



(12) 发明专利申请



(10) 申请公布号 CN 119522285 A

(43) 申请公布日 2025. 02. 25

(21) 申请号 202380038015.4

(22) 申请日 2023.03.03

(30) 优先权数据

63/316219 2022.03.03 US

(85) PCT国际申请进入国家阶段日

2024.11.01

(86) PCT国际申请的申请数据

PCT/US2023/063681 2023.03.03

(87) PCT国际申请的公布数据

W02023/168405 EN 2023.09.07

(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

(72) 发明人 J·M·威尔逊 C·欣德尔

窟内真

(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 彭昶

(51) Int.Cl.

C12N 15/86 (2006.01)

A61K 48/00 (2006.01)

权利要求书2页 说明书35页
序列表(电子公布) 附图18页

(54) 发明名称

编码GLP-2受体激动剂融合物的病毒载体及其在治疗短肠综合征中的用途

(57) 摘要

提供了用于治疗受试者的短肠综合征的组合物和方法。提供了一种病毒载体,所述病毒载体包含核酸分子,所述核酸分子包括编码GLP-2受体激动剂融合蛋白的序列和引导其表达的调控序列。

1. 一种组合物,其包括核酸,所述核酸包括编码融合蛋白的序列,所述融合蛋白包括GLP-2蛋白和IgG4 Fc、白蛋白或XTEN多肽,其中所述融合蛋白具有SEQ ID NO:13、15、17、19、21或23的序列或与其至少99%相同的序列。

2. 根据权利要求1至7中任一项所述的组合物,其中编码所述融合蛋白的所述序列是SEQ ID NO:14、16、18、20、22或24或共享与其至少75%相同的序列。

3. 所述的组合物,其包括病毒载体,所述病毒载体包括:

(a) 腺相关病毒(AAV)衣壳;以及

(b) 包装在所述AAV衣壳中的载体基因组,所述载体基因组包括AAV反向末端重复序列(ITR),包括GLP-2蛋白和IgG4 Fc、白蛋白或XTEN多肽的编码序列,其中所述融合蛋白具有SEQ ID NO:13、15、17、19、21或23的序列或与其至少99%相同的序列,以及引导所述融合蛋白的表达的调控序列。

4. 根据权利要求1至3中任一项所述的组合物,其中所述病毒载体是具有AAVrh91或AAVhu68的所述AAV衣壳的rAAV。

5. 根据权利要求1至4中的一项所述的组合物,其中所述融合蛋白处于诱导型基因表达系统的控制下。

6. 根据权利要求5所述的组合物,其中所述诱导型基因表达系统包括可调控启动子、激活结构域和DNA结合结构域。

7. 根据权利要求3至6中任一项所述的组合物,其中所述AAV反向末端重复序列(ITR)是位于所述融合蛋白编码序列和所述调控序列的侧翼的AAV2 5' ITR和AAV2 3' ITR。

8. 根据权利要求3至7中任一项所述的组合物,其中所述载体基因组包括CB7启动子和兔珠蛋白poly A。

9. 根据权利要求5至8中任一项所述的组合物,其中所述诱导型基因表达系统包括

(a) 激活结构域,所述激活结构域包括反式激活结构域和FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域;

(b) DNA结合结构域,所述DNA结合结构域包括锌指同源结构域(ZFHD)和一个、两个或三个FK506结合蛋白结构域(FKBP)亚基基因;以及

(c) ZFHD结合位点的至少一个拷贝,随后是最小启动子,以及

(d) 可调控启动子。

10. 根据权利要求9所述的组合物,其中所述诱导型基因表达系统包括在一个载体中。

11. 根据权利要求9所述的组合物,其中所述诱导型基因表达系统包括在两个载体中。

12. 根据权利要求9至11中任一项所述的组合物,其中所述反式激活结构域包括NF- κ B p65的一部分。

13. 根据权利要求9至12中任一项所述的组合物,其中所述可调控启动子是组成型启动子。

14. 根据权利要求12所述的组合物,其中所述可调控启动子是CMV启动子。

15. 根据权利要求9至14中任一项所述的组合物,其进一步包括IRES或2A。

16. 根据权利要求9至15中任一项所述的组合物,其包括所述ZFHD结合位点的至少8个拷贝。

17. 根据权利要求5至16中任一项所述的组合物,其包括可调控启动子;激活结构域,所

述激活结构域包括p65反式激活结构域和FKBP12-雷帕霉素相关蛋白 (FRAP) 的FKBP12-雷帕霉素结合 (FRB) 结构域;DNA结合结构域,所述DNA结合结构域包括锌指同源结构域 (ZFHD) 和三个FK506结合蛋白结合域 (FKBP) 亚基基因;所述ZFHD结合位点的12个拷贝;以及编码包括GLP-2类似物和人IgG4 Fc的融合蛋白的序列。

18. 一种适用于治疗受试者的代谢性疾病的药物组合物,所述药物组合物包括水性液体和根据权利要求1至17中任一项所述的组合物。

19. 根据权利要求1至18中任一项所述的组合物,其在用于治疗患有代谢性疾病的受试者的方法中使用。

20. 一种根据权利要求1至19中任一项所述的组合物的用途,其用于制备用于治疗患有代谢性疾病的受试者的药物。

21. 根据权利要求1至20中任一项所述的组合物或用途,其中所述组合物被调配用于以 1×10^9 GC/kg至 5×10^{13} GC/kg的所述rAAV的剂量施用。

22. 根据权利要求1至20中任一项所述的组合物或用途,其中患者是人并且向所述患者施用 1×10^{10} GC至 1.5×10^{15} GC的所述rAAV的剂量。

23. 根据权利要求1至22中任一项所述的组合物或用途,其中所述rAAV是肌肉或静脉内递送的。

24. 一种治疗患有代谢性疾病的受试者的方法,所述方法包括向所述受试者递送重组腺相关病毒 (rAAV),所述rAAV具有来自腺相关病毒rh91或hu68的AAV衣壳和包装在所述AAV衣壳中的载体基因组,所述载体基因组包括AAV反向末端重复序列 (ITR)、编码包括GLP-2类似物和人IgG4 Fc的融合蛋白的序列以及引导所述融合蛋白的表达的调控序列。

25. 根据权利要求24所述的方法,其中向患者施用根据权利要求1至18中任一项所述的组合物。

26. 根据权利要求24或25所述的方法,其中向所述患者施用 1×10^9 GC/kg至 5×10^{13} GC/kg体重的所述rAAV的剂量。

27. 根据权利要求24至26中任一项所述的方法,其中所述rAAV是肌肉或静脉内递送的。

28. 根据权利要求1至18中任一项所述的组合物,其用于治疗人的糖尿病。

编码GLP-2受体激动剂融合物的病毒载体及其在治疗短肠综合征中的用途

[0001] 电子序列表的引用

[0002] 电子序列表的内容(22-10018PCT_Seq-Listing.xml;大小:161kb;以及创建日期:2023年3月3日)通过引用整体并入本文。

背景技术

[0003] 短肠综合征(SBS)是一种罕见的器官衰竭,由先天或获得性原因导致的肠道外科手术切除引起。SBS是肠衰竭(IF)的主要原因,肠道功能持续下降到低于吸收大量营养物质和/或水和电解质所需的最低水平。SBS/IF的标准治疗包含终身每日肠胃外营养(PN),即静脉内输注特殊形式的食物,以支持患者每日所需的营养。除了与每天进行PN相关联的负担和风险外,PN还会引起并发症,包含感染、肠道发育不全、肝病、肾功能障碍和骨骼疾病,对SBS患者的生活质量产生严重的负面影响。在胰高血糖素样肽-2(GLP-2)激动剂替度鲁肽(Teduglutide)开发并获得FDA和EMA批准之前,用于SBS的药物一直是仅治疗与SBS/IF相关联的症状的药物。

[0004] GLP-2是一种33个氨基酸的肠肽激素,具有强效的促肠生长效应。GLP-2确实增加了小肠的长度和肠上皮的绒毛长度,引起营养的有效吸收(即使SBS患者的残余肠较短),并减少了PN体积和数量。然而,由于DPP IV的切割,天然GLP-2的血清半衰期非常短,因此不能用作有效的注射剂。DPP IV抗性长效GLP-2激动剂,包含与IgG Fc结构域和血清白蛋白的融合物,已被开发为SBS的疾病修饰剂。替度鲁肽是一种GLP-2样肽,具有针对DPP IV抗性的A2G突变,将血清半衰期从5分钟延长到1.5小时。每天皮下注射0.05mg/kg的替度鲁肽可在整个2年的治疗过程中持续并连续减少PN体积。在极好安全性概况的情况下,通过这种治疗,90%的患者每周的PN体积相较于基线减少了20%以上,并且70%的患者每周获得了相较于基线大于另外1天的PN中止(Schwartz等人,2016)。为了维持减少的PN体积和数量,需要终身每日注射替度鲁肽(Compher等人,2011)。

[0005] 需要用于SBS的改进的治疗。

发明内容

[0006] 本文提供了编码胰高血糖素样肽2(GLP-2)受体激动剂融合蛋白构建体的病毒载体。在一些实施例中,与不具有融合配偶体的GLP-2受体激动剂的载体介导的递送相比,这些病毒载体可以实现GLP-2受体激动剂在受试者中的持续表达和/或增加的循环半衰期。进一步提供了制备和使用此类病毒载体的方法。

[0007] 一方面,提供了一种病毒载体,所述病毒载体包含核酸,所述核酸包括编码融合蛋白的多核苷酸序列。所述融合蛋白包含(a)前导序列,所述前导序列包括分泌信号肽,(b)胰高血糖素样肽-1(GLP-2)受体激动剂,以及(c)融合结构域,所述融合结构域包括(i)IgG Fc或其功能变体,(ii)白蛋白或其功能变体,或(iii)XTEN多肽(Podust等人,240:52-66(2016年10月))。在一个实施例中,所述载体是腺相关病毒载体。

[0008] 在一个实施例中, (i) 所述前导序列的所述分泌信号肽包括凝血酶信号肽; (ii) 所述前导序列包括凝血酶前肽; 和/或 (iii) 所述前导序列包括凝血酶前导序列。在另一个实施例中, 所述前导序列包括IL-2前导序列。在一个实施例中, GLP融合物选自SEQ ID NO:13、15、17、19、21或23及其功能变体。

[0009] 在一个实施例中, 所述融合结构域是人IgG4 Fc, 所述人IgG4 Fc具有SEQ ID NO:8的序列或与其共享至少90%同一性的序列或其功能变体。在另一个实施例中, 所述融合结构域是人白蛋白, 所述人白蛋白具有SEQ ID NO:11的序列或与其共享至少90%同一性的序列或其功能变体。在一个实施例中, 所述融合结构域是恒河猴IgG4 Fc, 所述恒河猴IgG4 Fc具有SEQ ID NO:9的序列或与其共享至少90%同一性的序列或其功能变体。

[0010] 另一方面, 所述病毒载体包含AAV衣壳和包装在所述AAV衣壳中的载体基因组, 所述载体基因组包括AAV反向末端重复序列 (ITR)、编码所述融合蛋白的所述多核苷酸序列以及引导所述融合蛋白的表达的调控序列。

[0011] 另一方面, 提供了一种适用于治疗受试者的短肠综合征的药物组合物。所述组合物包含水性液体和如本文所述的病毒载体。在一个实施例中, 所述受试者是人。

[0012] 在又另一方面中, 提供了如本文所述的病毒载体用于制造用于治疗患有短肠综合征, 任选地糖尿病的受试者的药物的用途。

[0013] 另一方面, 提供了一种治疗患有短肠综合征的受试者的方法。所述方法包含向所述受试者施用有效量的如本文所述的病毒载体或组合物。

[0014] 根据本发明的以下详细描述, 本发明的其它方面和优点将变得显而易见。

附图说明

[0015] 图1是胰高血糖素原在体内加工的示意图。示出了GLP-1 (SEQ ID NO:4) 和GLP-2 (SEQ ID NO:1) 的序列。

[0016] 图2是示出了GLP-2和类似物的序列的表。

[0017] 图3A是GLP2.G2.Fc构建体的示意图, 所述构建体包含GLP2氨基酸序列中的A2G取代、凝血酶前导序列和IgG4 Fc融合物。

[0018] 图3B是GLP2.G2.SA构建体的示意图, 所述构建体包含GLP2氨基酸序列中的A2G取代、凝血酶前导序列和血清白蛋白融合物。

[0019] 图3C是人 (SEQ ID NO:1)、食蟹猴 (SEQ ID NO:5) 和小鼠 (SEQ ID NO:6) GLP2序列的比对。

[0020] 图4A示出了使用albuquerque柱纯化的hGLP2-SA蛋白的洗脱曲线。

[0021] 图4B示出了具有图4A中染色的hGLP2-SA蛋白的凝胶。

[0022] 图5A示出了以 1×10^{11} gc/小鼠肌内给药AAVrh91.CI.hGLP2.G2.SA.rBG的RagK0小鼠的hGLP2-SA水平。

[0023] 图5B是两张图表, 示出了如图5A所描述的经媒介剂和载体处理的小鼠的小肠长度和重量。

[0024] 图5C示出了如图5A所描述的经媒介剂和载体处理的小鼠的肠组织学。与经媒介剂处理的小鼠相比, 经载体处理的小鼠肠显示出更健康的绒毛。

[0025] 图6A示出了以 1×10^{11} gc/小鼠的剂量用AAVrh91.CI.hGLP2.G2.SA.rBG或

AAVrh91.CI.hGLP2.G2.Fc.rBG处理的小鼠的hGLP2-SA和hGLP2-Fc水平。Fc构建体的GLP-2水平在约研究第7天后较高。

[0026] 图6B示出了hGLP2和hGLP2.Fc的效力测定。hGLP2的EC50确定为0.4nM,而融合蛋白的EC=13.6nM。

[0027] 图7示出了如实例4中描述的实验的研究设计。

[0028] 图8A-8C示出了如实例4中描述的实验的结果。通过肌肉注射(IM)以 1×10^{13} (1e13) GC/kg (E185NG) 的剂量和 5×10^{10} (1e10) GC/kg (BM239H) 的剂量向2个NHP施用AAVrh91.CB7.CI.hGLP-2-Fc.rBG。图8A示出了GLP-2-Fc融合蛋白的血浆水平。图8B示出了血清瓜氨酸,一种肠道表面积的生物标志物。图8C示出了以1:100稀释的NHP血清中的抗GLP-2-Fc抗体的检测。

[0029] 图9A-9F示出了如实例5中描述的实验的结果。Rag1KO雌性小鼠通过IM施用途径以 1×10^{10} GC/小鼠、 3×10^{10} GC/小鼠或 1×10^{11} GC/小鼠的剂量注射载体AAVrh91.CB7.CI.hGLP-2-Fc.rBG来进行处理。研究设计示出于图9A中。图9B示出了血清GLP2水平,图9C示出了随时间推移的体重。图9D示出了第56天的体重。图9E示出了小肠(SI)长度,而图9F示出了SI重量。

具体实施方式

[0030] 本文描述了表达GLP-2激动剂的腺相关病毒(AAV)载体,以通过单次肌肉载体施用治疗SBS/IF患者。转基因GLP-2激动剂包含针对DPP-IV抗性的A2G突变,以及用于进一步延长血清半衰期的与人IgG Fc结构域或血清白蛋白的融合物。与这些半衰期延长技术组合,添加凝血酶前肽能够使得GLP-2激动剂的表达高于治疗水平,其中载体剂量显著降低(即, 1×10^{10} 至 1×10^{12} GC/kg)。本文描述了用于通过施用激活GLP-2激动剂序列的转录的小分子药物来组成性地或以受控方式表达这些蛋白质的表达盒。

[0031] 描述了通过多种途径,并且特别是通过由如rAAV载体等重组载体介导的体内表达将这些构建体递送到有需要的受试者。还提供了在用于治疗有需要的受试者的短肠综合征并且增加受试者的GLP-2的半衰期的方案中使用这些构建体的方法。另外,提供了用于增强受试者的GLP-2的活性的方法。

[0032] GLP-2融合蛋白

[0033] 胰高血糖素原的翻译后加工与GLP-1一起产生胰高血糖素样肽-2(GLP-2),一种33个氨基酸的促肠生长肽激素。GLP-2发挥减缓胃排空,减少胃分泌物,并增加肠血流量的作用。GLP-2还至少通过增强隐窝细胞增殖和绒毛长度来刺激大肠和小肠的生长,从而增加粘膜上皮的表面积。这些作用表明GLP-2可用于治疗各种胃肠道病状。然而,对人类患者单独施用GLP-2还没有显示出前景。GLP-2的半衰期短,限制了其作为治疗剂的用途,因为二肽基肽酶IV(DPP-IV)在体内快速切割GLP-2会产生基本无活性的肽。人GLP-2的氨基酸序列是HADGFSDEMNTILDNLAARDFINWLIQTKITD(SEQ ID NO:1)。

[0034] 如上所述,已经开发出一种名为替度鲁肽的GLP-2类似物,其中氨基酸残基2(丙氨酸)已被甘氨酸取代。此GLP-2类似物的序列在SEQ ID NO:2HGDFGFSDEMNTILDNLAARDFINWLIQTKITD中示出。

[0035] 本公开提供了包括GLP-2受体激动剂的一个或多个拷贝的融合蛋白,以及编码此

类融合蛋白的多核苷酸和载体。在一些实施例中,融合蛋白包括编码融合蛋白的多核苷酸序列,所述融合蛋白包括(a)前导序列,所述前导序列包括分泌信号肽;(b)胰高血糖素样肽-2(GLP-2)受体激动剂;以及(c)融合结构域。在一个实施例中,GLP-2受体激动剂包括凝血酶前导序列、GLP-2受体激动剂以及IgG Fc或其功能变体。在另一个实施例中,融合蛋白包括凝血酶前导序列、GLP-2受体激动剂以及白蛋白或其功能变体。在另一个实施例中,融合蛋白包括凝血酶前导序列、GLP-2受体激动剂的两个拷贝和白蛋白或其功能变体。

[0036] 在一些实施例中,GLP-2受体激动剂包含变体,所述变体可以包含与本文所述的或本领域已知的GLP-2核酸或氨基酸序列至多约10%变化,所述变体保留野生型序列的功能。如本文所使用的,“保留功能”意指核酸或氨基酸以与野生型序列相同的方式起作用,尽管不一定在同一表达或活性水平下。例如,在一个实施例中,与野生型序列相比,功能变体具有增加的表达或活性。在另一个实施例中,与野生型序列相比,功能变体具有降低的表达或活性。在一个实施例中,与野生型序列相比,功能变体的表达或活性具有10%、20%、30%、40%、50%、60%、70%、80%、90%或更大的增加或减少。

[0037] 已经开发了其它GLP-2类似物并且包含格列帕鲁肽(glepaglutide)(SEQ ID NO: 3)、阿普拉鲁肽(apraglutide)和其它类似物,如下文和图2所示。

药物	类型	公司	阶段	T1/2(小时)	EC50 (/GLP2)	备注
替度鲁肽	肽	生锐 (Shire)	获批	1.5(0.1)	1	
格列帕鲁肽	肽	西兰 (Zealand)	P3	50	4	
阿普拉鲁肽	肽	辉凌 (Ferring)	P3	30	0.5	
[0038] HM-15912	Fc 融合物	韩美 (Hanmi)	P2	?	5	
GX-G8	Fc 融合物	Genexine 公司 (Genexine)	P1	?	?	
TAK-681	Fc 融合物	武田 (Takeda) (生锐)	P1	100	5	
ZP7570	肽	西兰	P1	?	10	双作用
NB-1002	XTEN 融合物	亚穆尼克斯公司 (Amunix)	预先	240	100	

[0039] 在一个实施例中,融合物包括GLP-2类似物与异源序列的组合。GLP-2类似物意指与天然人GLP-2(SEQ ID NO:1)或GLP-2-A2G(SEQ ID NO:2)共享至少90%、95%、97%、98%、99%或100%同一性的多肽。在一个实施例中,与天然序列相比,GLP-2类似物具有最多1个、2个或3个氨基酸取代。在另一个实施例中,GLP-2序列源自除人以外的物种。例如,GLP-2可以来自非人灵长类动物、狗、猫、小鼠、大鼠、羊、牛、马等。在一些实施例中,期望改变天然GLP-2序列以优化其一个或多个特征。在某些实施例中,GLP-2具有氨基酸残基2(丙氨酸)已被甘氨酸取代的序列,例如SEQ ID NO:2。例如,在其它实施例中,与天然序列相比,GLP-2类似物含有选自A2G、D3E、S5T、D8S、M10L、N11A、N16A、N24A、Q28A的一个、两个或三个、4个、5个、6个、7个、8个或至多9个氨基酸取代。这些取代已被证明可以提高GLP-2临床概况的功效,包含防止DPP-4失活(A2G)。在一个实施例中,GLP-2类似物是GLP-2的DPP-IV抗性变体。在一个实施例中,GLP-2类似物具有包括SEQ ID NO:2或由其组成的序列。在一个实施例中,变体与SEQ ID NO:2共享至少90%同一性、95%同一性、97%同一性、98%同一性、99%

同一性或100%同一性。

[0040] 融合蛋白可以包括前导序列,所述前导序列可以包括分泌信号肽。如本文所使用的,术语“前导序列”是指多肽的任何N末端序列。

[0041] 前导序列可以源自最终打算施用的相同物种,例如,人。如本文所使用的,术语“源性”或“源自”是指序列或蛋白质来源于特定受试者物种或与来源于特定受试者物种的蛋白质或序列共享同一序列。例如,“源自”人的前导序列与如在人中表达的相同前导序列共享同一序列(或其变体,如本文所定义的)。然而,指定的核酸或氨基酸实际上不需要来源于人。本领域已知能够产生期望序列的各种技术,包含诱变类似蛋白质(例如,同源物)或人工产生核酸或氨基酸序列。“衍生的”核酸或氨基酸在其“所衍生”的物种中保留相同核酸或氨基酸的功能,而不管衍生序列的实际来源。

[0042] 术语“氨基酸取代”和其同义词旨在涵盖通过用另一个取代氨基酸置换一个氨基酸来修饰氨基酸序列。取代可以是保守取代。所述取代还可以是非保守取代。关于两个氨基酸,术语保守旨在意指氨基酸共享本领域技术人员公认的共同性质。例如,具有疏水性非酸性侧链的氨基酸、具有疏水性酸性侧链的氨基酸、具有亲水性非酸性侧链的氨基酸、具有亲水性酸性侧链的氨基酸以及具有亲水性碱性侧链的氨基酸。共同性质还可以是具有疏水性侧链的氨基酸、具有脂肪族疏水性侧链的氨基酸、具有芳香族疏水性侧链的氨基酸、具有极性中性侧链的氨基酸、具有带电侧链的氨基酸、具有带电酸性侧链的氨基酸以及具有带电碱性侧链的氨基酸。天然存在的氨基酸和非天然天然存在的氨基酸两者在本领域中都是已知的并且可以在各实施例用作取代氨基酸。用于置换氨基酸的方法是本领域技术人员熟知的,并且包含但不限于编码氨基酸序列的核苷酸序列的突变。在本文中提及“一个或多个”旨在涵盖例如1个、2个、3个、4个、5个、6个或更多个的单独实施例。

[0043] 在一个实施例中,前导序列是人凝血酶(因子II)序列。在一个实施例中,凝血酶前导序列具有带有最多1个、2个或3个氨基酸取代的SEQ ID NO:7:MAHVRGLQLPGCLALAALCSLVHSQHVFLAPQQARSLLQVRVR或其功能变体中所示的序列。

[0044] 在一个实施例中,期望的前导序列的功能变体包含变体,所述变体可以包含与本文所述的或本领域已知的前导核酸或氨基酸序列至多约10%变化,所述变体保留野生型序列的功能。

[0045] 在一些实施例中,前肽和GLP-2肽两者的编码区被掺入到单个核酸序列中,而在前肽的编码序列与GLP-2的编码序列之间没有连接子。

[0046] 融合蛋白进一步包含融合结构域。在一个实施例中,融合结构域是人IgG Fc片段或其功能变体。免疫球蛋白通常具有较长的体内循环半衰期。通过将GLP-2受体激动剂(和前导序列)与IgG Fc融合,融合蛋白的循环时间被延长,同时GLP-2的功能被保留。在另一个实施例中,融合结构域是恒河猴IgG Fc片段或其功能变体。

[0047] 如本文所使用的,免疫球蛋白的Fc部分具有免疫学领域中通常给予所述术语的含义。具体地,此术语是指不含有来自抗体的两个抗原结合区(Fab片段)的抗体片段。Fc部分由来自两条重链的抗体的恒定区组成,所述两条重链通过非共价相互作用和二硫键缔合。Fc部分可以包含铰链区并且通过CH2和CH3结构域延伸到抗体的c末端。Fc部分可以进一步包含一个或多个糖基化位点。在一个实施例中,融合结构域是人IgG Fc。高度保守的四个亚类IgG1、IgG2、IgG3和IgG4在其恒定区方面,特别是在其铰链和上部CH2结构域方面有所不

同。参见,Vidarsson等人,IgG亚型和同种型:从结构到效应子功能(IgG Subclasses and Allotypes:From Structure to Effector Functions),《免疫学前沿(Front Immunol.)》2014年10月;5:520,所述文献通过引用并入本文。Fc结构域可以源自任何人IgG,包含人IgG1、人IgG2、人IgG3或人IgG4。在一个实施例中,人IgG Fc是IgG4 Fc。在一个实施例中,人IgG Fc是SEQ ID NO:8:

[0048] AESKYGPPCPPAPEAAGGPSVFLFPPKPKDTLMISRTPEVTCVVDVVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTIKAKGQPREPQVYTLPPSQEEMTKN QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYT QKSLSLSLG。在另一个实施例中,人IgG Fc与SEQ ID NO:8共享至少90%同一性、至少95%同一性、至少99%同一性或至少100%同一性。

[0049] 在另一个实施例中,融合结构域是恒河猴IgG Fc。Fc结构域可以源自任何恒河猴IgG,包含恒河猴IgG1、恒河猴IgG2、恒河猴IgG3或恒河猴IgG4。在一个实施例中,恒河猴IgG Fc是IgG4 Fc。在一个实施例中,恒河猴IgG Fc是SEQ ID NO:9:

[0050] PPCPPCPAPE LLGGPSVFLF PPKPKDTLMI SRTPEVTCVVDVVSQEDPEV QFNWYVDGVE VHNAQTKPRE RQFNSTYRVV SVLTVTHQDW LNGKEYTCKV SNKGLPAPIE KTISKAKGQP REPQVYILPP PQEELTKNQV SLTCLVTGFY PSDIAVEWES NGQPENTYKT TPPVLDSDGS YLLYSKLTVN KSRWQPGNIF TCSVMHEALH NHYTQKSLSV SPGK。在另一个实施例中,恒河猴IgG Fc与SEQ ID NO:9共享至少90%同一性、至少95%同一性、至少99%同一性或至少100%同一性。在一个实施例中,恒河猴IgG进一步包括铰链序列。

[0051] 在另一个实施例中,融合结构域是人白蛋白或其功能变体。在一个实施例中,人白蛋白是

[0052] SEQ ID NO:10:

[0053] DAHKSEVAHRFKDLGEEHFKGLVLAFAQYLQQCPFEHVKLNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFHQKDDNPNLPLVRPEVDMCTAFHDNEETFLKKYLYEVAR RHPYFYAPELFFAKRYKAAAFTECCQAADKAAACLLPKLDELREDEGKASSAKQRLKASLQKFGRAFKA WAVARLS QRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVEND EMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLR LAKTYETTLEKCCAAADPHECYAK VFDEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCKHPEAKRMP C AEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIK KQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCKADDKETCF AE EGKLV AASQAALGL。在另一个实施例中,人白蛋白与SEQ ID NO:10共享至少90%同一性、至少95%同一性、至少99%同一性或至少100%同一性。

[0054] 在另一个实施例中,融合结构域是恒河猴白蛋白或其功能变体。在一个实施例中,恒河猴白蛋白是

[0055] SEQ ID NO:11:

[0056] DTHKSEVAHRFKDLGEEHFKGLVLAFAQYLQQCPFEHVKLNEVTEFAKTCVADESAENCDKSLHT LFGDKLCTVATLRETYGEMADCCAKQEPERNECFHQKDDNPNLPLVRPEVDMCTAFHDNEATFLKKYLYEVAR RHPYFYAPELFFAARYKAAFAECCQAADKAAACLLPKLDELREDEGKASSAKQRLKASLQKFGDRAFKAWAVARLS QKFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYMCENQDSISSKLKECCDKPLLEKSHCLAEVEND

EMPADLPSLAADYVESKDVCKNYAEAKDVFGLMFLYFYARRHPDYSVMLLLRLAKAYEATLEKCCAAADPHECYAK
VFDEFQPLVEEPQNLVKQNCLEFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGAKCKLPEAKRMPC
AEDYLSVVLNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTLSEKEKQVK
KQTALVELVKHKPKATKEQLKGVMDNFAAFVEKCKADDKEACFAEEGPKFVAASQAALA。在另一个实施例中，恒河猴白蛋白与SEQ ID NO:11共享至少90%同一性、至少95%同一性、至少99%同一性或至少100%同一性。

[0057] 本公开的融合蛋白的体内功能和稳定性可以通过添加小肽连接子来优化，例如，以防止潜在地不想要的结构域相互作用或出于其它原因。另外，富含甘氨酸的连接子可以提供一些结构柔性，使得GLP-2类似物部分可以与如胰腺的 β 细胞等靶细胞上的GLP-2受体有效地相互作用。因此，在一个实施例中，GLP-2类似物的C末端和融合蛋白的融合结构域的N末端通过连接子融合。在一个实施例中，连接子包含具有序列GGGGSGGGSGGGGS (SEQ ID NO:12)的富含G的肽连接子的1个、1.5个或2个重复序列。

[0058] 在一个实施例中，融合蛋白包括(a)人凝血酶前导序列；(b)GLP-2(A2G)的DPP-IV抗性变体；连接子；以及(c)人IgG Fc。在一个实施例中，融合蛋白具有SEQ ID NO:13的序列或与其至少90%、至少95%、至少98%或至少99%相同的序列。

[0059] SEQ ID NO:13

[0060] MAHVRGLQLPGCLALAAALCSLVHSQHVFLAPQQARSLLRVRRHGDGFSDEMNTILDNLAARDFINW
LIQTKITDGGGGGGSGGGSGGGGSAESKYPPCPPAPEAAGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQ
EDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTIKAKGQPR
EPQVYITLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSGFSFLYSRLTVDKSRWQEG
NVFSCSVMHEALHNHYTQKSLSLSLG*

[0061] 在一个实施例中，编码融合蛋白的序列是SEQ ID NO:14或与其至少75%、至少85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0062] SEQ ID NO:14:

[0063] ATGGCTCACGTTCCGGGACTGCAGCTGCCTGGATGTCTGGCTCTGCCGCTCTGTGTAGCCTGGTGCAC
AGCCAGCACGTGTTTCTGGCTCCTCAGCAAGCCAGATCACTGCTGCAGAGAGTTAGAAGGCACGGCAGCCAGCTT
CAGCGACGAGATGAACACCATCCTGGACAACCTGGCCGCCAGAGACTTCATCAACTGGCTGATCCAGACCAAGATCA
CCGACGGTGGCGGAGGCGGAGGATCTGGTGGTGGTGGATCTGGCGGCGGAGGTTCTGCCGAGTCTAAGTACGGACCT
CCTTGTCCTCCCTGTCTGCTCCAGAAGCTGCTGGCGGCCATCCGTGTTTCTGTTCCCTCAAAGCCTAAGGACAC
CCTGATGATCAGCAGAACCCTGAAGTGACCTGCGTGGTGGTTCGACGTGTCCAAGAGGATCCTGAGGTGCAGTTCA
ATTGGTACGTGGACGGCGTGAAGTGCACAACGCCAAGACCAAGCCTAGAGAGGAACAGTTCAACAGCACCTACAGA
GTGGTGTCCGTGCTGACCGTGTGACCCAGGATTGGCTGAACGGCAAAGAGTACAAGTGAAGGTGTCCAACAAGGG
CCTGCCTAGCAGCATCGAGAAAACCATCAGCAAGGCCAAGGGCCAGCCAAGAGAACCCAGGTGTACACACTGCCTC
CAAGCCAAGAGGAAATGACCAAGAACCAGGTGTCCCTGACCTGCCTGGTCAAGGGCTTCTACCCTTCCGATATCGCC
GTGGAATGGGAGAGCAACGGCCAGCCTGAGAACAACACTACAAGACCACACCTCCTGTGCTGGACTCCGATGGCTCATT
CTTCTGTACAGCAGACTGACCGTGGACAAGAGCAGGTGGCAAGAGGGCAACGTGTTTCAGCTGCAGCGTGATGCACG
AGGCCCTGCACAACCACTACACCCAGAAAAGCCTGAGCCTGTCTCTGGGCTAA

[0064] 在一个实施例中，融合蛋白包括(a)人凝血酶前导序列；(b)GLP-2(A2G)的DPP-IV抗性变体；连接子；以及(c)人血清白蛋白。在一个实施例中，融合蛋白具有SEQ ID NO:15的

序列或与其至少90%、至少95%、至少98%或至少99%相同的序列。SEQ ID NO:15:

[0065] MAHVRGLQLPGCLALAAALCSLVHSQHVFLAPQQARSLLRVRRHGDGFSFSDMNTILDNLAARDFINW
LIQTKITDGGGGGGSGGGGSDAHKSEVAHRFKDLGEENFKALVLI AFAQYLQQCPFEDHVKLVNEVTEFAK
TCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNL PRLV RPEVDVMCTAF
HDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQAADKAAACLLPKLDEL RDEGKASSAKQRLKASL
QKFGERAFAKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECC
EKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLRLAKTYET
TLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLG
KVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFSALEVDETYVPKEFNAETF
TFHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCKADDKETCF AEEGKKLVAASQAA
LG*

[0066] 在一个实施例中,编码融合蛋白的序列是SEQ ID NO:16或与其至少75%、至少85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0067] SEQ ID NO:16:

[0068] ATGGCTCACGTTTCGGGGACTGCAGCTGCCTGGATGCTGGCTCTGCCGCTCTGTGTAGCCTGGTGCAC
AGCCAGCACGTGTTTCTGGCTCCTCAGCAAGCCAGATCACTGCTGCAGAGAGTTAGAAGGCACGGCGACGGCAGCTT
CAGCGACGAGATGAACACCATCCTGGACAACCTGGCCGCCAGAGACTTCATCAACTGGCTGATCCAGACCAAGATCA
CCGACGGTGGCGGAGGCGGAGGATCTGGTGGTGGTGGATCTGGCGGCGGAGGTTCTGACGCCACAAATCTGAAGTG
GCCCACCGGTTCAAGGACCTGGGCGAAGAGAATTTCAAGGCCCTGGTGTGCTGATCGCCTTCGCTCAGTACCTGCAGCA
GTGCCCTTCGAGGACCACGTGAAGCTGGTCAACGAAGTGACCGAGTTCGCCAAGACCTGCGTGGCCGACGAGAGCG
CCGAGAACTGTGATAAGAGCCTGCACACCCTGTTTCGGCGACAAGCTGTGTACAGTGGCCACACTGAGAGAAACCTAC
GGCGAGATGGCCGACTGCTGCGCCAAGCAAGAGCCGAGAGAAACGAGTGCTTCCTGCAGCACAAGGACGACAACCC
CAACCTGCCTAGACTCGTGC GGCCCTGAAGTGGACGTGATGTGCACCGCCTTCCACGACAACGAGGAAACCTTCTGA
AGAAGTACCTGTACGAGATCGCCAGACGGCACCCCTACTTTTACGCCCTGAGCTGCTGTTCTTCGCCAAGCGGTAT
AAGGCCGCTTACCGAGTGTGTCAGGCCGCTGATAAGGCTGCCTGCCTGCTGCCTAAGCTGGACGAGCTTAGAGA
CGAGGGCAAAGCCAGCTCCGCCAAGCAGAGACTGAAGTGTGCCAGCCTGCAGAAGTTCGGCGAGAGAGCCTTTAAGG
CCTGGGCCGTTGCTAGACTGAGCCAGAGATTTCCCAAGGCCGAGTTTGCCGAGGTGTCCAAGCTCGTGACCGACCTG
ACAAAGGTGCACACCGAGTGTGCCACGGCGACCTGCTGGAATGCGCCGACGATAGAGCCGACCTGGCCAAGTACAT
CTGCGAGAACCAGGACAGCATCAGCAGCAAGCTGAAAGAGTGCTGCGAGAAGCCTCTGCTGAAAAGAGCCACTGTA
TCGCCGAGGTGAAAACGACGAGATGCCC GCGATCTGCCTTCTTGGCCGCGATTTTGTGAAAAGCAAGGACGTG
TGCAAGAACTACGCCGAGGCCAAGGACGTGTTCTGGGCATGTTTCTGTACGAGTACGCCCGCAGACACCCCGACTA
CTCTGTTGTGCTGCTGCTGAGACTGGCCAAAACCTACGAGACAACCCTGAAAAGTGCTGTGCCGCGCTGATCCTC
ACGAGTGTTACGCCAAGGTGTTTCGACGAGTTCAAGCCACTGGTGAAGAACCCCGAACCCTGATCAAGCAGAACTGC
GAGCTGTTTCGAGCAGCTGGGCGAGTACAAGTTCAGAACGCCCTGCTCGTGCAGTACACCAAGAAGGTGCCCCAGGT
TTCCACACCTACACTGGTTGAGGTGTCCCGAACCTGGGCAAAGTGGGCAGCAAGTGTGCAAGCACCCCTGAGGCCA
AGAGAATGCCCTGCGCCGAGGATTACCTGAGCGTGGTGTGAATCAGCTGTGCGTGCTGCACGAGAAAACCCCTGTG
TCCGACAGAGTGACCAAGTGTGTACCGAGAGCCTGGTCAACAGACGGCCTTGCTTTAGCGCCCTCGAGGTGGACGA
GACATACGTGCCCAAAGAGTTCAACGCCGAGACATTCACCTTCCACGCCGACATCTGTACCCTGAGCGAGAAAAGAGC
GGCAGATCAAGAAACAGACTGCCCTGGTGAAGTGGTCAAGCACAAGCCCAAGGCCACCAAGAACAGCTGAAGGCC

GTGATGGACGACTTCGCCGCTTCGTGGAAAAGTGCTGCAAGGCCGACGACAAAGAGACCTGCTTCGCCGAAGAGGG
CAAGAACTGGTGGCCGCTTCTCAGGCTGCTCTGGGATAA

[0069] 在一个实施例中,融合蛋白包括(a)人凝血酶前导序列;(b)GLP-2(A2G)的DPP-IV
抗性变体;连接子;以及(c)XTEN多肽。在一个实施例中,融合蛋白具有SEQ ID NO:17的序列
或与其至少90%、至少95%、至少98%或至少99%相同的序列。

[0070] SEQ ID NO:17:

[0071] MAHVRGLQLPGCLALAAALCSLVHSQHVFAPQQARSLLRVRRRHGDGFSDEMNTILDNLAARDFINW
LIQTKITDGGGGGGSGGGSSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTST
EPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATP
ESGPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGP
GTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSE
SATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSE
GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAP
GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSE
SATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSE
SATPESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSE
GSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSE
SATPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATP
ESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEE
GTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSE
SATPESGPGTSESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSE
GSAPGTSTEPSEGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG*

[0072] 在一个实施例中,编码融合蛋白的序列是SEQ ID NO:18或与其至少75%、至少
85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0073] SEQ ID NO:18:

[0074] ATGGCTCACGTTCCGGGACTGCAGCTGCCTGGATGTCTGGCTCTTGCCGCTCTGTGTAGCCTGGTGCAC
AGCCAGCACGTGTTTCTGGCTCCTCAGCAAGCCAGATCACTGCTGCAGAGAGTTAGAAGGCACGGCGACGGCAGCTT
CAGCGACGAGATGAACACCATCCTGGACAACCTGGCCGCCAGAGACTTCATCAACTGGCTGATCCAGACCAAGATCA
CCGACGGTGGCGGAGGCGGAGGATCTGGTGGTGGTGGATCTGGCGGCGGAGGAAGTTCTCCTGCTGGCAGCCCTACA
AGCACCGAGGAAGGCACAAGCGAGTCTGCCACACCTGAGTCTGGCCCTGGCACATCTACAGAGCCTAGCGAAGGATC
TGCCCCAGGATCTCCTGCCGCTCTCCAACATCTACCGAAGAGGGAACCAGCACCGAGCCATCTGAGGGATCTGCTC
CCGGAACAAGCACAGAGCCTTCAGAAGGATCCGCTCCTGGCACCTCTGAAAGCGCCACACCAGAAAGCGGACCTGGA
TCTGAGCCTGCCACAAGCGGATCTGAGACACCTGGAAGCGAGCCAGCCACATCTGGCAGCGAAACACCTGGTTCTCC
AGCCGGATCTCCCACCAGCACAGAAGAGGGCACATCCGAATCTGCTACCCCTGAATCTGGACCAGGCACCTCCACAG
AACCTTCCGAGGGTTCTGCCCTGGAACCTCTACCGAACCATCAGAAGGCAGCGCTCCAGGTTACCAGCTGGAAGC
CCAACCTCTACAGAGGAAGGGACATCCACTGAGCCAAGCGAGGGAAGCGCTCCCGGCACTAGTACAGAACCGAGCGA
GGGAGTGTCTGGAACCAGCGAATCCGCTACTCCAGAGAGTGGCCCAGGCACCAGTACTGAACCCTCTGAGGGTA
GCGCACCCGGAACATCTGAGAGCGCTACTCCGAATCAGGCCAGGCTCTGAACCTGCTACCAGCGGAAGTGAAACA
CCCGGCACCTCTACTGAGCCCTCCGAAGGCTCAGCACCTGGCACAAGCACTGAACCATCAGAGGGCTCCGCACCAGG
CACCAGCGAAAGTGCTACACCAGAGTCAGGACCCGGAACCTCCGAAAGTGCAACTCCTGAGAGCGGACCAGGCTCTC
CCGCTGGATCTCCTACATCAACTGAAGAAGGGACCTCCGAGAGCGCAACCCAGAGTCTGGTCCAGGATCAGAACCT
GCCACCTCCGGCTCTGAAACCCAGGCACTTCTGAGTCCGCCACGCCAGAATCTGGTCTGGGACTAGCACCGAACC

GAGTGAAGGTTTCAGCTCCCGGGACTTCTACGGAACCCAGTGAAGGATCTGCACCCGGCACATCAACCGAACCGTCAG
 AGGGATCAGCCCCTGGGACTTCCACAGAGCCGTCTGAGGGCAGCGCCCCAGGGACGTCTACAGAACCATCTGAAGGA
 TCAGCACCAGGGACCTCTACCGAGCCAAGTGAAGGCAGTGCACCCGGGAAGTCCAGCAGGCTCCCCTACAAGTACTGA
 AGAGGGTACTAGCACGGAACCCAGCGAGGGTTCCGCTCCAGGGACATCTGAATCCGCAACTCCGGAATCCGGACCTG
 GCAGTGAACCAGCTACATCCGGATCCGAGACTCCGGGAACCTCAGAATCAGCTACACCCGAGAGTGGACCTGGCTCC
 GAACCAGCAACTAGCGGCTCAGAACTCCTGGGACAAGCGAGAGTGCAACACCCGAATCTGGACCTGGAACAAGTAC
 TGAGCCCAGCGAAGGCAGCGCCCCTGGAACTTCTGAATCTGCCACTCTGAAAGTGGCCCTGGAAGCCCTGCAGGCT
 CACCCACATCCACAGAAGAAGGATCACCAGCAGGCAGCCCCACTTCAACGGAAGAGGGATCCCAGCTGGATCCCCA
 ACTAGTACGGAAGAAGGCACCTCAGAAAAGCGCTACGCCGAGTCCGGTCTTGGCACTTCTACTGAACCATCCGAGGG
 AAGTGGCCCTGGCACTTCCGAGAGTGCTACACCTGAAAGCGGTCCCGGCTCTGAACCAGCCACTTCTGGATCTGAAA
 CGCCCCGGACATCCGAGTCAGCAACGCCGAAAAGCGGCCAGGTTCCGAGCCGGCTACTAGTGGTTCAGAGACTCCA
 GGGACTTCCGAGTCTGCTACTCCTGAGTCCGGACCGGGAACATCAACCGAGCCTTCCGAAGGATCTGCACCTGGAAG
 CCCTGCCGATCTCCTACCAGTACTGAGGAAGGCACCTCAGAGTCTGCCACTCCAGAGTCAGGTCCTGGAAGCGAAC
 CTGCAACAAGCGGCAGCGAAACTCCAGGCACTAGCGAGTCAGCTACCCAGAATCAGGACCTGGATCTCCAGCAGGG
 TCCCCAACATCTACTGAGGAAGGCTCTCCTGCTGGCTCCCCTACCTTACCGAAGAGGGGACCTCAACAGAGCCATC
 CGAGGGGAGCGCACCTGGTACATCAGAGTCCGCAACTCCGAGTCTGGCCCCGGAAGTCCGAATCTGCAACCCCGG
 AAAGTGGACCCGGGACGAGTGAATCAGCCACACCTGAATCCGGTCCAGGATCCGAGCCTGCAACTTCTGGAAGCGAG
 ACACCAGGATCTGAGCCAGCTACGTCTGGCTCTGAGACTCCTGGATCTCCTGCTGGTAGTCCCACCTCCACTGAAGA
 GGGAACTTCCACCGAACCGAGCGAGGGATCAGCACCAGGCACTAGCACAGAACCGTCCGAAGGATCTGCTCCAGGCT
 CTGAACCCGCAACCTCCGGATCAGAAACCCCTGGAACATCCGAAAGCGCTACACCGAAAGTGGCCCCGGAACCTCT
 ACAGAACCTAGCGAGGGAAGCGCACCCAGGATAA

[0075] 在一个实施例中,融合蛋白包括(a)恒河猴凝血酶前导序列;(b)GLP-2(A2G)的DPP-IV抗性变体;连接子;以及(c)恒河猴IgG Fc。在一个实施例中,融合蛋白具有SEQ ID NO:19的序列或与其至少90%、至少95%、至少98%或至少99%相同的序列。

[0076] SEQ ID NO:19

[0077] MAHVRGLQLPGCLALALCSLVHSQHVFLAPQQALSLLQVRRRHGDGFSDEMNTVLVDNLATRDFIN
 WLIQTKITDGGGGGGSGGGGSAEFTPPCPPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQ
 EDPEVQFNWYVDGVEVHNAQTKPRERQFNSTYRVVSVLTVTHQDWLNGKEYTCKVSNKGLPAPIEKTISKAKGQPR
 EPQVYILPPPEELTKNQVSLTCLVTGFYPSDIAVEWESNGQPENTYKTPPVLDSDGSYLKSLTVNKSRLWQPG
 NIFTCSVMHEALHNHYTQKSLSVSPG*

[0078] 在一个实施例中,编码融合蛋白的序列是SEQ ID NO:20或与其至少75%、至少85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0079] SEQ ID NO:20:

[0080] ATGGCTCACGTTCCGGGACTGCAGCTGCCTGGATGTCTGGCTCTTGCCGCTCTGTGTAGCCTGGTGCAC
 AGCCAGCATGTGTTTCTGGCTCCTCAGCAGGCCCTGAGCCTGCTGCAAAGAGTTAGAAGGCACGGCGACGGCAGCTT
 CAGCGACGAGATGAATACCGTGTCTGGTGGACAACCTGGCCACCAGAGACTTCATCAACTGGCTGATCCAGACCAAGA
 TCACCGACGGTGGTGGCGGAGGCGGAGGATCTGGTGGCGGTGGTTCTGGCGGTGGCGGATCTGCTGAGTTTACCCT
 CCTGTCTCCTCCTGTCTGCTCCAGAACTGCTCGCGGACCTCCGTGTCTCTGTTTCCCAAAGCCTAAGGACAC
 CCTGATGATCAGCAGAACCCTGAAGTGACCTGCGTGGTGGTGGACGTGTCCAAGAGGATCCTGAGGTGCAGTTCA

ATTGGTACGTGGACGGCGTGGAAGTGCACAACGCCAGACAAAGCCCAGAGAGCGGCAGTTCAACAGCACCTACAGA
GTGGTGTCCGTGCTGACCGTGACACACCAGGATTGGCTGAACGGCAAAGAGTACACCTGTAAAGTCTCCAACAAGGG
CCTGCCTGCTCCTATCGAGAAAACCATCAGCAAGGCCAAGGGCCAGCCTAGAGAACCCAGGTGTACATCCTGCCTC
CACCTCAAGAGGAACTGACCAAGAACCAGGTGTCCCTGACCTGTCTGGTCACCGGCTTCTACCCTTCCGATATCGCC
GTGGAGTGGGAGAGCAACGGACAGCCCGAGAACACCTACAAGACCACACCTCCAGTGTGGACAGCGACGGCTCTTA
CCTGCTGTACTCCAAGCTGACAGTGAACAAGAGCCGGTGGCAGCCCGGCAACATCTTCACCTGTTCTGTGATGCACG
AGGCCCTGCACAACCACTACACCCAGAAAAGCCTGAGCGTGTCCCCTGGATAA

[0081] 在一个实施例中,融合蛋白包括(a)恒河猴凝血酶前导序列;(b)GLP-2(A2G)的DPP-IV抗性变体;连接子;以及(c)恒河猴血清白蛋白。在一个实施例中,融合蛋白具有SEQ ID NO:21的序列或与其至少90%、至少95%、至少98%或至少99%相同的序列。

[0082] SEQ ID NO:21

[0083] MAHVRGLQLPGCLALAAALCSLVHSQHVFAPQQALSLLQVRRRHGDGFSDEMNTVLVDNLATRDFINW
LIQTKITGGGGGGSGGGGSDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQCPFEHVKLKLVNEVTEFAKT
CVADESAENCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPLVRPEVDVMCTAFHD
NEATFLKKYLYEVARRHPYFYAPELLFFAARYKAAFAECCQAADKAACLLPKLDEL RDEGKASSAKQRLKCASLQKF
GDRAFKA WAVARLSQKFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYMCENQDSISSKLECCDKPL
LEKSHCLAEVENDEMPADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVMLLLRLAKAYEATLEKC
CAAADPHECYAKVFEDEFQPLVEEPQNLVKQNCLEFEQLGEYKFNALLVRYTKKVPQVSTPTLVEVSRNLGKVGAKC
CKLPEAKRMPCAEDYLSVVLNRLCVLHEKTPVSEKVTCKCTESLVNRRPCFSALELDEAYVPKAFNAETFTFHADMC
TLSEKEKQVKKQTALVELVKHKPKATKEQLKGVMDNFAAFVEKCKADDKEACFAEEGPKFVAASQAALA

[0084] 在一个实施例中,编码融合蛋白的序列是SEQ ID NO:或与其至少75%、至少85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0085] SEQ ID NO:22:

[0086] ATGGCTCACGTTCCGGGACTGCAGCTGCCTGGATGTCTGGCTCTTGCCGCTCTGTGTAGCCTGGTGCAC
AGCCAGCATGTGTTTCTGGCTCCTCAGCAAGCCCTGAGCCTGCTGCAAAGAGTTAGAAGGCACGGCGACGGCAGCTT
CAGCGACGAGATGAATACCGTGTGGTGGACAACCTGGCCACCAGAGACTTCATCAACTGGCTGATCCAGACCAAGA
TCACCGGTGGTGGCGGAGGCGGAGGATCTGGTGGCGGTGGTCTGGCGGTGGCGGATCTGATACACACAAGTCTGAG
GTGGCCCACCGTTCAAGGACCTGGGCGAAGAACACTTCAAAGGCCTGGTGTGGTGCCTTCAGCCAGTACCTGCA
GCAGTGCCCTTTTCGAGGAACACGTGAAGCTGGTCAACGAAGTGACCGAGTTCGCCAAGACCTGCGTGGCCGACGAGA
GCGCCGAGAACTGTGATAAGAGCCTGCACACCCTGTTCCGGCACAAGCTGTGTACAGTGGCCACACTGAGAGAAACC
TACGGCGAGATGGCCGACTGCTGCGCCAAGCAAGAGCCCGAGAGAAAACGAGTGCTTCTGCAGCACAAGGACGACAA
CCCCAACCTGCCTCCACTCGTCAGACCCGAAGTGGACGTGATGTGCACCGCCTTCCACGACAATGAGGCCACCTTCC
TGAAGAAATACCTGTACGAGGTGGCCAGACGGCACCCCTACTTTTACGCCCTGAACTGCTGTTCTTTGCCGCCAGG
TACAAGGCCGCTTCCCGAATGTTGTCAGGCCGCTGATAAGGCCGCTTGCCTGCTGCCTAAGCTGGACGAGCTTAG
AGACGAGGGCAAAGCCAGCTCCGCCAAGCAGAGACTGAAGTGTGCCAGCCTGCAGAAGTTCGGCGATAGAGCCTTTA
AGGCCTGGGCCGCTCGCTAGACTGAGCCAGAAGTTTCCCAAGGCCGAGTTTGGCGAGGTGTCCAAGCTCGTGACCGAC
CTGACAAAGGTGCACACCGAGTGTGTCACGGCGACCTGCTGGAATGCGCCGACGATAGAGCCGACCTGGCCAAGTA
CATGTGCGAGAACCAGGACAGCATCAGCAGCAAGCTGAAAGAGTGTGCGACAAGCCTCTGCTGGAAAAGAGCCACT
GTCTGGCCGAGGTGGAAAACGACGAGATGCCCGCCGATCTGCCTTCTCTGGCCGCCGATTACGTGGAAAGCAAGGAC

GTGTGCAAGAACTACGCCGAGGCCAAGGACGTGTTCCCTGGGCATGTTTCTGTACGAGTACGCCCGCAGACACCCCGA
CTACTCTGTTATGCTGCTGCTGAGACTGGCCAAGGCCTACGAGGCCACTCTGGAAAAGTGTGTGCCGCCGCTGATC
CCCACGAGTGTTACGCCAAAGTGTTCGACGAGTTCAGCCACTGGTGAAGAACCCAGAACCTGGTCAAGCAGAAC
TGCGAGCTGTTGACGAGCTGGGCGAGTACAAGTTCAGAACGCCCTGCTCGTGCGGTACACCAAGAAGGTGCCCA
GGTTTCCACACCTACACTGGTTGAGGTGTCCCGAACCTGGGAAAAGTGGGCGCCAAGTGTGCAAGCTGCCTGAGG
CCAAGAGAATGCCCTGCGCCGAGGATTACCTGAGCGTGGTGCTGAACAGACTGTGCGTGTGCACGAGAAAACCCCT
GTGTCCGAGAAAAGTGACCAAGTGTGTACCGAGAGCCTGGTCAATCGGAGGCCTTGCTTTAGCGCCCTGGAAGTGA
CGAGGCCTACGTGCCCAAGGCCTTCAACGCCGAGACATTCACCTCCACGCCGACATGTGTACCCTGAGCGAGAAAAG
AAAAGCAAGTGAAGAAAACAGACAGCCCTGGTCGAGCTGGTTAAGCACAAGCCTAAGGCCACCAAAGAACAAGTGAAG
GGCGTGATGGACAACTTCGCCGCTTTGTGGAAAAATGCTGCAAGGCCGACGACAAAGAGGCCTGCTTCGCAGAAGA
GGGCCCTAAGTTTGTGGCCGCTCTCAAGCTGCTCTGGCTTAA

[0087] 在一个实施例中,融合蛋白包括(a)恒河猴凝血酶前导序列;(b)GLP-2(A2G)的DPP-IV抗性变体;连接子;以及(c)XTEN多肽。在一个实施例中,融合蛋白具有SEQ ID NO:23的序列或与其至少90%、至少95%、至少98%或至少99%相同的序列。

[0088] SEQ ID NO:23

[0089] MAHVRGLQLPGCLALALCSLVHSQHVFAPQQALSLLQVRRRHGDGFSDEMNTVLVDNLATRDFINW
LIQTKITDGGGGGGSGGGSSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTE
PSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSESATPES
GPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTS
TEPSEGSAPGTSESATPESGPGSEPATSGSETPGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGTSESATP
ESGPGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSTEPSEGSAPG
TSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSESA
TPESGPGSEPATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGTSESATPESG
PGSPAGSPTSTEEGSPAGSPTSTEEGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGTSESATPESGPGSE
ATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSESATPE
SGPGSEPATSGSETPGTSESATPESGPGSPAGSPTSTEEGSPAGSPTSTEEGTSTEPSEGSAPGTSESATPESGPGT
SESATPESGPGTSESATPESGPGSEPATSGSETPGSEPATSGSETPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPS
EGSAPGSEPATSGSETPGTSESATPESGPGTSTEPSEGSAPG

[0090] 在一个实施例中,编码融合蛋白的序列是SEQ ID NO:24或与其至少75%、至少85%、至少90%、至少95%、至少98%或至少99%相同的序列。

[0091] SEQ ID NO:24:

[0092] ATGGCTCACGTTTCGGGACTGCAGCTGCCTGGATGCTGGCTCTTGCCGCTCTGTGTAGCCTGGTGAC
AGCCAGCATGTGTTTCTGGCTCCTCAGCAAGCCCTGAGCCTGCTGCAAAGAGTTAGAAGGCACGGCGACGGCAGCTT
CAGCGACGAGATGAATACCGTGTGGTGGACAACCTGGCCACCAGAGACTTCATCAACTGGCTGATCCAGACCAAGA
TCACCGACGGTGGCGGAGGCGGAGGATCTGGTGGTGGTGGATCTGGCGCGGAGGAAGTTCTCTGCTGGCAGCCCT
ACAAGCACCGAGGAAGGCACAAGCGAGTCTGCCACACCTGAGTCTGGCCCTGGCACATCTACAGAGCCTAGCGAAGG
ATCTGCCCCAGGATCTCCTGCCGGCTCTCCAACATCTACCGAAGAGGGAACCAGCACCGAGCCATCTGAGGGATCTG
CTCCCGGAACAAGCACAGAGCCTTCAGAAGGATCCGCTCCTGGCACCTCTGAAAGCGCCACACCAGAAAGCGGACCT
GGCTCTGAACCTGCCACAAGCGGATCTGAGACACCTGGAAGCGAGCCAGCCACATCTGGCAGCGAAACACCTGGATC

ACCAGCCGGATCTCCACCTCTACCGAGGAAGGGACATCCGAGAGCGCTACCCAGAATCTGGACCAGGCACCAGCA
CAGAACCCTCTGAAGGTTTCAGCCCCTGGAACCTCTACCGAACCATCAGAAGGCAGCGCTCCAGGTTCTCCCGCTGGA
TCCCCTACATCCACAGAAGAGGGCACCTCCACTGAACCTAGCGAGGGAAGTGCTCCCGGCACTTCCACAGAACCATC
CGAGGGCAGTGACCTGGAACCAGCGAATCTGCTACCCCTGAGAGTGGACCCGGAACATCCACTGAGCCCTCCGAGG
GTTTCAGCTCCAGGCACATCAGAATCCGCCACTCCAGAGTCCGGACCAGGATCTGAGCCAGCTACCAGCGGCTCTGAA
ACACCCGGCACTAGTACCGAGCCAAGCGAGGGTAGCGCACCAGGGACAAGTACCGAACCGTCTGAGGGCTCCGCACC
AGGCACTTCCGAAAGTGCTACTCCTGAAAAGCGGCCAGGCACTAGCGAATCCGCAACACCCGAGAGCGGTCTTGAA
GTCCTGCAGGTTACCTACCAGCACTGAAGAGGGGACTAGCGAGAGCGCAACTCCTGAATCAGGCCCTGGATCCGAA
CCTGCTACCTCCGGAAGTGAAACCCCTGGGACAAGCGAAAGTGCAACGCCCGAGTCAGGACCCGGGACTAGCACGGA
ACCCAGTGAAGGATCTGCACCCGGGACATCTACCGAGCCGTCAGAAGGTTCTGCTCCAGGGACTAGTACTGAGCCTT
CCGAAGGTTCTGCACCTGGAACCTTCCACAGAGCCCAGTGAAGGCAGTGCCCTGGCACAAGCACTGAACCGTCCGAA
GGCAGTGCTCCCGGGACCAGTACAGAACCAGCGAGGGCTCTGCTCCTGGTAGTCCAGCAGGATCTCCAAGTACGAC
CGAAGAAGGGACTTCCACCGAGCCTTCCGAGGGAAGCGTCTGGAACATCCGAGTCCGCTACGCCAGAGAGTGGCC
CAGGTTCTGAACCCGCTACTTCCGGCTCAGAGACTCCTGGGACTTCTGAGTCTGCAACCCCGAAAGTGGTCTGGT
AGCGAACCCAGCAACTAGCGGAAGCGAGACACCCGAACTCAGAGAGTGTACACCGAATCCGGTCCAGGGACGTC
TACGGAACCGTCTGAAGGATCAGCTCCCGGCACTAGCGAAAGCGCTACACCTGAAAGTGGTCCCGGATCTCCAGCAG
GCAGCCCAACCTCTACTGAAGAAGGTTCCCAGCTGGAAGCCCCTCCACTGAGGAAGGCTCTCCCGCAGGCTCA
CCCCTAGTACGGAAGAAGGCACGTCCGAGTCTGCTACTCCGAATCCGGACCTGGAAGTACTGAGCAAGCGA
AGGATCAGCACCCGGAACCTCTGAGTCCGCCACACCAGAATCTGGTCTGGTCCGAGCCTGCCACTTCAGGATCAG
AAACCCCGGGCACGAGTGAATCAGCAACGCCGGAATCTGGCCCCGGAAGCGAACCCGGTACGTCTGGATCTGAAACG
CCAGGGACCTCCGAATCAGCTACGCCTGAGTCTGGTCCAGGGACATCCACCGAACCTAGTGAAGGCTCCGCACCTGG
AAGCCCTGCTGGAAGCCCAACGAGTACTGAAGAGGGCACTTCTGAGAGCGCTACGCCTGAGTCAGGACCTGGAAGCG
AACCTGCAACATCCGGCTCAGAAACACCAGGGACCAGCGAAAGCGCAACCCAGAGAGTGGACCTGGATCTCCAGCT
GGCTCTCCTACTAGTACAGAGGAAGGCAGCCCTGCTGGCTCCCCAACGTCAACAGAAGAAGGTAAGTACTAGCACAGAGCC
CAGCGAGGGTCCGCTCCGGGAACTTCTGAATCTGCTACACCCGAGTCCAGTCCGTTACAAGCGAGTCAGCTACGC
CCGAAAGTGGACCTGGCACCTCAGAGTCTGCAACTCCTGAGAGCGGTCCAGGATCAGAACCAGCCACCTCTGGCTCT
GAGACACCAGGTTCTGAGCCTGCAACGTCCGGAAGCGAAACACCAGGCAGTCCGCGGAAGTCCACTTCAACCGA
AGAGGGGACCTCTACAGAGCCATCAGAGGGCTCTGCACCCGGGCACCTCAACAGAACCATCTGAAGGATCCGCACCGG
GCTCTGAGCCTGCTACTAGTGAAGCGAAACTCCTGGCACCAGTGAATCCGCTACTCCCGAGTCTGGCCCCGGAACG
TCTACTGAACCATCTGAGGGAAGTGCCCCAGGCTAA

[0093] 当期望前导序列、GLP-2受体激动剂或融合结构域的变体或片段时,可以使用野生型核酸序列的定点诱变来产生这些肽的编码序列。可替代地或另外,基于网络的或可商购获得的计算机程序以及基于服务的公司可以用于将氨基酸序列反向翻译为核酸编码序列,包含RNA和/或cDNA两者。参见,例如通过EMBOSS, ebi.ac.uk/Tools/st/; Gene Infinity (geneinfinity.org/sms-/sms_backtranslation.html); ExPasy (expasy.org/tools/)的backtranseq。在一个实施例中, RNA和/或cDNA编码序列被设计成在最终打算施用的受试者物种(例如,人类)中进行最佳表达。

[0094] 可以使用密码子优化将编码序列设计用于最佳表达。密码子优化的编码区可以通过各种不同的方法来设计。可以使用在线可获得的方法、公开的方法或提供密码子优化服

务的公司来进行此优化。例如在国际专利申请公开第WO 2015/012924号中描述了一种密码子优化方法,所述国际专利申请公开通过引用并入本文。简而言之,用同义密码子序列修饰编码产物的核酸序列。适合地,对产物的开放阅读框(ORF)的整个长度进行修饰。然而,在一些实施例中,仅ORF的片段可以被改变。通过使用这些方法中的一种方法,可以将频率应用于任何给定的多肽序列,并且产生编码多肽的经密码子优化的编码区的核酸片段。

[0095] 除了本文所提供的前导序列、GLP-2受体激动剂、融合结构域和融合蛋白之外,还提供了编码这些多肽的核酸序列。在一个实施例中,提供了一种编码本文所述的GLP-2肽的核酸序列。在一些实施例中,所述核酸序列可以包含编码SEQ ID NO:1的GLP-2序列的任何核酸序列。在另一个实施例中,所述核酸序列包含任何核酸,所述任何核酸包含SEQ ID NO:2的GLP-2序列。

[0096] 在一个实施例中,提供了一种编码本文所述的GLP-2融合蛋白的核酸序列。在另一个实施例中,所述核酸序列包含编码SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的任何核酸序列。

[0097] 表达盒

[0098] 另一方面,本文提供了一种表达盒,所述表达盒包括编码如本文所述的GLP-2融合蛋白的核酸。如本文所使用的,“表达盒”是指包括生物学上有用的核酸序列(例如,编码蛋白质、酶或其它有用的基因产物的基因cDNA、mRNA等)和与其可操作地连接的调控序列的核酸分子,所述调控序列引导或调节核酸序列和其基因产物的转录、翻译和/或表达。如本文所使用的,“可操作地连接的”序列包含与核酸序列邻接或非邻接的调控序列(也称为元件)以及以反式或顺式核酸序列起作用的调控序列两者。此类调控序列通常包含例如以下中的一者或多者:启动子、增强子、转录因子、转录终止子、内含子、增强翻译效率的序列(即,Kozak共有序列)、有效的RNA加工信号(如切片和聚腺苷酸化序列)、稳定细胞质mRNA的序列(例如,土拨鼠肝炎病毒(WHP)翻译后调控元件(WPRE))和TATA信号。表达盒可以含有基因序列上游(5'处)的调控序列,例如,启动子、增强子、内含子等中的一者或多者以及增强子中的一个或多个,或基因序列下游(3'处)的调控序列,例如,包括聚腺苷酸化位点的3'非翻译区(3'UTR),以及其它元件。在某些实施例中,调控序列与基因产物的核酸序列可操作地连接,其中调控序列通过介于中间的核酸序列,即5'非翻译区(5'UTR)与基因产物的核酸序列分开。在某些实施例中,表达盒包括一种或多种基因产物的核酸序列。在一些实施例中,表达盒可以是单顺反子表达盒或双顺反子表达盒。在其它实施例中,术语“转基因”是指来自外源的插入到靶细胞中的一个或多个DNA序列。

[0099] 在一个实施例中,表达盒是指包括GLP-2构建体编码序列(例如,GLP-2融合蛋白的编码序列)、启动子的核酸分子,并且可以包含其的其它调控序列,所述盒可以被工程化到遗传元件中和/或包装到病毒载体(例如,病毒颗粒)的衣壳中。通常,用于产生病毒载体的这种表达盒含有本文所述的GLP-2构建体序列和其它表达控制序列(如本文所述的表达控制序列),所述GLP-2构建体序列侧接有病毒基因组(并且被称为“载体基因组”)的包装信号。可以使用本领域已知的技术针对特定物种优化任何表达控制序列,所述技术包含例如密码子优化,如本文所述。

[0100] 在某些实施例中,表达盒包含组成型启动子。在另一个实施例中,使用CB7启动子。CB7是具有巨细胞病毒增强子元件的鸡 β -肌动蛋白启动子。在一些实施例中,CB7启动子具

有SEQ ID NO:25的核酸序列。在一个实施例中,启动子是CMV启动子。在一些实施例中,CMV启动子是SEQ ID NO:26的核酸序列。

[0101] 在另一个实施例中,使用组织特异性启动子。可替代地,可以使用其它肝脏特异性启动子,如在冷泉港的肝脏特异性基因启动子数据库(the Liver Specific Gene Promoter Database,Cold Spring Harbor)(rulai.schl.edu/LSPD)中列出的肝脏特异性启动子,并且包含但不限于 α 1抗胰蛋白酶(A1AT);人白蛋白(Miyatake等人,《病毒学杂志(J.Virol.)》,71:5124-32(1997)),humAlb;乙型肝炎病毒核心启动子(Sandig等人,《基因疗法(Gene Ther.)》,3:1002-9(1996));TTR最小增强子/启动子, α -抗胰蛋白酶启动子,肝脏特异性启动子(LSP)(Wu等人《分子疗法(Mol Ther.)》16:280-289(2008)),TBG肝脏特异性启动子。如病毒启动子、组成型启动子、可调控启动子(参见例如,WO 2011/126808和WO 2013/04943)等其它启动子或对生理学线索有应答的启动子可以用于本文所述的载体中。

[0102] 在一个实施例中,启动子包括在诱导型基因表达系统中。诱导型基因调控/表达系统含有至少以下组分:与编码本文所述的GLP-2融合蛋白的转基因可操作地连接的启动子(也称为可调控启动子)、激活结构域、DNA结合结构域和锌指同源结构域结合位点。在其它实施例中,另外的组分可以包含在表达系统中,如本文进一步描述的。

[0103] 系统包括GLP-2融合蛋白的编码序列上游的启动子。可以使用本文所述的启动子,如CMV和CB7启动子。在一个实施例中,启动子是CMV启动子,如SEQ ID NO:26中所示的启动子。在另一个实施例中,启动子是普遍存在的诱导型启动子Z12I,其包括ZFHD1和IL2最小启动子的结合位点的12个重复拷贝。参见例如,Chen等人,《人类基因疗法方法(Hum Gene Ther Methods.)》2013年8月;24(4):270-278,所述文献并入本文。

[0104] 表达系统包括激活结构域,所述激活结构域优选地位于DNA结合结构域的上游。在一个实施例中,激活结构域是来自NF- κ B的p65亚基的羧基末端与FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域的融合物。在一个实施例中,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域。

[0105] 在一个实施例中,在反式激活结构域与DNA结合结构域之间存在连接子,所述连接子可以是F2A或IRES。在一个实施例中,连接子选自IRES或2A肽。

[0106] DNA结合结构域由锌指同源结构域1(ZFHD1)与FK506结合蛋白(FKBP)的至多三个拷贝连接的DNA结合融合物构成。在存在诱导剂(例如,雷帕霉素类似物,如雷帕霉素)的情况下,DNA结合结构域和激活结构域通过其FKBP和FRB结构域的相互作用而二聚化,从而导致转基因的转录激活。在一些实施例中,ZFHD1包含在具有GT2A或IRES的框架中。

[0107] 表达系统被设计成具有FKBP序列的一个、两个或三个拷贝。这些在本文中称为FKBP亚基。在一个实施例中,亚基被设计成表达相同的蛋白质但具有彼此不同的核酸,以使重组最小化。例如,SEQ ID NO:27提供了FKBP的3个“摆动的”编码序列,所述编码序列中的每个序列编码SEQ ID NO:28:GVQVETISPGDGRTPFKRGQTCVVHYTGMLDGGKFDSSRDRNKPFKMLGKQEVIR GWEEGVAQMSVGRRAKLTISPDIYAYGATGHPGIIPPHATLVFDVELLKLE中所示的序列。

[0108] 表达系统进一步包括锌指同源结构域结合位点。核酸分子含有至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个或12个ZFHD结合位点。在一个实施例中,表达系统含有8(八)个锌指同源结构域结合位点(结合配偶体)(8XZFHD)。然而,本发明涵盖具有锌指结

合位点的二至约十二个拷贝的表达系统。ZFHD结合位点的单个拷贝的实例是：aatgatgggcgctcgagt (SEQ ID NO:29)。

[0109] 在一些实施例中,在锌指同源结构域结合位点的下游存在最小IL2启动子。示例性IL2启动子示出于SEQ ID NO:30。

[0110] 此类诱导型系统在本领域中是已知的,并且包含例如由以下文献描述的雷帕霉素诱导型系统:例如,Rivera等人,用于基因表达的药理学控制的人源化系统(A humanized system for pharmacologic control of gene expression),《自然医学(Nature Medicine)》第2卷,第1028-1032页(1996年9月)以及Rivera等人,在AAV介导的基因转移后灵长类动物的促红细胞生成素的长期药理学调控的表达(Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer),《血液(Blood)》,2005年2月15日,第105卷,第4期,这两个文献均通过引用并入本文。在一个实施例中,诱导型基因表达系统包括CMV启动子,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域、GT2A肽、ZFHD1 DNA结合结构域、三个FKBP亚基、hGH poly A、8XZFHD和最小sIL2启动子。这些序列除了GLP-2融合蛋白的编码序列以及任选地其它调控序列之外。

[0111] 除了启动子之外,表达盒和/或载体还可以含有其它适当的转录起始、终止、增强子序列、如剪接和聚腺苷酸化(polyA)信号等的有效RNA加工信号;稳定细胞质mRNA的序列;增强翻译效率的序列(即,Kozak共有序列);增强蛋白质稳定性的序列;以及在期望时,增强经过编码的产物的分泌的序列。合适的polyA序列的实例包含例如SV40、牛生长激素(bGH)、人生长激素(hGH)、SV40、兔 β -珠蛋白(也称为兔珠蛋白polyA;RGB)、经修饰的RGB(mRGB)和TK polyA。合适的增强子的实例包含例如 α -胎蛋白增强子、TTR最小启动子/增强子、LSP (TH结合球蛋白启动子/ α 1微球蛋白/双库尼茨抑制剂增强子)等。在一个实施例中,polyA是兔珠蛋白polyA。

[0112] 这些控制序列与GLP-2构建体序列“可操作地连接”。如本文所使用的,术语“可操作地连接”是指与所关注基因邻接的表达控制序列以及以反式或在远处起作用以控制所关注基因的表达控制序列两者。

[0113] 在一个实施例中,提供了rAAV,所述rAAV包含5' ITR、CB7启动子、鸡 β -肌动蛋白内含子、SEQ ID NO:13、15、17、19、21或23的融合蛋白的编码序列、兔珠蛋白poly A和3' ITR。在另一个实施例中,rAAV包括包含CMV启动子的多核苷酸,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域、IRES、ZFHD1 DNA结合结构域、三个FKBP亚基、hGH poly A、8XZFHD、最小sIL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。

[0114] 在另一个实施例中,提供了一种二载体诱导型系统。第一rAAV包括12XZFHD、最小IL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。可以包含填充序列以增加载体的包装大小。第二rAAV包括多核苷酸,所述多核苷酸包括CMV启动子、内含子,所述激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白(FRAP)的FKBP12-雷帕霉素结合(FRB)结构域、IRES、ZFHD1 DNA结合结构域和polyA。

[0115] 在一个实施例中,提供了一种表达盒,所述表达盒包含多核苷酸,所述多核苷酸包括CB7启动子、鸡 β -肌动蛋白内含子、SEQ ID NO:13、15、17、19、21或23的融合蛋白的编码序列和兔珠蛋白poly A。在一个实施例中,表达盒是发现于SEQ ID NO:X中的表达盒,或与其共享至少70%、75%、80%、85%、90%、95%、99%或100%同一性的序列。在另一个实施例中,提供了一种载体基因组,其中SEQ ID NO:X-X或与其共享至少70%、75%、80%、85%、90%、95%、99%或100%同一性的序列侧接有5'和3' AAV ITR。

[0116] 在另一个实施例中,提供了一种表达盒,所述表达盒包含多核苷酸,所述多核苷酸包括CB7启动子、鸡 β -肌动蛋白内含子、SEQ ID NO:13、15、17、19、21或23的融合蛋白的编码序列和兔珠蛋白poly A。

[0117] 病毒载体

[0118] 另一方面,提供了包含本文所述的表达盒的病毒载体。在本文所述的病毒载体的某些实施例中,病毒载体是腺相关病毒(AAV)病毒载体或重组AAV(rAAV)。如本文所使用的,术语“重组AAV”或“rAAV”是指天然存在的腺相关病毒、本领域技术人员可获得的和/或鉴于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 DNA酶抗性颗粒,其中包装有侧接有AAV反向末端重复序列(ITR)(一起称为“载体基因组”)的用于递送到靶细胞的表达盒。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,取决于所选择的AAV。可以选择各种AAV作为如上文所鉴定的AAV病毒载体的衣壳的来源。在一个实施例中,AAV衣壳是AAVrh91衣壳或其变体。在某些实施例中,衣壳蛋白由rAAV载体名称中的术语“AAV”之后的数字或数字和字母的组合指定。除非另外指明,否则本文所述的AAV衣壳、ITR和其它所选择的AAV组分可以容易地选自任何AAV,包括但不限于鉴定为AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAVrh10、AAVhu37、AAVrh32.33、AAVAnc80、AAV10、AAV11、AAV12、AAVrh8、AAVrh74、AAV-DJ8、AAV-DJ、AAVhu.37、AAVrh.64R1和AAVhu68的AAV。参见例如,美国公开专利申请第2007-0036760-A1号;美国公开专利申请第2009-0197338-A1号;EP 1310571。还参见WO 2003/042397(AAV7和其它猿猴AAV)、美国专利7790449和美国专利7282199(AAV8)、WO 2005/033321和US 7,906,111(AAV9)和WO 2006/110689以及WO2003/042397(rh.10)、WO 2005/033321、WO 2018/160582(AAVhu68),所述文献通过引用并入本文。其它合适的AAV可以包含但不限于AAVrh90[于2020年4月28日提交的PCT/US20/30273]、AAVrh91[于2020年4月28日提交的PCT/US20/030266,现在是在于2020年11月5日公开的公开WO 2020/223231]、AAVrh92、AAVrh93、AAVrh91.93[于2020年4月28日提交的PCT/US20/30281],所述文献均通过引用并入本文。其它合适的AAV包含于2019年10月21日提交的美国临时专利申请第62/924,112号和于2020年5月15日提交的美国临时专利申请第63/025,753号中描述的AAV3B变体,所述文献描述了AAV3B.AR2.01、AAV3B.AR2.02、AAV3B.AR2.03、AAV3B.AR2.04、AAV3B.AR2.05、AAV3B.AR2.06、AAV3B.AR2.07、AAV3B.AR2.08、AAV3B.AR2.10、AAV3B.AR2.11、AAV3B.AR2.12、AAV3B.AR2.13、AAV3B.AR2.14、AAV3B.AR2.15、AAV3B.AR2.16或AAV3B.AR2.17,所述文献均通过引用并入本文。还参见于2021年8月13日提交的国际专利申请第PCT/US21/45945号、于2020年8月14日提交的美国临时专利申请第63/065,616号和于2020年11月4日提交的美国临时专利申请第63/109,734号,所述专利申请全都通过引用整体并入本文。这些文档还描述了可以选择用

于产生rAAV的其它AAV衣壳,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。

[0119] 如本文所使用的,关于AAV,术语“变体”意指源自已知AAV序列的任何AAV序列,包含具有保守氨基酸置换的AAV序列以及与氨基酸或核酸序列共享至少90%、至少95%、至少97%、至少99%或更高的序列同一性的AAV序列。在另一个实施例中,AAV衣壳包含变体,所述变体可以包含与任何所描述或已知的AAV衣壳序列至多约10%变化。也就是说,AAV衣壳与本文所提供的和/或本领域已知的AAV衣壳共享约90%同一性至约99.9%同一性、约95%至约99%同一性或约97%至约98%同一性。在一个实施例中,AAV衣壳与AAV衣壳共享至少95%同一性。当确定AAV衣壳的同一性百分比时,可以对任何可变蛋白质(例如,vp1、vp2或vp3)进行比较。

[0120] 在一个实施例中,病毒载体是具有AAV8的衣壳或其功能变体的rAAV。在一个实施例中,病毒载体是具有AAVrh91的衣壳或其功能变体的rAAV。在一个实施例中,病毒载体是具有AAV3.AR.2.12的衣壳或其功能变体的rAAV。在一个实施例中,病毒载体是具有选自AAV9、AAVrh64R1、AAVhu37或AAVrh10的衣壳的rAAV。

[0121] 在某些实施例中,提供了新型分离的AAVrh91衣壳。在SEQ ID NO:31中提供了编码AAVrh91衣壳的核酸序列,并且在SEQ ID NO:32中提供了经编码的氨基酸序列。本文提供了一种rAAV,所述rAAV包括AAVrh91(SEQ ID NO:32)的vp1、vp2和vp3中的至少一个。本文还提供了rAAV,所述rAAV包括由AAVrh91(SEQ ID NO:31)的vp1、vp2和vp3中的至少一个编码的AAV衣壳。在又一个实施例中,在SEQ ID NO:19中提供了编码AAVrh91氨基酸序列的核酸序列,并且在SEQ ID NO:32中提供了经编码的氨基酸序列。本文还提供了rAAV,所述rAAV包括由AAVrh91eng(SEQ ID NO:19)的vp1、vp2和vp3中的至少一个编码的AAV衣壳。在某些实施例中,vp1、vp2和/或vp3是AAVrh91(SEQ ID NO:32)的全长衣壳蛋白。在其它实施例中,vp1、vp2和/或vp3具有N末端和/或C末端截短(例如,约1至约10个氨基酸的截短)。

[0122] 在某些实施例中,AAVrh91衣壳由以下的一种或多种表征:(1) AAVrh91衣壳蛋白,所述AAVrh91衣壳蛋白包括:选自以下的AAVrh91 vp1蛋白的异质群体:通过由编码SEQ ID NO:32的1至736的预测的氨基酸序列的核酸序列表达产生的vp1蛋白、由SEQ ID NO:31产生的vp1蛋白或由与编码SEQ ID NO:32的1至736的预测的氨基酸序列的SEQ ID NO:31至少70%相同的核酸序列产生的vp1蛋白,选自以下的AAVrh91 vp2蛋白的异质群体:通过由编码SEQ ID NO:32的至少约氨基酸138至736的预测的氨基酸序列的核酸序列表达产生的vp2蛋白、由包括SEQ ID NO:31的至少核苷酸412至2208的序列产生的vp2蛋白或由与编码SEQ ID NO:32的至少约氨基酸138至736的预测的氨基酸序列的SEQ ID NO:31的至少核苷酸412至2208至少70%相同的核酸序列产生的vp2蛋白,选自以下的AAVrh91 vp3蛋白的异质群体:通过由编码SEQ ID NO:32的至少约氨基酸203至736的预测的氨基酸序列的核酸序列表达产生的vp3蛋白、由包括SEQ ID NO:31的至少核苷酸607至2208的序列产生的vp3蛋白或由与编码SEQ ID NO:32的至少约氨基酸203至736的预测的氨基酸序列的SEQ ID NO:31的至少核苷酸607至2208至少70%相同的核酸序列产生的vp3蛋白;和/或(2) vp1蛋白的异质群体,所述vp1蛋白是编码SEQ ID NO:32的氨基酸序列的核酸序列的产物,vp2蛋白的异质群体,所述vp2蛋白是编码SEQ ID NO:32的至少约氨基酸138至736的氨基酸序列的核酸序

列的产物,以及vp3蛋白的异质群体,所述vp3蛋白是编码SEQ ID NO:32的至少氨基酸203至736的核酸序列的产物,其中:所述vp1、vp2和vp3蛋白含有具有氨基酸修饰的亚群体,所述氨基酸修饰包括SEQ ID NO:32中的天冬酰胺-甘氨酸对中的至少两个高度脱酰胺化的天冬酰胺(N),并且任选地进一步包括包含其它脱酰胺化的氨基酸的亚群体,其中脱酰胺化引起氨基酸变化;以及(B)AAVrh91衣壳中的载体基因组,所述载体基因组包括核酸分子,所述核酸分子包括AAV反向末端重复序列和非AAV核酸序列,所述非AAV核酸序列编码与引导产物在宿主细胞中的表达的序列可操作地连接的产物。

[0123] 在某些实施例中,AAVrh91衣壳由以下的一种或多种表征:(1)AAVrh91衣壳蛋白,所述AAVrh91衣壳蛋白包括:选自以下的AAVrh91 vp1蛋白的异质群体:通过由编码SEQ ID NO:32的1至736的预测的氨基酸序列的核酸序列表达产生的vp1蛋白、由SEQ ID NO:19产生的vp1蛋白或由与编码SEQ ID NO:32的1至736的预测的氨基酸序列的SEQ ID NO:19至少70%相同的核酸序列产生的vp1蛋白,选自以下的AAVrh91 vp2蛋白的异质群体:通过由编码SEQ ID NO:32的至少约氨基酸138至736的预测的氨基酸序列的核酸序列表达产生的vp2蛋白、由包括SEQ ID NO:19的至少核苷酸412至2208的序列产生的vp2蛋白或由与编码SEQ ID NO:32的至少约氨基酸138至736的预测的氨基酸序列的SEQ ID NO:19的至少核苷酸412至2208至少70%相同的核酸序列产生的vp2蛋白,选自以下的AAVrh91 vp3蛋白的异质群体:通过由编码SEQ ID NO:32的至少约氨基酸203至736的预测的氨基酸序列的核酸序列表达产生的vp3蛋白、由包括SEQ ID NO:19的至少核苷酸607至2208的序列产生的vp3蛋白或由与编码SEQ ID NO:32的至少约氨基酸203至736的预测的氨基酸序列的SEQ ID NO:19的至少核苷酸607至2208至少70%相同的核酸序列产生的vp3蛋白;和/或(2)vp1蛋白的异质群体,所述vp1蛋白是编码SEQ ID NO:32的氨基酸序列的核酸序列的产物,vp2蛋白的异质群体,所述vp2蛋白是编码SEQ ID NO:32的至少约氨基酸138至736的氨基酸序列的核酸序列的产物,以及vp3蛋白的异质群体,所述vp3蛋白是编码SEQ ID NO:32的至少氨基酸203至736的核酸序列的产物,其中:所述vp1、vp2和vp3蛋白含有具有氨基酸修饰的亚群体,所述氨基酸修饰包括SEQ ID NO:32中的天冬酰胺-甘氨酸对中的至少两个高度脱酰胺化的天冬酰胺(N),并且任选地进一步包括包含其它脱酰胺化的氨基酸的亚群体,其中脱酰胺化引起氨基酸变化;以及(B)AAVrh91衣壳中的载体基因组,所述载体基因组包括核酸分子,所述核酸分子包括AAV反向末端重复序列和非AAV核酸序列,所述非AAV核酸序列编码与引导产物在宿主细胞中的表达的序列可操作地连接的产物。

[0124] 在某些实施例中,AAVrh91 vp1、vp2和vp3蛋白含有具有氨基酸修饰的亚群体,所述氨基酸修饰包括SEQ ID NO:32中的天冬酰胺-甘氨酸对中的至少两个高度脱酰胺化的天冬酰胺(N),并且任选地进一步包括包含其它脱酰胺化的氨基酸的亚群体,其中脱酰胺化引起氨基酸变化。相对于SEQ ID NO:32的编号,在N-G对N57、N383和/或N512处观察到高水平的脱酰胺化。在其它残基中已经观察到脱酰胺化。在某些实施例中,AAVrh91可以具有其它脱酰胺化的残基,例如通常小于10%,和/或可以具有其它修饰,所述其它修饰包含磷酸化(例如,在存在的情况下,在约2%至约30%、或约2%至约20%、或约2%至约10%的范围内)(例如,在S149处)或氧化(例如,在~W22、~M211、W247、M403、M435、M471、W478、W503、~M537、~M541、~M559、~M599、M635和/或W695中的一个或多个处)。任选地,W可以氧化成犬尿氨酸。

[0125] 表A-AAVrh91脱酰胺化

基于VP1编号的AAVrh91脱酰胺化	脱酰胺化%
N57+脱酰胺化	65-90、70-95、80-95、75-100、80-100或90-100
N94+脱酰胺化	2-15或2-5
N303+脱酰胺化	2-15或5-10
N383+脱酰胺化	65-90、70-95、80-95、75-100、80-100或90-100
N497+脱酰胺化	2-15或5-10
N512+脱酰胺化	65-90、70-95、80-95、75-100、80-100或90-100
~N691+脱酰胺化	2-15、2-10或5-10

[0127] 在某些实施例中,在所提供的如使用胰蛋白酶使用质谱所确定的范围内,在上表中鉴定的位置中的一个或多个位置中修饰AAVrh91衣壳。在某些实施例中,如本文所述的修饰位置中的一个或多个位置或N之后的甘氨酸。残基数量基于本文所提供的AAVrh91序列。参见SEQ ID NO:32。

[0128] 在某些实施例中,AAVrh91衣壳包括:vp1蛋白的异质群体,所述vp1蛋白是编码SEQ ID NO:32的氨基酸序列的核酸序列的产物;vp2蛋白的异质群体,所述vp2蛋白是编码SEQ ID NO:32的至少约氨基酸138至736的氨基酸序列的核酸序列的产物;以及vp3蛋白的异质群体,所述vp3蛋白是编码SEQ ID NO:32的至少氨基酸203至736的核酸序列的产物。

[0129] 在某些实施例中,经修饰的AAVrh91核酸序列用于产生具有比天然AAVrh91衣壳脱酰胺化程度更低的衣壳的突变rAAV。此类突变rAAV可以具有降低的免疫原性和/或增加储存时的稳定性,特别是以悬浮液形式储存时的稳定性。

[0130] 一方面,提供了一种重组AAV(rAAV)。rAAV包含来自腺相关病毒rh91的AAV衣壳和包装在所述AAV衣壳中的载体基因组,所述载体基因组包括AAV反向末端重复序列(ITR)、SEQ ID NO:14的GLP-2受体激动剂的编码序列以及引导GLP-2受体激动剂表达的调控序列。

[0131] 在某些实施例中,AAV68衣壳的另外的特征在于以下中的一种或多种。AAV hu68衣壳蛋白包括:通过由编码SEQ ID NO:55的1至736的预测的氨基酸序列的核酸序列表达产生的AAVhu68 vp1蛋白、由SEQ ID NO:53或54产生的vp1蛋白或由与编码SEQ ID NO:55的1至736的预测的氨基酸序列的SEQ ID NO:53或54至少70%相同的核酸序列产生的vp1蛋白;通过由编码SEQ ID NO:55的至少约氨基酸138至736的预测的氨基酸序列的核酸序列表达产生的AAVhu68 vp2蛋白、由包括SEQ ID NO:53或54的至少核苷酸412至2211的序列产生的vp2蛋白或由与编码SEQ ID NO:55的至少约氨基酸138至736的预测的氨基酸序列的SEQ ID NO:53或54的至少核苷酸412至2211至少70%相同的核酸序列产生的vp2蛋白;和/或通过由编码SEQ ID NO:55的至少约氨基酸203至736的预测的氨基酸序列的核酸序列表达产生的AAVhu68 vp3蛋白、由包括SEQ ID NO:53或54的至少核苷酸607至2211的序列产生的vp3蛋白或由与编码SEQ ID NO:55的至少约氨基酸203至736的预测的氨基酸序列的SEQ ID NO:53或54的至少核苷酸607至2211至少70%相同的核酸序列产生的vp3蛋白。

[0132] 另外地或可替代地,提供了一种AAV衣壳,所述AAV衣壳包括任选地包括位置157处的缬氨酸的vp1蛋白的异质群体、任选地包括位置157处的缬氨酸的vp2蛋白的异质群体以及vp3蛋白的异质群体,其中基于SEQ ID NO:55的vp1衣壳的编号,vp1和vp2蛋白的至少一个亚群体包括位置157处的缬氨酸并且任选地进一步包括位置67处的谷氨酸。另外地或可

替代地,提供了一种AAVhu68衣壳,所述AAVhu68衣壳包括:vp1蛋白的异质群体,所述vp1蛋白是编码SEQ ID NO:55的氨基酸序列的核酸序列的产物;vp2蛋白的异质群体,所述vp2蛋白是编码SEQ ID NO:55的至少约氨基酸138至736的氨基酸序列的核酸序列的产物;以及vp3蛋白的异质群体,所述vp3蛋白是编码SEQ ID NO:55的至少氨基酸203至736的核酸序列的产物,其中:所述vp1、vp2和vp3蛋白含有具有氨基酸修饰的亚群体。

[0133] AAVhu68 vp1、vp2和vp3蛋白通常表示为由编码SEQ ID NO:55的全长vp1氨基酸序列(氨基酸1至736)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:编码SEQ ID NO:55的AAVhu68 vp3氨基酸序列(约aa 203至736)的核酸序列,所述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:53或54的约nt 607至约nt 2211)或与SEQ ID NO:53或54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%)相同的序列,所述SEQ ID NO:53或54编码SEQ ID NO:55的aa 203至736。另外地或可替代地,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:编码SEQ ID NO:55的AAVhu68 vp2氨基酸序列(约aa 138至736)的核酸序列,所述AAVhu68 vp2氨基酸序列没有vp1独特区(约aa 1至约137)或与其互补的链,即对应的mRNA或tRNA(SEQ ID NO:53或54的nt 412至2212)或与SEQ ID NO:53或54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%)相同的序列,所述SEQ ID NO:53或54编码SEQ ID NO:55的约aa 138至736。

[0134] 如本文所描述的,rAAVhu68具有在从编码SEQ ID NO:55的vp1氨基酸序列的AAVhu68核酸中表达衣壳的产生系统中产生的rAAVhu68衣壳,以及任选地例如编码不含vp1和/或vp2独特区的vp3蛋白的另外的核酸序列。使用单个核酸序列vp1从生产中产生的rAAVhu68产生vp1蛋白、vp2蛋白和vp3蛋白的异质群体。更具体地,AAVhu68衣壳含有具有来自SEQ ID NO:55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。

[0135] 在一个实施例中,AAVhu68 vp1核酸序列具有SEQ ID NO:53或54的序列或与其互补的链,例如对应的mRNA或tRNA。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,还提供了编码SEQ ID NO:55的AAVhu68 vp3氨基酸序列(约aa 203至736)的核酸序列,所述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:53或54的约nt 607至约nt 2211)。在某些实施例中,还提供了编码SEQ ID NO:55的AAVhu68 vp2氨基酸序列(约aa 138至736)的核酸序列,所述AAVhu68 vp2氨基酸序列没有vp1独特区(约aa 1至约137)或与其互补的链,即对应的mRNA或tRNA(SEQ ID NO:53或54的nt 412至2211)。

[0136] 然而,可以选择编码SEQ ID NO:55的氨基酸序列的其它核酸序列用于产生rAAVhu68衣壳。在某些实施例中,核酸序列具有SEQ ID NO:53或54的核酸序列或与编码SEQ ID NO:55的SEQ ID NO:53或54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%相同的序列。在某些实施例中,核酸序列具有SEQ ID

N0:53或54的核酸序列或与编码SEQ ID N0:55的vp2衣壳蛋白(约aa 138至736)的SEQ ID N0:53或54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%、至少99%相同的序列。在某些实施例中,核酸序列具有SEQ ID N0:53或54的约nt 607至约nt 2211的核酸序列或与编码SEQ ID N0:55的vp3衣壳蛋白(约aa 203至736)的nt SEQ ID N0:53或54至少70%至99%、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。

[0137]

脱酰胺化 基于预测的 AAVHu68[SEQ ID NO: 55]	基于 AAVhu68 衣壳中的 VP1/VP2/VP3 蛋白的平均%	
脱酰胺化残基 +1 (相邻的 AA)	宽范围的百分比 (%)	窄范围 (%)
N57 (N-G)	78 至 100%	80 至 100、85 至 97
N66 (N-E)	0 至 5	0、1 至 5
N94 (N-H)	0 至 15	0、1 至 15、5 至 12、8
N113 (N-L)	0 至 2	0、1 至 2
约 N253 (N-N)	10 至 25	15 至 22
Q259 (Q-I)	8 至 42	10 至 40、20 至 35
约 N270 (N-D)	12 至 30	15 至 28
约 N304 (N-N) (位置 303 也为 N)	0 至 5	1 至 4
N319 (N-I)	0 至 5	0、1 至 5、1 至 3
N329* (N-G) * (位置 328 也为 N)	65 至 100	70 至 95、85 至 95、80 至 100、 85 至 100
N336 (N-N)	0 至 100	0、1 至 10、25 至 100、30 至 100、30 至 95
约 N409 (N-N)	15 至 30	20 至 25
N452 (N-G)	75 至 100	80 至 100、90 至 100、95 至 100
N477 (N-Y)	0 至 8	0、1 至 5
N512 (N-G)	65 至 100	70 至 95、85 至 95、80 至 100、 85 至 100
约 N515 (N-S)	0 至 25	0、1 至 10、5 至 25、15 至 25

[0138]	脱酰胺化 基于预测的 AAVHu68[SEQ ID NO: 55]	基于 AAVhu68 衣壳中的 VP1/VP2/VP3 蛋白的平均%	
	脱酰胺化残基 + 1 (相邻的 AA)	宽范围的百分比 (%)	窄范围 (%)
	约 Q599 (Asn-Q-Gly)	1 至 20	2 至 20、5 至 15
	N628 (N-F)	0 至 10	0、1 至 10、2 至 8
	N651 (N-T)	0 至 3	0、1 至 3
	N663 (N-K)	0 至 5	0、1 至 5、2 至 4
	N709 (N-N)	0 至 25	0、1 至 22、15 至 25
	N735	0 至 40	0、1 至 35、5 至 50、20 至 35

[0139] 在某些实施例中,AAVhu68衣壳的特征在于具有衣壳蛋白,其中基于SEQ ID NO:55的氨基酸序列的编号,至少45%的N残基在位置N57、N329、N452和/或N512中的至少一个位置处脱酰胺化。在某些实施例中,在这些N-G位置(即,基于SEQ ID NO:55的氨基酸序列的编号,N57、N329、N452和/或N512)中的一个或多个位置处的N残基的至少约60%、至少约70%、至少约80%或至少约90%被脱酰胺化。在这些和其它实施例中,AAVhu68衣壳的特征进一步在于具有蛋白质群体,其中N残基的约1%至约20%在以下位置中的一个或多个位置处具有脱酰胺化:基于SEQ ID NO:55的氨基酸序列的编号,N94、N253、N270、N304、N409、N477和/或Q599。在某些实施例中,AAVhu68包括vp1、vp2和/或vp3蛋白的至少一个亚群体,基于SEQ ID NO:55的氨基酸序列的编号或其组合,所述蛋白在以下位置中的一个或多个位置处被脱酰胺化:N35、N57、N66、N94、N113、N252、N253、Q259、N270、N303、N304、N305、N319、N328、N329、N336、N409、N410、N452、N477、N515、N598、Q599、N628、N651、N663、N709、N735。在某些实施例中,衣壳蛋白可以具有一个或多个酰胺化的氨基酸。

[0140] 在另一个实施例中,提供了一种重组腺相关病毒(rAAV),其具有AAVhu68衣壳和载体基因组,其中(a)AAV hu68衣壳包括AAVhu68 vp1蛋白的异质群体、AAVhu68 vp2蛋白的异质群体;以及AAVhu68 vp3蛋白的异质群体,其中所述异质AAVhu68 vp1、AAVhu68 vp2和AAVhu68 vp3蛋白含有具有氨基酸修饰的亚群体,所述氨基酸修饰包括如使用质谱法确定的SEQ ID NO:55的N57、N329、N452、N512中的两个或更多个中的天冬酰胺-甘氨酸对中的至少两个天冬酰胺(N)中50%至100%的脱酰胺化,并且任选地进一步包括包含其它脱酰胺化的氨基酸的亚群体,其中脱酰胺化导致氨基酸变化,其中脱酰胺化的天冬酰胺被脱酰胺化为天冬氨酸、异天冬氨酸、相互转化的天冬氨酸/异天冬氨酸对或其组合,其中AAVhu68衣壳进一步包括具有以下一种或多种的亚群体:

[0141] 所述vp1蛋白的基于SEQ ID NO:55的编号定位于位置N57处的天冬酰胺-甘氨酸对中的脱酰胺化的至少65%的天冬酰胺(N);

[0142] 所述vp1、v2和vp3蛋白的基于SEQ ID NO:55的氨基酸序列的残基编号位于位置N329中的天冬酰胺-甘氨酸对中的脱酰胺化的至少75%的N;

[0143] 所述vp1、v2和vp3蛋白的基于SEQ ID NO:55的氨基酸序列的残基编号位于位置

N452中的天冬酰胺-甘氨酸对中的脱酰胺化的至少50%的N;和/或

[0144] 所述vp1、v2和vp3蛋白的基于SEQ ID NO:55的氨基酸序列的残基编号位于位置N512中的天冬酰胺-甘氨酸对中的脱酰胺化的至少75%的N;以及

[0145] 所述AAVhu68衣壳中的载体基因组,所述载体基因组包括核酸分子,所述核酸分子包括AAV反向末端重复序列和非AAV核酸序列,所述非AAV核酸序列对如本文所述的可操作地连接到序列的GLP-2融合物进行编码,所述序列引导所述GLP-2融合物在靶细胞中的表达。

[0146] 在一个实施例中,rAAV是scAAV。缩写“sc”是指自身互补的。“自身互补的AAV”是指具有其中重组AAV核酸序列携带的编码区已经被设计成形成分子内双链DNA模板的表达盒的质粒或载体。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,“自身互补的重组腺相关病毒(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》,(2001年8月),第8卷,第16期,第1248-1254页。在例如美国专利第6,596,535号、第7,125,717号和第7,456,683号中描述了自身互补的AAV,所述美国专利通过引用整体并入本文。

[0147] 在一个实施例中,编码本文所述的GLP-2构建体的核酸序列被工程化到任何合适的遗传元件中,例如,裸DNA、噬菌体、转座子、粘粒、RNA分子(例如,mRNA)、附加体等,所述遗传元件将其上携带的GLP-2序列转移至宿主细胞,例如,用于在包装宿主细胞中产生携带DNA或RNA的纳米颗粒、病毒载体和/或用于递送到受试者的宿主细胞。在一个实施例中,遗传元件是质粒。所选择的基因元件可以通过任何合适的方法递送,包含转染、电穿孔、脂质体递送、膜融合技术、高速DNA包被的团粒、病毒感染和原生质体融合。用于制备此类构建体的方法对于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)。

[0148] 如本文所使用的,术语“宿主细胞”可以指包装细胞系,其中由生产质粒产生载体(例如,重组AAV或rAAV)。在替代方案中,术语“宿主细胞”可以指期望表达本文所述的基因产物的任何靶细胞。因此,“宿主细胞”是指原核或真核细胞(例如,细菌细胞、人类细胞或昆虫细胞),所述原核或真核细胞含有通过任何方式(例如,电穿孔、磷酸钙沉淀、显微注射、转化、病毒感染、转染、脂质体递送、膜融合技术、高速DNA包被的团粒、病毒感染和原生质体融合)引入到细胞中的外源性或异源性DNA。在本文中的某些实施例中,术语“宿主细胞”是指用于体外评估本文所述的组合物的各种哺乳动物物种的细胞的培养物。在本文中的其它实施例中,术语“宿主细胞”是指用于产生和包装病毒载体或重组病毒的细胞。在另外的实施例中,术语“宿主细胞”是肠细胞、小肠细胞、胰细胞、肝脏细胞。

[0149] 如本文所使用的,术语“靶细胞”是指期望表达异源性核酸序列或蛋白质的任何靶细胞。在某些实施例中,靶细胞是肝脏细胞。在其它实施例中,靶细胞是肌细胞。

[0150] 在一个实施例中,提供了rAAV,所述rAAV包括包含表达盒的载体基因组,其中所述表达盒包括CMV启动子,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人

FKBP12-雷帕霉素相关蛋白 (FRAP) 的FKBP12-雷帕霉素结合 (FRB) 结构域、GT2A_V1肽、ZFHD1 DNA结合结构域、三个FKBP亚基、hGH poly A、12XZFHD、最小sIL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。在另一个实施例中,提供了rAAV,所述rAAV包括包含表达盒的载体基因组,其中所述表达盒包括CMV启动子,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白 (FRAP) 的FKBP12-雷帕霉素结合 (FRB) 结构域、GT2A_V2肽、ZFHD1 DNA结合结构域、三个FKBP亚基、hGH poly A、12XZFHD、最小sIL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。

[0151] 在一个实施例中,提供了rAAV,所述rAAV包括包含表达盒的载体基因组,其中所述表达盒包括CB7启动子、鸡 β -肌动蛋白内含子、SEQ ID NO:13、15、17、19、21或23的融合蛋白的编码序列、兔珠蛋白poly A和3' ITR。在另一个实施例中,提供了rAAV,所述rAAV包括包含表达盒的载体基因组,其中所述表达盒包括CMV启动子,激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白 (FRAP) 的FKBP12-雷帕霉素结合 (FRB) 结构域、IRES、ZFHD1 DNA结合结构域、三个FKBP亚基、hGH poly A、8XZFHD、最小sIL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。

[0152] 在另一个实施例中,提供了一种二载体诱导型系统。在一个实施例中,提供了第一rAAV,所述第一rAAV包括包含表达盒的载体基因组,其中所述表达盒包括12XZFHD、最小IL2启动子、SEQ ID NO:13、15、17、19、21或23的GLP-2融合蛋白的编码序列和兔 β 珠蛋白polyA。可以包含填充序列以增加载体的包装大小。第二rAAV包括包含表达盒的载体基因组,其中所述表达盒包括多核苷酸,所述多核苷酸包括CMV启动子、内含子,所述激活结构域是与来自人的NF- κ B的p65亚基的羧基末端融合的人FKBP12-雷帕霉素相关蛋白 (FRAP) 的FKBP12-雷帕霉素结合 (FRB) 结构域、IRES、ZFHD1 DNA结合结构域和polyA。

[0153] 在一个实施例中,rAAV包含载体基因组,所述载体基因组包含表达盒,所述表达盒包含多核苷酸,所述多核苷酸包括CB7启动子、鸡 β -肌动蛋白内含子、SEQ ID NO:13、15、17、19、21或23的融合蛋白的编码序列和兔珠蛋白poly A。

[0154] 将表达盒包装到AAV病毒颗粒中所需的最小序列是与衣壳具有相同AAV起源或具有不同AAV起源 (以产生AAV假型) 的AAV 5' 和3' ITR。在一个实施例中,来自AAV2的ITR序列或其删除版本 (Δ ITR) 是为方便起见采用的,并且为了加快监管审批。然而,可以选择来自其它AAV来源的ITR。优选地,ITR的来源与Rep蛋白的来源相同,所述Rep蛋白以反式提供用于生产。通常,AAV载体的表达盒包括AAV 5' ITR、GLP-2融合蛋白编码序列和任何调控序列以及AAV 3' ITR。然而,这些元件的其它构型可以是合适的。已经描述了被称为 Δ ITR的5' ITR的缩短版本,其中缺失了D序列和末端解析位点 (trs)。在其它实施例中,使用了全长AAV 5' 和3' ITR。

[0155] 为了将表达盒包装到病毒粒子中,ITR是在与基因相同的构建体中顺式所需的唯一AAV组分。在一个实施例中,用于复制 (rep) 和/或衣壳 (cap) 的编码序列从AAV基因组中移除并以反式供应或由包装细胞系供应,以产生AAV载体。例如,如上所述,假型AAV可以含有来自与AAV衣壳的来源不同的来源的ITR。在一个实施例中,可以利用嵌合AAV衣壳。可以选择仍其它的AAV组分。此类AAV序列的来源在本文中有所描述并且还可以从学术、商业或公共来源分离或获得 (例如,维吉尼亚州马纳萨斯的美国典型培养物保藏中心 (the American

Type Culture Collection, Manassas, VA)。AAV序列可以通过合成或其它合适的方式通过参考公开的序列(如在文献中或在如例如GenBank®、PubMed®等数据库中可获得的公开的序列)而获得。

[0156] 用于产生和分离适合于向受试者递送的AAV病毒载体的方法在本领域中是已知的。[参见例如,美国专利7790449;美国专利7282199;WO 2003/042397;WO 2005/033321;WO 2006/110689;以及US 7588772 B2]。在一个系统中,用编码侧接有ITR的转基因的构建体和编码rep和cap的构建体瞬时转染生产细胞系。在第二系统中,用编码侧接有ITR的转基因的构建体瞬时转染稳定供应rep和cap的包装细胞系。在这些系统的每种系统中,响应于用辅助腺病毒或疱疹病毒感染而产生AAV病毒粒子,从而需要从污染的病毒中分离rAAV。最近,已开发了不需要用辅助病毒感染来回收AAV的系统—通过所述系统还反式地提供所需辅助功能(即,腺病毒E1、E2a、VA和E4或疱疹病毒UL5、UL8、UL52和UL29,以及疱疹病毒聚合酶)。在这些较新的系统中,可以通过用编码所需辅助功能的构建体瞬时转染细胞来提供辅助功能,或者所述细胞可以被工程化成稳定地含有编码辅助功能的基因,所述基因的表达可以被控制在转录水平或转录后水平下。在又一个系统中,通过用基于杆状病毒的载体进行感染来将侧接有ITR的转基因和rep/cap基因引入到昆虫细胞中。关于这些产生系统的综述,通常参见例如Zhang等人,2009,“用于大规模重组腺相关病毒产生的腺病毒-腺相关病毒杂合体(Adenovirus-Adeno-associated virus hybrid for large-scale recombinant Adeno-associated virus production)”,《人类基因疗法(Human Gene Therapy)》20:922-929,所述参考文献中的每个参考文献的内容通过引用整体并入本文。在以下美国专利中也描述了制备和使用这些及其它AAV产生系统的方法,所述美国专利中的每个美国专利的内容通过引用整体并入本文:5,139,941;5,741,683;6,057,152;6,204,059;6,268,213;6,491,907;6,660,514;6,951,753;7,094,604;7,172,893;7,201,898;7,229,823和7,439,065。通常参见例如,Grieger和Samulski,2005,“腺相关病毒作为基因疗法载体:载体研发、产生及临床应用(Adeno-associated virus as a gene therapy vector:Vector development,production and clinical applications)”,《生物化学工程/生物技术进展(Adv.Biochem.Engin/Biotechnol.)》99:119-145;Buning等人,2008,“腺相关病毒载体技术的最新进展(Recent developments in Adeno-associated virus vector technology)”,《基因医学杂志(J.Gene Med.)》10:717-733;以及下文引用的参考文献,所述参考文献中的每个参考文献通过引用整体并入本文。用于构建本发明的任何实施例的方法对于核酸操作技术人员是已知的并且包含遗传工程、重组工程以及合成技术。参见例如,Green和Sambrook等人,《分子克隆:实验室手册》,纽约市冷泉港的冷泉港实验室出版社(2012)。类似地,产生rAAV病毒粒子的方法是熟知的,并且对合适的方法的选择不是对本发明的限制。参见例如,K.Fisher等人,(1993)《病毒学杂志》,70:520-532和美国专利第5,478,745号。

[0157] 本文所述的rAAV包括具有包装在内部的载体基因组的所选择的衣壳。载体基因组(或rAAV基因组)包括5'和3' AAV反向末端重复序列(ITR)、编码融合蛋白的多核苷酸序列以及引导编码融合蛋白的多核苷酸序列插入宿主细胞的基因组的调控序列。在一个实施例中,载体基因组是SEQ ID NO:16中所示的序列或与其共享至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或至少99%同一性的序列。

[0158] 如本文所使用的,“载体基因组”是指包装在形成病毒颗粒的细小病毒(例如,rAAV)衣壳内部的核酸序列。此类核酸序列含有AAV反向末端重复序列(ITR)。在本文的实例中,载体基因组至少含有从5'至3'的AAV 5' ITR、编码序列(即,转基因)和AAV 3' ITR。可以选择来自AAV2(与衣壳来源不同的AAV)或除了非全长ITR之外的ITR。在某些实施例中,ITR来自与在产生或反式补充AAV期间提供rep功能的AAV来源相同的AAV。另外,可以使用其它ITR,例如,自身互补的(scAAV) ITR。单链AAV和自身互补的(sc)AAV两者都涵盖在rAAV内。转基因是与载体序列异源的核酸编码序列,所述核酸编码序列编码多肽、蛋白质、功能性RNA分子(例如,miRNA、miRNA抑制剂)或其它所关注基因产物。核酸编码序列以允许转基因在靶组织的细胞中转录、翻译和/或表达的方式与调控组分可操作地连接。在本文中更详细地讨论了载体基因组的合适组分。在一个实例中,“载体基因组”至少含有从5'至3'的载体特异性序列、编码GLP-2构建体的核酸序列,所述核酸序列与调控控制序列(所述调控控制序列引导其在靶细胞中表达)可操作地连接,其中载体特异性序列可以是将载体基因组特异性地包装到病毒载体衣壳或包膜蛋白中的末端重复序列。例如,AAV反向末端重复序列用于包装到AAV和某些其它细小病毒衣壳中。

[0159] 载体的AAV序列通常包括顺式作用的5'和3'反向末端重复序列(参见例如,B.J.Carter,《细小病毒手册(Handbook of Parvoviruses)》,P.Tijsser编辑,CRC出版社,第155-168页(1990))。ITR序列的长度为约145bp。优选地,在分子中使用编码ITR的基本上整个序列,但是允许对这些序列进行某种程度的微小修饰。修饰这些ITR序列的能力在本领域的技术范围内。(参见例如,文本,如Sambrook等人,《分子克隆:实验室手册》,第2版,纽约的冷泉港实验室(1989);以及K.Fisher等人,《病毒学杂志》,70:520532(1996))。在本发明中采用的此类分子的实例是含有转基因的“顺式作用”质粒,其中所选择的转基因序列和相关调控元件侧接有5'和3'AAV ITR序列。在一个实施例中,ITR来自与供应衣壳的AAV不同的AAV。在一个实施例中,来自AAV2的ITR序列。然而,可以选择来自其它AAV来源的ITR。已经描述了被称为 Δ ITR的5' ITR的缩短版本,其中缺失了D序列和末端解析位点(trs)。在某些实施例中,载体基因组包含130个碱基对的缩短的AAV2 ITR,其中外部A元件缺失。不希望受理论束缚,据信在使用内部(A')元件作为模板的载体DNA扩增期间,缩短的ITR回复到145个碱基对的野生型长度。在其它实施例中,使用了全长AAV 5'和3' 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。然而,这些元件的其它构型可以是合适的。

[0160] 任选地,本文所述的GLP-2构建体可以通过除了rAAV之外的病毒载体递送。此类其它病毒载体可以包含适用于基因疗法的任何病毒,包括但不限于腺病毒;疱疹病毒;慢病毒;逆转录病毒等。适当地,在产生这些其它载体中的一种载体时,其产生为复制缺陷型病毒载体。

[0161] “复制缺陷型病毒”或“病毒载体”是指其中含有所关注基因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,所述合成或人工病毒颗粒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编码复制所需的酶的基因(基因组可以被工程化成“无肠的(gutless)”—仅含有所关注转基因,其侧接有扩增和包装人工基因组所需的信号),但是这些基因可以在产生期间供应。因此,这被认为可以安全地用于

基因疗法,因为除非存在复制所需的病毒酶,否则不会发生通过子代病毒粒子进行的复制和感染。

[0162] 还提供了包含本文所述的病毒载体构建体的组合物。本文所述的药物组合物被设计成用于通过任何合适的途径或不同途径的组合递送到有需要的受试者。直接递送到肝脏(任选地通过静脉内、通过肝动脉或通过移植)、口服施用途径、吸入施用途径、鼻内施用途径、气管内施用途径、动脉内施用途径、眼内施用途径、静脉内施用途径、肌内施用途径、皮下施用途径、皮内施用途径以及其它母体施用途径。本文所述的病毒载体可以以单一组合物或多种组合物的形式递送。任选地,可以递送两种或更多种不同的AAV,或多种病毒[参见例如,WO 2011/126808和WO 2013/049493]。在另一个实施例中,多种病毒可以含有不同的复制缺陷型病毒(例如,AAV和腺病毒)。在一个实施例中,施用是肌内进行的。在另一个实施例中,施用是静脉内进行的。

[0163] 复制缺陷型病毒可以与生理上可接受的载剂一起调配,用于基因转移和基因疗法应用。在AAV病毒载体的情况下,基因组拷贝(“GC”)的定量可以用作调配物中所含有的剂量的量度。可以使用本领域已知的任何方法来确定本发明的复制缺陷型病毒组合物的基因组拷贝(GC)数量。用于进行AAV GC数量滴定的一种方法如下:经纯化的AAV载体样品首先用DNA酶进行处理,以消除来自产生过程的未衣壳化的AAV基因组DNA或污染质粒DNA。然后使核酸酶抗性颗粒经受热处理,以从衣壳中释放基因组。然后使用靶向病毒基因组特定区域(通常为polyA信号)的引物/探针组通过实时PCR来定量经释放的基因组。用于确定基因组拷贝的另一种合适的方法是定量PCR(qPCR),特别是经优化的qPCR或数字液滴PCR[Lock Martin等人,《人类基因疗法方法(Human Gene Therapy Methods.)》2014年4月,25(2):115-125.doi:10.1089/hgtb.2013.131,2013年12月13日编辑之前在线发布]。

[0164] 而且,复制缺陷型病毒组合物可以以一定剂量单位调配以含有在约 1.0×10^9 GC至约 1.0×10^{15} GC的范围内的复制缺陷型病毒量。在另一个实施例中,这一量的病毒基因组可以以分剂量的形式递送。在一个实施例中,对于约70kg的平均人受试者,剂量为约 1.0×10^{10} GC至约 3.0×10^{14} GC。在另一个实施例中,剂量约 1×10^9 GC。例如,AAV病毒的剂量可以为约 1×10^{10} GC、 1×10^{11} GC、约 5×10^{11} GC、约 1×10^{12} GC、约 5×10^{12} GC或约 1×10^{13} GC。在另一个实施例中,对于人受试者,剂量为约 1.0×10^9 GC/kg至约 3.0×10^{14} GC/kg。在另一个实施例中,剂量约 1×10^9 GC/kg。例如,AAV病毒的剂量可以为约 1×10^{10} GC/kg、 1×10^{11} GC/kg、约 5×10^{11} GC/kg、约 1×10^{12} GC/kg、约 5×10^{12} GC/kg或约 1×10^{13} GC/kg。在一个实施例中,构建体可以以 $1 \mu\text{L}$ 至约100mL的体积递送。如本文所使用的,术语“剂量”或“量”可以指在治疗过程中向受试者递送的总剂量或量或以单一单位(或多单位或分剂量)施用递送的剂量或量。

[0165] 可以根据所公开的方法将上文所述的重组载体递送到宿主细胞。可以将优选地悬浮于生理学相容的载剂中的rAAV施用于包含人在内的期望的受试者。鉴于转移病毒所针对的适应症,本领域技术人员可以容易地选择合适的载剂。例如,一种合适的载剂包含盐水,其可以与各种缓冲溶液(例如,磷酸盐缓冲盐水)一起调配。其它示例性载剂包含无菌盐水、乳糖、蔗糖、磷酸钙、明胶、葡聚糖、琼脂、果胶、花生油、芝麻油和水。对载剂的选择不是对本发明的限制。

[0166] 在另一个实施例中,组合物包含载剂、稀释剂、赋形剂和/或佐剂。在某些实施例中,为了施用于人类患者,将rAAV合适地悬浮于含有盐水、表面活性剂和药学上和/或生理

学上相容的盐或盐的混合物的水溶液中。合适地,将调配物调整至生理上可接受的pH,例如,在pH 6至9、或pH 6.0至7.5、或pH 6.2至7.7、或pH 6.5至7.5、pH 7.0至7.7或pH 7.2至7.8或约7.0的范围内。在某些实施例中,将调配物的pH调整至约6.0、约6.1、约6.2、约6.3、约6.4、约6.5、约6.6、约6.7、约6.8、约6.9、约7.0、约7.1、约7.2、约7.3、约7.4、约7.5、约7.6、约7.7或约7.8。在某些实施例中,可能期望的pH为约7.28至约7.32、约6.0至约7.5、约6.2至约7.7、约7.5至约7.8、约6.0、约6.1、约6.2、约6.3、约6.4、约6.5、约6.6、约6.7、约6.8、约6.9、约7.0、约7.1、约7.2、约7.3、约7.4、约7.5、约7.6、约7.7或约7.8。在某些实施例中,对于静脉内递送,可能期望的pH为约6.8至约7.2。然而,可以选择最宽范围和这些子范围内的其它pH用于其它递送途径。

[0167] 任选地,除了rAAV和/或变体以及载剂之外,本发明的组合物还可以含有其它常规药物成分,如防腐剂或化学稳定剂。合适的示例性防腐剂包含氯丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸甲酯、乙基香草醛、甘油、苯酚和对氯苯酚。合适的化学稳定剂包含明胶和白蛋白。

[0168] 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、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲液、载剂溶液、悬浮液、胶质物等。此类介质和药剂用于药物活性物质的用途在本领域中是熟知的。还可以将补充性活性成分掺入到组合物中。短语“药学上可接受的”是指当施用于宿主时不会产生过敏或类似不良反应的分子实体和组合物。如脂质体、纳米胶囊、微颗粒、微球、脂质颗粒、囊泡等递送媒剂可以用于将本发明的组合物引入到合适的宿主细胞中。具体地,rAAV载体递送的转基因可以被调配成用于递送或封装在脂质颗粒、脂质体、囊泡、纳米球或纳米颗粒等中。

[0169] 在一个实施例中,组合物包含适合于递送到受试者的最终调配物,所述组合物是例如缓冲到生理上相容的pH和盐浓度的水性液体悬浮液。任选地,调配物中存在一种或多种表面活性剂。在另一个实施例中,组合物可以以浓缩物的形式运输,所述浓缩物被稀释以施用于受试者。在其它实施例中,组合物可以在施用前冻干并重构。

[0170] 可以从无毒的非离子表面活性剂中选择合适的表面活性剂或表面活性剂的组合。在一个实施例中,选择终止于伯羟基的双官能嵌段共聚物表面活性剂,例如Pluronic®F68 [BASF],也被称为泊洛沙姆(Poloxamer) 188,其具有中性pH,平均分子量为8400。可以选择其它表面活性剂和其它泊洛沙姆,即非离子型三嵌段共聚物,所述非离子型三嵌段共聚物由侧接有聚氧乙烯(聚(环氧乙烷))的两个亲水链的聚氧丙烯(聚(环氧丙烷))的中心疏水链、SOLUTOL HS15(聚乙二醇-15羟基硬脂酸酯)、LABRASOL(聚氧辛酸甘油酯)、聚氧10油醚、TWEEN(聚氧乙烯山梨聚糖脂肪酸酯)、乙醇和聚乙二醇构成。在一个实施例中,调配物含有泊洛沙姆。这些共聚物通常以字母“P”(对于泊洛沙姆)命名,后跟三个数字:前两位数字x 100给出了聚氧丙烯核的近似分子质量,并且最后一位数字x 10给出了聚氧乙烯含量的百分比。在一个实施例中,选择了泊洛沙姆188。表面活性剂可以以悬浮液的至多约0.0005%至约0.001%的量存在。

[0171] 载体的剂量主要取决于如所治疗的病状、患者的年龄、体重和健康状况等因素,并且因此在患者之间可能有所不同。例如,病毒载体的治疗有效人剂量通常在约25微升至约1000微升至约100mL溶液的范围,所述溶液含有浓度为约 1×10^9 至 1×10^{16} 个基因组病毒载体(以治疗平均体重为70kg的受试者),包含所述范围内的整数或分数量,并且对于人类

患者,优选地为 1.0×10^{12} GC至 1.0×10^{13} GC。本发明的组合物可以以约 $0.1 \mu\text{L}$ 至约 10mL 的体积递送,包含所述范围内的所有数字,这取决于待治疗区域的大小、所使用的病毒滴度、施用途径以及所述方法的期望效果。在一个实施例中,体积为约 $50 \mu\text{L}$ 。在另一个实施例中,体积为约 $70 \mu\text{L}$ 。在另一个实施例中,体积为约 $100 \mu\text{L}$ 。在另一个实施例中,体积为约 $125 \mu\text{L}$ 。在另一个实施例中,体积为约 $150 \mu\text{L}$ 。在另一个实施例中,体积为约 $175 \mu\text{L}$ 。在又一个实施例中,体积为约 $200 \mu\text{L}$ 。在另一个实施例中,体积为约 $250 \mu\text{L}$ 。在另一个实施例中,体积为约 $300 \mu\text{L}$ 。在另一个实施例中,体积为约 $450 \mu\text{L}$ 。在另一个实施例中,体积为约 $500 \mu\text{L}$ 。在另一个实施例中,体积为约 $600 \mu\text{L}$ 。在另一个实施例中,体积为约 $750 \mu\text{L}$ 。在另一个实施例中,体积为约 $850 \mu\text{L}$ 。在另一个实施例中,体积为约 $1000 \mu\text{L}$ 。在另一个实施例中,体积为约 1.5mL 。在另一个实施例中,体积为约 2mL 。在另一个实施例中,体积为约 2.5mL 。在另一个实施例中,体积为约 3mL 。在另一个实施例中,体积为约 3.5mL 。在另一个实施例中,体积为约 4mL 。在另一个实施例中,体积为约 5mL 。在另一个实施例中,体积为约 5.5mL 。在另一个实施例中,体积为约 6mL 。在另一个实施例中,体积为约 6.5mL 。在另一个实施例中,体积为约 7mL 。在另一个实施例中,体积为约 8mL 。在另一个实施例中,体积为约 8.5mL 。在另一个实施例中,体积为约 9mL 。在另一个实施例中,体积为约 9.5mL 。在另一个实施例中,体积为约 10mL 。

[0172] 在一些实施例中,在组合物中,携带编码期望转基因的核酸序列的重组腺相关病毒在调控序列的控制下的浓度令人期望地在约 10^7 和 10^{14} 个基因组拷贝/毫升(GC/mL)的范围内。

[0173] 在一个实施例中,组合物中的rAAV的剂量为约 1.0×10^9 GC/kg体重至约 1.5×10^{13} GC/kg。在一个实施例中,剂量为约 1.0×10^{10} GC/kg。在一个实施例中,剂量为约 1.0×10^{11} GC/kg。在一个实施例中,剂量为约 1.0×10^{12} GC/kg。在一个实施例中,剂量为约 5.0×10^{12} GC/kg。在一个实施例中,剂量为约 1.0×10^{13} GC/kg。本文所述的所有范围都包含端点。

[0174] 在一个实施例中,有效剂量(所递送的总基因组拷贝)为约 10^7 至 10^{13} 个基因组拷贝。在一个实施例中,总剂量为约 10^8 个基因组拷贝。在一个实施例中,总剂量为约 10^9 个基因组拷贝。在一个实施例中,总剂量为约 10^{10} 个基因组拷贝。在一个实施例中,总剂量为约 10^{11} 个基因组拷贝。在一个实施例中,总剂量为约 10^{12} 个基因组拷贝。在一个实施例中,总剂量为约 10^{13} 个基因组拷贝。在一个实施例中,总剂量为约 10^{14} 个基因组拷贝。在一个实施例中,总剂量为约 10^{15} 个基因组拷贝。

[0175] 令人期望的是利用最低有效浓度的病毒以降低如毒性等不期望效果的风险。在这些范围内的仍其它剂量和施用体积可以由主治医师考虑所治疗的受试者(优选地人)的身体状态、受试者的年龄、特定病症和病症(如果进行性的话)已经发展的程度来选择。

[0176] 在某些实施例中,组合物包括rAAV,所述rAAV包括诱导型GLP-2激动剂构建体。在某些实施例中,诱导剂或分子是雷帕霉素或雷帕霉素类似物。在某些实施例中,诱导剂是雷帕霉素,并且在包括rAAV的组合物之后施用至少一次或多次、至少两次或更多次、至少三次或更多次。在一些实施例中,雷帕霉素以至少约 4nM 至至少约 40nM 的剂量施用。在某些实施例中,诱导剂(即,雷帕霉素)以至少约 0.1mg/kg 至至少约 3.0mg/kg 的剂量施用。在某些实施例中,诱导剂(即,雷帕霉素)以至少约 0.5mg/kg 至至少约 2.0mg/kg 的剂量施用。

[0177] 本文所述的病毒载体和其它构建体可以用于制备药物,所述药物用于将GLP-2融

合蛋白构建体递送到有需要的受试者,将具有增加的半衰期的GLP-2供应到受试者,和/或用于治疗受试者的SBS。因此,另一方面,提供了一种治疗SBS的方法。所述方法包含向有需要的受试者施用如本文所述的组合物。在一个实施例中,组合物包含含有GLP-2融合蛋白表达盒的病毒载体,如本文所述。

[0178] 如本文所使用的,术语“治疗(treatment)”或“治疗(treating)”被定义为涵盖出于改善短肠综合征的一种或多种症状的目的向受试者施用本文所述的一种或多种化合物或组合物。因此,“治疗”可以包含降低给定受试者的SBS的进展、降低给定受试者的症状的严重程度、消除给定受试者的疾病症状、延迟给定受试者的疾病的进展或增加给定受试者的疗法功效中的一种或多种。

[0179] 在另一个实施例中,提供了一种用于治疗受试者的SBS的方法。所述方法包含施用病毒载体,所述病毒载体包括核酸分子,所述核酸分子包括编码如本文所述的融合蛋白的序列。在一个实施例中,所述受试者是人。

[0180] 治疗过程可以任选地涉及重复施用相同的病毒载体(例如,AAVrh91载体)或不同的病毒载体(例如,AAVrh91和AAV3B.AR2.12)。可以使用本文所述的病毒载体选择仍其它组合。任选地,本文所述的组合物可以在涉及营养疗法(肠内或肠胃外营养)、药物(如用于控制胃酸、减少腹泻或改善肠道吸收的药物)或GLP-2类似物或手术的方案中组合。在某些实施例中,AAV载体和组合疗法是基本上同时施用的。在其它实施例中,AAV载体是首先施用的。在其它实施例中,组合疗法是首先施用的。

[0181] 在一个实施例中,组合物与有效量的GLP-2类似物施用。本领域已知各种可商购获得的GLP-2产品,包含但不限于替度鲁肽、格列帕鲁肽和阿普拉鲁肽。

[0182] 在一些实施例中,与用病毒载体进行治疗之前相比,本文所述的rAAV与GLP-2类似物的组合降低了受试者的GLP-2类似物剂量要求。此类剂量要求可以降低10%或更多、20%或更多、30%或更多、40%或更多、50%或更多、60%或更多、70%或更多、80%或更多、或90%或更多。治疗医师可以确定受试者所需的GLP-2类似物的正确剂量。例如,可以使用GLP-2类似物或其它疗法来治疗受试者,所述治疗医师可以在施用AAV载体后继续进行治疗。此类GLP-2类似物或其它协同疗法可以随后根据需要进行、减少或中断。

[0183] 在一个实施例中,将包括表达盒、载体基因组、rAAV的组合物或本文所述的用于基因疗法的其它组合物以单剂量/患者递送。在一个实施例中,向受试者递送治疗有效量的本文所述的组合物。如本文所使用的,“治疗有效量”是指表达盒或载体或其组合在靶细胞中递送和表达足以达到治疗目标的GLP-2-Fc的量的量。治疗有效量可以由治疗医师选择,或基于先前确定的指南指导。例如,替度鲁肽可以以0.05mg/kg的初始剂量每日皮下注射。对于患有中度至重度肾损伤的受试者,剂量可以以0.025mg/kg的增量增加。可以将rAAV递送到受试者,并且然后根据需要服用口服或皮下替度鲁肽、或其它药物补充,以达到每天0.05mg/kg的期望剂量的当量。

[0184] 在某些实施例中,治疗目标是改善或治疗SBS的症状中的一种或多种。治疗有效量可以基于动物模型而不是人类患者来确定。

[0185] 如本文所使用的,当用于指vp衣壳蛋白时,术语“异质”或其任何语法变型是指由不相同的元件组成的群体,例如具有带有不同的经过修饰的氨基酸序列的vp1、vp2或vp3单体(蛋白质)。SEQ ID NO:32提供了AAVrh91 vp1蛋白的经编码的氨基酸序列。与vp1、vp2和

vp3蛋白(可替代地被称为同种型)结合使用的术语“异质的”是指衣壳内的vp1、vp2和vp3蛋白的氨基酸序列中的差异。AAV衣壳含有具有来自预测的氨基酸残基的修饰的vp1蛋白内、vp2蛋白内和vp3蛋白内的亚群体。这些亚群体至少包含某些脱酰胺化的天冬酰胺(N或Asn)残基。例如,某些亚群体包括天冬酰胺-甘氨酸对中的至少一个、两个、三个或四个高度脱酰胺化的天冬酰胺(N)位置,并且任选地进一步包括其它脱酰胺化的氨基酸,其中脱酰胺化引起氨基酸变化和其它任选的修饰。

[0186] 如本文所使用的,除非另有说明,否则vp蛋白的“亚群体”是指一组vp蛋白,所述一组vp蛋白具有至少一个限定的共同特性,并且由至少一个组成员到少于参考组的所有成员组成。例如,除非另有说明,否则vp1蛋白的“亚群体”是至少一种(1)vp1蛋白,并且少于组装的AAV衣壳中的所有vp1蛋白。除非另有说明,否则vp3蛋白的“亚群体”可以是少于组装的AAV衣壳中的所有vp3蛋白的一种(1)vp3蛋白。例如, vp1蛋白可以是vp蛋白的亚群体;vp2蛋白可以是vp蛋白的单独亚群体,并且vp3是组装的AAV衣壳中的vp蛋白的另外的亚群体。在另一个实例中, vp1、vp2和vp3蛋白可以含有具有不同修饰的亚群体,例如,至少一种、两种、三种或四种高度脱酰胺化的天冬酰胺,例如在天冬酰胺-甘氨酸对处。

[0187] 如本文所使用的,rAAV的“原液”是指rAAV的群体。尽管由于脱酰胺作用,其衣壳蛋白具有异质性,但是rAAV在原液中被预期与5共享相同的载体基因组。原液可以包含具有衣壳的rAAV,所述衣壳具有例如所选择的AAV衣壳蛋白和所选择的产生系统的特有的异质脱酰胺模式。原液可以从单个产生系统中产生,或者从产生系统的多次运行中池化。可以选择各种产生系统,包含但不限于本文所述的产生系统。如本文所使用的,术语“GLP-2构建体”、“GLP-2表达构建体”和同义词包含如本文所述的GLP-2序列与前导序列和融合结构域的组合。术语“GLP-2构建体”、“GLP-2表达构建体”和同义词可以用于指编码GLP-2融合蛋白或其表达产物的核酸序列。

[0188] 在核酸序列的上下文中,术语“同一性百分比(%)”、“序列同一性”、“序列同一性百分比”或“相同百分比”是指两个序列中的碱基在比对以获得对应性时是相同的。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100至150个核苷酸的片段,或根据需要。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至24个核苷酸、至少约28至32个核苷酸、至少约36个或更多个核苷酸。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”、“CAP序列组装”、“BLAST”、“MAP”和“MEME”,这些程序可通过因特网上的Web服务器进行访问。此类程序的其它来源是本领域技术人员已知的。可替代地,也使用了载体NTI实用程序。本领域已知的许多算法可以用于测量核苷酸序列同一性,包含上述程序中包含的那些。作为另一个实例,可以使用GCG 6.1版本的程序Fasta™比较多核苷酸序列。Fasta™提供了查询序列与搜索序列之间最佳重叠区的比对和序列同一性百分比。例如,核酸序列之间的序列同一性百分比可以是使用如GCG 6.1版本中所提供的采用其默认参数(字号6和评分矩阵的NOPAM系数)的Fasta™所确定的,所述程序通过引用并入本文。

[0189] 术语“高度保守的”意指至少80%同一性、优选地至少90%同一性,并且更优选地超过97%同一性。通过使用本领域技术人员已知的算法和计算机程序,本领域技术人员可以容易地确定同一性。

[0190] 除非由较高范围另行规定,否则应当理解,同一性百分比是同一性的最低水平,并

且涵盖同一性的所有较高水平,直至与参考序列的100%同一性。除非另有说明,否则应当理解,同一性的百分比是同一性的最低水平,并且涵盖同一性的所有较高水平,直至与参考序列的100%同一性。例如,“95%同一性”和“至少95%同一性”可以互换地使用,并且包含与参考序列的95%、96%、97%、98%、99%以及直至100%同一性,以及其间的所有分数。

[0191] 在氨基酸序列的上下文中,术语“同一性百分比(%)”、“序列同一性”、“序列同一性百分比”或“相同百分比”是指两个序列中的残基在比对以获得对应性时是相同的。可以容易地确定蛋白质全长、多肽、约70个氨基酸至约100个氨基酸或其肽片段或对应的核酸序列编码序列上的氨基酸序列的同一性百分比。合适的氨基酸片段的长度可以为至少约8个氨基酸,并且可以为至多约150个氨基酸。通常,当提及两种不同序列之间的“同一性”、“同源性”或“类似性”时,参考“比对”序列来确定“同一性”、“同源性”或“类似性”。“比对”序列或“比对”是指与参考序列相比,通常含有对丢失的或另外的碱基或氨基酸的校正的多个核酸序列或蛋白质(氨基酸)序列。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”、“MAP”、“PIMA”、“MSA”、“BLOCKMAKER”、“MEME”和“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,所述算法或计算机程序提供至少与参考算法和程序所提供的同一性或比对水平相同的同一性或比对。参见例如,J.D.Thomson等人,《核酸研究(Nucl.Acids.Res.)》,“多个序列比对的综合比较(A comprehensive comparison of multiple sequence alignments)”,27(13):2682-2690(1999)。

[0192] 应注意的是,术语“一个(a)”或“一种(an)”是指一个或多个/一种或多种。如此,术语“一个”(或“一种”)、“一个或多个”和“至少一个”在本文可互换地使用。

[0193] 词语“包括(comprise)”、“包括(comprises)”和“包括(comprising)”将被解释为是包含性而非排他性的。词语“由……组成(consist/consisting)”和其变型将被解释为是排他性的而非包含性的。虽然说明书中的各个实施例是使用“包括”语言来呈现的,但是在其它情况下,也旨在使用“由……组成”或“基本上由……组成”的语言来解释和描述相关实施例。

[0194] 如本文所使用的,“患者”或“受试者”是指哺乳动物,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。在另一个实施例中,受试者不是猫科动物。

[0195] 如本文所使用的,除非另有说明,否则术语“约”意指相对于给定参考的10%($\pm 10\%$,例如, ± 1 、 ± 2 、 ± 3 、 ± 4 、 ± 5 、 ± 6 、 ± 7 、 ± 8 、 ± 9 、 ± 10 或其之间的值)的变化性。

[0196] 在某些情况下,术语“E+#”或术语“e+#”用于指代指数。例如,“5E10”或“5e10”是 5×10^{10} 。这些术语可以互换地使用。

[0197] 如本文所使用的,术语“调控”或其变型是指组合物抑制生物通路的一个或多个组分的能力。

[0198] 如本文所使用的,“疾病”、“病症”和“病状”可互换地用于指示受试者的异常状态。

[0199] 除非在本说明书中另有定义,否则本文所使用的技术术语和科学术语具有与本领域的普通技术人员和参照公开文本所通常理解相同含义,这为本领域的技术人员提供了本申请中使用的许多术语的通用指南。

[0200] 在描述实施例时对“一个实施例”或“另一个实施例”的引用并不意味着所引用的实施例与另一个实施例(例如,在所引用实施例之前描述的实施例)相互排斥,除非另有明确规定。

[0201] 实例

[0202] 提供以下实例来说明本发明的各个实施例。实例并非旨在以任何方式限制本发明。

[0203] 实例1-GLP-2载体的构建

[0204] GLP-2激动剂通过腺相关病毒(AAV)表达是具有挑战性的。GLP-2通常由胰高血糖素前体蛋白表达,其需要组织特异性蛋白酶并且产生不想要的蛋白质。使用传统的异源性信号肽的表达系统产生低表达。使用具有通用蛋白酶切割位点的异源性前肽的表达系统产生可以作为T细胞的靶标的外源蛋白序列。

[0205] 更具体地,构建了载体,其中前导序列置于若干GLP-2受体激动剂氨基酸序列之一的上游,随后是融合结构域。参见例如,图3。将所得蛋白质序列回译,随后添加kozak共有序列、终止密码子和克隆位点。产生序列,并且将序列在诱导型表达系统的控制下克隆到含有CMV启动子的表达载体中。表达构建体侧接有AAV2 ITR。所得质粒被称为pAAV.Z12I.hGLP2.G2.Fc.rBG。

[0206] 在第二构建体中,产生了序列,并且将其克隆到含有CB7组成型启动子的表达载体中。

[0207] 实例2-体外表达

[0208] 如先前所述,通过三重转染将以下构建体包装到AAVrh91载体中。

[0209] AAVrh91.CB7.CI.hGLP-2-Fc.rBG

[0210] AAVrh91.CB7.CI.hGLP-2-SA.rBG

[0211] 在用具有人凝血酶信号序列的诱导型人GLP-2-SA的载体转染的HEK293细胞的培养上清液中测量GLP-2-SA融合物。GLP2-SA通过使用spyroRuby染色的凝胶电泳进行鉴定。图4A和图4B。

[0212] 实例3-Rag1KO小鼠体内的中试表达

[0213] Rag1KO雌性小鼠通过IM施用途注射载体AAVrh91.CB7.CI.hGLP-2-SA.rBG(1×10^{11} GC/小鼠)或AAVrh91.CB7.CI.hGLP-2-Fc.rBG(1×10^{11} GC/小鼠)来进行处理。通过在含有5微升DPP-IV抑制剂(密理博公司(Millipore))的血清分离管中分离全血来连续收集血清,并且如上所述测定活性GLP-2表达和活性。在第0天注射载体,并且在第56天对小鼠进行尸检。血清GLP-2浓度(nM)如图5A所示并且是基于Fc融合物标准的估计值。血清表达水平在给药后第14天达到增加。

[0214] 在尸检时对小肠进行称重和测量。与对照动物相比,小肠的长度和重量显著增加(图5B)。另外,与经媒剂处理的动物相比,经载体处理的肠显示出健康的肠上皮细胞生长。(图5C)。测量了血清瓜氨酸水平(μ M),一种肠道表面积的生物标志物。(图5D)。

[0215] 将GLP-2的表达与白蛋白融合构建体进行比较。Fc融合构建体显示出更大的GLP-2表达(图6A)。与hGLP($EC_{50} = 0.4$ nM)相比,GLP.G2.Fc融合显示出显著较低的效力($EC_{50} = 13.6$ nM)(图6B)。这些数据表明,AAV介导的GLP-2激动剂的表达显示出肠道表面积的显著和持久增加,这应对短肠综合征具有治疗作用。

[0216] 实例4-NHP中的剂量研究

[0217] 在这项研究中,检查了人GLP-2在非人类灵长类动物(NHP;即,恒河猴)体内的表达。图7示出了研究的概况。简而言之,通过肌肉注射(IM)以 1×10^{13} (1e13)GC/kg (E185NG)的剂量和 5×10^{10} (1e10)GC/kg (BM239H)的剂量向2个NHP施用AAVrh91.CB7.CI.hGLP-2-Fc.rBG。

[0218] 对GLP-2表达和效力、肝酶和瓜氨酸水平进行了测量。在第60天进行尸检。图8A示出了GLP-2-Fc融合蛋白的血浆水平。图8B示出了血清瓜氨酸,一种肠道表面积的生物标志物。图8C示出了以1:100稀释的NHP血清中的抗GLP-2-Fc抗体的检测。AAV介导的GLP-2-Fc融合物的表达表明,在高剂量下,肠道表面积显著增加,直到抗GLP-2-Fc抗体降低其表达为止,证明了其对短肠综合征的治疗功效。

[0219] 实例5-Rag1KO小鼠的长期功效和安全性研究

[0220] Rag1KO雌性小鼠通过IM施用途径以 1×10^{10} GC/小鼠、 3×10^{10} GC/小鼠或 1×10^{11} GC/小鼠的剂量注射载体AAVrh91.CB7.CI.hGLP-2-Fc.rBG来进行处理。研究设计示出于图9A中。在整个研究过程中测量了hGLP2-Fc水平和体重。最高载体剂量显示出最高血清GLP2水平(图9B),而所有组间的体重相对一致,其中媒介剂趋势最低(图9C、D)。在第28天尸检时,测量小肠长度和重量。虽然与媒介剂相比,SI长度显示出随剂量增加而适度增加(图9E),但两个最高剂量组的SI重量显著增加(图9F)。

[0221] 本说明书中引用的所有文件通过引用并入本文。于2022年3月3日提交的美国临时专利申请第63/316,219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。

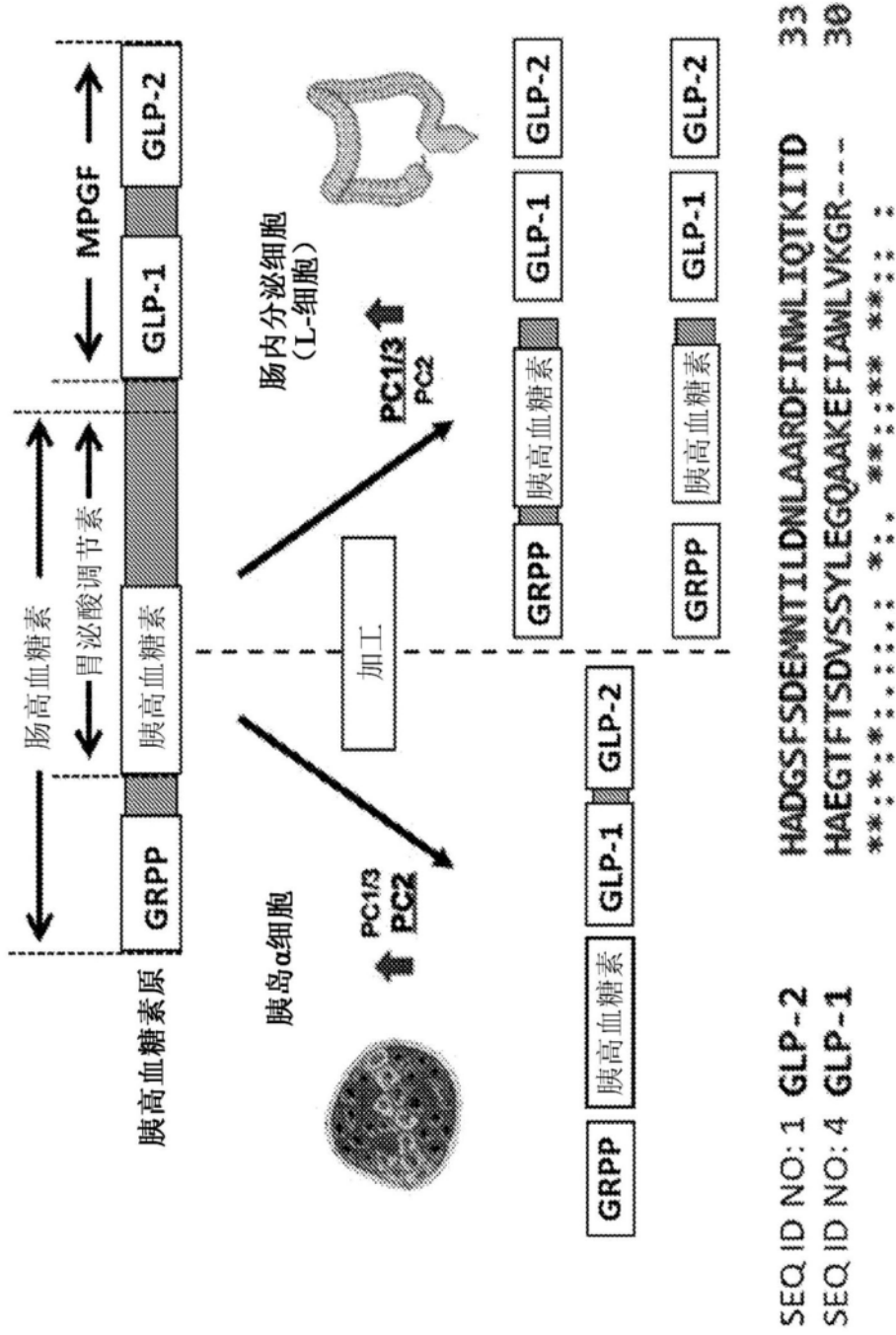


图1

表1
hGLP-2和GLP-2类似物的氨基酸序列和分子量
相对于天然hGLP-2的氨基酸取代以粗体加下划线字母示出

肽	分子量	氨基酸序列																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
SEQ ID NO: 1 hGLP-2	3766	H	A	D	G	S	F	S	D	E	M	N	T	I	L	D	N	L
SEQ ID NO: 2 格列帕鲁肽	4316	H	G	E	G	I	F	S	S	E	L	A	T	I	L	D	A	L
SEQ ID NO: 3 替度鲁肽	3752	H	G	D	G	S	F	S	D	E	M	N	T	I	L	D	N	L
SEQ ID NO: 1 hGLP-2	3766	A	A	R	D	F	I	N	W	L	I	Q	T	K	I	T	D	OH
SEQ ID NO: 2 格列帕鲁肽	4316	A	A	R	D	F	I	A	W	L	I	A	T	K	I	T	D	KKKKKKKNH₂
SEQ ID NO: 3 替度鲁肽	3752	A	A	R	D	F	I	N	W	L	I	Q	T	K	I	T	D	OH

f, D-苯丙氨酸。

图2

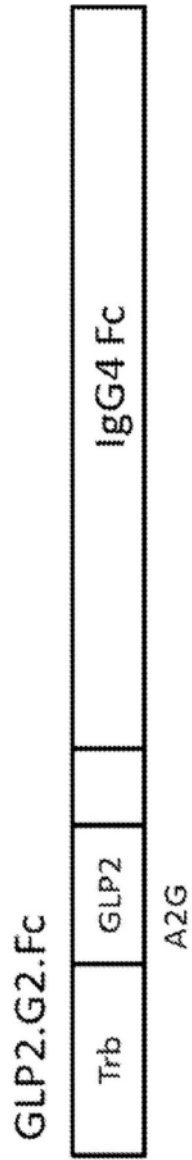


图3A

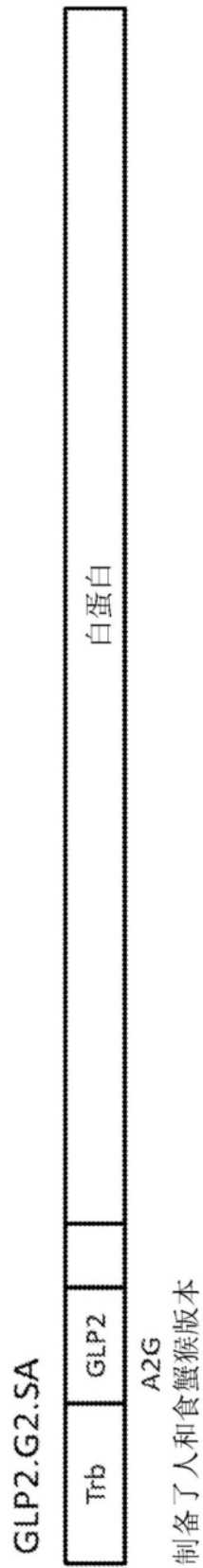


图3B

GLP2在物种之间不相同

人	HADGSFDEMNTILDNLAARDFINWLIQTKITD	33	SEQ ID NO: 1
食蟹猴	HADGSFDEMNTVLDNLAARDFINWLIQTKITD	33	SEQ ID NO: 5
小鼠	HADGSFDEMSTILDNLAARDFINWLIQTKITD	33	SEQ ID NO: 6

 *****;
 *****;

图3C

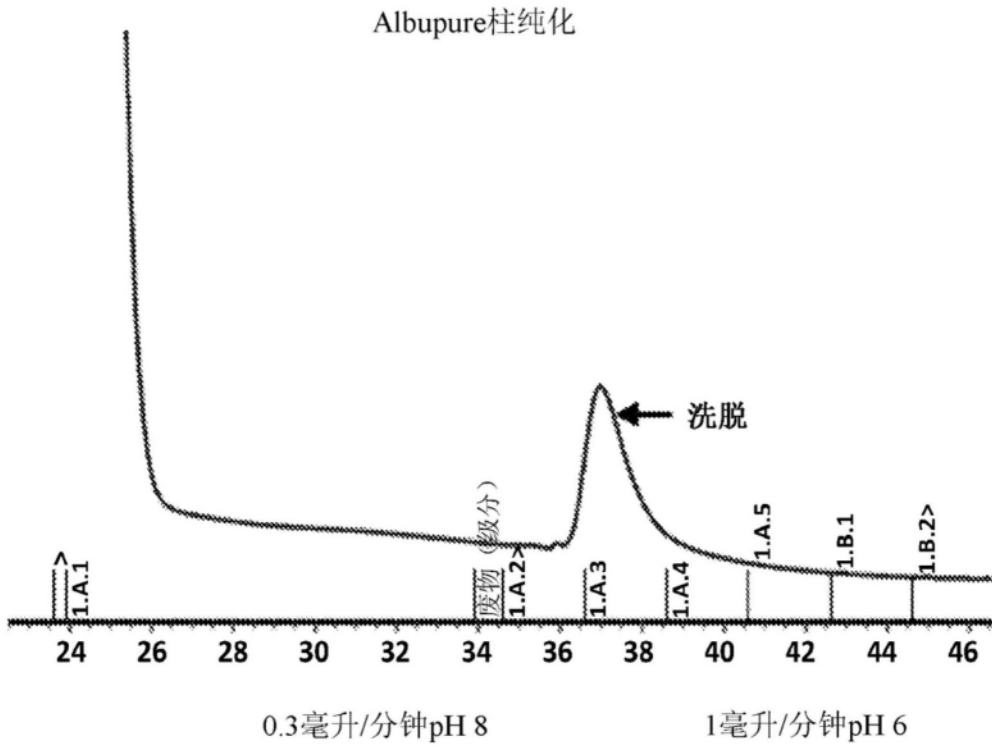


图4A

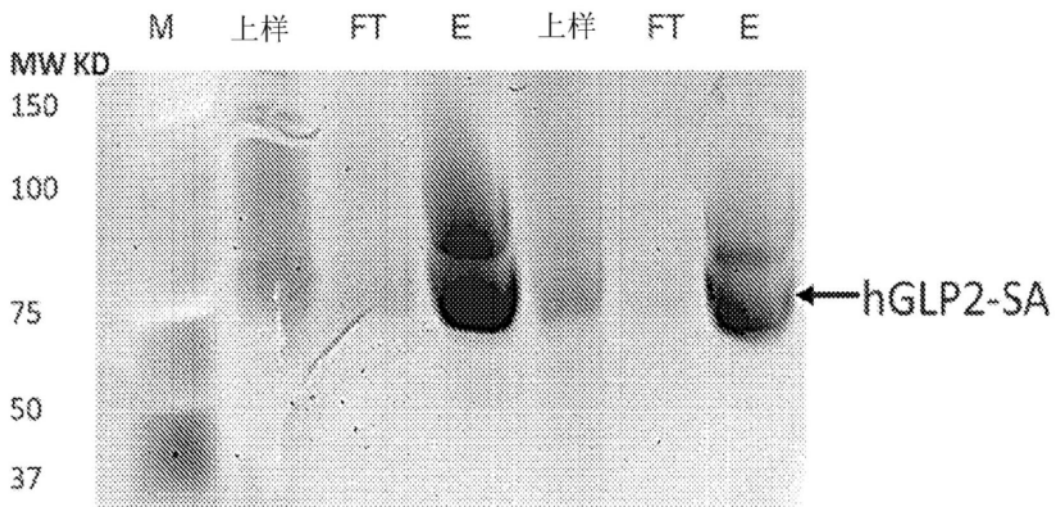


图4B

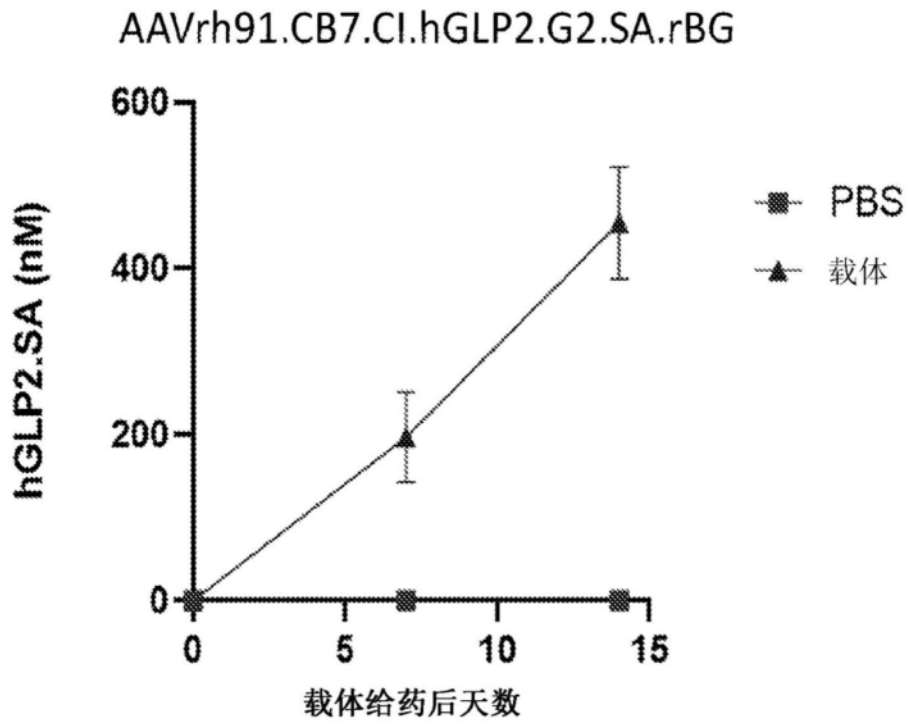


图5A

第56天尸检后

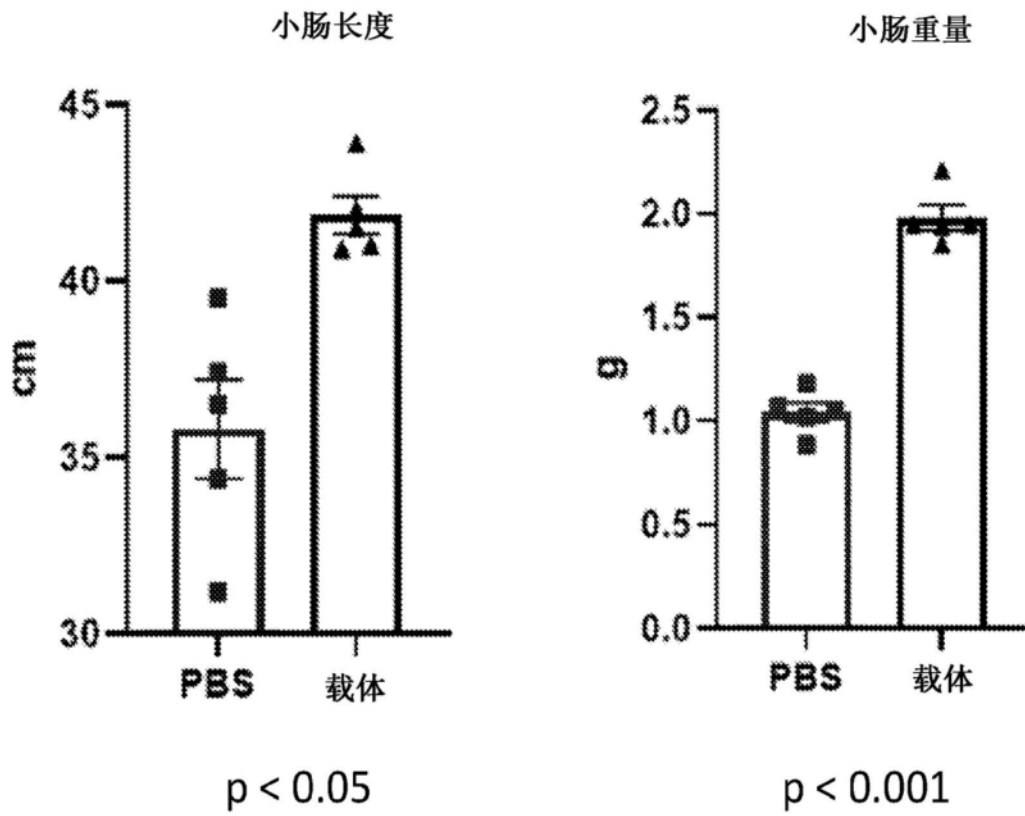


图5B

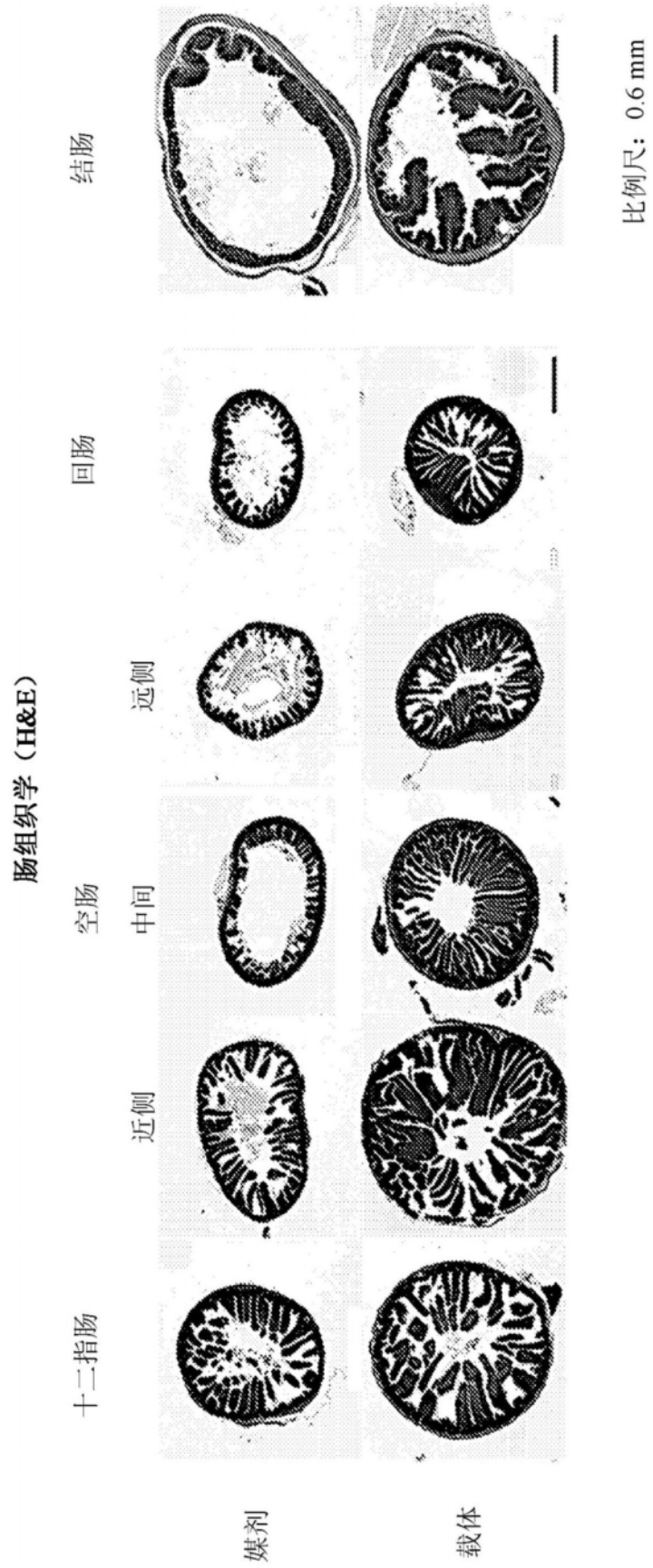


图5C

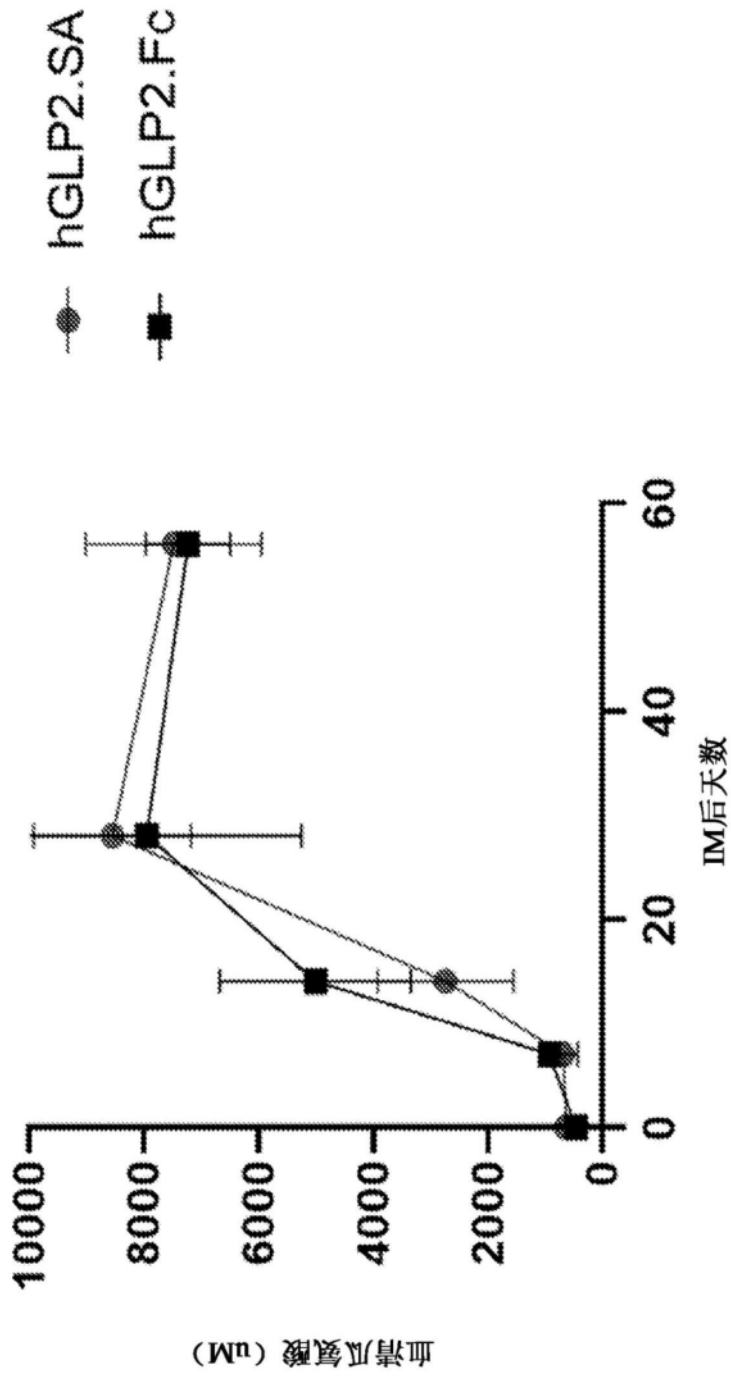


图5D

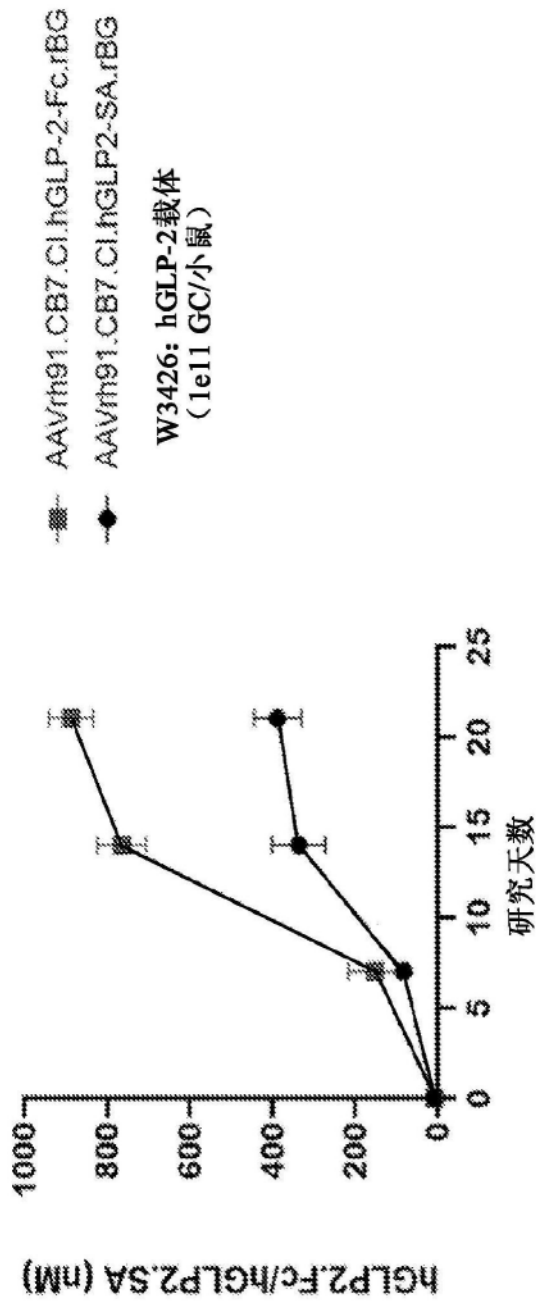


图6A

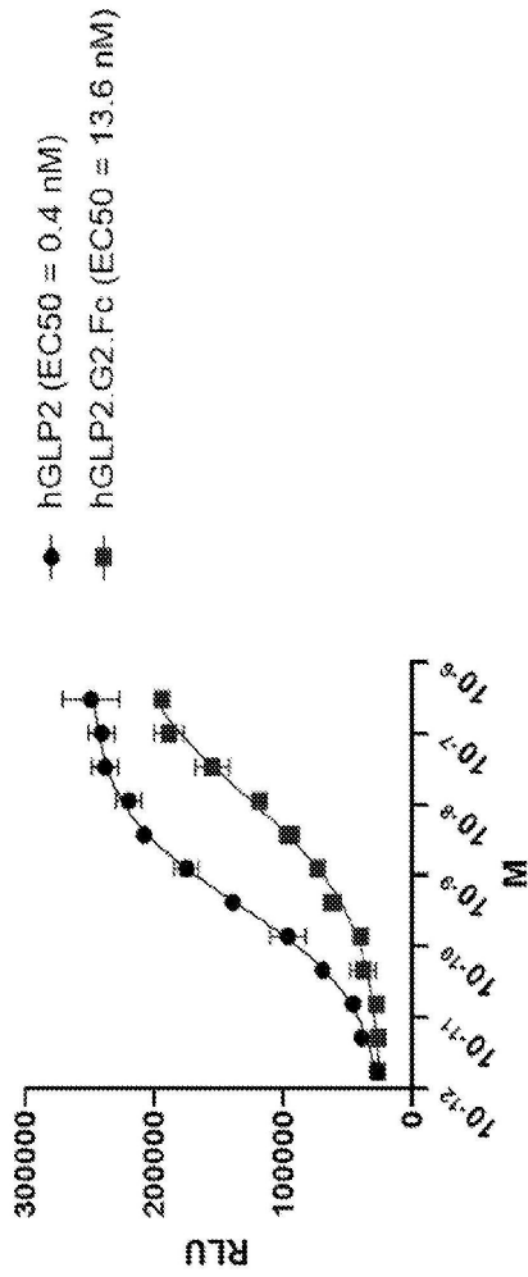


图6B

NHP研究设计

NHP	1	2
载体	AAVrh91.CB7.Cl.hGLP-2-Fc.rBG	
ROA	IM	
剂量 (GC/kg)	1.00E+12	1.00E+10
采血	第0天、第7天、第14天、第21天、第28天、第60天	
读出	TG表达和效力、ADA、生物标志物（瓜氨酸*）	
尸检	第60天（肠组织学**）	

*：血浆/血清瓜氨酸与肠上皮细胞质量相关

**：针对绒毛长度的H&E染色（GLP-1-Fc研究中的NHP将用作对照）

图7

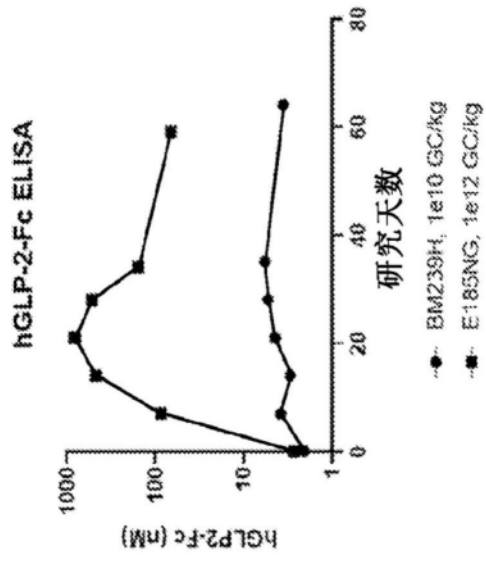


图8A

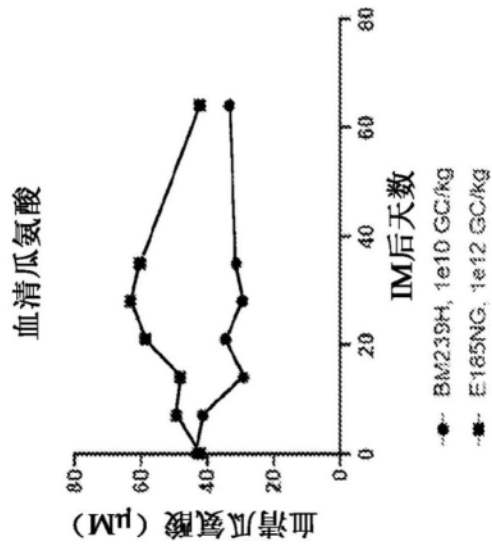


图8B

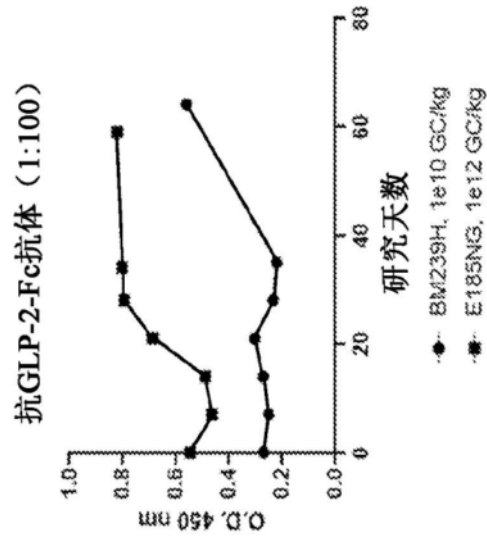


图8C

组	剂量 (GC/动物)	尸检日			
		1个月	3个月	6个月	12个月
1	0	5	5	10	10
2	1e10	5	5	10	10
3	3e10	5	5	10	10
4	1e11	5	5	10	10

图9A

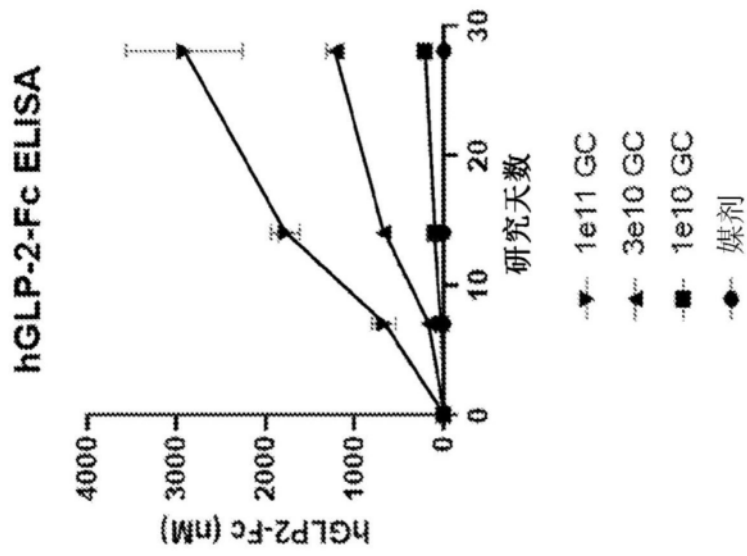


图9B

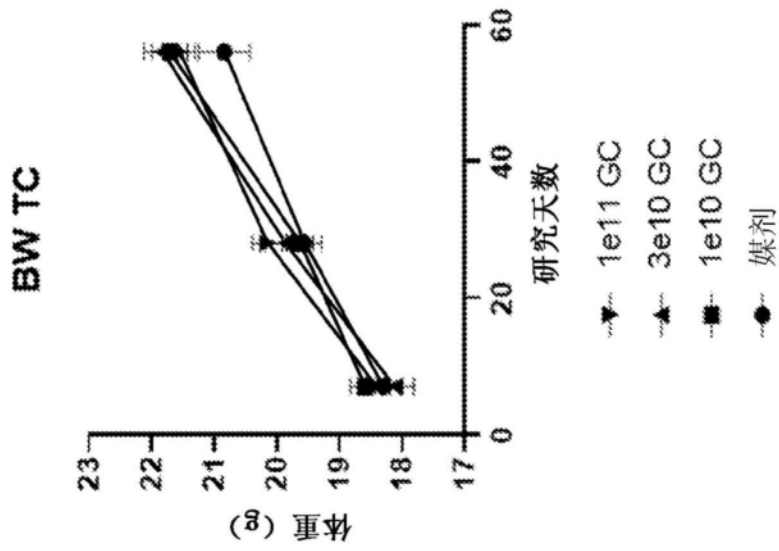


图9C

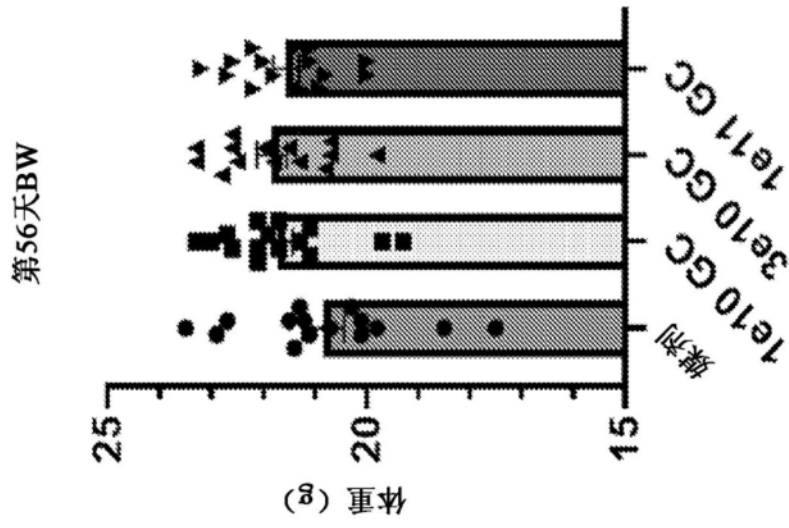


图9D

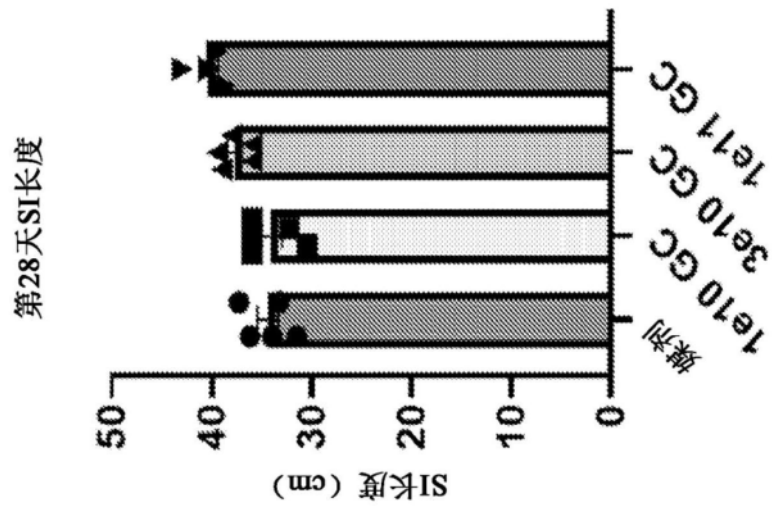


图9E

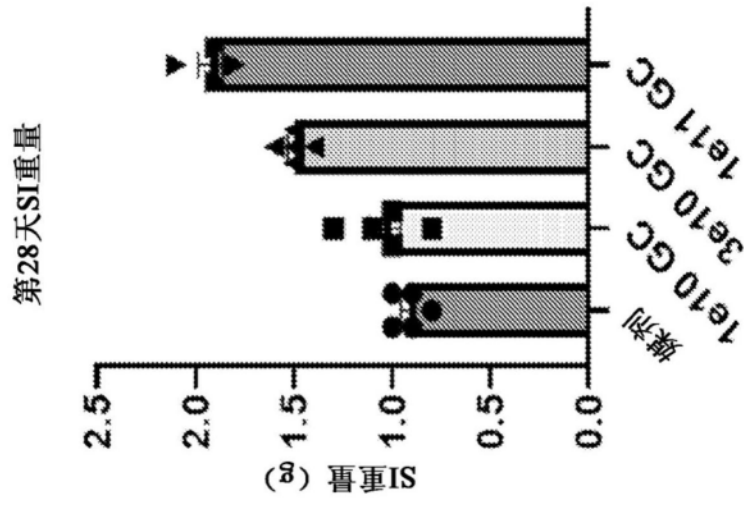


图9F