IV施用的或双重途径施用(ICV LD+IV LC)的载体没有显著的益处。然而,高剂量IV和ICV的组合的施用早在注射后第30天就将强度恢复到野生型水平,并且在第180天该组合具有增加的益处(图27)。

[0275] 对不同组的肌肉病理进行研究,以观察庞贝病的相关发现。来自庞贝病患者和6^{neo} GAAKO Pompe小鼠模型的肌肉的特征在于存在结构异常,如纤维萎缩、红细胞大小不均、自噬积聚和中枢成核(图48)。

[0276] 中央核量化

[0277] 在六个月后治疗的小鼠中,在治疗前已经发生了退化/再生循环(图49)。

[0278] LC3b量化

[0279] 对具有预先存在的病理的Pompe小鼠进行治疗导致了在IV HD($1x10^{12}$GC=$5x10^{13}$GC/kg)和ICV+IV HD(ICV $1x10^{11}$GC和IV $1x10^{12}$GC)组中自噬体积聚的逆转(图50)。

[0280] 四头肌肌纤维直径量化

[0281] 在六个月后治疗的小鼠中,当肌肉萎缩已经明显时,小纤维(S)的比例在ICV+IV HD治疗组(ICV $1x10^{11}$GC和IV $1x10^{12}$GC=$1x10^{13}$GC/kg)中降低,并且大纤维(L)的比例在IV HD和ICV+IV HD治疗组中提高(图51)。在该治疗组中,肌纤维大小分布与WT小鼠相似,并且与PBS对照Pompe小鼠相比,在统计学上有显著的挽救效果,表明在该晚期疾病对症后治疗

范例中预先存在的病理的挽救。结果与同一组中握力恢复至WT水平相关，如图27所示。HD IV治疗还导致病理改善，大纤维比例增加，并且萎缩纤维呈减少趋势(图51)。

[0282] 这些发现支持双重施用途径优选靶向疾病的所有方面。载体的递送逆转了老年的症状后Pompe小鼠中预先存在的肌纤维病理，包含典型的治疗抗性的发现，如纤维萎缩和自噬积聚。

[0283] 实例5:向非人灵长类动物施用DRG去靶向基因疗法载体

[0284] 进行了NHP灵长类动物研究，以评估毒性并评估CAG.BiP-IGF2-hGAAcoV780I或CAG.BiP-IGF2-hGAAcoV780I-4xmir183向AAVhu68衣壳中的ICM递送。载体以$3x10^{13}$GC/kg进行ICM注射，并在第35天处死动物。

[0285] 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，在腰椎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靶序列的存在没有改变转基因在运动神经元中的表达(图19)，这表明施用载体将有利于减少庞贝病患者的运动神经元中的糖原贮积。此外，在递送含有miR183靶序列的构建体后，在心脏中的转基因表达没有减少(图20)。事实上，在心脏中的表达似乎有所增加，表明对于庞贝病患者的心脏病治疗，功效将提高。值得注意的是，miR183靶序列的串联重复序列降低了来自颈段和胸段的DRG感觉神经元中的毒性(图21A和图21B)。在该剂量水平下，腰段中的毒性没有降低(图21C)，这可能是由于如图18所示的腰椎水平的残留蛋白表达。

[0286] 在下表中提供了用于进一步评估将具有miR183靶序列的构建体IV递送至NHP的研究设计。该研究包含恒河猴GAA(rhGAA)序列，以评估非自身免疫应答的潜在影响。

[0287]

组名称	1	2
N/组	6	3
施用途径	IV	IV
载体剂量	1e13 GC/kg	1e13 GC/kg
载体	AAVhu68.CAG.BiP.vlGF2.hGAAco(V780l)-4xmiR183	AAVhu68.CAG.BiP.vlGF2.rhGAAco-4xmiR183
持续时间，终点	60 天。药理学（hGAA 表达水平）、毒理学（）心脏标志物+经典组)、组织病理学	

[0288] 此外，使用剂量范围研究来评估CAG.BiP-IGF2-hGAAcoV780I-4xmir183载体的安全性概况。对剂量范围研究中的NHP小鼠施用不同剂量的ICM，包含$3x10^{12}$GC、$6x10^{12}$GC和$1x10^{13}$GC。

[0289]

组名称	1	2	3
N/组	4	4	4
施用途径	ICM	ICM	ICM
载体剂量	3E12GC	6E12GC	1E13GC
载体	AAVhu68.CAG.BiP.vlGF2.hGAAco (V780I) 4xmi R183		
持续时间，终点	60 天*药理学(hGAA 表达水平)、毒理学(心脏标志物+经典组+ NCV)、组织病理学		

[0290] 实例6:非人灵长类动物中的施用途径研究

[0291] 进行了NHP灵长类动物研究，以评估毒性并评价载体施用的替代或组合途径。AAVhu68.CAG.BiP-IGF2-hGAAcoV780I以$5x10^{13}$GC/kg(高剂量)或$1x10^{13}$GC/kg(低剂量)IV施用或以$3x10^{13}$GC(高剂量)或$1x10^{13}$GC(低剂量)ICM施用。还评估了双重施用途径的可行性和

毒性，例如，通过IV和ICM剂量联合(IV $5x10^{13}$GC/kg+ICM$3x10^{13}$GC或IV $1x10^{14}$GC/kg+ICM $1x10^{13}$GC/kg)施用。IV和ICM剂量的组合可以显示在治疗庞贝病患者中将是有益的协同效应。

[0292] 在图31中提供了用于评估施用途径和剂量的研究设计。初步研究显示，低剂量IV注射的动物在四头肌和心脏中具有hGAA的表达(图37)。与ICM注射的动物相比，IV注射的动物也表现出较低等级的脊髓轴突病变(图33D至图33F)。还通过组织学在低剂量ICM注射的动物的脊髓中观察到hGAA的表达(图37)。

[0293] 在对人类GAA转基因进行免疫染色的载玻片上，对脊髓段中表达hGAA的运动神经元进行定量。结果显示，只有ICM给药和ICM+IV给药的动物具有有意义数量的表达转基因的脊髓下部运动神经元。这一结果表明，直接CSF施用将有利于纠正运动神经元病理。单独的IV LD剂量没有导致任何hGAA阳性运动神经元，而IV HD导致稀疏的hGAA阳性运动神经元(图55)。

[0294] ICM注射的动物中的DRG变性和脊髓轴突病变不是剂量依赖性的(图33A至图33F)。此外，一只IV低剂量动物(RA3607:$1x10^{13}$GC/Kg)和一只IV+ICM动物(180717:IV $5x10^{13}$GC/kG+ICM $3x10^{13}$GC)显示，与IV高剂量注射的动物相比，DRG变性、脊髓轴突病变和心脏炎症应答增加。然而，在ICM施用具有miR靶序列的构建体后，并且在没有炎症的情况下，观察到增加的心脏GAA表达(图20)。

[0295] (序列表自由文本)

[0296] 以下信息是为数字标识符<223>下含有自由文本的序列提供的。

[0297]

SEQ ID NO: （含有自由文本）	在<223>下的自由文本

[0298]

3	<223>合成构建体 <220> <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 76kD GAA 蛋白 <220> <221>MISC_FEATURE <222> (204)..(952) <223>具有 V780I 的 70kD GAA 蛋白
4	<223>工程改造的 hGAAI 编码序列
6	<223>包括 hGAA780I 的融合蛋白
7	<223>编码包括 GAAV780I 的融合蛋白的工程改造的序列 <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>工程改造的 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>工程改造的 hGAAV780I 成熟蛋白 <220> <221>CDS <222> (1)..(2649)

[0299]

15	<223>合成构建体
16	<223>hGAA780I 123-890 的工程改造的 DNA <220> <221>CDS <222> (1)..(2304)
17	<223>合成构建体
18	<223>工程改造的 hGAA 70kD cDNA <220> <221>CDS <222> (1)..(2247)
19	<223>合成构建体
20	<223>hGAAV780I 76 kD 蛋白的工程改造的 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 延伸段 <220> <221>misc_feature <222> (271)..(2952) <223>hGAA 61 - 952 780I 的工程改造的 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) <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 延伸段 <220> <221>misc_feature <222> (271)..(2952) <223>hGAA 61-952 V780 DNA

[0300]

	<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> 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
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> enhancer <222> (195)..(437) <223>CMV IE 增强子

[0301]

	<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>misc_feature <222> (4113)..(4134) <223>miR-183 靶标 <220> <221>misc_feature <222> (4139)..(4160) <223>miR-183 靶标 <220> <221>misc_feature <222> (4167)..(4188) <223>miR-183 靶标 <220> <221>misc_feature <222> (4195)..(4216) <223>miR-183 靶标 <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
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

[0302]

42	<223>IGF2 Δ1-6、Y27L、K65R
43	<223>IGF2 Δ1-7、Y27L、K65R
44	<223>IGF2 Δ1-4、E6R、Y27L、K65R
45	<223>IGF2 Δ1-4、E6R、Y27L
46	<223>IGF2 E6R
48	<223>vIGF2 Δ1-4、E6R、Y27L、K65R
50	<223>修饰的 BiP-1
51	<223>修饰的 BiP-2
52	<223>修饰的 BiP-3
53	<223>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

[0303] 本说明书中引用的所有文献均通过引用并入本文。随此提交的标有“21-9596PCT_ST25”的序列表和其中的序列和文本通过引用并入。2019年4月30日提交的美国临时专利申请第62/840,911号、2019年10月10日提交的美国临时专利申请第62/913,401号、2020年4月29日提交的国际专利申请第PCT/US20/30484号、2020年4月29日提交的国际专利申请第PCT/US20/30493号、2020年5月14日提交的美国临时专利申请第63/024,941号、2020年11月4日提交的美国临时专利申请第63/109,677号和2021年4月27日提交的美国临时专利申请第63/180,379号通过引用并入本文。虽然已经参考特定实施例描述了本发明，但是应当理解，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。

<110> 宾夕法尼亚大学的受托人AMICUS THERAPEUTICS公司
<120> 可用于治疗庞贝病的组合物
<130> UPN-21-9596.PCT
<150> US 63/024,941
<151> 2020-05-14
<150> US 63/109,677
<151> 2020-11-04
<150> US 63/180,379
<151> 2021-04-27
<160> 64
<170> PatentIn第3.5版
<210> 1
<211> 2211
<212> DNA
<213> 腺相关病毒hu.68
<220>
<221> CDS
<222> (1)..(2211)
<400> 1

```
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
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
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

<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

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
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
725 730 735

<210> 3
<211> 952
<212> PRT
<213> 人工序列
<220>
<223> 合成构建体
<220>
<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76kD GAA蛋白

<220>
<221> MISC_FEATURE
<222> (204)..(952)
<223> 具有V780I的70kD GAA蛋白
<400> 3

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

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
<220>
<223> 工程改造的hGAAI编码序列
<400> 4

atgggcgtta gacaccctcc ttgctctcac agactgctgg ccgtgtgtgc tctggtgtct 60
ctggctacag ctgccctgct gggacatatc ctgctgcacg actttctgct ggtgcccaga 120
gagctgtctg gcagctctcc tgtgctggaa gagacacacc ctgcacatca gcagggcgcc 180
tctagacctg gacctagaga cgctcaggcc catcctggca gacctagagc tgtgcccaca 240
cagtgtgacg tgccacctaa cagcagattc gactgcgccc ctgacaaggc catcacacaa 300
gagcagtgtg aagccagagg ctgctgctac atccctgcca aacaaggact gcagggcgct 360
cagatgggac agccctggtg cttcttccca ccatcttacc ccagctacaa gctggaaaac 420
ctgagcagca gcgagatggg ctacaccgcc acactgacca gaaccacacc tacattcttc 480
ccgaaggaca tcctgacact gcggctggac gtgatgatgg aaaccgagaa ccggctgcac 540
ttcaccatca aggaccccgc caatcggaga tacgaggtgc cactggaaac ccctcacgtg 600
cactctagag ccccatctcc actgtacagc gtggaatttt ccgaggaacc cttcggcgtg 660
atcgtgcgga gacagctgga tggaagagtg ctgctgaaca ccacagtggc ccctctgttc 720
ttcgccgacc agtttctgca gctgagcacc agcctgccta gccagtatat cacaggcctg 780
gccgagcacc tgtctccact gatgctgagc acatcctgga ccagaatcac cctgtggaac 840
agagatctgg cccctacacc tggcgccaac ctgtacggct ctcacccctt ttatctggcc 900
ctggaagacg gcggatctgc ccacggtgtt tttctgctga actccaacgc catggacgtg 960
gtgctgcagc catctcctgc tctgtcttgg agaagcacag gcggcatcct ggacgtgtac 1020
atctttctgg gccccgagcc taagagcgtg gtgcagcagt atctggacgt cgtgggctac 1080
cccttcatgc ctccttattg gggcctgggc ttccacctgt gcaggtgggg atacagcagc 1140
accgccatca ccagacaggt ggtggaaaac atgacccggg ctcacttccc actggacgtg 1200
cagtggaacg acctggacta catggacagc agacgggact tcaccttcaa caaggacggc 1260
ttcagagact tccccgccat ggtgcaagaa ctgcaccaag gcggcagacg gtacatgatg 1320
attgtggacc cagccatcag ctctagcggc cctgccggaa gctacagacc ttacgatgag 1380
ggcctgagaa gaggcgtgtt catcaccaac gagacaggcc agcctctgat cggcaaagtg 1440

tggcctggca gcacagcctt tccagacttc acaaacccca ccgctctggc ttggtgggaa 1500
gatatggtgg ccgagttcca cgatcaggtg cccttcgatg gcatgtggat cgacatgaac 1560
gagcccagca acttcatccg gggcagcgag gacggctgcc ccaacaacga actggaaaat 1620
cctccttacg tgcccggcgt tgtcggcgga acacttcagg ccgctacaat ctgtgccagc 1680
agccaccagt tcctcagcac ccactacaac ctgcacaacc tgtacggcct gaccgaggcc 1740
attgcctctc atagagccct ggttaaggcc agaggcaccc ggcctttgt gatcagcaga 1800
agcacattcg ccggccacgg cagatacgcc ggacattgga caggcgacgt gtggtctagt 1860
tgggagcagc tggctagcag cgtgccagag atcctgcagt tcaatctgct gggcgtgcca 1920
ctcgtgggag ccgacgtttg tggcttcctg ggcaacacca gcgaggaact gtgtgtgcgt 1980
tggacacagc tgggcgcctt ctatcccttc atgagaaacc acaacagcct gctgagcctg 2040
cctcaagagc cctacagctt tagcgagcct gcacagcagg ccatgagaaa ggccctgact 2100
ctgagatacg ctctgctgcc ccacctgtac accctgtttc accaggctca cgtggccggg 2160
gagacagtgg ctagacctct gttcctggaa tttcccaagg acagctccac ctggaccgtg 2220
gatcatcagc tgctgtgggg agaagccctg ctcatcacac ctgttctgca ggccggaaag 2280
gccgaagtga ccggctattt tcctctcggc acttggtacg acctgcagac cgtgcctatt 2340
gaggccctgg gatctcttcc tccacctcct gctgctccta gagagcctgc catccactct 2400
gaaggccagt gggttacact gcccgctcct ctggacacca tcaacgtgca cctgagagct 2460
ggctacatca tcccactgca aggccctggc ctgaccacaa ccgaatctag acagcagccc 2520
atggctctgg ccgtggctct tacaaaaggc ggagaggcta gaggcgagct gttctgggac 2580
gacggcgagt ctctggaagt gctggaacgg ggcgcttata cccaagtgat cttcctggcc 2640
agaaacaaca ccatcgtgaa cgaactcgtg cgcgtgacca gtgaaggtgc tggactgcaa 2700
ctgcagaaag tgaccgtgct cggagtggcc acagctcctc aacaggtgct gtctaacggc 2760
gtgcccgtgt ccaacttcac atacagcccc gacaccaagg tcctggacat ctgtgtgtca 2820
ctgctgatgg gcgagcagtt cctggtgtcc tggtgc 2856

<210> 5
<211> 2859
<212> DNA
<213> 智人
<400> 5

atgggagtga ggcacccgcc ctgctcccac cggctcctgg ccgtctgcgc cctcgtgtcc 60
ttggcaaccg ctgcactcct ggggcacatc ctactccatg atttcctgct ggttccccga 120
gagctgagtg gctcctcccc agtcctggag gagactcacc cagctcacca gcagggagcc 180
agtagaccag ggccccggga tgcccaggca caccccggcc gtcccagagc agtgcccaca 240
cagtgcgacg tccccccaa cagccgcttc gattgcgccc ctgacaaggc catcacccag 300
gaacagtgcg aggcccgcgg ctgttgctac atccctgcaa agcaggggct gcagggagcc 360
cagatggggc agccctggtg cttcttccca cccagctacc ccagctacaa gctggagaac 420
ctgagctcct ctgaaatggg ctacacggcc accctgaccc gtaccacccc caccttcttc 480
cccaaggaca tcctgaccct gcggctggac gtgatgatgg agactgagaa ccgcctccac 540
ttcacgatca aagatccagc taacaggcgc tacgaggtgc ccttggagac cccgcatgtc 600

cacagccggg caccgtcccc actctacagc gtggagttct ccgaggagcc cttcggggtg 660
atcgtgcgcc ggcagctgga cggccgcgtg ctgctgaaca cgacggtggc gcccctgttc 720
tttgcggacc agttccttca gctgtccacc tcgctgccct cgcagtatat cacaggcctc 780
gccgagcacc tcagtccccct gatgctcagc accagctgga ccaggatcac cctgtggaac 840
cgggaccttg cgcccacgcc cggtgcgaac ctctacgggt ctcacccttt ctacctggcg 900
ctggaggacg gcgggtcggc acacggggtg ttcctgctaa acagcaatgc catggatgtg 960
gtcctgcagc cgagccctgc ccttagctgg aggtcgacag gtgggatcct ggatgtctac 1020
atcttcctgg gcccagagcc caagagcgtg gtgcagcagt acctggacgt tgtgggatac 1080
ccgttcatgc cgccatactg gggcctgggc ttccacctgt gccgctgggg ctactcctcc 1140
accgctatca cccgccaggt ggtggagaac atgaccaggg cccacttccc cctggacgtc 1200
cagtggaacg acctggacta catggactcc cggagggact tcacgttcaa caaggatggc 1260
ttccgggact tcccggccat ggtgcaggag ctgcaccagg gcggccggcg ctacatgatg 1320
atcgtggatc ctgccatcag cagctcgggc cctgccggga gctacaggcc ctacgacgag 1380
ggtctgcgga gggggtttt catcaccaac gagaccggcc agccgctgat tgggaaggta 1440
tggcccgggt ccactgcctt ccccgacttc accaacccca cagccctggc ctggtgggag 1500
gacatggtgg ctgagttcca tgaccaggtg cccttcgacg gcatgtggat tgacatgaac 1560
gagccttcca acttcatcag gggctctgag gacggctgcc ccaacaatga gctggagaac 1620
ccaccctacg tgcctggggt ggttgggggg accctccagg cggccaccat ctgtgcctcc 1680
agccaccagt ttctctccac acactacaac ctgcacaacc tctacggcct gaccgaagcc 1740
atcgcctccc acagggcgct ggtgaaggct cgggggacac gcccatttgt gatctcccgc 1800
tcgacctttg ctggccacgg ccgatacgcc ggccactgga cgggggacgt gtggagctcc 1860
tgggagcagc tcgcctcctc cgtgccagaa atcctgcagt ttaacctgct gggggtgcct 1920
ctggtcgggg ccgacgtctg cggcttcctg ggcaacacct cagaggagct gtgtgtgcgc 1980
tggacccagc tgggggcctt ctacccctttc atgcggaacc acaacagcct gctcagtctg 2040
ccccaggagc cgtacagctt cagcgagccg gcccagcagg ccatgaggaa ggccctcacc 2100
ctgcgctacg cactcctccc ccacctctac acactgttcc accaggccca cgtcgcgggg 2160
gagaccgtgg cccggcccct cttcctggag ttccccaagg actctagcac ctggactgtg 2220
gaccaccagc tcctgtgggg ggaggccctg ctcatcaccc cagtgctcca ggccgggaag 2280
gccgaagtga ctggctactt ccccttgggc acatggtacg acctgcagac ggtgccaata 2340
gaggcccttg gcagcctccc acccccacct gcagctcccc gtgagccagc catccacagc 2400
gaggggcagt gggtgacgct gccggccccc ctggacacca tcaacgtcca cctccgggct 2460
gggtacatca tccccctgca gggccctggc ctcacaacca cagagtcccg ccagcagccc 2520
atggccctgg ctgtggccct gaccaagggt ggggaggccc gaggggagct tttctgggac 2580
gatggagaga gcctggaagt gctggagcga ggggcctaca cacaggtcat cttcctggcc 2640
aggaataaca cgatcgtgaa tgagctggta cgtgtgacca gtgagggagc tggcctgcag 2700
ctgcagaagg tgactgtcct gggcgtggcc acggcgcccc agcaggtcct ctccaacggt 2760
gtccctgtct ccaacttcac ctacagcccc gacaccaagg tcctggacat ctgtgtctcg 2820
ctgttgatgg gagagcagtt tctcgtcagc tggtgttag 2859

〈210〉 6

<211> 982
<212> PRT
<213> 人工序列
<220>
<223> 包括hGAA780I的融合蛋白
<400> 6

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

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
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> 7
<211> 2952
<212> DNA
<213> 人工序列
<220>
<223> 编码包括GAAV780I的融合蛋白的工程改造的序列
<220>
<221> misc_feature
<222> (810)..(810)
<223> V810I
<400> 7

atgaagctgt ctctggtggc tgctatgctg ctgctcctgt ctgccgccag agccagcaga 60
acactttgtg gcggagagct 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 gagcagtgtg aagccagagg ctgctgctac 420
atccctgcca aacaaggact gcagggcgcc cagatgggac agccttggtg cttcttccca 480
ccatcttacc ccagctacaa gctggaaaac ctgagcagct ccgagatggg ctacaccgcc 540
acactgacca gaaccacacc tacattcttc ccgaaggaca tcctgacact gcggctggac 600
gtgatgatgg aaaccgagaa ccggctgcac ttcaccatca aggaccccgc caatcggaga 660
tacgaggtgc ccctggaaac accccacgtg cactctagag caccctctcc actgtacagc 720
gtggaatttt ccgaggaacc cttcggcgtg atcgtgcgga gacagctgga tggcagagtg 780
ctcctgaata ccacagtggc ccctctgttc ttcgccgacc agtttctgca gctgagcacc 840
agcctgccta gccagtatat cacaggcctg gccgagcatc tgagccctct gatgctgagc 900

acatcctgga ccagaatcac cctgtggaac cgcgacctgg ctcctacacc tggcgccaat 960
ctgtacggct ctcacccctt ctacctggca ctggaagacg gtggatctgc ccacggtgtc 1020
tttctgctga atagcaacgc catggacgtg gtgctgcagc cctctcctgc actgtcttgg 1080
agatctacag gcggcatcct ggacgtgtac atctttctgg gccccgagcc taagagcgtg 1140
gtgcagcagt atctggacgt cgtgggctac cccttcatgc ctccttattg gggcctgggc 1200
ttccacctgt gtaggtgggg ctacagcagc accgccatca ccagacaggt ggtggaaaac 1260
atgacccggg ctcacttccc actggacgtg cagtggaacg acctggacta catggacagc 1320
agacgggact tcaccttcaa caaggacggc ttcagagact tccccgccat 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 gatatggtgg ccgagttcca cgatcaggtg 1620
cccttcgatg gcatgtggat cgacatgaac gagcccagca acttcatccg gggcagcgag 1680
gatggctgcc ccaacaacga actagaaaat cctccttacg tgcccggcgt tgtcggcgga 1740
acacttcagg ccgctacaat ctgtgccagc agccatcagt ttctgagcac ccactacaac 1800
ctgcacaacc tgtacggcct gaccgaggcc attgcctctc atagagccct ggttaaggcc 1860
agaggcaccc ggccttttgt gatcagcaga agcacattcg ccggccacgg cagatacgca 1920
ggacattgga caggcgacgt gtggtctagt tgggagcagc tggctagcag cgtgccagag 1980
atcctgcagt tcaatctgct gggcgtgcca ctcgtgggag ccgacgtttg tggcttcctg 2040
ggcaacacca gcgaggaact gtgtgtgcgt tggacacagc tgggcgcctt ctatcccttc 2100
atgagaaacc acaacagcct gctgagcctg cctcaagagc cctacagctt tagcgagcct 2160
gcacagcagg ccatgagaaa ggccctgact ctgagatacg ccctgctgcc tcacctgtac 2220
accctgtttc atcaggccca cgtggcaggc gagacagtgg ctagacctct gttcctggaa 2280
tttcccaagg acagctccac ctggaccgtg gatcatcagc tgctgtgggg agaagccctg 2340
ctgattacac cagtgctgca ggccggaaag gccgaagtga caggctattt ccctctcggc 2400
acttggtacg acctgcagac cgtgcctatc gaggctctgg gatctcttcc tccacctcct 2460
gccgctccta gagagcctgc cattcactct gaaggccagt gggttaccct gcctgctcct 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
acagctcctc agcaggttct gtctaatggc gtgcccgtgt ccaacttcac atacagcccc 2880
gacaccaagg tcctggacat ctgtgtgtcc ctgcttatgg gcgagcagtt cctggtgtcc 2940
tggtgctgat aa 2952

<210> 8
<211> 934
<212> DNA
<213> 人工序列

<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中的杂合内含子(β-肌动蛋白内含子和兔β1-珠蛋白
<220>
<221> misc_feature
<222> (526)..(934)
<223> 杂合内含子
<400> 8

attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt tccattgacg 60
tcaatgggtg gagtatttac ggtaaactgc ccacttggca gtacatcaag tgtatcatat 120
gccaagtacg ccccctattg acgtcaatga cggtaaatgg cccgcctggc attatgccca 180
gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag tcatcgctat 240
taccatggtc gaggtgagcc ccacgttctg cttcactctc cccatctccc cccccctccc 300
accccaatt ttgtatttat ttatttttta attattttgt gcagcgatgg gggcgggggg 360
gggggggggg cgcgcgccag gcggggcggg gcggggcgag gggcggggcg gggcgaggcg 420
gagaggtgcg gcggcagcca atcagagcgg cgcgctccga aagtttcctt ttatggcgag 480
gcggcggcgg cggcggccct ataaaaagcg aagcgcgcgg cgggcgggag tcgctgcgcg 540
ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc ggctctgact 600
gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg gctgtaatta 660
gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc ttgaggggct 720
ccgggagggc cctttgtgcg gggggagcgg ctcggggctg tccgcggggg gacggctgcc 780
ttcggggggg acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag 840
cctctgctaa ccatgttcat gccttcttct ttttcctaca gctcctgggc aacgtgctgg 900
ttattgtgct gtctcatcat tttggcaaag aatt 934

<210> 9
<211> 127
<212> DNA
<213> 人工序列

<220>
<223> 兔珠蛋白polyA
<400> 9
gatctttttc cctctgccaa aaattatggg gacatcatga agcccccttga gcatctgact 60
tctggctaat aaaggaaatt tattttcatt gcaatagtgt gttggaattt tttgtgtctc 120
tcactcg 127
<210> 10
<211> 188
<212> DNA
<213> 腺相关病毒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
<400> 11
tcttcctaga gcatggctac gtagataagt agcatggcgg gttaatcatt aactacaagg 60
aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 120
ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag 180
cgcgcagc 188
<210> 12
<211> 81
<212> DNA
<213> 人工序列
<220>
<223> 工程改造的hGAAV780I信号肽
<220>
<221> 信号_肽
<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
<220>
<223> 合成构建体
<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
<220>
<223> 工程改造的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

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
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

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

tcc tgg tgc 2649
Ser Trp Cys

<210> 15
<211> 883
<212> PRT

<213> 人工序列

<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
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

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
<220>
<223> hGAA780I 123-890的工程改造的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

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

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

<210> 17
<211> 768
<212> PRT
<213> 人工序列

<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
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
625 630 635 640

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

<210> 18
<211> 2247
<212> DNA
<213> 人工序列
<220>
<223> 工程改造的hGAA 70kD 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

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 1152
His Asn Leu Tyr Gly Leu Thr Glu Ala Ile Ala Ser His Arg Ala Leu
370 375 380

gtt aag gcc aga ggc acc cgg cct ttt gtg atc agc aga agc aca ttc 1200
Val Lys Ala Arg Gly Thr Arg Pro Phe Val Ile Ser Arg Ser Thr Phe
385 390 395 400

gcc ggc cac ggc aga tac gcc gga cat tgg aca ggc gac gtg tgg tct 1248
Ala Gly His Gly Arg Tyr Ala Gly His Trp Thr Gly Asp Val Trp Ser
405 410 415

agt tgg gag cag ctg gct agc agc gtg cca gag atc ctg cag ttc aat 1296
Ser Trp Glu Gln Leu Ala Ser Ser Val Pro Glu Ile Leu Gln Phe Asn
420 425 430

ctg ctg ggc gtg cca ctc gtg gga gcc gac gtt tgt ggc ttc ctg ggc 1344
Leu Leu Gly Val Pro Leu Val Gly Ala Asp Val Cys Gly Phe Leu Gly
435 440 445

aac acc agc gag gaa ctg tgt gtg cgt tgg aca cag ctg ggc gcc ttc 1392
Asn Thr Ser Glu Glu Leu Cys Val Arg Trp Thr Gln Leu Gly Ala Phe

```
    450                 455                 460
tat ccc ttc atg aga aac cac aac agc ctg ctg agc ctg cct caa gag      1440
Tyr Pro Phe Met Arg Asn His Asn Ser Leu Leu Ser Leu Pro Gln Glu
465                 470                 475                 480
ccc tac agc ttt agc gag cct gca cag cag gcc atg aga aag gcc ctg      1488
Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln Ala Met Arg Lys Ala Leu
                485                 490                 495
act ctg aga tac gct ctg ctg ccc cac ctg tac acc ctg ttt cac cag      1536
Thr Leu Arg Tyr Ala Leu Leu Pro His Leu Tyr Thr Leu Phe His Gln
            500                 505                 510
gct cac gtg gcc ggg gag aca gtg gct aga cct ctg ttc ctg gaa ttt      1584
Ala His Val Ala Gly Glu Thr Val Ala Arg Pro Leu Phe Leu Glu Phe
        515                 520                 525
ccc aag gac agc tcc acc tgg acc gtg gat cat cag ctg ctg tgg gga      1632
Pro Lys Asp Ser Ser Thr Trp Thr Val Asp His Gln Leu Leu Trp Gly
    530                 535                 540
gaa gcc ctg ctc atc aca cct gtt ctg cag gcc gga aag gcc gaa gtg      1680
Glu Ala Leu Leu Ile Thr Pro Val Leu Gln Ala Gly Lys Ala Glu Val
545                 550                 555                 560
acc ggc tat ttt cct ctc ggc act tgg tac gac ctg cag acc gtg cct      1728
Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr Asp Leu Gln Thr Val Pro
                565                 570                 575
att gag gcc ctg gga tct ctt cct cca cct cct gct gct cct aga gag      1776
Ile Glu Ala Leu Gly Ser Leu Pro Pro Pro Pro Ala Ala Pro Arg Glu
            580                 585                 590
cct gcc atc cac tct gaa ggc cag tgg gtt aca ctg ccc gct cct ctg      1824
Pro Ala Ile His Ser Glu Gly Gln Trp Val Thr Leu Pro Ala Pro Leu
        595                 600                 605
gac acc atc aac gtg cac ctg aga gct ggc tac atc atc cca ctg caa      1872
Asp Thr Ile Asn Val His Leu Arg Ala Gly Tyr Ile Ile Pro Leu Gln
    610                 615                 620
ggc cct ggc ctg acc aca acc gaa tct aga cag cag ccc atg gct ctg      1920
Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg Gln Gln Pro Met Ala Leu
625                 630                 635                 640
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
<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
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
290 295 300
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
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
<220>
<223> hGAAV780I 76 kD蛋白的工程改造的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
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

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

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
<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
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

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
625 630 635 640
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
<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延伸段
<220>
<221> misc_feature
<222> (271)..(2952)
<223> hGAA 61 - 952 780I的工程改造的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
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

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

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
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
<220>
<223> 合成构建体
<400> 23

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

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

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> 24

<211> 2952

<212> DNA

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> CDS

<222> (1)..(2952)
<220>
<221> misc_feature
<222> (1)..(270)
<223> BiP-vIGF peptide
<220>
<221> misc_feature
<222> (1)..(270)
<223> BiP信号肽+ vIGF2+2GS延伸段
<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
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
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
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
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
<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
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
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

<210> 26
<211> 22
<212> DNA
<213> 人工序列
<220>
<223> miRNA靶序列
<400> 26
agtgaattct accagtgcca ta 22

<210> 27
<211> 24
<212> DNA
<213> 人工序列
<220>
<223> miRNA靶序列
<400> 27
agtgtgagtt ctaccattgc caaa 24

<210> 28
<211> 4581
<212> DNA
<213> 人工序列
<220>
<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中的杂合内含子
<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_信号
<222> (4161)..(4287)
<223> 兔β-珠蛋白poly a
<220>
<221> misc_feature
<222> (4452)..(4581)
<223> 3’ ITR
<400> 28

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg taccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg 240
actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat 300
caagtgtatc atatgccaag tacgcccccc attgacgtca atgacggtaa atggcccgcc 360
tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta 420
ttagtcatcg ctattaccat ggtcgaggtg agccccacgt tctgcttcac tctccccatc 480
tcccccccct ccccaccccc aattttgtat ttatttattt tttaattatt ttgtgcagcg 540
atgggggcgg gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg 600
ggcggggcga ggcggagagg tgcggcggca gccaatcaga gcggcgcgct ccgaaagttt 660
ccttttatgg cgaggcggcg gcggcggcgg ccctataaaa agcgaagcgc gcggcgggcg 720
ggagtcgctg cgcgctgcct tcgccccgtg ccccgctccg ccgccgcctc gcgccgcccg 780
ccccggctct gactgaccgc gttactccca caggtgagcg ggcgggacgg cccttctcct 840
ccgggctgta attagcgctt ggtttaatga cggcttgttt cttttctgtg gctgcgtgaa 900
agccttgagg ggctccggga gggccctttg tgcgggggga gcggctcggg gctgtccgcg 960
gggggacggc tgccttcggg ggggacgggg cagggcgggg ttcggcttct ggcgtgtgac 1020
cggcggctct agagcctctg ctaaccatgt tcatgccttc ttctttttcc tacagctcct 1080
gggcaacgtg ctggttattg tgctgtctca tcattttggc 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
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

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

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 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 tctagagtcg 4122
Phe Leu Val Ser Trp Cys
980
acccgggcgg cctcgaggac ggggtgaact acgcctgaga tctttttccc tctgccaaaa 4182
attatgggga catcatgaag ccccttgagc atctgacttc tggctaataa aggaaattta 4242
ttttcattgc aatagtgtgt tggaattttt tgtgtctctc actcggttaa caacaacaat 4302
tgcattcatt ttatgtttca ggttcagggg gagatgtggg aggtttttta aagcaagtaa 4362
aacctctaca aatgtggtaa aatcgataag gatcttccta gagcatggct acgtagataa 4422
gtagcatggc gggttaatca ttaactacaa ggaaccccta gtgatggagt tggccactcc 4482
ctctctgcgc gctcgctcgc tcactgaggc cgggcgacca aaggtcgccc gacgcccggg 4542
ctttgcccgg gcggcctcag tgagcgagcg agcgcgcag 4581

<210> 29
<211> 982
<212> PRT
<213> 人工序列
<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
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
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> 30
<211> 4687
<212> DNA
<213> 人工序列
<220>
<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中的杂合内含子
<220>
<221> CDS
<222> (1141)..(4092)
<223> BiP-vIGF2-hGAAco
<220>
<221> misc_feature
<222> (3568)..(3570)
<223> Ile密码子
<220>
<221> misc_feature
<222> (4113)..(4134)
<223> miR-183靶标
<220>
<221> misc_feature
<222> (4139)..(4160)
<223> miR-183靶标
<220>
<221> misc_feature
<222> (4167)..(4188)
<223> miR-183靶标
<220>
<221> misc_feature
<222> (4195)..(4216)
<223> miR-183靶标

<220>
<221> polyA_信号
<222> (4267)..(4393)
<223> 兔β-珠蛋白poly a
<220>
<221> misc_feature
<222> (4558)..(4687)
<223> 3’ ITR
<400> 30

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg taccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg 240
actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat 300
caagtgtatc atatgccaag tacgcccccc attgacgtca atgacggtaa atggcccgcc 360
tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta 420
ttagtcatcg ctattaccat ggtcgaggtg agccccacgt tctgcttcac tctccccatc 480
tccccccccc ccccacccccc aattttgtat ttatttattt tttaattatt ttgtgcagcg 540
atgggggcgg gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg 600
ggcggggcga ggcggagagg tgcggcggca gccaatcaga gcggcgcgct ccgaaagttt 660
ccttttatgg cgaggcggcg gcggcggcgg ccctataaaa agcgaagcgc gcggcgggcg 720
ggagtcgctg cgcgctgcct tcgccccgtg ccccgctccg ccgccgcctc gcgccgcccg 780
ccccggctct gactgaccgc gttactccca caggtgagcg ggcgggacgg cccttctcct 840
ccgggctgta attagcgctt ggtttaatga cggcttgttt cttttctgtg gctgcgtgaa 900
agccttgagg ggctccggga gggccctttg tgcgggggga gcggctcggg gctgtccgcg 960
gggggacggc tgccttcggg ggggacgggg cagggcgggg ttcggcttct ggcgtgtgac 1020
cggcggctct agagcctctg ctaaccatgt tcatgccttc ttctttttcc tacagctcct 1080
gggcaacgtg ctggttattg tgctgtctca tcattttggc 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 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

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

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

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 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

accagtgcca taggatagtg aattctacca gtgccataca cgtgagtgaa ttctaccagt 4182
gccatagcat gcagtgaatt ctaccagtgc catagcggcc gcctcgaccc gggcggcctc 4242
gaggacgggg tgaactacgc ctgagatctt tttccctctg ccaaaaatta tggggacatc 4302
atgaagcccc ttgagcatct gacttctggc taataaagga aatttatttt cattgcaata 4362
gtgtgttgga atttttgtg tctctcactc ggttaacaac aacaattgca ttcattttat 4422
gtttcaggtt caggggggaga tgtgggaggt tttttaaagc aagtaaaacc tctacaaatg 4482
tggtaaaatc gataaggatc ttcctagagc atggctacgt agataagtag catggcgggt 4542
taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct ctgcgcgctc 4602
gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt gcccgggcgg 4662
cctcagtgag cgagcgagcg cgcag 4687

<210> 31
<211> 982
<212> PRT
<213> 人工序列
<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
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
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
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
<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
<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
<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
<220>
<223> V43L
<400> 35
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 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
<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
<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
<220>
<223> IGF2 S50I
<400> 38

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 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
<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
65
<210> 40
<211> 67
<212> PRT
<213> 人工序列
<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
<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
<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
<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
<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
<220>
<223> IGF2 Δ1-4, 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
<220>
<223> IGF2 E6R
<400> 46

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
<211> 201
<212> DNA
<213> 智人
<400> 47

gcttaccgcc ccagtgagac cctgtgcggc ggggagctgg tggacaccct ccagttcgtc 60
tgtggggacc gcggcttcta cttcagcagg cccgcaagcc gtgtgagccg tcgcagccgt 120
ggcatcgttg aggagtgctg tttccgcagc tgtgacctgg ccctcctgga gacgtactgt 180
gctacccccg ccaagtccga g 201

<210> 48
<211> 189
<212> DNA
<213> 人工序列
<220>
<223> vIGF2 Δ1-4, E6R, Y27L, K65R
<400> 48

tctagaacac tgtgcggagg ggagcttgta gacactcttc agttcgtgtg tggagatcgc 60
gggttcctct tctctcgccc cgcttccaga gtttcacgga ggtctagggg tatagtagag 120
gagtgttgtt tcaggtcctg tgacttggcg ctcctcgaga cctattgcgc gacgccagcc 180
aggtccgaa 189

<210> 49
<211> 18
<212> PRT
<213> 智人

<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
<220>
<223> 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
<211> 25
<212> PRT
<213> 人工序列
<220>
<223> 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
<220>
<223> 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
<220>
<223> 修饰的BiP-4
<400> 53
Met Lys Leu Ser Leu Val Ala Ala Met Leu Leu Leu Leu Ala Leu Val
1 5 10 15
Ala Leu Leu Leu Leu Ser Ala Ala Arg Ala
20 25
<210> 54
<211> 17
<212> PRT
<213> 海洋桡脚类动物(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
<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
<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

<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
<220>
<223> 接头序列
<400> 58
Gly Gly Gly Gly Ser Gly Gly Gly Ser
1 5
<210> 59
<211> 9
<212> PRT
<213> 人工序列
<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
<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
<400> 61
Trp Ile Asp Met Asn Glu
1 5

<210> 62
<211> 7
<212> PRT
<213> 智人
<400> 62
Thr Val Pro Ile Glu Ala Leu
1 5
<210> 63
<211> 9
<212> PRT
<213> 智人
<400> 63
Gln Thr Val Pro Ile Glu Ala Leu Gly
1 5
<210> 64
<211> 25
<212> PRT
<213> 智人
<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

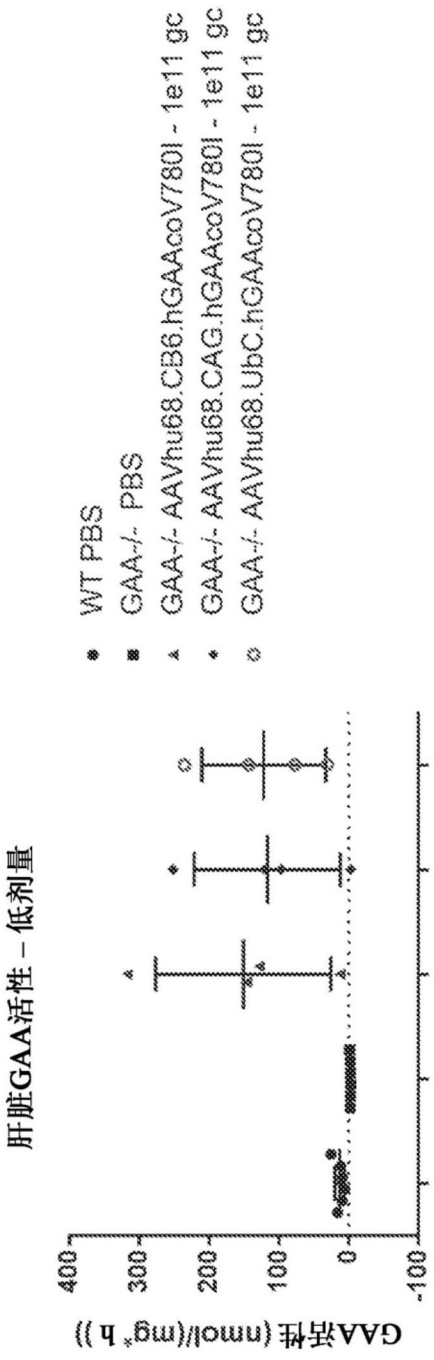

肝脏GAA活性 – 低剂量
GAA活性 (nmol/(mg*h))
400
300
200
100
0
-100
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e11 gc

图1A

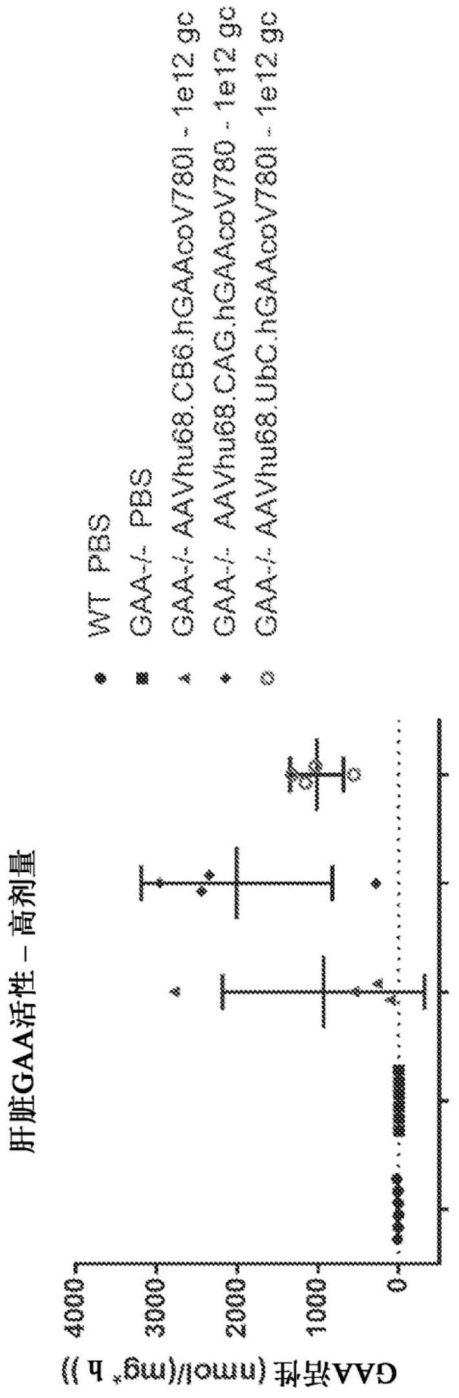

肝脏GAA活性 – 高剂量
GAA活性 (nmol/(mg*h))
4000
3000
2000
1000
0
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e12 gc
GAA-/- AAVhu68.CAG.hGAAcoV780 - 1e12 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e12 gc

图1B

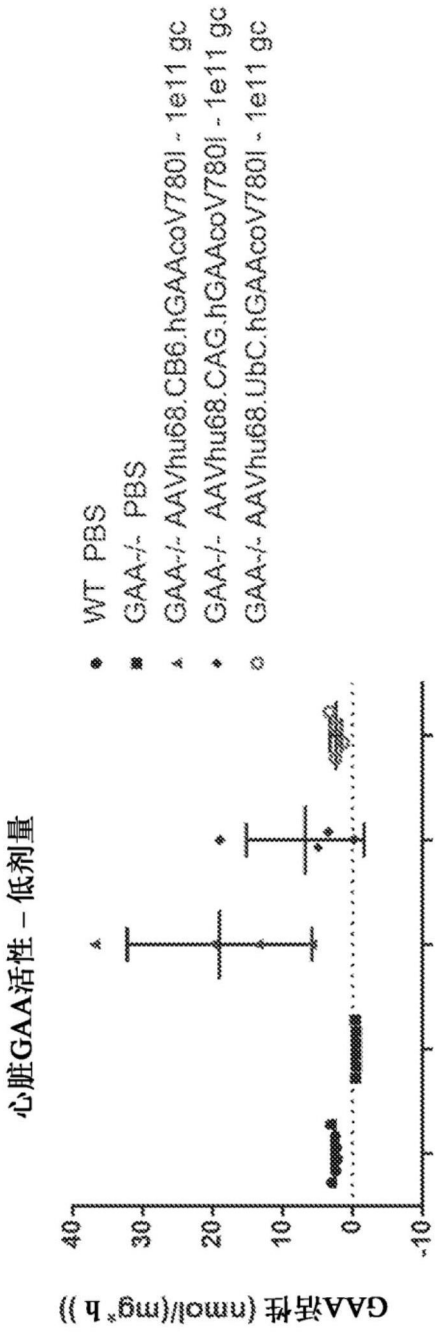

心脏GAA活性 – 低剂量
GAA活性 (nmol/(mg*h))
40
30
20
10
0
-10
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e11 gc

图2A

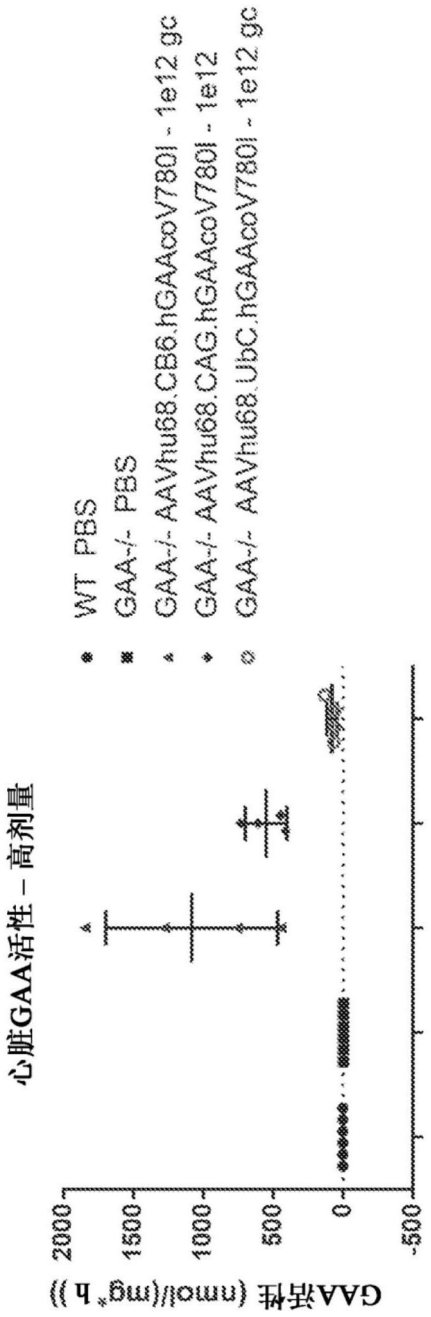

心脏GAA活性 – 高剂量
GAA活性 (nmol/(mg*h))
2000
1500
1000
500
0
-500
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e12 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e12
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e12 gc

图2B

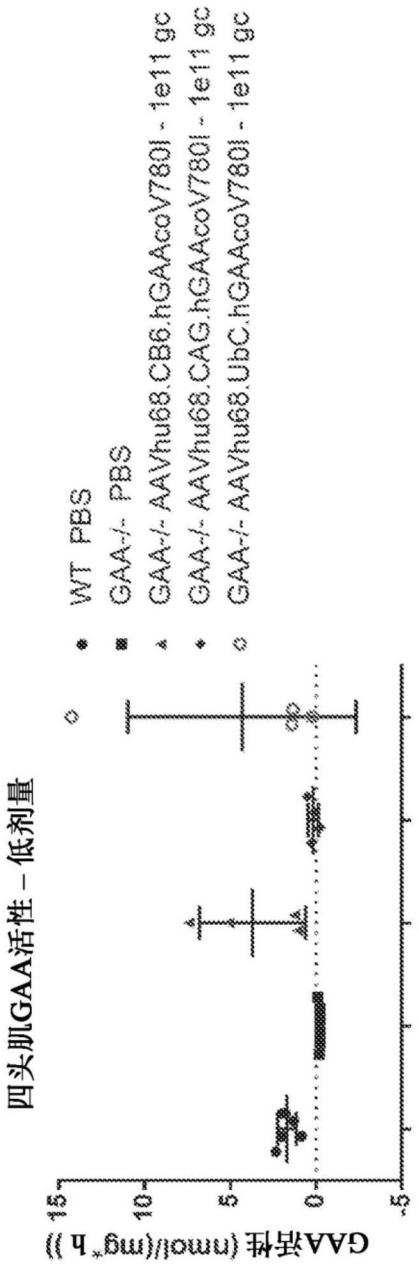

四头肌GAA活性 – 低剂量
GAA活性 (nmol/(mg*h))
15
10
5
0
-5
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e11 gc

图3A

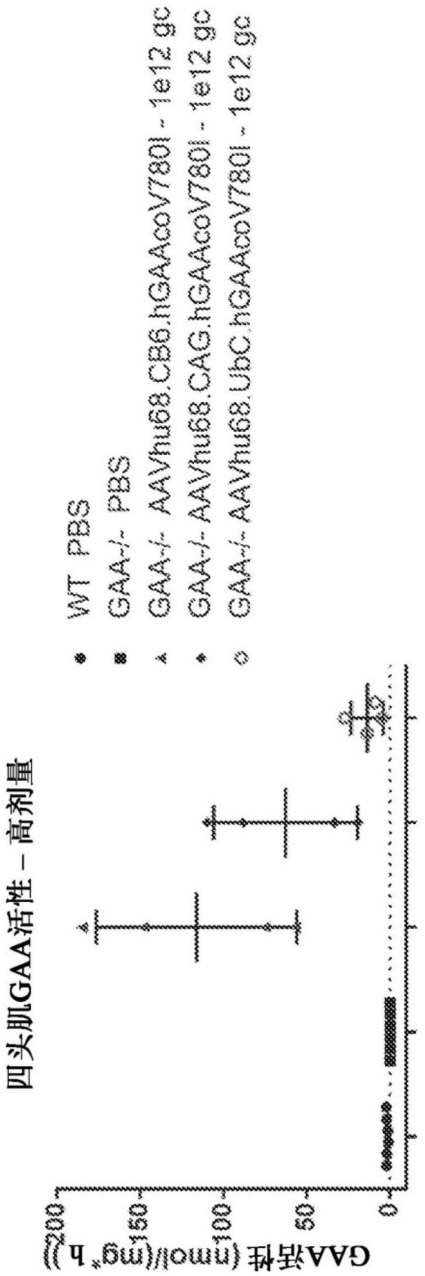

四头肌GAA活性 – 高剂量
GAA活性 (nmol/(mg*h))
200
150
100
50
0
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e12 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e12 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e12 gc

图3B

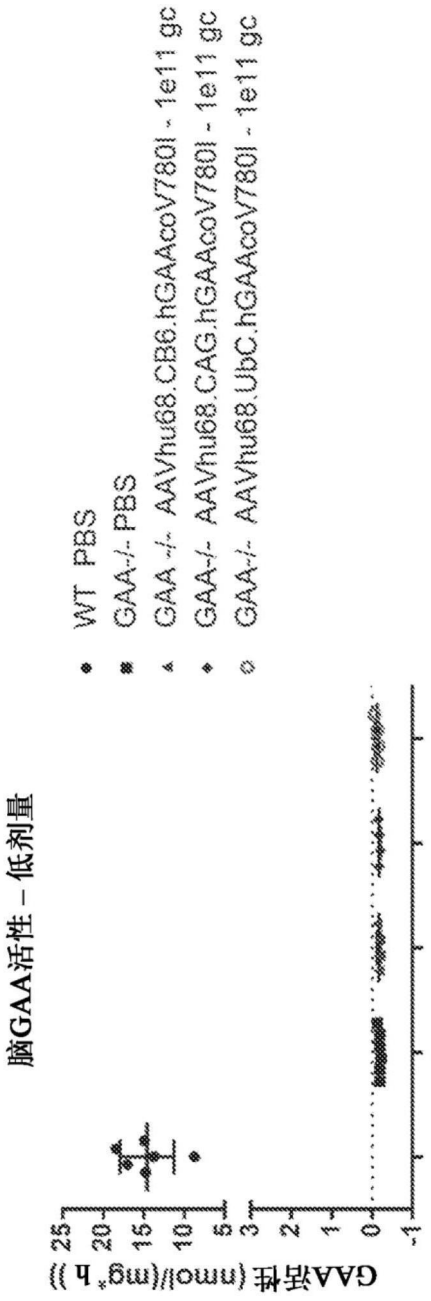

脑GAA活性 – 低剂量
GAA活性 (nmol/(mg*h))
25
20
15
10
5
3
2
1
0
-1
WT PBS
GAA-/- PBS
GAA -/- AAVhu68.CB6.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.CAG.hGAAcoV780I - 1e11 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e11 gc

图4A

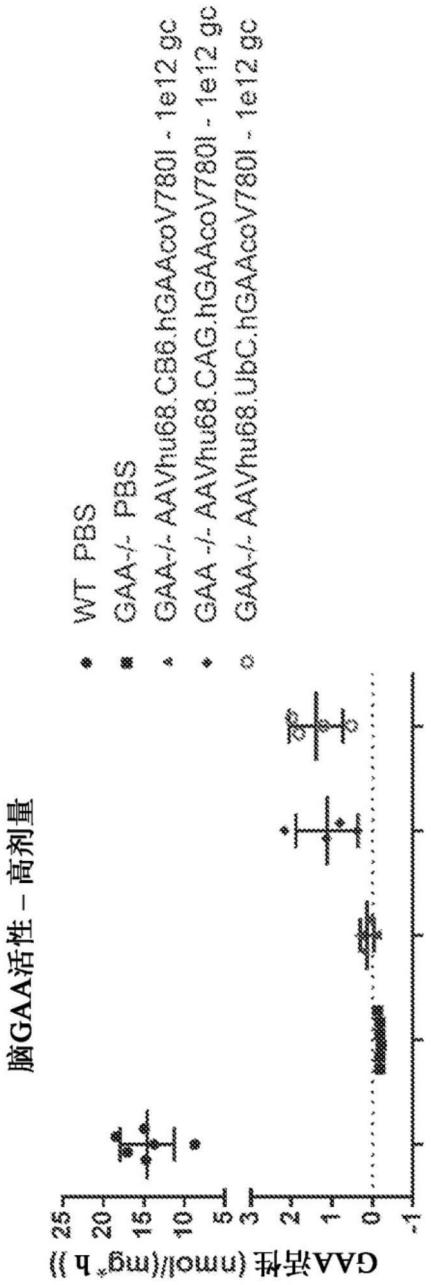

脑GAA活性 – 高剂量
GAA活性 (nmol/(mg*h))
25
20
15
10
5
3
2
1
0
-1
WT PBS
GAA-/- PBS
GAA-/- AAVhu68.CB6.hGAAcoV780I - 1e12 gc
GAA -/- AAVhu68.CAG.hGAAcoV780I - 1e12 gc
GAA-/- AAVhu68.UbC.hGAAcoV780I - 1e12 gc

图4B

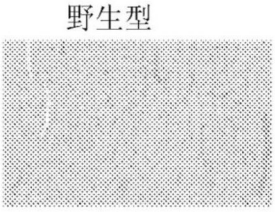

野生型

图5A

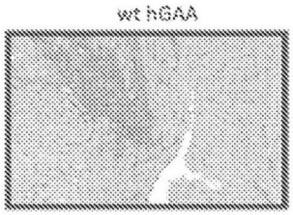


图5B

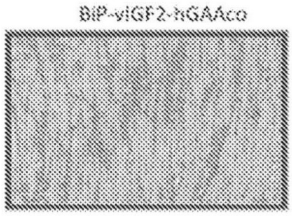


图5C

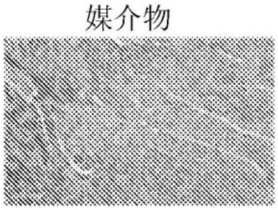


图5D

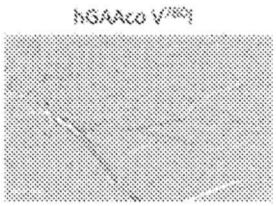


图5E

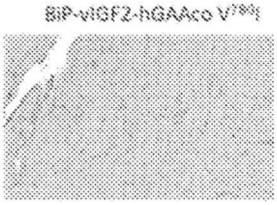


图5F

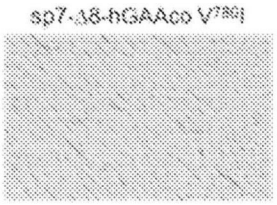

sp7-Δ8-hGAAco V780I

图5G

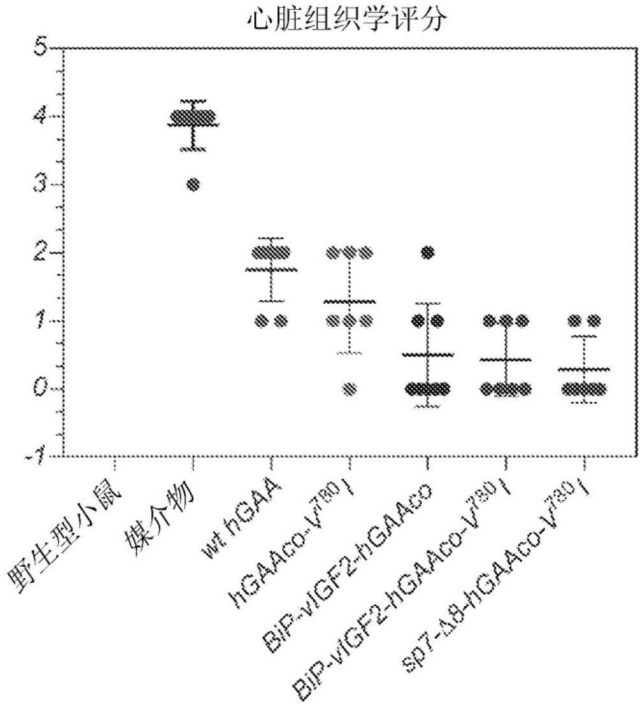

心脏组织学评分
5
4
3
2
1
0
-1
野生型小鼠
媒介物
wt hGAA
hGAAco-V780I
BiP-vIGF2-hGAAco
BiP-vIGF2-hGAAco-V780I
sp7-Δ8-hGAAco-V780I

图5H

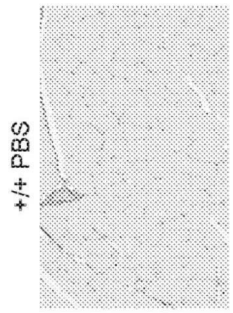

+/+ PBS

图6A

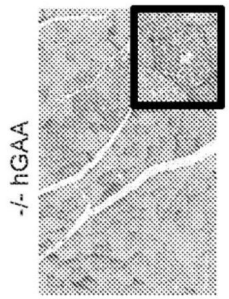


图6B

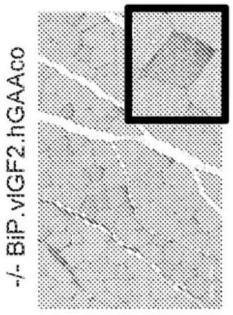


图6C

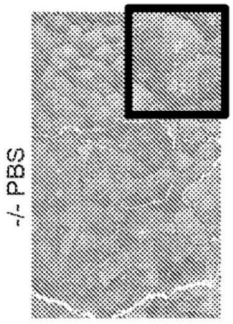


图6D

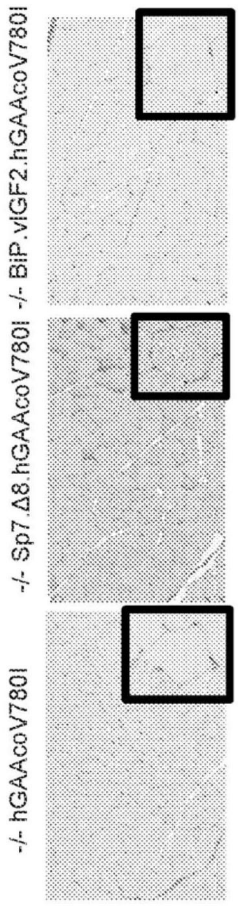


图6E 图6F 图6G

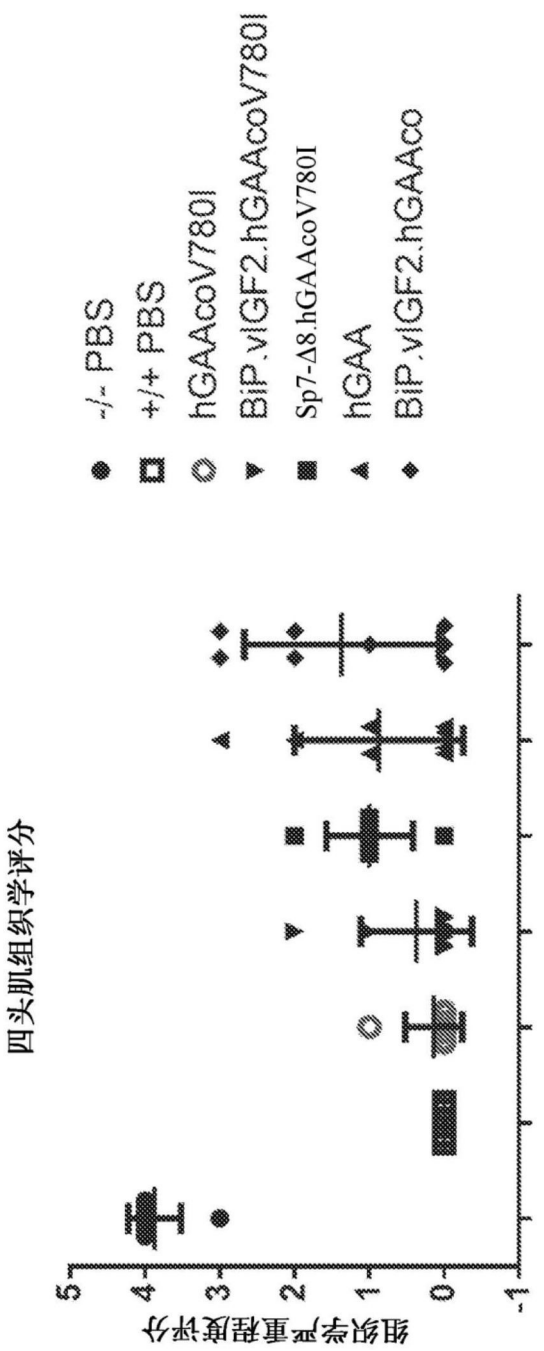

-/- PBS
+/+ PBS
hGAAcoV780I
BiP.vIGF2.hGAAcoV780I
Sp7-Δ8.hGAAcoV780I
hGAA
BiP.vIGF2.hGAAco
四头肌组织学评分
组织学严重程度评分
5
4
3
2
1
0
-1

图6H

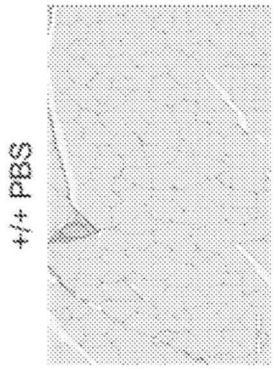


图7A

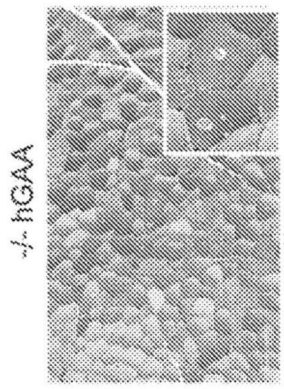


图7B

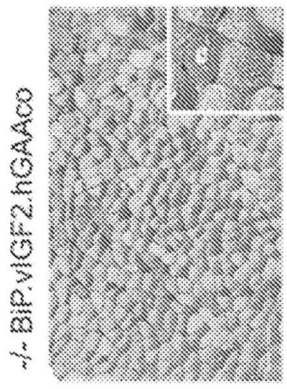


图7C

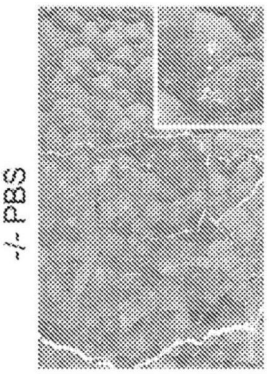


图7D

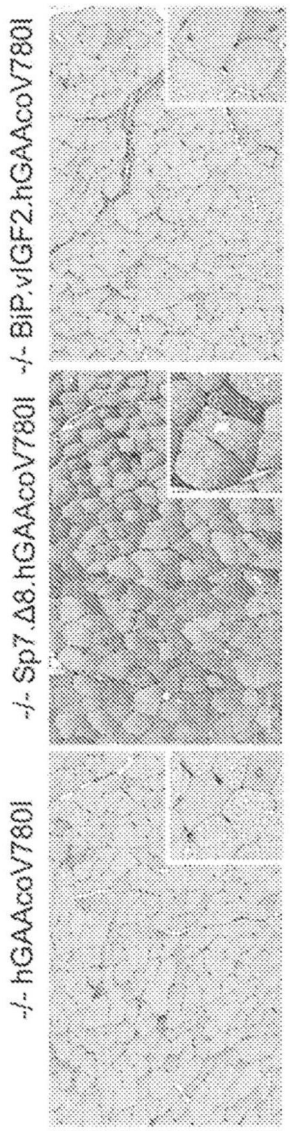


图7E 图7F 图7G

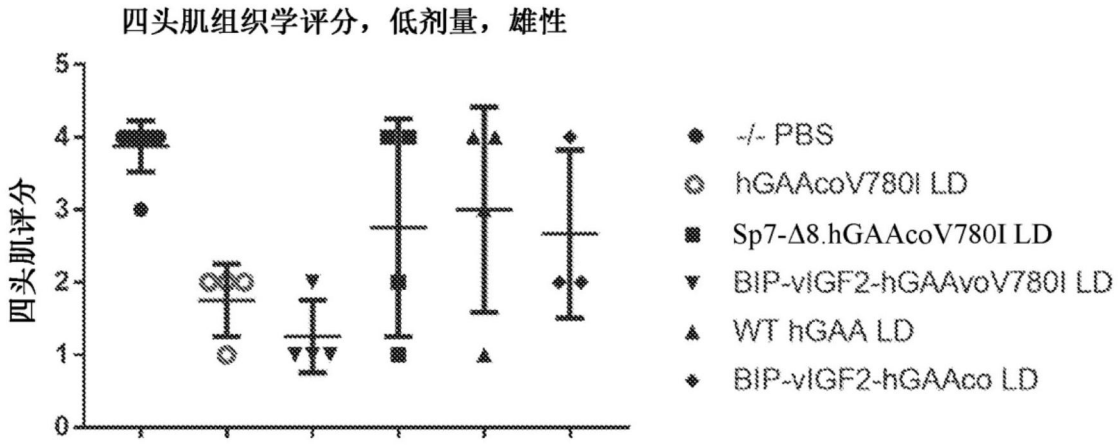

四头肌组织学评分，低剂量，雄性
四头肌评分
5
4
3
2
1
0
-/- PBS
hGAAcoV780I LD
Sp7-Δ8.hGAAcoV780I LD
BIP-vIGF2-hGAAvoV780I LD
WT hGAA LD
BIP-vIGF2-hGAAco LD

图7H

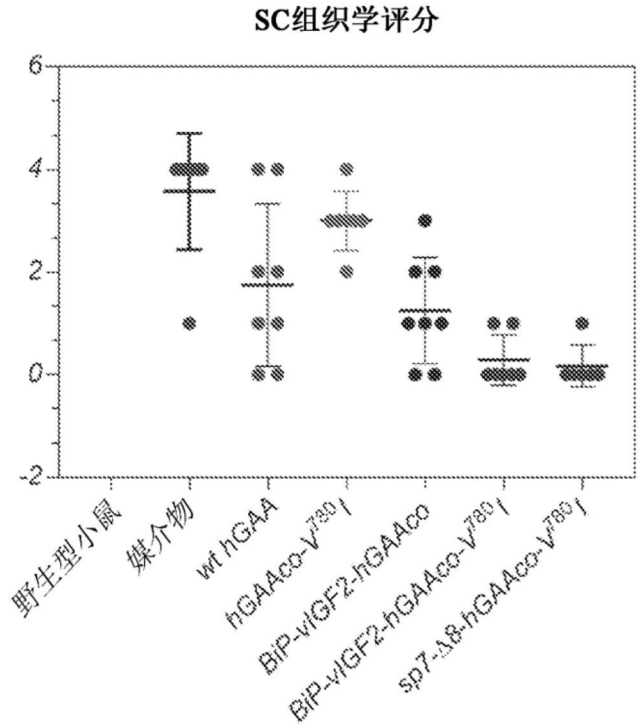

SC组织学评分
6
4
2
0
-2
野生型小鼠
媒介物
wt hGAA
hGAAco-V780I
BiP-vIGF2-hGAAco
BiP-vIGF2-hGAAco-V780I
sp7-Δ8-hGAAco-V780I

图8

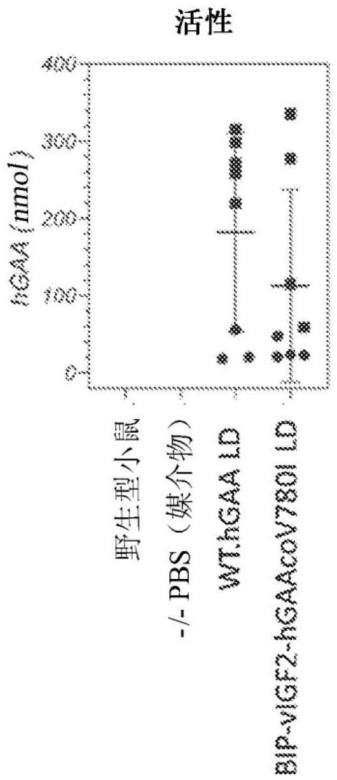

活性
hGAA (nmol)
0
10
20
30
40
野生型小鼠
-/- PBS（媒介物）
WT.hGAA LD
BIP-vIGF2-hGAAcoV780I LD

图9A

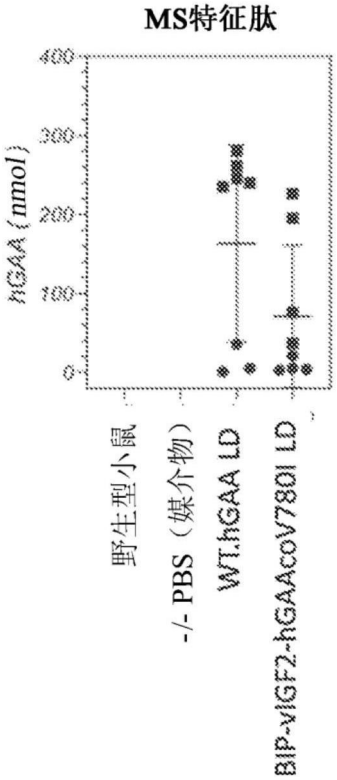

MS特征肽
hGAA (nmol)
0
10
20
30
40
野生型小鼠
-/- PBS（媒介物）
WT.hGAA LD
BIP-vIGF2-hGAAcoV780I LD

图9B

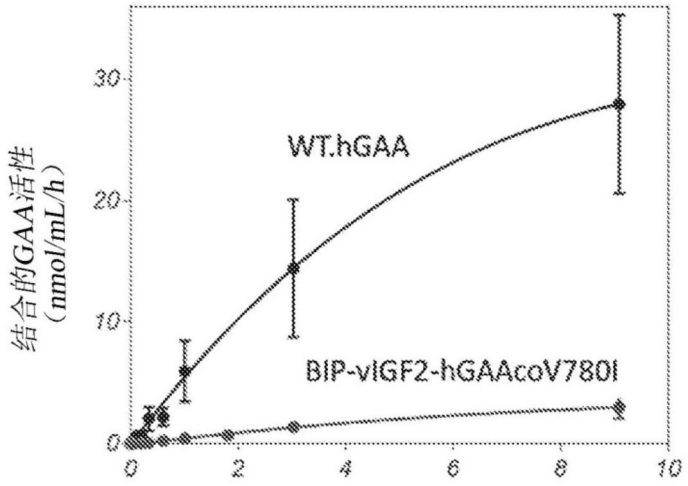

结合的GAA活性
（nmol/mL/h）
30
20
10
0
0
2
4
6
8
10
WT.hGAA
BIP-vIGF2-hGAAcoV780I

图9C

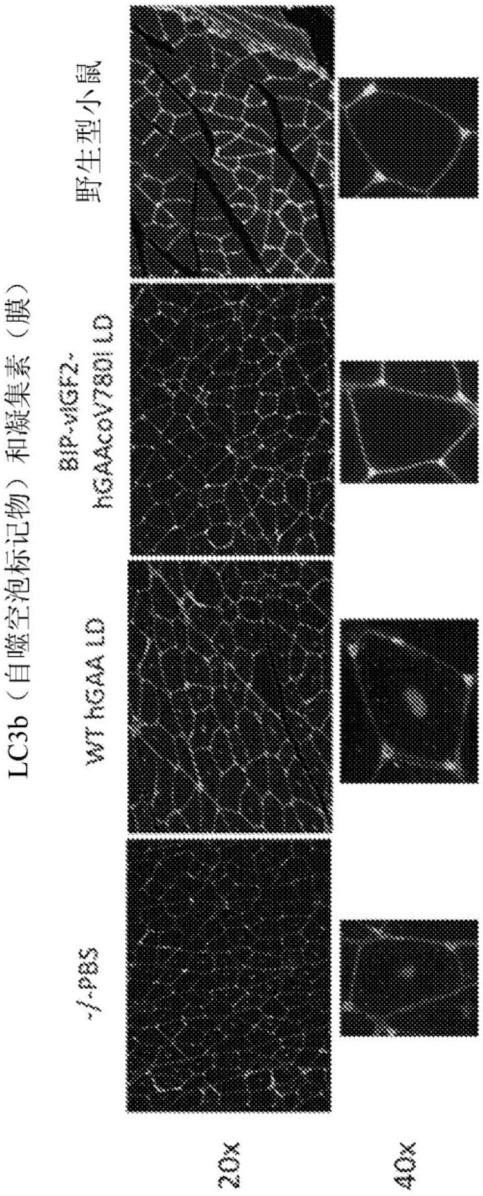

LC3b（自噬空泡标记物）和凝集素（膜）
-/-PBS
WT hGAA LD
BIP-sIGF2-hGAAcoV780i LD
野生型小鼠
20x
40x

图10

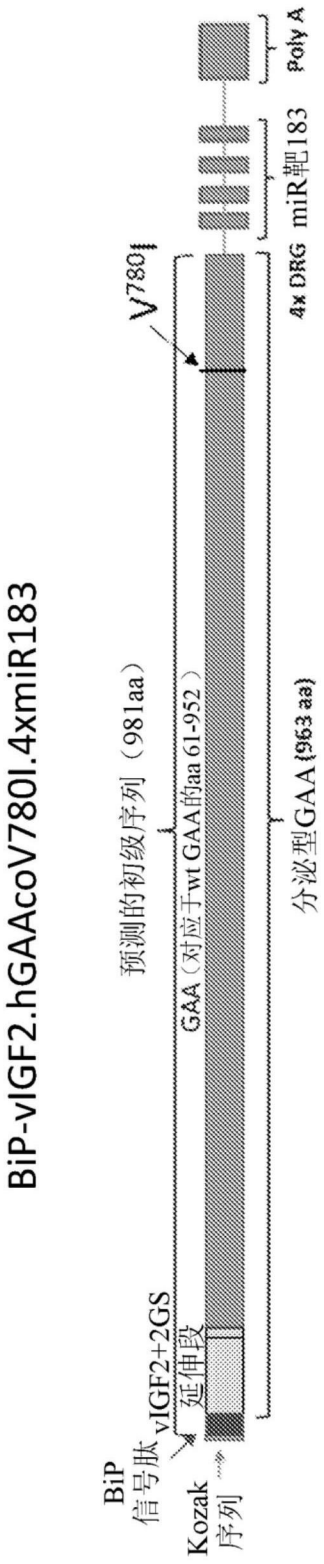

BiP-vIGF2.hGAAcoV780I.4xmiR183
BiP
信号肽
Kozak
序列
vIGF2+2GS
延伸段
预测的初级序列（981aa）
GAA（对应于wt GAA的aa 61-952）
V780I
分泌型GAA (963 aa)
4x DRG
miR靶183
Poly A

图11

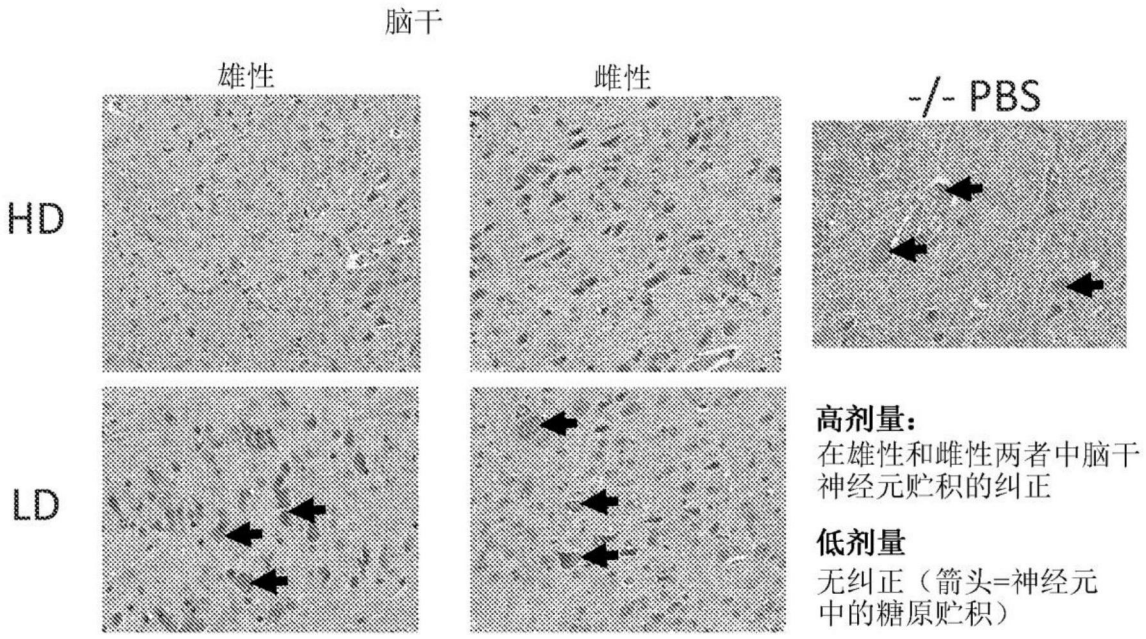

脑干
雄性
雌性
-/- PBS
HD
LD
高剂量：
在雄性和雌性两者中脑干神经元贮积的纠正
低剂量
无纠正（箭头=神经元中的糖原贮积）

图12

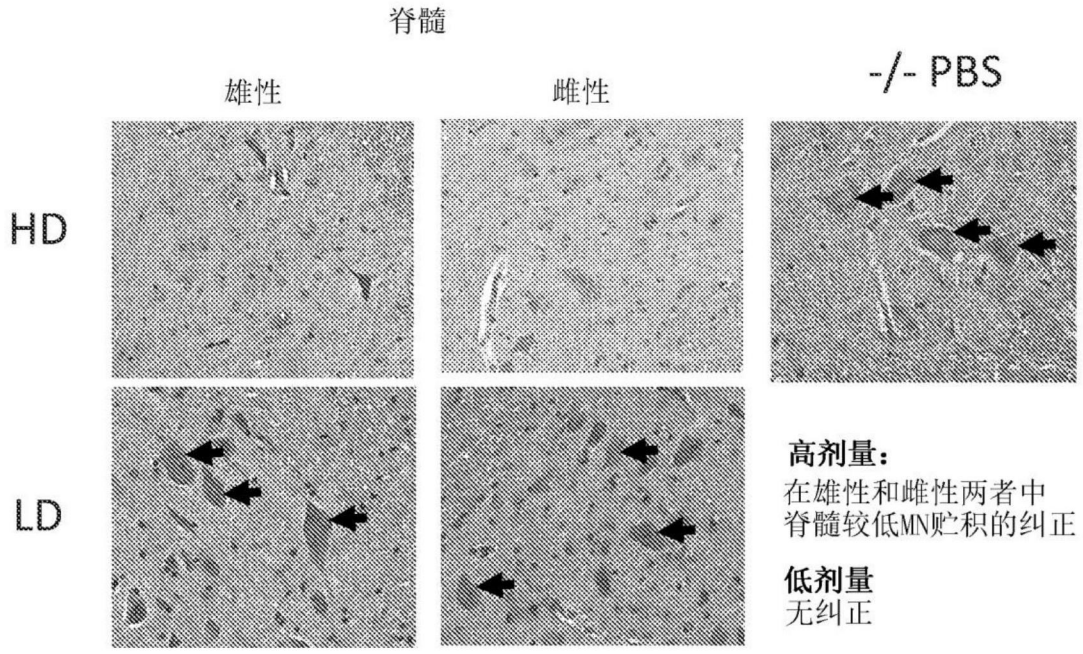

脊髓
雄性
雌性
-/- PBS
HD
LD
高剂量：
在雄性和雌性两者中脊髓较低MN贮积的纠正
低剂量
无纠正

图13

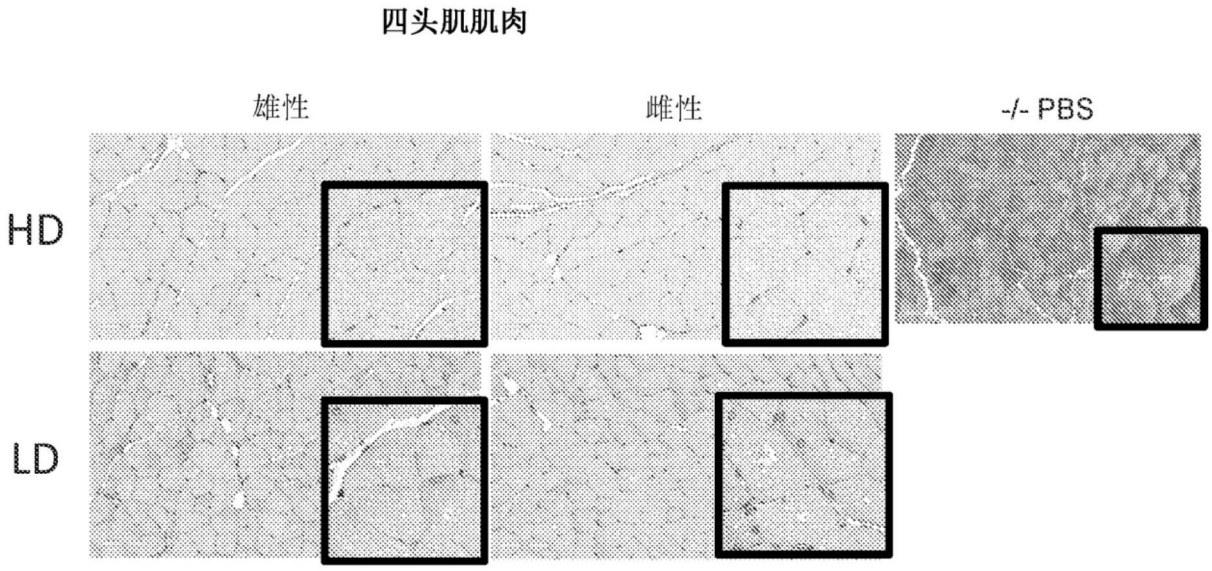


图14

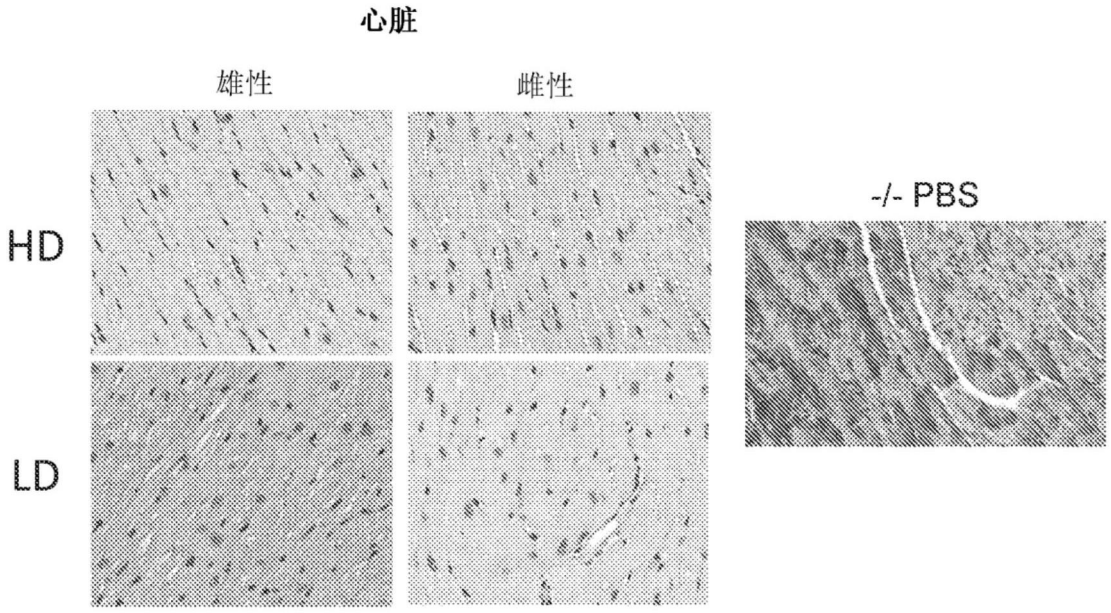


图15

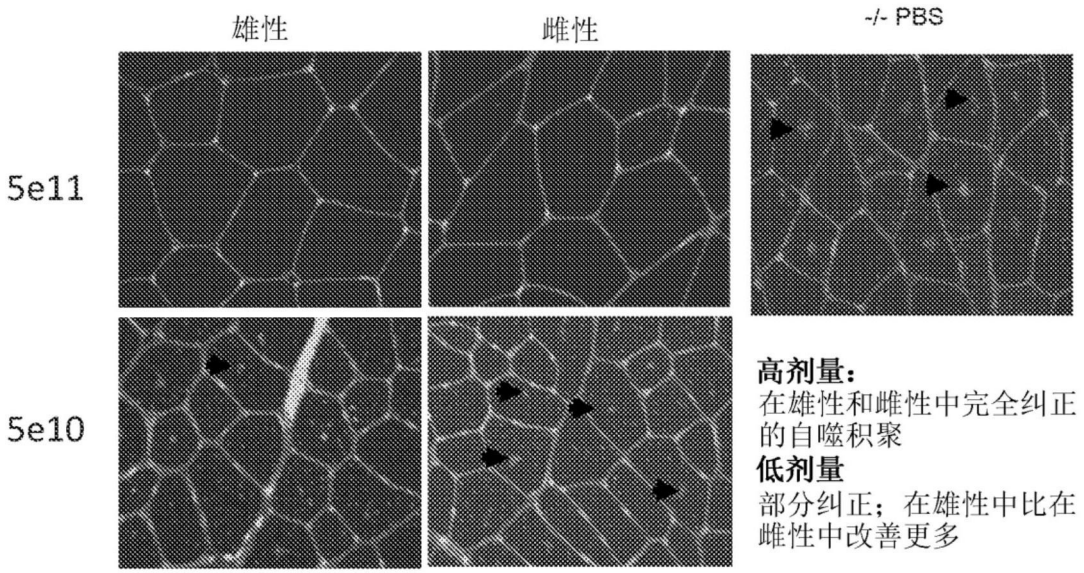

雄性
雌性
-/- PBS
5e11
5e10
高剂量：
在雄性和雌性中完全纠正的自噬积聚
低剂量
部分纠正；在雄性中比在雌性中改善更多

图16

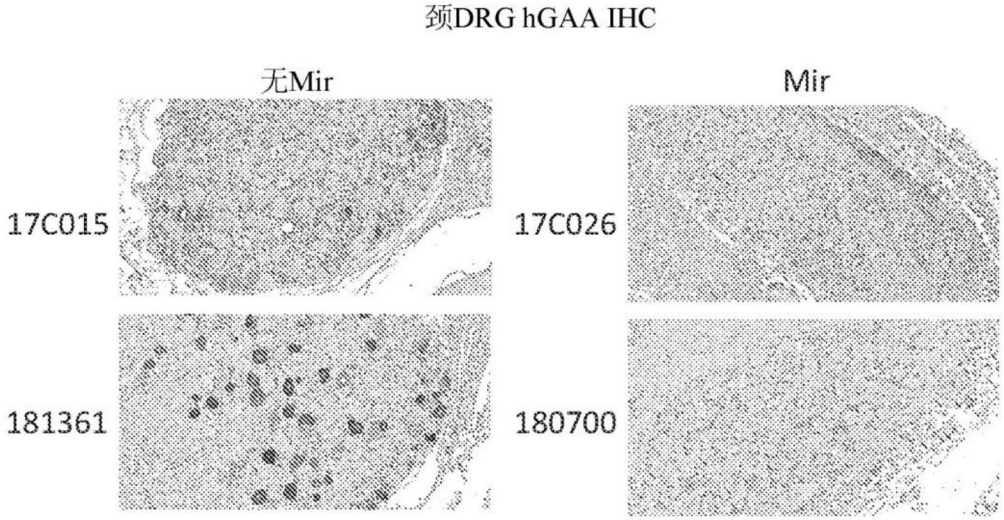

颈DRG hGAA IHC
无Mir
Mir
17C015
17C026
181361
180700

图17

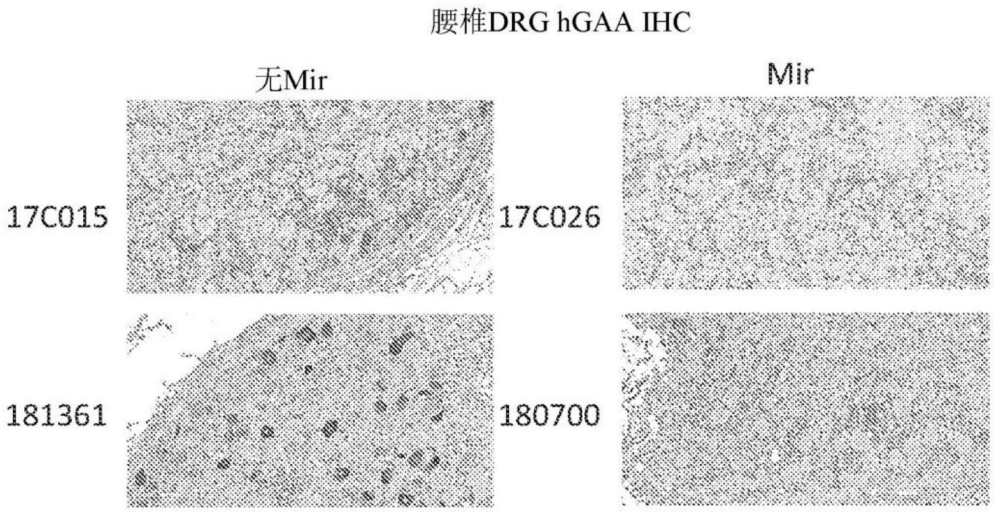

腰椎DRG hGAA IHC
无Mir
Mir
17C015
17C026
181361
180700

图18

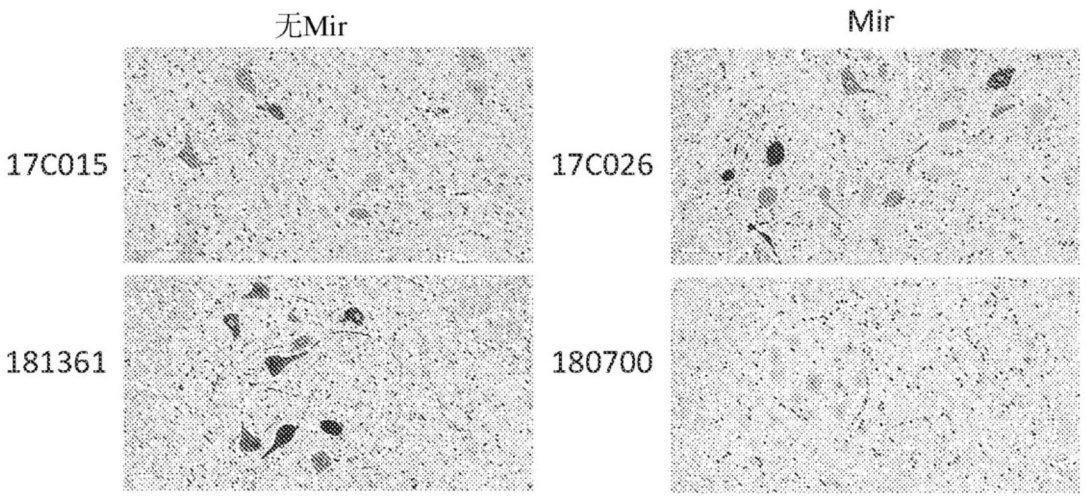

颈脊髓
较低的MN hGAA IHC
无Mir
Mir
17C015
17C026
181361
180700

图19

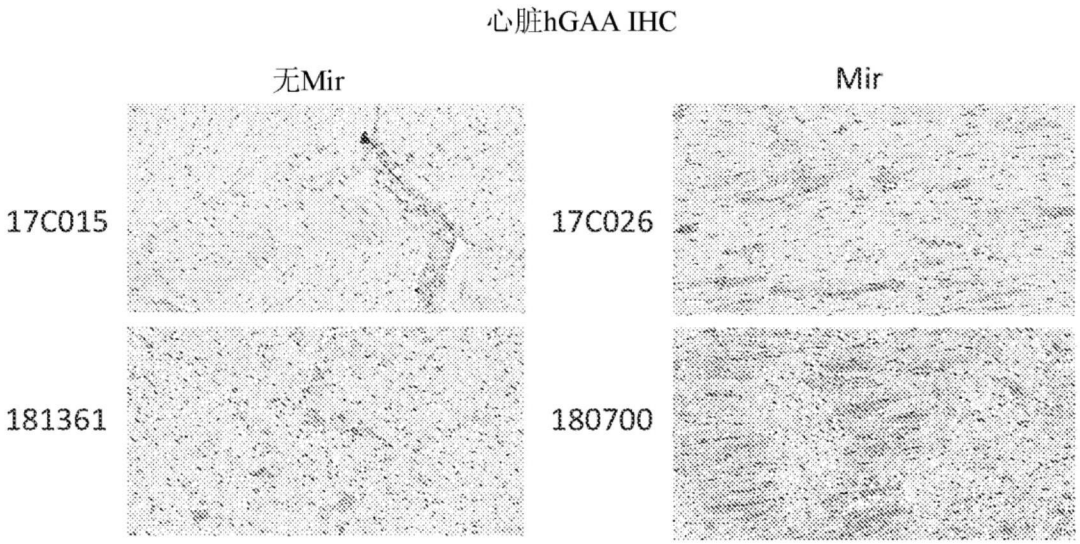

心脏hGAA IHC
无Mir
Mir
17C015
17C026
181361
180700

图20

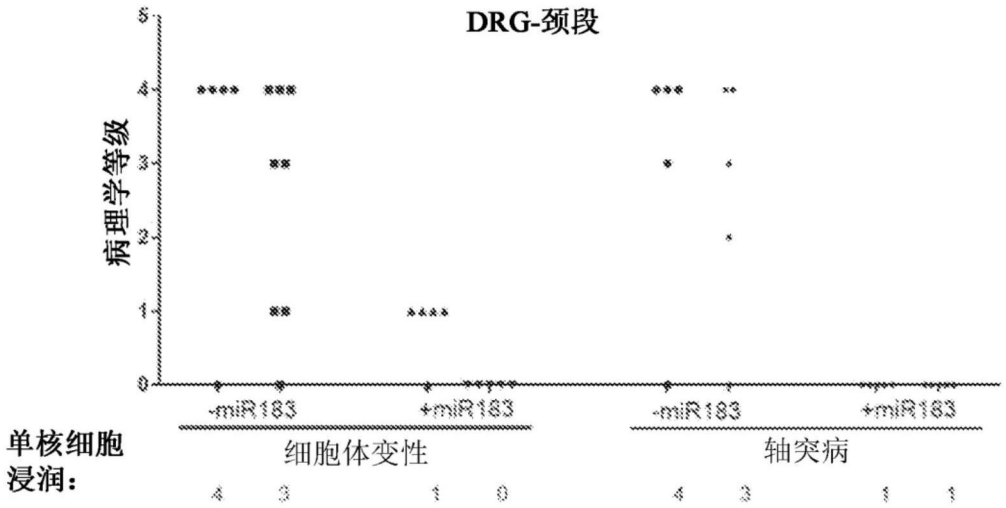

DRG-颈段
病理学等级
-miR183
+miR183
-miR183
+miR183
细胞体变性
轴突病
单核细胞浸润:
4
3
1
0
4
3
1
1

图21A

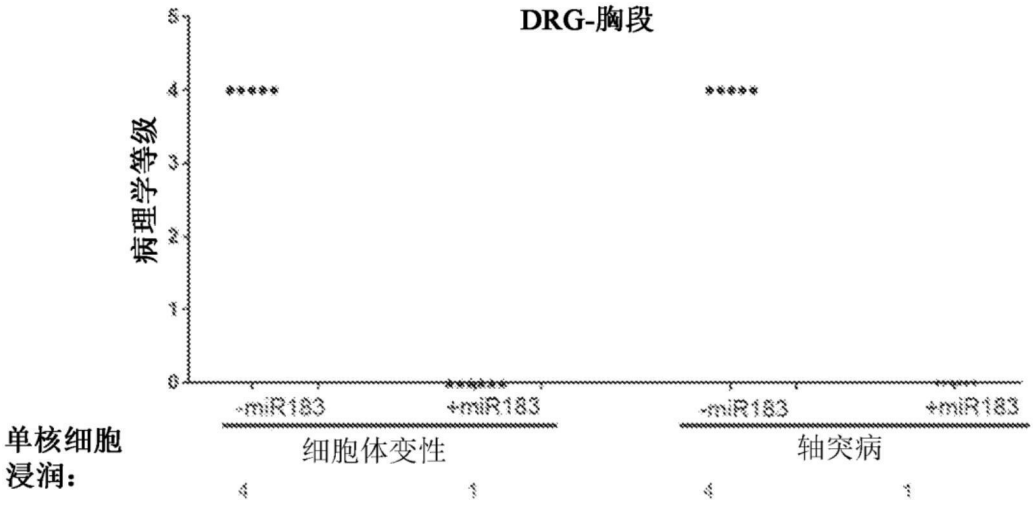

DRG-胸段
病理学等级
-miR183
+miR183
-miR183
+miR183
细胞体变性
轴突病
单核细胞浸润:
4
1
4
1

图21B

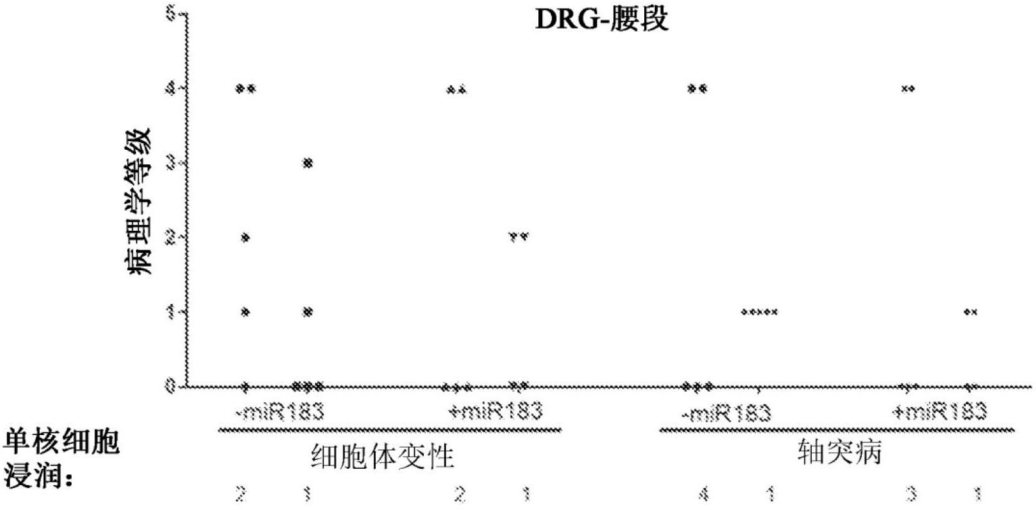

DRG-腰段
病理学等级
-miR183
+miR183
-miR183
+miR183
细胞体变性
轴突病
单核细胞浸润:
2
1
2
1
4
1
3
1

图21C

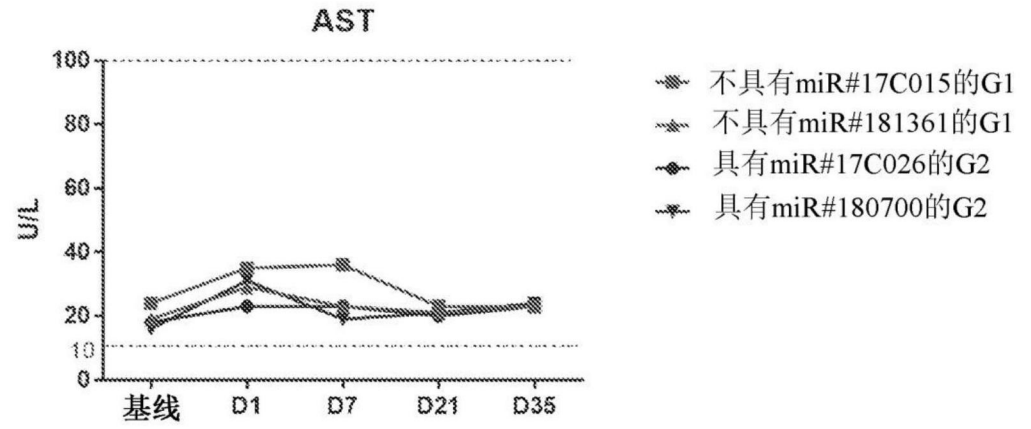


图22A

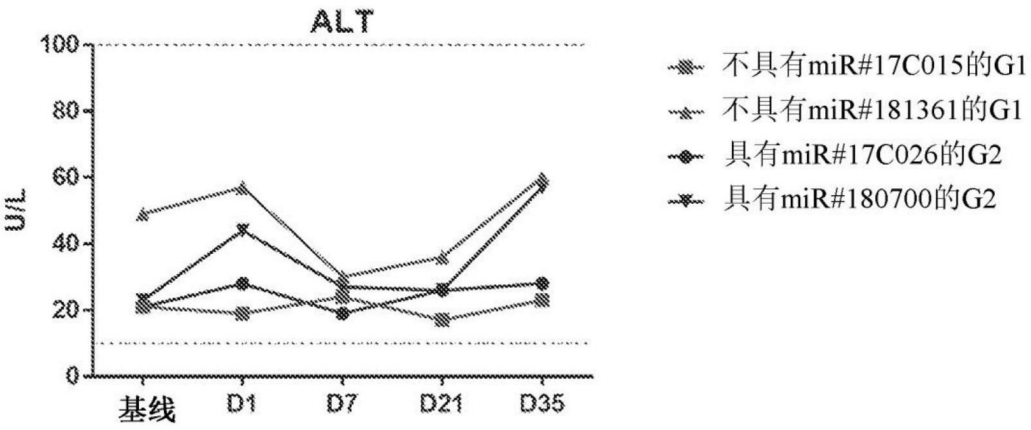


图22B

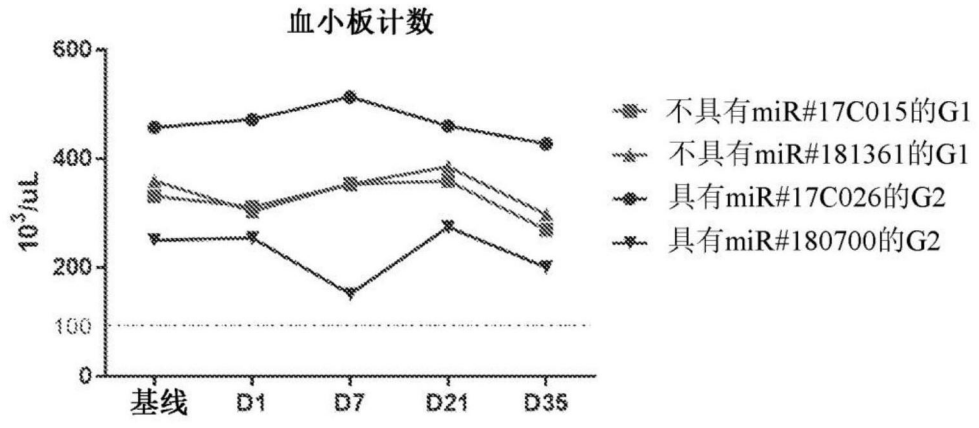


图22C

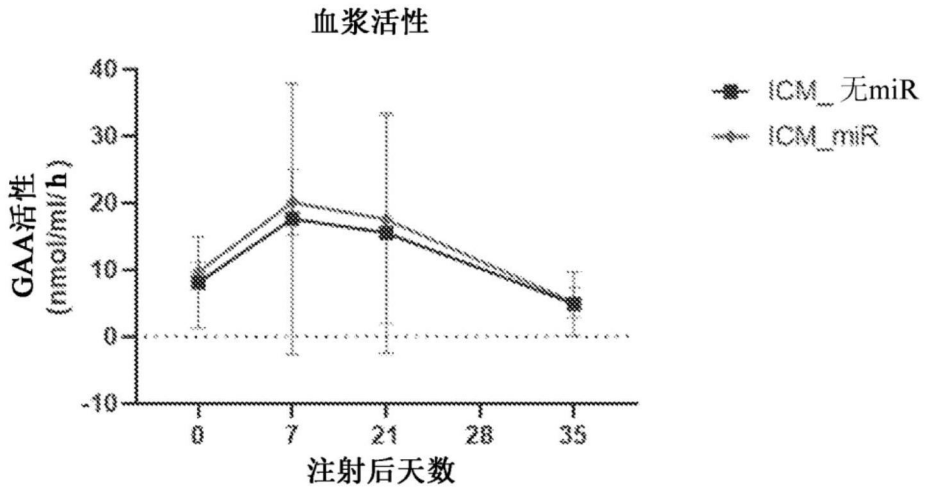

血浆活性
GAA活性
(nmol/ml/h)
40
30
20
10
0
-10
0
7
21
28
35
注射后天数
ICM_无miR
ICM_miR

图23

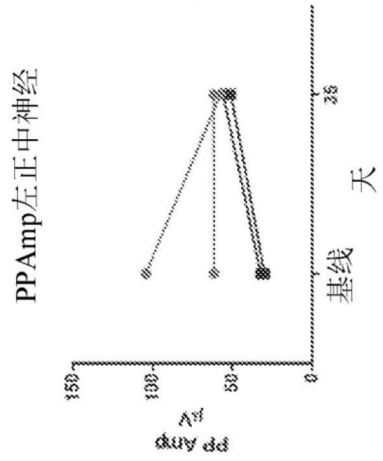

PPAmp左正中神经
基线
天
PP Amp
μV
150
100
50
0
35

图24A

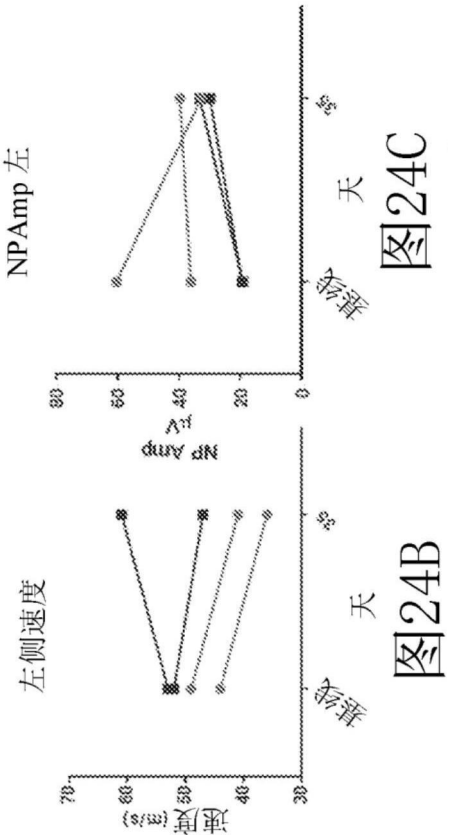


图24C

图24B

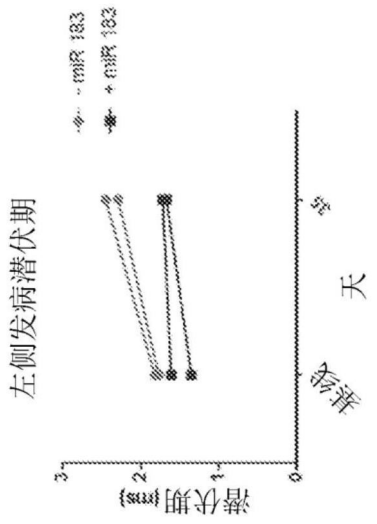
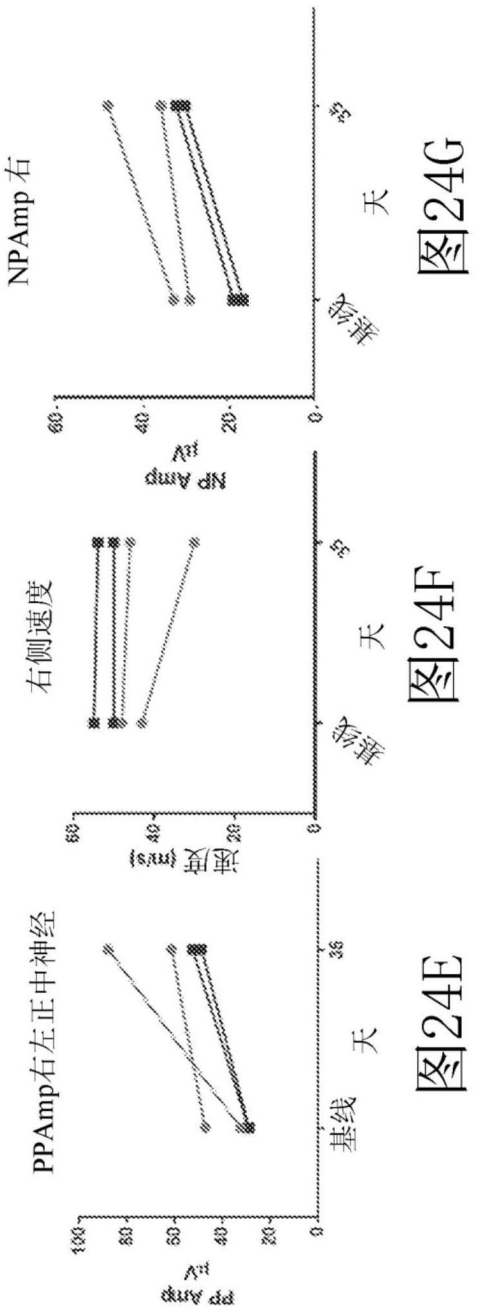


图24D


PPAmp右左正中神经
PP Amp μV
100
80
60
40
20
0
基线
天
36
图24E
右侧速度
速度 (m/s)
80
60
40
20
0
基线
天
56
图24F
NPAmp 右
NP Amp μV
60
40
20
0
基线
天
56
图24G

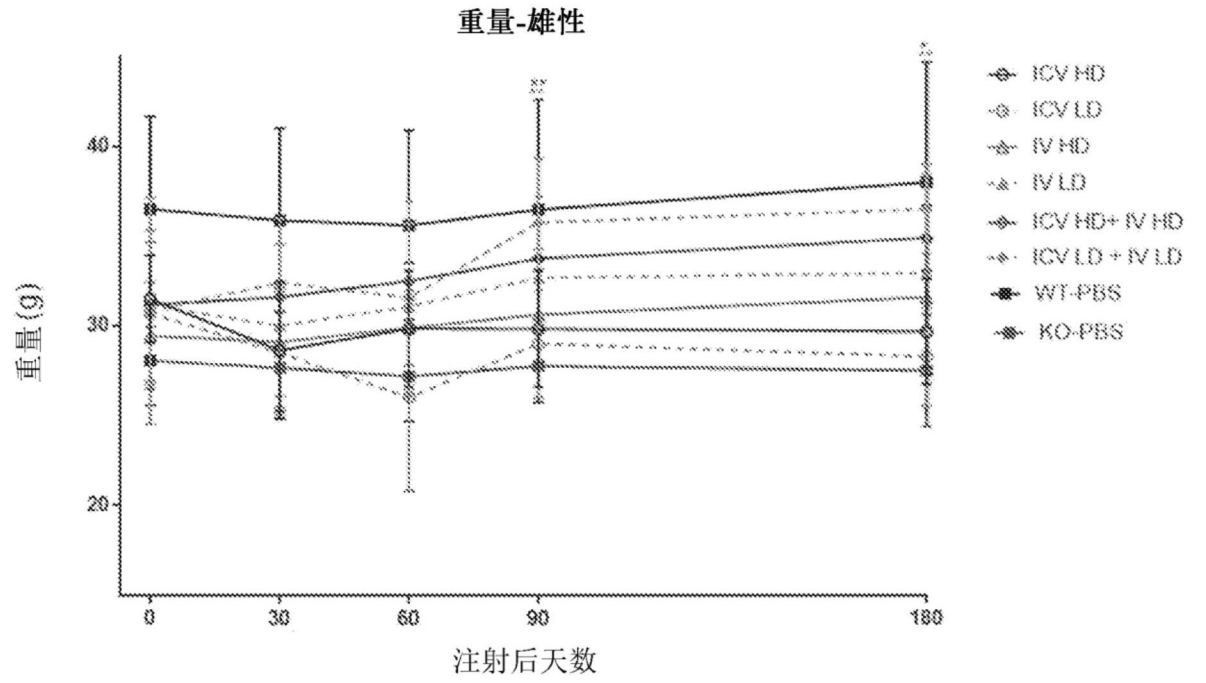

重量-雄性
重量(g)
40
30
20
0
30
60
90
180
注射后天数
ICV HD
ICV LD
IV HD
IV LD
ICV HD+ IV HD
ICV LD + IV LD
WT-PBS
KO-PBS

图25A

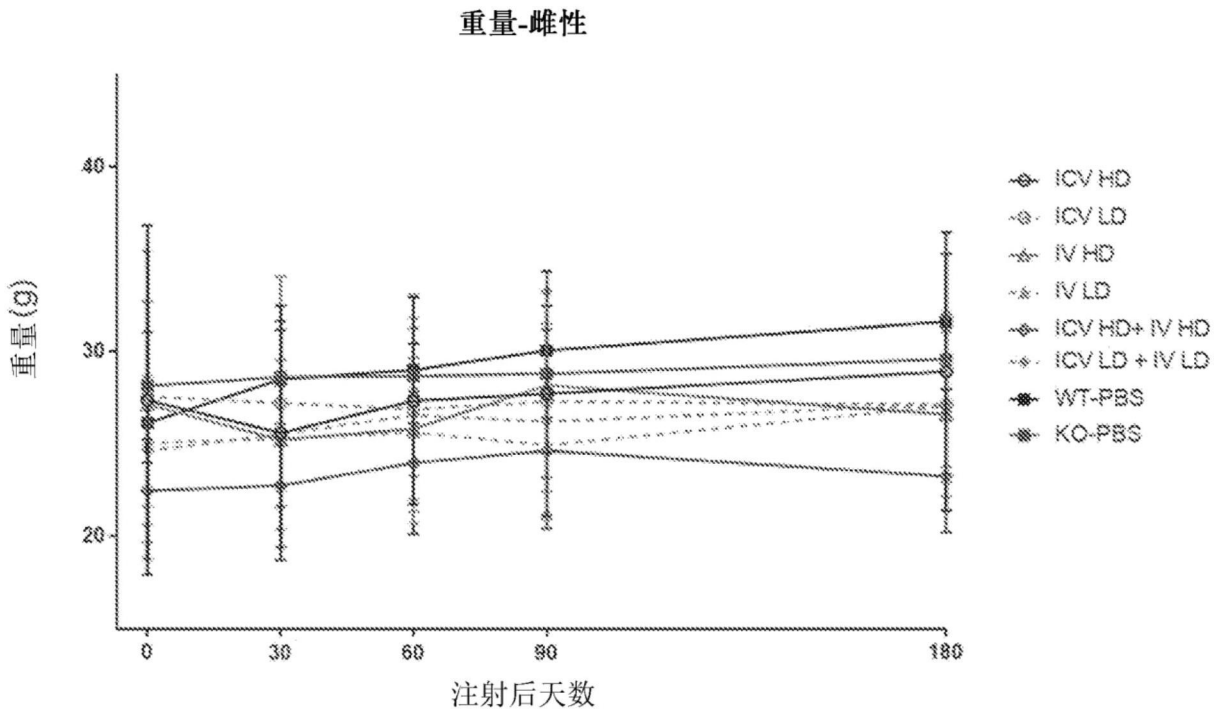

重量-雌性
重量(g)
40
30
20
0
30
60
90
180
注射后天数
ICV HD
ICV LD
IV HD
IV LD
ICV HD+ IV HD
ICV LD + IV LD
WT-PBS
KO-PBS

图25B

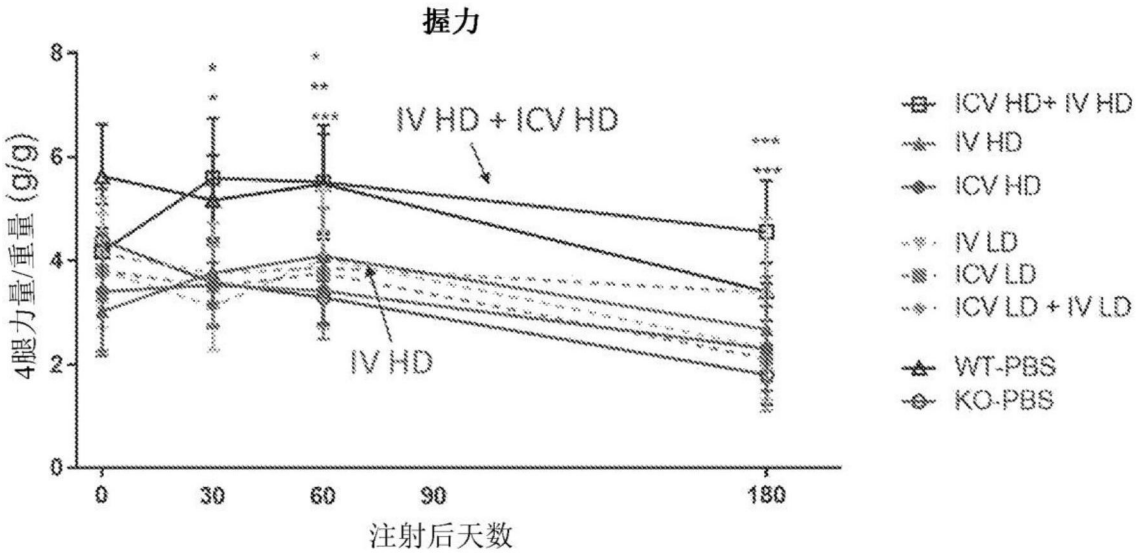

握力
4腿力量/重量 (g/g)
注射后天数
IV HD + ICV HD
IV HD
ICV HD+ IV HD
IV HD
ICV HD
IV LD
ICV LD
ICV LD + IV LD
WT-PBS
KO-PBS

图26

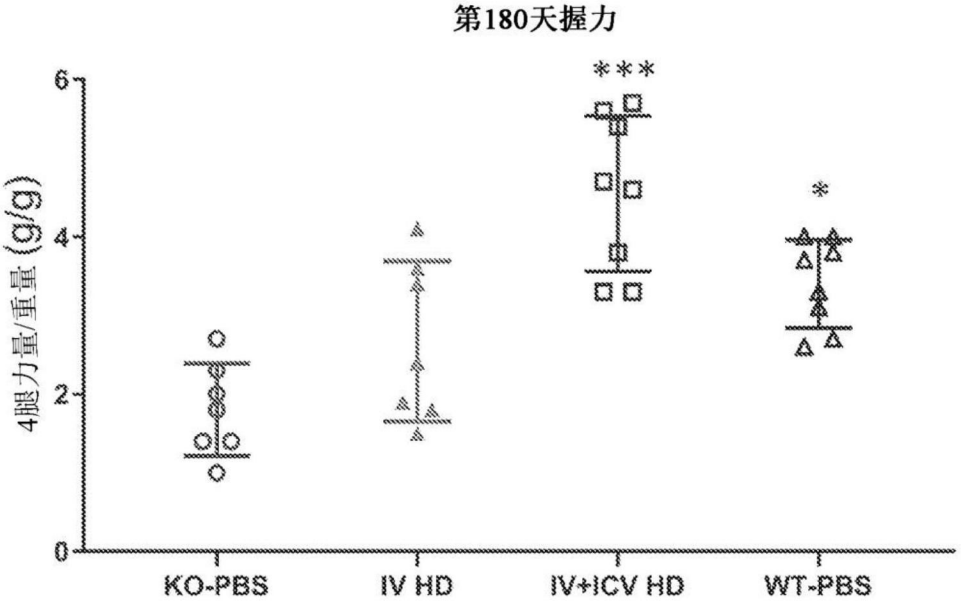

第180天握力
4腿力量/重量 (g/g)
KO-PBS
IV HD
IV+ICV HD
WT-PBS

图27

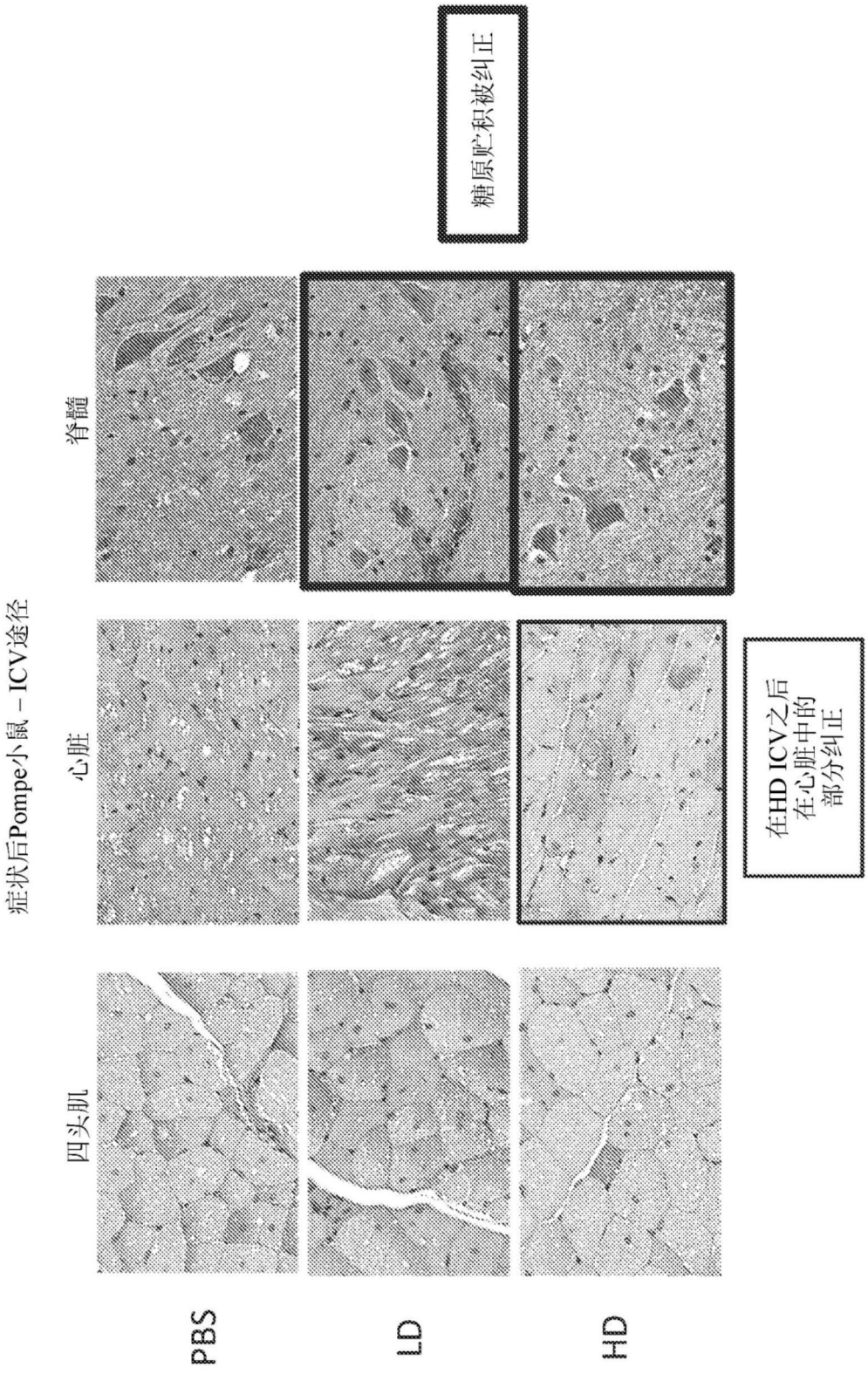

症状后Pompe小鼠 – ICV途径
四头肌
心脏
脊髓
糖原贮积被纠正
在HD ICV之后
在心脏中的
部分纠正
PBS
LD
HD

图28

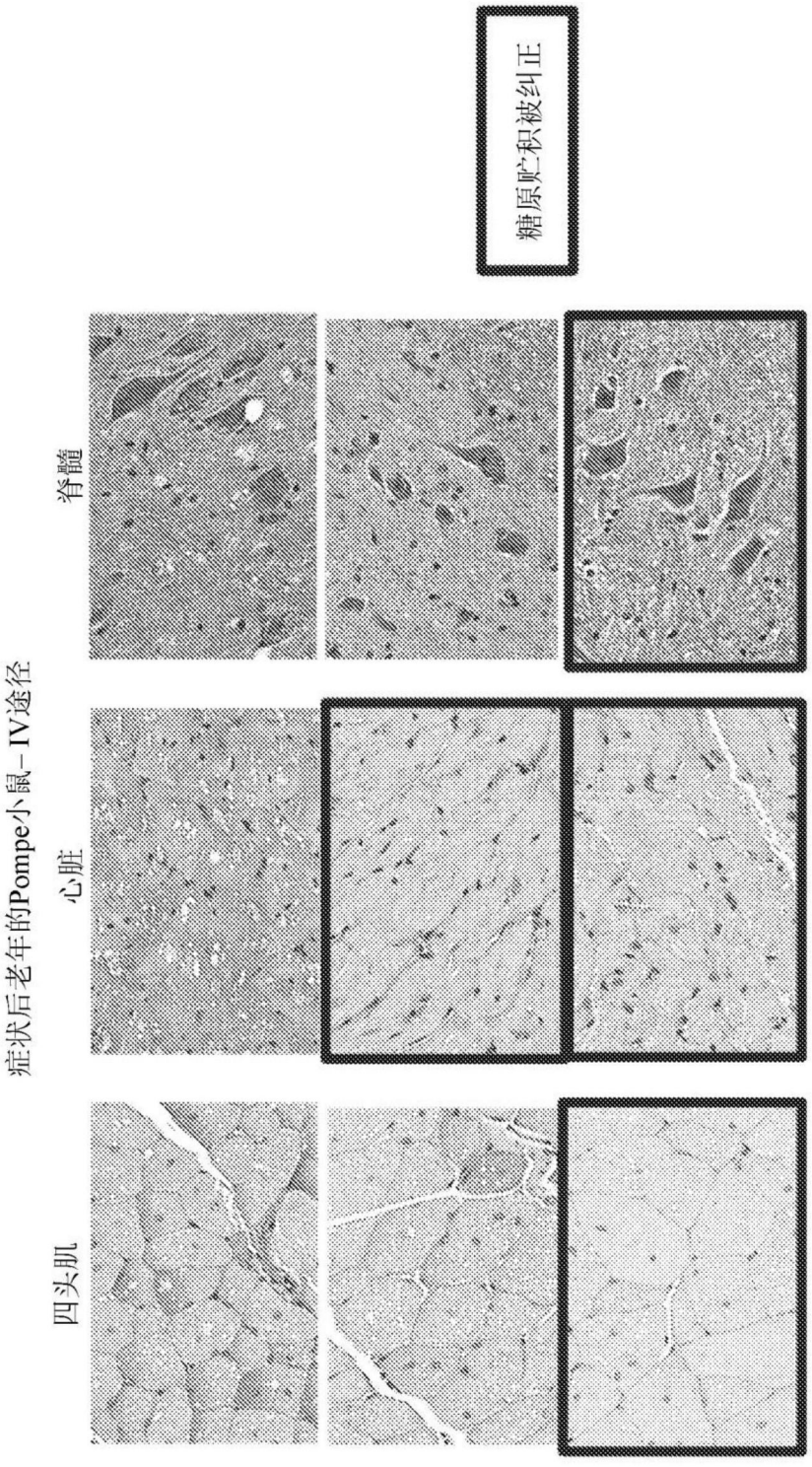


图29

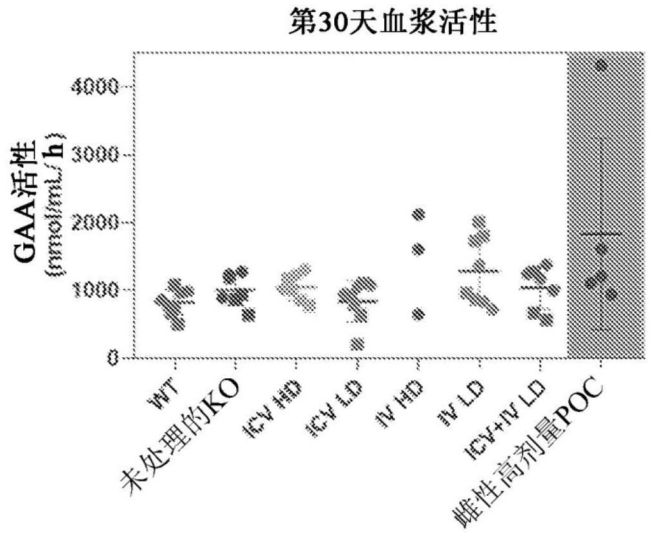

第30天血浆活性
GAA活性
(nmol/mL/h)
4000
3000
2000
1000
0
WT
未处理的KO
ICV HD
ICV LD
IV HD
IV LD
ICV+IV LD
雌性高剂量POC

图30A

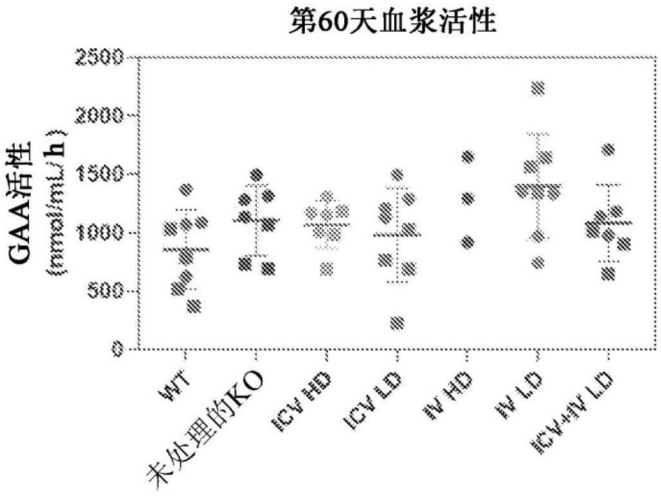

第60天血浆活性
GAA活性
(nmol/mL/h)
2500
2000
1500
1000
500
0
WT
未处理的KO
ICV HD
ICV LD
IV HD
IV LD
ICV+IV LD

图30B

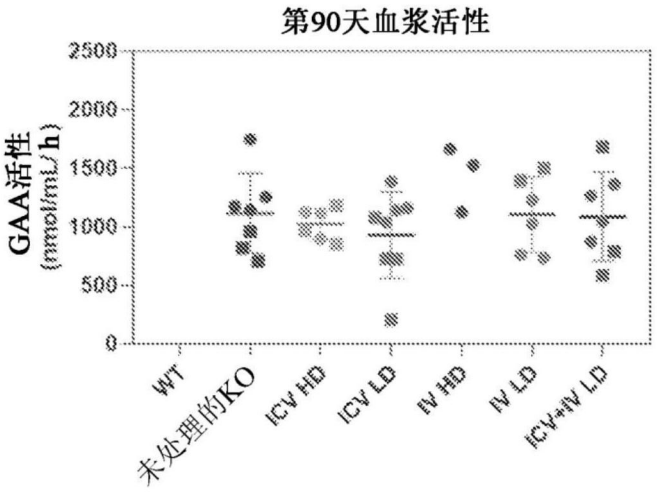

第90天血浆活性
GAA活性
(nmol/mL/h)
2500
2000
1500
1000
500
0
WT
未处理的KO
ICV HD
ICV LD
IV HD
IV LD
ICV+IV LD

图30C

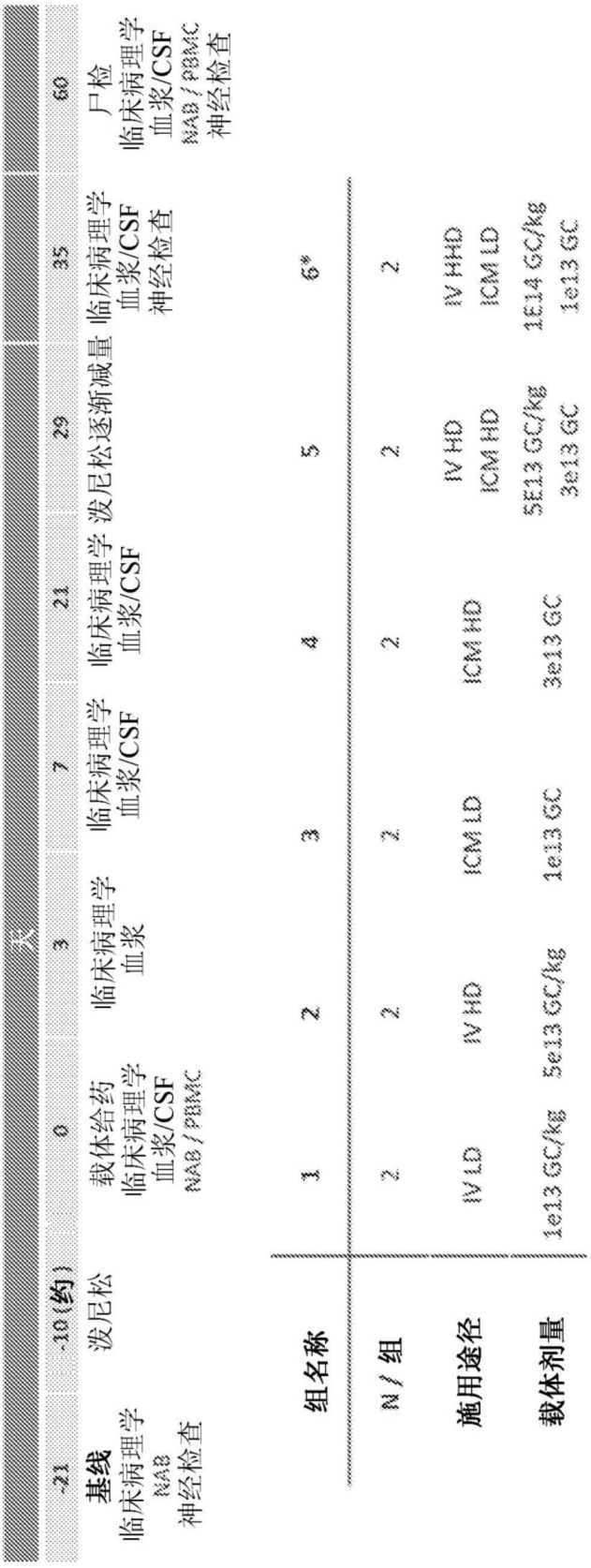

天
-21
-10(约)
0
3
7
21
29
35
60
基线
临床病理学
NAB
神经检查
泼尼松
载体给药
临床病理学
血浆/CSF
NAB / PBMC
临床病理学
血浆
临床病理学
血浆/CSF
临床病理学
血浆/CSF
泼尼松逐渐减量
临床病理学
血浆/CSF
神经检查
尸检
临床病理学
血浆/CSF
NAB / PBMC
神经检查
组名称
1
2
3
4
5
6*
N / 组
2
2
2
2
2
2
施用途径
IV LD
IV HD
ICM LD
ICM HD
IV HD
ICM HD
IV HHD
ICM LD
载体剂量
1e13 GC/kg
5e13 GC/kg
1e13 GC
3e13 GC
5E13 GC/kg
3e13 GC
1E14 GC/kg
1e13 GC

图31

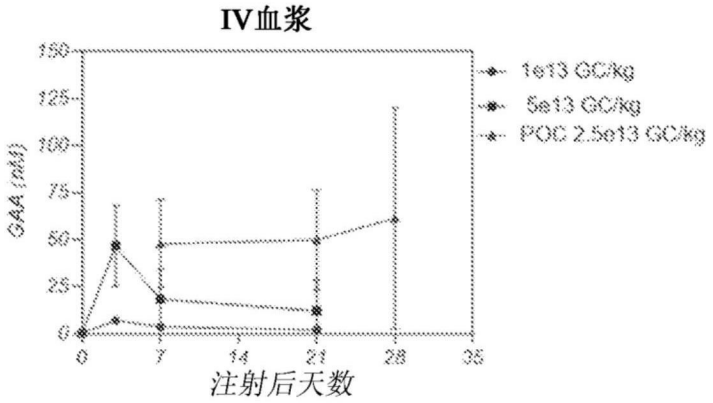


图32A

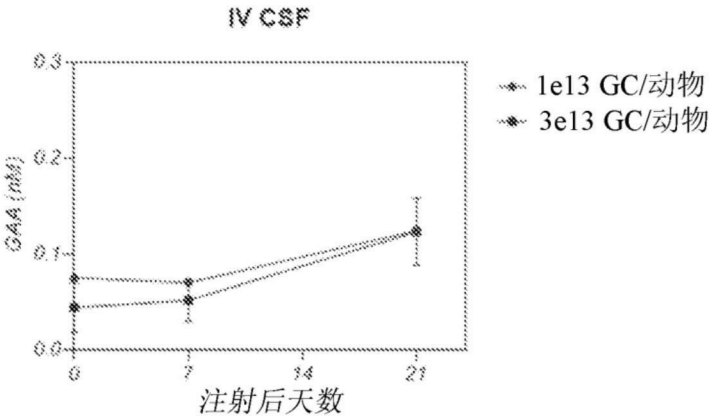


图32B

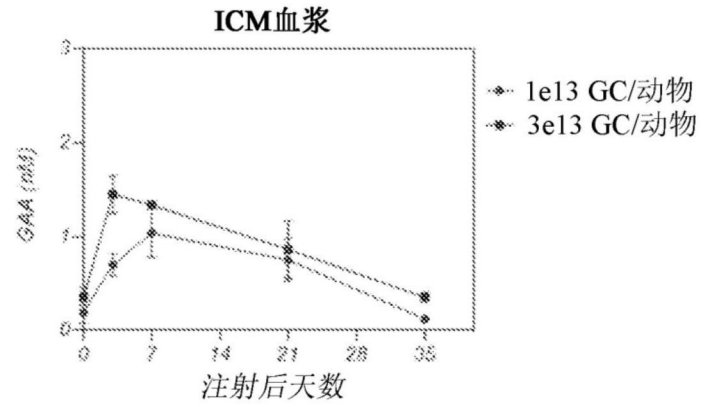


图32C

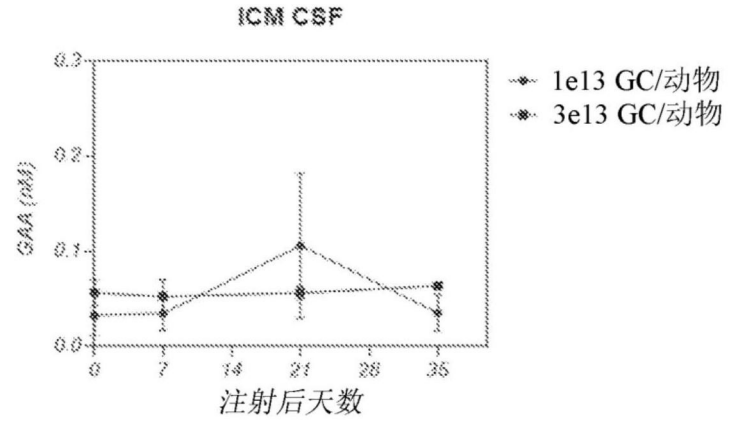


图32D

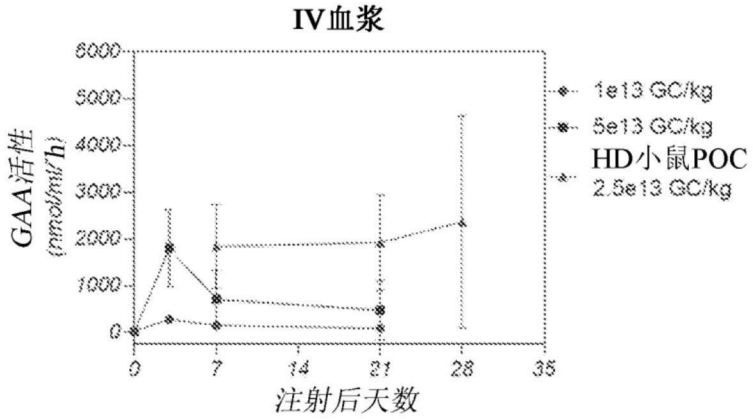


图32E

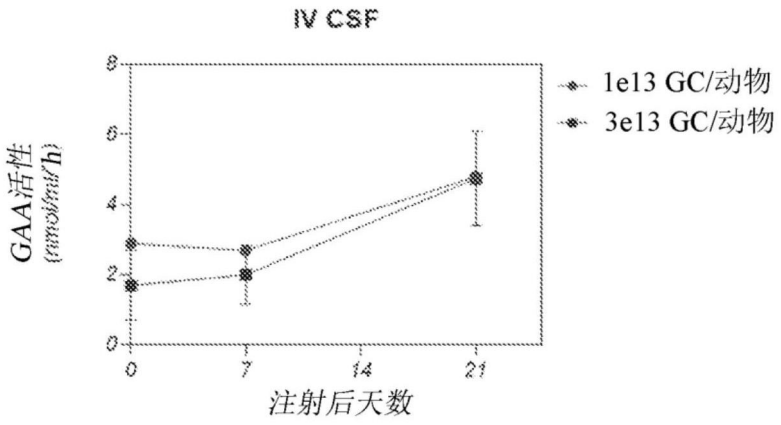


图32F

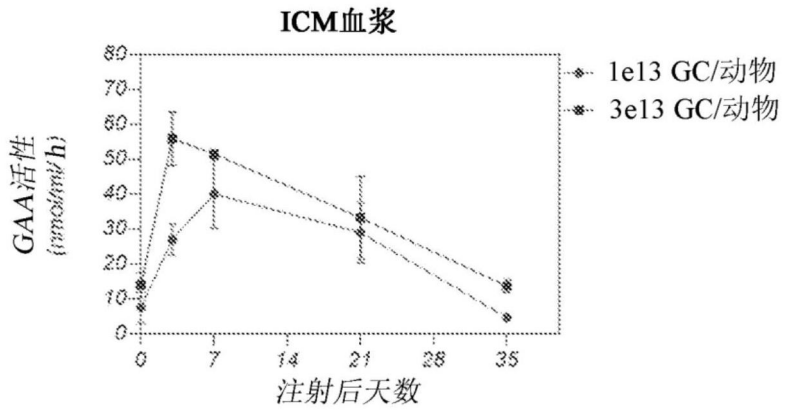


图32G

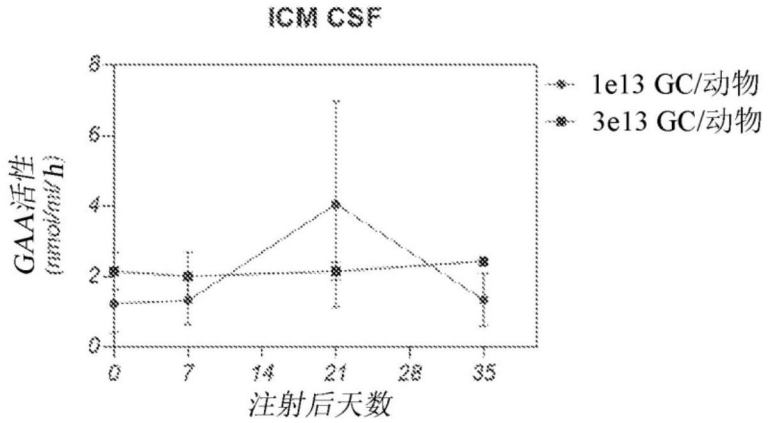


图32H

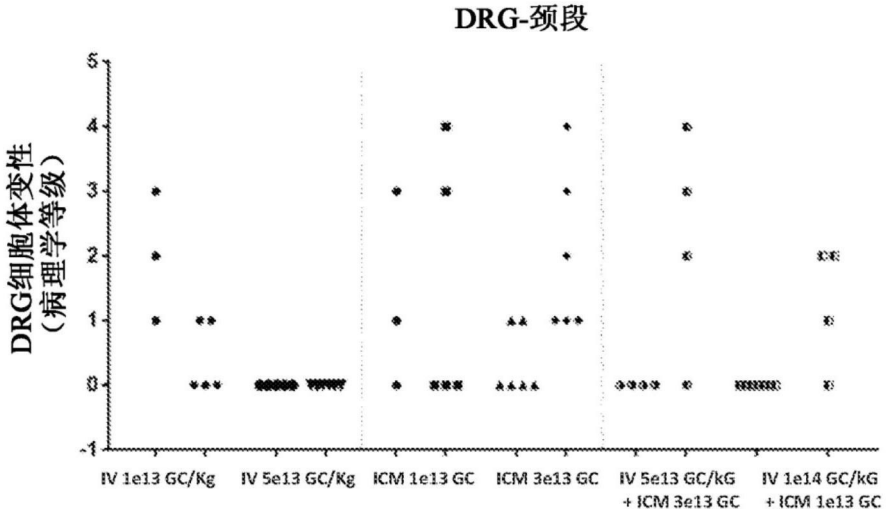


图33A

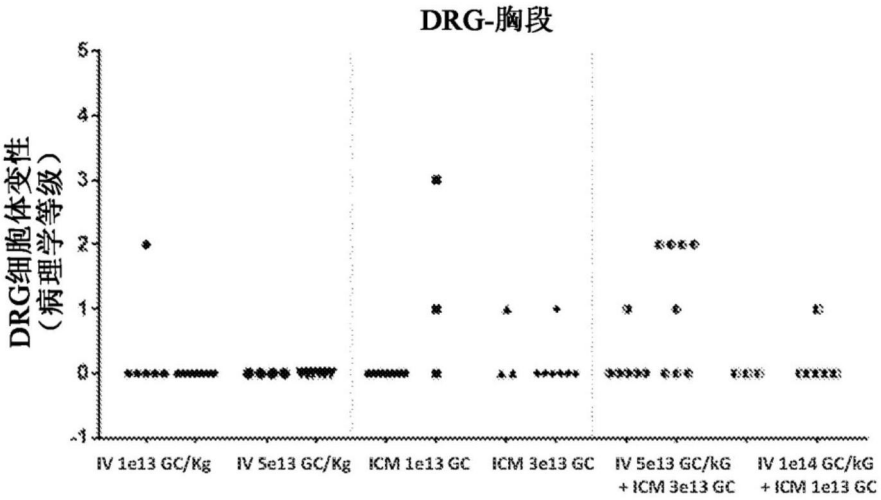

DRG-胸段
DRG细胞胞体变性（病理学等级）
IV 1e13 GC/Kg
IV 5e13 GC/Kg
ICM 1e13 GC
ICM 3e13 GC
IV 5e13 GC/kG + ICM 3e13 GC
IV 1e14 GC/kG + ICM 1e13 GC

图33B

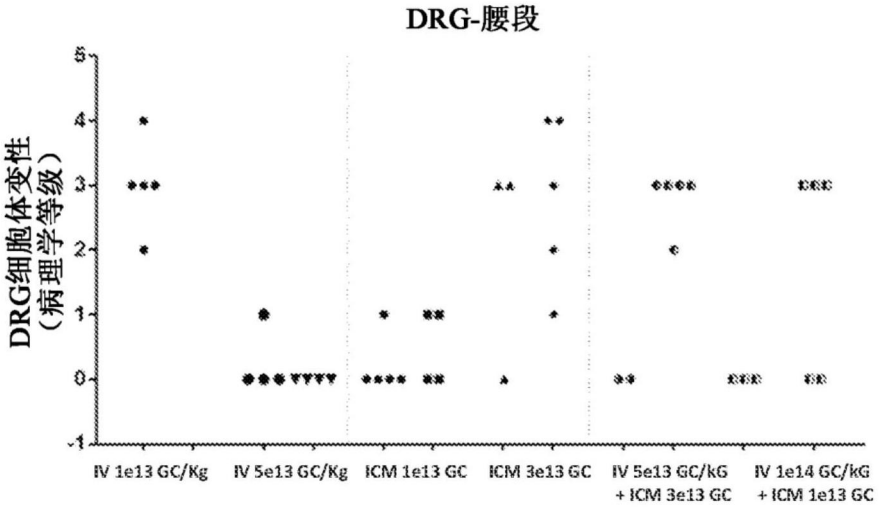

DRG-腰段
DRG细胞胞体变性（病理学等级）
IV 1e13 GC/Kg
IV 5e13 GC/Kg
ICM 1e13 GC
ICM 3e13 GC
IV 5e13 GC/kG + ICM 3e13 GC
IV 1e14 GC/kG + ICM 1e13 GC

图33C

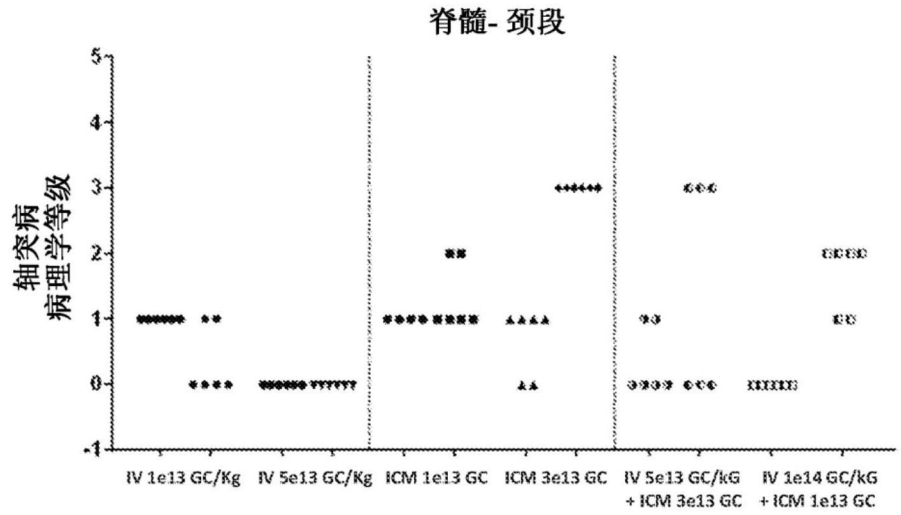

脊髓- 颈段
轴突病
病理学等级
IV 1e13 GC/Kg
IV 5e13 GC/Kg
ICM 1e13 GC
ICM 3e13 GC
IV 5e13 GC/kG
+ ICM 3e13 GC
IV 1e14 GC/kG
+ ICM 1e13 GC

图33D

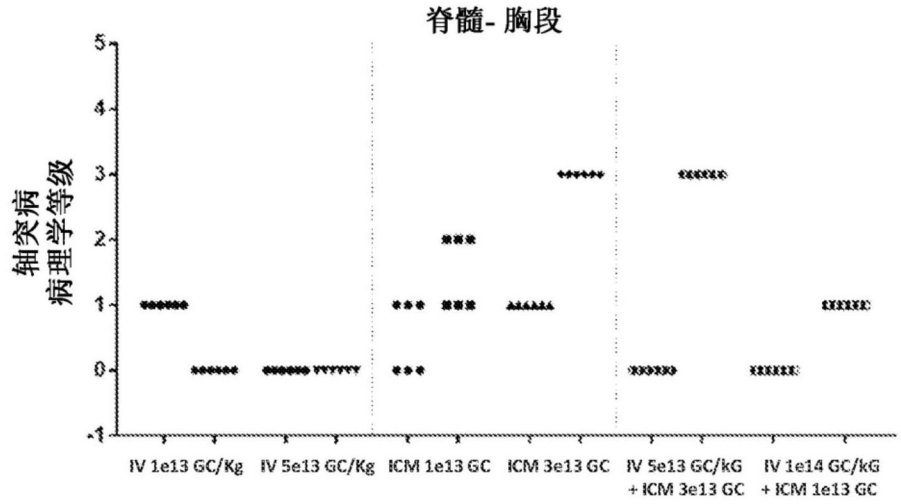

脊髓- 胸段
轴突病
病理学等级
IV 1e13 GC/Kg
IV 5e13 GC/Kg
ICM 1e13 GC
ICM 3e13 GC
IV 5e13 GC/kG
+ ICM 3e13 GC
IV 1e14 GC/kG
+ ICM 1e13 GC

图33E

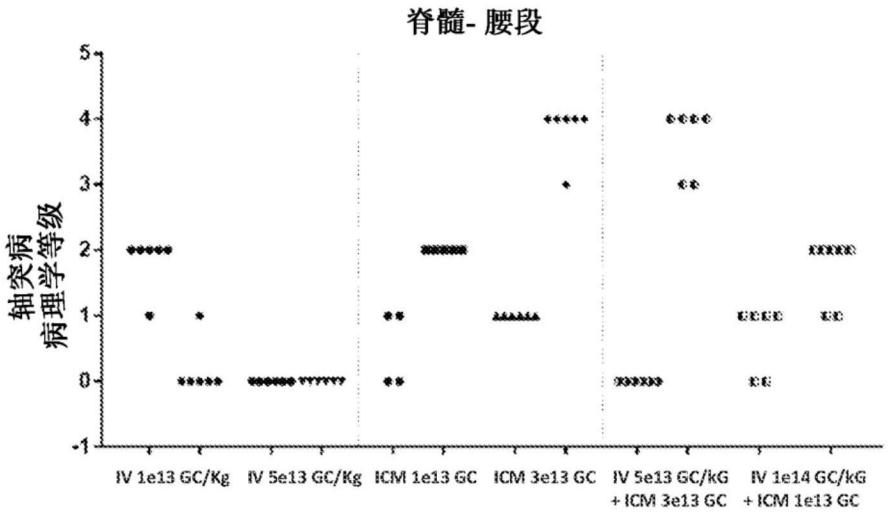

脊髓- 腰段
轴突病
病理学等级
5
4
3
2
1
0
-1
IV 1e13 GC/Kg
IV 5e13 GC/Kg
ICM 1e13 GC
ICM 3e13 GC
IV 5e13 GC/kG
+ ICM 3e13 GC
IV 1e14 GC/kG
+ ICM 1e13 GC

图33F

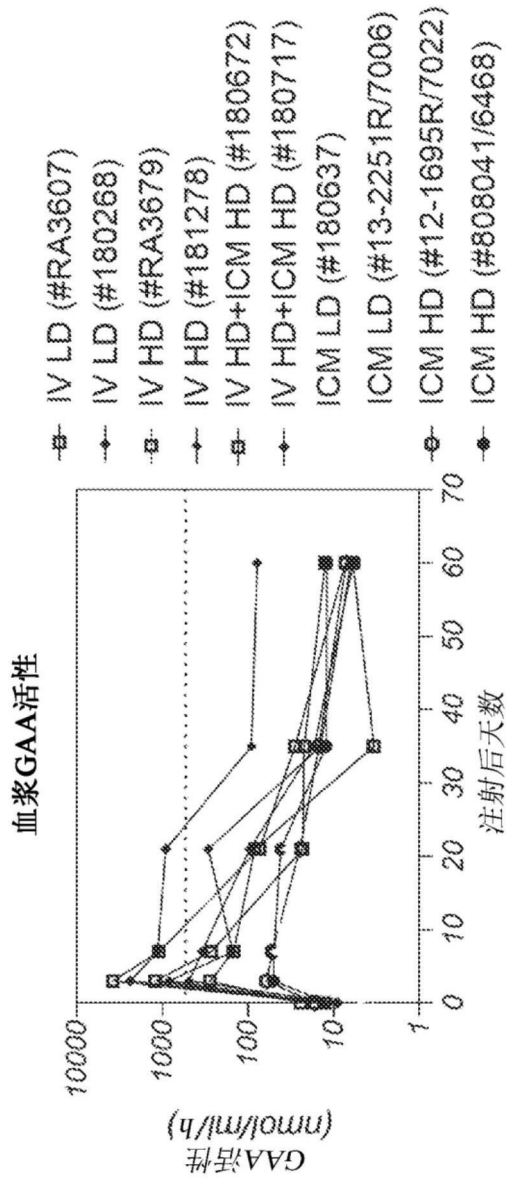

血浆GAA活性
GAA活性
(nmol/ml/h)
10000
1000
100
10
1
0
10
20
30
40
50
60
70
注射后天数
IV LD (#RA3607)
IV LD (#180268)
IV HD (#RA3679)
IV HD (#181278)
IV HD+ICM HD (#180672)
IV HD+ICM HD (#180717)
ICM LD (#180637)
ICM LD (#13-2251R/7006)
ICM HD (#12-1695R/7022)
ICM HD (#808041/6468)

图34

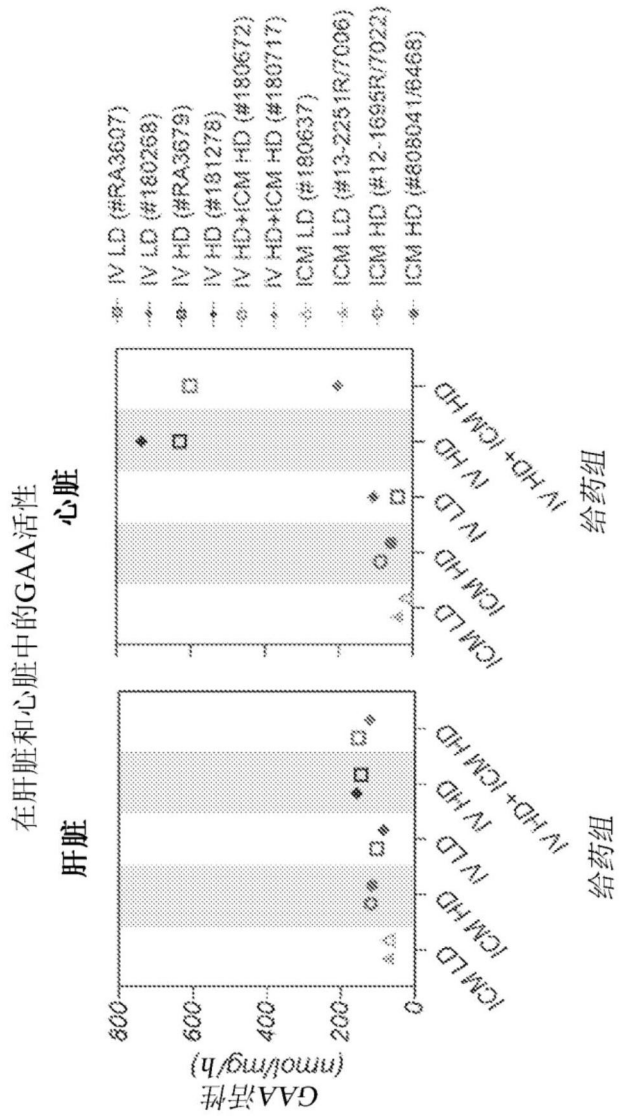

在肝脏和心脏中的GAA活性
心脏
肝脏
GAA活性
(nmol/mg/h)
600
400
200
0
ICM LD
ICM HD
IV LD
IV HD
IV HD+ICM HD
给药组
IV LD (#RA3607)
IV LD (#180258)
IV HD (#RA3679)
IV HD (#181278)
IV HD+ICM HD (#180672)
IV HD+ICM HD (#180717)
ICM LD (#180637)
ICM LD (#13-2251R/7006)
ICM HD (#12-1695R/7022)
ICM HD (#808041/6468)

图35A

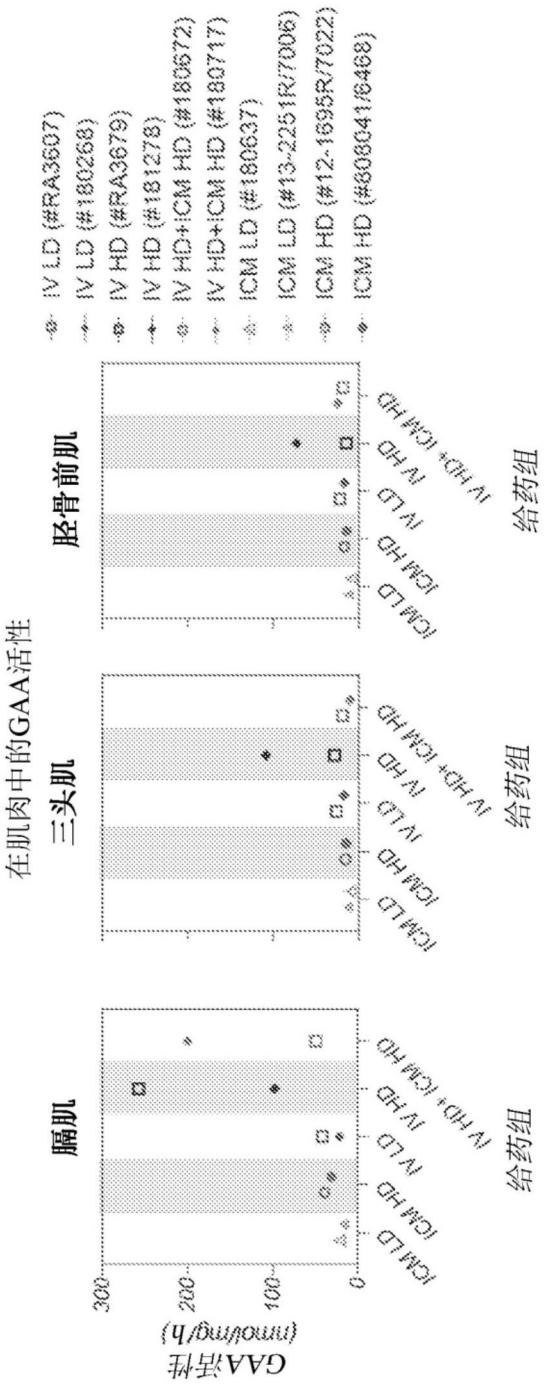

在肌肉中的GAA活性
膈肌
三头肌
胫骨前肌
GAA活性
(nmol/mg/h)
300
200
100
0
ICM LD
ICM HD
IV LD
IV HD
IV HD+ ICM HD
给药组
IV LD (#RA3607)
IV LD (#180268)
IV HD (#RA3679)
IV HD (#181278)
IV HD+ICM HD (#180672)
IV HD+ICM HD (#180717)
ICM LD (#180637)
ICM LD (#13-2251R/7006)
ICM HD (#12-1695R/7022)
ICM HD (#6080416468)

图35B

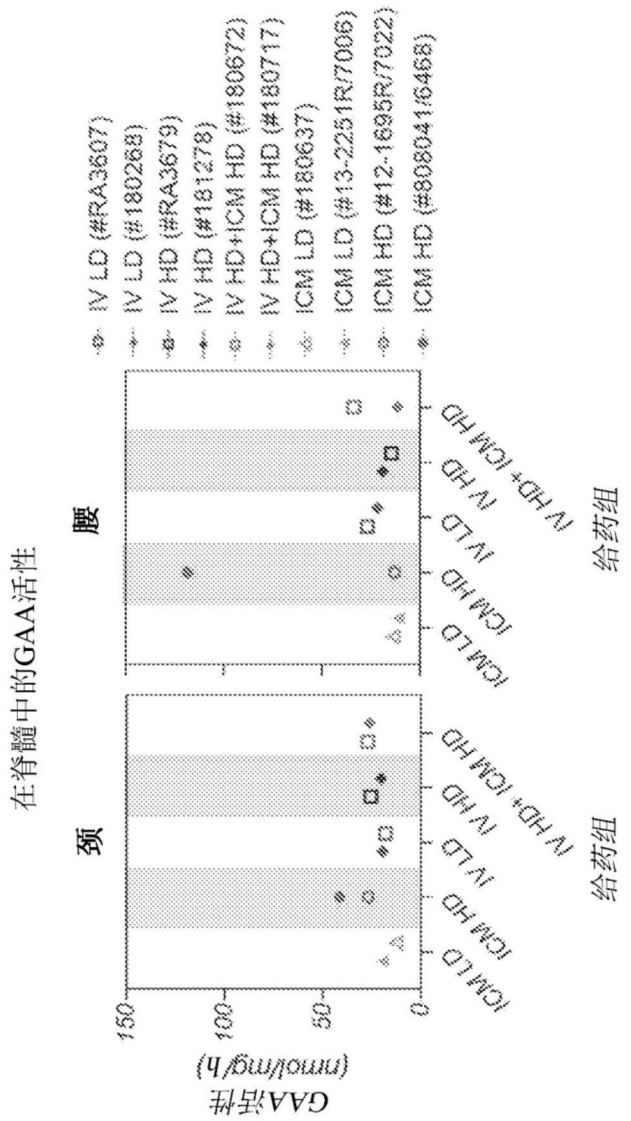

在脊髓中的GAA活性
颈
腰
GAA活性
(nmol/mg/h)
0
50
100
150
ICM LD
ICM HD
IV LD
IV HD
IV HD+ ICM HD
给药组
IV LD (#RA3607)
IV LD (#180268)
IV HD (#RA3679)
IV HD (#181278)
IV HD+ICM HD (#180672)
IV HD+ICM HD (#180717)
ICM LD (#180637)
ICM LD (#13-2251R/7006)
ICM HD (#12-1695R/7022)
ICM HD (#8080411/6468)

图35C

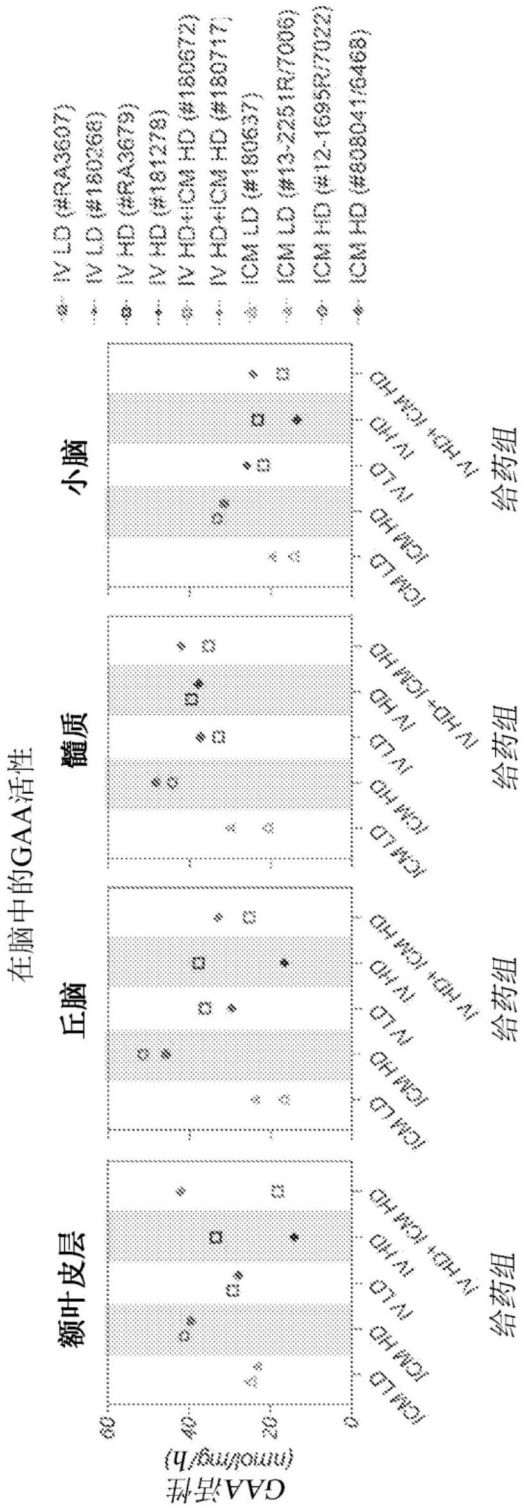

在脑中的GAA活性
额叶皮层
丘脑
髓质
小脑
GAA活性
(nmol/mg/h)
0
20
40
60
ICM LD
ICM HD
IV LD
IV HD
IV HD+ ICM HD
给药组
IV LD (#RA3607)
IV LD (#180368)
IV HD (#RA3679)
IV HD (#181278)
IV HD+ICM HD (#180672)
IV HD+ICM HD (#180717)
ICM LD (#180637)
ICM LD (#13-2251R/7006)
ICM HD (#12-1695R/7022)
ICM HD (#8080411/64468)

图35D

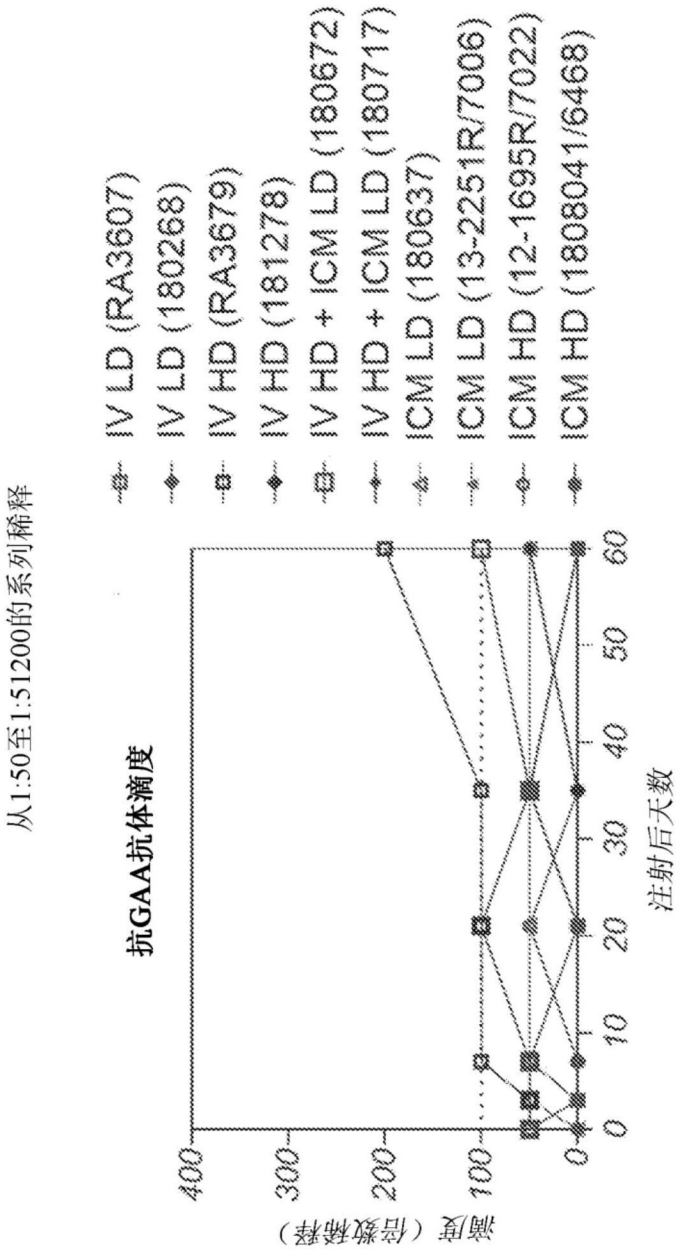

从1:50至1:51200的系列稀释
抗GAA抗体滴度
滴度（倍数稀释）
注射后天数
0
100
200
300
400
0
10
20
30
40
50
60
IV LD (RA3607)
IV LD (180268)
IV HD (RA3679)
IV HD (181278)
IV HD + ICM LD (180672)
IV HD + ICM LD (180717)
ICM LD (180637)
ICM LD (13-2251R/7006)
ICM HD (12-1695R/7022)
ICM HD (1808041/6468)

图36

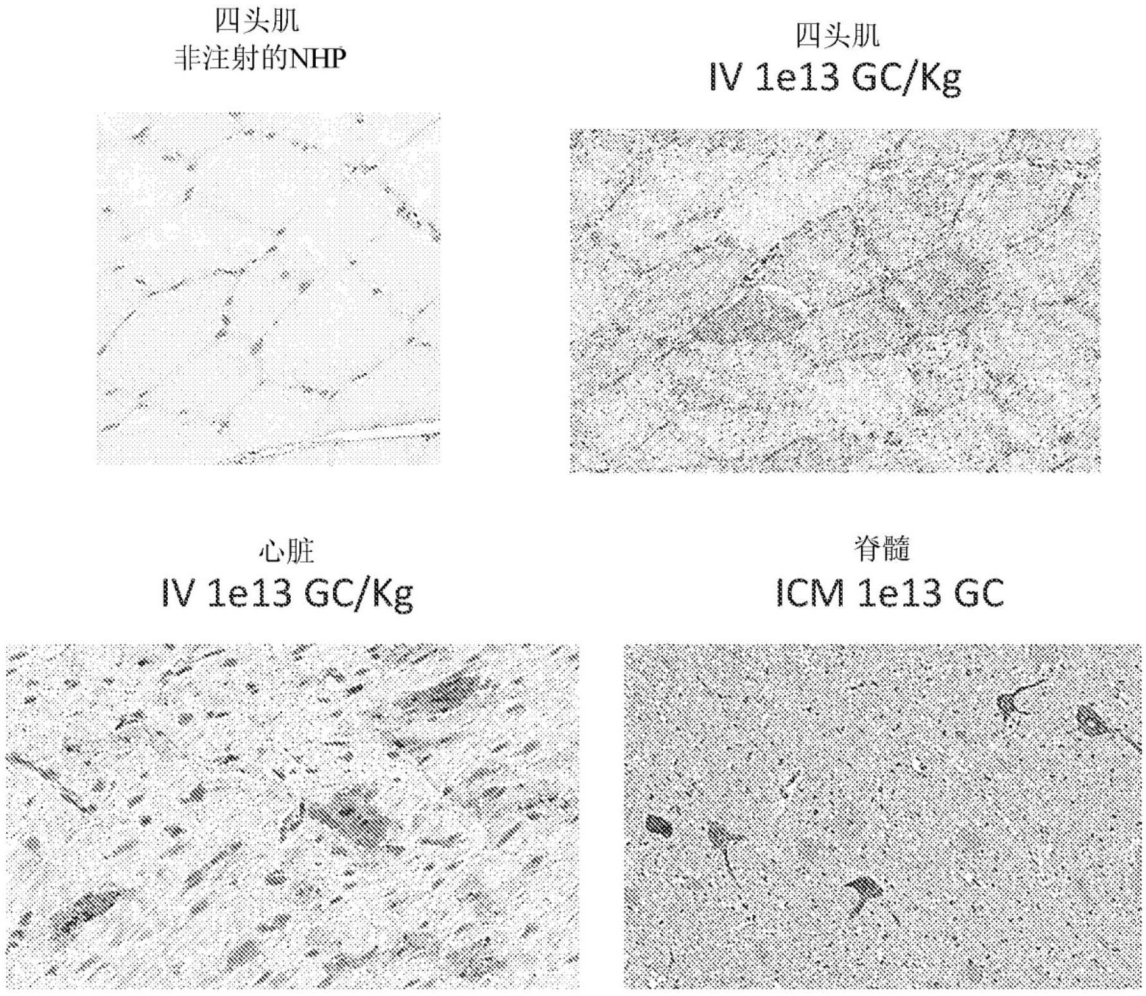


图37

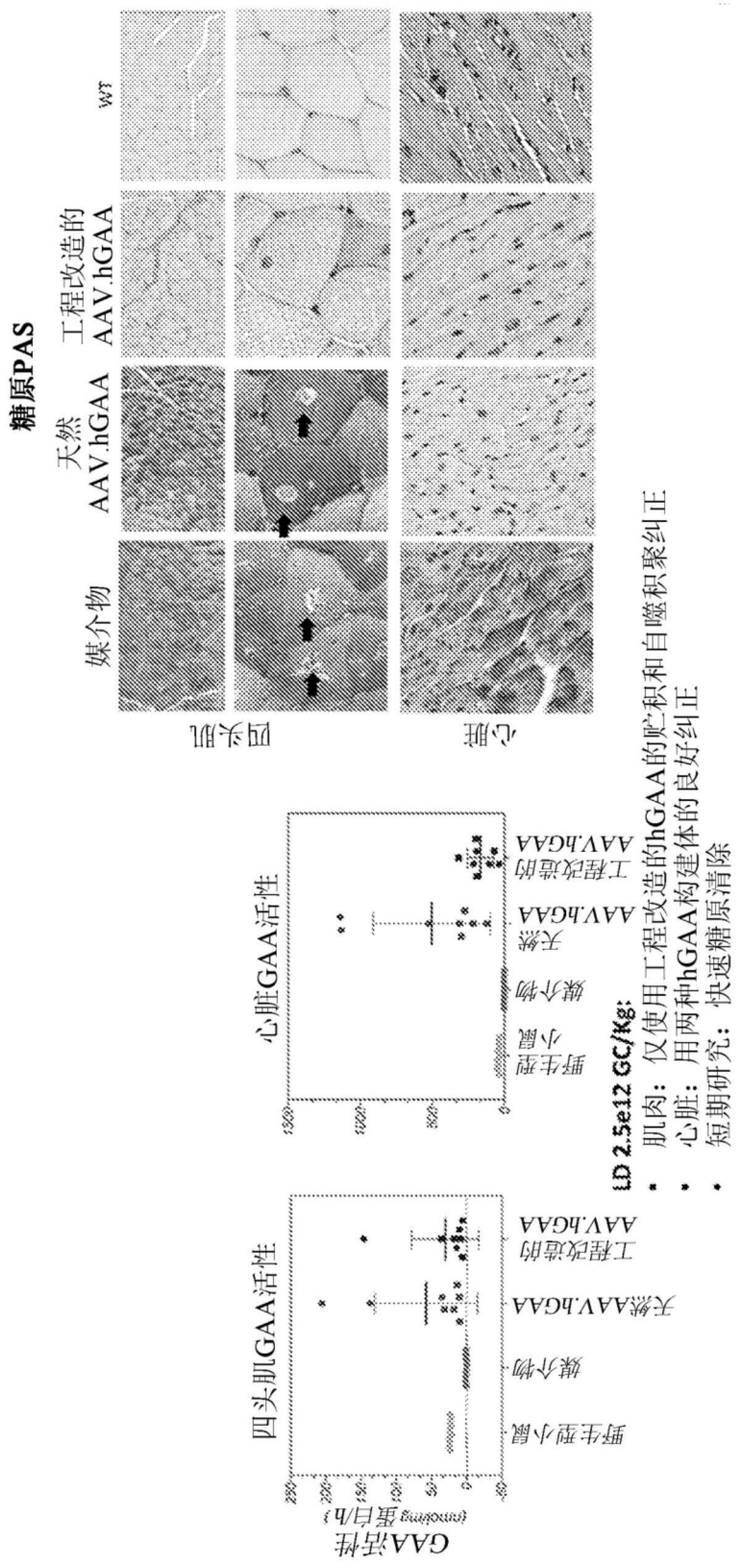

糖原PAS
媒介物
天然 AAV.hGAA
工程改造的 AAV.hGAA
WT
四头肌
心脏
GAA活性 (nmol/mg 蛋白/h)
四头肌GAA活性
心脏GAA活性
野生型小鼠
媒介物
天然AAV.hGAA
工程改造的 AAV.hGAA
LD 2.5e12 GC/Kg:
肌肉：仅使用工程改造的hGAA的贮积和自噬积聚纠正
心脏：用两种hGAA构建体的良好纠正
短期研究：快速糖原清除

图38

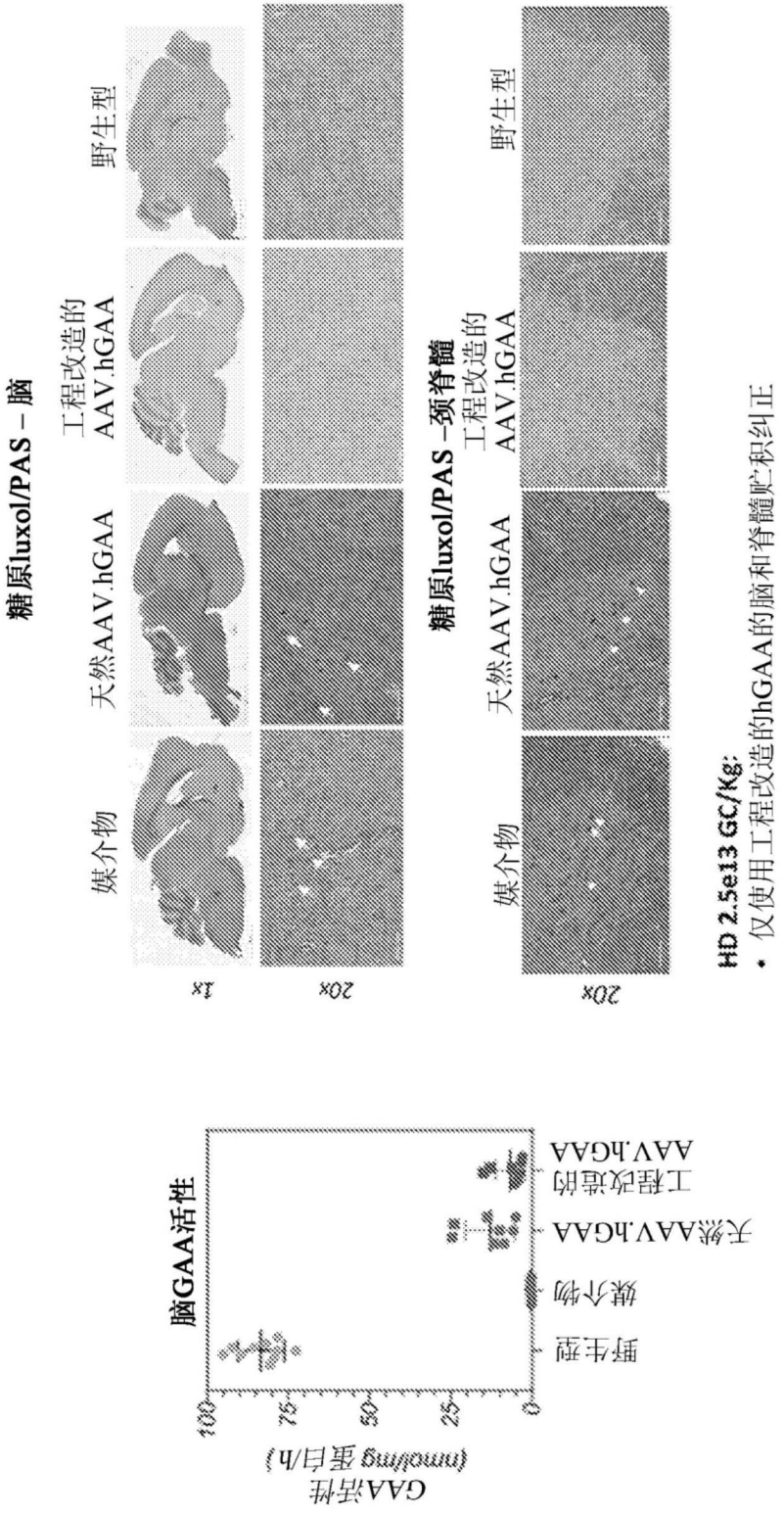

糖原luxol/PAS – 脑
媒介物
天然AAV.hGAA
工程改造的
AAV.hGAA
野生型
1x
20x
糖原luxol/PAS –颈脊髓
媒介物
天然AAV.hGAA
工程改造的
AAV.hGAA
野生型
20x
HD 2.5e13 GC/Kg:
• 仅使用工程改造的hGAA的脑和脊髓贮积纠正
脑GAA活性
GAA活性
(nmol/mg 蛋白/h)
100
75
50
25
0
野生型
媒介物
天然AAV.hGAA
工程改造的
AAV.hGAA

图39

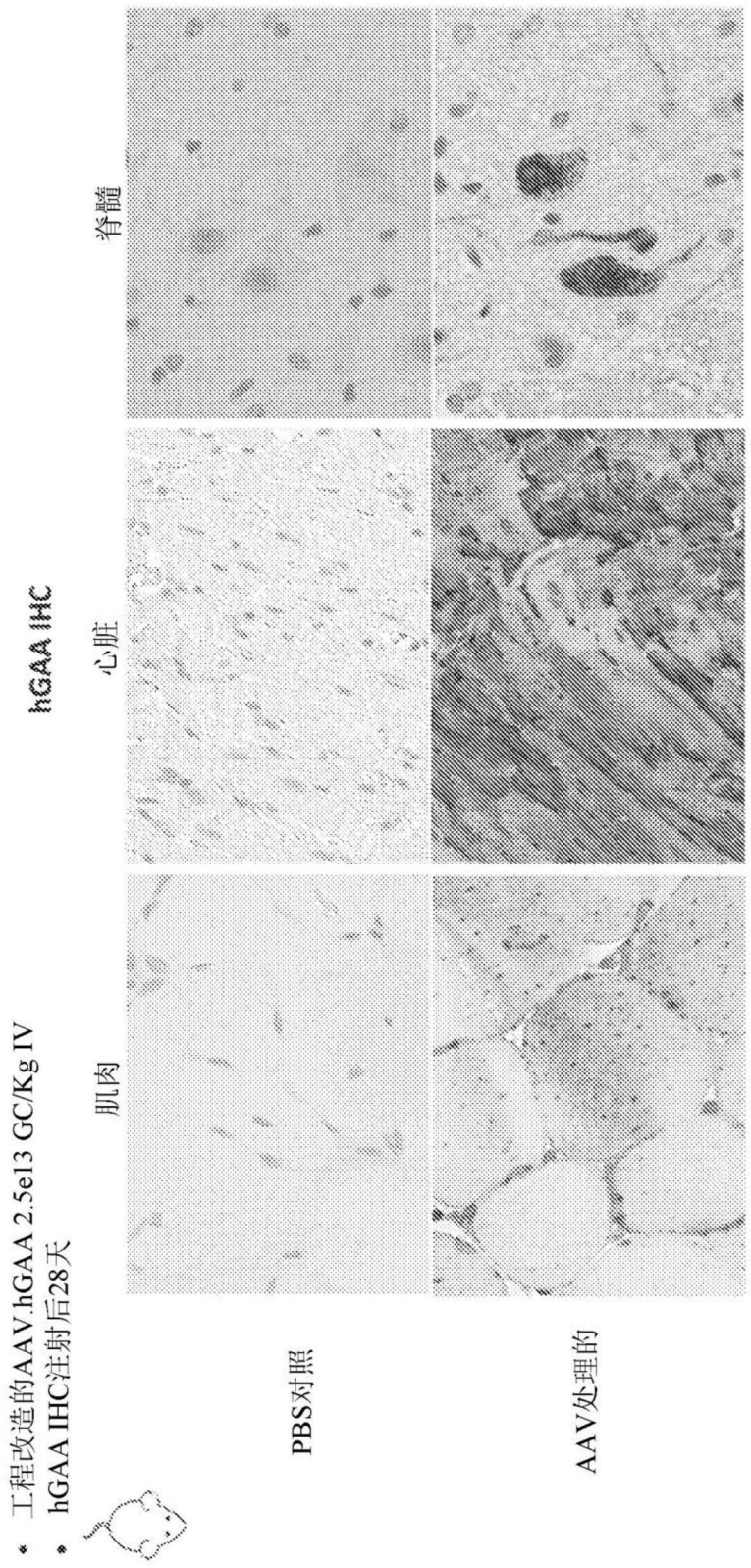

工程改造的AAV.hGAA 2.5e13 GC/Kg IV
hGAA IHC注射后28天
hGAA IHC
肌肉
心脏
脊髓
PBS对照
AAV处理的

图40

Pompe小鼠1-1.5个月大
工程改造的AAV.hGAA
尸检时的年龄
3-4.5个月
IV第0天
PBS
1e11 GC (5e12 GC/Kg); n = 4M & 4F
2e11 GC (1e13 GC/Kg); n = 4M & 4F
5e11 GC (2.5e13 GC/Kg); n = 4M & 4F
1e12 GC (5e13 GC/Kg); n = 4M & 4F
5e12 GC (2.5e14 GC/Kg); n = 4M & 4F
第60天
安乐死
组织学

图41

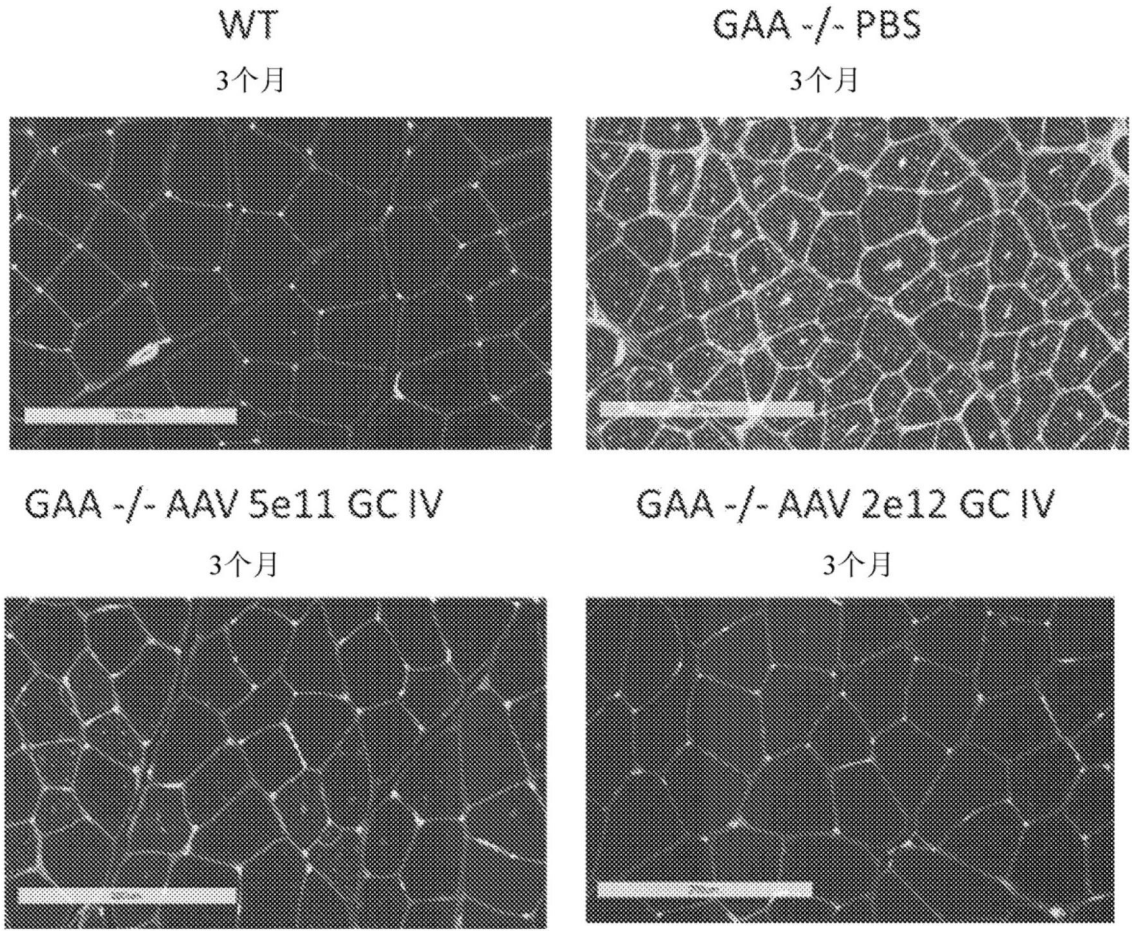

WT
3个月
GAA -/- PBS
3个月
GAA -/- AAV 5e11 GC IV
3个月
GAA -/- AAV 2e12 GC IV
3个月

图42

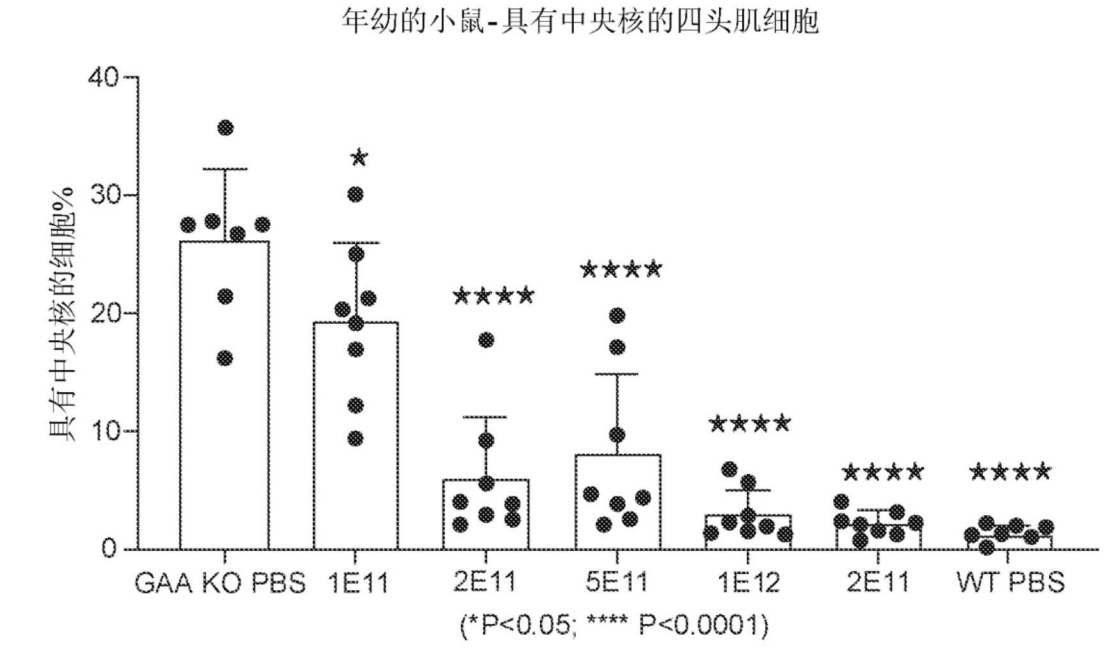

年幼的小鼠-具有中央核的四头肌细胞
具有中央核的细胞%
40
30
20
10
0
GAA KO PBS
1E11
2E11
5E11
1E12
2E11
WT PBS
(*P<0.05; **** P<0.0001)

图43

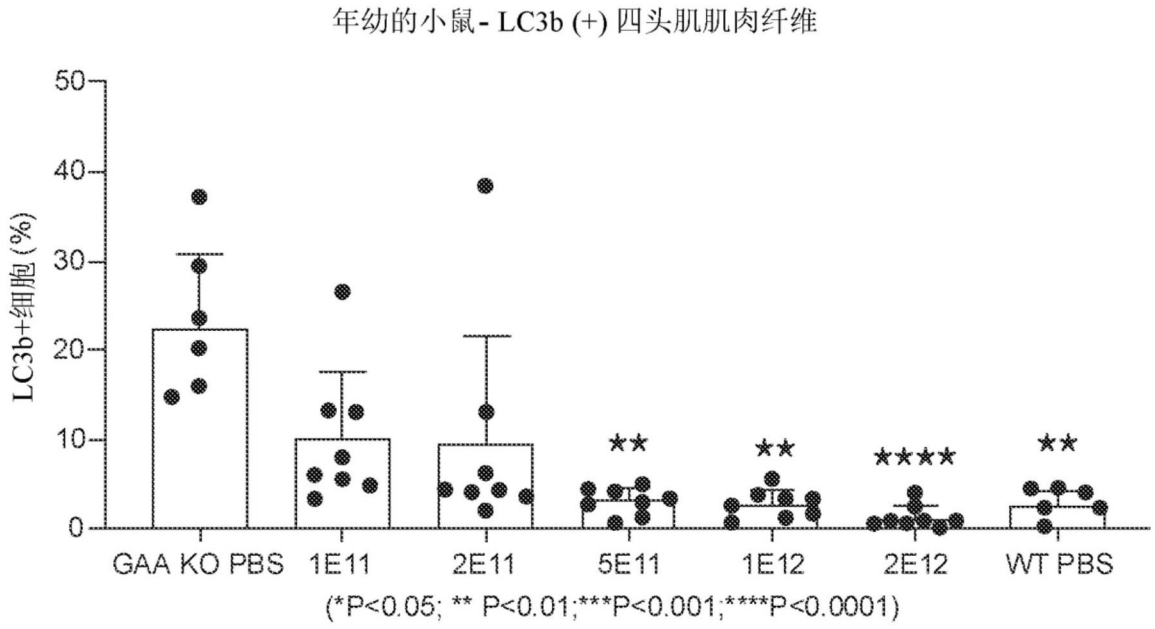

年幼的小鼠- LC3b (+) 四头肌肌肉纤维
LC3b+细胞 (%)
50
40
30
20
10
0
GAA KO PBS
1E11
2E11
5E11
1E12
2E12
WT PBS
(*P<0.05; ** P<0.01;***P<0.001;****P<0.0001)

图44

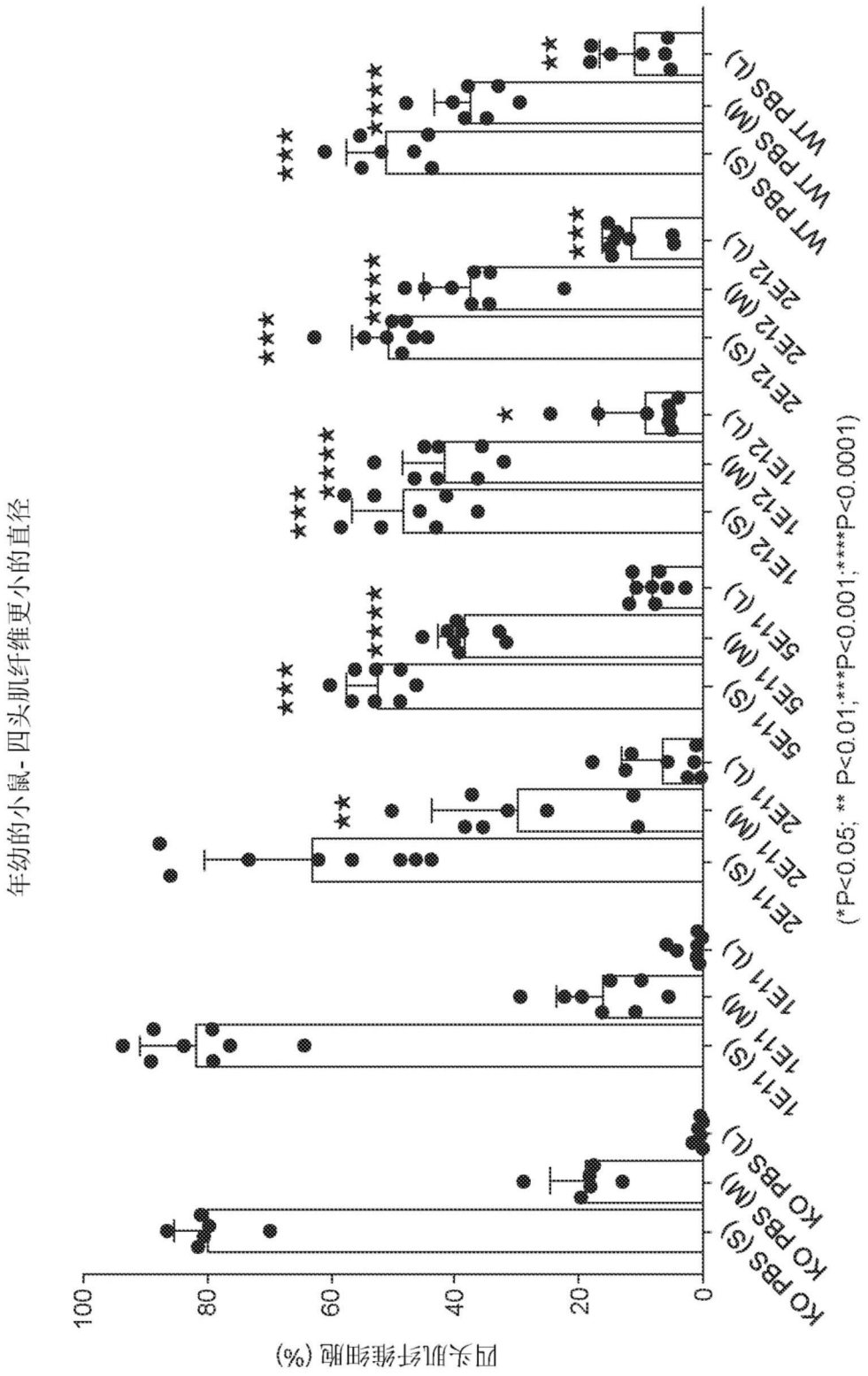

年幼的小鼠- 四头肌纤维更小的直径
四头肌纤维细胞 (%)
100
80
60
40
20
0
KO PBS (S)
KO PBS (M)
KO PBS (L)
1E11 (S)
1E11 (M)
1E11 (L)
2E11 (S)
2E11 (M)
2E11 (L)
5E11 (S)
5E11 (M)
5E11 (L)
1E12 (S)
1E12 (M)
1E12 (L)
2E12 (S)
2E12 (M)
2E12 (L)
WT PBS (S)
WT PBS (M)
WT PBS (L)
(*P<0.05; ** P<0.01;***P<0.001;****P<0.0001)

图45

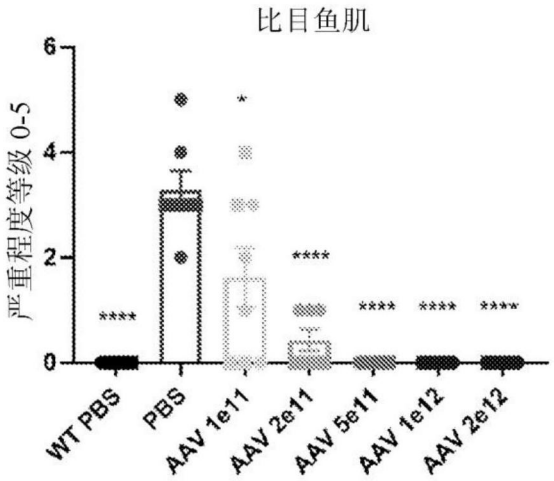


图46A

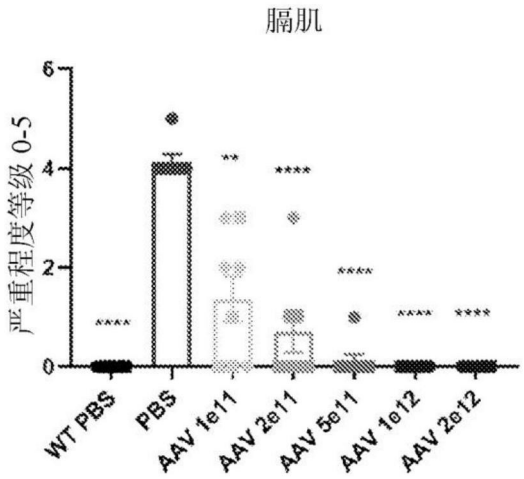


图46B

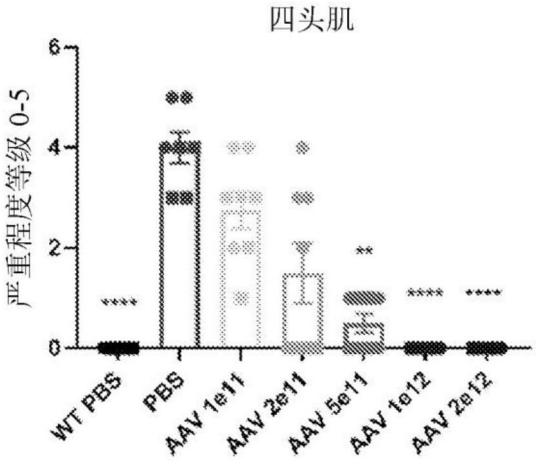


图46C

三头肌

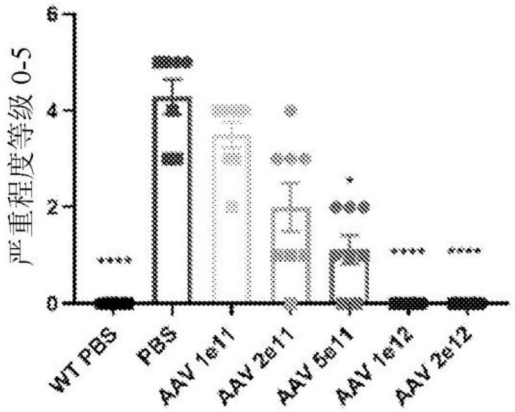


图46D

腓肠肌

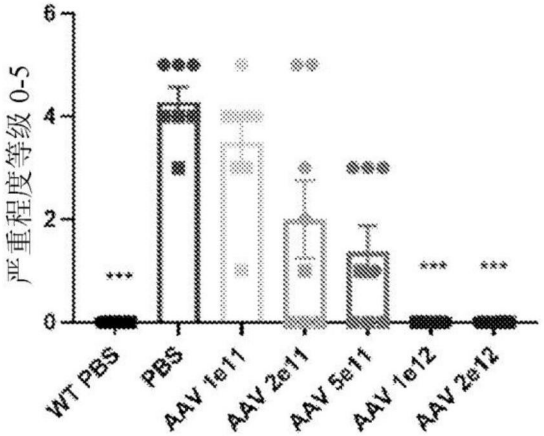


图46E

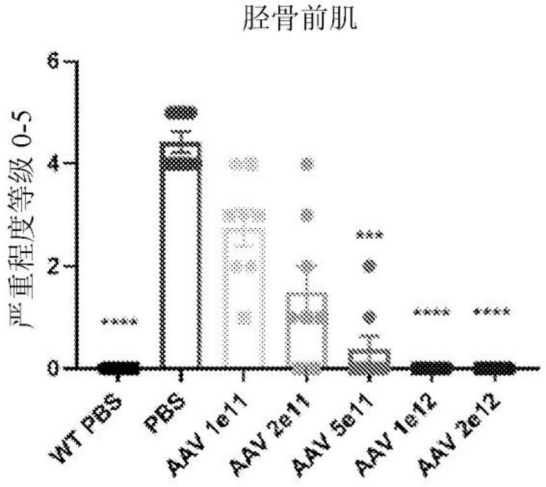


图46F

- Pompe小鼠**6-7个月大**
- **工程改造的AAV.hGAA**

图47

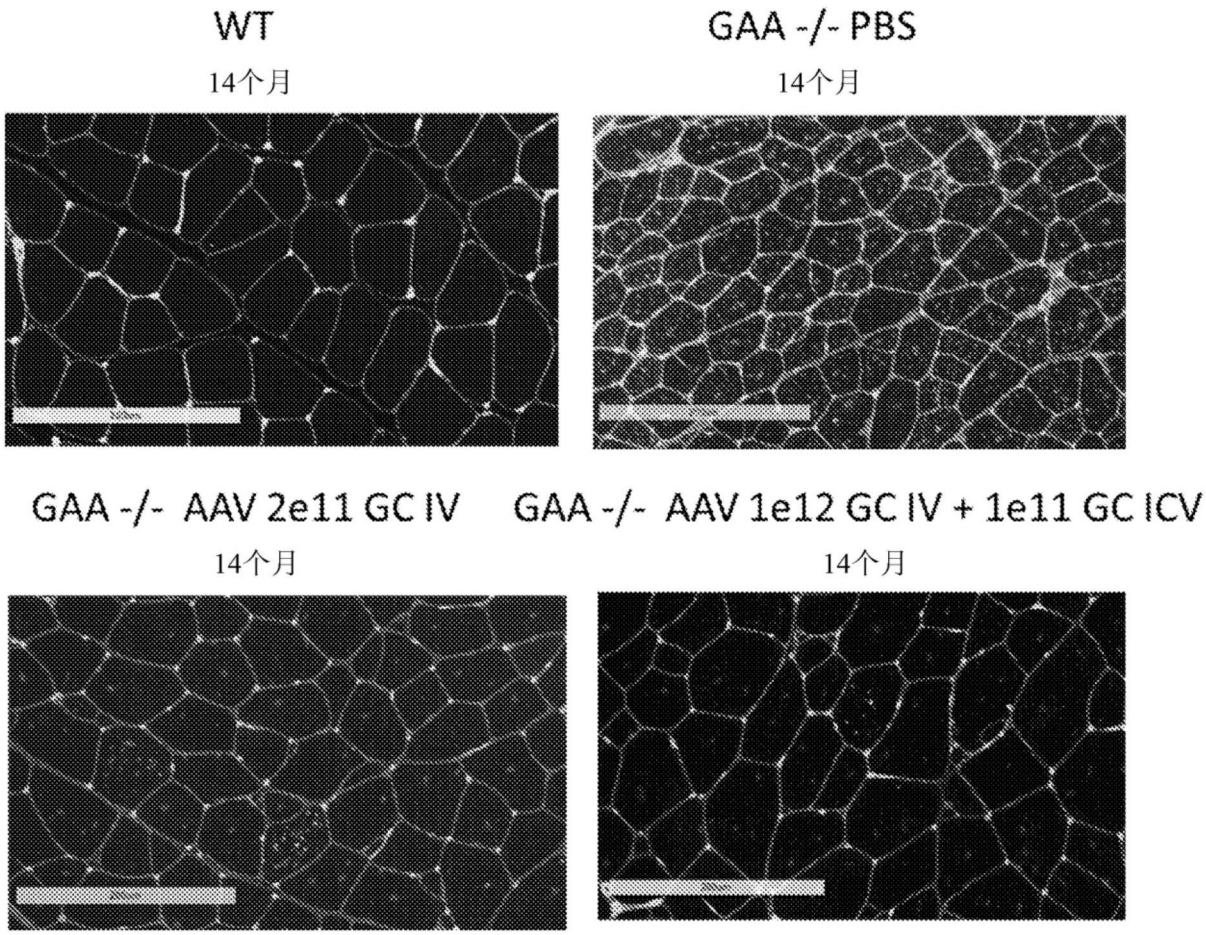

WT
14个月
GAA -/- PBS
14个月
GAA -/- AAV 2e11 GC IV
14个月
GAA -/- AAV 1e12 GC IV + 1e11 GC ICV
14个月

图48

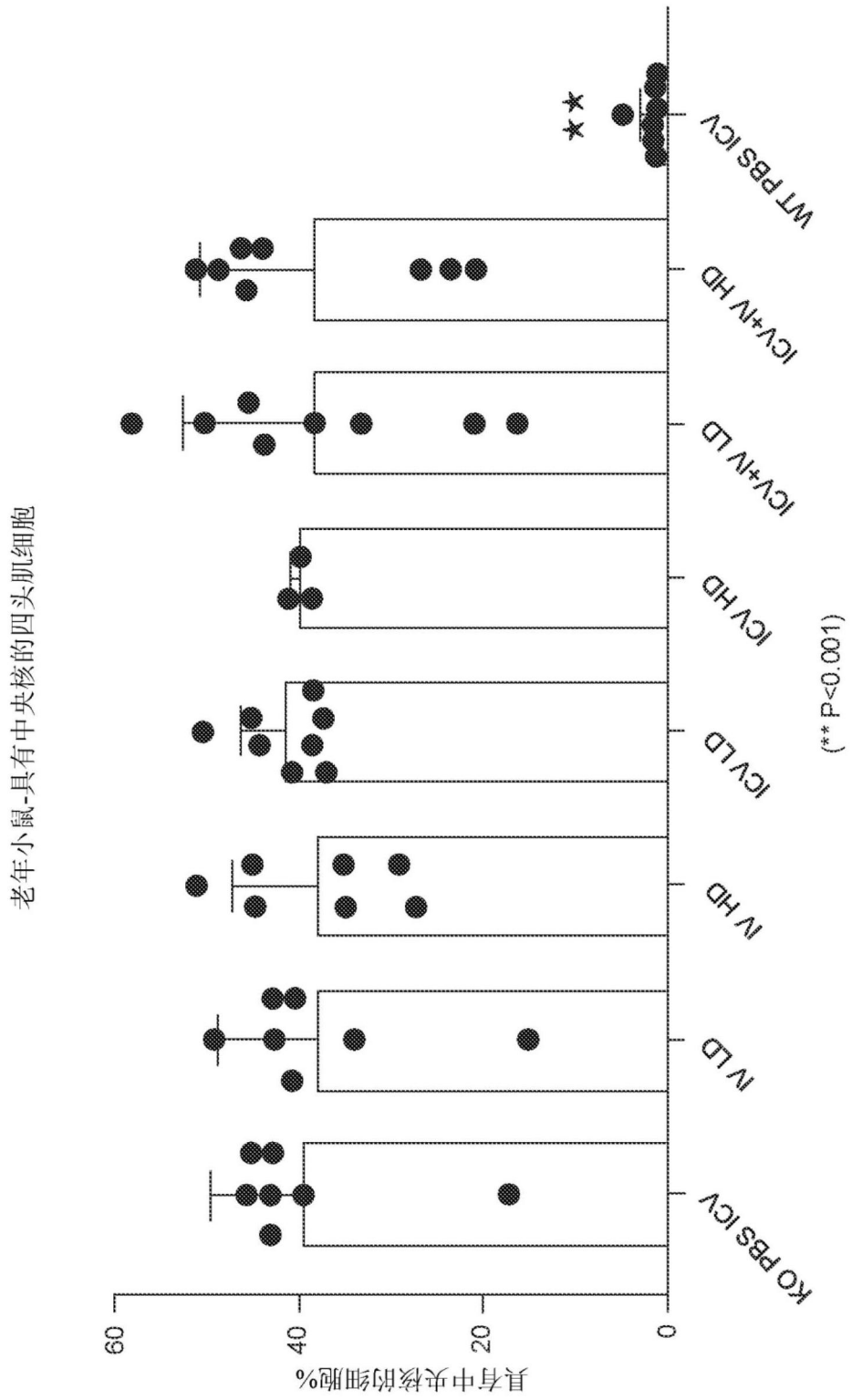

老年小鼠-具有中央核的四头肌细胞
具有中央核的细胞%
0
20
40
60
KO PBS ICV
IV LD
IV HD
ICV LD
ICV HD
ICV+IV LD
ICV+IV HD
WT PBS ICV
(** P<0.001)

图49

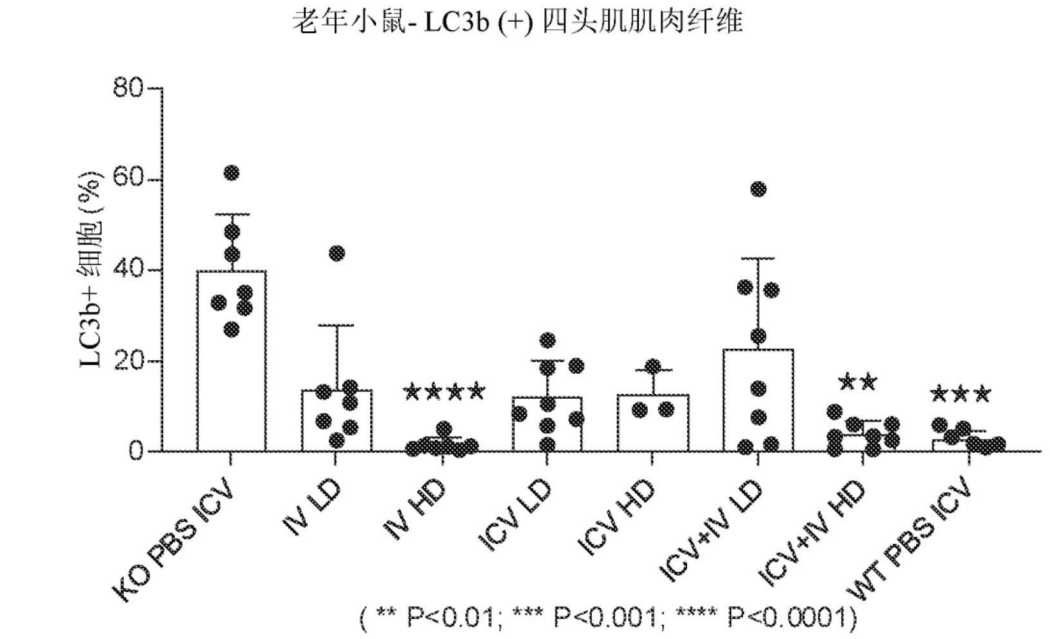

老年小鼠- LC3b (+) 四头肌肌肉纤维
LC3b+ 细胞 (%)
80
60
40
20
0
****
**
***
KO PBS ICV
IV LD
IV HD
ICV LD
ICV HD
ICV+IV LD
ICV+IV HD
WT PBS ICV
(** P<0.01; *** P<0.001; **** P<0.0001)

图50

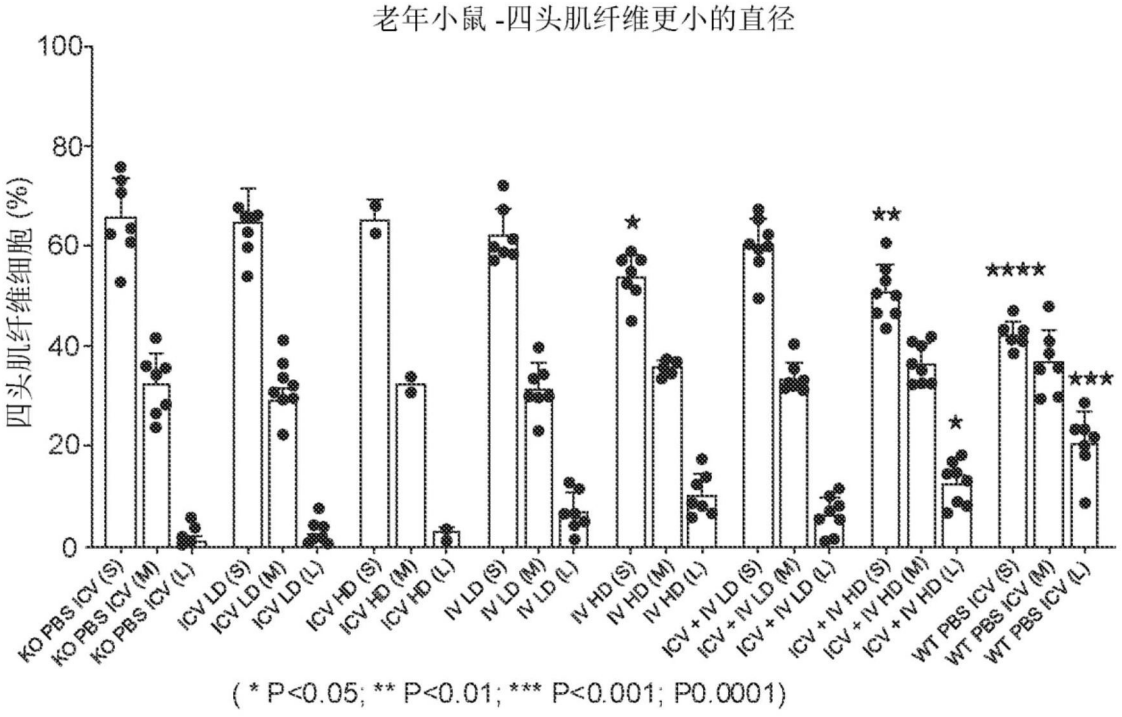

老年小鼠 -四头肌纤维更小的直径
四头肌纤维细胞 (%)
100
80
60
40
20
0
KO PBS ICV (S)
KO PBS ICV (M)
KO PBS ICV (L)
ICV LD (S)
ICV LD (M)
ICV LD (L)
ICV HD (S)
ICV HD (M)
ICV HD (L)
IV LD (S)
IV LD (M)
IV LD (L)
IV HD (S)
IV HD (M)
IV HD (L)
ICV + IV LD (S)
ICV + IV LD (M)
ICV + IV LD (L)
ICV + IV HD (S)
ICV + IV HD (M)
ICV + IV HD (L)
WT PBS ICV (S)
WT PBS ICV (M)
WT PBS ICV (L)
(* P<0.05; ** P<0.01; *** P<0.001; P0.0001)

图51

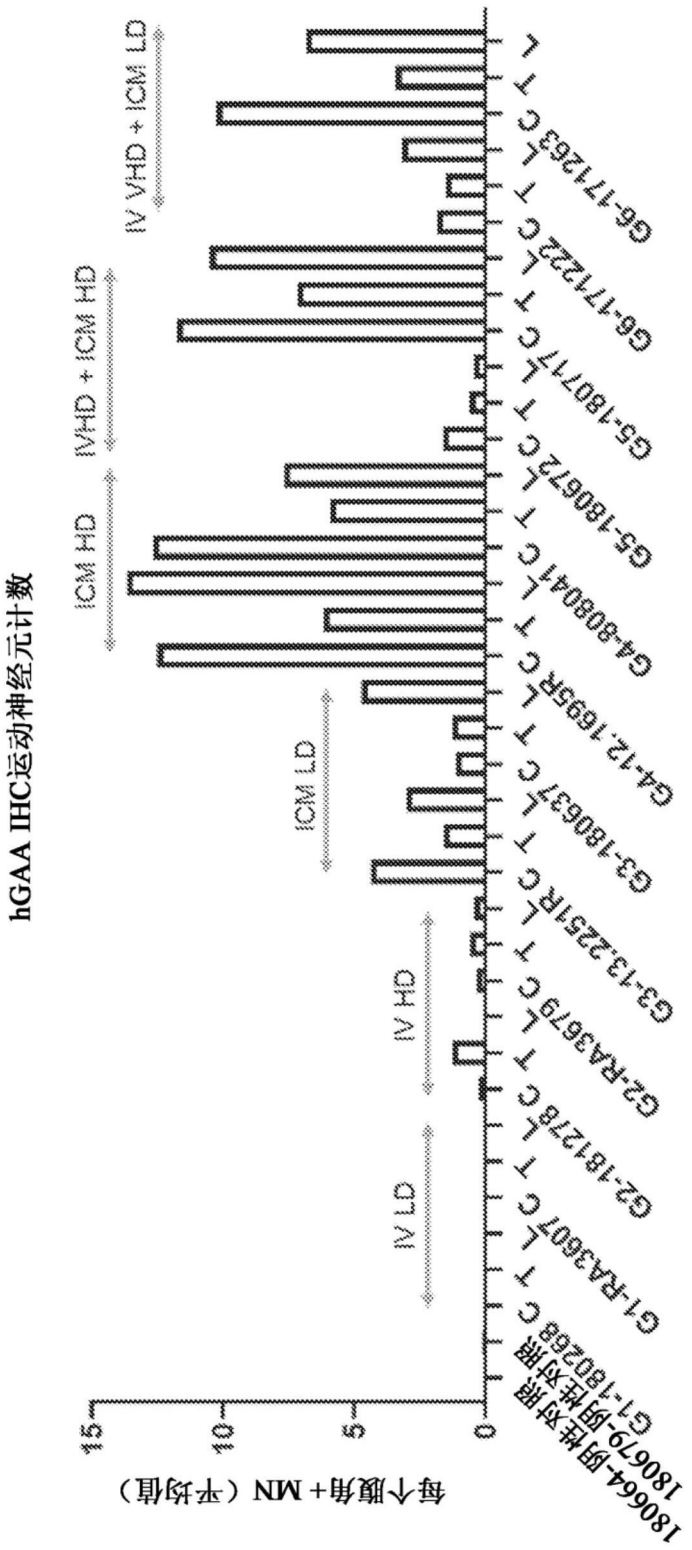

hGAA IHC运动神经元计数
每个腹角+ MN（平均值）
15
10
5
0
IV VHD + ICM LD
IVHD + ICM HD
ICM HD
ICM LD
IV HD
IV LD
18064-阴性对照
18079-阴性对照
G1-1802688 C T
G1-RA3607 C T
G2-18127B C T
G2-RA3679 C T
G3-13.225IR C T
G3-18063 C T
G4-12.1695R C T
G4-8080041 C T
G5-18067I2 C T
G5-18077 C T
G6-17122 C T
G6-17126 C T

图52