



(12) 发明专利申请

(10) 申请公布号 CN 116171325 A

(43) 申请公布日 2023.05.26

(21) 申请号 202180056421.4

D·里克斯

(22) 申请日 2021.07.21

(74) 专利代理机构 中国贸促会专利商标事务所

(30) 优先权数据

有限公司 11038

63/055,775 2020.07.23 US

专利代理师 刘晓东

(85) PCT国际申请进入国家阶段日

(51) Int.Cl.

2023.02.10

C12N 15/86 (2006.01)

(86) PCT国际申请的申请数据

PCT/EP2021/070455 2021.07.21

(87) PCT国际申请的公布数据

W02022/018171 EN 2022.01.27

(71) 申请人 UCL商业有限责任公司

地址 英国伦敦

(72) 发明人 S·N·沃丁顿 R·卡达

C·D·赫尔佐格 J·恩格

权利要求书5页 说明书37页

C·萨克拉门多 S·肖尔格

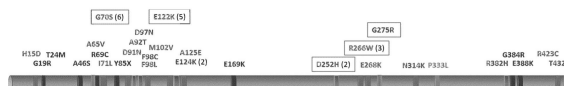
序列表44页 附图19页

(54) 发明名称

用于eEF1A2的基因疗法载体及其用途

(57) 摘要

本文提供了一种神经系统疾病的基因疗法，其使用重组腺相关病毒(rAAV)病毒粒子作为载体来表达eEF1A2蛋白或其功能变体。所述rAAV病毒粒子可以使用神经元特异性启动子，例如，人突触素1(hSYN)启动子。所述衣壳可以是AAV9衣壳或其功能变体。可以使用其它启动子或衣壳。进一步提供了治疗方法，如通过脑内和/或静脉内rAAV病毒粒子，以及其它组合物和方法。



1. 一种重组腺相关病毒 (rAAV) 病毒粒子, 其包括衣壳和载体基因组, 其中所述载体基因组包括与启动子操作性地连接的编码eEF1A2蛋白或其功能变体的多核苷酸序列。
2. 根据权利要求1所述的rAAV病毒粒子, 其中所述启动子是神经元特异性启动子。
3. 根据权利要求1或权利要求2所述的rAAV病毒粒子, 其中所述启动子是泛神经元启动子。
4. 根据权利要求1至3中任一项所述的rAAV病毒粒子, 其中所述启动子是突触素1启动子。
5. 根据权利要求4所述的rAAV病毒粒子, 其中所述突触素1启动子是人突触素1 (hSYN) 启动子。
6. 根据权利要求5所述的rAAV病毒粒子, 其中所述hSYN启动子包括与SEQ ID NO:3具有至少70%、至少80%或至少90%同一性的多核苷酸序列。
7. 根据权利要求6所述的rAAV病毒粒子, 其中所述hSYN启动子包括与SEQ ID NO:3具有至少95%或至少99%同一性的多核苷酸序列。
8. 根据权利要求7所述的rAAV病毒粒子, 其中所述hSYN启动子包括SEQ ID NO:3的多核苷酸序列。
9. 根据权利要求1至3中任一项所述的rAAV病毒粒子, 其中所述启动子是eSYN启动子。
10. 根据权利要求9所述的rAAV病毒粒子, 其中所述eSYN启动子包括与SEQ ID NO:64具有至少70%、至少80%、至少90%、至少95%、至少99%或100%同一性的多核苷酸序列。
11. 根据权利要求1至10中任一项所述的rAAV病毒粒子, 其中编码所述eEF1A2蛋白的所述多核苷酸序列与SEQ ID NO:2具有至少70%、至少80%或至少90%的同一性。
12. 根据权利要求11所述的rAAV病毒粒子, 其中编码所述eEF1A2蛋白的所述多核苷酸序列与SEQ ID NO:2具有至少95%或至少99%的同一性。
13. 根据权利要求12所述的rAAV病毒粒子, 其中编码所述eEF1A2蛋白的所述多核苷酸序列包括SEQ ID NO:2的多核苷酸序列。
14. 根据权利要求1至13中任一项所述的rAAV病毒粒子, 其中所述eEF1A2蛋白与SEQ ID NO:1具有至少70%、至少80%或至少90%的同一性。
15. 根据权利要求14所述的rAAV病毒粒子, 其中所述eEF1A2蛋白与SEQ ID NO:1具有至少95%或至少99%的同一性。
16. 根据权利要求15所述的rAAV病毒粒子, 其中所述eEF1A2蛋白包括SEQ ID NO:1的多核苷酸序列。
17. 根据权利要求15所述的rAAV病毒粒子, 其中所述启动子是突触素1启动子。
18. 根据权利要求17所述的rAAV病毒粒子, 其中所述突触素1启动子是人突触素1 (hSYN) 启动子。
19. 根据权利要求18所述的rAAV病毒粒子, 其中所述hSYN启动子包括与SEQ ID NO:3具有至少95%同一性的多核苷酸序列。
20. 根据权利要求1或11至15中任一项所述的rAAV病毒粒子, 其中所述启动子是组成型启动子。
21. 根据权利要求1、11至15或20中任一项所述的rAAV病毒粒子, 其中所述启动子是CAG启动子, 其中任选地所述CAG启动子与SEQ ID NO:14具有至少95%的同一性。

22. 根据权利要求1、11至15或20中任一项所述的rAAV病毒粒子,其中所述启动子是CMV启动子,其中任选地所述CMV启动子包括与SEQ ID NO:16或17至少95%相同的序列。

23. 根据权利要求1至22中任一项所述的rAAV病毒粒子,其中所述载体基因组包括聚腺苷酸化(polyA)位点。

24. 根据权利要求23所述的rAAV病毒粒子,其中所述polyA序列是bGH聚腺苷酸化位点。

25. 根据权利要求24所述的rAAV病毒粒子,其中所述bGH聚腺苷酸化位点与SEQ ID NO:53具有至少95%的同一性。

26. 根据权利要求23所述的rAAV病毒粒子,其中所述polyA序列是hGH聚腺苷酸化位点。

27. 根据权利要求26所述的rAAV病毒粒子,其中所述hGH聚腺苷酸化位点与SEQ ID NO:54具有至少95%的同一性。

28. 根据权利要求1至27中任一项所述的rAAV病毒粒子,其中所述载体基因组包括WPRE(x)元件。

29. 根据权利要求28所述的rAAV病毒粒子,其中所述WPRE(x)元件与SEQ ID NO:42具有至少95%的同一性。

30. 根据权利要求28所述的rAAV病毒粒子,其中所述WPRE(x)元件与SEQ ID NO:41或SEQ ID NO:43具有至少95%的同一性。

31. 根据权利要求1至30中任一项所述的rAAV病毒粒子,其中所述载体基因组包括Kozak序列。

32. 根据权利要求31所述的rAAV病毒粒子,其中所述Kozak序列是SEQ NO:10。

33. 根据权利要求1至32中任一项所述的rAAV病毒粒子,其中所述载体基因组包括5'非翻译区(UTR),其与SEQ ID NO:32-40中的一者或多者具有至少95%的同一性。

34. 根据权利要求1至33中任一项所述的rAAV病毒粒子,其中所述载体基因组包括3'非翻译区(UTR),其与SEQ ID NO:41-49中的一者或多者具有至少95%的同一性。

35. 根据权利要求1至34中任一项所述的rAAV病毒粒子,其中所述载体基因组包括5'倒置末端重复序列(ITR),其具有与SEQ ID NO:19或SEQ ID NO:20至少95%相同的序列。

36. 根据权利要求1至35中任一项所述的rAAV病毒粒子,其中所述载体基因组包括3'倒置末端重复序列(ITR),其具有与SEQ ID NO:21或SEQ ID NO:63至少95%相同的序列。

37. 根据权利要求21所述的rAAV病毒粒子,其中所述CAG启动子与SEQ ID NO:14具有至少95%的同一性。

38. 根据权利要求1至36中任一项所述的rAAV病毒粒子,其中所述衣壳是AAV9衣壳或其功能变体。

39. 根据权利要求1至38中任一项所述的rAAV病毒粒子,其中所述衣壳与SEQ ID NO:15具有至少98%、99%或100%的同一性。

40. 根据权利要求1至39中任一项所述的rAAV病毒粒子,其中所述载体基因组包括表达盒,所述表达盒按5'至3'顺序包括:

- a. HuBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pA珠蛋白-0c;
- b. CMV启动子、TPL-eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)和pA珠蛋白-0c;
- c. Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)、3'UTR(珠蛋白)和

pAGH-Bt;

d. CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Bt;

e. EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列和pA珠蛋白-0c;

f. HuBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pAGH-Bt;

g. Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x)、3'UTR (珠蛋白) 和pAGH-Hs;

h. CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (r) 和pAGH-Hs;

i. CMV启动子、TPL-eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (r) 和pAGH-Hs;

j. HuBA启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs;

k. CMV启动子、TPL/eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR (珠蛋白) 和pAGH-Bt;

l. EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (r) 和pAGH-Bt;

m. Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pA珠蛋白-0c;

n. CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pA珠蛋白-0c;

o. CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x)、3'UTR (珠蛋白) 和pAGH-Hs;

p. CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、3'UTR (珠蛋白) 和pA珠蛋白-0c;

q. CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pAGH-Bt;

r. EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR (珠蛋白) 和pAGH-Hs;

s. CMV启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR (珠蛋白) 和pAGH-Hs;

t. CMV启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs;

u. hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Bt;

v. hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Hs;

w. hSYN启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Hs;

x. CAG启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Hs;

y. CAG启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Hs;

z. hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE (x) 和pAGH-Bt;

aa. hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs;

bb. hSYN启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs;

cc. CAG启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs; 或

dd. CAG启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

41. 根据权利要求1至40中任一项所述的rAAV病毒粒子, 其中所述载体基因组包括与SEQ ID NO:55-58或65-68中任一者具有至少90%、95%、99%或100%同一性的多核苷酸序列, 基本上由所述多核苷酸序列组成或由所述多核苷酸序列组成。

42. 一种治疗和/或预防有需要的受试者的神经系统疾病或病症的方法, 所述方法包括

向所述受试者施用根据权利要求1至41中任一项所述的rAAV病毒粒子。

43. 根据权利要求42所述的方法,其中所述受试者具有或疑似具有一个或多个EEF1A2基因突变。

44. 根据权利要求42或权利要求43所述的方法,其中所述神经系统疾病或病症包括癫痫。

45. 根据权利要求42至44中任一项所述的方法,其中所述神经系统疾病或病症包括智力残疾。

46. 根据权利要求42至45中任一项所述的方法,其中所述神经系统疾病或病症包括自闭症。

47. 根据权利要求42至46中任一项所述的方法,其中所述施用步骤包括脑室内施用。

48. 根据权利要求42至47中任一项所述的方法,其中所述施用步骤包括静脉内施用。

49. 根据权利要求42至48中任一项所述的方法,其中所述施用步骤包括同时或按顺序的脑室内施用和静脉内施用。

50. 根据权利要求42至49中任一项所述的方法,其中所述受试者是哺乳动物。

51. 根据权利要求42至50中任一项所述的方法,其中所述方法防止所述受试者与未经治疗的对照受试者相比的神经性能下降。

52. 根据权利要求24至51中任一项所述的方法,其中所述方法改善肌肉力量和/或运动技能,其中任选地使用倒置网格测试或转棒测试来评估肌肉力量和/或运动技能。

53. 一种在有需要的受试者的脑中表达eEF1A2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41中任一项所述的rAAV病毒粒子。

54. 根据权利要求53所述的方法,其中所述方法在前脑中引起更高的表达,其中任选地将所述表达与包括载体基因组的参考载体进行比较,所述载体基因组包括hSYN启动子、Kozak序列、eEF1A2转基因和/或人球蛋白聚腺苷酸化序列(hGH)。

55. 根据权利要求53所述的方法,其中所述方法在皮层中引起更高的表达,其中任选地将所述表达与包括载体基因组的参考载体进行比较,所述载体基因组包括hSYN启动子、Kozak序列、eEF1A2转基因和/或人球蛋白聚腺苷酸化序列(hGH)。

56. 根据权利要求53至55中任一项所述的方法,其中所述rAAV病毒粒子的所述载体基因组不包括Kozak序列。

57. 根据权利要求53至56中任一项所述的方法,其中所述rAAV病毒粒子的所述载体基因组不包括人球蛋白聚腺苷酸化序列(hGH)。

58. 根据权利要求53至57中任一项所述的方法,其中所述rAAV病毒粒子的所述载体基因组包括牛球蛋白聚腺苷酸化序列(bGH)。

59. 根据权利要求53至58中任一项所述的方法,其中所述受试者具有或疑似具有一个或多个EEF1A2基因突变。

60. 根据权利要求53至59中任一项所述的方法,其中所述受试者患有神经系统疾病或病症或有患神经系统疾病或病症的风险。

61. 根据权利要求60所述的方法,其中所述神经系统疾病或病症包括癫痫。

62. 根据权利要求60或权利要求61所述的方法,其中所述神经系统疾病或病症包括智力残疾。

63. 根据权利要求60至62中任一项所述的方法,其中所述神经系统疾病或病症包括自闭症。

64. 根据权利要求53至63中任一项所述的方法,其中所述施用步骤包括脑室内施用。

65. 根据权利要求53至63中任一项所述的方法,其中所述施用步骤包括静脉内施用。

66. 根据权利要求53至63中任一项所述的方法,其中所述施用步骤包括同时或按顺序的脑室内施用和静脉内施用。

67. 根据权利要求53至67中任一项所述的方法,其中所述受试者是哺乳动物。

68. 根据权利要求53至67中任一项所述的方法,其中所述方法防止所述受试者与未经治疗的对照受试者相比的神经性能下降。

69. 根据权利要求53至68中任一项所述的方法,其中所述方法改善肌肉力量和/或运动技能,其中任选地使用倒置网格测试或转棒测试来评估肌肉力量和/或运动技能。

70. 一种药物组合物,其包括根据权利要求1至41中任一项所述的rAAV病毒粒子。

71. 一种试剂盒,其包括根据权利要求1至41中任一项所述的rAAV病毒粒子和使用说明书。

用于eEF1A2的基因疗法载体及其用途

[0001] 相关申请交叉引用

[0002] 本申请要求于2020年7月23日提交的美国临时专利申请序列号63/055,775的优先权的权益,所述美国临时专利申请在此通过引用整体并入。

[0003] 关于序列表的声明

[0004] 与本申请相关的序列表以文本形式提供以代替纸质副本,并且特此通过引用并入本说明书中。含有序列表的文本文件的名称为ROPA_019_02WO_ST25.txt。所述文本文件为约101千字节,于2021年7月20日创建,并通过EFS-Web以电子方式提交。

背景技术

[0005] EEF1A2基因编码真核伸长因子1, α -2 (eEF1A2),一种参与蛋白质合成、抑制细胞凋亡并调节肌动蛋白功能和细胞骨架结构的蛋白质。小鼠和人类直系同源物在463个氨基酸位置中的462个位置处具有同一性。EEF1A2是一种潜在的癌基因,因为它在卵巢癌中过表达。在卵巢癌研究中,实验性地使用编码EEF1A2的慢病毒载体转导永生化的卵巢表面上皮(IOSE)细胞,从而证明eEF1A2促进非致瘤前体细胞中的肿瘤发生。Sun等人《国际癌症杂志(Int J Cancer.)》123(8):1761-176(2008)。

[0006] EEF1A2在中枢神经系统(CNS)以及心脏和肌肉中高度表达。小鼠中Eef1a2的完全丧失导致运动神经元变性,这种表型被称为“废弃”,其基因型被称为wst。Davies等人《科学报告(Sci Rep.)》7:46019(2017)。人类EEF1A2基因中的点突变最近已被证明可导致癫痫、智力残疾和/或自闭症。Cao等人《人类分子遗传学(Human Molecular Genetics.)》26(18):3545-3552(2017);Lam等人《分子遗传学与基因组医学(Mol Genet Genomic Med.)》4(4):465-74(2016);Nakajima等人《临床遗传学(Clin Genet.)》87(4):356-61(2015)。

[0007] 利用携带野生型Eef1a2的转基因小鼠在细菌人工染色体(BAC)上的实验已经证实,当在发育过程中存在时,野生型Eef1a2补充了wst基因型。Newbury等人《生物化学杂志(J.Bio.Chem.)》282:2891-50(2007)。

[0008] EEF1A2相关疾病很少见。全世界只有大约100人被鉴定为具有EEF1A2突变。对疾病的病因仍然知之甚少。因此,尚不清楚是否可以通过出生后表达野生型EEF1A2来挽救疾病表型。此外,向CNS提供基因疗法具有挑战性且不可预测。

[0009] 对用于EEF1A2相关疾病的疗法存在未满足的需求。本文提供的基因疗法满足了这一需求。

发明内容

[0010] 本发明总体上涉及神经疾病或病症的基因疗法,其使用编码eEF1A2或其功能变体的多核苷酸的基于腺相关病毒(AAV)的递送。

[0011] 在一方面,本公开提供了一种重组腺相关病毒(rAAV)病毒粒子,其包括衣壳和载体基因组,其中所述载体基因组包括与启动子操作性地连接的编码eEF1A2蛋白或其功能变体的多核苷酸序列。所述启动子可以是神经元特异性启动子,例如,人突触素1(hSYN)启动

子。所述衣壳可以是AAV9衣壳或其功能变体。可以使用其它启动子或衣壳。

[0012] 在另一方面,本公开提供了一种治疗和/或预防有需要的受试者的神经疾病或病症的方法,所述方法包括向所述受试者施用本公开的rAAV病毒粒子或其药物组合物。所述rAAV病毒粒子可以脑内和/或静脉内施用。

[0013] 在另一方面,本公开提供了一种在有需要的受试者的脑中表达eEF1A2的方法,所述方法包括向所述受试者施用本公开的rAAV病毒粒子或其药物组合物。所述rAAV病毒粒子可以脑内和/或静脉内施用。

[0014] 在另外的方面,本公开提供了多核苷酸(例如,载体基因组)、药物组合物、试剂盒和其它组合物和方法。

[0015] 下文的详细描述中公开了各个其它方面和实施例。本发明仅由所附权利要求限制。

附图说明

[0016] 图1示出了eEF1A2的结构域图,所述结构域图示出了与疾病相关的点突变。

[0017] 图2示出了载体基因组的非限制性实例的载体图。

[0018] 图3示出了载体基因组的非限制性实例的载体图。

[0019] 图4示出了载体基因组的非限制性实例的载体图。

[0020] 图5示出了载体基因组的非限制性实例的载体图。

[0021] 图6示出了载体基因组的非限制性实例的载体图。

[0022] 图7示出了在新生儿注射、脑内(IC)或静脉内(IV)施用AAV9-hSyn-eEF1A2-2A-eGFP或对照后,小鼠的免疫荧光显微镜检查。比例尺,300 μ m。

[0023] 图8A示出了在新生儿注射、脑内(IC)或静脉内(IV)施用AAV9-hSyn-eEF1A2-2A-eGFP或对照后,小鼠的免疫组织化学分析。图8B示出了相同载玻片的放大视图。比例尺,300 μ m。

[0024] 图9A示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的存活期。

[0025] 图9B示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的重量减轻。

[0026] 图9C示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的转棒仪测试。

[0027] 图9D示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的倒置网格测试。

[0028] 图9E示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的eEF1A2表达。比例尺,125 μ m。

[0029] 图9F示出了与脑内(IC)、静脉内(IV)或两者的组合(IC+IV)治疗的小鼠相比,未经治疗的wst/wst(无效)小鼠中的eEF1A2表达。

[0030] 图10A-10K示出了以 2×10^{11} vg/动物施用的AAV9载体的比较,包括图2(“V1”)、图3(“V2”)、图4(“V3”)和图6(“V4”)中所示的载体基因组。

[0031] 图10A示出了FBS治疗的野生型、wst/wst、脑室内治疗的wst/wst动物用V1、V2、V3

和V4基因疗法治疗的Kaplan-Meier存活图。

[0032] 图10B示出了小鼠的重量(数据均值±S.E.M.)。每天称重动物直到出生后35岁,此后每周称重,直到定时处死点P60或人道终点15%的重量减轻。

[0033] 图10C示出了在第P15天通过倒置网格进行的肌肉力量评估。

[0034] 图10D示出了在第P15天通过转棒进行的肌肉力量评估。

[0035] 图10E示出了在第P23天通过倒置网格进行的肌肉力量评估。

[0036] 图10F示出了在第P23天通过转棒进行的肌肉力量评估。

[0037] 图10G示出了以野生型FBS(调配物缓冲溶液)为生理参考通过自由浮动免疫组化对整个脑的eEF1A2的代表性免疫染色(每组n=4-5,比例尺250μm)。

[0038] 图10H示出了在共表达eEF1A2的神经元(NeuN标志物)皮质脑区中的代表性免疫组织荧光(每组n=4.5,200μm)。

[0039] 图10I示出了脑中eEF1A2的代表性免疫印迹,其中定量示出了eEF1A2在整个脑中的表达,与V4载体相比,所有基因疗法载体在中脑、小脑和后脑区域中的表达更高(数据均值±S.E.M.、双向ANOVA)。

[0040] 图10J示出了前脑中的人eEF1A2转录本表达的qPCR,其示出V1载体的最高mRNA表达(数据均值±S.E.M.、双向ANOVA)。

[0041] 图10K示出了皮层中的人eEF1A2皮层表达的qPCR,其示出V1载体的最高mRNA表达(数据均值±S.E.M.、双向ANOVA)。

[0042] 图11A-11C

[0043] 图11A示出了FBS治疗的野生型、FBS治疗的D252H/+、敲除(D252H/-)、脑室内V3基因疗法治疗的D252H/-的Kaplan-Meier存活图(野生型FBS,n=12,D252H/+FBS,n=5,D252H/-FBS,n=4和 2×10^{11} vg/幼崽,V3治疗,n=5)。

[0044] 图11B示出了体重随时间的变化(数据均值±S.E.M.、单向ANOVA和邓尼特多重比较)。

[0045] 图11C示出了通过转棒进行的运动评估(数据均值±S.E.M.、双向ANOVA和邓尼特多重比较)。

[0046] 图12A示出了接受V3高剂量的De122ex3(2×10^{11} vg/幼崽,n=5)、接受V3低剂量的De122ex3(2×10^{10} vg/幼崽,n=5)、接受调配物缓冲溶液的De122ex3对照(n=3)和接受调配物缓冲溶液的野生型对照(n=6)的Kaplan-Meier存活图。

[0047] 图12B示出了体重随时间的变化(数据均值±S.E.M.)。

[0048] 图12C示出了通过握力测压P22-25进行的运动评估(数据均值±S.E.M.)

[0049] 图12D示出了P23处的握力测压(数据均值±S.E.M.、双向ANOVA和Tukey多重比较)。

[0050] 图12E示出了通过转棒仪P22-25进行的运动评估(数据均值±S.E.M.)

[0051] 图12F示出了P24处的转棒仪数据(数据均值±S.E.M.、双向ANOVA和Tukey多重比较)。

[0052] 图12G示出了来自P21-25的神经学评分。

具体实施方式

[0053] 小节标题仅是为了组织目的并且不应该被解释为将所描述的主题限于特定方面或实施例。

[0054] 除非另外定义,否则本文所使用的所有技术术语和科学术语的含义与本发明涉及的领域的普通技术人员通常理解的含义相同。虽然类似于或等同于本文所描述的那些方法和材料的方法和材料可以用于本发明的实践中,但是下面描述了合适的方法和材料。本文所提及的所有出版物、专利申请、专利和其它参考文献通过引用以其全部内容明确地并入。在发生冲突的情况下,应以本说明书(包含定义)为准。另外,本文所描述的材料、方法和实例仅是说明性的,而非限制性的。

[0055] 本文提及的所有出版物和专利均通过引用以其全文并入本文,就如同每个单独的出版物或专利被明确地且单独地指出通过引用并入一样。在冲突的情况下,以本申请(包含本文中的任何定义)为准。然而,对本文所引用的任何参考文献、文章、出版物、专利、专利出版物和专利申请的提及不被视为并且不应当被视为对其构成有效的现有技术或形成世界上任何国家的公知常识的一部分的承认或任何形式暗示。

[0056] 在本说明书中,任何浓度范围、百分比范围、比率范围或整数范围应当理解为包含所述范围内的任何整数的值,以及在适当时其分数(如整数的十分之一和百分之一),除非另有说明。当紧接在一个数字或数值之前时,术语“约”意指所述数字或数值的范围为加或减10%。应当理解,如本文所使用的,术语“一个/种(a和an)”是指所列举的组分中的“一个或多个”,除非另有说明。替代方案(例如,“或”)的使用应理解为意指替代方案中的任一者、两者或其任何组合。术语“和/或”应理解为意指替代方案中的一者或两者。如本文所使用的,术语“包含”和“包括”同义地使用。

[0057] 如本文所使用的,术语“同一性”和“相同”相对于多肽或多核苷酸序列是指“查询”序列与“受试者”序列的比对,如提供BLAST算法生成的比对中精确匹配残基的百分比。除非另有说明,否则跨受试者序列的全长计算同一性。因此,如果当查询序列与受试者序列比对时,受试者序列中的残基的至少x%(向下四舍五入)作为精确匹配与查询序列中的对应残基比对,则查询序列“与受试者序列具有至少x%的同一性”。当受试者序列具有可变位置(例如,表示为X的残基)时,与查询序列中的任何残基的比对被计算为匹配。

[0058] 如本文所使用的,“AAV载体”或“rAAV载体”是指包括一种或多种所关注的多核苷酸(或转基因)的重组载体,所关注的多核苷酸侧接AAV末端重复序列(ITR)。当存在于已被对rep和cap基因产物进行编码和表达的质粒转染的宿主细胞中时,这种AAV载体可以被复制并包装成感染性病毒颗粒。可替代地,可以使用已经被稳定地工程化以表达rep和cap基因的宿主细胞将AAV载体包装成感染性颗粒。

[0059] 如本文所使用的,“AAV病毒粒子”或“AAV病毒颗粒”或“AAV载体颗粒”是指由至少一种AAV衣壳蛋白和衣壳化的多核苷酸AAV载体构成的病毒颗粒。如本文所使用的,如果所述颗粒包括异源多核苷酸(即,除野生型AAV基因组以外的多核苷酸,如递送给哺乳动物细胞的转基因),则其通常被称为“AAV载体颗粒”或简单地称为“AAV载体”。因此,AAV载体颗粒的产生必然包含产生AAV载体,如此AAV载体颗粒内含有载体。

[0060] 如本文所使用的,“启动子”是指能够促进从真核细胞中的多核苷酸起始RNA转录的多核苷酸序列。

[0061] 如本文所使用的,“载体基因组”是指由载体(例如,rAAV病毒粒子)包装的多核苷酸序列,包含侧接序列(在AAV中,反向末端重复序列)。术语“表达盒”和“多核苷酸盒”是指侧接ITR序列之间的载体基因组的部分。“表达盒”意味着载体基因组包括编码与驱动表达的元件(例如,启动子)可操作地连接的基因产物的至少一个基因。

[0062] 如本文所使用的,术语“有需要的患者”或“有需要的受试者”是指有疾病、病症或病状风险或患有疾病、病症或病状的患者或受试者,所述疾病、病症或病状适于用本文公开的重组基因疗法载体或基因编辑系统治疗或改善。例如,有需要的患者或受试者可以是被诊断患有与中枢神经系统相关的病症。受试者可能具有EEF1A2基因突变或EEF1A2基因的全部或部分缺失,或基因调控序列缺失,这些导致eEF1A2蛋白表达异常。“受试者”和“患者”在本文中可互换使用。通过本文所描述的方法治疗的受试者可以是成人或儿童。受试者可能年龄不等。

[0063] 如本文所使用的,术语“变体”或“功能变体”可互换地是指与保留亲本蛋白的一种或多种期望活性的亲本蛋白相比具有一个或多个氨基酸取代、插入或缺失的蛋白质。

[0064] 如本文所使用的,“遗传破坏”是指基因中的部分或完全功能丧失或异常活动。例如,受试者可能遭受EEF1A2基因表达或功能的遗传破坏,该基因在受试者的至少某些细胞(例如,神经元)中降低表达或导致eEF1A2蛋白功能丧失或异常。

[0065] 如本文所使用的,“治疗”是指改善疾病或病症的一种或多种症状。术语“预防”是指延迟或中断疾病或病症的一种或多种症状的发作,或减缓eEF1A2相关神经系统疾病或病症的进展。

[0066] EEF1A2蛋白或多核苷酸

[0067] 本公开考虑了与伸长因子1- α 2(eEF1A2)蛋白相关的组合物和使用方法。已知EEF1A2中的各种突变(如图1所示)与神经系统病症有关,包含癫痫、智力残疾和/或自闭症。已经观察到遗传突变和新生突变。在一些情况下,杂合错义突变足以引起疾病。

[0068] eEF1A2的多肽序列如下:

[0069] MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKKAERER
GITIDISLWKFETTKYYITI IDAPGHRDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTRHALLAYTLGVK
QLIVGVNKMDSTEPAYSEKRYDEIVKEVSAYIKKIGYNPATVPFVPI SGWHGDNMLEPSNPMPWFKGWKVERKEGNA
SGVSLLEALDITLPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGILRPGMVVTFAPVNITTEVKSVMHHEALS
EALPGDNVGFNVKNSVKDIRRGVCGDSKSDPPQEAAQFTSQV IILNHPGQISAGYSPVIDCHTAHIACKFAELKE
KIDRRSGKKLEDNPKSLKSGDAAIVEMVPGKPMCVESFSQYPLGRFAVRDMRQTVAVGV IKNVEKKS GGAGKVTKS
AQAQKAGK

[0070] (SEQ ID NO:1)。

[0071] 在一些实施例中,eEF1A2蛋白包括与SEQ ID NO:1至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多肽序列。

[0072] 在一些实施例中,本公开提供了一种重组腺相关病毒(rAAV)病毒粒子,其包括衣壳和载体基因组,其中所述载体基因组包括与启动子操作性地连接的编码eEF1A2蛋白或其功能变体的多核苷酸序列。在一些实施例中,本公开提供了一种重组腺相关病毒(rAAV)病毒粒子,其包括衣壳和载体基因组,其中所述载体基因组包括与启动子操作性地连接的编码eEF1A2蛋白的多核苷酸序列。编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少

75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：

[0073] ATGGGCAAGGAGAAGACCCACATCAACATCGTGGTCATCGGCCACGTGGACTCCGGAAAGTCCACCACCACGGGCCACCTCATCTACAAATGCGGAGGTATTGACAAAAGGACCATTGAGAAGTTCGAGAAGGAGGCGGCTGAGATGGGGAAGGGATCCTTCAAGTATGCCTGGGTGCTGGACAAGCTGAAGGCGGAGCGTGAGCGCGGCATCACCATCGACATCTCCCTCTGGAAGTTCGAGACCACCAAGTACTACATCACCATCATCGATGCCCCGGCCACCGCGACTTCATCAAGAACATGATCACGGGTACATCCCAGGCGGACTGCGCAGTGCTGATCGTGGCGGGCGGGCGTGGGCGAGTTCGAGGCGGGCATCTCCAAGAATGGGCAGACGCGGGAGCATGCCCTGCTGGCCTACACGCTGGGTGTGAAGCAGCTCATCGTGGGCGTGAACAAAATGGACTCCACAGAGCCGGCTACACGAGAAAGCGCTACGACGAGATCGTCAAGGAAGTCAGCGCCTACATCAAGAAGATCGGCTACAACCCGGCCACCGTGCCCTTTGTGCCATCTCCGGCTGGCACGGTGACAACATGCTGGA GCCCTCCCCAACATGCCGTGGTTCAAGGGCTGGAAGGTGGAGCGTAAGGAGGGCAACGCAAGCGGCGTGTCCCTGCTGGAGGCCCTGGACACCATCCTGCCCCCACGCGCCCCACGGACAAGCCCCTGCGCCTGCCGCTGCAGGACGTGTAC AAGATTGGCGGCATTGGCACGGTGCCCGTGGGCCGGGTGGAGACCGGCATCCTGCGGCCGGGCATGGTGGTGACCTT TGCGCCAGTGAACATCACCCTGAGGTGAAGTCAGTGGAGATGCACCACGAGGCTCTGAGCGAAGCTCTGCCCCGGCG ACAACGTCGGCTTCAATGTGAAGAACGTGTGCGTGAAGGACATCCGGCGGGGCAACGTGTGTGGGGACAGCAAGTCT GACCCGCCGAGGAGGCTGCTCAGTTCACCTCCAGGTCATCATCCTGAACCACCCGGGGCAGATTAGCGCCGGCTA CTCCCCGGTCATCGACTGCCACACAGCCCACATCGCCTGCAAGTTTGGCGAGCTGAAGGAGAAGATTGACCGGCGCT CTGGCAAGAAGCTGGAGGACAACCCCAAGTCCCTGAAGTCTGGAGACGCGCCATCGTGGAGATGGTGCCGGGAAAG CCCATGTGTGTGGAGAGCTTCTCCAGTACCCGCTCTCGGCCGCTTCGCCGTGCGCGACATGAGGCAGACGGTGGC CGTAGGCGTCATCAAGAACGTGGAGAAGAAGAGCGGGCGCGCCGCAAGGTCACCAAGTCGGCGCAGAAGGCGCAGA AGGCGGGCAAG

[0074] (SEQ ID NO:2)。

[0075] 编码eEF1A2蛋白的多核苷酸序列可以是经密码子优化的。

[0076] 编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：ATGGGTAAAGA AAAACACATATTAATATAGTAGTAATCGGTCATGTTGACTCTGGAAAATCTACTACTACAGGACATTTGATTTATA AATGTGGAGGAATTGATAAAAAGAACAATAGAAAAATTTGAAAAAGAAGCTGCTGAAATGGGTAAAGGTAGTTTTAAA TATGCTTGGGTTTTGGATAAATTGAAAAGCTGAAAGAGAAAGAGGAATTACAATTGATATTTCTTTGTGGAAATTTGA AACTACAAAATATTATATAACAATAATAGATGCTCCTGGACATAGAGATTTTATTAATAATATGATTACAGGAACTT CTCAAGCAGATTGTGCTGTTTTGATAGTAGCAGCAGGAGTTGGTGAATTCGAAGCAGGCATTTCTAAAAATGGACAA ACTAGAGAACATGCTTTGTTGGCTTATACATTGGGCGTAAAACAATTGATTGTAGGAGTTAATAAAATGGATTCTAC TGAACCTGCATATTCTGAAAAAGATATGATGAAATAGTAAAAGAAGTTTCTGCTTATATTAATAAAATTTGGTTATA ATCCTGCTACAGTTCCATTTGTTCTATTTCTGGATGGCATGGAGATAATATGTTGGAACCTAGTCCTAATATGCCT TGGTTTTAAAGGATGGAAAGTTGAAAGGAAAAGAAGGAAATGCATCAGGAGTCTCCTTGTGGAAGCTTTGGATACAAT CTTGCCTCCAACAAGACCTACAGATAAACCTTTGAGATTGCCTCTCAAGATGTATATAAAATAGGAGGAATAGGAA CAGTGCCAGTTGGAAGAGTAGAAAACAGGTATATTGAGACCTGGAATGGTTGTAACATTTGCACCAGTTAATATAACT ACTGAAGTAAAATCTGTTGAAATGCATCATGAAGCTTTGTCTGAAGCTTCTCCTGGAGATAATGTAGGATTTAATGT TAAAAATGTAAGTGTAAAAGATATAAGAAGAGGAAATGTATGTGGTGATAGTAAATCAGATCCACCTCAAGAAGCAG CTCAATTTACATCACAAGTAATAATATTGAATCATCCTGGACAAATTTCTGCAGGATATTCACCAGTAATAGATTGT

CATACAGCACATATAGCTTGTAAATTTGCTGAATTGAAAGAAAAAATTGATAGAAGAAGTGGAAAAAACTTGAAGA
TAATCCTAAATCATTGAAATCAGGAGATGCAGCTATTGTAGAAATGGTACCTGGAAAACCAATGTGTGTAGAATCTT
TTTCTCAATATCCACCTCTCGGAAGATTTGCTGTTAGAGATATGAGACAAACAGTTGCAGTAGGAGTTATTAATAAAT
GTAGAAAAAAAAGCGGAGGTGCAGGAAAGGTTACAAAATCCGCACAAAAAGCTCAAAAAGCTGGTAAATAA

[0077] (SEQ ID NO:4)。

[0078] 编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：

[0079] ATGGGCAAAGAAAAACACATATAAACATTGTCGTTATCGGACACGTTGATTCTGGTAAAAGTACAACA
ACCGGTCACCTTGATATACAAATGCGGGGTATAGACAAAACGCACTATTGAAAAGTTCGAGAAAGAAGCTGCGGAGAT
GGGCAAAGGCTCATTCAAGTACGCGTGGTACTCGATAAGTTGAAAGCTGAACGCGAGAGGGGAATCACCATAGACA
TCTCACTTTGGAAATTCGAGACAACCAAGTATTACATAACTATTATAGATGCCCCAGGCCACAGGGATTTTATTA
AATATGATAACCGGCACATCTCAAGCCGATTGCGCCGACTCATCGTCGCCGCTGGTGTGGGTGAGTTCGAGGCAGG
TATTTCTAAAAATGGCCAGACACGCAACATGCTCTTCTGGCTTATACACTCGGGGTTAACAGCTCATAGTAGGAG
TGAATAAGATGGACTCCACTGAACCCGCTATTAGAGAAAGCGCTATGACGAAATTGTAAAGGAGGTCTCAGCATAT
ATTAATAAATTTGGCTATAACCCAGCCACGGTGCCATTCGTCCCATTAGTGGATGGCATGGTGACAATATGCTGGA
ACCAAGTCCCAATATGCCTTGGTTTAAGGGTTGAAAAGTAGAGCGGAAAGAGGGTAATGCTTCCGGCGTGTATTGC
TGGAGGCGCTTGACACGATACTCCCACCCACAAGGCCAACTGATAAGCCACTCCGATTGCCCTTGCAGGACGTGTAC
AAGATTGGGGGAATTGGGACTGTGCCCGTCGGGCGGTGGAGACGGGCATCCTCAGACCTGGGATGGTAGTCACTTT
TGCCCCCGTCAACATAACGACTGAAGTTAAATCAGTGAAAATGCATCACGAAGCTTTGAGTGAGGCGCTTCCCGGAG
ATAACGTTGGATTTAATGTCAAAAATGTCTCCGTTAAAGATATAAGAAGAGGAAACGTCTGCGGTGACTCAAAGTCA
GACCCACCACAGGAGGCTGCTCAATTTACGAGTCAAGTAATAATTCTGAATCACCCCTGGGCAAATAAGTGCGGGATA
CTCTCCAGTCATCGATTGTCACACCGCCCATATTGCATGTAAGTTCGCAGAACTTAAGGAAAAGATCGACCGAAGAA
GCGGAAAAAATTGGAAGATAATCCGAAAAGTTGAAAAGCGGTGACGCGCGATTGTAGAGATGGTCCCTGGCAAA
CCGATGTGTGTGGAGTCTTTCAGTCAATATCCACCACTCGGTCGCTTTGCCGTGCGGGATATGCGACAGACCGTTGC
TGTCGGCGTAATAAAAAACGTCGAAAAAAGAGCGGTGGGGCTGGAAAAGTTACAAAATCCGCTCAAAAGGCACAGA
AGGCGGGCAAGTGA

[0080] (SEQ ID NO:5)。

[0081] 编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：

[0082] ATGGGTAAAGAAAAGACCCACATTAACATAGTAGTAATCGGTCATGTTGACTCTGGGAAAAGCACTACT
ACCGGACATTTGATCTATAAATGTGGGGGCATCGACAAAAGAACGATAGAGAAGTTTGAGAAGGAGGCGGCGGAGAT
GGGTAAAGGTAGTTTTAAGTACGCTTGGGTTTTGGACAAATGAAAGCCGAGCGGAGCGCGCATTACCATTGACA
TTTCTCTCTGGAAATTCGAAACTACGAAGTATTATATAACAATAATAGACCCCCGGCCATCGGGACTTTATTA
AACATGATTACAGGAACTAGCCAAGCAGATTGTGCTGTGCTGATAGTAGCGGCAGGGGTGCGGGAGTTCGAAGCAGG
CATCTCTAAAAATGGACAACTCGAGAGCACGCCTTGTGGCTTATACCTTGGGCGTAAAGCAGCTGATCGTAGGAG
TTAATAAATGGATTCCACTGAACCCGCATATAGCGAAAAGCGATATGACGAAATAGTAAAGGAAGTCTCAGCTTAT
ATCAAGAAAATCGGTTACAATCCTGCGACGGTCCATTCGTTCCCTATCTCCGGGTGGCACGGCGATAATATGCTTGA
GCCAGTCCCAATATGCCCTGGTTCAAGGGGTGGAAGGTTGAGAGGAAGGAAGGCAATGCATCAGGCGTCAGCTTGT
TGGAAGCTCTCGACACCATCTGCCGCCACGAGGCCACAGACAAACCGTTGCGACTGCCTCTTCAAGATGTATAC

AAAATAGGCGGGATAGGAACCGTGCCGGTTGGACGAGTAGAGACGGGTATACTGCGGCCCGGAATGGTCGTGACGTT
TGCACCCGTGAATATACTACTGAGGTGAAGAGCGTCGAGATGCACCATGAAGCGCTGAGTGAAGCTCTCCCTGGCG
ATAACGTAGGGTTCAACGTGAAAAACGTAAGTGTAAAGGATATAAGGCGCGGAAATGTATGTGGTGACAGTAAAAGC
GACCCGCCGCAAGAGGCGGCGCAATTCACATCACAGGTAATAATATTGAATCACCCCGCCAAATTTCCGCAGGCTA
CTCACCAGTCATAGATTGCCACACCGCCACATAGCTTGTAAAGTTCGCTGAGTTGAAAGAGAAGATTGATAGACGAA
GTGGGAAGAACTTGAAGACAATCCGAAGTCCCTGAAGTCCGGTGACGCAGCGATTGTAGAAATGGTACCGGGCAAG
CCAATGTGTGTAGAGTCTTTCAGCCAGTACCCACCACTGGGGCGGTTTCGCGGTGCGAGACATGAGGCAAACGGTTGC
GGTCGGCGTCATTAATAAATGTCGAAAAAAGAGTGGCGGTGACAGGTAAGGTCACAAAAGCGCACAAAAGGCCCAGA
AAGCCGGTAAGTGA

[0083] (SEQ ID NO:6)。

[0084] 编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：

[0085] ATGGGAAAGGAAAAACTCACATAAACATTGTCGTCATCGGTCACGTAGACAGTGGCAAATCAACGACC
ACTGGACATCTCATCTATAAGTGTGGCGGTATTGACAAAACGCACTATCGAGAAATTCGAAAAGGAGGCTGCTGAGAT
GGGCAAAGGCTCTTTCAAGTACGCATGGGTCTTGATAAGCTGAAAGCGGAGCGAGAGAGAGGGATCACCATCGATA
TATCTCTGTGGAAATTTGAAACCACCAAGTACTACATCACAATTATTGATGCCCCAGGTCATAGGGATTTTATCAAG
AACATGATCACCGGACAAGCCAAGCCGACTGCGCAGTTCTCATAGTGGCGGCTGGAGTAGGGGAGTTTGAAGCAGG
GATATCTAAGAATGGACAGACCCGCGAGCACGCCTTGCTGGCCTACACCCTGGGAGTGAAGCAGCTCATAGTTGGCG
TCAATAAGATGGACAGCACCGAACCCGCTACAGTGAGAAGAGGTATGACGAGATTGTGAAGGAGGTTTCTGCTTAC
ATTAATAAAGATTGGCTATAACCCAGCTACTGTCCCATTCGTTCCAATCAGCGGCTGGCACGGTGATAACATGCTGGA
GCCTAGTCCCAACATGCCGTGGTTCAAGGGGTGGAAGGTTGAACGCAAGGAGGGGAATGCCTCAGGCGTTTCCCTGC
TGGAGGCCCTCGATACAATACTCCCCCGACCCGGCCTACAGATAAACCGCTGCGACTGCCTCTTCAGGACGTGTAT
AAAATCGGGGAATCGGCACAGTGGCCGTGGGCAGGGTAGAGACTGGCATCTTGCGGCCTGGAATGGTAGTCACCTT
TGCCCCGGTTAATATCACAACGGAGGTGAAATCTGTGGAGATGCATCACGAAGCACTGAGCGAGGCTCTGCCTGGTG
ACAACGTGGGATTTAACGTCAAAAAACGTGTCAGTCAAGGACATCCGCCGCGTAACGTTTGCAGGATTCTAAGTCC
GATCCCCCAGGAGGCAGCCCAATTTACCTCCCAAGTGATCATTCTGAATCACCCAGGCCAAATTTCCGCCGGTA
TTCCCCTGTGATTGACTGTACACAGCACACATCGCATGCAAATTCGCCGAACCTAAGGAGAAAATTGATCGGAGAA
GCGGTAAAAAACTGGAGGACAACCCAAAGTCCCTCAAGTCTGGGGATGCCGCCATCGTGGAGATGGTACCAGGCAAA
CCTATGTGCGTGGAAAGTTTTAGCCAGTACCCTCCACTGGGTCGCTTGTCTGTTCCGGATATGCGGCAGACAGTAGC
GGTTGGGGTCATAAAAAACGTCGAGAAAAAGAGCGGAGGAGCTGGGAAAGTTACCAAATCCGCACAGAAGGCACAAA
AAGCCGGAATGA

[0086] (SEQ ID NO:7)。

[0087] 编码eEF1A2蛋白的多核苷酸可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列：

[0088] ATGGGCAAAGAGAAAACACATATTAACATTGTTGTTATCGGGCACGTTGATAGCGGCAAGTCCACTACC
ACTGGCCATCTGATTTACAAGTGCGGCGGAATCGATAAACGAACTATTGAAAAGTTCGAAAAAGAAGCCGCCGAGAT
GGGAAAGGGCTCCTTTAAATACGCTTGGGTCTCGATAAACTCAAAGCAGAACGGGAGAGAGGAATCACCATCGATA
TATCCTTGTGGAAGTTGAAACTACAAAATATTACATTACCATCATTGATGCGCCTGGGCACCGCGACTTCATTAAG
AACATGATTACTGGCACCTCTCAAGCCGACTGCGCAGTGTGATCGTAGCCGAGGCGTCGGGGAGTTCGAAGCTGG

GATCAGCAAGAACGGGCAGACTAGGGAACACGCTCTGCTCGCATATACTCTTGGCGTGAAACAGTTGATCGTTGGCG
TGAACAAGATGGATTCAACTGAGCCTGCCTATTCTGAGAAAACGATACGACGAGATTGTGAAAGAGGTTTCAGCTTAC
ATCAAGAAAATTGGGTATAATCCCGCAACAGTTCCCTTCGTGCCCATCTCTGGGTGGCAGCGGACAACATGCTCGA
ACCATCCCCAAATATGCCATGGTTCAAGGGATGGAAGGTGGAGCGCAAAGAAGGCAACGCCTCCGGAGTGTCTCTGC
TCGAGGCCCTGGACACCATTCTGCCCCAACACGACCCACTGATAAGCCTCTGAGACTGCCACTGCAAGACGTTTAC
AAAATTGGGGGAATTGGAACCGTGCCTGTGGGTGGGTGGAAACCGGAATCCTCAGACCCGGCATGGTGGTCACCTT
CGCACCAGTGAATATAACGACAGAGGTCAAATCTGTGGAGATGCACCATGAGGCATTGAGCGAGGCACTCCCAGGAG
ACAACGTGGGTTTCAACGTGAAAAATGTCTCAGTTAAGGACATCCGACGCGGCAACGTGTGCGGAGATAGCAAATCT
GACCCCCCAGGAGGCCGCTCAATTCACAAGTCAGGTTATCATCCTTAATCACCCCTGGCCAAATATCTGCAGGCTA
CAGCCCCGTGATCGATTGTCACACAGCTCATATCGCCTGTAAATTTGCTGAACTCAAAGAAAAGATTGACCGCAGAT
CAGGAAAAAAGCTGGAGGACAACCCATAAAGTCTGAAAGTCCGGCGACGCTGCCATCGTGGAGATGGTCCCTGGGAAA
CCCATGTGCGTGGAGTCCTTTTCTCAGTACCCCCCTCTGGGACGATTGCGCGTGCACGACATGAGACAGACTGTGCG
CGTGGGCGTCATTAATAAATGTGGAATAAATAATCAGGAGGTGCAGGAAAGTGACAAAGAGTGCCAGAAAGCACAGA
AGGCTGGCAAGTGA

[0089] (SEQ ID NO:8)。

[0090] 任选地,编码载体基因组的多核苷酸序列可以包括Kozak序列,包含但不限于GCCACCATGG (SEQ ID NO:10)。Kozak序列可以与编码eEF1A2蛋白或其功能变体的多核苷酸序列重叠。例如,载体基因组可以包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列(Kozak带有下划线):

[0091] gccaccATGGGCAAGGAGAAGACCCACATCAACATCGTGGTCATCGGCCACGTGGACTCCGAAAGTC
CACCACCACGGCCACCTCATCTACAAATGCGGAGGTATTGACAAAAGGACCATTGAGAAGTTCGAGAAGGAGGCGG
CTGAGATGGGGAAGGGATCCTTCAAGTATGCCTGGGTGCTGGACAAGCTGAAGGCGGAGCGTGAGCGCGGCATCACC
ATCGACATCTCCCTCTGGAAGTTCGAGACCACCAAGTACTACATCACCATCATCGATGCCCCCGCCACCGCGACTT
CATCAAGAACATGATCACGGGTACATCCCAGGCGGACTGCGCAGTGCTGATCGTGGCGGCGGGCGTGGGCGAGTTCCG
AGGCGGGCATCTCCAAGAATGGGCAGACGCGGGAGCATGCCCTGCTGGCTACACGCTGGGTGTGAAGCAGCTCATC
GTGGGCGTGAACAAAATGACTCCACAGACCGGCCTACAGCGAGAAGCGCTACGACGAGATCGTCAAGGAAGTCAG
CGCCTACATCAAGAAGATCGGCTACAACCCGGCCACCGTGCCTTTGTGCCATCTCCGGCTGGCACGGTGACAACA
TGCTGGAGCCCTCCCCAACATGCCGTGGTTCAAGGGCTGGAAGGTGGAGCGTAAGGAGGGCAACGCAAGCGGCGTG
TCCCTGCTGGAGGCCCTGGACACCATCCTGCCCCCACGCGCCCCACGGACAAGCCCCTGCGCCTGCCGCTGCAGGA
CGTGTACAAGATTGGCGGCATTGGCACGGTGCCGTGGGCGGGTGGAGACCGGCATCCTGCGGCCGGGCATGGTGG
TGACCTTTGCGCCAGTGAACATCACCACTGAGGTGAAGTCAGTGGAGATGCACCACGAGGCTCTGAGCGAAGCTCTG
CCCGGCGACAACGTGGCTTCAATGTGAAGAACGTGTCGGTGAAGGACATCCGGCGGGCAACGTGTGTGGGGACAG
CAAGTCTGACCCGCCGAGGAGGCTGCTCAGTTCACCTCCCAGGTCATCATCCTGAACCACCCGGGGCAGATTAGCG
CCGGCTACTCCCGGTCATCGACTGCCACACAGCCACATCGCCTGCAAGTTTGGGAGCTGAAGGAGAAGATTGAC
CGGCGCTCTGGCAAGAAGCTGGAGGACAACCCAAAGTCCCTGAAGTCTGGAGACGCGGCCATCGTGGAGATGGTGC
GGGAAAGCCCATGTGTGTGGAGAGCTTCTCCAGTACCCGCTCTCGGCCGCTTCGCCGTGCGCGACATGAGGCAGA
CGGTGGCCGTAGGCGTCATCAAGAACGTGGAGAAGAAGAGCGGCGGCCGGCAAGGTCACCAAGTCGGCGCAGAAG
GCGCAGAAGGCGGGCAAG

[0092] (SEQ ID NO:9)。

[0093] 在一些实施例中,所述Kozak序列是替代的Kozak序列,其包括以下任一项或由以下任一项组成:

[0094] (gcc)gccRccAUGG(SEQ ID NO:11);

[0095] AGNNAUGN;

[0096] ANNAUGG;

[0097] ACCAUGG;

[0098] GACACCAUGG(SEQ ID NO:12)。

[0099] 在一些实施例中,载体基因组不包括Kozak序列。

[0100] 载体基因组

[0101] 本公开的AAV病毒粒子包括载体基因组。载体基因组可以包括表达盒(或用于不需要多核苷酸序列的表达的基因编辑应用的多核苷酸盒)。可以使用任何合适的反向末端重复序列(ITR)。ITR可以来自与衣壳相同的血清型或不同的血清型(例如,可以使用AAV2 ITR)。

[0102] 在一些实施例中,5' ITR包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0103] CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCCT

[0104] (SEQ ID NO:18)

[0105] 在一些实施例中,5' ITR包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0106] GCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCCTTGTAGTTAATGATTAACCGCCATGCTACTTATCTACGTA

[0107] (SEQ ID NO:19)

[0108] 在一些实施例中,5' ITR包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0109] CTGCGCGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCCTTGTAGTTAATGATTAAACCGCCATGCTACTTATCTACGTA

[0110] (SEQ ID NO:20)

[0111] 在一些实施例中,3' ITR包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0112] AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCGGGCTTTGCCCGGGCGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG

[0113] (SEQ ID NO:21)

[0114] 在一些实施例中,3' ITR包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0115] TACGTAGATAAGTAGCATGGCGGGTTAATCATTAACCTACAAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCCGACGCCGGGCTTTGCCCGGGCGGC

CTCAGTGAGCGAGCGAGCGCGC

[0116] (SEQ ID NO:63)

[0117] 在一些实施例中,载体基因组包括与以下序列例如至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的一个或多个填充序列:

[0118] GCGGCAATTCAGTCGATAACTATAACGGTCCTAAGGTAGCGATTTAAATACGCGCTCTCTTAAGGTAGCCCCGGGACGCGTCAATTGACTACAAACCGAGTATCTGCAGAGGGCCCTGCGTATG (SEQ ID NO:22);

[0119] CTTCTGAGGCGGAAAGAACCAGATCCTCTCTTAAGGTAGCATCGAGATTTAAATTAGGGATAACAGGGTAATGGCGCGGGCCGC (SEQ ID NO:23);或

[0120] GTTACCCAGGCTGGAGTGCAGTGGCACATTTCTGCTCACTGCAACCTCCTCCTCCCTGGGTTC (SEQ ID NO:24)。

[0121] 启动子

[0122] 在一些实施例中,编码eEF1A2蛋白或其功能变体的多核苷酸序列与启动子可操作地连接。

[0123] 本公开考虑了各种启动子的用途。可用于本公开的实施例中的启动子包括但不限于巨细胞病毒(CMV)启动子、磷酸甘油酸激酶(PGK)启动子或包括CMV增强子和鸡 β -肌动蛋白启动子和兔 β -珠蛋白基因(CAG)的部分的启动子序列。在一些情况下,启动子可以是合成启动子。Schlabach等人《美国国家科学院院刊(PNAS USA.)》107(6):2538-43(2010)提供了示例性合成启动子。在一些实施例中,启动子包括与以下序列至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列:

[0124] ACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTA
TGTTCCCATAGTAACCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGG
CAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT
GCCCAGTACATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGA
GGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGTATTTATTTATTTT
AATTATTTTGTGCAGCGATGGGGCGGGGGGGGGGGGGCGCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGG
CGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGC
GGCGGCGGCGCCCTATAAAAAGCGAAGCGCGCGGGCGGGCGG

[0125] (SEQ ID NO:14)

[0126] 在一些实施例中,编码eEF1A2蛋白或其功能变体的多核苷酸序列与诱导型启动子操作性地连接。诱导型启动子可以被配置成响应于药剂的添加或累积或响应于药剂的去除、降解或稀释而使多核苷酸序列被转录表达或不被转录表达。所述药剂可以是药物。所述药剂可以是四环素或其衍生物中的一种,包含但不限于多西环素(doxycycline)。在一些情况下,诱导型启动子是tet-on启动子、tet-off启动子、化学调节启动子、物理调节启动子(即,响应于光的存在或不存在或响应于低温或高温的启动子)。诱导型启动子包含重金属离子诱导型启动子(如小鼠乳腺肿瘤病毒(mMTV)启动子或各种生长激素启动子)以及在存在T7 RNA聚合酶的情况下具有活性的来自T7噬菌体的启动子。所述诱导型启动子的列表是非限制性的。

[0127] 在一些情况下,启动子是组织特异性启动子,如能够在神经元中比在非神经元细胞中更大程度地驱动表达的启动子。在一些实施例中,组织特异性启动子选自任何各种神

神经元特异性启动子,包含但不限于:hSYN1(人突触蛋白)、INA(α -丝联蛋白)、NES(巢蛋白)、TH(酪氨酸羟化酶)、FOXA2(叉头盒A2)、CaMKII(钙调蛋白依赖性蛋白激酶II)和NSE(神经元特异性烯醇化酶)。在一些情况下,启动子是普遍存在的启动子。“普遍存在的启动子”是指在实验或临床条件下非组织特异性的启动子。在一些情况下,普遍存在的启动子是以下中的任一种:CMV、CAG、UBC、PGK、EF1- α 、GAPDH、SV40、HBV、鸡 β -肌动蛋白和人 β -肌动蛋白启动子。

[0128] 在一些实施例中,所述启动子序列选自表3。在一些实施例中,启动子包括与SEQ ID NO:3、14、16-17和25-30中任一者至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列。

[0129] 表3

[0130]

启动子	SEQ ID NO:
人 β -肌动蛋白(HuBa)	25
鸡 β -肌动蛋白(CBA)	26
巨细胞病毒(CMV)	16
巨细胞病毒(CMV)(第二个版本)	17
CAG启动子	14
人EF1- α (EF1- α)	27
人突触蛋白1(Syn),短版本	28
人突触蛋白1(Syn),具有3'延伸	3
人突触蛋白1(Syn),具有5'延伸	29
人CamKIIa(CaMKIIa)	30
eSYN启动子	64

[0131] 在优选的实施例中,载体基因组包括与SEQ ID NO:3至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列。

[0132] 启动子的另外的说明性实例是来自猿猴病毒40的SV40晚期启动子、杆状病毒多面体增强子/启动子元件、单纯性疱疹病毒胸苷激酶(HSV tk)、来自巨细胞病毒(CMV)的立即早期启动子和包含LTR元件的各种逆转录病毒启动子。各种各样的其它启动子是已知的,并且在本领域中通常是可获得的,并且许多这种启动子的序列可在如基因库数据库等序列数据库中获取。

[0133] 其它调节元件

[0134] 在一些情况下,本公开的载体进一步包括选自由以下组成的组的一种或多种调节元件:增强子、内含子、poly-A信号、2A肽编码序列、土拨鼠肝炎病毒转录后调节元件(WPRE)和乙型肝炎转录后调节元件(HPRE)。

[0135] 在一些实施例中,载体包括CMV增强子。

[0136] 在某些实施例中,载体包括一种或多种增强子。在特定实施例中,增强子是CMV增强子序列、GAPDH增强子序列、 β -肌动蛋白增强子序列或EF1- α 增强子序列。前述的序列在本领域中是已知的。例如,CMV立即早期(IE)增强子的序列是:

[0137] ACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTA
TGTTCCCATAGTAACCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTGG

CAGTACATCAAGTGTATCATATGCCAAGTACGCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT
GCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA

[0138] (SEQ ID NO:31)

[0139] 在某些实施例中,载体包括一个或多个内含子。在特定实施例中,内含子是兔珠蛋白内含子序列、鸡 β -肌动蛋白内含子序列、合成内含子序列或EF1- α 内含子序列。

[0140] 在某些实施例中,载体包括polyA序列。在特定实施例中,polyA序列是兔珠蛋白polyA序列、人生长激素polyA序列、牛生长激素polyA序列、PGK polyA序列、SV40 polyA序列或TK polyA序列。在一些实施例中,poly-A信号可以是牛生长激素多腺苷酸化信号(bGHpA)。

[0141] 在某些实施例中,载体包括一个或多个转录稳定元件。在特定实施例中,转录稳定元件是WPRE序列、HPRE序列、支架附着区、3'UTR或5'UTR。在特定实施例中,载体包括5'UTR和3'UTR两者。

[0142] 在一些实施例中,载体包括选自表4的5'非翻译区(UTR)。在一些实施例中,载体基因组包括与SEQ ID NO:32-40中任一者至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列。

[0143] 表4

5'非翻译区	SEQ ID NO:
人 β -肌动蛋白外显子/内含子	32
鸡 β -肌动蛋白外显子/内含子+兔珠蛋白内含子	33
5'UTR-Syn1 Hs	34
CMV IE外显子	35
TPL-eMLP(腺病毒源性增强子元件)	36
人EF1- α 内含子/外显子	37
人EF1- α ,内含子A	38
5'UTR人CamKIIa	39
β -珠蛋白内含子	40

[0145] 在一些实施例中,载体包括选自表5的3'非翻译区。在一些实施例中,载体基因组包括与SEQ ID NO:41-49中任一者至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列。

[0146] 表5

3'非翻译区	SEQ ID NO:
WPRE(x)(突变的土拨鼠肝炎调节元件-版本1)	41
WPRE(x)(突变的土拨鼠肝炎调节元件-版本2)	42
WPRE(x)(突变的土拨鼠肝炎调节元件-版本3)	43
CAAX	44
EES	45
HPRE	46
R2V17(HepB源性增强子元件)	47
3'UTR(珠蛋白)	48

WPRE (r)	49
----------	----

[0148] 在一些实施例中,载体包括选自表6的多腺苷酸化 (polyA) 信号。在一些实施例中, polyA信号包括与SEQ ID NO:50-54中任一者至少75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、99%或100%相同的多核苷酸序列。

[0149] 表6

多腺苷酸化位点	SEQ ID NO:
兔珠蛋白 (pA珠蛋白-0c)	50
牛生长激素 (pAGH-Bt-版本1)	51
牛生长激素 (pAGH-Bt-版本2)	52
牛生长激素 (pAGH-Bt-版本3)	53
人生长激素 (pAGH-Hs)	54

[0151] 说明性载体基因组在图2-5中描绘并被提供为SEQ ID NO:55-58或65-68。在一些实施例中,载体基因组包括与SEQ ID NO:55-58或65-68中任一者具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的多核苷酸序列、基本上由其组成或由其组成,任选地具有或不具有小写的ITR序列。编码序列带有下划线。替代性载体基因组序列以SEQ ID NO:65-68的形式提供。

[0152] V1-载体基因组-3,144bp(图2) (SEQ ID NO:55)

cctgcagggcagctgcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgctc
gggcgacctttggtcgccccggcctcagtgagcgagcgagcgcgagagagggagtggcca
actccatcactaggggttcctGCGGCAATTCAGTCGATAACTATAACGGTCCTAAGGTAG
CGATTTAAATACGCGCTCTCTTAAGGTAGCCCCGGGACGCGTCAATTGACTACAAACCGA
GTATCTGCAGAGGGCCCTGCGTATGAGTGCAAGTGGGTTTTAGGACCAGGATGAGGCGGG
GTGGGGGTGCCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCCAACCCCATTC
CCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAACAGGATGCGGCGAGGCGCGT
GCGCACTGCCAGCTTCAGCACCGCGGACAGTGCCTTCGCCCCGCCTGGCGGCGCGCGCC
ACCGCCGCCTCAGCACTGAAGGCGCGCTGACGTCACTCGCCGGTCCCCCGCAAACCTCCCC
TTCCCGGCCACCTTGGTCGCGTCCGCGCCGCGCCGGCCAGCCGGACCGCACCCACGCGA
GGCGCGAGATAGGGGGGCACGGGCGCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGC
TGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCGTGCCTGAGAGCGCAGATGGGCA
[0153] AGGAGAAGACCCACATCAACATCGTGGTCATCGGCCACGTGGACTCCGGAAAGTCCACCA
CCACGGGCCACCTCATCTACAAATGCGGAGGTATTGACAAAAGGACCATTGAGAAGTTCG
AGAAGGAGGCGGCTGAGATGGGGAAGGGATCCTTCAAGTATGCCTGGGTGCTGGACAAGC
TGAAGGCGGAGCGTGAGCGCGGCATCACCATCGACATCTCCCTCTGGAAGTTCGAGACCA
CCAAGTACTACATCACCATCATCGATGCCCCCGGCCACCGCGACTTCATCAAGAACATGA
TCACGGGTACATCCCAGGCGGACTGCGCAGTGCTGATCGTGGCGGCGGGCGTGGGCGAGT
TCGAGGCGGGCATCTCCAAGAATGGGCAGACGCGGGAGCATGCCCTGCTGGCCTACACGC
TGGGTGTGAAGCAGCTCATCGTGGGCGTGAACAAAATGGACTCCACAGAGCCGGCCTACA
GCGAGAAGCGCTACGACGAGATCGTCAAGGAAGTCAGCGCCTACATCAAGAAGATCGGCT
ACAACCCGGCCACCGTGCCCTTTGTGCCATCTCCGGCTGGCACGGTGACAACATGCTGG
AGCCCTCCCCAACATGCCGTGGTTCAAGGGCTGGAAGGTGGAGCGTAAGGAGGGCAACG
CAAGCGGCGTGTCCCTGCTGGAGGCCCTGGACACCATCCTGCCCCCACGCGCCCCACGG
ACAAGCCCCTGCGCCTGCCGCTGCAGGACGTGTACAAGATTGGCGGCATTGGCACGGTGC

CCGTGGGCCGGGTGGAGACCGGCATCCTGCGGCCGGGCATGGTGGTGACCTTTGCGCCAG
TGAACATCACCCTGAGGTGAAGTCAGTGGAGATGCACCACGAGGCTCTGAGCGAAGCTC
TGCCCGGCGACAACGTCGGCTTCAATGTGAAGAACGTGTCGGTGAAGGACATCCGGCGGG
GCAACGTGTGTGGGGACAGCAAGTCTGACCCGCCGCAGGAGGCTGCTCAGTTCACCTCCC
AGGTCATCATCCTGAACCACCCGGGGCAGATTAGCGCCGGCTACTCCCCGGTCATCGACT
GCCACACAGCCACATCGCCTGCAAGTTTGC GGAGCTGAAGGAGAAGATTGACCGGCGCT
CTGGCAAGAAGCTGGAGGACAACCCCAAGTCCCTGAAGTCTGGAGACGCGGCCATCGTGG
AGATGGTGCCGGGAAAGCCCATGTGTGTGGAGAGCTTCTCCAGTACCCGCCTCTCGGCC
GCTTCGCCGTGCGCGACATGAGGCAGACGGTGGCCGTAGGCGTCATCAAGAACGTGGAGA
AGAAGAGCGGCGGCCGGCAAGGTCACCAAGTCGGCGCAGAAGGCGCAGAAGGCGGGCA
AGTGAAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATG
TTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTT
CCCGTATGGCTTTTCATTTTCTCCTTGTATAAATCCTGGTTGCTGTCTTTATGAGG
AGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTGCTGACGCAACCC
[0154] CCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTTCGCTTTCCCC
TCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTC
GGCTGTTGGGCACTGACAATTCGCTGGTGTGTCGGGGAAATCATCGTCCTTTCTTTGGC
TGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCTTTCGG
CCCTCAATCCAGCGGACCTTCCTTCCC GCGCCTGCTGCCGGCTCTGCGGCCTCTTCCGC
GTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCCTGTGC
CTTCTAGTTGCCAGCCATCTGTTGTTTGCCCTCCCCCGTGCCTTCTTGACCCTGGAAG
GTGCCACTCCCCTGTCCTTTCTAATAAAATGAGGAAATGCATCGCATTGTCTGAGTA
GGTGTCAATCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAG
ACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCA
GATCCTCTCTTAAGGTAGCATCGAGATTTAAATTAGGGATAACAGGGTAATGGCGCGGGC
CGCaggaaccctagtgatggagttggccactccctctctgcgcgctcgctcgctcactg
aggccgggcgaccaaaggtcgcccgacgcccgggctttgcccgggcggcctcagtgagcg
agcgagcgcgagctgcctgcagg

[0155] V2-载体基因组-3,035bp(图3) (SEQ ID NO:56)

gcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggogtcgggcgacctttgg

[0156] tcgcccggcctcagtgagcgagcgagcgcgagagagggagtgcccaactccatcactag
gggttccttgtagttaatgattaaccgcatgctacttatctacgtaAGTGCAAGTGGG

TTTTAGGACCAGGATGAGGCGGGGTGGGGTGCCTACCTGACGACCGACCCCGACCCACT
 GGACAAGCACCCAACCCCATTCCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGG
 AAACAGGATGCGGCGAGGCGCGTGCGCACTGCCAGCTTCAGCACCGCGGACAGTGCCTTC
 GCCCCGCCTGGCGGCGCGGCCACCGCCGCCTCAGCACTGAAGGCGCGCTGACGTCACT
 CGCCGGTCCCCGCAAACCTCCCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGCCGCCGG
 CCCAGCCGGACCGCACACCACGCGAGGCGGAGATAGGGGGGCACGGGCGCGACCATCTGCG
 CTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTC
 GTGCCTGAGAGCGCAGGCCACCATGGGCAAGGAGAAGACCCACATCAACATCGTGGTCAT
 CGGCCACGTGGACTCCGGAAAGTCCACCACCACGGGCCACCTCATCTACAAATGCGGAGG
 TATTGACAAAAGGACCATTGAGAAGTTCGAGAAGGAGGCGGCTGAGATGGGGAAGGGATC
 CTTCAAGTATGCCTGGGTGCTGGACAAGCTGAAGGCGGAGCGTGAGCGCGGCATCACCAT
 CGACATCTCCCTCTGGAAGTTCGAGACCACCAAGTACTACATCACCATCATCGATGCCCC
 CGGCCACCGCGACTTCATCAAGAACATGATCACGGGTACATCCCAGGCGGACTGCGCAGT
 GCTGATCGTGGCGGCGGGCGTGGGCGAGTTCGAGGCGGGCATCTCCAAGAAATGGGCAGAC
 GCGGGAGCATGCCCTGCTGGCCTACACGCTGGGTGTGAAGCAGCTCATCGTGGGCGTGAA
 CAAAATGGACTCCACAGAGCCGGCCTACAGCGAGAAGCGCTACGACGAGATCGTCAAGGA
 AGTCAGCGCCTACATCAAGAAGATCGGCTACAACCCGCGCCACCGTGCCTTTGTGCCCAT
 CTCCGGCTGGCACGGTGACAACATGCTGGAGCCCTCCCCAACATGCCGTGGTTCAAGGG
 CTGGAAGGTGGAGCGTAAGGAGGGCAACGCAAGCGGCGTGTCCCTGCTGGAGGCCCTGGA
 CACCATCCTGCCCCCACGCGCCCCACGGACAAGCCCCGCGCCTGCCGCTGCAGGACGT
 GTACAAGATTGGCGGCATTGGCACGGTGCCCGTGGGCGGGTGGAGACCGGCATCCTGCG
 GCCGGGCATGGTGGTGACCTTTGCGCCAGTGAACATCACCACTGAGGTGAAGTCAGTGGA
 GATGCACCACGAGGCTCTGAGCGAAGCTCTGCCCGGCGACAACGTGGCTTCAATGTGAA
 GAACGTGTGCGGTGAAGGACATCCGGCGGGGCAACGTGTGTGGGGACAGCAAGTCTGACCC
 GCCGCAGGAGGCTGCTCAGTTCACCTCCCAGGTCATCATCCTGAACCACCCGGGGCAGAT
 TAGCGCCGGCTACTCCCCGGTCATCGACTGCCACACAGCCCACATCGCCTGCAAGTTTGC
 GGAGCTGAAGGAGAAGATTGACCGGCGCTCTGGCAAGAAGCTGGAGGACAACCCCAAGTC
 CCTGAAGTCTGGAGACGCGGCCATCGTGGAGATGGTGCCGGGAAAGCCCATGTGTGTGGA
 GAGCTTCTCCAGTACCCGCCTCTCGGCCGCTTCGCCGTGCGCGACATGAGGCAGACGGT
 GGCCGTAGGCGTCATCAAGAACGTGGAGAAGAAGAGCGGCGGCGCCGGCAAGGTCACCAA
 GTCGGCGCAGAAGGCGCAGAAGGCGGGCAAGTGATCAACCTCTGGATTACAAAATTTGTG
 AAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTT

[0157]

TAATGCCTTTGTATCATGCTATTGCTTCCCCTATGGCTTTCATTTTCTCCTCCTTGTATA
AATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGG
TGTGCACTGTGTTTGTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGC
TCCTTTCCGGGACTTTCGCTTTCCTCCCTATGCCACGGCGGAACATCGCCGCCT
GCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGT
CGGGGAAATCATCGTCCTTTCCTTGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCG
GGACGTCCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCC
TGCTGCCGGCTCTGCGGCCTCTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCT
[0158] CCCTTTGGGCCGCTCCCCGCACTGCCCGGGTGGCATCCCTGTGACCCCTCCCCAGTGCC
TCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTAATAAAAATTAAG
TTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGTGGT
ATGGAGCAAGGGGCCAAGTTGGGAAGAAACCTGTAGGGCCTGCGTTACCCAGGCTGGAG
TGCAGTGGCACATTTCTGCTCACTGCAACCTCCTCCTCCCTGGGTTctacgtagataagt
agcatggcgggttaatcattaactacaaggaaccctagtgatggagttggccactccct
ctctgcgcgctcgctcgctcactgaggccgggaccaaaggtegcccgacgcccgggct
ttgcccgggcggcctcagtgagcgagcgagcgcg
[0159] V3-载体基因组-3,263bp(图4)(SEQ ID NO:57)
gcgcgctcgctcgctcactgaggccgcccgggcaaagcccgggcgctcgggacacctttgg
tcgcccggcctcagtgagcgagcgagcgcgagagagggagtgcccaactccatcactag
gggttcctttagttaaataacccgcatgctacttatctacgtaAGTGCAAGTGGG
TTTTAGGACCAGGATGAGGCGGGGTGGGGTGCCTACCTGACGACCGACCCGACCCACT
GGACAAGCACCCAACCCCATTCCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGG
AAACAGGATGCGGCGAGGCGCGTGCACACTGCCAGCTTCAGCACCGCGGACAGTGCCTTC
GCCCCGCTGGCGGCGCGGCCACCGCCGCTCAGCACTGAAGGCGGCTGACGTCACT
[0160] CGCCGGTCCCCGCAAACCTCCCCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGCGCCGG
CCCAGCCGGACCGCACACGCGAGGCGGAGATAGGGGGCACGGGCGGACCATCTGCG
CTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCGAGGAGTTCGTGTC
GTGCCTGAGAGCGCAGAGTCTGCGGTGGGCGAGGAGTTCGTGTCGTGCCTGAGAGCG
CAGCTGTGCTCCTGGGACCGCGCAGTCCGCCCCGCGGCTCCTGGCCAGACCACCCCTA
GGACCCCTGCCCAAGTCGCAGCCACCATGGGCAAGGAGAAGACCACATCAACATCGT
GGTCATCGGCCACGTGGACTCCGGAAAGTCCACCACCACGGGCCACCTCATCTACAAATG
CGGAGGTATTGACAAAAGGACCATTGAGAAGTTCGAGAAGGAGGCGGCTGAGATGGGGAA

GGGATCCTTCAAGTATGCCTGGGTGCTGGACAAGCTGAAGGCGGAGCGTGAGCGCGGCAT
CACCATCGACATCTCCCTCTGGAAGTTCGAGACCACCAAGTACTACATCACCATCATCGA
TGCCCCGGCCACCGCGACTTCATCAAGAACATGATCACGGGTACATCCCAGGCGGACTG
CGCAGTGCTGATCGTGGCGGCGGGCGTGGGCGAGTTCGAGGCGGGCATCTCCAAGAATGG
GCAGACGCGGGAGCATGCCCTGCTGGCCTACACGCTGGGTGTGAAGCAGCTCATCGTGGG
CGTGAACAAAATGGACTCCACAGAGCCGGCCTACAGCGAGAAGCGCTACGACGAGATCGT
CAAGGAAGTCAGCGCCTACATCAAGAAGATCGGCTACAACCCGGCCACCGTGCCCTTTGT
GCCATCTCCGGCTGGCACGGTGACAACATGCTGGAGCCCTCCCCAACATGCCGTGGTT
CAAGGGCTGGAAGGTGGAGCGTAAGGAGGGCAACGCAAGCGGCGTGTCCCTGCTGGAGGC
CCTGGACACCATCCTGCCCCCACGCGCCCCACGGACAAGCCCCTGCGCCTGCCGCTGCA
GGACGTGTACAAGATTGGCGGCATTGGCACGGTGCCCGTGGGCCGGGTGGAGACCGGCAT
CCTGCGGCCGGGCATGGTGGTGACCTTTGCGCCAGTGAACATCACCCTGAGGTGAAGTC
AGTGGAGATGCACCACGAGGCTCTGAGCGAAGCTCTGCCCGGCGACAACGTCGGCTTCAA
TGTGAAGAACGTGTTCGGTGAAGGACATCCGGCGGGGCAACGTGTGTGGGGACAGCAAGTC
TGACCCGCCGCAGGAGGCTGCTCAGTTCACCTCCAGGTCATCATCCTGAACCACCCGGG
GCAGATTAGCGCCGGCTACTCCCCGGTTCATCGACTGCCACACAGCCACATCGCCTGCAA
GTTTGC GGAGCTGAAGGAGAAGATTGACCGGCGCTCTGGCAAGAAGCTGGAGGACAACCC
CAAGTCCCTGAAGTCTGGAGACGCGGCCATCGTGGAGATGGTGCCGGGAAAGCCCATGTG
TGTGGAGAGCTTCTCCAGTACCCGCCTCTCGGCCGCTTCGCCGTGCGCGACATGAGGCA
GACGGTGGCCGTAGGCGTCATCAAGAACGTGGAGAAGAAGAGCGGCGGCGCCGGCAAGGT
CACCAAGTCGGCGCAGAAGGCGCAGAAGGCGGGCAAGTGATCAACCTCTGGATTACAAAA
TTTGTGAAAGATTGACTGGTATTCTTAACCTATGTTGCTCCTTTTACGCTATGTGGATACG
CTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCT
TGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCGTTGTCAGGCAACGTG
GCGTGGTGTGCACTGTGTTTGTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCT
GTCAGCTCCTTTCCGGGACTTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACCTCATCG
CCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
TGTGTTCGGGAAATCATCGTCCTTTTCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTTC
TGCGCGGGACGTCCTTCTGCTACGTCCTTTCGGCCCTCAATCCAGCGGACCTTCCCTCCC
GCGGCCTGCTGCCGGCTCTGCGGCCTTTCGGCGTCTTCGCCTTCGCCCTCAGACGAGTC
GGATCTCCCTTTGGGCCGCTCCCCGCAGCTGGAGCCTCGGTAGCCGTTCCCTCCTGCCCG
CTGGGCCTCCCAACGGGCCCTCCTCCCCTCCTTGCACCGGCCCTTCCCTGGTCTTTGAATA

[0161]

[0162] AATTCATTGCCTGCCCGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCT
GGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCTAATAAAAATTAAGTTGCATCATT
GTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGG
GCCCAAGTTGGGAAGAAACCTGTAGGGCCTGCGTTACCCAGGCTGGAGTGCAGTGGCACA
TTTCTGCTCACTGCAACCTCCTCCTCCCTGGGTTCtacgtagataagtagcatggcgggt
taatcattaactacaaggaaccctagtgatggagttggccactccctctctgcgcgctc
gctcgctcactgaggccgggaccaaaggtcgcccagcggcgggctttgcccgggacg
cctcagtgagcgagcgagcgcg

[0163] V4-载体基因组-4,299bp(图5)(SEQ ID NO:58)
gcgcgctcgctcgctcactgaggccgggcaaagcccgggagtcgggacgacctttgg
tcggccgacctcagtgagcgagcgagcgcgagagagggagtgcccaactccatcactag
gggttccttgtagttaatgattaaccgcatgctacttatctacgtaCTCTGGAGACGC
GTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCATTG
ACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTCCATTGACGTCAA
TGGGTGGAGTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCA
AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTAC
ATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC
ATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCTCCCCACCC
CCAATTTTGTATTTATTTATTTTAAATTTTGTGCAGCGATGGGGCGGGGGGGGGG
GGGGCGCGCCAGGCGGGGCGGGGCGAGGGGCGGGGCGAGGCGGGGCGAGGCGGAGAGG

[0164] TCGGCGGCAGCCAATCAGAGCGGCGGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCG
GCGGCGGCGGCCCTATAAAAAGCGAAGCGCGGCGGGCGGGAGTCGCTGCGCGCTGCCT
TCGCCCCGTGCCCGCTCCGCCGCCGCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGC
GTTACTCCCACAGGTGAGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTT
GGTTAATGACGGCTTGTCTTTCTGTGGCTGCGTGAAAGCCTTGAGGGGCTCCGGGA
GGGCCCTTTGTGCGGGGGAGCGGCTCGGGGGTGCCTGCGTGTGTGTGTGCGTGGGGAG
CGCCGCGTGC GGCTCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCGGGGCT
TTGTGCGCTCCGCAGTGTGCGGAGGGGAGCGCGGCCGGGGCGGTGCCCGCGGTGCGG
GGGGGCTGCGAGGGGAACAAAGGCTGCGTGC GGGTGTGTGCGTGGGGGGTGTGAGCAGG
GGGTGTGGGCGGTCGGTCCGGCTGCAACCCCCCTGCACCCCCCTCCCCGAGTTGCTGA
GCACGGCCCGGCTTCGGGTGCGGGGCTCCGTACGGGGCGTGGCGGGGGCTCGCCGTGCC
GGGCGGGGGTGGCGGCAGGTGGGGGTGCCGGGCGGGGCGGGGCCCTCGGGCCGGGA

GGGCTCGGGGAGGGGCGCGCGGCCCGGAGCGCCGGCGGCTGTCGAGGCGCGGCGAG
CCGCAGCCATTGCCTTTTATGGTAATCGTGCGAGAGGGGCGCAGGGACTTCCTTTGTCCCA
AATCTGTGCGGAGCCGAAATCTGGGAGGCGCCGCGCACCCCTCTAGCGGGCGCGGGGC
GAAGCGGTGCGGCGCCGGCAGGAAGGAAATGGGCGGGGAGGGCCTTCGTGCGTCGCCGCG
CCGCCGTCCCCTTCTCCCTCTCCAGCCTCGGGGCTGTCCGCGGGGGGACGGCTGCCTTCG
GGGGGACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACCGGCGGCTCTAGAGCCTC
TGCTAACCATGTTTCATGCCTTCTTCTTTTCTACAGCGCCACCATGGGCAAGGAGAAGA
CCCACATCAACATCGTGGTCATCGGCCACGTGGACTCCGGAAAGTCCACCACCACGGGCC
ACCTCATCTACAAATGCGGAGGTATTGACAAAAGGACCATTGAGAAGTTCGAGAAGGAGG
CGGCTGAGATGGGGAAGGGATCCTTCAAGTATGCCTGGGTGCTGGACAAGCTGAAGGCGG
AGCGTGAGCGCGGCATCACCATCGACATCTCCCTCTGGAAGTTCGAGACCACCAAGTACT
ACATCACCATCATCGATGCCCGGCCACCGCGACTTCATCAAGAACATGATCACGGGTA
CATCCAGGCGGACTGCGCAGTGCTGATCGTGGCGGGCGGGCGTGGGCGAGTTCGAGGCGG
GCATCTCCAAGAATGGGCAGACGCGGGAGCATGCCCTGCTGGCCTACACGCTGGGTGTGA
AGCAGCTCATCGTGGGCGTGAACAAAATGGACTCCACAGAGCCGGCCTACAGCGAGAAGC
GCTACGACGAGATCGTCAAGGAAGTCAGCGCCTACATCAAGAAGATCGGCTACAACCCGG
CCACCGTGCCCTTTGTGCCATCTCCGGCTGGCACGGTGACAACATGCTGGAGCCCTCCC
CCAACATGCCGTGGTTCAAGGGCTGGAAGGTGGAGCGTAAGGAGGGCAACGCAAGCGGCG
TGTCCCTGCTGGAGGCCCTGGACACCATCCTGCCCGCCACGCGCCCGCACGGACAAGCCCC
TGCGCCTGCCGCTGCAGGACGTGTACAAGATTGGCGGCATTGGCACGGTGCCCGTGGGCC
GGGTGGAGACCGGCATCCTGCGGCCGGGCATGGTGGTGACCTTTGCGCCAGTGAACATCA
CCACTGAGGTGAAGTCAGTGGAGATGCACCACGAGGCTCTGAGCGAAGCTCTGCCCGGCG
ACAACGTCCGGCTTCAATGTGAAGAACGTGTCGGTGAAGGACATCCGGCGGGGCAACGTGT
GTGGGGACAGCAAGTCTGACCCCGCGCAGGAGGCTGCTCAGTTCACCTCCAGGTCATCA
TCCTGAACCACCGGGGCGAGATTAGCGCCGGCTACTCCCGGTCATCGACTGCCACACAG
CCCACATCGCCTGCAAGTTTTCGGAGCTGAAGGAGAAGATTGACCGGCGCTCTGGCAAGA
AGCTGGAGGACAACCCCAAGTCCCTGAAGTCTGGAGACGCGGCCATCGTGGAGATGGTGC
CGGAAAGCCCATGTGTGTGGAGAGCTTCTCCAGTACCCGCCTCTCGGCCGCTTCGCCG
TGCGCGACATGAGGCAGACGGTGGCCGTAGGCGTCATCAAGAACGTGGAGAAGAAGAGCG
GCGGCGCCGGCAAGGTCACCAAGTCGGCGCAGAAGGCGCAGAAGGCGGGCAAGTGATCAA
CCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTT
ACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCT

[0165]

TTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCC
 GTTGTGACGGCAACGTGGCGTGGTGTGCACTGTGTTTGTGCTGACGCAACCCCCACTGGTTGG
 GGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCTCCCTCCCTATTGCC
 ACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGC
 ACTGACAATTCCGTGGTGTGTCGGGGAAATCATCGTCTTTCCTTGGCTGCTCGCCTGT
 GTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCA
 GCGGACCTTCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTT
 CGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCAGCTGGAGCCTCGGTAG
 [0166] CCGTTCCCTCCTGCCCGCTGGGCCTCCCAACGGGCCCTCCTCCCTCCTTGCACCGGCCCT
 TCCTGGTCTTTGAATAAATTCATTGCCTGCCCGGGTGGCATCCCTGTGACCCCTCCCCAG
 TGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCCCTAATAAAAT
 TAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGG
 TGGTATGGAGCAAGGGGCCCAAGTTGGGAAGAAACCTGTAGGGCCTGCGTTACCCAGGCT
 GGAGTGCACTGGCACATTTCTGCTCACTGCAACCTCCTCCTCCCTGGGTTTctacgtagat
 aagtagcatggcgggttaatcattaactacaaggaaccctagtgatggagttggccact
 ccctctctgcgcgctcgctcgctcactgaggccgggacccaaaggtcgccccgacgcccg
 ggctttgcccgggcggcctcagtgagcgagcgagcgcgcg

[0167] 在一个实施例中,表达盒按5'至3'顺序包括HuBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pA珠蛋白-0c。

[0168] 在一个实施例中,表达盒按5'至3'顺序包括CMV启动子、TPL-eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)和pA珠蛋白-0c。

[0169] 在一个实施例中,表达盒按5'至3'顺序包括Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)、3'UTR(珠蛋白)和pAGH-Bt。

[0170] 在一个实施例中,表达盒按5'至3'顺序包括CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Bt。

[0171] 在一个实施例中,表达盒按5'至3'顺序包括EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列和pA珠蛋白-0c。

[0172] 在一个实施例中,表达盒按5'至3'顺序包括HuBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pAGH-Bt。

[0173] 在一个实施例中,表达盒按5'至3'顺序包括Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)、3'UTR(珠蛋白)和pAGH-Hs。

[0174] 在一个实施例中,表达盒按5'至3'顺序包括CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)和pAGH-Hs。

[0175] 在一个实施例中,表达盒按5'至3'顺序包括CMV启动子、TPL-eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)和pAGH-Hs。

[0176] 在一个实施例中,表达盒按5'至3'顺序包括HuBA启动子、编码eEF1A2或其功能变

体的多核苷酸序列和pAGH-Hs。

[0177] 在一个实施例中,表达盒按5'至3'顺序包括CMV启动子、TPL-eMLP增强子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR(珠蛋白)和pAGH-Bt。

[0178] 在一个实施例中,表达盒按5'至3'顺序包括EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(r)和pAGH-Bt。

[0179] 在一个实施例中,表达盒按5'至3'顺序包括Syn启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pA珠蛋白-0c。

[0180] 在一个实施例中,表达盒按5'至3'顺序包括CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pA珠蛋白-0c。

[0181] 在一个实施例中,表达盒按5'至3'顺序包括CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)、3'UTR(珠蛋白)和pAGH-Hs。

[0182] 在一个实施例中,表达盒按5'至3'顺序包括CBA启动子、编码eEF1A2或其功能变体的多核苷酸序列、3'UTR(珠蛋白)和pA珠蛋白-0c。

[0183] 在一个实施例中,表达盒按5'至3'顺序包括CaMKIIa启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17和pAGH-Bt。

[0184] 在一个实施例中,表达盒按5'至3'顺序包括EF1 α 启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR(珠蛋白)和pAGH-Hs。

[0185] 在一个实施例中,表达盒按5'至3'顺序包括CMV启动子、编码eEF1A2或其功能变体的多核苷酸序列、R2V17、3'UTR(珠蛋白)和pAGH-Hs。

[0186] 在一个实施例中,表达盒按5'至3'顺序包括CMV启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

[0187] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Bt。

[0188] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Hs。

[0189] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Hs。

[0190] 在一个实施例中,表达盒按5'至3'顺序包括CAG启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Hs。

[0191] 在一个实施例中,表达盒按5'至3'顺序包括CAG启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Hs。

[0192] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列、WPRE(x)和pAGH-Bt。

[0193] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

[0194] 在一个实施例中,表达盒按5'至3'顺序包括hSYN启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

[0195] 在一个实施例中,表达盒按5'至3'顺序包括CAG启动子、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

[0196] 在一个实施例中,表达盒按5'至3'顺序包括CAG启动子、Kozak、编码eEF1A2或其功能变体的多核苷酸序列和pAGH-Hs。

[0197] 腺相关病毒载体

[0198] 腺相关病毒(AAV)是复制缺陷型细小病毒,其单链DNA基因组长度为约4.7kb,包含两个~145个核苷酸的反向末端重复(ITR)。存在多种已知的AAV变体,当按抗原表位分类时,所述AAV变体有时也被称为血清型。AAV血清型的基因组的核苷酸序列是已知的。例如,AAV-1的完整基因组在GenBank登录号NC_002077中提供;AAV-2的完整基因组在GenBank登录号NC_001401和Srivastava等人,《病毒学杂志(J.Virol.)》,45:555-564(1983)中提供;AAV-3的完整基因组在GenBank登录号NC_1829中提供;AAV-4的完整基因组在GenBank登录号NC_001829中提供;AAV-5基因组在GenBank登录号AF085716中提供;AAV-6的完整基因组在GenBank登录号NC_001862中提供;AAV-7和AAV-8基因组的至少部分分别在GenBank登录号AX753246和AX753249中提供;AAV-9基因组在Gao等人,《病毒学杂志》,78:6381-6388(2004)中提供;AAV-10基因组在《分子疗法》,13(1):67-76(2006)中提供;并且AAV-11基因组在《病毒学(Virology)》,330(2):375-383(2004)中提供。美国专利9,434,928中提供了AAVrh.74基因组的序列,其通过引用并入本文。指导病毒DNA复制(rep)、衣壳化/包装和宿主细胞染色体整合的顺式作用序列包含在AAV ITR中。三个AAV启动子(其相对图谱位置命名为p5、p19和p40)驱动编码rep和cap基因的两个AAV内部开放阅读框的表达。与单个AAV内含子的差异剪接(在核苷酸2107和2227处)偶联的两个rep启动子(p5和p19)引起从rep基因产生四种rep蛋白(rep78、rep68、rep52和rep40)。Rep蛋白具有多种酶特性,所述酶特性最终负责复制病毒基因组。cap基因由p40启动子表达,并且其编码三种衣壳蛋白VP1、VP2和VP3。替代性剪接和非共有翻译起始位点负责产生三种相关的衣壳蛋白。单一共有多腺苷酸化位点位于AAV基因组的图谱位置95处。AAV的生命周期和遗传学在Muzyczka,《当前微生物学和免疫学的话题(Current Topics in Microbiology and Immunology)》,158:97-129(1992)中评论。

[0199] AAV具有独特的特征,这使其作为例如在基因疗法中将外源DNA递送到细胞的载体具有吸引力。培养中细胞的AAV感染是非细胞病变的,并且人类和其它动物的自然感染是沉默的和无症状的。而且,AAV感染许多哺乳动物细胞,从而允许体内靶向许多不同组织的可能性。此外,AAV转导缓慢分裂和非分裂细胞,并且可以作为转录活性核附加体(染色体外元件)基本上持续这些细胞的寿命。将AAV原病毒基因组作为质粒中的克隆DNA插入,这使得重组基因组的构建成为可能。此外,由于指导AAV复制和基因组衣壳化的信号包含在AAV基因组的ITR中,因此部分或全部内部约4.3kb的基因组(编码复制和结构衣壳蛋白,rep-cap)可以用外源DNA替代。为了产生AAV载体,rep蛋白和cap蛋白可以以反式提供。AAV的另一个显著特征是其是极其稳定且强健的病毒。它易于承受用于使腺病毒失活的条件(56°C到65°C,持续数小时),使AAV的冷保存不太重要。甚至可以将AAV冻干。最后,AAV感染的细胞不耐受重复感染。

[0200] rAAV基因组中的AAV DNA可以来自任何可源自重组病毒的AAV变体或血清型,包括但不限于AAV变体或血清型AAV-1、AAV-2、AAV-3、AAV-4、AAV-5、AAV-6、AAV-7、AAV-8、AAV-9、AAV-10、AAV-11、AAV-12、AAV-13和AAVrh10。在例如WO 01/83692中公开了假型rAAV的产生。还考虑了其它类型的rAAV变体,例如具有衣壳突变的rAAV。参见例如Marsic等人,《分子疗

法》,22(11):1900-1909(2014)。各种AAV血清型的基因组的核苷酸序列是本领域已知的。

[0201] 在一些情况下,rAAV包括自身互补型基因组。如本文所定义的,包括“自身互补型”或“双链”基因组的rAAV是指已经被工程化使得所述rAAV的编码区被配置成形成分子内双链DNA模板的rAAV,如在McCarty等人自身互补型重组腺相关病毒(scAAV)载体促进独立于DNA合成的有效转导(Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis.)《基因疗法(Gene Therapy.)》8(16):1248-54(2001)中所描述的。在一些情况下,本公开考虑了包括自身互补型基因组的rAAV的用途,因为在感染(这种转导)而不是等待rAAV基因组的第二链的细胞介导的合成时,scAAV的两个互补的半部分将缔合以形成一个准备好立即复制和转录的双链DNA(dsDNA)单元。应当理解,包括自身互补型基因组的rAAV仅可能保持所述量的约一半(≈ 2.4 kb),而不是在rAAV中发现的全部编码能力(4.7-6kb)。

[0202] 在其它情况下,rAAV载体包括单链基因组。如本文所定义的,“单一标准”基因组是指非自身互补型基因组。在大多数情况下,非重组AAV具有单链DNA基因组。存在rAAV应是scAAV以实现细胞的有效转导的一些指示。然而,本公开考虑了可能具有单链基因组而不是自身互补型基因组的rAAV载体,其中理解的是,rAAV载体的其它基因修饰可能有利于在靶细胞中获得最佳基因转录。在一些情况下,本公开涉及能够实现向小鼠眼睛中的前段进行有效基因转移的单链rAAV载体。参见Wang等人单链腺相关病毒实现了向小鼠眼睛中的前段进行有效基因转移(Single stranded adeno-associated virus achieves efficient gene transfer to anterior segment in the mouse eye.)《公共科学图书馆期刊(PLoS ONE)》12(8):e0182473(2017)。

[0203] 在一些情况下,rAAV载体为血清型AAV1、AAV2、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAV13、AAVrh10或AAVrh74。在例如WO 01/83692中公开了假型rAAV的产生。还考虑了其它类型的rAAV变体,例如具有衣壳突变的rAAV。参见例如Marsic等人,《分子疗法》,22(11):1900-1909(2014)。在一些情况下,rAAV载体是血清型AAV9。在一些实施例中,所述rAAV载体是血清型AAV9,并且包括单链基因组。在一些实施例中,所述rAAV载体是血清型AAV9,并且包括自身互补型基因组。在一些实施例中,rAAV载体包括AAV2的反向末端重复(ITR)序列。在一些实施例中,rAAV载体包括AAV2基因组,使得rAAV载体是AAV-2/9载体、AAV-2/6载体或AAV-2/8载体。

[0204] 美国专利第8,524,446号提供了大部分已知AAV的全长序列和衣壳基因序列,所述美国专利通过全文引用的方式并入本文中。

[0205] AAV载体可以包括野生型AAV序列,或其可以包括对野生型AAV序列的一个或多个修饰。在某些实施例中,AAV载体包括衣壳蛋白(例如,VP1、VP2和/或VP3)内的一个或多个氨基酸修饰,例如取代、缺失或插入。在特定实施例中,当将AAV载体提供给受试者时,所述修饰提供了降低的免疫原性。

[0206] 可以对rAAV的衣壳蛋白进行修饰,使得rAAV靶向所关注特定靶组织,如神经元或更具体地是多巴胺能神经元。参见例如,Albert等人AAV载体介导的向黑质多巴胺神经元的基因递送:对基因疗法和疾病模型的意义(AAV Vector-Mediated Gene Delivery to Substantia Nigra Dopamine Neurons:Implications for Gene Therapy and Disease

Models.)《基因(Genes)》2017年2月8日;还参见美国专利第6,180,613号和美国专利公开第US20120082650A1号,所述两个专利的公开内容通过引用并入本文。在一些实施例中,将rAAV直接注射到受试者的黑质中。

[0207] 在一些实施例中,rAAV病毒粒子是AAV2 rAAV病毒粒子。衣壳可以是AAV2衣壳或其功能变体。在一些实施例中,AAV2衣壳与参考AAV2衣壳(例如,以下序列)具有至少98%、99%或100%的同源性:

[0208] MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNLDKGEVNEAD
AAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPV
EHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGN
SSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFSRDPWQRLI
NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGY
LTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPPHSSYAHSQSLDRLMNPLIDQYLYLSRTNTPSGT
TTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPANPGPAMASHKDDEE
KFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQD
RDVYLQGPWIWAKIPHTDGHFHPSPMLGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWE
LQKENS KRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

[0209] (SEQ ID NO:59)

[0210] 在一些实施例中,rAAV病毒粒子是AAV9 rAAV病毒粒子。衣壳可以是AAV9衣壳或其功能变体。在一些实施例中,AAV9衣壳与参考AAV9衣壳(例如,以下序列)具有至少98%、99%或100%的同源性:

[0211] MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPFGNLDKGEVNAAD
AAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRLEPLGLVEEAAKTAPGKKRPV
EQSPQEPDSSAGIGKSGAQAQPAKKRLNFGQTGDATESVDPDQPIGEPAAAPSGVGS LTMASGGGAPVADNNEGADGVGS
SSGNWHCDSQWLGDRVITTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDNRFHCHFSRDPWQR
LINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQY
GYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPPHSSYAHSQSLDRLMNPLIDQYLYLSKTINGS
GQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGE
DRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYQVATNHQSAQAQAQTGWVQNGILPGMVWQ
DRDVYLQGPWIWAKIPHTDGNFHPSPMLGGFGMKHPPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWE
ELQKENS KRWNPEIQYTSNYKSNVFAVNTGCVYSEPRPIGTRYLTRNL

[0212] (SEQ ID NO:15)

[0213] 在一些实施例中,rAAV病毒粒子是AAV-PHP.B rAAV病毒粒子或其神经营养性变体,如但不限于国际专利公开号WO 2015/038958 A1和WO 2017/100671 A1中公开的那些。例如,AAV衣壳可以包括来自序列TLAVPFK(SEQ ID NO:61)或KFPVALT(SEQ ID NO:62)的至少4个连续氨基酸,例如,插入在编码AAV9的氨基酸588和589的序列之间。

[0214] 衣壳可以是AAV-PHP.B衣壳或其功能变体。在一些实施例中,AAV-PHP.B衣壳与参考AAV-PHP.B衣壳(例如,以下序列)具有至少98%、99%或100%的同源性:

[0215] MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGPFGNLDKGEVNAAD
AAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRLEPLGLVEEAAKTAPGKKRPV

EQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGS
SSGNWHCDSQWLGDRVITTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQR
LINNNWGFPRKRLNFKLFNIQVKEVTDNNGVKTIANLNTSTVQVFTDSDYQLPYVLGSAHEGCLPPFPADVFMIPQY
GYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTINGS
GQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNEFAWPGASSWALNGRNSLMNPGPAMASHKEGE
DRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGVATNHQSAQTLAVPFKAQAQTGWVQNQGI
LPGMVWQDRDVYLQGPWIWAKIPHTDGNFHPSPMLGGFGMKHPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQ
VSVEIEWELQKENS KRWNPEIQYTSNYYKSNVFAVNTG VYSEPRPIGTRYLTRNL

[0216] (SEQ ID NO:60)

[0217] 本公开的rAAV病毒粒子中使用的其它AAV衣壳包含专利公开号W0 2009/012176 A2和W0 2015/168666 A2中公开的那些。

[0218] 药物组合物和试剂盒

[0219] 一方面,本公开提供了药物组合物,其包括本公开的rAAV病毒粒子和一种或多种药学上可接受的载体、稀释剂或赋形剂。

[0220] 出于施用的目的,例如通过注射,可以采用如无菌水性溶液等各种溶液。如果需要的话,这种水性溶液可以被缓冲,并且液体稀释剂首先用盐水或葡萄糖进行等渗。作为游离酸的rAAV溶液(DNA含有酸性磷酸酯基团)或药理学上可接受的盐可以在以0.001%或0.01%适当地与表面活性剂(如Pluronic™ F-68)混合的水中制备。还可以在甘油、液体聚乙二醇、及其混合物中以及在油中制备rAAV的分散体。在通常的储存和使用条件下,这些制剂含有防腐剂以预防微生物的生长。在这方面,采用的无菌含水介质都可通过本领域技术人员熟知的标准技术容易地获得。

[0221] 适于可注射使用的药物形式包含但不限于无菌水溶液或分散体以及用于临时制备无菌可注射溶液或分散体的无菌粉末。在所有情况下,所述形式是无菌的并且应具有达到易于注射的程度的流动性。其必须在制造和储存条件下稳定,并且必须防止如细菌和真菌等微生物的污染作用。载体可以是溶剂或分散介质,其含有例如水;乙醇;多元醇(例如,甘油、丙二醇、液体聚乙二醇等);其合适的混合物和植物油。可以例如通过使用如卵磷脂等包衣、通过在分散液的情况下维持所需的粒度以及通过使用表面活性剂来维持适当的流动性。通过各种抗细菌剂和抗真菌剂,例如,对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸、硫柳汞等可以带来防止微生物的作用。在许多情况下,将优选的是包含等渗剂,例如糖或氯化钠。可以通过使用延迟吸收剂(例如,单硬脂酸铝和明胶)延长可注射组合物的吸收。

[0222] 通过将rAAV以所需的量根据与以上枚举的各种其它成分一起并入适当的溶剂中、然后进行过滤灭菌来制备无菌可注射溶液。通常,通过将灭菌的活性成分并入到无菌媒剂中来制备分散体,所述无菌媒剂含有基础分散介质以及来自以上枚举的那些的所需其它成分。在用于制备无菌可注射溶液的无菌粉末的情况下,优选的制备方法是真空干燥和冷冻干燥技术,所述方法从其先前经过无菌过滤的溶液中产生活性成分加上任何另外所期望的粉末。

[0223] 另一方面,本公开包括试剂盒,其包括本公开的rAAV病毒粒子和使用说明。

[0224] 使用方法

[0225] 一方面,本公开提供了一种增加细胞中的eEF1A2活性的方法,所述方法包括使所

述细胞与本公开的rAAV接触。另一方面,本公开提供了一种增加受试者的eEF1A2活性的方法,所述方法包括向施用本公开的rAAV。在一些实施例中,细胞和/或受试者缺乏eEF1A2表达水平和/或活性和/或包括eEF1A2的功能缺失突变。细胞可以是神经元,例如多巴胺能神经元。

[0226] 在一些实施例中,所述方法促进神经元在细胞培养物和/或在体内的存活。

[0227] 治疗方法

[0228] 另一方面,本公开提供了一种治疗有需要的受试者的疾病或病症的方法,所述方法包括向所述受试者施用有效量的本公开的rAAV病毒粒子。在一些实施例中,所述疾病或病症是神经系统疾病或病症。在一些实施例中,受试者患有eEF1A2表达或功能的遗传破坏。在一些实施例中,所述疾病或病症是eEF1A2缺乏症和/或eEF1A2相关神经系统疾病(OMIM# 617309、616393、616409)表型谱,如智力残疾、智力迟钝、癫痫性脑病和自闭症谱系障碍。

[0229] AAV介导的eEF1A2蛋白向CNS的递送可延长寿命,预防神经元变性,预防或减轻神经行为缺陷、退行性癫痫运动障碍性脑病、癫痫和肌张力障碍。

[0230] 本发明还考虑了组合疗法。具体考虑了本发明的方法与标准医学治疗(例如,皮质类固醇或局部减压药物)的组合以及与新颖疗法的组合。在一些情况下,可以用类固醇治疗受试者,以防止或减少对施用本文所描述的rAAV的免疫应答。

[0231] 治疗有效量的rAAV载体,例如用于脑室内(ICV)或小脑延髓池内(ICM)注射的载体,是按脑重量计范围为约 $1e12$ vg/kg至约 $5e12$ vg/kg、或约 $1e13$ vg/kg至约 $5e13$ vg/kg、或约 $1e14$ vg/kg至约 $5e14$ vg/kg、或约 $1e15$ vg/kg至约 $5e15$ vg/kg的rAAV剂量。或按体重计范围为 $1213-1e14$ vg/kg的静脉递送剂量。本发明还包括包含这些范围的rAAV载体的组合物。

[0232] 例如,在特定实施例中,rAAV载体的治疗有效量是约 $1e10$ vg、约 $2e10$ vg、约 $3e10$ vg、约 $4e10$ vg、约 $5e10$ vg、约 $6e10$ vg、约 $7e10$ vg、约 $8e10$ vg、约 $9e10$ vg、约 $1e12$ vg、约 $2e12$ vg、约 $3e12$ vg、约 $4e12$ vg或约 $5e12$ vg的剂量。本发明还包括包含这些剂量的rAAV载体的组合物。

[0233] 在一些实施例中,例如在进行ICV注射的情况下,rAAV载体的治疗有效量是在 $1e10$ vg/半球至 $1e13$ vg/半球、或约 $1e10$ vg/半球、约 $1e11$ vg/半球、约 $1e12$ vg/半球、或约 $1e13$ vg/半球范围内的剂量。在一些实施例中,例如在进行ICM注射的情况下,rAAV载体的治疗有效量是在 $1e10$ vg总量至 $1e14$ vg总量、或约 $1e10$ vg总量、约 $1e11$ vg总量、约 $1e12$ vg总量、约 $1e13$ vg总量、或约 $1e14$ vg总量范围内的剂量。

[0234] 在一些实施例中,治疗组合物包括每体积所注射的治疗组合物多于约 $1e9$ 、 $1e10$ 或 $1e11$ 个rAAV载体基因组。在实施例情况下,治疗组合物包括每mL多于大约 $1e10$ 、 $1e11$ 、 $1e12$ 或 $1e13$ 个rAAV载体基因组。在某些实施例中,治疗组合物包括每mL少于约 $1e14$ 、 $1e13$ 或 $1e12$ 个rAAV载体基因组。

[0235] 在患者中功能改善的证据、临床益处或疗效可以通过使用UK-WHO儿科头围、身高和体重百分位数图表分析癫痫发作(肌阵挛性和全身强直阵挛发作)频率降低、脑生长和身体生长的替代标志物来进行评估。使用标准疾病评定量表(如儿童癫痫发作量表和药物日志)测量认知、运动、言语和语言功能。认知和发育评估,包含《peabody运动发育量表(Peabody Developmental Motor Scales)》第2版(PDMS-2)和《贝利婴儿发育量表(Bayley Scales of Infant Development)》第3版,适用于儿童的残疾程度。粗大运动功能测量

(GFMF-88)、残疾量表儿科评估(PEDI)。这些或类似的量表,以及患者报告的生活质量结果,如3分量表上的护理人员癫痫发作持续时间变化的全局印象(CGICSD)(平均持续时间减少、无变化或增加)、儿科生活质量量表(PedsQL™)和Vineland适应性行为量表-第2版,可能证明疾病成分的改善。基线和后治疗脑磁共振成像可能显示髓鞘形成和脑容量的改善。在患有常染色体显性遗传性EEF1A2相关神经发育障碍的患者中观察到心脏缺陷,包含心肌病、主动脉缺损和室间隔缺损(Kaneko等人,2021;Carvill等人,2020;McLachlan等人,2019)。在具有全球发育迟缓、癫痫、生长迟缓、扩张型心肌病和过早死亡的单亲中鉴定出EEF1A2的纯合变体(Cao等人,2017)。心脏状态的测量可以通过基线心电图和超声心动图监测。

[0236] 临床益处可以被观察为寿命的增加、满足正常的神经发育里程碑、癫痫发作活动(包含肌阵挛性、阵挛性、全身强直阵挛和/或癫痫性痉挛)的频率或幅度降低、发育中的肌张力减退或运动障碍(如舞蹈手足徐动症、肌张力障碍和/或共济失调)的改善或缺乏。神经保护和/或神经修复作用的证据可能在磁共振成像(MRI)上通过表征发育中的髓鞘化程度、胼胝体厚度和皮质和/或小脑萎缩程度而明显。脑电图(EEG)活动的有益变化将通过多灶性放电和/或全身尖峰活动的减少来证明。

[0237] 在一些实施例中,例如在进行静脉内施用的情况下,rAAV载体的治疗有效量是按受试者的总体重计在约 $1e12$ vg/kg至 $1e14$ vg/kg范围内的剂量。例如,在特定实施例中,rAAV载体的治疗有效量是约 $1e12$ vg/kg、约 $2e12$ vg/kg、约 $3e12$ vg/kg、约 $4e12$ vg/kg、约 $5e12$ vg/kg、约 $6e12$ vg/kg、约 $7e12$ vg/kg、约 $8e12$ vg/kg、约 $9e12$ vg/kg、约 $1e13$ vg/kg、约 $2e13$ vg/kg、约 $3e13$ vg/kg、约 $4e13$ vg/kg、约 $5e13$ vg/kg、约 $6e13$ vg/kg、约 $7e13$ vg/kg、约 $8e13$ vg/kg、约 $9e13$ vg/kg或约 $1e14$ vg的剂量。心脏获益的证据可能包含超声心动图显示心脏功能稳定。

[0238] 组合物的施用

[0239] 可以通过本领域标准的途径施用有效剂量的组合物,包括但不限于全身、局部、直接注射、静脉内、脑、脑脊髓、鞘内、脑池内、壳核内、海马内、纹状体(壳核和/或尾状核)内、皮质内或脑室内施用。在一些情况下,施用包括静脉内、脑、脑脊髓、鞘内、脑池内、壳核内、海马内、纹状体(壳核和/或尾状核)内或脑室内注射。可以通过使用或不使用特伦德伦伯(Trendelenburg)倾斜进行鞘内注射来施用。

[0240] 在一些实施例中,本公开提供了有效剂量的本发明的rAAV和组合物的局部施用和全身施用。例如,全身施用可以是对循环系统的施用,从而影响整个身体。全身施用包含通过注射、输注或植入进行亲本施用。

[0241] 具体地,可以通过使用将rAAV重组载体转运到动物的靶组织中的任何物理方法来完成本发明的rAAV的施用。施用包括但不限于注射到中枢神经系统(CNS)或脑脊髓液(CSF)中和/或直接注射到脑中。

[0242] 在一些实施例中,本公开的方法包括脑室内、小脑延髓池内、鞘内或脑实质内递送。可以使用专用套管、导管、注射器/针头,使用输注泵进行输注。任选地,可以通过MRI引导的成像来实现对注射位点的靶向。施用可以包括将有效量的rAAV病毒粒子或包括rAAV病毒粒子的药物组合物递送至CNS。这些可以例如通过单侧脑室内注射、双侧脑室内注射、采用特伦德伦伯倾斜程序的小脑延髓池内输注、或不采用特伦德伦伯倾斜程序的小脑延髓池内输注、采用特伦德伦伯倾斜程序的鞘内输注或不采用特伦德伦伯倾斜程序的鞘内输注来

实现。本公开的组合物可以进一步静脉内施用。

[0243] 直接递送至CNS可以涉及靶向脑室内空间(无论是单侧还是双侧)、特定神经元区或含有神经元靶标的更一般脑区。通过使用多种成像技术(MRI、CT、CT与MRI合并组合)以及使用任何数量的软件规划程序(例如,隐形系统(Stealth System)、Clearpoint神经导航系统(Clearpoint Neuronavigation System)、博医来(Brainlab)、Neuroinspire等),可以实现个体患者的脑室内空间、脑区和/或神经元靶标选择以及随后的AAV手术中递送。脑室内空间或脑区靶向和递送可以涉及使用标准立体定向框架(Leksell, CRW)或使用无框方法,伴随或不伴随手术MRI。AAV的实际递送可以通过针头或套管注射,内腔内衬或不内衬材料,以防止AAV载体的吸附(例如,Smartflow套管、MRI介入套管)。递送装置以预编程的输注速率和体积与注射器和自动输注或微量输注泵接口连接。注射器/针的组合或仅针可以与立体定向框架直接接口连接。输注可以包含恒定流动速率或变化的加强对流传递速率。

[0244] 实例

[0245] 实例1:启动子选择

[0246] 在野生型新生小鼠中进行生物分布研究以选择启动子以恢复eEF1A2在神经元中的表达。与所有其它候选启动子相比,人突触素(hSYN)启动子对神经系统表现出优越的选择性和强烈的神经元表达,如表1所示。令人惊讶的是,与eSYN和其它测试的启动子相比,hSYN启动子显示出更大的神经元选择性。

[0247] 表6

启动子	eEF1A2转基因的脑细胞嗜性
CMV	神经元和神经胶质
EF1a	强神经元
增强突触素(eSYN)	神经元和神经胶质
人突触素(hSYN)	强神经元

[0249] 实例2:EEF1A2基因敲除小鼠模型的AAV9基因疗法挽救

[0250] 我们已经为受EEF1A2基因突变影响的受试者(例如,儿童)开发了一种新的治疗方法。真核翻译伸长因子1 α 2(eEF1A2)对于将氨酰基转移RNA递送到核糖体以进行蛋白质合成至关重要。EEF1A2基因突变与严重的智力残疾、自闭症和癫痫有关。目前没有有效的治疗方法。EEF1A2敲除小鼠模型(废弃小鼠)已得到很好的表征。废弃(wst/wst)小鼠在断奶后表现出步态障碍和震颤,然后在23日龄时出现麻痹和运动神经元变性。使用这种小鼠模型,发明者测试了蛋白质的功能是否可以通过基因疗法恢复。我们使用泛神经元启动子人突触素设计了一种腺相关病毒9(AAV9),来驱动人EEF1A2 cDNA(hSyn-eEF1A2)的表达。包含了eGFP标记基因以跟踪构建体在体内的表达。免疫荧光(图7)显示了新生儿IC或IV注射AAV9-hSyn-eEF1A2-T2A-eGFP后的神经元靶向。免疫组织化学染色(图8)证实了来自单次注射rAAV的两种施用途径后CNS中广泛表达转基因(对于eEF1A2-2A-eGFP或单独的eGFP标志物)。

[0251] 基因疗法载体被证明可有效治疗废弃(wst/wst)小鼠。Eef1a2^{-/-}敲除小鼠(wst/wst)在注射IC时大多存活(3/4),并在注射IC和IV时全部存活(图9A)。未经治疗的小鼠死于P23。与WT小鼠相比,IC或IC/IV小鼠同样没有显示重量减轻,而未经治疗的对照小鼠表现出重量减轻,从而导致P23死亡(图9B)。转棒仪和倒置网格分析表明,在治疗后的性能没有下降(图9C和图9D)。双向ANOVA和邓尼特多重比较测试的结果都很显著。

[0252] 在野生型、IC和组合治疗(图9E和图9F)中,在整个脑中观察到eEF1A2表达。eEF1A2表达存在于野生型、IC和组合治疗组的脊髓组织中。然而,在未经治疗的废弃组和IV治疗组中没有表达(F)。

[0253] 实例3:EEF1A2^{D252H}或EEF1A2^{G70S}或EEF1A2^{E122K}小鼠模型的AAV9基因疗法挽救

[0254] 对图2-5和图6所示的载体设计的功效以及各种密码子优化进行比较,以鉴定出具有优越功效的载体。实验在小鼠模型中进行,这些模型概括了在人类(D252H,G70S和或E122K)和/或具有严重神经退行性表型(De1.22.ex3)的小鼠模型中发现的三种突变。在新生小鼠和成人发育的后期阶段进行实验,以确认编码eEF1A2的AAV载体可以挽救存活期、重量减轻和行为表型。对神经行为测试的有益影响的评估包含对旋转圆柱体(转棒仪)的性能、附着在悬挂电线表面的能力(电线悬挂测试)、脚部故障测试、倒置网格行为、握力和在开放场地中的行为,包含观察正常的探索活动或异常行为(例如“抽搐行为”)。可以获得一系列测试的集体神经评分,以评估注射AAV9-eEF1A2的小鼠相对于未经治疗的对照的神经行为功能,其包含后肢扣紧、步态、后凸和沿壁架行走的能力的分析。De1.22.ex3(或其它脑电图(EEG)异常的小鼠品系)小鼠癫痫发作频率和脑电图的改变延长了寿命,这揭示了减轻CNS中异常电活动的有益作用。生化和组织学分析证实,载体设计对CNS内eEF1A2表达水平和分布具有卓越的功效。组织分析包含通过蛋白质印迹和ELISA检测新鲜组织中的mRNA、DNA、载体拷贝数和转基因蛋白表达,以及通过免疫标记检测CNS固定切片中的mRNA、DNA、载体拷贝数和转基因蛋白表达。

[0255] 实例4:在EEF1A2不足的WST/WST小鼠模型中,AAV9介导的人EEF1A2蛋白向中枢神经系统的递送的影响

[0256] eEF1A2相关病症的一个主要模型是废弃小鼠模型,其中EEF1A2基因的第一外显子和所有启动子元件的自发缺失导致eEF1A2无效(wst/wst)。在未经治疗的动物中,31%的wst/wst小鼠在P20-22之间死亡,并且存活的wst/wst小鼠表现出重量衰减、震颤随后重量减轻,其余的在第24天全部死亡。未经治疗的动物也可能表现出握力受损和转棒性能受损,动物在最长的发育性震颤、进行性瘫痪和重量减轻中幸存下来。我们的wst/wst集群中的动物看起来更严重,衰退更剧烈,死亡更早。通过P23处的倒置网格掉棒潜伏期,我们已经观察到握力受损。由于wst/wst动物只能存活到P24,我们没有观察到转棒性能的恶化,以显著区分wst/wst与野生型动物。

[0257] 材料和方法

[0258] 动物福利

[0259] 所有动物实验均按照英国内政部和1986年的动物(科学程序)法案进行,并在伦敦大学学院伦理审查委员会的指导方针内进行。本研究中使用的wst/wst eEF1A2无效小鼠模型之前已经描述过(Chambers D等人《美国科学院院报(PNAS)》95:4463-8(1998)。)。对杂合小鼠进行时间交配以产生混合基因型窝。使用引物在P0对幼崽进行基因分型(引物EEF1A2 Mut F 5'ACCAGTGGTTTCACCTGCTC 3'、EEF1A2 Common R 5'CACTGTGGGGCTCTGGTTT 3'、EEF1A2 WT F 5'CAGAGCTTCACTCAGTCTG 3')。

[0260] AAV9-eEF1A2载体或对照品的施用是通过在P0处双侧脑室内注射到新生纯合wst/wst或WT同窝幼崽进行的,并随访动物至人道终点(重量减轻 \geq 15%)或定时处死点P60。如前所述,脑室内注射被定向到P0-1小鼠的侧脑室(Newbery HJ等人《神经病理学与实验神经

病学杂志(J Neuropathol Exp Neurol)》64:295-303(2005))。将33号针(汉密尔顿)垂直插入注射部位至3mm深度,并在5秒内将5 μ l载体施用到侧心室中。幼崽被迅速送回母鼠。基因治疗治疗的wst/wst小鼠的组大小为6,7个窝中14-16个对照同窝。

[0261] 行为研究

[0262] 定期对小鼠进行称重并评估其总体幸福感的变化和达到人道终点。在P23处进行行为测试转棒仪和倒置网格测试。所有行为测试均由对动物治疗组不知情的研究人员进行。将小鼠放置在转棒仪(Harvard Apparatus®)上,从4-40r.p.m.连续加速,持续最多5分钟。记录小鼠从棒上掉下来的时间,在每天的测试中对每只动物进行3次试验(掉棒潜伏期)。倒置网格测试涉及将小鼠放在不锈钢网格(41x25cm)上,该网格放置在30cm高架塑料透明盒上。记录从倒置网格的掉棒潜伏期,最长为5分钟。在测试的每一天,对每只小鼠重复3次倒置网格测试。

[0263] 小鼠组织的组织学和免疫组织化学分析

[0264] 使用PBS通过末端经心灌注剔除小鼠。将收集的组织(脑和内脏器官)减半,以允许不同的处理技术。将用于免疫组织化学的脑在4% PFA中后固定48小时,并转移到30%蔗糖溶液中,在4°C下进行冷冻保护直至切片。将脑封固在冷冻切片机(ThermoFisher®HM430)上,厚度为40 μ m,处于任一冠状面。进行基于自由浮动的免疫组化分析,以240 μ m的间隔选择脑切片进行全脑免疫组织化学。简而言之,将自由漂浮的切片在室温下在15%正常山羊血清(Vector Laboratories®)-tris缓冲盐水中用0.1%triton-X(TBS-T)(Sigma®)封闭1小时,并在初级抗体抗兔eEF1A2(Proteintech®)中在10%正常山羊血清-TBS-T中在4°C下温育过夜。第二天将切片与各自的物种特异性二级抗体(Vector Laboratories®)一起在室温下温育1小时,在TBS中洗涤,然后与Vectatain亲和素-生物素溶液(Vector Laboratories®)一起温育。将反应用3,3'-二氨基联苯胺(DAB)(Sigma®)可视化。使用冰冷的1x TBS停止DAB反应,并在封固到双涂层糊化载玻片上之前清洗切片。将封固的切片风干并在100%乙醇中脱水10分钟,并用脱水溶液(Histoclear™,National Diagnostics®)脱水30分钟,然后用封固剂(DPX,VWR International®)覆盖以进行盖玻封片。

[0265] 对于免疫荧光,将脑切片在15%山羊血清中封闭30分钟,然后在4°C下与初级抗体(兔eEF1A2 1:1000 Proteintech®和小鼠NeuN 1:1000 Milipore)一起在10%正态山羊血清TBS-T 0.3%中温育过夜。将切片在1xTBS中洗涤,并与分别用Alexa 488和Alexa 594(均来自Invitrogen®,稀释在10%正常山羊血清中)标记的物种特异性二级抗体一起在室温下温育2小时。将NV4ei用DAPI(Sigma Aldrich®)染色2分钟。将脑切片封固在双涂层载玻片上,并使用Fluoromount G™(Thermofisher Scientific®)进行盖玻封片。

[0266] 光学显微镜和荧光成像均使用徕卡DM 4000与徕卡DFC420相机系统联动进行。使用徕卡TCS SP5 AOBs共聚焦显微镜捕获共聚焦图像。使用Image J软件(美国国立卫生研究院)分析图像。

[0267] 免疫印迹

[0268] 在补充有蛋白酶抑制剂(Roche®)的冰冷的0.32M蔗糖中,使用Qiagen®组织裂

解器从小鼠脑组织中提取蛋白质,并在4度下离心15分钟。使用Pierce BCA蛋白质测定试剂盒(Thermo Scientific®)测量蛋白质浓度:用Laemmli缓冲液(Bio-Rad Laboratories®)和二硫苏糖醇(DTT)使10 μ g蛋白质变性。使用Mini-PROTEAN TGX™免染色凝胶(Bio-Rad Laboratories®)分离蛋白质,并将其转移到Trans-Blot Turbo转印膜(Bio-Rad Laboratories®)上。在室温下在Biorad®封闭缓冲液中封闭1小时后,将膜与初级抗体兔eEF1A2(Proteintech®,1:1000)和小鼠GAPDH(Ab Cam®,1:10,000)一起在4°C下温育过夜。然后将膜与二级StarBright™Blue 520山羊抗兔IgG(1:3000)和StarBright™Blue 700山羊抗小鼠IgG(1:3000)一起孵育。将免疫反应蛋白用Chemidoc MP(Bio-Rad Laboratories®)可视化。上样n=4-5个生物重复

[0269] qRT-PCR mRNA转录本表达分析

[0270] 按照制造商的说明,用RNeasy™迷你试剂盒(Qiagen®)从脑匀浆中提取RNA(前脑,皮层每组n=4-5个生物重复),并在Omega Fluostar™上定量。使用DNase I纯化试剂盒(NEB®)从总RNA(1 μ g)中去除污染DNA,然后使用大容量cDNA逆转录试剂盒(Applied Bioscience®)进行逆转录。然后使用10ng的DNA或合成的cDNA进行多重hEEF1A2和mGAPDH RT-qPCR(eEF1A2__Fwd1:ATCGTGGGCGTGAACAAA,eEF1A2_Rev1:GGTTGTAGCCGATCTTCTTGAT,eEF1A2_探针:ATCGTCAAGGAAGTCAGCGCCTAC和小鼠GAPDH,针对:ACGGCAAATTCAACGGCAC,Rev:TAGTGGGGTCTCGCTCCTGG,探针:TTGTCATCAACGGGAAGCCCATCA Luna Taqman™预混液(NEB®)在Quantstudio™实时荧光定量PCR系统(Applied Biosystems®)中。将GAPDH作为内源性对照,计算相对倍数变化。

[0271] 统计学分析

[0272] 使用GraphPad Prism™版本8对每个实验进行量身定制的统计分析。体内实验设计和样品量采用NC3R指导和功效计算进行设计。对于大多数动物实验分析,单向或双向ANOVA是用Bonferroni或Tukey多重比较进行的。

[0273] 结果

[0274] 图10A-10K示出了以 2×10^{11} vg/动物施用的AAV9载体的比较,包括图2(“V1”;SEQ ID NO:55)、图3(“V2”;SEQ ID NO:56)、图4(“V3”;SEQ ID NO:57)和图6(“V4”;SEQ ID NO:58)中所示的载体基因组。

[0275] 图10A FBS治疗的野生型、wst/wst、脑室内治疗的wst/wst动物用V1、V2、V3和V4基因疗法治疗的Kaplan-Meier存活图(2×10^{11} vg/幼崽,所有基因疗法治疗组,n=6,野生型FBS,n=14,wst/wst FBS=17)。所有基因疗法治疗的动物均显示出显著的生命延长(平均延长V1=4.4天,p=0.001,V2=3.8天,p=0.0014,V3 10.7天,p<0.0001,V4=4.4天,p=0.001,对数秩曼特尔-考克斯检验(Logrank Mantel-Cox test))。图10B小鼠的重量(数据均值 \pm S.E.M.)。每天称重动物直到出生后35岁,此后每周称重,直到定时处死点P60或人道终点15%的重量减轻。图10C、图10D、图10E和图10F通过倒置网格测试和转棒进行的肌肉力量评估(每组n=4-7,每只动物在15和23日龄时一式三份地进行测试)。这些测试在转棒上观察到wst/wst和野生型FBS对照组之间没有显著差异。与WT相比,在P23处观察到wst/wst FBS动物的握力显著降低,并且基因疗法载体对这些行为结果没有影响(数据均值 \pm S.E.M.、双向ANOVA)。图10G以野生型FBS为生理参考通过自由浮动免疫组织化学对整个脑

的eEF1A2的代表性免疫染色(每组n=4-5,比例尺250 μ m)。所有注射V1-4的wst/wst动物在整个脑中表达eEF1A2,其中V3观察到的最暗的染色和最高的表达。图10H在共表达eEF1A2的神经元(NeuN标志物)皮质脑区中的代表性免疫组织荧光(每组n=4.5,200 μ m)。与野生型相比,所有指标均表现出更高的神经元表达谱。通过荧光强度观察到的最高表达是V3显示贯穿所有皮质层的表达。所有图像均以恒定的成像设置捕获。在V2和V4中观察到皮质层4和5中更离散的表达。图10I脑中eEF1A2的代表性免疫印迹,其中定量示出了eEF1A2在整个脑中的表达,与V4载体相比,所有基因疗法载体在中脑、小脑和后脑区域中的表达更高(数据均值 \pm S.E.M.、双向ANOVA)。图10J前脑中的人eEF1A2转录本表达的qPCR,其示出V1载体的最高mRNA表达(数据均值 \pm S.E.M.、双向ANOVA)。图10K前脑中的人eEF1A2皮层表达的qPCR,其示出V1载体的最高mRNA表达(数据均值 \pm S.E.M.、双向ANOVA)。

[0276] 结论

[0277] 这些实验表明,所有四种载体都能够在具有EEF1A2基因纯合无效突变(在小鼠中称为wst)的小鼠模型中恢复eEF1A2的表达。令人惊讶的是,载体V1、V2和V3能够有效且显著地将wst/wst小鼠的存活期延长超过P23。(图10A)。此外,该实验出人意料地表明,eEF1A2转基因在脑中的表达,特别是前脑(图10J)和皮层(图10K),在V3中大于V1、V2或V4。不受理论约束,似乎与包含5'UTR(SEQ ID NO:34)和/或3'UTR(SEQ ID NO:48)在内的预期相反,在治疗上增加基因表达,从而延长存活期。

[0278] 实例5:在患有EEF1A2疾病的D252H小鼠模型中,AAV9介导的人EEF1A2蛋白向中枢神经系统的递送的影响

[0279] EEF1A2中的杂合新生突变,编码组织特异性翻译伸长因子eEF1A2,已被证明可引起神经发育障碍,通常包含严重的癫痫和智力残疾。大约有50种不同的错义突变被鉴定出来,但没有明显的功能丧失突变,尽管已知大的杂合缺失与生命相容。在相同的遗传背景下,携带引起错义D252H突变的疾病的敲入eEF1A2小鼠模型比无效纯合子受到更严重的影响,其显示为重量减轻、神经评分增加和P23死亡。错义突变杂合的小鼠没有表现出行为异常,但在体重和运动功能方面确实存在性别特异性缺陷,握力短暂受损。这种D252H新型小鼠的表型与del22ex3无效小鼠模型一起支持了D252H突变导致功能增益(Davies,Faith CJ等人《人类分子遗传学(Human Molecular Genetics)》(2020))。该实例描述了杂合和纯合D252H小鼠的基因疗法研究。

[0280] 材料和方法

[0281] 在发育过程中监测经治疗小鼠和未经治疗小鼠的存活期和重量。在P18-24日龄之间也进行了行为测试。

[0282] 动物福利

[0283] 所有动物实验均按照英国内政部和1986年的动物(科学程序)法案进行,并在伦敦大学学院伦理审查委员会的指导方针内进行。将杂合D252HeEF1A2敲入小鼠进行时间交配以产生混合基因型窝。在P0处将幼崽使用引物5'-3'AGGCTACCCCTTAGGCAGGT、TGAACAAATGGTAGGTGGGAGG进行基因分型。PCR扩增后,将样品通过Hin1III(赛默飞世尔公司(Thermo Fisher))进行限制性酶切。

[0284] 测试品和控制品的施用

[0285] 通过向新生纯合敲入小鼠(D252H-/-)单侧脑室内注射 1.8×10^{11} vg/幼崽的剂量来

实现5 μ L的V3载体或调配物缓冲液(FB)制品的施用。使用以下文献中划定的注射部位坐标,通过33号汉密尔顿针(Fisher Scientific®,英国拉夫堡)对野生型和杂合D252H施用测试或FB对照制品:Kim等人Kim,J.Y.等人《可视实验杂志(J Vis Exp)》91:51863(2014)。Lambdoid缝合线在新生幼崽中是可鉴定的,并且预期的注射部位是从lambdoid缝合线到眼睛的2/5,位于矢状缝合线约0.8mm-1mm的外侧,介于 λ 和前囟之间。注射后,幼崽被送回母鼠。

[0286] 体重

[0287] 从P1-P32收集个体动物体重,此后每周一次。重量减轻 $\geq 15\%$ 将达到安乐死的人道终点标准。

[0288] 行为测试

[0289] 所有行为测试测定均由对动物治疗组不知情的研究人员进行。转棒测试:从P18-24进行转棒仪训练/测试。将小鼠放置在转棒仪(哈佛仪器公司(Harvard Apparatus))上,从4-40r.p.m.连续加速,持续最多2分钟。记录小鼠从棒上掉下来的时间,在每天的测试中对每只动物进行3次试验(掉棒潜伏期)。倒置网格测试:倒置网格测试涉及将小鼠放在不锈钢网格(41x25cm)上,该网格放置在30cm高架塑料透明盒上。记录从倒置网格的掉棒潜伏期,最长为2分钟。在测试的每一天,对每只小鼠重复3次倒置网格测试。

[0290] 结果

[0291] 图11A-11C示出了未经治疗的D252H^{-/-}小鼠的表型。图11A:FBS治疗的野生型、敲除(D252H^{-/-}),脑室内V3基因疗法治疗的D252H^{-/-}的Kaplan-Meier存活图(2×10^{11} vg/幼崽,V3治疗的=2,V4治疗的=1,野生型FBS,n=5,D252H^{-/-}-FBS=3。图11B小鼠体重(数据均值 \pm S.E.M.)。图11C通过转棒仪进行的运动评估(数据均值 \pm S.E.M.、双向ANOVA和邓尼特多重比较)。

[0292] 结论

[0293] 这些实验表明,与未经治疗的D252H^{-/-}对照相比,AAV9介导的eEF1A2在具有V3的D252H^{-/-}小鼠中的表达可以增加存活期。此外,尽管D252H^{-/-}纯合小鼠的转棒性能始终较差,但在D252H^{-/-}-V3治疗的小鼠中,随着时间的推移一直观察到改善的趋势。在先验定义的长期安全性/耐受性研究中,对D252H^{+/+}杂合小鼠脑室内施用V3后,eEF1A2的过表达似乎对存活期或功能结果测量没有有害影响。在纯合和杂合D252H小鼠中进行的分析将进一步表征AAV9介导的eEF1A2过表达在eEF1A2缺陷模型中的长期影响。

[0294] 实例6:在患有EEF1A2疾病的DEL22EX3小鼠模型中,AAV9介导的人EEF1A2蛋白向中枢神经系统的递送的影响

[0295] 生成了CRISPR/Cas9生成的Del.22.ex.3eEF1A2小鼠模型以敲除eEF1A2表达(Davies,Faith CJ等人《人类分子遗传学》2020)。由CRISPR/Cas9诱变产生的Eef1a2外显子3内的22个碱基对缺失导致无效突变。这些Del122ex3小鼠呈现严重的表型,使得小鼠在疾病发作后(~ 21 -25天)存活时间不长,患有早发性运动神经元变性伴瘫痪,伴有致命性癫痫发作的其它临床相对症状。该实例描述了纯合Del122ex3 eEF1A2无效小鼠的基因疗法研究。

[0296] 材料和方法

[0297] 在发育过程中监测经治疗小鼠和未经治疗小鼠的存活期和体重。在P21-25日龄的关键窗口之间也进行了行为测试。

[0298] 动物福利

[0299] 所有动物实验均按照英国内政部和1986年的动物(科学程序)法案进行,并在伦敦大学学院伦理审查委员会的指导方针内进行。将杂合De122ex3小鼠进行时间交配以产生混合基因型窝。在P0处将幼崽使用引物5'-3'5'-TGAGTTGTGCCTCTACCCTT-3'和5'-TACAGGCACATCCCAGGTGT-3'进行基因分型。

[0300] 测试品和控制品的施用

[0301] 以 2×10^{11} vg/幼崽(V3高剂量)或 2×10^{10} vg/幼崽(V3低剂量)的剂量向新生纯合De122ex3小鼠(De122ex3)幼崽施用10 μ L(每个半球5 μ L,双侧)V3载体或调配物缓冲液(FBS)的脑室内注射。使用以下文献中划定的注射部位坐标,通过33号汉密尔顿针(Fisher Scientific®,英国拉夫堡)向野生型和纯合De122ex3小鼠施用调配物缓冲溶液(FBS,5 μ L双侧)作为对照:Kim等人Kim,J.Y.等人《可视实验杂志》91:51863(2014)。Lambdoid缝合线在新生幼崽中是可鉴定的,并且预期的注射部位是从lambdoid缝合线到眼睛的2/5,位于矢状缝合线约0.8mm-1mm的外侧,介于 λ 和前囟之间。注射后,幼崽被送回母鼠。

[0302] 体重

[0303] 每天从P1-P30收集个体动物体重,此后每周一次。重量减轻 $\geq 15\%$ 被用作安乐死的人道终点标准。

[0304] 行为测试

[0305] 转棒仪训练和测试发生在P21-25日(P21被认为是训练日)。所有行为测定均由对动物治疗组不知情的研究人员进行。将小鼠放置在转棒仪(哈佛仪器公司(Harvard Apparatus))上,从4-40r.p.m.连续加速,持续最多2分钟。记录小鼠从棒上掉下来的时间,在每天的测试中对每只动物进行3次试验(掉棒潜伏期)。

[0306] 使用握力计(**Bioseb®**)测量P21-25之间的肢体肌肉力量(P21被认为是训练日)。一式三份地测量所有四肢或前肢的握力,每次测试之间休息1分钟,让小鼠休息。对于每次测试,小鼠被握住被尾巴的底部并降低到网格上,直到它用前爪或所有四只爪子抓住。

[0307] 神经评分

[0308] 壁架测试:将每只小鼠放在空笼子的壁架上并允许其自由探索。观察小鼠沿着笼子的边缘行走并将自己放进笼子里,并相应地评分:0=自信的行走和良好的着陆,1=行走时的绊倒和摇摆,2=绊倒和摇晃,从壁架滑倒但恢复;3=无法沿壁架行走

[0309] 后肢扣紧测试:抓住小鼠尾巴底部附近并悬浮在空中10秒。观察后肢的位置并评分如下:0=后肢始终向外指向远离腹部,1=后肢在超过50%的悬挂时间内略微向身体拉动,2=后肢在超过50%的悬挂时间内向下指向腹部,3=后肢在超过50%的悬挂时间内完全缩回并接触腹部。

[0310] 步态测试:将动物放在平坦的表面上,头部背对研究者,然后从后面观察它走路时的行为,其行为评分如下:0=小鼠正常移动,四肢支撑体重,腹部不接触地面,双后肢均匀参与,1=观察到轻微震颤,骨盆略微抬高或轻微蹒跚,2=严重震颤,骨盆凸起或明显的蹒跚,3=运动脱节,骨盆凸起和严重蹒跚。小鼠可能根本移动不了多少。

[0311] 脊柱后凸测试:将小鼠放在平坦的表面上,并在行走时从侧面观察,其评分如下:0=行走时容易伸直脊椎,1=轻度脊柱后凸(脊柱曲率),但大多能够在行走时伸直自己,2=无法完全伸直脊柱并保持轻度但持续的脊柱后凸,3=在行走或坐着时保持明显的后凸。

[0312] 握力测压:在P21-25岁时测量肢体肌肉力量,并将再次在P30和P60处使用握力计(Bioseb®)测量。一式三份地测量前肢的握力,每次测试之间休息1分钟,让小鼠休息。对于每次测试,小鼠被握住被尾巴的底部并降低到网格上,直到它用前爪抓住。

[0313] 结果

[0314] 图12A接受V3高剂量的De122ex3 (2×10^{11} vg/幼崽, $n=5$)、接受V3低剂量的De122ex3 (2×10^{10} vg/幼崽, $n=5$)、接受调配物缓冲溶液的De122ex3对照 ($n=3$) 和接受调配物缓冲溶液的野生型对照 ($n=6$) 的Kaplan-Meier存活图。图12B体重随时间的变化(数据均值 \pm S.E.M.)。图12C通过握力测压P22-25进行的运动评估(数据均值 \pm S.E.M)

[0315] 图12D P23处的握力测压。(数据均值 \pm S.E.M.、双向ANOVA和Tukey多重比较)。图12E通过转棒仪P22-25进行的运动评估(数据均值 \pm S.E.M)

[0316] 图12F P24处的转棒仪(数据均值 \pm S.E.M.、双向ANOVA和Tukey多重比较)。

[0317] 图12G来自P21-25的神经学评分。

[0318] 结论

[0319] 该实验证明了AAV9介导的eEF1A2在具有V3的纯合De122ex3小鼠中的表达的剂量相关益处。治疗效果通过以下来证明:存活年龄提高到出生后第36天(迄今为止评估的最后时间点),通过握力和转棒仪测量的体重增加和肌肉力量以及运动行为,以及神经学评分的正常恶化改善。与未经治疗的纯合De122ex3对照相比,在V3高剂量治疗的动物中观察到更长的存活期,直到出生后第36天,而对照仅存活到出生后第25天。与对照组相比,V3低剂量组的存活期长度也增加,直到出生后第28天。在V3高剂量组中,体重增加与野生型同窝对照相当。在未经治疗的对照组中,握力测压显示肌肉力量降低,随着野生型动物年龄的增长,握力增加。与未治疗的对照相比,V3高剂量治疗的动物的握力增加,从P22-25持续。在出生后第23天,V3高剂量治疗的小鼠表现出显著更强的握力测压($p=0.0046$),与野生型同窝动物模棱两可。V3低剂量没有效果。与未治疗和V3低剂量相比,在出生后22-25天之间,转棒仪掉棒潜伏期性能在V3高剂量下表现出持续性能,两者都显示出随着年龄的增长而下降的趋势。在出生后第24天,与未经治疗的动物相比,V3高剂量治疗的动物表现出显著更高的转棒仪性能($p=0.0016$),并且与野生型同窝动物相当。V3低剂量组没有效果。AAV9-eEF1A2益处的其它证据可以在V3治疗的动物的神经学评分中找到。De122.ex3未经治疗的对照组的神经评分随着年龄的增长而从P21-25增加,与神经行为下降一致。在V3高剂量治疗的动物中没有观察到这种情况,因为它们出生在出生后第21-25天之间与野生型同窝动物相当。与De122ex3未经治疗的动物相比,在V3低剂量组中也观察到持续较低的神经评分,至少通过P24。

序列表

<110> UCL商业有限责任公司(UCL Business, Ltd)

<120> 用于eEF1A2的基因疗法载体及其用途

<130> ROPA-019/02W0 329592-2245

<150> 63/055,775

<151> 2020-07-23

<160> 68

<170> PatentIn 3.5版

<210> 1

<211> 463

<212> PRT

<213> 智人

<400> 1

```

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

```

195	200	205
Pro Trp Phe Lys Gly Trp Lys Val Glu Arg Lys Glu Gly Asn Ala Ser		
210	215	220
Gly Val Ser Leu Leu Glu Ala Leu Asp Thr Ile Leu Pro Pro Thr Arg		
225	230	235
Pro Thr Asp Lys Pro Leu Arg Leu Pro Leu Gln Asp Val Tyr Lys Ile		
245	250	255
Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Ile Leu		
260	265	270
Arg Pro Gly Met Val Val Thr Phe Ala Pro Val Asn Ile Thr Thr Glu		
275	280	285
Val Lys Ser Val Glu Met His His Glu Ala Leu Ser Glu Ala Leu Pro		
290	295	300
Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Asp Ile		
305	310	315
Arg Arg Gly Asn Val Cys Gly Asp Ser Lys Ser Asp Pro Pro Gln Glu		
325	330	335
Ala Ala Gln Phe Thr Ser Gln Val Ile Ile Leu Asn His Pro Gly Gln		
340	345	350
Ile Ser Ala Gly Tyr Ser Pro Val Ile Asp Cys His Thr Ala His Ile		
355	360	365
Ala Cys Lys Phe Ala Glu Leu Lys Glu Lys Ile Asp Arg Arg Ser Gly		
370	375	380
Lys Lys Leu Glu Asp Asn Pro Lys Ser Leu Lys Ser Gly Asp Ala Ala		
385	390	395
Ile Val Glu Met Val Pro Gly Lys Pro Met Cys Val Glu Ser Phe Ser		
405	410	415
Gln Tyr Pro Pro Leu Gly Arg Phe Ala Val Arg Asp Met Arg Gln Thr		
420	425	430
Val Ala Val Gly Val Ile Lys Asn Val Glu Lys Lys Ser Gly Gly Ala		
435	440	445
Gly Lys Val Thr Lys Ser Ala Gln Lys Ala Gln Lys Ala Gly Lys		
450	455	460

<210> 2

<211> 1389

<212> DNA

<213> 智人

<400> 2

atgggcaagg agaagacca catcaacatc gtggtcatcg gccacgtgga ctccgaaag

60

tccaccacca	cgggccacct	catctacaaa	tgcggaghta	ttgacaaaag	gaccattgag	120
aagttcgaga	aggaggcggc	tgagatgggg	aaggatcct	tcaagtatgc	ctgggtgctg	180
gacaagctga	aggcggagcg	tgagcgcggc	atcacatcg	acatctccct	ctggaagttc	240
gagaccacca	agtactacat	caccatcatc	gatgcccccg	gccaccgca	cttcatcaag	300
aacatgatca	cgggtacatc	ccaggcggac	tgcgcagtgc	tgatcgtggc	ggcgggctg	360
ggcgagtctg	aggcgggcat	ctccaagaat	gggcagacgc	gggagcatgc	cctgctggcc	420
tacacgctgg	gtgtgaagca	gctcatcgtg	ggcgtgaaca	aatggactc	cacagagccg	480
gcctacagcg	agaagcgcta	cgacgagatc	gtcaaggaag	tcagcgccta	catcaagaag	540
atcggctaca	acccggccac	cgtgcccttt	gtgccatct	ccggctggca	cggtgacaac	600
atgctggagc	cctcccccaa	catgccgtgg	ttcaagggt	ggaagggtga	gcgtaaggag	660
ggcaacgcaa	gcggcgtgtc	cctgctggag	gcctggaca	ccatctgcc	ccccacgcgc	720
cccacggaca	agccccctgcg	cctgccgctg	caggacgtgt	acaagattgg	cggcattggc	780
acgggtgccc	tgggccgggt	ggagaccggc	atcctgcggc	cgggcatggt	ggtgaccttt	840
gcgccagtga	acatcaccac	tgaggtgaag	tcagtggaga	tgcaccacga	ggctctgagc	900
gaagctctgc	ccggcgacaa	cgtcggcttc	aatgtgaaga	acgtgtcgg	gaaggacatc	960
cggcggggca	acgtgtgtgg	ggacagcaag	tctgaccgc	cgcaggaggc	tgctcagttc	1020
acctcccagg	tcatcatcct	gaaccacccg	gggcagatta	gcgccggcta	ctccccggtc	1080
atcgactgcc	acacagccca	catgcctgc	aagtttgcgg	agctgaagga	gaagattgac	1140
cggcgcctctg	gcaagaagct	ggaggacaac	ccaagtccc	tgaagtctgg	agacgcggcc	1200
atcgtggaga	tggtgccggg	aaagcccatg	tgtgtggaga	gcttctcca	gtacccgctt	1260
ctcggccgct	tcgccgtgcg	cgacatgagg	cagacggtgg	ccgtaggctg	catcaagaac	1320
gtggagaaga	agagcggcgg	cgccggcaag	gtaccaagt	cggcgcagaa	ggcgcagaag	1380
gcgggcaag						1389
<210>	3					
<211>	448					
<212>	DNA					
<213>	智人					
<400>	3					
agtgcaagtg	ggttttagga	ccaggatgag	gcggggtggg	ggtgcctacc	tgacgaccga	60
ccccgaccca	ctggacaage	acccaacccc	cattcccaa	attgcgcatc	ccctatcaga	120
gagggggagg	ggaaacagga	tgcggcgagg	cgcgtgcga	ctgccagctt	cagcaccgcg	180
gacagtgcct	tcgccccgc	ctggcggcgc	gcgccaccgc	cgcctcagca	ctgaaggcgc	240
gctgacgtca	ctcgcggtc	ccccgcaaac	tcccctccc	ggccaccttg	gtcgcgtccg	300
cgccgccgcc	ggcccagccg	gaccgcacca	cgcgaggcgc	gagatagggg	ggcacgggcg	360
cgaccatctg	cgctgcggcg	ccggcgactc	agcgtgcct	cagtctgcgg	tgggcagcgg	420
aggagtctgtg	tcgtgcctga	gagcgcag				448
<210>	4					
<211>	1392					
<212>	DNA					

<213> 人工序列

<220>

<223> 密码子优化的eEF1A2多核苷酸变体

<400> 4

atgggtaaag aaaaaacaca tattaatata gtagtaatcg gtcattgtga ctctggaaaa	60
tctactacta caggacattt gatttataaa tgtggaggaa ttgataaaag aacaatagaa	120
aaatttgaaa aagaagctgc tgaaatgggt aaaggtagtt ttaaatatgc ttgggttttg	180
gataaattga aagctgaaaag agaaaagagga attacaattg atatttcttt gtggaaattt	240
gaaactacaa aatattatat aacaataata gatgctcctg gacatagaga ttttattaaa	300
aatatgatta caggaacttc tcaagcagat tgtgctgttt tgatagtagc agcaggagtt	360
ggatgaattcg aagcaggcat ttctaaaaat ggacaaacta gagaacatgc tttgttggct	420
tatacattgg gcgtaaaaaca attgattgta ggagttaata aatggattc tactgaacct	480
gcatattctg aaaaaagata tgatgaaata gtaaaagaag tttctgctta tattaaaaaa	540
attggttata atcctgctac agttccattt gttcctattt ctggatggca tggagataat	600
atgttggaac ctatgcctaa tatgccttgg tttaaaggat ggaaagtga aaggaaagaa	660
ggaaatgcat caggagtctc cttgttggaa gctttggata caatcttgcc tccaacaaga	720
cctacagata aacctttgag attgcctctt caagatgtat ataaaatagg aggaatagga	780
acagtgccag ttggaagagt agaaacaggt atattgagac ctggaatggt tgtaacattt	840
gcaccagtta atataactac tgaagtaaaa tctgttgaaa tgcattcatga agctttgtct	900
gaagctcttc ctggagataa ttaggattt aatgttaaaa atgtaagtgt aaaagatata	960
agaagaggaa atgtatgtgg tgatagtaa tcagatccac ctcaagaagc agctcaattt	1020
acatcacaag taataatatt gaatcatcct ggacaaattt ctgcaggata ttcaccagta	1080
atagattgtc atacagcaca tatagcttgt aaatttgctg aattgaaaga aaaattgat	1140
agaagaagtg gaaaaaact tgaagataat cctaaatcat tgaaatcagg agatgcagct	1200
attgtagaaa tggtagctgg aaaaccaatg tgtgtagaat ctttttctca atatccacct	1260
ctcggaaagat ttgctgttag agatatgaga caaacagttg cagtaggagt tattaaaaat	1320
gtagaaaaaa aaagcggagg tgcaggaaag gttacaaaat ccgcacaaaa agctcaaaaa	1380
gctggtaaat aa	1392

<210> 5

<211> 1392

<212> DNA

<213> 人工序列

<220>

<223> 密码子优化的eEF1A2多核苷酸变体

<400> 5

atgggcacaag aaaaaacaca tataaacatt gtcgttatcg gacacgttga ttctggtaaa	60
agtacaacaa ccggtcactt gatatacaaa tgcgggggta tagacaaacg cactattgaa	120
aagttcgaga aagaagctgc ggagatgggc aaaggtcat tcaagtacgc gtgggtactc	180
gataagttga aagctgaacg cgagagggga atcacatag acatctcact ttggaattc	240

gagacaacca agtattacat aactattata gatgccccag gccacagga tttcattaaa	300
aatatgataa ccggcacatc tcaagccgat tgcgccgtac tcatcgtcgc cgctgggtgtg	360
ggtgagttcg aggcaggtat ttctaaaaat ggccagacac gcgaacatgc tcttctggct	420
tatacactcg gggttaaaca gctcatagta ggagtgaata agatggactc cactgaaccc	480
gcctattcag agaagcgcta tgacgaaatt gtaaaggagg tctcagcata tattaaaaaa	540
attggctata acccagccac ggtgccattc gtcccgatta gtggatggca tggtgacaat	600
atgctggaac caagtcccaa tatgccttgg ttttaagggtt ggaaagtaga gcggaaagag	660
ggtaatgctt ccggcgtgtc attgctggag gcgcttgaca cgatactccc acccacaagg	720
ccaactgata agccactccg attgcccttg caggacgtgt acaagattgg gggaattggg	780
actgtgcccc tccggcgcgt ggagacgggc atcctcagac ctgggatggg agtcactttt	840
gccccgtca acataacgac tgaagttaa tcagtggaaa tgcatacga agctttgagt	900
gaggcgcttc ccggagataa cgttggattt aatgtcaaaa atgtctcctg taaagatata	960
agaagaggaa acgtctgcgg tgactcaaag tcagaccac cacaggaggc tgctcaattt	1020
acgagtcaag taataattct gaatcaccct gggcaataa gtgcgggata ctctccagtc	1080
atcgattgtc acaccgcca tattgcatgt aagttcgcag aacttaagga aaagatcgac	1140
cgaagaagcg gaaaaaaatt ggaagataat ccgaaaagtt tgaaaagcgg tgacgcggcg	1200
attgtagaga tggtccttgg caaacgatg tgtgtggagt ctttcagtca atatccacca	1260
ctcggtcgct ttgccgtgcg ggatatgca cagaccgttg ctgtcggcgt aataaaaaac	1320
gtcgaaaaaa agagcgggtg ggctggaaaa gttacaaaat ccgctcaaaa ggcacagaag	1380
gcgggcaagt ga	1392
<210> 6	
<211> 1392	
<212> DNA	
<213> 人工序列	
<220>	
<223> 密码子优化的eEF1A2多核苷酸变体	
<400> 6	
atgggtaaag aaaagaccca cattaacata gtagtaatcg gtcattgtga ctctgggaaa	60
agcactacta ccggacattt gatctataaa tgtgggggca tcgacaaaag aacgatagag	120
aagtttgaga aggaggcggc ggagatgggt aaaggtagtt ttaagtacgc ttgggttttg	180
gacaaattga aagccgagcg cgagcgcggc attaccattg acatttctct ctggaaattc	240
gaaactacga agtattatat aacaataata gacgccccg gccatcggga ctttattaaa	300
aacatgatta caggaactag ccaagcagat tgtgctgtgc tgatagtagc ggcaggggtc	360
ggggagtctg aagcaggcat ctctaaaaat ggacaaactc gagagcacgc cttgttggct	420
tataccttgg gcgtaaagca gctgatcgta ggagttaata aatggattc cactgaaccc	480
gcataatagc gaaagcgata tgacgaaata gtaaaggaag tctcagctta tatcaagaaa	540
atcggttaca atcctgcgac ggttccattc gttcctatct ccgggtggca cggcgataat	600
atgcttgagc ccagtcccaa tatgccttgg ttcaagggtt ggaaggttga gaggaaggaa	660
ggcaatgcat caggcgtcag cttgttggaa gctctcgaca ccatcctgcc gccacgagg	720

cccacagaca aaccgttgcg actgcctctt caagatgtat acaaaatagg cgggatagga	780
accgtgccgg ttggacgagt agagacgggt atactgcggc cgggaatggt cgtgacgttt	840
gcacccgtga atataactac tgaggtgaag agcgtcgaga tgcacatga agcgtgagt	900
gaagctctcc ctggcgataa cgtagggttc aacgtgaaaa acgtaagtgt aaaggatata	960
aggcgcggaa atgtatgtgg tgacagtaaa agcgaccgcg cgcaagaggc ggcgcaattc	1020
acatcacagg taataatatt gaatcacccc ggccaaattt cgcaggcta ctcaccagtc	1080
atagattgcc acaccgccc catagcttgt aagttcgctg agttgaaaga gaagattgat	1140
agacgaagtg ggaagaaaact tgaagacaat ccgaagtccc tgaagtccgg tgacgcagcg	1200
attgtagaaa tggtagcggg caagccaatg tgtgtagagt ctttcagcca gtaccacca	1260
ctggggcggg tcgcggtgcg agacatgagg caaacggttg cggtcggcgt cattaanaat	1320
gtcgaaaaaa agagtggcgg tgcaggtaag gtcacaaaaa gcgcacaaaa ggcccagaaa	1380
gccgtaagt ga	1392
<210> 7	
<211> 1392	
<212> DNA	
<213> 人工序列	
<220>	
<223> 密码子优化的eEF1A2多核苷酸变体	
<400> 7	
atgggaaagg aaaaaactca cataaacatt gtcgtcatcg gtcacgtaga cagtggcaaa	60
tcaacgacca ctggacatct catctataag tgtggcggta ttgacaaacg cactatcgag	120
aaattcgaaa aggaggctgc tgagatgggc aaaggctctt tcaagtacgc atgggtcctg	180
gataagctga aagcggagcg agagagagg atcaccatcg atatatctct gtggaaattt	240
gaaaccacca agtactacat cacaattatt gatgccccag gtcataggga ttttatcaag	300
aacatgatca ccgggacaag ccaagccgac tgcgcagttc tcatagtggc ggctggagta	360
ggggagtttg aagcagggat atctaagaat ggacagacc gcgagcacgc cttgctggcc	420
tacaccctgg gagtgaagca gtcatagtt ggcgtcaata agatggacag caccgaacct	480
gcctacagtg agaagagta tgacgagatt gtgaaggagg tttctgctta cattaanaag	540
attggctata acccagctac tgtcccattc gttccaatca gcggctggca cggtgataac	600
atgctggagc ctagtcccaa catgccgtgg ttcaaggggt ggaaggttga acgcaaggag	660
gggaatgcct caggcgttcc cctgctggag gcctcgata caatactccc cccgaccggg	720
cctacagata aaccgtgcg actgcctctt caggacgtgt ataaaatcgg gggaatcggc	780
acagtccccg tgggcagggt agagactggc atcttgcggc ctggaatggt agtcaccttt	840
gccccgtta atatacacaac ggaggtgaaa tctgtggaga tgcatacga agcactgagc	900
gaggctctgc ctggtgacaa cgtgggattt aacgtcaaaa acgtgtcagt caaggacatc	960
cgcccggtta acgtttgcgg agattctaag tccgatcccc cccaggaggc agcccaattt	1020
acctccaag tgatcattct gaatcaccca ggccaaattt ccgccggta tccccctgtg	1080
attgactgtc acacagcaca catcgcagtc aaattcgccg aactcaagga gaaaattgat	1140
cggagaagcg gtaaaaaact ggaggacaac ccaaagtccc tcaagtctgg ggatgccgcc	1200

atcgtggaga tggtagcagg caaacctatg tgcgtggaaa gttttagcca gtaccctcca	1260
ctgggtcgct ttgctgttcg ggatatgcgg cagacagtag cggttgggggt cataaaaaac	1320
gtcgagaaaa agagcggagg agctgggaaa gttaccaaat ccgcacagaa ggcacaaaaa	1380
gccggaaaaat ga	1392
<210> 8	
<211> 1392	
<212> DNA	
<213> 人工序列	
<220>	
<223> 密码子优化的eEF1A2多核苷酸变体	
<400> 8	
atgggcaaaag agaaaacaca tattaacatt gttgttatcg ggcacgttga tagcggcaag	60
tccactacca ctggccatct gatttacaag tgcggcgaa tcgataaacg aactattgaa	120
aagttcgaaa aagaagccgc cgagatggga aagggtcct ttaaatacgc ttgggtcctc	180
gataaaactca aagcagaacg ggagagagga atcaccatcg atatatcctt gtggaagttc	240
gaaactacaa aatattacat taccatcatt gatgcgctg ggcaccgca cttcattaag	300
aacatgatta ctggcacctc tcaagccgac tgcgcagtgt tgatcgtagc cgcaggcgtc	360
ggggagttcg aagctgggat cagcaagaac ggcagacta ggaacacgc tctgctcgca	420
tatactcttg gcgtgaaaca gttgatcgtt ggcgtgaaca agatggattc aactgagcct	480
gcctattctg agaaacgata cgacgagatt gtgaaagagg tttcagctta catcaagaaa	540
attgggtata atcccgaac agttccctc gtgcccctc ctgggtggca cggcgacaac	600
atgctcgaac catccccaaa tatgccatgg ttcaaggat ggaaggtgga gcgcaaagaa	660
ggcaacgcct ccggagtgc tctgctcgag gccctggaca ccattctgcc cccaacacga	720
cccactgata agcctctgag actgccactg caagacgttt aaaaattgg gggaattgga	780
accgtgcctg tgggtcgggt ggaaaccgga atcctcagac ccggcatggt ggtcaccttc	840
gcaccagtga atataacgac agaggtcaaa tctgtggaga tgcaccatga ggcattgagc	900
gaggcactcc caggagacaa cgtgggtttc aacgtgaaaa atgtctcagt taaggacatc	960
cgacgcggca acgtgtcgg agatagcaaa tctgaccccc ccaggaggc cgctcaattc	1020
acaagtcagg ttatcatcct taatcacct ggccaaatat ctgcaggcta cagccccgtg	1080
atcgattgtc acacagctca tatcgctgt aaatttgctg aactcaaaga aaagattgac	1140
cgcagatcag gaaaaaagct ggaggacaac ctaaaagtc tgaagtccgg cgacgctgcc	1200
atcgtggaga tggcccttg gaaacccatg tgcgtggagt ctttttctca gtacccccct	1260
ctgggacgat tcgccgtgcg cgacatgaga cagactgtcg ccgtgggcgt cattaaaaat	1320
gtggaaaaaaa aatcaggagg tgcagggaaa gtgacaaaga gtgccagaa agcacagaag	1380
gctggcaagt ga	1392
<210> 9	
<211> 1395	
<212> DNA	
<213> 人工序列	

<220>		
<223>	具有Kozak序列的人eEF1A2	
<400>	9	
	gccaccatgg gcaaggagaa gaccacatc aacatcgtgg tcatcggcca cgtggactcc	60
	ggaaagtcca ccaccacggg ccacctcatc tacaaatgcg gaggtattga caaaaggacc	120
	attgagaagt tcgagaagga ggcggctgag atggggaagg gatccttcaa gtatgcctgg	180
	gtgctggaca agctgaaggc ggagcgtgag cgcggcatca ccatcgacat ctccctctgg	240
	aagttcgaga ccaccaagta ctacatcacc atcatcgatg cccccggcca ccgcgacttc	300
	atcaagaaca tgatcacggg tacatcccag gcggaactgc cagtgtgat cgtggcggcg	360
	ggcgtgggcg agttcgaggc gggcatctcc aagaatgggc agacgcggga gcatgccctg	420
	ctggcctaca cgctgggtgt gaagcagctc atcgtgggcg tgaacaaaat ggactccaca	480
	gagccggcct acagcgagaa gcgctacgac gagatcgtea aggaagtcag cgcctacatc	540
	aagaagatcg gctacaaccc ggccaccgtg ccctttgtgc ccatctccgg ctggcacggt	600
	gacaacatgc tggagccctc cccaacatg ccgtggttca agggctggaa ggtggagcgt	660
	aaggagggca acgcaagcgg cgtgtccctg ctggaggccc tggacacat cctgcccccc	720
	acgcgccccca cggacaagcc cctgcgcctg ccgctgcagg acgtgtacaa gattggcggc	780
	attggcacgg tgcccgtggg ccgggtggag accggcatcc tgcggccggg catggtggtg	840
	acctttgcgc cagtgaacat caccactgag gtgaagtcag tggagatgca ccacgaggct	900
	ctgagcgaag ctctgcccgg cgacaacgtc ggcttcaatg tgaagaacgt gtcggtgaag	960
	gacatccggc ggggcaacgt gtgtggggac agcaagtctg acccgccgca ggaggctgct	1020
	cagttcacct cccaggtcat catcctgaac caccgggggc agattagcgc cggctactcc	1080
	ccggtcatcg actgccacac agcccacatc gcctgcaagt ttgcggagct gaaggagaag	1140
	attgaccggc gctctggcaa gaagctggag gacaacccca agtcctgaa gtctggagac	1200
	gcggccatcg tggagatggt gccgggaaag cccatgtgtg tggagagctt ctcccagtac	1260
	ccgctctcgc gccgcttcgc cgtgcgcgac atgaggcaga cgggtggccgt aggcgtcatc	1320
	aagaacgtgg agaagaagag cggcggcgcc ggcaaggtca ccaagtcggc gcagaaggcg	1380
	cagaaggcgg gcaag	1395
<210>	10	
<211>	10	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	Kozak序列基序	
<400>	10	
	gccaccatgg	10
<210>	11	
<211>	13	
<212>	RNA	
<213>	人工序列	

<220>		
<223>	替代性Kozak序列基序	
<400>	11	
	gccgccrcca ugg	13
<210>	12	
<211>	10	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	替代性Kozak序列基序	
<400>	12	
	gacaccaugg	10
<210>	13	
<211>	168	
<212>	DNA	
<213>	腺相关病毒	
<400>	13	
	tacgtagata agtagcatgg cgggtaatc attaactaca aggaaccct agtgatggag	60
	ttggcactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc	120
	cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgc	168
<210>	14	
<211>	573	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	实验室制造-部分人β疱疹病毒5中的CAG启动子	
<400>	14	
	acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat	60
	aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga	120
	gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc	180
	ccctattgac gtcaatgacg gtaaattggc cgctggcat tatgcccagt acatgacctt	240
	atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtcga	300
	ggtgagcccc acgttctgct tcaactctcc catctcccc cctctcccc ccccaatddd	360
	gtatdddattt atdddttaat tatdddgtgc agcgatgggg gcgggggggg ggggggcgcg	420
	cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg	480
	cagccaatca gagcggcgcg ctccgaaagt ttcctdddttat ggcgaggcgg cggcggcggc	540
	ggccctataa aaagcgaagc gcgcggcggg cgg	573
<210>	15	
<211>	736	

<212> PRT
 <213> 腺相关病毒9
 <400> 15
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
 260 265 270
 Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285

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

595	600	605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His		
610	615	620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met		
625	630	635
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala		
645	650	655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr		
660	665	670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln		
675	680	685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn		
690	695	700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val		
705	710	715
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu		
725	730	735

<210> 16

<211> 220

<212> DNA

<213> 人β疱疹病毒5

<400> 16

tggtgatgcg gttttggcag tacaccaatg ggcgtggata gcggtttgac tcacggggat	60
ttccaagtct ccacccatt gacgtcaatg ggagtttggt ttggcaccaa aatcaacggg	120
actttccaaa atgtcgtaat aaccccgcc cgttgacgca aatgggcggt aggcgtgtac	180
ggtgggaggt ctatataagc agagctcgtt tagtgaaccg	220

<210> 17

<211> 583

<212> DNA

<213> 人β疱疹病毒5

<400> 17

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggcccgcc tggetgaccg cccaacgacc cccgccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	180
atgggtggag tatttacggt aaactgccc cttggcagta catcaagtgt atcatatgcc	240
aagtacgcc cctattgacg tcaatgacgg taaatggccc gectggcatt atgcccagta	300
catgacctta tgggactttc ctacttggea gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccaccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480

ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt	540
acgggtgggag gtctatataa gcagagctgg ttagtgaac cgt	583
<210> 18	
<211> 141	
<212> DNA	
<213> 腺相关病毒	
<400> 18	
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
actccatcac taggggttcc t	141
<210> 19	
<211> 168	
<212> DNA	
<213> 腺相关病毒	
<400> 19	
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg	60
tcgccggcc tcagtgagcg agcgagcgcg cagagagga gtggccaact ccatcactag	120
gggttccttg tagttaatga ttaaccgcc atgctactta tctacgta	168
<210> 20	
<211> 170	
<212> DNA	
<213> 腺相关病毒	
<400> 20	
ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc ccgggcgctc ggcgaccttt	60
ggtcgcccgg cctcagtgag cgagcgagcg cgagagagg gtagtgccaa ctccatcact	120
aggggttcct ttagtgaat gattaaccgc ccatgctact tatctacgta	170
<210> 21	
<211> 141	
<212> DNA	
<213> 腺相关病毒	
<400> 21	
aggaaccctt agtgatggag ttggcactc cctctctgcg cgctcgctcg ctactgagg	60
ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc	120
gagcgcgcag ctgcctgcag g	141
<210> 22	
<211> 124	
<212> DNA	
<213> 人工序列	
<220>	

<223> 实验室制造-载体填充序列	
<400> 22	
gcggaattc agtcgataac tataacggtc ctaaggtagc gatttaaata cgcgctctct	60
taaggtagcc ccgggacgcg tcaattgact acaaaccgag tatctgcaga gggccctgcg	120
tatg	124
<210> 23	
<211> 84	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-载体填充序列	
<400> 23	
cttctgaggc ggaaagaacc agatcctctc ttaaggtagc atcgagattt aaattagga	60
taacagggta atggcgcggg ccgc	84
<210> 24	
<211> 63	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-载体填充序列	
<400> 24	
gttaccagg ctggagtgca gtggcacatt tctgctcact gcaacctcct cctccctggg	60
ttc	63
<210> 25	
<211> 253	
<212> DNA	
<213> 智人	
<400> 25	
gcccagcacc ccaaggcggc caacgccaaa actctcctc ctctcttcc tcaatctcgc	60
tctcgtcttt ttttttttc gcaaaaggag gggagagggg gtaaaaaaat gctgcactgt	120
gcggcgaagc cggtagtga gcggcgcggg gccaatcagc gtgcgccgtt ccgaaagttg	180
ccttttatgg ctcgagcggc cgcggcggcg ccctataaaa ccagcggcg cgacgcgcca	240
ccaccgccga gtc	253
<210> 26	
<211> 281	
<212> DNA	
<213> 原鸡(Gallus gallus)	
<400> 26	
ggtcgaggtg agccccacgt tctgcttcac tctccccate tccccccct cccaccccc	60

aatTTTgtat ttatttattt tTtaattatt ttgtgcagcg atgggggcgg gggggggggg	120
ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg	180
cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc	240
ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcggg a	281
<210> 27	
<211> 455	
<212> DNA	
<213> 智人	
<400> 27	
caacctttgg agctaagcca gcaatggtag agggaagatt ctgcacgtcc cttccaggcg	60
gcctccccgt caccaccccc cccaaccgc cccgaccgga gctgagagta attcatacaa	120
aaggactcgc ccctgccttg gggaatcca gggaccgtcg ttaaactccc actaacgtag	180
aaccagaga tcgctgcgtt cccgccccct caccgcccc ctctcgtcat cactgaggtg	240
gagaatagca tgcgtgaggc tccggtgcc gtcagtgggc agagcgcaca tcgcccacag	300
tccccagaaa gttgggggga ggggtcggca attgaacggg tgcctagaga aggtggcgcg	360
gggtaaactg ggaaagtgat gtcgtgtact ggctccgct ttttcccgag ggtgggggag	420
aaccgtatat aagtgcagta gtcgccgtga acgtt	455
<210> 28	
<211> 401	
<212> DNA	
<213> 智人	
<400> 28	
agtgcaagtg ggttttagga ccaggatgag gcgggggtgg ggtgcctacc tgacgaccga	60
ccccgaccca ctggacaagc acccaacccc cattcccaa attg'gcatc ccctatcaga	120
gagggggagg ggaaacagga tgcggcgagg cgcgtgcga ctgccagctt cagcaccgcg	180
gacagtgcct tcgccccgc ctggcggcgc gcgccaccgc cgcctcagca ctgaaggcgc	240
gctgacgtca ctgcgggtc ccccgcaaac tccccttccc ggccaccttg gtcgcgtccg	300
cgccgccgcc ggcccagccg gaccgcacca cgcgaggcgc gagatagggg ggcacggcgc	360
cgaccatctg cgctgcggcg ccggcgactc agcgtgcct c	401
<210> 29	
<211> 422	
<212> DNA	
<213> 智人	
<400> 29	
ctgcagaggg ccctgcgtat gagtgcaagt gggttttagg accaggatga ggcggggtgg	60
gggtgcctac ctgacgaccg accccgacc actggacaag cacccaaccc ccattccca	120
aattgcgcat cccctatcag agagggggag gggaacagg atgcggcgag gcgcgtgcgc	180
actgccagct tcagcaccgc ggacagtgcc ttcgccccg cctggcggcg cgcgccaccg	240
ccgcctcagc actgaaggcg cgctgacgtc actgcgggt cccccgaaa ctccccttc	300

cgccacctt ggtcgcgtcc gcgccgccgc cggcccagcc ggaccgcacc acgcgaggcg	360
cgagataggg gggcacgggc gcgaccatct gcgctgcggc gccggcgact cagcgctgcc	420
tc	422
<210> 30	
<211> 281	
<212> DNA	
<213> 智人	
<400> 30	
acttgtggac aaagtttct ctattccacc tectccaggc cctccttggg tccatcacc	60
caggggtgct ggggccatcc cacccccagg cccacacagg cttgcagtat tgtgtgcggt	120
atggtcaggg cgtccgagag caggtttcgc agtggaggc aggcaggtgt tggggaggca	180
gttaccgggg caacgggaac agggcgtttt ggaggtggtt gccatgggga cctggatgct	240
gacgaaggct cgcgaggctg tgagcagcca cagtgcctg c	281
<210> 31	
<211> 293	
<212> DNA	
<213> 人β疱疹病毒5	
<400> 31	
acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgccat tgacgtcaat	60
aatgacgtat gttcccatag taacgccaat aggactttc cattgacgtc aatgggtgga	120
gtatttacgg taaactgccc acttggcagt acatcaagt tatcatatgc caagtacgcc	180
ccctattgac gtcaatgacg gtaaattggc cgctggcat tatgcccagt acatgacctt	240
atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta cca	293
<210> 32	
<211> 953	
<212> DNA	
<213> 智人	
<400> 32	
cgcgctccgcc cgcgagcaca gagcctcgcc tttgccgatc cgccgccgt ccacaccgc	60
cgccaggtaa gcccggccag ccgaccgggg catgcggccg cggcccttcg cccgtgcaga	120
gccgccgtct gggccgcagc ggggggcgca tggggcgga cggaccgcc gtggggggcg	180
cgggagaagc ccctgggcct ccggagatgg gggacacccc acgccagttc gcaggcgcga	240
ggccgcgctc gggcgggcgc gctccggggg tgccgctctc ggggcggggg caaccggcgg	300
ggtctttgtc tgagccgggc tcttgccaat ggggatcgca cggtaggggc ggcgtagccc	360
ccgtcaggcc cggtaggggc tggggcgcca tgcgctgctc cgctggtcct ttgggcgcta	420
actgcgtgct cgctgggaat tggcgcta tgcgctgctc cgctgggact caatggcgt	480
aatcgcgct gcgttctggg gcccgggcgc ttgcgccact tctgcccga gccgctggcg	540
cccgaggggtg tggccgctgc gtgcgcgcgc gcgaccgggt cgctgtttga accgggcgga	600
ggcggggctg gcgcccgtt gggagggggg tggggcctgg cttcctgccg cgcgccgcgg	660

ggacgcctcc gaccagtgtt tgccttttat ggtaataacg cggccggccc ggcttccttt	720
gtccccaatc tgggcgcgcg ccggcgcccc ctggcggcct aaggactcgg cgcgccggaa	780
gtggccaggg cggcagcggc tgctcttggc ggccccgagg tgactatagc cttcttttgt	840
gtcttgatag ttcgccagcc tctgctaacc atgttcatgc cttcttcttt ttcctacagc	900
tcctgggcaa cgtgctggtt attgtgctgt ctcatcattt tggcaaagaa ttc	953
<210> 33	
<211> 1068	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-鸡β-肌动蛋白外显子/内含子加兔珠蛋白内含子	
<400> 33	
gtcgtgcgc gctgccttcg ccccggtccc cgetccgccg ccgctcgcg ccgccccccc	60
cggtcttgac tgaccgcgtt actcccacag gtgagcgggc gggacggccc ttctctccg	120
ggctgtaatt agcgccttgg ttaatgacgg cttgtttctt ttctgtggct gcgtgaaagc	180
cttgaggggc tccgggaggg ccctttgtgc ggggggagcg gctcgggggg tgcgtgcgtg	240
tgtgtgtgcg tggggagcgc cgcgtgcggc tccgcgtgc ccggcggctg tgagcgtgc	300
gggcgcggcg cggggctttg tgcgtccgc agtgtgcgc aggggagcgc ggccgggggc	360
ggtgccccgc ggtgcggggg gggctgcgag gggaacaaag gctgcgtgcg ggtgtgtgc	420
gtgggggggt gagcaggggg tgtgggcgcg tcggtcgggc tgcaaccccc cctgcacccc	480
cctccccgag ttgctgagca cggcccggct tcgggtgcgg ggctccgtac ggggcgtggc	540
gcggggctcg ccgtgccggg cggggggtgg cggcaggtgg gggtgccggg cggggcgggg	600
ccgcctcggg ccggggaggg ctccggggag gggcgcggcg gccccggag cgccggcggc	660
tgtcgaggcg cggcgagccg cagccattgc cttttatggt aatcgtgcga gagggcgcag	720
ggacttcctt tgtcccaaat ctgtgcggag ccgaaatctg ggagcgcgg ccgcaccccc	780
tctagcgggc gcggggcgaa gcggtgcggc gccggcagga aggaaatggg cggggagggc	840
cttcgtgcgt cgccgcgccg ccgtcccctt ctccctctcc agcctcgggg ctgtcccgcg	900
ggggacggct gccttcgggg gggacggggc agggcggggt tcggcttctg gcgtgtgacc	960
ggcggctcta gagcctctgc taaccatggt catgcettet tcttttctt acagctcctg	1020
ggcaacgtgc tggttattgt gctgtctcat cattttggca aagaattc	1068
<210> 34	
<211> 126	
<212> DNA	
<213> 智人	
<400> 34	
agtctgcggt gggcagcggg ggagtcgtgt cgtgcctgag agcgcagctg tgctcctggg	60
caccgcgcag tccgccccg cggtccttgg ccagaccacc cctaggacc cctgccccaa	120
gtcgca	126
<210> 35	

<211>	121		
<212>	DNA		
<213>	人β疱疹病毒5		
<400>	35		
	tcagatcgcc tggagaggcc atccacgctg ttttgacctc catagtggac accgggaccg		60
	atccagcctc cgcggccggg aacggtgcat tggaacgcgg attccccgtg ccaagagtga		120
	c		121
<210>	36		
<211>	512		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	实验室制造-腺病毒衍生的增强子元件		
<400>	36		
	ctcactctct tccgcatcgc tgtctgcgag ggccagctgt tgggctcgcg gttgaggaca		60
	aactcttcgc ggtctttcca gtactcttgg atcggaaacc cgtcggcctc cgaacggtac		120
	tccgccaccg agggacctga gcgagtcgc atcgaccgga tcggaaaacc tctcgagaaa		180
	ggcgtctaac cagtcacagt cgcaaggtag gctgagcacc gtggcgggcg gcagcgggtg		240
	gcggtcgggg ttgtttctgg cggaggtgct gctgatgatg taattaaagt aggcggtctt		300
	gagacggcgg atggtcgagg tgaggtgtgg caggcttgag atccagctgt tggggtgagt		360
	actccctctc aaaagcgggc attacttctg cgctaagatt gtcagtttcc aaaacgagg		420
	aggatttgat attcacctgg cccgatctgg ccatacactt gagtgacaat gacatccact		480
	ttgcctttct ctccacaggt gtccactccc ag		512
<210>	37		
<211>	956		
<212>	DNA		
<213>	智人		
<400>	37		
	ctttttcgca acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg		60
	cctggcctct ttacgggta tggeccttgc gtgccttgaa ttacttccac ctggctccag		120
	tacgtgattc ttgatcccga gctggagcca ggggcgggccc ttgcgcttta ggagcccctt		180
	cgctcgtgc ttgagttgag gcctggcctg ggcgctgggg ccgccgcgtg cgaatctggt		240
	ggcaccttcg cgctgtctc gctgcttctg ataagtctct agccatttaa aatTTTTgat		300
	gacgtgctgc gacgcttttt ttctggcaag atagtcttgt aatgcgggc caggatctgc		360
	acactggat ttcggttttt gggcccgcgg ccggcgacgg ggcccgtgcg tcccagcgc		420
	catgttcggc gaggcggggc ctgcgagcgc ggccaccgag aatcgacgg gggtagtctc		480
	aagctggccg gcctgctctg gtgcctggcc tcgcgccgcc gtgtatgcc ccgccctggg		540
	cggcaaggct ggcccgtgc gcaccagttg cgtgagcggg aagatggccg cttcccggcc		600
	ctgctccagg gggctcaaaa tggaggacgc ggcgctcggg agagcgggcg ggtgagtcac		660

ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc ttcatgtgac tccacggagt	720
accgggcgcc gtccaggcac ctcgattagt tctggagctt ttggagtacg tcgtctttag	780
gttgggggga ggggttttat gcgatggagt ttccccacac tgagtgggtg gagactgaag	840
ttaggccagc ttggcacttg atgtaattct ctttgggaatt tggccttttt gagtttggat	900
cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca tttcag	956
<210> 38	
<211> 939	
<212> DNA	
<213> 智人	
<400> 38	
gtaagtgccg tgtgtggttc ccgcgggcct ggctctttha cgggttatgg cccttgcgtg	60
cccttgaatta cttccacctg gctgcagtac gtgattcttg atcccagact tcgggttggg	120
agtgggtggg agagttcgag gccttgcgct taaggagccc cttegcctcg tgctttagtt	180
gaggcctggc ctgggcgctg gggccgccgc gtgcgaatct ggtggcacct tcgcgcctgt	240
ctcgtgctt tcgataagtc tctagccatt taaaattttt gatgacctgc tgcgacgctt	300
tttttctggc aagatagtct tgtaaatgcg ggccaagatc tgcacactgg tatttcggtt	360
tttggggccg cgggcggcga cggggcccgt gcgtcccagc gcacatgttc ggcgaggcgg	420
ggcctgcgag cgcggccacc gagaatcgga cgggggtagt ctcaagctgg ccggcctgct	480
ctgggtgcctg gcctcgcgcc gccgtgtatc gccccgcctt gggcggcaag gctggcccgg	540
tcggcaccag ttgcgtgagc ggaaagatgg ccgcttcccg gccctgctgc agggagctca	600
aaatggagga cgcggcgctc gggagagcgg gcgggtgagt cacccacaca aaggaaaagg	660
gcctttccgt cctcagccgt cgcttcatgt gactccacgg agtaccgggc gccgtccagg	720
cacctcgatt agttctcgag cttttggagt acgtcgtctt taggttgggg ggaggggttt	780
tatgcgatgg agtttcccca cactgagtgg gtggagactg aagttaggcc agcttggcac	840
ttgatgtaat tctccttggg atttgcctt tttgagtttg gatcttggtt cattctcaag	900
cctcagacag tggttcaaag ttttttctt ccatttcag	939
<210> 39	
<211> 83	
<212> DNA	
<213> 智人	
<400> 39	
tcagaagccc cgggctcgtc agtcaaaccg gttctctggt tgcactcggc agcacgggca	60
ggcaagtggg ccctaggttc ggg	83
<210> 40	
<211> 476	
<212> DNA	
<213> 智人	
<400> 40	
gtgagtctat gggacccttg atgttttctt tccccttctt ttctatggtt aagttcatgt	60

cataggaagg ggagaagtaa caggtacac atattgacca aatcagggtta attttgcatt	120
tgtaatttta aaaaatgctt tcttcttta atatactttt ttgtttatct tattttcta	180
actttcccta atctctttct ttcagggtcaa taatgataca atgtatcatg cctctttgca	240
ccattctaaa gaataacagt gataatttct gggtaaggc aatagcaata tttctgcata	300
taaataatttc tgcatataaa ttgtaactga tgtaagaggt ttcataattgc taatagcagc	360
tacaatccag ctaccattct gcttttattt tatggttggg ataaggctgg attattctga	420
gtccaagcta ggcccttttg ctaatcatgt tcataacctt tatcttctc ccacag	476
<210> 41	
<211> 589	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-突变的土拨鼠肝炎调节元件	
<400> 41	
aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct	60
ccttttacgc tatgtggata cgtgcttta atgcctttgt atcatgctat tgcttccgt	120
atggctttca ttttctctc cttgtataaa tcttggttg tgtctcttta tgaggagttg	180
tggcccgttg tcaggcaacg tggcgtggtg tgcactgtgt ttgctgacgc aacccccact	240
ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgctt cccctcct	300
attgccacgg cggaactcat cgccgcctgc cttgcccgct gctggacagg ggctcggctg	360
ttgggcaactg acaattccgt ggtgtgtcg gggaaatcat cgtccttcc ttggctgctc	420
gcctgtgttg ccacctggat tctgcgcggg acgtccttct gctacgtccc ttcggccctc	480
aatccagcgg accttcttc ccgcggcctg ctgccgctc tgcggcctct tccgcgtctt	540
cgcttctgcc ctacagacgag tcggatctcc ctttgggccg cctccccgc	589
<210> 42	
<211> 588	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-突变的土拨鼠肝炎调节元件	
<400> 42	
tcaacctctg gattacaaaa tttgtgaaag attgactggt attcttaact atgttgctcc	60
ttttacgcta tgtggatacg ctgctttaat gcctttgtat catgctattg ctcccgtat	120
ggctttcatt ttctctctct tgtataaatc ctggttgctg tctctttatg aggagttgtg	180
gcccgttgtc aggcaacgtg gcgtggtgtg cactgtgttt gctgacgcaa cccccactgg	240
ttggggcatt gccaccacct gtcagctcct ttccgggact ttcgcttcc cctcctat	300
tgccacggcg gaactcatcg ccgcctgcct tgcccgtgc tggacagggg ctcggtgtt	360
gggcaactgac aattccgtgg tgttgcggg gaaatcatcg tccttctctt ggctgctcgc	420
ctgtgttgcc acctggattc tgcgcgggac gtccttctgc tacgtcctt cggccctcaa	480

tccagcggac cttccttccc gcggcctgct gccggetctg cgccctcttc cgcgtcttcg	540
ccttcgccct cagacgagtc ggatctccct ttgggcccgc tccccgca	588
<210> 43	
<211> 755	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-突变的土拨鼠肝炎调节元件	
<400> 43	
ttcctgttaa tcaacctctg gattacaaaa tttgtgaaag attgactggt attcttaact	60
atgttgctcc ttttacgcta tgtggatacg ctgctttaat gcctttgtat catgctattg	120
cttcccgtat ggctttcatt ttctctctct tgtataaate ctggttgctg tctctttatg	180
aggagtgtg gcccgtgtc aggcaacgtg gcgtggtgtg cactgtgttt gctgacgcaa	240
ccccactgg ttggggcatt gccaccacct gtcagctctt ttccgggact ttcgctttcc	300
ccctccctat tgccacggcg gaactcatcg ccgctgctt tgcccgtgc tggacagggg	360
ctcggctgtt gggcactgac aattccgtgg tgttgcggg gaagctgacg tcctttccgc	420
ggctgctcgc ctgtgttgcc acctggattc tgcgcgggac gtccttctgc tacgtccctt	480
cggccctcaa tccagcggac cttccttccc gcggcctgct gccggetctg cgccctcttc	540
cgcctcttcg ccttcgccct cagacgagtc ggatctccct ttgggcccgc tccccgcca	600
tgtatctttt tcacctgtgc cttgtttttg cctgtgttcc gcgtcctact tttcaagcct	660
ccaagctgtg ccttgggicgg ctttggggca tggacataga tcctataaa gaatttggtt	720
catcttatca gttgttgaat tttcttcctt tggac	755
<210> 44	
<211> 12	
<212> DNA	
<213> 人工序列	
<220>	
<223> CAAX基序	
<400> 44	
tgtgtgataa tg	12
<210> 45	
<211> 810	
<212> DNA	
<213> 智人	
<400> 45	
ctgttctcat cacatcatat caagttata taccatcaat attgccacag atgttactta	60
gccttttaat atttctctaa tttagtgtat atgcaatgat agttctctga tttctgagat	120
tgagtttctc atgtgtaatg attatrtaga gtttctcttt catctgttca aatrtttgtc	180
tagtrtttatt ttttactgat ttgtaagact tcttrttata atctgcatat tacaattctc	240

ttactgggg tgttgcaaat attttctgtc attctatggc ctgacttttc ttaatggttt 300
 ttaattttta aaaataagtc ttaatattca tgcaatctaa ttaacaatct tttctttgtg 360
 gttaggactt tgagtcataa gaaatttttc tctacactga agtcatgatg gcatgcttct 420
 atattatfff ctaaagatt taaagttttg ctttctccat ttagacttat aattcactgg 480
 aatftttttg tgtgtatgg atgacatatg ggttcccttt tttttttac atataaatat 540
 atttccctgt ttttctaaaa aagaaaaaga tcatcatttt cccattgtaa aatgccatat 600
 ttttttcata ggtcacttac atatatcaat gggctctgtt ctgagctcta ctctatftta 660
 tcagcctcac tgtctatccc cacacatctc atgctttgct ctaaactttg atatfttagtg 720
 gaacattctt tcccattttg ttctacaaga atatftttgt tattgtcttt gggctttcta 780
 tatacatttt gaaatgaggt tgacaagtta 810

<210> 46

<211> 726

<212> DNA

<213> 乙型肝炎病毒

<400> 46

ataacaggcc tattgattgg aaagtttgtc aacgaattgt gggctttttg gggtttgctg 60
 ccccttttac gcaatgtgga tatcctgctt taatgccttt atatgcatgt atacaagcaa 120
 aacaggcttt tactttctcg ccaacttaca aggctttct cagtaaacag tatatgacct 180
 tttaccccg tgcctggcaa cggcctggc tgtgccaagt gtttgcctgac gcaaccccca 240
 ctggttgggg cttggccata ggccatcagc gcatgcctgg aacctttgtg tctcctctgc 300
 cgatccatac tgcggaactc cttagcctgt gttttgctcg cagcaggtct ggagcaaacc 360
 tcatcgggac cgacaattct gtcgtactct cccgcaagta tacatcgttt ccatggctgc 420
 taggctgtgc tgccaactgg atcctgcgcg ggacgtcctt tgtttacgtc ccgtcggcgc 480
 tgaatcccgc ggacgacccc tcccggggcc gcttggggct ctaccgccc cttctccgtc 540
 tgccgtaccg tccgaccacg gggcgcacct ctctttacgc ggactcccc tctgtgcctt 600
 ctcatctgcc ggaccgtgtg cacttcgctt cacctctgca cgtcgcctgg aggccacct 660
 gaacgcccac cgaacctgc ccaaggtctt gcataagagg actcttgac tttcagcaat 720
 gtcac 726

<210> 47

<211> 755

<212> DNA

<213> 人工序列

<220>

<223> 实验室制造-HepB衍生的增强子元件

<400> 47

ttctgtaaa caggcctatt gattggaaag tttgtcaacg aattgtgggt cttttggggt 60
 ttgctgcccc ttttacgcaa tgtggatata ctgctttaat gcctttatat gcatgtatac 120
 aagcaaaaca ggcttttact ttctcgccaa cttacaaggc ctttctcagt aaacagtata 180
 tgacccttta ccccgttgc cggcaacggc ctggtctgtg ccaagtgttt gctgacgcaa 240

ccccactgg ttggggcttg gccataggcc atcagecgc atcgcgtggaacc tttgtgtctc	300
ctctgccgat ccatactgcg gaactcctag ccgcttgttt tgctcgcagc tggactggag	360
caaacctcat cgggaccgac aattctgtcg tactctccc caagcactca ccgtttccgc	420
ggctgctcgc ctgtgttgcc acctggattc tgcgcgggac gtccttctgc tacgtccctt	480
cggccctcaa tccagcggac cttccttccc gcggcctgct gccggctctg cggcctcttc	540
cgctctctcg ccttcgccct cagacgagtc ggatctccct ttgggccgcc tccccgcca	600
tgtatctttt tcacctgtgc cttgtttttg cctgtgttcc gcgtcctact tttcaagcct	660
ccaagctgtg ccttgggcgg ctttggggca tggacataga tccctataaa gaatttggtt	720
catcttatca gttgttgaat tttcttctt tggac	755
<210> 48	
<211> 94	
<212> DNA	
<213> 智人	
<400> 48	
gctggagcct cggtagccgt tctcctgcc cgctgggect cccaacgggc cctcctcccc	60
tccttgcacc ggcccttctt ggtctttgaa taaa	94
<210> 49	
<211> 596	
<212> DNA	
<213> 土拨鼠肝炎病毒	
<400> 49	
attcgagcat cttaccgcca ttattccca tatttgttct gtttttcttg atttgggtat	60
acatttaaat gtaataaaa caaatggtg gggcaatcat ttacattttt aggatattgt	120
aattactagt tcaggtgtat tgccacaaga caacatggt aagaaacttt cccgttattt	180
acgctctggt cctgttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgatat	240
tcttaactat gttgctcctt ttacgctgtg tggatattgct gctttaatgc ctctgtatca	300
tgctattgct tcccgtacgg ctttcgtttt ctctccttg tataaatcct gtttgctgct	360
tctttatgag gagttgtggc ccgttgtccg tcaacgtggc gtggtgtgct ctgtgtttgc	420
tgacgcaacc cccactggct ggggcattgc caccactgt caactcctt ctgggacttt	480
cgctttcccc ctcccgatcg ccacggcaga actcatgcc gctgccttg cccgctgctg	540
gacaggggct aggttctggt gcaactgataa ttccgtggtg ttgtcgggga agggcc	596
<210> 50	
<211> 387	
<212> DNA	
<213> 穴兔 (<i>Oryctolagus cuniculus</i>)	
<400> 50	
tggttaataa aggaaattta ttttcattgc aatagtgtgt tggattttt tgtgtctctc	60
actcggaaga acatatggga gggcaaatca tttaaaacat cagaatgagt atttggttta	120
gagtttggca acatatgccc atatgctggc tgccatgaac aaagttggc tataaagagg	180

tcatacgtat atgaaacagc cccctgctgt ccattcctta ttccatagaa aagccttgac	240
ttgaggttag atttttttta tattttgitt tgtgttattt ttttcttta catccctaaa	300
attttcctta catgttttac tagccagatt tttcctcctc tectgactac tcccagtcac	360
agctgtccct cttctcttat ggagatc	387
<210> 51	
<211> 251	
<212> DNA	
<213> 牛(Bos taurus)	
<400> 51	
ttgccagcca tctgttgitt gcccctccc cgtgccttc ttgacctgg aaggtgccac	60
tcccactgtc ctttctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca	120
ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aatacaatag	180
caggcatgct ggggatgcgg tgggctctat gggtagccag gtgctgaaga attgacctgg	240
ttcctcctgg g	251
<210> 52	
<211> 251	
<212> DNA	
<213> 牛(Bos taurus)	
<400> 52	
ttgccagcca tctgttgitt gcccctccc cgtgccttc ttgacctgg aaggtgccac	60
tcccactgtc ctttctaat aaaatgagga aattgcatcg cattgtctga gtaggtgtca	120
ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg aagacaatag	180
caggcatgct ggggatgcgg tgggctctat gggtagccag gtgctgaaga attgacctgg	240
ttcctcctgg g	251
<210> 53	
<211> 225	
<212> DNA	
<213> 牛(Bos taurus)	
<400> 53	
ctgtgccttc tagttgccag ccactctgtt tttgcccctc ccccgctcct tccttgacct	60
tggaaggtgc cactcccact gtccttctct aataaaatga ggaaattgca tcgcattgtc	120
tgagtaggtg tcattctatt ctggggggtg ggggtggggca ggacagcaag ggggaggatt	180
gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatgg	225
<210> 54	
<211> 202	
<212> DNA	
<213> 智人	
<400> 54	
ctgcccgggt ggcacccctg tgaccctcc ccagtgcctc tectggcctt ggaagttgcc	60

actccagtgc ccaccagcct tgtcctaata aaattaagtt gcatcatttt gtctgactag	120
gtgtccttct ataatattat ggggtggagg ggggtggtat ggagcaaggg gcccaagttg	180
ggaagaaacc tgtagggcct gc	202
<210> 55	
<211> 3144	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-载体基因组构建体	
<400> 55	
cctgcaggca gctgcgcgct cgtcctctca ctgaggccgc ccgggcaaag cccgggcgtc	60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
actccatcac taggggttcc tgcggcaatt cagtcgataa ctataacggt cctaaggtag	180
cgatttaaat acgcgctctc ttaaggtagc cccgggacgc gtcaattgac tacaaccga	240
gtatctgcag agggccctgc gtatgagtgc aagtgggttt taggaccagg atgaggcggg	300
gtgggggtgc ctacctgacg accgaccccc acccactgga caagcaccca accccattc	360
cccaaattgc gcatccccta tcagagaggg ggaggggaaa caggatgcgg cgaggcgcgt	420
gcgcaactgcc agcttcagca ccgcggacag tgccttcgcc cccgcctggc ggcgcgcgcc	480
accgccgcct cagcaactgaa ggcgcgctga cgtcaactgc cggccccccg caaactcccc	540
ttcccggcca ccttggctgc gtccgcgccg ccgccggccc agccggaccg caccacgcga	600
ggcgcgagat agggggggcac gggcgcgacc atctgcgctg cggcgccggc gactcagcgc	660
tgccctcagtc tgcggtgggc agcggaggag tcgtgtcgtg cctgagagcg cagatgggca	720
aggagaagac ccacatcaac atcgttgtca tcggccactg ggactccgga aagtccacca	780
ccacgggcca cctcatctac aaatgcggag gtattgaaa aaggaccatt gagaagttcg	840
agaaggaggc ggctgagatg ggaagggat cttcaagta tgccctgggtg ctggacaagc	900
tgaaggcgga gcgtgagcgc ggcatcacca tcgacatctc cctctggaag ttcgagacca	960
ccaagtacta catcaccatc atcgtatgcc ccggccaccg cgacttcacg aagaacatga	1020
tcacgggtac atcccaggcg gactgcgcag tgctgatcgt ggccggcggc gtgggcgagt	1080
tcgaggcggg catctccaag aatgggcaga cgcgggagca tgccctgctg gcctacacgc	1140
tgggtgtgaa gcagctcacc gtgggcgtga acaaatgga ctccacagag ccggcctaca	1200
gcgagaagcg ctacgacgag atcgtcaagg aagtcagcgc ctacatcaag aagatcggct	1260
acaaccggc caccgtgcc tttgtgcca tctccggctg gcacggtgac aacatgctgg	1320
agccctcccc caacatgcc tggttcaagg gctggaagg ggagcgtaag gagggcaacg	1380
caagcggcgt gtccctgctg gaggccctgg acaccatcct gccccccacg cccccacgg	1440
acaagcccct gcgcctgcc ctgcaggacg tgtacaagat tggcggcatt ggcacggtgc	1500
ccgtggggccg ggtggagacc ggcacccctgc ggccgggcat ggtggtgacc tttgcgccg	1560
tgaacatcac cactgaggtg aagtcagtgg agatgcacca cgaggctctg agcgaagctc	1620
tgcccggcga caacgtcggc ttcaatgtga agaactgtc ggtgaaggac atccggcggg	1680
gcaacgtgtg tggggacagc aagtctgacc cgcgcgagga ggctgctcag ttcacctccc	1740

aggatcatcat cctgaaccac ccggggcaga ttagcgcccg ctactccccg gtcactgact	1800
gccacacagc ccacatcgcc tgcaagtttg cggagctgaa ggagaagatt gaccggcgct	1860
ctggcaagaa gctggaggac aaccccaagt ccctgaagtc tggagacgcg gccatcgtgg	1920
agatgggtgcc gggaaagccc atgtgtgtgg agagcttctc ccagtaccgg cctctcggcc	1980
gcttcgccgt gcgacatg aggcagacgg tggccgtagg cgatcatcaag aacgtggaga	2040
agaagagcgg cggcgccggc aaggtcacca agtcggcgca gaaggcgag aaggcgggca	2100
agtgaaatca acctctggat tacaaaattt gtgaaagatt gactggtatt cttactatg	2160
ttgctccttt tacgctatgt ggatacgtg ctttaatgcc tttgtatcat gctattgctt	2220
cccgtatggc tttcattttc tctccttgt ataaatcctg gttgctgtct ctttatgagg	2280
agttgtggcc cgttgtcagg caacgtggcg tgggtgtgac tgtgtttgct gacgcaacc	2340
ccactggttg gggcattgcc accacctgtc agctccttc cgggacttcc gctttcccc	2400
tccctattgc cacggcgga ctcacgccc cctgccttgc ccgctgctgg acaggggctc	2460
ggctgttggg cactgacaat tccgtggtgt tgcggggaa atcactgtcc tttccttggc	2520
tgctcgctg tgttgcacc tggattctgc gcgggacgtc cttctgctac gtccttcgg	2580
ccctcaatcc agcggacctt cctccccgg cctgctgccc ggctctgccc cctcttcgc	2640
gtcttcgct tcgccctcag acgagtcgga tctcccttg gcccgctcc ccgctgtgc	2700
cttctagttg ccagccatct gttgtttgcc cctccccgt gcttctctg acctggaag	2760
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgc atgtctgagta	2820
gggtgtcattc tattctgggg ggtgggggtg ggcaggacag caagggggag gattgggaag	2880
acaatagcag gcatgctggg gatgcggtg gctctatggc ttctgaggcg gaaagaacca	2940
gatcctctct taaggtagca tcgagattta aattaggat aacagggtaa tggcgcgggc	3000
cgcaggaacc cctagtgat gagttggcca ctccctctct gcgctcgc tcgctcactg	3060
aggccggggc accaaagtc gccgacgcc cgggcttgc ccggcgggc tcagtgagcg	3120
agcgagcgcg cagctgcctg cagg	3144
<210> 56	
<211> 3035	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-载体基因组构建体	
<400> 56	
gcgctcgc tcgctcactg aggccgccc ggcaaagccc gggcgtcggg cgaccttgg	60
tcgccggcc tcagtgagcg agcgagcgc cagagagga gtggccaact ccactactag	120
gggttccttg tagttaatga ttaaccgcc atgctactta tctacgtaag tgcaagtggg	180
ttttaggacc aggatgagc ggggtggggg tgctacctg acgaccgacc ccgaccact	240
ggacaagcac ccaaccccc tccccaaat tgcgcatccc ctatcagaga gggggagggg	300
aaacaggatg cggcgaggcg cgtgcgcaact gccagcttca gcaccgcca cagtgccttc	360
gccccgctt ggcgcgcg gccaccgccc cctcagcact gaaggcgcg tgacgtcact	420
cgccggtccc ccgcaaacct ccttccccg ccaccttggc cgcgtccgc ccgccgccc	480

cccagccgga	ccgcaccacg	cgaggcgcga	gatagggggg	cacgggcgcg	accatctgcg	540
ctgccggcgc	ggcgactcag	cgctgcctca	gtctgcggtg	ggcagcggag	gagtcgtgtc	600
gtgcctgaga	gcgcaggcca	ccatgggcaa	ggagaagacc	cacatcaaca	tcgtggtcat	660
cggccacgtg	gactccgga	agtccaccac	cacgggccac	ctcatctaca	aatgcggagg	720
tattgacaaa	aggaccattg	agaagttcga	gaaggaggcg	gctgagatgg	ggaagggatc	780
cttcaagtat	gcctgggtgc	tggacaagct	gaaggcggag	cgtgagcgcg	gcatcaccat	840
cgacatctcc	ctctggaagt	tcgagaccac	caagtactac	atcacatca	tcgatcccc	900
cggccaccgc	gacttcatca	agaacatgat	cacgggtaca	tcccaggcgg	actgcgcagt	960
gctgatcgtg	gcggcgggcg	tgggcgagtt	cgaggcgggc	atctccaaga	atgggcagac	1020
gcgggagcat	gccctgctgg	cctacacgct	gggtgtgaag	cagctcatcg	tgggcgtgaa	1080
caaaatggac	tccacagagc	cggcctacag	cgagaagcgc	tacgacgaga	tcgtcaagga	1140
agtcagcgc	tacatcaaga	agatcggeta	caaccggcc	accgtgcct	ttgtgccc	1200
ctccggctgg	cacggtgaca	acatgctgga	gcctcccc	aacatgccgt	ggttcaaggg	1260
ctggaaggtg	gagcgttaag	agggcaacgc	aagcggcgtg	tcctgctgg	aggcctgga	1320
caccatcctg	ccccccacgc	gccccacgga	caagcccctg	cgctgccgc	tgcaggacgt	1380
gtacaagatt	ggcggcattg	gcacggtgcc	cgtgggccgg	gtggagaccg	gcatcctgcg	1440
gccgggcatg	gtggtgacct	ttgcgccagt	gaacatcacc	actgaggtga	agtcagtgga	1500
gatgcaccac	gaggctctga	gcgaagctct	gcccggcgac	aacgtcggct	tcaatgtgaa	1560
gaacgtgtcg	gtgaaggaca	tccggcgggg	caacgtgtgt	ggggacagca	agtctgacct	1620
gccgcaggag	gctgctcagt	tcacctcca	ggtcatcctc	ctgaaccacc	cggggcagat	1680
tagcgccggc	tactccccgg	tcctcgactg	ccacacagcc	cacatgcct	gcaagtttgc	1740
ggagctgaag	gagaagattg	accggcgctc	tggcaagaag	ctggaggaca	acccaagtc	1800
cctgaagtct	ggagacgcgg	ccatcgtgga	gatggtgccg	ggaaagcca	tgtgtgtgga	1860
gagcttctcc	cagtaccgc	ctctcggccg	cttcgccgtg	cgcgacatga	ggcagacggt	1920
ggccgtaggc	gtcatcaaga	acgtggagaa	gaagagcggc	ggcgccggca	aggtcaccaa	1980
gtcggcgcag	aaggcgcaga	aggcgggcaa	gtgatcaacc	tctggattac	aaaatttgtg	2040
aaagattgac	tggatttctt	aactatggtg	ctccttttac	gctatgtgga	tacgctgctt	2100
taatgccttt	gtatcatgct	attgcttccc	gtatggcttt	cattttctcc	tccttgata	2160
aatcctgggt	gctgtctctt	tatgaggagt	tgtggcccgt	tgtcaggcaa	cgtggcgtgg	2220
tgtgcaactg	gtttgctgac	gcaaccccc	ctggttgggg	cattgccacc	acctgtcagc	2280
tcctttccgg	gactttcgt	ttccccctec	ctattgccac	ggcggaactc	atgccgcct	2340
gccttgcccc	ctgctggaca	ggggtctggc	tgttgggcac	tgacaattcc	gtggtgttgt	2400
cggggaaatc	atcgtccttt	ccttggctgc	tcgctgtgt	tgccacctgg	attctgcgcg	2460
ggacgtcctt	ctgctacgtc	ccttcggccc	tcaatccagc	ggaccttct	tcccgcggcc	2520
tgtgcccggc	tctgcggcct	cttccgcgtc	ttcgcttctg	ccctcagacg	agtcggatct	2580
ccctttgggc	cgcctccccg	cactgcccgg	gtggcatccc	tgtgacctcc	ccccagtgcc	2640
tctcctggcc	ctggaagttg	ccactccagt	gcccaccagc	cttgtcctaa	taaaattaag	2700
ttgcatcatt	ttgtctgact	aggtgtcctt	ctataatatt	atggggtgga	ggggggtggt	2760
atggagcaag	gggcccgaag	tgggaagaaa	cctgtagggc	ctgcgttacc	caggctggag	2820

tgcaagtggca catttctgct cactgcaacc tcctcctccc tgggttctac gtagataagt	2880
agcatggcgg gttaatcatt aactacaagg aaccctagtg gatggagttg gccactccct	2940
ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaa ggtcgcccga cgcccgggct	3000
ttgcccgggc ggcctcagtg agcgagcgag cgcg	3035
<210> 57	
<211> 3263	
<212> DNA	
<213> 人工序列	
<220>	
<223> 实验室制造-载体基因组构建体	
<400> 57	
gcgcgctcgc tcgctcactg aggccgcccc ggcaaagccc gggcgtcggg cgacctttgg	60
tcgcccggcc tcagtgagcg agcgagcgcg cagagagga gtggccaact ccatcactag	120
gggttccttg tagttaatga ttaaccggcc atgctactta tctacgtaag tgcaagtggg	180
ttttaggacc aggatgaggc ggggtggggg tgectactg acgaccgacc ccgaccact	240
ggacaagcac ccaaccccca ttccccaaat tgcgcatccc ctatcagaga gggggagggg	300
aaacaggatg cggcgagcg cgtgcgcaact gccagcttca gcaccgcgga cagtgccttc	360
gccccgcct ggcggcgcgc gccaccgccc cctcagcact gaaggcgcgc tgacgtcact	420
cgccgggtccc ccgaaaactc cccttcccgg ccaccttggg cgcgtcccg cgcccggcg	480
cccagccgga ccgaccacg cgaggcgca gatagggggg cacgggcgcg accatctgcg	540
ctgcggcgcc ggcgactcag cgctgcctca gtctgcggtg ggcagcggag gactcgtgtc	600
gtgcctgaga gcgcagagtc tgcggtgggc agcggaggag tcgtgtcgtg cctgagagcg	660
cagctgtgct cctgggcacc gcgcagtccg cccccgggc tcctggccag accacccta	720
ggacccccctg ccccaagtgc cagccacat gggcaaggag aagaccaca tcaacatcgt	780
ggatcatcggc cacgtggact ccggaagtc caccaccag ggccacctca tctacaaatg	840
cggaggtatt gacaaaagga ccattgagaa gttcgagaag gaggcggctg agatggggaa	900
gggatccttc aagtatgcct gggtgctgga caagctgaag gcggagcgtg agcgcgcat	960
caccatcgac atctccctct ggaagttcga gaccaccaag tactacatca ccatcatcga	1020
tgcccccggc caccgcgact tcatcaagaa catgatcag ggtacatccc aggcggactg	1080
cgcaagtctg atcgtggcgg cgggcgtggg cgagttcgag gcgggcatct ccaagaatgg	1140
gcagacgcgg gagcatgcc tgctggccta cacgtgggt gtgaagcagc tcatcgtggg	1200
cgtgaacaaa atggactcca cagagccggc ctacagcgag aagcgtacg acgagatcgt	1260
caaggaagtc agcgcctaca tcaagaagat cggtacaaac ccggccaccg tgccccttgt	1320
gcccattctc ggctggcacg gtgacaacat getggagccc tccccaca tgccgtggtt	1380
caagggtgag aaggtggagc gtaaggagg caacgcaagc ggcgtgtccc tgctggaggc	1440
cctggacacc atcctgcccc ccacgcgcc cacggacaag cccctgcgcc tgccgctgca	1500
ggacgtgtac aagattggcg gcattggcac ggtgcccgtg ggccgggtg agaccggcat	1560
cctgcggccg ggcatggtg tgacctttgc gccagtgaac atcaccactg aggtgaagtc	1620
agtggagatg caccacgagg ctctgagcga agctctgccc ggcgacaac tcggcttcaa	1680

tgtgaagaac	gtgtcgggta	aggacatccg	gcggggcaac	gtgtgtgggg	acagcaagtc	1740
tgacccgccg	caggaggctg	ctcagttcac	ctcccaggtc	atcatcctga	accacccggg	1800
gcagattagc	gccggctact	ccccggtcat	cgactgccac	acagcccaca	tcgacctgcaa	1860
gtttgcggag	ctgaaggaga	agattgaccg	gcgctctggc	aagaagctgg	aggacaaccc	1920
caagtccctg	aagtctggag	acgcggccat	cgtggagatg	gtgccgggaa	agcccatgtg	1980
tgtggagagc	ttctcccagt	acccgcctct	cggccgcttc	gccgtgcgcg	acatgaggca	2040
gacggtggcc	gtaggcgtca	tcaagaacgt	ggagaagaag	agcggcggcg	ccggcaaggt	2100
caccaagtcg	gcgcagaagg	cgcagaaggc	gggcaagtga	tcaacctctg	gattacaaaa	2160
tttgtgaaaag	attgactggg	attcttaact	atgttgctcc	ttttacgcta	tgtggatacg	2220
ctgctttaat	gcctttgtat	catgctattg	cttcccgtat	ggctttcatt	ttctcctcct	2280
tgtataaaatc	ctggttctg	tctctttatg	aggagtgtg	gcccgttgtc	aggcaacgtg	2340
gcgtgggtg	cactgtgttt	gctgacgcaa	ccccactgg	ttggggcatt	gccaccacct	2400
gtcagctcct	ttccgggact	ttegettcc	ccctccctat	tgccacggcg	gaactcatcg	2460
ccgcctgcct	tgcccgctgc	tggacagggg	ctcggtgtt	gggcactgac	aattccgtgg	2520
tgttgcggg	gaaatcatcg	tcctttcett	ggctgctcgc	ctgtgttgcc	acctggattc	2580
tgcgcgggac	gtccttctgc	tacgtccett	cggccctcaa	tccagcggac	cttcttccc	2640
gcggcctgct	gccggctctg	cggcctcttc	cgcgtcttcg	ccttcgccct	cagacgagtc	2700
ggatctccct	ttgggccgcc	tccccgcagc	tggagcctcg	gtagccgttc	ctcctgcccg	2760
ctgggcctcc	caacggggcc	tcctcccctc	cttgcaccgg	cccttctctg	tctttgaata	2820
aattcattgc	ctgcccgggt	ggcatccctg	tgaccctcc	ccagtgcctc	tcttgccct	2880
ggaagtggcc	actccagtgc	ccaccagcct	tgtcctaata	aaattaagtt	gcatcatttt	2940
gtctgactag	gtgtccttct	ataatattat	ggggtggagg	ggggtggat	ggagcaaggg	3000
gccaagtgtg	ggaagaaacc	tgtagggcct	gcgttaccca	ggctggagtg	cagtggcaca	3060
tttctgctca	ctgcaacctc	ctcctccctg	ggttctacgt	agataagtag	catggcgggt	3120
taatcattaa	ctacaaggaa	cccctagtga	tggagtggc	cactccctct	ctgcgcgctc	3180
gctcgctcac	tgaggccggg	cgaccaaagg	tcgcccagc	cccgggcttt	gcccggcgcg	3240
cctcagtgag	cgagcgagcg	cgc				3263
<210>	58					
<211>	4299					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	实验室制造-载体基因组构建体					
<400>	58					
gcgcgctcgc	tcgctcactg	aggccgcccc	ggcaaagccc	gggcgtcggg	cgacctttgg	60
tcgcccggcc	tcagtgagcg	agcgagcgcg	cagagaggga	gtggccaact	ccatcactag	120
gggttccttg	tagttaatga	ttaaccggcc	atgctactta	tctacgtact	ctggagacgc	180
gttacataac	ttacggtaaa	tggcccgcct	ggctgaccgc	ccaacgacc	ccgccattg	240
acgtcaataa	tgacgtatgt	tcccatagta	acgccaatag	ggactttcca	ttgacgtcaa	300

tgggtggagt atttacgta aactgcccac ttggcagtac atcaagtgta tcatatgcca	360
agtacgcccc ctattgacgt caatgacggt aaatggcccc cctggcatta tgcccagtac	420
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc	480
atggctcgagg tgagccccac gttctgcttc actctcccca tctccccccc ctccccaccc	540
ccaattttgt atttatttat tttttaatta ttttgtgcag cgatgggggc gggggggggg	600
ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg ggccgggcca ggccgagagg	660
tgcggcggca gccaatcaga gcggcgcgct ccgaaagttt cttttatgg cgaggcggcg	720
gcggcggcgg ccctataaaa agcgaagcgc gcggcgggcg ggagtcgctg cgcgctgcct	780
tcgccccgtg ccccgtccg ccgccgctc gcgcccccc ccccggtct gactgaccgc	840
gttactccca caggtgagcg ggcgggacgg cccttctct cgggctgta attagcgctt	900
ggtttaatga cgcttggtt cttttctgtg gctgcgtgaa agccttgagg ggctccggga	960
gggccctttg tgcgggggga gcggctcggg gggctgcgtgc gtgtgtgtgt gcgtggggag	1020
cgcccgctgc ggctccgcgc tgcccggcgg ctgtgagcgc tgcgggcgcg gcgcggggct	1080
ttgtgcgctc cgcagtgtgc gcgaggggag cgcggccggg ggcggtgcc cgcggtgcgg	1140
ggggggctgc gaggggaaca aaggctgcgt gcggggtgtg tgcgtggggg ggtgagcagg	1200
gggtgtgggc gcgtcggtcg ggctgcaacc ccccctgcac ccccctccc gagttgctga	1260
gcacggcccc gcttcgggtg cggggctccg tacggggcgt ggccgggggc tcgccgtgcc	1320
gggcgggggg tggcggcagg tgggggtgcc gggcggggcg gggccgcctc gggccgggga	1380
gggctcgggg gaggggcgcg gcggccccc gagcgcgggc ggctgtcgag gcgcggcgag	1440
ccgcagccat tgccttttat ggtaatcgtg cgagagggcg cagggacttc ctttgtccca	1500
aatctgtgcg gagccgaaat ctgggaggcg ccgccgacc ccctctagcg ggccgggggc	1560
gaagcgggtc ggcccccga ggaaggaaat gggcggggag ggccctcgtg cgtcgccgcg	1620
ccgccgtccc ctctccctc tccagcctcg gggctgtccg cggggggacg gctgccttcg	1680
ggggggacgg ggcagggcgg ggttcggctt ctggcgtgtg accggcggct ctagagcctc	1740
tgctaaccat gttcatgcct tcttctttt cctacagcgc caccatgggc aaggagaaga	1800
cccacatcaa catcgtggtc atcggccacg tggactccgg aaagtccacc accacgggcc	1860
acctcatcta caaatgcgga ggtattgaca aaaggacat tgagaagttc gagaaggagg	1920
cggctgagat ggggaaggga tccttcaagt atgcctgggt gctggacaag ctgaaggcgg	1980
agcgtgagcg cgcatcacc atcgacatct ccctctggaa gttcgagacc accaagtact	2040
acatcaccat catcgatgcc cccggccacc gcgacttcat caagaacatg atcacgggta	2100
catcccaggc ggactgcgca gtgctgatcg tggcggcggg cgtgggcgag ttcgaggcgg	2160
gcacttccaa gaatgggcag acgcgggagc atgcctgct ggccctacacg ctgggtgtga	2220
agcagctcat cgtgggcgtg aacaaaatgg actccacaga gccggcctac agcgagaagc	2280
gctacgacga gatcgtcaag gaagtcagcg cctacatcaa gaagatcggc tacaaccggg	2340
ccaccgtgcc ctttgtgccc atctccggtt ggcacggtga caacatgctg gagccctccc	2400
ccaacatgcc gtggttcaag ggctggaagg tggagcgtaa ggagggcaac gcaagcggcg	2460
tgtccctgct ggaggccctg gacaccatcc tgccccccac gcgccccacg gacaagcccc	2520
tgcgcctgcc gctgcaggac gtgtacaaga ttggcggcat tggcacggtg cccgtgggcc	2580
gggtggagac cggcatcctg cggccgggca tggtggtgac ctttgcgcca gtgaacatca	2640

ccactgaggt gaagtcagtg gagatgcacc acgaggctct gagcgaagct ctgcccggcg	2700
acaacgtcgg cttcaatgtg aagaacgtgt cggatgaagga catccggcgg ggcaacgtgt	2760
gtggggacag caagtctgac ccgccgcagg aggctgctca gttcacctcc caggtcatca	2820
tcctgaacca cccggggcag attagcgccg gctactcccc ggtcatcgac tgccacacag	2880
cccacatcgc ctgcaagttt gcggagctga aggagaagat tgaccggcgc tctggcaaga	2940
agctggagga caacccaag tccctgaagt ctggagacgc ggccatcgtg gagatggtgc	3000
cgggaaagcc catgtgtgtg gagagcttct cccagtaccc gcctctcggc cgcttcgccg	3060
tgcgcgacat gaggcagacg gtggccgtag gcgtcatcaa gaacgtggag aagaagagcg	3120
gcggcgcgg caaggtcacc aagtcggcgc agaaggcgca gaaggcggc aagtgatcaa	3180
cctctggatt acaaaaattg tgaagattg actggtatc ttaactatgt tgctcctttt	3240
acgctatgtg gatacgtgc tttaatgcct ttgtatcatg ctattgcttc ccgtatggct	3300
ttcattttct cctccttcta taaatcctgg ttgctgtctc tttatgagga gttgtggccc	3360
gttgtcaggc aacgtggcgt ggtgtgcaact gtgtttgctg acgcaacccc cactggttgg	3420
ggcattgcca ccacctgtca gctccttcc gggactttcg ctttccccct ccctattgcc	3480
acggcggaac tcacgcgcg ctgccttgcc cgtctctgga caggggctcg gctgttgggc	3540
actgacaatt ccgtgggtgt gtcggggaaa tcacgtcctt tccttggct gctcgcctgt	3600
gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc cctcaatcca	3660
gcggaccttc cttcccgcgg cctgctgccg gctctgcggc ctcttcgcg tcttcgctt	3720
cgccctcaga cgagtcggat ctcccttgg gccgcctccc cgcagctgga gcctcggtag	3780
ccgttcctcc tgcccgtgg gcctcccaac gggccctcct ccctccttg caccggcct	3840
tcctggctct tgaataaatt cattgcctgc ccgggtggca tcctgtgac ccctccccag	3900
tgctctcct ggccctggaa gttgccactc cagtgccac cagccttgct ctaataaaat	3960
taagttgcat cattttgtct gactaggtgt cttctataa tattatgggg tggagggggg	4020
tggtatggag caaggggccc aagttgggaa gaaacctgta gggcctgcgt taccaggct	4080
ggagtgcagt ggcacatttc tgctcactgc aacctctcc tcctgggtt ctacgtagat	4140
aagtagcatg gcgggttaat cattaactac aaggaacccc tagtgatgga gttggccact	4200
ccctctctgc gcgctcgtc gctcactgag gccgggcgac caaaggtcgc ccgacgccc	4260
ggctttgccc gggcggcctc agtgagcgag cgagcgcg	4299

<210> 59

<211> 735

<212> PRT

<213> 腺相关病毒2

<400> 59

Met	Ala	Ala	Asp	Gly	Tyr	Leu	Pro	Asp	Trp	Leu	Glu	Asp	Thr	Leu	Ser
1				5					10					15	
Glu	Gly	Ile	Arg	Gln	Trp	Trp	Lys	Leu	Lys	Pro	Gly	Pro	Pro	Pro	Pro
				20					25					30	
Lys	Pro	Ala	Glu	Arg	His	Lys	Asp	Asp	Ser	Arg	Gly	Leu	Val	Leu	Pro
				35					40					45	

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

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

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
 675 680 685
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
 690 695 700
 Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr
 705 710 715 720
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735
 <210> 60
 <211> 743
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体-AAV9变体
 <400> 60
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190

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

	500		505		510										
Gly	Arg	Asn	Ser	Leu	Met	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His	Lys
	515		520		525										
Glu	Gly	Glu	Asp	Arg	Phe	Phe	Pro	Leu	Ser	Gly	Ser	Leu	Ile	Phe	Gly
	530		535		540										
Lys	Gln	Gly	Thr	Gly	Arg	Asp	Asn	Val	Asp	Ala	Asp	Lys	Val	Met	Ile
545			550		555										560
Thr	Asn	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Ser
			565		570										575
Tyr	Gly	Gln	Val	Ala	Thr	Asn	His	Gln	Ser	Ala	Gln	Thr	Leu	Ala	Val
			580		585										590
Pro	Phe	Lys	Ala	Gln	Ala	Gln	Thr	Gly	Trp	Val	Gln	Asn	Gln	Gly	Ile
			595		600										605
Leu	Pro	Gly	Met	Val	Trp	Gln	Asp	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro
			610		615										620
Ile	Trp	Ala	Lys	Ile	Pro	His	Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro
625			630		635										640
Leu	Met	Gly	Gly	Phe	Gly	Met	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile
			645		650										655
Lys	Asn	Thr	Pro	Val	Pro	Ala	Asp	Pro	Pro	Thr	Ala	Phe	Asn	Lys	Asp
			660		665										670
Lys	Leu	Asn	Ser	Phe	Ile	Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val
			675		680										685
Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro
			690		695										700
Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Tyr	Lys	Ser	Asn	Asn	Val	Glu	Phe
705			710		715										720
Ala	Val	Asn	Thr	Glu	Gly	Val	Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr
			725		730										735
Arg	Tyr	Leu	Thr	Arg	Asn	Leu									
			740												
<210>	61														
<211>	7														
<212>	PRT														
<213>	人工序列														
<220>															
<223>	肽插入														
<400>	61														
Thr	Leu	Ala	Val	Pro	Phe	Lys									

1	5	
<210>	62	
<211>	7	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	肽插入	
<400>	62	
Lys Phe Pro Val Ala Leu Thr		
1	5	
<210>	63	
<211>	168	
<212>	DNA	
<213>	腺相关病毒	
<400>	63	
tacgtagata agtagcatgg cgggttaatc attaactaca aggaaccct agtgatggag		60
ttggcactc cctctctgcg cgctcgtcgc ctactgagg ccgggcgacc aaaggtcgcc		120
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgc		168
<210>	64	
<211>	851	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	实验室制造-eSYN启动子多核苷酸	
<400>	64	
gacattgatt attgactagt tattaatagt aatcaattac ggggtcatta gttcatagcc		60
catatatgga gttccgcgtt acataactta cggtaaattgg cccgcctggc tgaccgccc		120
acgacccccg cccattgacg tcaataatga cgtatgttc catagtaacg ccaataggga		180
ctttccattg acgtcaatgg gtggactatt tacggtaaac tgcccacttg gcagtacatc		240
aagtgtatca tatgccaagt acgccccta ttgacgtcaa tgacggtaa tggcccgcct		300
ggcattatgc ccagtacatg accttatggg actttctac ttggcagtac atctacgtat		360
tagtcatcgc tattaccatg gctgcagagg gcctgcgta tgagtgaag tgggttttag		420
gaccaggatg aggcggggtg ggggtgccta cctgacgacc gacccgacc cactggacaa		480
gcaccaacc cccattcccc aaattgcgca tcccctatca gagaggggga ggggaaacag		540
gatgcggcga ggcgcgtcgc gactgccagc ttcagaccg cggacagtgc cttegcccc		600
gcctggcggc gcgcgccacc gccgcctcag cactgaaggc gcgctgacgt cactcgccgg		660
tccccgcaa actccccttc ccggccacct tggtcgcgtc cgcgccgccc ccggcccagc		720
cggaccgcac cacgcgaggc gcgagatagg ggggcacggg cgcgaccatc tgcgctgcgg		780
cgccggcgac tcagcgtgc ctcaagtctgc ggtgggcagc ggaggagtgc tgtcgtgcct		840

gagagcgcag g	851
<210> 65	
<211> 3014	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工载体基因组	
<400> 65	
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc	60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
actccatcac taggggttcc tgcggcaatt cagtcgataa ctataacggt cctaaggacg	180
cgtagtgcaa gtgggtttta ggaccaggat gaggcggggt gggggtgcct acctgacgac	240
cgacccccgac ccaactggaca agcaccacaac cccattccc caaattgcgc atcccctatc	300
agagagggggg aggggaaaca ggatgcggcg aggcgcgtgc gcaactgccag cttcagcacc	360
gcggacagtg ccttcgcccc cgcttggcgg cgcgcgccac cgccgcctca gcaactgaagg	420
cgcgctgacg tcaactcgccg gtcccccgca aactcccctt cccggccacc ttggtcgcgt	480
ccgcgccgcc gccggcccag ccggaccgca ccacgcgagg cgcgagatag gggggcacgg	540
gcgcgaccat ctgcgctgcg gcgccggcga ctcagcgctg cctcagtctg cggtgggcag	600
cggaggagtc gtgtcgtgcc tgagagcgca gatgggcaag gagaagacc acatcaacat	660
cgtggtcatc ggccacgtgg actccgaaa gtccaccacc acgggccacc tcatctaaa	720
atgcggagggt attgacaaaa ggaccattga gaagtctgag aaggaggcgg ctgagatggg	780
gaagggatcc ttcaagtatg cctgggtgct ggacaagctg aaggcggagc gtgagcgcgg	840
catcaccatc gacatctccc tctggaagtt cgagaccacc aagtactaca tcaccatcat	900
cgatgcccc ggccaccgcg acttcatcaa gaacatgac acgggtacat cccagcggga	960
ctgcgcagtg ctgatcgtgg cggcgggctt gggcgagttc gaggcgggca tctccaagaa	1020
tgggcagacg cgggagcatg ccctgctggc ctacacgctg ggtgtgaagc agctcatcgt	1080
ggcgctgaac aaaatggact ccacagagcc ggcctacagc gagaagcgt acgacgagat	1140
cgtcaaggaa gtcagcgcct acatcaagaa gatcggctac aaccggcca ccgtgccctt	1200
tgtgccccatc tccgctggc acggtgaaa catgctggag ccctccccca acatgccgtg	1260
gttcaagggc tggaaagtgg agcgtaaagga gggcaacgca agcggcgtgt ccctgctgga	1320
ggccctggac accatcctgc cccccacgcg cccacggac aagcccctgc gcctgccgct	1380
gcaggacgtg tacaagattg gcggcattgg cacggtgcc gtgggcccggg tggagaccgg	1440
catcctgcgg ccgggcatgg tggtagcctt tgcgccagt aacatcaca ctgaggtgaa	1500
gtcagtggag atgcaccacg aggetctgag cgaagctctg cccggcgaca acgtcggctt	1560
caatgtgaag aacgtgtcgg tgaaggacat ccggcggggc aacgtgtgtg gggacagcaa	1620
gtctgaccgg ccgcaggagg ctgctcagtt cacctcccag gtcacatcc tgaaccacc	1680
ggggcagatt agcgcggct actccccggt catcgactgc cacacagccc acatcgctg	1740
caagtttgcg gagctgaagg agaagattga ccggcgtct ggcaagaagc tggaggacaa	1800
ccccaagtcc ctgaagtctg gagacgcggc catcgtggag atggtgccgg gaaagccat	1860

gtgtgtggag agcttctccc agtaccgcc tctcgccgc ttcgccgtgc gcgacatgag	1920
gcagacgggtg gccgtaggcg tcatcaagaa cgtggagaag aagagcggcg gcgccggcaa	1980
ggtcaccaag tcggcgcaga aggcgcagaa ggccggcaag tgatcaacct ctggattaca	2040
aaattttgtga aagattgact ggtattctta actatgttgc tccttttacg ctatgtggat	2100
acgctgcttt aatgcctttg tatcatgcta ttgcttcccg tatggctttc attttctcct	2160
ccttgataaa atcctggttg ctgtctcttt atgaggagtt gtggcccgtt gtcaggcaac	2220
gtggcgtggt gtgcactgtg tttgctgacg caacccccac tggttggggc attgccacca	2280
cctgtcagct cctttccggg actttcgctt tccccctcc tattgccacg gcggaactca	2340
tcgccgcctg ccttgcccgc tgctggacag gggtcggct gttgggcact gacaattccg	2400
tggtgttgtc ggggaaatca tcgtcctttc cttggctgct cgctgtgtt gccacctgga	2460
ttctgcgcgg gacgtccttc tgetacgtc cttcgccct caatccagcg gaccttcctt	2520
cccgcggcct gctgccggt ctgcggctc ttccgcgtct tcgccttcgc cctcagacga	2580
gtcggatctc cctttgggcc gcctccccgc ctgtgccttc tagttgccag ccatctgttg	2640
tttgcccctc ccccgtgct tccttgacc tggaaggtgc cactcccact gtcttttct	2700
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctggggggtg	2760
gggtggggca ggacagcaag ggggaggatt gggaagaaa tagcaggcat gctggggatg	2820
cggtgggctc tatggattta aattaggat aacagggtaa tggcgcgggc cgcaggaacc	2880
cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg	2940
accaaaggtc gcccagacc cgggctttgc ccggcggcc tcagtgagcg agcgagcgcg	3000
cagctgcctg cagg	3014
<210> 66	
<211> 3049	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工载体基因组	
<400> 66	
gcgcgctcgc tcgctcactg aggccgccc ggcaaagccc gggcgtcggg cgacctttgg	60
tcgcccgcc tcagtgagcg agcgagcgc cagagagga gtggccaact ccatcactag	120
gggttccttg tagttaatga ttaaccgcc atgctactta tctacgtact ctggagacgc	180
gtagtgcaag tgggttttag gaccaggatg aggcggggtg ggggtgccta cctgacgacc	240
gaccccgacc cactggacaa gcacccaacc cccattccc aaattgcgca tcccctatca	300
gagaggggga ggggaaacag gatgcggcga ggccgctgcg cactgccagc ttcagcaccg	360
cggacagtgc cttcgcccc gcctggcggc gcgcgccacc gccgctcag cactgaaggc	420
gcgctgacgt cactcgccgg tccccgcaa actcccctc ccggccacct tggtcgcgtc	480
cgcgccgcc ccggcccagc cggaccgcac cacgcgagc gcgagatagg ggggcacggg	540
cgcgaccatc tgcgctcgg cgccggcgac tcagcgtgc ctcagtctgc ggtgggcagc	600
ggaggagtgc tgtcgtgcct gagagcgcag gccacatgg gcaaggagaa gaccacatc	660
aacatcgtgg tcatcgcca cgtggactcc gaaagtcca ccaccacggg ccacctatc	720

tacaaatgcg	gaggtattga	caaaaggacc	attgagaagt	tcgagaagga	ggcggctgag	780
atggggaagg	gatccttcaa	gtatgcctgg	gtgctggaca	agctgaaggc	ggagcgtgag	840
cgcgcatca	ccatcgacat	ctccctctgg	aagtctgaga	ccaccaagta	ctacatcacc	900
atcatcgatg	ccccggcca	ccgcgacttc	atcaagaaca	tgatcacggg	tacatcccag	960
gcggactgcg	cagtgctgat	cgtggcggcg	ggcgtgggcg	agttcgaggc	gggcatctcc	1020
aagaatgggc	agacgcggga	gcatgccctg	ctggcctaca	cgctgggtgt	gaagcagctc	1080
atcgtgggcg	tgaacaaaat	ggactccaca	gagccggcct	acagcgagaa	gcgctacgac	1140
gagatcgta	aggaagtcag	cgctacatc	aagaagatcg	gctacaacc	ggccaccgtg	1200
ccctttgtgc	ccatctccgg	ctggcacggt	gacaacatgc	tggagccctc	ccccaacatg	1260
ccgtggttca	agggtctgaa	ggtggagcgt	aaggagggca	acgcaagcgg	cgtgtccctg	1320
ctggaggccc	tggacacat	cctgcccc	acgcgcccc	cggacaagcc	cctgcgcctg	1380
ccgctgcagg	acgtgtacaa	gattggcggc	attggcacgg	tgcccgtggg	ccgggtggag	1440
accggcatcc	tgcggccggg	catggtggtg	acctttgcgc	cagtgaacat	caccactgag	1500
gtgaagtcag	tggagatgca	ccacgaggct	ctgagcgaag	ctctgcccgg	cgacaacgtc	1560
ggcttcaatg	tgaagaacgt	gtcgggtgaag	gacatccggc	ggggcaacgt	gtgtggggac	1620
agcaagtctg	acccgccgca	ggaggctgct	cagttcacct	cccaggtcat	catcctgaac	1680
caccgggggc	agattagcgc	cggctactcc	ccggtcatcg	actgccacac	agccccatc	1740
gcctgcaagt	ttgcggagct	gaaggagaag	attgaccggc	gctctggcaa	gaagctggag	1800
gacaacccca	agtccctgaa	gtctggagac	gcggccatcg	tggagatggt	gccgggaaag	1860
cccatgtgtg	tggagagctt	ctcccagtac	ccgcctctcg	gccgcttcgc	cgtgcgcgac	1920
atgaggcaga	cggtggccgt	aggcgtcatc	aagaacgtgg	agaagaagag	cggcggcgcc	1980
ggcaaggcca	ccaagtcggc	gcagaaggcg	cagaaggcgg	gcaagtgatc	aacctctgga	2040
ttacaaaatt	tgtgaaagat	tgactggtat	tcttaactat	gttgctcctt	ttacgctatg	2100
tggatacgct	gctttaatgc	ctttgtatca	tgctattgct	tcccgtatgg	ctttcatttt	2160
ctcctccttg	tataaatcct	ggttgctgtc	tctttatgag	gagttgtggc	ccgttgctcag	2220
gcaacgtggc	gtggtgtgca	ctgtgtttgc	tgacgcaacc	cccactggtt	ggggcattgc	2280
caccacctgt	cagctccttt	ccgggacttt	cgctttcccc	ctccctattg	ccacggcgga	2340
actcatcgcc	gcctgccttg	cccgtctgtg	gacaggggct	cggctgttgg	gactgacaaa	2400
ttccgtggtg	ttgtcgggga	aatcatcgtc	ctttccttgg	ctgctcgcct	gtgttgccac	2460
ctggattctg	cgcgggacgt	ccttctgcta	cgctccttcg	gccctcaatc	cagcggacct	2520
tccttcccgc	ggcctgctgc	cggtctctgc	gcctcttcgc	cgtcttcgcc	ttcgcctca	2580
gacgagtcgg	atctcccttt	gggccgctc	cccgcactgc	ccgggtggca	tcctgtgac	2640
ccctccccag	tgctctcct	ggcctggaa	gttgccactc	cagtgccac	cagccttgct	2700
ctaataaaaat	taagttgcat	cattttgtct	gactaggtgt	ccttctataa	tattatgggg	2760
tggagggggg	tggatggag	caaggggccc	aagttgggaa	gaaacctgta	ggcctgcgt	2820
taccaggct	ggagtgcagt	ggcacatttc	tgtcactgc	aacctctcc	tcctgggtt	2880
ctacgtagat	aagtagcatg	gcgggttaat	cattaactac	aaggaacccc	tagtgatgga	2940
gttggccact	ccctctctgc	gcgctcgtc	gctcactgag	gccgggcgac	caaaggtcgc	3000
ccgacgcccg	ggctttgccc	ggcgccctc	agtgagcgag	cgagcgcgc		3049

<210>	67								
<211>	3277								
<212>	DNA								
<213>	人工序列								
<220>									
<223>	人工载体基因组								
<400>	67								
gcgcgctcgc	tcgctcactg	aggccgcccc	ggcaaagccc	gggcgtcggg	cgacctttgg				60
tcgccccggc	tcagtgagcg	agcgagcgcg	cagagagggg	gtggccaact	ccatcactag				120
gggttccttg	tagttaatga	ttaaccgcc	atgctactta	tctacgtact	ctggagacgc				180
gtagtgcaag	tgggttttag	gaccaggatg	aggcggggtg	ggggtgcta	cctgacgacc				240
gaccccgacc	cactggacaa	gcacccaacc	ccattcccc	aaattgcgca	tcccctatca				300
gagaggggga	ggggaaacag	gatgcggcga	ggcgcgtgcg	cactgccagc	ttcagcaccg				360
cggacagtgc	cttcgcccc	gcctggcggc	gcgcgccacc	gccgcctcag	cactgaaggc				420
gcgctgacgt	cactcgccgg	tccccgcaa	actccccttc	ccggccacct	tggtcgcgtc				480
cgcgccgccc	ccggcccagc	cggaccgcac	cacgcgaggc	gcgagatagg	ggggcacggg				540
cgcgaccatc	tgcgctgcgg	cgccggcgac	tcagcgtgct	ctcagtctgc	ggtgggcagc				600
ggaggagtgc	tgtcgtgcct	gagagcgag	agtctgcggt	gggcagcggg	ggagtctgtt				660
cgtgcctgag	agcgcagctg	tgtcctggg	caccgcgag	tccgccccg	cggctcctgg				720
ccagaccacc	cctaggacc	cctgccccaa	gtcgcagcca	ccatgggcaa	ggagaagacc				780
cacatcaaca	tcgtggtcat	cggccacgtg	gactccggaa	agtccaccac	cacgggccac				840
ctcatctaca	aatgcggagg	tattgacaaa	aggaccattg	agaagtctga	gaaggaggcg				900
gctgagatgg	ggaagggatc	cttcaagtat	gcctgggtgc	tggacaagct	gaaggcggag				960
cgtgagcgcg	gcatcaccat	cgacatctcc	ctctggaagt	tcgagaccac	caagtactac				1020
atcaccatca	tcgatgcccc	cggccaccgc	gacttcatca	agaacatgat	cacgggtaca				1080
tcccaggcgg	actgcgcagt	gctgatcgtg	gcggcgggcg	tggcgagtt	cgaggcgggc				1140
atctccaaga	atgggcagac	gcgggagcat	gccctgctgg	cctacacgct	gggtgtgaag				1200
cagctcatcg	tggcgtgaa	caaatggac	tccacagagc	cggcctacag	cgagaagcgc				1260
tacgacgaga	tcgtcaagga	agtcagcgcc	tacatcaaga	agatcggtta	caaccgggcc				1320
accgtgccct	ttgtgccc	ctccggctgg	cacggtgaca	acatgctgga	gccctcccc				1380
aacatgccgt	ggttcaagg	ctggaaggtg	gagcgtaagg	agggaacgc	aagcggcgtg				1440
tccctgctgg	aggccctgga	caccatctg	ccccccacgc	gccccacgga	caagcccctg				1500
cgcttcccgc	tgcaggacgt	gtacaagatt	ggcggcattg	gcacggtgcc	cgtgggccgg				1560
gtggagaccg	gcatcctgcg	gccgggcatg	gtggtgacct	ttgcgccagt	gaacatcacc				1620
actgaggtga	agtcagtgga	gatgcaccac	gaggctctga	gcgaagctct	gcccggcgac				1680
aacgtcggct	tcaatgtgaa	gaacgtgtcg	gtgaaggaca	tccggcgggg	caacgtgtgt				1740
ggggacagca	agtctgacct	gccgcaggag	gctgctcagt	tcacctcca	ggtcatcacc				1800
ctgaaccacc	cggggcagat	tagcggcggc	tactccccgg	tcacgactg	ccacacagcc				1860
cacatcgctc	gcaagtttgc	ggagctgaag	gagaagattg	accggcgctc	tggcaagaag				1920

ctggaggaca accccaagtc cctgaagtct ggagacgcgg ccatcgtgga gatggtgccg	1980
ggaaagccca tgtgtgtgga gagcttctcc cagtaccgcg ctctcggccg cttcgccgtg	2040
cgcgacatga ggcagacggt ggccgtaggc gtcatcaaga acgtggagaa gaagagcggc	2100
ggcgccggca aggtcaccaa gtcggcgag aaggcgcaga aggcgggcaa gtgatcaacc	2160
tctggattac aaaatttgtg aaagattgac tggatttctt aactatgttg ctccttttac	2220
gctatgtgga tacgctgctt taatgccttt gtatcatgct attgcttccc gtatggcttt	2280
cattttctcc tccttgata aatcctggtt gctgtctctt tatgaggagt tgtggcccg	2340
tgtcaggcaa cgtggcgtgg tgtgcaactgt gtttctgac gcaaccccca ctggttgggg	2400
cattgccacc acctgtcagc tcctttccgg gactttcget ttccccctcc ctattgccac	2460
ggcggaactc atcgccgctt gccttgcctg ctgctggaca ggggctcggc tgttggggcac	2520
tgacaattcc gtggtgttgt cggggaaatc atcgtccttt ccttggctgc tcgctgtgt	2580
tgccacctgg attctgcgcg ggacgtcctt ctgctacgct ccttcggccc tcaatccagc	2640
ggaccttctt tcccgcggcc tgetgcccgc tctgcggcct ctcccgctc ttcgcttgc	2700
ccctcagacg agtcggatct ccctttgggc cgcctccccg cagctggagc ctcggtagcc	2760
gttctctctg cccgctgggc ctcccaacgg gccctctctc cctccttga ccggcccttc	2820
ctggtctttg aataaattca ttgcctgcc gggtggcatc cctgtgacc ctccccagt	2880
cctctctctg ccctggaagt tgccactcca gtgccacca gccttgtcct aataaatta	2940
agttgcatca ttttgtctga ctaggtgtcc ttctataata ttatggggtg gaggggggtg	3000
gtatggagca aggggcccga gttggaaga aacctgtagg gcctgcgta cccaggctgg	3060
agtgcagtgg cacatttctg ctactgcaa cctcctctc cctgggttct acgtagataa	3120
gtagcatggc ggggttaatca ttaactaaa ggaacccta gtgatggagt tggccactcc	3180
ctctctgcgc gctcgtcgc tactgaggc cggcgacca aaggtcgcc gacgcccggg	3240
ctttgcccgg gcggcctcag tgagcgagcg agcgcgc	3277
<210> 68	
<211> 3498	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工载体基因组	
<400> 68	
gcgcgctcgc tcgctcactg aggccgccc ggcaaagccc gggcgtcggg cgacctttgg	60
tcgcccggcc tcagtgagcg agcagcgcg cagagagga gtggccaact ccatcactag	120
gggttccttg tagttaatga ttaaccgccc atgctactta tctacgtact ctggagacgc	180
gttacataac ttacggtaaa tggcccgcct ggetgaccgc ccaacgacc ccgcccattg	240
acgtcaatag taacgccaat agggacttcc cattgacgct aatgggtgga gtatttacgg	300
taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc ccctattgac	360
gtcaatgacg gtaaatggcc cgctggcat tgtgccagc acatgacctt atgggacttt	420
cctacttggc agtacatcta cgtattagtc atcgtatta ccatggtcga ggtgagcccc	480
acgttctgct tcaactctcc catctcccc cctccccac cccaatttt gtattttatt	540

atTTTTtaat	tatTTTTgtgc	agcgatgggg	gcgggggggg	ggggggggcg	cgcgccaggc	600
ggggcggggc	ggggcgaggg	gcggggcggg	gcgaggcgga	gaggtgcggc	ggcagccaat	660
cagagcggcg	cgctccgaaa	gtttcctttt	atggcgaggg	ggcggcggcg	gcggccctat	720
aaaaagcgaa	gcgcgcggcg	ggcgggagtc	gctgcgcgct	gccttcgccc	cgtgccccgc	780
tccgccgccg	cctcgcgccg	cccggcccgg	ctctgactga	ccgcgttact	cccacagggtg	840
agcgggcggg	acggcccttc	tcctccgggc	tgtaattagc	tgagcaagag	gtaagggttt	900
aagggatggt	tggttggtgg	ggtattaatg	ttaattacc	tggagcacct	gcctgaaatc	960
actTTTTttc	aggttgggcc	accatgggca	aggagaagac	ccacatcaac	atcgtggtca	1020
tcggccacgt	ggactccgga	aagtccacca	ccacgggcca	cctcatctac	aaatgcggag	1080
gtattgacaa	aaggaccatt	gagaagtctg	agaaggaggc	ggctgagatg	gggaagggat	1140
ccttcaagta	tgctgggtg	ctggacaagc	tgaaggcgga	gcgtgagcgc	ggcatcacca	1200
tcgacatctc	cctctggaag	ttcgagacca	ccaagtacta	catcaccatc	atcgatgccc	1260
ccggccaccg	cgacttcate	aagaacatga	tcacgggtac	atcccaggcg	gactgcgcag	1320
tgctgatcgt	ggcggcgggc	gtgggcgagt	tcgaggcggg	catctccaag	aatgggcaga	1380
cgcgggagca	tgccctgctg	gcctacacgc	tggtgtgtaa	gcagctcatc	gtgggcgtga	1440
acaaaatgga	ctccacagag	ccggcctaca	gcgagaagcg	ctacgacgag	atcgtcaagg	1500
aagtcagcgc	ctacatcaag	aagatcggct	acaaccggc	caccgtgccc	tttgtgccc	1560
tctccggctg	gcacggtgac	aacatgctgg	agccctcccc	caacatgccg	tggttcaagg	1620
gctggaaggt	ggagcgtaag	gagggcaacg	caagcggcgt	gtccctgctg	gaggccctgg	1680
acaccatcct	gccccccacg	cgccccacgg	acaagcccct	gcgcctgccg	ctgcaggacg	1740
tgtacaagat	tggcggcatt	ggcacggtgc	ccgtgggccg	ggtggagacc	ggcatcctgc	1800
ggccgggcat	ggtggtgacc	tttgcgccag	tgaacatcac	cactgaggtg	aagtcagtgg	1860
agatgcacca	cgaggctctg	agcgaagctc	tgcccggcga	caacgtcggc	ttcaatgtga	1920
agaacgtgtc	ggtgaaggac	atccggcggg	gcaacgtgtg	tggggacagc	aagtctgacc	1980
cgccgcagga	ggctgctcag	ttcacctccc	aggtcatcat	cctgaaccac	ccggggcaga	2040
ttagcggcgg	ctactccccg	gtcatcgact	gccacacagc	ccacatcgcc	tgcaagtttg	2100
cggagctgaa	ggagaagatt	gaccggcgct	ctggcaagaa	gctggaggac	aacccaagt	2160
ccctgaagtc	tggagacgcg	gccatcgtgg	agatggtgcc	gggaaagccc	atgtgtgtgg	2220
agagcttctc	ccagtaccgg	cctctcggcc	gcttcgccgt	gcgcgacatg	aggcagacgg	2280
tggccgtagg	cgcatcaag	aacgtggaga	agaagagcgg	cggcgccggc	aaggtcacca	2340
agtcggcgca	gaaggcgcag	aaggcgggca	agtgatcaac	ctctggatta	caaaatttgt	2400
gaaagattga	ctggtattct	taactatggt	gctcctttta	cgctatgtgg	atacgctgct	2460
ttaatgcctt	tgtatcatgc	tattgcttcc	cgtatggctt	tcattttctc	ctccttgtat	2520
aaatcctggt	tgctgtctct	ttatgaggag	ttgtggcccc	ttgtcaggca	acgtggcgtg	2580
gtgtgcaactg	tgtttgcgta	cgcaaccccc	actggttggg	gcattgccac	cacctgtcag	2640
ctcctttccg	ggactttcgc	tttccccctc	ctattgcca	cggcggaact	catcgccgcc	2700
tgccttgccc	gctgctggac	aggggctcgg	ctgttgggca	ctgacaattc	cgtggtgttg	2760
tcggggaaat	catcgtcctt	tccttgctg	ctcgcctgtg	ttgccacctg	gattctgcgc	2820
gggacgtcct	tctgctacgt	cccttcggcc	ctcaatccag	cggaccttcc	ttcccgcggc	2880

ctgctgccgg	ctctgccc	tcttccgct	cttcgcttc	gccctcagac	gagtcggatc	2940
tccctttggg	ccgcctcccc	gcagctggag	cctcggtagc	cgttctctct	gcccgtggg	3000
cctcccaacg	ggccctcctc	ccctccttgc	accggccctt	cctggtcttt	gaataaatc	3060
attgcctgcc	cgggtggcat	ccctgtgacc	cctccccagt	gcctctctctg	gccctggaag	3120
ttgccactcc	agtgccacc	agccttgtcc	taataaaatt	aagttgcatc	atthttgtctg	3180
actaggtgtc	cttctataat	attatgggggt	ggaggggggt	ggtatggagc	aaggggcca	3240
agttgggaag	aaacctgtag	ggcctgcgtt	accaggctg	gagtgagtg	gcacatttct	3300
gctcactgca	acctctctct	ccctgggttc	tacgtagata	agtagcatgg	cgggttaatc	3360
attaactaca	aggaaccct	agtgatggag	ttggcactc	cctctctgcg	cgctcgctcg	3420
ctcactgagg	ccgggcgacc	aaaggtcgcc	cgacgcccgg	gctttgcccg	ggcggcctca	3480
gtgagcgagc	gagcgcgc					3498

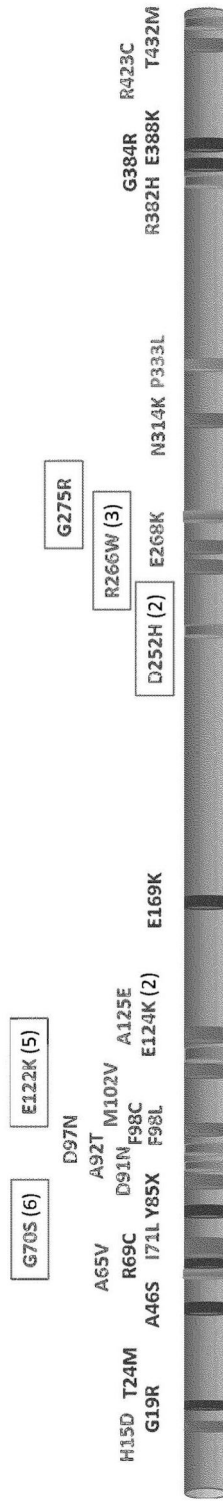


图 1

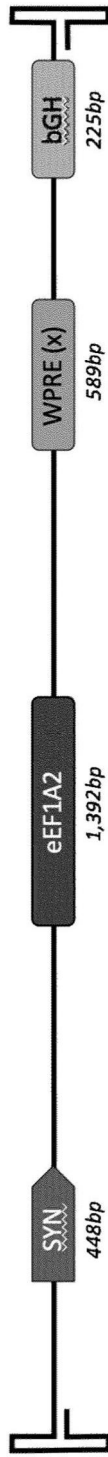


图2

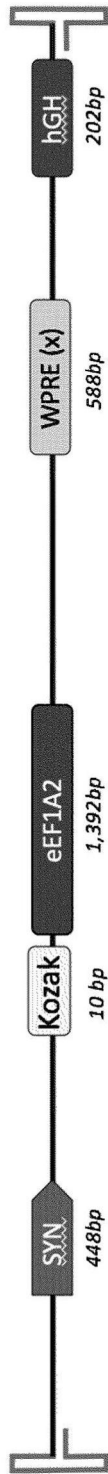


图3

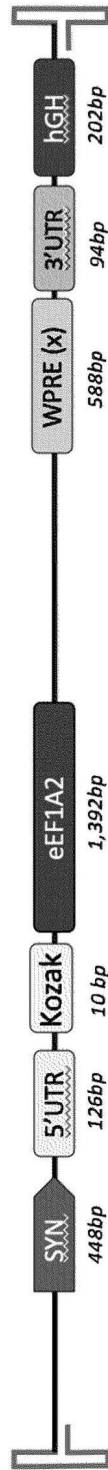


图4



图5



图6

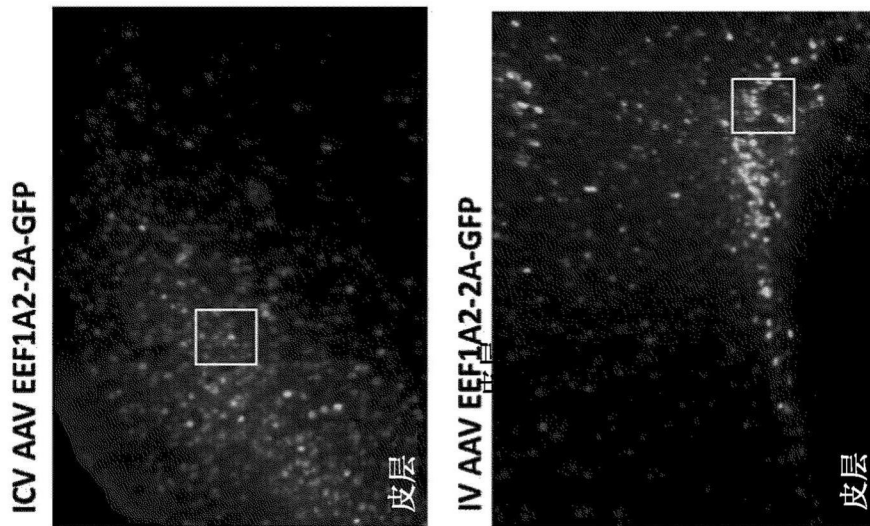
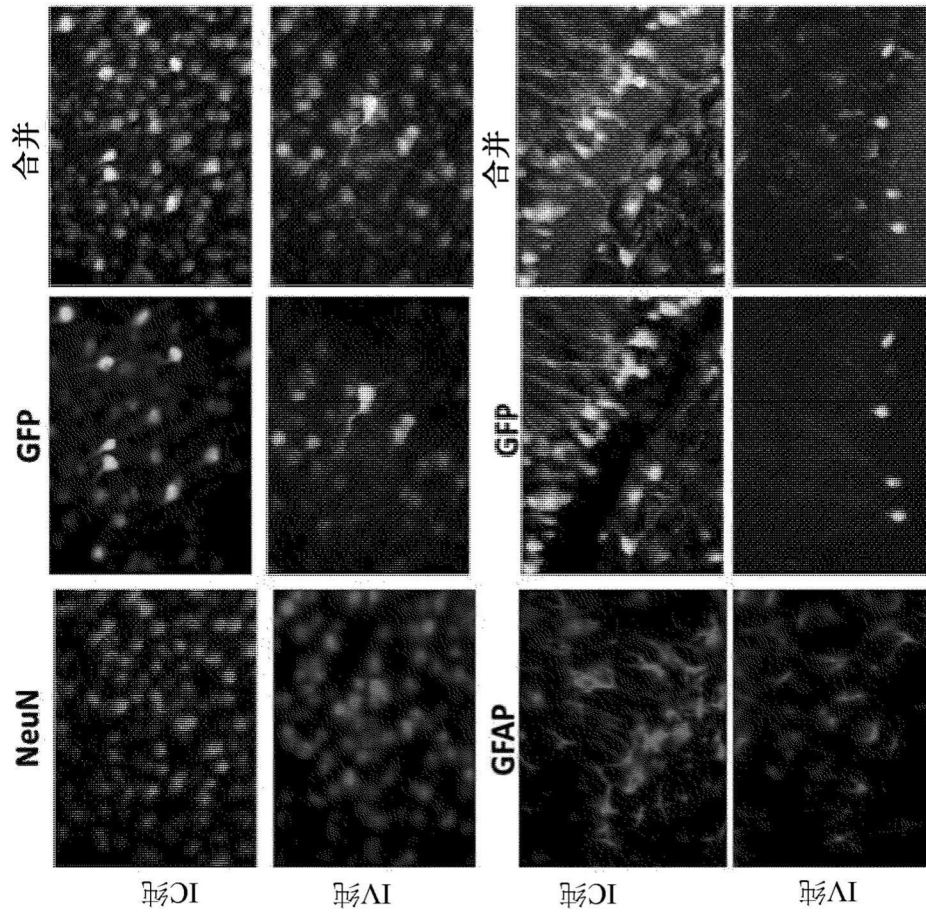


图7

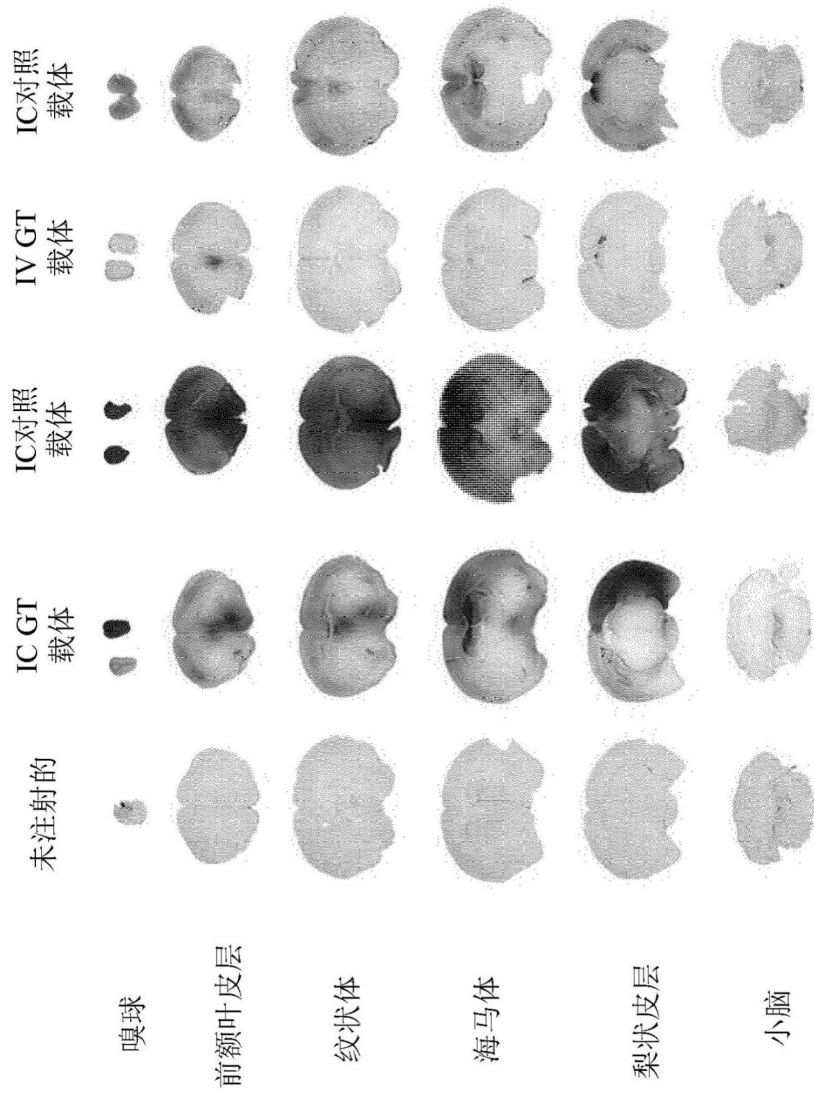


图8A

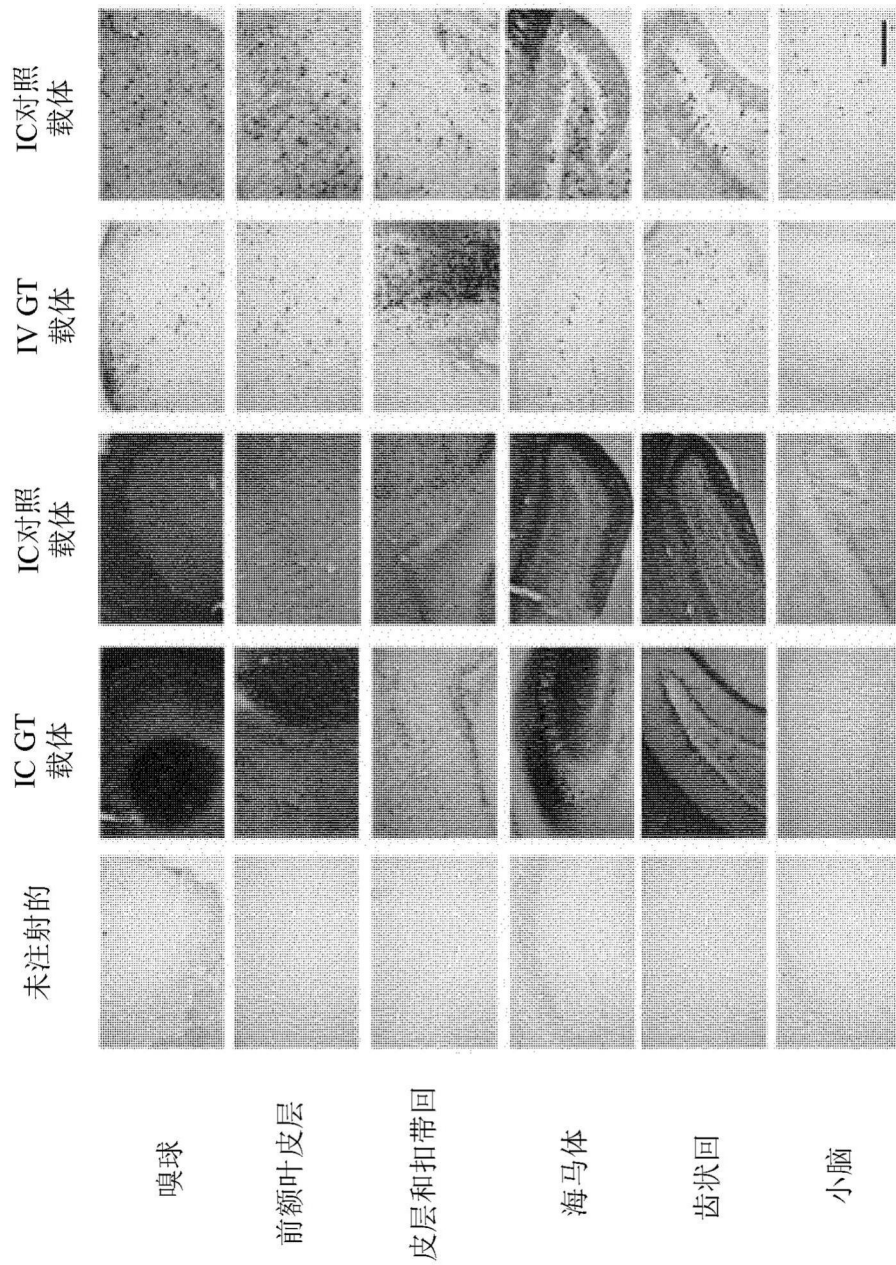


图8B

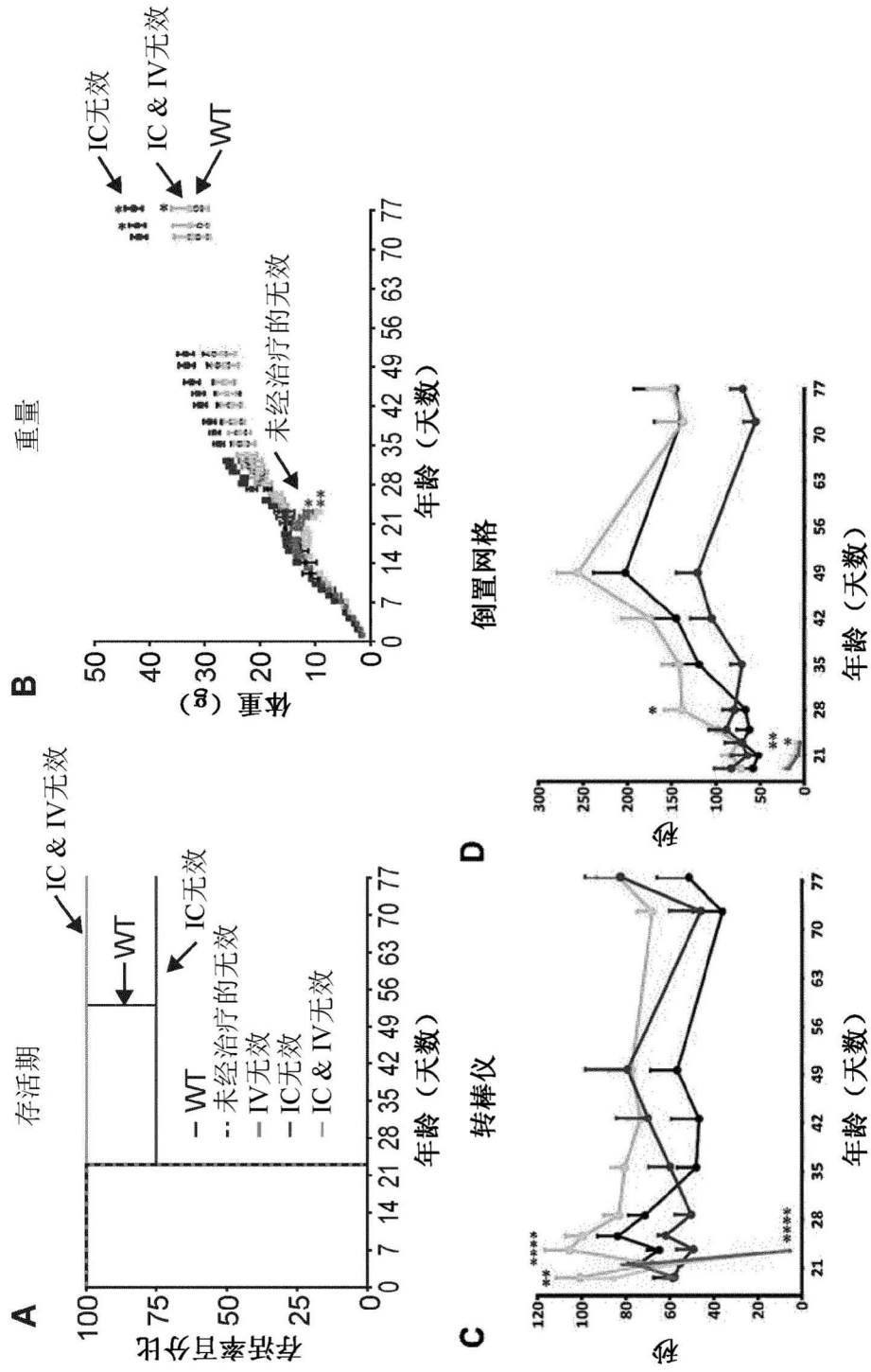


图9

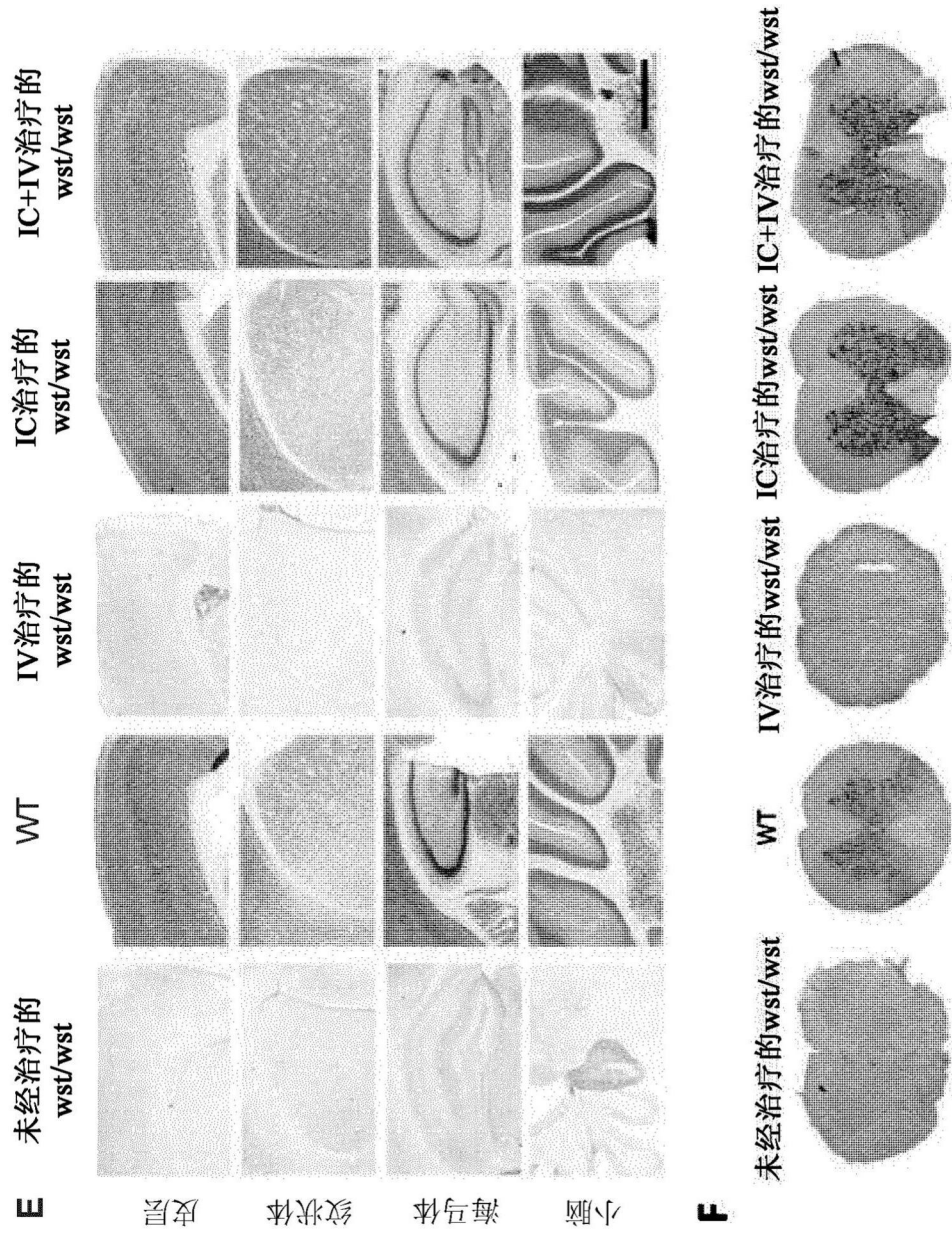


图9(续)

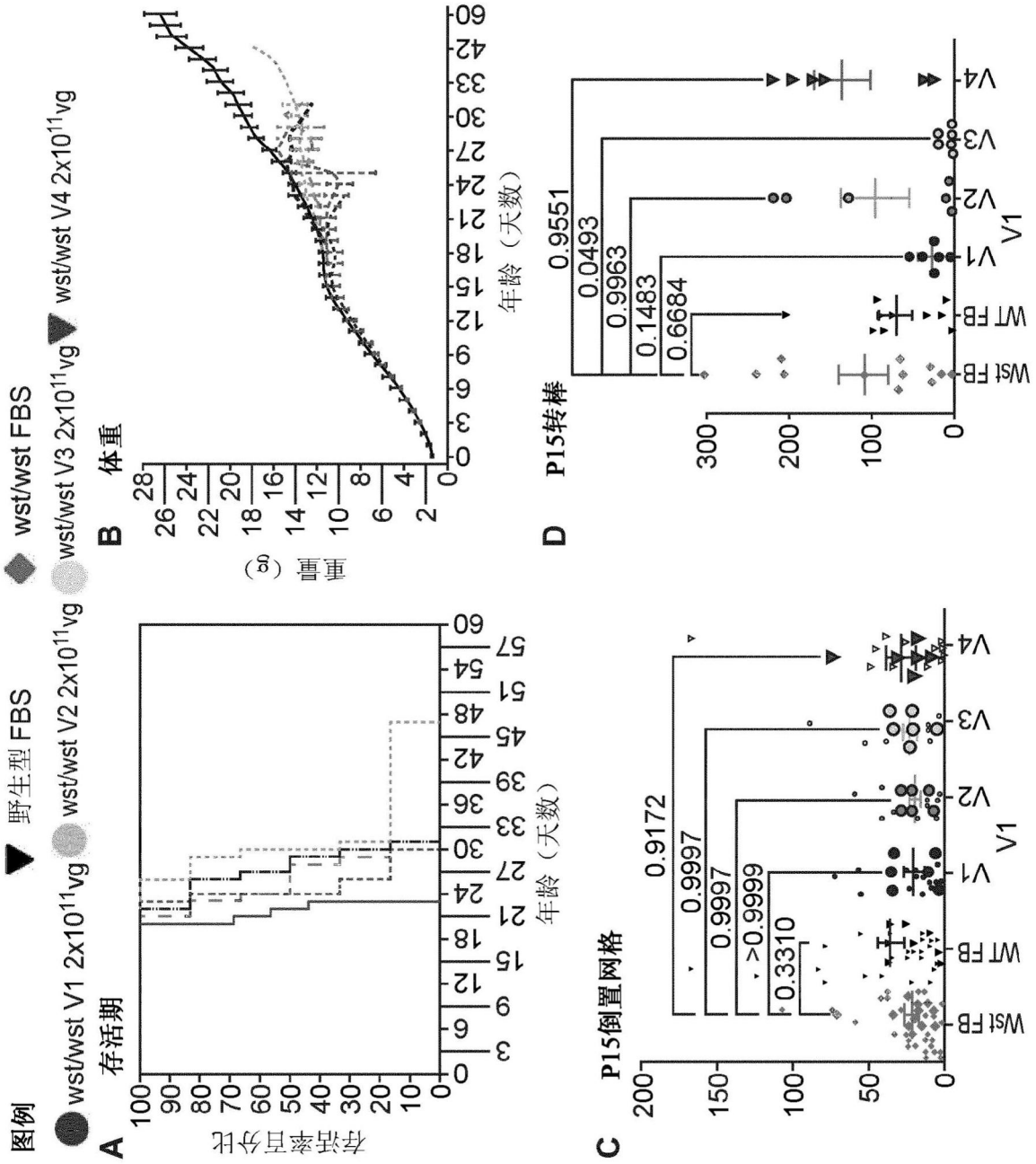


图10

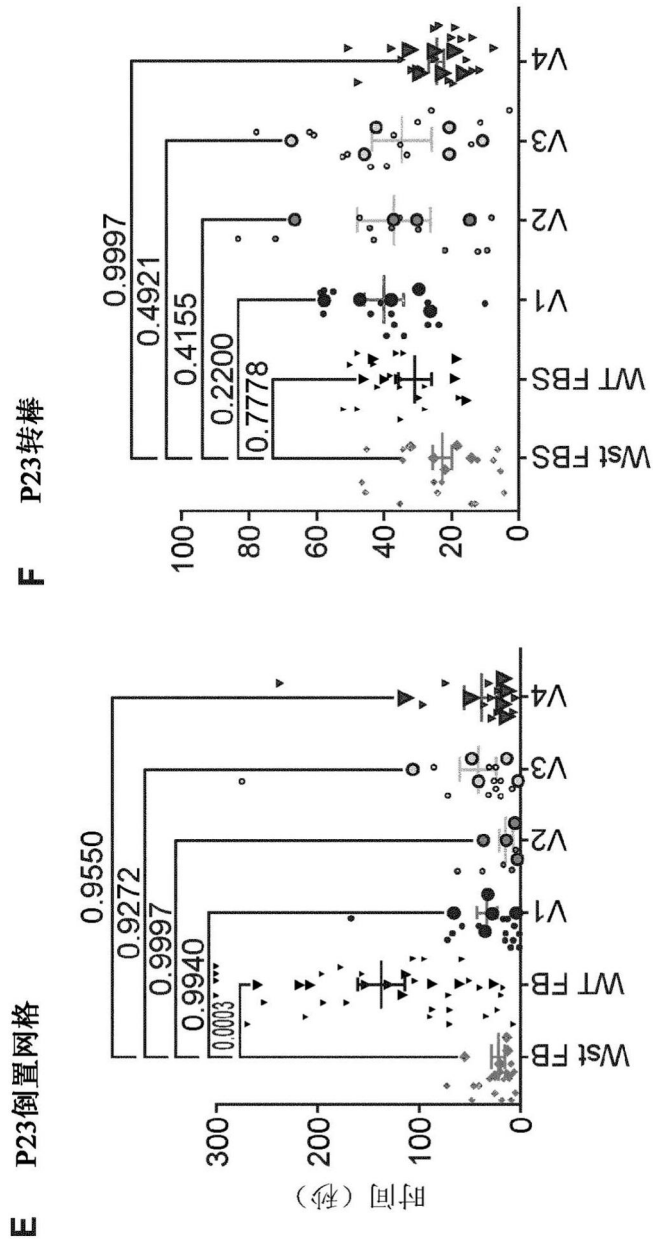
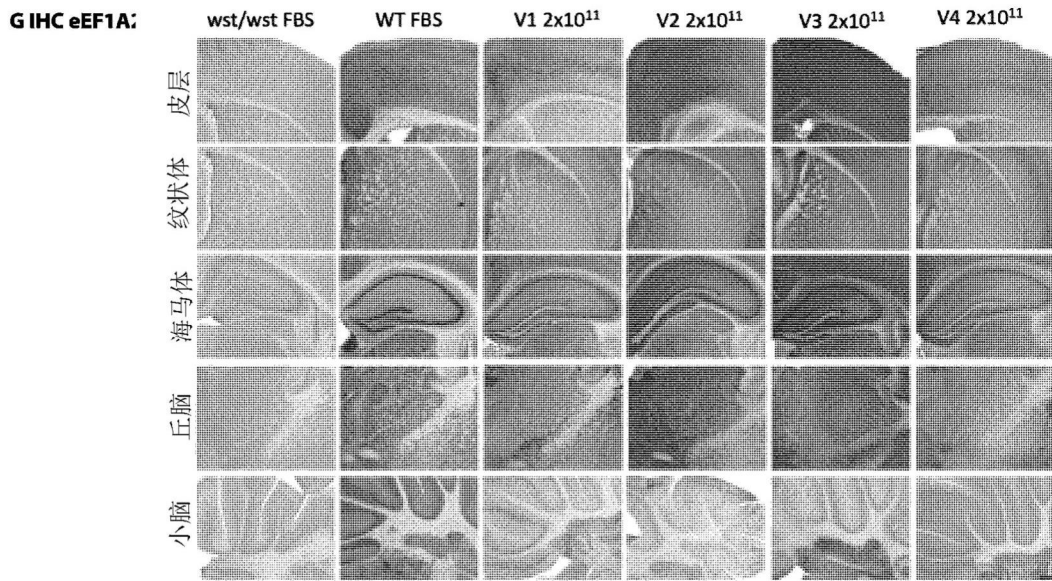


图10(续1)



H IHF eEF1A2: NeuN

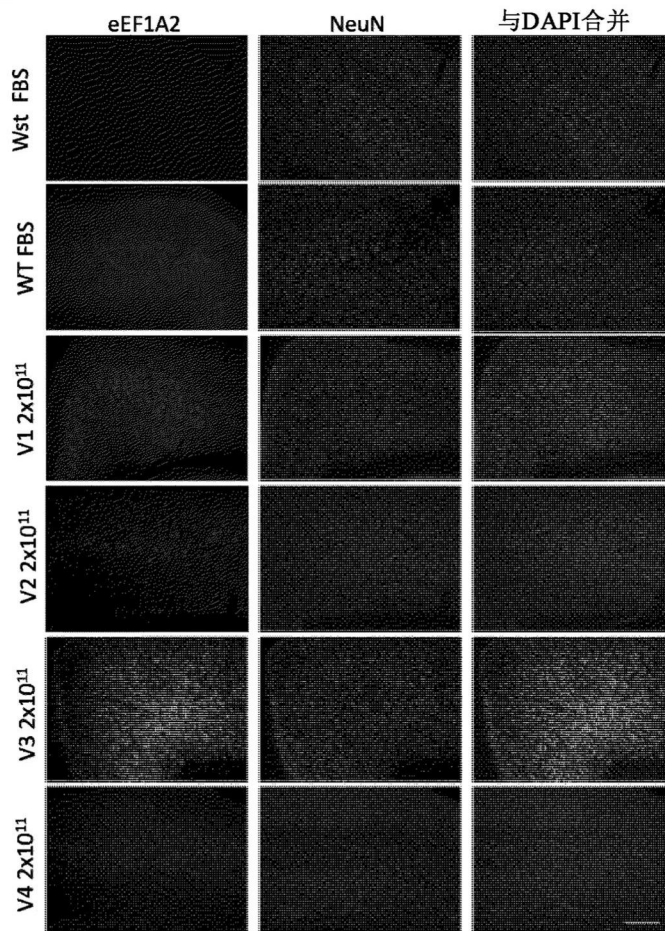


图10G-H

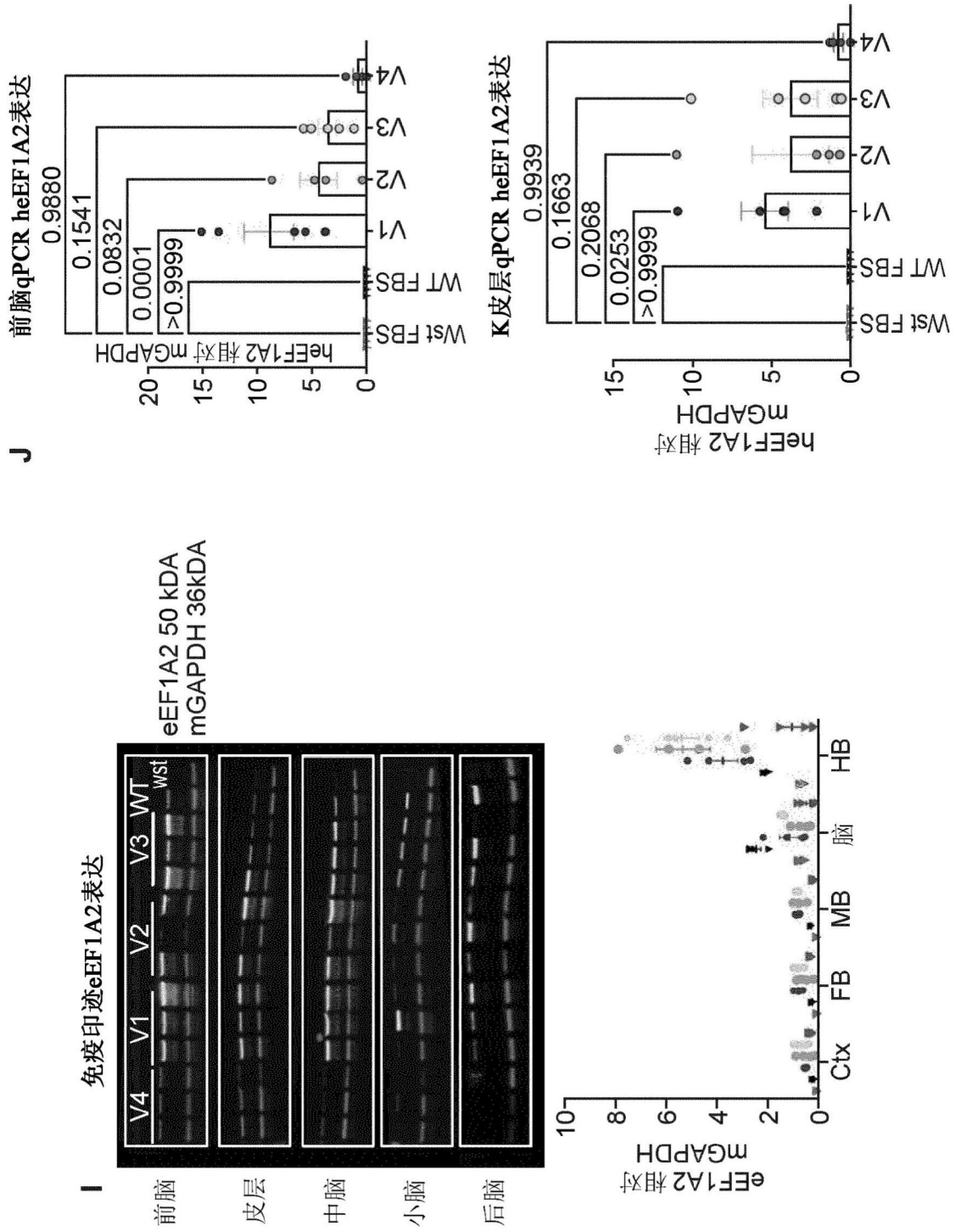


图10(续2)

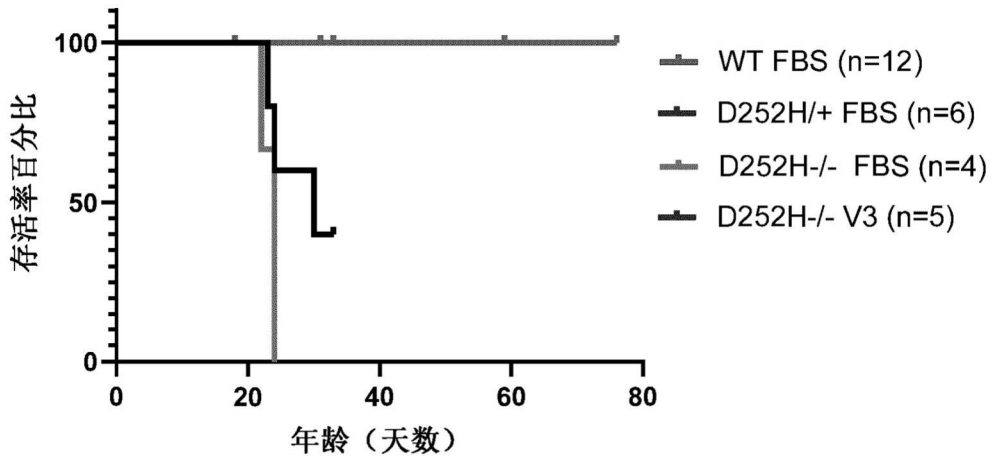


图11A

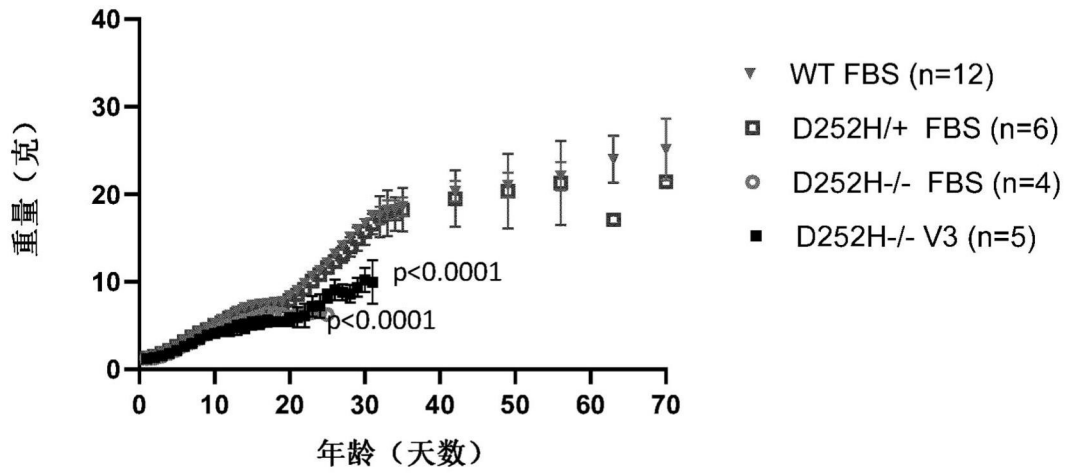


图11B

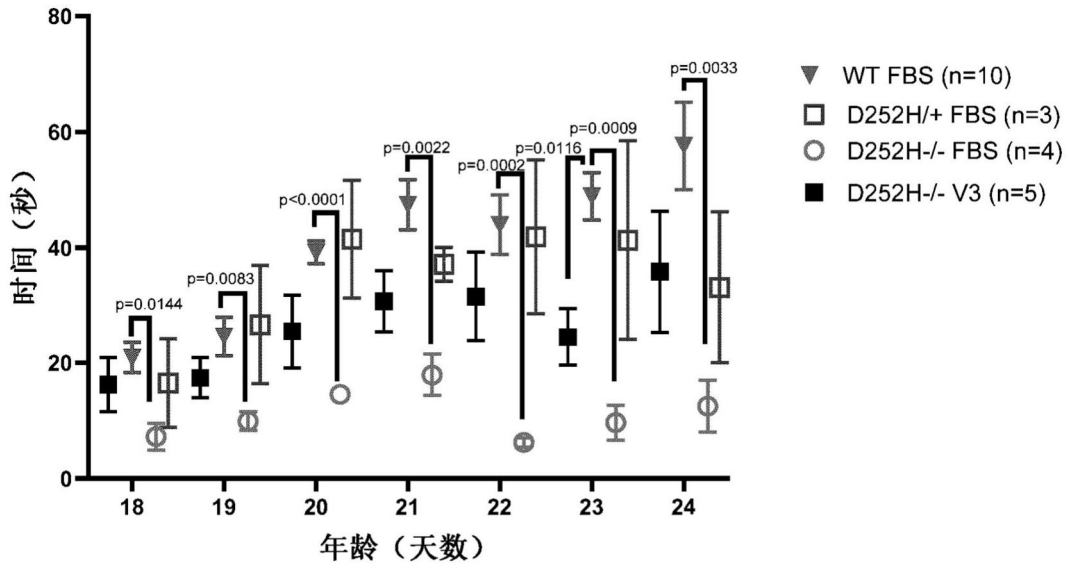


图11C

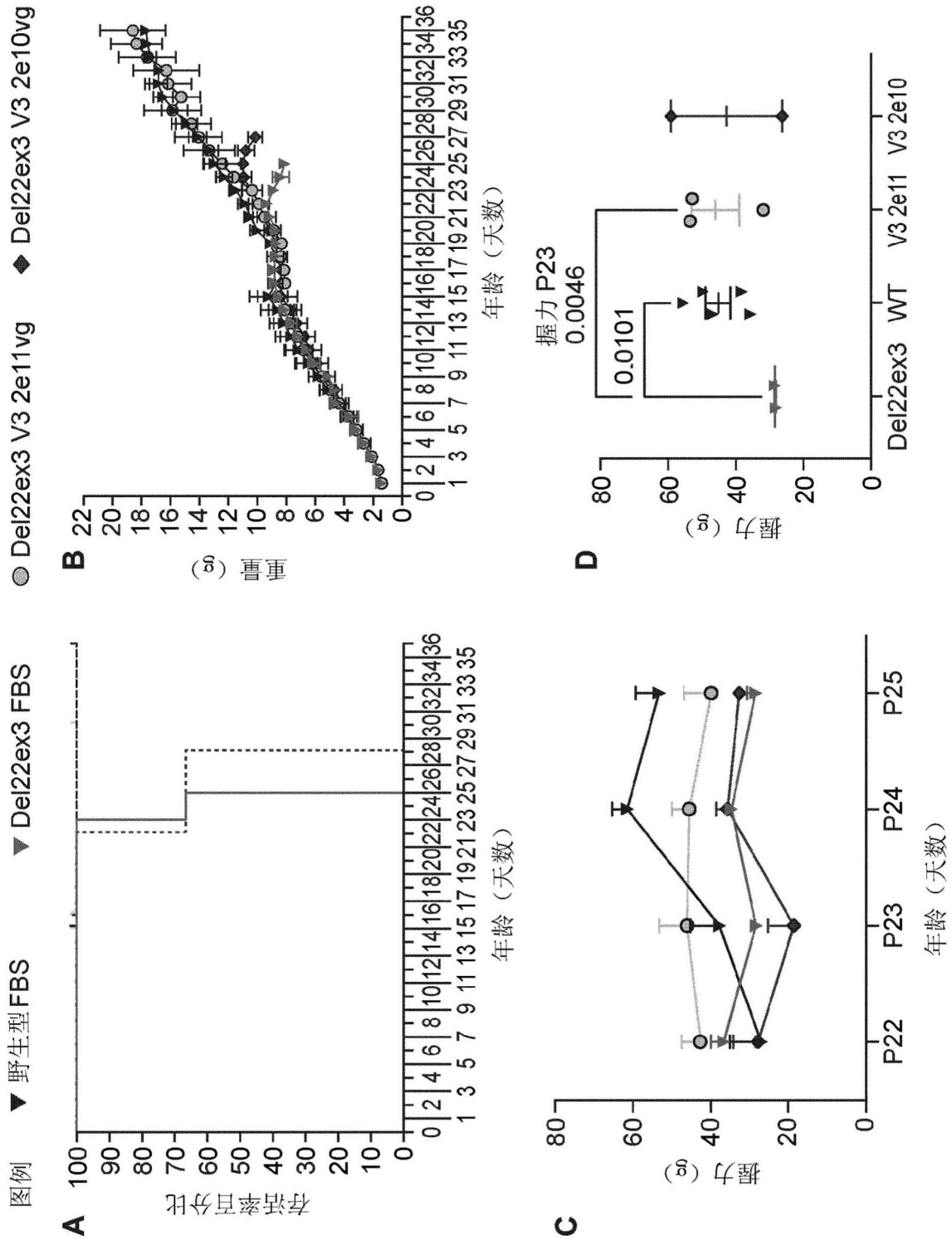


图12

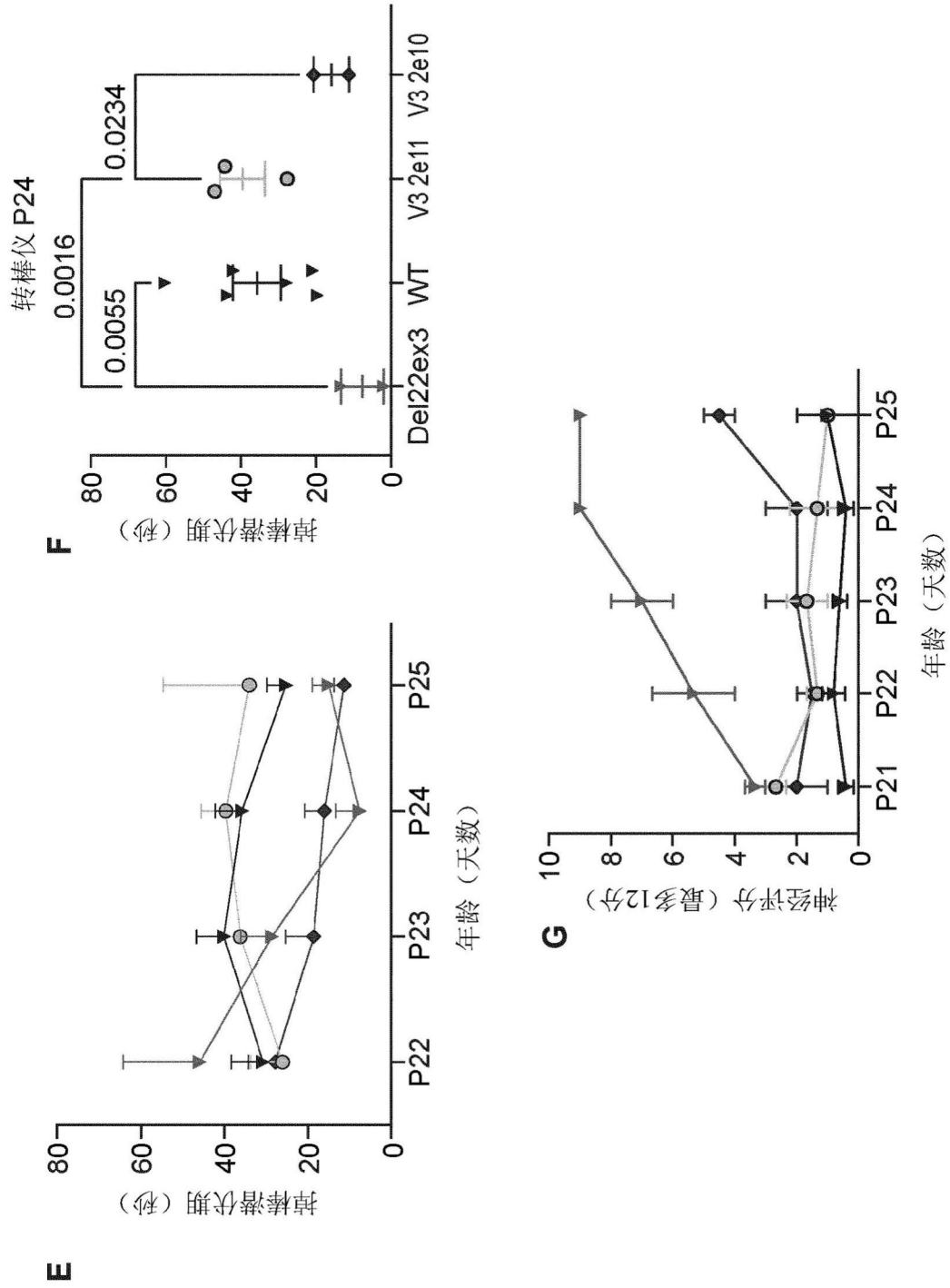


图12(续)