



(12) 发明专利申请

(10) 申请公布号 CN 116457026 A

(43) 申请公布日 2023. 07. 18

(21) 申请号 202180072951.8

辛克莱·库珀

(22) 申请日 2021.09.03

(74) 专利代理机构 北京品源专利代理有限公司

(30) 优先权数据

11332

2013940.8 2020.09.04 GB

专利代理师 刘明海 胡彬

(85) PCT国际申请进入国家阶段日

(51) Int.Cl.

2023.04.25

A61K 48/00 (2006.01)

(86) PCT国际申请的申请数据

PCT/GB2021/052273 2021.09.03

(87) PCT国际申请的公布数据

W02022/049385 EN 2022.03.10

(71) 申请人 阿斯克肋匹奥生物制药公司

地址 美国北卡罗来纳州

(72) 发明人 安妮·布拉埃

胡安·曼纽尔·伊格莱西亚斯

迈克尔·L·罗伯茨

权利要求书3页 说明书96页

序列表18页 附图2页

(54) 发明名称

调节核酸序列

(57) 摘要

本发明涉及调节核酸序列,特别是NS特异性启动子、细胞特异性启动子、多细胞特异性启动子及其元件。本发明还涉及包含此类启动子的表达构建体、载体、病毒体、药物组合物和细胞,以及它们的使用方法。调节核酸序列对基因治疗应用特别有用。

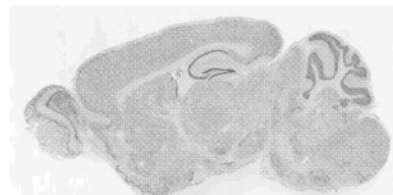


图 1A

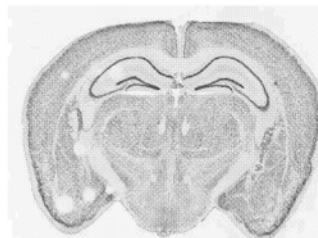


图 1B

1. 一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:

- CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- CRE0002\_S100B(SEQ ID NO:24)或其功能变体;
- CRE0005\_GFAP(SEQ ID NO:19)或其功能变体;
- CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- CRE0009\_S100B(SEQ ID NO:23)或其功能变体;
- CRE0012\_Arc(SEQ ID NO:21)或其功能变体;
- CRE0006\_GFAP(SEQ ID NO:27)或其功能变体;
- CRE0008\_GFAP(SEQ ID NO:28)或其功能变体;
- CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- CRE0008\_AQP4(SEQ ID NO:38)或其功能变体

可操作地连接至选自以下的启动子元件:

- CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- CRE0002\_CEND1\_mp(SEQ ID NO:30)或其功能变体;
- CRE0008\_SYN1\_mp(SEQ ID NO:31)或其功能变体;
- CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体。
- CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或
- CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。

2. 根据权利要求1所述的合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:

- CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- CRE0009\_S100B(SEQ ID NO:23)或其功能变体;
- CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- CRE0008\_AQP4(SEQ ID NO:38)或其功能变体

可操作地连接至选自以下的启动子元件:

- CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或
- CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。

3. 根据权利要求1或2所述的合成的NS特异性启动子,其包括以下CRE或其功能变体中的至少一个或由其组成:

- CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;

-CRE0009\_S100B(SEQ ID NO:23)或其功能变体;  
-CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和  
-CRE0008\_AQP4(SEQ ID NO:38)或其功能变体

可操作地连接至选自以下的启动子元件:

-CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;  
-CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;  
-CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;  
-CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或  
-CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。

4. 根据权利要求3所述的合成的NS特异性启动子,其中所述功能变体包括与SEQ ID NO:20、22-23、25-26、37-41中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

5. 根据权利要求1-3所述的合成的NS特异性启动子,其包括根据SEQ ID NO:1-14、33-36中的任何一个的序列或其功能变体或由其组成。

6. 根据权利要求5所述的合成的NS特异性启动子,其包括与SEQ ID NO:1-14、33-36中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

7. 一种NS特异性顺式调节元件(CRE),其包括根据SEQ ID NO:19-24、27、28、37、38中的任何一个的序列或其功能变体。

8. 根据权利要求7所述的NS特异性顺式调节元件(CRE),其包括与SEQ ID NO:19-24、27、28、37、38中的任何一个具有至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。

9. 一种合成的NS特异性顺式调节模块(CRM),其包括根据SEQ ID NO:15-18中的任何一个的序列,或包括根据权利要求7或8所述的CRE。

10. 一种合成的NS特异性启动子,其包括根据权利要求7或8所述的CRE或根据权利要求9所述的CRM。

11. 一种分离的启动子元件,其包括根据SEQ ID NO:25-26、29-32、39-41中的任何一个的序列,或其功能变体。

12. 根据权利要求11所述的分离的启动子元件,其包括与SEQ ID NO:25-26、29-32、39-41中的任何一个具有至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。

13. 一种合成的NS特异性启动子,其包括根据权利要求11或12所述的启动子元件。

14. 一种表达盒,其包括与编码表达产物的核酸序列可操作地连接的根据权利要求1-6、10或13中任一项所述的合成的NS特异性启动子。

15. 一种载体,其包括根据权利要求1-6、10或13中任一项所述的合成的NS特异性启动子或根据权利要求14所述的表达盒。

16. 根据权利要求15所述的载体,其中所述载体是病毒载体,例如AAV载体、腺病毒载体、逆转录病毒载体或慢病毒载体。

17. 根据权利要求16所述的载体,其中所述载体是慢病毒载体。

18. 根据权利要求16所述的载体,其中所述载体是AAV载体。

19. 一种病毒体,其包括根据权利要求15-18中任一项所述的载体。

20. 一种药物组合物,其包括根据权利要求1-6、10或13所述的合成的NS特异性启动子、根据权利要求14所述的表达盒、根据权利要求15-18中任一项所述的载体或根据权利要求19所述的病毒体。

21. 根据权利要求1-6、10或13所述的合成的NS特异性启动子、根据权利要求14所述的表达盒、根据权利要求15-18中任一项所述的载体、根据权利要求19所述的病毒体或根据权利要求20所述的药物组合物,其用作药物。

22. 一种细胞,其包括根据权利要求1-6、10或13所述的合成的NS特异性启动子、根据权利要求14所述的表达盒、根据权利要求15-18中任一项所述的载体或根据权利要求19所述的病毒体。

23. 根据权利要求1-6、10或13所述的合成的NS特异性启动子、根据权利要求14所述的表达盒、根据权利要求15-18中任一项所述的载体、根据权利要求19所述的病毒体或根据权利要求20所述的药物组合物,其用于制造用于治疗医学病况或疾病的药物组合物。

24. 一种生产表达产物的方法,所述方法包括在NS细胞中提供根据权利要求14所述的合成的NS特异性表达盒,以及表达所述合成的NS特异性表达盒中存在的表达产物。

25. 一种在NS细胞中表达治疗性转基因的方法,所述方法包括将根据权利要求14所述的合成的NS特异性表达盒、根据权利要求15-18中任一项所述的载体或根据权利要求19所述的病毒体引入NS细胞。

26. 根据权利要求25所述的在NS细胞中表达治疗性转基因的方法,其中所述表达盒、所述载体或所述病毒体是通过静脉内或注射引入的。

27. 根据权利要求26所述的方法,其中所述注射是在头静脉、正中静脉或贵要静脉中的一个中注射。

28. 一种治疗有需要的受试者,优选人类的方法,所述方法包括:

- 向所述受试者施用根据权利要求14所述的表达盒、根据权利要求15-18中任一项所述的载体、根据权利要求19所述的病毒体或根据权利要求20所述的药物组合物,其包括与根据权利要求1-6、10或13中任一项所述的启动子可操作地连接的编码治疗性产物的序列;和

- 在所述受试者的NS中表达治疗量的治疗性产物。

## 调节核酸序列

### 技术领域

[0001] 本发明涉及调节核酸序列,特别是神经系统特异性启动子、细胞特异性启动子、多细胞特异性启动子及其元件。本发明还涉及包含此类启动子的表达构建体、载体、病毒体、药物组合物和细胞,以及它们的使用方法。所述调节核酸序列对基因治疗应用特别有用。

### 背景技术

[0002] 提供以下讨论以帮助读者理解本公开,并不构成对现有技术的内容或相关性的任何承认。

[0003] 在对体内基因调节的内部机制进行广泛研究后,研究重点最近转向了通过将外源核酸序列引入细胞来调节基因表达。

[0004] 这在研究和生物工艺中是常规做法,其中将与启动子可操作地连接的所需表达产物的核酸序列引入生产细胞系中,通常是以载体的形式。

[0005] 在基因治疗领域,这对遗传性病症如单基因病症(或孟德尔病症)特别有意义,这些病症是由患者的细胞中存在缺陷基因引起的。将与启动子可操作地连接的缺陷基因的野生型等位基因的核酸序列引入患者的细胞是一种有利的治疗选择,因为它可以治愈病况,而传统药物只能解决症状。

[0006] 另外地或替代地,当缺陷基因具有对野生型等位基因产生拮抗作用的显性负突变时,将能够抑制缺陷基因的核酸引入患者的细胞可能是有益的。最后,将能够抵消或减轻缺陷基因的有害影响的核酸序列引入患者的细胞也是替代性或辅助性的治疗选择。

[0007] 在基因治疗中,控制已被引入细胞的外源核酸的表达对患者的健康和安全是最重要的。表达产物的水平不仅需要在治疗窗口内,而且表达需要在所需的组织内,甚至在所需组织的特定区域内,以使治疗有效。另外地,表达可能需要限制在特定的细胞类型或多种细胞类型,以避免副作用。在治疗窗口(即较低或较高)以外的表达或在治疗区域以外的表达,甚至在特定细胞或细胞组合以外的表达,可能没有治疗作用,甚至是有害的。

[0008] 亨廷顿氏病是由亨廷顿基因的两个遗传拷贝中的一个上的常染色体显性突变引起的。这种神经退行性疾病与逐渐恶化的运动功能障碍和由神经元死亡引起的认知能力损害有关。亨廷顿氏病没有治愈方法,治疗的目的是缓解症状。由于该病是由对体内存在的野生型亨廷顿等位基因有拮抗作用的常染色体显性突变引起的,因此通过引入野生型等位基因进行替代治疗不太可能特别有益。

[0009] 亨廷顿氏病的一个治疗选择是引入一种核酸,这种核酸可以抑制有缺陷的亨廷顿基因进入患者的细胞。在非疾病状态下,亨廷顿蛋白在全身表达,并且在CNS中的表达非常高,在CNS中它存在于所有的神经元中,也存在于胶质细胞中(Schulte和Littleton,2011)。由于有缺陷的亨廷顿基因的常染色体显性突变导致的功能丧失和/或有缺陷的亨廷顿基因的功能获得,不仅会导致神经元死亡,还会导致胶质细胞功能障碍。为了支持这一点,除了神经元死亡外,在亨廷顿患者的大脑中还观察到反应性小胶质细胞和激活的星形胶质细胞和少突胶质细胞(Schulte和Littleton,2011)。因此,希望表达不仅能在所有神经元中,而

且能在一些或所有胶质细胞中抑制有缺陷的亨廷顿基因的核酸。

[0010] 因此,需要驱动在神经系统(NS),特别是CNS中,特别是在特定的细胞类型,如神经元以及神经元支持细胞,包括一些或所有胶质细胞、少突胶质细胞、中间神经元等中的表达的启动子。在神经系统(NS)中的表达包括在中枢神经系统(CNS)和/或周围神经系统(PNS)的细胞中的表达。在CNS中的表达包括在大脑和脊髓的细胞中的表达,并且包括神经元细胞以及非神经元细胞、中间神经元或神经元支持细胞。在PNS中的表达包括在周围神经系统的细胞中的表达,并且包括神经元细胞以及非神经元细胞、中间神经元或神经元支持细胞。

[0011] 另一个治疗选择是将能够抵消或减轻有缺陷的亨廷顿基因的有害影响的核酸序列引入患者的细胞中。亨廷顿氏病与胆固醇代谢异常有关,尽管合成减少,但胆固醇在大脑中持续积累(Boussicault等人,2016)。由于血脑屏障对胆固醇不具渗透性,CNS的胆固醇稳态与身体的其他部分是分开的,因此可以单独靶向。在CNS中,星形胶质细胞合成胆固醇,神经元通过神经元特异性酶CYP46A1代谢胆固醇(Boussicault等人,2016)。另外地,胆固醇在少突胶质细胞中代谢,产生髓鞘碱性蛋白(MBP)。在亨廷顿氏病中,CYP46A1表达水平降低,CYP46A1酶水平显然不足以克服大脑中的胆固醇积累(Boussicault等人,2016)。为了支持这一点,发现在小鼠亨廷顿氏病模型中,通过泛启动子CMV或CAG表达CYP46A1可以调节大脑中的胆固醇代谢(Kacher等人,2019)。因此,在神经细胞中表达额外的CYP46A1和/或将CYP46A1(以及因此代谢胆固醇的能力)引入胶质细胞,能够调节大脑中的胆固醇代谢。

[0012] 虽然CYP46A1在CNS中的表达似乎是亨廷顿氏病的一个有吸引力的治疗选择,但在人类的基因治疗环境中,优选通过微创的全身施用,如静脉内或动脉内(如颈动脉内)施用,将编码CYP46A1的核酸引入NS中。然而,由于CYP46A1在非NS细胞中的表达,与泛启动子可操作地连接的编码CYP46A1酶的核酸的全身施用可能对NS以外的胆固醇代谢具有不想要的副作用。

[0013] 因此,需要NS特异性启动子驱动在神经元和一些或所有胶质细胞中的表达。

[0014] 其他NS的疾病也是合适的基因治疗靶标。在一些这样的疾病中,可能需要在特定的NS细胞中定向表达治疗基因,而在其他疾病中,可能适合在NS,优选在CNS中进行更普遍的、非特异性的表达。

[0015] 本发明的一个或多个方面旨在解决上述问题中的一个或多个。

## 发明内容

[0016] 在本发明的第一方面,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:

[0017] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;

[0018] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体;

[0019] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体;

[0020] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;

[0021] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体;

[0022] -CRE0012\_Arc(SEQ ID NO:21)或其功能变体;

[0023] -CRE0006\_GFAP(SEQ ID NO:27)或其功能变体;

[0024] -CRE0008\_GFAP(SEQ ID NO:28)或其功能变体;

- [0025] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- [0026] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0027] 可操作地连接至启动子元件,任选地,其中所述启动子元件选自:
- [0028] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- [0029] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- [0030] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0031] -CRE0002\_CEND1\_mp(SEQ ID NO:30)或其功能变体;
- [0032] -CRE0008\_SYN1\_mp(SEQ ID NO:31)或其功能变体;
- [0033] -CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0034] -CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- [0035] -CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或
- [0036] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。
- [0037] 在一些实施方案中,神经系统(NS)特异性启动子在中枢神经系统(CNS)和周围神经系统(PNS)中都具有活性。在一些实施方案中,NS特异性启动子在CNS中具有活性,即CNS特异性启动子。在一些实施方案中,NS特异性启动子在CNS中具有活性,但在PNS中没有活性。在一些实施方案中,NS特异性启动子在PNS中具有活性,即PNS特异性启动子。在一些实施方案中,NS特异性启动子在PNS中具有活性,但在CNS中没有活性。
- [0038] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0039] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0040] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体;
- [0041] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体;
- [0042] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- [0043] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体;
- [0044] -CRE0012\_Arc(SEQ ID NO:21)或其功能变体;
- [0045] -CRE0006\_GFAP(SEQ ID NO:27)或其功能变体;
- [0046] -CRE0008\_GFAP(SEQ ID NO:28)或其功能变体;
- [0047] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- [0048] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0049] 可操作地连接至选自以下的启动子元件:
- [0050] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- [0051] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- [0052] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0053] -CRE0002\_CEND1\_mp(SEQ ID NO:30)或其功能变体;
- [0054] -CRE0008\_SYN1\_mp(SEQ ID NO:31)或其功能变体;
- [0055] -CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0056] -CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- [0057] -CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或
- [0058] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。

[0059] 功能变体的定义见下文。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 19-32、37-41中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地,合成的NS特异性启动子可以包括与SEQ ID NO:19、20、21、22、23、24、27、28、37、38中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列,所述序列可操作地连接至与SEQ ID NO:25、26、29-32、39-41中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0060] 在一个实施方案中,提供了一种NS特异性启动子,其包括以下顺式调节元件(CRE)中的至少一个或由其组成:

[0061] -CRE0001\_S100B(SEQ ID NO:22)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0062] -CRE0007\_GFAP(SEQ ID NO:20)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0063] -CRE0009\_S100B(SEQ ID NO:23)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0064] -CRE0002\_S100B(SEQ ID NO:24)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0065] -CRE0005\_GFAP(SEQ ID NO:19)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0066] -CRE0012\_Arc(SEQ ID NO:21)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0067] -CRE0006\_GFAP(SEQ ID NO:27)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0068] -CRE0008\_GFAP(SEQ ID NO:28)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0069] -CRE0006\_AQP4(SEQ ID NO:37)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;和

[0070] -CRE0008\_AQP4(SEQ ID NO:38)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列

[0071] 与选自以下的启动子元件可操作地连接:

[0072] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0073] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0074] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0075] -CRE0002\_CEND1\_mp(SEQ ID NO:30)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0076] -CRE0008\_SYN1\_mp(SEQ ID NO:31)或与其具有至少60%、65%、70%、75%、80%、

90%、95%、96%、97%、98%或99%同一性的序列；

[0077] -CRE0013\_SYN1\_mp (SEQ ID NO:32) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；

[0078] -CRE0015\_GAP43\_mp (SEQ ID NO:39) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；

[0079] -CRE0016\_Eno2\_mp (SEQ ID NO:40) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；和/或

[0080] -CRE0004\_GFAP\_mp (SEQ ID NO:41) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地，合成的NS特异性启动子包括以下或由其组成：

[0081] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体；

[0082] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体；

[0083] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体；

[0084] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体；

[0085] -CRE0001\_S100B (SEQ ID NO:22) ，或其功能变体，可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30) 或其功能变体；

[0086] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0008\_SYN1\_mp (SEQ ID NO:31) 或其功能变体；

[0087] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体；

[0088] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体；

[0089] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体；

[0090] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体；

[0091] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体；

[0092] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体；

[0093] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体；

[0094] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30) 或其功能变体；

[0095] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0008\_SYN1\_mp

(SEQ ID NO:31)或其功能变体;

[0096] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,可操作地连接至CRE0015\_GAP43\_  
mp (SEQ ID NO:39)或其功能变体;

[0097] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,可操作地连接至CRE0016\_Eno2\_  
mp (SEQ ID NO:40)或其功能变体;

[0098] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,可操作地连接至CRE0004\_GFAP\_  
mp (SEQ ID NO:41)或其功能变体;

[0099] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0001\_CEND1\_  
mp (SEQ ID NO:29)或其功能变体;

[0100] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0013\_SYN1\_  
mp (SEQ ID NO:32)或其功能变体;

[0101] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0003\_CEND1\_  
mp (SEQ ID NO:25)或其功能变体;

[0102] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0009\_SYN1\_  
mp (SEQ ID NO:26)或其功能变体;

[0103] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0002\_CEND1\_  
mp (SEQ ID NO:30)或其功能变体;

[0104] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0008\_SYN1\_  
mp (SEQ ID NO:31)或其功能变体;

[0105] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0015\_GAP43\_  
mp (SEQ ID NO:39)或其功能变体;

[0106] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0016\_Eno2\_  
mp (SEQ ID NO:40)或其功能变体;

[0107] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,可操作地连接至CRE0004\_GFAP\_  
mp (SEQ ID NO:41)或其功能变体;

[0108] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0001\_CEND1\_  
mp (SEQ ID NO:29)或其功能变体;

[0109] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0013\_SYN1\_  
mp (SEQ ID NO:32)或其功能变体;

[0110] -CRE0009\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0003\_CEND1\_  
mp (SEQ ID NO:25)或其功能变体;

[0111] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0009\_SYN1\_  
mp (SEQ ID NO:26)或其功能变体;

[0112] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0002\_CEND1\_  
mp (SEQ ID NO:30)或其功能变体;

[0113] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0008\_SYN1\_  
mp (SEQ ID NO:31)或其功能变体;

[0114] -CRE0002\_S100B (SEQ ID NO:24)或其功能变体,可操作地连接至CRE0015\_GAP43\_  
mp (SEQ ID NO:39)或其功能变体;

- [0115] -CRE0002\_S100B (SEQ ID NO:24) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体;
- [0116] -CRE0002\_S100B (SEQ ID NO:24) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;
- [0117] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;
- [0118] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;
- [0119] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;
- [0120] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体;
- [0121] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30) 或其功能变体;
- [0122] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0008\_SYN1\_mp (SEQ ID NO:31) 或其功能变体;
- [0123] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体;
- [0124] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体;
- [0125] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;
- [0126] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;
- [0127] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;
- [0128] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;
- [0129] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体;
- [0130] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30) 或其功能变体;
- [0131] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0008\_SYN1\_mp (SEQ ID NO:31) 或其功能变体;
- [0132] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体;
- [0133] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体;
- [0134] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp

- (SEQ ID NO:41)或其功能变体;
- [0135] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29)或其功能变体;
- [0136] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32)或其功能变体;
- [0137] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25)或其功能变体;
- [0138] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26)或其功能变体;
- [0139] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30)或其功能变体;
- [0140] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0008\_SYN1\_mp (SEQ ID NO:31)或其功能变体;
- [0141] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39)或其功能变体;
- [0142] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0143] -CRE0006\_GFAP (SEQ ID NO:27)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0144] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29)或其功能变体;
- [0145] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32)或其功能变体;
- [0146] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25)或其功能变体;
- [0147] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26)或其功能变体;
- [0148] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0002\_CEND1\_mp (SEQ ID NO:30)或其功能变体;
- [0149] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0008\_SYN1\_mp (SEQ ID NO:31)或其功能变体;
- [0150] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39)或其功能变体;
- [0151] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0152] -CRE0008\_GFAP (SEQ ID NO:28)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0153] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29)或其功能变体;

- [0154] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0155] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- [0156] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- [0157] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0002\_CEND1\_mp(SEQ ID NO:30)或其功能变体;
- [0158] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0008\_SYN1\_mp(SEQ ID NO:31)或其功能变体;
- [0159] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- [0160] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;
- [0161] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;
- [0162] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0163] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0164] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- [0165] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- [0166] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0002\_CEND1\_mp(SEQ ID NO:30)或其功能变体;
- [0167] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0008\_SYN1\_mp(SEQ ID NO:31)或其功能变体;
- [0168] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- [0169] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;或
- [0170] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。
- [0171] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0172] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0173] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- [0174] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体;

- [0175] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- [0176] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0177] 可操作地连接至启动子元件,任选地,其中所述启动子元件选自:
- [0178] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0179] -CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0180] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;
- [0181] -CRE0015\_GAP43\_mp(SEQ ID NO:40)或其功能变体;和/或
- [0182] -CRE0016\_Eno2\_mp(SEQ ID NO:39)或其功能变体。
- [0183] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0184] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0185] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- [0186] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体;
- [0187] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- [0188] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0189] 可操作地连接至选自以下的启动子元件:
- [0190] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0191] -CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0192] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;
- [0193] -CRE0015\_GAP43\_mp(SEQ ID NO:40)或其功能变体;和/或
- [0194] -CRE0016\_Eno2\_mp(SEQ ID NO:39)或其功能变体。
- [0195] 功能变体的定义见下文。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22、23、37、38、29、32、41、40、39中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22、37、38或23中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列,所述序列可操作地连接至与SEQ ID NO: 29、41、40、39或32中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。
- [0196] 在一个实施方案中,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)中的至少一个或由其组成:
- [0197] -CRE0001\_S100B(SEQ ID NO:22)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0198] -CRE0007\_GFAP(SEQ ID NO:20)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0199] -CRE0009\_S100B(SEQ ID NO:23)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0200] -CRE0006\_AQP4(SEQ ID NO:37)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0201] -CRE0008\_AQP4(SEQ ID NO:38)或与其具有至少60%、65%、70%、75%、80%、

90%、95%、96%、97%、98%或99%同一性的序列

[0202] 与选自以下的启动子元件可操作地连接:

[0203] -CRE0001\_CEND1\_mp (SEQ ID NO:29) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0204] -CRE0013\_SYN1\_mp (SEQ ID NO:32) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0205] -CRE0004\_GFAP\_mp (SEQ ID NO:41) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0206] -CRE0015\_GAP43\_mp (SEQ ID NO:40) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;和

[0207] -CRE0016\_Eno2\_mp (SEQ ID NO:39) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0208] 适当地,合成的NS特异性启动子包括以下或由其组成:

[0209] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;

[0210] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;

[0211] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;

[0212] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:40) 或其功能变体;

[0213] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:39) 或其功能变体;

[0214] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;

[0215] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;

[0216] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;

[0217] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:40) 或其功能变体;

[0218] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:39) 或其功能变体;

[0219] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;

[0220] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;

[0221] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;

- [0222] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:40)或其功能变体;
- [0223] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:39)或其功能变体;
- [0224] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0225] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0226] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;
- [0227] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:40)或其功能变体;
- [0228] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:39)或其功能变体;
- [0229] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;
- [0230] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;
- [0231] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;
- [0232] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:40)或其功能变体;或
- [0233] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:39)或其功能变体。
- [0234] 适当地,提供了一种合成的CNS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0235] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0236] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;和
- [0237] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体
- [0238] 可操作地连接至启动子元件,任选地,其中所述启动子元件选自:
- [0239] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;和/或
- [0240] -CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体。
- [0241] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0242] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0243] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;和
- [0244] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体
- [0245] 可操作地连接至选自以下的启动子元件:
- [0246] -CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;和/或

[0247] -CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体。

[0248] 功能变体的定义见下文。适当地,合成的NS特异性启动子可以包括与SEQ ID NO:20、22、23、29、32中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地,合成的NS特异性启动子可以包括与SEQ ID NO:20、22或23中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列,所述序列可操作地连接至与SEQ ID NO:29或32中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0249] 在一个实施方案中,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)中的至少一个或由其组成:

[0250] -CRE0001\_S100B (SEQ ID NO:22) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0251] -CRE0007\_GFAP (SEQ ID NO:20) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0252] -CRE0009\_S100B (SEQ ID NO:23) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0253] 可操作地连接至选自以下的启动子元件:

[0254] -CRE0001\_CEND1\_mp (SEQ ID NO:29) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0255] -CRE0013\_SYN1\_mp (SEQ ID NO:32) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0256] 适当地,合成的NS特异性启动子包括或由以下部分组成:

[0257] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;

[0258] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;

[0259] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;

[0260] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体;

[0261] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0001\_CEND1\_mp (SEQ ID NO:29) 或其功能变体;或

[0262] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0013\_SYN1\_mp (SEQ ID NO:32) 或其功能变体。

[0263] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:

[0264] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体;

[0265] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体;

[0266] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体;

[0267] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体;和

- [0268] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0269] 可操作地连接至启动子元件,任选地,其中所述启动子元件选自:
- [0270] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能元件;
- [0271] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能元件;
- [0272] -CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能元件;
- [0273] -CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能元件;和/或
- [0274] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能元件。
- [0275] 适当地,提供了一种合成的NS特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0276] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体;
- [0277] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体;
- [0278] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体;
- [0279] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体;和
- [0280] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体
- [0281] 可操作地连接至选自以下的启动子元件:
- [0282] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;
- [0283] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;
- [0284] -CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;
- [0285] -CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;和/或
- [0286] -CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。
- [0287] 功能变体的定义见下文。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22、23、37、38、25、26、39、40、41中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22、37、38或23中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列,所述序列可操作地连接至与SEQ ID NO: 25、39、40、41或26中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。
- [0288] 在一个实施方案中,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)中的至少一个或由其组成:
- [0289] -CRE0001\_S100B(SEQ ID NO:22)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0290] -CRE0007\_GFAP(SEQ ID NO:20)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0291] -CRE0009\_S100B(SEQ ID NO:23)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;
- [0292] -CRE0006\_AQP4(SEQ ID NO:37)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;和
- [0293] -CRE0008\_AQP4(SEQ ID NO:38)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

- [0294] 可操作地连接至选自以下的启动子元件：
- [0295] -CRE0003\_CEND1\_mp (SEQ ID NO:25) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；
- [0296] -CRE0009\_SYN1\_mp (SEQ ID NO:26) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；
- [0297] -CRE0015\_GAP43\_mp (SEQ ID NO:39) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；
- [0298] -CRE0016\_Eno2\_mp (SEQ ID NO:40) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列；和/或
- [0299] -CRE0004\_GFAP\_mp (SEQ ID NO:41) 或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。
- [0300] 适当地，合成的NS特异性启动子包括以下或由其组成：
- [0301] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体；
- [0302] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体；
- [0303] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体；
- [0304] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体；
- [0305] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体，可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体；
- [0306] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体；
- [0307] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体；
- [0308] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体；
- [0309] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体；
- [0310] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体，可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体；
- [0311] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体，可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体；
- [0312] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体，可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体；
- [0313] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体，可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体；
- [0314] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体，可操作地连接至CRE0016\_Eno2\_

- mp (SEQ ID NO:40) 或其功能变体;
- [0315] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;
- [0316] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;
- [0317] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体;
- [0318] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体;
- [0319] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体;
- [0320] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体;
- [0321] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;
- [0322] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体;
- [0323] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQ ID NO:39) 或其功能变体;
- [0324] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40) 或其功能变体;
- [0325] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41) 或其功能变体。
- [0326] 适当地,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0327] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体;
- [0328] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体;和
- [0329] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体
- [0330] 可操作地连接至启动子元件,任选地,其中所述启动子元件选自:
- [0331] -CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;和/或
- [0332] -CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体。
- [0333] 适当地,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)或其功能变体中的至少一个或由其组成:
- [0334] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体;
- [0335] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体;和
- [0336] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体
- [0337] 可操作地连接至选自以下的启动子元件:
- [0338] -CRE0003\_CEND1\_mp (SEQ ID NO:25) 或其功能变体;和/或
- [0339] -CRE0009\_SYN1\_mp (SEQ ID NO:26) 或其功能变体。

[0340] 功能变体的定义见下文。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22、23、25-26中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。适当地,合成的NS特异性启动子可以包括与SEQ ID NO: 20、22或23中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列,所述序列可操作地连接至与SEQ ID NO:25或26中的任何一个具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0341] 在一个实施方案中,提供了一种合成的神经系统(NS)特异性启动子,其包括以下顺式调节元件(CRE)中的至少一个或由其组成:

[0342] -CRE0001\_S100B(SEQ ID NO:22)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0343] -CRE0007\_GFAP(SEQ ID NO:20)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0344] -CRE0009\_S100B(SEQ ID NO:23)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;

[0345] 可操作地连接至选自以下的启动子元件:

[0346] -CRE0003\_CEND1\_mp(SEQ ID NO:25)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列;和/或

[0347] -CRE0009\_SYN1\_mp(SEQ ID NO:26)或与其具有至少60%、65%、70%、75%、80%、90%、95%、96%、97%、98%或99%同一性的序列。

[0348] 适当地,合成的NS特异性启动子包括以下或由其组成:

[0349] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0350] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0351] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0352] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0353] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;或

[0354] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体。

[0355] NS特异性启动子可以在其他细胞中表达。然而,它在NS细胞如大脑、脊髓和周围神经系统中的神经元细胞,以及位于大脑、脊髓和周围神经系统中的非神经元细胞、中间神经元或神经元支持细胞中的表达程度更高。例如,NS特异性启动子在位于NS中的细胞,包括位于大脑、脊髓和周围神经系统中的神经元和非神经元细胞以及中间神经元中表达的基因比位于CNS和/或PNS以外的细胞高至少25%,或至少35%,或至少45%,或至少55%,或至少65%,或至少75%,或至少80%,或至少85%,或至少90%,或至少95%,或25%-95%之间的任何整数。

[0356] 适当地,合成的NS特异性启动子可以包括一个或多个额外的CRE或其功能变体。这些CRE可以是根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)、其功能变体或其他CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自本领域已知的任何CRE。例如,在一个实施方案中,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8,来自WO 2020/076614A1的SEQ ID NO:1-7,和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。在一些优选实施方案中,合成的NS特异性启动子可以另外包括CRE0012\_Arc(SEQ ID NO:21)或其功能变体。

[0357] 适当地,合成的NS特异性启动子包括以下或由其组成:

[0358] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;

[0359] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;

[0360] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;

[0361] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体;

[0362] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0001\_CEND1\_mp(SEQ ID NO:29)或其功能变体;或

[0363] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0013\_SYN1\_mp(SEQ ID NO:32)或其功能变体。

[0364] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0365] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0366] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0367] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0368] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;或

[0369] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体。

[0370] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQ ID NO:39)或其功能变体;

[0371] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;

[0372] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;

- [0373] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0374] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0375] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0376] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0377] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0378] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0379] -CRE0001\_S100B (SEQ ID NO:22)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0380] -CRE0001\_S100B (SEQ ID NO:22)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0381] -CRE0001\_S100B (SEQ ID NO:22)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0382] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0383] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0384] -CRE0007\_GFAP (SEQ ID NO:20)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0385] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0386] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0387] -CRE0009\_S100B (SEQ ID NO:23)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp (SEQ ID NO:41)或其功能变体;
- [0388] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp (SEQ ID NO:25)或其功能变体;
- [0389] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp (SEQ ID NO:26)或其功能变体;
- [0390] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp (SEQID NO:39)或其功能变体;
- [0391] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp (SEQ ID NO:40)或其功能变体;
- [0392] -CRE0006\_AQP4 (SEQ ID NO:37)或其功能变体,和CRE0012\_Arc (SEQID NO:21)或其

其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体;

[0393] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQID NO:25)或其功能变体;

[0394] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0395] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQID NO:21)或其功能变体,可操作地连接至CRE0015\_GAP43\_mp(SEQID NO:39)或其功能变体;

[0396] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQID NO:21)或其功能变体,可操作地连接至CRE0016\_Eno2\_mp(SEQ ID NO:40)或其功能变体;或

[0397] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQID NO:21)或其功能变体,可操作地连接至CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。

[0398] 适当地,合成的NS特异性启动子可以包括CRE0002\_S100B(SEQ ID NO:24)或其功能变体。CRE0002\_S100B(SEQ ID NO:24)包括CRE0009\_S100B(SEQ ID NO:23)和额外的序列,即CRE0009\_S100B(SEQ ID NO:23)是CRE0002\_S100B(SEQ ID NO:24)的较短版本。因此,CRE0002\_S100B(SEQ ID NO:24)可以被看作是CRE0009\_S100B(SEQ ID NO:23)的(较长)功能变体。在本文公开的方面或实施方案中,可以提供CRE0002\_S100B(SEQ ID NO:24)来代替CRE0009\_S100B(SEQ ID NO:23)。

[0399] 适当地,合成的NS特异性启动子可以包括CRE0005\_GFAP(SEQ ID NO:19)或其功能变体。CRE0005\_GFAP(SEQ ID NO:19)包括CRE0007\_GFAP(SEQ ID NO:20)和额外的序列,即CRE0007\_GFAP(SEQ ID NO:20)是CRE0005\_GFAP(SEQ ID NO:19)的较短版本。因此,CRE0005\_GFAP(SEQ ID NO:19)可以被看作是CRE0007\_GFAP(SEQ ID NO:20)的(较长)功能变体。在本文公开的方面或实施方案中,可以提供CRE0005\_GFAP(SEQ ID NO:19)来代替CRE0007\_GFAP(SEQ ID NO:20)。

[0400] 适当地,合成的NS特异性启动子可以包括CRE0006\_GFAP(SEQ ID NO:27)。CRE0006\_GFAP(SEQ ID NO:27)包括CRE0007\_GFAP(SEQ ID NO:20)和额外的序列,即CRE0007\_GFAP(SEQ ID NO:20)是CRE0006\_GFAP(SEQ ID NO:27)的较短版本。因此,CRE0006\_GFAP(SEQ ID NO:27)可以被看作是CRE0007\_GFAP(SEQ ID NO:20)的(较长)功能变体。在本文公开的方面或实施方案中,可以提供CRE0006\_GFAP(SEQ ID NO:27)以代替CRE0007\_GFAP(SEQ ID NO:20)。

[0401] 适当地,合成的NS特异性启动子可以包括CRE0008\_GFAP(SEQ ID NO:28)。CRE0008\_GFAP(SEQ ID NO:28)包括CRE0007\_GFAP(SEQ ID NO:20)和额外的序列,即CRE0007\_GFAP(SEQ ID NO:20)是CRE0008\_GFAP(SEQ ID NO:28)的较短版本。因此,CRE0008\_GFAP(SEQ ID NO:28)可以被看作是CRE0007\_GFAP(SEQ ID NO:20)的(较长)功能变体。在本文公开的方面或实施方案中,可以提供CRE0008\_GFAP(SEQ ID NO:28)以代替CRE0007\_GFAP(SEQ ID NO:20)。

[0402] 适当地,合成的NS特异性启动子包括以下或由其组成:

[0403] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0404] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp

(SEQ ID NO:26)或其功能变体;

[0405] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;或

[0406] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体。

[0407] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;

[0408] -CRE0005\_GFAP(SEQ ID NO:19)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体;

[0409] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0003\_CEND1\_mp(SEQ ID NO:25)或其功能变体;或

[0410] -CRE0002\_S100B(SEQ ID NO:24)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体,可操作地连接至CRE0009\_SYN1\_mp(SEQ ID NO:26)或其功能变体。

[0411] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0005\_GFAP(SEQ ID NO:19)。

[0412] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0005\_GFAP(SEQ ID NO:19)。

[0413] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0012\_Arc(SEQ ID NO:21)和CRE0005\_GFAP(SEQ ID NO:19)。

[0414] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0012\_Arc(SEQ ID NO:21)和CRE0005\_GFAP(SEQ ID NO:19)。

[0415] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0007\_GFAP(SEQ ID NO:20)。

[0416] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0007\_GFAP(SEQ ID NO:20)。

[0417] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0009\_S100B(SEQ ID NO:23)。

[0418] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0009\_S100B(SEQ ID NO:23)。

[0419] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0002\_S100B(SEQ ID NO:24)。

[0420] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0001\_S100B(SEQ ID NO:22)。

[0421] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0002\_S100B(SEQ ID NO:24)。

[0422] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0001\_S100B(SEQ ID NO:22)。

[0423] 在一个实施方案中,NS特异性启动子包括CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0002\_S100B(SEQ ID NO:24)和CRE0012\_Arc(SEQ ID NO:21)。

[0424] 在一个实施方案中,NS特异性启动子包括CRE0009\_SYN1\_mp (SEQ ID NO:26)、CRE0002\_S100B (SEQ ID NO:24) 和CRE0012\_Arc (SEQ ID NO:21)。

[0425] 在一个实施方案中,NS特异性启动子包括CRE0006\_AQP4 (SEQ ID NO:37) 和CRE0004\_GFAP\_mp (SEQ ID NO:41)。

[0426] 在一个实施方案中,NS特异性启动子包括CRE0008\_AQP4 (SEQ ID NO:38) 和CRE0015\_GAP43\_mp (SEQ ID NO:39)。

[0427] 在一个实施方案中,NS特异性启动子包括CRE0006\_AQP4 (SEQ ID NO:37) 和CRE0015\_GAP43\_mp (SEQ ID NO:39)。

[0428] 在一个实施方案中,NS特异性启动子包括CRE0006\_AQP4 (SEQ ID NO:37) 和CRE0016\_Eno2\_mp (SEQ ID NO:40)。

[0429] 通常优选的是,根据本发明的启动子,其包括SEQ ID NO:19-24、27、28、37、38中的任何一个的变体CRE,保留参考CRE的至少25%、50%、75%、80%、85%、90%、95%或100%的活性。可以通过比较合适报告基因在参考启动子控制下的表达与包含替换的CRE的其他相同启动子在等同条件下的表达来评估活性的保留。适当地,所述活性可以使用本文所述的实施例进行评估,但也可以使用其他方法。

[0430] 通常优选的是,根据本发明的启动子,其包括SEQ ID NO:25-26、29-32、39-41中的任何一个的变体启动子元件,保留参考启动子元件的至少25%、50%、75%、80%、85%、90%、95%或100%的活性。适当地,所述活性使用本文所述的实施例进行评估,但也可以使用其他方法。

[0431] 适当地,CRE是神经系统特异性CRE,适当地是CNS特异性顺式调节元件和/或PNS特异性顺式调节元件。

[0432] 适当地,启动子元件是最小或近端启动子。优选地,当存在时,近端启动子是NS特异性启动子元件,适当地是CNS特异性近端启动子和/或PNS特异性近端启动子。

[0433] 在一些实施方案中,合成的NS特异性启动子包括根据SEQ ID NO:1-14、33-36中的任何一个的序列或其功能变体或由其组成。

[0434] 在一些实施方案中,合成的NS特异性启动子包括与SEQ ID NO:1-14、33-36中的任何一个具有至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列或由其组成。

[0435] 因此,本发明提供了各种合成的NS特异性启动子及其功能变体。通常优选的是,根据本发明的启动子,其是SEQ ID NO:1-14、33-36中的任何一个的变体,保留参考启动子的至少25%、50%、75%、80%、85%、90%、95%或100%的活性。适当地,所述活性使用本文所述的实施例进行评估,但也可以使用其他方法。

[0436] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:1或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0437] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:2或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0438] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:3或其功能变体或由其组成,并在兴奋性神经元和星形胶质细胞中具有活性。

[0439] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:4或其功能变体或由其组成。

其组成,并在兴奋性神经元和星形胶质细胞中具有活性。

[0440] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:5或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0441] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:6或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0442] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:7或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0443] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:8或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0444] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:9或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0445] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:10或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0446] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:11或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0447] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:12或其功能变体或由其组成,并在神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0448] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:13或其功能变体或由其组成,并在兴奋性神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0449] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:14或其功能变体或由其组成,并在兴奋性神经元、星形胶质细胞和少突胶质细胞中具有活性。

[0450] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:33或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0451] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:34或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0452] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:35或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0453] 在一个实施方案中,合成的NS特异性启动子包括SEQ ID NO:36或其功能变体或由其组成,并在神经元和星形胶质细胞中具有活性。

[0454] 在一些实施方案中,合成的NS特异性启动子SP0013 (SEQ ID NO:1)、SP0014 (SEQ ID NO:2)、SP0030 (SEQ ID NO:5)、SP0031 (SEQ ID NO:6)、SP0032 (SEQ ID NO:7)、SP0019 (SEQ ID NO:9)、SP0020 (SEQ ID NO:10)、SP0021 (SEQ ID NO:11)、SP0022 (SEQ ID NO:12)、SP0011 (SEQ ID NO:33)、SP0034 (SEQ ID NO:34)、SP0035 (SEQ ID NO:35) 和SP0036 (SEQ ID NO:36) 是特别优选的。在一些实施方案中,合成的NS特异性启动子包括SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:33、SEQ ID NO:34、SEQ ID NO:35或SEQ ID NO:36或由其组成。

[0455] 在一个实施方案中,合成的NS特异性启动子包括在星形胶质细胞中具有活性的CNS特异性CRE,适当地,CRE可以选自:CRE0001\_S100B (SEQ ID NO:22)、CRE0009\_S100B (SEQ

ID NO:23)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0002\_S100B(SEQ ID NO:24)、CRE0006\_AQP4(SEQ ID NO:37)、CRE0008\_AQP4(SEQ ID NO:38)和CRE0005\_GFAP(SEQ ID NO:19)。

[0456] 在一个实施方案中,合成的NS特异性启动子包括在少突胶质细胞中具有活性的CNS特异性CRE,适当地,CRE可以选自CRE0001\_S100B(SEQ ID NO:22)、CRE0002\_S100B(SEQ ID NO:24)或CRE0009\_S100B(SEQ ID NO:23)。

[0457] 在一个实施方案中,合成的NS特异性启动子包括在兴奋性神经元中具有活性的CRE,适当地,CRE可以是CRE0012\_Arc(SEQ ID NO:21)。

[0458] 在一个实施方案中,合成的NS特异性启动子包括在神经元中具有活性的启动子元件,适当地,启动子元件可以选自CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0015\_GAP43\_mp(SEQ ID NO:39)、CRE0016\_Eno2\_mp(SEQ ID NO:40)或CRE0013\_SYN1\_mp(SEQ ID NO:32)。

[0459] 在一个实施方案中,合成的NS特异性启动子包括在星形胶质细胞中具有活性的启动子元件,适当地,启动子元件可以是CRE0004\_GFAP\_mp(SEQ ID NO:41)。

[0460] 在本发明的另一个方面,提供了一种NS特异性顺式调节元件(CRE),其包括根据SEQ ID NO:19-24、27、28、37、38中的任何一个的序列或其任何一个的功能变体或由其组成。在一些实施方案中,NS特异性CRE包括与SEQ ID NO:19-24、27、28、37、38中的任何一个具有至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列。

[0461] 当与启动子元件可操作地连接时,NS特异性CRE可以在CNS中具有活性,即CNS特异性CRE。当与启动子元件可操作地连接时,NS特异性CRE可以在PNS中具有活性,即PNS特异性CRE。当与启动子元件可操作地连接时,NS特异性CRE可以在CNS和PNS中具有活性。

[0462] 一般来说,根据本发明的NS特异性CRE,其是SEQ ID NO:19-24、27、28、37、38中的任何一个的变体,保留参考CRE的至少25%、50%、75%、80%、85%、90%、95%或100%的活性。可以通过比较合适报告基因在参考启动子控制下的表达与包含替换的CRE的其他相同启动子在等同条件下的表达来评估活性的保留。适当地,所述活性是使用本文所述的实施例进行评估的,但也可以使用其他方法。

[0463] 适当地,根据本发明的CRE或其功能变体可以尽可能短,同时仍保留活性。适当地,根据本发明的CRE或其功能变体可以被定义为“缩短的”,并且可以具有低于600bp、低于550bp、低于500bp、低于450bp或低于400bp的长度。优选地,根据本发明的CRE或其功能变体的长度为500或更少的核苷酸,更优选450或更少的核苷酸。

[0464] 适当地,根据本发明的CRE可以与启动子元件可操作地连接。启动子元件可以是最小或近端启动子。当存在时,优选地,近端启动子是NS特异性近端启动子,适当地是CNS特异性近端启动子。适当地,启动子元件可以是CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0015\_GAP43\_mp(SEQ ID NO:39)、CRE0016\_Eno2\_mp(SEQ ID NO:40)、CRE0004\_GFAP\_mp(SEQ ID NO:41)或CRE0013\_SYN1\_mp(SEQ ID NO:32),或其功能变体。

[0465] 在与最小或近端启动子可操作地连接时,CNS特异性顺式调节元件(CRE)可以在星形胶质细胞中具有活性,适当地,CRE可以选自:CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0002\_S100B(SEQ ID NO:24)、CRE0006\_AQP4(SEQ ID NO:37)、CRE0008\_AQP4(SEQ ID NO:38)和CRE0005\_GFAP(SEQ ID NO:19)。在一些实施方案中,当与最小或近端启动子可操作地连接时,在星形胶质细胞中具有活性的CRE,或其功能变体,可以具有低于600bp、低于550bp、低于500bp、低于450bp或低于400bp的长度。当与最小或近端启动子可操作地连接时,在星形胶质细胞中具有活性的CRE,或其功能变体,优选具有500或更少的核苷酸,更优选400或更少的核苷酸的长度。根据本发明的CNS特异性CRE可以在星形胶质细胞中具有活性并且是缩短的,即选自:CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)和CRE0002\_S100B(SEQ ID NO:24)。

[0466] 在与最小或近端启动子可操作地连接时,CNS特异性顺式调节元件(CRE)可以在少突胶质细胞中具有活性,适当地,CRE可以选自CRE0001\_S100B(SEQ ID NO:22)、CRE0002\_S100B(SEQ ID NO:24)或CRE0009\_S100B(SEQ ID NO:23)。在一些实施方案中,当与最小或近端启动子可操作地连接时,在少突胶质细胞中具有活性的CRE,或其功能变体,可以具有低于600bp、低于550bp、低于500bp、低于450bp或低于400bp的长度。当与最小或近端启动子可操作地连接时,在少突胶质细胞中具有活性的CRE,或其功能变体,优选为500或更少的核苷酸,更优选400或更少的核苷酸。根据本发明的CNS特异性CRE可以在少突胶质细胞中具有活性并且是缩短的,即选自:CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)和CRE0002\_S100B(SEQ ID NO:24)。

[0467] 当与最小或近端启动子可操作地连接时,NS特异性顺式调节元件(CRE)可以在兴奋性神经元中具有活性,适当地,CRE可以是CRE0012\_Arc(SEQ ID NO:21)。

[0468] 根据本发明的NS特异性CRE优选是缩短的,即选自:CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)和CRE0002\_S100B(SEQ ID NO:24)。

[0469] 适当地,根据本发明的CRE可以与一个或多个额外的CRE结合,以形成顺式调节模块(CRM)。适当地,所述一个或多个额外的CRE可以根据SEQ ID NO:19-24、27、28的CRE或其功能变体,或者它们可以是其他CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8,来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。适当地,额外的CRE是NS特异性的,适当地是CNS特异性的和/或PNS特异性的。

[0470] 在本发明的另一个方面,提供了一种合成的NS特异性启动子,其包括根据SEQ ID NO:19-24、27、28、37、38中的任何一个的CRE或其功能变体或由其组成。在一些实施方案中,CRE可以与启动子元件可操作地连接。在一些实施方案中,启动子元件可以是最小或近端启动子。优选地,近端启动子是NS特异性近端启动子,适当地是CNS特异性和/或PNS特异性近端启动子。适当地,启动子元件可以是CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0015\_GAP43\_mp(SEQ ID NO:39)、CRE0016\_

Eno2\_mp (SEQ ID NO:40)、CRE0004\_GFAP\_mp (SEQ ID NO:41) 或 CRE0013\_SYN1\_mp (SEQ ID NO:32), 或其功能变体。

[0471] 在本发明的另一个方面, 提供了一种包含根据本发明的CRE的NS特异性顺式调节模块(CRM)。适当地, CRM包括SEQ ID NO:15-18中的任何一个或其功能变体或由其组成。

[0472] 当与启动子元件可操作地连接时, NS特异性CRM可以在CNS中具有活性, 即CNS特异性CRM。当与启动子元件可操作地连接时, NS特异性CRM可以在PNS中具有活性, 即PNS特异性CRM。当与启动子元件可操作地连接时, NS特异性CRM可以在CNS和PNS中具有活性。

[0473] 适当地, 合成的NS特异性CRM包括与SEQ ID NO:15-18中的任何一个具有至少70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同一性的序列或由其组成。

[0474] 根据本发明的CRM可以与启动子元件可操作地连接。启动子元件可以是最小或近端启动子。当存在时, 优选地, 近端启动子是NS特异性近端启动子, 适当地是CNS特异性和/或PNS特异性近端启动子。适当地, 启动子元件可以是CRE0003\_CEND1\_mp (SEQ ID NO:25)、CRE0009\_SYN1\_mp (SEQ ID NO:26)、CRE0001\_CEND1\_mp (SEQ ID NO:29)、CRE0002\_CEND1\_mp (SEQ ID NO:30)、CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0015\_GAP43\_mp (SEQ ID NO:39)、CRE0016\_Eno2\_mp (SEQ ID NO:40)、CRE0004\_GFAP\_mp (SEQ ID NO:41) 或 CRE0013\_SYN1\_mp (SEQ ID NO:32), 或其功能变体。

[0475] 根据本发明的合成的NS特异性启动子可以包括与启动子元件(通常是最小或近端启动子)可操作地连接的根据本发明的CRM。近端启动子优选是NS特异性近端启动子, 适当地是CNS特异性和/或PNS特异性启动子元件。适当地, 启动子元件可以是CRE0003\_CEND1\_mp (SEQ ID NO:25)、CRE0009\_SYN1\_mp (SEQ ID NO:26)、CRE0001\_CEND1\_mp (SEQ ID NO:29)、CRE0002\_CEND1\_mp (SEQ ID NO:30)、CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0015\_GAP43\_mp (SEQ ID NO:39)、CRE0016\_Eno2\_mp (SEQ ID NO:40)、CRE0004\_GFAP\_mp (SEQ ID NO:41) 或 RE0013\_SYN1\_mp (SEQ ID NO:32), 或其功能变体。

[0476] 适当地, CRM包括以下顺式调节元件(CRE)或其功能变体中的至少一个:

[0477] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体;

[0478] -CRE0002\_S100B (SEQ ID NO:24) 或其功能变体;

[0479] -CRE0005\_GFAP (SEQ ID NO:19) 或其功能变体;

[0480] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体;

[0481] -CRE0009\_S100B (SEQ ID NO:23) 或其功能变体;

[0482] -CRE0012\_Arc (SEQ ID NO:21) 或其功能变体;

[0483] -CRE0006\_GFAP (SEQ ID NO:27) 或其功能变体;

[0484] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体;

[0485] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体; 和

[0486] -CRE0008\_GFAP (SEQ ID NO:28) 或其功能变体。

[0487] 适当地, CRM包括以下顺式调节元件(CRE)或其功能变体中的至少一个:

[0488] -CRE0001\_S100B (SEQ ID NO:22) 或其功能变体;

[0489] -CRE0007\_GFAP (SEQ ID NO:20) 或其功能变体;

[0490] -CRE0006\_AQP4 (SEQ ID NO:37) 或其功能变体;

[0491] -CRE0008\_AQP4 (SEQ ID NO:38) 或其功能变体; 和

[0492] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体。

[0493] 适当地,CRM可以包括CRE0002\_S100B(SEQ ID NO:24)或其功能变体、CRE0005\_GFAP(SEQ ID NO:19)或其功能变体、CRE0006\_GFAP(SEQ ID NO:27)或其功能变体、CRE0008\_GFAP(SEQ ID NO:28)或其功能变体、CRE0006\_AQP4(SEQ ID NO:37)或其功能变体、CRE0008\_AQP4(SEQ ID NO:38)或其功能变体。

[0494] 适当地,CRM可以包括根据本发明的一个或多个额外的CRE或其他NS特异性CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8、来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。适当地,CRM可以另外包括CRE0012\_Arc(SEQ ID NO:21)或其功能变体。

[0495] 适当地,合成的NS特异性CRM包括以下或由其组成:

[0496] -CRE0001\_S100B(SEQ ID NO:22)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体;

[0497] -CRE0007\_GFAP(SEQ ID NO:20)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体;

[0498] -CRE0006\_AQP4(SEQ ID NO:37)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体;

[0499] -CRE0008\_AQP4(SEQ ID NO:38)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体;或

[0500] -CRE0009\_S100B(SEQ ID NO:23)或其功能变体,和CRE0012\_Arc(SEQ ID NO:21)或其功能变体。

[0501] 在一个实施方案中,NS特异性CRM包括CRE0012\_Arc(SEQ ID NO:21)和CRE0005\_GFAP(SEQ ID NO:19)。适当地,在这样的实施方案中,CRM可以包括SEQ ID NO:15或SEQ ID NO:16或由其组成。适当地,当与最小启动子可操作地连接时,NS特异性CRM在兴奋性神经元和星形胶质细胞中具有活性。

[0502] 在一个实施方案中,NS特异性CRM包括CRE0002\_S100B(SEQ ID NO:24)和CRE0012\_Arc(SEQ ID NO:21)。适当地,在这样的实施方案中,CRM可以包括SEQ ID NO:17或SEQ ID NO:18或由其组成。适当地,当与最小启动子可操作地连接时,NS特异性CRM在兴奋性神经元、少突胶质细胞和星形胶质细胞中具有活性。

[0503] 在一个实施方案中,合成的NS特异性CRM包括在星形胶质细胞中具有活性的CRE,适当地,CRE可以选自:CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0002\_S100B(SEQ ID NO:24)、CRE0006\_AQP4(SEQ ID NO:37)、CRE0008\_AQP4(SEQ ID NO:38)和CRE0005\_GFAP(SEQ ID NO:19)。

[0504] 在一个实施方案中,合成的NS特异性CRM包括在少突胶质细胞中具有活性的CRE,适当地,CRE可以选自CRE0001\_S100B(SEQ ID NO:22)、CRE0002\_S100B(SEQ ID NO:24)或CRE0009\_S100B(SEQ ID NO:23)。

[0505] 在一个实施方案中,合成的NS特异性CRM包括在兴奋性神经元中具有活性的CRE,适当地,CRE可以是CRE0012\_Arc(SEQ ID NO:21)。

[0506] 在本发明的另一个方面,提供了一种启动子元件(最小或近端启动子),其包括根

据SEQ ID NO:25-26、29-32、39-41中的任何一个的序列或其功能变体或由其组成。

[0507] 启动子元件可以在CNS中具有活性,即CNS特异性启动子元件。启动子元件可以在PNS中具有活性,即PNS特异性启动子元件。启动子元件可以在PNS和CNS中具有活性。

[0508] 在本发明的另一个方面,提供了一种包含所述最小或近端启动子的合成启动子,适当地包含所述最小或近端启动子的合成的NS特异性启动子。适当地,最小或近端启动子的功能变体包括与SEQ ID NO:25-26、29-32、39-41具有至少60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%同一性的序列。

[0509] 适当地,根据本发明的启动子元件或其功能变体可以尽可能短,同时仍然保留活性。适当地,根据本发明的启动子元件或其功能变体可被定义为“缩短的”,并且可以具有小于500bp、小于450bp、小于400bp、小于350bp、小于300bp、小于250bp、小于200bp、小于150bp或小于100bp的长度。优选地,根据本发明的启动子元件或其功能变体的长度为350或更少的核苷酸,更优选300或更少的核苷酸。

[0510] 最小或近端启动子可以在神经元中具有活性,适当地,启动子元件可以选自CRE0003\_CEND1\_mp(SEQ ID NO:25)或CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0015\_GAP43\_mp(SEQ ID NO:39)、CRE0016\_Eno2\_mp(SEQ ID NO:40)或CRE0013\_SYN1\_mp(SEQ ID NO:32),或其功能变体。适当地,神经元可以是CNS神经元和/或PNS神经元。适当地,在神经元中具有活性的最小或近端启动子或其功能变体的长度小于500bp、小于450bp、小于400bp、小于350bp、小于300bp、小于250bp、小于200bp、小于150bp或小于100bp。优选地,在神经元中具有活性的最小或近端启动子或其功能变体的长度为300或更少的核苷酸,更优选250或更少的核苷酸。

[0511] 最小或近端启动子可以在星形胶质细胞中具有活性,适当地,启动子元件可以是CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体。适当地,在星形胶质细胞中具有活性的最小或近端启动子或其功能变体的长度小于500bp、小于450bp、小于400bp、小于350bp、小于300bp、小于250bp、小于200bp、小于150bp或小于100bp。优选地,在星形胶质细胞中具有活性的最小或近端启动子或其功能变体的长度为300或更少的核苷酸,更优选250或更少的核苷酸。

[0512] 根据本发明的启动子元件优选是缩短的,即选自:CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0013\_SYN1\_mp(SEQ ID NO:32)。

[0513] 根据本发明的启动子元件可以由CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)或CRE0003\_CEND1\_mp(SEQ ID NO:25)组成。CRE0003\_CEND1\_mp(SEQ ID NO:25)包括CRE0001\_CEND1\_mp(SEQ ID NO:29)和额外的序列,即CRE0001\_CEND1\_mp(SEQ ID NO:29)是CRE0003\_CEND1\_mp(SEQ ID NO:25)的较短版本。CRE0002\_CEND1\_mp(SEQ ID NO:30)包括CRE0001\_CEND1\_mp(SEQ ID NO:29)和额外的序列,即CRE0001\_CEND1\_mp(SEQ ID NO:29)是CRE0002\_CEND1\_mp(SEQ ID NO:30)的较短版本。同样,CRE0002\_CEND1\_mp(SEQ ID NO:30)包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和额外的序列,即CRE0003\_CEND1\_mp(SEQ ID NO:25)是CRE0002\_CEND1\_mp(SEQ ID NO:30)的较短版本。在本文公开的方面或实施方案中,可以提供CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)或CRE0003\_CEND1\_mp(SEQ ID NO:25)来代替彼此。

[0514] 根据本发明的启动子元件可以由CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0013\_SYN1\_mp (SEQ ID NO:32)或CRE0009\_SYN1\_mp (SEQ ID NO:26)组成。CRE0009\_SYN1\_mp (SEQ ID NO:26)包括CRE0013\_SYN1\_mp (SEQ ID NO:32)和额外的序列,即CRE0013\_SYN1\_mp (SEQ ID NO:32)是CRE0009\_SYN1\_mp (SEQ ID NO:26)的较短版本。CRE0008\_SYN1\_mp (SEQ ID NO:31)包括CRE0013\_SYN1\_mp (SEQ ID NO:32)和额外的序列,即CRE0013\_SYN1\_mp (SEQ ID NO:32)是CRE0008\_SYN1\_mp (SEQ ID NO:31)的较短版本。同样,CRE0008\_SYN1\_mp (SEQ ID NO:31)包括CRE0009\_SYN1\_mp (SEQ ID NO:26)和额外的序列,即CRE0009\_SYN1\_mp (SEQ ID NO:26)是CRE0008\_SYN1\_mp (SEQ ID NO:31)的较短版本。在本文公开的方面或实施方案中,可以提供CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0013\_SYN1\_mp (SEQ ID NO:32)或CRE0009\_SYN1\_mp (SEQ ID NO:26)来代替彼此。

[0515] 适当地,最小或近端启动子可以与CRE或CRM可操作地连接。所述CRE可以是根据本发明的CRE或任何其他CRE。适当地,当所述CRE不是根据本发明的CRE时,所述CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8、来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。所述CRM可以是根据本发明的CRM,或可以包括根据本发明的CRE。适当地,所述CRE或CRM是NS特异性的,适当地是CNS特异性和/或PNS特异性的。

[0516] 适当地,根据本发明的近端启动子可以与一个或多个近端启动子可操作地连接。根据本发明的合成的NS特异性启动子可以包括两个近端启动子或由其组成。适当地,根据本发明的合成的NS特异性启动子可以包括两个或更多个近端启动子或由其组成。适当地,近端启动子是NS特异性近端启动子,适当地是CNS特异性近端启动子和/或PNS特异性近端启动子。适当地,近端启动子可以选自CRE0003\_CEND1\_mp (SEQ ID NO:25)、CRE0009\_SYN1\_mp (SEQ ID NO:26)、CRE0001\_CEND1\_mp (SEQ ID NO:29)、CRE0002\_CEND1\_mp (SEQ ID NO:30)、CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0015\_GAP43\_mp (SEQ ID NO:39)、CRE0016\_Eno2\_mp (SEQ ID NO:40)、CRE0004\_GFAP\_mp (SEQ ID NO:41)或CRE0013\_SYN1\_mp (SEQ ID NO:32),或其功能变体。适当地,根据本发明的合成的NS特异性启动子可以包括以下组中的任何两个或由其组成:CRE0003\_CEND1\_mp (SEQ ID NO:25)、CRE0009\_SYN1\_mp (SEQ ID NO:26)、CRE0001\_CEND1\_mp (SEQ ID NO:29)、CRE0002\_CEND1\_mp (SEQ ID NO:30)、CRE0008\_SYN1\_mp (SEQ ID NO:31)、CRE0015\_GAP43\_mp (SEQ ID NO:39)、CRE0016\_Eno2\_mp (SEQ ID NO:40)、CRE0004\_GFAP\_mp (SEQ ID NO:41)和CRE0013\_SYN1\_mp (SEQ ID NO:32),或其功能变体。适当地,所述至少两个近端启动子可以与根据本发明的CRE或CRM或其他CRE或CRM可操作地连接。

[0517] 本发明的CRE、最小/近端启动子或启动子可以在CNS和/或PNS中具有活性。在一些实施方案中,本发明的CRE、最小/近端启动子或合成启动子在CNS中具有活性,即CNS特异性CRE、最小/近端启动子或合成启动子。在一些实施方案中,本发明的CRE、最小/近端启动子或合成启动子在PNS中具有活性,即PNS特异性CRE、最小/近端启动子或合成启动子。在一些实施方案中,本发明的CRE、最小/近端启动子或启动子在CNS和PNS中具有活性。

[0518] 本发明的CRE、最小/近端启动子或启动子可以在CNS的特定区域,优选在CNS的特定区域,或在特定的CNS细胞类型,或在CNS细胞类型的组合或两者的组合中具有活性。因此,适当地,本发明的CRE、最小/近端启动子或启动子是CNS特异性的。

[0519] 本发明的CRE、最小/近端启动子或启动子可以在PNS的特定区域,优选在PNS的特定区域,或在特定的PNS细胞类型,或在PNS细胞类型的组合或两者的组合中具有活性。因此,适当地,本发明的CRE、最小/近端启动子或启动子是PNS特异性的。

[0520] 本发明的CRE、最小/近端启动子或启动子可以在NS的一个或多个不同部分,适当地在CNS的一个或多个不同部分和/或PNS的一个或多个不同部分中具有活性。

[0521] 本发明的CRE、最小/近端启动子或启动子可以在CNS的一个或多个不同部分中具有活性。CNS主要由大脑和脊髓组成。视网膜、视神经、嗅神经和嗅觉上皮细胞有时被认为是与大脑和脊髓一起的CNS的一部分。这是因为它们直接与脑组织相连,没有中间的神经纤维。适当地,本发明的CRE、最小/近端启动子或启动子可以在大脑和脊髓中具有活性。适当地,本发明的CRE、最小/近端启动子或启动子可以在大脑中具有活性,但在脊髓或CNS的任何其他部分中没有活性。适当地,本发明的CRE、最小/近端启动子或启动子可以在脊髓中具有活性,但在大脑中没有活性。优选地,本发明的CRE、最小/近端启动子或启动子可以在大脑和脊髓中具有活性。适当地,本发明的CRE、最小/近端启动子或启动子可以在大脑中具有活性。适当地,本发明的CRE、最小/近端启动子或启动子可以在大脑内的一个或多个不同区域中具有活性。

[0522] 大脑区域的非限制性实例包括:额叶、旁叶、枕叶、颞叶(其包括海马体和杏仁核)、小脑、中脑(其包括基底神经节,而基底神经节又包括纹状体和黑质)、脑桥、延髓和间脑(其包括丘脑和下丘脑)。脊髓区域的非限制性实例包括:颈椎、胸椎、腰椎、骶椎和尾椎。在一些实施方案中,可能希望本发明的CRE、CRM、最小/近端启动子或启动子在大脑中显示广泛的活性。在一些实施方案中,可能希望本发明的CRE、CRM、最小/近端启动子或启动子在大脑和脊髓中显示广泛的活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在大脑或CNS(泛CNS)的所有部分,优选在大脑的所有区域中都具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在大脑中具有活性,但在CNS的其他部分如脊髓中没有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在脊髓中具有活性,但在CNS的其他部分如大脑中没有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述脊髓的1、2、3、4或5个区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在脊髓的大部分区域,即上述5个大脑区域中的至少3个、至少4个或全部5个中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述大脑的1、2、3、4、5、6、7、8或9个区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述大脑的1、2、3、4、5、6、7、8或9个区域和脊髓中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在大脑的大部分区域,即上述9个大脑区域中的至少5个、至少6个、至少7个、至少8个或全部9个中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述大脑的4至6个区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述大脑的2至4个区域中具有活性。适当地,本发明的CRE、CRM、最小/近端启动子或启动子优选在以下任何或所有大脑区域中具有活性:额叶、旁叶、枕叶、颞叶(其包括海马体)、小脑、中脑(其包括基底神经节,而基底神经节又包括纹状体和黑质)和间脑(其包括丘脑和下丘脑)。这可能是特别优选的,因为纹状体、黑质、大脑皮层的第3、5和6层(大脑皮层包括额叶、旁叶、枕叶和颞叶)、海马体、小脑、下丘脑和丘脑,这些都是亨廷顿

氏病中受影响最严重的大脑区域。适当地,本发明的CRE、CRM、最小/近端启动子或启动子优选在脊髓和以下任何或所有大脑区域中具有活性:额叶、旁叶、枕叶、颞叶(其包括海马体)、小脑、中脑(其包括基底神经节,而基底神经节又包括纹状体和黑质)和间脑(其包括丘脑和下丘脑)。

[0523] 本发明的CRE、最小/近端启动子或启动子可以在PNS的一个或多个不同部分中具有活性。PNS是指神经系统中在大脑和脊髓以外的部分。周围神经系统的非限制性实例包括颅神经、臂丛神经、胸腹神经、腰丛神经、骶丛神经和神经肌肉接头。在一些实施方案中,可能希望本发明的CRE、CRM、最小/近端启动子或启动子在PNS中显示广泛的活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述PNS的1、2、3、4、5或6个区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在PNS的大部分区域,即上述PNS的6个区域中的至少4个、至少5个或全部6个中具有活性。

[0524] 本发明的CRE、最小/近端启动子或启动子可以在上述PNS和CNS的一个或多个不同区域的组合中具有活性。

[0525] 在一些实施方案中,可能希望本发明的CRE、最小/最大启动子或启动子在CNS的一个区域,适当地在大脑的一个区域,或在脊髓的一个区域中显示主要活性。适当地,可能希望本发明的CRE、最小/近端启动子或启动子在大脑的一个区域显示出活性,但在大脑、脊髓、CNS或PNS的其余部分没有活性或只有极少的活性。适当地,可能希望本发明的CRE、最小/近端启动子或启动子在脊髓的一个区域显示出活性,但在脊髓、大脑、CNS或PNS的其余部分没有活性或只有极少的活性。在一些实施方案中,本发明的CRE、最小/近端启动子或启动子只在上述大脑或脊髓区域中的一个区域具有活性。

[0526] 在一些实施方案中,可能希望本发明的CRE、最小/近端启动子或启动子在PNS的一个区域显示主要活性。适当地,可能希望本发明的CRE、最小/近端启动子或启动子在PNS的一个区域显示出活性,但在大脑、CNS或PNS的其他部分没有活性或只有极少的活性。在一些实施方案中,本发明的CRE、最小/近端启动子或启动子只在上述PNS区域中的一个区域具有活性。

[0527] 本发明的CRE、最小/近端启动子或启动子可以在神经系统的各种细胞中具有活性。这些细胞可以是CNS细胞和/或PNS细胞。

[0528] 本发明的CRE、最小/近端启动子或启动子可以在PNS的各种细胞中具有活性。PNS中的主要细胞类型是神经元、卫星细胞和许旺细胞。其他细胞类型也可能存在,特别是在炎症状态下。在一些实施方案中,可能希望启动子在许多不同的细胞类型中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子基本上在PNS的所有细胞(如神经元、卫星细胞和许旺细胞)中都具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述PNS细胞类型中的至少两种PNS细胞类型,如神经元和卫星细胞中具有活性。

[0529] 本发明的CRE、最小/近端启动子或启动子可以在CNS的各种细胞中具有活性。CNS中的主要细胞类型是神经元、星形胶质细胞、少突胶质细胞、小胶质细胞和室管膜细胞。其他细胞类型也可能存在,特别是在炎症状态下。在一些实施方案中,可能希望启动子在许多不同的细胞类型中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子基本上在CNS的所有细胞(如神经元、星形胶质细胞、少突胶质细胞、小胶质细胞、室管

膜细胞)中都具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的至少四种CNS细胞类型,如神经元、星形胶质细胞、小胶质细胞和少突胶质细胞中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的至少三种CNS细胞类型,如神经元、星形胶质细胞和少突胶质细胞中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的至少两种CNS细胞类型,如神经元和星形胶质细胞中具有活性。

[0530] 本发明的CRE、最小/近端启动子或启动子可以在上述的PNS和/或CNS的任何细胞组合中具有活性。本发明的CRE、最小/近端启动子或启动子可以在神经元中具有活性。

[0531] 在一些实施方案中,可能希望启动子在有限的CNS细胞类型中具有活性,或在不超过一种CNS细胞类型中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的不超过4、3、2或1种CNS细胞类型中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的不超过三种CNS细胞类型,如神经元、星形胶质细胞和少突胶质细胞中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述CNS细胞类型中的不超过两种CNS细胞类型,如神经元和星形胶质细胞中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子只在上述CNS细胞类型中的一种CNS细胞类型,如神经元中具有活性。

[0532] 在一些实施方案中,可能希望启动子在有限的PNS细胞类型中具有活性,或在不超过一种PNS细胞类型中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述PNS细胞类型中的不超过2种或1种PNS细胞类型中具有活性。

[0533] 在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述PNS细胞类型中的不超过两种PNS细胞类型,如神经元和卫星细胞中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子只在上述PNS细胞类型中的一种PNS细胞类型,如神经元中具有活性。

[0534] 在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在CNS细胞的特定亚型中具有活性,例如棘状GABA能投射神经元(发现于纹状体)、多巴胺能神经元(发现于黑质)、锥体神经元(发现于大脑皮层的第3、5和6层)、海马神经元(发现于海马体)、浦肯野细胞(发现于小脑)、侧管核神经元(发现于下丘脑/丘脑)。在一些具体的优选实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述任何或所有神经元亚型中具有活性。在一些优选实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述任何或所有神经元亚型中具有活性,但在其他CNS或PNS细胞类型或亚型中没有活性。在一些优选实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述任何或所有神经元亚型和其他CNS细胞类型或其他CNS细胞亚型中具有活性。

[0535] 在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在GABA能或谷氨酸能神经元中具有活性。

[0536] 在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在PNS细胞的特定亚型,如运动神经元、PNS交感神经元、PNS肠道神经元和感觉神经元中具有活性。在一些具体的优选实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述任何或所有PNS神经元亚型中具有活性,但在其他神经系统细胞类型或亚型中没有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在上述任何或所有PNS神经元亚型和其

他CNS和/或PNS细胞类型或其他CNS细胞亚型中具有活性。

[0537] 在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在特定类型的CNS细胞或CNS细胞亚型中,以及在CNS的特定区域,适当地在大脑的特定区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在PNS细胞的特定类型或PNS细胞的亚类型中,以及在PNS的特定区域中具有活性。在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在CNS和/或PNS细胞的一个或多个特定类型或亚类型中,以及在CNS和/或PNS的特定区域中具有活性。

[0538] 本发明的CRE、CRM、最小/近端启动子或启动子在NS以外的组织中可能有也可能没有活性。NS以外的组织的非限制性实例是:心脏、肝脏、肾脏、骨骼肌和脾脏。适当地,在一些实施方案中,本发明的CRE、CRM、最小/近端启动子或启动子在NS以外的组织或细胞中没有活性或只有最小的活性。在一些实施方案中,与在NS以外的组织或细胞中的表达相比,本发明的CRE、CRM、最小/近端启动子或启动子在NS中表达更高。适当地,本发明的CRE、CRM、最小/近端启动子或启动子在上述NS以外的组织中的不超过1、2、3或4种组织中具有活性,适当地在全身递送如静脉内递送中具有活性。

[0539] 适当地,在一些实施方案中,可能希望本发明的CRE、CRM、最小/近端启动子或启动子在NS中具有活性,但在NS以外的其他组织中也具有活性。适当地,本发明的CRE、CRM、最小/近端启动子或启动子可以在上述NS以外的组织中的至少1、2、3、4或5种组织中具有活性,适当地在全身递送如静脉内递送中具有活性。

[0540] 由本发明的启动子驱动的表达在所需组织或细胞中可持续至少1小时、2小时、3小时、4小时、5小时、6小时、7小时、8小时、9小时、10小时、11小时、12小时、13小时、14小时、15小时、16小时、17小时、18小时、19小时、20小时、21小时、22小时、23小时、1天、2天、3天、4天、5天、6天、1周、8天、9天、10天、11天、12天、13天、2周、15天、16天、17天、18天、19天、20天、3周、22天、23天、24天、25天、26天、27天、28天、29天、30天、31天、1个月、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月、1年、13个月、14个月、15个月、16个月、17个月、18个月、19个月、20个月、21个月、22个月、23个月、2年、3年、4年、5年、6年、7年、8年、9年、10年、15年、20年、30年、40年、50年、60年、70年、80年、90年、100年。由本发明的启动子驱动的表达在所需组织或细胞中可持续超过1个月、2个月、3个月、4个月、5个月、6个月、7个月、8个月、9个月、10个月、11个月、1年、13个月、14个月、15个月、16个月、17个月、18个月、19个月、20个月、21个月、22个月、23个月、2年、3年、4年、5年、6年、7年、8年、9年、10年、20年、30年、40年、50年、60年、70年、80年、90年、100年。表达可持续1-5小时、1-12小时、1-2天、1-5天、1-2周、1-3周、1-4周、1-2个月、1-4个月、1-6个月、2-6个月、3-6个月、3-9个月、4-8个月、6-12个月、1-2年、1-5年、2-5年、3-6年、3-8年、4-8年、5-10年、10-15年、15-20年、20-30年、30-40年、40-50年、50-60年、60-70年、80-90年或90-100年。

[0541] 在本发明的另一个方面,提供了一种表达盒,其包括与编码表达产物的序列可操作地连接的本发明任何方面的合成的NS特异性启动子。适当地,所述表达产物是基因,例如转基因。在一些实施方案中,所述表达产物是治疗性表达产物。

[0542] 在另一个方面,提供了一种载体,其包括根据本发明的合成的NS特异性启动子或表达盒。在一些实施方案中,所述载体是表达载体。在一些实施方案中,所述载体是病毒载体。在一些实施方案中,所述载体是基因治疗载体,适当地是AAV载体、腺病毒载体、逆转录

病毒载体、单纯疱疹病毒载体或慢病毒载体。慢病毒载体已被广泛用作CNS的基因转移工具,并且已知能够成功地转导神经元、星形胶质细胞和少突胶质细胞(Jakobsson和Lundberg,2006)。它们是有利的,因为它们有相对大的克隆能力,而且病毒基因不被表达。特别优选的慢病毒载体系统是基于HIV-1的(Jakobsson和Lundberg,2006)。单纯疱疹病毒载体和腺病毒载体也显示出用作CNS的基因转移工具的潜力,因为它们显示出对CNS细胞的成功转导,但由于其毒性而不太优选。

[0543] AAV载体在本领域已得到广泛的讨论。AAV载体特别令人感兴趣,因为AAV载体通常不会整合到基因组中,并且不会引起免疫反应。AAV血清型1、2、4、5、8、9、rh10、DJ8、2i8、2.5、2i8G9和2g9(AAV1、AAV2、AAV4、AAV5、AAV8、AAV9、AAVrh10、AAVDJ8、AAV2i8、AAV2.5、AAV2i8G9和AAV2g9)已被指出在NS中实现高效转导。因此,AAV1、AAV2、AAV4、AAV5、AAV8、AAV9、AAVrh10、AAVDJ8、AAV2i8、AAV2.5、AAV2i8G9、AAV2g9及其衍生物是特别优选的AAV血清型。在一些实施方案中,AAV9是特别优选的AAV载体。在其他实施方案中,AAV2g9是特别优选的AAV载体(WO2014/144229)。在另外的其他实施方案中,特别优选的AAV载体是AAVDJ8。在一些实施方案中,AAVrh10是特别优选的AAV载体。适当地,AAV载体包括病毒基因组,所述病毒基因组包括定位在两个倒置末端重复(ITR)之间的本发明的核酸序列。例如,WO2019/028306公开了可用于CNS的各种野生型和修饰的AAV载体。在一个实施方案中,AAV载体递送后能够穿透血脑屏障。在一个实施方案中,本发明的AAV载体是重组AAV病毒载体,其复制有缺陷,在其病毒基因组内缺乏编码功能性Rep和Cap蛋白的序列。这些有缺陷的AAV载体可能缺乏大部分或所有的亲本编码序列,基本上只携带一个或两个AAV ITR序列和目的核酸,用于递送给细胞、组织、器官或生物体。适当地,本文使用的AAV载体包括病毒,所述病毒已被减少到转导核酸有效载荷或目的货物所需的最小组分。以这种方式,AAV载体被工程化为特定的递送载体,同时缺乏野生型病毒中发现的有害复制和/或整合特征。在一个实施方案中,本发明的AAV颗粒是scAAV。在另一个实施方案中,本发明的AAV颗粒是ssAAV。本领域广泛公开了生产和/或修饰AAV颗粒的方法(参见例如WO2000/28004;WO2001/23001;WO2004/112727;WO2005/005610和WO2005/072364,这些都通过引用并入本文)。在一个实施方案中,AAV载体包括允许在血管内(例如静脉内或动脉内)施用后进行血脑屏障渗透的衣壳(参见例如WO2014/144229,其中讨论了例如为有效穿越血脑屏障而工程化的衣壳,例如包括VOY101、VOY201、AAVPHP.N、AAVPHP.A、AAVPHP.B、PHP.B2、PHP.B3、G2A3、G2B4、G2B5、PHP.S及其变体的衣壳或肽插入物)。

[0544] 制备AAV载体的方法在本领域中是众所周知的,并在例如以下文件中进行了描述:美国专利号US6204059、US5756283、US6258595、US6261551、US6270996、US6281010、US6365394、US6475769、US6482634、US6485966、US6943019、US6953690、US7022519、US7238526、US7291498和US7491508、US5064764、US6194191、US6566118、US8137948;或国际公开号W01996039530、W01998010088、W01999014354、W01999/015685、W01999/047691、W02000/055342、W02000/075353和W02001/023597;Methods In Molecular Biology, ed.Richard, Humana Press, NJ(1995);O'Reilly et al, Baculovirus Expression Vectors, A Laboratory Manual, Oxford Univ. Press(1994);Samulski et al., J. Vir. 63: 3822-8(1989);Kajigaya et al, Proc. Nat'l. Acad. Sci. USA 88:4646-50(1991);Ruffing et al., J. Vir. 66:6922-30(1992);Kimbauer et al, Vir., 219:37-44(1996);Zhao et al,

Vir.272:382-93(2000);其中每一项的内容通过引用并入本文。常用于生产重组AAV病毒颗粒的病毒复制细胞包括但不限于HEK293细胞、COS细胞、HeLa细胞、KB细胞和其他哺乳动物细胞系。

[0545] 应该理解的是,病毒表达系统将进一步被修改,以包括在使用本文所述的方法生产给定病毒载体期间补充其所需的任何必要元件。例如,在某些实施方案中,核酸盒两侧为末端重复序列。在一个实施方案中,为了生产rAAV载体,AAV表达系统将进一步包括重组AAV质粒、表达Rep的质粒、表达Cap的质粒和腺病毒辅助质粒中的至少一种。给定病毒载体的补充元件是本领域众所周知的,熟练的从业者将能够相应地修饰本文所述的病毒表达系统。

[0546] 用于制造AAV载体的病毒表达系统(例如,AAV表达系统)可以进一步包括反式的复制(Rep)基因和/或衣壳(Cap)基因,例如,在诱导型启动子的控制下。Rep和Cap的表达可以在一个诱导型启动子的控制下,使这些基因的表达一起“打开”,或者在两个单独的诱导型启动子的控制下,其由不同的诱导剂“打开”。在AAV基因组的左侧(靠近5'端)有两个启动子,称为p5和p19,从这两个启动子可以产生两个不同长度的重叠的信使核糖核酸(mRNA)。这些中的每一个都包含内含子,所述内含子可以被剪掉或不剪掉,从而产生四个潜在的Rep基因;Rep78、Rep68、Rep52和Rep40。Rep基因(特别是Rep78和Rep68)在自我引发行中与倒置末端重复(ITR)形成的发夹结合,并在发夹内指定的末端解析位点上进行切割。它们是AAV基因组的AAVS1特异性整合所必需的。所有四个Rep蛋白都被证明能与ATP结合并具有螺旋酶活性。阳性感应AAV基因组的右侧(靠近3'端)编码三个衣壳蛋白VP1、VP2和VP3的重叠序列,所述衣壳蛋白与一个启动子(指定为p40)可操作地连接。cap基因产生额外的非结构蛋白,称为组装激活蛋白(AAP)。该蛋白由开放阅读框2(ORF2)产生,并且对衣壳组装过程至关重要。制造AAV载体的必要元件在本领域是已知的,并且可以进一步查阅,例如美国专利号US5478745A;US5622856A;US5658776A;US6440742B1;US6632670B1;US6156303A;US8007780B2;US6521225B1;US7629322B2;US6943019B2;US5872005A;和美国专利申请号US2017/0130245;US20050266567A1;US20050287122A1;其中每一项的内容通过引用整体并入本文。

[0547] 在一个实施方案中,用于生产AAV载体的细胞在悬浮液中培养。在另一个实施方案中,细胞在无动物组分的条件下培养。无动物组分培养基可以是与给定细胞系(例如HEK293细胞)兼容的任何无动物组分培养基(例如,无血清培养基)。本领域已知的能够传播AAV载体的任何细胞系都可以使用本文所述的方法用于AAV生产。可用于产生AAV载体的示例性细胞系包括但不限于HEK293、CHO、Cos-7和NS0。

[0548] 在一个实施方案中,用于生产AAV载体的细胞系稳定地表达AAV载体生产所需的任何或所有组分,例如Rep、Cap、VP1等。在一个实施方案中,用于生产AAV载体的细胞系瞬时地表达AAV载体生产所需的任何或所有组分,例如Rep、Cap、VP1等。

[0549] 在用于生产AAV载体的细胞系不稳定或瞬时表达rep或cap的情况下,将这些序列提供给AAV表达系统。AAV rep和cap序列可以通过本领域已知的任何方法提供。目前的方案通常在单个质粒上表达AAV rep和/或cap基因。AAV复制和包装序列不需要一起提供,尽管这样做可能很方便。AAV rep和/或cap序列可以由任何病毒或非病毒载体提供。例如,rep和/或cap序列可以由杂交的腺病毒或疱疹病毒载体提供(例如,插入到删除的腺病毒载体的E1a或E3区域)。也可以采用EBV载体来表达AAV cap和rep基因。这种方法的一个优点是

EBV载体是游离体,但在连续的细胞分裂过程中会保持高拷贝数(即arc作为染色体外元件稳定地整合到细胞中,被称为“基于EBV的核游离体”,参见Margolski, Curr.Top.Microbial.Immun.158:67(1992))。

[0550] 通常情况下,AAV rep/cap序列两侧将不会有TR,以防止这些序列的救援和/或包装维持。

[0551] 使用本文所述的方法制造慢病毒的病毒表达系统将进一步包括核酸盒两侧的长末端重复(LTR)。LTR是相同的DNA序列,在逆转录转座子或由逆转录病毒RNA逆转录形成的前病毒DNA的任一端重复数百或数千次。LTR通过宿主染色体中的LTR特异性整合酶介导逆转录病毒DNA的整合。LTR和制造慢病毒载体的方法在例如美国专利号US7083981B2; US6207455B1; US6555107B2; US8349606B2; US7262049B2; 和美国专利申请号US20070025970A1; US20170067079A1; US20110028694A1中进一步描述;各专利的内容通过引用整体并入本文。

[0552] 使用本文所述的方法制造腺病毒的病毒表达系统将进一步包括核酸盒两侧的大约90-140个碱基对(确切长度取决于血清型)的相同的反向末端重复(ITR)。病毒的复制起源正好在基因组末端的ITR内。腺病毒基因组是大约36000个碱基对的线性双链DNA分子。通常,用于基因治疗的腺病毒载体在E1区具有缺失,在这里可以引入新的遗传信息;E1的缺失使重组病毒的复制有缺陷。ITR和制造腺病毒载体的方法在例如美国专利号US7510875B2; US7820440B2; US7749493B2; US7820440B2; US10041049B2; 国际专利申请号W02000070071A1; 以及美国专利申请号W02000070071A1; US20030022356A1; US20080050770A1中进一步描述,各专利的内容通过引用整体并入本文。

[0553] 在一个实施方案中,病毒表达系统可以是宿主细胞,例如病毒、哺乳动物细胞或昆虫细胞。示例性的昆虫细胞包括但不限于Sf9、Sf21、Hi-5和S2昆虫细胞系。例如,用于制造AAV载体的病毒表达系统可以进一步包括杆状病毒表达系统,例如,如果病毒表达系统是昆虫细胞。杆状病毒表达系统是为高效大规模病毒生产和表达来自杆状病毒感染的昆虫细胞的重组蛋白而设计的。杆状病毒表达系统在例如美国专利号US6919085B2; US6225060B1; US5194376A中进一步描述;各专利的内容通过引用整体并入本文。

[0554] 在另一个实施方案中,病毒表达系统是无细胞系统。用于病毒载体生产的无细胞系统在例如Cerqueira A., et al. Journal of Virology, 2016; Sheng J., et al. The Royal Society of Chemistry, 2017; 和Svitkin Y.V., and Sonenberg N. Journal of Virology, 2003中进一步描述;各专利的内容通过引用整体并入本文。

[0555] 在细胞中产生的病毒载体可以用任何标准技术释放(即从产生载体的细胞中释放出来)。例如,病毒载体可以通过机械方法如微流控、离心或超声,或化学方法如通过加入裂解缓冲液和去污剂释放。然后用本领域的标准方法回收(即收集)和纯化释放的病毒载体,以获得纯的群体。例如,病毒载体可以通过纯化方法从它们被释放的缓冲液中回收,包括使用深度过滤或切向流过滤(TFF)的澄清步骤。如本文的实施例所述,病毒载体可以通过超声从细胞中释放出来,并通过使用柱色谱对澄清的裂解液进行纯化而回收。

[0556] 在一些实施方案中,所述载体是非病毒载体,例如使用阳离子聚合物或阳离子脂质,如在本领域已知的。各种非病毒载体在Selene Ingusci等人(Gene Therapy Tools for Brain Diseases. Front. Pharmacol. 10:724. doi:10.3389)中进行了讨论。

[0557] 在另一个方面,提供了一种病毒体(病毒颗粒),其包括根据本发明的载体,适当地病毒载体。在一些实施方案中,所述病毒体是AAV病毒体。

[0558] 在另一个方面,提供了一种药物组合物,其包括根据本发明的合成的NS特异性启动子、表达盒、载体或病毒体。

[0559] 例如,AAV载体颗粒可被制备为药物组合物。可以理解的是,这种组合物必然包括一种或多种活性成分,并且最经常包括药学上可接受的赋形剂。

[0560] 根据本公开的药物组合物可以以散装、单一单位剂量和/或多个单一单位剂量的形式来制备、包装和/或销售。如本文所用,“单位剂量”是指包含预定量的活性成分的药物组合物的离散量。活性成分的量通常等于向受试者施用的活性成分的剂量和/或这种剂量的方便的分数,例如这种剂量的二分之一或三分之一。

[0561] 在另一个方面,提供了根据本发明的合成的NS特异性启动子、表达盒、载体、病毒体或药物组合物,其用作药物。

[0562] 在另一个方面,提供了根据本发明的合成的NS特异性启动子、表达盒、载体、病毒体或药物组合物,其用于治疗,即预防或治疗医学病况或疾病。

[0563] 在另一个方面,提供了本文所述的合成的NS特异性启动子、NS特异性顺式调节元件(CRE)、CRM、启动子元件、表达盒、载体、病毒体或药物组合物,其用于制造用于治疗医学病况或疾病的药物组合物。下文讨论了与本方面有关的示例性医学病况或疾病。

[0564] 适当地,所述医学病况或疾病与异常基因表达有关,任选地与在神经系统,适当地在CNS和/或PNS组织或细胞中的异常基因表达有关。适当地,所述用途是用于基因治疗,优选地用于治疗涉及异常基因表达的疾病。适当地,涉及异常基因表达的医学病况或疾病可以是CNS和/或PNS的疾病。适当地,所述医学病况或疾病可以是CNS和/或PNS的单基因病症。适当地,基因治疗涉及在CNS和/或PNS细胞或组织中表达治疗性表达产物。适当地,表达产物可以是缺陷基因的野生型等位基因、防止缺陷基因表达的表达产物或抵消缺陷基因的有害影响的表达产物。下文讨论了与本方面有关的示例性医学病况或疾病。

[0565] 在一个实施方案中,所述医学病况或疾病是亨廷顿氏病。

[0566] 在另一个方面,提供了一种包含本发明的合成的NS特异性启动子、表达盒、载体或病毒体的细胞。在一些实施方案中,所述细胞是哺乳动物细胞,任选地是人类细胞。适当地,所述细胞是CNS或PNS细胞。适当地,所述细胞可以是神经元、中间神经元或神经元支持细胞,例如星形胶质细胞、少突胶质细胞、室管膜细胞或小胶质细胞。适当地,所述细胞可以是人类神经元、星形胶质细胞、少突胶质细胞、室管膜细胞、小胶质细胞、许旺细胞或卫星细胞。合成的NS特异性启动子可以是游离体,或者可以在细胞的基因组中。

[0567] 在一个实施方案中,所述细胞是人类神经元、人类星形胶质细胞或人类少突胶质细胞。

[0568] 在另一个方面,提供了一种生产表达产物的方法,所述方法包括在NS细胞或组织中提供本发明的合成的NS特异性表达盒、载体或病毒体,并表达存在于合成的NS特异性表达盒、载体或病毒体中的目的基因。所述方法可以是体外的或离体的,也可以是体内的。

[0569] 在另一个方面,提供了一种在NS细胞中表达治疗性转基因的方法,所述方法包括将本文所述的合成的NS特异性表达盒、载体或病毒体引入NS细胞,并表达存在于合成的NS特异性表达盒、载体或病毒体中的表达产物(例如目的基因)。NS细胞可以是CNS(神经元、星

形胶质细胞、少突胶质细胞、室管膜细胞或小胶质细胞)或PNS(神经元、许旺细胞或卫星细胞)的细胞。NS细胞可以是神经元、中间神经元或神经元支持细胞,例如星形胶质细胞、少突胶质细胞、室管膜细胞、小胶质细胞、许旺细胞或卫星细胞。

[0570] 在一个实施方案中,CNS细胞是神经元、星形胶质细胞或少突胶质细胞。

[0571] 在另一个方面,提供了一种治疗有需要的受试者,优选人类的方法,所述方法包括:

[0572] -向受试者施用本文所述的表达盒、载体、病毒体或药物组合物,其中包括编码与根据本发明的启动子可操作地连接的治疗性产物的序列;和

[0573] -在所述受试者的NS中表达治疗量的治疗性产物。

[0574] 在一个实施方案中,治疗量的治疗性产物在CNS中表达。在一个实施方案中,治疗量的治疗性产物在PNS中表达。在一个实施方案中,治疗量的治疗性产物在CNS和PNS中表达。

[0575] 适当地,所述方法用于治疗、预防、缓解或改善神经性疾病和/或病症。下文了讨论与本方面有关的示例性医学病况或疾病。

[0576] 合适的施用方法可以是肠内(如口服、舌下和直肠)或肠外(如注射),包括静脉内、动脉内、脑室内、颅内、肌肉内、皮下、关节内、鞘内和皮内注射。优选的施用方法是静脉内、动脉内、脑室内、颅内和鞘内注射。

[0577] 在一些实施方案中,所述方法包括将本文所述的表达盒、载体、病毒体或药物组合物引入受试者的NS,其中包括编码治疗性产物的基因。在CNS中引入表达盒、载体、病毒体或药物组合物的一个特定困难是血脑屏障。血脑屏障是内皮细胞的半透性边界,它阻止血流中的某些化学物质和分子进入中枢神经系统的细胞外液。在动物研究中,这一障碍已经通过直接注射到动物的大脑中来克服,如颅内注射,适当地脑室内(ICV)注射(参见例如 Keiser et al.,Curr.Protoc.Mouse Biol.2018Dec;8(4):e57)。这种施用方法对于人类的基因治疗可能是不利的,因为它很难执行,并且对受试者来说可能是危险的。

[0578] 相反,在人类的基因治疗环境中,优选的是通过静脉内、鼻内或动脉内(如颈动脉内)施用包含本文所述的表达盒的病毒载体将表达盒引入NS中。适当地,病毒载体是AAV载体。一些血清型的AAV的静脉或动脉内施用可以使AAV载体渗透到NS中。由于根据本发明的合成的NS特异性启动子的NS特异性,预计在非NS组织和细胞中的表达量最小。此外,预计随着专门用于CNS渗透的改进型AAV衣壳的开发,AAV载体的渗透性将得到改善。静脉内或动脉内施用比颅内施用更安全,创伤更小,同时仍可渗透通过血脑屏障。

[0579] 在一些实施方案中,AAV载体被直接施用到CSF中。例如,AAV载体可以通过鞘内(IT)或脑室内(ICV)施用来施用。“鞘内”施用用于表示鞘内注射,即以脑脊髓轴的任何水平施用到脑脊液中,包括注射到脑室中。这是通过脊髓膜注射到蛛网膜下腔中,使注射物到达脑脊液(CSF)。因为注射到大脑或脊髓中需要复杂的脑部手术,所以鞘内递送方法被认为比注射到CNS组织本身中创伤更小。鞘内递送可以在不需要专门的大脑注射中心的情况下进行。

[0580] 在一个实施方案中,AAV载体是通过腰部施用来施用的。腰内鞘内施用优选在选自自由L4至L5、L3至L4、L1至L2和L2至L3脊髓节段的组的位置处。最优选地,AAV载体通过鞘内注射在L4和L5之间施用。在一个实施方案中,AAV载体在小脑延髓池中施用。在一个实施方案

中,AAV载体在小脑延髓池中和/或通过腰部施用来施用。在另一个实施方案中,施用到池(cisterna)中与腰部施用联合进行,由此施用到池中是在腰部施用之前、与之同时或之后进行的。在该实施方案中,腰部施用是在鞘内区域。“大池”或“小脑髓质小管”是蛛网膜下腔的三个主要开口之一,位于大脑周围的蛛网膜和脑膜的软膜之间。这些开口被统称为箱。池位于小脑和髓质的背侧表面之间。第四脑室中产生的脑脊液通过外侧和正中的端口排入小脑延髓池。

[0581] 在一些实施方案中,AAV载体可与血脑屏障破坏的手段联合施用。适当地,允许AAV载体更容易渗透血脑屏障。

[0582] 适当地,所述医学病况或疾病是CNS和/或PNS的医学病况或疾病。

[0583] 适当地,所述医学病况或疾病是PNS的医学病况或疾病。适当地,所述医学病况或疾病可以选自,例如:附神经病症、酒精性多神经病、痛性感觉缺失、抗MAG周围神经病、自身免疫性自主神经节病、自主神经反射障碍、自主神经病、腋下神经功能障碍、腋下神经麻痹、Charcot-Marie-Tooth病、化疗引起的周围神经病、慢性溶剂引起的脑病、CMV多发性神经根脊髓病、先天性痛觉不敏感伴无汗症、去神经支配、糖尿病性神经病、自主神经异常(dysautonomia)、红斑性肢痛症、面神经麻痹、家族性自主神经异常、Guillain-Barré综合征、遗传性感觉和自主神经病、Horner综合征、神经压迫综合征、神经损伤、神经失用症(neurapraxia)、神经炎、直立性低血压、直立性不耐受、阵发性交感神经亢进、周围单神经病、周围神经病、梨状肌综合征、神经丛病、多神经病、体位性直立性心动过速综合征、原发性自主神经衰竭、Pronator teres综合征、近端糖尿病性神经病、阴部神经卡压、单纯自主神经功能衰竭、四边孔综合征、桡神经功能障碍、桡神经病、辐射诱发的腰骶神经丛病变、神经根病、坐骨神经痛、小纤维周围神经病、胸廓出口综合征、尺骨神经病、血管炎性神经病、Villaret综合征、Wartenberg综合征和翼状肩胛。

[0584] 适当地,所述医学病况或疾病是CNS的医学病况或疾病,例如,神经性疾病和/或病症。适当地,所述医学病况或疾病可以选自,例如:多巴胺转运体缺乏综合征、注意力缺陷/多动障碍(ADHD)、双相障碍、癫痫、多发性硬化症、tau病、阿尔茨海默病、亨廷顿氏病、帕金森病、Krabbe病、肾上腺脑白质营养不良、运动神经元病(MND)、原发性侧索硬化症(PLS)、脊髓性肌肉萎缩症(SMA)、Kennedy病、大脑性麻痹、Gaucher病、Tay Sachs病、Rett综合征、Sandhoff病、Charcot-Marie-Tooth病、Angelman综合征、可传播性海绵状脑病(TSE)(包括Creutzfeldt-Jakob病(CJD)、库鲁病、致命性家族性失眠症(FFI)和Gerstmann-Straussler-Scheinker病(GSS))、晚期婴儿神经元蜡样脂褐质沉积症、粘多糖病IIIA、粘多糖病IIIB、脑白质营养不良(例如异染性脑白质营养不良、Krabbe病、肾上腺脑白质营养不良、Pelizaeus-Merzbacher病(PMD)、Canavan病、儿童共济失调伴中枢神经系统低髓鞘化或CACH、Alexander病、Refsum病和脑腱黄瘤病)、遗传性溶酶体储存疾病如Niemann-Pick病C1型或神经节苷脂贮积病,和/或神经元蜡样脂褐质沉积症,如Batten病、进行性核上性麻痹(PSP)、皮质基底节综合征,以及脑或脊髓肿瘤或癌症(包括星形细胞瘤和胶质母细胞瘤)或各种形式的CNS癌症,如原发性CNS淋巴瘤、转移性或继发性脑肿瘤、原发性脊髓肿瘤、家族性周期性瘫痪、Friedreich共济失调、毛细血管扩张共济失调、1型、2型和3型脊髓小脑共济失调、多巴反应性肌张力障碍脆性X染色体综合征、额颞叶痴呆、额颞叶痴呆伴运动神经元病(FTD-MND)、遗传性痉挛性截瘫(HSP)或家族性痉挛性截瘫(FSP)、脊髓损伤(SCI)、创伤性

脑损伤(TBI)、中风、脊髓梗死、精神疾病(严重抑郁症,强迫症)、巨轴索神经病(GAN)、疼痛(包括癌痛、舌咽神经痛(GN)和三叉神经痛(TN))、闭锁综合征、视神经脊髓炎(NMO)、橄榄体脑桥小脑萎缩(OPCA)、遗传性和非遗传性共济失调(如被称为Machado-Joseph病的遗传性脊髓小脑共济失调)和多系统萎缩(MSA)、Schilder病。在一个实施方案中,所述医学病况或疾病是亨廷顿氏病。

[0585] 适当地,编码治疗性产物的序列可以是选自由以下组成的组的基因之一: NPC1、EAAT2、NPY、CYP46A1、GLB1、APOE(或APOE2)、HEX、CLN1、CLN2、CLN3、CLN4、CLN5、CLN6、SUMF1、DCTN1、PRPH、SOD1、SMN、NEFH、GBA、IDUA、NAGLU、GUSB、ARSA、MANB、AADC、GDNF、NTN、ASP、MECP2、PTCHD1、GJB1、UBE3A、HEXA、FXN和MOG。在一些实施方案中,编码治疗性产物的序列可以是选自由以下组成的组的基因之一: 天冬氨酰氨基葡萄糖苷酶、 $\alpha$ -半乳糖苷酶A、棕榈酰蛋白硫酯酶、三肽基肽酶、溶酶体跨膜蛋白、半胱氨酸转运体、酸性神经酰胺酶、酸性 $\alpha$ -L-岩藻糖苷酶、保护蛋白/组织蛋白酶A、酸性 $\beta$ -葡萄糖苷酶或葡萄糖脑苷脂酶、酸性 $\beta$ -半乳糖苷酶、艾杜糖醛酸-2-硫酸酯酶、 $\alpha$ -L-艾杜糖醛酸酶、半乳糖脑苷酯酶、酸性 $\alpha$ -甘露糖苷酶、酸性 $\beta$ -甘露糖苷酶、芳基硫酸酯酶B、芳基硫酸酯酶A、N-乙酰半乳糖胺-6-硫酸盐硫酸酯酶、酸性 $\beta$ -半乳糖苷酶、N-乙酰氨基葡萄糖-1-磷酸转移酶、酸性鞘磷脂酶、NPC-1、酸性 $\alpha$ -葡萄糖苷酶、 $\beta$ -己糖胺酶B、乙酰肝素N-硫酸酯酶、 $\alpha$ -N-乙酰葡萄糖胺糖苷酶、乙酰CoA: $\alpha$ -氨基葡萄糖苷N-乙酰基转移酶、N-乙酰葡萄糖胺-6-硫酸盐硫酸酯酶、 $\alpha$ -N-乙酰半乳糖胺酶、 $\alpha$ -N-乙酰半乳糖胺酶、 $\alpha$ -神经酰胺酶、 $\beta$ -葡萄糖醛酸酶、 $\beta$ -己糖胺酶A、酸性脂肪酶、神经营养因子如神经生长因子(NGF)、神经营养素-3(NT-3)、神经营养素-4/5(NT-4/5)、脑源性神经营养因子(BDNF)、脑多巴胺神经营养因子(CDNF)、神经胶质源性的神经营养因子(GDNF)、睫状神经营养因子(CNTF)、胰岛素样生长因子(IGF-1)、亨廷顿基因(也称为HTT、HD和IT15)。

[0586] 在一个实施方案中,编码治疗性产物的序列是CYP46A1基因。

[0587] 本文所公开的示例性基因产物和用AAV载体(作为治疗或预防)治疗的CNS疾病的医学病况或疾病是: 用于治疗ALS的GDNF或IGF-1, 用于治疗SMA-1的GDNF或IGF, 用于下调不完整的脊髓运动神经元存活蛋白(SMN)基因以治疗SMA的miRNA, 等位基因特异性下调以治疗亨廷顿氏病或下调HTT的两个等位基因, 用于治疗亨廷顿氏病的GDNF, 用于治疗MSA的GDNF, 用于治疗Fabry病的 $\alpha$ -半乳糖苷酶A, 用于治疗小儿巴顿氏病的溶酶体跨膜蛋白(CNL3), 高雪氏病1型、2型和3型的酸性 $\beta$ -葡萄糖苷酶或葡萄糖脑苷脂酶, 用于治疗Hunter综合征的艾杜糖醛酸-2-硫酸酯酶, 用于治疗Pompe病的酸性 $\alpha$ -葡萄糖苷酶, 治疗Sanfilippo综合征A型选自由用于治疗San Filippo综合征B型的乙酰肝素N-硫酸酯酶A和N-乙酰葡萄糖胺糖苷酶组成的组。

[0588] 另外地或替代地, 治疗性产物可以是抗体、抗体片段或类似抗体的支架蛋白。

[0589] 另外地或替代地, 治疗性产物可以是针对疾病等位基因的基因编辑系统(如CRISPR-Cas9系统、TALEN、ZFN等)。

[0590] 另外地或替代地, 治疗性产物可以是一种或多种调节多核苷酸, 例如, 作为治疗剂的RNA或DNA分子。例如, 调节多核苷酸可以是miRNA或siRNA。靶基因可以是与任何神经性疾病有关的任何基因, 例如但不限于本文所列的那些。例如, siRNA双链体或编码的dsRNA可以减少或沉默NS细胞中靶基因的表达, 从而改善神经性疾病的症状。在一个非限制性实例中, 靶基因是亨廷顿(HTT)。在另一个非限制性实例中, 靶基因是微管相关蛋白tau(MAPT)。在一

个实施方案中,治疗性产物是靶向HTT基因的调节多核苷酸。

[0591] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:1或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0592] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:2或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0593] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:3或由其组成。适当地,表达产物在兴奋性神经元和星形胶质细胞中表达。

[0594] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:4或由其组成。适当地,表达产物在兴奋性神经元和星形胶质细胞中表达。

[0595] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:5或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0596] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:6或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0597] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:7或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0598] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:8或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0599] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:9或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0600] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:10或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0601] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操

作地连接的SEQ ID NO:11或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0602] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:12或由其组成。适当地,表达产物在神经元、少突胶质细胞和星形胶质细胞中表达。

[0603] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:13或由其组成。适当地,表达产物在兴奋性神经元、少突胶质细胞和星形胶质细胞中表达。

[0604] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:14或由其组成。适当地,表达产物在兴奋性神经元、少突胶质细胞和星形胶质细胞中表达。

[0605] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:33或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0606] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:34或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0607] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:35或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

[0608] 在另一个方面,提供了一种在NS中表达表达产物的方法,所述方法包括向NS细胞中引入表达盒,所述表达盒包括合成的NS特异性启动子,所述启动子包括与表达产物可操作地连接的SEQ ID NO:36或由其组成。适当地,表达产物在神经元和星形胶质细胞中表达。

## 附图说明

[0609] 图1A显示了HTT基因在成年小鼠大脑矢状切面上的表达模式(取自Allen小鼠脑图谱;mouse.brain-map.org)。HTT(亨廷顿)在整个大脑中高度表达。

[0610] 图1B显示了CYP46A1基因在成年小鼠大脑冠状切面上的表达模式(取自Allen小鼠脑图谱;mouse.brain-map.org)。CYP46A1在大脑中广泛表达。

[0611] 图2A显示了合成的NS特异性启动子SP0013、SP0014、SP0030、SP0031、SP0032、SP0019、SP0020、SP0021、SP0022、SP0011、SP0034、SP0035、SP0036和对照启动子Synapsin-1相对于对照启动子CAG在神经母细胞瘤来源的SH-SY5Y细胞中的GFP表达中值。NTC表示非感染的细胞。数据收集自三个生物重复,每个重复是两个技术重复的平均值。误差条为标准误差。

[0612] 图2B显示了用与GFP可操作地连接的合成的NS特异性启动子SP0013、SP0014、SP0030、SP0031、SP0032、SP0019、SP0020、SP0021、SP0022、SP0011、SP0034、SP0035、SP0036

或对照启动子Synapsin-1和CAG转染时,在神经母细胞瘤来源的SH-SY5Y细胞中的转染效率。NTC表示非转染的细胞。数据收集自三个生物重复,每个重复是两个技术重复的平均值。误差条为标准误差。GFP阳性%表示所有细胞中GFP阳性的百分比。

## 具体实施方式

### [0613] CRE及其功能变体

[0614] 本文公开了可用于构建NS特异性启动子的各种CRE。适当地,CRE是NS特异性的,适当地是CNS特异性的和/或PNS特异性的。这些CRE通常来源于基因组启动子和增强子序列,但它们在本文中用于与其天然基因组环境完全不同的环境中。一般来说,CRE构成了更大的基因组调节结构域的一小部分,这些调节结构域控制着与它们通常相关的基因的表达。令人惊讶地发现,这些CRE,其中许多非常小,可以从它们的正常环境中分离出来,并保留NS特异性调节活性。这是令人惊讶的,因为从基因组中复杂的和“三维”天然环境中移除调节序列通常会导致活性显著丧失,因此没有理由期望给定的CRE在从他们的天然环境中移除时保留曾经观察到的活性水平。当CRE在AAV载体中保留NS特异性活性时,就更令人惊讶了。这是特别的情况,因为AAV载体包括反向末端重复(ITR)且与基因组相比具有不同的DNA结构,并且已知ITR和DNA结构都影响CRE的活性。

[0615] 应该注意的是,可以改变本发明的CRE的序列而不会导致活性的显著丧失。CRE的功能变体可以通过修饰CRE的序列来制备,前提是避免对CRE活性明显有害的修饰。鉴于本公开中提供的信息,修饰CRE以提供功能变体是简单的。此外,本公开提供了用于简单地评估任何给定CRE变体的功能的方法。

[0616] 根据本发明的某些CRE的相对较小的尺寸是有益的,因为它允许在载体中提供CRE,更具体地,包含它们的启动子,同时占用最小量的载体有效载荷。当CRE用于容量有限的载体如基于AAV的载体时,这一点尤其重要。

[0617] 本发明的CRE包含某些NS特异性转录因子结合位点(TFBS)。通常希望在CRE的功能变体中,这些NS特异性TFBS保留功能。技术人员很清楚TFBS序列可以变化但仍保留功能。鉴于此,TFBS的序列通常由共有序列来说明,其中通常存在某种程度的变异。可以使用位置权重矩阵(PWM)来说明有关TFBS中发生的变异的更多信息,该矩阵表示给定核苷酸在共有序列中的给定位置处通常发现的频率。TF共有序列和相关位置权重矩阵的详细信息可以在例如Jaspar或Transfac数据库(<http://jaspar.genereg.net/>和<http://gene-regulation.com/pub/databases.html>)中找到。该信息允许技术人员以保留并且在一些情况下甚至增加CRE功能的方式修饰CRE的任何给定TFBS中的序列。鉴于此,技术人员对如何修饰任何给定TF的TFBS同时保留结合所需TF的能力有充分的指导;例如,Jaspar系统将根据其与给定PWM的相似性对推定的TFBS进行评分。此外,可以针对JASPAR数据库中的所有PWM扫描CRE,以鉴定/分析所有TFBS。技术人员当然可以在文献中找到另外的指导,此外,可以使用常规实验来确认TF与任何变体CRE中的推定TFBS的结合。很明显,即使在CRE中的TFBS内,CRE中的显著序列修饰也可以在保留功能的同时进行。本发明的CRE可以与范围广泛的合适的最小启动子或NS特异性近端启动子,适当地CNS特异性近端启动子和/或PNS特异性近端启动子组合使用。

[0618] CRE的功能变体包括与参考CRE元件不同的序列,但它们基本上保留了作为NS特异

性CRE的活性。技术人员可以理解,可以改变CRE的序列,同时保留其募集合适的NS特异性转录因子(TF)的能力,从而提高表达。CRE的功能变体可以包括与参考CRE相比的替换、缺失和/或插入,只要它们不会使CRE基本上无功能。

[0619] 在一些实施方案中,CRE的功能变体可以被看作是CRE,当它取代启动子中的参考CRE时,基本上保留其活性。例如,包含给定CRE的功能变体的NS特异性启动子优选保留其活性的至少80%,更优选保留其活性的至少90%,更优选保留其活性的至少95%,还更优选保留其活性的100%(与包含未经修饰的CRE的参考启动子相比)。

[0620] 适当地,CRE的功能变体与参考CRE保持显著水平的序列同一性。适当地,功能变体包括与参考CRE具有至少70%同一性的序列,更优选与参考CRE具有至少80%、90%、95%或99%同一性的序列。

[0621] 可以通过比较合适报告基因在参考启动子控制下的表达与包含替换的CRE的其他相同启动子在等同条件下的表达来评估活性的保留。用于评估NS特异性启动子活性的合适测定法在本文中公开,例如在实施例1中。

[0622] 在一些实施方案中,CRE可以与一个或多个额外的CRE结合以形成顺式调节模块(CRM)。额外的CRE可以在根据本发明的CRE的上游提供,或在根据本发明的CRE的下游提供。所述一个或多个额外的CRE可以是本文公开的CRE,或者可以是本领域已知的其他CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8,来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。适当地,额外的CRE是NS特异性的。

[0623] 根据本发明的CRE或包含根据本发明的CRE的CRM可以包括一个或多个额外的调节元件。例如,它们可以包括可诱导或可抑制元件、边界控制元件、绝缘体、基因座控制区、响应元件、结合位点、末端重复片段、响应位点、稳定元件、去稳定元件和剪接元件等,只要它们不会使CRE或CRM基本上无功能。

[0624] 包含根据本发明CRE的启动子可以包括CRM和最小或近端启动子之间和/或CRE之间的间隔区。另外地或替代性,在CRM的5'端可以存在间隔区。

[0625] 显然,根据本发明的CRE或包含根据本发明的CRE的CRM,或其功能变体,可以与任何合适的启动子元件结合,以提供根据本发明的合成的NS特异性启动子。适当地,启动子元件是NS特异性近端启动子。

[0626] 在许多情况下,较短的启动子序列是优选的,特别是在载体(例如病毒载体,如AAV)容量有限的情况下使用。因此,在一些实施方案中,根据本发明的CRE或其功能变体的长度为600个或更少的核苷酸,例如600、500、450、400、350、300、250、200、150、100、75、60、50个或更少的核苷酸。适当地,包含根据SEQ ID NO:19-24、27、28、37、38的CRE中的至少一个的合成的NS特异性CRM或其功能变体的长度为1000个或更少的核苷酸,例如950、900、850、800、750、700、650、600、550、500、450、400、350、300、250、200、150、100、75、60、50个或更少个核苷酸。

[0627] CRM及其功能变体

[0628] 本文公开了各种合成的NS特异性CRM,其可用于合成的NS特异性启动子的构建。本发明的CRM可以与范围广泛的合适的最小启动子或NS特异性近端启动子组合使用。

[0629] CRM的功能变体包括与参考CRM元件不同但基本上保留了作为NS特异性CRM的活性的序列。技术人员将理解,可以改变CRM的序列,同时保留其募集合适的NS特异性转录因子(TF)从而增强表达的能力。与参考CRM相比,CRM的功能变体可以包含替换、缺失和/或插入,只要它们不会使CRM基本上无功能。

[0630] 在一些实施方案中,CRM的功能变体可以被视为当在启动子中取代参考CRM时基本上保留其活性的CRM。例如,包含给定CRM的功能变体的NS特异性启动子优选保留其活性的至少80%,更优选保留其活性的至少90%,更优选保留其活性的至少95%,还更优选保留其活性的100%(与包含未经修饰的CRM的参考启动子相比)。

[0631] 适当地,CRM的功能变体与参考CRM保持显著水平的序列同一性。适当地,功能变体包括与参考CRM具有至少70%同一性的序列,更优选与参考CRM具有至少80%、90%、95%或99%同一性的序列。

[0632] 可以通过比较合适报告基因在参考启动子控制下的表达与包含替换的CRM的其他相同启动子在等同条件下的表达来评估活性的保留。用于评估NS特异性启动子活性的合适测定法在本文中公开,例如在实施例中的。

[0633] 在一些实施方案中,给定CRM的功能变体可以包含参考CRM中存在的一个或多个CRE的功能变体。例如,给定CRM的功能变体可以包含参考CRM中存在的1或2个CRE的功能变体。

[0634] 在一些实施方案中,给定CRM的功能变体可以包含与参考CRM相同的组合CRE,但是CRE可以以与参考CRM不同的顺序存在。通常优选的是,CRE以与参考CRM相同的顺序存在(因此,适当地CRM的功能变体包括与参考CRM中列出的CRE相同的排列)。

[0635] 在一些实施方案中,给定CRM的功能变体可以包含一个或多个额外的CRE,而不是参考CRM中存在的那些。可以在参考CRM中存在的CRE的上游、参考CRM中存在的CRE的下游和/或参考CRM中存在的CRE之间提供额外的CRE。所述一个或多个额外的CRE可以是本文公开的CRE,或者可以是本领域已知的其他CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8,来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。通常,优选给定CRM的功能变体包含相同的CRE(或其功能变体)并且不包含额外的CRE。

[0636] 与参考CRM相比,给定CRM的功能变体可以包含一个或多个额外的调节元件。例如,它们可以包括可诱导或可抑制元件、边界控制元件、绝缘体、基因座控制区、响应元件、结合位点、末端重复片段、响应位点、稳定元件、去稳定元件和剪接元件等,只要它们不会使CRM基本上无功能。

[0637] 给定CRM的功能变体可以包含相邻CRE之间的额外的间隔区,或者,如果参考CRM中存在一个或多个间隔区,则所述一个或多个间隔区可以比参考CRM中更长或更短。参考CRM中存在的间隔区可以在功能变体中被移除。

[0638] 显然,本文公开的CRM或其功能变体可以与任何合适的启动子元件组合以提供根据本发明的合成的NS特异性启动子。适当地,启动子元件是NS特异性近端启动子。

[0639] 在许多情况下,较短的启动子序列是优选的,特别是在载体(例如病毒载体,如AAV)容量有限的情况下使用。因此,在一些实施方案中,合成的NS特异性CRM的长度为500个

或更少的核苷酸,例如450、400、350、300、250、200、150、100、75、60、50个或更少的核苷酸。

[0640] 合成的NS特异性启动子及其功能变体

[0641] 本文公开了多种合成的NS特异性启动子。参考合成NS特异性启动子的功能变体是包含不同于参考合成NS特异性启动子但基本上保留NS特异性启动子活性的序列的启动子。技术人员将理解,可以改变合成的NS特异性启动子的序列,同时保留其募集合适的NS特异性转录因子(TF)和募集RNA聚合酶II以提供可操作地连接的序列(例如开放阅读框)的NS特异性表达的能力。与参考启动子相比,合成的NS特异性启动子的功能变体可以包括替换、缺失和/或插入,只要这种替换、缺失和/或插入不会使合成的NS特异性启动子与参考启动子相比基本上无功能。

[0642] 因此,在一些实施方案中,合成的NS特异性启动子的功能变体可以被视为基本上保留参考启动子的NS特异性启动子活性的变体。例如,合成的NS特异性启动子的功能变体优选保留参考启动子的活性的至少70%,更优选保留其活性的至少80%,更优选保留其活性的至少90%,更优选保留其活性的至少95%,还更优选保留其活性的100%。

[0643] 合成的NS特异性启动子的功能变体通常与参考合成NS特异性启动子保持显著水平的序列相似性。在一些实施方案中,功能变体包括与参考合成NS特异性启动子具有至少70%同一性的序列,更优选具有与参考合成NS特异性启动子具有至少80%、90%、95%或99%同一性的序列。

[0644] 功能变体中的活性可以通过比较合适报告基因在参考合成NS特异性启动子控制下的表达与推定的功能变体在等同条件下的表达来评估。用于评估NS特异性启动子活性的合适测定法在本文中公开,例如在实施例1中。

[0645] 给定的合成的NS特异性启动子的功能变体可以包括参考合成NS特异性启动子中存在的CRE的功能变体。给定的合成的NS特异性启动子的功能变体可以包括参考合成NS特异性启动子中存在的CRM的功能变体。给定的合成的NS特异性启动子的功能变体可以包括启动子元件的功能变体,或者当与参考合成NS特异性启动子相比时不同的启动子元件。

[0646] 给定的合成的NS特异性启动子的功能变体可以包括与参考合成NS特异性启动子中存在的一个或多个额外的CRE。例如,额外的CRE可以在参考合成NS特异性启动子中存在的CRE的上游或在参考合成NS特异性启动子中存在的CRE的下游提供。所述一个或多个额外的CRE可以是本文公开的CRE,或者可以是本领域已知的其他CRE。适当地,当所述一个或多个额外的CRE不是根据本发明的CRE时,所述一个或多个额外的CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8、来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。

[0647] 给定的合成的NS特异性启动子的功能变体可以包括相邻元件(CRE、CRM或启动子元件)之间的额外的间隔区,或者,如果参考合成NS特异性启动子中存在一个或多个间隔区,则所述一个或多个间隔区可以比参考合成NS特异性启动子中更长或更短。另外,如果在参考合成NS特异性启动子中存在一个或多个间隔区,则这些间隔区可以在功能变体中被移除。

[0648] 显然,本发明的合成的NS特异性启动子可以包括本发明的CRE或包含本发明的CRE和额外的调节序列的CRM。例如,它们可以包括一个或多个额外的CRE、可诱导或可抑制元件、边界控制元件、绝缘体、基因座控制区、响应元件、结合位点、末端重复片段、响应位点、

稳定元件、去稳定元件和剪接元件等,只要它们不会使启动子基本上无功能。

[0649] 本发明优选的合成的NS特异性启动子表现出NS特异性启动子活性,所述活性是Synapsin-1、Camk2a、GFAP、MBP、IBA1或NSE启动子在NS细胞中表现出的活性的至少15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、175%、200%、250%、300%、350%或400%。在许多情况下,更高水平的启动子活性是优选的,但这并不总是如此;因此,在一些情况下,更适度的表达水平可能是优选的。在一些情况下,希望有一系列不同活性水平的启动子,以使表达水平符合要求;本发明提供了具有这种活性范围的启动子。与已知的启动子相比,本发明的给定的合成的NS特异性启动子的活性可以通过比较报告基因在合成的NS特异性启动子控制下的NS特异性表达与同一报告基因在已知启动子控制下的表达来评估,当这两个启动子在其他等同的表达构建体中并在等同条件下提供时。

[0650] 除了不同的活性水平外,在一些情况下,希望有一系列在不同区域具有活性的启动子,例如NS或PNS的不同区域,适当地CNS,适当地大脑或脊髓的不同区域。另外,希望有一系列在不同区域具有不同活性水平的启动子,以使表达水平符合要求。在一些情况下,希望在特定区域中表达。在一些实施方案中,希望在大脑中表达,而在CNS的其他部位、PNS或身体的其他部位中很少或没有表达。在一些实施方案中,希望在脊髓中表达,而在CNS的其他部位、PNS或身体的其他部位中很少或没有表达。适当地,可能需要在大脑或脊髓内的多个区域表达。在一些优选实施方案中,根据本发明的NS特异性启动子在以下任何或所有大脑区域显示出活性:纹状体、黑质、大脑皮层的第3、5和6层(大脑皮层包括额叶、旁叶、枕叶和颞叶)、海马体、小脑、下丘脑和丘脑。在一些实施方案中,根据本发明的NS特异性启动子在上述大脑区域显示出活性,而在大脑的其他区域、CNS的其他区域、PNS或身体的其他区域很少或没有活性。

[0651] 除了不同的活性水平和不同的活性区域外,在一些情况下,希望有一系列在不同细胞或细胞组合中具有活性的启动子,例如CNS细胞和/或PNS细胞。在一些优选实施方案中,根据本发明的NS特异性启动子在PNS和/或PNS神经元中显示出活性。在一些优选实施方案中,根据本发明的NS特异性启动子在神经元和星形胶质细胞中显示出活性。在一些优选实施方案中,根据本发明的NS特异性启动子在神经元和少突胶质细胞中显示出活性。在一些优选实施方案中,根据本发明的NS特异性启动子在神经元、星形胶质细胞和少突胶质细胞中显示出活性。在一些实施方案中,根据本发明的NS特异性启动子在CNS神经元中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子在CNS和PNS神经元中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子在PNS神经元中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子在神经元和少突胶质细胞中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子在神经元和星形胶质细胞中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子在神经元、少突胶质细胞和星形胶质细胞中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。在一些实施方案中,根据本发明的NS特异性启动子

在神经元和卫星细胞中显示出活性,而在其他CNS或PNS细胞类型或CNS或PNS亚型中很少或没有表达。

[0652] 另外,可能优选在CNS和/或PNS的所有或几乎所有区域广泛表达。在一些实施方案中,希望在CNS的所有或几乎所有区域,适当地在大脑的所有区域广泛表达。例如,在治疗诸如亨廷顿氏病等疾病时,希望在整个CNS,适当地在整个大脑中表达,可能就是这种情况。

[0653] 在一个实施方案中,根据本发明的NS特异性启动子在大脑几乎所有区域的神经元、星形胶质细胞和/或少突胶质细胞中显示出活性,适当地在纹状体、黑质、大脑皮层的第3、5和6层(大脑皮层包括额叶、旁叶、枕叶和颞叶)、海马体、小脑、下丘脑和丘脑中。

[0654] 本发明优选的合成的NS特异性启动子表现出的神经元特异性启动子活性是Synapsin-1在神经元中表现出的活性的至少15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、175%、200%、250%、300%、350%或400%。与Synapsin-1相比,本发明的给定的合成的NS特异性启动子的活性可以通过比较报告基因在合成的NS特异性启动子控制下的神经元特异性表达与同一报告基因在Synapsin-1启动子控制下的在神经元中的表达来评估,当这两个启动子在其他等同的表达构建体中并在等同条件下提供时。在一些实施方案中,本发明的合成的NS特异性启动子能够使受试者的神经元中的基因(例如治疗性基因或目的基因)的表达相对于已知的神经元特异性启动子(适当地为Synapsin-1)增加至少20%、至少40%、至少60%、至少80%、至少100%、至少200%、至少300%、至少500%、至少1000%或以上。

[0655] 本发明优选的合成的NS特异性启动子表现出的星形胶质细胞特异性启动子活性是GFAP启动子在星形胶质细胞中表现出的活性的至少15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、175%、200%、250%、300%、350%或400%。与GFAP相比,本发明的给定的合成的NS特异性启动子的活性可以通过比较报告基因在合成的NS特异性启动子控制下的星形胶质细胞特异性表达与同一报告基因在GFAP启动子控制下的在星形胶质细胞中的表达来评估,当这两个启动子在其他等同的表达构建体中并在等同条件下提供时。在一些实施方案中,本发明的合成的NS特异性启动子能够使受试者的星形胶质细胞中的基因(例如治疗性基因或目的基因)的表达相对于已知的星形胶质细胞特异性启动子(适当地为GFAP)增加至少20%、至少40%、至少60%、至少80%、至少100%、至少200%、至少300%、至少500%、至少1000%或以上。

[0656] 本发明优选的合成的NS特异性启动子表现出的少突胶质细胞特异性启动子活性是MBP启动子在少突胶质细胞中表现出的活性的至少15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、175%、200%、250%、300%、350%或400%。与MBP相比,本发明的给定的合成的NS特异性启动子的活性可以通过比较报告基因在合成的NS特异性启动子控制下的少突胶质细胞特异性表达与同一报告基因在MBP启动子控制下的在星形胶质细胞中的表达来评估,当这两个启动子在其他等同的表达构建体中并在等同条件下提供时。在一些实施方案中,本发明的合成的NS特异性启动子能够使受试者的少突胶质细胞中的基因(例如治疗性基因或目的基因)的表达相对于已知的少突胶质细胞特异性启动子(适当地为MBP)增加至少20%、至少40%、至少60%、至少80%、至少100%、至少200%、至少300%、至少500%、至少1000%或以上。

[0657] 在一些实施方案中,本发明的合成的NS特异性启动子能够使受试者的CNS和/或

PNS中或CNS和/或PNS细胞中的基因(例如治疗性基因或目的基因)的表达相对于已知的NS特异性启动子(适当地为Syn1、Camk2a或NSE启动子)增加至少20%、至少40%、至少60%、至少80%、至少100%、至少200%、至少300%、至少500%、至少1000%或以上。

[0658] 本发明优选的合成的NS特异性启动子在非NS细胞(例如Huh7和HEK293细胞)中表现出的活性与CMV-IE相比为50%或更低,优选为CMV-IE的25%或更低,更优选为CMV-IE的10%或更低,在一些情况下为CMV-IE的5%或更低,或CMV-IE的1%或更低。

[0659] 在许多情况下,较短的启动子序列是优选的,特别是在载体(例如病毒载体,如AAV)容量有限的情况下使用。因此,在一些实施方案中,合成的NS特异性启动子的长度为1000个或更少的核苷酸,例如900、800、700、600、500、450、400、350、300、250、200、150、100个或更少的核苷酸。

[0660] 特别优选的合成的NS特异性启动子是那些既短又能表现出高水平活性的启动子。

[0661] 当NS特异性启动子在AAV载体中保留NS特异性活性时,是令人惊讶的,因为已知AAV载体的ITR和与基因组相比不同的DNA结构会影响启动子的活性,往往ITR和不同的DNA结构对启动子的活性有负面影响。

[0662] 其他CRE

[0663] 合成的NS特异性启动子可以包括一个或多个额外的CRE或其功能变体。这些CRE可以是根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)、其功能变体或本领域已知的其他CRE。

[0664] 根据本发明的CRE可以与一个或多个额外的CRE结合,以形成顺式调节模块(CRM)。适当地,所述一个或多个额外的CRE可以是根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)或其功能变体,或本领域已知的其他CRE。

[0665] 根据本发明的CRM可以包括一个或多个额外的CRE。适当地,所述一个或多个额外的CRE可以是根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)或其功能变体,或本领域已知的其他CRE。

[0666] 根据本发明的最小或近端启动子可以可操作地连接一个或多个CRE或一个CRM。所述一个或多个CRE可以是根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)或其功能变体,或本领域已知的其他CRE。所述CRM可以包括根据本发明的CRE(SEQ ID NO:19-24、27、28、37、38中的任何一个)或其功能变体,或本领域已知的其他CRE。

[0667] 优选地,当所述一个或多个CRE不是根据本发明的CRE(本文称为其他CRE)时,所述其他CRE是NS特异性CRE。优选地,当所述一个或多个CRE不是根据本发明的CRE时,所述一个或多个CRE可以选自:来自WO 2019/199867A1的SEQ ID NO:1-8、来自WO 2020/076614A1的SEQ ID NO:1-7和来自WO 2020/097121的SEQ ID NO:25-51、177-178、188。根据来自WO 2019/199867A1的SEQ ID NO:1-8的CRE是对抑制性GABA能中间神经元具有特异性的CRE。根据来自WO 2020/076614A1的SEQ ID NO:1-7的CRE是对前脑GABA能中间神经元具有特异性的CRE。根据来自WO 2020/097121的SEQ ID NO:25-34、36-51、177-178、188的CRE是对谷氨酸能神经元具有特异性的CRE。根据来自WO 2020/097121的SEQ ID NO:35的CRE是对谷氨酸能和GABA能神经元具有特异性的CRE。

[0668] 合成的NS特异性表达盒

[0669] 本发明还提供了一种合成的NS特异性表达盒,其包括本发明的合成的NS特异性启

动子,所述启动子可操作地连接至编码表达产物的序列,适当地为基因(例如转基因)。

[0670] 因此,适当地,表达产物可以是基因。在所述基因编码蛋白质的情况下,它基本上可以是任何类型的蛋白质。作为非限制性实例,所述蛋白质可以是酶、抗体或抗体片段(例如单克隆抗体)、病毒蛋白(例如REP-CAP、REV、VSV-G或RD114)、治疗性蛋白质或毒性蛋白(例如半胱天冬酶3、8或9)。

[0671] 在本发明的一些优选实施方案中,所述基因编码治疗性表达产物,优选治疗性多肽,适用于治疗与异常基因表达相关的疾病或病况,任选地在NS中,适当地在CNS和/或PNS中。

[0672] 在一些实施方案中,治疗性表达产物包括那些对治疗CNS和/或PNS疾病有用的产物。术语“CNS疾病”和“PNS疾病”原则上由技术人员理解。所述术语涉及可通过向CNS和/或PNS,特别是向CNS和/或PNS细胞施用活性化合物而治疗和/或预防的疾病。在一些实施方案中,CNS和/或PNS疾病是神经性疾病和/或病症。

[0673] 所述疾病可以是CNS和/或PNS的疾病。

[0674] 作为非限制性实例,CNS疾病可以选自:透明隔缺如、酸性脂肪酶病、酸性麦芽糖酶缺乏症、获得性癫痫样失语症、急性播散性脑脊髓炎、注意力缺陷-多动障碍(ADHD)、艾迪瞳孔、艾迪综合征、肾上腺脑白质营养不良、胼胝体发育不全、认识不能、Aicardi综合征、Aicardi-Goutieres综合征、艾滋病-神经系统并发症、亚历山大病、Alpers病、交替性偏瘫、阿尔茨海默病、肌萎缩侧索硬化症(ALS)、无脑症、动脉瘤、Angelman综合征、血管瘤病、缺氧症、抗磷脂综合征、失语症、失用症、蛛网膜囊肿、蛛网膜炎、Arnold-Chiari畸形、动静脉畸形、阿斯伯格综合征、共济失调、共济失调毛细血管扩张症、共济失调和小脑或脊髓小脑变性、心房颤动和中风、注意力缺陷-多动障碍、自闭症谱系障碍、自主神经功能障碍、背痛、Barth综合征、Batten病、Becker肌强直、Behcet病、贝尔氏面瘫、良性原发性眼睑痉挛、良性局灶性肌萎缩症、良性颅内高压、Bernhardt-Roth综合征、Binswanger病、眼睑痉挛、Bloch-Sulzberger综合征、分娩性臂丛神经损伤、臂丛神经损伤、Bradbury-Eggleston综合征、脑和脊髓肿瘤、脑动脉瘤、脑损伤、Brown-Sequard综合征、脊髓延髓性肌萎缩、常染色体显性脑动脉病伴皮质下梗死和白质脑病(CADASIL)、Canavan病、腕管综合征、灼痛、海绵状血管瘤(Cavernomas)、海绵状血管瘤(Cavernous Angioma)、海绵状血管畸形(Cavernous Malformation)、中央颈髓综合征、中央脊髓综合征、中枢性疼痛综合征、脑桥中央髓鞘溶解症、头部病症、神经酰胺酶缺乏症、小脑变性、小脑发育不良、脑动脉瘤、脑动脉硬化、脑萎缩、脑型脚气病、脑海绵状血管畸形、脑性巨人症、脑缺氧、大脑性麻痹、脑-眼-面-骨骼综合征(COFS)、Charcot-Marie-Tooth病、Chiari畸形、胆固醇脂沉积病、舞蹈症、舞蹈症-棘红细胞增多症、慢性炎症性脱髓鞘多神经病(CIDP)、慢性直立不耐症、慢性疼痛、Cockayne综合征II型、Coffin Lowry综合征、空洞脑、昏迷、复杂的局部疼痛综合征、先天性双侧面瘫、先天性肌无力、先天性肌病、先天性血管性海绵状畸形、皮质基底节变性、颅动脉炎、颅缝早闭、Cree脑炎、Creutzfeldt-Jakob病、累积性创伤病症、库欣综合征、巨细胞包涵体病、巨细胞病毒感染、舞蹈眼-舞蹈脚综合征、Dandy-Walker综合征、Dawson病、De Morsier's综合征、Dejerine-Klumpke瘫痪、痴呆症、多发性梗死性痴呆症、语义性痴呆症、皮层下痴呆症、路易体痴呆症、齿状小脑共济失调、齿状红核萎缩、皮炎、发展性运用障碍、Devic综合征、弥漫性硬化、Dravet综合征、书写障碍、阅读障碍、吞咽障碍、运动障碍、肌阵挛性小脑协调

障碍、进行性小脑协调障碍、肌张力异常、早期婴儿癫痫性脑病、空蝶鞍综合征、脑炎、流行性脑炎、脑膨出、脑病、脑病(家族性婴儿)、脑三叉神经血管瘤病、癫痫、癫痫性偏瘫、Erb瘫痪、Erb-Duchenne和Dejerine-Klumpke瘫痪、特发性震颤、脑桥外髓鞘溶解症、Fabry病、Fahr综合征、昏厥、家族性血管瘤、家族性特发性基底神经节钙化、家族性周期性瘫痪、家族性痉挛性瘫痪、Farber病、高热惊厥、肌纤维发育不良、Fisher综合征、Floppy婴儿综合征、足下垂、Friedreich共济失调、额颞叶痴呆、Gaucher病、普遍性神经节苷脂沉积症、Gerstmann综合征、Gerstmann-Straussler-Scheinker病、巨轴索神经病、巨细胞动脉炎、巨细胞包涵体病、球状细胞脑白质营养不良、舌咽神经痛、糖原贮积病、Hallervorden-Spatz病、头部损伤、头痛、连续性偏侧头痛、半侧面部痉挛、偏瘫Alterans、遗传性神经病、遗传性痉挛性截瘫、多神经炎型遗传性共济失调、带状疱疹、耳带状疱疹、Hirayama综合征、Holmes-Adie综合征、半叶前脑无裂畸形(Holoprosencephal)、HTLV-1相关脊髓病、Hughes综合征、亨廷顿氏病、积水性无脑、脑积水、脑积水-正常压力、脊髓积水(Hydromyelia)、皮质醇增多症、嗜睡、肌张力过高、肌张力过低、缺氧、免疫介导的脑脊髓炎、包涵体肌炎、色素失调症、婴儿肌张力过低、婴儿神经轴索性营养不良、婴儿植烷酸贮积病、婴儿Refsum病、婴儿痉挛症、炎症性肌病、枕骨裂脑露畸形(Iniencephaly)、肠道脂肪营养不良、颅内囊肿、颅内高压、Isaacs综合征、Joubert综合征、Kearns-Sayre综合征、Kennedy病、Kinsbourne综合征、Kleine-Levin综合征、Klippel-Feil综合征、Klippel-Trenaunay综合征(KTS)、Kliiver-Bucy综合征、Korsakoff失忆综合征、Krabbe病、Kugelberg-Welander病、库鲁病、Lambert-Eaton肌无力综合征、Landau-Kleffner综合征、股外侧皮神经卡压、延髓外侧综合征、学习障碍、Leigh病、Lennox-Gastaut综合征、Lesch-Nyhan综合征、脑白质营养不良、Levin-Critchley综合征、路易体痴呆、脂质储存疾病、类脂蛋白沉积症、无脑回畸形、闭锁综合征、Lou Gehrig病、狼疮-神经系统后遗症、Lyme病-神经系统并发症、Machado-Joseph病、巨脑症(Macrencephaly)、巨脑症(Megalencephaly)、Melkersson-Rosenthal综合征、脑膜炎、脑膜炎和脑炎、Menkes病、感觉异常性股痛、异染性脑白质营养不良、头小畸形、偏头痛、Miller Fisher综合征、小中风、线粒体肌病、Moebius综合征、单肢肌萎缩、运动神经元疾病、烟雾病、粘脂贮积病(Mucopolysaccharidoses)、粘多糖贮积病、多发性梗死性痴呆、多灶性运动神经病、多发性硬化、多系统萎缩、多系统萎缩伴直立性低血压、肌营养不良症、先天性肌无力、重症肌无力、髓鞘裂解性弥漫性硬化、婴儿肌阵挛性脑病、肌阵挛、肌病、先天性肌病、甲状腺毒性肌病、肌强直、先天性肌强直、嗜睡症、神经棘红细胞增多症、伴有脑铁积累的神经变性、神经纤维瘤病、神经阻滞剂恶性综合征、艾滋病的神经系统并发症、莱姆病的神经系统并发症、巨细胞病毒感染的神经系统后果、庞贝病的神经系统表现、狼疮的神经系统后遗症、视神经脊髓炎、神经性肌强直、神经元蜡样脂褐质沉积症、神经元迁移障碍、遗传性神经病、神经系统结节病(Neurosarcoidosis)、神经梅毒、神经毒性、海绵状痣、Niemann-Pick病、O'Sullivan-McLeod综合征、枕神经痛、Ohtahara综合征、橄榄体脑桥小脑萎缩、眼阵挛、直立性低血压、过度使用综合征、慢性疼痛、泛酸激酶相关的神经变性、副肿瘤综合征、感觉异常、帕金森病、阵发性舞蹈手足徐动症、阵发性偏头痛、Parry-Romberg、Pelizaeus-Merzbacher病、Pena Shokeir II综合征、围神经囊肿、周期性瘫痪、脑室周围白质软化、持续性植物状态、广泛性发育障碍、植烷酸贮积病、Pick病、神经挟捏、垂体瘤、多发性肌炎、庞贝病、孔洞脑、小儿麻痹症后期综合症、带状疱疹后遗神经痛、感染后脑脊髓炎、体位性低血

压、体位性直立性心动过速综合征、体位性心动过速综合征、原发性齿状萎缩症 (Primary Dentatum Atrophy)、原发性侧索硬化症、原发性进行性失语症、朊病毒病、进行性半侧面部萎缩症、进行性运动性共济失调、进行性多灶性白质脑病、进行性硬化性灰质萎缩、进行性核上性麻痹、脸盲症、Pseudo-Torch综合征、假性弓形虫病综合征、假性脑瘤、精神性运动、Ramsay Hunt综合征I、Ramsay Hunt综合征II、Rasmussen脑炎、反射性交感神经营养不良综合征、Refsum病、婴儿Refsum病、重复性运动障碍、重复性应激损伤、不宁腿综合征、逆转录病毒相关性骨髓病、Rett综合征、Reye综合征、风湿性脑炎、Riley-Day综合征、骶神经根囊肿、Saint Vitus舞蹈病、唾液腺疾病、Sandhoff病、Schilder病、脑裂畸形、Seitelberger病、癫痫、语义性痴呆、视(神经)中隔发育不良、婴儿严重肌阵挛性癫痫(SMEI)、摇晃婴儿综合征、带状疱疹、Shy-Drager综合征、Sjogren综合征、睡眠呼吸暂停、昏睡症、Sotos综合征、痉挛、脊柱裂、脊髓梗死、脊髓损伤、脊髓肿瘤、脊髓性肌萎缩、小脑萎缩、小脑退化、Steele-Richardson-Olszewski综合征、僵人综合征、纹状体黑质变性、中风、Sturge-Weber综合征、亚急性硬化性全脑炎、皮层下动脉硬化性脑病、短暂单侧神经痛样头痛、吞咽障碍、Sydenham舞蹈症、晕厥、梅毒性脊柱硬化症、脊髓空洞积水症(Syringohydromyelia)、脊髓空洞症、系统性红斑狼疮、脊髓痨、迟发性运动障碍、Tarlov囊肿、Tay-Sachs病、颞动脉炎、脊髓拴系综合征、Thomsen肌强直、胸廓出口综合征、甲状腺素毒性肌病、痛性抽搐(Tic Douloureux)、Todd瘫痪、Tourette综合征、短暂性脑缺血发作、传染性海绵状脑病、横贯性脊髓炎、创伤性脑损伤、震颤、三叉神经痛、热带痉挛性下肢轻瘫、Troyer综合征、结节性硬化症、血管性勃起肿瘤、中枢和周围神经系统血管炎综合征、Von Economo病、Von Hippel-Lindau病(VHL)、Von Recklinghausen病、Wallenberg综合征、Werdnig-Hoffman病、Wernicke-Korsakoff综合征、West综合征、Whiplash、Whipple病、Williams综合征、Wilson病、Wolman病、X连锁脊柱和球状肌肉萎缩症。

[0675] 作为非限制性实例,PNS疾病可以选自:附神经病症、酒精性多神经病、痛性感觉缺失、抗MAG周围神经病、自身免疫性自主神经节病、自主神经反射障碍、自主神经病、腋下神经功能障碍、腋下神经麻痹、Charcot-Marie-Tooth病、化疗引起的周围神经病、慢性溶剂引起的脑病、CMV多发性神经根脊髓病、先天性痛觉不敏感伴无汗症、去神经支配、糖尿病性神经病、自主神经异常(dysautonomia)、红斑性肢痛症、面神经麻痹、家族性自主神经异常、Guillain-Barre综合征、遗传性感觉和自主神经病、Horner综合征、神经压迫综合征、神经损伤、神经失用症(neurapraxia)、神经炎、直立性低血压、直立性不耐受、阵发性交感神经亢进、周围单神经病、周围神经病、梨状肌综合征、神经丛病、多神经病、体位性直立性心动过速综合征、原发性自主神经衰竭、Pronator teres综合征、近端糖尿病性神经病、阴部神经卡压、单纯自主神经功能衰竭、四边孔综合征、桡神经功能障碍、桡神经病、辐射诱发的腰骶神经丛病变、神经根病、坐骨神经痛、小纤维周围神经病、胸廓出口综合征、尺骨神经病、血管炎性神经病、Villaret综合征、Wartenberg综合征和翼状肩胛。

[0676] 一些疾病可以是CNS和PNS两者的疾病。

[0677] 在一些实施方案中,CNS疾病选自由以下组成的列表:多巴胺转运体缺乏综合征、注意力缺陷/多动障碍(ADHD)、双相障碍、癫痫、多发性硬化症、tau病、阿尔茨海默病、亨廷顿氏病、帕金森病、Krabbe病、肾上腺脑白质营养不良、运动神经元病、大脑性麻痹、Batten病、Gaucher病、Tay Sachs病、Rett综合征、Sandhoff病、Charcot-Marie-Tooth病、Angelman

综合征、Canavan病、晚期婴儿神经元蜡样脂褐质沉积症、粘多糖病ⅢA、粘多糖病ⅢB、异染性脑白质营养不良、遗传性溶酶体储存疾病如Niemann-Pick病C1型,和/或神经元蜡样脂褐质沉积症,如Batten病、进行性核上性麻痹、皮质基底节综合征,以及脑癌(包括星形细胞瘤和胶质母细胞瘤)。

[0678] 在一个实施方案中,CNS疾病是亨廷顿氏病。

[0679] 本领域已经描述了适合治疗上述病况的各种表达产物。适当地,编码与根据本发明的合成的NS特异性启动子可操作地连接的表达产物的序列可以是选自以下组成的组的基因之一: NPC1、EAAT2、NPY、CYP46A1、GLB1、APOE(例如, ApoE2、ApoE3或ApoE4)、HEX、CLN1、CLN2、CLN3、CLN4、CLN5、CLN6、SUMF1、DCTN1、PRPH、SOD1、NEFH、GBA、IDUA、NAGLU、GUSB、ARSA、MANB、AADC、GDNF、NTN、ASP、MECP2、PTCHD1、GJB1、UBE3A、HEXA、MOG。另外地或替代地,与根据本发明的合成的NS特异性启动子可操作地连接的表达产物可以是针对疾病等位基因的miRNA/CRISPR Cas9。适当地,编码与根据本发明的合成的NS特异性启动子可操作地连接的表达产物的序列可以是基因CYP46A1。

[0680] CYP46A1是胆固醇降解的限速酶,已经发现它在多种CNS疾病中发挥有益作用。CYP46A1的抑制可能有助于通过增加病毒胆固醇的数量来诱发和/或加重阿尔茨海默病,如(Djel'ti等人,2015)中所述,该文献通过引用并入本文。还发现CYP46A1在亨廷顿氏病中具有神经保护作用,如(Boussicault等人,2016)中所述,该文献通过引用并入本文。因此,CYP46A1基因是特别优选的编码表达产物的序列。在一些优选实施方案中,CYP46A1基因与根据本发明的合成的NS特异性启动子可操作地连接。适当地,CYP46A1基因与在CNS的所有区域(泛CNS)都有活性的合成的启动子或在上述大脑的5、6、7、8或9个以上的区域有活性的启动子可操作地连接。CYP46A1在CNS的所有区域或上述大脑的5、6、7、8或9个以上区域的表达可能是有益的,因为发现由泛启动子CMV或CAG进行的CYP46A1表达对小鼠亨廷顿氏病模型是有益的(Kacher等人,2019)。适当地,CYP46A1基因与由SEQ ID NO:1-14、33-36中的任何一个组成或包含SEQ ID NO:1-14、33-36中的任何一个的合成启动子可操作地连接。

[0681] 在一些实施方案中,有用的表达产物包括肌营养不良蛋白(包括微肌营养不良蛋白)、 $\beta$ 1,4-正乙酰半乳糖胺半乳糖基转移酶(GALGT2)、氨基甲酰合成酶I、 $\alpha$ -1抗胰蛋白酶、鸟氨酸转氨甲酰酶、精氨酸琥珀酸合成酶、精氨酸琥珀酸裂解酶、精氨酸酶、富马酸乙酰乙酸水解酶、苯丙氨酸羟化酶、葡萄糖-6-磷酸酶、胆色素原脱氨酶、胱硫醚 $\beta$ -合成酶、支链酮酸脱羧酶、白蛋白、异戊酰-CoA脱氢酶、丙酰CoA羧化酶、甲基丙二酰CoA变位酶、戊二酰CoA脱氢酶、胰岛素、 $\beta$ -葡萄糖苷酶、丙酮酸羧酸盐、肝脏磷酸化酶、磷酸化酶激酶、甘氨酸脱羧酶、H蛋白、T蛋白和囊性纤维化跨膜调节蛋白(CFTR)。

[0682] 其他有用的表达产物包括在酶替代疗法中有用的酶,这些酶对因酶的活性不足而导致的各种病况都有用。例如,含有甘露糖-6-磷酸的酶可用于溶酶体储存疾病的治疗(例如,合适的基因包括编码 $\beta$ -葡萄糖醛酸酶(GUSB)的基因)。

[0683] 在一些实施方案中,示例性的多肽表达产物包括神经保护性多肽和抗血管生成多肽。合适的多肽包括但不限于神经胶质源性的神经营养因子(GDNF)、成纤维细胞生长因子2(FGF-2)、nurturin、睫状神经营养因子(CNTF)、神经生长因子(NGF;例如神经生长因子- $\beta$ )、脑源性神经营养因子(BDNF)、神经营养素-3(NT-3)、神经营养素-4(NT-4)、神经营养素-6(NT-6)、表皮生长因子(EGF)、色素上皮衍生因子(PEDF)、Wnt多肽、可溶性Fit-1、血管抑素、

内皮抑素、VEGF、抗VEGF抗体、可溶性VEGFR、因子VIII (FVIII)、因子IX (FIX) 和刺猬家族的成员(声波刺猬、印度刺猬和沙漠刺猬等)。

[0684] 在一些实施方案中,有用的治疗性表达产物包括激素和生长及分化因子,包括但不限于胰岛素、胰高血糖素、生长激素(GH)、甲状旁腺素(PTH)、生长激素释放因子(GRF)、卵泡刺激素(FSH)、黄体生成素(LH)、人绒毛膜促性腺激素(hCG)、血管内皮生长因子(VEGF)、血管生成素、血管抑素、粒细胞集落刺激因子(GCSF)、促红细胞生成素(EPO)、结缔组织生长因子(CTGF)、碱性成纤维细胞生长因子(bFGF)、酸性成纤维细胞生长因子(aFGF)、表皮生长因子(EGF)、血小板衍生生长因子(PDGF)、胰岛素生长因子I和II(IGF-I和IGF-II)、转化生长因子 $\alpha$ 超家族中的任何一个(包括TGF $\alpha$ )、激活素、抑制素或骨形态发生蛋白(BMP) BMP 1-15中的任何一个、生长因子的heregulin/神经调节蛋白/ARIA/neu分化因子(NDF)家族中的任何一个、神经生长因子(NGF)、脑源性神经营养因子(BDNF)、神经营养素NT-3和NT-4/5、睫状神经营养因子(CNTF)、神经胶质细胞系源性的神经营养因子(GDNF)、神经秩蛋白(neurturin)、agrin、semaphorins/collapsins家族中的任何一个、netrin-1和netrin-2、肝细胞生长因子(HGF)、ephrins、noggin、声波刺猬和酪氨酸羟化酶。

[0685] 在一些实施方案中,有用的表达产物包括调节免疫系统的蛋白质,包括但不限于细胞因子和淋巴因子,如促血小板生成素(TPO)、白细胞介素(IL) IL-1至IL-25(包括IL-2、IL-4、IL-12和IL-18)、单核细胞趋化蛋白、白血病抑制因子、粒细胞-巨噬细胞集落刺激因子、Fas配体、肿瘤坏死因子 $\alpha$ 和 $\beta$ 、干扰素( $\alpha$ 、 $\beta$ 和 $\gamma$ )、干细胞因子、flk-2/flt3配体。由免疫系统产生的基因产物在本发明中也是有用的。这些包括但不限于免疫球蛋白IgG、IgM、IgA、IgD和IgE、嵌合免疫球蛋白、人源化抗体、单链抗体、T细胞受体、嵌合T细胞受体、单链T细胞受体、I类和II类MHC分子以及工程化的免疫球蛋白和MHC分子。有用的基因产物还包括补体调节蛋白,如补体调节蛋白、膜辅助因子蛋白(MCP)、衰变加速因子(DAF)、CR1、CF2和CD59。

[0686] 在一些实施方案中,有用的表达产物包括激素、生长因子、细胞因子、淋巴因子、调节蛋白和免疫系统蛋白的受体中的任何一种。有用的异源核酸序列还包括胆固醇调节和/或脂质调节的受体,包括低密度脂蛋白(LDL)受体、高密度脂蛋白(HDL)受体、极低密度脂蛋白(VLDL)受体和清道夫受体。本发明还包括基因产物的使用,如类固醇激素受体超家族成员,包括糖皮质激素受体和雌激素受体、维生素D受体和其他核受体。此外,有用的基因产物包括转录因子,如jun、fos、max、mad、血清反应因子(SRF)、AP-1、AP-2、myb、MyoD和肌细胞生成素(myogenin)、含有ETS盒的蛋白、TFE3、E2F、ATF1、ATF2、ATF3、ATF4、ZF5、NFAT、CREB、HNF-4、C/EBP、SP1、CCAAT盒结合蛋白、干扰素调节因子(IRF-1)、维尔姆斯肿瘤蛋白、ETS结合蛋白、STAT、GATA盒结合蛋白如GATA-3以及叉头家族的翼状螺旋蛋白。

[0687] 在一些实施方案中,有用的表达产物包括非天然存在的多肽,例如具有含有插入、缺失或氨基酸替换的非天然存在的氨基酸序列的嵌合或杂交多肽。

[0688] 其他合适的表达产物包括微RNA(miRNA)、干扰RNA、反义RNA、核酶和适体。

[0689] 在本发明的一些实施方案中,合成的NS特异性表达盒包括用于基因编辑的基因,例如编码位点特异性核酸酶如大范围核酸酶、锌指核酸酶(ZFN)、基于转录激活因子样效应物的核酸酶(TALEN)或规律间隔成簇短回文重复系统(CRISPR-Cas)的基因。适当地,位点特异性核酸酶适于通过制造切口(通常是位点特异性双链断裂)来编辑期望的靶基因组位点,然后通过非同源末端连接(NHEJ)或同源依赖修复(HDR)进行修复,从而得到期望的编辑。所

述编辑可以是有功能障碍的基因的部分或完全修复,或者是功能正常的基因的敲低或敲除。替代地,所述编辑可以使用本领域已知的合适系统通过碱基编辑或先导编辑进行。适当地,对于亨廷顿氏病的治疗,表达产物可以抑制患者细胞中有缺陷的亨廷顿基因和/或可以抵消或减轻患者细胞中有缺陷的亨廷顿基因的有害影响。

[0690] 适当地,合成的NS特异性表达盒包括提供或编码核糖体结合位点、起始密码子、终止密码子和转录终止序列中的一个或多个(优选全部)的序列。适当地,所述表达盒包括编码转录后调节元件的核酸。适当地,所述表达盒包括编码polyA元件的核酸。

[0691] 载体和病毒颗粒

[0692] 本发明进一步提供了包含根据本发明的合成的NS特异性启动子或表达盒的载体。

[0693] 在本发明的一些实施方案中,所述载体是质粒。这种质粒可以包括多种其他功能性核酸序列,例如一种或多种可选择标志物、一种或多种复制起点、多克隆位点等。在本发明的一些实施方案中,所述载体是病毒载体。

[0694] 在本发明的一些实施方案中,所述载体是用于在真核细胞中表达的表达载体。真核表达载体的实例包括但不限于可从Stratagene获得的pW-LNE0、pSV2CAT、pOG44、pXT1和pSG;可从Amersham Pharmacia Biotech获得的pSVK3、pBPV、pMSG和pSVL;和可从Clontech获得的pCMVDsRed2-express、pIRES2-DsRed2、pDsRed2-Mito、pCMV-EGFP。许多其他载体是众所周知的并且是可商购的。对于哺乳动物细胞腺病毒载体,pSV和pCMV系列载体是特别众所周知的非限制性实例。有许多众所周知的酵母表达载体,包括但不限于酵母整合质粒(YIp)和酵母复制质粒(YRp)。对于植物,农杆菌的Ti质粒是示例性的表达载体,植物病毒也提供合适的表达载体,例如烟草花叶病毒(TMV)、马铃薯病毒X和豇豆花叶病毒。

[0695] 在一些优选实施方案中,所述载体是基因治疗载体。本领域已知各种基因治疗载体,可以提及AAV载体、腺病毒载体、逆转录病毒载体和慢病毒载体。当所述载体是基因治疗载体时,所述载体优选包含与本发明的合成的NS特异性启动子可操作地连接的核酸序列,所述核酸序列编码治疗性产物,适当地是治疗性蛋白质。治疗性蛋白质可以是可分泌的蛋白质。上文讨论了可分泌的蛋白质的非限制性实例,示例性的可分泌的治疗性蛋白质包括凝血因子,例如因子VIII或因子IX、胰岛素、促红细胞生成素、脂蛋白脂肪酶、抗体或纳米抗体、生长因子、细胞因子、趋化因子、血浆因子、毒性蛋白等。

[0696] 在本发明的一些实施方案中,所述载体是病毒载体,如逆转录病毒、慢病毒、腺病毒、单纯疱疹或腺相关病毒(AAV)载体。在一些优选实施方案中,所述载体是慢病毒载体,适当地是基于HIV-1的慢病毒载体。在一些优选实施方案中,所述载体是AAV载体。在一些优选实施方案中,AAV具有适合或专门优化用于CNS转导的血清型。为了转导PNS的细胞,不需要克服血脑屏障,所以PNS细胞的转导可以由任何合适的AAV血清型完成。在一些实施方案中,AAV选自由以下组成的组:AAV1、AAV2、AAV4、AAV5、AAV8、AAV9、AAVrh10、AAVDJ8、AAV2i8、AAV2.5、AAV2i8G9和AAV2g9,或其衍生物。

[0697] AAV载体优选用作自互补的双链AAV载体(scAAV),以克服AAV转导中的一个限制性步骤(即单链到双链AAV的转换),尽管单链AAV载体(ssAAV)的使用也包括在本文中。在本发明的一些实施方案中,AAV载体是嵌合的,即它包括至少两个AAV血清型的组分,如AAV2的ITR和AAV5的衣壳蛋白。已知AAV9能特别有效地转导CNS细胞和组织,因此AAV9及其衍生物对靶向CNS细胞和组织特别有意义。已知AAV2g9能特别有效地转导CNS细胞和组织,因此

AAV2g9及其衍生物对靶向CNS细胞和组织特别有意义。已知AAVrh10能特别有效地转导CNS细胞和组织,因此AAVrh10及其衍生物对靶向CNS细胞和组织特别有意义。AAVrh10是特别优选的,因为AAVrh10的全身或静脉内递送已被发现在中枢神经系统中提供高的转基因表达,如(Tanguy等人,2015)中所述,该文献通过引用并入本文。已知AAVDJ8能特别有效地转导CNS细胞和组织,因此AAVDJ8及其衍生物对靶向CNS细胞和组织特别有意义。AAVDJ8是优选的,因为它已被证明能有效地靶向大脑的多个区域,并能有效地靶向星形胶质细胞,如(Hammond等人,2017)中所述,该文献通过引用并入本文。还已知AAV1、AAV2、AAV4、AAV5、AAV2i8、AAV2.5、AAV2i8G9和AAV8可靶向CNS细胞和组织,因此这些AAV血清型及其衍生物对靶向CNS细胞和组织也特别有意义。

[0698] 本发明进一步提供了包含上述载体的重组病毒体(病毒颗粒)。

[0699] 药物组合物

[0700] 本发明的载体或病毒体可以与药学上可接受的赋形剂,即一种或多种药学上可接受的载体物质和/或添加剂,例如缓冲剂、载体、赋形剂、稳定剂等一起配制在药物组合物中。可以以试剂盒的形式提供所述药物组合物。药物组合物和适用于AAV载体的递送系统和方法和本领域已知。

[0701] 因此,本发明的另一方面提供了一种包含如本文所述的载体或病毒体的药物组合物。

[0702] 根据本公开的药物组合物中的活性成分(例如AAV载体颗粒)、药学上可接受的赋形剂和/或任何额外成分的相对量可以不同,取决于接受治疗的受试者的身份、大小和/或状况,并进一步取决于组合物的施用途径。例如,所述组合物可以包括0.1%至99%(w/w)的活性成分。举例来说,所述组合物可以包括0.1%至100%,例如,0.5%至50%、1-30%、5-80%、至少80%(w/w)的活性成分。

[0703] 药物组合物可以使用一种或多种赋形剂或稀释剂来配制,以(1)增加稳定性;(2)增加细胞转染或转导;(3)允许有效载荷的持续或延迟释放;(4)改变生物分布(例如,将病毒颗粒靶向至特定组织或细胞类型);(5)增加编码的蛋白质的翻译;(6)改变编码的蛋白质的释放曲线和/或(7)允许本发明的有效载荷的可调节表达。在一些实施方案中,药学上可接受的赋形剂可以是至少95%、至少96%、至少97%、至少98%、至少99%或100%的纯度。在一些实施方案中,赋形剂被批准用于人类和用于兽医用途。在一些实施方案中,赋形剂可由美国食品和药物管理局批准。在一些实施方案中,赋形剂可以是医药级的。在一些实施方案中,赋形剂可以满足美国药典(USP)、欧洲药典(EP)、英国药典和/或国际药典的标准。本文所用的赋形剂包括但不限于任何和所有溶剂、分散介质、稀释剂或其他液体载体、分散或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂等,以适合于所需的特定剂型。用于配制药物组合物的各种赋形剂和用于制备组合物的技术在本领域是已知的(参见Remington:The Science and Practice of Pharmacy,21st Edition,A.R.Gennaro,Lippincott,Williams and Wilkins,Baltimore,MD,2006;通过引用整体并入本文)。常规赋形剂介质的使用可以在本公开的范围考虑,除非任何常规赋形剂介质可能与物质或其衍生物不相容,例如产生任何不希望的生物效应或以与药物组合物的任何其他组分有害的方式相互作用。

[0704] 治疗性和其他方法及用途

[0705] 本发明还提供了根据本发明各方面的合成的NS特异性启动子、表达盒、载体、病毒体或药物组合物,其用于治疗疾病,优选与异常基因表达相关的疾病,任选地在NS中(例如遗传性NS疾病)。上文讨论了相关的病状、疾病和治疗性表达产物。

[0706] 本发明还提供了根据本发明各方面的合成的NS特异性启动子、表达盒、载体、病毒体,其用作药物。

[0707] 本发明还提供了根据本发明各方面的合成的NS特异性启动子、表达盒、载体、病毒体,其用于制造用于治疗本文所述的任何病况或疾病的药物组合物。

[0708] 本发明进一步提供了一种包含根据本发明各方面的合成的NS特异性启动子、表达盒、载体、病毒体的细胞。适当地,所述细胞是真核细胞。所述真核细胞可以适当地是动物(后生动物)细胞(例如哺乳动物细胞)。适当地,所述细胞是人类细胞。

[0709] 在本发明的一些实施方案中,所述细胞是离体的,例如在细胞培养中。在本发明的其他实施方案中,所述细胞可以是组织或多细胞生物体的一部分。

[0710] 在优选实施方案中,所述细胞是NS细胞,它可以是离体的或体内的。NS细胞可以是CNS细胞或PNS细胞。PNS细胞可以是初级神经元、许旺细胞或卫星细胞。CNS细胞可以是初级神经元、星形胶质细胞、少突胶质细胞、小胶质细胞或室管膜细胞。替代地,NS细胞可以是NS来源的细胞系,例如永生化细胞系。在一个实施方案中,CNS细胞是神经元、星形胶质细胞或少突胶质细胞。

[0711] 所述细胞可以存在于NS组织环境中(例如在动物的NS中),或者可以从NS组织中分离出来,例如,它可以在细胞培养中。适当地,原代细胞或细胞系是人类细胞。

[0712] 根据本发明的合成的NS特异性启动子、表达盒或载体可以插入到细胞的基因组中,或者它可以是游离体(例如存在于游离载体中)。

[0713] 在另一个方面,本发明提供了一种生产表达产物的方法,所述方法包括在细胞,优选NS细胞中提供根据本发明的合成的NS特异性表达盒(优选在如上所述的载体中),并表达存在于合成的NS特异性表达盒中的基因。所述方法适当地包括在合适的条件下维持所述NS细胞以表达所述基因。在培养中,这可能包括在合适的培养条件下孵育所述细胞或包含所述细胞的组织。当然,表达可以在体内进行,例如在受试者的NS的一个或多个细胞中。

[0714] 适当地,所述方法包括将合成的NS特异性表达盒引入NS细胞的步骤。转染NS细胞的范围广泛的方法在本领域中是众所周知的。优选的转染NS细胞的方法是用包含合成的NS特异性表达盒的病毒载体(例如AAV载体)转导细胞。

[0715] 对技术人员来说,显而易见的是,根据本发明各方面的合成的NS特异性启动子、表达盒、载体或病毒体可用于基因治疗。因此,这种核酸构建体在基因治疗中的使用构成了本发明的一部分。

[0716] 因此,本发明在一些实施方案中提供了根据本发明的表达盒、载体或病毒体,其用于受试者的基因治疗,优选通过治疗性基因的NS特异性表达进行的基因治疗。所述治疗可涉及通过从NS细胞分泌治疗性产物来治疗疾病,适当地是涉及NS中异常基因表达的疾病,如上所述。

[0717] 本发明还提供了一种在NS细胞中表达治疗性转基因的方法,所述方法包括将根据本发明的表达盒或载体引入NS细胞。NS细胞可以是体内的或离体的。

[0718] 本发明还提供了一种对有需要的受试者(优选人类)进行基因治疗的方法,所述方

法包括：

[0719] 向受试者施用(适当地引入受试者的NS)本发明的合成的NS特异性表达盒、载体、病毒体或药物组合物,其包括编码治疗性产物的基因。

[0720] 所述方法适当地包括在所述受试者的NS中表达治疗量的所述基因的治疗性产物。上文讨论了可以治疗的各种病况和疾病。编码合适的治疗性产物的基因已在上文进行了讨论。

[0721] 所述方法适当地包括向受试者施用根据本发明的载体或病毒体。适当地,所述载体是病毒基因治疗载体,例如AAV载体。

[0722] 在一些实施方案中,所述方法包括全身施用基因治疗载体。全身施用可以是肠内(例如口服、舌下和直肠)或肠外(例如注射)的。优选的注射途径包括静脉内、肌肉内、皮下、动脉内、关节内、鞘内和皮内注射。在一个实施方案中,基因治疗载体可以通过注射到CSF通路来递送。向CSF通路递送的非限制性实例包括鞘内(脊柱内注射、腰部注射或注射到蛛网膜下腔)和脑室内施用。在一些实施方案中,基因治疗载体可通过鼻内递送来递送。

[0723] 包含根据本发明的合成的NS特异性启动子或表达盒的AAV载体或病毒体的特别优选的施用途径是血管内。适当地,包含根据本发明的合成的NS特异性启动子或表达盒的AAV载体或病毒体可以在手背静脉或前臂前部静脉中施用。前臂前部中的合适静脉是头静脉、正中静脉或基底静脉。这是因为这种施用途径对患者来说通常是安全的,同时仍然允许一些渗透通过血脑屏障(进入CNS)。

[0724] 在一些实施方案中,病毒基因治疗载体可以与一种或多种额外的治疗剂或一种或多种被设计为防止载体被网状内皮系统清除的饱和剂同时或依次施用。

[0725] 当载体是AAV载体时,载体的剂量可以是从小于 $1 \times 10^{10}$  gc/kg到 $1 \times 10^{15}$  gc/kg或更多,适当地从小于 $1 \times 10^{12}$  gc/kg到 $1 \times 10^{14}$  gc/kg,适当地从小于 $5 \times 10^{12}$  gc/kg到 $5 \times 10^{13}$  gc/kg。

[0726] 一般来说,对其有需要的受试者将是哺乳动物,优选灵长类动物,更优选人类。通常,对其有需要的受试者将显示出疾病的特征症状。所述方法通常包括通过表达治疗量的治疗性产物来改善对其有需要的受试者所显示的症状。在一个实施方案中,本发明的治疗方法可用于减少功能能力和日常生活活动的下降,如通过标准评估系统来衡量,例如但不限于总功能能力(TFC)量表。在一个实施方案中,本发明的方法可用于改善用于衡量神经性疾病症状的任何评估的表现。此类评估包括但不限于ADAS-cog(阿尔茨海默病评估量表-认知)、MMSE(简易精神状态检查)、GDS(老年抑郁症量表)、FAQ(功能活动问卷)、ADL(日常生活活动)、GPCOG(全科医生认知评估)、Mini-Cog、AMTS(简略心理测试评分)、画钟测试(Clock-drawing test)、6-CIT(6项认知障碍测试)、TYM(测试你的记忆)、MoCa(蒙特利尔认知评估)、ACE-R(Addenbrookes认知评估)、MIS(记忆障碍筛查)、BADLS(Bristol日常生活活动量表)、巴塞尔指数、功能独立性评测、工具性日常生活活动、IQCODE(老年人认知功能下降知情者问卷)、神经精神症状问卷、Cohen-Mansfield激越问卷、BEHAVE-AD、EuroQol、Short Form-36和/或MBR Caregiver Strain Instrument,或Sheehan B(Ther Adv Neurol Disord.5(6):349-358(2012))中所述的任何其他测试,其内容通过引用整体并入本文。

[0727] 在体外和体内在靶细胞中进行治疗性基因表达的基因治疗方案在本领域中是众所周知的,在此不作详细讨论。简而言之,它们包括质粒DNA载体(裸露的或在脂质体中)或病毒载体的静脉内或动脉内施用(例如颈动脉内、肝动脉内、肝静脉内)、脑室内、颅内施用、

肌肉内注射、间质注射、气道内灌注、应用于内皮细胞和肝实质细胞内。已经开发了各种装置来提高DNA对靶细胞的可用性。一个简单的方法是用含有相关载体的导管或可植入材料与靶细胞进行物理接触,而更复杂的方法可以使用喷射装置或类似装置。基因转移到哺乳动物NS细胞中可以使用离体和体内程序进行。离体方法通常需要收获NS细胞,用合适的表达载体进行体外转导,然后将转导的NS细胞重新引入NS中。由于在NS中收获和重新引入NS细胞的难度和危险性,这种方法通常不太优选。体内基因转移已通过将DNA或病毒载体直接注射到NS中而实现,例如通过颅内注射,或通过静脉内或动脉内注射病毒载体。

[0728] 在一个实施方案中,可以将基因治疗载体以治疗有效量施用于受试者(例如,施用于受试者的NS),以减少受试者的神经性疾病症状(例如,使用已知评价方法确定)。在一些实施方案中,基因治疗载体和包含基因治疗载体的组合物可以以允许它们穿过血脑屏障、血管屏障或其他上皮屏障的方式施用。

[0729] 基因治疗载体可以与一种或多种其他治疗性、预防性、研究性或诊断性药剂联合使用。所谓“联合使用”,并不旨在暗示这些药剂必须同时施用和/或配制用于一起递送,尽管这些递送方法在本发明的范围内。组合物可以与一种或多种其他所需的治疗剂或医疗程序同时、在其之前或之后施用。在一些实施方案中,当第二种治疗(例如一种或多种治疗剂)的递送开始时,一种治疗(例如基因治疗载体)的递送仍在进行,因此在施用方面存在重叠。这在本文中有时被称为“同时”或“并行递送”。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前结束。在任一种情况的一些实施方案中,由于联合施用,治疗更有效。例如,第二种治疗更有效,例如,用较少的第二种治疗就能看到相当的效果,或者第二种治疗能在比在没有第一种治疗的情况下施用第二种治疗时所看到的更大的程度上减轻症状,或者用第一种治疗看到的类似的情况。在一些实施方案中,递送是这样的:症状或与病症有关的其他参数的减少程度大于在没有另一种治疗的情况下递送一种治疗所能观察到的情况。两种治疗的效果可以是部分相加、完全相加或大于相加。递送可以是这样的:当递送第二种治疗时,递送的第一种治疗的效果仍可检测到。本文所述的组合物和至少一种额外的治疗可以同时、在同一组合物或在单独的组合物中、或依次施用。对于依次施用,本文所述的基因治疗载体可以先施用,一种或多种治疗剂可以后施用,或者施用的顺序可以颠倒。基因治疗载体和一种或多种治疗剂可以在病症活动期间、或在疾病缓解期间或不太活跃的时期施用。基因治疗载体可以在另一种治疗前、与治疗同时、在治疗后或在病症缓解期间施用。

[0730] 当联合施用时,基因治疗载体和一种或多种治疗剂(例如第二或第三种治疗剂)或所有的治疗方法可以以高于、低于或等于单独使用的每种治疗方法(例如作为单一疗法)的量或剂量施用。在某些实施方案中,基因治疗载体和一种或多种治疗剂(例如第二或第三剂)或所有的治疗方法的施用量或剂量低于(例如至少20%、至少30%、至少40%或至少50%)单独使用的每种治疗方法的量或剂量。在其他实施方案中,产生预期效果(例如治疗NS疾病或病症)的基因治疗载体和一种或多种治疗剂(例如第二或第三剂)或所有的治疗方法的量或剂量低于为达到相同治疗效果而单独使用的每一种治疗方法所需的量或剂量(例如低于至少20%、至少30%、至少40%或至少50%)。

[0731] 可以与本文所述的AAV颗粒联合使用的化合物包括但不限于:胆碱酯酶抑制剂(多奈哌齐(donepezil)、利凡斯的明(rivastigmine)、加兰他敏(galantamine))、NMDA受体拮

抗剂如美金刚(memantine),抗精神病药,抗抑郁药,抗惊厥药(例如,用于肌阵挛的丙戊酸钠和左乙拉西坦(levetiracetam)),分泌酶抑制剂,淀粉样聚集抑制剂,铜或锌调节剂,BACE抑制剂,tau聚集抑制剂如亚甲蓝、吩噻嗪类、葱醌类、正苯胺类或罗丹明类,微管稳定剂如NAP、他克唑(taxol)或紫杉醇,激酶或磷酸酶抑制剂如靶向GSK3(锂)或PP2A的抑制剂,用淀粉样β肽或tau磷酸化表位免疫,抗tau或抗淀粉样抗体,多巴胺耗竭剂(例如,用于舞蹈症的四苯喹啉),苯二氮平类药物(例如,用于肌阵挛、舞蹈症、肌张力障碍、僵直和/或痉挛的氯硝西泮(clonazepam)),多巴胺的氨基酸前体(例如,用于僵直的左旋多巴),骨骼肌松弛剂(例如,用于僵直和/或痉挛的巴氯芬(baclofen)、替扎尼丁(tizanidine)),用于在神经肌肉接头处释放乙酰胆碱以引起肌肉麻痹的抑制剂(例如,用于磨牙症和/或肌张力障碍的肉毒杆菌毒素),非典型神经抑制剂(例如,用于精神病和/或烦躁症的奥氮平和喹硫平,用于精神病、舞蹈症和/或烦躁症的利培酮、舒必利和氟哌啶醇,用于耐药性精神病的氯氮平,用于有显著阴性症状的精神病的阿立哌唑),选择性5-羟色胺再摄取抑制剂(SSRI)(例如,用于抑郁症、焦虑症、强迫性行为和/或烦躁症的西酞普兰、氟西汀、帕罗西汀、舍曲林、米氮平、文拉法辛(venlafaxine)),催眠药(例如,用于改变睡眠-觉醒周期的佐匹克隆(xopiclone)和/或唑吡坦(Zolpidem)),抗惊厥剂(例如,用于躁狂症或轻度躁狂症的丙戊酸钠和卡马西平(轻度躁狂))和情绪稳定剂(例如,用于躁狂症或轻度躁狂症的锂)。

[0732] 根据一些优选实施方案,上述方法可用于治疗患有上文讨论的NS相关疾病(适当地为CNS疾病,如亨廷顿氏病)的受试者。

[0733] 定义和一般要点

[0734] 尽管下面详细讨论了本发明的各种实施方案的制备和使用,但是应该理解,本发明提供了许多可以在特定上下文的各种变化中体现的可应用的发明概念。本文所讨论的具体实施方案仅用于说明制备和使用本发明的具体方式,并不限制本发明的范围。

[0735] 为了便于理解本发明,下文定义了一些术语。本文定义的术语具有与本发明相关领域的普通技术人员通常理解的含义。诸如“一个”、“一种”和“所述”等术语并不旨在仅指单数实体,而是包括可用于说明的具体实施例的一般类别。本文的术语被用来描述本发明的具体实施方案,但它们的使用并不限定本发明,除非在权利要求书中概述。

[0736] 本文包括对本发明背景的讨论以解释本发明的上下文。这不应被视为承认所提及的任何材料在任何权利要求的优先权日之前已在任何国家/地区公布、已知或为公知常识的一部分。

[0737] 在整个本公开中,各种出版物、专利和公开的专利说明书通过识别引用来引用。本说明书中引用的所有文件均通过引用整体并入本文。特别地,本文特别提及的此类文件的教导或部分通过引用并入。

[0738] 除非另有说明,否则本发明的实施将采用细胞生物学、细胞培养、分子生物学、转基因生物学、微生物学、重组DNA和免疫学的常规技术,这些技术在本领域的技术范围内。此类技术在文献中有充分的解释。例如,参见Current Protocols in Molecular Biology (Ausubel,2000,Wiley and son Inc,Library of Congress,USA);Molecular Cloning:A Laboratory Manual,Third Edition,(Sambrook et al,2001,Cold Spring Harbor,New York:Cold Spring Harbor Laboratory Press);Oligonucleotide Synthesis(M.J.Gait ed.,1984);美国专利号4,683,195;Nucleic Acid Hybridization(Harries and Higgins

eds.1984);Transcription and Translation(Hames and Higgins eds.1984);Culture of Animal Cells(Freshney,Alan R.Liss,Inc.,1987);Immobilized Cells and Enzymes (IRL Press,1986);Perbal,A Practical Guide to Molecular Cloning(1984);the series,Methods in Enzymology(Abelson and Simon,eds.-in-chief,Academic Press, Inc.,New York),特别是Vols.154-155(Wu et al.eds.)和Vol.185,"Gene Expression Technology"(Goeddel,ed.);Gene Transfer Vectors For Mammalian Cells(Miller and Calos eds.,1987,Cold Spring Harbor Laboratory);Immunochemical Methods in Cell and Molecular Biology(Mayer and Walker,eds.,Academic Press,London,1987); Handbook of Experimental Immunology,Vols.I-IV(Weir and Blackwell,eds.,1986); 和Manipulating the Mouse Embryo,(Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,1986)。

[0739] 术语“神经系统”或“NS”被技术人员充分理解。NS由周围神经系统(PNS)和中枢神经系统(CNS)组成。

[0740] 术语“中枢神经系统”或“CNS”被技术人员充分理解。CNS由大脑和脊髓组成。优选地,合成的NS特异性启动子在大脑中具有活性。本发明的启动子可以在大脑和/或脊髓中具有活性。优选地,CNS是哺乳动物的CNS,甚至更优选是人类受试者的CNS。

[0741] 术语“周围神经系统”或“PNS”被技术人员充分理解。PNS是指大脑和脊髓以外的神经系统部分。本发明的启动子可以在PNS中具有活性。优选地,PNS是哺乳动物的PNS,甚至更优选是人类受试者的PNS。

[0742] 术语“一个CNS细胞”或“多个CNS细胞”涉及在CNS(CNS组织)中发现的细胞或来源于CNS组织的细胞。CNS细胞可以是原代细胞或细胞系(例如SH-Sy5y、Neuro2A、U87-MG)。CNS细胞可以是体内的(例如在CNS组织中)或体外的(例如在细胞培养中)。CNS细胞包括神经元、星形胶质细胞、少突胶质细胞、小胶质细胞和室管膜细胞。在CNS组织中发现的神经元包括细胞体、长轴和突触末端。神经元通过其长轴将细胞体中接收到的电信号传递给靠近其突触末端的其他细胞。少突胶质细胞是CNS中的一种神经胶质细胞,它产生的髓鞘包裹着神经元轴突以加快电信号的传导。星形胶质细胞是星形的,是CNS中最丰富的细胞类型。它们有多种作用,帮助和调节CNS内电脉冲的传输和神经元功能。小胶质细胞是CNS中的常驻巨噬细胞,参与免疫防御。室管膜细胞形成脑室的上皮层。如本文所用,术语“一个CNS细胞”或“多个CNS细胞”包括神经元、星形胶质细胞、少突胶质细胞、小胶质细胞和/或室管膜细胞。本发明的启动子可以在任何CNS细胞(例如神经元)中具有活性。本发明的启动子可以在两种类型的CNS细胞(例如神经元和星形胶质细胞)中具有活性。本发明的启动子可以在三种或更多类型的CNS细胞(例如神经元、少突胶质细胞和星形胶质细胞)中具有活性。本发明的启动子可以在所有类型的CNS细胞(神经元、星形胶质细胞、少突胶质细胞、小胶质细胞和室管膜细胞)中具有活性。另外,本发明的合成的CNS特异性启动子可以在某一类型的CNS细胞的亚型如多巴胺能神经元、GABA能神经元或成熟的少突胶质细胞中具有活性。在一些实施方案中,本发明的合成的CNS特异性启动子可以只在某一类型的CNS细胞的亚型如多巴胺能神经元或成熟的少突胶质细胞中具有活性。本发明的CRE、近端/最小启动子和启动子可以在CNS的特定区域、特定的CNS细胞和/或CNS细胞亚型或两者中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在CNS的所有区域内的特定CNS细胞类型

如神经元中具有活性。在其他实施方案中,本发明的CRE、近端/最小启动子和启动子可以在CNS的不超过一个区域内的特定CNS细胞类型如神经元中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在CNS的所有区域的所有CNS细胞中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在CNS的不超过一个区域的所有CNS细胞中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在CNS的多个区域的多个CNS细胞中具有活性。

[0743] 术语“一个PNS细胞”或“多个PNS细胞”涉及在PNS(PNS组织)中发现的细胞或来源于PNS组织的细胞。PNS细胞可以是原代细胞或细胞系。PNS细胞可以是体内的(例如在PNS组织中)或体外的(例如在细胞培养中)。PNS细胞包括神经元、许旺细胞和卫星细胞。在PNS组织中发现的神经元包括细胞体、长轴和突触末端。神经元通过其长轴将细胞体中接收到的电信号传递给靠近其突触末端的其他细胞。许旺细胞是PNS中的一种神经胶质细胞,它产生的髓鞘包裹着神经元轴突,以加快电信号的传导。卫星细胞(卫星神经胶质细胞)是覆盖在周围神经系统的神经节中的神经元细胞体表面的神经胶质细胞,支持神经元细胞。它们有多种作用,包括控制交感神经节的微环境。如本文所用,术语“一个PNS细胞”或“多个PNS细胞”包括神经元、许旺细胞和/或卫星细胞。本发明的启动子可以在任何一种PNS细胞(例如神经元)中具有活性。本发明的启动子可以在两种类型的PNS细胞(例如神经元和卫星细胞)中具有活性。本发明的启动子可以在所有类型的PNS细胞(例如神经元、许旺细胞和卫星细胞)中具有活性。另外,本发明的合成的NS特异性启动子可以在某一类型的PNS细胞的亚型如感觉神经元中具有活性。在一些实施方案中,本发明的合成的NS特异性启动子可以只在某一类型的PNS细胞的亚型如感觉神经元中具有活性。本发明的CRE、近端/最小启动子和启动子可以在PNS的特定区域、特定的PNS细胞和/或PNS细胞亚型或两者中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在PNS的所有区域内的特定PNS细胞类型如神经元中具有活性。在其他实施方案中,本发明的CRE、近端/最小启动子和启动子可以在PNS的不超过一个区域内的特定PNS细胞类型如神经元中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在PNS所有区域的所有PNS细胞中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在PNS的不超过一个区域的所有PNS细胞中具有活性。在一些实施方案中,本发明的CRE、近端/最小启动子和启动子可以在PNS的多个区域的多个PNS细胞中具有活性。

[0744] 术语“顺式调节元件”或“CRE”是技术人员众所周知的术语,是指可以调控或调节相邻基因(即顺式)的转录的核酸序列,例如增强子、启动子、绝缘子或沉默子。CRE存在于它们调节的基因附近。CRE通常通过与TF结合来调节基因转录,即它们包括转录因子结合位点(TFBS)。单个TF可以与许多CRE结合,从而控制许多基因的表达(多效性)。CRE通常但不总是位于它们调节的基因的转录起始位点(TSS)的上游。本文中的“增强子”是增强(即上调)与它们可操作地相关的基因的转录的CRE,并且可以存在于在它们调节的基因的上游、下游甚至内含子中。多个增强子可以协同作用来调节一个基因的转录。在本文中,“沉默子”涉及与称为阻遏物的TF结合的CRE,阻遏物起到阻止或下调基因转录的作用。术语“沉默子”也可以指信使RNA的3'非翻译区域中的区域,它结合抑制该mRNA分子翻译的蛋白质,但这种用法不同于它在描述CRE时的用途。通常,本发明的CRE是NS特异性的增强子元件(通常称为NS特异性CRE,或NS特异性CRE增强子等)。在本文中,优选CRE位于距转录起始位点(TSS)2500个核

核苷酸或更少,更优选距TSS 2000个核苷酸或更少,更优选距TSS 1500个核苷酸或更少,适当地距TSS 1000、750、500、250、200、150或100个核苷酸或更少。本发明的CRE优选长度较短,优选长度为1000个核苷酸或更短,例如它们的长度可以是800、700、600、500、400、300、200、175、150、90、80、70、60或50个核苷酸或更短。本发明的CRE通常与可操作地连接的启动子元件组合提供,所述启动子元件可以是最小启动子或近端启动子;本发明的CRE可以增强启动子元件的NS特异性活性。

[0745] 术语“顺式调节模块”或“CRM”是指功能性调节核酸模块,其通常包括两个或更多个CRE;在本发明中,CRE通常是NS特异性增强子,因此CRM是合成的NS特异性调节核酸。CRM可以包括多个NS特异性CRE。适当地,CRM中包含的至少一个CRE是根据SEQ ID NO:19-24、27、28、37、38的CRE或其功能变体。通常,CRM中的多个CRE共同作用(例如,相加或协同)以增强与包含CRM的启动子可操作地相关的基因的转录。在CRM中对CRE进行洗牌(即重新排序)、使CRE反转(即逆转方向)和更改CRE的间距有相当大的范围。因此,本发明的CRM的功能变体尤其包括所引用的CRM的变体,其中它们中的CRE已被洗牌和/或反转,和/或CRE之间的间距已被改变。

[0746] 如本文所用,短语“启动子”是指通常位于要转录的核酸序列上游的DNA区域,该区域是发生转录所需的,即其起始转录。启动子允许在其控制下正确激活或抑制编码序列的转录。启动子通常包含被多个TF识别和结合的特定序列。TF与启动子序列结合并导致RNA聚合酶的募集,RNA聚合酶是一种从基因编码区合成RNA的酶。许多不同的启动子是本领域已知的。

[0747] 如本文所用,术语“合成启动子”涉及自然界中不存在的启动子。在本文中,它通常包含与最小(或核心)启动子或NS特异性近端启动子(启动子元件)可操作地连接的本发明的CRE和/或CRM。本发明的CRE和/或CRM用于增强与合成启动子可操作地连接的基因的NS特异性转录。部分合成启动子可以是天然存在的(例如最小启动子或启动子中的一个或多个CRE),但合成启动子作为一个实体不是天然存在的。替代地,合成启动子可以是自然界中出现的启动子的较短的、截短的版本。

[0748] 如本文所用,“最小启动子”(也称为“核心启动子”)是指通常短的DNA区段,其本身无活性或大部分无活性,但当与其他转录调节元件组合时可介导转录。最小启动子序列可以来自各种不同的来源,包括原核和真核基因。最小启动子的实例包括多巴胺 $\beta$ -羟化酶基因最小启动子、巨细胞病毒(CMV)立即早期基因最小启动子(CMV-MP)和疱疹胸苷激酶最小启动子(MinTK)。最小启动子通常包含转录起始位点(TSS)和直接上游的元件、RNA聚合酶II的结合位点和一般转录因子结合位点(通常是TATA框)。最小启动子还可以包括TSS下游的一些元件,但这些元件通常在没有另外的调节元件的情况下几乎没有功能。

[0749] 如本文所用,“近端启动子”涉及最小启动子加上至少一些额外的调节序列,通常是趋向于含有初级调节元件的基因上游的近端序列。它通常在TSS上游延伸大约250个碱基对,并包括特定的TFBS。近端启动子还可以包括TSS下游的一种或多种调节元件,例如UTR或内含子。在当前情况下,适当地近端启动子可以是天然存在的NS特异性近端启动子的较短的、截短的版本。本发明的近端启动子可以与本发明的一种或多种CRE或CRM组合。然而,近端启动子也可以是合成的。

[0750] 如本文所用,“启动子元件”是指如上定义的最小启动子或近端启动子。在本发明

的上下文中,启动子元件可以与一个或多个CRE组合以提供本发明的合成的NS特异性启动子。

[0751] 在本发明的上下文中,CRE、CRM、启动子元件、启动子或其他调节核酸的“功能变体”是参考序列的变体,其保留以与参考序列相同的方式发挥作用的能力,例如,作为NS特异性CRE、NS特异性CRM或NS特异性启动子。此类功能变体的替代术语包括“生物等效物”或“等效物”。

[0752] 应当理解,给定的CRE、CRM、启动子或其他调节序列作为NS特异性增强子发挥作用的能力显著取决于该序列结合与参考序列结合的共同NS特异性TF的能力。因此,在大多数情况下,CRE或CRM的功能变体将包含与参考CRE、CRM或启动子相同的大部分或全部TF的TFBS。优选但不是必需的,功能变体的TFBS与参考CRE、CRM或启动子处于相同的相对位置(即顺序和一般位置)。同样优选但不是必需的,功能变体的TFBS与参考序列的方向相同(应注意,TFBS在一些情况下可以反向存在,例如作为参考序列中的反向互补vis-à-vis序列)。同样优选但不是必需的,功能变体的TFBS与参考序列在同一条链上。因此,在优选实施方案中,功能变体包含与参考序列相同的TF、相同的顺序、相同的位置、相同的方向和相同的链上的TFBS。还将理解,位于TFBS之间的序列(在一些情况下称为间隔区序列等)对CRE或CRM的功能影响较小。这样的序列通常可以有很大的变化,并且它们的长度可以改变。然而,在优选实施方案中,功能变体中的间距(即相邻TFBS之间的距离)与参考序列中的间距基本上相同(例如,其变化不超过20%,优选不超过10%,更优选的是近似相同)。很明显,在一些情况下,CRE的功能变体可以以相反的方向存在,例如它可以是上述CRE的反向补体,或其变体。

[0753] 功能变体和参考序列之间的序列同一性水平也可以是保留的功能的指标。CRE、CRM或启动子的TFBS中高水平的序列同一性通常比间隔区序列(其中很少需要或不需要任何序列保守性)中的序列同一性更重要。然而,应当理解,即使在TFBS内,也可以适应相当程度的序列变异,因为功能性TFBS的序列不需要与共有序列完全匹配。

[0754] 一种或多种TF与给定功能变体中的TFBS结合的能力可以通过本领域已知的任何相关手段来确定,包括但不限于电迁移率测定(EMSA)、结合测定、染色质免疫沉淀(ChIP)和ChIP测序(ChIP-seq)。在优选实施方案中,一种或多种TF结合给定功能变体的能力由EMSA确定。执行EMSA的方法在本领域中是众所周知的。Sambrook等人(上文引用的)描述了合适的方法。许多描述此过程的相关文章都可获得,例如Hellman和Fried,Nat Protoc.2007;2(8):1849-1861。

[0755] “NS特异性”或“NS特异性表达”是指顺式调节元件、顺式调节模块或启动子以与其他组织(例如肝脏、肾脏、脾脏、心脏、肌肉和肺)相比优先或占优势的方式增强或驱动基因在NS细胞(或NS来源的细胞)中的表达的能力。基因的表达可以以mRNA或蛋白质的形式。在优选实施方案中,NS特异性表达使得在其他(即非NS)组织或细胞中的表达可以忽略不计,即表达是高度NS特异性的。NS特异性表达可以是CNS特异性和/或PNS特异性表达。

[0756] “CNS特异性”或“CNS特异性表达”是指顺式调节元件、顺式调节模块或启动子以与其他组织(例如肝脏、肾脏、脾脏、心脏、肌肉和肺)相比优先或占优势的方式增强或驱动基因在CNS细胞(或CNS来源的细胞)中的表达的能力。基因的表达可以以mRNA或蛋白质的形式。在优选实施方案中,CNS特异性表达使得在其他(即非CNS)组织或细胞中的表达可以忽

略不计,即表达是高度CNS特异性的。

[0757] “PNS特异性”或“PNS特异性表达”是指顺式调节元件、顺式调节模块或启动子以与其他组织(例如肝脏、肾脏、脾脏、心脏、肌肉和肺)相比优先或占优势的方式增强或驱动基因在PNS细胞(或PNS来源的细胞)中的表达的能力。基因的表达可以以mRNA或蛋白质的形式。在优选实施方案中,PNS特异性表达使得在其他(即非PNS)组织或细胞中的表达可以忽略不计,即表达是高度PNS特异性的。

[0758] 技术人员可以容易地评估CRE、CRM或启动子作为NS特异性CRE、CRM或启动子起作用的能力。因此,技术人员可以容易地确定以上所述的特异性CRE、CRM或启动子的任何变体是否保留功能(即,它是如上定义的功能变体)。例如,任何要评估的给定CRM可以与最小启动子(例如,位于CMV-MP上游)可操作地连接,并测量顺式调节元件驱动基因(通常为报告基因)的NS特异性表达的能力。替代地,可以将CRE或CRM的变体替换为合成的NS特异性启动子来代替参考CRE或CRM,并且可以确定对由所述经修饰的启动子驱动的NS特异性表达的影响,并与未经修饰的形式进行比较。类似地,技术人员可以容易地评估启动子驱动NS特异性表达的能力(例如,如以下实施例所述)。可以将由参考启动子的变体驱动的基因的表达水平与由参考启动子驱动的表达水平进行比较。在一些实施方案中,其中由变体启动子驱动的NS特异性表达水平为由参考启动子驱动的表达水平的至少50%、至少60%、至少70%、至少80%、至少90%或至少100%,可以说所述变体保留功能。可以容易地构建用于评估NS特异性表达增强的合适的核酸构建体和报告基因测定,并且下面列出的实施例给出了合适的方法学。

[0759] 可以鉴定NS特异性,其中基因(例如治疗性基因或报告基因)的表达优先或占优势地发生在NS来源的细胞中。优先或占优势的表达可以被定义为,例如,在NS来源的细胞中的表达水平明显高于其他类型的细胞(即非NS来源的细胞)。例如,在NS来源的细胞中的表达适当地比非NS细胞至少高5倍,优选比非NS细胞至少高10倍,在一些情况下可能高50倍或更多。为方便起见,适当地可以通过比较不同的非NS细胞系,例如原代NS细胞或NS来源的细胞系,如SH-Sy5y、Neuro2A、U87-MG中的表达水平,与肌肉来源的细胞系如C2C12或H2K细胞(骨骼肌)或H9C2细胞(心脏)、肝脏来源的细胞系(如Huh7或HepG)、肾脏来源的细胞系(如HEK-293)、宫颈组织来源的细胞系(如HeLa)和/或肺来源的细胞系(如A549)中的表达水平来证明NS特异性表达。

[0760] 本发明的合成的NS特异性启动子与非组织特异性启动子如CMV-IE相比,优选在非NS来源的细胞中,适当地在C2C12、H9C2、Huh7、HEK-293、HeLa和/或A549细胞中表达减少。本发明的合成的NS特异性启动子优选在非NS来源的细胞中(适当地在C2C12、H9C2、Huh7、HEK-293、HeLa和/或A549中)的活性比CMV-IE启动子低50%或更低,适当地25%或更低、20%或更低、15%或更低、10%或更低、5%或更低或1%或更低。一般来说,优选使在非NS来源的细胞中的表达最小化,但在一些情况下,这可能是不必要的。即使本发明的合成的NS特异性启动子在例如一个或两个非NS细胞中具有较高的表达,只要它通常在一系列NS细胞中相对于非NS细胞总体上具有较高的表达,它仍然可以是NS特异性启动子。

[0761] 本发明的合成的NS特异性启动子优选适合于促进受试者NS中的表达,例如驱动转基因,优选治疗性转基因的NS特异性表达。本发明优选的合成的NS特异性启动子适合于促进NS特异性转基因的表达,并且在NS细胞中的活性是Synapsin-1启动子活性的至少15%、

20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、175%、200%、250%、300%、350%或400%。在一些实施方案中,本发明的合成的NS特异性启动子适合于促进NS特异性转基因的表达,其水平为Synapsin-1启动子活性的至少100%,优选为Synapsin-1启动子活性的150%、200%、300%或500%。这样的NS特异性表达适当地在NS来源的细胞中确定,例如SH-Sy5y、Neuro2A、U87-MG细胞系或原代NS细胞(适当地原代人类神经元、星形胶质细胞、少突胶质细胞、小胶质细胞、许旺细胞、卫星细胞和/或室管膜细胞)。

[0762] 与CMV-IE相比,本发明的合成的NS特异性启动子还能够在NS来源的细胞,例如SH-Sy5y、Neuro2A、U87-MG细胞系或原代NS细胞(适当地原代人类神经元、星形胶质细胞、少突胶质细胞、小胶质细胞、许旺细胞、卫星细胞和/或室管膜细胞)中以至少50%、100%、150%或200%的水平促进基因的NS特异性表达。

[0763] 如本文所用,术语“核酸”通常是指基本上由核苷酸组成的任何长度的低聚物或聚合物(优选线性聚合物)。核苷酸单元通常包括杂环碱基、糖基团和至少一个(例如一个、两个或三个)磷酸基团,包括修饰的或替换的磷酸基团。杂环碱基尤其可以包括嘌呤和嘧啶碱基,例如广泛存在于天然存在的核酸中的腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)、胸腺嘧啶(T)和尿嘧啶(U),其他天然存在的碱基(例如黄嘌呤、肌苷、次黄嘌呤)以及化学或生物化学修饰的(例如甲基化的)、非天然的或衍生的碱基。糖基团尤其可以包括戊糖(呋喃戊糖)基团,例如优选地在天然存在的核酸中常见的核糖和/或2-脱氧核糖,或阿拉伯糖、2-脱氧阿拉伯糖、苏阿糖或己糖糖基团,以及修饰或替换的糖基团。如本文所指的核酸可以包括天然存在的核苷酸、修饰的核苷酸或其混合物。修饰的核苷酸可以包括修饰的杂环碱基、修饰的糖部分、修饰的磷酸基团或其组合。可以引入磷酸基团或糖的修饰以提高稳定性、对酶降解的抗性或一些其他有用的性质。术语“核酸”进一步优选包括DNA、RNA和DNA RNA杂合分子,具体包括hnRNA、前mRNA、mRNA、cDNA、基因组DNA、扩增产物、寡核苷酸和合成的(例如化学合成的)DNA、RNA或DNA RNA杂合体。核酸可以是天然存在的,例如,存在于自然界中或从自然界中分离;或者可以是非天然存在的,例如重组的,即通过重组DNA技术产生的,和/或部分或全部化学或生物化学合成的。“核酸”可以是双链的、部分双链的或单链的。在单链的情况下,核酸可以是有义链或反义链。此外,核酸可以是环状的或线性的。

[0764] 当提及核酸时,“分离的”是指完全或部分缺乏通常在自然界中与其相关的序列的核酸分子或核酸序列;或一种序列,虽然它在自然界中存在,但具有与之相关的异源序列;或与染色体分离的分子。

[0765] 术语“同一性”和“相同”等是指两个聚合分子之间,例如两个核酸分子之间,例如两个DNA分子之间的序列相似性。可以进行序列比对和序列同一性的确定,例如,使用Altschul等人1990(*J Mol Biol* 215:403-10)最初描述的基本局部比对搜索工具(BLAST),例如Tatusova和Madden 1999(*FEMS Microbiol Lett* 174:247-250)描述的“Blast 2序列”算法。

[0766] 用于比对序列以进行比较的方法是本领域众所周知的。各种程序和比对算法描述于,例如:Smith and Waterman(1981)*Adv. Appl. Math.* 2:482;Needleman and Wunsch(1970)*J. Mol. Biol.* 48:443;Pearson and Lipman(1988)*Proc. Natl. Acad. Sci. U.S.A.* 85:2444;Higgins and Sharp(1988)*Gene* 73:237-44;Higgins and Sharp(1989)*CABIOS* 5:151-3;Corpet et al. (1988)*Nucleic Acids Res.* 16:10881-90;Huang et al. (1992)

Comp. Appl. Biosci. 8:155-65; Pearson et al. (1994) Methods Mol. Biol. 24:307-31; Tatiana et al. (1999) FEMS Microbiol. Lett. 174:247-50中。序列比对方法和同源性计算的详细考虑可以在例如 Altschul et al. (1990) J. Mol. Biol. 215:403-10中找到。

[0767] 美国国家生物技术信息中心 (NCBI) 基本局部比对搜索工具 (BLAST<sup>TM</sup>; Altschul 等人 (1990)) 可以从多个来源获得, 包括美国国家生物技术信息中心 (Bethesda, MD) 和互联网, 用于与多个序列分析程序结合使用。互联网上 BLAST<sup>TM</sup> 的“帮助”部分下提供了有关如何使用该程序确定序列同一性的说明。对于核酸序列的比较, 可以使用 BLAST<sup>TM</sup> (Blastn) 程序的“Blast 2 序列”功能, 使用默认参数。当通过该方法评估时, 与参考序列具有更大相似性的核酸序列将显示出更高的同一性百分比。通常, 在整个序列长度上计算序列同一性百分比。

[0768] 例如, 适当地通过 Needleman-Wunsch 算法使用以下评分参数找到全局最优比对: 匹配得分: +2, 错配得分: -3; 空位罚分: 空位打开 5, 空位延伸 2。适当地, 通过比对上的碱基数与比对总长度的比率计算所得最佳全局比对的同一性百分比, 其中比对长度包括匹配和错配, 乘以 100。

[0769] 术语“转录因子结合位点”(TFBS) 在本领域中是众所周知的。对技术人员来说, TFBS 序列显然可以被修饰, 只要它们被预期的转录因子 (TF) 结合。本文公开的各种 TFBS 的共有序列在本领域中是已知的, 并且技术人员可以容易地使用该信息来确定替代的 TFBS。此外, TF 与给定推定序列结合的能力可以容易地由技术人员通过实验确定 (例如, 通过 EMSA 和本领域众所周知的和本文讨论的其他方法)。

[0770] “共有序列”的含义在本领域中是众所周知的。在本申请中, 除非上下文另有说明, 否则以下符号用于共有序列。考虑以下示例性 DNA 序列:

[0771] A[CT]N{A}YR

[0772] A 表示始终在该位置发现 A; [CT] 代表该位置为 C 或 T; N 代表该位置为任意碱基; {A} 表示在该位置发现除 A 之外的任意碱基。Y 代表任意嘧啶, R 代表任意嘌呤。

[0773] 本申请中的“合成的”是指自然界中不存在的核酸分子。本发明的合成核酸是人工产生的, 通常通过重组技术或从头合成。此类合成核酸可以包含天然存在的序列 (例如启动子、增强子、内含子和其他此类调节序列), 但这些存在于非天然存在的环境中。例如, 合成基因 (或基因的一部分) 通常包含一个或多个在性质上不连续的核酸序列 (嵌合序列), 和/或可以涵盖替换、插入和缺失及其组合。

[0774] 如本文所用, “互补”或“互补性”是指两个核酸序列的 Watson-Crick 碱基配对。例如, 序列 5' -AGT-3' 与互补序列 3' -TCA-5' 结合。两个核酸序列之间的互补性可能是“部分的”, 其中只有一些碱基与其互补物结合, 或者当序列中的每个碱基与其互补碱基结合时, 它可能是完全的。

[0775] 如本文所用, 术语“施用”是指将外来物质引入人体或动物体内。例如, 施用可以是静脉内、动脉内或颅内。

[0776] 本申请中的“转染”泛指有意将核酸引入细胞的任何过程, 包括病毒和非病毒载体的引入, 包括或等同于转化、转导等术语和过程。实例包括但不限于: 用病毒载体转染; 用质粒载体转化; 电穿孔 (Fromm 等人 (1986) Nature 319:791-3); 脂转染 (Feigner 等人 (1987) Proc. Natl. Acad. Sci. USA 84:7413-7); 显微注射 (Mueller 等人 (1978) Cell 15:579-85);

农杆菌介导的转移 (Fraley等人 (1983) Proc. Natl. Acad. Sci. USA 80:4803-7); 直接DNA吸收; 晶须介导的转化; 和微粒轰击 (Klein等人 (1987) Nature 327:70)。

[0777] 如本文所用, 短语“转基因”是指外源核酸序列。在一个实例中, 转基因是编码工业上或药学上有用的化合物的基因, 或编码所需性状的基因。在又一个实例中, 转基因编码有用的核酸, 例如反义核酸序列, 其中反义核酸序列的表达抑制靶核酸序列的表达。转基因优选编码治疗性产物, 例如蛋白质。

[0778] 术语“载体”在本领域中是众所周知的, 并且如本文所用, 是指核酸分子, 例如双链DNA, 其可能已插入根据本发明的核酸序列。适当地, 载体用于将插入的核酸分子转运到合适的宿主细胞中。载体通常包含允许使插入的核酸分子转录并且优选地将转录物翻译成多肽的所有必要元件。载体通常包含所有必要元件, 这样, 一旦载体进入宿主细胞, 载体就可以独立于宿主染色体DNA复制, 或与宿主染色体DNA同时复制; 可以产生数个拷贝的载体及其插入的核酸分子。本发明的载体可以是游离型载体 (即, 不整合到宿主细胞基因组中), 或者可以是整合到宿主细胞基因组中的载体。该定义包括非病毒和病毒载体。非病毒载体包括但不限于质粒载体 (例如pMA-RQ、pUC载体、bluescript载体 (pBS) 和pBR322或其不含细菌序列的衍生物 (小环))、基于转座子的载体 (例如PiggyBac (PB) 载体或睡美人 (SB) 载体) 等。较大的载体如人工染色体 (细菌 (BAC)、酵母 (YAC) 或人类 (HAC)) 可用于容纳较大的插入片段。病毒载体来源于病毒, 包括但不限于逆转录病毒、慢病毒、腺相关病毒、腺病毒、疱疹病毒、肝炎病毒载体等。通常, 但不一定, 病毒载体是复制缺陷型的, 因为它们已经失去了在给定细胞中繁殖的能力, 因为复制所必需的病毒基因已从病毒载体中消除。然而, 一些病毒载体也可以适应在给定细胞如癌细胞中特异性复制, 通常用于引发 (癌症) 细胞特异性 (肿瘤) 溶解。病毒体是包含病毒和非病毒元件的载体的非限制性实例, 特别是它们将脂质体与灭活的HIV或流感病毒结合 (Yamada等人, 2003)。另一个实例包括与阳离子脂质混合的病毒载体。

[0779] 如本文所用, 术语“可操作地连接”或等效表达是指各种核酸元件相对于彼此的排列, 使得所述元件在功能上连接并且能够以预期的方式彼此相互作用。此类元件可以包括但不限于启动子、CRE (例如增强子或其他调节元件)、启动子元件、多聚腺苷酸化序列、一个或多个内含子和/或外显子以及待表达的目的基因的编码序列。当正确定向或可操作地连接时, 核酸序列元件共同作用以调节彼此的活性, 并最终可能影响表达产物的表达水平。调节是指增加、降低或维持特定元件的活性水平。每个元件相对于其他元件的位置可以用每个元件的5'末端和3'末端或它们在另一个元件或位置 (例如TSS或启动子元件) 上游或下游的位置来表示, 并且任何特定元件之间的距离可以通过元件之间的插入核苷酸或碱基对的数量来参考。如技术人员所理解的, 可操作地连接意味着功能活性, 并且不一定与自然位置连接相关。实际上, 当在核酸表达盒中使用时, CRE通常将位于启动子元件的紧上游 (尽管通常是这种情况, 但绝对不应将其解释为对核酸表达盒内位置的限制或排除), 但在体内不一定是这种情况, 例如, 当位于启动子上游时, 天然存在于基因下游的调节元件序列能够以相同的方式发挥作用, 该基因的转录受其影响。因此, 根据一个具体实施方案, 调节元件的调节或增强效果可以是位置无关的。

[0780] 如本文所用, “间隔区序列”或“间隔区”是分隔两个功能性核酸序列 (例如TFBS、CRE、CRM、启动子元件等) 的核酸序列。它基本上可以具有任何序列, 只要它不阻止功能性核

酸序列(例如顺式调节元件)按需要发挥作用(例如,如果它包括沉默子序列、阻止所需转录因子的结合等,则可能发生这种情况)。通常,它是非功能性的,因为它的存在只是为了将相邻的功能性核酸序列彼此隔开。在一些实施方案中,间隔区可以具有75、50、40、30、30或10个核苷酸或更少的长度。

[0781] 如本文所用,术语“药学上可接受的”与本领域一致并且意味着与药物组合物的其他成分相容并且对其接受者无害。

[0782] “治疗有效量”和类似短语是指在受试者中提供所需特定药理作用(例如在NS中表达治疗性基因)的剂量或血浆浓度。治疗有效量在治疗本文所述的病况中可能并不总是有效,即使这样的剂量被本领域技术人员认为是治疗有效量。治疗有效量可以根据施用途径和剂型、受试者的年龄和体重和/或所治疗的疾病或病况而变化。

[0783] 如本文所用,术语“AAV载体”在本领域中是众所周知的,一般是指包括各种核酸序列的AAV载体核酸序列。如本文所用,AAV载体通常包含非AAV来源的异源核酸序列作为载体的一部分。该异源核酸序列通常包含本文公开的启动子以及用于细胞遗传转化的其他目的序列。通常,异源核酸序列的两侧是至少一个,通常是两个AAV反向末端重复序列(ITR)。“AAV病毒体”或“AAV病毒”或“AAV病毒颗粒”或“AAV载体颗粒”是指由至少一种AAV衣壳多肽(包括变体AAV衣壳多肽和非变体亲本衣壳多肽两者)组成的病毒颗粒以及包被的多核苷酸AAV载体。如果所述颗粒包含异源核酸(即除野生型AAV基因组之外的多核苷酸,例如要递送至哺乳动物细胞的转基因),它可以被称为“AAV载体颗粒”或简称为“AAV载体”。因此,AAV病毒体或AAV颗粒的产生必然包括AAV载体的产生,因为这样的载体包含在AAV病毒体或AAV颗粒中。ITR可以来自与衣壳相同的血清型,选自表1中所列的任何血清型,或者可以来自与衣壳不同的血清型。AAV载体通常有一个以上的ITR。在非限制性实例中,AAV载体具有包含两个ITR的病毒基因组。在一个实施方案中,这些ITR彼此属于相同的血清型。在另一个实施方案中,ITR属于不同的血清型。非限制性实例包括零、一个或两个具有与衣壳相同的血清型的ITR。独立而言,每个ITR的长度可以是约100至约150个核苷酸。ITR的长度可以是约100-105个核苷酸、106-110个核苷酸、111-115个核苷酸、116-120个核苷酸、121-125个核苷酸、126-130个核苷酸、131-135个核苷酸、136-140个核苷酸、141-145个核苷酸或146-150个核苷酸。在一个实施方案中,ITR的长度为140-142个核苷酸。ITR长度的非限制性实例是长度为102、105、130、140、141、142、145个核苷酸。

[0784] 如本文所用,术语“微小RNA”是指任何类型的干扰RNA,包括但不限于内源性微小RNA和人工微小RNA(例如,合成miRNA)。内源性微小RNA是基因组中天然编码的小RNA,能够调节mRNA的生产性利用。人工微小RNA可以是除内源性微小RNA之外的能够调节mRNA活性的任何类型的RNA序列。微小RNA序列可以由任何一种或多种这些序列组成的RNA分子。微小RNA(或“miRNA”)序列已在出版物中有所描述,例如Lim, et al, 2003, *Genes&Development*, 17, 991-1008; Lim et al, 2003, *Science*, 299, 1540; Lee and Ambrose, 2001, *Science*, 294, 862; Lau et al, 2001, *Science* 294, 858-861; Lagos-Quintana et al, 2002, *Current Biology*, 12, 735-739; Lagos-Quintana et al., 2001, *Science*, 294, 853-857; 和Lagos-Quintana et al., 2003, *RNA*, 9, 175-179。微小RNA的实例包括较大RNA的任何RNA片段,或者是miRNA、siRNA、stRNA、sncRNA、tncRNA、snoRNA、smRNA、shRNA、snRNA或其他小的非编码RNA。参见,例如,美国专利申请20050272923、20050266552、20050142581和20050075492。

“微小RNA前体”（或“前miRNA”）是指具有茎环结构的核酸，其中并入了微小RNA序列。“成熟微小RNA”（或“成熟miRNA”）包括从微小RNA前体（“前miRNA”）裂解的微小RNA或为合成的（例如，通过无细胞合成在实验室中合成），长度为从约19个核苷酸到约27个核苷酸，例如，成熟微小RNA可以具有19nt、20nt、21nt、22nt、23nt、24nt、25nt、26nt或27nt的长度。成熟微小RNA可以与靶mRNA结合并抑制靶mRNA的翻译。

[0785] 术语“治疗”是指减少、改善或消除疾病或病况的一种或多种体征、症状或影响。因此，如本文所用，“治疗”包括对哺乳动物，特别是人类的疾病的任何治疗，并且包括：(a) 防止疾病在易患疾病或有患上疾病的风险但尚未被诊断出患有所述疾病的受试者中发生；(b) 抑制所述疾病，即阻止其发展；和(c) 缓解所述疾病，即导致所述疾病消退。

[0786] 向受试者“施用”药剂包括将药剂引入或递送给受试者以执行其预期功能的任何途径。可以通过任何合适的途径进行施用，包括口服、鼻内、眼内、眼部、肠外（血管内、肌肉内、腹膜内或皮下）或局部。施用包括自我施用和他人施用。静脉内或动脉内施用在本发明中是特别感兴趣的。

[0787] 术语“个体”、“受试者”和“患者”可互换使用，是指患有需要治疗的疾病或病况的任何个体受试者。出于本公开的目的，受试者可以是灵长类动物，优选人类，或另一种哺乳动物，例如狗、猫、马、猪、山羊或牛等。

[0788] 术语“在某一区域或组织中具有特异性活性”是指启动子在该区域或组织中占优势地具有活性，即在该区域或组织中比在其他区域或组织中更具有活性。

[0789] 实施例1

[0790] 序列信息

[0791] 表1-NS特异性启动子

[0792]

名称	序列	长度
SP0013 (SEQ ID NO: 1)	AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGICTGCA AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC CAGTGAATGACTCACCTTGGCACAGACACAATGTTCTGGGGTGGGCACAGTGCC TGCTTCCCGCCGCACCCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA GCAGGGGGCTTGCATTGCACCCCAGCCTGACAGCCTGGCATCTTGGGATAAAA GCAGCACAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGA GAACAAGGCTCTATTCAGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCCA	795

[0793]

	<p>GGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGA                  AGTCCAAGGACACAAATGGGTGAGGGGAGGGCTAGGCCTGCGCACCCACCCA                  CCGACCCCTCACCACCGACCCGTCACCCACCGACCAAGGGGCACCCTGGCCT                  AGAGGGGATGCTGAGCGGGACCCGCTCCTGCCTCTGGCAGTCCCAGATGGG                  ACTTGGACCCCGCAGTTGCTCTCTCGGACCCTAAGTTTCTACCCCTGGATCTAA                  GCGGAGCTGGGTTTGC GGATCCCACGGTTCCCGGCGGGGCGGGGCCCGGTC                  GCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAGCAGTGCATTGTGGGAACT                  CCCGA</p>	
<p>SP0014                  (SEQ                  ID NO:                  2)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC                  CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA                  AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC                  CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC                  TGCTTCCCGCCGACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA                  GCAGGGGGCTTGCAATGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA                  GCAGCACAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGA                  GAACAAGGCTCTATTCAGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCCA                  GGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGA                  AGTCCAAGGACACAAATGGGTGAGGGGATGCGGCGAGGCGCGTGC GCACTGC                  CAGCTTCAGCACCGCGGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGGCCAC                  CGCCGCTCAGCACTGAAGGCGCGCTGACGTC ACTCGCCGGTCCCCGCAAAC                  TCCCCTTCCCGGCCACCTTGGTTCGCGTCCGCGCCGCGCCGCGCCAGCCGGAC                  CGCACCACGCGAGGCGGAGATAGGGGGGACGGGCGGACCATCTGCGCTG                  CGGCGCCGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTC                  GTGTCGTGCCTGAGAGCGCAG</p>	<p>810</p>
<p>SP0026                  (SEQ                  ID NO:</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC                  CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA                  AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC                  CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC</p>	<p>916</p>

[0794]

3)	<p>TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA  GCAGGGGGCTTGCATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA  GCAGCACAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGA  GAACAAGGCTCTATTACGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCCA  GGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGA  AGTCCAAGGACACAAATGGGTGAGGGGAAGCGCGCAGAGTCTGCATGCGTGA  GGAAGCTCCTGGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATG  GCTCCGGAAGTGAGCTGGGGTTGCTGGCAGCCTGGCTGGCACTGGGCTAGGCC  TGCGCACCCACCCACCGACCCCTCACCCACCGACCCGTCACCCACCGACCAAG  GGCACCCCTGGCCTAGAGGGGATGCTGAGCGGGACCCGCTCCTGCCTCTGGC  AGTCCCAGATGGGACTTGGACCCCGCAGTTGCTCTCTCGGACCCTAAGTTTCT  ACCCCTGGATCTAAGGCGGAGCTGGGTTTGC GGATCCCACGGTTCCCGGCGGG  GCGGGGCCCGTCCGCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAGCAGTGC  ATTGTGGGAAACTCCCGA</p>	
<p>SP0027  (SEQ  ID NO:  4)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC  CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA  AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC  CAGTGAATGACTCACCTTGGCACAGACACAATGTTCCGGGGTGGGCACAGTGCC  TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA  GCAGGGGGCTTGCATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA  GCAGCACAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGA  GAACAAGGCTCTATTACGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCCA  GGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGA  AGTCCAAGGACACAAATGGGTGAGGGGAAGCGCGCAGAGTCTGCATGCGTGA  GGAAGCTCCTGGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATG  GCTCCGGAAGTGAGCTGGGGTTGCTGGCAGCCTGGCTGGCACTTGGCGGAGG  CGCGTGC GCACTGCCAGCTCAGCACCGCGGACAGTGCCTTCGCCCCCGCCTG  GCGGCGCGGCCACCGCCGCTCAGCACTGAAGGCGCGCTGACGTCACCTCGCC</p>	931

[0795]

	<p>GGTCCCCGCAAAC'TCCC'TTCCCGGCCACCTTGGTCGCGTCCGCGCCGCCGC                  CGGCCAGCCGGACCGCACACGCGAGGCGGAGATAGGGGGGCACGGGCGC                  GACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGG                  CAGCGGAGGAGTCGTGTCGTGCCTGAGAGCGCAG</p>	
<p>SP0030 (SEQ ID NO: 5)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC                  CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGICTGCA                  AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC                  CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC                  TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA                  GCAGGGGGCTTGCAATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA                  GCAGCACGGGCTAGGCCTGCGCACCCACCCACCGACCCCTCACCCACCGACCC                  GTCACCCACCGACCAAGGGGCACCCTGGCCTAGAGGGGATGCTGAGCGGGAC                  CCGCCTCCTGCCTCTGGCAGTCCCAGATGGGACTTGGACCCCGCAGTTGCTCTC                  TCGGACCCTAAGTTTCTACCCCTGGATCTAAGGCGGAGCTGGGTTTGC GGATC                  CCACGGTTCCCGCGGGGCGGGGCCCGGTGCGCCCTCCCCCTCCCCGCCCTCC                  TGCGCCGGGAGCAGTGCATTGTGGGAAACTCCCGA</p>	<p>617</p>
<p>SP0031 (SEQ ID NO: 6)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC                  CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGICTGCA                  AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC                  CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC                  TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA                  GCAGGGGGCTTGCAATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA                  GCAGCACTGCGGCGAGGCGCGTGC GCACTGCCAGCTTCAGCACCGCGGACAG                  TGCTTCCGCCCCGCCTGGCGGCGCGGCCACCGCCGCCTCAGCACTGAAGGC                  GCGCTGACGTCACTCGCCGGTCCCCCGCAAAC'TCCC'TTCCCGGCCACCTTGGT                  CGCGTCCGCGCCGCCCGGCCAGCCGGACCGCACACGCGAGGCGCGAGA                  TAGGGGGGCACGGGCGCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCT</p>	<p>632</p>

[0796]

	GCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCGTGCCTGAGAGCGCAG	
SP0032 (SEQ ID NO: 7)	TTCAGAGGTCGGATCTGAATCCAGCTCCAAGGCCCCAGCACCCAAGCCCTGAC CCCATGGCTGCCTGCTGGCTGGGAGTGGCATTCTTTAGAACAGGAAAGGCAGT GAGTTCTCATTGCATCAATACTTGCATTTGCTACAACAGAAGCTTTTTGTCTCT TCCTCACATTCCCTTLAGAACACAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCA CTGGAGTCAGCCTCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGG AGAAAACAATGTTGCTGCCAGGCCTTCTGGAATGACCCCTTCCCACCCCT GACCAGCCCAGCAAGGCCCGGGGCTGGCTGCCTAGTTGATTGGAGCCAAGA GTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGGGCTAGGCCTGCGCA CCCACCCACCGACCCCTCACCCACCGACCCGTCACCCACCGACCAAGGGGCAC CCTGGCCTAGAGGGGATGCTGAGCGGGACCCGCTCCTGCCTCTGGCAGTCCC AGATGGGACTTGGACCCCGCAGTTGCTCTCTCGGACCCTAAGTTTCTACCCCTG GATCTAAGGCGGAGCTGGGTTTGGCGATCCCACGGTTCCCGGCGGGGCGGGG CCCGGTCGCCCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAGCAGTGCATTGTG GGAAACTCCCGA	701
SP0033 (SEQ ID NO: 8)	TTCAGAGGTCGGATCTGAATCCAGCTCCAAGGCCCCAGCACCCAAGCC CTGACCCCATGGCTGCCTGCTGGCTGGGAGTGGCATTCTTTAGAACAG GAAAGGCAGTGAGTTCTCATTGCATCAATACTTGCATTTGCTACAACA GAAGCTTTTTGTCTCTTCTCACATTCCCTTLAGAACACAAGCCTCCTTTT CTGCCTGTGGCCATCGGTCACTGGAGTCAGCCTCGTGAGTGGCTTGGT GGCGGATGGCACCGTCCTTTGTGGGAGAAAACAATGTTGCTGCCAGG CCTTTCTGGAATGACCCCTTCCCACCCCTGACCAGCCCAGCAAGGC CCGGGGCTGGCTGCCTAGTTGATTGGAGCCAAGAGTTTGTGAATGGA TGAAGGGAGAAGGGACATCCAATGCGGCGAGGCGCGTGCGCACTGCC AGCTTACGACCCGCGACAGTGCCTTCGCCCCCGCCTGGCGGCGCGCG CCACCGCCGCTCAGCACTGAAGGCGCGCTGACGTCACCTGCCGGTCC CCCGAAACTCCCCTTCCCGGCCACCTTGGTCGCGTCCGCGCCGCGCC GGCCCAGCCGGACCGCACCCACGCGAGGCGCGAGATAGGGGGGCACGG	716

[0797]

	GCGCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTC TGCGGTGGGCAGCGGAGGAGTCGTGTTCGTGCCTGAGAGCGCAG	
SP0019 (SEQ ID NO: 9)	TCAGGGGTGCAGCTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTG TCGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAG GCCCCAGCACCCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGT GGCATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTGCATCAATAC TTGCATTTGCTACAACAGAAGCTTTTTGTCTCTTCTCACATTCCTTTAG AACACAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCACTGGAGTCAGCC TCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAAC AATGTTGCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGAC CAGCCCCAGCAAGGCCCGGGCTGGCTGCCTAGTTGATTGGAGCCAAG AGTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGTACTGAAG AAGATCTTAACAGAAGGGCTAGGCCTGCGCACCCACCCACCGACCCCT CACCCACCGACCCGTCACCCACCGACCAAGGGGCACCCTGGCCTAGAG GGGATGCTGAGCGGGACCCGCTCCTGCCTCTGGCAGTCCAGATGGG ACTTGGACCCCGCAGTTGCTCTCTCGGACCCTAAGTTTCTACCCCTGGA TCTAAGGCGGAGCTGGGTTTGC GGATCCCACGGTTCCCGGCGGGGCGG GGCCCGGTCGCCCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAGCAGTG CATTGTGGGAAACTCCCGA	792
SP0020 (SEQ ID NO: 10)	CTTCTCTTATATTTACCAAGACTCAGTTCCTGAGCAAGAAACCACAG GCACAGCAAGTGCCATGAAAAGCGGCTTTGTGTGGGGTGGGCTCTTCA CACTCCAATCTCCACTTCTTCTCAAGGCCTCAAAAAAAGTTGAAAAA TGAAAACAAAAGCCCTGCTGTGTTGAGCTGGGCTCTGGCGTTGCCATG GACCCAGGGCAAACAGCGGTGCTCCTGCTCTGCCCCGGCTCAGCTCA TGCTGGGCCTGCACTTCTGGAAGGGAGCATGGACTTTGGAATGACTGG TTAGAACCCAAATGAATTAATGGAATTTGACATAGTTCAAAAATAATA AAATGTGATACCCATGAAATGCTGATATTCTGCCTTAATTTGCCAGATT GGGGGCCGGGCTAGGCCTGCGCACCCACCCACCGACCCCTCACCCACC	686

[0798]

	<p>GACCCGTCACCCACCGACCAAGGGGCACCCTGGCCTAGAGGGGATGCT                  GAGCGGGACCCGCCTCCTGCCTCTGGCAGTCCCAGATGGGACTTGGAC                  CCCGCAGTTGCTCTCTCGGACCCTAAGTTTCTACCCCTGGATCTAAGGC                  GGAGCTGGGTTTGC GGATCCCACGGTTC CCGGCGGGGCGGGGCCCGGT                  CGCCCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAGCAGTGCATTGTGG                  GAAACTCCCGA</p>	
<p>SP0021                  (SEQ                  ID NO:                  11)</p>	<p>TCAGGGGTGCAGCTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTG                  TCGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAG                  GCCCCAGCACCCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGT                  GGCATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTGCATCAATAC                  TTGCATTTGCTACAACAGAAGCTTTTTGTCTCTTCTCACATTCTTTAG                  AACACAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCACTGGAGTCAGCC                  TCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAAC                  AATGTTGCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGAC                  CAGCCCCAGCAAGGCCCGGGGCTGGCTGCCTAGTTGATTGGAGCCAAG                  AGTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGTACTGAAG                  AAGATCTTAACAGAATGCGGCGAGGCGCGTGCGCACTGCCAGCTTCAG                  CACCGCGGACAGTGCCTTCGCCCCCGCTGGCGGCGCGCGCCACCGCC                  GCCTCAGCACTGAAGGCGCGCTGACGTCACCTCGCCGGTCCCCGCAAA                  CTCCCCTTCCCGCCACCTTGGTCGCGTCCGCGCCGCCGCCGCCCAGC                  CGGACCGCACACGCGAGGCGGAGATAGGGGGGCACGGGCGCGACC                  ATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTTGG                  GCAGCGGAGGAGTCGTGTCGTGCCTGAGAGCGCAG</p>	<p>807</p>
<p>SP0022                  (SEQ                  ID NO:                  12)</p>	<p>CTTCCTTATATTTACCAAGACTCAGTTCCTGAGCAAGAAACCACAG                  GCACAGCAAGTGCCATGAAAAGCGGCTTTGTGTGGGGTGGGCTCTTCA                  CACTCCAATCTCCACTTCCTTCTCAAGGCCTCAAAAAAAGTTGAAAAA                  TGAAAACAAAAGCCCTGCTGTGTTGAGCTGGGCTCTGGCGTTGCCATG                  GACCCAGGGCAAACAGCGGTGCTCCTGCTCTGCCCCGGCTCAGCTCA</p>	<p>701</p>

[0799]

	<p>TGCTGGGCCTGCACTTCTGGAAGGGAGCATGGACTTTGGAATGACTGG  TTAGAACCCAAATGAATTAATGGAATTTGACATAGTTCAAAAATAATA  AAATGTGATACCCATGAAATGCTGATATTCTGCCTTAATTTGCCAGATT  GGGGGCCTGCGGCGAGGCGCGTGCGCACTGCCAGCTTCAGCACCGCG  GACAGTGCCTTCGCCCCCGCTGGCGGCGCGGCCACCGCCGCTCAG  CACTGAAGGCGCGCTGACGTCACTCGCCGGTCCCCGCAAACCTCCCCT  TCCCGGCCACCTTGGTCGCGTCCGCGCCGCCGCCGCCCAGCCGGACC  GCACCACGCGAGGCGGAGATAGGGGGGCACGGGCGCGACCATCTGC  GCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGGTGGGCAGCG  GAGGAGTCGTGTCGTGCCTGAGAGCGCAG</p>	
<p>SP0028  (SEQ  ID NO:  13)</p>	<p>TCAGGGGTGCAGCTTTTTTCTGTCTTTACTCAGCCTGAGAAAGGTTG  TCGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAG  GCCCCAGCACCCAAGCCCTGACCCATGGCTGCCTGCTGGCTGGGAGT  GGCATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTGCATCAATAC  TTGCATTTGCTACAACAGAAGCTTTTTGTCTCTTCCTCACATTCCTTAG  AACACAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCACTGGAGTCAGCC  TCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAAC  AATGTTGCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGAC  CAGCCCCAGCAAGGCCCGGGGCTGGCTGCCTAGTTGATTGGAGCCAAG  AGTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGTACTGAAG  AAGATCTTAACAGAAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCC  TGGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTC  CGGAAGTGAGCTGGGGTTGCTGGCAGCCTGGCTGGCACTGGGCTAGGC  CTGCGCACCCACCCACCGACCCCTCACCCACCGACCCGTCACCCACCG  ACCAAGGGGCACCCTGGCCTAGAGGGGATGCTGAGCGGGACCCGCT  CCTGCCTCTGGCAGTCCCAGATGGGACTTGGACCCCGCAGTTGCTCTCT  CGGACCCTAAGTTTCTACCCCTGGATCTAAGGCGGAGCTGGGTTTGCG  GATCCCACGGTCCCAGCGGGGCGGGGCCCGGTCGCCCTCCCCCTCC</p>	<p>913</p>

[0800]

	CCGCCCTCCTGCGCCGGGAGCAGTGCATTGTGGGAAACTCCCGA	
SP0029 (SEQ ID NO: 14)	TCAGGGGTGCAGCTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTG TCGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAG GCCCCAGCACCCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGT GGCATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTGCATCAATAC TTGCATTTGCTACAACAGAAGCTTTTTGTCTCTTCTCACATTCCTTTAG AACACAAGCTCCTTTTCTGCCTGTGGCCATCGGTCACTGGAGTCAGCC TCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAAC AATGTTGCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGAC CAGCCCCAGCAAGGCCCGGGCTGGCTGCCTAGTTGATTGGAGCCAAG AGTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGTACTGAAG AAGATCTTAACAGAAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCC TGGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTC CGGAAGTGAGCTGGGGTTGCTGGCAGCCTGGCTGGCACTTGC GGCGAG GCGCGTGCGCACTGCCAGCTCAGCACCGCGGACAGTGCCTTCGCCCC CGCCTGGCGGGCGCGCCACCGCCGCTCAGCACTGAAGGCGCGCTGA CGTCACTCGCCGGTCCCCCGAAACTCCCCCTCCCGGCCACCTTGGTCG CGTCCGCGCCGCCGCCGGCCCAGCCGGACCGCACACGCGAGGCGCG AGATAGGGGGGCACGGGCGCGACCATCTGCGCTGCGGCGCCGGCGAC TCAGCGCTGCCTCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCGTGC CTGAGAGCGCAG	928
SP0011 (SEQ ID NO: 33)	CTAGCCACAGGAAATGTCTGTCTATATCCAGGCAAGTACCTTGCTCAT TGGACCAACCCGAAACTGTTTCAGGGAAGATCAGGGAAATCAACTCAGT TACAAATGGGATAATCATGCCAGTAAAAACTACCTGTGGTGAATAAA GAGTTAACCCCTGTTCCATCTTAGGTCACTATGCAGAGTACCAATGAGT ACAAGAGATGGTGCCAAAGAGGGTGGCCCCCTCCCTAGCTGGGAACAG TCAACCCTTAGGAACTAGACTGTCAACACATCAGCCAGCCAGAGACAA GGGAAACCGTGGCAACCAAGTGTGCTGGCACATTGTGAGGTGGTGAT	529

[0801]

	GGGAACTGCAGAGGCCCTGCACAGCATGCTAATGAGCCCAGGCAAAC ATGAGCTCTCCCATAGCTGGGCTGCGGCCAACCCACCCCCTCAGG CTATGCCAGGGGGTGTGTCAGGGGCACCCGGGCATCGCCAGTCTAGC CCACTCCTTCATAAAGCCCTCGCATCCCAGGAGCGAGCAGAGCCAGAG C	
SP0034 (SEQ ID NO: 34)	TCCAAAGAAAAGCCAGATAAGTAGCTGATTATTGCATAGAGCTGACAG TATCACAGGAAGATCAGTAGTAGCAGCTCAAGTACAAAAAGTTAATT AGCAATACTTAATAAGAAAACTACCTCTGGCAGGTGAAGAGTTAATC CCTGGTCAATTTTAAGCTACTCTGCTGAGAGTACTAATAAGTGTAGGG GTTGGAGCCAATGAGGGTGACCCCTTCCTTGATGGGAACAGTCATCCC TTAGGAACTGCCCTGGAAAGCATCAGCCAGCCAGAAAACAGGGAAAG AGGCTGAGAAACCGTGGTAACCAAGTTTTGCTGGCACTTTGTAAAATG GTAAGTCAACTGCCGAGGCTGTGCAGAGAATGCTAATAAGCCTAGGA CAACCTGTAAAGAGTGGACCTAGAAAATGTCCACCCGCTAGAGAGAG GGAGCGAGCATGTGCGATGAGCAATAGCTGTGGACCTTACAGTTGCTG CTAACTGCCCTGGTGTGTGTGAGGGAGAGAGAGGGAGGGAGGGAGAG AGAGCGCGCTAGCGGAGAGAGCGAGTGAGCAAGCGAGCAGAAAAG AGGTGGAGAGGGGGGAATAAGAAAGAGAGAGAAGGAAAGGAGAGA AGGCAGGAAGAAGGCAAGGGACGAGACAA	646
SP0035 (SEQ ID NO: 35)	CTAGCCCACAGGAAATGTCTGTCTATATCCAGGCAAGTACCTTGCTCAT TGGACCAACCCGAAACTGTTTCAGGGAAGATCAGGGAAATCAACTCAGT TACAAATGGGATAATCATGCCAGTAAAACTACCTGTGGTGAATAAA GAGTTAACCCCTGTTCCATCTTAGGTCCTATGCAGAGTACCAATGAGT ACAAGAGATGGTGCCAAAGAGGGTGGCCCCCTCCCTAGCTGGGAACAG TCAACCCTTAGGAACTAGACTGTCAACACATCAGCCAGCCAGAGACAA GGGAAACCGTGGCAACCAAGTGTGCTGGCACATTGTGAGGTGGTGTGAT GGGAACTGCAGAGGCCCTGCACAGCATGCTAATGAGCCCAGGCAAAC ATCGCTAGAGAGAGGGAGCGAGCATGTGCGATGAGCAATAGCTGTGG	614

[0802]

	ACCTTACAGTTGCTGCTAACTGCCCTGGTGTGTGTGAGGGAGAGAGAG GGAGGGAGGGAGAGAGAGCGCGCTAGCGCGAGAGAGCGAGTGAGCA AGCGAGCAGAAAAGAGGTGGAGAGGGGGGAATAAGAAAGAGAGAG AAGGAAAGGAGAGAAGGCAGGAAGAAGGCAAGGGACGAGACAA	
SP0036 (SEQ ID NO: 36)	CTAGCCACAGGAAATGTCTGTCTATATCCAGGCAAGTACCTTGCTCAT TGGACCAACCCGAAACTGTTCAGGGAAGATCAGGGAAATCAACTCAGT TACAAATGGGATAATCATGCCAGTAAAAACTACCTGTGGTGAATAAA GAGTTAACCCCTGTTCCATCTTAGGTCACTATGCAGAGTACCAATGAGT ACAAGAGATGGTGCCAAAGAGGGTGGCCCCCTCCCTAGCTGGGAACAG TCAACCCTTAGGAACTAGACTGTCAACACATCAGCCAGCCAGAGACAA GGGAAACCGTGGCAACCAAGTGTGCTGGCACATTGTGAGGTGGTGTAT GGGAACTGCAGAGGCCCTGCACAGCATGCTAATGAGCCAGGCAAAC ATTCGAGTTGGCTGGACAAGGTTATGAGCATCCGTGTACTTATGGGGT TGCCAGCTTGGTCCTGGATCGCCCGGGCCCTTCCCCACCCGTTCCGGTT CCCCACCACCACCCGCGCTCGTACGTGCGTCTCCGCTGCAGCTCTTGA CTCATCGGGGCCCCCGGGTCACATGCGCTCGCTCGGCTCTATAGGCG CCGCCCCCTGCCACCCCCCGCCGCGCTGGGAGCCGCAGCCGCCGCC ACTCCTGCTCTCTGCGCCGCCCGTCAACCACGCCACCGCCACCGG CTGAGTCTGCAGTCCTCGAG	695

[0803] 表2-来自表1的启动子的CRM

[0804]

名称	序列
CRM_SP0026 (SEQ ID NO: 15)	AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGCCT GAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCAAG CAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTTTCATGCCAG TGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCCTGCT TCCCGCCGACCCCAGCCCCCTCAAATGCCTTCCGAGAAGCCATTGAGCAGG GGGCTTGCATTGCACCCCAGCCTGACAGCCTGGCATCTTGGGATAAAAGCAGCA

[0805]

	<p>CAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGAGAACAAG  GCTCTATTCAGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCAGGCATGGA  CAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGAAGTCCAAGG  ACACAAATGGGTGAGGGGAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCCT  GGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTCCGGAAGT  GAGCTGGGGTTGCTGGCAGCCTGGCTGGCACT</p>
<p>CRM_SP0027  (SEQ ID NO:  16)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGCCT  GAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCAAG  CAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTTTCATGCCAG  TGAATGACTCACCTTGGCACAGACACAATGTTCCGGGTGGGCACAGTGCCTGCT  TCCC GCCG CACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCATTGAGCAGG  GGGCTTGCAATGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAAAGCAGCA  CAGCCCCCTAGGGGCTGCCCTTGCTGTGTGGCGCCACCGGCGGTGGAGAACAAG  GCTCTATTCAGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCAGGCATGGA  CAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGAAGTCCAAGG  ACACAAATGGGTGAGGGGAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCCT  GGGCGCGTCACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTCCGGAAGT  GAGCTGGGGTTGCTGGCAGCCTGGCTGGCACT</p>
<p>CRM_SP0028  (SEQ ID NO:  17)</p>	<p>TCAGGGGTGCAGCTTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTGT  CGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAGGC  CCCAGCACCCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGTGGC  ATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTCATCAATACTTGC  ATTTGCTACAACAGAAGCTTTTTGTCTCTTCCTCACATTCCTTTAGAACA  CAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCAGTGGAGTCAGCCTCGTG  AGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAACAATGTT  GCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGACCAGCCCC  AGCAAGGCCCGGGCTGGCTGCCTAGTTGATTGGAGCCAAGAGTTTGCT  GAATGGATGAAGGGAGAAGGGACATCCAAGTGAAGGAAGATCTTA</p>

	<p>ACAGAAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCCTGGGCGCGT CACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTCCGGAAGTGAG CTGGGGTTGCTGGCAGCCTGGCTGGCACT</p>
<p>[0806]</p>	<p>CRM_SP0029 (SEQ ID NO: 18)</p> <p>TCAGGGGTGCAGCTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTGT CGTTTGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAGGC CCCAGCACCCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGTGGC ATTCTTTAGAACAGGAAAGGCAGTGAGTTCTCATTGCATCAATACTTGC ATTTGCTACAACAGAAGCTTTTTGTCTCTTCCTCACATTCCTTTAGAACA CAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCCTGAGGAGTCAGCCTCGTG AGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGGAGAAAACAATGTT GCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCTGACCAGCCCC AGCAAGGCCCGGGGCTGGCTGCCTAGTTGATTGGAGCCAAGAGTTTGCT GAATGGATGAAGGGAGAAGGGACATCCAAGTGAAGGAGATCTTA ACAGAAAGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCCTGGGCGCGT CACAGCCGCGCTATTCTCAGCGTCTCTCCTTTTATGGCTCCGGAAGTGAG CTGGGGTTGCTGGCAGCCTGGCTGGCACT</p>

[0807] 表3-表1的启动子中包含的顺式调节元件

名称	序列
<p>[0808]</p> <p>CRE0005_GFAP (SYNP_AST_CR E0005_GFAP) (SEQ ID NO: 19)</p>	<p>AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTTTCATGCC CAGTGAATGACTCACCTTGGCACAGACACAATGTTCCGGGTGGGCACAGTGCC TGCTTCCCGCCGACCCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA GCAGGGGGCTTGCAATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA GCAGCACAGCCCCCTAGGGGCTGCCCTTGTGTGTGGCGCCACCGCGGTGGA GAACAAGGCTCTATTACAGCCTGTGCCAGGAAAGGGGATCAGGGGATGCCCA</p>

[0809]

	GGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCTGTCTGCTTCCCAGA AGTCCAAGGACACAAATGGGTGAGGGGA
CRE0007_GFAP (SYNP_AST_CR E0007_GFAP) (SEQ ID NO: 20)	AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC TGCTTCCCGCCGACCCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA GCAGGGGGCTTGCAATGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA GCAGCAC
CRE0012_Arc (SYNP_NEU_C RE0012_Arc) (SEQ ID NO: 21)	AGCGCGCAGAGTCTGCATGCGTGAGGAAGCTCCTGGGCGCGTCACAGCCGCG CTATTCTCAGCGTCTCTCCTTTTATGGCTCCGGAAGTGAGCTGGGGTTGCTGGC AGCCTGGCTGGCACT
CRE0001_S100B (SYNP_AST_CR E0001_S100B) (SEQ ID NO: 22)	CTTCCTTTATATTTACCAAGACTCAGTTCCTGAGCAAGAAACCACAGGCAC AGCAAGTGCCATGAAAAGCGGCTTTGTGTGGGGTGGGCTCTTCACACTCCAAT CTCCACTTCCTTCTCAAGGCCTCAAAAAAAGTTGAAAAATGAAAACAAAAGCC CTGCTGTGTGAGCTGGGCTCTGGCGTTGCCATGGACCCAGGGCAAACAGCGG TGCTCCTGCTCTGCCCCGGCTCAGCTCATGCTGGGCCTGCACTTCTGGAAGGG AGCATGGACTTTGGAATGACTGGTTAGAACCCAAATGAATTAATGGAATTGA CATAGTTCAAAAATAATAAAATGTGATAACCCATGAAATGCTGATATTCTGCCT TAATTTGCCAGATTGGGGGCC
CRE0009_S100B (SYNP_AST_CR E0009_S100B) (SEQ ID NO: 23)	TTCAGAGGTCGGATCTGAATCCAGCTCCAAGGCCCCAGCACCCAAGCCCTGAC CCCATGGCTGCCTGCTGGCTGGGAGTGGCATTCTTTAGAACAGGAAAGGCAGT GAGTTCTCATTGCATCAATACTTGCATTTGCTACAACAGAAGCTTTTTGTCTCT TCCTCACATTCCTTTAGAACACAAGCCTCCTTTTCTGCCTGTGGCCATCGGTCA CTGGAGTCAGCCTCGTGAGTGGCTTGGTGGCGGATGGCACCGTCCTTTGTGGG

	<p>AGAAAACAATGTTGCTGCCAGGCCTTTCTGGAATGACCCCTTCCCACCCCT                  GACCAGCCCCAGCAAGGCCCGGGCTGGCTGCCTAGTTGATTGGAGCCAAGA                  GTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAA</p>
	<p>TCAGGGGTGCAGCTTTTTTCTGTCTTTTACTCAGCCTGAGAAAGGTTGTCGTT                  TGACAAGGTTTGTTCAGAGGTCGGATCTGAATCCAGCTCCAAGGCCCCAGCAC                  CCAAGCCCTGACCCCATGGCTGCCTGCTGGCTGGGAGTGGCATTCTTTAGAAC                  AGGAAAGGCAGTGAGTTCTCATTGCATCAATACTTGCATTTGCTACAACAGAA                  GCTTTTTGTCTCTTCCCTCACATTCCTTTAGAACACAAGCCTCCTTTCTGCCTGT                  GGCCATCGGTCACTGGAGTCAGCCTCGTGAGTGGCTTGGTGGCGGATGGCACC                  GTCCTTTGTGGGAGAAAACAATGTTGCTGCCAGGCCTTTCTGGAATGACCC                  TCCCACCCCTGACCAGCCCCAGCAAGGCCCGGGCTGGCTGCCTAGTTGAT                  TGGAGCCAAGAGTTTGCTGAATGGATGAAGGGAGAAGGGACATCCAAGTGAC                  TGAAGAAGATCTTAACAGAA</p>
[0810]	<p>CRE0002_S100B                  (SYNP_AST_CR                  E0002_S100B)                  (SEQ ID NO: 24)</p> <p>CTAGCCACAGGAAATGTCTGTCTATATCCAGGCAAGTACCTTGCTCATTGGA                  CCAACCCGAAACTGTTCAAGGAAGATCAGGGAAATCAACTCAGTTACAAATG                  GGATAATCATGCCAGTAAAACTACCTGTGGTGAATAAAGAGTTAACCCCTG                  TTCCATCTTAGGTCATATGCAGAGTACCAATGAGTACAAGAGATGGTGCCAA                  AGAGGGTGGCCCTCCCTAGCTGGGAACAGTCAACCCTTAGGAACTAGACTGT                  CAACACATCAGCCAGCCAGAGACAAGGGAAACCGTGGCAACCAAGTGTGCT                  GGCACATTGTGAGGTGGTGTATGGGAACTGCAGAGGCCCTGCACAGCATGCTA                  ATGAGCCCAGGCAAACAT</p>
	<p>CRE0006_AQP4                  (SYNP_AST_CR                  E0006_AQP4)                  (SEQ ID NO: 37)</p> <p>TCCAAAGAAAAGCCAGATAAGTAGCTGATTATTGCATAGAGCTGACAGTATCA                  CAGGAAGATCAGTAGTAGCAGCTCAAGTACAAAAAGGTTAATTAGCAATACT                  TAATAAGAAAACCTCTGGCAGGTGAAGAGTTAATCCCTGGTCAATTTTA                  AGCTACTCTGCTGAGAGTACTAATAAGTGTAGGGGTTGGAGCCAATGAGGGTG                  ACCCTTCCTTGATGGGAACAGTCATCCCTTAGGAACTGCCCTGGAAAGCATC                  AGCCAGCCAGAAAACAGGGAAAGAGGCTGAGAAAACCGTGGTAACCAAGTTTT</p>
	<p>CRE0008_AQP4                  (SYNP_AST_TO                  _CRE0008v2_A                  QP4)                  (SEQ ID NO: 38)</p> <p>GCTGGCACTTTGTAAAATGGTAACTGCAACTGCCGAGGCTGTGCAGAGAATGC                  TAATAAGCCTAGGACAACCTGTAAAGAGTGGACCTAGAAAATGTCCACC</p>
[0811]	

[0812] 表4-表1的启动子中包含的最小/近端启动子

名称	序列
CRE0003_CEND 1_mp (SYNP_NEU_C RE0003_CEND1 _mp) (SEQ ID NO: 25)	GGGCTAGGCCTGCGCACCCACCCACCGACCCCTCACCCACCGACCCGTCACCCA CCGACCAAGGGGCACCCTGGCCTAGAGGGGATGCTGAGCGGGACCCGCCTCCT GCCTCTGGCAGTCCCAGATGGGACTTGGACCCCGCAGTTGCTCTCTCGGACCCT AAGTTTCTACCCCTGGATCTAAGGCGGAGCTGGGTTTGC GGATCCCACGGTTCC CGGCGGGGCGGGGCCCGGTCGCCCTCCCCCTCCCCGCCCTCCTGCGCCGGGAG CAGTGCATTGTGGGAAACTCCCGA
[0813] CRE0009_SYN1 _mp (SYNP_NEU_C RE0009_SYN1_ mp) (SEQ ID NO: 26)	TGC GGC GAG GCG CGT GCG CACT GCC AGCTT CAG CAC CGG ACAGT GCCTTCGC CCCC GCCTGGCGGCGCGGCCACCGCCGCCTCAGCACTGAAGGCGCGCTGACGT CACTCGCCGGTCCCCCGAAACTCCCCTTCCCGGCCACCTTGGTCGCGTCCGCGC CGCCGCCGGCCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGGCACG GGCGCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTCTGCGG TGGGCAGCGGAGGAGTCGTGTCGTGCTGAGAGCGCAG
CRE0015_GAP4 3_mp (SYNP_NEU_C RE0015_GAP43	CGCTAGAGAGAGGGAGCGAGCATGTGCGATGAGCAATAGCTGTGGACCTTACA GTTGCTGCTAACTGCCCTGGTGTGTGTGAGGGAGAGAGAGGGAGGGAGGGAGA GAGAGCGCGCTAGCGCGAGAGAGCGAGTGAGCAAGCGAGCAGAAAAGAGGTG GAGAGGGGGGAATAAGAAAGAGAGAGAAGGAAAGGAGAGAAGGCAGGAAG

[0814]

_mp) (SEQ ID NO: 39)	AAGGCAAGGGACGAGACAA
CRE0016_Eno2_ mp (SYNP_NEU_C RE0016_Eno2_m p) (SEQ ID NO: 40)	TCGAGTTGGCTGGACAAGGTTATGAGCATCCGTGTACTTATGGGGTTGCCAGCT TGGTCCTGGATCGCCCGGGCCCTCCCCACCCGTTTCGGTTCCCCACCACCACCC GCGCTCGTACGTGCGTCTCCGCCTGCAGCTCTTGACTCATCGGGGCCCCCGGGT CACATGCGCTCGCTCGGCTCTATAGGCGCCGCCCCCTGCCACCCCCGCCCCG GCTGGGAGCCGCAGCCGCCCACTCCTGCTCTCTCTGCGCCGCCCGCCTCACC ACCGCCACCGCCACCGGCTGAGTCTGCAGTCTCGAG
CRE0004_GFAP _mp (SYNP_AST_CR E0004_GFAP_m p) (SEQ ID NO: 41)	GAGCTCTCCCATAGCTGGGCTGCGGCCCAACCCACCCCTCAGGCTATGCCA GGGGGTGTTGCCAGGGGCACCCGGGCATCGCCAGTCTAGCCCACTCCTTCATAA AGCCCTCGCATCCCAGGAGCGAGCAGAGCCAGAGC

[0815] 表5-合成启动子概述

[0816]

启动子名称	最小/近端启动子	CRE	CRE
SP0013 (SYNP_NA_SP0013)	CRE0003_CEND1_mp	CRE0005_GFAP	
SP0014 (SYNP_NA_SP0014)	CRE0009_SYN1_mp	CRE0005_GFAP	
SP0026	CRE0003_CEND1_mp	CRE0012_Arc	CRE0005_GFAP

[0817]

SP0027	CRE0009_SYN1_mp	CRE0012_Arc	CRE0005_GFAP
SP0030 (SYNP_NA_SP0030)	CRE0003_CEND1_mp	CRE0007_GFAP	
SP0031 (SYNP_NA_SP0031)	CRE0009_SYN1_mp	CRE0007_GFAP	
SP0032 (SYNP_NA_SP0032)	CRE0003_CEND1_mp	CRE0009_S100B	
SP0033	CRE0009_SYN1_mp	CRE0009_S100B	
SP0019 (SYNP_NAO_SP0019)	CRE0003_CEND1_mp	CRE0002_S100B	
SP0020 (SYNP_NAO_SP0020)	CRE0003_CEND1_mp	CRE0001_S100B	
SP0021 (SYNP_NAO_SP0021)	CRE0009_SYN1_mp	CRE0002_S100B	
SP0022 (SYNP_NAO_SP0022)	CRE0009_SYN1_mp	CRE0001_S100B	
SP0028	CRE0003_CEND1_mp	CRE0012_Arc	CRE0002_S100B
SP0029	CRE0009_SYN1_mp	CRE0012_Arc	CRE0002_S100B

[0818]

SP0011 (SYNP_AST_SP0011)	CRE004_GFAP_mp	CRE0006_AQP4	
SP0034 (SYNP_NA_SP0034)	CRE0015_GAP43_mp	CRE0008_AQP4	
SP0035 (SYNP_NA_SP0035)	CRE0015_GAP43_mp	CRE0006_AQP4	
SP0036 (SYNP_NA_SP0036)	CRE0016_Eno2_mp	CRE0006_AQP4	

[0819] 表6-额外的CRE和启动子元件

[0820]

名称	序列
CRE0006_GFAP (SYNP_AST_CRE 0006_GFAP) (SEQ ID NO: 27)	AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA GCAGGGGGCTTGCATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA GCAGCACAGCCCCCTAGGGGCTGCCCTTGTGTGTGGCGCCA
CRE0008_GFAP (SYNP_AST_CR E0008_GFAP) (SEQ ID NO: 28)	AACATATCCTGGTGTGGAGTAGGGGACGCTGCTCTGACAGAGGCTCGGGGGC CTGAGCTGGCTCTGTGAGCTGGGGAGGAGGCAGACAGCCAGGCCTTGTCTGCA AGCAGACCTGGCAGCATTGGGCTGGCCGCCCCCAGGGCCTCCTCTTCATGCC CAGTGAATGACTCACCTTGGCACAGACACAATGTTTCGGGGTGGGCACAGTGCC TGCTTCCCGCCGCACCCAGCCCCCTCAAATGCCTTCCGAGAAGCCCATTGA

[0821]

	GCAGGGGGCTTGCATTGCACCCAGCCTGACAGCCTGGCATCTTGGGATAAAA GCACCGGCGGTGGAGAACAAGGCTCTATTCAGCCTGTGCCAGGAAAGGGGA TCAGGGGATGCCAGGCATGGACAGTGGGTGGCAGGGGGGAGAGGAGGGCT GTCTGCTTCCCAGAAGTCCAAGGACACAAATGGGTGAGGGGA
CRE0001_CEND 1_mp (SYNP_NEU_C RE0001_CEND1 _mp) (SEQ ID NO: 29)	GATCCCACGGTTCCCGGCGGGGCGGGGCCCGGTCGCCCCCTCCCCCTCCCCGCC CTCCTGCGCCGGGAGCAGTGCATTGTGGGAAACTC
CRE0002_CEND 1_mp (SYNP_NEU_C RE0002_CEND1 _mp) (SEQ ID NO: 30)	GGGCTAGGCCTGCGCACCCACCCACCGACCCCTCACCCACCGACCCGTCACCC ACCGACCAAGGGGCACCCTGGCCTAGAGGGGATGCTGAGCGGGACCCGCCTC CTGCCTCTGGCAGTCCCAGATGGGACTTGGACCCCGCAGTTGCTCTCTCGGAC CCTAAGTTTCTACCCCTGGATCTAAGGCGGAGCTGGGTTTGCGGATCCCACGG TCCCCGGCGGGGCGGGGCCCGGTCGCCCCCTCCCCCTCCCCGCCCTCCTGCGCC GGGAGCAGTGCATTGTGGGAAACTCCCGAGCTCTCCTCCGCGTTCGCAGCCGC CGTCATCCCCGCGGAGGAGCGCGCAGCCCC
CRE0008_SYN1 _mp (SYNP_NEU_C RE0008_SYN1_ mp) (SEQ ID NO: 31)	TGCAGAGGGCCCTGCGTATGAGTGCAAGTGGGTTTTAGGACCAGGATGAGGC GGGGTGGGGGTGCTACCTGACGACCGACCCCGACCCACTGGACAAGCACCC AACCCCATTCCCCAAATTGCGCATCCCCTATCAGAGAGGGGGAGGGGAAAC AGGATGCGGCGAGGCGCGTGCACACTGCCAGCTTCAGCACCGCGGACAGTGC CTTCGCCCCCGCTGGCGGCGCGGCCACCGCCGCTCAGCACTGAAGGCGCG CTGACGTCACCTGCGCGGTCCCCCGAAACTCCCCTTCCCGGCCACCTTGGTTCGC GTCCGCGCCGCCCGGCCAGCCGGACCGCACCACGCGAGGCGCGAGATAG GGGGCACGGGCGGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCC TCAGTCTGCGGTGGGCAGCGGAGGAGTCGTGTCGTGCTGAGAGCGCAG
CRE0013_SYN1	TGCGGCGAGGCGCGTGCACACTGCCAGCTTCAGCACCGCGGACAGTGCCTTCG

	_mp	CCCCCGCCTGGCGGCGCGCGCCACCGCCGCCTCAGCACTGAAGGCGCGCTGAC
(SYNP_NEU_C		GTCACTCGCCGGTCCCCCGCAAACCTCCCCTTCCCGGCCACCTTGGTCGCGTCCG
[0822] RE0013_SYN1_		CGCCGCCCGCCGGCCAGCCGGACCGCACCACGCGAGGCGCGAGATAGGGGGG
mp)		CACGGGCGCGACCATCTGCGCTGCGGCGCCGGCGACTCAGCGCTGCCTCAGTC
(SEQ ID NO: 32)		T

[0823] 实施例2

[0824] 根据本发明的合成的NS特异性启动子是通过查阅科学文献以确定在NS细胞中具有高度活性的基因及其各自的启动子而设计的。

[0825] 在设计这些启动子的过程中,考虑了已知NS特异性启动子的特殊缺点。首先,对NS细胞类型具有特异性的已知NS特异性启动子(例如Synapsin-1、CAMKIIa和GFAP)并不在整个细胞群中表达(例如不在所有神经元/星型胶质细胞中表达)。(Zhang等人,2019)已经证明了GFAP的情况,从Allen脑图谱的神经元中Syn-1的分布可以看出。其次,大多数已知的CRE、启动子元件和启动子都太大,无法包含在自补AAV载体(scAAV)中(根据转基因的大小,启动子的大小可能需要小于1000bp,优选小于900bp,更优选小于800bp,最优选小于700bp)。另外,可能需要在特定的细胞类型或整个NS的细胞类型的组合中表达,适当地在整个CNS或整个大脑中。

[0826] 目前已知的启动子不能解决这些缺点,在基因治疗中需要开发短的、细胞类型的NS特异性启动子,既要有针对性的局部表达,也要在整个NS中表达。例如,图1A和图1B所示的成年小鼠大脑中HTT(亨廷顿)和CYP46A1基因的表达模式突出了对在整个NS(例如整个大脑)中表达的要求。由于HTT(亨廷顿)基因在整个大脑中都有表达,因此抑制有缺陷的亨廷顿基因和/或抵消或减轻有缺陷的亨廷顿的有害影响的任何潜在表达产物在整个大脑中表达可能是有益的。同样,由于CYP46A1基因在整个大脑中表达,任何潜在的补充性CYP46A1在整个大脑中表达可能是有益的。

[0827] 在整个CNS的所有神经元以及星形胶质细胞和/或少突胶质细胞中的基因表达可能是治疗某些疾病如亨廷顿氏病所需要的。在星形胶质细胞和少突胶质细胞中的表达可能是有益的,因为神经胶质细胞与亨廷顿氏病有关(Shin等人,2005)。另外,在治疗方面,对于AAV-RNAi介导的降低HTT的治疗剂来拯救YAC128小鼠模型中的行为表型,发现星形胶质细胞转导是必需的(Stanek,Bu和Shihabuddin,2019)。

[0828] 因此,本发明着手设计串联的NS启动子,这些启动子在多种NS细胞类型中具有活性,同时解决上述的一些缺点。例如,启动子的设计涉及将一个或多个CRE与启动子元件结合在一起,以便与单独的CRE/启动子元件相比扩大细胞趋向性,从而创造出在多种NS细胞类型中具有活性的启动子,同时解决已知启动子不能在整个细胞群中表达的缺点。另外,为了解决已知的CRE、启动子元件和启动子过大而不能包含在AAV载体如自补AAV载体(scAAV)中的缺点,本发明的一些CRE和启动子元件已经用生物信息学分析、文献检索和公开的可以获得的基因组数据库进行了缩短,但仍期望成为有活性的CRE和启动子元件。例如,CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)和CRE0007\_GFAP(SEQ ID NO:20)已被缩短,但当与启动子元件可操作地连接时,预测其为赋予星形胶质细胞特异性

活性的活性调节序列。同样,CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)和CRE0002\_S100B(SEQ ID NO:24)已被缩短,但当与启动子元件可操作地连接时,预测其为赋予星形胶质细胞特异性和少突胶质细胞特异性活性的活性调节序列。根据本发明的NS特异性CRE优选被缩短,即CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0001\_S100B(SEQ ID NO:22)、CRE0009\_S100B(SEQ ID NO:23)和CRE0002\_S100B(SEQ ID NO:24)。由于其短的尺寸和预测的NS特异性活性,这些是特别优选的CRE。另外,CRE0003\_CEND1\_mp(SEQ ID NO:25)已被缩短,但预测其为神经元细胞中的活性调节序列。最后,CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0013\_SYN1\_mp(SEQ ID NO:32)已被缩短,但预测其为神经元细胞中的活性调节序列。根据本发明的启动子元件优选被缩短,即由于其短的尺寸和预测的神经元活性,CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0013\_SYN1\_mp(SEQ ID NO:32)是特别优选的启动子元件。

[0829] 根据本发明的合成的NS特异性启动子与编码CYP46A1转基因的核酸序列和人流感血凝素(HA)标签可操作地连接,并在野生型C57BL6/J小鼠中进行实验测试。与编码CYP46A1转基因的核酸序列和HA标签可操作地连接的根据本发明的合成的NS特异性启动子在病毒载体中静脉内施用。将通过qPCR分析来评估病毒转基因CYP46A1在大脑和脊髓组织切片中的载体拷贝数,并将其与内部基因组DNA拷贝数对照进行标准化,以确认等同的注射剂量。将进行蛋白质印迹来评估大脑和脊髓组织中HA标记的转基因的蛋白质表达。最后,将在大脑和脊髓组织切片上进行免疫荧光染色,以评估转基因在CNS细胞类型中的表达。同样,可以在PNS组织切片上进行免疫荧光染色,以评估转基因在PNS细胞类型中的表达。具体来说,将使用HA标签来标记CYP46A1基因的表达和神经元、星形胶质细胞、少突胶质细胞和小胶质细胞的标准标志物进行双重染色。

[0830] 预测SP0013(SEQ ID NO:1)在神经元和星形胶质细胞中具有活性。预测SP0014(SEQ ID NO:2)在神经元和星形胶质细胞中具有活性。预测SP0026(SEQ ID NO:3)在兴奋性神经元和星形胶质细胞中具有活性。预测SP0027(SEQ ID NO:4)在兴奋性神经元和星形胶质细胞中具有活性。预测SP0030(SEQ ID NO:5)在神经元和星形胶质细胞中具有活性。预测SP0031(SEQ ID NO:6)在神经元和星形胶质细胞中具有活性。预测SP0032(SEQ ID NO:7)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0033(SEQ ID NO:8)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0019(SEQ ID NO:9)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0020(SEQ ID NO:10)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0021(SEQ ID NO:11)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0022(SEQ ID NO:12)在神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0028(SEQ ID NO:13)在兴奋性神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0029(SEQ ID NO:14)在兴奋性神经元、星形胶质细胞和少突胶质细胞中具有活性。预测SP0011(SEQ ID NO:33)在神经元和星形胶质细胞中具有活性。预测SP0034(SEQ ID NO:34)在神经元和星形胶质细胞中具有活性。预测SP0035(SEQ ID NO:35)在神经元和星形胶质细胞中具有活性。预测SP0036(SEQ ID NO:36)在神经元和星形胶质细胞中具有活性。

[0831] 对RNA测序数据的生物信息学分析预测,与本发明的CRE和/或启动子元件相关的

一些基因(aqp4、cend1、eno2、gfap、s100B、syn1)在背根神经节和胫神经中表达。因此,与这些基因相关的CRE和/或启动子元件被预测为在PNS的细胞中表达。CRE0001\_S100B(SEQ ID NO:22)、CRE0002\_S100B(SEQ ID NO:24)、CRE0005\_GFAP(SEQ ID NO:19)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0009\_S100B(SEQ ID NO:23)、CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0006\_AQP4(SEQ ID NO:37)、CRE0008\_AQP4(SEQ ID NO:38)、CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0013\_SYN1\_mp(SEQ ID NO:32)、CRE0016\_Eno2\_mp(SEQ ID NO:40)和CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体被预测为在PNS的细胞中具有活性。对单细胞RNA测序数据的生物信息学分析预测,与本发明的CRE和/或启动子元件相关的一些基因(aqp4、cend1、eno2、gfap、s100B、syn1)在感觉神经元、PNS交感神经元和PNS肠道神经元中表达。因此,与这些基因相关的CRE和/或启动子元件被预测为在感觉神经元、PNS交感神经元和PNS肠道神经元中表达。CRE0001\_S100B(SEQ ID NO:22)、CRE0002\_S100B(SEQ ID NO:24)、CRE0005\_GFAP(SEQ ID NO:19)、CRE0007\_GFAP(SEQ ID NO:20)、CRE0009\_S100B(SEQ ID NO:23)、CRE0006\_GFAP(SEQ ID NO:27)、CRE0008\_GFAP(SEQ ID NO:28)、CRE0006\_AQP4(SEQ ID NO:37)、CRE0008\_AQP4(SEQ ID NO:38)、CRE0003\_CEND1\_mp(SEQ ID NO:25)、CRE0009\_SYN1\_mp(SEQ ID NO:26)、CRE0001\_CEND1\_mp(SEQ ID NO:29)、CRE0002\_CEND1\_mp(SEQ ID NO:30)、CRE0008\_SYN1\_mp(SEQ ID NO:31)、CRE0013\_SYN1\_mp(SEQ ID NO:32)、CRE0016\_Eno2\_mp(SEQ ID NO:40)和CRE0004\_GFAP\_mp(SEQ ID NO:41)或其功能变体被预测为在感觉神经元、PNS交感神经元和/或PNS肠道神经元中具有活性。

#### [0832] 实施例3

[0833] 为了评估根据本发明的CRE、CRM、启动子元件或合成的NS特异性启动子在NS中的活性,希望将合成的NS特异性启动子或包含根据本发明的CRE、CRM、启动子元件的合成的NS特异性启动子与报告基因(例如GFP)可操作地连接起来。将包含上述内容的表达构建体以病毒载体的形式静脉内施用,渗透到动物(例如小鼠或大鼠)的CNS。在一段时间后(例如1周、2周、3周、1个月、2个月、3个月),处死动物。将动物的CNS和PNS处理成组织切片,并且可以进行免疫染色(荧光和比色)以评估报告基因在整个CNS和PNS以及不同CNS和PNS细胞类型中的表达。如果需要定量,可以在CNS和PNS组织上进行报告子的蛋白质印迹。进行这些技术是常规和已知的。

#### [0834] 实施例4

[0835] 本文描述的是一种从Pro10/HEK293细胞制造病毒载体的方法,这些细胞已经被工程化为稳定地整合CYP46A1基因。

[0836] 美国专利号9,441,206中描述的稳定细胞系Pro10/HEK293对于AAV载体的可扩大生产是理想的。该细胞系可以与包含与根据SEQ ID NO:1-14、33-36中的任何一个的NS特异性启动子或其变体可操作地连接的CYP46A1基因的表达载体接触。使用本领域众所周知的方法选择CYP46A1整合到其基因组的克隆群体。

[0837] 用编码Rep2和血清型特异性Cap2的包装质粒AAV-Rep/Cap以及Ad-辅助质粒(XX680:编码腺病毒辅助序列)转染稳定地包含CYP46A1基因的Pro10/HEK293细胞。

[0838] 转染。在转染当天,用ViCell XR活力分析仪(Beckman Coulter)对细胞进行计数,

并稀释用于转染。为了混合转染混合物,将以下试剂按该顺序加入锥形管中:质粒DNA、OPTIMEM® I(Gibco)或OptiPro SFM(Gibco),或其他不含血清的兼容转染培养基,然后按质粒DNA的特定比例加入转染试剂。在室温下孵化之前,将混合物倒置以混合。将转染混合物用移液器移入烧瓶中,然后放回摇床/恒温箱中。所有的优化研究都是在30mL培养体积下进行的,然后在更大的培养体积下进行验证。转染后48小时收获细胞。

[0839] 使用Wave生物反应器系统生产rAAV。在转染前2天对波浪袋(wave bag)进行播种。波浪袋播种后两天,对细胞培养物进行计数,然后在转染前对细胞培养物进行扩大/稀释。然后对波浪生物反应器的细胞培养物进行转染。在转染后至少48小时,从波浪生物反应器袋中收获细胞培养物。

[0840] 滴度:AAV的滴度是在DNase消化后用qPCR对照标准曲线(AAV ITR特异性)和CYP46A1基因特异性引物计算的。

[0841] 从摇床瓶和60波浪生物反应器袋中收获悬浮细胞。转染后48小时,通过从摇床瓶中倒出或从波纹生物反应器袋中抽出,将细胞培养物收集到500mL聚丙烯锥形管(Corning)中。然后用Sorvall RC3C plus离心机和H6000A转子以655xg将细胞培养物离心10分钟。丢弃上清液,将细胞重新悬浮在1xPBS中,转移到50mL锥形管中,并在655xg下离心10分钟。在该点,沉淀物可以储存在NLT-60°C或通过纯化继续。

[0842] 使用qPCR测定细胞裂解液中的rAAV浓度。取出10mL的细胞培养物,用Sorvall RC3C plus离心机和H6000A转子以655xg离心10分钟。从沉淀物中倒出上清液。然后将细胞沉淀物重新悬浮在5mL的DNase缓冲液(5mM CaCl<sub>2</sub>, 5mM MgCl<sub>2</sub>, 50mM Tris-HCl pH 8.0)中,接着超声处理以有效裂解细胞。然后取出300μL,放入1.5mL微孔管中。然后向每个样品中加入140个单位的DNase I,并在37°C下孵育1小时。为了确定DNase消化的效果,将4-5mg CYP46A1质粒加入到添加和不添加DNase的非转染细胞裂解液中。向每个试管中加入50μL EDTA/Sarkosyl溶液(6.3% sarkosyl, 62.5mM EDTA pH 8.0),并在70°C下孵育20分钟。然后加入50μL蛋白酶K(10mg/mL),在55°C下孵育至少2小时。将样品煮沸15分钟,使蛋白酶K失活。从每个样品中取出等分试样,用qPCR进行分析。进行两个qPCR反应,以有效地确定每个细胞产生多少rAAV载体。一个qPCR反应使用一组被设计成与质粒XX680、pXR2和CYP46A1的骨架上的同源序列结合的引物进行。第二个qPCR反应使用一组结合和扩增CYP46A1基因内的一个区域的引物进行。qPCR使用罗氏的Sybr green试剂和Light cycler 480进行。样品在95°C下变性10分钟,然后进行45个循环(90°C 10秒, 62°C 10秒, 72°C 10秒)和熔解曲线(1个循环99°C 30秒, 65°C持续1分钟)。

[0843] 从粗制的裂解液中纯化rAAV。每个细胞沉淀物被调整到最终体积为10mL。将沉淀物进行短暂的涡旋,并以30%的产率在一秒开、一秒关的情况下超声处理4分钟。超声处理后,加入550U的DNase,并在37°C下孵育45分钟。然后使用Sorvall RCSB离心机和HS-4转子以9400xg离心沉淀物,以沉淀细胞碎片,并将澄清的裂解液转移到Type70Ti离心管(Beckman 361625)。关于收获和裂解HEK293细胞悬液以分离rAAV,本领域技术人员可以使用机械方法如微流控或化学方法如去污剂等,然后使用深度过滤或切向流过滤(TFF)进行澄清步骤。

[0844] AAV载体纯化。通过柱色谱方法纯化澄清的AAV裂解液,正如本领域技术人员所知道的,并在以下稿件中描述(Allay等人, Davidoff等人, Kaludov等人, Zolotukhin等人,

Zolotukin等人,等等),这些稿件通过引用整体并入本文。

[0845] 实施例5-启动子的体外测试

[0846] 在神经母细胞瘤来源的SH-SY5Y细胞中测试了根据本发明选择的NS特异性启动子。

[0847] 材料和方法

[0848] 细胞维持和转染

[0849] SH-SY5Y细胞在HAM F12培养基中培养,该培养基含有1mM L-谷氨酰胺(Gibco 11765-054)、15%热灭活的FBS(ThermoFisher 10500064)、1%非必需氨基酸(Merck M1745-100ML)和1%青霉素/链霉素(ThermoFisher 15140122)。细胞在1:3和1:4之间每周传代两次,以保持70-80%的健康细胞密度。细胞保持在代数20以下。为了进行转染,将细胞以 $10^5$ 个细胞/孔播种到粘附的48孔板中。播种后24小时,使用Lipofectamine3000试剂(ThermoFisher L3000008)将300ng质粒转染到细胞中。

[0850] 转染到SHSY5Y细胞系中的质粒包括与GFP可操作地连接的SP0013、SP0014、SP0030、SP0031、SP0032、SP0019、SP0020、SP0021、SP0033、SP0011、SP0034、SP0035或SP0036。

[0851] 流式细胞术

[0852] 转染后48小时,SH-SY5Y细胞用PBS清洗,然后用0.05%胰蛋白酶进行分离。收集细胞并重新悬浮在90% PBS、10% FBS溶液中。使用Attune Nxt声学聚焦细胞仪通过流式细胞术评估细胞的GFP表达。细胞活力染料7-AAD(ThermoFisher 00-6993-50)与对照细胞群混合,以识别和排除死细胞。使用带通滤波器510/10nm处的蓝色488nm激光测量活的单细胞群中GFP的表达。未转染的细胞被用来设置GFP阴性和GFP阳性细胞的闸门。用Attune Nxt软件计算GFP阳性单细胞的数量和所有GFP阳性细胞的GFP荧光中值。

[0853] 实验结果

[0854] 该实验的结果显示在图2A和2B中。通过流式细胞术评估用包含与GFP可操作地连接的SP0013、SP0014、SP0030、SP0031、SP0032、SP0019、SP0020、SP0021、SP0022、SP0011、SP0034、SP0035或SP0036的表达盒转染的神经母细胞来源的SH-SY5Y细胞的GFP表达中值和GFP阳性细胞的百分比。包含与GFP可操作地连接的已知启动子Synapsin-1和CAG的表达盒被用作对照。所有测试的启动子与神经元特异性对照启动子Synapsin-1具有相当的转染效率和GFP表达中值(见图2A和2B)。与对照启动子Synapsin-1和测试的合成的NS特异性启动子(图2A)相比,对照启动子CAG的转染效率高2至3倍(图2B),GFP表达中值高2.5左右。

[0855] 合成的NS特异性启动子SP0028(SEQ ID NO:13)与合成的NS特异性启动子SP0019(SEQ ID NO:9)的设计类似,因为两者都包括CRE0003\_CEND1\_mp(SEQ ID NO:25)和CRE0002\_S100B(SEQ ID NO:24)。除了这些元件外,SP0028(SEQ ID NO:13)还包括CRE0012\_Arc(SEQ ID NO:21)。因此,预期SP0028(SEQ ID NO:13)与SP0019(SEQ ID NO:9)的表现类似。

[0856] 合成的NS特异性启动子SP0029(SEQ ID NO:14)与合成的NS特异性启动子SP0021(SEQ ID NO:11)的设计类似,因为两者都包括CRE0009\_SYN1\_mp(SEQ ID NO:26)和CRE0002\_S100B(SEQ ID NO:24)。除了这些元件外,SP0029(SEQ ID NO:14)还包括CRE0012\_Arc(SEQ ID NO:21)。因此,预期SP0029(SEQ ID NO:14)与SP0021(SEQ ID NO:11)的表现类

似。

[0857] 合成的NS特异性启动子SP0026 (SEQ ID NO:3) 与合成的NS特异性启动子SP0013 (SEQ ID NO:1) 的设计类似,因为两者都包括CRE0003\_CEND1\_mp (SEQ ID NO:25) 和 CRE0005\_GFAP (SEQ ID NO:19)。除了这些元件外,SP0026 (SEQ ID NO:3) 还包括CRE0012\_Arc (SEQ ID NO:21)。因此,预期SP0026 (SEQ ID NO:3) 与SP0013 (SEQ ID NO:1) 的表现类似。

[0858] 合成的NS特异性启动子SP0027 (SEQ ID NO:4) 与合成的NS特异性启动子SP0014 (SEQ ID NO:2) 的设计类似,因为两者都包括CRE0009\_SYN1\_mp (SEQ ID NO:26) 和 CRE0005\_GFAP (SEQ ID NO:19)。除了这些元件外,SP0027 (SEQ ID NO:4) 还包括CRE0012\_Arc (SEQ ID NO:21)。因此,预期SP0027 (SEQ ID NO:4) 与SP0014 (SEQ ID NO:2) 的表现类似。

[0859] 合成的NS特异性启动子SP0033 (SEQ ID NO:8) 与SP0021 (SEQ ID NO:11) 的设计类似,因为两者都包括CRE0009\_SYN1\_mp (SEQ ID NO:26)。SP0033 (SEQ ID NO:8) 包括 CRE0009\_S100B (SEQ ID NO:23),它是CRE0002\_S100B (SEQ ID NO:24) 的较短版本,而后者又包含在SP0021 (SEQ ID NO:11) 中。因此,SP0033 (SEQ ID NO:8) 是SP0021 (SEQ ID NO:11) 的较短版本,因此,可以预期其表现类似。

[0860] 参考文献

[0861] Boussicault,L.et al. (2016) 'CYP46A1,the rate-limiting enzyme for cholesterol degradation,is neuroprotective in Huntington's disease',Brain,139 (3),pp.953-970.doi:10.1093/brain/awv384.

[0862] Djelti,F.et al. (2015) 'CYP46A1 inhibition,brain cholesterol accumulation and neurodegeneration pave the way for Alzheimer's disease', Brain,138(8),pp.2383-2398.doi:10.1093/brain/awv166.

[0863] Hammond,S.L.et al. (2017) 'Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection',PLoS ONE,12(12),pp.1-22.doi:10.1371/journal.pone.0188830.

[0864] Jakobsson,J.and Lundberg,C. (2006) 'Lentiviral vectors for use in the central nervous system',Molecular Therapy.The American Society of Gene Therapy,13(3),pp.484-493.doi:10.1016/j.ymthe.2005.11.012.

[0865] Kacher,R.et al. (2019) 'CYP46A1 gene therapy deciphers the role of brain cholesterol metabolism in Huntington's disease',Brain:a journal of neurology,142(8),pp.2432-2450.doi:10.1093/brain/awz174.

[0866] Schulte,J.and Littleton,J.T. (2011) 'The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology.', Current trends in neurology,5,pp.65-78.Available at:http://www.ncbi.nlm.nih.gov/pubmed/22180703%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3237673.

[0867] Shin,J.et al. (2005) 'Expression of mutant huntingtin in glial cells contributes to neuronal excitotoxicity',171(6),pp.1001-1012.doi:10.1083/jcb.200508072.Stanek,L.M.,Bu,J.and Shihabuddin,L.S. (2019) 'Neurobiology of

Disease Astrocyte transduction is required for rescue of behavioral phenotypes in the YAC128 mouse model with AAV-RNAi mediated HTT lowering therapeutics', *Neurobiology of Disease*. Elsevier, 129 (November 2018), pp. 29-37. doi:10.1016/j.nbd.2019.04.015.

[0868] Tanguy, Y. et al. (2015) 'Systemic AAVrh10 provides higher transgene expression than AAV9 in the brain and the spinal cord of neonatal mice', *Frontiers in Molecular Neuroscience*, 8 (JULY), pp. 1-10. doi:10.3389/fnmol.2015.00036.

[0869] Zhang, Z. et al. (2019) 'The Appropriate Marker for Astrocytes: Comparing the Distribution and Expression of Three Astrocytic Markers in Different Mouse Cerebral Regions', *BioMed Research International*, 2019. doi:10.1155/2019/9605265.

## 序列表

<110>	塞普洛麦克斯有限公司 阿斯克肋匹奥生物制药公司	
<120>	调节核酸序列	
<130>	P298164W0	
<150>	GB2013940.8	
<151>	2020-09-04	
<160>	41	
<170>	PatentIn version 3.5	
<210>	1	
<211>	795	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0013	
<400>	1	
	aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
	gctctgtgag ctggggagga ggacagacagc caggcettgt ctgcaagcag acctggcagc	120
	attgggctgg ccgccccca gggcctcctc ttcatgccc gtgaatgact caccttgga	180
	cagacacaat gttcgggtg ggacacagtgc ctgcttccc ccgaccccc gccccctca	240
	aatgccttcc gagaagccc ttgagcagg ggcttgatt gcaccccagc ctgacagcct	300
	ggcatccttg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
	cggcggtgga gaacaaggct ctattcagcc tgtgccagc aaaggggatc aggggatgcc	420
	caggcatgga cagtgggtg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
	aggacacaaa tgggtgaggg gagggctagg cctgcgacc caccaccga cccctcacc	540
	accgaccgt caccaccga ccaaggggca ccctggccta gaggggatgc tgagcgggac	600
	ccgctcctg cctctggcag tcccagatgg gacttgacc ccgagttgc tctctcgac	660
	cctaagtffc taccctgga tctaaggcgg agctgggttt gcggatcca cggttcccg	720
	cggggcgggg ccggtcgcc cctccccctc cccgcctcc tgcgccgga gcagtgcatt	780
	gtgggaaact cccga	795
<210>	2	
<211>	810	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0014	
<400>	2	
	aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60

gctctgtgag ctggggagga ggacagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctcctc ttcatgceca gtgaatgact caccttggca	180
cagacacaat gttcgggggtg ggacacagtgc ctgcttcccg ccgcacccca gccccctca	240
aatgccttcc gagaagccca ttgagcaggg ggcttgatt gcaccccagc ctgacagcct	300
ggcatcttgg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
cggcggtgga gaacaaggct ctattcagcc tgtgccagc aaaggggatc aggggatgcc	420
caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg gatgcggcga ggcgcgtgc cactgccagc ttcagaccg	540
cggacagtgc cttcgcccc gctggcggc gcgcgccacc gccgcctcag cactgaaggc	600
gcgctgacgt cactcgccgg tccccgcaa actcccctc ccggccacct tggtcgctc	660
cgcgccgccc ccggcccagc cggaccgac cacgcgaggg gcgagatagg ggggcacggg	720
cgcgaccatc tgcgctcgg cgccggcgac tcagcgtgc ctcaagtctgc ggtgggcagc	780
ggaggagtgc tgtcgtgcct gagagcgcag	810
<210> 3	
<211> 916	
<212> DNA	
<213> 人工序列	
<220>	
<223> SP0026	
<400> 3	
aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggacagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctcctc ttcatgceca gtgaatgact caccttggca	180
cagacacaat gttcgggggtg ggacacagtgc ctgcttcccg ccgcacccca gccccctca	240
aatgccttcc gagaagccca ttgagcaggg ggcttgatt gcaccccagc ctgacagcct	300
ggcatcttgg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
cggcggtgga gaacaaggct ctattcagcc tgtgccagc aaaggggatc aggggatgcc	420
caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg gaagcgcga gactctgcat gcgtgaggaa gctcctgggc	540
gcgtcacagc cgcgctatc tcagcgtctc tcttttatg gctccggaag tgagctgggg	600
ttgctggcag cctggctggc actgggctag gctgcgcac ccaccaccg acccctcacc	660
caccgaccg tcaccaccg accaagggg accctggcct agaggggatg ctgagcggga	720
cccgcctcct gcctctggca gtcccagatg ggacttgac ccgcagttg ctctctcgga	780
ccctaagttt ctaccctgg atctaaggcg gagctgggtt tgcggatccc acggttccc	840
gcggggcggg gcccggtgc cctccccct cccgcctc ctgcgccgg agcagtgc	900
tgtgggaaac tcccga	916
<210> 4	
<211> 931	
<212> DNA	

<213> 人工序列

<220>

<223> SP0027

<400> 4

aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttgga	180
cagacacaat gttcggggtg ggcacagtgc ctgcttccc cgcaccca gccccctca	240
aatgccttcc gagaagccc ttgagcagg ggcttgatt gcaccccagc ctgacagcct	300
ggcatccttg gataaaagca gcacagccc ctaggggctg cccttgctgt gtgggccac	360
cggcggtgga gaacaagget ctattcagcc tgtgccagg aaaggggatc aggggatgcc	420
caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg gaagcgcgca gagtctgcat gcgtgaggaa gctcctgggc	540
gcgtcacagc cgcgctattc tcagcgtctc tccttttatg gctccggaag tgagctgggg	600
ttgctggcag cctggctggc acttgcgggc aggcgcgtgc gactgccag cttcagcacc	660
gcggacagtg ccttcgcccc cgcctggcgg cgcgcgccac cgcgcctca gactgaagg	720
cgcgctgacg tcaactgcgg gtccccgca aactcccctt cccggccacc ttggtcgcgt	780
ccgcgccgcc gccggcccag ccggaccgca ccacgcgagg cgcgagatag gggggcacgg	840
gcgcgaccat ctgcgctgcg gcgccggcga ctacgcgctg cctcagtctg cggtgggcag	900
cggaggagtc gtgtcgtgcc tgagagcgca g	931

<210> 5

<211> 617

<212> DNA

<213> 人工序列

<220>

<223> SP0030

<400> 5

aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttgga	180
cagacacaat gttcggggtg ggcacagtgc ctgcttccc cgcaccca gccccctca	240
aatgccttcc gagaagccc ttgagcagg ggcttgatt gcaccccagc ctgacagcct	300
ggcatccttg gataaaagca gcacgggcta ggctgcgca cccaccacc gaccctcac	360
ccaccgacc gtcaccacc gaccaagggg caccctggcc tagaggggat gctgagcggg	420
accgcctcc tgctctggc agtcccagat gggacttga cccgcagtt gctctctcgg	480
accctaagtt tctaccctg gatctaaggc ggagctgggt ttgcggatcc cacggttccc	540
ggcggggcgg ggcccgtcg cccctcccc tccccccct cctgcgccg gagcagtga	600
ttgtgggaaa ctcccga	617

<210> 6

<211>	632	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0031	
<400>	6	
aacatatcct	ggtgtggagt	aggggacgct gctctgacag aggctcgggg gcctgagctg 60
gctctgtgag	ctggggagga	ggcagacagc caggcettgt ctgcaagcag acctggcagc 120
attgggctgg	ccgccccca	gggectctc ttcatgcca gtgaatgact caccttgca 180
cagacacaat	gttcggggtg	ggcacagtgc ctgcttccc cgcaccca gccccctca 240
aatgccttcc	gagaagccca	ttgagcaggg ggcttgcaatt gcacccagc ctgacagcct 300
ggcatcttgg	gataaaagca	gcaactgcggc gaggcgctg cgcactgcca gcttcagcac 360
cgcggacagt	gccttcgccc	ccgcttggcg gcgcgcgcca ccgccgctc agcactgaag 420
gcgcgctgac	gtcactcgcc	ggtccccgc aaactccct tcccggccac cttggtcgcg 480
tccgcgccgc	cgccggccca	gccggaccgc accacgcgag gcgcgagata ggggggcacg 540
ggcgcgacca	tctgcgctgc	ggcgccggcg actcagcgt gcctcagtct gcggtgggca 600
gcggaggagt	cgtgtcgtgc	ctgagagcgc ag 632
<210>	7	
<211>	701	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0032	
<400>	7	
ttcagaggtc	ggatctgaat	ccagctccaa ggccccagca cccaagcct gacccatgg 60
ctgcctgctg	gctgggagtg	gcattcttta gaacaggaaa ggcaatgagt tctcattgca 120
tcaataacttg	catttgctac	aacagaagct tttgtctct tcctcacatt ctttagaac 180
acaagcctcc	ttttctgctt	gtggccatcg gtcactggag tcagcctcgt gagggttg 240
gtggcggatg	gcaccgtcct	ttgtgggaga aaacaatggt gctgcccagg cttttctgga 300
atgaccctt	cccacccct	gaccagccc agcaaggccc ggggctggct gcctagttaga 360
ttggagccaa	gagtttgctg	aatggatgaa gggagaaggg acatccaagg gctaggcctg 420
cgcacccacc	caccgaccc	tcaccaccg acccgteacc caccgacca ggggcacct 480
ggcctagagg	ggatgctgag	cgggaccgc ctctgcctc tggcagtccc agatgggact 540
tggacccgc	agttgctctc	tcggacceta agtttctacc cctggatcta aggcggagct 600
gggtttgcgg	atcccacggt	tcccggcggg gcggggcccg gtcgcccctc ccctccccg 660
ccctcctgcg	ccgggagcag	tgcattgtgg gaaactccc a 701
<210>	8	
<211>	716	
<212>	DNA	

<213> 人工序列

<220>

<223> SP0033

<400> 8

ttcagaggtc	ggatctgaat	ccagctccaa	ggccccagca	ccaagccct	gaccccatgg	60
ctgcctgctg	gctgggagt	gcattcttta	gaacaggaaa	ggcagtgagt	tctcattgca	120
tcaatacttg	catttgctac	aacagaagct	tttgtctct	tcctcacatt	cctttagaac	180
acaagcctcc	ttttctgct	gtggccatcg	gtcactggag	tcagcctcgt	gagtggcttg	240
gtggcggatg	gcaccgtcct	ttgtgggaga	aaacaatgtt	gctgcccagg	cctttctgga	300
atgaccctt	cccaccctt	gaccagcccc	agcaaggccc	ggggctggct	gcctagttag	360
ttggagccaa	gagtttctg	aatggatgaa	gggagaaggg	acatccaatg	cggcgaggcg	420
cgtgcgcact	gccagcttca	gcaccgcgga	cagtgccttc	gccccgcct	ggcggcgcgc	480
gccaccgccg	cctcagcact	gaaggcgcgc	tgacgtcact	cgccggtecc	ccgcaaactc	540
cccttcccgg	ccaccttgg	cgcgtccgcg	ccgccgccgg	cccagccgga	ccgcaccacg	600
cgaggcgcga	gatagggggg	cacgggcgcg	accatctgcg	ctgcggcgcc	ggcgaactcag	660
cgctgcctca	gtctgcggtg	ggcagcggag	gagtcgtgtc	gtgcctgaga	gcgcag	716

<210> 9

<211> 792

<212> DNA

<213> 人工序列

<220>

<223> SP0019

<400> 9

tcaggggtgc	agcttttttt	ctgtctttta	ctcagcctga	gaaaggttgt	cgtttgacaa	60
ggtttgttca	gaggtcggat	ctgaatccag	ctccaaggcc	ccagcaccca	agccctgacc	120
ccatggctgc	ctgctggctg	ggagtggcat	tctttagaac	aggaaaggca	gtgagttctc	180
attgcatcaa	tacttgcat	tgctacaaca	gaagcttttt	gtctcttct	cacattcctt	240
tagaacacaa	gcctcctttt	ctgcctgtgg	ccatcggtea	ctggagtcag	cctcgtgagt	300
ggcttggtgg	cggatggcac	ctgcctttgt	gggagaaaac	aatgttgctg	cccaggcctt	360
tctggaatga	ccccttccca	cccctgacc	agccccagca	aggcccgggg	ctggctgcct	420
agttgattgg	agccaagagt	ttgctgaatg	gatgaaggga	gaaggacat	ccaagtgact	480
gaagaagatc	ttaacagaag	ggctaggect	gcgcaccacc	ccaccgacc	ctcaccacc	540
gaccctcac	ccaccgacca	aggggcacc	tggcctagag	gggatgctga	gcgggaccgc	600
cctcctgcct	ctggcagtc	cagatgggac	ttggaccccg	cagttgctct	ctcggaccct	660
aagtttctac	ccctggatct	aaggcggagc	tgggtttgcg	gateccacgg	ttcccggcgg	720
ggcggggccc	ggtcgcccct	ccccctccc	gcctcctgc	gccgggagca	gtgcattgtg	780
ggaaactccc	ga					792

<210> 10

<211> 686

<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0020	
<400>	10	
cttcctctta	tatttcacca	agactcagtt cctgagcaag aaaccacagg cacagcaagt 60
gccatgaaaa	gcggtttgt	gtgggtggg ctcttcacac tccaatctcc acttccttct 120
caaggcctca	aaaaaagtg	aaaaatgaaa acaaaagccc tgctgtgttg agctgggctc 180
tgcggttgcc	atggaccag	ggcaaacagc ggtgctctg ctctgcccc ggctcagctc 240
atgctgggcc	tgcaactctg	gaagggagca tggactttgg aatgactggg tagaacccaa 300
atgaattaat	ggaatttgac	atagttcaaa aataataaaa tgtgataccc atgaaatgct 360
gatattctgc	cttaatttgc	cagattgggg gccgggctag gctgcgcac ccaccaccg 420
accctcacc	caccgaccg	tcaccaccg accaaggggc accctggcct agaggggatg 480
ctgagcggga	cccgcctcct	gcctctggca gtcccagatg ggacttgac cccgcagttg 540
ctctctcgga	ccctaagttt	ctaccctgg atctaaggcg gagctgggtt tgcggatccc 600
acggttcccg	gcggggcggg	gcccggtcgc ccctccccct ccccgcctc ctgcgccggg 660
agcagtgc	atgtgggaaac	tcccga 686
<210>	11	
<211>	807	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0021	
<400>	11	
tcaggggtgc	agcttttttt	ctgtctttta ctcagcctga gaaaggttgt cgtttgacaa 60
ggtttgttca	gaggtcggat	ctgaatccag ctccaaggcc ccagcaccca agccctgacc 120
ccatggctgc	ctgctggctg	ggagtggcat tctttagaac aggaaaggca gtgagttctc 180
attgcatcaa	tacttgcat	tgctacaaca gaagcttttt gtctcttctt cacattcctt 240
tagaacacaa	gcctcctttt	ctgctgtgg ccatcggtea ctggagtcag cctcgtgagt 300
ggcttggtgg	cggatggcac	cgtcctttgt gggagaaaac aatgttgctg cccaggcctt 360
tctggaatga	ccccttccca	cccctgacc agccccagca aggccccggg ctggctgcct 420
agttgattgg	agccaagagt	ttgctgaatg gatgaaggga gaaggacat ccaagtgact 480
gaagaagatc	ttaacagaat	gcggcgaggc gcgtgcgac tgccagcttc agcaccgcgg 540
acagtgcctt	cgccccgc	tggcggcgcg cgccaccgc gcctcagcac tgaaggcgcg 600
ctgacgtcac	tcgccgtcc	cccgeaaact ccccttccc gccaccttg tgcgctccgc 660
gccgccgccg	gcccagccgg	accgcaccac gcgagcgcg agataggggg gcacggcgcg 720
gaccatctgc	gctgcggcg	cggcgactca gcgctgcctc agtctgcggt gggcagcgga 780
ggagtcgtgt	cgtgcctgag	agcgag 807
<210>	12	

<211>	701	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0022	
<400>	12	
cttcctctta	tatttcacca	agactcagtt cctgagcaag aaaccacagg cacagcaagt 60
gccatgaaaa	gcggttttgt	gtgggtggg ctcttcacac tccaatctcc acttccttct 120
caaggcctca	aaaaaagttg	aaaaatgaaa acaaaagccc tgctgtgttg agctgggctc 180
tggcgttgcc	atggaccag	ggcaaacagc ggtgctctg ctctgcccc ggctcagctc 240
atgctgggcc	tgcaactctg	gaagggagca tggactttgg aatgactggg tagaacccaa 300
atgaattaat	ggaatttgac	atagttcaaa aataataaaa tgtgataccc atgaaatgct 360
gatattctgc	cttaatttgc	cagattgggg gctgcggcg aggcgcgtgc gcaactgccag 420
cttcagcacc	gcggacagtg	ccttcgcccc cgctggcgg cgcgcgccac cgccgcctca 480
gcaactgaagg	cgcgctgacg	tactcgcgg gtccccgca aactcccctt cccggccacc 540
ttggtcgcgt	ccgcgccgcc	gccggcccag ccggaccgca ccacgcgagg cgcgagatag 600
gggggcacgg	gcgcgaccat	ctgcgctgcg gcgccggcga ctcagcgtg cctcagtctg 660
cggtgggcag	cggaggagtc	gtgtcgtgcc tgagagcgca g 701
<210>	13	
<211>	913	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SP0028	
<400>	13	
tcaggggtgc	agcttttttt	ctgtctttta ctcagcctga gaaaggttgt cgtttgacaa 60
ggtttgttca	gaggtcggat	ctgaatccag ctccaaggcc ccagcaccca agccctgacc 120
ccatggctgc	ctgctggctg	ggagtggcat tctttagaac aggaaaggca gtgagttctc 180
attgcatcaa	tacttgcat	tgctacaaca gaagcttttt gtctcttctt cacattcctt 240
tagaacacaa	gcctcctttt	ctgctgtgg ccatcggtca ctggagtcag cctcgtgagt 300
ggcttggtgg	cggatggcac	cgctccttgt gggagaaaac aatgttgctg cccaggcctt 360
tctggaatga	ccccctccca	ccccctgacc agccccagca aggccccggg ctggctgcct 420
agttgattgg	agccaagagt	ttgctgaatg gatgaaggga gaaggacat ccaagtgact 480
gaagaagatc	ttaacagaaa	gcgcgcagag tctgcatgcg tgaggaagct cctgggcgcg 540
tcacagccgc	gctattctca	gcgtctctcc ttttatggct ccggaagtga gctggggttg 600
ctggcagcct	ggctggcact	gggctaggcc tgcgcacca cccaccgacc cctcaccac 660
cgaccctca	cccaccgacc	aaggggcacc ctggcctaga ggggatgctg agcgggacc 720
gcctcctgcc	tctggcagtc	ccagatggga cttggacccc gcagttgctc tctcggacc 780
taagtttcta	cccctggatc	taaggcggag ctgggtttgc ggatcccacg gttcccggcg 840

gggcggggcc cggtcgcccc tccccctccc cgccctcctg cgccgggagc agtgcattgt	900
gggaaactcc cga	913
<210> 14	
<211> 928	
<212> DNA	
<213> 人工序列	
<220>	
<223> SP0029	
<400> 14	
tcaggggtgc agcttttttt ctgtctttta ctcagcctga gaaaggttgt cgtttgacaa	60
ggtttgttca gaggtcggat ctgaatccag ctccaaggcc ccagcaccca agccctgacc	120
ccatggctgc ctgctggetg ggagtggcat tctttagaac aggaaaggca gtgagttctc	180
attgcatcaa tacttgcat tgetacaaca gaagcttttt gtctcttctt cacattcctt	240
tagaacacaa gcctcctttt ctgcctgtgg ccatcggtca ctggagtcag cctcgtgagt	300
ggcttggtgg cggatggcac cgtcctttgt gggagaaaac aatgttgctg cccaggcctt	360
tctggaatga ccccttccca cccctgacc agccccagca aggcccgggg ctggctgcct	420
agttgattgg agccaagagt ttgctgaatg gatgaaggga gaaggacat ccaagtgact	480
gaagaagatc ttaacagaaa gcgcgagag tctgcatgcg tgaggaagct cctgggcgcg	540
tcacagccgc gctattctca gcgtctctcc ttttatggct ccggaagtga gctggggttg	600
ctggcagcct ggctggcact tgcggcgagg cgctgcgca ctgccagctt cagcaccgcg	660
gacagtgcct tcgccccgc ctggcggcgc gcgccaccgc cgctcagca ctgaaggcgc	720
gctgacgtca ctgcccgtc ccccgaaac tcccctccc ggccacctg gtcgctccg	780
cgccgccgcc ggcccagccg gaccgcacca cgcgaggcgc gagatagggg ggcacgggcg	840
cgaccatctg cgctgcggcg ccggcgactc agcgtgcct cagtctgcgg tgggcagcgg	900
aggagtcgtg tcgtgcctga gagcgag	928
<210> 15	
<211> 623	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRM_SP0026	
<400> 15	
aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttggea	180
cagacacaaat gttcgggggtg ggcacagtgc ctgettccc ccgcaccca gccccctca	240
aatgccttcc gagaagccca ttgagcaggg ggcttgcat gcaccccagc ctgacagcct	300
ggcatcttgg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
cggcgggtgga gaacaaggct ctattcagcc tgtgcccagg aaaggggatc aggggatgcc	420

caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg gaagcgcgca gagtctgcat gcgtgaggaa gtccttgggc	540
gcgtcacagc cgcgctattc tcagcgtctc tccttttatg gctccggaag tgagctgggg	600
ttgctggcag cctggctggc act	623
<210> 16	
<211> 623	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRM_SP0027	
<400> 16	
aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctcctc ttcattgccc gtgaatgact caccttggea	180
cagacacaaat gttcgggggtg ggcacagtgc ctgcttccc cgcacccca gccccctca	240
aatgccttcc gagaagccca ttgagcaggg ggcttgcttgcatt gcacccagc ctgacagcct	300
ggcatcttgg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
cggcgggtgga gaacaaggct ctattcagcc tgtgcccagg aaaggggatc aggggatgcc	420
caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg gaagcgcgca gagtctgcat gcgtgaggaa gtccttgggc	540
gcgtcacagc cgcgctattc tcagcgtctc tccttttatg gctccggaag tgagctgggg	600
ttgctggcag cctggctggc act	623
<210> 17	
<211> 620	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRM_SP0028	
<400> 17	
tcaggggtgc agcttttttt ctgtctttta ctcagcctga gaaaggttgt cgtttgacaa	60
ggtttgttca gaggtcggat ctgaatccag ctccaaggcc ccagcaccca agccctgacc	120
ccatggctgc ctgctggctg ggagtggcat tctttagaac aggaaaggca gtgagttctc	180
attgcatcaa tacttgcttgcatt tgcataca gaagcttttt gtctcttctt cacattcctt	240
tagaacacaa gcctcctttt ctgcctgtgg ccatcggtea ctggagtcag cctcgtgagt	300
ggcttgggtgg cggatggcac cgtcctttgt gggagaaaac aatgcttctg cccaggcctt	360
tctggaatga ccccttccca cccctgacc agccccagca agccccgggg ctggctgctt	420
agttgattgg agccaagagt ttgctgaatg gatgaaggga gaaggacat ccaagtgact	480
gaagaagatc ttaacagaaa gcgcgcagag tctgcatgct tgaggaagct cctgggcgcg	540
tcacagccgc gctattctca gcgtctctcc ttttatggct ccggaagtga gctggggttg	600

ctggcagcct ggctggcact	620
<210> 18	
<211> 620	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRM_SP0029	
<400> 18	
tcaggggtgc agcttttttt ctgtctttta ctcagcctga gaaaggttgt cgtttgacaa	60
ggtttgttca gaggtcggat ctgaatccag ctccaaggcc ccagcaccca agccctgacc	120
ccatggctgc ctgctggetg ggagtggcat tctttagaac aggaaaggca gtgagttctc	180
attgcatcaa tacttgcatt tgetacaaca gaagcttttt gtctcttctc cacattcctt	240
tagaacacaa gcctcctttt ctgcctgttg ccatcggtea ctggagttag cctcgtgagt	300
ggcttggtgg cggatggcac cgtcctttgt gggagaaaac aatggtgctg cccaggcctt	360
tctggaatga ccccttccca cccctgacc agccccagca aggccccggg ctggctgctt	420
agttgattgg agccaagagt ttgctgaatg gatgaaggga gaaggacat ccaagtgact	480
gaagaagatc ttaacagaaa gcgcgcagag tctgcatgag tgaggaagct cctgggcgag	540
tcacagccgc gctattctca gcgtctctcc ttttatgget ccggaagtga gctggggttg	600
ctggcagcct ggctggcact	620
<210> 19	
<211> 502	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0005_GFAP (SYNP_AST_CRE0005_GFAP)	
<400> 19	
aacatatacct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
attgggctgg ccgccccca gggcctctc ttcatgcca gtgaatgact caccttgcca	180
cagacacaaat gttcgggggtg ggcacagtgc ctgcttccc cgcaccca gccccctca	240
aatgccttcc gagaagcca ttgagcaggg ggcttgcatt gcaccccagc ctgacagcct	300
ggcatcttgg gataaaagca gcacagcccc ctaggggctg cccttgctgt gtggcgccac	360
cggcgggtgga gaacaaggct ctattcagcc tgtgcccagg aaaggggatc aggggatgcc	420
caggcatgga cagtgggtgg caggggggga gaggaggct gtctgcttcc cagaagtcca	480
aggacacaaa tgggtgaggg ga	502
<210> 20	
<211> 324	
<212> DNA	
<213> 人工序列	

<220>		
<223>	CRE0007_GFAP (SYNP_AST_CRE0007_GFAP)	
<400>	20	
	aacatatcct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
	gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
	attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttgca	180
	cagacacaat gttcgggggtg ggcacagtgc ctgcttccc cgcacccca gccccctca	240
	aatgccttcc gagaagccca ttgagcagg ggttgcatt gcacccagc ctgacagcct	300
	ggcatcttgg gataaaaagca gcac	324
<210>	21	
<211>	121	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0012_Arc (SYNP_NEU_CRE0012_Arc)	
<400>	21	
	agcgcgcaga gtctgcatgc gtgaggaagc tcctgggagc gtcacagccg cgctattctc	60
	agcgtctctc cttttatggc tccggaagtg agctggggtt gctggcagcc tggctggcac	120
	t	121
<210>	22	
<211>	393	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0001_S100B (SYNP_AST_CRE0001_S100B)	
<400>	22	
	cttcctctta tatttcacca agactcagtt cctgagcaag aaaccacagg cacagcaagt	60
	gccatgaaaa gcggctttgt gtggggtggg ctcttcacac tccaatctcc acttccttct	120
	caaggcctca aaaaaagttg aaaaatgaaa acaaaagccc tgctgtgttg agctgggctc	180
	tggcgttgcc atggaccag ggcacaacagc ggtgctctctg ctctgcccc ggctcagctc	240
	atgctgggcc tgcacttctg gaaggagca tggactttgg aatgactggt tagaacccaa	300
	atgaattaat ggaatttgac atagttcaaa aataataaaa tgtgataccc atgaaatgct	360
	gatattctgc cttaatttgc cagattgggg gcc	393
<210>	23	
<211>	408	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0009_S100B (SYNP_AST_CRE0009_S100B)	

<400> 23		
ttcagaggtc ggatctgaat ccagctccaa ggccccagca cccaagccct gaccccatgg	60	
ctgcctgctg gctgggagtg gcattcttta gaacaggaaa ggcagtgagt tctcattgca	120	
tcaatacttg catttgctac aacagaagct tttgtctct tcctcacatt cttttagaac	180	
acaagcctcc ttttctgcct gtggccatcg gtcactggag tcagcctcgt gagggttg	240	
gtggcggatg gcaccgtcct ttgtgggaga aaacaatgtt gctgcccagg cttttctgga	300	
atgaccctt cccaccctt gaccagcccc agcaaggccc gggctggct gcctagttag	360	
ttggagccaa gaggttgctg aatggatgaa gggagaaggg acatccaa	408	
<210> 24		
<211> 499		
<212> DNA		
<213> 人工序列		
<220>		
<223> CRE0002_S100B (SYNP_AST_CRE0002_S100B)		
<400> 24		
tcaggggtgc agcttttttt ctgtctttta ctcagcctga gaaagttgt cgtttgacaa	60	
ggtttgttca gaggtcggat ctgaatccag ctccaaggcc ccagcaccca agccctgacc	120	
ccatggctgc ctgctggctg ggagtggcat tctttagaac aggaaaggca gtgagttctc	180	
attgcatcaa tacttgatt tgctacaaca gaagcttttt gtctcttctt cacattcctt	240	
tagaacacaa gcctcctttt ctgcctgtgg ccatcggta ctggagtcag cctcgtgagt	300	
ggcttggtgg cggatggcac cgtcctttgt gggagaaaac aatggtgctg cccaggcctt	360	
tctggaatga ccccttccca cccctgacc agccccagca aggcccgggg ctggctgcct	420	
agttgattgg agccaagagt ttgctgaatg gatgaaggga gaaggacat ccaagtgact	480	
gaagaagatc ttaacagaa	499	
<210> 25		
<211> 293		
<212> DNA		
<213> 人工序列		
<220>		
<223> CRE0003_CEND1_mp (SYNP_NEU_CRE0003_CEND1_mp)		
<400> 25		
gggctaggcc tgcgcaccca cccaccgacc cctcaccac cgaccctga cccaccgacc	60	
aaggggcacc ctggcctaga ggggatgctg agcgggacc gcctcctgcc tctggcagtc	120	
ccagatggga cttggacccc gcagttgctc tctcgaccc taagtttcta cccctggatc	180	
taaggcggag ctgggtttgc ggatcccacg gttcccggcg gggcggggcc cggtcgcccc	240	
tccccctccc cgccctcctg cgccgggagc agtgcattgt gggaaactcc cga	293	
<210> 26		
<211> 308		
<212> DNA		

<213>	人工序列	
<220>		
<223>	CRE0009_SYN1_mp (SYNP_NEU_CRE0009_SYN1_mp)	
<400>	26	
	tgcggcgagg cgcgtgcgca ctgccagctt cagcaccgcg gacagtgcct tcgcccccg	60
	ctggcggcgc gcgccaccgc cgcctcagca ctgaaggcgc gctgacgtca ctgccggtc	120
	ccccgcaaac tccccttccc ggccaccttg gtcgctccg cggcccgcc ggcccagccg	180
	gaccgcacca cgcgaggcgc gagatagggg ggcacggcgc cgaccatctg cgctgcggcg	240
	ccggcgactc agcgtgcct cagtctgagg tgggcagcgg aggagtcgtg tcgtgcctga	300
	gagcgag	308
<210>	27	
<211>	359	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0006_GFAP (SYNP_AST_CRE0006_GFAP)	
<400>	27	
	aacatatacct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
	gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
	attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttgca	180
	cagacacaat gttcggggtg ggcacagtgc ctgcttccc cgcaccca gccccctca	240
	aatgccttcc gagaagcca ttgagcagg gcttgcatt gcacccagc ctgacagcct	300
	ggcatcttgg gataaaagca gcacagccc ctaggggctg cccttgctgt gtggcgcca	359
<210>	28	
<211>	463	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0008_GFAP (SYNP_AST_CRE0008_GFAP)	
<400>	28	
	aacatatacct ggtgtggagt aggggacgct gctctgacag aggctcgggg gcctgagctg	60
	gctctgtgag ctggggagga ggcagacagc caggccttgt ctgcaagcag acctggcagc	120
	attgggctgg ccgccccca gggcctctc ttcatgccc gtgaatgact caccttgca	180
	cagacacaat gttcggggtg ggcacagtgc ctgcttccc cgcaccca gccccctca	240
	aatgccttcc gagaagcca ttgagcagg gcttgcatt gcacccagc ctgacagcct	300
	ggcatcttgg gataaaagca ccggcggttg agaacaaggc tctattcagc ctgtgcccag	360
	gaaaggggat caggggatgc ccaggcatgg acagtgggtg gcagggggg agaggaggc	420
	tgtctgcttc ccagaagtcc aaggacaca atgggtgagg gga	463
<210>	29	

<211>	88	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0001_CEND1_mp (SYNP_NEU_CRE0001_CEND1_mp)	
<400>	29	
	gatccccacgg ttcccggcgg ggcggggccc ggtcgcccct ccccctcccc gccctcctgc	60
	gccgggagca gtgcattgtg ggaaactc	88
<210>	30	
<211>	346	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0002_CEND1_mp (SYNP_NEU_CRE0002_CEND1_mp)	
<400>	30	
	gggctaggcc tgcgcacca cccaccgacc cctcaccac cgaccctca cccaccgacc	60
	aaggggcacc ctggcctaga ggggatgctg agcgggacc gcctcctgcc tctggcagtc	120
	ccagatggga cttggacccc gcagttgctc tctcggacc taagtttcta ccctggatc	180
	taaggcggag ctgggtttgc ggatcccacg gttcccggcg gggcggggcc cggtcgcccc	240
	tccccctccc cgccctcctg cgccgggagc agtgcattgt gggaaactcc cgagctctcc	300
	tccgcgttcg cagccgccgt catcccgcgg aggagcgcgc agcccc	346
<210>	31	
<211>	468	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	CRE0008_SYN1_mp (SYNP_NEU_CRE0008_SYN1_mp)	
<400>	31	
	tgcagagggc cctgcgtatg agtgcaagtg ggttttagga ccaggatgag gcggggtggg	60
	ggtgcctacc tgacgaccga ccccgacca ctggacaagc acccaacccc cattccccaa	120
	attgcgcatc ccctatcaga gagggggagg ggaacagga tgcggcgagg cgcgtgcgca	180
	ctgccagctt cagcaccgcg gacagtgcct tcgccccgc ctggcggcgc gcgccaccgc	240
	cgcctcagca ctgaaggcgc gctgacgtca ctgcgggtc ccccgaaac tccccttccc	300
	ggccaccttg gtcgcgtccg cgccgccgc ggcccagccg gaccgacca cgcgaggcgc	360
	gagatagggg ggcacgggcg cgaccatctg cgctgcggcg ccggcgactc agcgtgcct	420
	cagtctgcgg tgggcagcgg aggagtctg tcgtgcctga gagcgcag	468
<210>	32	
<211>	266	
<212>	DNA	

<213> 人工序列	
<220>	
<223> CRE0013_SYN1_mp (SYNP_NEU_CRE0013_SYN1_mp)	
<400> 32	
tgcggcgagg cgcgtgcgca ctgccagctt cagcaccgcg gacagtgcct tcgccccgc	60
ctggcggcgc gcgccaccgc cgcctcagca ctgaaggcgc gctgacgtca ctgccggtc	120
ccccgcaaac tccccttccc ggccacctg gtcgctccg cgccgccgcc ggcccagccg	180
gaccgcacca cgcgaggcgc gagatagggg ggcacggcgc cgaccatctg cgctgcggcg	240
ccggcgactc agcgtgcct cagtct	266
<210> 33	
<211> 529	
<212> DNA	
<213> 人工序列	
<220>	
<223> SP0011	
<400> 33	
ctagcccaca ggaaatgtct gtctatatcc aggcaagtac cttgctcatt ggaccaaccc	60
gaaactgttc agggaagatc agggaaatca actcagttac aatgggata atcatgcca	120
gtaaaaaacta cctgtggtga ataaagagtt aaccctgtt ccatcttagg tcaactatgca	180
gagtaccaat gagtacaaga gatggtgcca aagagggtgg ccctcccta gctgggaaca	240
gtcaaccctt aggaactaga ctgtcaacac atcagccagc cagagacaag ggaaaccgtg	300
gcaaccaagt gttgctggca cattgtgagg tggatgagg aactgcagag gccctgcaca	360
gcatgctaat gagcccaggc aaacatgagc tctcccata gctgggctgc ggcccaccc	420
cacccccctca ggctatgcca gggggtgtg ccaggggcac cgggcatcg ccagtctagc	480
ccactccttc ataaagccct cgcatcccag gagcgagcag agccagagc	529
<210> 34	
<211> 646	
<212> DNA	
<213> 人工序列	
<220>	
<223> SP0034	
<400> 34	
tccaaagaaa agccagataa gtagctgatt attgcataga gctgacagta tcacaggaag	60
atcagtagta gcagctcaag taaaaaagg ttaattagca atacttaata agaaaaacta	120
cctctggcag gtgaagagtt aatccctggt caattttaag ctactctgct gagagtacta	180
ataagtgtag gggttggagc caatgagggt gacccttcc ttgatgggaa cagtcatecc	240
ttaggaactg ccctggaaag catcagccag ccagaaaaca gggaaagagg ctgagaaacc	300
gtggtaacca agttttgctg gcactttgta aaatggtaac tgcaactgcc gaggtgtgc	360
agagaatgct aataagccta ggacaacctg taaagagtgg acctagaaaa tgtccaccg	420

ctagagagag	ggagcgagca	tgtgcatga	gcaatagctg	tggaccttac	agttgctgct	480
aactgccctg	gtgtgtgtga	gggagagaga	gggagggagg	gagagagagc	gcgctagcgc	540
gagagagcga	gtgagcaagc	gagcagaaaa	gaggtggaga	gggggggaat	aagaaagaga	600
gagaaggaaa	ggagagaagg	caggaagaag	gcaagggacg	agacaa		646
<210>	35					
<211>	614					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	SP0035					
<400>	35					
ctagcccaca	ggaaatgtct	gtctatatec	aggcaagtac	cttgctcatt	ggaccaaccc	60
gaaactgttc	agggaagatc	agggaaatca	actcagttac	aatgggata	atcatgccc	120
gtaaaaaacta	cctgtggtga	ataaagagtt	aaccctgtt	ccatcttagg	tcactatgca	180
gagtaccaat	gagtacaaga	gatggtgcca	aagagggtgg	cccctccta	gctgggaaca	240
gtcaaccctt	aggaactaga	ctgtcaacac	atcagccagc	cagagacaag	ggaaaccgtg	300
gcaaccaagt	gttgctggca	cattgtgagg	tggtgatggg	aactgcagag	gccctgcaca	360
gcatgctaat	gagcccaggc	aaacatcgct	agagagaggg	agcgagcatg	tgcgatgagc	420
aatagctgtg	gaccttacag	ttgctgctaa	ctgccctggt	gtgtgtgagg	gagagagagg	480
gagggagggga	gagagagcgc	gctagcgcga	gagagcgagt	gagcaagcga	gcagaaaaga	540
ggtggagagg	gggggaataa	gaaagagaga	gaaggaaagg	agagaaggca	ggaagaaggc	600
aagggacgag	acaa					614
<210>	36					
<211>	695					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	SP0036					
<400>	36					
ctagcccaca	ggaaatgtct	gtctatatec	aggcaagtac	cttgctcatt	ggaccaaccc	60
gaaactgttc	agggaagatc	agggaaatca	actcagttac	aatgggata	atcatgccc	120
gtaaaaaacta	cctgtggtga	ataaagagtt	aaccctgtt	ccatcttagg	tcactatgca	180
gagtaccaat	gagtacaaga	gatggtgcca	aagagggtgg	cccctccta	gctgggaaca	240
gtcaaccctt	aggaactaga	ctgtcaacac	atcagccagc	cagagacaag	ggaaaccgtg	300
gcaaccaagt	gttgctggca	cattgtgagg	tggtgatggg	aactgcagag	gccctgcaca	360
gcatgctaat	gagcccaggc	aaacattcga	gttggtgga	caaggttatg	agcatccgtg	420
tacttatggg	gttgccagct	tggtcctgga	tcgcccgggc	ccttccccca	ccggttcggt	480
tccccaccac	caccgcgct	cgtacgtgcg	tctccgctg	cagctcttga	ctcatcgggg	540
ccccccgggt	cacatgcgct	cgctcggctc	tatagcgcc	gccccctgcc	cacccccgc	600

ccgcgctggg agccgcagcc gccgccactc ctgctctctc tgcgccgccc ccgtcaccac	660
cgccaccgcc accggctgag tctgcagtcc tcgag	695
<210> 37	
<211> 386	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0006_AQP4 (SYNP_AST_CRE0006_AQP4)	
<400> 37	
ctagcccaca ggaaatgtct gtctatatec aggcaagtac cttgctcatt ggaccaaccc	60
gaaactgttc agggaagatc agggaaatca actcagttac aaatgggata atcatgcccc	120
gtaaaaaacta cctgtggtga ataaagagtt aaccctgtt ccatcttagg tcaactatgca	180
gagtaccaat gagtacaaga gatggtgcca aagagggtgg ccctcccta gctgggaaca	240
gtcaaccctt aggaactaga ctgtcaacac atcagccagc cagagacaag ggaaaccgtg	300
gcaaccaagt gttgctggca cattgtgagg tggatgatgg aactgcagag gccctgcaca	360
gcatgctaata gagcccaggc aaacat	386
<210> 38	
<211> 418	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0008_AQP4 (SYNP_AST_TO_CRE0008v2_AQP4)	
<400> 38	
tccaaagaaa agccagataa gtagctgatt attgcataga gctgacagta tcacaggaag	60
atcagtagta gcagctcaag tacaaaaagg ttaattagca atacttaata agaaaaacta	120
cctctggcag gtgaagagtt aatccctggt caattttaag ctactctgct gagagtacta	180
ataagtgtag gggttggagc caatgagggt gacccttcc ttgatgggaa cagtcatccc	240
ttaggaactg ccctggaaag catcagccag ccagaaaaca gggaaagagg ctgagaaacc	300
gtggtaacca agttttgctg gcactttgta aatggtaac tgcaactgcc gaggctgtgc	360
agagaatgct aataagccta ggacaacctg taaagagtgg acctagaaaa tgtccacc	418
<210> 39	
<211> 228	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0015_GAP43_mp (SYNP_NEU_CRE0015_GAP43_mp)	
<400> 39	
cgctagagag agggagcgag catgtgcgat gagcaatagc tgtggacctt acagttgctg	60
ctaactgccc tgggtgtgtg gagggagaga gagggagga gggagagaga gcgcgctagc	120

gcgagagagc gaggagcaa ggcagcagaa aagaggtgga gaggggggga ataagaaaga	180
gagagaagga aaggagagaa ggcaggaaga aggcaaggga cgagacaa	228
<210> 40	
<211> 309	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0016_Eno2_mp (SYNP_NEU_CRE0016_Eno2_mp)	
<400> 40	
tcgagttggc tggacaaggt tatgagcadc cgtgtactta tggggttgcc agcttgggtcc	60
tggatcgccc gggcccttcc cccaccctgt cggttcccca ccaccaccgc cgctcgtagc	120
tgcgtctccg cctgcagctc ttgactcadc ggggcccccc gggtcacatg cgctcgctcg	180
gctctatagg cgccgcccc tgcccacccc ccgcccgcgc tgggagccgc agccgcccgc	240
actcctgctc tctctgcgcc gccgcccgtc ccaccgccac cgccaccggc tgagtctgca	300
gtcctcgag	309
<210> 41	
<211> 143	
<212> DNA	
<213> 人工序列	
<220>	
<223> CRE0004_GFAP_mp (SYNP_AST_CRE0004_GFAP_mp)	
<400> 41	
gagctctccc catagctggg ctgcggccca accccacccc ctcaggctat gccagggggt	60
gttgccaggg gcacccgggc atgccagtc tagcccactc cttcataaag ccctcgcatc	120
ccaggagcga gcagagccag agc	143



图1A

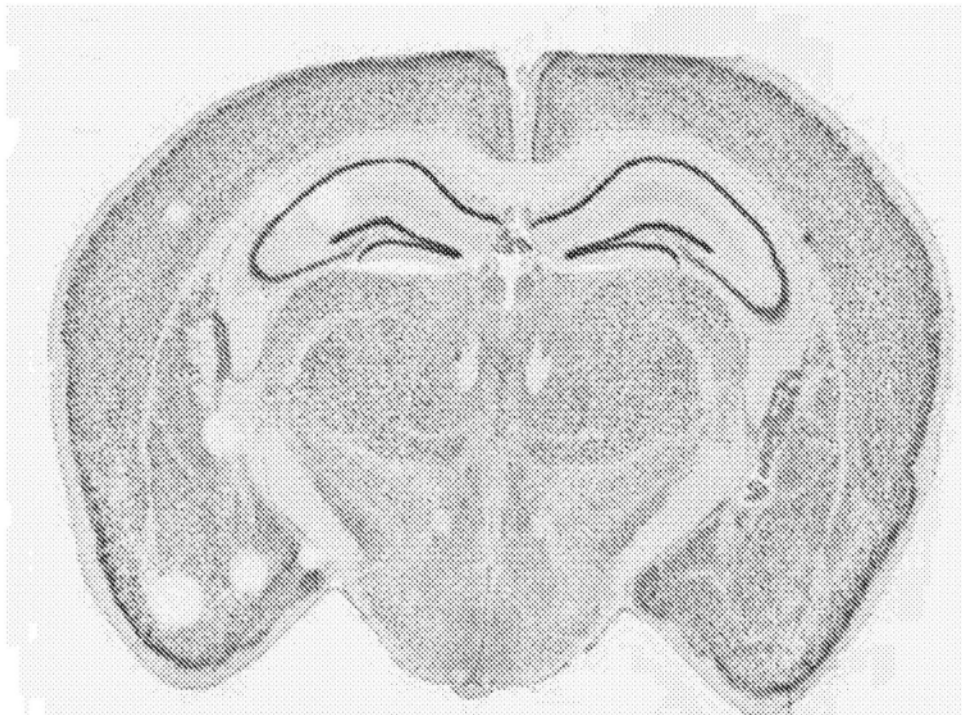


图1B

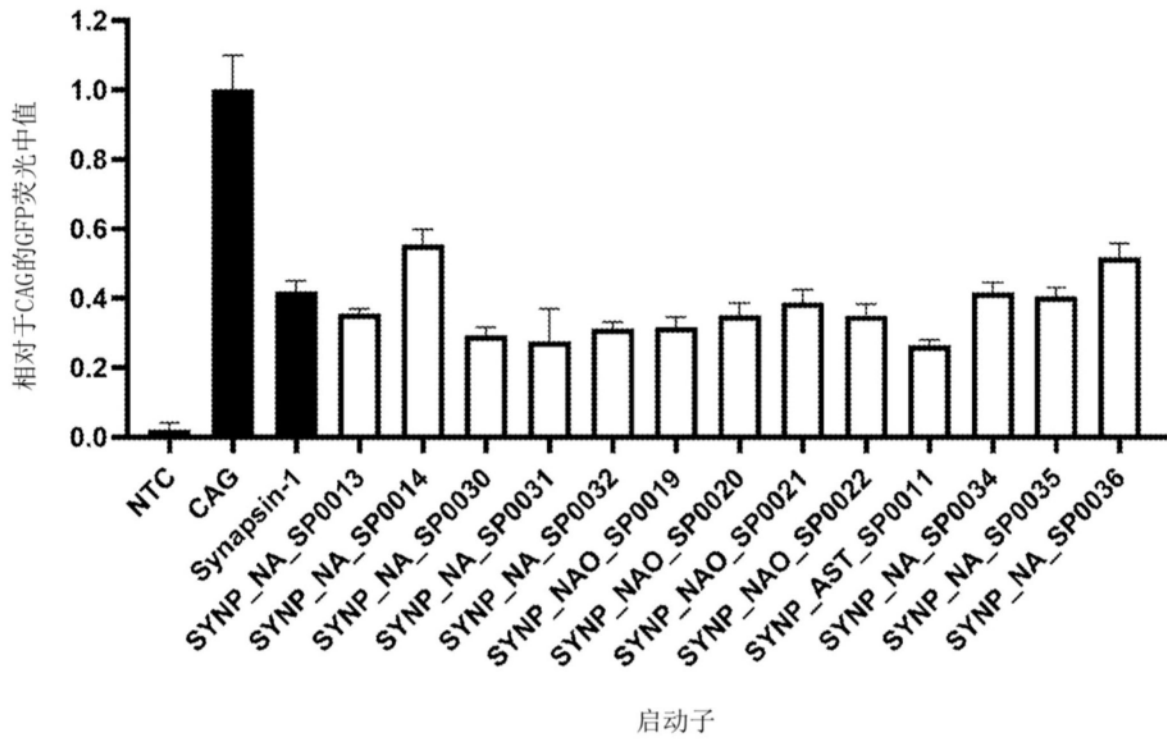


图2A

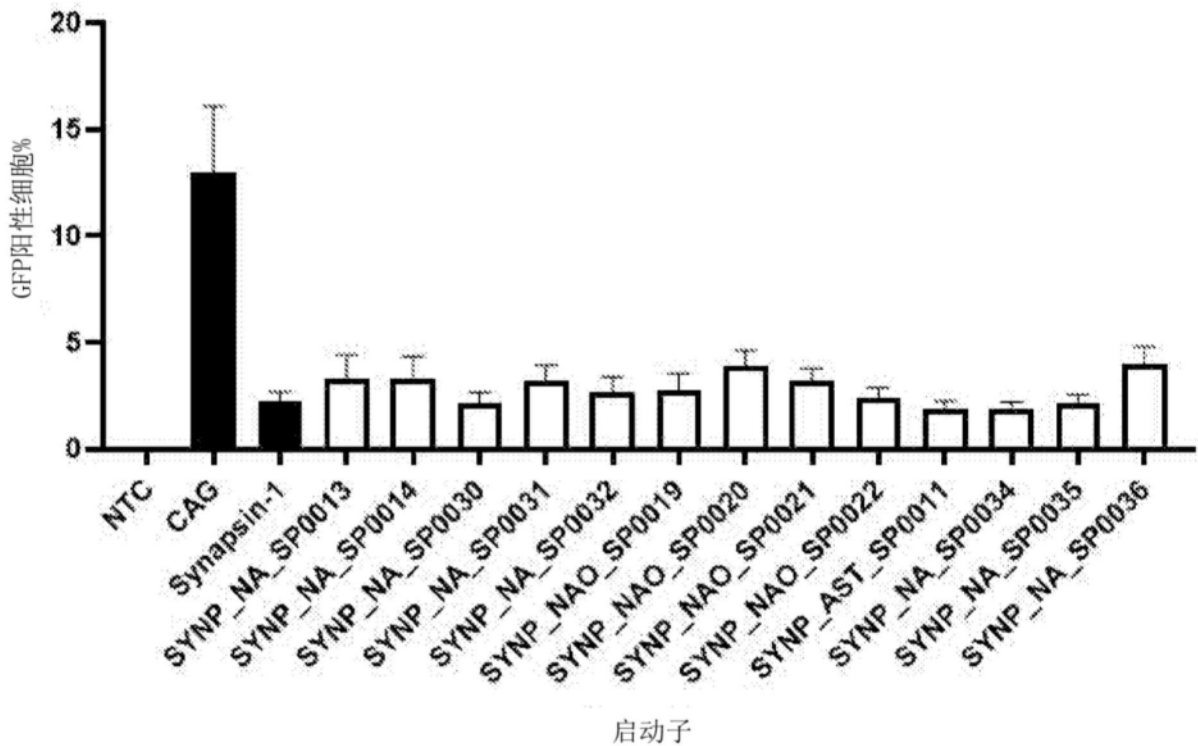


图2B