(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117881696 A
(43)申请公布日 2024.04.12

(21)申请号 202280057094.9
(22)申请日 2022.08.22
(30)优先权数据
63/236,215 2021.08.23 US
(85)PCT国际申请进入国家阶段日
2024.02.21
(86)PCT国际申请的申请数据
PCT/US2022/075280 2022.08.22
(87)PCT国际申请的公布数据
WO2023/028455 EN 2023.03.02
(71)申请人 比奥维拉迪维治疗股份有限公司
地址 美国马萨诸塞州
(72)发明人 A·马戈迪亚 C·缪勒 刘童瑶
(74)专利代理机构 北京坤瑞律师事务所 11494
专利代理师 封新琴
(51)Int.Cl.
C07K 14/755(2006.01)
C12N 15/86(2006.01)

权利要求书5页 说明书81页
序列表（电子公布） 附图57页

(54)发明名称

具有反向末端重复序列的封闭末端DNA产生

(57)摘要

本公开文本提供了核酸分子,其包含第一反向末端重复序列(ITR)、第二ITR和编码靶序列的基因盒。在一些实施方案中,第一ITR和/或第二ITR是人博卡病毒的ITR。还公开了在基因疗法应用中使用所述核酸分子的方法。

CN 117881696 A

1.一种核酸分子,所述核酸分子包含位于包含异源多核苷酸序列的基因盒两侧的第一反向末端重复序列(ITR)和第二ITR,其中:

所述第一ITR包含与SEQ ID NO:1至少约75%相同的多核苷酸序列,并且所述第二ITR包含与SEQ ID NO:2至少约75%相同的多核苷酸序列。

2.根据权利要求1所述的核酸分子,其中所述第一ITR包含与SEQ ID NO:1至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%相同的多核苷酸序列。

3.根据权利要求1或2所述的核酸分子,其中所述第一ITR包含SEQ ID NO:1中所示的多核苷酸序列。

4.根据权利要求1-3中任一项所述的核酸分子,其中所述第二ITR包含与SEQ ID NO:2至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%相同的多核苷酸序列。

5.根据权利要求1-4中任一项所述的核酸分子,其中所述第二ITR包含SEQ ID NO:2中所示的多核苷酸序列。

6.根据权利要求5所述的核酸分子,其中所述第一ITR包含SEQ ID NO:1中所示的多核苷酸序列,并且所述第二ITR包含SEQ ID NO:2中所示的多核苷酸序列。

7.根据权利要求1-6中任一项所述的核酸分子,其还包含启动子。

8.根据权利要求7所述的核酸分子,其中所述启动子是组织特异性启动子。

9.根据权利要求7或8所述的核酸分子,其中所述启动子驱动所述异源多核苷酸序列在器官或组织中的表达,其中所述器官或组织包括肌肉、中枢神经系统(CNS)、眼睛、肝脏、心脏、肾脏、胰腺、肺、皮肤、膀胱、泌尿道、脾脏、髓系细胞谱系和淋巴样细胞谱系或其任何组合。

10.根据权利要求7-9中任一项所述的核酸分子,其中所述启动子驱动所述异源多核苷酸序列在以下细胞中的表达:肝细胞、上皮细胞、内皮细胞、心肌细胞、骨骼肌细胞、窦状细胞、传入神经元、传出神经元、中间神经元、神经胶质细胞、星形胶质细胞、少突胶质细胞、小神经胶质细胞、室管膜细胞、肺上皮细胞、许旺细胞、卫星细胞、感光细胞、视网膜神经节细胞、T细胞、B细胞、NK细胞、巨噬细胞、树突细胞或其任何组合。

11.根据权利要求7-10中任一项所述的核酸分子,其中所述启动子定位于所述异源多核苷酸序列的5'。

12.根据权利要求7-11中任一项所述的核酸分子,其中所述启动子是小鼠甲状腺素转运蛋白启动子(mTTR)、天然人因子VIII启动子、人α-1-抗胰蛋白酶启动子(hAAT)、人白蛋白最小启动子、小鼠白蛋白启动子、三重四脯氨酸(TTP)启动子、CASI启动子、CAG启动子、巨细胞病毒(CMV)启动子、α1-抗胰蛋白酶(AAT)启动子、肌肉肌酸激酶(MCK)启动子、肌球蛋白重链α(αMHC)启动子、肌红蛋白(MB)启动子、结蛋白(DES)启动子、SPc5-12启动子、2R5Sc5-12启动子、dMCK启动子、tMCK启动子、磷酸甘油酸激酶(PGK)启动子或A1AT启动子。

13.根据权利要求1至12中任一项所述的核酸分子,其中所述异源多核苷酸序列还包含内含子序列。

14.根据权利要求13所述的核酸分子,其中所述内含子序列定位于所述异源多核苷酸序列的5'。

15.根据权利要求13或14所述的核酸分子,其中所述内含子序列定位于所述启动子的3'。

16.根据权利要求13-15中任一项所述的核酸分子,其中所述内含子序列包含合成内含子序列。

17.根据权利要求1至16中任一项所述的核酸分子,其中所述基因盒还包含转录后调节元件。

18.根据权利要求17所述的核酸分子,其中所述转录后调节元件定位于所述异源多核苷酸序列的3'。

19.根据权利要求17或18所述的核酸分子,其中所述转录后调节元件包含土拨鼠肝炎病毒调节元件(WPRE)、微小RNA结合位点、DNA核靶向序列、TLR9抑制序列或其任何突变。

20.根据权利要求1至19中任一项所述的核酸分子,其中所述基因盒还包含3'UTR多聚(A)尾序列。

21.根据权利要求20所述的核酸分子,其中所述3'UTR多聚(A)尾序列选自bGH多聚(A)、肌动蛋白多聚(A)、血红蛋白多聚(A)及其任何组合。

22.根据权利要求1至21中任一项所述的核酸分子,其中所述基因盒还包含增强子序列。

23.根据权利要求22所述的核酸分子,其中所述增强子序列定位于所述第一ITR与所述第二ITR之间。

24.根据权利要求1至23中任一项所述的核酸分子,其中所述核酸分子从5'至3'包含:所述第一ITR、所述基因盒和所述第二ITR,其中所述基因盒包含组织特异性启动子序列、内含子序列、异源多核苷酸序列、转录后调节元件和3'UTR多聚(A)尾序列。

25.根据权利要求24所述的核酸分子,其中所述基因盒从5'至3'包含:组织特异性启动子序列、内含子序列、所述异源多核苷酸序列、转录后调节元件和3'UTR多聚(A)尾序列。

26.根据权利要求25所述的核酸分子,其中所述基因盒包含SEQ ID NO:3的核苷酸序列。

27.根据权利要求1至26中任一项所述的核酸分子,其中所述基因盒是单链核酸。

28.根据权利要求1至26中任一项所述的核酸分子,其中所述基因盒是双链核酸。

29.根据权利要求1至28中任一项所述的核酸分子,其中所述异源多核苷酸序列编码治疗性蛋白质。

30.根据权利要求1至29中任一项所述的核酸分子,其中所述异源多核苷酸序列编码凝血因子、生长因子、激素、细胞因子、抗体、其片段或其任何组合。

31.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码凝血因子。

32.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码生长因子。

33.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码激素。

34.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码细胞因子。

35.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码抗体或其片段。

36.根据权利要求1至30中任一项所述的核酸分子,其中所述异源多核苷酸序列编码X连锁抗肌萎缩蛋白、MTM1(肌微管素)、酪氨酸羟化酶、AADC、环水解酶、SMN1、FXN(共济蛋白)、GUCY2D、RS1、CFH、HTRA、ARMS、CFB/CC2、CNGA/CNGB、Prf65、ARSA、PSAP、IDUA(MPS I)、IDS(MPS II)、PAH、GAA(酸性α-葡糖苷酶)、GALT、OTC、CMD1A、LAMA2或其任何组合。

37.根据权利要求1至28中任一项所述的核酸分子,其中所述异源多核苷酸序列编码微小RNA(miRNA)。

38.根据权利要求37所述的核酸分子,其中所述miRNA下调包含以下的靶基因的表达:SOD1、HTT、RHO、CD38或其任何组合。

39.根据权利要求31所述的核酸分子,其中所述凝血因子包括因子I(FI)、因子II(FII)、因子III(FIII)、因子IV(FIV)、因子V(FV)、因子VI(FVI)、因子VII(FVII)、因子VIII(FVIII)、因子IX(FIX)、因子X(FX)、因子XI(FXI)、因子XII(FXII)、因子XIII(FXIII)、血管性血友病因子(VWF)、前激肽释放酶、高分子量激肽原、纤连蛋白、抗凝血酶III、肝素辅因子II、蛋白质C、蛋白质S、蛋白质Z、蛋白质Z相关蛋白酶抑制剂(ZPI)、纤溶酶原、α2-抗纤溶酶、组织纤溶酶原激活物(tPA)、尿激酶、纤溶酶原激活物抑制剂-1(PAI-1)、纤溶酶原激活物抑制剂-2(PAI-2)或其任何组合。

40.根据权利要求1至39中任一项所述的核酸分子,其中所述异源多核苷酸序列经密码子优化。

41.根据权利要求40所述的核酸分子,其中所述异源多核苷酸序列经密码子优化以用于在人体内表达。

42.根据权利要求1至41中任一项所述的核酸分子,其中将所述核酸分子与递送剂一起配制。

43.根据权利要求42所述的核酸分子,其中所述递送剂包括脂质纳米颗粒(LNP)。

44.根据权利要求42所述的核酸分子,其中所述递送剂包括脂质体、非脂质聚合物分子、内体或其任何组合。

45.根据权利要求1至44中任一项所述的核酸分子,其中所述核酸分子被配制用于静脉内、经皮、皮内、神经内、眼内、鞘内、皮下、肺部或口服施用或其任何组合。

46.根据权利要求45所述的核酸分子,其中所述核酸分子被配制用于静脉内施用。

47.根据权利要求1至44中任一项所述的核酸分子,其中所述核酸分子被配制用于通过原位注射施用。

48.根据权利要求1至44中任一项所述的核酸分子,其中所述核酸分子被配制用于通过吸入施用。

49.一种载体,所述载体包含根据权利要求1至41中任一项所述的核酸分子。

50.一种宿主细胞,所述宿主细胞包含根据权利要求1至41中任一项所述的核酸分子或根据权利要求49所述的载体。

51.根据权利要求50所述的宿主细胞,其中所述宿主细胞是昆虫细胞。

52.一种药物组合物,所述药物组合物包含根据权利要求1至41中任一项所述的核酸。

53.一种药物组合物,所述药物组合物包含根据权利要求49所述的载体和药学上可接

受的赋形剂。

54.一种药物组合物,所述药物组合物包含根据权利要求50所述的宿主细胞和药学上可接受的赋形剂。

55.一种试剂盒,所述试剂盒包含根据权利要求1至48中任一项所述的核酸分子和用于向有需要的受试者施用所述核酸分子的说明书。

56.一种杆状病毒系统,所述杆状病毒系统用于产生根据权利要求1至41中任一项所述的核酸分子。

57.根据权利要求56所述的杆状病毒系统,其中根据权利要求1-41中任一项所述的核酸分子是在昆虫细胞中产生的。

58.根据权利要求56或57所述的杆状病毒系统,其还包含重组杆粒,其中所述重组杆粒包含变体VP80基因,使得所述杆粒展现出其编码的蛋白质的降低的表达。

59.一种纳米颗粒递送系统,所述纳米颗粒递送系统包含根据权利要求1至48中任一项所述的核酸分子。

60.一种在有需要的受试者体内表达异源多核苷酸序列的方法,所述方法包括向所述受试者施用根据权利要求1至48中任一项所述的核酸分子、根据权利要求49所述的载体或根据权利要求52-54中任一项所述的药物组合物。

61.一种治疗有需要的受试者的疾病或障碍的方法,所述方法包括向所述受试者施用根据权利要求1至48中任一项所述的核酸分子、根据权利要求49所述的载体或根据权利要求52-54中任一项所述的药物组合物。

62.一种治疗出血障碍的方法,其包括:

向有需要的受试者施用核酸分子,

其中所述核酸分子包含SEQ ID NO:3的核苷酸序列。

63.根据权利要求61或62所述的方法,其中所述障碍是血友病A。

64.根据权利要求60-63中任一项所述的方法,其中将所述核酸分子通过静脉内、经皮、皮内、神经内、眼内、鞘内、皮下、口服、肺部或其任何组合施用。

65.根据权利要求64所述的方法,其中将所述核酸分子通过静脉内施用。

66.根据权利要求60-63中任一项所述的方法,其中将所述核酸分子通过原位注射施用。

67.根据权利要求60-63中任一项所述的方法,其中将所述核酸分子通过吸入施用。

68.根据权利要求60-66中任一项所述的方法,其中所述受试者是哺乳动物。

69.根据权利要求67所述的方法,其中所述哺乳动物是人。

70.一种重组杆粒,所述重组杆粒包含:

编码HBoV1 Rep的序列,其中插入的HBoV1 Rep序列破坏报告基因或其功能部分的阅读框;以及

包含核酸分子的异源序列,所述核酸分子包含SEQ ID NO:3、9、14、33或35的核苷酸序列。

71.重组杆粒组,所述重组杆粒组包含第一杆粒和第二杆粒,

其中所述第一杆粒包含编码Rep的序列,其中插入的Rep破坏报告基因或其功能部分的阅读框;以及

其中所述第二杆粒包含异源序列,所述异源序列包含含有SEQ ID NO:3、9、14、33或35的核苷酸序列的核酸分子。

72.一种包含SEQ ID NO:3、9、14、33或35的核酸序列的稳定细胞系,其中所述核酸序列稳定地整合到所述稳定细胞系的基因组中。

73.一种产生封闭末端DNA(ceDNA)分子的方法,其包括用包含根据权利要求69所述的杆粒的重组杆状病毒感染昆虫细胞。

74.一种产生封闭末端DNA(ceDNA)分子的方法,其包括用包含根据权利要求70所述的重组杆粒组的重组杆状病毒感染昆虫细胞。

75.一种产生封闭末端DNA(ceDNA)分子的方法,其包括将编码Rep蛋白的杆状病毒引入根据权利要求71所述的稳定细胞系中。

具有反向末端重复序列的封闭末端DNA产生

相关申请

[0001] 本申请要求2021年8月23日提交的美国临时专利申请号63/236,215的优先权,其公开内容通过引用以其整体并入本文。

电子提交的序列表的引用

[0002] 以XML格式电子提交的序列表(名称:SA9-481_SeqListing.xml;大小:117,285字节;以及创建日期:2022年8月22日)的内容通过引用以其整体并入本文。

背景技术

[0003] 基因疗法提供了治疗多种疾病的持久手段的可能性。过去,许多基因疗法通常依赖于使用病毒载体。有多种可以选择用于这个目的的病毒剂,每种病毒剂具有不同特性,这使得它们更适合于或更不适合于基因疗法。然而,一些病毒载体的不期望的特性已经导致了临床安全性问题并且限制了它们的治疗用途。

[0004] 腺相关病毒(AAV)是一种常见的基因疗法载体,但是它并非没有缺点。AAV基因组的编码序列侧接反向末端重复序列(ITR),所述反向末端重复序列是病毒复制和包装以及转基因表达所需。AAV ITR的T形发夹结构易于由宿主细胞蛋白质结合,抑制AAV载体中的转基因表达。需要提供靶序列的有效且持久的表达,同时避免现有AAV载体技术的局限性。

发明内容

[0005] 本文公开了核酸分子及其用途,所述核酸分子包含位于包含异源多核苷酸序列的基因盒两侧的第一反向末端重复序列(ITR)和/或第二ITR。

[0006] 在一方面,本文提供了包含位于包含异源多核苷酸序列的基因盒两侧的第一ITR和第二ITR的核酸分子,其中所述第一ITR和第二ITR是博卡病毒ITR或其片段/衍生物(例如人博卡病毒1ITR)。在另一方面,本文提供了包含第一ITR和第二ITR的核酸分子,其中第一ITR包含与SEQ ID NO:1至少约75%相同的多核苷酸序列,并且第二ITR包含与SEQ ID NO:2至少约75%相同的多核苷酸序列。

[0007] 在一些实施方案中,第一ITR包含与SEQ ID NO:1至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%相同的多核苷酸序列。在一些实施方案中,第一ITR包含SEQ ID NO:1中所示的多核苷酸序列。在一些实施方案中,第一ITR包含与SEQ ID NO:1至少约50%相同的多核苷酸序列。

[0008] 在一些实施方案中,第二ITR包含与SEQ ID NO:2至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%相同的多核苷酸序列。在一些实施方案中,第二ITR包含SEQ ID NO:2中所示的多核苷酸序列。在一些实施方案中,第一ITR包含与SEQ ID NO:2至少约50%相同的多核苷酸序列。

[0009] 在一些实施方案中,第一ITR包含SEQ ID NO:1中所示的多核苷酸序列,并且第二ITR包含SEQ ID NO:2中所示的多核苷酸序列。

[0010] 在一些实施方案中,核酸分子还包含基因盒,所述基因盒包含异源多核苷酸序列

和至少一个表达控制序列,如启动子、增强子、内含子、转录终止信号或转录后调节元件。

[0011] 在一些实施方案中,基因盒还包含启动子。在一些实施方案中,所述启动子是组织特异性启动子。在一些实施方案中,所述启动子驱动所述异源多核苷酸序列在器官中的表达,其中所述器官包括肌肉、中枢神经系统(CNS)、眼睛、肝脏、心脏、肾脏、胰腺、肺、皮肤、膀胱、泌尿道、脾脏、髓系细胞谱系和淋巴样细胞谱系或其任何组合。在一些实施方案中,所述启动子驱动所述异源多核苷酸序列在以下细胞中的表达:肝细胞、上皮细胞、内皮细胞、心肌细胞、骨骼肌细胞、窦状细胞、传入神经元、传出神经元、中间神经元、神经胶质细胞、星形胶质细胞、少突胶质细胞、小神经胶质细胞、室管膜细胞、肺上皮细胞、许旺细胞、卫星细胞、感光细胞、视网膜神经节细胞、T细胞、B细胞、NK细胞、巨噬细胞、树突细胞或其任何组合。在一些实施方案中,所述启动子定位于所述异源多核苷酸序列的5'。在一些实施方案中,所述启动子是小鼠甲状腺素转运蛋白启动子(mTTR)、天然人因子VIII启动子、人α-1-抗胰蛋白酶启动子(hAAT)、人白蛋白最小启动子、小鼠白蛋白启动子、三重四脯氨酸(TTP)启动子、CASI启动子、CAG启动子、巨细胞病毒(CMV)启动子、α1-抗胰蛋白酶(AAT)、肌肉肌酸激酶(MCK)、肌球蛋白重链α(αMHC)、肌红蛋白(MB)、结蛋白(DES)、SPc5-12、2R5Sc5-12、dMCK、tMCK或磷酸甘油酸激酶(PGK)启动子。

[0012] 在一些实施方案中,所述基因盒还包含内含子序列。在一些实施方案中,所述内含子序列定位于所述异源多核苷酸序列的5'。在一些实施方案中,所述内含子序列定位于所述启动子的3'。在一些实施方案中,所述内含子序列包含合成内含子序列。

[0013] 在一些实施方案中,所述基因盒还包含转录后调节元件。在一些实施方案中,所述调节元件定位于异源多核苷酸序列的3'。在一些实施方案中,所述调节元件包含突变的土拨鼠肝炎病毒转录后调节元件(WPRE)、微小RNA结合位点、DNA核靶向序列或其任何组合。

[0014] 在一些实施方案中,所述基因盒还包含3'UTR多聚(A)尾序列。在一些实施方案中,所述3'UTR多聚(A)尾序列选自bGH多聚(A)、肌动蛋白多聚(A)、血红蛋白多聚(A)及其任何组合。

[0015] 在一些实施方案中,所述基因盒还包含增强子序列。在一些实施例中,所述增强子序列定位于第一ITR与第二ITR之间。

[0016] 在一些实施方案中,所述核酸分子从5'至3'包含:所述第一ITR、所述基因盒和所述第二ITR;其中所述基因盒包含组织特异性启动子序列、内含子序列、所述异源多核苷酸序列、转录后调节元件和3'UTR多聚(A)尾序列。

[0017] 在一些实施方案中,所述基因盒从5'至3'包含:组织特异性启动子序列、内含子序列、所述异源多核苷酸序列、转录后调节元件和3'UTR多聚(A)尾序列。

[0018] 在一些实施方案中,所述基因盒是单链核酸。在一些实施方案中,所述基因盒是双链核酸。

[0019] 在一些实施方案中,所述异源多核苷酸序列编码治疗性蛋白质。

[0020] 在一些实施方案中,所述异源多核苷酸序列编码凝血因子、生长因子、激素、细胞因子、抗体、其片段或其任何组合。在一些实施方案中,所述异源多核苷酸序列编码凝血因子。在一些实施方案中,所述异源多核苷酸序列编码生长因子。在一些实施方案中,所述异源多核苷酸序列编码激素。在一些实施方案中,所述异源多核苷酸序列编码细胞因子。

[0021] 在一些实施方案中,所述异源多核苷酸序列编码FVIII蛋白。

[0022] 在一些实施方案中,所述异源多核苷酸序列编码X连锁抗肌萎缩蛋白、MTM1(肌微管素)、酪氨酸羟化酶、AADC、环水解酶、SMN1、FXN(共济蛋白(frataxin))、GUCY2D、RS1、CFH、HTRA、ARMS、CFB/CC2、CNGA/CNGB、Prf65、ARSA、PSAP、IDUA(MPS I)、IDS(MPS II)、PAH、GAA(酸性α-葡糖苷酶)、GALT、OTC、CMD1A、LAMA2或其任何组合。

[0023] 在一些实施方案中,所述异源多核苷酸序列编码微小RNA(miRNA)。在一些实施方案中,所述miRNA下调包括以下的靶基因的表达:SOD1、HTT、RHO、CD38或其任何组合。

[0024] 在一些实施方案中,所述异源多核苷酸序列编码凝血因子,其中所述凝血因子包含因子I(FI)、因子II(FII)、因子III(FIII)、因子IV(FIV)、因子V(FV)、因子VI(FVI)、因子VII(FVII)、因子VIII(FVIII)、因子IX(FIX)、因子X(FX)、因子XI(FXI)、因子XII(FXII)、因子XIII(FXIII)、血管性血友病因子(VWF)、前激肽释放酶、高分子量激肽原、纤连蛋白、抗凝血酶III、肝素辅因子II、蛋白质C、蛋白质S、蛋白质Z、蛋白质Z相关蛋白酶抑制剂(ZPI)、纤溶酶原、α2-抗纤溶酶、组织纤溶酶原激活物(tPA)、尿激酶、纤溶酶原激活物抑制剂-1(PAI-1)、纤溶酶原激活物抑制剂-2(PAI2)或其任何组合。

[0025] 在一些实施方案中,所述异源多核苷酸序列经密码子优化。在一些实施方案中,所述异源多核苷酸序列经密码子优化以用于在人体内表达。

[0026] 在一些实施方案中,将所述核酸分子与递送剂一起配制。在一些实施方案中,所述递送剂包括脂质纳米颗粒。在一些实施方案中,所述脂质纳米颗粒是可电离的。在一些实施方案中,所述递送剂包括脂质体、非脂质聚合物分子、内体或其任何组合。

[0027] 在一些实施方案中,所述核酸分子被配制用于静脉内、经皮、皮内、皮下、肺部、神经内、眼内、鞘内、口服施用或其任何组合。在一些实施方案中,所述核酸分子被配制用于静脉内施用。在一些实施方案中,所述核酸分子被配制用于通过原位注射施用。在一些实施方案中,所述核酸分子被配制用于通过吸入施用。

[0028] 在另一方面,本文提供了一种载体,所述载体包含本文所述的核酸分子。

[0029] 在另一方面,本文提供了一种宿主细胞,所述宿主细胞包含本文所述的核酸分子或本文所述的载体。在一些实施方案中,宿主细胞是昆虫细胞。

[0030] 在另一方面,本文提供了一种药物组合物,所述药物组合物包含本文所述的核酸分子。

[0031] 在另一方面,本文提供了一种药物组合物,所述药物组合物包含本文所述的载体和药学上可接受的赋形剂。

[0032] 在另一方面,本文提供了一种药物组合物,所述药物组合物包含本文所述的宿主细胞和药学上可接受的赋形剂。

[0033] 在另一方面,本文提供了一种试剂盒,所述试剂盒包含本文所述的核酸分子和用于向有需要的受试者施用所述核酸分子的说明书。

[0034] 在另一方面,本文提供了一种杆状病毒系统,所述杆状病毒系统用于产生本文所述的核酸分子。

[0035] 在一些实施方案中,所述核酸分子是在昆虫细胞中产生的。

[0036] 在另一方面,本文提供了一种纳米颗粒递送系统,所述纳米颗粒递送系统包含本文所述的核酸分子。

[0037] 在另一方面,本文提供了一种在有需要的受试者体内表达异源多核苷酸序列的方

法,所述方法包括向所述受试者施用本文所述的核酸分子、本文所述的载体或本文所述的药物组合物。

[0038] 在另一方面,本文提供了一种治疗有需要的受试者的疾病或障碍的方法,所述方法包括向所述受试者施用本文所述的核酸分子、本文所述的载体或本文所述的药物组合物。

[0039] 在一些实施方案中,将所述核酸分子通过静脉内、经皮、皮内、皮下、口服、肺部、神经内、眼内、鞘内或其任何组合施用。在一些实施方案中,所述核酸分子是静脉内施用的。在一些实施方案中,所述核酸分子是通过原位注射施用。在一些实施方案中,所述核酸分子是通过吸入施用。

[0040] 在一些实施方案中,所述受试者是哺乳动物。在一些实施方案中,所述受试者是人。

附图说明

[0041] 图1A-图1C是根据本发明的一个实施方案的用于在杆状病毒系统中产生ceDNA的方法的示意图。图1A示出了单BAC方法的示意图,其中将在不同基因座处编码FVIIIXTEN和Rep基因的单一重组BEV用于在Sf9细胞中进行感染以产生ceDNA。图1B示出了双BAC方法的示意图,其中将Sf9细胞用编码FVIIIXTEN和/或Rep基因的重组BEV共感染以产生ceDNA。图1C示出了稳定细胞系方法的示意图,其中FVIIIXTEN表达盒稳定地整合到Sf9细胞基因组中并且通过感染编码Rep基因的重组BEV以产生ceDNA而被挽救。

[0042] 图2A-图2B是人FVIIIXTEN表达构建体的示意图。图2A示出了根据本发明的一个实施方案的表达构建体的示意性线性图谱,所述表达构建体包含与XTEN 144肽融合的B结构域缺失(BDD)的经密码子优化的人因子VIII(coFVIII)(FVIIIXTEN),其在肝特异性修饰小鼠甲状腺素转运蛋白(mTTR)启动子(mTTR482)和增强子元件(A1MB2)、杂合合成内含子(嵌合内含子)、土拨鼠转录后调节元件(WPRE)和牛生长激素多腺苷酸化(bGHpA)信号的调节下。FVIIIXTEN表达盒两侧为人博卡病毒1型(HBoV1)野生型(WT)ITR。(SEQ ID NO:1和SEQ ID NO:2)。图2B示出了通过将FVIIIXTEN表达盒(SEQ ID NO:3)插入pFastBac1载体(Invitrogen)中而制备的根据本发明的一个实施方案的Tn7转移载体的示意性图谱。

[0043] 图3A-图3C是根据本发明的实施方案的复制(Rep)基因表达构建体的示意图。图3A示出了合成DNA的线性示意性图谱,所述合成DNA编码在AcMNPV多角体蛋白启动子下的Sf密码子优化的HBoV1 NS1基因,随后是SV40多腺苷酸化信号(SV40 PAS)。图3B示出了通过将HBoV1 NS1合成DNA(SEQ ID NO:4)插入pFastBac1载体(Invitrogen)中而制备的根据本发明实施方案的Tn7转移载体的示意性图谱。图3C示出了通过将HBoV1 NS1合成DNA(SEQ ID NO:4)插入Cre-LoxP供体载体中而制备的根据本发明实施方案的Cre-LoxP供体载体的示意性图谱,所述Cre-LoxP供体载体如在杆状病毒表达系统(美国专利申请号63/069,073)中所描述而产生,将所述专利申请出于引用以其整体并入本文。

[0044] 图4A-图4C是根据本发明的实施方案的复制(Rep)基因表达构建体的示意图。图4A示出了合成DNA的线性示意性图谱,所述合成DNA编码在AcMNPV即时早期1(pIE1)启动子下的Sf密码子优化的HBoV1 NS1基因,所述启动子之前是AcMNPV转录增强子hr5元件,所述基因之后是SV40多腺苷酸化信号(SV40 PAS)。图4B示出了通过将HBoV1NS1合成DNA(SEQ ID

NO:4)插入pFastBac1载体(Invitrogen)中而制备的根据本发明实施方案的Tn7转移载体的示意性图谱。图4C示出了通过将HBoV1 NS1合成DNA(SEQ ID NO:4)插入Cre-LoxP供体载体中而制备的根据本发明实施方案的Cre-LoxP供体载体的示意性图谱。

[0045] 图5A-图5D是根据本发明的实施方案的复制(Rep)基因表达构建体的示意图。图5A示出了合成DNA的线性示意性图谱,所述合成DNA编码在OpMNPV即时早期2(OpIE2)启动子下的Sf密码子优化的HBoV1 NS1基因,随后是SV40多腺苷酸化信号(SV40 PAS)。图5B示出了合成DNA的线性示意性图谱,所述合成DNA编码在AcMNPV即时早期1(pIE1)启动子下的Sf密码子优化的HBoV1 NS1基因,随后是SV40多腺苷酸化信号(SV40 PAS)。图5C示出了通过将HBoV1 NS1合成DNA(SEQ ID NO:4)插入pFastBac1载体(Invitrogen)中而制备的根据本发明实施方案的Tn7转移载体的示意性图谱。图5D示出了通过将在AcMNPV即时早期1(pIE1)启动子下的HBoV1 NS1合成DNA(SEQ ID NO:4)插入如美国专利申请号63/069,073所述产生的Cre-LoxP供体载体中而制备的根据本发明实施方案的Cre-LoxP供体载体的示意性图谱。

[0046] 图6A-图6B示出了编码具有HBoV1 ITR的人FVIIIXTEN的重组杆状病毒表达载体(BEV)的产生。图6A是编码具有HBoV1 ITR的人FVIIIXTEN的重组BIVVBac杆粒克隆(BIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7})的限制酶作图的琼脂糖凝胶电泳图像。图6B是编码如所指示的两侧为HBoV1 ITR的FVIIIXTEN表达盒(SEQ ID NO:3)的重组BEV(AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7})的示意图。

[0047] 图7A-图7F是包含人FVIIIXTEN和Rep基因表达盒的单BAC的示意图及其确认研究。图7A-图7C是通过外侧/内侧PCR引物(SEQ ID NO:5和SEQ ID NO:6)筛选的BIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}(图7A)、BIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1^{LoxP}(图7B)、和BIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP}(图7C)的重组杆粒克隆的琼脂糖凝胶电泳图像。如图7D-图7F中的红色箭头所指示。图7D示出了编码如所指示的在AcMNPV多角体蛋白(pPolh)启动子下的HBoV1 NS1和两侧为HBoV1 ITR的FVIIIXTEN表达盒的重组杆状病毒表达载体(BEV)的示意性图谱(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP})。图7E示出了编码如所指示的在AcMNPV即时早期1(pIE1)启动子下的HBoV1 NS1、和两侧为HBoV1 ITR的FVIIIXTEN表达盒的重组BEV的示意性图谱(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1^{LoxP})。图7F示出了编码如所指示的在AcMNPV即时早期1启动子(之前是AcMNPV转录增强子hr5元件,pHR5.IE1)下的HBoV1 NS1、和两侧为HBoV1 ITR的FVIIIXTEN表达盒的重组BEV的示意性图谱(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP})。

[0048] 图8A-图8C示出了根据本发明的一个实施方案使用单BAC方法产生人FVIIIXTEN ceDNA载体。图8A是使用重组BEV在Sf9细胞中产生FVIIIXTEN ceDNA载体的单BAC方法的示意图,所述重组BEV编码在AcMNPV多角体蛋白启动子下的HBoV1NS1基因和两侧为HBoV1 ITR的人FVIIIXTEN表达盒(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP})。图8B示出了AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} BEV的示意性图谱。图8C是从用(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP})BEV的滴定病毒原液(P2)感染的Sf9细胞分离的ceDNA载体的琼脂糖凝胶电泳图像。对应于FVIIIXTEN ceDNA(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指

示。

[0049] 图9A-图9D示出了包含编码HBoV1 NS1的序列的重组杆状病毒表达载体(BEV)的示意图及其确认研究。图9A是在AcMNPV多角体蛋白即时早期1(之前是AcMNPV转录增强子hr5元件)下或OpMNPV即时早期2启动子下的HBoV1.NS1的重组杆粒克隆的限制酶作图的琼脂糖凝胶电泳图像(分别为BIVVBac.Polh.HBoV1.NS1^{Tn7}、BIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}和BIVVBac.OpIE2.HBoV1.NS1^{Tn7})。图9B示出了AcBIVVBac.Polh.HBoV1.NS1^{Tn7}的示意性图谱。图9C示出了AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}的示意性图谱。图9D示出了AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}的示意性图谱。

[0050] 图10A-图10C示出了根据本发明的一个实施方案使用双BAC方法产生人FVIIIXTEN ceDNA载体。图10A是产生FVIIIXTEN ceDNA载体的双BAC方法的示意图,其中用重组BEV共感染Sf9细胞,所述重组BEV编码两侧为HBoV1 ITR的FVIIIXTEN表达盒(AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7})和/或编码在AcMNPV多角体蛋白启动子下的HBoV1 NS1基因(AcBIVVBac.Polh.HBoV1.NS1^{Tn7})。图10B示出了AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV的示意性图谱。图10C是从Sf9细胞分离的ceDNA载体的琼脂糖凝胶电泳图像,所述Sf9细胞用所指示的AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV以恒定比率的不同MOI或恒定MOI的不同比率共感染。对应于FVIIIXTEN ceDNA载体(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指示。

[0051] 图11A-图11C示出了根据本发明的一个实施方案使用双BAC方法产生人FVIIIXTEN ceDNA载体。图11A是产生FVIIIXTEN ceDNA载体的双BAC方法的示意图,其中用重组BEV共感染Sf9细胞,所述重组BEV编码两侧为HBoV1 ITR的FVIIIXTEN表达盒(AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7})和/或编码在AcMNPV多角体蛋白启动子下的HBoV1 NS1基因(AcBIVVBac.Polh.HBoV1.NS1^{Tn7})或在即时早期1启动子(之前是AcMNPV转录增强子hr5元件)下的HBoV1 NS1基因(AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7})。图11B示出了AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}BEV的示意性图谱。图11C是从Sf9细胞分离的ceDNA载体的琼脂糖凝胶电泳图像,所述Sf9细胞用AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7}(左图)或AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}(右图)BEV的不同MOI共感染。对应于FVIIIXTEN ceDNA载体(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指示。

[0052] 图12A-图12C示出了用于产生编码两侧为HBoV1 ITR的FVIIIXTEN表达盒的稳定细胞系的材料。图12A示出了质粒的示意性图谱,所述质粒编码在AcMNPV即时早期1(IE1)启动子下的新霉素抗性标记,所述启动子之前是AcMNPV转录增强子hr5元件,所述新霉素抗性标记之后是AcMNPV p10多腺苷酸化信号(P10 PAS)。图12B示出了质粒的示意性图谱,所述质粒编码在AcMNPV即时早期1(IE1)启动子下的增强的绿色荧光蛋白(eGFP)标记,所述启动子之前是AcMNPV转录增强子hr5元件,所述eGFP标记之后是AcMNPV p10多腺苷酸化信号(P10 PAS)。图12C示出了两侧为HBoV1 ITR的FVIIIXTEN表达盒的示意性图谱(SEQ ID NO:3),其被稳定地整合到Sf9细胞基因组中以产生稳定的细胞系。

[0053] 图13A-图13E示出了根据本发明的一个实施方案使用双BAC方法生产和纯化FVIIIXTEN ceDNA载体的工作流程。图13A示出了Sf9细胞扩增和持续时间(第0-2天)的示意

图,其中将细胞从小规模(0.5L)到大规模培养(1.5L)烧瓶依次放大以在无血清ESF921培养基中实现2.5至3.0x 10^6/mL的细胞密度。图13B示出了Sf9大型培养(1.5L)烧瓶感染和孵育持续时间(第2-6天)的示意图,其中在MOI分别为0.1和0.01噬斑形成单位(pfu)/细胞的情况下,用重组BEV共感染细胞,所述重组BEV编码两侧为HBoV1 ITR的FVIIIXTEN表达盒(AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7})和/或编码在AcMNPV多角体蛋白启动子下的HBoV1 NS1基因(AcBIVVBac.Polh.HBoV1.NS1^{Tn7})。图13C示出了质粒Giga Prep纯化试剂盒和琼脂糖凝胶电泳以及处理持续时间(第6-7天)的图像,其中每日测量感染细胞的细胞密度和活力,并且一旦细胞活力达到70%-80%,就通过低速离心沉淀细胞。通过PureLink™ HiPure Expi质粒Gigaprep试剂盒(Invitrogen)将FVIIIXTEN ceDNA载体从感染细胞沉淀中纯化,并且在琼脂糖凝胶电泳上运行等分试样以确定FVIIIXTEN ceDNA(ceDNA)、杆状病毒DNA(vDNA)和/或Sf9细胞基因组DNA(gDNA)的生产力。图13D示出了Bio-Rad Model 491Prep Cell和琼脂糖凝胶电泳以及处理持续时间(第7-12天)的图像,其中将Giga-prep纯化的DNA加载到Prep Cell中的制备型琼脂糖凝胶上以将FVIIIXTEN ceDNA(约8.5kb片段)从高分子量DNA中分离。在0.8%至1.2%琼脂糖凝胶上分析以70-80min间隔从制备型琼脂糖凝胶电泳中收集的洗脱级分,以确定FVIIIXTEN ceDNA的纯度。图13E示出了琼脂糖凝胶电泳的图像,其中将从Prep Cell收集的级分合并,并且用1/10体积的3M NaOAc(pH 5.5)和3体积的100%乙醇沉淀以获得纯化的FVIIIXTEN ceDNA。凝胶图像示出了与起始材料相比FVIIIXTEN ceDNA的纯度,其中箭头指示对应于FVIIIXTEN ceDNA载体(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带。

[0054] 图14A-图14B示出了通过Chromogenix Coatest®SP因子VIII生色测定测量的血浆FVIII活性水平的图示。图14A示出了在以不同间隔从hFVIIIR593C$^{+/+}$/HemA小鼠收集的血液样品中测量的血浆FVIII活性水平的图形曲线,所述小鼠经由流体动力学尾静脉注射被全身注射了1600或400μg/kg单链FVIIIXTEN HBoV1 ITRs DNA(ssDNA)。图14B示出了在以不同间隔从hFVIIIR593C$^{+/+}$/HemA小鼠收集的血液样品中测量的血浆FVIII活性水平的图形曲线,所述小鼠经由流体动力学尾静脉注射被全身注射了80、40或12μg/kg FVIIIXTEN HBoV1 ITRs ceDNA(ceDNA)。误差条表示标准偏差。

[0055] 图15A-图15C示出了根据本发明的一个实施方案用AcBIVVBac.Polh.HBoV1.NS1^{Tn7}杆粒DNA和VP80 sgRNA共转染的Sf9细胞的红色荧光(上部小图)或亮视野(下部小图)显微图像。图15A示出了用AcBIVVBac.Polh.HBoV1.NS1^{Tn7}杆粒DNA和单独Cas9共转染的细胞的显微图像。图15B示出了用AcBIVVBac.Polh.HBoV1.NS1^{Tn7}杆粒DNA和sgRNA.VP80.T1共转染的细胞的显微图像。图15C示出了用AcBIVVBac.Polh.HBoV1.NS1^{Tn7}杆粒DNA和sgRNA.VP80.T2共转染的细胞的显微图像。

[0056] 图16A-图16C示出了VP80KO BEV的产生。图16A和图16B示出了AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7}和AcBIVVBac.OpIE2.HBoV1.NS1ΔVP80^{Tn7}克隆BEV的TIDE分析,以确定由CRISPR/Cas9诱导的插入缺失(indel)。图16C是从Sf9细胞分离的FVIIIXTEN ceDNA载体的琼脂糖凝胶电泳图像,所述Sf9细胞用如所指示的AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7}或AcBIVVBac.OpIE2.HBoV1.NS1ΔVP80^{Tn7} BEV以不同MOI共感染。对应于FVIIIXTEN ceDNA载体(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指示。

[0057] 图17A-图17C示出了使用双BAC方法产生人FVIIIXTEN HBoV1 ceDNA。图17A是产生FVIIIXTEN ceDNA载体的双BAC方法的示意图。图17B示出了AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV的示意性图谱。图17C是从Sf9细胞分离的FVIIIXTEN ceDNA载体的琼脂糖凝胶电泳图像,对所述Sf9细胞在1.0、2.0、3.0、4.0或5.0MOI下用AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV共感染。对应于FVIIIXTEN ceDNA载体(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指示。

[0058] 图18A-图18D示出了使用单BAC方法产生人FVIIIXTEN HBoV1 ceDNA载体。图18A是在Sf9细胞中产生FVIIIXTEN ceDNA载体的单BAC方法的示意图。图18B示出了AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} BEV的示意性图谱。图18C是从在Sf9细胞中扩增到P2的克隆HBoV1单BAC BEV分离的ceDNA的琼脂糖凝胶电泳图像。图18D是从在0.1、0.2、0.3、0.4或0.5MOI下用HBoV1单BAC BEV感染的Sf9细胞分离的ceDNA的琼脂糖凝胶电泳图像。对应于FVIIIXTEN ceDNA(ceDNA)、杆状病毒DNA(vDNA)和Sf9细胞基因组DNA(gDNA)的大小的DNA条带由箭头指示。

[0059] 图19A-图19C示出了体内HBoV1 ssDNA和ceDNA的产生和测试。图19A是单链DNA(ssDNA)FVIIIXTEN HBoV1的琼脂糖凝胶电泳图像。图19B是FVIIIXTEN HBoV1ceDNA的琼脂糖凝胶电泳图像。图19C示出了ssFVIIIXTEN和ceFVIIIXTEN的归一化为占正常值百分比的FVIII表达水平。误差条表示标准偏差。

[0060] 图20A-图20B示出了FVIIIXTEN HBoV1 ceDNA的单体和多聚体形式的测试。图20A是FVIIIXTEN HBoV1 ceDNA的单体和多聚体形式的琼脂糖凝胶电泳图像。图20B示出了在用FVIIIXTEN HBoV1 ceDNA的单体和多聚体形式注射的小鼠中归一化为占正常值百分比的FVIII表达水平。误差条表示标准偏差。

[0061] 图21A-图21C示出了肝特异性mTTR和人A1AT启动子驱动FVIIIXTEN在HBoV1 ITR构建体中表达的测试。图21A是两侧为HBoV1 WT ITR的具有肝特异性mTTR或A1AT启动子的FVIIIXTEN表达盒的示意图。图21B是如所述通过限制酶消化产生的单链DNA(ssDNA)FVIIIXTEN HBoV1的琼脂糖凝胶电泳图像。图21C示出了在用图21A中所示的mTTR或A1AT启动子构建体注射的小鼠中归一化为占正常值百分比的FVIII表达水平。误差条表示标准偏差。

具体实施方式

[0062] 本文公开了核酸分子及其用途,所述核酸分子包含经修饰的第一反向末端重复序列(ITR)和/或经修饰的第二ITR,所述经修饰的第一反向末端重复序列(ITR)和/或经修饰的第二ITR侧接包含异源多核苷酸序列的基因盒。在一些实施方案中,第一和/或第二ITR衍生自人博卡病毒1(HBoV1)。

[0063] 在附图和序列表中展示了本公开文本的示例性构建体。为了提供对说明书和权利要求的清晰理解,下文提供了以下定义。

定义

[0064] 应注意,术语“一个/一种(a)”或“一个/一种(an)”实体是指一个/一种或多个/多种所述实体:例如,“一个核苷酸序列”应理解为代表一个或多个核苷酸序列。类似地,“一种

治疗性蛋白质”和“一种miRNA”应理解为分别代表一种或多种治疗性蛋白质和一种或多种miRNA。因此,术语“一个/一种(a)”(或“一个/一种(an)”)、“一个/一种或多个/多种”和“至少一个/一种”在本文中可互换使用。

[0065] 术语“约”在本文中用于意指大约、大致、大概或在……左右。在术语“约”结合数字范围使用时,其通过扩展所述数值上下的边界来修饰所述范围。通常,术语“约”在本文中用于向上或向下(较高或较低)以10%的差异修饰高于和低于所述值的数值。

[0066] 同样如本文所用,“和/或”是指并且涵盖一个或多个相关列示项目的任何和所有可能组合,以及在替代方案(“或”)中解释时组合的缺少。

[0067] “核酸”、“核酸分子”、“核苷酸”、“一个或多个核苷酸的序列”和“多核苷酸”可互换使用并且是指呈单链形式或双链螺旋的核糖核苷(腺苷、鸟苷、尿苷或胞苷;“RNA分子”)或脱氧核糖核苷(脱氧腺苷、脱氧鸟苷、脱氧胸苷或脱氧胞苷;“DNA分子”)的磷酸酯聚合形式或其任何磷酸酯类似物,如硫代磷酸酯和硫代酸酯。单链核酸序列是指单链DNA(ssDNA)或单链RNA(ssRNA)。双链DNA-DNA、DNA-RNA和RNA-RNA螺旋是可能的。术语核酸分子、特别是DNA或RNA分子仅仅是指分子的一级和二级结构,并且不将其限制为任何特定三级形式。因此,这个术语包括尤其在线性或环状DNA分子(例如,限制性片段)、质粒、超螺旋化DNA和染色体中发现的双链DNA。在讨论特定双链DNA分子的结构时,在本文中可以根据常规习惯描述序列,仅沿DNA的非转录链(即,具有与mRNA同源的序列的链)以5’至3’方向给出序列。“重组DNA分子”是已经经历分子生物学操纵的DNA分子。DNA包括但不限于cDNA、基因组DNA、质粒DNA、合成DNA和半合成DNA。本公开文本的“核酸组合物”包含如本文所述的一种或多种核酸。

[0068] 如本文所用,“反向末端重复序列”(或“ITR”)是指位于单链核酸序列的5’端或3’端的核酸子序列,其包含一组核苷酸(初始序列),之后下游是其反向互补体,即回文序列。初始序列与反向互补体之间的插入核苷酸序列可以具有任何长度,包括零。在一个实施方案中,可用于本公开文本的ITR包含一个或多个“回文序列”。ITR可以具有任何数目的功能。在一些实施方案中,本文所述的ITR形成发夹结构。在一些实施方案中,ITR形成T形发夹结构。在一些实施方案中,ITR形成非T形发夹结构,例如U形发夹结构。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的长期存活。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的永久存活(例如,持续细胞的整个寿命)。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的稳定性。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的保留。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的持久性。在一些实施方案中,ITR抑制或防止核酸分子在细胞的细胞核中的降解。

[0069] 在一个实施方案中,ITR的初始序列和/或反向互补体包含约2-600个核苷酸、约2-550个核苷酸、约2-500个核苷酸、约2-450个核苷酸、约2-400个核苷酸、约2-350个核苷酸、约2-300个核苷酸或约2-250个核苷酸。在一些实施方案中,初始序列和/或反向互补体包含约5-600个核苷酸、约10-600个核苷酸、约15-600个核苷酸、约20-600个核苷酸、约25-600个核苷酸、约30-600个核苷酸、约35-600个核苷酸、约40-600个核苷酸、约45-600个核苷酸、约50-600个核苷酸、约60-600个核苷酸、约70-600个核苷酸、约80-600个核苷酸、约90-600个核苷酸、约100-600个核苷酸、约150-600个核苷酸、约200-600个核苷酸、约300-600个核苷酸、约350-600个核苷酸、约400-600个核苷酸、约450-600个核苷酸、约500-600个核苷酸或

约550-600个核苷酸。在一些实施方案中,初始序列和/或反向互补体包含约5-550个核苷酸、约5至500个核苷酸、约5-450个核苷酸、约5至400个核苷酸、约5-350个核苷酸、约5至300个核苷酸或约5-250个核苷酸。在一些实施方案中,初始序列和/或反向互补体包含约10-550个核苷酸、约15-500个核苷酸、约20-450个核苷酸、约25-400个核苷酸、约30-350个核苷酸、约35-300个核苷酸或约40-250个核苷酸。在某些实施方案中,初始序列和/或反向互补体包含约225个核苷酸、约250个核苷酸、约275个核苷酸、约300个核苷酸、约325个核苷酸、约350个核苷酸、约375个核苷酸、约400个核苷酸、约425个核苷酸、约450个核苷酸、约475个核苷酸、约500个核苷酸、约525个核苷酸、约550个核苷酸、约575个核苷酸或约600个核苷酸。在特定实施方案中,初始序列和/或反向互补体包含约400个核苷酸。

[0070] 在其他实施方案中,ITR的初始序列和/或反向互补体包含约2-200个核苷酸、约5-200个核苷酸、约10-200个核苷酸、约20-200个核苷酸、约30-200个核苷酸、约40-200个核苷酸、约50-200个核苷酸、约60-200个核苷酸、约70-200个核苷酸、约80-200个核苷酸、约90-200个核苷酸、约100-200个核苷酸、约125-200个核苷酸、约150-200个核苷酸或约175-200个核苷酸。在其他实施方案中,初始序列和/或反向互补体包含约2-150个核苷酸、约5-150个核苷酸、约10-150个核苷酸、约20-150个核苷酸、约30-150个核苷酸、约40-150个核苷酸、约50-150个核苷酸、约75-150个核苷酸、约100-150个核苷酸或约125-150个核苷酸。在其他实施方案中,初始序列和/或反向互补体包含约2-100个核苷酸、约5-100个核苷酸、约10-100个核苷酸、约20-100个核苷酸、约30-100个核苷酸、约40-100个核苷酸、约50-100个核苷酸或约75-100个核苷酸。在其他实施方案中,初始序列和/或反向互补体包含约2-50个核苷酸、约10-50个核苷酸、约20-50个核苷酸、约30-50个核苷酸、约40-50个核苷酸、约3-30个核苷酸、约4-20个核苷酸或约5-10个核苷酸。在另一个实施方案中,初始序列和/或反向互补体由两个核苷酸、三个核苷酸、四个核苷酸、五个核苷酸、六个核苷酸、七个核苷酸、八个核苷酸、九个核苷酸、十个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸或20个核苷酸组成。在其他实施方案中,初始序列与反向互补体之间的插入核苷酸是(例如,由以下组成)0个核苷酸、1个核苷酸、两个核苷酸、三个核苷酸、四个核苷酸、五个核苷酸、六个核苷酸、七个核苷酸、八个核苷酸、九个核苷酸、10个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸或20个核苷酸。

[0071] 因此,如本文所用的“ITR”可以折叠回于自身并形成双链区段。例如,在折叠形成双螺旋时,序列GATCXXXXGATC包含GATC的初始序列及其互补体(3’CTAG5’)。在一些实施方案中,ITR包含初始序列与反向互补体之间的连续回文序列(例如,GATCGATC)。在一些实施方案中,ITR包含初始序列与反向互补体之间的中断的回文序列(例如,GATCXXXXGATC)。在一些实施方案中,连续或中断的回文序列的互补部分彼此相互作用以形成“发夹环”结构。如本文所用,在单链核苷酸分子上的至少两个互补序列碱基配对形成双链部分时,产生“发夹环”结构。在一些实施方案中,仅ITR的一部分形成发夹环。在其他实施方案中,整个ITR形成发夹环。在一些实施方案中,ITR保留衍生出它的野生型ITR的Rep结合元件(RBE)。RBE的保留对于ITR的稳定性和制造目的可能是重要的。

[0072] 如本文所用的术语“细小病毒”涵盖细小病毒科,包括但不限于自主复制的细小病毒和依赖病毒。自主细小病毒包括例如以下属的成员:博卡病毒属(Bocavirus)、依赖病毒

属、红病毒属、阿留申病毒属(Amdovirus)、细小病毒属(Parvovirus)、浓核病毒属(Densovirus)、重复病毒属(Iteravirus)、康特拉病毒属(Contravirus)、禽细小病毒属(Aveparvovirus)、反刍类细小病毒属(Copiparvovirus)、原细小病毒属(Protoparvovirus)、四型细小病毒属(Tetraparvovirus)、双义浓核病毒属(Ambidensovirus)、短浓核病毒属(Brevidensovirus)、肝胰浓核病毒属(Hepandensovirus)和对虾浓核病毒属(Penstyldensovirus)。示例性自主细小病毒包括但不限于人博卡病毒1(HBoV1)、猪细小病毒、小鼠微小病毒、犬细小病毒、水貂肠炎病毒、牛细小病毒、鸡细小病毒、猫瘟病毒、猫细小病毒、鹅细小病毒(GPV)、H1细小病毒、番鸭细小病毒、蛇细小病毒和B19病毒。其他自主细小病毒是本领域技术人员已知的。参见例如Fields等人Virology,第2卷,第69章(第4版,Lippincott-Raven Publishers)。

[0073] 如本文所用的术语“非AAV”涵盖来自细小病毒科的核酸、蛋白质和病毒,不包括细小病毒科的任何腺相关病毒(AAV)。“非AAV”包括但不限于以下属的自主复制成员:博卡病毒属、依赖病毒属、红病毒属、阿留申病毒属、细小病毒属、浓核病毒属、重复病毒属、康特拉病毒属、禽细小病毒属、反刍类细小病毒属、原细小病毒属、四型细小病毒属、双义浓核病毒属、短浓核病毒属、肝胰浓核病毒属和对虾浓核病毒属。

[0074] 如本文所用,术语“腺相关病毒”(AAV)包括但不限于AAV 1型、AAV 2型、AAV 3型(包括3A型和3B型)、AAV 4型、AAV 5型、AAV 6型、AAV 7型、AAV 8型、AAV 9型、AAV 10型、AAV 11型、AAV 12型、AAV 13型、蛇AAV、禽AAV、牛AAV、犬AAV、马AAV、绵羊AAV、山羊AAV、虾AAV、Gao等人(J.Virol.78:6381(2004))和Moris等人(Virol.33:375(2004))所披露的那些AAV血清型和进化枝以及目前已知或将来发现的任何其他AAV。参见例如,FIELDS等人VIROLOGY,第2卷,第69章(第4版,Lippincott-Raven Publishers)。

[0075] 如本文所用,术语“衍生自”是指组分从指定分子或生物体分离或使用指定分子或生物体制备,或者信息(例如,氨基酸或核酸序列)来自指定分子或生物体。例如,衍生自第二核酸序列(例如,ITR)的核酸序列(例如,ITR)可以包括与第二核酸序列的核苷酸序列相同或基本上相似的核苷酸序列。在核苷酸或多肽的情况下,衍生的物质可以通过例如天然存在的诱变、人工定向诱变或人工随机诱变来获得。用于衍生核苷酸或多肽的诱变可以是有意定向的或有意随机的,或者是各自的混合。诱变核苷酸或多肽以产生衍生自第一核苷酸或多肽的不同核苷酸或多肽可以是随机事件(例如,由于聚合酶失真而引起),并且所衍生的核苷酸或多肽的鉴定可以通过适当的筛选方法(例如,如本文所讨论)来进行。多肽的诱变通常需要操纵编码多肽的多核苷酸。

[0076] “无衣壳(capsid-free)”或“无衣壳(capsid-less)”载体或核酸分子是指不具有衣壳的载体构建体。

[0077] 如本文所用,“编码区”或“编码序列”是多核苷酸中由可翻译成氨基酸的密码子组成的部分。虽然“终止密码子”(TAG、TGA或TAA)通常不翻译成氨基酸,但是可以认为其是编码区的一部分,但是任何侧接序列(例如启动子、核糖体结合位点、转录终止子、内含子等)不是编码区的一部分。编码区的边界通常由5’末端的起始密码子(编码所得多肽的氨基末端)和3’末端的翻译终止密码子(编码所得多肽的羧基末端)来决定。两个或更多个编码区可以存在于单一多核苷酸构建体中(例如,在单一载体上),或者存在于单独的多核苷酸构建体中(例如,在单独的(不同的)载体上)。然后,结果就是单一载体可以仅含有单个编码

区,或包含两个或更多个编码区。

[0078] 哺乳动物细胞分泌的某些蛋白质与分泌信号肽相关,一旦生长中的蛋白质链跨越糙面内质网的输出已经起始,所述分泌信号肽便从成熟蛋白质被切割下来。本领域普通技术人员知道,信号肽通常融合至多肽的N末端,并且从完整或"全长"多肽被切割下来以产生多肽的分泌或"成熟"形式。在某些实施方案中,天然信号肽或该序列的保留了指导多肽分泌的能力的功能衍生物与所述多肽可操作地缔合。可替代地,可以使用异源哺乳动物信号肽(例如,人组织纤溶酶原激活物(TPA)或小鼠β-葡糖醛酸糖苷酶信号肽)或其功能衍生物。

[0079] 术语"下游"是指核苷酸序列位于参考核苷酸序列的3'。在某些实施方案中,下游核苷酸序列是指转录起点之后的序列。例如,基因的翻译起始密码子位于转录起始位点的下游。

[0080] 术语"上游"是指核苷酸序列位于参考核苷酸序列的5'。在某些实施方案中,上游核苷酸序列是指位于编码区或转录起点的5'侧的序列。例如,大多数启动子位于转录起始位点的上游。

[0081] 如本文所用,术语"基因盒"是指能够指导特定多核苷酸序列在适当宿主细胞中表达的DNA序列,其包含与目的多核苷酸序列可操作地连接的启动子。基因盒可以涵盖位于编码区上游(5'非编码序列)、内部或下游(3'非编码序列),且影响相关编码区的转录、RNA加工、稳定性或翻译的核苷酸序列。如果意图在真核细胞中表达编码区,则多腺苷酸化信号和转录终止序列通常将位于编码序列的3'。在一些实施方案中,基因盒包含编码基因产物的多核苷酸。在一些实施方案中,基因盒包含编码miRNA的多核苷酸。在一些实施方案中,基因盒包含异源多核苷酸序列。

[0082] 编码产物(例如,miRNA或基因产物(例如,多肽,如治疗性蛋白质))的多核苷酸可以包括与一个或多个编码区可操作地缔合的启动子和/或其他表达(例如,转录或翻译)控制序列。在可操作缔合中,以将基因产物的表达置于一个或多个调节区的影响或控制下的方式将基因产物(例如,多肽)的编码区与所述一个或多个调节区缔合。例如,如果启动子功能的诱导引起编码由编码区所编码的基因产物的mRNA的转录,并且如果启动子与编码区之间的连接的性质不干扰启动子指导基因产物表达的能力或者不干扰DNA模板被转录的能力,则编码区和启动子"可操作地缔合"。除了启动子以外的其他表达控制序列(例如,增强子、操纵子、阻遏物和转录终止信号)也可以与编码区可操作地缔合以指导基因产物表达。

[0083] "表达控制序列"是指提供编码序列在宿主细胞中的表达的调节核苷酸序列,如启动子、增强子、终止子等。表达控制序列通常涵盖有助于与其可操作地连接的编码核酸的有效转录和翻译的任何调节核苷酸序列。表达控制序列的非限制性例子包括启动子、增强子、翻译前导序列、内含子、多腺苷酸化识别序列、RNA加工位点、效应子结合位点或茎环结构。多种表达控制序列是本领域技术人员已知的。这些包括而不限于在脊椎动物细胞中起作用的表达控制序列,如但不限于来自巨细胞病毒的启动子和增强子片段(即时早期启动子,与内含子A结合)、猿猴病毒40(早期启动子)和逆转录病毒(如劳斯肉瘤病毒)。其他表达控制序列包括衍生自脊椎动物基因的那些,如肌动蛋白、热休克蛋白、牛生长激素和兔β珠蛋白,以及能够控制真核细胞中基因表达的其他序列。另外的合适的表达控制序列包括组织特异性启动子和增强子以及淋巴因子诱导型启动子(例如可由干扰素或白细胞介素诱导的启动子)。其他表达控制序列包括内含子序列、转录后调节元件和多腺苷酸化信号。在本公开文

本的其他地方讨论了另外的示例性表达控制序列。

[0084] 类似地,多种翻译控制元件是本领域普通技术人员已知的。这些翻译控制元件包括但不限于核糖体结合位点、翻译起始和终止密码子以及衍生自小核糖核酸病毒的元件(特别是内部核糖体进入位点,或IRES)。

[0085] 如本文所用的术语“表达”是指多核苷酸产生基因产物(例如RNA或多肽)的过程。它包括而不限于将多核苷酸转录成信使RNA(mRNA)、转移RNA(tRNA)、小发夹RNA(shRNA)、小干扰RNA(siRNA)或任何其他RNA产物以及将mRNA翻译为多肽。表达产生“基因产物”。如本文所用,基因产物可以是核酸,例如通过基因转录产生的信使RNA,或者是从转录物翻译的多肽。本文描述的基因产物还包括经过转录后修饰(例如多腺苷酸化或剪接)的核酸,或经过翻译后修饰(例如甲基化、糖基化、脂质的添加、与其他蛋白质亚基缔合或蛋白水解切割)的多肽。如本文所用,术语“产量”是指通过基因表达产生的多肽的量。

[0086] “载体”是指用于将核酸克隆和/或转移至宿主细胞中的任何媒介物。载体可以是复制子,另一核酸区段可以与其连接以实现所连接区段的复制。“复制子”是指在体内起自主复制单元的作用(即,能够在其自身控制下复制)的任何基因元件(例如,质粒、噬菌体、粘粒、染色体、病毒)。术语“载体”包括在体外、离体或在体内将核酸引入细胞中的媒介物。很多载体是本领域已知和使用的,包括例如质粒、经修饰的真核病毒或经修饰的细菌病毒。通过将适当的多核苷酸片段连接到具有互补粘性末端的选择载体中,可以实现将多核苷酸插入合适的载体中。

[0087] 可以对载体进行工程化以编码选择性标记或报告物,其提供对已经掺入了载体的细胞的选择或鉴定。选择性标记或报告物的表达允许鉴定和/或选择掺入了并表达含于载体上的其他编码区的宿主细胞。本领域中已知并使用的选择性标记基因的例子包括:提供针对氨苄青霉素、链霉素、庆大霉素、卡那霉素、潮霉素、双丙氨磷除草剂、磺酰胺等的抗性的基因;和用作表型标记的基因,即花色苷调节基因、异戊烷基转移酶基因等。本领域中已知并使用的报告物的例子包括:萤光素酶(Luc)、绿色荧光蛋白(GFP)、氯霉素乙酰转移酶(CAT)、β-半乳糖苷酶(LacZ)、β-葡糖醛酸糖苷酶(Gus)等。还可以将选择性标记视为报告物。

[0088] 如本文所用的术语“宿主细胞”是指例如可以或已经用作ssDNA或载体的接受者的微生物、酵母细胞、昆虫细胞和哺乳动物细胞。所述术语包括已经转导的初始细胞的后代。因此,如本文所用的“宿主细胞”通常是指已经用外源DNA序列转导的细胞。应理解,由于天然、偶然或刻意突变,单个亲代细胞的后代的形态或基因组或总DNA互补体可能不一定与初始亲代完全相同。在一些实施方案中,宿主细胞可以是体外宿主细胞。

[0089] 术语“选择性标记”是指能够基于标记基因的效应(即,对抗生素的抗性、对除草剂的抗性、比色标记、酶、荧光标记等)加以选择的鉴定因子(通常是抗生素或化学抗性基因),其中所述效应用于跟踪目的核酸的遗传和/或鉴定已经遗传目的核酸的细胞或生物体。本领域中已知并使用的选择性标记基因的例子包括:提供针对氨苄青霉素、链霉素、庆大霉素、卡那霉素、潮霉素、双丙氨磷除草剂、磺酰胺等的抗性的基因;和用作表型标记的基因,即花色苷调节基因、异戊烷基转移酶基因等。

[0090] 术语“报告基因”是指编码能够基于报告基因的效应加以鉴定的鉴定因子的核酸,其中所述效应用于跟踪目的核酸的遗传性、鉴定已经遗传目的核酸的细胞或生物体和/或

测量基因表达诱导或转录。本领域中已知并使用的报告基因的例子包括:萤光素酶(Luc)、绿色荧光蛋白(GFP)、氯霉素乙酰转移酶(CAT)、β-半乳糖苷酶(LacZ)、β-葡糖醛酸糖苷酶(Gus)等。也可以将选择性标记基因视为报告基因。

[0091] “启动子”与“启动子序列”可互换使用并且是指能够控制编码序列或功能RNA的表达的DNA序列。通常,编码序列位于启动子序列的3'。启动子可以整体衍生自天然基因,或者由衍生自在自然界中发现的不同启动子的不同元件构成,或者甚至包含合成DNA区段。本领域技术人员应理解,不同的启动子可以指导基因在不同组织或细胞类型中、或在不同发育阶段、或响应不同环境或生理条件而表达。使基因在大多数细胞类型中在大多数时间表达的启动子一般被称为“组成型启动子”。使基因在特定细胞类型中表达的启动子一般被称为“细胞特异性启动子”或“组织特异性启动子”。使基因在发育或细胞分化的特定阶段表达的启动子一般被称为“发育特异性启动子”或“细胞分化特异性启动子”。在用诱导启动子的试剂、生物分子、化学品、配体、光等暴露或处理细胞后被诱导并且使基因表达的启动子一般被称为“诱导型启动子”或“可调型启动子”。还认识到,由于在大多数情况下,尚未完全界定调节序列的确切边界,所以不同长度的DNA片段可以具有相同的启动子活性。在本公开文本的其他地方讨论了另外的示例性启动子。

[0092] 启动子序列通常在其3'末端以转录起始位点为界,并向上游(5'方向)延伸以包括以高于背景的可检测水平起始转录所必需的最小数目的碱基或元件。在启动子序列内将发现转录起始位点(例如通过用核酸酶S1标示(mapping)方便地界定)以及负责RNA聚合酶的结合的蛋白质结合结构域(共有序列)。

[0093] 在一些实施方案中,核酸分子包含组织特异性启动子。在某些实施方案中,组织特异性启动子驱动治疗性蛋白质在肝脏中、在肝细胞和/或内皮细胞中的表达。在一个特定实施方案中,启动子包含TTP启动子。在一个特定实施方案中,启动子包含mTTR启动子。在一个特定实施方案中,启动子包含A1AT启动子。

[0094] 术语“质粒”是指染色体外元件,其通常携带不是细胞中心代谢的一部分的基因,并且通常是环状双链DNA分子的形式。此类元件可以是衍生自任何来源的单链或双链DNA或RNA的线性、环状或超螺旋化的自主复制序列、基因组整合序列、噬菌体或核苷酸序列,其中多个核苷酸序列已经接合或重组至独特构建体中,所述构建体能够将用于所选基因产物的启动子片段和DNA序列以及适当的3'非翻译序列引入细胞中。

[0095] 可以使用的真核病毒载体包括但不限于腺病毒载体、逆转录病毒载体、腺相关病毒载体、痘病毒(例如,痘苗病毒载体)、杆状病毒载体或疱疹病毒载体。非病毒载体包括质粒、脂质体、带电脂质(细胞转染剂)、DNA-蛋白质复合物和生物聚合物。

[0096] “克隆载体”是指“复制子”,其是单位长度的连续复制的核酸并且其包含复制起点,如质粒、噬菌体或粘粒,另一核酸区段可以与所述复制子连接以实现所连接区段的复制。某些克隆载体能够在一种细胞类型(例如,细菌)中复制,并在另一细胞类型(例如,真核细胞)中表达。克隆载体通常包含一个或多个可以用于选择包含载体的细胞的序列和/或一个或多个用于插入目的核酸序列的多克隆位点。

[0097] 术语“表达载体”是指设计为使得所插入核酸序列在插入宿主细胞中之后能够表达的媒介物。将所插入核酸序列以如上所述的与调节区可操作地缔合地放置。

[0098] 通过本领域中熟知的方法将载体引入宿主细胞中,所述方法是例如转染、电穿孔、

显微注射、转导、细胞融合、DEAE葡聚糖、磷酸钙沉淀、脂转染(溶酶体融合)、使用基因枪或DNA载体转运蛋白。如本文所用,“培养”(“culture”,“to culture”和“culturing”)意指在体外条件下孵育细胞,所述条件允许细胞生长或分裂或使细胞维持存活状态。如本文所用,“培养的细胞”是指体外繁殖的细胞。

[0099] 如本文所用,术语“多肽”旨在涵盖单数“多肽”以及复数“多肽”,并且是指由通过酰胺键(也称为肽键)线性连接的单体(氨基酸)构成的分子。术语“多肽”是指两个或更多个氨基酸的任何一条或多条链,并且不涉及产物的具体长度。因此,肽、二肽、三肽、寡肽、“蛋白质”、“氨基酸链”或用于指两个或更多个氨基酸的一条或多条链的任何其他术语包括于“多肽”的定义内,并且术语“多肽”可以代替这些术语中的任一个术语或与其互换使用。术语“多肽”还意图指多肽的表达后修饰的产物,所述修饰包括但不限于糖基化、乙酰化、磷酸化、酰胺化、通过已知的保护/阻断基团衍生、蛋白质水解切割或通过非天然存在的氨基酸修饰。多肽可以从天然生物来源衍生或通过重组技术产生,但不一定是从指定的核酸序列翻译。其能以任何方式产生,包括通过化学合成产生。

[0100] 术语“氨基酸”包括丙氨酸(Ala或A);精氨酸(Arg或R);天冬酰胺(Asn或N);天冬氨酸(Asp或D);半胱氨酸(Cys或C);谷氨酰胺(Gln或Q);谷氨酸(Glu或E);甘氨酸(Gly或G);组氨酸(His或H);异亮氨酸(Ile或I);亮氨酸(Leu或L);赖氨酸(Lys或K);甲硫氨酸(Met或M);苯丙氨酸(Phe或F);脯氨酸(Pro或P);丝氨酸(Ser或S);苏氨酸(Thr或T);色氨酸(Trp或W);酪氨酸(Tyr或Y);和缬氨酸(Val或V)。非传统氨基酸也在本公开文本的范围内,并且包括正亮氨酸、鸟氨酸、正缬氨酸、高丝氨酸和其他氨基酸残基类似物,如Ellman等人Meth.Enzym.202:301-336(1991)中所述的那些。为了产生这种非天然存在的氨基酸残基,可使用Noren等人Science244:182(1989)和Ellman等人,同上的操作。简单来说,这些操作涉及用非天然存在的氨基酸残基化学激活抑制剂tRNA,之后进行RNA的体外转录和翻译。引入非传统氨基酸也可以使用本领域中已知的肽化学来实现。如本文所用,术语“极性氨基酸”包括具有零个净电荷,但是在其侧链的不同部分具有非零部分电荷的氨基酸(例如,M、F、W、S、Y、N、Q、C)。这些氨基酸可以参与疏水相互作用和静电相互作用。如本文所用,术语“带电氨基酸”包括在其侧链上具有非零净电荷的氨基酸(例如,R、K、H、E、D)。这些氨基酸可以参与疏水相互作用和静电相互作用。

[0101] 本公开文本还包括多肽的片段或变体及其任何组合。术语“片段”或“变体”在涉及本公开文本的多肽结合结构域或结合分子时包括保留参考多肽的至少一些特性(例如,对FcRn结合结构域或Fc变体的FcRn结合亲和力、对FVIII变体的凝结活性或对VWF片段的FVIII结合活性)的任何多肽。除了本文中其他地方讨论的特定抗体片段以外,多肽片段包括蛋白质水解片段以及缺失片段,但是不包括天然存在的全长多肽(或成熟多肽)。本公开文本的多肽结合结构域或结合分子的变体包括如上所述的片段,以及由于氨基酸取代、缺失或插入而具有改变的氨基酸序列的多肽。变体可以是天然存在的或非天然存在的。非天然存在的变体可以使用本领域中已知的诱变技术来产生。变体多肽可以包含保守或非保守氨基酸取代、缺失或添加。

[0102] “保守氨基酸取代”是用具有类似侧链的氨基酸残基替代氨基酸残基的取代。本领域已经定义了具有相似侧链的氨基酸残基家族,包括碱性侧链(例如赖氨酸、精氨酸、组氨酸)、酸性侧链(例如天冬氨酸、谷氨酸)、不带电荷的极性侧链(例如甘氨酸、天冬酰胺、谷氨

酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸)、非极性的侧链(例如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸)、β-分支侧链(例如苏氨酸、缬氨酸、异亮氨酸)、以及芳香族侧链(例如酪氨酸、苯丙氨酸、色氨酸、组氨酸)。因此,如果多肽中的氨基酸被来自相同侧链家族的另一种氨基酸替代,则将取代视为保守的。在另一个实施方案中,氨基酸串可以用结构上类似但侧链家族成员的顺序和/或组成不同的串来保守地替代。

[0103] 如本领域中已知的术语"同一性百分比"是两个或更多个多肽序列或者两个或更多个多核苷酸序列之间的关系,如通过比较序列所确定。在本领域中,根据具体情况,"同一性"也意指多肽或多核苷酸序列之间的序列关联性程度,如通过此类序列的串之间的匹配所确定。"同一性"可以通过已知方法容易地计算,所述已知方法包括但不限于:Computational Molecular Biology(Lesk,A.M.编)Oxford University Press,纽约(1988);Biocomputing:Informatics and Genome Projects(Smith,D.W.编)Academic Press,纽约(1993);Computer Analysis of Sequence Data,Part I(Griffin,A.M.和Griffin,H.G.编)Humana Press,New Jersey(1994);Sequence Analysis in Molecular Biology(von Heinje,G.编)Academic Press(1987);以及Sequence Analysis Primer(Gribskov,M.和Devereux,J.编)Stockton Press,纽约(1991)。确定同一性的优选方法设计为给出所测试序列之间的最佳匹配。确定同一性的方法被编纂于可公开获得的电脑程序中。序列比对和同一性百分比计算可以使用序列分析软件来进行,所述序列分析软件是例如LASERGENE生物信息学计算套件的Megalign程序(DNASTAR Inc.,威斯康星州麦迪逊市)、GCG程序套件(Wisconsin Package 9.0版,Genetics Computer Group(GCG),威斯康星州麦迪逊市)、BLASTP、BLASTN、BLASTX(Altschul等人,J.Mol.Biol.215:403(1990))和DNASTAR(DNASTAR,Inc.1228S.Park St.威斯康星州麦迪逊市53715美国)。在本申请的上下文中将理解,如果使用序列分析软件进行分析,除非另外指定,否则分析结果将基于所参考程序的"缺省值"。如本文所用的"缺省值"将意指在首次初始化时最初用软件加载的任一组值或参数。出于确定查询序列(例如核酸序列)与参考序列之间的百分比同一性的目的,仅使用与参考序列中的核苷酸匹配的查询序列中的核苷酸来计算同一性百分比。因此,在确定查询序列或其指定部分(例如,核苷酸1-522)与参考序列之间的同一性百分比,通过将匹配的核苷酸的数量除以完全查询序列中的核苷酸的总数来计算同一性百分比。

[0104] 如本文所用,对应于本公开文本的特定序列中的核苷酸的核苷酸是通过比对本公开文本的序列以最大化与参考序列的同一性来鉴定。用于鉴定参考序列中的等效氨基酸的编号是基于用于鉴定本公开文本的序列中的相应氨基酸的编号。

[0105] 如本文所用,"治疗"("Treat"、"treatment"、"treating")是指,例如,疾病或病症严重程度的降低;病程持续时间的缩短;与疾病或病症有关的一种或多种症状的改善;为患有疾病或病症的受试者提供有益的效果而不一定治愈所述疾病或病症或与疾病或病症相关的一种或多种症状的预防。

[0106] 如本文所用,"施用"意指经由药学上可接受的途径将本公开文本的药学上可接受的核酸分子、从所述核酸分子表达的多肽或包含所述核酸分子的载体给予至受试者。施用途径可以是静脉内的,例如静脉内注射和静脉内输注。另外的施用途径包括例如皮下、肌内、口服、经鼻和经肺施用。核酸分子、多肽和载体可以作为包含至少一种赋形剂的药物组合物的一部分来施用。

[0107] 如本文所用,“脂质纳米颗粒”是指具有至少一个纳米级维度(例如,1nm至1,000nm)的颗粒,其包含一种或多种阳离子脂质。在一些实施方案中,脂质纳米颗粒包含在可用于将活性剂或治疗剂如核酸(例如mRNA)递送至相关靶位点(例如细胞、组织、器官、肿瘤等)的配制品中。在一些实施方案中,本文公开的脂质纳米颗粒包含核酸。此类脂质纳米颗粒通常包含一种或多种选自中性脂质、带电脂质、类固醇和聚合物缀合脂质的赋形剂。在一些实施方案中,可以将活性剂或治疗剂如核酸包封在脂质纳米颗粒的脂质部分中,或包封在由脂质纳米颗粒的一些或全部脂质部分包封的水性空间中,从而保护其免受酶降解或由宿主生物体或细胞的机制触发的其他不期望的效应,如不利的免疫应答。

[0108] 如本文所用的术语“药学上可接受的”是指分子实体和组合物在施用至人时在生理上可耐受并且通常不会产生毒性或过敏反应或类似的不良反应(如胃部不适、头晕等)。任选地,如本文所用,术语“药学上可接受的”意指由联邦政府或州政府的管理机构批准或者在美国药典或其他公认药典中列示用于动物,更特别是用于人。

[0109] 如本文所用,词组“有需要的受试者”包括将受益于本公开文本的核酸分子、多肽或载体的施用的受试者(如哺乳动物受试者)。在一些实施方案中,受试者是人受试者。在一些实施方案中,受试者是患有血友病的个体。受试者可以是成年人或未成年人(例如,小于12岁)。

[0110] 如本文所用,术语“治疗性蛋白质”是指本领域中已知可以向受试者施用的任何多肽。在一些实施方案中,治疗性蛋白质包含选自以下的蛋白质:凝血因子、生长因子、抗体、其功能片段、或它们的组合。如本文所用,术语“凝血因子”是指天然存在的或重组产生的分子或其类似物,其预防受试者的出血发作或缩短出血发作的持续时间。换句话说,其意指具有促凝血活性(即,负责将纤维蛋白原转化为不溶性纤维蛋白的网状物,从而引起血液凝结或凝血)的分子。如本文所用的“凝血因子”包括激活的凝血因子、其酶原或可激活的凝血因子。“可激活的凝血因子”是呈无活性形式(例如,呈其酶原形式)的凝血因子,其能够被转化为活性形式。术语“凝血因子”包括但不限于因子I(FI)、因子II(FII)、因子III(FIII)、因子IV(FIV)、因子V(FV)、因子VI(FVI)、因子VII(FVII)、因子VIII(FVIII)、因子IX(FIX)、因子X(FX)、因子XI(FXI)、因子XII(FXII)、因子XIII(FXIII)、血管性血友病因子(VWF)、前激肽释放酶、高分子量激肽原、纤连蛋白、抗凝血酶III、肝素辅因子II、蛋白质C、蛋白质S、蛋白质Z、蛋白质Z相关蛋白酶抑制剂(ZPI)、纤溶酶原、α2-抗纤溶酶、组织纤溶酶原激活物(tPA)、尿激酶、纤溶酶原激活物抑制剂-1(PAI-1)、纤溶酶原激活物抑制剂-2(PAI2)、其酶原、其活化形式或其任何组合。

[0111] 如本文所用的“凝血活性”意指参与生化反应级联的能力,所述级联以形成纤维蛋白凝块告终和/或降低出血症或出血发作的严重程度、持续时间或频率。

[0112] 如本文所用,“生长因子”包括本领域中已知的包括细胞因子和激素在内的任何生长因子。

[0113] 如本文所用,术语“异源的”或“外源的”是指此类分子通常在给定背景下(例如,在细胞中或在多肽中)未发现。例如,可以将外源或异源分子引入细胞中并且仅在例如通过转染或其他形式的基因工程化操纵细胞之后存在,或者异源氨基酸序列可以存在于其天然不存在的蛋白质中。

[0114] “参考核苷酸序列”在本文中用作与本公开文本的核苷酸序列的比较时,是与本公

开文本的核苷酸序列基本上相同的多核苷酸序列,但是对应于FVIII序列的部分未经优化。在一些实施方案中,本文公开的核酸分子的参考核苷酸序列是SEQ ID NO:32。

[0115] 如本文所用,关于核苷酸序列的术语“优化”是指编码多肽的多核苷酸序列,其中所述多核苷酸序列已经突变以增强所述多核苷酸序列的特性。在一些实施方案中,进行优化以增加转录水平、增加翻译水平、增加稳态mRNA水平、增加或减少调节蛋白(如通用转录因子)的结合、增加或减少剪接或增加多核苷酸序列产生的多肽的产量。可以对多核苷酸序列进行以使其优化的变化的例子包括密码子优化、G/C含量优化、去除重复序列、去除富含AT的元件、去除隐蔽剪接位点、去除阻遏转录或翻译的顺式作用元件、添加或去除多聚T或多聚A序列、在转录起始位点周围添加增强转录的序列(如Kozak共有序列)、去除可以形成茎环结构的序列、去除去稳定序列、去除CpG基序及其两个或更多个组合。

核酸分子

[0116] 本公开文本的某些方面旨在克服AAV载体用于基因疗法的缺陷。特别地,本公开文本的某些方面涉及包含第一ITR、第二ITR和基因盒的核酸分子。在一些实施方案中,基因盒编码治疗性蛋白质和/或miRNA。在一些实施方案中,第一ITR和第二ITR侧接包含异源多核苷酸序列的基因盒。在一些实施方案中,核酸分子不包含编码衣壳蛋白、复制蛋白和/或组装蛋白的基因。在一些实施方案中,基因盒编码治疗性蛋白质。在一些实施方案中,治疗性蛋白质包含凝血因子。在一些实施方案中,基因盒编码miRNA。在某些实施方案中,基因盒定位于第一ITR与第二ITR之间。在一些实施方案中,核酸分子还包含一个或多个非编码区。在某些实施方案中,所述一个或多个非编码区包含启动子序列、内含子、转录后调节元件、3’UTR多聚(A)序列或其任何组合。

[0117] 在一个实施方案中,基因盒是单链核酸。在另一个实施方案中,基因盒是双链核酸。在另一个实施方案中,基因盒是封闭末端双链核酸(ceDNA)。

[0118] 在一些实施方案中,核酸分子包含:(a)第一ITR,其是衍生自细小病毒科的非AAV家族成员的ITR(例如,HBoV1 ITR);(b)组织特异性启动子序列,例如TTP或TTR启动子;(c)内含子,例如合成内含子;(d)编码miRNA或治疗性蛋白质(例如凝血因子)的核苷酸;(e)转录后调节元件,例如WPRE;(f)3’UTR多聚(A)尾序列,例如bGHpA;(g)第二ITR,其是衍生自细小病毒科的非AAV家族成员的ITR(例如,HBoV1 ITR)。在一些实施方案中,核酸分子包含:(a)第一ITR,其是衍生自细小病毒科的非AAV家族成员的ITR(例如,HBoV1 ITR);(b)组织特异性启动子序列,例如mTTR启动子;(c)内含子,例如合成内含子;(d)编码miRNA或治疗性蛋白质(例如凝血因子)的核苷酸;(e)转录后调节元件,例如WPRE;(f)3’UTR多聚(A)尾序列,例如bGHpA;(g)第二ITR,其是衍生自细小病毒科的非AAV家族成员的ITR(例如,HBoV1 ITR)。在一些实施方案中,组织特异性启动子是人α-1-抗胰蛋白酶(A1AT)启动子。在一些实施方案中,组织特异性启动子包含SEQ ID NO:36的核苷酸序列。

[0119] 在一些实施方案中,本文公开了包含基因盒的分离的核酸分子,所述基因盒包含与SEQ ID NO:9至少约75%相同的核苷酸序列。在一些实施方案中,本文公开了包含与SEQ ID NO:9具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或100%序列同一性的核苷酸序列的核酸分子。

[0120] 在一些实施方案中,本文公开了包含基因盒的分离的核酸分子,所述基因盒包含与SEQ ID NO:33至少约75%相同的核苷酸序列。在一些实施方案中,本文公开了包含与SEQ

ID NO:33具有至少50％、55％、60％、65％、70％、75％、80％、85％、90％、95％或100％序列同一性的核苷酸序列的核酸分子。

[0121] 在一些实施方案中,本文公开了包含基因盒的分离的核酸分子,所述基因盒包含与SEQ ID NO:14至少约75％相同的核苷酸序列。在一些实施方案中,本文公开了包含与SEQ ID NO:14具有至少50％、55％、60％、65％、70％、75％、80％、85％、90％、95％或100％序列同一性的核苷酸序列的核酸分子。

[0122] 在另一方面,本文公开了包含表达因子VIII(FVIII)多肽的基因盒的分离的核酸分子,其中所述基因盒包含与SEQ ID NO:35至少85％相同的核苷酸序列。在一些实施方案中,基因盒包含与SEQ ID NO:35至少90％相同的核苷酸序列。在一些实施方案中,基因盒包含与SEQ ID NO:35至少95％、至少96％、至少97％、至少98％、至少99％或100％相同的核苷酸序列。在一些实施方案中,核苷酸序列与SEQ ID NO:35至少50％相同。

[0123] 本文还公开了包含表达因子VIII(FVIII)多肽的基因盒的分离的核酸分子,其中所述基因盒包含SEQ ID NO:35的核苷酸序列。

[0124] 在某些实施方案中,本文公开的核酸分子包含来自人博卡病毒1(HBoV1)的ITR序列。在某些实施方案中,本文公开的核酸分子包含与SEQ ID NO:1或SEQ ID NO:2至少约75％相同的第一ITR。

A.反向末端重复序列(ITR)

[0125] 本公开文本的某些方面涉及一种核酸分子,其包含第一ITR(例如,5'ITR)和第二ITR(例如,3'ITR)。通常,ITR参与细小病毒(例如,AAV)从原核质粒进行DNA复制和拯救或切除(Samulski等人,1983,1987;Senapathy等人,1984;Gottlieb和Muzyczka,1988)。另外,ITR似乎是AAV原病毒整合和将AAV DNA包装至病毒粒子中所需的最小序列(McLaughlin等人,1988;Samulski等人,1989)。这些元件是细小病毒基因组的有效操纵必需的。假设对于ITR功能不可或缺的最小定义元件是Rep结合位点和末端解离位点加允许发夹形成的可变回文序列。回文核苷酸区域通常作为DNA复制的起点以及作为病毒的包装信号以顺式一起作用。ITR中的互补序列在DNA复制期间折叠成发夹结构。在一些实施方案中,ITR折叠成发夹T形结构。在其他实施方案中,ITR折叠成非T形发夹结构,例如折叠成U形发夹结构。数据表明,AAV ITR的T形发夹结构可以抑制两侧为ITR的转基因的表达。参见例如,Zhou等人(2017)Scientific Reports 7:5432。通过利用不形成T形发夹结构的ITR,可以避免这种形式的抑制。因此,在某些方面中,与包含形成T形发夹的AAV ITR的多核苷酸相比,包含非AAV ITR的多核苷酸具有改善的转基因表达。

[0126] 如本文所用,“反向末端重复序列”(或“ITR”)是指位于单链核酸序列的5'端或3'端的核酸子序列,其包含一组核苷酸(初始序列),之后下游是其反向互补体,即回文序列。初始序列与反向互补体之间的插入核苷酸序列可以具有任何长度,包括零。在一个实施方案中,可用于本公开文本的ITR包含一个或多个“回文序列”。ITR可以具有任何数目的功能。在一些实施方案中,本文所述的ITR形成发夹结构。在一些实施方案中,ITR形成T形发夹结构。在一些实施方案中,ITR形成非T形发夹结构,例如U形发夹结构。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的长期存活。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的永久存活(例如,持续细胞的整个寿命)。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的稳定性。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的

保留。在一些实施方案中,ITR促进核酸分子在细胞的细胞核中的持久性。在一些实施方案中,ITR抑制或防止核酸分子在细胞的细胞核中的降解。

[0127] 因此,如本文所用的“ITR”可以折叠回于自身并形成双链区段。例如,在折叠形成双螺旋时,序列GATCXXXXGATC包含GATC的初始序列及其互补体(3'CTAG5')。在一些实施方案中,ITR包含初始序列与反向互补体之间的连续回文序列(例如,GATCGATC)。在一些实施方案中,ITR包含初始序列与反向互补体之间的中断的回文序列(例如,GATCXXXXGATC)。在一些实施方案中,连续或中断的回文序列的互补部分彼此相互作用以形成“发夹环”结构。如本文所用,在单链核苷酸分子上的至少两个互补序列碱基配对形成双链部分时,产生“发夹环”结构。在一些实施方案中,仅ITR的一部分形成发夹环。在其他实施方案中,整个ITR形成发夹环。

[0128] 在一些实施方案中,ITR包含天然存在的ITR,例如ITR包含衍生自细小病毒科的成员的ITR的全部或一部分。在一些实施方案中,ITR包含合成序列。在一个实施方案中,第一ITR或第二ITR包含合成序列。在另一个实施方案中,第一ITR和第二ITR各自包含合成序列。在一些实施方案中,第一ITR或第二ITR包含天然存在的序列。在另一个实施方案中,第一ITR和第二ITR各自包含天然存在的序列。

[0129] 在一些实施方案中,ITR包含天然存在的ITR的一部分(例如,截短型ITR)或由所述部分组成。在一些实施方案中,ITR包含天然存在的ITR的片段或由所述片段组成,其中所述片段包含至少约5个核苷酸、至少约10个核苷酸、至少约15个核苷酸、至少约20个核苷酸、至少约25个核苷酸、至少约30个核苷酸、至少约35个核苷酸、至少约40个核苷酸、至少约45个核苷酸、至少约50个核苷酸、至少约55个核苷酸、至少约60个核苷酸、至少约65个核苷酸、至少约70个核苷酸、至少约75个核苷酸、至少约80个核苷酸、至少约85个核苷酸、至少约90个核苷酸、至少约95个核苷酸、至少约100个核苷酸、至少约125个核苷酸、至少约150个核苷酸、至少约175个核苷酸、至少约200个核苷酸、至少约225个核苷酸、至少约250个核苷酸、至少约275个核苷酸、至少约300个核苷酸、至少约325个核苷酸、至少约350个核苷酸、至少约375个核苷酸、至少约400个核苷酸、至少约425个核苷酸、至少约450个核苷酸、至少约475个核苷酸、至少约500个核苷酸、至少约525个核苷酸、至少约550个核苷酸、至少约575个核苷酸或至少约600个核苷酸;其中ITR保留天然存在的ITR的功能特性。在某些实施方案中,ITR包含天然存在的ITR的片段或由所述片段组成,其中所述片段包含至少约129个核苷酸;其中ITR保留天然存在的ITR的功能特性。在某些实施方案中,ITR包含天然存在的ITR的片段或由所述片段组成,其中所述片段包含至少约102个核苷酸;其中ITR保留天然存在的ITR的功能特性。在一些实施方案中,ITR保留衍生出它的野生型ITR的Rep结合元件(RBE)。在一些实施方案中,ITR保留衍生出它的野生型ITR的RBE中的至少一个。在一些实施方案中,ITR保留衍生出它的野生型ITR的RBE或其功能部分中的至少一个。RBE的保留对于ITR的稳定性和制造目的可能是重要的。

[0130] 在一些实施方案中,ITR包含天然存在的ITR的一部分或由所述部分组成,其中所述片段包含天然存在的ITR的长度的至少约5%、至少约10%、至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%;其中所述片段保留天然存

在的ITR的功能特性。在一些实施方案中,第一ITR和/或第二ITR衍生自野生型HBoV1 ITR。在一些实施方案中,第一ITR和/或第二ITR衍生自野生型B19 ITR。在一些实施方案中,第一ITR和/或第二ITR衍生自野生型GPV ITR。

[0131] 在某些实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少50%、至少51%、至少52%、至少53%、至少54%、至少55%、至少56%、至少57%、至少58%、至少59%、至少60%、至少61%、至少62%、至少63%、至少64%、至少65%、至少66%、至少67%、至少68%、至少69%、至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%的序列同一性;其中ITR保留天然存在的ITR的功能特性。在其他实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少90%的序列同一性;其中ITR保留天然存在的ITR的功能特性。在一些实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少80%的序列同一性;其中ITR保留天然存在的ITR的功能特性。在一些实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少70%的序列同一性;其中ITR保留天然存在的ITR的功能特性。在一些实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少60%的序列同一性;其中ITR保留天然存在的ITR的功能特性。在一些实施方案中,ITR包含如下序列或由如下序列组成,所述序列在正确比对时与天然存在的ITR的同源部分具有至少50%的序列同一性;其中ITR保留天然存在的ITR的功能特性。

[0132] 在一些实施方案中,ITR包含来自AAV基因组的ITR。在一些实施方案中,ITR是选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10和AAV11及其任何组合的AAV基因组的ITR。在一些实施方案中,ITR是本领域技术人员已知的任何AAV基因组的ITR,包括天然分离物,例如天然人分离物。在一个特定实施方案中,ITR是AAV2基因组的ITR。在另一个实施方案中,ITR是经基因工程化以在其5′和3′端包括衍生自一个或多个AAV基因组的ITR的合成序列。

[0133] 在一些实施方案中,ITR不是衍生自AAV基因组(即ITR衍生自不是AAV的病毒)。在一些实施方案中,ITR是非AAV的ITR。在一些实施方案中,ITR是来自选自但不限于以下的病毒科细小病毒科的非AAV基因组的ITR:博卡病毒属、依赖病毒属、红病毒属、阿留申病毒属、细小病毒属、浓核病毒属、重复病毒属、康特拉病毒属、禽细小病毒属、反刍类细小病毒属、原细小病毒属、四型细小病毒属、双义浓核病毒属、短浓核病毒属、肝胰浓核病毒属、对虾浓核病毒属及其任何组合。在某些实施方案中,ITR衍生自人博卡病毒1(HBoV1)。在另一个实施方案中,ITR衍生自红病毒属细小病毒B19(人病毒)。在另一个实施方案中,ITR衍生自番鸭细小病毒(MDPV)株。在某些实施方案中,MDPV株被减弱,例如MDPV株FZ91-30。在其他实施方案中,MDPV株是病原性的,例如MDPV株YY。在一些实施方案中,ITR衍生自猪细小病毒,例如猪细小病毒U44978。在一些实施方案中,ITR衍生自小鼠微小病毒,例如小鼠微小病毒U34256。在一些实施方案中,ITR衍生自犬细小病毒,例如犬细小病毒M19296。在一些实施方案中,ITR衍生自水貂肠炎病毒,例如水貂肠炎病毒D00765。在一些实施方案中,ITR衍生自

依赖细小病毒(Dependoparvovirus)。在一个实施方案中,依赖细小病毒是依赖病毒属鹅细小病毒(GPV)株。在一个具体实施方案中,GPV株被减弱,例如GPV株82-0321V。在另一个具体实施方案中,GPV株是病原性的,例如GPV株B。

[0134] 核酸分子的第一ITR和第二ITR可以衍生自相同基因组(例如,衍生自相同病毒的基因组),或衍生自不同基因组(例如,衍生自两种或更多种不同病毒基因组的基因组)。在某些实施方案中,第一ITR和第二ITR衍生自相同的AAV基因组。在一个具体实施方案中,存在于本发明的核酸分子中的这两个ITR是相同的,并且特别地可以是AAV2ITR。在其他实施方案中,第一ITR衍生自AAV基因组并且第二ITR并非衍生自AAV基因组(例如,非AAV基因组)。在其他实施方案中,第一ITR并非衍生自AAV基因组(例如,非AAV基因组)并且第二ITR衍生自AAV基因组。在再其他实施方案中,第一ITR和第二ITR二者都并非衍生自AAV基因组(例如,非AAV基因组)。在一个特定实施方案中,第一ITR和第二ITR是相同的。

[0135] 在一些实施方案中,第一ITR衍生自非AAV基因组,并且第二ITR衍生自非AAV基因组,其中第一ITR和第二ITR衍生自相同的基因组。非AAV病毒基因组的非限制性例子来自:博卡病毒属、依赖病毒属、红病毒属、阿留申病毒属、细小病毒属、浓核病毒属、重复病毒属、康特拉病毒属、禽细小病毒属、反刍类细小病毒属、原细小病毒属、四型细小病毒属、双义浓核病毒属、短浓核病毒属、肝胰浓核病毒属和对虾浓核病毒属。在一些实施方案中,第一ITR衍生自非AAV基因组,并且第二ITR衍生自非AAV基因组,其中第一ITR和第二ITR衍生自不同的病毒基因组。

[0136] 在一些实施方案中,第一ITR衍生自AAV基因组,并且第二ITR衍生自人博卡病毒1(HBoV1)。在其他实施方案中,第二ITR衍生自AAV基因组,并且第一ITR衍生自人博卡病毒1(HBoV1)。

[0137] 在一些实施方案中,第一ITR包含衍生自AAV或非AAV基因组的ITR的全部或一部分或由所述ITR的全部或一部分组成,并且第二ITR包含衍生自AAV或非AAV基因组的ITR的全部或一部分或由所述ITR的全部或一部分组成。在一些实施方案中,衍生自AAV或非AAV基因组的ITR的一部分是衍生自AAV或非AAV基因组的天然存在的ITR的截短形式。在一些实施方案中,衍生自AAV或非AAV基因组的ITR的一部分包含衍生自AAV或非AAV基因组的天然存在的ITR的部分。例如,衍生自AAV或非AAV基因组的ITR的一部分包含衍生自AAV或非AAV基因组的天然存在的ITR的部分,其中至少一个RBE或其功能部分被保留。

[0138] 在某些实施方案中,第一ITR和/或第二ITR包含衍生自HBoV1的ITR的全部或一部分,或由所述ITR的全部或一部分组成。在某些实施方案中,第一ITR和/或第二ITR包含衍生自HBoV1的ITR的全部或一部分,或由所述ITR的全部或一部分组成。在一些实施方案中,第二ITR是第一ITR的反向互补体。在一些实施方案中,第一ITR是第二ITR的反向互补体。在一些实施方案中,衍生自HBoV1的第一ITR和/或第二ITR能够形成发夹结构。在某些实施方案中,发夹结构不包含T形发夹。

[0139] 在一些实施方案中,第一ITR和/或第二ITR包含如下核苷酸序列或由如下核苷酸序列组成,所述核苷酸序列与SEQ ID NO:1或2中所示的核苷酸序列至少约50%、至少约55%、至少约60%、至少约65%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或100%相同,其中第一ITR和/或第二ITR保留衍生出它的HBoV1 ITR的功能特性。在一些实施方案

中,第一ITR和/或第二ITR包含如下核苷酸序列或由如下核苷酸序列组成,所述核苷酸序列与选自SEQ ID NO:1或2的核苷酸序列至少约50%、至少约55%、至少约60%、至少约65%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或100%相同,其中第一ITR和/或第二ITR能够形成发夹结构。在某些实施方案中,发夹结构不包含T形发夹。

[0140] 在一些实施方案中,第一ITR和/或第二ITR包含SEQ ID NO:1的核苷酸序列或由其组成。在一些实施方案中,第一ITR和/或第二ITR包含SEQ ID NO:2的核苷酸序列或由其组成。在一些实施方案中,第一ITR包含SEQ ID NO:1中所示的核苷酸序列或由其组成。在一些实施方案中,第二ITR包含SEQ ID NO:2中所示的核苷酸序列或由其组成。在一些实施方案中,第一ITR包含SEQ ID NO:1中所示的核苷酸序列或由其组成,并且第二ITR包含SEQ ID NO:2中所示的核苷酸序列或由其组成。

[0141] 对于本领域技术人员将理解,本文所述的任何第一ITR序列可以与本文所述的任何第二ITR序列匹配。在一些实施方案中,本文所述的第一ITR序列是5'ITR序列。在一些实施方案中,本文所述的第二ITR序列是3'ITR序列。在一些实施方案中,本文所述的第二ITR序列是5'ITR序列。在一些实施方案中,本文所述的第一ITR序列是3'ITR序列。本领域技术人员将能够确定关于基因盒的架构的第一和第二ITR的合适的取向。

[0142] 在另一个特定实施方案中,ITR是经基因工程化以在其5'和3'端包括并非衍生自AAV基因组的ITR的合成序列。在另一个特定实施方案中,ITR是经基因工程化以在其5'和3'端包括衍生自一个或多个非AAV基因组的ITR的合成序列。存在于本发明的核酸分子中的两个ITR可以是相同或不同的非AAV基因组。具体地,ITR可以衍生自相同的非AAV基因组。在一个具体实施方案中,存在于本发明的核酸分子中的这两个ITR是相同的,并且特别地可以是AAV2 ITR。

[0143] 在一些实施方案中,ITR序列包含一个或多个回文序列。本文所公开的ITR的回文序列包括但不限于天然回文序列(即,在自然界中发现的序列)、合成序列(即,未在自然界中发现的序列)(如伪回文序列)及其组合或经修饰形式。"伪回文序列"是包括不完美的回文序列在内的回文DNA序列,其与形成二级结构的天然AAV或非AAV回文序列中的序列共享小于80%(包括小于70%、60%、50%、40%、30%、20%、10%或5%或无)核酸序列同一性。天然回文序列可以从本文所公开的任何基因组获得或衍生。合成回文序列可以基于本文所公开的任何基因组。

[0144] 回文序列可以是连续或中断的。在一些实施方案中,回文序列是中断的,其中回文序列包含第二序列的插入。在一些实施方案中,第二序列包含启动子、增强子、整合酶的整合位点(例如,Cre或Flp重组酶的位点)、基因产物的开放阅读框或其组合。

[0145] 在一些实施方案中,ITR形成发夹环结构。在一个实施方案中,第一ITR形成发夹结构。在另一个实施方案中,第二ITR形成发夹结构。仍然在另一个实施方案中,第一ITR和第二ITR二者都形成发夹结构。在一些实施方案中,第一ITR和/或第二ITR不形成T形发夹结构。在某些实施方案中,第一ITR和/或第二ITR形成非T形发夹结构。在一些实施方案中,非T形发夹结构包含U形发夹结构。

[0146] 在一些实施方案中,本文所述的核酸分子中的ITR可以是转录激活的ITR。转录激活的ITR可以包含已经通过包括至少一个转录活性元件而被转录激活的野生型ITR的全部

或一部分。多种类型的转录活性元件适用于这个背景。在一些实施方案中,转录活性元件是组成型转录活性元件。组成型转录活性元件提供持续水平的基因转录,并且在需要持续表达转基因时是优选的。在其他实施方案中,转录活性元件是诱导型转录活性元件。诱导型转录活性元件通常在诱导物(或诱导条件)不存在下展现低活性,并且在诱导物存在下(或切换至诱导条件)上调。当仅在某些时间或在某些位置需要表达时或当需要使用诱导剂逐步增加表达水平时,诱导型转录活性元件可能是优选的。转录活性元件也可以具有组织特异性;也就是说,其仅在某些组织或细胞类型中展现活性。

[0147] 可以以多种方式将转录活性元件掺入ITR中。在一些实施方案中,将转录活性元件掺入ITR的任何部分的5′或ITR的任何部分的3′。在其他实施方案中,转录激活的ITR的转录活性元件位于两个ITR序列之间。如果转录活性元件包含两个或更多个必须被间隔开的元件,则那些元件可以与ITR的部分交替。在一些实施方案中,ITR的发夹结构缺失并用转录元件的反向重复来替代。这后一种排列将产生模拟结构中的缺失部分的发夹。多个串联转录活性元件也可以存在于转录激活的ITR中,并且这些元件可以相邻或间隔开。另外,可以将蛋白质结合位点(例如,Rep结合位点)引入转录激活的ITR的转录活性元件中。转录活性元件可以包含使得能够通过RNA聚合酶受控转录DNA以形成RNA的任何序列,并且可以包含例如如下文所定义的转录活性元件。

[0148] 转录激活的ITR向具有相对受限的核苷酸序列长度的核酸分子提供转录激活和ITR功能二者,这有效地使可以携带并从核酸分子表达的转基因的长度最大化。将转录活性元件掺入ITR中可以以多种方式来完成。对ITR序列和转录活性元件的序列需求的比较可以提供对编码ITR内的元件的方式的了解。例如,可以通过在复制转录活性元件的功能元件的ITR序列中引入特定变化将转录活性添加至ITR。本领域中存在多种技术可有效添加、缺失和/或改变特定位点的具体核苷酸序列(参见例如,Deng和Nickoloff(1992)Anal.Biochem.200:81-88)。产生转录激活的ITR的另一方式涉及在ITR中的所需位置引入限制位点。另外,可以使用本领域中已知的方法将多个转录活性元件掺入转录激活的ITR中。

[0149] 通过说明的方式,可以通过包括一个或多个诸如以下的转录活性元件来产生转录激活的ITR:TATA盒、GC盒、CCAAT盒、Sp1位点、Inr区、CRE(cAMP调节元件)位点、ATF-1/CRE位点、APBβ盒、APBα盒、CArG盒、CCAC盒或如本领域中已知参与转录的任何其他元件。

B.治疗性蛋白质

[0150] 本公开文本的某些方面涉及一种包含第一ITR、第二ITR和编码靶序列的基因盒的核酸分子,其中靶序列编码治疗性蛋白质。在一些实施方案中,基因盒编码一种治疗性蛋白质。在一些实施方案中,基因盒编码多于一种治疗性蛋白质。在一些实施方案中,基因盒编码相同治疗性蛋白质的两个或更多个拷贝。在一些实施方案中,基因盒编码相同治疗性蛋白质的两种或更多种变体。在一些实施方案中,基因盒编码两种或更多种不同的治疗性蛋白质。

[0151] 本公开文本的某些实施方案涉及一种包含第一ITR、第二ITR和编码治疗性蛋白质的基因盒的核酸分子,其中治疗性蛋白质包含凝血因子。在一些实施方案中,凝血因子选自FI、FII、FIII、FIV、FV、FVI、FVII、FVIII、FIX、FX、FXI、FXII、FXIII、VWF、前激肽释放酶、高分子量激肽原、纤连蛋白、抗凝血酶III、肝素辅因子II、蛋白质C、蛋白质S、蛋白质Z、蛋白质Z

相关蛋白酶抑制剂(ZPI)、纤溶酶原、α2-抗纤维蛋白溶酶、组织纤溶酶原激活物(tPA)、尿激酶、纤溶酶原激活物抑制剂-1(PAI-1)、纤溶酶原激活物抑制剂-2(PAI2)、其任何酶原、其任何活性形式及其任何组合。在一个实施方案中,凝血因子包含FVIII或其变体或片段。在另一个实施方案中,凝血因子包含FIX或其变体或片段。在另一个实施方案中,凝血因子包含FVII或其变体或片段。在另一个实施方案中,凝血因子包含VWF或其变体或片段。

[0152] 在一些实施方案中,核酸分子包含第一ITR、第二ITR和编码靶序列的基因盒,其中靶序列编码治疗性蛋白质,其中治疗性蛋白质包含因子VIII多肽。除非另外指定,否则如本文所用,"因子VIII"在整个本申请中缩写为"FVIII",意指具有其在凝结中的正常作用的功能性FVIII多肽。因此,术语FVIII包括功能性变体多肽。"FVIII蛋白"可与FVIII多肽(或蛋白质)或FVIII互换使用。FVIII功能的例子包括但不限于激活凝结的能力、作为因子IX的辅因子作用的能力或在Ca^{2+}和磷脂存在下与因子IX形成因子X酶(tenase)复合物的能力,所述复合物然后将因子X转化为激活形式Xa。

[0153] 本文所用的治疗性蛋白质中的FVIII部分具有FVIII活性。FVIII活性可以通过本领域中任何已知方法来测量。许多测试可用于评估凝血系统的功能:激活部分促凝血酶原激酶时间(aPTT)测试、生色测定、ROTEM测定、凝血酶原时间(PT)测试(也用于确定INR)、纤维蛋白原测试(通常通过克劳斯法进行)、血小板计数、血小板功能测试(通常通过PFA-100进行)、TCT、出血时间、混合测试(如果将患者的血浆与正常血浆混合,是否矫正异常)、凝结因子测定、抗磷脂抗体、D-二聚体、遗传测试(例如,因子V Leiden、凝血酶原突变G20210A)、稀释鲁塞尔蝰蛇毒时间(dRVVT)、杂项血小板功能测试、凝血弹性描记法(TEG或Sonoclot)、凝血弹性测量法(TEM®,例如ROTEM®)或优球蛋白溶解时间(ELT)。

[0154] aPTT测试是测量"固有"(也称为接触激活途径)和常见凝结途径二者的功效的性能指示物。这个测试一般用于测量市售重组凝血因子(例如,FVIII)的凝血活性。其与测量外在途径的凝血酶原时间(PT)结合使用。

[0155] ROTEM分析提供关于止血的整体动力学的信息:凝血时间、凝块形成、凝块稳定性和溶解。凝血弹性测量法中的不同参数依赖于血浆凝结系统的活性、血小板功能、纤维蛋白溶解或影响这些相互作用的许多因素。这个分析可以提供次级止血的全面见解。

[0156] 生色测定机制是基于血液凝结级联的原理,其中在激活的因子IX、磷脂和钙离子的存在下,激活的FVIII加速因子X转化为因子Xa。因子Xa活性是通过水解对因子Xa具有特异性的对硝基苯胺(pNA)底物来评估。在405nM下测量的对硝基苯胺的初始释放速率与因子Xa活性成正比,并且因此与样品中的FVIII活性成正比。生色测定得到了国际血栓形成与止血学会(ISTH)的科学标准化委员会(SSC)的FVIII和因子IX小组委员会的推荐。从1994年以来,生色测定也已经是欧洲药典中用于指定FVIII浓缩物效力的参考方法。

[0157] 在一些实施方案中,基因盒包含编码FVIII多肽的核苷酸序列,其中核苷酸序列经密码子优化。在一些实施方案中,基因盒包含由mTTR启动子驱动、编码密码子优化的FVIII的核苷酸序列和合成内含子。在一些实施方案中,基因盒包含国际申请号PCT/US2017/015879中披露的核苷酸序列,将所述申请通过引用以其整体并入。在一些实施方案中,基因盒是如PCT/US2017/015879中所述的"hFVIIIco6XTEN"基因盒。在一些实施方案中,基因盒包含SEQ ID NO:32。

[0158] 在一些实施方案中,基因盒包含经密码子优化的cDNA,其编码与XTEN 144肽融合

的B结构域缺失(BDD)的经密码子优化的人因子VIII(BDDcoFVIII)。在一些实施方案中,基因盒包含SEQ ID NO:9所示的核苷酸序列。在一些实施方案中,基因盒包含SEQ ID NO:33所示的核苷酸序列。在一些实施方案中,基因盒包含SEQ ID NO:14所示的核苷酸序列。在一些实施方案中,基因盒具有SEQ ID NO:14的核苷酸序列。在一些实施方案中,核酸分子包含SEQ ID NO:35的核苷酸序列。

[0159] 在一些实施方案中,基因盒包含由mTTR启动子驱动、编码密码子优化的FVIII的核苷酸序列。在一些实施方案中,基因盒还包含A1MB2增强子元件。在一些实施方案中,基因盒还包含嵌合或合成内含子。在一些实施方案中,基因盒还包含土拨鼠转录后调节元件(WPRE)。在一些实施方案中,基因盒还包含牛生长激素多腺苷酸化(bGHpA)信号。

[0160] 在一些实施方案中,本公开文本涉及编码具有FVIII活性的多肽的经密码子优化的核酸分子。在一些实施方案中,多核苷酸编码全长FVIII多肽。在其他实施方案中,核酸分子编码B结构域缺失的(BDD)FVIII多肽,其中FVIII的B结构域的全部或一部分缺失。

[0161] 在其他实施方案中,通过去除一个或多个CpG基序和/或甲基化至少一个CpG基序来进一步优化本文公开的核酸分子。如本文所用,“CpG基序”是指含有通过磷酸键与鸟苷连接的非甲基化胞嘧啶的二核苷酸序列。术语“CpG基序”包括甲基化和非甲基化CpG二核苷酸。非甲基化CpG基序在细菌和病毒来源的核酸(例如,质粒DNA)中是常见的,但在脊椎动物DNA中被抑制并大部分甲基化。因此,非甲基化CpG基序刺激哺乳动物宿主产生快速炎症反应。Klinman,等人(1996).PNAS 93:2879-2883。在Yew,N.S.,等人(2002).Mol Ther.5(6):731-738和国际申请号PCT/US2001/010309中描述了CpG去除的示例性方法。在一些实施方案中,本文公开的核酸分子已被修饰为含有更少的CpG基序(即CpG减少或CpG缺失)。在一个实施方案中,位于所选氨基酸的密码子三联体内的CpG基序被改变为缺少CpG基序的相同氨基酸的密码子三联体。在一些实施方案中,本文公开的核酸分子已被优化以降低先天免疫应答。

[0162] 在一个特定实施方案中,核酸分子编码包含与SEQ ID NO:10具有至少约80%、至少约85%、至少约86%、至少约87%、至少约88%、至少约89%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的氨基酸序列的多肽或其片段。在一些实施方案中,本公开文本的核酸分子编码包含信号肽的FVIII多肽或其片段。在其他实施方案中,核酸分子编码缺少信号肽的FVIII多肽。在一些实施方案中,所述信号肽包含SEQ ID NO:11的氨基酸序列。在一些实施方案中,所述信号肽包含SEQ ID NO:10的氨基酸1-19。

[0163] 在一些实施方案中,核酸分子包含第一ITR、第二ITR和编码靶序列的基因盒,其中所述靶序列编码治疗性蛋白质,并且其中所述治疗性蛋白质包含生长因子。生长因子可以选自本领域中已知的任何生长因子。在一些实施方案中,生长因子是激素。在其他实施方案中,生长因子是细胞因子。在一些实施方案中,生长因子是趋化因子。

[0164] 在一些实施方案中,生长因子是肾上腺髓质素(AM)。在一些实施方案中,生长因子是血管生成素(Ang)。在一些实施方案中,生长因子是自分泌运动因子。在一些实施方案中,生长因子是骨形态发生蛋白(BMP)。在一些实施方案中,BMP选自BMP2、BMP4、BMP5和BMP7。在一些实施方案中,生长因子是睫状神经营养因子家族成员。在一些实施方案中,睫状神经营养因子家族成员选自睫状神经营养因子(CNTF)、白血病抑制因子(LIF)、白介素-6(IL-6)。

在一些实施方案中,生长因子是集落刺激因子。在一些实施方案中,集落刺激因子选自巨噬细胞集落刺激因子(m-CSF)、粒细胞集落刺激因子(G-CSF)和粒细胞巨噬细胞集落刺激因子(GM-CSF)。在一些实施方案中,生长因子是表皮生长因子(EGF)。在一些实施方案中,生长因子是肝配蛋白。在一些实施方案中,肝配蛋白选自肝配蛋白A1、肝配蛋白A2、肝配蛋白A3、肝配蛋白A4、肝配蛋白A5、肝配蛋白B1、肝配蛋白B2和肝配蛋白B3。在一些实施方案中,生长因子是促红细胞生成素(EPO)。在一些实施方案中,生长因子是成纤维细胞生长因子(FGF)。在一些实施方案中,FGF选自FGF1、FGF2、FGF3、FGF4、FGF5、FGF6、FGF7、FGF8、FGF9、FGF10、FGF11、FGF12、FGF13、FGF14、FGF15、FGF16、FGF17、FGF18、FGF19、FGF20、FGF21、FGF22和FGF23。在一些实施方案中,生长因子是胎牛促体素(FBS)。在一些实施方案中,生长因子是GDNF家族成员。在一些实施方案中,GDNF家族成员选自神经胶质细胞系衍生的神经营养因子(GDNF)、神经秩蛋白、普塞芬蛋白和阿特敏蛋白。在一些实施方案中,生长因子是生长分化因子-9(GDF9)。在一些实施方案中,生长因子是肝细胞生长因子(HGF)。在一些实施方案中,生长因子是肝细胞瘤衍生的生长因子(HDGF)。在一些实施方案中,生长因子是胰岛素。在一些实施方案中,生长因子是胰岛素样生长因子。在一些实施方案中,胰岛素样生长因子是胰岛素样生长因子-1(IGF-1)或IGF-2。在一些实施方案中,生长因子是白介素(IL)。在一些实施方案中,IL选自IL-1、IL-2、IL-3、IL-4、IL-5、IL-6和IL-7。在一些实施方案中,生长因子是角质形成细胞生长因子(KGF)。在一些实施方案中,生长因子是迁移刺激因子(MSF)。在一些实施方案中,生长因子是巨噬细胞刺激蛋白(MSP或肝细胞生长因子样蛋白(HGFLP)。在一些实施方案中,生长因子是肌生成抑制蛋白(GDF-8)。在一些实施方案中,生长因子是神经调节蛋白。在一些实施方案中,神经调节蛋白选自神经调节蛋白1(NRG1)、NRG2、NRG3和NRG4。在一些实施方案中,生长因子是神经营养蛋白。在一些实施方案中,生长因子是脑衍生的神经营养因子(BDNF)。在一些实施方案中,生长因子是神经生长因子(NGF)。在一些实施方案中,NGF是神经营养蛋白-3(NT-3)或NT-4。在一些实施方案中,生长因子是胎盘生长因子(PGF)。在一些实施方案中,生长因子是血小板衍生的生长因子(PDGF)。在一些实施方案中,生长因子是肾胺酶(RNLS)。在一些实施方案中,生长因子是T细胞生长因子(TCGF)。在一些实施方案中,生长因子是血小板生成素(TPO)。在一些实施方案中,生长因子是转化生长因子。在一些实施方案中,转化生长因子是转化生长因子α(TGF-α)或TGF-β。在一些实施方案中,生长因子是肿瘤坏死因子-α(TNF-α)。在一些实施方案中,生长因子是血管内皮生长因子(VEGF)。

C.表达控制序列

[0165] 在一些实施方案中,本公开文本的核酸分子或载体还包含至少一个表达控制序列。例如,本公开文本的分离的核酸分子可以与至少一个表达控制序列可操作地连接。例如,表达控制序列可以是启动子序列或启动子-增强子组合。

[0166] 组成型哺乳动物启动子包括但不限于以下基因的启动子:次黄嘌呤磷酸核糖基转移酶(HPRT)、腺苷脱氨酶、丙酮酸激酶、β-肌动蛋白启动子和其他组成型启动子。在真核细胞中组成性地发挥作用的示例性病毒启动子包括例如来自以下病毒的启动子:巨细胞病毒(CMV)、猿猴病毒(例如,SV40)、乳头状瘤病毒、腺病毒、人免疫缺陷病毒(HIV)、劳斯肉瘤病毒、巨细胞病毒、莫洛尼白血病病毒的长末端重复(LTR)和其他逆转录病毒、以及单纯疱疹病毒的胸苷激酶启动子。其他组成型启动子是本领域普通技术人员已知的。可用作本公开

文本的基因表达序列的启动子还包括诱导型启动子。诱导型启动子是在诱导剂的存在下表达。例如,在某些金属离子的存在下诱导金属硫蛋白启动子以促进转录和翻译。其他诱导型启动子是本领域普通技术人员已知的。

[0167] 在一个实施方案中,本公开文本包括转基因在组织特异性启动子和/或增强子控制下的表达。在另一个实施方案中,启动子或其他表达控制序列选择性地增强转基因在肝细胞中的表达。在某些实施方案中,启动子或其他表达控制序列选择性地增强转基因在肝细胞、窦状细胞和/或内皮细胞中的表达。在一个特定实施方案中,启动子或其他表达控制序列选择性地增强转基因在内皮细胞中的表达。在某些实施方案中,启动子或其他表达控制序列选择性地增强转基因在以下中的表达:肌肉细胞、中枢神经系统、眼睛、肝脏、心脏或其任何组合。肝特异性启动子的例子包括但不限于小鼠甲状腺素转运蛋白启动子(mTTR)、天然人因子VIII启动子、人α-1-抗胰蛋白酶启动子(hAAT)、人白蛋白最小启动子和小鼠白蛋白启动子。在一些实施方案中,本文公开的核酸分子包含mTTR启动子。mTTR启动子描述于Costa等人(1986)Mol.Cell.Biol.6:4697中。FVIII启动子描述于Figueiredo和Brownlee,1995,J.Biol.Chem.270:11828-11838中。在一些实施方案中,启动子选自肝特异性启动子(例如,α1-抗胰蛋白酶(AAT))、肌肉特异性启动子(例如,肌肉肌酸激酶(MCK)、肌球蛋白重链α(αMHC)、肌红蛋白(MB)和结蛋白(DES))、合成启动子(例如,SPc5-12、2R5Sc5-12、dMCK和tMCK)或其任何组合。

[0168] 在一些实施方案中,转基因表达靶向肝脏。在某些实施方案中,转基因表达靶向肝细胞。在其他实施方案中,转基因表达靶向内皮细胞。在一个特定实施方案中,转基因表达靶向天然表达内源FVIII的任何组织。在一些实施方案中,转基因表达靶向中枢神经系统。在某些实施方案中,转基因表达靶向神经元。在一些实施方案中,转基因表达靶向传入神经元。在一些实施方案中,转基因表达靶向传出神经元。在一些实施方案中,转基因表达靶向中间神经元。在一些实施方案中,转基因表达靶向神经胶质细胞。在一些实施方案中,转基因表达靶向星形胶质细胞。在一些实施方案中,转基因表达靶向少突胶质细胞。在一些实施方案中,转基因表达靶向小神经胶质细胞。在一些实施方案中,转基因表达靶向室管膜细胞。在一些实施方案中,转基因表达靶向许旺细胞。在一些实施方案中,转基因表达靶向卫星细胞。在一些实施方案中,转基因表达靶向肌肉组织。在一些实施方案中,转基因表达靶向平滑肌。在一些实施方案中,转基因表达靶向心肌。在一些实施方案中,转基因表达靶向骨骼肌。在一些实施方案中,转基因表达靶向眼睛。在一些实施方案中,转基因表达靶向感光细胞。在一些实施方案中,转基因表达靶向视网膜神经节细胞。

[0169] 可用于本文公开的核酸分子的其他启动子包括小鼠甲状腺素转运蛋白启动子(mTTR)、天然人FVIII启动子、人α-1-抗胰蛋白酶启动子(hAAT)、人白蛋白最小启动子、小鼠白蛋白启动子、三重四脯氨酸(TTP)启动子、CASI启动子、CAG启动子、巨细胞病毒(CMV)启动子、α1-抗胰蛋白酶(AAT)启动子、肌肉肌酸激酶(MCK)启动子、肌球蛋白重链α(αMHC)启动子、肌红蛋白(MB)启动子、结蛋白(DES)启动子、SPc5-12启动子、2R5Sc5-12启动子、dMCK启动子、tMCK启动子、磷酸甘油酸激酶(PGK)启动子或人α-1-抗胰蛋白酶(A1AT)启动子或其任何组合。

[0170] 在一些实施方案中,本文公开的核酸分子包含甲状腺素转运蛋白(TTR)启动子。在一些实施方案中,启动子是小鼠甲状腺素转运蛋白(mTTR)启动子。mTTR启动子的非限制性

例子包括mTTR202启动子、mTTR202opt启动子和mTTR482启动子,如美国公开号US2019/0048362中披露,其通过引用以其整体并入。在一些实施方案中,启动子是肝特异性修饰小鼠甲状腺素转运蛋白(mTTR)启动子。在一些实施方案中,启动子是肝特异性修饰小鼠甲状腺素转运蛋白(mTTR)启动子mTTR482。mTTR482启动子的例子描述于Kyostio-Moore等人(2016)Mol Ther Methods Clin Dev.3:16006,和Nambiar B.等人(2017)Hum Gene Ther Methods,28(1):23-28中。在一些实施方案中,启动子是包含SEQ ID NO:16的核酸序列的肝特异性修饰小鼠甲状腺素转运蛋白(mTTR)启动子。在一些实施方案中,组织特异性启动子是人α-1-抗胰蛋白酶(A1AT)启动子。在一些实施方案中,组织特异性启动子包含SEQ ID NO:36的核苷酸序列。

[0171] 可以使用一种或多种增强子元件进一步增强表达水平以实现治疗功效。一种或多种增强子可以单独提供,或与一种或多种启动子元件一起提供。通常,表达控制序列包含多个增强子元件和组织特异性启动子。在一个实施方案中,增强子包含一个或多个拷贝的α-1-微球蛋白/双库尼茨抑制剂(bikunin)增强子(Rouet等人(1992)J.Biol.Chem.267:20765-20773;Rouet等人(1995),Nucleic Acids Res.23:395-404;Rouet等人(1998)Biochem.J.334:577-584;Ill等人(1997)Blood Coagulation Fibrinolysis 8:S23-S30)。在一些实施方案中,增强子衍生自肝特异性转录因子结合位点,如EBP、DBP、HNF1、HNF3、HNF4、HNF6和Enh1,包括HNF1、(有义)-HNF3、(有义)-HNF4、(反义)-HNF1、(反义)-HNF6、(有义)-EBP、(反义)-HNF4(反义)。

[0172] 在一些实施方案中,增强子元件包含一个或两个修饰的凝血酶原增强子(pPrT2)、一个或两个α1-微双库尼茨抑制剂(microbikunin)增强子(A1MB2)、修饰的小鼠白蛋白增强子(mEalb)、乙型肝炎病毒增强子II(HE11)或CRM8增强子。在一些实施方案中,A1MB2增强子是国际申请号PCT/US2019/055917中披露的增强子。在一些实施方案中,增强子元件是A1MB2。在一些实施方案中,增强子元件包括A1MB2增强子序列的多个拷贝。在一些实施方案中,A1MB2增强子定位于编码FVIII多肽的核酸序列的5'。在一些实施方案中,A1MB2增强子定位于启动子序列如mTTR启动子的5'。在一些实施方案中,增强子元件是包含SEQ ID NO:15的核酸序列的A1MB2增强子。

[0173] 在一些实施方案中,本文公开的核酸分子包含内含子或内含子序列。在一些实施方案中,内含子序列是天然存在的内含子序列。在一些实施方案中,内含子序列是合成序列。在一些实施方案中,内含子序列衍生自天然存在的内含子序列。在一些实施方案中,内含子序列是杂合合成内含子或嵌合内含子。在一些实施方案中,内含子序列是嵌合内含子,所述嵌合内含子由鸡β-肌动蛋白/兔β-珠蛋白内含子组成并且已经修饰以消除现有五个ATG序列,从而减少假翻译起始。在一些实施方案中,嵌合内含子包含SEQ ID NO:17的核酸序列。在一些实施方案中,内含子序列定位于编码FVIII多肽的核酸序列的5'。在一些实施方案中,嵌合内含子定位于启动子序列如mTTR启动子的5'。

[0174] 在一些实施方案中,本文公开的核酸分子包含转录后调节元件。在某些实施方案中,转录后调节元件包含突变的土拨鼠肝炎病毒调节元件(WPRE)。WPRE被认为增强病毒载体递送的转基因的表达。WPRE的例子描述于Zufferey等人(1999)J Virol.,73(4):2886-2892;Loeb等人(1999)Hum Gene Ther.10(14):2295-2305中。在一些实施方案中,WPRE定位于编码FVIII多肽的核酸序列的3'。在一些实施方案中,WPRE包含SEQ ID NO:18的核酸序

列。

[0175] 在一些实施方案中,本文公开的核酸分子包含转录终止子。在一些实施方案中,转录终止子是多腺苷酸化(poly(A))序列。转录终止子的非限制性例子包括衍生自牛生长激素多腺苷酸化信号(BGHpA)、猿猴病毒40多腺苷酸化信号(SV40pA)或合成多腺苷酸化信号的那些。在一个实施方案中,3'UTR多聚(A)尾包含肌动蛋白多聚(A)位点。在一个实施方案中,3'UTR多聚(A)尾包含血红蛋白多聚(A)位点。在一些实施方案中,转录终止子是BGHpA。Woychik等人(1984)PNAS 81:3944-3948描述了BGHpA转录终止子的例子。在一些实施方案中,转录终止子定位于对编码FVIII多肽的核酸序列进行编码的基因盒的3'端。在一些实施方案中,转录终止子是包含SEQ ID NO:19的核酸序列的BGHpA。

[0176] 在一些实施方案中,文中公开的核酸分子包含一个或多个DNA核靶向序列(DTS)。DTS促进含有此类序列的DNA分子易位至核中。在某些实施方案中,DTS包含SV40增强子序列。在某些实施方案中,DTS包含c-Myc增强子序列。在一些实施方案中,核酸分子包含位于第一ITR与第二ITR之间的DTS。在一些实施方案中,核酸分子包含位于第一ITR的3'和转基因(例如FVIII蛋白)的5'的DTS。在一些实施方案中,核酸分子包含位于核酸分子上转基因3'和第二ITR 5'的DTS。

[0177] 在一些实施方案中,本文公开的核酸分子包含toll样受体9(TLR9)抑制序列。示例性TLR9抑制序列描述于例如Trieu等人(2006)Crit Rev Immunol.26(6):527-44;Ashman等人Int'l Immunology 23(3):203-14中。

[0178] 在一些实施方案中,本文公开的核酸分子包含编码HBoV1的非结构蛋白的核酸序列。"非结构蛋白"是指由HBoV1表达的六种蛋白即NS1、NS1-70、NS2、NS3、NS4和NP1中的任一种。非结构蛋白由mRNA转录物表达,所述mRNA转录物通过单个病毒前体mRNA的选择性剪接和多腺苷酸化产生。NS1至NS4蛋白被编码于相同开放阅读框(ORF)的不同区域中。NS1结合HBoV1复制起点,并且推测在滚卡复制(rolling-hairpin replication)过程中对起点的单链DNA(ssDNA)产生切口。NS1在真核细胞中HBoV1 ITR介导的载体生产的表达中起重要作用。在一个实施方案中,产生表达HBoV1 NS1的表达构建体。在一些实施方案中,本文公开的核酸分子编码Shen等人(2015)J Virology 89(19):10097-10109中描述的非结构蛋白。

[0179] 在一些实施方案中,核酸分子包含微小RNA(miRNA)结合位点。在一个实施方案中,miRNA结合位点是miR-142-3p的miRNA结合位点。在其他实施方案中,miRNA结合位点是Rennie等人(2016)RNA Biol.13(6):554-560描述的miRNA结合位点。

杆状病毒中ceDNA的产生

[0180] 杆状病毒是最突出的感染昆虫的病毒。已鉴定出超过500种杆状病毒分离株,其中大部分源自鳞翅目(Lepidoptera)的昆虫。两种最常见的分离株是苜蓿银纹夜蛾(Autographa californica)多核多角体病毒(AcMNPV)和家蚕(Bombyx mori)核多角体病毒(BmNPV)。在表达载体中,杆状病毒因其超大的基因载荷容量(高达几十kb,有一些报道为高达100kb)而脱颖而出。这种转基因容量已经用于产生重组AAV载体(高达38kb表达盒)。然而,在生产用于基因疗法的病毒或非病毒载体时,通常需要将几种杆状病毒表达载体感染到昆虫宿主细胞中。每种杆状病毒表达载体的产生是耗时的,并且增加了生产成本,这代表了大多数杆状病毒表达载体系统的显著缺点。然而,产生了一种新的多功能杆状病毒穿梭载体(杆粒),其被专门设计用于容纳多个转基因(其可使用现有杆粒工具实现)。这种多功

能杆粒(称为“BIVVBac”)还可用于体内基因疗法的rAAV载体产生,以及任何所需蛋白质(例如重组蛋白质)的产生。该杆粒表达系统进一步描述于美国专利申请号63/069,073中,将其通过引用以其整体特此并入。

[0181] 在某些实施方案中,使用包含“BIVVBac”重组杆粒的杆状病毒表达载体系统产生所公开的核酸分子。在某些实施方案中,BIVVBac是经基因修饰的AcMNPV,其包含至少两个外来序列插入位点。包含含有至少两个外来序列插入位点的杆粒的杆状病毒表达载体系统允许减少需要产生的杆状病毒表达载体的总数。

[0182] 在某些实施方案中,BIVVBac包含第一外来序列插入位点和第二外来序列插入位点。第一外来序列插入位点和第二外来序列插入位点可以是不同的,从而利用不同的机制来驱动外来序列(例如,异源序列、异源基因)的插入。外来序列的插入可以由本领域已知的任何方法驱动。例如,可以通过转座或位点特异性重组来插入外来序列。可以将外来序列插入位点设计为包含在报告基因内,使得在插入外来序列后,报告基因被破坏。报告基因的破坏可有助于鉴定其中插入了外来序列的杆粒克隆。在此类实施方案中,将外来序列插入位点与报告基因框内融合,或者将报告基因与外来序列插入位点框内融合。

[0183] 在某些实施方案中,第一外来序列插入位点允许经由转座插入外来序列。在某些实施方案中,第一外来序列插入位点包含用于插入转座子的优先靶位点。在某些实施方案中,第一外来序列插入位点是用于插入转座子的优先靶位点。在某些实施方案中,第一外来序列插入位点是作为细菌转座子的附着位点的优先靶位点。合适的细菌转座子及其相应的附着位点是本领域技术人员已知的。例如,转座子Tn7以其高频率转座到细菌染色体的特定位点(attTn7)的能力而闻名。因此,在某些实施方案中,第一外来序列插入位点是作为Tn7转座子的附着位点的优先靶位点(例如,attTn7)。在一些实施方案中,第一外来序列插入位点是作为mini-Tn7转座子的附着位点的优先靶位点(例如,mini-attTn7,即Tn7转座因子识别和插入Tn7转座子所需的最小DNA序列)。

[0184] 在某些实施方案中,第二外来序列插入位点允许经由位点特异性重组插入外来序列。在某些实施方案中,第二外来序列插入位点包含能够介导位点特异性重组事件的优先靶位点。各种位点特异性重组酶技术是本领域技术人员已知的。例如,Cre-loxP系统经由Cre重组酶介导位点特异性重组,所述Cre重组酶能够识别称为loxP位点的34个碱基对的DNA序列。因此,第二外来序列插入位点是用于Cre介导的重组的优先靶位点。在某些实施方案中,第二外来序列插入位点是包含能够被Cre重组酶识别的loxP位点或其变体的优先靶位点。

[0185] 在一些实施方案中,重组杆粒包含变体VP80基因,使得杆粒展现出其编码的蛋白质的降低的表达。例如,本文公开了包含由于VP80基因座中的插入和/或缺失而失活的VP80基因的杆状病毒DNA骨架。在一些实施方案中,重组杆粒包含美国专利申请号US 63/069,115中披露的杆粒。

[0186] 在某些实施方案中,使用单个杆状病毒表达载体产生ceDNA。在这种“单BAC”方法中,单一杆状病毒表达载体(例如,BIVVBac)编码在杆状病毒系统中产生ceDNA所需的所有必需元件,并且有可能在任何杆状病毒容许性细胞系中用于产生ceDNA。这种方法描绘于图1A中。

[0187] 在某些实施方案中,使用多种杆状病毒表达载体产生ceDNA。在这种“双BAC”方法

中,将ceDNA产生所需的必需元件插入到两种不同的杆状病毒(例如,两种BIVVBac杆粒)中,并且有可能用于在容许杆状病毒感染的任何细胞系中进行共感染。这种方法描绘于图1B中。

[0188] 在某些实施方案中,通过稳定细胞系产生ceDNA。在这种方法中,将产生ceDNA所需的必需元件插入在杆状病毒系统的两个组分中。该方法描绘于图1C中。稳定细胞系可以通过以下方式来产生:在杆状病毒基因启动子(例如,杆状病毒组成型基因启动子)的控制下稳定地整合蛋白质编码序列。在某些实施方案中,稳定细胞系是稳定的昆虫细胞系。

[0189] 用于将核酸稳定地整合到多种宿主细胞系中的方法在本领域中是已知的。例如,重复选择(例如,通过使用选择性标记)可以用于选择已整合含有选择性标记(和AAV cap和rep基因和/或rAAV基因组)的核酸的细胞。在其他实施方案中,核酸能以位点特异性方式整合到细胞系中以产生生产细胞系。若干种位点特异性重组系统是本领域已知的,如FLP/FRT(参见例如O'Gorman,S.等人(1991)Science 251:1351-1355)、Cre/loxP(参见例如Sauer,B.和Henderson,N.(1988)Proc.Natl.Acad.Sci.85:5166-5170)以及phi C31-att(参见例如Groth,A.C.等人(2000)Proc.Natl.Acad.Sci.97:5995-6000)。

[0190] 本公开文本还提供了由本公开文本的核酸分子编码的多肽。在一些实施方案中,本公开文本的多肽由包含本文公开的分离的核酸分子的载体编码。在又其他实施方案中,本公开文本的多肽由包含本文公开的分离的核酸分子的宿主细胞产生。

宿主细胞

[0191] 本公开文本还提供了一种包含本公开文本的核酸分子或载体的宿主细胞。如本文所用,术语"转化"应在广义上使用,是指将DNA引入受体宿主细胞中,这改变了基因型并因此导致受体细胞的变化。

[0192] "宿主细胞"是指已经用使用重组DNA技术构建并且编码至少一种异源基因的载体转化的细胞。本公开文本的宿主细胞优选地具有哺乳动物来源;最优选地具有人或小鼠来源。相信本领域技术人员有能力优先地确定最适合于其目的的特定宿主细胞系。示例性宿主细胞系包括但不限于CHO、DG44和DUXB11(中国仓鼠卵巢系,DHFR-)、HELA(人宫颈癌)、CVI(猴肾系)、COS(具有SV40 T抗原的CVI的衍生物)、R1610(中国仓鼠成纤维细胞)、BALBC/3T3(小鼠成纤维细胞)、HAK(仓鼠肾系)、SP2/0(小鼠骨髓瘤)、P3x63-Ag8.653(小鼠骨髓瘤)、BFA-1c1BPT(牛内皮细胞)、RAJI(人淋巴细胞)、PER.C6®、NS0、CAP、BHK21和HEK 293(人肾)。在一个特定实施方案中,宿主细胞选自:CHO细胞、HEK293细胞、BHK21细胞、PER.C6®细胞、NS0细胞、CAP细胞及其任何组合。在一些实施方案中,本公开文本的宿主细胞具有昆虫来源。在一个特定实施方案中,宿主细胞是SF9细胞。宿主细胞系典型地可从商业服务、美国组织培养物保藏中心(American Tissue Culture Collection)或公开的文献中获得。

[0193] 将本公开文本的核酸分子或载体引入宿主细胞中可以通过本领域技术人员熟知的多种技术来完成。这些技术包括但不限于转染(包括电泳和电穿孔)、原生质体融合、磷酸钙沉淀、与包膜DNA的细胞融合、显微注射和完整病毒感染。参见Ridgway,A.A.G."Mammalian Expression Vectors"第24.2章,第470-472页Vectors,Rodriguez和Denhardt编(Butterworths,Boston,Mass.1988)。最优选地,通过电穿孔将质粒引入宿主中。使转化的细胞在适合于产生轻链和重链的条件下生长,并且测定重链和/或轻链蛋白质合成。示例性测定技术包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)或荧光激活细胞分选仪分

析(FACS)、免疫组织化学等。

[0194] 使包含本公开文本的分离的核酸分子或载体的宿主细胞在适当生长培养基中生长。如本文所用,术语“适当生长培养基”是指含有细胞生长所需的营养物的培养基。细胞生长所需的营养素可以包括碳源、氮源、必需氨基酸、维生素、矿物质和生长因子。任选地,培养基可以含有一种或多种选择因子。任选地,培养基可以含有小牛血清或胎牛血清(FCS)。在一个实施方案中,培养基基本上不含IgG。生长培养基通常将例如通过药物选择或必需营养素的缺乏来选择含有DNA构建体的细胞,所述营养素通过DNA构建体上或与DNA构建体共转染的选择性标记来补充。所培养的哺乳动物细胞通常在市售的含血清或无血清培养基(例如,MEM、DMEM、DMEM/F12)中生长。在一个实施方案中,培养基是CDoptiCHO(Invitrogen,加利福尼亚州卡尔斯巴德)。在另一个实施方案中,培养基是CD17(Invitrogen,加利福尼亚州卡尔斯巴德)。适合于所用特定细胞系的培养基的选择在本领域普通技术人员的水平内。

[0195] 本公开文本的多个方面提供了一种克隆本文所述的核酸分子的方法,所述方法包括将能够形成复杂二级结构的核酸分子插入合适的载体中,并且将所得载体引入合适的细菌宿主菌株中。如本领域中已知,核酸的复杂二级结构(例如,长回文区)可能不稳定并且难以在细菌宿主菌株中克隆。例如,本公开文本的包含第一ITR和第二ITR(例如,非AAV细小病毒ITR,例如HBoV1 ITR)的核酸分子可能难以使用常规方法来克隆。长DNA回文序列抑制DNA复制并且在大肠杆菌(E.coli)、芽孢杆菌属(Bacillus)、链球菌属(Streptococcus)、链霉菌属(Streptomyces)、酿酒酵母(S.cerevisiae)、小鼠和人的基因组中不稳定。这些效应是由于通过链内碱基配对形成发夹或十字形结构所致。在大肠杆菌中,在SbcC或SbcD突变体中可以显著克服对DNA复制的抑制。SbcD是核酸酶亚基,并且SbcC是SbcCD复合物的ATP酶亚基。大肠杆菌SbcCD复合物是负责防止长回文序列复制的外切核酸酶复合物。SbcCD复合物是具有ATP依赖性双链DNA外切核酸酶活性和ATP非依赖性单链DNA内切核酸酶活性的核复合物。SbcCD可以识别DNA回文序列并通过攻击所产生的发夹结构来瓦解复制叉。

[0196] 在某些实施方案中,合适的细菌宿主菌株不能拆分十字形DNA结构。在某些实施方案中,合适的细菌宿主菌株包含SbcCD复合物中的破坏。在一些实施方案中,SbcCD复合物中的破坏包含SbcC基因和/或SbcD基因中的基因破坏。在某些实施方案中,SbcCD复合物中的破坏包含SbcC基因中的基因破坏。包含SbcC基因中的基因破坏的各种细菌宿主菌株是本领域中已知的。例如而不限于,细菌宿主菌株PMC103包含基因型sbcC、recD、mcrA、ΔmcrBCF;细菌宿主菌株PMC107包含基因型recBC、recJ、sbcBC、mcrA、ΔmcrBCF;并且细菌宿主菌株SURE包含基因型recB、recJ、sbcC、mcrA、ΔmcrBCF、umuC、uvrC。因此,在一些实施方案中,克隆本文所述的核酸分子的方法包括将能够形成复杂二级结构的核酸分子插入合适的载体中,并且将所得载体引入宿主菌株PMC103、PMC107或SURE中。在某些实施方案中,克隆本文所述的核酸分子的方法包括将能够形成复杂二级结构的核酸分子插入合适的载体中,并且将所得载体引入宿主菌株PMC103中。

[0197] 合适的载体是本领域已知的。在某些实施方案中,用于本公开文本的克隆方法中的合适的载体是低拷贝载体。在某些实施方案中,用于本公开文本的克隆方法中的合适的载体是pBR322。

[0198] 因此,本公开文本提供了克隆核酸分子的方法,其包括将能够形成复杂二级结构的核酸分子插入合适的载体中,和将所得载体引入包含SbcCD复合物中的破坏的细菌宿主

菌株中,其中核酸分子包含第一反向末端重复序列(ITR)和第二ITR,其中第一ITR和/或第二ITR包含与SEQ ID NO.1或2中所示的核苷酸序列至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或100%相同的核苷酸序列或其功能衍生物。

FVIII多肽的产生

[0199] 本公开文本还提供了由本公开文本的核酸分子编码的多肽。在一些实施方案中,本公开文本的多肽由包含本文公开的分离的核酸分子的载体编码。在又其他实施方案中,本公开文本的多肽由包含本公开文本的分离的核酸分子的宿主细胞产生。

[0200] 多种方法可用于从本公开文本的优化的核酸分子重组产生FVIII蛋白。所需序列的多核苷酸可通过从头固相DNA合成或通过PCR诱变早期制备的多核苷酸来产生。寡核苷酸介导的诱变是制备核苷酸序列中的取代、插入、缺失或改变(例如,改变密码子)的一种方法。例如,通过将编码所需突变的寡核苷酸与单链DNA模板杂交来改变起始DNA。杂交后,DNA聚合酶用于合成整合了寡核苷酸引物的模板的完整第二互补链。在一个实施方案中,基因工程例如基于引物的PCR诱变足以掺入本文所定义的改变以产生本公开文本的多核苷酸。

[0201] 对于重组蛋白生产,将本公开文本的编码FVIII蛋白的优化的多核苷酸序列插入适当的表达媒介物中,即含有所插入编码序列的转录和翻译所必需的元件的载体,或在RNA病毒载体的情况下为含有复制和翻译所必需的元件的载体。

[0202] 将本公开文本的多核苷酸序列插入载体的适当阅读框中。然后将表达载体转染到将表达多肽的合适靶细胞中。本领域已知的转染技术包括但不限于磷酸钙沉淀(Wigler等人1978,Cell 14:725)和电穿孔(Neumann等人1982,EMBO,J.1:841)。多种宿主表达载体系统可用于在真核细胞中表达本文所述的FVIII蛋白。在一个实施方案中,真核细胞是动物细胞,包括哺乳动物细胞(例如,HEK293细胞、PER.C6®、CHO、BHK、Cos、HeLa细胞)。本公开文本的多核苷酸序列还可编码允许FVIII蛋白分泌的信号序列。本领域技术人员将理解,当FVIII蛋白被翻译时,信号序列被细胞裂解以形成成熟蛋白。本领域已知各种信号序列,例如天然因子VII信号序列、天然因子IX信号序列和小鼠IgK轻链信号序列。可替代地,当不包括信号序列时,可通过裂解细胞回收FVIII蛋白。

[0203] 本公开文本的FVIII蛋白可以在转基因动物(如啮齿类动物、山羊、绵羊、猪或牛)体内合成。术语“转基因动物”是指已经将外来基因掺入基因组中的非人动物。因为这个基因是在种系组织中存在,其从亲代传递给后代。将外源基因引入单细胞胚胎中(Brinster等人1985,Proc.Natl.Acad.Sci.USA 82:4438)。产生转基因动物的方法是本领域已知的,包括产生免疫球蛋白分子的转基因(Wagner等人1981,Proc.Natl.Acad.Sci.USA 78:6376;McKnight等人1983,Cell 34:335;Brinster等人1983,Nature 306:332;Ritchie等人1984,Nature 312:517;Baldassarre等人2003,Theriogenology 59:831;Robl等人2003,Theriogenology 59:107;Malassagne等人2003,Xenotransplantation 10(3):267)。

[0204] 表达载体可以编码允许容易纯化或鉴定重组产生的蛋白质的标签。例子包括但不限于载体pUR278(Ruther等人1983,EMBO J.2:1791),其中本文所述的FVIII蛋白编码序列可以与lac Z编码区同框连接到载体中,从而产生杂合蛋白;pGEX载体可用于表达具有谷胱甘肽S-转移酶(GST)标签的蛋白质。这些蛋白质通常是可溶的,并且可以通过吸附到谷胱甘肽-琼脂糖珠上,随后在游离谷胱甘肽的存在下洗脱而容易地从细胞中纯化。载体包括切割

位点(例如,PreCission Protease(Pharmacia,新泽西州皮帕克)),以便在纯化后容易除去标签。

[0205] 出于本公开文本的目的,可以使用许多表达载体系统。这些表达载体典型地可作为附加体或作为宿主染色体DNA的组成部分在宿主生物中复制。表达载体可包括表达控制序列,包括但不限于启动子(例如天然关联的或异源启动子)、增强子、信号序列、剪接信号、增强子元件和转录终止序列。优选地,表达控制序列是能够转化或转染真核宿主细胞的载体中的真核启动子系统。表达载体还可以利用衍生自动物病毒的DNA元件,所述动物病毒如牛乳头瘤病毒、多瘤病毒、腺病毒、痘苗病毒、杆状病毒、逆转录病毒(RSV、MMTV或MOMLV)、巨细胞病毒(CMV)或SV40病毒。其他涉及具有内部核糖体结合位点的多顺反子系统的使用。

[0206] 一般来说,表达载体含有选择标记(例如氨苄青霉素抗性、潮霉素抗性、四环素抗性或新霉素抗性)以允许检测用所需DNA序列转化的那些细胞(参见例如,Itakura等人,美国专利4,704,362)。可以通过引入一种或多种标记来选择已将DNA整合至其染色体中的细胞,所述一种或多种标记允许选择经转染的宿主细胞。所述标记可以提供对营养缺陷型宿主的原营养、杀生物剂抗性(例如抗生素)或对重金属如铜的抗性。选择性标记基因可以直接与待表达的DNA序列连接,或通过共转化引入同一细胞中。

[0207] 可用于表达优化的FVIII序列的载体的例子是NEOSPLA(美国专利号6,159,730)。此载体含有巨细胞病毒启动子/增强子、小鼠β珠蛋白主要启动子、SV40复制起点、牛生长激素多腺苷酸化序列、新霉素磷酸转移酶外显子1和外显子2、二氢叶酸还原酶基因和前导序列。已发现此载体在掺入可变区和恒定区基因、细胞中转染、随后在含有G418的培养基中选择和甲氨蝶呤扩增后,会导致非常高水平的抗体表达。在美国专利号5,736,137和5,658,570中也传授了载体系统,这些文献各自通过引用以其整体并入本文。此系统提供高表达水平,例如>30pg/细胞/天。其他示例性载体系统披露于例如美国专利号6,413,777中。

[0208] 在其他实施方案中,可使用多顺反子构建体表达本公开文本的多肽。在这些表达系统中,可以从单个多顺反子构建体产生多个目的基因产物,如多聚体结合蛋白的多个多肽。这些系统有利地使用内部核糖体进入位点(IRES)以在真核宿主细胞中提供相对较高水平的多肽。在美国专利号6,193,980中披露了相容的IRES序列,该专利也并入本文。

[0209] 更一般地,一旦制备了编码多肽的载体或DNA序列,就可以将表达载体引入适当的宿主细胞中。也就是说,所述宿主细胞可被转化。如上所述,可以通过本领域技术人员熟知的各种技术来实现将质粒引入宿主细胞中。转化的细胞在适于产生FVIII多肽的条件下生长,并进行FVIII多肽合成测定。示例性测定技术包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)或荧光激活细胞分选仪分析(FACS)、免疫组织化学等。

[0210] 在描述用于从重组宿主分离多肽的过程时,术语“细胞”和“细胞培养物”可互换使用以表示多肽来源,除非另有明确说明。换言之,从“细胞”中回收多肽可以意指从离心沉淀的全细胞,或从含有培养基和悬浮细胞二者的细胞培养物中回收。

[0211] 用于蛋白质表达的宿主细胞系优选是哺乳动物来源的;最优选是人或小鼠来源的,因为本公开文本的分离的核酸已被优化用于在人细胞中表达。上文已经描述了示例性宿主细胞系。在产生具有FVIII活性的多肽的方法的一个实施方案中,所述宿主细胞是HEK293细胞。在产生具有FVIII活性的多肽的方法的另一个实施方案中,所述宿主细胞是CHO细胞。

[0212] 编码本公开文本的多肽的基因也可以在非哺乳动物细胞(如细菌或酵母或植物细胞)中表达。在这一方面,应当理解,也可以转化多种单细胞非哺乳动物微生物如细菌;即那些能够在培养或发酵中生长的那些微生物。易于转化的细菌包括以下的成员:肠杆菌科,如大肠杆菌(Escherichia coli)或沙门氏菌属(Salmonella)的菌株;芽孢杆菌科,如枯草芽孢杆菌(Bacillus subtilis);肺炎球菌属(Pneumococcus);链球菌属(Streptococcus)和流感嗜血杆菌(Haemophilus influenzae)。还应理解,当在细菌中表达时,多肽典型地会成为内含体的一部分。必须分离、纯化多肽,然后将其组装成功能性分子。

[0213] 可替代地,可将本公开文本的优化的核苷酸序列掺入转基因中以引入转基因动物的基因组中并随后在转基因动物的乳汁中表达(参见,例如Deboer等人,US 5,741,957,Rosen,US 5,304,489,和Meade等人,US 5,849,992)。合适的转基因包括与来自乳腺特异性基因的启动子和增强子可操作连接的多肽的编码序列(如酪蛋白或β乳球蛋白)。

[0214] 体外生产允许按比例放大以得到大量的所希望的多肽。在组织培养条件下用于哺乳动物细胞培养的技术是本领域已知的并且包括同质悬浮培养(例如在气升式反应器或连续搅拌反应器中),或在琼脂糖上微珠或陶瓷盒上的固定化或包埋的细胞培养(例如在中空纤维、微胶囊中)。如果必要和/或需要,可以通过常规色谱方法纯化多肽溶液,例如凝胶过滤,离子交换色谱,DEAE-纤维素色谱或(免疫)亲和色谱,例如,在合成铰链区多肽的优选生物合成后或在本文所述的HIC色谱步骤之前或之后。亲和标签序列(例如His(6)标签)可任选地附接或包含在多肽序列内以促进下游纯化。

[0215] 一旦表达后,所述FVIII蛋白可以根据本领域的标准操作纯化,包括硫酸铵沉淀、亲和柱色谱、HPLC纯化、凝胶电泳等(一般参见Scopes,Protein Purification(Springer-Verlag,N.Y.,(1982)))。对于药物用途,至少约90%至95%均一性的基本上纯的蛋白质是优选的,98%至99%或更高均一性是最优选的。

药物组合物

[0216] 含有本公开文本的分离的核酸分子、核酸分子编码的具有FVIII活性的多肽、载体或宿主细胞的组合物可以含有合适的药学上可接受的载体。例如,所述组合物可以含有有助于将活性化合物加工成设计用于递送至作用位点的制剂的赋形剂和/或助剂。

[0217] 药物组合物可以被配制用于通过推注注射的肠胃外施用(即静脉内、皮下或肌内)。注射用配制品能以单位剂型存在,例如在添加有防腐剂的安瓿中或多剂量容器中。组合物可以采取诸如于油性或水性媒介物中的悬浮液、溶液或乳液等形式,并且含有诸如助悬剂、稳定剂和/或分散剂等配制剂。可替代地,活性成分可以呈粉末形式,用于用合适的媒介物(例如,无热原水)来构造。

[0218] 适合于肠胃外施用的配制品还包括呈水溶性形式(例如,水溶性盐)的活性化合物的水溶液。另外,可以施用作为适当油性注射悬浮液的活性化合物的悬浮液。合适的亲脂溶剂或媒介物包括脂肪油(例如,芝麻油)或合成脂肪酸酯(例如,油酸乙酯或甘油三酯)。水性注射悬浮液可以含有增加悬浮液粘度的物质,包括例如羧甲基纤维素钠、山梨醇和葡聚糖。任选地,悬浮液还可以含有稳定剂。脂质体还可以用于包封本公开文本的分子,用于递送至细胞或间质空间中。示例性药学上可接受的载体是生理上相容的溶剂、分散介质、包衣、抗细菌和抗真菌剂、等渗和吸收延迟剂、水、盐水、磷酸盐缓冲盐水、右旋糖、甘油、乙醇等。在一些实施方案中,组合物包含等渗剂,例如糖、多元醇(如甘露醇、山梨醇)或氯化钠。在其他

实施方案中,组合物包含增强活性成分的贮存期限或有效性的药学上可接受的物质(如润湿剂)或少量辅助物质(如润湿剂或乳化剂、防腐剂或缓冲液)。

[0219] 本公开文本的组合物可以呈多种形式,包括例如液体(例如,可注射和可输注的溶液)、分散液、悬浮液、半固体和固体剂型。优选形式取决于施用方式和治疗应用。

[0220] 组合物可以配制为溶液、微乳液、分散液、脂质体或其他适合于高药物浓度的有序结构。无菌可注射溶液可以通过以下方式制备:将活性成分以所需量掺入视需要具有上文所列举成分中的一种或组合的适当溶剂中,之后过滤灭菌。通常,分散液是通过以下方式来制备:将活性成分掺入无菌媒介物中,所述无菌媒介物含有基础分散介质和来自上文所列举的那些成分的所需其他成分。在用于制备无菌可注射溶液的无菌粉末的情况下,优选制备方法是真空干燥和冷冻干燥,这些方法从先前无菌过滤的溶液产生活性成分和任何其他所需成分的粉末。可以例如通过以下方式维持溶液的适当流动性:通过使用诸如卵磷脂等包衣,在分散液的情况下通过维持所需粒度,以及通过使用表面活性剂。可注射组合物的延长吸收可以通过以下方式来实现:在组合物中包括延迟吸收的药剂,例如单硬脂酸盐和明胶。

[0221] 活性成分可以用受控释放配制品或装置来配制。此类配制品和装置的例子包括植入物、经皮贴剂和微囊化递送系统。可以使用生物可降解的生物相容性聚合物,例如乙烯乙酸乙烯酯、聚酐、聚乙醇酸、胶原蛋白、聚原酸酯和聚乳酸。制备此类配制品和装置的方法是本领域已知的。参见例如,Sustained and Controlled Release Drug Delivery Systems,J.R.Robinson编,Marcel Dekker,Inc.,纽约,1978。

[0222] 可注射长效配制品可以通过以下方式来制备:形成药物于生物可降解聚合物(如聚丙交酯-聚乙交酯)中的微囊化基质。根据药物与聚合物的比率以及所采用聚合物的性质,可以控制药物释放速率。其他示例性生物可降解聚合物是聚原酸酯和聚酐。可注射长效配制品还可以通过将药物捕集于脂质体或微乳液中来制备。

[0223] 可以将补充性活性化合物掺入组合物中。在一个实施方案中,用另一凝血因子或其变体、片段、类似物或衍生物配制本公开文本的嵌合蛋白。例如,凝血因子包括但不限于因子V、因子VII、因子VIII、因子IX、因子X、因子XI、因子XII、因子XIII、凝血酶原、纤维蛋白原、血管性血友病因子或重组可溶性组织因子(rsTF)或前述任一种的激活形式。凝血因子或止血剂还可以包括抗纤维蛋白溶解药,例如ε-氨基己酸、氨甲环酸。

[0224] 可以调节剂量方案以提供最佳所需反应。例如,可以施用单次推注,可以随时间施用若干次分开剂量,或者可以如治疗情况的紧迫性所示的成比例地减少或增加剂量。有利的是以剂量单位形式配制肠胃外组合物,以便于施用和剂量均匀。参见例如,Remington's Pharmaceutical Sciences(Mack Pub.Co.,宾夕法尼亚州伊斯顿1980)。

[0225] 除了活性化合物以外,液体剂型还可以含有惰性成分,如水、乙醇、碳酸乙酯、乙酸乙酯、苯甲醇、苯甲酸苄酯、丙二醇、1,3-丁二醇、二甲基甲酰胺、油、甘油、四氢糠醇、聚乙二醇和山梨聚糖的脂肪酸酯。

[0226] 合适的药物载体的非限制性例子也描述于E.W.Martin的Remington's Pharmaceutical Sciences中。赋形剂的一些例子包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、水稻、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、脱脂乳粉、甘油、丙烯、二醇、水、乙醇等。组合物还可以含有pH缓冲试剂以及润湿剂或乳化剂。

[0227] 对于口服施用,药物组合物可以采取通过常规手段制备的片剂或胶囊的形式。还可以将组合物制备为液体,例如糖浆或悬浮液。液体可以包括助悬剂(例如,山梨醇糖浆、纤维素衍生物或氢化食用脂肪)、乳化剂(卵磷脂或阿拉伯胶)、非水性媒介物(例如,扁桃仁油、油性酯、乙醇或分馏植物油)和防腐剂(例如,对羟基苯甲酸甲酯或对羟基苯甲酸丙酯或山梨酸)。制剂还可以包括调味剂、着色剂和甜味剂。可替代地,组合物可以作为干产品存在,用于用水或另一种合适的媒介物来构造。

[0228] 对于经颊施用,组合物可以采取根据常规方案的片剂或锭剂的形式。

[0229] 对于通过吸入施用,根据本公开文本使用的化合物便捷地以含有或不含赋形剂的雾化气溶胶的形式或以气溶胶喷雾剂的形式从任选地具有推进剂的加压包或雾化器递送,所述推进剂是例如二氯二氟甲烷、三氯氟甲烷、二氯四氟甲烷、二氧化碳或其他合适的气体。在加压气溶胶的情况下,剂量单位可以通过提供递送计量量的阀来确定。例如用于吸入器或吹入器中的明胶的胶囊和药筒可以配制含有化合物与诸如乳糖或淀粉等合适的粉末基质的粉末混合物。

[0230] 药物组合物还可以被配制用于作为例如含有常规栓剂基质(如可可脂或其他甘油酯)的栓剂或保留灌肠剂经直肠施用。

[0231] 在一个实施方案中,药物组合物包含具有因子VIII活性的多肽、编码所述具有因子VIII活性的多肽的经优化的核酸分子、包含所述核酸分子的载体、或包含所述载体的宿主细胞和药学上可接受的载体。在一些实施方案中,组合物是通过选自以下的途径来施用:局部施用、眼内施用、肠胃外施用、鞘内施用、硬膜下施用和口服施用。肠胃外施用可以是静脉内或皮下施用。

治疗方法

[0232] 在一些方面中,本公开文本涉及治疗有需要的受试者的疾病或病症的方法,其包括施用本文所公开的核酸分子、载体、多肽或药物组合物。

[0233] 在一些实施方案中,本公开文本涉及治疗出血障碍的方法。在一些实施方案中,本公开文本涉及治疗血友病A的方法。

[0234] 分离的核酸分子、载体或多肽可以用以下方式施用:静脉内、皮下、肌内或通过任何粘膜表面,例如口服、舌下、经颊、舌下、经鼻、经直肠、经阴道或通过经肺途径。可以将凝血因子蛋白植入生物聚合物固相支持物内或与所述固相支持物连接,从而允许将嵌合蛋白缓慢释放至所需部位。

[0235] 在一个实施方案中,分离的核酸分子、载体或多肽的施用途径是肠胃外的。如本文所用的术语肠胃外包括静脉内、动脉内、腹膜内、肌内、皮下、直肠或阴道施用。在一些实施方案中,将分离的核酸分子、载体或多肽静脉内施用。虽然所有这些施用形式都明确地考虑在本公开文本的范围内,但是施用形式可以是注射用溶液,特别是用于静脉内或动脉内注射或滴注。

[0236] 用于治疗病症的本公开文本的组合物的有效剂量根据许多不同因素而变化,所述因素包括施用方式、靶位点、患者的生理状态、患者是人还是动物、施用的其他药物以及治疗是预防性的还是治疗性的。通常,患者是人,但是也可以治疗非人哺乳动物,包括转基因哺乳动物。治疗剂量可以使用本领域技术人员已知的常规方法逐步调整,以优化安全性和功效。

[0237] 本公开文本的核酸分子、载体或多肽可以任选地与在需要治疗(例如,预防性或治疗性)的障碍或病症的治疗中有效的其他药剂组合施用。

[0238] 如本文所用,与辅助疗法结合或组合的本公开文本的分离的核酸分子、载体或多肽的施用意指依序、同时、同延、并行、伴随或同期施用或应用所述疗法和所公开的多肽。本领域技术人员将理解,可以定时施用或施加组合的治疗性方案的各种组分以增强治疗的总体有效性。基于所选辅助疗法和本说明书的传授内容,技术人员(例如,医师)在不进行过度实验的情况下便能够容易地辨别有效的组合治疗方案。

[0239] 应进一步了解,本公开文本的分离的核酸分子、载体或多肽可以与一种或多种药剂结合或组合使用(例如,以提供组合治疗方案)。可以与本公开文本的多肽或多核苷酸组合的示例性药剂包括代表用于所治疗的特定障碍的当前护理标准的药剂。此类药剂在本质上可以是化学的或生物的。术语“生物剂(biologic)”或“生物剂(biologic agent)”是指预期用作治疗剂的从活的生物体和/或其产物制备的任何药学活性剂。

[0240] 要与本公开文本的多核苷酸或多肽组合使用的药剂的量可以随受试者变化,或者可以根据本领域的知识来施用。参见例如,Bruce A Chabner等人,Antineoplastic Agents,Goodman&Gilman's The Pharmacological Basis of Therapeutics 1233-1287(Joel G.Hardman等人编,第9版1996)。在另一个实施方案中,施用符合护理标准的量的这种药剂。

[0241] 在一个实施方案中,本文还公开了一种试剂盒,所述试剂盒包含本文公开的核酸分子以及向有需要的受试者施用所述核酸分子的说明书。在另一个实施方案中,本文公开了一种用于产生本文所提供的核酸分子的杆状病毒系统。所述核酸分子是在昆虫细胞中产生。在另一个实施方案中,提供了一种针对表达构建体的纳米颗粒递送系统。表达构建体包含本文公开的核酸分子。

基因疗法

[0242] 在一些实施方案中,本文公开的核酸分子用于基因疗法。本文公开的优化的FVIII核酸分子可用于需要表达FVIII的任何情境。在一些实施方案中,核酸分子包含SEQ ID NO:2的核苷酸序列。在一些实施方案中,核酸分子包含SEQ ID NO:1的核苷酸序列。

[0243] 例如,已经探索出体基因疗法作为血友病A的可能治疗。基因疗法是血友病的特别有吸引力的治疗,因为其可能在载体的单次施用后通过连续内源性产生FVIII来治愈疾病。血友病A非常适合于基因替代方法,因为其临床表现可完全归因于缺少以微小量(200ng/ml)在血浆中循环的单一基因产物(FVIII)。

[0244] 在一方面,本文所述的核酸分子可用于AAV基因疗法。AAV能够感染许多哺乳动物细胞。参见例如,Tratschin等人(1985)Mol.Cell Biol.5:3251-3260和Grimm等人(1999)Hum.Gene Ther.10:2445-2450。rAAV载体携带编码目的基因或其片段的核酸序列,其处于指导基因产物在细胞中表达的调节序列的控制下。在一些实施方案中,将rAAV与载体和适于施用的其他组分一起配制。

[0245] 在另一方面,本文所述的核酸分子可用于慢病毒基因疗法。慢病毒是RNA病毒,其中病毒基因组是RNA。当宿主细胞被慢病毒感染时,基因组RNA被逆转录成DNA中间体,所述中间体被非常有效地整合到所感染细胞的染色体DNA中。在一些实施方案中,将慢病毒与载体和适于施用的其他组分一起配制。在另一方面,本文所述的核酸分子可用于腺病毒疗法。

腺病毒用于基因疗法的综述可参见例如Wold等人(2013)Curr Gene Ther.13(6):421-33)。在另一方面,本文所述的核酸分子可用于非病毒基因疗法。本公开文本的优化的FVIII蛋白可以在哺乳动物(例如,人患者)体内产生,使用基因治疗方法对于选自以下的出血疾病或障碍的治疗将是治疗上有益的:出血凝结障碍、关节积血、肌肉出血、口腔出血、出血症、出血至肌肉中、口腔出血症、创伤、头部创伤、胃肠出血、颅内出血、腹内出血、胸腔内出血、骨折、中枢神经系统出血、咽后间隙出血、腹膜后间隙出血和髂腰肌鞘出血。在一个实施方案中,出血疾病或障碍是血友病。在另一个实施方案中,所述出血性疾病或障碍是血友病A。这涉及施用与合适的表达控制序列可操作地连接的经优化的FVIII编码核酸。在某些实施方案中,这些序列被掺入病毒载体中。用于此类基因疗法的合适病毒载体包括腺病毒载体、慢病毒载体、杆状病毒载体、EB病毒载体、乳多空病毒载体、痘苗病毒载体、单纯疱疹病毒载体和腺相关病毒(AAV)载体。所述病毒载体可以是复制缺陷型病毒载体。在其他实施方案中,腺病毒载体在其E1基因或E3基因中有缺失。在其他实施方案中,所述序列被掺入本领域技术人员已知的非病毒载体中。

[0246] 在另一方面,本文公开的核酸分子用于特异性改变活生物体的遗传信息(例如基因组)。如本文所用,术语“改变”或“遗传信息的改变”是指细胞基因组的任何改变。在治疗遗传障碍的上下文中,改变可包括但不限于插入、缺失和/或校正。

[0247] 在一些方面,改变还可以包括基因敲入、敲除或敲低。如本文所用,术语“敲入”是指将DNA序列或其片段添加到基因组中。这种待敲入的DNA序列可以包括一个或多个完整的基因,可以包括与前述的基因或任何部分或片段关联的调节序列。例如,可以将编码野生型蛋白质的cDNA插入携带突变基因的细胞的基因组中。敲入策略不需要替换全部或部分缺陷基因。在一些情况下,敲入策略可进一步涉及用所提供的序列取代现有序列,例如用野生型拷贝取代突变等位基因。术语“敲除”是指消除基因或基因的表达。例如,可以通过缺失或添加核苷酸序列从而导致阅读框破坏来敲除基因。作为另一个例子,可以通过用不相关序列替代基因的一部分来敲除基因。如本文所用的术语“敲低”是指减少基因或其一种或多种基因产物的表达。作为基因敲低的结果,可以减弱蛋白质活性或功能,或者可以降低或消除蛋白质水平。

[0248] 在一些实施方案中,本文公开的核酸序列用于基因组编辑。基因组编辑通常是指优选以精确或预先确定的方式修饰基因组的核苷酸序列的过程。本文所述的基因组编辑方法的例子包括使用定点核酸酶在基因组中的精确靶位置切割脱氧核糖核酸(DNA),从而在基因组中的特定位置产生单链或双链DNA断裂的方法。如Cox等人(2015).Nature Medicine 21(2):121-31中最近所综述的,这样的断裂可以并定期地通过天然的内源性细胞过程如同源定向修复(HDR)和非同源末端连接(NHEJ)来修复。这两个主要的DNA修复过程由一系列替代途径组成。NHEJ直接连接由双链断裂导致的DNA末端,有时会丢失或添加核苷酸序列,这可能会破坏或增强基因表达。HDR利用同源序列或供体序列作为模板,用于在断裂点插入限定的DNA序列。同源序列可以在内源性基因组中,如姐妹染色单体。可替代地,供体可以是外源核酸,如质粒、单链寡核苷酸、双链寡核苷酸、双链体寡核苷酸或病毒,其具有与核酸酶切割的基因座高度同源的区域,但还可以含有额外的序列或序列变化,包括可以掺入切割的靶基因座中的缺失。第三种修复机制可以是微同源介导的末端连接(MMEJ),也称为“替代NHEJ”,其中遗传结局与NHEJ类似,因为在切割位点可发生小的缺失和插入。MMEJ可以利用

DNA断裂位点两侧的几个碱基对的同源序列来驱动更有利的DNA末端连接修复结局,并且最近的报道进一步阐明了该过程的分子机制,参见例如Cho和Greenberg(2015).Nature 518,174-76。在一些情况下,可以基于对DNA断裂位点处的潜在微同源性的分析来预测可能的修复结局。

[0249] 这些基因组编辑机制中的每一种都可用于产生所需的基因组改变。基因组编辑过程中的一个步骤可以是在预期突变位点附近的靶基因座中产生一个或两个DNA断裂,后者作为双链断裂或作为两个单链断裂。这可以通过使用定点多肽,如CRISPR内切核酸酶系统等来实现。

[0250] 在另一方面,本文所述的核酸分子可用于脂质纳米颗粒(LNP)介导的FVIII ceDNA递送。由阳离子脂质与其他脂质组分如中性脂质、胆固醇、PEG、PEG化脂质和寡核苷酸形成的脂质纳米颗粒已用于阻断血浆中核酸的降解并促进细胞对寡核苷酸的摄取。此类脂质纳米颗粒可用于将本文所述的核酸分子递送给受试者。

[0251] 本公开文本提供了增加受试者中具有FVIII活性的多肽的表达的方法,其包括向有需要的受试者施用本公开文本的分离的核酸分子,其中所述多肽的表达相对于包含SEQ ID NO:6的参考核酸分子而言增加。本公开文本还提供了增加受试者中具有FVIII活性的多肽的表达的方法,其包括向有需要的受试者施用本公开文本的载体,其中所述多肽的表达相对于包含参考核酸分子的载体增加。

[0252] 本文所述的所有各个方面、实施方案和选择都能以任何和所有变化组合。

[0253] 本说明书中提到的所有出版物、专利和专利申请都通过引用并入本文,并入程度就像每个单独的出版物、专利或专利申请被具体地且单独地指出以通过引用并入一样。

实施例

[0254] 在提供了上述公开文本之后,可以通过参考本文提供的实施例获得进一步的理解。这些实施例仅用于说明的目的,而不旨在是限制性的。

实施例1:产生ceDNA的方法

[0255] 在杆状病毒-昆虫细胞系统中,重组BEV递送在强启动子下的感兴趣基因,并且提供病毒在昆虫细胞中复制所必需的转录复合物。该系统提供了将感兴趣转基因以稳定细胞系的形式插入在杆状病毒基因组和/或昆虫细胞基因组中的灵活性。利用杆状病毒-昆虫细胞系统的这些优点,设计了三种不同的ceDNA生产方法,以根据可扩展性的难易度提供广泛的选择。

1.单BAC

[0256] 为了研究单Bac方法用于转基因表达的用途,通过Tn7转座将优化的FVIIIXTEN表达盒与细小病毒ITR插入在BIVVBac中多角体蛋白基因座中的mini-attTn7位点处,并且在相同骨架中,通过Cre-LoxP重组将ITR特异性复制(Rep)基因表达盒插入在EGT基因座中的LoxP位点处。然后产生重组BEV并将其用于在Sf9细胞中感染以产生FVIIIXTEN ceDNA,如图1A中所描绘。使用用于控制Rep表达水平的不同启动子来证明单Bac方法用于ceDNA产生的概念,如下所述。

2.双BAC:

[0257] 为了研究双Bac方法用于转基因表达的用途,通过Tn7转座将优化的FVIIIXTEN表达盒与细小病毒ITR、以及将ITR特异性复制(Rep)基因表达盒插入在两种不同BIVVBac杆粒

中的多角体蛋白基因座中的mini-attTn7位点处。然后产生重组BEV并将其用于在Sf9细胞中共感染以产生FVIIIXTEN ceDNA,如图1B中所描绘。使用两种杆状病毒的不同感染复数(MOI)比率和对Rep表达水平的微调以获得可重复的ceDNA生产力来研究与双BAC方法相关的挑战,如以下实验中所述。

3.稳定细胞系:

[0258] 为了研究稳定细胞系方法用于转基因表达的用途,用优化的具有细小病毒ITR的FVIIIXTEN表达盒产生稳定细胞系。重组杆粒还通过以下方式产生:通过Tn7转座将ITR特异性复制(Rep)基因表达盒插入在BIVVBac杆粒中多角体蛋白基因座中的mini-attTn7位点处。然后产生重组Rep.BEV并将其用于在FVIIIXTEN稳定细胞系中感染以产生FVIIIXTEN ceDNA,如图1C中所描绘。通过使用GFP作为代表的FACS细胞分选富集FVIIIXTEN转化体以加快产生稳定细胞系的过程来研究与稳定细胞系方法相关的挑战,如以下实验中所述。

实施例2:FVIIIXTEN HBoV1 ITR表达构建体

[0259] 人博卡病毒1(HBoV1)(一种自主细小病毒)是支持野生型腺相关病毒2(AAV2)复制的辅助病毒。AAV以及非AAV细小病毒ITR用于产生FVIIIXTEN ceDNA的用途已经在杆状病毒系统中得到证实(参见例如,美国专利申请号63/069,073)。与其他细小病毒ITR相比,HBoV1 ITR具有独特的大小和形式。HBoV1的5'(REH)ITR为140bp长(SEQ ID NO:1)并形成具有完全碱基配对的"U"形发夹,而3'(LEH)ITR为200bp长(SEQ ID NO:2)并形成具有三向连接的环,这使得其本质上是不对称的,不同于其他细小病毒ITR的末端区域(图2A)。

[0260] 研究HboV1 ITR用于产生FVIIIXTEN ceDNA。假设不对称ITR可以通过稳定转基因而增强长期持续表达。为了验证该假设,通过GenScript®(Piscataway,NJ)®合成DNA构建体以产生SEQ ID NO:3所示的核酸序列(图2A),所述DNA构建体包含在肝特异性修饰小鼠甲状腺素转运蛋白(mTTR)启动子(mTTR482)和增强子元件(A1MB2)、杂合合成内含子(嵌合内含子)、土拨鼠转录后调节元件(WPRE)、牛生长激素多腺苷酸化(bGHpA)信号的调节下的包含XTEN 144肽(FVIIIXTEN)的B结构域缺失(BDD)的经密码子优化的人因子VIII(BDDcoFVIII),其两侧为人HBoV1 5'和3'ITR。将该合成DNA克隆到pFastBac1(Invitrogen)载体中以产生pFastBac.mTTR.FVIIIXTEN.HBoV1.ITRs转移载体(图2B)。然后将该载体转化到BIVVBac^{DH10B}大肠杆菌中以产生重组BEV,AcBIVVBac.Polh.GPV.Rep^{Tn7},如下所述。

实施例3:HBoV1 NS1(非结构的)表达构建体

HBoV1 NS1 Tn7转移载体

[0261] 已知HBoV1通过mRNA转录物表达五种非结构蛋白,即NS1、NS2、NS3、NS4和NP1,所述mRNA转录物通过单个病毒前体mRNA的选择性剪接和多腺苷酸化产生。NS1至NS4蛋白被编码于相同开放阅读框(ORF)的不同区域中。NS1由分别在N末端、中间和C末端的起始结合/核酸内切酶结构域(OBD)、解旋酶结构域和推定的反式激活结构域(TAD)组成。NS1结合HBoV1复制起点,并且推测在滚卡复制(rolling-hairpin replication)过程中对起点的单链DNA(ssDNA)产生切口。

[0262] 为了研究NS1在真核细胞中的ITR介导的载体产生中并从Sf9细胞"拯救"两侧为HBoV1 ITR的FVIIIXTEN ceDNA载体基因组的作用,产生HBoV1 NS1表达构建体并将其插入BIVVBac中以在Sf9细胞中产生表达HBoV1 NS1的重组BEV。

[0263] 为了产生表达载体,从HBoV1基因组(GenBank登录号:JQ923422)获得HBoV1NS1的

编码序列并针对Sf细胞基因组进行密码子优化,然后通过GenScript®合成以产生如SEQ ID NO:4所示的核酸序列。然后在AcMNPV多角体蛋白启动子的控制下,将合成的HBoV1 NS1 DNA克隆到pFastBac1(Invitrogen)载体中(图3A),以产生pFastBac.Polh.HBoV1.NS1(图3B)转移载体。还将合成的HBoV1 NS1 DNA克隆到在即时早期1(IE1)启动子控制下的pFastBac1(Invitrogen)载体中,所述启动子之前是AcMNPV转录因子hr5元件(图4A),以产生pFastBac.HR5.IE1.HBoV1.NS1(图4B)转移载体。还将合成的HBoV1 NS1 DNA克隆到在OpMNPV即时早期2(IE2)启动子下的pFastBac1(Invitrogen)载体中(图5A),以产生pFastBac.OpIE2.HBoV1.NS1(图5C)转移载体。然后将这些载体转化到BIVVBac^{DH10B}大肠杆菌中以产生重组BEV:分别为AcBIVVBac.Polh.HBoV1.NS1^{Tn7}、AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}或AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}。

[0264] 然后将这些重组BEV与FVIIIXTEN BEV一起用于Sf9细胞的共感染(双BAC),以产生FVIIIXTEN ceDNA载体。

HBoV1 NS1 Cre-LoxP供体载体

[0265] 为了研究单Bac方法用于产生ceDNA的用途,将HBoV1 NS1基因在LoxP位点插入多角体蛋白基因座中的Tn7位点处编码FVIIIXTEN表达盒的重组BIVVBac中。在这些位点插入这两种基因的基本原理是避免在FVIIIXTEN两侧的反向末端重复序列(ITR)与也是回文重复序列的LoxP序列的干扰。

[0266] 此外,为了解决与上述单BAC系统相关的挑战,测试了在感染周期期间在不同时间和水平表达的杆状病毒基因的不同启动子,以控制编码FVIIIXTEN HBoV1 ITR和NS1的单BAC中HBoV1 NS1表达的水平。

[0267] 在AcMNPV多角体蛋白启动子(图3A)或前面有(图4A)和没有(图5B)AcMNPV转录增强子hr5元件的即时早期1启动子的控制下,将合成的Sf密码子优化的HBoV1 NS1 DNA克隆到Cre-LoxP供体载体(如美国专利申请号63/069,073所述)中,分别产生pCL.Polh.HBoV1.NS1(图3C)、pCL.HR5.IE1.HBoV1.NS1(图4C)和pCL.IE1.HBoV1.NS1(图5D)Cre-LoxP供体载体。这些构建体用前缀“pCL”表示“质粒Cre-LoxP”。(见图3C、图4C、图5D)。然后将所得Cre-LoxP供体载体在LoxP位点插入Tn7位点处编码FVIIIXTEN HBoV1 ITR的BIVVBac杆粒中(图6B),如下所述。

实施例4:FVIIIXTEN HBoV1 ITR杆状病毒表达载体(BEV)

[0268] 为了产生编码具有HBoV1 ITR的FVIIIXTEN表达盒(图2A)的重组BEV,首先,用Tn7转移载体pFastBac.mTTR.FVIIIXTEN.HBoV1.ITRs(图2B)超转化BIVVBac^{DH10B}大肠杆菌(美国专利申请号63/069,073中所述)。在卡那霉素、庆大霉素、X-Gal和IPTG上选择转化体。BIVVBac中在mini-attTn7插入位点处的FVIIIXTEN表达盒和庆大霉素抗性基因的位点特异性转座破坏了LacZα(与mini-attTn7框内融合)并且在X-Gal介导的双重抗生素选择上产生白色大肠杆菌菌落。通过碱裂解小量制备从白色大肠杆菌菌落中分离重组杆粒DNA,并且用限制酶消化以确定正确的遗传结构。限制酶作图的结果显示每种重组杆粒的预期片段,表明具有HBoV1 ITR的FVIIIXTEN表达盒在BIVVBac的多角体蛋白基因座中的位点特异性转座(图6A)。通过使用转移质粒内部和外部的引物对跨越预期插入位点的区域进行PCR扩增并对所得扩增子进行测序,获得了进一步的确认(数据未显示)。

[0269] 根据制造商的说明书,将编码具有HBoV1 ITR的FVIIIXTEN表达盒的正确重组杆粒

大量制备纯化并使用Cellfectin®(Invitrogen)转染试剂转染到Sf9细胞中。在转染后4-5天,收获子代杆状病毒并在Sf9细胞中进行噬斑纯化,如前所述。Jarvis等人(2014),Methods Enzymol.,536:149-163。重组BEV,AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}(图6B)的六个经噬斑纯化的RFP+克隆在Sf9细胞中扩增至P1(传代1),所述Sf9细胞以0.5×10^6/mL接种在T25烧瓶中的补充有10%热灭活胎牛血清(FBS)的ESF-921培养基中。在感染后4-5天,对于每个克隆BEV,所有克隆显示通过RFP+细胞数确定的感染进展,表明病毒能够正常复制,并且在杆状病毒基因组中插入具有HBoV1 ITR的FVIIIXTEN转基因对子代病毒产生没有不利影响。选择最高的RFP+克隆用于在Sf9细胞中进一步扩增以产生工作BEV原液(P2),然后用于在双BAC系统中与HBoV1 NS1 BEV共感染以产生ceDNA载体。

实施例5:FVIIIXTEN HBoV1 ITRs+HBoV1 NS1杆状病毒表达载体(BEV)

[0270] 为了测试BIVVBac是否可用于容纳多个转基因,产生了在不同启动子控制下编码两个转基因表达盒的衍生载体家族:1)FVIIIXTEN HBoV1 ITRs和2)HBoV1 NS1。这些BEV分两步产生。首先,如上所述,将具有HBoV1 ITR的FVIIIXTEN表达盒经由Tn7转座插入在BIVVBac中多角体蛋白基因座中的mini-attTn7位点。然后,使用Cre重组酶(New England Biolabs)经由体外Cre-LoxP重组将所得杆粒BIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs(图6B)用于在EGT基因座的LoxP位点处插入HBoV1 NS1表达盒。

[0271] 在该方法中,将编码在AcMNPV多角体蛋白启动子(图3C)或前面有(图4C)和没有(图5D)AcMNPV转录增强子hr5元件的即时早期1(IE1)启动子下的HBoV1 NS1的Cre-LoxP供体载体插入BIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs(图6B)杆粒中。将重组反应物在DH10B大肠杆菌中转化,并且在卡那霉素、庆大霉素和氨苄青霉素上选择转化体。通过以下方式来筛选三重抗生素抗性菌落:通过限制酶作图和/或使用转移质粒内部和外部的引物对跨越预期插入位点的区域进行PCR扩增(图7A、图7B和图7C),并对所得扩增子进行测序。

[0272] 大量制备法纯化编码两个转基因盒的正确重组杆粒,并且使用Cellfectin®(Invitrogen)转染试剂在Sf9细胞中转染。在转染后4-5天,收获子代杆状病毒并在Sf9细胞中进行噬斑纯化。将每种重组BEV(AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP}:图7D;AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1^{LoxP}:图7E;和AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1^{LoxP}:图7F)的六个经噬斑纯化的RFP+和GFP+克隆在Sf9细胞中扩增至P1(传代1),所述Sf9细胞以0.5×10^6/mL接种在T25烧瓶中的补充有10%热灭活胎牛血清(FBS)的ESF921培养基中。在感染后4-5天,所有克隆对于每种重组BEV均显示出感染进展(通过GFP+和RFP+细胞的数量确定),这表明病毒能够正常复制并且在同一杆状病毒基因组中插入多个转基因对子代病毒产生没有不利影响。通过低速离心收获P1病毒,并处理感染的细胞沉淀,以通过免疫印迹检测HBoV1 NS1。最后,进一步扩增每种BEV的最高HBoV1 NS1表达克隆以产生工作BEV原液(P2),随后在Sf9细胞中滴定。如下所述,将滴定的BEV用于Sf9细胞中的感染以产生FVIIIXTEN ceDNA载体。

实施例6:从单BAC产生FVIIIXTEN HBoV1 ITRs ceDNA载体

[0273] 在Sf9细胞中测试编码FVIIIXTEN HBoV1 ITR和HBoV1 NS1基因(图7D-图7F)的单BAC BEV的FVIIIXTEN ceDNA产生。用每种BEV的滴定工作原液(P2)以0.1、0.5、1.0、2.0或3.0噬斑形成单位(pfu)/细胞的感染复数(MOI)感染约2.5×10^6/mL细胞(图8A)。将细胞悬浮于50mL无血清ESF-921培养基中,然后在28℃振荡培养箱中孵育72-96h或直至细胞活力

达到60％-70％。在感染后约96h,收获感染细胞,并且处理沉淀,以用于根据制造商的说明书通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)进行FVIIIXTEN ceDNA载体分离。在0.8％至1.2％琼脂糖凝胶电泳上分析最终洗脱级分,以确定FVIIIXTEN ceDNA载体的生产力。

[0274] 编码具有HBoV1 ITR的FVIIIXTEN和多角体蛋白驱动的HBoV1-NS1的AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1^{LoxP} BEV(图8B)的琼脂糖凝胶分析示于图8C中。结果显示在所有测试剂量中,DNA条带对应于FVIIIXTEN HBoV1ITRs(约8.5kb)ceDNA的大小,随着MOI增加生产力增加。

[0275] 该结果与用AAV2 ITRs单BAC获得的ceDNA生产力(其中先前观察到生产力随着病毒载量的增加而降低)相反。不受理论的约束,HBoV1-NS1蛋白可以由于REH和LEH ITR的独特结构而在HBoV1 ITR的末端解离位点具有独特的结合机制和核酸内切酶活性以用于DNA复制(图2A)。

[0276] 总之,这些实验显示,单BAC方法确实证明了从编码具有HBoV1 ITR的FVIIIXTEN和NS1转基因的单一重组BEV产生ceDNA的概念。它还显示了在杆状病毒穿梭载体(BIVVBac)中不同基因座处插入多个转基因的可行性和功能性,以及其用于在杆状病毒昆虫细胞系统中产生重组AAV载体的潜在用途。

实施例7:HBoV1 NS1(非结构的)杆状病毒表达载体(BEV)

[0277] 唯一的在结构上表征的细小病毒NS1 N末端核酸酶结构域来自AAV2 Rep,其结合复制起点(Ori)中的连续四核苷酸重复序列。然而,这种四核苷酸重复序列对AAV是特异性的并且不存在于HBoV1中。实际上,HBoV1基因组的LEH(3'ITR)形成具有三向连接的环,而REH(5'ITR)是具有完全碱基配对的发夹(图2A),它们在博卡病毒中是保守的,但与AAV和细小病毒B19(B19V)基因组的末端区域不同。这些发现表明,HBoV1中Ori的NS1识别模式可能与AAV中的不同。此外,已知AAV不引起人疾病并且是依赖病毒,因为病毒复制需要辅助病毒如疱疹病毒或腺病毒。HBoV1 NS1与AAV Rep共享少至14％的序列同一性。已经证明,HBoV1-NS1含有带正电荷的表面,所述表面是Ori的推定结合位点并直接支持HBoV DNA复制,如以针对细小病毒提出的常见滚卡机制。

[0278] HBoV1-NS1显现是真核细胞中ITR介导的载体生产所必需的。为了研究两侧为ITR的FVIIIXTEN载体基因组从Sf9细胞或FVIIIXTEN BEV中的潜在“拯救”,在不同杆状病毒启动子下产生编码HBoV1-NS1的重组BEV以优化Sf9细胞中NS1表达的水平。

[0279] 为了产生这些BEV,用Tn7转移载体pFastBac.Polh.HBoV1-NS1(图3B)、pFastBac.HR5.IE1.HBoV1-NS1(图4B)或pFastBac.OpIE2.HBoV1.NS1(图5C)超转化BIVVBac^{DH10B}大肠杆菌(参见美国专利申请号63/069,073)。在卡那霉素、庆大霉素、X-Gal和IPTG上选择转化体。在BIVVBac中在mini-attTn7插入位点处HBoV1-NS1表达盒和庆大霉素抗性基因的位点特异性转座破坏LacZα(与mini-attTn7框内融合),并切在X-Gal介导的双重抗生素选择上产生白色大肠杆菌菌落。因此,通过碱裂解小量制备从白色大肠杆菌菌落中分离重组杆粒DNA,并且用限制酶消化以确定正确的遗传结构。限制酶作图的结果显示了每种重组杆粒的预期片段,这表明HBoV1-NS1在BIVVBac的多角体蛋白基因座中的位点特异性转座(图9A)。通过使用转移质粒内部和外部的引物对跨越预期插入位点的区域进行PCR扩增并对所得扩增子进行测序,获得了进一步的确认(数据未显示)。

[0280] 根据制造商的说明书,使用Cellfectin®(Invitrogen)转染试剂将编码Polh.HBoV1-NS1、HR5.IE1.HBoV1-NS1或OpIE2-HBoV1-NS1的确认的正确重组杆粒转染到Sf9细胞中。在转染后4-5天,收获子代杆状病毒并在Sf9细胞中进行噬斑纯化,如前所述。Jarvis等人(2014),Methods Enzymol.,536:149-163。将每种重组BEV,AcBIVVBac.Polh.HBoV1-NS1^{Tn7}(图9B)、AcBIVVBac.HR5.IE1.HBoV1.NS1^{Tn7}(图9C)和AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}(图9D)的六个经噬斑纯化的RFP+克隆在Sf9细胞中扩增至P1(传代1),所述Sf9细胞以0.5×10^6/mL接种在T25烧瓶中的补充有10%热灭活胎牛血清(FBS)的ESF-921培养基中。在感染后4-5天,对于每个克隆BEV,所有克隆显示通过RFP+细胞数确定的感染进展,表明病毒能够正常复制,并且在杆状病毒基因组中插入HBoV1-NS1对子代病毒产生没有不利影响。

[0281] 选择最高的RFP+克隆用于在Sf9细胞中进一步扩增以产生工作BEV原液(P2)。然后将滴定的病毒原液用于与FVIIIXTEN BEV在双BAC系统或FVIIIXTEN HBoV1 ITRs稳定细胞系中共感染以产生ceDNA载体。

实施例8:从双BAC产生FVIIIXTEN HBoV1 ITRs ceDNA载体

[0282] 为了研究用于转基因表达的双BAC方法,测试编码FVIIIXTEN HBoV1 ITRs的克隆重组BEV与多角体蛋白驱动的HBoV1-NS1 BEV,以用于以1:10和1:5比率的不同MOI或以0.3、1.0、3.0和5.0pfu/细胞的MOI的不同比率共感染以在Sf9细胞中产生FVIIIXTEN ceDNA载体(图10A)。具体地,将约2.0×10^6/mL细胞接种在50mL无血清ESF-921培养基中,并且分别用MOI为0.1、0.3、0.5、1.0、3.0、5.0pfu/细胞的AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7} BEV与MOI为0.01、0.03、0.05、0.1、0.3、0.5pfu/细胞来保持恒定1:10比率或MOI为0.02、0.06、0.1、0.2、0.6、1.0pfu/细胞来保持恒定1:5比率的AcBIVVBac.Polh.HBoV1-NS1^{Tn7} BEV的滴定工作原液(P2)共感染。类似地,对细胞还以1:1、1:2、1:5或1:10比率以0.3、1.0、3.0或5.0pfu/细胞的恒定MOI共感染(图10B)。在每种情况下,不去除病毒接种物,并且将细胞在28℃摇床培养箱中孵育直到活力达到60%-70%。在感染后约96h,收获感染细胞,并且处理沉淀,以用于根据制造商的说明书通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)进行FVIIIXTEN ceDNA载体分离。在0.8%至1.2%琼脂糖凝胶电泳上分析最终洗脱级分,以确定ceDNA生产力。

[0283] 如所预期的,琼脂糖凝胶分析显示出不同条件下FVIIIXTEN ceDNA生产力的不同程度。然而,以3.0pfu/细胞的MOI共感染的双BAC显示FVIIIXTEN ceDNA生产力水平随病毒载量比率增加而增加,与测试的其他条件相比,1:10比率下的生产力水平是最高的(图10C)。较高的病毒载量显现提高了FVIIIXTEN HBoV1 ITRs ceDNA的生产力,这与在单BAC BEV中的观察结果一致(参见实施例6)。这进一步表明在Sf9细胞中HBoV1-ITR依赖性FVIIIXTEN ceDNA复制需要更高水平的HBoV1-NS1。

[0284] 单BAC或双BAC的结果表明,HBoV1-NS1复制水平对杆状病毒系统中FVIIIXTEN ceDNA生产力具有显著影响。

[0285] 作为测试上文讨论的几种不同共感染条件的替代方案,通过利用杆状病毒基因组的不同启动子探索了提高FVIIIXTEN ceDNA生产力的其他方法。杆状病毒基因启动子根据其在感染周期中的转录开始分为即时早期、早期、晚期和极晚期启动子。其中,如名称所指示,即时早期(ie)基因启动子在病毒感染后立即启动,并在整个感染周期中保持活性。然

而,晚期或极晚期基因启动子如多角体蛋白保持沉默直到病毒达到感染的晚期。

[0286] 为了利用从杆状病毒基因组中启动子的这种宽范围选择,测试了HBoV1-NS1的即时早期1(IE1)启动子。显示增加Sf9细胞中表达水平的转录增强子hr5元件被包括在IE1启动子之前。这产生了编码在之前是AcMNPV转录增强子hr5元件的AcMNPV即时早期1(IE1)启动子控制下的HBoV1-NS1的重组BEV,如图9C所描绘。

[0287] 基于图10C中获得的结果,通过保持恒定比率1:10,以不同MOI用编码FVIIIXTEN HBoV1 ITRs和hr5.IE1驱动的HBoV1-NS1的BEV共感染Sf9细胞。作为阳性对照,包括多角体蛋白驱动的HBoV1-NS1 BEV,并在同一组实验中再次测试。更具体地,将约2.0×10^{6}/mL Sf9细胞用MOI为0.1、0.3、0.5、1.0、3.0、5.0pfu/细胞的AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7} BEV(图10B)与MOI为0.01、0.03、0.05、0.1、0.3、0.5pfu/细胞来保持恒定1:10比率的AcBIVVBac.Polh.HBoV1-NS1^{Tn7}或AcBIVVBac.hr5.IE1.HBoV1-NS1^{Tn7} BEV(图11B)的滴定工作原液(P2)共感染。其余操作如上所述(参见实施例6)。

[0288] 在0.8%至1.2%琼脂糖凝胶电泳上分析最终洗脱级分,以确定ceDNA生产力。多角体蛋白驱动的HBoV1-NS1共感染显示FVIIIXTEN ceDNA生产力水平随着MOI增加而增加,这进一步证实了用于ceDNA生产的双BAC方法的再现性。然而,出人意料的是,如图10C中观察到的,hr5.IE1驱动的HBoV1-NS1共感染显示出几乎不可检测水平的FVIIIXTEN ceDNA,生产力没有随着MOI增加而明显增加。

[0289] 该数据表明HBoV1-NS1的早期发作对于拯救具有HBoV1 ITR的FVIIIXTEN ceDNA可能不是关键的。相反,感染后期较高水平的表达可能是具有HBoV1 ITR的FVIIIXTEN ceDNA的有效拯救和生产力所需要的。这些结果进一步证实了在Sf9细胞中HBoV1-ITR依赖性FVIIIXTEN ceDNA复制需要较高水平的HBoV1-NS1。

[0290] 总之,这些实验已经表明,双BAC方法确实证明了从编码具有HBoV1 ITR的FVIIIXTEN和/或NS1转基因的两种重组BEV产生ceDNA的概念。这些实验也证明了最佳MOI比率和/或启动子对于在Sf9细胞中实现FVIIIXTEN ceDNA的较高生产力的重要性。

实施例9:FVIIIXTEN HBoV1 ITR稳定细胞系

[0291] 假设昆虫细胞基因组可能被修饰以产生ceDNA,以用于杆状病毒感染后的治疗应用。为了检验该假设,通过GenScript®(Piscataway,NJ)合成质粒,所述质粒编码在AcMNPV即时早期(ie1)启动子控制下的新霉素抗性标记(pUC57.HR5.IE1.NeoR.P10PAS:SEQ ID NO:7)(图12A)或增强的绿色荧光蛋白(eGFP)(pUC57.HR5.IE1.eGFP.P10PAS:SEQ ID NO:8)(图12B),所述启动子之前是转录增强子hr5元件,所述新霉素抗性标记或所述增强的绿色荧光蛋白之后是AcMNPV p10多腺苷酸化信号。

[0292] 使用修改的磷酸钙转染方法,将这些质粒与编码具有HBoV1 ITR的FVIIIXTEN的质粒(Sf.mTTR.FVIIIXTEN.HBoV1.ITRs)(图12C)共转染到Sf9细胞中。在转染后24h,在荧光显微镜下可视化细胞以确定转染效率,并且结果显示>80% GFP+细胞,这表明更高的转染效率。在转染后72h,用G418抗生素(Sigma Aldrich)以1.0mg/mL终浓度选择细胞,所述G418抗生素悬浮于完全TNMFH培养基(补充有10% FBS+0.1% Pluronic F68的Grace昆虫培养基)中。在约一周的选择之后,回收了约50%的转化细胞,这表明新霉素抗性标记稳定地整合到该细胞群中。将存活细胞从选择培养基中取出并且用新鲜的完全TNMFH培养基培养直到汇合生长。随着汇合细胞继续分裂,将它们作为贴壁培养物逐渐扩增到更大的培养器皿中。随

后,通过以下方式使多克隆细胞群适应悬浮培养:在摇瓶中在完全TNMFH中培养1代并且在补充有10% FBS的ESF-921培养基中生长1代。最后,使细胞作为悬浮培养物适应摇瓶中的无血清ESF-921。通常将这些摇瓶培养物维持在无血清ESF-921培养基中,每四天传代一次,并且监测细胞生长。

实施例10:FVIIIXTEN ceDNA纯化

[0293] 在杆状病毒-昆虫细胞系统中,重组BEV递送在强启动子下的感兴趣基因,并且提供病毒在昆虫细胞中复制所必需的转录复合物。通常,杆状病毒DNA基因组在细胞核中复制并产生数以千万个子代病毒颗粒,每个子代病毒颗粒含有全长DNA基因组。已经证明杆状病毒基因组DNA与ceDNA共纯化,同时使用基于质粒DNA的纯化方法(如硅胶柱)从昆虫细胞中分离DNA。商业质粒DNA试剂盒柱通常不被设计用于基于分子量分离DNA,并且因此,通常样品中存在的所有形式的DNA均可以与这些柱结合。此外,大分子量DNA的结合能力可能与低分子量DNA不同,并且基于阴离子交换的试剂盒柱未基于不同大小的DNA的结合效率进行优化。

[0294] 假设在ceDNA制备物中观察到的高分子量DNA(>20kb)最可能是与低分子量FVIIIXTEN ceDNA(约8.5kb)共纯化的杆状病毒和/或Sf9细胞基因组DNA(参见例如图8C、图10C和图11C)。

[0295] 先前,采用间接方法以通过敲除杆状病毒衣壳基因(如VP80)来减少杆状病毒DNA,所述杆状病毒衣壳基因是产生感染性子代病毒产生所需的。这种方法显示出从敲除BEV获得的ceDNA制备物中杆状病毒DNA的显著减少(参见美国专利申请号63/069,115)。虽然这种方法在减少杆状病毒DNA污染方面是有效的,但它无法减少细胞基因组DNA,其在从感染细胞沉淀中获得的总DNA中以显著量(约60%)存在。

[0296] 因此,采用了将FVIIIXTEN ceDNA与其余不需要的DNA分离的直接方法,并且证明所述直接方法从来自感染细胞沉淀的总DNA制备物中高效获得纯化的FVIIIXTEN(>95%纯度)。这种新颖的方法利用了制备型电泳,其广泛用于根据大小和电荷分离不同的蛋白质分子。参见例如Michov,B.(2020)Electrophoresis.Berlin,Boston:De Gruyter,第405-424页。例如,Bio-Rad Model 491prep cell或其他此类单元可用于基于大小分离复杂分子。

[0297] ceDNA纯化的整个工作流程如图13中所示,其中过程开始于将无血清昆虫细胞培养基中的Sf9细胞培养物从0.5L放大至1.5L或更高体积(图13A)。在达到约2.5x 10^6/mL的所需细胞密度后,通常在接种密度为约1.3x10^6/mL的情况下孵育2天后,将细胞用单BAC或双BAC BEV(取决于用于产生ceDNA的方法)以优化的MOI感染,并且将细胞在28℃摇床培养箱中孵育,直到活力达到约60%-70%,这通常需要约4天(图13B)。一旦活力达到约70%,就收获细胞并且对其进行处理,以根据制造商的说明书通过阴离子交换色谱试剂盒柱(如PureLink HiPure Expi质粒Gigaprep纯化试剂盒(Invitrogen))进行总DNA纯化。在0.8%至1.2%琼脂糖凝胶电泳上检查纯化DNA材料的等分试样以确定DNA生产力和完整性(图13C)。

[0298] 然后将纯化的材料加载到制备型琼脂糖凝胶电泳单元上,所述单元含有0.5%制备型琼脂糖凝胶和0.25%堆积琼脂糖凝胶,根据制造商的说明书进行组装。在缓冲液再循环流速为约50mL/min且洗脱缓冲液速率为50μL/min的情况下,将样品在4℃下以低电压(约40恒定伏特)运行6-7天,以在70-80min将每个级分收集在级分收集室中。在连续洗脱电泳

之后,在0.8%至1.2%琼脂糖凝胶电泳上检查20μL的每个级分以确定FVIIIXTEN ceDNA的纯度(图13D)。将所需级分合并以在-20C下用3M NaOAc(pH 5.5)和100%EtOH沉淀1-2h。最后,将沉淀的FVIIIXTEN ceDNA以高速沉淀并且用70% EtOH洗涤一次,然后再悬浮于TE(pH 8.0)缓冲液中。将纯化的FVIIIXTEN ceDNA再次在0.8%至1.2%琼脂糖凝胶电泳上检查以确认纯度和完整性,然后注射到动物中以进行体内功效研究(图13E)。

实施例11:FVIIIXTEN HBoV1 ITR体内功效

ssFVIIIXTEN HBoV1 ITR(单链DNA)

[0299] 假设在ITR区域内形成的发夹能够实现较高水平的长期持续转基因表达。为了在体内研究HBoV1 ITR的功能性,在hFVIIIR593C$^{+/+}$/HemA小鼠中测试包含具有预先形成的HBoV1 ITR的经密码子优化的人FVIIIXTEN的单链DNA(ssDNA)。这些小鼠含有人FVIII-R593C转基因,所述转基因被设计为具有鼠白蛋白(Alb)启动子,所述启动子驱动含有在轻度血友病A患者中经常观察到的突变的改变的人凝结因子VIII(FVIII)cDNA的表达。这些小鼠还携带FVIII基因的敲除并且缺乏内源性FVIII蛋白。这些双突变小鼠耐受人FVIII注射并且没有FVIII活性。它们在用人FVIII处理后产生非常少的抑制性抗体并且缺乏FVIII应答性T细胞或B细胞。hFVIIIR593C$^{+/+}$/HemA小鼠进一步描述于Bril,等人(2006)Thromb.Haemost.95(2):341-7中。

[0300] 通过在95℃使PmlI消化的双链DNA片段产物(FVIII表达盒和质粒骨架)变性,然后在4℃冷却以允许回文ITR序列折叠来产生具有预先形成的HBoV1 ITR的单链FVIIIXTEN(ssFVIIIXTEN)。然后,将ssFVIIIXTEN以10μg或40μg/小鼠(分别相当于400μg或1600μg/kg)经由流体动力学尾静脉注射全身注射。以7天的间隔从注射小鼠中收集血浆样品5.5个月。根据制造商的说明书,通过ChromogenixCoatest®SP因子VIII生色测定来测量血浆FVIII活性。

[0301] 对于ssFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图14A中。结果显示在5.5个月的历程内在HemA小鼠中具有剂量依赖性反应,在两个测试剂量下均具有FVIII表达的超生理水平(>1000%正常值)。然而,直到第56天观察到FVIII表达的初始下降,但随后在第168天之前水平稳定,表明来自注射动物肝脏的两侧为HBoV1 ITR的ssFVIIIXTEN的持续表达。因此,这些结果验证了HBoV1 ITR用于FVIIIXTEN在体内的长期持续表达的功能性。

ceFVIIIXTEN HBoV1 ITR(封闭末端DNA)

[0302] 为了验证HBoV1 ITR在ceDNA中的功能性,将如上所述从感染的Sf9细胞沉淀纯化的ceFVIIIXTEN以0.3μg、1.0μg或2.0μg/小鼠(分别相当于12μg、40μg和80μg/kg)经由流体动力学尾静脉注射全身注射到hFVIIIR593C$^{+/+}$/HemA小鼠中。以7天间隔收集来自注射小鼠的血浆样品,并且通过生色测定测量FVIII活性,如上所述。

[0303] 对于ceFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图14B中。该研究结果显示在HemA小鼠中的剂量依赖性反应,在注射后直至第56天,在最高测试剂量下观察到FVIII表达的超生理水平(>500%正常值)。有趣的是,当对小鼠注射1600μg/kg的ssFVIIIXTEN时实现类似的表达水平,其比ceFVIIIXTEN剂量(80μg/kg)高至少20倍(图14A-图14B)。该数据表明与ssDNA形式相比,ceDNA提供更高水平的FVIII表达。

[0304] 总之,这些体内研究验证了HBoV1 ITR在ssDNA或ceDNA形式中的功能性,并证明

HBoV1 ITR可用于在杆状病毒昆虫细胞系统中产生编码目的转基因的功能性ceDNA。

实施例12:使用CRISPR Cas敲除HBoV1 NS1 BEV中的VP80提高ceDNA载体纯度

[0305] 在AcMNPV多角体蛋白启动子下表达的HBoV1 NS1确实能够拯救两侧为HBoV1 ITR的FVIIIXTEN,并证明在杆状病毒系统中用HBoV1 ITR产生ceDNA的概念。然而,在ceDNA制备物中观察到显著水平的杆状病毒DNA(vDNA)污染,可能是由于与AAV2 Rep-BEV相比需要更高的病毒载量来实现更高的ceDNA生产力。在这些ceDNA制备物中观察到的高分子量DNA(>20kb)(图8C,图10C)最可能是与低分子量ceDNA(约8kb)共纯化的杆状病毒基因组DNA。

[0306] 为了减少ceDNA制备物中的杆状病毒DNA污染,实施了敲除VP80的间接方法,VP80是在昆虫细胞(Sf9)中产生感染性病毒颗粒所需的杆状病毒基因组的必需基因。使用Alt-R CRISPR-Cas9系统(参见美国专利申请号63/069,115)在所有三种NS1 BEV中敲除VP80(图9B、图9C和图9D)。这种方法可能减少子代病毒颗粒的数量,并最终减少ceDNA制备物中的杆状病毒DNA污染。

CRISPR-Cas9敲除AcMNPV VP80基因:

[0307] 如前所述(参见例如,国际申请号PCT/US2021/047202),选择编码在AcMNPV多角体蛋白(图9B)或OpMNPV OpIE2(图9C)启动子下的HBoV1 NS1的重组BEV,以通过CRISPR-Cas9系统敲除vp80基因。

[0308] 简言之,根据制造商的说明书,设计两个靶向编码序列的crRNA并用于使用Alt-R CRISPR-Cas9系统(Integrated DNA Technology™)产生功能性sgRNA。然后使用Cellfectin®(Invitrogen™)转染试剂,将每个sgRNA与SpCas9核酸酶和AcBIVVBac.Polh.HBoV1.NS1^{Tn7}或AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7}杆粒DNA共转染Sf9细胞,所述细胞以0.5×106/mL接种在T25烧瓶中的无血清ESF-921培养基中。转染后4-5天,在荧光显微镜下观察细胞,结果显示对于两种sgRNA靶标,约10% RFP+细胞。感染细胞的示例性荧光显微镜图像示于图15中。在单独Cas9情况下用AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV感染的细胞显示出预期的感染进展(图15A),然而,相比之下,sgRNA.VP80.T1(图15B)或sgRNA.VP80.T2(图15C)处理的细胞显示出对单独细胞的限制性感染,这可能是由于VP80的敲除。

[0309] 为了确定每种sgRNA诱导的插入缺失,收获子代杆状病毒并在互补Sf.39K.VP80细胞系中纯化噬斑,如前所述。Jarvis等人(2014),Methods Enzymol.,536:149-163。在感染后5-6天,将十二个经噬斑纯化的RFP+克隆在Sf.39K.VP80细胞中扩增到P1,所述细胞以0.5×10^6/mL接种在T25烧瓶中的补充有10% FBS的ESF-921培养基中。扩增克隆的荧光显微镜观察显示约80%的RFP+细胞,这表明Sf.39K.VP80细胞系能够以反式互补VP80功能用于产生子代病毒。通过低速离心收获每个克隆BEV,然后根据制造商的说明书,通过Qiagen的DNeasy血液和组织基因组DNA分离试剂盒(目录号69506)将细胞沉淀用于总DNA分离。将得到的DNA用作模板,以用AcMNPV vp80编码序列特异的引物对每个靶序列进行PCR扩增。然后凝胶纯化PCR扩增物并通过Genewiz测序设施直接测序。通过TIDE(通过DEcomposition追踪插入缺失)程序(tide.deskgen.com)使用默认设置分析所得序列以确定每种sgRNA诱导的插入缺失。TIDE分析显示,在vp80编码序列中,在sgRNA.T1处理的AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV克隆#4中具有最高(97.1%)-15bp缺失的移码突变(图16A)和AcBIVVBac.OpIE2.HBoV1.NS1^{Tn7} BEV克隆#4中具有最高(37.4%)-4bp和(26.9%)-3bp缺失的移码突变(图16B),没有可检测的插入。将每个克隆扩增到P2以产生工作BEV原

液,随后在Sf.39K.VP80细胞中滴定,如前所述。Jarvis等人(2014),Methods Enzymol.,536:149-163。然后将滴定的vp80KO BEV的工作原液用于共感染双BAC系统以产生FVIIIXTEN HBoV1 ceDNA载体。

使用vp80KO BEV产生人FVIIIXTEN ceDNA:

[0310] 将约2.0×10^6个细胞接种于100mL无血清ESF-921培养基中,并以1.0、2.0、3.0、4.0和5.0pfu/细胞的MOI用AcBIVVBac.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1ΔVP80^{Tn7}或AcMNPV.OpIE2.HBoV1.NS1ΔVP80^{Tn7} BEV的滴定工作PP1P2原液共感染。在每种情况下,不去除病毒接种物,并且将细胞在28℃摇床培养箱中孵育直到活力达到60%-70%。在感染后约96h,收获感染细胞,并且处理沉淀,以用于根据制造商的说明书通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)进行FVIIIXTEN HBoV1 ceDNA分离。

[0311] 通过0.8%至1.2%琼脂糖凝胶电泳分析最终洗脱级分,以确定ceDNA生产力和纯度。琼脂糖凝胶分析显示在表达在AcMNPV多角体蛋白或OpMNPV OpIE2启动子下的HBoV1 NS1的vp80KO BEV中非常低至无可检测的高分子量(>20kb)杆状病毒DNA(vDNA)污染(图16C)。这表明当以2.0、3.0或4.0pfu/细胞的MOI共感染细胞时,vp80KO方法能够减少污染性杆状病毒DNA并同时提高FVIIIXTEN HboV1 ceDNA产量(图11C,图16C)。

实施例13:从双BAC系统产生FVIIIXTEN HBoV1 ceDNA载体

[0312] 遗传不稳定性是杆状病毒学领域中并且尤其是在Sf9细胞中重组杆状病毒的若干传代后的主要关注点之一。此外,杆状病毒基因组含有几个同源区(hr),它们易于在Sf9细胞中传代时重组,并可能在重组BEV中丢失转基因。反向末端重复序列(ITR)也是回文重复序列,并且考虑到杆状病毒DNA的大尺寸,可能在杆状病毒基因组中的不同基因座处重组。因此,为了确定编码具有HBoV1 WT ITR的在肝特异性mTTR启动子下的FVIIIXTEN基因的重组BEV的遗传稳定性,通过以0.1pfu/细胞的MOI感染Sf9细胞来依序扩增BEV,如前所述。Jarvis等人(2014),Methods Enzymol.,536:149-163。测试所得重组BEV的使用双BAC系统的FVIIIXTEN HBoV1 ceDNA产生(参见图17A,和图17B中描绘的构建体)。

[0313] 将约2.0×10^6个细胞接种于100mL无血清ESF-921培养基中,并以1.0、2.0、3.0、4.0和5.0pfu/细胞的MOI用AcBIVVBac.mTTR.FVIIIXTEN.HBoV1.ITRs^{Tn7}和AcBIVVBac.Polh.HBoV1.NS1^{Tn7} BEV的滴定工作原液(P3或P4)共感染。在每种情况下,不去除病毒接种物,并且将细胞在28℃摇床培养箱中孵育直到活力达到60%-70%。在感染后约96h,收获感染细胞,并且处理沉淀,以用于根据制造商的说明书通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)进行FVIIIXTEN HBoV1 ceDNA分离。

[0314] 通过0.8%至1.2%琼脂糖凝胶电泳分析最终洗脱级分,以确定ceDNA生产力。如图17C中所示的琼脂糖凝胶分析显示与P3或P4(和P5,数据未显示)BEV几乎相等水平的FVIIIXTEN HBoV1 ceDNA生产力,表明编码FVIIIXTEN HBoV1 ITR的重组BEV在Sf9细胞中的随后传代中是遗传稳定的。

实施例14:从单BAC系统产生FVIIIXTEN HBoV1 ceDNA载体

[0315] HBoV1单BAC系统已显示在Sf9细胞中产生FVIIIXTEN HBoV1 ceDNA载体(参见例如图8C)。然而,使用多克隆重组BEV实现了概念验证。为了支持更大规模的生产,需要产生克隆BEV。因此,在该研究中,将HBoV1单BAC多克隆BEV进行噬斑纯化并在Sf9细胞中扩增(参见

图18A)。然后筛选这些克隆单BAC BEV,以用于在Sf9细胞中产生FVIIIXTEN HBoV1 ceDNA载体。

[0316] 如前所述进行重组HBoV1单BAC BEV的噬斑纯化和扩增。Jarvis等人(2014),Methods Enzymol.,536:149-163。通过在补充有10%胎牛血清的100mL ESF-921培养基中感染约1.0×106/mL Sf9细胞,将六个经噬斑纯化的克隆扩增到P2,并在28℃振荡培养箱中孵育4-5天或直到细胞活力达到60%-70%。在感染后4-5天,收获无细胞上清液并作为P2工作原液储存,并根据制造商的说明书,通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)处理细胞沉淀用于FVIIIXTEN HBoV1 ceDNA分离。

[0317] 通过0.8%至1.2%琼脂糖凝胶电泳分析最终洗脱级分,以确定FVIIIXTEN HBoV1 ceDNA载体生产力。图18C示出了编码具有HBoV1 ITR的FVIIIXTEN和多角体蛋白驱动的HBoV1-NS1的HBoV1单BAC(图18B所描绘的构建体)的琼脂糖凝胶分析。结果显示出不同克隆的HBoV1 ceDNA生产力的不同程度,与其他测试的克隆相比,克隆#2和克隆#4是HBoV1 ceDNA的更高生产者(图18C)。该结果显示出从相同原液获得的不同杆状病毒克隆的变异性,并突出了使用克隆重组BEV用于大规模ceDNA制造的重要性。

[0318] 为了确定克隆HBoV1单BAC BEV的最佳生产力,以0.1、0.2、0.3、0.4、0.5、1.0、2.0、3.0、4.0或5.0pfu/细胞的MOI用HBoV1单BAC BEV克隆#5的滴定工作原液(P2)感染约2.0×10^6个细胞。在每种情况下,不去除病毒接种物,并且将细胞在28℃摇床培养箱中孵育直到细胞活力达到60%-70%。在感染后约96h,收获感染细胞,并且处理沉淀,以用于根据制造商的说明书通过PureLink Maxi Prep DNA分离试剂盒(Invitrogen)进行FVIIIXTEN HBoV1 ceDNA分离。通过0.8%至1.2%琼脂糖凝胶电泳分析最终洗脱级分,以确定FVIIIXTEN HBoV1 ceDNA生产力。

[0319] 琼脂糖凝胶分析显示在所有测试剂量中,DNA条带对应于FVIIIXTEN HBoV1ceDNA(约8.5kb)的大小,随着MOI增加生产力增加。该结果与用AAV2 ITR单BAC获得的ceDNA生产力(其中观察到生产力随病毒载量增加而降低)相反。该HBoV1单BAC方法证明了从编码具有HBoV1 ITR的FVIIIXTEN和NS1转基因的单个重组BEV产生ceDNA的概念。还显示在杆状病毒穿梭载体(BIVVBac)的不同基因座处插入多个转基因的可行性和功能性。

实施例15:FVIIIXTEN HBoV1 ssDNA与ceDNA体内功效

ssFVIIIXTEN HBoV1 ITR(单链DNA)

[0320] 假设在HBoV1 ITR区域内形成的发夹以更高水平驱动转基因的长期持续表达。为了在体内验证HBoV1 ITR的功能性,在hFVIIIR593C+/+/HemA小鼠中测试包含经密码子优化的人FVIIIXTEN(ssFVIIIXTEN)和预先形成的HBoV1 ITR的单链DNA(ssDNA)。

[0321] 通过在95℃使PmlI消化的双链DNA(dsDNA)片段产物(FVIII表达盒和质粒骨架)变性,然后在4℃冷却以允许回文ITR序列折叠来产生具有预先形成的HBoV1 ITR的ssFVIIIXTEN。所得ssFVIIIXTEN通过0.8%至1.2%琼脂糖凝胶电泳证实。凝胶分析显示ssFVIIIXTEN的dsDNA大小的一半,表明有效的发夹形成(图19A)。将ssFVIIIXTEN以10μg或40μg/小鼠(分别相当于400μg或1600μg/kg)经由流体动力学尾静脉注射全身注射。以7天的间隔从注射小鼠中收集血浆样品5.5个月。根据制造商的说明书,通过Chromogenix Coatest®SP因子VIII生色测定来测量血浆FVIII活性。

[0322] 对于ssFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图

19C中。结果显示在5.5个月的历程内在HemA小鼠中具有剂量依赖性反应,在较高剂量队列中具有FVIII表达的超生理水平(>1000%正常值)。然而,直到第56天观察到FVIII表达的初始下降,并在第140天之前水平稳定,表明来自肝脏的两侧为HBoV1 ITR的ssFVIIIXTEN的持续表达。这些结果验证了HBoV1 ITR用于FVIIIXTEN在体内的长期持续表达的功能性。

ceFVIIIXTEN HBoV1 ITR(封闭末端DNA)

[0323] 在封闭末端DNA(ceDNA)和单链DNA(ssDNA)之间存在主要的结构差异,其中分别地,前者是双链,而后者是单链。这种差异可能影响表达水平以及核酸分子的稳定性。该研究显示出HBoV1 ITR在ceDNA形式中的体内功能性。为了对此进行测试,使用如实施例8中所述的双Bac方法获得ceFVIIIXTEN,将其从感染的Sf9细胞沉淀纯化,并通过0.8%至1.2%琼脂糖凝胶电泳测定质量。琼脂糖凝胶分析显示ceFVIIIXTEN纯度>90%,没有可检测的污染DNA(图19B)。

[0324] 将所得的ceFVIIIXTEN以0.3μg、1.0μg或2.0μg/小鼠(分别相当于12μg、40μg和80μg/kg)经由流体动力学尾静脉注射全身注射到hFVIIIR593C+/+/HemA小鼠中。以7天间隔收集来自注射小鼠的血浆样品,并通过生色分析测量FVIII活性,如上所述。对于ceFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图19C中。

[0325] 结果显示出在HemA小鼠中具有剂量依赖性反应,在对于ceFVIIIXTEN测试的最高剂量(80μg/kg)下FVIII表达具有超生理水平(>500%正常值)。ceDNA的FVIII表达的最高水平比1600μg/kg下ssFVIIIXTEN实现的最高水平低2倍。然而,以高得多的量给予ssDNA以实现FVIII表达的这些高水平。ceDNA显现每剂量量提供更高水平的FVIII表达。例如,400μg/kg的ssDNA和40μg/kg的ceDNA的FVIII表达水平是可比的(图19C)。这些体内研究验证了HBoV1 ITR在ssDNA或ceDNA形式中的功能性,并显示HBoV1 ITR可用于在杆状病毒昆虫细胞系统中产生编码目的转基因的功能性ceDNA。

实施例16:FVIIIXTEN HBoV1单体与多聚体ceDNA体内功效

[0326] 已显示重组AAV基因组在附加体中持续存在,并且认为它们的附加体存在与长期转基因表达相关。这些基因组显现通过单体环化过程产生,产生头对尾AAV环状基因组。然而,随着时间的推移,单体环状中间体下降,有利于高分子量环状多联体。其他细节披露于Duan等人(1998),J Virol.72(11);8568-8577中。目前关于封闭末端DNA(ceDNA)的附加体存在和在体内ceDNA的单体形式优于多联体形式的益处知之甚少。

[0327] 进行该研究以通过在hFVIIIR593C+/+/HemA小鼠中经由流体动力学尾静脉注射测试两种形式的FVIIIXTEN HBoV1 ceDNA(ceFVIIIXTEN)来确定单体形式和多聚体形式的ceDNA的体内影响。

[0328] 通过PAGE纯化产生ceFVIIIXTEN的单体和多聚体形式,如前所述(参见国际申请号PCT/US2021/047218)。通过0.8%至1.2%琼脂糖凝胶电泳测定ceFVIIIXTEN的多联体形式的质量,结果显示大多数物种是ceFVIIIXTEN的单体或多聚体形式(图20A)。将纯化的单体或多聚体ceFVIIIXTEN经由流体动力学尾静脉注射以40μg/kg全身注射到hFVIIIR593C+/+/HemA小鼠中。以7天的间隔从注射小鼠中收集血浆样品约3个月。根据制造商的说明书,通过ChromogenixCoatest®SP因子VIII生色测定来测量血浆FVIII活性。

[0329] 对于ceFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图20B中。结果显示在3个月的历程中,单体或多聚体形式的ceFVIIIXTEN之间在FVIII表达水

平上没有显著差异。该数据表明,单体和多聚体形式的ceFVIIIXTEN具有可比的体内效力和稳定性。

实施例17:FVIIIXTEN HBoV1 mTTR与A1AT ssDNA体内功效

[0330] 用于上文公开的实验的FVIIIXTEN表达盒含有mTTR启动子和增强子元件(V2.0,图1)。该启动子是小鼠肝特异性的,但其肝特异性表达尚未在大型动物模型或人受试者中研究。因此,在该研究中,通过用V2.0表达盒中的人肝特异性α-1-抗胰蛋白酶(A1AT)启动子(SEQ ID NO:36)替代mTTR启动子和增强子元件来产生V3.0 FVIIIXTEN表达盒(SEQ ID NO:35)(图1)。

[0331] 为了在体内验证mTTR相对于A1AT启动子的功能性,在hFVIIIR593C+/+/HemA小鼠中测试包含经密码子优化的人FVIIIXTEN(ssFVIIIXTEN)与预先形成的HBoV1 ITR的单链DNA(ssDNA)(图21A)。通过在95℃使PmlI消化的双链DNA(dsDNA)片段产物(mTTR或A1AT FVIII表达盒和质粒骨架)变性,然后在4℃冷却以允许回文ITR序列折叠来产生具有预先形成的HBoV1 ITR的ssFVIIIXTEN。将所得ssFVIIIXTEN通过0.8%至1.2%琼脂糖凝胶电泳检测。凝胶分析显示ssFVIIIXTEN的dsDNA大小的一半,表明有效的发夹形成(图21B)。将ssFVIIIXTEN以10μg/小鼠经由流体动力学尾静脉注射全身注射到hFVIIIR593C+/+/HemA小鼠中。以7天的间隔从注射小鼠中收集血浆样品5.5个月。根据制造商的说明书,通过ChromogenixCoatest®SP因子VIII生色测定来测量血浆FVIII活性。

[0332] 对于ssFVIIIXTEN注射的动物,归一化为占正常值百分比的血浆FVIII活性示于图21C中。这些结果显示,直至注射后第21天,FVIII表达水平相当,表明hFVIIIR593C+/+/HemA小鼠动物模型中mTTR或A1AT启动子表达的FVIIIXTEN水平没有显著差异。

序列

表2:其他核苷酸和氨基酸序列

SEQ ID NO / 描述	核苷酸或氨基酸序列
SEQ ID NO.1: HBoV1 5' ITR	GTGGTTGTACAGACGCCATCTTGGAATCCAATATGTCTGCCGGCTCAGT CATGCCTGCGCTGCGCGCAGCGCGCTGCGCGCGCGCATGATCTAATCGC CGGCAGACATATTGGATTCCAAGATGGCGTCTGTACAACCAC
SEQ ID NO.2: HBoV1 3' ITR	TTGCTTATGCAATCGCGAAACTCTATATCTTTTAATGTGTTGTTGTTGTA CATGCGCCATCTTAGTTTTATATCAGCTGGCGCCTTAGTTATATAACATG CATGTTATATAACTAAGGCGCCAGCTGATATAAAACTAAGATGGCGCAT GTACAACAACAACACATTAAAAGATATAGAGTTTCGCGATTGCATAAGC AA
SEQ ID NO.3: HBoV1-5'ITR-mTTR482-内含子-coBDDFVIIIXTEN (V2.0)-WPRE-bGHPolyA-HBoV1-3'ITR	GTATACCTGCAGGCTAGCCACGTGTTGTTGTTGTACATGCGCCATCTTAG TTTTATATCAGCTGGCGCCTTAGTTATATAACATGCATGTTATATAACTA AGGCGCCAGCTGATATAAAACTAAGATGGCGCATGTACAACAACAACA CATTAAAAGATATAGAGTTTCGCGATTGCAAGCTTGGCCCCAGGTTAAT TTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTTGGCAGCATTTACTCT CTCTATTGACTTTGGTTAATAATCTCAGGAGCACAAACATTCCTGGAGG CAGGAGAAGAAATCAACATCCTGGACTTATCCTCTGGGCCTCTCCCCAC CTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTG GCCCTTGGCAGCATTTACTCTCTCTATTGACTTTGGTTAATAATCTCAGG AGCACAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTT ATCCTCTGGGCCTCTCCCCACCGATATCTACCTGCTGATCGCCCGGCCCC TGTTCAAACATGTCCTAATACTCTGTCGGGGCAAAGGTCGGCAGTAGTT TTCCATCTTACTCAACATCCTCCCAGTGTACGTAGGATCCTGTCTGTCTG CACATTTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGGCAAAG GTCGTATTGACTTAGGTTACTTATTCTCCTTTTGTTGACTAAGTCAATAA TCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGG GTTGGAAGGAGGGGGTATAAAAGCCCCTTCACCAGGAGAAGCCGTCAC ACAGATCCACAAGCTCCTGCTAGGAATTCTCAGGAGCACAAACATTCCT GGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCTGGGCCTCTC CCCACCGATATCTACCTGCTGATCGCCCGGCCCCTGTTCAAACATGTCCT AATACTCTGTCGGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCAAC ATCCTCCCAGTGTACGTAGGATCCTGTCTGTCTGCACATTTCGTAGAGCG AGTGTTCCGATACTCTAATCTCCCGGGGCAAAGGTCGTATTGACTTAGG TTACTTATTCTCCTTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTT TGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGT ATAAAAGCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTC CTGCTAGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCCCCGCTCCGCCGC CGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAG GTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGG TTTATTGACGGCTTGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTGAGGGG CTCCGGGAAGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGC GTGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCTCCGCGCTGCCCGGCG GCTGTGAGCGCTGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGT GCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGG GCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGT GAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACCCCCCCTGCACC CCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCG TACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA GGTGGGGGTGCCGGGCGGGGCGGGGCCGCCTCGGGCCGGGGAGGGCTC GGGGGAGGGGCGCGGCGGCCCCCGGAGCGCCGGCGGCTGTCGAGGCGC GGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAG GGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGAGGCGCC GCCGCACCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCGGCA

	GGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTC CCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCT TCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGG CGGCTCTAGAGCCTCTGCTAACCTTGTTCTTGCCTTCTTCTTTTTCCTACA GCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAG AATTACTCGAGGCCACCATGCAGATTGAACTGTCCACTTGCTTCTTCCTG TGCCTCCTGCGGTTTTGCTTCTCGGCCACCCGCCGGTATTACTTAGGTGC TGTGGAACTGAGCTGGGACTACATGCAGTCCGACCTGGGAGAACTGCC GGTGGACGCGAGATTCCCACCTAGAGTCCCGAAGTCCTTCCCATTCAAC ACCTCCGTGGTCTACAAAAAGACCCTGTTCGTGGAGTTCACTGACCACC TTTTCAATATTGCCAAGCCGCGCCCCCCCTGGATGGGCCTGCTTGGTCCT ACGATCCAAGCAGAGGTCTACGACACCGTGGTCATCACACTGAAGAAC ATGGCCTCACACCCCGTGTCGCTGCATGCTGTGGGAGTGTCCTACTGGA AGGCCTCAGAGGGTGCCGAATATGATGACCAGACCAGCCAGAGGGAAA AGGAGGATGACAAAGTGTTCCCGGGTGGCAGCCACACTTACGTGTGGC AAGTGCTGAAGGAAAACGGGCCTATGGCGTCGGACCCCCTATGCCTGA CCTACTCCTACCTGTCCCATGTGGACCTTGTGAAGGATCTCAACTCGGG ACTGATCGGCGCCCTCTTGGTGTGCAGAGAAGGCAGCCTGGCGAAGGA AAAGACTCAGACCCTGCACAAGTTCATTCTGTTGTTTGCTGTGTTCGATG AAGGAAAGTCCTGGCACTCAGAAACCAAGAACTCGCTGATGCAGGATA GAGATGCGGCCTCGGCCAGAGCCTGGCCTAAAATGCACACCGTCAACG GATATGTGAACAGGTCGCTCCCTGGCCTCATCGGCTGCCACAGAAAGTC CGTGTATTGGCATGTGATCGGCATGGGTACTACTCCGGAAGTGCATAGT ATCTTTCTGGAGGGCCATACCTTCTTGGTGCGCAACCACAGACAGGCCT CGCTGGAAATCTCGCCTATCACTTTCTTGACTGCGCAGACCCTCCTTATG GACCTTGGACAGTTCCTGCTGTTCTGTCACATCAGCTCCCATCAGCATGA TGGGATGGAGGCCTATGTCAAAGTGGACTCCTGCCCTGAGGAGCCACA GCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGACGACGACCT GACTGACAGCGAAATGGACGTCGTGCGATTCGATGACGACAACAGCCC GTCCTTCATCCAAATTAGATCAGTGGCGAAGAAGCACCCCAAGACCTGG GTGCACTACATTGCCGCCGAGGAAGAGGACTGGGACTACGCGCCGCTG GTGCTGGCGCCAGACGACAGGAGCTACAAGTCCCAGTACCTCAACAAC GGGCCGCAGCGCATTGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCC TACACTGATGAAACCTTCAAGACGAGGGAAGCCATCCAGCACGAGTCA GGCATCCTGGGACCGCTCCTTTACGGCGAAGTCGGGGATACCCTGCTCA TCATTTTCAAGAACCAGGCATCGCGGCCCTACAACATCTACCCTCACGG GATCACAGACGTGCGCCCGCTCTACTCCCGCCGGCTGCCCAAGGGAGTG AAGCACCTGAAGGATTTTCCCATCCTGCCGGGAGAAATCTTCAAGTACA AGTGGACCGTGACTGTGGAAGATGGCCCTACCAAGTCGGACCCTCGCTG TCTGACCCGGTACTATTCCTCGTTTGTGAACATGGAGCGCGACCTGGCC TCGGGGCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACC AGCGCGGGAACCAGATCATGTCCGACAAGCGCAACGTGATCCTGTTCTC TGTCTTTGATGAAAACAGATCGTGGTACTTGACTGAGAATATCCAGCGG TTCCTGCCCAACCCAGCGGGAGTGCAACTGGAGGACCCGGAGTTCCAG GCCTCAAACATTATGCACTCTATCAACGGCTATGTGTTCGACTCGCTCCA ACTGAGCGTGTGCCTGCATGAAGTGGCATACTGGTACATTCTGTCCATC GGAGCCCAGACCGACTTCCTGTCCGTGTTCTTCTCCGGATACACCTTCAA GCATAAGATGGTGTACGAGGACACTCTGACCCTCTTCCCATTTTCCGGA GAAACTGTGTTCATGTCAATGGAAAACCCGGGCTTGTGGATTCTGGGTT GCCATAACTCGGACTTCCGGAATAGAGGGATGACCGCCCTGCTGAAAGT GTCCAGCTGTGACAAGAATACCGGCGATTACTACGAGGACAGCTATGA GGACATCTCCGCTTATCTGCTGTCCAAGAACAACGCCATTGAACCCAGG TCCTTCTCCCAAAACGGTGCACCGACCTCCGAAAGCGCCACCCCAGAGT CAGGACCTGGCTCGGAACCGGCTACCTCGGGCTCAGAGACACCGGGGA CTTCCGAGTCCGCAACCCCCGAGAGTGGACCCGGATCCGAACCAGCAA CCTCAGGATCAGAAACCCCGGGAACTTCGGAATCCGCCACTCCCGAGTC GGGACCAGGCACCTCCACTGAGCCTTCCGAGGGAAGCGCCCCCGGATC CCCTGCTGGATCCCCTACCAGCACTGAAGAAGGCACCTCAGAATCCGCG ACCCCTGAGTCCGGCCCTGGAAGCGAACCCGCCACCTCCGGTTCCGAAA

	CCCCTGGGACTAGCGAGAGCGCCACTCCGGAATCGGGCCCAGGAAGCC CTGCCGGATCCCCGACCAGCACCGAGGAGGGAAGCCCCGCCGGGTCAC CGACTTCCACTGAGGAGGGAGCCTCATCCCCCCCCGTGCTGAAGCGGCA TCAAAGAGAGATCACCAGGACCACTCTCCAGTCCGATCAGGAAGAAAT TGACTACGACGATACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGA CATCTACGATGAGGATGAGAACCAGTCCCCTCGGAGCTTTCAGAAGAA AACCCGCCACTACTTCATCGCTGCCGTGGAGCGGCTGTGGGATTACGGG ATGTCCAGCTCACCGCATGTGCTGCGGAATAGAGCGCAGTCAGGATCGG TGCCCCAGTTCAAGAAGGTCGTGTTCCAAGAGTTCACCGACGGGTCCTT CACTCAACCCCTGTACCGGGGCGAACTCAACGAACACCTGGGACTGCTT GGGCCGTATATCAGGGCAGAAGTGGAAGATAACATCATGGTCACCTTCC GCAACCAGGCCTCCCGGCCGTACAGCTTCTACTCTTCACTGATCTCCTAC GAGGAAGATCAGCGGCAGGGAGCCGAGCCCCGGAAGAACTTCGTCAAG CCTAACGAAACTAAGACCTACTTTTGGAAGGTCCAGCATCACATGGCCC CGACCAAAGACGAGTTCGACTGTAAAGCCTGGGCCTACTTCTCCGATGT GGACCTGGAGAAGGACGTGCACTCGGGACTCATTGGCCCGCTCCTTGTG TGCCATACTAATACCCTGAACCCTGCTCACGGTCGCCAAGTCACAGTGC AGGAGTTCGCCCTCTTCTTCACCATCTTCGATGAAACAAAGTCCTGGTA CTTTACTGAGAACATGGAACGCAATTGCAGGGCACCCTGCAACATCCAG ATGGAAGATCCCACCTTCAAGGAAAACTACCGGTTTCATGCCATTAACG GCTACATAATGGACACGTTGCCAGGACTGGTCATGGCCCAGGACCAGA GAATCCGGTGGTATCTGCTCTCCATGGGCTCCAACGAAAACATTCACAG CATTCATTTTTCCGGCCATGTGTTCACCGTCCGGAAGAAGGAAGAGTAC AAGATGGCTCTGTACAACCTCTACCCTGGAGTGTTCGAGACTGTGGAAA TGCTGCCTAGCAAGGCCGGCATTTGGAGAGTGGAATGCCTGATCGGAG AGCATTTGCACGCCGGAATGTCCACCCTGTTTCTTGTGTACTCCAACAA GTGCCAGACCCCGCTGGGAATGGCCTCAGGTCATATTAGGGATTTCCAG ATCACTGCTTCGGGGCAGTACGGGCAGTGGGCACCTAAGTTGGCCCGGC TGCACTACTCTGGCTCCATCAATGCCTGGTCCACCAAGGAACCCTTCTCC TGGATTAAGGTGGACCTCCTGGCCCCAATGATTATTCACGGTATTAAGA CCCAGGGTGCCCGACAGAAGTTCTCCTCACTCTACATCTCGCAATTCAT CATAATGTACAGCCTGGATGGGAAGAAGTGGCAGACCTACCGGGGAAA CTCCACTGGAACGCTCATGGTGTTTTTCGGCAACGTGGACTCCTCCGGC ATTAAGCACAACATCTTCAACCCTCCGATCATTGCTCGGTACATCCGGC TGCACCCAACTCACTACAGCATCCGGTCCACCCTGCGGATGGAACTGAT GGGTTGTGACCTGAACTCCTGCTCCATGCCCCTTGGGATGGAATCCAAG GCCATTAGCGATGCACAGATCACCGCCTCTTCATACTTCACCAACATGT TCGCGACCTGGTCCCCGTCGAAGGCCCGCCTGCACCTCCAAGGTCGCTC CAATGCGTGGCGGCCTCAAGTGAACAACCCCAAGGAGTGGCTCCAGGT CGACTTCCAAAAGACCATGAAGGTCACCGGAGTGACCACCCAGGGCGT GAAGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCCTCATCTCCTCAA GCCAAGACGGACATCAGTGGACCCTGTTCTTCCAAAACGGAAAAGTCA AAGTATTCCAGGGCAACCAGGACTCCTTCACCCCTGTGGTCAACAGCCT GGACCCCCCATTGCTGACCCGCTACCTCCGCATCCACCCCCAAAGCTGG GTCCACCAGATCGCACTGCGCATGGAGGTCCTTGGATGCGAAGCCCAAG ATCTGTACTAAGCGGCCGCTCATAATCAACCTCTGGATTACAAAATTTG TGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTG GATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCT TTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAG TTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTG ACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTC CGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCG CCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAA TTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCT GTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCG GCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGC GGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTT TGGGCCGCCTCCCCGCTGCCTAGGCGACTGTGCCTTCTAGTTGCCAGCC ATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCA

	CTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAG GGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAAGACCATGGG CGCGCCAGGCCTGTCGACGCCCGGGCGGTACCGCGATCGCTCGCGACGC ATAAAGTATATGTGACGTGGTTGTACAGACGCCATCTTGGAATCCAATA TGTCTGCCGGCGATTAGATCATGCGCGCGCGCAGCGCGCTGCGCGCAGC GCAGGCATGACTGAGCCGGCAGACATATTGGATTCCAAGATGGCGTCTG TACAACCACGTGCTTAAGCTGCAGACTAGTGAGCTCGTTAAC
SEQ ID NO.4: Sf密码子优化 HBoV1 NS1	GCGGCCGCGGATCCGCCACCATGGCATTCAATCCGCCCGTAATACGCGC ATTTTCACAACCCGCCTTTACGTATGTCTTTAAGTTTCCGTACCCTCAAT GGAAAGAGAAAGAGTGGCTACTGCACGCGTTGCTTGCCCACGGCACCG AGCAGTCCATGATTCAATTACGTAACTGTGCCCCACACCCGGACGAGGA TATTATCCGGGACGATCTTCTAATAAGTTTGGAAGATAGGCATTTCGGG GCGGTCCTGTGTAAAGCGGTATACATGGCTACTACCACGTTGATGTCTC ACAAGCAACGCAATATGTTCCCAAGGTGCGACATAATCGTTCAGTCAGA GTTAGGTGAAAAAAATTTACATTGTCATATTATCGTTGGAGGCGAAGGC CTATCAAAGAGAAACGCTAAGAGCTCTTGCGCTCAGTTTTACGGACTTA TATTAGCAGAAATTATCCAGCGCTGTAAGAGTTTACTAGCCACCCGTCC GTTTGAGCCGGAAGAAGCGGATATATTTCATACGTTGAAGAAAGCGGA GCGCGAGGCCTGGGGTGGAGTTACTGGCGGTAACATGCAAATCTTACA ATACAGGGACCGTCGGGGTGACCTGCATGCACAGACTGTTGATCCCCTC AGATTCTTCAAAAATTATTTGTTACCGAAGAACCGATGCATAAGTAGTT ACAGCAAACCTGATGTCTGTACTAGCCCTGATAACTGGTTCATTCTGGC CGAAAAAACGTACTCGCATACACTTATCAATGGATTGCCGCTTCCCGAG CACTATCGAAAAAACTATCATGCCACCCTGGATAATGAAGTTATACCTG GACCACAGACTATGGCGTATGGAGGGAGAGGCCCTTGGGAACATTTAC CCGAGGTGGGTGACCAGAGGCTTGCCGCAAGTTCCGTGAGCACTACGTA TAAGCCAAACAAGAAGGAGAAGCTAATGCTCAACCTCCTCGACAAGTG TAAGGAGTTGAATCTTCTAGTTTATGAGGATCTTGTAGCGAACTGCCCA GAGCTGCTGCTCATGCTAGAAGGCCAACCTGGAGGTGCTCGACTCATCG AGCAAGTACTAGGAATGCATCACATCAATGTATGCTCGAATTTCACCGC GCTAACGTACCTCTTCCATCTGCATCCGGTGACATCGCTGGATAGTGAC AACAAAGCGTTACAGCTTTTACTAATTCAAGGGTACAACCCCCTGGCAG TGGGGCATGCTCTCTGTTGTGTGTTAAACAAACAATTTGGTAAACAGAA CACAGTCTGTTTTTACGGGCCAGCATCTACTGGGAAAACAAATATGGCA AAAGCGATTGTGCAGGGAATCCGGCTATATGGCTGCGTCAACCATCTTA ACAAGGGTTTTGTTTTCAATGATTGTCGACAACGCCTCGTAGTCTGGTG GGAGGAATGCCTAATGCACCAGGACTGGGTGGAGCCAGCAAAGTGTAT TCTTGGCGGGACCGAATGTCGTATCGACGTCAAGCACAGAGATTCTGTC CTATTGACACAAACGCCTGTAATAATTTCGACTAATCACGACATTTACG CCGTCGTGGGAGGGAATTCGGTGTCTCACGTTCACGCTGCGCCTCTCAA AGAACGGGTTATTCAGCTGAATTTTATGAAACAACTCCCCCAAACTTTT GGTGAGATAACCGCCACAGAAATCGCTGCTCTGCTACAGTGGTGCTTTA ATGAATATGACTGCACCCTGACAGGTTTCAAACAGAAGTGGAATTTGGA CAAGATACCTAACTCATTCCCGTTGGGGGTATTGTGCCCAACACATTCC CAAGATTTCACACTTCACGAAAATGGGTATTGCACGGACTGCGGGGGCT ACCTTCCCCACTCCGCTGATAATTCAATGTATACCGATCGGGCTAGCGA AACATCCACCGGCGACATAACGCCCTCCAAATGATTCGAATCTAGAGCC TGCAGTCTCGAGGCATGCGGTACC
SEQ ID NO.5: 外侧引物	GTGGACGTGAAAGAAACC
SEQ ID NO.6: 内侧引物	GGTCATAGCTGTTTCCTGTG
SEQ ID NO.7:	ATTAAGCTTCCGCGTAAAACACAATCAAGTATGAGTCATAAGCTGATGT

hr5.ie1.neo.p10PAS	CATGTTTTGCACACGGCTCATAACCGAACTGGCTTTACGAGTAGAATTC TACTTGTAACGCACGATCAGTGGATGATGTCATTTGTTTTTCAAATCGAG ATGATGTCATGTTTTGCACACGGCTCATAAACTCGCTTTACGGGTAGAA TTCTACGTGTAACGCACGATCGATTGATGAGTCATTTGTTTTGCAATATG ATATCATACAATATGACTCATTTGTTTTTCAAAACCGAACTTGATTTACG GGTAGAATTCTACTTGTAAAGCACAATCAAAAAGATGATGTCATTTGTT TTTCAAAACTGAACTCGCTTTACGAGTAGAATTCTACGTGTAAAACACA ATCAAGAAATGATGTCATTTGTTATAAAAATAAAAGCTGATGTCATGTT TTGCACATGGCTCATAACTAAACTCGCTTTACGGGTAGAATTCTACGCG CGTCGATGTCTTTGTGATGCGCGCGACATTTTTGTAGGTTATTGATAAAA TGAACGGATACGTTGCCCGACATTATCATTAAATCCTTGGCGTAGAATT TGTCGGGTCCATTGTCCGTGTGCGCTAGCATGCCCGTAACGGACCTCGT ACTTTTGGCTTCAAAGGTTTTGCGCACAGACAAAATGTGCCACACTTGC AGCTCTGCATGTGTGCGCGTTACCACAAATCCCAACGGCGCAGTGTACT TGTTGTATGCAAATAAATCTCGATAAAGGCGCGGCGCGCGAATGCAGCT GATCACGTACGCTCCTCGTGTTCCGTTCAAGGACGGTGTTATCGACCTC AGATTAATGTTTATCGGCCGACTGTTTTCGTATCCGCTCACCAAACGCGT TTTTGCATTAACATTGTATGTCGGCGGATGTTCTATATCTAATTTGAATA AATAAACGATAACCGCGTTGGTTTTAGAGGGCATAATAAAAGAAATATT GTTATCGTGTTCGCCATTAGGGCAGTATAAATTGACGTTCATGTTGGAT ATTGTTTCAGTTGCAAGTTGACACTGGCGGCGACAAGATCGTGAACAAC CAAGTGACGCGGCCGCATTTGTAAAAAAAAAATAAATAAAAATGATCG AGCAGGACGGCCTGCACGCTGGTTCTCCAGCTGCTTGGGTCGAGCGTCT GTTCGGTTACGACTGGGCTCAGCAGACCATCGGTTGCTCCGACGCTGCT GTGTTCCGTCTGTCCGCTCAGGGTCGTCCCGTGCTGTTCGTCAAGACCGA CCTGTCCGGTGCTCTGAACGAGCTGCAGGACGAGGCTGCTCGTCTGTCC TGGCTGGCTACCACTGGTGTCCCTTGCGCTGCTGTCCTGGACGTGGTCAC TGAGGCTGGTCGTGACTGGCTGCTGCTGGGAGAAGTGCCTGGACAGGA CCTGCTGTCCAGCCACCTGGCTCCAGCTGAGAAGGTGTCCATCATGGCT GACGCTATGCGTCGTCTGCACACCCTGGACCCTGCTACCTGCCCCTTCG ACCACCAAGCTAAGCACCGTATCGAGCGTGCTCGTACCCGTATGGAAGC TGGCCTGGTGGACCAGGACGACCTGGACGAAGAACACCAGGGACTGGC CCCTGCTGAGCTGTTCGCTCGTCTGAAGGCTCGTATGCCCGACGGCGAG GACCTGGTGGTTACTCACGGCGACGCTTGCCTGCCCAACATCATGGTCG AGAACGGTCGTTTCTCCGGTTTCATCGACTGCGGTCGTCTGGGTGTCGCT GACCGTTACCAGGATATCGCTCTGGCTACCCGTGATATCGCTGAGGAAC TGGGTGGCGAGTGGGCTGACAGATTCCTGGTGCTGTACGGTATCGCTGC TCCCGACTCCCAGCGTATCGCTTTCTACCGTCTGCTGGACGAGTTCTTCT AAGCCCCTTGTAAACGCCACAATTGTGTTTGTTGCAAATAAACCCATGA TTATTTGATTAAAATTGTTGTTTTCTTTGTTCATAGACAATAGTGTGTTTT GCCTAAACGGGTACC
SEQ ID NO.8: hr5.ie1.eGFP.p10PAS	ATTAAGCTTCCGCGTAAAACACAATCAAGTATGAGTCATAAGCTGATGT CATGTTTTGCACACGGCTCATAACCGAACTGGCTTTACGAGTAGAATTC TACTTGTAACGCACGATCAGTGGATGATGTCATTTGTTTTTCAAATCGAG ATGATGTCATGTTTTGCACACGGCTCATAAACTCGCTTTACGGGTAGAA TTCTACGTGTAACGCACGATCGATTGATGAGTCATTTGTTTTGCAATATG ATATCATACAATATGACTCATTTGTTTTTCAAAACCGAACTTGATTTACG GGTAGAATTCTACTTGTAAAGCACAATCAAAAAGATGATGTCATTTGTT TTTCAAAACTGAACTCGCTTTACGAGTAGAATTCTACGTGTAAAACACA ATCAAGAAATGATGTCATTTGTTATAAAAATAAAAGCTGATGTCATGTT TTGCACATGGCTCATAACTAAACTCGCTTTACGGGTAGAATTCTACGCG CGTCGATGTCTTTGTGATGCGCGCGACATTTTTGTAGGTTATTGATAAAA TGAACGGATACGTTGCCCGACATTATCATTAAATCCTTGGCGTAGAATT TGTCGGGTCCATTGTCCGTGTGCGCTAGCATGCCCGTAACGGACCTCGT ACTTTTGGCTTCAAAGGTTTTGCGCACAGACAAAATGTGCCACACTTGC AGCTCTGCATGTGTGCGCGTTACCACAAATCCCAACGGCGCAGTGTACT TGTTGTATGCAAATAAATCTCGATAAAGGCGCGGCGCGCGAATGCAGCT

	GATCACGTACGCTCCTCGTGTTCCGTTCAAGGACGGTGTTATCGACCTC AGATTAATGTTTATCGGCCGACTGTTTTCGTATCCGCTCACCAAACGCGT TTTTGCATTAACATTGTATGTCGGCGGATGTTCTATATCTAATTTGAATA AATAAACGATAACCGCGTTGGTTTTAGAGGGCATAATAAAAGAAATATT GTTATCGTGTTCGCCATTAGGGCAGTATAAATTGACGTTCATGTTGGAT ATTGTTTCAGTTGCAAGTTGACACTGGCGGCGACAAGATCGTGAACAAC CAAGTGACGCGGCCGCATTTGTAAAAAAAAAATAAATAAAAATGGTGT CCAAGGGCGAGGAACTGTTCACCGGTGTCGTGCCCATCCTGGTCGAACT GGACGGCGACGTGAACGGTCACAAGTTCTCCGTGTCTGGCGAAGGCGA GGGCGACGCTACCTACGGAAAGCTGACCCTGAAGTTCATCTGCACCACC GGCAAGCTGCCCGTGCCTTGGCCTACCCTGGTCACCACTCTGACCTACG GTGTCCAGTGCTTCTCCCGTTACCCCGACCACATGAAGCAGCACGATTT CTTCAAGTCCGCTATGCCCGAGGGTTACGTGCAAGAGCGTACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGTGCTGAAGTGAAGTTCGAA GGCGACACCCTCGTGAACCGTATCGAGCTGAAGGGTATCGACTTCAAGG AAGATGGAAACATCCTGGGCCACAAGCTCGAGTACAACTACAACTCCC ACAACGTGTACATCATGGCCGACAAGCAAAAGAACGGCATCAAAGTGA ACTTCAAGATCCGCCACAACATCGAGGACGGTTCCGTGCAGCTGGCTGA CCACTACCAGCAGAACACCCCCATCGGCGACGGTCCTGTGCTGCTGCCT GACAACCACTACCTGTCCACCCAGTCCGCTCTGTCCAAGGACCCCAACG AGAAGCGTGACCACATGGTGCTGCTCGAGTTCGTGACCGCTGCTGGTAT CACCCTGGGCATGGACGAGCTGTACAAGTAAGCCCCTTGTAAACGCCAC AATTGTGTTTGTTGCAAATAAACCCATGATTATTTGATTAAAATTGTTGT TTTCTTTGTTCATAGACAATAGTGTGTTTTGCCTAAACGGGTACC
SEQ ID NO: 9 编码coBDDFVIIIXTEN (V2.0)的核苷酸序列	ATGCAGATTGAACTGTCCACTTGCTTCTTCCTGTGCCTCCTGCGGTTTTG CTTCTCGGCCACCCGCCGGTATTACTTAGGTGCTGTGGAACTGAGCTGG GACTACATGCAGTCCGACCTGGGAGAACTGCCGGTGGACGCGAGATTC CCACCTAGAGTCCCGAAGTCCTTCCCATTCAACACCTCCGTGGTCTACA AAAAGACCCTGTTCGTGGAGTTCACTGACCACCTTTTCAATATTGCCAA GCCGCGCCCCCCCTGGATGGGCCTGCTTGGTCCTACGATCCAAGCAGAG GTCTACGACACCGTGGTCATCACACTGAAGAACATGGCCTCACACCCCG TGTCGCTGCATGCTGTGGGAGTGTCCTACTGGAAGGCCTCAGAGGGTGC CGAATATGATGACCAGACCAGCCAGAGGGAAAAGGAGGATGACAAAGT GTTCCCGGGTGGCAGCCACACTTACGTGTGGCAAGTGCTGAAGGAAAA CGGGCCTATGGCGTCGGACCCCCTATGCCTGACCTACTCCTACCTGTCCC ATGTGGACCTTGTGAAGGATCTCAACTCGGGACTGATCGGCGCCCTCTT GGTGTGCAGAGAAGGCAGCCTGGCGAAGGAAAAGACTCAGACCCTGCA CAAGTTCATTCTGTTGTTTGCTGTGTTCGATGAAGGAAAGTCCTGGCACT CAGAAACCAAGAACTCGCTGATGCAGGATAGAGATGCGGCCTCGGCCA GAGCCTGGCCTAAAATGCACACCGTCAACGGATATGTGAACAGGTCGCT CCCTGGCCTCATCGGCTGCCACAGAAAGTCCGTGTATTGGCATGTGATC GGCATGGGTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGGGCCATA CCTTCTTGGTGCGCAACCACAGACAGGCCTCGCTGGAAATCTCGCCTAT CACTTTCTTGACTGCGCAGACCCTCCTTATGGACCTTGGACAGTTCCTGC TGTTCTGTCACATCAGCTCCCATCAGCATGATGGGATGGAGGCCTATGT CAAAGTGGACTCCTGCCCTGAGGAGCCACAGCTCCGGATGAAGAACAA TGAGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGA CGTCGTGCGATTCGATGACGACAACAGCCCGTCCTTCATCCAAATTAGA TCAGTGGCGAAGAAGCACCCCAAGACCTGGGTGCACTACATTGCCGCC GAGGAAGAGGACTGGGACTACGCGCCGCTGGTGCTGGCGCCAGACGAC AGGAGCTACAAGTCCCAGTACCTCAACAACGGGCCGCAGCGCATTGGC AGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAACCTTCA AGACGAGGGAAGCCATCCAGCACGAGTCAGGCATCCTGGGACCGCTCC TTTACGGCGAAGTCGGGGATACCCTGCTCATCATTTTCAAGAACCAGGC ATCGCGGCCCTACAACATCTACCCTCACGGGATCACAGACGTGCGCCCG CTCTACTCCCGCCGGCTGCCCAAGGGAGTGAAGCACCTGAAGGATTTTC CCATCCTGCCGGGAGAAATCTTCAAGTACAAGTGGACCGTGACTGTGGA

	AGATGGCCCTACCAAGTCGGACCCTCGCTGTCTGACCCGGTACTATTCC TCGTTTGTGAACATGGAGCGCGACCTGGCCTCGGGGCTGATTGGTCCGC TGCTGATCTGCTACAAGGAGTCCGTGGACCAGCGCGGGAACCAGATCAT GTCCGACAAGCGCAACGTGATCCTGTTCTCTGTCTTTGATGAAAACAGA TCGTGGTACTTGACTGAGAATATCCAGCGGTTCCTGCCCAACCCAGCGG GAGTGCAACTGGAGGACCCGGAGTTCCAGGCCTCAAACATTATGCACTC TATCAACGGCTATGTGTTCGACTCGCTCCAACTGAGCGTGTGCCTGCAT GAAGTGGCATACTGGTACATTCTGTCCATCGGAGCCCAGACCGACTTCC TGTCCGTGTTCTTCTCCGGATACACCTTCAAGCATAAGATGGTGTACGA GGACACTCTGACCCTCTTCCCATTTTCCGGAGAAACTGTGTTCATGTCAA TGGAAAACCCGGGCTTGTGGATTCTGGGTTGCCATAACTCGGACTTCCG GAATAGAGGGATGACCGCCCTGCTGAAAGTGTCCAGCTGTGACAAGAA TACCGGCGATTACTACGAGGACAGCTATGAGGACATCTCCGCTTATCTG CTGTCCAAGAACAACGCCATTGAACCCAGGTCCTTCTCCCAAAACGGTG CACCGACCTCCGAAAGCGCCACCCCAGAGTCAGGACCTGGCTCGGAAC CGGCTACCTCGGGCTCAGAGACACCGGGGACTTCCGAGTCCGCAACCCC CGAGAGTGGACCCGGATCCGAACCAGCAACCTCAGGATCAGAAACCCC GGGAACTTCGGAATCCGCCACTCCCGAGTCGGGACCAGGCACCTCCACT GAGCCTTCCGAGGGAAGCGCCCCCGGATCCCCTGCTGGATCCCCTACCA GCACTGAAGAAGGCACCTCAGAATCCGCGACCCCTGAGTCCGGCCCTG GAAGCGAACCCGCCACCTCCGGTTCCGAAACCCCTGGGACTAGCGAGA GCGCCACTCCGGAATCGGGCCCAGGAAGCCCTGCCGGATCCCCGACCA GCACCGAGGAGGGAAGCCCCGCCGGGTCACCGACTTCCACTGAGGAGG GAGCCTCATCCCCCCCCGTGCTGAAGCGGCATCAAAGAGAGATCACCA GGACCACTCTCCAGTCCGATCAGGAAGAAATTGACTACGACGATACTAT CAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATGAGGATGA GAACCAGTCCCCTCGGAGCTTTCAGAAGAAAACCCGCCACTACTTCATC GCTGCCGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTCACCGCATG TGCTGCGGAATAGAGCGCAGTCAGGATCGGTGCCCCAGTTCAAGAAGG TCGTGTTCCAAGAGTTCACCGACGGGTCCTTCACTCAACCCCTGTACCG GGGCGAACTCAACGAACACCTGGGACTGCTTGGGCCGTATATCAGGGC AGAAGTGGAAGATAACATCATGGTCACCTTCCGCAACCAGGCCTCCCGG CCGTACAGCTTCTACTCTTCACTGATCTCCTACGAGGAAGATCAGCGGC AGGGAGCCGAGCCCCGGAAGAACTTCGTCAAGCCTAACGAAACTAAGA CCTACTTTTGGAAGGTCCAGCATCACATGGCCCCGACCAAAGACGAGTT CGACTGTAAAGCCTGGGCCTACTTCTCCGATGTGGACCTGGAGAAGGAC GTGCACTCGGGACTCATTGGCCCGCTCCTTGTGTGCCATACTAATACCCT GAACCCTGCTCACGGTCGCCAAGTCACAGTGCAGGAGTTCGCCCTCTTC TTCACCATCTTCGATGAAACAAAGTCCTGGTACTTTACTGAGAACATGG AACGCAATTGCAGGGCACCCTGCAACATCCAGATGGAAGATCCCACCTT CAAGGAAAACTACCGGTTTCATGCCATTAACGGCTACATAATGGACACG TTGCCAGGACTGGTCATGGCCCAGGACCAGAGAATCCGGTGGTATCTGC TCTCCATGGGCTCCAACGAAAACATTCACAGCATTCATTTTTCCGGCCAT GTGTTCACCGTCCGGAAGAAGGAAGAGTACAAGATGGCTCTGTACAAC CTCTACCCTGGAGTGTTCGAGACTGTGGAAATGCTGCCTAGCAAGGCCG GCATTTGGAGAGTGGAATGCCTGATCGGAGAGCATTTGCACGCCGGAAT GTCCACCCTGTTTCTTGTGTACTCCAACAAGTGCCAGACCCCGCTGGGA ATGGCCTCAGGTCATATTAGGGATTTCCAGATCACTGCTTCGGGGCAGT ACGGGCAGTGGGCACCTAAGTTGGCCCGGCTGCACTACTCTGGCTCCAT CAATGCCTGGTCCACCAAGGAACCCTTCTCCTGGATTAAGGTGGACCTC CTGGCCCCAATGATTATTCACGGTATTAAGACCCAGGGTGCCCGACAGA AGTTCTCCTCACTCTACATCTCGCAATTCATCATAATGTACAGCCTGGAT GGGAAGAAGTGGCAGACCTACCGGGGAAACTCCACTGGAACGCTCATG GTGTTTTTCGGCAACGTGGACTCCTCCGGCATTAAGCACAACATCTTCA ACCCTCCGATCATTGCTCGGTACATCCGGCTGCACCCAACTCACTACAG CATCCGGTCCACCCTGCGGATGGAACTGATGGGTTGTGACCTGAACTCC TGCTCCATGCCCCTTGGGATGGAATCCAAGGCCATTAGCGATGCACAGA TCACCGCCTCTTCATACTTCACCAACATGTTCGCGACCTGGTCCCCGTCG AAGGCCCGCCTGCACCTCCAAGGTCGCTCCAATGCGTGGCGGCCTCAAG

	TGAACAACCCCAAGGAGTGGCTCCAGGTCGACTTCCAAAAGACCATGA AGGTCACCGGAGTGACCACCCAGGGCGTGAAGTCCCTGCTGACCTCTAT GTACGTTAAGGAGTTCCTCATCTCCTCAAGCCAAGACGGACATCAGTGG ACCCTGTTCTTCCAAAACGGAAAAGTCAAAGTATTCCAGGGCAACCAGG ACTCCTTCACCCCTGTGGTCAACAGCCTGGACCCCCCATTGCTGACCCG CTACCTCCGCATCCACCCCCAAAGCTGGGTCCACCAGATCGCACTGCGC ATGGAGGTCCTTGGATGCGAAGCCCAAGATCTGTACTAA
SEQ ID NO: 10 coBDDFVIIIX TEN (V2.0)的 氨基酸序列	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTL FVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAV GVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMAS DPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAV FDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHR KSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMD LGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTD SEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAP DDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLL YGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPG EIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKES VDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQ ASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHK MVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSC DKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNGAPTSESATPESGPGSEP ATSGSETPGTSESATPESGPGSEPATSGSETPGTSESATPESGPGTSTEPSEGS APGSPAGSPTSTEEGTSESATPESGPGSEPATSGSETPGTSESATPESGPGSPA GSPTSTEEGSPAGSPTSTEEGASSPPVLKRHQREITRTTLQSDQEEIDYDDTIS VEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLR NRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNI MVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQH HMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNTLNPAHGRQV TVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAING YIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKM ALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTP LGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDL LAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFF GNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWL QVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV KVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDL Y
SEQ ID NO: 11 coBDDFVIIIX TEN (V2.0)的 信号肽	MQIELSTCFFLCLLRFCFS
SEQ ID NO: 12 BDD成熟人 FVIII的氨基 酸序列	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTL FVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAV GVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMAS DPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAV FDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHR KSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMD LGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTD SEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAP DDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLL YGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPG

	EIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKES VDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQ ASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHK MVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSC DKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNPPVLKRHQREITRTTLQS DQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWD YGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGSFTQPLYRGELNEHLGLL GPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNE TKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVHSGLIGPLLVCHT NTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPT FKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHV FTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMST LFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINAWS TKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYR GNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGC DLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAW RPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGH QWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALR MEVLGCEAQDLY
SEQ ID NO: 13 编码BDD成熟人FVIII的核苷酸序列	ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTG CTTTAGTGCCACCAGAAGATACTACCTGGGTGCAGTGGAACTGTCATGG GACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTTC CTCCTAGAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAA AAAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAACATCGCTAAG CCAAGGCCACCCTGGATGGGTCTGCTAGGTCCTACCATCCAGGCTGAGG TTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCCATCCTGTC AGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTG AATATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGATAAAGTCT TCCCTGGTGGAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAATGG TCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCATG TGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGT ATGTAGAGAAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAA ATTTATACTACTTTTTGCTGTATTTGATGAAGGGAAAAGTTGGCACTCAG AAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTCGGGC CTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCA GGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAA TGGGCACCACTCCTGAAGTGCACTCAATATTCCTCGAAGGTCACACATT TCTTGTGAGGAACCATCGCCAGGCGTCCTTGGAAATCTCGCCAATAACT TTCCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTT TTGTCATATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAA GTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATAATGAA GAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGG TCAGGTTTGATGATGACAACTCTCCTTCCTTTATCCAAATTCGCTCAGTT GCCAAGAAGCATCCTAAAACTTGGGTACATTACATTGCTGCTGAAGAGG AGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCCGATGACAGAAGTTA TAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTAC AAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTCGTG AAGCTATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGA AGTTGGAGACACACTGTTGATTATATTTAAGAATCAAGCAAGCAGACCA TATAACATCTACCCTCACGGAATCACTGATGTCCGTCCTTTGTATTCAAG GAGATTACCAAAAGGTGTAAAACATTTGAAGGATTTTCCAATTCTGCCA GGAGAAATATTCAAATATAAATGGACAGTGACTGTAGAAGATGGGCCA ACTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAA TATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATCTGCT ACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCAGACAAGA GGAATGTCATCCTGTTTTCTGTATTTGATGAGAACCGAAGCTGGTACCTC ACAGAGAATATACAACGCTTTCTCCCCAATCCAGCTGGAGTGCAGCTTG AGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTA

	TGTTTTTGATAGTTTGCAGTTGTCAGTTTGTTTGCATGAGGTGGCATACT GGTACATTCTAAGCATTGGAGCACAGACTGACTTCCTTTCTGTCTTCTTC TCTGGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTCACCC TATTCCCATTCTCAGGAGAAACTGTCTTCATGTCGATGGAAAACCCAGG TCTATGGATTCTGGGGTGCCACAACTCAGACTTTCGGAACAGAGGCATG ACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATT ACGAGGACAGTTATGAAGATATTTCAGCATACTTGCTGAGTAAAAACAA TGCCATTGAACCAAGAAGCTTCTCTCAAAACCCACCAGTCTTGAAACGC CATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAA TTGACTATGATGATACCATATCAGTTGAAATGAAGAAGGAAGATTTTGA CATTTATGATGAGGATGAAAATCAGAGCCCCCGCAGCTTTCAAAAGAA AACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATGGG ATGAGTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTG TCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCCTTT ACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGG GGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTTCAG AAATCAGGCCTCTCGTCCCTATTCCTTCTATTCTAGCCTTATTTCTTATGA GGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCC TAATGAAACCAAAACTTACTTTTGGAAAGTGCAACATCATATGGCACCC ACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTG ACCTGGAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTG CCACACTAACACACTGAACCCTGCTCATGGGAGACAAGTGACAGTACA GGAATTTGCTCTGTTTTTCACCATCTTTGATGAGACCAAAAGCTGGTACT TCACTGAAAATATGGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGAT GGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGC TACATAATGGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAGGA TTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTCTAT TCATTTCAGTGGACATGTGTTCACTGTACGAAAAAAAGAGGAGTATAAA ATGGCACTGTACAATCTCTATCCAGGTGTTTTTGAGACAGTGGAAATGT TACCATCCAAAGCTGGAATTTGGCGGGTGGAATGCCTTATTGGCGAGCA TCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGT CAGACTCCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTA CAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGACTTC ATTATTCCGGATCAATCAATGCCTGGAGCACCAAGGAGCCCTTTTCTTG GATCAAGGTGGATCTGTTGGCACCAATGATTATTCACGGCATCAAGACC CAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCAT CATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCC ACTGGAACCTTAATGGTCTTCTTTGGCAATGTGGATTCATCTGGGATAA AACACAATATTTTTAACCCTCCAATTATTGCTCGATACATCCGTTTGCAC CCAACTCATTATAGCATTCGCAGCACTCTTCGCATGGAGTTGATGGGCT GTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCAAT ATCAGATGCACAGATTACTGCTTCATCCTACTTTACCAATATGTTTGCCA CCTGGTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGC CTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTT CCAGAAGACAATGAAAGTCACAGGAGTAACTACTCAGGGAGTAAAATC TCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGCAGTCAA GATGGCCATCAGTGGACTCTCTTTTTTCAGAATGGCAAAGTAAAGGTTT TTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAACTCTCTAGACCC ACCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTGGGTGCAC CAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTCT AC
SEQ ID NO: 14 V2.0 表达盒 mTTR482-内含子-	GGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTTG GCAGCATTTACTCTCTCTATTGACTTTGGTTAATAATCTCAGGAGCACAA ACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCTG GGCCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTC AAAGGTCAAAGTGGCCCTTGGCAGCATTTACTCTCTCTATTGACTTTGGT TAATAATCTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAGAAATCA ACATCCTGGACTTATCCTCTGGGCCTCTCCCCACCGATATCTACCTGCTG

coBDDFVIIIXTEN (V2.0)-WPRE-bGHPolyA	ATCGCCCGGCCCCTGTTCAAACATGTCCTAATACTCTGTCGGGGCAAAG GTCGGCAGTAGTTTTCCATCTTACTCAACATCCTCCCAGTGTACGTAGGA TCCTGTCTGTCTGCACATTTCGTAGAGCGAGTGTTCCGATACTCTAATCT CCCGGGGCAAAGGTCGTATTGACTTAGGTTACTTATTCTCCTTTTGTTGA CTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGAT CAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAAGCCCCTTCACCAGGA GAAGCCGTCACACAGATCCACAAGCTCCTGCTAGGAATTCTCAGGAGCA CAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCT CTGGGCCTCTCCCCACCGATATCTACCTGCTGATCGCCCGGCCCCTGTTC AAACATGTCCTAATACTCTGTCGGGGCAAAGGTCGGCAGTAGTTTTCCA TCTTACTCAACATCCTCCCAGTGTACGTAGGATCCTGTCTGTCTGCACAT TTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGGCAAAGGTCGT ATTGACTTAGGTTACTTATTCTCCTTTTGTTGACTAAGTCAATAATCAGA ATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGA AGGAGGGGGTATAAAAGCCCCTTCACCAGGAGAAGCCGTCACACAGAT CCACAAGCTCCTGCTAGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCCCC GCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTT ACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAAT TAGCGCTTGGTTTATTGACGGCTTGTTTCTTTTCTGTGGCTGCGTGAAAG CCTTGAGGGGCTCCGGGAAGGCCCTTTGTGCGGGGGGAGCGGCTCGGG GGGTGCGTGCGTGTGTGTGTGCGTGGGGAGCGCCGCGTGCGGCTCCGCG CTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCGGGGCTTTGTGCGC TCCGCAGTGTGCGCGAGGGGAGCGCGGCCGGGGGCGGTGCCCCGCGGT GCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCG TGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTCGGGCTGCAACCCC CCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCCCGGCTTCGGGTGC GGGGCTCCGTACGGGGCGTGGCGCGGGGCTCGCCGTGCCGGGCGGGGG GTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCGCCTCGGGCCGG GGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGGAGCGCCGGCGGCTG TCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATGGTAATCGTGCGAG AGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCGGAGCCGAAATCTG GGAGGCGCCGCCGCACCCCCTCTAGCGGGCGCGGGGCGAAGCGGTGCG GCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGC GCCGCCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGA CGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCG TGTGACCGGCGGCTCTAGAGCCTCTGCTAACCTTGTTCTTGCCTTCTTCT TTTTCCTACAGCTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCAT TTTGGCAAAGAATTACTCGAGGCCACCATGCAGATTGAACTGTCCACTT GCTTCTTCCTGTGCCTCCTGCGGTTTTGCTTCTCGGCCACCCGCCGGTAT TACTTAGGTGCTGTGGAACTGAGCTGGGACTACATGCAGTCCGACCTGG GAGAACTGCCGGTGGACGCGAGATTCCCACCTAGAGTCCCGAAGTCCTT CCCATTCAACACCTCCGTGGTCTACAAAAAGACCCTGTTCGTGGAGTTC ACTGACCACCTTTTCAATATTGCCAAGCCGCGCCCCCCCTGGATGGGCC TGCTTGGTCCTACGATCCAAGCAGAGGTCTACGACACCGTGGTCATCAC ACTGAAGAACATGGCCTCACACCCCGTGTCGCTGCATGCTGTGGGAGTG TCCTACTGGAAGGCCTCAGAGGGTGCCGAATATGATGACCAGACCAGC CAGAGGGAAAAGGAGGATGACAAAGTGTTCCCGGGTGGCAGCCACACT TACGTGTGGCAAGTGCTGAAGGAAAACGGGCCTATGGCGTCGGACCCC CTATGCCTGACCTACTCCTACCTGTCCCATGTGGACCTTGTGAAGGATCT CAACTCGGGACTGATCGGCGCCCTCTTGGTGTGCAGAGAAGGCAGCCTG GCGAAGGAAAAGACTCAGACCCTGCACAAGTTCATTCTGTTGTTTGCTG TGTTCGATGAAGGAAAGTCCTGGCACTCAGAAACCAAGAACTCGCTGAT GCAGGATAGAGATGCGGCCTCGGCCAGAGCCTGGCCTAAAATGCACAC CGTCAACGGATATGTGAACAGGTCGCTCCCTGGCCTCATCGGCTGCCAC AGAAAGTCCGTGTATTGGCATGTGATCGGCATGGGTACTACTCCGGAAG TGCATAGTATCTTTCTGGAGGGCCATACCTTCTTGGTGCGCAACCACAG ACAGGCCTCGCTGGAAATCTCGCCTATCACTTTCTTGACTGCGCAGACC CTCCTTATGGACCTTGGACAGTTCCTGCTGTTCTGTCACATCAGCTCCCA TCAGCATGATGGGATGGAGGCCTATGTCAAAGTGGACTCCTGCCCTGAG

	GAGCCACAGCTCCGGATGAAGAACAATGAGGAAGCGGAGGATTACGAC GACGACCTGACTGACAGCGAAATGGACGTCGTGCGATTCGATGACGAC AACAGCCCGTCCTTCATCCAAATTAGATCAGTGGCGAAGAAGCACCCCA AGACCTGGGTGCACTACATTGCCGCCGAGGAAGAGGACTGGGACTACG CGCCGCTGGTGCTGGCGCCAGACGACAGGAGCTACAAGTCCCAGTACCT CAACAACGGGCCGCAGCGCATTGGCAGGAAGTACAAGAAAGTCCGCTT CATGGCCTACACTGATGAAACCTTCAAGACGAGGGAAGCCATCCAGCA CGAGTCAGGCATCCTGGGACCGCTCCTTTACGGCGAAGTCGGGGATACC CTGCTCATCATTTTCAAGAACCAGGCATCGCGGCCCTACAACATCTACC CTCACGGGATCACAGACGTGCGCCCGCTCTACTCCCGCCGGCTGCCCAA GGGAGTGAAGCACCTGAAGGATTTTCCCATCCTGCCGGGAGAAATCTTC AAGTACAAGTGGACCGTGACTGTGGAAGATGGCCCTACCAAGTCGGAC CCTCGCTGTCTGACCCGGTACTATTCCTCGTTTGTGAACATGGAGCGCG ACCTGGCCTCGGGGCTGATTGGTCCGCTGCTGATCTGCTACAAGGAGTC CGTGGACCAGCGCGGGAACCAGATCATGTCCGACAAGCGCAACGTGAT CCTGTTCTCTGTCTTTGATGAAAACAGATCGTGGTACTTGACTGAGAAT ATCCAGCGGTTCCTGCCCAACCCAGCGGGAGTGCAACTGGAGGACCCG GAGTTCCAGGCCTCAAACATTATGCACTCTATCAACGGCTATGTGTTCG ACTCGCTCCAACTGAGCGTGTGCCTGCATGAAGTGGCATACTGGTACAT TCTGTCCATCGGAGCCCAGACCGACTTCCTGTCCGTGTTCTTCTCCGGAT ACACCTTCAAGCATAAGATGGTGTACGAGGACACTCTGACCCTCTTCCC ATTTTCCGGAGAAACTGTGTTCATGTCAATGGAAAACCCGGGCTTGTGG ATTCTGGGTTGCCATAACTCGGACTTCCGGAATAGAGGGATGACCGCCC TGCTGAAAGTGTCCAGCTGTGACAAGAATACCGGCGATTACTACGAGG ACAGCTATGAGGACATCTCCGCTTATCTGCTGTCCAAGAACAACGCCAT TGAACCCAGGTCCTTCTCCCAAAACGGTGCACCGACCTCCGAAAGCGCC ACCCCAGAGTCAGGACCTGGCTCGGAACCGGCTACCTCGGGCTCAGAG ACACCGGGGACTTCCGAGTCCGCAACCCCCGAGAGTGGACCCGGATCC GAACCAGCAACCTCAGGATCAGAAACCCCGGGAACTTCGGAATCCGCC ACTCCCGAGTCGGGACCAGGCACCTCCACTGAGCCTTCCGAGGGAAGC GCCCCCGGATCCCCTGCTGGATCCCCTACCAGCACTGAAGAAGGCACCT CAGAATCCGCGACCCCTGAGTCCGGCCCTGGAAGCGAACCCGCCACCTC CGGTTCCGAAACCCCTGGGACTAGCGAGAGCGCCACTCCGGAATCGGG CCCAGGAAGCCCTGCCGGATCCCCGACCAGCACCGAGGAGGGAAGCCC CGCCGGGTCACCGACTTCCACTGAGGAGGGAGCCTCATCCCCCCCCGTG CTGAAGCGGCATCAAAGAGAGATCACCAGGACCACTCTCCAGTCCGAT CAGGAAGAAATTGACTACGACGATACTATCAGCGTGGAGATGAAGAAG GAGGACTTCGACATCTACGATGAGGATGAGAACCAGTCCCCTCGGAGCT TTCAGAAGAAAACCCGCCACTACTTCATCGCTGCCGTGGAGCGGCTGTG GGATTACGGGATGTCCAGCTCACCGCATGTGCTGCGGAATAGAGCGCA GTCAGGATCGGTGCCCCAGTTCAAGAAGGTCGTGTTCCAAGAGTTCACC GACGGGTCCTTCACTCAACCCCTGTACCGGGGCGAACTCAACGAACACC TGGGACTGCTTGGGCCGTATATCAGGGCAGAAGTGGAAGATAACATCA TGGTCACCTTCCGCAACCAGGCCTCCCGGCCGTACAGCTTCTACTCTTCA CTGATCTCCTACGAGGAAGATCAGCGGCAGGGAGCCGAGCCCCGGAAG AACTTCGTCAAGCCTAACGAAACTAAGACCTACTTTTGGAAGGTCCAGC ATCACATGGCCCCGACCAAAGACGAGTTCGACTGTAAAGCCTGGGCCTA CTTCTCCGATGTGGACCTGGAGAAGGACGTGCACTCGGGACTCATTGGC CCGCTCCTTGTGTGCCATACTAATACCCTGAACCCTGCTCACGGTCGCCA AGTCACAGTGCAGGAGTTCGCCCTCTTCTTCACCATCTTCGATGAAACA AAGTCCTGGTACTTTACTGAGAACATGGAACGCAATTGCAGGGCACCCT GCAACATCCAGATGGAAGATCCCACCTTCAAGGAAAACTACCGGTTTCA TGCCATTAACGGCTACATAATGGACACGTTGCCAGGACTGGTCATGGCC CAGGACCAGAGAATCCGGTGGTATCTGCTCTCCATGGGCTCCAACGAAA ACATTCACAGCATTCATTTTTCCGGCCATGTGTTCACCGTCCGGAAGAA GGAAGAGTACAAGATGGCTCTGTACAACCTCTACCCTGGAGTGTTCGAG ACTGTGGAAATGCTGCCTAGCAAGGCCGGCATTTGGAGAGTGGAATGC CTGATCGGAGAGCATTTGCACGCCGGAATGTCCACCCTGTTTCTTGTGT ACTCCAACAAGTGCCAGACCCCGCTGGGAATGGCCTCAGGTCATATTAG

	GGATTTCCAGATCACTGCTTCGGGGCAGTACGGGCAGTGGGCACCTAAG TTGGCCCGGCTGCACTACTCTGGCTCCATCAATGCCTGGTCCACCAAGG AACCCTTCTCCTGGATTAAGGTGGACCTCCTGGCCCCAATGATTATTCAC GGTATTAAGACCCAGGGTGCCCGACAGAAGTTCTCCTCACTCTACATCT CGCAATTCATCATAATGTACAGCCTGGATGGGAAGAAGTGGCAGACCT ACCGGGGAAACTCCACTGGAACGCTCATGGTGTTTTTCGGCAACGTGGA CTCCTCCGGCATTAAGCACAACATCTTCAACCCTCCGATCATTGCTCGGT ACATCCGGCTGCACCCAACTCACTACAGCATCCGGTCCACCCTGCGGAT GGAACTGATGGGTTGTGACCTGAACTCCTGCTCCATGCCCCTTGGGATG GAATCCAAGGCCATTAGCGATGCACAGATCACCGCCTCTTCATACTTCA CCAACATGTTCGCGACCTGGTCCCCGTCGAAGGCCCGCCTGCACCTCCA AGGTCGCTCCAATGCGTGGCGGCCTCAAGTGAACAACCCCAAGGAGTG GCTCCAGGTCGACTTCCAAAAGACCATGAAGGTCACCGGAGTGACCAC CCAGGGCGTGAAGTCCCTGCTGACCTCTATGTACGTTAAGGAGTTCCTC ATCTCCTCAAGCCAAGACGGACATCAGTGGACCCTGTTCTTCCAAAACG GAAAAGTCAAAGTATTCCAGGGCAACCAGGACTCCTTCACCCCTGTGGT CAACAGCCTGGACCCCCCATTGCTGACCCGCTACCTCCGCATCCACCCC CAAAGCTGGGTCCACCAGATCGCACTGCGCATGGAGGTCCTTGGATGCG AAGCCCAAGATCTGTACTAAGCGGCCGCTCATAATCAACCTCTGGATTA CAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTA CGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCC CGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTT TATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTG TGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCA GCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAAC TCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGG CACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGC TGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTAC GTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCC GGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGA TCTCCCTTTGGGCCGCCTCCCCGCTGCCTAGGCGACTGTGCCTTCTAGTT GCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAA GGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGC ATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGA CAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAAG ACCATGGGCGCGCCAGGCCTGTCGACGCCCGGGCGGTACCGCGATCGCT CGCGACGCATAAAG
SEQ ID NO: 15 A1MB2增强子	GGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAGGTCAAAGTGGCCCTTG GCAGCATTTACTCTCTCTATTGACTTTGGTTAATAATCTCAGGAGCACAA ACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTATCCTCTG GGCCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTC AAAGGTCAAAGTGGCCCTTGGCAGCATTTACTCTCTCTATTGACTTTGGT TAATAATCTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAGAAATCA ACATCCTGGACTTATCCTCTGGGCCTCTCCCCACC
SEQ ID NO: 16 mTTR启动子	GATATCTACCTGCTGATCGCCCGGCCCCTGTTCAAACATGTCCTAATACT CTGTCGGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCAACATCCTC CCAGTGTACGTAGGATCCTGTCTGTCTGCACATTTCGTAGAGCGAGTGT TCCGATACTCTAATCTCCCGGGGCAAAGGTCGTATTGACTTAGGTTACTT ATTCTCCTTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGT CAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAA GCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGCTAG
SEQ ID NO: 17 嵌合内含子	TCAGGAGCACAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTG GACTTATCCTCTGGGCCTCTCCCCACCGATATCTACCTGCTGATCGCCCG GCCCCTGTTCAAACATGTCCTAATACTCTGTCGGGGCAAAGGTCGGCAG TAGTTTTCCATCTTACTCAACATCCTCCCAGTGTACGTAGGATCCTGTCT GTCTGCACATTTCGTAGAGCGAGTGTTCCGATACTCTAATCTCCCGGGG

	CAAAGGTCGTATTGACTTAGGTTACTTATTCTCCTTTTGTTGACTAAGTC AATAATCAGAATCAGCAGGTTTGGAGTCAGCTTGGCAGGGATCAGCAG CCTGGGTTGGAAGGAGGGGGTATAAAAGCCCCTTCACCAGGAGAAGCC GTCACACAGATCCACAAGCTCCTGCTAGAGTCGCTGCGCGCTGCCTTCG CCCCGTGCCCCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGA CTGACCGCGTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTC CGGGCTGTAATTAGCGCTTGGTTTATTGACGGCTTGTTTCTTTTCTGTGG CTGCGTGAAAGCCTTGAGGGGCTCCGGGAAGGCCCTTTGTGCGGGGGG AGCGGCTCGGGGGGTGCGTGCGTGTGTGTGTGCGTGGGGAGCGCCGCG TGCGGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGCGCG GGGCTTTGTGCGCTCCGCAGTGTGCGCGAGGGGAGCGCGGCCGGGGGC GGTGCCCCGCGGTGCGGGGGGGGCTGCGAGGGGAACAAAGGCTGCGTG CGGGGTGTGTGCGTGGGGGGGTGAGCAGGGGGTGTGGGCGCGTCGGTC GGGCTGCAACCCCCCCCTGCACCCCCCTCCCCGAGTTGCTGAGCACGGCC CGGCTTCGGGTGCGGGGCTCCGTACGGGGCGTGGCGCGGGGCTCGCCGT GCCGGGCGGGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGC CGCCTCGGGCCGGGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCCCGGA GCGCCGGCGGCTGTCGAGGCGCGGCGAGCCGCAGCCATTGCCTTTTATG GTAATCGTGCGAGAGGGCGCAGGGACTTCCTTTGTCCCAAATCTGTGCG GAGCCGAAATCTGGGAGGCGCCGCCGCACCCCCTCTAGCGGGCGCGGG GCGAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTT CGTGCGTCGCCGCGCCGCCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTG TCCGCGGGGGGACGGCTGCCTTCGGGGGGGACGGGGCAGGGCGGGGTT CGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTAACCTTGTTC TTGCCTTCTTCTTTTTCCTACAGCTCCTGGGCAACGTGCTGGTTATTGTG CTGTCTCATCATTTTGGCAAAGAATTA
SEQ ID NO: 18 WPRE	TCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATT CTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCC TTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTA TAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGC AACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTG GGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCC TCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTG GACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGG AAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCT GCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGAC CTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCG CCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCTG
SEQ ID NO: 19 bGHpA	CGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTG CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAA TGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGG GGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAG CAGGCATGCTGGGGA
SEQ ID NO: 20 野生型人成熟 FVIII蛋白的 氨基酸序列	ATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFNTSVVYKKTL FVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAV GVSYWKASEGAEYDDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMAS DPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAV FDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHR KSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMD LGQFLLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTD SEMDVVRFDDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDYAPLVLAP DDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLL YGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPG EIFKYKWTVTVEDGPTKSDPRCLTRYYSSFVNMERDLASGLIGPLLICYKES

	VDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQ ASNIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYTFKHK MVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALLKVSSC DKNTGDYYEDSYEDISAYLLSKNNAIEPRSFSQNSRHPSTRQKQFNATTIPE NDIEKTDPWFAHRTPMPKIQNVSSSDLLMLLRQSPTPHGLSLSDLQEAKYET FSDDPSPGAIDSNNSLSEMTHFRPQLHHSGDMVFTPESGLQLRLNEKLGTTA ATELKKLDFKVSSTSNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDT TLFGKKSSPLTESGGPLSLSEENNDSKLLESGLMNSQESSWGKNVSSTESGR LFKGKRAHGPALLTKDNALFKVSISLLKTNKTSNNSATNRKTHIDGPSLLIE NSPSVWQNILESDTEFKKVTPLIHDRMLMDKNATALRLNHMSNKTTSSKN MEMVQQKKEGPIPPDAQNPDMSFFKMLFLPESARWIQRTHGKNSLNSGQG PSPKQLVSLGPEKSVEGQNFLSEKNKVVVGKGEFTKDVGLKEMVFPSSRNL FLTNLDNLHENNTHNQEKKIQEEIEKKETLIQENVVLPQIHTVTGTKNFMKN LFLLSTRQNVEGSYDGAYAPVLQDFRSLNDSTNRTKKHTAHFSKKGEEENL EGLGNQTKQIVEKYACTTRISPNTSQQNFVTQRSKRALKQFRLPLEETELEK RIIVDDTSTQWSKNMKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHSIPQ ANRSPLPIAKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFL QGAKKNNLSLAILTLEMTGDQREVGSLGTSATNSVTYKKVENTVLPKPDLP KTSGKVELLPKVHIYQKDLFPTETSNGSPGHLDLVEGSLLQGTEGAIKWNE ANRPGKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPE KTAFKKKDTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNP PVLKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQ KKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKKVVFQEFTDGS FTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEED QRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLE KDVHSGLIGPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENM ERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLS MGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWR VECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAP KLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQF IIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPT HYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSP SKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTS MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRY LRIHPQSWVHQIALRMEVLGCEAQDLY
SEQ ID NO: 21 B19 WT 5’	CCAAATCAGATGCCGCCGGTCGCCGCCGGTAGGCGGGACTTCCGGTACA AGATGGCGGACAATTACGTCATTTCCTGTGACGTCATTTCCTGTGACGTC ACTTCCGGTGGGCGGGACTTCCGGAATTAGGGTTGGCTCTGGGCCAGCT TGCTTGGGGTTGCCTTGACACTAAGACAAGCGGCGCGCCGCTTGATCTT AGTGGCACGTCAACCCCAAGCGCTGGCCCAGAGCCAACCCTAATTCCGG AAGTCCCGCCCACCGGAAGTGACGTCACAGGAAATGACGTCACAGGAA ATGACGTAATTGTCCGCCATCTTGTACCGGAAGTCCCGCCTACCGGCGG CGACCGGCGGCATCTGATTTGGTGTCTTCTTTTAAATTTT
SEQ ID NO: 22 B19 WT 3’	AAAATTTAAAAGAAGACACCAAATCAGATGCCGCCGGTCGCCGCCGGT AGGCGGGACTTCCGGTACAAGATGGCGGACAATTACGTCATTTCCTGTG ACGTCATTTCCTGTGACGTCACTTCCGGTGGGCGGGACTTCCGGAATTA GGGTTGGCTCTGGGCCAGCGCTTGGGGTTGACGTGCCACTAAGATCAAG CGGCGCGCCGCTTGTCTTAGTGTCAAGGCAACCCCAAGCAAGCTGGCCC AGAGCCAACCCTAATTCCGGAAGTCCCGCCCACCGGAAGTGACGTCAC AGGAAATGACGTCACAGGAAATGACGTAATTGTCCGCCATCTTGTACCG GAAGTCCCGCCTACCGGCGGCGACCGGCGGCATCTGATTTGG
SEQ ID NO: 23 5’_B19_最小	CCAAATCAGATGCCGCCGGTCGCCGCCGGTAGGCGGGACTTCCGGTACA GCGCGCCGCTGTACCGGAAGTCCCGCCTACCGGCGGCGACCGGCGGCA TCTGATTTGGTGTCTTCTTTTAAATTTT

SEQ ID NO: 24 3'_B19_最小	AAAATTTAAAAGAAGACACCAAATCAGATGCCGCCGGTCGCCGCCGGT AGGCGGGACTTCCGGTACAGCGGCGCGCTGTACCGGAAGTCCCGCCTAC CGGCGGCGACCGGCGGCATCTGATTTGG
SEQ ID NO: 25 5'_GPV_最小	CTCATTGGAGGGTTCGTTCGTTCGAACGTTCGTTCGCATGCGAACGAAC GTTCGAACGAACGAACCCTCCAATGAGACTCAAGGACAAGAGGATATT TTGCGCGCCAGGAAGTG
SEQ ID NO: 26 3'_GPV_最小	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGG TTCGTTCGTTCGAACGTTCGTTCGCATGCGAACGAACGTTCGAACGAAC GAACCCTCCAATGAG
SEQ ID NO: 27 5'_GPV_Δ186	CTCATTGGAGGGTTCGTTCGTTCGAACCAGCCAATCAGGGGAGGGGGA AGTGACGCAAGTTCCGGTCACATGCTTCCGGTGACGCACATCCGGTGAC GTAGTTCGCATGCCTGTCTATCGCCTACCCATCCCTGTCTGAGATCAAGG GCGTGATCGTGCACAGACTGGAGAGCGTGTCCTATAATATCGGCTCTCA GGAGTGGAGCACCACAGTGCCCAGATACGTGGCCACCCAGGGCTATCT GATCTCCAACTTCGACGCATGCGAACTACGTCACCGGATGTGCGTCACC GGAAGCATGTGACCGGAACTTGCGTCACTTCCCCCTCCCCTGATTGGCT GGTTCGAACGAACGAACCCTCCAATGAGACTCAAGGACAAGAGGATAT TTTGCGCGCCAGGAAGTG
SEQ ID NO: 28 3'_GPV_Δ186	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGG TTCGTTCGTTCGAACCAGCCAATCAGGGGAGGGGGAAGTGACGCAAGT TCCGGTCACATGCTTCCGGTGACGCACATCCGGTGACGTAGTTCGCATG CCTGTCTATCGCCTACCCATCCCTGTCTGAGATCAAGGGCGTGATCGTG CACAGACTGGAGAGCGTGTCCTATAATATCGGCTCTCAGGAGTGGAGCA CCACAGTGCCCAGATACGTGGCCACCCAGGGCTATCTGATCTCCAACTT CGACGCATGCGAACTACGTCACCGGATGTGCGTCACCGGAAGCATGTG ACCGGAACTTGCGTCACTTCCCCCTCCCCTGATTGGCTGGTTCGAACGA ACGAACCCTCCAATGAG
SEQ ID NO: 29 5'_GPV_Δ120	CTCATTGGAGGGTTCGTTCGTTCGAACCAGCCAATCAGGGGAGGGGGA AGTGACGCAAGTTCCGGTCACATGCTTCCGGTGACGCACATCCGGTGAC GTAGTTCCGGTCACGTGCTTCCTGTCACGTGTTTCCGGTCGCATGCCTGT CTATCGCCTACCCATCCCTGTCTGAGATCAAGGGCGTGATCGTGCACAG ACTGGAGAGCGTGTCCTATAATATCGGCTCTCAGGAGTGGAGCACCACA GTGCCCAGATACGTGGCCACCCAGGGCTATCTGATCTCCAACTTCGACG CATGCTCACGTGACCGGAAACACGTGACAGGAAGCACGTGACCGGAAC TACGTCACCGGATGTGCGTCACCGGAAGCATGTGACCGGAACTTGCGTC ACTTCCCCCTCCCCTGATTGGCTGGTTCGAACGAACGAACCCTCCAATG AGACTCAAGGACAAGAGGATATTTTGCGCGCCAGGAAGTG
SEQ ID NO: 30 3'_GPV_Δ120	CACTTCCTGGCGCGCAAAATATCCTCTTGTCCTTGAGTCTCATTGGAGGG TTCGTTCGTTCGAACCAGCCAATCAGGGGAGGGGGAAGTGACGCAAGT TCCGGTCACATGCTTCCGGTGACGCACATCCGGTGACGTAGTTCCGGTC ACGTGCTTCCTGTCACGTGTTTCCGGTCACGTGAGCATGCCTGTCTATCG CCTACCCATCCCTGTCTGAGATCAAGGGCGTGATCGTGCACAGACTGGA GAGCGTGTCCTATAATATCGGCTCTCAGGAGTGGAGCACCACAGTGCCC AGATACGTGGCCACCCAGGGCTATCTGATCTCCAACTTCGACGGCATGC GACCGGAAACACGTGACAGGAAGCACGTGACCGGAACTACGTCACCGG

	ATGTGCGTCACCGGAAGCATGTGACCGGAACTTGCGTCACTTCCCCCTC CCCTGATTGGCTGGTTCGAACGAACGAACCCTCCAATGAG
SEQ ID NO: 31 野生型人FVIII的核酸序列	ATGCAAATAGAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCGATTCTG CTTTAGTGCCACCAGAAGATACTACCTGGGTGCAGTGGAACTGTCATGG GACTATATGCAAAGTGATCTCGGTGAGCTGCCTGTGGACGCAAGATTTC CTCCTAGAGTGCCAAAATCTTTTCCATTCAACACCTCAGTCGTGTACAA AAAGACTCTGTTTGTAGAATTCACGGATCACCTTTTCAACATCGCTAAG CCAAGGCCACCCTGGATGGGTCTGCTAGGTCCTACCATCCAGGCTGAGG TTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCCATCCTGTC AGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTG AATATGATGATCAGACCAGTCAAAGGGAGAAAGAAGATGATAAAGTCT TCCCTGGTGGAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAATGG TCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCATG TGGACCTGGTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGT ATGTAGAGAAGGGAGTCTGGCCAAGGAAAAGACACAGACCTTGCACAA ATTTATACTACTTTTTGCTGTATTTGATGAAGGGAAAAGTTGGCACTCAG AAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCTGCTCGGGC CTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCA GGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAA TGGGCACCACTCCTGAAGTGCACTCAATATTCCTCGAAGGTCACACATT TCTTGTGAGGAACCATCGCCAGGCGTCCTTGGAAATCTCGCCAATAACT TTCCTTACTGCTCAAACACTCTTGATGGACCTTGGACAGTTTCTACTGTT TTGTCATATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAA GTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATAATGAA GAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGG TCAGGTTTGATGATGACAACTCTCCTTCCTTTATCCAAATTCGCTCAGTT GCCAAGAAGCATCCTAAAACTTGGGTACATTACATTGCTGCTGAAGAGG AGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCCGATGACAGAAGTTA TAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTAC AAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTCGTG AAGCTATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGA AGTTGGAGACACACTGTTGATTATATTTAAGAATCAAGCAAGCAGACCA TATAACATCTACCCTCACGGAATCACTGATGTCCGTCCTTTGTATTCAAG GAGATTACCAAAAGGTGTAAAACATTTGAAGGATTTTCCAATTCTGCCA GGAGAAATATTCAAATATAAATGGACAGTGACTGTAGAAGATGGGCCA ACTAAATCAGATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAA TATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCCTCTCCTCATCTGCT ACAAAGAATCTGTAGATCAAAGAGGAAACCAGATAATGTCAGACAAGA GGAATGTCATCCTGTTTTCTGTATTTGATGAGAACCGAAGCTGGTACCTC ACAGAGAATATACAACGCTTTCTCCCCAATCCAGCTGGAGTGCAGCTTG AGGATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTA TGTTTTTGATAGTTTGCAGTTGTCAGTTTGTTTGCATGAGGTGGCATACT GGTACATTCTAAGCATTGGAGCACAGACTGACTTCCTTTCTGTCTTCTTC TCTGGATATACCTTCAAACACAAAATGGTCTATGAAGACACACTCACCC TATTCCCATTCTCAGGAGAAACTGTCTTCATGTCGATGGAAAACCCAGG TCTATGGATTCTGGGGTGCCACAACTCAGACTTTCGGAACAGAGGCATG ACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTGATTATT ACGAGGACAGTTATGAAGATATTTCAGCATACTTGCTGAGTAAAAACAA TGCCATTGAACCAAGAAGCTTCTCCCAGAATTCAAGACACCCTAGCACT AGGCAAAAGCAATTTAATGCCACCACAATTCCAGAAAATGACATAGAG AAGACTGACCCTTGGTTTGCACACAGAACACCTATGCCTAAAATACAAA ATGTCTCCTCTAGTGATTTGTTGATGCTCTTGCGACAGAGTCCTACTCCA CATGGGCTATCCTTATCTGATCTCCAAGAAGCCAAATATGAGACTTTTTC TGATGATCCATCACCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAA ATGACACACTTCAGGCCACAGCTCCATCACAGTGGGGACATGGTATTTA CCCCTGAGTCAGGCCTCCAATTAAGATTAAATGAGAAACTGGGGACAA CTGCAGCAACAGAGTTGAAGAAACTTGATTTCAAAGTTTCTAGTACATC

	AAATAATCTGATTTCAACAATTCCATCAGACAATTTGGCAGCAGGTACT GATAATACAAGTTCCTTAGGACCCCCAAGTATGCCAGTTCATTATGATA GTCAATTAGATACCACTCTATTTGGCAAAAAGTCATCTCCCCTTACTGA GTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATAATGATTCAAAGTTG TTAGAATCAGGTTTAATGAATAGCCAAGAAAGTTCATGGGGAAAAAAT GTATCGTCAACAGAGAGTGGTAGGTTATTTAAAGGGAAAAGAGCTCAT GGACCTGCTTTGTTGACTAAAGATAATGCCTTATTCAAAGTTAGCATCTC TTTGTTAAAGACAAACAAAACTTCCAATAATTCAGCAACTAATAGAAAG ACTCACATTGATGGCCCATCATTATTAATTGAGAATAGTCCATCAGTCT GGCAAAATATATTAGAAAGTGACACTGAGTTTAAAAAAGTGACACCTTT GATTCATGACAGAATGCTTATGGACAAAAATGCTACAGCTTTGAGGCTA AATCATATGTCAAATAAAACTACTTCATCAAAAAACATGGAAATGGTCC AACAGAAAAAAGAGGGCCCCATTCCACCAGATGCACAAAATCCAGATA TGTCGTTCTTTAAGATGCTATTCTTGCCAGAATCAGCAAGGTGGATACA AAGGACTCATGGAAAGAACTCTCTGAACTCTGGGCAAGGCCCCAGTCC AAAGCAATTAGTATCCTTAGGACCAGAAAAATCTGTGGAAGGTCAGAA TTTCTTGTCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACA AAGGACGTAGGACTCAAAGAGATGGTTTTTCCAAGCAGCAGAAACCTA TTTCTTACTAACTTGGATAATTTACATGAAAATAATACACACAATCAAG AAAAAAAAATTCAGGAAGAAATAGAAAAGAAGGAAACATTAATCCAA GAGAATGTAGTTTTGCCTCAGATACATACAGTGACTGGCACTAAGAATT TCATGAAGAACCTTTTCTTACTGAGCACTAGGCAAAATGTAGAAGGTTC ATATGACGGGGCATATGCTCCAGTACTTCAAGATTTTAGGTCATTAAAT GATTCAACAAATAGAACAAAGAAACACACAGCTCATTTCTCAAAAAAA GGGGAGGAAGAAAACTTGGAAGGCTTGGGAAATCAAACCAAGCAAATT GTAGAGAAATATGCATGCACCACAAGGATATCTCCTAATACAAGCCAG CAGAATTTTGTCACGCAACGTAGTAAGAGAGCTTTGAAACAATTCAGAC TCCCACTAGAAGAAACAGAACTTGAAAAAAGGATAATTGTGGATGACA CCTCAACCCAGTGGTCCAAAAACATGAAACATTTGACCCCGAGCACCCT CACACAGATAGACTACAATGAGAAGGAGAAAGGGGCCATTACTCAGTC TCCCTTATCAGATTGCCTTACGAGGAGTCATAGCATCCCTCAAGCAAAT AGATCTCCATTACCCATTGCAAAGGTATCATCATTTCCATCTATTAGACC TATATATCTGACCAGGGTCCTATTCCAAGACAACTCTTCTCATCTTCCAG CAGCATCTTATAGAAAGAAAGATTCTGGGGTCCAAGAAAGCAGTCATTT CTTACAAGGAGCCAAAAAAAATAACCTTTCTTTAGCCATTCTAACCTTG GAGATGACTGGTGATCAAAGAGAGGTTGGCTCCCTGGGGACAAGTGCC ACAAATTCAGTCACATACAAGAAAGTTGAGAACACTGTTCTCCCGAAAC CAGACTTGCCCAAAACATCTGGCAAAGTTGAATTGCTTCCAAAAGTTCA CATTTATCAGAAGGACCTATTCCCTACGGAAACTAGCAATGGGTCTCCT GGCCATCTGGATCTCGTGGAAGGGAGCCTTCTTCAGGGAACAGAGGGA GCGATTAAGTGGAATGAAGCAAACAGACCTGGAAAAGTTCCCTTTCTGA GAGTAGCAACAGAAAGCTCTGCAAAGACTCCCTCCAAGCTATTGGATCC TCTTGCTTGGGATAACCACTATGGTACTCAGATACCAAAAGAAGAGTGG AAATCCCAAGAGAAGTCACCAGAAAAAACAGCTTTTAAGAAAAAGGAT ACCATTTTGTCCCTGAACGCTTGTGAAAGCAATCATGCAATAGCAGCAA TAAATGAGGGACAAAATAAGCCCGAAATAGAAGTCACCTGGGCAAAGC AAGGTAGGACTGAAAGGCTGTGCTCTCAAAACCCACCAGTCTTGAAAC GCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGA AATTGACTATGATGATACCATATCAGTTGAAATGAAGAAGGAAGATTTT GACATTTATGATGAGGATGAAAATCAGAGCCCCCGCAGCTTTCAAAAG AAAACACGACACTATTTTATTGCTGCAGTGGAGAGGCTCTGGGATTATG GGATGAGTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCA GTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACTGATGGCTCC TTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCC TGGGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTT CAGAAATCAGGCCTCTCGTCCCTATTCCTTCTATTCTAGCCTTATTTCTT ATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCA AGCCTAATGAAACCAAAACTTACTTTTGGAAAGTGCAACATCATATGGC ACCCACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGAT

	GTTGACCTGGAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGG TCTGCCACACTAACACACTGAACCCTGCTCATGGGAGACAAGTGACAGT ACAGGAATTTGCTCTGTTTTTCACCATCTTTGATGAGACCAAAAGCTGGT ACTTCACTGAAAATATGGAAAGAAACTGCAGGGCTCCCTGCAATATCCA GATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAAT GGCTACATAATGGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAA GGATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTC TATTCATTTCAGTGGACATGTGTTCACTGTACGAAAAAAAGAGGAGTAT AAAATGGCACTGTACAATCTCTATCCAGGTGTTTTTGAGACAGTGGAAA TGTTACCATCCAAAGCTGGAATTTGGCGGGTGGAATGCCTTATTGGCGA GCATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAG TGTCAGACTCCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGA TTACAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGACT TCATTATTCCGGATCAATCAATGCCTGGAGCACCAAGGAGCCCTTTTCTT GGATCAAGGTGGATCTGTTGGCACCAATGATTATTCACGGCATCAAGAC CCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCA TCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTC CACTGGAACCTTAATGGTCTTCTTTGGCAATGTGGATTCATCTGGGATA AAACACAATATTTTTAACCCTCCAATTATTGCTCGATACATCCGTTTGCA CCCAACTCATTATAGCATTCGCAGCACTCTTCGCATGGAGTTGATGGGC TGTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGAGTAAAGCAA TATCAGATGCACAGATTACTGCTTCATCCTACTTTACCAATATGTTTGCC ACCTGGTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATG CCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACT TCCAGAAGACAATGAAAGTCACAGGAGTAACTACTCAGGGAGTAAAAT CTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGCAGTCA AGATGGCCATCAGTGGACTCTCTTTTTTCAGAATGGCAAAGTAAAGGTT TTTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAACTCTCTAGACC CACCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTGGGTGCA CCAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTC TAC
SEQ ID NO: 32 编码 BDD-co6FVIII (V1.0)的核苷酸序列 (无XTEN)	GCCACTCGCCGGTACTACCTTGGAGCCGTGGAGCTTTCATGGGACTACA TGCAGAGCGACCTGGGCGAACTCCCCGTGGATGCCAGATTCCCCCCCCG CGTGCCAAAGTCCTTCCCCTTTAACACCTCCGTGGTGTACAAGAAAACC CTCTTTGTCGAGTTCACTGACCACCTGTTCAACATCGCCAAGCCGCGCCC ACCTTGGATGGGCCTCCTGGGACCGACCATTCAAGCTGAAGTGTACGAC ACCGTGGTGATCACCCTGAAGAACATGGCGTCCCACCCCGTGTCCCTGC ATGCGGTCGGAGTGTCCTACTGGAAGGCCTCCGAAGGAGCTGAGTACG ACGACCAGACTAGCCAGCGGGAAAAGGAGGACGATAAAGTGTTCCCGG GCGGCTCGCATACTTACGTGTGGCAAGTCCTGAAGGAAAACGGACCTAT GGCATCCGATCCTCTGTGCCTGACTTACTCCTACCTTTCCCATGTGGACC TCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCACTTCTCGTGTGCCG CGAAGGTTCGCTCGCTAAGGAAAAGACCCAGACCCTCCATAAGTTCATC CTTTTGTTCGCTGTGTTCGATGAAGGAAAGTCATGGCATTCCGAAACTA AGAACTCGCTGATGCAGGACCGGGATGCCGCCTCAGCCCGCGCCTGGCC TAAAATGCATACAGTCAACGGATACGTGAATCGGTCACTGCCCGGGCTC ATCGGTTGTCACAGAAAGTCCGTGTACTGGCACGTCATCGGCATGGGCA CTACGCCTGAAGTGCACTCCATCTTCCTGGAAGGGCACACCTTCCTCGT GCGCAACCACCGCCAGGCCTCTCTGGAAATCTCCCCGATTACCTTTCTG ACCGCCCAGACTCTGCTCATGGACCTGGGGCAGTTCCTTCTCTTCTGCCA CATCTCCAGCCATCAGCACGACGGAATGGAGGCCTACGTGAAGGTGGA CTCATGCCCGGAAGAACCTCAGTTGCGGATGAAGAACAACGAGGAGGC CGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACGTCGTGCGG TTCGATGACGACAACAGCCCCAGCTTCATCCAGATTCGCAGCGTGGCCA AGAAGCACCCCAAAACCTGGGTGCACTACATCGCGGCCGAGGAAGAAG ATTGGGACTACGCCCCGTTGGTGCTGGCACCCGATGACCGGTCGTACAA GTCCCAGTATCTGAACAATGGTCCGCAGCGGATTGGCAGAAAGTACAA

	GAAAGTGCGGTTCATGGCGTACACTGACGAAACGTTTAAGACCCGGGA GGCCATTCAACATGAGAGCGGCATTCTGGGACCACTGCTGTACGGAGA GGTCGGCGATACCCTGCTCATCATCTTCAAAAACCAGGCCTCCCGGCCT TACAACATCTACCCTCACGGAATCACCGACGTGCGGCCACTCTACTCGC GGCGCCTGCCGAAGGGCGTCAAGCACCTGAAAGACTTCCCTATCCTGCC GGGCGAAATCTTCAAGTATAAGTGGACCGTCACCGTGGAGGACGGGCC CACCAAGAGCGATCCTAGGTGTCTGACTCGGTACTACTCCAGCTTCGTG AACATGGAACGGGACCTGGCATCGGGACTCATTGGACCGCTGCTGATCT GCTACAAAGAGTCGGTGGATCAACGCGGCAACCAGATCATGTCCGACA AGCGCAACGTGATCCTGTTCTCCGTGTTTGATGAAAACAGATCCTGGTA CCTCACTGAAAACATCCAGAGGTTCCTCCCAAACCCCGCAGGAGTGCAA CTGGAGGACCCTGAGTTTCAGGCCTCGAATATCATGCACTCGATTAACG GTTACGTGTTCGACTCGCTGCAGCTGAGCGTGTGCCTCCATGAAGTCGC TTACTGGTACATTCTGTCCATCGGCGCCCAGACTGACTTCCTGAGCGTGT TCTTTTCCGGTTACACCTTTAAGCACAAGATGGTGTACGAAGATACCCT GACCCTGTTCCCTTTCTCCGGCGAAACGGTGTTCATGTCGATGGAGAAC CCGGGTCTGTGGATTCTGGGATGCCACAACAGCGACTTTCGGAACCGCG GAATGACTGCCCTGCTGAAGGTGTCCTCATGCGACAAGAACACCGGAG ACTACTACGAGGACTCCTACGAGGATATCTCAGCCTACCTCCTGTCCAA GAACAACGCGATCGAGCCGCGCAGCTTCAGCCAGAACCCGCCTGTGCT GAAGAGGCACCAGCGAGAAATTACCCGGACCACCCTCCAATCGGATCA GGAGGAAATCGACTACGACGACACCATCTCGGTGGAAATGAAGAAGGA AGATTTCGATATCTACGACGAGGACGAAAATCAGTCCCCTCGCTCATTC CAAAAGAAAACTAGACACTACTTTATCGCCGCGGTGGAAAGACTGTGG GACTATGGAATGTCATCCAGCCCTCACGTCCTTCGGAACCGGGCCCAGA GCGGATCGGTGCCTCAGTTCAAGAAAGTGGTGTTCCAGGAGTTCACCGA CGGCAGCTTCACCCAGCCGCTGTACCGGGGAGAACTGAACGAACACCT GGGCCTGCTCGGTCCCTACATCCGCGCGGAAGTGGAGGATAACATCATG GTGACCTTCCGTAACCAAGCATCCAGACCTTACTCCTTCTATTCCTCCCT GATCTCATACGAGGAGGACCAGCGCCAAGGCGCCGAGCCCCGCAAGAA CTTCGTCAAGCCCAACGAGACTAAGACCTACTTCTGGAAGGTCCAACAC CATATGGCCCCGACCAAGGATGAGTTTGACTGCAAGGCCTGGGCCTACT TCTCCGACGTGGACCTTGAGAAGGATGTCCATTCCGGCCTGATCGGGCC GCTGCTCGTGTGTCACACCAACACCCTGAACCCAGCGCATGGACGCCAG GTCACCGTCCAGGAGTTTGCTCTGTTCTTCACCATTTTTGACGAAACTAA GTCCTGGTACTTCACCGAGAATATGGAGCGAAACTGTAGAGCGCCCTGC AATATCCAGATGGAAGATCCGACTTTCAAGGAGAACTATAGATTCCACG CCATCAACGGGTACATCATGGATACTCTGCCGGGGCTGGTCATGGCCCA GGATCAGAGGATTCGGTGGTACTTGCTGTCAATGGGATCGAACGAAAA CATTCACTCCATTCACTTCTCCGGTCACGTGTTCACTGTGCGCAAGAAGG AGGAGTACAAGATGGCGCTGTACAATCTGTACCCCGGGGGTGTTCGAAAC TGTGGAGATGCTGCCGTCCAAGGCCGGCATCTGGAGAGTGGAGTGCCTG ATCGGAGAGCACCTCCACGCGGGGATGTCCACCCTCTTCCTGGTGTACT CGAATAAGTGCCAGACCCCGCTGGGCATGGCCTCGGGCCACATCAGAG ACTTCCAGATCACAGCAAGCGGACAATACGGCCAATGGGCGCCGAAGC TGGCCCGCTTGCACTACTCCGGATCGATCAACGCATGGTCCACCAAGGA ACCGTTCTCGTGGATTAAGGTGGACCTCCTGGCCCCTATGATTATCCAC GGAATTAAGACCCAGGGCGCCAGGCAGAAGTTCTCCTCCCTGTACATCT CGCAATTCATCATCATGTACAGCCTGGACGGGAAGAAGTGGCAGACTTA CAGGGGAAACTCCACCGGCACCCTGATGGTCTTTTTCGGCAACGTGGAT TCCTCCGGCATTAAGCACAACATCTTCAACCCACCGATCATAGCCAGAT ATATTAGGCTCCACCCCACTCACTACTCAATCCGCTCAACTCTTCGGATG GAACTCATGGGGTGCGACCTGAACTCCTGCTCCATGCCGTTGGGGATGG AATCAAAGGCTATTAGCGACGCCCAGATCACCGCGAGCTCCTACTTCAC TAACATGTTCGCCACCTGGAGCCCCTCCAAGGCCAGGCTGCACTTGCAG GGACGGTCAAATGCCTGGCGGCCGCAAGTGAACAATCCGAAGGAATGG CTTCAAGTGGATTTCCAAAAGACCATGAAAGTGACCGGAGTCACCACCC AGGGAGTGAAGTCCCTTCTGACCTCGATGTATGTGAAGGAGTTCCTGAT TAGCAGCAGCCAGGACGGGCACCAGTGGACCCTGTTCTTCCAAAACGG

	AAAGGTCAAGGTGTTCCAGGGGAACCAGGACTCGTTCACACCCGTGGT GAACTCCCTGGACCCCCCACTGCTGACGCGGTACTTGAGGATTCATCCT CAGTCCTGGGTCCATCAGATTGCATTGCGAATGGAAGTCCTGGGCTGCG AGGCCCAGGACCTGTACTGA
SEQ ID NO: 33 编码 coBDDFVIII 的核苷酸序列 (V2.0) (无XTEN)	GCCACCCGCCGGTATTACTTAGGTGCTGTGGAACTGAGCTGGGACTACA TGCAGTCCGACCTGGGAGAACTGCCGGTGGACGCGAGATTCCCACCTAG AGTCCCGAAGTCCTTCCCATTCAACACCTCCGTGGTCTACAAAAAGACC CTGTTCGTGGAGTTCACTGACCACCTTTTCAATATTGCCAAGCCGCGCCC CCCCTGGATGGGCCTGCTTGGTCCTACGATCCAAGCAGAGGTCTACGAC ACCGTGGTCATCACACTGAAGAACATGGCCTCACACCCCGTGTCGCTGC ATGCTGTGGGAGTGTCCTACTGGAAGGCCTCAGAGGGTGCCGAATATGA TGACCAGACCAGCCAGAGGGAAAAGGAGGATGACAAAGTGTTCCCGGG TGGCAGCCACACTTACGTGTGGCAAGTGCTGAAGGAAAACGGGCCTAT GGCGTCGGACCCCCTATGCCTGACCTACTCCTACCTGTCCCATGTGGAC CTTGTGAAGGATCTCAACTCGGGACTGATCGGCGCCCTCTTGGTGTGCA GAGAAGGCAGCCTGGCGAAGGAAAAGACTCAGACCCTGCACAAGTTCA TTCTGTTGTTTGCTGTGTTCGATGAAGGAAAGTCCTGGCACTCAGAAAC CAAGAACTCGCTGATGCAGGATAGAGATGCGGCCTCGGCCAGAGCCTG GCCTAAAATGCACACCGTCAACGGATATGTGAACAGGTCGCTCCCTGGC CTCATCGGCTGCCACAGAAAGTCCGTGTATTGGCATGTGATCGGCATGG GTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGGGCCATACCTTCTTG GTGCGCAACCACAGACAGGCCTCGCTGGAAATCTCGCCTATCACTTTCT TGACTGCGCAGACCCTCCTTATGGACCTTGGACAGTTCCTGCTGTTCTGT CACATCAGCTCCCATCAGCATGATGGGATGGAGGCCTATGTCAAAGTGG ACTCCTGCCCTGAGGAGCCACAGCTCCGGATGAAGAACAATGAGGAAG CGGAGGATTACGACGACGACCTGACTGACAGCGAAATGGACGTCGTGC GATTCGATGACGACAACAGCCCGTCCTTCATCCAAATTAGATCAGTGGC GAAGAAGCACCCCAAGACCTGGGTGCACTACATTGCCGCCGAGGAAGA GGACTGGGACTACGCGCCGCTGGTGCTGGCGCCAGACGACAGGAGCTA CAAGTCCCAGTACCTCAACAACGGGCCGCAGCGCATTGGCAGGAAGTA CAAGAAAGTCCGCTTCATGGCCTACACTGATGAAACCTTCAAGACGAGG GAAGCCATCCAGCACGAGTCAGGCATCCTGGGACCGCTCCTTTACGGCG AAGTCGGGGATACCCTGCTCATCATTTTCAAGAACCAGGCATCGCGGCC CTACAACATCTACCCTCACGGGATCACAGACGTGCGCCCGCTCTACTCC CGCCGGCTGCCCAAGGGAGTGAAGCACCTGAAGGATTTTCCCATCCTGC CGGGAGAAATCTTCAAGTACAAGTGGACCGTGACTGTGGAAGATGGCC CTACCAAGTCGGACCCTCGCTGTCTGACCCGGTACTATTCCTCGTTTGTG AACATGGAGCGCGACCTGGCCTCGGGGCTGATTGGTCCGCTGCTGATCT GCTACAAGGAGTCCGTGGACCAGCGCGGGAACCAGATCATGTCCGACA AGCGCAACGTGATCCTGTTCTCTGTCTTTGATGAAAACAGATCGTGGTA CTTGACTGAGAATATCCAGCGGTTCCTGCCCAACCCAGCGGGAGTGCAA CTGGAGGACCCGGAGTTCCAGGCCTCAAACATTATGCACTCTATCAACG GCTATGTGTTCGACTCGCTCCAACTGAGCGTGTGCCTGCATGAAGTGGC ATACTGGTACATTCTGTCCATCGGAGCCCAGACCGACTTCCTGTCCGTGT TCTTCTCCGGATACACCTTCAAGCATAAGATGGTGTACGAGGACACTCT GACCCTCTTCCCATTTTCCGGAGAAACTGTGTTCATGTCAATGGAAAAC CCGGGCTTGTGGATTCTGGGTTGCCATAACTCGGACTTCCGGAATAGAG GGATGACCGCCCTGCTGAAAGTGTCCAGCTGTGACAAGAATACCGGCG ATTACTACGAGGACAGCTATGAGGACATCTCCGCTTATCTGCTGTCCAA GAACAACGCCATTGAACCCAGGTCCTTCTCCCAAAACGGTGCACCGGCC TCATCCCCCCCCCGTGCTGAAGCGGCATCAAAGAGAGATCACCAGGACC ACTCTCCAGTCCGATCAGGAAGAAATTGACTACGACGATACTATCAGCG TGGAGATGAAGAAGGAGGACTTCGACATCTACGATGAGGATGAGAACC AGTCCCCTCGGAGCTTTCAGAAGAAAACCCGCCACTACTTCATCGCTGC CGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTCACCGCATGTGCTG CGGAATAGAGCGCAGTCAGGATCGGTGCCCCAGTTCAAGAAGGTCGTG TTCCAAGAGTTCACCGACGGGTCCTTCACTCAACCCCTGTACCGGGGCG AACTCAACGAACACCTGGGACTGCTTGGGCCGTATATCAGGGCAGAAG

	TGGAAGATAACATCATGGTCACCTTCCGCAACCAGGCCTCCCGGCCGTA CAGCTTCTACTCTTCACTGATCTCCTACGAGGAAGATCAGCGGCAGGGA GCCGAGCCCCGGAAGAACTTCGTCAAGCCTAACGAAACTAAGACCTAC TTTTGGAAGGTCCAGCATCACATGGCCCCGACCAAAGACGAGTTCGACT GTAAAGCCTGGGCCTACTTCTCCGATGTGGACCTGGAGAAGGACGTGCA CTCGGGACTCATTGGCCCGCTCCTTGTGTGCCATACTAATACCCTGAACC CTGCTCACGGTCGCCAAGTCACAGTGCAGGAGTTCGCCCTCTTCTTCAC CATCTTCGATGAAACAAAGTCCTGGTACTTTACTGAGAACATGGAACGC AATTGCAGGGCACCCTGCAACATCCAGATGGAAGATCCCACCTTCAAGG AAAACTACCGGTTTCATGCCATTAACGGCTACATAATGGACACGTTGCC AGGACTGGTCATGGCCCAGGACCAGAGAATCCGGTGGTATCTGCTCTCC ATGGGCTCCAACGAAAACATTCACAGCATTCATTTTTCCGGCCATGTGT TCACCGTCCGGAAGAAGGAAGAGTACAAGATGGCTCTGTACAACCTCT ACCCTGGAGTGTTCGAGACTGTGGAAATGCTGCCTAGCAAGGCCGGCAT TTGGAGAGTGGAATGCCTGATCGGAGAGCATTTGCACGCCGGAATGTCC ACCCTGTTTCTTGTGTACTCCAACAAGTGCCAGACCCCGCTGGGAATGG CCTCAGGTCATATTAGGGATTTCCAGATCACTGCTTCGGGGCAGTACGG GCAGTGGGCACCTAAGTTGGCCCGGCTGCACTACTCTGGCTCCATCAAT GCCTGGTCCACCAAGGAACCCTTCTCCTGGATTAAGGTGGACCTCCTGG CCCCAATGATTATTCACGGTATTAAGACCCAGGGTGCCCGACAGAAGTT CTCCTCACTCTACATCTCGCAATTCATCATAATGTACAGCCTGGATGGG AAGAAGTGGCAGACCTACCGGGGAAACTCCACTGGAACGCTCATGGTG TTTTTCGGCAACGTGGACTCCTCCGGCATTAAGCACAACATCTTCAACC CTCCGATCATTGCTCGGTACATCCGGCTGCACCCAACTCACTACAGCAT CCGGTCCACCCTGCGGATGGAACTGATGGGTTGTGACCTGAACTCCTGC TCCATGCCCCTTGGGATGGAATCCAAGGCCATTAGCGATGCACAGATCA CCGCCTCTTCATACTTCACCAACATGTTCGCGACCTGGTCCCCGTCGAAG GCCCGCCTGCACCTCCAAGGTCGCTCCAATGCGTGGCGGCCTCAAGTGA ACAACCCCAAGGAGTGGCTCCAGGTCGACTTCCAAAAGACCATGAAGG TCACCGGAGTGACCACCCAGGGCGTGAAGTCCCTGCTGACCTCTATGTA CGTTAAGGAGTTCCTCATCTCCTCAAGCCAAGACGGACATCAGTGGACC CTGTTCTTCCAAAACGGAAAAGTCAAAGTATTCCAGGGCAACCAGGACT CCTTCACCCCTGTGGTCAACAGCCTGGACCCCCCATTGCTGACCCGCTA CCTCCGCATCCACCCCCAAAGCTGGGTCCACCAGATCGCACTGCGCATG GAGGTCCTTGGATGCGAAGCCCAAGATCTGTACTAA
SEQ ID NO: 34 V1.0 表达盒 TTP-内含子-BDDFVIIIco6 XTEN (V1.0)-WPRE-bGHPolyA	ATGCAGATTGAGCTGTCCACTTGTTTCTTCCTGTGCCTCCTGCGCTTCTG TTTCTCCGCCACTCGCCGGTACTACCTTGGAGCCGTGGAGCTTTCATGGG ACTACATGCAGAGCGACCTGGGCGAACTCCCCGTGGATGCCAGATTCCC CCCCCGCGTGCCAAAGTCCTTCCCCTTTAACACCTCCGTGGTGTACAAG AAAACCCTCTTTGTCGAGTTCACTGACCACCTGTTCAACATCGCCAAGC CGCGCCCACCTTGGATGGGCCTCCTGGGACCGACCATTCAAGCTGAAGT GTACGACACCGTGGTGATCACCCTGAAGAACATGGCGTCCCACCCCGTG TCCCTGCATGCGGTCGGAGTGTCCTACTGGAAGGCCTCCGAAGGAGCTG AGTACGACGACCAGACTAGCCAGCGGGAAAAGGAGGACGATAAAGTGT TCCCGGGCGGCTCGCATACTTACGTGTGGCAAGTCCTGAAGGAAAACGG ACCTATGGCATCCGATCCTCTGTGCCTGACTTACTCCTACCTTTCCCATG TGGACCTCGTGAAGGACCTGAACAGCGGGCTGATTGGTGCACTTCTCGT GTGCCGCGAAGGTTCGCTCGCTAAGGAAAAGACCCAGACCCTCCATAA GTTCATCCTTTTGTTCGCTGTGTTCGATGAAGGAAAGTCATGGCATTCCG AAACTAAGAACTCGCTGATGCAGGACCGGGATGCCGCCTCAGCCCGCG CCTGGCCTAAAATGCATACAGTCAACGGATACGTGAATCGGTCACTGCC CGGGCTCATCGGTTGTCACAGAAAGTCCGTGTACTGGCACGTCATCGGC ATGGGCACTACGCCTGAAGTGCACTCCATCTTCCTGGAAGGGCACACCT TCCTCGTGCGCAACCACCGCCAGGCCTCTCTGGAAATCTCCCCGATTAC CTTTCTGACCGCCCAGACTCTGCTCATGGACCTGGGGCAGTTCCTTCTCT TCTGCCACATCTCCAGCCATCAGCACGACGGAATGGAGGCCTACGTGAA GGTGGACTCATGCCCGGAAGAACCTCAGTTGCGGATGAAGAACAACGA GGAGGCCGAGGACTATGACGACGATTTGACTGACTCCGAGATGGACGT

	CGTGCGGTTCGATGACGACAACAGCCCCAGCTTCATCCAGATTCGCAGC GTGGCCAAGAAGCACCCCAAAACCTGGGTGCACTACATCGCGGCCGAG GAAGAAGATTGGGACTACGCCCCGTTGGTGCTGGCACCCGATGACCGGT CGTACAAGTCCCAGTATCTGAACAATGGTCCGCAGCGGATTGGCAGAA AGTACAAGAAAGTGCGGTTCATGGCGTACACTGACGAAACGTTTAAGA CCCGGGAGGCCATTCAACATGAGAGCGGCATTCTGGGACCACTGCTGTA CGGAGAGGTCGGCGATACCCTGCTCATCATCTTCAAAAACCAGGCCTCC CGGCCTTACAACATCTACCCTCACGGAATCACCGACGTGCGGCCACTCT ACTCGCGGCGCCTGCCGAAGGGCGTCAAGCACCTGAAAGACTTCCCTAT CCTGCCGGGCGAAATCTTCAAGTATAAGTGGACCGTCACCGTGGAGGAC GGGCCCACCAAGAGCGATCCTAGGTGTCTGACTCGGTACTACTCCAGCT TCGTGAACATGGAACGGGACCTGGCATCGGGACTCATTGGACCGCTGCT GATCTGCTACAAAGAGTCGGTGGATCAACGCGGCAACCAGATCATGTCC GACAAGCGCAACGTGATCCTGTTCTCCGTGTTTGATGAAAACAGATCCT GGTACCTCACTGAAAACATCCAGAGGTTCCTCCCAAACCCCGCAGGAGT GCAACTGGAGGACCCTGAGTTTCAGGCCTCGAATATCATGCACTCGATT AACGGTTACGTGTTCGACTCGCTGCAACTGAGCGTGTGCCTCCATGAAG TCGCTTACTGGTACATTCTGTCCATCGGCGCCCAGACTGACTTCCTGAGC GTGTTCTTTTCCGGTTACACCTTTAAGCACAAGATGGTGTACGAAGATA CCCTGACCCTGTTCCCTTTCTCCGGCGAAACGGTGTTCATGTCGATGGAG AACCCGGGTCTGTGGATTCTGGGATGCCACAACAGCGACTTTCGGAACC GCGGAATGACTGCCCTGCTGAAGGTGTCCTCATGCGACAAGAACACCG GAGACTACTACGAGGACTCCTACGAGGATATCTCAGCCTACCTCCTGTC CAAGAACAACGCGATCGAGCCGCGCAGCTTCAGCCAGAACGGCGCGCC AACATCAGAGAGCGCCACCCCTGAAAGTGGTCCCGGGAGCGAGCCAGC CACATCTGGGTCGGAAACGCCAGGCACAAGTGAGTCTGCAACTCCCGA GTCCGGACCTGGCTCCGAGCCTGCCACTAGCGGCTCCGAGACTCCGGGA ACTTCCGAGAGCGCTACACCAGAAAGCGGACCCGGAACCAGTACCGAA CCTAGCGAGGGCTCTGCTCCGGGCAGCCCAGCCGGCTCTCCTACATCCA CGGAGGAGGGCACTTCCGAATCCGCCACCCCGGAGTCAGGGCCAGGAT CTGAACCCGCTACCTCAGGCAGTGAGACGCCAGGAACGAGCGAGTCCG CTACACCGGAGAGTGGGCCAGGGAGCCCTGCTGGATCTCCTACGTCCAC TGAGGAAGGGTCACCAGCGGGCTCGCCCACCAGCACTGAAGAAGGTGC CTCGAGCCCGCCTGTGCTGAAGAGGCACCAGCGAGAAATTACCCGGAC CACCCTCCAATCGGATCAGGAGGAAATCGACTACGACGACACCATCTCG GTGGAAATGAAGAAGGAAGATTTCGATATCTACGACGAGGACGAAAAT CAGTCCCCTCGCTCATTCCAAAAGAAAACTAGACACTACTTTATCGCCG CGGTGGAAAGACTGTGGGACTATGGAATGTCATCCAGCCCTCACGTCCT TCGGAACCGGGCCCAGAGCGGATCGGTGCCTCAGTTCAAGAAAGTGGT GTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCGCTGTACCGGGGA GAACTGAACGAACACCTGGGCCTGCTCGGTCCCTACATCCGCGCGGAAG TGGAGGATAACATCATGGTGACCTTCCGTAACCAAGCATCCAGACCTTA CTCCTTCTATTCCTCCCTGATCTCATACGAGGAGGACCAGCGCCAAGGC GCCGAGCCCCGCAAGAACTTCGTCAAGCCCAACGAGACTAAGACCTAC TTCTGGAAGGTCCAACACCATATGGCCCCGACCAAGGATGAGTTTGACT GCAAGGCCTGGGCCTACTTCTCCGACGTGGACCTTGAGAAGGATGTCCA TTCCGGCCTGATCGGGCCGCTGCTCGTGTGTCACACCAACACCCTGAAC CCAGCGCATGGACGCCAGGTCACCGTCCAGGAGTTTGCTCTGTTCTTCA CCATTTTTGACGAAACTAAGTCCTGGTACTTCACCGAGAATATGGAGCG AAACTGTAGAGCGCCCTGCAATATCCAGATGGAAGATCCGACTTTCAAG GAGAACTATAGATTCCACGCCATCAACGGGTACATCATGGATACTCTGC CGGGGCTGGTCATGGCCCAGGATCAGAGGATTCGGTGGTACTTGCTGTC AATGGGATCGAACGAAAACATTCACTCCATTCACTTCTCCGGTCACGTG TTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCGCTGTACAATCTGT ACCCCGGGGTGTTCGAAACTGTGGAGATGCTGCCGTCCAAGGCCGGCAT CTGGAGAGTGGAGTGCCTGATCGGAGAGCACCTCCACGCGGGGATGTC CACCCTCTTCCTGGTGTACTCGAATAAGTGCCAGACCCCGCTGGGCATG GCCTCGGGCCACATCAGAGACTTCCAGATCACAGCAAGCGGACAATAC GGCCAATGGGCGCCGAAGCTGGCCCGCTTGCACTACTCCGGATCGATCA

	ACGCATGGTCCACCAAGGAACCGTTCTCGTGGATTAAGGTGGACCTCCT GGCCCCTATGATTATCCACGGAATTAAGACCCAGGGCGCCAGGCAGAA GTTCTCCTCCCTGTACATCTCGCAATTCATCATCATGTACAGCCTGGACG GGAAGAAGTGGCAGACTTACAGGGGAAACTCCACCGGCACCCTGATGG TCTTTTTCGGCAACGTGGATTCCTCCGGCATTAAGCACAACATCTTCAAC CCACCGATCATAGCCAGATATATTAGGCTCCACCCCACTCACTACTCAA TCCGCTCAACTCTTCGGATGGAACTCATGGGGTGCGACCTGAACTCCTG CTCCATGCCGTTGGGGATGGAATCAAAGGCTATTAGCGACGCCCAGATC ACCGCGAGCTCCTACTTCACTAACATGTTCGCCACCTGGAGCCCCTCCA AGGCCAGGCTGCACTTGCAGGGACGGTCAAATGCCTGGCGGCCGCAAG TGAACAATCCGAAGGAATGGCTTCAAGTGGATTTCCAAAAGACCATGA AAGTGACCGGAGTCACCACCCAGGGAGTGAAGTCCCTTCTGACCTCGAT GTATGTGAAGGAGTTCCTGATTAGCAGCAGCCAGGACGGGCACCAGTG GACCCTGTTCTTCCAAAACGGAAAGGTCAAGGTGTTCCAGGGGAACCA GGACTCGTTCACACCCGTGGTGAACTCCCTGGACCCCCCACTGCTGACG CGGTACTTGAGGATTCATCCTCAGTCCTGGGTCCATCAGATTGCATTGC GAATGGAAGTCCTGGGCTGCGAGGCCCAGGACCTGTACTGA
SEQ ID NO: 35 V3.0 表达盒 人密码子优化 A1AT-内含子- BDDFVIIIXTEN-WPRE-bGHPolyA	ATCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGG CCCTTGGCAGCATTTACTCTCTCTGTTTGCTCTGGTTAATAATCTCAGGA GCACAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTA TCCTCTGGGCCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAA GCAGTCAAAAGTCCAAGTGGCCCTTGGCAGCATTTACTCTCTCTGTTTGC TCTGGTTAATAATCTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAG AAATCAACATCCTGGACTTATCCTCTGGGCCTCTCCCCACCTTCGAACTA GCCACTAGCCTGAGGCTGGTCAAAATTGAACCTCCTCCTGCTCTGAGCA GCCTGGGGGGCAGACTAAGCAGAGGGCTGTGCAGACCCACATAAAGAG CCTACTGTGTGCCAGGCACTTCACCCGAGGCACTTCACAAGCATGCTTG GGAATGAAACTTCCAACTCTTTGGGATGCAGGTGAAACAGTTCCTGGTT CAGAGAGGTGAAGCGGCCTGCCTGAGGCAGCACAGCTCTTCTTTACAGA TGTGCTTCCCCACCTCTACCCTGTCTCACGGCCCCCCATGCCAGCCTGAC GGTTGTGTCTGCCTCAGTCATGCTCCATTTTTCCATCGGGACCATCAAGA GGGTGTTTGTGTCTAAGGCTGACTGGGTAACTTTGGATGAGCGGTCTCT CCGCTCTGAGCCTGTTTCCTCATCTGTCAAATGGGCTCTAACCCACTCTG ATCTCCCAGGGCGGCAGTAAGTCTTCAGCATCAGGCATTTTGGGGTGAC TCAGTAAATGGTAGATCTTGCTACCAGTGGAACAGCCACTAAGGATTCT GCAGTGAGAGCAGAGGGCCAGCTAAGTGGTACTCTCCCAGAGACTGTC TGACTCACGCCACCCCCTCCACCTTGGACACAGGACGCTGTGGTTTCTG AGCCAGGTACAATGACTCCTTTCGGTAAGTGCAGTGGAAGCTGTACACT GCCCAGGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCAG CCAGTGGACTTAGCCCCTGTTTGCTCCTCCGATAACTGGGGTGACCTTG GTTAATATTCACCAGCAGCCTCCCCCGTTGCCCCTCTGGATCCACTGCTT AAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTCAGGCACCACCA CTGACCTGGGACAGGAATTCTCAGGAGCACAAACATTCCTGGAGGCAG GAGAAGAAATCAACATCCTGGACTTATCCTCTGGGCCTCTCCCCACCGA TATCTACCTGCTGATCGCCCGGCCCCTGTTCAAACATGTCCTAATACTCT GTCGGGGCAAAGGTCGGCAGTAGTTTTCCATCTTACTCAACATCCTCCC AGTGTACGTAGGATCCTGTCTGTCTGCACATTTCGTAGAGCGAGTGTTC CGATACTCTAATCTCCCGGGGCAAAGGTCGTATTGACTTAGGTTACTTA TTCTCCTTTTGTTGACTAAGTCAATAATCAGAATCAGCAGGTTTGGAGTC AGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGGGTATAAAAG CCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTCCTGCTAGA GTCGCTGCGCGCTGCCTTCGCCCCGTGCCCCGCTCCGCCGCCGCCTCGC GCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGG GCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTATTGAC GGCTTGTTTCTTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGGA AGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGGTGCGTGCGTGTGTGTG TGCGTGGGGAGCGCCGCGTGCGGCTCCGCGCTGCCCGGCGGCTGTGAGC GCTGCGGGCGCGGCGCGGGGCTTTGTGCGCTCCGCAGTGTGCGCGAGG

	GGAGCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGGCTGCGAGG GGAACAAAGGCTGCGTGCGGGGTGTGTGCGTGGGGGGGTGAGCAGGGG GTGTGGGCGCGTCGGTCGGGCTGCAACCCCCCCTGCACCCCCCTCCCCG AGTTGCTGAGCACGGCCCGGCTTCGGGTGCGGGGCTCCGTACGGGGCGT GGCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCAGGTGGGGGTG CCGGGCGGGGCGGGGCCGCCTCGGGCCGGGGAGGGCTCGGGGGAGGG GCGCGGCGGCCCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAGCCG CAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCCT TTGTCCCAAATCTGTGCGGAGCCGAAATCTGGGAGGCGCCGCCGCACCC CCTCTAGCGGGCGCGGGGCGAAGCGGTGCGGCGCCGGCAGGAAGGAAA TGGGCGGGGAGGGCCTTCGTGCGTCGCCGCGCCGCCGTCCCCTTCTCCC TCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCGGGGGGGA CGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGA GCCTCTGCTAACCTTGTTCTTGCCTTCTTCTTTTTCCTACAGCTCCTGGGC AACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAGAATTACTCGA GGCCACCATGCAGATTGAACTGTCCACTTGCTTCTTCCTGTGCCTCCTGC GGTTTTGCTTCTCGGCCACCCGCCGGTATTACTTAGGTGCTGTGGAACTG AGCTGGGACTACATGCAGTCCGACCTGGGAGAACTGCCGGTGGACGCG AGATTCCCACCTAGAGTCCCGAAGTCCTTCCCATTCAACACCTCCGTGG TCTACAAAAAGACCCTGTTCGTGGAGTTCACTGACCACCTTTTCAATATT GCCAAGCCGCGCCCCCCCTGGATGGGCCTGCTTGGTCCTACGATCCAAG CAGAGGTCTACGACACCGTGGTCATCACACTGAAGAACATGGCCTCACA CCCCGTGTCGCTGCATGCTGTGGGAGTGTCCTACTGGAAGGCCTCAGAG GGTGCCGAATATGATGACCAGACCAGCCAGAGGGAAAAGGAGGATGAC AAAGTGTTCCCGGGTGGCAGCCACACTTACGTGTGGCAAGTGCTGAAGG AAAACGGGCCTATGGCGTCGGACCCCCTATGCCTGACCTACTCCTACCT GTCCCATGTGGACCTTGTGAAGGATCTCAACTCGGGACTGATCGGCGCC CTCTTGGTGTGCAGAGAAGGCAGCCTGGCGAAGGAAAAGACTCAGACC CTGCACAAGTTCATTCTGTTGTTTGCTGTGTTCGATGAAGGAAAGTCCTG GCACTCAGAAACCAAGAACTCGCTGATGCAGGATAGAGATGCGGCCTC GGCCAGAGCCTGGCCTAAAATGCACACCGTCAACGGATATGTGAACAG GTCGCTCCCTGGCCTCATCGGCTGCCACAGAAAGTCCGTGTATTGGCAT GTGATCGGCATGGGTACTACTCCGGAAGTGCATAGTATCTTTCTGGAGG GCCATACCTTCTTGGTGCGCAACCACAGACAGGCCTCGCTGGAAATCTC GCCTATCACTTTCTTGACTGCGCAGACCCTCCTTATGGACCTTGGACAGT TCCTGCTGTTCTGTCACATCAGCTCCCATCAGCATGATGGGATGGAGGC CTATGTCAAAGTGGACTCCTGCCCTGAGGAGCCACAGCTCCGGATGAAG AACAATGAGGAAGCGGAGGATTACGACGACGACCTGACTGACAGCGAA ATGGACGTCGTGCGATTCGATGACGACAACAGCCCGTCCTTCATCCAAA TTAGATCAGTGGCGAAGAAGCACCCCAAGACCTGGGTGCACTACATTGC CGCCGAGGAAGAGGACTGGGACTACGCGCCGCTGGTGCTGGCGCCAGA CGACAGGAGCTACAAGTCCCAGTACCTCAACAACGGGCCGCAGCGCAT TGGCAGGAAGTACAAGAAAGTCCGCTTCATGGCCTACACTGATGAAAC CTTCAAGACGAGGGAAGCCATCCAGCACGAGTCAGGCATCCTGGGACC GCTCCTTTACGGCGAAGTCGGGGATACCCTGCTCATCATTTTCAAGAAC CAGGCATCGCGGCCCTACAACATCTACCCTCACGGGATCACAGACGTGC GCCCGCTCTACTCCCGCCGGCTGCCCAAGGGAGTGAAGCACCTGAAGG ATTTTCCCATCCTGCCGGGAGAAATCTTCAAGTACAAGTGGACCGTGAC TGTGGAAGATGGCCCTACCAAGTCGGACCCTCGCTGTCTGACCCGGTAC TATTCCTCGTTTGTGAACATGGAGCGCGACCTGGCCTCGGGGCTGATTG GTCCGCTGCTGATCTGCTACAAGGAGTCCGTGGACCAGCGCGGGAACCA GATCATGTCCGACAAGCGCAACGTGATCCTGTTCTCTGTCTTTGATGAA AACAGATCGTGGTACTTGACTGAGAATATCCAGCGGTTCCTGCCCAACC CAGCGGGAGTGCAACTGGAGGACCCGGAGTTCCAGGCCTCAAACATTA TGCACTCTATCAACGGCTATGTGTTCGACTCGCTCCAACTGAGCGTGTG CCTGCATGAAGTGGCATACTGGTACATTCTGTCCATCGGAGCCCAGACC GACTTCCTGTCCGTGTTCTTCTCCGGATACACCTTCAAGCATAAGATGGT GTACGAGGACACTCTGACCCTCTTCCCATTTTCCGGAGAAACTGTGTTC ATGTCAATGGAAAACCCGGGCTTGTGGATTCTGGGTTGCCATAACTCGG

	ACTTCCGGAATAGAGGGATGACCGCCCTGCTGAAAGTGTCCAGCTGTGA CAAGAATACCGGCGATTACTACGAGGACAGCTATGAGGACATCTCCGCT TATCTGCTGTCCAAGAACAACGCCATTGAACCCAGGTCCTTCTCCCAAA ACGGTGCACCGACCTCCGAAAGCGCCACCCCAGAGTCAGGACCTGGCT CGGAACCGGCTACCTCGGGCTCAGAGACACCGGGGACTTCCGAGTCCG CAACCCCCGAGAGTGGACCCGGATCCGAACCAGCAACCTCAGGATCAG AAACCCCGGGAACTTCGGAATCCGCCACTCCCGAGTCGGGACCAGGCA CCTCCACTGAGCCTTCCGAGGGAAGCGCCCCCGGATCCCCTGCTGGATC CCCTACCAGCACTGAAGAAGGCACCTCAGAATCCGCGACCCCTGAGTCC GGCCCTGGAAGCGAACCCGCCACCTCCGGTTCCGAAACCCCTGGGACTA GCGAGAGCGCCACTCCGGAATCGGGCCCAGGAAGCCCTGCCGGATCCC CGACCAGCACCGAGGAGGGAAGCCCCGCCGGGTCACCGACTTCCACTG AGGAGGGAGCCTCATCCCCCCCCGTGCTGAAGCGGCATCAAAGAGAGA TCACCAGGACCACTCTCCAGTCCGATCAGGAAGAAATTGACTACGACGA TACTATCAGCGTGGAGATGAAGAAGGAGGACTTCGACATCTACGATGA GGATGAGAACCAGTCCCCTCGGAGCTTTCAGAAGAAAACCCGCCACTA CTTCATCGCTGCCGTGGAGCGGCTGTGGGATTACGGGATGTCCAGCTCA CCGCATGTGCTGCGGAATAGAGCGCAGTCAGGATCGGTGCCCCAGTTCA AGAAGGTCGTGTTCCAAGAGTTCACCGACGGGTCCTTCACTCAACCCCT GTACCGGGGCGAACTCAACGAACACCTGGGACTGCTTGGGCCGTATATC AGGGCAGAAGTGGAAGATAACATCATGGTCACCTTCCGCAACCAGGCC TCCCGGCCGTACAGCTTCTACTCTTCACTGATCTCCTACGAGGAAGATC AGCGGCAGGGAGCCGAGCCCCGGAAGAACTTCGTCAAGCCTAACGAAA CTAAGACCTACTTTTGGAAGGTCCAGCATCACATGGCCCCGACCAAAGA CGAGTTCGACTGTAAAGCCTGGGCCTACTTCTCCGATGTGGACCTGGAG AAGGACGTGCACTCGGGACTCATTGGCCCGCTCCTTGTGTGCCATACTA ATACCCTGAACCCTGCTCACGGTCGCCAAGTCACAGTGCAGGAGTTCGC CCTCTTCTTCACCATCTTCGATGAAACAAAGTCCTGGTACTTTACTGAGA ACATGGAACGCAATTGCAGGGCACCCTGCAACATCCAGATGGAAGATC CCACCTTCAAGGAAAACTACCGGTTTCATGCCATTAACGGCTACATAAT GGACACGTTGCCAGGACTGGTCATGGCCCAGGACCAGAGAATCCGGTG GTATCTGCTCTCCATGGGCTCCAACGAAAACATTCACAGCATTCATTTTT CCGGCCATGTGTTCACCGTCCGGAAGAAGGAAGAGTACAAGATGGCTC TGTACAACCTCTACCCTGGAGTGTTCGAGACTGTGGAAATGCTGCCTAG CAAGGCCGGCATTTGGAGAGTGGAATGCCTGATCGGAGAGCATTTGCA CGCCGGAATGTCCACCCTGTTTCTTGTGTACTCCAACAAGTGCCAGACC CCGCTGGGAATGGCCTCAGGTCATATTAGGGATTTCCAGATCACTGCTT CGGGGCAGTACGGGCAGTGGGCACCTAAGTTGGCCCGGCTGCACTACTC TGGCTCCATCAATGCCTGGTCCACCAAGGAACCCTTCTCCTGGATTAAG GTGGACCTCCTGGCCCCAATGATTATTCACGGTATTAAGACCCAGGGTG CCCGACAGAAGTTCTCCTCACTCTACATCTCGCAATTCATCATAATGTAC AGCCTGGATGGGAAGAAGTGGCAGACCTACCGGGGAAACTCCACTGGA ACGCTCATGGTGTTTTTCGGCAACGTGGACTCCTCCGGCATTAAGCACA ACATCTTCAACCCTCCGATCATTGCTCGGTACATCCGGCTGCACCCAACT CACTACAGCATCCGGTCCACCCTGCGGATGGAACTGATGGGTTGTGACC TGAACTCCTGCTCCATGCCCCTTGGGATGGAATCCAAGGCCATTAGCGA TGCACAGATCACCGCCTCTTCATACTTCACCAACATGTTCGCGACCTGGT CCCCGTCGAAGGCCCGCCTGCACCTCCAAGGTCGCTCCAATGCGTGGCG GCCTCAAGTGAACAACCCCAAGGAGTGGCTCCAGGTCGACTTCCAAAA GACCATGAAGGTCACCGGAGTGACCACCCAGGGCGTGAAGTCCCTGCT GACCTCTATGTACGTTAAGGAGTTCCTCATCTCCTCAAGCCAAGACGGA CATCAGTGGACCCTGTTCTTCCAAAACGGAAAAGTCAAAGTATTCCAGG GCAACCAGGACTCCTTCACCCCTGTGGTCAACAGCCTGGACCCCCCATT GCTGACCCGCTACCTCCGCATCCACCCCCAAAGCTGGGTCCACCAGATC GCACTGCGCATGGAGGTCCTTGGATGCGAAGCCCAAGATCTGTACTAAG CGGCCGCTCATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGAC TGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTT TAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCT CCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTT

	GTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCA CTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCT TTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCG CTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTG TCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTG GATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAG CGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGT CTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCC GCTGCCTAGGCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCC CTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTT CCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTC TATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA AGACAATAGCAGGCATGCTGGGGAAGACCATGGGCGCGCCAGGCCTGT CGACGCCCGGGCGGTACCGCGATCGCTCGCGACGCATAAAG
SEQ ID NO: 36 人α-1-抗胰蛋白酶（A1AT）启动子	ATCGATGGCCCCAGGTTAATTTTTAAAAAGCAGTCAAAAGTCCAAGTGG CCCTTGGCAGCATTTACTCTCTCTGTTTGCTCTGGTTAATAATCTCAGGA GCACAAACATTCCTGGAGGCAGGAGAAGAAATCAACATCCTGGACTTA TCCTCTGGGCCTCTCCCCACCTTCGATGGCCCCAGGTTAATTTTTAAAAA GCAGTCAAAAGTCCAAGTGGCCCTTGGCAGCATTTACTCTCTCTGTTTGC TCTGGTTAATAATCTCAGGAGCACAAACATTCCTGGAGGCAGGAGAAG AAATCAACATCCTGGACTTATCCTCTGGGCCTCTCCCCACCTTCGAACTA GCCACTAGCCTGAGGCTGGTCAAAATTGAACCTCCTCCTGCTCTGAGCA GCCTGGGGGGCAGACTAAGCAGAGGGCTGTGCAGACCCACATAAAGAG CCTACTGTGTGCCAGGCACTTCACCCGAGGCACTTCACAAGCATGCTTG GGAATGAAACTTCCAACTCTTTGGGATGCAGGTGAAACAGTTCCTGGTT CAGAGAGGTGAAGCGGCCTGCCTGAGGCAGCACAGCTCTTCTTTACAGA TGTGCTTCCCCACCTCTACCCTGTCTCACGGCCCCCCATGCCAGCCTGAC GGTTGTGTCTGCCTCAGTCATGCTCCATTTTTCCATCGGGACCATCAAGA GGGTGTTTGTGTCTAAGGCTGACTGGGTAACTTTGGATGAGCGGTCTCT CCGCTCTGAGCCTGTTTCCTCATCTGTCAAATGGGCTCTAACCCACTCTG ATCTCCCAGGGCGGCAGTAAGTCTTCAGCATCAGGCATTTTGGGGTGAC TCAGTAAATGGTAGATCTTGCTACCAGTGGAACAGCCACTAAGGATTCT GCAGTGAGAGCAGAGGGCCAGCTAAGTGGTACTCTCCCAGAGACTGTC TGACTCACGCCACCCCCTCCACCTTGGACACAGGACGCTGTGGTTTCTG AGCCAGGTACAATGACTCCTTTCGGTAAGTGCAGTGGAAGCTGTACACT GCCCAGGCAAAGCGTCCGGGCAGCGTAGGCGGGCGACTCAGATCCCAG CCAGTGGACTTAGCCCCTGTTTGCTCCTCCGATAACTGGGGTGACCTTG GTTAATATTCACCAGCAGCCTCCCCCGTTGCCCCTCTGGATCCACTGCTT AAATACGGACGAGGACAGGGCCCTGTCTCCTCAGCTTCAGGCACCACCA CTGACCTGGGACAG

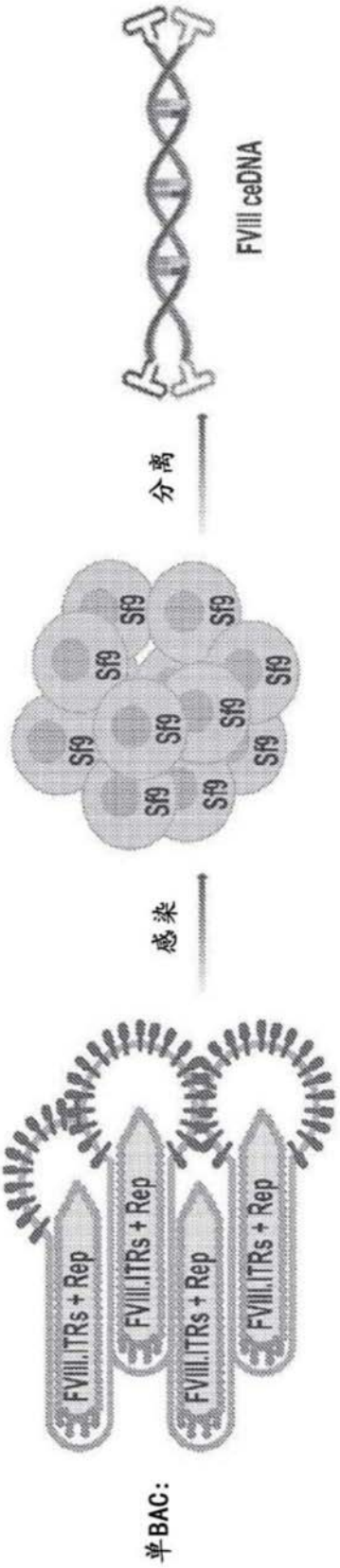

FVIII ceDNA
分离
Sf9
感染
单BAC:
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep

图1A

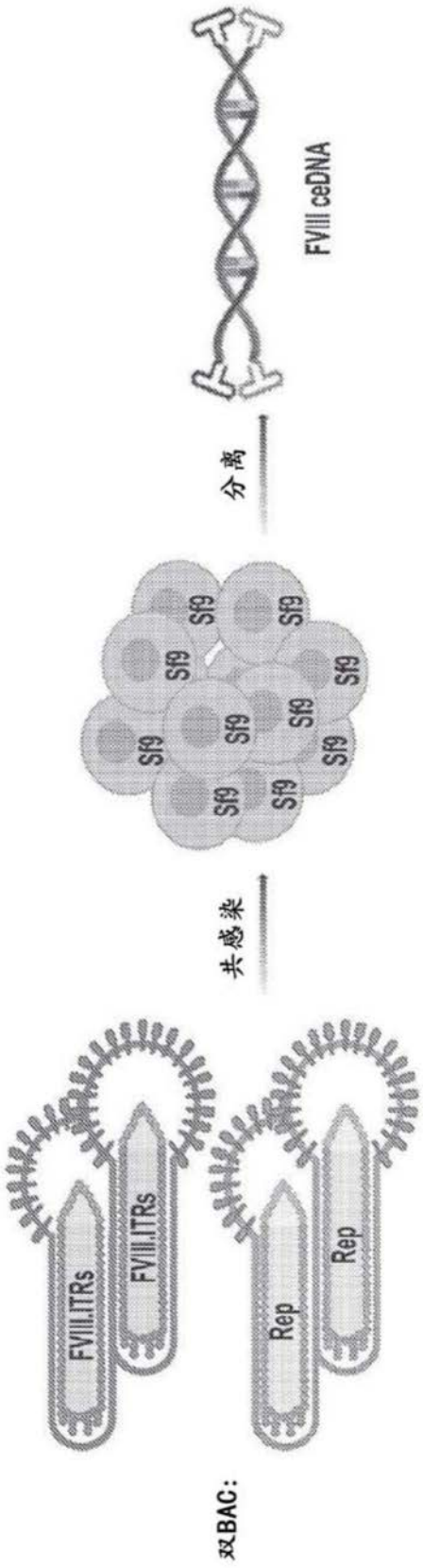


图1B

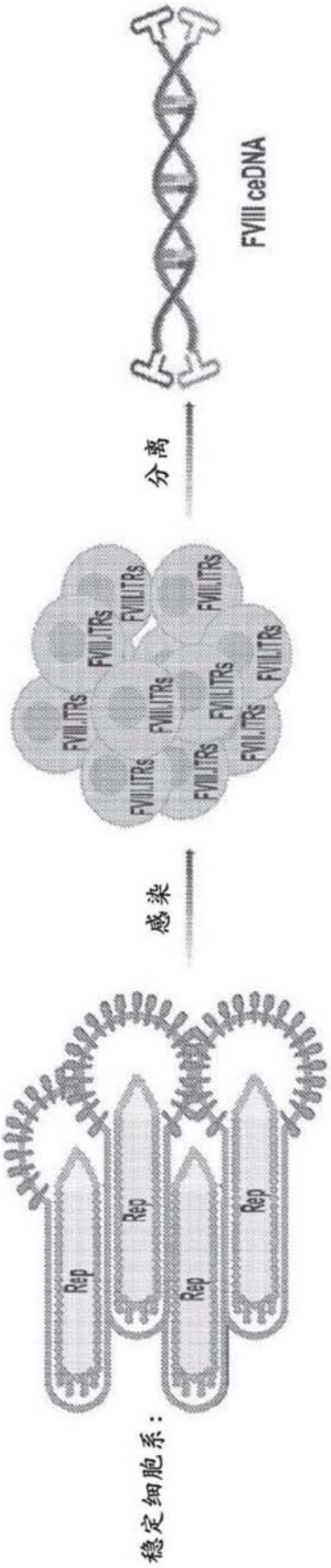


图1C

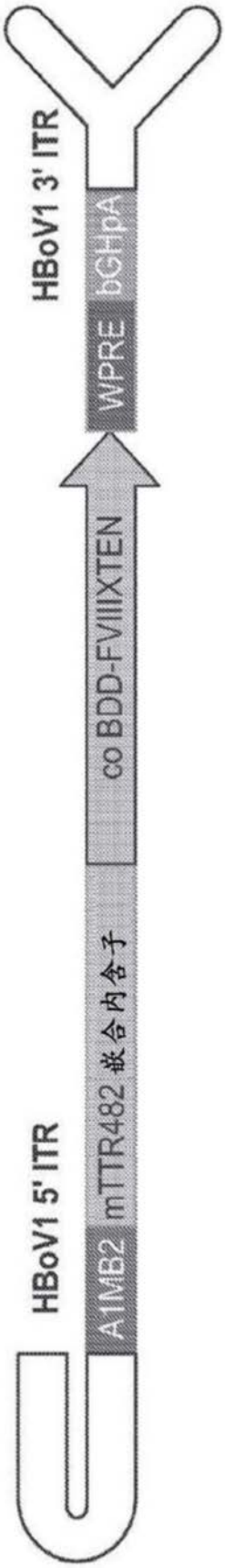

HBoV1 3' ITR
bGHpA
WPRE
co BDD-FVIIIXTEN
嵌合内含子
mTTR482
A1MB2
HBoV1 5' ITR

图2A

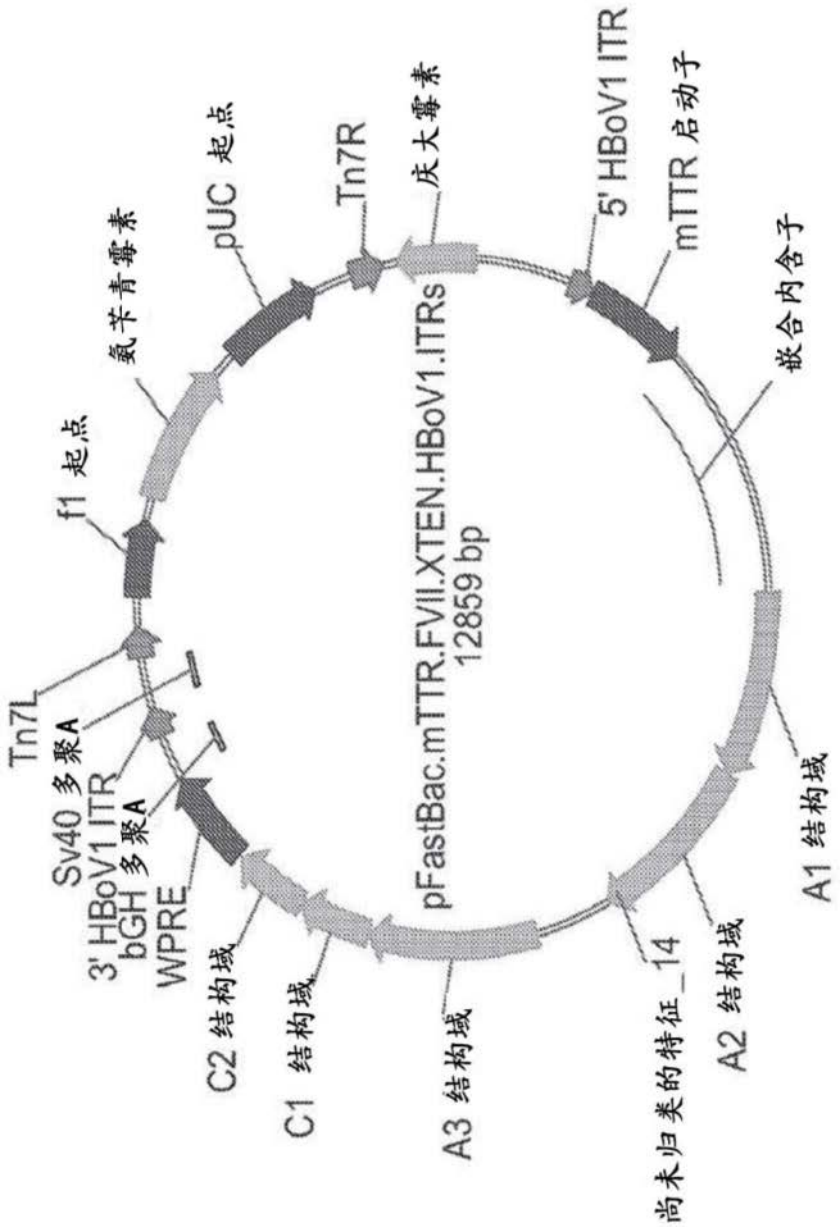

pFastBac.mTTR.FVII.XTEN.HBoV1.ITRs
12859 bp
嵌合内含子
mTTR 启动子
5' HBoV1 ITR
庆大霉素
Tn7R
pUC 起点
氨苄青霉素
f1 起点
Tn7L
Sv40 多聚A
3' HBoV1 ITR
bGH 多聚A
WPRE
C2 结构域
C1 结构域
A3 结构域
尚未归类的特征_14
A2 结构域
A1 结构域

图2B

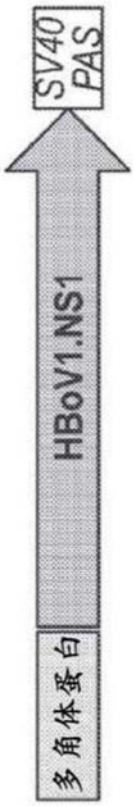


图3A

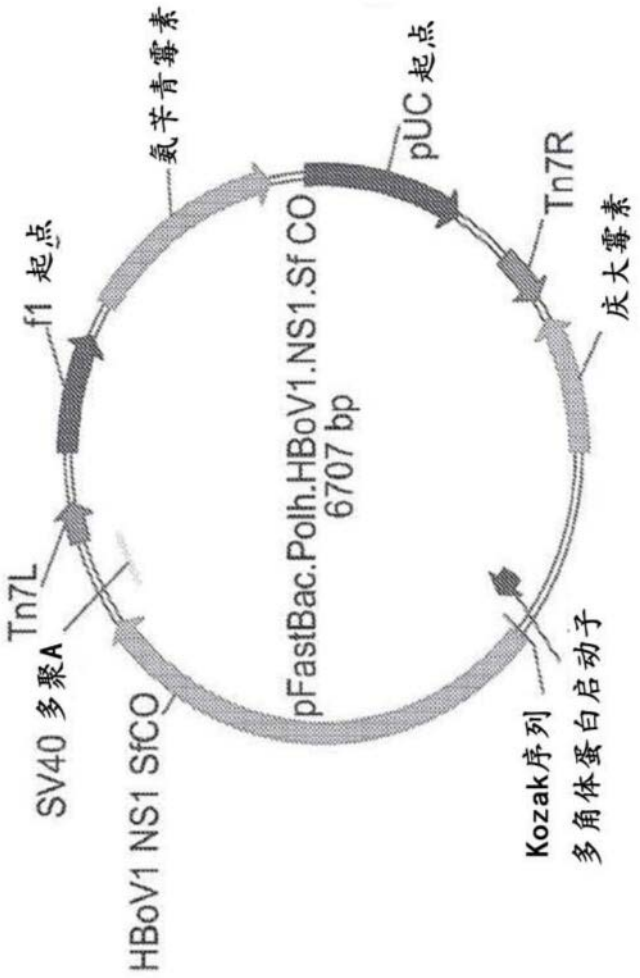

氨苄青霉素
pUC 起点
Tn7R
庆大霉素
Kozak序列
多角体蛋白启动子
pFastBac.Polh.HBoV1.NS1.Sf CO
6707 bp
HBoV1 NS1 SfCO
SV40 多聚A
Tn7L
f1 起点

图3B

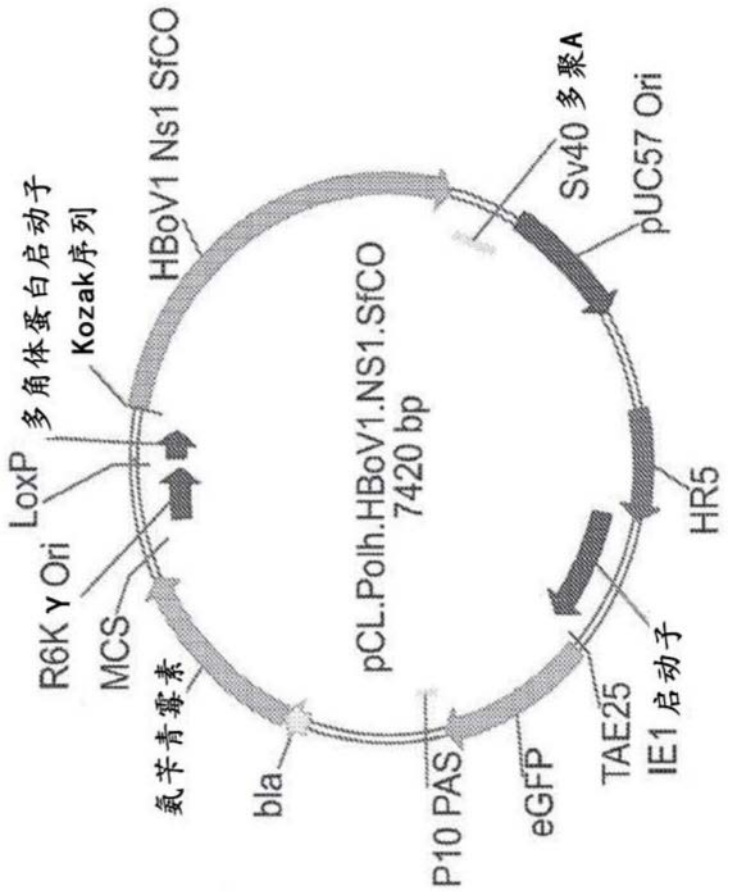
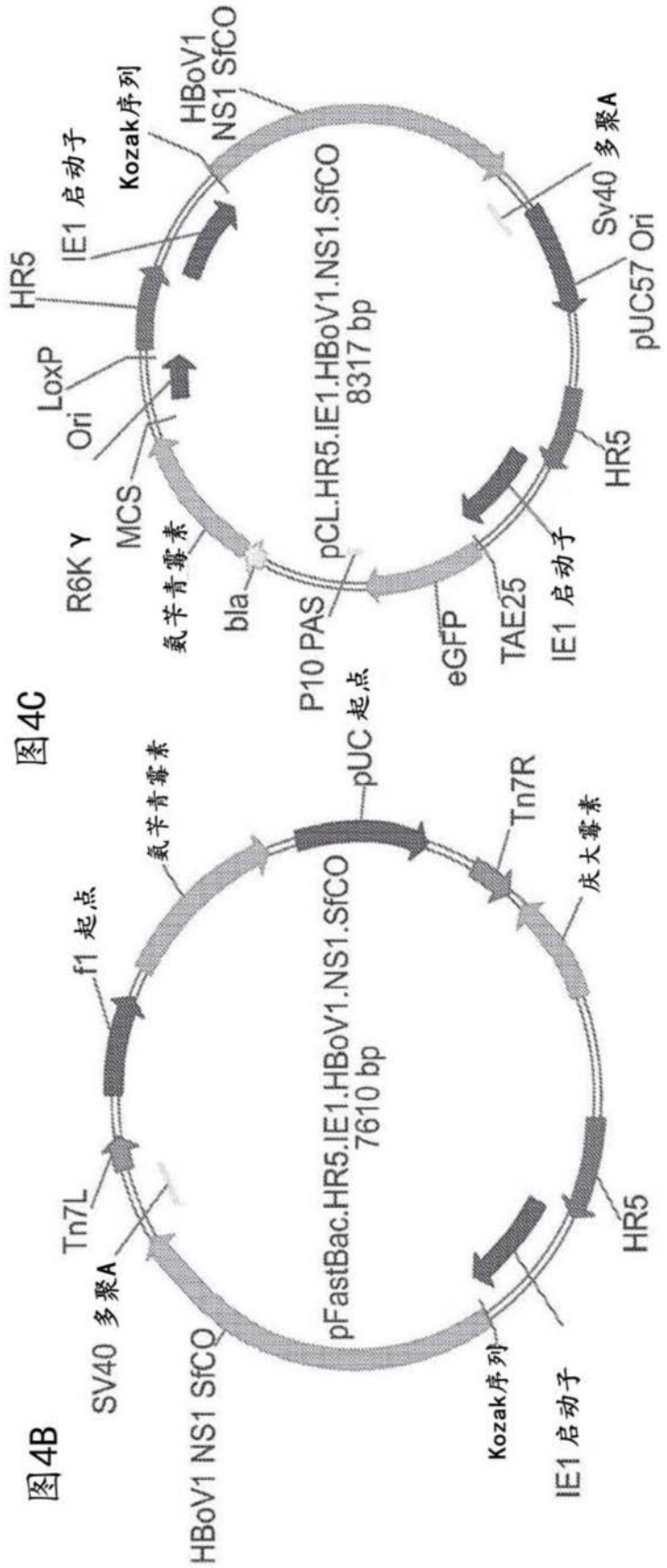

pCL.Polh.HBoV1.NS1.SfCO
7420 bp
HBoV1 Ns1 SfCO
多角体蛋白启动子
Kozak序列
LoxP
R6K γ Ori
MCS
氨苄青霉素
bla
P10 PAS
eGFP
TAE25
IE1 启动子
HR5
Sv40 多聚A
pUC57 Ori

图3C



图4A


图4B
pFastBac.HR5.IE1.HBoV1.NS1.SfCO
7610 bp
HBoV1 NS1 SfCO
SV40 多聚A
Tn7L
f1 起点
氨苄青霉素
pUC 起点
Tn7R
庆大霉素
HR5
IE1 启动子
Kozak序列
图4C
pCL.HR5.IE1.HBoV1.NS1.SfCO
8317 bp
HBoV1 NS1 SfCO
Kozak序列
IE1 启动子
HR5
LoxP
Ori
MCS
R6K γ
氨苄青霉素
bla
P10 PAS
eGFP
TAE25
IE1 启动子
HR5
Sv40 多聚A
pUC57 Ori

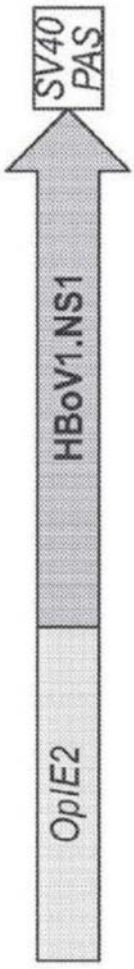
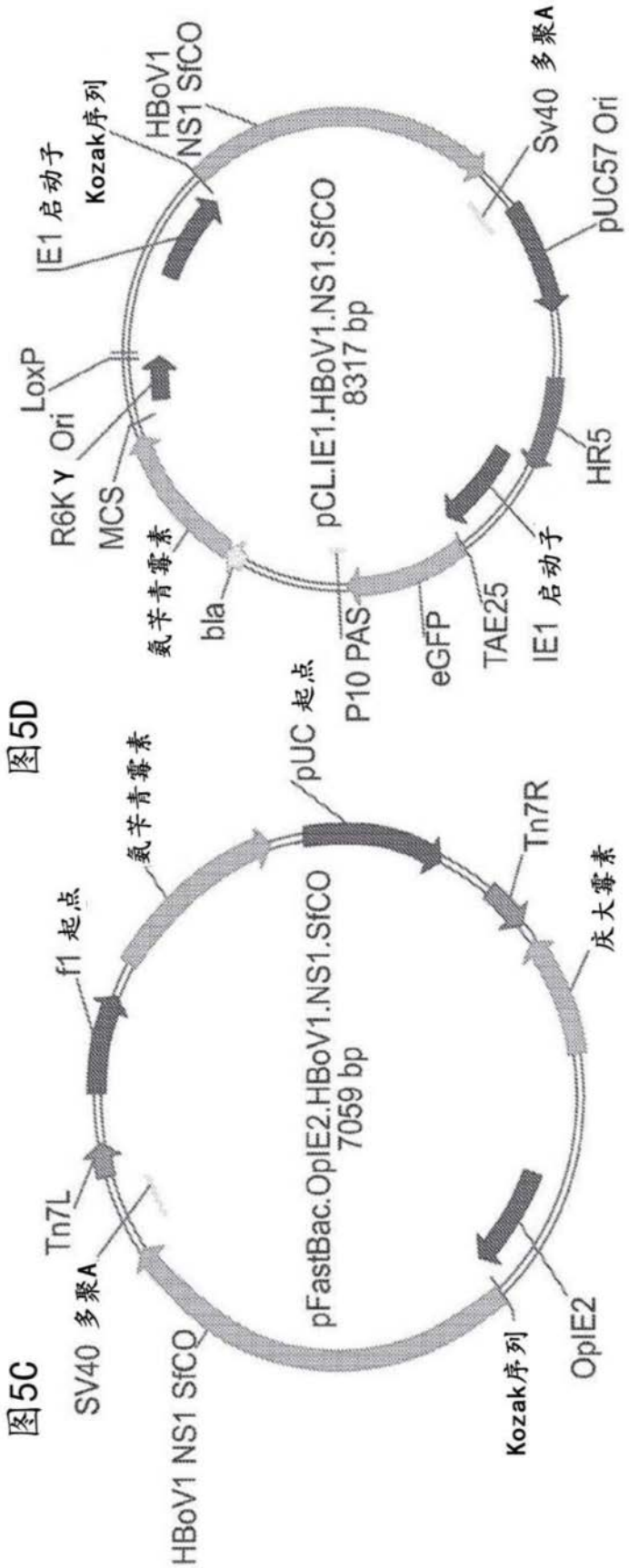


图5A



图5B


图5C
pFastBac.OpIE2.HBoV1.NS1.SfCO
7059 bp
Tn7L
SV40 多聚A
HBoV1 NS1 SfCO
f1 起点
氨苄青霉素
pUC 起点
Tn7R
庆大霉素
Kozak序列
OpIE2
图5D
pCL.IE1.HBoV1.NS1.SfCO
8317 bp
LoxP
R6K γ Ori
MCS
氨苄青霉素
bla
IE1 启动子
Kozak序列
HBoV1 NS1 SfCO
P10 PAS
eGFP
TAE25
IE1 启动子
HR5
Sv40 多聚A
pUC57 Ori

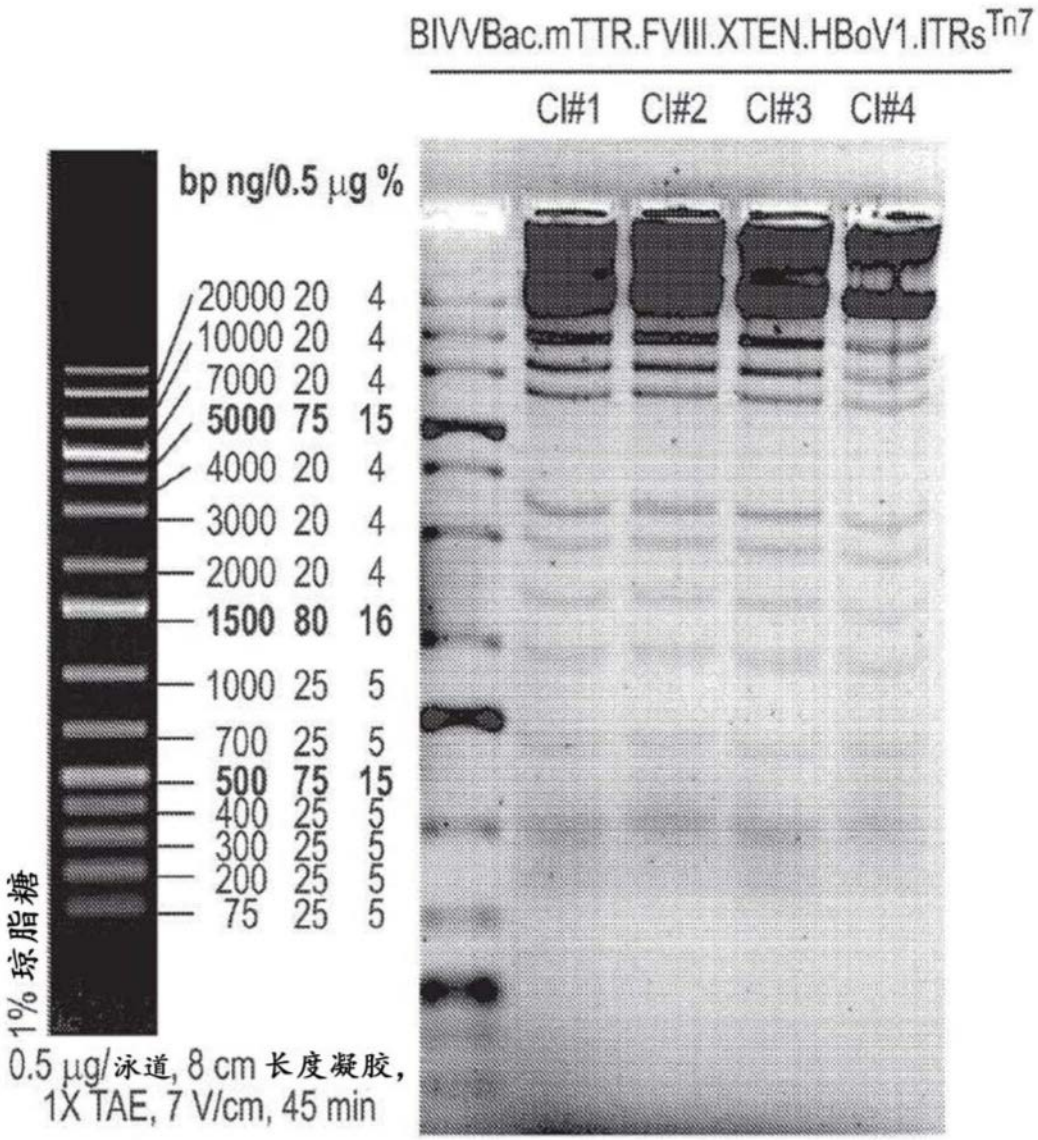

BIVVBac.mTTR.FVIII.XTEN.HBoV1.ITRs^{Tn7}
Cl#1
Cl#2
Cl#3
Cl#4
bp ng/0.5 μg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75 25 5
1% 琼脂糖
0.5 μg/泳道, 8 cm 长度凝胶,
1X TAE, 7 V/cm, 45 min

图6A

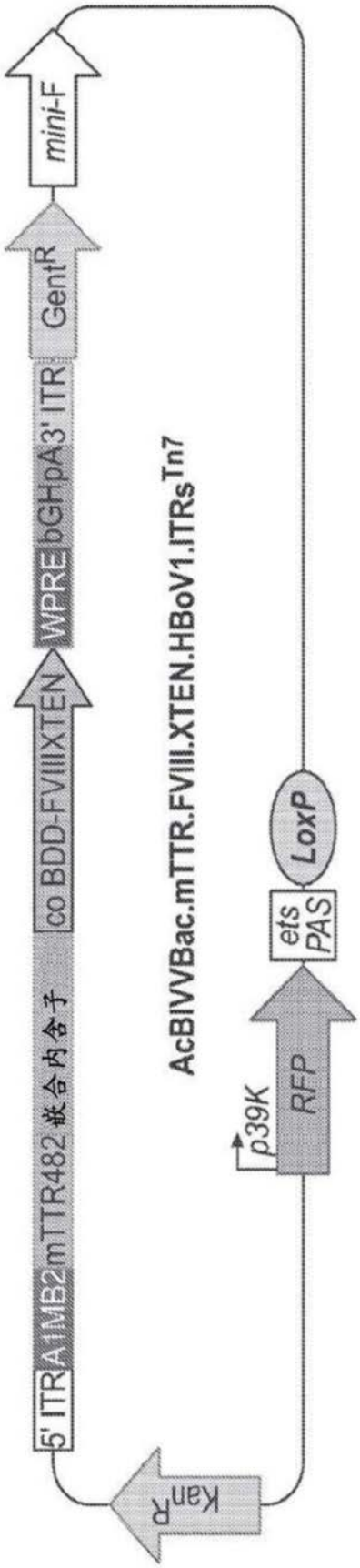

mini-F
GentR
WPRE bGHpA3' ITR
co BDD-FVIIIXTEN
5' ITR A1MB2 mTTR482 嵌合内含子
AcBIVVBac.mTTR.FVIII.XTEN.HBoV1.ITRs Tn7
LoxP
ets PAS
p39K
RFP
KanR

图6B

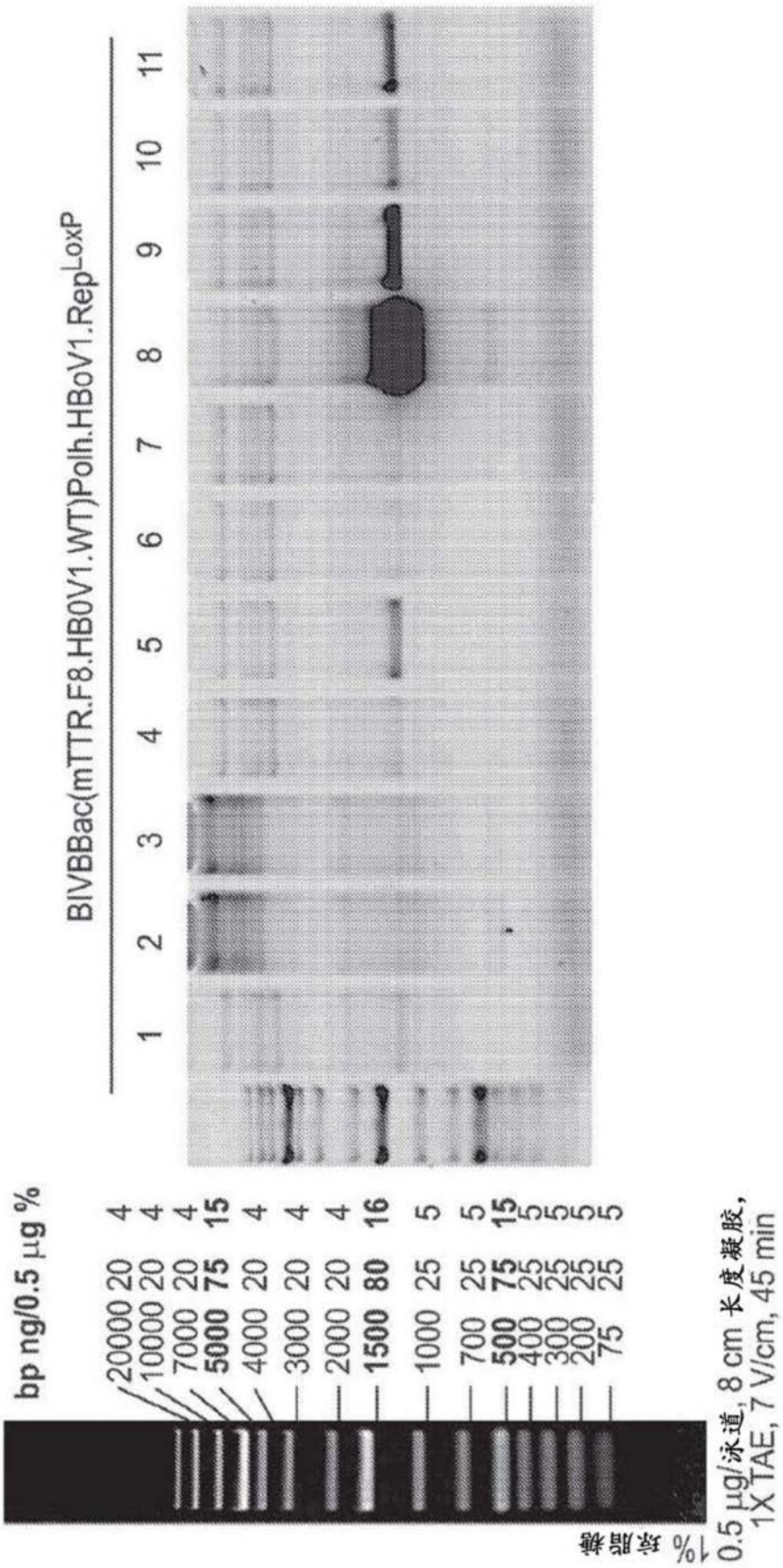

BIVBBac(mTTR.F8.HB0V1.WT)Polh.HB0V1.Rep^LoxP
1 2 3 4 5 6 7 8 9 10 11
bp ng/0.5 µg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75 25 5
1% 琼脂糖
0.5 µg/泳道, 8 cm 长度凝胶,
1X TAE, 7 V/cm, 45 min

图7A

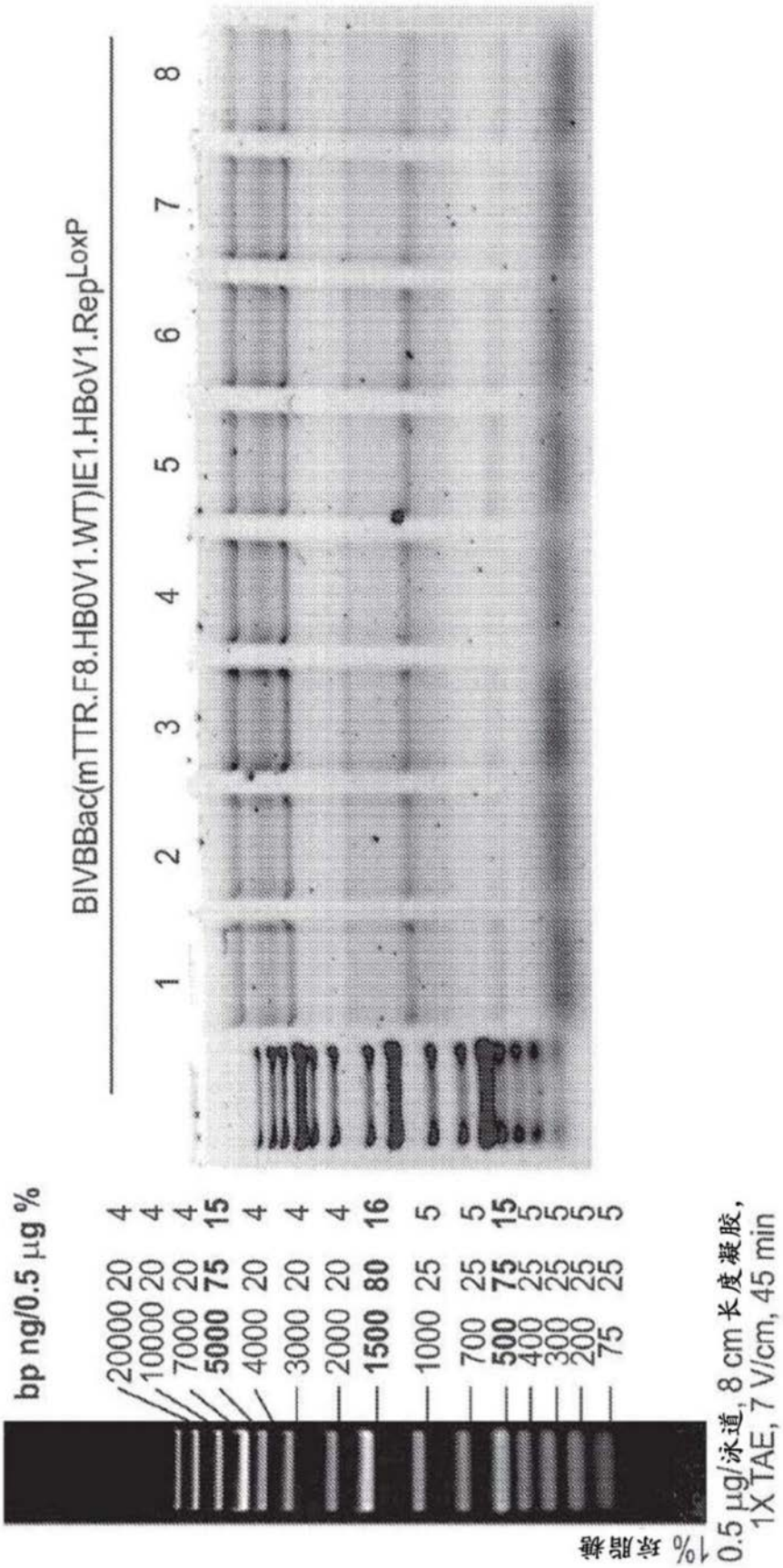

BIVBBac(mTTR.F8.HB0V1.WT)IE1.HBoV1.RepLoxP
1
2
3
4
5
6
7
8
bp ng/0.5 μg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75 25 5
1% 琼脂糖
0.5 μg/泳道, 8 cm 长度凝胶,
1X TAE, 7 V/cm, 45 min

图7B

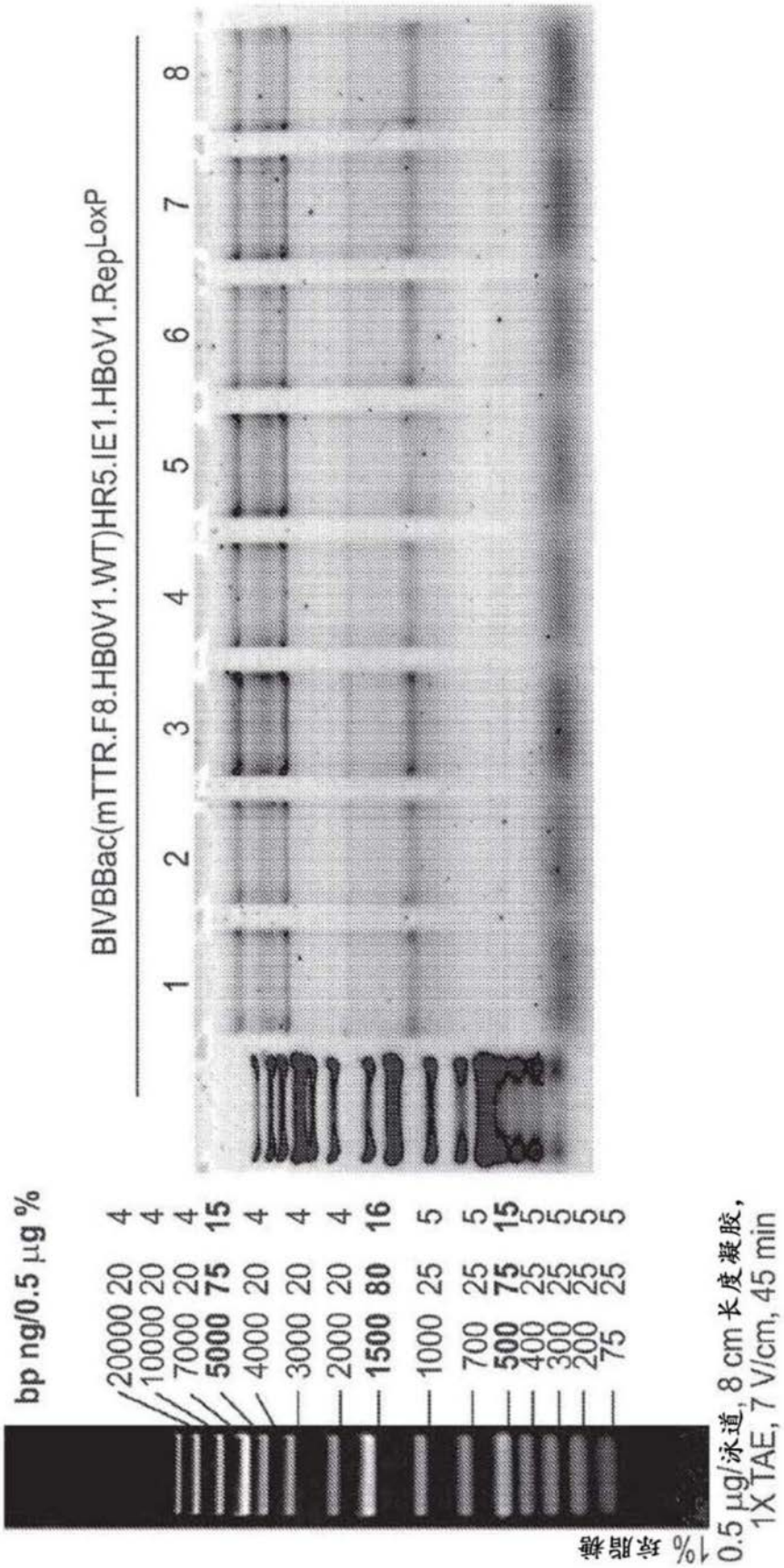

BIVBBac(mTTR.F8.HB0V1.WT)HR5.IE1.HBoV1.RepLoxP
1
2
3
4
5
6
7
8
bp ng/0.5 μg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75
1% 琼脂糖
0.5 μg/泳道, 8 cm 长度凝胶,
1X TAE, 7 V/cm, 45 min

图7C

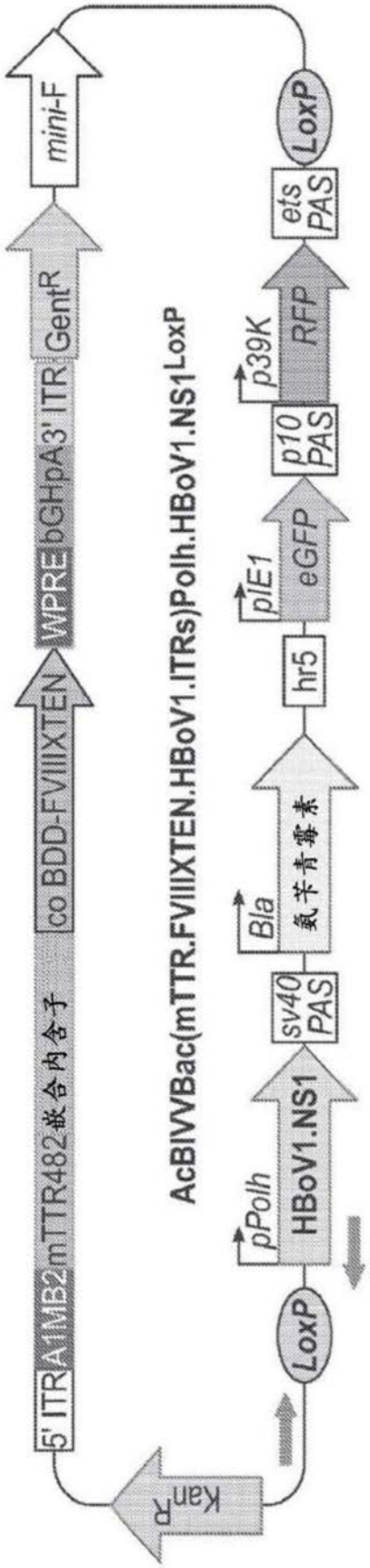

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1LoxP
mini-F
GentR
WPRE
bGHpA3' ITR
co BDD-FVIIIXTEN
5' ITR
A1MB2mTTR482嵌合内含子
LoxP
pPolh
HBoV1.NS1
sv40 PAS
Bla
氯苄青霉素
hr5
pIE1
eGFP
p10 PAS
p39K
RFP
ets PAS
LoxP
KanR

图7D

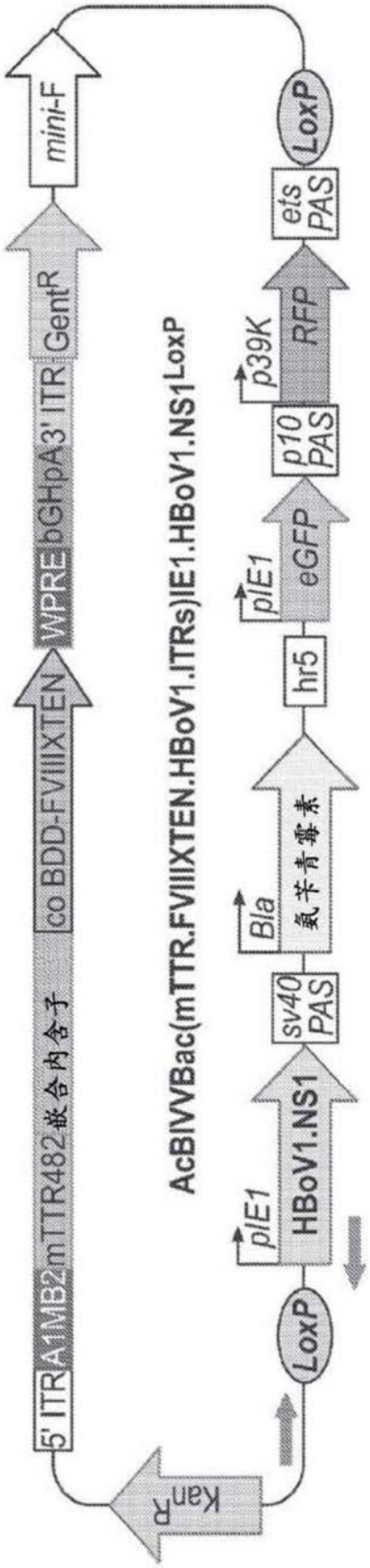

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)IE1.HBoV1.NS1LoxP
mini-F
GentR
3' ITR
bGHpA
WPRE
co BDD-FVIIIXTEN
mTTR482嵌合内含子
B2
A1
5' ITR
LoxP
pIE1
HBoV1.NS1
sv40 PAS
Bla
氯苄青霉素
hr5
pIE1
eGFP
p10 PAS
p39K
RFP
ets PAS
LoxP
KanR

图7E

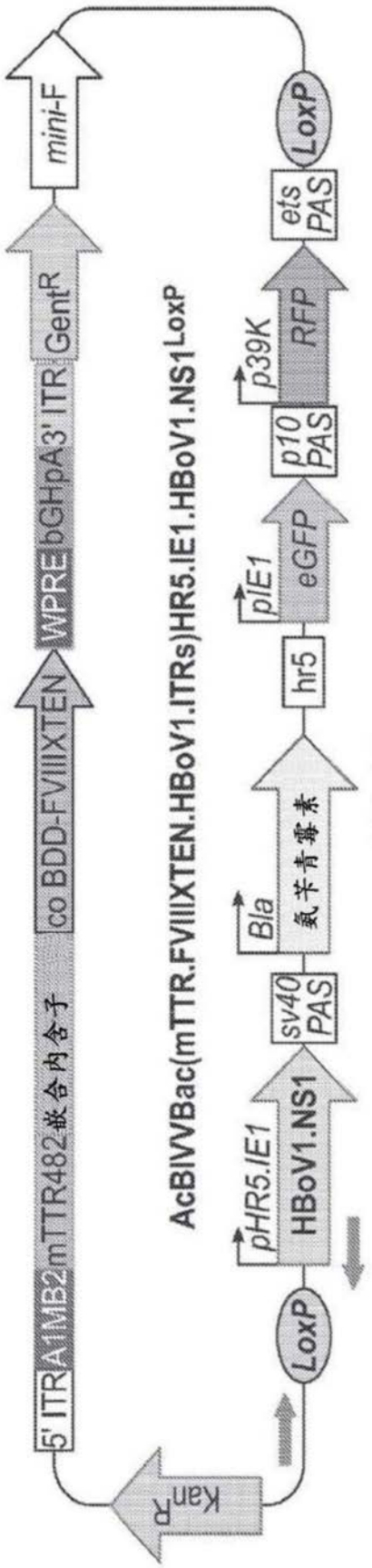

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)HR5.IE1.HBoV1.NS1LoxP
mini-F
GentR
WPRE
bGHpA3' ITR
co BDD-FVIIIXTEN
5' ITR
A1MB2mTTR482嵌合内含子
LoxP
pHR5.IE1
HBoV1.NS1
sv40 PAS
Bla
氨苄青霉素
hr5
pIE1
eGFP
p10 PAS
p39K
RFP
ets PAS
LoxP
KanR

图7F

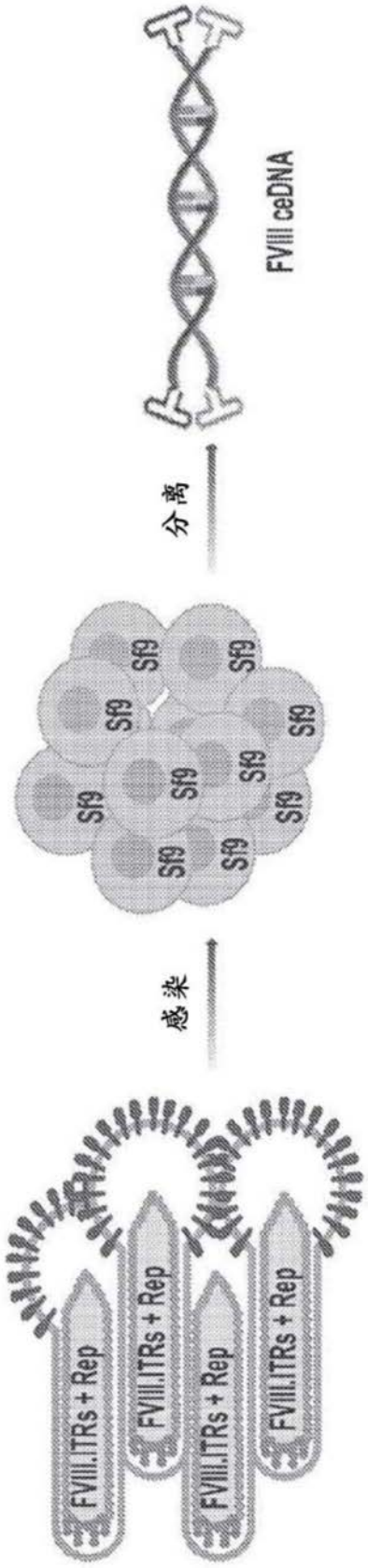

FVIII ceDNA
分离
Sf9
感染
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep

图8A

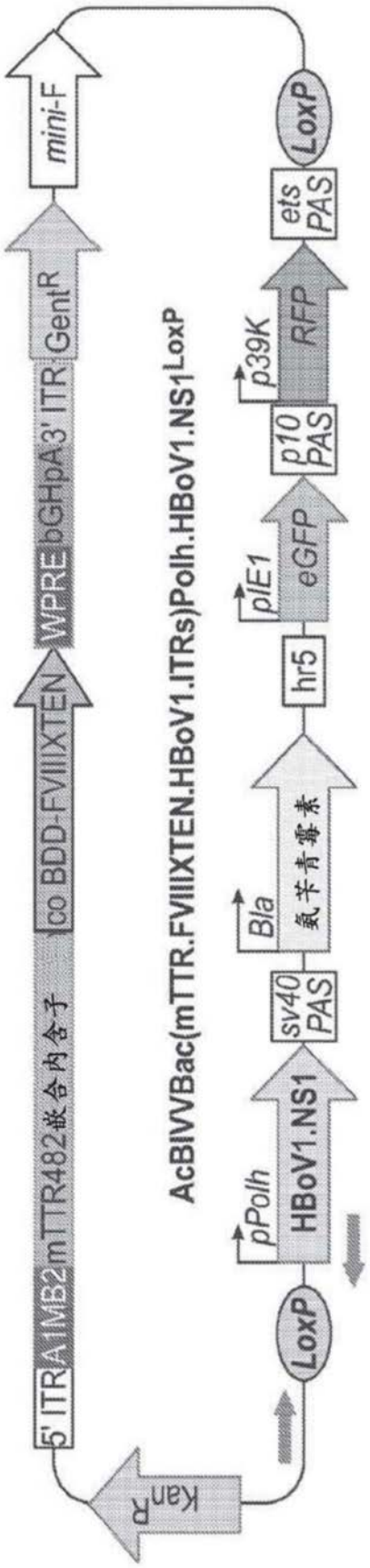

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)Polh.HBoV1.NS1LoxP
KanR
5' ITR
A1MB2
mTTR482嵌合内含子
co BDD-FVIIIXTEN
WPRE
bGHpA
3' ITR
GentR
mini-F
LoxP
pPolh
HBoV1.NS1
sv40 PAS
Bla
氯苄青霉素
hr5
pIE1
eGFP
p10 PAS
p39K
RFP
ets PAS
LoxP

图8B

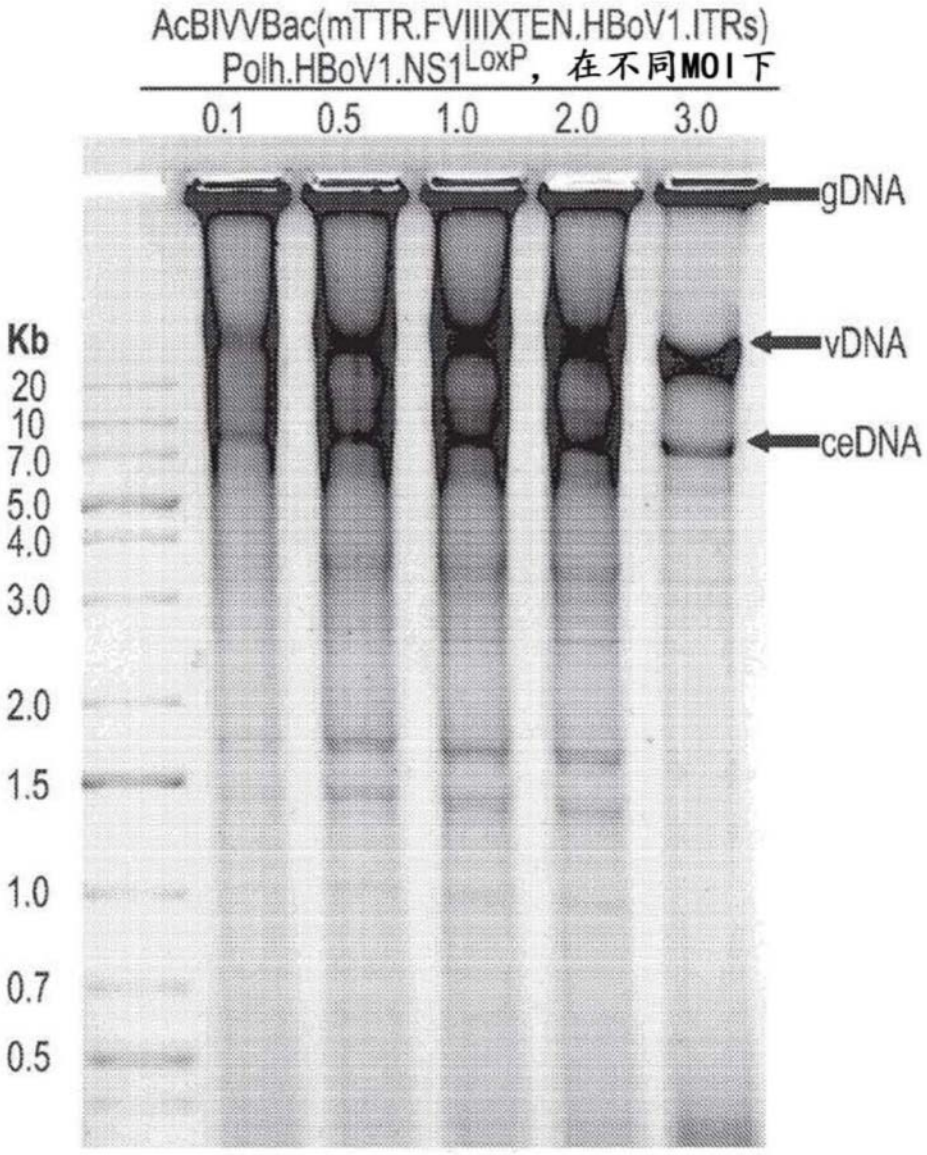

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)
Polh.HBoV1.NS1LoxP，在不同MOI下
0.1
0.5
1.0
2.0
3.0
gDNA
vDNA
ceDNA
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5

图8C

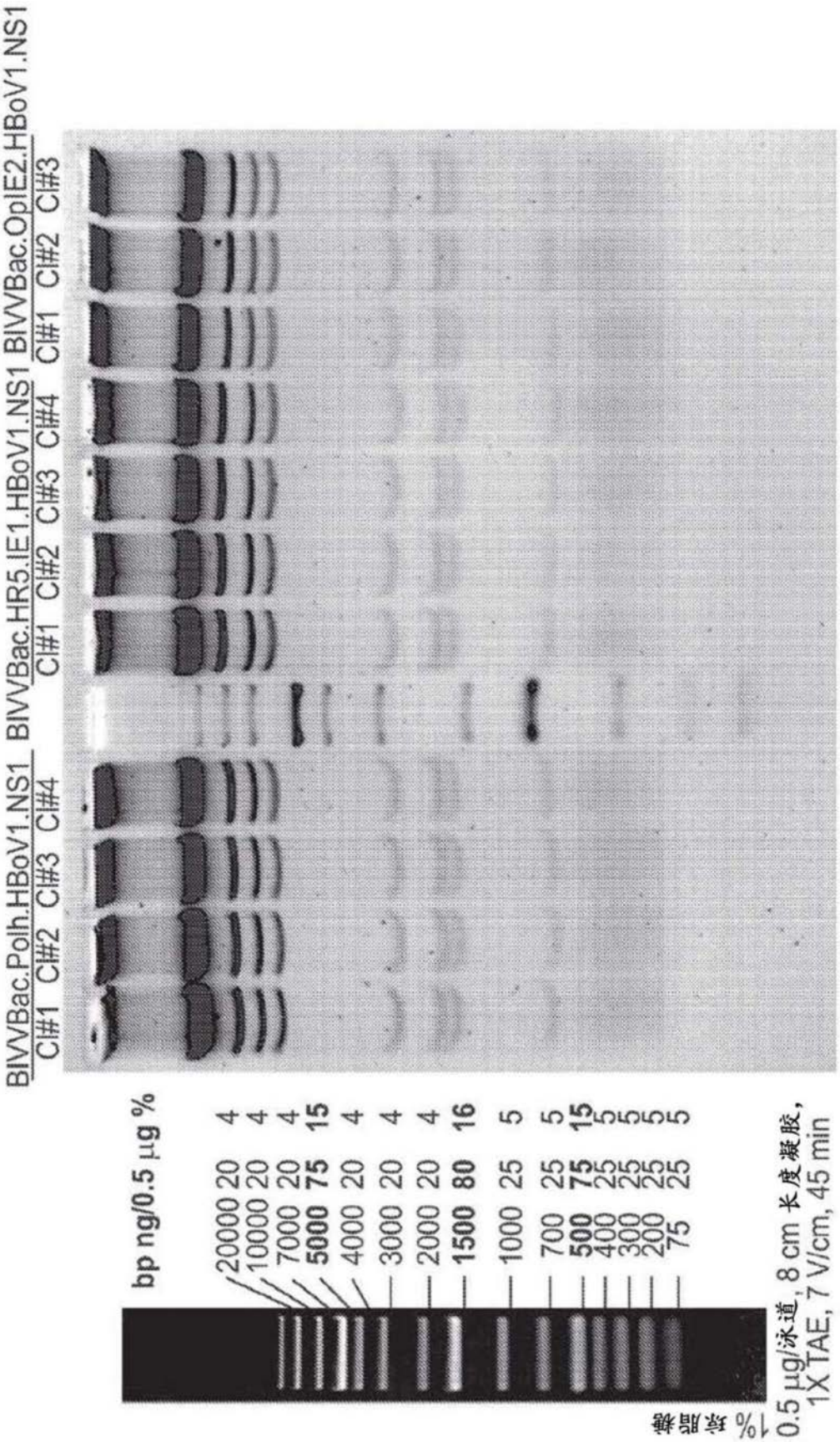

BIVVBac.PolH.HBoV1.NS1
Cl#1 Cl#2 Cl#3 Cl#4
BIVVBac.HR5.IE1.HBoV1.NS1
Cl#1 Cl#2 Cl#3 Cl#4
BIVVBac.OplE2.HBoV1.NS1
Cl#1 Cl#2 Cl#3
bp ng/0.5 μg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75
1% 琼脂糖
0.5 μg/泳道, 8 cm 长度凝胶,
1X TAE, 7 V/cm, 45 min

图9A

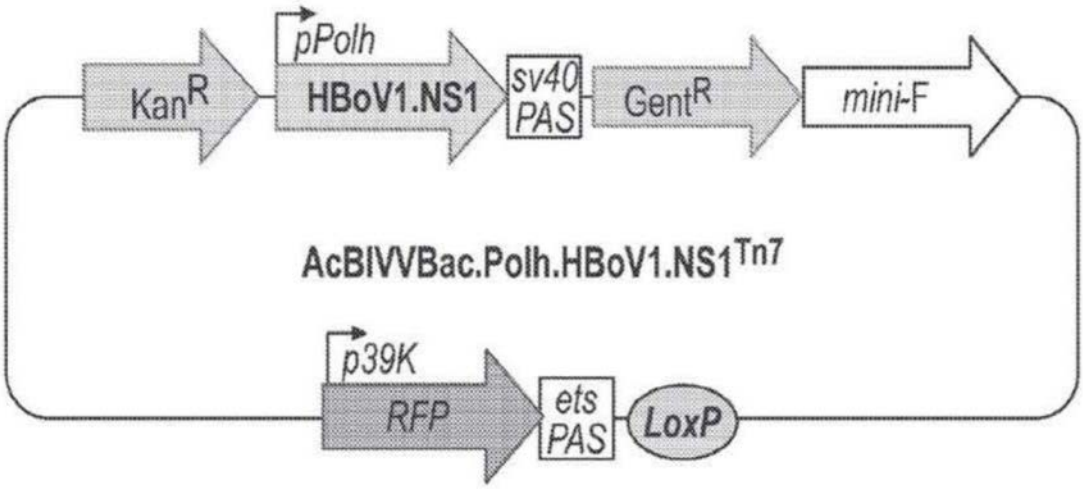

pPolh
Kan^R
HBoV1.NS1
sv40 PAS
Gent^R
mini-F
AcBlVVBac.Polh.HBoV1.NS1^Tn7
p39K
RFP
ets PAS
LoxP

图9B

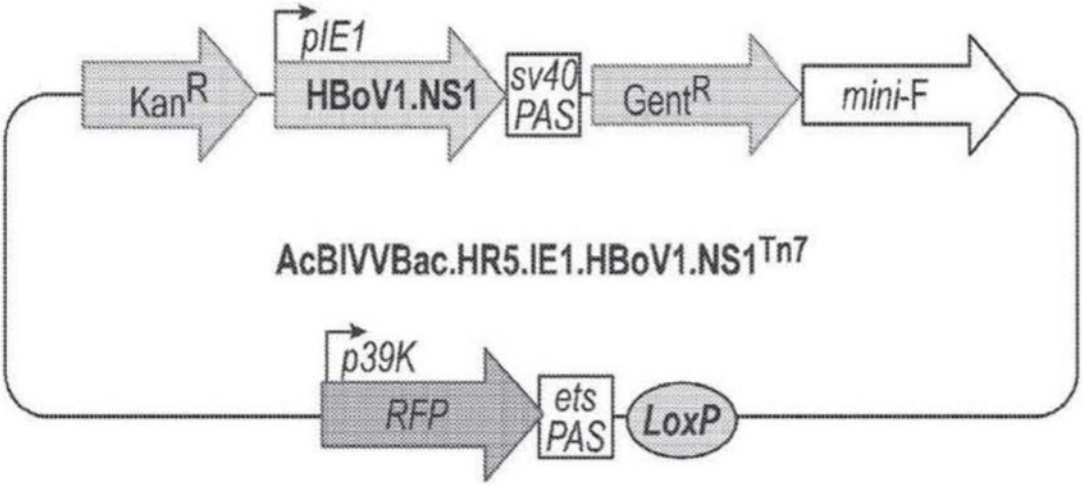

pIE1
Kan^R
HBoV1.NS1
sv40 PAS
Gent^R
mini-F
AcBlVVBac.HR5.IE1.HBoV1.NS1^Tn7
p39K
RFP
ets PAS
LoxP

图9C

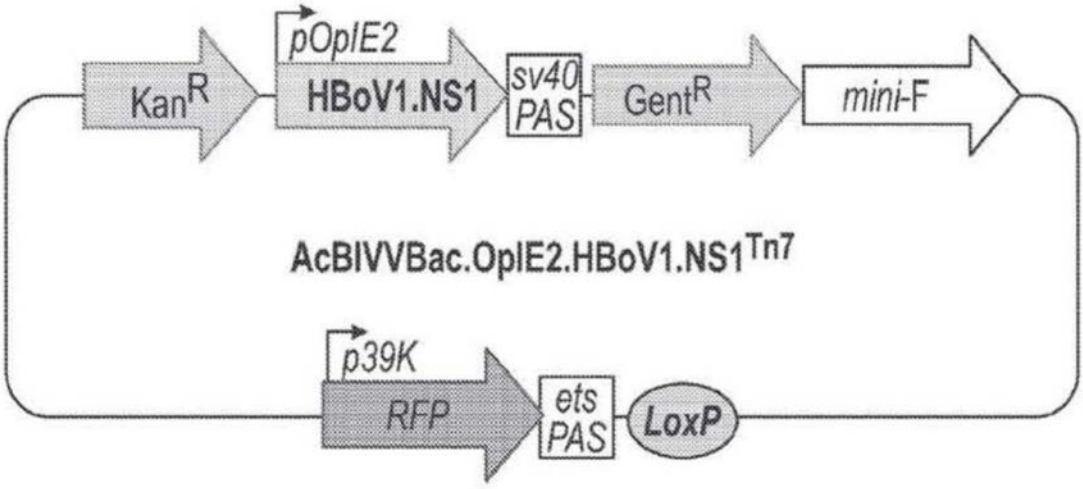

pOplE2
Kan^R
HBoV1.NS1
sv40 PAS
Gent^R
mini-F
AcBlVVBac.OplE2.HBoV1.NS1^Tn7
p39K
RFP
ets PAS
LoxP

图9D

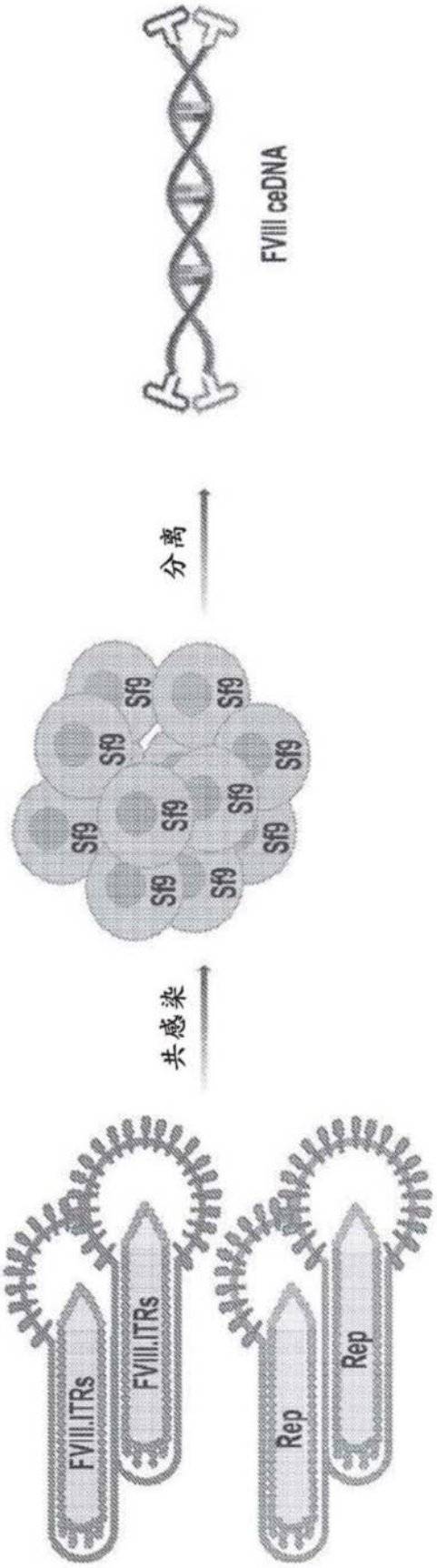


图10A

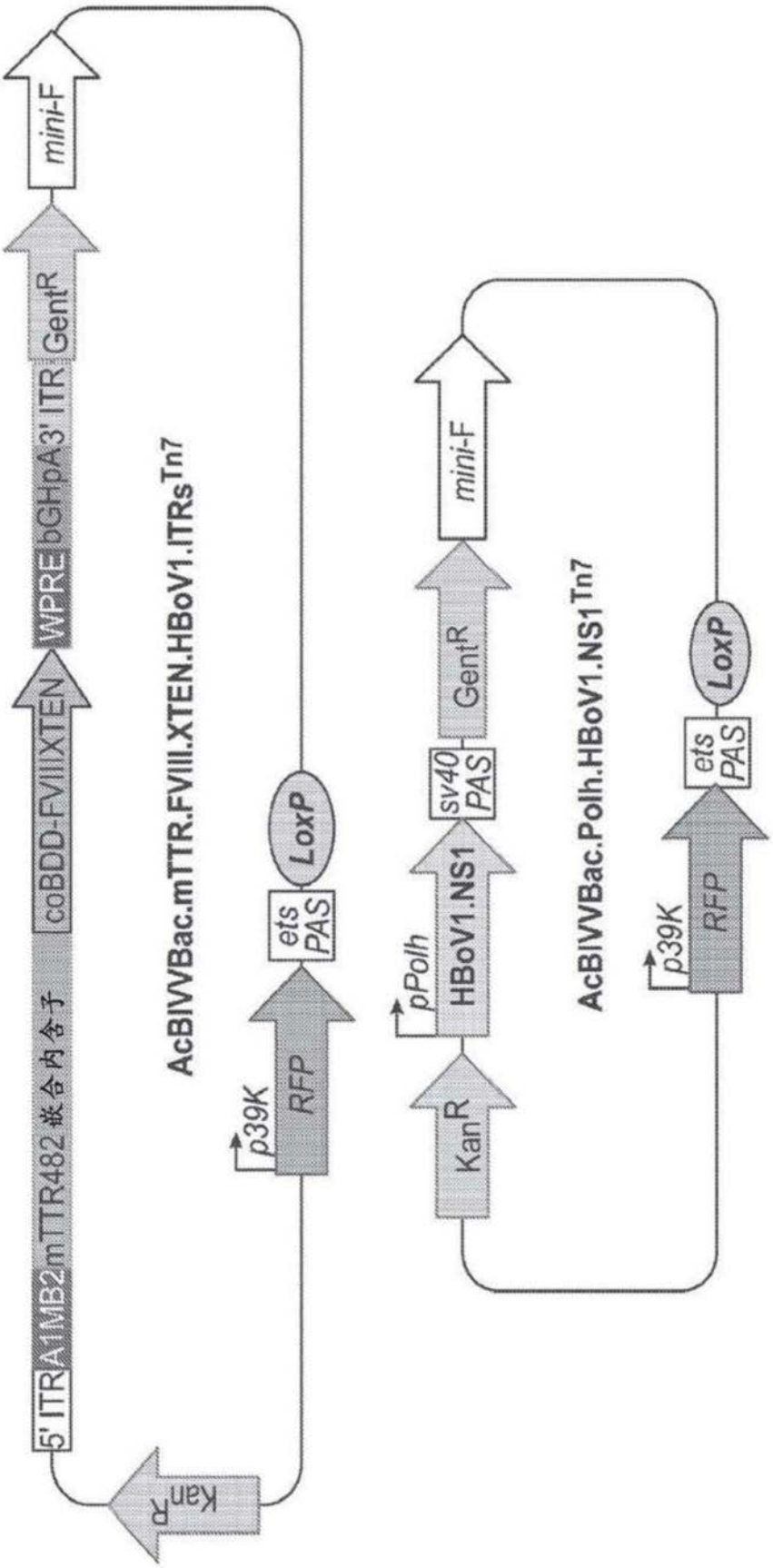

mini-F
GentR
WPRE bGHpA3' ITR
coBDD-FVIIIXTEN
5' ITR A1MB2mTTR482 嵌合内含子
KanR
AcBIVVBac.mTTR.FVIII.XTEN.HBoV1.ITRsTn7
p39K
RFP
ets PAS
LoxP
mini-F
GentR
sv40 PAS
pPolh
HBoV1.NS1
KanR
AcBIVVBac.Polh.HBoV1.NS1Tn7
p39K
RFP
ets PAS
LoxP

图10B

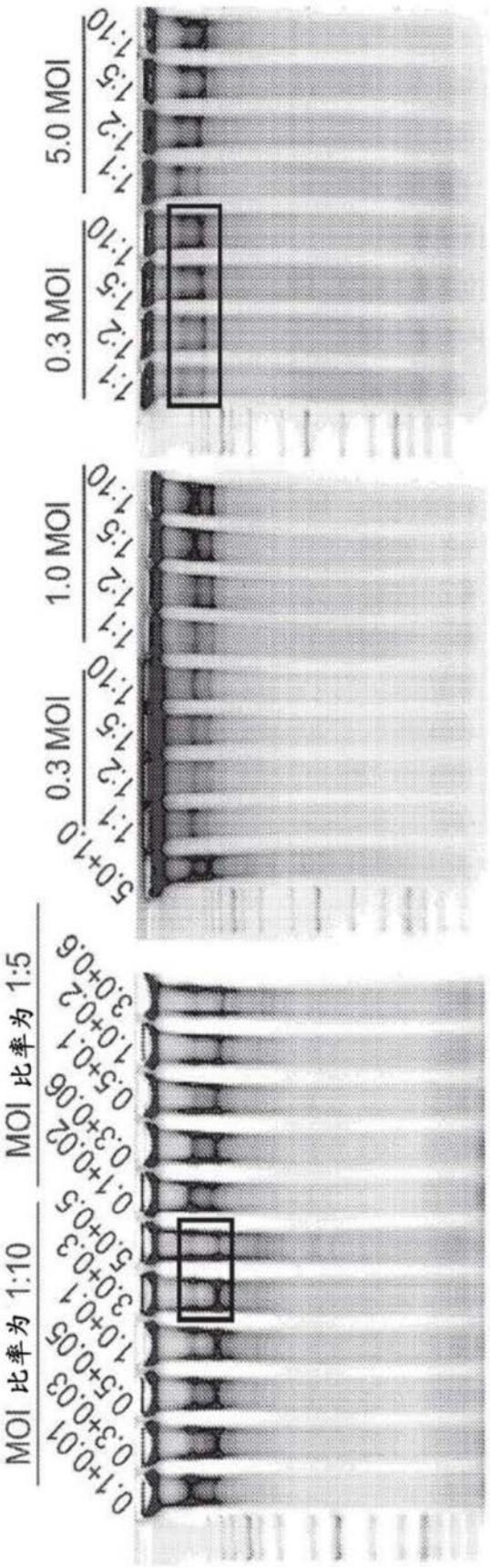


图10C

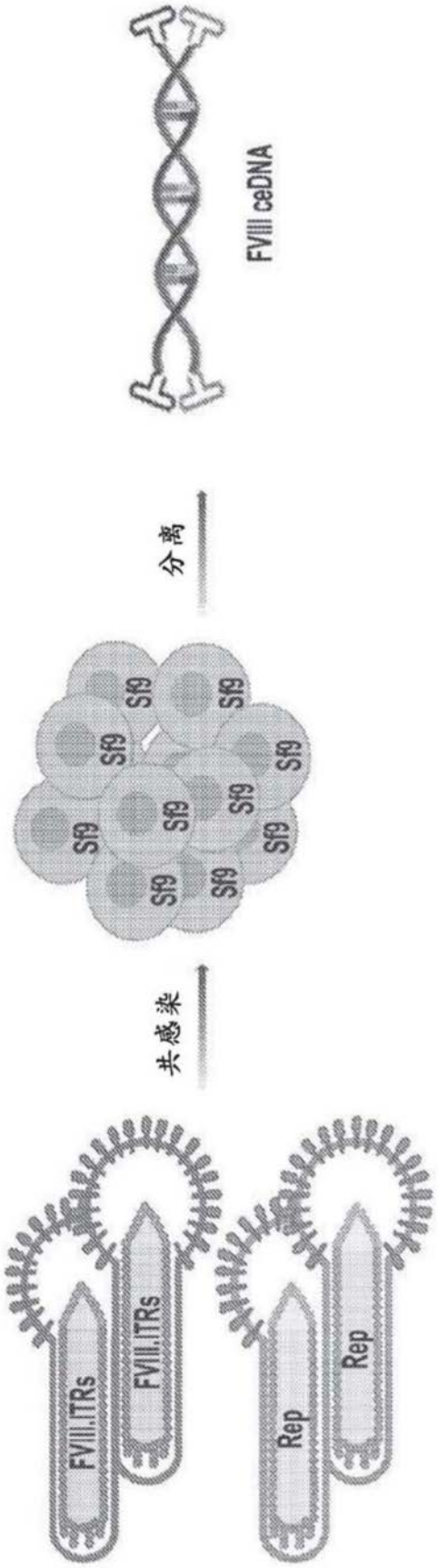


图11A

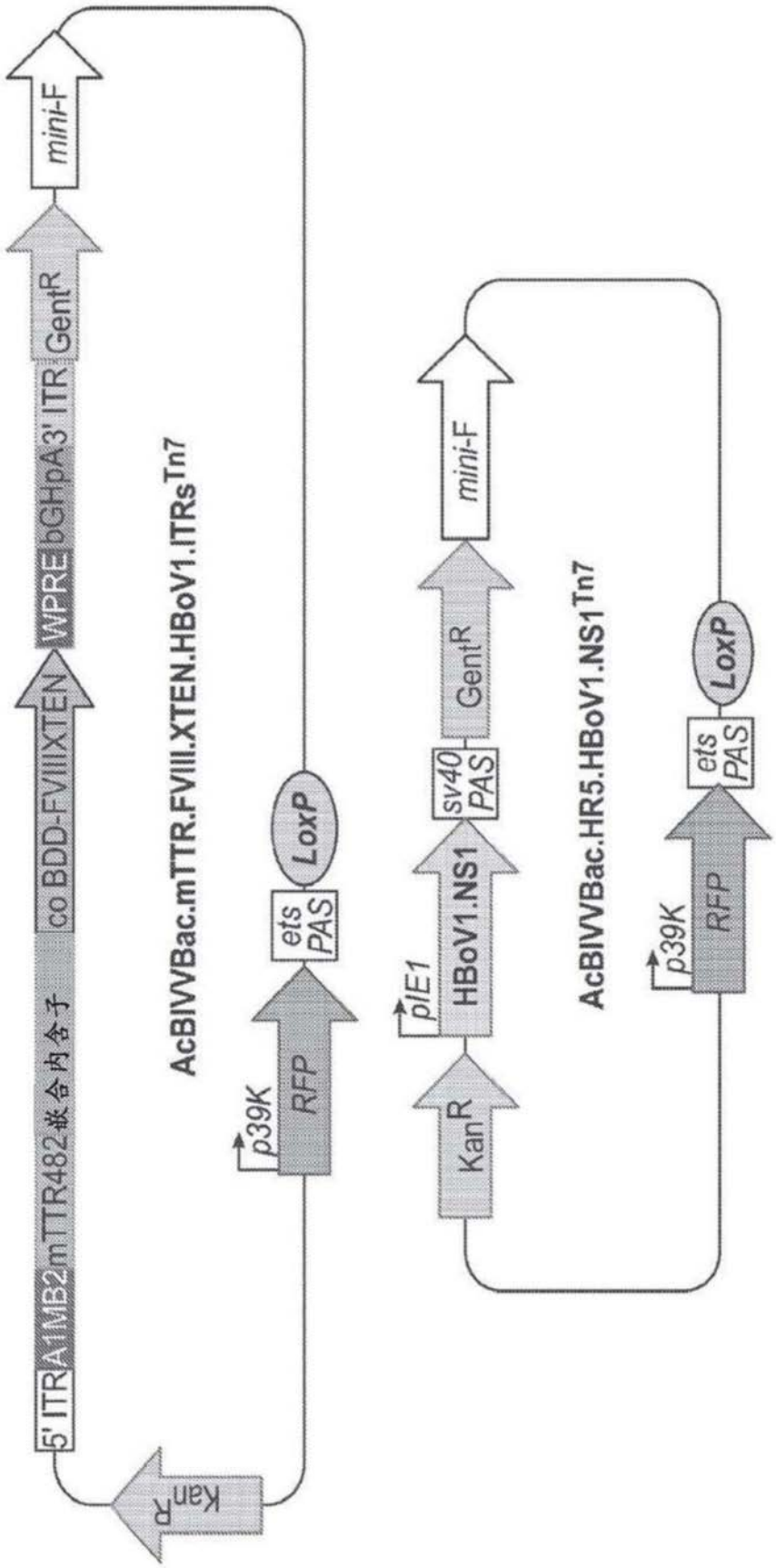

mini-F
GentR
WPRE bGHpA3' ITR
co BDD-FVIIIXTEN
5' ITR A1MB2mTTR482嵌合内含子
KanR
AcBIVVBac.mTTR.FVIII.XTEN.HBoV1.ITRsTn7
p39K
RFP
ets PAS
LoxP
mini-F
GentR
sv40 PAS
pIE1
HBoV1.NS1
KanR
AcBIVVBac.HR5.HBoV1.NS1Tn7
p39K
RFP
ets PAS
LoxP

图11B

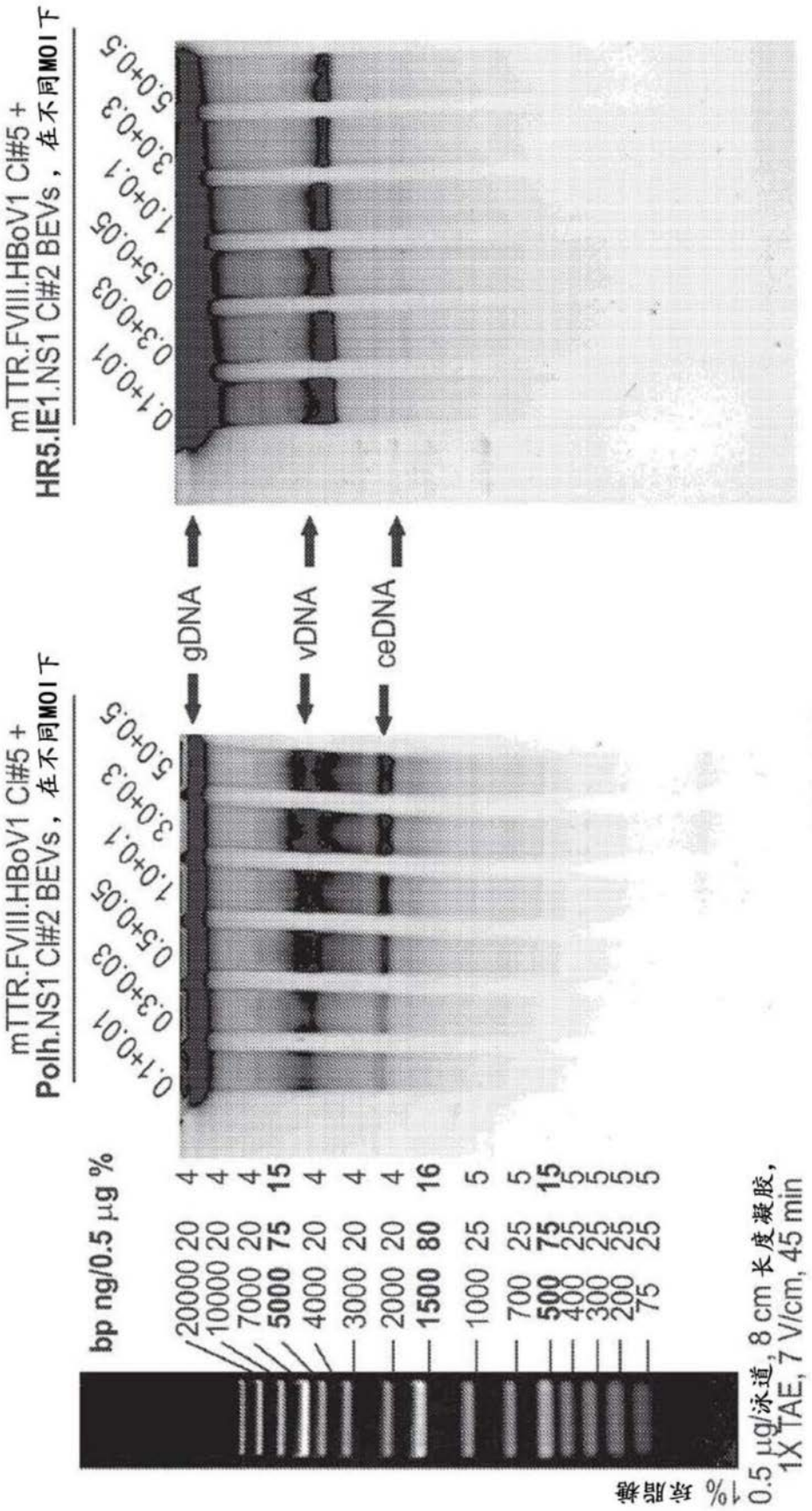

mTTR.FVIII.HBoV1 C#5 +
HR5.IE1.NS1 C#2 BEVs，在不同MOI下
0.1+0.01
0.3+0.03
0.5+0.05
1.0+0.1
3.0+0.3
5.0+0.5
gDNA
vDNA
ceDNA
mTTR.FVIII.HBoV1 C#5 +
PolIh.NS1 C#2 BEVs，在不同MOI下
0.1+0.01
0.3+0.03
0.5+0.05
1.0+0.1
3.0+0.3
5.0+0.5
bp ng/0.5 μg %
20000 20 4
10000 20 4
7000 20 4
5000 75 15
4000 20 4
3000 20 4
2000 20 4
1500 80 16
1000 25 5
700 25 5
500 75 15
400 25 5
300 25 5
200 25 5
75 25
1% 琼脂糖
0.5 μg/泳道，8 cm 长度凝胶，
1X TAE, 7 V/cm, 45 min

图11C

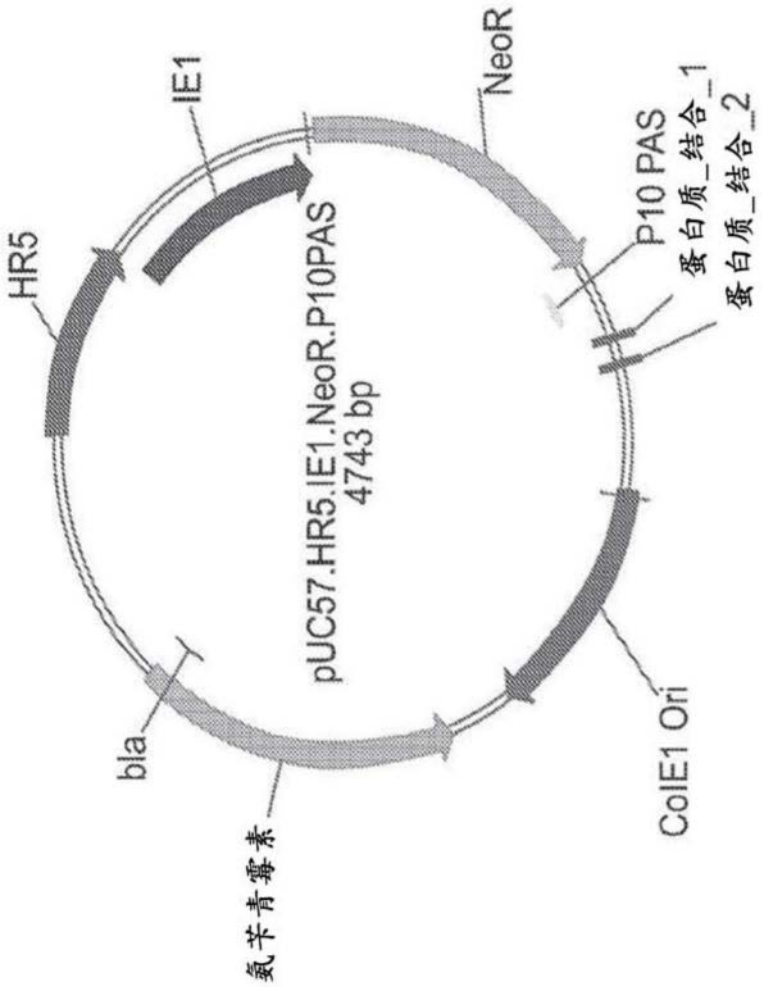

IE1
NeoR
HR5
P10 PAS
蛋白质_结合_1
蛋白质_结合_2
pUC57.HR5.IE1.NeoR.P10PAS
4743 bp
bla
ColE1 Ori
氨苄青霉素

图12A

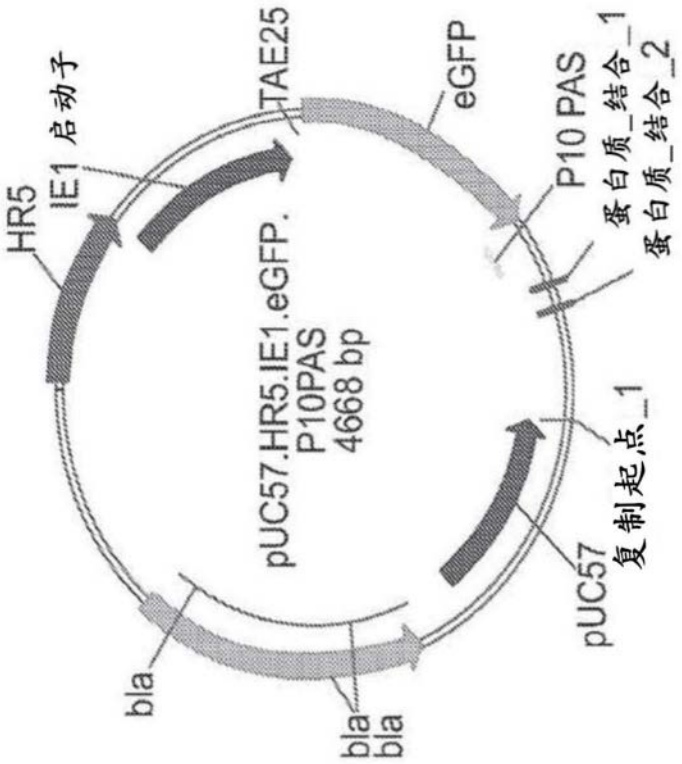

HR5
IE1 启动子
TAE25
eGFP
P10 PAS
蛋白质_结合_1
蛋白质_结合_2
pUC57.HR5.IE1.eGFP.
P10PAS
4668 bp
pUC57
复制起点_1
bla
bla
bla

图12B

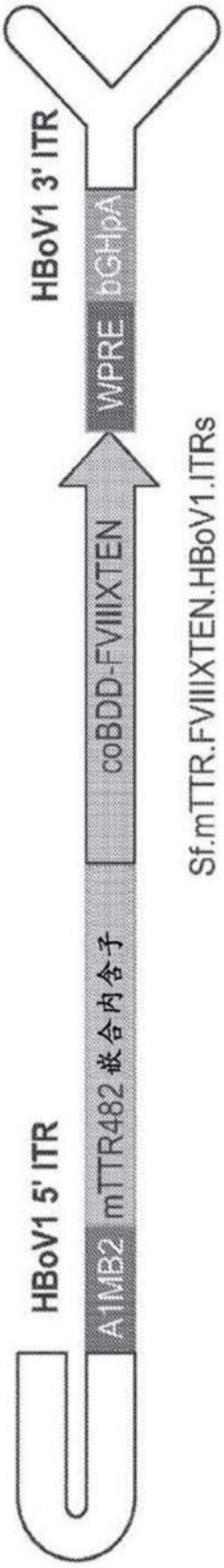
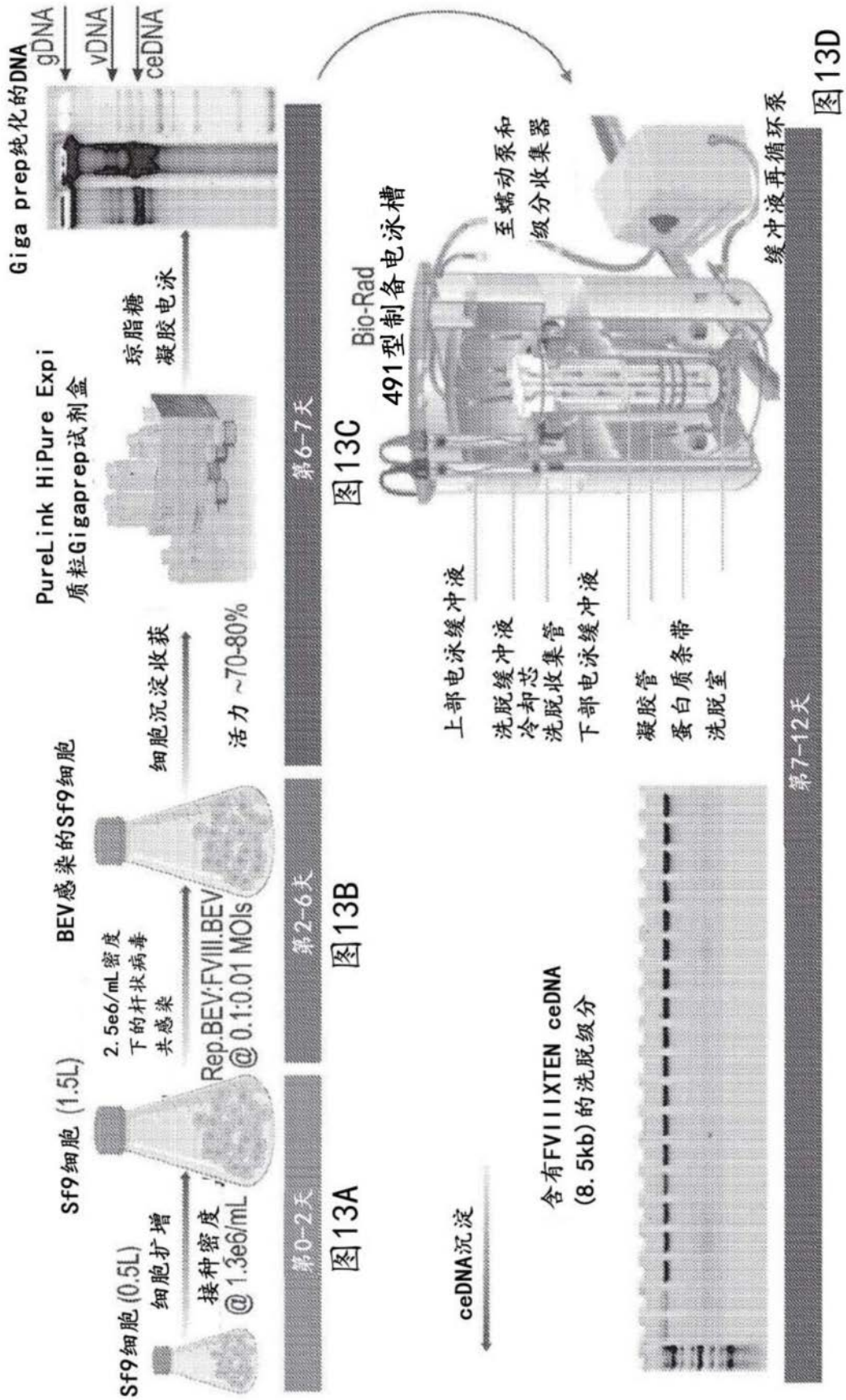

HBoV1 3' ITR
bGHpA
WPRE
coBDD-FVIIIXTEN
Sf.mTTR.FVIIIXTEN.HBoV1.ITRs
嵌合内含子
mTTR482
A1MB2
HBoV1 5' ITR

图12C


Sf9细胞 (0.5L)
细胞扩增
接种密度
@1.3e6/mL
Sf9细胞 (1.5L)
BEV感染的Sf9细胞
2.5e6/mL密度
下的杆状病毒
共感染
Rep.BEV:FVIII.BEV
@0.1:0.01 MOIs
细胞沉淀收获
活力 ~70-80%
PureLink HiPure Expi
质粒Gigaprep试剂盒
琼脂糖
凝胶电泳
Giga prep纯化的DNA
gDNA
vDNA
ceDNA
第0-2天
第2-6天
第6-7天
第7-12天
图13A
图13B
图13C
图13D
Bio-Rad
491型制备电泳槽
至蠕动泵和
级分收集器
上部电泳缓冲液
洗脱缓冲液
冷却芯
洗脱收集管
下部电泳缓冲液
凝胶管
蛋白质条带
洗脱室
缓冲液再循环泵
ceDNA沉淀
含有FVIIIXTEN ceDNA
(8.5kb)的洗脱级分

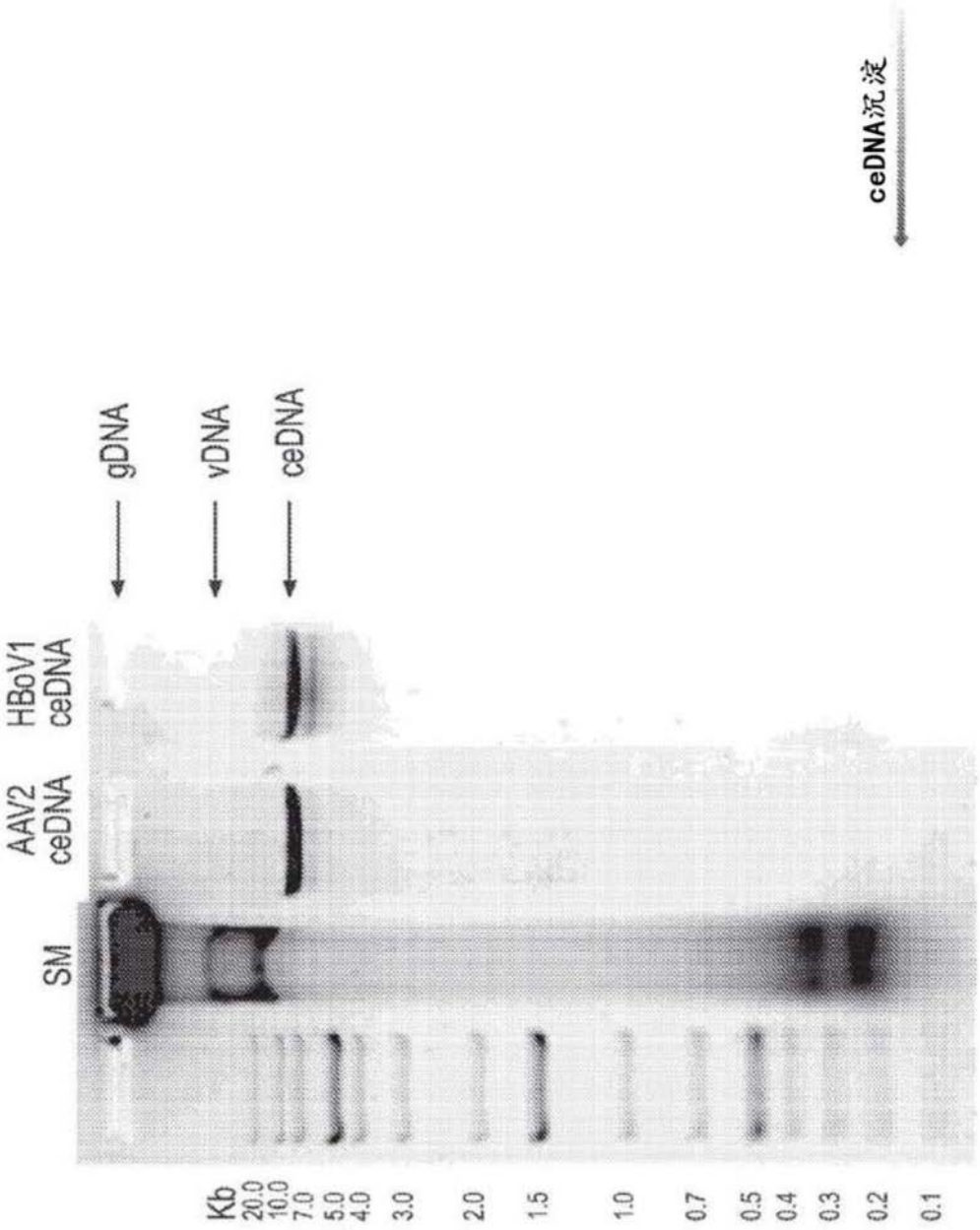

ceDNA沉淀
gDNA
vDNA
ceDNA
HBoV1 ceDNA
AAV2 ceDNA
SM
Kb
20.0
10.0
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5
0.4
0.3
0.2
0.1

图13E

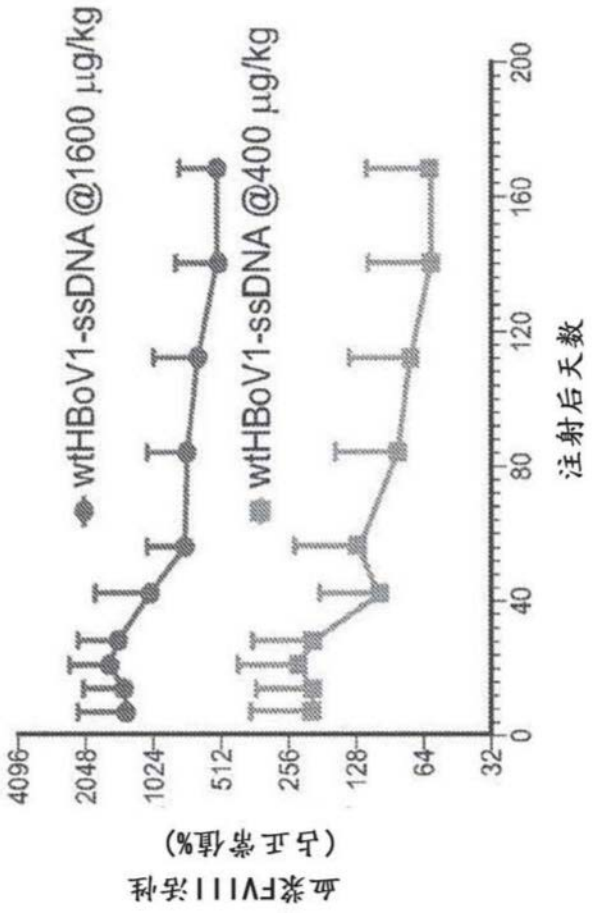

wtHBoV1-ssDNA @1600 μg/kg
wtHBoV1-ssDNA @400 μg/kg
血浆FVIII活性
(占正常值%)
4096
2048
1024
512
256
128
64
32
0
40
80
120
160
200
注射后天数

图14A

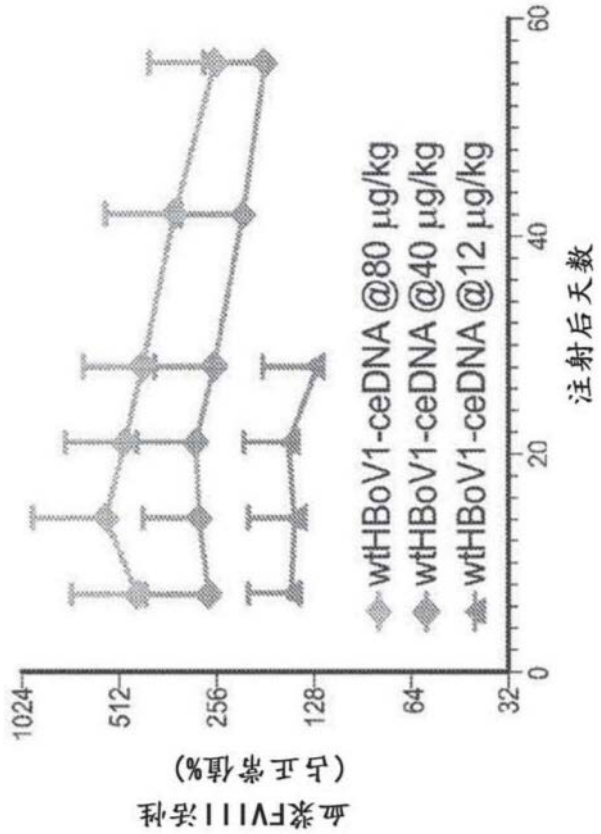

血浆FVIII活性
(占正常值%)
1024
512
256
128
64
32
0
20
40
60
注射后天数
wtHBoV1-ceDNA @80 μg/kg
wtHBoV1-ceDNA @40 μg/kg
wtHBoV1-ceDNA @12 μg/kg

图14B

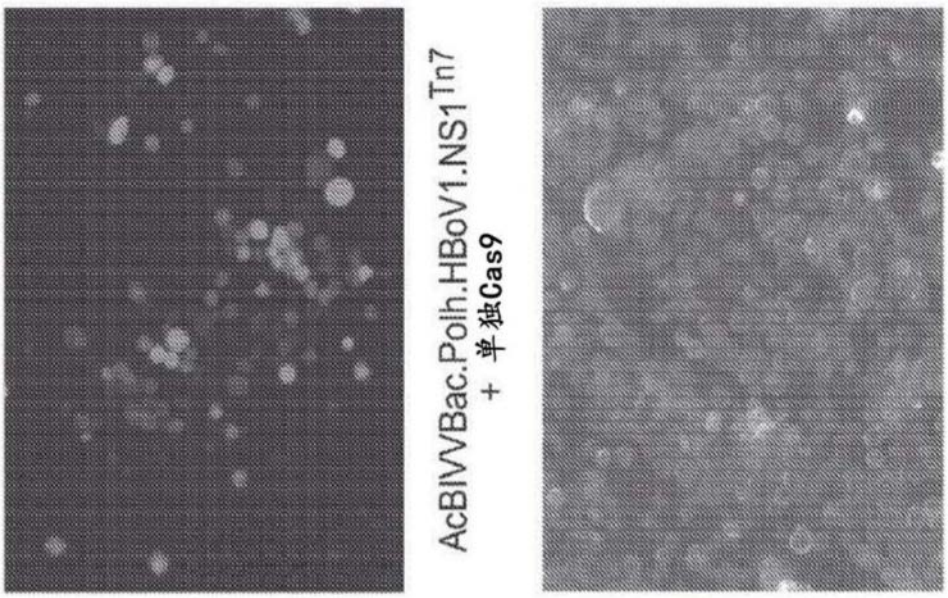

AcBIVVBac.Polh.HBoV1.NS1 Tn7
+ 单独Cas9

图15A

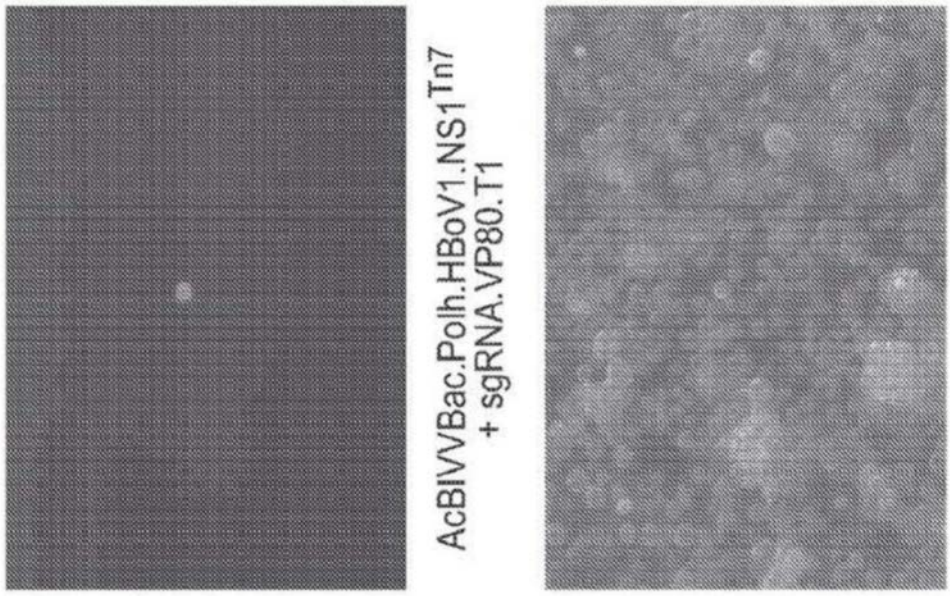

AcBIVVBac.PolhHBoV1.NS1Tn7
+ sgRNA.VP80.T1

图15B

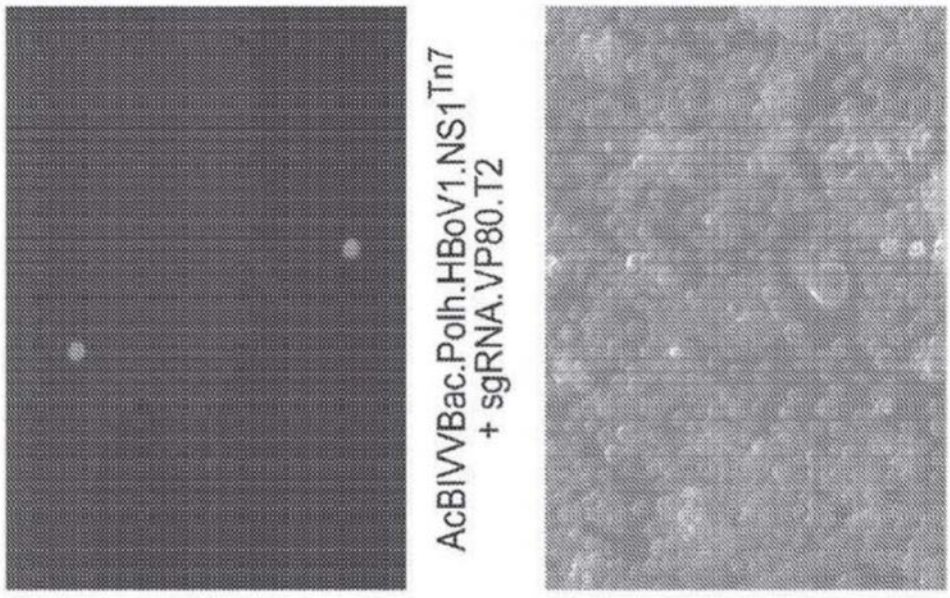

AcBIVVBac.PolhHBoV1.NS1Tn7
+ sgRNA.VP80.T2

图15C

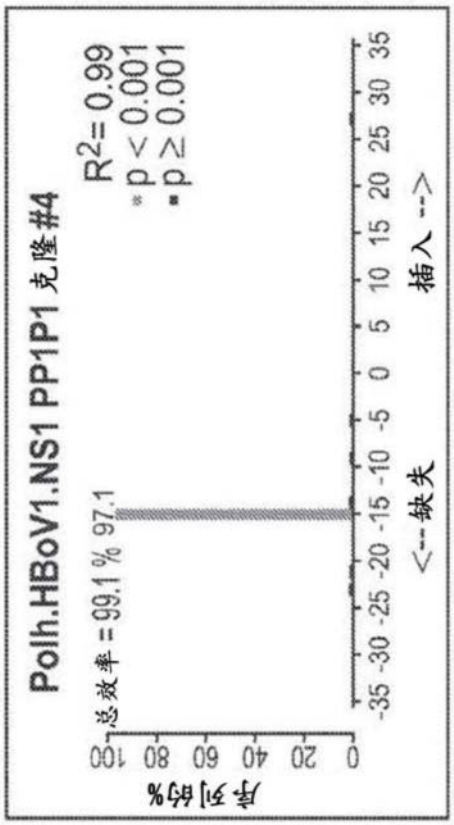

PolhHBoV1.NS1 PP1P1 克隆#4
R2= 0.99
* p < 0.001
▪ p ≥ 0.001
总效率 = 99.1 % 97.1
序列的%
0 20 40 60 80 100
-35 -30 -25 -20 -15 -10 -5 0 5 10 15 20 25 30 35
<-- 缺失 插入 -->

图16A

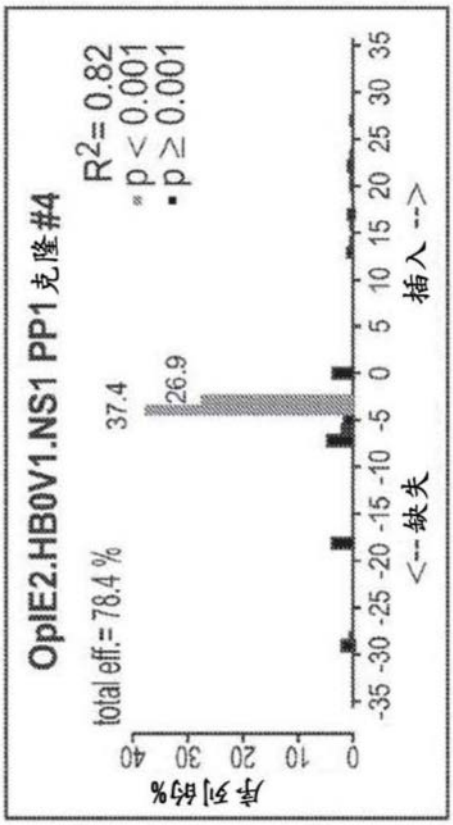

OpIE2.HB0V1.NS1 PP1克隆#4
R^2= 0.82
p < 0.001
p ≥ 0.001
total eff.= 78.4 %
37.4
26.9
<-- 缺失
插入 -->
序列的%

图16B

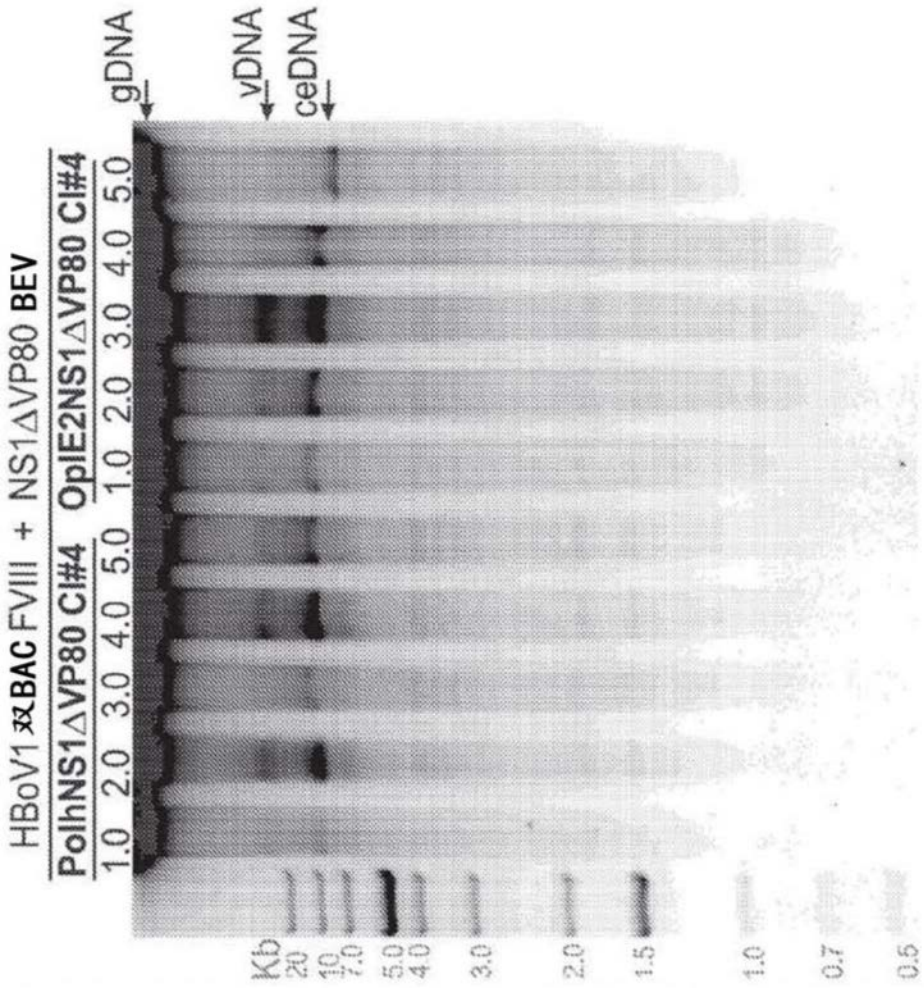

gDNA
vDNA
ceDNA
HBoV1 双BAC FVIII + NS1ΔVP80 BEV
PolhNS1ΔVP80 Cl#4
OplE2NS1ΔVP80 Cl#4
1.0 2.0 3.0 4.0 5.0 1.0 2.0 3.0 4.0 5.0
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5

图16C

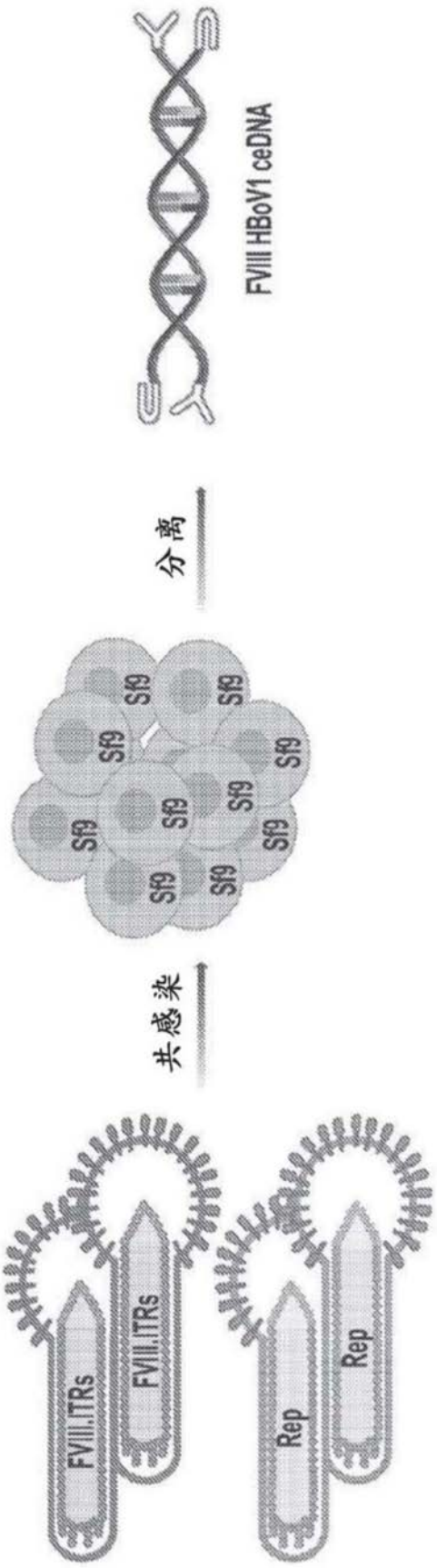


图17A

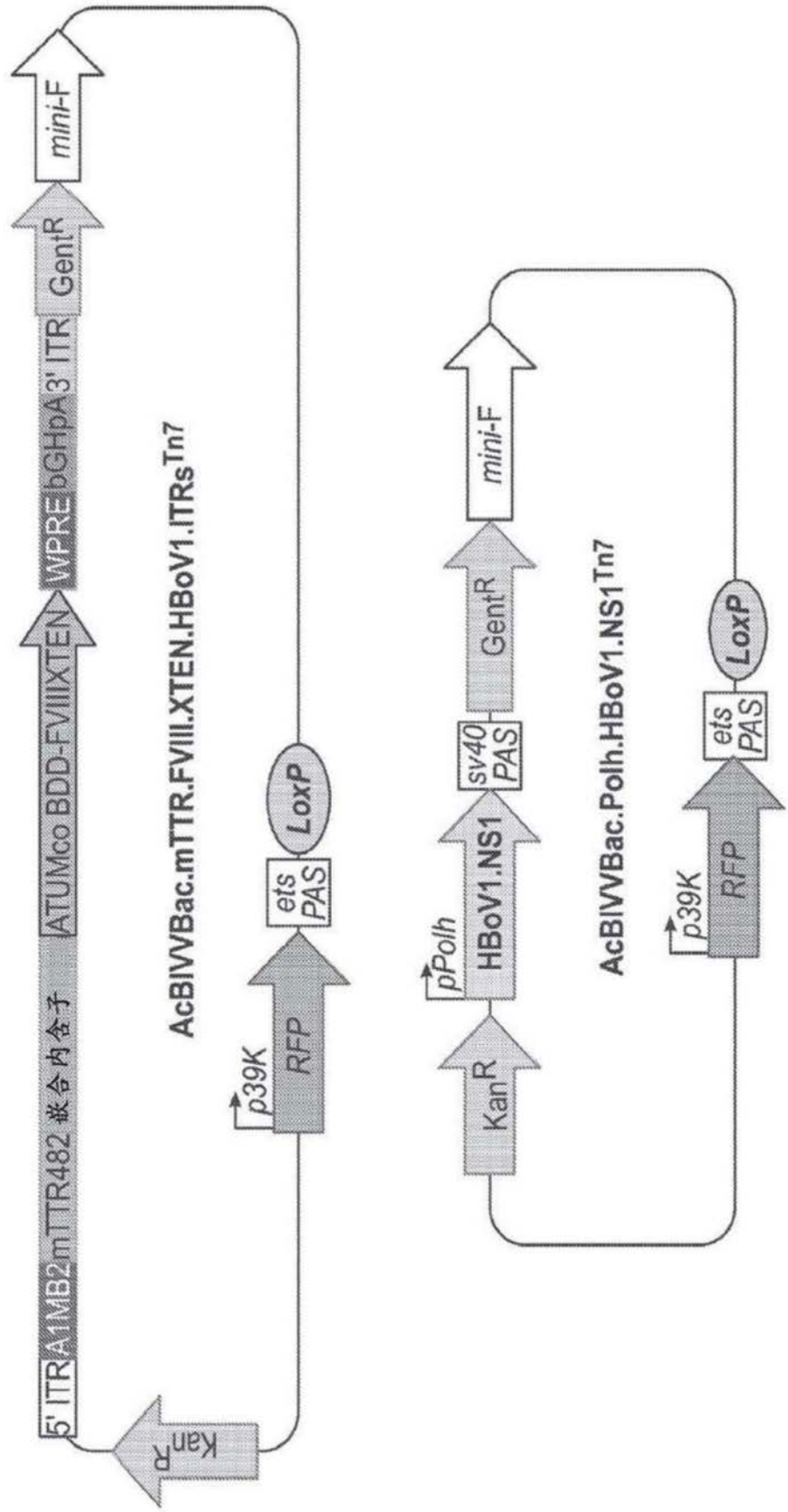

AcBlVVBac.mTTR.FVIII.XTEN.HBoV1.ITRs Tn7
5' ITR
A1MB2mTTR482 嵌合内含子
ATUMco BDD-FVIIIXTEN
WPRE
bGHpA3' ITR
GentR
mini-F
Kan R
p39K
RFP
ets PAS
LoxP
AcBlVVBac.Polh.HBoV1.NS1 Tn7
KanR
pPolh
HBoV1.NS1
sv40 PAS
GentR
mini-F
p39K
RFP
ets PAS
LoxP

图17B

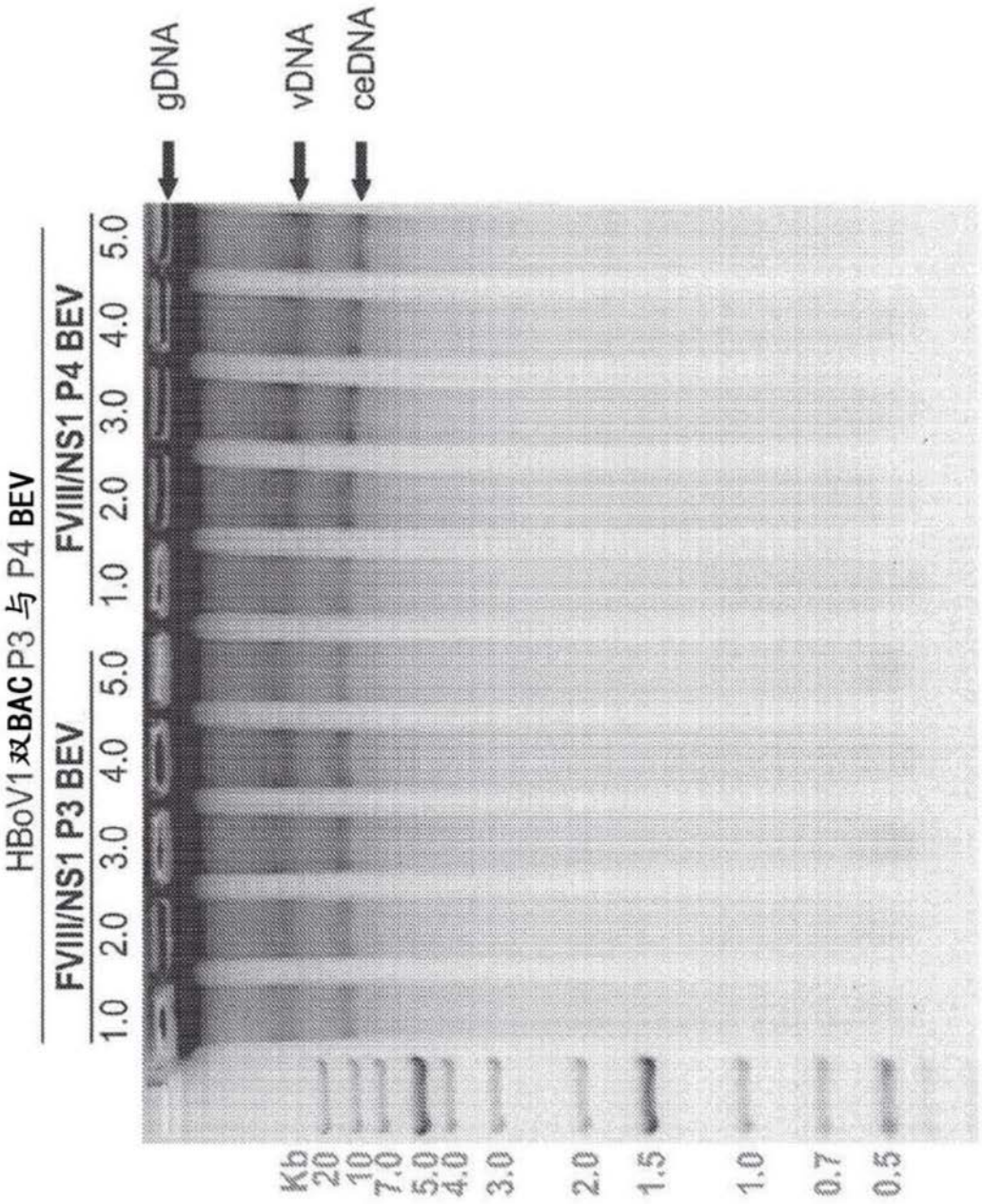

gDNA
vDNA
ceDNA
HBoV1双BAC P3与P4 BEV
FVIII/NS1 P3 BEV
FVIII/NS1 P4 BEV
1.0
2.0
3.0
4.0
5.0
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5

图17C

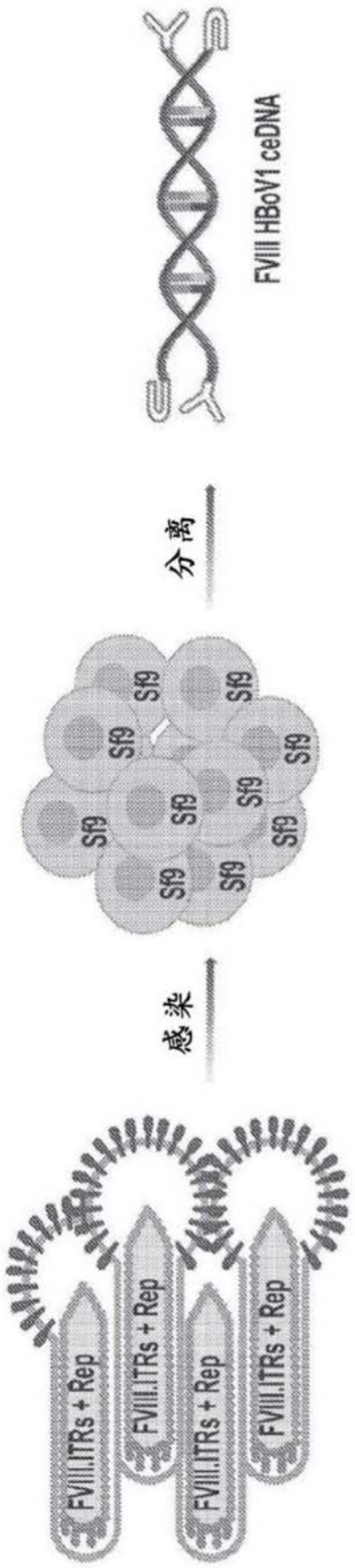

FVIII HBoV1 ceDNA
分离
Sf9
感染
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep
FVIII.ITRs + Rep

图18A

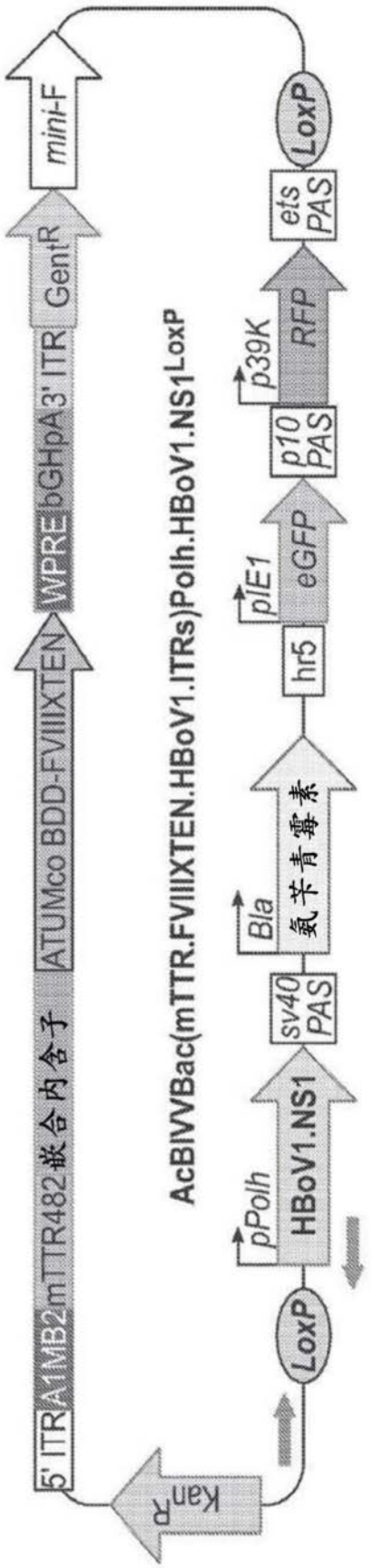

AcBIVVBac(mTTR.FVIIIXTEN.HBoV1.ITRs)PolH.HBoV1.NS1LoxP
mini-F
GentR
WPRE bGHpA 3' ITR
ATUMco BDD-FVIIIXTEN
5' ITR A1MB2mTTR482嵌合内含子
KanR
LoxP
pPolh
HBoV1.NS1
sv40 PAS
Bla
氯苄青霉素
hr5
pIE1
eGFP
p10 PAS
p39K
RFP
ets PAS
LoxP

图18B

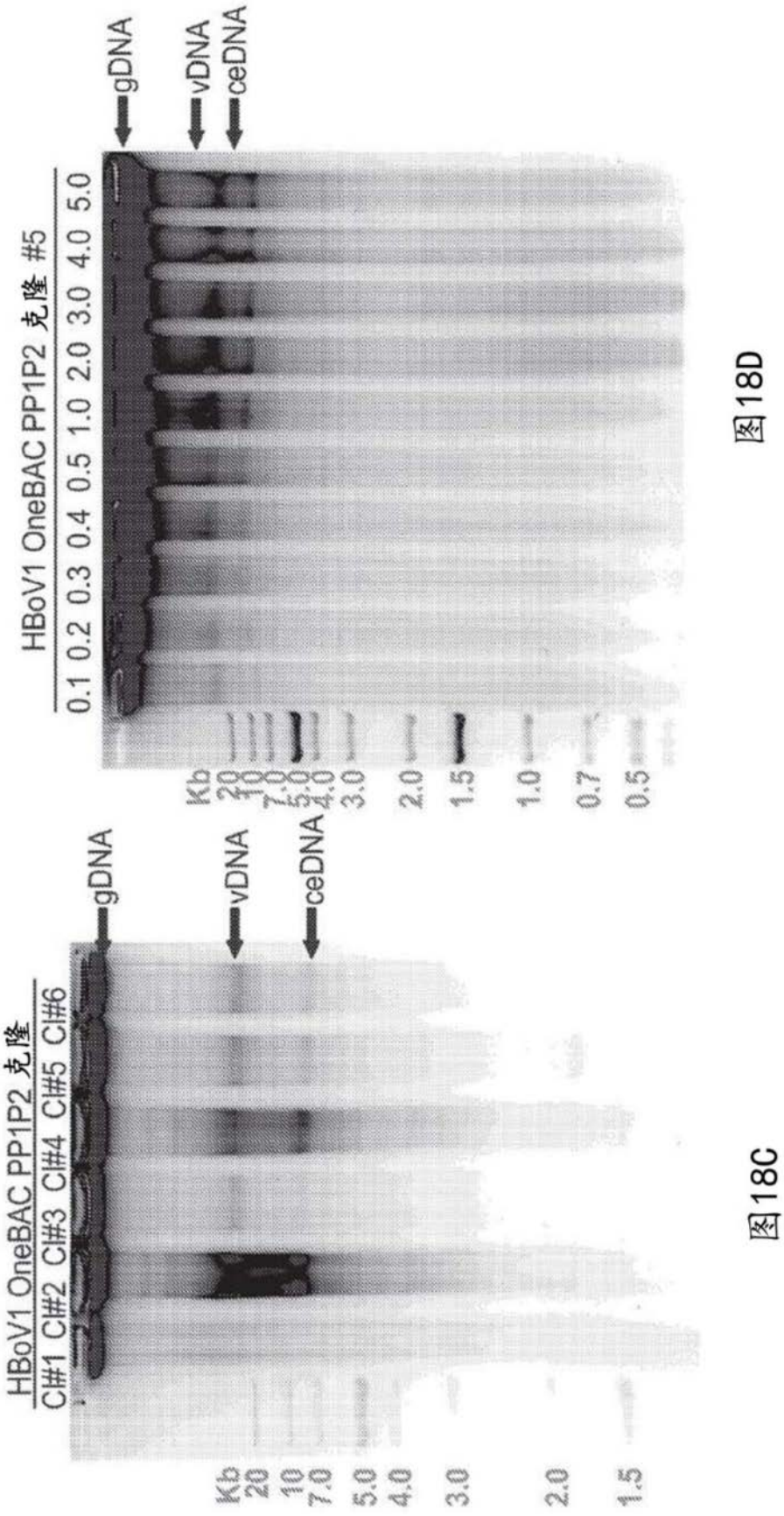

HBoV1 OneBAC PP1P2 克隆 #5
0.1 0.2 0.3 0.4 0.5 1.0 2.0 3.0 4.0 5.0
gDNA
vDNA
ceDNA
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5
HBoV1 OneBAC PP1P2 克隆
C#1 C#2 C#3 C#4 C#5 C#6
gDNA
vDNA
ceDNA
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5

图18D

图18C

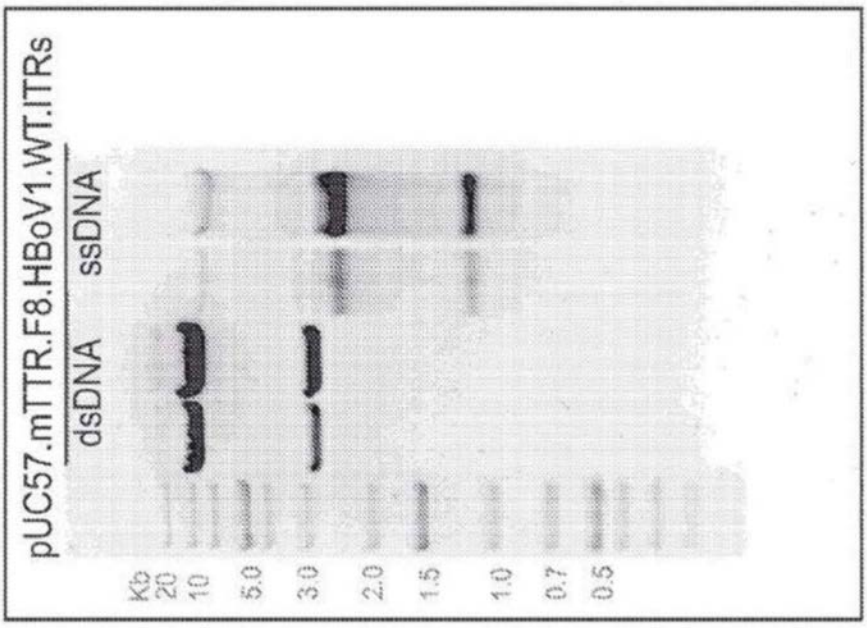

pUC57.mTTR.F8.HBoV1.WT.ITRs
dsDNA
ssDNA
Kb
20
10
5.0
3.0
2.0
1.5
1.0
0.7
0.5

图19A

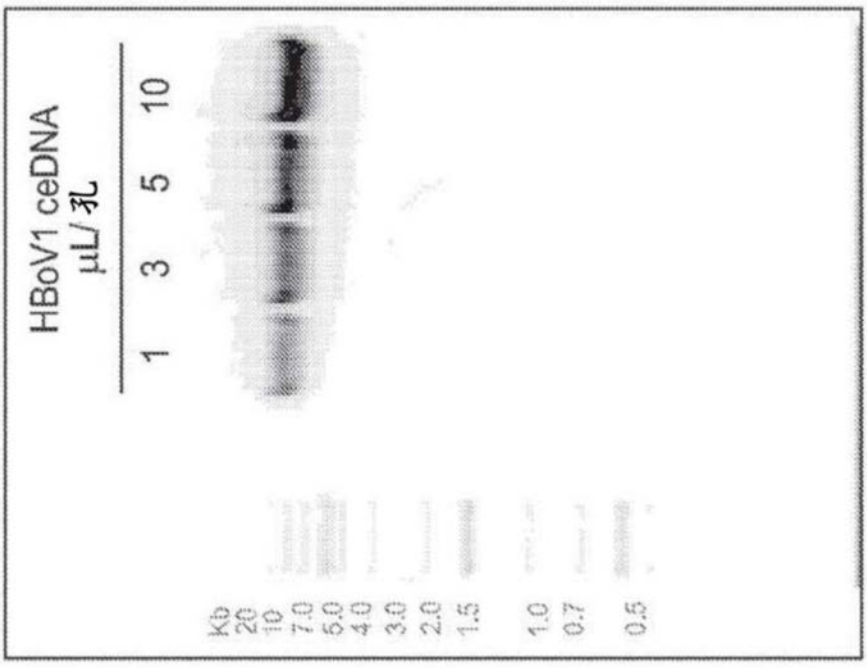

HBoV1 ceDNA
μL/孔
1
3
5
10
Kb
20
10
7.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5

图19B

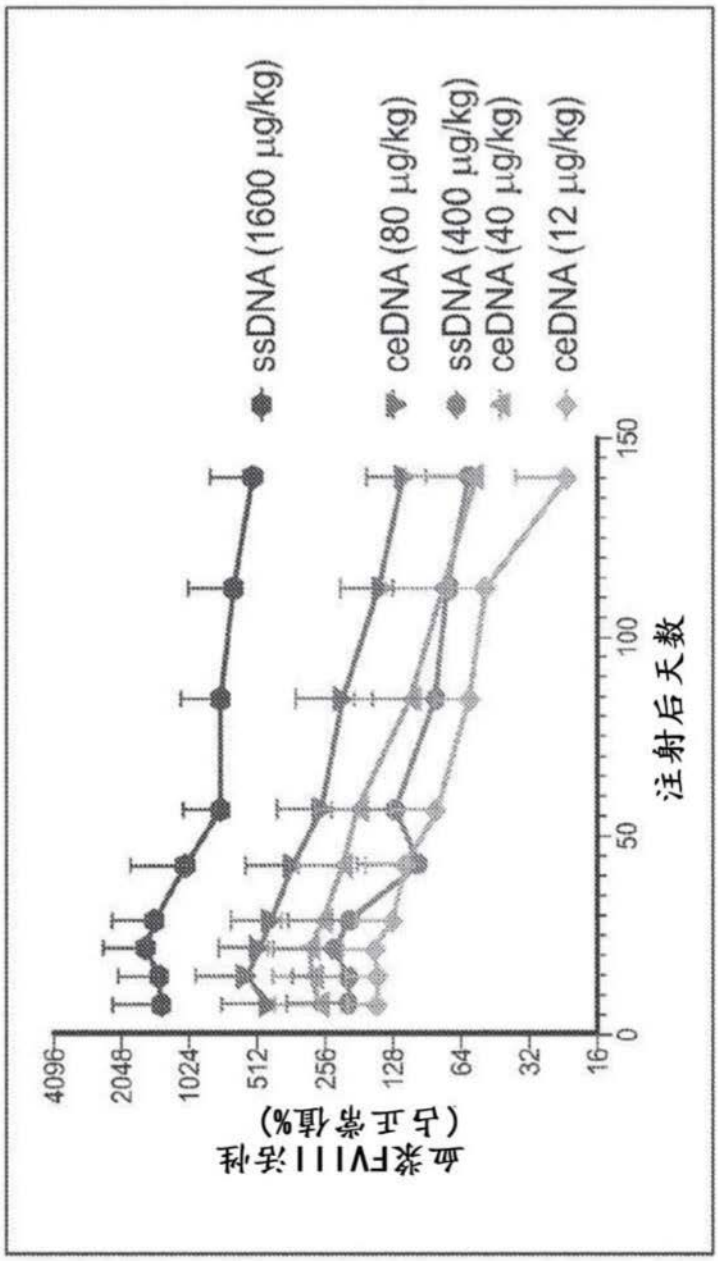

ssDNA (1600 μg/kg)
ceDNA (80 μg/kg)
ssDNA (400 μg/kg)
ceDNA (40 μg/kg)
ceDNA (12 μg/kg)
血浆FVIII活性
(占正常值%)
4096
2048
1024
512
256
128
64
32
16
0
50
100
150
注射后天数

图19C

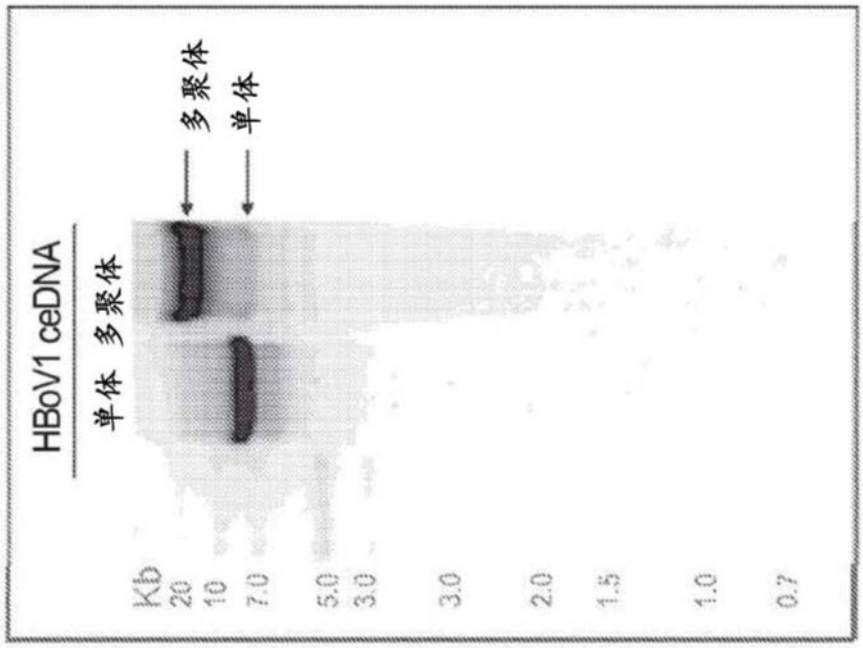


图20A

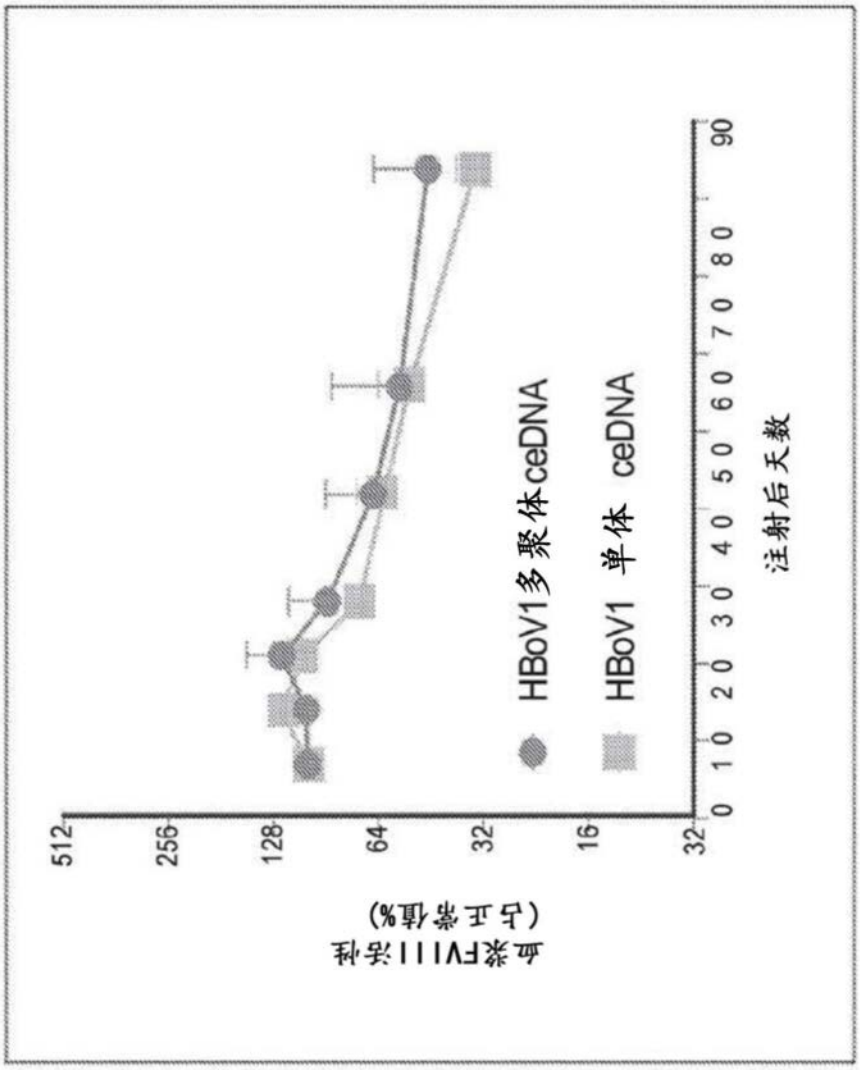

血浆FVIII活性
(占正常值%)
512
256
128
64
32
16
32
0 10 20 30 40 50 60 70 80 90
注射后天数
HBoV1多聚体ceDNA
HBoV1 单体 ceDNA

图20B

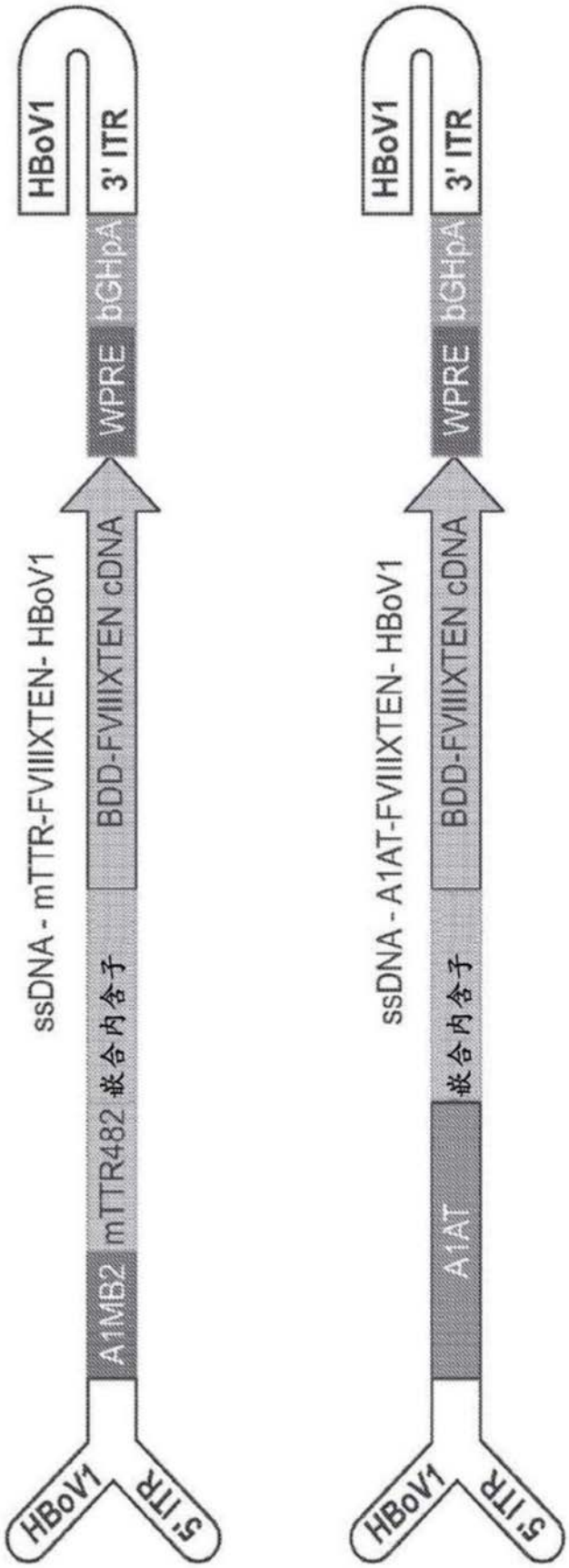

HBoV1
3' ITR
WPRE
bGHpA
BDD-FVIIIXTEN cDNA
ssDNA - mTTR-FVIIIXTEN- HBoV1
A1MB2
mTTR482 嵌合内含子
HBoV1
5' ITR
HBoV1
3' ITR
WPRE
bGHpA
BDD-FVIIIXTEN cDNA
ssDNA - A1AT-FVIIIXTEN- HBoV1
A1AT
嵌合内含子
HBoV1
5' ITR

图21A

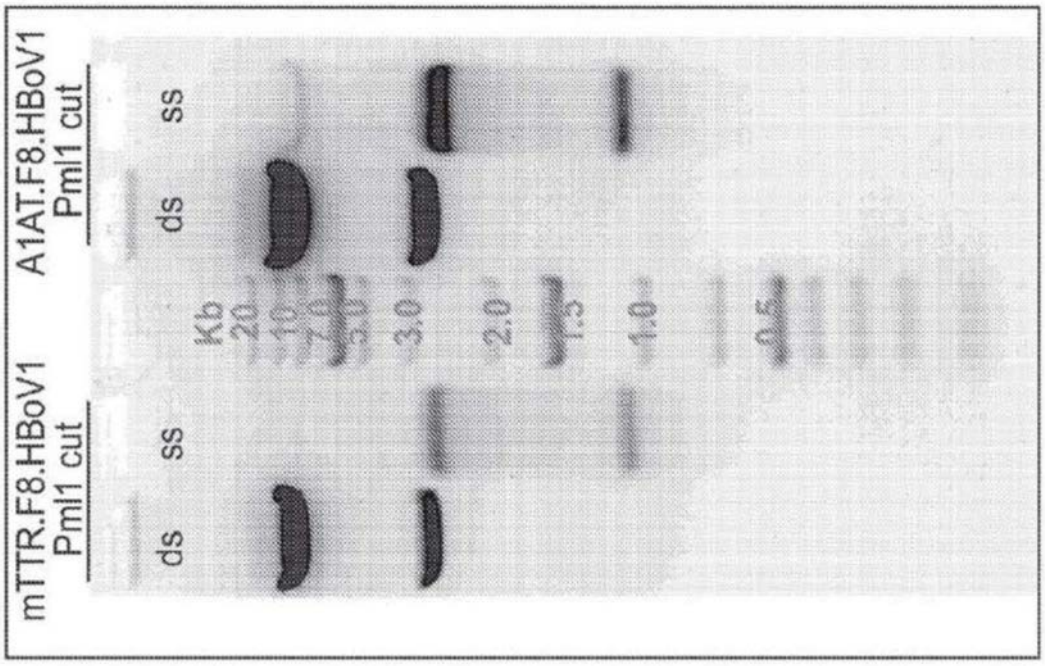

mTTR.F8.HBoV1
Pml1 cut
ds
ss
A1AT.F8.HBoV1
Pml1 cut
ds
ss
Kb
20
10
7.0
5.0
3.0
2.0
1.5
1.0
0.5

图21B

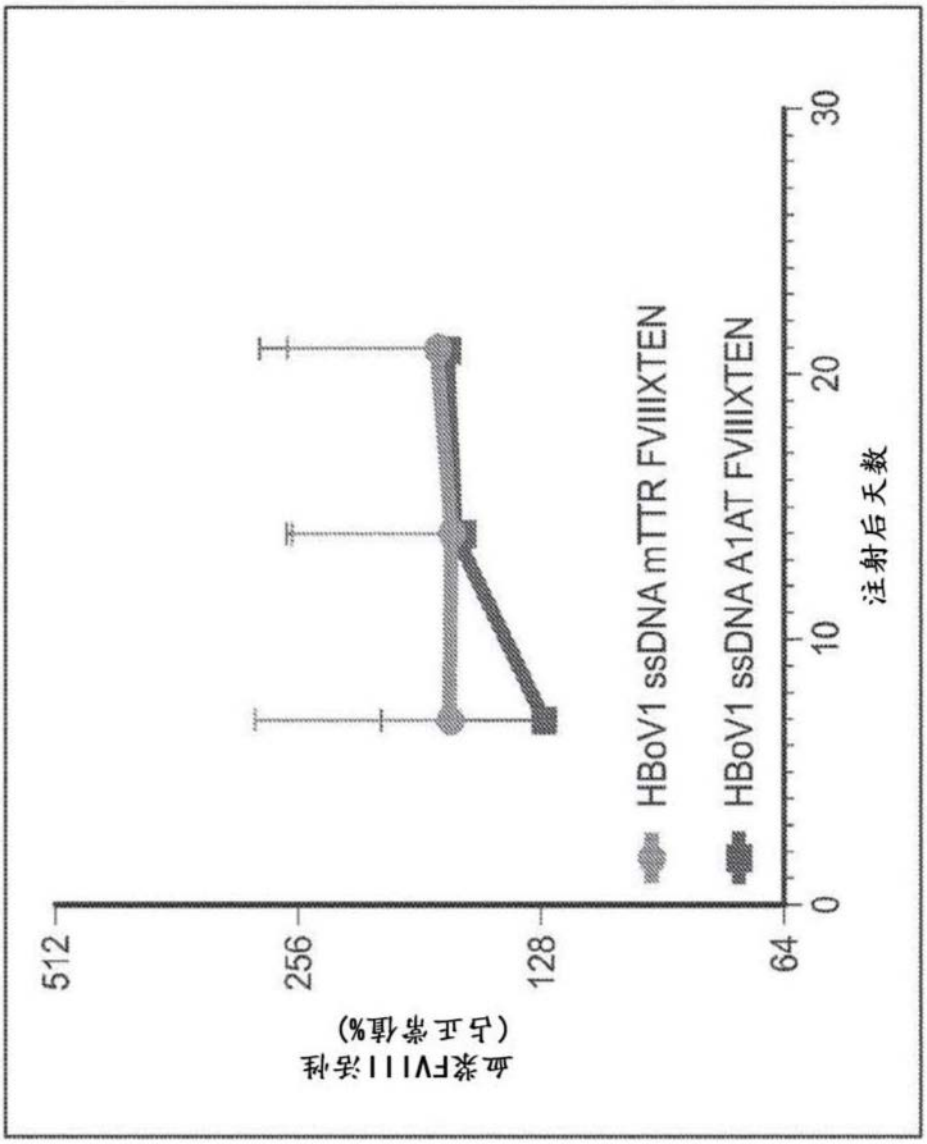

血浆FVIII活性
(占正常值%)
512
256
128
64
0
10
20
30
注射后天数
HBoV1 ssDNA mTTR FVIIIXTEN
HBoV1 ssDNA A1AT FVIIIXTEN

图21C