



(12) 发明专利申请

(10) 申请公布号 CN 113164560 A

(43) 申请公布日 2021.07.23

(21) 申请号 201980070614.8

(22) 申请日 2019.10.28

(30) 优先权数据

18202888.6 2018.10.26 EP

(85) PCT国际申请进入国家阶段日

2021.04.25

(86) PCT国际申请的申请数据

PCT/EP2019/079351 2019.10.28

(87) PCT国际申请的公布数据

W02020/084161 EN 2020.04.30

(71) 申请人 VRIJE布鲁塞尔大学

地址 比利时布鲁塞尔

(72) 发明人 T.范登德里斯奇 L.K.蔡

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 涂滔 邹宗亮

(51) Int.Cl.

A61K 38/38 (2006.01)

C07K 14/745 (2006.01)

权利要求书2页 说明书45页

序列表44页 附图16页

(54) 发明名称

改善基因疗法的新工具及其用途

(57) 摘要

本发明涉及一种用于增加由转基因编码的蛋白质或多肽的水平和/或活性的编码人白蛋白的核酸分子,及其在含有肝脏特异性调节元件和密码子优化的因子IX、因子VIII、因子VII或因子VIIa转基因的核酸表达盒和载体中的用途,以及采用这些表达盒和载体的方法及其用途,所述核酸分子包含由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列。本发明特别适用于使用肝脏定向基因疗法的应用,尤其适用于治疗血友病A、血友病B或因子VII缺陷。

1. 一种编码人白蛋白的密码子优化的核酸分子,其包含由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列,优选由SEQ ID NO:14定义的序列。

2. 根据权利要求1所述的核酸分子,其包含与SEQ ID NO:14定义的所述序列或与SEQ ID NO:14定义的所述序列具有至少80%序列同一性的所述序列融合的转基因,优选地,其中所述转基因是密码子优化的转基因,任选地,其中所述转基因位于SEQ ID NO:14所定义的所述序列或与所述序列具有至少80%序列同一性的所述序列的5'末端。

3. 根据权利要求2所述的核酸分子,其中所述转基因通过编码一个或多个多肽或肽接头的序列与由SEQ ID NO:14定义的所述序列或与所述序列具有至少80%序列同一性的所述序列分隔开,所述编码一个或多个多肽或肽接头的序列优选编码由SEQ ID NO:18定义的肽接头。

4. 根据权利要求1至3中任一项所述的核酸分子,其用于增加由转基因编码的蛋白质或多肽的表达和/或循环水平和/或活性。

5. 一种核酸表达盒,其包含根据权利要求1至3中任一项所述的核酸分子,所述核酸分子可操作地连接至启动子。

6. 根据权利要求5所述的核酸表达盒,其包含

-至少一种组织特异性核酸调节元件,其可操作地连接至所述启动子和如权利要求1至3中任一项所定义的核酸分子;

-小鼠微小病毒(MVM)内含子,优选由SEQ ID NO:20定义的MVM内含子;和/或

-转录终止信号,优选多腺苷酸化信号,更优选由SEQ ID NO:21定义的合成的多腺苷酸化信号或SEQ ID NO:23定义的猿猴病毒40(SV40)多腺苷酸化信号。

7. 根据权利要求5或6中任一项所述的核酸表达盒,其中所述转基因编码分泌型治疗蛋白或分泌型免疫蛋白,优选地,所述转基因编码选自下列的分泌型治疗蛋白:因子IX、因子VII、因子VIIa、因子VIII、肝细胞生长因子(HGF)、组织因子(TF)、组织因子途径抑制剂(TFPI)、ADAMTS13、血管内皮生长因子(VEGF)、胎盘生长因子(PLGF)、成纤维细胞生长因子(FGF)、可溶性fms样酪氨酸激酶1(sFLT1)、 α 1-抗胰蛋白酶(AAT)、胰岛素、胰岛素原、因子X、血管性血友病因子、C1酯酶抑制剂(C1-INH)、溶酶体酶、溶酶体酶艾杜糖醛酸-2-硫酸酯酶(I2S)、促红细胞生成素(EPO)、干扰素- α 、干扰素- β 、干扰素- γ 、白介素1(IL-1)、白介素2(IL-2)、白介素3(IL-3)、白介素4(IL-4)、白介素5(IL-5)、白介素6(IL-6)、白介素7(IL-7)、白介素8(IL-8)、白介素9(IL-9)、白介素10(IL-10)、白介素11(IL-11)、白介素12(IL-12)、趋化因子(C-X-C基序)配体5(CXCL5)、粒细胞集落刺激因子(G-CSF)、粒细胞-巨噬细胞集落刺激因子(GM-CSF)、巨噬细胞集落刺激因子(M-CSF)、干细胞因子(SCF)、角质细胞生长因子(KGF)、单核细胞趋化蛋白-1(MCP-1)、肿瘤坏死因子(TNF)、afamin(AFM)、 α 1-抗胰蛋白酶、 α -半乳糖苷酶A、 α -L-艾杜糖醛酸酶、脂蛋白脂肪酶、载脂蛋白(apolipoprotein)、低密度脂蛋白受体(LDL-R)、白蛋白、耐二肽基肽酶4(DPP-4)的胰高血糖素样肽1(GLP-1)、GLP-2、胰高血糖素、生长激素(GH)、干扰素(例如,IFN α -2b)、 β -利尿钠肽、IL-1Ra、exendin-4、胃酸调节素、卵泡抑素、抗体和纳米抗体;优选地,其中所述转基因编码用于治疗 and/或预防肝脏相关疾病的治疗性蛋白,所述肝脏相关疾病优选血友病A、血友病B或因子VII缺陷。

8. 根据权利要求7所述的核酸表达盒,其中所述转基因编码

-凝血因子IX(FIX),优选地,其中所述凝血因子FIX包含超激活突变,更优选地,其中所

述超激活突变对应于R338L氨基酸取代,更优选地,其中所述转基因编码具有SEQ ID NO:11所定义的核酸序列的凝血因子IX;

-凝血因子VIII (FVIII),优选地,其中所述转基因是密码子优化的凝血因子FVIII,或其中所述凝血因子VIII具有B结构域的缺失,优选地,其中所述FVIII的所述B结构域被SEQ ID NO:15定义的接头替换,更优选地,其中所述转基因编码具有SEQ ID NO:16所定义的核酸序列的凝血因子VIII;或

-凝血因子VII (FVII) 或因子FVIIa (FVIIa) 的轻链和重链,任选地,其中所述FVII或FVIIa的轻链通过一个或多个可切割的多肽或肽接头与所述FVII或FVIIa的重链偶联,优选地,其中所述转基因编码SEQ ID NO:34所定义的氨基酸序列。

9. 根据权利要求7或8中任一项所述的核酸表达盒,其中所述至少一种组织特异性核酸调节元件是至少一种肝脏特异性核酸调节元件。

10. 根据权利要求9所述的核酸表达盒,其中所述至少一种肝脏特异性核酸调节元件由SEQ ID NO:25定义的Serpín增强子或与所述序列具有至少95%同一性的序列组成,优选由三重重复、优选串联排列的SEQ ID NO:25定义的Serpín增强子或与所述序列具有至少95%同一性的序列组成。

11. 根据权利要求7至10中任一项所述的核酸表达盒,其中所述启动子是肝脏特异性启动子,优选地,肝脏特异性启动子选自包含以下的群组:甲状腺素运载蛋白(TTR)启动子、最小TTR启动子(TTRm)、AAT启动子、白蛋白(ALB)启动子或最小启动子、载脂蛋白A1(APOA1)启动子或最小启动子、补体因子B(CFB)启动子、己酮糖激酶(KHK)启动子、血色素结合蛋白(HPX)启动子或最小启动子、烟酰胺N甲基转移酶(NNMT)启动子或最小启动子、(肝)羧酸酯酶1(CES1)启动子或最小启动子、蛋白C(PROC)启动子或最小启动子、载脂蛋白C3(APOC3)启动子或最小启动子、甘露聚糖结合凝集素丝氨酸蛋白酶2(MASP2)启动子或最小启动子、铁调素抗菌肽(HAMP)启动子或最小启动子,或丝氨酸蛋白酶抑制剂C类(抗凝血酶)成员1(SERPINC1)启动子或最小启动子,优选TTRm。

12. 一种载体,其包含根据权利要求5至11中任一项所述的核酸表达盒,优选地,其中所述载体是病毒载体,更优选地,其中所述载体源自腺相关病毒(AAV),优选地,其中所述载体是单链AAV。

13. 根据权利要求12所述的载体,其具有SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8,优选SEQ ID NO:8。

14. 一种药物组合物,其包含根据权利要求12或13所述的载体和药学上可接受的载体。

15. 根据权利要求12或13所述的载体或根据权利要求14所述的药物组合物,其用于医学,优选基因疗法,更优选肝脏定向基因疗法。

16. 根据权利要求12或13所述的载体或根据权利要求14所述的药物组合物,其用于治疗肝脏相关病症,若所述转基因是FIX、FVIII、FVII或FVIIa,则所述肝脏相关病症优选血友病。

17. 一种用于在肝细胞中表达转基因产物的体外或离体方法,其包括:

-将根据权利要求5至11中任一项所述的核酸表达盒或根据权利要求12或13所述的载体引入到所述肝细胞中;

-在所述肝细胞中表达所述转基因产物。

改善基因疗法的新工具及其用途

技术领域

[0001] 本发明涉及能够增强(转)基因的组织特异性表达和/或活性的新工具、采用这些工具的方法及其用途。本发明还包括表达系统和包含这些表达系统的药物组合物。本发明特别适用于使用基因疗法、更具体而言组织定向基因疗法的应用,以及适用于疫苗接种目的。

背景技术

[0002] 血友病B是由凝血因子IX (FIX) 缺乏引起的X连锁隐性出血性疾病。血友病A是一种严重的出血性疾病,是缺乏或完全不存在由肝脏产生的凝血因子VIII (FVIII) 所引起。血友病A和B的临床表现以自发性和长期出血发作为特征。据估计,5,000人中有1人患有血友病A,20,000人中有1人患有血友病B。目前,血友病A和B通过使用血浆衍生的或重组的FVIII或FIX的蛋白质替代疗法进行治疗。尽管蛋白质替代大幅提高了血友病患者的预期寿命,他们仍然面临严重出血发作和慢性关节损伤的风险,因为预防性治疗受到半衰期短、有效性有限和高成本的纯化凝血因子限制,纯化凝血因子的成本接近100,000美元/患者/年。此外,使用从污染的血源获得的血浆衍生因子会增加病毒传播的风险。基因疗法提供了治疗血友病B的新方法的希望,因为治疗窗口相对较宽,略高于正常生理水平1%的水平就能产生疗效。如果成功,基因疗法可以在肝脏中提供恒定的FVIII或FIX合成,从而可导致治愈这种疾病。血友病基因疗法的不同模式已进行了广泛的论述,例如在Chuah等人,2012a,2012b,2012c;VandenDriessche等人,2012;High2001,2011;和Matrai等人,2010a,2010b中。

[0003] 国际血栓形成和止血学会科学和标准化委员会(the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis) 因子VIII和因子IX小组委员会(the subcommittee on Factor VIII and Factor IX) 分别根据FVIII水平和FIX水平将血友病A和血友病B的严重程度分为三级:1) 重度(FVIII或FIX水平低于0.01国际单位(IU)/ml,即低于正常FVIII或FIX水平的1%);2) 中度(FVIII或FIX水平为0.01到0.05IU/ml,即为正常FVIII或FIX水平的1至5%);以及3) 轻度(FVIII或FIX水平高于0.05至0.4IU/ml,即高于正常FVIII或FIX水平的5至40%)。血友病A是最常见的遗传性凝血疾病,其发病率接近1/5000名男性。

[0004] 使用纯化或重组的FVIII和FIX的蛋白替代疗法(PST)显著改善了患者的生活质量。然而,PST不能治愈,患者仍有可能发生危及生命的出血和严重的关节炎。不幸的是,许多患有血友病A(高达40%)的患者在进行PST后产生了FVIII的中和抗体(即“抑制剂”)。同样,估计10%血友病B的患者会产生FIX的“抑制剂”。这些抑制剂使出血发作的管理复杂化,并可能使PST进一步失效。这些血友病患者可以用因子VIIa治疗,因子VIIa即使面对FVIII或FIX的中和抗体也能进行止血矫正。PST的这些局限性证明了基因疗法发展为血友病治疗的潜在替代疗法的合理性。此外,仅需要适度增加FIX或FVIII血浆浓度就能产生治疗益处,超过正常水平的1%的水平就能够显著降低自发性出血和长期关节病的发生率。

[0005] 肝脏是FIX和FVIII合成的主要生理部位,因此,肝细胞是非常适合血友病基因治

疗的靶细胞。FIX或FVIII蛋白很容易从这个位置进入血液循环。此外,肝脏壁龛(niche)可有利于诱导对转基因产物的免疫耐受(Annoni等人,2007;Follenzi等人,2004;Brown等人,2007;Herzog等人,1999;Matrai等人,2011;Matsui等人,2009)。血友病肝脏定向基因疗法可以用不同的病毒载体完成,包括逆转录病毒载体(Axelrod等人,1990;Kay等人,1992;VandenDriessche等人,1999;Xu等人,2003,2005)、慢病毒载体(Ward等人,2011;Brown等人,2007;matrai等人,2011)、腺相关病毒载体(AAV)(Herzog等人,1999)和腺病毒载体(Brown等人,2004;Ehrhardt&Kay,2002)。AAV是一种具有单链DNA基因组的天然存在的复制缺陷型非致病性病毒。AAV载体具有良好的安全性,并且能够实现持续的转基因表达。长期表达主要由游离留存的AAV基因组介导。超过90%的稳定转导的载体基因组存在于染色体外,主要组织成高分子量串联体(concatamer)。因此,插入性肿瘤形成的风险很小,尤其是在血友病基因疗法的背景下,预计不会发生转导细胞的选择性扩增。AAV载体的主要局限在于载体颗粒有限的包装能力(即约5.0kb,包括AAV反向末端重复序列),从而限制了转基因表达盒的大小以获得功能性载体(Jiang等人,2006)。已从人和非人灵长类动物中分离出几种免疫学上不同的AAV血清型(Gao等人,2002;Gao等人2004),尽管大多数血友病基因治疗的载体最初来自最普遍的AAV2血清型。基于AAV的基因疗法在治疗先天性失明上取得的首次临床成功突出地显示了这种基因转移技术的潜力(Bainbridge等人,2008)。

[0006] 在患有血友病的鼠和犬模型或非人类灵长类动物中用AAV载体进行的临床前研究表明,持续的治疗性表达导致血友病模型中出血表型的部分或完全的矫正(Snyder等人,1997,1999;Wang等人,1999,2000;Mount等人,2002;Nathwani等人,2002)。特别地,肝脏转导方便地诱导对FIX的免疫耐受,其需要调节性T细胞(Treg)的诱导(Mingozzi等人,2003;Dobrzynski等人,2006)。在用肝脏导向的AAV2-FIX基因疗法治疗的倾向于产生抑制剂的无效突变血友病B犬中实现了没有抑制剂产生的血友病表型的长期矫正(Mount等人,2002)。为了进一步减少载体剂量,已经开发了更有效的FIX表达盒。这可以通过使用更强的启动子/增强子元件、密码子优化的FIX或自互补双链AAV载体(scAAV)来实现,这些载体克服了AAV转导中的限制步骤之一(即单链向双链AAV转化)(McCarty,2001,2003;Nathwani等人,2002,2006,2011;Wu等人,2008)。可以使用替代的AAV血清型(例如AAV8或AAV5),从而增加向肝细胞的转导,改善核内载体的导入,并可降低T细胞活化的风险(Gao等人,2002;Vandenbergh等人,2006),尽管不确定这是否一定也会适用于人类受试者,因为表位在不同的AAV血清型之间是保守的(Mingozzi等人,2007)。与使用慢病毒载体相比,至少在小鼠中使用AAV8或AAV9对血友病B的肝脏定向基因疗法更有效,并且导致更少的炎症(VandenDriessche等人,2007,2002)。此外,研究表明表面暴露的酪氨酸残基的突变允许载体颗粒避开磷酸化和随后的泛素化,并因此防止蛋白酶体介导的降解,从而导致FIX在小鼠中的肝脏表达增加10倍(Zhong等人,2008)。

[0007] 这些肝脏定向的临床前研究为将AAV载体用于临床基因治疗患有重度血友病B患者铺平了道路。在通过肝动脉插管接受AAV-FIX的受试者中,AAV-FIX载体的肝递送产生瞬时的治疗性FIX水平(最大为正常水平的12%)(Kay等人,2000)。最近,血友病的基因疗法已经向前迈出了重要的一步(Nathwani等人,2012;Nathwani等人,2014;VandenDriessche&Chuah的评论,2012)。患有重度血友病B(<1%FIX)的受试者静脉注射自互补(sc)AAV8载体,所述载体由肝脏特异性启动子表达密码子优化的FIX。FIX表达水平在大约3年的时间内在

正常水平的1%至6%之间变化,具有载体剂量依赖性效应。值得注意的是,高剂量队列中的所有6名患者的FIX表达水平都达到正常FIX的5%。FIX表达水平与FIX使用量和每年出血发作次数的减少相一致,在最高剂量队列中观察到最高的相对减少(即94%)。尽管如此,受试者仍有出血的风险,包括创伤引起的出血,因此需要保证间歇性FIX覆盖,在循环FIX水平较低的受试者中尤其如此。特别令人鼓舞的是,基因疗法总体具有良好的耐受性。与先前的肝脏定向AAV试验的主要区别在于,基因治疗后首次可以达到持续有疗效的FIX水平。然而,最常研究的相关不良事件是循环肝酶(即丙氨酸转氨酶,ALT)水平的无症状升高,在高剂量队列的6名患者中,有4名在载体输注后约7至10周出现该症状。这使人联想到在超过10年的血友病的初始基因治疗临床试验中使用基于AAV2的载体也观察到过的情况(Manno等人,2006)。为试图限制这种载体特异性免疫应答,使用短疗程糖皮质激素进行瞬时免疫抑制。在最近的一项临床试验中使用了相同的载体设计(ClinicalTrials.gov编号:#NCT02396342),但是在这种情况下使用另一种衣壳血清型(即AAV5)来包装治疗性FIX基因盒(Miesbach等人,2018)。这是基于以下理由,即AAV5表现出更有利的免疫谱,即与AAV2或AAV8相比更少出现中和抗体(Nab),而且至少在非人灵长类动物中能够转导肝细胞到与AAV8相似的程度。然而,使用标准化验证分析需要比较性研究,以更好地明白不同AAV血清型的总的血清阳性率,因为地区和特定人群会影响读出。先前的一项基于AAV5的临床试验表明,AAV5似乎没有引发针对衣壳的细胞免疫应答(D'Avola等人,2015)。尽管如此,基于所涉及的患者数量相对有限,得出使用AAV5的可能的免疫后果优于AAV8的任何明确的结论似乎还为时过早。

[0008] 这项开放标签临床研究包括10名患有重度血友病B的成年人,且其中没有人具有预先存在的抗AAV5的中和抗体。研究参与者被纳入低(5×10^{12} vg/kg)和高(2×10^{13} vg/kg)载体剂量队列。在低剂量队列的患者中平均FIX表达水平对应于4.4%,在高剂量队列中平均FIX表达水平增加到6.9%。这与FIX使用量和年度自发性出血率的降低是一致的,尽管创伤引起的出血的风险与基因疗法干预之前一样保持很高。有意思的是,在基因疗法前接受预防性治疗的9名患者中有8名不再需要预防性治疗。在低剂量和高剂量队列中,FIX表达分别持续一年和约6个月。在3名试验参与者中检测到无症状和短暂的肝转氨酶升高,但这似乎与任何AAV衣壳特异性T细胞应答或FIX活性降低无关。研究表明,在基于AAV的基因疗法中,将血清型从AAV5改变为AAV8并不能预防这种不良事件。因此,患者接受了糖皮质激素的短暂免疫抑制治疗,以阻止这些不必要的免疫应答和转氨酶升高。

[0009] 在产生用于通过基因疗法治疗血友病A的有效病毒基因递送系统中的一个显著限制是FVIII cDNA的较大尺寸。先前针对血友病A的基于病毒载体的基因疗法研究通常依赖于使用小但弱的启动子,需要过高的载体剂量,这些剂量与临床不相关或严重累及载体滴度。探索了其它几种临时策略,例如使用分裂(split)载体或双载体设计来克服AAV的包装限制,但是这些方法总体上相对低效,并且引起了额外的免疫原性问题(在Petrus等人,2010中进行了论述)。已发现FVIII B结构域对于促凝活性是可有可无的。因此,缺失了B结构域的FVIII构建体可用于基因转移目的,因为它们的较小尺寸更容易掺入载体中。此外,已显示B结构域的缺失导致mRNA和主要翻译产物增加17倍。其中B结构域缺失并用14个氨基酸的短接头替换的FVIII目前以重组产品的形式生产,并以Refacto®销售以供临床使用(惠氏制药公司)(Sandberg等人,2001)。Miao等人(2004)将短B结构域序列添加回B结构域

缺失的FV VIII,最佳为226个氨基酸并保留6个位点用于N-连接的糖基化,以改善分泌。McIntosh等人(2013)用17个氨基酸的肽代替Miao等人的226个氨基酸间隔区,在17个氨基酸的肽中将来自B结构域的六个糖基化三重体并置。然而,生产仍不足以用于治疗目的。人类因子VIII cDNA的密码子优化也导致高水平表达。在用由密码子优化的cDNA序列表达FV VIII的慢病毒载体转导的新生患血友病A小鼠的血浆中检测到活性FV VIII蛋白具有显著更高的水平(最多增加44倍,超过200%的正常人的水平),从而成功地矫正了疾病模型(Ward等人,2011)。

[0010] 用于FIX的肝脏特异性表达的示例性现有载体在W0 2009/130208中进行了描述,并且由包含驱动因子cDNA的TTR/Serp调节序列的单链AAV载体组成。载体中包含FIX第一内含子以及多腺苷酸化信号。使用所述改进的载体产生约25-30%稳定的循环因子IX。

[0011] 为了将基于病毒载体的血友病基因疗法转化到临床上,必须解决与向肝脏施用大的载体剂量相关的安全性问题以及制造大量临床级载体的需求。提高基因转移载体的效能(效力/剂量)对于实现这些目标至关重要。它将允许使用较低的剂量来获得疗效,从而减少与体内给药相关的潜在毒性和免疫激活,并缓解制造需求。

[0012] 提高效率的一种方法是设计转基因序列本身,以使每个载体拷贝的表达和生物学活性最大化。本发明人先前已经证实,优化密码子使用并携带与凝血过度活跃和易栓症相关的R338L氨基酸取代的FIX转基因(Simioni等人,2009)使得在血友病B小鼠中使用慢病毒载体或AAV载体的基因治疗的效力提高到15倍,没有检测到不良反应,从而显著降低了达到治疗效果所需的剂量要求,并因此促进未来的规模扩大及其临床转化(Cantore等人,2012; Nair等人,2014)。

[0013] 已基于FIX-R338L Padua变体进行了两项独立的近期基因治疗临床试验,结果令人鼓舞。此外,其它几项试验正在准备之中,这些试验已从战略上将其重点从使用野生型FIX转移到高活性FIX-R338L变体,这实际上表明FIX-R338L变体现在正成为血友病B基因疗法的金标准。在第一次FIX-R338L Padua试验中(Baxalta,现在部分属于Shire; NCT01687608;同样参见Monahan等人(2015)和Horling等人(2017)),重度血友病B患者通过单次静脉注射自互补(sc)AAV8载体进行治疗,所述载体表达密码子优化的FIX-R338L Padua变体(称为BAX335)。在这次试验中,采用了3个剂量队列,共治疗7名患者。其中两名患者的瞬时FIX活性水平超过50%,最终降至基础水平。来自中等剂量队列(10^{12} vg/kg)的一名患者显示FIX水平超过一年保持在20%活性范围内。然而,在本试验纳入的大多数患者中, FIX表达要么无法检测到,要么无法持续。根本原因分析表明, FIX表达降低的潜在原因可能是由于过高的免疫原性所致。实际上, BAX335在FIX编码序列中含有大量CpG二核苷酸基序,其可能有助于通过以Toll样受体9(TLR9)依赖性方式刺激先天免疫系统来增加免疫原性(Faust等人,2013)。

[0014] 与BAX335试验的结果相比,第二次FIX-R338L Padua试验(Spark Therapeutics/Pfizer; NCT02484092)在患者中产生了更一致的响应。设计了单链AAV载体(称为SPK-9001),其从肝细胞特异性启动子表达密码子优化的FIX-F338L Padua变体,所述肝细胞特异性启动子由载脂蛋白E基因肝控制区(APOE)和肝脏特异性 α 1-抗胰蛋白酶(hAAT)启动子组成(George等人,2017)。基于其有利的血清阳性谱,使用另一突变的AAV衣壳(称为AAV-Spark100)包装AAV载体。给10名重度血友病B患者静脉注射了相对低剂量的载体(5×10^{11} vg/

kg)。在载体施用后14周,10名参与者达到了约 $33.7 \pm 18.5\%$ 的稳态FIX活性水平,这与每年的年出血率(从11.1到0.4次出血事件/年)和年输注次数(从67.5到1.2)均下降是一致的。单次静脉注射基因治疗载体后,FIX活性持续了1年。这促使通过蛋白替代疗法中止预防治疗。与所有其它血友病B基因疗法试验一样,这里也有两名试验参与者显示肝转氨酶的短暂升高,这与AAV衣壳特异性T细胞免疫应答相吻合,需要逐渐减少口服皮质类固醇的剂量的治疗。在通过瞬时免疫抑制治疗的一个试验参与者中,FIX表达稳定在正常FIX活性的70-90%范围内。然而,在第二名患者中,免疫抑制不足以防止FIX水平的显著降低。这表明,在转氨酶升高开始之前,尽早使用口服皮质类固醇作为预防性治疗可能是必要且明智的。

[0015] 基于血友病小鼠和非人灵长类动物的令人鼓舞的结果(Bunting等人,2018),给9名重度血友病A患者静脉注射编码密码子优化的B结构域缺失的人因子VIII的AAV5载体(AAV5-hFVIII-SQ)(Rangarajan等人,2017)。在低剂量(1例患者; 6×10^{12} vg/kg)和中剂量(1例患者; 2×10^{13} vg/kg)下,未检测到循环FVIII水平。然而,用最高载体剂量 6×10^{13} vg/kg治疗的7名患者中有6名显示FVIII活性水平显著增加,在20周后达到正常FVIII活性水平的50%以上。该试验中使用的载体剂量似乎大大高于任何基于AAV的血友病B试验中使用的剂量。然而,在没有任何标准的情况下,有必要谨慎比较载体剂量和试验结果。最重要的是,载体注射一年后,接受最高剂量的试验参与者的平均FVIII活性水平达到了 $93 \pm 48\%$ 。然而,在四名患者中,达到了正常FVIII活性水平的150%以上,峰值为正常水平的201%至349%。这引起了一些关于可能增加血栓风险的担忧,但这些超生理水平没有持续,并且在78周时FVIII水平进一步下降并落在生理范围内。先前接受FVIII预防治疗的6例接受最高剂量的患者显示,基因治疗后的年出血率从每年发生16次降至1次。同样,基因治疗后,年度FVIII输注率中位数从每年输注138次降至2次。与B型血友病试验一样,高剂量队列中的一些患者显示转氨酶水平显著升高,需要逐渐减少糖皮质激素治疗的剂量。这表明基于AAV5的基因疗法导致肝脏炎症,尽管没有证据表明是T细胞介导的对AAV5衣壳的免疫应答。这使人联想到上述基于AAV5的血友病B试验中获得的结果(Miesbach等人,2018)。

[0016] 在W02014/064277中描述了将强固型Serpine增强子与编码FIX或FVIII的密码子优化的转基因组合的表达载体,分别导致FIX和FVIII的肝脏特异性表达的增加。

[0017] 然而,在最大化基因表达方面,可以达到诸如启动子工程和转基因密码子优化之类的策略所能达到的极限。最近在蛋白质工程方面的成功证明,将FIX蛋白与白蛋白或免疫球蛋白Fc结构域融合可以显著延长FIX蛋白的半衰期(Peters等人,2010;Powell等人,2013;Metzner等人,2009;Santagostino等人,2016)。因此,与使用标准重组FIX蛋白进行常规蛋白替代疗法以维持足够的止血作用相比,血友病B患者需要输注这些长效因子的频率要低得多。

[0018] 本发明的目的是进一步提高针对血友病A和B的肝脏定向基因疗法的功效和安全性。

发明内容

[0019] 虽然先前显示将FIX蛋白与白蛋白融合显著延长了FIX蛋白的半衰期(Metzner等人,2009;Santagostino等人,2016),但这些研究涉及生产重组FIX-Alb融合蛋白和将所述重组FIX-Alb融合蛋白(而不是载体)施用给受试者。本发明人发现,当施用表达现有技术的

编码FIX和人白蛋白融合蛋白的非密码子优化的融合基因的病毒载体时,与单独表达FIX转基因的病毒载体相比,融合蛋白的循环水平和/或活性没有显著差异,并且没有检测到表达水平的显著增加。这表明将病毒载体用于基因治疗目的不能重复使用具有人白蛋白的重组融合基因所报道的成功。然而,当施用表达包含根据SEQ ID NO:14的密码子优化的人白蛋白和FIX转基因的融合基因的载体时,观察到融合基因的基因表达以及FIX蛋白的循环水平和活性显著增加,所述融合基因编码FIX-人白蛋白融合蛋白。

[0020] 因此,本发明提供了以下方面:

[0021] 方面1.一种编码人白蛋白的密码子优化的核酸分子,其包含由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列,优选由SEQ ID NO:14定义的序列。

[0022] 方面2.根据方面1所述的核酸分子,其包含与SEQ ID NO:14定义的所述序列或与所述序列具有至少80%序列同一性的所述序列融合的转基因,优选地,其中所述转基因是密码子优化的转基因。

[0023] 方面3.根据方面1或2所述的核酸分子,其中所述转基因位于SEQ ID NO:14所定义的所述序列或与所述序列具有至少80%序列同一性的所述序列的5'末端。

[0024] 方面4.根据方面1至3种任一项所述的核酸分子,其中所述转基因通过编码一个或多个多肽或肽接头的序列与由SEQ ID NO:14定义的所述序列或与所述序列具有至少80%序列同一性的所述序列分隔开,所述编码一个或多个多肽或肽接头的序列优选编码由SEQ ID NO:18定义的肽接头。

[0025] 方面5.根据权利要求1至4中任一项所述的核酸分子,其用于增加由转基因编码的蛋白质或多肽的表达和/或循环水平和/或活性。

[0026] 方面6.一种核酸表达盒,其包含根据方面1至4中任一项的核酸分子,所述核酸分子可操作地连接至启动子。

[0027] 方面7.根据方面6所述的核酸表达盒,其包含至少一种组织特异性核酸调节元件,所述调节元件可操作地连接至所述启动子和如方面1至5中任一项所定义的核酸分子。

[0028] 方面8.根据方面6或7所述的核酸表达盒,其包含小鼠微小病毒(MVM)内含子,优选由SEQ ID NO:20定义的MVM内含子。

[0029] 方面9.根据方面6至8中任一项所述的核酸表达盒,其包含转录终止信号,优选多腺苷酸化信号,更优选由SEQ ID NO:21定义的合成的多腺苷酸化信号或SEQ ID NO:23定义的猿猴病毒40(SV40)多腺苷酸化信号。

[0030] 方面10.根据方面6至9中任一项所述的核酸表达盒,其中所述转基因编码分泌型治疗蛋白或分泌型免疫蛋白,优选地,所述转基因编码选自下列的分泌型治疗蛋白:因子IX、因子VIIa、因子VIII、肝细胞生长因子(HGF)、组织因子(TF)、组织因子途径抑制剂(TFPI)、ADAMTS13、血管内皮生长因子(VEGF)、胎盘生长因子(PLGF)、成纤维细胞生长因子(FGF)、可溶性fms样酪氨酸激酶1(sFLT1)、 α 1-抗胰蛋白酶(AAT)、胰岛素、胰岛素原、因子X、血管性血友病因子、C1酯酶抑制剂(C1-INH)、溶酶体酶、溶酶体酶艾杜糖醛酸-2-硫酸酯酶(I2S)、促红细胞生成素(EPO)、干扰素- α 、干扰素- β 、干扰素- γ 、白介素1(IL-1)、白介素2(IL-2)、白介素3(IL-3)、白介素4(IL-4)、白介素5(IL-5)、白介素6(IL-6)、白介素7(IL-7)、白介素8(IL-8)、白介素9(IL-9)、白介素10(IL-10)、白介素11(IL-11)、白介素12(IL-12)、趋化因子(C-X-C基序)配体5(CXCL5)、粒细胞集落刺激因子(G-CSF)、粒细胞-巨噬细胞集落

刺激因子(GM-CSF)、巨噬细胞集落刺激因子(M-CSF)、干细胞因子(SCF)、角质细胞生长因子(KGF)、单核细胞趋化蛋白-1(MCP-1)、肿瘤坏死因子(TNF)、afamin(AFM)、 α 1-抗胰蛋白酶、 α -半乳糖苷酶A、 α -L-艾杜糖醛酸酶、脂蛋白脂肪酶、载脂蛋白(apolipoprotein)、低密度脂蛋白受体(LDL-R)、白蛋白、耐二肽基肽酶4(DPP-4)的胰高血糖素样肽1(GLP-1)、GLP-2、胰高血糖素、生长激素(GH)、干扰素(例如,IFN α 1 α -2b)、 β -利尿钠肽、IL-1Ra、exendin-4、胃酸调节素、卵泡抑素、抗体和纳米抗体。

[0031] 方面11.根据方面10所述的核酸表达盒,其中所述转基因编码用于治疗 and/或预防肝脏相关疾病的治疗性蛋白,所述肝脏相关疾病优选血友病A、血友病B或因子VII缺陷。

[0032] 方面12.根据方面11所述的核酸表达盒,其中所述转基因编码凝血因子FIX(FIX),优选地,其中所述凝血因子FIX包含超激活突变,更优选地,其中所述超激活突变对应于R338L氨基酸取代,更优选地,其中所述转基因编码具有SEQ ID NO:11所定义的核酸序列的凝血因子IX。

[0033] 方面13.根据方面12所述的核酸表达盒,其中所述转基因编码凝血因子VIII(FVIII),优选地,其中所述转基因是密码子优化的凝血因子FVIII,或其中所述凝血因子VIII具有B结构域的缺失,优选地,其中所述FVIII的所述B结构域被SEQ ID NO:15定义的接头替换,更优选地,其中所述转基因编码具有SEQ ID NO:16所定义的核酸序列的凝血因子VIII。

[0034] 方面14.根据方面12所述的核酸表达盒,其中所述转基因编码凝血因子VII(FVII)或因子FVIIa(FVIIa)的轻链和重链,其中所述FVII的轻链通过一个或多个可切割的多肽或肽接头与所述FVII或FVIIa的重链偶联,优选地,其中所述转基因编码SEQ ID NO:34所定义的氨基酸序列。

[0035] 方面15.根据方面11至14中任一项所述的核酸表达盒,其中所述至少一种组织特异性核酸调节元件是至少一种肝脏特异性核酸调节元件。

[0036] 方面16.根据方面15所述的核酸表达盒,其中所述至少一种肝脏特异性核酸调节元件包含由SEQ ID NO:25定义的Serp_{in}增强子或与所述序列具有至少95%同一性的序列。

[0037] 方面17.根据方面16所述的核酸表达盒,其包含三重重复、优选串联排列的SEQ ID NO:25定义的Serp_{in}增强子或与所述序列具有至少95%同一性的序列。

[0038] 方面18.根据方面11至17中任一项所述的核酸表达盒,其中所述启动子是肝脏特异性启动子,优选地,肝脏特异性启动子选自包含以下的群组:甲状腺素运载蛋白(TTR)启动子、最小TTR启动子(TTR_m)、AAT启动子、白蛋白(ALB)启动子或最小启动子、载脂蛋白A1(APOA1)启动子或最小启动子、补体因子B(CFB)启动子、己酮糖激酶(KHK)启动子、血色素结合蛋白(HPX)启动子或最小启动子、烟酰胺N甲基转移酶(NNMT)启动子或最小启动子、(肝)羧酸酯酶1(CES1)启动子或最小启动子、蛋白C(PROC)启动子或最小启动子、载脂蛋白C3(APOC3)启动子或最小启动子、甘露聚糖结合凝集素丝氨酸蛋白酶2(MASP2)启动子或最小启动子、铁调素抗菌肽(HAMP)启动子或最小启动子,或丝氨酸蛋白酶抑制剂C类(抗凝血酶)成员1(SERPINC1)启动子或最小启动子,优选TTR_m。

[0039] 方面19.一种载体,其包含根据方面6至18中任一项所述的核酸表达盒,优选地,其中所述载体是病毒载体,更优选地,其中所述载体源自腺相关病毒(AAV)。

[0040] 方面20.根据方面19所述的载体,其中所述载体是单链AAV。

[0041] 方面21.根据方面20所述的载体,其具有SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8,优选SEQ ID NO:8。

[0042] 方面22.一种药物组合物,其包含根据权利要求19至21中任一项所述的载体和药学上可接受的载体。

[0043] 方面23.根据方面19至21中任一项所述的载体或根据方面22所述的药物组合物,其用于医学,优选基因疗法,更优选肝脏定向基因疗法。

[0044] 方面24.根据方面19至21中任一项所述的载体或根据方面22所述的药物组合物,其用于治疗肝脏相关病症,如果所述转基因是FIX(血友病B)、FVIII(血友病A)或FVII(血友病A、血友病B、优选具有FVIII或FIX抑制剂的患者;FVII缺陷),则所述肝脏相关病症优选血友病A、血友病B或FVII缺陷。

[0045] 方面25.一种在需要这种治疗的受试者中治疗肝脏相关病症,优选血友病的方法,所述方法包括向所述受试者施用治疗有效量的根据方面19至21中任一项所述的载体或根据方面22所述的药物组合物。

[0046] 方面26.根据方面19至21中任一项所述的载体或根据方面22所述的药物组合物在制造用于治疗受试者的肝脏相关病症(优选血友病)的药物中的用途。

[0047] 方面27.一种用于在肝细胞中表达转基因产物的体外或离体方法,其包括:

[0048] -将根据方面6至18中任一项所述的核酸表达盒或根据方面19至21中任一项所述的载体引入到所述肝细胞中;

[0049] -在所述肝细胞中表达所述转基因产物。

[0050] 方面28.根据方面1至4中任一项所述的核酸分子、根据方面6至18中任一项所述的核酸表达盒或根据方面19至21中任一项所述的载体在增加由转基因编码的蛋白质或多肽的表达和/或循环水平和/或活性中的用途,优选地,其中所述用途是体外用途。

[0051] 方面29.一种增加由转基因编码的蛋白质或多肽的表达和/或循环水平和/或活性的方法,其使用根据方面1至4中任一项所述的核酸分子、根据方面6至18中任一项所述的核酸表达盒或根据方面19至21中任一项所述的载体。

附图说明

[0052] 本发明通过以下附图来说明,这些附图仅出于说明目的,并且绝不将本发明限制于其中公开的实施方案:

[0053] 图1:pAAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA载体的质粒图谱

[0054] 图2:pAAVss-1XSERP-mTTR-MVM-hFIXco-Alb-SV40pA载体的质粒图谱

[0055] 图3:pAAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体的质粒图谱

[0056] 图4:pAAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA载体的质粒图谱

[0057] 图5:pAAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA载体的质粒图谱

[0058] 图6:pAAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA载体的质粒图谱

[0059] 图7:pAAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA载体的制粒图谱

[0060] 图8:pAAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA载体的质粒图谱

[0061] 图9:在FIX敲除(KO)小鼠中转导所述载体后的FIX蛋白质水平和活性。9A):以 5×10^9 vg/小鼠转导AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA和AAVss-1XSERP-mTTR-

MVM-hFIXcoPadua-Alb-SV40pA后获得的FIX蛋白质表达水平随时间的变化;9B):以 5×10^9 vg/小鼠转导AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA和AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA后获得的FIX蛋白活性随时间的变化;9C):以 5×10^9 vg/小鼠转导AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA和AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA后获得的蛋白质活性随时间的变化。

[0062] 图10:在鼠中导所述载体后的FIX蛋白质水平和mRNA表达。10A):以 1×10^9 vg/小鼠转导AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA、AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA和AAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA后获得的FIX蛋白表达水平随时间的变化;10B):以 1×10^9 vg/小鼠转导AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA、AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA和AAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA后获得的相对于对照的mRNA表达。

[0063] 图11:在FIX敲除(KO)鼠中导所述载体后的FIX蛋白和活性水平。11A):以 5×10^9 vg/小鼠转导AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA和AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA后获得的蛋白表达水平随时间的变化;11B):以 5×10^9 vg/小鼠转导AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA和AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA后获得的蛋白活性随时间的变化。

[0064] 图12:在FIX敲除(KO)鼠中,以 5×10^8 vg/小鼠、 1×10^9 vg/小鼠或 5×10^9 vg/小鼠转导所述载体后,所获得的FIX抗原水平和FIX活性水平随时间的变化(1周和3周)。

[0065] 图13:在FIX敲除(KO)鼠中,以 5×10^8 vg/小鼠(A)、 1×10^9 vg/小鼠(B;D)或 5×10^9 vg/小鼠(C;E)转导所述载体后,所获得的FIX活性水平随时间的变化(在注射后的数个星期内所表达的)。

[0066] 发明详述

[0067] 本文所用的单数形式“一”、“一个”和“该”包括单数和复数指示物,除非上下文另有明确规定。

[0068] 本文所用的术语“包含”与“包括”或“含有”同义,并且是广义的或开放式的,并且不排除附加的、未列举的成员、元件或方法步骤。这些术语还涵盖“由...组成”和“基本上由...组成”,它们在专利术语中具有公认的含义。

[0069] 用端点对数值范围的描述包括在各自范围内纳入的所有数字和部分,以及所述的端点。

[0070] 当提及诸如参数、量、时间周期等可测量值时,本文所用的术语“约”或“大约”意味着涵盖特定值的变化,例如为特定值的 $\pm 10\%$ 或更小的变化,优选 $\pm 5\%$ 或更少,更优选 $\pm 1\%$ 或更少,且还更优选 $\pm 0.1\%$ 或更少,只要这样的变化适合于在所公开的本发明中实现。应当理解,修饰语“约”所指的本身也明确并且优选地公开。

[0071] 尽管术语“一个或多个”或“至少一个”(例如一个或多个成员或一组成员中的至少一个成员)本身是清楚的,但通过进一步的举例,该术语尤其涵盖对所述成员中的任何一个成员的引用,或所述成员中的任何两个或更多个成员的引用,例如,所述成员中的任何 ≥ 3 个、 ≥ 4 个、 ≥ 5 个、 ≥ 6 个或 ≥ 7 个成员,等等,直至对所有所述成员的引用。在另一实例中,“一个或多个”或“至少一个”可以指1、2、3、4、5、6、7或更多。

[0072] 这里包括对本发明背景的讨论,以解释本发明的上下文。这不应视为承认,自权利

要求书中任一项的优先权日起,所提及的任何材料在任何国家/地区已被发布,知晓或是常识的一部分。

[0073] 在本公开中,各种出版物、专利和已发布的专利说明书通过识别的引文来引用。本说明书中引用的所有文献均通过引用的方式全文并入本文中。具体地,本文明确提及的此类文件的教导或部分通过引用的方式并入。

[0074] 除非另有定义,公开本发明使用的包括技术和科学术语的所有术语具有与本发明所属技术领域的普通技术人员通常所理解的相同的含义。通过进一步的指导,本发明包括术语定义以更好地理解本发明的教导。当结合本发明的特定方面或本发明的特定实施方案来定义特定术语时,除非另有定义,这样的涵义意味着可应用于整个说明书,即也适用于本发明的其它方面或实施方案的上下文中。

[0075] 在下面的段落中,更详细地定义了本发明的不同方面或实施方案。这样定义的每个方面或实施方案可以与任何其它方面或实施方案组合,除非有明确相反的指示。具体而言,表示为优选或有利的任何特征可以与表示为优选或有利的任何其它特征组合。

[0076] 在本说明书中,“一个实施方案”、“实施方案”是指结合该实施方案描述的特定特征、结构或特性包括在本发明的至少一个实施方案中。因此,在本说明书的不同地方出现的短语“在一个实施方案中”或“在一实施方案中”不一定都指同一实施方案,但可以指同一实施方案。此外,在一个或多个实施方案中,对于本领域技术人员显而易见的是,具体特征、结构或特性可以以任何合适的方式组合。此外,本领域技术人员应明白,虽然本文描述的一些实施方案包括一些但在其它实施方案中包括的其它特征,但是不同实施方案的特征的组合意味着在本发明的范围内,并且形成不同的实施方案。例如,在所附权利要求书中,所要求保护的实施方案中的任何一个可以任意组合使用。

[0077] 本发明将针对特定实施方案并参照特定的附图进行描述,但本发明并不限于此而仅由权利要求书限定。权利要求中的任何参考符号不应解释为限制范围。所描述的附图仅是示意性的,而非限制性的。在附图中,为了说明的目的,一些元件的尺寸可能会被放大且未按比例绘制。

[0078] 本文提供的术语或定义是为了帮助理解本发明。除非在本文中特别定义,否则本文中使用的术语具有与本发明所述领域技术人员所理解的相同的含义。实践者关于本领域的定义和术语特别参考 Sambrook 等人,《分子克隆:实验室手册 (Molecular Cloning: A Laboratory Manual)》,第二版,冷泉港出版社,Plainsview,纽约(1989);和 Ausubel 等人,《分子生物学的当前方案 (Current Protocols in Molecular Biology)》(增刊47),John Wiley&Sons,纽约(1999)。

[0079] 本文提供的定义不应被解释为具有小于本领域普通技术人员所理解的范围。

[0080] 本发明人意外地发现密码子优化的人白蛋白(本文中也称为“Albco”)、优选具有由 SEQ ID NO:14 定义的序列或与所述序列具有至少 80% 序列同一性、优选至少 85% 序列同一性的序列的密码子优化的人白蛋白,在与转基因进行基因融合时,可用于增强所述转基因的基因表达和/或增加由所述转基因编码的蛋白质或多肽的水平和/或活性。更具体地,密码子优化的人白蛋白可用于制备编码例如人凝血因子 IX (hFIX) -Alb、人凝血因子 VIII (hFVIII) -Alb 或人凝血因子 FVII -Alb 的基因融合的转基因(在本文中也称为融合基因),以分别增强 hFIX、hFVIII 或 hFVII 的基因表达,和/或在体外和体内分别增加 hFIX、hFVIII 或

hFVII的表达或循环水平和/或活性。相反,如果非密码子优化的(本文也称为“野生型”)人白蛋白(本文也称为“A1b”)与转基因进行基因融合,则未观察到对所述转基因所编码的蛋白质或多肽的基因表达和/或水平和/或活性产生有益影响。

[0081] 因此,第一方面提供了编码人白蛋白的密码子优化的核酸分子以用于增强转基因的基因表达和/或用于增加转基因编码的蛋白质或多肽的水平或/或活性,所述核酸分子包含由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列,优选包含由SEQ ID NO:14定义的序列。

[0082] 本文所用的术语“核酸分子”或“核酸”通常是指基本上由核苷酸组成的任何长度的寡聚物或聚合物(优选线性聚合物)。核苷酸单元通常包括杂环碱基、糖基团和至少一个(例如一个、两个或三个)磷酸基团,包括修饰或取代的磷酸基团。杂环碱基可包括尤其是嘌呤和嘧啶碱基,例如腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)、胸腺嘧啶(T)和尿嘧啶(U),它们广泛存在于天然核酸中;其它天然存在的碱基(例如,黄嘌呤,肌昔,次黄嘌呤),以及化学或生物化学修饰的(例如,甲基化)非天然或衍生的碱基。糖基团可包括尤其是戊糖(戊呋喃糖)基团,例如优选在天然存在的核酸中常见的核糖和/或2-脱氧核糖,或阿拉伯糖、2-脱氧阿拉伯糖、苏糖或己糖糖基团,以及修饰或取代的糖基团。本文意指的核酸可包括天然存在的核苷酸、修饰的核苷酸或其混合物。修饰的核苷酸可以包括修饰的杂环碱基、修饰的糖部分、修饰的磷酸基团或它们的组合。可以引入对磷酸基团或糖的修饰以提高稳定性、抗酶促降解性或其它一些有用的特性。术语“核酸”进一步优选地包括DNA、RNA和DNA/RNA杂合分子,具体地包括hnRNA、前mRNA、cDNA、基因组DNA、扩增产物、寡核苷酸和合成的(例如,化学合成的)DNA、RNA或DNA/RNA杂合体。核酸可以是天然存在的,例如存在于自然界中或从自然界中分离出来;或者可以是非天然存在的,例如重组的,即通过重组DNA技术产生,和/或部分或完全地通过化学或生物化学合成。“核酸”可以是双链、部分双链或单链。在为单链的情况下,核酸可以是有义链或反义链。此外,核酸可以是环状或线性的。这种核酸分子或核酸可以适当地分离。

[0083] 由SEQ ID NO:14定义的核苷酸序列表示如SEQ ID NO:28中定义的野生型或非密码子优化的人白蛋白cDNA的密码子优化形式,所述人白蛋白cDNA的前体形式在NCBI Genbank (<http://www.ncbi.nlm.nih.gov/>) 登录号NM_000477.6下进行了注释。

[0084] 如本文教导的由密码子优化的核酸分子编码的人白蛋白蛋白质可以是如SEQ ID NO:27中定义的野生型人白蛋白蛋白质,其前体形式的氨基酸序列在NCBI参考序列NP_000468.1下进行了注释;或者可以是所述人白蛋白蛋白质的变体或突变体,该变体或突变体相对于相应的天然蛋白质携带氨基酸序列变异,例如氨基酸缺失、添加和/或取代。例如,如Andersen等人,2014和Strohl等人,2015中所述,人白蛋白还可以包括白蛋白的K573P突变(即,其中氨基酸残基位置573上的赖氨酸被脯氨酸取代)。

[0085] 结合转基因使用的术语“密码子优化”是指修饰转基因的密码子而不改变由所述转基因编码的蛋白质或多肽的氨基酸序列。通常,转基因中的稀有密码子(即在要表达转基因的宿主中很少使用的密码子)被宿主生物体的转基因中丰度更大的密码子替换。

[0086] 本文使用的术语“密码子”是指特定的信使RNA(mRNA)分子或编码蛋白质或多肽中的特定氨基酸残基或用于终止翻译(“终止密码子”)的编码DNA(cDNA)中的三个连续核苷酸碱基的任何组。术语“密码子”还包括DNA链中的三联碱基。密码子优化的转基因的计算机预

测可以通过本领域已知的任何方法获得,例如使用商业密码子优化工具,如OptimumGene™基因设计系统(GenScript)或密码子优化工具(OMICS_23398,omicotools)。当然,这些只是预测工具,每个密码子优化的实际效果仍然不确定,需要进行广泛的测试。

[0087] 本领域技术人员将理解,由SEQ ID NO:14定义的核苷酸序列中的有限变异(相对于(即相比于)相应核酸所作的一个或多个核苷酸添加、缺失或取代)仍会如本文所述的SEQ ID NO:14定义的核苷酸序列那样产生编码相同蛋白质的核苷酸序列。

[0088] 因此,在特定实施方案中,本文教导的核酸分子包含由SEQ ID NO:14定义的序列或与SEQ ID NO:14具有至少约80%同一性的序列,例如,优选至少约85%同一性;例如,更优选至少约90%同一性,例如,至少91%同一性、92%同一性;甚至更优选至少约93%同一性,例如,至少94%同一性;甚至更优选至少约95%同一性,例如至少96%同一性;还更优选至少约97%同一性,例如至少98%同一性,并且最优选至少99%同一性。

[0089] 如本文所用,术语“同一性”和“相同”等是指两个聚合物分子之间(例如两个核酸分子(例如两个DNA分子)之间)的序列相似性。序列比对和序列同一性的确定可以例如使用Altschul等人,1990(J Mol Biol 215:403-10)最初所描述的基本局部比对搜索工具(Basic Local Alignment Search Tool, BLAST)、例如Tatusova和Madden 1999 (EMS Microbiol Lett 174:247-250)描述的“Blast 2序列”算法进行。通常,在序列的整个长度上计算序列同一性百分比。如本文所用,术语“基本上相同”表示至少90%、优选至少95%、例如95%、96%、97%、98%或99%的序列同一性。

[0090] 在本说明书中广泛使用的术语“蛋白质”通常包括包含一个或多个多肽链的大分子,即通过肽键连接的氨基酸残基的聚合链。该术语可以包括天然、重组、半合成或合成产生的蛋白质。该术语还包括携带多肽链的一种或多种共表达或后表达类型修饰的蛋白质,所述修饰例如但不限于糖基化、乙酰化、磷酸化、磺化、甲基化、泛素化、信号肽去除、N-末端甲基去除,将酶原或前激素转化为活性形式等。该术语还包括蛋白质变体或突变体,其相对于相应的天然蛋白质携带氨基酸序列变异,例如氨基酸缺失、添加和/或取代。该术语同时考虑到全长蛋白质和蛋白质部分或片段,例如,因加工这种全长蛋白质而产生的天然蛋白质部分。

[0091] 在本说明书中广泛使用的术语“多肽”通常包括包含通过肽键连接的氨基酸残基的聚合链。因此,特别是当蛋白质仅由单个多肽链组成时,术语“蛋白质”和“多肽”在本文中可互换使用以表示这种蛋白质。该术语不受到多肽链的任何最小长度的限制。该术语可以包括天然、重组、半合成或合成产生的多肽。该术语还包括携带多肽链的一种或多种共表达或后表达类型修饰的多肽,例如但不限于糖基化、乙酰化、磷酸化、磺化、甲基化、泛素化、信号肽去除、N-末端甲基去除,将酶原或前激素转化为活性形式等。该术语还包括多肽变体或突变体,其相对于相应的天然多肽携带氨基酸序列变异,例如氨基酸缺失、添加和/或取代。该术语同时考虑到全长多肽和多肽部分或片段,例如,加工这种全长多肽而产生的天然多肽部分。

[0092] 在本说明书中使用的术语“肽”优选地是指如本文所用的多肽,其基本上由50个或更少的氨基酸组成,例如45个氨基酸或更少,优选40个氨基酸或更少,例如35个氨基酸或更少;更优选30个氨基酸或更少,例如,25个或更少、20个或更少、15个或更少、10个或更少或5个或更少的氨基酸。

[0093] 这种蛋白质、多肽或肽可以适当地分离。与特定组分(例如核酸、蛋白质、多肽或肽)有关的术语“分离”通常表示这样的组分以分离的形式存在,例如,该组分已与其自然环境的一种或多种其它成分分离或以分离的形式制备和/或保持。例如,分离的人或动物蛋白或复合物可能与其天然所在的人或动物体分离存在。

[0094] 在特定的实施方案中,本文教导的核酸分子包含与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合的转基因,优选地,其中所述转基因是密码子优化的转基因。

[0095] 在本发明的上下文中,本文使用的术语“融合”与“连接”、“结合”、“偶联”、“接合”同义,并且是指至少两个元件或组分之间的物理连接。在两种基因或编码蛋白质的核苷酸序列的上下文中,“融合”是指编码序列的融合(称为“基因融合”),从而在表达时产生融合蛋白。

[0096] 转基因与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合通常导致形成编码融合蛋白或多肽的融合基因或嵌合基因。术语“融合蛋白”或“融合多肽”表示通过编码融合产物的单个连续多核苷酸分子遗传表达的基因融合的产物,其中两个或更多个蛋白质、多肽或其变体或片段经由它们各自的多肽骨架通过共线性的共价键连接。通常,为了产生编码融合产物的连续多核苷酸分子,将每一个均编码特定多肽片段的两个或更多个开放阅读框(ORF)连接以形成连续的长ORF,其中每个原始ORF都保持正确的阅读框。在所得的重组融合多肽中,由原始PRF编码的两个或更多个多肽片段连接在相同多肽分子中,而它们通常事实上不是这样连接的。虽然因此使得阅读框在整个融合基因片段中是连续的,但是如此融合的多肽片段可以通过例如框内多肽或肽接头在物理上或空间上分隔开,所述框内多肽或肽接头可以是可切割的,也可以是不可切割的。

[0097] 本领域技术人员应明白,取决于转基因和SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的所述序列的顺序,可能需要去除位于最上游序列的终止密码子以避免截断由融合基因编码的融合蛋白或多肽。此外,本领域技术人员还应明白,转基因和SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的所述序列需要在框架内,以便有效地制备由融合基因编码的蛋白质或多肽。

[0098] 如本文所用,术语“转基因”或“(转)基因”是指编码将在其所插入的细胞中表达的多肽或多肽的一部分的特定核酸序列。然而,也可以将转基因表示为RNA,通常是为了降低特定核酸序列所插入的细胞中特定多肽的量。这些RNA分子包括但不限于通过RNA干扰(shRNA、RNAi)、微RNA调节(miRNA)、催化RNA、反义RNA、RNA适配体等发挥其功能的分子。如何将核酸序列引入到细胞中对于本发明不是至关重要的,它可以例如通过整合在基因组中或作为附加体质粒,或借助于病毒或非病毒载体引入。值得注意的是,转基因的表达可能受限于其中插入核酸序列的细胞的子集。术语“转基因”是指包括(1)细胞中非天然存在的核酸序列(即异源核酸序列);(2)作为其中引入了核酸序列的细胞中天然存在的核酸序列的突变形式的核酸序列;(3)用于增加其中引入了核酸序列的细胞中天然存在的相同(即同源的)或相似的核酸序列的额外拷贝的核酸序列;或(4)沉默的天然存在的或同源的核酸序列,其表达在其中引入了核酸序列的细胞中被诱导。“突变形式”是指包含一个或多个与野生型或天然存在的序列不同的核苷酸的核酸序列,即,所述突变核酸序列包含一个或多个核苷酸取代、缺失和/或插入。在某些情况下,转基因还可以包括编码前导肽或信号序列的

序列,使得转基因产物将从细胞分泌。

[0099] 在特定的实施方案中,转基因是密码子优化的转基因。

[0100] 在特定的实施方案中,转基因编码分泌型蛋白质。

[0101] 在特定的实施方案中,转基因编码治疗性蛋白或免疫原性蛋白,优选分泌型治疗蛋白或分泌型免疫原性蛋白。

[0102] 本文使用的术语“分泌型蛋白质”是指在特定细胞或特定组织(例如肝脏)中表达,然后输出到血流以运输到身体其它部分的蛋白质。

[0103] 在实施方案中,转基因编码分泌型治疗蛋白,例如激素、细胞因子、趋化因子、生长因子、胞外酶(例如葡萄糖苷酶、脂蛋白脂酶、 α 1-抗胰蛋白酶)、血浆因子、凝血因子、促红细胞生成素、抗体和纳米抗体。

[0104] 分泌型治疗蛋白的非限制实例包括因子IX、因子VIII、因子VII、因子VIIa(FVIIa)、肝细胞生长因子(HGF)、组织因子(TF)、组织因子途径抑制剂(TFPI)、ADAMTS13、血管内皮生长因子(VEGF)、胎盘生长因子(PLGF)、成纤维细胞生长因子(FGF)、可溶性fms样酪氨酸激酶1(sFLT1)、 α 1-抗胰蛋白酶(AAT)、胰岛素、胰岛素原、因子X、血管性血友病因子、C1酯酶抑制剂(C1-INH)、溶酶体酶、溶酶体酶艾杜糖醛酸-2-硫酸酯酶(I2S)、促红细胞生成素(EPO)、干扰素- α 、干扰素- β 、干扰素- γ 、白介素1(IL-1)、白介素2(IL-2)、白介素3(IL-3)、白介素4(IL-4)、白介素5(IL-5)、白介素6(IL-6)、白介素7(IL-7)、白介素8(IL-8)、白介素9(IL-9)、白介素10(IL-10)、白介素11(IL-11)、白介素12(IL-12)、趋化因子(C-X-C基序)配体5(CXCL5)、粒细胞集落刺激因子(G-CSF)、粒细胞-巨噬细胞集落刺激因子(GM-CSF)、巨噬细胞集落刺激因子(M-CSF)、干细胞因子(SCF)、角质细胞生长因子(KGF)、单核细胞趋化蛋白-1(MCP-1)、肿瘤坏死因子(TNF)、afamin(AFM)、 α 1-抗胰蛋白酶、 α -半乳糖苷酶A、 α -L-艾杜糖醛酸酶、脂蛋白脂肪酶、载脂蛋白(apolipoprotein)(例如载脂蛋白A-I(apoA-I))、低密度脂蛋白受体(LDL-R)、白蛋白、耐二肽基肽酶4(DPP-4)的胰高血糖素样肽1(GLP-1)、GLP-2、胰高血糖素、生长激素(GH)、干扰素(例如,IFN α -2b)、 β -利尿钠肽、IL-1Ra、exendin-4、胃泌酸调节素、卵泡抑素、转基因编码抗体、纳米抗体,极其片段、亚基或突变体,等等。

[0105] 在实施方案中,转基因编码分泌型免疫原性蛋白。分泌型免疫原性蛋白或亚基的非限制性实例包括源自癌细胞(HER2)、病毒(HPV、HBV)、细菌(百日咳、白喉、破伤风)和真菌或寄生虫(疟疾)的抗原。

[0106] 本文所用的术语“免疫原性”是指能够刺激引发免疫应答的物质或组合物。

[0107] 在特定的实施方案中,所述转基因编码分泌型治疗蛋白,以用于治疗 and/或预防与肝脏相关的病症,优选血友病。

[0108] 通常,本文所述的核酸分子、表达盒和载体中的转基因编码凝血因子IX、凝血因子VIII或凝血因子VII(或凝血因子VIIa),优选人凝血因子IX、凝血因子VIII或凝血因子VII(或凝血因子VIIa),更优选人凝血因子IX。

[0109] 术语“凝血因子IX”具有本领域已知的含义。凝血因子IX的同义词是“FIX”或“克雷马斯因子(Christmas factor)”或“F9”,并且可以互换使用。具体而言,术语“凝血因子IX”包括由Genbank登录号NM_000133中所定义的mRNA序列编码的人蛋白。

[0110] 优选地,所述FIX是突变的FIX,其与野生型FIX相比具有高活性或高功能性。修饰人凝血因子的功能活性可以通过生物工程来完成,例如引入点突变。通过这种方法,报道了

一种高活性R338A变体,其在体外活化部分凝血活酶时间测定(APPT)中与野生型人FIX相比,显示出3倍增加的凝血活性(Chang等人,1998)且在用突变的FIX基因转导的血友病B小鼠中显示出2至6倍更高的特异性活性(Schuettrumpf等人,2005)。已经描述了具有甚至更高凝血活性的其它的示例性FIX的点突变体或结构域交换突变体:FIX,其EGF-1结构域用FVII的EGF-1结构域单独替换或联合以下突变体:R338A点突变(Brunetti-Pierrri等人,2009)、V86A/E277A/R338A三重突变体(Lin等人,2010)、Y259F、K265T和/或Y345T单、双或三重突变体(Milanov等人,2012)和G190V点突变体(Kao等人,2010),这些全部通过引用的方式并入本文中。在特别优选的实施方案中,FIX突变体是Simioni等人2009年所述并命名为“因子IX Padua”突变体的突变体,其引起X连锁易栓症。所述突变因子IX是高活性的,并携带R338L氨基酸取代。在本发明的优选实施方案中,本文所述的核酸表达盒和表达载体中使用的FIX转基因编码人FIX蛋白,最优选地是,FIX转基因编码人FIX蛋白的Padua突变体。因此,在本发明的特别优选的实施方案中,所述转基因具有SEQ ID NO:11(即编码人FIX蛋白的Padua突变体的密码子优化的转基因)。

[0111] 术语“凝血因子VIII”具有本领域已知的含义。凝血因子VIII的同义词是“FVIII”或“抗凝血因子”或“AHF”,并且可以互换使用。术语“凝血因子VIII”包括例如具有在Uniprot登录号P00451中定义的氨基酸序列的人蛋白。

[0112] 在实施方案中,所述FVIII是其中B结构域缺失的FVIII(即B结构域缺失的FVIII,在本文中也称为BDD FVIII或FVIII Δ B或FVIIIdeltaB)。术语“B结构域缺失的FVIII”包括例如但不限于其中B结构域的全部或部分缺失的FVIII突变体和其中B结构域被接头替换的FVIII突变体。B结构域缺失的FVIII的非限制性实例在Ward等人(2011)和WO 2011/005968中进行了描述,它们尤其通过全文引用的方式并入本文中。

[0113] 在优选的实施方案中,所述FVIII是B结构域缺失的FVIII,其中B结构域被具有以下序列的接头替换:SFSQNPPVLRHQR(SEQ ID NO:15)(即Ward等人(2011)中定义的SQ FVIII)。在特别优选的实施方案中,所述编码FVIII的转基因具有SEQ ID NO:16(即编码B结构域缺失的人FVIII的密码子优化的转基因,在本文中也称为(h)FVIIIcopt或co(h)FVIIIdeltaB或co(h)FVIIIdeltaB转基因),其同样也在WO 2011/005968中公开,该案通过全文引用的方式并入本文中。

[0114] 在特定实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码凝血因子IX(FIX),优选地,其中所述凝血因子FIX包含超激活突变,更优选地,其中所述超激活突变对应于R338L氨基酸取代,更优选地,其中所述转基因编码具有SEQ ID NO:11所定义的核酸序列的凝血因子IX。

[0115] 在特定实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码凝血因子VIII(FVIII),优选地,其中所述转基因是密码子优化的凝血因子FVIII,或其中所述凝血因子VIII具有B结构域的缺失,优选地,其中所述FVIII的所述B结构域被SEQ ID NO:15定义的接头替换,更优选地,其中所述转基因编码具有SEQ ID NO:16所定义的核酸序列的凝血因子VIII。

[0116] 术语“凝血因子VII”具有本领域已知的含义。凝血因子VII的同义词是“FVII”且可互换使用。术语“凝血因子VIII”包括例如具有在Uniprot登录号P08709中定义的氨基酸序列的人蛋白。凝血FVII通常通过在例如由SEQ ID NO:30定义的氨基酸序列的152位氨基酸

残基(精氨酸)和163位氨基酸残基(异亮氨酸)之间的单肽键的蛋白水解而转变为其活性形式“FVIIa”,从而形成两个多肽链,即,例如由SEQ ID NO:31定义的氨基酸序列的N-端轻链,以及例如由SEQ ID NO:32定义的氨基酸序列的C-端重链,它们通过二硫键保持在一起。先前已表明,凝血因子VII蛋白或凝血因子VIIa蛋白与白蛋白的融合显著延长FVII或FVIIa蛋白的半衰期(Schulte,2008;N0081te,2016;Herzog等人,2014;Zollner等人,2014;Metzner等人,2013;Golor等人,2013)。

[0117] 在特定实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码凝血因子VII,优选地,其中编码FVII的所述转基因是密码子优化的。本文所用的术语“凝血因子VII”是指包含凝血因子VII的轻链(例如SEQ ID NO:31所定义)和重链(例如SEQ ID NO:32所定义)的多肽或蛋白质。激活后,所述轻链和重链通过二硫键彼此偶联(即,产生“活化的凝血因子VII”或“FVIIa”)。因此,在特定的实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码凝血因子VII的轻链和重链,优选地,其中凝血因子VII的轻链和重链通过一个或多个可切割的多肽或肽接头彼此分隔开。

[0118] 在特定的实施方案中,所述一个或多个可切割多肽或肽接头包含至少氨基酸序列RKRRKR(SEQ ID NO:33)、RKR或PRPSRKRR(SEQ ID NO:35),优选RKRRKR(SEQ ID NO:33),如Margaritis等人(2004)先前所述。

[0119] 在优选的实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码由SEQ ID NO:31定义的因子VII的轻链、由SEQ ID NO:32定义的重链,以及包含SEQ ID NO:33所定义的序列的一个或多个可切割多肽或肽接头。在更优选的实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因编码SEQ ID NO:34定义的氨基酸序列。

[0120] 在特定的实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因位于由SEQ ID NO:14定义的所述序列或与所述序列具有至少80%序列同一性的序列的5'端(即上游)。

[0121] 在特定的实施方案中,由转基因编码的蛋白质或多肽和由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列编码的人白蛋白可以通过例如框内多肽或肽接头在物理或空间上分隔开。

[0122] 因此,在特定的实施方案中,本文所述的核酸分子、表达盒和载体中的所述转基因与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列通过编码一个或多个多肽或肽接头的核酸序列分隔开。

[0123] 本文所使用的术语“接头”是指用于连接其它元件的连接元件。优选地,转基因与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列之间的连接可以是水解稳定的连接,即,在适用的pH值下,包括特别是在生理条件下,在水中基本稳定,优选地,其中所述转基因是密码子优化的转基因。

[0124] 在特定的实施方案中,接头是一个或多个氨基酸的肽接头。更具体而言,肽接头可以为1至50个氨基酸长,或2至50个氨基酸长,或1至45个氨基酸长,或2至45个氨基酸长;优选1至40个氨基酸长,或2至40个氨基酸长,或1至35个氨基酸长,或2至35个氨基酸长;更优选1至30个氨基酸长,或2至30个氨基酸长。进一步优选地,接头可以为5至25个氨基酸长或5至20个氨基酸长。尤其优选地,接头可以为5至15个氨基酸长或7至15个氨基酸长。因此,在某些实施方案中,接头可以为1、2、3或4个氨基酸长。在其它实施方案中,接头可以为5、6、7、

8或9个氨基酸长。在另外的实施方案中,接头可以为10、11、12、13或14个氨基酸长。在仍然另外的实施方案中,接头可以为15、16、17、18或19个氨基酸长。在其它实施方案中,接头可以为20、21、22、23、24或25个氨基酸长。

[0125] 构成接头的氨基酸的性质不是特别有相关性,只要由其连接的多肽片段的生物活性基本上没有受到损害,并且接头提供由转基因编码的蛋白质或多肽与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列编码的人白蛋白在空间上想要的分离,优选地,其中所述转基因是密码子优化的转基因。优选的接头基本上是非免疫原性的。

[0126] 在某些优选的实施方案中,肽接头可包含以下氨基酸、由以下氨基酸组成、或基本上由以下氨基酸组成:甘氨酸(G)、丝氨酸(S)、丙氨酸(A)、苏氨酸(T)及其组合。在甚至更优选的实施方案中,肽接头可包含以下氨基酸、由以下氨基酸组成、或基本上由以下氨基酸组成:甘氨酸、丝氨酸及其组合。这种接头提供了特别好的灵活性。在某些实施方案中,接头可仅由甘氨酸残基组成。在某些实施方案中,接头可仅由丝氨酸残基组成。例如,多肽或肽接头包含、组成为或基本上组成为:SSGSGSGSGSGSGSGSGSGSGSGS (SEQ ID NO:17) 或其片段。

[0127] 在特定的实施方案中,如果转基因编码如本文其它地方所述的凝血因子VII (或因子FVIIa)的轻链和重链,则白蛋白与凝血因子VII的轻链和重链之间的接头优选由SEQ ID NO:17定义的接头。

[0128] 在特定的实施方案中,所述多肽或肽接头是可切割的接头。

[0129] 本文所用的术语“可切割的”(有时也称为“可生物降解的”)是指被分离或分裂的能力,更具体而言,将复杂分子分裂成更简单的分子的能力。蛋白水解是蛋白质分解成较小的多肽或氨基酸。蛋白水解通常由称为蛋白酶的细胞酶催化。低pH或高温下无需酶也会引起蛋白水解。根据本发明的融合蛋白的体内降解导致由转基因编码的蛋白质或多肽的释放,其随后可以在目标区域中诱导其生物学效应。

[0130] 在特定的实施方案中,多肽或肽接头可被一种或多种蛋白酶切割,所述蛋白酶对由转基因编码的特定蛋白质的活化具有特异性,优选地,其中所述蛋白质是如本文其它地方所述的分泌型治疗蛋白。

[0131] 在特定的实施方案中,例如,如果转基因是FIX,则多肽或肽接头可被一种或多种也能激活野生型FIX的蛋白酶切割。在特定的更多实施方案中,如果转基因是FIX,则多肽或肽接头源自由FIX轻链的C端和FIX活化肽的N端构成的切割位点,从而允许在FIX激活的同时去除人白蛋白以增加特异性活性,如Metzner等人,2009中所述。在更特定的实施方案中,如果转基因是FIX,则多肽或肽接头包含、基本上组成为、或组成为:SVSQTSKLTRAETVFPDVGDS (SEQ ID NO:18) 或其片段。由SEQ ID NO:18定义的氨基酸序列组成的多肽或肽接头可以通过SEQ ID NO:29定义的序列编码。

[0132] 在核酸分子的特定实施方案中,所述转基因与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列通过SEQ ID NO:29定义的序列或与该序列具有至少85%、优选至少90%、更优选至少95% (例如96%、97%、98%或99%) 同一性的序列或其功能片段分隔开。

[0133] 另一方面提供了本文所教导的由核酸分子编码的融合蛋白。本文所教导的融合蛋白通常包含白蛋白和转基因(或转基因和coA1b),它们任选地通过一个或多个多肽或肽接

头连接。

[0134] 另一方面提供了用于增强转基因的基因表达和/或用于增加由转基因编码的蛋白质或多肽的水平和/或活性的核酸表达盒。更具体地,本发明提供了核酸表达盒,其包含如本文教导的与启动子可操作地连接的核酸分子。

[0135] 本文所用的术语“核酸表达盒”是指包括一种或多种转录控制元件(例如但不限于启动子、增强子和/或调节元件、多腺苷酸化序列,和内含子)的核酸分子,所述转录控制元件指导在一种或多种所需细胞类型、组织或器官中的(转)基因表达。通常,本文所述的核酸表达盒将包含本文教导的核酸分子。

[0136] 本文所用的术语“可操作地连接”是指各种核酸分子元件相对于彼此的排列,使得所述元件彼此功能性地连接并且能够相互作用。这样的元件可以包括但不限于启动子、增强子和/或调节元件、多腺苷酸化序列、一种或多种内含子和/或外显子以及要表达的目的基因(即转基因)的编码序列或要表达的目的融合基因(例如,与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合的转基因)的编码序列。当核酸序列元件适当定向或可操作地连接时,它们一起作用以调节彼此的活性,并最终影响转基因的表达水平。调节是指增加、减少或维持特定元件的活性水平。每个元件相对于其它元件的位置可以用每个元件的5'末端和3'末端来表示,并且任何特定元件之间的距离可以通过介于元件之间的核苷酸或碱基对的数量来表示。

[0137] 本申请中使用的术语“启动子”是指核酸序列,其直接或间接调节与其可操作连接的相应核酸编码序列(例如转基因或内源基因)的转录。启动子可以单独起作用以调节转录,或者可以与一个或多个其它调节序列(例如,增强子或沉默子)协同作用。在本申请的上下文中,启动子通常可操作地连接到调节元件以调节转基因和/或融合基因的转录。启动子可以是同源的(即来自与用核酸表达盒转染的动物、特别是哺乳动物相同的物种)或异源的(即来自与用核酸表达盒转染的动物,特别是哺乳动物不同物种的来源)。因此,启动子的来源可以是任何病毒、任何单细胞原核或真核生物体、任何脊椎动物或无脊椎动物生物体,或任何植物,或者甚至可以是合成启动子(即具有非天然存在的序列),前提是启动子与本文所述的调节元件组合起作用。在优选的实施方案中,启动子是哺乳动物启动子,尤其是鼠或人启动子。启动子的非限制性实例包括逆转录病毒LTR启动子,特别是Rous肉瘤病毒或小鼠白血病病毒LTR、巨细胞病毒(CMV)启动子、SV40启动子、二氢叶酸还原酶启动子、 β -肌动蛋白启动子、磷酸甘油激酶(PGK)启动子和EF1 α 启动子。

[0138] 启动子可以是诱导型或组成型启动子。

[0139] 在特定的实施方案中,包含在本文公开的核酸表达盒和载体中的启动子是组织特异性启动子。组织特异性启动子在特定类型的细胞或组织中具有活性,例如在肝脏、B细胞、T细胞、造血细胞、单核细胞、白细胞、巨噬细胞、肌肉、胰腺腺泡或 β 细胞、内皮细胞、星形胶质细胞、神经元或肺中具有活性。

[0140] 在特定的实施方案中,包含在本文公开的核酸表达盒和载体中的启动子是肝脏特异性启动子,优选本文其它地方描述的肝脏特异性启动子。这是为了增加肝脏特异性和/或避免其它组织中表达的泄漏。

[0141] 术语“肝脏特异性启动子”包括赋予(转)基因肝脏特异性表达的任何启动子。肝脏特异性启动子的非限制性实例提供在肝脏特异性启动子数据库(LSPD, <http://>

rulai.cshl.edu/LSPD/)中,并且例如包括:甲状腺素运载蛋白(transthyretin,TTR)启动子或最小TTR启动子(TTRm)、 α 1-抗胰蛋白酶(AAT)启动子、白蛋白(ALB)启动子或最小启动子、载脂蛋白A1(APOA1)启动子或最小启动子、补体因子B(CFB)启动子、己酮糖激酶(KHK)启动子、血色素结合蛋白(HPX)启动子或最小启动子、烟酰胺N甲基转移酶(NNMT)启动子或最小启动子、(肝)羧酸酯酶1(CES1)启动子或最小启动子、蛋白C(PROC)启动子或最小启动子、载脂蛋白C3(APOC3)启动子或最小启动子、甘露聚糖结合凝集素丝氨酸蛋白酶2(MASP2)启动子或最小启动子、铁调素抗菌肽(HAMP)启动子或最小启动子,或丝氨酸蛋白酶抑制剂(serpin)C类(抗凝血酶)成员1(SERPINC1)启动子或最小启动子。

[0142] 本申请中使用的术语“肝脏特异性表达”是指与其它(即非肝脏)组织或细胞相比,(转)基因(如RNA和/或多肽)在肝脏、肝组织或肝细胞中的优先或主要表达。根据特定实施方案,至少50%、更具体而言至少60%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少98%、至少99%或100%的(转)基因表达发生在肝脏组织或肝细胞内。根据特定的实施方案,肝脏特异性表达需要表达的基因产物没有“泄漏”到肝脏以外的其它器官或组织中,例如肺、肌肉、脑、肾和/或脾中。加以必要的修改,这同样适用于肝细胞特异性表达和肝母细胞特异性表达,这些可认为是肝脏特异性表达的特定形式。在整个申请书中,当在表达的上下文中提及了肝脏特异性时,也明确设想了肝细胞特异性表达和肝母细胞特异性表达。

[0143] 本文所用的术语“肝细胞”包括主要占据肝脏的细胞,并且主要包括肝细胞、卵圆细胞、肝窦内皮细胞(LSEC)和胆管细胞(形成胆管的上皮细胞)。

[0144] 本文所用的术语“肝细胞”是指已经从祖肝母细胞分化从而能够在适当条件下表达肝脏特异性表型的细胞。术语“肝细胞”也指去分化的肝细胞。该术语包括体内细胞和离体培养的细胞,而不管这种细胞是原代的还是传代的。

[0145] 本文所用的术语“肝母细胞”是指中胚层中的胚胎细胞,其分化产生肝细胞、卵圆细胞或胆管细胞。该术语包括体内细胞和离体培养的细胞,而不管这种细胞是原代的还是传代的。

[0146] 在特别优选的实施方案中,启动子是哺乳动物肝脏特异性启动子,尤其是鼠或人肝脏特异性启动子。

[0147] 根据另一特定实施方案,肝脏特异性启动子来自转甲状腺素蛋白(TTR)基因或来自 α 1-抗胰蛋白酶(AAT)基因。根据又一特定实施方案,TTR启动子是最小启动子(本文也称为TTRm或TTRmin),更具体而言是由SEQ ID NO:19定义的最小TTR启动子。

[0148] 根据特定的实施方案,本文公开的核酸表达盒和载体中的启动子是最小启动子。

[0149] 本文所用的“最小启动子”是全尺寸启动子的一部分,其仍能够驱动表达但缺少有助于调节(例如组织特异性)表达的序列中的至少一部分。该定义涵盖了已删除(组织特异性)调节元件的启动子,该启动子能够驱动基因表达但丧失了其以组织特异性方式表达该基因的能力;和已删除(组织特异性)调节元件的启动子,该启动子能够驱动基因(可能降低的)表达,但未必失去了其以组织特异性方式表达该基因的能力。在本领域中已广泛记录了最小启动子,本说明书中提供了最小启动子的非限制性清单。

[0150] 其它序列也可以并入本文公开的核酸表达盒中,通常是为了进一步增加或稳定转基因和/或融合基因产物的表达(例如,内含子和/或多腺苷酸化序列)。

[0151] 任何内含子都可用于本文所述的表达盒中。术语“内含子”包括整个内含子的任何部分,该部分足够大以能够被核剪接装置识别和剪接。通常,短的功能性的内含子序列是优选的,以便使表达盒的大小尽可能小,从而有利于表达盒的构建和操作。在一些实施方案中,所述内含子从编码由所述表达盒内的编码序列编码的蛋白质的基因获得。内含子可以位于编码序列的5'端,编码序列的3'端,或在该编码序列内。内含子位于编码序列的5'端的优点是使内含子干扰多腺苷酸化信号功能的机会最小化。在实施方案中,本文公开的核酸表达盒还包含内含子。合适的内含子的非限制性实例为小鼠微小病毒(MVM)内含子、 β -珠蛋白内含子(betaIVS-II)、因子IX(FIX)内含子A、猿猴病毒40(SV40)小t内含子和 β -肌动蛋白内含子。优选地,所述内含子是MVM内含子,更优选由SEQ ID NO:20定义的MVM微型内含子。将MVM内含子克隆到本文所述的核酸表达盒中显示出导致与其可操作地连接的转基因的出乎意料的高表达水平。

[0152] 因此,在特定实施方案中,核酸表达盒包含小鼠微小病毒(MVM)内含子,优选由SEQ ID NO:20定义的MVM内含子。

[0153] 指导多聚A尾的合成的任何多腺苷酸化信号可用于本文所述的表达盒中,其实例是本领域技术人员所熟知的。示例性的多腺苷酸化信号包括但不限于源自以下的多聚A序列:猿猴病毒40(SV40)晚期基因、牛生长激素(BGH)多腺苷酸化信号、最小兔 β 珠蛋白(mRBG)基因,以及合成多聚A(SPA位点),如Levitt等人 β (1989, Genes Dev 3:1019-1025)(SEQ ID NO:21)所述。优选地,所述多腺苷酸化信号为牛生长激素(BGH)多腺苷酸化信号(SEQ ID NO:22)或猿猴病毒40(SV40)多腺苷酸化信号(SEQ ID NO:23)。

[0154] 因此,在特定实施方案中,核酸表达盒包含转录终止信号,优选多腺苷酸化信号,更优选合成的多腺苷酸化信号(SEQ ID NO:21)或猿猴病毒40(SV40)多腺苷酸化信号(SEQ ID NO:23)。

[0155] 通常,根据本发明的核酸表达盒包含如本文教导的启动子、增强子、核酸分子(例如,包含转基因和由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列的融合基因)和转录终止子。

[0156] 在特定实施方案中,本文教导的核酸表达盒包含至少一个(例如一个、两个、三个、四个、五个或六个,优选三个(串联)重复的)核酸调节元件,所述核酸调节元件可操作地连接至启动子和转基因,所述转基因与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合。

[0157] 在特定的实施方案中,本文教导的核酸表达盒包含至少一个(例如一个、两个、三个、四个、五个或六个,优选三个(串联)重复的)组织特异性核酸调节元件,所述组织特异性核酸调节元件可操作地连接至启动子和转基因,所述转基因与由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合,优选地,其中所述至少一个组织特异性核酸调节元件是可操作地连接至启动子和转基因的肝脏特异性核酸调节元件。

[0158] 本文所用的“调节元件”是指能够调节和/或控制基因的转录、特别是基因的组织特异性转录的转录控制元件,特别是非编码顺式作用转录控制元件。调节元件包括至少一个转录因子结合位点(TFBS),更特别地至少一个组织特异性转录因子的结合位点,例如至少一个肝脏特异性转录因子的结合位点。通常,当与单独由启动子而没有调节元件转基因相比时,本文所使用的调节元件增加或增强启动子驱动基因的表达。因此,调节元件特别

地包含增强子序列,但是应当明白,增强转录的调节元件不限于通常在远端上游的增强子序列,而是可以出现在它们所调节的基因的任何距离处。实际上,本领域已知调节转录的序列可以位于它们在体内调节的基因的上游(例如在启动子区域中)或下游(例如在3'UTR中),并且可位于基因附近或更远的地方。值得注意的是,尽管本文所公开的调节元件通常是天然存在的序列,但这类调节元件(的部分)或调节元件的数个拷贝(即非天然存在的序列)的组合本身也被认为是调节元件。本文所用的调节元件可以是参与转录控制的较大序列的一部分,例如启动子序列的一部分。然而,仅调节元件通常不足以启动转录,为此而需要启动子。

[0159] 包含在本文公开的核酸表达盒和载体中的一种或多种调节元件优选是细胞或组织特异性的,例如对以下细胞或组织具有特异性:肝脏、B细胞、T细胞、造血细胞、单核细胞、白细胞、巨噬细胞、肌肉(例如,WO 2015/110449中所公开的调节元件,其通过全文引用的方式并入本文中)、隔膜(例如,WO 2018/178067中所公开的调节元件,其通过全文引用的方式并入本文中)、胰腺腺泡或β细胞、内皮细胞(例如,如WO 2017/109039中所公开的调节元件,其通过全文引用的方式并入本文中)、星形胶质细胞、神经元或肺。

[0160] 在特定的实施方案中,包含在本文公开的核酸表达盒和载体中的一个或多个调节元件是肝脏特异性的。肝脏特异性调节元件的非限制性实例公开于2009/130208和/或WO 2016/146757中,其具体通过引用的方式并入本文中。肝脏特异性调节元件的另一个实例是源自转甲状腺素运载蛋白(TTR)基因的调节元件,例如SEQ ID NO:24定义的调节元件,在本文中也称为“TTRe”或“TTREnh”(Wu等人,2008)。本申请中使用的术语“肝脏特异性表达”是指与其它(即非肝脏)组织相比,(转)基因(如RNA和/或多肽)在肝脏中优先或主要表达。根据特定实施方案,至少50%(转)基因表达发生在肝脏内。根据特定实施方案,至少60%、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100%的(转)基因表达发生在肝脏内。根据特定的实施方案,肝脏特异性表达需要表达的基因产物没有“泄漏”到其它器官,例如肺、肌肉、心脏和/或肺。这加以必要的修改同样适用于肝细胞特异性表达,其可认为是肝脏特异性表达的特定形式。在整个申请书中,当在表达的上下文中提及了肝脏特异性时,也明确设想了肝细胞特异性表达。类似地,在本申请中使用组织特异性表达时,还设想了主要组成该组织的细胞类型的细胞类型特异性表达。

[0161] 优选地,本文公开的核酸表达盒和载体中的一种或多种调节元件在仅具有有限的长度的同时具有完全的功能性。这允许它在载体或核酸表达盒中使用,而不会过度限制它们的有效负载能力。因此,在实施方案中,本文公开的表达盒和载体中的一种或多种调节元件是具有1000个核苷酸或更少、800个核苷酸或更少、或600个核苷酸或更少、优选400个核苷酸或更少、例如300个核苷酸或更少、200个核苷酸或更少、150个核苷酸或更少、或100个核苷酸或更少的核酸(即,核酸调节元件具有1000个核苷酸、800个核苷酸、600个核苷酸、400个核苷酸、300个核苷酸、200个核苷酸、150个核苷酸或100个核苷酸的最大长度)。然而,应当理解,所公开的核酸调节元件保留调节活性(即关于特异性和/或转录活性),因此它们尤其具有20个核苷酸、25个核苷酸、30个核苷酸、35个核苷酸、40个核苷酸、45个核苷酸、50个核苷酸、55个核苷酸、60个核苷酸、65个核苷酸,或70个核苷酸的最小长度。

[0162] 在优选的实施方案中,本文公开的核酸表达盒和载体中的一种或多种调节元件包含来自SERPINA1调节元件(即控制体内SERPINA1基因表达的调节元件)的序列。所述调节元

件优选包含、基本上组成为或组成为:SEQ ID NO:25定义的序列、与所述序列具有至少85%、优选至少90%、更优选至少95% (例如96%、97%、98%或99%) 同一性的序列,或其功能片段。还优选地是,所述调节元件具有150个核苷酸或更少,优选100个核苷酸或更少的最大长度,并且包含、基本上组成为或组成为:SEQ ID NO:25定义的序列、与所述序列具有至少85%、优选至少90%、更优选至少95% (例如96%、97%、98%或99%) 同一性的序列,或其功能片段。由SEQ ID NO:25组成的肝脏特异性核酸调节元件在本文中称为“SerpIn增强子”、“SerpEnh”或“Serp”。

[0163] 本申请中使用的术语“功能片段”是指本文公开的序列的片段,其保留了调节肝脏特异性表达的能力,即,它们仍然赋予组织特异性,并且它们能够以与它们所源自的序列相同的方式(尽管可能程度不相同)调节(转)基因的表达。片段包含来自其所源自的序列的至少10个连续的核苷酸。在另一特定的实施方案中,片段包含来自其所源自的序列的至少15个、至少20个、至少25个、至少30个、至少35个或至少40个连续的核苷酸。还优选地是,功能片段可包含至少1个、更优选至少2个、至少3个或至少4个、甚至更优选至少5个、至少10个或至少15个存在于其所源自的序列中的转录因子结合位点(TFBS)。

[0164] 在特别优选的实施方案中,本文公开的核酸表达盒和载体包含以下肝脏特异性调节元件的两个或更多个(例如两个、三个、四个、五个或六个)、优选三个(串联)重复序列:包含、基本上组成为或组成为:SEQ ID NO:25或与该序列具有至少85%、优选至少90%、更优选至少95% (例如96%、97%、98%或99%) 同一性的序列的肝脏特异性调节元件;更优选地,具有150个核苷酸或更少、优选100个核苷酸或更少、更优选80个核苷酸或更少的肝脏特异性调节元件,其包含、基本上组成为或组成为:SEQ ID NO:25或与该序列具有至少85%、优选至少90%、更优选至少95% (例如96%、97%、98%或99%) 同一性的序列。优选的包含SEQ ID NO:25的三个串联重复的核酸调节元件在本文中称为“3xSerp”,且由SEQ ID NO:26定义。

[0165] 在更特定的实施方案中,本文教导的核酸表达盒包含至少一种肝脏特异性核酸调节元件,其中所述至少一种肝脏特异性核酸调节元件由SEQ ID NO:25定义的SerpIn增强子或与所述序列具有至少95%同一性的序列组成。在甚至更特定的实施方案中,本文教导的核酸表达盒包含三重重复、优选串联排列的由SEQ ID NO:25定义的SerpIn增强子或与所述序列具有至少95%同一性的序列。

[0166] 当如本文所述的调节元件与启动子和转基因和/或融合基因可操作地连接时,调节元件可以(1)赋予转基因和/或融合基因在体内(和/或在体外的肝细胞/肝细胞系中)显著程度的肝脏特异性表达,和/或(2)可以增加转基因和/或融合基因在肝脏中(和/或在体外的肝细胞/肝细胞系中)的表达水平。

[0167] 在本发明的一个典型实施方案中,公开了核酸表达盒,其包含:

[0168] -肝脏特异性核酸调节元件,优选包含一个或三个串联重复的SerpIn增强子(例如SEQ ID NO:25)的调节元件,

[0169] -肝脏特异性启动子,优选TTRm启动子(例如SEQ ID NO:19所定义),

[0170] -内含子,优选MVM内含子,例如SEQ ID NO:20所定义,以及

[0171] -(转)基因,优选FIX,更优选SEQ ID NO:9定义的密码子优化的因子IX cDNA,甚至更优选地,如SEQ ID NO:11定义的hFIXcoPadua;

[0172] -核酸序列,其编码一种或多种多肽或肽接头,优选编码SEQ ID NO:18中定义的肽接头;

[0173] -由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列,以及

[0174] -转录终止子,优选多腺苷酸化信号,例如SEQ ID NO:23定义的猿猴空泡病毒40或猿猴病毒40(SV40)多腺苷酸化信号或SEQ ID NO:21定义的合成多聚A位点。

[0175] 作为非限制性实例,这样的载体由SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8定义,优选SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8,更优选SEQ ID NO:6或SEQ ID NO:8,甚至更优选SEQ ID NO:8。

[0176] 本文公开的表达盒因此可以这样使用,或者通常,它们可以是核酸载体的一部分。因此,另一方面涉及本文所述的核酸表达盒在载体、特别是核酸载体中的用途。

[0177] 另一方面提供了载体,其包含如本文教导的核酸表达盒,优选地其中所述载体是病毒载体,更优选地其中所述载体源自腺相关病毒(AAV)。

[0178] 本申请中使用的术语“载体”是指核酸分子,如单链或双链DNA,其中可能已经插入另一核酸分子(插入的核酸分子),例如但不限于cDNA分子。载体用于将插入的核酸分子运输到合适的宿主细胞中。载体可以包含允许转录插入的核酸分子和任选地将转录物翻译成多肽的必要元件。插入的核酸分子可以来源于宿主细胞,或者可以来源于不同的细胞或生物体。一旦进入宿主细胞,载体就可以独立于宿主染色体DNA复制或与宿主染色体DNA同时复制,并且可以生成载体及其插入的核酸分子的数个拷贝。

[0179] 因此,术语“载体”也可以定义为促进基因转运到靶细胞中的基因递送载体。该定义包括非病毒载体和病毒载体两者。非病毒载体包括但不限于阳离子脂质、脂质体、纳米颗粒、PEG、PEI等。病毒载体源于病毒,包括但不限于:逆转录病毒、慢病毒、腺相关病毒、腺病毒、疱疹病毒、肝炎病毒等。或者,基因递送系统可用于组合病毒和非病毒组分,例如纳米颗粒或病毒体(Yamada等人,2003)。

[0180] 通常,但不是必需的,病毒载体由于已经失去了在特定细胞中繁殖的能力而是复制缺陷型的,这是因为对复制至关重要的病毒基因已经从病毒载体中消除。然而,有些病毒载体也可适于在特定细胞(例如癌细胞)中特异性复制,并且通常用于触发(癌症)细胞特异性(瘤细胞)溶解。优选的载体源自腺相关病毒、腺病毒、逆转录病毒和慢病毒。

[0181] 逆转录病毒和慢病毒是RNA病毒,其在感染后具有将其基因插入宿主细胞染色体中的能力。已经开发了逆转录病毒和慢病毒载体,它们缺乏编码病毒蛋白的基因,但保留了感染细胞并将其基因插入靶细胞的染色体中的能力(Miller,1990;Naldini等人,1996;VandenDriessche等人,1999)。慢病毒和基于莫洛尼氏鼠白血病毒(MLV)的经典逆转录病毒载体之间的区别在于,慢病毒载体可以转导分裂细胞和非分裂细胞,而基于MLV的逆转录病毒载体只能转导分裂细胞。

[0182] 腺病毒载体被设计为直接施用至活的受试者。与逆转录病毒载体不同,大多数腺病毒载体基因组不整合到宿主细胞的染色体中。实际上,使用腺病毒载体引入细胞中的基因作为染色体外元件(附加体(episome))保留在细胞核中,并存留延长的时间。腺病毒载体将在体内许多不同组织(包括气道上皮细胞、内皮细胞、肝细胞和各种肿瘤)中转导分裂和非分裂细胞(Trapnell,1993;Chuah等人,2003)。另一种病毒载体来源于单纯疱疹病毒,它

是一种大型的双链DNA病毒。牛痘病毒(另一种dsDNA病毒)的重组形式可以容纳较大的插入物,并通过同源重组产生。

[0183] 腺相关病毒(AAV)是一种小型ssDNA病毒,可感染人类和其它灵长类动物,未发现会引起疾病,因此仅引起非常轻微的免疫应答。AAV可以感染分裂细胞和非分裂细胞,且可将其基因组整合到宿主细胞的基因组中。这些特征使AAV成为创建用于基因治疗的病毒载体的非常有吸引力的候选者,尽管载体的克隆能力相对有限。因此,在本发明的优选实施方案中,所使用的载体来源于腺相关病毒(即AAV载体)。

[0184] 已经分离并表征了不同的AAV血清型,例如AAV2血清型、AAV5血清型、AAV8血清型和AAV9血清型,并且本文考虑了所有AAV血清型。特别地,包含如本文公开的FIX转基因的AAV载体优选为AAV9血清型载体,并且包含如本文公开的FVIII转基因的AAV载体优选为AAV8血清型载体。AAV载体也可以是非天然存在的AAV载体,例如衍生自天然存在的载体的AAV载体,其包含修饰的衣壳,从而通过抗AAV抗体影响嗜性或中和。

[0185] 本文公开的AAV载体可以是单链(即ssAAV载体)或自互补的(即scAAV载体)。特别地,包含如本文公开的FIX转基因的AAV载体优选为自互补的,并且包含如本文公开的FVIII转基因的AAV载体优选为单链的。术语“自互补AAV”在本文中指重组的衍生自AAV的载体,其中编码区已被设计成形成分子内双链DNA模板。

[0186] 使用本文公开的腺相关病毒载体的基因疗法显示出诱导对载体中所包含的转基因的免疫耐受。

[0187] 在实施方案中,根据本发明的载体包含以下元件(参考图3、6和8):

[0188] -反向末端重复序列(ITR),任选突变的,

[0189] -肝脏特异性调节元件,优选包含一个或三个串联重复的Serpin增强子(“Serp”或“SerpEnh”)的调节元件(例如,包含SEQ ID NO:25定义的核酸片段的调节元件),

[0190] -启动子,优选如SEQ ID NO:19所定义的最小TTRm启动子,

[0191] -内含子,优选由SEQ ID NO:20定义的MVM内含子,

[0192] -(转)基因,优选SEQ ID NO:9定义的密码子优化的因子IX cDNA,甚至更优选SEQ ID NO:11定义的密码子优化的因子IX Padua cDNA;

[0193] -核酸序列,其编码一种或多种多肽或肽接头,优选编码SEQ ID NO:18中定义的肽接头;

[0194] -由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列,

[0195] -转录终止子,优选多腺苷酸化信号,例如如SEQ ID NO:23定义的猿猴空泡病毒40或猿猴病毒40(SV40)多腺苷酸化信号或SEQ ID NO:21定义的合成多聚A位点,以及

[0196] -反向末端重复序列(ITR)。

[0197] 所述元件的组合导致受试者肝脏中密码子优化的FIX或具有Padua突变的密码子优化FIX的出乎意料的高蛋白水平和活性,而当使用包含野生型非密码子优化形式的人白蛋白cDNA而不是SEQ ID NO:14所定义的序列或与所述序列具有至少80%序列同一性的序列的载体时,未观察到这一现象。

[0198] 优选地,载体是源自腺相关病毒的载体,更优选自互补AAV载体,甚至更优选自互补AAV9血清型载体,例如由SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8定义的载体,优选SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8定义的载体,更优选SEQ ID NO:6或

SEQ ID NO:8定义的载体,甚至更优选SEQ ID NO:8定义的载体。

[0199] 在特定的实施方案中,所述载体是单链AAV。

[0200] 在特定实施方案中,提供了以下质粒/载体:

[0201] pAAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (SEQ ID NO:2),

[0202] pAAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (SEQ ID NO:6),

[0203] pAAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (SEQ ID NO:3),和

[0204] pAAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (SEQ ID NO:8)。

[0205] 在特定实施方案中,本文教导的载体具有SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8,优选SEQ ID NO:3、SEQ ID NO:6或SEQ ID NO:8,更优选SEQ ID NO:6或SEQ ID NO:8,甚至更优选SEQ ID NO:8。在其它实施方案中,所述载体是非病毒载体,例如基于转座子的载体。优选地,所述基于转座子的载体源自Sleeping Beauty (SB) 或PiggyBac (PB)。Ivics等人(1997)中已经描述了优选的SB转座子及其高活性的版本,包括SB100X,如Mates等人(2009)所述。基于PiggyBac的转座子是安全的载体,因为它们不会增加致癌风险。此外,使用这些载体的肝脏定向基因疗法显示出诱导对载体中包含的转基因、特别是hFIX或hFVIII转基因的免疫耐受。

[0206] 基于转座子的载体优选与编码转座酶的载体组合施用以用于基因疗法。例如,基于源自PiggyBac的转座子的载体可以与野生型PiggyBac转座酶(Pbase)或小鼠密码子优化的PiggyBac转座酶(mPBase)一起施用。优选地,所述转座酶是高活性转座酶,例如,含有七个氨基酸取代(I30V,S103P,G165S,M282V,S509G,N538K,N570S)的高活性PB(hyPB)转座酶,如Yusa等人,(2011)中所述,其具体通过引用的方式并入本文中。

[0207] 转座子/转座酶构建体可以通过水动力注射或使用非病毒纳米颗粒递送以转染细胞,例如肝细胞。

[0208] 另一方面提供了药物组合物或药物制剂,其包含如本文所教导的核酸分子、核酸表达盒或载体和药学上可接受的载体,即一种或多种药学上可接受的载体物质和/或添加剂,例如缓冲液、载体、赋形剂、稳定剂等。所述药物组合物可以试剂盒的形式提供。

[0209] 本文使用的术语“药学上可接受的”与本领域意义一致,并且与药物组合物的其它成分相容并且对其接受者无害。本文使用的术语“药学上可接受的盐”是指无机酸加成盐,如盐酸盐、硫酸盐和磷酸盐;或有机酸加成盐,如乙酸盐、马来酸盐、富马酸盐、酒石酸盐和柠檬酸盐。药学上可接受的金属盐的实例是碱金属盐,如钠盐和钾盐;碱土金属盐,如镁盐和钙盐、铝盐和锌盐。药学上可接受的铵盐的实例是铵盐和四甲基铵盐。药学上可接受的有机胺加成盐的实例是含有吗啉和哌啶的盐。药学上可接受的有机胺加成盐的实例是含有赖氨酸、甘氨酸和苯丙氨酸的盐。根据本发明的药物组合物可以口服给药,例如以丸剂、片剂、涂漆片剂、糖衣片剂、颗粒剂、硬明胶胶囊和软明胶胶囊、水性、酒精或油性溶液、糖浆、乳剂或悬浮液的形式,或通过直肠,例如以栓剂的形式。也可以进行肠胃外给药,例如通过皮下、肌肉内或静脉内以注射或输注溶液的形式进行给药。其它合适的给药形式为(例如)经皮或局部给药,例如以软膏、酏剂、喷雾剂或透皮治疗系统的形式,或以鼻喷雾剂或气雾剂混合物的形式吸入给药,或例如以微胶囊、植入物或棒的形式给药。所述药物组合物可以按照本领域技术人员本身已知的方式制备。为此目的,将如本文所定义的核酸表达盒或表达载体、一种或多种固态或液态药学上可接受的赋形剂(如果需要的话)与其它药物活性化合物组

合在一起形成适合的给药形式或剂型,然后可以用作人类医学或兽医学中的药物。

[0210] 根据另一方面,提供了药物组合物,其包含核酸分子,所述核酸分子包含编码治疗性蛋白的转基因,所述转基因与本文教导的SEQ ID NO:14定义的序列与所述序列具有至少80%序列同一性的序列融合;或包含本文教导的这样的核酸分子的核酸表达盒,以及药学上可接受的载体。

[0211] 根据另一方面,所述药物组合物包括包含核酸表达盒的载体,以及药学上可接受的载体,所述核酸表达盒包含核酸分子,所述核酸分子包含编码治疗性蛋白的转基因,所述转基因与本文教导的SEQ ID NO:14定义的序列与所述序列具有至少80%序列同一性的序列融合。根据另一特定的实施方案,所述转基因编码因子IX,并且所述药物组合物用于治疗血友病B,或所述转基因编码因子VIII,并且所述药物组合物用于治疗血友病A。

[0212] 另一方面提供了如本文所教导的核酸分子、核酸表达盒、载体、药物组合物在增强转基因的基因表达和/或增加由转基因编码的蛋白质或多肽的水平和/或活性中的用途,其中所述用途是体外、离体或体内用途,优选地,其中所述用途是体外用途。

[0213] 在特定的实施方案中,与由编码融合蛋白的核酸分子中存在的相同转基因(其未与白蛋白在体外或体内融合)编码的蛋白质或多肽的水平相比,由包含与本文教导的SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列体外或体内融合的转基因的核酸分子编码的融合蛋白的水平为前者的约2倍至约5倍或更多。蛋白质或多肽的水平可以通过任何本领域公认的方法来测定,例如通过基于抗体的测定法,例如蛋白质印迹或ELISA测定法,以评估是否实现了基因产物的治疗性表达。基因产物的表达也可以在生物测定中测量,所述生物测定检测本文其它地方所述的基因产物的酶促活性或生物活性。

[0214] 在特定的实施方案中,与由编码融合蛋白的核酸分子中存在的相同转基因(其未与白蛋白融合)编码的蛋白质或多肽的活性相比,由包含与本文教导的SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合的转基因的核酸分子编码的融合蛋白的活性为前者的约1.5倍至约4倍或更多。提及的多肽或蛋白质的“活性”通常可以包括多肽或蛋白质的生物学活性的任何一个或多个方面,例如但不限于其生化活性、酶活性、信号传导活性、相互作用活性、配体活性、和/或结构(例如在细胞、组织、器官或生物体内的)活性的一个或多个方面。蛋白质或多肽的活性可以通过本领域已知的任何方法确定,并且取决于蛋白质或多肽的类型和活性的类型。例如,如果蛋白质或多肽是FIX,则可以使用显色法(HYPHEN BioMed,Andresy,France)来确定活性。

[0215] 在特定的实施方案中,本文所教导的核酸分子、核酸表达盒、载体和药物组合物增加了由包含转基因的核酸分子编码的融合蛋白的体外半衰期(其可以由较高的稳态蛋白质水平反映)以及体内循环半衰期(其可以由较高的稳态蛋白质水平反映),因此也增加了由所述转基因编码的蛋白质或多肽的半衰期,所述转基因与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合。更具体地说,与由融合蛋白中相同的转基因编码但未与白蛋白融合的蛋白质或多肽的稳态蛋白水平相比,本文所教导的融合蛋白的稳态水平(以及由转基因编码的蛋白或多肽的稳态水平)是前者的约1.5倍至5倍或更多。蛋白质或多肽的半衰期可以通过本领域已知的任何方法,例如通过药代动力学研究来确定。

[0216] 白蛋白与目的蛋白质或多肽(例如治疗性蛋白)的基因融合改善了所述目的蛋白质或多肽的药代动力学特性,更具体而言,延长了其半衰期。与未与人白蛋白融合的蛋白质

或多肽相比,本文所教导的融合蛋白可以认为是长效融合蛋白。

[0217] 此外,本文所述的表达盒和载体指导治疗量的融合基因产物的长期表达。通常,预期治疗性表达持续至少20天、至少50天、至少100天、至少200天、至少300天、至少1年、至少2年、至少3年、至少4年、至少5年,至少6年,至少7年,至少8年,至少9年,在某些情况下为10年或更长时间。在另一个方面,本文所述的核酸分子、核酸表达盒和载体可用于基因治疗。

[0218] 本领域已经广泛地描述了旨在在靶细胞中(在体外,但也尤其是在体内)实现治疗性基因产物表达的基因治疗方案。这些包括但不限于,肌肉注射质粒DNA(裸露的或在脂质体中),间质注射,气道滴注,施用于内皮、肝内实质,以及静脉内或动脉内给药(例如肝动脉内,肝静脉内)。已经开发了各种装置来增强DNA对靶细胞的可用性。一种简单的方法是将靶细胞与含有DNA的导管或可植入材料物理接触。另一种方法是利用无针喷射装置,该装置在高压下将管柱内液体直接投射到目标组织中。这些递送的范例也可用于递送病毒载体。靶基因递送的另一种方法是使用分子缀合物,所述分子缀合物由蛋白质或合成配体组成,所述蛋白质或合成配体上已附着核酸或DNA结合剂,以将核酸特异性靶向细胞(Cristiano等人,1993)。

[0219] 根据特定的实施方案,预期将本文所述的核酸分子、核酸表达盒和载体用于特定类型的细胞或组织的基因疗法(例如,肝脏(即肝脏定向基因疗法)、肌肉(即肌肉定向基因疗法)、内皮细胞(即内皮特异性基因治疗)、优选肝细胞的基因疗法(即肝脏定向基因疗法)。根据另一个特定实施方案,将所述核酸分子、表达盒或载体用于基因疗法,特别是体内肝脏定向基因疗法。根据另一个特定实施方案,将基因治疗的方法、特别是肝脏定向基因疗法用于治疗血友病、特别是治疗血友病B或血友病A。

[0220] 已经使用离体和体内程序进行了向哺乳动物肝细胞的基因转移。离体方法需要收获肝细胞,用长期表达载体进行体外转导,并将转导的肝细胞重新引入门脉循环(Kay等人,1992;Chowdhury等,1991)。体内靶向已经通过将DNA或病毒载体注射到肝实质、肝动脉或门静脉中以及通过转录靶向完成(Kuriyama等人,1991;Kistner等人,1996)。最近的方法还包括裸DNA的门静脉递送(Budker等人,1996)和水动力尾静脉转染(Liu等人,1999;Zhang等人,1999)。

[0221] 根据另一方面,提供了用于在细胞、优选肝细胞中表达蛋白质的方法,所述方法包括以下步骤:

[0222] -在细胞、优选肝细胞中引入如本文所述的核酸表达盒或载体,以及

[0223] -在细胞、优选肝细胞中表达融合基因产物。

[0224] 这些方法在体外和体内都可进行。

[0225] 另一方面提供了如本文所教导的核酸分子、核酸表达盒、载体、药物组合物用于治疗疾病或病症、优选通过基因疗法治疗疾病或病症的用途。

[0226] 还提供了用于有需要的受试者的基因治疗的方法,所述方法包括在受试者的器官、优选肝脏中引入核酸表达盒的步骤,所述核酸表达盒包含核酸分子,所述核酸分子包含与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合的转基因,其中所述转基因编码治疗性蛋白,并在所述器官、优选肝脏中表达治疗量的治疗性蛋白。根据另一实施方案,所述方法包括在受试者的器官、优选肝脏中引入载体的步骤,所述载体包含核酸表达盒,所述核酸表达盒包含核酸分子,所述核酸分子包含与SEQ ID NO:14定义的

序列或与所述序列具有至少80%序列同一性的序列融合的转基因,其中所述转基因编码治疗性蛋白,并在所述器官、优选肝脏中表达治疗量的治疗性蛋白。

[0227] 可受益于使用如本文所教导的编码人白蛋白的核酸分子、核酸表达盒、载体或药物组合物的基因疗法的示例性疾病和病症包括:肝脏疾病、肝脏相关病症如血友病(包括血友病A和B)、肌小管肌病(MTM)、Pompe症、肌营养不良(例如Duchenne型肌营养不良(DMD)/Becker型肌营养不良(BMD))、肌强直性营养不良、强直性肌营养不良(DM)、Miyoshi肌病、福山型先天性肌营养不良、肌营养不良性神经肌肉疾病、运动神经元疾病(MND)、如Charcot-Marie-Tooth症(CMT)、脊髓性肌萎缩症(SMA)以及肌萎缩性侧索硬化症(ALS)、Emery-Dreifuss肌营养不良症、面肩肱型肌营养不良症(FSHD)、先天性肌营养不良症、先天性肌肉疾病、肢带肌营养不良症、代谢性肌病、肌肉炎性疾病、肌无力、线粒体肌病、离子通道异常、核被膜疾病、心肌病、心脏肥大、心力衰竭、远端肌病、心血管疾病、血管性血友病、微血管血栓形成、血栓性血小板减少性紫癜、外周血管疾病、冠状动脉疾病、动脉粥样硬化疾病、中风、心脏病、糖尿病、胰岛素抵抗、慢性肾衰竭、肿瘤生长、转移、静脉血栓形成、缺血、肿瘤生长、肿瘤血管化、癌症和病毒感染性疾病、例如埃博拉、登革热、登革出血热、自身免疫性疾病(例如克罗恩氏病、多发性硬化症和溶酶体贮积症)。

[0228] 在特定的实施方案中,所述疾病或病症是肝脏相关的疾病或病症。

[0229] 本文使用的术语“肝脏相关疾病或病症”是指与肝脏中改变的基因表达相关的疾病或病症。肝脏相关的疾病或病症包括狭义的肝脏疾病以及一些遗传性疾病,这些遗传性疾病不会直接导致肝脏疾病,但主要在身体的其它地方表现出来。肝脏相关的疾病或病症的非限制性实例包括血友病(包括血友病A和B)、肝炎、癌症和肝硬化、多囊性肝病(PLD)、血友病A或B、家族性高胆固醇血症、溶酶体贮积症,鸟氨酸转氨甲酰酶缺乏症和 α -抗胰蛋白酶缺乏症。

[0230] 转基因以及组织特异性启动子和/或组织特异性调节元件的选择通常与旨在使用本文教导的载体或药物组合物治疗的疾病或病症相关。例如,如果要治疗的疾病或病症是肝脏相关的疾病,例如血友病,则所使用的启动子和调节元件优选是肝脏或肝细胞特异性的,并且转基因优选是FIX、FVIII或FVII,更优选FIX或FVIII。

[0231] 根据非常特定的实施方案,由核酸表达盒或载体中的转基因编码的治疗性蛋白是因子IX,并且所述方法是治疗血友病B的方法。通过基因疗法在肝脏中表达因子IX,可以治疗血友病B(Snyder等人,1999)。根据另一个非常特定的实施方案,由核酸表达盒或载体中的转基因编码的治疗性蛋白是因子VIII,且所述方法是治疗血友病A的方法。根据另一个非常特定的实施方案,由核酸表达盒或载体中的转基因编码的治疗性蛋白为因子VII(或凝血因子VIIa),且所述方法为治疗血友病A、血友病B或FVII缺陷的方法。

[0232] 本文还提供了在需要这种治疗的受试者中治疗如本文其它地方所述的可受益于治疗基因疗法的疾病或病症的方法,所述病症优选与肝脏相关的病症,更优选血友病,所述方法包括向受试者施用治疗有效量的如本文教导的载体或药物组合物。

[0233] 除非另有说明,术语“受试者”或“患者”可互换使用,并且指动物,优选脊椎动物,更优选哺乳动物,且具体而言包括人类患者和非人类哺乳动物,例如小鼠。优选的患者或受试者是人类受试者。

[0234] 本文所用的术语“诊治”或“治疗”是指治疗性治疗和预防性或防护性措施,其中目

标是预防或减缓(减轻)不希望的生理变化或病症,例如增生性疾病(例如癌症)的发展或扩散。有益的或期望的临床结果包括但不限于症状缓解、疾病程度减弱、疾病状态稳定(即不恶化)、疾病进展延迟或减缓、疾病状态改善或缓解,以及康复(无论是部分还是全部),无论其是可检测的还是不可检测的。“治疗”也可以表示与不接受治疗时的预期生存期相比延长生存期。

[0235] 本文所用的诸如“需要治疗的受试者”的短语包括将受益于特定病症的治疗的受试者,例如哺乳动物受试者,所述特定病症例如血友病B或血友病A。这些受试者通常包括但不限于那些已经被诊断出患有所述病症的受试者、容易患有或发展所述病症的受试者和/或要预防该病症的受试者。

[0236] 术语“治疗有效量”是指有效治疗受试者中的特定病症,即,获得期望的局部或全身效果和性能的化合物或药物组合物的量。因此,该术语是指研究人员、兽医、医生或其它临床医生正寻求的在组织、系统、动物或人中引起生物或药物反应的化合物或药物组合物的量,所述反应包括缓解正在治疗的疾病或病症的症状。特定言之,这些术语是指以单剂量或多剂量形式预防、治愈、改善或至少最小化与特定病症(例如如果转基因编码的治疗性蛋白是因子IX或VIII,则所述病症为血友病)相关的临床损伤、症状或并发症所必需的根据本发明的化合物或药物组合物的量,。

[0237] 在特定的实施方案中,如果由转基因编码的治疗性蛋白是因子IX(并且该转基因与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合),那么该术语意味着血浆中因子IX的水平等于或高于生理活性的至少约1%的治疗浓度,即10mU/ml(毫单位/毫升)血浆;至少5%生理活性或50mU/ml血浆、至少10%生理活性或100mU/ml血浆、至少15%生理活性或150mU/ml血浆、至少20%生理活性或200mU/ml血浆、至少25%生理活性或250mU/ml血浆、至少30%生理活性或300mU/ml、至少35%生理活性或350mU/ml、至少40%生理活性或400mU/ml、至少45%生理活性或400mU/ml、至少45%生理活性或450mU/ml、至少50%生理活性或500mU/ml、至少65%生理活性或650mU/ml、至少70%生理活性或700mU/ml、至少75%生理活性或750mU/ml、至少80%生理活性或800mU/ml、至少85%生理活性或850mU/ml、至少95%生理活性或950mU/ml、或至少100%生理活性或1000mU/ml可以在受试者中通过将根据本文所述的任何一个实施方案的载体转导或转染到该受试者中获得。由于本发明的核酸表达盒和载体具有非常高的效率和/或由此获得的融合基因产物具有更长的半衰期,因此即使施用相对低剂量的载体,也可以在受试者中获得这种高治疗水平的因子IX。

[0238] 在另一个特定的实施方案中,如果由转基因编码的治疗蛋白是因子VIII(并且该转基因与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合),那么该术语意味着等于或高于10mU/ml(毫单位/毫升)血浆、50mU/ml血浆、100mU/ml血浆、150mU/ml血浆、200mU/ml血浆、250mU/ml血浆、300mU/ml血浆、350mU/ml血浆、400mU/ml血浆、450mU/ml血浆、500mU/ml血浆、550mU/ml血浆、600mU/ml血浆、650mU/ml血浆、750mU/ml血浆、800mU/ml血浆、850mU/ml血浆、900mU/ml血浆、950mU/ml血浆或更高的治疗浓度的血浆中因子VIII的低谷水平可以通过将本文公开的任何载体转导或转染到受试者中来获得。由于本文所述的载体和核酸表达盒具有非常高的效率和/或由此获得的融合基因产物具有更长的半衰期,因此即使施用相对低剂量的载体,也可以在受试者中获得这种高治疗水平

的因子VIII。

[0239] 在另一个特定的实施方案中,如果由转基因编码的治疗蛋白是因子VII或凝血因子VIIa(并且该转基因与SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合),那么该术语意味着等于或高于10U/ml(单位/毫升)血浆、50U/ml血浆、100U/ml血浆、或更高的治疗浓度(或等于或高于如Abshire等人,2004中所述的治疗浓度)的血浆中因子VII或因子VIIa的低谷水平可通过将本文公开的任何载体转导或转染到受试者中来获得。由于本文所述的载体和核酸表达盒具有非常高的效率和/或由此获得的融合基因产物(即融合蛋白)具有更长的半衰期,因此即使施用相对低剂量的载体,也可以在受试者中获得这种高治疗水平的因子VII(或因子VIIa)。

[0240] 另一方面提供了如本文教导的载体或药物组合物,其用作药物。

[0241] 另一方面提供了如本文教导的载体或药物组合物,其用于治疗如本文其它地方所述的可受益于基因治疗的疾病或病症,优选肝脏相关病症,如果转基因编码FIX、FVIII或FVII,则更优选血友病。

[0242] 本文还提供了如本文教导的载体或药物组合物在制备用于治疗受试者中如本文其它地方所述的可受益于基因疗法的疾病或病症的药物中的用途,所述病症优选肝脏相关病症,更优选血友病。

[0243] 在特定的实施方案中,如果转基因编码因子IX或因子VIII,则可以以低于 6×10^{13} vg/kg(病毒基因组/千克)的剂量将根据本文定义的任一实施方案的载体转导到受试者中,以在受试者中获得100mU/ml血浆或更高的治疗性因子IX水平。例如,可以以低于 5×10^{11} vg/kg的剂量在受试者中实现300mU/ml血浆或更高的因子IX水平。

[0244] 对于血友病疗法,例如可以通过评估受试者因血友病引起的出血来测量治疗的功效。还可以进行体外测试,例如但不限于体外活化部分凝血活酶时间测定(APPT)、测试因子IX显色活性测定、凝血时间、因子IX或人因子VIII特异性ELISA。当然可以使用本领域已知的用于评估治疗效果的任何其它测试。

[0245] 对于特定的病症,本发明的核酸表达盒、载体或药物组合物可以单独使用或与任何已知疗法组合使用。例如,已知的血友病疗法包括施用重组或纯化的凝血因子。因此,本发明的核酸表达盒、载体或药物组合物可以单独施用或与一种或多种活性化合物组合施用。后者可以在施用所述试剂之前、之后或同时施用。

[0246] 还预期本文所公开的核酸分子、核酸表达盒和载体组分在制备用于治疗特定病症,优选肝脏相关疾病,更优选血友病,甚至更优选血友病B或血友病A的药物组合物中的用途。

[0247] 在一个替代实例中,本文公开的核酸分子、表达盒和载体可用于表达免疫量的基因产物(例如多肽,特别是免疫原性蛋白或RNA),以用于疫苗接种目的。

[0248] 在实施方案中,所述药物组合物可以是疫苗。所述疫苗可进一步包含一种或多种用于增强免疫应答的佐剂。合适的佐剂包括例如但不限于皂苷、矿物凝胶如氢氧化铝、表面活性物质如溶血磷脂、复合多元醇、聚阴离子、肽、油或羟乳液、卡介菌(BCG)、短棒状杆菌和合成的佐剂QS-21。任选地,疫苗可进一步包含一种或多种免疫刺激分子。免疫刺激分子的非限制性实例包括具有免疫刺激、免疫增强和促炎活性的各种细胞因子、淋巴因子和趋化因子,例如白介素(例如IL-1、IL-2、IL-3、IL-4、IL-12、IL-13);生长因子(例如,粒细胞-巨

噬细胞 (GM) - 集落刺激因子 (CSF)) ; 和其它免疫刺激分子, 如巨噬细胞炎性因子、Flt3配体、B7.1、B7.2等。

[0249] 在实施方案中, 本文所述的核酸分子、核酸表达盒、载体或药物组合物可以用作疫苗, 尤其用作预防性疫苗。

[0250] 本文还公开了本文所述的核酸分子、核酸调节元件、核酸表达盒、载体或药物组合物在制备疫苗, 特别是在制备预防性疫苗中的用途。

[0251] 本文还公开了对需要所述疫苗接种的受试者进行疫苗接种、特别是预防性疫苗接种的方法, 所述方法包括:

[0252] - 在受试者中, 尤其是在受试者的肝脏中引入本文教导的核酸表达盒、载体或药物组合物; 其中所述核酸表达盒、载体或药物组合物包含至少一种转基因, 所述转基因与可操作地连接至启动子的由SEQ ID NO:14定义的序列或与所述序列具有至少80%序列同一性的序列融合; 以及

[0253] - 在受试者中, 尤其是在受试者的肝脏中表达免疫有效量的融合基因产物 (即融合蛋白)。

[0254] 本文所用的“免疫有效量”是指有效增强受试者对随后暴露于由 (转) 基因编码的免疫原的免疫应答的 (转) 基因产物的量。例如, 可以通过测量中和分泌和/或血清抗体的量, 例如通过空斑中和、补体结合、酶联免疫吸附或微量中和法来确定诱导免疫的水平。

[0255] 应当理解的是, 尽管本文已经针对根据本发明的方法和应用讨论了特定的实施方案、特定的构造和配置以及材料, 但是在不脱离本发明的范围和精神的情况下, 可以在形式和细节上进行各种改变或修改。

[0256] 提供以下实例是为了更好地说明特定实施方案, 并且它们不应被视为限制本申请。本申请仅受权利要求书的限制。

实施例

[0257] 实施例1: 人白蛋白的密码子优化

[0258] 通过GeneArt (GeneArt AG) 使用其内部专用软件GeneOptimizer进行的使用密码子适应指数分析已知人白蛋白cDNA (SEQ ID NO:28) 和适应于智人密码子偏好的密码子用法, 产生密码子优化的序列。

[0259] 尽可能消除可能对表达产生负面影响的负顺式作用位点 (例如剪接供体和受体位点、内部TATA盒、chi位点和核糖体进入位点、RNA不稳定性基序、重复序列和RNA二级结构等)。

[0260] 调整GC含量以延长mRNA半衰期。更具体而言, 尽可能避免极高 (>80%) 或极低 (<30%) GC含量的区域。

[0261] 密码子用法导致密码子适应指数 (CAI) 值为0.96 (参见Nucleic Acids Res.1987Feb 11;15 (3):1281-95;The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications.;Sharp PM, Li WH)。CAI描述了密码子与目标生物体的密码子用法偏好的匹配程度, 并且优选>0.9。

[0262] 实施例2: hFIXco-Albco融合导致稳态hFIX水平和活性的强劲增加

[0263] 1. 研究设计

[0264] 通过常规克隆和合成基因组生成以下载体构建体:

[0265] 1) AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA (SEQ ID NO:1)

[0266] 2) AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (SEQ ID NO:2)

[0267] 3) AAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (SEQ ID NO:3)

[0268] 4) AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA (SEQ ID NO:4)

[0269] 5) AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA (SEQ ID NO:5)

[0270] 6) AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (SEQ ID NO:6)

[0271] 7) AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA (SEQ ID NO:7)

[0272] 8) AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (SEQ ID NO:8)

[0273] 相应的载体质粒图谱和序列如图1-8和12中所示。AAVss对应于单链(ss) AAV载体主链,如先前在VandenDriessche等人,2007中所述。SERP对应于源自SERPINA1基因的顺式调节元件,与先前出版物中的HS-CRM8相同(Nair等人,2014; Chuah等人,2014)。载体包含该SERP元件的单个副本(称为1XSERP)或三重重复(称为3XSERP),如所示。mTTR对应于最小甲状腺素运载蛋白启动子,MVM对应于小鼠微小病毒内含子。载体还包含TTR基因的5'非翻译区(5' UTR),其位于TTR最小启动子(mTTR)的下游。hFIXco对应于先前描述的密码子优化的人(h)FIX基因(Nair等人,2014; Chuah等人,2014)。Padua指的是R338L获得功能性的FIX突变,最初由Simioni及其同事在易栓症患者中描述(Simioni等人,2009)。Alb指的是野生型非密码子优化的白蛋白序列,而Albco指的是相应的密码子优化的白蛋白序列。SV40 pA对应于SV40多腺苷酸化位点。hFIXcoAlbco、hFIXcoPadua-Alb和hFIXcoPadua-Albco融合构建体包含接头(SEQ ID NO:18),其允许融合蛋白的合成,如先前所述(Metzner等人,2009; Santagostino等人,2016)。具有高活性Padua突变(即FIX-Padua)和肝细胞特异性启动子的密码子优化的FIX(coFIX)的产生和初始表征先前已作描述(Cantore等人,2012; Nair等人,2014)。

[0274] 不同FIX基因的序列如下:hFIXco(SEQ ID NO:9);hFIXco-Albco(SEQ ID NO:10);hFIXcoPadua(SEQ ID NO:11);hFIXcoPadua-Alb(SEQ ID NO:12);hFIXcoPadua-Albco(SEQ ID NO:13)。

[0275] 通过用含有AAV载体的质粒和编码AAV8-DJ衣壳的辅助构建体(Grimm等人,2008; Gao等人,2004)共转染293T细胞来产生AAV载体,如Nair等人,2014; Chuah等人,2014所述。通过氯化铯超离心纯化载体,并通过用载体特异性引物定量实时PCR测定载体滴度,如Nair等人,2014; Chuah等人,2014所述。以指示的AAV载体剂量静脉注射成年C57B16和血友病B小鼠(Wang等人,1997)。通过酶联免疫吸附测定(ELISA)测定FIX抗原水平,并如Nair等人,2014; Chuah等人,2014所述且根据制造商的用法说明,通过显色法(HYPHEN BioMed, Andresy, France)测定FIX活性。动物实验得到了大学动物伦理委员会的批准。mRNA表达水平分别通过定量实时PCR和定量实时逆转录酶PCR测定,如Nair等人,2014; Chuah等人,2014所述。

[0276] 2. 结果

[0277] 申请人首先评估了密码子优化与非密码子优化的白蛋白融合对循环人FIX水平和活性的影响。向血友病FIX缺陷小鼠(FIX敲除或FIX KO)以 5×10^9 vg/小鼠静脉注射(i.v.) AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA、AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-

Alb-SV40pA、AAVss-1X SERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA载体。所有载体以AAV8-DJ衣壳包装。结果表明,与编码未与白蛋白融合的hFIXcoPadua蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA)相比,编码密码子优化的hFIXcoPadua-Albco融合蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA)产生了明显更高的FIX活性(约3倍)(图9A)。在治疗的动物中FIX活性水平得以维持,这与缺乏抗hFIX或抗hAlb免疫应答相一致。

[0278] 相比之下,以 5×10^9 vg/小鼠静脉注射AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA与AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA载体的血友病FIX缺陷型小鼠的循环FIX抗原水平没有显著差异(图9B)。同样,与使用hFIXcoPadua-Albco融合基因相比,以 5×10^9 vg/小鼠静脉注射AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA载体的血友病FIX缺陷型小鼠中循环FIX活性水平与用AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA载体所获得的水平相似或甚至略低(图9C)。总的来说,这些结果表明,只要对白蛋白融合本身进行适当的密码子优化,FIX-白蛋白融合就会导致更高的FIX活性水平。

[0279] 为了确认白蛋白融合的作用,随后使用非Padua hFIXco作为治疗性转基因评估其对野生型C57B16小鼠中FIX水平的影响。因此,向正常C57B16小鼠以 10^9 vg/小鼠注射AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA或AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体。结果表明,与非融合对照(即AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA)相比,注射AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体后白蛋白融合显著增加了循环FIX抗原水平(图10A)。因此,使用FIX-白蛋白融合的基因治疗的增加的功效可以基于两种不同的蛋白质,即高活性Padua FIX-R338L(图9)和野生型FIX(图10)获得,且因此与Padua FIX-R338L突变无关。

[0280] 此外,图10B显示了注射AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA或AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体的小鼠之间肝脏中FIX的mRNA表达基本相同,这表明与非融合对照(即AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA)(图10A)相比,注射AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体后循环FIX抗原水平的增加可归因于hFIX-白蛋白融合蛋白的半衰期增加,而不是其mRNA表达增加。

[0281] 为了进一步提高FIX-白蛋白基因治疗方法的治疗效果,我们评估了包括SERP元件的多个拷贝是否会进一步提高稳态FIX水平。因此,向正常C57B16小鼠以 10^9 vg/小鼠注射AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA或AAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA载体。结果表明,SERP元件的多个拷贝增强了hFIX-白蛋白融合蛋白的循环水平(图10A),这与hFIXco-Albco mRNA水平的显著增加相一致(图10B)。

[0282] 最后,使用基于多个SERP元件的优化的载体设计,我们证实了hFIXcoPadua-Albco转基因与hFIXcoPadua-Alb相比具有更优越的治疗效果。因此,向血友病FIX缺陷小鼠以 5×10^9 vg/小鼠注射AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA或AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA载体。结果表明,与编码未与白蛋白融合的hFIXcoPadua蛋白的AAV载体(即AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-SV40pA)相比,编码密码子优化的hFIXcoPadua-Albco融合蛋白的AAV载体(即AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA)产生了明显更高的FIX抗原和活性水平(约2至4倍)(图11A和图11B)。

[0283] 为了进一步证实白蛋白融合的效果,随后使用非Padua hFIXco或Padua-hFIXco (“hFIXcoPadua”)作为治疗转基因,通过综合剂量-响应分析,评估其在血友病FIX中对FIX抗原水平和FIX活性的影响。因此,向血友病FIX缺陷小鼠以 5×10^8 vg/小鼠、 1×10^9 vg/小鼠或 5×10^9 vg/小鼠注射AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA (SEQ ID NO:1)、AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (SEQ ID NO:2)、AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA (SEQ ID NO:4)或AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (SEQ ID NO:6)载体。在载体注射后1周后和3周测量FIX抗原和活性水平(图12和13)。

[0284] 结果表明,与用编码未与白蛋白或密码子优化的白蛋白 (“Albco”)融合的hFIXcoPadua蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA)所获得的抗原和活性水平相比,编码密码子优化的hFIXcoPadua-Albco融合蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA)产生了明显更高的FIX抗原和活性水平(图12和图13A-C)。同样,结果还表明,与用编码未与白蛋白或密码子优化的白蛋白 (“Albco”)融合的hFIXco蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA)所获得的抗原和活性水平相比,编码密码子优化的hFIXcoPadua-Albco融合蛋白的AAV载体(即AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA)产生了明显更高的FIX抗原和活性水平(图12和图13D-E)。此外,用编码hFIXcoPadua的AAV载体获得的FIX活性水平明显高于用编码非高活性hFIXco的AAV载体获得的活性水平(图12和图13A-E)。这与在白蛋白融合或非融合对照中观察到的相一致。FIX抗原和活性水平随着载体剂量的增加而增加(图12和图13A-E)。以 5×10^9 vg/小鼠注射AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA后达到最高FIX抗原和活性水平(1000-1200%的FIX活性:10到12倍的生理FIX水平)(图12和图13C)。

[0285] AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA白蛋白融合与非融合AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA对照构建体之间的FIX抗原和活性水平的差异在最低载体剂量(5×10^8 vg/小鼠)(10至19倍)(图12和图13A)和中间载体剂量(10^9 vg/小鼠)(4至9倍)(图12和图13B)下比在最高载体剂量下(5×10^9 vg/小鼠)(图12和图13C)更显著,这表明可能存在饱和效应。同样,AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA白蛋白融合与非融合AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA对照构建体之间的FIX抗原和活性水平的差异在中间载体剂量(10^9 vg/小鼠)(2至5倍)(图12和图13D)下也比在最高载体剂量下(5×10^9 vg/小鼠)(图12和图13E)更显著。总而言之,这些结果证实,使用FIX-白蛋白融合的基因治疗的增加的功效可以基于两种不同的蛋白质获得,即高活性Padua FIX-R338L和野生型FIX,且因此与Padua FIX-R338L突变无关。

[0286] 参考文献

[0287] ABSHIRE T, KENET G (2004). Recombinant factor VIIa: a review of efficacy, dosing regimen and safety in patients with congenital and acquired factor VIII or IX inhibitors. *J Thromb Haemost.* 2(6):899-909.

[0288] ANDERSEN JT, DALHUS B, VIUFF D, ET AL. Extending serum half-life of albumin by engineering neonatal Fc receptor (FcRn) binding (2014). *J Biol Chem*; 289:13492-13502.

[0289] ANNONI A, BROWN BD, CANTORE A, SERGI LS, NALDINI L, and RONCAROLO MG.

(2009). In vivo delivery of a microRNA-regulated transgene induces antigen-specific regulatory T cells and promotes immunologic tolerance. *Blood* 114, 5152-5161

[0290] ARRUDA VR, STEDMAN HH, HAURIGOT V, and BUCHLIS G. (2010). Peripheral transvenular delivery of adeno-associated viral vectors to skeletal muscle as a novel therapy for hemophilia B. *Blood* 115, 4678-88.

[0291] AXELROD JH, READ MS, BRINKHOUS KM, and VERMA IM. (1990). Phenotypic correction of factor IX deficiency in skin fibroblasts of hemophilic dogs. *Proc Natl Acad Sci USA*; 87, 5173-7.

[0292] BROWN BD, SHI CX, POWELL S, HURLBUT D, GRAHAM FL, and LILLICRAP D. (2004). Helper-dependent adenoviral vectors mediate therapeutic factor VIII expression for several months with minimal accompanying toxicity in a canine model of severe hemophilia A. *Blood* 103, 804-10.

[0293] BROWN BD, CANTORE A, ANNONI A, SERGI LS, LOMBARDO A, DELLAVALLE P, D'ANGELO A, and NALDINI L. (2007). A microRNA-regulated lentiviral vector mediates stable correction of hemophilia B mice. *Blood* 110, 4144-52.

[0294] BRUNETTI-PIERRI N, GROVE NC, ZUO Y, EDWARDS R, PALMER D, CERULLO V, TERUYA J, NG P. Bioengineered factor IX molecules with increased catalytic activity improve the therapeutic index of gene therapy vectors for hemophilia B. *Hum Gene Ther*. 2009 May; 20(5): 479-85.

[0295] BUCHLIS G, PODSAKOFF GM, RADU A, HAWK SM, FLAKE AW, MINGOZZI F, and HIGH KA. (2012). Factor IX expression in skeletal muscle of a severe hemophilia B patient 10 years after AAV-mediated gene transfer. *Blood* 119, 3038-41.

[0296] Budker V, Zhang G, Knechtle S, Wolff JA. Naked DNA delivered intraportally expresses efficiently in hepatocytes. (1996) *Gene Ther*. Jul; 3(7): 593-8.

[0297] BUNTING S, ZHANG L, XIE L, BULLENS S, MAHIMKAR R, FONG S, SANDZA K, HARMON D, YATES B, HANDYSIDE B, SIHN CR, GALICIA N, TSURUDA L, O'NEILL CA, BAGRI A, COLOSI P, LONG S, VEহার G, CARTER B. Gene Therapy with BMN 270 Results in Therapeutic Levels of FVIII in Mice and Primates and Normalization of Bleeding in Hemophilic Mice. *Mol Ther*. 2018 Feb 7; 26(2): 496-509.

[0298] Cantore A, Nair N, Della Valle P, Di Matteo M, Mâtrai J, Sanvito F, Brombin C, Di Serio C, D'Angelo A, Chuah M, Naldini L, Vandendriessche T. Hyper-functional coagulation factor IX improves the efficacy of gene therapy in hemophilic mice. *Blood*. 2012. Oct 4.

[0299] Chang, J., Jin, J., Lollar, P., Bode, W., Brandstetter, H., Hamaguchi, N., Straight, D.L. & Stafford, D.W. (1998). Changing residue 338 in human factor IX from arginine to alanine causes an increase in catalytic activity. *J Biol Chem* 273 (20): 12089-12094.

- [0300] Chowdhury JR, Grossman M, Gupta S, Chowdhury NR, Baker JR Jr, Wilson JM. (1991) Long-term improvement of hypercholesterolemia after ex vivo gene therapy in LDLR-deficient rabbits. *Science*. Dec 20; 254(5039):1802-5.
- [0301] CHUAH MK, SCHIEDNER G, THORREZ L, BROWN B, JOHNSTON M, GILLIJNS V, HERTEL S, VAN ROOIJEN N, LILLICRAP D, COLLEN D, VANDENDRIESSCHE T, and KOCHANNEK S. (2003). Therapeutic factor VIII levels and negligible toxicity in mouse and dog models of hemophilia A following gene therapy with high-capacity adenoviral vectors. *Blood* 101,1734-43.
- [0302] CHUAH MK, NAIR N, VANDENDRIESSCHE T. Recent progress in gene therapy for hemophilia. *Hum Gene Ther*. 2012a Jun; 23(6):557-65.
- [0303] CHUAH MK, NAIR N, VANDENDRIESSCHE T. Recent progress in gene therapy for hemophilia. *Hum Gene Ther*. 2012b Jun; 23(6):557-65.
- [0304] CHUAH MK, VANDENDRIESSCHE T. Platelet-directed gene therapy overcomes inhibitory antibodies to factor VIII. *J Thromb Haemost*. 2012c Aug; 10(8):1566-9
- [0305] CHUAH MK, PETRUS I, DE BLESER P, ET AL. Liver-specific transcriptional modules identified by genome-wide in silico analysis enable efficient gene therapy in mice and non-human primates. *Mol Ther*. 2014; 22(9):1605-1613.
- [0306] DONSANTE A, MILLER DG, LI Y, VOGLER C, BRUNT EM, RUSSELL DW, and SANDS MS. (2007). AAV vector integration sites in mouse hepatocellular carcinoma. *Science* 317,477.
- [0307] DOBRZYNSKI E, FITZGERALD JC, CAO O, MINGOZZI F, WANG L, and HERZOG RW (2006) Prevention of cytotoxic T lymphocyte responses to factor IX-expressing hepatocytes by gene transfer-induced regulatory T cells. *Proc Natl Acad Sci USA* 103,4592-4597.
- [0308] EHRHARDT A, and KAY MA. (2002). A new adenoviral helper-dependent vector results in long-term therapeutic levels of human coagulation factor IX at low doses in vivo. *Blood* 99,3923-30.
- [0309] FAUST SM, BELL P, CUTLER BJ, ASHLEY SN, ZHU Y, RABINOWITZ JE, WILSON JM. CpG-depleted adeno-associated virus vectors evade immune detection. *J Clin Invest*. 2013 Jul; 123(7):2994-3001.
- [0310] FIELDS PA, ARRUDA VR, ARMSTRONG E, KIRK CHU, MINGOZZI F, HAGSTROM J, HERZOG R, HIGH KA. (2001). Risk and prevention of anti-factor IX formation in AAV-mediated gene transfer in the context of a large deletion of F9. *Mol Ther*. 4, 201-210.
- [0311] FINN JD, NICHOLS TC, SVORONOS N, ET AL. The efficacy and the risk of immunogenicity of FIX Padua (R338L) in hemophilia B dogs treated by AAV muscle gene therapy. *Blood*. 2012 Nov 29; 120(23):4521-3.
- [0312] FOLLENZI A, BATTAGLIA M, LOMBARDO A, ANNONI A, RONCAROLO MG, and NALDINI L. (2004). Targeting lentiviral vector expression to hepatocytes limits

transgene-specific immune response and establishes long-term expression of human antihemophilic factor IX in mice. *Blood* 103,3700-9.

[0313] GAO GP, ALVIRA MR, WANG L, JOHNSTON J, WILSON JM. (2002). Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci USA* 99,11854-9.

[0314] GAO G, VANDENBERGH LH, ALVIRA MR, LU Y, CALCEDO R, ZHOU X, and WILSON JM. (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. *J. Viro* 178,6381-6388.

[0315] GEORGE LA, SULLIVAN SK, GIERMASZ A, ET AL. Hemophilia B Gene Therapy with a High-Specific-Activity Factor IX Variant. *N Engl J Med*. 2017;377(23):2215-2227.

[0316] GOLOR G, BENSEN-KENNEDY D, HAFFNER S, EASTON R, JUNG K, MOISES T, LAWO JP, JOCH C, VELDMAN A (2013). *J Thromb Haemost*. 11(11):1977-85

[0317] GRIMM D, LEE JS, WANG L, ET AL. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J Virol*. 2008;82(12):5887-5911.

[0318] HERZOG RW, YANG EY, COUTO LB, HAGSTROM JN, ELWELL D, FIELDS PA, BURTON M, BELLINGER DA, READ MS, BRINKHOUS KM, PODSAKOFF GM, NICHOLS TC, KURTZMAN GJ, and HIGH KA. (1999). Long-term correction of canine hemophilia B by gene transfer of blood coagulation factor IX mediated by adeno-associated viral vector. *Nat Med*. 5,56-63.

[0319] HERZOG RW, MOUNT JD, ARRUDA VR, HIGH KA, and LOTHROP CD Jr. (2001). Muscle-directed gene transfer and transient immune suppression result in sustained partial correction of canine hemophilia B caused by a null mutation. *Mol Ther*. 4,192-200.

[0320] HERZOG RW, HAGSTROM JN, KUNG SH, TAI SJ, WILSON JM, FISHER KJ, and HIGH KA. (1997) Stable gene transfer and expression of human blood coagulation factor IX after intramuscular injection of recombinant adeno-associated virus. *Proc Natl Acad Sci USA*. 94,5804-5809.

[0321] HERZOG RW, FIELDS PA, ARRUDA VR, BRUBAKER JO, ARMSTRONG E, MCCLINTOCK D, BELLINGER DA, COUTO LB, NICHOLS TC, HIGH KA (2002) Influence of vector dose on factor IX-specific T and B cell responses in muscle-directed gene therapy. *Hum Gene Ther* 13,1281-1291.

[0322] HERZOG E, HARRIS S, MC EWEN A, HENSEN C, PRAGST I, DICKNEITE G, SCHULTE S, ZOLLNER S (2014) *Thromb Res* 134(2):495-502

[0323] HIGH KA. (2001). Gene Transfer as an approach to treating Hemophilia. *Circ Res*. 88,137-144.

[0324] HIGH KA. (2011) Gene therapy for hemophilia: a long and winding road. *J Thromb Haemost*. 9 Suppl.1:2-11.

- [0325] HORLING F,FALKER F,CHUAH M. (2017)Development of SHP648,Shire's high performing AAV8-based FIX gene therapy vector.Human Gene Therapy.28(12):A18
- [0326] BAINBRIDGE J,SMITH AJ,BARKER S,et al. (2008)Effect of Gene Therapy on Visual Function in Leber's Congenital Amaurosis.N Engl J Med.358,2231-2239.
- [0327] JIANG H,LILLICRAP D,and PATARROYO-WHITE S. (2006) .Multiyear therapeutic benefit of AAV serotypes 2,6,and 8 delivering factor VIII to hemophilia A mice and dogs.Blood.108,107-15.
- [0328] Kao,C.Y.,Lin,C.N.,Yu,I.S.,Tao,M.H.,Wu,H.L.,Shi,G.Y.,Yang,Y.L.,Kao, J.T.&Lin,S.W. (2010) .FIX-Triple,a gain-of-function factor IX variant,improves haemostasis in mouse models without increased risk of thrombosis.Thromb Haemost 104(2):355-365.
- [0329] Kay MA,Baley P,Rothenberg S,Leland F,Fleming L,Ponder KP,Liu T, Finegold M,Darlington G,Pokorny W,Woo SLC. (1992)Expression of human alpha 1-antitrypsin in dogs after autologous transplantation of retroviral transduced hepatocytes.Proc Natl Acad Sci U S A.Jan 1;89(1):89-93.
- [0330] KAY MA,MANNO CS,RAGNI MV,COUTO LB,MCCLELLAND A,GLADER B,CHEW AJ,TAI SJ,HERZOG RW,ARRUDA V,JOHNSON F,SCALLAN C,SKARSGARD E,FLAKE AW,and HIGH KA. (2000).Evidence for gene transfer and expression of factor IX in hemophilia B patients treated with an AAV vector.Nat Genet.24,257-61.
- [0331] Kistner A,Gossen M,Zimmermann F,Jerecic J,Ullmer C,Lybbert H,Bujard H. (1996)Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice.Proc Natl Acad Sci U S A.Oct 1;93(20):10933-8.
- [0332] KREN BT,UNGER GM,SJEKLOCHA L,TROSSEN AA,KORMAN V,DIETHELEM-OKITA BM, REDING MT,and STEER CJ. (2009) .Nanocapsule-delivered Sleeping Beauty mediates therapeutic Factor VIII expression in liver sinusoidal endothelial cells of hemophilia A mice.J Clin Invest.19,2086-99.
- [0333] Kuriyama S,Yoshikawa M,Ishizaka S,Tsujii T,Ikenaka K,Kagawa T,Morita N,Mikoshiha K. (1991)A potential approach for gene therapy targeting hepatoma using a liver-specific promoter on a retroviral vector.Cell Struct Funct.Dec; 16(6):503-10.
- [0334] LI H,MALANI N,HAMILTON SR,SCHLACHTERMAN A,BUSSADORI G,EDMONSON SE, SHAH R,ARRUDA VR,MINGOZZI F,WRIGHT JF,BUSHMAN FD,and HIGH KA. (2011) .Assessing the potential for AAV vector genotoxicity in a murine model.Blood.117,3311-9.
- [0335] Lin,C.N.,Kao,C.Y.,Miao,C.H.,Hamaguchi,N.,Wu,H.L.,Shi,G.Y.,Liu,Y.L., High,K.A.&Lin,S.W. (2010) .Generation of a novel factor IX with augmented clotting activities in vitro and in vivo.J Thromb Haemost 8(8):1773-1783.
- [0336] Liu F,Song Y,Liu D. (1999)Hydrodynamics-based transfection in animals by systemic administration of plasmid DNA.Gene Ther.Jul;6(7):1258-66.
- [0337] MANNO CS,PIERCE GF,and ARRUDA VR. (2006) .Successful transduction of

liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med.*12,342-7.

[0338] MÁTÉS L, CHUAH MK, BELAY E, JERCHOW B, MANOJ N, ACOSTA-SANCHEZ A, GRZELA DP, SCHMITT A, BECKER K, MATRAI J, MA L, SAMARA-KUKO E, GYSEMANS C, PRYPUTNIEWICZ D, MISKEY C, FLETCHER B, VANDENDRIESSCHE T, IVICS Z, and IZSVAK Z. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet.*41,753-61.

[0339] MARGARTITIS P, VALDER R A, ALJAMALI M, CAMIRE R M, SCHLACHTERMAN A, HIGH K A (2004) *J Clin Invest.* Novel therapeutic approach for hemophilia using gene delivery of an engineered secreted activated Factor VII.113(7):1025-1031.

[0340] MÁTRAI J, CHUAH MK, and VANDENDRIESSCHE T. (2010a). Pre clinical and clinical progress in hemophilia gene therapy. *Curr Opin Hematol.*17,387-92.

[0341] MÁTRAI J, CHUAH MK, and VANDENDRIESSCHE T. (2010b). Recent advances in lentiviral vector development and applications. *Mol Ther.*18,477-90.

[0342] MÁTRAI J, CANTORE A, BARTHOLOMAE CC, ANNONI A, WANG W, ACOSTA-SANCHEZ A, SAMARA-KUKO E, DE WAELE L, MA L, GENOVESE P, DAMO M, ARENS A, GOUDY K, NICHOLS TC, VON KALLE C, L CHUAH MK, RONCAROLO MG, SCHMIDT M, VANDENDRIESSCHE T, and NALDINI L. (2011). Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatology* 53,1696-707.

[0343] MATSUI H, SHIBATAM, BROWN B, LABELLE A, HEGADRON C, ANDREWS C, CHUAH M, VANDENDRIESSCHE T, MIAO CH, HOUGH C, and LILLICRAP D. (2009). A murine model for induction of long-term immunologic tolerance to factor VIII does not require persistent detectable levels of plasma factor VIII and involves contributions from Foxp3+T regulatory cells. *Blood.*114,677-85.

[0344] MATSUI H, HEGADORN C, OZELO M, BURNETT E, TUTTLE A, LABELLE A, McCARY PB Jr., NALDINI L, BROWN B, HOUGH C, and LILLICRAP D. (2011). A microRNA-regulated and GP64-pseudotyped lentiviral vector mediates stable expression of FVIII in a murine model of Hemophilia A. *Mol Ther.*19,723-30.

[0345] McCARTY DM, MONAHAN PE, and SAMULSKI RJ. (2001). Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther.*8,1248-54.

[0346] McCARTY DM, FU H, MONAHAN PE, TOULSON CE, NAIK P, and SAMULSKI RJ. (2003). Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.*10,2112-8.

[0347] MCINTOSH, J. ET AL. Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant. *Blood* (2013).

- [0348] MENDELL JR,AL-ZAIDY S,SHELL R,ET AL.Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy.N Engl J Med.2017;377(18):1713-1722.
- [0349] METZNER HJ,WEIMER T,KRONTHALER U,ET AL.Genetic fusion to albumin improves the pharmacokinetic properties of factor IX.Thromb Haemost.2009;102(4):634-644.
- [0350] METZNER HJ,PIPE SW,WEIMER T,SCHULTE S(2013)Extending the pharmacokinetic half-life of coagulation factors by fusion to recombinant albumin.Thromb Haemost.110(5):931-939
- [0351] MIAO CH,OHASHI K,PATIJN GA,MEUSE L,YE X,THOMPSON AR,and KAY MA . (2000).Inclusion of the hepatic locus control region,an intron,and untranslated region increases and stabilizes hepatic factor IX gene expression in vivo but not in vitro.Mol Ther.1,522-32.
- [0352] MIAO H.Z.,SIRACHAINAN N.,PALMER L.,et al.(2004).Bioengineering of coagulation factor VIII for improved secretion.Blood 103(9):3412-3419.
- [0353] MIESBACH W,MEIJER K,COPPENS M,ET AL.Gene therapy with adeno-associated virus vector 5-human factor IX in adults with hemophilia B.Blood.2018;131:1022-1031.
- [0354] MIESBACH W,MEIJER K,COPPENS M,KAMPMANN P,KLAMROTH R,SCHUTGENS R,TANGELDER M,CASTAMAN G,SCHWÄBLE J,BONIG H,SEIFRIED E,CATTANEO F,MEYER C,LEE BEEK FWG.Gene therapy with adeno-associated virus vector 5-human factor IX in adults with hemophilia B.Blood.2018 Mar 1;131(9):1022-1031.
- [0355] Milanov,et al.,2012 Engineered factor IX variants bypass FVIII and correct hemophilia A phenotype in mice Blood 119:602-611.Miller AD. (1990) Retrovirus packaging cells.HumGene Ther.Spring;1(1):5-14.
- [0356] MINGOZZI F,LIU YL,DOBZYNSKI E,KAUFHOLD A,LIU JH,WANG Y,ARRUDA VR,HIGH KA,and HERZOG RW. (2003).Induction of immune tolerance to coagulation factor IX antigen by in vivo hepatic gene transfer.J Clin Invest.111,1347-56.
- [0357] MINGOZZI F,MAUS MV,HUI DJ,SABATINO DE,MURPHY SL,RASKO JE,RAGINI MV,MANNO CS,SOMMER J,JIANG H,PIERCE GF,ERTL HC,and HIGH KA. (2007).CD8(+) T-cell responses to adeno-associated virus capsid in humans.Nat Med.13,419-22.
- [0358] MONAHAN PE,SUN J,GUI T,ET AL.Employing a gain-of-function factor IX variant R338L to advance the efficacy and safety of hemophilia B human gene therapy:preclinical evaluation supporting an ongoing adeno-associated virus clinical trial.Hum Gene Ther.2015 Feb;26(2):69-81.
- [0359] MONAHAN P,WALSH C,POWELL J.S. (2015)Update on a phase 1/2 open-label trial of BAX335,an adeno-associated virus 8(AAV8)vector-based gene therapy program for hemophilia B.Journal of Thrombosis and Haemostasis.13:87
- [0360] MOUNT JD,HERZOG RW,TILLSON DM,GOODMAN SA,ROBINSON N,MCCLELAND ML,BELLINGER D,NICHOLS TC,ARRUDA VR,LOTHROP CD JR,and HIGH KA. (2002).Sustained

phenotypic correction of hemophilia B dogs with a factor IX null mutation by liver-directed gene therapy. *Blood* 99,2670-6.

[0361] Nair N, Rincon MY, Evens H, Sarcar S, Dastidar S, Samara-Kuko E, Ghandeharian O, Man Viecelli H, Thöny B, De Bleser P, VandenDriessche T, Chuah MK. (2014). Computationally designed liver-specific transcriptional modules and hyperactive factor IX improve hepatic gene therapy. *Blood* 123,3195-9.

[0362] NAIR N, RINCON MY, EVENS H, ET AL. Computationally designed liver-specific transcriptional modules and hyperactive factor IX improve hepatic gene therapy. *Blood*. 2014;123(20):3195-3199.

[0363] CRUDELE JM, FINN JD, SINER JI, ET AL. AAV liver expression of FIX-Padua prevents and eradicates FIX inhibitor without increasing thrombogenicity in hemophilia B dogs and mice. *Blood*. 2015 Mar 5;125(10):1553-61.

[0364] Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D. (1996) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*. Apr 12;272(5259):263-7.

[0365] NATHWANI AC, DAVIDOFF AM, HANAWAH, YUNYU HU, HOFFER FA, NIKANOROV A, SLAUGHTER C, NG CYC, ZHOU J, LOZIER J, MANDRELLTD, VANIN EF, and NIENHUIS AW. (2002). Sustained high-level expression of human factor IX (hFIX) after liver-targeted delivery of recombinant adeno-associated virus encoding the hFIX gene in rhesus macaques. *Blood* 100,1662-1669.

[0366] NATHWANI AC, GRAY JT, NG CY, ZHOU J, SPENCE Y, WADDINGTON SN, TUDDENHAM EG, KEMBALL COOK G, MCINTOSH J, BOON-SPIJKER M, MERTENS K, DAVIDOFF AM. (2006). Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood* 107,2653-61.

[0367] NATHWANI AC, TUDDENHAM EG, RANGARAJAN S, ROSALES C, MCINTOSH J, LINCH DC, CHOWDARY P, RIDDELL A, PIE AJ, HARRINGTON C, O'BEIRNE J, SMITH K, PASI J, GLADER B, RUSTAGI P, NG CY, KAY MA, ZHOU J, SPENCE Y, MORTON CL, ALLAY J, COLEMAN J, SLEEP S, CUNNINGHAM JM, SRIVASTAVA D, BASNER-TSCHAKARJAN E, MINGOZZI F, HIGH KA, GRAY JT, REISS UM, NIENHUIS AW, and DAVIDOFF AM. (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med*. 365,2357-2365.

[0368] NATHWANI AC, REISS UM, TUDDENHAM EG, ET AL. Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N Engl J Med*. 2014;371(21):1994-2004.

[0369] OHLFEST JR, FRANSEN JL, FRITZ S, LOBITZ PD, PERKINSON SG, CLARK KJ, NELSESTUEN G, KEY NS, MCLVOR RS, HACKETT PB, and LARGAESPADA DA. (2004). Phenotypic correction and long-term expression of factor VIII in hemophilic mice by immunotolerization and nonviral gene transfer using the Sleeping Beauty transposon system. *Blood* 105,2691-8.

- [0370] NATHWANI AC, REISS UM, TUDDENHAM EG, ROSALES C, CHOWDARY P, MCINTOSH J, DELLA PERUTA M, LHERITEAU E, PATEL N, RAJ D, RIDDELL A, PIE J, RANGARAJAN S, BEVAN D, RECHT M, SHEN YM, HALKA KG, BASNER-TSCHAKARJAN E, MINGOZZI F, HIGH KA, ALLAY J, KAY MA, NG CY, ZHOU J, CANCIO M, MORTON CL, GRAY JT, SRIVASTAVA D, NIENHUIS AW, DAVIDOFF AM. Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N Engl J Med*. 2014 Nov 20; 371(21):1994-2004.
- [0371] NEGRIER C. (2016) Entering new areas in known fields: recombinant fusion protein linking recombinant factor VIIa with recombinant albumin (rVIIa-FP) - advancing the journey. *Thromb Res*. 141(3):S9-S12.
- [0372] PETERS RT, LOW SC, KAMPHAUS GD, DUMONT JA, ET AL. Prolonged activity of factor IX as a monomeric Fc fusion protein. *Blood*. 2010; 115(10):2057-2064.
- [0373] PETRUS, I., CHUAH, M. & VANDENDRIESSCHE, T. Gene therapy strategies for hemophilia: benefits versus risks. *J Gene Med* 12, 797-809 (2010). SANDBERG H, ALMSTEDT A, BRANDT J, et al. (2001). Structural and functional characteristics of the B domain-deleted recombinant factor VIII protein, r-VIII SQ. *Thromb Haemost*. 85(1):93-100.
- [0374] POWELL JS, PASI KJ, RAGNI MV, ET AL. Phase 3 study of recombinant factor IX Fc fusion protein in hemophilia B. *N Engl J Med*. 2013; 369(24):2313-2323.
- [0375] RANGARAJAN S, WALSH L, LESTER W, ET AL. AAV5-factor VIII gene transfer in severe hemophilia A. *N Engl J Med*. 2017; 377:2519-2530.
- [0376] SANTAGOSTINO E, MARTINOWITZ U, LISSITCHKOV T, ET AL. Long-acting recombinant coagulation factor IX albumin fusion protein (rIX-FP) in hemophilia B: results of a phase 3 trial. *Blood*. 2016; 127(14):1761-1769.
- [0377] SCHULTE (2008) Use of albumin fusion technology to prolong the half-life of recombinant factor VIIa. *Thromb Res*. 122(4):S14-19.
- [0378] Schuettrumpf, J., Herzog, R.W., Schlachterman, A., Kaufhold, A., Stafford, D.W. & Arruda, V.R. (2005). Factor IX variants improve gene therapy efficacy for hemophilia B. *Blood* 105(6):2316-2323.
- [0379] Simioni, P., Tormene, D., Tognin, G., Gavasso, S., Bulato, C., Iacobelli, N.P., Finn, J.D., Spiezia, L., Radu, C. & Arruda, V.R. (2009). X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N Engl J Med* 361(17):1671-1675.
- [0380] SNYDER RO, MIAO CH, PATIJN GA, SPRATT SK, DANOS O, NAGY D, GOWN AM, WINTHER B, MEUSE L, COHEN LK, THOMPSON AR, and KAY MA. (1997). Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors. *Nat Genet*. 16, 270-276.
- [0381] SNYDER RO, MIAO C, MEUSE L, TUBB J, DONAHUE BA, HUI-FENG LIN, STAFFORD DW, PATEL S, THOMPSON AR, NICHOLS T, READ MS, BELLINGER DA, BRINKHOUS KM, and KAY MA. (1999). Correction of hemophilia B in canine and murine models using recombinant adeno-associated viral vectors. *Nat Med*. 5, 64-70.

- [0382] STROHL WR (2015) Fusion proteins for half-life extension fo biologics as a strategy to make biobetters. *Biodrugs*.29 (4) :215-239.
- [0383] Trapnell BC. (1993) Adenoviral vectors for gene transfer. *Adv. Drug Del. Rev.*12:185-199.
- [0384] VANDENBERGHE LH, WANG L, SOMANATHAN S, ZHI Y, FIGUEREDO J, CALCEDO R, SANMIGUEL J, DESAI RA, CHEN CS, JOHNSTON J, GRANT RL, GAO G, and WILSON JM. (2006) .Heparin binding directs activation of T cells against adeno-associated virus serotype 2 capsid. *Nat Med*.12,967-71.
- [0385] VANDENDRIESSCHE T, VANSLEMBROUCK V, GOOVAERTS I, ZWINNEN H, VANDERHAEGHEN ML, COLLEN D, and CHUAH MK. (1999) .Long-term expression of human coagulation factor VIII and correction of hemophilia A after in vivo retroviral gene transfer in factor VIII-deficient mice. *Proc Natl Acad Sci USA*.96,10379-84.
- [0386] VANDENDRIESSCHE T, THORREZ L, NALDINI L, FOLLENZI A, MOONS L, ZWI BERNEMAN, COLLEN D, and CHUAH MK. (2002) .Lentiviral vectors containing the human immunodeficiency virus type-1 central polypurine tract can efficiently transduce nondividing hepatocytes and antigen-presenting cells in vivo. *Blood* 100,813-22.
- [0387] VANDENDRIESSCHE T, THORREZ L, ACOSTA-SANCHEZ A, PETRUS I, WANG L, MA L, DE WAELE L, IWASAKI Y, GILLIJNS V, WILSON JM, COLLEN D, and CHUAH MK. (2007) .Efficacy and safety of adeno-associated viral vectors based on serotype 8 and 9 vs.lentiviral vectors for hemophilia B gene therapy. *J Thromb Haemost*.5,16-24.
- [0388] VANDENDRIESSCHE T, IVICS Z, IZSVÁK Z, and CHUAH MK. (2009) .Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. *Blood* 114,1461-8.
- [0389] VANDENDRIESSCHE T, and CHUAH MK. (2012) .Clinical progress in gene therapy:sustained partial correction of the bleeding disorder in patients suffering from severe hemophilia B. *Hum Gene Ther*.23,4-6.
- [0390] Wang L, Zoppè M, Hackeng TM, Griffin JH, Lee KF, Verma IM. A factor IX-deficient mouse model for hemophilia B gene therapy. *Proc Natl Acad Sci USA*.1997;94 (21) :11563-11566.
- [0391] WANG L, TAKABE K, BIDLINGMAIER SM, ILL CR, and VERMA IM. (1999) .Sustained correction of bleeding disorder in hemophilia B mice by gene therapy. *Proc Natl Acad Sci USA* 96,3906-3910.
- [0392] WANG L, NICHOLS TC, READ MS, BELLINGER DA, and VERMA IM. (2000) .Sustained expression of therapeutic level of factor IX in hemophilia B dogs by AAV-mediated gene therapy in liver. *Mol Ther*.1,154-158.
- [0393] WANG L, CAO O, SWALM B, DOBRZYNSKI E, MINGOZZI F, and HERZOG RW (2005) Major role of local immune responses in antibody formation to factor IX in AAV gene transfer. *Gene Ther* 12,1453-464.

- [0394] WARD NJ, BUCKLEY SM, WADDINGTON SN, VANDENDRIESSCHE T, CHUAH MK, NATHWANI AC, McLINTOSH J, TUDDENHAM EG, KINNON C, THRASHER AJ, and McVEY JH (2010) Codon optimization of human factor VIII cDNAs leads to high-level expression. *Blood* 117, 798-807.
- [0395] WARD, N.J. ET AL. Codon optimization of human factor VIII cDNAs leads to high-level expression. *Blood* 117, 798-807 (2011).
- [0396] WEILLER M, SCHUSTER M, COULIBALY S, ET AL. Biopotency and efficacy of SHP648, a next-generation fix gene therapy vector. *Mol Ther.* 2018; 26(suppl): Abstract 822.
- [0397] WU Z, SUN J, ZHANG T, YIN C, YIN F, VAN DYKE T, SAMULSKI RJ, and MONAHAN PE. (2008). Optimization of self-complementary AAV vectors for liver-directed expression results in sustained correction of hemophilia B at low vector dose. *Mol Ther.* 16, 280-9.
- [0398] XU L, GAO C, and SANDS MS. (2003). Neonatal or hepatocyte growth factor-potentiated adult gene therapy with a retroviral vector results in therapeutic levels of canine factor IX for hemophilia B. *Blood* 101, 3924-3932.
- [0399] XU L, NICHOLS TC, SARKAR R, Mc CORQUODALE S, BELLINGER DA, PONDER KP. (2005). Absence of a desmopressin response after therapeutic expression of factor VIII in hemophilia A dogs with liver-directed neonatal gene therapy. *Proc Natl Acad Sci USA* 102, 6080-6085.
- [0400] Yamada T, Iwasaki Y, Tada H, Iwabuki H, Chuah MK, VandenDriessche T, Fukuda H, Kondo A, Ueda M, Seno M, Tanizawa K, Kuroda S. (2003) Nanoparticles for the delivery of genes and drugs to human hepatocytes. *Nat Biotechnol.* Aug; 21(8): 885-90.
- [0401] YANT SR, MEUSE L, CHIU W, IVICS Z, IZSVAK Z, and KAY MA. (2000). Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat Genet.* 25, 35-41.
- [0402] Yusa et al. A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci U S A.* 2011; 108(4): 1531-6.
- [0403] Zhang G, Budker V, Wolff JA. (1999) High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA. *Hum Gene Ther.* Jul 1; 10(10): 1735-7.
- [0404] ZHONG L, LI B, MAH CS, GOVINDASAMY L, AGBANDJE-MCKENNA, COOPER M, HERZOG RW, ZOLOTUKHIN I, WARRINGTON JR. KH, WEIGEL-VAN AKEN K, HOBBS JA, ZOLOTUKHIN S, MUZYCZKA N, and SRIVASTAVA A (2008). Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci USA* 105, 7827-32.
- [0405] ZOLLNER S, SCHUERMANN D, RAQUET E, MUELLER-COHRIS J, WEIMER T, PRAGST I, DICKNEITE G, SCHULTE S. (2014). Pharmacological characteristics of a novel,

recombinant fusion protein linking coagulation factor VIIa with albumin (rVIIa-FP). *J Thromb Haemost.* 12 (2) :220-228.

序列表

<110> VRIJE布鲁塞尔大学

<120> 改善基因疗法的新工具及其用途

<130> VUB-075-PCT

<150> 18202888.6

<151> 2018-10-26

<160> 35

<170> PatentIn version 3.5

<210> 1

<211> 4989

<212> DNA

<213> 人工序列

<220>

<223> AAV_{SS}-1XSERP-mTTR-MVM-hFIX_{co}-SV40pA

<400> 1

```
cctgcaggca gctgcgcgct cgtcgcctca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc tgcggccgcg gtaccggcgc gccgggggag gctgctggtg 180
aatattaacc aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacacgcg 240
tggtagcgtc tgtctgcaca tttcgtagag cgagtgttcc gatactctaa tctccctagg 300
caaggttcat atttgtgtag gttacttatt ctctttttgt tgactaagtc aataatcaga 360
atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttgaa ggagggggta 420
taaaagcccc ttcaccagga gaagccgtca cacagatcca caagctcctg aagaggtaa 480
ggtttaaggg atggttggtt ggtggggat taatgtttaa ttacctggag cacctgcctg 540
aaatcacttt ttttcaggtt ggctagttct agcccacat gcagcgcgtg aacatgatca 600
tggccgagag ccccgccctg atcaccatct gcctgctggg ctacctgctg agcggcagat 660
gcaccgtggt cctggaccac gagaacgcca acaagatcct gaaccgcccc aagcgtaca 720
acagcggcaa gctggaggag ttcgtgcagg gcaacctgga gcgcgagtc atggaggaga 780
agtgcagctt cgaggaggcc cgcgaggtgt tcgagaacac cgagcgcacc accgagttct 840
ggaagcagta cgtggacggc gaccagtgcg agagcaacc ctgcctgaac ggcggcagct 900
gcaaggacga catcaacagc tacgagtgtt ggtgcccctt cggcttcgag ggcaagaact 960
gcgagctgga cgtgacctgc aacatcaaga acggccgctg cgagcagttc tgcaagaaca 1020
gcgccgacaa caaggtggtg tgcagctgca ccgagggcta ccgctggcc gagaaccaga 1080
agagctgcga gcccgcctg ccttcccc gcggccgcgt gagcgtgagc cagaccagca 1140
agctgacctg cggcggagcc gtgttcccc acgtggacta cgtgaacagc accgagccg 1200
agaccatcct ggacaacatc acccagagca cccagagctt caacgacttc acccgcgtgg 1260
tgggcggcga ggacgccaag cccggccagt tcccctggca ggtggtgctg aacggcaagg 1320
tggacgcctt ctgcggcggc agcatcgtga acgagaagtg gatcgtgacc gccgccact 1380
```

gcgtggagac cggcgtgaag atcacctggt tggccggcga gcacaacatc gaggagaccg 1440
 agcacaccga gcagaagcgc aacgtgatcc gcatcatccc ccaccacaac tacaacgccg 1500
 ccatcaacaa gtacaaccac gacatcgccc tgctggagct ggacgagccc ctggtgctga 1560
 acagctacgt gacccccatc tgcctcgccc acaaggagta caccaacatc ttcctgaagt 1620
 tcggcagcgg ctacgtgagc ggctggggcc gcgtgttcca caagggccgc agcgccttg 1680
 tgctgcagta cctgcgcgtg cccctggtgg accgcgccac ctgcctgcgc agcaccaagt 1740
 tcaccatcta caacaacatg ttctgcgccg gcttccacga gggcggccgc gacagctgcc 1800
 agggcgacag cggcggcccc cacgtgaccg aggtggaggg caccagcttc ctgaccggca 1860
 tcatcagctg gggcgaggag tgcgccatga agggcaagta cggcatctac accaaggtga 1920
 gccgctacgt gaactggatc aaggagaaga ccaagctgac ctaatgaaag atggatttcc 1980
 aaggttaatt cattggaatt gaaaattaac agcccccccc ccccccccc tgcagatctc 2040
 aagcttcgaa ttctgcagtc gacaccggtg gataactgat cggatctagg ctgcacatgc 2100
 tttatttgtg aaatttgtga tgetattgct ttatttgtaa ccattataag ctgcaataaa 2160
 caagttaaca acaacaattg cattcatttt atgtttcagg ttcaggggga ggtgtgggag 2220
 gttttttaa ctcgagatcc actagggccg caggaacccc tagtgatgga gttggccact 2280
 ccctctctgc gcgctcgtc gctcactgag gccgggcgac caaaggtcgc ccgacccccg 2340
 ggctttgccc gggcggcctc agtgagcagc cgagcgcgca gctgcctgca ggggcgcctg 2400
 atgcggtatt ttctccttac gcatctgtgc ggtatttcac accgcatacg tcaaagcaac 2460
 catagtacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 2520
 tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 2580
 tcgccacgtt cgccggcttt cccgtcaag ctctaaatcg ggggctccct ttagggttcc 2640
 gatttagtgc ttacggcac ctgcaccca aaaaacttga tttgggtgat ggttcacgta 2700
 gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 2760
 atagtggact cttgttcaa actggaacaa cactcaacce tatctcgggc tattcttttg 2820
 atttataagg gattttgccg atttcggcct attggttaaa aatgagctg atttaacaaa 2880
 aatttaacgc gaattttaac aaaatattaa cgtttacaat tttatggtgc actctcagta 2940
 caatctgctc tgatgccgca tagttaagcc agccccgaca cccgccaaca cccgctgacg 3000
 cgccctgacg ggcttgtctg ctcccggcat ccgcttacag acaagctgtg accgtctccg 3060
 ggagctgcat gtgtcagagg ttttaccgt catcaccgaa acgcgcgaga cgaaagggcc 3120
 tcgtgatacg cctattttta taggttaatg tcatgataat aatggtttct tagacgtcag 3180
 gtggcacttt tcggggaat gtgcgcggaa cccctatttg tttatttttc taaatacatt 3240
 caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa 3300
 ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 3360
 gccttctgt ttttgtcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 3420
 tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaaagatc cttgagagtt 3480
 ttgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 3540
 tattatcccg tattgacgcc gggcaagagc aactcggteg ccgcatacac tattctcaga 3600
 atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 3660
 gagaattatg cagtgtgcc ataaccatga gtgataaac tgcggccaac ttacttctga 3720

caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 3780
 ctgcgcttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 3840
 ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 3900
 ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 3960
 ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 4020
 gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 4080
 ttatctacac gacggggagt caggcaacta tggatgaacg aatagacag atcgtgaga 4140
 taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 4200
 agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 4260
 atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 4320
 aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa 4380
 caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 4440
 ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc 4500
 cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 4560
 tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttgactcaa 4620
 gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc 4680
 ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 4740
 gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa 4800
 caggagagcg cacgaggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 4860
 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc 4920
 tatggaaaaa cgccagcaac gcggcctttt tacggttctt ggccttttgc tggccttttg 4980
 ctcacatgt 4989

<210> 2

<211> 6728

<212> DNA

<213> 人工序列

<220>

<223> AAV_{ss}-1XSERP-mTTR-MVM-hFIX_{co}-Alb_{co}-SV40pA

<400> 2

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgct 60
 gggcgacctt tggtcgcccg gcctcagtga gcgagcagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgcg gtaccggcgc gccgggggag gctgctggtg 180
 aatattaacc aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacacgcg 240
 tggtagcgtc tgtctgcaca tttcgtagag cgagtgttcc gatactctaa tctccctagg 300
 caaggttcat atttgtgtag gttacttatt ctcttttgt tgactaagtc aataatcaga 360
 atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttgga ggagggggta 420
 taaaagcccc ttcaccagga gaagccgtca cacagatcca caagctcctg aagaggtaa 480
 ggtttaaggg atggttggtt ggtggggtat taatgtttaa ttacctggag cacctgcctg 540
 aaatcacttt ttttcagggt ggctagttct agaagatctc aagcttgcta gcatgcagcg 600

cgtgaacatg atcatggccg agagccccgg cctgatcacc atctgcctgc tgggctacct 660
 gctgagcgcc gaggtcaccg tgttcctgga ccacgagaac gccaacaaga tcctgaaccg 720
 ccccaagcgc tacaacagcg gcaagctgga ggagttcgtg cagggcaacc tggagcgcga 780
 gtgcatggag gagaagtgca gcttcgagga ggcccgcgag gtgttcgaga acaccgagcg 840
 caccaccgag ttctggaagc agtacgtgga cggcgaccag tgcgagagca acccctgcct 900
 gaacggcggc agctgcaagg acgacatcaa cagctacgag tgctggtgcc ctttcggctt 960
 cgagggcaag aactgcgagc tggacgtgac ctgcaacatc aagaacggcc gctgagcgca 1020
 gttctgcaag aacagcgccg acaacaaggt ggtgtgcagc tgcaccgagg gctaccgcct 1080
 ggccgagaac cagaagagct gcgagccccg cgtgcccttc ccctgcggcc gcgtgagcgt 1140
 gagccagacc agcaagctga cccgcgccga ggccgtgttc cccgacgtgg actacgtgaa 1200
 cagcaccgag gccgagacca tcttgacaa catcaccag agcaccaga gcttcaacga 1260
 cttcaccgcg gtggtgggcg gcgaggacgc caagcccggc cagtccctt ggcaggtggt 1320
 gctgaacggc aaggtggacg cttctgctgg cggcagcatc gtgaacgaga agtggatcgt 1380
 gaccgccgcc cactgctgag agaccggcgt gaagatcacc gtggtggccg gcgagcaca 1440
 catcgaggag accgagcaca ccgagcagaa gcgcaacgtg atccgcatca tccccacca 1500
 caactacaac gccgccatca acaagtacaa ccacgacatc gccctgctgg agctggacga 1560
 gccctggtg ctgaacagct acgtgacccc catctgcatc gccgacaagg agtacaccaa 1620
 catcttctg aagttcggca gcggctacgt gagcggctgg ggccgcgtgt tccacaagg 1680
 ccgcagcgcc ctggtgctgc agtacctgcg cgtgccctg gtggaccgcg ccacctgcct 1740
 gcgcagcacc aagttcacca tctacaacaa catgttctgc gccggcttcc acgagggcgg 1800
 ccgcgacagc tgccaggcg acagcggcg cccccacgtg accgaggtgg agggcaccag 1860
 cttcctgacc ggcatcatca gctggggcga ggagtgcgcc atgaaggca agtacggcat 1920
 ctacaccaag gtgagccgct acgtgaactg gatcaaggag aagaccaagc tgacctctgt 1980
 gagccagacc tccaagctca ccagggccga gactgtcttc cctgatgtgg acggatccga 2040
 cgcccacaag agcagagtg gccacagatt caaggacctg ggcgaggaaa acttcaaggc 2100
 tctggtgctg atcgcttctg cccagctact gcagcagtg cccttcgagg accacgtgaa 2160
 gctggtcaac gaagtgaccg agttcgccaa gacctgctg gccgacgaga gcgccgagaa 2220
 ctgagacaag agcctgcaca ccctgttcgg cgacaagctg tgcaccgtgg ccaccctgcg 2280
 ggaaacctac ggcgagatgg ccgactgctg cgccaagcag gaaccgagc ggaacgagtg 2340
 cttcctgcag cacaaggacg ataaccceaa cctgccccgg ctctgctggc ccgaggtgga 2400
 cgtgatgtgc accgccttcc acgacaacga ggaaacctc ctgaagaagt acctgtacga 2460
 gatcgccaga cggcaccctt acttctacgc ccccgagctg ctgttcttcg ccaagcggta 2520
 caaggccgcc ttcaccgagt gctgccagc gcgccataag gccgcctgcc tgctgcccga 2580
 gctggatgag ctgagggacg agggcaaggc cagctccgcc aagcagagac tgaagtgcgc 2640
 cagcctgcag aagttcggcg agagagcctt taaggcctgg gctgtggccc ggctgagcca 2700
 gagattcccc aaggccgagt ttgccgaggt gtccaagctg gtcaccgacc tcaccaaggt 2760
 gcacaccgag tgtgtcacg gcgacctgct ggaatgcgcc gacgacagag ccgacctggc 2820
 caagtacatc tgcgagaacc aggacagcat cagctccaag ctgaaagagt gctgcgagaa 2880
 gccctgctg gaaaagagcc actgtatcgc cgaggtggaa aacgacgaga tgcccgccga 2940

cctgcccagc ctggccgccc acttcgtgga aagcaaggac gtgtgcaaga actacgccga 3000
ggccaaggat gtgttcctgg gcatgttcct gtatgagtac gcccgcagac accccgacta 3060
cagcgtgggtg ctgctgctgc ggctggccaa gacctacgag acaaccctgg aaaagtgtctg 3120
cgccgctgcc gacccccacg agtgctacgc caaggtgttc gacgagttca agccactggt 3180
ggaagaacct cagaacctga tcaagcagaa ttgagagctg ttcgagcagc tgggcgagta 3240
caagttccag aacgccctgc tcgtgcggta caccaagaaa gtgccccagg tgtccacccc 3300
caccctgggtg gaagtgctcc ggaacctggg caaagtgggc agcaagtgtc gcaagcacc 3360
tgaggccaag cggatgccct gcgccgagga ctacctgagc gtggtcctga accagctgtg 3420
cgtgctgcac gagaaaacct ccgtgtccga cagagtgtacc aagtgtgtga ccgagagcct 3480
ggtcaacaga cggccctgct tctccgccct ggaagtggac gagacatacg tgcccaaaga 3540
gttcaacgcc gagacattca cttccacgc cgacatctgc acctgagcg agaaagagcg 3600
gcagatcaag aagcagaccg cctgggtcga gctgggtcaag cacaagccca aggccacca 3660
agaacagctg aaggccgtga tggacgactt cgccgccttc gtcgagaagt gttgcaaggc 3720
cgacgacaaa gagacatgct tcgccgaaga gggcaaaaag ctggtggccg cctctcaggc 3780
cgccctggga ctctaagtgc acaccggtag ataactgatc ggatctaggc tcgacatgct 3840
ttatttgtga aatttgtgat gctattgctt tatttghtaac cattataagc tgcaataaac 3900
aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag gtgtgggagg 3960
ttttttaaac tcgagatcca ctagggccgc aggaaccctt agtgatggag ttggccactc 4020
cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg 4080
gctttgcccg ggccgcctca gtgagcgagc gagcgcgcag ctgcctgcag gggcgcctga 4140
tgccgtattt tctccttacg catctgtgcg gtatttcaca ccgcatacgt caaagcaacc 4200
atagtacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt 4260
gaccgctaca cttgccagcg ccctagcgcc cgtctcttcc gctttcttcc cttcctttct 4320
cgccacgctt gccggctttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg 4380
atntagtgct ttacggcacc tcgaccccaa aaaacttgat ttgggtgatg gttcacgtag 4440
tgggccatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttaa 4500
tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcgggct attcttttga 4560
tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga ttttaacaaa 4620
atttaacgcg aattttaaca aaatattaac gtttacaatt ttatggtgca ctctcagtac 4680
aatctgctct gatgccgat agttaagcca gcccgcacac ccgccaacac ccgctgacgc 4740
gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg 4800
gagctgcatg tgcagaggt tttaccgctc ataccgaaa cgcgcgagac gaaagggcct 4860
cgtgatacgc ctatttttat aggttaatgt catgataata atggtttctt agacgtcagg 4920
tggcactttt cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc 4980
aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag 5040
gaagagtatg agtattcaac atttccgtgt cgccttatt cccttttttg cggcattttg 5100
ccttctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt 5160
gggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt 5220
tcgccccgaa gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt 5280

attatcccgt attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa 5340
 tgacttgggtt gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag 5400
 agaattatgc agtgctgcc aaccatgag tgataacact gcggccaact tacttctgac 5460
 aacgatcggg ggaccgaagg agctaaccgc ttttttgac aacatggggg atcatgtaac 5520
 tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaagcagc agcgtgacac 5580
 cacgatgcct gtagcaatgg caacaacggt gcgcaacta ttaactggcg aactacttac 5640
 tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact 5700
 tctgcgctcg gcccttccgg ctggctgggt tattgctgat aatctggag ccggtgagcg 5760
 tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt 5820
 tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat 5880
 aggtgcctca ctgattaagc attggtaact gtcagacca gtttactcat atatacttta 5940
 gattgattta aaacttcatt ttaatttaa aaggatctag gtgaagatcc tttttgataa 6000
 tctcatgacc aaaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga 6060
 aaagatcaaa ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac 6120
 aaaaaaacca ccgctaccag cggtggtttg tttgccgat caagagctac caactctttt 6180
 tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc tagttagacc 6240
 gtagttaggc caccacttca agaactctgt agcaccgctt acatacctcg ctctgctaata 6300
 cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag 6360
 acgatagtta ccgataaagg cgcagcgtc gggctgaac gggggttcgt gcacacagcc 6420
 cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag 6480
 cgccacgctt cccgaaggga gaaaggcggg caggtatccg gtaagcggca gggctcggaa 6540
 aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata gtctgtcgg 6600
 gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct 6660
 atggaaaaac gccagcaacg cggccttttt acggttctg gccttttgct ggccttttgc 6720
 tcacatgt 6728

<210> 3

<211> 6851

<212> DNA

<213> 人工序列

<220>

<223> AAV_{SS}-3XSERP-mTTR-MVM-hFIX_{co}-Alb_{co}-SV40pA

<400> 3

cctgcaggca gctgcgcgct cgtcgtctca ctgaggccgc ccgggcaaag cccgggcgct 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgcg gtagggggag gctgctggtg aatattaacc 180
 aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacggggg aggctgctgg 240
 tgaatattaa ccaaggtcac cccagttatc ggaggagcaa acaggggcta agtccacggg 300
 ggaggctgct ggtgaatatt aaccaaggtc accccagtta tcggaggagc aaacaggggc 360
 taagtccacc gtctgtctgc acatttcgta gagcagtggt tccgatactc taatctcct 420

aggcaagggtt catatttgtg taggttactt attctccttt tgttgactaa gtcaataatc 480
 agaatcagca ggtttggagt cagcttggca gggatcagca gcctggggtg gaaggagggg 540
 gtataaaagc cccttcacca ggagaagccg tcacacagat ccacaagctc ctgaagaggt 600
 aagggtttaa gggatggttg gttggtgggg tattaatgtt taattacctg gagcacctgc 660
 ctgaaatcac tttttttcag gttggctagt tctagaagat ctcaagcttg ctagcatgca 720
 gcgcgtgaac atgatcatgg ccgagagccc cggcctgac accatctgcc tgctgggcta 780
 cctgctgagc gccgagtga ccgtgttctt ggaccacgag aacgccaaca agatcctgaa 840
 ccgccccaaag cgctacaaca gcggcaagct ggaggagtcc gtgcagggca acctggagcg 900
 cgagtgcattg gaggagaagt gcagcttcca ggaggcccg gaggtgttcg agaacaccga 960
 gcgcaccacc gatttctgga agcagctacgt ggacggcgac cagtgcgaga gcaacccttg 1020
 cctgaacggc gccagctgca aggacgacat caacagctac gattgctggt gcccttctcg 1080
 cttcgagggc aagaactgcg agctggacgt gacctgcaac atcaagaacg gccgctgca 1140
 gcgattctgc aagaacagcg ccgacaaca ggtggtgtgc agctgcaccg agggctaccg 1200
 cctggccgag aaccagaaga gctgcgagcc cgccgtgcc ttcccctgcg gccgcgtgag 1260
 cgtgagccag accagcaagc tgaccgcgc cgaggccgtg ttcccgcagc tggactacct 1320
 gaacagcacc gaggccgaga ccatcctgga caacatcacc cagagcacc agagcttcaa 1380
 cgacttcacc cgctggttg gcggcgagga cgccaagccc ggccagttcc cctggcaggt 1440
 ggtgctgaac ggcaagggtg acgccttctg cggcggcagc atcgtgaacg agaagtggat 1500
 cgtgaccgcc gccactgcg tggagaccgg cgtgaagat accgtggtgg ccggcgagca 1560
 caacatcgag gagaccgagc acaccgagca gaagcgcaac gtgatccgca tcatcccca 1620
 ccacaactac aacgccgcca tcaacaagta caaccagac atcgccctgc tggagctgga 1680
 cgagcccctg gtgctgaaca gctacgtgac cccatctgc atcgccgaca aggagtacac 1740
 caacatcttc ctgaagtctg gcagcggtta cgtgagcggc tggggcccgc tgttccaca 1800
 gggccgcagc gccctggtgc tgcactacct gcgcgtgcc ctggtggacc gcgccacctg 1860
 cctgcgcagc accaagtcca ccatctaca caacatgtt tgcgccggt tccacgaggg 1920
 cggccgcgac agctgccagg gcgacagcgg cggccccac gtgaccgagg tggagggcac 1980
 cagcttcttg accggcatca tcagctgggg cgaggagtgc gccatgaagg gcaagtacgg 2040
 catctacacc aagtgagcc gctacgtgaa ctggatcaag gagaagacca agctgacctc 2100
 tgtgagccag acctccaagc tcaccagggc cgagactgtc ttccctgatg tggacggatc 2160
 cgacgccac aagagcgagg tggcccacag attcaaggac ctgggcgagg aaaacttcaa 2220
 ggctctggtg ctgacgcct tcgcccagta cctgcagcag tgtcccttcg aggaccacct 2280
 gaagctggtc aacgaagtga ccgagttcgc caagacctgc gtggccgacg agagcgccga 2340
 gaactgcgac aagagcctgc acacctgtt cggcgacaag ctgtgcaccg tggccacct 2400
 gcgggaaacc tacggcgaga tggccgactg ctgcgccaag caggaaccg agcggaacga 2460
 gtgcttctg cagcacaagg acgataacc caacctgcc cggctcgtgc ggcccaggt 2520
 ggagctgatg tgaccgcct tccacgaca cgaggaaacc ttctgaaga agtacctgta 2580
 cgagatgcc agacggcacc cctactteta cggccccag ctgctgttct tcgccaagcg 2640
 gtacaaggcc gccttcacc agtgctgcca ggccgccgat aaggccgct gcctgctgcc 2700
 caagctggat gagctgaggg acgagggcaa ggccagctcc gccaaagcaga gactgaagtg 2760

cgccagcctg cagaagttcg gcgagagagc ctttaaggcc tgggctgtgg cccggctgag 2820
ccagagattc cccaaggccg agtttgccga ggtgtccaag ctggtcaccg acctcacaa 2880
ggtgcacacc gagtgttgtc acggcgacct gctggaatgc gccgacgaca gagccgacct 2940
ggccaagtac atctgcgaga accaggacag catcagctcc aagctgaaag agtgctgcca 3000
gaagcccctg ctggaaaaga gccactgtat cgccgaggtg gaaaacgacg agatgcccgc 3060
cgacctgccc agcctggccg ccgacttctg ggaaagcaag gacgtgtgca agaactacgc 3120
cgaggccaag gatgtgttcc tgggcatgtt cctgtatgag tacgcccgca gacaccccga 3180
ctacagcgtg gtgctgctgc tgcggctggc caagacctac gagacaacc tggaaaagt 3240
ctgcgccgct gccgaccccc acgagtgtca cgccaaggtg ttcgacgagt tcaagccact 3300
ggtggaagaa ccccagaacc tgatcaagca gaattgcgag ctgttcgagc agctgggcca 3360
gtacaagttc cagaacgccc tgetctgtgc gtacaccaag aaagtgcccc aggtgtccac 3420
ccccaccctg gtggaagtgt cccggaacct gggcaaagt ggcagcaagt gctgcaagca 3480
ccctgaggcc aagcggatgc cctgcgccga ggactacct agcgtggtcc tgaaccagct 3540
gtgcgtgctg cacgagaaaa cccccgtgtc cgacagagt accaagtgt gtaccgagag 3600
cctggtcaac agacggccct gttctccgc cctggaagt gacgagacat acgtgcccga 3660
agagttcaac gccgagacat tcacctcca cgccgacatc tgcaccctga gcgagaaaga 3720
gcggcagatc aagaagcaga ccgccctggt cgagctggtc aagcacaagc ccaaggccac 3780
caaagaacag ctgaaggccg tgatggacga cttcgccgc ttcgtcgaga agtgttgcaa 3840
ggccgacgac aaagagacat gttcgccga agaggcaaa aagctggtg ccgcctctca 3900
ggccgccctg ggactctaag tcgacaccgg tagataactg atcggatcta ggctcgacat 3960
gctttatttg taaaatttgt gatgctattg ctttatttgt aaccattata agctgcaata 4020
aacaagttaa caacaacaat tgcatcatt ttatgtttca ggttcagggg gaggtgtggg 4080
aggtttttta aactcgagat ccactagggc cgcaggaacc cctagtgatg gagttggcca 4140
ctccctctct gcgcgctcgc tcgctcactg aggccggcg accaaaggtc gcccgacgcc 4200
cgggctttgc cgggcgccg tcagtgagcg agcgagcgcg cagctgcctg cagggcgcc 4260
tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgata cgtcaaagca 4320
accatagtac gcgccctgta gcggcgcat aagcgcggcg ggtgtggtg ttacgcgcag 4380
cgtgaccgct aacttgcca gcgccctagc gccgctcct ttcgctttct tcccttctt 4440
tctcgccacg ttcgccggtt ttccccgta agctctaaat cgggggctcc ctttagggtt 4500
ccgatttagt gctttacgac actcgacc caaaaaactt gatttgggtg atggttcacg 4560
tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttgagat ccacgttctt 4620
taatagtgga ctcttgttcc aaactggaac aactcaac cctatctcgg gctattcttt 4680
tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca 4740
aaaatttaac gcgaatttta acaaaatatt aacgtttaca attttatggt gactctcag 4800
tacaatctgc tctgatgccg catagttaag ccagccccga ccccccaa caccgctga 4860
cgcgccctga cgggcttgtc tgetccccgc atccgcttac agacaagctg tgaccgtctc 4920
cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgca gacgaaagg 4980
cctcgtgata cgctatttt tataggttaa tgtcatgata ataatggtt cttagacgtc 5040
agggtggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatttt tctaaataca 5100

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 5160
 aaggaagagt atgagtattc aacatttccg tgtegccectt attccctttt ttgcggcatt 5220
 ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 5280
 gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 5340
 ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 5400
 ggtattatcc cgtattgacg ccgggcaaga gcaactcggg cgccgcatac actattctca 5460
 gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 5520
 aagagaatta tgcagtgctg ccataacatc gatgataac actgcggcca acttacttct 5580
 gacaacgacg ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 5640
 aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 5700
 caccacgatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg gcgaactact 5760
 tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 5820
 acttctgcgc tcggcccttc cggetggetg gtttattgct gataaatctg gagccggtga 5880
 gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 5940
 agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga 6000
 gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact catatatact 6060
 ttagattgat ttaaaaactc atttttaatt taaaaggatc taggtgaaga tcctttttga 6120
 taatctcatg accaaaatcc cttaacgtga gttttcgctc cactgagcgt cagaccccg 6180
 agaaaagatc aaaggatcct cttgagatcc ttttttctg cgcgtaatct gctgcttgca 6240
 aacaaaaaaaa ccaccgctac cagcgggtgt ttgtttgccg gatcaagagc taccaactct 6300
 ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgt 6360
 gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 6420
 aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc 6480
 aagacgatag ttaccggata aggcgagcgc gtcgggctga acggggggtt cgtgcacaca 6540
 gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 6600
 aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggttaagc gcagggtcgg 6660
 aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatcttt atagtcctgt 6720
 cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 6780
 cctatggaaa aacgccagca acgcgccctt tttacggttc ctggcctttt gctggccttt 6840
 tgctcacatg t 6851
 <210> 4
 <211> 4989
 <212> DNA
 <213> 人工序列
 <220>
 <223> AAV_{ss}-1XSERP-mTTR-MVM-hFIX_{co}Padua-SV40pA
 <400> 4
 cctgcaggca gctgcgcgct cgctcgtcct ctgaggccgc ccgggcaaag cccgggcgctc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgcg gtaccggcgc gccgggggag gctgctggtg 180
 aatattaacc aaggtcaccc cagttatcgg aggagcaaac aggggctaag tccacacgcg 240
 tggtagcgtc tgtctgcaca tttcgtagag cgagtgttcc gataactctaa tctccctagg 300
 caaggttcat atttgtgtag gttacttatt ctcttttgtg tgactaagtc aataatcaga 360
 atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttgga gagggggta 420
 taaaagcccc ttcaccagga gaagccgtca cacagatcca caagctcctg aagaggtaag 480
 ggtttaaggg atggttggtt ggtggggtat taatgtttaa ttacctggag cacctgcctg 540
 aaatcacttt ttttcaggtt ggctagttct agcccacat gcagcgcgtg aacatgatca 600
 tggccgagag ccccgccctg atcaccatct gctgctggg ctacctgctg agcggcagct 660
 gcaccgtgtt cctggaccac gagaacgcca acaagatcct gaaccgcccc aagcgtaca 720
 acagcggcaa gctggaggag ttcgtgcagg gcaacctgga gcgcgagtgc atggaggaga 780
 agtgcagctt cgaggaggcc cgcgaggtgt tcgagaacac cgagcgcacc accgagttct 840
 ggaagcagta cgtggacggc gaccagtgcg agagcaaccc ctgcctgaac ggcggcagct 900
 gcaaggacga catcaacagc tacgagtgtt ggtgcccctt cggcttcgag ggcaagaact 960
 gcgagctgga cgtgacctgc aacatcaaga acggccgctg cgagcagttc tgcaagaaca 1020
 gcgccgacaa caaggtggtg tgcagctgca ccgagggcta ccgctggcc gagaaccaga 1080
 agagctgcga gcccgccgtg cccttcccct gcggccgctg gagcgtgagc cagaccagca 1140
 agctgaccgg cgccgaggcc gtgttccccg acgtggacta cgtgaacagc accgaggccg 1200
 agaccatcct ggacaacatc accagagca cccagagctt caacgacttc acccgcgtgg 1260
 tgggcccgca ggacgccaag cccggccagt tcccctggca ggtggtgctg aacggcaagg 1320
 tggacgcctt ctgcggcggc agcatcgtga acgagaagtg gatcgtgacc gccgcccact 1380
 gcgtggagac cggcgtgaag atcacctggg tggcccgca gcacaacatc gaggagaccg 1440
 agcacaccga gcagaagcgc aacgtgatcc gcatcatccc ccaccacaac tacaacgccg 1500
 ccatcaacaa gtacaaccac gacatcgccc tgctggagct ggacgagccc ctggtgctga 1560
 acagctacgt gacccccatc tgcacgcccg acaaggagta caccaacatc ttcctgaagt 1620
 tcggcagcgg ctacgtgagc ggctggggcc gcgtgttcca caaggccgc agcggcctgg 1680
 tgctgcagta cctgcgcgtg cccctggtgg accgcgccac ctgcctgctg agcaccaagt 1740
 tcaccatcta caacaacatg ttctgcgccc gcttccacga gggcggccgc gacagctgcc 1800
 agggcgacag cggcggcccc cacgtgaccg aggtggaggg caccagcttc ctgaccggca 1860
 tcacagctg gggcgaggag tgcgcatga agggcaagta cggcatctac accaaggtga 1920
 gccgctacgt gaactggatc aaggagaaga ccaagctgac ctaatgaaag atggatttcc 1980
 aaggttaatt cattggaatt gaaaattaac agcccccccc ccccccccc tgcagatctc 2040
 aagcttcgaa ttctgcagtc gacaccgta gataactgat cggatctagg ctgcacatgc 2100
 tttatgtgtg aaatttgtga tgetattgtt ttatgtgtaa ccattataag ctgcaataaa 2160
 caagttaaca acaacaattg cattcatttt atgtttcagg ttcaggggga ggtgtgggag 2220
 gttttttaa ctcgagatcc actagggccg caggaaaccc tagtgatgga gttggccact 2280
 ccctctctgc gcgctcgtc gctcactgag gccgggcgac caaaggtcgc ccgacgcccg 2340
 ggctttgccc gggcggcctc agtgagcagc cgagcgcgca gctgcctgca ggggcgctg 2400
 atgcggtatt ttctccttac gcatctgtgc ggtatttcc accgcatacg tcaaagcaac 2460

catagtacgc gccctgtagc ggcgcattaa gcgcggcggg tgtgggtggtt acgcgcagcg 2520
tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ctttcctttc 2580
tcgccacggt cgccggcttt ccccgtaag ctctaaatcg ggggctccct ttagggttcc 2640
gatttagtgc tttacggcac ctgcaccca aaaaacttga tttgggtgat gtttcacgta 2700
gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 2760
atagtggact ctgtttccaa actggaacaa cactcaacc tatctcgggc tattcttttg 2820
atttataagg gattttgccg atttcggcct attggttaaa aatgagctg atttaacaaa 2880
aatttaacgc gaattttaac aaaatattaa cgtttacaat tttatgggtgc actctcagta 2940
caatctgctc tgatgccgca tagttaagcc agccccgaca cccgccaaca cccgctgacg 3000
cgccctgacg ggcttgtctg ctcccggcat ccgcttacag acaagctgtg accgtctccg 3060
ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc 3120
tcgtgatacg cctattttta taggttaatg tcatgataat aatggtttct tagacgtcag 3180
gtggcacttt tcggggaaat gtgcgcggaa ccctatttg tttatttttc taaatacatt 3240
caaatatgta tccgctcatg agacaataac cctgataaat gttcaataa tattgaaaa 3300
ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 3360
gccttcctgt ttttgtcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 3420
tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaaagatc cttgagagtt 3480
ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg 3540
tattatcccg tattgacgcc gggcaagagc aactcggctc ccgcatacac tattctcaga 3600
atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc atgacagtaa 3660
gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac ttacttctga 3720
caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg gatcatgtaa 3780
ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 3840
ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc gaactactta 3900
ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt gcaggaccac 3960
ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga gccggtgagc 4020
gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 4080
ttatctacac gacggggagt caggcaacta tggatgaacg aatagacag atcgtgaga 4140
taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca tatatacttt 4200
agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc ctttttgata 4260
atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca gaccccgtag 4320
aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa 4380
caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt 4440
ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc 4500
cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc gctctgctaa 4560
tctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa 4620
gacgatagtt accggataag gcgcagcggc cgggctgaac ggggggttcg tgcacacagc 4680
ccagcttggg gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa 4740
gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggctcgaa 4800

caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg 4860
 ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctctgcaggg gggcggagcc 4920
 tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg 4980
 ctcacatgt 4989
 <210> 5
 <211> 6728
 <212> DNA
 <213> 人工序列
 <220>
 <223> AAV_{SS}-1XSERP-mTTR-MVM-hFIX_{co}Padua-Alb-SV40pA
 <400> 5
 cctgcaggca gctgcgcgct cgtcctctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgcg gtaccggcgc gccgggggag gctgctggtg 180
 aatattaacc aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacacgcg 240
 tggtagcgtc tgtctgcaca tttcgtagag cgagtgttcc gatactctaa tctccctagg 300
 caaggttcat atttgtgtag gttacttatt ctctttttgt tgactaagtc aataatcaga 360
 atcagcaggt ttggagtcag cttggcaggg atcagcagcc tgggttgga ggagggggta 420
 taaaagcccc ttcaccagga gaagccgtca cacagatcca caagctcctg aagaggtaa 480
 ggtttaaggg atggttggtt ggtggggtat taatgtttaa ttacctggag cacctgcctg 540
 aaatcacttt ttttcaggtt ggctagttct agaagatctc aagcttgcta gcatgcagcg 600
 cgtgaacatg atcatggccg agagccccgg cctgatcacc atctgcctgc tgggctacct 660
 gctgagcgcc gactgcaccg tgttcttgga ccacgagaac gccaacaaga tctgaaccg 720
 cccaagcgc tacaacagcg gcaagctgga ggagtctgtg cagggaacc tggagcgcga 780
 gtgcatggag gagaagtga gcttcgagga ggcccgcgag gtgttcgaga acaccgagcg 840
 caccaccgag ttctggaagc agtacgtgga cggcgaccag tgcgagagca acccctgcct 900
 gaacggcggc agctgcaagg acgacatcaa cagctacgag tgctggtgcc cttcggcctt 960
 cgagggcaag aactgcgagc tggacgtgac ctgcaacatc aagaacggcc gctgcgagca 1020
 gttctgcaag aacagcgcgg acaacaaggt ggtgtgcagc tgcaccgagg gctaccgcct 1080
 ggccgagaac cagaagagct gcgagcccgc cgtgcccttc ccctgcggcc gcgtgagcgt 1140
 gagccagacc agcaagctga cccgcgccga ggccgtgttc cccgacgtgg actacgtgaa 1200
 cagcaccgag gccgagacca tcttgacaa catcaccag agcaccaga gcttcaacga 1260
 cttcaccgcg gtggtgggag gcgaggacgc caagcccggc cagttcccct ggcaggtggt 1320
 gctgaacggc aaggtggacg cttctgaggc cggcagcatc gtgaacgaga agtggtatcgt 1380
 gaccgccgcc cactgcgtgg agaccggcgt gaagatcacc gtggtggccg gcgagcaca 1440
 catcgaggag accgagcaca ccgagcagaa gcgcaacgtg atccgcatca tccccacca 1500
 caactacaac gccgcatca acaagtaca ccacgacatc gccctgctgg agctggacga 1560
 gccctgggtg ctgaacagct acgtgacccc catctgcatc gccgacaagg agtacacaa 1620
 catcttctg aagttcggca gcggctacgt gagcggctgg ggccgcgtgt tccacaagg 1680

ccgcagcgcc ctggtgctgc agtacctgcg cgtgcccctg gtggaccgcg ccacctgcct 1740
gctgagcacc aagttcacca tctacaacia catgttctgc gccggcttcc acgagggcgg 1800
ccgcgacagc tgccagggcg acagcggcgg cccccacgtg accgaggtgg agggcaccag 1860
cttcctgacc ggcatcatca gctggggcga ggagtgcgcc atgaaggca agtacggcat 1920
ctacaccaag gtgagccgct acgtgaactg gatcaaggag aagaccaagc tgacctctgt 1980
gagccagacc tccaagctca ccagggccga gactgtcttc cctgatgtgg acggatccga 2040
tgcacacaag agtgaggttg ctcatcggtt taaagatttg ggagaagaaa atttcaaagc 2100
cttgggtgtg attgcctttg ctcaagtatc tcagcagtgt ccatttgaag atcatgtaa 2160
attagtgaat gaagtaactg aatttgcaaa aacatgtgtt gctgatgagt cagctgaaaa 2220
ttgtgacaaa tcaactcata ccttttttgg agacaaatta tgcacagttg caactcttcg 2280
tgaaacctat ggtgaaatgg ctgactgctg tgcaaaacia gaacctgaga gaaatgaatg 2340
cttcttgcaa cacaaagatg acaacccaaa cctccccga ttggtgagac cagaggttga 2400
tgtgatgtgc actgcttttc atgacaatga agagacattt ttgaaaaat acttatatga 2460
aattgccaga agacatcctt acttttatgc cccggaactc cttttctttg ctaaaaggta 2520
taaagctgct ttacagaaat gttgccaagc tgctgataaa gctgcctgcc tgttgccaaa 2580
gctcgatgaa cttcgggatg aagggaaggc ttcgtctgcc aacagagac tcaagtgtgc 2640
cagtctcaa aaatttgag aaagagcttt caaagcatgg gcagtagctc gcctgagcca 2700
gagatttccc aaagctgagt ttgcagaagt ttccaagtta gtgacagatc ttaccaaagt 2760
ccacacggaa tgctgccatg gagatctgct tgaatgtgct gatgacaggc cggaccttgc 2820
caagtatatc tgtgaaaatc aagattcgat ctccagtaaa ctgaaggaaat gctgtgaaaa 2880
acctctgttg gaaaaatccc actgcattgc cgaagtggaa aatgatgaga tgcctgctga 2940
cttgccttca ttagctgctg attttgttga aagtaaggat gtttgcaaaa actatgctga 3000
ggcaaaggat gtcttcttgg gcatgttttt gtatgaatat gcaagaaggc atcctgatta 3060
ctctgtcgtg ctgctgctga gacttgccaa gacatatgaa accactctag agaagtgctg 3120
tgccgctgca gatcctcatg aatgctatgc caaagtgttc gatgaattta aacctcttgt 3180
ggaagagcct cagaatttaa tcaacaaaa ttgtgagctt ttgagcagc ttggagagta 3240
caaattccag aatgcgctat tagttcgta caccaagaaa gtacccaag tgtcaactcc 3300
aactcttgta gaggtctcaa gaaacctagg aaaagtgggc agcaaatgtt gtaaacatcc 3360
tgaagcaaaa agaatgccct gtgcagaaga ctatctatcc gtggtcctga accagttatg 3420
tgtgttgcag gagaaaacgc cagtaagtga cagagtcacc aatgctgca cagaatcctt 3480
ggtgaacagg cgacctgct tttcagctct ggaagtcgat gaaacatacg ttcccaaaga 3540
gtttaatgct gaaacattca cttccatgc agatatatgc aactttctg agaaggagag 3600
acaaatcaag aaacaaactg cacttgttga gctcgtgaaa cacaagccca aggcaacaaa 3660
agagcaactg aaagctgtta tggatgattt cgcagctttt gtagagaagt gctgcaaggc 3720
tgacgataag gagacctgct ttgccgagga gggtaaaaaa cttgttctg caagtcaagc 3780
tgacctaggc ttataggtcg acaccggtag ataactgatc ggatctaggc tcgacatgct 3840
ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc tgcaataaac 3900
aagttaacia caacaattgc attcatttta tgtttcaggt tcagggggag gtgtgggagg 3960
ttttttaaac tcgagatcca ctagggccgc aggaaccctc agtgatggag ttggccactc 4020

cctctctgcg cgctcgcctcg ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg 4080
gctttgcccg ggcggcctca gtgagcgagc gagecgcgag ctgcctgcag gggcgcctga 4140
tgcggtatth tctccttacg catctgtgcg gtatttcaca ccgcatacgt caaagcaacc 4200
atagtacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt 4260
gaccgctaca cttgccagcg ccctagcgcc cgctccttc gctttcttcc cttcctttct 4320
cgccacgttc gccggccttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg 4380
athtagtgct ttacggcacc tcgaccccaa aaaacttgat ttgggtgatg gttcacgtag 4440
tgggcatcgc ccctgataga cggtttttcg cccttgacg ttggagtcca cgttctttaa 4500
tagtggactc ttgttccaaa ctggaacaac actcaacct atctcgggct attcttttga 4560
tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga tttacaacaaa 4620
atttaacgcg aattttaaca aaatattaac gtttacaatt ttatggtgca ctctcagtac 4680
aatctgctct gatgccgat agttaagcca gcccgcacac ccgccaacac ccgctgacgc 4740
gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg 4800
gagctgcatg tgtcagaggt tttcacgctc atcaccgaaa cgcgcgagac gaaagggcct 4860
cgtgatacgc ctatthttat aggttaatgt catgataata atggtttctt agacgtcagg 4920
tggcacttht cggggaaaat tgccgcggaac ccctatthgt ttatthttct aaatacattc 4980
aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag 5040
gaagagtatg agtattcaac atthccgtgt cgccttatt ccctthtttg cggcattthg 5100
ccttctgtt thtgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt 5160
gggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagtht 5220
tcgccccgaa gaacgthttc caatgatgag cactthttaa gthctgctat gtggcgcggt 5280
attatcccgt attgacccg ggcaagagca actcggctgc cgcatacact atthctcagaa 5340
tgactthgtt gactactcac cagtcacaga aaagcatctt acggatggca tgacagtaag 5400
agaattatgc agtgctgcca taacctgag tgataacact gcggccaact tactthctgac 5460
aacgatcggg gaaccgaagg agctaaccgc thttthgac aacatggggg atcatgtaac 5520
tcgcttgat cgtthgggaa cggagctgaa tgaagccata ccaaacgacg agcgtgacac 5580
cacgatgcct gtagcaatgg caacaacgth gcgcaacta ttaactggcg aactactthc 5640
tctagctthc cgccaacaat taatagactg gatggaggcg gataaagthg caggaccact 5700
tctgcgctcg gccctthccg ctggctggtt thttgctgat aatctggag ccggtgagcg 5760
tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctcc gtatcgtagt 5820
tatctacacg accgggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat 5880
aggthcctca ctgattaagc atthgtaact gtcagaccaa gthtactcat atatacttht 5940
gattgatttht aaactthcatt thtaatttht aaggatctag gtgaagatcc thttthgata 6000
tctcatgacc aaaaatccctt aacgtgagth ttcgthccac tgagcgtcag acccctgata 6060
aaagatcaaa ggatctthctt gagatcctth thttctgcgc gtaatctgct gctthgcaac 6120
aaaaaaacca ccgctaccag ccgthggtthg thtgccggt caagagctac caactcttht 6180
tccgaaggta actggtthca gcagagcgc gataccaaat actgthctthc tagthtagcc 6240
gtagthtagc caccactthca agaactctgt agcaccgct acatacctcg ctctgctaat 6300
cctgthtaca gtggctgctg ccagthgcca taagctgthg cthaccgggt thgactcaag 6360

acgatagtta cccgataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc 6420
 cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag 6480
 cgccacgctt cccgaaggga gaaaggcggga caggtatccg gtaagcggca gggtcggaac 6540
 aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg 6600
 gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct 6660
 atggaaaaaac gccagcaacg cggccttttt acggttctct gccttttgct ggccttttgc 6720
 tcacatgt 6728

<210> 6

<211> 6728

<212> DNA

<213> 人工序列

<220>

<223> AAV_{ss}-1XSERP-mTTR-MVM-hFIX_{co}Padua-Albco-SV40pA

<400> 6

cctgcaggca gctgcgcgct cgtcgcctca ctgaggccgc ccgggcaaag cccgggcgct 60
 gggcgacctt tggtcgcccg gcctcagtga gcgagcagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgcg gtaccggcgc gccgggggag gctgctggtg 180
 aatattaacc aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacacgcg 240
 tggtagcgtc tgtctgcaca tttcgtagag cgagtgttcc gatactctaa tctccctagg 300
 caaggttcat atttgtgtag gttacttatt ctcttttgt tgactaagtc aataatcaga 360
 atcagcaggt ttggagtcag cttggcagg atcagcagcc tgggttgaa ggaggggta 420
 taaaagcccc ttcaccagga gaagccgtca cacagatcca caagctcctg aagaggtaa 480
 ggtttaagg atggttggtt ggtggggtat taatgtttaa ttacctggag cacctgcctg 540
 aaatcacttt ttttcaggtt ggctagtctt agaagatctc aagcttgcta gcatgcagc 600
 cgtgaacatg atcatggccg agagccccgg cctgatcacc atctgcctgc tgggctacct 660
 gctgagcgc gagtgcaccg tttcctgga ccacgagaac gccaacaaga tcctgaaccg 720
 cccaagcgc tacaacagcg gcaagctgga ggagttcgtg cagggaacc tggagcgcga 780
 gtgcatggag gagaagtga gttcagga ggcccgcgag gtgttcgaga acaccgagc 840
 caccaccgag ttctggaagc agtacgtgga cggcgaccag tgcgagagca acccctgcct 900
 gaacggcggc agctgcaagg acgacatcaa cagctacgag tgctggtgcc cttcggctt 960
 cgagggcaag aactgcgagc tggacgtgac ctgcaacatc aagaacggcc gctgcgagca 1020
 gttctgcaag aacagcgcg acaacaaggt ggtgtgcagc tgcaccgagg gctaccgctt 1080
 ggccgagaac cagaagagct gcgagcccgc cgtgcccttc cctgcggcc gcgtgagcgt 1140
 gagccagacc agcaagctga cccgcgccga ggccgtgttc cccgacgtgg actacgtgaa 1200
 cagcaccgag gccgagacca tcctggaaa catcaccag agcaccaga gcttcaacga 1260
 cttcaccgc gtggtgggcg gcgaggacgc caagcccgc cagttcccct ggcaggtggt 1320
 gctgaacggc aaggtggacg cttctgcgg cggcagcatc gtgaacgaga agtggatcgt 1380
 gaccgccgc cactgcgtgg agaccggcgt gaagatcacc gtggtggccg gcgagcaaa 1440
 catcgaggag accgagcaca ccgagcagaa gcgcaacgtg atccgcatca tccccacca 1500

caactacaac gccgcatca acaagtacaa ccacgacatc gccctgctgg agctggacga 1560
gcccctggtg ctgaacagct acgtgacccc catctgcatc gccgacaagg agtacaccaa 1620
catcttcctg aagttcggca gcggtacgt gagcggctgg ggccgcgtgt tccacaaggg 1680
ccgcagcgcg ctggtgctgc agtacctgcg cgtgccctg gtggaccgcg ccacctgcct 1740
gctgagcacc aagttcacca tctacaacaa catgttctgc gccggcttcc acgagggcgg 1800
ccgcgacagc tgccagggcg acagcggcg cccccacgtg accgaggtgg agggcaccag 1860
cttcctgacc ggcatcatca gctggggcga ggagtgcgcc atgaaggga agtacggcat 1920
ctacaccaag gtgagccgct acgtgaactg gatcaaggag aagaccaagc tgacctctgt 1980
gagccagacc tccaagctca ccagggccga gactgtcttc cctgatgtgg acggatccga 2040
cgcccacaag agcaggttg cccacagatt caaggacctg ggcgaggaaa acttcaaggc 2100
tctggtgctg atcgcttcg cccagtaact gcagcagtgt cccttcgagg accacgtgaa 2160
gctggtcaac gaagtgaccg agttcgccaa gacctgcgtg gccgacgaga gcgccgagaa 2220
ctgcgacaag agcctgcaca cctgttctgg cgacaagctg tgcaccgtgg ccacctgcg 2280
ggaaacctac ggcgagatgg ccgactgctg cgccaagcag gaacctgagc ggaacgagtg 2340
cttcctgcag cacaaggacg ataaccccaa cctgccccgg ctctgtcggc ccgaggtgga 2400
cgtgatgtgc accgccttc acgacaacga ggaaacctc ctgaagaagt acctgtacga 2460
gatcgccaga cggcaccctt acttctacgc ccccgagctg ctgttcttcg ccaagcggta 2520
caaggccgcc ttcaccgagt gctgccaggc cgccgataag gccgcctgcc tgctgcccga 2580
gctggatgag ctgagggacg agggcaaggc cagctccgcc aagcagagac tgaagtgcgc 2640
cagcctgcag aagttcggcg agagaccctt taaggcctgg gctgtggccc ggctgagcca 2700
gagattccc aaggccgagt ttgccgaggt gtccaagctg gtcaccgacc tcaccaaggt 2760
gcacaccgag tgttgtcacg gcgacctgct ggaatgcgcc gacgacagag ccgacctggc 2820
caagtacatc tgcgagaacc aggacagcat cagctccaag ctgaaagagt gctgcgagaa 2880
gcccctgctg gaaaagagcc actgtatcgc cgaggtggaa aacgacgaga tgcccgccga 2940
cctgcccagc ctggccgcc acttctgtgga aagcaaggac gtgtgcaaga actacgccga 3000
ggccaaggat gtgttcttg gcatgttctt gtatgagtac gcccgacagc accccgacta 3060
cagcgtggtg ctgctgctgc ggctggccaa gacctacgag acaaccttg aaaagtgctg 3120
cgccgctgcc gacccccacg agtgctacgc caaggtgttc gacgagttca agccactggt 3180
ggaagaacct cagaacctga tcaagcagaa ttgcgagctg ttcgagcagc tgggcgagta 3240
caagttccag aacgcctgc tctgtcggta caccaagaaa gtgccccagg tgtccacccc 3300
caccctggtg gaagtgtccc ggaacctggg caaagtgggc agcaagtgct gcaagcacc 3360
tgaggccaag cggatgcctt gcgccgagga ctacctgagc gtggtctctga accagctgtg 3420
cgtgctgcac gagaaaacct ccgtgtccga cagagtgacc aagtgtgta ccgagagcct 3480
ggtcaacaga cggccttget tctccgctt ggaagtggac gagacatacg tgcccagaaga 3540
gttcaacgcc gagacattca ccttccacgc cgacatctgc accctgagcg agaaagagcg 3600
gcagatcaag aagcagaccg ccttggtcga gctggtcaag cacaagccca aggccaccaa 3660
agaacagctg aaggccgtga tggacgactt cgccgccttc gtcgagaagt gttgcaaggc 3720
cgacgacaaa gagacatgct tcgccgaaga gggcaaaaag ctggtggccg cctctcaggc 3780
cgccctggga ctctaagtgc acaccggtag ataactgatc ggatctaggc tcgacatgct 3840

ttatttgtga aatttgtgat gctattgctt tatttghtaac cattataagc tgcaataaac 3900
 aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag gtgtgggagg 3960
 ttttttaaac tcgagatcca ctagggccgc aggaaccctt agtgatggag ttggccactc 4020
 cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc cgacgcccgg 4080
 gctttgcccc ggcggcctca gtgagcgagc gagcgcgcag ctgcctgcag gggcgcctga 4140
 tgcggtatatt tctccttacg catctgtgcg gtatttcaca ccgcatacgt caaagcaacc 4200
 atagtacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt 4260
 gaccgctaca cttgccagcg ccctagcgcc cgctccttc gctttcttcc ctccctttct 4320
 cgccacgttc gccggcttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg 4380
 atttagtgct ttacggcacc tcgaccccaa aaaacttgat ttgggtgatg gttcacgtag 4440
 tgggccatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca cgttctttaa 4500
 tagtggactc ttgttccaaa ctggaacaac actcaacct atctcgggct attcttttga 4560
 tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga ttttaacaaaa 4620
 atttaacgcg aattttaaca aaatattaac gtttacaatt ttatggtgca ctctcagtac 4680
 aatctgctct gatgccgat agttaagcca gccccgacac ccgccaacac ccgctgacgc 4740
 gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg 4800
 gagctgcatg tgcagaggt tttcacgctc atcaccgaaa cgcgcgagac gaaagggcct 4860
 cgtgatacgc ctatttttat aggttaatgt catgataata atggtttctt agacgtcagg 4920
 tggcactttt cggggaaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc 4980
 aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag 5040
 gaagagtatg agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg 5100
 ccttctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt 5160
 ggggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt 5220
 tcgccccgaa gaacgttttc caatgatgag cactttttaa gttctgctat gtggcgcggt 5280
 attatcccgt attgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa 5340
 tgacttgggt gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag 5400
 agaattatgc agtgctgcca taaccatgag tgataaact cggccaact tacttctgac 5460
 aacgatcgga ggaccgaagg agctaaccgc tttttgac aacatggggg atcatgtaac 5520
 tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaagcagc agcgtgacac 5580
 cacgatgcct gtagcaatgg caacaacgct gcgcaacta ttaactggcg aactacttac 5640
 tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact 5700
 tctgcgctcg gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg 5760
 tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctcc gtatcgtagt 5820
 tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat 5880
 aggtgcctca ctgattaagc attggttaact gtcagaccaa gtttactcat atatacttta 5940
 gattgattta aaacttcatt ttttaatttaa aaggatctag gtgaagatcc tttttgataa 6000
 tctcatgacc aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga 6060
 aaagatcaaa ggatcttctt gagatccttt tttctgcgc gtaatctgct gcttgcaaac 6120
 aaaaaaacca ccgctaccag cggtggtttg tttgccgat caagagctac caactctttt 6180

tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc 6240
 gtagttaggc caccacttca agaactctgt agcaccgctt acatacctcg ctctgctaata 6300
 cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag 6360
 acgatagtta ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc 6420
 cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag 6480
 cgccacgctt cccgaaggga gaaaggcgga caggtatccg gtaagcggca ggttcggaac 6540
 aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg 6600
 gtttcgccac ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct 6660
 atggaaaaaac gccagcaacg cggccttttt acggttctg gccttttgct ggccttttgc 6720
 tcacatgt 6728

<210> 7

<211> 6851

<212> DNA

<213> 人工序列

<220>

<223> AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA

<400> 7

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgct 60
 gggcgacctt tggctgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc tgcggccgcg gtagggggag gctgctggtg aatattaacc 180
 aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacggggg aggctgctgg 240
 tgaatattaa ccaaggtcac cccagttatc ggaggagcaa acaggggcta agtccacggg 300
 ggaggctgct ggtgaatatt aaccaagtc accccagtta tcggaggagc aaacaggggc 360
 taagtccacc gtctgtctgc acatttcgta gagcagtggt tccgatactc taatctccct 420
 aggcaagggt catatattgt taggttactt attctccttt tgttgactaa gtcaataatc 480
 agaatcagca ggtttgaggt cagcttggca gggatcagca gcctggggtg gaaggagggg 540
 gtataaaaag cccttcacca ggagaagccg tcacacagat ccacaagctc ctgaagaggt 600
 aagggtttaa gggatggtg gttggtgggg tattaatggt taattacctg gagcacctgc 660
 ctgaaatcac tttttttcag gttggctagt tctagaagat ctcaagcttg ctagcatgca 720
 gcgcgtgaac atgatcatgg ccgagagccc cggcctgatc accatctgcc tgctgggcta 780
 cctgctgagc gccgagtgca ccgtgttctt ggaccacgag aacgccaaca agatcctgaa 840
 ccgccccaaag cgctacaaca gcggcaagct ggaggagttc gtgcagggca acctggagcg 900
 cgagtgcatt gaggagaagt gcagcttcca ggaggccgc gaggtgttcg agaacaccga 960
 gcgcaccacc gatttctgga agcagtagct ggacggcgac cagtgcgaga gcaaccctg 1020
 cctgaacggc ggcagctgca aggacgacat caacagctac gactgctggt gcccttcgg 1080
 cttcgagggc aagaactgag agctggacgt gacctgcaac atcaagaacg gccgctgca 1140
 gcagttctgc aagaacagcg ccgacaaca ggtggtgtgc agctgcaccg agggtaccg 1200
 cctggccgag aaccagaaga gctgcgagcc cgccgtgcc ttcccctgag gccgctgag 1260
 cgtgagccag accagcaagc tgaccgcgc cgaggccgtg ttcccgcagc tggactacgt 1320

gaacagcacc gaggccgaga ccatcctgga caacatcacc cagagcacc agagcttcaa 1380
 cgacttcacc cgcgtggtgg gcggcgagga cgccaagccc ggccagttcc cctggcaggt 1440
 ggtgctgaac ggcaaggtgg acgccttctg cggcggcagc atcgtgaacg agaagtggat 1500
 cgtgaccgcc gccactgcg tggagaccgg cgtgaagatc accgtggtgg ccggcgagca 1560
 caacatcgag gagaccgagc acaccgagca gaagcgcaac gtgatccgca tcatcccca 1620
 ccacaactac aacgccgcca tcaacaagta caaccagac atcgccctgc tggagctgga 1680
 cgagcccctg gtgctgaaca gctacgtgac ccccatctgc atcgccgaca aggagtacac 1740
 caacatcttc ctgaagtctg gcagcggcta cgtgagcggc tggggccgcg tgttcacaa 1800
 gggccgcagc gccctggtgc tgcagtacct gcgcgtgcc ctggtggacc gcgccacctg 1860
 cctgctgagc accaagttca ccatctaaa caacatgttc tgcgccggct tccacgaggg 1920
 cggccgcgac agctgccagg gcgacagcgg cggcccccac gtgaccgagg tggagggcac 1980
 cagcttcctg accggcatca tcagctgggg cgaggagtgc gccatgaagg gcaagtacgg 2040
 catctacacc aagtgagcc gctacgtgaa ctggatcaag gagaagacca agctgacctc 2100
 tgtgagccag acctccaagc tcaccagggc cgagactgtc ttcctgatg tggacggatc 2160
 cgatgcacac aagagtgagg ttgctcatcg gtttaaatg ttgggagaag aaaatttcaa 2220
 agccttggtg ttgattgctt ttgctcagta tcttcagcag tgtccatttg aagatcatgt 2280
 aaaattagtg aatgaagtaa ctgaatttgc aaaaacatgt gttgctgatg agtcagctga 2340
 aaattgtgac aaatcacttc ataccctttt tggagacaaa ttatgcacag ttgcaactct 2400
 tcgtgaaacc tatggtgaaa tggctgactg ctgtgcaaaa caagaacctg agagaaatga 2460
 atgcttcttg caacacaaag atgacaacc aaacctccc cgattggtga gaccagaggt 2520
 tgatgtgatg tgcactgctt tcatgacaa tgaagagaca tttttgaaaa aatacttata 2580
 tgaaattgcc agaagacatc ctactttta tgccccgaa ctcttttct ttgctaaaag 2640
 gtataaagct gcttttacag aatgttgcca agctgctgat aaagctgcct gcctgttgcc 2700
 aaagctcgat gaacttcggg atgaaggaa ggcttcgtct gccaaacaga gactcaagtg 2760
 tgccagtctc caaaaattg gagaaagagc tttcaaagca tgggcagtag ctgcctgag 2820
 ccagagattt cccaaagctg agtttgaga agtttccaag ttagtgacag atcttaccaa 2880
 agtccacacg gaatgctgcc atggagatct gcttgaatgt gctgatgaca gggcggacct 2940
 tgccaagtat atctgtgaaa atcaagattc gatctccagt aaactgaagg aatgctgtga 3000
 aaaacctctg ttggaanaat cccactgcat tgccgaagtg gaaaatgatg agatgcctgc 3060
 tgacttgctt tcattagctg ctgattttgt tgaaagtaag gatgtttgca aaaactatgc 3120
 tgaggcaaaag gatgtcttcc tgggcatgtt tttgtatgaa tatgcaagaa ggcacctga 3180
 ttactctgtc gtgctgctgc tgagacttgc caagacatat gaaaccactc tagagaagtg 3240
 ctgtgccgct gcagatctc atgaatgcta tgccaaagtg ttcgatgaat ttaaacctct 3300
 tgtggaagag cctcagaatt taatcaaca aaattgtgag ctttttgagc agcttgagga 3360
 gtacaaattc cagaatgcgc tattagtctg ttacaccaag aaagtacccc aagtgtcaac 3420
 tccaactctt gtagaggtct caagaaacct aggaaaagtg ggcagcaaat gttgtaaca 3480
 tcctgaagca aaaagaatgc cctgtgcaga agactatcta tccgtggtcc tgaaccagtt 3540
 atgtgtgttg catgagaaaa cgccagtaag tgacagagtc accaaatgct gcacagaatc 3600
 cttggtgaac aggcgacct gcttttcagc tctggaagtc gatgaaacat acgttcccaa 3660

agagtttaat gctgaaacat tcaccttcca tgcagatata tgcacacttt ctgagaagga 3720
gagacaaatc aagaaacaaa ctgcacttgt tgagctcgtg aaacacaagc ccaaggcaac 3780
aaaagagcaa ctgaaagctg ttatggatga tttcgcagct tttgtagaga agtgctgcaa 3840
ggctgacgat aaggagacct gctttgccga ggagggtaaa aaacttgttg ctgcaagtca 3900
agctgcctta ggcttatagc tcgacaccgg tagataactg atcggatcta ggctcgacat 3960
gctttatttg tgaattttgt gatgctattg ctttatttgt aaccattata agctgcaata 4020
aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg gaggtgtggg 4080
aggtttttta aactcgagat ccactagggc cgcaggaacc cctagtgatg gagttggcca 4140
ctccctctct gcgcgctcgc tcgctcactg aggccggcg accaaaggtc gcccgacgcc 4200
cgggctttgc cggggcgcc tcagtgagcg agcgagcgcg cagctgcctg cagggcgcc 4260
tgatgcggtta ttttctcett acgcatctgt gcggtatttc acaccgata cgtcaaagca 4320
accatagtac gcgccctgta gcggcgctt aagcgcggcg ggtgtggtgg ttacgcgcag 4380
cgtgaccgct aacttgcca gcgccctagc gcccgctct ttcgctttct tcccttctt 4440
tctcgccacg ttcgccggtt ttccccgta agctctaaat cgggggctcc ctttagggtt 4500
ccgatttagt gctttacggc acctcgacc caaaaaactt gatttgggtg atggttcacg 4560
tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttgagt ccacgttctt 4620
taatagtga ctcttgttcc aaactggaac aacactcaac cctatctcgg gctattcttt 4680
tgatttataa gggattttgc cgatttcggc ctatttggtta aaaaatgagc tgatttaaca 4740
aaaatttaac gcgaatttta acaaaatatt aacgtttaca attttatggt gcaactctcag 4800
tacaatctgc tctgatgcc catagttaag ccagccccga caccgcca caccgctga 4860
cgcgccctga cgggcttgc tgctcccggc atccgcttac agacaagctg tgaccgtctc 4920
cgggagctgc atgtgtcaga ggtttcacc gtcatcacc aaacgcgca gacgaaagg 4980
cctcgtgata cgcctatfff tataggtaa tgcctatgata ataatggttt cttagacgtc 5040
aggttggcact tttcggggaa atgtgcgcgg aaccctatt tgtttatfff tctaaataca 5100
ttcaaataatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 5160
aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctfff ttgcggcatt 5220
ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 5280
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tcttgagag 5340
ttttcgccc gaagaacgtt ttccaatgat gagcactfff aaagttctgc tatgtggcgc 5400
ggtattatcc cgtattgacg ccgggcaaga gcaactcggc cccgcatac actattctca 5460
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 5520
aagagaatta tgcagtgtc ccataacat gagtgataac actgcggcca acttacttct 5580
gacaacgatc ggaggaccga aggagctaac cgctttttg cacaacatgg gggatcatgt 5640
aactgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 5700
caccacgatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg gcgaactact 5760
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 5820
acttctgcgc tcggccctc cggttggtg gtttattgct gataaatctg gagccggtga 5880
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 5940
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgtga 6000

gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 6060
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 6120
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 6180
agaaaagatc aaaggatcctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 6240
aacaiaaaaaa ccaccgctac cagcgggtgg ttgtttgccg gatcaagagc taccaactct 6300
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgtg 6360
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 6420
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc 6480
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 6540
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 6600
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc gcagggtcgg 6660
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt atagtctctgt 6720
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 6780
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 6840
tgctcacatg t 6851

<210> 8

<211> 6851

<212> DNA

<213> 人工序列

<220>

<223> AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA

<400> 8

cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc ccgggcaaag cccgggcgct 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcagc gcgagagag ggagtggcca 120
actccatcac taggggttcc tgcggccgcg gtagggggag gctgctggtg aatattaacc 180
aaggtcacc cagttatcgg aggagcaaac aggggctaag tccacggggg aggctgctgg 240
tgaatattaa ccaaggtcac cccagttatc ggaggagcaa acaggggcta agtccacggg 300
ggaggctgct ggtgaatatt aaccaaggtc accccagtta tcggaggagc aaacaggggc 360
taagtccacc gtctgtctgc acatttcgta gagcagtggt tccgatactc taatctccct 420
aggcaaggtt catatttgtg taggttactt attctctttt tgttgactaa gtcaataatc 480
agaatcagca ggtttgaggt cagcttggca gggatcagca gcctggggtt gaaggagggg 540
gtataaaagc cccttcacca ggagaagccg tcacacagat ccacaagctc ctgaagaggt 600
aagggtttaa gggatggttg gttggtgggg tattaatggt taattacctg gagcacctgc 660
ctgaaatcac tttttttcag gttggetagt tctagaagat ctcaagcttg ctagcatgca 720
gcgcgtgaac atgatcatgg ccgagagccc cggcctgatc accatctgcc tgctgggcta 780
cctgctgagc gccgagtgca ccgtgttccct ggaccacgag aacgccaaca agatcctgaa 840
ccgccccaaag cgctacaaca gcggcaagct ggaggagttc gtgcagggca acctggagcg 900
cgagtgcagtg gaggagaagt gcagcttcca ggaggcccgc gaggtgttcg agaacaccga 960
gcgcaccacc gagtcttggg agcagtacgt ggacggcgac cagtgcgaga gcaaccctg 1020

cctgaacggc ggcagctgca aggacgacat caacagctac gagtgtggt gcccttcgg 1080
cttcgagggc aagaactgcg agctggacgt gacctgcaac atcaagaacg gccgctgca 1140
gcagttctgc aagaacagcg ccgacaaca ggtggtgtgc agctgcaccg agggctaccg 1200
cctggccgag aaccagaaga gctgcgagcc cgccgtgccc ttcccctgcg gccgctgag 1260
cgtgagccag accagcaagc tgacccgcg cgaggccgtg ttcccgcagc tggactacgt 1320
gaacagcacc gaggccgaga ccatacctgga caacatcacc cagagcacc agagcttcaa 1380
cgacttcacc cgctgtgtgg gcggcgagga cgccaagccc ggccagttcc cctggcaggt 1440
ggtgctgaac ggcaaggtgg acgccttctg cggcggcagc atcgtgaacg agaagtggat 1500
cgtgaccgcc gccactgcg tggagaccgg cgtgaagatc accgtggtgg ccggcgagca 1560
caacatcgag gagaccgagc acaccgagca gaagcgcaac gtgatccgca tcatcccca 1620
ccacaactac aacgccgcca tcaacaagta caaccagc atcgccctgc tggagctgga 1680
cgagcccctg gtgctgaaca gctacgtgac cccatctgc atcgccgaca aggagtacac 1740
caacatcttc ctgaagttcg gcagcggeta cgtgagcggc tggggccgcg tttccacaa 1800
gggcccagc gccctggtgc tgcagtacct gcgctgccc ctggtggacc gcgccacctg 1860
cctgtgagc accaagtcca ccatacaca caacatgttc tgcgccggt tccacgagg 1920
cggcccgcgac agctgccagg gcgacagcgg cgccccccac gtgaccgagg tggagggcac 1980
cagcttctg accggcatca tcagctgggg cgaggagtgc gccatgaagg gcaagtacgg 2040
catctacacc aaggtgagcc gctacgtgaa ctggatcaag gagaagacca agctgacctc 2100
tgtgagccag acctccaagc tcaccagggc cgagactgtc ttccctgatg tggacggatc 2160
cgacgccac aagagcgagg tggcccacag attcaaggac ctgggcgagg aaaacttcaa 2220
ggctctggtg ctgatgcct tcgcccagta cctgcagcag tgtcccttcg aggaccacct 2280
gaagctggtc aacgaagtga ccgagttcgc caagacctgc gtggccgacg agagcgccga 2340
gaactgcgac aagagcctgc acaccctgtt cggcgacaag ctgtgcaccg tggccacct 2400
gcgggaaacc tacggcgaga tggccgactg ctgcgccaag caggaaccg agcggaacga 2460
gtgcttctg cagcacaagg acgataacc caacctgccc cggctcgtgc ggcccaggt 2520
ggacgtgatg tgaccgcct tccacgaca cgaggaaacc ttctgaaga agtacctgta 2580
cgagatgcc agacggcacc cctacttcta cgccccgag ctgctgttct tcgccaagcg 2640
gtacaaggcc gccttcaccg agtgctgcca ggccgccgat aaggccgct gcctgctgcc 2700
caagctggat gagctgagg acgagggcaa ggccagctcc gccaaagcaga gactgaagt 2760
cgccagcctg cagaagttcg gcgagagagc ctttaaggcc tgggctgtgg cccggctgag 2820
ccagagattc cccaaggccg agtttgccga ggtgtccaag ctggtcaccg acctaccaa 2880
ggtgcacacc gagtgttgc acggcgacct gctggaatgc gccgacgaca gagccacct 2940
ggccaagtac atctgcgaga accaggacag catcagctcc aagctgaaag agtgctgca 3000
gaagcccctg ctggaaaaga gccactgtat cgccgaggtg gaaaacgacg agatgcccgc 3060
cgacctgcc agcctggccg ccgacttctg ggaaagcaag gacgtgtgca agaactacgc 3120
cgaggccaag gatgtgttcc tgggcatgtt cctgtatgag tacgcccga gacacccga 3180
ctacagcgtg gtgctgctgc tgcgctggc caagacctac gagacaacc tggaaaagt 3240
ctgcgccgct gccgaccccc acgagtgtca cgccaaggtg ttcgacgagt tcaagccact 3300
ggtggaagaa ccccagaacc tgatcaagca gaattgcgag ctgttcgagc agctgggcga 3360

gtacaagttc cagaacgccc tgctcgtgcg gtacaccaag aaagtgcccc aggtgtccac 3420
ccccaccctg gtggaagtgt cccggaacct gggcaaagtg ggcagcaagt gctgcaagca 3480
ccctgaggcc aagcggatgc cctgcgccga ggactacctg agcgttgtcc tgaaccagct 3540
gtgcgtgctg cacgagaaaa cccccgtgtc cgacagagtg accaagtgct gtaccgagag 3600
cctggtcaac agacggccct gtttctccgc cctggaagtg gacgagacat acgtgcccac 3660
agagttcaac gccgagacat tcacctcca cgccgacatc tgcaccctga gcgagaaaga 3720
gcggcagatc aagaagcaga ccgccctggt cgagctggtc aagcacaagc ccaaggccac 3780
caaagaacag ctgaaggccg tgatggacga cttcgccgcc ttcgtcgaga agtgttgcaa 3840
ggccgacgac aaagagacat gttcgcgca agagggcaaa aagctggtgg ccgctctca 3900
ggccgccctg ggactctaag tcgacaccgg tagataactg atcggatcta ggctcgacat 3960
gctttatattg tgaattttgt gatgctattg ctttattttg aaccattata agctgcaata 4020
aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg gaggtgtggg 4080
aggtttttta aactcgagat ccactagggc cgcaggaacc cctagtgatg gagttggcca 4140
ctccctctct gcgcgctcgc tcgctcactg aggccggcg accaaaggtc gcccgacgcc 4200
cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg cagctgcctg caggggcgcc 4260
tgatgcggta ttttctcett acgcatctgt gcggtatttc acaccgata cgtcaaagca 4320
accatagtac gcgccctgta gcggcgcatc aagcgcgcg ggtgtggtgg ttacgcgcag 4380
cgtgaccgct acaactgcca gcgccctagc gcccgctcct ttcgctttct tcccttcctt 4440
tctcgccacg ttcgccggtt ttccccgca agctctaaat cgggggctcc ctttaggggt 4500
ccgatttagt gctttacggc acctcgacc caaaaaactt gatttggtg atggttcacg 4560
tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttgagt ccacgttctt 4620
taatagtgga ctcttgttcc aaactggaac aacctcaac cctatctcgg gctattcttt 4680
tgatttataa gggattttgc cgatttcggc ctattggta aaaaatgagc tgatttaaca 4740
aaaatttaac gcgaatttta acaaaatatt aacgtttaca attttatggt gcaactctcag 4800
tacaatctgc tctgatgcc catagttaag ccagccccga caccgcca caccgctga 4860
cgcgccctga cgggcttgc tgctcccggc atccgcttac agacaagctg tgaccgtctc 4920
cgggagctgc atgtgtcaga gttttcacc gtcatcacc aaacgcgca gacgaaagg 4980
cctcgtgata cgctatttt tataggtaa tgtcatgata ataatggtt cttagacgtc 5040
aggtagcact tttcgggaa atgtgcgcg aacctatt tgtttattt tctaaataca 5100
ttcaaatatg tatccgctca tgagacaata acctgataa atgcttcaat aatattgaaa 5160
aaggaagagt atgagtattc aacatttccg tgtcgccctt attcccttt ttgcggcatt 5220
ttgccctcct gtttttgc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 5280
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tcttgagag 5340
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 5400
ggtattatcc cgtattgacg ccgggcaaga gcaactcggc cgccgatac actattctca 5460
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 5520
aagagaatta tgcagtgtc ccataacct gagtgataac actgcggcca acttacttct 5580
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 5640
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaacg acgagcgtga 5700

caccacgatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg gcgaactact 5760
 tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 5820
 acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 5880
 gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 5940
 agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgcgtga 6000
 gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact catatatact 6060
 ttagattgat ttaaaacttc attttaatt taaaaggatc taggtgaaga tcctttttga 6120
 taatctcatg accaaaaatcc cttaacgtga gttttcgctt cactgagcgt cagaccccgt 6180
 agaaaagatc aaaggatcct cttgagatcc ttttttctg cgcgtaatct gctgcttga 6240
 aacaaaaaaaa ccaccgctac cagcgggtgg ttgtttgccg gatcaagagc taccaactct 6300
 ttttccgaag gtaactgget tcagcagagc gcagatacca aatactgtcc ttctagtga 6360
 gccgtagtta ggccaccaact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 6420
 aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc 6480
 aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 6540
 gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 6600
 aagcggcacg cttcccgaag ggagaaaggc ggacaggtat ccgtaagcg gcagggtcgg 6660
 aacaggagag cgcacgaggg agcttcagg gggaaacgcc tggatcttt atagtcctgt 6720
 cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 6780
 cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 6840
 tgctcacatg t 6851

<210> 9

<211> 1389

<212> DNA

<213> 人工序列

<220>

<223> hFIXco

<400> 9

atgcagcgcg tgaacatgat catggccgag agccccggcc tgatcacat ctgcctgctg 60
 ggctacctgc tgagcggcga gtcaccgtg ttcttgacc acgagaacgc caacaagatc 120
 ctgaaccgcc ccaagcgcta caacagcggc aagctggagg agttcgtgca gggcaacctg 180
 gagcgcgagt gcatggagga gaagtgcagc ttcgaggagg cccgcgaggt gttcgagaac 240
 accgagcgcga ccaccgagtt ctggaagcag tacgtggacg gcgaccagtg cgagagcaac 300
 ccctgcctga acggcggcag ctgcaaggac gacatcaaca gctacgagtg ctggtgcccc 360
 ttcggtctcg agggcaagaa ctgcgagctg gacgtgacct gcaacatcaa gaacggccgc 420
 tgcgagcagt tctgcaagaa cagcggccgac aacaaggtgg tgtgcagctg caccgagggc 480
 taccgcctgg ccgagaacca gaagagctgc gagccccccg tgcccttccc ctgcggccgc 540
 gtgagcgtga gccagaccag caagctgacc cgcgccgagg ccgtgttccc cgacgtggac 600
 tacgtgaaca gcaccgaggc cgagaccatc ctggacaaca tcaccagag caccagagc 660
 ttcaacgact tcaccgcgt ggtgggcggc gaggacgcca agccccgcca gttcccctgg 720

caggtggtgc tgaacggcaa ggtggacgcc ttctgcggcg gcagcatcgt gaacgagaag 780
 tggatcgtga ccgccgccca ctgcgtggag accggcgtga agatcacctg gttggccggc 840
 gagcacaaca tcgaggagac cgagcacacc gagcagaagc gcaacgtgat ccgcatcatc 900
 ccccaccaca actacaacgc cgccatcaac aagtacaacc acgacatcgc cctgctggag 960
 ctggacgagc ccctggtgct gaacagctac gtgaccccca tctgcatcgc cgacaaggag 1020
 tacaccaaca tcttcctgaa gttcggcagc ggctacgtga gcggctgggg ccgcgtgttc 1080
 cacaagggcc gcagcgcctt ggtgctgcag tacctgcgcg tccccctggt ggaccgcgcc 1140
 acctgcctgc gcagcaccaa gttcaccatc tacaacaaca tgttctgcgc cggcttccac 1200
 gagggcggcc gcgacagctg ccagggcgac agcggcggcc cccacgtgac cgaggtggag 1260
 ggcaccagct tcctgaccgg catcatcagc tggggcgagg agtgcgcat gaagggcaag 1320
 tacggcatct acaccaaggt gagccgctac gtgaactgga tcaaggagaa gaccaagctg 1380
 acctaata 1389

<210> 10

<211> 3204

<212> DNA

<213> 人工序列

<220>

<223> hFIXco-Albco

<400> 10

atgcagcgcg tgaacatgat catggccgag agccccggcc tgatcacctat ctgcctgctg 60
 ggctacctgc tgagcgccga gtgcaccgtg ttcttgacc acgagaacgc caacaagatc 120
 ctgaaccgcc ccaagcgcta caacagcggc aagctggagg agttcgtgca gggcaacctg 180
 gagcgcgagt gcatggagga gaagtgcagc ttcgaggagg cccgcgaggt gttcgagaac 240
 accgagcgcg caaccgagtt ctggaagcag tacgtggacg gcgaccagtg cgagagcaac 300
 ccctgcctga acggcggcag ctgcaaggac gacatcaaca gctacgagtg ctggtgcccc 360
 ttcggcttcg agggcaagaa ctgcgagctg gacgtgacct gcaacatcaa gaacggccgc 420
 tgcgagcagt tctgcaagaa cagcgcggac aacaaggtgg tgtgcagctg caccgagggc 480
 taccgcctgg ccgagaacca gaagagctgc gagccccggc tgcccttccc ctgcggccgc 540
 gtgagcgtga gccagaccag caagctgacc cgcgcccagg ccgtgttccc cgacgtggac 600
 tacgtgaaca gcaccgagc cgagaccatc ctggacaaca tcaccagag caccagagc 660
 ttcaacgact tcacccgcgt ggtgggcggc gaggacgcca agccccgcca gttcccctgg 720
 caggtggtgc tgaacggcaa ggtggacgcc ttctgcggcg gcagcatcgt gaacgagaag 780
 tggatcgtga ccgccgccca ctgcgtggag accggcgtga agatcacctg gttggccggc 840
 gagcacaaca tcgaggagac cgagcacacc gagcagaagc gcaacgtgat ccgcatcatc 900
 ccccaccaca actacaacgc cgccatcaac aagtacaacc acgacatcgc cctgctggag 960
 ctggacgagc ccctggtgct gaacagctac gtgaccccca tctgcatcgc cgacaaggag 1020
 tacaccaaca tcttcctgaa gttcggcagc ggctacgtga gcggctgggg ccgcgtgttc 1080
 cacaagggcc gcagcgcctt ggtgctgcag tacctgcgcg tccccctggt ggaccgcgcc 1140
 acctgcctgc gcagcaccaa gttcaccatc tacaacaaca tgttctgcgc cggcttccac 1200

gagggcggcc gcgacagctg ccagggcgac agcggcggcc cccacgtgac cgaggtggag 1260
 ggcaccagct tcctgaccgg catcatcagc tggggcgagg agtgcgccat gaagggcaag 1320
 tacggcatct acaccaaggt gagccgctac gtgaactgga tcaaggagaa gaccaagctg 1380
 acctctgtga gccagacctc caagctcacc agggccgaga ctgtcttccc tgatgtggac 1440
 ggatccgacg cccacaagag cgaggtggcc cacagattca aggacctggg cgaggaaaac 1500
 ttcaaggctc tgggtgctgat cgccttcgcc cagtacctgc agcagtgctc cttcgaggac 1560
 cacgtgaagc tggtaacga agtgaccgag ttcccaaga cctgcgtggc cgacgagagc 1620
 gccgagaact gcgacaagag cctgcacacc ctgttcggcg acaagctgtg caccgtggcc 1680
 accctgcggg aaacctacgg cgagatggcc gactgctgcg ccaagcagga acccgagcgg 1740
 aacgagtgtc tcctgcagca caaggacgat aacccaacc tgccccggct cgtgcggccc 1800
 gaggtggacg tgatgtgcac cgccttccac gacaacgagg aaaccttct gaagaagtac 1860
 ctgtacgaga tcgccagacg gcacctac ttctacgcc ccgagctgct gttcttcgcc 1920
 aagcggatac aggccgctt caccgagtgc tgccaggccg ccgataaggc cgctgctctg 1980
 ctgccaagc tggatgagct gagggacgag ggcaaggcca gctccgcaa gcagagactg 2040
 aagtgcgcca gcctgcagaa gttcggcgag agagccttta aggctgggc tgtggcccgg 2100
 ctgagccaga gattcccaa ggccgagttt gccgaggtgt ccaagctggt caccgacctc 2160
 accaagggtc acaccgagtg ttgtcacggc gacctgctgg aatgcgccga cgacagagcc 2220
 gacctggcca agtacatctg cgagaaccag gacagcatca gctccaagct gaaagagtgc 2280
 tgcgagaagc ccctgctgga aaagagccac tgtatcgcc aggtgaaaa cgacgagatg 2340
 cccgccgacc tgcccagcct ggccgccgac ttcgtgaaa gcaaggacgt gtgcaagaac 2400
 tacgccgagg ccaaggatgt gttcctgggc atgttctgt atgagtacg ccgagacac 2460
 cccgactaca gcgtggtgct gctgctgcgg ctggccaaga cctacgagac aaccctggaa 2520
 aagtgctgcg ccgtgcca cccccagag tgctacgcca aggtgttcga cgagttcaag 2580
 ccaactggtg aagaaccca gaacctgac aagcagaatt gcgagctgtt cgagcagctg 2640
 ggcgagtaca agttccagaa cgccctgctc gtgcggtaca ccaagaaagt gccccaggtg 2700
 tccaccccca ccctggtgga agtgtcccgg aacctgggca aagtggcag caagtgtgc 2760
 aagcaccctg aggccaagcg gatgccctgc gccgaggact acctgagcgt ggtcctgaac 2820
 cagctgtgcg tgctgcacga gaaaaccccc gtgtccgaca gactgaccaa gtgctgtacc 2880
 gagagcctgg tcaacagacg gcctgcttcc tccgccctgg aagtggacga gacatacgtg 2940
 cccaaagagt tcaacgccga gacattcacc ttccacgcc acatctgcac cctgagcgag 3000
 aaagagcggc agatcaagaa gcagaccgcc ctggctgagc tggtaagca caagcccaag 3060
 gccaccaaag aacagctgaa ggccgtgat gacgacttcg ccgccttcgt cgagaagtgt 3120
 tgcaaggccg acgacaaaaga gacatgcttc gccgaagagg gcaaaaagct ggtggcccgc 3180
 tctcaggccg ccctgggact ctaa 3204

<210> 11

<211> 1389

<212> DNA

<213> 人工序列

<220>

<223> hFIXcoPadua

<400> 11

atgcagcgcg tgaacatgat catggccgag agccccggcc tgatcacat ctgcctgctg 60
 ggctacctgc tgagcgccga gtgcaccgtg ttcttgacc acgagaacgc caacaagatc 120
 ctgaaccgcc ccaagcgcta caacagcggc aagctggagg agttcgtgca gggcaacctg 180
 gagcgcgagt gcatggagga gaagtgcagc ttcgaggagg cccgcgaggt gttcgagaac 240
 accgagcgca ccaccgagtt ctggaagcag tacgtggacg gcgaccagtg cgagagcaac 300
 ccctgcctga acggcggcag ctgcaaggac gacatcaaca gctacgagtg ctggtgcccc 360
 ttcggttcg agggcaagaa ctgcgagctg gacgtgacct gcaacatcaa gaacggccgc 420
 tgcgagcagt tctgcaagaa cagcgccgac aacaaggtag tgtgcagctg caccgagggc 480
 taccgcctgg ccgagaacca gaagagctgc gagccccgcg tgcccttccc ctgcggccgc 540
 gtgagcgtga gccagaccag caagctgacc cgcgccgagg ccgtgttccc cgacgtggac 600
 tacgtgaaca gcaccgagge cgagaccatc ctggacaaca tcaccagag caccagagc 660
 ttcaacgact tcaccgcgt ggtgggcggc gaggacgcca agccccgcca gttcccctgg 720
 caggtggtgc tgaacggcaa ggtggacgcc ttctgcggcg gcagcatcgt gaacgagaag 780
 tggatcgtga ccgccgcca ctgcgtggag accggcgtga agatcacctg ggtggccggc 840
 gagcacaaca tcgaggagac cgagcacacc gagcagaagc gcaacgtgat ccgcatcctc 900
 ccccaccaca actacaacgc cgccatcaac aagtacaacc acgacatcgc cctgctggag 960
 ctggacgagc ccctggtgct gaacagctac gtgaccccc tctgcatcgc cgacaaggag 1020
 tacaccaaca tcttctgaa gttcggcagc ggctacgtga gcggctgggg ccgcgtgttc 1080
 cacaagggcc gcagcgccct ggtgctgcag tacctgcgcg tgcccctggt ggaccgcgcc 1140
 acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgcgc cggcttccac 1200
 gagggcggcc gcgacagctg ccagggcgac agcggcggcc cccacgtgac cgaggtggag 1260
 ggcaccagct tctgaccgg catcatcagc tggggcgagg agtgcgcat gaagggaag 1320
 tacggcatct acaccaaggt gagccgctac gtgaactgga tcaaggagaa gaccaagctg 1380
 acctaata 1389

<210> 12

<211> 3204

<212> DNA

<213> 人工序列

<220>

<223> hFIXcoPadua-Alb

<400> 12

atgcagcgcg tgaacatgat catggccgag agccccggcc tgatcacat ctgcctgctg 60
 ggctacctgc tgagcgccga gtgcaccgtg ttcttgacc acgagaacgc caacaagatc 120
 ctgaaccgcc ccaagcgcta caacagcggc aagctggagg agttcgtgca gggcaacctg 180
 gagcgcgagt gcatggagga gaagtgcagc ttcgaggagg cccgcgaggt gttcgagaac 240
 accgagcgca ccaccgagtt ctggaagcag tacgtggacg gcgaccagtg cgagagcaac 300
 ccctgcctga acggcggcag ctgcaaggac gacatcaaca gctacgagtg ctggtgcccc 360

ttcggcttcg agggcaagaa ctgcgagctg gacgtgacct gcaacatcaa gaacggccgc 420
 tgcgagcagt tctgcaagaa cagcgccgac aacaaggtgg tgtgcagctg caccgagggc 480
 taccgcctgg ccgagaacca gaagagctgc gagccccgagc tgcccttccc ctgcggccgc 540
 gtgagcgtga gccagaccag caagctgacc cgcgccgagg cegtgttccc cgacgtggac 600
 tacgtgaaca gcaccgaggc cgagaccatc ctggacaaca tcaccagag caccagagc 660
 ttcaacgact tcacccgcgt ggtggcggc gaggacgcca agcccggcca gttcccctgg 720
 caggtggtgc tgaacggcaa ggtggacgcc ttctgcggcg gcagcatcgt gaacgagaag 780
 tggatcgtga ccgccgcca ctgcgtggag accggcgtga agatcacctg gttggccggc 840
 gagcacaaca tcgaggagac cgagcacacc gagcagaagc gcaacgtgat ccgcatcctc 900
 ccccaccaca actacaacgc cgccatcaac aagtacaacc acgacatcgc cctgctggag 960
 ctggacgagc ccctggtgct gaacagctac gtgacccca tctgcatcgc cgacaaggag 1020
 tacaccaaca tcttctgaa gttcggcagc ggctacgtga gcggctgggg ccgctgttcc 1080
 cacaagggcc gcagcgccct ggtgctgcag tacctgcgag tgcccctggt ggaccgccc 1140
 acctgcctgc tgagcaccaa gttcaccatc tacaacaaca tgttctgcgc cggcttccac 1200
 gagggcgccc gcgacagctg ccagggcgac agcggcgccc cccacgtgac cgaggtggag 1260
 ggcaccagct tcctgaccgg catcatcagc tggggcgagg agtgcgcat gaagggaag 1320
 tacggcatct acaccaaggt gagccgctac gtgaactgga tcaaggagaa gaccaagctg 1380
 acctctgtga gccagacctc caagctcacc agggccgaga ctgtcttccc tgatgtggac 1440
 ggatccgatg cacacaagag tgaggttgcct catcggtta aagatttggg agaagaaaat 1500
 ttcaaagcct tgggtgtgat tgcctttgct cagtatcttc agcagtgtcc atttgaagat 1560
 catgtaaaat tagtgaatga agtaactgaa tttgcaaaaa catgtgttgc tgatgagtca 1620
 gctgaaaatt gtgacaaatc acttcatacc ctttttgag acaaattatg cacagttgca 1680
 actcttcgtg aaacctatgg tgaatggct gactgctgtg caaacaaga acctgagaga 1740
 aatgaatgct tcttgcaaca caaagatgac aacccaaacc tccccgatt ggtgagacca 1800
 gaggttgatg tgatgtgcac tgcttttcat gacaatgaag agacattttt gaaaaaatac 1860
 ttatatgaaa ttgccagaag acatccttac ttttatgcc cggaactcct tttctttgct 1920
 aaaaggtata aagctgcttt tacagaatgt tgccaagctg ctgataaagc tgccctgcctg 1980
 ttgccaaagc tcgatgaact tcgggatgaa gggaaggctt cgtctgcca acagagactc 2040
 aagtgtgcca gtctccaaaa atttgagaa agagctttca aagcatgggc agtagctgc 2100
 ctgagccaga gatttcccaa agctgagttt gcagaagttt ccaagttagt gacagatctt 2160
 accaaagtcc acacggaatg ctgcatgga gatctgctg aatgtgctga tgacagggcg 2220
 gaccttgcca agtatatctg tgaaaatcaa gattcgatct ccagtaaact gaaggaatgc 2280
 tgtgaaaaac ctctgttga aaaatccac tgcattgccg aagtggaaa tgatgagatg 2340
 cctgctgact tgccctcatt agctgctgat tttgttga gtaaggatgt ttgcaaaaac 2400
 tatgctgagg caaaggatgt cttcctgggc atgtttttgt atgaatatgc aagaaggcat 2460
 cctgattact ctgtcgtgct gctgctgaga cttgccaaga catatgaaac cactctagag 2520
 aagtgctgtg ccgctgcaga tcctcatgaa tgctatgcca aagtgttcca tgaatttaaa 2580
 cctcttgagg aagagcctca gaatttaac aaacaaaatt gtgagctttt tgagcagctt 2640
 ggagagtaca aattccagaa tgcgctatta gttcgttaca ccaagaaagt accccaagtg 2700

tcaactccaa ctctttaga ggtctcaaga aacctaggaa aagtggcag caaatgttgt 2760
 aaacatcctg aagcaaaaag aatgccctgt gcagaagact atctatccgt ggtcctgaac 2820
 cagttatgtg tgttgcatga gaaaacgcca gtaagtgaca gagtcaccaa atgctgcaca 2880
 gaatccttgg tgaacaggcg accatgcttt tcagctctgg aagtcgatga aacatacgtt 2940
 cccaaagagt ttaatgctga aacattcacc ttccatgcag atatatgcac actttctgag 3000
 aaggagagac aaatcaagaa acaaactgca cttgttgagc tcgtgaaaca caagcccaag 3060
 gcaacaaaag agcaactgaa agctgttatg gatgatttcg cagcttttgt agagaagtgc 3120
 tgcaaggctg acgataagga gacctgcttt gccgaggagg gtaaaaaact tgttgctgca 3180
 agtcaagctg ccttaggctt atag 3204

<210> 13

<211> 3204

<212> DNA

<213> 人工序列

<220>

<223> hFIXcoPadua-Albco

<400> 13

atgcagcgcg tgaacatgat catggccgag agccccggcc tgatcaccat ctgcctgctg 60
 ggctacctgc tgagcgccga gtgcaccgtg ttctggacc acgagaacgc caacaagatc 120
 ctgaaccgcc ccaagcgcta caacagcggc aagctggagg agttcgtgca gggcaacctg 180
 gagcgcgagt gcatggagga gaagtgcagc ttcgaggagg cccgcgaggt gttcgagaac 240
 accgagcgca ccaccgagtt ctggaagcag tacgtggacg gcgaccagtg cgagagcaac 300
 ccctgcctga acggcggcag ctgcaaggac gacatcaaca gctacgagtg ctggtgcccc 360
 ttcggcttcg agggcaagaa ctgcgagctg gacgtgacct gcaacatcaa gaacggccgc 420
 tgcgagcagt tctgcaagaa cagcgcggac aacaaggtgg tgtgcagctg caccgagggc 480
 taccgcctgg ccgagaacca gaagagctgc gagccccccg tgcccttccc ctgcggccgc 540
 gtgagcgtga gccagaccag caagctgacc cgcgccgagg ccgtgttccc cgacgtggac 600
 tacgtgaaca gcaccgaggc cgagaccatc ctggacaaca tcaccagag caccagagc 660
 ttcaacgact tcaccgcgt ggtggcgccg gaggacgcca agccccgcca gttcccctgg 720
 caggtggtgc tgaacggcaa ggtggacgcc ttctgcggcg gcagcatcgt gaacgagaag 780
 tggatcgtga ccgccgccc ctgcgtggag accggcgtga agatcacctg ggtggccggc 840
 gagcacaaca tcgaggagac cgagcacacc gagcagaagc gcaacgtgat ccgcatcatc 900
 ccccaccaca actacaacgc cgccatcaac aagtacaacc acgacatcgc cctgctggag 960
 ctggacgagc ccctggtgct gaacagctac gtgaccccc tctgcatcgc cgacaaggag 1020
 tacaccaaca tcttctgaa gttcggcagc ggctacgtga gcggctgggg ccgcgtgttc 1080
 cacaagggcc gcagcgcct ggtgctgcag tacctgcgcg tgcccctggt ggaccgcgcc 1140
 acctgcctgc tgagaccaa gttcaccatc tacaacaaca tgttctgcgc cggcttccac 1200
 gagggcggcc gcgacagctg ccagggcgac agcggcggcc cccacgtgac cgaggtggag 1260
 ggcaccagct tcctgaccgg catcatcagc tggggcgagg agtgcgcat gaagggcaag 1320
 tacggcatct acaccaaggt gagccgctac gtgaactgga tcaaggagaa gaccaagctg 1380

acctctgtga gccagacctc caagctcacc agggccgaga ctgtcttccc tgatgtggac 1440
 ggatccgacg cccacaagag cgaggtggcc cacagattca aggacctggg cgaggaaaac 1500
 ttcaaggctc tgggtgctgat cgccttcgcc cagtacctgc agcagtgctc cttcgaggac 1560
 cacgtgaagc tggtaacga agtgaccgag ttcgccaaga cctgcgtggc cgacgagagc 1620
 gccgagaact gcgacaagag cctgcacacc ctgttcggcg acaagctgtg caccgtggcc 1680
 accctgcggg aaacctacgg cgagatggcc gactgctgcg ccaagcagga acccgagcgg 1740
 aacgagtgct tcctgcagca caaggacgat aacccaacc tgccccggct cgtgcggccc 1800
 gaggtggacg tgatgtgcac cgccttccac gacaacgagg aaaccttctt gaagaagtac 1860
 ctgtacgaga tcgccagacg gcaccctac ttctacgcc cggagctgct gttcttcgcc 1920
 aagcggatca aggccgctt caccgagtgc tgccaggccg ccgataaggc cgcctgcctg 1980
 ctgcccaagc tggatgagct gagggacgag ggcaaggcca gctccgcca gacagactg 2040
 aagtgcgcca gcctgcagaa gttcggcgag agagccttta aggctgggc tgtggcccgg 2100
 ctgagccaga gattcccaaa ggccgagttt gccgaggtgt ccaagctggt caccgacctc 2160
 accaagggtc acaccgagtg ttgtcacggc gacctgctgg aatgcgccga cgacagagcc 2220
 gacctggcca agtacatctg cgagaaccag gacagcatca gctccaagct gaaagagtgc 2280
 tgcgagaagc ccctgctgga aaagagccac tgtatcgccg aggtggaaaa cgacgagatg 2340
 cccgccgacc tgcccagcct ggccgccgac ttcgtggaaa gcaaggacgt gtgcaagaac 2400
 tacgccgagg ccaaggatgt gttcctgggc atgttctgt atgagtacgc ccgcagacac 2460
 cccgactaca gcgtggtgct gctgctgctg ctggccaaga cctacgagac aacctggaa 2520
 aagtgctgcg ccgctgccga cccccacgag tgctacgcca agtggttcga cgagttcaag 2580
 ccaactggtg aagaacccca gaacctgatc aagcagaatt gcgagctggt cgagcagctg 2640
 ggccgagtaca agttccagaa cgcctgctc gtgcggtaca ccaagaaagt gccccaggtg 2700
 tccaccccca ccctggtgga agtgtcccgg aacctgggca aagtgggacg caagtgctgc 2760
 aagcaccctg aggccaagcg gatgccctgc gccgaggact acctgagcgt ggtcctgaac 2820
 cagctgtgcg tgctgcacga gaaaacccc gtgtccgaca gactgaccaa gtgctgtacc 2880
 gagagcctgg tcaacagacg gccctgctt tccgccctgg aagtggacga gacatacgtg 2940
 cccaaagagt tcaacgccga gacattcacc ttccacgccg acatctgcac cctgagcgag 3000
 aaagagcggc agatcaagaa gcagaccgcc ctggctgagc tggtaagca caagcccaag 3060
 gccaccaaag aacagctgaa ggccgtgatg gacgacttcg ccgccttcgt cgagaagtgt 3120
 tgcaaggccg acgacaaaaga gacatgctt gccgaagagg gcaaaaagct ggtggccgcc 3180
 tctcaggccg ccctgggact ctaa 3204
 <210> 14
 <211> 1758
 <212> DNA
 <213> 人工序列
 <220>
 <223> Albc0
 <400> 14
 gacgccca agagcgaggt ggccacaga ttcaaggacc tgggcgagga aaacttcaag 60

gctctggtgc tgatgcctt cgcccagtac ctgcagcagt gtcccttcga ggaccacgtg 120
 aagctggtca acgaagtgac cgagttcgcc aagacctgcg tggccgacga gagcgccgag 180
 aactgcgaca agagcctgca caccctgttc ggcgacaagc tgtgcaccgt ggccaccctg 240
 cgggaaacct acggcgagat ggccgactgc tgcgccaagc aggaaccga gcggaacgag 300
 tgcttcctgc agcacaagga cgataacccc aacctgcccc ggctcgtgcg gcccgaggtg 360
 gacgtgatgt gcaccgcctt ccacgacaac gaggaaacct tcctgaagaa gtacctgtac 420
 gagatcgcca gacggcacc cttacttttac gccccgagc tgctgttctt cgccaagcgg 480
 tacaaggccg ccttcaccga gtgctgccag gccgcccata aggccgctg cctgctgccc 540
 aagctggatg agctgagga cgagggcaag gccagctccg ccaagcagag actgaagtgc 600
 gccagcctgc agaagtctgg cgagagagcc ttttaaggcct gggctgtggc ccggtgagc 660
 cagagattcc ccaaggccga gtttgccgag gtgtccaagc tggtcaccga cctcaccagg 720
 gtgcacaccg agtgttgtca cggcgacctg ctggaatgag ccgacgacag agccgacctg 780
 gccaaagtaca tctgcgagaa ccaggacagc atcagctcca agctgaaaga gtgctgagc 840
 aagccccctgc tggaaaagag ccaactgtatc gccgaggtgg aaaacgacga gatgcccccc 900
 gacctgcccga gcctggccgc cgacttctgt gaaagcaagg acgtgtgcaa gaactacgcc 960
 gaggccaagg atgtgttctt gggcatgttc ctgtatgagt acgcccgcag acacccccgac 1020
 tacagcgtgg tgctgctgct gcggtggtcc aagacctac agacaaccct ggaaaagtgc 1080
 tgcgccgctg ccgacccccga cgagtgtctac gccaaaggtgt tcgacgagtt caagccactg 1140
 gtggaagaac ccagaaacct gatcaagcag aattgagcgc tgttcgagca gctgggagc 1200
 tacaagttcc agaacgcctt gctcgtgagg tacaccaaga aagtgccccca ggtgtccacc 1260
 cccaccctgg tggaaagtgc ccggaacctg ggcaaagtgg gcagcaagtg ctgcaagcac 1320
 cctgaggcca agcggatgcc ctgcccagag gactacctga gcgtggtcct gaaccagctg 1380
 tgcgtgctgc acgagaaaac cccgtgtcc gacagagtga ccaagtgtg taccgagagc 1440
 ctggtcaaca gacggccctg cttctccgcc ctggaagtgg acgagacata cgtgccccaa 1500
 gagttcaacg ccgagacatt cacctccac gccgacatct gcaccctgag cgagaaagag 1560
 cggcagatca agaagcagac cgccctggtc gagctggtca agcacaagcc caaggccacc 1620
 aaagaacagc tgaaggccgt gatggagcgc ttccgcccct tcgtcgagaa gtgttgcaag 1680
 gccgacgaca aagagacatg cttcgcgcaa gagggcaaaa agctggtggc cgcctctcag 1740
 gccgcccctgg gactctaa 1758

<210> 15

<211> 14

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 15

Ser Phe Ser Gln Asn Pro Pro Val Leu Thr Arg His Gln Arg

1

5

10

<210> 16

<211> 4377

<212> DNA

<213> 人工序列

<220>

<223> coFVIIIdeltaB

<400> 16

```

atgcagatcg agctgtccac ctgctttttt ctgtgctgc tgcggttctg cttcagcgcc 60
acccggcggt actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg 120
ggcgagctgc ccgtggacgc ccggttcccc cccagagtgc ccaagagctt ccccttcaac 180
accagcgtgg tgtacaagaa aacctgttc gtggagtcca ccgaccacct gttcaatata 240
gccaagccca ggccccctg gatgggctg ctgggcccc ccatccaggc cgaggtgtac 300
gacaccgtgg tgatcacct gaagaacatg gccagccacc ccgtgacct gcacgcctg 360
ggcgtgagct actggaaggc cagcgagggc gccgagtac acgaccagac cagccagcgg 420
gagaaagaag atgacaaggt gttccctggc ggcagccaca cctacgtgtg gcaggtgctg 480
aaagaaaacg gccccatggc ctccgacccc ctgtgctga cctacagcta cctgagccac 540
gtggacctgg tgaaggacct gaacagcggc ctgatcggc ctctgctcgt ctgccgggag 600
ggcagcctgg ccaaagagaa aaccagacc ctgcacaagt tcctctgct gttcgcctg 660
ttcgacgagg gcaagagctg gcacagcag acaaagaaca gcctgatgca ggaccgggac 720
gccgcctctg ccagagcctg gcccaagatg cacaccgtga acggctacgt gaacagaagc 780
ctgcccggcc tgattggctg ccaccggaag agcgtgtact ggcacgtgat cggcatgggc 840
accacacccg aggtgcacag catctttctg gaagggcaca ctttctggt ccggaaccac 900
cggcaggcca gcctggaaat cagccctatc accttctga ccgccagac actgctgatg 960
gacctgggcc agttcctgct gtttgccac atcagctctc accagcacga cggcatggaa 1020
gcctacgtga aggtggactc ttgccccgag gaacccagc tgcggatgaa gaacaacgag 1080
gaagccgagg actacgacga cgacctgacc gacagcgaga tggacgtggt gcggttcgac 1140
gacgacaaca gccccagctt catccagatc agaagcgtgg ccaagaagca cccaagacc 1200
tgggtgcaat atatcgccg caggaagag gactgggact acgccccct ggtgctggcc 1260
cccagcaca gaagctaca gagccagtac ctgaacaatg gccccagcg gatcggccgg 1320
aagtacaaga aagtgcggtt catggcctac accgacgaga cattcaagac ccgggaggcc 1380
atccagcacg agagcggcat cctgggcccc ctgctgtacg gcaagtgagg cgacacactg 1440
ctgatcatct tcaagaacca ggctagccgg cctacaaca tctaccccc cggcatcacc 1500
gacgtgcggc ccctgtacag caggcggctg cccaaggcg tgaagacct gaaggacttc 1560
cccatcctgc ccggcgagat cttcaagtac aagtggaccg tgacctgga ggacggcccc 1620
accaagagcg accccagatg cctgaccgg tactacagca gcttcgtgaa catggaacgg 1680
gacctggcct ccgggctgat cggacctctg ctgatctgct acaaagaaag cgtggaccag 1740
cggggcaacc agatcatgag cgacaagcgg aacgtgatcc tgttcagcgt gttcagatgag 1800
aaccggtcct ggtatctgac cgagaacatc cagcgttttc tgcccaacc tgccggcgtg 1860
cagctggaag atcccagatt ccaggccagc aacatcatgc actccatcaa tggctacgtg 1920
ttcgactctc tgcagctctc cgtgtgtctg cacgaggtgg cctactggta catcctgagc 1980

```


atcggcgccc agaccgactt cctgagcgtg ttcttcagcg gctacacctt caagcacaag 2040
atgggtgtacg aggacaccct gacctgttc ctttcagcg gcgagacagt gttcatgagc 2100
atggaaaacc ccggcctgtg gattctgggc tgccacaaca gcgacttccg gaaccggggc 2160
atgaccgccc tgctgaaggt gtccagctgc gacaagaaca ccggcgacta ctacgaggac 2220
agctacgagg atatcagcgc ctacctgctg tccaagaaca acgccatcga accccggagc 2280
ttcagccaga accccccctg gctgacgctg caccagcggg agatcacccg gacaaccctg 2340
cagtccgacc aggaagagat cgattacgac gacaccatca gcgtggagat gaagaaagag 2400
gatttcgata tctacgacga ggacgagaac cagagcccca gaagcttcca gaagaaaacc 2460
cggcactact tcattgccgc cgtggagagg ctgtgggact acggcatgag ttctagcccc 2520
cacgtgctgc ggaaccgggc ccagagcggc agcgtgcccc agttcaagaa agtgggtgttc 2580
caggaattca cagacggcag cttcacccag cctctgtata gaggcgagct gaacgagcac 2640
ctggggctgc tggggcccta catcagggcc gaagtggagg acaacatcat ggtgaccttc 2700
cggaatcagg ccagcagacc ctactcette tacagcagcc tgatcagcta cgaagaggac 2760
cagcggcagg gcgccgaacc ccggaagaac ttctgtgaag ccaacgaaac caagacctac 2820
ttctggaaaag tgcagacca catggcccc accaaggacg agttcgactg caaggcctgg 2880
gcctacttca gcgacgtgga tctggaaaag gacgtgact ctggactgat tggcccactc 2940
ctgggtctgcc aactaaacac cctcaacccc gccacggcc gccaggtgac cgtgcaggaa 3000
ttcgccctgt tcttcacat cttcagcag acaaagtctt ggtacttcac cgagaatatg 3060
gaacggaact gcagagcccc ctgcaacatc cagatggaag atcctacctt caaagagaac 3120
taccggttcc acgcatcaa cggctacatc atggacacce tgccctggcct ggtgatggcc 3180
caggaccaga gaatccggtg gtatctgctg tccatgggca gcaacgagaa tatccacagc 3240
atccacttca gcggccacgt gttcacctg cggaagaaag aagagtacaa gatggccctg 3300
tacaacctgt accccggcgt gttcagaca gtggagatgc tgcccagcaa ggccggcatc 3360
tggcgggtgg agtgtctgat cggcgagcac ctgcacgctg gcatgagcac cctgtttctg 3420
gtgtacagca acaagtgcc aacccactg ggcatggcct ctggccacat ccgggacttc 3480
cagatcaccc cctccggcca gtacggccag tgggccccca agctggccag actgcactac 3540
agcggcagca tcaacgcctg gtccacaaa gagcccttca gctggatcaa ggtggacctg 3600
ctggccccta tgatcatcca cggcattaag acccaggcg ccaggcagaa gttcagcagc 3660
ctgtacatca gccagttcat catcatgtac agcctggacg gcaagaagtg gcagacctac 3720
cggggcaaca gcaccggcac cctgatggtg ttcttcggca atgtggacag cagcggcatc 3780
aagcacaaca tcttcaacce cccatcatt gcccggtaca tccggctgca cccacccac 3840
tacagcatta gatccacact gagaatggaa ctgatgggct gcgacctgaa ctctgcagc 3900
atgcctctgg gcatggaaaag caaggccatc agcagcgcce agatcacagc cagcagctac 3960
ttaccaaca tgttcgccac ctggtcccc tccaaggcca ggctgcacct gcagggccgg 4020
tccaacgcct ggcggcctca ggtcaacaac cccaaagaat ggctgcaggt ggactttcag 4080
aaaaccatga aggtgaccgg cgtgaccacc cagggcgtga aaagcctgct gaccagcatg 4140
tacgtgaaag agtttctgat cagcagctct caggatggcc accagtggac cctgtttctt 4200
cagaacggca aggtgaaaag gttccagggc aaccaggact cttcacccc cgtggtgaac 4260
tccctggacc cccccctgct gaccgcctac ctgagaatcc acccccagtc ttgggtgcac 4320

cagatcgccc tcaggatgga agtcctggga tgtgaggccc aggatctgta ctgatga 4377

<210> 17

<211> 25

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 17

Ser Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser Gly Gly

1 5 10 15

Ser Gly Gly Ser Gly Gly Ser Gly Ser

20 25

<210> 18

<211> 21

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 18

Ser Val Ser Gln Thr Ser Lys Leu Thr Arg Ala Glu Thr Val Phe Pro

1 5 10 15

Asp Val Asp Gly Ser

20

<210> 19

<211> 202

<212> DNA

<213> 人工序列

<220>

<223> TTRm (TTR最小启动子)

<400> 19

gtctgtctgc acatttcgta gagcgagtgt tccgatactc taatctccct aggcaagggtt 60

catatattgtg taggttactt attctccttt tgttgactaa gtcaataatc agaatcagca 120

ggttttggagt cagcttggca gggatcagca gcctgggttg gaaggagggg gtataaaagc 180

cccttcacca ggagaagccg tc 202

<210> 20

<211> 92

<212> DNA

<213> 人工序列

<220>

<223> MVM内含子

<400> 20

aagaggtaag ggtttaaggg atggttggtt ggtgggtat taatgtttaa ttacctggag 60
cacctgcctg aaatcacttt tttcaggtt gg 92

<210> 21

<211> 49

<212> DNA

<213> 人工序列

<220>

<223> 合成多聚A

<400> 21

aataaaagat ctttattttc attagatctg tgtgttggtt ttttgtgtg 49

<210> 22

<211> 295

<212> DNA

<213> 人工序列

<220>

<223> BGH多聚A

<400> 22

gatctgagcc gaattcctgc agcccggggg atcagcctcg actgtgcctt ctagttgccca 60
gccatctggt gtttgcccct cccccgtgcc ttcttgacc ctggaagtg ccaactccac 120
tgtcctttcc taataaaatg aggaaattgc atcgattgt ctgagtaggt gtcattctat 180
tctggggggg ggggtggggc aggacagcaa gggggaggat tgggaagaca atagcaggca 240
tgctggggat gcggtgggct ctatggcttc tgaggcggaa agaaccagct gggga 295

<210> 23

<211> 134

<212> DNA

<213> 人工序列

<220>

<223> SV40多聚A

<400> 23

atgctttatt tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa 60
taaacaagtt aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg 120
ggaggttttt taaa 134

<210> 24

<211> 100

<212> DNA

<213> 人工序列

<220>

<223> TTRe
 <400> 24
 cactgggagg atgttgagta agatggaaaa ctactgatga cccttgcaga gacagagtat 60
 taggacatgt ttgaacaggg gccgggcat cagcaggtag 100
 <210> 25
 <211> 72
 <212> DNA
 <213> 人工序列
 <220>
 <223> SerpEnh
 <400> 25
 gggggaggct gctggtgaat attaaccaag gtcacccag ttatcggagg agcaaacagg 60
 ggctaagtcc ac 72
 <210> 26
 <211> 218
 <212> DNA
 <213> 人工序列
 <220>
 <223> 3xSerp
 <400> 26
 gggggaggct gctggtgaat attaaccaag gtcacccag ttatcggagg agcaaacagg 60
 ggctaagtcc accgggggag gctgctggtg aatattaacc aaggtcacc cagttatcgg 120
 aggagcaaac aggggctaag tccaccgggg gaggctgctg gtgaatatta accaaggtca 180
 cccagttat cggaggagca aacaggggct aagtccac 218
 <210> 27
 <211> 585
 <212> PRT
 <213> 智人
 <400> 27
 Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
 1 5 10 15
 Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
 20 25 30
 Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
 35 40 45
 Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
 50 55 60
 Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
 65 70 75 80

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

385	390	395	400
Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro			
	405	410	415
Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys			
	420	425	430
Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys			
	435	440	445
Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His			
	450	455	460
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser			
465	470	475	480
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr			
	485	490	495
Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp			
	500	505	510
Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala			
	515	520	525
Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu			
	530	535	540
Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys			
545	550	555	560
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val			
	565	570	575
Ala Ala Ser Gln Ala Ala Leu Gly Leu			
	580	585	

<210> 28

<211> 1758

<212> DNA

<213> 智人

<400> 28

```

gatgcacaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaatttcaaa 60
gccttggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta 120
aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa 180
aattgtgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt 240
cgtgaaacct atggtgaaat ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa 300
tgcttcttgc aacacaaaaga tgacaacceca aacctceccc gattggtgag accagaggtt 360
gatgtgatgt gcaactgctt tcatgacaat gaagagacat ttttgaaaaa atacttatat 420
gaaattgcca gaagacatcc ttacttttat gccccggaac tccttttctt tgctaaaagg 480
tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca 540

```

aagctcgatg aacttcggga tgaaggaag gcttcgtctg ccaaacagag actcaagtgt 600
 gccagtctcc aaaaatttgg agaaagagct ttcaaagcat gggcagtagc tcgcctgagc 660
 cagagatttc ccaaagctga gtttgagaa gttccaagt tagtgacaga tcttaccaa 720
 gtccacacgg aatgctgcc tggagatctg ctggaatgtg ctgatgacag ggcggacctt 780
 gccaaagtata tctgtgaaaa tcaagattcg atctccagta aactgaagga atgctgtgaa 840
 aaacctctgt tggaaaaatc cactgcatt gccgaagtgg aaaatgatga gatgcctgct 900
 gacttgacctt cattagctgc tgattttgtt gaaagtaagg atgtttgaa aaactatgct 960
 gaggcaaagg atgtcttctt gggcatgttt ttgtatgaat atgcaagaag gcacacctgat 1020
 tactctgtcg tgctgtctgt gagacttgcc aagacatag aaaccactct agagaagtgc 1080
 tgtgccgctg cagatcctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt 1140
 gtggaagagc ctcaagaattt aatcaaaca aattgtgagc tttttgagca gcttgagag 1200
 taaaaattcc agaatgcgct attagttcgt tacaccaaga aagtaccca agtgtcaact 1260
 ccaactcttg tagaggtctc aagaaaccta ggaaaagtgg gcagcaaatg ttgtaaacat 1320
 cctgaagcaa aaagaatgcc ctgtgcagaa gactatctat ccgtggtcct gaaccagtta 1380
 tgtgtgttgc atgagaaaaac gccagtaagt gacagagtc ccaaatgctg cacagaatcc 1440
 ttggtgaaca ggcgacctg cttttcagct ctggaagtgc atgaaacata cgttcccaa 1500
 gagttaaatg ctgaaacatt caccttccat gcagatata gcacactttc tgagaaggag 1560
 agacaaatca agaaacaaac tgcacttggt gagctcgtga aacacaagcc caaggcaaca 1620
 aaagagcaac taaaagctgt tatggatgat ttcgcagctt ttgtagagaa gtgctgcaag 1680
 gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttgttgc tgcaagtcaa 1740
 gctgccttag gcttatag 1758

<210> 29

<211> 63

<212> DNA

<213> 人工序列

<220>

<223> 接头

<400> 29

tctgtgagcc agacctcaa gtcaccagg gccgagactg tcttcctga tgtggacgga 60
tcc 63

<210> 30

<211> 406

<212> PRT

<213> 智人

<400> 30

Ala Asn Ala Phe Leu Glu Glu Leu Arg Pro Gly Ser Leu Glu Arg Glu
 1 5 10 15
 Cys Lys Glu Glu Gln Cys Ser Phe Glu Glu Ala Arg Glu Ile Phe Lys
 20 25 30

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

	340		345		350
Arg Gly Thr Trp Tyr Leu Thr Gly Ile Val Ser Trp Gly Gln Gly Cys					
	355		360		365
Ala Thr Val Gly His Phe Gly Val Tyr Thr Arg Val Ser Gln Tyr Ile					
	370		375		380
Glu Trp Leu Gln Lys Leu Met Arg Ser Glu Pro Arg Pro Gly Val Leu					
385		390		395	400
Leu Arg Ala Pro Phe Pro					
	405				

<210> 31

<211> 152

<212> PRT

<213> 智人

<400> 31

Ala Asn Ala Phe Leu Glu Glu Leu Arg Pro Gly Ser Leu Glu Arg Glu					
1	5		10		15
Cys Lys Glu Glu Gln Cys Ser Phe Glu Glu Ala Arg Glu Ile Phe Lys					
	20		25		30
Asp Ala Glu Arg Thr Lys Leu Phe Trp Ile Ser Tyr Ser Asp Gly Asp					
	35		40		45
Gln Cys Ala Ser Ser Pro Cys Gln Asn Gly Gly Ser Cys Lys Asp Gln					
	50		55		60
Leu Gln Ser Tyr Ile Cys Phe Cys Leu Pro Ala Phe Glu Gly Arg Asn					
65		70		75	80
Cys Glu Thr His Lys Asp Asp Gln Leu Ile Cys Val Asn Glu Asn Gly					
	85		90		95
Gly Cys Glu Gln Tyr Cys Ser Asp His Thr Gly Thr Lys Arg Ser Cys					
	100		105		110
Arg Cys His Glu Gly Tyr Ser Leu Leu Ala Asp Gly Val Ser Cys Thr					
	115		120		125
Pro Thr Val Glu Tyr Pro Cys Gly Lys Ile Pro Ile Leu Glu Lys Arg					
	130		135		140
Asn Ala Ser Lys Pro Gln Gly Arg					
145		150			

<210> 32

<211> 254

<212> PRT

<213> 智人

<400> 32

Ile Val Gly Gly Lys Val Cys Pro Lys Gly Glu Cys Pro Trp Gln Val
 1 5 10 15
 Leu Leu Leu Val Asn Gly Ala Gln Leu Cys Gly Gly Thr Leu Ile Asn
 20 25 30
 Thr Ile Trp Val Val Ser Ala Ala His Cys Phe Asp Lys Ile Lys Asn
 35 40 45
 Trp Arg Asn Leu Ile Ala Val Leu Gly Glu His Asp Leu Ser Glu His
 50 55 60
 Asp Gly Asp Glu Gln Ser Arg Arg Val Ala Gln Val Ile Ile Pro Ser
 65 70 75 80
 Thr Tyr Val Pro Gly Thr Thr Asn His Asp Ile Ala Leu Leu Arg Leu
 85 90 95
 His Gln Pro Val Val Leu Thr Asp His Val Val Pro Leu Cys Leu Pro
 100 105 110
 Glu Arg Thr Phe Ser Glu Arg Thr Leu Ala Phe Val Arg Phe Ser Leu
 115 120 125
 Val Ser Gly Trp Gly Gln Leu Leu Asp Arg Gly Ala Thr Ala Leu Glu
 130 135 140
 Leu Met Val Leu Asn Val Pro Arg Leu Met Thr Gln Asp Cys Leu Gln
 145 150 155 160
 Gln Ser Arg Lys Val Gly Asp Ser Pro Asn Ile Thr Glu Tyr Met Phe
 165 170 175
 Cys Ala Gly Tyr Ser Asp Gly Ser Lys Asp Ser Cys Lys Gly Asp Ser
 180 185 190
 Gly Gly Pro His Ala Thr His Tyr Arg Gly Thr Trp Tyr Leu Thr Gly
 195 200 205
 Ile Val Ser Trp Gly Gln Gly Cys Ala Thr Val Gly His Phe Gly Val
 210 215 220
 Tyr Thr Arg Val Ser Gln Tyr Ile Glu Trp Leu Gln Lys Leu Met Arg
 225 230 235 240
 Ser Glu Pro Arg Pro Gly Val Leu Leu Arg Ala Pro Phe Pro
 245 250

<210> 33

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 33

Arg Lys Arg Arg Lys Arg
 1 5
 <210> 34
 <211> 412
 <212> PRT
 <213> 人工序列
 <220>
 <223> hFVII-2RKR
 <400> 34
 Ala Asn Ala Phe Leu Glu Glu Leu Arg Pro Gly Ser Leu Glu Arg Glu
 1 5 10 15
 Cys Lys Glu Glu Gln Cys Ser Phe Glu Glu Ala Arg Glu Ile Phe Lys
 20 25 30
 Asp Ala Glu Arg Thr Lys Leu Phe Trp Ile Ser Tyr Ser Asp Gly Asp
 35 40 45
 Gln Cys Ala Ser Ser Pro Cys Gln Asn Gly Gly Ser Cys Lys Asp Gln
 50 55 60
 Leu Gln Ser Tyr Ile Cys Phe Cys Leu Pro Ala Phe Glu Gly Arg Asn
 65 70 75 80
 Cys Glu Thr His Lys Asp Asp Gln Leu Ile Cys Val Asn Glu Asn Gly
 85 90 95
 Gly Cys Glu Gln Tyr Cys Ser Asp His Thr Gly Thr Lys Arg Ser Cys
 100 105 110
 Arg Cys His Glu Gly Tyr Ser Leu Leu Ala Asp Gly Val Ser Cys Thr
 115 120 125
 Pro Thr Val Glu Tyr Pro Cys Gly Lys Ile Pro Ile Leu Glu Lys Arg
 130 135 140
 Asn Ala Ser Lys Pro Gln Gly Arg Arg Lys Arg Arg Lys Arg Ile Val
 145 150 155 160
 Gly Gly Lys Val Cys Pro Lys Gly Glu Cys Pro Trp Gln Val Leu Leu
 165 170 175
 Leu Val Asn Gly Ala Gln Leu Cys Gly Gly Thr Leu Ile Asn Thr Ile
 180 185 190
 Trp Val Val Ser Ala Ala His Cys Phe Asp Lys Ile Lys Asn Trp Arg
 195 200 205
 Asn Leu Ile Ala Val Leu Gly Glu His Asp Leu Ser Glu His Asp Gly
 210 215 220
 Asp Glu Gln Ser Arg Arg Val Ala Gln Val Ile Ile Pro Ser Thr Tyr
 225 230 235 240

Val Pro Gly Thr Thr Asn His Asp Ile Ala Leu Leu Arg Leu His Gln
245 250 255
Pro Val Val Leu Thr Asp His Val Val Pro Leu Cys Leu Pro Glu Arg
260 265 270
Thr Phe Ser Glu Arg Thr Leu Ala Phe Val Arg Phe Ser Leu Val Ser
275 280 285
Gly Trp Gly Gln Leu Leu Asp Arg Gly Ala Thr Ala Leu Glu Leu Met
290 295 300
Val Leu Asn Val Pro Arg Leu Met Thr Gln Asp Cys Leu Gln Gln Ser
305 310 315 320
Arg Lys Val Gly Asp Ser Pro Asn Ile Thr Glu Tyr Met Phe Cys Ala
325 330 335
Gly Tyr Ser Asp Gly Ser Lys Asp Ser Cys Lys Gly Asp Ser Gly Gly
340 345 350
Pro His Ala Thr His Tyr Arg Gly Thr Trp Tyr Leu Thr Gly Ile Val
355 360 365
Ser Trp Gly Gln Gly Cys Ala Thr Val Gly His Phe Gly Val Tyr Thr
370 375 380
Arg Val Ser Gln Tyr Ile Glu Trp Leu Gln Lys Leu Met Arg Ser Glu
385 390 395 400
Pro Arg Pro Gly Val Leu Leu Arg Ala Pro Phe Pro
405 410

<210> 35

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 接头

<400> 35

Pro Arg Pro Ser Arg Lys Arg Arg

1

5

AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA (4989 bp)

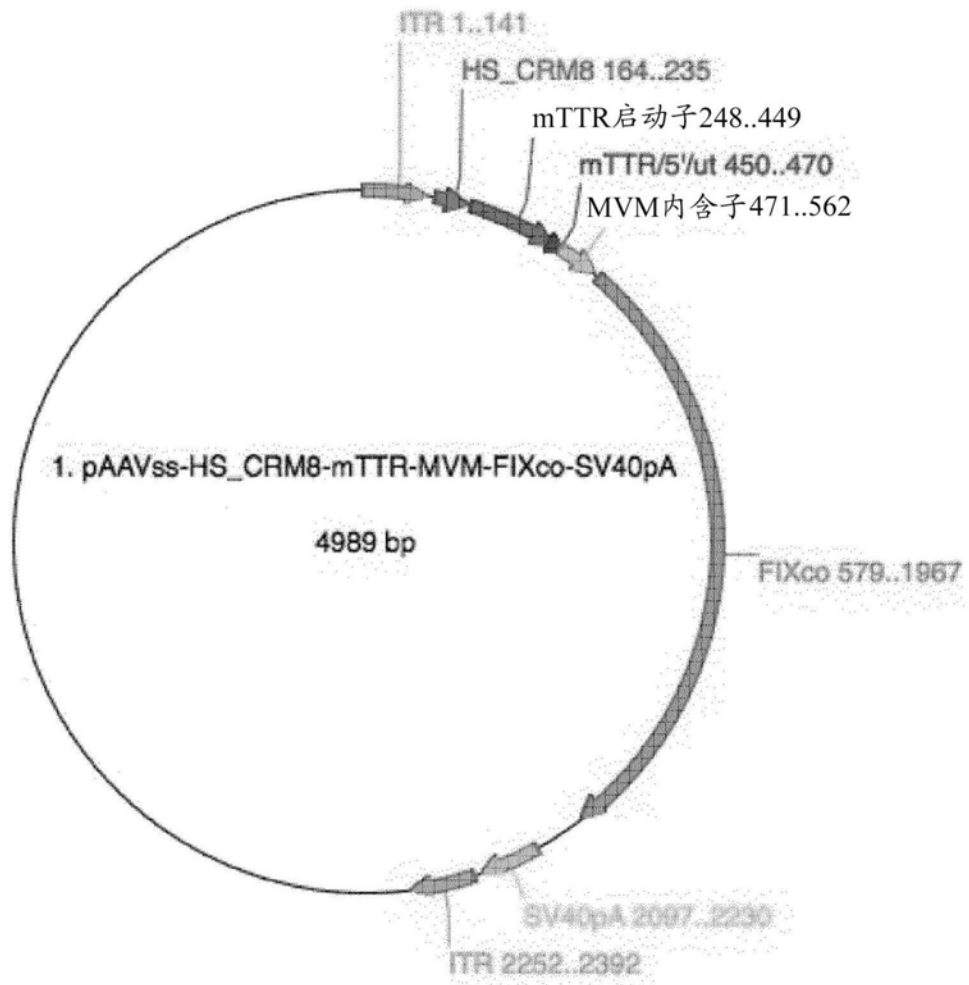


图1

AAVss-1XSERP-mTTR-MVM-hFIXco-Alb-SV40pA (6728 bp)

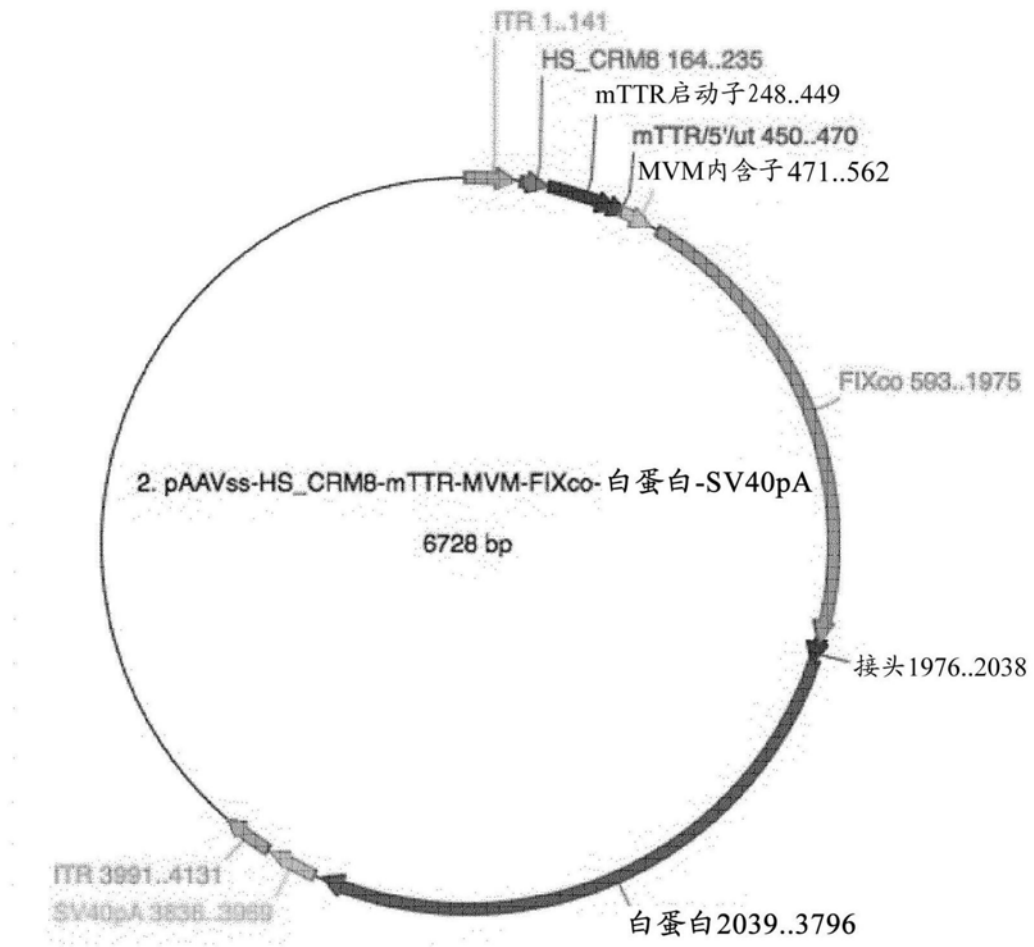


图2

AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA (6728 bp)

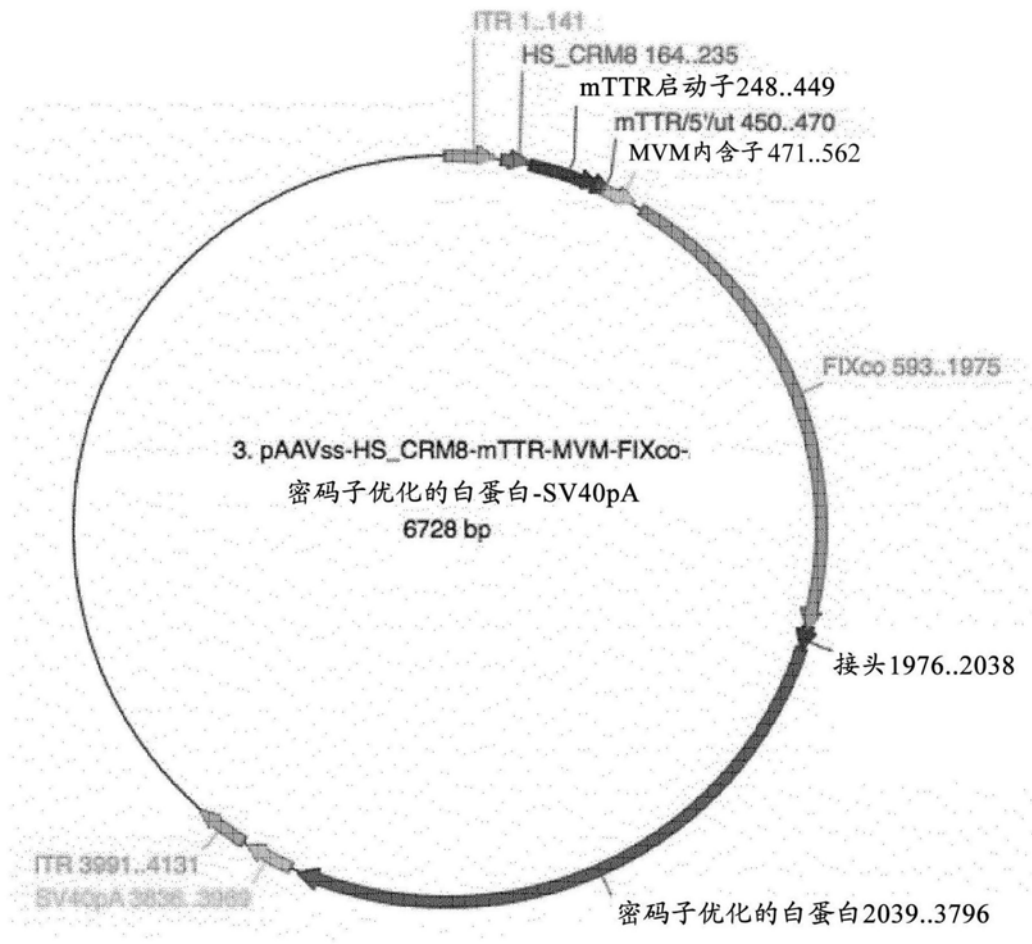


图3

AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-SV40pA (4989 bp)

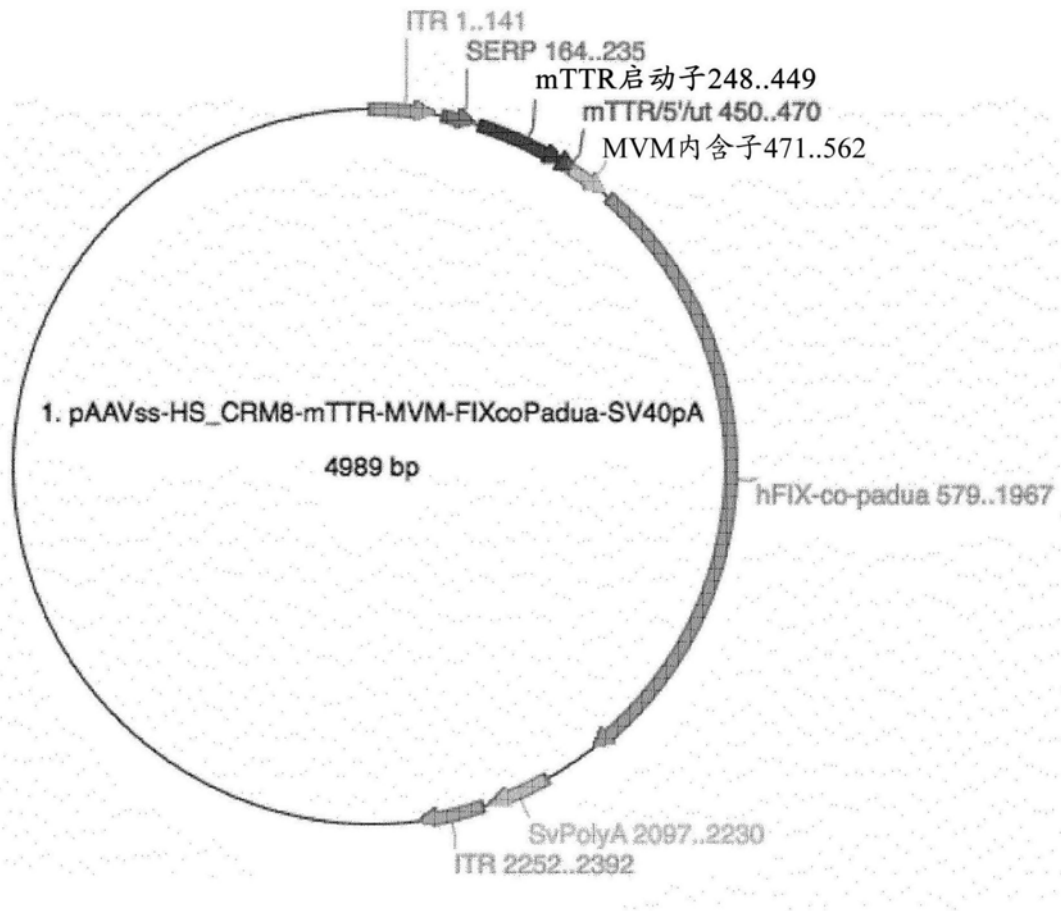


图4

AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA (6728 bp)

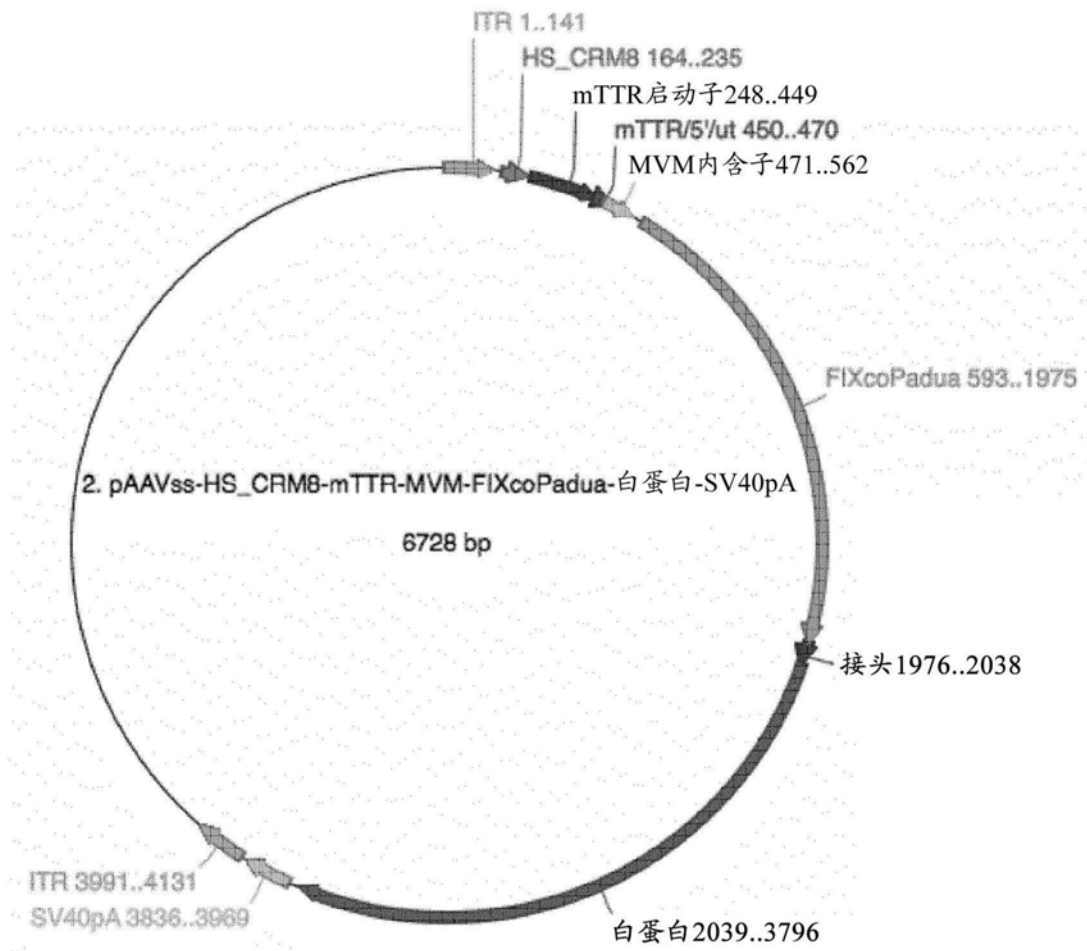


图5

AAVss-1XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (6728 bp)

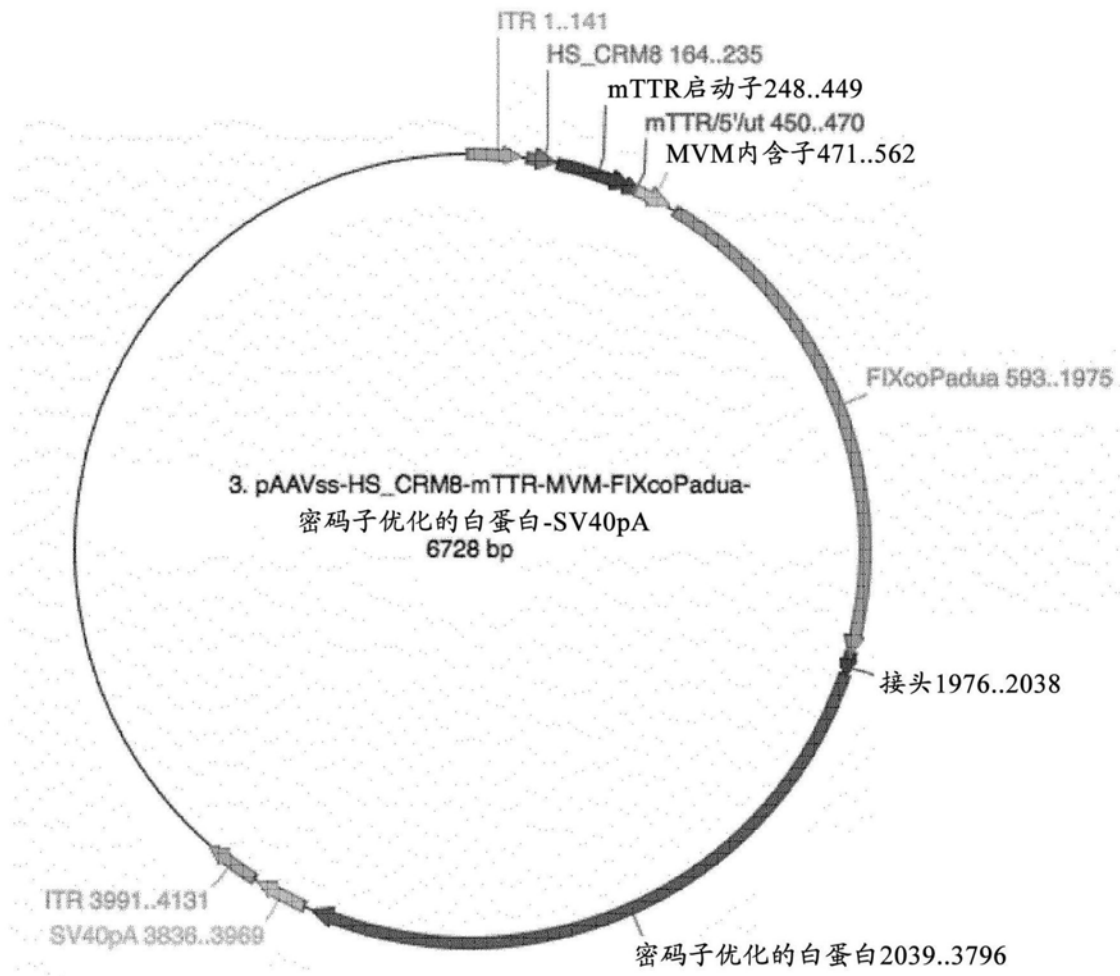


图6

AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA (6851 bp)

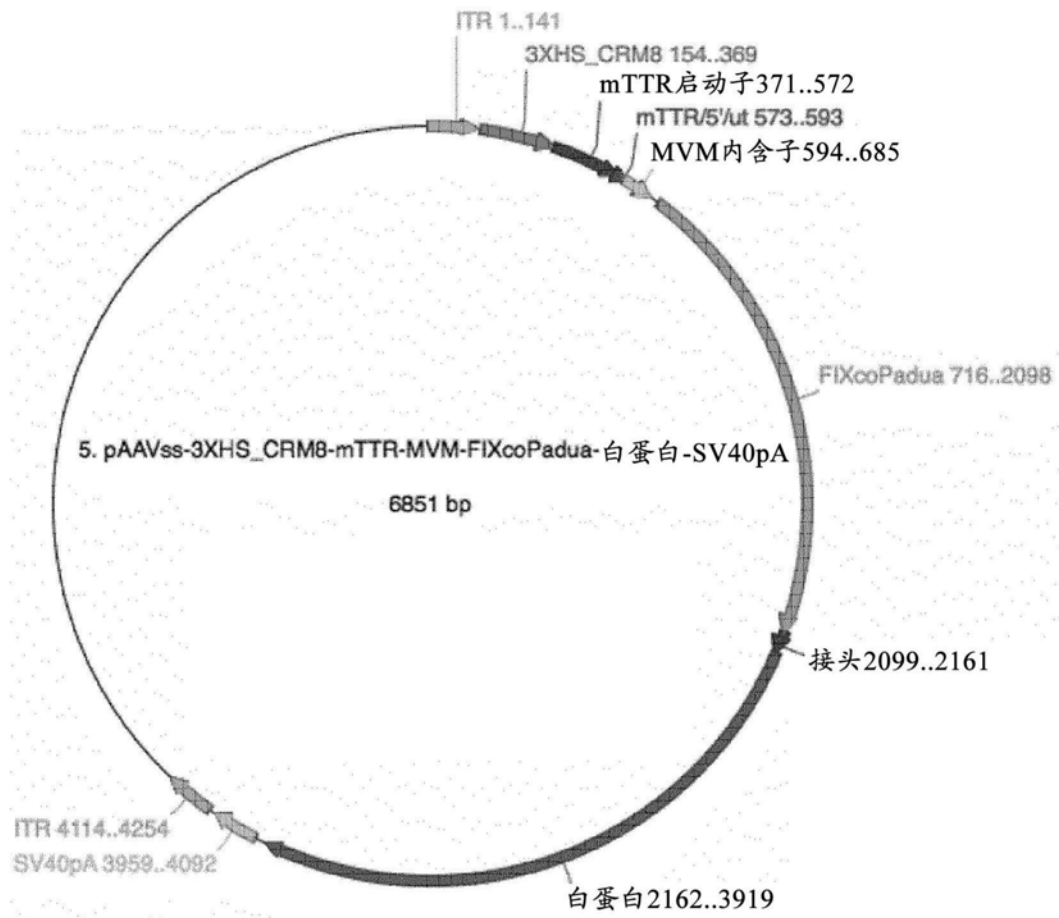


图7

AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA (6851 bp)

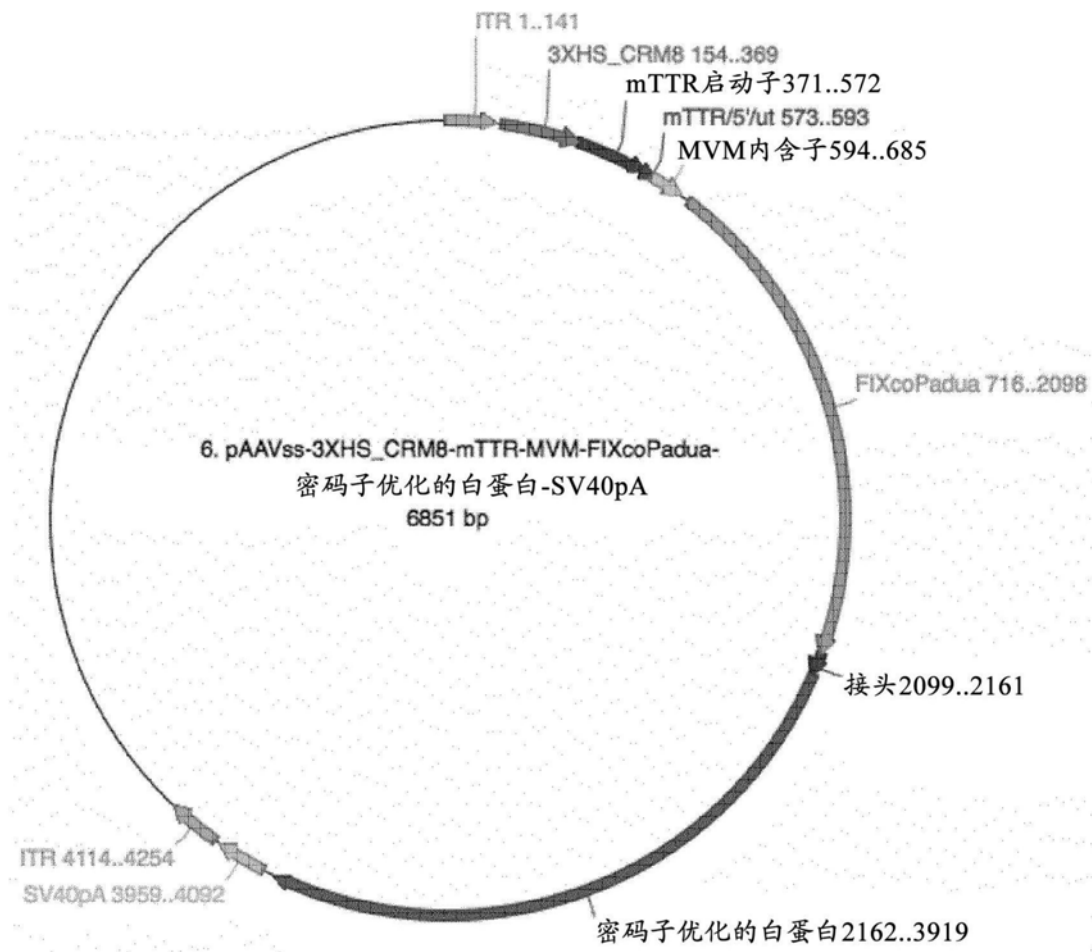


图8

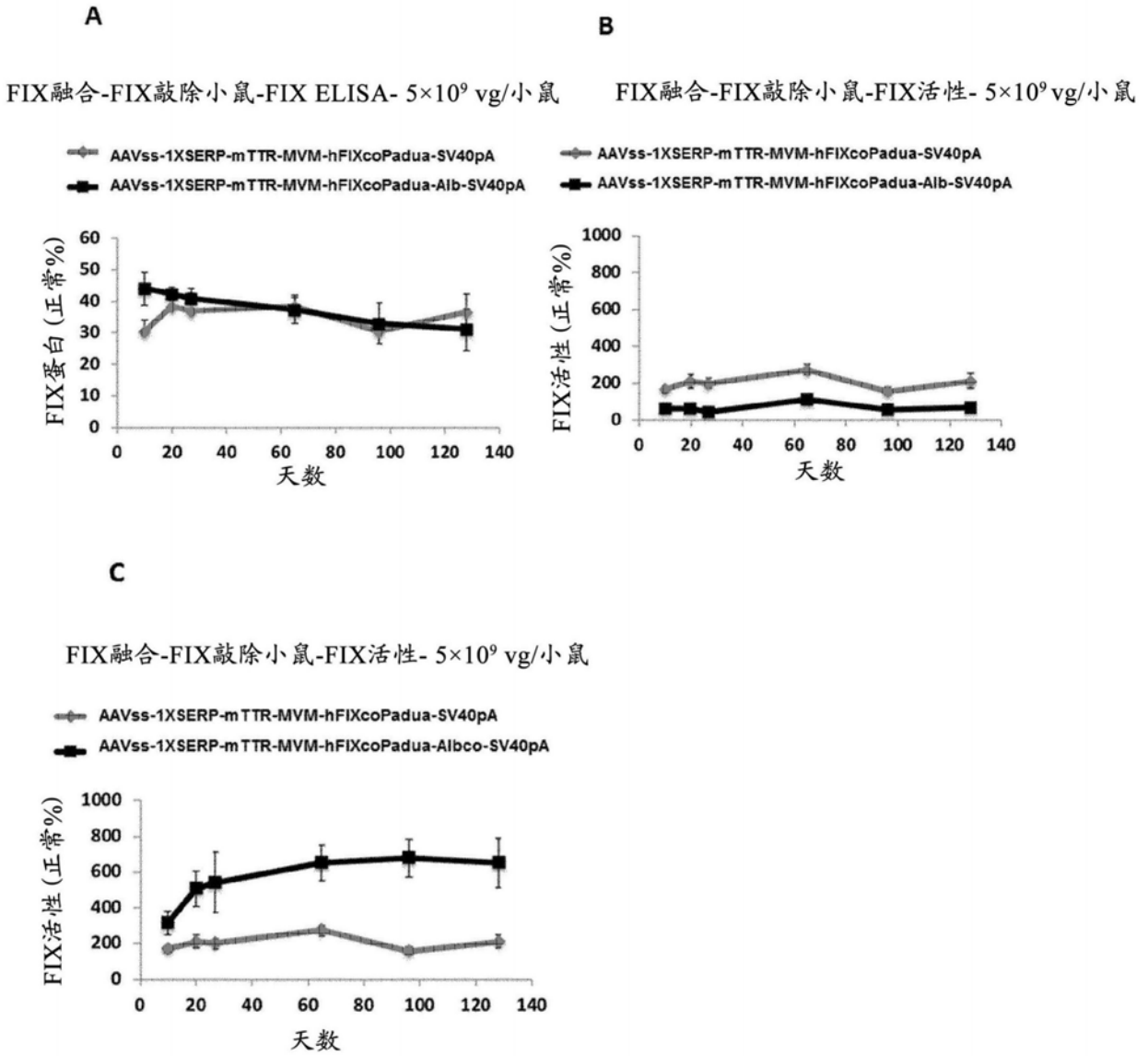


图9

A

FIX ELISA-C57BL/6小鼠 -1×10^9 vg/小鼠

- ◆ AAVss-1XSERP-mTTR-MVM-hFIXco-SV40pA
- AAVss-1XSERP-mTTR-MVM-hFIXco-Albco-SV40pA
- ▲ AAVss-3XSERP-mTTR-MVM-hFIXco-Albco-SV40pA

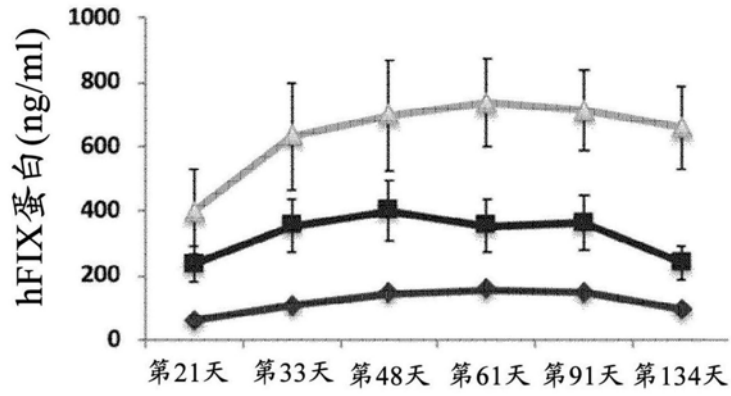


图10

B 肝脏中的mRNA表达-C57BL/6小鼠 -1 × 10⁹ vg/小鼠

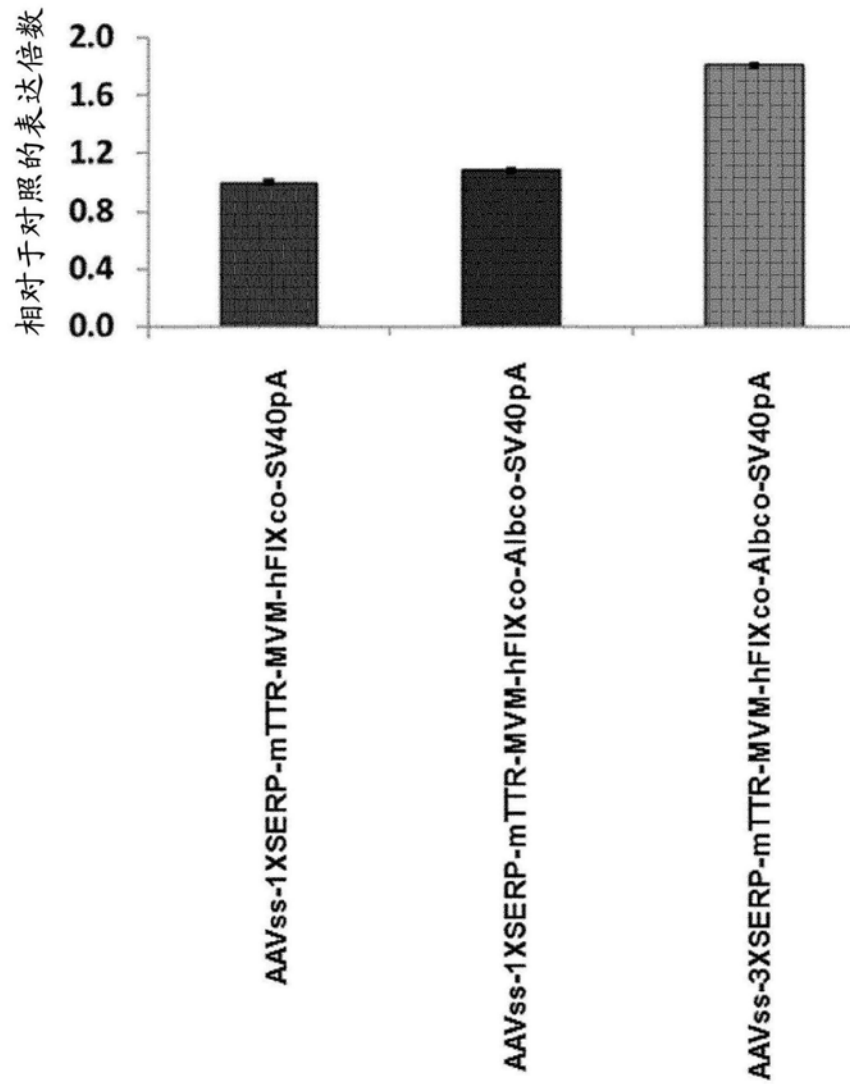
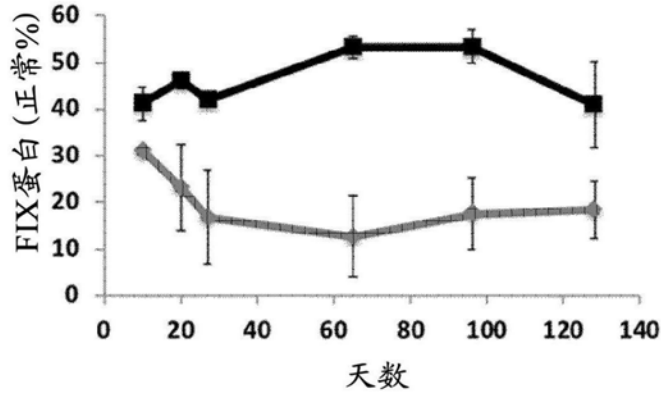


图10 (续)

A

FIX融合-FIX敲除小鼠-FIX ELISA - 5×10^9 vg/小鼠

- ◆ AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA
- AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA



B

FIX融合-FIX敲除小鼠-FIX活性- 5×10^9 vg/小鼠

- ◆ AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Alb-SV40pA
- AAVss-3XSERP-mTTR-MVM-hFIXcoPadua-Albco-SV40pA

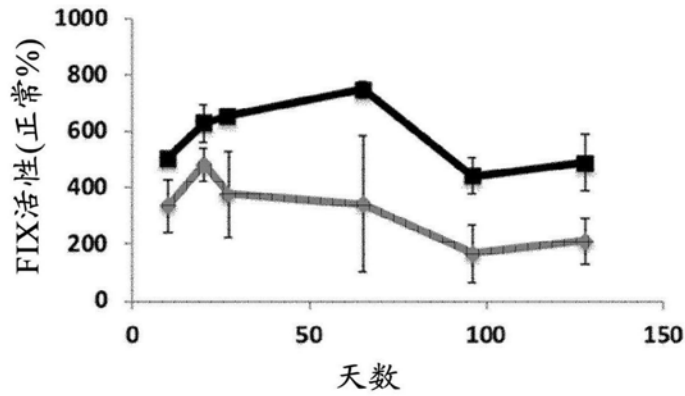


图11

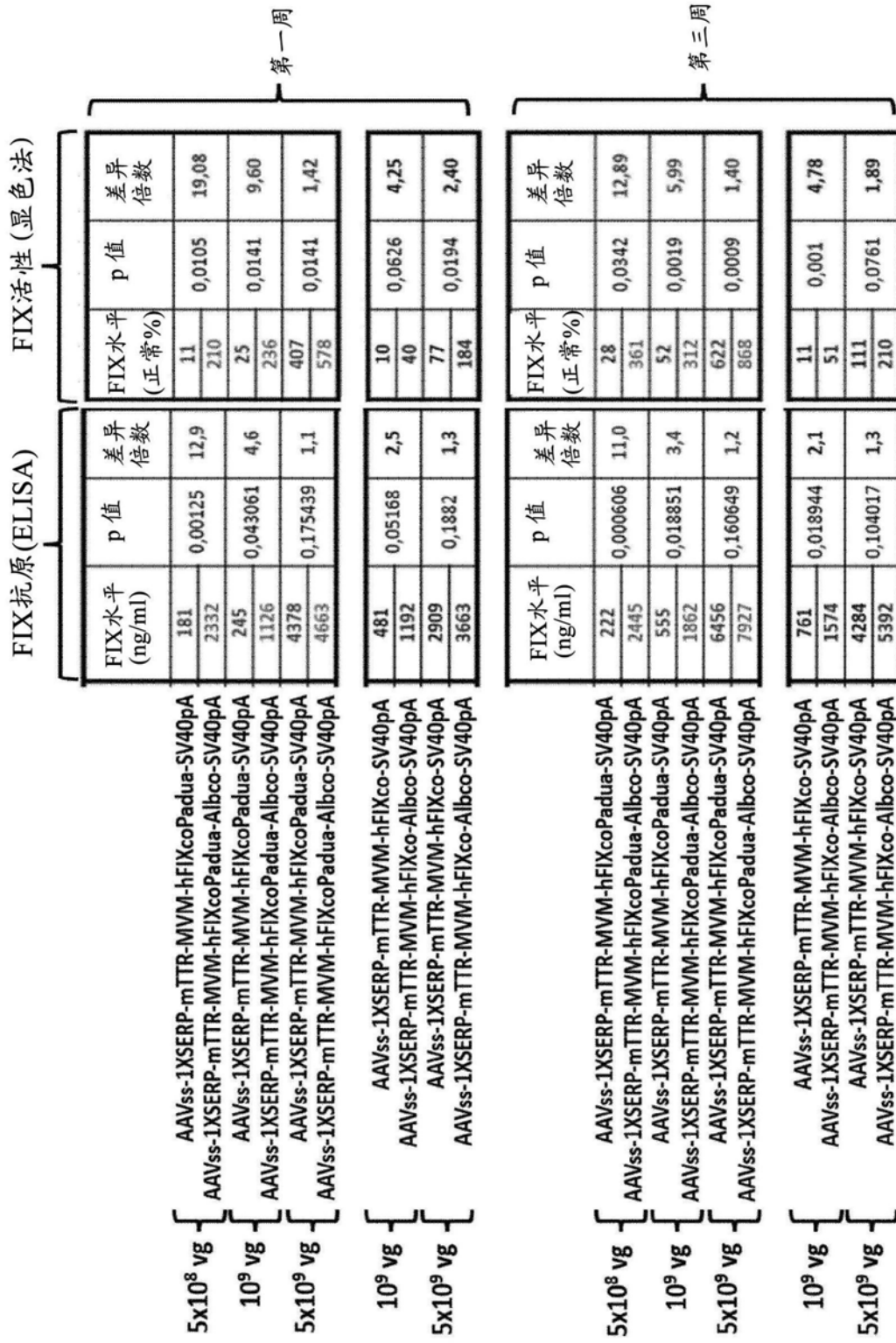
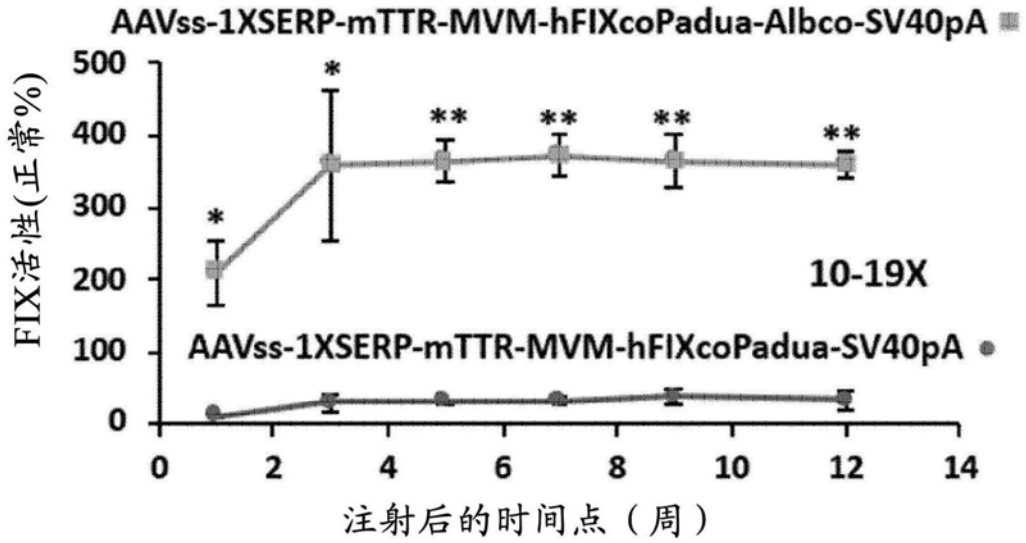


图12

A

剂量- 5×10^8 vg/小鼠



B

剂量- 5×10^9 vg/小鼠

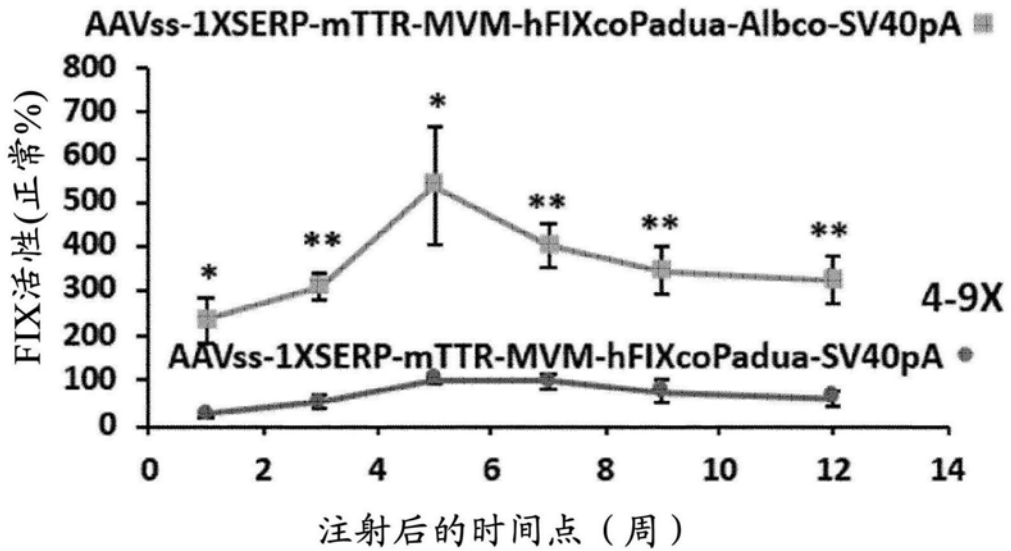
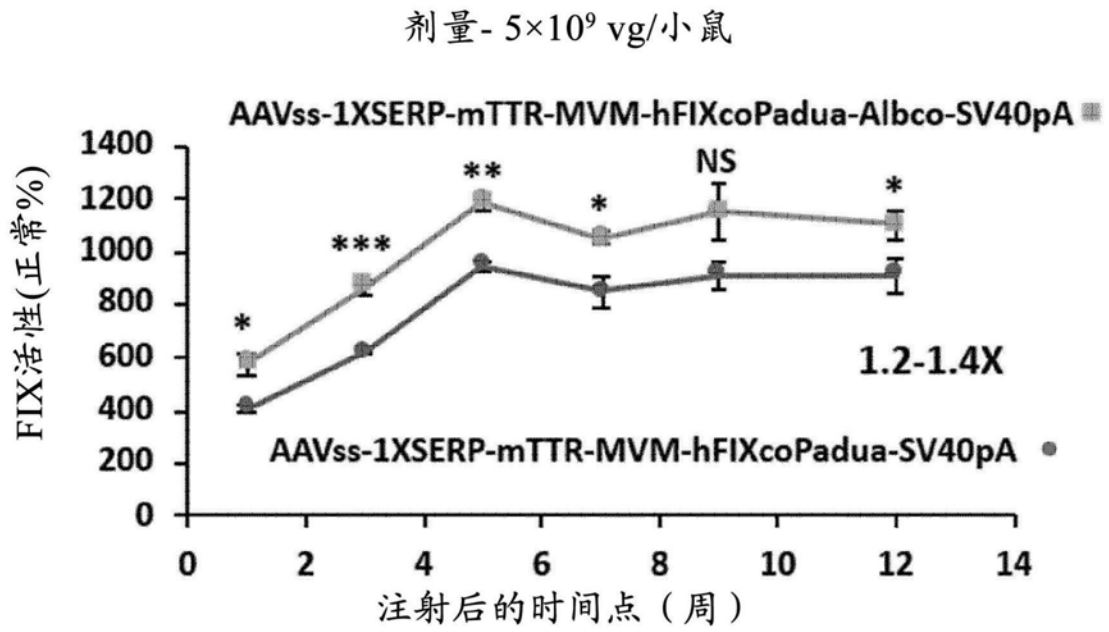


图13

C



D

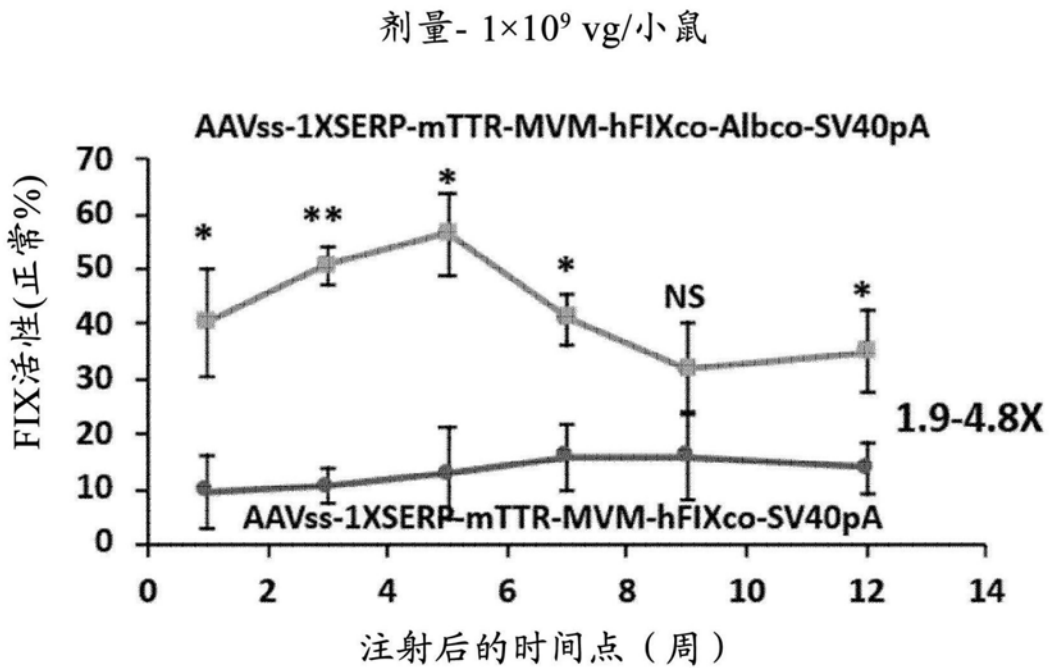


图13(续)

E

剂量- 5×10^9 vg/小鼠

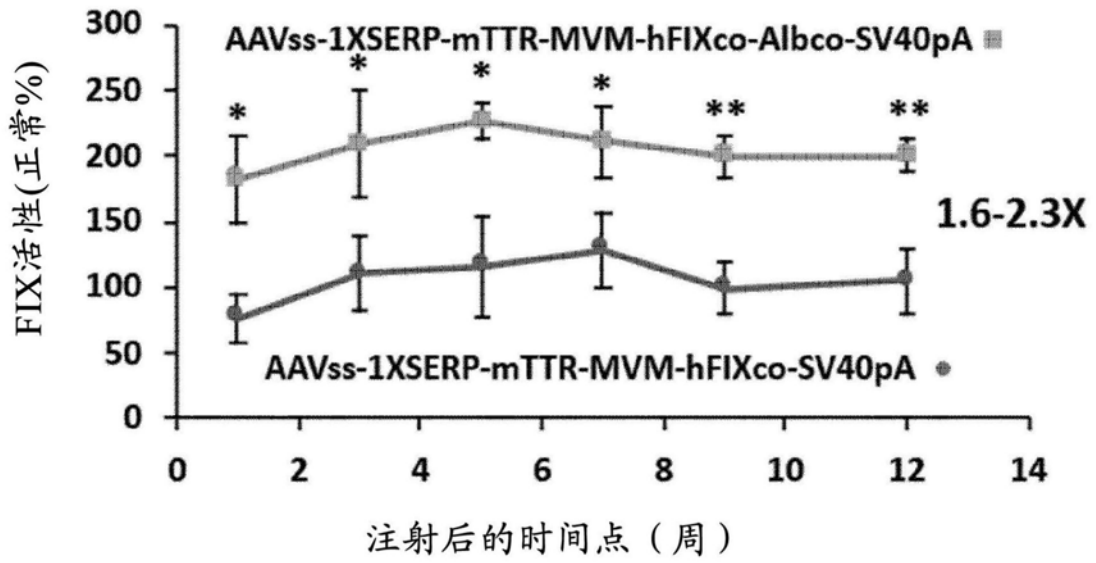


图13(续)