



(12) 发明专利申请

(10) 申请公布号 CN 116745409 A

(43) 申请公布日 2023. 09. 12

(21) 申请号 202180063852.3

(22) 申请日 2021.08.19

(30) 优先权数据

63/067,668 2020.08.19 US

(85) PCT国际申请进入国家阶段日

2023.03.17

(86) PCT国际申请的申请数据

PCT/US2021/046699 2021.08.19

(87) PCT国际申请的公布数据

W02022/040432 EN 2022.02.24

(71) 申请人 萨雷普塔治疗公司

地址 美国马塞诸塞州

(72) 发明人 C·奥巴尼昂

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

专利代理师 罗文锋 彭昶

(51) Int.Cl.

C12N 7/00 (2006.01)

权利要求书4页 说明书31页
序列表36页 附图21页

(54) 发明名称

用于治疗雷特综合征的腺相关病毒载体

(57) 摘要

本公开提供了在用于治疗与mecp2基因中的突变相关的疾病如雷特综合征和/或延缓其发作的方法中使用的核酸(包含AAV表达盒)、AAV载体和组合物。此外,本文提供了用于治疗脑源性神经营养因子(BDNF)相关疾病和/或延缓其发作的方法。



1. 一种核酸,所述核酸包含腺相关病毒(AAV)表达盒,其中所述AAV表达盒从5'至3'包含:

5'反向末端重复序列(ITR);
合成的活性依赖性启动子;
雷特综合征相关基因;以及
3' ITR。

2. 如权利要求1所述的核酸,其中所述启动子驱动所述雷特综合征相关基因的表达。

3. 如权利要求1或2所述的核酸,其中所述启动子是MECP2非依赖性启动子。

4. 如权利要求1至3中任一项所述的核酸,其中所述启动子包含源自神经元即早期基因的启动子的核酸序列。

5. 如权利要求4所述的核酸,其中所述神经元即早期基因选自由以下组成的组:Arc基因、c-fos基因和egr-1基因。

6. 如权利要求1至5中任一项所述的核酸,其中所述启动子包含最小Arc基因启动子(ArcMin)。

7. 如权利要求6所述的核酸,其中所述ArcMin是人ArcMin(hArcMin)。

8. 如权利要求7所述的核酸,其中所述hArcMin包含SEQ ID NO:12的核酸序列,或与其至少90%一致的序列。

9. 如权利要求1至8中任一项所述的核酸,其中所述启动子包含环AMP应答元件(CRE)、血清应答元件(SRE)、突触活动应答元件(SARE)、MEF2应答元件或其组合。

10. 如权利要求9所述的核酸,其中所述启动子包含突触活动应答元件(SARE)。

11. 如权利要求10所述的核酸,其中所述突触活动应答元件(SARE)是人突触活动应答元件(hSARE)。

12. 如权利要求11所述的核酸,其中所述hSARE包含SEQ ID NO:11的核酸序列,或与其至少90%一致的序列。

13. 如权利要求1至12中任一项所述的核酸,其中所述启动子包含人ArcMin(hArcMin)和至少一个hSARE。

14. 如权利要求13所述的核酸,其中所述启动子包含hArcMin和一个hSARE。

15. 如权利要求14所述的核酸,其中所述启动子包含SEQ ID NO:6的核酸序列,或与其至少90%一致的序列。

16. 如权利要求13所述的核酸,其中所述启动子包含hArcMin和五个hSARE。

17. 如权利要求16所述的核酸,其中所述启动子包含SEQ ID NO:16的核酸序列,或与其至少90%一致的序列。

18. 如权利要求1至17中任一项所述的核酸,其中所述启动子结合神经元活性依赖性转录因子。

19. 如权利要求18所述的核酸,其中所述神经元活性依赖性转录因子是cAMP应答元件结合蛋白(CREB)、肌细胞增强因子2(MEF2)、血清应答因子(SRF)或Elk-1。

20. 如权利要求1至19中任一项所述的核酸,其中所述雷特综合征相关基因编码脑源性神经营养因子(BDNF)、胰岛素样生长因子1(IGF1)、甲基-CpG-结合蛋白2(MECP2)、亨廷顿蛋白、亨廷顿相关蛋白1、正小齿同源框2(OTX-2)、含FXFD结构域的离子转运调节因子1

(FXVD1)、神经毒素-2- α (NRXN2)或蛋白激酶C γ (PRKCG)。

21. 如权利要求20所述的核酸,其中所述雷特综合征相关基因编码脑源性神经营养因子(BDNF)。

22. 如权利要求21所述的核酸,其中所述BDNF是人BDNF。

23. 如权利要求22所述的核酸,其中所述BDNF由SEQ ID NO:7的核酸列序或与其至少90%一致的序列编码。

24. 如权利要求1至23中任一项所述的核酸,其中所述5' ITR和所述3' ITR中的至少一者的长度为约110至约160个核苷酸。

25. 如权利要求1至24中任一项所述的核酸,其中所述5' ITR与所述3' ITR的长度相同。

26. 如权利要求1至24中任一项所述的核酸,其中所述5' ITR和所述3' ITR具有不同的长度。

27. 如权利要求1至26中任一项所述的核酸,其中所述5' ITR和所述3' ITR中的至少一者从AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的基因组中分离或来源于所述基因组。

28. 如权利要求1至23中任一项所述的核酸,其中所述5' ITR包含SEQ ID NO:1的序列。

29. 如权利要求1至23中任一项所述的核酸,其中所述3' ITR包含SEQ ID NO:2的序列。

30. 如权利要求1至29中任一项所述的核酸,其中所述AAV盒包含脑源性神经营养因子(BDNF)短3' UTR,或BDNF长3' UTR。

31. 如权利要求30所述的核酸,其中所述BDNF短3' UTR或所述BDNF长3' UTR位于雷特综合征相关基因和所述3' ITR之间。

32. 如权利要求30或权利要求31所述的核酸,其中所述AAV盒包含BDNF短3' UTR。

33. 如权利要求32所述的核酸,其中所述BDNF短3' UTR包含SEQ ID NO:8的核酸序列,或与其至少90%一致的序列。

34. 如权利要求30或权利要求31所述的核酸,其中所述AAV盒包含BDNF长3' UTR。

35. 如权利要求34所述的核酸,其中所述BDNF长3' UTR包含SEQ ID NO:10的核酸序列,或与其至少90%一致的序列。

36. 如权利要求1至35中任一项所述的核酸,其中所述AAV盒包含聚腺苷酸化信号。

37. 如权利要求36所述的核酸,其中所述聚腺苷酸化信号是从以下基因中的一者或多者分离或来源于以下基因中的一者或多者的聚腺苷酸化信号:猿猴病毒40(SV40)、rBG、 α -珠蛋白、 β -珠蛋白、人胶原、人生长激素(hGH)、多瘤病毒、人生长激素(hGH)或牛生长激素(bGH)。

38. 如权利要求36所述的核酸,其中所述AAV盒包含bGH聚腺苷酸化信号。

39. 如权利要求38所述的核酸,其中所述bGH聚腺苷酸化信号包含SEQ ID NO:9的核酸序列,或与其至少90%一致的序列。

40. 如权利要求1至39中任一项所述的核酸,其中所述AAV盒包含至少一个填充序列。

41. 如权利要求40所述的核酸,其中所述至少一个填充序列包含SEQ ID NO:13的核酸序列,或与其至少90%一致的序列。

42. 如权利要求1至41中任一项所述的核酸,其中所述AAV表达盒包含Kozak序列,其中所述Kozak序列与所述雷特综合征相关基因的起始密码子重叠。

43. 如权利要求42所述的核酸,其中所述Kozak序列包含SEQ ID NO:14的核酸序列,或与其至少90%一致的序列;或SEQ ID NO:15的核酸序列,或与其至少90%一致的序列。

44. 如权利要求1至43中任一项所述的核酸,其中所述AAV表达盒包含核酸序列SEQ ID NO:3,或与其至少90%一致的序列;SEQ ID NO:4或与其至少90%一致的序列;或SEQ ID NO:5或与其至少90%一致的序列。

45. 一种质粒,所述质粒包含如权利要求1至44中任一项所述的核酸。

46. 一种细胞,所述细胞包含如权利要求1至44中任一项所述的核酸或如权利要求45所述的质粒。

47. 一种产生重组AAV载体的方法,所述方法包括使AAV生产细胞与如权利要求1至44中任一项所述的核酸或如权利要求45所述的质粒接触。

48. 一种重组AAV载体,所述重组AAV载体通过如权利要求47所述的方法产生。

49. 如权利要求48所述的重组AAV载体,其中所述载体具有选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV和牛类AAV的血清型。

50. 如权利要求48或权利要求49所述的重组AAV载体,其中所述重组AAV载体是单链AAV(ssAAV)。

51. 如权利要求48至50中任一项所述的重组AAV载体,其中所述重组AAV载体是自身互补AAV(scAAV)。

52. 如权利要求48至51中任一项所述的重组AAV载体,其中所述AAV载体包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的衣壳蛋白。

53. 如权利要求48至52中任一项所述的重组AAV载体,其中所述AAV载体包含与野生型AAV衣壳蛋白相比具有一个或一个取代或突变的衣壳蛋白。

54. 一种组合物,所述组合物包含:(a) 如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞或如权利要求48至53中任一项所述的重组AAV载体;以及(b) 药学上可接受的载剂。

55. 一种用于在有需要的受试者中表达雷特综合征相关基因的方法,所述方法包括向所述受试者施用治疗有效量的如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞、如权利要求48至53中任一项所述的重组AAV载体或如权利要求54所述的组合物。

56. 如权利要求55所述的方法,其中所述受试者患有雷特综合征。

57. 一种用于在受试者中治疗雷特综合征或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞、如权利要求48至53中任一项所述的重组AAV载体或如权利要求54所述的组合物。

58. 一种用于在有需要的受试者中表达脑源性神经营养因子(BDNF)的方法,所述方法包括向所述受试者施用治疗有效量的如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞、如权利要求48至53中任一项所述的重组AAV载体或如权利要求54所述的组合物。

59. 如权利要求58所述的方法,其中所述受试者患有认知障碍或压力相关障碍。

60. 如权利要求58所述的方法,其中所述受试者患有抑郁症、强迫症、阿尔茨海默病、亨廷顿病和痴呆、神经性厌食症和神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、肥胖症、雷特综合征或化疗后认知障碍。

61. 一种在受试者中治疗BDNF相关疾病或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞、如权利要求48至53中任一项所述的重组AAV载体或如权利要求54所述的组合物。

62. 如权利要求61所述的方法,其中所述BDNF相关疾病是认知障碍和/或压力相关障碍。

63. 如权利要求61所述的方法,其中所述BDNF相关疾病为抑郁症、强迫症、阿尔茨海默病、亨廷顿病和痴呆、神经性厌食症、神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、双相性精神障碍、雷特综合征、重度抑郁症或化疗后认知障碍。

64. 一种在受试者中治疗MECP2相关疾病或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如权利要求1至44中任一项所述的核酸、如权利要求45所述的质粒、如权利要求46所述的细胞、如权利要求48至53中任一项所述的重组AAV载体或如权利要求54所述的组合物。

65. 如权利要求64所述的方法,其中所述MECP2相关疾病为MECP2重复综合征、MECP2相关重症新生儿脑病、PPM-X综合征或雷特综合征。

66. 如权利要求55至65中任一项所述的方法,其中所述受试者为人受试者。

67. 如权利要求55至66中任一项所述的方法,其中所述核酸、所述质粒、所述细胞、所述重组AAV载体或组合物通过注射到中枢神经系统中来施用。

68. 如权利要求55至67中任一项所述的方法,其中所述雷特综合征相关基因在所述受试者的神经元中表达。

69. 如权利要求68所述的方法,其中所述神经元是活动神经元。

用于治疗雷特综合征的腺相关病毒载体

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年8月19日提交的美国临时申请第63/067,668号的优先权,该申请的内容以全文引用的方式并入本文中。

技术领域

[0003] 本公开涉及分子生物学和基因治疗领域。更具体地,本公开涉及用于产生重组病毒载体的组合物和方法。

[0004] 序列表的并入

[0005] 以电子方式提交的文本文件的内容以全文引用的方式并入本文中:序列表的计算机可读格式副本(STRD_020_01W0_SeqList_ST25.txt,记录日期:2021年8月19日,文件大小:约89,555字节)。

背景技术

[0006] 雷特综合征(Rett Syndrome)是主要在女孩和妇女中表现为约6个月至18个月年龄的遗传性神经障碍。雷特综合征是一种罕见的疾病,女性发病率为约1/8500。患有雷特综合征的患者通常表现出广泛的症状,包括语言障碍、协调能力下降、小头畸形、重复性运动、癫痫发作、脊柱侧凸和认知障碍。虽然患有雷特综合征的女性的预期寿命为40多岁,但患有雷特综合征的男性常常在婴儿期死亡。

[0007] 雷特综合征由mecp2基因突变引起,并与神经元功能缺陷相关。雷特综合征的许多病例是由mecp2基因的新突变引起的。然而,在少数病例中,mecp2突变显示X连锁的显性遗传模式。mecp2的突变也会引起其他疾病,诸如MECP2重复综合征和PPM-X综合征,并且可能与自闭症谱系障碍有关。

[0008] 尚无已知的雷特综合征或其他mecp2相关疾病的治疗方法。雷特综合征症状通常使用抗惊厥药、特殊教育、物理疗法和支架来控制。基因治疗是治疗雷特综合征的潜在有效方法。然而,靶向神经系统的基因治疗经常面临几个技术障碍,包括鉴定在表达时将减轻症状的正确的神经元靶基因,以受调节的方式促进这些基因的高水平表达,以及以靶向方式将基因治疗构建体递送至神经元。因此,对于治疗雷特综合征和其他mecp2相关疾病的组合物和方法存在未满足的需求。

发明内容

[0009] 本公开提供了包含腺相关病毒(AAV)表达盒的核酸,其中AAV表达盒从5'到3'包含:5'反向末端重复序列(ITR);合成的活性依赖性启动子;雷特综合征相关基因;以及3' ITR。在一些实施方案中,启动子驱动雷特综合征相关基因的表达。在一些实施方案中,启动子是MECP2非依赖性启动子。在一些实施方案中,启动子包含源自神经元即早期基因的启动子的核酸序列。在一些实施方案中,神经元即早期基因选自以下组成的组:Arc基因、c-fos基因和egr-1基因。在一些实施方案中,启动子包含最小Arc基因启动子(ArcMin)。

[0010] 在一些实施方案中, ArcMin是人ArcMin(hArcMin)。在一些实施方案中, hArcMin包含SEQ ID NO:12的核酸序列, 或与其至少90%一致的序列。在一些实施方案中, 启动子包含环AMP应答元件(CRE)、血清应答元件(SRE)、突触活动应答元件(SARE)、MEF2应答元件或其组合。在一些实施方案中, 启动子包含突触活动应答元件(SARE)。在一些实施方案中, 突触活动应答元件(SARE)是人突触活动应答元件(hSARE)。在一些实施方案中, hSARE包含SEQ ID NO:11的核酸序列, 或与其至少90%一致的序列。在一些实施方案中, 启动子包含人ArcMin(hArcMin)和至少一个hSARE。在一些实施方案中, 启动子包含hArcMin和一个hSARE。在一些实施方案中, 启动子包含SEQ ID NO:6的核酸序列, 或与其至少90%一致的序列。在一些实施方案中, 启动子包含hArcMin和五个hSARE。在一些实施方案中, 启动子包含SEQ ID NO:16的核酸序列, 或与其至少90%一致的序列。

[0011] 在一些实施方案中, 启动子结合神经元活性依赖性转录因子。在一些实施方案中, 神经元活性依赖性转录因子是cAMP应答元件结合蛋白(CREB)、肌细胞增强因子2(MEF2)、血清应答因子(SRF)或Elk-1。在一些实施方案中, 雷特综合征相关基因编码脑源性神经营养因子(BDNF)、胰岛素样生长因子1(IGF1)、甲基-CpG-结合蛋白2(MECP2)、亨廷顿蛋白(Huntingtin protein)、亨廷顿相关蛋白1、正小齿同源框2(Orthodenticle homeobox 2)(OTX-2)、含FXFD结构域的离子转运调节因子1(FXFD1)、神经毒素-2- α (NRXN2)或蛋白激酶C γ (PRKCG)。在一些实施方案中, 雷特综合征相关基因编码脑源性神经营养因子(BDNF)。在一些实施方案中, BDNF是人BDNF。在一些实施方案中, BDNF由SEQ ID NO:7的核酸序列或与其至少90%一致的序列编码。

[0012] 在一些实施方案中, 5' ITR和3' ITR中的至少一者的长度为约110至约160个核苷酸。在一些实施方案中, 5' ITR与3' ITR的长度相同。在一些实施方案中, 5' ITR和3' ITR具有不同的长度。在一些实施方案中, 5' ITR和3' ITR中的至少一者从AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的基因组中分离或来源于该基因组。在一些实施方案中, 5' ITR包含SEQ ID NO:1的序列。在一些实施方案中, 3' ITR包含SEQ ID NO:2的序列。

[0013] 在一些实施方案中, AAV盒包含脑源性神经营养因子(BDNF)短3' UTR, 或BDNF长3' UTR。在一些实施方案中, BDNF短3' UTR或BDNF长3' UTR位于雷特综合征相关基因和3' ITR之间。在一些实施方案中, AAV盒包含BDNF短3' UTR。在一些实施方案中, BDNF短3' UTR包含SEQ ID NO:8的核酸序列, 或与其至少90%一致的序列。在一些实施方案中, AAV盒包含BDNF长3' UTR。在一些实施方案中, BDNF长3' UTR包含SEQ ID NO:10的核酸序列, 或与其至少90%一致的序列。

[0014] 在一些实施方案中, AAV盒包含聚腺苷酸化信号。在一些实施方案中, 聚腺苷酸化信号是从以下基因中的一者或多者分离或来源于以下基因中的一者或多者的聚腺苷酸化信号: 猿猴病毒40(SV40)、rBG、 α -珠蛋白、 β -珠蛋白、人胶原、人生长激素(hGH)、多瘤病毒、人生长激素(hGH)或牛生长激素(bGH)。在一些实施方案中, AAV盒包含bGH聚腺苷酸化信号。在一些实施方案中, bGH聚腺苷酸化信号包含SEQ ID NO:9的核酸序列, 或与其至少90%一致的序列。

[0015] 在一些实施方案中, AAV盒包含至少一个填充序列。在一些实施方案中, 至少一个填充序列包含SEQ ID NO:13的核酸序列, 或与其至少90%一致的序列。在一些实施方案中,

AAV表达盒包含Kozak序列,其中Kozak序列与雷特综合征相关基因的起始密码子重叠。在一些实施方案中,Kozak序列包含SEQ ID NO:14的核酸序列,或与其至少90%一致的序列;或SEQ ID NO:15的核酸序列,或与其至少90%一致的序列。在一些实施方案中,AAV表达盒包含核酸序列SEQ ID NO:3,或与其至少90%一致的序列;SEQ ID NO:4或与其至少90%一致的序列;或SEQ ID NO:5或与其至少90%一致的序列。

[0016] 本公开提供了包含本文公开的任一种核酸的质粒。本公开还提供了包含本文公开的任一种核酸或本文公开的任一种质粒的细胞。本公开还提供了生产重组AAV载体的方法,该方法包括使AAV生产细胞与本文公开的任一种核酸或本文公开的任一种质粒接触。本公开提供了通过本文公开的产生重组AAV载体的任一种方法产生的重组AAV载体。在一些实施方案中,载体具有选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV和牛类AAV的血清型。在一些实施方案中,重组AAV载体是单链AAV(ssAAV)。在一些实施方案中,重组AAV载体是自身互补AAV(scAAV)。在一些实施方案中,AAV载体包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的衣壳蛋白。在一些实施方案中,AAV载体包含与野生型AAV衣壳蛋白相比具有一个或一个取代或突变的衣壳蛋白。

[0017] 本公开提供了组合物,该组合物包含:(a) 本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体;以及(b) 药学上可接受的载剂。本公开提供了用于在有需要的受试者中表达雷特综合征相关基因的方法,该方法包括向受试者施用治疗有效量的本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体,或本文公开的任一种组合物。在一些实施方案中,受试者患有雷特综合征。

[0018] 本公开提供了用于在受试者中治疗雷特综合征或延缓其发作的方法,该方法包括向受试者施用治疗有效量的本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体,或本文公开的任一种组合物。本公开提供了在有需要的受试者中表达脑源性神经营养因子(BDNF)的方法,该方法包括向受试者施用治疗有效量的本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体,或本文公开的任一种组合物。

[0019] 在一些实施方案中,受试者患有认知障碍或压力相关障碍。在一些实施方案中,受试者患有抑郁症、强迫症、阿尔茨海默病(Alzheimer's disease)、亨廷顿病(Huntington's disease)和痴呆、神经性厌食症和神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、肥胖症、雷特综合征或化疗后认知障碍。

[0020] 本公开提供了用于在受试者中治疗BDNF相关疾病或延缓其发作的方法,该方法包括向受试者施用治疗有效量的本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体,或本文公开的任一种组合物。在一些实施方案中,BDNF相关疾病是认知障碍和/或压力相关障碍。在一些实施方案中,BDNF相关疾病是抑郁症、强迫症、阿尔茨海默病、亨廷顿病和痴呆、神经性厌食症、神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、双相性精神障碍、雷特综合征、重度抑郁症或化疗后认知障碍。

[0021] 本公开提供了用于在受试者中治疗MECP2相关疾病或延缓其发作的方法,该方法包括向受试者施用治疗有效量的本文公开的任一种核酸、本文公开的任一种质粒、本文公开的任一种细胞或本文公开的任一种重组AAV载体,或本文公开的任一种组合物。在一些实施方案中,MECP2相关疾病是MECP2重复综合征、MECP2相关重症新生儿脑病、PPM-X综合征或雷特综合征。在一些实施方案中,受试者是人受试者。在一些实施方案中,核酸、质粒、细胞、重组AAV载体或组合物通过注射到中枢神经系统中来施用。在一些实施方案中,雷特综合征相关基因在受试者的神经元中表达。在一些实施方案中,神经元是活动神经元。

[0022] 这些和其他实施方案将在下面的详细描述中更详细地阐述。

附图说明

[0023] 该专利或申请文件包含至少一张彩色附图。本专利或专利申请公开的副本以及彩色附图将由专利局在收到要求并支付必要费用后提供。

[0024] 图1示出了AAV表达盒,其包含hSARE-hArcMin启动子(SEQ ID NO:6)、BDNF基因(SEQ ID NO:7)、BDNF短3'UTR(SEQ ID NO:8)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0025] 图2示出了AAV表达盒,其包含hSARE-hArcMin启动子(SEQ ID NO:6)、BDNF基因(SEQ ID NO:7)、bGH聚A信号(SEQ ID NO:9)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0026] 图3示出了AAV表达盒,其包含hSARE-hArcMin启动子(SEQ ID NO:6)、BDNF基因(SEQ ID NO:7)和BDNF长3'UTR(SEQ ID NO:10),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0027] 图4示出了用于报道基因表达的AAV表达盒的示意图,该表达盒包含组成型(“C”)启动子hSyn或活性依赖性(“AD”)启动子hSARE-hArcMin。表达盒的元件也列在表3中。

[0028] 图5A是显示未用河豚毒素(TTX)处理的细胞中的报道蛋白荧光相对于用2 μ M TTX处理的细胞的荧光的条形图。图5B是显示用150mM KCl处理的细胞中的报道蛋白荧光相对于用2 μ M TTX处理的细胞的荧光的条形图。图5C是显示用30 μ M荷包牡丹碱(BIC)处理的细胞中的报道蛋白荧光相对于用2 μ M TTX处理的细胞的荧光的条形图。每个条显示了用包含所示AAV表达盒的AAV载体转导的细胞在X轴上所示的感染复数(MOI)下的相对荧光。

[0029] 图6A是显示用包含所示AAV表达盒的AAV载体转导的细胞中相对报道基因表达的图,该表达盒用30 μ M BIC处理,并进一步用2 μ M TTX处理,或如所示未用TTX处理。图6B是显示在用包含活性依赖型(AD) AAV表达盒的AAV载体转导的细胞中相对报道基因表达的条形图,该表达盒用30 μ M BIC处理,并进一步用2 μ M TTX处理,或如所示未用TTX处理。

[0030] 图7描绘了显示mScarlet荧光(顶部图像)和用神经元标记物,用包含所示组成型(C) AAV表达盒的AAV载体转导的神经元的抗BIII微管蛋白抗体染色(底部图像)的显微图像。

[0031] 图8描绘了显微图像,显示了:用包含“AD-dmScar-bGH” AAV表达盒的AAV载体转导的神经元的Hoescht DNA染色、dmScarlet荧光、抗BIII微管蛋白抗体染色和叠加图像。上图显示用2 μ M TTX和30 μ M BIC处理的细胞的图像,并且下图显示未用TTX处理但用30 μ M BIC处理的细胞的图像。

[0032] 图9A显示了AAV表达盒,其包含SEQ ID NO:26的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、mScarlet基因(SEQ ID NO:36)、BDNF短3'UTR(SEQ ID NO:8)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图9B显示了AAV表达盒,其包含SEQ ID NO:27的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、mScarlet基因(SEQ ID NO:36)和BDNF长3'UTR(SEQ ID NO:10),所有这些元件的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图9C显示了AAV表达盒,其包含SEQ ID NO:28的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、mScarlet基因(SEQ ID NO:36)、bGH聚A信号(SEQ ID NO:9)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0033] 图10A显示了AAV表达盒,其包含SEQ ID NO:29的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、dmScarlet基因(SEQ ID NO:35)、BDNF短3'UTR(SEQ ID NO:8)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图10B显示了AAV表达盒,其包含SEQ ID NO:30的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、dmScarlet基因(SEQ ID NO:35)和BDNF长3'UTR(SEQ ID NO:10),所有这些元件的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图10C显示了AAV表达盒,其包含SEQ ID NO:31的核酸序列,并且包含以下元件:hSyn启动子(SEQ ID NO:38)、dmScarlet基因(SEQ ID NO:35)、bGH聚A信号(SEQ ID NO:9)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0034] 图11A显示了AAV表达盒,其包含SEQ ID NO:32的核酸序列,并且包含以下元件:hSARE-hArcMin启动子(SEQ ID NO:6)、dmScarlet基因(SEQ ID NO:35)、BDNF短3'UTR(SEQ ID NO:8)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图11B显示了AAV表达盒,其包含SEQ ID NO:33的核酸序列,并且包含以下元件:hSARE-hArcMin启动子(SEQ ID NO:6)、dmScarlet基因(SEQ ID NO:35)和BDNF长3'UTR(SEQ ID NO:10),所有这些元件的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。图11C显示了AAV表达盒,其包含SEQ ID NO:34的核酸序列,并且包含以下元件:hSARE-hArcMin启动子(SEQ ID NO:6)、dmScarlet基因(SEQ ID NO:35)、bGH聚A信号(SEQ ID NO:9)和填充序列(SEQ ID NO:13),所有这些序列的侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。

[0035] 图12描绘了显微图像,显示了:用包含“AD-dmScar-长UTR”AAV表达盒的AAV载体转导的神经元的Hoescht DNA染色、dmScarlet荧光、抗βIII微管蛋白抗体染色和叠加图像。上图显示用2μM TTX和30μM BIC处理的细胞的图像,并且下图显示未用TTX处理但用30μMBIC处理的细胞的图像。

具体实施方式

[0036] 雷特综合征是一种罕见的遗传性神经系统疾病,主要发生在女孩中。该综合征的症状在6个月龄至18个月龄后明显,可包括以下中的一者或多者:语言问题、协调问题、重复性运动、生长缓慢、行走问题、头部尺寸变小、癫痫发作、脊柱侧凸、认知障碍和睡眠问题。患有雷特综合征的患者可以表现出这些症状的任何组合,每种症状的严重程度不同。

[0037] 雷特综合征由编码甲基CpG结合蛋白2 (MECP2) 蛋白的mecp2基因中的遗传突变引

起。参见UniProt登记号P51608,其以全文引用的方式并入本文中。在患有雷特综合征的女性中,已经鉴定出mecp2基因中超过620个突变,其中许多是单碱基对插入或缺失。MECP2有助于神经细胞的正常功能,并在成熟神经细胞中以高水平存在。MECP2结构域的甲基-CpG结合(MBD)可以识别具有5-甲基胞嘧啶修饰的DNA区。mecp2基因的突变也可以引起其他疾病,诸如MECP2重复综合征、MECP2相关重症新生儿脑病和PPM-X综合征。没有被批准的疗法来治疗任何与mecp2基因突变相关的疾病。

[0038] 本公开提供了在用于治疗与mecp2基因中的突变相关的疾病如雷特综合征和/或延缓其发作的方法中使用的核酸(包含AAV表达盒)、AAV载体和组合物。此外,本文提供了用于治疗脑源性神经营养因子(BDNF)相关疾病和/或延缓其发作的方法。

[0039] 定义

[0040] 在本文的说明书和所附权利要求中使用以下术语:

[0041] 单数形式的“一个(a)”、“一种(an)”和“该”也旨在包括复数形式,除非上下文中另有明确指示。

[0042] 此外,如本文所用,术语“约”当指可测量的值,诸如多核苷酸或多肽序列的长度、剂量、时间、温度等的量时,意指涵盖指定量的 $\pm 20\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、 $\pm 1\%$ 、 $\pm 0.5\%$ 或甚至 $\pm 0.1\%$ 的变化。

[0043] 此外,如本文所用,“和/或”是指并涵盖一个或多个相关列出项目的任何和所有可能的组合,以及当以替代方式(“或”)解释时缺少组合。

[0044] 术语“野生型”是本领域技术人员所理解的术语,并且意指生物体、菌株、基因、蛋白质或其在自然界中的特征的典型形式,区别于突变体或变体形式。例如,野生型蛋白质是该蛋白质在自然界中的典型形式。

[0045] 术语“突变蛋白”是本领域技术人员所理解的术语,并且是指基于氨基酸修饰(例如氨基酸取代、插入和/或缺失)的存在而区别于野生型蛋白形式的蛋白。术语“突变基因”是本领域技术人员所理解的术语,并且是指基于核酸修饰(例如核酸取代、插入和/或缺失)的存在而区别于野生型基因形式的基因。在一些实施方案中,突变基因编码突变蛋白。

[0046] “核酸”或“多核苷酸”是核苷酸碱基的序列,例如RNA、DNA或DNA-RNA杂合序列(包括天然存在和非天然存在的核苷酸)。在一些实施方案中,本公开的核酸是单链或双链DNA序列。核酸的长度可以为1-1,000、1,000-10,000、10,000-100,000、100,000-100万或大于100万个核苷酸。核酸将通常含有磷酸二酯键,尽管在某些情况下包括可能具有交替主链的核酸类似物,包括例如磷酰胺、硫代磷酸酯、二硫代磷酸酯、O-甲基磷酰亚胺键和肽核酸主链和键。其他类似的核酸包括具有阳性主链、非离子主链和非核糖主链的那些核酸。含有一个或多个碳环糖的核酸也包括在核酸的定义中。核糖-磷酸主链的这些修饰可以促进标记的添加,或者增加此类分子在生理环境中的稳定性和半衰期。本公开的核酸可以是线性的,或者可以是环状的(例如,质粒)。

[0047] 如本文所用,术语“启动子”是指指导可操作地连接的核酸转录的一个或多个核酸控制序列。启动子可以包括转录起始位点附近的核酸序列,诸如TATA元件。启动子也可以包括能被转录因子结合的顺式作用多核苷酸序列。

[0048] “组成型”启动子是在大多数环境和发育条件下有活性的启动子。“诱导型”启动子是在环境或发育调节下有活性的启动子。术语“可操作地连接”是指核酸表达控制序列(诸

如启动子或转录因子结合位点阵列)与第二核酸序列之间的功能性连接,其中表达控制序列指导对应于第二序列的核酸的转录。

[0049] “AAV表达盒”是被包装到重组AAV载体中的核酸,并且包含编码一个或多个转基因的序列。当AAV载体与靶细胞接触时,转基因由靶细胞表达。

[0050] 如本文所用,术语“病毒载体(virus vector)”、“病毒载体(viral vector)”或“基因递送载体”是指起核酸递送载体作用的病毒颗粒,并且其包含包装在病毒颗粒内的核酸(例如,AAV表达盒)。本公开的示例性病毒载体包括腺病毒载体、腺相关病毒载体(AAV)、慢病毒载体和逆转录病毒载体。

[0051] 如本文所用,术语“腺相关病毒”(AAV),包括但不限于AAV 1型、AAV 2型、AAV 3型(包括3A型和3B型)、AAV4型、AAV 5型、AAV 6型、AAV 7型、AAV 8型、AAV 9型、AAV 10型、AAV 11型、AAV 12型、AAV 13型、AAV rh32.33型、AAV rh8型、AAV rh10型、AAV rh74型、AAVhu.68型、鸟类AAV、牛类AAV、犬类AAV、马类AAV、绵羊类AAV、蛇类AAV、鬃狮蜥AAV、AAV2i8、AAV2g9、AAV-LK03、AAV7m8、AAV Anc80、AAV PHP.B,以及现在已知或以后发现的任何其他AAV。参见例如表1。

[0052] 表1:腺相关病毒血清型

[0053]

	GenBank 登 录号		GenBank 登 录号		GenBank 登 录号
完整基因组		进化枝 C		Rh57	AY530569
腺相关病毒 1	NC_002077, AF063497	Hu9	AY530629	Rh50	AY530563
腺相关病毒 2	NC_001401	Hu10	AY530576	Rh49	AY530562
腺相关病毒 3	NC_001729	Hu11	AY530577	Hu39	AY530601
腺相关病毒 3B	NC_001863	Hu53	AY530615	Rh58	AY530570
腺相关病毒 4	NC_001829	Hu55	AY530617	Rh61	AY530572
腺相关病毒 5	Y18065, AF085716	Hu54	AY530616	Rh52	AY530565
腺相关病毒 6	NC_001862, AAB95450.1	Hu7	AY530628	Rh53	AY530566
鸟 类 AAV ATCC VR-865	AY186198, AY629583, NC_004828	Hu18	AY530583	Rh51	AY530564
鸟类 AAV 毒株 DA-1	NC_006263, AY629583	Hu15	AY530580	Rh64	AY530574
牛类 AAV	NC_005889, AY388617, AAR26465	Hu16	AY530581	Rh43	AY530560
AAV11	AAT46339, AY631966	Hu25	AY530591	AAV8	AF513852
AAV12	ABI16639, DQ813647	Hu60	AY530622	Rh8	AY242997
进化枝 A		Ch5	AY243021	Rh1	AY530556
AAV1	NC_002077, AF063497	Hu3	AY530595	进化枝 F	
AAV6	NC_001862	Hu1	AY530575	Hu14 (AAV9)	AY530579
Hu.48	AY530611	Hu4	AY530602	Hu31	AY530596
Hu 43	AY530606	Hu2	AY530585	Hu32	AY530597
Hu 44	AY530607	Hu61	AY530623	HSC1	MI332400.1
Hu 46	AY530609	进 化 枝 D		HSC2	MI332401.1
进化枝 B		Rh62	AY530573	HSC3	MI332402.1
Hu. 19	AY530584	Rh48	AY530561	HSC4	MI332403.1
Hu. 20	AY530586	Rh54	AY530567	HSC5	MI332405.1
Hu 23	AY530589	Rh55	AY530568	HSC6	MI332404.1
Hu22	AY530588	Cy2	AY243020	HSC7	MI332407.1
Hu24	AY530590	AAV7	AF513851	HSC8	MI332408.1
Hu21	AY530587	Rh35	AY243000	HSC9	MI332409.1
Hu27	AY530592	Rh37	AY242998	HSC11	MI332406.1
Hu28	AY530593	Rh36	AY242999	HSC12	MI332410.1

[0054]

Hu 29	AY530594	Cy6	AY243016	HSC13	MI332411.1
Hu63	AY530624	Cy4	AY243018	HSC14	MI332412.1
Hu64	AY530625	Cy3	AY243019	HSC15	MI332413.1
Hu13	AY530578	Cy5	AY243017	HSC16	MI332414.1
Hu56	AY530618	Rh13	AY243013	HSC17	MI332415.1
Hu57	AY530619	进化枝 E		Hu68	
Hu49	AY530612	Rh38	AY530558	克隆分离物	
Hu58	AY530620	Hu66	AY530626	AAV5	Y18065, AF085716
Hu34	AY530598	Hu42	AY530605	AAV 3	NC_001729
Hu35	AY530599	Hu67	AY530627	AAV 3B	NC_001863
AAV2	NC_001401	Hu40	AY530603	AAV4	NC_001829
Hu45	AY530608	Hu41	AY530604	Rh34	AY243001
Hu47	AY530610	Hu37	AY530600	Rh33	AY243002
Hu51	AY530613	Rh40	AY530559	Rh32	AY243003
Hu52	AY530614	Rh2	AY243007	其他的	
Hu T41	AY695378	Bb1	AY243023	Rh74	
Hu S17	AY695376	Bb2	AY243022	鬃狮蜥 AAV	
Hu T88	AY695375	Rh10	AY243015	蛇类 AAV	NC_006148.1
Hu T71	AY695374	Hu17	AY530582		
Hu T70	AY695373	Hu6	AY530621		
Hu T40	AY695372	Rh25	AY530557		
Hu T32	AY695371	Pi2	AY530554		
Hu T17	AY695370	Pi1	AY530553		
Hu LG15	AY695377	Pi3	AY530555		

[0055] 术语“病毒生产细胞(viral production cell)”、“病毒生产细胞系”或“病毒生产细胞(viral producer cell)”是指用于生产病毒载体的细胞。HEK293和293T细胞是常见的病毒生产细胞系。下表2列出了各种病毒载体的示例性病毒生产细胞系。

[0056] 表2:示例性病毒生产细胞系

[0057]

病毒载体	示例性病毒生产细胞系
腺病毒	HEK293, 911, pTG6559, PER.C6, GH329, N52.E6, HeLa-E1, UR, VLI-293
腺相关病毒(AAV)	HEK293, Sf9
逆转录病毒	HEK293
慢病毒	293T

[0058] “HEK293”是指最初来源于在组织培养物中生长的人胚肾细胞的细胞系。HEK293细胞系易于在培养物中生长,并且通常用于病毒生产。如本文所用,“HEK293”也可指一种或多种变体HEK293细胞系,即来源于原始HEK293细胞系的另外包含一种或多种遗传改变的细胞系。针对一种或多种特定应用,已经开发并优化了许多变体HEK293系。例如,293T细胞系含有SV40大T抗原,其允许含有SV40复制起点的转染质粒的附加型复制,导致所需基因产物的表达增加。

[0059] “Sf9”是指昆虫细胞系,其是来源于亲本草地贪夜蛾细胞系IPLB-Sf-21-AE的克隆分离物。Sf9细胞可以在不存在血清的情况下生长,并且可以附着或悬浮培养。

[0060] “转染试剂”是指增强核酸转移到细胞中的组合物。本领域常用的一些转染试剂包括一种或多种与核酸和细胞表面结合的脂质(例如,LipofectamineTM)。

[0061] 确定两个或多个氨基酸序列之间序列相似性或一致性的方法是本领域已知的。序列相似性或一致性可以使用本领域已知的标准技术来确定,包括但不限于Smith&Waterman,Adv.Appl.Math.2,482(1981)的局部序列一致性算法、通过Needleman&Wunsch,J Mol.Biol.48,443(1970)的序列一致性比对算法、通过Pearson&Lipman,Proc.Natl.Acad.Sci.USA 85,2444(1988)的相似性搜索法、通过这些算法的计算机化实施方式(遗传学软件包中的GAP、BESTFIT、FASTA和TFASTA,Genetics Computer Group,575Science Drive,Madison,WI),Devereux等人,Nucl.Acid Res.12,387-395(1984)描述的最佳拟合序列程序,或通过检查。

[0062] 另一种合适的算法是BLAST算法,在Altschul等人,J Mol.Biol.215,403-410,(1990)和Karlin等人,Proc.Natl.Acad.Sci.USA90,5873-5787(1993)中描述。特别有用的BLAST程序是WU-BLAST-2程序,得自Altschul等人,Methods in Enzymology,266,460-480(1996);<http://blast.wustl.edu/blast/README.html>。WU-BLAST-2使用几个搜索参数,这些参数任选地设置为默认值。参数是动态值,并且由程序本身根据特定序列的组成和搜索所关注的序列所针对的特定数据库的组成来确定;然而,可以调整这些值以增加灵敏度。

[0063] 此外,如Altschul等人,(1997)Nucleic Acids Res.25,3389-3402所报道,另一种有用的算法是gapped BLAST。

[0064] 如本文所用,“治疗(treatment)”或“治疗(treating)”或“减轻”或“改善”可互换使用。这些术语是指获得有益的或期望结果的方法,包括但不限于治疗益处和/或预防益处。治疗益处是指在治疗中对一种或多种疾病、病况或症状的任何治疗相关的改善或效果。对于预防益处,可以将组合物施用于处于发展特定疾病、病况或症状的风险中的受试者,或者施用于报告了一种或多种疾病的生理症状的受试者,即使疾病、病况或症状可能尚未显现。

[0065] 术语“受试者”、“个体”和“患者”在本文中可互换使用,指脊椎动物,诸如哺乳动物。哺乳动物可以是例如小鼠、大鼠、兔、猫、狗、猪、羊、马、非人灵长类动物(例如食蟹猴、黑猩猩)或人。还包括体内获得的或体外培养的受试者的组织、细胞或其衍生物。人受试者可以是成人、青少年、儿童(2岁至14岁)、婴儿(1个月至24个月)或新生儿(最多1个月)。在一些实施方案中,成人是约65岁或更老,或约60岁或更老的老年人。在一些实施方案中,受试者是孕妇或打算怀孕的妇女。

[0066] 术语“有效量”或“治疗有效量”是指足以实现结果,例如,实现有益的或期望的结果的药剂的量。治疗有效量可以根据以下中的一种或多种而变化:所治疗的受试者和疾病状况、受试者的体重和年龄、疾病状况的严重程度、施用方式等,本领域普通技术人员可以容易地确定这些因素。具体剂量可以根据以下中的一种或多种而变化:所选择的具体药剂、所遵循的给药方案、该药剂是否与其他化合物联合施用、施用时机、待成像的组织以及携带该药剂的物理递送系统。

[0067] 如本文所用,术语“基因治疗”是指将遗传物质引入细胞以补偿异常基因或产生治疗性蛋白质的过程。

[0068] 如本文所用,当神经元质膜的至少一部分去极化时,神经元被称为“活动”神经元或具有“神经元活性”。在一些实施方案中,去极化是由电压门控 Na^+ 通道的打开引起的。在一些实施方案中,神经元质膜的去极化导致 Ca^{2+} 离子流入细胞,并触发即早期基因(IEG)的下游转录。在体内,突触后神经元的质膜的至少一部分的去极化可以由来自突触前神经元的信号触发。在体外,通过用KCl处理神经元可以诱导去极化。通过使用小分子(例如,荷包牡丹碱)灭活信号受体(诸如 γ -氨基丁酸A型(GABAA)受体)来促进体外去极化,其通常通过打开 Cl^- 或 K^+ 通道来减缓去极化。

[0069] 如本文所用,当没有穿过神经元质膜的净离子流时,神经元被称为“静息”神经元。在一些实施方案中,没有部分的静息神经元质膜去极化。在一些实施方案中,静息神经元具有约-70mV的静息膜电位。

[0070] AAV表达盒

[0071] 本公开提供了包含一个或多个腺相关病毒(AAV)表达盒的核酸序列。在一些实施方案中,AAV表达盒包含5'反向末端重复序列(ITR)、启动子、转基因和3'ITR。在一些实施方案中,转基因是雷特综合征相关基因。在一些实施方案中,AAV表达盒包含Kozak序列、聚腺苷酸化序列和/或填充序列。

[0072] 在一些实施方案中,AAV表达盒包含SEQ ID NO:3的核酸序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。在一些实施方案中,AAV表达盒包含SEQ ID NO:4的核酸序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。在一些实施方案中,AAV表达盒包含SEQ ID NO:5的核酸序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。

[0073] (i) 反向末端重复序列

[0074] 反向末端重复序列或ITR序列是介导AAV前病毒整合和将AAVDNA包装成病毒体的序列。ITR参与了AAV生命周期中的各种活动。例如,可以形成发夹结构的ITR序列在转染后从质粒切除、载体基因组的复制以及从宿主细胞基因组整合和拯救中起作用。

[0075] 本公开的AAV表达盒可以包含5'ITR和3'ITR。ITR序列的长度可以为约110至约160个核苷酸,例如110、111、112、113、114、115、116、117、118、119、120、121、122、123、124、125、126、127、128、129、130、131、132、133、134、135、136、137、138、139、140、141、142、143、144、145、146、147、148、149、150、151、152、153、154、155、156、157、158、159或160个核苷酸的长度。在一些实施方案中,ITR序列的长度可以为约141个核苷酸。在一些实施方案中,5'ITR与3'ITR的长度相同。在一些实施方案中,5'ITR和3'ITR具有不同的长度。在一些实施方案中,5'ITR比3'ITR长,并且在其他实施方案中,3'ITR比5'ITR长。

[0076] ITR可以从任何AAV的基因组中分离或来源于该基因组,例如表1中列出的AAV。在一些实施方案中,5'ITR和3'ITR中的至少一者从AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的基因组中分离或来源于该基因组。在一些实施方案中,5'ITR和3'ITR中的至少一者可

以是野生型或突变型ITR,其从除AAV以外的另一种细小病毒物种的成员中分离或来源于该成员。例如,在一些实施方案中,ITR可以从博卡病毒(bocavirus)或细小病毒B19分离或来源于以上病毒的野生型或突变型ITR。

[0077] 在一些实施方案中,ITR包含促进scAAV产生的修饰。在一些实施方案中,促进scAAV产生的修饰是从ITR中缺失末端解析序列(TRS)。在一些实施方案中,5' ITR是野生型ITR,并且3' ITR是缺乏末端解析序列的突变型ITR。在一些实施方案中,3' ITR是野生型ITR,并且5' ITR是缺乏末端解析序列的突变型ITR。在一些实施方案中,末端解析序列不存在于5' ITR和3' ITR中。在其他实施方案中,促进scAAV产生的修饰是用不同的发夹形成序列如shRNA形成序列替换ITR。

[0078] 在一些实施方案中,5' ITR可以包含SEQ ID NO:1的序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。在一些实施方案中,3' ITR可以包含SEQ ID NO:2的序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。在一些实施方案中,5' ITR包含SEQ ID NO:1的序列,并且3' ITR包含SEQ ID NO:2的序列。

[0079] 在一些实施方案中,AAV表达盒包含一个或多个“替代”ITR,即起与ITR相同作用的非ITR序列。参见例如.,Xie,J.等人,Mol.Ther.,25(6):1363-1374(2017)。在一些实施方案中,AAV表达盒中的ITR被替代ITR取代。在一些实施方案中,替代ITR包含发夹形成序列。在一些实施方案中,替代ITR是短发夹(sh)RNA形成序列。

[0080] (ii)活性依赖性启动子

[0081] 在一些实施方案中,本文所述的AAV表达盒包含启动子。在一些实施方案中,启动子是组织特异性启动子。在一些实施方案中,启动子是合成启动子。在一些实施方案中,启动子可以包含来源于内源启动子和/或内源增强子的核酸序列。在一些实施方案中,启动子可以包含来源于基因起始密码子上游区域的核酸序列。核酸序列可以位于起始密码子上游1个碱基对(bp)内至约8000个碱基对内,例如,位于起始密码子上游约10个碱基对、50个碱基对、100个碱基对、500个碱基对、1000个碱基对、1500个碱基对、2000个碱基对、2500个碱基对、3000个碱基对、3500个碱基对、4000个碱基对、4500个碱基对、5000个碱基对、5500个碱基对、6000个碱基对、6500个碱基对、7000个碱基对、7500个碱基对或8000个碱基对,包括位于其间的所有子范围和值内。

[0082] 不受任何理论的束缚,认为MECP2功能有助于参与神经元功能的靶基因的表达。当MECP2突变时,这些靶基因的表达受到影响。以MECP2非依赖性方式恢复这些靶基因的表达可以缓解与MECP2突变相关或由该突变引起的疾病的症状。因此,在一些实施方案中,启动子是MECP2非依赖性启动子。如本文所用,“MECP2非依赖性启动子”是指其功能不受MECP2功能降低或MECP2功能丧失而受损的启动子。

[0083] 不受任何理论的束缚,认为参与神经元功能的靶基因的组成型表达可能是有害的。因此,在一些实施方案中,启动子是诱导型启动子。在一些实施方案中,启动子的功能依赖于一种或多种刺激物的存在。在一些实施方案中,启动子是活性依赖性启动子。如本文所用,可与“神经元活性依赖性启动子”互换使用的“活性依赖性启动子”的功能由神经元活性

诱导。在一些实施方案中,活性依赖性启动子通过钙离子流入神经元来诱导。在一些实施方案中,活性依赖性启动子用于驱动任何靶基因的表达,当在细胞(诸如神经元)中组成型表达时,该靶基因是有害的。

[0084] 在一些实施方案中,启动子包含来源于神经元即早期基因的启动子和/或增强子的核酸序列。神经元即早期基因的身份不受限制,并且可以指本领域已知的或将来鉴定的任何神经元即早期基因。神经元即早期基因的非限制性实例包括Arc基因、c-fos基因和egr-1基因。因此,在一些实施方案中,启动子包含来源于Arc基因、c-fos基因或egr-1基因的启动子和/或增强子的核酸序列。在一些实施方案中,启动子包含来源于神经营养蛋白基因的启动子和/或增强子的核酸序列。神经营养蛋白基因的非限制性实例包括bdnf(脑源性神经营养因子)基因、ngf(神经生长因子)基因、神经营养因子-3基因和神经营养因子-4基因。

[0085] 在一些实施方案中,启动子能够结合一种或多种神经元活性依赖性转录因子。如本文所用,“神经元活性依赖性转录因子”是响应于神经元活性而被激活的转录因子。在一些实施方案中,神经元活性依赖性转录因子响应于神经元活性而促进基因表达。在一些实施方案中,神经元活性依赖性转录因子响应于神经元活性而抑制基因表达。在一些实施方案中,神经元活性依赖性转录因子被钙依赖性激酶级联激活。

[0086] 在一些实施方案中,启动子包含一种或多种能够结合一种或多种神经元活性依赖性转录因子的核酸序列。神经元活性依赖性转录因子的非限制性实例包括cAMP应答元件结合蛋白(CREB)、肌细胞增强因子2(MEF2)、血清应答因子(SRF)或Elk-1。在一些实施方案中,启动子能够结合多于一种神经元活性依赖性转录因子,例如,启动子能够结合2、3、4或5种神经元活性依赖性转录因子。在一些实施方案中,启动子能够结合两种神经元活性依赖性转录因子。在一些实施方案中,启动子能够结合三种神经元活性依赖性转录因子。在一些实施方案中,启动子能够结合以下神经元活性依赖性转录因子:CREB、MEF2和SRF。

[0087] 在一些实施方案中,启动子包含一个或多个应答元件。如本文所用,“应答元件”是在刺激物的存在下有助于驱动基因表达的启动子区域。在一些实施方案中,应答元件对于启动子在刺激物的存在下驱动基因表达是至关重要的。刺激物的非限制性实例包括激素;环境线索,诸如热或光;以及化学离子,诸如钙。在一些实施方案中,启动子包含在钙的存在下驱动基因表达的应答元件。在一些实施方案中,启动子包含存在于神经元即早期基因启动子中的应答元件。在一些实施方案中,启动子包含存在于Arc基因、c-fos基因和/或egr-1基因的启动子中的应答元件。

[0088] 在一些实施方案中,应答元件包括以下中的一者或多者:环AMP应答元件(CRE)、血清应答元件(SRE)、突触活动应答元件(SARE)和MEF2应答元件(MRE)。在一些实施方案中,启动子包含一个或多个CRE;一个或多个SRE;一个或多个SARE;一个或多个MRE,或其任何组合。在一些实施方案中,应答元件结合一种或多种神经元活性依赖性转录因子。在一些实施方案中,CRE与CREB结合。在一些实施方案中,SRE与SRF结合。在一些实施方案中,MRE与MEF2结合。在一些实施方案中,SARE与CREB、MEF2和SRF结合。

[0089] 在一些实施方案中,启动子包含1-20个SARE,例如1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个SARE。在一些实施方案中,启动子包含1个SARE。在一些实施方案中,启动子包含5个SARE。在一些实施方案中,SARE是小鼠SARE。在一些实施方案中,

SARE是人突触活动应答元件(hSARE)。在一些实施方案中,hSARE包含与SEQ ID NO:11的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,hSARE包含SEQ ID NO:11的核酸序列,或与其至少90%一致的序列。

[0090] 在一些实施方案中,启动子包含来源于Arc基因的启动子的核酸序列。在一些实施方案中,Arc基因是小鼠Arc基因。在一些实施方案中,Arc基因是人Arc基因。在一些实施方案中,启动子包含跨越Arc基因的核酸-300至+300的区域,或其任何子区域。在一些实施方案中,启动子包含最小Arc基因启动子(ArcMin)。在一些实施方案中,人ArcMin(hArcMin)包含Arc基因的短上游序列和5'UTR(-276至+208)。在一些实施方案中,小鼠ArcMin(mArcMin)包含Arc基因的短上游序列和5'UTR(-222至+198)。在一些实施方案中,hArcMin包含与SEQ ID NO:12的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,hArcMin包含SEQ ID NO:12的核酸序列,或与其至少90%一致的序列。

[0091] 在一些实施方案中,启动子包含hArcMin和至少一个hSARE。启动子中hSARE的数量没有限制,并且可以在1个hSARE到20个hSARE的范围内,诸如1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个hSARE。在一些实施方案中,启动子包含hArcMin和一个hSARE(称为SARE-ArcMin)。在一些实施方案中,SARE-ArcMin包含与SEQ ID NO:6的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,SARE-ArcMin包含SEQ ID NO:6的核酸序列,或与其至少90%一致的序列。在一些实施方案中,启动子包含hArcMin和5个hSARE(称为E-SARE)。在一些实施方案中,E-SARE包含与SEQ ID NO:16的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,E-SARE包含SEQ ID NO:16的核酸序列,或与其至少90%一致的序列。关于hArcMin和hSARE的进一步细节在Kawashima等人,NatMethods 10,889-895(2013)、Kawashima等人,FrontNeural Circuits2014年4月23日;8:37,以及Kawashima等人,PNAS1月6日,2009106(1)316-32中提供,这些文献以全文引用的方式并入本文中。

[0092] 在一些实施方案中,启动子诱导的基因表达水平高于神经元即早期基因如Arc基因、c-fos基因或egr-1基因的内源启动子。在一些实施方案中,启动子诱导的基因表达水平比神经元即早期基因如Arc基因、c-fos基因或egr-1基因的内源启动子高至少1.5倍。在一些实施方案中,启动子诱导的基因表达水平比神经元即早期基因如Arc基因、c-fos基因或egr-1基因的内源启动子高约1.5倍至约100倍(例如,约2倍、约3倍、约4倍、约5倍、约10倍、约15倍、约20倍、约30倍、约40倍、约50倍、约60倍、约70倍、约80倍、约90倍或约100倍)。

[0093] 在一些实施方案中,启动子包含来源于本领域常用于基因表达的一种或多种启动子的核酸序列。例如,在一些实施方案中,启动子还包含来源于CMV启动子、SV40早期启动子、SV40晚期启动子、金属硫蛋白启动子、鼠乳腺肿瘤病毒(MMTV)启动子、劳斯肉瘤(Rous

sarcoma) 病毒 (RSV) 启动子、多角体蛋白启动子、鸡 β -肌动蛋白 (CBA) 启动子、二氢叶酸还原酶 (DHFR) 启动子和磷酸甘油激酶 (PGK) 启动子的核酸序列。在一些实施方案中,启动子包含来源于鸡 β -肌动蛋白 (CBA) 启动子、EF-1 α 启动子或EF-1 α 短启动子的核酸序列。在一些实施方案中,启动子包含选自SEQ ID NO:17-20,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列中的任一者的序列。

[0094] 在一些实施方案中,本文所述的AAV表达盒还包含增强子。增强子可以是例如CMV增强子。在一些实施方案中,增强子包含SEQ ID NO:21的序列,或与其至少70%一致(例如,与其至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致,包括位于其间的所有值和子范围)的序列。

[0095] 在一些实施方案中,启动子还包含来源于任何一种或多种以下启动子的核酸序列:HMG-COA还原酶启动子;甾醇调节元件1 (SRE-1);磷酸烯醇丙酮酸羧激酶 (PEPCK) 启动子;人C反应蛋白 (CRP) 启动子;人葡萄糖激酶启动子;胆固醇7- α 羟化酶 (CYP-7) 启动子; β -半乳糖苷酶 α -2,6唾液酸转移酶启动子;胰岛素样生长因子结合蛋白 (IGFBP-1) 启动子;醛缩酶B启动子;人转铁蛋白启动子;I型胶原蛋白启动子;前列腺酸性磷酸酶 (PAP) 启动子;前列腺分泌蛋白94 (PSP 94) 启动子;前列腺特异性抗原复合物启动子;人腺激肽释放酶基因启动子 (hgt-1);肌细胞特异性增强子结合因子MEF-2;肌肉肌酸激酶启动子;胰腺炎相关蛋白启动子 (PAP);弹性蛋白酶1转录增强子;胰腺特异性淀粉酶和弹性蛋白酶增强子启动子;胰腺胆固醇酯酶基因启动子;子宫珠蛋白启动子;胆固醇侧链裂解 (SCC) 启动子; γ - γ 烯醇化酶(神经元特异性烯醇化酶,NSE) 启动子;神经丝重链 (NF-H) 启动子;人CGL-1/粒酶B启动子;末端脱氧转移酶 (TdT)、 λ 5、VpreB和Ick (淋巴细胞特异性酪氨酸蛋白激酶p56lck) 启动子;人CD2启动子及其3'转录增强子;人NK和T细胞特异性激活 (NKG5) 启动子;pp60c-src酪氨酸激酶启动子;器官特异性新抗原 (OSN)、mw 40kDa (p40) 启动子;结肠特异性抗原-P启动子;人 α -乳清蛋白启动子;磷酸烯醇丙酮酸羧激酶 (PEPCK) 启动子、HER2/neu启动子、酪蛋白启动子、IgG启动子、绒毛膜胚胎抗原启动子、弹性蛋白酶启动子、胆色素原脱氨酶启动子、胰岛素启动子、生长激素因子启动子、酪氨酸羟化酶启动子、白蛋白启动子、甲胎蛋白启动子、乙酰胆碱受体启动子、乙醇脱氢酶启动子、 α 或 β 珠蛋白启动子、T细胞受体启动子、骨钙素启动子、IL-2启动子、IL-2受体启动子、乳清 (wap) 启动子和II类MHC启动子。

[0096] (iii) 雷特综合征相关基因

[0097] 如本文所用,“雷特综合征相关基因”是指患有雷特综合征的受试者中的任何基因,该受试者可以通过基因治疗靶向以减轻雷特综合征的至少一种体征或症状。在一些实施方案中,在患有雷特综合征的受试者中,由雷特综合征相关基因编码的蛋白质的水平降低或不可检测。在一些实施方案中,由雷特综合征相关基因编码的蛋白质的水平在雷特综合征的至少一种症状发作时降低。在一些实施方案中,由雷特综合征相关基因编码的蛋白质的水平在正常出生后发育和/或神经元发育时增加。在一些实施方案中,雷特综合征相关基因编码有助于正常神经元功能的蛋白质。在一些实施方案中,雷特综合征相关基因编码参与促进和/或维持突触可塑性的蛋白质。在一些实施方案中,雷特综合征相关基因是神经营养蛋白。在一些实施方案中,在患有雷特综合征的受试者中存在雷特综合征相关基因的突变或雷特综合征相关基因的功能丧失。在一些实施方案中,雷特综合征相关基因的突变

或雷特综合征相关基因的功能丧失引起雷特综合征。在一些实施方案中,雷特综合征相关基因的突变或雷特综合征相关基因的功能丧失,特别是在前脑兴奋性神经元中,引起类似于mecp2基因丧失的症状。突变的类型不受限制,并且可以是插入、缺失、复制和/或取代。

[0098] 本公开提供了包含雷特综合征相关基因的AAV表达盒。在一些实施方案中,AAV表达盒包含雷特综合征相关基因,该基因编码蛋白质,包括治疗性(例如,用于医学或兽医用途)或免疫原性(例如,用于疫苗)多肽。在一些实施方案中,AAV表达盒包含哺乳动物雷特综合征相关基因。在一些实施方案中,AAV表达盒包含人雷特综合征相关基因。在一些实施方案中,AAV表达盒包含雷特综合征相关基因,该雷特综合征相关基因编码脑源性神经营养因子(BDNF)、胰岛素样生长因子1(IGF1)、甲基-CpG-结合蛋白2(MECP2)、亨廷顿蛋白、亨廷顿相关蛋白1、正小齿同源框2(OTX-2)、含FXFD结构域的离子转运调节因子1(FXYD1)、神经毒素-2- α (NRXN2)或蛋白激酶C γ (PRKCG)。在一些实施方案中,雷特综合征相关基因是选自由以下组成的组的基因:KCNA1(钾电压门控通道亚家族A成员1)、GABRA1(γ -氨基丁酸A型受体亚单位 α 1)、MAPK1(促分裂原活化蛋白激酶1)、NRXN2(神经连接蛋白2)、RBFox1(RNA结合Fox-1同源物1)、GNAO1(G蛋白亚单位 α 01)、NCAN(神经蛋白聚糖)、PRKCG(蛋白激酶C γ)、KCNJ4(钾向内整流通道亚家族J成员4)、CAMK2B(钙/钙调蛋白依赖性蛋白激酶II β)、EFNB3(肝配蛋白B3)、GABBR1(γ -氨基丁酸B型受体亚单位1)、LY6H(淋巴细胞抗原6家族成员H)、KCNA2(钾电压门控通道亚家族A成员2)和NEFL(神经丝轻链)。

[0099] 在一些实施方案中,雷特综合征相关基因编码脑源性神经营养因子(BDNF)。BDNF是在支持神经元存活和生长、突触发育和可塑性中起作用的神经营养蛋白。BDNF表达在雷特综合征中改变。bdnf的功能丧失导致与mecp2功能丧失相似的表型,并且BDNF功能在mecp2功能丧失时受到影响。关于BDNF功能的更多细节在Eduardo E. Benarroch, Neurology 2015年4月, 84(16) 1693-1704; Chang等人. Neuron. 2006; 49(3): 341-348; 以及Zhou等人, Neuron 52, 255-269, 2006年10月19日中提供, 这些文献各自以全文引用的方式并入本文中。在一些实施方案中, BDNF是人BDNF。在一些实施方案中, BDNF由与SEQ ID NO:7的核酸序列具有至少70%一致性(例如, 至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性, 包括位于其间的所有值和子范围)的核酸序列编码。在一些实施方案中, BDNF由SEQ ID NO:7的核酸序列或与其至少90%一致的序列编码。

[0100] 在BDNF中存在两个交替的聚腺苷酸化转录终止位点, 它们产生具有短3'UTR(约0.35kb长)或长3'UTR(约2.85kb长)的两个不同的bdnf mRNA群体。具有短3'UTR的bdnf mRNA变体主要定位于神经元的细胞体, 而具有长3'UTR的bdnf mRNA变体也定位于树突。在一些实施方案中, AAV表达盒包含BDNF短3'UTR, 或BDNF长3'UTR。在一些实施方案中, BDNF短3'UTR或BDNF长3'UTR存在于终止密码子和3' ITR之间。

[0101] 在一些实施方案中, BDNF短3'UTR包含与SEQ ID NO:8的核酸序列具有至少70%一致性(例如, 至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性, 包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中, BDNF短3'UTR包含SEQ ID NO:8的核酸序列, 或与其至少90%一致的序列。在一些实施方案中, BDNF短3'UTR包含与SEQ ID NO:10的核酸序列具有至少70%一致性(例如, 至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少

98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,BDNF长3'UTR包含SEQ ID NO:10的核酸序列,或与其至少90%一致的序列。

[0102] 在一些实施方案中,AAV表达盒包含Kozak序列。kozak序列是在许多真核mRNA转录物中起蛋白质翻译起始位点作用的核酸序列。在一些实施方案中,Kozak序列与起始密码子重叠。在一些实施方案中,Kozak序列包含与SEQ ID NO:14或SEQ ID NO:15的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,Kozak序列包含SEQ ID NO:14的核酸序列,或与其至少90%一致的序列;或SEQ ID NO:15的核酸序列,或与其至少90%一致的序列。

[0103] (iv) 聚腺苷酸化(聚A)信号

[0104] 聚腺苷酸化信号是在几乎所有哺乳动物基因中发现的核苷酸序列,并且控制将一串约200个腺苷残基(聚(A)尾)添加到基因转录物的3'末端。聚(A)尾有助于mRNA的稳定性,而缺少聚(A)尾的mRNA被迅速降解。也有证据表明,聚(A)尾的存在通过影响翻译的起始而积极地有助于mRNA的可翻译性。

[0105] 在一些实施方案中,本公开的AAV表达盒包含聚腺苷酸化信号。聚腺苷酸化信号可选自猿猴病毒40(SV40)、rBG、 α -珠蛋白、 β -珠蛋白、人胶原、人生长激素(hGH)、多瘤病毒、人生长激素(hGH)和牛生长激素(bGH)的聚腺苷酸化信号。

[0106] 在一些实施方案中,AAV表达盒包含bGH聚腺苷酸化信号。在一些实施方案中,bGH聚腺苷酸化信号包含与SEQ ID NO:9的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,bGH聚腺苷酸化信号包含SEQ ID NO:9的核酸序列,或与其至少90%一致的序列。

[0107] 在一些实施方案中,聚腺苷酸化信号是SV40聚腺苷酸化信号。在一些实施方案中,聚腺苷酸化信号是rBG聚腺苷酸化信号。在一些实施方案中,聚腺苷酸化信号包含SEQ ID NO:22或SEQ ID NO:23的序列。在一些实施方案中,聚腺苷酸化信号包含与SEQ ID NO:22或SEQ ID NO:23的序列至少90%、至少95%、至少96%、至少97%、至少98%或至少99%一致的序列。

[0108] (v) 填充序列

[0109] AAV载体通常接受具有限定大小范围的DNA插入片段,该大小范围通常为约4kb至约5.2kb,或稍大一些。因此,对于较短的序列,可能需要在插入片段中包含额外的核酸,以达到AAV载体可接受的所需长度。因此,在一些实施方案中,本公开的AAV表达盒可以包含填充序列。填充序列可以是例如长度为1-10、10-20、20-30、30-40、40-50、50-60、60-75、75-100、100-150、150-200、200-250、250-300、300-400、400-500、500-750、750-1,000、1,000-1,500、1,500-2,000、2,000-2,500、2,500-3,000、3,000-3,500、3,500-4,000、4,000-4,500或4,500-5,000或更多个核苷酸的序列。填充序列可以位于盒中任何所需的位置,使得它不会妨碍载体的功能或活性。在一些实施方案中,填充序列位于3'UTR的上游。例如,在一些实施方案中,填充序列存在于BDNF短3'UTR和3' ITR之间,或bGH聚A信号和3' ITR之间。

[0110] 在一些实施方案中,AAV盒包含至少一个填充序列。在一些实施方案中,填充序列

包含与SEQ ID NO:13的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。在一些实施方案中,填充序列包含SEQ ID NO:13的核酸序列,或与其至少90%一致的序列。

[0111] (vi)内含子序列

[0112] 在一些实施方案中,本公开的AAV表达盒可以包含内含子序列。与不存在内含子序列的表达相比,包含内含子序列可以增强表达。

[0113] 在一些实施方案中,内含子序列是杂合或嵌合序列。在一些实施方案中,内含子序列从SV40、 β -珠蛋白、鸡 β -肌动蛋白、小鼠微小病毒(MVM)、因子IX和/或人IgG(重链或轻链)中的一者或多者的内含子序列分离或来源于上述内含子序列。在一些实施方案中,内含子序列是嵌合的。在一些实施方案中,内含子序列包含SEQ ID NO:24或SEQ ID NO:25的序列。在一些实施方案中,内含子序列包含与SEQ ID NO:24或SEQ ID NO:25的核酸序列具有至少70%一致性(例如,至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%或100%一致性,包括位于其间的所有值和子范围)的核酸序列。

[0114] (vii)示例性AAV表达盒

[0115] 在一些实施方案中,AAV表达盒包含5'反向末端重复序列(ITR)、转基因和3' ITR。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:7(BDNF)和SEQ ID NO:2(3' ITR)。

[0116] 在一些实施方案中,AAV表达盒包含5' ITR、启动子、转基因和3' ITR。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:6(hSARE-ArcMin)、SEQ ID NO:7(BDNF)和SEQ ID NO:2(3' ITR)。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:16(E-SARE)、SEQ ID NO:7(BDNF)和SEQ ID NO:2(3' ITR)。

[0117] 在一些实施方案中,AAV表达盒包含5' ITR、启动子、转基因、聚腺苷酸化序列和3' ITR。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:6(hSARE-ArcMin)、SEQ ID NO:7(BDNF)、SEQ ID NO:9(bGH聚腺苷酸化序列)和SEQ ID NO:2(3' ITR)。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:16(E-SARE)、SEQ ID NO:7(BDNF)、SEQ ID NO:9(bGH聚腺苷酸化序列)和SEQ ID NO:2(3' ITR)。

[0118] 在一些实施方案中,AAV表达盒包含5' ITR、启动子、Kozak序列、转基因和3' ITR。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:6(hSARE-ArcMin)、SEQ ID NO:14或15(Kozak序列)、SEQ ID NO:7(BDNF)和SEQ ID NO:2(3' ITR)。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:16(E-SARE)、SEQ ID NO:14或15(Kozak序列)、SEQ ID NO:7(BDNF)和SEQ ID NO:2(3' ITR)。

[0119] 在一些实施方案中,AAV表达盒包含5' ITR、启动子、Kozak序列、转基因、聚腺苷酸化序列和3' ITR。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:6(hSARE-ArcMin)、SEQ ID NO:14或15(Kozak序列)、SEQ ID NO:7(BDNF)、SEQ ID NO:9(bGH聚腺苷酸化序列)和SEQ ID NO:2(3' ITR)。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:16(E-SARE)、SEQ ID NO:14或15(Kozak序列)、SEQ ID NO:7(BDNF)、SEQ ID NO:9(bGH聚腺苷酸化序列)和SEQ ID NO:2(3' ITR)。

[0120] 在一些实施方案中, AAV表达盒包含5' ITR、启动子、转基因、3' UTR和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8 (BDNF短3' UTR) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8 (BDNF短3' UTR) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:7 (BDNF)、SEQ ID NO:10 (BDNF长3' UTR) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:7 (BDNF)、SEQ ID NO:10 (BDNF长3' UTR) 和SEQ ID NO:2 (3' ITR)。

[0121] 在一些实施方案中, AAV表达盒包含5' ITR、启动子、转基因、3' UTR、聚腺苷酸化序列和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。

[0122] 在一些实施方案中, AAV表达盒包含5' ITR、启动子、Kozak序列、转基因、3' UTR和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR) 和SEQ ID NO:2 (3' ITR)。

[0123] 在一些实施方案中, AAV表达盒包含5' ITR、启动子、Kozak序列、转基因、3' UTR、聚腺苷酸化序列和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。

[0124] 在一些实施方案中, AAV表达盒包含5' ITR、填充序列、启动子、Kozak序列、转基因、3' 非翻译区 (3' UTR)、聚腺苷酸化序列和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:13 (填充序列)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:13 (填充序列)、SEQ ID NO:16 (E-SARE)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7 (BDNF)、SEQ ID NO:8或10 (BDNF短或长3' UTR)、SEQ ID NO:9 (bGH聚腺苷酸化序列) 和SEQ ID NO:2 (3' ITR)。

[0125] 在一些实施方案中, AAV表达盒包含5' ITR、启动子、Kozak序列、转基因、3' 非翻译区 (3' UTR)、聚腺苷酸化序列、填充序列和3' ITR。在一些实施方案中, AAV表达盒包含SEQ ID NO:1 (5' ITR)、SEQ ID NO:6 (hSARE-ArcMin)、SEQ ID NO:14或15 (Kozak序列)、SEQ ID NO:7

(BDNF)、SEQ ID NO:8或10(BDNF短或长3'UTR)、SEQ ID NO:9(bGH聚腺苷酸化序列)、SEQ ID NO:13(填充序列)和SEQ ID NO:2(3' ITR)。在一些实施方案中,AAV表达盒包含SEQ ID NO:1(5' ITR)、SEQ ID NO:13(填充序列)、SEQ ID NO:16(E-SARE)、SEQ ID NO:14或15(Kozak序列)、SEQ ID NO:7(BDNF)、SEQ ID NO:8或10(BDNF短或长3'UTR)、SEQ ID NO:9(bGH聚腺苷酸化序列)、SEQ ID NO:13(填充序列)和SEQ ID NO:2(3' ITR)。

[0126] AAV生产方法

[0127] 可以使用标准分子生物学技术将本文所述的AAV表达盒并入载体(例如,质粒或杆粒)中。本公开提供了包含本文所述的任何一种AAV表达盒的载体。载体(例如,质粒或杆粒)还可包含在AAV产生过程中使用的一个或多个遗传元件,包括例如AAV rep和cap基因,以及辅助病毒蛋白序列。

[0128] AAV表达盒和包含本文所述的AAV表达盒的载体(例如质粒)可以用于产生重组AAV载体。

[0129] 本公开提供了用于生产重组AAV载体的方法,该方法包括使AAV生产细胞(例如,HEK293细胞)与本公开的AAV表达盒或载体(例如,质粒)接触。本公开还提供了包含本文公开的任何一种AAV表达盒或载体的细胞。在一些实施方案中,该方法还包括使AAV生产细胞与一种或多种编码例如AAV rep和cap基因以及辅助病毒蛋白序列的额外质粒接触。在一些实施方案中,用于产生重组AAV载体的方法包括使AAV生产细胞(例如,昆虫细胞,诸如Sf9细胞)与至少一种包含本公开的AAV表达盒的昆虫细胞相容性载体接触。“昆虫细胞相容性载体”是有助于用核酸转化或转染昆虫细胞的任何化合物或制剂(生物的或化学的)。在一些实施方案中,昆虫细胞相容性载体是杆状病毒载体。在一些实施方案中,该方法还包括在产生AAV的条件下维持昆虫细胞。

[0130] 本公开提供了使用本文公开的任一种方法产生的重组AAV载体。产生的重组AAV载体可以具有任何血清型,例如AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV。在一些实施方案中,与天然AAV衣壳相比,产生的重组AAV载体可以包含一个或多个氨基酸修饰(例如,取代和/或缺失)。例如,重组AAV载体可以是来源于AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV和牛类AAV的修饰的AAV载体。在一些实施方案中,重组AAV载体是单链AAV(ssAAV)。在一些实施方案中,重组AAV载体是自身互补AAV(scAAV)。

[0131] 在一些实施方案中,AAV载体包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的衣壳蛋白。在一些实施方案中,AAV载体包含与野生型AAV衣壳蛋白相比具有一个或一个取代或突变的衣壳蛋白。本文公开的重组AAV载体可用于用转基因序列转导靶细胞,例如通过使重组AAV载体与靶细胞接触。

[0132] 表达和治疗方法

[0133] 本公开提供了包含本文公开的核酸、AAV表达盒、质粒、细胞或重组AAV载体中任一者的组合物。在一些实施方案中,本文公开的组合物包含至少一种药学上可接受的载剂、赋形剂和/或媒介物,例如溶剂、缓冲液、溶液、分散介质、包衣、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂。在一些实施方案中,药学上可接受的载剂、赋形剂和/或媒介物可以包括盐

水、缓冲盐水、右旋糖、水、甘油、无菌等渗水性缓冲液及其组合。在一些实施方案中,药学上可接受的载剂、赋形剂和/或媒介物包含磷酸盐缓冲盐水、无菌盐水、乳糖、蔗糖、磷酸钙、葡聚糖、琼脂、果胶、花生油、芝麻油、药物级甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁、多元醇(例如,甘油、丙二醇和液体聚乙二醇等)或其合适的混合物。在一些实施方案中,本文公开的组合物还包含少量的乳化剂或润湿剂,或pH缓冲剂。

[0134] 在一些实施方案中,本文公开的组合物还包含其他常规药物成分,诸如防腐剂,或化学稳定剂,诸如氯丁醇、山梨酸钾、山梨酸、二氧化硫、没食子酸丙酯、对羟基苯甲酸酯、乙基香草醛、甘油、苯酚、对氯苯酚或白蛋白。在一些实施方案中,本文公开的组合物还可以包含抗细菌剂和抗真菌剂,诸如对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸或硫柳汞;等渗剂,诸如糖或氯化钠,和/或延迟吸收的试剂,诸如单硬脂酸铝和明胶。

[0135] 本公开提供了通过使细胞与本文公开的核酸、AAV表达盒、质粒、重组AAV载体或组合物中的任一者接触而将雷特综合征相关基因递送到细胞中的方法。在一些实施方案中,细胞是分裂细胞,诸如细胞培养物中的培养细胞。在一些实施方案中,细胞是非分裂细胞。在一些实施方案中,将雷特综合征相关基因体外递送至细胞,例如,在体外产生雷特综合征相关多肽或用于离体基因治疗。

[0136] 在一些实施方案中,将雷特综合征相关基因递送至有需要的受试者,例如,以表达免疫原性或治疗性多肽。以这种方式,可以在受试者体内产生多肽或功能性RNA。因此,本公开提供了在有需要的受试者中表达雷特综合征相关基因的方法,该方法包括向受试者施用治疗有效量的本文公开的核酸、AAV表达盒、质粒、细胞、重组AAV载体或组合物中的任一者。本公开还提供了用于在受试者中治疗至少一种雷特综合征的症状和/或延缓其发作的方法。在一些实施方案中,该方法包括向受试者施用治疗有效量的本文公开的核酸、AAV表达盒、质粒、细胞、重组AAV载体或组合物中任一者的任一种核酸。在一些实施方案中,受试者患有雷特综合征。在一些实施方案中,受试者具有发展为雷特综合征的高风险;例如,受试者是被鉴定为携带mecp2基因突变的新生儿。在一些实施方案中,通过基因疗法靶向雷特综合征相关基因以增加其表达和/或功能。在一些实施方案中,通过基因疗法靶向雷特综合征相关基因以降低其表达和/或功能。

[0137] 本公开提供了在受试者中治疗MECP2相关疾病和/或延缓其发作的方法,该方法包括向受试者施用治疗有效量的本文公开的核酸、AAV表达盒、质粒、细胞、重组AAV载体或组合物中的任一者。如本文所用,“MECP2相关疾病”是与野生型mecp2基因相比,与mecp2基因的遗传改变(例如,一个或多个缺失、插入、复制和/或取代),和/或与野生型MECP2蛋白相比,与MECP2蛋白的表达和/或活性改变相关或由其引起的疾病。

[0138] 在一些实施方案中,MECP2相关疾病是MECP2重复综合征、MECP2相关重症新生儿脑病、PPM-X综合征或雷特综合征。MECP2重复综合征由MECP2基因重复引起,并且以智力残疾、发育迟缓和癫痫发作为特征。MECP2基因的重复导致额外的MECP2蛋白的产生和蛋白功能的增加,从而导致异常的神经元功能。MECP2相关重症新生儿脑病是由mecp2基因突变引起的,其中大多数是单碱基对插入、缺失或取代。这种病况几乎只影响男性,并且其特征在于头部尺寸变小(小头畸形)、运动障碍、呼吸问题和癫痫发作。mecp2基因的突变可以改变MECP2蛋白的结构或减少产生的蛋白质的量。PPM-X综合征是以轻度至重度智力残疾、双相情感障碍和运动异常模式为特征的疾病。所有PPM-X综合征病例的大约一半是由mecp2基因中的八个

突变之一引起的。这些突变要么引起MECP2蛋白中氨基酸的插入、缺失或取代，要么在mecp2 mRNA中产生过早终止信号。

[0139] 本公开还提供了在受试者中治疗BDNF相关疾病和/或延缓其发作的方法，该方法包括向受试者施用治疗有效量的本文公开的核酸、AAV表达盒、质粒、细胞、重组AAV载体或组合物中的任一者。如本文所用，“BDNF相关疾病”是与bdnf基因的遗传改变和/或BDNF蛋白的表达和/或活性的改变相关或由其引起的疾病。在一些实施方案中，BDNF相关疾病是认知障碍和/或压力相关障碍。在一些实施方案中，BDNF相关疾病是抑郁症、强迫症、阿尔茨海默病、亨廷顿病、痴呆、神经性厌食症、神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、双相性精神障碍、雷特综合征、重度抑郁症或化疗后认知障碍。

[0140] 在一些实施方案中，雷特综合征相关基因在受试者的神经元中表达。在一些实施方案中，雷特综合征相关基因在神经元中的表达水平高于在身体的非神经元细胞中的表达水平。在一些实施方案中，雷特综合征相关基因的表达在非神经元细胞中是不可检测的。在一些实施方案中，雷特综合征相关基因在神经元中的表达水平相对于在身体的非神经元细胞中的表达水平高至少约1.2倍（例如，约1.5倍、约2倍、约2.5倍、约3倍、约3.5倍、约4倍、约4.5倍、约5倍、约5.5倍、约6倍、约6.5倍、约7倍、约7.5倍、约8倍、约8.5倍、约9倍、约9.5倍、约10倍、约15倍、约20倍、约30倍、约40倍、约50倍、约60倍、约70倍、约80倍、约90倍或约100倍，包括位于其间的所有值和子范围）。

[0141] 在一些实施方案中，神经元是活动神经元。在一些实施方案中，雷特综合征相关基因在活动神经元中的表达水平高于在身体的静息神经元中的表达水平。在一些实施方案中，雷特综合征相关基因的表达在静息神经元中是不可检测的。在一些实施方案中，雷特综合征相关基因在活动神经元中的表达水平相对于在静息神经元中的表达水平高至少约1.2倍（例如，约1.5倍、约2倍、约2.5倍、约3倍、约3.5倍、约4倍、约4.5倍、约5倍、约5.5倍、约6倍、约6.5倍、约7倍、约7.5倍、约8倍、约8.5倍、约9倍、约9.5倍、约10倍、约15倍、约20倍、约30倍、约40倍、约50倍、约60倍、约70倍、约80倍、约90倍或约100倍，包括位于其间的所有值和子范围）。

[0142] 在一些实施方案中，雷特综合征相关基因在受试者的中枢神经系统 (CNS) 神经元中表达。在一些实施方案中，雷特综合征相关基因在CNS神经元中的表达水平相对于在身体的非神经元细胞中的表达水平高至少约1.2倍（例如，约1.5倍、约2倍、约2.5倍、约3倍、约3.5倍、约4倍、约4.5倍、约5倍、约5.5倍、约6倍、约6.5倍、约7倍、约7.5倍、约8倍、约8.5倍、约9倍、约9.5倍、约10倍、约15倍、约20倍、约30倍、约40倍、约50倍、约60倍、约70倍、约80倍、约90倍或约100倍，包括位于其间的所有值和子范围）。

[0143] 在一些实施方案中，雷特综合征相关基因在CNS神经元中的表达水平高于在其他非CNS神经元（诸如身体的外周神经系统 (PNS) 神经元）中的表达水平。在一些实施方案中，雷特综合征相关基因的表达在PNS神经元中是不可检测的。在一些实施方案中，雷特综合征相关基因在CNS神经元中的表达水平相对于其他非CNS神经元（诸如身体的外周神经系统 (PNS) 神经元）中的表达水平高至少约1.2倍（例如，约1.5倍、约2倍、约2.5倍、约3倍、约3.5倍、约4倍、约4.5倍、约5倍、约5.5倍、约6倍、约6.5倍、约7倍、约7.5倍、约8倍、约8.5倍、约9倍、约9.5倍、约10倍、约15倍、约20倍、约30倍、约40倍、约50倍、约60倍、约70倍、约80倍、约90倍或约100倍，包括位于其间的所有值和子范围）。

[0144] 待施用于受试者的重组AAV载体的剂量取决于施用方式、待治疗和/或预防的疾病或病况、个体受试者的病况、特定的病毒载体或衣壳和待递送的核酸等,并且可以以常规方式确定。用于实现治疗效果的示例性剂量是至少约 10^5 、约 10^6 、约 10^7 、约 10^8 、约 10^9 、约 10^{10} 、约 10^{11} 、约 10^{12} 、约 10^{13} 、约 10^{14} 、约 10^{15} 个转导单位,任选地约 10^8 至约 10^{13} 个转导单位的滴度。

[0145] 在特定的实施方案中,可以采用多于一次的施用(例如,两次、三次、四次或更多次施用)以在不同的间隔期(例如,每天、每周、每月、每年等)内达到所需的基因表达水平。

[0146] 示例性的施用模式包括口服、直肠、经粘膜、鼻内、吸入(例如,经由气雾剂)、经颊(例如,舌下)、阴道、鞘内、眼内、经皮、子宫内(或卵内)、肠胃外(例如,静脉内、皮下、皮内、肌肉[包括向骨骼、隔膜和/或心肌施用]、皮内、胸膜内、脑内和关节内)、局部(例如,向皮肤和粘膜表面,包括气道表面,以及经皮施用)、淋巴管内等,以及直接组织或器官注射(例如,向肝脏、骨骼肌、心肌、隔膜肌或脑)。在一些实施方案中,施用是通过注射到中枢神经系统中。在任何给定的情况下,最合适的途径将取决于所治疗和/或预防的病况的性质和严重程度,以及所使用的特定载体的性质。

[0147] 向靶组织的递送也可以通过递送包含病毒载体和/或衣壳的贮库来实现。在代表性实施方案中,将包含病毒载体和/或衣壳的贮库植入骨骼、心脏和/或隔膜肌组织中,或者该组织可以与包含病毒载体和/或衣壳的膜或其他基质接触。

[0148] 在一些实施方案中,本文公开的方法可以包括向受试者施用治疗有效量的本文公开的核酸、AAV表达盒、质粒、细胞、重组AAV载体或组合物中的任一者与靶向雷特综合征的一种或多种二级治疗的组合。在一些实施方案中,本文公开的在受试者中治疗雷特综合征的至少一种症状和/或延迟其发作的方法还可以包括施用一种或多种靶向雷特综合征的二级治疗。在一些实施方案中,二级治疗包括施用醋酸格拉替雷(glatiramer acetate)。关于使用醋酸格拉替雷的更多细节在jukic A等人,PediatrNeurol.2016年8月;61:51-7中提供,该文献的内容以全文引用的方式并入本文中。

[0149] 在一些实施方案中,二级治疗包括:施用药物以治疗癫痫发作、施用药物以治疗反流,或其组合。在一些实施方案中,二级治疗包括施用药物来治疗癫痫发作。治疗癫痫发作的药物的非限制性实例包括左乙拉西坦(levetiracetam)、丙戊酸、奥卡西平(oxcarbazepine)和拉莫三嗪(lamotrigine)。在一些实施方案中,二级治疗包括施用药物来治疗反流。治疗反流的药物的非限制性实例包括H2阻断剂,诸如西咪替丁(cimetidine)、雷尼替丁(ranitidine)、尼扎替丁(nizatidine)和法莫替丁(famotidine);质子泵抑制剂,诸如奥美拉唑(omeprazole)、艾美拉唑(esomeprazole)、兰索拉唑(lansoprazole)、雷贝拉唑(rabeprazole)、泮托拉唑(pantoprazole)和右兰索拉唑(dexlansoprazole);以及运动剂,诸如低剂量红霉素(erythromycin)、苯甲酰胺、多潘立酮(domperidone)和利那洛肽(linaclotide)。

[0150] 如本文所用,术语“组合”施用应理解是指是在受试者患有病症(诸如雷特综合征)的过程中,向受试者递送两种(或多种)不同的治疗,使得治疗对患者的效果在某一时间点重叠。在某些实施方案中,一种治疗的递送在第二种治疗的递送开始时仍在进行,因此在施用方面存在重叠。这在本文中有时被称为“同时”或“并发”递送。在其他实施方案中,一种治疗的递送在另一种治疗的递送开始之前结束,这可以被称为“顺序”递送。

[0151] 在一些实施方案中,由于组合施用,治疗更有效。例如,第二种治疗更有效;例如,

与在没有第一种治疗的情况下施用第二种治疗相比,用较少的第二种治疗观察到等同的效果,或者第二种治疗在更大程度上减轻症状,或者用第一种治疗观察到类似的情况。两种治疗的效果可以是部分相加的、完全相加的或大于相加的(协同的)。

[0152] 本说明书中引用的所有论文、出版物和专利都在此以引用的方式并入,如同每个单独的论文、出版物或专利被具体且单独地指明通过引用并入,并且通过引用并入本文以公开和描述与所引用的出版物相关的方法和/或材料。然而,本文引用的任何参考文献、文章、出版物、专利、专利公布和专利申请的提及不是,并且不应该被视为承认或以任何形式暗示它们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。

[0153] 除非上下文中另有说明,否则本文所述的各种特征可以任何组合使用。

[0154] 除非另有定义,否则本文所用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常所理解相同的含义。

[0155] 应当理解,上文的描述以及以下的实施例旨在说明而不是限制本发明的范围。本发明范围内的其他方面、优点和修改对于本发明所属领域的技术人员来说是显而易见的。

[0156] 实施例

[0157] 包括在本文中的以下实施例仅用于说明目的,并不旨在是限制性的。

[0158] 实施例1:AAV表达盒的制备

[0159] 使用标准克隆技术产生三个AAV表达盒,该表达盒包含活性依赖性hSARE-hArcMin启动子(SEQ ID NO:6)和BDNF基因(SEQ ID NO:7),侧翼为5' ITR(SEQ ID NO:1)和3' ITR(SEQ ID NO:2)。参见图1-图3,显示了三个盒的示意图。第一个盒包含BDNF短3'UTR(SEQ ID NO:8)和填充序列(SEQ ID NO:13)。第二个盒包含bGH聚A信号(SEQ ID NO:9)和填充序列(SEQ ID NO:13)。第三个盒包含BDNF长3'UTR(SEQ ID NO:10)。三个AAV表达盒分别包含SEQ ID NO:3(图1)、SEQ ID NO:4(图2)和SEQ ID NO:5(图3)的核酸序列。

[0160] 实施例2:哺乳动物细胞中重组AAV载体的制备

[0161] 将三个AAV表达盒中的每一个掺入质粒中,得到包含三个AAV表达盒中的每一个的三种质粒。使用适当的转染试剂(例如,Lipofectamine™)将三种质粒中的每一种连同编码Rep和Cap基因的Rep/Cap质粒,以及包含AAV产生所需的各种辅助序列(E4、E2a和VA)的辅助质粒一起转染到病毒生产细胞(例如,HEK293)中。在37℃下温育预定时间段后,从培养基中收集AAV颗粒,或者裂解细胞以释放AAV颗粒。然后对AAV颗粒进行纯化并滴定,并且可以储存在-80℃下以备后用。

[0162] 实施例3:昆虫细胞中重组AAV载体的制备

[0163] 将三个AAV表达盒中的每一个掺入到杆状病毒载体中,得到包含三个AAV表达盒中的每一个的三种杆状病毒载体。将昆虫细胞(例如Sf9)在悬浮培养物中与三种杆状病毒载体中的每一种以及至少一种包含编码AAV Rep和Cap蛋白的序列的额外重组杆状病毒载体共感染。在28℃下温育预定时间段后,从培养基中收集AAV颗粒,或者裂解细胞以释放AAV颗粒。然后对AAV颗粒进行纯化并滴定,并且可以储存在-80℃下以备后用。

[0164] 实施例4:hSARE-hArcMin启动子驱动活性依赖性报道基因表达

[0165] 为了测试hSARE-hArcMin启动子是否能够以依赖于神经元活性的方式驱动基因表达,进行了以下实验。

[0166] 使用标准克隆技术产生AAV表达盒,该表达盒包含驱动神经元长期表达的组成型

启动子(hSyn;人突触蛋白1基因启动子,包含SEQ ID NO:38的核酸序列)或活性依赖性启动子(hSARE-hArcMin,SEQ ID NO:6)(参见表3)。图4、图9A-图9C、图10A-图10C和图11A-图11C示出了这些盒的示意图。这些盒包含以下3'UTR序列之一:BDNF短3'UTR(SEQ ID NO:8)、bGH聚A信号(称为“bGHpA”或“bGH”,SEQ ID NO:9)、或BDNF长3'UTR(SEQ ID NO:10)。这些盒还包含与组成型或活性依赖性启动子可操作连接的报告基因。报告基因编码单体红色荧光蛋白(mScarlet,由SEQ ID NO:36的核酸序列编码),或mScarlet的不稳定形式(dmScarlet,由SEQ ID NO:35的核酸序列编码)。mScarlet的去稳定形式包含C末端PEST降解决定子信号(由核酸序列SEQ ID NO:37编码),与mScarlet相比,其促进由蛋白酶体降解驱动的更快的周转率。因此,dmScarlet荧光在细胞中的累积程度低于mScarlet荧光。

[0167] 表3:用于报道基因表达的表达盒

构建体编号	构建体名称	SEQ ID NO:	示意图	启动子	报道基因	编码区下游的序列
[0168]	1 C-mScar-短 UTR	26	图 9A	hSyn (组成型或“C”)	mScarlet-I (mScar)	短 3'UTR
	2 C-mScar-长 UTR	27	图 9B	hSyn (组成型或“C”)	mScarlet-I (mScar)	长 3'UTR
	3 C-mScar-bGH	28	图 9C	hSyn (组成型或“C”)	mScarlet-I (mScar)	bGHpA
	4 C-dmScar-短 UTR	26	图 10A	hSyn (组成型或“C”)	dmScarlet-I (dmScar)	短 3'UTR
	5 C-dmScar-长 UTR	27	图 10B	hSyn (组成型或“C”)	dmScarlet-I (dmScar)	长 3'UTR
	6 C-dmScar-bGH	28	图 10C	hSyn (组成型或“C”)	dmScarlet-I	bGHpA
[0169]					(dmScar)	
	7 AD-dmScar-短 UTR	32	图 11A	hSARE-hArcMin (活性依赖性或“AD”)	dmScarlet-I (dmScar)	短 3'UTR
	8 AD-dmScar-长 UTR	33	图 11B	hSARE-hArcMin (活性依赖性或“AD”)	dmScarlet-I (dmScar)	长 3'UTR
	9 AD-dmScar-bGH	34	图 11C	hSARE-hArcMin (活性依赖性或“AD”)	dmScarlet-I (dmScar)	bGHpA

[0170] 如本文所述制备包含表3中列出的AAV表达盒1、2、3、7、8和9的重组AAV载体。用这些重组AAV载体中的每一种转导野生型小鼠原代神经元。将转导的细胞用2 μ M钠通道抑制剂河豚毒素(TTX)处理过夜。用TTX处理抑制神经元活性。此后,将培养基完全吸出并用以下中的任一种代替:(a)仅培养基、(b)培养基+TTX、(c)培养基+150mMKCl、或(d)培养基+30 μ m荷包牡丹碱(BIC)。当KCl使神经元去极化时,BIC是 γ -氨基丁酸A型(GABAA)受体的竞争性拮抗剂,并诱导神经元的去抑制。因此,用KCl或BIC处理神经元促进神经元活性。单独放置在培养基中的对照神经元(按照上述处理(a))既不被抑制也不被激活,并且是静止的,而放置在含有TTX的培养基中的神经元(按照上述处理(b))的活性被抑制。将细胞在37 $^{\circ}$ C下进一步温育2小时。然后将细胞固定并用Hoescht 33342染色,并使用荧光显微镜成像。

[0171] 图5A显示了相对于用2 μ M TTX处理的细胞的荧光,未用TTX处理的细胞(即,其中神经元活性既未被促进也未被抑制的细胞)中的报道蛋白荧光,其被归一化为1。如图5A所示,当神经元活性既不被促进也不被抑制时,用使用hSARE-hArcMin启动子驱动dmScarlet表达

的AAV载体转导的细胞的相对荧光与用使用组成型启动子hSyn驱动mScarlet表达的AAV载体转导的细胞的相对荧光相当。

[0172] 图5B显示了相对于用2 μ M TTX处理的细胞的荧光,未用TTX处理而是用150mMKCl处理的细胞(即,神经元活性被促进的细胞)中的报道蛋白荧光,其被归一化为1。图5C显示了相对于用2 μ M TTX处理的细胞的荧光,未用TTX处理而是用30 μ M荷包牡丹碱(BIC)处理的细胞(即,其中神经元活性被促进的细胞)中的报道蛋白荧光,其被归一化为1。

[0173] 如图5B和图5C所示,当使用KCl或BIC处理来促进神经元活性时,用使用活性依赖性启动子hSARE-hArcMin驱动dmScarlet表达的AAV载体转导的细胞的相对荧光高于用使用组成型启动子(hSyn)驱动mScarlet表达的AAV载体转导的细胞的相对荧光。

[0174] 这些结果表明hSARE-hArcMin启动子能够以依赖于神经元活性的方式驱动基因表达。此外,包含bGH或长3'UTR的dmScarlet构建体类似地由hSARE-hArcMin启动子表达,如图5B和图5C所示。

[0175] 图6A-B进一步示出了hSARE-hArcMin启动子的活性依赖性表达。如图6A所示,组成型启动子hSyn驱动高水平的mScarlet表达,而与TTX的存在无关。引人注目的是,图6B显示,当细胞不用TTX处理,而是用30 μ M BIC处理以促进神经元活性时,hSARE-hArcMin启动子驱动高水平的报道基因表达。值得注意的是,图6B显示,与包含hSARE-hArcMin启动子与bGHpA的组合的AAV表达构建体相比,包含hSARE-hArcMin启动子与长3'UTR的组合的AAV表达构建体在TTX的存在下导致约5倍高水平的dmScarlet表达。这些结果进一步证明了hSARE-hArcMin启动子的神经元活性依赖性诱导。

[0176] 为了确定表达的荧光蛋白的亚细胞定位和水平,使用显微镜观察在组成型hSyn启动子或活性依赖性hSARE-hArcMin启动子下表达报道基因的神经元。

[0177] 如图7所示,使用组成型启动子hSyn表达mScarlet的细胞是荧光的,然而,荧光水平根据表达盒的3'UTR区或聚腺苷酸化信号而变化。例如,包含长3'UTR的表达盒导致较高水平的mScarlet基因表达,随后是包含短3'UTR的表达盒。包含短3'UTR的表达盒导致相当较低水平的mScarlet基因表达。相比之下,在活性依赖性hSARE-hArcMin启动子下,报道蛋白的表达不根据所用3'UTR的类型而变化。具体地讲,图5和图6显示包含hSARE-hArcMin启动子与bGHpA或长3'UTR的组合的AAV表达构建体在神经元活性的存在下导致相似水平的dmScarlet表达。

[0178] 此外,图7显示由组成型启动子表达的mScarlet的亚细胞定位根据表达盒的3'UTR区或聚腺苷酸化信号而变化。虽然包含长3'UTR和bGHpA的表达盒导致mScarlet在细胞体和神经突中表达,但包含短3'UTR的表达盒导致mScarlet仅在细胞体中表达。

[0179] 如图8和图12所示,包含hSARE-hArcMin启动子与bGHpA(图8)或长3'UTR(图12)的组合的AAV表达构建体在神经元活性的存在下导致相似水平的dmScarlet表达。此外,使用活性依赖性hSARE-hArcMin启动子表达dmScarlet的细胞在神经元激活时(在不存在TTX和存在BIC的情况下)在细胞体和神经突中显示荧光信号。

[0180] 总之,上述结果表明,与受抑制或静止的神经元相比,合成的hSARE-hArcMin启动子在活动神经元中促进更高水平的基因表达。因此,合成的hSARE-hArcMin启动子可用于驱动任何靶基因(诸如本文所述的雷特综合征相关基因)的神经元活性依赖性表达。

[0181] 实施例5:包含BDNF的AAV表达盒的表征

[0182] 通过注射到表现出类似于雷特综合征的症状的mecp2突变小鼠中来施用实施例2或3中产生的AAV颗粒,以测试使用活性依赖性启动子hSARE-hArcMin表达BDNF是否降低一种或多种症状的严重程度。还测试了活性依赖性BDNF表达延迟新生mecp2突变小鼠中雷特综合征症状发作的能力。

[0183] 上述内容是对本发明的说明,而不应被解释为对本发明的限制。

[0184] 编号的实施方案

[0185] 以下实施例列表仅出于说明目的而包括在本文中,并且不旨在是全面的或限制性的。要求保护的主体明确不限于以下实施方案。

[0186] 实施方案1.一种核酸,所述核酸包含腺相关病毒(AAV)表达盒,其中所述AAV表达盒从5'至3'包含:

[0187] 5'反向末端重复序列(ITR);

[0188] 合成的活性依赖性启动子;

[0189] 雷特综合征相关基因;以及

[0190] 3' ITR。

[0191] 实施方案2.如实施方案1所述的核酸,其中所述启动子驱动所述雷特综合征相关基因的表达。

[0192] 实施方案3.如实施方案1或2所述的核酸,其中所述启动子是MECP2非依赖性启动子。

[0193] 实施方案4.如实施方案1至3中任一项所述的核酸,其中所述启动子包含源自神经元即早期基因的启动子的核酸序列。

[0194] 实施方案5.如实施方案4所述的核酸,其中所述神经元即早期基因选自由以下组成的组:Arc基因、c-fos基因和egr-1基因。

[0195] 实施方案6.如实施方案1至5中任一项所述的核酸,其中所述启动子包含最小Arc基因启动子(ArcMin)。

[0196] 实施方案7.如实施方案6所述的核酸,其中所述ArcMin是人ArcMin(hArcMin)。

[0197] 实施方案8.如实施方案7所述的核酸,其中所述hArcMin包含SEQ ID NO:12的核酸序列,或与其至少90%一致的序列。

[0198] 实施方案9.如实施方案1至8中任一项所述的核酸,其中所述启动子包含环AMP应答元件(CRE)、血清应答元件(SRE)、突触活动应答元件(SARE)、MEF2应答元件或其组合。

[0199] 实施方案10.如实施方案9所述的核酸,其中所述启动子包含突触活动应答元件(SARE)。

[0200] 实施方案11.如实施方案10所述的核酸,其中所述突触活动应答元件(SARE)是人突触活动应答元件(hSARE)。

[0201] 实施方案12.如实施方案11所述的核酸,其中所述hSARE包含SEQ ID NO:11的核酸序列,或与其至少90%一致的序列。

[0202] 实施方案13.如实施方案1至12中任一项所述的核酸,其中所述启动子包含人ArcMin(hArcMin)和至少一个hSARE。

[0203] 实施方案14.如实施方案13所述的核酸,其中所述启动子包含hArcMin和一个hSARE。

[0204] 实施方案15.如实施方案14所述的核酸,其中所述启动子包含SEQ ID NO:6的核酸序列,或与其至少90%一致的序列。

[0205] 实施方案16.如实施方案13所述的核酸,其中所述启动子包含hArcMin和五个hSARE。

[0206] 实施方案17.如实施方案16所述的核酸,其中所述启动子包含SEQ ID NO:16的核酸序列,或与其至少90%一致的序列。

[0207] 实施方案18.如实施方案1至17中任一项所述的核酸,其中所述启动子结合神经元活性依赖性转录因子。

[0208] 实施方案19.如实施方案18所述的核酸,其中所述神经元活性依赖性转录因子是cAMP应答元件结合蛋白(CREB)、肌细胞增强因子2(MEF2)、血清应答因子(SRF)或Elk-1。

[0209] 实施方案20.如实施方案1至19中任一项所述的核酸,其中所述雷特综合征相关基因编码脑源性神经营养因子(BDNF)、胰岛素样生长因子1(IGF1)、甲基-CpG-结合蛋白2(MECP2)、亨廷顿蛋白、亨廷顿相关蛋白1、正小齿同源框2(OTX-2)、含FXFD结构域的离子转运调节因子1(FXFD1)、神经毒素-2- α (NRXN2)或蛋白激酶C γ (PRKCG)。

[0210] 实施方案21.如实施方案20所述的核酸,其中所述雷特综合征相关基因编码脑源性神经营养因子(BDNF)。

[0211] 实施方案22.如实施方案21所述的核酸,其中所述BDNF是人BDNF。

[0212] 实施方案23.如实施方案22所述的核酸,其中所述BDNF由SEQ ID NO:7的核酸序列或与其至少90%一致的序列编码。

[0213] 实施方案24.如实施方案1至23中任一项所述的核酸,其中所述5' ITR和所述3' ITR中的至少一者的长度为约110至约160个核苷酸。

[0214] 实施方案25.如实施方案1至24中任一项所述的核酸,其中所述5' ITR与所述3' ITR的长度相同。

[0215] 实施方案26.如实施方案1至24中任一项所述的核酸,其中所述5' ITR和所述3' ITR具有不同的长度。

[0216] 实施方案27.如实施方案1至26中任一项所述的核酸,其中所述5' ITR和所述3' ITR中的至少一者从AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的基因组中分离或来源于所述基因组。

[0217] 实施方案28.如实施方案1至23中任一项所述的核酸,其中所述5' ITR包含SEQ ID NO:1的序列。

[0218] 实施方案29.如实施方案1至23中任一项所述的核酸,其中所述3' ITR包含SEQ ID NO:2的序列。

[0219] 实施方案30.如实施方案1至29中任一项所述的核酸,其中所述AAV盒包含脑源性神经营养因子(BDNF)短3' UTR,或BDNF长3' UTR。

[0220] 实施方案31.如实施方案30所述的核酸,其中所述BDNF短3' UTR或所述BDNF长3' UTR位于雷特综合征相关基因和所述3' ITR之间。

[0221] 实施方案32.如实施方案30或实施方案31所述的核酸,其中所述AAV盒包含BDNF短3' UTR。

[0222] 实施方案33.如实施方案32所述的核酸,其中所述BDNF短3'UTR包含SEQ ID NO:8的核酸序列,或与其至少90%一致的序列。

[0223] 实施方案34.如实施方案30或实施方案31所述的核酸,其中所述AAV盒包含BDNF长3'UTR。

[0224] 实施方案35.如实施方案34所述的核酸,其中所述BDNF长3'UTR包含SEQ ID NO:10的核酸序列,或与其至少90%一致的序列。

[0225] 实施方案36.如实施方案1至35中任一项所述的核酸,其中所述AAV盒包含聚腺苷酸化信号。

[0226] 实施方案37.如实施方案36所述的核酸,其中所述聚腺苷酸化信号是从以下基因中的一者或多者分离或来源于以下基因中的一者或多者的聚腺苷酸化信号:猿猴病毒40(SV40)、rBG、 α -珠蛋白、 β -珠蛋白、人胶原、人生长激素(hGH)、多瘤病毒、人生长激素(hGH)或牛生长激素(bGH)。

[0227] 实施方案38.如实施方案36所述的核酸,其中所述AAV盒包含bGH聚腺苷酸化信号。

[0228] 实施方案39.如实施方案38所述的核酸,其中所述bGH聚腺苷酸化信号包含SEQ ID NO:9的核酸序列,或与其至少90%一致的序列。

[0229] 实施方案40.如实施方案1至39中任一项所述的核酸,其中所述AAV盒包含至少一个填充序列。

[0230] 实施方案41.如实施方案40所述的核酸,其中所述至少一个填充序列包含SEQ ID NO:13的核酸序列,或与其至少90%一致的序列。

[0231] 实施方案42.如实施方案1至41中任一项所述的核酸,其中所述AAV表达盒包含Kozak序列,其中所述Kozak序列与所述雷特综合征相关基因的起始密码子重叠。

[0232] 实施方案43.如实施方案42所述的核酸,其中所述Kozak序列包含SEQ ID NO:14的核酸序列,或与其至少90%一致的序列;或SEQ ID NO:15的核酸序列,或与其至少90%一致的序列。

[0233] 实施方案44.如实施方案1至43中任一项所述的核酸,其中所述AAV表达盒包含核酸序列SEQ ID NO:3,或与其至少90%一致的序列;SEQ ID NO:4或与其至少90%一致的序列;或SEQ ID NO:5或与其至少90%一致的序列。

[0234] 实施方案45.一种质粒,所述质粒包含如实施方案1至44中任一项所述的核酸。

[0235] 实施方案46.一种细胞,所述细胞包含如实施方案1至44中任一项所述的核酸或如实施方案45所述的质粒。

[0236] 实施方案47.一种产生重组AAV载体的方法,所述方法包括使AAV生产细胞与如实施方案1至44中任一项所述的核酸或如实施方案45所述的质粒接触。

[0237] 实施方案48.一种重组AAV载体,所述重组AAV载体通过如实施方案47所述的方法产生。

[0238] 实施方案49.如实施方案48所述的重组AAV载体,其中所述载体具有选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV和牛类AAV的血清型。

[0239] 实施方案50.如实施方案48或实施方案49所述的重组AAV载体,其中所述重组AAV载体是单链AAV(ssAAV)。

[0240] 实施方案51.如实施方案48至50中任一项所述的重组AAV载体,其中所述重组AAV载体是自身互补AAV(scAAV)。

[0241] 实施方案52.如实施方案48至51中任一项所述的重组AAV载体,其中所述AAV载体包含AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAVrh32.33、AAVrh74、鸟类AAV或牛类AAV的衣壳蛋白。

[0242] 实施方案53.如实施方案48至52中任一项所述的重组AAV载体,其中所述AAV载体包含与野生型AAV衣壳蛋白相比具有一个或一个取代或突变的衣壳蛋白。

[0243] 实施方案54.一种组合物,所述组合物包含:(a)如实施方案1至44中任一项所述的核酸、如实施方案45所述的质粒、如实施方案46所述的细胞或如实施方案48至53中任一项所述的重组AAV载体;以及(b)药学上可接受的载剂。

[0244] 实施方案55.一种用于在有需要的受试者中表达雷特综合征相关基因的方法,所述方法包括向所述受试者施用治疗有效量的如实施方案1至44中任一项所述的核酸、如实施方案45所述的质粒、如实施方案46所述的细胞、如实施方案48至53中任一项所述的重组AAV载体或如实施方案54所述的组合物。

[0245] 实施方案56.如实施方案55所述的方法,其中所述受试者患有雷特综合征。

[0246] 实施方案57.一种用于在受试者中治疗雷特综合征或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如实施方案1至44中任一项所述的核酸、如实施方案45所述的质粒、如实施方案46所述的细胞、如实施方案48至53中任一项所述的重组AAV载体或如实施方案54所述的组合物。

[0247] 实施方案58.一种用于在有需要的受试者中表达脑源性神经营养因子(BDNF)的方法,所述方法包括向所述受试者施用治疗有效量的如实施方案1至44中任一项所述的核酸、如实施方案45所述的质粒、如实施方案46所述的细胞、如实施方案48至53中任一项所述的重组AAV载体或如实施方案54所述的组合物。

[0248] 实施方案59.如实施方案58所述的方法,其中所述受试者患有认知障碍或压力相关障碍。

[0249] 实施方案60.如实施方案58所述的方法,其中所述受试者患有抑郁症、强迫症、阿尔茨海默病、亨廷顿病和痴呆、神经性厌食症和神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、肥胖症、雷特综合征或化疗后认知障碍。

[0250] 实施方案61.一种在受试者中治疗BDNF相关疾病或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如实施方案1至44中任一项所述的核酸、如实施方案45所述的质粒、如实施方案46所述的细胞、如实施方案48至53中任一项所述的重组AAV载体或如实施方案54所述的组合物。

[0251] 实施方案62.如实施方案61所述的方法,其中所述BDNF相关疾病是认知障碍和/或压力相关障碍。

[0252] 实施方案63.如实施方案61所述的方法,其中所述BDNF相关疾病为抑郁症、强迫症、阿尔茨海默病、亨廷顿病和痴呆、神经性厌食症、神经性贪食症、精神分裂症、癫痫、创伤后应激障碍、双相性精神障碍、雷特综合征、重度抑郁症或化疗后认知障碍。

[0253] 实施方案64.一种在受试者中治疗MECP2相关疾病或延缓其发作的方法,所述方法包括向所述受试者施用治疗有效量的如实施方案1至44中任一项所述的核酸、如实施方案

45所述的质粒、如实施方案46所述的细胞、如实施方案48至53中任一项所述的重组AAV载体或如实施方案54所述的组合物。

[0254] 实施方案65.如实施方案64所述的方法,其中所述MECP2相关疾病为MECP2重复综合征、MECP2相关重症新生儿脑病、PPM-X综合征或雷特综合征。

[0255] 实施方案66.如实施方案55至65中任一项所述的方法,其中所述受试者为人受试者。

[0256] 实施方案67.如实施方案55至66中任一项所述的方法,其中所述核酸、所述质粒、所述细胞、所述重组AAV载体或组合物通过注射到中枢神经系统中来施用。

[0257] 实施方案68.如实施方案55至67中任一项所述的方法,其中所述雷特综合征相关基因在所述受试者的神经元中表达。

[0258] 实施方案69.如实施方案68所述的方法,其中所述神经元是活动神经。

序列表

<110> 萨雷普塔治疗公司

<120> 用于治疗雷特综合征的腺相关病毒载体

<130> STRD-020/01W0 331843-2220

<150> US 63/067,668

<151> 2020-08-19

<160> 38

<170> PatentIn版本3.5

<210> 1

<211> 141

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 5' ITR

<400> 1

```
cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc t 141
```

<210> 2

<211> 166

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 3' ITR

<400> 2

```
tcgagccatg ggcgcgccat cgatgaggaa cccctagtga tggagttggc cactccctct 60
ctgcgcgctc gctcgtcac tgaggccggg cgaccaaagg tcgcccgcac cccgggcttt 120
gcccgggcgg cctcagttag cgagcgagcg cgcagctgcc tgcagg 166
```

<210> 3

<211> 4200

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hSARE-ArcMin-BDNF-短3' UTR

<400> 3

```
cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc ttgcgtcgac aggcctccta ggcggaccgc ttgcatgcac 180
```

gcgtgggtgg tgcagagtct gcatgctga ggagctcctg ggcgcgtcac agccgcgcta 240
 ttctcagcgt ctctcctttt atggctccgg aagtgagctg gggttgctgg cagccgtccc 300
 tcgggtccccg caggctgaat gagccccgcc gtccgggggtg ggaggtacgg agcggaggga 360
 ttagtcatcc tggccccgcc atgtgtgcag ggggtgggac ctcccgctcc ttgccggccg 420
 ggagccgagt gggcgagggg cggggcctgc aggggcgcgg gcggggcgct ggcggggagc 480
 ctgcggctgg gccaatgaga aaccgggctc ggcgagccgg ccgcccacgg gcctcgctgg 540
 ctgcataaag agccggcggc caggactcag cgcagagctc gggcgcgggc tcctccctcc 600
 gcagcagccg agccggacct gcctccccgg gcgtgctccg ccggccccgc cgccggccccg 660
 cagcgacaga caggcgctcc ccgcagctcc gcacgggacc caggccgccc gacccagcg 720
 ccggaccacc ctctgtccgc ccgaggagt ttgccgctg ccggagcacc tgcgcacaga 780
 ccggtcgcca ccatgacct ctttttctt actatggtta tttcatactt tggttgcatg 840
 aaggctgccc ccatgaaaga agcaaacatc cgaggacaag gtggcttggc ctaccaggt 900
 gtgcggaccc atgggactct ggagagcgtg aatgggccc aggcaggttc aagaggcttg 960
 acatcattgg ctgacacttt cgaacacgtg atagaagagc tgttgatga ggaccagaaa 1020
 gttcgcccca atgaagaaaa caataaggac gcagacttgt acacgtccag ggtgatgctc 1080
 agtagtcaag tgcctttgga gcctcctctt ctctttctgc tggaggaata caaaaattac 1140
 ctagatgctg caaacatgtc catgagggtc cggcgccact ctgaccctgc ccgccaggg 1200
 gagctgagcg tgtgtgacag tattagttag tgggtaacgg cggcagacaa aaagactgca 1260
 gtggacatgt cgggcgggac ggtcacagtc cttgaaaagg tccctgtatc aaaaggccaa 1320
 ctgaagcaat acttctacga gaccaagtgc aatcccatgg gttacacaaa agaaggctgc 1380
 aggggcatag acaaaaggca ttggaactcc cagtgcgaa ctaccagtc gtacgtgcgg 1440
 gcccttacca tggatagcaa aaagagaatt ggctggcgat tcataaggat agacacttct 1500
 tgtgtatgta cattgacct taaaagggga agatagggtta cctggattta tgtgtatag 1560
 attagattat attgagacaa aaattatcta tttgtatata tacataacag ggtaaattat 1620
 tcagttaaga aaaaaataat tttatgaact gcatgtataa atgaagtta tacagtacag 1680
 tggttctaca atctatttat tggacatgtc catgaccaga agggaaacag tcatttgcgc 1740
 acaacttaaa aagtctgcat tacattcctt gataatgttg tggtttgttg ccgttgccaa 1800
 gaactgaaaa cataaaaagt taataaaat aataaattgc atgctgcttt aattgtgaat 1860
 tgataataag cgcgccatcg atgtgaagt ggtaaccttt atttcccttc tttttctctt 1920
 tagctcggct tattccaggg gtgtgtttcg tcgagatgca cacaagagtg aggttgctca 1980
 tcggtttaaa gatattgggag aagaaaattt caaagccttg gtgttgattg cctttgctca 2040
 gtatcttcag cagtgtccat ttgaagatca tgtaaaatta gtgaatgaag taactgaatt 2100
 tgcaaaaaca tgtgttgctg atgagtcagc tgaaaattgt gacaaatcac ttcataacct 2160
 ttttgagac aaattatgca cagttgcaac tcttcgtgaa acctatggtg aaatggctga 2220
 ctgctgtgca aaacaagaac ctgagagaaa tgaatgcttc ttgcaacaca aagatgacaa 2280
 cccaaacctc ccccgattgg tgagaccaga ggttgatgtg atgtgactg cttttcatga 2340
 caatgaagag acctttttga aaaaatactt atacgaaatt gccagaagac atccttactt 2400
 ttatgccccg gaactccttt tctttgctaa aaggtataaa gctgctttta cagaatgttg 2460
 ccaagctgct gataaagctg cctgcctgtt gccaaagctc gatgaacttc gggatgaagg 2520

gaaggcttctg tctgccaaac agagactcaa gtgtgccagt ctccaaaaat ttggagaaag 2580
 agcttttcaaa gcatgggcag tagctcgcct gagccagaga tttcccaaag ctgagtttgc 2640
 agaagtttcc aagttagtga cagatcttac caaagtccac acggaatgct gccatggaga 2700
 tctgcttgaa tgtgctgatg acagggcgga ccttgccaag tatatctgtg aaaatcaaga 2760
 ttcgatctcc agtaaaactga aggaatgctg tgaaaaacct ctgttggaag aatcccactg 2820
 cattgccgaa gtggaaaatg atgagatgcc tgctgacttg ctttcattag ctgctgattt 2880
 tgttgaaagt aaggatgttt gcaaaaacta tgctgaggca aaggatgtct tcctgggcat 2940
 gtttttgtat gaatatgcaa gaaggcatcc tgattactct gtcgtgctgc tgctgagact 3000
 tgccaagacc tatgaaacca ctctagagaa gtgctgtgcc gctgcagatc ctcatgaatg 3060
 ctatgccaaa gtgttcgatg aatttaaacc tcttgaggaa gagcctcaga atttaataca 3120
 acaaaattgt gagctttttg agcagcttgg agagtacaaa ttccagaatg cgctattagt 3180
 tcgttacacc aagaaagtac cccaagtgtc aactccaact cttgtagagg tctcaagaaa 3240
 cctaggaaaa gtgggcagca aatgttgtaa acatcctgaa gcaaaaagaa tgccctgtgc 3300
 agaagactat ctatccgtgg tcctgaacca gttatgtgtg ttgcatgaga aaacgccagt 3360
 aagtgcagaa gtcaccaaat gctgcacaga atccttggtg aacaggcgac catgtcgtga 3420
 aacctatggt gaaatggctg actgctgtgc aaaacaagaa cctgagagaa atgaatgctt 3480
 cttgcaacac aaagatgaca acccaaacct ccccgattg gtgagaccag aggttgatgt 3540
 gatgtgcact gcttttcatg acaatgaaga gacctttttg aaaaaatact tatacgaaat 3600
 tgccagaaga catccttact tttatgcccc ggaactcctt ttctttgcta aaaggtataa 3660
 agctgctttt acagaatgtt gccaagctgc tgataaagct gcctgcctgt tgccaaagct 3720
 cgatgaactt cgggatgaag ggaaggcttc gtctgccaaa cagagactca agtgtgccag 3780
 tctccaaaaa tttggagaaa gagctttcaa agcatgggca gtagctcgcc tgagccagac 3840
 ttttcagctc tggaagtcga tgaaacatac gttcccaaag agtttaatgc tgaaacattc 3900
 accttccatg cagatatatg cacactttct gagaaggaga gacaaatcaa gaaacaaact 3960
 gcacttggtg agctcgtgaa acacaagccc aaggcaacaa aagagcaact gaaagctggt 4020
 atggatgatt tcgctcgagc catgggcgcg ccatcgatga ggaacccta gtgatggagt 4080
 tggccactcc ctctctgcgc gctcgtcgc tcactgaggc cgggcgacca aaggtcgccc 4140
 gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgagc tgccctgcagg 4200
 <210> 4
 <211> 4098
 <212> DNA
 <213> 人工序列(Artificial Sequence)
 <220>
 <223> hSARE-ArcMin-BDNF-bGHpA-hAlb
 <400> 4
 cctgcaggca gctgcgcgct cgctcgtcga ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggcggaccgc ttgcatgcac 180
 gcgtgggtgg tgcagagtct gcatgcgtga ggagctcctg ggcgcgtcac agccgcgcta 240

```

ttctcagcgt ctctcctttt atggctccgg aagtgagctg gggttgctgg cagccgtccc 300
tcgggtccccg caggctgaat gagccccgcc gtccgggggtg ggaggtacgg agcggaggga 360
ttagtcatcc tggccccgcc atgtgtgcag ggggtgggac ctcccgtccc ttgccgcccg 420
ggagccgagt gggcgagggg cggggcctgc agggcgcgcg gcggggcgct ggcggggagc 480
ctgcggctgg gccaatgaga aaccgggctc ggcgagccgg ccgcccacgg gcctcgctgg 540
ctgcataaag agccggcggc caggactcag cgcagagctc gggcgcgggc tcctccctcc 600
gcagcagccg agccggacct gcctccccgg gcgtgctccg ccggccccgc cgccggccccg 660
cagcgacaga caggcgctcc ccgcagctcc gcacgggacc caggccgccc gaccccagcg 720
ccggaccacc ctctgtccgc cccgaggagt ttgccgctg ccggagcacc tgcgcacaga 780
ccggtcgcca ccatgaccat ctttttctt actatggta tttcatactt tggttgcatg 840
aaggctgccc ccatgaaaga agcaaacatc cgaggacaag gtggcttggc ctaccaggt 900
gtgcggaccc atgggactct ggagagcgtg aatgggccc aggcaggttc aagaggcttg 960
acatcattgg ctgacacttt cgaacacgtg atagaagagc tgttgatga ggaccagaaa 1020
gttcggccca atgaagaaaa caataaggac gcagacttgt acacgtccag ggtgatgctc 1080
agtagtcaag tgcttttga gcctcctctt ctctttctgc tggaggaata caaaaattac 1140
ctagatgctg caaacatgtc catgagggtc cggcgccact ctgaccctgc ccgccaggg 1200
gagctgagcg tgtgtgacag tattagttag tgggtaacgg cggcagacaa aaagactgca 1260
gtggacatgt cgggcgggac ggtcacagtc cttgaaaagg tccctgtatc aaaaggccaa 1320
ctgaagcaat acttctacga gaccaagtgc aatcccatgg gttacacaaa agaaggctgc 1380
aggggcatag acaaaaggca ttggaactcc cagtgccgaa ctaccagtc gtacgtgcgg 1440
gcccttacca tggatagcaa aaagagaatt ggctggcgat tcataaggat agacacttct 1500
tgtgtatgta cattgaccat taaaagggga agatagggtc ccctgtgcct tctagttgcc 1560
agccatctgt tgtttgcccc tccccgtgc cttccttgac cctggaaggt gccactccca 1620
ctgtcctttc ctaataaaat gaggaaattg catcgcatg tctgagtagg tgtcattcta 1680
ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac aatagcaggc 1740
atgctgggga tgcggtgggc tctatgggcg cgccatcgat gtgaagtggg taacctttat 1800
ttcccttctt tttctcttta gtcgggctta ttccagggt gtgtttcgtc gagatgcaca 1860
caagagttag gttgctcatc ggtttaaaga tttgggagaa gaaaatttca aagccttggg 1920
gttgattgcc tttgctcagt atcttcagca gtgtccattt gaagatcatg taaaattagt 1980
gaatgaagta actgaatttg caaaaacatg tgttgctgat gagttagctg aaaatttgta 2040
caaatcactt catacccttt ttggagacaa attatgcaca gttgcaactc ttcgtgaaac 2100
ctatggtgaa atggctgact gctgtgcaaa acaagaacct gagagaaatg aatgcttctt 2160
gcaacacaaa gatgacaacc caaacctccc ccgattgggt agaccagagg ttgatgtgat 2220
gtgcactgct tttcatgaca atgaagagac ctttttgaaa aaatacttat acgaaattgc 2280
cagaagacat ccttactttt atgccccgga actccttttc tttgctaaaa ggtataaagc 2340
tgctttttaca gaatgttgcc aagctgctga taaagctgcc tgctgttgcc caaagctcga 2400
tgaacttcgg gatgaaggga aggtttcgtc tgccaaacag agactcaagt gtgccagtct 2460
ccaaaaattt ggagaaagag ctttcaaagc atgggcagta gctcgcctga gccagagatt 2520
tcccaaagct gagtttgcag aagtttccaa gttagtgcag gatcttacca aagtccacac 2580

```

ggaatgctgc catggagatc tgcttgaatg tgctgatgac agggcggacc ttgccaaagta 2640
 tatctgtgaa aatcaagatt cgatctccag taaactgaag gaatgctgtg aaaaacctct 2700
 gttggaaaaa tcccactgca ttgccgaagt ggaaaatgat gagatgcctg ctgacttgcc 2760
 ttcattagct gctgattttg ttgaaagtaa ggatgtttgc aaaaactatg ctgaggcaaa 2820
 ggatgtcttc ctgggcatgt ttttgtatga atatgcaaga aggcacctg attactctgt 2880
 cgtgctgctg ctgagacttg ccaagacctg tgaaccact ctagagaagt gctgtgccgc 2940
 tgcagatcct catgaatgct atgccaaagt gttcgatgaa tttaaacctc ttgtggaaga 3000
 gcctcagaat ttaatcaaac aaaattgtga gctttttgag cagcttggag agtacaaatt 3060
 ccagaatgcg ctattagtgc gttacaccaa gaaagtaccc caagtgtcaa ctccaactct 3120
 tgtagaggtc tcaagaaacc taggaaaagt gggcagcaaa tgttgtaaac atcctgaagc 3180
 aaaaagaatg ccctgtgcag aagactatct atccgtggtc ctgaaccagt tatgtgtgtt 3240
 gcatgagaaa acgccagtaa gtgacagagt caccaaatgc tgcacagaat ccttggtgaa 3300
 caggcgacca tgtcgtgaaa cctatggtga aatggctgac tgctgtgcaa aacaagaacc 3360
 tgagagaaat gaatgcttct tgcaacacaa agatgacaac ccaaacctcc cccgattggt 3420
 gagaccagag gttgatgtga tgtgcaactgc ttttcatgac aatgaagaga cttttttgaa 3480
 aaaatactta tacgaaattg ccagaagaca tccttacttt tatgccccgg aactcctttt 3540
 ctttgctaaa aggtataaag ctgctttttac agaattgtgc caagctgctg ataaagctgc 3600
 ctgcctgttg ccaaagctcg atgaacttcg ggatgaaggg aaggcttcgt ctgccaaaca 3660
 gagactcaag tgtgccagtc tccaaaaatt tggagaaaga gctttcaaag catgggcagt 3720
 agctcgctg agccagactt ttcagctctg gaagtcgatg aaacatacgt tcccaaagag 3780
 tttaatgctg aaacattcac cttccatgca gatatatgca cactttctga gaaggagaga 3840
 caaatcaaga acaaaactgc acttggtgag ctctgaaac acaagcccaa ggcaacaaaa 3900
 gagcaactga aagctgttat ggatgatttc gctcgagcca tgggcgcgcc atcgatgagg 3960
 aacccttagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 4020
 ggcgaccaa ggctgcccga cgcccgggct ttgcccgggc ggcctcagt agcgagcgag 4080
 cgcgagctg cctgcagg 4098

<210> 5

<211> 4630

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hSARE-ArcMin-BDNF-长3' UTR

<400> 5

cctgcaggca gctgcgcgt cgtcgtctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggcggaccgc ttgcatgcag 180
 ggtggtgcag agtctgcatg cgtgaggagc tcctgggcgc gtcacagccg cgctattctc 240
 agcgtctctc cttttatggc tccggaagtg agctgggggtt gctggcagcc gtccctcggt 300
 ccccgagggc tgaatgagcc ccgccgtccg ggggtgggagg tacggagcgg agggattagt 360

catcctggcc ccgccatgtg tgcaggggggt gggacctccc gtcccttgcc gcccgggagc 420
cgagtgggcg cagggcgggg cctgcagggg cgcgggcggg gcgctggcgg ggagcctgcg 480
gctgggccaa tgagaaaccg ggctcggcga gccggccgcc cacgggcctc gctggctgca 540
taaagagccg gcggccagga ctacagcgcag agctcgggcg cggcgtcctc cctccgcagc 600
agccgagccg gacctgcctc cccgggctgt ctccgccggc cccgccgccg gcccgagcgc 660
acagacaggc gctccccgca gctccgcacg ggaccaggc cgcgggaccc cagcgccgga 720
ccaccctctg tccgccccga ggagtttgcc gcctgccgga gcacctgcgc acagaccggt 780
cgccaccatg accatccttt tccttactat ggttatttca tacttttggtt gcatgaaggc 840
tgcccccatg aaagaagcaa acatccgagg acaaggtggc ttggcctacc caggtgtgcg 900
gacccatggg actctggaga gcgtgaatgg gcccaaggca ggttcaagag gcttgacatc 960
attggctgac actttcgaac acgtgataga agagctgttg gatgaggacc agaaagtctg 1020
gccaatgaa gaaaacaata aggacgcaga cttgtacacg tccagggtga tgctcagtag 1080
tcaagtgcct ttggagcctc ctcttctctt tctgctggag gaatacaaaa attacctaga 1140
tgctgcaaac atgtccatga ggggccggcg ccactctgac cctgcccgcc gaggggagct 1200
gagcgtgtgt gacagtatta gtgagtgggt aacggcgcca gacaaaaaga ctgcagtgga 1260
catgtcgggc gggacgggtc cagtccttga aaaggtccct gtatcaaaag gccaaactgaa 1320
gcaatacttc tacgagacca agtgcaatcc catgggttac acaaaagaag gctgcagggg 1380
catagacaaa aggcatgtga actcccagtg ccgaactacc cagtcgtacg tgcgggccct 1440
taccatggat agcaaaaaaga gaattggctg gcgattcata aggatagaca cttcttgtgt 1500
atgtacattg accattaaaa ggggaagata gggtagcttg atttatgttg tatagattag 1560
attatatgtg gacaaaaatt atctatttgt atatatacat aacagggtaa attattcagt 1620
taagaaaaaa ataattttat gaactgcatg tataaatgaa gtttatacag tacagtgggt 1680
ctacaatcta tttattggac atgtccatga ccagaaggga aacagtcatt tgcgcacaac 1740
ttaaaaagtc tgcattacat tccttgataa tgttgtgggt tgttgccgtt gccagaact 1800
gaaaacataa aaagttaaaa aaaataataa attgcatgct gctttaattg tgaattgata 1860
ataaactgtc ctctttcaga aaacagaaaa aaacacacac acacacaaca aaaatttgaa 1920
ccaaaacatt ccgtttacat tttagacagt aagtatcttc gttcttggtt gtactatatc 1980
tgttttactg cttttaactt ctgtagcgt tggaattaaa acaatgtcaa ggtgctgttg 2040
tcattgcttt actggcttag gggatggggg atggggggta tatttttggt tgttttgtgt 2100
ttttttttcg tttgtttgtt ttgtttttta gttccacag ggagtagaga tggggaaaga 2160
attcctacaa tatatatctt ggctgataaa agatacattt gtatgttggt aagatgtttg 2220
caatatcgat cagatgacta gaaagtgaat aaaaattaag gcaactgaac aaaaaaatgc 2280
tcacactcca catcccgtga tgcacctccc aggccccgt cattctttgg gcgttggtca 2340
gagtaagctg cttttgacgg aaggacctat gtttgctcag aacacattct tccccccct 2400
ccccctctgg tctcctcttt gttttgtttt aaggaagaaa aatcagttgc gcgttctgaa 2460
atattttacc actgctgtga acaagtgaac acattgtgtc acatcatgac actcgtataa 2520
gcatggagaa cagtgtttt tttttagaac agaaaacaac aaaaaataac cccaaatga 2580
agattatttt ttatgaggag tgaacatttg ggtaaactcat ggctaagctt aaaaaaaact 2640
catggtgagg cttacaatg tcttgtaagc aaaaggtaga gccctgtatc aaccagaaa 2700

cacctagatc agaacaggaa tccacattgc cagtgcacatg agactgaaca gccaaatgga 2760
 ggctatgtgg agttggcatt gcatttaccg gcagtgcggg aggaatttct gaggggccat 2820
 cccaaggtct aggtggaggt ggggcatggt atttgagaca ttccaaaacg aaggcctctg 2880
 aaggaccctt cagaggtggc tctggaatga catgtgtcaa gctgcttgga cctcgtgctt 2940
 taagtgccta cattatctaa ctgtgctcaa gaggttctcg actggaggac cacactcaag 3000
 ccgacttatg cccaccatcc cacctctgga taattttgca taaaattgga ttagcctgga 3060
 gcaggttggg agccaaatgt ggcattttgtg atcatgagat tgatgcaatg agatagaaga 3120
 tgtttgctac ctgaacactt attgctttga aactagactt gaggaacca gggtttatct 3180
 tttgagaact tttggttaagg gaaaagggaa caggaaaaga aaccccaaac tcaggccgaa 3240
 tgatcaaggg gaccatagg aaatcttgtc cagagacaag acttcgggaa ggtgtctgga 3300
 cattcagaac accaagactt gaaggtgcct tgctcaatgg aagaggccag gacagagctg 3360
 acaaaatttt gctccccagt gaaggccaca gcaaccttct gccatcctg tctgttcatg 3420
 gagagggtcc ctgcctcacc tctgccattt tgggttagga gaagtcaagt tgggagcctg 3480
 aaatagtggg tcttggaata atggatcccc agtgaaaact agagctctaa gccattcag 3540
 cccatttcac acctgaaaat gttagtgtc accacttgga ccagcatcct taagtatcag 3600
 aaagcccaaa gcaattgctg catcttagta gggtaggga taagcaaaag aggatgttca 3660
 ccataacca ggaatgaaga taccatcagc aaagaatttc aatttggtca gtctttcatt 3720
 tagagctagt ctttcacagt accatctgaa tacctctttg aaagaaggaa gactttacgt 3780
 agtgtagatt tgttttgtgt tgtttgaaaa tattatcttt gtaattattt ttaatatgta 3840
 aggaatgctt ggaatatctg ctatatgtca actttatgca gcttcctttt gagggacaaa 3900
 tttaaaacaa acaaccccc atcacaaact taaaggattg caagggccag atctgttaag 3960
 tggtttcata ggagacacat ccagcaattg tgtggtcagt ggctctttta cccaataaga 4020
 tacatcacag tcacatgctt gatggtttat gttgacctaa gatttatattt gttaaaatct 4080
 ctctctgttg tgttcgttct tgttctgttt tgtttgttt tttaaagtct tgctgtggtc 4140
 tctttgtggc agaagtgttt catgcatggc agcaggcctg ttgctttttt atggcgattc 4200
 ccattgaaaa tgtaagtaaa tgtctgtggc cttgttctct ctatggtaaa gatattattc 4260
 accatgtaaa acaaaaaaca atatttattg tatttttagta tatttatata attatgttat 4320
 tgaaaaaaat tggcattaaa acttaaccgc atcagaacct attgtaaata caagttctat 4380
 ttaagtgtac taattaacat ataatatatg ttttaaatat agaattttta atgtttttta 4440
 atatattttc aaagtacata aaactcgagc catgggcgcg ccatcgatga ggaaccctta 4500
 gtgatggagt tggccactcc ctctctgcgc gctcgtcgc tcactgaggc cgggcgacca 4560
 aaggtcgccc gacgccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgagc 4620
 tgcctgcagg 4630

<210> 6

<211> 595

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hSARE-ArcMin

<400> 6

gggtggtgca gagtctgcat gcgtgaggag ctcttgggcg cgtcacagcc gcgctattct 60
cagcgtctct ccttttatgg ctccggaagt gagctggggt tgctggcagc cgtccctcgg 120
tccccgcagg ctgaatgagc cccgccgtcc ggggtgggag gtacggagcg gagggattag 180
tcattcctggc cccgccatgt gtgcaggggg tgggacctcc cgtcccttgc cgcccgaggag 240
ccgagtgggc gcagggcggg gcctgcaggg gcgcggggcg ggcgctggcg gggagcctgc 300
ggctgggcca atgagaaacc gggctcggcg agccggccgc ccacgggcct cgctggctgc 360
ataaagagcc ggcgccagg actcagcgca gagctcgggc gcggcgctcct ccctccgcag 420
cagccgagcc ggacctgcct ccccgggcgt gctccgccgg ccccgccgcc ggcccgagc 480
gacagacagg cgctccccgc agtccgcac gggaccagg ccgccggacc ccagcgccgg 540
accacctct gtccgccccg aggagtttgc cgctgccgg agcacctgcg cacag 595

<210> 7

<211> 744

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> BDNF

<400> 7

atgaccatcc ttttccttac tatggttatt tcatactttg gttgcatgaa ggctgcccc 60
atgaaagaag caaacatccg aggacaaggt ggcttggcct acccaggtgt gcggacccat 120
gggactctgg agagcgtgaa tgggccaag gcaggttcaa gaggttgac atcattggct 180
gacactttcg aacacgtgat agaagagctg ttggatgagg accagaaagt tcggcccaat 240
gaagaaaaca ataaggacgc agacttgtac acgtccaggg tgatgctcag tagtcaagt 300
cctttggagc ctctcttct ctttctgctg gaggaataca aaaattacct agatgctgca 360
aacatgtcca tgagggtccg gcgccactct gaccctgccc gccgagggga gctgagcgtg 420
tgtgacagta ttagtgagt ggtaacggcg gcagacaaaa agactgcagt ggacatgtcg 480
ggcgggacgg tcacagtcct tgaaaaggct cctgtatcaa aaggccaact gaagcaatac 540
ttctacgaga ccaagtgcaa tcccatgggt tacacaaaag aaggctgcag gggcatagac 600
aaaaggcatt ggaactccca gtgccgaact acccagtcgt acgtgcgggc cttaccatg 660
gatagcaaaa agagaattgg ctggcgattc ataaggatag acatttcttg tgtatgtaca 720
ttgaccatta aaagggaag atag 744

<210> 8

<211> 327

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> BDNF短3' UTR

<400> 8

tggatttatg ttgtatagat tagattatat tgagacaaaa attatctatt tgtatatata 60

cataacaggg taaattattc agttaagaaa aaaataattt tatgaactgc atgtataaat 120
 gaagtttata cagtacagtg gttctacaat ctattttattg gacatgtcca tgaccagaag 180
 ggaaacagtc atttgcgcac aacttaaaaa gtctgcatta cattccttga taatgttgtg 240
 gtttgttgcc gttgccaaga actgaaaaca taaaaagtta aaaaaataa taaattgcat 300
 gctgctttta ttgtgaattg ataataa 327

<210> 9

<211> 225

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> bGH聚(A)

<400> 9

ctgtgccttc tagttgccag ccattctgtt tttgccctc ccccgctgcct tccttgacct 60
 tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc 120
 tgagtaggtg tcattctatt ctggggggtg ggggtggggca ggacagcaag ggggaggatt 180
 gggaagacaa tagcaggcat gctggggatg cggtgggctc tatgg 225

<210> 10

<211> 2926

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> BDNF长3' UTR

<400> 10

tggatttatg ttgtatagat tagattatat tgagacaaaa attatctatt tgtatatata 60
 cataacaggg taaattattc agttaagaaa aaaataattt tatgaactgc atgtataaat 120
 gaagtttata cagtacagtg gttctacaat ctattttattg gacatgtcca tgaccagaag 180
 ggaaacagtc atttgcgcac aacttaaaaa gtctgcatta cattccttga taatgttgtg 240
 gtttgttgcc gttgccaaga actgaaaaca taaaaagtta aaaaaataa taaattgcat 300
 gctgctttta ttgtgaattg ataataaact gtctctcttc agaaaacaga aaaaaacaca 360
 cacacacaca acaaaaaattt gaaccaaaac attccgttta catttttagac agtaagtatc 420
 ttcgttcttg ttagtactat atctgtttta ctgcttttaa cttctgatag cgttggaatt 480
 aaaacaatgt caaggtgctg ttgtcattgc tttactggct taggggatgg gggatggggg 540
 gtatatTTTT gtttgttttg tgtttttttt tcgtttgttt gttttgtttt ttagttccca 600
 caggagtag agatggggaa agaattccta caatatatat tctggctgat aaaagataca 660
 tttgtatgtt gtgaagatgt ttgcaatatc gatcagatga ctagaaagtg aataaaaatt 720
 aaggcaactg aacaaaaaaa tgctcacact ccacatcccg tgatgcacct cccagcccc 780
 gctcattctt tgggcgttgg tcagagtaag ctgcttttga cggaaggacc tatgtttgct 840
 cagaacacat tctttcccc cctccccctc tggtctctc tttgttttgt ttttaaggaag 900
 aaaaatcagt tgcgcgttct gaaatatttt accactgctg tgaacaagtg aacacattgt 960

```

gtcacatcat gacactcgta taagcatgga gaacagtgat ttttttttag aacagaaaac 1020
aacaaaaaat aacccccaaa tgaagattat tttttatgag gagtgaacat ttgggtaaat 1080
catggctaag cttaaaaaaa actcatgggtg aggcttaaca atgtcttgta agcaaaaggt 1140
agagccctgt atcaaccag aaacacctag atcagaacag gaatccacat tgccagtgc 1200
atgagactga acagccaaat ggaggctatg tggagttggc attgcattta ccggcagtg 1260
gggaggaatt tctgagtggc catcccaagg tctaggtgga ggtggggcat ggtatttgag 1320
acattccaaa acgaaggcct ctgaaggacc cttcagaggt ggctctggaa tgacatgtgt 1380
caagctgctt ggacctcgtg ctttaagtgc ctacattatc taactgtgct caagaggttc 1440
tcgactggag gaccacactc aagccgactt atgcccacca tcccacctct ggataatttt 1500
gcataaaatt ggattagcct ggagcagggtt gggagccaaa tgtggcattt gtgatcatga 1560
gattgatgca atgagataga agatgtttgc tacctgaaca cttattgctt tgaaactaga 1620
cttgaggaaa ccagggttta tcttttgaga acttttggtg agggaaaagg gaacaggaaa 1680
agaaaccca aactcaggcc gaatgatcaa ggggacccat aggaaatctt gtccagagac 1740
aagacttcgg gaaggtgtct ggacattcag aacaccaaga cttgaaggtg ctttgc tcaa 1800
tggaagaggc caggacagag ctgacaaaat tttgctcccc agtgaaggcc acagcaacct 1860
tctgcccata ctgtctgttc atggagaggg tccctgcctc acctctgcca ttttgggtta 1920
ggagaagtca agttgggagc ctgaaatagt ggttcttgga aaaatggatc cccagtga 1980
actagagctc taagcccatt cagcccattt cacacctgaa aatgttagtg atcaccactt 2040
ggaccagcat ccttaagtat cagaaagccc caagcaattg ctgcatctta gtagggtag 2100
ggataagcaa aagaggatgt tcaccataac ccaggaatga agataccatc agcaaagaat 2160
ttcaatttgt tcagtccttc atttagagct agtctttcac agtaccatct gaatacctct 2220
ttgaaagaag gaagacttta cgtagtgtag atttgtttg tgttgtttga aaatattatc 2280
tttgtaatta tttttaatat gtaaggaatg cttggaatat ctgctatatg tcaactttat 2340
gcagcttcct tttgagggac aaatttaaaa caaacaacce cccatcaca acttaaagga 2400
ttgcaagggc cagatctgtt aagtgggttc ataggagaca catccagcaa ttgtgtggtc 2460
agtggctctt ttacccaata agatacatca cagtcacatg cttgatgggt tatgttgacc 2520
taagatttat tttgttaaaa tctctctctg ttgtgttcgt tcttgttctg ttttgtttg 2580
ttttttaaaag tcttgctgtg gtctctttgt ggcagaagtg tttcatgcat ggcagcaggc 2640
ctgttgcttt tttatggcga ttcccattga aaatgtaagt aaatgtctgt ggccttggtc 2700
tctctatggg aaagatatta ttcaccatgt aaaacaaaa acaatattta ttgtatttta 2760
gtatatattat ataattatgt tattgaaaaa aattggcatt aaaacttaac cgcatcagaa 2820
cctattgtaa atacaagttc tatttaagtg tactaattaa catataatat atgttttaaa 2880
tatagaattt ttaatgtttt taaatatatt ttcaaagtac ataaaa 2926

```

<210> 11

<211> 111

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hSARE

<220>

<221> misc_feature

<222> (40) .. (49)

<223> cAMP应答元件

<220>

<221> misc_feature

<222> (53) .. (63)

<223> MEF2结合位点

<220>

<221> misc_feature

<222> (69) .. (88)

<223> 血清应答元件

<400> 11

gggtggtgca gagtctgcat gcgtgaggag ctcttgggcg cgtcacagcc gcgctattct 60
cagcgtctct ccttttatgg ctccggaagt gagctggggt tgctggcagc c 111

<210> 12

<211> 484

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hArcMin

<400> 12

gtccctcggc cccgcaggc tgaatgagcc ccgccgtccg ggggtgggagg tacggagcgg 60
agggattagt catcctggcc ccgcatgtg tgcaggggggt gggacctccc gtcccttgcc 120
gcccgggagc cgagtgggcg cagggcgggg cctgcagggg cgcgggcggg gcgctggcgg 180
ggagcctgcg gctgggcca tgaagaaaccg ggctcggcga gccggccgcc cacgggcctc 240
gctggctgca taaagagccg gcggccagga ctacgcgcag agctcgggcg cggcgtcctc 300
cctccgcagc agccgagccg gacctgcctc cccgggcgtg ctccgccggc cccgccgccg 360
gcccgcagcg acagacaggc gtcctccgca gtcgcgcag ggacctaggc cgcgggacct 420
cagcgccgga ccacctctg tccgccccga ggagtttgcc gcctgccgga gcacctgcgc 480
acag 484

<210> 13

<211> 2165

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hAlbumin填充序列

<400> 13

gcgcgccatc gatgtgaagt gggtaacctt tatttccctt ctttttctct ttagctcggc 60

```

ttattccagg ggtgtgtttc gtcgagatgc acacaagagt gaggttgctc atcggtttaa 120
agatttgga gaagaaaatt tcaaagcctt ggtgttgatt gcctttgctc agtatcttca 180
gcagtgtcca tttgaagatc atgtaaaatt agtgaatgaa gtaactgaat ttgcaaaaac 240
atgtgttgct gatgagtcag ctgaaaattg tgacaaatca cttcataccc tttttggaga 300
caaattatgc acagttgcaa ctcttcgtga aacctatggt gaaatggctg actgctgtgc 360
aaaacaagaa cctgagagaa atgaatgctt cttgcaacac aaagatgaca acccaaacct 420
cccccgattg gtgagaccag aggttgatgt gatgtgact gcttttcatg acaatgaaga 480
gacctttttg aaaaaatact tatacgaaat tgccagaaga catccttact tttatgcccc 540
ggaactcctt ttctttgcta aaaggtataa agctgctttt acagaatgtt gccaaactgc 600
tgataaagct gcctgcctgt tgccaaagct cgatgaactt cgggatgaag ggaaggcttc 660
gtctgccaaa cagagactca agtgtgccag tctccaaaaa tttggagaaa gagctttcaa 720
agcatgggca gtagctcgcc tgagccagag atttccaaa gctgagtttg cagaagtttc 780
caagttagtg acagatctta ccaaagtcca cacggaatgc tgccatggag atctgcttga 840
atgtgctgat gacagggcgg accttgccaa gtatatctgt gaaaatcaag attcgatctc 900
cagtaaaactg aaggaatgct gtgaaaaacc tctgttgaa aaatcccact gcattgccga 960
agtggaaaat gatgagatgc ctgctgactt gccttcatta gctgctgatt ttgttgaaag 1020
taaggatgtt tgcaaaaact atgctgaggc aaaggatgtc ttcctgggca tgtttttgta 1080
tgaatatgca agaaggcatc ctgattactc tgtcgtgctg ctgctgagac ttgccaagac 1140
ctatgaaacc actctagaga agtgctgtgc cgctgcagat cctcatgaat gctatgccaa 1200
agtgttcgat gaatttaaac ctcttgtaga agagcctcag aatttaatca aacaaaattg 1260
tgagcttttt gagcagcttg gagagtacaa attccagaat gcgctattag ttcgttacac 1320
caagaaagta cccaagtgt caactccaac tcttgtagag gtctcaagaa acctaggaaa 1380
agtgggcagc aaatgttgta aacatcctga agcaaaaaga atgccctgtg cagaagacta 1440
tctatccgtg gtcctgaacc agttatgtgt gttgcatgag aaaacgccag taagtgcag 1500
agtcacaaaa tgctgcacag aatccttggt gaacaggcga ccatgtcgtg aaacctatgg 1560
tgaaatggct gactgctgtg caaaacaaga acctgagaga aatgaatgct tcttgcaaca 1620
caaagatgac aacccaaacc tccccgatt ggtgagacca gaggttgatg tgatgtgcac 1680
tgcttttcat gacaatgaag agacctttt gaaaaatac ttatacгаа ttgccagaag 1740
acatccttac ttttatgccc cggaactcct tttctttgct aaaaggtata aagctgcttt 1800
tacagaatgt tgccaagctg ctgataaagc tgctgctg ttgccaagc tcgatgaact 1860
tcgggatgaa gggaaggctt cgtctgccaa acagagactc aagtgtgcca gtctccaaaa 1920
atltggagaa agagctttca aagcatgggc agtagctgc ctgagccaga cttttcagct 1980
ctggaagtcg atgaaacata cgttccaaa gagtttaatg ctgaaacatt caccttccat 2040
gcagatatat gcacactttc tgagaaggag agacaaatca agaaacaaac tgcacttggt 2100
gagctcgtga aacacaagcc caaggcaaca aaagagcaac tgaaagctgt tatggatgat 2160
ttcgc 2165
<210> 14
<211> 12
<212> DNA

```

<213> 人工序列(Artificial Sequence)

<220>

<223> Kozak序列

<400> 14

acccggagca gc 12

<210> 15

<211> 9

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> Kozak序列

<400> 15

acagccacc 9

<210> 16

<211> 1039

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> E-SARE-ArcMin序列

<400> 16

gggtggtgca gagtctgcat gcgtgaggag ctcttgggcg cgtcacagcc gcgctattct 60
cagcgtctct ccttttatgg ctccggaagt gagctggggt tgctggcagc cgggtggtgc 120
agagtctgca tgcgtgagga gctcctgggc gcgtcacagc cgcgctattc tcagcgtctc 180
tccttttatg gctccggaag tgagctgggg ttgctggcag ccgggtggtg cagagtctgc 240
atgcgtgagg agctcctggg cgcgtcacag ccgcgctatt ctcagcgtct ctccttttat 300
ggctccgga gtagctggg gttgctggca gccgggtggt gcagagtctg catgcgtgag 360
gagctcctgg gcgcgtcaca gccgcgctat tctcagctc tctcctttta tggctccgga 420
agtgagctgg ggttgctggc agccgggtgg tgcagagtct gcatgcgtga ggagctcctg 480
ggcgcgtcac agccgcgcta ttctcagcgt ctctcctttt atggctccgg aagtgagctg 540
gggttgctgg cagccgtccc tcggtccccg caggtgaat gagccccgcc gtccgggggtg 600
ggaggtacgg agcggaggga ttagtcatcc tggccccgcc atgtgtgcag ggggtgggac 660
ctcccgctcc ttgccgcccc ggagccgagt gggcgagggg cggggcctgc aggggcgcgg 720
gcggggcgct ggcggggagc ctgcggtggt gccaatgaga aaccgggctc ggcgagccgg 780
ccgccacagg gcctcgctgg ctgcataaag agccggcggc caggactcag cgcagagctc 840
gggcgcggcg tcctccctcc gcagcagccg agccggacct gcctccccgg gcgtgctccg 900
ccggccccgc cgccggcccc cagcgacaga caggcgtcc ccgcagctcc gcacgggacc 960
caggccgccg gaccccagcg ccggaccacc ctctgtccgc cccgaggagt ttgccgctg 1020
ccggagcacc tgcgcacag 1039

<210> 17

<211> 278

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 启动子序列

<400> 17

```
tcgaggtgag ccccacgttc tgcttcactc tccccatctc cccccctcc ccacccccaa 60
ttttgtatatt atttatTTTT taattatTTT gtgcagcgat gggggcgggg gggggggggg 120
ggcgcgcgcc aggcggggcg gggcggggcg agggcggggg cggggcgagg cggagaggtg 180
cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc 240
ggcgcgcgcc ctataaaaag cgaagcgcg cgcgggcg 278
```

<210> 18

<211> 240

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 启动子序列

<400> 18

```
cggctggggc tgagggtgag ggtcccgttt ccccaaaggc ctagcctggg gttccagcca 60
caagccctac cgggcagcgc ccggccccgc ccctccaggc ctggcactcg tctcaacca 120
agatggcgcg gatggcttca ggcgcatcac gacaccggcg cgtcacgca cccgccctac 180
gggcacctcc cgcgcttttc ttagcgccgc agacggtggc cgagcggggg accgggaagc 240
```

<210> 19

<211> 379

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 启动子序列

<400> 19

```
attcctgctg ggaaaagcaa gtggaggtgc tccttgaaga aacaggggga tcccaccgat 60
ctcaggggtt ctgttctggc ctgcggccct ggatcgtcca gcctgggtcg gggtagggag 120
cagacctcgc cttatcggc tggggctgag ggtgagggtc ccgtttcccc aaaggcctag 180
cctggggttc cagccacaag ccctaccggg cagcgcccg ccccgcccct ccaggcctgg 240
cactcgtcct caaccaagat ggcgcggtat gttcaggcg catcacgaca ccggcgcgtc 300
acgcgacccg ccctacgggc acctcccgcg cttttcttag cgccgcagac ggtggtcgag 360
cggggggaccg ggaagctta 379
```

<210> 20

<211> 146

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 启动子序列

<400> 20

atgacacaaa ccccgcccag cgtcttgta ttggcgaatt cgaacacgca gatgcagtcg 60
gggcggcgcg gtcccaggtc cacttcgcat attaaggtga cgcgtgtggc ctggaacacc 120
gagcgaccct gcagcgaccc gcttaa 146

<210> 21

<211> 301

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 增强子序列

<400> 21

tacataactt acggtaaatt gcccgccctgg ctgaccgccc aacgaccccc gccattgac 60
gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg 120
ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaa 180
tacgccccct attgacgtca atgacggtaa atggcccgcc tggcattatg ccagttacat 240
gaccttatgg gactttccta cttggcagta catctacgta ttagtcatcg ctattacat 300
g 301

<210> 22

<211> 122

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 聚A信号

<400> 22

taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaataa aaatgcttta 60
tttgtgaaat ttgtgatgct attgctttat ttgtaacct tataagctgc aataaacaag 120
tt 122

<210> 23

<211> 56

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 聚A信号

<400> 23

aataaaggaa atttattttc attgcaatag tgtgttgga ttttttgtgt ctctca 56

<210> 24

<211> 97

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 内含子序列

<400> 24

gtaagttag tctttttgtc ttttatttca ggtcccgat ccggtggtg tgcaaatcaa 60
agaactgctc ctcatggat gttgccttta cttctag 97

<210> 25

<211> 133

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 内含子序列

<400> 25

gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg cttgtcgaga 60
cagagaagac tcttgcgttt ctgataggca cctattggtc ttactgacat ccactttgcc 120
tttctctcca cag 133

<210> 26

<211> 4008

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-mScar-shortUTR

<400> 26

cctgcaggca gctgcgcgct cgctcgtca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccgc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc ttgcgtcgac aggcctccta ggccggaccgc ttgcatgcac 180
gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
ccgaccccca cccactggac aagcacccaa cccccattcc ccaaattgcg catcccctat 300
cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgcaactgcca gcttcagcac 360
cgccgacagt gccttcgccc ccgcctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
gcgcgctgac gtcaactgcc ggtccccgc aaactcccct tcccggccac cttggtcgcg 480
tccgcgccgc cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
ggcgcgacca tctgcgctgc ggcgccggcg actcagcgct gcctcagtct gcggtgggca 600
gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccaccatggt gagcaagggc 660
gaggcagtga tcaaggagtt catgcggttc aaggtgcaca tggagggtc catgaacggc 720
cacgagttcg agatcgaggg cgagggcgag ggccgccctt acgagggcac ccagaccgcc 780
aagctgaagg tgaccaaggg tggccccctg cccttctcct gggacatcct gtcccctcag 840

```

ttcatgtacg gctccagggc cttcatcaag caccgccg acatccccga ctactataag 900
cagtccttcc ccgagggcctt caagtgggag cgcgtgatga acttcgagga cggcggcgcc 960
gtgaccgtga cccaggacac ctccctggag gacggcacc tgatctacaa ggtgaagctc 1020
cgcggcacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080
gcgtccaccg agcggttgta ccccaggag ggcgtgctga agggcgacat taagatggcc 1140
ctgcgcctga aggacggcgg ccgctacctg gcggacttca agaccaccta caaggccaag 1200
aagcccggtc agatgcccgg cgcctacaac gtcgaccgca agttggacat cacctccac 1260
aacgaggact acaccgtggt ggaacagtac gaacgctccg agggccgcca ctccaccggc 1320
ggcatggacg agctgtacaa gtagggtagc tggatttatg ttgtatagat tagattatat 1380
tgagacaaaa attatctatt tgtatatata cataacaggg taaattattc agttaagaaa 1440
aaaataattt tatgaactgc atgtataaat gaagtttata cagtacagtg gttctacaat 1500
ctatttattg gacatgtcca tgaccagaag ggaaacagtc atttgccac aacttaaaaa 1560
gtctgcatta cattccttga taatgtttg gtttgttgc gttgccaaga actgaaaaca 1620
taaaaagtta aaaaaataa taaattgcat gctgcttta ttgtgaattg ataataagcg 1680
cgccatcgat gtgaagtggg taacctttat ttcccttctt tttctcttta gctcggtta 1740
ttccaggggt gtgtttcgtc gagatgcaca caagagttag gttgctcatc ggtttaaaga 1800
tttgggagaa gaaaatttca aagccttggt gttgattgcc tttgctcagt atcttcagca 1860
gtgtccattt gaagatcatg taaaattagt gaatgaagta actgaatttg caaaaacatg 1920
tgttgctgat gtagcagctg aaaattgtga caaatcactt catacccttt ttggagacaa 1980
attatgcaca gttgcaactc ttcgtgaaac ctatggtgaa atggctgact gctgtgcaaa 2040
acaagaacct gagagaaatg aatgcttctt gcaacacaaa gatgacaacc caaacctccc 2100
ccgattgggtg agaccagagg ttgatgtgat gtgactgct tttcatgaca atgaagagac 2160
ctttttgaaa aaatacttat acgaaattgc cagaagacat cttactttt atgccccgga 2220
actccttttc tttgctaaaa ggtataaagc tgcttttaca gaatgttgcc aagctgctga 2280
taaagctgcc tgcctgttgc caaagctcga tgaacttcgg gatgaaggga aggcttcgtc 2340
tgccaaacag agactcaagt gtgccagtct ccaaaaattt ggagaaagag ctttcaaagc 2400
atgggcagta gctcgcctga gccagagatt tcccaaagct gagtttgtag aagtttccaa 2460
gttagtgaca gatcttacca aagtccacac ggaatgctgc catggagatc tgcttgaatg 2520
tgctgatgac agggcggacc ttgccaagta tatctgtgaa aatcaagatt cgatctccag 2580
taaactgaag gaatgctgtg aaaaacctct gttggaaaaa tccactgca ttgccgaagt 2640
ggaaaatgat gagatgcctg ctgacttgcc ttcattagct gctgattttg ttgaaagtaa 2700
ggatgtttgc aaaaactatg ctgaggcaaa ggatgtcttc ctgggcatgt tttgtatga 2760
atatgcaaga aggcacacct attactctgt cgtgctgctg ctgagacttg ccaagaccta 2820
tgaaccact ctagagaagt gctgtgccgc tgcagatcct catgaatgct atgccaaagt 2880
gttcgatgaa tttaaacctc ttgtggaaga gcctcagaat ttaatcaaac aaaattgtga 2940
gctttttgag cagcttgag agtacaaatt ccagaatgcg ctattagtgc gttacaccaa 3000
gaaagtaccc caagtgtcaa ctccaactct tgtagaggtc tcaagaaacc taggaaaagt 3060
gggcagcaaa tgttgtaaac atcctgaagc aaaaagaatg ccctgtgcag aagactatct 3120
atccgtgggtc ctgaaccagt tatgtgtgtt gcatgagaaa acgccagtaa gtgacagagt 3180

```

caccaaattgc tgcacagaat ccttggtgaa caggcgacca tgtcgtgaaa cctatggtga 3240
 aatggctgac tgctgtgcaa aacaagaacc tgagagaaat gaatgcttct tgcaacacaa 3300
 agatgacaac ccaaacctcc cccgattggt gagaccagag gttgatgtga tgtgcactgc 3360
 ttttcatgac aatgaagaga cctttttgaa aaaatactta tacgaaattg ccagaagaca 3420
 tccttacttt tatgccccgg aactcctttt ctttgctaaa aggtataaag ctgcttttac 3480
 agaatgttgc caagctgctg ataaagctgc ctgcctgttg ccaaagctcg atgaacttcg 3540
 ggatgaaggg aaggcttcgt ctgccaaaca gagactcaag tgtgccagtc tccaaaaatt 3600
 tggagaaaaga gctttcaaag catgggcagt agctcgctg agccagactt ttcagctctg 3660
 gaagtcgatg aaacatacgt tcccaaagag tttaatgctg aaacattcac cttccatgca 3720
 gatatatgca cactttctga gaaggagaga caaatcaaga acaaaactgc acttggtgag 3780
 ctcgtgaaac acaagcccaa ggcaacaaaa gagcaactga aagctgttat ggatgatttc 3840
 gctcgagcca tgggcgcgcc atcgatgagg aaccctagt gatggagttg gccactccct 3900
 ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaa ggtcgcccga cgccggggt 3960
 ttgcccgggc ggcctcagt agcgagcgag cgcgagctg cctgcagg 4008

<210> 27

<211> 4443

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-mScar-longUTR

<400> 27

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggctccta ggccgaccgc ttgcatgcac 180
 gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
 ccgaccccga cccactggac aagcacccaa ccccatctc ccaaattgag catcccctat 300
 cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgactgcca gcttcagcac 360
 cgccgacagt gccttcgccc ccgcctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
 gcgcgctgac gtcactcgcc ggtccccgc aaactccct tccgggccac cttggtcgcg 480
 tccgcgcgcg cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
 ggccgcacca tctgcgctgc ggccggcgcg actcagcgct gcctcagtct gcggtgggca 600
 gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccaccatggt gagcaagggc 660
 gaggcagtga tcaaggagtt catgcggttc aaggtgcaca tggagggtc catgaacggc 720
 cacgagttcg agatcgaggg cgagggcgag ggccgcccct acgagggcac ccagaccgcc 780
 aagctgaagg tgaccaaggg tggccccctg cccttctcct gggacatcct gtcccctcag 840
 ttcatgtacg gctccagggc cttcatcaag caccgcgcg acatccccga ctactataag 900
 cagtccttcc ccgagggtt caagtgggag gcgctgatga acttcgagga cgccggcgcc 960
 gtgaccgtga cccaggacac ctccctggag gacggcacc tgatctacaa ggtgaagctc 1020
 cgccggacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080

gcgtccaccg agcggttgta ccccgaggac ggcgtgctga agggcgacat taagatggcc 1140
 ctgcgcctga aggacggcgg ccgctacctg gcggacttca agaccaccta caaggccaag 1200
 aagcccggtc agatgcccgg cgcctacaac gtcgaccgca agttggacat cacctcccac 1260
 aacgaggact acaccgtggt ggaacagtac gaacgctccg agggccgcca ctccaccggc 1320
 ggcatggacg agctgtacaa gtaggggtacc tggatttatg ttgtatagat tagattatat 1380
 ttagacaaaa attatctatt tgtatatata cataacaggg taaattattc agttaagaaa 1440
 aaaataattht tatgaactgc atgtataaat gaagtttata cagtacagtg gttctacaat 1500
 ctatthtattg gacatgtcca tgaccagaag ggaaacagtc atthtgcgcac aacttaaaaa 1560
 gtctgcatta cattccttga taatgtttgt gtttgttgcc gttgccaaga actgaaaaca 1620
 taaaaagtta aaaaaaataa taaattgcat gctgctthta ttgtgaattg ataataaact 1680
 gtctctthtc agaaaacaga aaaaaacaca cacacacaca acaaaaatth gaacaaaaac 1740
 attccgthta cattthtagac agtaagtatc thcgtthctg ttagtactat atctgththta 1800
 ctgctththta cthctgatag cgttggaatt aaaacaatgt caaggtgctg ttgtcattgc 1860
 thtactggct taggggatgg gggatggggg gtataththt gthtgtthtg tgtthththt 1920
 tcgtthgtth gththgttht ttagththcca cagggatag agatggggaa agaaththcta 1980
 caatatatat tctggctgat aaaagataca thtgtatgt gtgaagatgt ttgcaatath 2040
 gatcagatga ctagaaagt aataaaaatt aaggcaactg aacaaaaaaa tgctcacact 2100
 ccacatcccg tgatgcacct cccaggcccc gctcattctt tgggcgttg tgagagtaag 2160
 ctgctththga cggaaggacc tatgtthgtc cagaacacat thththcccc cthccccctc 2220
 tggthctctc thtgtthtgt thtaaggaag aaaaatcagt tgcgcgttht gaaataththt 2280
 accactgctg tgaacaagt aacacattgt gtcacatcat gacactcgta taagcatgga 2340
 gaacagtgat ththththtag aacagaaaac acaaaaaat aacccccaaa tgaagattat 2400
 thththtatg gagtgaacat ttgggtaaat catggctaag cthaaaaaaa actcatggtg 2460
 aggtthtaaca atgtcttgta agcaaaaggt agagccctgt atcaaccag aaacacctag 2520
 atcagaacag gaatccacat tgccagtga atgagactga acagccaaat ggaggctatg 2580
 tggagtthgc attgcattta ccggcagtgc gggaggaatt tctgagtggc catcccaagg 2640
 tctaggtgga ggtggggcat ggtaththg acaththcaa acgaaggcct ctgaaggacc 2700
 cthcagaggt ggctctggaa tgacatgtgt caagctgctt ggacctcgth cthtaagtgc 2760
 ctacattath taactgtgct caagaggttc tcgactggag gaccacactc aagccgactt 2820
 atgcccacca thccacctct ggataathth gcataaaat ggattagcct ggagcaggtt 2880
 gggagccaaa tgtggcattt gtgatcatga gattgatgca atgagataga agatgtthgc 2940
 tactgaaca cthattgtct tgaaactaga cthgaggaaa ccagggtthta thththtgaga 3000
 actththtgta agggaaaagg gaacaggaaa agaaaccca aactcaggcc gaatgatcaa 3060
 ggggacctat aggaathctt gtccagagac aagactthcg gaaggtgtct ggacattcag 3120
 aacaccaaga cthgaaggtg cctthgtcaa tggaagaggc caggacagag ctgacaaaat 3180
 thtgtcccc agtgaaggcc acagcaacct thtgcctatc ctgtctgttc atggagaggg 3240
 thctgcctc acctctgcca ththgggtta ggagaagtca agthgggagc ctgaaatagt 3300
 ggtthctgga aaaaatggatc cccagtgaat actagagctc taagccatt cagccattth 3360
 cacacctgaa aatgttagtg atcaccactt ggaccagcat cthtaagtat cagaaagccc 3420

caagcaattg ctgcatctta gtagggtag ggataagcaa aagaggatgt tcaccataac 3480
 ccaggaatga agataccatc agcaaagaat ttcaatttgt tcagtctttc atttagagct 3540
 agtctttcac agtaccatct gaatacctct ttgaaagaag gaagacttta cgtagtgtag 3600
 atttgttttg tgttgtttga aaatattatc tttgtaatta tttttaatat gtaaggaatg 3660
 cttggaatat ctgctatatg tcaactttat gcagcttctt tttgaggac aaatttaaaa 3720
 caaacaaccc cccatcacaa acttaaagga ttgcaagggc cagatctgtt aagtggtttc 3780
 ataggagaca catccagcaa ttgtgtggc agtggctctt ttaccaata agatacatca 3840
 cagtcacatg cttgatgggt tatgttgacc taagatttat tttgttaaaa tctctctctg 3900
 ttgtgttcgt tcttgttctg ttttgttttg ttttttaaag tcttgctgtg gtctctttgt 3960
 ggcagaagtg tttcatgcat ggcagcaggc ctgttgcttt tttatggcga ttcccattga 4020
 aaatgtaagt aaatgtctgt ggcttggctc tctctatggt aaagatatta ttcacatgt 4080
 aaaacaaaaa acaatattta ttgtatttta gtatatttat ataattatgt tattgaaaaa 4140
 aattggcatt aaaacttaac cgcacagaa cctattgtaa atacaagttc tatttaagtg 4200
 tactaattaa catataatat atgttttaaa tatagaattt ttaatgtttt taaatatatt 4260
 ttcaaagtac ataaaactcg agccatgggc gcgccatcga tgaggaacct ctagtgatgg 4320
 agttggccac tccctctctg cgcgctcgtc cgctcactga ggccgggcga ccaaaggtcg 4380
 ccgacgccc gggtttgcc cgggcggcct cagtgcgcga gcgagcgcgc agctgcctgc 4440
 agg 4443

<210> 28

<211> 3906

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-mScar-bGH

<400> 28

cctgcaggca gctgcgcgtc cgctcgtca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggctgcccgc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggccggaccgc ttgcatgcac 180
 gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
 ccgaccccga cccactggac aagcacccaa ccccatcctc ccaaattgcg catcccctat 300
 cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgcaactgcca gcttcagcac 360
 cgccggacagt gccttcgccc ccgcctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
 gcgcgctgac gtcactcgcc ggtccccgc aaactccctt tcccggccac cttggtcgcg 480
 tccgcgccgc cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
 ggccgcacca tctgcgctgc ggccggcg actcagcgt gcctcagtct gcggtgggca 600
 gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccacatggt gagcaagggc 660
 gaggcagtga tcaaggagtt catgcggtt aaggtgcaca tggagggtc catgaacggc 720
 cacgagttcg agatcgaggc cgaggcgag ggccgccctt acgagggcac ccagaccgcc 780
 aagctgaagg tgaccaaggg tggccccctg cccttctcct gggacatcct gtcccctcag 840

```

ttcatgtacg gctccagggc cttcatcaag caccgccgacg acatccccga ctactataag 900
cagtccttcc ccgagggcctt caagtgggag cgcgtgatga acttcgagga cggcgggcgc 960
gtgaccgtga cccaggacac ctccctggag gacggcaccg tgatctacaa ggtgaagctc 1020
cgcggcacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080
gcgtccaccg agcgggttgta ccccaggagc ggcgtgctga agggcgacat taagatggcc 1140
ctgcgcctga aggacggcgg ccgctacctg gcggacttca agaccaccta caaggccaag 1200
aagcccggtc agatgcccgg cgcctacaac gtcgaccgca agttggacat cacctccac 1260
aacgaggact acaccgtggt ggaacagtac gaacgctccg agggccgcca ctccaccggc 1320
ggcatggacg agctgtacaa gtagggtagc ctgtgccttc tagttgccag ccatctgttg 1380
tttgccccctc ccccgctgcct tccttgacct tggaagggtg cactcccact gtcctttcct 1440
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg 1500
gggtggggca ggacagcaag ggggaggatt gggaagaca tagcaggcat gctggggatg 1560
cggtgggctc tatggggcgc ccacgatgt gaagtgggta acctttattt cccttctttt 1620
tctcttttagc tcggcttatt ccaggggtgt gtttcgtcga gatgcacaca agagttaggt 1680
tgctcatcgg tttaaagatt tgggagaaga aaatttcaa gccttggtgt tgattgcctt 1740
tgctcagtat cttcagcagt gtccatttga agatcatgta aaattagtga atgaagtaac 1800
tgaatttgca aaacatgtg ttgctgatga gtcagctgaa aattgtgaca aatcacttca 1860
tacccttttt ggagacaaat tatgcacagt tgcaactctt cgtgaaacct atggtgaaat 1920
ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa tgcttcttgc aacacaaaga 1980
tgacaacca aacctcccc gattggtgag accagagggt gatgtgatgt gactgcttt 2040
tcatgacaat gaagagacct ttttgaaaa atacttatac gaaattgcca gaagacatcc 2100
ttacttttat gccccggaac tccttttctt tgctaaaagg tataaagctg cttttacaga 2160
atgttgccaa gctgctgata aagctgcctg cctgttgcca aagctcgatg aacttcggga 2220
tgaagggaag gcttcgtctg ccaaacagag actcaagtgt gccagtctcc aaaaatttgg 2280
agaaagagct ttcaaagcat gggcagtagc tcgcctgagc cagagatttc ccaaagctga 2340
gtttgcagaa gtttccaagt tagtgacaga tcttaccaa gtccacacgg aatgctgcca 2400
tgagatctg cttgaatgtg ctgatgacag ggcggacctt gccagtata tctgtgaaaa 2460
tcaagattcg atctccagta aactgaagga atgctgtgaa aaacctctgt tggaaaaatc 2520
ccactgcatt gccgaagtgg aaaatgatga gatgcctgct gacttgcttt cattagctgc 2580
tgattttgtt gaaaagtaag atgtttgcaa aaactatgct gaggcaaagg atgtcttcct 2640
gggcatgttt ttgtatgaat atgcaagaag gcactctgat tactctgtcg tgctgctgct 2700
gagacttgcc aagacctatg aaaccactct agagaagtgc tgtgccgctg cagatcctca 2760
tgaatgctat gccaaaagtgt tcgatgaatt taaacctctt gtggaagagc ctcagaattt 2820
aatcaacaa aattgtgagc tttttgagca gcttgagag tacaatttc agaattgcgt 2880
attagttcgt tacaccaaga aagtaccca agtgtcaact ccaactcttg tagaggtctc 2940
aagaaacctg gaaaaagtgg gcagcaaatg ttgtaaacat cctgaagcaa aaagaatgcc 3000
ctgtgcagaa gactatctat ccgtggtcct gaaccagtta tgtgtgttgc atgagaaaac 3060
gccagtaagt gacagagtca ccaaagtctg cacagaatcc ttggtgaaca ggcgacctg 3120
tcgtgaaacc tatggtgaaa tggctgactg ctgtgcaaaa caagaacctg agagaaatga 3180

```

atgcttcttg caacacaaag atgacaaccc aaacctcccc cgattggtga gaccagaggt 3240
 tgatgtgatg tgcactgctt ttcattgaca tgaagagacc tttttgaaaa aatacttata 3300
 cgaaattgcc agaagacatc cttactttta tgccccggaa ctccttttct ttgctaaaag 3360
 gtataaagct gctttttacag aatgttgcca agctgctgat aaagctgcct gcctgttgcc 3420
 aaagctcgat gaacttcggg atgaaggga ggcttcgtct gccaaacaga gactcaagtg 3480
 tgccagtctc caaaaatttg gagaaagagc tttcaaagca tgggcagtag ctgcctgag 3540
 ccagactttt cagctctgga agtcgatgaa acatacgttc ccaaagagtt taatgctgaa 3600
 acattcacct tccatgcaga tatatgcaca ctttctgaga aggagagaca aatcaagaaa 3660
 caaactgcac ttgttgagct cgtgaaacac aagcccaagg caacaaaaga gcaactgaaa 3720
 gctgttatgg atgatttcgc tcgagccatg ggcgcgccat cgatgaggaa cccctagtga 3780
 tggagttagc cactccctct ctgcgcgtc gctcgtcac tgaggccggg cgaccaaagg 3840
 tcgcccgcg cccgggcttt gcccgggcgg cctcagttag cgagcgagcg cgcagctgcc 3900
 tgcagg 3906

<210> 29

<211> 4140

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-dmScar-shortUTR

<400> 29

cctgcaggca gctgcgcgtc cgctcgtca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggccgaccgc ttgcatgcac 180
 gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
 ccgaccccca cccactggac aagcacccaa cccccattcc ccaaattgcg catcccctat 300
 cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgcaactgcca gcttcagcac 360
 cgccgacagt gccttcgccc ccgcctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
 gcgcgctgac gtcactcgcc ggtccccgc aaactcccct tcccggccac cttggtcgcg 480
 tccgcgccgc cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
 ggccgcacca tctgcgtgc ggccggcg actcagcgtc gcctcagtct gcggtgggca 600
 gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccaccatggt gagcaagggc 660
 gaggcagtga tcaaggagtt catgcggttc aaggtgcaca tggagggtc catgaacggc 720
 cacgagttcg agatcgaggg cgagggcgag ggccgcccct acgagggcac ccagaccgcc 780
 aagctgaagg tgaccaaggg tggccccctg cccttctcct gggacatcct gtcccctcag 840
 ttcattgtac gctccagggc cttcatcaag cccccgccg acatccccga ctactataag 900
 cagtccttcc ccgagggtt caagtgggag cgcgtgatga acttcgagga cgccggcgcc 960
 gtgaccgtga cccaggacac ctccctggag gacggcacc tgatctacaa ggtgaagctc 1020
 cgccgcacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080
 gcgtccaccg agcggttgta ccccaggac ggcgtgctga agggcgacat taagatggcc 1140

ctgcgcctga aggacggcgg ccgctacctg gcggacttca agaccaccta caaggccaag 1200
aagcccggtgc agatgcccg cgcctacaac gtgcaccgca agttggacat cacctccac 1260
aacgaggact acaccgtggt ggaacagtac gaacgctccg agggccgcca ctccaccggc 1320
ggcatggacg agctgtacaa gaagcttccg cggagccatg gcttcccggc ggcggtggcg 1380
gcgcaggatg atggcacgct gcccattgtt tgtgcccagg agagcgggat ggaccgtcac 1440
cctgcagcct gtgcttctgc taggatcaat gtgtagggtta cctggattta tgttgtatag 1500
attagattat attgagacaa aaattatcta tttgtatata tacataacag ggtaaattat 1560
tcagttaaga aaaaaataat tttatgaact gcatgtataa atgaagttta tacagtacag 1620
tggttctaca atctatttat tggacatgtc catgaccaga agggaaacag tcatttgcgc 1680
acaacttaaa aagtctgcat tacattcctt gataatgttg tggtttgttg ccgttgccaa 1740
gaactgaaaa cataaaaaagt taaaaaaaat aataaattgc atgctgcttt aattgtgaat 1800
tgataataag cgcgccatcg atgtgaagtg ggtaaccttt atttcccttc tttttctctt 1860
tagctcggct tattccaggg gtgtgtttcg tcgagatgca cacaagagtg aggttgctca 1920
tcggtttaaa gatttgggag aagaaaattt caaagccttg gtgttgattg cttttgctca 1980
gtatcttcag cagtgtccat ttgaagatca tgtaaaatta gtgaatgaag taactgaatt 2040
tgcaaaaaca tgtgttgctg atgagtcagc tgaaaattgt gacaaatcac ttcataacct 2100
ttttggagac aaattatgca cagttgcaac tcttcgtgaa acctatggtg aaatggctga 2160
ctgctgtgca aaacaagaac ctgagagaaa tgaatgcttc ttgcaacaca aagatgacaa 2220
cccaaacctc ccccgattgg tgagaccaga ggttgatgtg atgtgactg cttttcatga 2280
caatgaagag acctttttga aaaaatactt atacgaaatt gccagaagac atccttactt 2340
ttatgccccg gaactccttt tctttgctaa aaggtataaa gctgctttta cagaatgttg 2400
ccaagctgct gataaagctg cctgcctgtt gccaaagctc gatgaacttc gggatgaagg 2460
gaaggcttcg tctgccaaac agagactcaa gtgtgccagt ctccaaaaat ttggagaaag 2520
agctttcaaa gcatgggcag tagctcgctt gagccagaga tttcccaaag ctgagtttgc 2580
agaagtttcc aagttagtga cagatcttac caaagtcac acggaatgct gccatggaga 2640
tctgcttgaa tgtgctgatg acagggcgga ccttgccaag tatatctgtg aaaatcaaga 2700
ttcgatctcc agtaaaactga aggaatgctg tgaaaaacct ctgttggaat aatccactg 2760
cattgccgaa gtggaaaatg atgagatgcc tgctgacttg ctttcattag ctgctgattt 2820
tgttgaaagt aaggatgttt gcaaaaacta tgctgaggca aaggatgtct tcctgggcat 2880
gtttttgtat gaatatgcaa gaaggcatcc tgattactct gtcgtgctgc tgctgagact 2940
tgccaagacc tatgaaacca ctctagagaa gtgctgtgcc gctgcagatc ctcatgaatg 3000
ctatgccaaa gtgttcgatg aatttaaac tcttgaggaa gagcctcaga atttaataca 3060
acaaaattgt gagctttttg agcagcttgg agagtacaaa ttccagaatg cgctattagt 3120
tcgttacacc aagaaagtac cccaagtgtc aactccaact cttgtagagg tctcaagaaa 3180
cctaggaaaa gtgggcagca aatgttgtaa acatcctgaa gcaaaaagaa tgccctgtgc 3240
agaagactat ctatccgtgg tcctgaacca gttatgtgtg ttgcatgaga aaacgccagt 3300
aagtacaga gtcaccaaact gctgcacaga atccttggtg aacaggcgac catgtcgtga 3360
aacctatggt gaaatggctg actgctgtgc aaaacaagaa cctgagagaa atgaatgctt 3420
cttgcaacac aaagatgaca acccaaacct ccccgattg gtgagaccag aggttgatgt 3480

gatgtgcact gcttttcatg acaatgaaga gacctttttg aaaaaatact tatacgaat 3540
 tgccagaaga catccttact tttatgcccc ggaactcctt ttctttgcta aaaggtataa 3600
 agctgctttt acagaatgtt gccaaagctgc tgataaagct gcctgcctgt tgccaaagct 3660
 cgatgaactt cgggatgaag ggaaggcttc gtctgcaaaa cagagactca agtgtgccag 3720
 tctccaaaaa tttggagaaa gagctttcaa agcatgggca gtagctcgcc tgagccagac 3780
 ttttcagctc tggaagtcga tgaaacatac gttcccaaag agtttaatgc tgaaacattc 3840
 accttccatg cagatatatg cacactttct gagaaggaga gacaaatcaa gaaacaaact 3900
 gcacttggtg agctcgtgaa acacaagccc aaggcaacaa aagagcaact gaaagctgtt 3960
 atggatgatt tcgctcgagc catgggcgcg ccatcgatga ggaacccta gtgatggagt 4020
 tggccactcc ctctctgcgc gctcgtcgc tcactgaggc cgggcgacca aaggtcgccc 4080
 gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgagc tgctgcagg 4140

<210> 30

<211> 4575

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-dmScar-longUTR

<400> 30

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgctc 60
 gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggcggaccgc ttgcatgcac 180
 gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
 ccgaccccgga cccactggac aagcacccaa cccccattcc ccaaattgcg catcccctat 300
 cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgcaactgcca gcttcagcac 360
 cgcgagacagt gccttcgccc ccgcctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
 gcgcgctgac gtcactcgcc ggtccccgc aaactcccct tcccggccac cttggtcgcg 480
 tccgcgccgc cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
 ggcgcgacca tctgcgctgc ggcgccggcg actcagcgct gcctcagtct gcggtgggca 600
 gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccaccatggt gagcaagggc 660
 gaggcagtga tcaaggagtt catgcggttc aaggtgcaca tggagggctc catgaacggc 720
 cacgagttcg agatcgaggg cgagggcgag ggccgcccct acgagggcac ccagaccgcc 780
 aagctgaagg tgaccaaggg tggccccctg cctttctcct gggacatcct gtcccctcag 840
 ttcatgtacg gctccagggc cttcatcaag caccgcgcg acatccccga ctactataag 900
 cagtccttcc ccgagggtt caagtgggag cgcgtgatga acttcgagga cgcgcgcgcc 960
 gtgaccgtga cccaggacac ctccctggag gacggcacc tgatctacaa ggtgaagctc 1020
 cgcggcacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080
 gcgtccaccg agcggttgta ccccaggac ggcgtgctga agggcgacat taagatggcc 1140
 ctgcgcctga aggacggcgg ccgctacctg gcggaattca agaccaccta caaggccaag 1200
 aagcccgtgc agatgcccgg cgcctacaac gtcgaccgca agttggacat cacctcccac 1260

aacgaggact acaccgtggt ggaacagtac gaacgctccg agggcccgcca ctccaccggc 1320
 ggcatggacg agctgtacaa gaagcttccg cggagccatg gcttcccggc ggcggtggcg 1380
 gcgcaggatg atggcacgct gcccattgtct tgtgcccagg agagcgggat ggaccgtcac 1440
 cctgcagcct gtgcttctgc taggatcaat gtgtagggtta cctggattta tgttgtatag 1500
 attagattat attgagacaa aaattatcta tttgtatata tacataacag ggtaaattat 1560
 tcagttaaga aaaaaataat tttatgaact gcatgtataa atgaagttaa tacagtacag 1620
 tggttctaca atctatttat tggacatgtc catgaccaga agggaaacag tcatttgcgc 1680
 acaacttaaa aagtctgcat tacattcctt gataatgttg tggtttggtg ccgttgccaa 1740
 gaactgaaaa cataaaaaagt taaaaaaaaat aataaattgc atgctgcttt aattgtgaat 1800
 tgataataaa ctgtcctctt tcagaaaaca gaaaaaaca cacacacaca caacaaaaat 1860
 ttgaacaaaa acattccgtt tacatttttag acagtaagta tcttcgttct tgttagtact 1920
 atatctgttt tactgctttt aacttctgat agcgttggaa ttaaaacaat gtcaaggtgc 1980
 tgttgtcatt gctttactgg cttaggggat gggggatggg ggggtatattt ttgtttgttt 2040
 tgtgtttttt tttcgtttgt ttgttttggt ttttagttcc cacaggagt agagatgggg 2100
 aaagaattcc tacaatatat attctggctg ataaaagata ctttgttatg ttgtgaagat 2160
 gtttgcaata tcgatcagat gactagaaaag tgaataaaaa ttaaggcaac tgaacaaaaa 2220
 aatgctcaca ctccacatcc cgtgatgcac ctcccaggcc ccgctcattc tttgggcgtt 2280
 ggtcagagta agctgctttt gacggaagga cctatgtttg ctcagaacac attctttccc 2340
 cccctcccc tctggtctcc tctttgtttt gttttaagga agaaaaatca gttgcgcgtt 2400
 ctgaaatatt ttaccactgc tgtgaacaag tgaacacatt gtgtcacatc atgacactcg 2460
 tataagcatg gagaacagtg attttttttt agaacagaaa acaacaaaaa ataaccccaa 2520
 aatgaagatt attttttatg aggagtgaac atttgggtta atcatggcta agcttaaaaa 2580
 aaactcatgg tgaggcttaa caatgtcttg taagcaaaag gtagagccct gtatcaacce 2640
 agaaacacct agatcagaac aggaatccac attgccagtg acatgagact gaacagccaa 2700
 atggaggcta tgtggagtgt gcattgcatt taccggcagt gcgggaggaa tttctgagt 2760
 gccatcccaa ggtctagggt gaggtggggc atggtatttg agacattcca aaacgaaggc 2820
 ctctgaagga cccttcagag gtggctctgg aatgacatgt gtcaagctgc ttggacctcg 2880
 tgctttaagt gcctacatta tctaactgtg ctcaagaggt tctcactgg aggaccacac 2940
 tcaagccgac ttatgcccac catcccact ctggataatt ttgcataaaa ttggattagc 3000
 ctggagcagg ttgggagcca aatgtggcat ttgtgatcat gagattgatg caatgagata 3060
 gaagatgttt gctacctgaa cacttattgc tttgaaacta gacttgagga aaccagggtt 3120
 tatcttttga gaacttttgg taagggaaaa gggaacagga aaagaaacc caaactcagg 3180
 ccgaatgatc aaggggaccc ataggaaatc ttgtccagag acaagacttc gggaaggtgt 3240
 ctggacattc agaacaccaa gacttgaagg tgcttgcctc aatggaagag gccaggacag 3300
 agctgacaaa attttgctcc ccagtgaagg ccacagcaac cttctgccc tctgtctgt 3360
 tcatggagag ggtccctgcc tcacctctgc cattttgggt taggagaagt caagttggga 3420
 gcctgaaata gtggttcttg gaaaaatgga tccccagtga aaactagagc tctaagccca 3480
 ttcagcccat ttcacacctg aaaatgttag tgatcaccac ttggaccagc atccttaagt 3540
 atcagaaagc cccaagcaat tgctgcatct tagtagggtg agggataagc aaaagaggat 3600

gttcaccata acccaggaat gaagatacca tcagcaaaga atttcaattt gttcagtcctt 3660
 tcatttagag ctagtccttc acagtacat ctgaatacct ctttgaaaga aggaagactt 3720
 tacgtagtgt agatttggtt tgtgttggtt gaaaatatta tctttgtaat tatttttaaat 3780
 atgtaaggaa tgcttggaat atctgctata tgtcaacttt atgcagcttc cttttgaggg 3840
 acaaatTTAA aacaacaac ccccatcac aaacttaaag gattgcaagg gccagatctg 3900
 ttaagtgggtt tcataggaga cacatccagc aattgtgtgg tcagtggctc ttttaccCAA 3960
 taagatacat cacagtcaca tgcttgatgg tttatgttga cctaagattt attttgTTAA 4020
 aatctctctc tgttggttc gttcttggtc tgTTTTGTTT tgTTTTTTAA agtcttgctg 4080
 tggctctctt gtggcagaag tgtttcatgc atggcagcag gcctgttgct tttttatggc 4140
 gattcccatT gaaaatgtaa gtaaagtgtc gtggccttgt tctctctatg gtaaagatat 4200
 tattcccat gtaaaacaaa aaacaatatt tattgtattt tagtatattt atataattat 4260
 gttattgaaa aaaattggca ttaaaactta accgcatcag aacctattgt aaatacaagt 4320
 tctatttaag tgtactaatt aacatataat atatgtttta aatatagaat ttttaatgtt 4380
 tttaaatata ttttcaaagt acataaaact cgagccatgg gcgcgccatc gatgaggAAC 4440
 ccctagtgat ggagttggcc actccctctc tgcgcgctcg ctcgctcact gaggccgggc 4500
 gaccaaaggT cgcccgacgc ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc 4560
 gcagctgcct gcagg 4575

<210> 31

<211> 4038

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C-dmScar-bGH

<400> 31

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggccgaccgc ttgcatgcac 180
 gcgtagtgca agtgggtttt aggaccagga tgaggcgggg tgggggtgcc tacctgacga 240
 ccgaccccgga cccactggac aagcacccaa ccccatctc ccaaattgcg catccctat 300
 cagagagggg gaggggaaac aggatgcggc gaggcgcgtg cgcaactgcca gcttcagcac 360
 cgccgacagt gccttcgccc ccgctggcg gcgcgcgcca ccgccgcctc agcactgaag 420
 gcgcgctgac gtactcgcc ggtccccgc aaactccct tcccggccac cttggtcgcg 480
 tccgcgccgc cgccggccca gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
 ggcgcgacca tctgcgctgc ggcgccggcg actcagcgt gcctcagtct gcggtgggca 600
 gcggaggagt cgtgtcgtgc ctgagagcgc agaccggtcg ccacatggt gagcaagggc 660
 gaggcagtga tcaaggagtt catgcggtt aaggtgcaca tggagggtc catgaacggc 720
 cacgagttcg agatcgagg cgaggcgag ggccgcccct acgagggcac ccagaccgcc 780
 aagctgaagg tgaccaaggg tggccccctg cccttctctt gggacatcct gtccctcag 840
 ttcatgtacg gctccagggc cttcatcaag cccccgccg acatccccga ctactataag 900

cagtccttcc ccgagggcctt caagtgggag cgcgtgatga acttcgagga cggcggcgcc 960
 gtgaccgtga cccaggacac ctccctggag gacggcacc tgatctacaa ggtgaagctc 1020
 cgcggcacca acttccctcc tgacggcccc gtaatgcaga agaagacaat gggctgggaa 1080
 gcgtccaccg agcggttgta ccccaggac ggcgtgctga agggcgacat taagatggcc 1140
 ctgcgcctga aggacggcgg ccgctacctg gcggacttca agaccaccta caaggccaag 1200
 aagcccgtgc agatgcccgg cgcctacaac gtcgaccgca agttggacat cacctcccac 1260
 aacgaggact acaccgtggt ggaacagtac gaacgctccg agggccgcca ctccaccggc 1320
 ggcattggacg agctgtacaa gaagcttccg cggagccatg gcttcccggc ggcggtggcg 1380
 gcgcaggatg atggcacgct gcccattgtct tgtgcccagg agagcgggat ggaccgtcac 1440
 cctgcagcct gtgcttctgc taggatcaat gtgtagggta ccctgtgcct tctagttgcc 1500
 agccatctgt tgtttgcccc tccccgtgc ctctcttgac cctggaaggt gccactccca 1560
 ctgtcctttc ctaataaaaat gaggaaattg catcgcatg tctgagtagg tgtcattcta 1620
 ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac aatagcaggc 1680
 atgctgggga tgcggtgggc tctatgggcg cgccatcgat gtgaagtggg taacctttat 1740
 ttcccttctt tttctcttta gctcggtta ttccagggt gtgtttcgtc gagatgcaca 1800
 caagagtgag gttgctcatc ggtttaaaga tttgggagaa gaaaatttca aagccttggc 1860
 gttgattgcc tttgctcagt atcttcagca gtgtccattt gaagatcatg taaaattagt 1920
 gaatgaagta actgaatttg caaaaacatg tgttgctgat gagtcatctg aaaattgtga 1980
 caaatcactt catacccttt ttggagacaa attatgcaca gttgcaactc ttcgtgaaac 2040
 ctatggtgaa atggctgact gctgtgcaaa acaagaacct gagagaaatg aatgcttctt 2100
 gcaacacaaa gatgacaacc caaacctccc ccgattgggt agaccagagg ttgatgtgat 2160
 gtgcactgct tttcatgaca atgaagagac ctttttgaaa aaatacttat acgaaattgc 2220
 cagaagacat ccttactttt atgccccgga actcctttt tttgctaaaa ggtataaagc 2280
 tgctttttaca gaatgttgcc aagctgctga taaagctgcc tgcctgttgcc caaagctcga 2340
 tgaacttcgg gatgaaggga aggcttcgtc tgccaaacag agactcaagt gtgccagtct 2400
 ccaaaaattt ggagaaagag ctttcaaagc atgggcagta gctcgcctga gccagagatt 2460
 tcccaaagct gagtttgcag aagtttccaa gttagtgaac gatcttacca aagtccacac 2520
 ggaatgctgc catggagatc tgcttgaatg tgctgatgac agggcggacc ttgccaagta 2580
 tatctgtgaa aatcaagatt cgatctccag taaactgaag gaatgctgtg aaaaacctct 2640
 gttggaaaaa tcccactgca ttgccgaagt ggaaaatgat gagatgcctg ctgacttgcc 2700
 ttcattagct gctgattttg ttgaaagtaa ggatgtttgc aaaaactatg ctgaggcaaa 2760
 ggatgtcttc ctgggcatgt ttttgatatga atatgcaaga aggcacctg attactctgt 2820
 cgtgctgctg ctgagacttg ccaagaccta tgaaaccact ctagagaagt gctgtgccgc 2880
 tgcagatcct catgaatgct atgccaaagt gttcgatgaa tttaaaccctc ttgtggaaga 2940
 gcctcagaat ttaatcaaac aaaattgtga gctttttgag cagcttgag agtacaaatt 3000
 ccagaatgcg ctattagtct gttacaccaa gaaagtaccc caagtgtcaa ctccaactct 3060
 thtagaggct tcaagaaacc taggaaaagt gggcagcaaa tgttgtaaac atcctgaagc 3120
 aaaaagaatg ccctgtgcag aagactatct atccgtggct ctgaaccagt tatgtgtgtt 3180
 gcatgagaaa acgccagtaa gtgacagagt caccaaagc tgcacagaat ccttggtgaa 3240

caggcgacca tgtcgtgaaa cctatggtga aatggctgac tgctgtgcaa aacaagaacc 3300
 tgagagaaat gaatgcttct tgcaacacaa agatgacaac ccaaacctcc cccgattggt 3360
 gagaccagag gttgatgtga tgtgcactgc ttttcatgac aatgaagaga cttttttgaa 3420
 aaaataactta tacgaaattg ccagaagaca tccttacttt tatgccccgg aactcctttt 3480
 ctttgctaaa aggtataaag ctgctttttac agaattgtgc caagctgctg ataaagctgc 3540
 ctgcctgttg ccaaagctcg atgaacttcg ggatgaaggg aaggcttcgt ctgccaaca 3600
 gagactcaag tgtgccagtc tccaaaaatt tggagaaaga gctttcaaag catgggcagt 3660
 agctcgccctg agccagactt ttcagctctg gaatcgatg aaacatacgt tcccaaagag 3720
 tttaatgctg aaacattcac cttccatgca gatatatgca cactttctga gaaggagaga 3780
 caaatcaaga aacaaactgc acttggtgag ctctgtaa acagcccaa ggcaacaaaa 3840
 gagcaactga aagctgttat ggatgatttc gctcgagcca tgggcgcgcc atcgatgagg 3900
 aacccttagt gatggagtgc gccactccct ctctgcgcgc tcgctcgctc actgaggccg 3960
 ggcgacaaa ggtcgcccga cgcccggtt ttgcccgggc ggcctcagt agcgagcgag 4020
 cgcgcagctg cctgcagg 4038

<210> 32

<211> 4287

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> AD-dmScar-shortUTR

<400> 32

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggcctccta ggcggaccgc ttgcatgcac 180
 gcgtgggtgg tgcagagtct gcatgcgtga ggagctcctg ggcgcgtcac agccgcgcta 240
 ttctcagcgt ctctcctttt atggctccgg aagtgcgtg gggttgctgg cagccgtccc 300
 tcgggtccccg caggctgaat gagccccgcc gtccgggggtg ggaggtacgg agcggaggga 360
 ttagtcatcc tggccccgcc atgtgtgcag ggggtgggac ctcccgtccc ttgccgcccg 420
 ggagccgagt gggcgagagg cggggcctgc aggggcgcgg gcggggcgct ggcggggagc 480
 ctgcggctgg gccaatgaga aaccgggctc ggcgagccgg ccgcccacgg gcctcgctgg 540
 ctgcataaag agccggcggc caggactcag cgcagagctc gggcgcgggc tcctccctcc 600
 gcagcagccg agccggacct gcctccccgg gcgtgctccg ccggccccgc cgccggccccg 660
 cagcgacaga caggcgtcc ccgcagctc gcacgggacc caggccgccc gaccccagcg 720
 ccggaccacc ctctgtccgc cccgaggagt ttgccgctg ccggagcacc tgcgcacaga 780
 ccggtcgcca ccatggtgag caagggcgag gcagtgatca aggagtcat gcggttcaag 840
 gtgcacatgg agggctccat gaacggccac gaggctcaga tcgagggcga gggcgagggc 900
 cgcccctacg agggcaccca gaccgccaag ctgaagtgga ccaagggtgg ccccctgccc 960
 ttctcctggg acatcctgtc ccctcagttc atgtacggct ccagggcctt catcaagcac 1020
 cccgccgaca tccccgacta ctataagcag tccttccccg agggcttcaa gtgggagcgc 1080

gtgatgaact tgcaggacgg cggcgccgtg accgtgaccc aggacacctc cctggaggac 1140
 ggcaccctga tctacaaggt gaagctccgc ggcaccaact tccctcctga cggccccgta 1200
 atgcagaaga agacaatggg ctgggaagcg tccaccgagc ggttgtagcc cgaggacggc 1260
 gtgctgaagg gcgacattaa gatggccctg cgcctgaagg acggcgggcg ctacctggcg 1320
 gacttcaaga ccacctacaa ggccaagaag cccgtgcaga tgcccggcgc ctacaacgtc 1380
 gaccgcaagt tggacatcac ctcccacaac gaggactaca ccgtggtgga acagtacgaa 1440
 cgctccgagg gccgccactc caccggcggc atggacgagc tgtacaagaa gcttccgcgg 1500
 agccatggct tcccgcggcg ggtggcgggc caggatgatg gcacgctgcc catgtcttgt 1560
 gcccgaggaga gcgggatgga ccgtcaccct gcagcctgtg cttctgctag gatcaatgtg 1620
 tagggtacct ggatttatgt tgtatagatt agattatatt gagacaaaaa ttatctatct 1680
 gtatatatac ataacagggt aaattattca gttaagaaaa aaataatttt atgaactgca 1740
 tgtataaatg aagtttatac agtacagtgg ttctacaatc tatttattgg acatgtccat 1800
 gaccagaagg gaaacagtca tttgcgcaca acttaaaaag tctgcattac attccttgat 1860
 aatgttggtg tttgttgccg ttgccaagaa ctgaaaacat aaaaagttaa aaaaaataat 1920
 aaattgcatg ctgctttaat tgtgaattga taataagcgc gccatcgatg tgaagtgggt 1980
 aacctttatt tcccttcttt ttctctttag ctcggttat tccaggggtg tgtttcgtcg 2040
 agatgcacac aagagtgagg ttgctcatcg gtttaaatg ttgggagaag aaaatttcaa 2100
 agccttggtg ttgattgcct ttgctcagta tcttcagcag tgtccatttg aagatcatgt 2160
 aaaattagtg aatgaagtaa ctgaatttgc aaaaacatgt gttgctgatg agtcagctga 2220
 aaattgtgac aaatcacttc ataccctttt tggagacaaa ttatgcacag ttgcaactct 2280
 tcgtgaaacc tatggtgaaa tggctgactg ctgtgcaaaa caagaacctg agagaaatga 2340
 atgcttcttg caacacaaag atgacaaccc aaacctcccc cgattggtga gaccagaggt 2400
 tgatgtgatg tgcactgctt ttcatgacaa tgaagagacc ttttgaaaaa aatacttata 2460
 cgaaattgcc agaagacatc ctacttttta tgccccggaa ctctttttct ttgctaaaag 2520
 gtataaagct gctttttacag aatgttgcca agctgctgat aaagctgcct gcctgttgcc 2580
 aaagctcgat gaacttcggg atgaaggga ggcttcgtct gccaaacaga gactcaagtg 2640
 tgccagtcct caaaaatttg gagaaagagc tttcaaagca tgggcagtag ctgcctgag 2700
 ccagagattt cccaaagctg agtttgaga agtttccaag ttagtgacag atcttaccaa 2760
 agtccacacg gaatgtgcc atggagatct gcttgaatgt gctgatgaca gggcgacact 2820
 tgccaagtat atctgtgaaa atcaagattc gatctccagt aaactgaagg aatgctgtga 2880
 aaaacctctg ttggaaaaat cccactgcat tgccgaagtg gaaaatgatg agatgcctgc 2940
 tgacttgctt tcattagctg ctgattttgt tgaaagtaag gatgtttgca aaaactatgc 3000
 tgaggcaaaag gatgtcttcc tgggcatgtt tttgtatgaa tatgcaagaa ggcactcctga 3060
 ttactctgtc gtgtgctgc tgagacttgc caagacctat gaaaccactc tagagaagtg 3120
 ctgtgccgct gcagatcctc atgaatgcta tgccaaagtg ttcgatgaat ttaaacctct 3180
 tgtggaagag cctcagaatt taatcaaaaa aaattgtgag ctttttgagc agcttgagga 3240
 gtacaaattc cagaatgcgc tattagtctg ttacaccaag aaagtacccc aagtgtcaac 3300
 tccaactctt gtagaggtct caagaaacct aggaaaagtg ggcagcaaat gttgtaaaaa 3360
 tcctgaagca aaaagaatgc cctgtgcaga agactatcta tccgtggtcc tgaaccagtt 3420

atgtgtgttg catgagaaaa cgccagtaag tgacagagtc accaaatgct gcacagaatc 3480
 cttggtgaac aggcgaccat gtcgtgaaac ctatggtgaa atggctgact gctgtgcaaa 3540
 acaagaacct gagagaaatg aatgcttctt gcaacacaaa gatgacaacc caaacctccc 3600
 ccgattggtg agaccagagg ttgatgtgat gtgcactgct tttcatgaca atgaagagac 3660
 ctttttgaaa aaatacttat acgaaattgc cagaagacat cttactttt atgccccgga 3720
 actccttttc ttgctaaaaa ggtataaagc tgcttttaca gaatgttgcc aagctgctga 3780
 taaagctgcc tgccgtgtgc caaagctcga tgaacttcgg gatgaaggga aggcttcgtc 3840
 tgccaaacag agactcaagt gtgccagtct ccaaaaattt ggagaaagag ctttcaaagc 3900
 atgggcagta gctcgctga gccagacttt tcagctctgg aagtcgatga aacatacgtt 3960
 cccaaagagt ttaatgctga aacattcacc ttccatgcag atatatgcac actttctgag 4020
 aaggagagac aaatcaagaa acaaactgca cttgttgagc tcgtgaaaca caagcccaag 4080
 gcaacaaaag agcaactgaa agctgttatg gatgatttcg ctcgagccat gggcgcgcca 4140
 tcgatgagga acccctagtg atggagttgg ccactccctc tctgcgcgct cgctcgctca 4200
 ctgaggccgg gcgaccaaag gtcgcccagc gcccgggctt tgccccggcg gcctcagtga 4260
 gcgagcgagc gcgcagctgc ctgcagg 4287

<210> 33

<211> 4722

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> AD-dmScar-longUTR

<400> 33

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
 gggcgacctt tggctgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggctccta ggcggaccgc ttgcatgcac 180
 gcgtgggtgg tgcagagtct gcatgcgtga ggagctcctg ggcgcgctac agccgcgcta 240
 ttctcagcgt ctctcctttt atggctccgg aagtgcgctg gggttgctgg cagccgtccc 300
 tcgggtcccc caggctgaat gagccccgcc gtccgggggtg ggaggtacgg agcggaggga 360
 ttagtcatcc tgccccgcc atgtgtgcag ggggtgggac ctcccgtccc ttgccgcccc 420
 ggagccgagt gggcgagagg cggggcctgc aggggcgcgg gcggggcgct ggcggggagc 480
 ctgcggctgg gccaatgaga aaccgggctc ggcgagccgg ccgcccacgg gcctcgctgg 540
 ctgcataaag agccggcggc caggactcag cgcagagctc gggcgcgggc tcctccctcc 600
 gcagcagccg agccggacct gcctccccgg gcgtgctccg ccggccccgc cgccggcccc 660
 cagcgacaga caggcgctcc ccgcagctcc gcacgggacc caggccgccc gaccccagcg 720
 ccggaccacc ctctgtccgc cccgaggagt ttgccgctg ccggagcacc tgcgcacaga 780
 ccggctgcca ccatggtgag caagggcgag gcagtgatca aggagtcat gcggttcaag 840
 gtgcacatgg agggctccat gaacggccac gatttcgaga tcgagggcga gggcgagggc 900
 cgcccctacg agggcaccca gaccgccaag ctgaaggtga ccaagggtgg ccccctgccc 960
 ttctcctggg acatcctgtc ccctcagttc atgtacggct ccagggcctt catcaagcac 1020

```

ccccgccgaca tccccgacta ctataagcag tccttccccg agggcttcaa gtgggagcgc 1080
gtgatgaact tcgaggacgg cggcgccgtg accgtgaccc aggacacctc cctggaggac 1140
ggcacccctga tctacaaggt gaagctccgc ggcaccaact tccctcctga cggccccgta 1200
atgcagaaga agacaatggg ctgggaagcg tccaccgagc ggttgtagcc cgaggacggc 1260
gtgctgaagg gcgacattaa gatggccctg cgcctgaagg acggcggccg ctacctggcg 1320
gacttcaaga ccacctacaa ggccaagaag cccgtgcaga tgcccggcgc ctacaacgtc 1380
gaccgcaagt tggacatcac ctcccacaac gaggactaca ccgtggtgga acagtacgaa 1440
cgctccgagg gccgccactc caccggcggc atggacgagc tgtacaagaa gcttccgagg 1500
agccatggct tcccgcgggc ggtggcggcg caggatgatg gcacgtgcc catgtcttgt 1560
gcccaggaga gcgggatgga ccgtcacctc gcagcctgtg cttctgctag gatcaatgtg 1620
tagggtacct ggatttatgt tgtatagatt agattatatt gagacaaaaa ttatctattt 1680
gtatatatac ataacagggt aaattattca gttaagaaaa aaataatfff atgaactgca 1740
tgtataaatg aagtttatac agtacagtgg ttctacaatc tatttattgg acatgtccat 1800
gaccagaagg gaaacagtca tttgcgcaca acttaaaaag tctgcattac attccttgat 1860
aatgttgtgg tttgttgccg ttgccaagaa ctgaaaacat aaaaagttaa aaaaaataat 1920
aaattgcatg ctgctttaat tgtgaattga taataaactg tcctctttca gaaaacagaa 1980
aaaaacacac acacacacaa caaaaatttg aaccaaaca ttccgtttac attttagaca 2040
gtaagtatct tcgttcttgt tagtactata tctgttttac tgcttttaac ttctgatagc 2100
gttggaatta aaacaatgtc aaggtgctgt tgtcattgct ttactggctt aggggatggg 2160
ggatgggggg tataatffff tttgttttgt gttttttttt cgtttgtttg ttttgttttt 2220
tagttccac agggagtaga gatggggaaa gaattcctac aatatatatt ctggctgata 2280
aaagatacat ttgtatgttg tgaagatgtt tgcaatatcg atcagatgac tagaaagtga 2340
ataaaaaatta aggcaactga acaaaaaaat gtcacactc cacatcccgt gatgcacctc 2400
ccaggccccg ctcatctttt gggcgtttgt cagagtaagc tgcttttgac ggaaggacct 2460
atgtttgctc agaacacatt ctttcccccc ctccccctct ggtctcctct ttgttttgtt 2520
ttaaggaaga aaaatcagtt gcgcgttctg aaatatfita ccactgctgt gaacaagtga 2580
acacattgtg tcacatcatg acactcgtat aagcatggag aacagtgatt tttttttaga 2640
acagaaaaaca acaaaaaata accccaaaat gaagattatt ttttatgagg agtgaacatt 2700
tgggtaaatac atggctaagc ttaaaaaaaaa ctcatggtga ggcttaacaa tgtcttgtaa 2760
gcaaaaggta gagccctgta tcaaccaga aacacctaga tcagaacagg aatccacatt 2820
gccagtgaca tgagactgaa cagccaaatg gaggtatgt ggagtggca ttgcatttac 2880
cggcagtgcg ggaggaatff ctgagtggcc atcccaaggt ctaggtggag gtggggcatg 2940
gtatfftgaga cattccaaaa cgaaggctc tgaaggacct ttcagagggt gctctggaat 3000
gacatgtgtc aagctgcttg gacctcgtgc tttaagtgc tacattatct aactgtgctc 3060
aagaggttct cgactggagg accacactca agccgactta tgcccacat cccacctctg 3120
gataatffff cataaaattg gattagcctg gagcaggttg ggagccaaat gtggcatttg 3180
tgatcatgag attgatgcaa tgagatagaa gatgtttgct acctgaacac ttattgctff 3240
gaaactagac ttgaggaaac cagggtttat cttttgagaa cttttggtta gggaaaaggg 3300
aacaggaaaa gaaaccccaa actcaggccg aatgatcaag gggaccata ggaaatcttg 3360

```

tccagagaca agacttcggg aaggtgtctg gacattcaga acaccaagac ttgaaggtgc 3420
 cttgctcaat ggaagaggcc aggacagagc tgacaaaatt ttgctcccca gtgaaggcca 3480
 cagcaacctt ctgcccattc tgtctgttca tggagagggt ccctgcctca cctctgccat 3540
 tttgggttag gagaagtcaa gttgggagcc tgaaatagtg gttcttgga aaatggatcc 3600
 ccagtgaaaa ctagagctct aagcccattc agcccatttc acacctgaaa atgttagtga 3660
 tcaccacttg gaccagcatc cttaagtatc agaaagcccc aagcaattgc tgcattctag 3720
 tagggtaggg gataagcaaa agaggatgtt caccataacc caggaatgaa gataccatca 3780
 gcaaagaatt tcaatttggt cagtctttca tttagagcta gtctttcaca gtaccatctg 3840
 aataacctct tgaaagaagg aagactttac gtagtgtaga tttgttttgt gttgtttgaa 3900
 aatattatct ttgtaattat ttttaatatg taaggaatgc ttggaatata tgctatatgt 3960
 caactttatg cagcttcctt ttgagggaca aatttaaaac aaacaacccc ccatcacaaa 4020
 cttaaaggat tgcaagggcc agatctgtta agtggtttca taggagacac atccagcaat 4080
 tgtgtggtca gtggctcttt tacccaataa gatacatcac agtcacatgc ttgatggttt 4140
 atgttgacct aagatttatt ttgttaaaat ctctctctgt tgtgttcgtt cttgttctgt 4200
 tttgttttgt tttttaaaagt cttgctgtgg tctctttgtg gcagaagtgt ttcatgcatg 4260
 gcagcaggcc tgttgctttt ttatggcgat tcccattgaa aatgtaagta aatgtctgtg 4320
 gccttgttct ctctatggta aagatattat tcaccatgta aaacaaaaaa caatatttat 4380
 tgtatttttag tataattata taattatgtt attgaaaaaa attggcatta aaacttaacc 4440
 gcatcagaac ctattgtaaa tacaagttct atttaagtgt actaattaac atataatata 4500
 tgttttaaat atagaatttt taatgttttt aaatatattt tcaaagtaca taaaactcga 4560
 gccatgggcg cgccatcgat gaggaacccc tagtgatgga gttggccact ccctctctgc 4620
 gcgctcgctc gctcactgag gccgggcgac caaaggtcgc ccgacgcccg ggctttgccc 4680
 gggcggcctc agtgagcgag cgagcgcgca gctgcctgca gg 4722

<210> 34

<211> 4185

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> AD-dmScar-bGH构建体

<400> 34

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgctc 60
 gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
 actccatcac taggggttcc ttgcgtcgac aggctccta ggcggaaccgc ttgcatgcac 180
 gcgtgggtgg tgcagagtct gcatgcgtga ggagctcctg ggcgcgtcac agccgcgcta 240
 ttctcagcgt ctctcctttt atggctccgg aagtgagctg gggttgctgg cagecgtccc 300
 tcgggtcccc caggctgaat gagccccgcc gtccgggggtg ggaggtacgg agcggaggga 360
 ttagtcatcc tggccccgcc atgtgtgcag ggggtgggac ctcccgtccc ttgccgcccc 420
 ggagccgagt gggcgagagg cggggcctgc aggggcgcgg gcggggcgct ggcggggagc 480
 ctgcggctgg gccaatgaga aaccgggctc ggcgagccgg ccgcccacgg gcctcgttgg 540

```

ctgcataaag agccggcggc caggactcag cgcagagctc gggcgcgggc tcctccctcc 600
gcagcagccg agccggacct gcctccccgg gcgtgctccg ccggccccgc cgccggcccc 660
cagcgacaga caggcgctcc ccgcagctcc gcacgggacc caggccgccc gaccccagcg 720
ccggaccacc ctctgtccgc cccgaggagt ttgccgcctg ccggagcacc tgcgcacaga 780
ccggtcgcca ccatggtgag caagggcgag gcagtgatca aggagtcat gcggttcaag 840
gtgcacatgg agggctccat gaacggccac gatttcgaga tcgagggcga gggcgagggc 900
cgccccctac agggcaccca gaccgccaag ctgaaggatga ccaagggtgg cccctgccc 960
ttctcctggg acatcctgtc cctcagttc atgtacggct ccagggcctt catcaagcac 1020
cccgcgaca tccccgacta ctataagcag tccttccccg agggcttcaa gtgggagcgc 1080
gtgatgaact tcgaggacgg cggcgccgtg accgtgacct aggacacct cctggaggac 1140
ggcacctga tctacaaggt gaagctccgc ggcaccaact tcctcctga cggccccgta 1200
atgcagaaga agacaatggg ctgggaagcg tccaccgagc ggttgtagcc cgaggacggc 1260
gtgctgaagg gcgacattaa gatggcctg ccctgaagg acggcgccg ctacctggcg 1320
gacttcaaga ccacctacaa ggccaagaag cccgtgcaga tgcccggcgc ctacaacgtc 1380
gaccgcaagt tggacatcac ctcccacaac gaggactaca ccgtggtgga acagtacgaa 1440
cgctccgagg gccgccactc caccggcggc atggacgagc tgtacaagaa gcttccgcgg 1500
agccatggct tcccgccggc ggtggcggcg caggatgatg gcacgtgcc catgtcttgt 1560
gcccaggaga gcgggatgga ccgtcaccct gcagcctgtg cttctgctag gatcaatgtg 1620
tagggtaccc tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt 1680
ccttgaccct ggaagggtgc actccactg tcctttccta ataaatgag gaaattgcat 1740
cgcatgtct gagtagtgt cattctattc tggggggtgg ggtggggcag gacagcaagg 1800
gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct atgggcgcgc 1860
catcgatgtg aagtgggtaa cctttatttc cttcttttt ctcttttagct cggcttattc 1920
caggggtgtg tttcgtcgag atgcacacaa gattgaggtt gctcatcggt ttaaagattt 1980
gggagaagaa aatttcaaag ccttggtgtt gattgccttt gctcagtatc ttcagcagtg 2040
tccatttgaa gatcatgtaa aattagttaa tgaagtaact gaatttgcaa aaacatgtgt 2100
tgctgatgag tcagctgaaa attgtgacaa atcacttcat accctttttg gagacaaatt 2160
atgcacagtt gcaactcttc gtgaaaccta tggtgaaatg gctgactgct gtgcaaaaca 2220
agaacctgag agaaatgaat gcttcttgca acacaaagat gacaaccaa acctcccccg 2280
attggtgaga ccagaggttg atgtgatgtg cactgctttt catgacaatg aagagacctt 2340
tttgaaaaaa tacttatacg aaattgccag aagacatcct tacttttatg ccccggaact 2400
ccttttcttt gctaaaaggt ataaagctgc ttttacagaa tgttgccaag ctgctgataa 2460
agctgcctgc ctgttgccaa agctcgatga acttcgggat gaagggaagg cttcgtctgc 2520
caaacagaga ctcaagtgtg ccagtctcca aaaatttgga gaaagagctt tcaaagcatg 2580
ggcagtagct cgctgagcc agagatttcc caaagctgag tttgcagaag tttccaagtt 2640
agtacagat cttaacaaag tccacacgga atgtgccaat ggagatctgc ttgaatgtgc 2700
tgatgacagg gcggaccttg ccaagtatat ctgtgaaaat caagattcga tctccagtaa 2760
actgaaggaa tgctgtgaaa aacctctgtt ggaaaaatcc cactgcattg ccgaagtgga 2820
aaatgatgag atgcctgctg acttgccctc attagctgct gatthttgtt aaagtaagga 2880

```

tgttttcaaa aactatgctg aggcaaagga tgtcttcctg ggcatgtttt tgtatgaata 2940
 tgcaagaagg catcctgatt actctgtcgt gctgctgctg agacttgcca agacctatga 3000
 aaccactcta gagaagtgtg gtgccgctgc agatcctcat gaatgctatg ccaaagtgtt 3060
 cgatgaattt aaacctcttg tggaagagcc tcagaattta atcaaacaaa attgtgagct 3120
 ttttgagcag cttggagagt acaaattcca gaatgcgcta ttagttcgtt acaccaagaa 3180
 agtaccceaa gtgtcaactc caactcttgt agaggtctca agaaacctag gaaaagtggg 3240
 cagcaaatgt tgtaaacatc ctgaagcaaa aagaatgccc tgtgcagaag actatctatc 3300
 cgtggtcctg aaccagttat gtgtgttgca tgagaaaacg ccagtaagtg acagagtcac 3360
 caaatgctgc acagaatcct tgggtgaacag gcgaccatgt cgtgaaacct atggtgaaat 3420
 ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa tgcttcttgc aacacaaaga 3480
 tgacaacca aacctcccc gattggtgag accagaggtt gatgtgatgt gcaactgttt 3540
 tcatgacaat gaagagacct ttttgaaaaa ataacttatac gaaattgcca gaagacatcc 3600
 ttacttttat gccccggaac tccttttctt tgctaaaagg tataaagctg cttttacaga 3660
 atgttgccaa gctgctgata aagctgcctg cctgttgcca aagctcgatg aacttcggga 3720
 tgaagggaag gcttcgtctg ccaaacagag actcaagtgt gccagtctcc aaaaatttgg 3780
 agaaagagct ttcaaagcat gggcagtagc tcgcctgagc cagacttttc agctctggaa 3840
 gtcgatgaaa catacgttcc caaagagttt aatgctgaaa cattcacctt ccatgcagat 3900
 atatgcacac tttctgagaa ggagagacaa atcaagaaac aaactgcact tgttgagctc 3960
 gtgaaacaca agcccaaggc aacaaaagag caactgaaag ctgttatgga tgatttcgct 4020
 cgagccatgg gcgcgccatc gatgaggaac ccctagtgat ggagttggcc actccctctc 4080
 tgcgcgctcg ctgcctcact gaggccgggc gaccaaaggt cgcccgacgc ccgggctttg 4140
 cccgggcggc ctgagtgagc gagcgagcgc gcagctgcct gcagg 4185

<210> 35

<211> 831

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> dmScarlet (具有C末端PEST降解决定子的mScarlet-I)

<400> 35

atggtgagca agggcgaggc agtgatcaag gagttcatgc ggttcaaggt gcacatggag 60
 ggctccatga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 120
 ggcacccaga ccgccaagct gaaggtgacc aagggtggcc ccctgccctt ctcttgggac 180
 atcctgtccc ctgagttcat gtacggtctc agggcettca tcaagcacc cgcgcacatc 240
 cccgactact ataagcagtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc 300
 gaggacggcg gcgccgtgac cgtgaccag gacacctccc tggaggacgg caccctgac 360
 tacaagtgta agctccgcgg caccaacttc ctcctgacg gccccgtaat gcagaagaag 420
 acaatgggct gggaagcgtc caccgagcgg ttgtacccc aggacggcgt gctgaagggc 480
 gacattaaga tggccctgcg cctgaaggac ggcgccgct acctggcgga cttcaagacc 540
 acctacaagg ccaagaagcc cgtgcagatg cccggcgcct acaacgtcga ccgcaagttg 600

gacatcacct cccacaacga ggactacacc gtggtggaac agtacgaacg ctccgagggc 660
cgccactcca ccggcggcat ggacgagctg tacaagaagc ttccgcggag ccatggcttc 720
ccgccggcgg tggcggcgca ggatgatggc acgctgccc tgtcttgtgc ccaggagagc 780
gggatggacc gtcaccctgc agcctgtgct tctgctagga tcaatgtgta g 831

<210> 36

<211> 696

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> mScarlet-I

<400> 36

atggtgagca agggcgaggc agtgatcaag gattcatgc ggttcaaggt gcacatggag 60
ggctccatga acggccacga gttcgagatc gagggcgagg gcgagggccg cccctacgag 120
ggcaccacaga ccgccaagct gaagtgacc aagggtggcc ccctgccctt ctcttgggac 180
atcctgtccc ctcatgtcat gtacggtcc agggccttca tcaagcacc cgccgacatc 240
cccgactact ataagcagtc cttccccgag ggcttcaagt gggagcgcgt gatgaacttc 300
gaggacggcg gcgccgtgac cgtgaccag gacacctccc tggaggacgg caccctgac 360
tacaaggtga agctccgcgg caccaacttc cctcctgacg gccccgtaat gcagaagaag 420
acaatgggct gggaagcgtc caccgagcgg ttgtacccc aggacggcgt gctgaagggc 480
gacattaaga tggccctgcg cctgaaggac ggcgccgct acctggcgga cttcaagacc 540
acctacaagg ccaagaagcc cgtgcagatg cccggcgctt acaacgtcga ccgcaagttg 600
gacatcacct cccacaacga ggactacacc gtggtggaac agtacgaacg ctccgagggc 660
cgccactcca ccggcggcat ggacgagctg tacaag 696

<210> 37

<211> 135

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> C末端PEST降解决定子

<400> 37

aagcttccgc ggagccatgg cttcccgcg gcggtggcgg cgcaggatga tggcacgctg 60
cccatgtctt gtgcccagga gagcgggatg gaccgtcacc ctgcagcctg tgcttctgct 120
aggatcaatg tgtag 135

<210> 38

<211> 448

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> hSyn

<400> 38

```
agtgcaagtg ggttttagga ccaggatgag gcggggtggg ggtgcctacc tgacgaccga 60
ccccgaccca ctggacaagc acccaacccc cattcccaa attgcgcac ccctatcaga 120
gagggggagg ggaaacagga tgcggcgagg cgcgtgcgca ctgccagctt cagcaccgcg 180
gacagtgcct tcgccccgc ctggcggcgc gcgccaccgc cgcctcagca ctgaaggcgc 240
gctgacgtca ctcgccggtc ccccgcaaac tccccttccc ggccaccttg gtcgcgtccg 300
cgccgccgcc ggcccagccg gaccgcacca cgcgaggcgc gagatagggg ggcacgggcg 360
cgaccatctg cgctgcggcg ccggcgactc agcgtgcct cagtctgcgg tgggcagcgg 420
aggagtcgtg tcgtgcctga gagcgcag 448
```

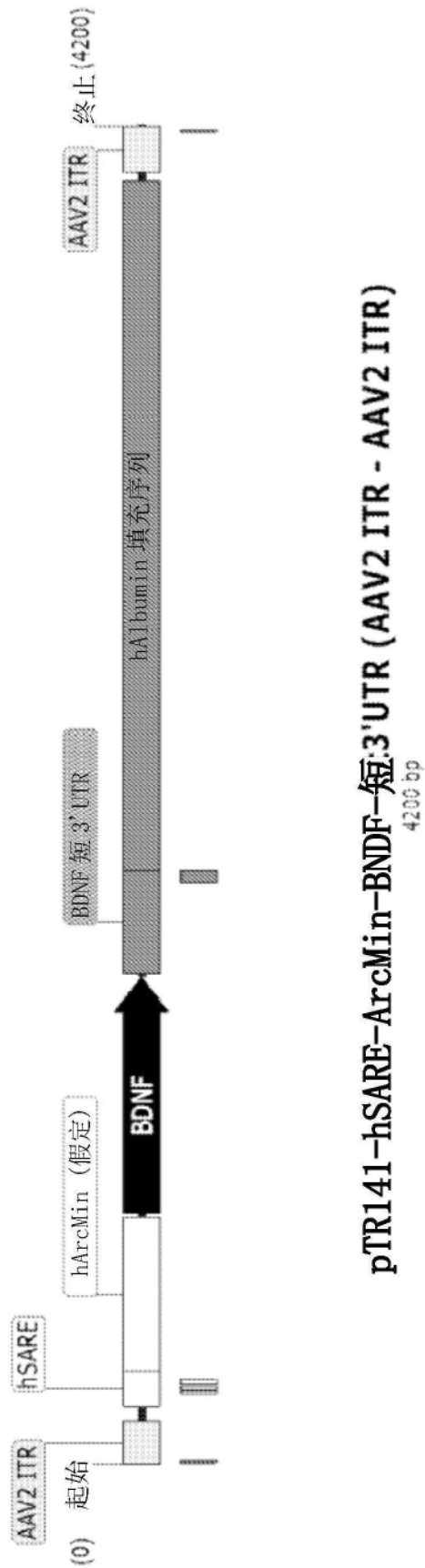


图1

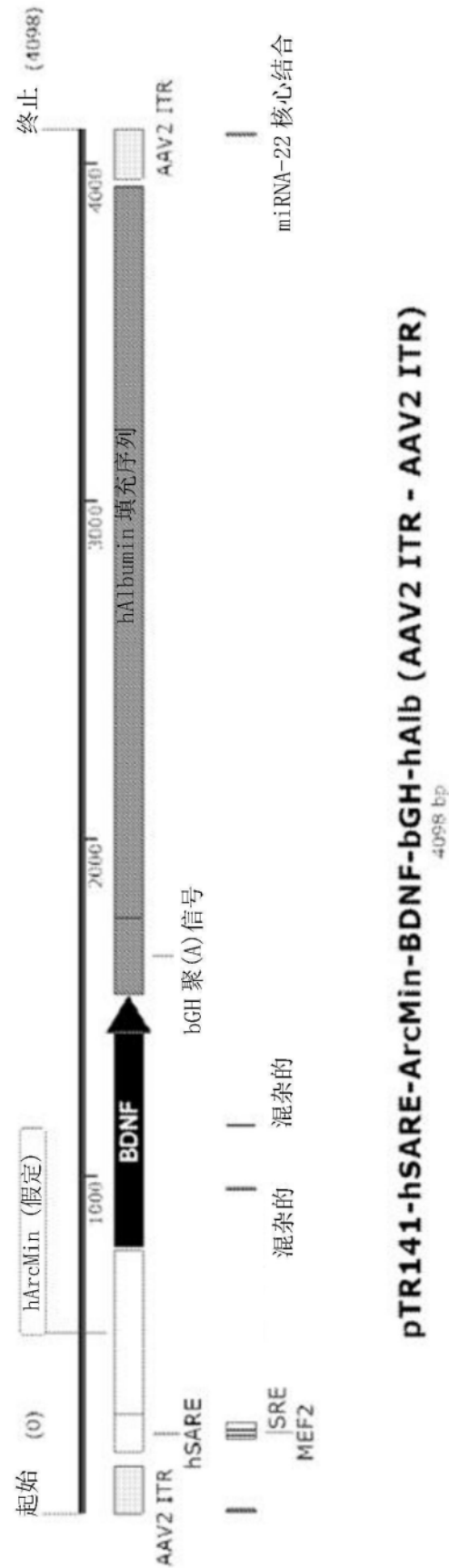


图2

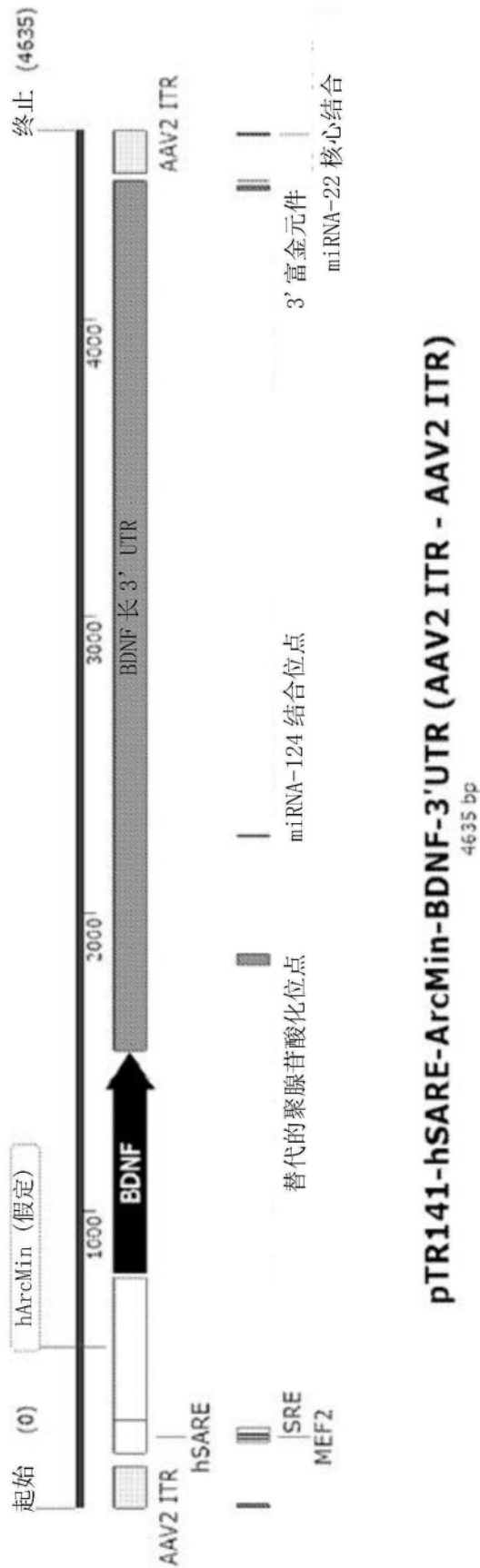


图3

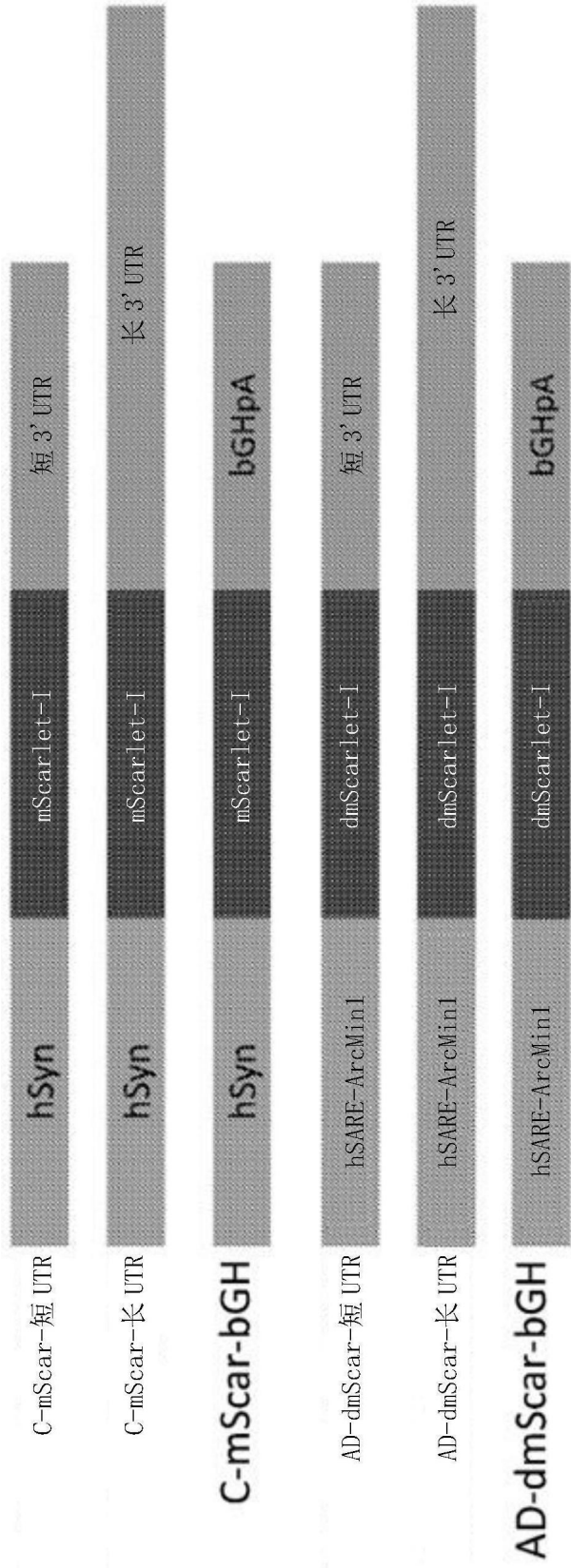


图4

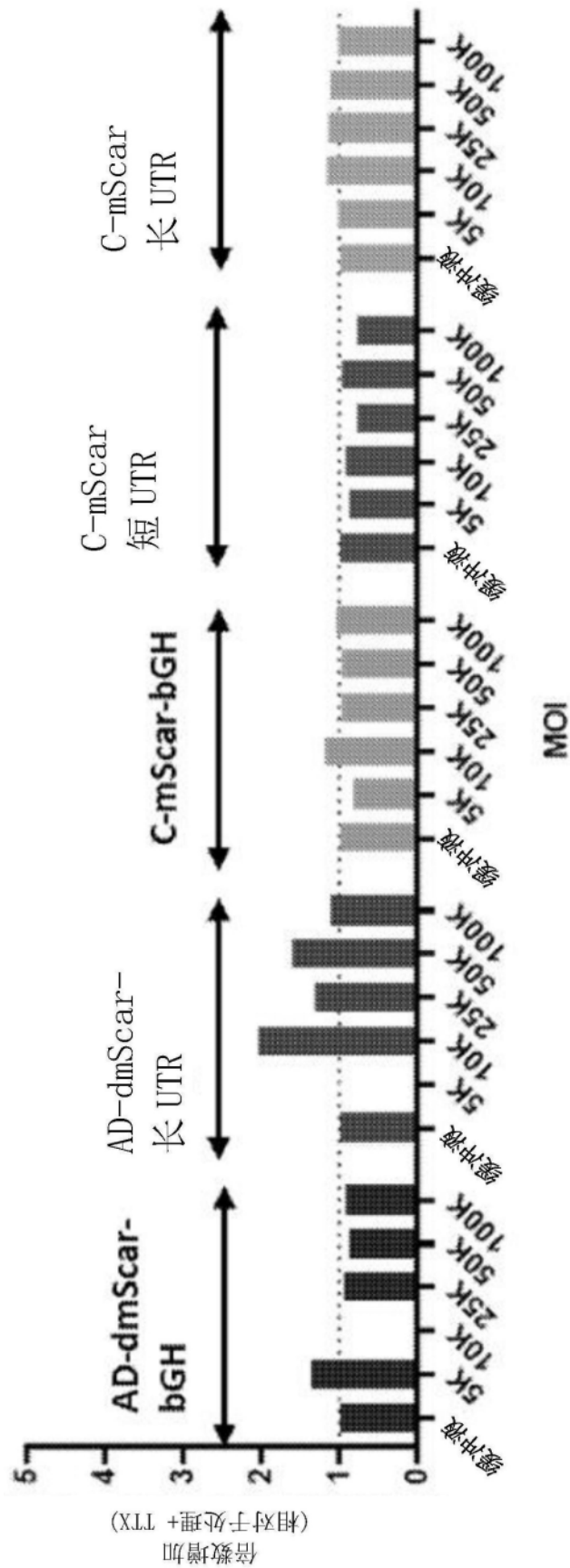


图5A

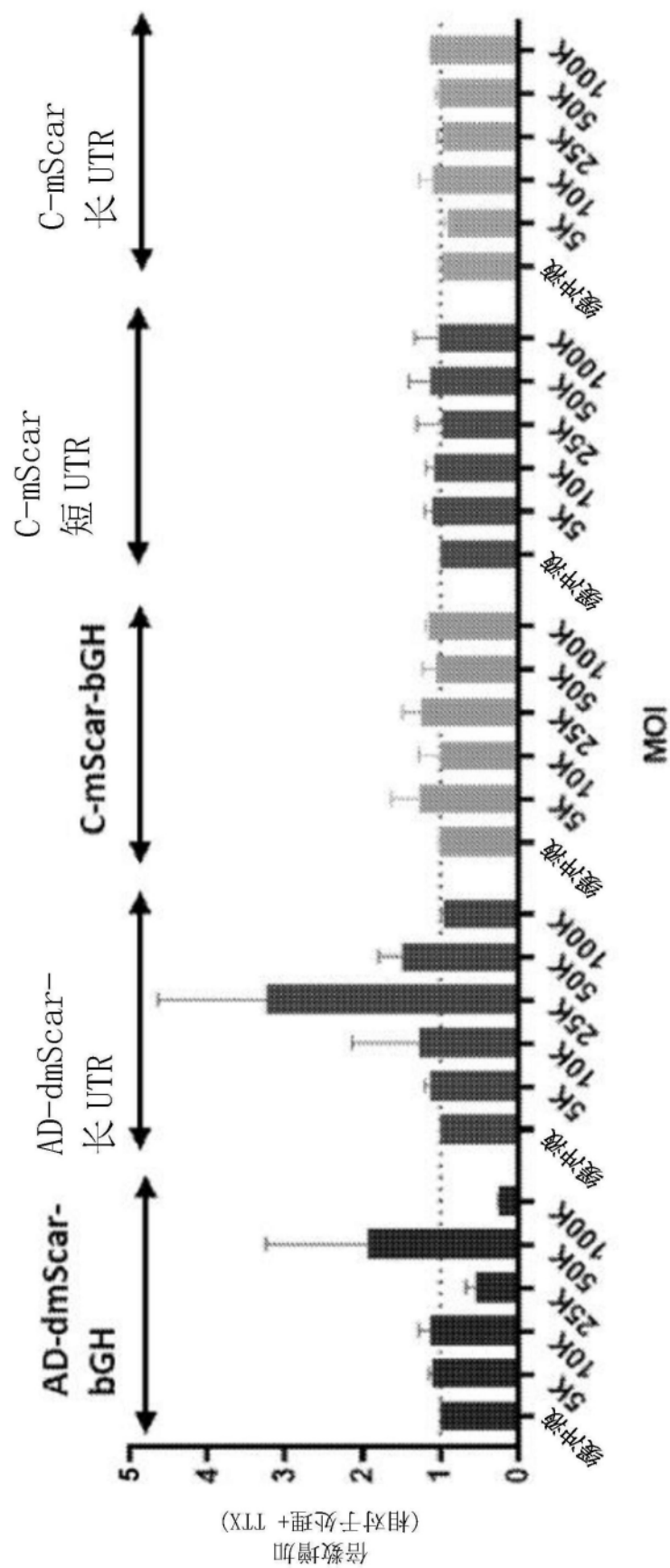


图5B

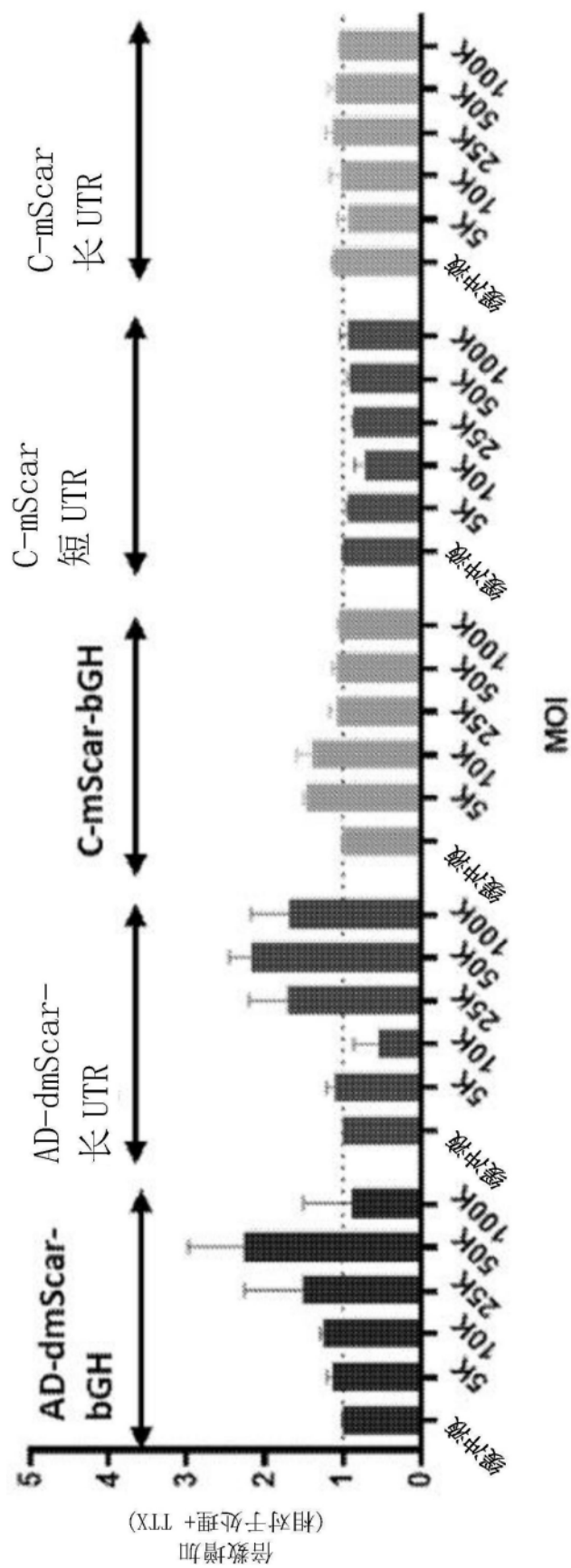


图5C

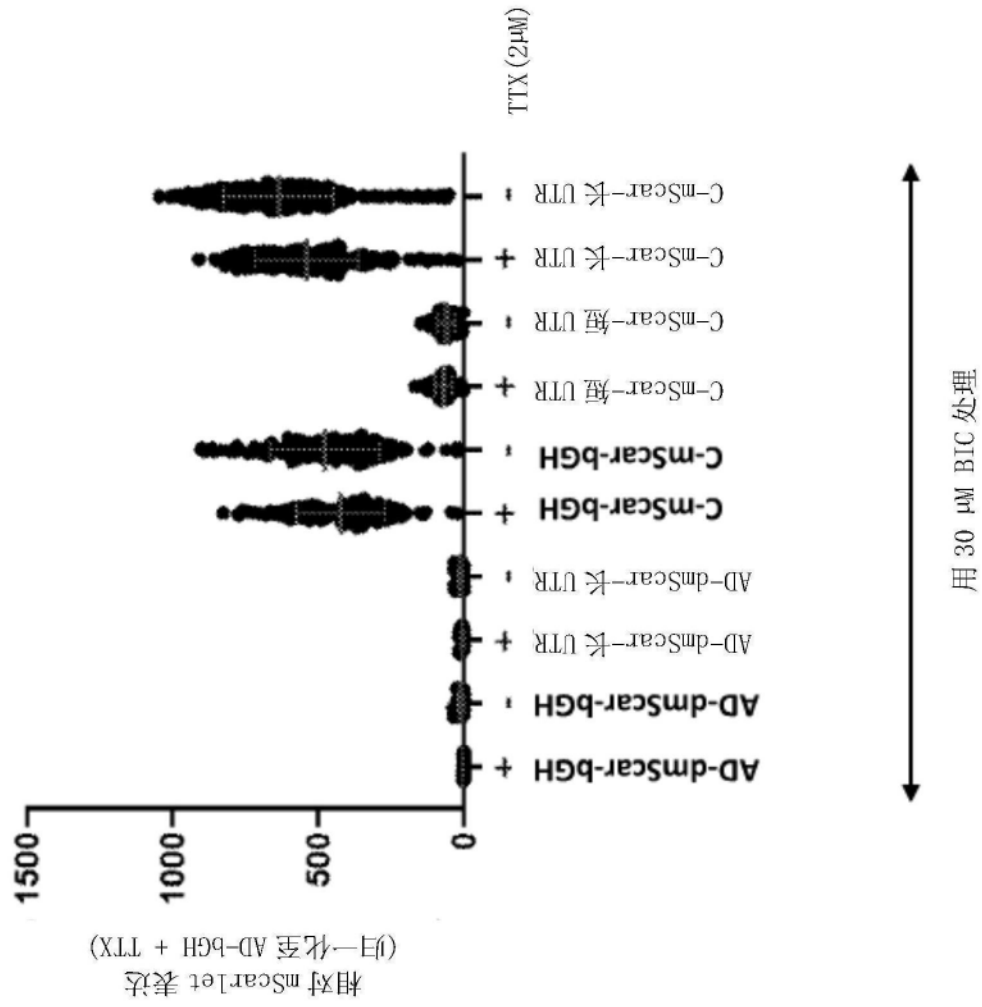


图6A

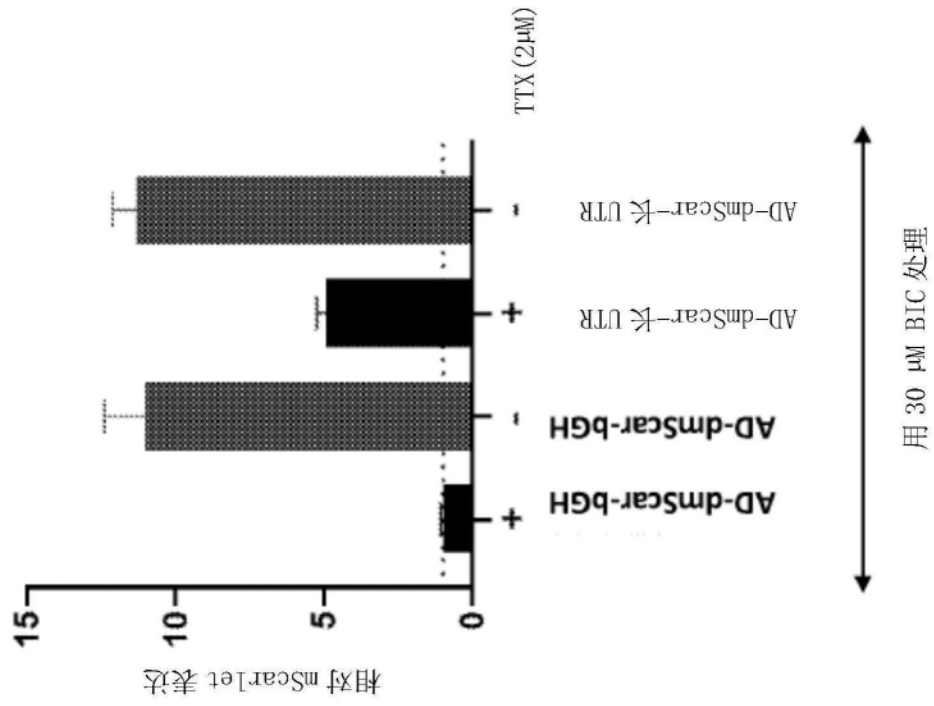


图6B

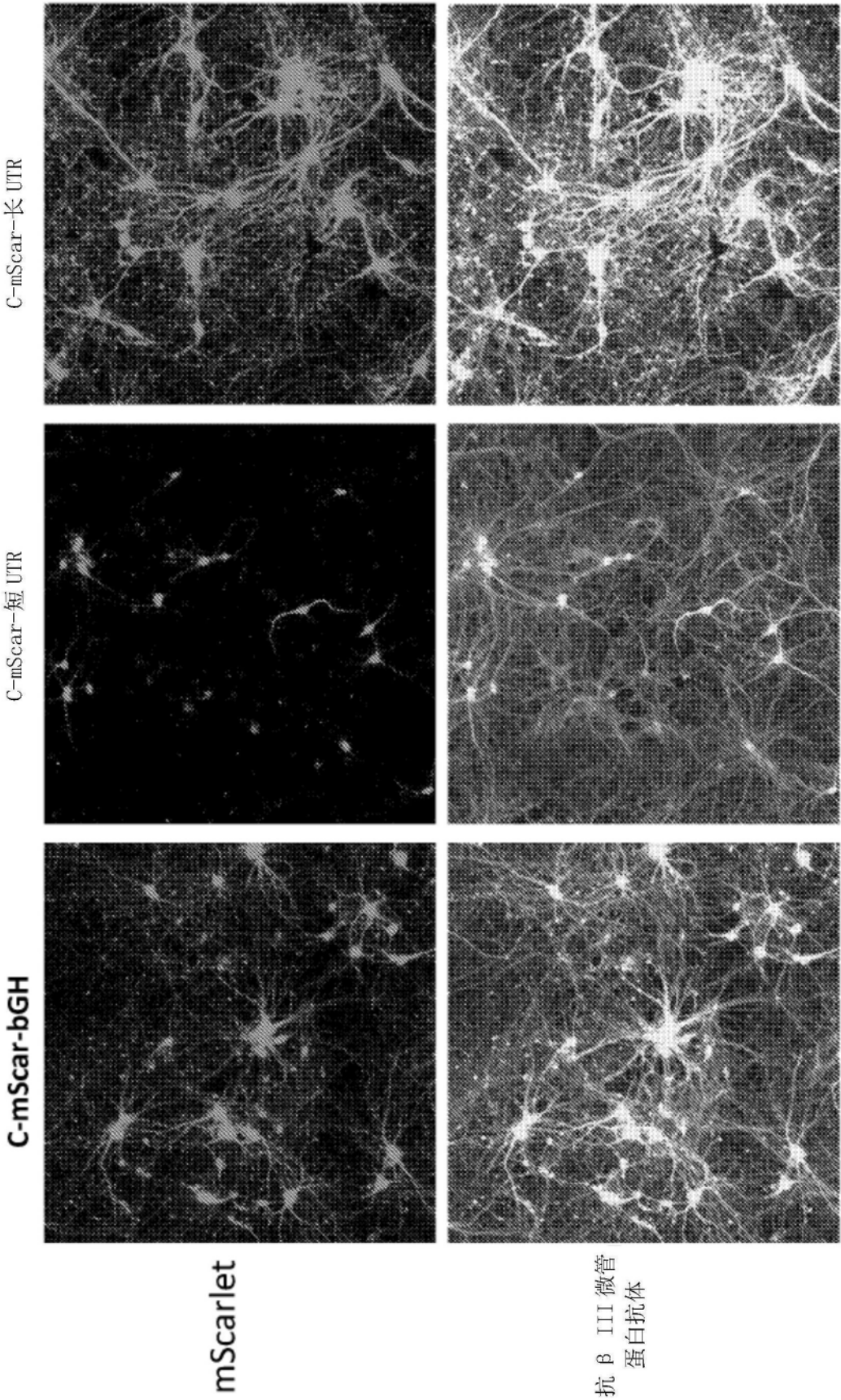


图7

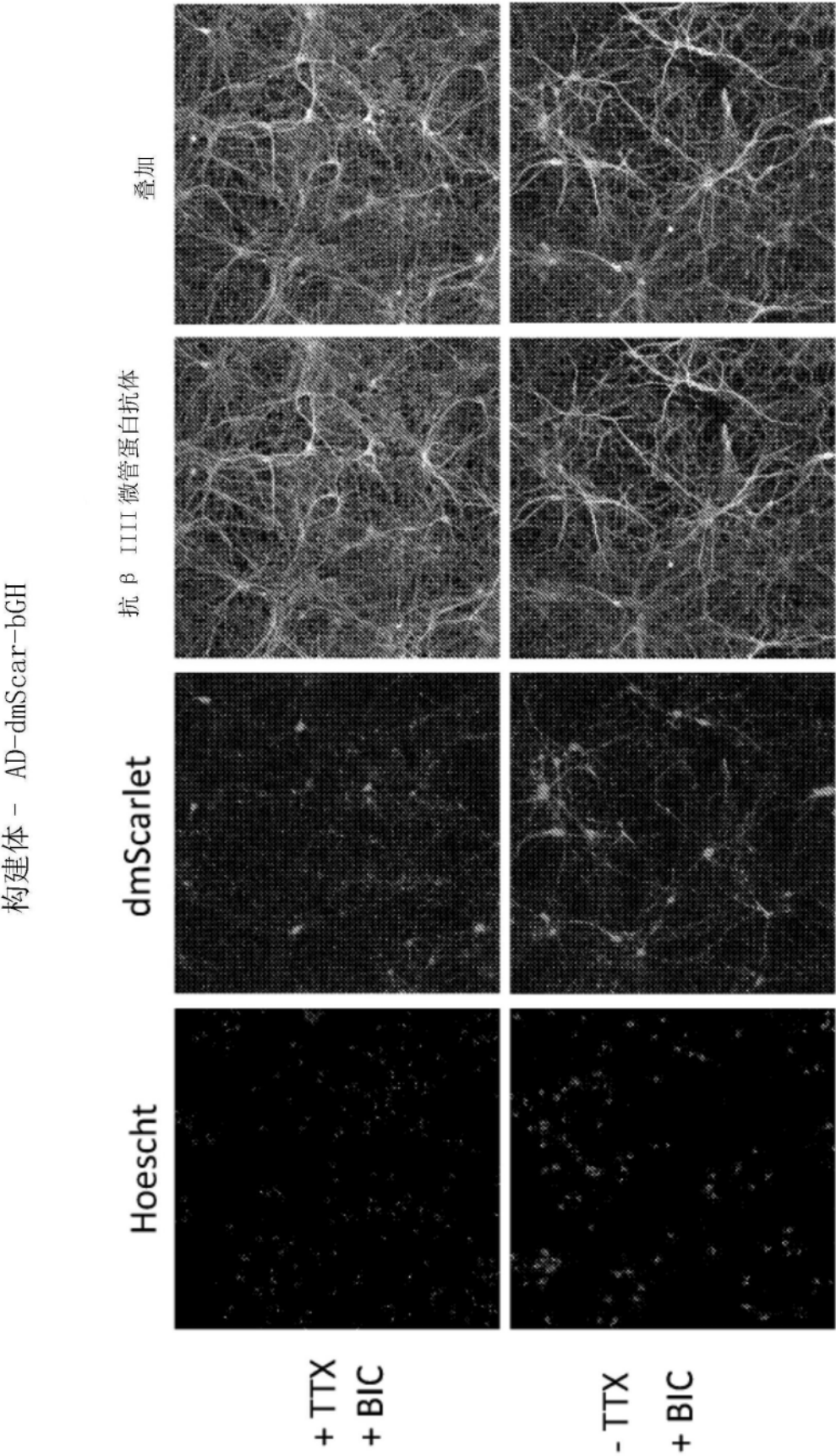


图8

C-mScarlet-短 UTR

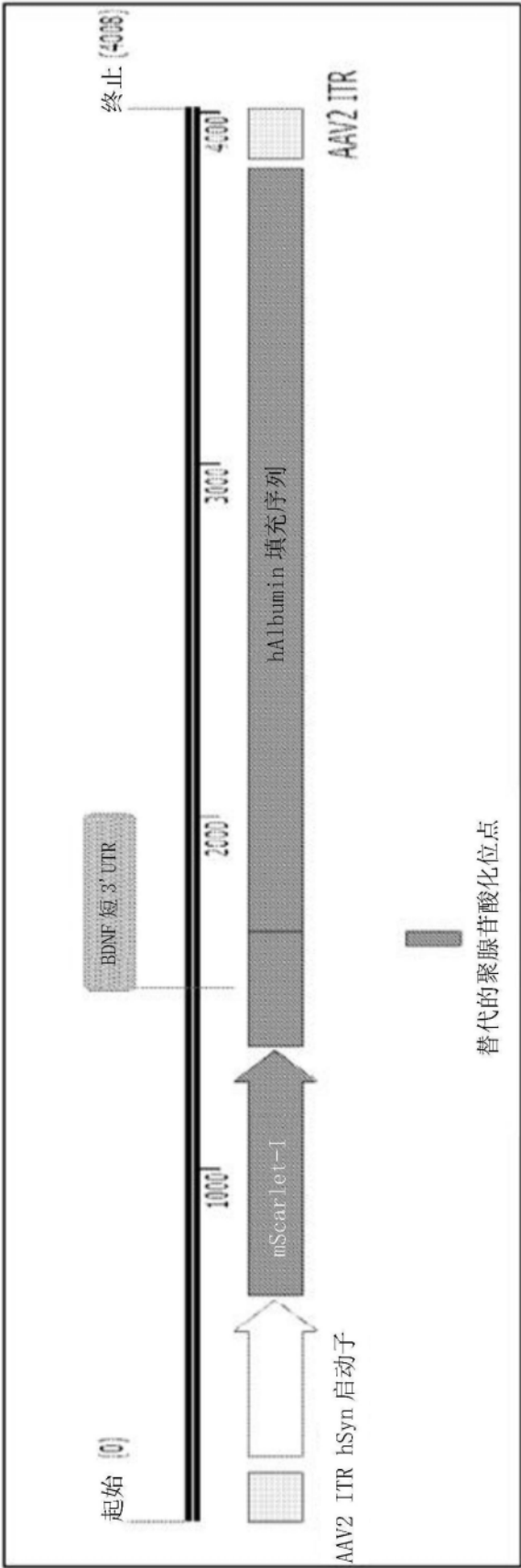


图9A

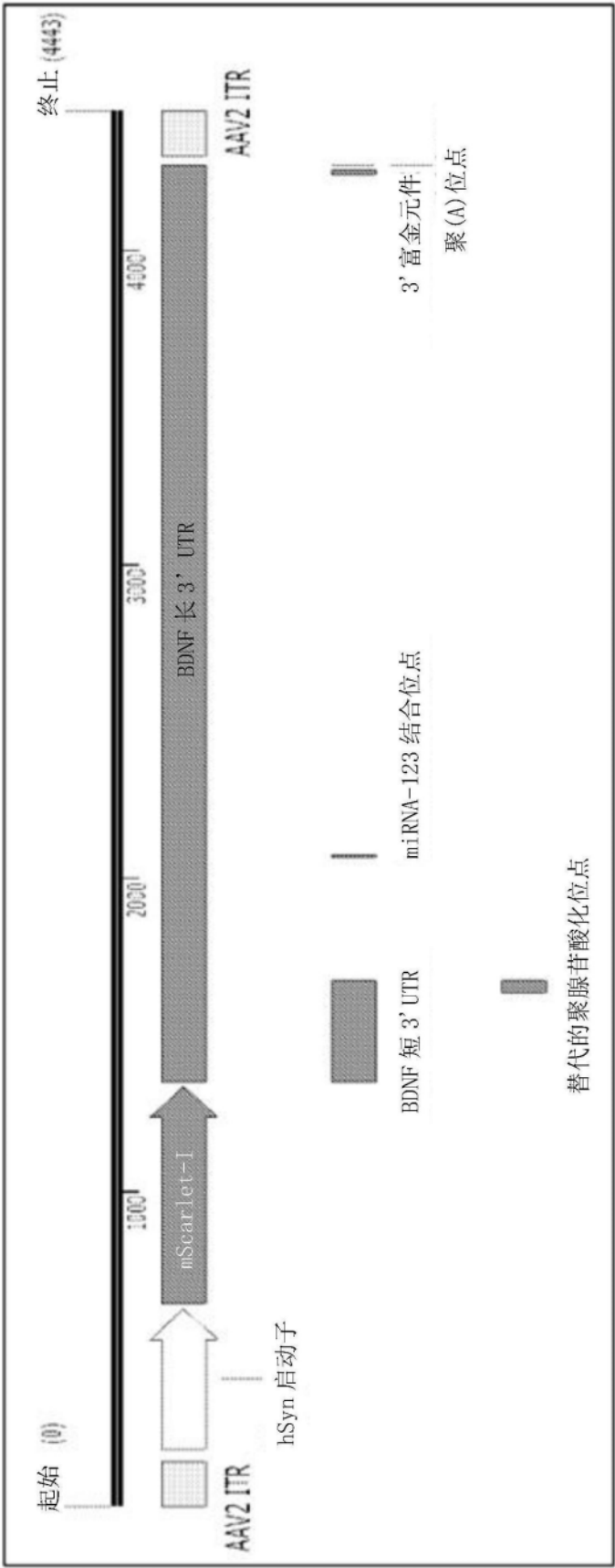


图9B

C-mScarlet-bGH

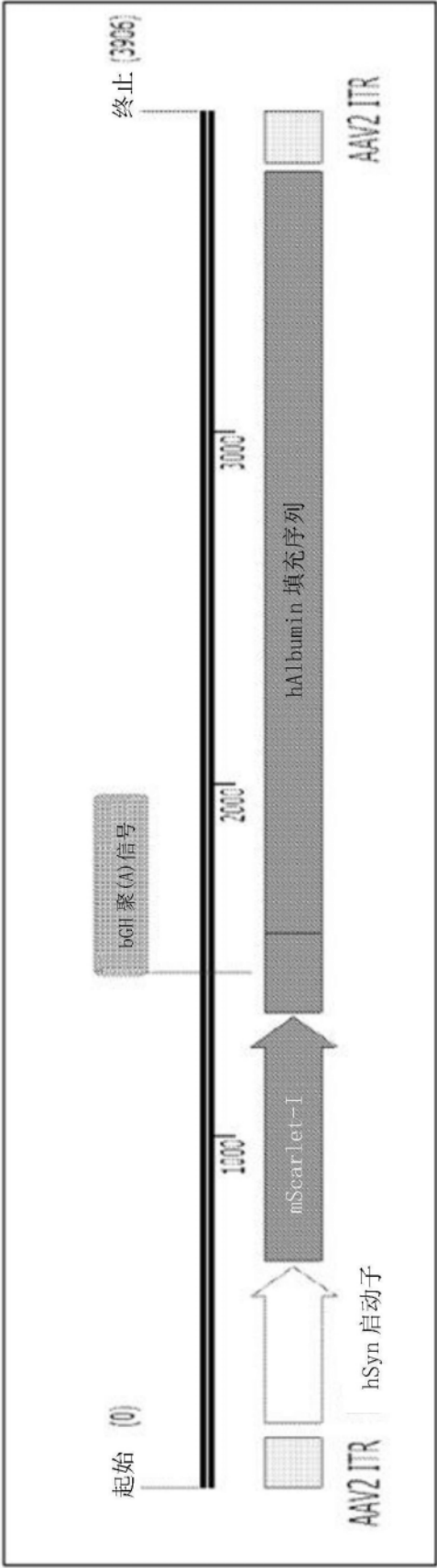


图9C

C-dmScarlet-短 UTR

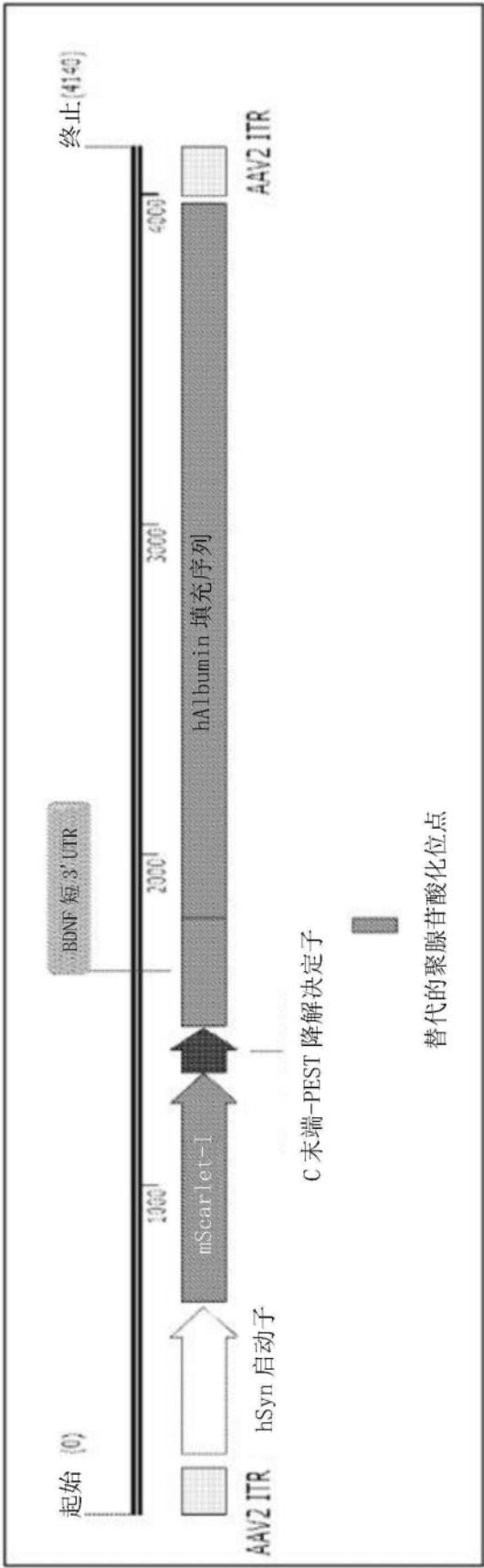


图10A

C-dmScarlet-长 UTR

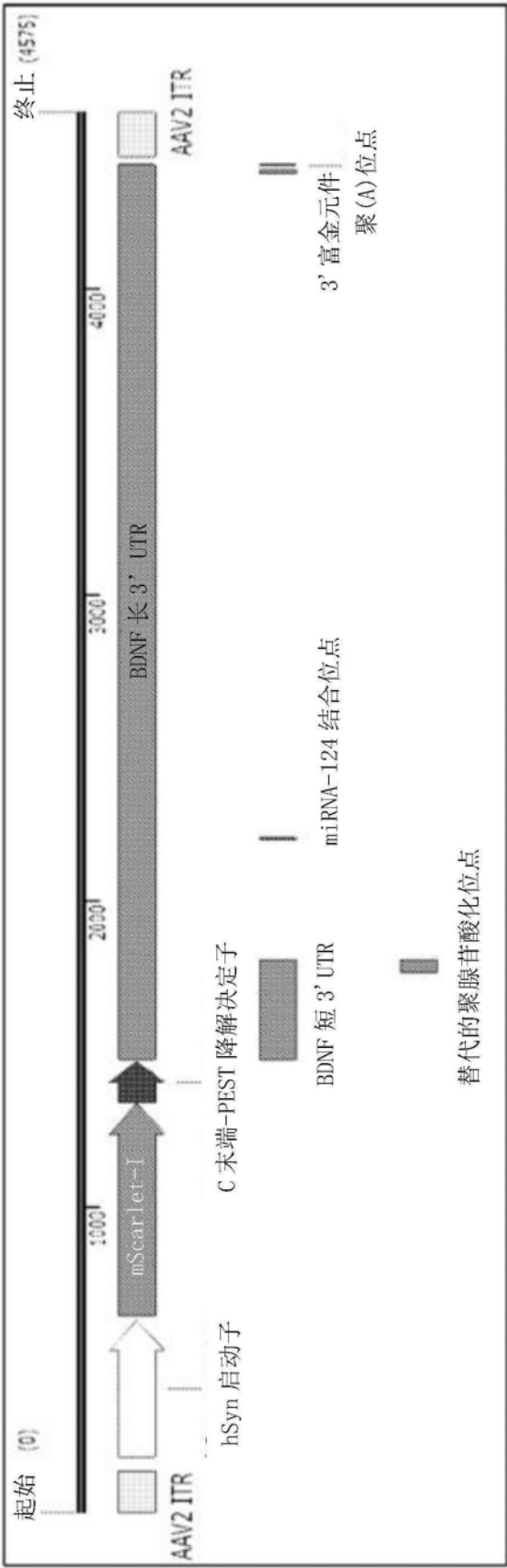


图10B

C-dmScar-bGH

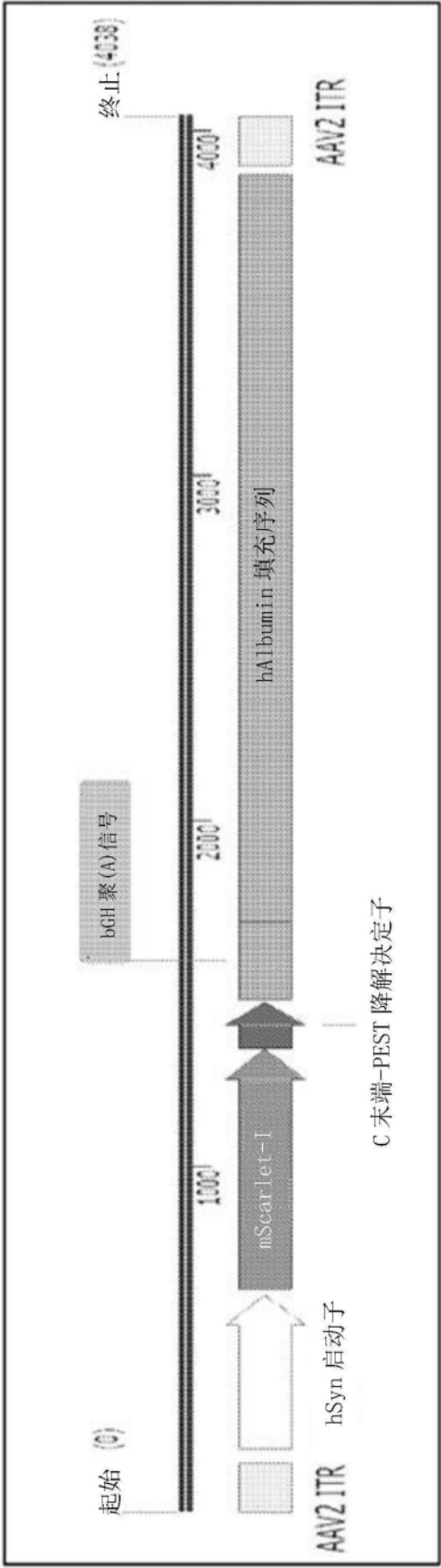


图10C

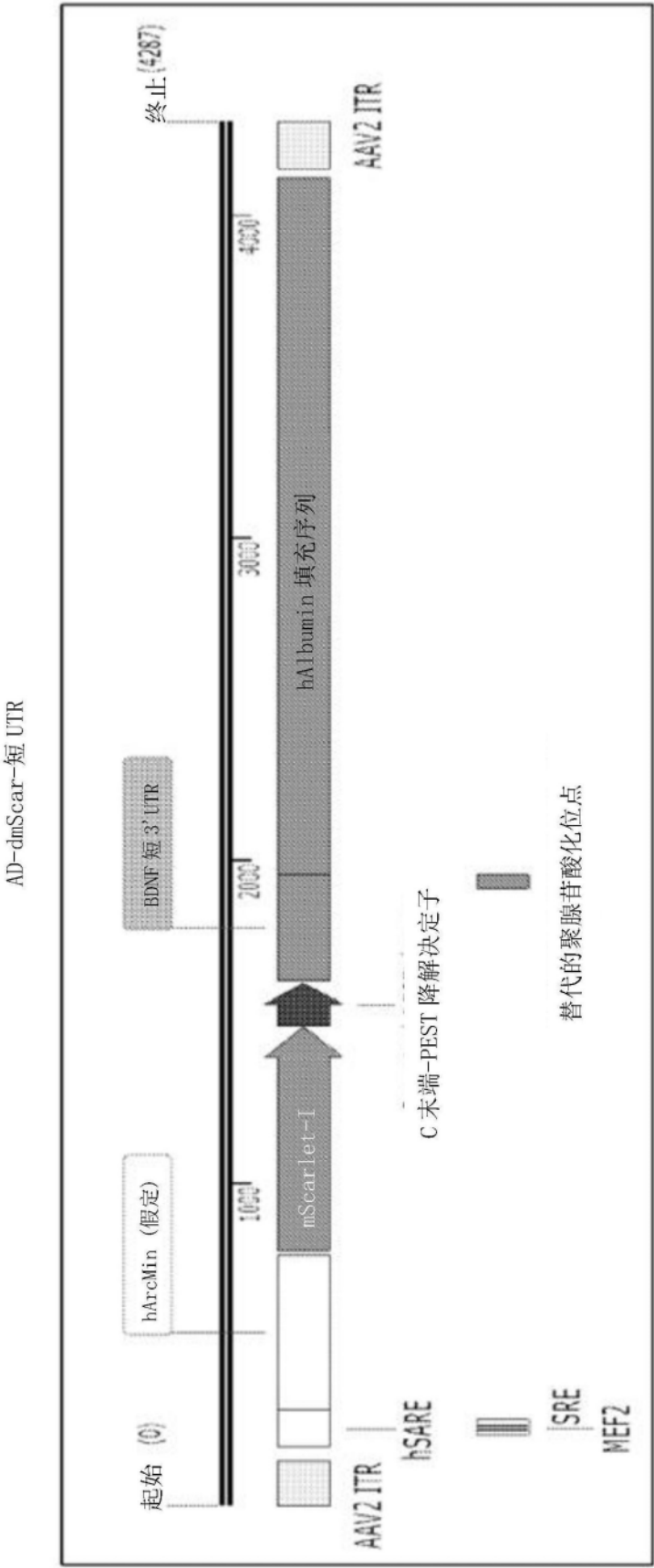


图11A

AD-dmScar--长 UTR

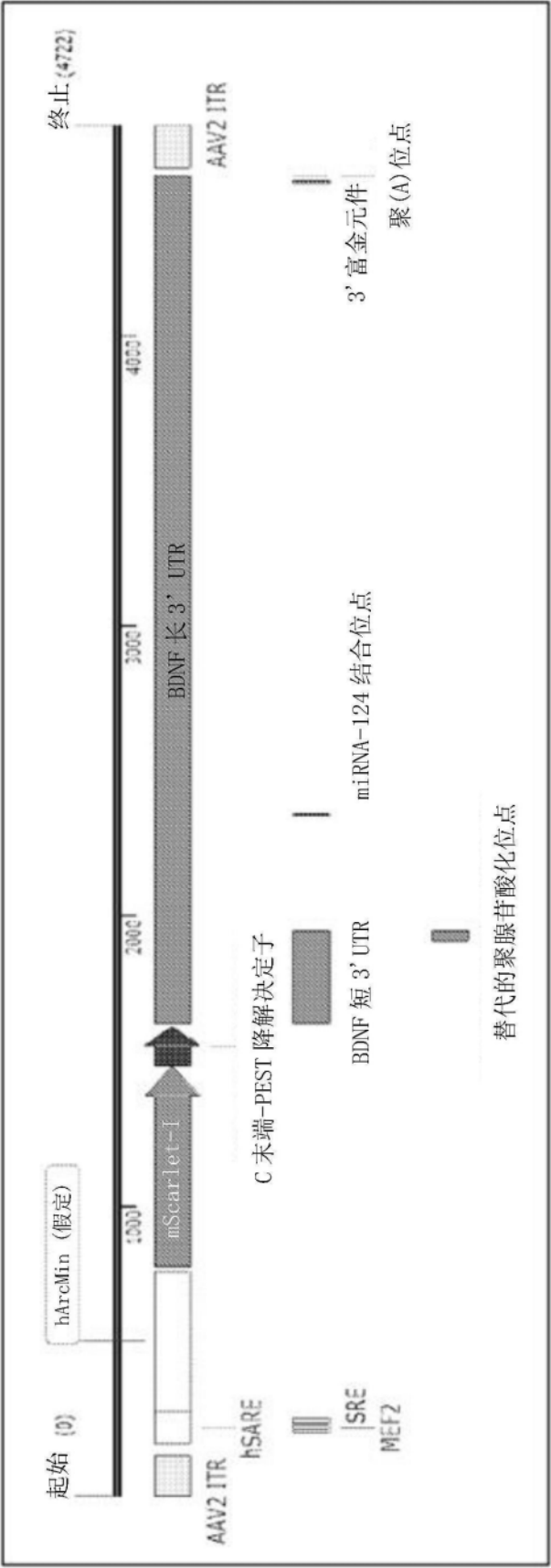


图11B

AD-dmScar-bGH

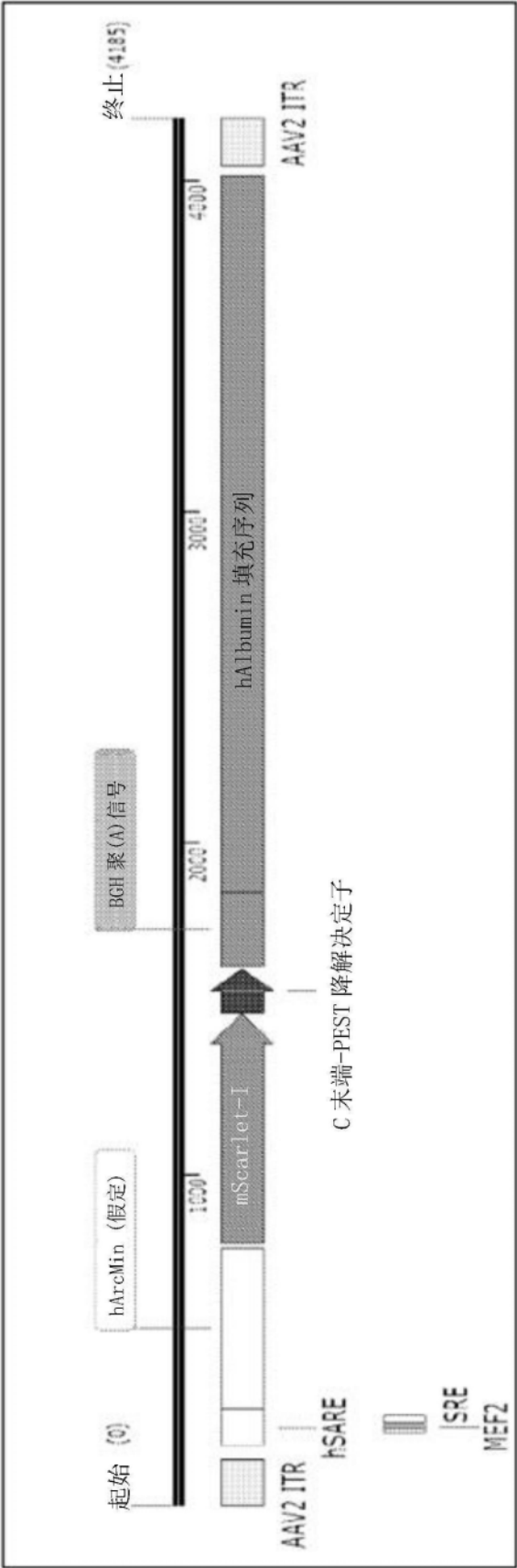


图11C

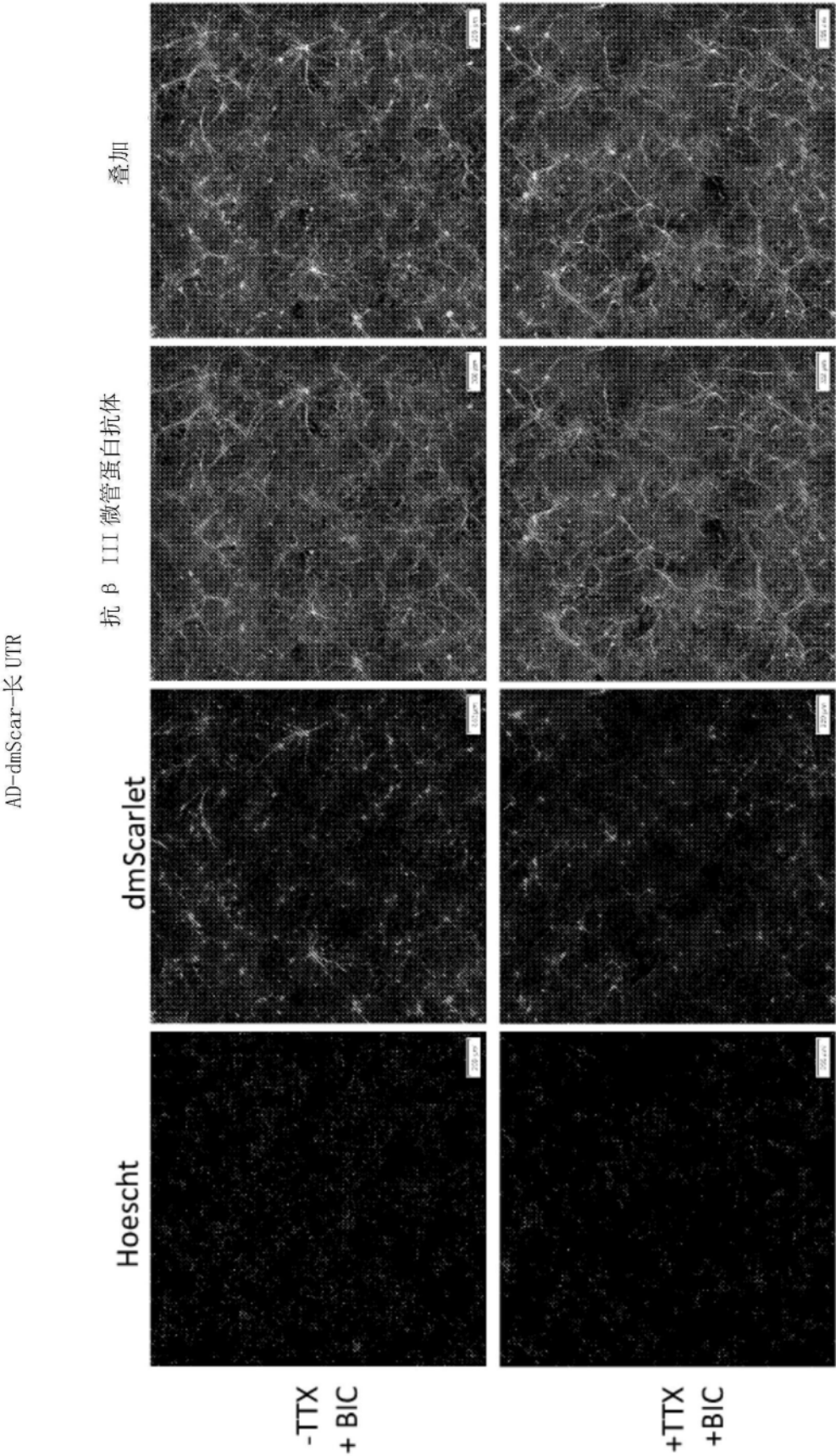


图12