



(12) 发明专利

(10) 授权公告号 CN 110730823 B

(45) 授权公告日 2023.12.29

(21) 申请号 201880036950.6

C12Q 1/6883 (2018.01)

(22) 申请日 2018.04.03

(56) 对比文件

(65) 同一申请的已公布的文献号

WO 2007078599 A2, 2007.07.12

申请公布号 CN 110730823 A

WO 2012087983 A1, 2012.06.28

(43) 申请公布日 2020.01.24

US 2004096885 A1, 2004.05.20

(30) 优先权数据

US 2014142160 A1, 2014.05.22

62/480,998 2017.04.03 US

US 2015353917 A1, 2015.12.10

(85) PCT国际申请进入国家阶段日

US 2011268747 A1, 2011.11.03

2019.12.03

Axel H. Meyer等. In Vivo Labeling of Parvalbumin-Positive Interneurons and Analysis of Electrical Coupling in Identified Neurons.《The Journal of Neuroscience》.2002, 第22卷(第16期), 第7055-7064页.

(86) PCT国际申请的申请数据

Axel H. Meyer等. In Vivo Labeling of Parvalbumin-Positive Interneurons and Analysis of Electrical Coupling in Identified Neurons.《The Journal of Neuroscience》.2002, 第22卷(第16期), 第7055-7064页.

(续)

PCT/US2018/025940 2018.04.03

审查员 赵秋歌

(87) PCT国际申请的公布数据

W02018/187363 EN 2018.10.11

(73) 专利权人 编码治疗公司

地址 美国加利福尼亚州

(72) 发明人 斯蒂芬妮·塔利亚泰拉

安德鲁·杨 陈思颖

卡提克·拉马穆尔蒂

大卫·奥伯科夫勒

(74) 专利代理机构 北京市金杜律师事务所

11256

专利代理人 陈文平

(51) Int.Cl.

C12N 15/63 (2006.01)

权利要求书2页 说明书63页

A01K 67/027 (2006.01)

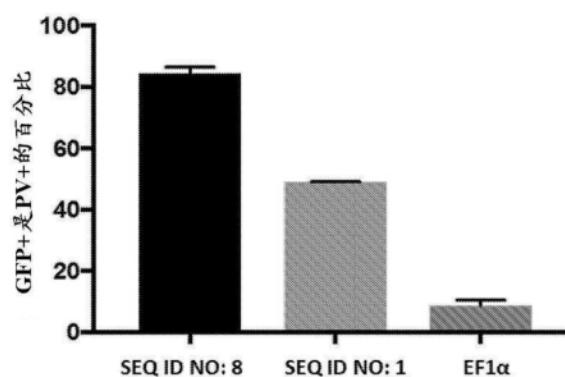
序列表31页 附图14页

(54) 发明名称

组织选择性转基因表达

(57) 摘要

本文提供了用于选择性地表达转基因的组合物和方法。用于选择性地表达转基因的组合物和方法包括一个或多个调控元件，该调控元件在与转基因可操作地连接时，相比于在至少一种或多种非靶细胞中，可促进转基因在靶细胞中的选择性表达(例如，细胞类型选择性表达)。



[转续页]

[接上页]

(56) 对比文件

NOBUAKI TAMAMAKI等.Green Fluorescent Protein Expression and Colocalization with Calretinin, Parvalbumin, and Somatostatin in the GAD67-GFP Knock-In Mouse.《THE JOURNAL OF COMPARATIVE

NEUROLOGY》.2003,第467卷第60-79页.

A N van den Pol等.Selective neuronal expression of green fluorescent protein with cytomegalovirus promoter reveals entire neuronal arbor in transgenic mice.《J Neurosci》.1998,第18卷(第24期),第10640-10651页.

1.一种核酸盒,其包含:

与转基因可操作地连接的调控元件,所述调控元件导致所述转基因在中枢神经系统(CNS)中的小清蛋白(PV)神经元中的表达相对于在CNS中的一种或多种非PV细胞中的表达选择性地增加,并且其中所述调控元件包含SEQ ID NO:1。

2.根据权利要求1所述的核酸盒,其中所述转基因编码离子通道亚基、神经递质调节因子、DNA结合域、基因编辑蛋白。

3.根据权利要求2所述的核酸盒,其中所述转基因编码离子通道亚基,并且所述离子通道亚基是钠离子通道的 α 亚基、钠离子通道的 β 亚基或钾离子通道的亚基。

4.根据权利要求3所述的核酸盒,其中所述转基因编码SEQ ID NO:37-43中任一项的序列。

5.根据权利要求3所述的核酸盒,其中所述转基因编码SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1或KV3.3。

6.根据权利要求2所述的核酸盒,其中所述转基因编码神经递质调节因子,并且所述神经递质调节因子包括STXBP1。

7.根据权利要求2所述的核酸盒,其中所述转基因编码DNA结合蛋白,并且所述DNA结合蛋白调节内源基因的表达。

8.根据权利要求7所述的核酸盒,其中所述内源基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3或STXBP1。

9.根据权利要求2所述的核酸盒,其中所述转基因编码所述基因编辑蛋白。

10.根据权利要求9所述的核酸盒,其中所述基因编辑蛋白是Cas蛋白。

11.根据权利要求1-10中的任一项所述的核酸盒,其中所述调控元件的大小小于2.5kb。

12.根据权利要求1-10中的任一项所述的核酸盒,其中所述非PV细胞是兴奋性神经元、多巴胺能神经元、星形胶质细胞、小胶质细胞或运动神经元。

13.根据权利要求1-10中的任一项所述的核酸盒,其中所述核酸盒是线性构建体。

14.根据权利要求1-10中的任一项所述的核酸盒,其中所述核酸盒是载体。

15.根据权利要求14所述的核酸盒,其中所述载体是质粒。

16.根据权利要求14所述的核酸盒,其中所述载体是病毒载体。

17.根据权利要求16所述的核酸盒,其中所述病毒载体是腺伴随病毒(AAV)载体。

18.根据权利要求17所述的核酸盒,其中所述AAV载体是AAV1、AAV8、AAV9、scAAV1、scAAV8或scAAV9。

19.根据权利要求16所述的核酸盒,其中所述病毒载体是慢病毒载体。

20.根据权利要求1-10中的任一项所述的核酸盒,其中所述调控元件包含来自GAD2、GAD1、SYN1、NKX2.1、DLX1、DLX5/6、SST、PV或VIP的转录起始位点的10kb以内的小于600bp的连续序列。

21.根据权利要求1-10中的任一项所述的核酸盒在制备用于治疗有需要的受试者的神经系统病症或病况的药物中的用途。

22.根据权利要求1-10中的任一项所述的核酸盒在制备用于在受试者的CNS中相对于一种或多种非PV细胞选择性地在PV神经元中表达转基因的药物中的用途。

23. 根据权利要求21所述的用途,其中所述神经系统病况或病症是癫痫、神经变性、滔蛋白病变或神经元兴奋不足。

24. 根据权利要求21所述的用途,其中所述神经系统病况或病症是Dravet综合征。

25. 根据权利要求21所述的用途,其中所述神经系统病况或病症是阿尔茨海默病。

组织选择性转基因表达

[0001] 交叉引用

[0002] 本申请要求于2017年4月3日提交的美国临时申请号62/480,998的权益，该临时申请通过引用全文并入本文。

[0003] 序列表

[0004] 本申请包含序列表，该序列表已经以ASCII格式电子提交，并且通过引用以其全文并入于此。所述ASCII副本创建于2018年3月29日，命名为46482-704_601_SL.txt，大小为78,248字节。

背景技术

[0005] 基因疗法在应对和治疗人类疾病方面的巨大潜力早已被认可。可通过直接靶向潜在的病因来治疗患者，尤其是具有潜在遗传因子的患者，而无需依靠于药物或手术。此外，通过靶向潜在的遗传原因，基因疗法能够有效治愈患者，或提供更长期的持续治疗。但尽管如此，基因疗法的临床应用仍需要在几个方面进行改进。脱靶效应是其中一个值得关注的领域。解决脱靶效应的一种有吸引力的方法是将基因疗法的基因表达靶向感兴趣的细胞类型或组织，或者靶向靶细胞类型或组织。因此，需要鉴定用于将基因疗法或基因表达靶向感兴趣的组织或细胞类型的元件及其使用方法。

发明内容

[0006] 存在相当大的将基因疗法及其基因/转基因表达在体内靶向所需的组织和/或细胞类型的需要，这可以降低脱靶效应，提高对靶组织和/或细胞类型的治疗功效，并通过降低达到功效所需的有效剂量而提高患者的安全性和耐受性。

[0007] 本文提供了相对于在一种或多种非靶组织或细胞类型中，使转基因在靶组织或细胞类型中选择性地表达的组合物和方法。用于选择性地表达转基因的组合物和方法包括一种或多种调控元件(RE)，当与转基因(例如，离子通道亚基或神经递质调节因子，或突触融合蛋白结合蛋白)可操作地连接时，相比于在一种或多种非靶细胞类型(例如，非PV细胞)中，该调控元件可促进或导致转基因在靶组织或细胞类型(例如，小清蛋白(PV)神经元)中选择性表达或优先表达。在一些情况下，RE是非天然存在的序列。在一些情况下，RE是来源于人类的调控元件。在一些情况下，RE包含来自非人类物种的序列，例如猴、狗、兔、小鼠的序列。在一些情况下，使用病毒载体和/或病毒颗粒，例如腺伴随病毒(AAV)或慢病毒，将本文所述的组合物在体内、离体或体外递送到细胞中。在一些情况下，本文所述的组合物作为基因疗法被递送到细胞中。本文还考虑到用于治疗与CNS(CNS)的遗传缺陷有关的神经系统病况或病症的方法和组合物。在一些情况下，受遗传缺陷影响的相关细胞类型或组织是PV细胞。在一些情况下，神经系统病况或疾病是Dravet综合征、阿尔茨海默病、癫痫和/或癫痫发作。在一些情况下，神经系统病况或疾病是精神障碍(例如，精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病)；自闭症谱系障碍(例如，脆性X综合征、雷特综合征)；癫痫(例如，慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结

节性硬化)；或神经变性(例如，阿尔茨海默病、帕金森病)。在一些情况下，神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。

[0008] 一方面，本公开内容考虑到一种核酸盒，其包含与转基因可操作地连接的一个或多个调控元件，相对于在一种或多种非靶细胞类型中或在CNS中的非PV细胞中，该调控元件导致在任何靶细胞类型，例如CNS中的PV神经元中的选择性表达。在一些情况下，每个调控元件包含(i) SEQ ID NO:1-32的序列，(ii) 其功能片段或组合，或(iii) 与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下，使用BLAST确定序列同一性。在一些情况下，至少一种调控元件是来源于人类的。在一些情况下，至少一种调控元件来源于非人类哺乳动物。在一些情况下，调控元件是非天然存在的。在一些情况下，如通过共定位测定所测量的，调控元件导致转基因在PV神经元中的选择性表达，该选择性表达高于相同转基因与非选择性调控元件可操作地连接时的表达。在一些情况下，非选择性调控元件是组成型启动子。在一些情况下，非选择性调控元件是CAG、EF1 α 、SV40、CMV、UBC、PGK和CBA中的任一个。在一些情况下，如通过共定位测定所测量的，所述调控元件导致转基因在PV神经元中的选择性表达，该选择性表达的水平是该转基因在与非选择性调控元件可操作地连接时在PV神经元中的选择性表达的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍或至少20倍。在一些情况下，调控元件导致在PV神经元中的选择性表达，该选择性表达比在转基因与非选择性调控元件可操作地连接时在PV神经元中的表达高至少2%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。在一些情况下，调控元件导致在PV神经元中的选择性表达，该选择性表达是CNS中PV神经元的自然分布的预期表达的约1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍、6倍、6.5倍、7倍、7.5倍、8倍、8.5倍、9倍、9.5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。在一些情况下，共定位测定是免疫组织化学测定。在一些情况下，免疫组织化学测定包括抗PV抗体。在一些情况下，如下文实施例5所示地进行共定位测定。在一些情况下，转基因编码离子通道亚基、神经递质调节因子、DNA结合域、基因编辑蛋白或其变体或功能片段。在一些情况下，离子通道亚基是钠离子通道的 α 亚基或 β 亚基，或钾离子通道的亚基。在一些情况下，转基因包括下列任一项：(i) SEQ ID NO:37-43；(ii) 其功能片段；或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下，使用BLAST确定序列同一性。在一些情况下，转基因包括(i) SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1或KV3.3；(ii) 其功能片段；或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下，转基因是神经递质调节因子，包含(i) STXBP1，(ii) 其功能片段，或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下，转基因包括调节内源基因的表达的DNA结合蛋白。在一些情况下，内源基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3或STXBP1。在一些情况下，转基因包括DNA结合蛋白，该DNA结合蛋白包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如，核酸酶、限制酶、重组酶等)的DNA结合域，例如，核酸酶失活的Cas (dCas)、失活的类转录激活因子效应物核酸酶，或核酸酶失活的锌指蛋白。在一些情况下，转基因包含与转录调节域(例如，转录激活或阻遏域)连接的DNA结合域。在一些情况下，

基因编辑蛋白是Cas蛋白。在一些情况下，组合的调控元件的大小小于2.5kb、小于2kb、小于1.5kb、小于1kb或小于500bp。在一些情况下，非PV细胞包括CNS中的一种或多种非PV细胞类型。在一些情况下，非PV细胞包括一种或多种兴奋性神经元、多巴胺能神经元、星形胶质细胞、小胶质细胞和运动神经元。在一些情况下，核酸盒是线性构建体。在一些情况下，核酸盒是载体。在一些情况下，载体是质粒。在一些情况下，载体是病毒载体。在一些情况下，病毒载体是腺伴随病毒(AAV)载体。在一些情况下，AAV载体是AAV1、AAV8、AAV9、scAAV1、scAAV8或scAAV9。在一些情况下，病毒载体是慢病毒载体。

[0009] 一方面，本文公开的任何核酸盒的调控元件包含来自GAD2、GAD1、SYN1、NKX2.1、DLX1、DLX5/6、SST、PV和/或VIP的转录起始位点的10kb以内的小于600bp的连续序列。

[0010] 一方面，一种治疗有需要的受试者的神经系统病症或病况的方法，包括递送治疗有效量的本文公开的任何核酸盒。在一些情况下，神经系统病症或病况是精神障碍(例如，精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病)；自闭症谱系障碍(例如，脆性X综合征、雷特综合征)；癫痫(例如，慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化)；或神经变性(例如，阿尔茨海默病、帕金森病)。在一些情况下，神经系统病症或病况为Dravet综合征或阿尔茨海默病。在一些情况下，神经系统疾病或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。

[0011] 一方面，一种增加转基因在CNS中的PV神经元中的选择性表达的方法包括使细胞与本文公开的核酸盒接触。

[0012] 在一些方面，本公开内容考虑到一种将任何转基因的表达靶向CNS中的PV神经元的方法，该方法包括将一个或多个PV神经元选择性调控元件与转基因可操作地连接。在一些情况下，每个调控元件包含(i)SEQ ID NO:1-32的序列，(ii)其功能片段或组合，或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下，使用BLAST确定序列同一性。在一些情况下，如通过共定位测定所测量的，调控元件导致转基因在PV神经元中的选择性表达，该选择性表达大于相同转基因在与非选择性调控元件可操作地连接时的表达。在一些情况下，免疫组织化学测定包括抗PV抗体(例如，如下面实施例5中所述)。在一些情况下，非选择性调控元件是组成型启动子。在一些情况下，非选择性调控元件是CAG、EF1 α 、SV40、CMV、UBC、PGK和CBA中的任一个。在一些情况下，如通过共定位测定所测量的，调控元件导致转基因在PV神经元中的选择性表达，该选择性表达的水平是非选择性调控元件与转基因可操作地连接时的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍或至少20倍。在一些情况下，调控元件导致在PV神经元中的选择性表达，该选择性表达比在转基因与非选择性调控元件可操作地连接时在PV神经元中的表达高至少2%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。在一些情况下，调控元件导致在PV神经元中的选择性表达该选择性表达是CNS中PV神经元的自然分布的预期表达的约1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍、6倍、6.5倍、高7倍、7.5倍、8倍、8.5倍、9倍、9.5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。在一些情况下，转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1、DNA结合蛋白、

基因编辑蛋白或其功能片段中的任一种。在一些情况下，调控元件和转基因在AAV中。在一些情况下，AAV是AAV1、AAV8、AAV9、scAAV1、scAAV8或scAAV9。

[0013] 另一方面，本公开内容考虑到一种治疗有需要的受试者的神经系统病况或病症的方法，该方法包括使细胞与核酸盒接触，该核酸盒包含：与转基因可操作地连接的一个或多个调控元件，该调控元件导致相对于在CNS中的一种或多种非PV细胞中，转基因在CNS中的PV神经元中选择性表达。在一些情况下，每个调控元件包含(i) SEQ ID NO:1-32的序列，(ii) 其功能片段或组合，或(iii) 与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下，使用BLAST确定序列同一性。在一些情况下，转基因是电压门控离子通道亚基，或其变体或功能片段。在一些情况下，该亚基是钠离子通道的β亚基。在一些情况下，该亚基是钠离子通道的α亚基。在一些情况下，该亚基为钾离子通道的亚基。在一些情况下，转基因是下列任一项：(i) SCN1A、SCN1B、SCN2B、KV3.1或KV3.3；(ii) 其功能片段；或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下，转基因是DNA结合蛋白。在一些情况下，该DNA结合蛋白调节内源基因。在一些情况下，该内源基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或STXBP1。在一些情况下，转基因是DNA结合蛋白，其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白（例如，核酸酶、限制酶、重组酶等）的DNA结合域，例如，核酸酶失活的Cas（dCas）、失活的类转录激活因子效应物核酸酶，或核酸酶失活的锌指蛋白。在一些情况下，转基因包括与转录调节域（例如，转录激活或阻遏域）连接的DNA结合域。在一些情况下，转基因是基因编辑蛋白。在一些情况下，基因编辑蛋白是Cas蛋白，例如Cas9。在一些情况下，神经系统病况或病症与SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或STXBP1中任一个的单倍性不足或突变有关。在一些情况下，神经系统病况或病症是癫痫、神经变性、滔蛋白病变或神经元兴奋不足。在一些情况下，神经系统病况或病症是Dravet综合征。在一些情况下，神经系统病况或病症是阿尔茨海默病。在一些情况下，神经系统病况或疾病是精神障碍（例如，精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病）；自闭症谱系障碍（例如，脆性X综合征、雷特综合征）；癫痫（例如，慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化）；或神经变性（例如，阿尔茨海默病、帕金森病）。在一些情况下，神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。在一些情况下，如通过共定位测定所测量的，本文公开的调控元件导致转基因在PV神经元中的选择性表达，该选择性表达高于相同转基因在与非选择性调控元件可操作地连接时的表达。在一些情况下，非选择性调控元件是组成型启动子。在一些情况下，非选择性调控元件是CAG、EF1 α 、SV40、CMV、UBC、PGK和CBA中的任一个。在一些情况下，如通过共定位测定所测量的，调控元件导致在PV神经元中的选择性表达，所述选择性表达的水平是非选择性调控元件与转基因可操作地连接时的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍或至少20倍。在一些情况下，调控元件导致在PV神经元中的选择性表达，该选择性表达比在转基因与非选择性调控元件可操作地连接时在PV神经元中的表达高至少2%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%。在一些情况下，调控元件

导致在PV神经元中的选择性表达,该选择性表达是CNS中PV神经元的自然分布的预期表达的约1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍、6倍、6.5倍、7倍、7.5倍、8倍、8.5倍、9倍、9.5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。在一些情况下,核酸盒在AAV中。在一些情况下,AAV是AAV1、AAV8、AAV9、scAAV1、scAAV8或scAAV9。

[0014] 在一方面,本公开内容提供了一种治疗Dravet综合征的方法,其包括使细胞与包含转基因的AAV接触,其中所述转基因是(i)SCN1A、SCN2A、SCN8A、SCN1B、SCN2B或DNA结合蛋白,(ii)其功能片段,或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下,使用BLAST测量序列同一性。在一些情况下,DNA结合蛋白调节内源基因。在一些情况下,DNA结合蛋白是转录调节物。在一些情况下,转基因是DNA结合蛋白,其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如,核酸酶、限制酶、重组酶等)的DNA结合域,例如,核酸酶失活的Cas(dCas)、失活的类转录激活因子效应物核酸酶,或核酸酶失活的锌指蛋白。在一些情况下,DNA结合域与转录调节域(例如,转录激活或阻遏域)连接。在一些情况下,转基因包括基因编辑蛋白,例如Cas蛋白Cas9。在一些情况下,内源基因是SCN1A、SCN2A、SCN8A、SCN1B或SCN2B。在一些情况下,AAV还包含与转基因可操作地连接的调控元件,其为一种或多种PV神经元选择性调控元件,或本文公开的一种或多种选择性调控元件。在一些情况下,每个调控元件独立地包含(i)SEQ ID NO:1-32的序列,(ii)其功能片段或组合,或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。

[0015] 另一方面,本公开内容提供了一种治疗阿尔茨海默病的方法,其包括使细胞与包含转基因的AAV接触,其中所述转基因是下列任一项:(i)SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1或DNA结合蛋白;(ii)其功能片段;或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下,使用BLAST测量序列同一性。在一些情况下,DNA结合蛋白调节内源基因。在一些情况下,内源基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或STXBP1。在一些情况下,转基因是包含转录调节物的DNA结合蛋白。在一些情况下,转基因是DNA结合蛋白,其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如,核酸酶、限制酶、重组酶等)的DNA结合域,例如,核酸酶失活的Cas(dCas)、失活的类转录激活因子效应物核酸酶,或核酸酶失活的锌指蛋白。在一些情况下,DNA结合域与转录调节域(例如,转录激活或阻遏域)连接。在一些情况下,转基因包括基因编辑蛋白,例如Cas蛋白Cas9。在一些情况下,AAV还包含与转基因可操作地连接的一种或多种PV神经元选择性调控元件或一种或多种本文公开的调控元件。在一些情况下,每个调控元件独立地包括(i)SEQ ID NO:1-32的序列,(ii)其功能片段或组合,或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。

[0016] 援引并入

[0017] 本说明书中提到的所有出版物、专利和专利申请均通过引用并入本文,其程度如

同明确且单独地指出每个单独的出版物、专利或专利申请通过引用并入。

附图说明

[0018] 本发明的新颖性特征在所附的权利要求书中具体阐述。通过参考以下对利用到本发明原理的说明性实施方案加以阐述的详细描述以及附图,将获得对本发明的特征和优点的更好的理解。在这些附图中:

[0019] 图1示出了在用重组AAVDJ载体治疗后,SCN1A杂合小鼠中癫痫发作的频率(癫痫发作/12小时间隔),该重组AAVDJ载体包含与调控元件可操作地连接的SCN1B或eGFP,该调控元件包含SEQ ID NO:32的序列。该图示出了每日记录的平均值,误差线代表平均值的标准误差。

[0020] 图2示出了以下不同小鼠的高 γ 功率(50-100Hz):野生型对照(WT)、未治疗的转基因APP/PS1小鼠(APP/PS1)或用包含SCN1B的rAAV治疗的转基因APP/PS1小鼠(APP/PS1+SCN1B),该SCN1B与包含SEQ ID NO:32序列的调控元件可操作地连接。

[0021] 图3A示出了幼体在接受包含eGFP转基因的AAV9的新生儿全身注射之后,CNS细胞的免疫荧光共定位测定,该eGFP转基因与包含SEQ ID NO:1或SEQ ID NO:8序列的调控元件可操作地连接。包含与CAG可操作地连接的eGFP转基因的AAV9被用作对照。下排的图像示出了eGFP+细胞。中间排的图像示出了用抗PV抗体染色的PV+细胞。上排的图像(合并)为PV+、eGFP+荧光(具有代表性的eGFP+和PV+细胞,示为白色或浅灰色的细胞,如箭头所指)和DAPI+的叠加。

[0022] 图3B示出了图3A所示的免疫荧光共定位研究的定量,其中如通过免疫荧光共定位测定所测量的,在PV细胞中的选择性表达以与CAG对照相比,同时是PV+的eGFP+细胞的百分比表示。

[0023] 图4A示出了在全身注射包含eGFP转基因的AAV9后,成年小鼠CNS细胞的免疫荧光共定位测定,该eGFP转基因与包含SEQ ID NO:1或SEQ ID NO:8序列的调控元件可操作地连接。包含与EF1 α 可操作地连接的eGFP转基因的AAV9被用作对照。下排的图像示出了eGFP+细胞。中间排的图像示出了用抗PV抗体染色的PV+细胞。上排的图像(合并)示出了PV+eGFP+荧光(具有代表性的eGFP+和PV+细胞,示为白色或浅灰色的细胞,如箭头所指)和DAPI+的叠加。

[0024] 图4B显示了图4A所示的免疫荧光共定位研究的定量,其中如通过免疫荧光共定位测定所测量的,在PV细胞中的选择性表达以与EF1 α 对照相比也是PV+的eGFP+细胞的百分比表示。

[0025] 图5A-5F示出了在直接CNS注射包含eGFP转基因的AAVDJ之后,成年小鼠CNS细胞的免疫荧光共定位测定,该eGFP转基因与包含SEQ ID NO:2-22序列的调控元件可操作地连接。下排的图像示出了eGFP+细胞。中间排的图像示出了用抗PV抗体染色的PV细胞。上排的图像(合并)示出了PV+、eGFP+荧光(具有代表性的eGFP+和PV+细胞,示为白色或浅灰色的细胞,如箭头所指)和DAPI+的叠加。图5A示出了用AAVDJ进行的免疫荧光共定位测定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4和SEQ ID NO:5之一。图5B示出了用AAVDJ进行的免疫荧光共定位测定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:6、SEQ ID NO:7和SEQ ID NO:9之一。图5C示出了用AAVDJ进行的免疫荧光共定位测

定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12和SEQ ID NO:13之一。图5D示出了用AAVDJ进行的免疫荧光共定位测定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16和SEQ ID NO:17之一。图5E示出了用AAVDJ进行的免疫荧光共定位测定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20和SEQ ID NO:21之一。图5F示出了用AAVDJ进行的免疫荧光共定位测定,该AAVDJ包含与eGFP可操作地连接的SEQ ID NO:22或SEQ ID NO:34,其中SEQ ID NO:34是先前已表征的非选择性调控元件,并用作对照以供比较。

[0026] 图6示出了图5A-5F所示的免疫荧光共定位研究的定量,其中如通过免疫荧光共定位测定测量的,在PV细胞中的选择性表达以与SEQ ID NO:34相比也是PV+的eGFP+细胞的百分比表示。

[0027] 图7示出了本公开内容的包含RE的表达盒的实例的示意图,该RE为,例如,增强子、启动子和稳定性元件。RE可以位于表达盒中转基因的上游和/或下游,表达盒可以是质粒、载体或病毒载体。

具体实施方式

[0028] 本公开内容考虑到组合物,和在基因疗法中使用该组合物来治疗与中枢神经系统(CNS)有关的疾病或病况的方法,该疾病或病况为,例如,Dravet综合征、阿尔茨海默病、癫痫和/或癫痫发作。

[0029] 基因疗法可以替代、修饰、删除或添加基因或特定核酸序列(例如表达盒),从而在细胞中发挥治疗作用。在一些情况下,基因疗法用于将产生或导致治疗效果的表达盒递送至细胞中。在一些情况下,包含含有表达盒的病毒载体的病毒,例如AAV,可用于将转基因递送至细胞中。该表达盒可包含在细胞中表达时提供治疗效果的转基因。

[0030] 基因疗法中的一个挑战是确保转基因在合适且感兴趣的细胞类型或靶细胞类型中表达,以实现或靶向基因表达。靶向基因疗法的传统方法通常依赖于递送方法和/或载体(例如改变所使用的病毒或病毒的衣壳序列)。除了相对于在一种或多种非靶细胞类型中,将表达盒靶向靶细胞类型或选择性地在靶细胞类型中表达以外,该领域的另一挑战是增加在靶细胞类型或组织中的基因表达以发挥治疗效果,尤其是当基因较大时。

[0031] 本公开内容提供了为非编码核苷酸序列的多个调控元件,其可以与任何转基因可操作地连接,以增加或改善CNS例如PV神经元中转基因表达的选择性。通过使用本文公开的一个或多个调控元件提高基因表达的选择性,可以改善基因疗法的功效,降低产生治疗效果所需的有效剂量,使不良作用或脱靶效应最小化和/或增加患者的安全性和/或耐受性。

[0032] 一方面,可将一个或多个调控元件与表达盒中的任何转基因可操作地连接,以调节细胞中的基因表达,例如相对于一种或多种非靶细胞类型或组织(例如非PV CNS细胞类型),将转基因的表达靶向靶细胞类型或组织(例如PV细胞)。在一些情况下,转基因在靶细胞类型或组织中的靶向表达包括在靶细胞类型或组织中基因表达的增加。与转基因可操作地连接的一个或多个调控元件可以是表达盒的一部分,该表达盒可以是线性或环状构建体、质粒、载体、病毒载体,例如腺伴随病毒(AAV)载体。该表达盒可适于基因疗法或递送至受试者(例如,人类、患者或哺乳动物)。在一些情况下,将一个或多个调控元件与基因可操作地连接导致该基因在CNS中的靶组织或细胞类型,例如小清蛋白(PV)神经元中的靶向表

达。在一些情况下,一个或多个调控元件(例如,SEQ ID N0:1-32、其功能片段或组合,或与其具有至少80%、至少90%、至少95%或至少99%序列同一性的序列)提高在CNS中的靶组织或细胞类型,例如PV神经元中基因表达的选择性。在一些情况下,基因疗法包含本文公开的一个或多个调控元件,其中该调控元件与转基因可操作地连接并驱动转基因在PV神经元中的选择性表达。

[0033] 在一些情况下,PV神经元中基因的选择性表达被用于治疗与内源基因的单倍性不足和/或遗传缺陷相关的疾病或病况,其中遗传缺陷可以是该基因突变或基因失调。这种遗传缺陷可导致基因产物的水平降低和/或基因产物的功能和/或活性受损。在一些情况下,表达盒包含基因、亚基、其变体或功能片段,其中来自表达盒的基因表达用于治疗与遗传缺陷、功能和/或活性受损和/或内源基因失调有关的疾病或病况。在一些情况下,该疾病或病况是Dravet综合征、阿尔茨海默病、癫痫症、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作。

[0034] 在一些情况下,转基因是离子通道或神经递质调节因子、DNA结合蛋白,或其亚基、变体或功能片段。在一些情况下,转基因是钠离子通道 α 亚基、钠离子通道 β 亚基或其变体或功能片段。在一些情况下,转基因是钾离子通道或其亚基。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白(例如,调节内源基因的表达的DNA结合蛋白)或其变体或功能片段。在一些情况下,转基因包括与SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白或其变体或功能片段中任一个具有至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性的序列。在一些情况下,转基因是调节内源基因的表达的DNA结合蛋白,如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3和STXBP1中的任一个。

[0035] 如本文所用,除非上下文另外明确指出,单数形式“一个”、“一种”和“该”包括复数个指代物。此外,就详细描述和/或权利要求中使用的术语“包括”、“具有”、“带有”或其变体而言,这些术语旨在是包含性的,类似于术语“包含”。

[0036] 术语“约”或“大约”是指在本领域普通技术人员确定的特定值的可接受误差范围内,该误差范围部分将取决于如何测量或确定该值,即测量系统的局限性。例如,按照本领域的实践,“约”可以指在一个或多个标准偏差内。或者,“约”可以指给定值的至多20%、至多15%、至多10%、至多5%或至多1%的范围。

[0037] 术语“确定”、“测量”、“评价”、“评估”、“测定”、“分析”及其语法等同形式在本文中可以互换使用,以指代任何形式的测量,并且包括确定元素是否存在(例如检测)。这些术语可以包括定量和/或定性确定。评估可以是相对的,也可以是绝对的。

[0038] 术语“表达”是指从DNA模板转录核酸序列或多核苷酸的过程(例如翻译成mRNA或其他RNA转录物)和/或转录的mRNA随后翻译成肽、多肽或蛋白质的过程。转录物和编码的多肽可统称为“基因产物”。如果多核苷酸来源于基因组DNA,则表达可包括在真核细胞中剪接mRNA。

[0039] 如本文所用,“可操作连接”、“可操作的连接”、“可操作地连接”或其语法等同物是指例如启动子、增强子、多聚腺苷酸化序列等遗传元件的并列,其中元件之间的关系允许它们以预期的方式发挥作用。例如,如果可包含启动子和/或增强子序列的调控元件帮助启动编码序列的转录,则该调控元件可操作地连接至编码区。只要维持这种功能关系,在调控元

件与编码区之间可以有间插的残基。

[0040] 如本文所用，“载体”是指包含多核苷酸或与多核苷酸有关，且可用于介导多核苷酸向细胞递送的大分子或大分子复合物。载体的实例包括质粒、病毒载体、脂质体和其他基因递送载体。载体通常包含与基因可操作地连接的遗传元件，例如调控元件，以促进基因在靶标中表达。调控元件和与其可操作地连接以供表达的一个或多个基因的组合被称为“表达盒”。

[0041] 术语“AAV”是腺伴随病毒的缩写，可用于指代病毒本身或其衍生物。该术语涵盖所有血清型、亚型、以及天然存在形式和重组形式，除非另有要求。缩写“rAAV”是指重组腺伴随病毒，也称为重组AAV载体（或“rAAV载体”）。术语“AAV”包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、rh10及其杂合物，禽AAV、牛AAV、犬AAV、马AAV、灵长类动物AAV、非灵长类动物AAV和羊AAV。AAV的各种血清型的基因组序列，以及天然末端重复序列（TR）、Rep蛋白和衣壳亚基的序列是本领域已知的。这样的序列可以在文献中或公共数据库如GenBank中找到。如本文所用，“rAAV载体”是指包含非AAV来源的多核苷酸序列（即与AAV异源的多核苷酸）的AAV载体，该序列通常是用于细胞遗传转化的感兴趣的序列。通常，异源多核苷酸的侧翼是至少一个，并且通常是两个，AAV反向末端重复序列（ITR）。术语rAAV载体涵盖rAAV载体颗粒和rAAV载体质粒二者。rAAV载体可以是单链(ssAAV)或自互补(scAAV)。“AAV病毒”或“AAV病毒颗粒”或“rAAV载体颗粒”是指由至少一个AAV衣壳蛋白和衣壳化的多核苷酸rAAV载体组成的病毒颗粒。如果颗粒包含异源多核苷酸（即野生型AAV基因组以外的多核苷酸，例如要递送至哺乳动物细胞的转基因），则其通常称为“rAAV载体颗粒”或简称为“rAAV载体”。因此，rAAV颗粒的产生必然包括rAAV载体的产生，因为这种载体包含在rAAV颗粒内。

[0042] 如本文所用，术语“治疗”、“疗法”等是指获得所需的药理和/或生理作用，包括但不限于减轻、延迟或减慢进展，减少影响或症状，预防发作，抑制、减轻疾病或病症的发作，获得关于疾病、病症或医学状况的有益或期望结果，例如治疗益处和/或预防益处。如本文所用，“治疗”涵盖对哺乳动物特别是人类的疾病的任何治疗，并且包括：(a) 在可能易患疾病或有发展出该疾病的风险但尚未被诊断为患有该疾病的受试者中预防该疾病的发生；(b) 抑制疾病，即阻止其发展；以及(c) 减轻疾病，即引起疾病消退。治疗益处包括根除或改善所治疗的潜在病症。同样，通过消除或改善与潜在病症相关的一种或多种生理症状而获得治疗益处，使得在受试者中观察到改善，尽管受试者可能仍患有该潜在病症。在一些情况下，出于预防的目的，将组合物施用于有发展出特定疾病的受试者，或报告疾病的一种或多种生理症状的受试者，即使可能尚未对该疾病作出诊断。本公开内容的方法可以用于任何哺乳动物。在一些情况下，该治疗可导致症状减轻或停止（例如，癫痫发作的频率或持续时间减少）。预防作用包括延迟或消除疾病或病况的出现，延迟或消除疾病或病况的症状的发作，减慢、停止或逆转疾病或病况的进展，或其任何组合。

[0043] 术语“有效量”或“治疗有效量”是指本文所述的组合物的量，其足以影响预期的应用，包括但不限于如下定义的疾病治疗。治疗有效量可以根据预期的治疗应用（体内）或所治疗的受试者和疾病状况而变化，例如，受试者的体重和年龄、疾病状况的严重程度、施用方式等，这很容易由本领域普通技术人员确定。该术语还适用于将在靶细胞中诱导特定应答的剂量。具体剂量将根据所选的特定组合物、要遵循的给药方案、是否与其他化合物联合

施用、施用时间、该剂量所施用至的组织以及携带该剂量的物理递送系统而变化。

[0044] 核苷酸或肽序列的“片段”是指小于被认为是“全长”序列的序列。

[0045] 分子的“变体”是指此类序列的等位基因变异,即,在结构和生物活性上与整个分子或其片段基本相似的序列。

[0046] 术语“功能片段”旨在包括分子的“片段”、“变体”、“类似物”或“化学衍生物”。

[0047] DNA或蛋白质序列的“功能片段”具有该序列的至少生物活性片段,该生物活性片段是指保留与全长DNA或蛋白质序列的生物活性基本相似的生物活性(在功能或结构上)的片段。DNA序列的生物活性可以是其以已知归因于全长序列的方式影响表达的能力。例如,调控元件的功能片段将保留其作为全长RE影响转录的能力。

[0048] 术语“受试者”和“个体”在本文可互换使用,以指脊椎动物,优选为哺乳动物,更优选为人类。“受试者”是指动物,例如哺乳动物,例如人类。本文所述的方法可用于人类治疗、兽医应用和/或疾病或病况的动物模型的临床前研究。在一些情况下,受试者是哺乳动物,并且在一些情况下,受试者是人类。

[0049] 术语“体内”是指在受试者体内发生的事件。

[0050] 术语“体外”是指在受试者体外发生的事件。例如,体外测定涵盖在受试者体外进行的任何测定运行。体外测定包括其中使用活细胞或死细胞的基于细胞的测定。体外测定还包括其中不使用完整细胞的无细胞测定。

[0051] 例如出于评估同一性、突变或测试序列的一个或多个位置相对于参考序列的一个或多个指定位置所处于的位置的目的,可以通过任何合适的比对算法来进行序列比较序列比对,包括但不限于Needleman-Wunsch算法(例如,参见www.ebi.ac.uk/Tools/psa/emboss_needle/上可用的EMBOSS Needle比对器,任选使用默认设置)、BLAST算法(例如,参见blast.ncbi.nlm.nih.gov/Blast.cgi上可用的BLAST比对工具,任选使用默认设置),以及Smith-Waterman算法(例如,参见www.ebi.ac.uk/Tools/psa/emboss_water/上可用的EMBOSS Water比对器,任选使用默认设置)。可以使用所选算法的任何合适参数,包括默认参数,来评估最佳比对。

[0052] 通常,可以互换使用的“序列同一性”或“序列同源性”是指两个多核苷酸或多肽序列的精确核苷酸-核苷酸或氨基酸-氨基酸对应性。通常,用于确定序列同一性的技术包括确定多核苷酸的核苷酸序列和/或确定由此编码的氨基酸序列,并将这些序列与第二核苷酸或氨基酸序列进行比较。两个或更多个序列(多核苷酸或氨基酸)可以通过确定它们的“同一性百分比”,也称为“同源性百分比”,来进行比较。与可能是较长分子(例如,多核苷酸或多肽)内的序列的参考序列(例如,核酸或氨基酸序列)的同一性百分比可以计算为两个最佳比对序列之间精确匹配的数目除以参考序列的长度并乘以100。同一性百分比还可以例如通过使用高级BLAST计算机程序比较序列信息来确定,该计算机程序包括2.2.9版,可从美国国立卫生研究院获得。BLAST程序基于Karlin和Altschul, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990) 的比对方法,有关讨论见Altschul等人,J. Mol. Biol. 215:403-410 (1990); Karlin和Altschul, Proc. Natl. Acad. Sci. USA 90:5873-5877 (1993); 和Altschul等人,Nucleic Acids Res. 25:3389-3402 (1997)。简而言之, BLAST程序将同一性定义为相比对符号(即核苷酸或氨基酸)的数目除以两个序列中较短者的符号总数。该程序可以用于确定被比较序列的整个长度上的同一性百分比。提供默认参数以优化短查询序列的搜

索,例如使用blastp程序。该程序还允许使用SEG过滤器来屏蔽查询序列区段,如由Wootton和Federhen,Computers and Chemistry 17:149-163(1993)的SEG程序确定的。所需的序列同一性程度的范围约为80%至100%,以及介于两者之间的整数值。通常,公开的序列与要求保护的序列之间的同一性百分比为至少80%、至少85%、至少90%、至少95%、至少98%或至少99%。通常,精确匹配表示在参考序列的长度上100%的同一性。在一些情况下,提及序列同一性百分比是指使用BLAST(基本局部比对搜索工具)测量的序列同一性。在其他情况下,ClustalW可用于多序列比对。

[0053] 除非另有说明,否则本文所用的所有术语具有对于本领域技术人员而言相同的含义,并且本发明的实施将采用分子生物学、微生物学和重组DNA技术中的常规技术,这些技术均属于本领域技术人员的知识范畴。

[0054] 调控元件

[0055] 调控元件是能够影响(例如增加或减少)基因表达和/或使基因(例如eGFP、转基因或治疗性基因等报告基因)在特定的兴趣组织或细胞类型中选择性表达的核酸序列或遗传元件。在一些情况下,调控元件可以是转基因、内含子、启动子、增强子、UTR、绝缘子、阻遏子、反向末端重复序列(ITR)、长末端重复序列(LTR)、稳定性元件、翻译后应答元件,或polyA序列或其组合。在一些情况下,调控元件是启动子或增强子或其组合。在一些情况下,调控元件来源于人类序列。

[0056] 在一些情况下,感兴趣的细胞类型是PV神经元。调控元件可以在DNA和/或RNA水平上起作用。调控元件可以在感兴趣的细胞类型中起到调节基因表达选择性的作用。调控元件可以在基因表达的转录阶段、转录后阶段或翻译阶段起到调节基因表达的作用。调控元件包括但不限于启动子、增强子、阻遏子、沉默子和绝缘子序列。在RNA水平下,调节可以在翻译(例如稳定mRNA以进行翻译的稳定性元件)、RNA切割、RNA剪接和/或转录终止的水平下发生。在一些情况下,调控元件可以将转录因子募集到编码区,该转录因子增加感兴趣的细胞类型中基因表达的选择性。在一些情况下,调控元件可以提高产生RNA转录物的速率,提高产生的RNA的稳定性,和/或提高从RNA转录物合成蛋白质的速率。在一些情况下,调控元件可以防止RNA降解和/或增加其稳定性以促进蛋白质合成。在一些情况下,调控元件抑制脱靶细胞类型中的转录和/或翻译过程。在一些情况下,脱靶细胞类型包括但不限于兴奋性神经元、非PV CNS细胞类型和非神经元CNS细胞类型。

[0057] 多种测定,包括但不限于DNA酶超敏反应、ATAC-Seq和ChIP-Seq,可用于鉴定假定的非编码调控元件(RE)。这些测定中每一种的酶促反应均优先靶向开放/可及的染色质状态,该状态被认为可预测调控元件。为了发现细胞类型选择性调控元件,可以测定感兴趣的目标细胞类型(例如小清蛋白神经元)的开放染色质序列,并将其与非目标细胞类型(例如兴奋性神经元)的开放染色质序列进行比较。可以应用其他过滤器来进一步优化靶标选择,包括与细胞类型选择性基因的接近性、物种保守性和/或序列基序,如转录因子结合位点。目标细胞类型中的独特鉴定出的DNA序列可以合成并克隆到表达载体中。调控元件的选择性可以使用免疫组织化学法确定,以将与已知细胞类型选择性蛋白的共定位定量。

[0058] 例如,一种分离细胞类型选择性调控元件的方法包括来自动物模型的脑组织或感兴趣的细胞类型中分离细胞核,这可以通过使用分离感兴趣的组织或细胞类型的亲和纯化方法(例如使用包被有抗PV抗体的珠来分离PV神经元)来实现;使用高通量天然引物和DNA

合成从细胞核中的开放染色质区域产生序列池；对序列池进行测序以鉴定在感兴趣的组织或细胞类型中驱动基因表达的推定序列；以及在体外细胞系和/或动物模型的报道基因系统中验证选择性表达。

[0059] 另一种用于鉴定在感兴趣的组织或细胞类型中具有选择性的候选调控元件的方法包括使用R26-CAG-LSL-Sun1-sfGFP-Myc基因敲入小鼠来收获感兴趣的组织或细胞类型；使用亲和纯化从该小鼠品系的新皮质分离GFP+/Myc+细胞核，例如，使用抗GFP或抗Myc抗体和蛋白G包被的磁珠从新皮质分离细胞核。可以将来自纯化细胞核或全部新皮层细胞核的核RNA转换为cDNA，并使用Nugen Ovation RNA-seq系统V2 (Nugen 7102) 进行扩增，然后使用Illumina HISEQ® 2500进行测序。可以将来自纯化细胞核的基因组DNA片段化并用于制作MethylC-seq文库，该文库可以使用Illumina HISEQ® 2000进行测序。为了产生ATAC-seq文库，使用Tn5转座酶 (Illumina FC-121-1030) 将与磁珠结合的细胞核转座。在9-12个循环的PCR扩增后，使用Illumina HISEQ® 2500对文库进行测序。为了产生ChIP-seq文库，可以使用微球菌核酸酶将兴奋性神经元细胞核消化成单核小体，然后用盐提取染色质，并构建可以在Illumina HISEQ® 2500上测序的天然ChIP和文库。在对这些文库进行测序后，映射序列以鉴定，例如，富含CG的区域中细胞类型特异性低甲基化的相关性、组蛋白修饰、转录因子结合位点以及与高度表达的转录因子相关的模式。来自上述多个测定和/或文库的重叠特征和相关性提供了证据来将这些基因组区域内的候选序列鉴定为与从新皮层分离的细胞中的选择性表达和/或高表达相关的潜在调控元件。例如，在methylC-seq文库中检测到的低甲基化、ChIP测定与同一区域中转录因子结合基序的富集之间强烈重叠为特征的基因组区域提供了表明该基因组区域包含对分离的组织或细胞类型具有选择性的推定调控元件序列的收敛数据。在另一个实例中，为了鉴定候选PV神经元选择性调控元件，可以分离PV神经元，并从分离的PV细胞中纯化细胞核，使得在上述多个序列测定中被鉴定为有活性的基因组序列很可能是PV细胞选择性调控元件，例如在ATAC-seq测定中被鉴定为有活性的基因组区域(对应于开放染色质区域)，在RNA-seq中被鉴定为活性的基因组区域(指示该区域中的活跃基因表达和低DNA甲基化模式)，以及在methylC-seq测定中被鉴定为有活性的基因组区域(从感兴趣的细胞类型产生单碱基分辨率的甲基化组图谱)。

[0060] 一旦将候选基因组区域鉴定为在感兴趣的细胞类型中具有选择性活性，就可以使用PCR方法产生该区域内的序列，并在体外和/或体内的额外测定中进行测试，以验证该序列的组织或细胞类型选择性。这样的验证测定包括免疫组织化学共定位测定，其中抗体或任何可检测的标记用于标记感兴趣的细胞类型，并且第二可检测的标记，例如荧光转基因，与推定的调控元件可操作地连接。包含这些元件的表达盒在体外和/或体内递送到细胞中。可以通过测量感兴趣的细胞类型(通过来自其标记的标记，例如抗PV抗体的可检测信号或荧光来测量)与对应于与调控元件可操作地连接的转基因(例如eGFP或RFP)的表达的第二可检测标记之间的重叠来验证由一个或多个假定调控元件驱动的选择性表达。如果观察到的重叠的量高于用CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)、非选择性调控元件或先前表征的非选择性调控元件等对照替换调控元件时观察到的重叠，则来自两个可检测标记的信号中的重叠表示在标记的细胞类型中的细胞类型选择性。适于在感兴趣的细胞类型中表达可检测标记的各种小鼠品系允许在体内验证调控元件的细胞类型选择性。例

如,可以使用在特定细胞类型中表达Cre的许多小鼠系,因为细胞类型选择性的Cre表达可以在感兴趣的细胞类型中驱动荧光蛋白例如RFP的Cre诱导表达。在体内标记这种感兴趣的细胞类型允许确定与可操作地连接至相同小鼠的荧光或报告基因转基因的推定调控元件相关的细胞类型选择性表达的水平。与共定位测定相似,来自两个标记的信号的重叠超过针对CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件检测到的重叠,则表明测试的调控元件具有细胞类型选择性。在一些情况下,使用的小鼠品系是B6 PV-Cre小鼠(Jackson Laboratory),它是B6 PV-Cre敲入小鼠,在表达小清蛋白的神经元(例如大脑中的中间神经元和背根神经节中的本体感受传入感觉神经元)中表达Cre重组酶,而不会破坏内源性小清蛋白的表达。

[0061] 在针对特定细胞类型验证调控元件的细胞类型选择性时,可以使用各种诱变方法,例如易错PCR方法,来改变此类调控元件的序列,以提高其选择性。在一些情况下,可以将具有细胞选择性的两个或更多个调控元件组合。在一些情况下,组合的调控元件在驱动感兴趣的细胞类型中的基因表达中表现出增强的细胞类型选择性。在一些情况下,将此类调控元件一次截短一个或多个碱基,以确定保留其细胞类型选择性的最小序列量。保留细胞类型选择性的较小调控元件有助于开发包含大转基因的基因疗法,或者有助于解决基因疗法中载体或质粒的克隆能力因希望使用该基因疗法递送的转基因大小而受到限制的情况。

[0062] 本公开内容提供了作为调控元件的多个核苷酸序列。在一些情况下,本文公开的任何一种或多种调控元件导致小清蛋白细胞中基因表达的选择性增加。在一些情况下,本文公开的调控元件是PV细胞选择性的。在一些情况下,PV细胞选择性调控元件与PV细胞中的选择性基因表达超过非PV CNS细胞类型中的表达相关。在一些情况下,PV细胞选择性调控元件与非PV CNS细胞类型中的基因表达降低相关。

[0063] 调控元件的非限制性实例包括SEQ ID NO:1-32,如以下表1所示。

[0064] 表1:本文公开的核酸序列的列表

[0065]

SEQ ID NO:	核酸序列	来源/ 基因组位 置
1	GGAGGAAGCCATCAACTAAACTACAATGACTGTAAGATA CAAAATTGGGAATGGTAACATATTTGAAGTTCTGTTGAC ATAAAGAACATGATATTAATGCCCATGGAAATGAAAGG GCGATCAACACTATGGTTGAAAAGGGGAAATTGTAGA GCACAGATGTGTCGTGGCAGTGTGCTCTAGCAA TACTCAGAGAAGAGAGAGAACATGAAATTCTGATTGGC CCCAGTGTGAGGCCAGATGAGGTTCAGCTGCCAACTTCT CTTTCACATCTTATGAAAGTCATTAAAGCACAACTAACTTT TTTTTTTTTTTTTTTGAGACAGAGTCTGCTCTGTTG CCCAGGACAGAGTCAGTAGTGAUTCAATCTCGGCTCACT GCAGCCTCCACCTCCTAGGCTCAAACGGCCTCCTGCATC AGCCTCCAAGTAGCTGAAATTACAGGAGTGGCCCACCAT GCCAGCTAATTGTATTAAATAGATACGGGGTTTC ACCATATCACCCAGGCTGGTCTCGAACTCCTGGCCTCAAG TGATCCACCTGCCTCGGCCTCCAAAGTGCTGGGATTATA GGCGTCAGCCACTATGCCCAACCCGACCAACCTTTTAA AATAAATATTAAAAATTGGTATTCACATATACTAG T	人类; hg19: chr2: 17162190 0-1716225 80
2	AGTTGGACAAGAACTATAGTTCTAGCTTCTCTGGGTCTC CACCTTGCAGAGAACATGCAAGCTTCATTATCTCATGAGCCA AACTCTCATCATCTCTTCCATATATCTGTCGGTGCCTTC CATGAGTAACACACACAGAAGGAGCACTACACA GGCTGTGTTCTCTTATTATCATAGCTGTTGTCAGAC ATGTGCATTCTGTTCTGCTCAATGCTAAAGGAGTCT CAGGATATGAGAACTGTACCAGCCGAGGCATCAGGAAAC ATGGGTGGAAATTCCCACAGTACTATTGTTCACTGTG ACCTTGGGCCAGTCACATCCCTTCCTGAGGGCTCGATTCC CCAAGCTATAAAAGAACATCTCTTAACCTTTTAGGTC ATGAGTCAGGCCAGCACACTCTCAGGGAGACTCATGAG AGTACAGATCATTCCCAGTAAAAACCATAGTTTATAT CCAGAGGCTTTCTGTAAG	小鼠; mm10:chr 2:3605385 8-3605435 9
3	GGTTCCAGTTCAAGAGGCAGAGCATTGGGGTCCCAGTCA GGAGCTTCCTCTCCGCTCCTAGTTCTCTCTTTAAA AAAAAAATGGGTGATAGTATAGAAAGGAAGCTCTGGGCTC GGGGACCAGGGCCCTGGGATCCCCGCTCCCAGCCACTCGC TCCTGACCCCTCCAGGGACAAGCTCCCCCCCACCCGTCC TTTCCAGGCTGCCACTAGAAGAGATGGGGACGCGTGGTCA	小鼠; chr2:36,09 1,144-36,0 91,966

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	<pre> GCCGCTTCTGTCGCCCCCAGGGAACGGTCTCACGCTGGA GGGGCAGTCCTCGAACAGGACAGTCAGCCAAAGCC AGCCAAGCGCGCGGACGTCTTCAACCGCAGAGCAATTG CAGGTACCCGGCAAGCCCCAAGCGTGTGGCGGGC TTCGGAGTGGCGTGGTTCTCGGGACTTGTGACTCCGCC CCTTGTGCGGGGACCCGCGTGAGGCCGCTCCAAGGATGAA GCTGCCTGGGGCGTGGCCTCGGACCCCTGAGCCTTGATTG GGCGGAGGTCTCAGGGCCCTTCTGCGCCCCACAGGTTATG CAGGCGCAGTCGCGCAGGACAACAACCCGGACCGCGCAG ACGCTGAGAAGCTGGCGGACATGACGGGCTCAGTCGC AGGGTCATCCAGGTGGGGCTCCGGGTCTCGGCCTTCAGG TCTAGGGTGAACCTTAGGAAAGCGCTGAAGCTCGTAGTGG TACGGATGGTCGCGCGTGCACGTGGCCGCCCCCTCCAGT GTGGCCTAAGGACCCCAGTCGGCACGGGTTGACCCTTTTC CTTGATTACTGAGAGTGCAGAGGCTGT </pre>	
[0066]	<pre> 4 TGGTGGGAAGACATGTCCAGGGAAAGAAATGGCCTCCAGA GGCCTGAGGTGGGGAAATGCTGGAGGTGGAGAGAGGAAC AACTGACTGAAAATGAGCTTCACTGTGGCTTAGTAGCCT ATACCAAGTCTAGAGTATAGGGTAGGAGAAGATTAGGAA AGCGATGGGTCTGAGAATGATGTGGCCTGTTGACTTTGT AAACCCAAAGCACCTGGACTAAACCTATGAACAGTGTG GTGCCACCAAAGACTATAATGAGCTCAGGGAACAGAATT CTGTGTGCATGGTATTAAAACTGCTAACTGC AGTCTGGGTGATGCATTGACAAACCAATCCTGGAAAGTAA GAGGCAAGGGCAGCTGGACGGTGAGAGGAGCCTGATGG GAACCAGGCCAAGCAGGGCAGCAGAGGGCATGAAGAGG ATGTGGTGCATCCAGAGACTCACTCATTAGCTGGAGGC CTGCTGGATAGGGTCTGAAGGTTCTGGTATCTGAGTTGGC GGGCTGGGTGAGTGGTGGCTCTGCTTCCCTGAACAGTGTGT GCAAGAGGAAACAGGGTTAAGGGCTAGGACAGTCACAGG TGAGTCAGCCTACAAGAGCAACCTCCCTAGTGCAGA </pre>	小鼠; chr2:36,09 5,396-36,0 96,028
	<pre> 5 GGAGGTCTCCTTGGCCCCGGTTCCAACAAGAGAACGCAA GGCTGTATCTCAATTCTTGAGCCTCTCTGTATTATAGAA GAAAAGTAGGGAAGCCATACGCCCTCTGAGCTTCAGTG TCTCTGTCTCTGCAAATGAGGCTGGGGAGGCTGGGGC GGCGTGAAAGAGGCCCGCGCCAAGCCACCCCCACCTC TGCCCCCTCCCCAGGTCAACAAACCTCATCTGGCACGTGCG GTGCCTCGAGTGCTCCGTGTGTCGCACATCGCTGAGGCAG CAGAATAGCTGCTACATCAAGAACAGGAGATCTACTGC AAGATGGACTACTCAGGTAGGCAGCGGCCATCCGCCAG CAAGCGCTGGAGCATGAACGCCCTGCACACGCGTGCCTAG GCCACTTGTGTGGCTGTGCTCTCCAATTCTGAGCCCTGC TGTTCAGAGTGCACAACGCCGCTCAGCGCACTGGCCGGC CCTCCTACTCAGCACGTCTTACACAGAAGGGAGGCCAGT CTCAGCCTGAGTTCTGGGGGGATCTGCCTCGGGTTCCATCTC CCGATCTGACAGGCCGCTGGCCACGGGCTGGTTCCATCTC TGGTCTTCTGGCCCCGAGCACCAGTGTGTTCTGTTGAGC </pre>	小鼠; mm10:chr 2:3610252 4-3610319 3

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	TCTGATGTCCGAGGCTCTGGCCCCGATCA	
6	<p>[0067]</p> <p>CTCTGGCTACCTCTTATCTTGGCATTACGACAATTCTA ATTGCAGGTAGTTGTGTGTGCGCGTGTGTTTTCCCC CTCAGAGGCTTGGATTGCAAAGGAACTAAGCGATTACTTC AAGAGCCACGGGTTAAGTGCAGGGAGAGGGGGAGAGAG AGGGAAAAAAACCCAATCAAATTCAAATTGCTTCATTAG AGAGACACCGCTTGTGGGAAGGGCTTAAATGCCAC TACAAAGTTAGGACTCATTGTTAGCGCCGGTTATATAA CAGGCAGGGGAGGCCTGGCTCTGACAGCTCCGAGCC AGTTCAGCAGCCGCCGTCGCCTGCATTCCCTCCCCCTCCCC CAGGTGATGGCCCAGCCAGGGTCCGGCTGCAAAGCGACC ACCCGCTGTCTCGAAGGGACCGCTCCGCCTGCCATGGTGA GTCCTTCGGTCCTGCTTCGGCCCCGAGTCCCCCAACAG CACAGGCCAGGGCTCTGGCTCAGCCTCCGGCTACCAAC CTCTACCCCTGCCTGGAAAAGTCCGATAGGAGCCGCCT CTCGTTGAGCCTGGTTTCTGGCCTGGAATGTGAGCTTT GGCTGCTCCTGCACCCAGGATGCGCTGTGTTAAAGTTG GGGGCCGTCCCTTCTCCAATAGGTCCCTTCATTCTTG ACTCCAGCCTAGGGCGCGACATCCCTGGCACATTGCGGTG TCAGTCGGTGCAGGAAACAGATTCAACTCTGAGTAC TCGGCTAACGCCTCGCTGTTCCCTCTCCCATTCAGGCT CAGTCAGACGCAGAGGCCTGGCAGGCCTCTGGACAAG GACGAAGGTAGAGCCTCCCATGTACGCCAGCACACCGT CTGTCTGCTGCCCTCTGCTGCCTCTCCGTGCCGTCT GCCGGCAAGAATATCTGCTCCAGTTGCGGTCTGGAGATCC TGGACCGGTATCTGCTCAAGGTGAGTCAGGGTAGGTGTG CTGCTTGCCTACGGGTGTTGAGCCCCAACAGACTGT GGCACATGCTCAGTGAATTG</p>	小鼠; mm10:chr 2:3610328 6-3610432 8
7	CAAGACTTTAAAAGTTAGATAAATAACAAACATTGA CGGCTTCCATCACATCTAGACTATAATCAAAGATCTAT ATGGTCCAAACGACTTACACTTAACTACCGTCTCCATA TGGCTTCTCCCCATCAGTCATTGCTCTCAGCCATAGTGG CCTCCCTGTTCCCTGGTACAAGGGAAACAACCTCCGTGAG AGGTTCCATTAGCTGCTGTTGCCTGAGATGCTCTGAGCCC ACACCATCTGCTCATTTCTCTCCTCACGTGTCAGTGATTAA GAGGCTGCTCCCTGGCCTCCGTCAAAATTACATCCCTGCC GCTTCCACTCTGCCTCTTATTTCTAAATAGAACTAA CTCACCACTACCCAACATTCTATATAATTGGATATCTGTCC TCTGTTAAATATAATTGACTCAAGAAAGAACGTTGT CACTGCCCTGTCAACCAGACTTTAAACAGTGCCTATCGTGT GGCACATGCTCAGTGAATTG	小鼠; mm10:chr 2:3611431 1-3611481 7
8	TCAACAGGGGACACTTGGAAAGAAGGATGGGGACAGA GCCGAGAGGACTGTTACACATTAGAGAAACATCAGTGACT GTGCCAGCTTGGGGTAGACTGCACAAAAGCCCTGAGGCA GCACAGGCAGGATCCAGTCTGCTGGTCCCAGGAAGCTAAC CGTCTCAGACAGAGCACAAAGCACCGAGACATGTGCCAC AAGGTTGTGTAGAGAGGGTCAGAGGACAGCGTACAGGTC CCAGAGATCAAACCTCACCAAGGCTTGGCAGCAAGC	小鼠; mm10:chr 15:781791 09-781796 10

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	CTTTACCAACCCACCCCCACCCACCCACCCCTGCACGCGC CCCTCTCCCCTCCCCATGGTCTCCCATGGCTATCTCACTTG GCCCTAAAATGTTAAGGATGACACTGGCTGCTGAGTGGAA AATGAGACAGCAGAACAGTAGATTAGGAAAGC CAGAGAAAAAGGCTGTGCTGTTAGAAAGCCAAGGG ACAAGCTAAGATAGGGCCCAAGTAAT	
9	[0068] AAATAGAACTGTGAGATAGGGGAGAGGGGGCAGGAAG GACAAGAGACCCCTGTCTCATTGTGATCCCCACCTGTCTG CTCTGTGGAGGGTACCCATGAGGGCCAGCCCACAGCCCT TAGGTGGACATTGTCTGGTCCTGTCTCACTGTCCCTCCCAG CAGCCCCAGAGGCCAGGAGACAGGGTCTCAGTCCTCACT GAGAGATGTGAAACTGAGGCCAGTGAATGTTGAGGGC CAGGGCATGCCCTGGTGGATGTGACCTGGGTCTCCTTC GCACGGGCTTCCTCCCCGAAGCCGAGCTGAGCATTTGGAG TTGAAATGTTCCGTACTTAGCAATCTGCTCCTCTATTCC CGGGCGGACTTCCGATACTCCGGCTTATGCTGCACTAG ATAAGATGGAGCAGGGAGAGGACACGGCACTACTATGT AACCGGCCTCTGAAAAATGGAGCAGCGGTAGGGCGGA ACAAGACGTCCCTCTACGCATCCCTCTCCTTCCCTGC TAAGGCTGCAGCTGGAGTCAGAGGCAGGGCTGTTCCAATC TGTCTTGATCAGTAACGCAGCCAGCCTCCAGCCTCCGTC AGCCTCCTCATGGCTGAGACCCGGCCTCAGTTCCCCAC TTACATCCCGAGGATCAGAGCCTGTGAGGATGAAATGGG ATAAGGTAGCTGGAACCGTCTGGCAGAGAGCGAGTCCTC AGGACTGTTGATGCCTGTGGCTGCCCTGGCTGACCCCAAG TGACCCCGCCTCCTCATCCTGCAGCAGGAGAA	小鼠; mm10:chr 15:781953 47-781961 34
10	TCTATAGAAATGTGCCCCAGCCTGTTTCCACACTGATA CGCAAGGAATGCATACCAACAGAGAGGGATGAGGGTAGCA TCCAGCCTGCTTCCTGTGTGTCGGGGCGCTACAGCCACAT CTCCCCAGTCCATCTCAGACCGTCACAGAGCTCGCCGAA TGTATAGCTTGTTCTGTGAGACAGGGAGACAGAGCC TTGGGAAGCATAGGTGCTTGCTTCTTGCCCCACTGAGTCTT AGCTGGACTTGACACACCACATGCCTCACAGCCGGCGCAC TTGCATTGTCACCCAGGCCAGTGATGATGGCTCTGCTTG CTTGCTTCTTGCCAACACAGCTCCAGCACCTGTGCC TGGGTTTCACTCCTTAGTTGAACACGTAGTTACTGGGGT TGTAGGGATGGAGCCTTCTGCTCCTCTGGCAAAGTCCT TAGCGGCCTGCTGCCGGGGTGGGGGTGTTCAGGGAGT GGTGTGATGAAGTATGACAG	小鼠; mm10:chr 15:781963 05-781968 06
11	TCTCCAGTTGGAGAAACAGATGCTGTAACGGGCCACAG TATAAAAGAGAGGCCAGACATTGAACACTGTCAACACAGAAG CCTGGCACACTGGAACCTGGCAGTCCAGCTGGAAACAAGG GGTAGAGGCTGAGGCCACTAAGTCACACTGAGGCAGGAGA CATAGGAGCTAAAGCAGCTGAAGGGTGCAGGACAGCTGG GGGGTCTGAAGTGGGCCTCATGCCAGAGCTATGAAGTCA GGGGCTGTAGCCTAGGAGCCTTGGAAAGCCAGCTGGCAAG CTGTGGCCCAAAGACGCTGACTCACCAGGAGGGGCAGC	小鼠; mm10:chr 15:782052 34-782057 66

SEQ ID NO:	核酸序列	来源/ 基因组位 置	
	TGGAGCCAGGCACTCCTAAGGTTCCAGGAAGGGCAGCCT TCCAGGGCTCAGCTAGGGGAGACAGTGTTGACAGCAAGTT GTCAGGCAACTTGAGCTACTGGGCAGCTGGGAAGCTGTCC CTTGGTCCCCAGTATCATCATCACCCCCAGACGCTGCCAC CTGCCTCAGGTCCCACACAGTGATCCTCCATCTTAACAC AACACATGACCAGAGAGA		
12	GTCACCCCTCCCCAAACAACCCCTTCTCTGGTTCGAG AAATTACAGGCATGAAAGATATAAAATCGGGATGCTTGACT TGGGAATATAAAATCACTAAAGCTTGGGGCAGGGGTGGG CGACCTTGTGACCGTCCTGTGCGTGCAGTAATCCTGT GGTCCAGGGGAGAAGAAAAGGCTGTGTTCTGCTCAC AAAGCTGCAGAAACCATTCTTAAGCCAAAAGCACTTCC AGAGAGAGCAGAGCATCCCCAGGCTGCTGGCTCAGCAAG TTCACTGTGCTCAATCTCAGGAAGTGAGGATAAGAGCAGT GCCTGGAGAGTGCCTGGTGTGAGCTGAGGGTTCTGAAC ACATTAAAGCAGGGAGCATGGACCGGGCCTCAGGAGGGG TGTTGAACATCCCTAGGCAGAGGAGTCTAGCTTCTGGGA AAAGATATCAGGTTAACGACACACATGTCCCTGGAATAA GATAATCTTCTGATCACACACTATACACACACACAAAAGCC TGCTC	小鼠; mm10:chr 15:782248 41-782253 64	
[0069]	13	GCCCTCTAGGCCACCTGACCAGGTCCCCCTCAGTCCCCCCC TTCCCACACTCCCACACTCAGCCCCCTCCCCCCCCCGA CCCCTGCAGGATTATCCTGTCTGTGTTCTGACTCAGCCTG GGAGCCACCTGGGCAGCAGGGGCAAGGGTGTCTAGAA GGGACCTGGAGTCCACGCTGGGCAAGCCTGCCCTTCTC CCTCTGTCTTCCGTCCCTGCTGCGGTTCTGCTGAATGTGG TTATTTCTCTGGCTCCTTTACAGAGAATGCTGCTGCTAAT TTTATGTGGAGCTTGAGGCAGTGTAAATTGGAAGCCAGAC ACCCCTGTCAGCAGTGGGCTCCGCTCCTGAGCTGCCATGCT TCCTGCTCTCCTCCGTCCGGCTCCTCATTTCATGCAGCC ACCTGTCCCAGGGAGAGAGGAGTCACCCAGGCCCTCAGT CCGCCCTTAAATAAGAAAGCCTCCGTTGCTCGGCACACA TACCAAGCAGCCGCTGGTGAATCT	小鼠; mm10:chr 15:782413 48-782418 56
14	GTGTTCTCCCTCCCCCTTGGACCCCCGAGACAAGCCAAT AAAATACTCGGCAGGGTGGCTCTCCTTTTTGCCAGT AATAAACAGACTCAGAGCAAGTTAAGGGTCTGGTCCAAG GTCATGGCTGGGATCAGTGACAGAGCCAGAAGAGAAC TGAGACTTCTGCTGAGCCAAGCTGGAGAGGACAGAAAG GAATGCGTCTACTCCATGCATGACCCCTGCCAGCTTGCT CCTTCCTAAGGGACCATGAACGATATGTGACACCCGCTCA TACGTATGTGCACACCTGCAAGAGGGAGGCATCCATGTAC ACCTATGAGACGCACAGAGAAACATATATGTAGCCATAG GCTAGAAATTCTTCTCTTAGGTCTGCCCTCTGCA	小鼠; mm10:chr 9:1073409 28-107341 325	
15	GGACCACTCAGTGTACACGGAATGTAGAATTGAGTCTGCC ATTGGTCTCCCTCAAAGTCTGGAGGGCTGGACTGATA TTGGGAGCATCTGGCAGAGAAGGCCACAAAGACAGGGT GGTTTTCTACACTGGGACATACTCGTGAGCATGCACAGA	小鼠; mm10:chr 9:1073492	

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	<pre>GGCGTGTCCCCAACTTCCCTGTCACCCCTGTCCTCTGCCGG CTAGAGGGGATCGGGGGTGGACATATGCTGCTATTGGGC AGATATCACATGTTAAGAGGTGGGGGGGGCTCAAGAGG CGGAGGGCTAGGAGCATCCCATGGGAGAGGTTCTGGTTT TCTTGCTGCCTCTAGCTGCTATAAAACGTTAGCAGTGAG CAACTGGAAAGCTCTGAGTAATTAGGATGCACAAAGCTG TAATTAACTCCAGCATCTCAGTGTGCGAGAGCATTAAAG ATGTAATTAAAGATGTTACACAAAGAGATTGGAGTCTGTG ACACTTGGGGTCAAACACCCAGGAAGGGACACAATGGG TGAGGTGAGGACTGTGGGAGGCCTGGGGACAGTCACTTG GATCCCAGCTATGAGATGGCAGGCCACCCAGCTGTTCTC CTTGGAAATGTTTGGCCTGGGGTTGGGGTGGGGCATC ACACTTGATATGGAGATGGGGCAACAAAGCTGCAATAT CTGGGGGTGGAGAGGTCAAGTGGATGGAGTCTTGAGAT CATGTCAGGAAGAGGGCTGATCCCCAAAATCATGGTGA CATATGGTGTCTGGGTTCACAGGAGCTATGTCTAAAAT ACAAAAGTAAA</pre>	27-107350 036
[0070]	16 <pre>TCTGCAGAACGCTGCCATTCCACCATTAAACCTGTGACT CCAGGCCTTAAGCCTGTTGAAGGTGAGTCCCAGAAGGGT CATATGTGCAACTGCCTAGGGAGAGTTCCACTCGCAGGG CCAAGAGGAGTCCCCCGGTCTGAGGTGTGGGGCGGGGA CGTGCACTGGCGCTGGGACCAACGGCTGGCTCAGGACT CGC</pre>	小鼠; mm10:chr 9:1073994 38-107399 639
	17 <pre>TGCCTCAGTTCTCGCCTAGAAAGCCGGTCTAAGGGTA CATGCCCTGATTCTTCTGGGGTGTCTGAATTAAACA ACACATACTGTTCTGGGCTGATGACAAGAGGAAGTACTGG TCGGTGGCTGATGGACATCCACCATGGTGGCAACTGGAGG GAGGGGGAACGGACGTTGAAACCCCTGCCCTCCTGGAATCT GTCGCATGCACGCACGTTGACAATGCTTGGCACTGGGAC AGGCTGGGATGGATGGAGCGGAGCGTGAGGAGGAGTGGG CATGCAGGCCCGAGTGTCTGTTGCTGATTGCTCCTTTG CTTCAAGGAGATTAAACTATTTAGTCATGCCACTGC TGGTAGAGACGCTGGAGGAAGCCTTCATCGTTGAGATT TCTGGAAGCTGCCAAGTGTGGCTTCAGCTCAATTCTGGG AGCCTCCCAGAGTGGAGGGAGGAACATTCCATCTGGG GGCTTCGGGGACAGGCTAAGATCTCCCTGGGTCTTGC TGCCTGGCCTCCTCAAACACACGCTGCCCTCAGCTGCATA AAGCAGTAATCTGATGTGCCGATGTTGTAACGCTGTGT TTAAAAAAAGTAATTATTTCTAATTATTCTGTCTTGC ATAACCATGCATTGCCAAAGTGTGCTATTAAAATATT ATCTCTCCACGCCGAGGAGCAGCTCTGGAGCGTGGAGGG GGAAGAAAATAAAAGTCCCGTGCCAGTCGCAGGCATATT ACTTGACTCGTCTGGCTTGACGTCCTGTAAT ACATTATTTCATTAGGACGTTCTGAGCTTGTGGCCCC CGGAGAGCGGAGTGATTACGCTGTTCATGCAAGCGATG CAATAGAGGGTACTCGCAGAATGACTTCCGCCAGAGC ATCCTGCGCCTGTCT</pre>	小鼠; mm10:chr 9:1074432 92-107444 228

SEQ ID NO:	核酸序列	来源/ 基因组位 置
[0071]	<p>18</p> <pre>TAAAATACCTTATTTTCCAGTCTCTAAACTGCTAATCT CCCAGGCTAACGGGATTCTGGGACAAGGCAAGGCCTGGA AGTGGAAATCTGAAAATTAGCTTCAGCGGTATTAGTGTG TGCAGTTGAAGATTGAAAAACTGCTTCCCAGGGCCTGAT TGGAGGCTCCACTCTCCTCCAGGAAGAGGCAAGGACTCTG GGCTGGCACTGAGGACAAATCCTGGGAGGCTGCTATGGG GCCTGGGAGGCCAGGCTGCCTGTGCTAGAGGCCTAGAGAG TGTCTGTGTCCCAGTCCAAAGCTCAGGCACAGCAGGTGCCAAGATCAC AGCTTTCCAGTTCTCAGGCACAGCAGGTGCCAAGATCAC GCTCTGGAGTCCAGCTGGGCCCCCTCCTCTTCTTTTTTT TTTTTTTTAAGACCTCCTGGACACTGTTCTCTCCCCCCC CCCGTGACCCCCCCCCCTCAGTCTCAAACACGTGAGGGTT GGGGGAGGGTCCACAGCCAGAGAGAGGGGCCAGCTG GTGCCTGTGGGTACGCCGCCGTATGGCCATCAGGCCT CTTGTGTGCTTGATTGCCTCTGATTGGCTGCAGCTGAATT AGCAAAAGCTATTATTGCCCTGATGAGCCAATCAGATG GCCTCATTGCCATTCAAGAGCAGGCACCGAACCTGAGGG TGGGGTGGGGGTGGGGATGGAGATGGACTCAGTGAG GGGGTGGGAAGCTCTAAAACAGATGCAAGGACCTGAGCCT GTCTGTGTCCACCACGACCTTCACACAGGTCACACCCCCCT TCCCCCTGACTTGTCAACCCAAACCAAGGGCTTGTGCCAA CCCCACCTCACAAATTCCCTCACTCTGTAACACCTTCCATA TACCTCTGCATGTCTAAACCAAGACTTGCTCTATGAAAT C</pre>	小鼠; mm10:chr 9:1074448 25-107445 746
	<p>19</p> <pre>AGACCCCTGCTTAGCACAGCTCTAGCGGGCCTTAGGGG GTCTCCCAGCGGGCCCAGTGGGAATGAGATAAGGAAGGA CACAGCTGTCCATTCTCCGTGCCTGCTAAGGAGGAAATG GGGCCGCCTACATAATTGGGCAATTGTTCCACTCTTGT CCTCCTGGTATCATGGCTATCACCCCTCCTGCTCAGGG GTCCTGATTGAGCGAGAAAGCTCAGGCCTCCCTCTCCCT CCTGCTGGGGGTGCTGAACAGAGGGTGTAGGAGCCATA GGCTCTGTCAGTGCTGAGATCTGCCAGATGTCTAGGCCAG GAGAAAATGAAAGGGCTAAGTCACAGCATATGTGGCCA CTCAGGCCTATAGCCCCAAATCTGCCTGGTAACCCATTAT GTCCCCAGAGAATTGCATGGCGGGACACCCCTCATGCCGG GTCTCAGTAAGGGAAAGGGTGGGAGGAAAAATATCCCT CCCCACCCCTGAATCTCCACCCCTCCCCCAGAAACTGAC ACTTGGCCTGTCTAAGGATGGTTCCAAAATCCTTCT GAAAAAAACAGAATTCAAGAGTCACTCCCTCCGGGTCTC AGCCTAGAACATATGCAGTATCCCTGACGTCCATAGGG</pre>	小鼠; mm10:chr 9:1074520 80-107452 718
	<p>20</p> <pre>AAACTGGCACAGTAATGGCGGGCTGACAGACAAGGGAGT CTGTAGCACCCGCTGCCTCCGCCACCCCTCTCCGAGCA ATTAAAAGGTGTTATGTGGGGCTGGCAGTGGCTCTGCC TCCCTCCATTACGAACATTAAGAGATCTGACCCCTCCAC TTTCCCCGCTTGTAAAGGAGCTGCAGACACGTGGAGCCA ATTAGGGCAGCGTGGCGCCAAGGGCCTGAGCAGCTT TTCTCCCTGATTGCGGCCATTACAGCTGATTATTCTCCCT</pre>	小鼠; mm10:chr 9:1074704 14-107471 129

[0072]

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	CACCCAAACAGTGCTGCTCCTGGCAAGGTGCCACCCAGA GGAGCCGGCTGGGGCCCCCTGGGGACAGGGGAGGGACTGG ATTAGTAAATGGGCATCTATCGAATGGCTTCATATGTGT GGCTGGAAGGGAGAAGGGTAGGGCCAGGAATGGTGGCAG CAAGGGCCCAGGTAGCAATGAGGGTTCTCTAACCCACCA TTTAGGGATAGCGATCAGAAAAGGGGCCCTCGAGGAGGTG ACCTAAATGTGTGTAGAAGCTGACGGCCACTACACACACA CACACACACACACACACATACACACATAAACAGCATCCTGTGCT TGGAGTCGGTCAGCATGAGCAAGAGAAAGATGTTCCCAG TGGCCATGAGAGTGGAGGCCCTCCCTACTTACATCCAG GTTGGATGGCCAGGAGATCCTGAGATCCTCAAGACTCC	
21	AAGCCACATCCTGGGTGAAATATATGGCTCAATTCCA CTCTTCCGGATGACCTCTGTGGGGAGGCCCTGGCTTCACCTT GGTCCAGCTTCATCCCTAGCCTCGCTGCCAGGAAGGCAG TGAGGTCAGAGGCTGGTGTGGCGTG	小鼠; mm10:chr 9:1074848 87-107485 033
22	CCTACCTGGTCCCCGCCAACATCTGGGGGCCATCCTGGCC AGGCCAGCGTGGTGAAGGCACTGTGCGCCGTGGTAC TGTTTCTCTACCTGCTTCCCTCGCTGTGGACACGGGCTGC CTGGCCGTCACCCCAGGCTACCTTCCCACCCAACCTCTG GATCTGGACCCCTGGCCACCCACGGGCTCATGGAACAGCAC GTGTGGGACGTGGCCATTAGCCTGGCCACAGTGGTTGTGG CGGGCGATTACTGGAGCCCCTCTGGGGAGCCTTGGAGCT GCTCATCTTCTCTC	小鼠; mm10:chr 9:1075344 90-107534 786
23	AAACGGACGGGCCTCCGCTGAACCAGTGAGGCCAGAC GTGCGATAAATAACCCCTGCGTGCACCCACCTGGGAGA GAGGGGGAGGACCACGGTAAAT	人类; hg19: chr2:1716 72063-171 672163
24	GGAGCGAGCGCATAGCAAAAGGGACGCGGGGTCTTTCTCTGCGGACTGGTAGCTGTGGCCAGGTGTGGTAC TTTGATGGGCCAGGGCTGGA	人类; hg19: chr2:1716 72697-171 672797
25	GCTCAAGGAAGCGTCGCAGGGTCACAGATCTGGGGAAAC CCCGGGGAAAAGCACTGAGGCAAAACCGCCGCTCGTCTC CTACAATATATGGGAGGGGGAGG	人类; hg19: chr2:1716 72918-171 673018
26	TTGAGTACGTTCTGGATTACTCATAGACCTTTTTTC CTTCCGGCGAAAACCGTGAGCTGGATTATAATCGCCC TATAAAGCTCCAGAGGCGGTAGGCACCTGCAGAGGAGC CCCGCCGCTCCGCCACTAGCTGCCCCCGAGCAACGGC CTCGTGATTTCCCCGCCATCGGTCCCCGCCCTCCCCACTC TGCCCCCGCCTACCCGGAGCCGTGCAGCCGCTCTCCGA	人类; hg19: chr2:1716 73150-171 673696

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	ATCTCTCTTCTCCTGGCGCTCGCGTGCAGAGAGGAACT AGCGAGAACGAGGAAGCAGCTGGAGGTGACGCCGGCAG ATTACGCCTGTCAGGGCCGAGCCGAGCGGATCGCTGGCG CTGTGCAGAGGAAAGGCAGGAGTGCCTGGCTCGCTGTAG CAGAGCCGAGGTGGTAAGCTAGCGACCACCTGGACTTCC CAGCGCCCAACCGTGGCTTCAGCCAGGTCTCTCCTCCCC GCGGCTTCTCAACCAACCCCACCTCCAGCGCCGGCACCCA ACCTCCCCGAAATGAGTGCTTCCTGCCCC	
27	CAGCAGCCGAAGGCAGCTACTAGGAACGGTAACCTGTTACT TTTCCAGGGGCCGTAGTCGACCCGCTGCCGAGTTGCTGT GCGACTGCGCGCGCGGGCTA	人类; hg19: chr2:1716 73900-171 674000
28	GAGTGCAAGGTGACTGTGGTTCTTCTCTGCCAAGTCCGA GGGAGAACGTAAGATATGGGCCTTTCCCCCTCTCACC TTGTCTCACCAAAGTCCCTAGTCCCCGGAGCAGTTAGCCT CTTTCTTCCAGGGAATTAGCCAGACACAACAAACGGGAAC CAGACACCGAACCGACATGCCGCCCCGTGCGCCCTCCCC C	人类; hg19: chr2:1716 74400-171 674600
29	GCTCGCTGCCTTCCTCCCTCTTGTCTCTCCAGAGCCGGAT CTTCAGGGGAGCCTCCGTGCCCGCTGCTCAGTCCCT CCGGTGTGCAGGACCCCGGAAGTCCTCCCCGACAGCTCT CGCTTCTCTTGCAAGCTGTTCTGCGCCGGACCAGTCGAG GACTCTGGACAGTAGAGGGCCCCGGGACGACCGAGCTG	人类; hg19: chr2:1716 74903-171 675101
30	AAACGGACGGGCTCCGCTGAACCAGTGAGGCCAGAC GTGCGCATATAAAACCCCTGCGTGCACCACTGGGA GAGGGGGAGGACCAACGGTAAATGGAGCGAGCGCATAGCA AAAGGGACGCGGGTCCTTCTCTGCCGGTGGCACTGGG TAGCTGTGGCCAGGTGTGGTACTTGATGGGGCCAGGGC TGGAGCTCAAGGAAGCGTCGCAGGGTCACAGATCTGGGG GAACCCCGGGAAAAGCACTGAGGCAAAACCGCCGCTCG TCTCCTACAATATGGGAGGGGGAGGTTGAGTACGTTCT GGATTACTCATAAGACCTTTTTTCCCTCCGGCGCAA AACCGTGAGCTGGATTATAATCGCCCTATAAGCTCCAG AGGCAGTCAGGCACCTGCGAGAGGGAGCCCGCCGCTCCGC CGACTAGCTGCCCGCGAGCAACGGCCTCGTGAATTCCC CGCCGATCCGGTCCCCGCCCTCCCCACTCTGCCCGCCTAC CCCGGAGCCGTGCAGCCGCTCTCCGAATCTCTCTTCTC CTGGCGCTCGCGTGCAGAGGGAACTAGCGAGAACGAGG AAGCAGCTGGAGGTGACGCCGGCAGATTACGCCTGTCA GGGCCGAGCCGAGCGGATCGCTGGCGCTGTGCAGAGGA AAGGCAGGGAGTGCCTGGCTCGCTGCGAGAGCCGAGGT GGGTAAGCTAGCGACCACCTGGACTTCCCAGCGCCCAACC GTGGCTTTCAAGCCAGGTCTCTCCTCCCGGGCTTCTCAA CCAACCCCATCCCAGCGCCGGCACCCAACCTCCGAAAT GAGTGCTTCTGCCAGCAGCCGAAGGGCTACTAGGAA	人类

[0073]

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	CGGTAACCTGTTACTTTCCAGGGGCCGTAGTCGACCCGC TGCCCAGTTGCTGTGCGACTGCGCGCGGGCTAGAGT GCAAGGTGACTGTGGTTCTCTGGCCAAGTCCGAGGGGA GAACGTAAAGATATGGGCCTTTCCCCCTCACCTTGTC TCACCAAAGTCCTAGTCCCCGGAGCAGTTAGCCTCTTCT TTCCAGGGAATTAGCCAGACACAACAAACGGGAACCAGAC ACCGAACCGAGACATGCCCGCCCCGTGCGCCCTCCCCGCTC GCTGCCCTTCCTCCCTTGTCTCTCCAGAGCCGGATCTTC AAGGGGAGCCTCCGTGCCCCGGCTGCTCAGTCCCTCCGG TGTGCAGGACCCCGGAAGTCCTCCCGCACAGCTCTCGCT TCTCTTGCAGCCTGTTCTGCGCCGGACCAGTCGAGGACT CTGGACAGTAGAGGCCCCGGACGACCGAGCTG	
[0074]	31 GGAGGAAGCCATCAACTAAACTACAATGACTGTAAGATA CAAAATTGGGAATGGTAACATATTGAAAGTTCTGTTGAC ATAAAGAATCATGATATTAATGCCATGAAATGAAAGG GCGATCAACACTATGGTTGAAAAGGGGAAATTGAGA GCACAGATGTGTTCGTGTGGCAGTGTGCTCTAGCAA TACTCAGAGAAGAGAGAGAACATGAAATTCTGATTGGC CCCAGTGTGAGGCCAGATGAGGTTAGCTGCCAACTTCT CTTCACATCTTATGAAAGTCATTAAGCACAACAACTT TTTTTTTTTTTTTTTTGAGACAGAGTCTGCTCTGTTG CCCAGGACAGAGTCAGTAGTGAUTCAATCTCGGCTCACT GCAGCCTCCACCTCCTAGGCTCAAACGGCTCTGCATC AGCCTCCCAAGTAGCTGGAATTACAGGAGTGGCCACCAT GCCAGCTAATTGTATTTAATAGATACGGGGTTTC ACCATATCACCCAGGCTGGCTCGAACCTCTGGCCTCAAG TGATCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTATA GGCGTCAAGCCACTATGCCAACCCGACCAACCTTTAA AATAAATATTAAAAAATTGGTATTCACATATATACTAG TATTACATTATCCACACAAACGGACGGGCTCCGCTG AACCAAGTGAGGCCAGACGTGCGCATAAATAACCCCTGC GTGCTGCACCACCTGGGGAGAGGGGGAGGACCAACGGTAA ATGGAGCGAGCGCATAGCAAAAGGGACGCGGGCTT TCTCTGCCGGTGGCACTGGTAGCTGTGCCAGGTGTGGT ACTTGATGGGCCAGGGCTGGAGCTCAAGGAAGCGTC GCAGGGTCACAGATCTGGGAAACCCGGGAAAGCAC TGAGGCAAAACGCCGCTCGTCTCCTACAATATATGGAG GGGGAGGTTGAGTACGTTCTGGATTACTCATAAGACCTT TTTTTCTCTCCGGCGCAAACCGTGAGCTGGATTATA ATCGCCCTATAAAGCTCCAGAGGCCGCTGGCAGGCACCTGCAG AGGAGCCCGCCGACTAGCTGCCCGCGAGC AACGGCCTCGTATTCCCCGCCATCCGGTCCCCGCC CCCACTCTGCCCGCCTACCCGGAGCCGTGAGCCGCC TCTCCGAATCTCTCTTCTGCCGCTGGCAGCGTGCAGAGAG GGAAGTAGCGAGAACGAGGAAGCAGCTGGAGGTGACGCC GGGCAGATTACGCCGTGCAAGGGCGAGCCGAGCGGATCG CTGGCGCTGTGCAAGAGGAAGGGAGTGCCTGGCT	人类

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	GCTGTCGCAGAGCCGAGGGTGGTAAGCTAGCGACCACCT GGACTTCCCAGCGCCAACCGTGGCTTCAGCCAGGTCC TCTCCTCCCAGCGGCTCTCAACCAACCCATCCCAGCGCC GGCCACCCAACCTCCCAGAAATGAGTGCTTCCTGCCAGC AGCCGAAGGCAGTACTAGGAACGGTAACCTGTTACTTTTC CAGGGGCCGTAGTCGACCCGCTGCCGAGTTGCTGTGCGA CTGCGCGCGGGGCTAGAGTGCAAGGTGACTGTGGTTCT TCTCTGGCCAAGTCCGAGGGAGAACGTAAGATAATGGGC CTTTTCCCCCTCTCACCTGTCTCACCAAAGTCCCTAGTC CCCGGAGCAGTTAGCCTCTTCTTCCAGGGATTAGCCA GACACAACAACGGGAAACCAGACACCGAACCGAACATGCC CGCCCCGTGCGCCCTCCCCGCTCGCTGCCTTCCTCCCTCT TGTCTCTCCAGAGCCGGATCTCAAGGGAGCCTCCGTGC CCCCGGCTGCTCAGTCCCTCCGGTGTGCAGGACCCGGAA GTCCTCCCCGCACAGCTCGCTCTTGCAGCCTGTT CTGCGCCGGACCAGTCGAGGACTCTGGACAGTAGAGGCC CCGGGACGACCGAGCTG	
[0075]	<p>32 TCAACAGGGGACACTGGGAAAGAACGGATGGGACAGA GCCGAGAGGACTGTTACACATTAGAGAAACATCAGTGACT GTGCCAGCTTGGGTAGACTGCACAAAAGCCCTGAGGCA GCACAGGCAGGATCCAGTCTGCTGGTCCCAGGAAGCTAAC CGTCTCAGACAGAGCACAAAGCACCGAGACATGTGCCAC AAGGCTTGTAGAGAGGTAGAGGACAGCGTACAGGTC CCAGAGATCAAACCTCACCTCACCCAGGCTTGGCAGCAAGC CTTACCAACCCACCCCCACCCACCCACCCCTGCACGCGC CCCTCTCCCCTCCCCATGGTCTCCCATGGCTATCTCACTTG GCCCTAAAATGTTAAGGATGACACTGGCTGCTGAGTGG AATGAGACAGCAGAAGTCAACAGTAGATTAGGAAAGC CAGAGAAAAAGGCTGTGCTGTTAGAAAGCCAAGGG ACAAGCTAAGATAGGGCCAAGTAATGCTAGTATTACAT TTATCCACACAAAACGGACGGGCCTCCGCTGAACCAGTGA GGCCCCAGACGTGCGCATAATAACCCCTGCGTGCAC CACCTGGGAGAGGGGAGGACACGGTAAATGGAGCGA GCGCATAGCAAAAGGGACGCGGGGTCTTCTGCCGG TGGCACTGGTAGCTGGCCAGGTGTGGTACTTTGATGG GGCCCAGGGCTGGAGCTCAAGGAAGCGTCGCAGGGTCAC AGATCTGGGGAAACCCGGGAAAAGCACTGAGGGAAAA CCGCCGCTCGTCTCCTACAATATGGGAGGGGAGGTTG AGTACGTTCTGGATTACTCATAAGACCTTTTTTCTT CCGGCGCAAACCGTGAGCTGGATTATAATGCCCTAT AAAGCTCCAGAGGCAGGTCAAGGCACCTGCAGAGGCC GCCGCTCCGCCACTAGCTGCCCGCGAGCAACGGCCTC GTGATTCCCCGCCATCGGTCCCCGCCCTCCCCACTCTGC CCCCGCCCTACCCGGAGCCGTGCAGCCGCTCTCGAATC TCTCTCTCTCCCTGGCGCTCGCGTGCAGAGGGAACTAGC GAGAACGAGGAAGCAGCTGGAGGGTACGCCGGGAGATT ACGCCTGTCAGGGCCGAGCCGAGCGGATCGCTGGCGCT </p>	人类和小鼠

SEQ ID NO:	核酸序列	来源/ 基因组位 置
	GTGCAGAGGAAGGCGGGAGTGCCCGGCTCGCTGTCGCA GAGCCGAGGTGGTAAGCTAGCGACCACCTGGACTTCCA GCGCCCAACCGTGGCTTCAGCCAGGT CCTCTCCTCCGC GGCTTCTCAACCAACCCATCCCAGCGCCGGCACCCAAC CTCCCAGAAATGAGTGCTCCTGCCAGCAGCCGAAGGCG CTACTAGGAACGGTAACCTGTTACTTTCCAGGGGCCGTA GTCGACCCGCTGCCAGTTGCTGTGCGACTGCGCGCG GGGCTAGAGTCAAGGTGACTGTGGTTCTCTGGCAA GTCCGAGGGAGAACGTAAGATATGGGCCTTTCCCCCT CTCACCTTGTCTCACCAAAGTCCCTAGTCCCAGGAGCAGT TAGCCTCTTCTTCCAGGGATTAGCCAGACACAAC GGGAACCAGACACCGAACAGACATGCCGCCCGTGC CCCTCCCCGCTCGCTGCCCTCCCTTGTCTCCAG AGCCGGATCTTCAAGGGGAGCCTCCGTGCCCGGCTGCT CAGTCCCTCCGGTGTGCAGGACCCCGGAAGTCTCCCCGC ACAGCTCTCGCTCTTGCAGCCTGTTCTGCGCCGGAC CAGTCGAGGACTCTGGACAGTAGAGGCCCGGGACGACC GAGCTG	
33	ATTTACATTATCCACACA	人类
34	TGCCGCTGGACTCTTCCAAGGAACCTAGGAGAACCAAGA TCCGTTTTCTGCCAAGGGCTGCCACGCCACGCCAAC CCCCTCACCCGATCCCCACAGAAAGAAATCTTGAGGTAG CTGGAGCTTCTCTGTGGGTGTGACAGGACTGCCATTCTCC TCTGTAGTCTGAGAACGCTGCCATTCCACCATTTAACCT GTGACTCCAGGCCTTAAGCCTGTTGAAGGTCGAGTCCCAG AAGGGTCATATGTGCAACTGCCCTAGGGAGAGTTCCCACTC GCAGGGCCAAGAGGAGTCCCCCGTCTGAGGTGTGGGG CGGGGACGTGCACTGGCGCTGGGACCA CGGCTGGGGCT CAGGACTCGCGAGCTTGGATTGGATCGGTTGCGCGAGC CAGTAGGGCAGGCTCCGGGTGAACGGGACGAGGGCG CGCGGGCACAGGCCGGCGCGTGAACGCCGGCGGGCG CGGAGGCAGGCCGGCGCAAGGAGAGGGAGGGAGGGAATG AGGGAGGGAGCGACAGGGGAGGGCGGCCGGCAGGTTG GCAGGGCGCCGCTATTGAGCGCAGGTCCCGGGCCAGGCGC TCAAAGCGCTGGAGCCAGCGCGGGAGATCGCTGC GCGCAGCCCGCAGAGGCCTGCCAGTGCAGCCCCGG AGGCCCGCGCGAGAACGGAGGTGGAGAACAGGCCGGCT TTCCGCCCGCCGCCGCCGGCCCCACCTCCATCCCGCG CCGCCGTCCCCCTCCCTCCCGCGCGCATCTTGAAT GGAAAC	小鼠; chr9:107,3 99,268-10 7,400,067
35	GAGTAATTACAAAAGGACTGCCCTGCCTGGGAA TCCCAGGGACCGTCGTTAAACTCCACTAACGTAGAACCC AGAGATCGCTCGCTCCGCCCTACCCGCCGCTCTC GTCATCACTGAGGTGGAGAACAGGCATGCGTGAGGCTCCG GTGCCCGTCAGTGGCAGAGCGCACATGCCACAGTCCC CGAGAACGTTGGGGAGGGTGGCAATTGAACCGGTGC	

[0076]

SEQ ID NO:	核酸序列	来源/ 基因组位 置
[0077]	CTAGAGAAGGTGGCGCGGGTAAACTGGGAAAGTGATGT CGTGTACTGGCTCCGCCTTTCCCAGGGTGGGGAGAA CCGTATATAAGTCAGTAGTCGCCGTGAACGTTCTTTCG CAACGGGTTGCCGCCAGAACACACAGTAAGTGCCGTGT GGTCCCCGCGGCCCTGGCCTCTTACGGTTATGCCCTTG CGTGCCTTGAATTACTCCACGCCCTGGCTGCAGTACGT GATTCTTGATCCCAGCTCGGGTTGGAAGTGGGTGGGAG AGTCGAGGCCTTGCCTTAAGGAGGCCCTCGCCTCGTG CTTGAGTTGAGGCCTGGCTTGGCGCTGGGCCGCCGCGT GCGAATCTGGTGGCACCTCGCCTGTCTCGCTGCTTTCG ATAAGTCTCTAGCCATTAAAATTTTGATGACCTGCTGCG ACGCTTTCTGGCAAGATAGTCTGTAATGCGGGCC AAGATCTGCACACTGGTATTCGGTTTGGGCCGCCGG CGCGACGGGGCCCGTGCCTCCAGCGCACATGTTGGCG AGGCGGGGCCTGCGAGCGCGGCCACCGAGAAATGGACGG GGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCC TCGCGCCCGTGTATGCCCGCCCTGGCGGCAAGGCT GGCCCGGTGGCACCAAGTTGCGTGAGCGGAAAGATGGCC GCTTCCCAGGCGCCCTGCTGCAGGGAGCTCAAATGGAGGACG CGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAA AGGAAAAGGGCCTTCCGTCCAGCCGTCGCTCATGTG ACTCCACGGAGTACCGGGCGCCGTCCAGGCACCTCGATT GTTCTCGAGCTTGGAGTACGTCGTCTTAGGTTGGGG GAGGGGTTTATGCGATGGAGTTCCCCACACTGAGTGGG TGGAGACTGAAGTTAGGCCAGCTGGCACTTGATGTAATT CTCCTTGAATTGCCCTTGTAGTTGGATCTGGTTCA TTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTCTCCA TTTCAGGTGTCGTGA	

[0078] 一方面,本文公开的调控元件是细胞类型选择性的。在一些情况下,本文公开的调控元件对PV神经元具有选择性。在一些情况下,本文公开的调控元件对CNS中的PV神经元具有选择性。在一些情况下,PV细胞选择性调控元件或本文公开的任何调控元件可导致相比于至少一种、两种、三种、四种、五种或更多种非PV CNS细胞类型,在PV神经元中的选择性基因表达。

[0079] 在一些情况下,本文公开的任何一种或多种调控元件与表达盒中的转基因可操作地连接,以导致在靶细胞类型例如PV神经元中的选择性表达。在一些情况下,本文任何实施方案的调控元件包含以下任一项或由以下任一项组成:(i) SEQ ID NO:1-33;(ii)其变体、功能片段或组合;或(iii)与(i)或(ii)中任一项具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列同一性的序列。在一些情况下,调控元件包含SEQ ID NO:1-32中的任一个。在一些情况下,序列同一性通过BLAST测量。

[0080] 在一些情况下,将SEQ ID NO:1-32中的两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个或十个或更多个,或功能片段或其组合,或与其具有至少80%、至少90%、至少95%或至少99%序列同一性的序列组合形成较大的调控元件,或将其与表达盒中的基因可操作地连接。在一些情

况下,使用1-50个核苷酸的接头序列将SEQ ID NO:1-32中的两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个或十个或更多个,或功能片段或其组合,或与其具有至少80%、至少90%、至少95%或至少99%的序列同一性的序列组合。在一些情况下,将SEQ ID NO:1-32中的两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个或十个或更多个,或功能片段或其组合,或与其具有至少80%、至少90%、至少95%或至少99%序列同一性的序列以无接头序列的方式组合。在一些情况下,SEQ ID NO:33的序列用作任何两个调控元件之间的接头。在一些情况下,任何两个调控元件之间的接头序列包含SEQ ID NO:33,或与SEQ ID NO:33具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%序列同一性的序列。在一些情况下,序列同一性通过BLAST测量。

[0081] 在一些情况下,当两个或更多个调控元件在表达盒中组合或使用时,调控元件不必在表达盒中相邻或连接。例如,一个调控元件可以位于转基因的上游,而第二调控元件和/或其他调控元件可以位于转基因的下游。在一些情况下,一个或多个调控元件可以位于转基因的上游。在一些情况下,一个或多个调控元件可以位于转基因的下游。

[0082] 在一些情况下,SEQ ID NO:1-22中的任一个或多个,或其功能片段或组合,或与其具有至少80%、至少90%、至少95%或至少99%序列同一性的序列可以与SEQ ID NO:23-30中的任一个或多个,或其功能片段或组合,或与其具有至少80%、至少90%、至少95%或至少99%的序列同一性的序列组合以形成较大的调控元件。例如,调控元件包含SEQ ID NO:1和SEQ ID NO:30。调控元件包含SEQ ID NO:8和SEQ ID NO:30。调控元件包含SEQ ID NO:1和SEQ ID NO:23-29。调控元件包含SEQ ID NO:8和SEQ ID NO:23-29。在一些情况下,调控元件包含SEQ ID NO:30或SEQ ID NO:23-29中的任一个或多个,或其功能片段或组合,或与其具有至少80%、至少90%、至少95%或至少99%序列同一性的序列。

[0083] 在一些情况下,本公开内容的一个或多个调控元件导致PV细胞中的选择性基因表达。在一些情况下,在靶细胞类型中显示选择性活性或功能的调控元件在一种或多种脱靶细胞类型,例如非PV CNS细胞类型、非抑制性神经元或兴奋性神经元中也显示出最小的活性或功能。

[0084] 在一些情况下,与基因可操作地连接的一个或多个调控元件调节细胞中的基因表达,包括但不限于相对于非靶细胞类型,在靶细胞类型中的选择性表达。靶细胞或细胞类型中的选择性表达也可以称为细胞选择性表达或细胞类型选择性表达。

[0085] 选择性表达通常是指与其他细胞(或非靶细胞类型)相比,在感兴趣的细胞类型(或靶细胞类型)的大部分细胞中表达。选择性表达也可以被视为相对于一种或多种非靶细胞或细胞类型,优先在靶细胞或靶细胞类型中表达。在一些情况下,将本公开内容的一种或多种调控元件的选择性表达与CAG、EF1 α 、组成型启动子(例如,SV40、CMV、UBC、PGK和CBA)或已知非选择性地驱动任何细胞或细胞类型中的表达的非选择性调控元件相比较。在一些情况下,将本公开内容的一种或多种调控元件的选择性表达与CAG、EF1 α 、组成型启动子(例如,SV40、CMV、UBC、PGK和CBA)、非选择性调控元件或没有调控元件的表达盒相比较。

[0086] 非靶细胞类型可以包括不同于靶细胞或靶细胞类型的细胞子集、亚型或类型,或所有非靶细胞类型。在一些情况下,与基因可操作地连接的一种或多种调控元件导致相对

于至少一种非靶细胞类型或至少两种、至少三种、至少四种、至少五种或五种以上的非靶细胞类型，在靶细胞类型中选择性表达。在一些情况下，非靶细胞类型是指不包括靶细胞类型的所有其他细胞类型。在一些情况下，非靶细胞类型是相关组织或器官内不包括靶细胞类型的所有其他细胞类型，例如CNS中的所有非靶细胞类型，海马体中的所有非靶细胞类型。在一些情况下，非靶细胞或非靶细胞类型涵盖不是靶细胞的细胞子集或亚型。例如，非PV CNS细胞类型可以包括表达钙网膜蛋白和/或促生长素抑制素而不是小清蛋白的GABA能细胞，或不表达小清蛋白的所有GABA能细胞。在一些情况下，细胞类型通过具有不同的细胞标志物、形态、表型、基因型、功能和/或用于划分细胞类型的任何其他方式来区分。

[0087] 由感兴趣的细胞或细胞类型中的调控元件驱动的表达的选择性可以以多种方式测量。通过将从与一个或多个调控元件可操作连接的基因表达可检测水平的转录物的靶细胞数与表达该基因的总细胞数进行比较，可以测量靶细胞类型相对于非靶细胞类型中基因表达的选择性。这种测量、检测和定量可以在体内或体外进行。

[0088] 在一些情况下，可以使用共定位测定确定对PV神经元的选择性。在一些情况下，共定位测定基于免疫组织化学。在一些情况下，将可检测的报道基因用作转基因，以允许检测和/或测量细胞中的基因表达。在一些情况下，特异性标记靶细胞的可检测标记，例如荧光标记或抗体，被用于检测和/或测量靶细胞。在一些情况下，共定位测定采用成像，例如荧光成像，来确定不同荧光标记之间的重叠，例如，指示靶细胞的荧光信号与指示基因表达的另一荧光信号之间的重叠。在一些情况下，用于共定位测定的荧光标记包括红色荧光蛋白(RFP)，如tdTomato报告基因，和绿色荧光报告蛋白，如eGFP。

[0089] 在一些情况下，与一个或多个调控元件可操作地连接的基因是荧光蛋白，例如eGFP或RFP，其中转基因的表达提供可检测的信号。在一些情况下，对组织进行eGFP染色，或使用荧光显微镜直接检测eGFP的荧光。具有不同荧光或可检测信号的第二荧光标记或报道基因可用于指示靶细胞，如鉴定靶细胞的抗体。例如，与PV神经元特异性相互作用的抗PV抗体可用于产生可检测的信号，该信号可与用于测量基因表达的荧光，例如红色荧光或红色染料相区别。因此，在eGFP是可操作地连接至驱动PV神经元中选择性表达的一个或多个调控元件的转基因，并且其中PV神经元被抗PV抗体标记的一个实例中，PV细胞中基因表达的选择性以也是PV+的eGFP+细胞的百分比来测量。在这样的测定中，通过两个荧光信号的重叠，即红色和绿色荧光的重叠，来指示也是eGFP+的PV+细胞。这样的测量、分析和/或检测可以通过目测或计算机来完成。

[0090] 在一些情况下，也可以测量表达转基因的感兴趣的细胞类型(或靶细胞类型)的比例，与表达该转基因的非靶细胞类型(或其他细胞)的比例相比较，以评估与转基因可操作地连接的一种或多种调控元件的选择性。类似地，也可以通过将表达与一个或多个调控元件可操作地连接的转基因的靶细胞数与表达该转基因的所有细胞的总数进行比较来测量表达的选择性。在这两种方法中，表达转基因的靶细胞数越多，调控元件对靶细胞就越具有选择性。在一些情况下，靶细胞是PV神经元。

[0091] 在一些情况下，本文公开的一种或多种调控元件导致基因在PV神经元中的基因表达的选择性增加。在一些情况下，本文公开的一种或多种调控元件导致相比于在非PV CNS细胞类型中，在PV神经元中的基因表达的选择性增加。在一些情况下，本文公开的一种或多种调控元件导致相比于在非PVGABA能细胞中，在PV神经元中的基因表达的选择性增加，其

中非PVGABA能细胞可以是表达钙网膜蛋白(CR)、促生长素抑制素(SOM)、胆囊收缩素(CCK)、神经肽Y(NPY)、血管活性肠多肽(VIP)、胆碱乙酰转移酶(ChAT)或其组合的GABA能细胞中的任何一种或多种。在一些情况下,本文公开的一种或多种调控元件导致相比于在至少一个、至少两个、至少三个、至少四个、至少五个或多于五个的非PVGABA能细胞亚型中,在PV神经元中的基因表达的选择性增加。在一些情况下,本文公开的一种或多种调控元件导致相比于在所有其他非PVGABA能细胞,或所有其他不表达PV的GABA能细胞,或所有其他不表达PV的CNS细胞,或所有其他不表达PV的神经元中,在PV神经元中的基因表达的选择性增加。在一些情况下,本文公开的一种或多种调控元件导致相比于在CNS中的所有非PV细胞或所有非PV神经元中,在PV神经元中的基因表达的选择性增加。

[0092] 在一些情况下,与转基因可操作地连接的一个或多个调控元件导致转基因在PV细胞中的选择性表达,其中表达转基因的PV细胞的百分比高于转基因与CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件如SEQ ID N0:34,或其功能片段,或与其具有至少80%序列同一性的序列可操作地连接时在PV细胞中的基因表达。在一些情况下,一种或多种调控元件导致PV神经元中的选择性表达,该选择性表达为基因与CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA),或非选择性调控元件如SEQ ID N0:34,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列可操作地连接时的表达的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少15倍、至少20倍、至少25倍或至少50倍。

[0093] 在一些方面,调控元件是来源于人类的或包含来源于人类的序列。在一些情况下,调控元件是来源于小鼠的或包含来源于小鼠的序列。在一些情况下,调控元件包含非天然存在的序列。在一些情况下,调控元件是非天然存在的。在一些情况下,将一个或多个来源于人类的调控元件与另一调控元件组合以产生非天然存在的调控元件。在一些情况下,将来源于人类的调控元件与来源于小鼠的调控元件组合。

[0094] 如本文所用,术语“来源于人类的”是指在人类基因组(或人类基因组构建体)中发现的序列,或与其同源的序列。同源序列可以是具有与人类基因组区域相比有至少80%序列同一性(例如,通过BLAST测量)的区域的序列。例如,与人类序列具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%同源性的序列被认为是来源于人类的。在一些情况下,调控元件包含来源于人类的序列和来源于非人类的序列,使得总体上该调控元件与人类基因组具有较低的序列同一性,而调控元件的一部分与人类基因组具有100%序列同一性(或局部序列同一性)。

[0095] 在一些情况下,来源于人类的调控元件是与人类序列100%相同的序列。在一些情况下,细胞类型选择性调控元件的序列是100%来源于人类的。

[0096] 在其他情况下,至少5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、98%或99%的调控元件序列是来源于人类的。例如,调控元件的50%序列可以是来源于人类的,而其余50%是来源于非人类的(例如,来源于小鼠的或完全合成的)。进一步举例来说,被认为是50%来源于人类并且具有300bp的调控元件可与人类基因组中的序列具有45%的总体序列同一性,同时RE的碱基对1-150可能与人类基因组中大小相似的区域具有90%的同一性(局部序列同一性)。

[0097] 在一些情况下,与来源于人类的调节序列同源的序列与人类序列至少90%相同。在一些情况下,本文的调控元件包含与SEQ ID N0:1、23-31和33中的任一个具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%同一性的序列。在一些情况下,序列同一性由BLAST测定。当调控元件包含与SEQ ID N0:23-29中任一个或多个或其功能片段或组合同源(例如,至少80%、至少85%、至少90%、至少95%或至少99%的序列同一性)的序列时,与没有该调控元件的类似载体相比,这种调控元件导致可操作地连接的转基因的更高表达(当表达载体或表达盒中也存在启动子时)。可以在例如HEK293T或CHO细胞中观察到转基因的这种更高表达。

[0098] 在一些情况下,一种或多种调控元件包含在具有或不具有接头序列的情况下与SEQ ID N0:1-22中的任一个或多个组合或结合使用的SEQ ID N0:23-29中的任一个或多个。在一些情况下,本公开内容的任何两个调控元件使用包含1-50个核苷酸的多核苷酸接头如SEQ ID N0:33或其变体连接在一起。在一些情况下,接头序列是来源于人类的序列。在一些情况下,接头序列是来源于小鼠的或非天然存在的。在一些情况下,两个调控元件在没有接头或没有任何间插序列的情况下相连。在一些情况下,接头包含1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个核苷酸。在一些情况下,接头序列是限制酶位点、连接、PCR和/或克隆的结果。

[0099] 在一些情况下,将来源于人类的调控元件与来源于小鼠的调控元件组合,如SEQ ID N0:32,其是SEQ ID N0:8(来源于小鼠的序列)和SEQ ID N0:23-29(来源于人类的序列)的组合。在一些情况下,细胞类型选择性调控元件无需额外的接头序列即可直接组合。在其他情况下,细胞类型选择性调控元件与一个或多个短接头序列组合,该短接头序列可以是特制或克隆的产物。在一些情况下,接头序列包含1-50个碱基。例如,SEQ ID N0:31包含SEQ ID N0:1和23-29的序列,以及紧接在SEQ ID N0:1序列之后的另外19bp的基因组序列(SEQ ID N0:33)。SEQ ID N0:32也包括这些19bp,但没有SEQ ID N0:1的序列。在其他实例中,组合的细胞类型选择性调控元件可以包括来自克隆质粒或限制酶识别位点的短序列,该短序列通常小于50bp、小于20bp、小于15bp或小于10b。

[0100] 在一些情况下,调控元件可以来源于非编码DNA序列。在一些情况下,来源于非编码DNA的调控元件与基因,如上游序列、内含子、3' 和5' 非翻译区(UTR)和/或下游区域相关。在其他情况下,来源于非编码DNA序列的调控元件与基因不相关。在一些情况下,调控元件来源于编码序列。在一些情况下,调控元件所源自的基因组区域与可操作地连接的转基因所源自的基因组区域不同。在一些情况下,RE来源于相对于转基因所源自的基因组区域或位置(例如转基因的天然存在或内源性形式)处于远端的基因组区域或位置。

[0101] 一方面,调控元件是调节基因表达,例如靶细胞中表达的选择性的任何非编码序列。在一些情况下,靶细胞是PV神经元。在一些情况下,调控元件来源于转录起始位点上游的基因组序列、5' UTR序列、外显子序列、内含子序列或3' UTR序列。在一些情况下,来源于人类的调控元件包含来源于人类的内含子序列。在一些情况下,调控元件包含增强子,并且相比于在没有增强子的情况下相同转基因通过启动子的表达,增强子在表达盒中与启动子一起的存在增加可操作地连接的转基因在靶细胞类型(例如PV神经元)中的表达。在一些情况下,增强子通过转录机制、转录后机制或两者来增加可操作地连接的转基因的表达。在一些情况下,调控元件包含增强子序列、启动子序列或增强子和启动子序列的组合。在一些情况

下,调控元件包含一个或多个来源于人类的增强子序列、来源于人类的启动子序列、来源于人类的内含子序列和/或其组合。

[0102] 在一些情况下,调控元件包含SEQ ID N0:1-32中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,调控元件包含SEQ ID N0:1-22中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,调控元件包含SEQ ID N0:23-29中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,调控元件包含SEQ ID N0:30-32中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,调控元件包含SEQ ID N0:30-32的序列,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。

[0103] 在一些情况下,调控元件来源于非人类DNA序列或人类因组序列和非人类基因组序列二者。在一些情况下,细胞类型选择性调控元件或其部分与哺乳动物基因组序列同源。在一些情况下,调控元件与哺乳动物基因组序列具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同一性。在一些情况下,调控元件来源于小鼠基因组序列。在一些情况下,调控元件或其片段与小鼠基因组序列或非人类哺乳动物基因组序列具有至少约80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的同一性。在一些情况下,序列同一性通过BLAST测量。在一些情况下,调控元件可包含SEQ ID N0:1-33中的任一个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。

[0104] 在一些情况下,细胞类型选择性调控元件较短。在一些情况下,调控元件的大小与载体,例如病毒载体或rAAV的克隆能力是相容的,使得转基因和一个或多个调控元件的组合大小不超过载体的克隆能力。在一些情况下,细胞类型选择性调控元件的长度至多约2050bp、2000bp、1900bp、1800bp、1700bp、1600bp、1500bp、1400bp、1300bp、1200bp、1100bp、1000bp、900bp、800bp、700bp、600bp、500bp、400bp、300bp、200bp或100bp。在一些情况下,细胞类型选择性调控元件的总长度不超过约20bp、30bp、40bp、50bp、60bp、70bp、80bp、90bp、100bp、200bp、300bp、400bp、500bp、600bp、700bp、800bp、900bp、1000bp、1010bp、1020bp、1030bp、1040bp、1050bp、1060bp、1070bp、1080bp、1090bp、1100bp、1200bp、1300bp、1500bp、1600bp、1700bp、1800bp、1900bp或2000bp。在一些情况下,细胞类型选择性调控元件的长度为约100bp-1100bp、100bp-1000bp、100bp-900bp、200bp-900bp、200bp-800bp、300bp-600bp、400bp-800bp、500bp-600bp或600bp-900bp。在一些情况下,调控元件在约400-600bp、400-600bp、400-700bp、400-800bp、400-900bp、400-1000bp或400-1500bp之间。在一些情况下,调控元件在约500-600bp、500-700bp、500-800bp、500-900bp、500-1000bp或500-1500bp之间。在一些情况下,将两个或更多个调控元件组合以形成1300-2500bp、1300-2060bp、约1350bp、约2050bp或约1880bp的较大的细胞类型选择性调控元件。

[0105] 在一些情况下,可以组合两个或更多个细胞类型选择性调控元件。例如,可以组合两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个细胞类型选择性调控元件。例如,SEQ ID N0:30包含来自七个调控元件的序列,即SEQ ID N0:23-29,所有这些序列均来

源于人类基因组序列。在一些情况下，细胞类型选择性调控元件是指PV神经元选择性调控元件。

[0106] 在一些情况下，将细胞类型选择性调控元件重复两次或更多次以形成组合的调控元件，该组合的调控元件也具有细胞类型选择性或具有增强的细胞类型选择性性质。在一些情况下，将具有不同细胞类型选择性的两个或更多个调控元件组合。在一些情况下，将细胞类型选择性调控元件与非选择性调控元件，例如驱动基因高表达的非选择性增强子元件组合。例如，可以将对靶细胞具有高选择性的启动子调控元件与具有高表达效率的调控元件组合。在一些情况下，将一个或多个细胞类型选择性调控元件与一个或多个高效率调控元件组合。例如，SEQ ID NO:1-32中的任一个或多个，或其片段或组合，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列可以与组成型启动子，如GAD2启动子、人突触蛋白启动子、*min*CMV启动子、TATA框、超核心启动子或EF1 α 启动子或其组合进行组合。

[0107] 在一些方面，本公开内容提供了可添加到任何基因疗法中的调控元件的列表，该调控元件导致相对于在一种或多种非靶细胞类型，如包括但不限于兴奋性细胞和/或非PVGABA能细胞的非PV CNS细胞中，在靶细胞类型如PV神经元中的选择性基因表达。

[0108] 在一些情况下，细胞类型选择性调控元件可以与其他调控元件，如高表达启动子或增加mRNA稳定性的序列组合。在一些情况下，将一个或多个细胞类型选择性调控元件与人类、非人类或非哺乳动物的序列，例如hSyn1启动子、CBA启动子、CMV启动子、EF1 α 启动子、polyA信号(例如SV40 polyA信号)或转录后调控元件如土拨鼠肝炎病毒转录后调控元件(WPRE)组合。

[0109] 在一些情况下，组合的调控元件可能来自不同物种。组合的调控元件可以来自物种内的不同基因组区域。在一些情况下，调控元件来源于远端基因组序列，例如，在一般情况下或自然状态下与彼此或感兴趣的细胞类型不相关的序列组合。在一些情况下，用于制备组合的调控元件的各个调控元件可能来自不同的人类染色体。

[0110] 一方面，本公开内容的调控元件包含SEQ ID NO:1-32中任一项的功能片段，或其组合，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。与没有调控元件的类似表达盒或载体相比，这种功能片段可以增加转基因在表达盒或载体中的表达。与没有功能片段的类似载体或盒相比，当该功能片段与转基因可操作地连接时，该功能片段可充当增强子以增加细胞类型选择性表达。片段的长度优选大于30、40、50或60bp。

[0111] 在一些情况下，PV细胞选择性调控元件或本公开内容的任何调控元件包含(i) SEQ ID NO:1-32中的任一个，(ii)与SEQ ID NO:1-32中的任一个具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(iii)(i)或(ii)的任何序列的功能片段，或(iv)(i)、(ii)和/或(iii)的任何序列的组合。在一些情况下，序列同一性通过BLAST测量。在一些情况下，SEQ ID NO:1-29中的两个或更多个，或其功能片段或组合，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列，被用作调控元件以与非PV CNS细胞相比增加在PV细胞中转基因的选择性表达，或与非靶细胞类型相比增加在任何靶细胞类型中转基因的选择性表达。在一些情况下，功能片段是一种导致相对于一种或多种非靶

细胞类型,在靶细胞类型中选择性表达的功能片段。

[0112] 在一些情况下,调控元件的两个或更多个拷贝可用于增强在靶细胞中的选择性表达,例如,SEQ ID NO:1-29中任一个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列的两个或更多个拷贝。

[0113] 在其他情况下,SEQ ID NO:1-32中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列,可操作地连接至另一调控元件,如启动子或增强子,以进一步增加在靶细胞中的选择性表达。在一些情况下,可以通过添加一个或多个本文公开的调控元件来改善或增加基因疗法在靶细胞中相对于在非靶细胞中的表达,从而增强任何基因疗法。在一些情况下,靶细胞是表达小清蛋白的PV神经元或GABA能细胞。

[0114] 在一些方面,本文公开的一个或多个调控元件(例如,SEQ ID NO:1-32中的任一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)显示出相对于至少一种、至少两种、至少三种、至少四种、至少五种或超过五种非靶细胞类型,对靶细胞类型的细胞类型选择性。在一些情况下,调控元件驱动在靶细胞亚型中相对于至少一种、至少两种、至少三种、至少四种、至少五种或超过五种非靶细胞亚型或全部其他已知细胞亚型中的选择性表达或优先表达。例如,GABA能细胞包括不同的亚型,包括PV细胞。在一些情况下,靶细胞类型是PV细胞。在一些情况下,相对于对至少一种、至少两种、至少三种、至少四种、至少五种或超过五种非靶细胞类型,一种或多种调控元件对PV细胞具有选择性。在一些情况下,相对于所有其他已知的CNS细胞类型,一种或多种调控元件对PV细胞具有选择性。

[0115] 在一些情况下,SEQ ID NO:1-32中的任一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列,可以与SEQ ID NO:1-32中的任一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列组合。在一些情况下,使用1-50个核苷酸的接头连接这种组合的调控元件。在一些情况下,这种组合的调控元件没有连接在一起。

[0116] 在一些情况下,本文公开的一种或多种调控元件在与任何转基因(例如报告基因转基因或治疗性转基因)可操作地连接时,在至少一种靶细胞类型中驱动选择性表达或优先表达,该表达的水平在统计学上显著高于CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件(例如SEQ ID NO:34,或其片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)在与相同转基因可操作地连接时驱动的表达或由不含调控元件的相同构建体驱动的表达。在一些情况下,统计学上显著更高意味着,调控元件在靶细胞类型中驱动选择性表达,该表达的水平是CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件在与相同转基因可操作地连接时的表达或由不含调控元件的相同构建体的表达的至少1.1、1.2、1.3、1.4、1.5、2、2.5、3、3.5、4、4.5、5、5.5、6、6.5、7、7.5、8、8.5、9、9.5、10、10.5、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100倍。在一些情况下,使用如本文所述的共定位测定来测定此类细胞类型选择性的表达。在一些情况下,靶细胞类型是小清蛋白细胞。在一些情况下,使用抗PV抗体进行这样的共定位测定。在一些情况下,使用本文公开的PV-Cre小鼠进行这样的共定位测定。在一些情况下,非选择性调控元件是SEQ ID

NO:34,或其片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。

[0117] 表达盒

[0118] 术语“表达盒”和“核酸盒”可互换使用,是指多核苷酸分子或核酸序列。在一些情况下,表达盒包含与转基因可操作地连接的一种或多种本文公开的调控元件。在一些情况下,表达盒包含一个或多个调控元件。在一些情况下,表达盒包含本文公开的一种或多种细胞类型选择性调控元件。在一些情况下,表达盒包含本文公开的一个或多个PV细胞选择性调控元件。在一些情况下,表达盒还包含启动子。在一些情况下,表达盒包含SEQ ID NO:1-32的一个或多个序列和/或其任何组合。在一些情况下,表达盒包含(i) SEQ ID NO:1-32中的一个或多个,(ii) 与SEQ ID NO:1-32中的任一个具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列;(iii) (i)或(ii)的任何序列的功能片段;或(iv) (i)、(ii)和/或(iii)的任何序列的组合。在一些情况下,序列同一性通过BLAST测量。在一些情况下,调控元件位于表达盒中转基因的上游。在一些情况下,调控元件位于表达盒中转基因的下游。在一些情况下,表达盒还包含启动子,例如,hSyn1启动子、CBA启动子、CMV启动子、EF1 α 启动子、polyA信号(例如SV40 polyA信号)或转录后调控元件,如土拨鼠肝炎病毒转录后调控元件(WPRE)。

[0119] 在一些方面,本文所述的一个或多个调控元件与表达盒中的转基因可操作地连接。在一些情况下,基因疗法包括表达盒,该表达盒包含与本公开内容的一个或多个、两个或更多个、三个或更多个、四个或更多个或五个或更多个调控元件可操作地连接的转基因,导致该转基因在靶组织或细胞类型如PV神经元中的选择性表达。在一些情况下,表达盒包含与转基因可操作地连接的一个或多个PV细胞选择性调控元件或本文公开的一个或多个调控元件,该转基因例如是报告基因、eGFP、SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白,或其变体或片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。

[0120] 在一些情况下,表达盒适于通过基因疗法递送。在一些情况下,表达盒是线性或环状构建体。在一些情况下,表达盒是质粒、载体、病毒载体或rAAV的一部分。

[0121] 在一些情况下,将基因疗法直接施用于有需要的受试者的CNS,或通过注射和/或输注全身施用。在一些情况下,此类受试者已确诊患有与单倍性不足或遗传突变相关的疾病或病况,例如以下任一种基因的单倍体功能不全或突变:SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3或STXBP1。在一些情况下,受试者有风险发展出,或患有Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作。在一些情况下,使用病毒或病毒载体如rAAV递送此类基因疗法。在一些情况下,使用对CNS细胞具有向性和/或穿过血脑屏障能力的AAV血清型,如AAV9或其变体。

[0122] 在一些情况下,表达盒中与转基因可操作地连接的一个或多个调控元件(例如,SEQ ID NO:1-32中的一个或多个,或其片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)导致相比于非PV CNS细胞或相比于对照元件,在PV细胞中的选择性基因表达,该对照元件例如是CAG、EF1 α 、组成型启动子(例如,SV40、CMV、UBC、PGK和CBA)或非选择性调控元件(例如SEQ ID NO:34,或其功能片段,或与其

具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)。在一些情况下,调控元件导致表达转基因的所有细胞中的至少50%、至少60%、至少70%、至少80%或至少90%是PV神经元。在一些情况下,调控元件导致在PV神经元中的选择性基因表达,该表达约为CNS中PV神经元自然分布的预期表达的约1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、7.5倍、8倍、9倍或10倍。在一些情况下,调控元件驱动在PV细胞中的选择性表达,其中表达转基因的PV细胞的百分比为CNS中PV细胞的预期分布的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍或至少10倍,或者比当转基因与CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或具有SEQ ID NO:34序列,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列的非选择性调控元件可操作地连接时PV细胞中的表达高至少1%-5%、5%-10%、10%-15%、15%-20%、20%-25%、25%-30%、30%-35%、35%-40%、40%-45%、45%-50%、50%-55%、55%-60%、65%-70%、70%-75%、75%-80%、80%-85%、85%-90%或90%-95%,如免疫组织化学共定位测定中所测得的。在一些情况下,表达盒中的调控元件或表达盒中此类调控元件的使用导致在PV细胞或CNS的PV细胞或PV神经元中的选择性基因表达,其中约10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%或95%的表达转基因的细胞为PV阳性。

[0123] 在一些情况下,表达盒或基因疗法包括一种或更多种、两种或更多种、三种或更多种、四种或更多种、五种或更多种、六种或更多种、七种或更多种、八种或更多种、九种或更多种,或十种或更多种表1中所述的调控元件,例如SEQ ID NO:1-32,或功能片段或其组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。

[0124] 在一些情况下,一种或多种PV细胞选择性调控元件或本文公开的一种或多种调控元件与表达盒中的任何转基因可操作地连接。在一些情况下,表达盒是基因疗法。在一些情况下,表达盒是载体或质粒,例如病毒载体或rAAV载体的一部分。在一些情况下,表达盒是AAV1、AAV8、AAV9或AAVDJ或其变体或杂合体的一部分。在一些情况下,表达盒包含与转基因可操作地连接的一种或多种PV细胞选择性调控元件或本文公开的一种或多种调控元件,其中该转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白(例如内源基因的转录调节物),或其变体或功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,相比于在非PV CNS细胞类型中,此类调控元件增加转基因在PV神经元中的选择性表达。在一些情况下,此类调控元件增加转基因在靶细胞类型如PV神经元中的选择性表达。在一些情况下,靶细胞类型是PV细胞。

[0125] 本文考虑到的用于体细胞基因疗法的技术包括通过病毒载体(例如逆转录病毒、腺病毒、AAV、辅助依赖性腺病毒系统、杂合腺病毒系统、单纯疱疹病毒、痘病毒、慢病毒和爱泼斯坦-巴尔病毒)和非病毒系统,如物理系统(裸露的DNA、DNA轰击、电穿孔、流体动力学、超声和磁转染)和化学系统(阳离子脂质、不同的阳离子聚合物,和脂质聚合物)进行递送。

[0126] 载体或病毒表达载体的克隆能力对表达大转基因是一项特殊挑战。例如,AAV载体通常具有约4.8kb的包装能力,慢病毒通常具有约8kb的能力,腺病毒通常具有约7.5kb的能力,而甲病毒通常具有约7.5kb的能力。一些病毒可以具有更大的包装能力,例如疱疹病毒的能力可以大于30kb,而牛痘病毒的能力可以为约25kb。将AAV用于基因疗法的优势包括低致病性、整合到宿主基因组中的频率极低、以及感染分裂和非分裂细胞的能力。

[0127] 为了解决某些病毒载体的大小限制或改善来自病毒载体的表达,本公开内容考虑到使用长度小于2.5kb、2kb、1.5kb、1kb、900bp、800bp、700bp、600bp、500bp、400bp、300bp、200bp、150bp或110bp,但至少为10bp、50bp或100bp的调控元件。在一些情况下,组合的调控元件的大小为约2500bp、2000bp、1500bp、1400bp、1300bp、1200bp、1100bp或1000bp。在一些情况下,每个组合的调控元件的总长度为约100bp、200bp、300bp、400bp、500bp、600bp、700bp、800bp、900bp、1000bp、1100bp、1200bp、1300bp、1400bp、1500bp、1600bp、1700bp、1800bp、1900bp、2000bp、2100bp、2200bp、2300bp、2400bp或2500bp。在一些情况下,组合的RE的大小的总长度为约200bp-3000bp、200bp-2500bp、200bp-2100bp、500bp-2500bp、1000bp-2500bp、1500bp-2500bp、1500b-2000bp或2000bp-2500bp。

[0128] 在一些情况下,本公开内容的调控元件优选是(i)选择性地驱动感兴趣的细胞类型如PV细胞中的表达的调控元件;(ii)包括来源于人类的序列,以及(iii)小于2.5kb、2kb、1.5kb或1kb。

[0129] 本文还考虑了表达盒,其可以是环状或线性核酸分子。在一些情况下,表达盒在载体(例如表达载体)中递送至细胞(例如,多种不同的细胞或细胞类型,包括靶细胞或细胞类型和/或非靶细胞类型)。载体可以是整合载体或非整合载体,该整合是指载体将表达盒和/或转基因整合到细胞基因组中的能力。整合载体或非整合载体都可以用于递送包含与调控元件可操作地连接的转基因的表达盒。载体的实例包括但不限于:(a)非病毒载体,如核酸载体,包括线性寡核苷酸和环状质粒;人工染色体,如人类人工染色体(HAC)、酵母人工染色体(YAC)和细菌人工染色体(BAC或PAC);游离型载体;转座子(例如PiggyBac);和(b)病毒载体,如逆转录病毒载体、慢病毒载体、腺病毒载体和AAV载体。病毒对递送核酸具有若干优点,包括对某些靶细胞或组织的高感染性和/或向性。在一些情况下,使用病毒递送核酸分子或表达盒,该核酸分子或表达盒包含与转基因可操作地连接的本文所述的一个或多个调控元件。

[0130] 病毒基因疗法载体或基因递送载体的优选特征包括可再现、稳定繁殖和纯化至高滴度的能力;介导靶向递送的能力(例如,将转基因特异性递送至感兴趣的组织或器官而没有在他处的广泛载体散播或脱靶递送);以及介导基因递送和/或转基因表达而不会引起有害的副作用或脱靶效应的能力。为避免潜在的有害副作用,可通过将转基因置于细胞类型选择性调控元件的控制下来实现靶向表达或组织/细胞类型选择性表达,该调控元件例如是SEQ ID NO:1-32中的一个或多个,或其功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列、增强子、启动子、稳定性元件、UTR或其组合。例如,可以将含有病毒载体的病毒颗粒设计成感染许多不同的细胞类型,但在感兴趣的细胞类型(例如PV神经元)中提高和/或优化转基因的表达,并且在其他非靶细胞类型(例如非PV CNS细胞)中降低和/或最小化转基因的表达。可以使用对一种或多种细胞类型具有选择性的不同转录因子或调控元件来控制、工程化或操纵转基因在不同细胞类型中的差异表达。在一些情况下,一种或多种调控元件,如启动子或增强子或其组合,与转基因可操作地连接以驱动转基因的组织或细胞选择性表达。在一些情况下,用于基因疗法或载体的一种或多种调控元件以细胞类型选择性的方式驱动基因表达,即赋予靶细胞、细胞类型或组织中的选择性基因表达,并且/或者不驱动基因在一种或多种(例如,至少一种、两种、三种或四种)脱靶细胞或细胞类型中的表达。在一些情况下,与转基因可操作地连接的

一种或多种调控元件增强转基因在靶细胞、细胞类型或组织中的选择性表达,同时该一种或多种调控元件抑制转基因在脱靶细胞、细胞类型或组织中的表达,或赋予一种或多种脱靶细胞、细胞类型或组织显著较低、最小或统计学上较低的基因表达。

[0131] 已将AAV、非致病性细小病毒的几种血清型工程化用于基因递送的目的,其中一些已知对某些组织或细胞类型具有向性。可将用于各种基因疗法应用的病毒工程化为复制缺陷型,或在受试者或宿主中具有低毒性和低致病性。可以通过从病毒基因组中删除全部或部分编码区,并完整保留对于例如将载体基因组包装到病毒衣壳中或将载体核酸(例如DNA)整合至宿主染色质中等功能而言所必需的序列(例如,反向末端重复序列),从而获得此类基于病毒的载体。例如,可以将包含转基因的表达盒克隆到病毒骨架,例如缺少病毒基因的经修饰或工程化的病毒骨架中,并与其他载体(例如包装载体)结合使用,该载体在例如共转染时产生重组病毒载体颗粒。在一些情况下,可穿过血脑屏障或感染CNS细胞的AAV血清型是优选的。在一些情况下,使用AAV9或其变体递送本公开内容的表达盒,该表达盒包含与转基因可操作地连接的一个或多个PV选择性调控元件。

[0132] 如本文所述,使用基因疗法例如rAAV递送本公开内容的表达盒的一个优点是,这种疗法可以提供更有针对性和随时间持续的治疗效果。另外,可以对病毒基因疗法进行工程化,以使其相对于非靶细胞类型或组织,对感兴趣的细胞类型或组织比具有向性。例如,可以对病毒基因疗法进行工程化,以将有效负载或治疗剂例如转录调节物或转基因感染和递送到CNS内的一个或多个区域、组织或细胞类型(例如PV细胞),同时对脱靶组织或细胞类型(例如非CNS组织或细胞类型、非PV CNS细胞)影响最小。在一些情况下,可以将病毒基因疗法工程化,以穿过血脑屏障递送转基因和/或靶向CNS内的特定区域或组织(例如海马体)或CNS内的细胞类型,例如PV细胞。

[0133] 在一些情况下,用于在体内或体外向细胞、细胞类型或组织中递送一种或多种调控元件和转基因的AAV载体或AAV病毒颗粒或病毒粒子优选是复制缺陷的。

[0134] 在一些情况下,表达盒被设计成由AAV或重组AAV(rAAV)递送。在一些情况下,使用慢病毒或慢病毒载体递送表达盒。在一些情况下,优选使用慢病毒或慢病毒载体递送较大的转基因,即超过AAV克隆能力的基因。

[0135] 在本文所述的组合物和方法中使用的AAV可以是任何血清型(例如,AAV1、AAV2、AAV5、AAV8、AAV9和AAVDJ),包括杂合或嵌合AAV血清型。在一些情况下,使用AAV来递送和/或表达与相比于对非PV CNS细胞,对PV神经元具有选择性的一个或多个调控元件可操作地连接的转基因。在一些情况下,使用对CNS细胞具有高向性和/或穿过血脑屏障的AAV。在一些情况下,使用AAV1、AAV8、AAV9和/或AAVDJ递送本文所述的表达盒。

[0136] 在一些情况下,表达盒包含与已知在体内表达不足的转基因可操作地连接的一种或多种PV细胞选择性调控元件或本文公开的一种或多种调控元件,用于例如与基因的单倍性不足相关的疾病或病况。在一些情况下,转基因是电压门控离子通道(例如,钠离子通道或钾离子通道)、神经递质调节因子或其亚基或功能片段。在一些方面,转基因是DNA结合蛋白、离子通道、神经递质调节因子或离子通道或神经递质调节因子的亚基。在一些情况下,转基因是包含一个或多个锌指的DNA结合蛋白。在一些情况下,DNA结合蛋白包含Cas9的结构域、Cas家族蛋白、核酸酶灭活的Cas9(或dCas9)、dCas家族蛋白或类转录激活因子效应物(TALE)。在一些情况下,转基因是DNA结合蛋白,其包含其中DNA切割结构域或核酸酶结构域

已失活的DNA结合蛋白或DNA切割蛋白(例如,核酸酶、限制酶、重组酶等)的DNA结合域,例如,核酸酶失活的Cas(dCas)、失活的类转录激活因子效应物核酸酶,或核酸酶失活的锌指蛋白。在一些情况下,DNA结合域与转录调节域(例如,转录激活或阻遏域)连接。在一些情况下,转基因包括基因编辑蛋白,例如Cas蛋白Cas9。

[0137] 在一些方面,转基因是电压门控离子通道或其亚基,如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2或KV3.3,或其功能片段或变体。在一些情况下,转基因是钠离子通道的 α 亚基。在一些情况下,转基因是钠离子通道的 β 亚基。在一些方面,神经递质调节因子是STXBP1或其功能片段或变体。

[0138] 在一些方面,表达盒以病毒载体如AAV递送。在一些方面,AAV是AAV1、AAV8、AAV9、AAV-DJ、scAAV1、scAAV8或scAAV9。在一些方面,将包含本公开内容的表达盒的基因疗法施用于有需要的受试者(例如,人类患者、哺乳动物、转基因动物或动物模型)。在一些情况下,有需要的受试者有阿尔茨海默病、Dravet综合征、癫痫症、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作的症状、已确诊上述疾病或有发展出上述疾病的风险。在一些情况下,有需要的受试者具有SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3和STXBP1中的任一种或多种的基因表达不足或突变。

[0139] 在一些情况下,使用基因疗法如rAAV9递送表达盒,该表达盒包含可操作地连接转基因的一个或多个PV细胞选择性调控元件或本文公开的一个或多个调控元件,其中转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白或其功能片段,或与SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1、DNA结合蛋白或其功能片段中的任一个具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。在一些情况下,转基因包含与SEQ ID NO:37-43中的任一个具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列,或其功能片段,如以下表2中所提供的。

[0140] 在一些情况下,转基因是SEQ ID NO:36-43中的任一个,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,本文公开的一个或多个调控元件在表达盒中与SEQ ID NO:36-43中的任一个,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列可操作地连接。

[0141] 表2本文公开的氨基酸序列的列表

SEQ ID NO :	基因	氨基酸序列
36	eGFP	MVSKGEELFTGVVPILVELGDVNGHKFSVSGEGEGDATYGKL TLKFICTTGKLPVPWPTLVTLTYGVQCFSRYPDHMKQHDFFKS AMPEGYYQERTIFFKDDGNYKTRAEVKFEGLTVNRIELKGIDF KEDGNILGHKLEYNNYNHNVYIMADKQKNGIKVNFKIRHNIED GSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALKDPNEKRD HMVLL
37	SCN1B	MGTLLALVVGAALVSSAWGGCDEVSDTEAVYGMTFKILCISC KRRSETTAETFTEWTFRQKGTEEFVKILRYENEVLQLEEDERFE GRVVWNGSRGTKDLQDLSIFITNVTYNHSGDYECHVYRLLFFD NYEHNTSVVKIHLLEVVDKANRDMASIVSEIMMYVLIVVLTIW LVAEMVYCYKKIAATEAAAQENASEYLAITSESKENCTGVQV AE
38	SCN2B	MHRDAWLPRPAFLTGLSLFFSLVPPGRSMEVTVPATLNVNGSD ARLPCTFNSCYTVNHKQFSLNWTYQECNNCSEEMFLQFRMKII NLKLERFQDRVEFGNPSKYDVSMLRNQPEDEGIYNCYIMN PPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVIL VLMVVVKCVRRKKEQKLSTDDLKTEEGKTGEGNPDDGAK
39	SCN1A	[0142] MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKD DENGPKPNSDLEAGKNLPIYGDIPPEMVSEPLEDLDPYYINKKT FIVLNKGKAIFRFSAATSALYILTPNPLRKIAKILVHSLFSMLIMC TILTNCVFMTMSNPPDWTKNVEYFTGIYTFESLIKIIARGFCLE DFTFLRDPWNWLDFITFAYVTEFVDLGNVSALRTFRVRLRAL KTISVIPGLKTIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFM GNLRNKCIQWPPTNASLEEHHSIEKNITVNYNGTLINETVFEFDW KSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGR NPNGYTSFDTFSWAFLSLFRLMTQDFWENLYQLTLRAAGKTY MIFFLVLFGLSFYLINLILAVVAMA YEEQNQATLEEAEQKEAEF QQMIEQLKKQQEAAQQAATASEHSREPSAAGRLSDSSSEAS KLSSKSAAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSSIRR KGFRFSIEGNRLTYEKRYSSPHQSLLSIRGSLFSPRRNSRTLSFS RGRAKDVGSENDFADDEHSTFEDNESRRDSLFPVPRRHGERRNS NLSQTSRSSRMLAVFPANGKMHTVDCNGVSVLVGGPSVPTSP VGQLLPEVIIDKPATDDNGTTETEMRKRRSSSFHVSMDFLED SQRQRAMSIA SILTNTVEELESRQKCPCWYKFSNIFIWDCSP YWLKVKHVVNLVVMDFVDLAITICIVLNTLFMAMEHYPMTD HFNNVLTGVNLVFTGIFTAEMFLKIIAMDPPYYFQEGWNIFDF IVTLSLVELGLANVEGLSVLRSFRLRVFKLAKSWPTLNMLIKII GNSVGALGNLTLVLAIIVFIFAVGMQLFGKSYKDCVCKIASDC QLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQAMCL TVFMMVMVIGNLVVLNLFLALLSSFSADNLAA TDDDNEMNN LQIAVDRMHKGVAYVKRKIYEFIQQSFIRKQKILDEIKPLDDLN NKKDSCMSNHTAEIGKLDYLKDVGNGTTSGIGTGSSVEKYIIDE SDYMSFINNPSLT TVVPIAVGESDFENLNTEDFSSESDEESKEK LNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACFTEGCVQRF KCCQINVEEGRGKQWWNLLRRTCFRIVEHNWFETFIVFMILLSSG ALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWKVAYGYQ TYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLAL

SEQ ID NO :	基因	氨基酸序列
[0143]		RPLRALSRFEGMRRVVVNALLGAIPISIMNVLLVCLIFWLIFSIMGVNL FAGKFYHCINTTGDRFDIEDVNNHTDCLKLIERNETARWK NVKVNFDNVGFGYLSLLQVATFKGWMMDIMYAAVDSRNVELQ PKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFGGQ DIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQGMVFDFVT RQVFDISIMILICLNMTMMVETDDQSEYVTTILSRINLVFIVLFT GECVLKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVS PTLFRVIRLARIGRILRIKGAKGIRTLLFALMMSLPALFNIGLLL FLVMFIYIAIFGMSNFAYVKREVGIDDMFNFTFGNSMICLFQITT SAGWDGLLAPILN SKPPDCDPNKVNPGSSVKGDCGNPSVGIFFF VSYIISFLVVVNMYIAVILENFSVATEESAEPLESEDDFEMFYEV WEKFDPDATQFMF EFEKLSQFAALEPPLNPQPNKLQLIAMDL PMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEERFMASNP SKVSYQPIITTLKRKQEEVS AIIQRAYRRHLLKRTVKQASFTY NKNKIKGGANLLIKE DMIIDRINENSITEKT DLT MSTAACPPSYD RVTKPIVEKHEQEGKDEKA KGK
40	STXBP 1	MAPIGLKAVVGEKIMHDVIKKVKKKGEWKVLVVDQLSMRML SSCCKMTDIMTEGITIVEDINKRREPLPSLEAVYLITPSEKSVHSL ISDFKDPTAKYRAAHVFFT DSCP DALF NEL VKSRAAKVIKTLT EINIAFLPYESQVYSLDSADSFSQSFYSPHKAQMKNPILERLAEQI ATLCATLKEYPAVRYRGEYKD N ALLAQLI QDKLDAYKADDPT MGEGPDKARSQ L I LD RGFDPSSPV LHELT FQAMS YD L L PIEND VYKYETSGIGEARVKEVLLDE D D DL WIALRHKHIAEVSQEVTR SLKDFSSSKRMNTGEKTTMRDLSQMLKKMPQYQKELSKYSTH LH LAEDCMKHYQGTVDKLCRVEQDLAMGTDAE GEKIKDPMR AIVPILL DANVSTYDKIRI ILLYIFLKNGITEENLNKLIQHAQIPPE DSEIITNMAHLGVPIVT D STLRRRSKPERKERISEQTYQLSRWTPI IKDIME DTIEDKLDTKHYPYISTRSSASF STT AVSARYGHWHKN KAPGEYRSGPRLIIFILGGVSLNEMRCAYEV TQANGKWEVLIGS THILTPTKFLMDLRHPDFRESSRVS FEDQAPTME
41	Kv3.1	MGQGDESERIVINVGGTRHQT RSTLPGTRLAWLAEPDAH SHFDYDPRADEFFFDRHPGVFAHILNYYRTGKLHCPADVCGPL YEEELAFWGIDETDVEPCCWMTYRQHRDAEEALDSFGGAPLD NSADDADADGP GD SGD GEDELEMTKRLALS D SPDGRPGGFWR RWQPRIWALFEDPYSSRYARYVAFASLFFILVSITFCLETHERF NPIVNKT E IENV RNGTQV RYYREAETEAFLTYIEGVCVWFTFE FLMRVIFCPNKVEFIKNSLNIIDFVAILPFYLEVGLSGLSSKAKD VLGLFLRVVRFVRILRIFKLTRHFVGLRVLGHTLRASTNEFL LIIF LALGVLIFATM IYYAERIGAQPNDPSASEHTHFKNIPIGFWWAV VTMTTLGYGDMPQTWSGMLVGALCALAGVLTIA MPVPVIVN NFGMYYSLAMAKQKLPKKKKHI PRPPQLGSPNYCKSVVNSP HHSTQS DTCPLAQEEILEINRAGRKPLRGMSI
42	Kv3.2	MGKIESNERVILNVGGTRHETYRSTLKTLPGTRLALLAS SEPQG DCLTAAGDKLQPLPPPLSPPP RPPPLSPVPSGC FEGGAGNCSSHG GNGGNNGGSDHPGGGREFFDRHPGVFAYVLNYYRTGKLHCPA DVCGPLFEEELAFWGIDETDVEPCCWMTYRQHRDAEEALDIFE

SEQ ID NO :	基因	氨基酸序列
		TPDLIGGDPGDDEDLAALKRLGIEDAAGLGGPDGKSGRWRKLQP RMWALFEDPYSSRAARFIAFASLFFILVSITTCLETHEAFNIVK NKTEPVINGTSPVLQYEIETDPALTVEGVCVVWFTFEFLVRIVF SPNKLEFIKNLLNIIDFVAILPFYLEVGLSGLSSKAAKDVLGFLRV VRFVRILRIFKLTRHFVGLRVLGHTLRASTNEFLLIIFLAGVLI FATMIYYAERVGAQPNDPSASEHTQFKNIPIGFWAVVTMTTL GYGDMYPQTWSGMLVGALCALAGVLTIAMPPVIVNNFGMY YSLAMAKQKLPRKRKKHIPPAPLASSPTFCKTELNMACNSTQS DTCLGKENRLLEHNRSVLSGDDSTGSEPPLSPPERLPIRRSSTRD KNRRGETCFLLTTGDYTCASDGIRKASTLEPMESTAQTKGDT RPEAHWNCAHLLNFGCPTGSSFPTL
[0144]	43	Kv3.3 MLSSVCVSSFRGRQGASKQQPAPPQPPESPPPPPQPPQQQQPAQ PGPAASPAGPPAPRGPGDRRAEPCPGLPAAAMGRHGGGGGDS GKIVINVGGVRHETYRSTLRTLPGTRLAGLTEPEAAARFDYDPG ADEFFFDRHPGVFAYVLNYYRTGKLHCPADVCGPLFEEELGFW GIDETDVEACCWMTRYRQHRDAEEALDSFEAPDPAGAANAANA AGAHDGGGLDEAGAGGGGLDGAGGELKRLCFQDAGGGAGGP PGGAGGAGGTWWRRWQPRVWALFEDPYSSRAARYVAFASLF FILISITTCLETHEGFIHISNKTVTQASPIPGAPPENITNVEVETEP FLTYVEGVCVVWFTFEFLMRITFCPDKVEFLKSSLNIIDCVAILP FYLEVGLSGLSSKAAKDVLGFLRVVRFVRILRIFKLTRHFVGLR VLGHTLRASTNEFLLIIFLAGVLIFATMIYYAERIGADPDDILG SNHTYFKNIPIGFWAVVTMTTLGYGDMYPKTWSGMLVGAL CALAGVLTIAMPPVIVNNFGMYYSLAMAKQKLKKKNKHPR PPQPGSPNYCKPDPPPPPHPHGSGGISPPPPITPPSMGVTVAG AYPAGPHTHPGLLRRGGAGGLGIMGLPPLPAPGEPCPLAQEEIEI NRADPRPNQDAAAALAHEDCPAIDQPAMSPEDKSPITPGSRGR YSRDRACFLTDYAPSPDGSIRKATGAPPLPPQDWRKPGPPSFL PDLNANAAAWISP

[0145] 在一些情况下,一个或多个PV细胞选择性调控元件包含SEQ ID NO:1-32的序列、其功能片段或组合,或与其具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性的序列。在一些情况下,通过BLAST测量序列同一性。在一些情况下,此类基因疗法被用于治疗癫痫、神经变性、滔蛋白病变、神经元兴奋不足、Dravet综合征和/或阿尔茨海默病。在一些情况下,此类基因疗法被用于治疗与Dravet综合征和/或阿尔茨海默病相关的癫痫和/或癫痫发作。在一些情况下,使用本文所述的基因疗法的治疗导致减少癫痫发作的频率和/或持续时间。在一些情况下,使用本文所述的基因疗法的治疗导致增加体内功能性钠离子通道、功能性钾离子通道或功能性神经递质调节机制的形成。

[0146] 在一些情况下,AAV血清型1、8和/或9,或其杂合物可与本文所述的表达盒一起用于靶向PV细胞中的选择性表达。在一些情况下,设计成通过AAV递送的表达盒包含5' ITR、一个或多个细胞类型选择性调控元件、任选的增强子、任选的最小启动子、转基因、一个或多个任选的内含子、任选的polyA信号和3' ITR。在一些情况下,表达盒可包含5' ITR、两个细胞类型选择性RE、基础启动子、转基因、一个或多个转录后RNA调控元件和3' ITR。

[0147] 示例性AAV表达盒在图7中示出。在一些情况下,表达盒包含5' AAV ITR、增强子(例

如,PV细胞选择性增强子或者一个或多个组合的调控元件)、启动子(例如,一个或多个PV细胞选择性启动子或调控元件)、转基因(例如,SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1或DNA结合蛋白)、转录后调控元件和3' AAV ITR。启动子可以是PV细胞选择性启动子,或组成型启动子。在一些情况下,转基因是报告基因,例如,eGFP、RFP或荧光标记的编码序列。在其他情况下,转基因是调节基因表达的DNA结合蛋白。

[0148] 在一些情况下,该转基因是治疗性转基因,例如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3、STXBP1或DNA结合蛋白的编码序列,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。转录后调控元件可以是影响蛋白质从mRNA表达或RNA稳定性的任何序列,例如,内含子、内部核糖体进入位点(IRES)或土拨鼠肝炎病毒转录后调控元件。在一些情况下,转录后调控元件是两个或更多个转录后调控元件的组合。

[0149] 表达盒可以设计成通过优化的治疗性逆转录病毒载体(例如慢病毒载体)来递送。逆转录病毒载体可以是慢病毒载体,包含左侧的(5')LTR;有助于病毒的包装和/或核导入的序列,至少一个细胞类型选择性调控元件,任选为慢病毒反向应答元件(RRE);任选的启动子或其活性部分;与一个或多个调控元件可操作地连接的转基因;任选地绝缘子;以及右侧的(3')逆转录病毒ITR。

[0150] 在一些情况下,表达盒包含本文公开的一个或多个细胞类型选择性调控元件。在一些情况下,表达盒包含组合的两个或更多个调控元件。在一些实例中,表达盒包含未组合的两个或更多个调控元件,例如,转基因上游的启动子和位于转基因下游的增强子或稳定性元件。

[0151] 在一些情况下,表达盒包含在感兴趣的细胞类型中具有选择性活性的推定的细胞类型选择性调控元件,例如,推定的PV细胞选择性调控元件。包含推定的调控元件的表达盒可包装在病毒载体中并转染到动物模型中,以评估推定的细胞类型选择性调控元件的活性。在一些情况下,通过将包含推定的细胞类型选择性调控元件的载体递送到包含靶细胞或细胞类型的多个细胞或细胞类型中,然后比较推定的调控元件与对照调控元件(如组成型启动子或调控元件,或先前已知的调控元件)的细胞类型选择活性,可以在体外或离体评估推定的细胞类型选择性调控元件。

[0152] 在一些情况下,选择性表达被用于在感兴趣的细胞类型(或感兴趣的组织类型),如CNS中的PV神经元中选择性地表达治疗部分或转基因。在一些情况下,包含与转基因可操作地连接的细胞类型选择性调控元件的载体导致相比于一种或多种(例如,至少两种、三种、四种或五种)其他细胞、细胞类型、组织或组织类型,该转基因在感兴趣的细胞类型中的选择性表达增加,或导致相比于一种或多种细胞或细胞类型,例如,至少一种、两种、三种、四种或五种非靶细胞类型,在感兴趣的细胞类型中优选表达转基因。

[0153] 任何已知技术均可用于将调控元件和转基因或包含调控元件和转基因的组合物递送至感兴趣的细胞(或靶细胞或细胞类型),从而以细胞类型选择性的方式在体外、体内或离体赋予或诱导转基因的表达。

[0154] 包含本公开内容的细胞类型选择性调控元件的表达盒还包含一种或多种转基因。转基因可以是蛋白质编码基因。在一些情况下,表达盒包含转基因。转基因可以替代缺失或有缺陷的基因,或补偿细胞内蛋白质的表达缺陷。转基因可参与细胞信号传导途径。在一些

情况下,转基因可以编码野生型蛋白质、其功能片段、具有增强的治疗特性,例如增强的活性的变体或突变体蛋白质。在一些情况下,转基因可编码DNA结合蛋白,该DNA结合蛋白包含一个或多个锌指,或dCas9的结构域、离子通道(如钾离子通道或钠离子通道)或其亚基单位、神经递质因子或神经递质调节因子。在一些情况下,转基因可以编码离子通道亚基、其变体或突变体。在一些情况下,转基因是DNA结合蛋白,其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如,核酸酶、限制酶、重组酶等)的DNA结合域,例如,核酸酶失活的Cas(dCas)、失活的类转录激活因子效应物核酸酶,或核酸酶失活的锌指蛋白。在一些情况下,DNA结合域与转录调节域(例如,转录激活或阻遏域)相连。在一些情况下,转基因包括基因编辑蛋白,例如Cas蛋白Cas9。

[0155] 本文公开的调控元件可以位于表达载体或表达盒内的任何位置。例如,调控元件可以位于增强子的上游、增强子的下游但在启动子的上游、转基因的5' UTR内、转基因中的内含子内、转基因的3' UTR内或转基因的下游。在一些情况下,一个或多个调控元件位于可操作连接的转基因的上游或下游。

[0156] 在一些实例中,本公开内容的调控元件导致可操作连接的转基因在靶细胞类型(例如PV细胞)中以至少0.5、1.0、1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、2.0、2.5或3IU/ml的水平选择性表达,如通过ELISA测量的。在一些情况下,可以在小鼠中评估调控元件增加转基因表达的能力,其中测量整个小鼠中转基因表达的总量和/或具有转基因表达的细胞类型或组织类型的总数。

[0157] 当评估表达盒或载体的活性时,该活性或表达可表示为单位剂量的活性或表达水平,或相对于所施用或递送至细胞、小鼠或受试者的表达盒或载体的剂量标准化。在一些情况下,将转基因的表达或活性相对于所用的质粒或DNA(例如,每只小鼠的 $\mu\text{g}/\text{kg}$)或病毒颗粒标准化(例如,相对于每个小鼠或受试者的基因拷贝数/kg的量标准化),以允许在具有或不具有调控元件的不同表达载体或表达盒之间进行比较。例如,当评估小鼠中的调控元件的活性时,所测定的PV细胞中的选择性表达或活性可相对于每只小鼠约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或大于10或720 μg 表达载体、表达盒或质粒的剂量标准化。在一些情况下,表达水平或活性可相对于每只小鼠 10^{10} 、 10^{11} 、 10^{12} 、 10^{13} 、 10^{14} 或 10^{15} 个gc/kg病毒颗粒(包含本文公开的表达载体或表达盒的)标准化。

[0158] 在一些方面,表达盒包含与转基因可操作地连接的本文公开的一个或多个调控元件(例如,SEQ ID NO:1-32中的任一个或多个,或功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列),导致相对于在至少一种、至少两种、至少三种、至少四种、至少五种或超过五种非靶细胞类型中,转基因在靶细胞类型中的细胞类型选择性表达或优先表达。在一些情况下,表达盒包含与转基因可操作地连接的一个或多个调控元件,导致相对于在至少一种、至少两种、至少三种、至少四种、至少五种或超过五种非靶细胞亚型或所有其他已知的细胞亚型中,在靶细胞亚型中的细胞类型选择性表达或优先表达。在一些情况下,该转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3和STXBP1中的任一个,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,转基因是调节内源基因(例如内源性SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.2、KV3.3或STXBP1)的DNA结合蛋白。在一些情况下,该转基因是SEQ ID NO:36-43中的任一个,或其功能片段,或与其具有至少80%、至

少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下，转基因是转录调节物。在一些情况下，转基因是DNA结合蛋白，其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白（例如，核酸酶、限制酶、重组酶等）的DNA结合域，例如，核酸酶失活的Cas（dCas）、失活的类转录激活因子效应物核酸酶，或核酸酶失活的锌指蛋白。在一些情况下，DNA结合域与转录调节域（例如，转录激活或阻遏域）相连。在一些情况下，该转基因是基因编辑蛋白，如Cas家族蛋白、Cas9、锌指核酸酶、锌指核酸酶或类转录激活因子效应物核酸酶。在一些情况下，转基因是报告基因或荧光标记。在一些情况下，本文公开的表达盒在病毒载体中。在一些情况下，本文公开的表达盒包装在rAAV如rAAV9或rAAVDJ中。在一些情况下，本文公开的表达盒作为基因疗法被递送到细胞中。在一些情况下，将本文公开的基因疗法递送给受试者，优选人或哺乳动物。在一些情况下，本文公开的表达盒被用于治疗神经系统病况或疾病，如癫痫、神经变性病、滔蛋白病变、神经元兴奋不足、Dravet综合征或阿尔茨海默病。

[0159] 在一些情况下，表达盒（例如，基因疗法、病毒载体、载体或质粒）包含与转基因可操作地连接的SEQ ID NO:1-32中的任一个或多个，或其功能片段或组合，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下，此类组合的调控元件使用1-50个核苷酸的接头连接。在一些情况下，此类组合的调控元件未被连接。在一些情况下，两个或更多个调控元件位于启动子的上游或下游。在一些情况下，两个或更多个调控元件位于转基因的上游或下游。

[0160] 在一些情况下，表达盒包含本文公开的一个或多个调控元件，其当与任何转基因（例如，报告基因转基因或治疗性转基因）可操作地连接时，驱动在至少一种靶细胞类型中的选择性表达或优先表达，该表达的水平在统计上显著高于CAG、EF1 α 、组成型启动子或非选择性调控元件在与相同转基因可操作地连接时驱动的表达，或由不含调控元件的相同构件驱动的表达。在一些情况下，统计学上显著更高意味着，调控元件在靶细胞类型中驱动选择性表达，该表达的水平是CAG、EF1 α 、组成型启动子或非选择性调控元件在与相同转基因可操作地连接时在靶细胞类型中的表达水平或不含调控元件的相同构建体的表达水平的至少1.1、1.2、1.3、1.4、1.5、2、2.5、3、3.5、4、5、5.5、6、6.5、7、7.5、8、8.5、9、9.5、10、10.5、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100倍。在一些情况下，使用本文所述的共定位测定来测定细胞类型选择表达。

[0161] 在其他方面，包含与本文公开的任何转基因可操作地连接的本文公开的一个或多个调控元件的表达盒导致相对于在至少一种、至少两种、至少三种、至少四种、至少五种或超过五种的非靶细胞类型或非靶亚型中，转基因在靶细胞类型中的选择性表达或优先表达。

[0162] 在一些实例中，靶细胞类型是PV细胞。在一些情况下，非靶细胞亚型是本文公开的非PVGABA能亚型中的至少一个、至少两个、至少三个或至少四个。在一些情况下，相对于对所有非PVGABA能细胞或所有非PV CNS细胞，包含本文公开的调控元件的表达盒对PV细胞具有选择性。在一些情况下，根据本文公开的共定位测定来测量细胞类型选择性。在一些情况下，使用在靶细胞类型中表达Cre的老鼠来测量细胞类型选择性。

[0163] 小清蛋白 (PV) 神经元

[0164] GABA能神经元产生 γ -氨基丁酸(GABA)，CNS中的一种主要抑制性神经递质。GABA

对降低整个神经系统的神经兴奋性很重要。GABA通过结合特定的跨膜受体并导致离子通道打开(负面影响膜极化),而作用于抑制性突触。这通常导致细胞的超极化,并增加触发动作电位所需的信号。GABA能神经元中的缺陷可导致兴奋性信号传导与抑制性信号传导失衡,并且牵涉到许多神经系统疾病,包括Dravet综合征、癫痫、神经变性、滔蛋白病变和阿尔茨海默病。其他牵涉到的神经系统病况或疾病包括精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);和/或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或与癫痫相关的病况或疾病。

[0165] 小清蛋白是一种钙结合蛋白,在体感皮层的GABA能中间神经元总数的约40%中表达。在CNS中,PV细胞通常被认为是GABA能细胞。各种研究还发现GABA能细胞包括不同的细胞亚型,包括表达PV、SOM、CR、CCK、NPY、VIP或其组合的细胞。

[0166] PV神经元与各种神经系统疾病或病况,如Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变和/或癫痫发作特别相关。在一些情况下,PV神经元相关的神经系统病况或疾病是精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。在许多方面,靶细胞是CNS中的PV细胞,或表达PV的GABA能细胞。

[0167] 在许多方面,表达PV的中间神经元也被称为篮细胞,其可以根据细胞体的大小(例如大篮细胞、小篮细胞和巢篮细胞)以及树突和轴突进一步细分。生理上,表达PV的篮细胞通常是快放电的(FS),其特征是具有低适应性的一系列高频的动作电位序列(AP)。普遍认为PV篮神经元支配兴奋性锥体神经元的胞体和近端树突。通过表达FS PV的篮神经元介导的前馈抑制可见于包括丘脑皮层回路、经椎板回路和区域间回路在内的若干皮层网络中。FS PV篮神经元强烈抑制邻近的兴奋性锥体神经元。已显示,共享共同的兴奋性输入的PV篮神经元和锥体神经元往往是相互连接的(反馈抑制)。这些连接可以用于调节精确的时间窗,在时间窗内,兴奋性神经元可响应于兴奋性驱动而产生放电。此外,通过介导活性依赖性前馈抑制的高频刺激来抑制对FS PV篮神经元的丘脑皮层和皮层内兴奋性输入。表达PV的篮细胞还支配包括其他篮细胞在内的其他中间神经元,并通过缝隙连接彼此电耦合。已提出,这一特征可能有助于产生和维持皮层网络的同步和振荡。

[0168] 在一些情况下,本文公开的一个或多个调控元件导致相比于在至少一种、至少两种、至少三种、至少四种或至少五种非PV表达神经元中,在PV神经元中的基因表达的选择性增加。在一些情况下,非PV细胞包括所有非PVGABA能细胞。在一些情况下,非PVGABA能神经元包括但不限于表达钙网膜蛋白(CR)、促生长素抑制素(SOM)、胆囊收缩素(CCK)、CR+SOM、CR+神经肽Y(NPY)、CR+血管活性肠多肽(VIP)、SOM+NPY、SOM+VIP、VIP+胆碱乙酰转移酶(ChAT)、CCK+NPY、CR+SOM+NPY以及CR+SOM+VIP的细胞。

[0169] 在一些情况下,CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或以非细胞类型选择性的方式驱动基因表达的非选择性调控元件中的任一个均可用于与本文公开

的PV选择性调控元件或任何细胞类型选择性调控元件进行比较。在一些情况下，调控元件导致在PV细胞中的选择性表达的水平高于与CAG或EF1 α 对照可操作地连接的基因的表达水平则表明对PV细胞的选择性。在一些情况下，本文公开的调控元件在PV细胞中显示选择性基因表达，该表达比与CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件(例如，SEQ ID NO:34，或其功能片段，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)可操作地连接的转基因的PV表达水平高至少2%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%或至少95%，如通过共定位测定所测量的。在一些情况下，本文公开的调控元件在PV细胞中显示出选择性基因表达，该表达是CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件(例如，SEQ ID NO:34，或其功能片段，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)下的表达水平的至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍或至少10倍，如通过本文所述的共定位测定所测量的。

[0170] 在优选的情况下，相对于在至少一种其他CNS细胞类型(例如，至少两种、至少三种、至少四种、至少五种非PV细胞，或两种或更多种、三种或更多种或四种或更多种非PV细胞和/或非PVGABA能神经元)中，本文所述的一种或多种调控元件选择性地驱动转基因在GABA能细胞(如表达小清蛋白的GABA能细胞)中的表达，。

[0171] 在一些情况下，靶细胞类型是表达小清蛋白的GABA能神经元，或PV细胞。

[0172] 在大脑的细胞亚群中选择性地表达转基因的一种方法是使用病毒载体，该病毒载体包含与细胞类型选择性调控元件或与在大脑的细胞亚群(例如PV细胞)中具有选择性(或具有选择活性)的调控元件可操作地连接的转基因。可以选择具有高感染性且对特定的细胞类型没有选择性的病毒载体，同时调控元件赋予选择性。例如，细胞类型选择性调控元件可以在PV神经元而不是其他神经元中驱动转基因的表达。

[0173] 在一些情况下，本公开内容涉及使用选择性驱动在PV神经元中表达的调控元件(即PV细胞选择性调控元件)。

[0174] GABA能细胞是产生 γ -氨基丁酸的抑制性神经元。通过谷氨酸脱羧酶2(GAD2)的表达可鉴定GABA能细胞。GABA能细胞的其他标记包括GAD1、NKX2.1、DLX1、DLX5、SST、PV和VIP。

[0175] 在一些情况下，非PV CNS细胞是兴奋性神经元、多巴胺能神经元、星形胶质细胞、小胶质细胞、运动神经元或血管细胞。在一些情况下，非GABA能神经元是不表达GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP中的一种或多种的细胞。在一些情况下，非PV神经元是不表达小清蛋白的GABA能神经元。在一些情况下，其他CNS细胞是指从未表达过PV、GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP中的任一种的CNS细胞类型。

[0176] 在一些情况下，相对于对至少一种、两种、三种、四种、五种或超过五种非PV CNS细胞类型，本文公开的调控元件对PV表达细胞具有选择性。在一些情况下，非PV细胞类型包括非PVGABA能细胞。在一些情况下，感兴趣的细胞类型是PV细胞。在一些情况下，对PV细胞具有选择性的RE被称为PV细胞选择性调控元件。

[0177] 在一些情况下，本文公开的PV细胞选择性调控元件包含SEQ ID NO:1-32或其任何组合。

[0178] 在一些情况下,一种或多种PV细胞选择性调控元件或本文公开的一种或多种调控元件被用于使PV表达细胞中转基因的表达与不含调控元件时的表达相比增加至少2、5、10、15、20、30、40、50、60、70、80、90、100倍或更多倍。在一些情况下,表达盒中的RE使基因表达与不含调控元件的表达相比增加至少1.5%、2%、5%、10%、15%、20%或50%,或增加超过1.5%、2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或100%。在一些情况下,组合物及其使用方法包括含有一个或多个调控元件的表达盒,该调控元件导致转基因表达(例如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KCNC1(也称为KV3.1)、KCNC3(也称为KV3.3)、STXBP1、DNA结合蛋白或其变体或功能片段或其蛋白质的表达)与不含调控元件时的水平相比或与非选择性调控元件(例如CAG、EF1 α 、组成型启动子或SEQ ID NO:34,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)相比增加10-500%。在一些情况下,与没有表达盒或调控元件时的水平相比,SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KCNC1、KCNC3和STXBP1中任一个的基因表达和/或蛋白质水平增加为1.5-5%、5%-10%、10-15%、15-20%、20-30%、30-40%、40-50%、50-60%、60-70%、70-80%、80-90%、90-100%、100-150%、150-200%、250-300%、300-350%、350-400%、400-450%、450-500%,或1.5-20%、20%-50%、50%-100%、100-200%、200-300%、300-400%或400-500%。在一些情况下,与包含对照(例如CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)或非选择性调控元件)或非细胞类型选择性调控元件(例如SEQ ID NO:34,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列)的表达盒的表达相比,该基因或蛋白质表达在PV细胞中是选择性的。

[0179] 在一些情况下,通过将表达PV和eGFP二者(与一个或多个调控元件可操作地连接的转基因)的细胞数除以表达eGFP的细胞总数,再乘以100来换算成百分数,即可计算出PV细胞中表达的选择性。本文所述的PV细胞选择性调控元件可在PV细胞中高选择性地表达。例如,PV细胞选择性调控元件或本文公开的一个或多个调控元件对PV神经元可表现出约50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或大于约99%的选择性。

[0180] 在一些情况下,PV细胞选择性调控元件或本文公开的任何调控元件赋予在PV神经元中表达转基因的选择性,该表达的水平在统计学上高于对照调控元件,例如EF1 α 或先前已知的调控元件。在一些实例中,PV细胞选择性调控元件与对照调控元件之间的统计学差异为至少2倍、5倍、10倍、20倍或超过2倍的差异,或超过5倍、10倍或20倍的差异,如通过本文所述的任一种方法如共定位测定所确定的。

[0181] 本公开内容包括对PV细胞有选择性的调控元件。这些PV细胞选择性RE或任何细胞类型选择性RE优选较短,优选小于约1100个碱基对、1000bp、900bp、800bp、700bp、600bp、500bp、400bp、300bp、200bp或小于约110bp。PV细胞选择性RE或任何细胞类型选择性RE可在1050bp至100bp之间、100bp至500bp之间或500bp至1050bp之间。PV细胞选择性调控元件的一些实例由SEQ ID NO:1-32或其功能片段或组合提供。本公开内容所考虑到的其他PV细胞选择性调控元件包含与本文所述的这些序列中的任一个具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列,或本文所述的序列之一的部分或片段。

[0182] 在一些情况下,PV细胞选择性调控元件或本文公开的任何调控元件与本文所述的

序列或本文所述的序列的片段具有至少约70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或超过99%的同一性。在一些情况下,PV细胞选择性调控元件与本文所述的序列或其功能片段中的至少约30%、40%、50%、60%、70%、80%、90%或95%具有至少约80%的同一性。

[0183] 在一些情况下,PV细胞选择性调控元件与SEQ ID NO:1-32中的任一个或多个,或其功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列具有至少80%的同一性。在一些情况下,PV细胞选择性调控元件与SEQ ID NO:1-22中的50%或更多的序列具有90%的同一性。在一些实例中,PV选择性调控元件是SEQ ID NO:1-32中的任一个的功能片段或其组合。在一些情况下,功能片段能够在PV细胞中选择性表达转基因,在PV细胞中表达的选择性为至少50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或大于95%。

[0184] 在一些情况下,将本公开内容的两个或更多个PV细胞选择性调控元件或本文公开的任何两个或更多个调控元件组合以形成组合调控元件。在一些情况下,将两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个,或十个或更多个PV细胞选择性调控元件或本文公开的多个调控元件组合。例如,SEQ ID NO:31是SEQ ID NO:1和SEQ ID NO:23-29的组合。另一个实例,SEQ ID NO:32是SEQ ID NO:8和SEQ ID NO:23-29的组合。在一些情况下,可以将两个或更多个PV选择性调控元件的片段组合以形成组合调控元件。例如,50%的SEQ ID NO:1可以与30%的SEQ ID NO:8和90%的SEQ ID NO:30组合以形成组合调控元件。

[0185] 在一些情况下,本公开内容的一个或多个PV细胞选择性调控元件或本文公开的任一个或多个调控元件相比于在一个或多个其他CNS细胞类型中,在PV神经元中选择性地表达可操作地连接的转基因。这种选择性表达可以通过计算表达可检测水平的所连转基因的PV神经元的数目与表达转基因的细胞总数(包括表达转基因的非PV神经元的数目)的百分比来定量。换言之,PV调控元件在特定细胞类型或靶细胞中的选择性可以通过测量和/或比较表达与调控元件可操作地连接的转基因的PV神经元(或靶细胞)的数目与表达该转基因的非靶细胞类型的数目(或表达该转基因的细胞总数)来确定。

[0186] 在一些情况下,PV细胞选择性调控元件对PV神经元、CNS中的PV神经元或同样表达PV的GABA能神经元表现出约50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或大于约99%的选择性。在一些情况下,与CAG或EF1 α 或非细胞类型选择元件相比,或与非PV CNS细胞相比,或与CNS中的至少一种、至少两种、至少三种、至少四种或至少五种其他非PVGABA能神经元亚型相比,本公开内容的一种或多种调控元件可表现出约50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或大于约99%的PV神经元选择性。

[0187] 在一些情况下,PV选择性调控元件赋予在PV神经元中表达转基因的选择性,该表达的水平在统计学上高于对照调控元件,例如,CAG、EF1 α 、组成型启动子(例如SV40、CMV、UBC、PGK和CBA)、非选择性调控元件或先前已知的调控元件。在一些情况下,如通过本文所述的任一种方法所确定的,PV细胞选择性RE与对照元件之间的统计学差异为至少2倍、5倍、10倍、20倍或超过2倍的差异,或超过5倍、10倍或20倍的差异。在一些情况下,使用本文所述的共定位测定来测量PV的选择性。

[0188] 在一些方面,本文所述的细胞类型选择性调控元件可用于相比于在其他CNS细胞类型中,选择性地调节转基因在CNS细胞类型中的表达。例如,本文所述的细胞类型选择性调控元件可用于相比于在包括其他类型的神经元在内的其他CNS细胞中,选择性地调节转基因在PV细胞中的表达。对于基因疗法,可能期望在靶细胞类型中选择性地表达转基因和/或在非靶细胞类型中最小化表达转基因。在非预期细胞类型(例如非靶细胞类型)中表达转基因可能对受试者产生负面影响。在非预期细胞类型中表达转基因可能抵消转基因在预期细胞类型中的治疗效果。例如,预期在PV细胞中表达的转基因如果在谷氨酸能神经元中表达,则可能对受试者产生负面影响。本文所述的细胞类型选择性调控元件可用于表达盒中,以确保转基因的适当表达和/或减少基因疗法的脱靶效应。

[0189] 本文中的细胞类型选择性调控元件可用于基因表达盒中,由此该调控元件与一个或多个转基因可操作地连接。使用这样的基因表达盒来将转基因递送到细胞中以进行表达。表达盒可包含如本文所述的细胞类型选择性调控元件、细胞类型选择性调控元件的组合,或与转基因可操作地连接的如本文所述的细胞类型选择性调控元件的片段。

[0190] 优选地,本文中的表达盒包含与转基因可操作地连接的一个或多个细胞类型选择性调控元件,二者不在体内的内源性环境中共同作用。例如,钠离子通道β亚基的转基因如SCN1B可以与不在体内的同一环境中起作用的或者不能可检测地且内源性地驱动SCN1B的表达的一个或多个调控元件可操作地连接。类似地,核酸盒可包含与调控元件可操作地连接的神经递质调节因子,如STXBP1,该调控元件不在相同的内源性或体内环境中起作用,或不在同一开放阅读框中,或不在同一人类染色体上,或不在体内可检测地驱动STXBP1的表达。在一些情况下,细胞类型选择性调控元件与转基因相连接,其中该细胞类型选择性调控元件在体内不调节与转基因相对应的内源基因。

[0191] 在一些方面,本文公开的细胞类型选择性调控元件来源于从人类染色体基因座分离的序列,该基因座不同于与转基因相对应的天然基因的基因座。因此,在一些情况下,表达盒包含具有来源于染色体的序列的细胞类型选择性调控元件,该染色体不同于与同一表达盒上的转基因相对应的染色体。在其他情况下,本公开内容的调控元件和在表达盒中可操作地连接的转基因来源于人类基因组中位置相距超过20kb的序列或来源于在远端基因组位置的序列。当在表达盒上使用两个或更多个来源于人类的调控元件时,该两个或更多个调控元件可以具有在人类基因组中位置相距超过5kb、超过10kb、超过15kb或超过20kb的序列,或者其中该两个或更多个调控元件在基因组中彼此不天然地相互作用。

[0192] 在一些情况下,包含PV选择性调控元件的表达盒可以排除来源于hSyn1或GAD2启动子序列的已知序列。在一些实例中,PV选择性调控元件不包含GAD2、GAD1、SYN1、NKX2.1、DLX1、DLX5、SST和VIP启动子中任一个的完整启动子序列。在一些情况下,PV选择性调控元件不包含来自GAD2、GAD1、SYN1、NKX2.1、DLX1、DLX5、SST和VIP启动子中的一个或多个的序列的超过500个连续碱基对。在一些情况下,PV选择性调控元件不包含位于GAD2、SYN1、NKX2.1、DLX1、DLX5、SST和VIP中任一个的转录起始位点的1kb、2kb、3kb、4kb、5kb、6kb、7kb、8kb、9kb或10kb内的序列。

[0193] 在一些情况下,转基因可用于与治疗感兴趣的特定细胞类型相关的疾病。在一些情况下,感兴趣的细胞类型是神经元、抑制性神经元、GABA能神经元或PV神经元。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3和STXBP1中的任一个或多

个。在一些情况下，转基因是调节基因表达的DNA结合蛋白(例如，调节内源基因表达的转录激活物或转录阻遏物)。在一些情况下，转基因是基因编辑蛋白，如锌指核酸酶、类转录激活因子效应物核酸酶、Cas家族蛋白。在一些情况下，转基因是报告基因或可检测标记，如eGFP、tdTomato或RFP。在一些情况下，转基因是Cas蛋白，如Cas9。

[0194] 可将用于治疗与PV神经元细胞相关的病况的转基因并入载体、核酸盒或本文所述的方法中。本文使用的转基因通常不包含内含子，或不包含超过一个内含子。转基因可以从cDNA序列而不是基因组序列中获得。在一些情况下，转基因可以包含其内源性内含子中的一些或全部。在一些实例中，这种转基因编码DNA结合域或离子通道。可在本公开内容的表达盒中编码的DNA结合域的实例包括锌指、Cas9、Cas家族蛋白、dCas9、dCas家族蛋白或类转录激活因子效应物(TALE)。在一些情况下，转基因是DNA结合蛋白，其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如，核酸酶、限制酶、重组酶等)的DNA结合域，例如，核酸酶失活的Cas(dCas)、失活的类转录激活因子效应物核酸酶，或核酸酶失活的锌指蛋白。在一些情况下，DNA结合域与转录调节域(例如，转录激活或阻遏域)相连。在一些情况下，转基因包括基因编辑蛋白，例如Cas蛋白Cas9。在一些情况下，转基因是离子通道或膜蛋白的亚基或组分，或是与本文公开的神经系统病况或疾病相关的基因。可用于本公开内容的表达盒中的离子通道转基因的实例包括电压门控和配体门控的离子通道。电压门控离子通道包括钠通道、钙通道、钾通道和质子通道。在一些情况下，转基因编码电压门控钠通道的亚基。电压门控钠通道亚基的实例包括SCN1B(NM_001037.4)、SCN1A(NM_001165963.1)和SCN2B(NM_004588.4)。

[0195] 在一些情况下，转基因编码电压门控钾通道的亚基。电压门控钾通道亚基的实例包括KCNC1(NM_001112741.1)和KCNC3(NM_004977.2)。在一些情况下，转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1中的任一个或多个、其变体和功能片段。在一些情况下，该转基因是与SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1中的任一个、其变体或功能片段具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%序列同一性的序列。在一些情况下，使用BLAST测量序列同一性。

[0196] 在一些情况下，本文公开的表达盒包含与转基因可操作地连接的一个或多个PV选择性调控元件或本公开内容的一个或多个调控元件，该转基因与编码SEQ ID NO:37-43中的任一个或其功能片段或变体的序列或对应于SEQ ID NO:37-43的GenBank序列具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%的序列同一性。在一些情况下，本文公开的表达盒包括转基因，该转基因具有根据以下各项的序列：(i) SEQ ID NO:37-43的序列，或(ii)其功能片段，或(iii)与(i)或(ii)具有至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或至少99.5%序列同一性的序列。

[0197] 在一些实例中，转基因是神经递质调节因子，或其变体或功能片段。神经递质调节因子可参与调节CNS中神经递质的产生或释放。例如，神经递质调节因子可有助于突触融合以释放神经递质。神经递质调节因子的实例是STXBP1(NM_001032221.3)，或其功能片段，或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。转基因

还可以是神经递质调节因子的一个亚基。

[0198] 在一些情况下,本公开内容的表达盒可以包含AAV2 5' ITR、PV选择性增强子、PV选择性启动子,或一个或多个PV选择性启动子和增强子的组合、SCN1B的cDNA、WPRE、hGH polyA信号、PV选择性调控元件以及AAV2 3' ITR。在一个实例中,表达盒包含AmpR启动子和AmpR编码序列、细菌复制起点、AAV2 ITR、SEQ ID NO:8、SEQ ID NO:23-29、转基因(编码序列、WPRE、人类生长激素polyA信号、AAV2 ITR和f1起点)。

[0199] 作为另一示例,本公开内容的表达盒可以包含AAV2 5' ITR、增强子、启动子、内源性SCN1A基因的转录激活物、WPRE、hGH polyA信号、调控元件和AAV2 3' ITR。在一些情况下,表达盒包含AAV2 5' ITR、启动子、内含子元件、转录修饰因子、合成polyA和AAV2 3' ITR。在一些情况下,本公开内容的表达盒可以包含AAV2 5' ITR、PV选择性增强子、PV选择性启动子、编码SCN1A或SCN1B的转录激活物的序列、WPRE、hGH polyA信号、PV选择性调控元件或本文公开的任何调控元件,以及AAV2 3' ITR。

[0200] 包含本公开内容的一个或多个调控元件的表达盒可用于治疗医学状况。在一些情况下,包含本公开内容的调控元件的表达盒被用于治疗神经系统病况或神经变性病况。神经系统病况可能由已知的遗传事件引起,或者可能有未知的原因。

[0201] 神经系统病况可能是与PV神经元相关的疾病。神经系统病况可能是与抑制性神经元如PV神经元相关的疾病。与PV神经元相关的疾病或病况可以通过将携带转基因和一个或多个PV细胞选择性调控元件或本文所述的任何调控元件的表达盒递送到体内细胞来治疗。在一些方面,包含与一个或多个PV细胞选择性调控元件或本文公开的任何调控元件可操作地连接的转基因的表达盒可用于治疗Dravet综合征、阿尔茨海默病、癫痫、神经变性病症、滔蛋白病变、神经元兴奋不足和/或癫痫发作。在一些情况下,本公开内容的表达盒被用于治疗精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,Dravet综合征、慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);和/或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。

[0202] 大多数Dravet综合征病例与SCN1A和/或SCN2A基因的突变有关。SCN1A的突变或异常也与癫痫发作病症、癫痫、自闭症、家族性偏瘫性偏头痛3型(FHM3)、全面性癫痫伴热性惊厥附加症(GEFS+)以及某些抗癫痫发作药物的有效性有关。例如,SCN1A的ICS5N+5G>A突变与抗癫痫发作药物苯妥英和卡马西平的最大安全量(剂量)有关。

[0203] 在一些阿尔茨海默病患者中,β-淀粉样蛋白(Aβ)的产生涉及许多可以影响神经元的兴奋性的肽和蛋白酶,导致癫痫发作和PV神经元中Nav1.1钠通道的下调。

[0204] 与功能失调的PV神经元相关的疾病,如SCN1A或Nav1.1的功能丧失突变引起的疾病包括:Dravet综合征、大田原综合征、癫痫、早期婴儿型癫痫性脑病6(EIEE6)、家族性热性惊厥3A(FEB3A)、难治性儿童癫痫伴全面性强直-阵挛发作(ICEGTC)、偏头痛、家族性偏瘫性偏头痛3型(FHM3)、Panayiotopoulos综合征、家族性心房颤动13(ATFB13)、全面性癫痫伴热性惊厥附加症1型(gefs+1型)、布鲁加达综合征、非特异性心脏传导缺陷、全面性癫痫伴热性惊厥附加症、良性家族性婴儿惊厥、早期婴儿型癫痫性脑病11(EIEE11)、良性家族性婴儿癫痫、神经变性、滔蛋白病变和阿尔茨海默病。在一些情况下,神经系统病况是Dravet综合

征。Dravet综合征与SCN1A和/或SCN2A基因的突变有关。在一些情况下,本公开内容的一个或多个调控元件被用于基因疗法或表达盒,以治疗与PV神经元相关的神经系统病况或疾病,例如,精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,Dravet综合征、慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,本公开内容的一个或多个调控元件(例如PV神经元选择性调控元件)被用于治疗Dravet综合征和/或阿尔茨海默病(例如,在表达盒、载体或基因疗法中)。在一些情况下,神经系统病况或疾病是牵涉PV神经元的任何与癫痫发作和/或癫痫相关的病况或疾病。

[0205] 本公开内容的方法和组合物可用于治疗已被诊断患有疾病例如神经系统疾病或神经变性疾病受试者。受试者可以是患有某种形式的癫痫的患者。在一些情况下,受试者是患有Dravet综合征的患者。受试者可以是患有神经变性疾病的患者,例如,阿尔茨海默病患者。在一些情况下,癫痫、脑病和/或癫痫发作与SCN8A内的基因突变有关。在一些情况下,SCN8A内的基因突变可能导致癫痫综合征,例如,Dravet综合征。在一些实例中,STXBP1内的基因突变与以反复癫痫发作为特征的癫痫性脑病相关。

[0206] 在一些情况下,用本文所述的一种或多种组合物治疗的受试者是被诊断为在离子通道或神经递质调节因子(例如,突触融合蛋白结合蛋白)中有突变或遗传异常。此类突变的实例包括SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KCNC1、KCNC3和/或STXBP1的突变或其组合。包含如本文所述的细胞类型选择性调控元件的表达盒可被递送给受试者以治疗或预防具有与特定细胞类型相关的症状的疾病。例如,包含与一个或多个PV细胞选择性调控元件可操作地连接的转基因的表达盒被递送给具有与PV神经元相关的症状或有发展出该症状的风险的受试者。

[0207] 在一些情况下,该治疗可以施用于患有Dravet综合征或有发展出Dravet综合征的风险的受试者。与Dravet综合征相关的症状包括癫痫发作、记忆缺陷、发育迟缓、肌张力差和/或认知问题。用本公开内容的表达盒治疗可导致一个或多个症状的改善,如癫痫发作的次数、持续时间和/或强度的减小。对有发展出Dravet综合征的风险的受试者施用本文所述的基因疗法可防止发展出一个或多个症状或减缓一个或多个症状的进展。

[0208] 在另一个实例中,该治疗可施用于患有阿尔茨海默病的受试者。与阿尔茨海默病相关的症状包括短期记忆丧失、认知困难、癫痫发作,以及语言、执行功能、感知(失认)和动作执行(失用)存在困难。用本公开内容的表达盒治疗可导致一个或多个阿尔茨海默病症状的改善,如记忆丧失的进展减慢,或防止一种或多种症状。在一些情况下,该治疗可以校正高 γ 功率的大脑活动。治疗可降低癫痫发作频率和/或发作严重程度,或将高 γ 功率活动减少10%、20%、30%、40%、50%、60%或70%。在一些情况下,治疗可以改善认知功能。学习和/或记忆可被改善10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或超过100%。

[0209] 本公开内容的方法和组合物可用于治疗有发展出疾病的受试者。可以已知受试者易患某种疾病,例如,神经系统疾病或与癫痫、癫痫发作、和/或脑病有关的疾病。受试者可能由于遗传事件或已知的风险因素而易患疾病。例如,受试者可能携带与Dravet综合征相关的SCN1A突变。在一些情况下,由于受试者的年龄,受试者可能易患例如阿尔茨海

默病等疾病。

[0210] 治疗可以减少或停止症状。例如,治疗可以改善学习、记忆、认知功能和/或运动功能;减少癫痫发作的频率和/或持续时间;和/或降低温度敏感性(或提高触发癫痫发作的温度阈值)。

[0211] 在一些情况下,本文公开的基因疗法或表达盒的靶细胞类型是PV细胞。在一些情况下,非靶细胞亚型是本文公开的非PVGABA能亚型中的至少一个、至少两个、至少三个或至少四个。在一些情况下,相对于对所有非PV CNS细胞,包含本文公开的调控元件的表达盒对PV细胞具有选择性。在一些情况下,根据本文公开的共定位测定来测量细胞类型选择性。在一些情况下,使用在靶细胞类型中表达Cre的小鼠来测量细胞类型选择性。

[0212] 在一些情况下,治疗不会对受试者产生不良反应。与使用包含与非选择性调控元件连接的相同转基因的类似基因疗法相比,使用包含PV选择性调控元件的基因疗法可对受试者造成更少或严重程度更低的不良反应。

[0213] 在许多方面,本文公开的任何表达盒均可适于或用于基因疗法(例如rAAV或rAAV9基因疗法)以治疗Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作中的任一种或多种。在一些情况下,基因疗法包含与转基因可操作地连接的SEQ ID N0:1-32中的任一个或多个、两个或更多个、三个或更多个、四个或更多个,五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个或十个或更多个,或其功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,基因疗法包括本公开内容的表达盒。在一些情况下,基因疗法包含与任何转基因(例如,报告基因转基因或治疗性转基因)可操作地连接的本文公开的一个或多个调控元件,使得调控元件驱动至少一种靶细胞类型中的选择性表达或优先表达,该表达的水平在统计上显著高于CAG、EF1 α 或非选择性调控元件在与相同转基因可操作地连接时所驱动的表达或由不含调控元件的相同构建体所驱动的表达。在一些情况下,在统计学上显著更高意味着,调控元件在靶细胞类型中驱动选择性表达,该表达的水平是CAG、EF1 α 、组成型启动子或非选择性调控元件在与相同转基因可操作地连接时在靶细胞类型中的表达水平或不含调控元件的相同构建体的表达水平的至少1.1、1.2、1.3、1.4、1.5、2、2.5、3、3.5、4、5、5.5、6、6.5、7、7.5、8、8.5、9、9.5、10、10.5、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100倍。在一些情况下,使用本文所述的共定位测定来测定这样的细胞类型选择表达。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1和其功能片段中的任一个或多个。在一些情况下,转基因是调节内源基因的表达的DNA结合蛋白,如转录调节物、转录激活物或转录阻遏物。在一些情况下,转基因是DNA结合蛋白,其包含其中DNA切割结构域或核酸酶结构域已失活的DNA结合蛋白或DNA切割蛋白(例如,核酸酶、限制酶、重组酶等)的DNA结合域,例如,核酸酶失活的Cas (dCas)、失活的类转录激活因子效应物核酸酶,或核酸酶失活的锌指蛋白。在一些情况下,DNA结合域与转录调节域(例如,转录激活或阻遏域)相连。在一些情况下,转基因是基因编辑蛋白,如锌指核酸酶或类转录激活因子效应物核酸酶。在一些情况下,转基因是报告基因或可检测标记,如eGFP、tdTomato或RFP。在一些情况下,转基因是Cas蛋白,如Cas9。

[0214] 在一些情况下,包括本文公开的表达盒的基因疗法被用于治疗神经系统病况或疾

病。在一些情况下,包含本文公开的表达盒的基因疗法被用于治疗神经系统病况或疾病,其中表达盒包含与转基因可操作地连接的SEQ ID NO:1-32中的任一个或多个、两个或更多个、三个或更多个、四个或更多个、五个或更多个、六个或更多个、七个或更多个、八个或更多个、九个或更多个或十个或更多个,或其功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1、DNA结合蛋白和其功能片段中的任一个或多个。在一些情况下,包含本文公开的表达盒的基因疗法被用于治疗Dravet综合征。在一些方面,包含本文公开的表达盒的基因疗法被用于治疗阿尔茨海默病。在一些情况下,包含本文公开的表达盒的基因疗法被用于治疗与Dravet综合征和/或阿尔茨海默病相关的癫痫和/或癫痫发作症状。在一些情况下,治疗Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作中的任一种包括将本公开内容的基因疗法递送或施用至需要的受试者的细胞。在一些情况下,有需要的受试者有发展出Dravet综合征、阿尔茨海默病、癫痫和/或癫痫发作中任一种的风险或患有上述任一种。在一些情况下,受试者是儿童或未成年人。在一些情况下,包含本文公开的表达盒的基因疗法被用于治疗被诊断患有Dravet综合征或有发展出Dravet综合征的风险的婴儿、儿童或未成年人。在一些情况下,包含本文公开的表达盒的基因疗法被用于治疗具有SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3和STXBP1中的任一个或多个中的突变或遗传缺陷的受试者。

[0215] 在一些方面,本公开内容提供了治疗Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作中任一种的方法,包括向受试者的细胞施用基因疗法,其中基因疗法包括本文公开的表达盒。在一些情况下,这种表达盒包含与转基因可操作地连接的SEQ ID NO:1-32中的任一个或多个,或功能片段或组合,或与其具有至少80%、至少85%、至少90%、至少95%或至少99%序列同一性的序列,其中转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1及其功能片段中的任一个。在一些情况下,转基因是钠离子通道或钾离子通道的亚基。在一些情况下,转基因为突触融合蛋白结合蛋白。在一些情况下,转基因为转录调节物,例如,转录激活物或转录阻遏物。在一些情况下,转基因是调节内源基因(例如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或STXBP1)表达的转录调节物。在一些情况下,转基因是基因编辑蛋白,如锌指核酸酶或类转录激活因子效应物核酸酶。在一些情况下,转基因是Cas蛋白,如Cas9。

[0216] 在其他方面,本公开内容提供了一种通过添加本文公开的一种或多种调控元件来修改设计用于治疗Dravet综合征、阿尔茨海默病、癫痫、神经变性、滔蛋白病变、神经元兴奋不足和/或癫痫发作的任何基因疗法,以改善该基因疗法的细胞类型选择性的方法。在一些情况下,基因疗法是rAAV基因疗法。

[0217] 在一些情况下,与未治疗对照或治疗前水平相比,用本文公开的表达盒治疗将癫痫发作的持续时间和/或频率(例如,与Dravet综合征相关的癫痫发作)减少至少1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%或95%。

[0218] 在一些情况下,与未治疗对照或治疗前水平相比,用本文公开的表达盒治疗将高

γ 功率活动(例如,与阿尔茨海默病相关的高 γ 功率活动)降低至少1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%或95%。

[0219] 一方面,本公开内容提供了本公开内容的核酸盒,该核酸盒包含与转基因可操作地连接的一个或多个调控元件,该控元件导致相对于在一个或多个非靶细胞类型中,在靶细胞类型中的选择性表达,例如,相对于在一个或多个非PV CNS细胞类型中,在CNS中的PV 神经元中的选择性表达。在一些情况下,每个调控元件包含(i) SEQ ID NO:1-32的序列,(ii)其功能片段或组合,或(iii)与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,序列同一性百分比可使用BLAST测量。在一些情况下,至少一个调控元件是来源于人类的。在一些情况下,至少一个调控元件来源于非人类哺乳动物。在一些情况下,调控元件是非天然存在的。在一些情况下,调控元件导致在PV细胞中的选择性表达,该表达大于与 CAG或EF1 α 或非选择性调控元件可操作性连接的转基因的表达,如通过共定位测定所测量的。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3和STXBP1中的任一个或多个,或其功能片段,或与其具有至少80%、至少85%、至少90%、至少95%或至少 99%序列同一性的序列。在一些情况下,转基因是调节基因(例如内源基因,如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3和STXBP1)的表达的DNA结合蛋白,如转录调节物、转录激活物或转录阻遏物。在一些情况下,转基因是基因编辑蛋白,如锌指核酸酶或类转录激活因子效应物核酸酶。在一些情况下,转基因是报告基因或可检测标记,如eGFP、tdTomato或RFP。在一些情况下,转基因是Cas蛋白,如Cas9。在一些情况下,如通过共定位测定所测量的,调控元件导致在PV细胞中的选择性表达,该表达的水平是CAG或EF1 α 或非选择性调控元件的表达的至少0.5倍、至少0.6倍、至少0.7倍、至少0.8倍、至少0.9倍、至少1.1倍、至少1.2倍、至少1.3倍、至少1.4倍、至少1.5倍、至少2倍、至少3倍、至少4倍、至少5倍、至少6倍、至少7倍、至少8倍、至少9倍、至少10倍、至少11倍、至少12倍、至少13倍、至少14倍、至少15倍、至少16倍、至少17倍、至少18倍、至少19倍、至少20倍、至少25倍、至少30倍、至少40倍、至少50倍、至少60倍、至少70倍、至少80倍、至少90倍、至少100倍。在一些情况下,倍数差异是指由一个或多个调控元件产生的eGFP+、PV+细胞的百分比与由非选择性调控元件产生的百分比之间的倍数差异。在一些情况下,共定位测定是免疫组织化学测定,如下面实施例5所述。在一些情况下,使用可商购获得的抗PV抗体进行共定位测定。在一些情况下,转基因编码离子通道亚基、神经递质调节因子,或其变体或功能片段。在一些情况下,离子通道亚基是钠离子通道的α亚基或β亚基,或钾离子通道的亚基。在一些情况下,转基因是以下的任一项:(i) SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或DNA结合蛋白;(ii)其功能片段;或(iii)与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,神经递质调节因子是(i) STXBP1,(ii)其功能片段,或(iii)与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,调控元件和可操作地连接的转基因位于不同的染色体上。在一些情况下,组合的调控元件的大小小于2.5kb、小于1.5kb、小于1kb或小于500bp。在一些情况下,非PV 细胞包括非PV CNS细胞类型中的任一种或多种,包括但不限于兴奋性神经元、多巴胺能神经元、星形胶质细胞、小胶质细胞或运动神经元。在一些情况下,核酸盒是线性构建体。在一些情况下,核酸盒是载体。在一些情况下,核酸盒是质粒。在一些情况下,载体为病毒载体。在

一些情况下,该病毒载体为腺伴随病毒(AAV)载体。在一些情况下,AAV载体是AAV1、AAV8、AAV9、scAAV1、scAAV8或scAAV9。在一些情况下,病毒载体为慢病毒载体。在一些情况下,调控元件包含来自GAD2、GAD1、SYN1、NKK2.1、DLX1、DLX5/6、SST、PV和/或VIP的转录起始位点的10kb以内的小于600bp的连续序列。

[0220] 在本文公开的各种实施方案中,调控元件小于2050bp、2000bp、1900bp、1800bp、1700bp、1600bp、1500bp、1400bp、1300bp、1200bp、1100bp、1000bp、900bp、800bp、700bp、600bp、500bp、400bp、300bp、200bp、100bp、90bp、80bp、70bp、60bp、50bp、40bp、30bp、20bp、10bp或5bp。在本文公开的各种实施方案中,表达盒包含比常规病毒载体(例如AAV)中的典型转基因大小更大的转基因。在一些方面,本文公开的任何实施方案的表达盒包含至少1kb、1.5kb、2kb、2.5kb、3kb、3.5kb、4kb、4.5kb、5kb、5.5kb、6kb、6.5kb、7kb、7.5kb或8kb的转基因。在一些方面,本文公开的任何实施方案包含表达盒(例如,AAV),该表达盒包含大小超过1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、2、2.1、2.2、2.3、2.4、2.5、2.6、2.7、2.8、2.9、3、3.1、3.2、3.3、3.4、3.5、3.6、3.7、3.8、3.9或4kb的转基因。在一些方面,本文公开的任何实施方案还可以包括一个或多个异源核酸序列或元件。

[0221] 一方面,治疗有需要的受试者的神经系统病症或病况的方法包括递送治疗有效量的本文所述的核酸盒。另一方面,增加转基因在PV神经元中的选择性表达的方法包括使细胞与本文所述的核酸盒接触。在一些情况下,使用本文公开的任何实施方案的方法来治疗神经系统病况或疾病,例如精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,Dravet综合征、慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,本文公开的任何实施方案的方法可用于治疗Dravet综合征。本文公开的任何实施方案的方法可用于治疗阿尔茨海默病。在一些情况下,本公开内容的方法和/或组合物可用于治疗与癫痫发作和/或癫痫相关的和/或其中牵涉PV神经元的任何神经系统病况或疾病。

[0222] 一方面,本文所述的神经系统病况通过基因疗法来治疗,该基因疗法优选导致相对于另一组织类型或细胞类型(例如PV神经元),在一种组织类型或细胞类型中优先表达的基因疗法,如通过共定位测定所确定的。在一些情况下,基因疗法是AAV。

[0223] 一方面,将任何转基因的表达靶向CNS中的PV神经元的方法包括将一个或多个PV神经元选择性调控元件与转基因可操作地连接。在一些情况下,调控元件包含SEQ ID NO:1-32的一个或多个序列,或与SEQ ID NO:1-32具有至少80%序列同一性的序列,或其功能片段。在一些情况下,如通过共定位测定所测量的,调控元件导致在PV神经元的选择性表达,该表达的水平是与转基因可操作连接的CAG或EF1 α 或非选择性调控元件的至少2倍、至少5倍、至少7倍或至少10倍。在一些情况下,转基因是SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1、DNA结合蛋白或其功能片段。在一些情况下,调控元件和转基因在AAV中。在一些情况下,AAV是AAV9。

[0224] 一方面,治疗有需要的受试者的神经系统病况或病症的方法包括使细胞与核酸盒接触,该核酸盒包含与转基因可操作地连接的一个或多个调控元件,该调控元件导致相对于在一种或多种非PV CNS细胞中,在PV神经元中的选择性表达。在一些情况下,该调控元件包含SEQ ID NO:1-32中的任一个或多个,或其功能片段或组合,或与其具有至少80%、至少

85%、至少90%、至少95%或至少99%序列同一性的序列。在一些情况下,转基因是电压门控离子通道亚基,或其变体或功能片段。在一些情况下,亚基是钠离子通道的 β 亚基。在一些情况下,亚基是钠离子通道的 α 亚基。在一些情况下,亚基是钾离子通道的亚基。在一些情况下,转基因是下列任一项:(i) SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、DNA结合蛋白或STXBP1;(ii) 其功能片段;或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,神经系统病况或病症与SCN1A、SCN1B、SCN2B、KV3.1和KV3.3中任一个的单倍性不足或突变有关。在一些情况下,神经系统病况或病症是Dravet综合征。在一些情况下,神经系统病况或病症是阿尔茨海默病。在一些情况下,神经系统病况或疾病是精神障碍(例如,精神分裂症、强迫症、成瘾、抑郁、焦虑、精神病);自闭症谱系障碍(例如,脆性X综合征、雷特综合征);癫痫(例如,慢性创伤性脑病、全面性癫痫伴热性惊厥附加症(GEFS+)、癫痫性脑病、颞叶癫痫、局灶性癫痫、结节性硬化);或神经变性(例如,阿尔茨海默病、帕金森病)。在一些情况下,神经系统病况或疾病是牵涉PV神经元的任何癫痫发作和/或癫痫相关的病况或疾病。在一些情况下,如通过共定位测定所测量的,核酸盒导致在PV神经元的选择性表达,该表达的水平是与转基因可操作连接的CAG或EF1 α 或非选择性调控元件的至少2倍、至少5倍,或至少7倍或至少10倍。在一些情况下,核酸盒在AAV中。在一些情况下,AAV是AAV9。

[0225] 一方面,治疗Dravet综合征的方法包括使细胞与AAV接触,该AAV包含是以下任一项的转基因:(i) SCN1A、SCN1B、SCN2B、DNA结合蛋白;(ii) 其功能片段;或(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,AAV还包含与转基因可操作地连接的一个或多个PV神经元选择性调控元件或本文公开的任何调控元件。在一些情况下,每个调控元件独立地包含一段序列,该序列包括SEQ ID NO:1-32中的任一个,或其任何功能片段或组合,或与SEQ ID NO:1-32中的任一个具有至少80%序列同一性的序列。

[0226] 另一方面,治疗阿尔茨海默病的方法包括使细胞与AAV接触,该AAV包含是以下任一项的转基因:(i) SCN1A、SCN2B、KV3.1、KV3.3和STXBP1,(ii) 其功能片段,和(iii) 与(i)或(ii)具有至少80%序列同一性的序列。在一些情况下,AAV还包含与转基因可操作地连接的一个或多个PV神经元选择性调控元件。在一些情况下,每个调控元件独立地包含一段序列,该序列包括SEQ ID NO:1-32中的任一个,或其任何功能片段或组合,或与SEQ ID NO:1-32中的任一个具有至少80%序列同一性的序列。

[0227] 实施例

[0228] 提供这些实施例仅出于说明的目的,并非限制本文提供的权利要求书的范围。

[0229] 实施例1

[0230] 鉴定推定的PV选择性调控元件

[0231] 为了鉴定和筛选对PV细胞有选择性的推定的调控元件,可以使用亲和纯化,例如使用抗GFP或抗Myc抗体和蛋白G包被的磁珠,从R26-CAG-LSL-Sun1-sfGFP-Myc基因敲入小鼠获得PV细胞。可通过使用抗PV抗体包被的珠或亲和纯化基质来富集PV细胞。然后从PV细胞中分离出细胞核。核RNA可从细胞核中纯化并转化为cDNA,然后用Nugen Ovation RNA-seq系统V2(Nugen 7102)进行扩增,随后使用Illumina HISEQ® 2500进行测序。基因组DNA可从细胞核中纯化、片段化,并用于制备methy1C-seq文库,该文库可使用Illumina HISEQ® 2000进行测序。为了产生ATAC-seq文库,使用Tn5转座酶(Illumina FC-121-

1030) 将与珠结合的细胞核转座。在9-12个PCR扩增循环后, 使用 Illumina HISEQ® 2500 对文库进行测序。为了产生ChIP-seq文库, 使用微球菌核酸酶将PV细胞的细胞核消化成单核小体, 然后用盐提取染色质, 并构建天然ChIP和文库, 它们可以在 Illumina Hiseq 2500 上进行测序。在对这些文库进行测序后, 序列被映射以鉴定富含CG的区域中低甲基化的相关性和模式、组蛋白修饰、转录因子结合位点, 以及与PV细胞中高表达的转录因子相关的模式。来自上述多个测定和/或文库的重叠特征和相关性为鉴定出是推定的PV选择性调控元件的候选序列提供了收敛的证据。推定的PV选择性调控元件可以使用如下面的实施例5所述的共定位测定进行进一步测试。推定的PV选择性调控元件也可以在B6 PV-Cre小鼠 (Jackson Laboratory) 中测试, 该小鼠是在小清蛋白表达细胞中表达Cre重组酶的B6 PV-Cre基因敲入小鼠, 如下面的实施例2所述。在验证调控元件的PV选择性之后, 可将调控元件与转基因可操作地连接, 以使表达选择性地靶向PV细胞, 而非至少一种、两种、三种、四种、五种或超过五种非PV细胞。

[0232] 实施例2

[0233] 对PV-Cre小鼠的PV神经元的选择性

[0234] 使用荧光成像可以确定对PV神经元的选择性。含有与(i)对照启动子(EF1a); 或(ii)在上述实施例1鉴定的PV选择性RE; 或(iii)选自SEQ ID NO:1-32的PV选择性RE可操作地连接的eGFP的AAV9载体; 以及含有Cre依赖性tdTomato的AAV9载体被共同注射到B6 PV-Cre小鼠 (Jackson Labs) 中。PV-Cre是一种基因敲入小鼠, 其在小清蛋白表达神经元(如脑内的中间神经元和背根神经节中的本体感受传入感觉神经元) 中表达Cre重组酶, 而不破坏内源性小清蛋白的表达。

[0235] 以 $0.3\mu\text{L}/\text{min}$ 的速率将 $1.5\mu\text{L}$ 的AAV9载体(5^{12} 至 1^{13} 个gc/ml) 双侧输注到小鼠的背侧和腹侧海马体中, 注射后休息4min。将小鼠麻醉以进行注射。将动物放置在立体定向框架(Kopf instruments, USA) 中, 使用以下坐标, 背侧海马体(AP-2.0mm, 侧向 ± 1.5 , DV-距硬脑膜1.4mm) 和腹侧海马体(AP-3.1mm, 侧向 ± 2.8 , DV-距硬脑膜3.8mm)。Hamilton注射器(型号80308; 10 μL 注射器和相应的30ga钝头针) 可与立体定向显微操作器一起使用, 用于指定和钻出钻孔。该钻头仅用于穿透骨头。钻孔后, 将输液管下放到大脑中注射所需位置的深度, 例如, 注射量: $1.5\mu\text{L}$; 注射速率: $0.3\mu\text{L}/\text{min}$ 。输注前, 让针平衡1分钟。一旦完成递送, 针留置4分钟, 然后经约1分钟取出。一旦所有的输注完成, 皮肤切口就用缝合线封闭, 并在施用术后止痛药。经治疗的小鼠在余下的研究中每天经历健康检查, 并每周称重一次以监测体重。

[0236] 为了收集组织, 通过过量的异氟醚对小鼠实施安乐死, 并灌注4%多聚甲醛(PFA)。提取一块含有海马体的脑组织, 并在4°C的4%PFA中放置至少12小时。然后将脑组织在4°C的30%蔗糖(在磷酸盐缓冲盐水中)中脱水, 直到组织沉入管底。将脑组织嵌入组织-Tek OCT中, 用于在低温恒温器中进行切片。使用标准的免疫组织化学程序, 用抗RFP多克隆兔抗体(Rockland Antibodies and Assay) 和抗eGFP多克隆鸡抗体(Aves Labs) 对切片的脑组织进行eGFP和tdTomato染色。使用荧光显微镜成像将细胞可视化。eGFP或绿色荧光对应于所有的基因表达。来自tdTomato的红色荧光对应于PV+细胞。可作为黄色或白色细胞可视化的这两种荧光信号的重叠代表表达eGFP转基因的PV+细胞。与对照启动子(EF1a)相比, 包含PV选择性调控元件的AAV9载体预期产生更多的eGFP+且PV+的细胞。例如, 对来自注射包含任一PV选择性RE(例如, 实施例1中鉴定的SEQ ID NO:1-32或推定的RE)的AAV9的小鼠的细

胞的荧光成像预期显示出更多同样为PV+的eGFP+细胞。对PV细胞的选择性可以定量为所有eGFP+细胞中同样是PV+的百分比。

[0237] 实施例3

[0238] 在Dravet小鼠模型中减少癫痫发作

[0239] B6 (Cg) - Scn1a^{tm1.1Dsf} /J 小鼠通过Jackson Laboratories从Dravet综合征欧洲联盟获得。这些小鼠在SCN1A的外显子24中含有Dravet综合征相关的突变(1783位的A变成V)。该小鼠还含有具有野生型序列的 floxed 外显子24。当不被处理时,该小鼠品系表达SCN1A的WT等位基因的两个拷贝。然而,在递送表达Cre重组酶的AAV后,AAV所靶向的任何细胞都将转变为表达突变体等位基因的一个拷贝。在突变体SCN1A亚基表达后,小鼠在10天内发展出自发性癫痫发作。

[0240] 如实施例2所述,B6 (Cg) - Scn1a^{tm1.1Dsf} /J 小鼠和对照C57B1/6小鼠被注射在EF1 α 启动子控制下表达CRE重组酶的AAV和包含驱动eGFP (SEQ ID NO:36) 或SCN1B (SEQ ID NO:37) 表达的PV细胞选择性调控元件SEQ ID NO:32的AAV。一旦完成所有四次输注,立即进行遥测植入(F20-EET, Data Sciences International)。从术后10天起,连续监测脑皮层电图数据持续14天。分析脑皮层电图数据,并记录所有癫痫发作事件,标注日期、开始时间、停止时间、癫痫发作持续时间和严重程度评分。图1示出了在治疗后14天内,12小时窗口内癫痫发作的频率。与对照动物相比,用SCN1B治疗的小鼠表现出癫痫发作频率较低的趋势。

[0241] 这一观察结果与以下观点一致:钠离子通道的 β 单位例如SCN1B可以有助于钠离子通道的运输和组装,并且选择性地增加PV神经元中 β 单位的表达可以导致Nav1.1通道的运输和组装的增加,从而导致用SCN1B基因疗法治疗的小鼠的癫痫发作频率和持续时间降低的趋势。

[0242] 实施例4

[0243] 在小鼠模型中治疗阿尔茨海默病

[0244] 研究中使用在**PSYCHOGENICS®**饲养的雌性APP/PS1小鼠和WT小鼠。APP/PS1小鼠含有携带瑞典突变(670G-T和671A-C)的淀粉样 β 前体蛋白(APP)和带有L166P突变的早老蛋白1(PSEN1)二者的人类转基因,二者均受Thy1启动子的控制。这些小鼠发展出阿尔茨海默病的症状,包括淀粉样斑块和记忆缺陷。关于这些小鼠的进一步描述可参见Radde等人,2006(Radde,Rebecca等人“ $A\beta$ 42-driven cerebral amyloidosis in transgenic mice reveals early and robust pathology.”EMBO reports 7.9(2006):940-946)。

[0245] 将APP/PS1小鼠用作模型来确定用RE控制下的SCN1B进行治疗对阿尔茨海默病症状的影响。APP/PS1小鼠和非转基因对照被注射表达eGFP的对照载体或表达SCN1B的治疗载体,二者均受SEQ ID NO:32的控制;并如实施例3所述地植入EET发射机。在术后4周,经2-4小时评估大脑活动。自动分析脑皮层电图数据,并比较不同频带的功率水平。图2示出非转基因对照(WT)小鼠、APP/PS1小鼠和用SCN1B治疗的APP/PS1小鼠中的高 γ 功率(50-100Hz)。在阿尔茨海默病患者和癫痫患者中,高 γ 功率活动的增加与癫痫发作有关。与对照小鼠相比,APP/PS1小鼠显示出更高水平的高 γ 功率活动。然而,在经治疗的小鼠中没有该增加,表明载体治疗有效。

[0246] 实施例5

[0247] 对C57BL/6J(WT)小鼠的PV神经元的选择性

[0248] 使用免疫组织化学法测试了本文公开的各种RE在PV神经元中的选择性基因表达的选择性。将C57BL/6J (WT) 小鼠系用于PV免疫组织化学测定。表达盒包括在AAV9构建体中与调控元件 (SEQ ID NO:1或SEQ ID NO:8) 或CAG启动子可操作地连接的报告基因转基因eGFP。

[0249] 幼体全身输注:使用带有31G针的300U胰岛素注射器,通过面部静脉注射向出生1天的C57BL/6J小鼠输注AA9载体 ($1E^{12}$ 至 $3E^{12}$)。为了收集组织,在输注后21天,通过过量的戊巴比妥钠(i.p.)将小鼠安乐死,并灌注肝素化(2.5IU/ml)生理盐水,然后灌注4%甲醛。取出大脑,然后在4摄氏度下在4%甲醛中浸泡固定24-48小时。然后将大脑放入含30%蔗糖的PBS中,让其在4摄氏度下下沉(约2-3天)。下沉后,将单个大脑半球冷冻在TISSUE-TEK® OCT中,中线朝下。将冷冻的大脑在低温恒温器上进行矢状切片,然后自由漂浮地置于PBS中。使用标准的免疫组织化学程序,用鸡抗GFP (Aves Lab, GFP-1020) 和小鼠抗PV (Sigma, P3088) 对切片进行eGFP和小清蛋白(PV)染色。

[0250] 成年体全身输注:通过尾静脉注射向4周龄C57BL/6小鼠输注60μL表达eGFP的AAV9载体 (4.9^{13} 至 1^{14} 个gc/ml)。为了收集组织,在输注后21天,通过过量的异氟醚将小鼠安乐死,然后提取整个大脑,用PBS清洗并放入含有冰冷的4%甲醛的单独的5ml管中。将组织在4摄氏度下固定过夜。第二天,将大脑放入含30%蔗糖的PBS中,让其在4摄氏度下下沉。下沉后,将单个大脑半球冷冻在Tissue-Tek OCT中,中线朝下。将冷冻的大脑在低温恒温器上进行矢状切片,然后自由漂浮地置于PBS中。使用标准的免疫组织化学程序,用鸡抗GFP (Aves Lab, GFP-1020) 和小鼠抗PV (Sigma, P3088) 对切片进行eGFP和小清蛋白(PV)染色。

[0251] 免疫组织化学方案:使用抗PV抗体采用免疫组织化学法分析eGFP信号与PV信号的共定位,其中信号的叠加在上图(合并)中表现为白色或浅灰色斑点,其中有代表性的叠加用箭头表示。eGFP和PV荧光的叠加表明在PV细胞中的表达。这些实验可用于确定在PV细胞中表达的选择性。为进行免疫组织化学实验,在室温下,用封闭缓冲溶液(在1X PBS中包含3%BSA、3%NGS、0.3%Triton X-100、0.2%Tween-20)将从每只小鼠获得的组织封闭1小时。然后在4C下,将组织与初级抗体在封闭缓冲液中温育过夜,用1mL 1X PBS洗涤三次,每次间隔5分钟。然后在室温下,将组织与二级抗体在封闭缓冲液中温育1小时,随后用1mL 1X PBS洗涤三次,每次间隔5分钟。将组织与DAPI (1:1000) 在PBS缓冲液中温育5分钟,并用1mL 1X PBS洗涤两次。将组织固定在载玻片上,用荧光显微镜成像和分析。使用Vectra 3成像系统(Perkin Elmer)拍摄图像,并使用inform-Tissue finder、高级图像分析软件或手工评分对eGFP和PV染色的共同标记进行定量。在确定共定位的百分比之前,在每个图中计数至少80个GFP阳性细胞。

[0252] 图3A-3B示出了在全身AAV9注射后,在幼体中进行的免疫组织化学实验的结果。图4A-4B示出了在AAV9注射后,在成年小鼠中进行的类似免疫组织化学实验的结果。

[0253] 图3A示出了在全身AAV9注射后,在幼体中进行的免疫组织化学实验的叠加。图3B示出了免疫组织化学实验的共定位的定量,其中对PV细胞的选择性被测量为与在CAG启动子控制下的eGFP表达相比,同时是PV+的GFP+细胞的百分比。

[0254] 图4A示出了在全身AAV9注射后,在成年小鼠中进行的免疫组织化学实验的叠加。图4B示出了免疫组织化学实验的共定位的定量,其中对PV细胞的选择性被测量为与在EF1α控制下的eGFP表达相比,同时是PV+的GFP+细胞的百分比。

[0255] 据估计,GABA能神经元组成约20%的CNS,而PV细胞组成约40%的GABA能神经元,这意味着PV细胞大约占CNS中所有神经元的8%。参见Pelkey,Ka等人,2017和Lee,S等人,2010。因此,可以预测,由非选择性调控元件(例如CAG、EF1 α 或组成型启动子)标记的细胞中约有8%是PV阳性的,或在这个范围内。因此,高于8%的PV细胞中的表达表明在PV细胞中的选择性增加。值得注意的是,包含调控元件SEQ ID NO:8的AAV9注射导致约60%的细胞呈PV阳性,比PV细胞分布的预期值高7.5倍。

[0256] 进行了与上文所述类似的免疫组织化学实验以确定与具有SEQ ID NO:34序列的非选择性调控元件相比,其他调控元件即SEQ ID NO:2-7和9-22的选择性表达,与上文实验的差别在于使用AAVDJ病毒载体将与调控元件可操作地连接的eGFP递送到C57BL/6J(WT)小鼠中。这种AAVDJ病毒直接注射到成年小鼠的海马体中的CNS中。在计算共定位的百分比或选择性(即GFP阳性细胞同时是PV阳性的百分比)之前,在每个实验中计数至少80个GFP阳性细胞。图5A-5F示出了用于确定共定位或选择性(测量为eGFP阳性细胞同时是PV阳性的百分比)的荧光成像,以及与非选择性调控元件SEQ ID NO:34的信号的比较。对标记呈阳性的细胞在图像中显示为白色/灰色细胞。合并图像示出了对应的eGFP和抗PV图像之间的重叠。对eGFP和PV二者都呈阳性的细胞在合并图像中显示为白色/浅灰色细胞。图6示出了共定位分析的定量,测量为eGFP+细胞同时是PV+的百分比。

[0257] 实施例6

[0258] 在不同小鼠系中治疗Dravet综合征

[0259] 使用本文所述的表达盒对Dravet综合征和/或其症状的治疗在各种小鼠系中测试,如上文所述的B6(Cg)-Scn1a^{t_{m1.1Dsf}}/J、Scn1a^{t_{m1Kea}}和Scn1a-R1470X小鼠系。这些小鼠系是已确立的Dravet综合征的小鼠模型。Scn1a^{t_{m1Kea}}和Scn1a-R1470X小鼠系不需要CRE重组酶。

[0260] Scn1a^{t_{m1Kea}}小鼠(可从Jackson Laboratory获得;在Hawkins等人,Scientific Reports,第7卷:15327(2017)中描述)具有SCN1A的第一编码外显子的缺失。SCN1A敲除等位基因纯合的小鼠的特征为震颤、运动失调、癫痫发作,并在出生后第16天死亡。C57BL/6背景下的杂合小鼠发展出自发性癫痫发作,并在数周内死亡。这样的小鼠品系可用于研究癫痫和Dravet综合征治疗的安全性和功效。其他信息请参见Miller等人,Genes Brain Behav.2014年2月;13(2):163-72。

[0261] Scn1a-R1470X小鼠是基因敲入小鼠,携带在SCN1A基因的外显子21中的提前终止密码子R1407X。在三个无关的SMEI患者中,同一突变已被鉴定为致病性突变。Scn1a^{RX/RX}幼体的特征是在出生后12天内出现反复的自发性癫痫发作,包括出生后12-16天的强直阵挛性癫痫发作和阵挛性癫痫发作,以及有节奏的抽动和不自主的肌肉收缩。其他信息请参见Ogiwara等人,Journal of Neuroscience,2007年5月30日,27(22)5903-5914。

[0262] 为了测试本文所述的组合物,如AAV基因疗法和使用该基因疗法的治疗,向上述每个小鼠品系的Dravet小鼠和对照小鼠(例如,该品系的野生型小鼠或未经治疗的Dravet小鼠)注射(例如,通过腹膜内注射施用)表达eGFP或另一报告基因的AAV,或包含与本文公开的转基因(如SCN1A、SCN1B或SCN2B,或SEQ ID NO:37-39中的任一个,或其变体或功能片段)可操作地连接的本文所述的一个或多个PV选择性RE(例如,SEQ ID NO:1-32)的表达盒。注射AAV后,随时间监测小鼠的存活率。每天监测所有小鼠的总体健康状况(例如体重、补水、理毛行为和活动能力),并记录死亡。可在AAV注射后立即进行遥测植入(F20-EET,Data

Sciences International)。从术后10天起,连续记录和监测脑皮层电图数据至少14天。所有癫痫发作事件可在AAV治疗后至少14天内记录,标注日期、开始时间、停止时间、持续时间和严重程度评分。与eGFP对照或未治疗对照相比,用上述AAV治疗后癫痫发作的频率和/或持续时间的减少表明基因疗法在减轻Dravet综合征的症状和/或严重程度方面的功效。

[0263] 在用AAV治疗小鼠后,可使用各种PCR和/或测序方法随时间监测转基因(例如SCN1A;SCN1B;SCN2B;DNA结合蛋白,如调节内源性SCN1A、SCN1B或SCN2B的转录激活物;SEQ ID NO:37-39中的任一个;或其任何变体或功能片段)的表达水平,以显示AAV治疗可导致PV细胞中基因表达的增加。还可使用RNA印迹分析和原位杂交来分析体内的转基因表达。还可在治疗后监测从该蛋白质表达的蛋白质的水平,以显示转基因表达的增加与体内相应蛋白质的增加相关。蛋白质水平可以用各种方法来测定,包括但不限于蛋白质印迹分析、免疫组织化学、荧光免疫组织化学和/或ELISA分析。还可使用电流钳分析法来测定功能性电压门控钠离子通道的形成。

[0264] 可评估高热诱导的癫痫发作,以将野生型小鼠和/或未经治疗的Dravet小鼠与经AAV基因疗法治疗的Dravet小鼠进行比较,该AAV基因疗法包含本文所述的表达盒(例如,包含与本公开内容的转基因可操作地连接的本公开内容的一个或多个RE的表达盒,该转基因是例如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、其功能片段,或调节内源性SCN1A、SCN2A、SCN8A、SCN1B或SCN2B的DNA结合蛋白)。在这些实验中,核心体温用RET-3直肠温度探头(Physitemp Instruments, Inc, New Jersey, USA)监测,并由连接到重新配置有Partlow 1160+控制器(West Control Solutions, Brighton, UK)的啮齿动物体温调节器(TCAT-2DF, Physitemp Instruments, Inc.)的加热灯来控制。每两分钟将体温升高0.5°C,直到第一次阵挛性抽搐发作。与未经治疗的Dravet小鼠相比,用AAV基因疗法治疗的Dravet小鼠在首次阵挛性抽搐发作前具有更高的阈值体温,并且/或者在最高测试温度下保持无癫痫发作的比例更高。

[0265] 可向小鼠施用不同剂量的包含表达盒的AAV以确定每种基因疗法治疗的安全性和功效特征。这些临床前研究还可以为用于治疗Dravet综合征的基因疗法提供最佳剂量。

[0266] 实施例7

[0267] 在小鼠中治疗阿尔茨海默病

[0268] 使用在**PSYCHOGENICS®**饲养并且是确立的阿尔茨海默病小鼠模型的雌性APP/PS1小鼠和野生型(WT)小鼠来研究包含一种或多种PV选择性RE的本文所述的组合物在治疗阿尔茨海默病中的安全性和功效。APP/PS1小鼠如上文实施例4所述。

[0269] 向APP/PS1小鼠和非转基因对照注射表达eGFP的对照AAV载体或包含与转基因可操作地连接的本文公开的一个或多个PV选择性RE(例如SEQ ID NO:1-32)的治疗AAV载体,该转基因是阿尔茨海默病中缺陷或受损的,如SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3、STXBP1、调节内源基因(例如,SCN1A、SCN2A、SCN8A、SCN1B、SCN2B、KV3.1、KV3.3或STXBP1)的DNA结合蛋白,或SEQ ID NO:37-43中的任一个,或其功能片段。

[0270] 注射AAV后,随时间监测小鼠的存活率。每天监测所有小鼠的总体健康状况(例如体重、补水、理毛行为和活动能力),并记录死亡。在注射AAV后,小鼠也被植入EET发射机,如上文实施例3所述。在术后至少4周内,经24小时记录和监测脑活动。自动分析脑皮层电图数据,并且可比较以下不同组之间在不同频带(50-100Hz)中的功率水平:WT小鼠、未经治疗的APP/PS1小鼠和经AAV治疗的APP/PS1小鼠,每次治疗均采用上文所述的AAV基因疗法。在阿

尔茨海默病患者和癫痫患者中,高 γ 功率活动的增加与癫痫发作有关。因此,未经治疗的APP/PS1小鼠预期会显示出比对照小鼠更高水平的高 γ 功率活动,而这种增加预期在经治疗的小鼠中不存在或减少,表明AAV基因疗法的治疗有效。

[0271] 在用AAV治疗小鼠后,可使用各种PCR和/或测序方法随时间监测转基因的表达水平,以显示AAV治疗可导致转基因的内源性表达的增加。还可使用RNA印迹分析和原位杂交来分析体内的基因表达。还可在治疗后监测从该转基因表达的蛋白质的水平,以显示基因表达的增加与蛋白质表达水平的增加相关。蛋白质水平可以用各种方法来测定,包括但不限于蛋白质印迹分析、免疫组织化学和/或ELISA分析。还可使用电流钳分析法来测定功能性电压门控钠或钾离子通道的形成。

[0272] 可向小鼠施用不同剂量的包含表达盒的AAV以确定每种基因疗法治疗的安全性和功效特征。这些临床前研究还可以为用于治疗阿尔茨海默病的基因疗法提供最佳剂量。

- [0001] 序列表
[0002] <110> 编码治疗公司
[0003] <120> 组织选择性转基因表达
[0004] <130> 46482-704.601
[0005] <140>
[0006] <141>
[0007] <150> 62/480,998
[0008] <151> 2017-04-03
[0009] <160> 43
[0010] <170> PatentIn版本3.5
[0011] <210> 1
[0012] <211> 681
[0013] <212> DNA
[0014] <213> 智人
[0015] <400> 1
[0016] ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
[0017] attttgaagt tctgttgaca taaagaatca tgatattaaat gcccatggaa atgaaaggc 120
[0018] gatcaacact atggttgaa aaggggaaa ttgttagagca cagatgtgtt cgtgtggcag 180
[0019] tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaattctg attggcccc 240
[0020] gtgtgagccc agatgaggtt cagctgccaa ctttctttt cacatcttat gaaagtctt 300
[0021] taagcacaac taactttttt tttttttttt tttttttttag acagagtctt gctctgttgc 360
[0022] ccaggacaga gtgcagtagt gactcaatct cggttcactg cagcctccac ctcctaggct 420
[0023] caaacggtcc tcctgcatca gcctcccaag tagctggaa tacaggagtg gccccaccatg 480
[0024] cccagctaattttt ttaatagata cgggggttcc accatatcac ccaggctgg 540
[0025] ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600
[0026] ggcgtcagcc actatgccca acccgaccaa ctttttttaa aataaatatt taaaaaatgg 660
[0027] gtatttcaca tatatactag t 681
[0028] <210> 2
[0029] <211> 502
[0030] <212> DNA
[0031] <213> 小家鼠
[0032] <400> 2
[0033] agtttggaca agaactatag ttctagctt ctctgggtct ccacccgtca gagaatgcag 60
[0034] ctttcattat ctcatgagcc aaactctcat catcttttc catatatctg tcgggtct 120
[0035] tccatgagta ctctaaccaca cacagaagga gcacttacac aggctgttgtt ttttctttta 180
[0036] ttatcatagc tgggttcag acatgtgcat tctgttcttgc ttgcttcaat gctaaaggag 240
[0037] tctcaggata tgagaactgtt accagccgag gcatcaggaa acatgggtgg aaattccac 300
[0038] agtactattt gttcactgtg tgaccttggg ccagtcacat cccttcctg aggcttcgt 360
[0039] tcggccaaatgtt ataaaagaag catcttttaa ctttttttaa ggtcatgagt caggccac 420
[0040] acactctcag ggagactcat gagagtacag atcatttccc atagaaaaac catagttta 480
[0041] tatccagagg ctttctgtta ag 502

[0042] <210> 3
 [0043] <211> 823
 [0044] <212> DNA
 [0045] <213> 小家鼠
 [0046] <400> 3
 [0047] gttccagtt cagaggcaga gcatttgggg ttcccagtca ggagcttcc tctctccgct 60
 [0048] ccttagtttc ctctcttaa aaaaaaatgg gtatgtat agaaaggaag ctctgggctc 120
 [0049] ggggaccagg gccctggat ccccgctccc agccactcgc tcctgaccct tccagggaca 180
 [0050] agctcccccc caccccgatcc tttccaggct gccactagaa gagatgggaa cgcgtggta 240
 [0051] gccgcttctg tcgccccca gggAACGGTC tcacgctgga gggggcagtgc ccctcgaaac 300
 [0052] aggacagtca gcccaagcca gccaaAGCGCG cgccggacgtc cttcaccgca gagcaattgc 360
 [0053] aggtaccccg ggcaagcccc gaagcgtgtg ggccgggctt cggagtggc gtggttgttc 420
 [0054] gggacttgtg actccgcccc ttgtgcgggg acccgctga ggccgctcca aggtatgaagc 480
 [0055] tgccctgggc gtggcctcg accctgagcc tctgattggg cggaggtctc agggcccttc 540
 [0056] tgcgcacccac aggttatgca ggccgcgttc gcgcaggaca acaaccggaa cgccgcagacg 600
 [0057] ctgcagaagc tggcggacat gacgggcctc agtcgcagggt tcattccaggt gggcgtccgg 660
 [0058] ggtctcgcc ttcaggctta gggtaaacct tagggaaagcg ctgaagctcg tagtggtacg 720
 [0059] gatggtcgacgtc cgtgcacgtg gccgccttc tccagtggtt cctaaggacc ccagtcggca 780
 [0060] cgggttggacc ctttcccttg attactgaga gtgcagaggc tgt 823
 [0061] <210> 4
 [0062] <211> 633
 [0063] <212> DNA
 [0064] <213> 小家鼠
 [0065] <400> 4
 [0066] tggtgggaag acatgtccag ggaagaaaatg gcctccagag gcctgagggtg gggaaatgct 60
 [0067] ggaggtggag agaggaacaa ctgactgaaa atgagcttcc actgtggctt agtagcctat 120
 [0068] accaagtcta gaggataggg taggagaaga ttaggaaagc gatgggtctg agaatgatgt 180
 [0069] ggcctgtga cttttgtaaa cccaaagcac cttggactaa accctatgaa cagtgtgg 240
 [0070] ccaccaaaga ctataatgag ctcaggaaac agaattctgt gtcatgggtt atttttttt 300
 [0071] ttttttctg ctaactgcag tctgggtat gcattgacca accaatcctg gaaagtaaga 360
 [0072] ggcaaggcga gctggacgg tgaggaggc ctgatggaa ccaggccaa cagggcagca 420
 [0073] gaggcgtatga agaggatgtg gtcatcccg agactcaattt cattagctgg aggcactgct 480
 [0074] ggatagggtc tgaaggctt ggtatctgat ttggcggctt gggtgagtgg tggctctgct 540
 [0075] tcctgaacag tgtgtcaag agggaaacagg gttaaaggctt aggacagtca caggtgagtc 600
 [0076] agcctcacaa gagcaacacctt cccctgtgc aga 633
 [0077] <210> 5
 [0078] <211> 670
 [0079] <212> DNA
 [0080] <213> 小家鼠
 [0081] <400> 5
 [0082] ggaggtctcc ttttgcggc gttccaacaa gagaatgcaaa ggctgtatct caatttcctt 60
 [0083] ggcctctct gtattataga agaaaagtag ggaagccata cgccttcctt gagcttcagt 120

[0084] gtctcttgt ctctgaaat gaggctgggg aggctgggg cggcggtgaa agaggccgc 180
 [0085] gccaagccga cccccaccc tc(ccc)ctcc ccaggtcaac aacctcatct ggcacgtgcg 240
 [0086] gtgcctcgag tgctccgtgt gtgcacatc gctgaggcag cagaatagct gctacatcaa 300
 [0087] gaacaaggag atctactgca agatggacta cttcaggtag gcagcggcca tcccggcagc 360
 [0088] aagcgctgga gcatgaacgc ctgcacacg cgtgcctagg ccacttgcgt ggcctgtgct 420
 [0089] ctccaattcc tgagccctgc tggtcagagt gcacaacgcg gtcagcgca ctggccggc 480
 [0090] ctcctactc agcacgtttt acacagaagg gagcgccagt ctcagcctga gttctggcgg 540
 [0091] gggatctgcc tcgggttcct ccgatctgac aggcgctggc cacgggtctg gttccatctc 600
 [0092] tggcttttc tggccccgag caccagtgtt ttctgttgag ctctgatgtc cgaggctctg 660
 [0093] gccggatca 670
 [0094] <210> 6
 [0095] <211> 1043
 [0096] <212> DNA
 [0097] <213> 小家鼠
 [0098] <400> 6
 [0099] ctctggctac ctcttatctt gggcattcac gacaatttct aattgcaggt agttgtgtg 60
 [0100] tgtgcgcgtg tttttttcc ccctcagagg ctggattgc aaaggaacta agcgattact 120
 [0101] tcaagagcca cgggttaagt gcagggagag gggagagag agggaaaaaaa acccaatcca 180
 [0102] aattcaaatt gcttcatttag agagacaccg ctgggtggg gaaggcttt aaatgcccac 240
 [0103] tacaaagtta ggactcattt ttcagcggcgtt gtttatataa caggcgaggg gaggcgctgg 300
 [0104] gctctgacag ctccgagcca gttcagcagc cgccgtcgcc tgcattccct cccctcccc 360
 [0105] caggtgatgg cccagccagg gtccggctgc aaagcgacca cccgctgtct cgaagggacc 420
 [0106] gctccgcctg ccatggtgag tccttcggc cctgcatttcg gccccgagtc ccccaacag 480
 [0107] cacaggccag ggcttctggc tcagcattcc ggctaccaac ctctaccct gcgctggaaa 540
 [0108] actgcccata ggagccgcct ctgcgttgcg cttgggtttt ctggcctgga atgtgagct 600
 [0109] tggctgcttc ctgcacccag gatgcgtgt gttaaaagtt gggggccgtc ctttcttc 660
 [0110] caatagggtcc tttcattctt gtactccagc cttagggcgcc acatccctgg cacatttcgg 720
 [0111] tgtcagtcgg tgcgcgagga aaccagattc aactctgagt actcggctaa ggcgttcgt 780
 [0112] gttcctctct cccatttcag gctcagtcag acgcagaggc ctggcaggc gctctggaca 840
 [0113] aggacgaagg tagagccctcc ccatgtacgc ccagcacacc gtctgtctgc tcggccccc 900
 [0114] ctgctgcctc ttccgtgccc tctggcggca agaatatctg ctccagttgc ggtctggaga 960
 [0115] tcctggaccg gtatctgctc aaggtaggtt gggtaggtt tgctgtctg cccacgggtg 1020
 [0116] tggtttgcag ccccaagagc tgt 1043
 [0117] <210> 7
 [0118] <211> 507
 [0119] <212> DNA
 [0120] <213> 小家鼠
 [0121] <400> 7
 [0122] caagactttt aaaagtttag ataaataaac aaacatttgc cggcttcca tcacatctag 60
 [0123] actataatcc aaagatctat atggcccaa acgacttaca cttactacc gtctccata 120
 [0124] tggcttcttc ccccatcagt cattgtcctc agccatagtg gcctccctgt tcctttgggt 180
 [0125] acaaggaaac aactccctga gaggtccat tagctgctgt tgctgtgat gctcttgagc 240

[0126] ccacaccatc tgctcatttc tctcctcactg tgtcagtgtat taagaggctg tccttggcct 300
 [0127] cccgtcaaaa ttacatccct gccgcttcc acttcttgcc ttcttatccc ctaaatagaa 360
 [0128] ctaactcacc actaccacac attctatata attggatatac tgtcctctgt ttaaatataa 420
 [0129] ttttgacttc aagaaagaac gttgtcactg ccctgtcacc agactttaa acagtgccta 480
 [0130] tcgtgtggca catgctcagt gaaattg 507
 [0131] <210> 8
 [0132] <211> 502
 [0133] <212> DNA
 [0134] <213> 小家鼠
 [0135] <400> 8
 [0136] tcaacagggg gacacttggg aaagaaggat ggggacagag ccgagaggac tgttacacat 60
 [0137] tagagaaaca tcagtactg tgccagctt gggtagact gcacaaaagc cctgaggcag 120
 [0138] cacaggcagg atccagtcgt ctggtcccg gaagctaacc gtctcagaca ggcacaaag 180
 [0139] caccgagaca tgtgccacaa ggcttgtta gagaggtcag aggacagcgt acaggtcccc 240
 [0140] gagatcaaacc tcaacctcac caggcttggc agcaaggctt taccacccca cccccacccc 300
 [0141] acccaccctg cacgcgcccc tctccctcc ccatggcttc ccatggctat ctcacttggc 360
 [0142] cctaaaatgt ttaaggatga cactggctgc tgagtggaaa tgagacagca gaagtcaaca 420
 [0143] gtagattttt gaaaaaggcag agaaaaaggc ttgtgcttttt tagaaagc caagggacaa 480
 [0144] gctaagatag ggcggcaagta at 502
 [0145] <210> 9
 [0146] <211> 788
 [0147] <212> DNA
 [0148] <213> 小家鼠
 [0149] <400> 9
 [0150] aaatagaact gtgagatagg gggagagggg gcaggaagga caagagaccc ctgtctcatt 60
 [0151] gtgatccccca cctgtctgt ctgtggagg gtacccatga gggccagccc acagccctta 120
 [0152] ggtggacatt gtctggcct gtctcactgt ccctcccgac agcccccagag gcccaggagac 180
 [0153] aggggtctca gtcctcaactg agagatgtgt aaactgaggc ccagtgaatg ttgagggcca 240
 [0154] gggcatgccc ttgggtggat gtgacctggg tctccttcgc acgggcttcc tccccgaagc 300
 [0155] cgagctgagc atttggagtt tgaaatgttt ccgtacttag caatctgctc ctctattccc 360
 [0156] gggcggactt ccgatagctc cggccttatg ctgcactaga taagatggag cagggagagg 420
 [0157] acacggcact acttatgtaa ccgcctctt gaaaaatgga gcagcggta gggcggaaaca 480
 [0158] agacgtcctc tctctacgca tccctctcct ttccctgcta aggctgcagc tggagtcaga 540
 [0159] ggcagggctg ttccaatctg tcttgcata gtaacgcagc cagcctccag cctccgtcag 600
 [0160] cctcctcatg gctgagaccc ggcctcagtt tccccactt acatcccggag gatcagagcc 660
 [0161] tgtgaggatg aaatggata agtagctgg aaccgtctgg cagagagcga gtcctcagga 720
 [0162] ctgttgcata ctgtggctgc ctgccttgcac cccaaatgcac cccgcctcct catcctgcag 780
 [0163] caggagaa 788
 [0164] <210> 10
 [0165] <211> 502
 [0166] <212> DNA
 [0167] <213> 小家鼠

- [0168] <400> 10
- [0169] tctatagaat gtgtccccag cctgtttc cacacttgat acgcaaggaa tgcataccac 60
- [0170] agagagggat gaggtagca tccagcctgc ttccctgtgt tcggggcgct acagccacat 120
- [0171] ctccccagtc catctcagac cgtcacagag cttcgccgaa tgtatagctt tttctctgt 180
- [0172] gcagacaggg agacagagcc ttggaaagca tagtgcttg cttcttgcc cactgagtct 240
- [0173] tagctggact tgcacaccac atgcctcaca gccgggcgca cttgcatttgc tcacccaggc 300
- [0174] ccagtgtatga tggctctgtc tgctttgtc tttgtgccaa ctacagctcc agcacctgt 360
- [0175] ccctgggtt tcactccccc agttgaacac gtatgtactg gggttgttagg gatggagcct 420
- [0176] ttctgcttcc ttctggcaaa gtccttagcg gcctgctgca ggggtggggg gtgttcaggg 480
- [0177] gagtggtgat gaagtatgac ag 502
- [0178] <210> 11
- [0179] <211> 533
- [0180] <212> DNA
- [0181] <213> 小家鼠
- [0182] <400> 11
- [0183] tctccagttt gagaacacaga tgctgttaact gggccacag tataaagaga gcccagacat 60
- [0184] tgaactgtca acacagaagc ctggcacact ggaactggca gtccagctgg gaacaagggg 120
- [0185] tagaggctga gccactaag tcaactgagg caggagacat aggagctaaa gcagctgaag 180
- [0186] ggtgcaggac agctgggggg tctgaagtgg gcctcatgcc cagagctatg aagtcagggg 240
- [0187] ctgtagcccta ggagccttgg aagccagctg gcaagctgtg gcccaaagac gctgactcac 300
- [0188] caggaggggg cagctggagc cagggactcc taaggttcc aggaaggca gccttccagg 360
- [0189] gctcagctag gggagacagt gttgacagca agttgtcagg caacttgagc tactggcag 420
- [0190] ctgggaagct gtcccttggt cccagatc atcatcaccc cagacgctgc ccacctgcct 480
- [0191] caggtcccac acagtgtatcc tcccatctt aacacaacac atgaccagag aga 533
- [0192] <210> 12
- [0193] <211> 524
- [0194] <212> DNA
- [0195] <213> 小家鼠
- [0196] <400> 12
- [0197] gtcaccctcc ccccaaacaa ccccttcttc tctggttcga gaaattacag gcatgaaaga 60
- [0198] tataaatcgg gatgcttgc ttggatat aaatcactaa agcttgggg caggggtggg 120
- [0199] cgaccttgcgtt gaccgtcctt gtgcgtgcca gtaaatcctg tggccaggg gagaagaaaa 180
- [0200] ggctgtgtgg cttctgctca caaagctgca gaaaccattc ttaagccca aaagcacttc 240
- [0201] cagagagagc agagcatccc caggctgctg gctcagcaag ttcactgtgc tcaatctcag 300
- [0202] gaagtgagga taagagcagt gcctggagag tgcctggc tgagctgagg gtttctgaac 360
- [0203] acattaaagc ggggagcatg gaccggcct caggaggggt gttgaacatc cctaggcaga 420
- [0204] ggagtcttagc ttccctggaa aagatatcgt gttaaagcaca cacatgtcct ctggataaag 480
- [0205] ataatcttc tgatcacaca ctatacacac acaaaagcct gctc 524
- [0206] <210> 13
- [0207] <211> 509
- [0208] <212> DNA
- [0209] <213> 小家鼠

- [0210] <400> 13
- [0211] gcctcttagg ccacactgacc aggtcccctc agtccccccc ttcccacact cccacactca 60
- [0212] gccccccctcc cccccccccc acccctgcag gattatcctg tctgtgttcc tgactcagcc 120
- [0213] tgggagccac ctgggcagca ggggccaagg gtgtcctaga agggacctgg agtccacgct 180
- [0214] gggccaagcc tgccctttct ccctctgtct tccgtccctg cttcggttc tgctgaatgt 240
- [0215] gtttatttct ctggctcctt ttacagagaa tgctgctgct aattttatgt ggagctctga 300
- [0216] ggcagtgtaa ttggaagcca gacaccctgt cagcagtggg ctcccttcct gagctgccat 360
- [0217] gtttcctgct ctccctccgt cccggctcct catttcatgc agccacctgt cccagggaga 420
- [0218] gaggagtcac ccaggcccct cagtcggccc cttaaataag aaaggctccg ttgctcggca 480
- [0219] cacataccaa gcagccgctg gtcaatct 509
- [0220] <210> 14
- [0221] <211> 398
- [0222] <212> DNA
- [0223] <213> 小家鼠
- [0224] <400> 14
- [0225] gtttcttcc cttcccttt ggaccccgaa gacaagccaa taaaatactc ggcagggtgg 60
- [0226] ctctctcct tttttgcca gtaataaaca gactcagagc aagtaaggg tctggtccaa 120
- [0227] ggtcatggct gggatcagt acagagccca gaagagaacc tgagacttct tgctgagcca 180
- [0228] agctggagag gacagaaagg aatgcgtcta ctccatgcat gaccctctgc cagtttgct 240
- [0229] cttcctaag ggaccatgaa cgatatgtgc acaccgctca tacgtatgtg cacacctgca 300
- [0230] agaggaggca tcccatgtac acctatgaga cgcacagaga aacatatatg tagccatagg 360
- [0231] ctagaaattc tttcttttc taggtctgcc cctctgca 398
- [0232] <210> 15
- [0233] <211> 810
- [0234] <212> DNA
- [0235] <213> 小家鼠
- [0236] <400> 15
- [0237] ggaccactca gtgtacacgg aatgtagaat tgagtctgcc attggcttc cctcaaagtc 60
- [0238] ttggaggctt gggactgata ttgggagcat ctggcagag aaggccacaa agacagggtg 120
- [0239] gttttctac actggacat actcgtgagc atgcacagag gcgtgtcccc aacttccctg 180
- [0240] tcacccctgt cctctgcccgg ctagagggga tgcgggggtg gacatatgct gctattggc 240
- [0241] agatatcaca tggtaagagg tggggggggg ctcaagagggc ggagggttag gaggatcccc 300
- [0242] tggggagagg ttctggttt ctgctgcct ctagctgcta taaatacggtt agcacttgag 360
- [0243] caactggaaa gctctgagta atttaggatg cacaagctg taatttaact ccagcatctc 420
- [0244] agtgtgcgag agcattaaag atgtaattaa gatgtttaca caaagagatt ggagtctgt 480
- [0245] acacttgggg tgcaaaaccc caggaaggga cacaatgggt gaggtgagga tctgtggag 540
- [0246] gcctggggac agtcacttgg atcccagcta tgagatggca ggccacccag ctgtttctcc 600
- [0247] ttggaaatgt tttggcctgg ggggtgggg tggggcatca cacttgata tggagatggg 660
- [0248] gcaacaaagc ctgcaatatc tgggggtgga gaggtcaagt ggatggagtc tttgagatc 720
- [0249] atgtcaggaa gagggctcga tccccaaaa tcatggtgac atatggtgc tcggggttca 780
- [0250] caggagctat gtctaaaata caaaagtaaa 810
- [0251] <210> 16

- [0252] <211> 202
 [0253] <212> DNA
 [0254] <213> 小家鼠
 [0255] <400> 16
 [0256] tctgcagaag cctgccattc caccatttaa acctgtgact ccaggccta agcctgttga 60
 [0257] aggtcgagtc ccagaagggt catatgtca actgcctagg gagagttccc actcgcaggg 120
 [0258] ccaagaggag tccccggtc tgaggtgtgg gggcggggac gtgcactggg cgctggacc 180
 [0259] acggctgggg ctcaggactc gc 202
 [0260] <210> 17
 [0261] <211> 937
 [0262] <212> DNA
 [0263] <213> 小家鼠
 [0264] <400> 17
 [0265] tgccctcagtt tcttcgccta gaaagccggg tctaaggta catgccctga ttctttctg 60
 [0266] ggggtgtctcg aattttaaac aacacatact gttctgggt gatgacaaga ggaagtactg 120
 [0267] gtcgggtggct gatggacatc caccatggtg gcaactggag ggagggggaa cggacgttga 180
 [0268] aaccctgccc tcctggaaatc tgtcgcatgc acgcacgtt acaatgctt gcactgggaa 240
 [0269] caggctggga tggatggagc ggagcgttag gaggagtggg catgcaggcc cgagtgtctg 300
 [0270] ttttgctgat tgctccttt gcttcaagg agattaaact attttagtc catgcctact 360
 [0271] gctgggtgaga cgctggagga agccttcca tcggtgagat tttctggaaag ctgccaagtg 420
 [0272] tggtcttcag ctaattctg ggagcctccc agagtggag ggaggaacat ttccatctgg 480
 [0273] gggcttcggg gacaggctaa gatttccct ggggtcctg ctgcgctggc ctcctcaaac 540
 [0274] cacgctgcct cggcctgcat aaagcagtaa tctgatgtgc ccgatgttt taacgctgt 600
 [0275] tttaaaaaaaa gtaatttatt ttctaattat tccttgtt gcataaccat gcattgcca 660
 [0276] agtgtcgcta tttaaaatat ttatctctcc acgccgcagg agcagctctg gagcgtggag 720
 [0277] gggaaagaaa taaaagtccg cgtgccagtc gcaggcatat tacttgact cgtcctgg 780
 [0278] gctttgacgt ctccctgtaa atacatttat ttttcattag gacgtttctg agcttgtggc 840
 [0279] ccccgagag cggagtgtt acgctgttca tctgcaagcg atgcaataga ggggtactcg 900
 [0280] cagaatgact tccgcccaga gcatcctgct cctgtct 937
 [0281] <210> 18
 [0282] <211> 922
 [0283] <212> DNA
 [0284] <213> 小家鼠
 [0285] <400> 18
 [0286] taaaataacct tattttttc cagtctctaa actgctaatac tcccaggcta agggattctg 60
 [0287] ggacaaaggc aaggcctgga agtgaaatc tgtaaaatta gtttcagcgg tatttagtgg 120
 [0288] tgcagttgaa gattgaaaaa ctgtttccc agggcctgat tggaggctcc actctcctcc 180
 [0289] aggaagaggc aaggactctg ggctggcact gaggacaaat cctggaggc tgctatggg 240
 [0290] cctgggagcc aggctgcctt gtcttagagg ccttagagat gtctgtgtcc caagtccaa 300
 [0291] gctaccccca gcagctaaca gctttccag ttctcaggca cagcagggtgc caagatcac 360
 [0292] ctctggagtc cagctgggcc cttcctt cttttttttt aagacctcct 420
 [0293] ggacactgtt cctctcccccc ccccccgtgac cccccccctc agttctcaaa cacgtgaggg 480

- [0294] ttggggagg gttccacagc cagagagagg ggccagctct ggtgcctgtg ggtaccccc 540
 [0295] cccgtatggc ccatcaggcc tcttgtgc ttgattgcct ctgattggct gcagctgaat 600
 [0296] tcagcaaaag ctattatttg cccttgcata gccaatcaga tggcctcatt ggccattcag 660
 [0297] agcaggcacc ggaacctgag ggtgggtgg ggggtgggg atggagatgg gactcagtga 720
 [0298] ggggtggga agctctaaaa cagatgcagg acctgagccct gtctgtgtcc accacgacct 780
 [0299] tcacacaggt cacacccct tcccctgact tgtcacccca aaccaggcgt tggtgcctaa 840
 [0300] ccccacctca caattccctc actctgtaac acctttccat atacctctgc atgtctaaac 900
 [0301] ccaagacttg ctctatgaaa tc 922
 [0302] <210> 19
 [0303] <211> 639
 [0304] <212> DNA
 [0305] <213> 小家鼠
 [0306] <400> 19
 [0307] agaccctgct tagcacagct ctttaggggt ccttttaggg gtctccagc gggcccagt 60
 [0308] ggaatgagat aaggaaggac acagctgtcc attctcccggt gcctgctaag gaggaaatgg 120
 [0309] ggccgcctta cataatttggg gcaatttggtt ccactcttgcgtt cctctggta tcacggctat 180
 [0310] cacccctcc ttgctcaggagtt agtccttgcgtt tgagcgagaa gctcaggcct ccctctctcc 240
 [0311] ctccctgctgg ggggtgcgtg acagagggtg taggagccat aggctctgtc actgctgaga 300
 [0312] tctgccagat gtctaggcca ggagaaaaatg gaaaggcgtt agtcacagca tatgtggcca 360
 [0313] ctcaggccta tagccccaaa tctgcctggt aaccattat gtcccccagag aatttgcatt 420
 [0314] ggccggacacc ctcatgccgg gtctcagtaa gggaggggt gggaggcaaa aatatccctc 480
 [0315] cccaccctga atctccaccc cctccccca gaaactgaca cttggccttgcgtt tctaaggatg 540
 [0316] ggtttccca aaatccttct gaaaaaaaaaca gaatttcaag agtcaactccc tccgggtctc 600
 [0317] agcctagaac atatgcagta tccctgacg tccataggg 639
 [0318] <210> 20
 [0319] <211> 716
 [0320] <212> DNA
 [0321] <213> 小家鼠
 [0322] <400> 20
 [0323] aaactggcac agtaatggcg ggctgacaga caagggagtc tgcgtaccc gctgcctcc 60
 [0324] cccaccctt ctccgagcaa taaaagggtt tttatgtggg gctggcagtg gcttctgcct 120
 [0325] cccttcatt acgaacatttta agagatcttgcgtt acccttccac tttccccgtt cttgaaaggta 180
 [0326] gctgcagaca cgtggagcca attaggcgca cgcgtggcg ccaagggcct gagcagctt 240
 [0327] ttctccctga ttgcggcggtt tacagctgtt tattctcccc tcacccaaac agtgcgtctt 300
 [0328] cctggcaagg tgccacccag aggagccggc tggggcccc tggggacagg ggaggactgg 360
 [0329] attagtaaat gggcatctat cgaatggctt tcatatgtgtt ggctgaaagg gagaaggta 420
 [0330] gggccaggaa tgggtggcagc aaggcccag gtagcaatga gggttcttctt aaccaccat 480
 [0331] ttagggatag cgatcagaaa agggccctcg aggaggtgac ctaaatgtgtt gtatagctg 540
 [0332] acggccacta cacacacaca cacacacaca cacacacata cacaagcatc cttgtccttgcgtt 600
 [0333] gagtcggta gcatgagcaa gagaaagatg ttcccagtgccgtt ccatgagagt ggagccctcc 660
 [0334] tccctactta catccaggtt ggatggccag gagatcctga gatccttcaa gactcc 716
 [0335] <210> 21

- [0336] <211> 147
[0337] <212> DNA
[0338] <213> 小家鼠
[0339] <400> 21
[0340] aagccacatc ctgggtggaa atatatggct tcaattccca ctcttccgga tgaccttgt 60
[0341] ggggagccct ggcttcacct tggccagct tcatccctta gcctcgctgc caggaaggca 120
[0342] gtgaggtcag aggctggtgc tggcgtg 147
[0343] <210> 22
[0344] <211> 297
[0345] <212> DNA
[0346] <213> 小家鼠
[0347] <400> 22
[0348] ctttacacttggt gcccgcac acatctgggc catcctggcc agcggcagcg tggtggtgaa 60
[0349] ggcactgtgc gccgtggtac tttttctcta cctgctttcc ttgcgtgtgg acacgggctg 120
[0350] cctggccgtc acccccaggct accttttccc acccaacttc tggatctgga ccctggccac 180
[0351] ccacgggctc atggaacagc acgtgtggaa cgtggccatt agcctggcca cagtggttgt 240
[0352] gccccggcga ttactggagc ccctctgggg agccttggag ctgctcatct tcttctc 297
[0353] <210> 23
[0354] <211> 101
[0355] <212> DNA
[0356] <213> 智人
[0357] <400> 23
[0358] aaacggacgg gcctccgctg aaccagttag gccccagacg tgccataaaa taacccctgc 60
[0359] gtgctgcacc acctggggag agggggagga ccacggtaaa t 101
[0360] <210> 24
[0361] <211> 101
[0362] <212> DNA
[0363] <213> 智人
[0364] <400> 24
[0365] ggagcgagcg catagaaaaa gggacgcggg gtcctttct ctgccggtag cactggtag 60
[0366] ctgtggccag gtgtggact ttgatgggc ccagggtgg a 101
[0367] <210> 25
[0368] <211> 101
[0369] <212> DNA
[0370] <213> 智人
[0371] <400> 25
[0372] gctcaaggaa gcgtcgagg gtcacagatc tggggaaacc ccggggaaaaa gcactgaggc 60
[0373] aaaaccggccg ctcgtctcct acaatatacg ggagggggag g 101
[0374] <210> 26
[0375] <211> 547
[0376] <212> DNA
[0377] <213> 智人

- [0378] <400> 26
- [0379] tttagtacgt tctggattac tcataagacc tttttttt cttccggc gcaaaaccgt 60
- [0380] gagctggatt tataatcgcc ctataaagct ccagaggcgg tcaggcacct gcagaggagc 120
- [0381] cccgcccgtc cgccgactag ctgccccgc gagcaacggc ctcgtgattt ccccgccgat 180
- [0382] ccggtccccg cctccccact ctgccccgc ctaccccgga gccgtgcage cgcctctccg 240
- [0383] aatctctctc ttctcctggc gtcgcgtgc gagagggAAC tagcgagaAC gaggaAGCAG 300
- [0384] ctggaggtga cgccggcag attacgcctg tcagggcga gccgagcggA tcgctggcgc 360
- [0385] ctgtcagag gaaaggcggg agtgcggc tcgctgtcgc agagggagg tggtaagct 420
- [0386] agcgaccacc tggacttccc agcgcccaac cgtggcttt cagccaggTC ctctcctccc 480
- [0387] gccgcttctc aaccaacccc atcccagcgc cggccaccca acctccgaa atgagtgcTT 540
- [0388] cctgccc 547
- [0389] <210> 27
- [0390] <211> 101
- [0391] <212> DNA
- [0392] <213> 智人
- [0393] <400> 27
- [0394] cagcagccga aggccgtact aggaacggta acctgttact tttcagggg ccgttagtcga 60
- [0395] cccgctgccc gagttgtgt gcgactgcgc ggcggggct a 101
- [0396] <210> 28
- [0397] <211> 201
- [0398] <212> DNA
- [0399] <213> 智人
- [0400] <400> 28
- [0401] gagtgcaagg tgactgtggt tcttctctgg ccaagtccga gggagaacgt aaagatatgg 60
- [0402] gccttttcc ccctctcacc ttgtctcacc aaagtcccta gtcccccggag cagttgcct 120
- [0403] ctttcttcc agggattAG ccagacacAA caacggAAC cagacaccGA accagacATG 180
- [0404] cccgccccgt gcgcctcccc c 201
- [0405] <210> 29
- [0406] <211> 199
- [0407] <212> DNA
- [0408] <213> 智人
- [0409] <400> 29
- [0410] gtcgcgtgcc ttccctccct ctgtctctc cagagccggA tcttaaggg gggccctccgt 60
- [0411] gccccggct gtcagtcctc tccgggtgtgc aggaccccgA aagtccctccc cgcacagctc 120
- [0412] tcgcttctct ttgcagcctg ttctgcgc ggaccagtcg aggactctgg acagtagagg 180
- [0413] ccccgggacg accgagctg 199
- [0414] <210> 30
- [0415] <211> 1351
- [0416] <212> DNA
- [0417] <213> 智人
- [0418] <400> 30
- [0419] aaacggacgg gcctccgctg aaccagttag gccccagacg tgccataaa taaccctgc 60

[0420] gtgctgcacc acctggggag agggggagga ccacggtaaa tggagcgagc gcatagcaaa 120
 [0421] agggacgcgg ggtcctttc tctgccgtg gcactggta gctgtggca ggtgtggta 180
 [0422] tttgatgggg cccaggcgtg gagctcaagg aagcgtcgca gggtcacaga tctggggaa 240
 [0423] ccccggggaa aagcaactgag gcaaaaaccgc cgctcgctc ctacaatata tgggaggggg 300
 [0424] aggtttagta cgttctggat tactcataag acctttttt ttccttccg ggcgcaaaac 360
 [0425] cgtgagctgg atttataatc gccctataaa gctccagagg cggtcaggca cctgcagagg 420
 [0426] agccccgccc ctccgcccac tagctgcccc cgcgagcaac ggcctcgta tttccccccc 480
 [0427] gatccggtcc ccgcctcccc actctgcccc cgccctacccc ggagccgtgc agccgcctct 540
 [0428] cccaatctct ctcttctcct ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag 600
 [0429] cagctggagg tgacgcccgg cagattacgc ctgtcaggc cgagccgagc ggatcgctgg 660
 [0430] ggcgtgtgca gaggaaaggc gggagtgcgg gcgtcgctgt cgccagagccg aggtggtaa 720
 [0431] getagcgacc acctggactt cccagcgccc aaccgtggct tttcagccag gtcctctcct 780
 [0432] ccccgccgtt ctcaaccaac cccatcccg cgccggccac ccaacctccc gaaatgagtg 840
 [0433] ctccctgccc cagcagccga aggccgtact aggaacggta acctgttact tttccagggg 900
 [0434] ccgtagtcga cccgctgccc gagttgctgt gcgactgcgc ggcggggct agagtcaag 960
 [0435] gtgactgtgg ttcttctctg gccaagtccg agggagaacg taaagatatg ggccttttc 1020
 [0436] cccctctcac cttgtctcac caaagtccct agtccccggc gcagttagcc tcttttttc 1080
 [0437] cagggattta gccagacaca acaacggaa ccagacaccc aaccagacat gcccccccc 1140
 [0438] tgccgcctcc cgcgctcgctg ctttcctcc ctcttgcctc tccagagccg gatcttcaag 1200
 [0439] gggagccctcc gtccccccgg ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc 1260
 [0440] cccgcacagc tctcgcttct ctttgcagcc tggggctcgcc ccggaccagt cgaggactct 1320
 [0441] ggacagttaga ggccccggga cgaccgagct g 1351
 [0442] <210> 31
 [0443] <211> 2051
 [0444] <212> DNA
 [0445] <213> 智人
 [0446] <400> 31
 [0447] ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60
 [0448] atttgaagt tctgttgaca taaagaatca tgatattaat gcccattggaa atgaaaggc 120
 [0449] gatcaacact atggttgaa aaggggaaa ttgttagagca cagatgtgtt cgtgtggcag 180
 [0450] tgtgctgtct ctagcaatac tcagagaaga gagagaacaa taaaattctg attggccccca 240
 [0451] gtgtgagccc agatgagggtt cagctgcca ctttctctt cacatctt gaaagtctt 300
 [0452] taagcacaac taactttttt tttttttt tttttttttag acagagtctt gctctgtgc 360
 [0453] ccaggacaga gtgcagtagt gactcaatct cggctcaactg cagcctccac ctccttaggt 420
 [0454] caaacggtcc tcctgcatca gcctcccaag tagctggaa tacaggagtgc gcccaccatg 480
 [0455] cccagctaattttt ttaatagata cgggggtttc accatatac ccaggctgg 540
 [0456] ctcgaactcc tggcctcaag tgatccacct gcctcgccct cccaaagtgc tgggattata 600
 [0457] ggcgtcagcc actatgccc acccgaccaa cttttttaa aataaatatt taaaaattt 660
 [0458] gtatttcaca tatataactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720
 [0459] aaccagttag gccccagacg tgccataaa taaccctgc gtgctgcacc acctggggag 780
 [0460] agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtcctttc 840
 [0461] tctgcccgtg gcactggta gctgtggca ggtgtggta tttgatgggg cccaggctg 900

[0462] gagctcaagg aagcgtcgca gggcacaga tctggggaa cccggggaa aagcactgag 960
 [0463] gcaaaaaccgc cgctcgctc ctacaatata tgggaggggg aggtttagta cgttctggat 1020
 [0464] tactcataag acctttttt tttccttccg ggcgc当地 cgtgagctgg atttataatc 1080
 [0465] gccctataaaa gctccagagg cggtcaggca cctgcagagg agccccgccc ctccgcccac 1140
 [0466] tagctgcccc cgcgagcaac ggctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200
 [0467] actctgcccc cgccctacccc ggagccgtgc agccgcctc ccgaatctct ctcttctct 1260
 [0468] ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320
 [0469] cagattacgc ctgtcagggc cgagccgagc ggatcgctgg ggcgctgtgc gaggaaaggc 1380
 [0470] gggagtgccc ggctcgctgt cgccagagccg aggtggtaa gctagcgacc acctggactt 1440
 [0471] cccagcgccc aaccgtggct tttcagccag gtcctctctt ccccgccgtt ctcaccaac 1500
 [0472] cccatcccag cgccggccac ccaacccc gaaatgagtg cttctgccc cagcagccga 1560
 [0473] aggcgctact aggaacggta acctgttact tttcagggg ccgtagtcga cccgctgccc 1620
 [0474] gagttgctgt gcgactgcgc gcgcggggct agagtgcag gtgactgtgg ttcttctctg 1680
 [0475] gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740
 [0476] caaagtccct agtccccggc gcagttagcc tctttcttcc cagggatta gccagacaca 1800
 [0477] acaacggaa ccagacaccc aaccagacat gcccggcccg tgcggccctcc cgcgtcgctg 1860
 [0478] ctttcctcc ctcttgcgtc tccagagccg gatcttcaag gggagccctcc gtgggggggg 1920
 [0479] ctgctcagtc cctccgggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980
 [0480] cttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagttaga ggccccggga 2040
 [0481] cgaccgagct g 2051
 [0482] <210> 32
 [0483] <211> 1878
 [0484] <212> DNA
 [0485] <213> 未知
 [0486] <220>
 [0487] <223> 对未知的描述：
 [0488] RE序列
 [0489] <400> 32
 [0490] tcaacagggg gacacttggg aaagaaggat ggggacagag ccgagaggac tgttacacat 60
 [0491] tagagaaaca tcagtactg tgccagctt gggtagact gcacaaaagc cctgaggcag 120
 [0492] cacaggcagg atccagtcg ctggcccag gaagctaacc gtctcagaca gaggcacaaag 180
 [0493] caccgagaca tgtgccacaa ggcttgcgtta gagaggtcag aggacagcgt acaggtcccc 240
 [0494] gagatcaaacc tcaacccac caggcttggc agcaagcctt taccaacccca ccccccacccc 300
 [0495] acccaccctg cacgc当地 tctccctcc ccatggctc ccatggctat ctcacttggc 360
 [0496] cctaaaatgt ttaaggatga cactggctgc tgagtggaaa tgagacagca gaagtcaaca 420
 [0497] gtagattttt gaaaaaggcag agaaaaaggc ttgtgctgtt tttagaaagc caagggacaa 480
 [0498] gctaagatag ggcccaagta atgctgtat ttacatttat ccacacaaaa cggacggggcc 540
 [0499] tccgctgaac cagtgaggcc ccagacgtgc gcataaataa cccctgcgtg ctgcaccacc 600
 [0500] tggggagagg gggaggacca cggtaaatgg agcgagcgca tagaaaaagg gacgcgggggt 660
 [0501] cttttctct cccgggtggca ctggtagct gtggccaggt gtggacttt gatggggccc 720
 [0502] agggctggag ctcaaggaag cgtcgcaggc tcacagatct ggggaaccc cggggaaaag 780
 [0503] cactgaggca aaaccggccgc tcgtctccta caatatatgg gagggggagg ttgagtagct 840

[0504] tctggattac tcataagacc tttttttt cttccggc gcaaaaccgt gagctggatt 900
 [0505] tataatcgcc ctataaaact ccagaggcgg tcagggcacct gcagaggagc cccgcccgtc 960
 [0506] cggcactag ctgccccgc gagacaacggc ctcgtgattt ccccgccat ccggccccg 1020
 [0507] cctccccact ctgccccgc ctaccccgga gccgtgcagc cgcctctccg aatctcttc 1080
 [0508] ttctcctggc gctcgctgc gagaggaaac tagcgagaac gaggaagcag ctggaggtga 1140
 [0509] cgccggcag attacgcctg tcagggccga gccgagcgga tcgctggcg ctgtgcagag 1200
 [0510] gaaaggcggg agtccccggc tcgctgtcgc agagccgagg tggtaagct agcgaccacc 1260
 [0511] tgacttccc agcgcacaac cgtggcttt cagccaggtc ctctcctccc gcggcttctc 1320
 [0512] aaccaacccc atcccagcgc cggccaccca acctcccgaa atgagtgcctt cctccccag 1380
 [0513] cagccgaagg cgctactagg aacggtaacc tgtaacttt ccagggccg tagtcgaccc 1440
 [0514] gctccccgag ttgctgtcgc actgcgcgcg cggggctaga gtgcaaggtg actgtggttc 1500
 [0515] ttctctggcc aagtccgagg gagaacgtaa agatatggc cttttcccc ctctcacctt 1560
 [0516] gtctcaccaa agtccctagt cccggagca gttagcctct ttcttccag ggaatttagcc 1620
 [0517] agacacaaca acgggaacca gacaccgaac cagacatgcc cgcgggtgc gccctccccg 1680
 [0518] ctcgctgcct ttcctccctc ttgtctctcc agagccggat cttcaagggg agcctccgtg 1740
 [0519] ccccccggctg ctcagtcct ccgggtgtcga ggaccccgga agtccctccc gcacagctct 1800
 [0520] cgttctctt tgcagcctgt ttctgcgcgc gaccagtgcg ggactctgga cagtagaggc 1860
 [0521] cccgggacga ccgagctg 1878
 [0522] <210> 33
 [0523] <211> 19
 [0524] <212> DNA
 [0525] <213> 智人
 [0526] <400> 33
 [0527] attacattt atccacaca 19
 [0528] <210> 34
 [0529] <211> 800
 [0530] <212> DNA
 [0531] <213> 小家鼠
 [0532] <400> 34
 [0533] tgccgctgga ctctttcca aggaactagg agaaccaaga tccgttttc tgccaaggc 60
 [0534] tgccccccccc acgcccccaa cccctcacc ccgatcccc cagaaagaaa tcttgaggtt 120
 [0535] gctggagctt cttctgtggg tgtacagga ctgcattct cctctgttagt ctgcagaagc 180
 [0536] ctgccattcc accatttaaa cctgtactc caggcattaa gcctgtgaa ggtcgagtcc 240
 [0537] cagaagggtc atatgtcaa ctgcctaggg agagttccca ctcgcaggc caagaggagt 300
 [0538] ccccccggctc gaggtgtggg ggccgggacg tgcactggc gctgggacca cggctggggc 360
 [0539] tcaggactcg cgagcttggc ttggatcgg tttgcgcag ccagtagggc aggctccggg 420
 [0540] gtgaacgggg acgaggggcgc cgccggcaca ggcggcgcgc tgaccgcgc gggggcgcgc 480
 [0541] ggaggcgggc cggccaagga gagggaggga gggaatgagg gagggagcga cagggaggg 540
 [0542] cggcgccggc aggttggcgg cggccgctat ttgagcgcag gtccgggcc aggcgcctaa 600
 [0543] agcgcttggc gccagcgcgg cggggagatc gctgcgcgc gcccgcagag ggcgcgcggc 660
 [0544] cagtgcagcc ccggaggccc cgcgcggaga aggaggtgga gaagaggccg gctttccgccc 720
 [0545] cggcccccgc gccccccac ctccatcccg cggccgcgt ccccccctccc tcccccgccg 780

- [0546] gcccgcattt gaatggaaac 800
 [0547] <210> 35
 [0548] <211> 1335
 [0549] <212> DNA
 [0550] <213> 人工序列
 [0551] <220>
 [0552] <223> 对人工序列的描述：合成多核苷酸
 [0553] <400> 35
 [0554] gagtaattca tacaaaagga ctcgccccctg cttggggaa tcccaaggac cgtcgtaaa 60
 [0555] ctccccactaa cgtagaaccc agagatcgct gcgttcccgcc cccctcaccgc gcccgcctc 120
 [0556] gtcatcaactg aggtggagaa gagcatgcgt gaggctccgg tgcccgtag tggcagagc 180
 [0557] gcacatcgcc cacagtcccc gagaagttgg ggggagggtt cggcaattga accggtgccc 240
 [0558] agagaaggtg gcgccgggta aactggaaa gtatgtcgt gtactggctc cgccttttc 300
 [0559] ccgagggtgg gggagaaccg tatataagtg cagtagtcgc cgtAACGTT cttttcgca 360
 [0560] acgggtttgc cgccagaaca caggttaagtgc cctgtgtgg ttcccgccggg cctggctc 420
 [0561] ttacgggtta tggcccttgc gtgccttgc ttacttccac gcccctggct gcagtaatgt 480
 [0562] attcttgc cccgagcttcg ggttggaaat ggggtggaga gttcgaggcc ttgcgtttaa 540
 [0563] ggagccctt cgcctcgat ttgagtttag gcctggcttgc ggcgcgggg ccgcgcgtg 600
 [0564] cgaatctggc ggcacccatcg cgcctgtctc gctgtttcg ataatctct agccatttaa 660
 [0565] aattttttat gacctgctgc gacgttttt ttctggcaag atatgtttgt aaatgcgggc 720
 [0566] caagatctgc acactggat ttgggtttttt gggccgcgg gcggcgacgg gcccgtgcg 780
 [0567] tcccagcgca catgttcggc gaggccgggc ctgcgagcgc ggcacccgg aatcgacgg 840
 [0568] ggtagtctc aagctggccg gcctgctctg gtgcctggcc tcgcgcggcc gtgtatcgcc 900
 [0569] ccgcctggg cggcaaggct ggccggctcg gcaccagttt cgtgagcggaa aagatggccg 960
 [0570] ctcccgccctt ctgctgcagg gagctaaaaa tggaggacgc ggcgcctggg agagcgggc 1020
 [0571] ggtgagtac ccacacaaag gaaaaggccccc ttccgtcct cagccgtcgc ttcatgtgac 1080
 [0572] tccacggagt accggcgcc gtcaggcac ctcgattagt tctcgagctt ttggagtgac 1140
 [0573] tcgtcttag gttggggggaa ggggtttat gcgatggagt ttccacac tgagtgggtg 1200
 [0574] gagactgaag ttaggccagc ttggacttg atgtattttt ctttggattt tgccctttt 1260
 [0575] gagtttggat ctgggttcat tctcaagcct cagacagtgg ttcaaagttt tttcttcca 1320
 [0576] ttccagggtt cgtga 1335
 [0577] <210> 36
 [0578] <211> 222
 [0579] <212> PRT
 [0580] <213> 人工序列
 [0581] <220>
 [0582] <223> 对人工序列的描述：合成多核苷酸
 [0583] <400> 36
 [0584] Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 [0585] 1 5 10 15
 [0586] Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 [0587] 20 25 30

[0588]	Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile		
[0589]	35	40	45
[0590]	Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr		
[0591]	50	55	60
[0592]	Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys		
[0593]	65	70	75
[0594]	Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu		
[0595]	85	90	95
[0596]	Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu		
[0597]	100	105	110
[0598]	Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly		
[0599]	115	120	125
[0600]	Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr		
[0601]	130	135	140
[0602]	Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn		
[0603]	145	150	155
[0604]	Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser		
[0605]	165	170	175
[0606]	Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly		
[0607]	180	185	190
[0608]	Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu		
[0609]	195	200	205
[0610]	Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu		
[0611]	210	215	220
[0612]	<210> 37		
[0613]	<211> 218		
[0614]	<212> PRT		
[0615]	<213> 未知		
[0616]	<220>		
[0617]	<223> 对未知的描述:		
[0618]	SCN1B序列		
[0619]	<400> 37		
[0620]	Met Gly Thr Leu Leu Ala Leu Val Val Gly Ala Ala Leu Val Ser Ser		
[0621]	1	5	10
[0622]	Ala Trp Gly Gly Cys Val Glu Val Asp Ser Asp Thr Glu Ala Val Tyr		
[0623]	20	25	30
[0624]	Gly Met Thr Phe Lys Ile Leu Cys Ile Ser Cys Lys Arg Arg Ser Glu		
[0625]	35	40	45
[0626]	Thr Thr Ala Glu Thr Phe Thr Glu Trp Thr Phe Arg Gln Lys Gly Thr		
[0627]	50	55	60
[0628]	Glu Glu Phe Val Lys Ile Leu Arg Tyr Glu Asn Glu Val Leu Gln Leu		
[0629]	65	70	75
			80

[0630]	Glu Glu Asp Glu Arg Phe Glu Gly Arg Val Val Trp Asn Gly Ser Arg		
[0631]	85	90	95
[0632]	Gly Thr Lys Asp Leu Gln Asp Leu Ser Ile Phe Ile Thr Asn Val Thr		
[0633]	100	105	110
[0634]	Tyr Asn His Ser Gly Asp Tyr Glu Cys His Val Tyr Arg Leu Leu Phe		
[0635]	115	120	125
[0636]	Phe Asp Asn Tyr Glu His Asn Thr Ser Val Val Lys Lys Ile His Leu		
[0637]	130	135	140
[0638]	Glu Val Val Asp Lys Ala Asn Arg Asp Met Ala Ser Ile Val Ser Glu		
[0639]	145	150	155
[0640]	Ile Met Met Tyr Val Leu Ile Val Val Leu Thr Ile Trp Leu Val Ala		
[0641]	165	170	175
[0642]	Glu Met Val Tyr Cys Tyr Lys Lys Ile Ala Ala Ala Thr Glu Ala Ala		
[0643]	180	185	190
[0644]	Ala Gln Glu Asn Ala Ser Glu Tyr Leu Ala Ile Thr Ser Glu Ser Lys		
[0645]	195	200	205
[0646]	Glu Asn Cys Thr Gly Val Gln Val Ala Glu		
[0647]	210	215	
[0648]	<210> 38		
[0649]	<211> 215		
[0650]	<212> PRT		
[0651]	<213> 未知		
[0652]	<220>		
[0653]	<223> 对未知的描述:		
[0654]	SCN2B序列		
[0655]	<400> 38		
[0656]	Met His Arg Asp Ala Trp Leu Pro Arg Pro Ala Phe Ser Leu Thr Gly		
[0657]	1 5 10 15		
[0658]	Leu Ser Leu Phe Phe Ser Leu Val Pro Pro Gly Arg Ser Met Glu Val		
[0659]	20 25 30		
[0660]	Thr Val Pro Ala Thr Leu Asn Val Leu Asn Gly Ser Asp Ala Arg Leu		
[0661]	35 40 45		
[0662]	Pro Cys Thr Phe Asn Ser Cys Tyr Thr Val Asn His Lys Gln Phe Ser		
[0663]	50 55 60		
[0664]	Leu Asn Trp Thr Tyr Gln Glu Cys Asn Asn Cys Ser Glu Glu Met Phe		
[0665]	65 70 75 80		
[0666]	Leu Gln Phe Arg Met Lys Ile Ile Asn Leu Lys Leu Glu Arg Phe Gln		
[0667]	85 90 95		
[0668]	Asp Arg Val Glu Phe Ser Gly Asn Pro Ser Lys Tyr Asp Val Ser Val		
[0669]	100 105 110		
[0670]	Met Leu Arg Asn Val Gln Pro Glu Asp Glu Gly Ile Tyr Asn Cys Tyr		
[0671]	115 120 125		

[0672]	Ile Met Asn Pro Pro Asp Arg His Arg Gly His Gly Lys Ile His Leu		
[0673]	130	135	140
[0674]	Gln Val Leu Met Glu Glu Pro Pro Glu Arg Asp Ser Thr Val Ala Val		
[0675]	145	150	155
[0676]	Ile Val Gly Ala Ser Val Gly Gly Phe Leu Ala Val Val Ile Leu Val		
[0677]	165	170	175
[0678]	Leu Met Val Val Lys Cys Val Arg Arg Lys Lys Glu Gln Lys Leu Ser		
[0679]	180	185	190
[0680]	Thr Asp Asp Leu Lys Thr Glu Glu Glu Gly Lys Thr Asp Gly Glu Gly		
[0681]	195	200	205
[0682]	Asn Pro Asp Asp Gly Ala Lys		
[0683]	210	215	
[0684]	<210> 39		
[0685]	<211> 2009		
[0686]	<212> PRT		
[0687]	<213> 未知		
[0688]	<220>		
[0689]	<223> 对未知的描述:		
[0690]	SCN1A序列		
[0691]	<400> 39		
[0692]	Met Glu Gln Thr Val Leu Val Pro Pro Gly Pro Asp Ser Phe Asn Phe		
[0693]	1	5	10
[0694]	Phe Thr Arg Glu Ser Leu Ala Ala Ile Glu Arg Arg Ile Ala Glu Glu		
[0695]	20	25	30
[0696]	Lys Ala Lys Asn Pro Lys Pro Asp Lys Lys Asp Asp Asp Glu Asn Gly		
[0697]	35	40	45
[0698]	Pro Lys Pro Asn Ser Asp Leu Glu Ala Gly Lys Asn Leu Pro Phe Ile		
[0699]	50	55	60
[0700]	Tyr Gly Asp Ile Pro Pro Glu Met Val Ser Glu Pro Leu Glu Asp Leu		
[0701]	65	70	75
[0702]	Asp Pro Tyr Tyr Ile Asn Lys Lys Thr Phe Ile Val Leu Asn Lys Gly		
[0703]	85	90	95
[0704]	Lys Ala Ile Phe Arg Phe Ser Ala Thr Ser Ala Leu Tyr Ile Leu Thr		
[0705]	100	105	110
[0706]	Pro Phe Asn Pro Leu Arg Lys Ile Ala Ile Lys Ile Leu Val His Ser		
[0707]	115	120	125
[0708]	Leu Phe Ser Met Leu Ile Met Cys Thr Ile Leu Thr Asn Cys Val Phe		
[0709]	130	135	140
[0710]	Met Thr Met Ser Asn Pro Pro Asp Trp Thr Lys Asn Val Glu Tyr Thr		
[0711]	145	150	155
[0712]	Phe Thr Gly Ile Tyr Thr Phe Glu Ser Leu Ile Lys Ile Ile Ala Arg		
[0713]	165	170	175

[0714]	Gly Phe Cys Leu Glu Asp Phe Thr Phe Leu Arg Asp Pro Trp Asn Trp		
[0715]	180	185	190
[0716]	Leu Asp Phe Thr Val Ile Thr Phe Ala Tyr Val Thr Glu Phe Val Asp		
[0717]	195	200	205
[0718]	Leu Gly Asn Val Ser Ala Leu Arg Thr Phe Arg Val Leu Arg Ala Leu		
[0719]	210	215	220
[0720]	Lys Thr Ile Ser Val Ile Pro Gly Leu Lys Thr Ile Val Gly Ala Leu		
[0721]	225	230	235
[0722]	Ile Gln Ser Val Lys Lys Leu Ser Asp Val Met Ile Leu Thr Val Phe		
[0723]	245	250	255
[0724]	Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe Met Gly Asn		
[0725]	260	265	270
[0726]	Leu Arg Asn Lys Cys Ile Gln Trp Pro Pro Thr Asn Ala Ser Leu Glu		
[0727]	275	280	285
[0728]	Glu His Ser Ile Glu Lys Asn Ile Thr Val Asn Tyr Asn Gly Thr Leu		
[0729]	290	295	300
[0730]	Ile Asn Glu Thr Val Phe Glu Phe Asp Trp Lys Ser Tyr Ile Gln Asp		
[0731]	305	310	315
[0732]	Ser Arg Tyr His Tyr Phe Leu Glu Gly Phe Leu Asp Ala Leu Leu Cys		
[0733]	325	330	335
[0734]	Gly Asn Ser Ser Asp Ala Gly Gln Cys Pro Glu Gly Tyr Met Cys Val		
[0735]	340	345	350
[0736]	Lys Ala Gly Arg Asn Pro Asn Tyr Gly Tyr Thr Ser Phe Asp Thr Phe		
[0737]	355	360	365
[0738]	Ser Trp Ala Phe Leu Ser Leu Phe Arg Leu Met Thr Gln Asp Phe Trp		
[0739]	370	375	380
[0740]	Glu Asn Leu Tyr Gln Leu Thr Leu Arg Ala Ala Gly Lys Thr Tyr Met		
[0741]	385	390	395
[0742]	Ile Phe Phe Val Leu Val Ile Phe Leu Gly Ser Phe Tyr Leu Ile Asn		
[0743]	405	410	415
[0744]	Leu Ile Leu Ala Val Val Ala Met Ala Tyr Glu Glu Gln Asn Gln Ala		
[0745]	420	425	430
[0746]	Thr Leu Glu Glu Ala Glu Gln Lys Glu Ala Glu Phe Gln Gln Met Ile		
[0747]	435	440	445
[0748]	Glu Gln Leu Lys Lys Gln Gln Glu Ala Ala Gln Gln Ala Ala Thr Ala		
[0749]	450	455	460
[0750]	Thr Ala Ser Glu His Ser Arg Glu Pro Ser Ala Ala Gly Arg Leu Ser		
[0751]	465	470	475
[0752]	Asp Ser Ser Ser Glu Ala Ser Lys Leu Ser Ser Lys Ser Ala Lys Glu		
[0753]	485	490	495
[0754]	Arg Arg Asn Arg Arg Lys Lys Arg Lys Gln Lys Glu Gln Ser Gly Gly		
[0755]	500	505	510

[0756]	Glu Glu Lys Asp Glu Asp Glu Phe Gln Lys Ser Glu Ser Glu Asp Ser		
[0757]	515	520	525
[0758]	Ile Arg Arg Lys Gly Phe Arg Phe Ser Ile Glu Gly Asn Arg Leu Thr		
[0759]	530	535	540
[0760]	Tyr Glu Lys Arg Tyr Ser Ser Pro His Gln Ser Leu Leu Ser Ile Arg		
[0761]	545	550	555
[0762]	Gly Ser Leu Phe Ser Pro Arg Arg Asn Ser Arg Thr Ser Leu Phe Ser		
[0763]	565	570	575
[0764]	Phe Arg Gly Arg Ala Lys Asp Val Gly Ser Glu Asn Asp Phe Ala Asp		
[0765]	580	585	590
[0766]	Asp Glu His Ser Thr Phe Glu Asp Asn Glu Ser Arg Arg Asp Ser Leu		
[0767]	595	600	605
[0768]	Phe Val Pro Arg Arg His Gly Glu Arg Arg Asn Ser Asn Leu Ser Gln		
[0769]	610	615	620
[0770]	Thr Ser Arg Ser Ser Arg Met Leu Ala Val Phe Pro Ala Asn Gly Lys		
[0771]	625	630	635
[0772]	Met His Ser Thr Val Asp Cys Asn Gly Val Val Ser Leu Val Gly Gly		
[0773]	645	650	655
[0774]	Pro Ser Val Pro Thr Ser Pro Val Gly Gln Leu Leu Pro Glu Val Ile		
[0775]	660	665	670
[0776]	Ile Asp Lys Pro Ala Thr Asp Asp Asn Gly Thr Thr Glu Thr Glu		
[0777]	675	680	685
[0778]	Met Arg Lys Arg Arg Ser Ser Ser Phe His Val Ser Met Asp Phe Leu		
[0779]	690	695	700
[0780]	Glu Asp Pro Ser Gln Arg Gln Arg Ala Met Ser Ile Ala Ser Ile Leu		
[0781]	705	710	715
[0782]	Thr Asn Thr Val Glu Glu Leu Glu Glu Ser Arg Gln Lys Cys Pro Pro		
[0783]	725	730	735
[0784]	Cys Trp Tyr Lys Phe Ser Asn Ile Phe Leu Ile Trp Asp Cys Ser Pro		
[0785]	740	745	750
[0786]	Tyr Trp Leu Lys Val Lys His Val Val Asn Leu Val Val Met Asp Pro		
[0787]	755	760	765
[0788]	Phe Val Asp Leu Ala Ile Thr Ile Cys Ile Val Leu Asn Thr Leu Phe		
[0789]	770	775	780
[0790]	Met Ala Met Glu His Tyr Pro Met Thr Asp His Phe Asn Asn Val Leu		
[0791]	785	790	795
[0792]	Thr Val Gly Asn Leu Val Phe Thr Gly Ile Phe Thr Ala Glu Met Phe		
[0793]	805	810	815
[0794]	Leu Lys Ile Ile Ala Met Asp Pro Tyr Tyr Tyr Phe Gln Glu Gly Trp		
[0795]	820	825	830
[0796]	Asn Ile Phe Asp Gly Phe Ile Val Thr Leu Ser Leu Val Glu Leu Gly		
[0797]	835	840	845

[0798]	Leu Ala Asn Val Glu Gly Leu Ser Val Leu Arg Ser Phe Arg Leu Leu		
[0799]	850	855	860
[0800]	Arg Val Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Met Leu Ile		
[0801]	865	870	875
[0802]	Lys Ile Ile Gly Asn Ser Val Gly Ala Leu Gly Asn Leu Thr Leu Val		
[0803]	885	890	895
[0804]	Leu Ala Ile Ile Val Phe Ile Phe Ala Val Val Gly Met Gln Leu Phe		
[0805]	900	905	910
[0806]	Gly Lys Ser Tyr Lys Asp Cys Val Cys Lys Ile Ala Ser Asp Cys Gln		
[0807]	915	920	925
[0808]	Leu Pro Arg Trp His Met Asn Asp Phe Phe His Ser Phe Leu Ile Val		
[0809]	930	935	940
[0810]	Phe Arg Val Leu Cys Gly Glu Trp Ile Glu Thr Met Trp Asp Cys Met		
[0811]	945	950	955
[0812]	Glu Val Ala Gly Gln Ala Met Cys Leu Thr Val Phe Met Met Val Met		
[0813]	965	970	975
[0814]	Val Ile Gly Asn Leu Val Val Leu Asn Leu Phe Leu Ala Leu Leu Leu		
[0815]	980	985	990
[0816]	Ser Ser Phe Ser Ala Asp Asn Leu Ala Ala Thr Asp Asp Asp Asn Glu		
[0817]	995	1000	1005
[0818]	Met Asn Asn Leu Gln Ile Ala Val Asp Arg Met His Lys Gly Val		
[0819]	1010	1015	1020
[0820]	Ala Tyr Val Lys Arg Lys Ile Tyr Glu Phe Ile Gln Gln Ser Phe		
[0821]	1025	1030	1035
[0822]	Ile Arg Lys Gln Lys Ile Leu Asp Glu Ile Lys Pro Leu Asp Asp		
[0823]	1040	1045	1050
[0824]	Leu Asn Asn Lys Lys Asp Ser Cys Met Ser Asn His Thr Ala Glu		
[0825]	1055	1060	1065
[0826]	Ile Gly Lys Asp Leu Asp Tyr Leu Lys Asp Val Asn Gly Thr Thr		
[0827]	1070	1075	1080
[0828]	Ser Gly Ile Gly Thr Gly Ser Ser Val Glu Lys Tyr Ile Ile Asp		
[0829]	1085	1090	1095
[0830]	Glu Ser Asp Tyr Met Ser Phe Ile Asn Asn Pro Ser Leu Thr Val		
[0831]	1100	1105	1110
[0832]	Thr Val Pro Ile Ala Val Gly Glu Ser Asp Phe Glu Asn Leu Asn		
[0833]	1115	1120	1125
[0834]	Thr Glu Asp Phe Ser Ser Glu Ser Asp Leu Glu Glu Ser Lys Glu		
[0835]	1130	1135	1140
[0836]	Lys Leu Asn Glu Ser Ser Ser Ser Glu Gly Ser Thr Val Asp		
[0837]	1145	1150	1155
[0838]	Ile Gly Ala Pro Val Glu Glu Gln Pro Val Val Glu Pro Glu Glu		
[0839]	1160	1165	1170

[0840]	Thr Leu Glu Pro Glu Ala Cys Phe Thr Glu Gly Cys Val Gln Arg		
[0841]	1175	1180	1185
[0842]	Phe Lys Cys Cys Gln Ile Asn Val Glu Glu Gly Arg Gly Lys Gln		
[0843]	1190	1195	1200
[0844]	Trp Trp Asn Leu Arg Arg Thr Cys Phe Arg Ile Val Glu His Asn		
[0845]	1205	1210	1215
[0846]	Trp Phe Glu Thr Phe Ile Val Phe Met Ile Leu Leu Ser Ser Gly		
[0847]	1220	1225	1230
[0848]	Ala Leu Ala Phe Glu Asp Ile Tyr Ile Asp Gln Arg Lys Thr Ile		
[0849]	1235	1240	1245
[0850]	Lys Thr Met Leu Glu Tyr Ala Asp Lys Val Phe Thr Tyr Ile Phe		
[0851]	1250	1255	1260
[0852]	Ile Leu Glu Met Leu Leu Lys Trp Val Ala Tyr Gly Tyr Gln Thr		
[0853]	1265	1270	1275
[0854]	Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu Ile Val Asp		
[0855]	1280	1285	1290
[0856]	Val Ser Leu Val Ser Leu Thr Ala Asn Ala Leu Gly Tyr Ser Glu		
[0857]	1295	1300	1305
[0858]	Leu Gly Ala Ile Lys Ser Leu Arg Thr Leu Arg Ala Leu Arg Pro		
[0859]	1310	1315	1320
[0860]	Leu Arg Ala Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn		
[0861]	1325	1330	1335
[0862]	Ala Leu Leu Gly Ala Ile Pro Ser Ile Met Asn Val Leu Leu Val		
[0863]	1340	1345	1350
[0864]	Cys Leu Ile Phe Trp Leu Ile Phe Ser Ile Met Gly Val Asn Leu		
[0865]	1355	1360	1365
[0866]	Phe Ala Gly Lys Phe Tyr His Cys Ile Asn Thr Thr Thr Gly Asp		
[0867]	1370	1375	1380
[0868]	Arg Phe Asp Ile Glu Asp Val Asn Asn His Thr Asp Cys Leu Lys		
[0869]	1385	1390	1395
[0870]	Leu Ile Glu Arg Asn Glu Thr Ala Arg Trp Lys Asn Val Lys Val		
[0871]	1400	1405	1410
[0872]	Asn Phe Asp Asn Val Gly Phe Gly Tyr Leu Ser Leu Leu Gln Val		
[0873]	1415	1420	1425
[0874]	Ala Thr Phe Lys Gly Trp Met Asp Ile Met Tyr Ala Ala Val Asp		
[0875]	1430	1435	1440
[0876]	Ser Arg Asn Val Glu Leu Gln Pro Lys Tyr Glu Glu Ser Leu Tyr		
[0877]	1445	1450	1455
[0878]	Met Tyr Leu Tyr Phe Val Ile Phe Ile Ile Phe Gly Ser Phe Phe		
[0879]	1460	1465	1470
[0880]	Thr Leu Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln		
[0881]	1475	1480	1485

[0882]	Gln Lys Lys Lys Phe Gly Gly Gln Asp Ile Phe Met Thr Glu Glu		
[0883]	1490	1495	1500
[0884]	Gln Lys Lys Tyr Tyr Asn Ala Met Lys Lys Leu Gly Ser Lys Lys		
[0885]	1505	1510	1515
[0886]	Pro Gln Lys Pro Ile Pro Arg Pro Gly Asn Lys Phe Gln Gly Met		
[0887]	1520	1525	1530
[0888]	Val Phe Asp Phe Val Thr Arg Gln Val Phe Asp Ile Ser Ile Met		
[0889]	1535	1540	1545
[0890]	Ile Leu Ile Cys Leu Asn Met Val Thr Met Met Val Glu Thr Asp		
[0891]	1550	1555	1560
[0892]	Asp Gln Ser Glu Tyr Val Thr Thr Ile Leu Ser Arg Ile Asn Leu		
[0893]	1565	1570	1575
[0894]	Val Phe Ile Val Leu Phe Thr Gly Glu Cys Val Leu Lys Leu Ile		
[0895]	1580	1585	1590
[0896]	Ser Leu Arg His Tyr Tyr Phe Thr Ile Gly Trp Asn Ile Phe Asp		
[0897]	1595	1600	1605
[0898]	Phe Val Val Val Ile Leu Ser Ile Val Gly Met Phe Leu Ala Glu		
[0899]	1610	1615	1620
[0900]	Leu Ile Glu Lys Tyr Phe Val Ser Pro Thr Leu Phe Arg Val Ile		
[0901]	1625	1630	1635
[0902]	Arg Leu Ala Arg Ile Gly Arg Ile Leu Arg Leu Ile Lys Gly Ala		
[0903]	1640	1645	1650
[0904]	Lys Gly Ile Arg Thr Leu Leu Phe Ala Leu Met Met Ser Leu Pro		
[0905]	1655	1660	1665
[0906]	Ala Leu Phe Asn Ile Gly Leu Leu Leu Phe Leu Val Met Phe Ile		
[0907]	1670	1675	1680
[0908]	Tyr Ala Ile Phe Gly Met Ser Asn Phe Ala Tyr Val Lys Arg Glu		
[0909]	1685	1690	1695
[0910]	Val Gly Ile Asp Asp Met Phe Asn Phe Glu Thr Phe Gly Asn Ser		
[0911]	1700	1705	1710
[0912]	Met Ile Cys Leu Phe Gln Ile Thr Thr Ser Ala Gly Trp Asp Gly		
[0913]	1715	1720	1725
[0914]	Leu Leu Ala Pro Ile Leu Asn Ser Lys Pro Pro Asp Cys Asp Pro		
[0915]	1730	1735	1740
[0916]	Asn Lys Val Asn Pro Gly Ser Ser Val Lys Gly Asp Cys Gly Asn		
[0917]	1745	1750	1755
[0918]	Pro Ser Val Gly Ile Phe Phe Val Ser Tyr Ile Ile Ile Ser		
[0919]	1760	1765	1770
[0920]	Phe Leu Val Val Val Asn Met Tyr Ile Ala Val Ile Leu Glu Asn		
[0921]	1775	1780	1785
[0922]	Phe Ser Val Ala Thr Glu Glu Ser Ala Glu Pro Leu Ser Glu Asp		
[0923]	1790	1795	1800

[0924]	Asp Phe Glu Met Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Asp		
[0925]	1805	1810	1815
[0926]	Ala Thr Gln Phe Met Glu Phe Glu Lys Leu Ser Gln Phe Ala Ala		
[0927]	1820	1825	1830
[0928]	Ala Leu Glu Pro Pro Leu Asn Leu Pro Gln Pro Asn Lys Leu Gln		
[0929]	1835	1840	1845
[0930]	Leu Ile Ala Met Asp Leu Pro Met Val Ser Gly Asp Arg Ile His		
[0931]	1850	1855	1860
[0932]	Cys Leu Asp Ile Leu Phe Ala Phe Thr Lys Arg Val Leu Gly Glu		
[0933]	1865	1870	1875
[0934]	Ser Gly Glu Met Asp Ala Leu Arg Ile Gln Met Glu Glu Arg Phe		
[0935]	1880	1885	1890
[0936]	Met Ala Ser Asn Pro Ser Lys Val Ser Tyr Gln Pro Ile Thr Thr		
[0937]	1895	1900	1905
[0938]	Thr Leu Lys Arg Lys Gln Glu Glu Val Ser Ala Val Ile Ile Gln		
[0939]	1910	1915	1920
[0940]	Arg Ala Tyr Arg Arg His Leu Leu Lys Arg Thr Val Lys Gln Ala		
[0941]	1925	1930	1935
[0942]	Ser Phe Thr Tyr Asn Lys Asn Lys Ile Lys Gly Gly Ala Asn Leu		
[0943]	1940	1945	1950
[0944]	Leu Ile Lys Glu Asp Met Ile Ile Asp Arg Ile Asn Glu Asn Ser		
[0945]	1955	1960	1965
[0946]	Ile Thr Glu Lys Thr Asp Leu Thr Met Ser Thr Ala Ala Cys Pro		
[0947]	1970	1975	1980
[0948]	Pro Ser Tyr Asp Arg Val Thr Lys Pro Ile Val Glu Lys His Glu		
[0949]	1985	1990	1995
[0950]	Gln Glu Gly Lys Asp Glu Lys Ala Lys Gly Lys		
[0951]	2000	2005	
[0952]	<210> 40		
[0953]	<211> 603		
[0954]	<212> PRT		
[0955]	<213> 未知		
[0956]	<220>		
[0957]	<223> 对未知的描述:		
[0958]	STXBP1序列		
[0959]	<400> 40		
[0960]	Met Ala Pro Ile Gly Leu Lys Ala Val Val Gly Glu Lys Ile Met His		
[0961]	1	5	10
[0962]	15		
[0962]	Asp Val Ile Lys Lys Val Lys Lys Gly Glu Trp Lys Val Leu Val		
[0963]	20	25	30
[0964]	Val Asp Gln Leu Ser Met Arg Met Leu Ser Ser Cys Cys Lys Met Thr		
[0965]	35	40	45

[0966]	Asp Ile Met Thr Glu Gly Ile Thr Ile Val Glu Asp Ile Asn Lys Arg		
[0967]	50	55	60
[0968]	Arg Glu Pro Leu Pro Ser Leu Glu Ala Val Tyr Leu Ile Thr Pro Ser		
[0969]	65	70	75
[0970]	Glu Lys Ser Val His Ser Leu Ile Ser Asp Phe Lys Asp Pro Pro Thr		80
[0971]	85	90	95
[0972]	Ala Lys Tyr Arg Ala Ala His Val Phe Phe Thr Asp Ser Cys Pro Asp		
[0973]	100	105	110
[0974]	Ala Leu Phe Asn Glu Leu Val Lys Ser Arg Ala Ala Lys Val Ile Lys		
[0975]	115	120	125
[0976]	Thr Leu Thr Glu Ile Asn Ile Ala Phe Leu Pro Tyr Glu Ser Gln Val		
[0977]	130	135	140
[0978]	Tyr Ser Leu Asp Ser Ala Asp Ser Phe Gln Ser Phe Tyr Ser Pro His		
[0979]	145	150	155
[0980]	Lys Ala Gln Met Lys Asn Pro Ile Leu Glu Arg Leu Ala Glu Gln Ile		160
[0981]	165	170	175
[0982]	Ala Thr Leu Cys Ala Thr Leu Lys Glu Tyr Pro Ala Val Arg Tyr Arg		
[0983]	180	185	190
[0984]	Gly Glu Tyr Lys Asp Asn Ala Leu Leu Ala Gln Leu Ile Gln Asp Lys		
[0985]	195	200	205
[0986]	Leu Asp Ala Tyr Lys Ala Asp Asp Pro Thr Met Gly Glu Gly Pro Asp		
[0987]	210	215	220
[0988]	Lys Ala Arg Ser Gln Leu Leu Ile Leu Asp Arg Gly Phe Asp Pro Ser		
[0989]	225	230	235
[0990]	Ser Pro Val Leu His Glu Leu Thr Phe Gln Ala Met Ser Tyr Asp Leu		240
[0991]	245	250	255
[0992]	Leu Pro Ile Glu Asn Asp Val Tyr Lys Tyr Glu Thr Ser Gly Ile Gly		
[0993]	260	265	270
[0994]	Glu Ala Arg Val Lys Glu Val Leu Leu Asp Glu Asp Asp Asp Leu Trp		
[0995]	275	280	285
[0996]	Ile Ala Leu Arg His Lys His Ile Ala Glu Val Ser Gln Glu Val Thr		
[0997]	290	295	300
[0998]	Arg Ser Leu Lys Asp Phe Ser Ser Ser Lys Arg Met Asn Thr Gly Glu		
[0999]	305	310	315
[1000]	Lys Thr Thr Met Arg Asp Leu Ser Gln Met Leu Lys Lys Met Pro Gln		320
[1001]	325	330	335
[1002]	Tyr Gln Lys Glu Leu Ser Lys Tyr Ser Thr His Leu His Leu Ala Glu		
[1003]	340	345	350
[1004]	Asp Cys Met Lys His Tyr Gln Gly Thr Val Asp Lys Leu Cys Arg Val		
[1005]	355	360	365
[1006]	Glu Gln Asp Leu Ala Met Gly Thr Asp Ala Glu Gly Glu Lys Ile Lys		
[1007]	370	375	380

[1008]	Asp Pro Met Arg Ala Ile Val Pro Ile Leu Leu Asp Ala Asn Val Ser			
[1009]	385	390	395	400
[1010]	Thr Tyr Asp Lys Ile Arg Ile Ile Leu Leu Tyr Ile Phe Leu Lys Asn			
[1011]	405		410	415
[1012]	Gly Ile Thr Glu Glu Asn Leu Asn Lys Leu Ile Gln His Ala Gln Ile			
[1013]	420	425	430	
[1014]	Pro Pro Glu Asp Ser Glu Ile Ile Thr Asn Met Ala His Leu Gly Val			
[1015]	435	440	445	
[1016]	Pro Ile Val Thr Asp Ser Thr Leu Arg Arg Arg Ser Lys Pro Glu Arg			
[1017]	450	455	460	
[1018]	Lys Glu Arg Ile Ser Glu Gln Thr Tyr Gln Leu Ser Arg Trp Thr Pro			
[1019]	465	470	475	480
[1020]	Ile Ile Lys Asp Ile Met Glu Asp Thr Ile Glu Asp Lys Leu Asp Thr			
[1021]	485	490	495	
[1022]	Lys His Tyr Pro Tyr Ile Ser Thr Arg Ser Ser Ala Ser Phe Ser Thr			
[1023]	500	505	510	
[1024]	Thr Ala Val Ser Ala Arg Tyr Gly His Trp His Lys Asn Lys Ala Pro			
[1025]	515	520	525	
[1026]	Gly Glu Tyr Arg Ser Gly Pro Arg Leu Ile Ile Phe Ile Leu Gly Gly			
[1027]	530	535	540	
[1028]	Val Ser Leu Asn Glu Met Arg Cys Ala Tyr Glu Val Thr Gln Ala Asn			
[1029]	545	550	555	560
[1030]	Gly Lys Trp Glu Val Leu Ile Gly Ser Thr His Ile Leu Thr Pro Thr			
[1031]	565	570	575	
[1032]	Lys Phe Leu Met Asp Leu Arg His Pro Asp Phe Arg Glu Ser Ser Arg			
[1033]	580	585	590	
[1034]	Val Ser Phe Glu Asp Gln Ala Pro Thr Met Glu			
[1035]	595	600		
[1036]	<210> 41			
[1037]	<211> 511			
[1038]	<212> PRT			
[1039]	<213> 未知			
[1040]	<220>			
[1041]	<223> 对未知的描述:			
[1042]	Kv3.1序列			
[1043]	<400> 41			
[1044]	Met Gly Gln Gly Asp Glu Ser Glu Arg Ile Val Ile Asn Val Gly Gly			
[1045]	1	5	10	15
[1046]	Thr Arg His Gln Thr Tyr Arg Ser Thr Leu Arg Thr Leu Pro Gly Thr			
[1047]	20	25	30	
[1048]	Arg Leu Ala Trp Leu Ala Glu Pro Asp Ala His Ser His Phe Asp Tyr			
[1049]	35	40	45	

[1050]	Asp Pro Arg Ala Asp Glu Phe Phe Phe Asp Arg His Pro Gly Val Phe		
[1051]	50	55	60
[1052]	Ala His Ile Leu Asn Tyr Tyr Arg Thr Gly Lys Leu His Cys Pro Ala		
[1053]	65	70	75
[1054]	Asp Val Cys Gly Pro Leu Tyr Glu Glu Leu Ala Phe Trp Gly Ile		
[1055]	85	90	95
[1056]	Asp Glu Thr Asp Val Glu Pro Cys Cys Trp Met Thr Tyr Arg Gln His		
[1057]	100	105	110
[1058]	Arg Asp Ala Glu Glu Ala Leu Asp Ser Phe Gly Gly Ala Pro Leu Asp		
[1059]	115	120	125
[1060]	Asn Ser Ala Asp Asp Ala Asp Gly Pro Gly Asp Ser Gly Asp		
[1061]	130	135	140
[1062]	Gly Glu Asp Glu Leu Glu Met Thr Lys Arg Leu Ala Leu Ser Asp Ser		
[1063]	145	150	155
[1064]	Pro Asp Gly Arg Pro Gly Gly Phe Trp Arg Arg Trp Gln Pro Arg Ile		
[1065]	165	170	175
[1066]	Trp Ala Leu Phe Glu Asp Pro Tyr Ser Ser Arg Tyr Ala Arg Tyr Val		
[1067]	180	185	190
[1068]	Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe Cys		
[1069]	195	200	205
[1070]	Leu Glu Thr His Glu Arg Phe Asn Pro Ile Val Asn Lys Thr Glu Ile		
[1071]	210	215	220
[1072]	Glu Asn Val Arg Asn Gly Thr Gln Val Arg Tyr Tyr Arg Glu Ala Glu		
[1073]	225	230	235
[1074]	Thr Glu Ala Phe Leu Thr Tyr Ile Glu Gly Val Cys Val Val Trp Phe		
[1075]	245	250	255
[1076]	Thr Phe Glu Phe Leu Met Arg Val Ile Phe Cys Pro Asn Lys Val Glu		
[1077]	260	265	270
[1078]	Phe Ile Lys Asn Ser Leu Asn Ile Ile Asp Phe Val Ala Ile Leu Pro		
[1079]	275	280	285
[1080]	Phe Tyr Leu Glu Val Gly Leu Ser Gly Leu Ser Ser Lys Ala Ala Lys		
[1081]	290	295	300
[1082]	Asp Val Leu Gly Phe Leu Arg Val Val Arg Phe Val Arg Ile Leu Arg		
[1083]	305	310	315
[1084]	Ile Phe Lys Leu Thr Arg His Phe Val Gly Leu Arg Val Leu Gly His		
[1085]	325	330	335
[1086]	Thr Leu Arg Ala Ser Thr Asn Glu Phe Leu Leu Leu Ile Ile Phe Leu		
[1087]	340	345	350
[1088]	Ala Leu Gly Val Leu Ile Phe Ala Thr Met Ile Tyr Tyr Ala Glu Arg		
[1089]	355	360	365
[1090]	Ile Gly Ala Gln Pro Asn Asp Pro Ser Ala Ser Glu His Thr His Phe		
[1091]	370	375	380

[1092]	Lys Asn Ile Pro Ile Gly Phe Trp Trp Ala Val Val Thr Met Thr Thr			
[1093]	385	390	395	400
[1094]	Leu Gly Tyr Gly Asp Met Tyr Pro Gln Thr Trp Ser Gly Met Leu Val			
[1095]	405	410	415	
[1096]	Gly Ala Leu Cys Ala Leu Ala Gly Val Leu Thr Ile Ala Met Pro Val			
[1097]	420	425	430	
[1098]	Pro Val Ile Val Asn Asn Phe Gly Met Tyr Tyr Ser Leu Ala Met Ala			
[1099]	435	440	445	
[1100]	Lys Gln Lys Leu Pro Lys Lys Lys Lys His Ile Pro Arg Pro Pro			
[1101]	450	455	460	
[1102]	Gln Leu Gly Ser Pro Asn Tyr Cys Lys Ser Val Val Asn Ser Pro His			
[1103]	465	470	475	480
[1104]	His Ser Thr Gln Ser Asp Thr Cys Pro Leu Ala Gln Glu Glu Ile Leu			
[1105]	485	490	495	
[1106]	Glu Ile Asn Arg Ala Gly Arg Lys Pro Leu Arg Gly Met Ser Ile			
[1107]	500	505	510	
[1108]	<210> 42			
[1109]	<211> 639			
[1110]	<212> PRT			
[1111]	<213> 未知			
[1112]	<220>			
[1113]	<223> 对未知的描述:			
[1114]	Kv3.2序列			
[1115]	<400> 42			
[1116]	Met Gly Lys Ile Glu Ser Asn Glu Arg Val Ile Leu Asn Val Gly Gly			
[1117]	1	5	10	15
[1118]	Thr Arg His Glu Thr Tyr Arg Ser Thr Leu Lys Thr Leu Pro Gly Thr			
[1119]	20	25	30	
[1120]	Arg Leu Ala Leu Leu Ala Ser Ser Glu Pro Gln Gly Asp Cys Leu Thr			
[1121]	35	40	45	
[1122]	Ala Ala Gly Asp Lys Leu Gln Pro Leu Pro Pro Leu Ser Pro Pro			
[1123]	50	55	60	
[1124]	Pro Arg Pro Pro Pro Leu Ser Pro Val Pro Ser Gly Cys Phe Glu Gly			
[1125]	65	70	75	80
[1126]	Gly Ala Gly Asn Cys Ser Ser His Gly Gly Asn Gly Gly Asn Gly Gly			
[1127]	85	90	95	
[1128]	Ser Asp His Pro Gly Gly Arg Glu Phe Phe Asp Arg His Pro			
[1129]	100	105	110	
[1130]	Gly Val Phe Ala Tyr Val Leu Asn Tyr Tyr Arg Thr Gly Lys Leu His			
[1131]	115	120	125	
[1132]	Cys Pro Ala Asp Val Cys Gly Pro Leu Phe Glu Glu Leu Ala Phe			
[1133]	130	135	140	

[1134]	Trp Gly Ile Asp Glu Thr Asp Val Glu Pro Cys Cys Trp Met Thr Tyr			
[1135]	145	150	155	160
[1136]	Arg Gln His Arg Asp Ala Glu Glu Ala Leu Asp Ile Phe Glu Thr Pro			
[1137]	165	170	175	
[1138]	Asp Leu Ile Gly Gly Asp Pro Gly Asp Asp Glu Asp Leu Ala Ala Lys			
[1139]	180	185	190	
[1140]	Arg Leu Gly Ile Glu Asp Ala Ala Gly Leu Gly Gly Pro Asp Gly Lys			
[1141]	195	200	205	
[1142]	Ser Gly Arg Trp Arg Lys Leu Gln Pro Arg Met Trp Ala Leu Phe Glu			
[1143]	210	215	220	
[1144]	Asp Pro Tyr Ser Ser Arg Ala Ala Arg Phe Ile Ala Phe Ala Ser Leu			
[1145]	225	230	235	240
[1146]	Phe Phe Ile Leu Val Ser Ile Thr Thr Phe Cys Leu Glu Thr His Glu			
[1147]	245	250	255	
[1148]	Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val Ile Asn Gly Thr			
[1149]	260	265	270	
[1150]	Ser Pro Val Leu Gln Tyr Glu Ile Glu Thr Asp Pro Ala Leu Thr Tyr			
[1151]	275	280	285	
[1152]	Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg			
[1153]	290	295	300	
[1154]	Ile Val Phe Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu Leu Asn			
[1155]	305	310	315	320
[1156]	Ile Ile Asp Phe Val Ala Ile Leu Pro Phe Tyr Leu Glu Val Gly Leu			
[1157]	325	330	335	
[1158]	Ser Gly Leu Ser Ser Lys Ala Ala Lys Asp Val Leu Gly Phe Leu Arg			
[1159]	340	345	350	
[1160]	Val Val Arg Phe Val Arg Ile Leu Arg Ile Phe Lys Leu Thr Arg His			
[1161]	355	360	365	
[1162]	Phe Val Gly Leu Arg Val Leu Gly His Thr Leu Arg Ala Ser Thr Asn			
[1163]	370	375	380	
[1164]	Glu Phe Leu Leu Leu Ile Ile Phe Leu Ala Leu Gly Val Leu Ile Phe			
[1165]	385	390	395	400
[1166]	Ala Thr Met Ile Tyr Tyr Ala Glu Arg Val Gly Ala Gln Pro Asn Asp			
[1167]	405	410	415	
[1168]	Pro Ser Ala Ser Glu His Thr Gln Phe Lys Asn Ile Pro Ile Gly Phe			
[1169]	420	425	430	
[1170]	Trp Trp Ala Val Val Thr Met Thr Thr Leu Gly Tyr Gly Asp Met Tyr			
[1171]	435	440	445	
[1172]	Pro Gln Thr Trp Ser Gly Met Leu Val Gly Ala Leu Cys Ala Leu Ala			
[1173]	450	455	460	
[1174]	Gly Val Leu Thr Ile Ala Met Pro Val Pro Val Ile Val Asn Asn Phe			
[1175]	465	470	475	480

[1176]	Gly Met Tyr Tyr Ser Leu Ala Met Ala Lys Gln Lys Leu Pro Arg Lys		
[1177]	485	490	495
[1178]	Arg Lys Lys His Ile Pro Pro Ala Pro Leu Ala Ser Ser Pro Thr Phe		
[1179]	500	505	510
[1180]	Cys Lys Thr Glu Leu Asn Met Ala Cys Asn Ser Thr Gln Ser Asp Thr		
[1181]	515	520	525
[1182]	Cys Leu Gly Lys Glu Asn Arg Leu Leu Glu His Asn Arg Ser Val Leu		
[1183]	530	535	540
[1184]	Ser Gly Asp Asp Ser Thr Gly Ser Glu Pro Pro Leu Ser Pro Pro Glu		
[1185]	545	550	555
[1186]	Arg Leu Pro Ile Arg Arg Ser Ser Thr Arg Asp Lys Asn Arg Arg Gly		
[1187]	565	570	575
[1188]	Glu Thr Cys Phe Leu Leu Thr Thr Gly Asp Tyr Thr Cys Ala Ser Asp		
[1189]	580	585	590
[1190]	Gly Gly Ile Arg Lys Ala Ser Thr Leu Glu Pro Met Glu Ser Thr Ala		
[1191]	595	600	605
[1192]	Gln Thr Lys Gly Asp Thr Arg Pro Glu Ala His Trp Asn Cys Ala His		
[1193]	610	615	620
[1194]	Leu Leu Asn Phe Gly Cys Pro Thr Gly Ser Ser Phe Pro Thr Leu		
[1195]	625	630	635
[1196]	<210> 43		
[1197]	<211> 757		
[1198]	<212> PRT		
[1199]	<213> 未知		
[1200]	<220>		
[1201]	<223> 对未知的描述:		
[1202]	Kv3.3序列		
[1203]	<400> 43		
[1204]	Met Leu Ser Ser Val Cys Val Ser Ser Phe Arg Gly Arg Gln Gly Ala		
[1205]	1	5	10
[1206]	15		
[1207]	Ser Lys Gln Gln Pro Ala Pro Pro Pro Gln Pro Pro Glu Ser Pro Pro		
[1208]	20	25	30
[1209]	Pro Pro Pro Leu Pro Pro Gln Gln Gln Gln Pro Ala Gln Pro Gly Pro		
[1210]	35	40	45
[1211]	Ala Ala Ser Pro Ala Gly Pro Pro Ala Pro Arg Gly Pro Gly Asp Arg		
[1212]	50	55	60
[1213]	60		
[1214]	Arg Ala Glu Pro Cys Pro Gly Leu Pro Ala Ala Met Gly Arg His		
[1215]	65	70	75
[1216]	75		
[1217]	Gly Gly Gly Gly Gly Asp Ser Gly Lys Ile Val Ile Asn Val Gly Gly		
[1218]	80		
[1219]	85	90	95
[1220]	95		
[1221]	Val Arg His Glu Thr Tyr Arg Ser Thr Leu Arg Thr Leu Pro Gly Thr		
[1222]	100	105	110

[1218]	Arg Leu Ala Gly Leu Thr Glu Pro Glu Ala Ala Ala Arg Phe Asp Tyr		
[1219]	115	120	125
[1220]	Asp Pro Gly Ala Asp Glu Phe Phe Phe Asp Arg His Pro Gly Val Phe		
[1221]	130	135	140
[1222]	Ala Tyr Val Leu Asn Tyr Tyr Arg Thr Gly Lys Leu His Cys Pro Ala		
[1223]	145	150	155
[1224]	Asp Val Cys Gly Pro Leu Phe Glu Glu Leu Gly Phe Trp Gly Ile		
[1225]	165	170	175
[1226]	Asp Glu Thr Asp Val Glu Ala Cys Cys Trp Met Thr Tyr Arg Gln His		
[1227]	180	185	190
[1228]	Arg Asp Ala Glu Glu Ala Leu Asp Ser Phe Glu Ala Pro Asp Pro Ala		
[1229]	195	200	205
[1230]	Gly Ala Ala Asn Ala Ala Asn Ala Gly Ala His Asp Gly Gly Leu		
[1231]	210	215	220
[1232]	Asp Asp Glu Ala Gly Ala Gly Gly Gly Leu Asp Gly Ala Gly Gly		
[1233]	225	230	235
[1234]	Glu Leu Lys Arg Leu Cys Phe Gln Asp Ala Gly Gly Gly Ala Gly Gly		
[1235]	245	250	255
[1236]	Pro Pro Gly Gly Ala Gly Gly Thr Trp Trp Arg Arg Trp		
[1237]	260	265	270
[1238]	Gln Pro Arg Val Trp Ala Leu Phe Glu Asp Pro Tyr Ser Ser Arg Ala		
[1239]	275	280	285
[1240]	Ala Arg Tyr Val Ala Phe Ala Ser Leu Phe Phe Ile Leu Ile Ser Ile		
[1241]	290	295	300
[1242]	Thr Thr Phe Cys Leu Glu Thr His Glu Gly Phe Ile His Ile Ser Asn		
[1243]	305	310	315
[1244]	Lys Thr Val Thr Gln Ala Ser Pro Ile Pro Gly Ala Pro Pro Glu Asn		
[1245]	325	330	335
[1246]	Ile Thr Asn Val Glu Val Glu Thr Glu Pro Phe Leu Thr Tyr Val Glu		
[1247]	340	345	350
[1248]	Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Met Arg Ile Thr		
[1249]	355	360	365
[1250]	Phe Cys Pro Asp Lys Val Glu Phe Leu Lys Ser Ser Leu Asn Ile Ile		
[1251]	370	375	380
[1252]	Asp Cys Val Ala Ile Leu Pro Phe Tyr Leu Glu Val Gly Leu Ser Gly		
[1253]	385	390	395
[1254]	Leu Ser Ser Lys Ala Ala Lys Asp Val Leu Gly Phe Leu Arg Val Val		
[1255]	405	410	415
[1256]	Arg Phe Val Arg Ile Leu Arg Ile Phe Lys Leu Thr Arg His Phe Val		
[1257]	420	425	430
[1258]	Gly Leu Arg Val Leu Gly His Thr Leu Arg Ala Ser Thr Asn Glu Phe		
[1259]	435	440	445

[1260]	Leu Leu Leu Ile Ile Phe Leu Ala Leu Gly Val Leu Ile Phe Ala Thr		
[1261]	450	455	460
[1262]	Met Ile Tyr Tyr Ala Glu Arg Ile Gly Ala Asp Pro Asp Asp Ile Leu		
[1263]	465	470	475
[1264]	Gly Ser Asn His Thr Tyr Phe Lys Asn Ile Pro Ile Gly Phe Trp Trp		
[1265]	485	490	495
[1266]	Ala Val Val Thr Met Thr Leu Gly Tyr Gly Asp Met Tyr Pro Lys		
[1267]	500	505	510
[1268]	Thr Trp Ser Gly Met Leu Val Gly Ala Leu Cys Ala Leu Ala Gly Val		
[1269]	515	520	525
[1270]	Leu Thr Ile Ala Met Pro Val Pro Val Ile Val Asn Asn Phe Gly Met		
[1271]	530	535	540
[1272]	Tyr Tyr Ser Leu Ala Met Ala Lys Gln Lys Leu Pro Lys Lys Lys Asn		
[1273]	545	550	555
[1274]	Lys His Ile Pro Arg Pro Pro Gln Pro Gly Ser Pro Asn Tyr Cys Lys		
[1275]	565	570	575
[1276]	Pro Asp Pro Pro Pro Pro Pro His Pro His His Gly Ser Gly		
[1277]	580	585	590
[1278]	Gly Ile Ser Pro Pro Pro Ile Thr Pro Pro Ser Met Gly Val Thr		
[1279]	595	600	605
[1280]	Val Ala Gly Ala Tyr Pro Ala Gly Pro His Thr His Pro Gly Leu Leu		
[1281]	610	615	620
[1282]	Arg Gly Gly Ala Gly Gly Leu Gly Ile Met Gly Leu Pro Pro Leu Pro		
[1283]	625	630	635
[1284]	Ala Pro Gly Glu Pro Cys Pro Leu Ala Gln Glu Glu Val Ile Glu Ile		
[1285]	645	650	655
[1286]	Asn Arg Ala Asp Pro Arg Pro Asn Gly Asp Pro Ala Ala Ala Leu		
[1287]	660	665	670
[1288]	Ala His Glu Asp Cys Pro Ala Ile Asp Gln Pro Ala Met Ser Pro Glu		
[1289]	675	680	685
[1290]	Asp Lys Ser Pro Ile Thr Pro Gly Ser Arg Gly Arg Tyr Ser Arg Asp		
[1291]	690	695	700
[1292]	Arg Ala Cys Phe Leu Leu Thr Asp Tyr Ala Pro Ser Pro Asp Gly Ser		
[1293]	705	710	715
[1294]	Ile Arg Lys Ala Thr Gly Ala Pro Pro Leu Pro Pro Gln Asp Trp Arg		
[1295]	725	730	735
[1296]	Lys Pro Gly Pro Pro Ser Phe Leu Pro Asp Leu Asn Ala Asn Ala Ala		
[1297]	740	745	750
[1298]	Ala Trp Ile Ser Pro		
[1299]	755		

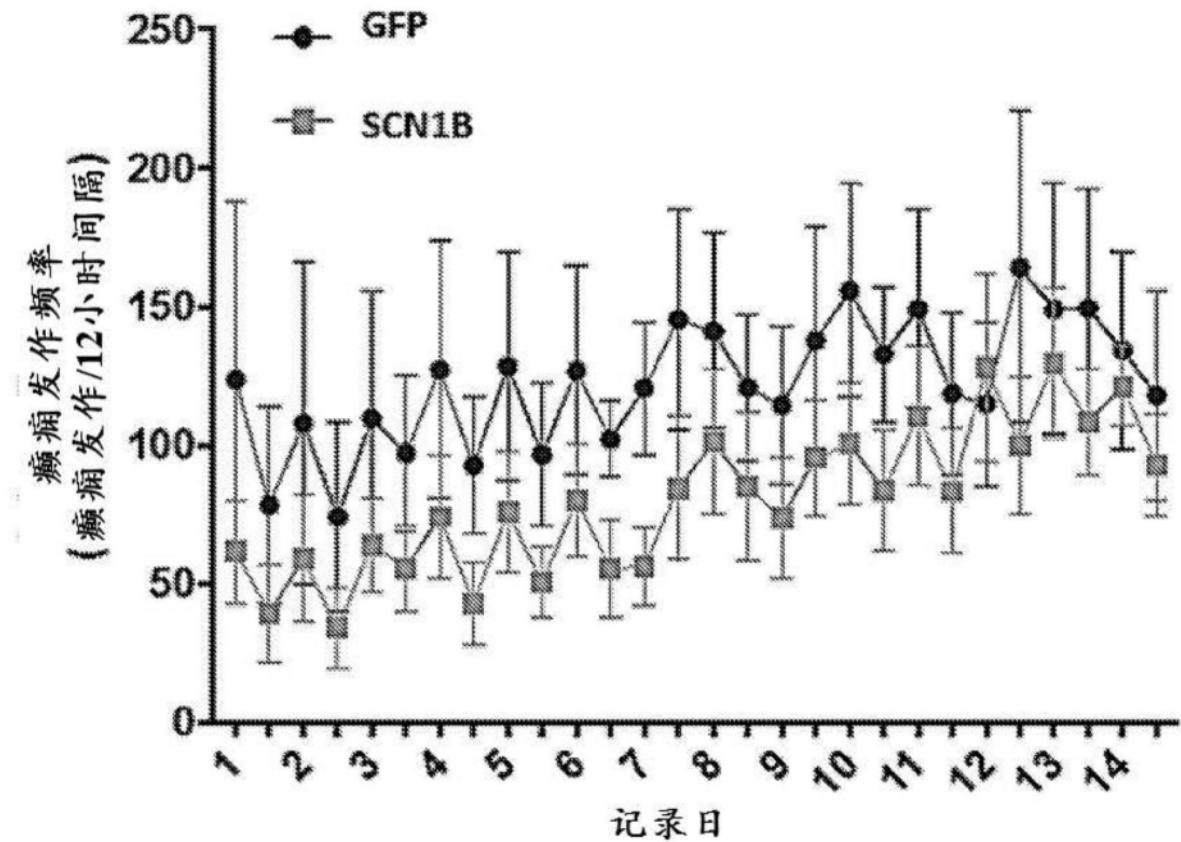


图1

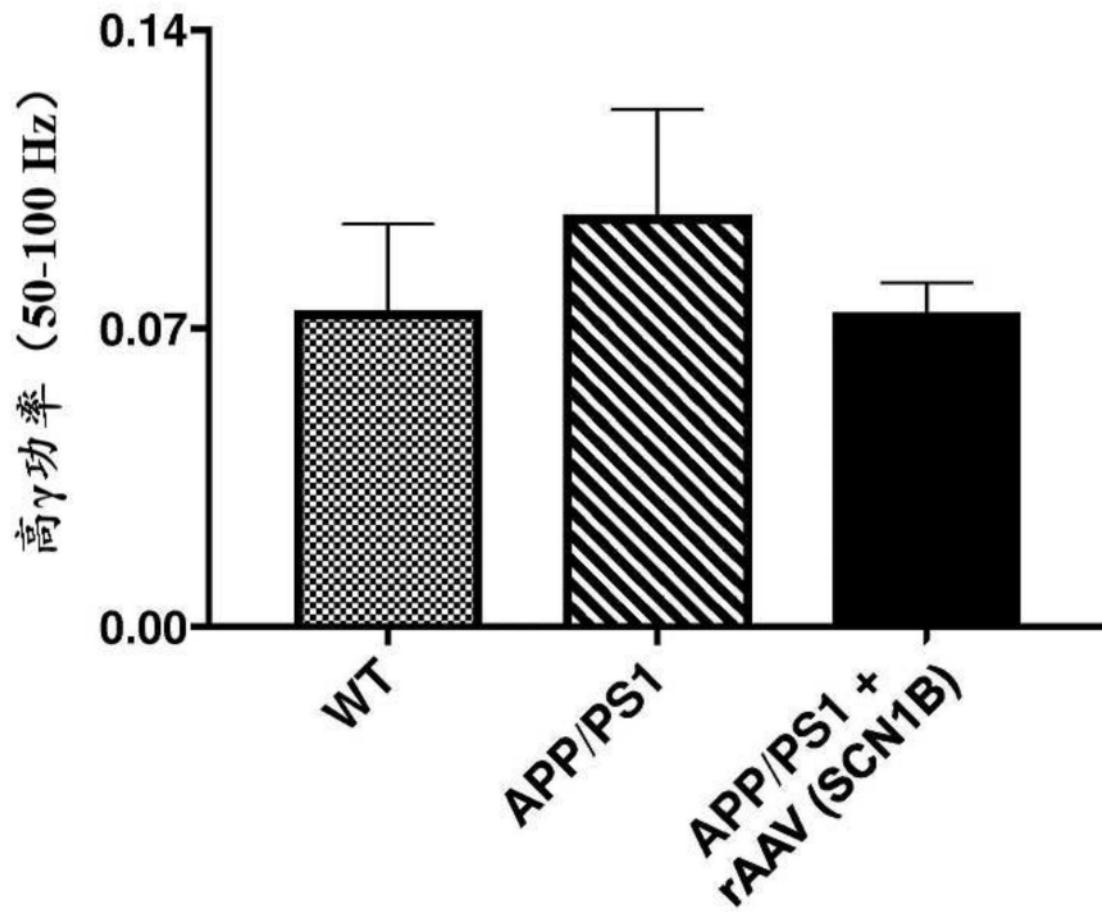


图2

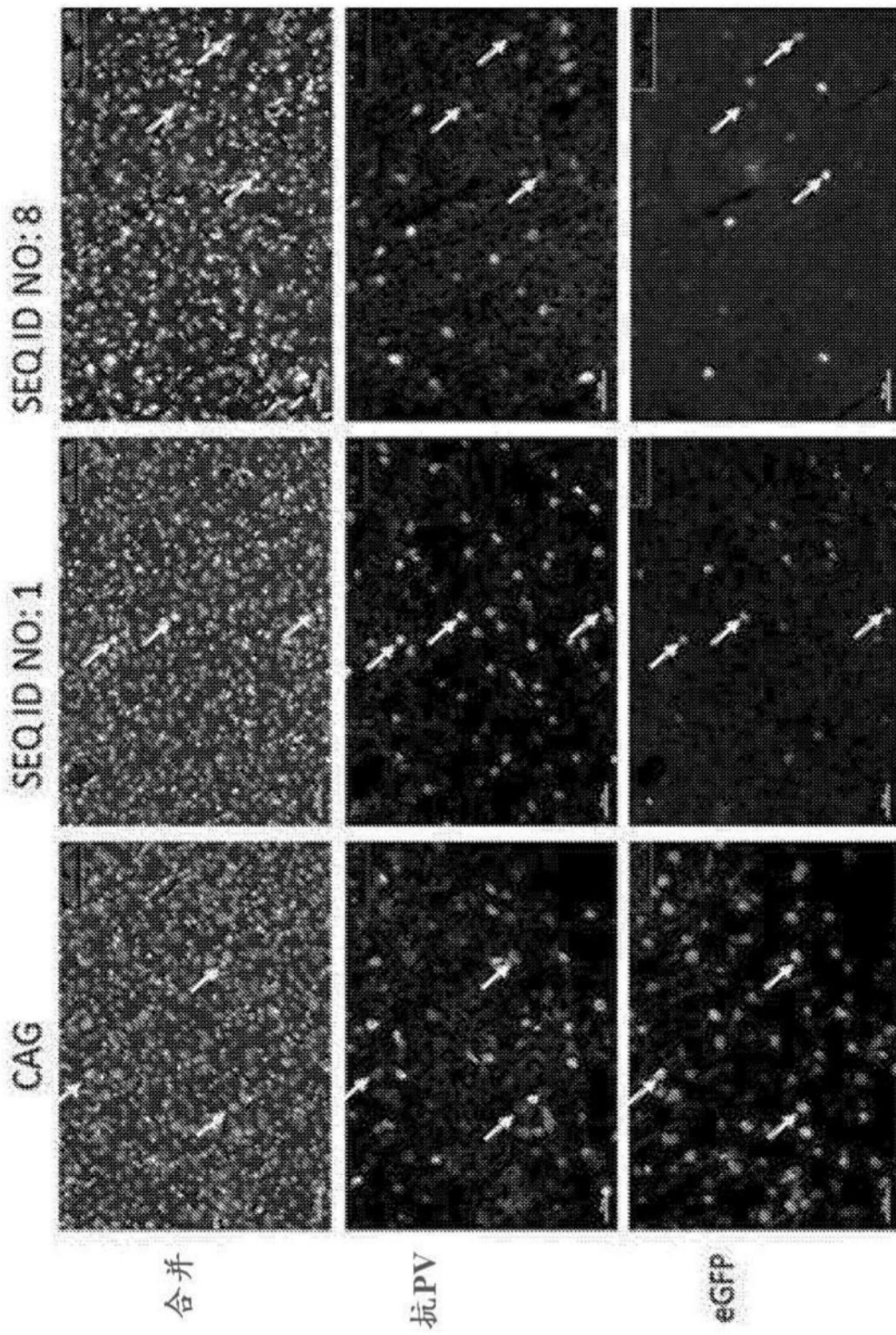


图3A

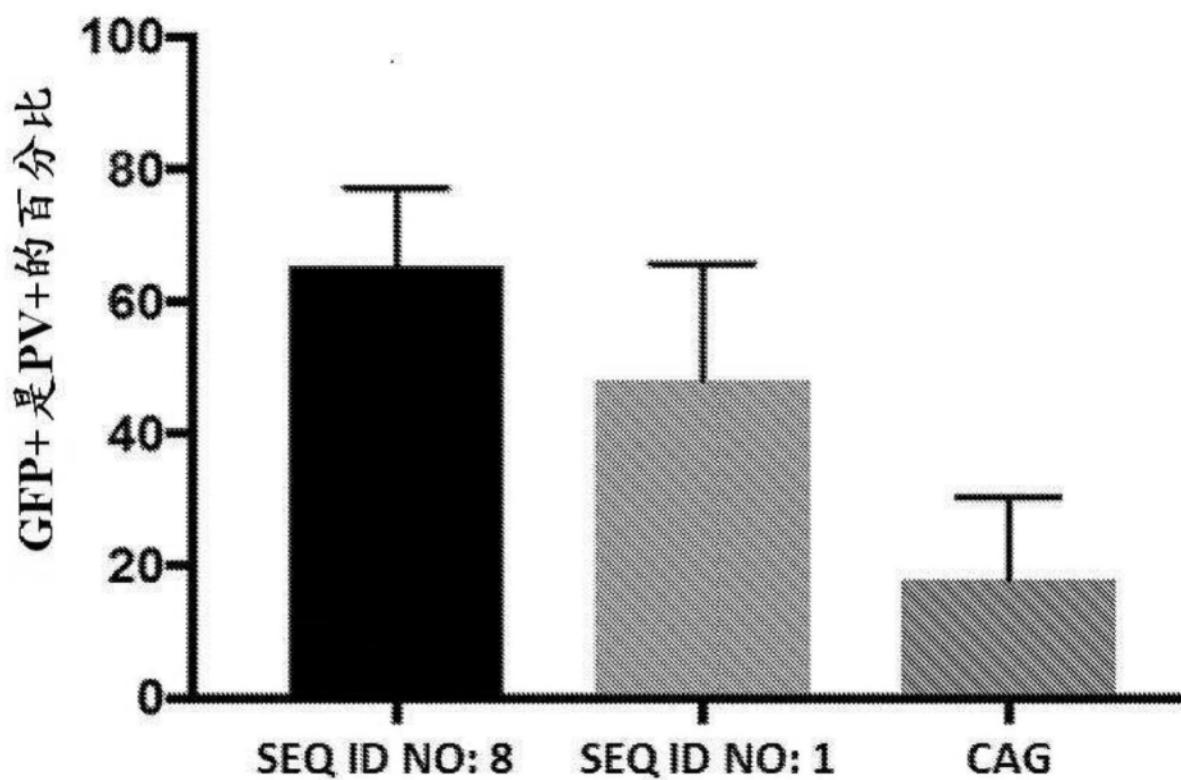


图3B

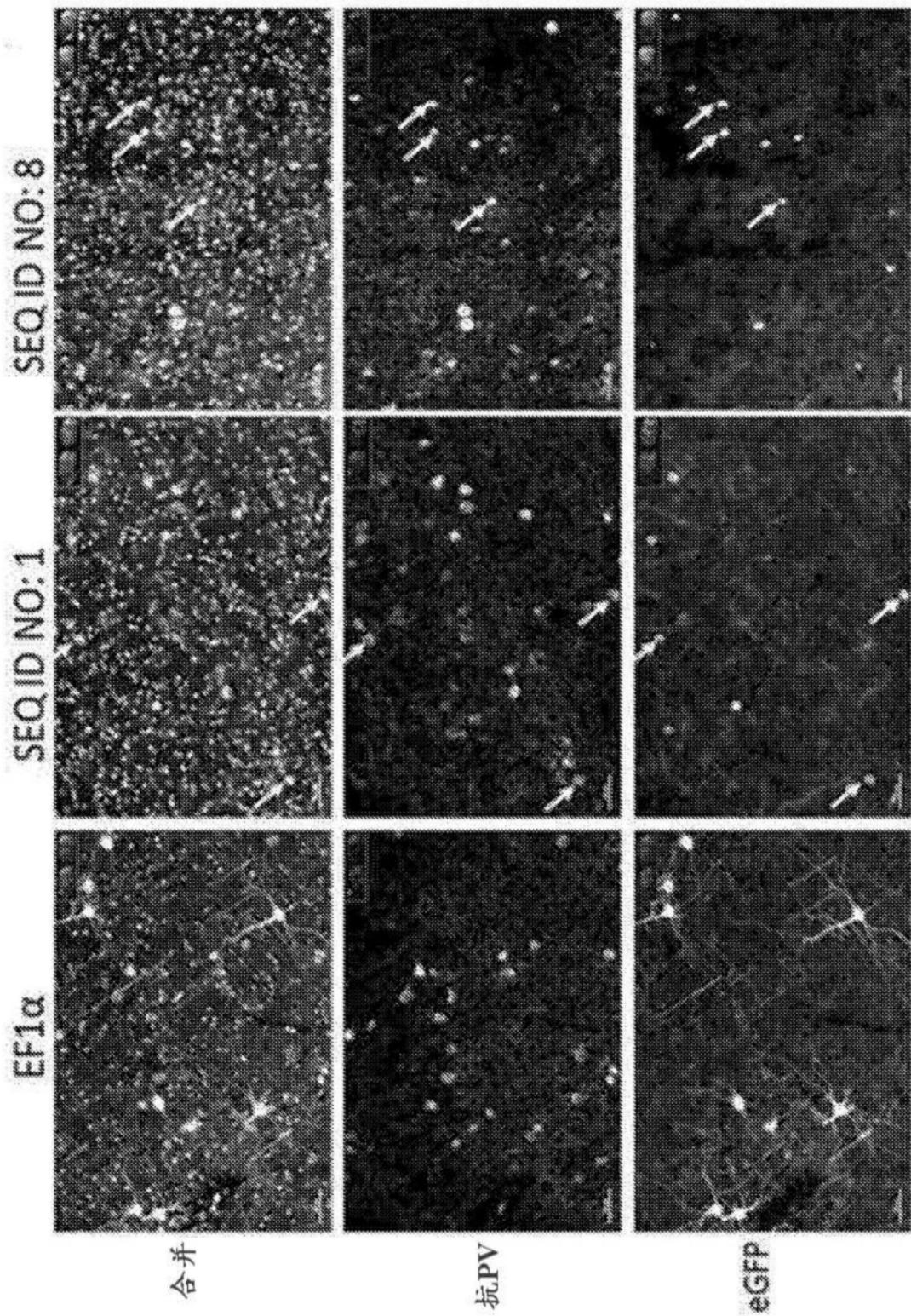


图4A

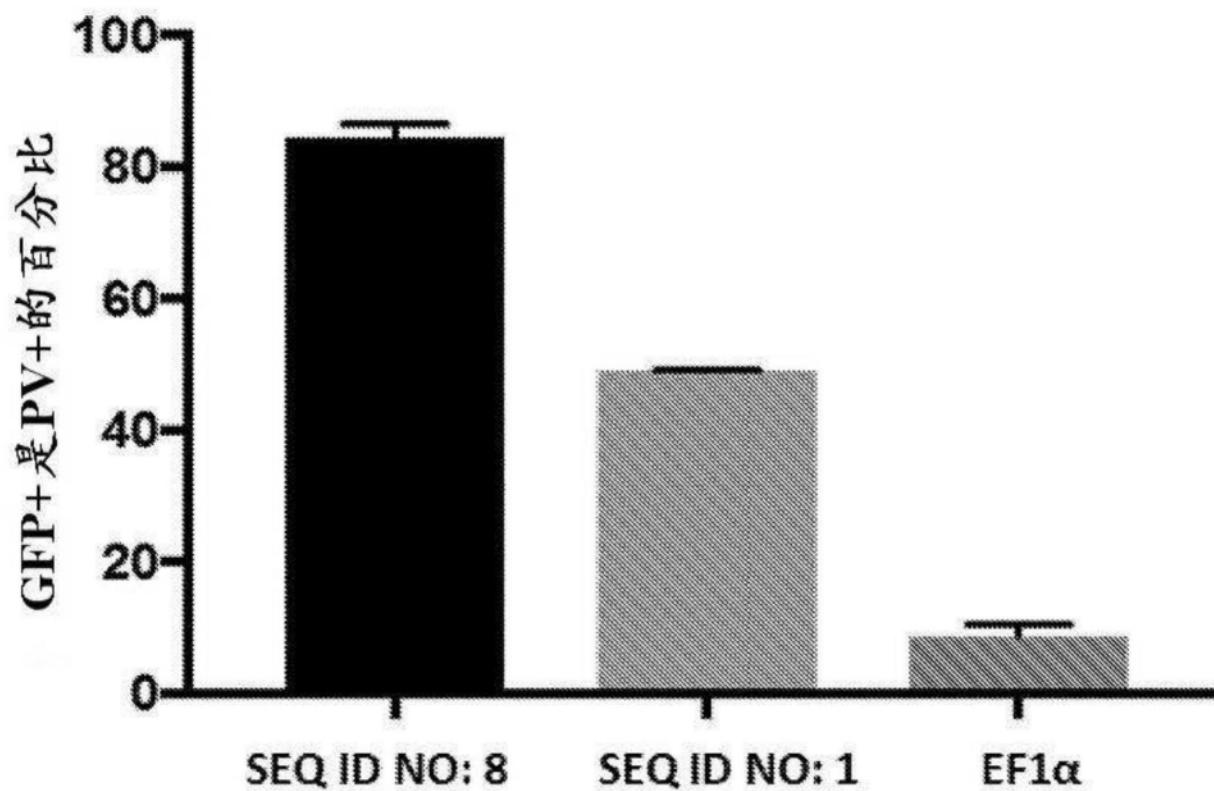


图4B

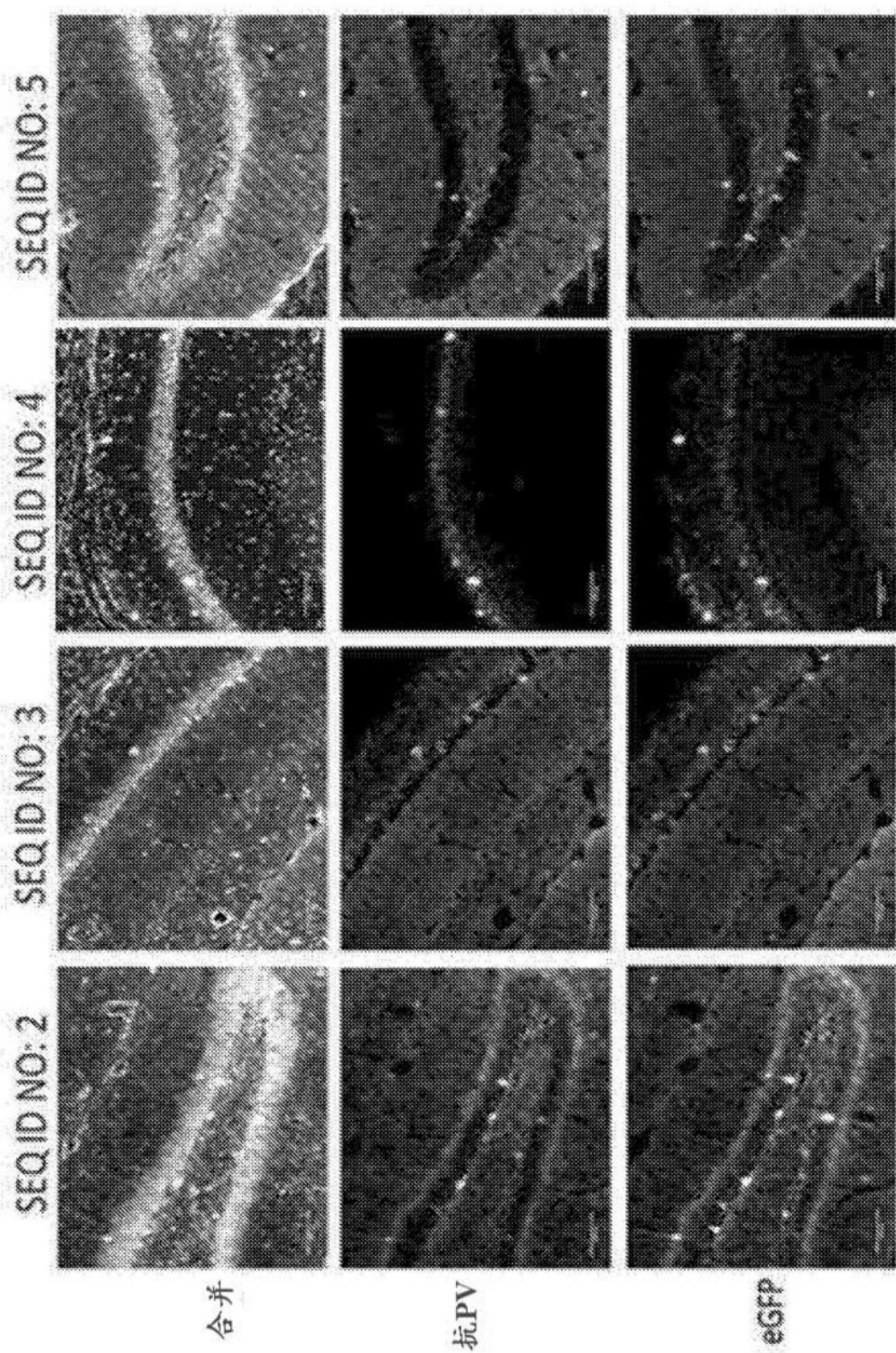


图5A

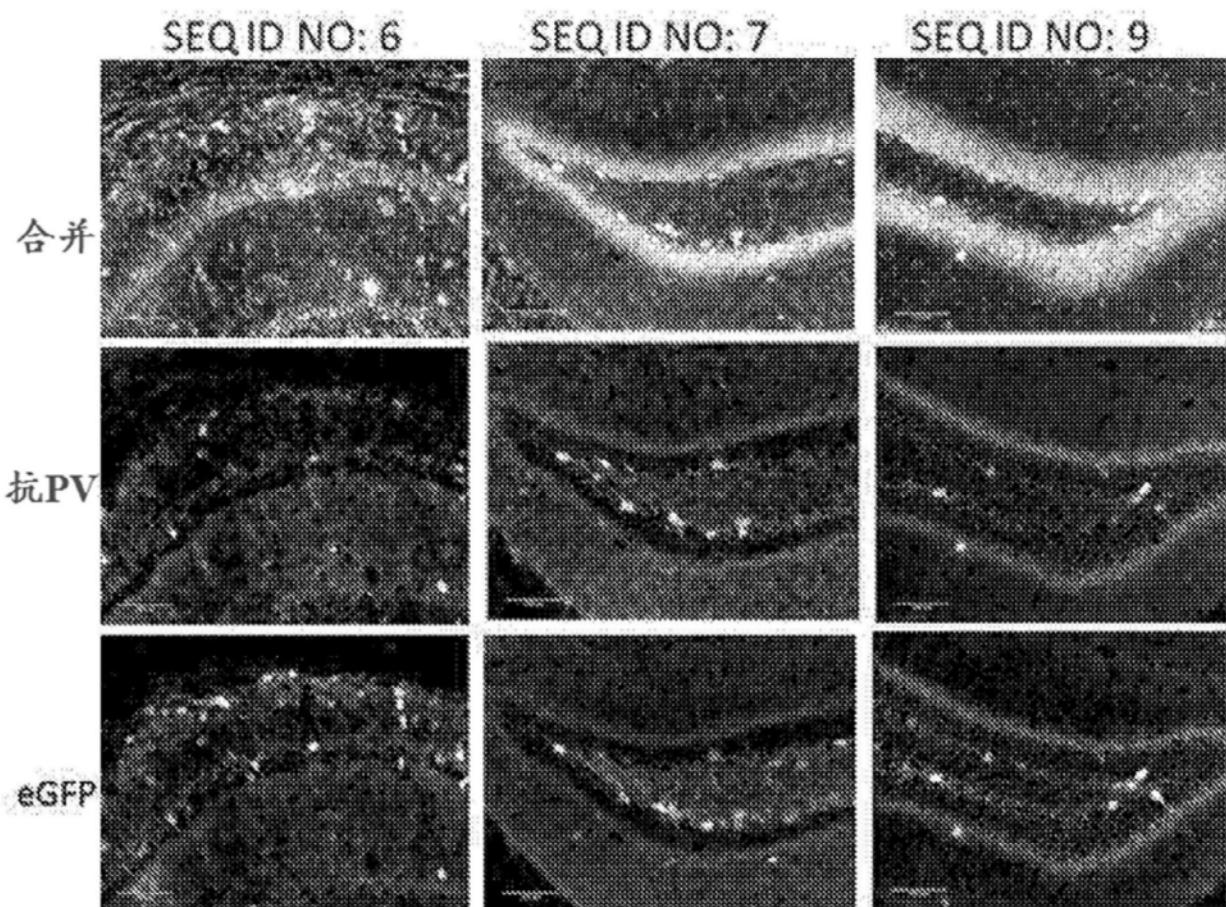


图5B

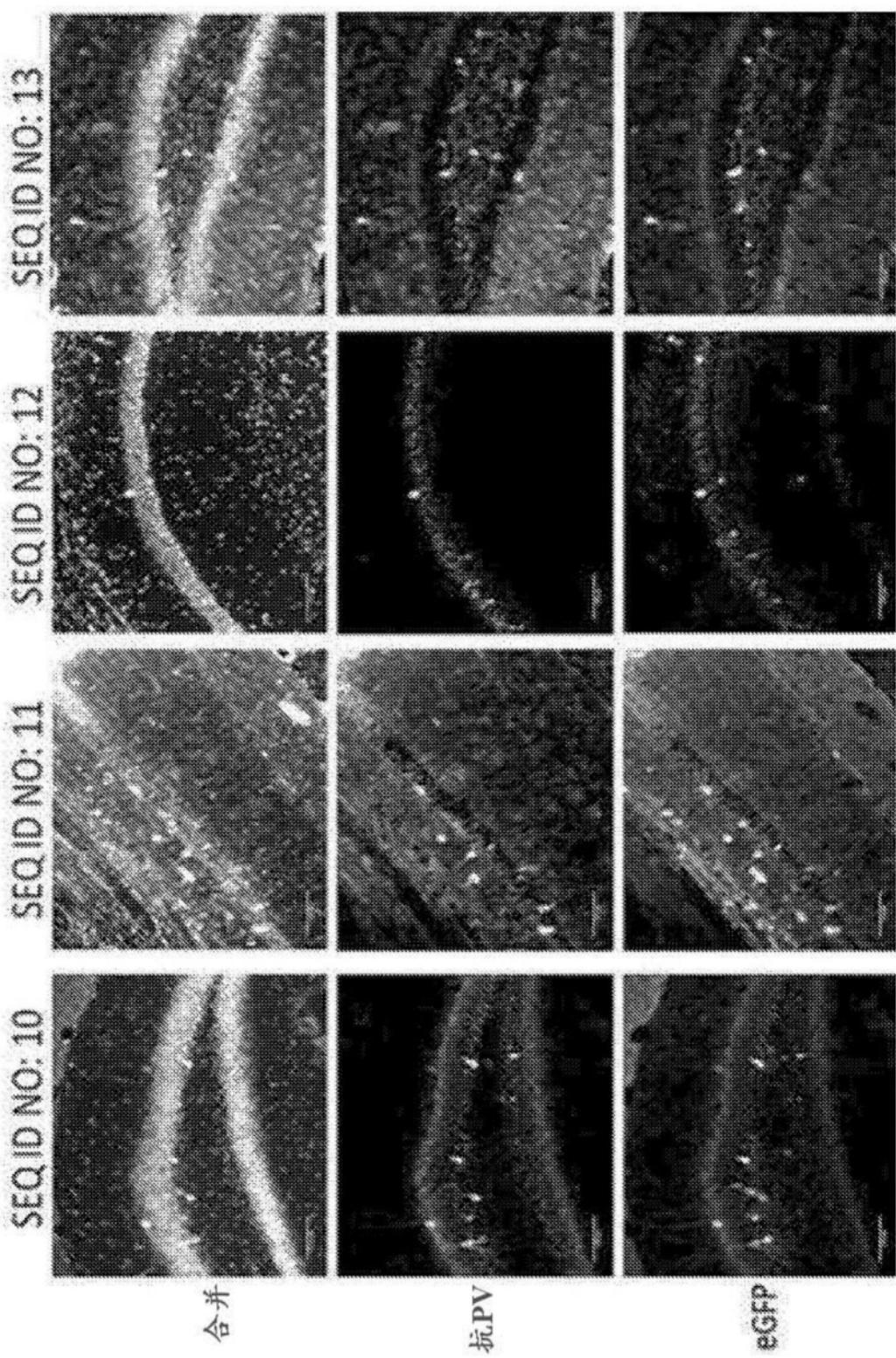


图5C

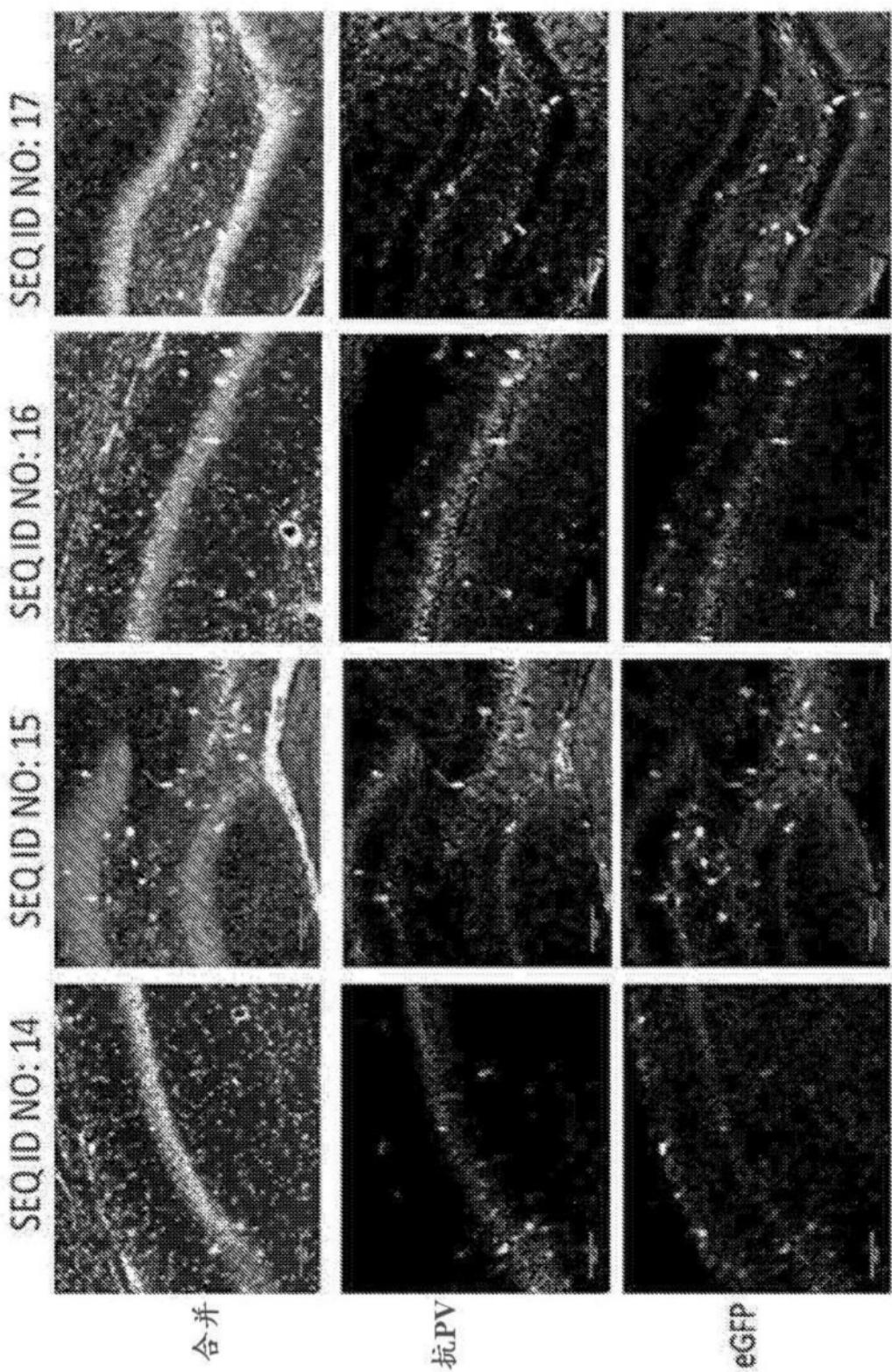


图5D

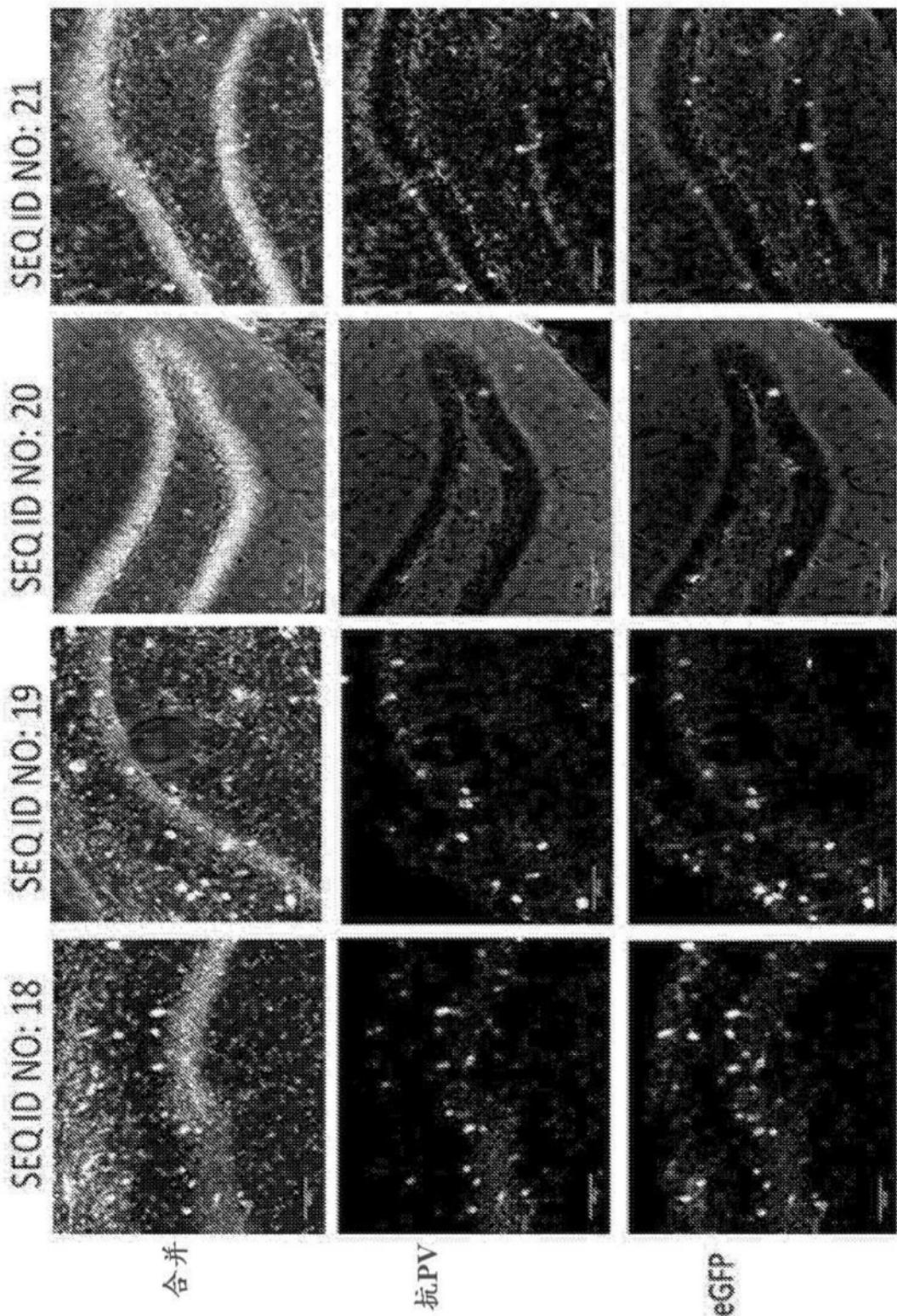


图5E

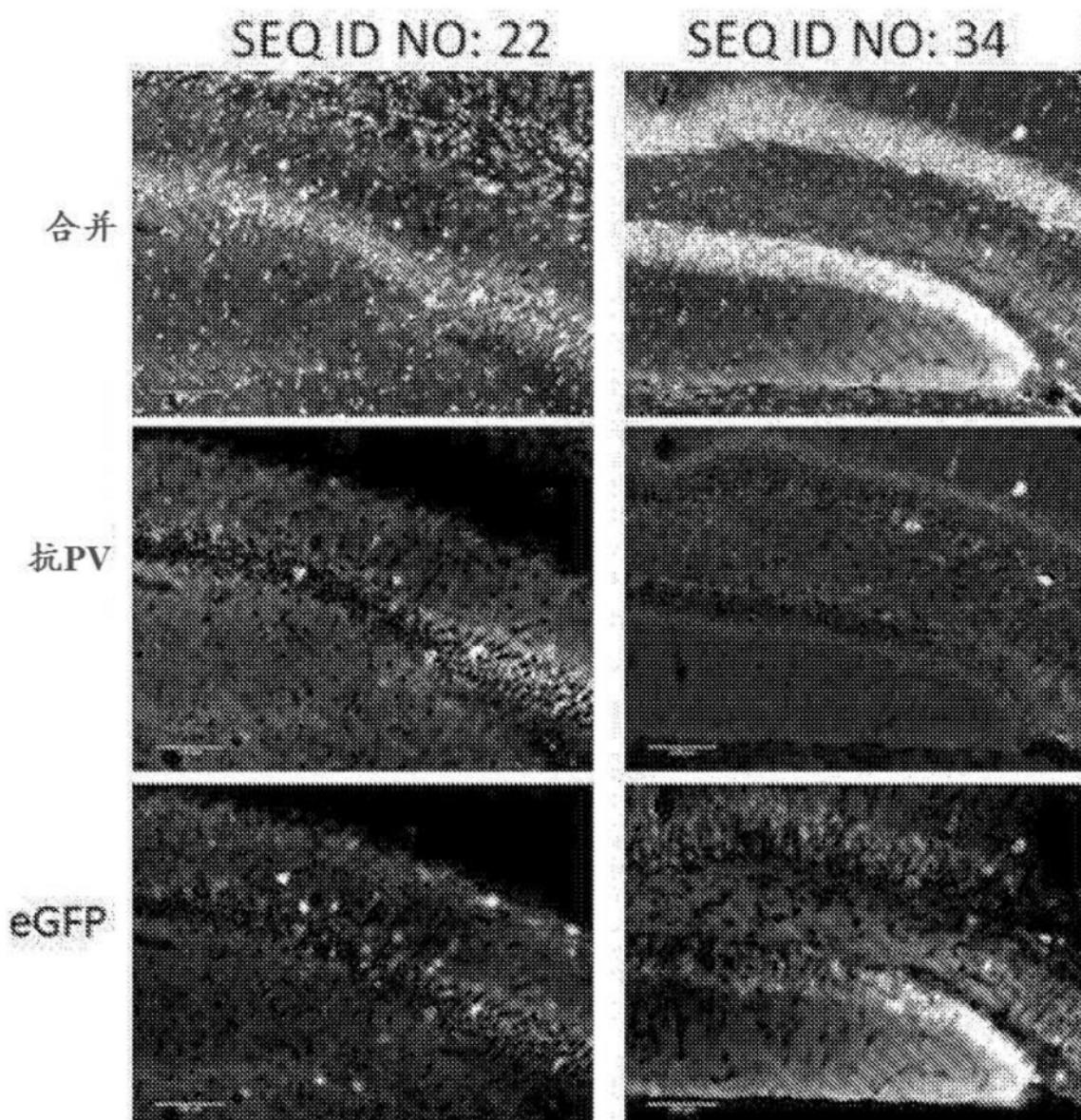


图5F

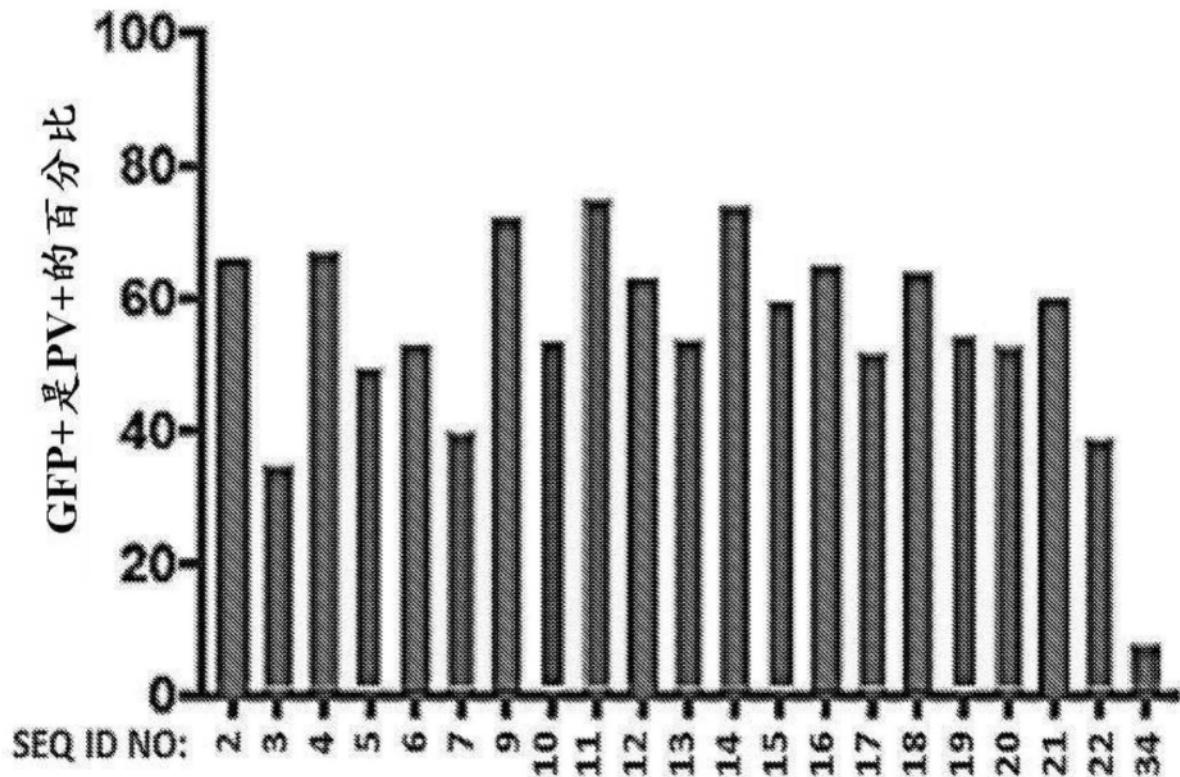


图6

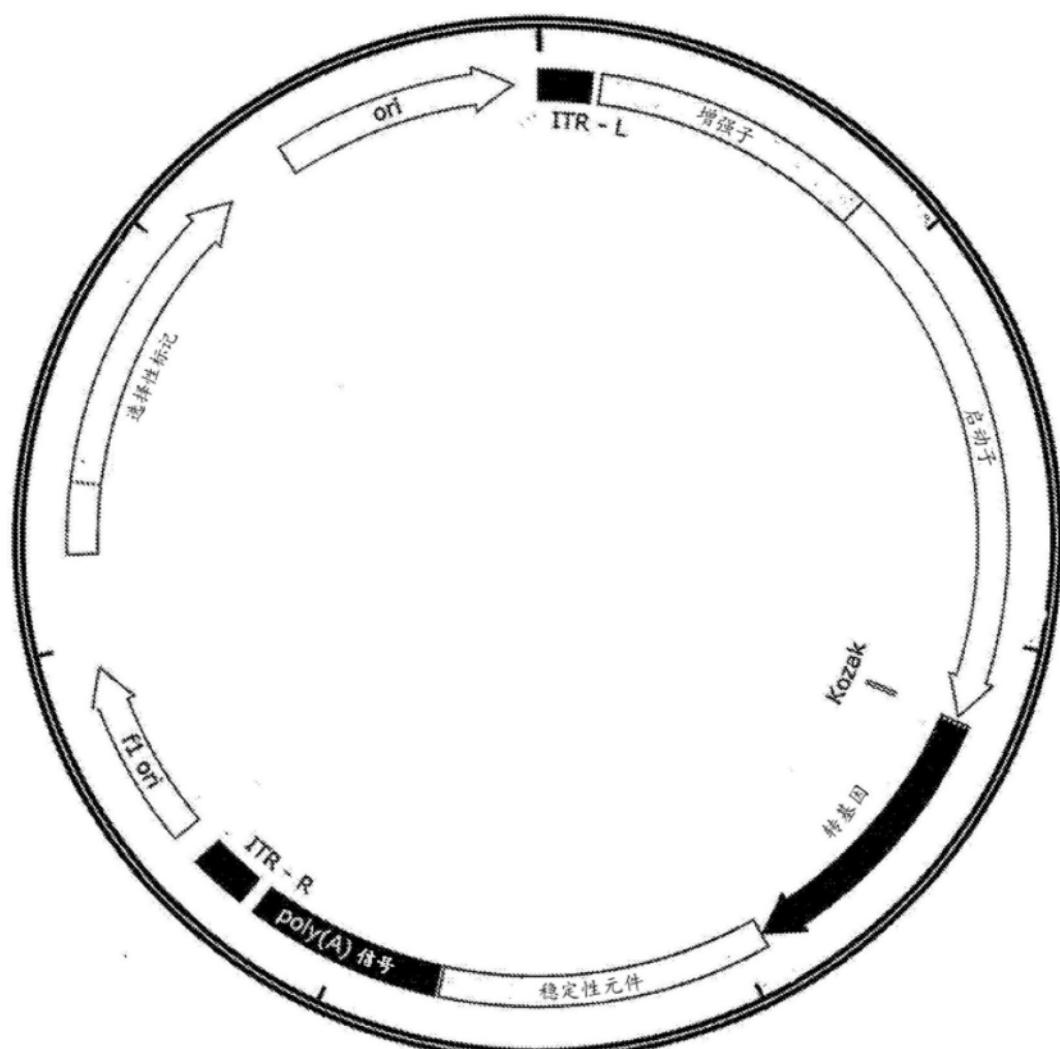


图7