



(12) 发明专利申请

(10) 申请公布号 CN 118345115 A

(43) 申请公布日 2024. 07. 16

(21) 申请号 202410447378.X

(22) 申请日 2018.11.30

(30) 优先权数据

62/593,824 2017.12.01 US
62/610,014 2017.12.22 US
62/618,966 2018.01.18 US
62/641,806 2018.03.12 US
62/646,198 2018.03.21 US
62/664,814 2018.04.30 US
62/664,817 2018.04.30 US

(62) 分案原申请数据

201880088375.4 2018.11.30

(71) 申请人 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

安妮·坦依豪斯

卡提克·拉马穆尔蒂 安德鲁·杨

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

专利代理师 陈文平 王北南

(51) Int.Cl.

C12N 15/85 (2006.01)

C12N 15/864 (2006.01)

C07K 14/47 (2006.01)

C07K 14/705 (2006.01)

A61K 38/00 (2006.01)

A61K 48/00 (2006.01)

A61P 37/04 (2006.01)

C12N 15/63 (2006.01)

权利要求书4页 说明书187页

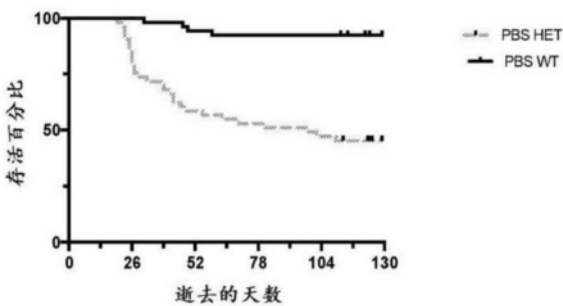
序列表(电子公布) 附图40页

(54) 发明名称

工程化的DNA结合蛋白

(57) 摘要

本文提供了包含调节内源基因表达的非天然存在的DNA结合蛋白的组合物及其使用方法。



1. 一种表达盒,其包含编码非天然存在的转录因子的序列,所述转录因子增加细胞中SCN1A基因的表达,其中所述转录因子与能够在转录激活试验中使SCN1A表达增加至少20%的靶位点结合,并且其中所述转录因子能够在高热癫痫发作试验中将Dravet综合征小鼠模型中的癫痫发作频率降低至少20%。

2. 根据权利要求1所述的表达盒,其中所述转录因子与具有18-27个核苷酸的基因组区域结合。

3. 根据权利要求1或2中任一项所述的表达盒,其中所述转录因子包含DNA结合域。

4. 根据权利要求1-3中任一项所述的表达盒,其中所述转录因子包含DNA结合域和转录激活域。

5. 根据权利要求3或4中任一项所述的表达盒,其中所述DNA结合域与其最接近的人类对应物具有至少80%的序列同一性。

6. 根据权利要求5所述的表达盒,其中所述DNA结合域与其最接近的人类对应物具有至少90%的序列同一性。

7. 根据权利要求4或5中任一项所述的表达盒,其中所述DNA结合域和转录激活域均与其最接近的人类对应物具有至少80%的序列同一性。

8. 根据权利要求7所述的表达盒,其中所述DNA结合域和转录激活域均与其最接近的人类对应物具有至少90%的序列同一性。

9. 根据权利要求3或4中任一项所述的表达盒,其中所述DNA结合域包含指导RNA和核酸酶灭活的Cas蛋白。

10. 根据权利要求9所述的表达盒,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。

11. 根据权利要求3-8中任一项所述的表达盒,其中所述DNA结合域包含锌指结构域。

12. 根据权利要求11所述的表达盒,其中所述DNA结合域包含6至9个锌指结构域。

13. 根据权利要求12所述的表达盒,其中所述DNA结合域包含6个锌指。

14. 根据权利要求13所述的表达盒,其中所述DNA结合域与具有18个核苷酸的基因组区域结合。

15. 根据权利要求12所述的表达盒,其中所述DNA结合域包含9个锌指。

16. 根据权利要求15所述的表达盒,其中所述DNA结合域与具有27个核苷酸的基因组区域结合。

17. 根据权利要求3-8或11-12中任一项所述的表达盒,其中所述DNA结合域包含与SEQ ID NO:135、371、372或376中的任一个具有至少95%序列同一性的序列。

18. 根据权利要求17所述的表达盒,其中所述DNA结合域包含具有SEQ ID NO:135、371、372或376中的任一个的序列。

19. 根据权利要求3-8或11-18中任一项所述的表达盒,其中所述DNA结合域衍生自人EGR1或人EGR3。

20. 根据权利要求4-8或11-19中任一项所述的表达盒,其中所述转录激活域包含VPR、VP64、CITED2或CITED4序列,或其功能片段。

21. 根据权利要求4-8或11-20中任一项所述的表达盒,其中所述转录激活域包含人CITED2或CITED4序列,或其功能片段。

22. 根据权利要求1-21中任一项所述的表达盒,其中所述表达盒进一步包含调节元件,所述调节元件在PV神经元中以比在其他细胞类型中更高的水平驱动所述转录因子的表达。

23. 根据权利要求22所述的表达盒,其中所述调节元件包含具有SEQ ID NO:183、184、185或417中的任一个的序列。

24. 根据权利要求23所述的表达盒,其中所述调节元件包含具有SEQ ID NO:183或185的序列。

25. 根据权利要求1-24中任一项所述的表达盒,其中所述表达盒是病毒载体的一部分。

26. 根据权利要求25所述的表达盒,其中所述病毒载体是AAV病毒。

27. 根据权利要求26所述的表达盒,其中所述AAV病毒是AAV9病毒或scAAV9病毒。

28. 一种通过施用权利要求1-27中任一项的表达盒来增加细胞中SCN1A的表达的方法。

29. 根据权利要求28所述的方法,其中所述细胞是PV神经元。

30. 根据权利要求28或29中任一项所述的方法,其中所述细胞在受试者内。

31. 根据权利要求30所述的方法,其中所述受试者是哺乳动物。

32. 根据权利要求31所述的方法,其中所述受试者是人。

33. 根据权利要求28-32中任一项所述的方法,其中增加SCN1A的表达治疗疾病、病症或症状。

34. 根据权利要求33所述的方法,其中所述病症是中枢神经系统病症。

35. 根据权利要求34所述的方法,其中所述病症是Dravet综合征。

36. 根据权利要求34或35中任一项所述的方法,其中所述中枢神经系统病症的症状是神经元活动过度。

37. 根据权利要求36所述的方法,其中治疗所述中枢神经系统病症包括减轻神经元活动过度。

38. 根据权利要求34或35中任一项所述的方法,其中所述中枢神经系统病症的症状是癫痫发作。

39. 根据权利要求38所述的方法,其中治疗所述中枢神经系统病症包括降低癫痫发作的频率。

40. 根据权利要求38或39中任一项所述的方法,其中治疗所述中枢神经系统病症包括降低癫痫发作的严重程度。

41. 一种表达盒,其包含编码增加细胞中GRN基因表达的非天然存在的转录因子的序列。

42. 根据权利要求41所述的表达盒,其中所述转录因子与具有18-27个核苷酸的基因组区域结合。

43. 根据权利要求41或42中任一项所述的表达盒,其中所述转录因子包含DNA结合域。

44. 根据权利要求41-43中任一项所述的表达盒,其中所述转录因子包含DNA结合域和转录激活域。

45. 根据权利要求43或44中任一项所述的表达盒,其中所述DNA结合域包含指导RNA和核酸酶灭活的Cas蛋白。

46. 根据权利要求45所述的表达盒,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。

47. 根据权利要求43或44中任一项所述的表达盒,其中所述DNA结合域包含锌指结构域。
48. 根据权利要求47所述的表达盒,其中所述DNA结合域包含6至9个锌指结构域。
49. 根据权利要求48所述的表达盒,其中所述DNA结合域包含6个锌指。
50. 根据权利要求49所述的表达盒,其中所述DNA结合域与具有18个核苷酸的基因组区域结合。
51. 根据权利要求48所述的表达盒,其中所述DNA结合域包含9个锌指。
52. 根据权利要求51所述的表达盒,其中所述DNA结合域与具有27个核苷酸的基因组区域结合。
53. 根据权利要求43-44或47-48中任一项所述的表达盒,其中所述DNA结合域包含与SEQ ID NO:171或412-416中的任一个具有至少95%序列同一性的序列。
54. 根据权利要求53所述的表达盒,其中所述DNA结合域包含具有SEQ ID NO:171或412-416中的任一个的序列。
55. 根据权利要求43-44或47-54中任一项所述的表达盒,其中所述DNA结合域衍生自人EGR1或人EGR3。
56. 根据权利要求44或47-55中任一项所述的表达盒,其中所述转录激活域包含VPR、VP64、CITED2或CITED4序列,或其功能片段。
57. 根据权利要求41-56中任一项所述的表达盒,其中所述表达盒进一步包含驱动所述转录因子的表达的调节元件。
58. 根据权利要求57所述的表达盒,其中所述调节元件是细胞类型选择性调节元件。
59. 根据权利要求41-58中任一项所述的表达盒,其中所述表达盒是病毒载体的一部分。
60. 根据权利要求59所述的表达盒,其中所述病毒载体是AAV病毒。
61. 根据权利要求60所述的表达盒,其中所述AAV病毒是AAV9病毒或scAAV9病毒。
62. 一种通过施用权利要求41-61中任一项的表达盒来增加细胞中GRN的表达的方法。
63. 一种表达盒,其编码包含DNA结合域的非天然存在的DNA结合蛋白,该DNA结合域包含三个或更多个锌指结构域,其中所述DNA结合蛋白将SCN1A以外的内源基因的表达提高或抑制至少5倍,并且其中所述DNA结合蛋白与其最接近的人类对应物具有90%或更高的序列同一性。
64. 根据权利要求63所述的表达盒,其中所述DNA结合蛋白与人类基因组中所述人类对应蛋白质天然不结合的靶位点结合。
65. 根据权利要求63或64所述的表达盒,其中所述人类对应物是EGR1。
66. 根据权利要求63或64所述的表达盒,其中所述人类对应物是EGR3。
67. 根据权利要求63-66中任一项所述的表达盒,其中所述DNA结合蛋白包含至少6个锌指结构域。
68. 根据权利要求67所述的表达盒,其中所述DNA结合蛋白包含9个锌指结构域。
69. 根据权利要求63-68中任一项所述的表达盒,其中所述DNA结合域在一个或多个锌指结构域的识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。
70. 根据权利要求63-69中任一项所述的表达盒,其中所述DNA结合蛋白包含一个或多

个衍生自EGR1或EGR3的锌指结构域。

71. 根据权利要求63-70中任一项所述的表达盒, 其中所述DNA结合蛋白进一步包含转录效应物结构域。

72. 根据权利要求71所述的表达盒, 其中所述转录效应物结构域衍生自CITED2或CITED4。

73. 根据权利要求71或72所述的表达盒, 其中所述转录效应物结构域位于所述DNA结合蛋白中的DNA结合域的C-末端。

74. 根据权利要求71或72所述的表达盒, 其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的N-末端。

75. 根据权利要求71或72所述的表达盒, 其中所述DNA结合蛋白包含位于所述DNA结合蛋白中的DNA结合域的N-末端的效应物结构域和位于所述DNA结合蛋白中的DNA结合域的C-末端的效应物结构域。

76. 根据权利要求63-75中任一项所述的表达盒, 其中所述DNA结合蛋白识别18bp或更长的靶结合位点。

77. 根据权利要求76所述的表达盒, 其中所述DNA结合蛋白识别27bp的靶结合位点。

78. 根据权利要求63-77中任一项所述的表达盒, 其中所述表达盒进一步包含细胞类型选择性调节元件。

79. 根据权利要求63-78中任一项所述的表达盒, 其中所述表达盒在病毒载体中。

80. 根据权利要求79所述的表达盒, 其中所述病毒载体是AAV。

81. 根据权利要求80所述的表达盒, 其中所述AAV是AAV9或scAAV9。

82. 一种治疗疾病或病况的方法, 所述方法包括施用包含根据权利要求63-81中任一项所述的表达盒的基因疗法。

83. 根据权利要求82所述的方法, 其中当在细胞中或在体内表达时, 所述DNA结合蛋白引发降低的或最小的免疫应答。

84. 根据权利要求82或83中任一项所述的方法, 其中所述细胞在受试者内。

85. 根据权利要求84所述的方法, 其中所述受试者是哺乳动物。

86. 根据权利要求85所述的方法, 其中所述受试者是人。

工程化的DNA结合蛋白

[0001] 本申请是申请日为2018年11月30日,申请号为“201880088375.4”,发明名称为“工程化的DNA结合蛋白”的申请的分案申请。

[0002] 交叉引用

[0003] 本申请要求2017年12月1日提交的第62/593,824号美国临时专利申请、2017年12月22日提交的第62/610,014号美国临时专利申请、2018年3月21日提交的第62/646,198号美国临时专利申请、2018年1月18日提交的第62/618,966号美国临时专利申请、2018年3月12日提交的第62/641,806号美国临时专利申请、2018年4月30日提交的第62/664,814号美国临时专利申请和2018年4月30日提交的第62/664,817号美国临时专利申请的权益,其中每个临时申请均通过引用整体并入本文。

[0004] 序列表

[0005] 本申请含有序列表,该序列表已经以ASCII格式电子提交,并且通过引用整体并入本文。创建于2018年11月30日的所述ASCII副本被命名为46482-711_601_SL.txt,大小为864,453个字节。

背景技术

[0006] 众多的人类疾病与基因的异常表达相关。在一些情况下,基因中的遗传突变会导致其调节异常、下调或根本不表达,从而导致单倍体不足。在一些情况下,基因中的遗传突变会导致其上调,从而导致基因的过表达。在治疗遗传性病症或疾病中存在许多挑战。一种方法是基因疗法,它涉及将核酸治疗性地递送到患者细胞中。然而,与基因疗法相关的各种挑战仍未得到解决,如基因疗法引发的不希望免疫应答、脱靶效应、对基因治疗媒介物(例如病毒)克隆能力的限制、在较长时段内维持治疗效果等。中枢神经系统(CNS)对开发解决基因和/或蛋白质表达潜在损害的疗法提出了许多独特的挑战。虽然有一些药物可以帮助控制CNS疾病/病症的症状,但许多CNS疾病/病症,例如Dravet综合征或额颞痴呆,缺乏特定的治疗或治愈方法。因此,需要能够调节任何内源基因的表达以帮助逆转疾病或病症的影响的新组合物和方法,特别是具有降低的免疫原性、降低的脱靶效应、增加的靶基因特异性和/或增加的治疗功效的疗法。

发明内容

[0007] 在一个方面,本申请提供了一种表达盒,其包含编码增加细胞中SCN1A基因表达的非天然存在的转录因子的序列。在某些实施方案中,所述增加SCN1A的表达的转录因子与能够在转录激活试验中使SCN1A表达增加至少20%的靶位点结合。在某些实施方案中,所述增加SCN1A的表达的转录因子能够在高热癫痫发作试验中将Dravet综合征小鼠模型中的癫痫发作频率降低至少20%。在某些实施方案中,所述增加SCN1A的表达的转录因子与能够在转录激活试验中使SCN1A表达增加至少20%的靶位点结合,并且所述转录因子能够在高热癫痫发作试验中将Dravet综合征小鼠模型中的癫痫发作频率降低至少20%。

[0008] 在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置

结合。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在SCN1A基因的转录起始位点上游或下游110kb内。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在位置166179652-165989571(参照GRCh38.p12)内。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在位置166128050-166127958(参照GRCh38.p12)内。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在位置166155414-166140590(参照GRCh38.p12)内。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在位置166179652-1661777272(参照GRCh38.p12)内。在某些实施方案中,所述增加SCN1A的表达的转录因子与染色体2上的基因组位置结合,该基因组位置在位置1659990246-165989592(参照GRCh38.p12)内。在某些实施方案中,所述增加SCN1A的表达的转录因子与基因组位置的200bp内的基因组区域结合,该基因组位置具有SEQ ID NO:35-37、105-111、136、195-211、224-238、240-267中的任一个的序列。在某些实施方案中,所述增加SCN1A的表达的转录因子与同基因组位置至少部分重叠的基因组区域结合,该基因组位置具有SEQ ID NO:35-37、105-111、136、195-211、224-238、240-267中的任一个的序列。在某些实施方案中,所述增加SCN1A的表达的转录因子与基因组区域结合,该基因组区域具有SEQ ID NO:35-37、105-111、136、195-211、224-238、240-267中的任一个的序列。在某些实施方案中,所述增加SCN1A的表达的转录因子与具有18-27个核苷酸(例如,至少12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30个核苷酸)的基因组区域结合。

[0009] 在某些实施方案中,所述增加SCN1A的表达的转录因子包含DNA结合域。在某些实施方案中,所述增加SCN1A的表达的转录因子包含DNA结合域和转录激活域。在某些实施方案中,所述DNA结合域与其最接近的人类对应物具有至少80%的序列同一性。在某些实施方案中,所述DNA结合域与其最接近的人类对应物具有至少90%的序列同一性。在某些实施方案中,所述DNA结合域和转录激活域均与其最接近的人类对应物具有至少80%的序列同一性。在某些实施方案中,所述DNA结合域和转录激活域均与其最接近的人类对应物具有至少90%的序列同一性。在某些实施方案中,所述DNA结合域包含指导RNA和核酸酶灭活的Cas蛋白。在某些实施方案中,所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。在某些实施方案中,所述DNA结合域包含锌指结构域。在某些实施方案中,所述DNA结合域包含6至9个锌指结构域。在某些实施方案中,所述DNA结合域包含6个锌指。在某些实施方案中,所述DNA结合域包含6个与具有18个核苷酸的基因组区域结合的指。在某些实施方案中,所述DNA结合域包含9个锌指。在某些实施方案中,所述DNA结合域包含9个与具有27个核苷酸的基因组区域结合的锌指。在某些实施方案中,所述DNA结合域包含与SEQ ID NO:135、371、372或376中的任一个具有至少95%序列同一性的序列。在某些实施方案中,所述DNA结合域包含具有SEQ ID NO:135、371、372或376中的任一个的序列。在某些实施方案中,所述DNA结合域衍生自人EGR1或人EGR3。在某些实施方案中,所述转录激活域包含VPR、VP64、CITED2或CITED4序列,或其功能片段。在某些实施方案中,所述转录激活域包含人CITED2或CITED4序列,或其功能片段。

[0010] 在某些实施方案中,所述编码增加SCN1A表达的转录因子的表达盒进一步包含调节元件,该调节元件在PV神经元中以比在其他细胞类型中更高的水平驱动所述转录因子的

表达。在某些实施方案中,所述调节元件包含具有SEQ ID NO:183、184、185或417中的任一个的序列。在某些实施方案中,所述调节元件包含具有SEQ ID NO:183或185的序列。

[0011] 在某些实施方案中,所述编码增加SCN1A表达的转录因子的表达盒是病毒载体的一部分。在某些实施方案中,所述病毒载体是AAV病毒。在某些实施方案中,所述AAV病毒是AAV9病毒或scAAV9病毒。

[0012] 在另一方面,本申请提供了一种通过施用本文所述的任何表达盒来增加细胞中SCN1A表达的方法,该表达盒编码增加SCN1A的表达的转录因子。在某些实施方案中,所述细胞是PV神经元。在某些实施方案中,所述细胞在受试者内。在某些实施方案中,所述受试者是哺乳动物。在某些实施方案中,所述受试者是人。在某些实施方案中,增加SCN1A的表达治疗疾病、病症或症状。在某些实施方案中,所述病症是中枢神经系统病症。在某些实施方案中,所述病症是Dravet综合征。在某些实施方案中,所述中枢神经系统病症是神经元活动过度。在某些实施方案中,治疗所述中枢神经系统病症包括减轻神经元活动过度。在某些实施方案中,所述中枢神经系统病症的症状是癫痫发作。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的频率。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的严重程度。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的严重程度和频率。

[0013] 在另一方面,本申请提供了一种表达盒,其包含编码增加细胞中GRN基因表达的非天然存在的转录因子的序列。

[0014] 在某些实施方案中,所述增加GRN的表达的转录因子与具有18-27个核苷酸的基因组区域结合。

[0015] 在某些实施方案中,所述增加GRN的表达的转录因子包含DNA结合域。在某些实施方案中,所述增加GRN的表达的转录因子包含DNA结合域和转录激活域。在某些实施方案中,所述DNA结合域包含指导RNA和核酸酶灭活的Cas蛋白。在某些实施方案中,所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。在某些实施方案中,所述DNA结合域包含锌指结构域。在某些实施方案中,所述DNA结合域包含6至9个锌指结构域。在某些实施方案中,所述DNA结合域包含6个锌指。在某些实施方案中,所述DNA结合域包含6个与具有18个核苷酸的基因组区域结合的锌指。在某些实施方案中,所述DNA结合域包含9个锌指。在某些实施方案中,所述DNA结合域包含9个与具有27个核苷酸的基因组区域结合的锌指。在某些实施方案中,所述DNA结合域包含与SEQ ID NO:171或412-416中的任一个具有至少95%序列同一性的序列。在某些实施方案中,所述DNA结合域包含具有SEQ ID NO:171或412-416中的任一个的序列。在某些实施方案中,所述DNA结合域衍生自人EGR1或人EGR3。在某些实施方案中,所述转录激活域包含VPR、VP64、CITED2或CITED4序列,或其功能片段。

[0016] 在某些实施方案中,所述编码增加GRN表达的转录因子的表达盒进一步包含调节元件,该调节元件驱动非天然存在的转录调节物的表达。在某些实施方案中,所述调节元件是细胞类型选择性调节元件。

[0017] 在某些实施方案中,所述编码增加GRN表达的转录因子的表达盒是病毒载体的一部分。在某些实施方案中,所述病毒载体是AAV病毒。在某些实施方案中,所述AAV病毒是AAV9病毒或scAAV9病毒。

[0018] 在另一方面,本申请提供了一种通过施用本文所述的任何表达盒来增加细胞中

GRN表达的方法,该表达盒编码增加GRN的表达的转录因子。在某些实施方案中,增加GRN的表达治疗疾病、病症或症状。在某些实施方案中,所述细胞在受试者内。在某些实施方案中,所述受试者是哺乳动物。在某些实施方案中,所述受试者是人。在某些实施方案中,所述细胞选自:中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞和纹状体细胞。在某些实施方案中,调节GRN的表达治疗疾病或病症。在某些实施方案中,所述病症是中枢神经系统病症。在某些实施方案中,所述病症是额颞变性(FTD)、帕金森病、阿尔茨海默病或动脉粥样硬化。在某些实施方案中,中枢神经系统病症的症状是Lewy小体的存在、颗粒蛋白前体(GRN)的单倍体不足、社交缺陷、溶酶体异常、记忆丧失、运动协调性丧失或肌肉震颤。在某些实施方案中,治疗中枢神经系统病症包括降低肌肉震颤的频率和/或降低肌肉震颤的严重程度。

[0019] 在另一方面,本申请提供了一种表达盒,其编码包含DNA结合域的非天然存在的DNA结合蛋白,该DNA结合域包含三个或更多个DNA结合域,其中所述DNA结合蛋白将内源基因的表达提高或抑制至少5倍,并且其中所述DNA结合蛋白与其最接近的人类对应物具有90%或更高的序列同一性。在示例性实施方案中,所述DNA结合域包含三个或更多个锌指结构域。

[0020] 在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白与人类基因组中人类对应蛋白质天然不结合的靶位点结合。在某些实施方案中,所述人类对应物是EGR1。在某些实施方案中,所述人类对应物是EGR3。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白包含至少6个锌指结构域。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白包含9个锌指结构域。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合域在一个或多个锌指结构域的识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白包含一个或多个衍生自EGR1或EGR3的锌指结构域。

[0021] 在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白进一步包含转录效应物结构域。在某些实施方案中,所述转录效应物结构域衍生自CITED2或CITED4。在某些实施方案中,所述转录效应物结构域位于所述DNA结合蛋白中的DNA结合域的C-末端。在某些实施方案中,所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的N-末端。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白包含位于所述DNA结合蛋白中的DNA结合域的N-末端的效应物结构域和位于所述DNA结合蛋白中的DNA结合域的C-末端的效应物结构域。

[0022] 在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白识别18bp或更长的靶结合位点。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白识别27bp的靶结合位点。

[0023] 在某些实施方案中,编码与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白的表达盒进一步包含细胞类型选择性调节元件。在某些实施方案中,编码与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白的表达盒进一步包含调节元件,该调节元件在PV神经元中以比在其他细胞类型中更高的水平驱动所述DNA结合蛋白

的表达。在某些实施方案中,所述调节元件包含具有SEQ ID NO:183、184、185或417中的任一个的序列。在某些实施方案中,所述调节元件包含具有SEQ ID NO:183或185的序列。

[0024] 在某些实施方案中,编码与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白的表达盒是病毒载体。在某些实施方案中,所述病毒载体是AAV。在某些实施方案中,所述AAV是AAV9或scAAV9。

[0025] 在另一方面,本申请提供了一种治疗疾病或病况的方法,该方法包括施用包含本文提供的任何表达盒的基因疗法,该表达盒编码与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白当在细胞中或在体内表达时,引发降低的或最小的免疫应答。在某些实施方案中,所述细胞在受试者内。在某些实施方案中,所述受试者是哺乳动物。在某些实施方案中,所述受试者是人。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白增加内源SCN1A的表达。在某些实施方案中,所述细胞是PV神经元。在某些实施方案中,增加SCN1A的表达治疗疾病、病症或症状。在某些实施方案中,所述病症是中枢神经系统病症。在某些实施方案中,所述病症是Dravet综合征。在某些实施方案中,所述中枢神经系统病症是神经元活动过度。在某些实施方案中,治疗所述中枢神经系统病症包括减轻神经元活动过度。在某些实施方案中,所述中枢神经系统病症的症状是癫痫发作。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的频率。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的严重程度。在某些实施方案中,治疗所述中枢神经系统病症包括降低癫痫发作的严重程度和频率。在某些实施方案中,所述与其最接近的人类对应物具有90%或更高序列同一性的DNA结合蛋白增加内源GRN的表达。在某些实施方案中,所述细胞选自:中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞和纹状体细胞。在某些实施方案中,调节GRN的表达治疗疾病或病症。在某些实施方案中,所述病症是中枢神经系统病症。在某些实施方案中,所述病症是额颞变性(FTD)、帕金森病、阿尔茨海默病或动脉粥样硬化。在某些实施方案中,中枢神经系统病症的症状是Lewy小体的存在、颗粒蛋白前体(GRN)的单倍体不足、社交缺陷、溶酶体异常、记忆丧失、运动协调性丧失或肌肉震颤。在某些实施方案中,治疗中枢神经系统病症包括降低肌肉震颤的频率和/或降低肌肉震颤的严重程度。

[0026] 援引并入

[0027] 本说明书中所提及的所有出版物、专利和专利申请均通过引用并入本文,其程度犹如具体地且单独地指出每个单独的出版物、专利或专利申请均通过引用而并入。

附图说明

[0028] 本发明的新颖特征在所附权利要求书中具体阐述。通过参考以下对利用本发明原理的说明性情况加以阐述的详细描述以及附图,将会对本发明的特征和优点获得更好的理解,在这些附图中:

[0029] 图1示出了以归一化的萤光素酶活性衡量的,不同调节元件对293T细胞中萤光素酶表达的影响。例如,与最小CMV(minCMV)启动子组合驱动萤光素酶表达的水平是单独minCMV启动子驱动的表达水平的约1.4倍,并且是SCP启动子驱动的表达水平的约60倍。在同一实验中,与minCMV启动子连接的SEQ ID NO:178驱动萤光素酶表达的水平是单独

minCMV启动子的约3.5倍,并且是SCP启动子的约140倍。

[0030] 图2示出了每种调节元件的大小归一化的活性(通过将图1所示的归一化的萤光素酶活性除以调节元件的碱基对长度来计算)。

[0031] 图3示出了在GABA能选择性调节元件(RE)(如对小白蛋白(PV)细胞具有选择性的RE)控制下的eGFP的选择性表达。将含有eGFP(在EF1 α 启动子或GABA能/PV选择性RE如SEQ ID NO:183——其包括SEQ ID NO:185——或SEQ ID NO:184的控制下)的AAVDJ载体注射到在GABA能/PV细胞中表达tdTomato(红色荧光蛋白(RFP))的小鼠中。顶行图像显示了从含有EF1 α 启动子(SEQ ID NO:186)、GABA能/PV选择性RE SEQ ID NO:183或SEQ ID NO:184的构建体表达的eGFP的模式。底行图像显示了GABA能/PV神经元,其为GAD2阳性。表达eGFP和tdTomato两者的细胞可以通过将顶行图像与底行图像重叠或通过鉴别来自顶行图像和底行图像中相同位置的荧光来鉴定。表达eGFP和tdTomato两者的细胞的实例用箭头指示。

[0032] 图4示出了图3中的GABA能/PV神经元表达eGFP的效率的定量(例如,以还表达eGFP的tdTomato表达细胞的百分比来测量的)。

[0033] 图5示出了图3中的GABA能/PV神经元表达eGFP的特异性的定量(例如,以还表达tdTomato(或RFP+)的eGFP表达细胞的百分比来测量的)。

[0034] 图6示出了使用与染色体2上的各个区域(参照GRCh38.p12)结合的工程化转录因子,内源SCN1A的上调。数据被呈现为SCN1A表达相对于对照(EGFP-KASH)条件的变化倍数。

[0035] 图7示出了以在Log₁₀上相对于eGFP对照的变化倍数呈现的,SCN1A的相对表达。

[0036] 图8示出了以相对于对照的变化倍数呈现的,内源SCN1A、TTC21B、SCN9A、SNCA、GRN、UTRN和TTN转录物的相对表达。

[0037] 图9A示出了本文公开的eTF的一个实施方案,其中天然存在的人转录因子(TF)的DNA结合域(DBD)已经被修饰、重新编程或工程改造,以识别在目的基因靶标处或附近的靶结合位点。上图示出了天然存在的转录因子,其对与内源目的基因或基因靶标相关的靶位点(用黑条表示)没有结合亲和力。下图示出了一种eTF,其中DBD包含多个锌指(以圆柱体表示),这些锌指已被修饰、工程改造或重新编程,以与目的基因的靶位点结合,例如通过双重复制或三重复制天然存在的DBD的锌指和/或在锌指中进行氨基酸置换。

[0038] 图9B示出了本文公开的eTF或杂合eTF的另一个实施方案,其包含与人共调节蛋白或转录辅因子的反式激活域(TAD)(例如,Cbp/p300相互作用反式激活蛋白如CITED2或CITED4的TAD)融合的衍生自天然存在的人TF的DBD。上图示出了衍生自人类蛋白质并且被工程改造为与基因组中的所需靶位点(黑条)特异性结合的DBD。下图示出了与TAD(例如,CITED2的TAD)融合以形成杂合eTF的DBD。这样的杂合eTF通过与基因组中的靶位点序列结合并经由eTF中的TAD调节(例如,激活)基因靶标的转录来驱动靶基因的表达。

[0039] 图10示出了各种转录调节蛋白质与天然存在的人类蛋白质之间的保守性(或序列同一性)百分比,及其以碱基对(bp)表示的大小。在各个实施方案中,本公开的eTF可以由1,000-1,800bp的核酸序列编码,并且与天然存在的人类蛋白质具有>90%的全局序列同一性(即,在eTF的全长氨基酸序列上的序列同一性),当递送到细胞中或在体内时,例如通过基因疗法,其可以降低免疫应答或免疫原性。相比之下,包含失活的Cas9(dCas9)结构域的基于dCas9的转录因子与内源人类蛋白质没有全局序列同一性,并且大小更大。常规构建的锌指蛋白(ZFP)或包含计算机设计的锌指的人工或合成DNA结合蛋白可具有更小的大小,但与

内源人类蛋白质的序列同一性较低(约55%)。

[0040] 图11示出了用于调节内源基因的各种方法。在顶部图中,转录调节蛋白包含与VP64的反式激活域(TAD)融合的包含多个合成锌指(ZF)如计算机设计的ZF的DBD。在第二幅图中,野生型TF是天然存在的,但不与目的基因的靶位点结合。在第三幅图中,eTF包含天然存在的TF,其DBD已被修饰或重新编程,以与目的基因处或附近的18nt靶位点结合,例如,其中野生型TF的DBD被合成锌指替代。在第四幅图中,eTF衍生自野生型TF,其中DBD被工程改造为识别目的基因处或附近的9nt靶标结合位点。在底部图中,eTF衍生自野生型TF,其中野生型TF的DBD已被双重复制并重新编程(例如,通过ZF中的氨基酸置换)以形成6锌指的DBD,该DBD识别目的基因处或附近的18nt靶结合位点。可以使用报道激活试验来测定eTF的这些各种实施方案,其中基因靶标是报道基因(例如,EGFP)。在顶部、第三、第四和第五幅图中的eTF导致报道基因的表达。

[0041] 图12A示出了SCN1A激活试验的示意图,其中HEK293细胞中SCN1A的表达指示与SCN1A特异性靶结合位点的结合以及eTF对SCN1A基因表达的激活。上图示出了一种eTF,其被工程改造为识别EGFP报道基因的TRE结合位点,其无法与SCN1A靶结合位点结合,因此无法表达SCN1A。下图示出了由天然存在的人类蛋白质衍生的eTF,其与该人类蛋白质具有至少98%的序列同一性,并且包含被工程改造为与18nt SCN1A靶结合位点结合的DBD,导致SCN1A基因激活,从而导致SCN1A表达。

[0042] 图12B示出了采用图12A中示出的示例性eTF,SCN1A激活试验的结果的定量。包含六个锌指的eTF(SEQ ID NO:6)——其与EGR1具有高全局或整体序列同一性,并且其中其DBD被工程改造为识别内源SCN1A基因处或附近的特异性结合位点,导致转染的HEK293细胞中的SCN1A表达2.5倍于SEQ ID NO:4,后者的DBD被工程改造为识别TRE结合位点,而不是SCN1A靶结合位点。

[0043] 图13A示出了SCN1A被两个示例性eTF激活的示意图,每个eTF衍生自天然存在的人类蛋白质。上图示出了eTF,其中天然存在的人类蛋白质(例如,EGR1或EGR3)的DBD被双重复制,以形成具有6个锌指的DBD,并且被修饰或重新编程为与内源SCN1A基因处或附近的靶结合位点结合。下图示出了eTF,其中天然存在的人类蛋白质的DBD被三重复制,以形成具有9个锌指的DBD,并且其被修饰或重新编程为与内源SCN1A基因处或附近的靶结合位点结合。

[0044] 图13B示出了被图13A中示出的每个eTF激活的相对SCN1A表达的定量。在转染的HEK293细胞中,与EGR1具有高序列同一性并且包含衍生自EGR1且被工程改造为识别SCN1A靶结合位点的6锌指DBD的eTF(SEQ ID NO:7),导致SCN1A表达约7倍于空载体对照。与EGR1具有高序列同一性并且包含衍生自EGR1且被工程改造为识别SCN1A靶结合位点的9锌指DBD的eTF(SEQ ID NO:9),导致SCN1A表达约20倍于对照。

[0045] 图14示出了杂合eTF的各种实施方案,所述杂合eTF包含与天然存在的人类蛋白质具有高序列同一性的人DBD,该人DBD与另一种蛋白质的TAD(如病毒激活域)或另一种人类蛋白质(如CITED2)的TAD融合。在各个实施方案中,eTF的DBD与天然存在的人类蛋白质的DBD之间的局部序列同一性为至少85%。这样的DBD可以融合至各种激活域,例如与人类蛋白质具有0%序列同一性的病毒激活域,如VP64;或CITED2激活域。与天然存在的人类蛋白质相比,在第三行中用人CITED序列代替VP64结构域提高了杂合eTF的整体序列同一性。

[0046] 图15示出了使用SCN1A特异性转录激活物,在HEK293细胞中内源SCN1A的相对表

达(参见实施例6中的表33-34)。数据以相对于对照条件的变化倍数呈现,并以 Log_{10} 尺度显示。

[0047] 图16A示出了使用SCN1A特异性转录激活物(SEQ ID NO:366),在GABA神经元中内源SCN1A的相对表达。数据以相对于对照条件(CBA-EGFP)的变化倍数呈现。

[0048] 图16B示出了使用SCN1A特异性转录激活物(SEQ ID NO:13或SEQ ID NO:15),在GABA神经元中内源SCN1A的相对表达。数据以在 Log_{10} 上相对于对照条件(CBA-EGFP)的变化倍数呈现。

[0049] 图17示出了由SCN1A特异性转录因子SEQ ID NO:366驱动的内源SCN1A和40个最近邻基因的相对表达。数据以在 Log_{10} 上相对于对照条件(CBA-EGFP-KASH)的变化倍数呈现。

[0050] 图18A和图18B示出了与表达eGFP的对照表达盒相比,SCN1A特异性转录激活物在体内的表达。图18A示出了在注射有对照eGFP表达盒或包含与靶位点结合的转录激活物的表达盒A的小鼠中,SCN1A基因的相对表达,表达盒A导致内源SCN1A基因的表达上调。图18B以平均eGFP百分比示出了SCN1A表达的变化。这些实验表明,表达盒A的转录激活物导致SCN1A表达上调约20-30%。

[0051] 图19A、图19B和图19C示出了使用若干种SCN1A特异性转录因子,野生型小鼠的海马体中内源SCN1A的上调。图19A:表达盒A;图19B:表达盒B;图19C:SEQ ID NO:365或366。将SCN1A转录因子与对照载体(EGFP-KASH)和CBA-EGFP-KASH进行比较。数据以相对于EGFP-KASH条件的变化倍数呈现。

[0052] 图20A、图20B、图20C、图20D和图20E示出了与对照相比,使用各种SCN1A特异性转录因子,在Dravet综合征小鼠模型中对高热癫痫发作的效果。向P1 Scn1a+/-小鼠输注AAV9-EGFP载体或表达SCN1A特异性转录因子(SEQ ID NO:305-309)的AAV9载体。在P26-P28,对输注的小鼠进行高热诱发的癫痫发作试验,并记录其经历强直-阵挛性癫痫发作的内部温度。

[0053] 图21A、图21B、图21C、图21D和图21E示出了在各种条件下在Dravet综合征小鼠模型中的存活。图21A示出了在存活试验中野生型与Scn1a+/-小鼠之间的比较。向P1 Scn1a+/- (N=53) 和Scn1a+/+ (N=54) 小鼠输注PBS。每天在其饲养笼中观察小鼠,如有任何死亡,则记录日期。Scn1a+/-与Scn1a+/+动物之间的存活率存在显著差异($P<0.0001$)。图21B-E示出了与对照相比,对于用各种SCN1A特异性转录因子治疗的小鼠,在Dravet综合征小鼠模型中对存活的影响。向P1 Scn1a+/-小鼠输注PBS或表达SCN1A特异性转录因子(SEQ ID NO:305-307或309)的AAV9载体。每天在其饲养笼中观察小鼠,如有任何死亡,则记录日期。

[0054] 图22示出了在实质内递送编码SCN1A特异性转录因子(SEQ ID NO:305)的AAV9载体后,在不同脑组织中的相对Scn1A mRNA表达,所述载体以 1.2×10^{12} gc/动物施用至两只食蟹猴,并相对于两只未经治疗的对照动物进行归一化。注射后28天处死所有动物,并通过Taqman PCR对组织样品中的Scn1A mRNA进行定量。数据被报告为目标mRNA在来自脑的不同组织切片中的归一化表达。用不同的一组Scn1a基因衍生的引物/探针也记录到相似的结果。

[0055] 图23示出了使用与各个基因组区域结合的工程化转录因子,内源GRN的上调。

[0056] 图24A和图24B示出了采用各种表达盒的体外实验,每个表达盒包含GRN的转录激活物。图24A示出了以相对于对照的变化倍数呈现的颗粒蛋白前体的相对表达。图24B示出

了采用颗粒蛋白前体的各种转录激活物的ELISA实验,以上清液中的hPGRN(ng/mL)表示。

[0057] 图25示出了各种调节元件(RE)对293T细胞中萤光素酶表达的归一化萤光素酶活性。例如,与minCMV启动子可操作地连接的调节元件SEQ ID NO:178驱动萤光素酶表达的水平是单独minCMV启动子的约3.5倍,并且是SCP启动子的约140倍。

[0058] 图26示出了在用表达EGFP或GRN特异性激活物(SEQ ID NO:10、16、59、60、63、64、72、73、77、L或342)的质粒转染的HEK293细胞中,通过qPCR测量的PGRN转录物的水平。数据以相对于对照条件的变化倍数呈现。

[0059] 图27示出了通过ELISA测量的,从用表达EGFP或GRN特异性激活物(SEQ ID NO:10或16)的质粒转染的HEK293细胞分泌的PGRN蛋白的水平。数据以相对于对照条件的变化倍数呈现。

[0060] 图28示出了在用CBA启动子控制下表达EGFP或GRN特异性激活物(SEQ ID NO:342)的AAV载体(血清型AAV-DJ)转染的GABA神经元中,通过qPCR测量的PGRN转录物的水平。数据以相对于对照条件的变化倍数呈现。

[0061] 图29示出了通过ELISA测量的,从以低剂量(0.5E11个基因组拷贝/孔)和高剂量(2E11个基因组拷贝/孔)用在CBA启动子控制下表达EGFP或GRN特异性激活物(SEQ ID NO:342)的AAV载体(血清型AAV-DJ)转染的GABA神经元分泌的PGRN蛋白的水平。数据以相对于对照条件的变化倍数呈现。

[0062] 图30示出了与PBS对照相比,对于以不同剂量用包含表达盒J或G(各自包含GRN的转录激活物)的AAV9处理的小鼠,以 $\mu\text{g/mL}$ 测量的血浆PGRN蛋白水平。

[0063] 图31A示出了采用各种eTF的报道分子(EGFP)激活试验的结果的定量,其被表示为与空载体对照相比的相对EGFP表达。包含强病毒激活域的人工转录因子(SEQ ID NO:51)导致70倍的相对EGFP表达。相对于对照(空载体),未修饰的人转录因子EGR1(SEQ ID NO:1)不激活报道基因。包含六个工程化锌指的衍生自EGR1的eTF(SEQ ID NO:2)导致10倍的EGFP表达。包含三个锌指、在其锌指中具有至少一个或多个氨基酸变化的衍生自EGR1的eTF(SEQ ID NO:3)导致3倍的相对EGFP表达。在EGR1指中锌指的双重复制(SEQ ID NO:4)导致15倍的相对EGFP表达。

[0064] 图31B示出了在图31A中定量的报道分子(EGFP)激活试验的荧光显微照片。表达EGFP的细胞显示为浅灰色。

[0065] 图32示出了衍生自EGR3并且被工程改造为识别导致EGFP报道基因表达的靶结合位点的eTF的各种实施方案。衍生自EGR3蛋白且被工程改造为在报道分子试验中识别TRE结合位点的包含六个锌指的eTF(SEQ ID NO:12)导致EGFP表达类似于包含衍生自VP64的TAD的衍生自EGR1的eTF(SEQ ID NO:51)。与包含经修饰的人DBD而没有激活域的对照(由SEQ ID NO:187编码)相比,这两种eTF均导致高EGFP表达。

[0066] 图33A示出了经工程改造为在EGFP报道分子激活试验中识别TRE结合序列的各种eTF的示意图。第一幅图示出了经修饰的人DBD,其被工程改造为识别TRE结合位点,但是缺少激活域(SEQ ID NO:50),其不导致EGFP表达的任何激活。在第二幅图中,将修饰的人DBD与强病毒激活域VP64(SEQ ID NO:95)融合,导致EGFP表达的激活。在第三幅和第四幅图中,将修饰的人DBD融合至人激活域,即CITED2(SEQ ID NO:96)或CITED4(SEQ ID NO:97),两者均激活EGFP表达。

[0067] 图33B示出了对于图33A中描述的每种eTF,相对EGFP表达的定量。SEQ ID NO:50导致与仅报道分子的对照相似的EGFP表达水平。与仅报道分子的对照相比,具有序列SEQ ID NO:51-53的eTF分别导致约40、15和10倍的相对EGFP表达。

[0068] 图33C示出了在图33A中显示并在图33B中定量的,每种eTF在EGFP报道分子激活试验中的荧光显微照片。

[0069] 图34示出了eTF的各种实施方案的概述,它们具有广泛的转录活性(或相对报道分子表达)和不同程度的与天然存在的人类蛋白质的序列同一性(或保守性)。

[0070] 图35A示出了用于检测本文公开的eTF的免疫原性的Elispot测定的示意图。Elispot测定用来检测来自正常健康供体的T细胞是否能够识别并响应于预测具有免疫原性的肽。T细胞与具有待测肽池的抗原呈递细胞(例如,树突细胞或DC)在elispot孔中混合。如果T细胞将肽识别为外来的,则T细胞被激活、增殖并分泌细胞因子(如干扰素- γ)。用显色试剂处理后,可以对干扰素- γ (IFN γ) 阳性斑点进行计数。与对照(例如,仅已知的非免疫原性肽或缓冲液)相比更高数目的斑点与较高的免疫原性相关。

[0071] 图35B示出了elispot结果的特写示例,其显示与对照(左侧圆圈)相比免疫原性增加(右侧圆圈)。IFN- γ 阳性斑点显示为黑色斑点。

[0072] 图36示出了包含表达盒的载体,该表达盒包含非天然存在的转录调节物,该转录调节物包含与VP64的激活域可操作地连接的转录因子的DNA结合域,用于增加SCN1A基因的表达。

[0073] 图37示出了包含表达盒的AAV载体,该表达盒包含非天然存在的转录调节物,如与VP64的转录激活域可操作地连接的核酸酶灭活的dSaCa9结构域。这样的表达盒还包含与dSaCas9转基因一起表达的指导RNA。示出的其他元件包括核定位信号(NLS)、启动子、AAV ITR、聚A信号和选择性标记。

具体实施方式

[0074] 本文提供了工程化的转录因子或eTF,它们是非天然存在的,并且已被设计成与基因组靶位点结合并调节目的内源基因的表达。可以将这样的eTF设计为上调或下调目的基因的表达(RNA和/或蛋白质表达)。

[0075] 在一方面,本申请提供了能够上调钠电压门控通道 α 亚单位1(SCN1A)基因的表达并增加其相应蛋白质产物Nav1.1的表达的eTF,以及用其治疗与Nav1.1缺乏相关的疾病或病症例如Dravet综合征的方法。

[0076] 在另一方面,本申请提供了能够上调颗粒蛋白前体(GRN)基因的表达并增加GRN蛋白的表达的eTF,以及用其治疗与GRN缺乏相关的疾病或病症例如额颞痴呆(FTD)的方法。

[0077] 在另一方面,本申请提供了与人类蛋白质具有高序列同一性百分比的eTF,该人类蛋白质可以被设计为结合基因组靶位点并调节任何目的基因的表达(上调或下调),包括例如SCN1A或GRN。相比于与人类蛋白质序列具有较低同一性百分比的eTF,这类eTF当施用于受试者时没有或几乎没有免疫原性,或具有降低的免疫原性。

[0078] 定义

[0079] 除非上下文另有明确规定,否则如本文所用的单数形式“一个”、“一种”和“该”也意欲包括复数形式。此外,就具体实施方式和/或权利要求书中使用的术语“包括”、“包含”、

“具有”或其变化形式而言,这些术语旨在以与术语“包含”类似的方式为包含性的。

[0080] 术语“约”或“大约”意指在本领域普通技术人员测定的特定值的可接受误差范围内,该可接受误差范围将部分取决于该值如何测量或测定,即,测量系统的局限性。例如,根据本领域中的实践,“约”可指在1个或大于1个标准偏差内。或者,“约”可指给定值的至多20%、至多15%、至多10%、至多5%或至多1%的范围。

[0081] 术语“确定”、“测量”、“评价”、“评估”、“测定”、“分析”及其语法等同语在本文中可互换使用,以指代任何测量形式,并且包括确定元素是否存在(例如,检测)。这些术语可包括定量和/或定性测定。评估可以是相对的或绝对的。

[0082] 术语“表达”是指核酸序列或多核苷酸从DNA模板转录(如转录成mRNA或其他RNA转录物)的过程,和/或所转录的mRNA随后被翻译成肽、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”。如果多核苷酸衍生自基因组DNA,则表达可以包括真核细胞中mRNA的剪接。

[0083] 如本文所用的,“可操作地连接的”、“可操作的连接”、“有效连接的”或其语法等同语是指例如启动子、增强子、多腺苷酸化序列等遗传元件并列,其中这些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/或增强子序列的调节元件帮助启动编码序列的转录,则该调节元件与编码区可操作地连接。在调节元件与编码区之间可以有间隔残基,只要能够维持这种功能关系即可。

[0084] 如本文所用的“载体”是指包含多核苷酸或与多核苷酸缔合,且可用于介导多核苷酸向细胞的递送的大分子或大分子缔合物。载体的实例包括质粒、病毒载体、脂质体和其他基因递送媒介物。载体通常包含与基因可操作地连接的遗传元件,例如调节元件,以促进该基因在靶标中的表达。

[0085] 如本文所用的,“表达盒”和“核酸盒”可互换使用,是指一起表达或可操作地连接以供表达的核酸序列或元件的组合。在一些情况下,表达盒是指调节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。

[0086] 术语“AAV”是腺伴随病毒的缩写,可用来指代该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及其杂合体,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蛋白和衣壳亚单位的序列是本领域已知的。这类序列可见于文献中或公共数据库如GenBank中。如本文所用的“rAAV载体”是指包含非AAV来源的多核苷酸序列(即,与AAV相比异源的多核苷酸)的AAV载体,该多核苷酸序列一般是用于细胞的遗传转化的目的序列。通常,异源多核苷酸的侧翼为至少一个、通常两个AAV反向末端重复序列(ITR)。rAAV载体可以是单链的(ssAAV)或自互补的(scAAV)。“AAV病毒”或“AAV病毒颗粒”是指由至少一个AAV衣壳蛋白和包壳的多核苷酸rAAV载体组成的病毒颗粒。如果该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,如待递送至哺乳动物细胞的转基因),则其一般被称为“rAAV病毒颗粒”或简称为“rAAV颗粒”。因此,rAAV颗粒的产生必然包括rAAV载体的产生,因为这样的载体包含在rAAV颗粒内。

[0087] 如本文所用的,术语“治疗”、“处理”、“疗法”等是指减轻、延迟或减慢疾病或病症

的进展,预防、减弱、减少影响或症状,防止其发作,抑制或减轻其发作。本公开的方法可用于任何哺乳动物。示例性的哺乳动物包括但不限于大鼠、猫、狗、马、牛、绵羊、猪,更优选人。治疗性益处包括正在被治疗的潜在疾病的根除或改善。另外,治疗性益处也可以如下实现:一种或多种与该基础病症相关的生理学症状得到根除或改善,使得在受试者中观察到起色,虽然该受试者可能仍受基础病症的折磨。在一些情况下,对于预防性益处,可以将治疗剂施用于具有发展成特定疾病的风险的受试者或报告了疾病的一个或多个生理学症状的受试者,即使可能尚未作出该疾病的诊断。本公开的方法可用于任何哺乳动物。在一些情况下,治疗可以导致症状的减少或停止(例如,癫痫发作频率、持续时间和/或严重程度的降低)。预防性效果包括延缓或消除疾病或病况的出现,延缓或消除疾病或病况的症状发作,减慢、终止或逆转疾病或病况的进展,或其任意组合。

[0088] 术语“有效量”或“治疗有效量”是指足以实现预期应用的本文所述组合物的量,该预期应用包括但不限于如下文定义的疾病治疗。治疗有效量可根据预期的治疗应用(体内),或所治疗的受试者和疾病状况,例如,受试者的体重和年龄、疾病状况的严重程度、给药方式等而变化,这些可以容易地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择的特定组合物、所遵循的给药方案、是否与其他化合物联合施用、给药时机、所施用的组织、携带该组合物的物理递送系统而变化。

[0089] 核苷酸或肽序列的“片段”是指短于参考或“全长”序列的序列。

[0090] 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。

[0091] DNA或蛋白质序列的“功能片段”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。

[0092] 术语“受试者”和“个体”在本文中可互换使用,是指脊椎动物,优选哺乳动物,更优选人。本文所述的方法可用于人类治疗、兽医应用和/或在疾病或病况的动物模型中的临床前研究。

[0093] 术语“体内”是指在受试者的身体中发生的事件。

[0094] 术语“体外”是指在受试者的身体外发生的事件。例如,体外试验包括在受试者体外进行的任何试验。体外试验包括其中采用活细胞或死细胞的基于细胞的试验。体外试验还包括其中不采用完整细胞的无细胞试验。

[0095] 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分别的确切核苷酸-核苷酸对应关系或氨基酸-氨基酸对应关系。一般而言,用于确定序列同一性的技术包括比较两个核苷酸或氨基酸序列,并确定其同一性百分比。例如为了评估同一性,序列比较可以通过任何合适的比对算法来进行,包括但不限于Needleman-Wunsch算法(参见,例如,可以在www.ebi.ac.uk/Tools/psa/emboss_needle/获得的EMBOSS Needle比对器(aligner),任选地采用默认设置)、BLAST算法(参见,例如,可以在blast.ncbi.nlm.nih.gov/Blast.cgi获得的BLAST比对工具,任选地采用默认设置)和Smith-Waterman算法(参见,例如,可以在www.ebi.ac.uk/Tools/psa/emboss_water/获得的EMBOSS Water比对器,任选地采用默认设置)。可以使用所选算法的任何合适的参数(包括默认参数)来评估最佳比对。两个序列之间的“同一性百分比”,也称为“同源性百分比”,

可被计算为两个最佳比对的序列之间的精确匹配数除以参考序列的长度并乘以100。例如,也可以通过使用可从国立卫生研究院(National Institutes of Health)得到的包括2.2.9版本在内的高级BLAST计算机程序比较序列信息来确定同一性百分比。BLAST程序基于Karlin和Altschul, *Proc. Natl. Acad. Sci. USA* 87:2264-2268 (1990) 的比对方法, 并如以下文献所讨论: 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)。简言之, BLAST程序将同一性定义为相同的比对标志(即核苷酸或氨基酸)的数目除以两个序列中较短者中的标志的总数。可以使用该程序来确定在所比较的序列的总长度上的同一性百分比。提供默认参数来优化采用短查询序列的搜索, 例如采用blastp程序。该程序还允许使用SEG过滤器来屏蔽通过Wootton和Federhen, *Computers and Chemistry* 17:149-163 (1993) 的SEG程序确定的查询序列区段。高序列同一性通常包括大约80%至100%的序列同一性范围, 及其之间的整数值。

[0096] 如本文所用的, 关于蛋白质的“工程化的”是指非天然存在的蛋白质, 包括但不限于衍生自天然存在的蛋白质的蛋白质, 或其中天然存在的蛋白质已被修饰或重新编程以具有某种性质的蛋白质。

[0097] 如本文所用的, “合成的”和“人工的”可互换使用, 是指与天然存在的人类蛋白质具有低序列同一性(例如, 小于50%的序列同一性)的蛋白质或其结构域。例如, VPR和VP64结构域是合成的反式激活域。

[0098] 如本文所用的, “工程化转录因子”或“eTF”是指经修饰或重新编程以结合特定靶结合位点和/或包括经修饰或替换的转录效应物结构域的非天然存在的DNA结合蛋白或非天然存在的转录调节物。

[0099] 如本文所用的, “DNA结合域”可以用来指代一个或多个DNA结合基序, 如锌指或碱性螺旋-环-螺旋(bHLH)基序, 其单独地或共同地作为DNA结合蛋白的一部分。

[0100] 术语“转录激活域”、“转录激活结构域”、“反式激活域”、“反式激活结构域”和“TAD”在本文中可互换使用, 并且是指这样的蛋白质结构域, 它与DNA结合域一起, 可以通过直接地或通过被称为共激活物的其他蛋白质接触转录机制(例如, 通用转录因子和/或RNA聚合酶)而激活从启动子的转录。

[0101] 术语“转录阻抑域”、“转录阻抑结构域”和“TRD”在本文中可互换使用, 并且是指这样的蛋白质结构域, 它与DNA结合域一起, 可以通过直接地或通过被称为共阻抑物的其他蛋白质接触转录机制(例如, 通用转录因子和/或RNA聚合酶)而阻抑从启动子的转录。

[0102] 术语“GRCh38.p12”是指Genome Reference Consortium Human Build38补丁版本12(GRCh38.p12), 其GenBank Assembly登录号为GCA_000001405.27, 日期为2017/12/21。

[0103] 除非另有说明, 否则本文使用的所有术语均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, 这些技术均在本领域技术人员的知识范畴内。

[0104] 工程化转录因子(eTF)

[0105] 转录因子(TF)是结合基因组中的特定序列并控制基因表达的蛋白质。本文提供的工程化转录因子或eTF是非天然存在的蛋白质, 其包含DNA结合域(DBD)和至少一个作为转录调节物的结构域, 例如转录激活域(TAD)或转录阻抑物结构域(TRD)。在一个实施方案中,

eTF可包含DBD和TAD(例如,TAD-DBD或DBD-TAD),其中该DBD和TAD可衍生自相同的蛋白质或不同的蛋白质。在另一个实施方案中,eTF可包含DBD和两个TAD,其中该DBD和TAD衍生自相同的蛋白质,该DBD衍生自第一蛋白质且两个TAD均衍生自第二蛋白质,该DBD和一个TAD衍生自第一蛋白质且第二个TAD衍生自第二蛋白质,或者该DBD衍生自第一蛋白质、一个TAD衍生自第二蛋白质且第二个TAD衍生自第三蛋白质(例如,TAD1-DBD-TAD1或TAD1-DBD-TAD2)。在一个实施方案中,eTF可包含DBD和TRD(例如,TRD-DBD或DBD-TRD),其中该DBD和TRD可衍生自相同的蛋白质或不同的蛋白质。在另一个实施方案中,eTF可包含DBD和两个TRD,其中该DBD和TRD衍生自相同的蛋白质,该DBD衍生自第一蛋白质且两个TRD均衍生自第二蛋白质,该DBD和一个TRD衍生自第一蛋白质且第二个TRD衍生自第二蛋白质,或者该DBD衍生自第一蛋白质、一个TRD衍生自第二蛋白质且第二个TRD衍生自第三蛋白质(例如,TRD1-DBD-TRD1或TRD1-DBD-TRD2)。在某些实施方案中,该DBD可以是包含来自多种蛋白质的结构域的合成构建体。

[0106] 在某些实施方案中,DBD和TAD可以直接缀合,例如,没有间插氨基酸序列。在其他实施方案中,可以使用肽连接体将DBD和TAD缀合。在某些实施方案中,DBD经由具有1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、60、70、75、80、90或100个氨基酸或者1-5、1-10、1-20、1-30、1-40、1-50、1-75、1-100、5-10、5-20、5-30、5-40、5-50、5-75、5-100、10-20、10-30、10-40、10-50、10-75、10-100、20-30、20-40、20-50、20-75或20-100个氨基酸的连接体与TAD缀合。在一些情况下,DBD和TAD经由在衍生出所述结构域的天然存在的蛋白质中发现的天然存在的间插残基缀合。在其他实施方案中,DBD和TAD经由合成的或外源的连接体序列缀合。合适的连接体可以是柔性的、可切割的、不可切割的、亲水的和/或疏水的。在某些实施方案中,DBD和TAD可经由包含多个甘氨酸和/或丝氨酸残基的连接体融合在一起。甘氨酸/丝氨酸肽连接体的实例包括[GS]_n、[GGGS]_n(SEQ ID NO: 436)、[GGGS]_n(SEQ ID NO: 437)、[GGSG]_n(SEQ ID NO: 438),其中n为等于或大于1的整数。在某些实施方案中,可用于缀合DBD和TAD的连接体是GGSGGGSG(SEQ ID NO: 410)。在某些实施方案中,当DBD缀合至两个TAD时,第一TAD和第二TAD可以通过相同或不同的连接体缀合至DBD,或者一个TAD可以通过连接体缀合至DBD,而另一个TAD直接缀合至DBD(例如,没有间插连接体序列),或者两个TAD可以直接缀合至DBD(例如,没有间插连接体序列)。

[0107] 本文提供的eTF具有与天然存在的转录因子不同的性质。在某些实施方案中,本文提供的eTF包含衍生自天然存在的蛋白质的DBD,该DBD已被修饰,使得与衍生出该DBD的天然存在的蛋白质相比,该DBD与不同的靶位点结合,与衍生出该DBD的天然存在的蛋白质相比,包含这种修饰的DBD的eTF调节从不同基因的表达,并且/或者与衍生出该DBD的天然存在的蛋白质相比,包含这种修饰的DBD的eTF不同地调节靶基因的表达(例如,上调与下调)。在其他实施方案中,本文提供的eTF包含衍生自天然存在的蛋白质的TAD,该TAD已被修饰,使得与衍生出该TAD的天然存在的蛋白质相比,包含这种修饰的TAD的eTF调节从不同基因的表达,并且/或者与衍生出该TAD的天然存在的蛋白质相比,包含这种修饰的TAD的eTF不同地调节靶基因的表达(例如,上调与下调)。在某些实施方案中,本文提供的eTF包含衍生自天然存在的蛋白质的DBD和衍生自天然存在的蛋白质(相同或不同的蛋白质)的TAD,其中该DBD和TAD均已被修饰。在这样的实施方案中,与衍生自天然存在的蛋白质相比,所述DBD可以与不同的靶位点结合,与衍生出结构域的天然存在的蛋白质相比,包含这种修

饰的DBD和TAD的eTF调节从不同基因的表达,并且/或者与衍生出DBD和TAD结构域的天然存在的蛋白质相比,包含这种修饰的DBD和TAD的eTF不同地调节靶基因的表达(例如,上调与下调)。

[0108] DNA结合域 (DBD)

[0109] 本文提供的eTF可包含与目的靶位点结合的任何合适的DBD。在某些实施方案中,该DBD可以是合成设计的DBD。在其他实施方案中,该DBD可以衍生自天然存在的蛋白质。DBD家族包括碱性螺旋-环-螺旋(bHLH)(例如,c-Myc)、碱性亮氨酸拉链(例如,C/EBP)、螺旋-转角-螺旋(例如,Oct-1)和锌指(例如,EGR3)。这些家族表现出广泛的DNA结合特异性和基因靶标。如本文所预期的,任何一种已知的人转录因子蛋白质都可以充当蛋白质平台,用于工程化和/或重新编程DBD,以识别特异性靶位点,从而导致目的内源基因表达的调节。在示例性实施方案中,本文提供的DBD包含锌指结构域、TALEN结合域或gRNA/Cas复合物。

[0110] 本文提供的DBD可以被设计为识别任何目的靶位点。在示例性实施方案中,DBD被工程改造为识别当被eTF结合时能够调节(例如,上调或下调)从目的基因的表达的靶位点。在示例性实施方案中,DBD被设计成识别基因组位置并当被eTF结合时调节内源基因的表达。当被eTF结合时能够调节目的内源基因表达的结合位点可以位于基因组中导致调节靶基因的基因表达的任何位置。在各个实施方案中,该结合位点可以位于与目的基因不同的染色体上、与目的基因相同的染色体上、在目的基因的转录起始位点(TSS)的上游、在目的基因的TSS的下游、在目的基因的TSS的近端、在目的基因的远端、在目的基因的编码区内、在目的基因的内含子内、在目的基因的聚A尾的下游、在调节目的基因的启动子序列内、在调节目的基因的增强子序列内或在调节目的基因的阻抑物序列内。

[0111] DBD可以被设计成与任何长度的靶结合位点结合,只要它提供DBD对靶结合位点序列的特异性识别,例如具有最小的或没有脱靶结合。在某些实施方案中,与其他所有基因相比,当被eTF结合时,靶结合位点可以以至少2倍、5倍、10倍、20倍、50倍、75倍、100倍、150倍、200倍、250倍、500倍或更高倍数的水平调节目的内源基因的表达。在某些实施方案中,与40个最近邻基因(例如,位于染色体上最靠近目的基因的编码序列上游或下游的40个基因)相比,当被eTF结合时,靶结合位点可以以至少2倍、5倍、10倍、20倍、50倍、75倍、100倍、150倍、200倍、250倍、500倍或更高倍数的水平调节目的内源基因的表达。在某些实施方案中,靶结合位点可以是至少5bp、10bp、15bp、20bp、25bp、30bp、35bp、40bp、45bp或50bp或更大。结合位点的具体长度将取决于eTF中的DBD类型。通常,结合位点的长度越长,对基因表达的结合和调节的特异性就越高(例如,较长的结合位点具有较低的脱靶作用)。在某些实施方案中,相对于使用具有与较短靶位点结合的DBD的eTF所观察到的脱靶效应,具有识别较长靶结合位点的DBD的eTF具有较低的与非特异性结合相关的脱靶效应(例如,调节脱靶基因或目的基因以外的其他基因的表达)。在一些情况下,脱靶结合的减少是与具有识别较短靶结合位点的DBD的相当eTF相比低至少1.2、1.3、1.4、1.5、2、3、4、5、6、7、8、9或10倍。

[0112] 在某些实施方案中,本文提供的DBD可以被修饰,以具有增加的结合亲和力,使其与更长的靶结合位点结合,使得与该DBD缀合的TAD能够募集更多转录因子和/或募集这类转录因子更长时间,以对内源目的基因的表达水平产生更大的影响。在某些实施方案中,可以对eTF进行工程改造,使得DBD稳定地结合靶结合位点并阻断转录或表达内源基因所需的转录机制。在某些实施方案中,可以修饰DBD以增加其与所需靶位点的特异性结合(或靶上

结合),和/或修饰DBD以减少其非特异性或脱靶结合。

[0113] 在各个实施方案中,可以使用各种方法确定DBD或eTF与靶结合位点之间的结合。在某些实施方案中,可以使用迁移率变化测定、DNA酶保护测定或本领域已知用于测定蛋白质-DNA结合的其他任何体外方法来确定DBD或eTF与靶结合位点之间的特异性结合。在其他实施方案中,可以使用功能测定,例如,通过测量当靶结合位点被eTF结合时基因的表达(RNA或蛋白质)来确定eTF与靶结合位点之间的特异性结合。例如,靶结合位点可以位于细胞中包含的或整合到细胞基因组中的载体上的报道基因(例如,eGFP)或目的靶基因的上游,其中该细胞表达eTF。与对照(例如,没有eTF或识别不同靶位点的eTF)相比,在eTF的存在下报道基因(或目的靶基因)的表达水平更高,表明该eTF的DBD与靶位点结合。合适的体外(例如,非基于细胞的)转录和翻译系统也可以以类似的方式使用。在某些实施方案中,与对照(例如,无eTF或识别不同靶位点的eTF)相比,与靶位点结合的eTF可具有至少2倍、3倍、5倍、10倍、15倍、20倍、30倍、50倍、75倍、100倍、150倍或更高的报道基因或目的靶基因的表达。

[0114] 在其他实施方案中,本文提供的eTF可以包含DBD,该DBD来自锌指蛋白,衍生自锌指蛋白,或者是核酸酶灭活的锌指蛋白。锌指是一种小的蛋白质结构基序,其特征在于一个或多个锌离子(Zn^{2+})的配位,以便稳定折叠。锌指(Znf)结构域是相对较小的蛋白质基序,其含有多个导致与DNA靶位点串联接触的手指状突起。锌指基序的模块性质允许大量的DNA序列组合以高度亲和力和特异性结合,因此理想地适合于工程改造可被靶向并结合特定DNA序列的蛋白质。许多工程化的锌指阵列基于鼠转录因子Zif268的锌指结构域。Zif268具有三个单独的锌指基序,它们共同地以高亲和力结合9bp序列。如本文进一步描述的,已经基于结构鉴定了多种锌指蛋白并将其表征为不同的类型。任何此类锌指蛋白都可与本文所述的DBD结合使用。

[0115] 多种设计锌指蛋白的方法是可用的。例如,描述了设计锌指蛋白以结合目的靶DNA序列的方法,参见,例如,Liu Q等人,Design of polydactyl zinc-finger proteins for unique addressing within complex genomes,Proc Natl Acad Sci USA.94(11):5525-30(1997);Wright DA等人,Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly,Nat Protoc.Nat Protoc.2006;1(3):1637-52;以及CA Gersbach和T Gaj,Synthetic Zinc Finger Proteins:The Advent of Targeted Gene Regulation and Genome Modification Technologies,Am Chem Soc 47:2309-2318(2014)。另外,用于设计锌指蛋白以与目的DNA靶序列结合的各种基于Web的工具是可公开获得的,参见,例如,可从万维网地址omictools.com/zfns-category处找到的OmicX的Zinc Finger Nuclease Design Software Tools and Genome Engineering Data Analysis网址;和可从万维网地址scripps.edu/barbas/zfdesign/zfdesignhome.php处找到的Scripps的Zinc Finger Tools设计网址。另外,用于设计锌指蛋白以与目的DNA靶序列结合的各种市售服务是可获得的,参见,例如,Creative Biolabs提供的市售服务或试剂盒(万维网地址creative-biolabs.com/Design-and-Synthesis-of-Artificial-Zinc-Finger-Proteins.html),可从Addgene获得的Zinc Finger Consortium Modular Assembly Kit(万维网地址addgene.org/kits/zfc-modular-assembly/),或来自Sigma Aldrich的CompoZr Custom ZFN Service(万维网地址sigmaaldrich.com/life-science/

zinc-finger-nuclease-technology/custom-zfn.html)。

[0116] 在某些实施方案中,本文提供的eTF包含含有一个或多个锌指的DBD,或衍生自锌指蛋白的DBD。在一些情况下,该DBD包含多个锌指,其中每个锌指在其N-末端或C-末端或均通过氨基酸连接体与另一个锌指或另一个结构域连接。在一些情况下,本文提供的DBD包含来自表5中描述的一种或多种锌指类型的一个或多个锌指。在一些情况下,本文提供的DBD包含多个锌指结构或基序,或多个具有SEQ ID NO:115-130或141-164中的一个或多个或其任何组合的锌指。在某些实施方案中,DBD包含 $X-[ZF-X]_n$ 和/或 $[X-ZF]_n-X$,其中ZF是具有表5中列出的任一基序(例如,SEQ ID NO:39-49中的任一个)的锌指结构域,X是包含1-50个氨基酸的氨基酸连接体,并且n是1-15的整数,例如1、2、3、4、5、6、7、8、9、10、11、12、13、14或15,其中每个ZF可以独立地具有与该DBD中的其他ZF序列相同的序列或不同的序列,并且其中每个连接体X可以独立地具有与该DBD中的其他X序列相同的序列或不同的序列。每个锌指可以在其C-末端、N-末端或两端与另一个序列、锌指或结构域连接。在DBD中,每个连接体X可以在序列、长度和/或性质(例如,柔性或电荷)上相同,或者在序列、长度和/或性质上不同。在一些情况下,两个或更多个连接体可以相同,而其他连接体则不同。在示例性实施方案中,该连接体可以获自或衍生自连接表5中提供的一种或多种天然存在的锌指蛋白中发现的锌指的序列。在其他实施方案中,合适的连接体序列包括例如长度为5个或更多个氨基酸的连接体。关于长度为6个或更多个氨基酸的示例性连接体序列,还参见美国专利6,479,626、6,903,185和7,153,949,其中每一篇均整体并入本文。本文提供的DBD蛋白可以包括该蛋白质的各个锌指之间的合适连接体的任何组合。本文描述的DBD蛋白可以包括该蛋白质的各个锌指之间的合适连接体的任何组合。

[0117] 在某些实施方案中,本文提供的DBD包含一个或多个经典锌指。经典的C2H2锌指在一条链中具有两个半胱氨酸,在另一条链中具有两个组氨酸残基,它们由锌离子配位。经典的锌指结构域具有两个 β -折叠和一个 α -螺旋,其中 α -螺旋与DNA分子相互作用并形成DBD与靶位点结合的基础,并且可以被称为“识别螺旋”。在示例性实施方案中,锌指的识别螺旋在位置-1、2、3或6处包含至少一个氨基酸置换,从而改变锌指结构域的结合特异性。在其他实施方案中,本文提供的DBD包含一个或多个非经典锌指,例如,C2-H2、C2-CH和C2-C2。

[0118] 在另一个实施方案中,本文提供的DBD结构域包含具有以下结构的锌指基序: LEPGEKP-[YKCPECGKSFS X HQRTH TGEKP] $_n$ -YKCPECGKSFS X HQRTH-TGKKTS (SEQ ID NO: 421),其中n为1-15的整数,例如1、2、3、4、5、6、7、8、9、10、11、12、13、14或15,并且每个X独立地是能够与靶序列的3bp结合的识别序列(例如识别螺旋)。在示例性实施方案中,n为3、6或9。在特别优选的实施方案中,n为6。在各个实施方案中,与DBD中的其他X序列相比,每个X可以独立地具有相同的氨基酸序列或不同的氨基酸序列。在示例性实施方案中,每个X是包含7个氨基酸的序列,该序列已经使用万维网地址scripps.edu/barbas/zfdesign/zfdesignhome.php处的Scripps的Zinger Finger Design Tool设计为与目的靶结合位点的3bp相互作用。

[0119] 由于DBD内的每个锌指都识别3bp,因此DBD中包含的锌指的数目会告知被DBD识别的结合位点的长度,例如,具有1个锌指的DBD将识别具有3bp的靶结合位点,具有2个锌指的DBD将识别具有6bp的靶结合位点,具有3个锌指的DBD将识别具有9bp的靶结合位点,具有4个锌指的DBD将识别具有12bp的靶结合位点,具有5个锌指的DBD将识别具有15bp的靶结合位点。

点,具有6个锌指的DBD将识别具有18bp的靶结合位点,具有9个锌指的DBD将识别具有27bp的靶结合位,依此类推。通常,识别更长靶结合位点的DBD将表现出更高的结合特异性(例如,更少的脱靶结合或非特异性结合)。

[0120] 在其他实施方案中,本文提供的DBD通过在锌指结构域的一个或多个识别螺旋中进行一个或多个氨基酸置换而衍生自天然存在的锌指蛋白,以便改变DBD的结合特异性(例如,改变被DBD识别的靶位点)。本文提供的DBD可以衍生自任何天然存在的锌指蛋白。在各个实施方案中,这样的DBD可以衍生自诸如小鼠、大鼠、人等任何物种的锌指蛋白。在示例性实施方案中,本文提供的DBD衍生自人锌指蛋白。在某些实施方案中,本文提供的DBD衍生自表5中列出的天然存在的蛋白质。在示例性实施方案中,本文提供的DBD蛋白衍生自人EGR锌指蛋白,例如,EGR1、EGR2、EGR3或EGR4。

[0121] 在某些实施方案中,本文提供的DBD通过修饰DBD而衍生自天然存在的蛋白质,该修饰通过在天然存在的蛋白质的DBD内重复一个或多个锌指而增加DBD蛋白中的锌指结构域的数目。在某些实施方案中,这样的修饰包括在天然存在的蛋白质的DBD内锌指的双重复制、三重复制、四重复制或进一步的倍增。在一些情况下,来自人类蛋白质的DBD的一个锌指被复制,例如,2、3、4、5、6、7、8、9、10、11、12个或更多个拷贝的相同锌指基序在eTF的DBD中被重复。在一些情况下,来自天然存在的蛋白质的DBD的一组锌指被复制,例如,来自天然存在的蛋白质的DBD的一组3个锌指被双重复制,以产生包含具有6个锌指的DBD的eTF,被三重复制,以产生具有9个锌指的eTF的DBD,或被四重复制,以产生具有12个锌指的eTF的DBD,依此类推。在一些情况下,来自天然存在的蛋白质的DBD的一组锌指被部分复制,以形成具有更大数目的锌指的eTF的DBD,例如,eTF的DBD包含四个锌指,其中所述锌指代表来自天然存在的蛋白质的第一锌指的一个拷贝、第二锌指的一个拷贝和第三锌指的两个拷贝,在eTF的DBD中共有四个锌指。然后通过锌指结构域的一个或多个识别螺旋中进行一个或多个氨基酸置换而进一步修饰这样的DBD,以便改变DBD的结合特异性(例如,改变被DBD识别的靶位点)。在示例性实施方案中,该DBD衍生自天然存在的人类蛋白质,如人EGR锌指蛋白,例如,EGR1、EGR2、EGR3或EGR4。

[0122] 人EGR1和EGR3的特征在于三指C2H2锌指DBD。锌指的通用结合规则规定,所有三个指与其具有相似几何形状的同源DNA序列相互作用,使用每个锌指的 α 螺旋中的相同氨基酸来确定靶结合位点序列的特异性或识别。这样的结合规则允许人们修饰EGR1或EGR3的DBD,以工程改造识别所需靶结合位点的DBD。在一些情况下,根据已发布的锌指设计规则,对EGR1或EGR3的锌指基序中的7个氨基酸的DNA识别螺旋进行修饰。在某些实施方案中,例如通过改变一个或多个识别螺旋的序列和/或通过增加DBD中锌指的数目来修饰EGR1或EGR3的三指DBD中的每个锌指。在某些实施方案中,将EGR1或EGR3重新编程,以识别在所需靶位点处至少9、12、15、18、21、24、27、30、33、36个或更多个碱基对的靶结合位点。在某些实施方案中,这类衍生自EGR1或EGR3的DBD包含至少3、4、5、6、7、8、9、10、11、12个或更多个锌指。在示例性实施方案中,该DBD中的一个或多个锌指在识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。

[0123] 在各个实施方案中,包含衍生自EGR1或EGR3的DBD的eTF具有与天然存在的EGR1或EGR3的结合特异性不同的DNA结合特异性,例如,该DBD识别具有与被未修饰的EGR1或EGR3识别的结合位点的以下序列不同的序列的靶结合位点:(GCG (T/G) GGGCG) (SEQ ID NO:

373)。

[0124] 在其他实施方案中,本文提供的eTF包含作为gRNA/Cas复合物的DBD。CRISPR(成簇的规律间隔的短回文重复序列)/Cas9是一种基因组编辑工具,其允许进行位点特异性基因组靶向。II型CRISPR/Cas系统是一种原核适应性免疫应答系统,其使用非编码RNA指导Cas9核酸酶诱导位点特异性DNA切割。已经利用CRISPR/Cas9系统创建了一种简单的、RNA可编程的方法来介导哺乳动物细胞中的基因组编辑。可以产生单个指导RNA(sgRNA),以将Cas9核酸酶引导至特定的基因组位置,然后被gRNA/Cas9复合物结合。可以使用多种方法和工具将gRNA设计为与目的靶位点结合。例如,用于设计gRNA以结合目的靶DNA序列的方法描述于Aach等人,Flexible algorithm for identifying specific Cas9 targets in genomes.BioRxiv,Cold Spring Harbor Labs.doi:http://dx.doi.org/10.1101/005074(2014);Bae等人,Cas-OFFinder:a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases.Bioinformatics.30(10):1473-1475(2014);Doench,J.G.等人,Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9.Nat Biotech 34,184-191(2016);Gratz等人,Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila.Genetics.196(4):961-971(2014);Heigwer等人,E-CRISP:fast CRISPR target site identification.Nat Methods.11(2):122-123(2014);Ma等人,A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes.Biomed Res Int.doi:http://doi.org/10.1155/2013/270805(2013);Montague等人,CHOPCHOP:a CRISPR/Cas9 and TALEN web tool for genome editing.Nucleic Acids Res.42(W1):W401-W407(2014);Liu等人,CRISPR-ERA:a comprehensive design tool for CRISPR-mediated gene editing,repression and activation.Bioinformatics.31(22):3676-3678(2015);Ran等人,In vivo genome editing using Staphylococcus aureus Cas9.Nature.520(7546):186-191(2015);Wu等人,Target specificity of the CRISPR-Cas9 system.Quant Biol.2(2):59-70(2015);Xiao等人,CasOT:a genome-wide Cas9/gRNA off-target searching tool.Bioinformatics.30(8):1180-1182(2014);Zetsche等人,Cpf1 is a single RNA-guided endonuclease of a Class 2CRISPR-Cas System.Cell.163(3):759-771(2015)。另外,用于设计gRNA以结合目的DNA靶序列的各种基于Web的工具是可公开获得的,参见,例如,可从万维网地址atum.bio/eCommerce/cas9/input?multipleContacts=false处找到的AUTM的CRISPR gRNA Design工具;可从万维网地址portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design-crisprai处找到的Broad Institute的CRISPRa/i gRNA设计工具;可从万维网地址e-crisp.org/E-CRISP/处找到的DKFZ German Cancer Research Center的E-CRISP设计工具;以及可从万维网地址design.synthego.com/#/处找到的Synthego的Knockout Guide Design工具。另外,用于设计gRNA以结合目的DNA靶序列的各种可商购服务是可以获得的,参见,例如,IDT(万维网地址idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE)、ThermoFisher(万维网地址thermofisher.com/order/custom-oligo/crispr)和GenScript(万维网地址genscript.com/gRNA-design-tool.html)提供的可商购服务。

[0125] 在示例性实施方案中,作为gRNA/Cas复合物的DBD包含核酸酶灭活的Cas蛋白或dCas,例如dCas9,如核酸酶灭活的金黄色葡萄球菌Cas9(dSaCas9)或核酸酶灭活的酿脓链球菌Cas9(dSpCas9)。提供gRNA作为包含靶向区域和支架区域的序列,该靶向区域将gRNA/Cas复合物靶向所需的靶位点,而该支架区域促进与Cas蛋白的相互作用。任何合适的gRNA支架可以与本文提供的gRNA一起使用。在示例性实施方案中,该gRNA是单个gRNA或sgRNA,并包含以下支架序列:5'-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTGGCGAGA-3'(SEQ ID NO:102)。指导RNA的靶向区域连接到支架序列的5'端,以形成完整的sgRNA。在某些实施方案中,gRNA和dCas蛋白可以从同一表达盒表达。在某些实施方案中,U6启动子用来表达gRNA。在其他实施方案中,例如,通过将dCas稳定地整合到基因组中或在染色体外稳定维持的质粒上,可以在经工程改造以稳定表达dCas-TAD蛋白的细胞中表达gRNA。

[0126] 在其他实施方案中,本文提供的eTF可以包含DBD,该DBD来自TALEN,衍生自TALEN,或者是核酸酶灭活的TALEN。转录激活物样效应物核酸酶(TALEN)是包含DBD和可被工程改造以切割特定DNA序列的核酸酶结构域的限制酶。通过将TAL效应物DNA结合域与DNA切割结构域(例如核酸酶)缀合来创建TALEN。转录激活物样效应物(TALE)可以被工程改造为与所需的靶DNA序列结合,从而将核酸酶结构域引导至特定位置。

[0127] TAL效应物是来自黄单胞菌属(xanthomonas)细菌的细菌蛋白质。DNA结合域包含重复的高度保守的33-34个氨基酸的序列,具有不同的第12个和第13个氨基酸。这两个位置,被称为重复可变二残基(RVD),是高度可变的,并且显示出与特定核苷酸识别的强相关性。氨基酸序列与DNA识别之间的这种直接关系允许通过选择含有适当RVD的重复区段的组合来工程改造特异性靶向所需序列的DBD。

[0128] 多种设计TALE的方法是可用的。例如,用于设计TALE以与目的靶DNA序列结合的方法描述于T.Cermak等人,Nucleic Acids Research.39(12):e82(2011);F.Zhang F等人,Nature Biotechnology.29(2):149-53(2011);R.Morbitzer等人,Nucleic Acids Research.39(13):5790-9(2011);T.Li等人,Nucleic Acids Research.39(14):6315-25(2011);R.Geissler等人,PLOS One.6(5):e19509(2011);和E.Weber等人,PLOS One.6(5):e19722(2011)。另外,用于设计TALE以与目的DNA靶序列结合的各种基于Web的工具是可公开获得的,参见,例如,可从万维网地址e-talen.org/E-TALEN/TAL处找到的E-Talen,以及可从万维网地址tale-nt.cac.cornell.edu/node/add/single-tale处找到的Effector Nucleotide Targeter 2.0工具。另外,用于设计TALE以与目的DNA靶序列结合的各种市售服务是可获得的,参见,例如,OmicX(万维网地址omictools.com/),Addgene(万维网地址addgene.org/talen/guide/)或ThermoFisher(万维网地址thermofisher.com/us/en/home/life-science/genome-editing/geneart-tals/tal-design-tool.html)提供的市售服务。另外,可以使用可公开获得的软件程序(DNAWorks)设计适合于装配TALE的寡核苷酸,参见,例如,D.Hoover D Methods in Molecular Biology.852:215-23(2012)。

[0129] 转录调节域

[0130] 本文提供的eTF可包含能够募集一种或多种蛋白质因子的任何合适的结构域,当eTF通过DBD(例如,锌指DBD、gRNA/Cas DBD或TALE DBD)与靶位点结合时,该蛋白质因子可以调节从目的基因的转录(例如,RNA聚合酶II、CBP/p300、CREB或KRAB)或基因表达水平。在

某些实施方案中,这样的结构域募集增加目的基因的转录或基因表达水平的蛋白质因子,并且是转录激活域(TAD)。在其他实施方案中,这样的结构域募集降低目的基因的转录或基因表达水平的蛋白质因子,并且是转录阻抑结构域(TRD)。在某些实施方案中,该转录调节域(TAD或TRD)可以是合成设计的结构域。在其他实施方案中,该转录调节域(TAD或TRD)可以衍生自天然存在的蛋白质,例如,转录因子、转录共激活物、转录共阻抑物或沉默子蛋白质。在各个实施方案中,转录调节域(TAD或TRD)可以衍生自诸如小鼠、大鼠、猴、病毒或人等任何物种的蛋白质。

[0131] 在一个示例性实施方案中,适于在本文提供的eTF中使用的TAD衍生自病毒蛋白质。衍生自病毒蛋白质的示例性TAD包括,例如,VP64 (SEQ ID NO:95)、VPR (SEQ ID NO:114)、VP16、VP128、p65、p300的TAD结构域,或其任何功能片段或变体,或与之具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0132] 在另一个示例性实施方案中,适于在本文提供的eTF中使用的TAD衍生自人类蛋白质。衍生自人类蛋白质的示例性TAD包括,例如,CBP/p300相互作用反式激活物2 (CITED2) (SEQ ID NO:96或98)、CBP/p300相互作用反式激活物4 (CITED4) (SEQ ID NO:97或100)、EGR1 (SEQ ID NO:1)或EGR3 (SEQ ID NO:422)的TAD结构域,或其任何功能片段或变体,或与之具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0133] 在另一个实施方案中,适于在本文提供的eTF中使用的TRD包含KRAB结构域,该结构域可以募集抑制转录的蛋白质,或其任何功能片段或变体,或与之具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0134] 在某些实施方案中,锌指DBD缀合至转录调节域,如TAD或TRD,如下文进一步描述。在各个实施方案中,锌指DBD可以缀合至来自病毒蛋白质如VP64或VPR的TAD或来自人类蛋白质如CITED2或CITED4的TAD。在某些实施方案中,衍生自人类蛋白质例如EGR1或EGR3的锌指DBD与衍生自人类蛋白质例如CITED2或CITED4的TAD缀合。在某些实施方案中,衍生自人类蛋白质例如EGR1或EGR3的锌指DBD与VP64或VPR TAD缀合。在某些实施方案中,与人类蛋白质例如EGR1或EGR3具有小于75%序列同一性的合成锌指DBD或锌指DBD与衍生自人类蛋白质例如CITED2或CITED4的TAD缀合。在某些实施方案中,与人类蛋白质例如EGR1或EGR3具有小于75%序列同一性的合成锌指DBD或锌指DBD与VP64或VPR TAD缀合。

[0135] 在某些实施方案中,dCas蛋白缀合至转录调节域,如TAD或TRD,如下文进一步描述。在各个实施方案中,dCas可以缀合至来自病毒蛋白质如VP64或VPR的TAD或来自人类蛋白质如CITED2或CITED4的TAD。在示例性实施方案中,dCas9与VP64或VPR TAD缀合。

[0136] 在某些实施方案中,TALE蛋白缀合至转录调节域,如TAD或TRD,如下文进一步描述。在各个实施方案中,TALE可以缀合至来自病毒蛋白质如VP64或VPR的TAD或来自人类蛋白质如CITED2或CITED4的TAD。在示例性实施方案中,TALE与VP64或VPR TAD缀合。

[0137] 上调SCN1A的eTF

[0138] 在一方面,本申请提供了能够上调钠电压门控通道 α 亚单位1 (SCN1A) 基因的表达并增加其相应蛋白质产物Nav1.1的表达的eTF。SCN1A基因属于编码用于装配钠通道的亚单位的基因家族。这些通道将带正电荷的钠离子转运到细胞内,在细胞产生并传输电信号的能力方面起关键作用。SCN1A基因编码被称为Nav1.1的钠通道的一部分(α 亚单位)。这些通

道主要存在于脑中,在脑中它们控制钠离子向细胞内的流动。Nav1.1通道参与将信号从一个神经细胞(或神经元)传输到另一个神经细胞(或神经元)。已经发现SCN1A基因中的数个突变会导致遗传性癫痫伴热性癫痫发作+(GEFS+),这是严重程度不同的一系列癫痫发作病症。这些病况包括单纯的热性(发热相关的)癫痫发作,其开始于婴儿期,通常到5岁时停止,以及热性癫痫发作+(FS+)。FS+涉及热性和其他类型的癫痫发作,包括与发热无关的癫痫发作(无发热癫痫发作),其持续到超过儿童期。GEFS+谱还包括其他病况,如Dravet综合征(也称为婴儿重度肌阵挛性癫痫或SMEI),其导致更严重的癫痫发作,持续时间更长,并且可能难以控制。这些反复的癫痫发作(癫痫)会随着时间的推移而恶化,并且常常伴有脑功能下降。其他许多突变与家族性偏瘫性偏头痛有关,家族性偏瘫性偏头痛是一种在家族中普遍存在的偏头痛形式,并且至少一种突变与某些抗癫痫药物的有效性相关。因此,本文提供的增加SCN1A表达的eTF可用来治疗与Nav1.1通道中的突变相关的多种疾病或病症。

[0139] 在某些实施方案中,本文公开的上调SCN1A的eTF识别以下大小的靶结合位点:至少9bp、12bp、15bp、18bp、21bp、24bp、27bp、30bp、33bp或36bp;超过9bp、12bp、15bp、18bp、21bp、24bp、27bp或30bp;或者9-33bp、9-30bp、9-27bp、9-24bp、9-21bp、9-18bp、9-15bp、9-12bp、12-33bp、12-30bp、12-27bp、12-24bp、12-21bp、12-18bp、12-15bp、15-33bp、15-30bp、15-27bp、15-24bp、15-21bp、15-18bp、18-33bp、18-30bp、18-27bp、18-24bp、18-21bp、21-33bp、21-30bp、21-27bp、21-24bp、24-33bp、24-30bp、24-27bp、27-33bp、27-30bp或30-33bp。在示例性实施方案中,本文公开的上调SCN1A的eTF识别18-27bp、18bp或27bp的靶结合位点。

[0140] 在某些实施方案中,本文公开的上调SCN1A的eTF识别位于染色体2上的靶结合位点。在某些实施方案中,本文公开的上调SCN1A的eTF识别位于染色体2上在SCN1A的TSS上游或下游110kb、100kb、90kb、80kb、70kb、60kb、50kb、40kb、30kb、20kb、10kb、5kb、4kb、3kb、2kb或1kb内的靶结合位点。在某些实施方案中,本文公开的上调SCN1A的eTF识别位于染色体2上在SCN1A的TSS上游110kb内的靶结合位点。在某些实施方案中,本文公开的上调SCN1A的eTF识别位于染色体2上在SCN1A的TSS下游110kb内的靶结合位点。在示例性实施方案中,这类靶结合位点是18-27bp、18bp或27bp。

[0141] 在某些实施方案中,本文公开的上调SCN1A的eTF识别位于染色体2上在位置166179652-165989571内、位置166128050-166127958内、位置166155414-166140590内、位置166179652-1661777272内或位置1659990246-165989592内(均参照GRCh38.p12)的靶结合位点。在示例性实施方案中,这类靶结合位点是18-27bp、18bp或27bp。

[0142] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i)为18-27bp、18bp或27bp,(ii)与染色体2上选自166178880、166178871、166177369、166177362、166177299、166177299、166155393、166155264、166149373、166149176、166149165、166149118、166148953、166148565、166142396、166142391、166142344、166142239、166141162、166140928、166140590、165990076、165989684、165989571、166155255、166155099、166148843、166148361、166142219、166141090、165990246、165990193、166149168、166127991、166128002、166128037或166128025(均参照GRCh38.p12)的位置重叠,并且(iii)当被本文公开的eTF结合时能够产生至少1.2倍的SCN1A表达增加。

[0143] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267中任一个或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少1.2倍的SCN1A表达增加。

[0144] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp, (ii) 与染色体2上选自166155255、166155099、166148843、166148361、166142219、166141090、165990246、165990193、166149168、166127991、166128002、166128037或166128025(均参照GRCh38.p12)的位置重叠,并且(iii) 当被本文公开的eTF结合时能够产生至少2倍的SCN1A表达增加。

[0145] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:35-36、108-109、136、209-210、226、228、233、236或247-248中任一个或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少2倍的SCN1A表达增加。

[0146] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp, (ii) 与染色体2上选自166149168、166127991、166128002、166128037或166128025(均参照GRCh38.p12)的位置重叠,并且(iii) 当被本文公开的eTF结合时能够产生至少5倍的SCN1A表达增加。

[0147] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:35、36、108、109或136中任一个或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少5倍的SCN1A表达增加。

[0148] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp, (ii) 与染色体2上选自166128002、166128037或166128025(均参照GRCh38.p12)的位置重叠,并且(iii) 当被本文公开的eTF结合时能够产生至少15倍的SCN1A表达增加。

[0149] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:36、108或109中任一个或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少15倍的SCN1A表达增加。

[0150] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp, (ii) 与染色体2上选自166128037或166128025(均参照GRCh38.p12)的位置重叠,并且(iii) 当被本文公开的eTF结合时能够产生至少20倍的SCN1A表达增加。

[0151] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:36或109中任一个或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少20倍的SCN1A表达增加。

[0152] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp, (ii) 与染色体2上位置166128025处的位置重叠,并且(iii) 当被本文公开的eTF结合时能够产生至少25倍的SCN1A表达增加。

[0153] 在某些实施方案中,本文公开的上调SCN1A的eTF (i) 与包含SEQ ID NO:36或由其组成的靶位点结合,并且(ii) 当被本文公开的eTF结合时能够产生至少25倍的SCN1A表达增加。

[0154] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i) 为18-27bp、18bp或27bp,并且(ii) 与基因组区域结合,该基因组区域在具有SEQ ID NO:35-

37、101、105-111、136、195-211、224-238或240-267中任一序列的基因组位置的至少1kb、750bp、500bp、400bp、300bp、200bp、100bp或50bp内。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表达增加。

[0155] 在某些实施方案中,本文公开的上调SCN1A的eTF识别这样的靶结合位点,其(i)为18-27bp、18bp或27bp,并且(ii)与基因组区域结合,该基因组区域与具有SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267中任一序列的基因组位置至少部分重叠。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表达增加。

[0156] 在某些实施方案中,本文公开的上调SCN1A的eTF识别具有任一个以下序列的靶结合位点:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表达增加。

[0157] 在某些实施方案中,本文公开的上调SCN1A的eTF导致与对照(例如,没有eTF或者不识别靶位点的eTF)相比,在细胞中或在体内至少1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、25倍、50倍、100倍或更高,或至少50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高的SCN1A表达(SCN1A RNA和/或Nav1.1蛋白)上调。在各个实施方案中,可以使用PCR方法、Western印迹法或免疫测定来检测SCN1A表达的上调。

[0158] 在某些实施方案中,本文公开的上调SCN1A的eTF与靶位点结合,在转录激活试验中,该靶位点相对于对照能够将SCN1A表达增加至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2倍、3倍、4倍、5倍、8倍、10倍、12倍、15倍、18倍、20倍、25倍、30倍、40倍、50倍、75倍、100倍或更高,或至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。本文在实施例3中提供了示例性的SCN1A转录激活试验。简言之,用携带eTF或对照eGFP报道基因构建体的质粒转染HEK293。转染后48h,收集细胞,分离RNA,进行逆转录,并通过qPCR(例如,使用具有SEQ ID NO:190和191的引物)分析所得的cDNA样品,以对内源SCN1A转录物的水平进行定量。GAPDH可用作参考基因,以确定SCN1A表达的相对水平。

[0159] 在某些实施方案中,本文公开的上调SCN1A的eTF具有最小的脱靶效应,例如,与非特异性结合相关的脱靶效应,例如,脱靶基因或除SCN1A以外的基因的表达的调节。在一个实施方案中,本文公开的上调SCN1A的eTF与对照相比特异性上调SCN1A,是该eTF与对照相比对于一个或多个脱靶基因产生的表达的至少5倍、10倍、15倍、20倍、25倍、30倍、40倍或50倍。在示例性实施方案中,本文公开的上调SCN1A的eTF与对照相比特异性上调从SCN1A基因的转录,是该eTF相对于对照产生的40个最近邻基因(例如,位于染色体2上SCN1A编码序列的40个最接近的基因)的转录的至少15倍,这些基因例如是PLA2R1、ITGB6、RBMS1、TANK、PSMD14、TBR1、SLC4A10、DPP4、FAP、IFIH1、GCA、FIGN、GRB14、COBLL1、SLC38A11、SCN3A、SCN2A、CSRNP3、GALNT3、TTC21B、SCN9A、SCN7A、B3GALT1、STK39、CERS6、NOSTRIN、SPC25、ABCB11、DHRS9、BBS5、KLHL41、FASTKD1、PP1G、CCDC173、PHOSPHO2、KLHL23、SSB、METTL5、UBR3和MYO3B(参见表33)。在各个实施方案中,可以使用PCR方法检测从SCN1A基因的转录的上调。

[0160] 在某些实施方案中,本文公开的上调SCN1A的eTF能够在Dravet综合征的Scn1a^{tm1kea}小鼠模型中的高热癫痫发作(HTS)试验中降低癫痫发作的频率。在某些实施方案中,与对照(例如,PBS处理的或用包含编码eGFP的序列的AAV载体治疗)相比,本文公开的eTF能够在HTS试验中将在42.6℃时的癫痫发作频率降低至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2.0倍或更多,或至少20%、30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%或更多。在某些实施方案中,本文公开的eTF能够在HTS试验中降低在42.6℃时的癫痫发作频率,使得该实验中分析的至少60%、62%、65%、70%、75%、76%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的小鼠在42.6℃时没有癫痫发作。本文在实施例11中描述了示例性的HTS试验。简言之,可以在P1经双侧ICV,向雄性Scn1a^{+/+}小鼠与雌性C57BL/6J小鼠杂交产生的幼仔给予如本文提供的编码上调SCN1A的eTF的AAV9载体或编码eGFP的对照载体。可以以约1.0E10-5.0E12 gc/小鼠向小鼠给药。通过每2分钟将小鼠的体温(在受控条件下并通过体温监测)提高约0.5℃,直到出现第一次强直-阵挛性癫痫发作伴姿势丧失或者直到体温达到43℃,在混合129Stac X C57BL/6背景中的P26-P28 SCN1A杂合小鼠和SCN1A野生型小鼠中进行HTS试验。如果在整个实验过程中没有检测到癫痫发作伴姿势丧失,则认为小鼠没有癫痫发作。

[0161] 在某些实施方案中,本文公开的上调SCN1A的eTF能够提高SCN1A杂合小鼠,例如Scn1a^{tm1kea}小鼠系的存活率。在某些实施方案中,与对照(例如,PBS处理的或用包含编码eGFP的序列的AAV载体治疗)相比,本文公开的eTF能够将P100时SCN1A杂合小鼠的存活率提高至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2.0倍或更多,或至少20%、30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%或更多。在某些实施方案中,本文公开的eTF能够提高P100时SCN1A杂合小鼠的存活率,使得该试验中分析的至少65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的小鼠在P100时仍然存活。本文在实施例12中描述了示例性的存活试验。简言之,可以在P1经双侧ICV,向雄性Scn1a^{+/+}小鼠与雌性C57BL/6J小鼠杂交产生的幼仔给予AAV9载体。可以以约1.0E10-5.0E12 gc/小鼠向小鼠给药。确定了存活到P100的小鼠的数目。

[0162] 在某些实施方案中,本文公开的上调SCN1A的eTF与一种或多种人类蛋白质具有高同一性百分比(如下文进一步描述)。在某些实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。在某些实施方案中,这类eTF可包含衍生自人EGR1或EGR3的DBD和衍生自人EGR1、EGR3、CITED2或CITED4的TAD。

[0163] 在某些实施方案中,本文公开的上调SCN1A的eTF包含具有一个或多个锌指结构域的DNA结合域,所述锌指结构域包含含有SEQ ID NO:115-130中任一个的识别螺旋。在某些实施方案中,本文公开的上调SCN1A的eTF包含具有至少一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个锌指结构域的DNA结合域,其中每个锌指结构域独

立地包含含有SEQ ID NO:115-130中任一个的识别螺旋。在某些实施方案中,本文公开的上调SCN1A的eTF包含具有六个锌指结构域的DNA结合域,其中每个锌指结构域独立地包含含有SEQ ID NO:115-130中任一个的识别螺旋。在某些实施方案中,本文公开的上调SCN1A的eTF包含具有九个锌指结构域的DNA结合域,其中每个锌指结构域独立地包含含有SEQ ID NO:115-130中任一个的识别螺旋。在示例性实施方案中,此类eTF包含具有SEQ ID NO:421的DNA结合域,其中每个X独立选自SEQ ID NO:115-130中的任一个,并且n为6或9。

[0164] 在某些实施方案中,本文公开的上调SCN1A的eTF包含具有以下任一序列的DNA结合域:(i)包含RSDNLVR x REDNLHT x RSDNLVR x QSGNLTE x TSGHLVR x QNSTLTE(SEQ ID NO:135)的序列,其中x可以是1-50个氨基酸的连接体,(ii)与SEQ ID NO:135具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:256的靶位点结合,并且将SCN1A的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0165] 在某些实施方案中,本文公开的上调SCN1A的eTF包含具有以下任一序列的DNA结合域:(i)包含RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT(SEQ ID NO:371)的序列,其中x可以是1-50个氨基酸的连接体,(ii)与SEQ ID NO:371具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:264的靶位点结合,并且将SCN1A的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0166] 在某些实施方案中,本文公开的上调SCN1A的eTF包含具有以下任一序列的DNA结合域:(i)包含RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT(SEQ ID NO:372)的序列,(ii)与SEQ ID NO:372具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:37的靶位点结合,并且将

SCN1A的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0167] 在某些实施方案中,本文公开的上调SCN1A的eTF包含具有以下任一序列的DNA结合域:(i)包含DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR (SEQ ID NO:376)的序列,(ii)与SEQ ID NO:376具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:136的靶位点结合,并且将SCN1A的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0168] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i)包含SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366中任一个的序列;(ii)包含SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376、391-409或423-435中任一个的序列;(iii)与(i)或(ii)的任一个序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0169] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i)包含SEQ ID NO:268-282、305-312或365-366中任一个的序列;(ii)包含SEQ ID NO:131-135、371-372、376、391-409或423-426中任一个的序列;(iii)与(i)或(ii)的任一个序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0170] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i)包含SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、295-299、313-325或364中任一个的序列;(ii)包含SEQ ID NO:22-25、29-31、84-85、88、90-92或427-435中任一个的序列;(iii)与(i)或(ii)的任一个序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少

20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0171] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:305的序列;(ii) 包含SEQ ID NO:423的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0172] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:306的序列;(ii) 包含SEQ ID NO:424的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0173] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:308的序列;(ii) 包含SEQ ID NO:425的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0174] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:313的序列;(ii) 包含SEQ ID NO:427的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、

88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0175] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:316的序列;(ii) 包含SEQ ID NO:430的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0176] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:317的序列;(ii) 包含SEQ ID NO:431的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:371并且与具有SEQ ID NO:264的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0177] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:315的序列;(ii) 包含SEQ ID NO:429的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、

88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0178] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:319的序列;(ii) 包含SEQ ID NO:433的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:371并且与具有SEQ ID NO:264的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0179] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:318的序列;(ii) 包含SEQ ID NO:432的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:371并且与具有SEQ ID NO:264的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0180] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i) 包含SEQ ID NO:314的序列;(ii) 包含SEQ ID NO:428的序列;(iii) 与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、

88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0181] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i)包含SEQ ID NO:440的序列;(ii)包含SEQ ID NO:441的序列;(iii)与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv)(i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:135并且与具有SEQ ID NO:256的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0182] 在某些实施方案中,本文公开的上调SCN1A的eTF包含以下任一序列:(i)包含SEQ ID NO:325的序列;(ii)包含SEQ ID NO:435的序列;(iii)与(i)或(ii)的序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv)(i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF包含SEQ ID NO:371并且与具有SEQ ID NO:264的靶位点结合。在示例性实施方案中,这类eTF能够将SCN1A表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0183] 在某些实施方案中,本文公开的上调SCN1A的eTF包含含有gRNA/Cas复合物的DBD,其中所述gRNA包含靶向序列,该靶向序列包含SEQ ID NO:105-111、113、195-211、224-238或240-251中的任何一个。gRNA的靶序列连接到具有以下序列的支架序列的5'端:5'-GTTT TAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATG CCGTGTATCTCGTCAACTTGTTGGCGAGA-3' (SEQ ID NO:102)。在示例性实施方案中,Cas蛋白是核酸酶灭活的Cas9蛋白。在某些实施方案中,这样的eTF进一步包含与Cas蛋白缀合的一个或多个TAD,其中该TAD选自VP64、VPR、CITED2或CITED4。在一个实施方案中,这样的eTF包含与Cas蛋白的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与Cas蛋白的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与Cas蛋白的N-末端、C-末端或N-末端

和C-末端缀合的CITED4 TAD。在示例性实施方案中,与对照相比,这类eTF能够将SCN1A表达上调至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高。

[0184] 上调颗粒蛋白前体 (GRN) 的eTF

[0185] 在另一方面,本申请提供了能够上调从颗粒蛋白前体 (GRN) 基因的表达并增加GRN蛋白的表达的eTF。颗粒蛋白前体 (GRN或PGRN),也称为前颗粒体蛋白,是富含半胱氨酸的分泌型糖蛋白,具有类似生长因子的性质。GRN位于染色体17q21.31上 (NCBI参考序列:NC_000017.11)。颗粒体蛋白在血管生成、伤口修复、细胞增殖和炎症中起作用。颗粒蛋白前体基因的突变和/或分泌的颗粒蛋白前体的缺乏与多种神经变性疾病和代谢疾病相关,如额颞叶退化;额颞变性或额颞痴呆——与部分颗粒蛋白前体缺乏相关的早发型神经变性疾病;进行性非流畅性失语;词义性痴呆;帕金森病;阿尔茨海默病;和神经元蜡样质脂褐质沉积病伴颗粒蛋白前体完全缺乏。颗粒蛋白前体的突变和/或缺乏还与动脉粥样硬化有关,动脉粥样硬化是一种进行性疾病,其特征由于脂质积累和细胞增殖而引起的炎症和动脉壁增厚。因此,本文提供的增加GRN表达的eTF可用来治疗与GRN中的突变相关的多种疾病或病症。

[0186] 在某些实施方案中,本文公开的上调GRN的eTF识别以下大小的靶结合位点:至少9bp、12bp、15bp、18bp、21bp、24bp、27bp、30bp、33bp或36bp;超过9bp、12bp、15bp、18bp、21bp、24bp、27bp或30bp;或者9-33bp、9-30bp、9-27bp、9-24bp、9-21bp、9-18bp、9-15bp、9-12bp、12-33bp、12-30bp、12-27bp、12-24bp、12-21bp、12-18bp、12-15bp、15-33bp、15-30bp、15-27bp、15-24bp、15-21bp、15-18bp、18-33bp、18-30bp、18-27bp、18-24bp、18-21bp、21-33bp、21-30bp、21-27bp、21-24bp、24-33bp、24-30bp、24-27bp、27-33bp、27-30bp或30-33bp。在示例性实施方案中,本文公开的上调GRN的eTF识别18-27bp、18bp或27bp的靶结合位点。

[0187] 在某些实施方案中,本文公开的上调GRN的eTF识别位于染色体17上的靶结合位点。在某些实施方案中,本文公开的上调GRN的eTF识别位于染色体17上在GRN的TSS上游或下游110kb、100kb、90kb、80kb、70kb、60kb、50kb、40kb、30kb、20kb、10kb、5kb、4kb、3kb、2kb或1kb内的靶结合位点。在示例性实施方案中,这类靶结合位点是18-27bp、18bp或27bp。

[0188] 在某些实施方案中,本文公开的上调GRN的eTF识别位于染色体17上在位置44,344,963-44,345,178 (参照GRCh38.p12) 内的靶结合位点。在示例性实施方案中,这类靶结合位点是18-27bp、18bp或27bp。

[0189] 在某些实施方案中,本文公开的上调GRN的eTF识别这样的靶结合位点,其(i)为18-27bp、18bp或27bp,并且(ii)与基因组区域结合,该基因组区域在具有SEQ ID NO:38、113或330-336中任一序列的基因组位置的至少1kb、750bp、500bp、400bp、300bp、200bp、100bp或50bp内。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的GRN表达增加。

[0190] 在某些实施方案中,本文公开的上调GRN的eTF识别这样的靶结合位点,其(i)为18-27bp、18bp或27bp,并且(ii)与基因组区域结合,该基因组区域与具有SEQ ID NO:38、113或330-336中任一序列的基因组位置至少部分重叠。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表达增加。

[0191] 在某些实施方案中,本文公开的上调GRN的eTF识别具有任一个以下序列的靶结合位点:SEQ ID NO:38、113、330-336。在某些实施方案中,当被本文公开的eTF结合时,靶结合位点能够产生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表达增加。

[0192] 在某些实施方案中,本文公开的上调GRN的eTF导致与对照(例如,没有eTF或者不识别靶位点的eTF)相比,在细胞中或在体内至少1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、25倍、50倍、100倍或更高,或20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高的GRN表达(GRN RNA和/或蛋白质)上调。在各个实施方案中,可以使用PCR方法、Western印迹法或免疫测定来检测GRN表达的上调。

[0193] 在某些实施方案中,本文公开的上调GRN的eTF具有最小的脱靶效应,例如,与非特异性结合相关的脱靶效应,例如,脱靶基因或除GRN以外的基因的表达的调节。在一个实施方案中,本文公开的上调GRN的eTF与对照相比特异性上调GRN,是该eTF与对照相比对于一个或多个脱靶基因产生的表达的至少5倍、10倍、15倍、20倍、25倍、30倍、40倍或50倍。在示例性实施方案中,本文公开的上调GRN的eTF与对照相比特异性上调从GRN基因的转录,是该eTF相对于对照产生的10、20、30、40或50个最近邻基因(例如,位于染色体17上GRN编码序列的10、20、30、40或50个最接近的基因)的转录的至少5倍、10倍、15倍、20倍或更高。在各个实施方案中,可以使用PCR方法检测从GRN基因的转录的上调。

[0194] 在某些实施方案中,本文公开的上调GRN的eTF与一种或多种人类蛋白质具有高同一性百分比(如下文进一步描述)。在某些实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。在某些实施方案中,这类eTF可包含衍生自人EGR1或EGR3的DBD和衍生自人EGR1、EGR3、CITED2或CITED4的TAD。

[0195] 在某些实施方案中,本文公开的上调GRN的eTF包含具有一个或多个锌指结构域的DNA结合域,所述锌指结构域包含含有SEQ ID NO:141-164中任一个的识别螺旋。在某些实施方案中,本文公开的上调GRN的eTF包含具有至少一个、两个、三个、四个、五个、六个、七个、八个、九个、十个、十一个或十二个锌指结构域的DNA结合域,其中每个锌指结构域独立地包含含有SEQ ID NO:141-164中任一个的识别螺旋。在某些实施方案中,本文公开的上调GRN的eTF包含具有六个锌指结构域的DNA结合域,其中每个锌指结构域独立地包含含有SEQ ID NO:141-164中任一个的识别螺旋。在某些实施方案中,本文公开的上调GRN的eTF包含具有九个锌指结构域的DNA结合域,其中每个锌指结构域独立地包含含有SEQ ID NO:141-164中任一个的识别螺旋。在示例性实施方案中,此类eTF包含具有SEQ ID NO:421的DNA结合域,其中每个X独立选自SEQ ID NO:141-164中的任一个,并且n为6或9。

[0196] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i)包含RNDLTLE x DPGALVR x TSGELVR x RSDNLVR x TSGELVR x TKNSLTE(SEQ ID NO:171)的序列,(ii)与SEQ ID NO:171具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i) 或(ii)的功能片段。在某些实施方案中,这样的eTF进一

步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:331的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0197] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i) 包含RSDNLVR x DPGHLVR x RSDHLTT x RSDELVR x RSDKLVR x TTGNLTV (SEQ ID NO:412)的序列,(ii) 与SEQ ID NO:412具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i) 或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:332的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0198] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i) 包含RSDHLTT x RSDELVR x RSDKLVR x TTGNLTV x QLAHLRA x TKNSLTE (SEQ ID NO:413)的序列,(ii) 与SEQ ID NO:413具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i) 或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:333的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0199] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i) 包含SPADLTR x DSGNLRV x QLAHLRA x QRANLRA x REDNLHT x RSDNLVR (SEQ ID NO:414)的序列,(ii) 与SEQ ID NO:414具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i) 或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD

的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:38的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0200] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i)包含TSHSLTE x HKNALQN x ERSHLRE x SKKALTE x QRANLRA x RKDNLKN x QSSNLVR x QSSSLVR x QAGHLAS (SEQ ID NO:415)的序列,(ii)与SEQ ID NO:415具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:334的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0201] 在某些实施方案中,本文公开的上调GRN的eTF包含具有以下任一序列的DNA结合域:(i)包含QSGDLRR x SPADLTR x DSGNLRV x QLAHLRA x QRANLRA x REDNLHT x RSDNLVR (SEQ ID NO:416)的序列,(ii)与SEQ ID NO:416具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列,或(ii) (i)或(ii)的功能片段。在某些实施方案中,这样的eTF进一步包含一个或多个选自VP64、VPR、CITED2或CITED4的TAD。在一个实施方案中,这样的eTF包含与DBD的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED2TAD。在某些实施方案中,这样的eTF包含与DBD的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在某些实施方案中,这样的eTF能够与具有SEQ ID NO:335的靶位点结合,并且将GRN的表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0202] 在某些实施方案中,本文公开的上调GRN的eTF包含以下任一序列:(i)包含SEQ ID NO:10、16、59-60、63-64、72-73、76-77或337-350中任一个的序列;(ii)包含SEQ ID NO:26、32、86-89、93、165-171、377-390或412-416中任一个的序列;(iii)与(i)或(ii)的任一个序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv) (i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF能够将GRN表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

[0203] 在某些实施方案中,本文公开的上调GRN的eTF包含含有gRNA/Cas复合物的DBD,其中所述gRNA包含含有SEQ ID NO:113的靶向序列。gRNA的靶序列连接到具有以下序列的支

架序列的5'端: 5'-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCCAAAATGCCGTGTTTATCTCGTCAACTTGTGGCGAGA-3' (SEQ ID NO:102)。在示例性实施方案中,Cas蛋白是核酸酶灭活的Cas9蛋白。在某些实施方案中,这样的eTF进一步包含与Cas蛋白缀合的一个或多个TAD,其中该TAD选自VP64、VPR、CITED2或CITED4。在一个实施方案中,这样的eTF包含与Cas蛋白的C-末端缀合的VPR TAD结构域。在某些实施方案中,这样的eTF包含与Cas蛋白的N-末端、C-末端或N-末端和C-末端缀合的CITED2 TAD。在某些实施方案中,这样的eTF包含与Cas蛋白的N-末端、C-末端或N-末端和C-末端缀合的CITED4 TAD。在示例性实施方案中,这类eTF能够将GRN表达上调与对照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。与人类蛋白质高度同源的eTF

[0204] 在另一方面,本申请提供了与一种或多种人类蛋白质具有高序列同一性百分比的eTF,该人类蛋白质可以被设计为结合基因组靶位点并调节任何目的基因的表达(上调或下调),包括例如SCN1A或GRN。相比于与人类蛋白质序列具有较低同一性百分比的eTF,这类eTF当施用于受试者时没有或几乎没有免疫原性,或具有降低的免疫原性。

[0205] 在某些实施方案中,本文提供的eTF与一种或多种人类蛋白质具有至少75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。当本文提供的eTF包含衍生自同一蛋白质的DBD和TAD时,与人类蛋白质的同一性百分比可以通过计算eTF中与其衍生自的人类蛋白质(例如EGR1或EGR3)相匹配的氨基酸残基的总数除以eTF中的氨基酸残基总数来确定。当提供的eTF包含来自一种人类蛋白质的DBD和衍生自不同人类蛋白质的TAD时,与人类蛋白质的同一性百分比可以通过分别计算每个结构域与人类蛋白质的同一性百分比并将二者相加来确定,例如,(i) 计算DBD中与其衍生自的人类蛋白质(例如,EGR1或EGR3)相匹配的氨基酸残基总数,除以eTF中的氨基酸残基总数;(ii) 计算TAD中与其衍生自的人类蛋白质(例如,CITED2或CITED4)相匹配的氨基酸残基总数,除以eTF中的氨基酸残基总数;以及(iii) 将(i)和(ii)的总和相加。在这样的实施方案中,结构如下划分:第一结构域从eTF的N-末端延伸到第二结构域的编码序列的起点,第二结构域从第二结构域的编码序列的起点延伸到eTF的C-末端(例如,对于具有构型NLS-DBD-连接体-NLS-TAD的eTF,第一结构域为NLS-DBD-连接体,而第二结构域为NLS-TAD)。当本文提供的eTF包含来自一种人类蛋白质的DBD和衍生自一种或多种不同人类蛋白质的两个TAD时,与人类蛋白质的同一性百分比可以通过分别计算每个结构域与人类蛋白质的同一性百分比并将全部三者相加来确定,例如,(i) 计算DBD中与其衍生自的人类蛋白质(例如,EGR1或EGR3)相匹配的氨基酸残基总数,除以eTF中的氨基酸残基总数;(ii) 计算第一TAD中与其衍生自的人类蛋白质(例如,CITED2或CITED4)相匹配的氨基酸残基总数,除以eTF中的氨基酸残基总数;(iii) 计算第二TAD中与其衍生自的人类蛋白质(例如,CITED2或CITED4)相匹配的氨基酸残基总数,除以eTF中的氨基酸残基总数;以及(iv) 将(i)、(ii)和(iii)的总和相加。在这样的实施方案中,结构如下划分:第一结构域从eTF的N-末端延伸到第二结构域的编码序列的起点,第二结构域从第二结构域的编码序列的起点延伸到第三结构域的编码序列的起点,第三结构域从第三结构域的编码序列的起点延伸到eTF的C-末端(例如,对于具有构型NLS-TAD1-连接

体-NLS-DBD-连接体-NLS-TAD2的eTF,第一结构域为-NLS-TAD1-连接体,第二结构域为-NLS-DBD-连接体,而第三结构域为-NLS-TAD2)。如本节所述,与一种或多种人类蛋白质的同一性百分比可以使用同一性百分比输出来确定,该同一性百分比输出使用可从NCBI获得的标准蛋白质BLAST工具(例如,blastp套件比对工具,使用具有默认参数的blastp(蛋白质->蛋白质)算法)来获得,该工具可以从万维网上的NCBI网址blast.ncbi.nlm.nih.gov/获得。

[0206] 在某些实施方案中,由于与天然存在的人类蛋白质的高度序列同一性,本文提供的eTF具有在人类受试者中引发极低、最小或不引发不良免疫应答的益处。在某些实施方案中,与使用与一种或多种人类蛋白质具有较低同一性百分比的eTF(例如,与一种或多种人类蛋白质具有低于50%、55%、65%或70%序列同一性的eTF)所观察到的免疫原性相比,本文提供的eTF引发降低的免疫原性,例如至少0.5、1、1.5、2、2.5、3、3.5、4、4.5、5、10、15、20、25、30、35、40、45或50倍或更多倍的免疫原性降低。在一些情况下,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。具有低或最小免疫原性的基因疗法具有几个优点,包括改善的患者耐受性、降低的达到治疗效果所需的剂量、一次给药后延长的治疗效果、根据需要多次施用或以多剂量施用的能力、每次给药在较长时段上持续的治疗功效、基因疗法的提高的安全性和/或提高的有效性。

[0207] 在某些实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含衍生自一种或多种天然存在的人类蛋白质的DBD和TAD。在某些实施方案中,这样的eTF可以包含衍生自包含DBD的任何天然存在的人类蛋白质的DBD。在示例性实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含衍生自天然存在的锌指蛋白的DBD,例如表7中列出的衍生自任何一种锌指蛋白的DBD。在某些实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含衍生自人类EGR蛋白如EGR1、EGR2、EGR3或EGR4的DBD。在示例性实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含衍生自人EGR1或EGR3的DBD。在各个实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含衍生自人锌指蛋白质的DBD,其中已经在DBD的一个或多个锌指结构域中进行了最小的氨基酸改变(例如1、2、3、4、5、6、7个,或1-2、1-3、1-4、1-5、1-6、1-7、2-3、2-4、2-5、2-6、2-7、3-4、3-5、36或3-7个氨基酸改变),以改变DBD识别目的靶结合位点的结合特异性。这类序列修饰优选地在DBD的锌指结构域的识别螺旋中进行,而其余的人锌指DBD或蛋白质(包括TAD)保持未修饰,从而与天然存在的人类蛋白质保持尽可能高的序列同一性。

[0208] 在某些实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含与衍生自人类蛋白质的DBD缀合的衍生自人类蛋白质的一个或多个转录调节域(例如,TAD或TRD)。在各个实施方案中,转录调节域可以衍生自任何天然存在的人类蛋白质,该人类蛋白质具有当eTF通过DBD与靶位点结合时能够募集可以调节从目的基因的转录(例如,RNA聚合酶II、共激活蛋白或共阻抑蛋白)或基因表达水平的一种或多种蛋白质因子的结构域。在示例性实施方案中,TAD衍生自人EGR蛋白,例如人EGR1、EGR2、EGR3或EGR4,或引用的人类蛋白质,例如人CITED2或CITED4蛋白。在示例性实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含来自人EGR1或EGR3蛋白的TAD。在另一个示例性实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含来自人CITED2或CITED4蛋白的TAD。

[0209] 在一个实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF可包含人DBD(hDBD)和人TAD(hTAD)(例如,hTAD-hDBD或hDBD-hTAD),其中hDBD和hTAD可以衍生自相同的人类蛋白质或不同的人类蛋白质。在另一个实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF可包含hDBD和两个hTAD,其中hDBD和hTAD衍生自相同的人类蛋白质,hDBD衍生自第一人类蛋白质并且两个hTAD均衍生自第二人类蛋白质,hDBD和一个hTAD衍生自第一人类蛋白质并且第二hTAD衍生自第二人类蛋白质,或者hDBD衍生自第一人类蛋白质、一个hTAD衍生自第二人类蛋白质并且第二hTAD衍生自第三人类蛋白质(例如,hTAD1-hDBD-hTAD1或hTAD1-hDBD-hTAD2)。

[0210] 在示例性实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含以下构型中的任一种:(i)均衍生自人EGR1的hDBD和hTAD;(ii)均衍生自人EGR3的hDBD和hTAD;(iii)衍生自人EGR1的hDBD和衍生自CITED2的hTAD(例如,hEGR1 DBD-hCITED2 TAD或hCITED2 TAD-hEGR1 DBD);(iv)衍生自人EGR1的hDBD和衍生自人CITED4的hTAD(例如,hEGR1 DBD-hCITED4 TAD或hCITED4 TAD-hEGR1 DBD);(v)衍生自人EGR3的hDBD和衍生自CITED2的hTAD(例如,hEGR3 DBD-hCITED2 TAD或hCITED2TAD-hEGR3 DBD);(vi)衍生自人EGR3的hDBD和衍生自人CITED4的hTAD(例如,hEGR3 DBD-hCITED4 TAD或hCITED4 TAD-hEGR3DBD);(vii)衍生自人EGR1的hDBD和两个衍生自CITED2的hTAD(例如,hCITED2 TAD-hEGR1 DBD-hCITED2 TAD);(viii)衍生自人EGR1的hDBD和两个衍生自人CITED4的hTAD(例如,hCITED4 TAD-hEGR1 DBD-hCITED4 TAD);(ix)衍生自人EGR3的hDBD和两个衍生自人CITED2的hTAD(例如,hCITED2 TAD-hEGR3 DBD-hCITED2TAD);(x)衍生自人EGR3的hDBD和两个衍生自人CITED4的hTAD(例如,hCITED4 TAD-hEGR3 DBD-hCITED4 TAD);(xi)衍生自人EGR1的hDBD,衍生自人CITED2的第一hTAD,衍生自人CITED4的第二hTAD(例如,hCITED2 TAD-hEGR1 DBD-hCITED4 TAD或hCITED4TAD-hEGR1 DBD-hCITED2 TAD);或(xii)衍生自人EGR3的hDBD,衍生自人CITED2的第一hTAD,衍生自人CITED4的第二hTAD(例如,hCITED2 TAD-hEGR3 DBD-hCITED4 TAD或hCITED4 TAD-hEGR3DBD-hCITED2 TAD)。

[0211] 在另一个实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF可以包含hDBD和hTRD(例如,hTRD-hDBD或hDBD-hTRD),其中hDBD和hTRD可以衍生自相同的蛋白质或不同的蛋白质。在另一个实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF可包含hDBD和两个hTRD,其中hDBD和hTRD衍生自相同的人类蛋白质,hDBD衍生自第一人类蛋白质并且两个hTRD均衍生自第二人类蛋白质,hDBD和一个hTRD衍生自第一人类蛋白质并且第二hTRD衍生自第二人类蛋白质,或者hDBD衍生自第一人类蛋白质、一个hTRD衍生自第二人类蛋白质并且第二hTRD衍生自第三人类蛋白质(例如,hTRD1-hDBD-hTRD1或hTRD1-hDBD-hTRD2)。

[0212] 在某些实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF包含以下任一种:(i)包含SEQ ID NO:6-10、13-16、57-64、67-77、295-299、313-325、345-350、364、436中任一个的序列;(ii)包含SEQ ID NO:22-26、29-32、84-93、385-390、406-409、427-435或437中任一个的序列;(iii)与(i)或(ii)的任一个序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列,或(iv)(i)、(ii)或(iii)中任一序列的功能片段或变体。在示例性实施方案中,这类eTF能够将SCN1A或GRN表达上调与对照相比至少2倍、5倍、10倍、15倍、20

倍、25倍、30倍、40倍、50倍或更高,或者与对照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性实施方案中,这类eTF与一种或多种人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整体序列同一性。在某些实施方案中,相比于与一种或多种人类蛋白质具有较低整体序列同一性百分比的eTF,这类eTF表现出降低的免疫原性。在各个实施方案中,免疫原性的降低可以使用elispot测定、免疫测定或计算机模拟方法来测量。

[0213] 在某些实施方案中,本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF除DBD和TAD结构域外还可以另外包含一个或多个氨基酸序列或结构域,如核定位信号或连接体等。另外,编码本文提供的与一种或多种人类蛋白质具有高序列同一性百分比的eTF的多核苷酸,除eTF的编码序列外,还可以另外包含一个或多个核酸序列,如启动子、增强子、聚A尾等。在这样的实施方案中,一个或多个另外的氨基酸序列和/或核酸序列优选地是人序列,衍生自人序列,或与人类蛋白质具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。

[0214] 多核苷酸

[0215] 在另一方面,本申请提供了编码本文公开的任何eTF的多核苷酸。

[0216] 在某些实施方案中,本申请提供了包含以下任一序列的多核苷酸:(i) 编码包含SEQ ID NO:6-10、13-16、57-64、67-77、268-282、305-325、337-350、364、295-299或365-366中任一个的eTF或其变体或功能片段的核酸序列;(ii) 编码具有SEQ ID NO:6-10、13-16、57-64、67-77、268-282、305-325、337-350、364、295-299或365-366中任一个的eTF的功能片段的核酸;或(iii) 编码eTF或其变体或功能片段的核酸,该eTF与具有SEQ ID NO:6-10、13-16、57-64、67-77、112、268-282、305-325、337-350、364、295-299、365-366中任一个的eTF具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该eTF能够上调SCN1A或GRN。

[0217] 在某些实施方案中,本申请提供了包含以下任一序列的多核苷酸:(i) 编码包含SEQ ID NO:22-26、29-32、84-93、131-135、165-171、377-409或423-435中任一个的DBD或其变体或功能片段的核酸序列;(ii) 编码具有SEQ ID NO:22-26、29-32、84-93、131-135、165-171、377-409、423-435中任一个的DBD的功能片段的核酸;或(iii) 编码DBD或其变体或功能片段的核酸,该DBD与具有SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、377-409或423-435中任一个的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该eTF能够与被SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、377-409、423-435中任一个结合的靶位点结合。

[0218] 在某些实施方案中,本申请提供了编码上调内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一序列:(i) 编码包含SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366中任一个的eTF的核酸序列;(ii) 编码具有SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366中任一个的eTF的功能片段的核酸;或(iii) 编码eTF的核酸,该eTF与具有SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366中任一个的eTF具有

至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该eTF能够上调SCN1A。

[0219] 在某些实施方案中,本申请提供了编码DBD的多核苷酸,该DBD与被本文公开的eTF结合时能够上调内源SCN1A的基因组靶位点结合,其中该多核苷酸包含以下任一个序列:

(i) 编码包含SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376、391-409或423-435中任一个的DBD的核酸序列;(ii) 编码具有SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376、391-409或423-435中任一个的DBD的功能片段的核酸;或(iii) 编码eTF的核酸,该eTF与具有SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376、391-409或423-435中任一个的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该DBD能够与被SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376、391-409或423-435中任一个结合的靶位点结合。

[0220] 在某些实施方案中,本申请提供了编码DBD的多核苷酸,该DBD与被本文公开的eTF结合时能够上调内源SCN1A的基因组靶位点结合,其中该多核苷酸包含以下任一个序列:

(i) 编码包含SEQ ID NO:135、371、372或376中任一个的DBD的核酸序列;(ii) 编码具有SEQ ID NO:135、371、372或376中任一个的DBD的功能片段的核酸;或(iii) 编码eTF的核酸,该eTF与具有SEQ ID NO:135、371、372或376中任一个的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该DBD能够与被SEQ ID NO:22-25、29-31、84-85、88、90-92、131-135、371-372、376或391-409中任一个结合的靶位点结合。

[0221] 在某些实施方案中,本申请提供了编码上调内源GRN的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i) 编码包含SEQ ID NO:10、16、59-60、63、64、72-73、76、77或337-350中任一个的eTF的核酸序列;(ii) 编码具有SEQ ID NO:10、16、59-60、63、64、72-73、76、77或337-350中任一个的eTF的功能片段的核酸;或(iii) 编码eTF的核酸,该eTF与具有SEQ ID NO:10、16、59-60、63、64、72-73、76、77或337-350中任一个的eTF具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该eTF能够上调GRN。

[0222] 在某些实施方案中,本申请提供了编码DBD的多核苷酸,该DBD与被本文公开的eTF结合时能够上调内源GRN的基因组靶位点结合,其中该多核苷酸包含以下任一个序列:(i) 编码包含SEQ ID NO:26、32、86-89、93、165-171或377-390中任一个的DBD的核酸序列;(ii) 编码具有SEQ ID NO:26、32、86-89、93、165-171或377-390中任一个的DBD的功能片段的核酸;或(iii) 编码eTF的核酸,该eTF与具有SEQ ID NO:26、32、86-89、93、165-171或377-390中任一个的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性,其中该DBD能够与被SEQ ID NO:26、32、86-89、93、165-171或377-390中任一个结合的靶位点结合。

[0223] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i) 具有SEQ ID NO:353-363或454中任一个的核酸序列;(ii) 具有SEQ ID NO:442-453中任一个的核酸序列;(iii) 具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、

86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸,其中该多核苷酸编码能够上调SCN1A的eTF。

[0224] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:353的核酸序列;(ii)具有SEQ ID NO:442的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:305的eTF,或其能够上调SCN1A的功能片段或变体。

[0225] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:354的核酸序列;(ii)具有SEQ ID NO:443的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:306的eTF,或其能够上调SCN1A的功能片段或变体。

[0226] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:355的核酸序列;(ii)具有SEQ ID NO:444的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:308的eTF,或其能够上调SCN1A的功能片段或变体。

[0227] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:356的核酸序列;(ii)具有SEQ ID NO:445的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:313的eTF,或其能够上调SCN1A的功能片段或变体。

[0228] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:357的核酸序列;(ii)具有SEQ ID NO:451的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:314的eTF,或其能够上调SCN1A的功能片段或变体。

[0229] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一个序列:(i)具有SEQ ID NO:358的核酸序列;(ii)具有SEQ ID

N0:448的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:315的eTF, 或其能够上调SCN1A的功能片段或变体。

[0230] 在某些实施方案中, 本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸, 其中该多核苷酸包含以下任一个序列: (i) 具有SEQ ID N0:359的核酸序列; (ii) 具有SEQ ID N0:446的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:316的eTF, 或其能够上调SCN1A的功能片段或变体。

[0231] 在某些实施方案中, 本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸, 其中该多核苷酸包含以下任一个序列: (i) 具有SEQ ID N0:360的核酸序列; (ii) 具有SEQ ID N0:447的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:317的eTF, 或其能够上调SCN1A的功能片段或变体。

[0232] 在某些实施方案中, 本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸, 其中该多核苷酸包含以下任一个序列: (i) 具有SEQ ID N0:361的核酸序列; (ii) 具有SEQ ID N0:450的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:318的eTF, 或其能够上调SCN1A的功能片段或变体。

[0233] 在某些实施方案中, 本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸, 其中该多核苷酸包含以下任一个序列: (i) 具有SEQ ID N0:362的核酸序列; (ii) 具有SEQ ID N0:449的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:319的eTF, 或其能够上调SCN1A的功能片段或变体。

[0234] 在某些实施方案中, 本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸, 其中该多核苷酸包含以下任一个序列: (i) 具有SEQ ID N0:454的核酸序列; (ii) 具有SEQ ID N0:452的核酸序列; (iii) 具有 (i) 或 (ii) 的任一序列的功能片段的核酸; 或 (iv) 与 (i)、(ii) 或 (iii) 的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中, 这类多核苷酸编码具有SEQ ID N0:440的eTF, 或其能够上调SCN1A的功能片段或

变体。

[0235] 在某些实施方案中,本申请提供了编码能够调节内源SCN1A的eTF的多核苷酸,其中该多核苷酸包含以下任一序列:(i)具有SEQ ID NO:363的核酸序列;(ii)具有SEQ ID NO:453的核酸序列;(iii)具有(i)或(ii)的任一序列的功能片段的核酸;或(iv)与(i)、(ii)或(iii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性实施方案中,这类多核苷酸编码具有SEQ ID NO:325的eTF,或其能够上调SCN1A的功能片段或变体。

[0236] 表达盒

[0237] 在另一方面,本申请提供了表达盒,其包含如本文提供的编码eTF的多核苷酸和一个或多个调节元件。

[0238] 在某些实施方案中,本文公开的编码eTF的多核苷酸是表达盒的一部分,该表达盒除了编码eTF的序列外还包含一个或多个调节元件。在示例性实施方案中,本文公开的编码eTF的多核苷酸是表达盒的一部分,该表达盒包含位于编码eTF的序列上游的启动子,以便能够驱动eTF在细胞中的表达。

[0239] 在某些实施方案中,本文公开的表达盒包含编码eTF的多核苷酸和位于编码eTF的序列上游的组成型启动子,以便能够驱动eTF在细胞中的表达。组成型启动子的实例包括GAD2启动子、人突触蛋白启动子、CBA启动子、CMV启动子、minCMV启动子、TATA盒、超核心启动子或EF1 α 启动子,或其组合。

[0240] 在某些实施方案中,本文公开的表达盒包含编码eTF的多核苷酸和能够驱动eTF在细胞中的表达的短启动子。在某些实施方案中,适合根据本文描述的核酸分子使用的短启动子包含小于500bp、450bp、400bp、350bp、300bp、250bp、225bp、200bp、175bp、150bp、145bp、140bp、135bp、130bp、125bp、120bp、115bp、110bp、105bp、100bp、95bp、90bp、85bp、80bp或75bp,或约80-300bp、80-275bp、80-250bp、80-200bp、80-150bp、80-125bp、80-120bp、80-115bp、80-110bp、80-105bp、80-100bp、85-300bp、85-275bp、85-250bp、85-200bp、85-150bp、85-125bp、85-120bp、85-115bp、85-110bp、85-105bp、85-100bp、90-300bp、90-275bp、90-250bp、90-200bp、90-150bp、90-125bp、90-120bp、90-115bp、90-110bp、90-105bp、90-100bp、95-300bp、95-275bp、95-250bp、95-200bp、95-150bp、95-125bp、95-120bp、95-115bp、95-110bp、95-105bp、95-100bp、100-300bp、100-275bp、100-250bp、100-200bp、100-150bp、100-125bp、100-120bp、100-115bp、100-110bp或100-105bp。在示例性实施方案中,适合根据本文描述的表达盒使用的短启动子包含约100-120bp、约117bp或约100bp。

[0241] 在某些实施方案中,本文公开的表达盒包含短启动子,该启动子包含以下任一序列或由以下任一序列组成:(i)SEQ ID NO:178;(ii)其变体或功能片段;或(iii)与(i)或(ii)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,该核酸序列与编码本文公开的任何一种eTF的多核苷酸可操作地连接。短启动子序列的其他实例可见于PCT公开WO 2018/213786。

[0242] 在某些实施方案中,本文公开的表达盒包含编码eTF的多核苷酸和位于编码eTF的

序列上游的细胞类型选择性启动子,以便能够驱动eTF在目的细胞中的选择性表达。在某些实施方案中,细胞类型选择性启动子可以对任何目的细胞类型(例如,心脏细胞、肝细胞、肌肉细胞、骨细胞、神经元或其亚群)具有选择性(例如,选择性地驱动在其中的表达)。在示例性实施方案中,本文公开的表达盒包含编码eTF的多核苷酸和位于编码eTF的序列上游的PV选择性调节元件(例如,启动子、增强子和/或启动子和增强子),以便能够驱动eTF在PV细胞中的选择性表达。PV选择性调节元件是指特异性调节在PV神经元中的基因表达的调节元件。在某些实施方案中,相对于一种或多种其他CNS细胞类型,PV选择性调节元件增强在PV神经元中的表达。在某些实施方案中,PV选择性调节元件抑制脱靶细胞类型中的转录和/或翻译过程。

[0243] 在某些实施方案中,与脱靶细胞类型相比,本文提供的PV选择性调节元件导致在PV细胞中的选择性基因表达。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。在某些实施方案中,PV选择性调节元件导致在PV神经元中相对于在至少一种、两种、三种、四种、五种或更多种非PV CNS细胞类型中的选择性基因表达。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、小神经胶质细胞、运动神经元、血管细胞或非GABA能神经元(例如,不表达GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP中的一种或多种的细胞)、非PV神经元(例如,不表达小白蛋白的GABA能神经元)或其他CNS细胞(例如,从未表示任何PV、GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP的CNS细胞类型)。在一些情况下,本文提供的PV选择性调节元件导致在PV神经元中与在非PVGABA能细胞中相比,基因表达的选择性增加。在一些情况下,通过具有不同的细胞标志物、形态学、表型、基因型、功能和/或用于细胞类型分类的其他任何手段来区分细胞类型。

[0244] 由PV选择性调节元件驱动的表达的选择性可以用多种方式来测量。在一个实施方案中,可以通过将从与PV选择性调节元件可操作地连接的基因表达可检测水平的转录物的PV细胞的数目与表达该基因的细胞总数进行比较(例如,PV与表达该基因的总细胞(PV+非PV细胞)之比),来测量PV细胞中相对于非PV细胞中基因表达的选择性。例如,可以使用基于免疫组织化学的共定位测定来确定对PV神经元的选择性,该共定位测定使用与PV选择性调节元件可操作地连接的编码荧光蛋白(例如,eGFP)的基因来测量基因表达,并使用与第二荧光标记物(例如,红色荧光蛋白)连接的鉴定PV细胞的抗体(例如,与PV神经元特异性相互作用的抗PV抗体)。PV细胞中的表达的选择性可以如下计算:将表达PV和eGFP两者的细胞(例如,PV细胞)数除以表达eGFP的细胞(例如,PV细胞和非PV细胞)总数,然后乘以100以转换成百分比。表达转基因的PV细胞的百分比越高,调节元件对PV细胞的选择性就越高。在某些实施方案中,本文提供的PV选择性调节元件对于在PV细胞中的表达可能是高度选择性的。例如,本文提供的PV选择性调节元件可表现出约50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或高于约99%的对PV神经元的选择性(例如,PV神经元/总细胞 $\times 100$)。

[0245] 在一些情况下,本文提供的PV选择性调节元件是短的。在一些情况下,PV选择性调节元件的大小与载体(例如,病毒载体或rAAV)的克隆容量兼容,使得转基因和一个或多个PV选择性调节元件的组合大小不超过载体的克隆容量。在一些情况下,PV选择性调节元件的长度可达约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,PV选择性调节元件为约500-600bp、500-700bp、500-800bp、500-900bp、500-1000bp或500-1500bp。

[0246] 在某些实施方案中,本文提供的PV选择性调节元件包含以下任一序列或由以下任一序列组成:(i)SEQ ID NO:183-185和417;(ii)变体、功能片段或其组合;或(iii)与(i)或(ii)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。在一些情况下,调节元件包含SEQ ID NO:183-185和417中的任一个。PV选择性调节元件的其他实例可见于PCT公开WO 2018/187363。

[0247] 在示例性实施方案中,本申请提供了表达盒,其包含在PV选择性调节元件的控制下编码SCN1A eTF的核酸序列。在某些实施方案中,本申请提供了表达盒,其包含编码SCN1A eTF的核酸序列,该SCN1A eTF包含具有以下任一序列的DBD:在具有SEQ ID NO:183-185或417中任一个的PV选择性调节元件控制下的SEQ ID NO:135、371-372或376。在某些实施方案中,本申请提供了表达盒,其包含编码SCN1A eTF的核酸序列,该SCN1A eTF包含以下任一序列:在具有SEQ ID NO:183-185或417中任一个的PV选择性调节元件控制下的SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366。在某些实施方案中,本申请提供了表达盒,其包含含有以下任一序列的核酸序列:在具有SEQ ID NO:183-185或417中任一个的PV选择性调节元件控制下的SEQ ID NO:353-363。在某些实施方案中,本申请提供了表达盒,其包含编码SCN1A eTF的核酸序列,该SCN1A eTF包含具有以下任一序列的DBD:在具有SEQ ID NO:183或185中任一个的PV选择性调节元件控制下的SEQ ID NO:135、371、372或376。在某些实施方案中,本申请提供了表达盒,其包含编码SCN1A eTF的核酸序列,该SCN1A eTF包含以下任一序列:在具有SEQ ID NO:183或185中任一个的PV选择性调节元件控制下的SEQ ID NO:6-9、13-15、57-58、61-62、67-71、74-75、268-282、295-299、305-325或364-366。在某些实施方案中,本申请提供了表达盒,其包含含有以下任一序列的核酸序列:在具有SEQ ID NO:183或185中任一个的PV选择性调节元件控制下的SEQ ID NO:353-363。

[0248] 在某些实施方案中,本申请提供了表达盒,其包含在细胞类型选择性调节元件的控制下编码GRN eTF的核酸序列。在某些实施方案中,该细胞类型选择性调节元件可以驱动上调GRN的eTF在选自下组的细胞中与其他细胞类型(例如,非CNS细胞)相比以更高水平表达:中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞、浦肯野细胞、锥体细胞(例如,Betz细胞)、运动神经元、大脑皮质神经元和纹状体细胞。

[0249] 在某些实施方案中,本文提供的表达盒除了启动子外还可包含一个或多个额外的调节元件,例如与转录起始或终止相关的序列、增强子序列和有效的RNA加工信号。示例性的调节元件包括,例如,内含子、增强子、UTR、稳定性元件、WPRE序列、Kozak共有序列、翻译后响应元件、微RNA结合位点或多腺苷酸化(聚A)序列或其组合。调节元件可以用来在基因表达的转录阶段、转录后阶段或翻译阶段调节基因表达。在RNA水平上,调节可以在翻译(例如,稳定mRNA以供翻译的稳定性元件)、RNA切割、RNA剪接和/或转录终止的水平上发生。在各个实施方案中,调节元件可以将转录因子募集到编码区,该转录因子增加在目的细胞类型中的基因表达选择性,提高产生RNA转录物的速率,提高产生的RNA的稳定性,并且/或者

提高从RNA转录物合成蛋白质的速率。

[0250] 在某些实施方案中,本文所述的表达盒进一步包含聚A序列。合适的聚A序列包括,例如,长度约为75bp的人工聚A(PA75)(参见,例如,WO 2018/126116)、牛生长激素聚A、SV40早期聚A信号、SV40晚期聚A信号、兔 β 珠蛋白聚A、HSV胸苷激酶聚A、鱼精蛋白基因聚A、腺病毒5E1b聚A、生长激素聚A或PBGD聚A。在示例性实施方案中,适用于本文提供的表达盒的聚A序列是hGH聚A(SEQ ID NO:327)或合成聚A(SEQ ID NO:326)。通常,聚A序列在本文所述的表达盒中位于编码eTF的多核苷酸的下游。

[0251] 在某些实施方案中,本文提供的表达盒进一步包含编码一种或多种核定位信号(NLS)的一种或多种核酸序列。可以使用任何有助于将与其附接的蛋白质输入到细胞核中的NLS肽。NLS的实例包括,例如,SV40大T抗原NLS、核质蛋白NLS、EGL-13NLS、c-Myc NLS和TUS蛋白NLS。参见,例如,C.Dingwall等人,J.Cell Biol.107:841-9(1988);J.P.Makkerh等人,Curr Biol.6:1025-7(1996);和M.Ray等人,Bioconjug.Chem.26:1004-7(2015)。NLS可以位于eTF蛋白序列上的任何位置,但是在优选的实施方案中,其与eTF的N-末端或eTF的结构域缀合。在示例性实施方案中,本文提供的核酸盒编码具有与eTF的N-末端融合的NLS的eTF。在其他实施方案中,本文提供的核酸盒编码具有与eTF的N-末端融合的第一NLS和位于eTF的DBD与TAD结构域之间的第二NLS的eTF。

[0252] 表达载体

[0253] 在某些实施方案中,本文所述的表达盒可以并入表达载体中。表达载体可以用来通过转染或转导将表达盒递送至靶细胞。载体可以是整合载体或非整合载体,是指载体将表达盒或转基因整合到宿主细胞基因组中的能力。表达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);附加型载体;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、腺病毒载体和腺伴随病毒载体。

[0254] 表达载体可以是线性寡核苷酸或环状质粒,并且可以通过各种转染方法(包括物理和化学方法)递送至细胞。物理方法通常是指利用物理力对抗细胞膜屏障以促进遗传物质的细胞内递送的递送方法。物理方法的实例包括使用针、弹道DNA、电穿孔、声孔效应、光穿孔、磁转染和氢穿孔(hydroporation)。化学方法通常是指其中化学运载体将核酸分子递送至细胞的方法,并且可以包括无机颗粒、基于脂质的载体、基于聚合物的载体和基于肽的载体。

[0255] 在一些实施方案中,使用阳离子脂质(例如,阳离子脂质体)将表达载体施用于靶细胞。已经研究了用于基因递送的各种类型的脂质,例如脂质纳米乳液(例如,其为一种不混溶液体在另一种不混溶液体中的分散体,被乳化剂稳定)或固体脂质纳米颗粒。

[0256] 在一些实施方案中,使用基于肽的递送媒介物将表达载体施用于靶细胞。基于肽的递送媒介物可具有保护待递送的遗传物质、靶向特定细胞受体、破坏内体膜和将遗传物质递送至细胞核中的优点。在一些实施方案中,使用基于聚合物的递送媒介物将表达载体施用于靶细胞。基于聚合物的递送媒介物可包含天然蛋白质、肽和/或多糖或合成聚合物。在一个实施方案中,基于聚合物的递送媒介物包含聚乙烯亚胺(PEI)。PEI可以将DNA浓缩成带正电荷的颗粒,该颗粒与阴离子细胞表面残基结合,并通过内吞作用被带入细胞。在其他

实施方案中,基于聚合物的递送媒介物可以包含聚-L-赖氨酸(PLL)、聚(DL-乳酸)(PLA)、聚(DL-丙交酯-共-糖苷)(PLGA)、聚鸟氨酸、聚精氨酸、组蛋白、鱼精蛋白、树状聚体、壳聚糖、葡聚糖的合成氨基衍生物和/或阳离子丙烯酸聚合物。在某些实施方案中,基于聚合物的递送媒介物可包含诸如PEG和PLL等聚合物的混合物。

[0257] 在某些实施方案中,表达载体可以是适合基因疗法的病毒载体。病毒基因疗法载体或基因递送载体的优选特征可包括可再现地且稳定地增殖并纯化至高滴度的能力;介导靶向递送(例如,将转基因特异性递送至目的组织或器官而没有在他处的广泛载体散播)的能力;以及介导基因递送和转基因表达而不会诱发有害副作用的能力。

[0258] 通过利用病毒感染途径但避免随后会导致复制和毒性的病毒基因表达,已经为了基因疗法的目的工程改造了几种类型的病毒,例如被称为腺伴随病毒的非致病性细小病毒。可以通过从病毒基因组中删除全部或部分编码区,但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可能必需的序列(例如,末端重复序列),从而获得这样的病毒载体。

[0259] 在各个实施方案中,合适的病毒载体包括逆转录病毒(例如,A型、B型、C型和D型病毒)、腺病毒、细小病毒(例如,腺伴随病毒或AAV)、冠状病毒、负链RNA病毒如正粘病毒(例如,流感病毒)、弹状病毒(例如狂犬病和水疱性口炎病毒)、副粘病毒(例如麻疹和仙台病毒)、正链RNA病毒如小RNA病毒和甲病毒,以及双链DNA病毒,包括腺病毒、疱疹病毒(例如单纯疱疹病毒1型和2型、EB病毒、巨细胞病毒)和痘病毒(例如牛痘、禽痘和金丝雀痘)。逆转录病毒的实例包括禽造血白细胞组织增生-肉瘤病毒、人嗜T-淋巴细胞病毒1型(HTLV-1)、牛白血病病毒(BLV)、慢病毒和泡沫病毒。例如,其他病毒包括Norwalk病毒、披膜病毒、黄病毒、呼肠孤病毒、乳多空病毒、嗜肝DNA病毒和肝炎病毒。根据整合到宿主基因组中的能力,病毒载体可被分类为两组——整合型和非整合型。致癌逆转录病毒和慢病毒可以整合到宿主细胞的染色质中,而腺病毒、腺伴随病毒和疱疹病毒则主要以染色体外附加体的形式存在于细胞核中。

[0260] 在某些实施方案中,合适的病毒载体是逆转录病毒载体。逆转录病毒是指逆转录病毒科的病毒。逆转录病毒的实例包括诸如鼠白血病病毒(MLV)等致癌逆转录病毒,以及诸如人类免疫缺陷病毒1(HIV-1)等慢病毒。逆转录病毒基因组是单链(ss)RNA,并且包含可以以顺式或反式提供的各种基因。例如,逆转录病毒基因组可以包含顺式作用序列,如两个长末端重复序列(LTR),以及用于基因表达、逆转录和整合到宿主染色体中的元件。其他组分包括用于将特定RNA包装到新形成的病毒体中的包装信号(ψ 或 ψ)和聚嘌呤束(PPT),后者是逆转录过程中正链DNA合成的起始位点。另外,逆转录病毒基因组可包含gag、pol和env基因。gag基因编码结构蛋白质,pol基因编码与ssRNA伴随并将病毒RNA逆转录为DNA的酶,而env基因编码病毒包膜。通常,gag、pol和env以反式提供,以用于病毒复制和包装。

[0261] 在某些实施方案中,本文提供的逆转录病毒载体可以是慢病毒载体。识别出慢病毒的至少五种血清组或血清型。不同血清型的病毒可以差异地感染某些细胞类型和/或宿主。例如,慢病毒包括灵长类动物逆转录病毒和非灵长类动物逆转录病毒。灵长类动物逆转录病毒包括HIV和猿猴免疫缺陷病毒(SIV)。非灵长类动物逆转录病毒包括猫免疫缺陷病毒(FIV)、牛免疫缺陷病毒(BIV)、山羊关节炎-脑炎病毒(CAEV)、马感染性贫血病毒(EIAV)和维斯纳病毒(visnavirus)。慢病毒或慢病毒载体可以能够转导休眠细胞。与致癌逆转录病

毒载体一样,慢病毒载体的设计可以基于顺式作用序列与反式作用序列的分离。

[0262] 在某些实施方案中,本申请提供了已被设计用于通过优化的治疗性逆转录病毒载体进行递送的表达载体。该逆转录病毒载体可以是慢病毒,其包含左(5')LTR;有助于病毒的包装和/或核输入的序列;启动子;可选的一种或多种额外的调节元件(例如,增强子或聚A序列);可选的慢病毒反向响应元件(RRE);包含与编码eTF的序列可操作地连接的PV选择性调节元件的构建体;可选的绝缘子;和右(3')逆转录病毒LTR。

[0263] 在示例性实施方案中,本文提供的病毒载体是腺伴随病毒(AAV)。AAV是一种小型、复制缺陷型、无包膜的动物病毒,其可感染人类和其他一些灵长类动物物种。尚不知道AAV是否会引起人类疾病,但会诱导轻度的免疫应答。AAV载体还可以感染分裂中的细胞和休眠细胞,而无需整合到宿主细胞基因组中。

[0264] AAV基因组由长度约为4.7kb的线性单链DNA组成。该基因组由侧翼是长度约为145bp的反向末端重复(ITR)序列的两个开放阅读框(ORF)组成。该ITR由5'端的核苷酸序列(5' ITR)和位于3'端的核苷酸序列(3' ITR)组成,它们包含回文序列。该ITR通过折叠经由互补碱基配对形成T形发夹结构而顺式起作用,该结构在第二链合成的DNA复制起始过程中起引物的作用。这两个开放阅读框编码参与病毒体复制和包装的rep和cap基因。在示例性实施方案中,本文提供的AAV载体不含rep或cap基因。这样的基因可以反式提供,以供产生如下文进一步描述的病毒体。

[0265] 在某些实施方案中,AAV载体可以包括填充核酸。在一些实施方案中,该填充核酸可以编码绿色荧光蛋白或抗生素抗性基因,如卡那霉素或氨苄青霉素抗性基因。在某些实施方案中,该填充核酸可以位于ITR序列之外(例如,与位于5' ITR序列与3' ITR序列之间的eTF转基因序列和调节序列相比)。

[0266] 存在各种AAV血清型,包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12和AAV13。这些血清型的向性或所感染的细胞类型不同。AAV可包含来自多种血清型(例如,假型)的基因组和衣壳。例如,AAV可以包含包装在来自血清型5或血清型9的衣壳中的血清型2的基因组(例如,ITR)。假型可以提高转导效率并改变向性。

[0267] 在一些情况下,可以穿过血脑屏障或感染CNS的细胞的AAV血清型是优选的。在一些情况下,AAV9或其变体用来递送本公开的表达盒,其包含与编码上调sCN1A的eTF的转基因可操作地连接的PV选择性调节元件。

[0268] 在示例性实施方案中,本申请提供了已被设计用于通过AAV进行递送的表达载体。该AAV可以是任何血清型,例如AAV1、AAV2、AAV3、AAV3b、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAV-DJ,或嵌合、杂合或变异AAV。该AAV也可以是自互补AAV(scAAV)。在某些实施方案中,设计用于通过AAV进行递送的表达载体包含5' ITR和3' ITR。在某些实施方案中,设计用于通过AAV进行递送的表达载体包含5' ITR、启动子、编码eTF的转基因和3' ITR。在某些实施方案中,设计用于通过AAV进行递送的表达载体包含5' ITR、增强子、启动子、编码eTF的转基因、聚A序列和3' ITR。示例性AAV表达载体在图36和图37中示出。

[0269] 宿主细胞

[0270] 在另一方面,本发明涉及包含编码本发明eTF的表达盒或表达载体的宿主细胞。宿主细胞可以是细菌细胞、酵母细胞、昆虫细胞或哺乳动物细胞。在示例性实施方案中,宿主细胞是指易被目的病毒感染并适于体外培养的任何细胞系。

[0271] 在某些实施方案中,本文提供的宿主细胞可用于离体基因疗法的目的。在这样的实施方案中,用包含编码本发明eTF的序列的核酸分子或表达载体转染细胞,随后将细胞移植到患者或受试者中。移植的细胞可以具有自体的、同种异体的或异源的来源。对于临床用途,通常在良好生产规范(GMP)条件下进行细胞分离。移植前,一般要检查细胞质量以及是否存在微生物或其他污染物,并可以进行预处理,例如采用放射和/或免疫抑制治疗。此外,可以将宿主细胞与生长因子一起移植,以刺激细胞增殖和/或分化。

[0272] 在某些实施方案中,宿主细胞可用于离体基因疗法。优选地,所述细胞是真核细胞,如哺乳动物细胞,该哺乳动物包括但不限于人,非人灵长类动物如猿、黑猩猩、猴和猩猩,驯养动物,包括狗和猫,以及牲畜,如马、牛、猪、绵羊和山羊,或其他哺乳动物物种,包括但不限于小鼠、大鼠、豚鼠、兔、仓鼠等。本领域技术人员将根据待移植的患者或受试者选择更合适的细胞。

[0273] 在某些实施方案中,本文提供的宿主细胞可以是具有自我更新和多能性性质的细胞,如干细胞或诱导多能性干细胞。干细胞优选为间充质干细胞。间充质干细胞(MSC)能够分化为成骨细胞、软骨细胞、脂肪细胞或肌细胞中的至少一种,并且可以从任何类型的组织中分离。通常,MSC将从骨髓、脂肪组织、脐带或外周血中分离。其获得方法是本领域技术人员公知的。诱导多能干细胞(也称为iPS细胞或iPSC)是一种类型的多能干细胞,其可以直接从成体细胞产生。Yamanaka等人通过将Oct3/4、Sox2、Klf4和c-Myc基因转移到小鼠和人成纤维细胞中并迫使所述细胞表达该基因,而诱导了iPS细胞(WO 2007/069666)。Thomson等人随后使用Nanog和Lin28代替Klf4和c-Myc产生了人iPS细胞(WO 2008/118820)。

[0274] 在示例性实施方案中,本文提供的宿主细胞是包装细胞。所述细胞可以是贴壁细胞或悬浮细胞。包装细胞和辅助载体或病毒或DNA构建体以反式一起提供了病毒载体的完全复制和包装所需的所有缺失功能。

[0275] 优选地,所述包装细胞是真核细胞,如哺乳动物细胞,包括猿猴、人、狗和啮齿动物细胞。人类细胞的实例是PER.C6细胞(W001/38362)、MRC-5(ATCC CCL-171)、WI-38(ATCC CCL-75)、HEK-293细胞(ATCC CRL-1573)、HeLa细胞(ATCC CCL2)和胚胎恒河猴肺细胞(ATCC CL-160)。非人灵长类动物细胞的实例是Vero细胞(ATCC CCL81)、COS-1细胞(ATCC CRL-1650)或COS-7细胞(ATCC CRL-1651)。狗细胞的实例是MDCK细胞(ATCC CCL-34)。啮齿动物细胞的实例是仓鼠细胞,如BHK21-F、HKCC细胞或CHO细胞。

[0276] 作为哺乳动物来源的替代,用于本发明的细胞系可以来源于禽类来源,如鸡、鸭、鹅、鹌鹑或雉鸡。禽细胞系的实例包括禽胚胎干细胞(W001/85938和W003/076601)、无限增殖化的鸭视网膜细胞(W02005/042728)和禽胚胎干细胞衍生的细胞,包括鸡细胞(W02006/108846)或鸭细胞,如EB66细胞系(W02008/129058和W02008/142124)。

[0277] 在另一个实施方案中,所述宿主细胞是昆虫细胞,如SF9细胞(ATCC CRL-1711)、Sf21细胞(IPLB-Sf21)、MG1细胞(BTI-TN-MG1)或High FiveTM细胞(BTI-TN-5B1-4)。

[0278] 在某些实施方案中,本文提供的包含本发明重组AAV载体/基因组(例如,包含编码eTF的序列)的宿主细胞可进一步包含一种或多种额外的核酸构建体,例如,(i) 编码rep和cap基因但不携带ITR序列的核酸构建体(例如,AAV辅助质粒),和/或(ii) 提供AAV复制所必需的腺病毒功能的核酸构建体(例如质粒)。在示例性实施方案中,本文提供的宿主细胞包含:i) 包含本发明的编码eTF的序列(即,重组AAV基因组)的表达载体;ii) 编码不携带ITR序

列的AAV rep和cap基因的核酸构建体;以及iii)包含腺病毒辅助基因的核酸构建体(如下文进一步描述)。

[0279] 在某些实施方案中,rep基因、cap基因和腺病毒辅助基因可以组合在一个质粒上(Blouin V等人.J Gene Med.2004;6(suppl):S223-S228;Grimm D.等人.Hum.Gene Ther.2003;7:839-850)。因此,在另一个示例性实施方案中,本文提供的宿主细胞包含:i)包含本发明的编码eTF的序列(即,重组AAV基因组)的表达载体;以及ii)编码不携带ITR序列的AAV rep和cap基因且进一步包含腺病毒辅助基因的质粒。

[0280] 在另一个实施方案中,本文提供的宿主细胞包含:a)包含本发明的编码eTF的序列(即,重组AAV基因组)的表达载体;b)编码不携带ITR序列的AAV rep和cap基因的质粒;以及c)包含腺病毒辅助基因E2a、E4和VA RNA的质粒;其中在组成型地表达并反式补充腺病毒E1基因的细胞,优选哺乳动物细胞,如HEK-293细胞(ATCC CRL-1573)中进行共转染。

[0281] 在某些实施方案中,适合于大规模生产AAV载体的宿主细胞是可以用重组杆状病毒的组合来感染的昆虫细胞(Urabe等人,Hum.Gene Ther.2002;13:1935-1943)。例如,SF9细胞可以用分别表达AAV rep、AAV cap和待包装的AAV载体的三种杆状病毒载体来共感染。重组杆状病毒载体将提供病毒复制和/或包装所需的病毒辅助基因功能。

[0282] 关于根据本发明用于基因疗法的病毒体的构建和生产的进一步指导可见于:Viral Vectors for Gene Therapy,Methods and Protocols.Series:Methods in Molecular Biology,Vol.737.Merten和Al-Rubeai(编),2011Humana Press(Springer);Gene Therapy.M.Giacca.2010Springer-Verlag;Heilbronn R.和Weger S.Viral Vectors for Gene Transfer:Current Status of Gene Therapeutics.于:Drug Delivery,Handbook of Experimental Pharmacology 197;M.Schafer-Korting(编).2010Springer-Verlag;pp.143-170;Adeno-Associated Virus:Methods and Protocols.R.O.Snyder和P.Moullier(编).2011Humana Press(Springer);Bunning H.等人,Recent developments in adeno-associated virus technology.J.Gene Med.2008;10:717-733;以及Adenovirus:Methods and Protocols.M.Chillon和A.Bosch(编);第三版.2014Humana Press(Springer)。

[0283] 病毒体和生产病毒体的方法

[0284] 在某些实施方案中,本申请提供了包含病毒载体的病毒颗粒,该病毒载体包含本发明的编码eTF的序列。术语“病毒颗粒”和“病毒体”在本文中可互换使用,并且涉及感染性的且通常是复制缺陷性的病毒颗粒,其包含包装在衣壳中的病毒基因组(例如,病毒表达载体),并且视情况而定,例如,对于逆转录病毒,其包含包围衣壳的脂质包膜。“衣壳”是指其中包装有病毒基因组的结构。衣壳由若干个寡聚体结构亚单位组成,该亚单位由蛋白质构成。例如,AAV具有由以下三种衣壳蛋白的相互作用形成的二十面体衣壳:VP1、VP2和VP3。在一个实施方案中,本文提供的病毒体是通过将包含与本文所述编码eTF的序列可操作地连接的PV选择性调节元件的AAV载体包装在蛋白质壳中而获得的重组AAV病毒体或rAAV病毒体。

[0285] 在某些实施方案中,可以通过将来源于特定AAV血清型的AAV基因组用衣壳包装在由天然Cap蛋白(对应于相同特定血清型的AAV)形成的病毒颗粒中来制备本文提供的重组AAV病毒体。在其他实施方案中,本文提供的AAV病毒颗粒包含以下病毒载体,其包含包装到

来自不同血清型的蛋白质中的给定AAV血清型的ITR。参见,例如,Bunning H等人,J Gene Med 2008;10:717-733。例如,具有来自给定AAV血清型的ITR的病毒载体可以包装到:a)由来源于相同或不同AAV血清型的衣壳蛋白(例如,AAV2 ITR和AAV9衣壳蛋白;AAV2 ITR和AAV8衣壳蛋白;等等)构成的病毒颗粒;b)由来自不同AAV血清型或突变体的衣壳蛋白混合物(例如,AAV2 ITR与AAV1和AAV9衣壳蛋白)构成的镶嵌型病毒颗粒;c)由已通过在不同AAV血清型或变体之间进行域交换而被截短的衣壳蛋白(例如,AAV2 ITR与具有AAV9域的AAV8衣壳蛋白)构成的嵌合病毒颗粒;或d)经工程改造而展示出选择性结合域,从而能够与靶细胞特异性受体发生严格相互作用的靶向病毒颗粒(例如,AAV5 ITR与通过插入肽配体而被遗传截短的AAV9衣壳蛋白;或通过将肽配体偶联至衣壳表面而非遗传修饰的AAV9衣壳蛋白)。

[0286] 技术人员将会理解,本文提供的AAV病毒体可包含任何AAV血清型的衣壳蛋白。在一个实施方案中,病毒颗粒包含来自选自AAV1、AAV2、AAV5、AAV8和AAV9的AAV血清型的衣壳蛋白,它们更适于递送至CNS(M.Hocquemiller等人,Hum Gene Ther 27(7):478-496(2016))。在特定实施方案中,病毒颗粒包含本发明的表达盒,其中该表达盒的5' ITR和3' ITR序列是AAV2血清型的而衣壳蛋白是AAV9血清型的。

[0287] 用于生产rAAV病毒体的许多方法是本领域已知的,包括转染、稳定细胞系生产以及包括腺病毒-AAV杂合体、疱疹病毒-AAV杂合体(Conway,J E等人(1997)J.Virology 71(11):8780-8789)和杆状病毒-AAV杂合体在内的感染性杂合病毒生产系统。用于生产rAAV病毒颗粒的rAAV生产培养物都需要:1)合适的宿主细胞,例如在杆状病毒生产系统的情况下,包括人源细胞系,如HeLa、A549或293细胞,或昆虫来源的细胞系,如SF-9;2)合适的辅助病毒功能,其由野生型或突变型腺病毒(如温度敏感性腺病毒)、疱疹病毒、杆状病毒或提供辅助功能的质粒构建体提供;3)AAV rep和cap基因及基因产物;4)侧翼为AAV ITR序列的转基因(例如,与本文所述的编码eTF的序列可操作地连接的启动子);以及5)支持rAAV生产的合适的培养基和培养基组分。

[0288] 在各个实施方案中,本文所述的宿主细胞包含以下三种组分:(1)rep基因和cap基因,(2)提供辅助功能的基因,以及(3)转基因(例如,侧翼为ITR的与本文所述的编码eTF的序列可操作地连接的启动子)。AAV rep基因、AAV cap基因和提供辅助功能的基因可通过将所述基因并入载体如质粒中并将所述载体引入宿主细胞中而引入细胞中。rep、cap和辅助功能基因可以并入到相同质粒或不同质粒中。在优选实施方案中,将AAV rep和cap基因并入一个质粒中,并将提供辅助功能的基因并入另一质粒中。可以通过使用本领域公知的任何合适的方法将用于创建产生病毒体的宿主细胞的各种质粒(例如,包含AAV rep和cap基因、辅助功能或转基因)引入细胞中。转染方法的实例包括但不限于磷酸钙共沉淀、DEAE-葡聚糖、聚凝胺、电穿孔、显微注射、脂质体介导的融合、脂质转染、逆转录病毒感染和生物射弹转染。在某些实施方案中,可将提供rep和cap基因、辅助功能和转基因(例如,侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子)的质粒同时引入细胞中。在另一个实施方案中,提供rep和cap基因以及辅助功能的质粒可以在引入包含转基因的质粒之前或之后被引入细胞中。在示例性实施方案中,用以下三种质粒同时转染细胞(例如,三重转染法):(1)包含转基因(例如,侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子)的质粒,(2)包含AAV rep和cap基因的质粒,以及(3)包含提供辅助功能的基因的质粒。

示例性宿主细胞可以是293、A549或HeLa细胞。

[0289] 在其他实施方案中, (1) AAV rep和cap基因、(2) 提供辅助功能的基因和(3) 转基因中的一个或多个可以由包装细胞游离地携带和/或整合到包装细胞的基因组中。在一个实施方案中, 宿主细胞可以是包装细胞, 其中AAV rep和cap基因以及辅助功能在宿主细胞中稳定地保留, 并且用含有转基因 (例如, 侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子) 的质粒瞬时转染宿主细胞。在另一个实施方案中, 宿主细胞是包装细胞, 其中AAV rep和cap基因在宿主细胞中稳定地保留, 并且用含有转基因 (例如, 侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子) 的质粒和包含辅助功能的质粒瞬时转染宿主细胞。在另一个实施方案中, 宿主细胞可以是包装细胞, 其中辅助功能在宿主细胞中稳定地保留, 并且用含有转基因 (例如, 侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子) 的质粒和包含rep和cap基因的质粒瞬时转染宿主细胞。在另一个实施方案中, 宿主细胞可以用rep和cap基因、辅助功能和转基因序列 (例如, 侧翼为ITR的与本文公开的编码eTF的序列可操作地连接的启动子) 稳定转染的生产细胞系。示例性包装和生产细胞可以来源于293、A549或HeLa细胞。

[0290] 在另一个实施方案中, 生产细胞系是用提供Rep和Cap蛋白的杆状病毒表达载体感染的昆虫细胞系 (通常是Sf9细胞)。该系统不需要腺病毒辅助基因 (Ayuso E等人, Curr. Gene Ther. 2010, 10: 423-436)。

[0291] 如本文所用的, 术语“cap蛋白”是指具有天然AAV Cap蛋白 (例如, VP1、VP2、VP3) 的至少一种功能活性的多肽。cap蛋白的功能活性的实例包括诱导衣壳形成、促进单链DNA积累、促进AAV DNA包装到衣壳中 (即衣壳化)、结合细胞受体以及促进病毒体进入宿主细胞中的能力。原则上, 任何Cap蛋白均可用于本发明。

[0292] 已经报道, Cap蛋白对宿主向性, 细胞、组织或器官特异性, 受体使用、感染效率和AAV病毒的免疫原性具有影响。因此, 可以考虑例如受试者的物种 (例如人类或非人类)、受试者的免疫学状态、受试者对于长期或短期治疗的适合性或具体的治疗应用 (例如, 具体疾病或病症的治疗, 或向特定细胞、组织或器官的递送) 来选择用于rAAV的AAV cap。在某些实施方案中, cap蛋白来源于选自AAV1、AAV2、AAV5、AAV8和AAV9血清型的AAV。在示例性实施方案中, cap蛋白来源于AAV9。

[0293] 在一些实施方案中, 用于本发明方法的AAV Cap可以通过一种上述AAV cap或其编码核酸的诱变 (即通过插入、缺失或置换) 来产生。在一些实施方案中, 该AAV cap与一种或多种上述AAV cap至少70%、75%、80%、85%、90%、95%、98%或99%或更高水平相似。

[0294] 在一些实施方案中, AAV cap是嵌合的, 包含来自两种、三种、四种或更多种上述AAV cap的结构域。在一些实施方案中, 该AAV cap是源自两种或三种不同AAV或重组AAV的VP1、VP2和VP3单体的嵌合体。在一些实施方案中, rAAV组合物包含超过一种上述cap。

[0295] 在一些实施方案中, 用于rAAV病毒体的AAV cap被工程改造为包含异源序列或其他修饰。例如, 可以将赋予选择性靶向或免疫逃避的肽或蛋白序列工程改造到cap蛋白中。备选地或另外地, 可以对cap进行化学修饰, 使得rAAV的表面被聚乙二醇化 (即peg化), 这可以促进免疫逃避。还可以对cap蛋白进行诱变 (例如, 以去除其天然受体结合或掩盖免疫原性表位)。

[0296] 如本文所用的, 术语“rep蛋白”是指具有天然AAV rep蛋白 (例如, rep 40、52、68、

78)的至少一种功能活性的多肽。rep蛋白的功能活性的实例包括与该蛋白质的生理功能相关的任何活性,包括通过识别来促进DNA复制、结合并切割DNA复制的AAV起点以及DNA解旋酶活性。其他功能包括调节从AAV(或其他异源)启动子的转录以及AAV DNA向宿主染色体中的位点特异性整合。在特定实施方案中,AAV rep基因可以来自血清型AAV1、AAV2、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10或AAVrh10;更优选地,来自选自AAV1、AAV2、AAV5、AAV8和AAV9的AAV血清型。

[0297] 在一些实施方案中,用于本发明方法的AAV rep蛋白可以通过一种上述AAV rep或其编码核酸的诱变(即通过插入、缺失或置换)来产生。在一些实施方案中,该AAV rep与一种或多种上述AAV rep至少70%、75%、80%、85%、90%、95%、98%或99%或更高水平相似。

[0298] 如本文所用的,表述“辅助功能”或“辅助基因”是指AAV复制所依赖的病毒蛋白质。辅助功能包括AAV复制所需的那些蛋白质,包括但不限于参与AAV基因转录激活、阶段特异性AAV mRNA剪接、AAV DNA复制、cap表达产物的合成和AAV衣壳装配的那些蛋白质。基于病毒的辅助功能可以来源于任何已知的辅助病毒,如腺病毒、疱疹病毒(除单纯疱疹病毒1型以外)和痘苗病毒。辅助功能包括但不限于腺病毒E1、E2a、VA和E4或疱疹病毒UL5、ULB、UL52和UL29以及疱疹病毒聚合酶。在优选的实施方案中,AAV复制所依赖的蛋白质来源于腺病毒。

[0299] 在一些实施方案中,用于本发明方法的AAV复制所依赖的病毒蛋白质可以通过一种上述病毒蛋白质或其编码核酸的诱变(即通过插入、缺失或置换)来产生。在一些实施方案中,该病毒蛋白质与一种或多种上述病毒蛋白质至少70%、75%、80%、85%、90%、95%、98%或99%或更高水平相似。

[0300] 测定AAV复制所依赖的cap蛋白、rep蛋白和病毒蛋白质的功能的方法是本领域公知的。

[0301] 表达目的转基因(例如,与编码eTF的序列可操作地连接的启动子)的宿主细胞可以在足以装配AAV病毒体的条件下生长。在某些实施方案中,宿主细胞生长一段合适的时间,以促进AAV病毒体的装配和病毒体向培养基中的释放。通常,细胞可以生长约24小时、约36小时、约48小时、约72小时、约4天、约5天、约6天、约7天、约8天、约9天或至多约10天。在约10天后(或更早,取决于培养条件和所用的具体宿主细胞),生产水平通常显著下降。通常,从病毒产生的点开始测量培养时间。例如,在AAV的情况下,病毒产生通常在如本文所述的合适的宿主细胞中提供辅助病毒功能时开始。通常,在辅助病毒感染后(或在病毒产生开始后)约48至约100、优选约48至约96、优选约72至约96、优选约68至约72小时收获细胞。

[0302] rAAV生产培养物可以在适合于所使用的具体宿主细胞的多种条件(在宽温度范围内、持续不同时间长度等)下生长。rAAV生产培养物包括附着依赖性培养物,其可以在合适的附着依赖性容器,例如滚瓶、中空纤维过滤器、微载体以及填充床或流化床生物反应器中培养。rAAV载体生产培养物还可以包含适应悬浮的宿主细胞,如HeLa、293和SF-9细胞,其可以通过多种方式培养,包括,例如旋转烧瓶、搅拌罐生物反应器和一次性系统,如Wave袋系统。

[0303] 本领域已知的合适的培养基可以用于产生rAAV病毒体。这些培养基包括但不限于Hyclone Laboratories和JRH生产的培养基,包括改良Eagle培养基(MEM)、Dulbecco改良

Eagle培养基(DMEM),其各自通过引用整体并入本文。在某些实施方案中,rAAV生产培养基可以补充有0.5%-20%(v/v或w/v)水平的血清或血清衍生重组蛋白。或者,可以在无血清条件下产生rAAV载体,该无血清条件也可以被称为没有动物来源的产物的培养基。

[0304] 在培养宿主细胞以允许产生AAV病毒体后,可以收获并纯化所得的病毒体。在某些实施方案中,(1)可以通过裂解宿主细胞从生产培养物的宿主细胞获得AAV病毒体,和/或(2)可以在转染一段时间(优选72小时)后从所述细胞的培养基中获得AAV病毒体。只要在导致rAAV病毒体从完整细胞释放到培养基中的条件下培养细胞,就可以从生产培养物中用过的培养基收获rAAV病毒体(参见,例如,美国专利6,566,118)。裂解细胞的合适方法在本领域中也是已知的,并且包括例如多个冷冻/融化循环、声处理、微流化和用化学物质如去污剂和/或蛋白酶处理。

[0305] 收获后,可以纯化rAAV病毒体。如本文所用的术语“纯化的”包括不含至少一些其他组分的rAAV病毒体的制品,这些其他组分可以存在于天然出现rAAV病毒体之处或最初制备rAAV病毒体之处。因此,例如,可以使用分离技术从来源混合物如培养裂解物或生产培养上清液中富集rAAV病毒体,来制备纯化的rAAV病毒体。富集可以通过多种方式来测量,例如通过溶液中存在的DNA酶抗性颗粒(DRP)或基因组拷贝(gc)的比例,或通过感染性,或者可以相对于来源混合物中存在的第二种潜在干扰性物质(例如污染物,包括生产培养污染物或过程中污染物,包括辅助病毒、培养基组分等)来测量。

[0306] 在某些实施方案中,可以将rAAV生产培养收获物进行澄清,以去除宿主细胞碎片。在一些实施方案中,可以使用多种标准技术,如离心或通过孔径为0.2 μ m或更大的过滤器(例如,乙酸纤维素过滤器或一系列深层过滤器)过滤,将生产培养收获物进行澄清。

[0307] 在某些实施方案中,将rAAV生产培养收获物进一步用BenzonaseTM处理,以消化生产培养物中存在的任何高分子量DNA。在一些实施方案中,BenzonaseTM消化在标准条件下进行,例如,BenzonaseTM的终浓度为1-2.5单位/ml,温度范围从环境温度至37 $^{\circ}$ C,时间为30分钟至数小时。

[0308] 在某些实施方案中,可以使用以下一种或多种纯化步骤来分离或纯化rAAV病毒体:平衡离心;流通式阴离子交换过滤;用于浓缩rAAV颗粒的切向流过滤(TFF);通过磷灰石色谱法捕获rAAV;辅助病毒热灭活;通过疏水相互作用色谱法捕获rAAV;通过大小排阻色谱法(SEC)进行缓冲液交换;纳米过滤;以及通过阴离子交换色谱法、阳离子交换色谱法或亲和色谱法捕获rAAV。这些步骤可以单独使用,以各种组合使用,或以不同的顺序使用。纯化rAAV颗粒的方法可见于,例如,Xiao等人,(1998)Journal of Virology 72:2224-2232;美国专利6,989,264和8,137,948;以及WO 2010/148143。

[0309] 在某些实施方案中,纯化的AAV病毒体可以用PBS透析、过滤并储存在-80 $^{\circ}$ C。病毒基因组的滴度可以使用线性化质粒DNA作为标准曲线,通过定量PCR来确定(参见,例如,Lock M等人,Hum.Gene Ther.2010;21:1273-1285)。

[0310] 药物组合物

[0311] 在某些实施方案中,本申请提供了包含编码eTF的序列和药学上可接受的载体的组合物。在其他实施方案中,本申请提供了包含编码eTF的序列和药学上可接受的载体的病毒体。在示例性实施方案中,这样的组合物适用于基因疗法应用。药物组合物优选地在制备和储存条件下是无菌的和稳定的。无菌溶液可以通过例如通过无菌滤膜过滤来实现。

[0312] 药物组合物中可接受的载体和赋形剂优选地在所采用的剂量和浓度下对接受者无毒。可接受的载体和赋形剂可包括缓冲液,如磷酸盐、柠檬酸盐、EPES和TAE,抗氧化剂,如抗坏血酸和甲硫氨酸,防腐剂,如氯化六甲双铵、十八烷基二甲基苄基氯化铵、间苯二酚和苯扎氯铵,蛋白质,如人血清白蛋白、明胶、葡聚糖和免疫球蛋白,亲水聚合物,如聚乙烯吡咯烷酮,氨基酸,如甘氨酸、谷氨酰胺、组氨酸和赖氨酸,以及碳水化合物,如葡萄糖、甘露糖、蔗糖和山梨醇。本公开的药物组合物可以以可注射制剂的形式肠胃外施用。可以使用无菌溶液或任何药学上可接受的液体作为媒介物来配制用于注射的药物组合物。药学上可接受的媒介物包括但不限于无菌水和生理盐水。

[0313] 本公开的药物组合物可以在微胶囊如羟甲基纤维素或明胶-微胶囊和聚甲基丙烯酸甲酯微胶囊中制备。本公开的药物组合物还可以在其他药物递送系统如脂质体、白蛋白微球、微乳剂、纳米颗粒和纳米胶囊中制备。用于基因疗法的药物组合物可以在可接受的稀释剂中,或者可以包含其中嵌有基因递送媒介物的缓释基质。

[0314] 本文提供的药物组合物可以被配制用于肠胃外给药、皮下给药、静脉内给药、肌肉内给药、动脉内给药、实质内给药、鞘内给药、小脑延髓池内给药、脑室内给药或腹膜内给药。药物组合物还可以被配制用于或通过鼻、喷雾、口、气雾剂、直肠或阴道给药。在一个实施方案中,本文提供的药物组合物被施用于CNS或脑脊液(CSF),即通过实质内注射、鞘内注射、小脑延髓池内注射或脑室内注射施用。组织靶标可以是特异性的,例如CNS特异性的,或者它可以是若干种组织例如肌肉和CNS组织的组合。示例性的组织或其他靶标可以包括肝、骨骼肌、心肌、脂肪沉积物、肾、肺、血管内皮、上皮、造血细胞、CNS和/或CSF。在优选的实施方案中,本文提供的包含上调SCN1A或GRN的ETF的药物组合物被施用于CNS或CSF,即通过实质内注射、鞘内注射、小脑延髓池内注射或脑室内注射施用。这些方法中的一种或多种可以用来施用本公开的药物组合物。

[0315] 在某些实施方案中,本文提供的药物组合物包含“有效量”或“治疗有效量”。如本文所用的,这样的量是指在达到所需治疗结果所必需的剂量和时间段下有效的量,该治疗结果例如是增加SCN1A表达的水平 and/或降低癫痫发作的频率和/或持续时间,或增加GRN表达的水平 and/或治疗与GRN相关的疾病或病症,如FTD。

[0316] 本公开的药物组合物的剂量取决于包括给药途径、待治疗的疾病和受试者的身体特性(例如年龄、体重、一般健康)在内的因素。可以调整剂量以提供最佳治疗反应。一般而言,剂量可以是有效治疗疾病而不诱发显著毒性的量。在一个实施方案中,本文提供的AAV载体可以以 5×10^{11} 至 1×10^{14} gc/kg(基因组拷贝/千克患者体重(gc/kg))范围内的量或剂量施用于患者,以治疗SCN1A缺乏(包括,例如,Dravet综合征),或治疗GRN缺乏(包括,例如,FTD)。在更具体的实施方案中,该AAV载体以约 5×10^{11} gc/kg至约 3×10^{13} gc/kg,或约 1×10^{12} 至约 1×10^{14} gc/kg,或约 1×10^{12} 至约 1×10^{13} gc/kg范围内包含的量,或约 5×10^{11} gc/kg、 1×10^{12} gc/kg、 1.5×10^{12} gc/kg、 2.0×10^{12} gc/kg、 2.5×10^{12} gc/kg、 3×10^{12} gc/kg、 3.5×10^{12} gc/kg、 4×10^{12} gc/kg、 4.5×10^{12} gc/kg、 5×10^{12} gc/kg、 5.5×10^{12} gc/kg、 6×10^{12} gc/kg、 6.5×10^{12} gc/kg、 7×10^{12} gc/kg、 7.5×10^{12} gc/kg、 8×10^{12} gc/kg、 8.5×10^{12} gc/kg、 9×10^{12} gc/kg或 9.5×10^{12} gc/kg施用。例如,可以通过qPCR或数字小液滴PCR(ddPCR)来确定gc/kg(参见,例如,M.Lock等人,Hum Gene Ther Methods.2014年4月;25(2):115-25)。在另一个实施方案中,本文提供的AAV载体可以以 1×10^9 至 1×10^{11} iu/kg(载体的感染单位(iu)/受试者或患者

体重(kg))范围内的量或剂量施用于患者,以治疗SCN1A缺乏(包括,例如,Dravet综合征)。在某些实施方案中,所述药物组合物可以根据需要以单位剂量形成。这样的单剂量单位可包含约 1×10^9 gc至约 1×10^{15} gc。

[0317] 本公开的药物组合物可以例如每天、每周、每月、每半年、每年一次或多次(例如,1-10次或更多次)或根据医学需要施用于有需要的受试者。在示例性实施方案中,单次给药是足够的。在一个实施方案中,包含编码上调SCN1A或GRN的eTF的表达盒的药物组合物适用于人类受试者,并且通过实质内注射、鞘内注射、小脑延髓池内注射或脑室内注射施用。在一个实施方案中,该药物组合物通过团注经由外周静脉递送。在其他实施方案中,该药物组合物通过经约10分钟(± 5 分钟)、经约20分钟(± 5 分钟)、经约30分钟(± 5 分钟)、经约60分钟(± 5 分钟)或经约90分钟(± 10 分钟)的输注而经由外周静脉递送。

[0318] 在另一方面,本申请进一步提供了一种试剂盒,其包含在一个或多个容器中的本文所述的核酸分子、载体、宿主细胞、病毒体或药物组合物。试剂盒可包括说明书或包装材料,其说明如何将试剂盒内包含的核酸分子、载体、宿主细胞或病毒体施用于患者。试剂盒的容器可以是任何合适的材料,例如玻璃、塑料、金属等,并且可以是任何合适的大小、形状或构造。在某些实施方案中,试剂盒可包括一个或多个安瓿或注射器,其中包含合适的液体或溶液形式的核酸分子、载体、宿主细胞、病毒体或药物组合物。

[0319] 治疗方法

[0320] 在各个实施方案中,本申请提供了使用本文公开的eTF的方法。在某些实施方案中,本申请提供了用于施用表达盒、表达载体或病毒颗粒的方法,该表达盒、表达载体或病毒颗粒包含编码本文公开的eTF的多核苷酸,以调节目的基因在细胞中的表达。在某些实施方案中,该eTF可以上调目的基因的表达。在其他实施方案中,该eTF可以下调目的基因的表达。在各个实施方案中,该eTF可以用来在体外、在体内或离体调节目的基因在细胞中的表达。

[0321] 在某些实施方案中,本申请提供了通过向有需要的受试者施用表达盒、表达载体或病毒颗粒来治疗与目的基因相关的疾病或病症的方法,该表达盒、表达载体或病毒颗粒包含编码调节目的基因表达的eTF的多核苷酸。在示例性实施方案中,该疾病或病症与目的基因的单倍体不足相关,并且该表达盒、表达载体或病毒颗粒包含编码上调目的基因表达的eTF的多核苷酸。在某些实施方案中,治疗疾病或病症的方法包括施用表达盒、表达载体或病毒颗粒,该表达盒、表达载体或病毒颗粒包含编码调节与这类疾病或病症相关的内源目的基因的表达的eTF的多核苷酸,从而纠正这类基因的过度表达或表达不足,使其处于健康个体的水平内,或处于医护标准所定义的正常范围内。在某些实施方案中,本文公开的方法用来治疗与内源基因相关的疾病或病症,该内源基因包含导致该基因的异常表达的一个或多个突变。在一些情况下,这样的内源基因是由于过大而无法在基因疗法中作为转基因递送的基因,或者是当作为转基因递送时不能有效表达的基因。

[0322] 在某些实施方案中,本申请提供了通过向有需要的受试者施用表达盒、表达载体或病毒颗粒来减轻与疾病或病症相关的症状的方法,该表达盒、表达载体或病毒颗粒包含编码调节与这类疾病或病症相关的基因表达的eTF的多核苷酸。

[0323] 在示例性实施方案中,本申请提供了通过向有需要的受试者施用包含编码上调SCN1A基因或其蛋白质产物Nav1.1表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒来

治疗具有SCN1A突变、Nav1.1缺乏和/或Nav1.1活性降低的疾病、病症或症状的方法。电压门控钠离子通道对于横纹肌和神经元组织中动作电位的产生和传播至关重要。电压门控钠离子通道是由大中央孔形成糖基化 α 亚单位和2个较小的辅助 β 亚单位组成的异聚复合物。由SCN1A基因编码的大 α 亚单位Nav1.1亚单位与多种疾病或病症如Dravet综合征有关。Nav1.1在神经元中表达,并且可以由包括由SCN1B基因表达的Nav β 1在内的各种 β 亚单位装配。

[0324] 在某些实施方案中,本申请提供了使用上调内源SCN1A基因表达的eTF治疗与SCN1A突变或Nav1.1活性降低相关的疾病的方法。与SCN1A突变相关的疾病和病症包括但不限于:Dravet综合征、大田原综合征、癫痫、早期婴儿型癫痫性脑病6(EIEE6)、家族性热性癫痫发作3A(FEB3A)、难治性儿童癫痫伴全身性强直-阵挛性癫痫发作(ICEGTC)、家族性偏瘫性偏头痛3(FHM3)、Panayiotopoulos综合征、家族性心房颤动13(ATFB13)、全身性癫痫伴热性癫痫发作+1型(gefs+1型)、Brugada综合征、非特异性心脏传导缺陷、全身性癫痫伴热性癫痫发作+、良性家族性婴儿癫痫发作、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、tau蛋白病变和阿尔茨海默病。在一些情况下,该神经系统病况是Dravet综合征。SCN1A的突变或异常还与癫痫病症、癫痫、自闭症、家族性偏瘫性偏头痛3型(FHM3)、遗传性癫痫伴热性癫痫发作+(GEFS+)和某些抗癫痫发作药物的有效性相关。例如,SCN1A中的ICS5N+5G>A突变与抗癫痫药苯妥英和卡马西平的最大安全量(剂量)相关。

[0325] 在某些实施方案中,本申请提供了通过施用包含编码上调SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒来治疗患有Dravet综合症或具有患Dravet综合症风险的受试者的方法。Dravet综合症的特征在于在儿童期第一年内延长的发热和非发热癫痫发作。这种疾病会进展为其他癫痫发作类型,例如肌阵挛性癫痫发作和部分癫痫发作、精神运动迟缓和共济失调。其特征认知损害、行为失常和运动缺陷。行为缺陷通常包括多动和冲动,在极少数情况下,包括孤独样行为。Dravet综合征还与包括嗜睡和失眠在内的睡眠障碍相关。在许多患者中,Dravet综合征是由导致产生无功能蛋白质的基因突变引起的。在治疗与遗传原因相关的病症中存在许多挑战。因此,大多数现有的治疗已被用于癫痫发作和其他症状的预防性医疗处理。

[0326] 在70-90%的患者中,Dravet综合征是由SCN1A基因中的无义突变引起的,该突变导致提前终止密码子,因而导致无功能性蛋白质。通常,钠通道孔的S5或S6区段中的错义突变导致通道功能的丧失和Dravet综合征的发展。SCN1A突变的杂合遗传是形成有缺陷的钠通道所必需的;Dravet综合征患者仍具有该基因的一个正常拷贝。因此,该疾病被表征为单倍体不足的疾病,并且增加SCN1A的功能性拷贝的表达可以恢复Nav1.1的正常产生水平。

[0327] 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。用本文所述的表达盒、表达载体或病毒颗粒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。向具有发生Dravet综合征的风险的受试者施用本文所述的基因疗法可以防止Dravet综合征的一个或多个症状的发展或减缓其进展。

[0328] 在某些实施方案中,与未治疗的对照相比或与治疗前水平相比,用本文所述的包含编码上调SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗将癫痫发作持续时间和/或频率,例如与Dravet综合征相关的癫痫发作降低至少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%、95%或更多。

[0329] 在一些阿尔茨海默病患者中,淀粉样蛋白 β (A β)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。在另一个实施方案中,本申请提供了通过施用包含编码上调SCN1A的eTF的多核苷酸的本文所述表达盒、表达载体或病毒颗粒来治疗罹患阿尔茨海默病的受试者的方法。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。用包含编码上调SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行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,该治疗可以导致高 γ 功率脑活动的校正。该治疗可以导致癫痫发作频率和/或癫痫发作严重程度降低,或与未治疗相比将高 γ 功率活动减少10%、20%、30%、40%、50%、60%、70%或更多。在一些情况下,该治疗可以导致认知功能改善。与未治疗或用本文公开的编码上调SCN1A的eTF的多核苷酸治疗前相比,学习和/或记忆可以改善至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或超过100%。

[0330] 在一些情况下,与未治疗的对照相比或与治疗前水平相比,用包含编码上调SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗将高 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%。

[0331] 帕金森症是指在帕金森病(PD)中发现的体征和症状的集合,包括缓慢(运动徐缓)、僵硬(强直)、震颤和失衡(姿势不稳)。在一些情况下,向具有发生或罹患帕金森病的风险的受试者施用本文所述包含编码上调SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可以预防其一种或多种症状的发展,或者与未治疗相比,将帕金森病的进展减慢至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%或至少90%。

[0332] 在某些实施方案中,本申请提供了可用于治疗具有患病风险的受试者的方法。可能已知受试者易患疾病,例如神经系统疾病或与癫痫、癫痫发作和/或脑病相关的疾病。由于遗传事件,或由于已知的危险因素,受试者可能易患疾病。例如,受试者可以在SCN1A中携带与Dravet综合征相关的突变。在一些情况下,由于受试者的年龄,受试者可能易患诸如阿尔茨海默病的疾病。在一些情况下,受试者可能具有不足量的SCN1A蛋白,并且治疗与SCN1A相关的疾病包括施用本文所述包含编码上调内源SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒。

[0333] 在某些实施方案中,使用本文提供的包含编码上调内源SCN1A的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗可导致与Dravet或其他SCN1A相关疾病或病症相关的症状的减轻或停止。例如,治疗可以改善学习、记忆、认知功能和/或运动功能;降低癫痫发作的频率和/或持续时间;并且/或者降低温度敏感性(或提高触发癫痫发作的温度阈值)。

[0334] 在另一个示例性实施方案中,本申请提供了通过向有需要的受试者施用包含编码

上调GRN基因或蛋白质产物表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒来治疗与GRN基因突变或GRN蛋白缺乏或活性降低相关的疾病或病症的方法。

[0335] 颗粒蛋白前体是富含半胱氨酸的分泌型糖蛋白,其具有生长因子样性质,并属于颗粒体蛋白家族。颗粒体蛋白(例如,颗粒体蛋白A至G和paragranulin)在血管生成、伤口修复、细胞增殖和炎症中起作用。颗粒蛋白前体基因的突变或分泌的颗粒蛋白前体的缺乏与多种神经变性疾病和代谢疾病相关。中枢神经系统(CNS)中分泌的颗粒体蛋白的缺乏可导致神经变性,包括额颞叶退化(FTLD);额颞变性或额颞痴呆(FTD)——一种与部分颗粒蛋白前体缺乏相关的早发型神经变性疾病;进行性非流畅性失语(PNFA);词义性痴呆;帕金森病;阿尔茨海默病;和神经元蜡样质脂褐质沉积病(NCL)伴颗粒蛋白前体完全缺乏。FTD是指一组复杂的神经变性疾病,其特征在于额颞叶的进行性恶化,这导致行为和/或言语改变,如运动协调性丧失、社交障碍、记忆丧失和肌肉震颤。

[0336] 通常,在一个拷贝的颗粒蛋白前体基因中具有一个或多个致病突变的患者会患上FTD,这是FTLD的一个亚型。FTLD的临床症状包括前额叶和/或前颞叶萎缩、行为性额颞痴呆(特征为社交障碍和执行功能障碍)、词义性痴呆和进行性原发性非流畅性失语,伴有早期语言障碍。FTD是一种成年发作型行为障碍,其后是额叶痴呆、帕金森症和肌肉萎缩。

[0337] 颗粒蛋白前体也与代谢疾病有关,并且被认为是与饮食诱发的肥胖和胰岛素抵抗有关的脂肪因子。颗粒蛋白前体的突变或缺乏与动脉粥样硬化有关,动脉粥样硬化是一种进行性疾病,其特征由于脂质积累和细胞增殖而引起的强炎性组成部分和动脉壁增厚。

[0338] GRN中的突变可包括无义突变、剪接位点突变;导致正常阅读框移位的插入和缺失;以及各种点突变。无义、剪接位点和移码突变可导致单倍体不足,这是由于mRNA无义介导的转录物的衰变或核降解引起的。在一些零星的FTLD、阿尔茨海默病和肌萎缩侧索硬化(ALS)患者中观察到错义突变。与疾病或病症相关的各种GRN突变是已知的,并且包括GRN编码区的突变,例如,-8+5G>C;-8+3A>T;2T>C,3G>A,26CA(A9D);63_64insC;90_91insCTGC;102ΔC;138+1G>A;154ΔA;234_235ΔAG;243ΔC;361ΔG;373C>T;380_381ΔCT;384_387ΔTAGT;388_391ΔCAGT;463-1G>A;468_474ΔCTGCTGT;675_676ΔCA;708+1G>A;707+1G>C;709-2A>G;759_760ΔTG;813_816ΔCACT;835_835+1insCTGA;836-1G>C;909ΔC;910_911insTG;911G>A;933+1G>A;942C>A;998ΔG;1095_1096ΔCT;1145ΔC;1157G>A;1201C>T;1231_1232ΔGT;1232_1233insGT;1252C>T;1395_1296insC;1402C>T;1414-15_1590Δ;1477C>T;或其组合。参见Eriksen JL,Mackenzie IR.Paragranulin:normal function and role in neurodegeneration.J Neurochem.2008年1月;104(2):287-97。目前,尚无治愈此类GRN相关的神经变性疾病和代谢疾病的方法。需要针对GRN的治疗选择,例如,增加体内颗粒蛋白前体水平和/或功能的疗法。

[0339] 在某些实施方案中,本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可用来治疗有需要的受试者,其中该受试者具有任何一种或多种上面列出的GRN突变。在一些情况下,有需要的受试者包括GRN的单倍体不足或GRN的缺乏。在一些情况下,有需要的受试者患有FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化,或处于以上疾病的危险之中。

[0340] 在某些实施方案中,将本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表

达盒、表达载体或病毒颗粒施用于患有FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化或具有患上疾病的风险的受试者。在另一个实例中,可以向罹患FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化的受试者施用治疗。用本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗可导致与FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化相关的一种或多种症状的改善,例如由FTD引起的与行为和/或言语改变相关的一种或多种症状的减轻,或动脉粥样硬化中动脉壁厚度的减小。与GRN相关的中枢神经系统病症有关的其他症状包括,例如,Lewy小体的存在、颗粒蛋白前体 (GRN) 的单倍体不足、社交缺陷、溶酶体异常、运动协调性丧失、肌肉震颤。在某些实施方案中,该治疗可导致GRN表达水平或功能与野生型GRN相比恢复至少5%、至少10%、至少20%、至少30%、至少40%、至少50%、至少60%或至少70%。在某些实施方案中,该治疗可导致与治疗前相比,认知功能改善,如言语和/或记忆改善至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少100%或超过100%。在某些实施方案中,用本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗可延迟与FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化相关的症状或终止其症状的进展。在一些情况下,与不治疗相比,用本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒进行治疗可使CNS、额皮质或纹状体中的细胞修复增加或使其中的细胞损伤逆转至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或至少100%。在一些情况下,治疗GRN相关的中枢神经系统病症包括降低肌肉震颤的频率和/或严重程度。

[0341] 在某些实施方案中,本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可用来治疗被诊断出患有疾病,例如神经系统或神经变性疾病或与GRN相关的代谢疾病的受试者。该受试者可以是罹患FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化的患者。在一些方面,该受试者是FTD患者。该受试者可以是罹患GRN相关疾病或病况如痴呆症或代谢疾病的患者。在一些情况下,可以使用基因检测针对GRN基因突变筛查患者。如果患者在GRN中包含致病突变,则可以使用本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒来预防、减轻或治疗神经变性或代谢疾病如FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化的进展。

[0342] 在某些实施方案中,用本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒治疗的受试者是被诊断出在GRN中具有突变或遗传异常的受试者。这样的突变可以是纯合的或杂合的,或影响体内蛋白质表达、分泌、溶解性、活性和/或蛋白水解切割的GRN基因中一个或多个碱基的异常插入、缺失或置换。

[0343] 在某些实施方案中,本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可用来治疗患有与GRN突变和/或缺乏相关的神经变性疾病或代谢疾病或具有患上疾病风险的受试者,以在体内增加GRN表达和/或功能,以便治疗、预防该疾病,例如FTLD、FTD、进行性非流畅性失语、词义性痴呆、帕金森病、阿尔茨海默病、NCL、糖尿病或动脉粥样硬化,或减轻其影响。在一些情况下,本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可用来治疗、改善、减轻或控制此类疾病的症状,例如言语障碍、社交缺陷、运动技能或运动协调性受损、肌肉萎缩、肌肉震颤、神经元萎缩、记忆丧失、帕金森症和/或溶酶体异常。在一些情况下,此类组合物用来在CNS中,或

者尤其是在额皮质或纹状体中,减轻炎症,减少细胞死亡,减少神经元萎缩,减少运动神经元萎缩,和/或增加细胞修复。

[0344] 在某些实施方案中,本文所述包含编码上调内源GRN表达的eTF的多核苷酸的表达盒、表达载体或病毒颗粒可用来治疗具有患神经变性疾病或代谢疾病风险的受试者。通过各种筛查或诊断方法,包括寻找GRN中的基因突变的方法或针对所分泌的GRN蛋白水平的血液检查,可以知晓受试者易患疾病,例如神经系统疾病或代谢疾病。由于遗传事件,或由于已知的危险因素,受试者可能易患疾病。例如,受试者可以携带与FTD、阿尔茨海默病、帕金森病和/或动脉粥样硬化相关的GRN突变。在一些情况下,受试者的GRN蛋白或其同种型的量可能不足。例如,可能已知受试者的GRN蛋白的量不足。

[0345] 术语“受试者”和“个体”在本文中可互换使用,是指脊椎动物,优选哺乳动物,更优选人。本文所述的方法可用于人类治疗、兽医应用和/或在疾病或病况的动物模型中的临床前研究。在各个实施方案中,可根据本文所述的方法治疗的受试者是哺乳动物,例如小鼠、大鼠、仓鼠、豚鼠、沙鼠、牛、绵羊、猪、山羊、驴、马、狗、猫、美洲驼、猴(例如,猕猴,如恒河猴或食蟹猴)或人。在示例性实施方案中,受试者是人。

[0346] 下表提供了本文公开的序列。

[0347] 表1:SCN1A和GRN的基因组基因座和蛋白质序列

目的内源基因的染色体位置	蛋白质SEQ ID NO.	氨基酸序列 (N 至 C 末端)	
		SCN1A	GRN
染色体 2 上的碱基对 165,989,160 至 166,149,216 (智人 (Homo sapiens) 注释版本 109, GRCh38.p1	SEQ ID NO: 180	MEQTVLVPPGPDSENFNFFTRESLAAIERRIAEEKAKNPKPDKKD	MEQTVLVPPGPDSENFNFFTRESLAAIERRIAEEKAKNPKPDKKD
		DDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLEDLDPYIIN	DDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLEDLDPYIIN
		KKTFIVLNKGKAIKFRFSATSALYILTPFNPLRKIAIKILVHSLFS	KKTFIVLNKGKAIKFRFSATSALYILTPFNPLRKIAIKILVHSLFS
		MLIMCTILTNCVFMTMSNPPDWTKNVEYTFGTGIYTFESLIKIIA	MLIMCTILTNCVFMTMSNPPDWTKNVEYTFGTGIYTFESLIKIIA
		RGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRT	RGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRT
		FRVLRALKTISVIPGLKTIVGALIQSVKKLSQVMILTVFCLSVF	FRVLRALKTISVIPGLKTIVGALIQSVKKLSQVMILTVFCLSVF
		ALIGLQLFMGNLRNKCICQWPPTNASLEEHSIEKNITVNYNGTLI	ALIGLQLFMGNLRNKCICQWPPTNASLEEHSIEKNITVNYNGTLI
		NETVFEFDWKSYSIQDSRYHYFLEGFLDALLCGNSSDAGQCPE	NETVFEFDWKSYSIQDSRYHYFLEGFLDALLCGNSSDAGQCPE
		GYMCVKAGRNPNYGYTSFDTFSWAFSLFRLMTQDFWENLY	GYMCVKAGRNPNYGYTSFDTFSWAFSLFRLMTQDFWENLY
		QLTLRAAGKTYMIFFFVLVIFLGSFYLINLILAVVAMAYEEQNQ	QLTLRAAGKTYMIFFFVLVIFLGSFYLINLILAVVAMAYEEQNQ
		ATLEEAQKEAEFQQMIEQLKKQQEAAQQAATATASEHSREP	ATLEEAQKEAEFQQMIEQLKKQQEAAQQAATATASEHSREP
		SAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEK	SAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEK
		DEDEFQKSESEDSIRRKGFRRFSIEGNRLTYEKRYSSPHQSLLSIR	DEDEFQKSESEDSIRRKGFRRFSIEGNRLTYEKRYSSPHQSLLSIR
		GSLFSPRRNSRTSLFSFRGRAKDVGSENDFAADDEHSTFEDNES	GSLFSPRRNSRTSLFSFRGRAKDVGSENDFAADDEHSTFEDNES

[0349]

2)		RRDSL FVPRRHGERRNSNLSQTSRSSRMLAVFPANGKMHSTV DCNGVVSLVGPSVPTSPVGQLLPEVIIDK PATDDNGTTTETE MRKRRSSSFHVSMDFLEDPSQRQRAMSIASILTNTVEELES QKCPPCWYKFSNIFLIWDCSPYWLKVHVNLVVMDFV DL AITICIVLNTLFMAMEHYPMTHFNNVLT VGNLVFTGIFTAEM FLKIIAMDPY YFQEGWNIFDGFIVT LSLVELGLANVEGLSVL RSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLT LVLAIIVF IFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIV FRVLCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNLVV LNLFLALLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVA YVKRKIYEFIQQS FIRKQKILDEIKPLDDLNNKKDSCMSNHTA EIGKDL DY LKDVNGTTS GIGTGSSVEKYIIDESDYMSFINNPSL TVTVPIAVGESDFENLNTEDFSSES DLEESKEKLNESSSSSSEGS TVDIGAPVEEQPVVEPEETLEPEACFTEGCVQRFKCCQINVEE GRGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIY IDQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGYQTYFTNA WCWLD FLIVDVSLVSLTANALGYSELGAIKSLRTLRLRPLR ALSRFEGMRVVVNALLGAIPSIMNVLLVCLIFWLIFSIMGVNL FAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKN VKVNF DN VGFGYLSLLQVATFKGWM DIMYAAVDSRNVELQ PKYEESLYMYLYFVIFIFGSFFTLNLFIGVIIDNFNQKKKFGG QDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFGQMVFD FVTRQVFDISIMILICLNMVTMMVETDDQSEYVTTILSRINLVF IVLFTGECVLKLISLRHYFTIGWNIFDFVVVILSIVGMFLAELI EKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTL LFALMMSLPA LFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFN FETFGN SMICLFQITTSAGWDGLLAPILNSKPPDCDPNKVNP GSSSVKGD CGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEP EDDFEMFYEVWEKFDPDATQFMEFEKLSQFAAALEPPLNL PQ PNKLQLIAMDLP MVSGDRIHCLDILFAFTKRVLGESGEMDAL RIQMEERFMASNP SKVSYQPITTT LKRKQEEVSAVIIQRAYRR HLLKRTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEK TDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK
GRN 染色体 17 上的碱基对 44,345,086 至 44,353,106 (智人注释 版本 109, GRCh38.p1 2)	SEQ ID NO: 181	MWTLVSWVALTAGLVAGTRCPDGQFCPVACCLDPGGASYSC CRPLLDKWPTTL SRHLGGPCQVDAHCSAGHSCIFTVSGTSSCC PFPEAVACGDGHHCCPRGFHCSADGRSCFQ RSGNNSVGAIQC PDSQFECPDFSTCCVMVDGSGWCCPMPQASCCEDRVHCCPH GAFCDLVHTRCITPTGTHPLAKKLPAQRTNRAVALSSSVMCP DARSRCPDGSTCCELP SGKYGCCPMPNATCCSDHLHCCPQDT VCDLIQSKCLSKENATTDLLTKLPAHTVGDVKCDMEVSCPDG YTCCRLQSGAWGCCPFTQAVCCEDHIHCCPAGFTCDTQKGTC EQGPHQVPWMEKAPAHLSLPDPQALKRDVPCDNVSSCPSSDT CCQLTSGEWGCCPIPEAVCCSDHQHCCPQGYTCVAEGQCQR GSEIVAGLEKMPARRASLSHPRDIGCDQHTSCPVGQTCCPSLG GSWACCQLPHAVCCEDRQHCCPAGYTCNVKARSCEKEVVSA QPATFLARSPHVGVKDVECGEGHFCHDNQTCCRDNRQGWA CCPYRQGVCCADRRHCCPAGFRCAARGTKCLRREAPRWDA LRDPALRQLL

[0350] 表2:与人EGR1/EGR3具有高序列同一性的eTF的序列。

[0351]

	氨基酸序列 (N-至 C-末端)
	与人 EGR1 具有高序列同一性的 eTF 的序列
SEQ ID NO: 1	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRDELTRHIRIHTGQ KPFQCRICMRNFSRSDHLTTHIRTHTGKPFACDICGRKFARSDERKRHTKI HLRQKDKKADKSVVASSATSSLSSYSPVATSYSPVTTSYSPATTSYSPSP VPTSFSSPGSSTYSPVHSGFSPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFS ASTGLSDMTATFSPRTIEIC
SEQ ID NO: 2	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHELEPEGEKPYKCPECGKSFSRRDELNVHQR HTGEKPYKCPECGKSFSRRTRCRAHQRTHTGEKPYKCPECGKSFSQSSNLV RHQRTHHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFS SGNLVRHQRTHHTGEKPYKCPECGKSFSHRRTLTNHQRTHHTGKKTSSKADK SVVASSATSSLSSYSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSP SPVHSGFSPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFS PRTIEIC
SEQ ID NO: 3	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSQLAHLRAHIRIHTG QKPFQCRICMRNFSTSGNLVRHIRTHTGEKPFACDICGRKFAHRTTLTNHTK IHLRQKDKKADKSVVASSATSSLSSYSPVATSYSPVTTSYSPATTSYSPSP VPTSFSSPGSSTYSPVHSGFSPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFS ASTGLSDMTATFSPRTIEIC
SEQ ID NO: 4	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFPGSAGTALQYPPPAYPAAKGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRRDELNVHIRIHTG QKPFQCRICMRNFSSRRTRCRAHIRTHHTGEKPFACDICGRKFAQSSNLVRHTK IHLRQKDRPYACPVESCDRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFST GNLVRHIRTHTGEKPFACDICGRKFAHRTTLTNHTKIHLRQKDKKADKSVV ASSATSSLSSYSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSPSPV

[0352]

	氨基酸序列 (N-至 C-末端)
	HSGFPSVSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 5	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRDELNVHIRIHTG QKPFQCRICMRNFSSRRTCRAHIRTHTGEKPFACDICGRKFAQSSNLVRHIR THTGEKPFACDICGRKFSQLAHLRAHIRIHTGQKPFQCRICMRNFSTSGNLV RHIRTHTGEKPFACDICGRKFAHRTTLTNHTKIHRLRQKDKKADKSVVASSA TSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSPVHSGF PSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 6	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTG QKPFQCRICMRNFSREDNLHTHIRHTHTGEKPFACDICGRKFARSDELVRHT KIHRLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFST SGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHRLRQKDKKADKSV VASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSPSP VHSGFPSVSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 7	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTG QKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHT KIHRLRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQS SSLVRHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLRQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSSTYSPSPV HSGFPSVSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 8	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTG QKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHIR THTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVR

[0353]

	氨基酸序列 (N-至 C-末端)
	HIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVVASSAT SSLSSYPSPVATSYPSPVTTSYPSPATTSYSPSPVPTSFSSPGSSTYSPSPVHSGFP SPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 9	MAAAKAEMQLMSPLQISDPFGSFPHTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRDELNVHIRIHTG QKPFQCRICMRNFSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHT KIHLRQKDRPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSH RTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDRPYACPV ESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKP FACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVVASSATSSLSSYPSPV ATSYPSPVTTSYPSPATTSYSPSPVPTSFSSPGSSTYSPSPVHSGFSPSPSVATTYSS VPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 10	MAAAKAEMQLMSPLQISDPFGSFPHTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSSPADLTHIRIHTGQ KPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGRKFAQLAHLRAHTKI HLRQKDRPYACPVESCDRRFSQRANLRAHIRIHTGQKPFQCRICMRNFSRE DNLHTHIRTHTGEKPFACDICGRKFARSDDLVRHTKIHLRQKDKKADKSVV ASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYSPSPVPTSFSSPGSSTYSPV HSGFSPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTI EIC
SEQ ID NO: 325	MAAAKAEMQLMSPLQISDPFGSFPHTMDNYPKLEEMMLLSNGAPQFLG AAGAPEGSGSNSSSSSSGGGGGGGGGSNSSSSSSTFNPQADTGEQPYEHLT AESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPL FSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSNDSSPIYSAAPT FPTPNTDIFPEPQSQAFFPGSAGTALQYPPPAYPAAGGGFQVPMIPDYLFPQQ QGDLGLGTPDQKPFQGLSRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQ SQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTG QKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHIR THTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVR HIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVVASSAT SSLSSYPSPVATSYPSPVTTSYPSPATTSYSPSPVPTSFSSPGSSTYSPSPVHSGFP SPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
与人 EGR3 具有高序列同一性的 eTF 的序列	
SEQ ID NO: 11	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSA LNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK

[0354]

	氨基酸序列 (N-至 C-末端)
	YPNRPSKTPLHERPHACPAEGCDRRFSQLAHLRAHLRIHTGHKPFQCRICM RSFSTSGNLVRHIRTHTGEKPFACEFCGRKFAHRTTLTNHAKIHLKQKEKK AEKGGAPSASSAPPVSLAPVVTTC
SEQ ID NO: 12	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK YPNRPSKTPLHERPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICM RSFSSRRTCRAHIRTHTGEKPFACEFCGRKFAQSSNLVRHAKIHLKQKEHA CPAEGCDRRFSQLAHLRAHLRIHTGHKPFQCRICMRSFSTSGNLVRHIRTHT GEKPFACEFCGRKFAHRTTLTNHAKIHLKQKEKKAEEKGGAPSASSAPPVSL APVVTTC
SEQ ID NO: 13	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK YPNRPSKTPLHERPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSREDNLHHTHIRTHTGEKPFACEFCGRKFARSDELVRHAKIHLKQKEHA CPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHT GEKPFACEFCGRKFAQNSTLTEHAKIHLKQKEKKAEEKGGAPSASSAPPVSL APVVTTC
SEQ ID NO: 14	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK YPNRPSKTPLHERPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICM RSFSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKEHA CPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRICMRSFSQSSSLVRHIRTHTG EKPFAFCGRKFAREDNLHTHAKIHLKQKEKKAEEKGGAPSASSAPPVSLA PVVTTC
SEQ ID NO: 15	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK YPNRPSKTPLHERPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICM RSFSRSDHLTNHIRTHTGEKPFACEFCGRKFARSDDLVRHAKIHLKQKEHA CPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSHRTTLTNHIRTHT GEKPFACEFCGRKFAREDNLHTHAKIHLKQKEHACPAEGCDRRFSTSHSLT EHLRIHTGHKPFQCRICMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFARE DNLHTHAKIHLKQKEKKAEEKGGAPSASSAPPVSLAPVVTTC
SEQ ID NO: 16	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHYNQMAT ENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWCQDNII SLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLY SEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMPIDYNLYHHPNDMGS IPEHKPFQGMPIRVNPPITPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRK

	氨基酸序列 (N-至 C-末端)
	YPNRPSKTPLHERPHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMR SFSDSGNLRVHIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIHLKQKEHAC PAEGCDRRFSQRANLRAHLRIHTGHKPFQCRICMRFSREDNLHTHIRTHTG EKPFACEFCGRKFARSDNLVRHAKIHLKQKEKKAEEKGGAPSASSAPPVSLA PVVTTCA

[0356] 表3:表2中公开的eTF的DBD序列

eTF	DBD SEQ ID NO:	DBD 的氨基酸序列(N-至 C-末端)
与 EGR1 具有高序列同一性的 eTF 的 DBD 序列		
SEQ ID NO: 1	SEQ ID NO: 17	RPYACPVESCDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRS DHLTTHIRHTHTGEKPFACDICGRKFARSDERKRHTKIHRLQKD
SEQ ID NO: 2	SEQ ID NO: 18	LEPGEEKPYKCPECCKSFSRRDELNVHQRTHHTGEKPYKCPECCKG SFSSRRTCRAHQRTHTGEKPYKCPECCKGSFSQSSNLVRHQRTHT GEKPYKCPECCKGSFSQLAHLRAHQRTHTGEKPYKCPECCKGSFS TSGNLVRHQRTHTGEKPYKCPECCKGSFSHRTTLTNHQRTHHTGK KTS
SEQ ID NO: 3	SEQ ID NO: 19	RPYACPVESCDRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFSTS GNLVRHIRHTHTGEKPFACDICGRKFAHRTTLTNHTKIHRLQKD
SEQ ID NO: 4	SEQ ID NO: 20	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRNFSSR RTCRAHIRHTHTGEKPFACDICGRKFAQSSNLVRHTKIHRLQKDR PYACPVESCDRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFSTS GNLVRHIRHTHTGEKPFACDICGRKFAHRTTLTNHTKIHRLQKD
SEQ ID NO: 5	SEQ ID NO: 21	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRNFSSR RTCRAHIRHTHTGEKPFACDICGRKFAQSSNLVRHIRHTHTGEKPF ACDICGRKFSQLAHLRAHIRIHTGQKPFQCRICMRNFSTSGNLV RHIRHTHTGEKPFACDICGRKFAHRTTLTNHTKIHRLQKD
SEQ ID NO: 6	SEQ ID NO: 22	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRE DNLHTHIRHTHTGEKPFACDICGRKFARSDDELVRHTKIHRLQKDR PYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSG HLVRHIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHRLQKD
SEQ ID NO: 7	SEQ ID NO: 23	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSSH RTTLTNHIRHTHTGEKPFACDICGRKFARNLHTHTKIHRLQKD RPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQS SSLVRHIRHTHTGEKPFACDICGRKFARNLHTHTKIHRLQKD
SEQ ID NO: 8	SEQ ID NO: 24	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSSH RTTLTNHIRHTHTGEKPFACDICGRKFARNLHTHTHTHTHTGEKPF ACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVR HIRHTHTGEKPFACDICGRKFARNLHTHTKIHRLQKD
SEQ ID NO: 9	SEQ ID NO: 25	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRNFSRS DHLTNHIRHTHTGEKPFACDICGRKFARSDDLVRHTKIHRLQKDR PYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHR TTLTNHIRHTHTGEKPFACDICGRKFARNLHTHTKIHRLQKDR PYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSS SLVRHIRHTHTGEKPFACDICGRKFARNLHTHTKIHRLQKD

[0358]

eTF	DBD SEQ ID NO:	DBD 的氨基酸序列(N-至 C-末端)
SEQ ID NO: 10	SEQ ID NO: 26	RPYACPVESCDRRFSSPADLTRHIRIHTGQKPFQCRICMRNFSDS GNLRVHIRTHTGEKPFACDICGRKFAQLAHLRAHTKIHLRQKD RPYACPVESCDRRFSQRANLRAHIRIHTGQKPFQCRICMRNFSR EDNLHTHIRTHTGEKPFACDICGRKFARSDNLVRHTKIHLRQKD
SEQ ID NO: 325	SEQ ID NO: 351	RKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQ KPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFARED NLHTHIRTHTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQC RICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTH TKIHLRQKDKKADKSVV
与 EGR3 具有高序列同一性的 eTF 的 DBD 序列		
SEQ ID NO: 11	SEQ ID NO: 27	PHACPAEGCDRRFSQLAHLRAHLRIHTGHKPFQCRICMRSFSTS GNLVRHIRTHTGEKPFACEFCGRKFAHRTTLTNHAKIHLKQKE
SEQ ID NO: 12	SEQ ID NO: 28	HACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRSFSSRR TCRAHIRTHTGEKPFACEFCGRKFAQSSNLVRHAKIHLKQKEHA CPAEGCDRRFSQLAHLRAHLRIHTGHKPFQCRICMRSFSTSGNL VRHIRTHTGEKPFACEFCGRKFAHRTTLTNHAKIHLKQKE
SEQ ID NO: 13	SEQ ID NO: 29	PHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSSRE DNLHTHIRTHTGEKPFACEFCGRKFARSDNLVRHAKIHLKQKEH ACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRICMRSFSTSGH LVRHIRTHTGEKPFACEFCGRKFAQNSTLTEHAKIHLKQKE
SEQ ID NO: 14	SEQ ID NO: 30	PHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSSHR TTLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKEH ACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRICMRSFSSQSSSL VRHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKE
SEQ ID NO: 15	SEQ ID NO: 31	PHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRSFSSRS DHLTNHIRTHTGEKPFACEFCGRKFARSDDLVRHAKIHLKQKE HACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSSHRT TLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKEHA CPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRICMRSFSSQSSSLV RHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHLKQKE
SEQ ID NO: 16	SEQ ID NO: 32	PHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFSDS GNLRVHIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIHLKQKE HACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRICMRSFSSRED NLHTHIRTHTGEKPFACEFCGRKFARSDNLVRHAKIHLKQKE

[0359] 表4:表2中公开的eTF所识别的靶位点序列

[0360]

eTF	用于 eTF 的 蛋白质 平台	靶位点 SEQ IDNO:	靶结合位点序列(5'-至-3')	基因靶标	与蛋白质平 台的 NCBI- BLAST 序列 同一性
SEQ ID NO: 1	WT EGR1	N/A	N/A	N/A	100%
SEQ ID NO: 2	EGR1	SEQ ID NO: 33	AGTGATAGAGAACGT ATG	报道基因	87%

[0361]

eTF	用于 eTF 的 蛋白质 平台	靶位点 SEQ IDNO:	靶结合位点序列(5'-至-3')	基因靶标	与蛋白质平 台的 NCBI- BLAST 序列 同一性
SEQ ID NO: 3	EGR1	SEQ ID NO: 34	AGTGATAGA	报道基因	97%
SEQ ID NO: 4	EGR1	SEQ ID NO: 33	AGTGATAGAGAACGT ATG	报道基因	96%
SEQ ID NO: 5	EGR1	SEQ ID NO: 33	AGTGATAGAGAACGT ATG	报道基因	94%
SEQ ID NO: 6	EGR1	SEQ ID NO: 35	CTAGGTCAAGTGTAGG AG	SCN1A	97%
SEQ ID NO: 7	EGR1	SEQ ID NO: 36	TAGGTACCATAGAGTG AG	SCN1A	96%
SEQ ID NO: 8	EGR1	SEQ ID NO: 36	TAGGTACCATAGAGTG AG	SCN1A	94%
SEQ ID NO: 9	EGR1	SEQ ID NO: 37	TAGGTACCATAGAGTG AGGCGAGGATG	SCN1A	97%
SEQ ID NO: 10	EGR1	SEQ ID NO: 38	GAGTAGAAAAGAAAC ACA	GRN	95%
SEQ ID NO: 11	EGR3	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因	94%
SEQ ID NO: 12	EGR3	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因	96%
SEQ ID NO: 13	EGR3	SEQ ID NO: 35	CTAGGTCAAGTGTAGG AG	SCN1A	97%
SEQ ID NO: 14	EGR3	SEQ ID NO: 36	TAGGTACCATAGAGTG AG	SCN1A	96%
SEQ ID NO: 15	EGR3	SEQ ID NO: 37	TAGGTACCATAGAGTG AGGCGAGGATG	SCN1A	97%
SEQ ID NO: 16	EGR3	SEQ ID NO: 38	GAGTAGAAAAGAAAC ACA	GRN	95%

[0362] 表5: 用于生成eTF的锌指结构和蛋白质的不同类型

[0363]

ZF 类型名称		ZF 结构 (其中每个 x 可以独立地 为任何残基)	可以充当 eTF 或本文公开的 eTF 的 DNA 结合域的蛋白质 平台的示例性蛋白质
锌指 C2H2-型 (ZNF)	SEQ ID NO: 39	C-x-C-x-H-x-H	KLF4, KLF5, EGR3, ZFP637, SLUG, ZNF750, ZNF281, ZBP89, GLIS1, GLIS3

[0364]

ZF 类型名称		ZF 结构 (其中每个 x 可以独立地为任何残基)	可以充当 eTF 或本文公开的 eTF 的 DNA 结合域的蛋白质平台的示例性蛋白质
环指蛋白(RNF)	SEQ ID NO: 40	C-x-C-x-C-x-H-xxx-C-x-C-x-C-x-C	MDM2, BRCA1, ZNF179
PHD 指蛋白 (PHF)	SEQ ID NO: 41	C-x-C-x-C-x-C-xxx-H-x-C-x-C-x-C	KDM2A, PHF1, ING1
含 LIM 结构域	SEQ ID NO: 42	C-x-C-x-H-x-C-x-C-x-C-x-C-x-(C,H,D)	ZNF185, LIMK1, PXN
核激素受体 (NR)	SEQ ID NO: 43	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	VDR, ESR1, NR4A1
锌指 CCCH-型 (ZC3H)	SEQ ID NO: 44	C-x-C-x-C-x-H	RC3H1, HELZ, MBNL1, ZFP36, ZFP36L1
锌指 FYVE-型 (ZFYVE)	SEQ ID NO: 43	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	EEA1, HGS, PIKFYVE
锌指 CCHC-型 (ZCCHC)	SEQ ID NO: 45	C-x-C-x-H-x-C	CNBP, SF1, LIN28A
锌指 DHHC-型 (ZDHHHC)	SEQ ID NO: 46	C-x-C-x-H-x-C-xxx-C-x-C-x-H-x-C	ZDHHHC2, ZDHHHC8, ZDHHHC9
锌指 MYND-型 (ZMYND)	SEQ ID NO: 47	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	PDCD2, RUNX1T1, SMYD2, SMYD1
锌指 RANBP2-型 (ZRANBP2)	SEQ ID NO: 48	C-x-C-x-C-x-C	YAF2, SHARPIN, EWSR1
锌指 ZZ-型 (ZZZ)	SEQ ID NO: 48	C-x-C-x-C-x-C	HERC2, NBR1, CREBBP
锌指 C2HC-型 (ZC2HC)	SEQ ID NO: 45	C-x-C-x-H-x-C	IKBKG, L3MBTL1, ZNF746
含 GATA 锌指结构域 (GATAD)	SEQ ID NO: 48	C-x-C-x-C-x-C	GATA4, GATA6, MTA1
ZF 类别同源框和假基因	SEQ ID NO: 39	C-x-C-x-H-x-H	ADNP, ZEB1, ZHX1
含 THAP 结构域 (THAP)	SEQ ID NO: 44	C-x-C-x-C-x-H	THAP1, THAP4, THAP11
锌指 CXXC-型 (CXXC)	SEQ ID NO: 43	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	CXXC1, CXXC5, MBD1, DNMT1
锌指 SWIM-型 (ZSWIM)	SEQ ID NO: 44	C-x-C-x-C-x-H	MAP3K1, ZSWIM5, ZSWIM6
锌指 AN1-型 (ZFAND)	SEQ ID NO: 49	C-x-C-x-C-x-C-xxx-C-x-H-x-H-x-C	ZFAND3, ZFAND6, IGHMBP2
锌指 3CxxC-型 (Z3CXXC)	SEQ ID NO: 45	C-x-C-x-H-x-C	ZAR1, RTP1, RTP4
锌指 CW-型 (ZCW)	SEQ ID NO: 48	C-x-C-x-C-x-C	MORC1, ZCWPW1, KDM1B
锌指 GRF-型 (ZGRF)	SEQ ID NO: 48	C-x-C-x-C-x-C	TTF2, NEIL3, TOP3A

[0365]

ZF 类型名称		ZF 结构 (其中每个 x 可以独立地为任何残基)	可以充当 eTF 或本文公开的 eTF 的 DNA 结合域的蛋白质平台的示例性蛋白质
锌指 MIZ-型 (ZMIZ)	SEQ ID NO: 45	C-x-C-x-H-x-C	PIAS1, PIAS3, PIAS4
锌指 BED-型 (ZBED)	SEQ ID NO: 39	C-x-C-x-H-x-H	ZBED1, ZBED4, ZBED6
锌指 HIT-型 (ZNHIT)	SEQ ID NO: 47	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	ZNHIT3, DDX59, INO80B
锌指 MYM-型 (ZMYM)	SEQ ID NO: 48	C-x-C-x-C-x-C	ZMYM2, ZMYM3, ZMYM4
锌指基质蛋白-型 (ZMAT)	SEQ ID NO: 39	C-x-C-x-H-x-H	ZNF638, ZMAT1, ZMAT3, ZMAT5
锌指 C2H2C-型	SEQ ID NO: 39	C-x-C-x-H-x-H	MYT1, MYT1L, ST18
锌指 DBF-型 (ZDBF)	SEQ ID NO: 39	C-x-C-x-H-x-H	DBF4, DBF4B, ZDBF2
锌指 PARP-型	SEQ ID NO: 45	C-x-C-x-H-x-C	LIG3, PARP1

[0366] 表6:eTF的全长氨基酸序列的实例

[0367]

	氨基酸序列(N-至 C-末端)
SEQ ID NO: 50	MAPKKKRKVGIIHGVPAALEPGKPYKCPECGKSFSRRDELNVHQRTHHTGE KPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLVRHQR THTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTSGNL VRHQRTHHTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGKKT SKRPAATKKA GQAKKKKGSYPYDVPDYA
SEQ ID NO: 51	MAPKKKRKVGIIHGVPAALEPGKPYKCPECGKSFSRRDELNVHQRTHHTGE KPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLVRHQR THTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTSGNL VRHQRTHHTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGKKT SKRPAATKKA GQAKKKKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLGSDAL DDFDLDMLGSDALDDFDLDM
SEQ ID NO: 52	MAPKKKRKVGIIHGVPAALEPGKPYKCPECGKSFSRRDELNVHQRTHHTGE KPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLVRHQR THTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTSGNL VRHQRTHHTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGKKT SKRPAATKKA GQAKKKKGSYPYDVPDYALEMSGLEMA DHMMAMNHGRFPDGTNGLHH HPAHRMGMGQFPSPHHHQQQPQHAFNALMGEHIHYGAGNMNATSGIRH AMGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLN NQYFNHHPYPHNHYMPDLHPAAGHQMN GTNQHFRDCNPKHSGGSSTPGG SGGSSTPGGSGSSSGGAGSSNSGGSGSGNMPASVAHVPAAMLPPNVIDT DFIDEEVLMSLVIEMGLDRIKELPELWLGQNEFD FMTDFVCKQQPSRVSC
SEQ ID NO: 53	MAPKKKRKVGIIHGVPAALEPGKPYKCPECGKSFSRRDELNVHQRTHHTGE KPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLVRHQR THTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTSGNL VRHQRTHHTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGKKT SKRPAATKKA

[0368]

	氨基酸序列(N-至 C-末端)
	GQAKKKKGSYPYDVPDYALEMADHMLAEGYRLVQRPPSAAAAHGH PHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVP PPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGG MDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS C
SEQ ID NO: 54	MAADHMLAEGYRLVQRPPSAAAAHGH PHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRA AAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFSRRDELNVHQRTHTGEKPYKCPECGKS FSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLVRHQRTHTGEKPYKC PECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTGNLVRHQRTHT GEKPYKCPECGKSFHRTTLTNHQRTHTGKKT SKRPAATKKAGQAKKKKGSYPYDVPDYALEMADHMLAEGYRLVQRPPS AAAAHGH PHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVA TPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAELIDEEALTS LELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 55	MQSQLIKPSRMRYPNRPSKTPHERPYACPVESCDRRFSRRDELNVH IRIHTGQKPFQCRICMRNFSSRRTCRAHIRTHTGEKPFACDICGRKFAQSSNLVRH TKIHLRQKDRPYACPVESCDRRFSQLAHLRAHIRTHTGQKPFQCRICMRNF STGNLVRHIRTHTGEKPFACDICGRKFAHRTTLTNHTKIHLRQKDKLEMSG LEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGMGQFSPHHHQQQQP QHAFNALMGEHIIHYGAGNMNATSGIRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHPAAG HQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSGSSGGGAGSSNSG GGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKEL ELWLGQNEFDFMTDFVCKQQPSRVSC
SEQ ID NO: 56	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGMGQFSPHHHQQQ QPQHAFNALMGEHIIHYGAGNMNATSGVRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP AAGHQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSGSSGGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQLIKPSRMRYPNRPSKTP HERPYACPVESCDRRFSRRDELNVHIRTHTGQKPFQCRICMRNFSSRRTCRA HIRTHTGEKPFACDICGRKFAQSSNLVRHTKIHLRQKDRPYACPVESCDRRF SQLAHLRAHIRTHTGQKPFQCRICMRNFSTGNLVRHIRTHTGEKPFACDIC GRKFAHRTTLTNHTKIHLRQKDK
SEQ ID NO: 57	MQSQLIKPSRMRYPNRPSKTPHERPYACPVESCDRRFSRSDNLVRH IRIHTGQKPFQCRICMRNFSDNLHIRTHTHTGQKPFACDICGRKFARSD ELVHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNF STSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDKLEMS GLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGMGQFSPHHHQQQQP QHAFNALMGEHIIHYGAGNMNATSGIRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHPA AGHQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSGSSGGGAGSSNS GGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKEL PELWLGQNEFDFMTDFVCKQQPSRVSC
SEQ ID NO: 58	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGMGQFSPHHHQQQ QPQHAFNALMGEHIIHYGAGNMNATSGVRHAMGPGTVNGGHPPSALAPAA

[0369]

	氨基酸序列(N-至 C-末端)
	RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP AAGHQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSSSSGGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQLIKPSRMRKYPNRPSKTPP HERPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRICMRNFSREDNLHT HIRTHTGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDRPYACPVESCDRRF SQSGNLTEHIRHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICG RKFAQNSTLTEHTKIHRLRQKDK
SEQ ID NO: 59	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSSPADLTRHIRH TGQKPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGRKFAQLAHLRA HTKIHRLRQKDRPYACPVESCDRRFSQRANLRAHIRHTGQKPFQCRICMRNF SREDNLHTHIRTHTGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDKLEMS GLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQQP QHAFNALMGEHIHYGAGNMNATSGXRHAMGPGTVNGGHPPSALAPAARF NNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHPAA GHQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSSSSGGGAGSSNS GGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIKEL PELWLGQNEFDFMTDFVCKQQPSRVSC (其中 X 可以是 I 或 V)
SEQ ID NO: 60	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQ QPQHAFNALMGEHIHYGAGNMNATSGVRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP AAGHQMNGTNQHFRDCNPKHSGGSSTPGSGGSSTPGSGSSSSGGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQLIKPSRMRKYPNRPSKTPP HERPYACPVESCDRRFSRRDELNVHIRHTGQKPFQCRICMRNFSSRRTCRA HIRTHTGEKPFACDICGRKFAQSSNLVRHTKIHRLRQKDRPYACPVESCDRRF SQLAHLRAHIRHTGQKPFQCRICMRNFSTSGNLVRHIRTHTGEKPFACDIC GRKFAHRTTLTNHTKIHRLRQKDK
SEQ ID NO: 61	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRH TGQKPFQCRICMRNFREDNLHTHIRTHTGEKPFACDICGRKFARSDNLVR HTKIHRLRQKDRPYACPVESCDRRFSQSGNLTEHIRHTGQKPFQCRICMRNF STSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHRLRQKDKLEMA DHLMLAEGYRLVQRPPSAAAAGHGPALRTLPPYAGPGLDSGLRPRGAPLG PPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPP NAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVR ELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 62	MAADHMLAEGYRLVQRPPSAAAAGHGPALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRK YPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRICMR NFSREDNLHTHIRTHTGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDRPY ACPVESCDRRFSQSGNLTEHIRHTGQKPFQCRICMRNFSTSGHLVRHIRT TGEKPFACDICGRKFAQNSTLTEHTKIHRLRQKDKLEMADHMLAEGYRLV QRPPSAAAAGHGPALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYG AFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAP

[0370]

	氨基酸序列(N-至 C-末端)
	SAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSC
SEQ ID NO: 63	MQSQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSSPADLTRHIRH TGQKPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGRKFAQLAHLRA HTKIHRLRQKDRPYACPVESCDRRFSQRANLRAHIRIHTGQKPFQCRICMRN SREDNLHTHIRTHTGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDKLEMA DHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLG PPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPP NAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVR ELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 64	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRY PNRPSKTPPHERPYACPVESCDRRFSSPADLTRHIRIHTGQKPFQCRICMR NFSDSGNLRVHIRTHTGEKPFACDICGRKFAQLAHLRAHTKIHRLRQKDRPY ACPVESCDRRFSQRANLRAHIRIHTGQKPFQCRICMRNFSREDNLHTHIRT TGEKPFACDICGRKFARSDNLVRHTKIHRLRQKDKLEMADHLMLAEGYRLV QRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYG AFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPPNAPGGPPGPQPAP SAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSC
SEQ ID NO: 65	MQSQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSRRDELNVHIRIH TGQKPFQCRICMRNFSRRRTCRAHIRTHTGEKPFACDICGRKFAQSSNLVRH TKIHRLRQKDRPYACPVESCDRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFS TSGNLVRHIRTHTGEKPFACDICGRKFAHRTTLTNHTKIHRLRQKDKLEMAD HLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLG PPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPPN APGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVREL PELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 66	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRY PNRPSKTPPHERPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMR NFSRRRTCRAHIRTHTGEKPFACDICGRKFAQSSNLVRHTKIHRLRQKDRPYA CPVESCDRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFSSTSGNLVRHIRTHT GEKPFACDICGRKFAHRTTLTNHTKIHRLRQKDKLEMADHLMLAEGYRLVQ RPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGA FGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPPNAPGGPPGPQPAPS AAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSC
SEQ ID NO: 67	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRY PNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLRQKDRPY

[0371]

	氨基酸序列(N-至 C-末端)
	ACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHT GEKPFACDICGRKFAREDNLHTHTKIHLRQKDKLEMADHMLAEGYRLVQ RPPSAAAAHGPALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGA FGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPPNAPGGPPGPQPAPS AAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCF SDLGSAPPAGSVSC
SEQ ID NO: 68	MAADHMLAEGYRLVQRPPSAAAAHGPALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSOLIKPSRMRK YPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDRPY ACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHT GEKPFACDICGRKFAREDNLHTHTKIHLRQKDK
SEQ ID NO: 69	MAADHMLAEGYRLVQRPPSAAAAHGPALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGGSGQSOLIKPSRMRK YPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSREDNLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLRQKDRPY ACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSSTSGHLVRHIRTHT TGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDK
SEQ ID NO: 70	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSREDNLHT HIRTHTGEKPFACEFCGRKFARSDELVRHAKIHLKQKEHACPAEGCDRRFS QSGNLTEHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFACEFCG RKFAQNSTLTEHAKIHLKQKEKKAEGGAPSASSAPPVSLAPVVTTCALAM SGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQ PQHAFNALMGEHIIHYGAGNMNATSGIRHAMGPGTVNGGHPPSALAPAA FNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHPA AGHQMNNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGAGSSN SGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKE LPELWLGQNEFDFMTDFVCKQQPSRVSC
SEQ ID NO: 71	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQ QPQHAFNALMGEHIIHYGAGNMNATSGVRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP AAGHQMNNGTNQHFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSCRPHACPAEGCDRRFSRSDNLV RHLRIHTGHKPFQCRICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARS DELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRIC MRSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTEHAKIHLKQKEK KAEGGAPSASSAPPVSLAPVVTCA
SEQ ID NO: 72	MRPHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFSDSGNLRV HIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIHLKQKEHACPAEGCDRRFS QRANLRAHLRIHTGHKPFQCRICMRSFSREDNLHTHIRTHTGEKPFACEFCG RKFARSDELVRHAKIHLKQKEKKAEGGAPSASSAPPVSLAPVVTTCALAM MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQ QPQHAFNALMGEHIIHYGAGNMNATSGXRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP

[0372]

	氨基酸序列(N-至 C-末端)
	AAGHQMNQGTNQHFRDCNPKHSGSGSSTPGGSGGSSTPGGSGSSSGGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSC (其中 X 可以是 I 或 V)
SEQ ID NO: 73	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPHHHQQQ QPQHAFNALMGEHIHYGAGNMNATSGVRHAMGPGTVNGGHPPSALAPAA RFNNSQFMGPPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDLHP AAGHQMNQGTNQHFRDCNPKHSGSGSSTPGGSGGSSTPGGSGSSSGGGAGSS NSGGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIK ELPELWLGQNEFDFMTDFVCKQQPSRVSCRPHACPAEGCDRRFSSPADLTR HLRIHTGHKPFQCRICMRSFSDSGNLRVHIRTHTGEKPFACEFCGRKFAQLA HLRAHAKIHLKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRIC MRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSDNLVRHAKIHLKQKEK KAEKGGAPSASSAPPVSLAPVVTTC
SEQ ID NO: 74	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSREDNLHT HIRTHTGEKPFACEFCGRKFARSDNLVRHAKIHLKQKEHACPAEGCDRRFS QSGNLTEHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFACEFCG RKFAQNSTLTEHAKIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAHAHG HALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFPFPA VPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHAL GGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGS VSC
SEQ ID NO: 75	MAADHMLAEGYRLVQRPPSAAAHAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFPFPAVPPPAAGIAHLQPVATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGSGRPHACPAEGCD RRFSRSDNLVRHLRIHTGHKPFQCRICMRSFSREDNLHTHIRTHTGEKPFAC EFCGRKFARSDNLVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLRIHT GHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTEHA KIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAHAHGPHALRTLPPYAGPG LDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFPFPAVPPPAAGIAHLQ VATPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEAL TSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 76	MRPHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFSDSGNLRV HIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIHLKQKEHACPAEGCDRRFS QRANLRAHLRIHTGHKPFQCRICMRSFSREDNLHTHIRTHTGEKPFACEFCG RKFAARSDNLVRHAKIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAHAHG PHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFPFPA AVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHA LGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSC
SEQ ID NO: 77	MAADHMLAEGYRLVQRPPSAAAHAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFPFPAVPPPAAGIAHLQPVATPYPGRA AAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGSGRPHACPAEGCD RRFSSPADLTRHLRIHTGHKPFQCRICMRSFSDSGNLRVHIRTHTGEKPFAC EFCGRKFAQLAHLRAHAKIHLKQKEHACPAEGCDRRFSQRANLRAHLRIH

[0373]

	氨基酸序列(N-至 C-末端)
	TGHKPFQCRICMRSFSREDNLHIRTHHTGKPFACEFCGRKFARSDNLVR HAKIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYA GGLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAH LQPVATPYPGRAAAPNAPGGPPGPQAPSAAPPPPAHALGGMDAELIDE EALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 78	MRPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRSFSSRRTCRA HIRTHHTGKPFACEFCGRKFAQSSNLVRHAKIHLKQKEHACPAEGCDRRFS QLAHLRAHLRIHTGHKPFQCRICMRSFSTSGNLVRHIRTHHTGKPFACEFCG RKFAHRTTLTNHAKIHLKQKEKLEMSGLEMAHMMAMNHGRFPDGTNG LHHHPAHRMGMGQFSPHHHQQQQPQHAFNALMGEHIHYGAGNMNATS GIRHAMGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQ KLNNQYFNHHPYPHNHYMPDLHPAAGHQMNQTNQHFRCNPKHSGGSST PGSGGSSTPGSGSSSGGAGSSNSGGSGSGNMPASVAHVPAAMLPPN VIDTDFIDEEVLMMLVIEMGLDRIKELPELWLGQNEFDFMTDFVCKQQPSR VSC
SEQ ID NO: 79	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPLDGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRA AAPNAPGGPPGPQAPSAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGSGGSGRPHACPAEGCD RRFSRRDELNVHLRIHTGHKPFQCRICMRSFSSRRTCRAHIRTHHTGKPFAC EFCGRKFAQSSNLVRHAKIHLKQKEHACPAEGCDRRFSQLAHLRAHLRIHT GHKPFQCRICMRSFSTSGNLVRHIRTHHTGKPFACEFCGRKFAHRTTLTNH AKIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGP GLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPNAPGGPPGPQAPSAAPPPPAHALGGMDAELIDEEA LTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 80	MRPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRSFSSRRTCRA HIRTHHTGKPFACEFCGRKFAQSSNLVRHAKIHLKQKEHACPAEGCDRRFS QLAHLRAHLRIHTGHKPFQCRICMRSFSTSGNLVRHIRTHHTGKPFACEFCG RKFAHRTTLTNHAKIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAAHG PHALRTLPPYAGPLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPF AVPPPAAGIAHLQPVATPYPGRAAAPNAPGGPPGPQAPSAAPPPPAHA LGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAG SVSC
SEQ ID NO: 268	MAPKKRKRKVGIVGPAALEPGEKPYKCPECCKGSFSDPGHLVRHQRTHTGE KPYKCPECCKGSFSTKNSLTEHQRTHTGEKPYKCPECCKGSFSRNDALTEHQ THTGEKPYKCPECCKGSFSERSHLREHQRTHTGEKPYKCPECCKGSFSKHL AEHQRTHTGEKPYKCPECCKGSFSTHLDIRHQRTHTGEKPYKCPECCKGSFS QSGNLTEHQRTHTGKKTSKRPAATKKAG
SEQ ID NO: 269	MAPKKRKRKVGIVGPAALEPGEKPYKCPECCKGSFSSRRTCRAHQRTHTGE KPYKCPECCKGSFSTTGALTEHQRTHTGEKPYKCPECCKGSFSRSDALVRHQ THTGEKPYKCPECCKGSFSRNDALTEHQRTHTGEKPYKCPECCKGSFSQSGDL RRHQRTHTGEKPYKCPECCKGSFSTSHSLTEHQRTHTGKKTSKRPAATKKA G
SEQ ID NO: 270	MAPKKRKRKVGIVGPAALEPGEKPYKCPECCKGSFSRKNLKNHQRTHTGE KPYKCPECCKGSFSDPGALVRHQRTHTGEKPYKCPECCKGSFSREDNLHQR THTGEKPYKCPECCKGSFSDPGALVRHQRTHTGEKPYKCPECCKGSFSTSGEL VRHQRTHTGEKPYKCPECCKGSFSRKNLKNHQRTHTGKKTSKRPAATKKA G

[0374]

	氨基酸序列(N-至 C-末端)
SEQ ID NO: 271	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGE KPYKCPECGKSFSREDNLHHTHQRTHTEKPYKCPECGKSFSRSDNLVRHQ RTHTEKPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKSFSSTSGHL VRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTSKRPAATKKA GQAKKKKGSYPYDVPDYALEDALDDFDLMLGSDALDDFDLMLGSDAL DDFDLMLGSDALDDFDLML
SEQ ID NO: 272	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSKKALTEHQRTHTEK KPYKCPECGKSFSPPADLTHQRTHTEKPYKCPECGKSFSRSDNLVRHQ RTHTEKPYKCPECGKSFSREDNLHHTHQRTHTEKPYKCPECGKSFSRSD NLVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKS FSSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKT SKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDDFDLMLGSDALDD FDLMLGSDALDDFDLML
SEQ ID NO: 273	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSPPADLTHQRTHTEK KPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHHT HQRTHTEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFS QSGNLTEHQRTHTEKPYKCPECGKSFSSTSGHLVRHQRTHTGKKT SKRPAATKKA G
SEQ ID NO: 274	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSKKALTEHQRTHTEK KPYKCPECGKSFSPPADLTHQRTHTEKPYKCPECGKSFSRSDNLVRHQ RTHTEKPYKCPECGKSFSREDNLHHTHQRTHTEKPYKCPECGKSFSRSD NLVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGKKT SKRPAATKKA G
SEQ ID NO: 275	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSDCRDLARHQRTHTGE KPYKCPECGKSFSRNDALTEHQRTHTEKPYKCPECGKSFSRNDALTEHQ RTHTEKPYKCPECGKSFSPPADLTHQRTHTEKPYKCPECGKSFSDPGNL VRHQRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKS FSQSSSLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGKKT SKRPAATKKAG
SEQ ID NO: 276	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSRNDALTEHQRTHTEK KPYKCPECGKSFSPPADLTHQRTHTEKPYKCPECGKSFSDPGNLVRHQ RTHTEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSQSS SLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGKKT SKRPAATKKA G
SEQ ID NO: 277	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSSTKNSLTEHQRTHTEK KPYKCPECGKSFSRADNLTEHQRTHTEKPYKCPECGKSFSQLAHLRAHQ RTHTEKPYKCPECGKSFSSTKNSLTEHQRTHTEKPYKCPECGKSFSQAGH LASHQRTHTEKPYKCPECGKSFSSTHLDIRHQRTHTGKKT SKRPAATKKA G
SEQ ID NO: 278	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSRNDALTEHQRTHTEK KPYKCPECGKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSSTSGELVRHQ RTHTEKPYKCPECGKSFSSTHLDIRHQRTHTGEKPYKCPECGKSFSKKAL TEHQRTHTEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKS FSRSDHLTNHQRTHTGKKT SKRPAATKKAG
SEQ ID NO: 279	MAPKKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGE KPYKCPECGKSFSHRTTLTNHQRTHTEKPYKCPECGKSFSREDNLHHT HQRTHTEKPYKCPECGKSFSSTSHSLTEHQRTHTEKPYKCPECGKSFS QSSSLVRHQRTHTGEKPYKCPECGKSFSREDNLHHTHQRTHTGKKT SKRPAATKKAG

	氨基酸序列(N-至 C-末端)
	QAKKKKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDM
SEQ ID NO: 280	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSDPGALVRHQRTHTGE KPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQR THTGEKPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCPECGKSFSSTSGNLV RHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTSKRPAATKKAG QAKKKKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLGSDALD DFDLDMLGSDALDDFDLDM
SEQ ID NO: 281	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSQAGHLASHQRTHTGE KPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSSTSGNLTEHQR THTGEKPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCPECGKSFSQKSSLI AHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTSKRPAATKKAG
SEQ ID NO: 282	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSSTGNLTVHQRTHTGE KPYKCPECGKSFSSTSGELVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQR THTGEKPYKCPECGKSFSSTSGNLTEHQRHTHTGEKPYKCPECGKSFSQSSSLV RHQRTHTGEKPYKCPECGKSFSORANLRAHQRTHTGKKTSKRPAATKKAG

[0376] 表7:表6中公开的eTF的DBD序列

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列	用于 DBD 的蛋白 平台	DBD 与 蛋白质平 台之间的 NCBI- BLAST 序列同一 性
SEQ ID NO: 50	SEQ ID NO: 81	LEPG EKPYKCPEC GKSFSRRDELNVHQRTHTG EKPYKCPEC GKSFSRRRTCRAHQRTHTGEKPY KCPEC GKSFSQSSNLVRHQRTHTGEKPYKCPE CGKSFSQLAHLRAHQRTHTGEKPYKCPEC GKS FSTSGNLVRHQRTHTGEKPYKCPEC GKSFSHR TTLTNHQRTHTGKKTS	合成的	67%
SEQ ID NO: 51	SEQ ID NO: 81	(同上)	合成的	67%
SEQ ID NO: 52	SEQ ID NO: 81	(同上)	合成的	67%
SEQ ID NO: 53	SEQ ID NO: 81	(同上)	合成的	67%
SEQ ID NO: 54	SEQ ID NO: 81	(同上)	合成的	67%

[0378]

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列	用于 DBD 的蛋白 平台	DBD 与 蛋白质平 台之间的 NCBI- BLAST 序列同一 性
SEQ ID NO: 55	SEQ ID NO: 82	MQSQLIKPSRMRKYPNRPSTPPHERPYACPV ESCDRRFSRRDELNVHIRIHTGQKPFQCRICMR NFSSRRTCRAHIRTHTGEKPFACDICGRKFAQS SNLVRHTKIHRLQKDRPYACPVESCDRRFSQL AHLRAHIRIHTGQKPFQCRICMRNFSTSGNLVR HIRTHTGEKPFACDICGRKFAHRTTLTNHTKIH LRQKDK	EGR1	86%
SEQ ID NO: 56	SEQ ID NO: 83	QSQLIKPSRMRKYPNRPSTPPHERPYACPVES CDRRFSRRDELNVHIRIHTGQKPFQCRICMRNF SSRRTCRAHIRTHTGEKPFACDICGRKFAQSSN LVRHTKIHRLQKDRPYACPVESCDRRFSQLAH LRAHIRIHTGQKPFQCRICMRNFSTSGNLVRHI RTHTGEKPFACDICGRKFAHRTTLTNHTKIHRL QKDK	EGR1	86%
SEQ ID NO: 57	SEQ ID NO: 84	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPF QCRICMRNFSREDNLHTHIRTHTGEKPFACDIC GRKFARSDLVHTKIHRLQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFS TSGHLVRHIRTHTGEKPFACDICGRKFAQNSTL TEHTKIHRLQKD	EGR1	93%
SEQ ID NO: 58	SEQ ID NO: 85	QSQLIKPSRMRKYPNRPSTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SREDNLHTHIRTHTGEKPFACDICGRKFARSDE LVRHTKIHRLQKDRPYACPVESCDRRFSQSGN LTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHI RTHTGEKPFACDICGRKFAQNSTLTEHTKIHRL QKDK	EGR1	93%
SEQ ID NO: 59	SEQ ID NO: 86	MQSQLIKPSRMRKYPNRPSTPPHERPYACPV ESCDRRFSPPADLTHIRIHTGQKPFQCRICMR NFSDSGNLRVHIRTHTGEKPFACDICGRKFAQL AHLRAHTKIHRLQKDRPYACPVESCDRRFSQR ANLRAHIRIHTGQKPFQCRICMRNFSREDNLHT HIRTHTGEKPFACDICGRKFARSDNLVRHTKIH LRQKDK	EGR1	84%
SEQ ID NO: 60	SEQ ID NO: 87	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPF QCRICMRNFSSRRTCRAHIRTHTGEKPFACDIC GRKFAQSSNLVRHTKIHRLQKDRPYACPVESC DRRFSQLAHLRAHIRIHTGQKPFQCRICMRNFS TSGNLVRHIRTHTGEKPFACDICGRKFAHRTTL TNHTKIHRLQKDK*	EGR1	84%
SEQ ID	SEQ ID	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPF QCRICMRNFSREDNLHTHIRTHTGEKPFACDIC	EGR1	93%

[0379]

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列	用于 DBD 的蛋白 平台	DBD 与 蛋白质平 台之间的 NCBI- BLAST 序列同一 性
NO: 61	NO: 88	GRKFARSDLV RHTKIHLRQKDRPYACPVESC DRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFS TSGHLVRHIRTHTGEKPFACDICGRKFAQNSTL TEHTKIHLRQKDK		
SEQ ID NO: 62	SEQ ID NO: 88	(同上)	EGR1	93%
SEQ ID NO: 63	SEQ ID NO: 89	RPYACPVESC DRRFSSPADLTRHIRIHTGQKPF QCRICMRNFSDSGNLRVHIRTHTGEKPFACDIC GRKFAQLAHLRAHTKIHLRQKDRPYACPVESC DRRFSQRANLRAHIRIHTGQKPFQCRICMRNFS REDNLHTHIRTHTGEKPFACDICGRKFARSDN LVRHTKIHLRQKDK	EGR1	84%
SEQ ID NO: 64	SEQ ID NO: 89	(同上)	EGR1	84%
SEQ ID NO: 65	SEQ ID NO: 82	(同上)	EGR1	86%
SEQ ID NO: 66	SEQ ID NO: 82	(同上)	EGR1	86%
SEQ ID NO: 67	SEQ ID NO: 90	QSQLIKPSRMRYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SHRTTLTNHIRTHTGEKPFACDICGRKFARNL LHTHTKIHLRQKDRPYACPVESC DRRFSTSHSL TEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRT HTGEKPFACDICGRKFARNLHTHTKIHLRQ KDK	EGR1	79%
SEQ ID NO: 68	SEQ ID NO: 90	(同上)	EGR1	79%
SEQ ID NO: 69	SEQ ID NO: 91	QSQLIKPSRMRYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SREDNLHTHIRTHTGEKPFACDICGRKFARSDE LVRHTKIHLRQKDRPYACPVESC DRRFSQSGN LTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHI	EGR1	81%

[0380]

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列	用于 DBD 的蛋白 平台	DBD 与 蛋白质平 台之间的 NCBI- BLAST 序列同一 性
		RTHTGEKPFACDICGRKFAQNSTLTEHTKIHRL QKDK		
SEQ ID NO: 70	SEQ ID NO: 92	HACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQ CRICMRSFSREDNLHTHIRHTHTGEKPFACEFCG RKFAKDELVRHAKIHLKQKEHACPAEGCDR RFSQSGNLTEHLRIHTGHKPFQCRICMRSFSTS GHLVRHIRTHTGEKPFACEFCGRKFAQNSTLT EHAKIHLKQKE	EGR3	92%
SEQ ID NO: 71	SEQ ID NO: 92	(同上)	EGR3	92%
SEQ ID NO: 72	SEQ ID NO: 93	HACPAEGCDRRFSSPADLTRHLRIHTGHKPFQ CRICMRSFSDSGNLRVHIRTHTGEKPFACEFCG RKFAQLAHLRAHAKIHLKQKEHACPAEGCDR RFSQRANLRAHLRIHTGHKPFQCRICMRSFSRE DNLHTHIRTHTGEKPFACEFCGRKFARSDNLV RHAKIHLKQKE	EGR3	80%
SEQ ID NO: 73	SEQ ID NO: 93	(同上)	EGR3	80%
SEQ ID NO: 74	SEQ ID NO: 92	(同上)	EGR3	92%
SEQ ID NO: 75	SEQ ID NO: 92	(同上)	EGR3	92%
SEQ ID NO: 76	SEQ ID NO: 93	(同上)	EGR3	80%
SEQ ID NO: 77	SEQ ID NO: 93	(同上)	EGR3	80%
SEQ ID NO: 78	SEQ ID NO: 94	HACPAEGCDRRFSRRDELNVHLRIHTGHKPFQ CRICMRSFSSRRTCRAHIRTHTGEKPFACEFCG RKFAQSSNLVRHAKIHLKQKEHACPAEGCDR RFSQLAHLRAHLRIHTGHKPFQCRICMRSFSTS	EGR3	45%

[0381]

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列	用于 DBD 的蛋白 质平台	DBD 与 蛋白质平 台之间的 NCBI- BLAST 序列同一 性
		GNLVRHIRTHTGEKPFACEFCGRKFAHRTTTLT NHAKIHLKQKE		
SEQ ID NO: 79	SEQ ID NO: 94	(同上)	EGR3	45%
SEQ ID NO: 80	SEQ ID NO: 94	(同上)	EGR3	45%

[0382]

表8:可与本文公开的eTF的任何DBD融合的示例性TAD的氨基酸序列

[0383]

TAD SEQ ID NO.	TAD 序列(N-至 C-末端)	TAD 蛋白质平台
SEQ ID NO: 95	DALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM	VP64 (病毒激活域)
SEQ ID NO: 114	DALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LINSRSSGSPKKR KVG SQYLPD TD DRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPRRIA VPSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFP SGQ ISQASALAPAPPQVLPQAPAPAPAPAMVSALA QAPAPV PVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDE DLGALLGNSTDP AVFTDLASVDNSEFQQLLNQGIPVAP HTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLP NGLLSGDEDFSSIADMDFSALLGSGSGSRDSREGMFLP KPEAGSAISDVFE GREVCQPKRIRPFHPPGSPWANRPLP ASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASH LLED PDEETSQAVKALREMA DTVIPQKEEA AICGQMD LSHPPPRGHLDELTTTLESMTEDLNLD SPLTPELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	VPR (病毒激活域)
SEQ ID NO: 96	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMG MGQFPSPHHHQQQQPQHAFNALMGEHIHYGAGNMNA TSGVRHAMGPGTVNGGHPPSALAPAARFNNSQFMGPP VASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMPDL HPAAGHQMN GTNQHFRDCNPKHSGGSSTPGGSGGSST	CITED2 (人类蛋白质结构域)

[0384]

TAD SEQ ID NO.	TAD 序列(N-至 C-末端)	TAD 蛋白质平台
	PGGSGSSSGGGAGSSNSGGGSGSGNMPASVAHVPAAML LPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQN EFDGMTDFVCKQQPSRVSC	
SEQ ID NO: 97	MADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYA GPGLDGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPF FPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGP QPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGL HRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	CITED4 (人类蛋白质结构域)
SEQ ID NO: 98	SGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGM GQFSPPHHHQQQQPQHAFNALMGEHIIHYGAGNMNAT SGVRHAMGPGTVNGGHPPSALAPAAARFNNSQFMGPPV ASQGGSLPASMQLQKLNNQYFNHHHPYPHNHYMPDLH PAAGHQMNNGTNQHFRDCNPKHSGGSSTPGGSGGSSTP GGSGSSSGGGAGSSNSGGGSGSGNMPASVAHVPAAML PPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQNE FDFMTDFVCKQQPSRVSC	CITED2 (人类蛋白质结构域)
SEQ ID NO: 99	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGM MGQFSPPHHHQQQQPQHAFNALMGEHIIHYGAGNMNA TSGVRHAMGPGTVNGGHPPSALAPAAARFNNSQFMGPP VASQGGSLPASMQLQKLNNQYFNHHHPYPHNHYMPDL HPAAGHQMNNGTNQHFRDCNPKHSGGSSTPGGSGGSST PGGSGSSSGGGAGSSNSGGGSGSGNMPASVAHVPAAM LPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELWLGQN EFDGMTDFVCKQQPSRVSC	CITED2 (人类蛋白质结构域)
SEQ ID NO: 100	AADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAG PGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPF PAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQ PAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGLH RVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	CITED4 (人类蛋白质结构域)

[0385] 表9: TAD与蛋白质平台之间的序列同一性百分比

[0386]

eTF SEQ ID NO:	DBD SEQ ID NO:	TAD SEQ ID NO:	用于 TAD 的蛋白质平台	TAD 相对于 DBD 的位置	TAD 与蛋白质平台的 NCBI-BLAST 序列同一性(%)
SEQ ID NO: 50	SEQ ID NO: 81	N/A	N/A	N/A	N/A
SEQ ID NO: 51	SEQ ID NO: 81	SEQ ID NO: 95	VP64	C-末端	0
SEQ ID NO: 52	SEQ ID NO: 81	SEQ ID NO: 96	CITED2	c-末端	100
SEQ ID NO: 53	SEQ ID NO: 81	SEQ ID NO: 97	CITED4	c-末端	100

[0387]

eTF SEQ ID NO:	DBD SEQ ID NO:	TAD SEQ ID NO:	用于 TAD 的蛋白质平台	TAD 相对于 DBD 的位置	TAD 与蛋白质平台的 NCBI-BLAST 序列同一性(%)
SEQ ID NO: 54	SEQ ID NO: 81	SEQ ID NO: 97	CITED4	N/C-末端	100
SEQ ID NO: 55	SEQ ID NO: 82	SEQ ID NO: 96	CITED2	c-末端	100
SEQ ID NO: 56	SEQ ID NO: 83	SEQ ID NO: 98	CITED2	n-末端	100
SEQ ID NO: 57	SEQ ID NO: 84	SEQ ID NO: 96	CITED2	c-末端	100
SEQ ID NO: 58	SEQ ID NO: 85	SEQ ID NO: 99	CITED2	n-末端	100
SEQ ID NO: 59	SEQ ID NO: 86	SEQ ID NO: 96	CITED2	c-末端	100
SEQ ID NO: 60	SEQ ID NO: 87	SEQ ID NO: 99	CITED2	n-末端	100
SEQ ID NO: 61	SEQ ID NO: 88	SEQ ID NO: 97	CITED4	c-末端	100
SEQ ID NO: 62	SEQ ID NO: 88	SEQ ID NO: 97	CITED4	n/c-末端	100
SEQ ID NO: 63	SEQ ID NO: 89	SEQ ID NO: 97	CITED4	c-末端	100
SEQ ID NO: 64	SEQ ID NO: 89	SEQ ID NO: 97	CITED4	n/C-末端	100
SEQ ID NO: 65	SEQ ID NO: 82	SEQ ID NO: 97	CITED4	C-末端	100
SEQ ID NO: 66	SEQ ID NO: 82	SEQ ID NO: 97	CITED4	N/C-末端	100
SEQ ID NO: 67	SEQ ID NO: 90	SEQ ID NO: 100	CITED4	N/C-末端	100
SEQ ID NO: 68	SEQ ID NO: 90	SEQ ID NO: 100	CITED4	N-末端	100
SEQ ID NO: 69	SEQ ID NO: 91	SEQ ID NO: 100	CITED4	N-末端	100
SEQ ID NO: 70	SEQ ID NO: 92	SEQ ID NO: 96	CITED2	c-末端	100
SEQ ID NO: 71	SEQ ID NO: 92	SEQ ID NO: 99	CITED2	n-末端	100
SEQ ID NO: 72	SEQ ID NO: 93	SEQ ID NO: 6	CITED2	c-末端	100
SEQ ID NO: 73	SEQ ID NO: 93	SEQ ID NO: 99	CITED2	n-末端	100
SEQ ID NO: 74	SEQ ID NO: 92	SEQ ID NO: 97	CITED4	c-末端	100
SEQ ID NO: 75	SEQ ID NO: 92	SEQ ID NO: 97	CITED4	n/c-末端	100

[0388]

eTF SEQ ID NO:	DBD SEQ ID NO:	TAD SEQ ID NO:	用于 TAD 的蛋白质平台	TAD 相对于 DBD 的位置	TAD 与蛋白质平台的 NCBI-BLAST 序列同一性(%)
SEQ ID NO: 76	SEQ ID NO: 93	SEQ ID NO: 97	CITED4	c-末端	100
SEQ ID NO: 77	SEQ ID NO: 93	SEQ ID NO: 97	CITED4	n/c-末端	100
SEQ ID NO: 78	SEQ ID NO: 94	SEQ ID NO: 96	CITED2	C-末端	100
SEQ ID NO: 79	SEQ ID NO: 94	SEQ ID NO: 97	CITED4	N/C-末端	100
SEQ ID NO: 80	SEQ ID NO: 94	SEQ ID NO: 97	CITED4	C-末端	100

[0389] 表10:eTF及其靶位点序列的实例

[0390]

eTF	被 eTF 识别的靶位点	靶位点序列(5'-至-3')	基因靶标
SEQ ID NO: 50	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 51	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 52	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 53	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 54	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 55	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 56	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 57	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 58	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 59	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 60	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 61	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 62	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 63	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN

[0391]

eTF	被 eTF 识别的靶位点	靶位点序列(5'-至-3')	基因靶标
SEQ ID NO: 64	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 65	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 66	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 67	SEQ ID NO: 36	TAGGTACCATAGAGTGAG	SCN1A
SEQ ID NO: 68	SEQ ID NO: 36	TAGGTACCATAGAGTGAG	SCN1A
SEQ ID NO: 69	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 70	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 71	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 72	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 73	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 74	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 75	SEQ ID NO: 35	CTAGGTCAAGTGTAGGAG	SCN1A
SEQ ID NO: 76	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 77	SEQ ID NO: 38	GAGTAGAAAAGAAACACA	PGRN
SEQ ID NO: 78	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 79	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因
SEQ ID NO: 80	SEQ ID NO: 33	AGT GAT AGA GAA CGT ATG	报道基因

[0392] 表11:用于靶向能够上调内源SCN1A基因的靶位点的dSaCas9蛋白序列和指导RNA的实例。

[0393]

dCas 支架	
SEQ ID NO: 102	GTTT TAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAA ATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
dSaCas9氨基酸序列	
SEQ ID NO: 103	KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGR RSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEAR VKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKE QISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEA

[0394]

	KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWK DIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVI TRDENEKLEYEYEFQIENVFQKQKKPTLKQIAKEILVNEEDIKGY RVTSTGKPEFTNLKVYHDIKDITARKEIENAELLDQIAKILTIYQS SEDIQEELTNLNSLTQEEIEQISNLKGYTGTHNLSLKAINLILDEL WHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKR SFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNR QTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDL LNNPFNYEVDHIIIPRSVSFDNSFNKVLVKQEEASKKGNRTPFQY LSSSDSKISYETFKKHILNLAAGKGRISKTKKEYLLEERDINRFSV QKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFT SFLRRKWKFKKERNKGYKHAEDALIINANADFIFKEWKLDKAK KVMENQMFEKQAESMPEIETEYKEIFITPHQIKHIKDFKDYK YSHRVDKKNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDND KLKLINKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKY YEETGNYLTKYSKDNGPVIKKIKYYGNKLNAHLDDITDDYPNSR NKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNS KCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLL NRIEVMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILG NLYEVKSKKHPQIIKKG
SEQ ID NO: 104 (或具有NLS和 HA标签的SEQ ID NO: 103)	MAPKKKRKVGIIHGVPAAKRNILGLAIGITSVGYGIIDYETRDVI DAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLF DYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRG VHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKD GEVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLE TRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKY AYNADLYNALNDLNNLVI TRDENEKLEYEYEFQIENVFQKQKK PTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARK EIIENAELLDQIAKILTIYQSSSEDIQEELTNLNSLTQEEIEQISNLK YTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQ KEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREK NSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLH DMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNSFNK VLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGR ISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRS YFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDA LIINANADFIFKEWKLDKAKKVMENQMFEKQAESMPEIETEY YKEIFITPHQIKHIKDFKDYKYSHRVDKKNRELINDTLYSTRKDD KGNTLIVNNLNGLYDKDNDKLKLINKSPEKLLMYHHDPPQTYQ KLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKDNGPVIKKIKY YGNKLNAHLDDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKF VTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNN DLIKINGELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPP RIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKGKRPAATKK AGQAKKKKGSYPYDVPDYALEDALDDFDLMLGSDALDDFDL DMLGSDALDDFDLMLGSDALDDFDLML
SaCas9指导RNA (gRNA)序列	
SEQ ID NO: 105	TGACCTAGACAGCCTTACATA
SEQ ID NO:	TAAGGCTGTCTAGGTCAAGTG

[0395]

106	
SEQ ID NO: 107	AAGGCTGTCTAGGTCAAGTGT
SEQ ID NO: 108	GATGAAGCCGAGAGGATACTG
SEQ ID NO: 109	GCTGATTTGTATTAGGTACCA
SpCas9指导RNA (gRNA)序列	
SEQ ID NO: 110	CTTCGTTATGTAAGGCTGTCT
SEQ ID NO: 107	AAGGCTGTCTAGGTCAAGTGT
SEQ ID NO: 111	CAAGTGTAGGAGACACACTGC

[0396] 表12: 可以在eTF中组合以调节SCN1A基因的锌指的实例。

[0397]

	序列		序列
SEQ ID NO: 115	RSDNLVR	SEQ ID NO: 123	QSGDLRR
SEQ ID NO: 116	REDNLHT	SEQ ID NO: 124	TSHSLTE
SEQ ID NO: 117	RSDELVR	SEQ ID NO: 125	THLDLIR
SEQ ID NO: 118	QSGNLTE	SEQ ID NO: 126	QSSSLVR
SEQ ID NO: 119	TSGHLVR	SEQ ID NO: 127	TSGNLVR
SEQ ID NO: 120	QNSTLTE	SEQ ID NO: 128	RRDELNV
SEQ ID NO: 121	DPGALVR	SEQ ID NO: 129	RSDDLVR
SEQ ID NO: 122	HRTTLTN	SEQ ID NO: 130	RSDLTN

[0398] 表13: 用于调节SCN1A的eTF的DBD的实例

[0399]

	DBD 的氨基酸序列(包括连接体)
SEQ ID NO: 131	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SREDNLHTHQRTHTEKPYKCPECGKSFSRSDNLVRHQRTHTE KPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKSFSSTSGH LVRHQRTHTEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS
SEQ ID NO: 132	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SHRTTLTNHQRTHTEKPYKCPECGKSFSREDNLHTHQRTHTE KPYKCPECGKSFSSTSHSLTEHQRTHTEKPYKCPECGKSFSQSSSL VRHQRTHTEKPYKCPECGKSFSREDNLHTHQRTHTGKKTS
SEQ ID NO: 133	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SRSDNLVRHQRTHTEKPYKCPECGKSFSQSGDLRRHQRTHTE KPYKCPECGKSFSSTHLDLIRHQRTHTEKPYKCPECGKSFSSTSGN LVRHQRTHTEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS
SEQ ID NO: 134	LEPGEKPYKCPECGKSFSRRDELNVHQRTHTEKPYKCPECGKSF SRSDHLTNHQRTHTEKPYKCPECGKSFSRDDLVRHQRTHTE KPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSFSHRTT LTNHQRTHTEKPYKCPECGKSFSREDNLHTHQRTHTEKPYKC PECGKSFSSTSHSLTEHQRTHTEKPYKCPECGKSFSQSSSLVRHQR THTEKPYKCPECGKSFSREDNLHTHQRTHTGKKTS

[0400]

SEQ ID NO: 135	RSDNLVR x REDNLHT x RSDDELVR x QSGNLTE x TSGHLVR x QNSTLTE (其中 X 表示包含 1-50 个氨基酸残基的连接体)
SEQ ID NO: 371	RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (其中 X 表示包含 1-50 个氨基酸残基的连接体)
SEQ ID NO: 372	RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT (其中 X 表示包含 1-50 个氨基酸残基的连接体)
SEQ ID NO: 376	DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR (其中 X 表示包含 1-50 个氨基酸残基的连接体)
SEQ ID NO: 423	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SREDNLHTHQRTHTEKPYKCPECGKSFSRSDDELVRHQRTHTE KPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKSFSSTSGH LVRHQRTHTEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS
SEQ ID NO: 424	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SREDNLHTHQRTHTEKPYKCPECGKSFSRSDDELVRHQRTHTE KPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKSFSSTSGH LVRHQRTHTEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS
SEQ ID NO: 425	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SREDNLHTHQRTHTEKPYKCPECGKSFSRSDDELVRHQRTHTE KPYKCPECGKSFSQSGNLTEHQRTHTEKPYKCPECGKSFSSTSGH LVRHQRTHTEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS
SEQ ID NO: 426	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTEKPYKCPECGKSF SHRTTLTNHQRTHTEKPYKCPECGKSFSRSDNLHTHQRTHTE KPYKCPECGKSFSSTSHSLTEHQRTHTEKPYKCPECGKSFSQSSSL VRHQRTHTEKPYKCPECGKSFSRSDNLHTHQRTHTGKKTS
SEQ ID NO: 427	RKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKP FQCRICMRNFSREDNLHTHIRTHTEKPFACDICGRKFARSDELV RHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHT KIHLRQKDKKADKSVV
SEQ ID NO: 428	RKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKP FQCRICMRNFSREDNLHTHIRTHTEKPFACDICGRKFARSDELV RHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHT KIHLRQKDKKADKSVV
SEQ ID NO: 429	RKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKP FQCRICMRNFSREDNLHTHIRTHTEKPFACDICGRKFARSDELV RHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHT KIHLRQKDK
SEQ ID NO: 430	RKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKP FQCRICMRNFSREDNLHTHIRTHTEKPFACDICGRKFARSDELV

[0401]

	RHTKIHRLQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSSTGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHT KIHRLQKDK
SEQ ID NO: 431	RKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLH THIRTHTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMR NFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQ KDKKADKSVV
SEQ ID NO: 432	QSQLIKPSRMRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLV RHIRIHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICG RKFAREDNLHTHTKIHRLQKDRPYACPVESCDRRFSTSHSLTEHI RIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKF AREDNLHTHTKIHRLQKDK
SEQ ID NO: 433	QSQLIKPSRMRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLV RHIRIHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICG RKFAREDNLHTHTKIHRLQKDRPYACPVESCDRRFSTSHSLTEHI RIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKF AREDNLHTHTKIH
SEQ ID NO: 434	QSQLIKPSRMRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLV RHIRIHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICG RKFAREDNLHTHTKIHRLQKDRPYACPVESCDRRFSTSHSLTEHI RIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKF AREDNLHTHTKIHRLQKDK
SEQ ID NO: 435	RKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLH THIRTHTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMR NFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQ KDKKADKSVV
SEQ ID NO: 441	RKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSDELV RHIRTHTGEKPFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRICM RNFSSTGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHRL QKDKKADKSVV
SEQ ID NO: 406	RKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLH THTKIHRLQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQC RICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTHT KIHRLQKDKKADKSVV
SEQ ID NO: 407	AAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGNSSSSSSGGGGGGGGNSSSSSSTFNPQA DTGEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRF SLEPAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSAS ASQSPPLSCAVPSNDSSPIYSAAPTFTPNTDIFPEPQSQAFPGSAG TALQYPPPAYPAAGGFQVPMIPDYLFPPQQGDLGLGTPDQKPF QGLESRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSR MRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQ KPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSDEL VRHTKIHRLQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF QCRICMRNFSSTGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTE HTKIHRLQKDKKADKSVVASSATSSLSSYPSPVATSYSPVTTSSYP

[0402]

	SPATTSYPSPVPTSFSSPGSSSTYPSPVHSGFSPSPSVATTYSSVPPAFP AQVSSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC
SEQ ID NO: 408	RPYACPVECDRRFSDPGALVRHRIHTGQKPFQCRICMRNFSRSD NLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKIHRLQKDRPY ACPVECDRRFSTHLDLIRHRIHTGQKPFQCRICMRNFSTSGNLV RHIRTHTGEKPFACDICGRKFARSNDLVRHTKIHRLQKDRPYACP VECDRRFSQSGHLTEHIRIHTGQKPFQCRICMRNFSERSHLREHI RTHTGEKPFACDICGRKFAQAGHLASHTKIHRLQKD
SEQ ID NO: 409	RKYPNRPSKTPPHERPYACPVECDRRFSRSDNLVRHRIHTGQKPF QCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSDELV RHIRTHTGEKPFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRICM RNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHRL QKDKKADKSVV
SEQ ID NO: 391	LEPGEKPYKCPECGKSFSDDPGHLVRHQRTHTGEKPYKCPECGKSF STKNSLTEHQRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGE KPYKCPECGKSFSERSHLREHQRTHTGEKPYKCPECGKSFSKKH LAEHQRTHTGEKPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCP ECGKSFSQSGNLTEHQRTHTGKKTS
SEQ ID NO: 392	LEPGEKPYKCPECGKSFSRRRTCTRAHQRTHTGEKPYKCPECGKSF STTGALTEHQRTHTGEKPYKCPECGKSFSRSDLVLRHQRTHTGEK PYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSQSGDL RRHQRTHTGEKPYKCPECGKSFSSTSHSLTEHQRTHTGKKTS
SEQ ID NO: 393	LEPGEKPYKCPECGKSFSRKDNLKNHQRTHTGEKPYKCPECGKS FSDPGALVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTG EKPYKCPECGKSFSDDPGALVRHQRTHTGEKPYKCPECGKSFSSTG ELVRHQRTHTGEKPYKCPECGKSFSRKDNLKNHQRTHTGKKTS
SEQ ID NO: 394	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSF SREDNLHTHQRTHTGEKPYKCPECGKSFSRSDLVLRHQRTHTGE KPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSSTSGH LVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS
SEQ ID NO: 395	LEPGEKPYKCPECGKSFSKKALTEHQRTHTGEKPYKCPECGKSF SSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGE KPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSD LVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKC PECGKSFSSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQ RTHTGKKTS
SEQ ID NO: 396	LEPGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSF SRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSRSDLVLRHQRTHTGEKPYKCPECGKSFSQSGN LTEHQRTHTGEKPYKCPECGKSFSSTSGHLVRHQRTHTGKKTS
SEQ ID NO: 397	LEPGEKPYKCPECGKSFSKKALTEHQRTHTGEKPYKCPECGKSF SSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGE KPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSD LVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGKKTS
SEQ ID NO: 398	LEPGEKPYKCPECGKSFSDCRDLARHQRTHTGEKPYKCPECGKSF SRNDALTEHQRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGE KPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDDPGN LVRHQRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKC PECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQ RTHTGKKTS

[0403]

SEQ ID NO: 399	LEPGKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSF SSPADLTRHQRTHTGEKPYKCPECGKSFSDPGNLVRHQRTHTGE KPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSQSSS LVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGKKTS
SEQ ID NO: 400	LEPGKPYKCPECGKSFSSTKNSLTEHQRTHTGEKPYKCPECGKSF SRADNLTEHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGE KPYKCPECGKSFSSTKNSLTEHQRTHTGEKPYKCPECGKSFSQAGH LASHQRTHTGEKPYKCPECGKSFSSTHLDLIRHQRTHTGKKTS
SEQ ID NO: 401	LEPGKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSF SDPGHLVRHQRTHTGEKPYKCPECGKSFSSTSGELVRHQRTHTGE KPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCPECGKSFSKKA LTEHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKC PECGKSFSRSDHLTNHQRTHTGKKTS
SEQ ID NO: 402	LEPGKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSF SHRTTLTNHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSSTSHSLTEHQRTHTGEKPYKCPECGKSFSQSSSL VRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGKKTS
SEQ ID NO: 403	LEPGKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSF SRSDNLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTGE KPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCPECGKSFSSTSGN LVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS
SEQ ID NO: 404	LEPGKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPECGKSF SREDNLHTHQRTHTGEKPYKCPECGKSFSSTSGNLTEHQRTHTGE KPYKCPECGKSFSSTHLDLIRHQRTHTGEKPYKCPECGKSFSQKSS LIAHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTS
SEQ ID NO: 405	LEPGKPYKCPECGKSFSSTTGNLTVHQRTHTGEKPYKCPECGKSF STSGELVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGE KPYKCPECGKSFSSTSGNLTEHQRTHTGEKPYKCPECGKSFSQSSS LVRHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGKKTS

[0404] 表14: 包含多个调节SCN1A的锌指的DNA结合域的实例

[0405]

DBD 中的 ZF 的数目	包含以下 ZF 的 DBD (N-至-C 末端):	ZF 序列:	DBD SEQ ID NO.
6	SEQ ID NO: 115 SEQ ID NO: 116 SEQ ID NO: 117 SEQ ID NO: 118 SEQ ID NO: 119 SEQ ID NO: 120	RSDNLVR; REDNLHT; RSDLVLR; QSGNLTE; TSGHLVR; QNSTLTE	SEQ ID NO: 131
6	SEQ ID NO: 115 SEQ ID NO: 122 SEQ ID NO: 116 SEQ ID NO: 124 SEQ ID NO: 126 SEQ ID NO: 116	RSDNLVR; HRTTLTN; REDNLHT; TSHSLTE; QSSSLVR; REDNLHT	SEQ ID NO: 132
6	SEQ ID NO: 121 SEQ ID NO: 115 SEQ ID NO: 123	DPGALVR; RSDNLVR; QSGDLRR;	SEQ ID NO: 133

[0406]

DBD 中的 ZF 的数目	包含以下 ZF 的 DBD (N-至-C 末端):	ZF 序列:	DBD SEQ ID NO.
	SEQ ID NO: 125 SEQ ID NO: 127 SEQ ID NO: 115	THLDLIR; TSGNLVR; RSDNLVR	
9	SEQ ID NO: 128 SEQ ID NO: 130 SEQ ID NO: 129 SEQ ID NO: 115 SEQ ID NO: 122 SEQ ID NO: 116 SEQ ID NO: 124 SEQ ID NO: 126 SEQ ID NO: 116	RRDELNV; RSDHLTN; RSDDLVR; RSDNLVR; HRTTLTN; REDNLHT; TSHSLTE; QSSSLVR; REDNLHT	SEQ ID NO: 134

[0407] 表15: 用于调节SCN1A的靶位点序列的实例

[0408]

靶序列 SEQ ID NO.	被 ZF 或 gRNA 识别的	被 eTF 识别的 SCN1A 靶位点序列
SEQ ID NO: 35	ZF	CTAGGTCAAGTGTAGGAG
SEQ ID NO: 36	ZF	TAGGTACCATAGAGTGAG
SEQ ID NO: 37	ZF	TAGGTACCATAGAGTGAGGCCGAGGATG
SEQ ID NO: 136	ZF	GAGGATACTGCAGAGGTC
SEQ ID NO: 252	ZF	CAAACTCCCAGCCTGCCTGGC
SEQ ID NO: 253	ZF	CCAGCACTGGTGCTTCGT
SEQ ID NO: 254	ZF	AAGGCTGTCTAGGTCAAG
SEQ ID NO: 256	ZF	CTAGGTCAAGTGTAGGAGACACAC
SEQ ID NO: 257	ZF	GGTCAAGTGTAGGAGACA
SEQ ID NO: 258	ZF	CAAGTGTAGGAGACACAC
SEQ ID NO: 259	ZF	AGTGTAGGAGACACACTGCTGGCC
SEQ ID NO: 260	ZF	AGTGTAGGAGACACACTG
SEQ ID NO: 261	ZF	ACTTGACCTAGACAGCCT
SEQ ID NO: 262	ZF	AGGAGACACACTGCTGGCCTG
SEQ ID NO: 264	ZF	TGAATAACTCATTAGTGA
SEQ ID NO: 265	ZF	AAAGTACATTAGGCTAAT
SEQ ID NO: 105	gRNA	TGACCTAGACAGCCTTACATA
SEQ ID NO: 106	gRNA	TAAGGCTGTCTAGGTCAAGTG
SEQ ID NO: 108	gRNA	GATGAAGCCGAGAGGATACTG
SEQ ID NO: 109	gRNA	GCTGATTTGTATTAGGTACCA
SEQ ID NO: 110	gRNA	CTTCGTTATGTAAGGCTGTCT
SEQ ID NO: 107	gRNA	AAGGCTGTCTAGGTCAAGTGT
SEQ ID NO: 111	gRNA	CAAGTGTAGGAGACACACTGC
SEQ ID NO: 195	gRNA	CACTCATTGAGGACAGAAGCA
SEQ ID NO: 196	gRNA	ACAAAGTGGAGAGACTACCCG
SEQ ID NO: 197	gRNA	GGTACGGGCAAAGATTTCTTG
SEQ ID NO: 198	gRNA	TTTGCCCGTACCAAAGTCTTG
SEQ ID NO: 199	gRNA	TCTTCAGCCACAAAAAAAAA
SEQ ID NO: 200	gRNA	GTGTGCAACCCATGAGGTCGC

[0409]

SEQ ID NO: 201	gRNA	ACACAATGAGCCACCTACAAG
SEQ ID NO: 202	gRNA	GTGGCTCATTGTGTGTGTGCC
SEQ ID NO: 203	gRNA	AGAAAGCTGATACAGATACAA
SEQ ID NO: 205	gRNA	AAACCAGCAACAAACAGTATA
SEQ ID NO: 206	gRNA	GAGAATCATCTAGGTCAGGAC
SEQ ID NO: 207	gRNA	TTCTCAGTTTTGAAATTAAAA
SEQ ID NO: 208	gRNA	CATATCCCTGCAGGTTTCAGAA
SEQ ID NO: 209	gRNA	TGGATTCTCTTCTGAACCTGC
SEQ ID NO: 210	gRNA	AGAGAGAGAGAGAGAGAGAGA
SEQ ID NO: 211	gRNA	TGGTCTCATTCTTTTTGTGGG
SEQ ID NO: 224	gRNA	TGTTCCCTCCAGATTAACTT
SEQ ID NO: 225	gRNA	ATTACAGTTCTGTCAGCATGC
SEQ ID NO: 226	gRNA	ATCATCTGTAACCATCAAGGA
SEQ ID NO: 227	gRNA	TCCTGCCTACTTAGTTTCAAG
SEQ ID NO: 228	gRNA	TGCTGAGGCAGGACACAGTGT
SEQ ID NO: 229	gRNA	ACAAAGTAAGTGTGTCAGTGTG
SEQ ID NO: 230	gRNA	CTGACACTTACTTTGTCTAAA
SEQ ID NO: 231	gRNA	ATAATAGTTGTGTCTTTATAA
SEQ ID NO: 232	gRNA	CGATATTTTCATGGATTCTT
SEQ ID NO: 233	gRNA	AAAACCTGGAACCGCATTCCTCA
SEQ ID NO: 234	gRNA	TGTACAAGCAGGGCTGCAAAG
SEQ ID NO: 235	gRNA	AAAGGGGAATGGGAACACCCG
SEQ ID NO: 236	gRNA	ATGTTCAAGGTGCAGAAGGAA
SEQ ID NO: 237	gRNA	TTCAACAAGCTCCCAAGAAGT
SEQ ID NO: 238	gRNA	GTTAACAAATACACTAAACAC
SEQ ID NO: 240	gRNA	AAAACCGTCAACCCCATACAG
SEQ ID NO: 241	gRNA	AGGAATCACTTGTGACACCTA
SEQ ID NO: 242	gRNA	GAGACTGTGCGTATAAAGGCA
SEQ ID NO: 243	gRNA	TACTGATACAGGAAGTGTGCC
SEQ ID NO: 244	gRNA	CTTGTCATGAAACATGAGCTA
SEQ ID NO: 245	gRNA	TATCTGTATTCTTTTATTTT
SEQ ID NO: 246	gRNA	ATAAATGGTAATTAAATGTG
SEQ ID NO: 247	gRNA	TGTTTGCTCAAACGTGCACCA
SEQ ID NO: 248	gRNA	AAATATGTACCACAAGAAATG
SEQ ID NO: 249	gRNA	TATCTGGTTTCTCTCACTGCT
SEQ ID NO: 250	gRNA	AAATAAGACATGAAAACAAGA
SEQ ID NO: 251	gRNA	ATTGCAAAGCATAATTTGGAT

[0410] 表16:可上调SCN1A表达的本文公开的示例性eTF的核酸序列

[0411]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
SEQ ID NO: 305	SEQ ID NO: 353	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaacatatatttgaa gttctgttgacataaagaatcatgatattaatgccatggaaatgaaagggcgatcaacactatggttg aaaagggggaaattgtagacacagatgtgtcgtgtggcagtgctgtctctagcaataactcagag aagagagagaacaatgaaattctgattggcccgatgtgagccagatgaggttcagctgccacttt

[illegible]

[0414]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		aaacggacgggctccgctgaaccagtgaggccccagacgtgcgcataataaccctgcgtgct gcaccacctggggagagggggaggaccacggtaaatggagcgagcgcatagcaaaaggagc cggggctcttttctgcccgtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaaagcgtgcaggggtcacagatctgggggaacccggggaaaagcactga ggcaaaaccgccgctcgtctctacaatatatgggagggggagggtgagtagcttctgattactcat aagaccttttttcttccggcgcaaaaccgtgagctggatttataatcgccctataaagctcaga ggcgggtcaggcacctgcagaggagccccgccgctccgccaactagctgccccgcgagcaacg gcctcgtgattccccgccgatccgggtccccgcctccccactctgccccgcctaccccgagccgt gcagccgcctctccgaatctctcttctcctggcgctcgcgtgcgagaggggaactagcgagaacga ggaagcagctggaggtgacgccgggcagattacgcctgtcagggccgagccgagcggatcgtg ggcgctgtgcagaggaaaggcgggagtgcccggctcgtgtgcagagccgaggtgggtaagct agcgaccacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctctcccgcgcttc tcaaccaaccccatcccagcgccggccaccaacctcccgaatgagtgttcttgcgccagcagc cgaaggcgctactaggaacgtaacctgttacttttccagggccgtagtcgacctgtcccaggt tgctgtgcactgcgcgcgggggtagagtgaaggtagtgtgttcttctgtgccaagtccga gggagaacgtaaatatgggctttttccccctctcacctgtctcaccaaagtccctagtcgccgga gcagtagcccttcttcttccaggaattagccagacacaacaacgggaaccagacaccgaaccaga catgccccccccgtgcgcctccccgctcgtgccttctctcccttctgtctcagagccggtactt caaggggagcctccgtgcccccggtgctcagtcctcgggtgtgcaggaccccggaagtcctcc ccgcacagctctcgttctcttgcagcctgttctgcgcgggaccagtcgaggactctggacagtag aggccccgggacgaccgagctgGAATTCGCCACCATGGCCCCAAAGAA GAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCG AACCAGGTGAAAAACCTTACAAATGTCTGAATGTGGGAAA TCATTAGTCGCAGCGACAACCTGGTGAGACATCAACGCAC CCATACAGGAGAAAAACCTTATAAATGTCCAGAATGTGGAA AGTCCTTCTCACGAGAGGATAACTTGCACACTCATCAACGAA CACATACTGGTGAAAAACCATACAAGTGTCCCGAATGTGGT AAAAGTTTTAGCCGGAGCGATGAACTTGTCCGACACCAACG AACCCATACAGGCGAGAAGCCTTACAAATGTCCCGAGTGTG GCAAGAGCTTCTCACAATCAGGGAATCTGACTGAGCATCAA CGAACTCATACCGGGGAAAAACCTTACAAGTGTCCAGAGTG TGGAAGAGCTTTTCCACAAGTGGACATCTGGTACGCCACCA GAGGACACATACAGGGGAGAAGCCCTACAAATGCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACCGAACACC AGCGAACACACACTGGGAAAAAAACGAGTAAAAGGCCGGC GGCCACGAAAAAGGCCGGCCAGGCAGGCAAAAAAGAAAAAGGGA TCCTACCCATACGACGTACCAGATTACGCTCTCGAGGACGCG CTGGACGATTTTCGATCTCGACATGCTGGGTTCGATGCCCTC GATGACTTTGACCTGGATATGTTGGGAAGCGACGCATTGGAT GACTTTGATCTGGACATGCTCGGCTCCGATGCTCTGGACGAT TTCGATCTCGATATGTTATAAATACTAGTaaagagaccggttcactgtgacagt aaaagagaccggttcactgtgagaatgaaagagaccggttcactgtgacggaagagaccggtt cactgtgagcgcccttgaaccagcagacaatgtagctcagtagaaaccagcagacaatgtagc tgaatggaaccagcagacaatgtagcttcggagaaccagcagacaatgtagctAAGCTT GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCC CTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTTAATA

[illegible]

[0417]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		<p>ggaagcagctggaggtgacgccgggagattacgcctgtcagggccgagccgagcggatcgctg ggcgctgtgcagaggaaggcgaggagtgcccgctcgctgtcgcagagccgaggtgggtaagct agcgaccacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctctcccgcggttc tcaaccaaccccatccagcgccggccaccaacctcccgaatgagtgttctctgccccagcagc cgaaggcgctactaggaacggtaacctgttacttttcagggggccgtatgcacccgtgcccaggt tgctgtgcgactgcgcgcggggctagagtgcagggtgactgtgttctctctgccaagtccga gggagAACgtaaagatatgggccttttccccctctcacctgtctcaccAAgtccctagtccccgga gcagttagcctcttcttccagggaattagccagacacaaacgggaaccagacaccgaaccaga catgccccccccgtgcgcctccccgtcgtcgtcttctctctctgtctctccagagccggatctt caaggggagcctccgtgccccgggtgctcagtcctccgggtgtgcaggaccccggaagtcctcc ccgcacagctctcgttctcttgcagcctgttctgcgcggaccagtcgaggactctggacagtag aggccccgggacgaccgagctgGAATTCGCCACCATGGccgcgccaaggccga gatgcagctgatgtccccgtgcagatctctgacccgttcggatcttctcactcgcaccatgga caactaccctaagctggaggagatgatgtctgagcaacggggctccccagttcctggcgccgc cggggccccagagggcagcggcagcaacagcagcagcagcagcagcggggcggtggaggc ggcgggggcggcagcaacagcagcagcagcagcagcaccttcaacctcaggcgacacggg cgagcagccctacgagcacctgaccgcagagcttttctgacatctctctgaacaacgagaaggtg ctggtggagaccagttacccagccaaaccactcgtcgtcccccatcacctatactggccgttttc cctggagcctgcacccaacagtggcaacacctgtggcccgagccccctctcagcttggctcagtggc ctagtgcagcatgaccaaccaccggcctcctcgtcctcagcaccatctccagcgccctctccgct ccgctccccagagcccaccctgagctgcgcagtgccatccaacgacagcagtccttactcag cggcaccacacctccccacgccgaacactgacatttccctgagccacaaaggcaggccttccgg gtcggcagggacagcgtccagtaccgctcctcgtcctaccctgccccaagggtggcttccag gttccatgatccccgactacctgttccacagcagcagggggatctgggcttgggacccccagac cagaagcccttccaggcctggagagccgcacccagcagccttcgtaaccctctgtctactatta aggcctttgccactcagtcgggctcccaggacctgaaggccctcaataaccagctaccgtccagct catcaaacccagccgatgcgcaagtacccaaccggccagcaagacgccccccacgaacgc ccttacgcttggcagtgaggctctgtgatcgccgttctccCGCAGCGACAACCTGG TGAGAcacatccgcatccacacaggccagaagcccttccagtgcgcatctgcatgAGAAa cttcagcCGAGAGGATAACTTGCACACTcacatccgcaccacacaggcgaaa agcccttcgcctgcgacatctgtggaagaaagtttgcCGGAGCGATGAACCTTGTC CGAcataccaagatccacttgcggcagaaggaccgccccttacgcttggcagtgaggctctgtga tcgccgttctccCAATCAGGGAATCTGACTGAGcacatccgcatccacacagg ccagaagcccttccagtgcgcgatctgcatgAGAAacttcagcACAAGTGGACATCT GGTACGCcacatccgcacccacacaggcgaaaagcccttcgctgcgacatctgtggaaga aagtttgcCAGAATAGTACCCTGACCGAAcataccaagatccacttgcggcag aaggacaagaaagcagacaaaagtgtgtGCTAGCtgggccacctctctctcttcttacc cgtccccgggtgctaccttaccgtccccgggttactaccttattccatccccggccaccacctata cccatccctgtgcccacctcttctctctccccggctcctcgacctaccatccctgtgcacagtg cttccctccccgtcgggtggccaccagctactcctctgttccccctgcttcccgccagggtcagca gcttcccttctcagctgtcaccactcctcagcgctccacagggttccggacatgacagcaacct tttctcccaggacaattgaaatttgctaaACTAGTaataaaagatctttatttcaatagatctgtgtg tggtttttgtgtg</p>
SEQ ID NO: 314	SEQ ID NO: 357	<p>ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaacatatttgaa gttctgtgacataaagaatcatgatattaatgccatggaaatgaaagggcgatcaacactatggttg aaaagggggaaattgtagacacagatgtgtcgtgtggcagtgctgtctctagcaataactcagag</p>

[illegible]

[0419]

[illegible]

[illegible]

[illegible]

[0426]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		ggcaaaaccgccgctcgtctcctacaatatatgggagggggaggttgagtacgttctggattactcat aagaccttttttttcttccggcgcaaaaccgtgagctggattataatcgcctataaagctcaga ggcggtagggcacctgcagaggagccccgccgtccgccaactagctccccgcgagcaacg gcctcgtgattccccgccgatccgggtccccgcctcccaactctgccccgcctaccccgagccgt gcagccgcctctccgaatctctcttctcctggcgctcgcgtgcgagagggaactagcgagaacga ggaagcagctggagggtgacgccgggcagattacgcctgtcaggggccgagccgagcgatcgctg ggcgctgtgcagaggaaaggcgggagtgcccggctcgtgtgcagagccgaggtgggtaagct agcgaccacctggacttcccagcggccaaccgtggcttttcagccaaggctcctcctcccgcggttc tcaaccaacccatccagcggcgccaccaacctcccgaatgagtgttctcgtccccagcagc cgaaggcgctactaggaacggtaacctgttactttccaggggccgtatgcgacctgtcccaggt tgctgtgcgactgcgcgcggggctagagtgcagggtgactgtgttctctctggccaagtccga gggagaacgtaagatatgggccttttccccctctcacctgtctcaccaaagtccctagtcggga gcagtagaccttcttccagggaattagccagacacaacaacgggaaccagacaccgaaccaga catgcccggccgtgcgcctccccgcctgccttctcctccttctgtctcctcagagccggaatt caaggggagcctccgtgccccggctgctcagtcctccgggtgtgcaggaccccggaagtcctcc ccgcacagctctcgttctcttgcagcctgtttctgcgcggaccagtcgaggactctggacagtag aggccccgggagaccgagctgGAATTCGCCACCATGGccggcggaaggccga gatgcagctgatgtcccgtgcagatctctgacctcggatccttctcactcgccaccatgga caactaccctaagctggaggagatgatgtgtgagcaacggggctccccagttcctggcgccgc cggggccccagaggcagcggcagcaacagcagcagcagcagcagcggggcggtggaggc ggcgggggcggcagcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc cgagcagccctacgagcagctgaccgagagcttctcctgacatctctctgaacaacgagaaggtg ctggtggagaccagttacccagccaaaccactgactgcccccatcacctatactggcgcctttc cctggagcctgcacccaacagtggcaacacctgtggcccgagccctctcagcttggctagtggc ctagtgcagatgaccaaccaccggcctcctcgtcctcagcaccatctccagcggcctcctccgct ccgctcctcagagcccacctgagctgcgcagtccatccaacgacagcagtcctatctactcag cggcaccacacctccccacgccgaacactgacatttccctgagccacaagccaggccttccgg gctcggcagggacagcgtccaglacccgcctcctgcctacctgcccgaagggtggcttccag gttccatgatccccgactacctgtttccacagcagcaggggagctgggctgggcacccagac cagaagcccttcagggcctggagagccgcaccagcagccttcgctaaccctctgtctactatta aggccttggcactcagtcgggtcccaggacctgaaggccctcaataccagctaccagtcctcagct catcaaacccagccgcatgcgcaagtacccaaccggccagcaagacgccccccacgaacgc ccttacgcttggccagtggaagtctgtgatcgccgcttctccCGCAGCGACAACCTGG TGAGAcacatccgcatccacacaggccagaagcccttcagtgccgcatctgcatgAGAAa cttcagcCGAGAGGATAACTTGCACACTcacatccgacccacacaggcgaaa agcccttcgctgcgacatctgtggaagaaagtgttccCGGAGCGATGAACCTTGTC CGACACATCCGGACACACACGGGAGAGAAGCCATTTCGCATGt GATATATGTGGGAGAAAGTTCTCCCAATCAGGGAATCTGACT GAGcacatccgcatccacacaggccagaagcccttcagtgccgcatctgcatgAGAAacttc agcACAAGTGGACATCTGGTACGCcacatccgacccacacaggcgaaaagc ccttcgctgcgacatctgtggaagaaagtgttgcCAGAATAGTACCCTGACCGA AcataccaagatccacttgcggcagaaggacaagaagcagacaaaagtgtgtgGCTAGCt cggccacctctctctcttcttaccgtccccgggtgtaccttaccgtccccgggttactacctt tatccatccccggccaccacctataccatccctgtgccacctctctctctccggctcctcg acctaccatccctgtgcagtggttccccctcccgtcgggtggccaccagctactcctgttccc cctgcttcccggcccaggctcagcagcttccctctcagctgtcaccacactcttcagcgctccaca

[0427]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		gggctttcggacatgacagcaaccttttctcccaggacaattgaaattgctaaACTAGTAAG CTTGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTG GCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCTTA ATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCT ATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGG CAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTG GGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTG CAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGC CTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCA GCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTACCATATT GGCCAGGCTGGTCTCCAACCTCCTAATCTCAGGTGATCTACCC ACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCAC TGCTCCCTTCCCTGTCCTT
	SEQ ID NO: 442	ATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTACAAATGTC CTGAATGTGGGAAATCATTAGTCGCAGCGACAACCTGGTG AGACATCAACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGATAACTTGCA CACTCATCAACGAACACATACTGGTGAAAAACCATACAAGT GTCCCGAATGTGGTAAAAGTTTTAGCCGGAGCGATGAACCTG TCCGACACCAACGAACCCATACAGGCGAGAAGCCTTACAAA TGTCCCGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATCTG ACTGAGCATCAACGAACTCATACCGGGGAAAAACCTTACAA GTGTCCAGAGTGTGGGAAGAGCTTTTCCACAAGTGGACATCT GGTACGCCACCAGAGGACACATACAGGGGAGAAGCCCTACA AATGCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAGTACCC TGACCGAACACCAGCGAACACACACTGGGAAAAAAACGAGT AAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAA AGAAAAAGGGATCCTACCCATACGACGTACCAGATTACGCT CTCGAGGAGGCCAGCGGTTCCGGACGGGCTGACGCATTGGA CGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGA TTTTGACCTTGACATGCTTGGTTCCGGATGCCCTTGATGACTTT GACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTTCGAC CTGGACATGCTGATTAACCTCTAGAAGTTCCGGATCTCCGAAA AAGAAACGCAAAGTTGGTAGCCAGTACCTGCCCCGACACCGA CGACCGGCACCGGATCGAGGAAAAGCGGAAGCGGACCTACG AGACATTCAAGAGCATCATGAAGAAGTCCCCCTTCAGCGGC CCCACCGACCCTAGACCTCCACCTAGAAGAATCGCCGTGCCC AGCAGATCCAGCGCCAGCGTGCCAAAACCTGCCCCCAGCC TTACCCCTTCACCAGCAGCCTGAGCACCATCAACTACGACGA GTTCCCTACCATGGTGTTCCCCAGCGGCCAGATCTCTCAGGC CTCTGCTCTGGCTCCAGCCCCCTCCTCAGGTGCTGCCTCAGGC TCCTGCTCCTGCACCAGCTCCAGCCATGGTGTCTGCACTGGC TCAGGCACCAGCACCCGTGCCTGTGCTGGCTCCTGGACCTCC ACAGGCTGTGGCTCCACCAGCCCCTAAACCTACACAGGCCG GCGAGGGCACACTGTCTGAAGCTCTGCTGCAGCTGCAGTTCC

[0428]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		ACGACGAGGATCTGGGAGCCCTGCTGGGAAACAGCACCGAT CCTGCCGTGTTACCGACCTGGCCAGCGTGGACAACAGCGA GTTCCAGCAGCTGCTGAACCAGGGCATCCCTGTGGCCCCTCA CACCACCGAGCCCATGCTGATGGAATACCCCGAGGCCATCA CCCGGCTCGTGACAGGCGCTCAGAGGCCTCCTGATCCAGCTC CTGCCCCCTCTGGGAGCACCAGGCCTGCCTAATGGACTGCTGT CTGGCGACGAGGACTTCAGCTCTATCGCCGATATGGATTTCT CAGCCTTGCTGGGCTCTGGCAGCGGCAGCCGGGATTCCAGG GAAGGGATGTTTTTGCCGAAGCCTGAGGCCGGCTCCGCTATT AGTGACGTGTTTGAGGGCCGCGAGGTGTGCCAGCCAAAACG AATCCGGCCATTTTCATCCTCCAGGAAGTCCATGGGCCAACCG CCCACTCCCCGCCAGCCTCGCACCAACACCAACCGGTCCAGT ACATGAGCCAGTCGGGTCACTGACCCCGGCACCAGTCCCTC AGCCACTGGATCCAGCGCCCGCAGTGACTCCCGAGGCCAGT CACCTGTTGGAGGATCCCGATGAAGAGACGAGCCAGGCTGT CAAAGCCCTTCGGGAGATGGCCGATACTGTGATTTCCCAGA AGGAAGAGGCTGCAATCTGTGGCCAAATGGACCTTTCCCATC CGCCCCAAGGGGCCATCTGGATGAGCTGACAACCACACTT GAGTCCATGACCGAGGATCTGAACCTGGACTCACCCCTGACC CCGGAATTGAACGAGATTCTGGATACCTTCTGAACGACGA GTGCCTCTTGATGCCATGCATATCAGCACAGGACTGTCCAT CTTCGACACATCTCTGTTT
	SEQ ID NO: 443	ATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTACAAATGTC CTGAATGTGGGAAATCATTTCAGTCGCAGCGACAACCTGGTG AGACATCAACGCACCCATACAGGAGAAAAACCTTATAAATG TCCAGAATGTGGAAAGTCTTCTCACGAGAGGATAACTTGCA CACTCATCAACGAACACATACTGGTGAAAAACCATACAAGT GTCCCGAATGTGGTAAAAGTTTTAGCCGGAGCGATGAACTTG TCCGACACCAACGAACCCATACAGGCGAGAAGCCTTACAAA TGTCCCGAGTGTGGCAAGAGCTTCTACAATCAGGGAATCTG ACTGAGCATCAACGAACATACACCGGGGAAAAACCTTACAA GTGTCCAGAGTGTGGGAAGAGCTTTTCCACAAGTGGACATCT GGTACGCCACCAGAGGACACATACAGGGGAGAAGCCCTACA AATGCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAGTACCC TGACCGAACACCAGCGAACACACACTGGGAAAAAAACGAGT AAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAA AGAAAAAGGGATCCTACCCATACGACGTACCAGATTACGCT CTCGAGGACGCGCTGGACGATTTTCGATCTCGACATGCTGGGT TCTGATGCCCTCGATGACTTTGACCTGGATATGTTGGGAAGC GACGCATTGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTTCGATCTCGATATGTTA
	SEQ ID NO: 444	ATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGT CCCAGCAGCCCTCGAACCAGGTGAAAAACCTTACAAATGTC CTGAATGTGGGAAATCATTTCAGTCGCAGCGACAACCTGGTG AGACATCAACGCACCCATACAGGAGAAAAACCTTATAAATG

[0429]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		<p>TCCAGAATGTGGAAAGTCCTTCTCACGAGAGGATAA CTTGCA CACTCATCAACGAACACATACTGGTGAAAAACCATA CAAGT GTCCCGAATGTGGTAAAAGTTTTAGCCGGAGCGATGA A CTTG TCCGACACCAACGAACCCATACAGGCGAGAAGCCTTACA AA TGTCCCGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATCTG ACTGAGCATCAACGAACTCATACCGGGGAAAAACCTTACA A GTGTCCAGAGTGTGGGAAGAGCTTTTCCACAAGTGGA CATCT GGTACGCCACCAGAGGACACATACAGGGGAGAAGCCCTACA AATGCCCCGAATGCGGTAAAAGTTTCTCTCAGAATAGTACCC TGACCGAACACCAGCGAACACACACTGGGAAAAAAAACGAGT AAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGC AAAA AGAAAAAGGGATCCTACCCATACGACGTACCAGATTACGCT CTCGAGGACGCGCTGGACGATTTTCGATCTCGACATGCTGGGT TCTGATGCCCTCGATGACTTTGACCTGGATATGTTGGGAAGC GACGCATTGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTTCGATCTCGATATGTTA</p>
	SEQ ID NO: 445	<p>ATGGccgcggccaaggccgagatgcagctgatgtccccgctgcagatctctgacccgttcggat cctttctcactcgccaccatggacaactaccctaagctggaggagatgatgctgctgagcaacgg ggctccccagttcctcggcgccgccggggccccagaggcagcggcagcaacagcagcagcag cagcagcggggcggtggaggcgccggggcgccagcaacagcagcagcagcagcagcagcacc ttcaacctcagcgggacacggcgagcagccctacgagcacctgaccgcagagcttttctgac atctcttgaacaacgagaaggtgctggtggagaccagttacccagccaaccactcgactgcc cccatcacctatactggccgttttccctggagcctgcaccaacagtggaacacctgtggcccg gcccccttcagcttggtcagtggtcctagtgagcatgaccaaccaccggcctcctcgtctcagc catctccagcggcctcctccgctccgctccagagccaccctgagctgcgcagtgccatcca acgacagcagtccttactcagcggcaccaccttccccacgccgaacctgacatttccctgag ccacaaagccaggccttcccggtcggcagggacagcgctccagtaccgcctcctgctaccct ggcgccaagggtggcttccaggttccatgatccccgactacctgttccacagcagcagggggatc tgggcctgggcacccagaccagaagccctccaggcctggagagccgcaccagcagccttc gtaacccctctgtctactattaaggccttggcactcagtcgggctccaggaacctgaaggccctca ataccagctaccagtcccagctcatcaaacccagccgatgcgcaagtacccaaccggccagc aagacgccccccacgaacgcccttacgcttggccagtggagtctgtgatcgccgttctccCG CAGCGACAACCTGGTGAGAcacatccgcatccacacaggccagaagcccttcca gtgccgatctgcatgAGAAacttcagcCGAGAGGATAA CTTGCACACTcac atccgacccacacaggcgaaaagcccttcgctgcgacatctgtgaagaaagtgttccCGGA GCGATGAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgccttac gcttggccagtggagtctgtgatcgccgttctccCAATCAGGGAATCTGACTGA GcacatccgcatccacacaggccagaagcccttcagtgccgatctgcatgAGAAacttcagc ACAAGTGGACATCTGGTACGCacatccgcacccacacaggcgaaaagccctt cgctgcgacatctgtgaagaaagtgttggcCAGAATAGTACCCTGACCGAAca taccaagatccacttgcggcagaaggacaagaaagcagacaaaagtgtgtGCTAGCtcgg ccacctctctctcttctacccgtccccgggtgctacctcttaccgtccccgggttactaccttctc catccccggccaccacctcataccatccccgtgcccacctcttctctctccccggtcctcgacct acctatccccgtgcacagtggcttccccctccccgtcggtggccaccagctactctctgttccccctg cttccccggccaggtcagcagcttcccttctcagctgtcaccaactcttcagcgctccacaggg cttctggacatgacagcaaccttttctccaggacaattgaaatttgc</p>

[0430]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
	SEQ ID NO: 446	<p>ATGGCCgccgaccacctgatgctcgccgagggctaccgctggtgcagaggccgctccg ccgcccgcgcccatggccctcatgcgtccggactctgccgctacgcgggcccgggctggac agtgggctgaggccgcgggggctccgctggggccgcgcgccccccaacccggggcctg gcgtacggggccttcgggcccgcgtcctcctccagccctttccggcgtgctccgcggccgcg ggcatcgcgcacctgcagcctgtggcgacgccgtacccggccgcgcggccgcccccaac gctccgggaggccccccggggccgcagccggccccaaagcgcgcagccccggccgcccgc gcacgccctggcgggcatggacgccgaactcatcgacgaggaggcgtgacgtcgtggagctg gagctggggctgcaccgcgtgcgcgagctgccgagctgttctgggcccagagcagttcgactg cttctggacttggggctccgcgcgcccggcggctccgtgagctgcggtggttctggtggttctg gtcagtcacagctcatcaaacccagccgatgcgaagtacccaaccggcccagcaagacgcc ccccacgaacgcccttacgcttgcacagtggagtcctgtgatcgccgttctccCGCAGCGA CAACCTGGTGAGAcacatccgatccacacaggccagaagcccttcagtgccgatct gcatgAGAAacttcagcCGAGAGGATAACTTGCACACTcacatccgaccca cacaggcgaaaagcccttcgctgcgacatctgtggaagaaagtgtccCGGAGCGATGA ACTTGTCCGAcataccaagatccacttgcggcagaaggaccgcccttacgcttgcacgtg gagtcctgtgatcgccgcttccCAATCAGGGAATCTGACTGAGcacatccgca tcacacaggccagaagcccttcagtgccgatctgcatgAGAAacttcagcACAAGTG GACATCTGGTACGCcacatccgacccacacaggcgaaaagcccttcgctgcgacat ctgtggaagaaagtgtccCAGAATAGTACCCTGACCGAAcataccaagatcca cttgcggcagaaggacaagCTCGAGatggccgaccacctgatgctcgccgagggtaccgc ctggtgcagaggccgctccgcgcgcccgcgccatggccctcatgcgtccggactctgccgc gtacgcggggcccgccctggacagtgggctgaggccgcgggggctccgctggggccgcgcg gccccccaacccggggccctggcgtaggggcttcggccgcccgtcctcctccagcccttcc ggcgtgcttccgcccgcgggcatcgccacctgcagcctgtggcgacgccgtaccccggc cgcgcggccgcgcccccaacgctccgggaggccccccggggccgcagccggccccaaagcgc cgacccccgcggcccccgcacgccctggcgggcatggagccgaactatcgacgagga ggcgctgacgtcgtggagctggagctgggctgcaccgcgtgcgcgagctcccagctgttcc tgggccagagcgagttcgactgcttctggacttgggtccgcgcgcccgcggctccgtgagct gc</p>
	SEQ ID NO: 447	<p>ATGGccgcggccaaggccgagatgcagctgatgtccccgctgcagatcttgacccgttcgat ccttctcactcgcccaccatggacaactaccctaagctggaggagatgatgctgctgagcaacgg ggctccccagttcctcggcgcgcccggggcccccagaggcagcggcagcaacagcagcagcag cagcagcgggggctggaggcgccggggcgccagcaacagcagcagcagcagcagcagcagc ttcaacctcagcggaacggcgagcagccctacgagcacctgaccgcagagcttttctgac atctcttgaacaacgagaaggtgctggtggagaccagttacccagccaaacctcagctgcc cccatcacctatactggccgctttccctggagcctgcacccaacagtggaacacctgtggccga gcccctctcagcttggcagtgccctagtgagcatgaccaaccaccggcctcctcctcagcac catctccagcgccctcctccgctccgctccagagccaccctgagctggcagtgccatcca acgacagcagctccatttactcagcggaacccaccttcccacgccgaacactgacatttctcctgag ccaaaagccaggccttcccgggctcggcagggacagcgtccagtaccgctcctgctacccct ggcgaaggggtggcttccaggttcccatgatccccgactacctgttccacagcagcaggggagc tgggcttgggcaccccagaccagaagccctccagggcctggagagccgacccagcagccttc gtaacccctctgtctactattaaggccttggcactcagtcgggctcccaggacctgaaggccctca ataccagctaccagtcccagctcatcaaacccagccgatgcgaagtacccaaccggcccagc aagacgccccccacgaacgcccttacgcttgcacagtggagtcctgtgatcgccgcttctccCG CTCAGACAACCTCGTTTCGAcacatccgatccacacaggccagaagcccttcag</p>

[0431]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		tgccgcatctgcatgAGAAacttcagcCACCGGACTACACTCACGAACcacat ccgcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtgttccAGAGA AGACAATCTCCATACTCACATCCGGACACACACGGGAGAGA AGCCATTTCGCATGtGATATATGTGGGAGAAAAGTTCTCCACCA GCCATTCTCTCACTGAACacatccgcatccacacaggccagaagcccttcagtgcc gcatctgcatgAGAAacttcagcCAGTCTAGCTCACTGGTGAGGcacatccgc acccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtgttccAGGGAGG ATAACCTGCATACGcataccaagatccacttgcggcagaaggacaagaaagcagaca aaagtgtgtgGCTAGCtggccacctctctctcttcttaccggtcccggttctaccttta cccgtccccggttactaccttattccatccccggccaccacctataccatccctgtgcccacctc ctctctctccccggtctctcgacctaccatccctgtgacagtggttccccctccccgtcggtggc caccagtgactctctgttccccctgtttcccgcccgagtcagcagcttccccctctcagctgtacc aactcttcagcgctccacagggtttcggacatgacagcaaccttttctccaggacaattgaaatt tgc
	SEQ ID NO: 448	ATGGCCgccgaccacctgatgctcgccgagggtaccgctggtgcagaggccgcccgtccg ccgcccgcgcccatggccctcatgcgtccggactctgccgccgtacgcgggcccgggctggac agtgggctgaggccgcggggggctccgtggggccgccgcccgcgaacccggggccctg gcgtacggggccttcggggccgccgtcctcctccagcccttccggcgtgctccgccggccgcg ggcatcgcgcacctgcagcctgtggcgacgccgtaccccgccgcgcggccgcgcccccaac gctccgggaggccccccggggccgcagccggccccagcgccgcagccccgccgcccgc gcacgccctggggcgcatggacgccgaactcatcgacgaggaggcgtgacgtcgtggagctg gagctggggctgcaccgctgcgcgagctgccgagctgttctggggcagagcgagttcgactg ctctcggacttggggtccgcgcggccggcgtccgtgagctgcggtgtgtctgtgtgtgtctg gtcagtcacagtcataaaccagccgcatgcgcaagtaccccaaccggcccagcaagacgcc cccacgaacgcccttacgttggccagtggagtctgtgatcgccgttctccCGCAGCGA CAACCTGGTGAGAcacatccgcatccacacaggccagaagcccttcagtgccgcatct gcatgAGAAacttcagcCGAGAGGATAACTTGCACACTcacatccgcacca cacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtgttccCGGAGCGATGA ACTTGTCCGAcataccaagatccacttgcggcagaaggaccgccccttacgttggccagtg gagtctgtgatcgccgttctccCAATCAGGGAATCTGACTGAGcacatccgca tcacacaggccagaagcccttcagtgccgcatctgcatgAGAAacttcagcACAAGTG GACATCTGGTACGCcacatccgcaccacacaggcgaaaagcccttcgcctgcgacat ctgtggaagaaagtgttccCAGAATAGTACCCTGACCGAAAcataccaagatcca ctgcccgcagaaggacaag
	SEQ ID NO: 449	ATGGCCgccgaccacctgatgctcgccgagggtaccgctggtgcagaggccgcccgtccg ccgcccgcgcccatggccctcatgcgtccggactctgccgccgtacgcgggcccgggctggac agtgggctgaggccgcggggggctccgtggggccgccgcccgcgaacccggggccctg gcgtacggggccttcggggccgccgtcctcctccagcccttccggcgtgctccgccggccgcg ggcatcgcgcacctgcagcctgtggcgacgccgtaccccgccgcgcggccgcgcccccaac gctccgggaggccccccggggccgcagccggccccagcgccgcagccccgccgcccgc gcacgccctggggcgcatggacgccgaactcatcgacgaggaggcgtgacgtcgtggagctg gagctggggctgcaccgctgcgcgagctgccgagctgttctggggcagagcgagttcgactg ctctcggacttggggtccgcgcggccggcgtccgtgagctgcggtgtgtctgtgtgtgtctg gtcagtcacagtcataaaccagccgcatgcgcaagtaccccaaccggcccagcaagacgcc cccacgaacgcccttacgttggccagtggagtctgtgatcgccgttctccCGCTCAGA CAACCTCGTTCGAcacatccgcatccacacaggccagaagcccttcagtgccgcatct

[0432]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		gcatgAGAAaacttcagcCACCGGACTACACTCACGAACcacatccgcacca cacaggcgaaaagcccttcgctgcgacatctgtggaagaaagtttgcAGAGAAGACA ATCTCCATACTcataccaagatccacttgcggcagaaggaccgcccttacgcttgccagt ggagtctgtgatcgccgcttccACCAGCCATTCTCTCACTGAAcacatccgc atccacacaggccagaagcccttcagtgccgcatctgcatgAGAAaacttcagcCAGTCTA GCTCACTGGTGAGGcacatccgcaccacacaggcgaaaagcccttcgctgcgaca tctgtggaagaaagtttgcAGGGAGGATAACCTGCATACGcataccaagatcc acttgcggcagaaggacaagCTCGAGatggccgaccacctgatgctcgccgagggtaccg cctggtgcagaggccgcccgcggccgcccgcctatgcccctcatgctcggactctgccgc cgtacgcggggcccgccctggacagtgggctgaggccgcggggggctcgcctggggccgccc cgccccccaacccggggccctggcgtacggggccttcgggccgcccgtcctcctccagcccttc cgccgtgctccgcccggccgcccgcacatgcgcacctgcagcctgtggcgacgccgtacccggg ccgcgcggccgcgcccccaacgctccgggaggccccccggggccgcagccggccccaaagcg ccgcagccccgcccggccgcccgcacgcccctgggcggcatggacgcccgaactcatgcagagg aggcgctgacgtcgtggagctggagctggggtgcaccgcgtgcgcgagctgcccgagctgttc ctgggccagagcgagttcgactgcttctcggacttgggtccgcgccgccgcccgggtcctgtgagc tgc
	SEQ ID NO: 450	ATGGCCgccgaccacctgatgctgcggagggtaccgctggtgcagaggccgcccgtccg ccgcgcggcccatggccctcatgcgctccggactctgccgccgtacgcggggccgggctggac agtgggctgaggccgcgggggctccgtggggccgcccggccccaacccggggccctg gcgtacggggccttcggggccgcccgtcctcctccagcccttcggccgtgctccgcccggccgcg ggcatcgcgcacctgcagcctgtggcgacgccgtacccggccgcgcggccgcgcccccaac gctccgggaggccccccggggccgcagccggccccaaagcgccgcagccccgcccggccgcg gcacgcccctgggcggcatggacgcccgaactcatgcagaggaggcgtgacgtcgtggagctg gagctggggctgcaccgctgcgcgagctgcccgagctgttctgggccagagcgagttcgactg cttctggacttgggtccgcggcccccgggtccgtgagctgcgggtggttctggtggtgttctg gtcagttccagctcatcaaaaccagccgatgcgcaagtaccccaaccggcccagcaagacgcc ccccacgaacgcccttacgcttgcccagtggaagtcctgtgatcgccgttctccCGTCTAGA CAACCTCGTTTCGAcacatccgatccacacaggccagaagcccttcagtgccgcatct gcatgAGAAaacttcagcCACCGGACTACACTCACGAACcacatccgcacca cacaggcgaaaagcccttcgctgcgacatctgtggaagaaagtttgcAGAGAAGACA ATCTCCATACTcataccaagatccacttgcggcagaaggaccgcccttacgcttgccagt ggagtctgtgatcgccgcttccACCAGCCATTCTCTCACTGAAcacatccgc atccacacaggccagaagcccttcagtgccgcatctgcatgAGAAaacttcagcCAGTCTA GCTCACTGGTGAGGcacatccgcaccacacaggcgaaaagcccttcgctgcgaca tctgtggaagaaagtttgcAGGGAGGATAACCTGCATACGcataccaagatcc acttgcggcagaaggacaag
	SEQ ID NO: 451	ATGGccgcggccaaggccgagatgcagctgatgtccccgctgcagatctctgaccggttcggat cctttcctcactcgccaccatggacaactaccctaagctggaggagatgatgctgctgagcaacgg ggctccccagttcctcggcgccgcccggggcccccagaggcagcggcagcaacagcagcagcag cagcagcggggggcgtggaggcggcggggggcggcagcaacagcagcagcagcagcagcacc tcaaccctcaggcggacacgggcgagcagccctacgagcaactgaccgcagagcttcttctgac atctcttgaacaacgagaaggtgctggtggagaccagttacccagccaaaccactgactgccc cccatcactatactggccgcttttccctggagcctgcaccaacagtggaacaccttgtggccga ggcccttctcagcttggctagtgccctagtgagcatgaccaaccaccggcctcctcgtcctcagcac catctccagcgccctcctccgctccgctccagagccaccctgagctgcgcagtgccatcca

[0433]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		acgacagcagtcctcatttactcagcggcaccaccttccccacgccgaacactgacattttccctgag ccacaaagccaggccttcccgggtcggcaggacagcgtccagtaccggcctctgctaccct gccgccaagggtggcttcagggtcccatgatccccgactacctgtttccacagcagcaggggagc tgggcttgggcaccccagaccagaagcccttccaggcctggagagccgcaccagcagccttc gtaacccctctgtctactattaaggccttggccactcagtcgggtccccaggacctgaaggccctca ataccagctaccagtcctcagctcatcaaacccagccgcatgcgcaagtacccaaccggcccagc aagacgccccccacgaacgcccttacgcttcccagtgagtcctgtgatcgccgttctccCG CAGCGACAACCTGGTGAGAcacatccgcatccacacaggccagaagcccttcca gtccgcatctgcatgAGAAacttcagcCGAGAGGATAACTTGCACACTcac atccgcaccacacaggcgaaaagcccttcgctgcgacatctgtgaagaaagtgttccCGGA GCGATGAACTTGTCCGACataccaagatccacttgcggcagaaggaccgccccttac gcttcccagtgagtcctgtgatcgccgcttctccCAATCAGGGAATCTGACTGA GcacatccgcatccacacaggccagaagcccttccagtcgcatctgcatgAGAAacttcagc ACAAGTGGACATCTGGTACGCcacatccgacccacacaggcgaaaagccctt cgctgcgacatctgtgaagaaagtgttccCAGAATAGTACCCTGACCGAAca taccagatccacttgcggcagaaggacaagaaagcagacaaaagtgtgtGCTAGCtgg ccacctctctctcttctaccggtccccggttctaccttaccggtccccggttactaccttctac catccccggccaccacctataccatccccctgtgccacctcttctctctccggctctctgacct accatccccctgtgcacagtggcttccccctcccgctgggtggccaccacgtactctctgttccccctg cttccccggcccaggtcagcagcttcccttctcagctgtcaccacactcttcagcgcctccacaggg cttcggacatgacagcaaccttttccaggacaattgaaatttgc
	SEQ ID NO: 452	ATGGccgcggccaaggccgagatgcagctgatgtccccgctgcagatctctgaccttccgat ccttctcactcgcaccatggacaactaccctaagctggaggagatgatgtgctgagcaacgg ggtccccagttctcggcgccgccggggccccagagggcagcggcagcaacagcagcagcag cagcagcggggggcgggtggaggcggcgggggcggcagcaacagcagcagcagcagcagcacc tcaacctcagcgggacacggcgagcagccctacgagcacctgaccgcagagtcttctctgac atctctctgaacaacgagaagggtgctggtggagaccagttacccagccaaaccactcgactgcc cccatcacctatactggccgctttccctggagcctgcaccaacagtggaacacctgtggccgga gccctcttcagcttggcagtggtcagtgagcatgaccaaccaccggcctctctgctcagcac catctccagcggcctctccgctccgctccagagccacccttgagctgcgcatgacctcca acgacagcagtccttactcagcggcaccacacttccccacgccgaacactgacattttccctgag ccacaaagccaggccttcccgggtcggcaggacagcgtccagtaccggcctctgctaccct gccgccaagggtggcttcagggtcccatgatccccgactacctgtttccacagcagcaggggagc tgggcttgggcaccccagaccagaagcccttccaggcctggagagccgcaccagcagccttc gtaacccctctgtctactattaaggccttggccactcagtcgggtccccaggacctgaaggccctca ataccagctaccagtcctcagctcatcaaacccagccgcatgcgcaagtacccaaccggcccagc aagacgccccccacgaacgcccttacgcttcccagtgagtcctgtgatcgccgttctccCG CAGCGACAACCTGGTGAGAcacatccgcatccacacaggccagaagcccttcca gtccgcatctgcatgAGAAacttcagcCGAGAGGATAACTTGCACACTcac atccgcaccacacaggcgaaaagcccttcgctgcgacatctgtgaagaaagtgttccCGGA GCGATGAACTTGTCCGACACATCCGGACACACACGGGAGAG AAGCCATTCGCATGtGATATATGTGGGAGAAAGTTCTCCCAA TCAGGGAATCTGACTGAGcacatccgcatccacacaggccagaagcccttccagt gccgcatctgcatgAGAAacttcagcACAAGTGGACATCTGGTACGCcacat ccgcaccacacaggcgaaaagcccttcgctgcgacatctgtgaagaaagtgttccCAGAA TAGTACCCTGACCGAAcataccaagatccacttgcggcagaaggacaagaagcag

[illegible]

[0434]

[0435]

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
		accaaccttttttaaataaataatttataaaaaattggtatttcacatatatactagttattacatttatccacaca aaacggacgggctccgctgaaccagttagggccccagacgtgcgcataataaacctgcgtgct gcaccacctggggagagggggaggaccacggtaaatggagcgcgcataagcaaaaggacg cggggctcttttctgccgggtggcactgggtagctgtggccaggtgtggtactttgatggggcccag ggctggagctcaaggaaagcgtcgcaggggtcacagatctgggggaaccccggggaaaagcactga ggcaaaaccgccgtcgtctctacaatatatgggagggggagggttgagtacgttctggtactcat aagaccttttttctcctccggcgcaaaaccgtgagctggatttataatcggcctataaagctccaga ggcggtcaggcacctgcagaggagccccgccgtccggcactagctccccgcgcagcaacg gcctcgtgattccccgccgatccgggtccccgcctccccactctgccccgcctaccccgagccgt gcagccgctctccgaatctctcttctcctggcgctcgcgtgcgcagagagggaactagcgcagaacga ggaagcagctggaggtgacggcgggcagattacgcctgtcagggccgagccgagcggatcgtg ggcgctgtgcagaggaaaggcgggagtgcccggtcgcgtgcgcagagccgaggtgggtaagct agcgaccacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccgcgcttc tcaaccaaccccatccagcgccggccaccaacctcccgaatgagtgttctcgtccccagcagc cgaaggcgctactaggaacgtaacctgttacttttcaggggccgtagtcgacccgctgcccgagt tgcgtgtgcgactgcgcgcggggctagagtgaagggtgactgtggttcttctgtgccaaagtcgga gggagaacgtaaagatatgggcctttttccccctctcacctgtctcaccaaaagtccttagtccccgga gcagttagcctcttcttccagggaattagccagacacaacaacgggaaccagacaccgaaccaga catgccccccccgtgcgccctccccgctcgtgccttctcctccttctgtctcctcagagccggtactt caaggggagcctccgtgcccccggtgctcagtcctccgggtgtgcaggaccccggaagtcctcc ccgcacagctctcgttctcttgcagcctgttctgcgcggaccagtcgaggactctggacagtag aggccccgggacgaccgagctgGAATTCGCCACCATGGcccgcgccaaggccga gatgcagctgatgtccccgctgcagatctctgaccttcggatcctttctcactcggccacctgga caactaccctaagctggaggagatgatgctgctgagcaacggggctccccagttcctcggcgccgc cggggccccagagggcagcggcagcaacagcagcagcagcagcagcggggcggtggaggc ggcgggggcggcagcaacagcagcagcagcagcagcacttcaacctcaggcgacacggg cgagcagccctacgagcacctgaccgcagagcttttctgacatctctctgaacaacgagaagggtg ctggtggagaccagttaccccagccaaaccactcgaactgcccccatcacctatactggcgcttttc cctggagcctgcacccaacagtggcaaaccttgtggcccgagccctctcagcttggtcagtggc ctagtgcagatgaccaaccaccggcctcctcgtcctcagcaccatctccagcgccctcctcgct ccgctccagagcccaccctgagctgcgcagtgccatccaacgacagcagtccttactcag cggcaccaccttccccacggcgaacactgacatttccctgagccacaaagccagccttccgg gctcggcagggacagcgtccagtaccgctcctgctacccgtggcccaagggtggcttccag gttcccatgatccccgactacctgtttccacagcagcagggggatctgggctgggcaccccgagc cagaagcccttccaggcgctggagagccgcacccagcagccttcgtaacctctgtctactatta aggcctttgccactcagtcgggtcccaggacctgaaggccctcaataccagctaccagtcaccagct catcaaacccagccgatgcgcaagtaccccaaccggccagcaagacgccccccacgaagc ccttacgcttggccagtggagtctgtgatcgccgttctccCGCAGCGACAACCTGG TGAGAcacatccgcatccacacaggccagaagcccttcagtgccgcatctgcatGAGaa cttagcCGAGAGGATAACTTGCACACTcacatccgcaccacacaggcgaaa agcccttcgctgcgacatctgtggaagaaagtttgcCGGAGCGATGAACTTGTC CGACACATCCGGACACACACGGGAGAGAAGCCATTTCGCATGt GATATATGTGGGAGAAAGTTCTCCCAATCAGGGAATCTGACT GAGcacatccgatccacacaggccagaagcccttcagtgccgcatctgcatGAGaaactt agcACAAGTGGACATCTGGTACGCcacatccgcaccacacaggcgaaaagc cttcgcctgcgacatctgtggaagaaagtttgcCAGAATAGTACCCTGACCGA

eTF 蛋 白质 SEQ ID NO:	核酸 SEQ ID NO:	核酸序列
[0436]		AcataccaagatccacttgcggcagaaggacaagaaagcagacaaaagtgtgtgGCTAGCt cggccacctctctctctctctctaccggtcccggtgctacctctaccggtcccggttactacctt tatccatcccgccaccacctataccatccctgtgccacctctctctctctccggctctcg acctaccatccctgtgcacagtggcttccctcccggtcggtggccaccacgtactctctgttccc cctgcttcccgccagggtcagcagcttcccttctcagctgtcaccaactcctcagcgctccaca gggcttccggacatgacagcaacctttctcccaggacaattgaaattgctaaACTAGTAAG CTGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTG GCCCTGGAAGTTGCCACTCCAGTGCCACCAGCCTTGTCTTA ATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCT ATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGG CAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTG GGAACCAAGCTGGAGTGCAGTGGCACAACTCTGGCTCACTG CAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGC CTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCA GCTAATTTTTGTTTTTTGGTAGAGACGGGGTTTCACCATATT GGCCAGGCTGGTCTCCAACCTCCTAATCTCAGGTGATCTACCC ACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCAC TGCTCCCTTCCCTGTCTT

[0437] 表17:上调SCN1A表达的eTF的实例

表达盒	靶位点 SEQ ID NO:	DBD SEQ ID NO:	TAD SEQ ID NO:
对照	N/A	无	无
A	SEQ ID NO: 35	SEQ ID NO: 131	SEQ ID NO: 95
B	SEQ ID NO: 35	SEQ ID NO: 131	SEQ ID NO: 114
C	SEQ ID NO: 36	SEQ ID NO: 132	SEQ ID NO: 95
D	SEQ ID NO: 36	SEQ ID NO: 132	SEQ ID NO: 114
E	SEQ ID NO: 37	SEQ ID NO: 133	SEQ ID NO: 95
F	SEQ ID NO: 136	SEQ ID NO: 134	SEQ ID NO: 114
G	SEQ ID NO: 107	SEQ ID NO: 103	SEQ ID NO: 95
H	SEQ ID NO: 108	SEQ ID NO: 103	SEQ ID NO: 114

[0439] 表18:上调SCN1A表达的eTF的序列

eTF SEQ ID NO:	eTF 蛋白质序列
[0440] SEQ ID NO: 305	MAPKKKRKVGIHGVPAALEPGEKPYKCPECCKSFSRSDNLVRHQR THTGEKPYKCPECCKSFSREDNLHQRTHQTHTGEKPYKCPECCKSFS RSDNLVRHQRTHQTHTGEKPYKCPECCKSFSQSGNLTEHQRTHQTHTGEKPY KCPECCKSFSSTSGHLVRHQRTHQTHTGEKPYKCPECCKSFSQNSTLTH QRTHTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEEASGS GRADALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDA LDDFDLMLINSRSSGSPKKRKRVGSQYLPDTRHRIIEKRRKRTY

[0441]

eTF SEQ ID NO:	eTF 蛋白质序列
	ETFKSIMKKSPFSGPTDPRPPRRRIA VPSRSSASVPKPAPQYPFTSSL STINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSA LAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDE DLGALLGNSTDP AVFTDLASVDNSEFQQLNQGIPVAPHTTEPMLM EYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMD FSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPF HPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVT PEASHLLED PDEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHP PPRGHLDELTTTLESMTE DLNLDSP LTPELNEILD TFLNDECLLHAM HISTGLSIFDTSLF
SEQ ID NO: 306	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSREDNLH THQRTHTGEKPYKCPECGKSFS RSD ELVRHQR THTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPY KCPECGKSFSSTSGHLVRHQR THTGEKPYKCPECGKSFSQNSTLTEH QR THTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDD FDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM L
SEQ ID NO: 307	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSREDNLH THQRTHTGEKPYKCPECGKSFS RSD ELVRHQR THTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPY KCPECGKSFSSTSGHLVRHQR THTGEKPYKCPECGKSFSQNSTLTEH QR THTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDD FDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM L
SEQ ID NO: 308	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSREDNLH THQRTHTGEKPYKCPECGKSFS RSD ELVRHQR THTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPY KCPECGKSFSSTSGHLVRHQR THTGEKPYKCPECGKSFSQNSTLTEH QR THTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDD FDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM L
SEQ ID NO: 309	MAPKKKRKVG IHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFS REDNLH THQRTHTGEKPYKCPECGKSFSSTSHSLTEHQRTHTGEKPY KCPECGKSFSQSSSLVRHQR THTGEKPYKCPECGKSFSREDNLH TH QR THTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEEASGS GRADALDDFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LGSDA LDDFDLDM LINSRSSGSPKKRKVG SQYLPD TDDRHRIEEK RKRTY ETFKSIMKKSPFSGPTDPRPPRRRIA VPSRSSASVPKPAPQYPFTSSL STINYDEFPTMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSA LAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDE DLGALLGNSTDP AVFTDLASVDNSEFQQLNQGIPVAPHTTEPMLM EYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMD FSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPF HPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVT PEASHLLED PDEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHP PPRGHLDELTTTLESMTE DLNLDSP LTPELNEILD TFLNDECLLHAM HISTGLSIFDTSLF

[0442]

eTF SEQ ID NO:	eTF 蛋白质序列
SEQ ID NO: 310	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGEKPYKCPECGKSFS REDNLHHTHQRTHHTGEKPYKCPECGKSFSSTSHSLTEHQRTHHTGEKPY KCPECGKSFSQSSSLVRHQRTHHTGEKPYKCPECGKSFSRSDNLHHT QRTHHTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDD FDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM L
SEQ ID NO: 311	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRRDELNVHQR THTGEKPYKCPECGKSFSRSDHLTNHQRTHHTGEKPYKCPECGKSFS RSDDLVRHQRTHHTGEKPYKCPECGKSFSRSDNLVRHQRTHHTGEK PYKCPECGKSFSHRTTLTNHQRTHHTGEKPYKCPECGKSFSREDNLHT HQRTHHTGEKPYKCPECGKSFSSTSHSLTEHQRTHHTGEKPYKCPECGK SFSQSSSLVRHQRTHHTGEKPYKCPECGKSFSREDNLHHTHQRTHHTGK KTSKRPAATKKAGQAKKKKGSYPYDVPDYALEEASGSGRADALD DFDLMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM MLINSRSSGSPKKKRKVGSQYLPDTPDRHRIEEKRKRTYETFKSIM KKSPFSGPTDPRPPPRRIA VPSRSSASVPKPAPQYPFTSSLSTINYDE FPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAMVSALAQA PVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDDELGALL GNSTDPAVFTDLASVDNSEFQQLNQGIPVAPHTTEPMLMEYPEAI TRLVTGAQRPPDPAPAPLGPGLPGLNGLSGDEDFSSSIADMDFSALL GSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGS PWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASH LLEDPEETSQAVKALREMA DTVIPQKEEA AICGQMDLSHPPPRGH LDELTTTLESMTEDLNLDSP LTPELNEILD TFLNDECLLHAMHISTG LSIFDTS LF
SEQ ID NO: 312	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRRDELNVHQR THTGEKPYKCPECGKSFSRSDHLTNHQRTHHTGEKPYKCPECGKSFS RSDDLVRHQRTHHTGEKPYKCPECGKSFSRSDNLVRHQRTHHTGEK PYKCPECGKSFSHRTTLTNHQRTHHTGEKPYKCPECGKSFSREDNLHT HQRTHHTGEKPYKCPECGKSFSSTSHSLTEHQRTHHTGEKPYKCPECGK SFSQSSSLVRHQRTHHTGEKPYKCPECGKSFSREDNLHHTHQRTHHTGK KTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM L
SEQ ID NO: 313	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSGGGGGGGGSGNSSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSTTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQA FPGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPPQQQDGLGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQS QLIKPSRMRYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHRIHTGQKPFQCRICMR NFSREDNLHHTHRTHTGEKPFACDICGRKFARSDELVRHTKIHLRQK DRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSG HLVRHRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDKKAD KSVVASSATSSLSSYPSPVATSYSPSVTTSYSPATTSYSPSVPTSFS PGSSTYSPVHSGFSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSA STGLSDMTATFSPRTIEIC

[0443]

eTF SEQ ID NO:	eTF 蛋白质序列
SEQ ID NO: 314	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSGGGGGGGGGSSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFPGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPQQQGDGLGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSREDNLHTHIRHTTGEKPFACDICGRKFARSDELVRHTKIHLRQK DRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSG HLVRHIRHTTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDKKAD KSVVASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFS PGSSTYSPVHSGFSPSPVATTYSSVPPAFPAQVSSFPSAVTNSFSA STGLSDMTATFSPRTIEIC
SEQ ID NO: 315	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAE LIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS CGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRR FSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNLHTHIRHTTGEK FACDICGRKFARSDELVRHTKIHLRQKDRPYACPVESCDRRFSQSG NLTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRHTTGEKPFACDIC GRKFAQNSTLTEHTKIHLRQKDK
SEQ ID NO: 316	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAE LIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS CGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRR FSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNLHTHIRHTTGEK FACDICGRKFARSDELVRHTKIHLRQKDRPYACPVESCDRRFSQSG NLTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRHTTGEKPFACDIC GRKFAQNSTLTEHTKIHLRQKDKLEMADHMLAEGYRLVQRPPSA AAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYG AFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPG PQAPSAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPE LFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 317	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSGGGGGGGGGSSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFPGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPQQQGDGLGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHIRHTTGEK PFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVRH IRHTTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVVA SSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSFGSST

[0444]

eTF SEQ ID NO:	eTF 蛋白质序列
	YPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLS DMTATFSPRTIEIC
SEQ ID NO: 318	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAE LIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS CGGSGGGSGQSQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRR FSRSDNLVRHIRHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKP FACDICGRKFAREDNLHTHTKIHLRQKDRPYACPVESCDRRFSTSH SLTEHIRHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDIC GRKFAREDNLHTHTKIHLRQKDK
SEQ ID NO: 319	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAE LIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS CGGSGGGSGQSQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRR FSRSDNLVRHIRHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKP FACDICGRKFAREDNLHTHTKIHLRQKDRPYACPVESCDRRFSTSH SLTEHIRHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDIC GRKFAREDNLHTHTKIHLRQKDKLEMAADHMLAEGYRLVQRPPS AAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAY GAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQAPSAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVREL PELFLGQSEFDCFSDLGSAPPAGSVSC
SEQ ID NO: 320	MAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSGGGGGGGGGSSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFGSAGTALQY PPPAYPAKGGFQVPMIPDYLFPPQQQDGLGLTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRYPNRP SKTPPHERPYACPVESCDRRFSRRDELNVHIRHTGQKPFQCRICMR NFSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIHLRQ KDRPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRICMRNFSH RTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDRP YACPVESCDRRFSTSHSLTEHIRHTGQKPFQCRICMRNFSQSSSLVR HIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSS TYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGL SDMTATFSPRTIEIC
SEQ ID NO: 321	MAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSGGGGGGGGGSSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFGSAGTALQY PPPAYPAKGGFQVPMIPDYLFPPQQQDGLGLTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRYPNRP SKTPPHERPYACPVESCDRRFSRRDELNVHIRHTGQKPFQCRICMR

[0445]

eTF SEQ ID NO:	eTF 蛋白质序列
	NFSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIHRLQ KDRPYACPVESCDRRFSRSDNLVRHRIHTGQKPFQCRICMRNFSH RTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQKDRP YACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVR HIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSS TYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFSSAVTNSFSASTGL SDMTATFSPRTIEIC
SEQ ID NO: 322	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSSGGGGGGGGGSSNSSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPPQQQDGLGLTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSOLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRRDELNVHRIHTGQKPFQCRICMR NFSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIHRLQ KDRPYACPVESCDRRFSRSDNLVRHRIHTGQKPFQCRICMRNFSH RTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQKDRP YACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRICMRNFSQSSSLVR HIRTHTGEKPFACDICGRKFAREDNLHTHTKIHRLQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSS TYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFSSAVTNSFSASTGL SDMTATFSPRTIEIC
SEQ ID NO: 323	MTGKLAELPVTMSSLLNQLPDNLYPEEIPSAALNLFSGSSDSVVHY NQMATENVMDIGLTNEKNPELSYSGSFQAPGNKTVTYLGKFAF DSPSNWCQDNIISLMSAGILGVPPASGALSTQTSTASMVQPPQGDV EAMYPALPPYSNCGDLYSEPVSFHDPPQGNPGLAYSPQDYQSAKPA LDSNLFPMIPDYNLYHHPNDMGSIPEHKPFQGMPIRVNPPPITPLE TIKAFKDKQIHGPGFSLPQPPLTLKPIRPRKYPNRPSKTPLHERPHAC PAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRSFREDNLHTHI RHTHTGEKPFACEFCGRKFARSDDELVRHAKIHLKQKEHACPAEGCD RRFSQSGNLTEHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGE KPFACEFCGRKFAQNSTLTEHAKIHLKQKEKKAEGGAPSASSAPP VSLAPVVTTC
SEQ ID NO: 324	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSG LRPRGAPLGPPPPRQPGALAYGAFGPPSSFQFPFVPPPAAGIAHLQ PVATPYPGRAAAPPNAPGGPPGPQAPSAAPPPPAHALGGMDAE LIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVS CQSQLIKPSRMRKYPNRPSKTTPPHERPYACPVESCDRRFSRSDNLV RHRIHTGQKPFQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGR KFAREDNLHTHTKIHRLQKDRPYACPVESCDRRFSTSHSLTEHIRI TGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFARED NLHTHTKIHRLQKDK
SEQ ID NO: 325	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSSGGGGGGGGGSSNSSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS

[0446]

eTF SEQ ID NO:	eTF 蛋白质序列
	PPLSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFFGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPPQQGDLGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRICMR NFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHIRTHTGEK PFACDICGRKFSTSHSLTEHIRHTGQKPFQCRICMRNFSQSSSLVRH IRTHTGEKPFACDICGRKFAREDNLHHTKIHLRQKDKKADKSVVA SSATSSLSSYPSPVATSYSPVTTTSYPSPATTSYSPSPVPTSFSSPGSST YSPVHSGFSPSPSVATTYSSVPPAFPAQVSSFPSAVTNSFSASTGLS DMTATFSPRTIEIC
SEQ ID NO: 365	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGEKPYKCPECGKSFS REDNLHHTHQRTHHTGEKPYKCPECGKSFSSTSHSLTEHQRTHHTGEKPY KCPECGKSFSQSSSLVRHQRTHHTGEKPYKCPECGKSFSREDNLHHT QRTHHTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEDALDD FDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM L
SEQ ID NO: 366	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRSDNLVRHQR THTGEKPYKCPECGKSFSHRTTLTNHQRTHHTGEKPYKCPECGKSFS REDNLHHTHQRTHHTGEKPYKCPECGKSFSSTSHSLTEHQRTHHTGEKPY KCPECGKSFSQSSSLVRHQRTHHTGEKPYKCPECGKSFSREDNLHHT QRTHHTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALEEASGS GRADALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLINRSGSPKKRKVGVSQYLPDTPDRHRIEKKRRTY ETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQYPFTSSL STINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAPAMVSA LAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDE DLGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLM EYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMD FSALLGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPF HPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVT PEASHLLEDPEETSQAVKALREMA DTVIPQKEEAICGQMDLSHP PPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDFTLNDECLLHAM HISTGLSIFDTSLF
SEQ ID NO: 364	MYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSGSGSAAAKAE MQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLGAA GAPEGSGSNSSSSSSGGGGGGGGGNSSSSSSTFNPQADTGEQPYE HLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSG NTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCA VPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFFGSAGTALQYPPPAYP AAKGGFQVPMIPDYLFPPQQGDLGLGTPDQKPFQGLSRTQQPSLT PLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSKTPPH ERPYPACPVESCDRRFSRSDNLVRHIRHTGQKPFQCRICMRNFSHRT TLTNHIRTHTGEKPFACDICGRKFAREDNLHHTKIHLRQKDRPYA CPVESCDRRFSTSHSLTEHIRHTGQKPFQCRICMRNFSQSSSLVRHI RTHHTGEKPFACDICGRKFAREDNLHHTKIHLRQKDKKADKSVVA SSATSSLSSYPSPVATSYSPVTTTSYPSPATTSYSPSPVPTSFSSPGSST YSPVHSGFSPSPSVATTYSSVPPAFPAQVSSFPSAVTNSFSASTGLS

[0447]

eTF SEQ ID NO:	eTF 蛋白质序列
	DMTATFSPRTIEIC
SEQ ID NO: 295	MYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAGSGSGSAAAKAE MQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGAPQFLGAA GAPEGSGSNSSSSSSSGGGGGGGGGSNSSSSSTFNPQADTGEQPYE HLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSG NTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCA VPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFPGSAGTALQYPPPAYP AAKGGFQVPMIPDYLFPQQQGDGLGTPDQKPFQGLSRTQQPSLT PLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRPSTPPH ERPYPACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRED NLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLRQKDRPYA CPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSSTSGHLVRH IRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDKKADKSVVA SSATSSLSSYSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSST YSPVHSGFSPSPVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLS DMTATFSPRTIEIC
SEQ ID NO: 296	MAAACAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSSGGGGGGGGGSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFPGSAGTALQY PPPAYPAKGGFQVPMIPDYLFPQQQGDGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSDPGALVRHIRIHTGQKPFQCRICMR NFSRSDNLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKIHLRQ KDRPYACPVESCDRRFSTHDLIRHIRIHTGQKPFQCRICMRNFSST GNLVRHIRTHTGEKPFACDICGRKFARSNDNLVRHTKIHLRQKDRPY ACPVESCDRRFSQSGHLTEHIRIHTGQKPFQCRICMRNFSERSHLRE HIRTHTGEKPFACDICGRKFAQAGHLASHTKIHLRQKDKKADKSV VASSATSSLSSYSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGS STYSPVHSGFSPSPVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTG LSDMTATFSPRTIEIC
SEQ ID NO: 297	MRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQKP FQCRICMRNFSHSTLTNHIRTHTGEKPFACDICGRKFARSNDNRKTH IRHTHTGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICMRNFS QSSSLTRHIRTHTGEKPFACDICGRKFARSNDNRKTHTKIHLRQKDK KADKSVVAS
SEQ ID NO: 298	MRKYPNRPSTPPHERPYACPVESCDRRFSRSDNLTRHIRIHTGQKP FQCRICMRNFSRSDNLTHIRTHTGEKPFACDICGRKFARSDEKRRH IRHTHTGEKPFACDICGRKFSSQSGNLTEHIRIHTGQKPFQCRICMRNFS TSGHLTRHIRTHTGEKPFACDICGRKFAQSSTRKEHTKIHLRQKDK KADKSVVAS
SEQ ID NO: 299	MAAACAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSSGGGGGGGGGSNSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFPGSAGTALQY PPPAYPAKGGFQVPMIPDYLFPQQQGDGLGTPDQKPFQGLSRT

eTF SEQ ID NO:	eTF 蛋白质序列
	QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSREDNLHTHIRHTHTGEKPFACDICGRKFARSDELVRHIRTHTGEK PFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRICMRNFSSTSGHLVR HIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHRLQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSS TYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFSSAVTNSFSASTGL SDMTATFSPRTIEIC
[0448] SEQ ID NO: 440	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMMLLSNGA PQFLGAAGAPEGSGSNSSSSSSSGGGGGGGGGSSSSSSSTFNPQADT GEQPYEHLTAESFPDISLNNEKVLVETSYPSQTTRLPPITYTGRFSLE PAPNSGNTLWPEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQS PPLSCAVPSNDSSPIYSAAPTFTPTNTDIFPEPQSQAFFGSAGTALQY PPPAYPAAKGGFQVPMIPDYLFPPQQQGDGLGTPDQKPFQGLSRT QQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLIKPSRMRKYPNRP SKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMR NFSREDNLHTHIRHTHTGEKPFACDICGRKFARSDELVRHIRTHTGEK PFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRICMRNFSSTSGHLVR HIRHTHTGEKPFACDICGRKFAQNSTLTEHTKIHRLQKDKKADKSVV ASSATSSLSSYPSPVATSYSPVTTSYSPATTSYSPVPTSFSSPGSS TYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFSSAVTNSFSASTGL SDMTATFSPRTIEIC

[0449] 表19:本文公开的聚A序列

	聚 A 序列
[0450] SEQ ID NO: 326 spA (合成的聚 A)	AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTT TTGTGTGCGGACCGCACGTG
SEQ ID NO: 327 hGH (人生长激素聚 A)	GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGC CCTGGAAGTTGCCACTCCAGTGCCCAACCAGCCTTGTCCTAA TAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCT ATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGG GCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTAT TGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCA CTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCT CAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAG GCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCA CCATATTGGCCAGGCTGGTCTCCAATCCTAATCTCAGGTG ATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGC GTGAACCACTGCTCCCTTCCCTGTCCTT

[0451] 表20:用于调节内源GRN基因的表达的锌指的实例

	氨基酸序列		氨基酸序列
[0452] SEQ ID NO: 141	RNDTLTE	SEQ ID NO: 153	RSDKLVR

[0453]

SEQ ID NO: 142	RSDNLVR	SEQ ID NO: 154	QLAHLRA
SEQ ID NO: 143	RSDHLTT	SEQ ID NO: 155	ERSHLRE
SEQ ID NO: 144	SPADLTR	SEQ ID NO: 156	TTGNLTV
SEQ ID NO: 145	TSHSLTE	SEQ ID NO: 157	QRANLRA
SEQ ID NO: 146	QSGDLRR	SEQ ID NO: 158	SKKALTE
SEQ ID NO: 147	DPGALVR	SEQ ID NO: 159	REDNLHT
SEQ ID NO: 148	DPGHLVR	SEQ ID NO: 160	TKNSLTE
SEQ ID NO: 149	RSDELVR	SEQ ID NO: 161	RKDNLKN
SEQ ID NO: 150	DSGNLRV	SEQ ID NO: 162	QSSNLVR
SEQ ID NO: 151	HKNALQN	SEQ ID NO: 163	QSSSLVR
SEQ ID NO: 152	TSGELVR	SEQ ID NO: 164	QAGHLAS

[0454]

表21: 调节GRN表达的eTF的DBD的实例, 其包含多个ZF

[0455]

	氨基酸序列
SEQ ID NO: 165	LEPGEKPYKCPECGKSFSRNDTLTEHQRTHTGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSSTSGELVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSSTSGELVRHQRTHTGEKPYKCPECGKSFSSTKNSLTEHQRTHTGKKTS
SEQ ID NO: 166	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSRSDHLTTHQRTHTGEKPYKCPECGKSFSRSDDELVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTGEKPYKCPECGKSFSSTTGNLTVHQRTHTGKKTS
SEQ ID NO: 167	LEPGEKPYKCPECGKSFSRSDHLTTHQRTHTGEKPYKCPECGKSFSDDELVRHQRTHTGEKPYKCPECGKSFSRSDKLVRHQRTHTGEKPYKCPECGKSFSSTTGNLTVHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSSTKNSLTEHQRTHTGKKTS
SEQ ID NO: 168	LEPGEKPYKCPECGKSFSFSPADLTRHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS
SEQ ID NO: 169	LEPGEKPYKCPECGKSFSSTSHSLTEHQRTHTGEKPYKCPECGKSFHKNALQNHQRTHTGEKPYKCPECGKSFSERSHLREHQRTHTGEKPYKCPECGKSFSSTKNSLTEHQRTHTGEKPYKCPECGKSFSQANLRAHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSQSSNLVRHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTS
SEQ ID NO: 170	LEPGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDSGNLRVHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS
SEQ ID NO: 171	RNDTLTE \times DPGALVR \times TSGELVR \times RSDNLVR \times TSGELVR \times TKNSLTE (其中X表示包含1-50个氨基酸残基的连接体)
SEQ ID NO: 412	RSDNLVR \times DPGHLVR \times RSDHLTT \times RSDELVR \times RSDKLVR \times TTGNLTV

[0456]

	氨基酸序列
	(其中X表示包含1-50个氨基酸残基的连接体)
SEQ ID NO: 413	RSDHLTT x RSDLVLR x RSDKLVR x TTGNLTV x QLAHLRA x TKNSLTE (其中X表示包含1-50个氨基酸残基的连接体)
SEQ ID NO: 414	SPADLTR x DSGNLRV x QLAHLRA x QRANLRA x REDNLHT x RSDNLVR (其中X表示包含1-50个氨基酸残基的连接体)
SEQ ID NO: 415	TSHSLTE x HKNALQN x ERSHLRE x SKKALTE x QRANLRA x RKDNLKN x QSSNLVR x QSSSLVR x QAGHLAS (其中X表示包含1-50个氨基酸残基的连接体)
SEQ ID NO: 416	QSGDLRR x SPADLTR x DSGNLRV x QLAHLRA x QRANLRA x REDNLHT x RSDNLVR (其中X表示包含1-50个氨基酸残基的连接体)

[0457] 表22: 用于调节GRN的表达的靶位点序列的实例

[0458]

靶位点 SEQ ID NO:	识别靶位点的 ZF 或 gRNA	靶位点序列	染色体 17 起始坐标
SEQ ID NO: 38	ZF	GAGTAGAAAAGAAACACA	44345153
SEQ ID NO: 330	ZF	CGCACTGTCAATGCCCA	44344963
SEQ ID NO: 331	ZF	CCTGCTGAGGCTGTCCCG	44345058
SEQ ID NO: 332	ZF	AATGGGGTGTGGGGCGAG	44345111
SEQ ID NO: 333	ZF	CCTAGAAATGGGGTGTGG	44345105
SEQ ID NO: 334	ZF	TGAGTAGAAAAGAAACACAGCA TTCCA	44345152
SEQ ID NO: 335	ZF	GAGTAGAAAAGAAACACAGCA	44345153
SEQ ID NO: 336	ZF	AGTAGAAAAGAAACACAGCAT	44345154
SEQ ID NO: 113	gRNA	CATGATCCCTAGAAATGGGGT	44345098

[0459] 表23: 调节GRN的表达的eTF的实例

[0460]

表达盒	靶位点SEQ ID NO:	DBD SEQ ID NO:	TAD SEQ ID NO:	调节元件 SEQ ID NO:
对照	N/A	无	无	SEQ ID NO: 178
A	SEQ ID NO: 331	SEQ ID NO: 165	SEQ ID NO: 95	SEQ ID NO: 178
B	SEQ ID NO: 331	SEQ ID NO: 165	SEQ ID NO: 114	SEQ ID NO: 178
C	SEQ ID NO: 332	SEQ ID NO: 166	SEQ ID NO: 95	SEQ ID NO: 178
D	SEQ ID NO: 333	SEQ ID NO: 167	SEQ ID NO: 95	SEQ ID NO: 178
E	SEQ ID NO: 333	SEQ ID NO: 167	SEQ ID NO: 114	SEQ ID NO: 178
F	SEQ ID NO: 38	SEQ ID NO: 168	SEQ ID NO: 95	SEQ ID NO: 178
G	SEQ ID NO: 38	SEQ ID NO: 168	SEQ ID NO: 114	SEQ ID NO: 178
H	SEQ ID NO: 334	SEQ ID NO: 169	SEQ ID NO: 114	SEQ ID NO: 178
I	SEQ ID NO: 335	SEQ ID NO: 170	SEQ ID NO: 114	SEQ ID NO: 179
J (K+eGFP)	SEQ ID NO: 38	SEQ ID NO: 168	SEQ ID NO: 95	SEQ ID NO: 179
K	SEQ ID NO: 38	SEQ ID NO: 168	SEQ ID NO: 95	SEQ ID NO: 179
L	SEQ ID NO: 38	SEQ ID NO: 168	SEQ ID NO: 114	SEQ ID NO: 179
M	SEQ ID NO: 113	SEQ ID NO: 112	SEQ ID NO: 95	SEQ ID NO: 178

[0461] 表24:本文公开的调节元件 (RE) 的实例

[0462]

	RE 的类型	序列
SEQ ID NO: 178	内含子的	GTGTGTATGCTCAGGGGCTGGGAAAGGAGGGGAGGGAGCTCCGGCTCAG
SEQ ID NO: 179	CBA 启动子	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCTCCCCACCCCCAATTTTGTATTTATTTATTTTAAATT

[0463]

	RE 的类型	序列
		ATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGCGC GCGCCAGGCGGGGCGGGGCGGGGCGAGGGGCGGGGCG GGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAG CGGCGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGCG GCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGGCGG CGGGAGTCGCTGCGTTGCCTTCGCCCCGTGCCCGCTCC GCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCG CGTTACTCCACAGGTGAGCGGGCGGGACGGCCCTTCTC CTCCGGGCTGTAATTAGCGCTTGTTTAAATGACGGCTCG TTTCTTTTCTGTGGCTGCGTGAAAGCCTTAAAGGGCTCC GGGAGGGCCCTTTGTGCGGGGGGAGCGGCTCGGGGGG TCGTGCGTGTGTGTGTGCGTGGGGAGCGCCGCGTGCG GCCCGCGCTGCCCGGCGGCTGTGAGCGCTGCGGGCGCG GCGCGGGGCTTTGTGCGCTCCGCGTGTGCGCGAGGGGA GCGCGGCCGGGGGCGGTGCCCCGCGGTGCGGGGGGGCT GCGAGGGGAACAAAGGCTGCGTGCGGGGTGTGTGCGTG GGGGGGTGAGCAGGGGGTGTGGGCGCGGCGGTCTGGGCT GTAACCCCCCTGCACCCCCCTCCCGAGTTGCTGAGC ACGGCCCGGCTTCGGGTGCGGGGCTCCGTGCGGGGCGT GCGCGGGGCTCGCCGTGCCGGGCGGGGGGTGGCGGCA GGTGGGGGTGCCGGGCGGGGCGGGGCCGCTCGGGCCG GGGAGGGCTCGGGGGAGGGGCGCGGCGGCCCGGAGC GCCGGCGGCTGTGAGGCGCGGCGAGCCGAGCCATTG CCTTTTATGGTAATCGTGCGAGAGGGCGCAGGGACTTCC TTTGTCCCAAATCTGGCGGAGCCGAAATCTGGGAGGCG CCGCCGCACCCCCTCTAGCGGGCGCGGGCGAAGCGGTG CGGCGCCGCGCAGGAAGGAAATGGGCGGGGAGGGCCTTC GTGCGTCGCCGCGCCGCGCTCCCTTCTCCATCTCCAGC CTCGGGGCTGCCGAGGGGGACGGCTGCCTTCGGGGGG GACGGGGCAGGGCGGGGTTCGGCTTCTGGCGTGTGACC GCGGGCTCTAGAGCCTCTGCTAACCATGTTTCATGCCTTC TTCTTTTCTCTACAGCTCCTGGGCAACGTGCTGGTTGTTG TGCTGTCTCATCATTTTGGCAAAGAATT
SEQ ID NO: 182	内含子的	GTAAGGTAAGAATTGAATTTCTCAGTTGAAGGATGCTTACACTCTTGTCCATCTAG
SEQ ID NO: 183	小白蛋白 (PV)-选择性	GGAGGAAGCCATCAACTAAACTACAATGACTGTAAGAT ACAAAATTGGGAATGGTAACATATTTTGAAGTTCTGTTG ACATAAAGAATCATGATATTAATGCCCATGGAAATGAA AGGGCGATCAACACTATGGTTTGAAAAGGGGGAAATTG TAGAGCACAGATGTGTTTCGTGTGGCAGTGTGCTGTCTCT AGCAATACTCAGAGAAGAGAGAGAACAATGAAATTCTG ATTGGCCCCAGTGTGAGCCCAGATGAGGTTACAGCTGCC AACTTTCTCTTTCACATCTTATGAAAGTCATTTAAGCAC AACTAACTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTC TTGCTCTGTTGCCAGGACAGAGTGCAGTAGTGACTCAA TCTCGGCTCACTGCAGCCTCCACCTCCTAGGCTCAAACG GTCCTCCTGCATCAGCCTCCCAAGTAGCTGGAATTACAG GAGTGGCCCACCATGCCAGCTAATTTTTGTATTTTAA

[0464]

	RE 的类型	序列
		TAGATACGGGGGTTTCACCATATCACCCAGGCTGGTCTC GAACTCCTGGCCTCAAGTGATCCACCTGCCTCGGCCTCC CAAAGTGCTGGGATTATAGGCGTCAGCCACTATGCCCA ACCCGACCAACCTTTTTTAAAATAAATATTTAAAAAATT GGTATTTACATATATACTAGTATTTACATTTATCCACA CAAAACGGACGGGCCTCCGCTGAACCAGTGAGGCCCA GACGTGCGCATAAATAACCCCTGCGTGCTGCACCACCTG GGGAGAGGGGGAGGACCACGGTAAATGGAGCGAGCGC ATAGCAAAAGGGACGCGGGGTCTTTTTCTCTGCCGGTG GCACTGGGTAGCTGTGGCCAGGTGTGGTACTTTGATGGG GCCCAGGGCTGGAGCTCAAGGAAGCGTCGAGGGTCAC AGATCTGGGGGAACCCCGGGGAAAAGCACTGAGGCAA AACCGCCGCTCGTCTCCTACAATATATGGGAGGGGGAG GTTGAGTACGTTCTGGATTACTCATAAGACCTTTTTTTTT TCCTTCCGGGCGCAAAACCGTGAGCTGGATTTATAATCG CCTATAAAGCTCCAGAGGCGGTGAGGCACCTGCAGAG GAGCCCCGCCGCTCCGCCGACTAGCTGCCCCCGCGAGC AACGGCCTCGTGATTTCCCCGCCGATCCGGTCCCCGCCT CCCCCTCTGCCCCCGCCTACCCCGGAGCCGTGCAGCCG CCTCTCCGAATCTCTCTCTTCTCCTGGCGCTCGCGTGCGA GAGGGAAGTAGCGAGAACGAGGAAGCAGCTGGAGGTG ACGCCGGGCAGATTACGCCTGTCAGGGCCGAGCCGAGC GGATCGCTGGGCGCTGTGCAGAGGAAAGGCGGGAGTGC CCGGCTCGCTGTGCGAGAGCCGAGGTGGGTAAGCTAGC GACCACCTGGACTTCCCAGCGCCCAACCGTGCTTTTCA GCCAGGTCTCTCCTCCCGCGGCTTCTCAACCAACCCCA TCCCAGCGCCGGCCACCCAACCTCCCGAAATGAGTGCTT CCTGCCCCAGCAGCCGAAGGCGCTACTAGGAACGGTAA CCTGTTACTTTTCCAGGGGGCCGTAGTCGACCCGCTGCCC GAGTTGCTGTGCGACTGCGCGCGCGGGGCTAGAGTGCA AGGTGACTGTGGTTCTTCTCTGGCCAAGTCCGAGGGAGA ACGTAAAGATATGGGCCTTTTTCCCCCTCTCACCTTGTCT CACCAAAGTCCCTAGTCCCCGGAGCAGTTAGCCTCTTTC TTTCCAGGGAATTAGCCAGACACAACAACGGGAACCAG ACACCGAACCAGACATGCCCCGCCCGTGCGCCCTCCCC GCTCGCTGCCTTTCCTCCCTCTTGTCTCTCCAGAGCCGGA TCTTCAAGGGGAGCCTCCGTGCCCCCGGCTGCTCAGTCC CTCCGGTGTGCAGGACCCCGGAAGTCCTCCCCGCACAG CTCTCGCTTCTCTTTGCAGCCTGTTTCTGCGCCGGACCAG TCGAGGACTCTGGACAGTAGAGGCCCCGGGACGACCGA GCTG
SEQ ID NO: 184	小白蛋白 (PV)-选择 性	TCAACAGGGGGGACACTTGGGAAAGAAGGATGGGGACA GAGCCGAGAGGACTGTTACACATTAGAGAAACATCAGT GACTGTGCCAGCTTTGGGGTAGACTGCACAAAAGCCCT GAGGCAGCACAGGCAGGATCCAGTCTGCTGGTCCCAGG AAGCTAACCGTCTCAGACAGAGCACAAAGCACCCGAGAC ATGTGCCACAAGGCTTGTGTAGAGAGGTCAGAGGACAG CGTACAGGTCCCAGAGATCAAACCTCAACCTCACCAGGC

[0465]

	RE 的类型	序列
		TTGGCAGCAAGCCTTTACCAACCCACCCCCACCCCACCC ACCCTGCACGCGCCCCTCTCCCCTCCCCATGGTCTCCCA TGGCTATCTCACTTGGCCCTAAAATGTTTAAGGATGACA CTGGCTGCTGAGTGGAATGAGACAGCAGAAGTCAACA GTAGATTTTAGGAAAGCCAGAGAAAAAGGCTTGTGCTG TTTTAGAAAGCCAAGGGACAAGCTAAGATAGGGCCCA AGTAATGCTAGTATTTACATTTATCCACACAAAACGGAC GGGCCTCCGCTGAACCAGTGAGGCCCCAGACGTGCGCA TAAATAACCCCTGCGTGCTGCACCACCTGGGGAGAGGG GGAGGACCACGGTAAATGGAGCGAGCGCATAGCAAAA GGGACGCGGGGTCTTTTCTCTGCCGGTGGCACTGGGTA GCTGTGGCCAGGTGTGGTACTTTGATGGGGCCCAGGGCT GGAGCTCAAGGAAGCGTCGCAGGGTACAGATCTGGGG GAACCCCGGGGAAAAGCACTGAGGCAAAACCGCCGCTC GTCTCCTACAATATATGGGAGGGGGAGGTTGAGTACGT TCTGGATTACTCATAAGACCTTTTTTTTTTCTCCGGGC GCAAAACCGTGAGCTGGATTATAATCGCCCTATAAAG CTCCAGAGGCGGTCAGGCACCTGCAGAGGAGCCCCGCC GCTCCGCCGACTAGCTGCCCCCGCAGCAACGGCCTCGT GATTTCCCCGCCGATCCGGTCCCCGCCTCCCCACTCTGC CCCCCGCTACCCCGGAGCCGTGCAGCCGCTCTCCGAAT CTCTCTCTCTCTCTGGCGCTCGCGTGCGAGAGGGAATA GCGAGAACGAGGAAGCAGCTGGAGGTGACGCCGGGCA GATTACGCCTGTGAGGGCCGAGCCGAGCGGATCGCTGG GCGCTGTGCAGAGGAAAGGCGGGAGTGCCCGGCTCGCT GTCGCAGAGCCGAGGTGGGTAAGCTAGCGACCACCTGG ACTTCCCAGCGCCCAACCGTGGCTTTTCAGCCAGGTCCT CTCCTCCCGCGGCTTCTCAACCAACCCCATCCCAGCGCC GGCCACCCAACCTCCCGAAATGAGTGCTTCTGCCCCAG CAGCCGAAGGCGCTACTAGGAACGGTAACCTGTTACTTT TCCAGGGGCCGTAGTCGACCCGCTGCCCCAGTTGCTGTG CGACTGCGCGCGCGGGGCTAGAGTGCAAGGTGACTGTG GTTCTTCTCTGGCCAAGTCCGAGGGAGAACGTAAAGAT ATGGGCCTTTTTCCCCCTCTCACCTTGTCTACCAAAGTC CCTAGTCCCCGGAGCAGTTAGCCTCTTTCTTTCCAGGGA ATTAGCCAGACACAACAACGGGAACCAGACACCGAACC AGACATGCCCGCCCCGTGCGCCCTCCCCGCTCGCTGCCT TTCCTCCCTCTTGTCTCTCCAGAGCCGGATCTTCAAGGG GAGCCTCCGTGCCCCCGGCTGCTCAGTCCCTCCGGTGTG CAGGACCCCGGAAGTCCTCCCCGCACAGCTCTCGTTCT CTTTGCAGCCTGTTTCTGCGCCGGACCAGTCGAGGACTC TGGACAGTAGAGGCCCCCGGACGACCGAGCTG
SEQ ID NO: 185	小白蛋白 (PV)-选择 性	GGAGGAAGCCATCAACTAACTACAATGACTGTAAGAT ACAAAATTGGGAATGGTAACATATTTTGAAGTTCTGTTG ACATAAAGAATCATGATATTAATGCCCATGGAAATGAA AGGGCGATCAACACTATGGTTTGAAGGGGGAAATTG TAGAGCACAGATGTGTTCTGTGGCAGTGTGCTGTCTCT AGCAATACTCAGAGAAGAGAGAGAACAATGAAATTCTG

[0466]

	RE 的类型	序列
		ATTGGCCCCAGTGTGAGCCCAGATGAGGTTTCAGCTGCC AACTTTCTCTTTACATCTTATGAAAGTCATTTAAGCAC AACTAACTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGTC TTGCTCTGTTGCCCAGGACAGAGTGCAGTAGTGACTION TCTCGGCTCACTGCAGCCTCCACCTCCTAGGCTCAAACG GTCCTCCTGCATCAGCCTCCCAAGTAGCTGGAATTACAG GAGTGGCCCCACCATGCCCAGCTAATTTTTGTATTTTTAA TAGATACGGGGGTTTCACCATATCACCCAGGCTGGTCTC GAACTCCTGGCCTCAAGTGATCCACCTGCCTCGGCCTCC CAAAGTGCTGGGATTATAGGCGTCAGCCACTATGCCCA ACCCGACCAACCTTTTTTAAAATAAATATTTAAAAAATT GGTATTTACATATATACTAGT
SEQ ID NO: 417		GCCCTCTAGGCCACCTGACCAGGTCCCCTCAGTCCCCC CTCCCACACTCCCACACTCAGCCCCCTCCCCCCCCC CGACCCCTGCAGGATTATCCTGTCTGTGTTCTGACTCA GCCTGGGAGCCACCTGGGCAGCAGGGGCCAAGGGTGTC CTAGAAGGGACCTGGAGTCCACGCTGGGCCAAGCCTGC CCTTTCTCCCTCTGTCTTCCGTCCCTGCTTGCGGTTCTGC TGAATGTGGTTATTTCTCTGGCTCCTTTTACAGAGAATG CTGCTGCTAATTTTATGTGGAGCTCTGAGGCAGTGTAAT TGAAGCCAGACACCCTGTCAGCAGTGGGCTCCCGTCTT GAGCTGCCATGCTTCCTGCTCTCCTCCCGTCCCGGCTCCT CATTTTCATGCAGCCACCTGTCCAGGGAGAGAGGAGTC ACCCAGGCCCTCAGTCCGCCCTTAAATAAGAAAGCCT CCGTTGCTCGGCACACATACCAAGCAGCCGCTGGTGCA ATCT

[0467] 表25:本文公开的额外的序列

[0468]

	序列
SEQ ID NO: 112 (用于调 节 GRN 的 dSaCas9 序列)	KRNYILGLAIGITSVGYGIIIDYETRDVIDAGVRLFKEANVENNEGRRSKR GARRLKRRRRHRIQVRVKLLFDYNLLTDHSELSGINPYEARVKGLSQKLS EEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYV AELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVKAYHQLDQSFID TYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVK YAYNADLYNALNDLNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLK QIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAELLD QIAKILTIYQSSEDIQEELTNLSELQEEIEQISNLKGYTGTHNLSLKAINL ILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRS FIQSIKVINAIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERI EEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNPNFNYEVD HIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHI LNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLM NLLRSYFRVNNLDVKVKSINGGFTSFLRRKWFKKERNKGYKHHAEDA LIANADFIFKEWKLDKAKKVMENQMFEKQAESMPEIETEQEYKEIFI TPHQIKHIKDFKDYKYSHRVDKKNRELINDTLYSTRKDDKGNTLIVNNL NGLYDKDNDKLKLINKSPEKLLMYHHPQTYQKLKLIMEQYGDEKNP

[0469]

	序列
	LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNHLITDDYPNSRN KVVKLSLKPYPYFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEA KKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITY REYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKK G
SEQ ID NO: 186 EF1a 启动子	GAGTAATTCATACAAAAGGACTCGCCCCTGCCTTGGGGAATCCCAGG GACCGTCGTAAACTCCCCTAACGTAGAACCCAGAGATCGCTGCGT TCCCGCCCCCTCACCCGCCGCTCTCGTCATCACTGAGGTGGAGAAGA GCATGCGTGAGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGC CCACAGTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGT GCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTA CTGGCTCCGCCTTTTTCCCGAGGGTGGGGGAGAACCCTATATAAGTG CAGTAGTCGCCGTGAACGTTCTTTTTTCGCAACGGGTTTGCCGCCAGAA CACAGGTAAGTGCCGTGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGG GTTATGGCCCTTGCGTGCCTTGAATTACTTCCACGCCCTGGCTGCAG TACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAG TTCGAGGCCTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGA GGCCTGGCTTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACC TTCGCGCCTGTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTT TTGATGACCTGCTGCGACGCTTTTTTCTGGCAAGATAGTCTTGTA TGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGCGGG CGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTTCGGCGAGGCGGG GCCTGCGAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCT GGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCG CCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAAGTTGCGTGAGCGGA AAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGGAGGA CGCGGCGCTCGGGAGAGCGGGCGGGTGAGTACCCACACAAAGGAA AAGGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTA CCGGGCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTAC GTCGTCTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCA CACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTGGCACTTGATGT AATTCTCCTTGGAATTTGCCCTTTTTGAGTTTGGATCTTGGTTCATTCT CAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTCAGGTGTC GTGA
SEQ ID NO: 187	GAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCC TATCAGTGATAGAGAACGATGTCGAGTTTACTCCCTATCAGTGATAG AGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTC GAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCT ATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAG AGAACGTATGTCGAGGTAGGCGGTACGGTGGGAGGCCTATATAAGC AGAGCTCGTTTAGTGAACCGTCAGATCGCCGAATTCGCCACCATGGT GAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCG AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTG CACCACCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCT GACCTACGGCGTGCAAGTCTCAGCCGCTACCCCGACCACATGAAGC AGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA GGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGG

[0470]

	序列
	GCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAG TACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA GAACGGCATCAAGGTGAAC TTCAAGATCCGCCACAACATCGAGGACG GCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGC GACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTC CGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGC TGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTG TACAAGTCCGGACTCAGATCTCGAGAGGAGGAGGAGGAGACAGACA GCAGGATGCCCCACCTCGACAGCCCCGGCAGCTCCCAGCCGAGACGC TCCTTCCTCTCAAGGGTGATCAGGGCAGCTCTACCGTTGCAGCTGCTT CTGCTGCTGCTGCTGCTCCTGGCCTGCCTGCTACCTGCCTCTGAAGAT GACTACAGCTGCACCCAGGCCAACAAC TTTGCCCCGATCCTTCTACCCC ATGCTGCGGTACACCAACGGGGCCACCTCCCACCTAGACTAGTaataaaaga tctttatttcattagatctgtgtgtgtgtttttgtgtg
SEQ ID NO: 188	GAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTACTCCC TATCAGTGATAGAGAACGATGTCGAGTTTACTCCCTATCAGTGATAG AGAACGTATGTCGAGTTTACTCCCTATCAGTGATAGAGAACGTATGTC GAGTTTACTCCCTATCAGTGATAGAGAACGTATGTCGAGTTTATCCCT ATCAGTGATAGAGAACGTATGTCGAGTTTACTCCCTATCAGTGATAG AGAACGTATGTCGAGGTAGGCGTGACGGTGGGAGGCCTATATAAGC AGAGCTCGTTTAGTGAACCGTCAGATCGCC
SEQ ID NO: 418	GCCACCATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCC CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG TGTCGGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTG AAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCCACCCTC GTGACCACCCTGACCTACGGCGTGCAAGTTCAGCCGCTACCCCGA CCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCT ACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAG ACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC ACAAGCTGGAGTACAAC TACAACAGCCACAACGTCTATATCATGGCC GACAAGCAGAAGAACGGCATCAAGGTGAAC TTCAAGATCCGCCACA ACATCGAGGACGGCAGCGTGCAAGTTCGAGGACCACTACCAGCAGAAC ACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCT GAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATC ACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGC ATGGACGAGCTGTACAAGTCCGGACTCAGATCTCGAGAGGAGGAGG AGGAGACAGACAGCAGGATGCCCCACCTCGACAGCCCCGGCAGCTCC CAGCCGAGACGCTCCTTCTCTCAAGGGTGATCAGGGCAGCTCTACC GTTGCAGCTGCTTCTGCTGCTGCTGCTCCTGGCCTGCCTGCTACCT GCCTCTGAAGATGACTACAGCTGCACCCAGGCCAACAAC TTTGCCCCG ATCCTTCTACCCCATGCTGCGGTACACCAACGGGGCCACCTCCCACCTA G
SEQ ID NO: 365	MAPKKKRKVGIHGVPAALEPGEKPYKCPECCKSFSRSDNLVRHQRTHTG EKPYKCPECCKSFSRHTLTNHRQRTHTGEKPYKCPECCKSFSREDNLH QRTHTGEKPYKCPECCKSFSSTSHSLTEHQRTHTGEKPYKCPECCKSFS SSLVRHQRTHTGEKPYKCPECCKSFSREDNLHHTHQRTHTGKKTSKRPAA TKKAGQAKKKKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDM LGSDALDDFDLDMLGSDALDDFDLDM

[0471]

	序列
SEQ ID NO: 366	MAPKKKRKVGIHGVPAALEPGEKPYKCPECCKSFSRSDNLVRHQRTHTG EKPYKCPECCKSFSHRTTLTNHQRTHTGEKPYKCPECCKSFSREDNLHTH QRTHTGEKPYKCPECCKSFSSTSHSLTEHQRTHTGEKPYKCPECCKSFSQS SSLVRHQRTHTGEKPYKCPECCKSFSREDNLHTHQRTHTGKKTSKRPAA TKKAGQAKKKKGSYPYDVPDYALEEASGSGRADALDDFDLDMLGSDA LDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINRSGSPKKKR KVGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAV PSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQ VLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEG TLSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLNQGIPV APHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDF SSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRI RPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTP EASHLLEDPEETSQAVKALREMADTVIPQKEEAICGQMDLSHPPPRG HLDLTTTLESMTEDLNLDSPLTPELNEILDFTLNDECLLHAMHISTGLSIF DTSLF
SEQ ID NO: 367 对照 载体 EGFP- KASH	GCGGCCGCACGCGTGGCGCGCCGTTTAACTTAATTAAGCTAGCCGT TACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC CCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAA TAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTA TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGggtcaggtgagccccacgttctgcttactctccccatctccccccctcccc acccccaattttgtattttatttttttaattttttgtgcagcgaatggggcgggggggggggggcgcgccagg cggggcgggggcgggggcgagggggcgggggcgggggcgagggcgagaggtgcggcgagccaatcagagcg gcgcgctccgaaagtcttctttatggcgagggcgggcgggcgggcgccctataaaaagcgagcgcgcgggcg gcgggagtcgctgcgttccttcgccccgtgccccgctccgcggcgctcgcgcggcgccccggctctgactg accggttactccacaggtgagcgggcggggacggcccttctctcggggtgtaattagcgcttggttaatgacg gctcgtttctttctgtggtgcgtgaaagccttaaaagggtccgggagggcccttgcgggggggagcgggctg gggggtgcgtgcgtgtgtgtgcgtggggagcgccgctgcggcccgcgctgcccggcggtgtgagcgctg cggggcgggcgcgggggcttgcgtcgcgtgcgtgaggggagcgcgggcgggggcggtgccccgcg gtgcgggggggctgcgaggggaacaaaggctgcgtgcgggggtgtgtgcgtgggggggtgagcaggggggtgtg ggcgcgggcggtgggctgtaacccccctgcacccccctccccagttgctgagcacggcccggttcgggtg cggggctccgtgcggggcggtggcgggggctgcgtgcggggcggggggtggcggcaggtgggggtgccg ggcgggggcgggggcgccctcggggcggggagggctcgggggagggcgcgcgggccccggagcgccggcg gctgtcagggcgggcgagccgcagccattgcctttatggtaatcgtgcgagagggcgagggacttccttgc ccaaatctggcgagccgaaatctgggagggcgccgcgccccctctagcgggcgcgggcggaagcggtgcg gcggcgaggaaggaatggcggggagggccttcgtgcgtgcggcgccgccgtcccccttctccatctccag cctcggggctgccgagggggagggctgccttcgggggggagggggcagggcggggggttcggcttctggcggtg gaccggcggtctagagcctctgctaaccatgttcattgccttcttcttctctacagctcctgggaacgtggtg ttgtgctgtctcatcttttgcaagaattGGTACCGCCACCATGGTGAGCAAGGGCGA GGAGCTGTTACACGGGGTGGTGCCCATCTGGTCGAGCTGGACGGCG ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA GCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGT GCAGTGCTTACGCCGCTACCCCGACCACATGAAGCAGCAGCACTTCT TCAAGTCCGCCATGCCCCAAGGCTACGTCCAGGAGCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGA

[0472]

[0473]

	序列
	ggcgcgcggtcgggctgtaacccccctgcacccccctccccagttgctgagcacggcccggttcgggtg cggggctccgtgcggggcggtggcgcggggctcgcgtgccggcgggggggtggcggcaggltgggggtgccg ggcggggcgggggccctcgggcccgggagggctcgggggaggggagggcgggcgccccggagcgccggcg gctgtcaggcgcgcgagccgagccattgcctttatggtaatcgtgcgagaggcgagggacttccttgtc ccaaatctggcgagccgaaatctgggagggcgccgcacccccctagcggcgcgggcggaagcggtgcg gcggcgaggaaggaaatggcggggagggccttcgtgcgtcgcgcggcgccgtcccttctccatccag cctcggggctgccgagggggacggctgccttcgggggggacggggcagggcggggttcggcttctggcgtgt gaccggcgctctagagcctctgtaacctgttcatgccttctcttttctacagctcctgggcaacgtgctgttg ttgtgtgtctcatctttggcaagaattGGTACCGCCACCATGGTGAGCAAGGGCGA GGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAA GCTGCCCCGTGCCCTGGCCACCCCTCGTGACCACCCTGACCTACGGCGT GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCT TCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGA GGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCA AGGAGGACGGCAACATCCTGGGGACAAGCTGGAGTACAATAACA CAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCA AGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAG CTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGT GCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCA AAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTG ACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGG ACTCAGATCTCGAGAGGAGGAGGAGGAGACAGACAGCAGGATGCCC CACCTCGACAGCCCCGGCAGCTCCAGCCGAGACGCTCCTTCTCTCA AGGGTGATCAGGGCAGCTCTACCGTTGCAGCTGCTTCTGCTGCTGCTG CTGCTCCTGGCCTGCCTGCTACCTGCCTCTGAAGATGACTACAGCTGC ACCCAGGCCAACAACTTTGCCCGATCCTTCTACCCCATGCTGCGGTAC ACCAACGGGCCACCTCCCACCTAGAGCGCTaatcaacctctggattacaaaatttgtaa agattgactgtattcttaactatgttgctcctttacgctatgtggatacgtgctttaatgcctttgtatcatgctattgctt cccgtatggctttcatttctcctcctgtataaatcctgggtgctgtctctttatgaggagttgtggccggtgtcaggcaa cgtggcggtgtgtcactgtgttgctgacgcaacccccactggttggggcattgccaccacctgtcagctccttcc gggactttcgtttccccctcctattgccacggcggaactatcgcgcctgccttggcgctgtggacaggggc tcggctgttgggactgacaattccgtggtgtgtcggggaaatcatgctccttcttggctgtcgcctatgttgcca cctggattctgcgcgggacgtccttctgtacgtcccttcggccctcaatccagcggaccttcttcccgggcctgc tgccggctctgcggccttccgcgtcttcgccttcgcctcagacgagtcggatcctccttggcgccgctccccgc AGATCTACGGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGG CCCTGGAAGTTGCCACTCCAGTGCCCAACAGCCTTGTCTAATAAAAT TAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGG GTGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACC TGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGC ACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCT CCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAG GCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTACCATATTG GCCAGGCTGGTCTCCAACCTCTAATCTCAGGTGATCTACCCACCTTGG CCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTCCCT GTCCTTCTGATTTTGTAGGTAACCACGTGCGGACCGAGCGGCCGC
SEQ ID	MTGKLAEKLPVTMSSLNQLPDNLYPEEIPSA LNLFSGSSDSV VHYNQ M

[0474]

	序列
NO: 422	ATENVMDIGLTNEKPNPELSYSGSFQAPGNKTVTYLGKFAFDSPSNWC
WT	QDNIISLMSAGILGVPPASGALSTQTSTASMVQPPQGDVEAMYPALPPYS
EGR3	NCGDLYSEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMIPDYNLY
	HHPNDMGSIPEHKPFQGMPIRVNPPPITPLETIKAFKDKQIHPGFGSLPQP
	PLTLKPIRPRKYPNRPSKTPLHERPHACPAEGCDRRFSRDELTRHLRIHT
	GHKPFQCRICMRSFSRSDHLTTHIRTHTGEKPFACEFCGRKFARSDERKR
	HAKIHLKQKEKKAEGGAPSASSAPPVSLAPVVTTC

[0475] 表26:被gRNA和dCas蛋白靶向的SCN1A基因组区域

[0476]

人类染色体 2 起始位置-靶序列所处的位置	gRNA 的 SEQ ID NO	指导 RNA 序列(对应于被 gRNA 识别的基因组序列)
166178880	197	ggtacgggcaaagatttcttg
166178871	198	tttggccgtaccaagctcttg
166177369	201	ACACAATGAGCCACCTACAAG
166177362	202	GTGGCTCATTGTGTGTGTGCC
166177299	203	AGAAAGCTGATACAGATACAA
166155393	207	TTCTCAGTTTGAATTAATAA
166155264	208	CATATCCCTGCAGGTTCAAG
166155255	209	TGGATTCTCTTCTGAACCTGC
166155099	210	agagagagagagagagagaga
166149373	211	TGGTCTCATTCTTTTGTGGG
166149176	107	aaggctgtctaggtaagtg
166149118	224	tggtctccagattaacactt
166148953	225	ATTACAGTTCTGTCAGCATGC
166148843	226	ATCATCTGTAACCATCAAGGA
166148565	227	TCCTGCCTACTTAGTTTCAAG
166148361	228	TGCTGAGGCAGGACACAGTGT
166142396	229	ACAAAGTAAGTGTGAGTGTGG
166142391	230	CTGACACTTACTTTGTCTAAA
166142344	231	ATAATAGTTGTGTCTTTATAA
166142239	232	CGATATTTTCATGGATTTCCTT
166142219	233	AAAACCTGGAACCGCATTCCCA
166141162	234	TGTACAAGCAGGGCTGCAAAG
166141145	235	AAAGGGGAATGGGAACACCCG
166141090	236	ATGTTCAAGGTGCAGAAGGAA
166140928	237	ttcaacaagctccaagaagt
166140590	238	GTTAACAATACTAAACAC
166128037	109	gctgatttgattaggtacca
166128002	108	gatgaagccgagaggatactg
165990246	247	TGTTTGCTCAAACGTGCACCA
165990193	248	AAATATGTACCACAAGAAATG
165990076	249	TATCTGGTTTCTCTCACTGCT
165989684	250	AAATAAGACATGAAAACAAGA
165989571	251	ATTGCAAAGCATAATTGGAT

[0477] 表27:被锌指eTF靶向的SCN1A基因组区域的实例

人类染色体 2 起始位置 - 靶序列所处的位置	靶位点的 SEQ ID NO	被 DNA 结合蛋白识别的靶位点序列	eTF 的 SEQ ID NO
166149168	35	ctagggtcaagtgtaggag	271
166149165	257	GGTCAAGTGTAGGAGACA	273
166128025	36	taggtaccatagagtgag	279
166127991	136	gaggatactgcagaggtc	280

[0479] 表28:识别SCN1A靶位点的eTF的氨基酸序列

eTF SEQ ID NO:	eTF 序列	识别的 靶位点
SEQ ID NO: 272	MAPKKRKRVGIHGVPAALEPGEKPYKCPECGKSFSSKKALTEHQ RTHTGEKPYKCPECGKSFSSPADLTRHQRTHRTHTGEKPYKCPECGKS FSRSDNLVRHQRTHRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTG EKPYKCPECGKSFSRSDNLVRHQRTHRTHTGEKPYKCPECGKSFSQSG NLTEHQRTHRTHTGEKPYKCPECGKSFSSTSGHLVRHQRTHRTHTGEKPYK CPECGKSFSQNSTLTEHQRTHRTHTGKKT SKRPAATKKAGQAKKKKG SYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLGSDALDD FDLDMLGSDALDDFDLDM	SEQ ID NO: 257
SEQ ID NO: 280	MAPKKRKRVGIHGVPAALEPGEKPYKCPECGKSFSDPGALVRHQ RTHTGEKPYKCPECGKSFSRSDNLVRHQRTHRTHTGEKPYKCPECGK SFSQSGDLRRHQRTHRTHTGEKPYKCPECGKSFSSTHLDLIRHQRTHTG EKPYKCPECGKSFSSTSGNLVRHQRTHRTHTGEKPYKCPECGKSFSRSD NLVRHQRTHRTHTGKKT SKRPAATKKAGQAKKKKGSYPYDVPDYAL EDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDAL DDFDLDM	SEQ ID NO: 136

[0481] 表29:eTF的DBD的氨基酸序列

[0482]	eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列
	SEQ ID NO: 270	SEQ ID NO: 393	LEPGEKPYKCPECGKSFSRKDNLKNHQRTHTEKPYKCPECGKS FSDPGALVRHQRTHTEKPYKCPECGKSFSREDNLHTHQRTHTG EKPYKCPECGKSFSDPGALVRHQRTHTEKPYKCPECGKSFSTS GELVRHQRTHTEKPYKCPECGKSFSRKDNLKNHQRTHTGKKT S
	SEQ ID NO: 272	SEQ ID NO: 395	LEPGEKPYKCPECGKSFSSKKALTEHQRTHTEKPYKCPECGKS FSSPADLTRHQRTHTEKPYKCPECGKSFSRSDNLVRHQRTHTG EKPYKCPECGKSFSREDNLHTHQRTHTEKPYKCPECGKSFSSRS DELVRHQRTHTEKPYKCPECGKSFSQSGNLTEHQRTHTEKPY KCPECGKSFSSTSGHLVRHQRTHTEKPYKCPECGKSFSQNSTLT EHQRTHTGKKT
	SEQ ID	SEQ ID	LEPGEKPYKCPECGKSFSRNDALTEHQRTHTEKPYKCPECGKS

[0483]

NO: 278	NO: 401	FSDPGHLVRHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTG EKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSSKK ALTEHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPY KCPECGKSFSRSDHLTNHQRTHTGKKTS
SEQ ID NO: 280	SEQ ID NO: 403	LEPGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKS FSRSDNLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTG EKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSTSG NLVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS
SEQ ID NO: 281	SEQ ID NO: 404	LEPGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPECGKS FSREDNLHHTHQRTHTGEKPYKCPECGKSFSTSGNLTEHQRTHTG EKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSQKS SLIAHQRTHTGEKPYKCPECGKSFSQAGHLASHQRTHTGKKTS
SEQ ID NO: 282	SEQ ID NO: 405	LEPGEKPYKCPECGKSFSTTGNLTVHQRTHTGEKPYKCPECGKS FSTSGELVRHQRTHTGEKPYKCPECGKSFSREDNLHHTHQRTHTG EKPYKCPECGKSFSTSGNLTEHQRTHTGEKPYKCPECGKSFSQSS SLVRHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGKKTS

[0484] 表30:GRN靶区域和用于上调GRN表达的eTF

[0485]

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
SEQ ID NO: 337	MAPKKRKRKVGIHGVPAALEPGEKPYKCPECGKSFSTSHSLTE HQRTHTGEKPYKCPECGKSFSDCRDLARHQRTHTGEKPYKC PECGKSFSTTGNLTVHQRTHTGEKPYKCPECGKSFSDPGALV RHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKC PECGKSFSHTGHLLEHQRTHTGKKTSKRPAATKKAGQAKKK KGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG DALDDFDLDMLGSDALDDFDLDM	SEQ ID NO: 330
SEQ ID NO: 338	MAPKKRKRKVGIHGVPAALEPGEKPYKCPECGKSFNRNDTLT EHQRTHTGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYK CPECGKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSRSDNL VRHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGEKPY KCPECGKSFSTKNSLTEHQRTHTGKKTSKRPAATKKAGQAK KKKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDM GSDALDDFDLDMLGSDALDDFDLDM	SEQ ID NO: 331
SEQ ID NO: 339	MAPKKRKRKVGIHGVPAALEPGEKPYKCPECGKSFSTSGNLTE HQRTHTGEKPYKCPECGKSFNRADNLTEHQRTHTGEKPYKC PECGKSFSSPADLTHQRTHTGEKPYKCPECGKSFSQSSNLV RHQRTHTGEKPYKCPECGKSFSQRANLRAHQRTHTGEKPYK CPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFHRTTL TNHQRTHTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYA LEEASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDALD DFDLMLGSDALDDFDLDMINSRSSGSPKKRKRKVGSYLP DTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAV PSRSSASVPKPAPQYPFTSSLSTINYDEFPTMVFPSPGQISQAS ALAPAPPQVLPQAPAPAPAMVSALAQAPAPVPVLPAGPPQ	SEQ ID No: 336

[0486]

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
	AVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDP AVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRL VTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSAL LGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPF HPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPA PAVTPEASHLLEDPEETSQAVKALREMA DTVIPQKEEAAIC GQMDLSHPPPRGHLDELTTTLESMTEDLNLDSP LTPELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	
SEQ ID NO: 340	MAPKKRKRKVG IHGVPAALEPGEKPYKCPECGKSFSSPADLTR HQRTH TGEKPYKCPECGKSFSDSGNLRVHQRTH TGEKPYKC PECGKSFSQLAHLRAHQRTH TGEKPYKCPECGKSFSQLANLR AHQRTH TGEKPYKCPECGKSFSDNLHQRTH TGEKPYK CPECGKSFSDNLVRHQRTH TGGKTSKRPAATKKAGQAKK KKGSYPYDVPDYALEDALDDFDLDM LGSDALDDFDLDM LG SDALDDFDLDM LGSDALDDFDLDM L	SEQ ID NO: 38
SEQ ID NO: 341	MAPKKRKRKVG IHGVPAALEPGEKPYKCPECGKSFSSPADLTR HQRTH TGEKPYKCPECGKSFSDSGNLRVHQRTH TGEKPYKC PECGKSFSQLAHLRAHQRTH TGEKPYKCPECGKSFSQLANLR AHQRTH TGEKPYKCPECGKSFSDNLHQRTH TGEKPYK CPECGKSFSDNLVRHQRTH TGGKTSKRPAATKKAGQAKK KKGSYPYDVPDYALEEASGSGRADALDDFDLDM LGSDALD DFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LINSRSSGS PKKKRKRKVG SQYLPD TDRHRIEEKRKRTYETFKSIMKKSPFS GPTDPRPPRR IAVPSRSSASVPKPAPQYPFTSSLSTINYDEF TMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSALAQA PAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDE DLGALLGNSTDP AVFTDLASVDNSEFQQLLNQGIPVAPHTTE PMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDE DFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGS LTPAPVPQPLDPAPAVTPEASHLLEDPEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTEDLN LDSPLTPELNEILD TFLNDECLLHAMHISTGLSIFDTSLF	SEQ ID NO: 38
SEQ ID NO: 342	MAPKKRKRKVG IHGVPAALEPGEKPYKCPECGKSFSSPADLTR HQRTH TGEKPYKCPECGKSFSDSGNLRVHQRTH TGEKPYKC PECGKSFSQLAHLRAHQRTH TGEKPYKCPECGKSFSQLANLR AHQRTH TGEKPYKCPECGKSFSDNLHQRTH TGEKPYK CPECGKSFSDNLVRHQRTH TGGKTSKRPAATKKAGQAKK KKGSYPYDVPDYALEEASGSGRADALDDFDLDM LGSDALD DFDLDM LGSDALDDFDLDM LGSDALDDFDLDM LINSRSSGS PKKKRKRKVG SQYLPD TDRHRIEEKRKRTYETFKSIMKKSPFS GPTDPRPPRR IAVPSRSSASVPKPAPQYPFTSSLSTINYDEF TMVFP SGQISQASALAPAPPQVLPQAPAPAPAMVSALAQA PAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDE	SEQ ID NO: 38

[0487]

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
	DLGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTE PMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDE DFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDVFE GREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGS LTPAPVPQPLDPAPAVTPEASHLLEDPEETSQAVKALREMA DTVIPQKEEAICGQMDLSHPPPRGHLDELTTTLESMTEDLN LDSPLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSLF	
SEQ ID NO: 343	MAPKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSSPADLTR HQRTHTEKPYKCPECGKSFSDSGNLRVHQRTHTEKPYKC PECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSQRANLR AHQRTHTEKPYKCPECGKSFSREDNLHTHQRTHTEKPYK CPECGKSFSRSDNLVRHQRTHTGKKTSKRPAATKKAGQAKK KKGSYPYDVPDYALEDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDMLGSDALDDFDLDMLG	SEQ ID NO: 38
SEQ ID NO: 344	MAPKKRKVGIIHGVPAALEPGEKPYKCPECGKSFSTSHSLTE HQRTHTEKPYKCPECGKSFHKNALQNHQRTHTEKPYKC PECGKSFSERSHLREHQRTHTEKPYKCPECGKSFSKKALT EHQRTHTEKPYKCPECGKSFSQRANLRAHQRTHTGEKPYK CPECGKSFSRKDNLKNHQRTHTEKPYKCPECGKSFSQSSNL VRHQRTHTEKPYKCPECGKSFSQSSSLVRHQRTHTEKPY KCPECGKSFSQAGHLASHQRTHTGKKTSKRPAATKKAGQAK KKKGSYPYDVPDYALEEASGSGRADALDDFDLDMLGSDAL DDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLINRSG SPKKRKVGVSQYLPDTDDRHRIEEKRTYETFKSIMKSPF SGPTDPRPPRRIAVPSRSSASVPKPAPQYPFTSSLSTINYDEF PTMVFPQGQISQASALAPAPPQVLPQAPAPAPAMVSALAQ APAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDD EDLGALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTT EPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGD EDFSSIADMDFSALLGSGSGSRDSREGMFLPKPEAGSAISDV EGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVG SLTPAPVPQPLDPAPAVTPEASHLLEDPEETSQAVKALREM ADTVIPQKEEAICGQMDLSHPPPRGHLDELTTTLESMTEDL NLSPLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSLF	SEQ ID NO: 335
SEQ ID NO: 345	MQSQLIKPSRMRYPNRPSKTPPHERPYACPVESCDRRFSSP ADLTRHIRIHTGQKPFQCRICMRNFSDSGNLRVHIRTHTEKP FACDICGRKFAQLAHLRAHTKIHLRQKDRPYACPVESCDRRF SQRANLRAHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTG EKFACDICGRKFARSDNLVRHTKIHLRQKDKLEMSGLEMA DHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPPHHHQQ QQPQHAFNALMGEHIHYGAGNMNATSGIRHAMGPGTVNNG HPPSALAPAAFNNNSQFMGPPVASQGGSLPASMQLQKLNNQ YFNHHPYPHNHYMPDLHPAAGHQMNNGTNQHFRDCNPKHSG GSSTPGGSGGSSTPGGSGSSSGGAGSSNSGGGSGSGNMPAS VAHVPAAMLPPNVIDTDFIDEVLMSLVIEMGLDRIKELPEL WLGQNEFDFMTDFVCKQQPSRVSC	SEQ ID NO: 38

[0488]

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
SEQ ID NO: 60	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQF PSPHHHQQQQPQHAFNALMGEHIHYGAGNMNATSGVRHA MGPGTVNGGHPPSALAPAARFNSQFMGPPVASQGGSLPAS MQLQKLNNQYFNHHPYPHNHYMPDLHPAAGHQMNNGTNQH FRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSGG GSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMG LDRIKELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQLIKPS RMRKYPNRPSKTPPHERPYACPVESCDRRFSRRDELNVHIRI HTGQKPFQCRICMRNFSSRRTCRAHIRTHTGEKPFACDICGR KFAQSSNLVRHTKIHLRQKDRPYACPVESCDRRFSQLAHLRA HIRIHTGQKPFQCRICMRNFSTSGNLVRHIRTHTGEKPFACDI CGRKFAHRTTLTNHTKIHLRQKDK	SEQ ID NO: 38
SEQ ID NO: 346	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQF PSPHHHQQQQPQHAFNALMGEHIHYGAGNMNATSGVRHA MGPGTVNGGHPPSALAPAARFNSQFMGPPVASQGGSLPAS MQLQKLNNQYFNHHPYPHNHYMPDLHPAAGHQMNNGTNQH FRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSGG GSGSGNMPASVAHVPAAMLPPNVIDTDFIDEVLMSLVIEMG LDRIKELPELWLGQNEFDFMTDFVCKQQPSRVSCQSQLIKPS RMRKYPNRPSKTPPHERPYACPVESCDRRFSRSPADLTRHIRI TGQKPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGRK FAQLAHLRAHTKIHLRQKDRPYACPVESCDRRFSQRANLRA HIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDI CGRKFARSDNLVRHTKIHLRQKDK	SEQ ID NO: 38
SEQ ID NO: 63	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRSP ADLTRHIRIHTGQKPFQCRICMRNFSDSGNLRVHIRTHTGEK FACDICGRKFAQLAHLRAHTKIHLRQKDRPYACPVESCDRRF SQRANLRAHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTG EKPFACDICGRKFARSDNLVRHTKIHLRQKDKLEMADHML AEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLRPRG APLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQP VATPYPGRAAAPPNAPGGPPGPQAPSAAAPPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGS APPAGSVSC	SEQ ID NO: 38
SEQ ID NO: 64	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPG LDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPP AAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQAPSAAAPPP PAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSE FDCFSDLGSAPPAGSVSCGGSGGGSGQSQLIKPSRMRKYPNR PSKTPPHERPYACPVESCDRRFSRSPADLTRHIRIHTGQKPFQ RICMRNFSDSGNLRVHIRTHTGEKPFACDICGRKFAQLAHLR AHTKIHLRQKDRPYACPVESCDRRFSQRANLRAHIRIHTGQK PFQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSD NLVRHTKIHLRQKDKLEMADHMLAEGYRLVQRPPSAAAA HGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYG	SEQ ID NO: 38

[0489]

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
	AFGPPSSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPNPAPG GPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELELGL HRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC	
SEQ ID NO: 347	MRPHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRS FSDSGNLRVHIRTHTGKPFACEFCGRKFAQLAHLRAHAKIH LKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRIC MRSFSREDNLHTHIRTHTGKPFACEFCGRKFARSDNLVRHA KIHLKQKEKKAEEKGGAPSASSAPPVSLAPVVTTCALEMSGLE MADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQFSPPHH HQQQQPQHAFNALMGEHIHYGAGNMNATSGIRHAMGPGTV NGGHPPSALAPAARFNNSQFMGPPVASQGGSLPASMQLQKL NNQYFNHHPYPHNHYMPDLHPAAGHQMNNGTNQHFRDCNP KHSGGSSTPGSGGSSTPGSGSGSSGGGAGSSNSGGSGSGN MPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKE LPELWLQGNEFDFMTDFVCKQQPSRVSC	SEQ ID NO: 38
SEQ ID NO: 348	MSGLEMAHMMAMNHGRFPDGTNGLHHHPAHRMGMGQF PSPHHHQQQQPQHAFNALMGEHIHYGAGNMNATSGVRHA MGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPAS MQLQKLNNQYFNHHPYPHNHYMPDLHPAAGHQMNNGTNQH FRDCNPKHSGGSSTPGSGGSSTPGSGSGSSGGGAGSSNSGG GSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIEMG LDRIKELPELWLQGNEFDFMTDFVCKQQPSRVSCRPHACPAE GCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFSDSGNLRV HIRTHTGKPFACEFCGRKFAQLAHLRAHAKIHLKQKEHACP AEGCDRRFSQRANLRAHLRIHTGHKPFQCRICMRSFSREDNL HTHIRTHTGKPFACEFCGRKFARSDNLVRHAKIHLKQKEKK AEKGGAPSASSAPPVSLAPVVTTCA	SEQ ID NO: 38
SEQ ID NO: 349	MRPHACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRS FSDSGNLRVHIRTHTGKPFACEFCGRKFAQLAHLRAHAKIH LKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRIC MRSFSREDNLHTHIRTHTGKPFACEFCGRKFARSDNLVRHA KIHLKQKEKLEMADHMLAEGYRLVQRPPSAAAAHGPHAL RTLPPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSS FQPFPAVPPPAAGIAHLQPVATPYPGRAAAPNPAPGGPPGPQP APSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRVREL PELFLGQSEFDCFSDLGSAPPAGSVSC	SEQ ID NO: 38
SEQ ID NO: 350	MAADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPG LDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPP AAGIAHLQPVATPYPGRAAAPNPAPGGPPGPQPAPSAAAPPP PAHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSE FDCFSDLGSAPPAGSVSCGGSGGGSGRPHACPAEGCDRRFSS PADLTRHLRIHTGHKPFQCRICMRSFSDSGNLRVHIRTHTGK PFACEFCGRKFAQLAHLRAHAKIHLKQKEHACPAEGCDRRF SQRANLRAHLRIHTGHKPFQCRICMRSFSREDNLHTHIRTHT	SEQ ID NO: 38

eTF SEQ ID NO:	eTF 序列	靶位点 SEQ ID NO:
[0490]	GEKPFACEFCGRKFARSDNLVRHAKIHLKQKEKLEMADHLM LAEGYRLVQRPPSAAAAHGPALRTLPPYAGPGLDSGLRPR GAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQ PVATPYPGRAAAPNAPGGPPGPQPAPSAAAPPPPAHALGGM DAELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDLGS APPAGSVSC	

[0491] 表31:用于上调GRN表达的eTF的示例性DBD

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列
SEQ ID NO: 337	SEQ ID NO: 377	LEPGEKPYKCPECCKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSDCRDLARHQRTHTGEKPYKCPECCKSFSTTGNLTV HQRTHTGEKPYKCPECCKSFSDPGALVRHQRTHTGEKPY KCPECCKSFSTHLDLIRHQRTHTGEKPYKCPECCKSFSHTG HLEHQRTHTGKKTS
SEQ ID NO: 338	SEQ ID NO: 378	LEPGEKPYKCPECCKSFSRNDLTHQRTHTGEKPYKCPE CGKSFSDPGALVRHQRTHTGEKPYKCPECCKSFSTSGELV RHQRTHTGEKPYKCPECCKSFSRSDNLVRHQRTHTGEKPY KCPECCKSFSTSGELVRHQRTHTGEKPYKCPECCKSFSSTK NSLTHQRTHTGKKTSKRPAATKKAGQAKKKK
SEQ ID NO: 339	SEQ ID NO: 379	LEPGEKPYKCPECCKSFSTSGNLTEHQRTHTGEKPYKCPE CGKSFSRADNLTHQRTHTGEKPYKCPECCKSFSSPADLT RHQRTHTGEKPYKCPECCKSFSQSSNLVRHQRTHTGEKPY KCPECCKSFSQRANLRAHQRTHTGEKPYKCPECCKSFSQL AHLRAHQRTHTGEKPYKCPECCKSFSHRTTLTNHQRTHT GKKTS
SEQ ID NO: 340	SEQ ID NO: 380	LEPGEKPYKCPECCKSFSSPADLTRHQRTHTGEKPYKCPE CGKSFSDSGNLVRVHQRTHTGEKPYKCPECCKSFSQLAHLR AHQRTHTGEKPYKCPECCKSFSQRANLRAHQRTHTGEKPY KCPECCKSFSREDNLHTHQRTHTGEKPYKCPECCKSFSR SDNLVRHQRTHTGKKTS
SEQ ID NO: 341	SEQ ID NO: 381	LEPGEKPYKCPECCKSFSSPADLTRHQRTHTGEKPYKCPE CGKSFSDSGNLVRVHQRTHTGEKPYKCPECCKSFSQLAHLR AHQRTHTGEKPYKCPECCKSFSQRANLRAHQRTHTGEKPY KCPECCKSFSREDNLHTHQRTHTGEKPYKCPECCKSFSR SDNLVRHQRTHTGKKTS
SEQ ID NO: 342	SEQ ID NO: 382	LEPGEKPYKCPECCKSFSSPADLTRHQRTHTGEKPYKCPE CGKSFSDSGNLVRVHQRTHTGEKPYKCPECCKSFSQLAHLR AHQRTHTGEKPYKCPECCKSFSQRANLRAHQRTHTGEKPY KCPECCKSFSREDNLHTHQRTHTGEKPYKCPECCKSFSR SDNLVRHQRTHTGKKTS
SEQ ID NO: 343	SEQ ID NO: 383	LEPGEKPYKCPECCKSFSSPADLTRHQRTHTGEKPYKCPE CGKSFSDSGNLVRVHQRTHTGEKPYKCPECCKSFSQLAHLR AHQRTHTGEKPYKCPECCKSFSQRANLRAHQRTHTGEKPY

[0493]

eTF SEQ ID NO:	DBD SEQ ID NO:	DBD 序列
		YKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSR SDNLVRHQRTHTGKKTS
SEQ ID NO: 344	SEQ ID NO: 384	LEPGKPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPEC GKSFSHKNALQNHQRTHTGEKPYKCPECGKSFSERSHLRE HQRTHTGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYK CPECGKSFSQRANLRAHQRTHTGEKPYKCPECGKSFSRKD NLKNHQRTHTGEKPYKCPECGKSFSQSSNLVRHQRTHTGE KPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFS QAGHLASHQRTHTGKKTS
SEQ ID NO: 345	SEQ ID NO: 385	RKYPNRPSTPPHERPYACPVESCDRRFSSPADLTRHIRIHT GQKPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGR KFAQLAHLRAHTKIHRLRQKDRPYACPVESCDRRFSQRANL RAHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPF ACDICGRKFARSDNLVRHTKIHRLRQKD
SEQ ID NO: 346	SEQ ID NO: 386	RKYPNRPSTPPHERPYACPVESCDRRFSSPADLTRHIRIHT GQKPFQCRICMRNFSDSGNLRVHIRTHTGEKPFACDICGR KFAQLAHLRAHTKIHRLRQKDRPYACPVESCDRRFSQRANL RAHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTGEKPF ACDICGRKFARSDNLVRHTKIHRLRQDK
SEQ ID NO: 347	SEQ ID NO: 387	HACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFS DSGNLRVHIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIH LKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRI CMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSDNLV RHAKIHLKQKE
SEQ ID NO: 348	SEQ ID NO: 388	HACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFS DSGNLRVHIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIH LKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRI CMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSDNLV RHAKIHLKQKE
SEQ ID NO: 349	SEQ ID NO: 389	HACPAEGCDRRFSSPADLTRHLRIHTGHKPFQCRICMRSFS DSGNLRVHIRTHTGEKPFACEFCGRKFAQLAHLRAHAKIH LKQKEHACPAEGCDRRFSQRANLRAHLRIHTGHKPFQCRI CMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSDNLV RHAKIHLKQKE
SEQ ID NO: 350	SEQ ID NO: 390	MADHMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPG LDGLRPRGAPLGPMPRQPGALAYGAFGPPSSFPFPAVP PPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQAPSAA APPPPAHALGGMDAELIDEEALTSLELELGLHRVRELPELF LGQSEFDCFSDLGSAPPAGSVSC

[0494] 实施例

[0495] 包括以下实施例以进一步描述本公开的一些方面,并且不应用来限制本发明的范围。

[0496] 实施例1

[0497] 增加HEK293T细胞中的基因表达

[0498] 用在几种不同调节元件(即无启动子对照;SCP;CMV;与minCMV可操作地连接的SEQ ID NO:178;与minCMV可操作地连接的SEQ ID NO:182;和CAG)之一的控制下含有萤光素酶基因的质粒DNA转染HEK293T细胞。每种构建体的归一化萤光素酶值在图1中示出。每种构建体的大小归一化的活性值在图2中示出。具有与SEQ ID NO:178或182的调节元件连接的

minCMV启动子的构建体比单独的minCMV和单独的SCP驱动更高水平的萤光素酶表达。

[0499] 该实验表明,SEQ ID NO:178和182是在细胞中驱动高基因表达的RE。可以将这样的RE添加到本文公开的包含SCN1A的非天然存在的转录激活物的表达盒中,以增加细胞中转录激活物的表达,这导致内源SCN1A基因的表达增加。在一些情况下,选自SEQ ID NO:178和/或182的一个或多个RE与转录激活物可操作地连接,以增加细胞中的SCN1A表达。可以将这类调节元件添加到表达盒中转录激活物的上游和/或下游。

[0500] 实施例2

[0501] 对小白蛋白神经元的选择性

[0502] 使用荧光成像确定对小白蛋白细胞——一种类型的GABA能神经元——的选择性。将含有与对照启动子(EF1 α)或PV选择性RE可操作地连接的eGFP的AAV载体以及含有CRE依赖性tdTomato的AAV载体共同注射到GAD2-IRES-CRE小鼠(Jackson Labs)中。

[0503] 以0.3 μ L/min的速率将1.5 μ L的AAV载体(5^{12} 至 1^{13} gc/ml)双侧输注到小鼠的背侧和腹侧海马体中,注射后有4min的休息期。用异氟烷(2%,800mL/min O₂)对小鼠进行麻醉以供注射。使用布比卡因/肾上腺素进行局部麻醉,而使用卡洛芬进行围手术期/术后麻醉。将动物放置在立体定位架(Kopf instruments,USA)中,使用以下坐标,背侧海马体(AP-2.0mm,侧向 \pm 1.5,DV-距硬脑膜1.4mm),腹侧海马体(AP-3.1mm,侧向 \pm 2.8,DV-距硬脑膜3.8mm)。Hamilton注射器(型号80308;配有相应的30ga钝头针的10 μ L注射器)与立体定向显微操作器一起使用,以指定并钻出钻孔。该钻头仅用于穿透骨骼。钻孔后,将输液套管下降到脑中所需位置的深度。具有micro4控制器的超微泵III(World Precision Instruments)的注射设置为:注射体积:1.5 μ L;注射速率:0.3 μ L/min。将针缓慢下降(经大约1分钟)至输注坐标以下DV-0.1mm(背侧海马体为-1.5mm,腹侧海马体为-3.9mm,然后分别升至DV-1.4或-3.8)。输注前,让针平衡1分钟。一旦完成递送,将针留置4分钟,然后在约1分钟内取出。一旦所有四个输注完成,就将皮肤切口用缝合线封闭。经治疗的小鼠在余下的研究中每天进行健康检查,并每周称重一次以监测体重。

[0504] 对于组织采集,通过过量的异氟烷对小鼠实施安乐死,并灌注4%低聚甲醛(PFA)。提取一块含有海马体的脑组织,并在4℃的4% PFA中放置至少12小时。然后将脑组织在4℃下在30%蔗糖中(在磷酸盐缓冲盐水中)脱水,直到组织沉入管底。将脑组织包埋到Tissue-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的染色。

[0505] 图3示出了用含有eGFP和PV选择元件(SEQ ID NO:183或SEQ ID NO:184)或组成型EF1 α 控制启动子(SEQ ID NO:186)的AAV载体以及含有tdTomato荧光报道基因(其表达依赖于Cre重组酶活性)的第二AAV载体处理的小鼠的代表性图像。GAD2在包括PV神经元在内的GABA能神经元中选择性表达,因此tdTomato表达可鉴别GABA能神经元(图3,下行)。

[0506] 图4示出了对于每种启动子或调节元件,在PV细胞中的表达效率。由EF1 α (组成型启动子)驱动的eGFP表达显示出约40%的表达效率。相反,由SEQ ID NO:183(其包含具有SEQ ID NO:185的序列的启动子)或SEQ ID NO:184驱动的eGFP表达显示出约90%的表达效率,显示出比EF1 α 启动子高得多的表达效率。

[0507] 图5示出了不同调节序列的表达特异性,与显示对小白蛋白(PV)细胞的特异性低

于10%的EF1 α 启动子相比,SEQ ID NO:183和184均以约75%的特异性驱动eGFP在PV细胞中的表达。

[0508] 这些实验证实了SEQ ID NO:183和184调节元件对GABA能/PV神经元的选择性,并且示出了可用来测量对GABA能/PV细胞的选择性的试验。本文所述的表达盒中可以包含SEQ ID NO:183-184调节元件中的一种或多种,以增加转录激活物在GABA能/PV细胞中与在非GABA能细胞中相比的选择性表达,从而使脱靶效应最小化,如使用以上描述的试验所测定的。在一些情况下,使用与以上描述的试验类似的试验,在表达盒中使用一种或多种对PV细胞具有选择性的RE,以增加转录激活物在PV细胞中与在非PV细胞中相比的选择性表达。

[0509] 实施例3

[0510] 使用SCN1A特异性转录激活物,能够上调SCN1A的靶区域的鉴定

[0511] 为了鉴定能够上调内源SCN1A表达的基因组区域,设计了各种工程化转录因子(锌指核酸酶或gRNA/daCas9构建体),它们靶向如以上表26-29所示的基因组的各个区域。对于gRNA/daCas9构建体,由于将gRNA设计为靶向互补基因组链,因此该gRNA具有与靶区域相同的序列。dCas9蛋白的序列为SEQ ID NO:104,也是具有NLS和HA标签的SEQ ID NO:103。

[0512] 按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3 μ g携带工程化转录因子或对照构建体的质粒进行转染(FugeneHD, Promega)。用表达以上表26-29中所示的构建体的质粒转染细胞。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并用DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3 μ g)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98 $^{\circ}$ C下30s, 40x[98 $^{\circ}$ C下10sec, 66 $^{\circ}$ C下15sec, 72 $^{\circ}$ C下15sec])。使用针对SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO:190); 5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO:191))来量化内源SCN1A转录物的水平,并以GAPDH为参考基因通过delta-delta Ct方法(5'-ACCACAGTCCATGCCATCAC'-3'(SEQ ID NO:192); 5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定SCN1A表达的相对水平。数据以相对于对照条件的变化倍数呈现。

[0513] 结果以SCN1A转录相对于对照条件(例如,EGFP-KASH报道构建体)的变化倍数示于图6和以下的表32中。

[0514] 表32:不同基因组靶位点和相应eTF对转录的影响

[0515]

靶位点 SEQ ID NO:	被 eTF 识别的靶序列	平均值	T 检验
197	ggtacgggcaaagatttcttg	1.36801947	0.48762102
198	tttgcccgtagcaaagtcttg	1.26748653	0.71262289
201	ACACAATGAGCCACCTACAAG	1.31187425	0.42605224
202	GTGGCTCATTGTGTGTGTGCC	1.25217773	0.26572657
203	AGAAAGCTGATACAGATACAA	1.7542842	0.34519408
203	AGAAAGCTGATACAGATACAA	1.45636874	0.44464045
207	TTCTCAGTTTTGAAATTAAAA	1.64498972	0.21705582
208	CATATCCCTGCAGGTTTCAGAA	1.75688991	0.28984533
209	TGGATTCTCTTCTGAACCTGC	2.27026665	0.43195546
210	agagagagagagagagagaga	2.05701745	0.42409102
211	TGGTCTCATTCTTTTTGTGGG	1.71471378	0.32104302
107	aaggctgtctaggtcaagtgt	1.36378425	0.18753821
35	ctaggtcaagtgtaggag	5.12090848	0.0096628
257	GGTCAAGTGTAGGAGACA	1.52068773	0.62403349
224	tgttctccagattaacactt	1.63040825	0.46710683
225	ATTACAGTTCTGTCAGCATGC	1.34500323	0.32186367
226	ATCATCTGTAACCATCAAGGA	2.58328006	0.0748197
227	TCCTGCCTACTTAGTTTCAAG	1.97097781	0.25980859
228	TGCTGAGGCAGGACACAGTGT	2.4290169	0.30364553
229	ACAAAGTAAGTGTCAGTGTGG	1.30739959	0.72725347
230	CTGACACTTACTTTGTCTAAA	1.95513108	0.02069095
231	ATAATAGTTGTGTCTTTATAA	1.55783618	0.29846459
232	CGATATTTTCATGGATTCCCTT	1.7735976	0.21954265
233	AAAACTGGAACCGCATTCCCA	2.08698135	0.0454403
234	TGTACAAGCAGGGCTGCAAAG	1.4663605	0.02946062
235	AAAGGGGAATGGGAACACCCG	1.16749665	0.65768346
236	ATGTTCAAGGTGCAGAAGGAA	2.04547409	0.09880194
237	ttcaacaagctccaagaagt	1.46929899	0.24465271

[0516]	靶位点 SEQ ID NO:	被 eTF 识别的靶序列	平均值	T 检验
	238	GTTAACAAATACACTAAACAC	1.37399196	0.33638238
	109	gctgattgtattaggtacca	22.4892633	0.09291316
	36	taggtaccatagagttag	25.4730028	0.14942042
	108	gatgaagccgagaggatactg	18.7579211	0.13148732
	136	gaggatactgcagaggtc	8.6766618	0.16432794
	247	TGTTTGCTCAAACGTGCACCA	2.13402102	0.25583999
	248	AAATATGTACCACAAGAAATG	2.29522738	0.41829497
	249	TATCTGGTTTCTCTCACTGCT	1.44542116	0.0947106
	250	AAATAAGACATGAAAACAAGA	1.27016182	0.32368695
	251	ATTGCAAAGCATAATTTGGAT	1.42246971	0.18117243

[0517] 实施例4

[0518] SCN1A从表达盒的相对表达

[0519] 该实施例描述了SCN1A从各种表达盒的相对表达,所述表达盒包含增加SCN1A基因表达的非天然存在的转录调节物。SCN1A基因属于提供建立钠离子通道的指令的基因家族。这些通道将带正电荷的钠原子(钠离子)转运到细胞内,在细胞产生并传输电信号的能力方面起关键作用。为了测试表达盒,按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3 μg质粒进行转染(PEI)。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并用DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3μg)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98℃下30s, 40x[98℃下10sec, 66℃下15sec, 72℃下15sec])。

[0520] 构建各种表达盒,包括包含eGFP报道转基因且缺乏转录激活物的对照表达盒;以及如以上表17所述的表达盒A-H。对于表达盒A-F,各自包含非天然存在的DNA结合域,该DNA结合域包含6至9个根据SEQ ID NO:103、131-134的锌指,所述锌指连接至具有SEQ ID NO:95的序列的VP64转录激活域或具有SEQ ID NO:114的序列的VPR转录激活域。工程改造此类非天然存在的转录激活物,以结合具有SEQ ID NO:35-37、107-108或136中任一序列的靶位点或包含SEQ ID NO:35-37、107-108或136中任一个的基因组区域。表达盒G和H各自包含具有SEQ ID NO:103的序列的dSaCas9 DNA结合域,该DNA结合域连接至具有SEQ ID NO:95的序列的VP64转录激活域。每个G和H表达盒还包含用于靶向非天然存在的转录激活物的gRNA,其中每个gRNA包含选自SEQ ID NO:107-108的序列。每个表达盒还包含具有SEQ ID NO:178的序列的RE。

[0521] 使用针对SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO:190);5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO:191))来量化内源SCN1A转录物的水平,并以GAPDH为参考基因通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC' 3'(SEQ ID NO:192);5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定SCN1A表达的相对水平。

[0522] 图7示出了用不同表达盒转染的细胞的SCN1A的相对表达,呈现为相对于对照eGFP构建体的变化倍数(Log₁₀)。相对于对照表达盒,非天然存在的转录激活物从表达盒A-H的表达导致HEK293细胞中的SCN1A基因表达增加。

[0523] 图8示出了包含非天然存在的转录激活物的表达盒A,该转录激活物用来增加SCN1A基因与其他基因相比在HEK293细胞中的表达。另外,使用针对TTC21B(5'-GGTCACGTACAGCTTCGCAT-3'(SEQ ID NO:283);5'-CTGGTTTCTGGCTCGTGGAG-3'(SEQ ID NO:284))、SCN9A(5'-AAGCCCCAAAGCCAAGCAG-3'(SEQ ID NO:285);5'-AGGTGTGGCATTGAAACGG-3'(SEQ ID NO:286))、GRN(5'-ATGGTCAGTTCTGCCCTGTG-3'(SEQ ID NO:287);5'-CGGTAAAGATGCAGGAGTGGC-3'(SEQ ID NO:288))、UTRN(5'-TGACAATGGGCAGAACGAAT-3'(SEQ ID NO:289);5'-TGCAGCACTCTGTTGACGTT(SEQ ID NO:290))和TTN(5'-TGTTGCCACTGGTGCTAAAG-3'(SEQ ID NO:291);5'-ACAGCAGTCTTCTCCGCTTC-3'(SEQ ID NO:292))的引物来量化内源TTC21B、SCN9A、GRN、UTRN和TTN转录物的水平,并使用具有SEQ ID NO:192-193的序列的引物,以GAPDH为参考基因,通过delta-delta Ct法确定这些基因的相对表达水平。

[0524] 图8示出了以相对于对照条件的变化倍数呈现的,内源SCN1A、TTC21B、SCN9A、GRN、UTRN和TTN转录物的相对表达。如上所述,与测试的其他基因相比,表达盒A能够特异性增加SCN1A基因或Nav1.1蛋白的表达。这表明被表达盒A的转录激活物识别的靶位点对于SCN1A基因是特异性的,因此导致HEK293细胞中SCN1A基因表达的增加。

[0525] 实施例5

[0526] SCN1A激活

[0527] 按照标准方法培养HEK293细胞,并用3ug的每种测试的激活物或EGFP对照进行转染(FuGene HD,Promega)。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并用DNA酶处理。使用OligoDT引物(Superscript IV,Invitrogen)对RNA(3ug)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98℃下30s,40x[98℃下10sec,68℃下15sec,72℃下15sec])。使用针对人SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO:190);5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO:191))来量化报道基因驱动的EGFP转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC-3'(SEQ ID NO:192);5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定EGFP表达的相对水平。数据以相对于对照条件的变化倍数呈现。参见图12B和图13B。SEQ ID NO:6、7和9是衍生自人EGR1蛋白的eTF,并且被工程改造为与内源SCN1A基因处或附近的靶位点结合,其中靶位点为SEQ ID NO:35-37。与报道EGFP对照相比,包含SEQ ID NO:6、7或9中任一个的eTF分别导致至少2倍、7倍或20倍的相对SCN1A表达。

[0528] 实施例6

[0529] 使用SCN1A特异性转录因子,HEK293细胞中内源SCN1A的上调

[0530] 按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3ug携带工程化转录因子或EGFP对照构建体的质粒进行转染(FugeneHD,Promega)。用含有以下表33中所示的构建体的质粒转染细胞。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并用DNA酶处理。使用OligoDT引物(Superscript IV,Invitrogen)对RNA(3ug)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98℃下30s,40x[98℃下10sec,66℃下15sec,72℃下15sec])。使用针对SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO:190);5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO:191))来量化内源SCN1A转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法

(5' -ACCACAGTCCATGCCATCAC' 3' (SEQ ID NO:192); 5' -TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO:193)) 确定SCN1A表达的相对水平。数据以相对于对照条件的变化倍数呈现(参见图15)。对照构建体由在具有根据SEQ ID NO:178的序列的启动子的控制下表达的EGFP组成。相对于EGFP条件,工程化转录因子的递送诱导内源SCN1A转录物的不同程度的上调。

[0531] 表33. 在实施例6和图15中使用的构建体。

[0532]	eTF SEQ ID NO:	
	表 17 中的表达盒 D	SEQ ID NO: 15
	表 17 中的表达盒 B	SEQ ID NO: 364
	表 17 中的表达盒 C	SEQ ID NO: 9
	表 17 中的表达盒 E	SEQ ID NO: 8
	SEQ ID NO: 67	SEQ ID NO: 7
	SEQ ID NO: 68	SEQ ID NO: 295
	SEQ ID NO: 88	SEQ ID NO: 14
	SEQ ID NO: 69	SEQ ID NO: 296
	SEQ ID NO: 87	SEQ ID NO: 297
	SEQ ID NO: 83	SEQ ID NO: 13
	SEQ ID NO: 84	SEQ ID NO: 298
	SEQ ID NO: 75	SEQ ID NO: 299
	SEQ ID NO: 74	
	SEQ ID NO: 70	
	SEQ ID NO: 71	

[0533] 表34: 在实施例6和图15中使用的构建体。

[0534]	eTF SEQ ID NO:	DBD SEQ ID NO:
	SEQ ID NO:364	SEQ ID NO:406
	SEQ ID NO:295	SEQ ID NO:407
	SEQ ID NO:296	SEQ ID NO:408
	SEQ ID NO:299	SEQ ID NO:409

[0535] 实施例7

[0536] 使用SCN1A特异性转录因子, GABA神经元中内源SCN1A的上调

[0537] 将iCell GABA神经元(Cellular Dynamics)接种于6孔板中(约1E6个细胞/孔), 并且按照制造商推荐的方案进行维持。接种后72h, 以大约2E11个基因组拷贝/孔, 将在普遍存在的启动子(CBA启动子)控制下表达EGFP或激活物(图16A中的SEQ ID NO:366或图16B中的SEQ ID NO:13或SEQ ID NO:15)的重组AAV(血清型AAV-DJ)添加到培养基中。感染后一周(图16A)或两周(图16B), 从培养的细胞中分离RNA(Qiagen RNeasy Mini试剂盒), 并进行DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对回收的RNA进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I, 通过qPCR分析cDNA样品: (98°C下30s, 40x[98°C下10sec, 66°C下15sec, 72°C下15sec])。使用针对SCN1A的引物(5' -TGTCTCGGCATTGAGAACATTC-3' (SEQ ID NO:190); 5' -ATTGGTGGGAGGCCATTGTAT-3' (SEQ ID NO:191)) 来量化内源SCN1A转录物的水平, 并以GAPDH为参考基因, 通过delta-delta Ct法(5' -ACCACAGTCCATGCCATCAC' 3' (SEQ ID NO:192); 5' -TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO:193)) 确定SCN1A表达的相对水平。数据以相对于对照条件的变化倍数呈现(参见图16A

和图16B)。AAV驱动的工程化转录因子的表达在培养的iPS衍生的GABA神经元中产生内源SCN1A转录物的显著上调。

[0538] 实施例8

[0539] 使用SCN1A特异性转录因子,GABA神经元中内源SCN1A的特异性上调

[0540] 将iCell GABA神经元(Cellular Dynamics)接种于6孔板中(约1E6个细胞/孔),并且按照制造商推荐的方案进行维持。接种后72h,以大约2E11个基因组拷贝/孔,将在CBA启动子控制下表达EGFP或激活物(SEQ ID NO:366,其包含与CBA启动子驱动的VPR TAD融合的锌指DBD)的重组AAV(血清型AAV-DJ)添加到培养基中。

[0541] 感染后一周,从培养的细胞中分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用TruSeq Stranded mRNA文库试剂盒(Illumina)从回收的RNA制备RNAseq文库,并在Illumina NextSeq上进行测序(2x 75个循环配对末端测序)。将测序读取与人类基因组(RNASTAR)进行比对,并使用DESeq2进行差异表达分析。数据以相对于对照(AAVDJ-CBA-EGFP)样品的变化倍数来呈现(参见图17)。结果示于表35中,图17示出了以相对于对照条件的变化倍数呈现的,内源SCN1A和40个最近邻基因转录物的相对表达。如上所述,与检查的其他基因相比,SEQ ID NO:366能够特异性增加SCN1A基因或Nav1.1蛋白的表达。这表明被SEQ ID NO:366的转录激活物识别的靶位点对于SCN1A基因是特异性的,因此导致GABA神经元中SCN1A基因表达的增加。

[0542] 表35.对用SCN1A特异性转录因子(SEQ ID NO:366)处理的GABA神经元中内源SCN1A和40个最近邻基因的转录的影响。

[0543]

基因名称	染色体 2 起 点	染色体 2 终点	染色体链	相对于对照的变化倍数
PLA2R1	160788518	160919121	-	0.16367458
ITGB6	160956176	161128399	-	0.20679884
RBMS1	161128661	161350305	-	1.63514667
TANK	161993418	162092732	+	0.90946407
PSMD14	162164548	162268228	+	0.92699237
TBR1	162272604	162282381	+	0.53199642

[0544]

基因名称	染色体 2 起 点	染色体 2 终点	染色体链	相对于对照的变化倍数
SLC4A10	162280842	162841792	+	1.89407328
DPP4	162848750	162931052	-	2.82345284
FAP	163027193	163101661	-	2.26977379
IFIH1	163123588	163175213	-	1.46146481
GCA	163175349	163228105	+	2.58702426
FIGN	164449905	164592522	-	0.46785861
GRB14	165349321	165478358	-	0.5631965
COBLL1	165510133	165700189	-	0.43199257
SLC38A11	165752695	165812035	-	4.06730119
SCN3A	165944031	166060577	-	1.0807866
SCN2A	166095911	166248818	+	1.24475196
CSRNP3	166326156	166545917	+	0.82971233
GALNT3	166604100	166651192	-	0.33804418
TTC21B	166713984	166810353	-	1.58661143
SCN1A	166845669	166984523	-	62.9552975
SCN9A	167051694	167232503	-	1.71659087
SCN7A	167260082	167350757	-	0.29331967
B3GALT1	168675181	168730551	+	0.64436013
STK39	168810529	169104651	-	1.19821739
CERS6	169312371	169631644	+	0.86828378
NOSTRIN	169643048	169722024	+	1.82142718
SPC25	169690641	169769881	-	0.86880697
ABCB11	169779447	169887832	-	3.1441368
DHRS9	169921298	169952677	+	1.10381777
BBS5	170335687	170382432	+	0.65476347
KLHL41	170366211	170382772	+	0.87373377
FASTKD1	170386258	170430385	-	1.02786927
PPIG	170440849	170497916	+	1.09866236
CCDC173	170501934	170550943	-	0.67290779
PHOSPHO2	170550974	170558218	+	0.91339152
KLHL23	170550997	170633499	+	0.73926347
SSB	170648442	170668574	+	1.00631994
METTL5	170666590	170681441	-	1.21271497
UBR3	170683967	170940641	+	1.21350908
MYO3B	171034654	171511681	+	0.52839217

[0545] 实施例9

[0546] SCN1A从表达盒的体内表达

[0547] 为了测试SCN1A的转录激活物的体内表达,由Vector Biolabs (Malvern,PA)生成了重组AAV9载体。向雄性C57B1/6小鼠(每组N=5,7-8周龄)的背侧海马体(AP-2.0mm,侧向±1.5,距硬脑膜DV-1.4mm)和腹侧海马体(AP-3.1mm,侧向±2.8,距硬脑膜DV-3.8mm)中双侧输注1.5ul纯化的AAV载体,总共4个注射部位。以0.3ul/分钟的速率递送AAV,每次注射后有4m的休息期。治疗后四周,对小鼠实施安乐死并解剖海马组织。对于每组,除了一只动物仅收集并均质化左侧海马体外,大多数动物(N=4)均收集来自左侧和右侧海马组织的组织,并合并以进行均质化。从匀浆中分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV,Invitrogen)对RNA(3μg)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品中的小鼠SCN1A表达:98℃下30s,40x[98℃下10sec,64℃下15sec,72℃下15sec]。使用针对小鼠SCN1A的引物(5'-CAAAAAGCCACAAAAGCCT-3'(SEQ ID NO:374);5'-TTAGCTCCGCAAGAAACATC-3'(SEQ ID NO:375))来量化内源SCN1A转录物的水平,并以GAPDH为参考基因,使用以上实施例1所述的相同GAPDH引物,通过delta-delta Ct法确定体内SCN1A表达的相对水平。

[0548] 图18A和图18B示出了五只动物的平均结果,每只动物注射AAV9构建体。eGFP对照构建体包含eGFP报道转基因。表达盒A(参见表17)包含转录激活物,该转录激活物识别包含SEQ ID NO:35的靶序列,如以上表17中所述。图18A示出了SCN1A在体内的相对表达。图18B以平均eGFP表达百分比示出了体内SCN1A表达的变化。这些结果表明,表达盒A的SCN1A转录激活物导致体内SCN1A表达上调约20-30%。

[0549] 此类表达盒可适合用于人类,以治疗Dravet综合征、癫痫、癫痫发作、阿尔茨海默病、帕金森病和/或与SCN1A缺乏和/或活性受损相关的其他任何疾病或病况。

[0550] 实施例10

[0551] 通过qPCR确定的,使用SCN1A特异性转录因子,在WT小鼠海马体中内源SCN1A转录物的上调

[0552] 重组AAV9载体由Vector Biolabs (Malvern,PA)产生。向雄性C57B1/6小鼠(每组N=5,7-8周龄)的背侧海马体(AP-2.0mm,侧向±1.5,距硬脑膜DV-1.4mm)和腹侧海马体(AP-3.1mm,侧向±2.8,距硬脑膜DV-3.8mm)中双侧输注1.5ul纯化的AAV载体,总共4个注射部位。使用包含各种SCN1A特异性转录因子的AAV载体:表17中的表达盒A(图19A),表达盒B(图19B),以及SEQ ID NO:366和SEQ ID NO:365(图19C),并与对照载体(由相同启动子驱动的EGFP-KASH)(图19A-C)和由CBA启动子驱动的EGFP-KASH(图19C)进行比较。

[0553] 以0.3ul/分钟的速率递送AAV,每次注射后有4分钟的休息期。治疗后四周,对小鼠实施安乐死并解剖海马组织。对于每组,除了一只动物仅收集并均质化左侧海马体外,大多数动物(N=4)均收集来自左侧和右侧海马组织的组织,并合并以进行均质化。从匀浆中分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV,Invitrogen)对RNA(3μg)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品中的小鼠SCN1A(mNav1.1正向引物:5'-CAAAAAGCCACAAAAGCCT-3'(SEQ ID NO:293)和mNav1.1反向引物:5'-TTAGCTCCGCAAGAAACATC-3'(SEQ ID NO:294)): (98℃下30s,40x[98℃下10sec,64℃下15sec,72℃下15sec])。

[0554] 直接注射携带每种工程化转录因子构建体的AAV导致小鼠海马组织中内源SCN1A转录物的上调。

[0555] 实施例11

[0556] 在Dravet综合征小鼠模型中的高热癫痫发作 (HTS) 试验

[0557] 在Scn1a^{tm1K^{ea}}小鼠品系中测试了使用表达盒对Dravet综合征和/或其症状的治疗。该小鼠系是已确立的Dravet综合征的小鼠模型。Scn1a^{tm1K^{ea}}小鼠系不需要CRE重组酶。Scn1a^{tm1K^{ea}}小鼠 (可从Jackson Laboratory获得;在Hawkins等人,Scientific Reports, vol.7:15327 (2017) 中描述) 包含SCN1A的第一个编码外显子的缺失。SCN1A敲除等位基因的纯合小鼠的特征为震颤、共济失调、癫痫发作,并在出生后第16天死亡。C57BL/6背景的杂合小鼠发展出自发性癫痫发作,并且大多数在数周内死亡。这样的小鼠品系可用来研究癫痫和Dravet综合征治疗的安全性和功效。其他信息参见Miller等人,Genes Brain Behav.2014Feb;13 (2):163-72。

[0558] 为了测试转录激活物在Scn1a^{tm1K^{ea}}小鼠品系中的功效,在P1经双侧ICV,向雄性Scn1a^{+/+}与雌性C57BL/6J杂交繁殖产生的幼仔给予AAV载体。向小鼠给予以下表36中总结的构建体。使母鼠不受干扰,直到在P18断奶,然后再次使小鼠不受干扰,直到P26-P28开始HTS试验。给药的P1小鼠的单独幼仔在P18断奶,并每天观察死亡率。在混合129Stac X C57BL/6背景中的P26-P28 HET和WT Scn1a小鼠中,进行高热癫痫发作诱导。在试验之前,向小鼠插入润滑的直肠温度探头 (Ret-4),其连接至与加热灯 (HL-1) 串联连接的温度控制模块 (TCAT 2DF,Physitemp)。然后将小鼠放入一个大玻璃烧杯中,使其短暂适应环境。此后,每2分钟将体温升高约0.5℃,直到出现第一次伴有姿势丧失的强直-阵挛性癫痫发作或达到43℃为止。如果小鼠发生癫痫发作伴姿势丧失,则结束实验并记录小鼠的体内温度。如果在整个实验过程中没有检测到癫痫发作伴姿势丧失,则认为该小鼠没有癫痫发作,并结束试验。在P1从小鼠获得组织样品,并且在实验过程中使用实时PCR对小鼠进行基因型分型。完成试验后,对基因型分型进行揭盲,并将小鼠的状态HET或WT与获得的数据进行关联。将数据绘制成Kaplan-Meier存活曲线,并通过Mantel-Cox检验确定显著性。结果示于表36和表37以及图20A-E中。

[0559] 表36. 实施例11中使用的条件的总结

[0560]	eTF 的 SEQ ID NO	剂量 (gc/小鼠)	调节元件	被 eTF 识别 的靶位点	eTF 的 DBD+TAD 结 构	聚 A
	SEQ ID NO: 305	5.8E+10	SEQ ID NO: 185	SEQ ID NO: 35	ZF DBD + VPR	sPA
	SEQ ID NO: 306	6.0E+10	SEQ ID NO: 185	SEQ ID NO: 35	ZF DBD + VP64	hGH
	SEQ ID NO: 307	1.4/1.7E+11	SEQ ID NO: 417	SEQ ID NO: 35	ZF DBD + VP64	sPA
	SEQ ID NO: 308	3.1E+11	SEQ ID NO: 185	SEQ ID NO: 35	ZF DBD + VP64	hGH
	SEQ ID NO: 309	4.3E+11	SEQ ID NO: 183	SEQ ID NO: 36	ZF DBD + VPR	sPA

[0561] 表37. 高热癫痫发作试验的结果的总结

[0562]	eTF 的 SEQ ID NO	对照动物数 (PBS 处理的)	治疗的 动物数	在 42.6 °C 时无 癫痫发 作的%	P 值
	EGFP 报 道分子	16	N/A	44%	
	SEQ ID NO: 305	16	14	93	P<0.001
	SEQ ID NO: 306	16	18	95	P<0.0001
	SEQ ID NO: 307	16	15	87	P<0.01
	SEQ ID NO: 308	16	21	76	P<0.05
	SEQ ID NO: 309	16	24	62.5	P<0.05

[0563] 实施例12

[0564] 在Dravet综合征小鼠模型中的存活试验

[0565] 为了测试转录激活物在Scn1a^{tm1K^{ea}}小鼠品系中的功效,在P1经双侧ICV,向雄性Scn1a^{+/+}与雌性C57Bl/6J杂交繁殖产生的幼仔给予AAV载体。使母鼠不受干扰,直到断奶。在P18断奶后,每天进行Scn1a^{+/+}小鼠的健康状态的观察。对于发现因任何原因死于饲养笼中的小鼠,记录日期。将数据绘制成Kaplan-Meier存活曲线,并通过Mantel-Cox检验确定显著性。

[0566] 结果示于表38和图21A-E中。

[0567] 表38:存活试验的条件和结果的总结

[0568]	SEQ ID	剂量 (gc/小鼠)	对照动物 数(PBS 处理的)	治疗的 动物数	P100 时的 %存活率 (*于 P83)	P 值
	PBS	N/A	53	N/A	49%	
	SEQ ID NO: 305	5.8E+10	53	29	76%	P<0.05
	SEQ ID NO: 306	6.0E+10	53	34	97%	P<0.0001
	SEQ ID NO: 307	1.7E+11	53	37	92%	P<0.0001
	SEQ ID NO: 309	4.3E+11	53	12	100%	P<0.001

[0569] 实施例13

[0570] 用编码SCN1A特异性转录因子的AAV处理后,非人灵长类动物中的SCN1A转录水平

[0571] 该研究使用了2至3岁的雄性食蟹猴(*macaca fascicularis*)。在加入研究之前,通过基于细胞的中和抗体试验,针对AAV9的交叉反应性抗体对动物进行预筛查。将表达SCN1A特异性转录因子(SEQ ID NO:305)的AAV9或对照在PBS中稀释,并以1.2E12 gc/动物进行实质内注射。为了注射,在每个半球中确定了三个不同的立体定位坐标,每只动物六个注射部

位。每个部位注射10ul体积。右半球的注射与左半球的注射对称。两只未经治疗的动物用作对照。

[0572] 为了评估Scn1A mRNA表达,进行了逆转录,随后进行qPCR方法。给药后28天尸检时,从对照和经治疗动物的大脑各个区域(额皮质、顶叶皮质、颞皮质、枕叶皮质、海马体、髓质,小脑;各200mg)采集组织切片,置于RNAlater中,然后冷冻。简言之,解剖30mg组织,提取RNA(使用Qiagen Rneasy Lipid tissue mini试剂盒,目录号1023539),通过逆转录(使用Applied Biosystems高容量cDNA逆转录试剂盒,目录号4368814)转化为cDNA,并使用针对Scn1A和管家基因GAPDH的引物/探针组(Applied Biosystems,目录号Rh02621745-gI FAM)进行qPCR。

[0573] 针对SCN1A的引物/探针组在下面给出。

[0574] 表39:在实施例13中使用的引物序列

基因	SEQ ID NO	序列(5'-3')	注意
[0575] Scn1A	SEQ ID NO: 328	CCATGGAACCTGGCTCGATTTCAC	F-引物
	SEQ ID NO: 439	ATTGGTGGGAGGCCACTGTAT	R-引物
	SEQ ID NO: 329	AGGCCTGAAAACCATTTGTGGGAGCCCT	探针(FAM)

[0576] 针对SCN1A的引物/探针组在下面给出。使用比较Ct (Δ Ct) 法,通过相对定量(RQ)确定每个测试样品中Scn1A的基因表达。该方法测量靶基因与管家基因之间的Ct差异(Δ Ct),然后将治疗样品的 Δ Ct值与对照样品进行比较。

[0577] Δ Ct = 靶基因的Ct平均值 - 管家基因的Ct平均值

[0578] $\Delta \Delta$ Ct = 治疗样品的 Δ Ct - 对照样品的 Δ Ct

[0579] 相对表达(治疗样品) = $2^{-\Delta \Delta Ct}$

[0580] 数据被报告为目标mRNA在来自脑的不同组织切片中的归一化表达(参见图22)。如图22所示,在实质内注射部位近端的大脑部位显示出最高水平的SCN1A转录物表达。

[0581] 实施例14

[0582] 在不同小鼠系中治疗Dravet综合征

[0583] 使用本文所述的表达盒对Dravet综合征和/或其症状的治疗可以在各种小鼠系中测试,如B6(Cg) - Scn1a^{tm1.1Dsf}/J、Scn1a^{tm1Kea}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^{tm1Kea}和Scn1a-R1470X小鼠系不需要CRE重组酶。

[0584] B6(Cg) - Scn1a^{tm1.1Dsf}/J小鼠可以通过Jackson Laboratories从Dravet Syndrome European Federation获得,以研究本文描述的SCN1A转录激活物组合物在治疗Dravet综合征中的安全性和功效。这些小鼠在SCN1A的外显子24中具有与Dravet综合征相关的突变(位置1783处的A至V)。该小鼠还含有具有野生型序列的侧翼的(floxed)外显子24。当不进行操作时,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或将该品系与表达Cre的小鼠品系杂交后,任何表达Cre的细胞都将转换为表达突变等位基因的一个拷贝。突变SCN1A亚单位的表达后,小鼠在4周内发展成自发癫痫发作。对于B6(Cg) - Scn1a^{tm1.1Dsf}/J小鼠实验,可以使用B6(Cg) - Scn1a^{tm1.1Dsf}/J小鼠和对照C57B16小鼠。

[0585] Scn1a^{tm1Kea}小鼠(可从Jackson Laboratory获得;在Hawkins等人,Scientific

Reports, vol. 7:15327 (2017) 中描述) 具有SCN1A的第一编码外显子的缺失。SCN1A敲除等位基因的纯合小鼠的特征为震颤、共济失调、癫痫发作,并在出生后第16天死亡。C57BL/6背景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来研究癫痫和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。

[0586] Scn1a-R1470X小鼠是基因敲入小鼠,其携带在SCN1A基因的外显子21中的提前终止密码子R1407X。在三名无关的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。其他信息参见Ogiwara等人, Journal of Neuroscience, May 30, 2007, 27(22) 5903-5914。

[0587] 为了测试包含在体内上调SCN1A表达的eTF的本文所述组合物,包括AAV基因疗法和使用这样的基因疗法的治疗,向上述每个小鼠品系的Dravet小鼠和对照小鼠(例如,该品系的具有野生型SCN1A的小鼠或未经治疗的Dravet小鼠)注射(例如,通过腹膜内注射施用)表达EGFP或另一报道基因的AAV,或包含与内源SCN1A近端的靶结合序列结合的eTF的表达盒(例如,本文公开的任何表达盒)。一些AAV可进一步包含一个或多个细胞类型选择性调节元件。

[0588] 注射AAV后,随时间监测小鼠的存活。每天监测所有小鼠的一般健康状况(例如体重、补水、理毛行为和活动能力),并记录死亡。

[0589] 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,连续记录并监测脑皮层电图数据至少14天。所有癫痫发作事件可在AAV治疗后记录至少14天,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对照或未经治疗的对照相比,用包含SCN1A转录激活物的AAV治疗后癫痫发作的频率和/或持续时间的减少表明SCN1A转录激活物在减轻Dravet综合征的症状和/或严重程度方面的功效。

[0590] 在用AAV处理小鼠后,可使用各种PCR和/或测序方法随时间监测SCN1A的表达水平,以显示采用SCN1A转录激活物的AAV处理可导致内源SCN1A表达的增加。还可以使用Northern印迹分析和原位杂交来分析体内SCN1A表达。还可以在处理后监测Nav1.1蛋白的水平,以显示SCN1A表达的增加与Nav1.1蛋白的增加相关。Nav1.1蛋白可以采用多种方法来测定,包括但不限于Western印迹分析、免疫组织化学、免疫荧光组织化学和/或ELISA测定。功能性电压门控钠离子通道的形成还可以使用电流钳分析来测定。

[0591] 可以评价高热诱发的癫痫发作,以将野生型小鼠和/或未经治疗的Dravet小鼠与用包含本文所述表达盒的AAV基因疗法治疗的Dravet小鼠进行比较。在这类实验中,核心体温用RET-3直肠温度探头(Physitemp Instruments, Inc, New Jersey, USA)来监测,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)的加热灯来控制。每两分钟将体温升高0.5℃,直到第一次阵挛性抽搐发作。与未经治疗的Dravet小鼠相比,用包含SCN1A转录激活物的AAV治疗的Dravet小鼠在首次阵挛性抽搐发作前预期具有更高的阈值体温。

[0592] 还可以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谱。这些临床前研究还可以为用于治疗Dravet综合征提供基因疗法的最佳剂量。

[0593] 实施例15**[0594] 小鼠中阿尔茨海默病的治疗**

[0595] 可以使用在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) 小鼠来研究本文所述的SCN1A转录激活物组合物在治疗阿尔茨海默病中的安全性和功效。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) 和具有L166P突变的早老蛋白1 (PSEN1) 二者的人类转基因,二者均在Thy1启动子的控制下。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见Radde等人,2006 (Radde, Rebecca等人“A β 42-driven cerebral amyloidosis in transgenic mice reveals early and robust pathology.”EMBO reports 7.9(2006):940-946)。

[0596] 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SCN1A转录激活物(例如,本文公开的任何表达盒)的治疗AAV载体。一些AAV可以进一步包含GABA能选择性和/或PV选择性调节元件。

[0597] 注射AAV后,随时间监测小鼠的存活。每天监测所有小鼠的一般健康状况(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,也向小鼠中植入EET发射机。术后至少4周,经24小时记录并监测脑活动。可以自动分析脑皮层电图数据,并且可以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,每次治疗均采用如上所述的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,表明采用包含SCN1A转录激活物的AAV基因疗法的治疗有效。

[0598] 在用AAV处理小鼠后,可使用各种PCR和/或测序方法随时间监测SCN1A的表达水平,以显示采用SCN1A转录激活物的AAV处理可导致内源SCN1A表达的增加。还可以使用Northern印迹分析和原位杂交来分析体内SCN1A表达。还可以在处理后监测Nav1.1蛋白的水平,以显示SCN1A表达的增加与Nav1.1蛋白的增加相关。Nav1.1蛋白可以采用多种方法来测定,包括但不限于Western印迹分析、免疫组织化学、免疫荧光组织化学和/或ELISA测定。功能性电压门控钠离子通道的形成还可以使用电流钳分析来测定。

[0599] 还可以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谱。这些临床前研究还可以为用于治疗阿尔茨海默病提供基因疗法的最佳剂量。

[0600] 实施例16**[0601] 使用GRN特异性转录激活物,能够上调GRN的靶区域的鉴定**

[0602] 为了鉴定能够上调内源GRN表达的基因组区域,设计了各种工程化转录因子(锌指核酸酶或gRNA/daCas9构建体),它们靶向如以下表40所示的基因组的各个区域。对于gRNA/daCas9构建体,由于将gRNA设计为靶向互补基因组链,因此该gRNA具有与靶区域相同的序列。按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3 μ g携带工程化转录因子或对照构建体的质粒进行转染(FugeneHD, Promega)。用表达以下表40和表41中所示的构建体的质粒转染细胞。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3 μ g)进行逆转录。使用

Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98℃下30s,40x[98℃下10sec,66℃下15sec,72℃下15sec])。使用针对SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO:190);5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO:191))来量化内源SCN1A转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC' 3'(SEQ ID NO:192);5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定SCN1A表达的相对水平。数据以相对于对照条件的变化倍数呈现。

[0603] 结果以GRN转录相对于对照条件(例如,EGFP-KASH对照)的变化倍数示于图23和以下的表40中。

[0604] 表40:GRN靶区域及使用指定的转录因子对内源GRN上调的影响。

[0605]

eTF SEQ ID NO:	eTF序列	靶位点SEQ ID NO:
SEQ ID NO:337	(见上)	SEQ ID NO:330
A	(见表23)	SEQ ID NO:331
C	(见表23)	SEQ ID NO:332
M	(见表23)	SEQ ID NO:113
D	(见表23)	SEQ ID NO:333
E	(见表23)	SEQ ID NO:333
F	(见表23)	SEQ ID NO:38
J	(见表23)	SEQ ID NO:38
G	(见表23)	SEQ ID NO:38
B	(见表23)	SEQ ID NO:331
SEQ ID NO:338	(见上)	SEQ ID NO:331
H	(见表23)	SEQ ID NO:334
I	(见表23)	SEQ ID NO:335
SEQ ID NO:339	(见上)	SEQ ID NO:336
SEQ ID NO:340	(见上)	SEQ ID NO:38
SEQ ID NO:341	(见上)	SEQ ID NO:38
K	(见表23)	SEQ ID NO:38
L	(见表23)	SEQ ID NO:38
SEQ ID NO:342	(见上)	SEQ ID NO:38
SEQ ID NO:343	(见上)	SEQ ID NO:38
SEQ ID NO:344	(见上)	SEQ ID NO:335
SEQ ID NO:10	(见上)	SEQ ID NO:38
SEQ ID NO:16	(见上)	SEQ ID NO:38
SEQ ID NO:46	(见上)	SEQ ID NO:38
SEQ ID NO:345	(见上)	SEQ ID NO:38
SEQ ID NO:60	(见上)	SEQ ID NO:38
SEQ ID NO:346	(见上)	SEQ ID NO:38
SEQ ID NO:63	(见上)	SEQ ID NO:38
SEQ ID NO:64	(见上)	SEQ ID NO:38

SEQ ID NO:347	(见上)	SEQ ID NO:38
SEQ ID NO:348	(见上)	SEQ ID NO:38
SEQ ID NO:349	(见上)	SEQ ID NO:38
SEQ ID NO:350	(见上)	SEQ ID NO:38

[0606] 表41:GRN DBD

eTF SEQ ID NO:	DBD SEQ ID NO:
SEQ ID NO: 337	SEQ ID NO: 377
SEQ ID NO: 338	SEQ ID NO: 378
SEQ ID NO: 339	SEQ ID NO: 379
SEQ ID NO: 340	SEQ ID NO: 380
SEQ ID NO: 341	SEQ ID NO: 381

[0607]

eTF SEQ ID NO:	DBD SEQ ID NO:
SEQ ID NO: 342	SEQ ID NO: 382
SEQ ID NO: 343	SEQ ID NO: 383
SEQ ID NO: 344	SEQ ID NO: 384
SEQ ID NO: 345	SEQ ID NO: 385
SEQ ID NO: 346	SEQ ID NO: 386
SEQ ID NO: 347	SEQ ID NO: 387
SEQ ID NO: 348	SEQ ID NO: 388
SEQ ID NO: 349	SEQ ID NO: 389
SEQ ID NO: 350	SEQ ID NO: 390

[0608]

[0609] 实施例17

[0610] GRN从表达盒的相对表达

[0611] 该实施例描述了GRN从表达盒的相对表达,该表达盒包含增加HEK293细胞中GRN基因表达(如通过GRN RNA测量的)的非天然存在的转录激活物。构建了表达盒A、B、C、D、E、F、J、K、L、G、H、I和M,其各自包含非天然存在的转录调节物,该转录调节物包含与VPR或VP64转录激活域(例如,SEQ ID NO:95或SEQ ID NO:114)连接的选自SEQ ID NO:165-170和SEQ ID NO:112的DNA结合域,如以上表23中所述。每个表达盒还包含具有SEQ ID NO:178或179的序列的调节元件。表达盒J与表达盒K相似,不同之处在于J进一步包含eGFP报道分子。

[0612] 为了测试GRN转录激活物,按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3 μ g质粒进行转染(PEI)。转染后48h,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3 μ g)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98 $^{\circ}$ C下30s,40x[98 $^{\circ}$ C下10sec,66 $^{\circ}$ C下15sec,72 $^{\circ}$ C下15sec])。使用针对GRN的引物(5'-ATGGTCAGTTCTGCCCTGTG-3'(SEQ ID NO:287);5'-CGGTAAAGATGCAGGAGTGGC-3'(SEQ ID NO:288))来量化内源GRN转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC-3'(SEQ ID NO:192);5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定GRN表达的相对水平。

[0613] 对照是指包含SEQ ID NO:XX并且缺乏转录激活物的表达盒。图24A示出了用每种

表达盒转染的HEK293细胞中GRN的相对表达,呈现为相对于对照条件的变化倍数。该实验表明,如本文所述的转录激活物导致从内源GRN基因转录的RNA的上调。测试的所有转录激活物表达盒均导致HEK293细胞中GRN表达的增加。

[0614] 还进行了ELISA实验,以评价包含GRN的各种转录激活物的表达盒增加细胞中分泌的颗粒蛋白前体(hPGRN)蛋白的能力。293T细胞用表23中描述的某些表达盒转染,并在转染后48小时收集培养基。短暂离心以去除细胞碎片后,使用来自R&D Systems的人颗粒蛋白前体ELISA试剂盒,根据制造商的说明对培养基进行ELISA分析。基于试剂盒手册中描述的标准曲线分析,计算颗粒蛋白前体的浓度。图24B示出了根据上清液中的人颗粒蛋白前体(hPGRN)(ng/mL)检测的,对于每种表达盒获得的三个独立ELISA结果的平均值。相对于对照表达盒,测试的所有转录激活物表达盒均导致细胞中分泌的hPGRN的增加。

[0615] 实施例18

[0616] 增加HEK293T细胞中的基因表达

[0617] 用在几种不同调节元件(即无启动子对照、SCP、CMV、与minCMV可操作地连接的SEQ ID NO:178和CAG)之一的控制下含有萤光素酶基因的质粒DNA转染HEK293T细胞。每种构建体的归一化萤光素酶值在图25中示出。与单独的minCMV和单独的SCP相比,与minCMV启动子可操作地连接的调节元件SEQ ID NO:XX驱动更高水平的萤光素酶表达。

[0618] 该实验表明,SEQ ID NO:178是在细胞中驱动高基因表达的RE。可以将这样的RE添加到本文公开的包含GRN的非天然存在的转录激活物的表达盒中,以增加细胞中转录激活物的表达,这导致内源GRN基因的表达增加。在一些情况下,具有SEQ ID NO:178的序列的一个或多个RE与转录激活物可操作地连接,以增加细胞中的GRN表达。可以将这类调节元件添加到表达盒中转录激活物的上游和/或下游。

[0619] 实施例19

[0620] 使用颗粒蛋白前体特异性转录因子,HEK293细胞中颗粒蛋白前体转录物和蛋白质水平的上调

[0621] 按照标准方法培养HEK293细胞,并在6孔板的每个孔中用3ug质粒进行转染(FugeneHD, Promega)。用含有以下表42中所示的构建体的质粒转染细胞。转染后48h,从细胞中分离RNA用于qPCR分析,并从感染的孔中收集培养基,以通过ELISA分析对分泌的PGRN水平进行定量。

[0622] 对于qPCR分析,收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3ug)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98℃下30s, 40x[98℃下10sec, 66℃下15sec, 72℃下15sec])。使用针对PGRN的引物(5'-ATGGTCAGTTCTGCCCTGTG-3'(SEQ ID NO:287); 5'-CGGTAAAGATGCAGGAGTGGC-3'(SEQ ID NO:288))来量化内源PGRN转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC' 3'(SEQ ID NO:192); 5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定PGRN表达的相对水平。数据以相对于对照条件的变化倍数呈现(参见图26)。相对于EGFP对照条件,工程化转录因子的递送诱导内源GRN转录物的不同程度的上调。

[0623] 对于ELISA分析,在病毒感染后的指定日期收集细胞培养基,并使用人PGRN ELISA试剂盒(R&D systems, DPGRN0)进行ELISA分析。ELISA程序是按照制造商的说明书进行的。

数据以相对于对照条件的变化倍数呈现 (参见图27)。这些数据证实,工程化转录因子对GRN的上调导致PGRN蛋白的分泌增加。

[0624] 表42. 在实施例19中使用的GRN构建体

[0625]	eTF的SEQ ID NO
	SEQ ID NO:10
	SEQ ID NO:16
	SEQ ID NO:59
	SEQ ID NO:341
	SEQ ID NO:63
	SEQ ID NO:60
	SEQ ID NO:64
	SEQ ID NO:335
	SEQ ID NO:336

[0626] 实施例20

[0627] 使用颗粒蛋白前体特异性转录因子,GABA神经元中颗粒蛋白前体转录物和蛋白质水平的上调

[0628] 将iCell GABA神经元 (Cellular Dynamics) 接种于6孔板中 (约1E6个细胞/孔), 并且按照制造商推荐的方案进行维持。接种后72h, 以大约2E11个基因组拷贝/孔, 将在CBA启动子控制下表达EGFP或激活物 (SEQ ID NO:342) 的重组AAV (血清型AAV-DJ) 添加到培养基中。感染后一周, 从细胞中分离RNA用于qPCR分析, 并从感染的孔中收集培养基, 以通过ELISA分析对分泌的PGRN水平进行定量。

[0629] 对于qPCR分析, 从培养的细胞中分离RNA (Qiagen RNeasy Mini试剂盒), 并进行DNA酶处理。使用OligoDT引物 (Superscript IV, Invitrogen) 对回收的RNA进行逆转录。使用Phusion聚合酶 (New England Biolabs) 和SYBR Green I, 通过qPCR分析cDNA样品: (98℃下30s, 40x [98℃下10sec, 66℃下15sec, 72℃下15sec])。使用针对PGRN的引物 (5' - ATGGTCAGTTCTGCCCTGTG-3' (SEQ ID NO:287); 5' - CGGTAAAGATGCAGGAGTGGC-3' (SEQ ID NO:288)) 来量化内源PGRN转录物的水平, 并以GAPDH为参考基因, 通过delta-delta Ct法 (5' - ACCACAGTCCATGCCATCAC' 3' (SEQ ID NO:192); 5' - TCCACCACCCTGTTGCTGTA-3' (SEQ ID NO:193)) 确定PGRN表达的相对水平。数据以相对于对照条件的变化倍数呈现 (参见图28)。相对于EGFP对照条件, SEQ ID NO:342的递送产生内源GRN转录物的稳健上调。

[0630] 对于ELISA分析, 以0.5E11个基因组拷贝/孔 (低剂量条件) 或2E11个基因组拷贝/孔 (高剂量条件) 在病毒感染后的指定日期收集细胞培养基, 并使用人PGRN ELISA试剂盒 (R&D systems, DPGRN0) 进行ELISA分析。ELISA程序是按照制造商的说明书进行的。数据参见图29所示。这些数据证实, SEQ ID NO:342对GRN的上调导致培养的神经元中PGRN蛋白的分泌增加。

[0631] 实施例21

[0632] 在体内增加血浆GRN蛋白

[0633] 为了在体内测试GRN的转录激活物的表达, 将包含以上表23中所述的表达盒J或G的AAV9注射到小鼠中。纯化的AAV9病毒是在VectorBioLabs产生的, 并在注射前重建为

200uL总体积(用PBS)。适应期后,使用6至9周龄的C57BL6雄性动物进行尾静脉注射。通过尾静脉放血进行寿命内血液采集,并使用K2EDTA法生成血浆样品。使用小鼠PGRN ELISA试剂盒(Thermo Fisher Scientific),根据制造商的手册进行ELISA分析。图30示出了对于用不同AAV9构建体或对照处理的小鼠,在注射后一周,以 $\mu\text{g/mL}$ 为单位的血浆GRN的ELISA结果。与PBS缓冲液对照相比,用包含GRN转录激活物的AAV9处理的所有小鼠均显示出血浆GRN蛋白水平的显著增加。测试了不同剂量的AAV9-J和AAV9-G:每只小鼠1E11个基因组拷贝(GC)和5E11个GC。对于AAV9-G,与1E11个GC/小鼠的较低剂量相比,增加每只小鼠注射的基因组拷贝导致血浆GRN蛋白至少增加两倍。

[0634] 实施例22

[0635] 推动神经胶质细胞中的高GRN表达

[0636] 可以使用荧光成像确定对神经胶质细胞,或者特别是对小神经胶质细胞的选择性。将含有eGFP报道基因的AAV载体与对照启动子(EF1 α)或神经胶质细胞或小神经胶质细胞选择性RE可操作地连接。将这类AAV载体与CRE依赖性tdTomato载体共注射到CRE小鼠中,其中Cre特别在神经胶质细胞或小神经胶质细胞中表达。

[0637] 以0.3 $\mu\text{L/min}$ 的速率将1.5 μL 的AAV载体(5^{12} 至 1^{13}gc/mL)双侧输注到小鼠的额皮质或纹状体中,注射后有4min的休息期。治疗后,每天监测小鼠的健康和体重。对于组织采集,通过过量的异氟烷对小鼠实施安乐死,并灌注4%低聚甲醛(PFA)。为了分析报道基因在神经胶质细胞或小神经胶质细胞中的选择性表达,从额皮质或纹状体中获取一块脑组织,切片,并使用抗RFP多克隆兔抗体(Rockland Antibodies and Assay)和抗eGFP多克隆鸡抗体(Aves Labs)以标准免疫组织化学程序进行eGFP和tdTomato染色。使用荧光显微术,与对照相比,eGFP报道基因的荧光与tdTomato荧光之间的强烈重叠表明在神经胶质细胞或小神经胶质细胞中的选择性表达。

[0638] 该方法可用于鉴定靶向受GRN缺乏影响的额皮质或其他CNS细胞如浦肯野细胞、锥体细胞(例如,Betz细胞)、运动神经元和大脑皮质神经元中的高基因表达的调节元件。

[0639] 一旦已经鉴定出对神经胶质细胞、小神经胶质细胞或CNS中受GRN缺乏负面影响的其他任何细胞类型具有选择性的调节元件,此类调节元件就可以与本文所述的表达盒中靶向GRN基因的转录激活物可操作地连接,以在体内选择性地增加在靶细胞或组织类型中的GRN表达。例如,对神经胶质细胞或小神经胶质细胞具有选择性的RE可以与本文所述的表达盒的任何一种转录激活物可操作地连接,以在体内选择性地增加GRN。

[0640] 实施例23

[0641] 在体内增加血浆GRN蛋白

[0642] 为了在体内测试GRN的转录激活物的表达,使用了包含本文所述的任何表达盒的AAV9。纯化的AAV9病毒可以在VectorBioLabs产生,并在注射前重建为200uL总体积(用PBS)。适应期后,可以使用6至9周龄的C57BL6雄性动物进行尾静脉注射。可以通过尾静脉放血进行寿命内血液采集,并且可以使用K2EDTA法生成血浆样品。可以使用小鼠GRN ELISA试剂盒(Thermo Fisher Scientific),根据制造商的手册进行ELISA分析。与PBS缓冲液对照相比,用包含GRN转录激活物的AAV9处理的所有小鼠均预期显示出血浆GRN蛋白水平的显著增加。可以测试不同剂量的AAV9-eTF,如每只小鼠1E11个基因组拷贝(GC)和5E11个GC。

[0643] 实施例24

[0644] 治疗小鼠的FTD症状

[0645] 表现出FTD症状,即GRN单倍体不足的杂合GRN敲除小鼠,提供了用于验证本文所述的组合物和方法的体内动物模型。将本文所述的AAV表达盒注射到小鼠中,以通过增加体内GRN的表达水平来拯救GRN单倍体不足。一旦通过注射/输注将表达盒递送到小鼠中,就可以使用各种方法随时监测GRN单倍体不足的拯救,这些方法例如是测定血液样品中分泌的GRN的水平,使用PCR方法测定GRN的转录,和/或测定与对照组和/或未治疗的组相比,经治疗的小鼠的行为改变(例如,运动协调、学习、与其他小鼠的社交行为以及其他认知功能)。

[0646] 类似的方法可以在阿尔茨海默病、帕金森病和动脉粥样硬化的小鼠模型中使用(例如,使用本文所述的表达盒测量小鼠治疗后动脉壁增厚的减少)。

[0647] 实施例25

[0648] 报道系统筛查

[0649] 按照标准方法培养HEK293细胞,并用1.5ug由四环素响应元件(TRE)紧密启动子控制下的EGFP组成的报道质粒和1.5ug每种测试激活物或空载体对照(pUC57)进行共转染(FuGene HD, Promega)。转染后48h,使用落射荧光显微镜针对GFP表达对细胞进行成像。收集细胞并分离RNA(Qiagen RNeasy Mini试剂盒),并进行DNA酶处理。使用OligoDT引物(Superscript IV, Invitrogen)对RNA(3ug)进行逆转录。使用Phusion聚合酶(New England Biolabs)和SYBR Green I,通过qPCR分析cDNA样品:(98°C下30s, 40x[98°C下10sec, 65°C下15sec, 72°C下15sec])。使用针对EGFP的引物(5'-GCTACCCCGACCACATGAAG-3'(SEQ ID NO:369); 5'-TCTGTAGTTGCCGTCGTCC-3'(SEQ ID NO:370))来量化报道基因驱动的EGFP转录物的水平,并以GAPDH为参考基因,通过delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC-3'(SEQ ID NO:192); 5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO:193))确定EGFP表达的相对水平。

[0650] 数据以相对于对照条件的变化倍数呈现。参见图31A和图31B,示出了与空载体对照相比,SEQ ID NO:2-4的eTF导致增加的EGFP表达。包含病毒VP64 TAD的SEQ ID NO:51的eTF导致最高水平的相对EGFP表达。如图32所示,与对照SEQ ID NO:187相比,包含SEQ ID NO:12的eTF导致至少5倍的相对EGFP表达。图33示出了与单独的报道分子或没有TAD的对照蛋白质(SEQ ID NO:50)相比,SEQ ID NO:51-53的eTF导致超过10倍的相对EGFP表达。

[0651] 本公开的各个实施方案参考以下编号的条目来限定:

[0652] 1.一种核酸盒,其编码包含三个或更多个DNA结合域的非天然存在的DNA结合蛋白,其中所述非天然存在的DNA结合蛋白将内源基因的表达提高或抑制至少1.5、2、3、4、5、6、7、8、9、10、20、25、50或100倍,其中所述非天然存在的DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性。

[0653] 2.一种编码非天然存在的DNA结合蛋白的核酸盒,所述非天然存在的DNA结合蛋白包含共同地与人类基因组中的至少21个碱基结合的多个结合域,并且与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性。

[0654] 3. 条目1或2的核酸盒,其中所述DNA结合蛋白与所述人类对应物不结合的靶位点结合。

[0655] 4. 条目1-3中任一项的核酸盒,其中所述人类对应物是EGR1或EGR3。

[0656] 5. 条目1-4中任一项的核酸盒,其中所述结合域是锌指。

[0657] 6. 条目1-5中任一项的核酸盒,其中所述结合域包含人类对应物的结合域的双重复制或三重复制。

[0658] 7. 条目1-5中任一项的核酸盒,其中所述结合域包含人类对应物的DNA结合域或锌指的2、3、4、5、6、7、8或9个拷贝。

[0659] 8. 条目1-7中任一项的核酸盒,其中所述非天然存在的DNA结合蛋白包含至少6、7、8、9、10、11或12个锌指。

[0660] 9. 条目1-8中任一项的核酸盒,其中所述内源基因是SCN1A或GRN。

[0661] 10. 条目1-9中任一项的核酸盒,其中所述DNA结合蛋白识别基因组位置处的结合位点,使得在该位点处的结合允许所述DNA结合蛋白调节所述内源基因的表达。

[0662] 11. 条目1-10中任一项的核酸盒,其中所述DNA结合蛋白识别位于内源SCN1A或GRN基因座处或附近的结合位点。

[0663] 12. 条目1-11中任一项的核酸盒,其中所述DNA结合蛋白识别具有以下序列的结合位点: (i) SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336; (ii) 其变体; (iii) 与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列;或(iv)在其5bp、10bp、50bp、100bp、200bp或500bp内的基因组区域。

[0664] 13. 条目1-11中任一项的核酸盒,其中所述DNA结合蛋白识别具有以下序列的结合位点: (i) SEQ ID NO:36或38; (ii) 其变体; (iii) 与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列;或(iv)在其5bp、10bp、50bp、100bp、200bp或500bp内的基因组区域。

[0665] 14. 条目1-12中任一项的核酸盒,其中所述DNA结合蛋白具有以下序列: (i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416; (ii) 其变体或功能片段;或(iii)与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0666] 15. 条目1-11和13中任一项的核酸盒,其中所述DNA结合蛋白具有以下序列: (i) SEQ ID NO:6-10、13-16、57-61、63、64、67-73、74-77、268-282、295-299、337-350或365-366; (ii) 其变体或功能片段;或(iii)与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0667] 16. 条目1-15中任一项的核酸盒,其中所述DNA结合蛋白包含衍生自EGR1或EGR3的转录效应物结构域。

[0668] 17. 条目1-15之一的核酸盒,其中所述DNA结合蛋白包含衍生自CITED2或CITED4的

转录效应物结构域。

[0669] 18. 条目16-17之一的核酸盒,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的C-末端。

[0670] 19. 条目16-17之一的核酸盒,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的N-末端。

[0671] 20. 条目1-19之一的核酸盒,其中所述DNA结合域共同包含以下序列:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0672] 21. 条目1-20之一的核酸盒,其中所述DNA结合蛋白包含具有以下序列的效应物结构域:(i) SEQ ID NO:95-100或114;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

[0673] 22. 条目1-21之一的核酸盒,其中所述DNA结合域在所述结构域中锌指的识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。

[0674] 23. 条目1-22之一的核酸盒,其中所述DNA结合蛋白识别18nt或更长、21nt或更长或者24nt或更长的靶结合位点。

[0675] 24. 条目1-23之一的核酸盒,其中所述序列同一性使用BLAST或ClustalW测量。

[0676] 25. 条目1-24之一的核酸盒,其中所述DNA结合蛋白在病毒载体中。

[0677] 26. 条目25的核酸盒,其中所述病毒载体是AAV。

[0678] 27. 条目26的核酸盒,其中所述AAV是AAV9。

[0679] 28. 条目1-27中任一项的核酸盒,其中如通过elispot分析所测定的,所述非天然存在的DNA结合蛋白具有与其人类对应物相当的免疫原性(例如,其中相当的是指在由所述天然存在的人类蛋白质产生的免疫应答的1%、2%、3%、4%、5%、6%、7%、8%、9%或10%内的免疫原性水平)。

[0680] 29. 一种AAV载体,其包含编码非天然存在的DNA结合蛋白的核酸盒,所述非天然存在的DNA结合蛋白与人类蛋白质具有>90%的序列同一性,或者与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性,并且所述AAV载体包含DNA结合域,所述结构域包含共同结合>9、>10、>11、>12、>13、>14、>15、>16、>17、>18、>19、>20、>21、>22、>23、>24、>25、>26、>27、>28、>29或>30个碱基的多个结合基序。

[0681] 30. 条目29的AAV载体,其中所述非天然存在的DNA结合蛋白与人类基因组中所述人类蛋白质不结合的靶位点结合。

[0682] 31. 条目29-30中任一项的AAV载体,其中所述人类蛋白质是EGR1或EGR3。

[0683] 32. 条目29-31中任一项的AAV载体,其中所述DNA结合域包含至少6、7、8、9、10、11或12个锌指。

[0684] 33. 条目29-32中任一项的AAV载体,其中所述DNA结合域包含所述人类蛋白质中的

DNA结合域的双重复制、三重复制或四重复制。

[0685] 34. 条目29-33中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白识别SCN1A或GRN的基因组基因座(例如,表1中所述的基因组基因座)处或附近的靶结合位点,或内源SCN1A或GRN的起始密码子上游的靶结合位点。

[0686] 35. 条目29-34中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白调节内源SCN1A或GRN的表达(例如,增加其表达)。

[0687] 36. 条目29-35中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白识别包含以下序列的结合位点:(i) SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336;(ii) 其变体;(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列;或(iv)在其5bp、10bp、50bp、100bp、200bp或500bp内的基因组区域。

[0688] 37. 条目29-36中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白包含以下序列:(i) SEQ ID NO:6-9、13-15、44-45、48-49、54-55、58-62、67-77、103、112、114、268-282、305-325、337-350、364、295-299、365-366;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0689] 38. 以上条目中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白包含以下序列:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0690] 39. 条目29-38中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白包含EGR1或EGR3的转录效应物结构域。

[0691] 40. 条目29-39中任一项的AAV载体,其中所述非天然存在的DNA结合蛋白包含CITED2或CITED4的转录效应物结构域。

[0692] 41. 条目39-40中任一项的AAV载体,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的C-末端。

[0693] 42. 条目39-41中任一项的AAV载体,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的N-末端。

[0694] 43. 条目29-42中任一项的AAV载体,其中所述DNA结合域包含以下序列:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0695] 44. 条目29-43中任一项的AAV载体,其中所述DNA结合蛋白包含具有以下序列的效应物结构域:(i) SEQ ID NO:95-100或114;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0696] 45. 条目29-44中任一项的AAV载体,其中所述DNA结合域在所述结构域中锌指的识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。

[0697] 46. 条目29-45中任一项的AAV载体,其中所述DNA结合蛋白识别18nt或更长、21nt

或更长或者24nt或更长的靶结合位点。

[0698] 47. 条目29-46中任一项的AAV载体,其中如通过elispot分析所测定的,所述非天然存在的DNA结合蛋白具有与所述人类蛋白质相当的免疫原性(例如,其中相当的是指在由所述天然存在的人类蛋白质产生的免疫应答的1%、2%、3%、4%、5%、6%、7%、8%、9%或10%内的免疫原性水平)。

[0699] 48. 条目29-47中任一项的AAV载体,其中所述序列同一性使用BLAST或ClustalW测量。

[0700] 49. 条目29-48中任一项的AAV载体,其中所述AAV是AAV9。

[0701] 50. 一种编码非天然存在的DNA结合蛋白的载体,所述非天然存在的DNA结合蛋白与人类蛋白质具有90%或更高的序列同一性,或者与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性,其中所述DNA结合蛋白能够调节内源人类基因的表达。

[0702] 51. 条目50的载体,其中所述非天然存在的DNA结合蛋白包含识别至少9bp长的靶结合位点的DNA结合域。

[0703] 52. 条目50-51任一项的载体,其中所述非天然存在的DNA结合蛋白包含识别至少12bp、15bp、18bp、21bp或24bp长的靶结合位点的DNA结合域。

[0704] 53. 条目50-52中任一项的载体,其中所述DNA结合蛋白与人类基因组中不被人EGR1或EGR3识别的靶位点结合。

[0705] 54. 条目50-53中任一项的载体,其中所述内源人类基因是SCN1A或GRN。

[0706] 55. 条目50-54中任一项的载体,其中所述DNA结合蛋白识别SCN1A或GRN的基因组基因座(例如,表1中所述的基因组基因座)处或附近的靶结合位点,或内源SCN1A或GRN的起始密码子上游的靶结合位点。

[0707] 56. 条目50-55中任一项的载体,其中所述DNA结合蛋白包含至少6、7、8、9、10、11或12个锌指。

[0708] 57. 条目50-56中任一项的载体,其中所述DNA结合蛋白包含DNA结合域的双重复制、三重复制或四重复制或EGR1或EGR3的一个或多个锌指。

[0709] 58. 条目50-57中任一项的载体,其中所述DNA结合蛋白识别包含以下序列的结合位点:(i) SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336;(ii) 其变体;(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列;或(iv)在其5bp、10bp、50bp、100bp、200bp或500bp内的基因组区域。

[0710] 59. 条目50-58中任一项的载体,其中所述DNA结合蛋白包含以下序列:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0711] 60. 上述条目中任一项的载体,其中所述DNA结合蛋白包含以下序列:(i) SEQ ID

N0:6-9、13-15、44-45、48-49、54-55、58-62、67-77、103、112、114、268-282、305-325、337-350、364、295-299或365-366; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0712] 61. 条目50-60中任一项的载体,其中所述序列同一性是指所述DNA结合蛋白中的DNA结合域与EGR1或EGR3的DNA结合域的局部序列同一性。

[0713] 62. 条目61的载体,其中所述DNA结合蛋白包含衍生自CITED2或CITED4的转录效应物结构域。

[0714] 63. 条目50-62中任一项的载体,其中所述序列同一性是指所述DNA结合蛋白与人EGR1或EGR3相比的全局序列同一性。

[0715] 64. 条目63的载体,其中所述DNA结合蛋白包含衍生自EGR1或EGR3的转录效应物结构域。

[0716] 65. 条目62或64的载体,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的C-末端。

[0717] 66. 条目62或64的载体,其中所述效应物结构域位于所述DNA结合蛋白中的DNA结合域的N-末端。

[0718] 67. 条目50-66中任一项的载体,其中所述DNA结合蛋白包含具有以下序列的DNA结合域: (i) SEQ ID N0:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0719] 68. 条目50-67中任一项的载体,其中所述效应物结构域具有以下序列: (i) SEQ ID N0:95-100或114; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0720] 69. 条目50-68中任一项的载体,其中所述锌指在所述锌指的识别螺旋的位置-1、2、3或6处包含至少一个氨基酸置换。

[0721] 70. 条目50-69中任一项的载体,其中如通过elispot分析所测定的,所述DNA结合蛋白具有与人EGR1或ERG3相当的免疫原性(例如,其中相当的是指在由所述天然存在的人类蛋白质产生的免疫应答的1%、2%、3%、4%、5%、6%、7%、8%、9%或10%内的水平)。

[0722] 71. 条目50-70中任一项的载体,其中所述序列同一性使用BLAST或ClustalW测量。

[0723] 72. 条目70-72中任一项的载体,其中所述载体是病毒载体。

[0724] 73. 条目72的载体,其中所述病毒载体是AAV。

[0725] 74. 条目73的载体,其中所述AAV是AAV9。

[0726] 75. 一种编码非天然DNA结合蛋白的核酸盒,所述非天然DNA结合蛋白与天然存在的人类蛋白质具有90%或更高的序列同一性,或者与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性,其中

所述DNA结合蛋白能够调节SCN1A的表达。

[0727] 76. 条目75的核酸盒,其中所述非天然存在的DNA结合蛋白选择性地结合具有SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336中任一个的区域,或其5bp、10bp、50bp、100bp、200bp或500bp内的区域。

[0728] 77. 条目75-76中任一项的核酸盒,其中所述非天然存在的DNA结合蛋白与对照相比能够使SCN1A的表达增加至少1.5、2、3、4、5、6、7、8、9、10、15、20、50或100倍。

[0729] 78. 条目75-77中任一项的核酸盒,其中所述非天然存在的DNA结合蛋白包含具有以下序列的DNA结合域:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0730] 79. 条目75-78中任一项的核酸盒,其中所述核酸盒是病毒载体。

[0731] 80. 条目75-79中任一项的核酸盒,其中所述核酸盒是AAV载体,如AAV9。

[0732] 81. 一种编码非天然DNA结合蛋白的核酸盒,所述非天然DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的全局序列同一性,其中所述DNA结合蛋白能够调节GRN的表达。

[0733] 82. 条目81的核酸盒,其中所述非天然存在的DNA结合蛋白选择性地结合具有SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336的区域,或其5bp、10bp、50bp、100bp、200bp或500bp内的区域。

[0734] 83. 条目81-82中任一项的核酸盒,其中所述非天然存在的DNA结合蛋白与对照相比能够使GRN的表达增加至少1.5、2、3、4、5、6、7、8、9、10、15、20、25、50或100倍。

[0735] 84. 条目81-83中任一项的核酸盒,其中所述非天然存在的DNA结合蛋白包含具有以下序列的DNA结合域:(i) SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0736] 85. 条目81-84中任一项的核酸盒,其中所述DNA结合蛋白包含以下序列:(i) SEQ ID NO:6-9、13-15、44-45、48-49、54-55、58-62、67-77、103、112、114、268-282、305-325、337-350、364、295-299、365-366;(ii) 其变体或功能片段;或(iii) 与(i)或(ii)中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0737] 86. 条目81-85中任一项的核酸盒,其中所述核酸盒是病毒载体。

[0738] 87. 条目81-86中任一项的核酸盒,其中所述核酸盒是AAV载体。

[0739] 88. 编码调节SCN1A的蛋白质的载体,其中所述蛋白质具有SEQ ID NO:6-9、13、15、57、58、61-62、67-71、74-75、268-282、295-299、305-325或364-366的氨基酸序列。

[0740] 89. 编码调节SCN1A的蛋白质的载体,其中所述蛋白质包含具有SEQ ID NO:22-25、

29-31、84-85、88、90-92、131-135、371-372、376或391-409的序列的DNA结合域。

[0741] 90. 条目88-89中任一项的载体,其中所述蛋白质识别包含SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267的序列的靶结合序列。

[0742] 91. 编码调节GRN的蛋白质的载体,其中所述蛋白质具有SEQ ID NO:10、16、59-60、63-64、72-73、76-77、112或337-350的氨基酸序列。

[0743] 92. 编码调节GRN的蛋白质的载体,其中所述蛋白质包含具有SEQ ID NO:26、32、86-89、93、165-171、112或377-390的序列的DNA结合域。

[0744] 93. 条目91-92中任一项的载体,其中所述蛋白质识别包含SEQ ID NO:38、113或330-336的序列的靶结合序列。

[0745] 94. 条目91-93中任一项的载体,其中所述DNA结合域与CITED2或CITED4的反式激活域融合。

[0746] 95. 一种治疗疾病或病况的方法,其包括使细胞或受试者与包含条目1-94中任一项的组合物接触。

[0747] 96. 条目95的方法,其中所述受试者是动物、哺乳动物或人。

[0748] 97. 一种治疗疾病或病况的方法,其包括施用基因疗法,该基因疗法包括编码包含DNA结合域的非天然存在的DNA结合蛋白的核酸盒,其中所述非天然存在的DNA结合蛋白调节内源基因的表达至少5倍,并且其中所述非天然存在的DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有全局序列同一性。

[0749] 98. 条目97的方法,其中当在细胞中或在体内表达时,所述非天然存在的DNA结合蛋白引发降低的或最小的免疫应答。

[0750] 99. 条目98的方法,其中所述非天然存在的DNA结合蛋白引发与天然存在的人类蛋白质相当的免疫应答,其中相当的是指在由所述天然存在的人类蛋白质(例如,EGR1或EGR3)产生的免疫应答的1%、2%、3%、4%、5%、6%、7%、8%、9%或10%内的水平。

[0751] 100. 条目92-99中任一项的方法,其中所述免疫应答使用elispot分析来测定。

[0752] 101. 条目97-100中任一项的方法,其中所述疾病或病况是Dravet综合征。

[0753] 102. 条目97-100中任一项的方法,其中所述疾病或病况是额颞痴呆或痴呆。

[0754] 103. 条目97-100中任一项的方法,其中所述疾病或病况选自:Dravet综合征、帕金森病、阿尔茨海默病、GABA能功能减退、神经元活动过度、癫痫和癫痫发作。

[0755] 104. 条目97-100中任一项的方法,其中所述疾病或病况选自:神经变性、神经炎症、额颞叶退化、额颞痴呆、痴呆、帕金森病、阿尔茨海默病和动脉粥样硬化。

[0756] 105. 条目97-104中任一项的方法,其中所述DNA结合蛋白具有SEQ ID NO:6-9、13-15、44-45、48-49、54-55、58-62、67-77、103、112、114、268-282、305-325、337-350、364、295-299、365-366的氨基酸序列。

[0757] 106. 条目97-105中任一项的方法,其中所述DNA结合蛋白具有SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416的氨基酸序列。

[0758] 107. 上述条目中任一项的方法,其中所述DNA结合蛋白包含具有SEQ ID NO:22-

- 26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416的序列的DNA结合域。
- [0759] 108. 条目97-107中任一项的方法,其中所述DNA结合蛋白识别包含SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336的序列的靶结合序列。
- [0760] 109. 上述条目中任一项的方法,其中所述内源基因是SCN1A或GRN。
- [0761] 110. 上述条目中任一项的方法,其中所述基因疗法是AAV。
- [0762] 111. 一种调节内源基因表达的方法,该方法包括施用编码非天然存在的DNA结合蛋白的核酸盒,所述非天然存在的DNA结合蛋白包含共同结合人类基因组中具有至少21个碱基的靶结合位点的多个DNA结合域,并且与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有全局序列同一性。
- [0763] 112. 条目111的方法,其中所述天然存在的蛋白质是人EGR1或EGR3。
- [0764] 113. 条目111-112中任一项的方法,其中所述DNA结合蛋白包含CITED2或CITED4的结构域。
- [0765] 114. 一种治疗疾病或病况的方法,该方法包括施用包含非天然存在的DNA结合蛋白的AAV载体,所述非天然存在的DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有全局序列同一性,并且包含至少6、7、8、9、10、11或12个锌指或DNA结合域。
- [0766] 115. 条目114的方法,其中所述人类蛋白质是EGR1或EGR3。
- [0767] 116. 条目114-115中任一项的方法,其中所述疾病或病况是CNS疾病或病况。
- [0768] 117. 条目116的方法,其中所述疾病或病况是Dravet综合征。
- [0769] 118. 条目116的方法,其中所述疾病或病况是痴呆或额颞痴呆。
- [0770] 119. 条目114-115中任一项的方法,其中所述序列同一性使用用于全局序列同一性的Needleman-Wunsch算法来测量。
- [0771] 120. 上述条目中任一项的方法,其中所述序列同一性使用用于局部序列同一性的Smith-Waterman算法来测量。
- [0772] 121. 一种治疗Dravet综合征的方法,其包括施用编码非天然存在的DNA结合蛋白的载体,所述非天然存在的DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有全局序列同一性,其中所述DNA结合蛋白能够激活或增加内源SCN1A的表达。
- [0773] 122. 条目121的方法,其中所述非天然存在的DNA结合蛋白选择性地结合具有SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336中任一个的区域,或其5bp、10bp、50bp、100bp、200bp或500bp内的区域。
- [0774] 123. 条目121-122中任一项的方法,其中所述非天然存在的DNA结合蛋白与对照相

比能够使SCN1A的表达增加至少1.5、2、3、4、5、6、7、8、9、10、15、20、25、50或100倍。

[0775] 124. 条目121-123中任一项的方法,其中所述非天然存在的DNA结合蛋白包含具有以下序列的DNA结合域: (i) SEQ ID NO: 22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0776] 125. 条目121-124中任一项的方法,其中所述载体是病毒载体。

[0777] 126. 条目125的方法,其中所述载体是AAV载体。

[0778] 127. 一种治疗额颞痴呆或痴呆的方法,其包括施用编码非天然存在的DNA结合蛋白的载体,所述非天然存在的DNA结合蛋白与天然存在的人类蛋白质的相应结构域(例如,在两种蛋白质之间的DNA结合域或转录效应物结构域之内)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%的局部序列同一性,或者与天然存在的人类蛋白质(例如,EGR1或EGR3)具有全局序列同一性,其中所述DNA结合蛋白能够激活或增加内源GRN或其同种型的表达。

[0779] 128. 条目127的方法,其中所述非天然存在的DNA结合蛋白选择性地结合具有SEQ ID NO: 38、113或330-336的区域,或其5bp、10bp、50bp、100bp、200bp或500bp内的区域。

[0780] 129. 条目127-128中任一项的方法,其中所述非天然存在的DNA结合蛋白与对照相比能够使GRN的表达增加至少1.5、2、3、4、5、6、7、8、9、10、15、20、25、50或100倍。

[0781] 130. 条目127-129中任一项的方法,其中所述非天然存在的DNA结合蛋白包含具有以下序列的DNA结合域: (i) SEQ ID NO: 26、32、86-89、93、165-171、165-171或377-390; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0782] 131. 条目127-130中任一项的方法,其中所述DNA结合蛋白包含以下序列: (i) SEQ ID NO: 10、16、59-60、63-64、72-73、76、77、112或337-350; (ii) 其变体或功能片段;或 (iii) 与 (i) 或 (ii) 中的任一个具有至少90%序列同一性(或至少91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。

[0783] 132. 条目127-131中任一项的方法,其中所述载体是病毒载体。

[0784] 133. 条目133的方法,其中所述载体是AAV载体。

[0785] 本文提供了包含非天然存在的转录调节物的表达盒,如增加SCN1A基因表达的表达盒,以及使用其治疗与SCN1A基因相关的病况的方法。本文还提供了调节神经蛋白质如SCN1A之表达的非天然存在的DNA结合蛋白。

[0786] 本公开的各个实施方案参考以下编号的条目来限定:

[0787] 1. 一种表达盒,其包含增加SCN1A基因表达的非天然存在的转录调节物。

[0788] 2. 条目1的表达盒,其中所述非天然存在的转录调节物结合选自下组的基因组区域: SEQ ID NO: 35-37、101、105-111、136、195-211、224-238或240-267,或其至少200bp内的区域。

[0789] 3. 条目1-2的表达盒,其中所述非天然存在的转录调节物结合选自下组的基因组区域: SEQ ID NO: 35-37、101、105-111、136、195-211、224-238或240-267,或其至少100bp内的区域。

- [0790] 4. 条目1-2的表达盒,其中所述非天然存在的转录调节物结合选自下组的基因组区域:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267,或其至少50bp内的区域。
- [0791] 5. 条目1-2的表达盒,其中所述非天然存在的转录调节物结合选自下组的基因组区域:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267,或其至少10bp内的区域。
- [0792] 6. 条目1-2的表达盒,其中所述非天然存在的转录调节物结合与hg19的区域相对应的基因组区域,该区域与选自下组的序列相匹配:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267。
- [0793] 7. 上述条目中任一项的表达盒,其中所述表达盒是病毒载体的一部分。
- [0794] 8. 条目7的表达盒,其中所述病毒载体是AAV病毒。
- [0795] 9. 条目8的表达盒,其中所述AAV病毒是AAV9病毒或scAAV9病毒。
- [0796] 10. 上述条目中任一项的表达盒,其中所述转录调节物包含DNA结合域和转录激活域。
- [0797] 11. 上述条目中任一项的表达盒,其中所述转录激活域包含选自下组的至少一种蛋白质的一部分:VPR、VP64、VP16、VP128和p300。
- [0798] 12. 上述条目中任一项的表达盒,其中所述转录激活域包含SEQ ID NO:95-100或114,或与之具有至少80%序列相似性的结构域。
- [0799] 13. 上述条目中任一项的表达盒,其中所述DNA结合域是锌指结构域。
- [0800] 14. 上述条目中任一项的表达盒,其中所述DNA结合域是Cas蛋白。
- [0801] 15. 上述条目中任一项的表达盒,其中所述表达盒进一步包含gRNA。
- [0802] 16. 上述条目中任一项的表达盒,其中所述gRNA包含选自下组的序列:SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336,或与之具有至少80%序列相似性的序列。
- [0803] 17. 条目16的表达盒,其中所述gRNA包含选自下组的序列:SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336,或与之具有至少90%序列相似性、与之具有至少95%序列相似性或与之具有至少99%序列相似性的序列。
- [0804] 18. 上述条目中任一项的表达盒,其中所述Cas蛋白是核酸酶灭活的Cas蛋白。
- [0805] 19. 条目18的表达盒,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。
- [0806] 20. 条目19的表达盒,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas3。
- [0807] 21. 上述条目中任一项的表达盒,其中所述DNA结合域是TAL效应物DNA结合域。
- [0808] 22. 上述条目中任一项的表达盒,其中所述表达盒进一步包含调节元件,该调节元件在GABA能神经元中以比在其他细胞类型中更高的水平驱动所述非天然存在的转录调节物的表达。
- [0809] 23. 上述条目中任一项的表达盒,其中所述表达盒进一步包含调节元件,该调节元件小于100bp,并驱动所述非天然存在的转录调节物的高表达。
- [0810] 24. 上述条目中任一项的表达盒,其中所述表达盒进一步包含至少一个选自下组的调节元件:SEQ ID NO:178-179、182-185或417,或与之具有至少80%、至少90%、至少95%或至少99%的序列同一性的序列。

[0811] 25. 上述条目中任一项的表达盒,其中所述非天然存在的转录调节物选自SEQ ID NO:6-9、13-15、44-45、48-49、54-55、58-62、67-77、103、112、114、268-282、305-325、337-350、364、295-299、365-366,或与之具有至少80%、至少90%、至少95%或至少99%的序列同一性。

[0812] 26. 上述条目中任一项的表达盒,其中所述非天然存在的转录调节物选自SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416。

[0813] 27. 一种非天然存在的DNA结合蛋白,其结合至少一个选自下组的基因组位置:SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336,或其200bp内、100bp内或50bp内的基因组位置。

[0814] 28. 上述任何条目的非天然存在的DNA结合蛋白,其中所述非天然存在的DNA结合蛋白结合至少一个选自下组的基因组位置:SEQ ID NO:35-38、105-111、113、136、195-211、224-238、240-267或330-336。

[0815] 29. 上述任何条目的非天然存在的DNA结合蛋白,其进一步包含转录调节域。

[0816] 30. 条目29的非天然存在的DNA结合蛋白,其中所述转录调节域是转录激活域。

[0817] 31. 条目29-30的非天然存在的DNA结合蛋白,其中所述转录激活域是转录因子的转录激活域。

[0818] 32. 上述条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含锌指转录因子的转录激活域。

[0819] 33. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含选自下组的至少一种蛋白质的一部分:VPR、VP64、VP16和VP128,或与之同源的蛋白质。

[0820] 34. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含SEQ ID NO:95-100或114中的任一个,或与之具有至少80%序列相似性的结构域。

[0821] 35. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含SEQ ID NO:95-100或114中的任一个,或与之具有至少90%、至少95%或至少99%序列相似性的结构域。

[0822] 36. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录调节域包含转录共激活蛋白,或转录共激活域的结构域。

[0823] 37. 条目36的非天然存在的DNA结合蛋白,其中所述转录共激活域包含p300的全部或部分。

[0824] 38. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含锌指结构域。

[0825] 39. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含四个锌指。

[0826] 40. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含五个锌指。

[0827] 41. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含六个锌指。

[0828] 42. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含七个锌指。

- [0829] 43. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含八个锌指。
- [0830] 44. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含九个锌指。
- [0831] 45. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含Cas蛋白。
- [0832] 46. 条目45的非天然存在的DNA结合蛋白,其中所述Cas蛋白是核酸酶灭活的Cas蛋白。
- [0833] 47. 条目46的非天然存在的DNA结合蛋白,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。
- [0834] 48. 条目46的非天然存在的DNA结合蛋白,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas3。
- [0835] 49. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含TAL效应物DNA结合域。
- [0836] 50. 上述任何条目的非天然存在的DNA结合蛋白,其中所述非天然存在的DNA结合蛋白选自SEQ ID NO:22-26、29-32、84-93、112、131-135、165-171、371-372、376-409或412-416,或与之具有至少90%、至少95%或至少99%的序列同一性。
- [0837] 51. 上述任何条目的非天然存在的DNA结合蛋白,其中当转染到细胞中时,所述非天然存在的DNA结合蛋白调节SCN1A的表达。
- [0838] 52. 一种通过施用非天然存在的转录调节物来调节细胞中SCN1A表达的方法,所述非天然存在的转录调节物结合至少一个选自下组的基因组位置:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267,或其200bp内、100bp内或20bp内的基因组位置。
- [0839] 53. 条目52的方法,其中所述非天然存在的转录调节物结合至少一个选自下组的基因组位置:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267。
- [0840] 54. 条目52-53的方法,其中所述非天然存在的转录调节物结合选自下组的序列:SEQ ID NO:35-37、101、105-111、136、195-211、224-238或240-267。
- [0841] 55. 一种通过施用编码非天然存在的转录调节物的表达盒来调节细胞中SCN1A表达的方法。
- [0842] 56. 条目55的方法,其中所述表达盒是病毒载体的一部分。
- [0843] 57. 条目56的方法,其中所述病毒载体是AAV病毒。
- [0844] 58. 条目57的方法,其中所述AAV病毒是AAV9病毒或scAAV9病毒。
- [0845] 59. 上述任何条目的方法,其中所述细胞是小白蛋白(PV)细胞。
- [0846] 60. 上述任何条目中任一项的方法,其中所述细胞在生物体内。
- [0847] 61. 条目60的方法,其中所述生物体是哺乳动物。
- [0848] 62. 条目60的方法,其中所述生物体是人。
- [0849] 63. 条目60的方法,其中调节SCN1A的表达治疗疾病或病症。
- [0850] 64. 条目60的方法,其中所述病症是中枢神经系统病症。
- [0851] 65. 条目60的方法,其中所述病症是帕金森病。
- [0852] 66. 条目60的方法,其中所述病症是Dravet综合征。

- [0853] 67. 条目60的方法,其中所述病症是阿尔茨海默病。
- [0854] 68. 条目60的方法,其中所述中枢神经系统病症的症状是GABA能功能减退。
- [0855] 69. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括增加PV细胞功能。
- [0856] 70. 上述任何条目的方法,其中所述中枢神经系统病症的症状是神经元活动过度。
- [0857] 71. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括减轻神经元活动过度。
- [0858] 72. 上述任何条目的方法,其中所述中枢神经系统病症的症状是癫痫发作。
- [0859] 73. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括降低癫痫发作的频率。
- [0860] 74. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括降低癫痫发作的严重程度。
- [0861] 75. 上述任何条目的方法,其中所述转录调节物包含DNA结合域和转录激活域。
- [0862] 76. 上述任何条目的方法,其中所述转录激活域包含转录因子的转录激活域。
- [0863] 77. 上述任何条目的方法,其中所述转录激活域包含锌指转录因子的转录激活域。
- [0864] 78. 上述任何条目的方法,其中所述转录激活域包含选自下组的至少一种蛋白质的一部分:VPR、VP64、VP16和VP128。
- [0865] 79. 上述任何条目的方法,其中所述转录激活域包含SEQ ID NO:95-100或114,或与之具有至少80%序列相似性的结构域。
- [0866] 80. 上述任何条目的方法,其中所述转录激活域包含转录共激活蛋白,或转录共激活域的结构域。
- [0867] 81. 上述任何条目的方法,其中所述转录激活域包含p300。
- [0868] 82. 上述任何条目的方法,其中所述DNA结合域是锌指结构域。
- [0869] 83. 上述任何条目的方法,其中所述DNA结合域是Cas蛋白。
- [0870] 84. 上述任何条目的方法,其中所述表达盒进一步包含gRNA。
- [0871] 85. 上述任何条目的方法,其中所述gRNA包含选自下组的序列:SEQ ID NO:105-111或113,或与之具有至少80%序列同一性的序列。
- [0872] 86. 上述任何条目的方法,其中所述Cas蛋白是核酸酶灭活的Cas蛋白。
- [0873] 87. 上述任何条目的方法,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。
- [0874] 88. 上述任何条目的方法,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas3。
- [0875] 89. 上述任何条目的方法,其中所述DNA结合域是TAL效应物DNA结合域。
- [0876] 90. 上述任何条目的方法,其中所述表达盒进一步包含对小白蛋白细胞具有特异性的调节元件。
- [0877] 91. 上述任何条目的方法,其中所述表达盒进一步包含调节元件,该调节元件小于100bp,并驱动高表达。
- [0878] 92. 上述任何条目的方法,其中所述表达盒进一步包含至少一个选自下组的调节元件:SEQ ID NO:178-179、182-185或417,或与之具有至少80%序列同一性的序列。
- [0879] 93. 上述任何条目的方法,其中调节SCN1A的表达包括增加SCN1A基因或Nav1.1蛋白的表达。
- [0880] 94. 上述任何条目的方法,其中SCN1A基因或Nav1.1蛋白的表达增加在小白蛋白细

胞中特异性发生。

[0881] 95. 上述任何条目的方法, 其中SCN1A基因或Nav1.1蛋白的表达增加在小白蛋白神经元中特异性发生。

[0882] 96. 上述条目中任一项的方法, 其中所述非天然存在的转录调节物选自SEQ ID NO:95-100或114, 或与之具有至少80%、至少90%、至少95%或至少99%的序列同一性。

[0883] 本文提供了表达盒, 其包含增加颗粒蛋白前体 (GRN) 或其功能片段或变体的表达的非天然存在的转录调节物, 以及使用此类组合物治疗与GRN基因相关的病况或疾病的方法。本文还提供了调节细胞中GRN之表达的非天然存在的DNA结合蛋白。

[0884] 本公开的各个实施方案参考以下编号的条目来限定:

[0885] 1. 一种表达盒, 其包含增加GRN基因表达的非天然存在的转录调节物。

[0886] 2. 条目1的表达盒, 其中所述非天然存在的转录调节物结合选自下组的基因组区域: SEQ ID NO:38、113或330-336, 或其至少200bp内、至少100bp内、至少50bp内或至少10bp内的区域。

[0887] 3. 上述任何条目的表达盒, 其中所述非天然存在的转录调节物结合与hg19的区域相对应的基因组区域, 该区域与选自下组的序列相匹配: SEQ ID NO:38、113或330-336。

[0888] 4. 上述任何条目的表达盒, 其中所述表达盒是病毒载体的一部分。

[0889] 5. 条目4的表达盒, 其中所述病毒载体是AAV病毒。

[0890] 6. 条目5的表达盒, 其中所述AAV病毒是AAV9病毒或scAAV9病毒。

[0891] 7. 上述任何条目的表达盒, 其中所述转录调节物包含DNA结合域和转录激活域。

[0892] 8. 上述任何条目的表达盒, 其中所述转录激活域包含选自下组的至少一种蛋白质的一部分: VPR、VP64、VP16、VP128、p65和p300。

[0893] 9. 上述任何条目的表达盒, 其中所述转录激活域包含SEQ ID NO:95-100或114, 或与之具有至少80%序列相似性的结构域。

[0894] 10. 上述任何条目的表达盒, 其中所述DNA结合域是锌指结构域。

[0895] 11. 上述任何条目的表达盒, 其中所述DNA结合域是Cas蛋白。

[0896] 12. 上述任何条目的表达盒, 其中所述表达盒进一步包含gRNA。

[0897] 13. 上述任何条目的表达盒, 其中所述gRNA包含选自下组的序列: SEQ ID NO:113, 或与之具有至少80%、至少90%、至少95%或至少99%的序列相似性的序列。

[0898] 14. 上述任何条目的表达盒, 其中所述Cas蛋白是核酸酶灭活的Cas蛋白 (dCas)。

[0899] 15. 上述任何条目的表达盒, 其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9 (dCas9) 或dSaCas9。

[0900] 16. 上述任何条目的表达盒, 其中所述DNA结合域是TAL效应物DNA结合域。

[0901] 17. 上述任何条目的表达盒, 其中所述表达盒进一步包含调节元件, 该调节元件在选自中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞和纹状体细胞的细胞中以比在其他细胞类型中更高的水平驱动所述非天然存在的转录调节物的表达。

[0902] 18. 上述任何条目的表达盒, 其中所述表达盒进一步包含调节元件, 该调节元件小于100bp或小于50bp, 并驱动所述非天然存在的转录调节物的高表达。

[0903] 19. 上述任何条目的表达盒, 其中所述表达盒进一步包含根据SEQ ID NO:178-179、182-185或417的序列, 或与之具有至少80%、至少90%、至少95%或至少99%的序列同

一性的序列。

[0904] 20. 上述任何条目的表达盒,其中所述非天然存在的转录调节物选自SEQ ID NO: 10、16、59-60、63-64、72-73、76-77、112或337-335,或与之具有至少80%、至少90%、至少95%或至少99%的序列同一性。

[0905] 21. 一种非天然存在的DNA结合蛋白,其结合至少一个选自下组的基因组位置:SEQ ID NO:38、113或330-336,或其200bp内、100bp内或50bp内或10bp内的基因组位置。

[0906] 22. 上述任何条目的非天然存在的DNA结合蛋白,其中所述非天然存在的DNA结合蛋白结合至少一个选自下组的基因组位置:SEQ ID NO:38、113或330-336。

[0907] 23. 上述任何条目的非天然存在的DNA结合蛋白,其进一步包含转录调节域。

[0908] 24. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录调节域是转录激活域。

[0909] 25. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域是转录因子的转录激活域。

[0910] 26. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含锌指转录因子的转录激活域。

[0911] 27. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含选自下组的至少一种蛋白质的一部分:VPR、VP64、VP16和VP128,或与之同源的蛋白质。

[0912] 28. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录激活域包含SEQ ID NO:95-100或114,或与之具有至少80%、至少90%、至少95%或至少99%序列相似性的结构域。

[0913] 29. 上述任何条目的非天然存在的DNA结合蛋白,其中所述转录调节域包含转录共激活蛋白,或转录共激活域的结构域。

[0914] 30. 条目29的非天然存在的DNA结合蛋白,其中所述转录共激活域包含p300的全部或部分。

[0915] 31. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含锌指结构域。

[0916] 32. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含四个锌指。

[0917] 33. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含五个锌指。

[0918] 34. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含六个锌指。

[0919] 35. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含七个锌指。

[0920] 36. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含八个锌指。

[0921] 37. 上述任何条目的非天然存在的DNA结合蛋白,其中所述锌指结构域包含九个锌指。

[0922] 38. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含Cas蛋

白。

[0923] 39. 上述任何条目的非天然存在的DNA结合蛋白,其中所述Cas蛋白是核酸酶灭活的Cas蛋白。

[0924] 40. 上述任何条目的非天然存在的DNA结合蛋白,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。

[0925] 41. 上述任何条目的非天然存在的DNA结合蛋白,其中所述DNA结合蛋白包含TAL效应物DNA结合域。

[0926] 42. 上述任何条目的非天然存在的DNA结合蛋白,其中所述非天然存在的DNA结合蛋白选自SEQ ID NO:26、32、86-89、93、165-171、112或377-390,或与之具有至少90%、至少95%或至少99%的序列同一性。

[0927] 43. 上述任何条目的非天然存在的DNA结合蛋白,其中当转染到细胞中时,所述非天然存在的DNA结合蛋白调节GRN的表达。

[0928] 44. 一种通过施用非天然存在的转录调节物来调节细胞中GRN表达的方法,所述非天然存在的转录调节物结合至少一个选自下组的基因组位置:SEQ ID NO:38、113或330-336,或其200bp内、100bp内或20bp内的基因组位置。

[0929] 45. 条目44的方法,其中所述非天然存在的转录调节物结合选自下组的序列:SEQ ID NO:38、113或330-336。

[0930] 46. 一种通过施用编码非天然存在的转录调节物的表达盒来调节细胞中GRN表达的方法。

[0931] 47. 条目46的方法,其中所述表达盒是病毒载体的一部分。

[0932] 48. 条目47的方法,其中所述病毒载体是AAV病毒。

[0933] 49. 条目48的方法,其中所述AAV病毒是AAV9病毒或scAAV9病毒。

[0934] 50. 上述任何条目的方法,其中所述细胞选自:中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞和纹状体细胞。

[0935] 51. 上述条目中任一项的方法,其中所述细胞在生物体内。

[0936] 52. 条目51的方法,其中所述生物体是哺乳动物。

[0937] 53. 条目52的方法,其中所述生物体是人。

[0938] 54. 上述任何条目的方法,其中调节GRN的表达治疗疾病或病症。

[0939] 55. 上述任何条目的方法,其中所述病症是中枢神经系统病症。

[0940] 56. 上述任何条目的方法,其中所述病症是额颞变性(FTD)。

[0941] 57. 上述任何条目的方法,其中所述病症是帕金森病。

[0942] 58. 上述任何条目的方法,其中所述病症是阿尔茨海默病。

[0943] 59. 上述任何条目的方法,其中所述病症是动脉粥样硬化。

[0944] 60. 上述任何条目的方法,其中所述中枢神经系统病症的症状是Lewy小体的存在。

[0945] 61. 上述任何条目的方法,其中所述中枢神经系统病症的症状是颗粒蛋白前体(GRN)的单倍体不足。

[0946] 62. 上述任何条目的方法,其中所述中枢神经系统病症的症状是社交缺陷。

[0947] 63. 上述任何条目的方法,其中所述中枢神经系统病症的症状是溶酶体异常。

[0948] 64. 上述任何条目的方法,其中所述中枢神经系统病症的症状是记忆丧失。

- [0949] 65. 上述任何条目的方法,其中所述中枢神经系统病症的症状是运动协调性丧失。
- [0950] 66. 上述任何条目的方法,其中所述中枢神经系统病症的症状是肌震颤。
- [0951] 67. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括降低肌震颤的频率。
- [0952] 68. 上述任何条目的方法,其中治疗所述中枢神经系统病症包括降低肌震颤的严重程度。
- [0953] 69. 上述任何条目的方法,其中所述转录调节物包含DNA结合域和转录激活域。
- [0954] 70. 上述任何条目的方法,其中所述转录激活域包含转录因子的转录激活域。
- [0955] 71. 上述任何条目的方法,其中所述转录激活域包含锌指转录因子的转录激活域。
- [0956] 72. 上述任何条目的方法,其中所述转录激活域包含选自下组的至少一种蛋白质的一部分:VPR、VP64、VP16和VP128。
- [0957] 73. 上述任何条目的方法,其中所述转录激活域包含SEQ ID NO:95-100或114,或与之具有至少80%序列相似性的结构域。
- [0958] 74. 上述任何条目的方法,其中所述转录激活域包含转录共激活蛋白,或转录共激活域的结构域。
- [0959] 75. 上述任何条目的方法,其中所述转录激活域包含p300。
- [0960] 76. 上述任何条目的方法,其中所述DNA结合域是锌指结构域。
- [0961] 77. 上述任何条目的方法,其中所述DNA结合域是Cas蛋白。
- [0962] 78. 上述任何条目的方法,其中所述表达盒进一步包含gRNA。
- [0963] 79. 上述任何条目的方法,其中所述gRNA包含选自下组的序列:SEQ ID NO:113,或与之具有至少80%序列同一性的序列。
- [0964] 80. 上述任何条目的方法,其中所述Cas蛋白是核酸酶灭活的Cas蛋白。
- [0965] 81. 上述任何条目的方法,其中所述核酸酶灭活的Cas蛋白是核酸酶灭活的Cas9。
- [0966] 82. 上述任何条目的方法,其中所述DNA结合域是TAL效应物DNA结合域。
- [0967] 83. 上述任何条目的方法,其中所述表达盒进一步包含非细胞特异性调节元件。
- [0968] 84. 上述任何条目的方法,其中所述表达盒进一步包含调节元件,该调节元件小于100bp,并驱动高表达。
- [0969] 85. 上述任何条目的方法,其中所述表达盒进一步包含调节元件,该调节元件包含根据SEQ ID NO:178-179、182-185或417的序列,或与之具有至少80%的序列同一性的序列。
- [0970] 86. 上述任何条目的方法,其中调节GRN的表达包括增加GRN基因或颗粒蛋白前体蛋白的表达。
- [0971] 87. 上述任何条目的方法,其中GRN基因或颗粒蛋白前体蛋白的表达增加在多种细胞类型中发生。
- [0972] 88. 上述任何条目的方法,其中GRN基因或颗粒蛋白前体蛋白的表达增加在选自中枢神经系统细胞、额皮质细胞、神经胶质细胞、小神经胶质细胞和纹状体细胞的细胞中相比于在其他细胞类型中特异性发生。
- [0973] 89. 上述任何条目的方法,其中所述非天然存在的转录调节物选自SEQ ID NO:10、16、59-60、72-72、76-77、112或337-350,或与之具有至少80%、至少90%、至少95%或至少

99%的序列同一性。

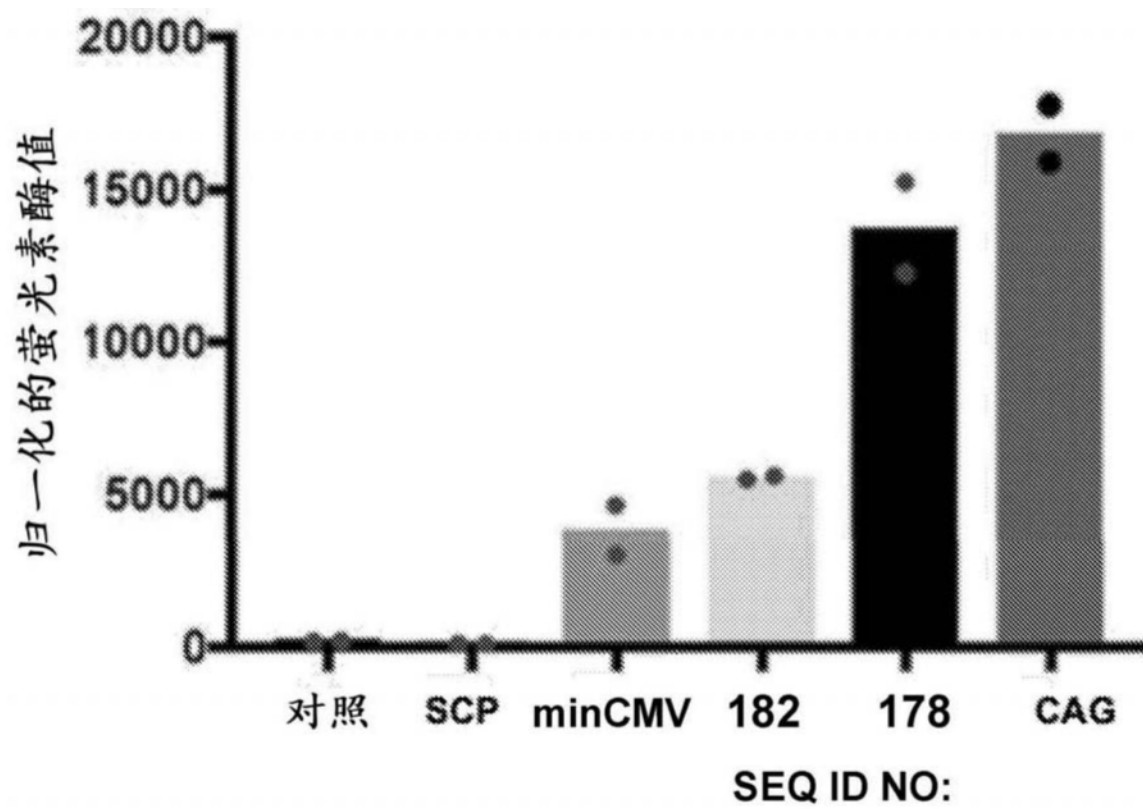


图1

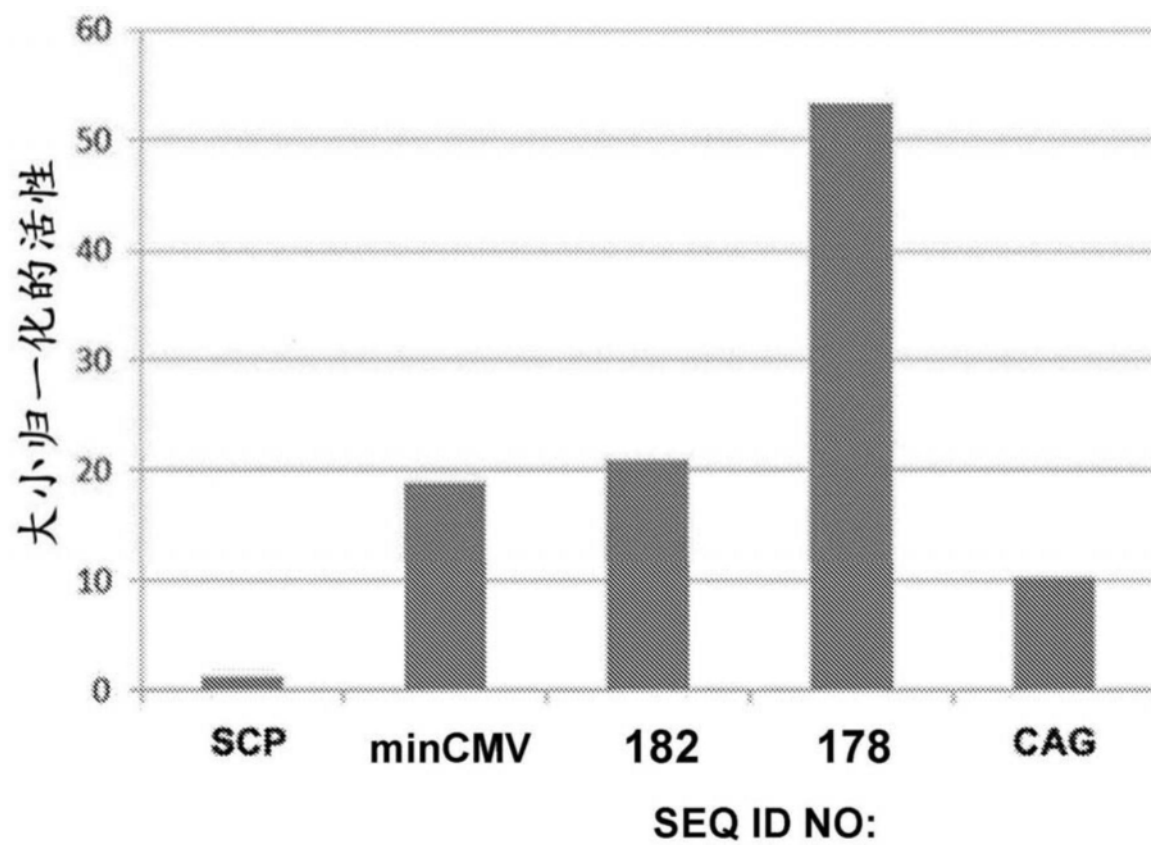


图2

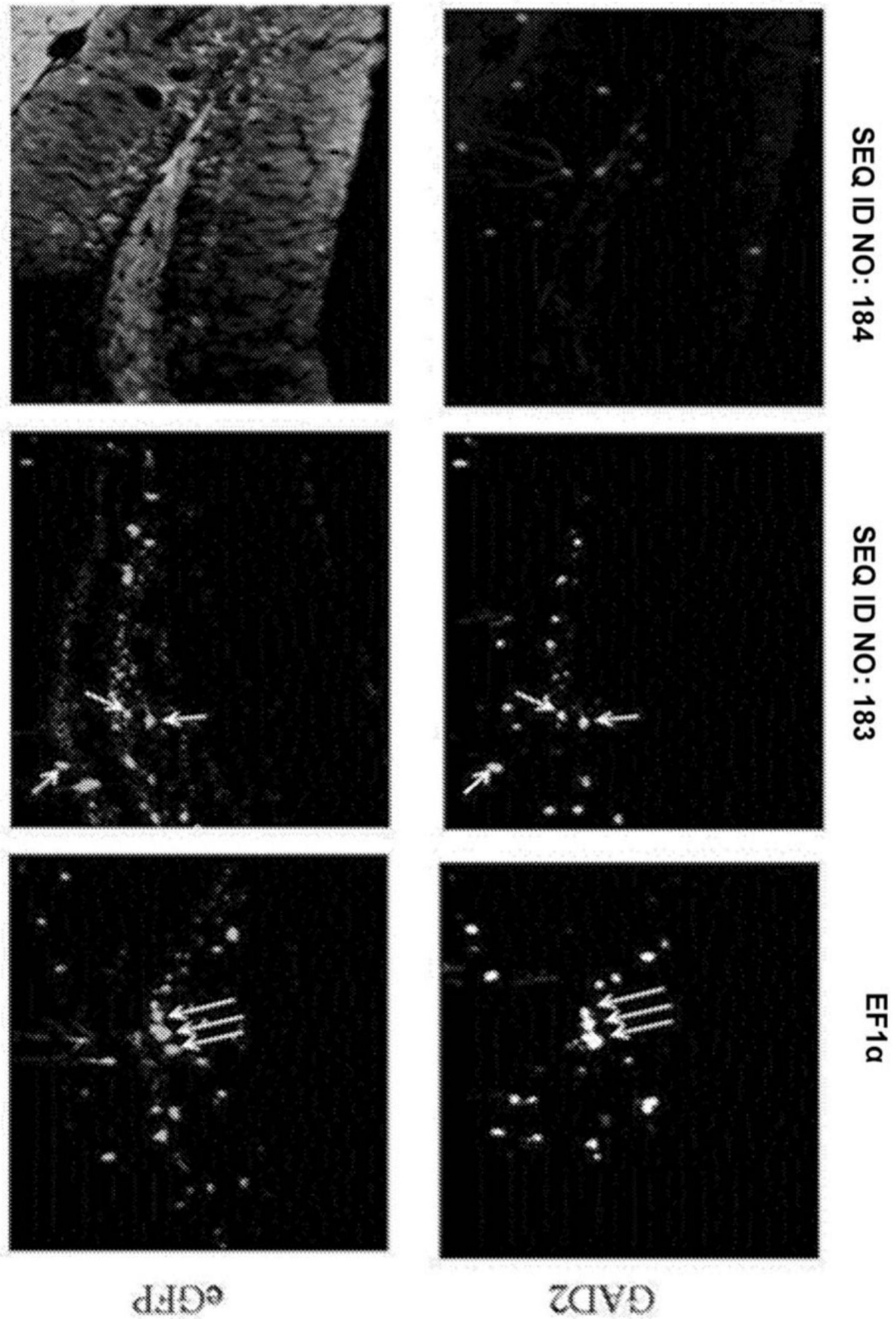


图3

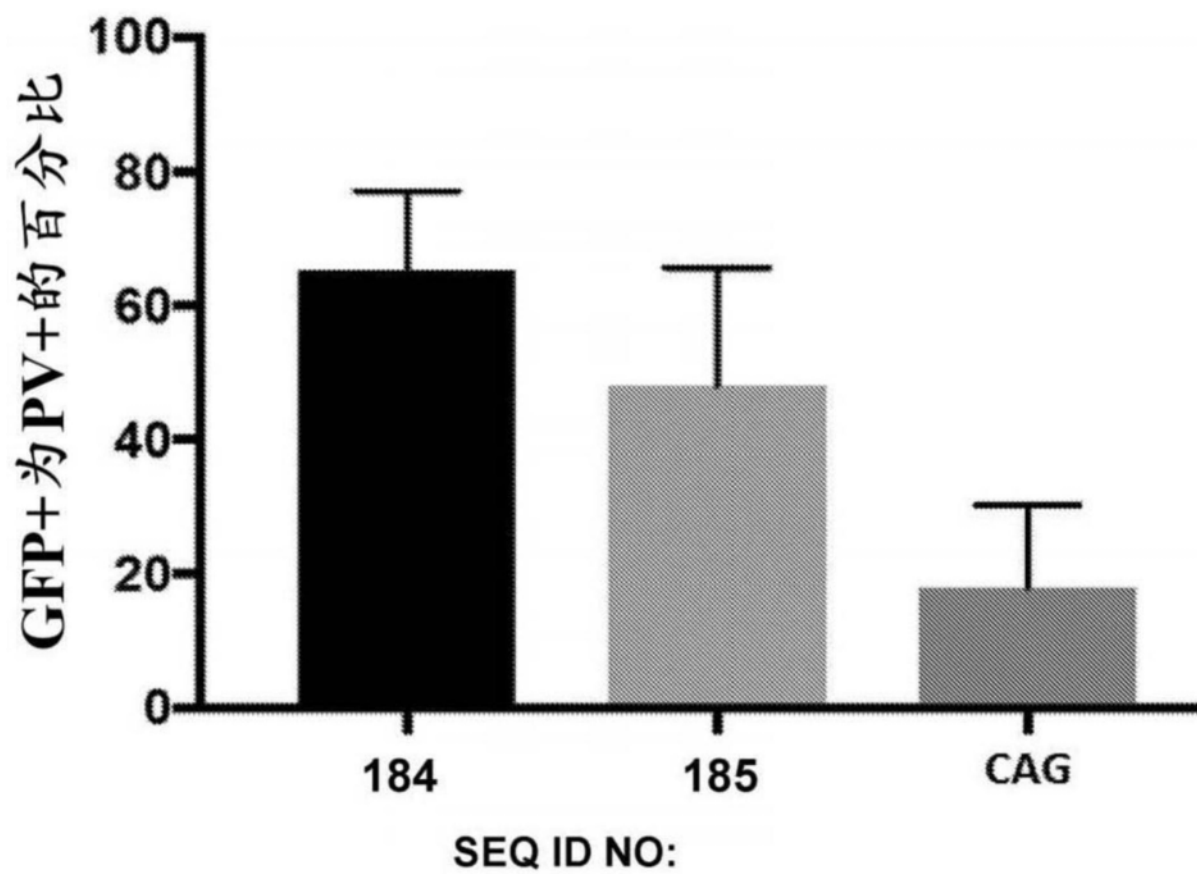


图4

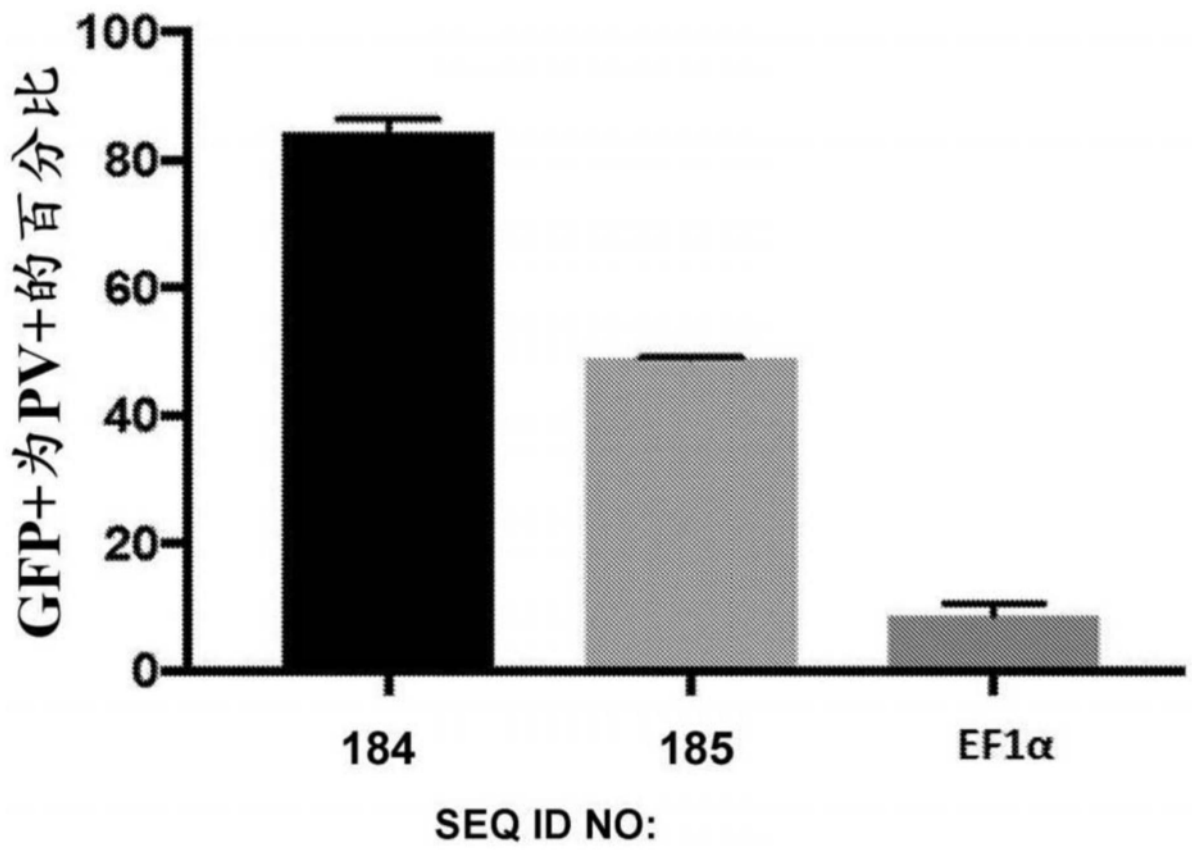


图5

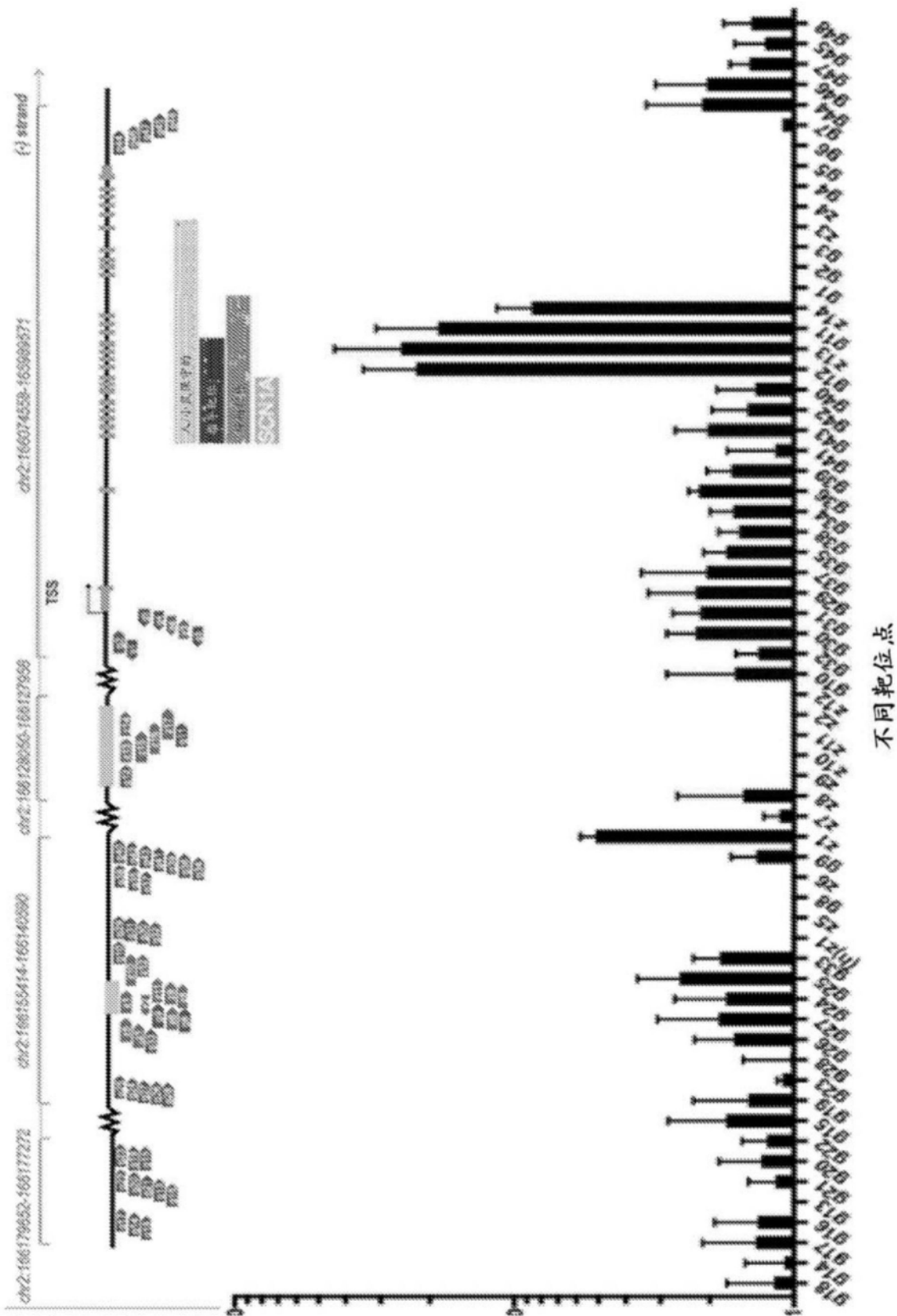


图6

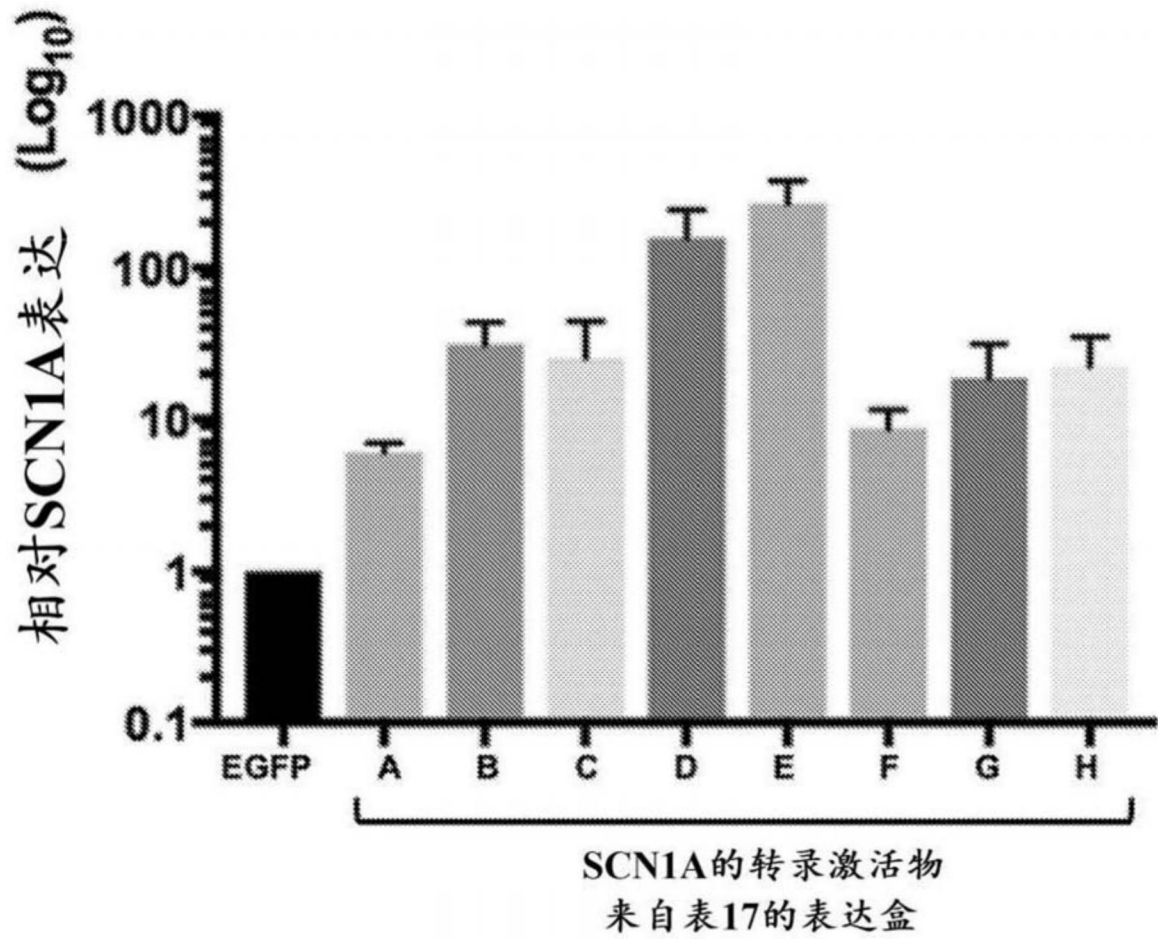


图7

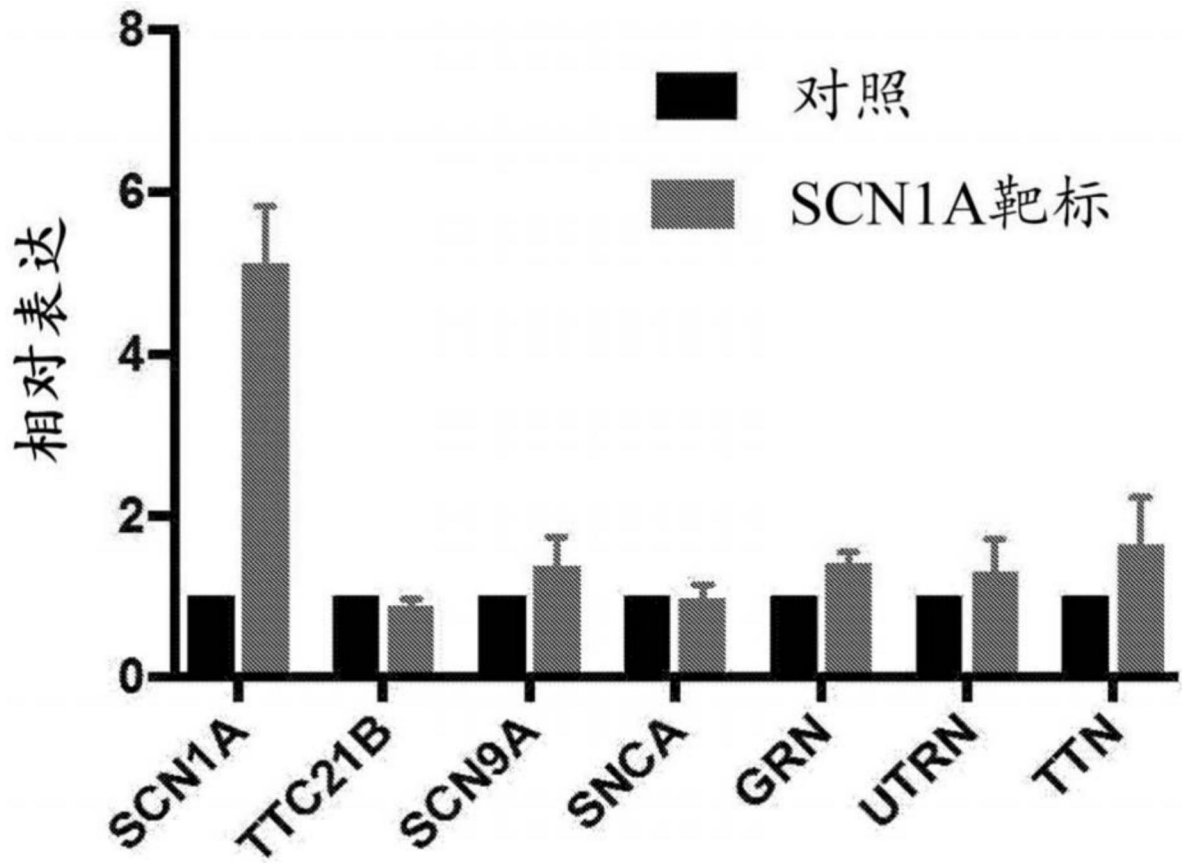
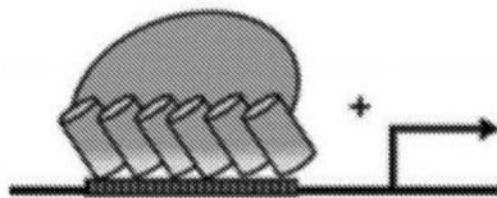


图8



未修饰的或野生型TF



具有重新编程的或工程改造的
DBD的eTF

图9A

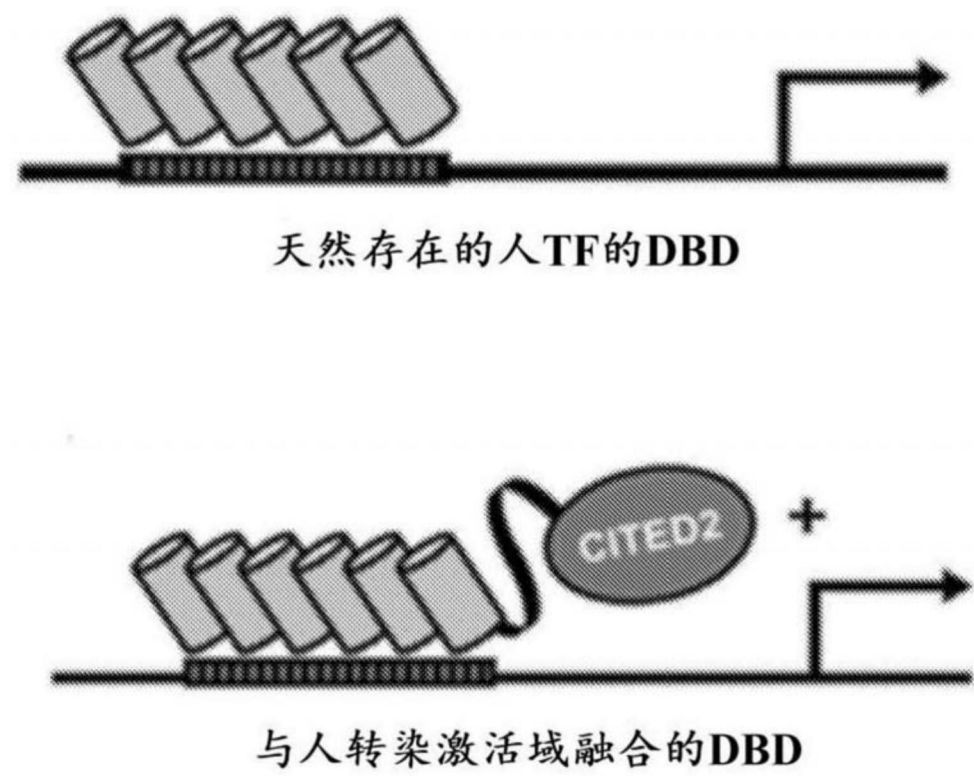


图9B

		保守百分比	大小 (bp)
eTF		>90%	1,000-1,800
dCas9		0%	3,900
ZFP		55%	800-900

图10

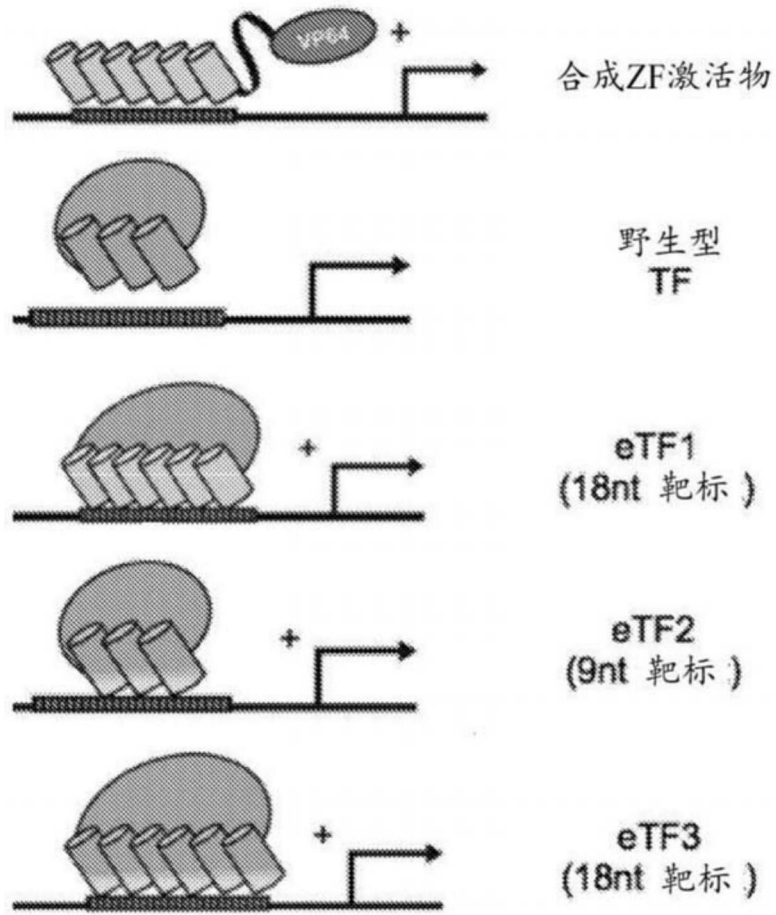


图11

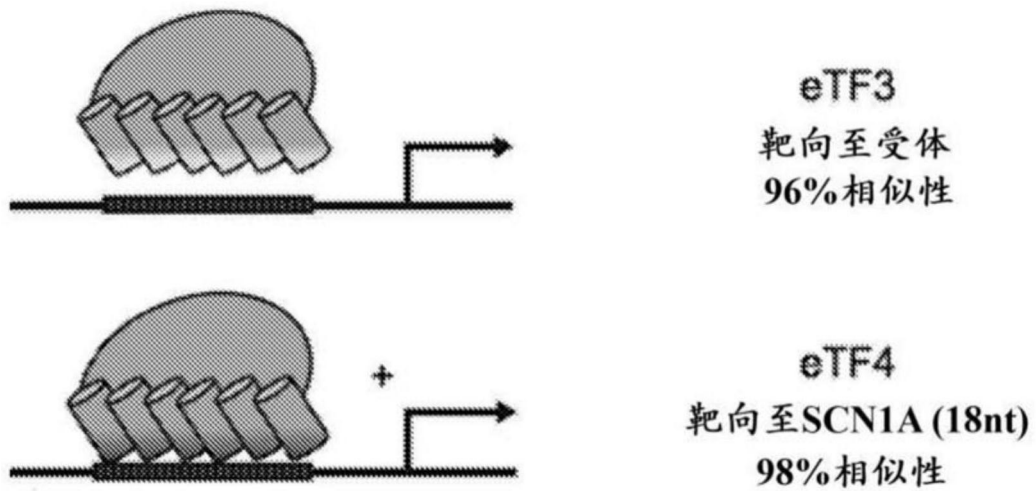


图12A

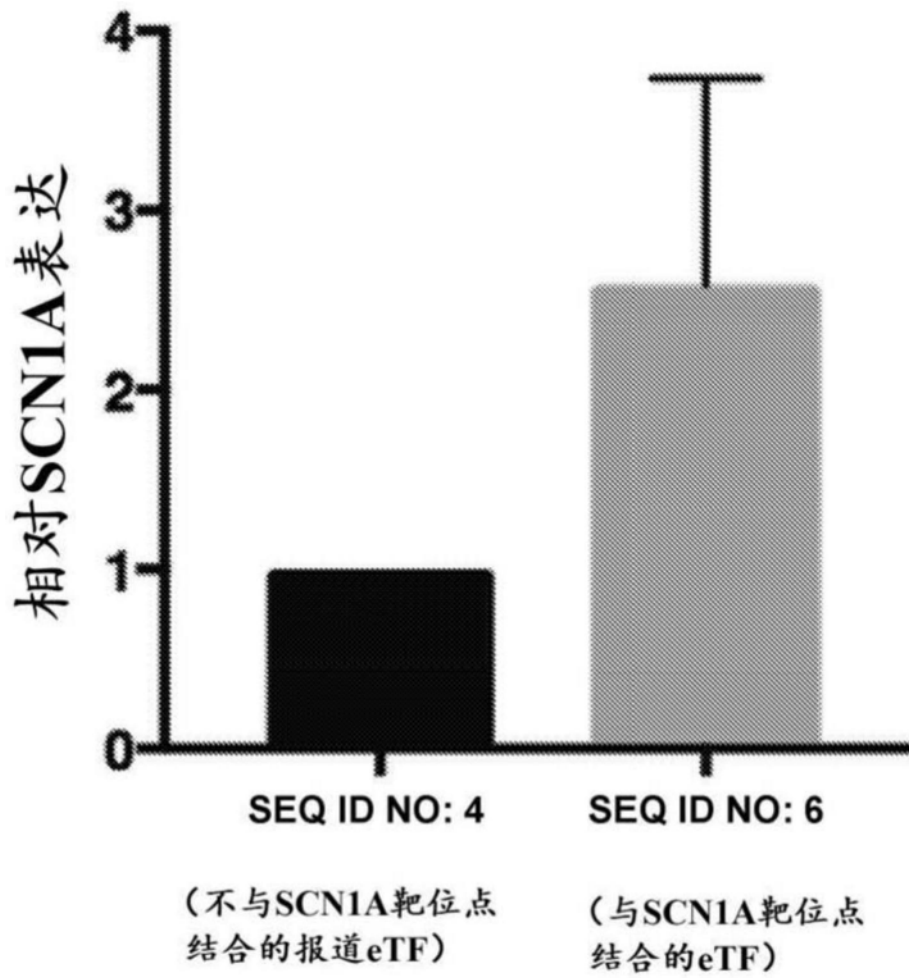


图12B

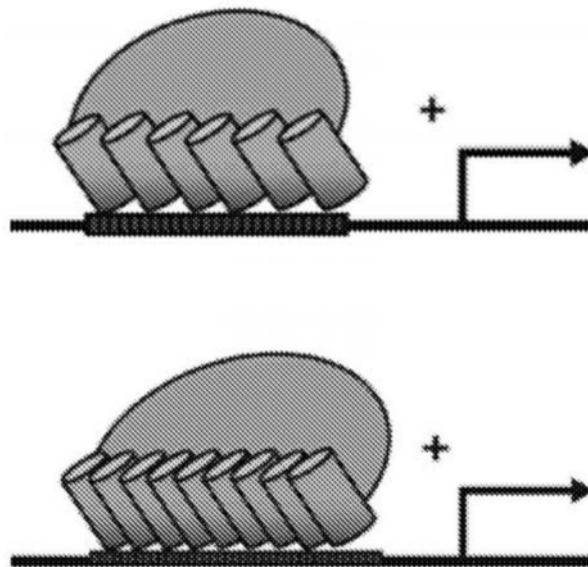


图13A

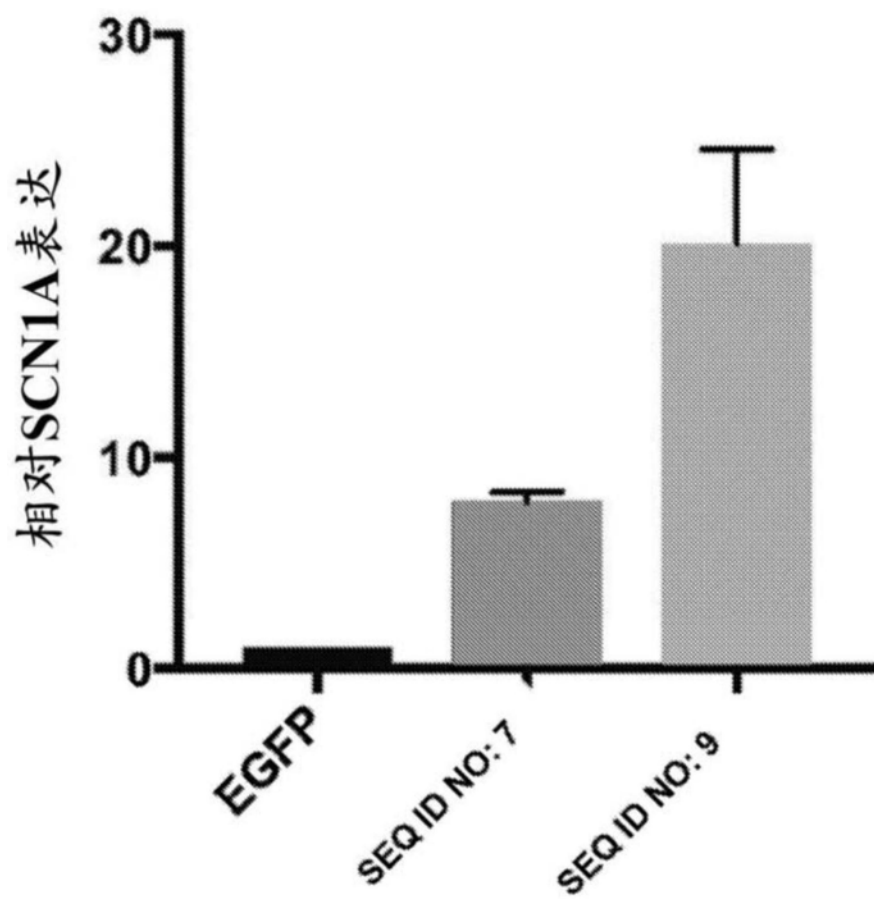


图13B




		DNA 结合域	激活域
人DNA结合域 无激活域		85%	N/A
人DNA结合域 病毒激活域		85%	0%
人DNA结合域 人激活域		85%	100%

图14

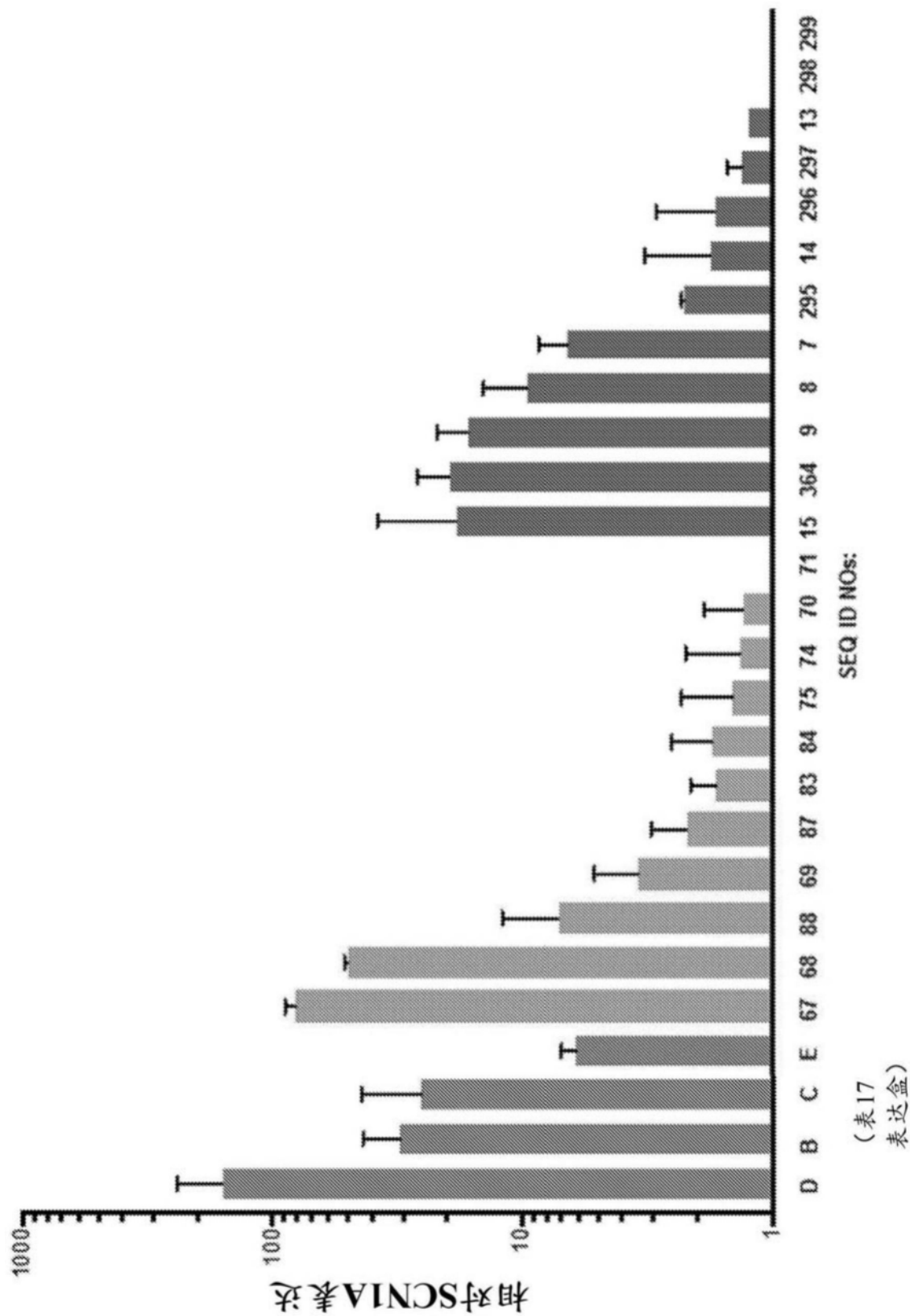


图15

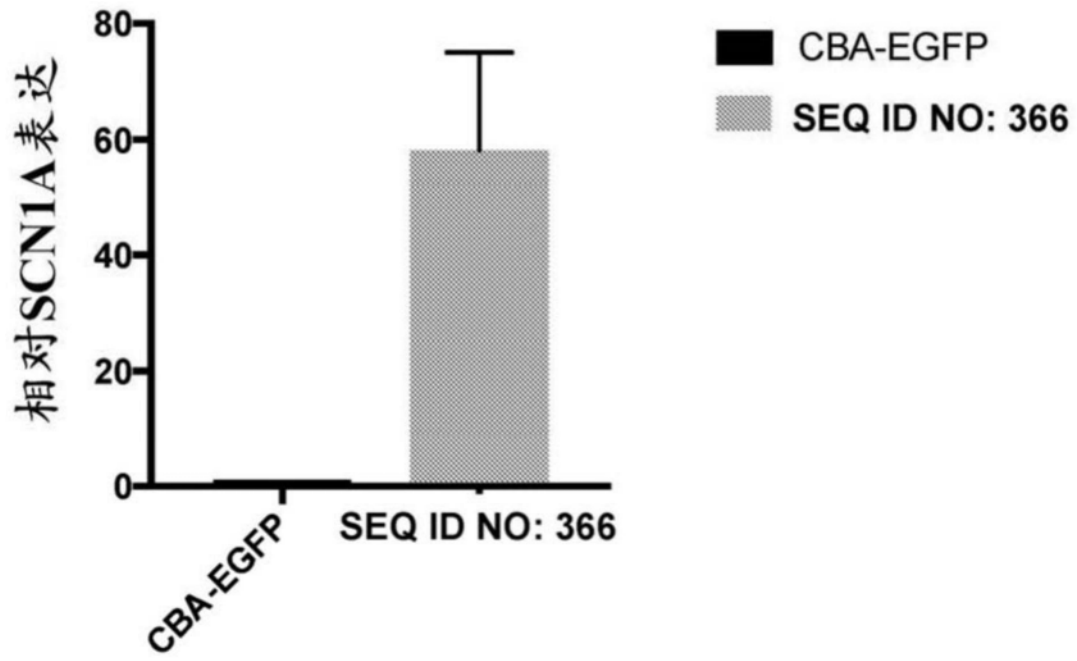


图16A

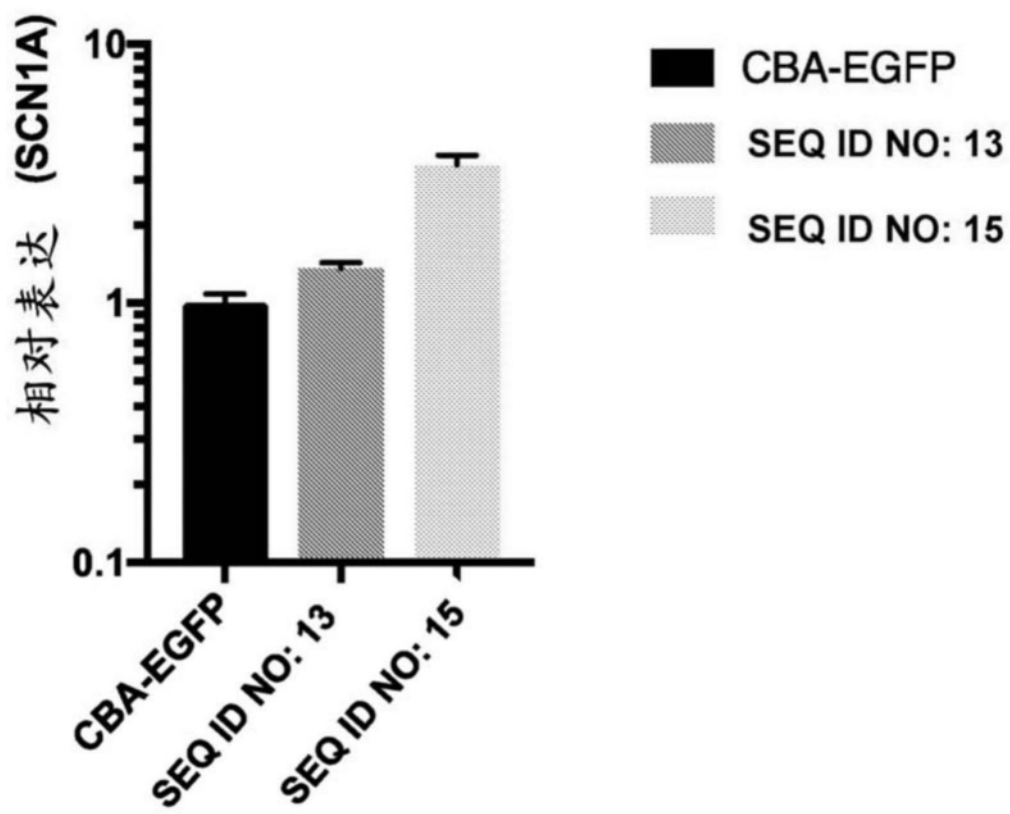


图16B

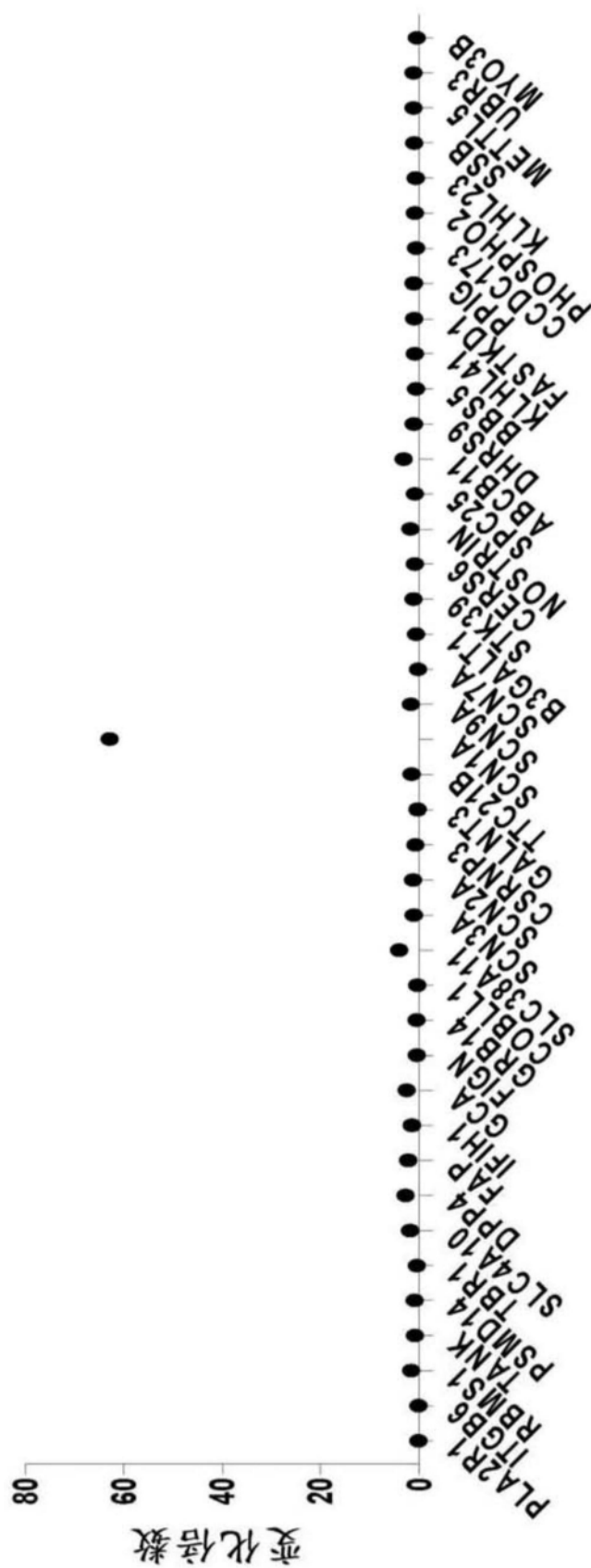


图17

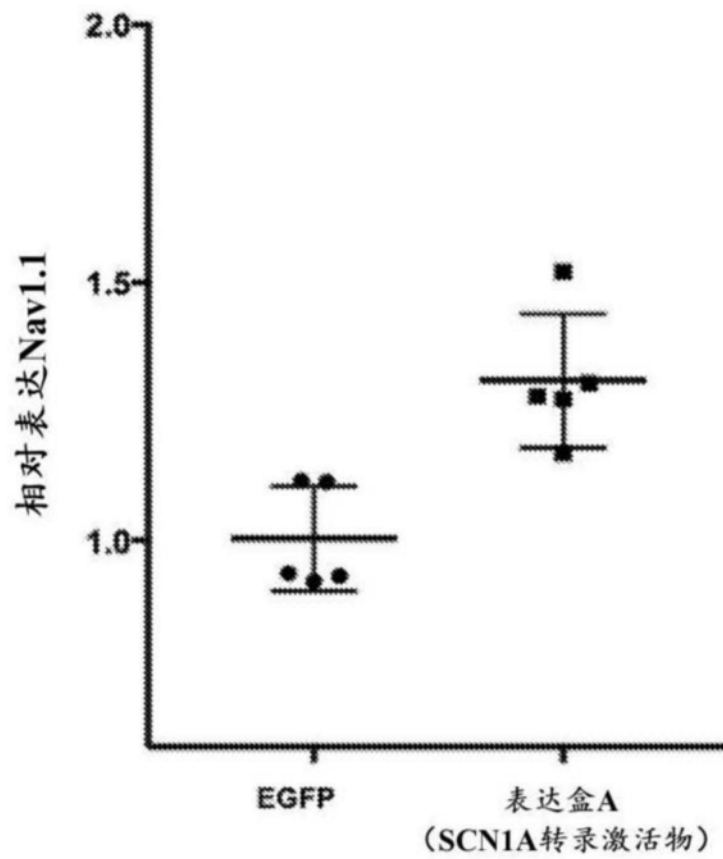


图18A

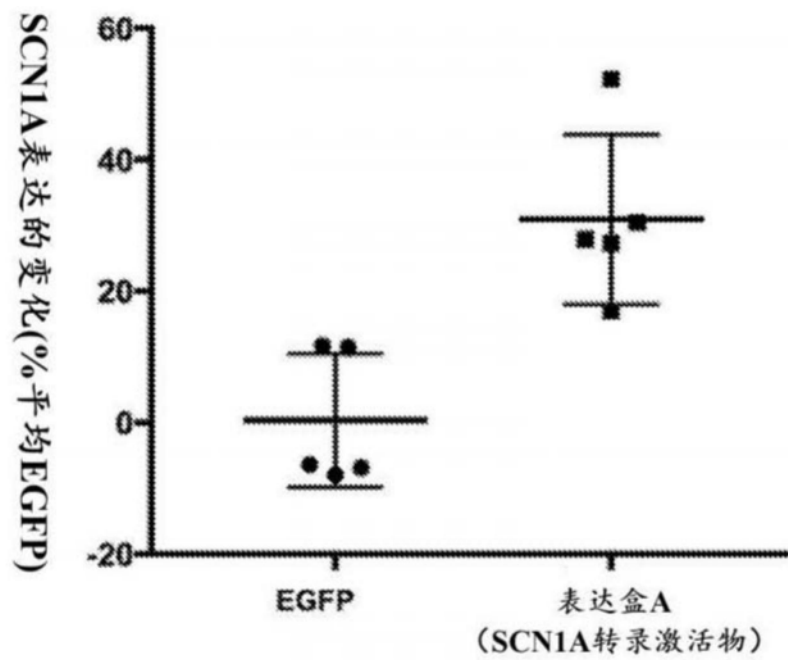


图18B

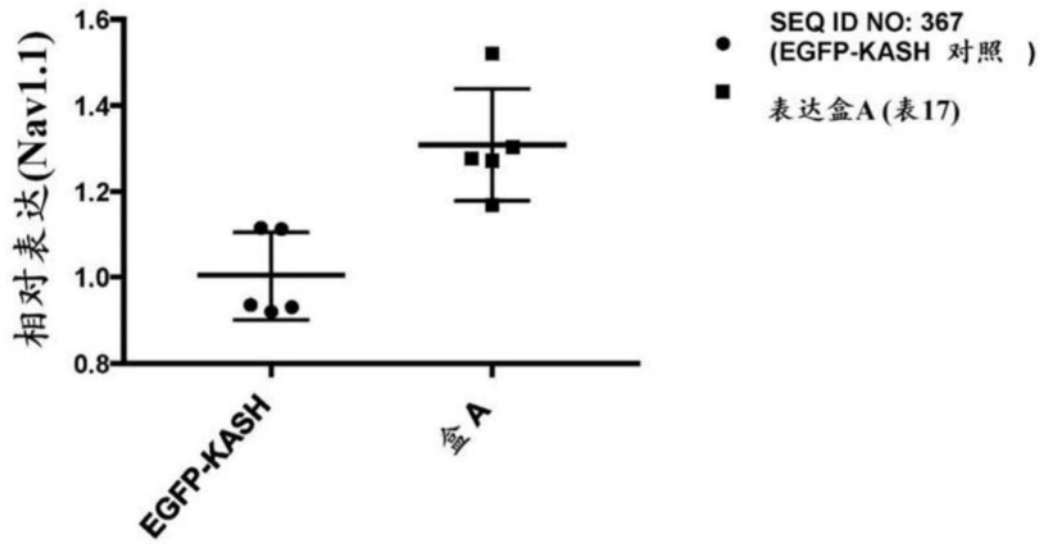


图19A

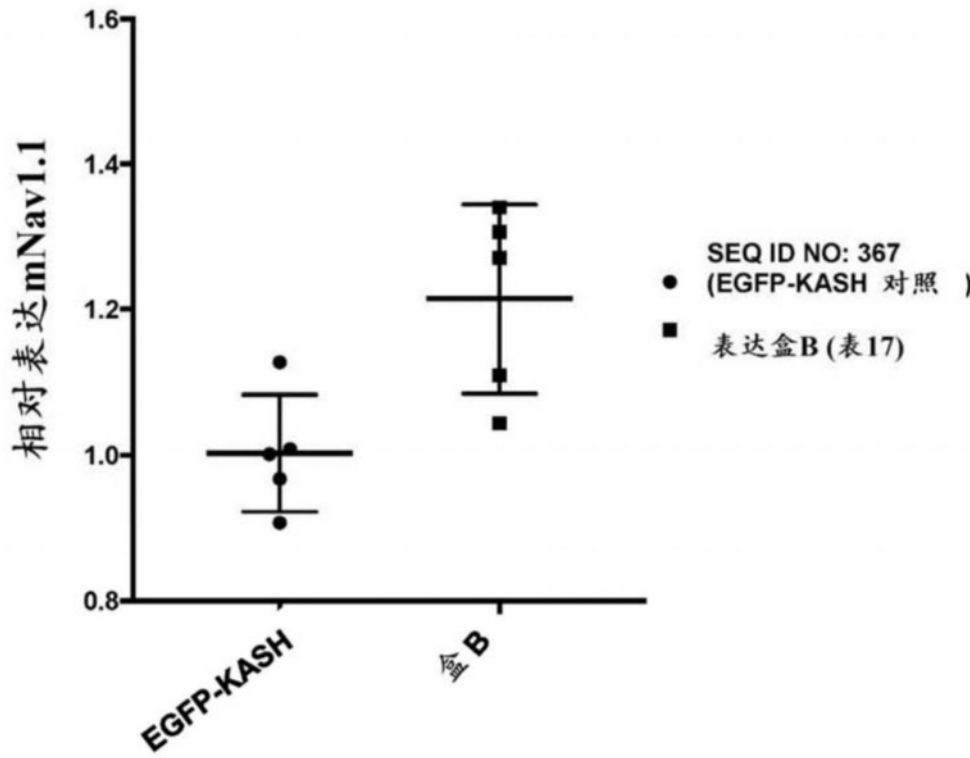


图19B

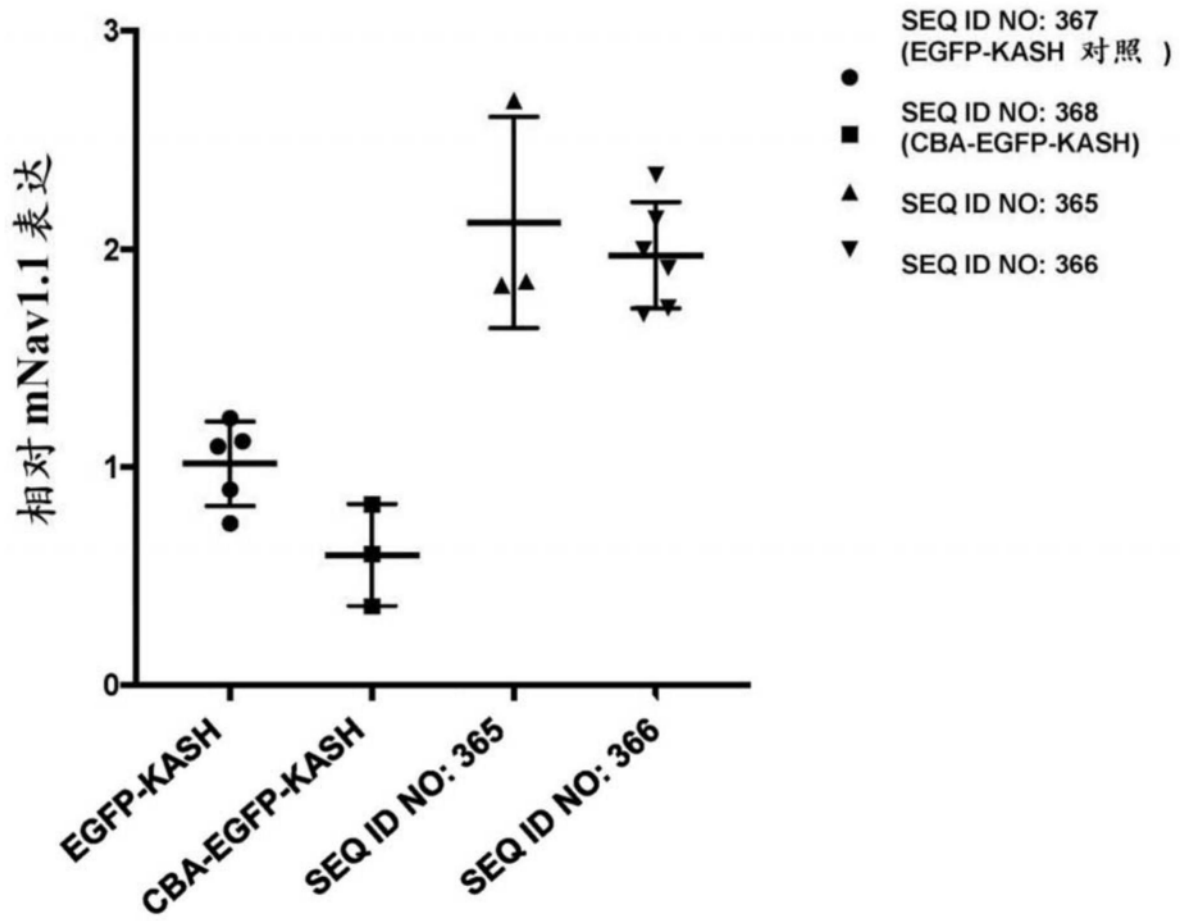


图19C

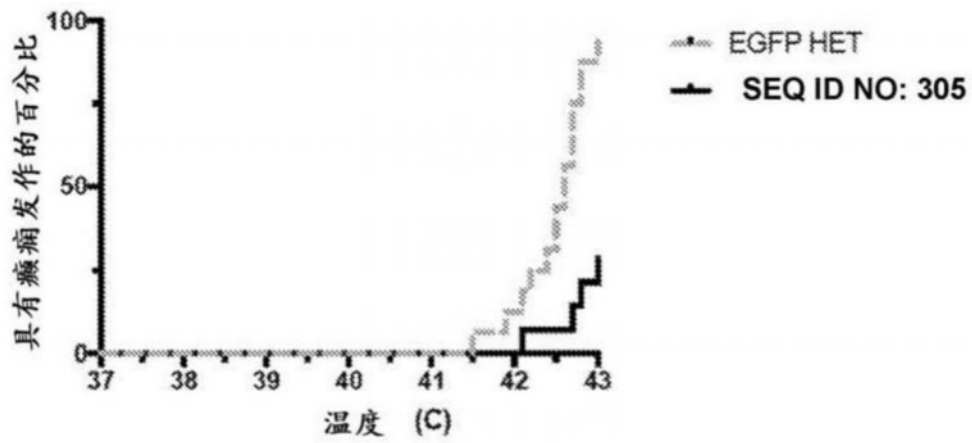


图20A

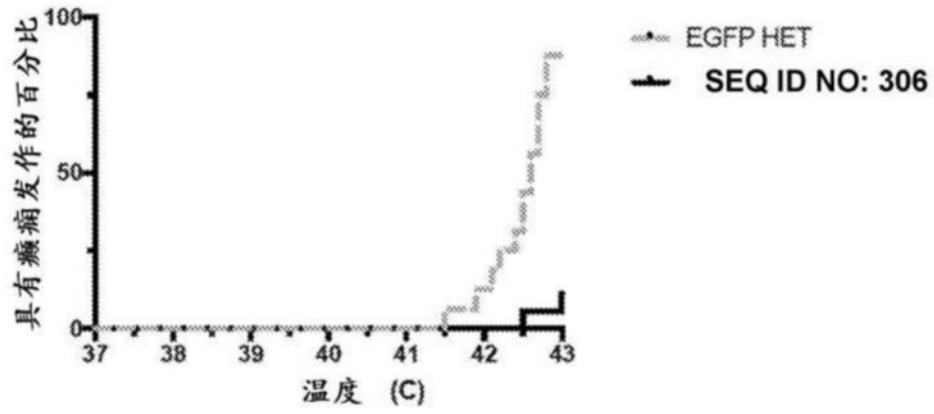


图20B

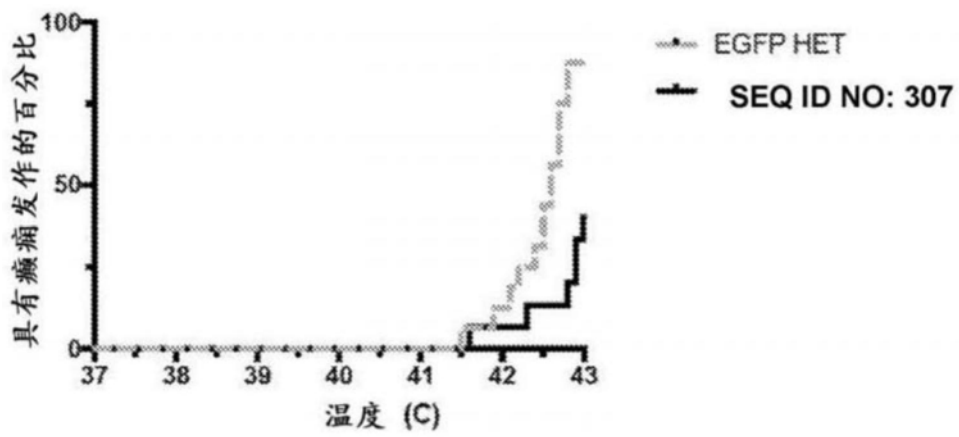


图20C

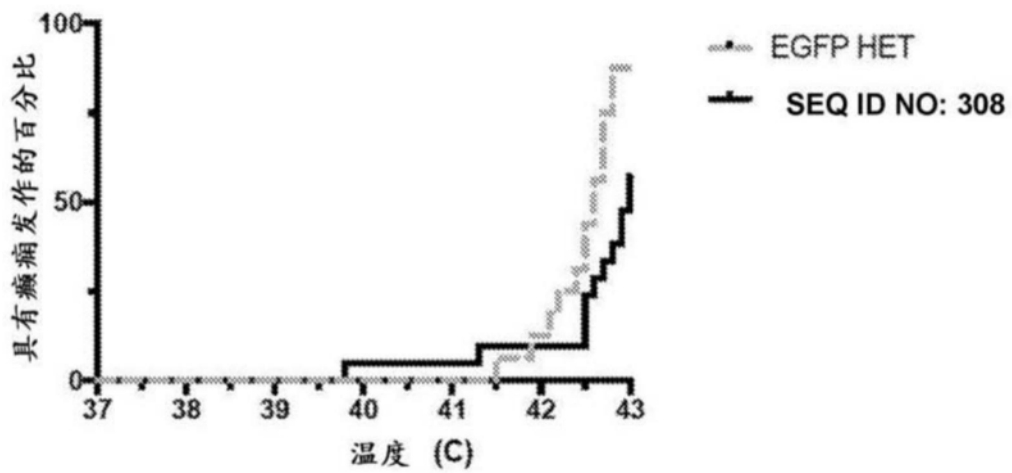


图20D

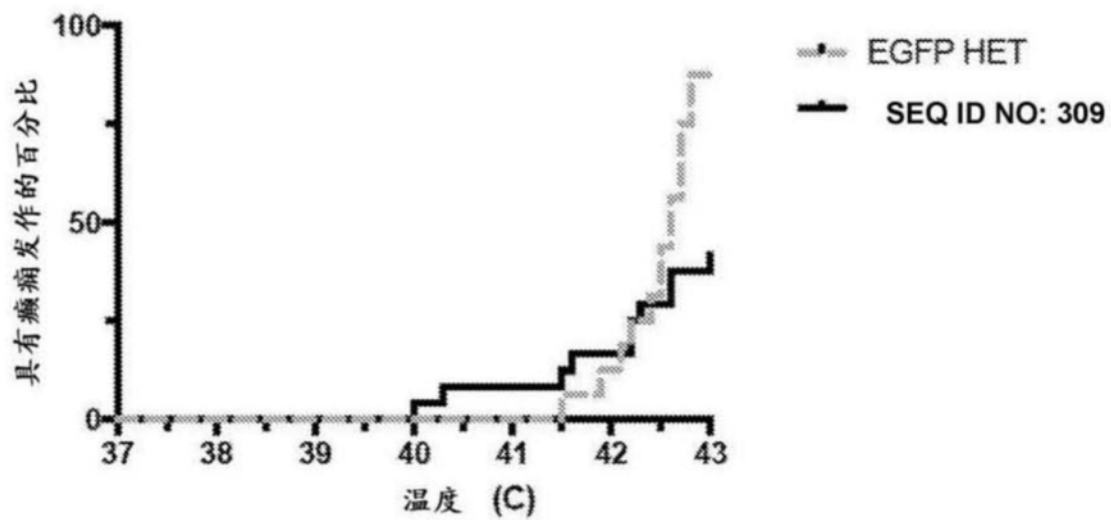


图20E

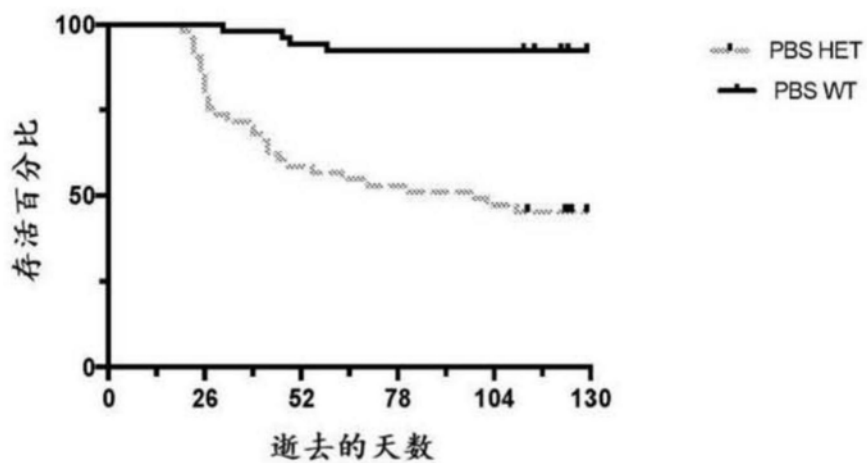


图21A

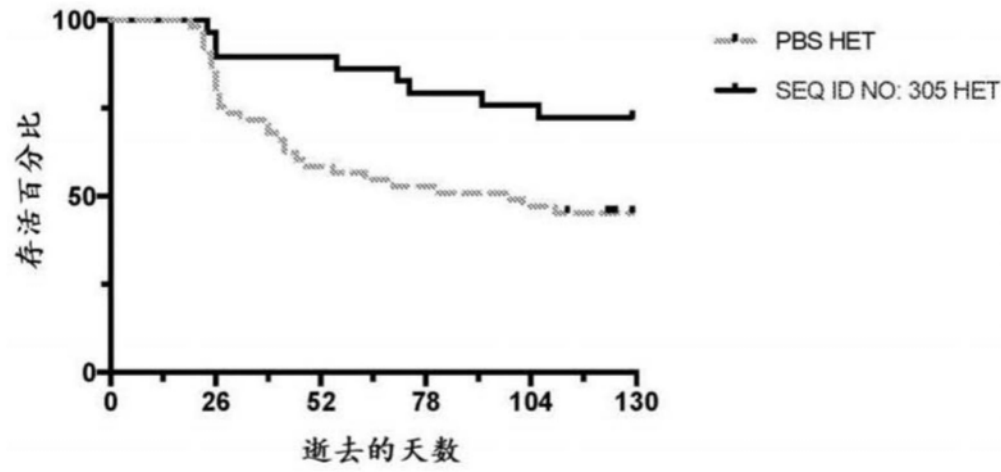


图21B

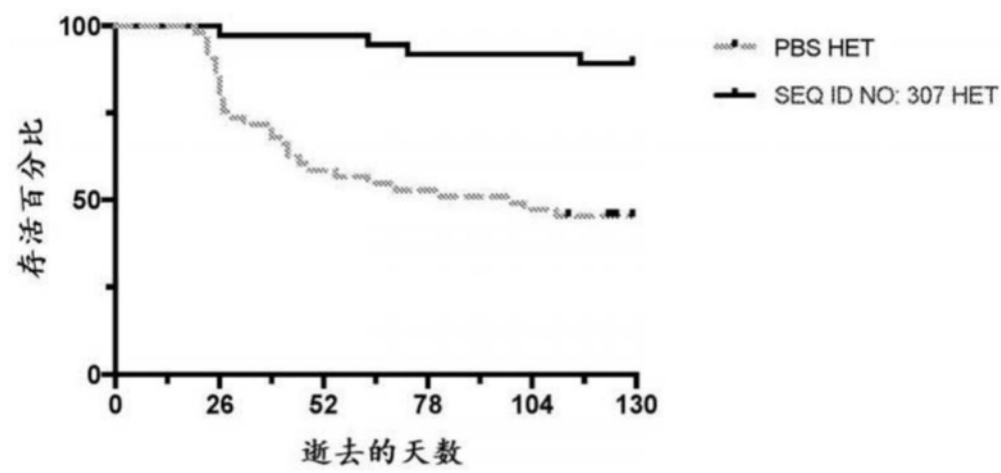


图21C

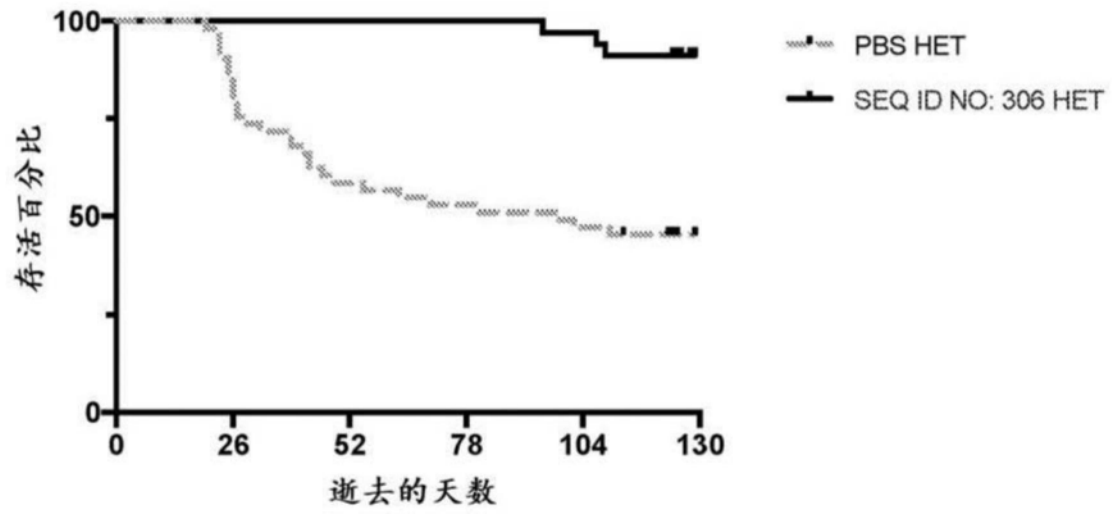


图21D

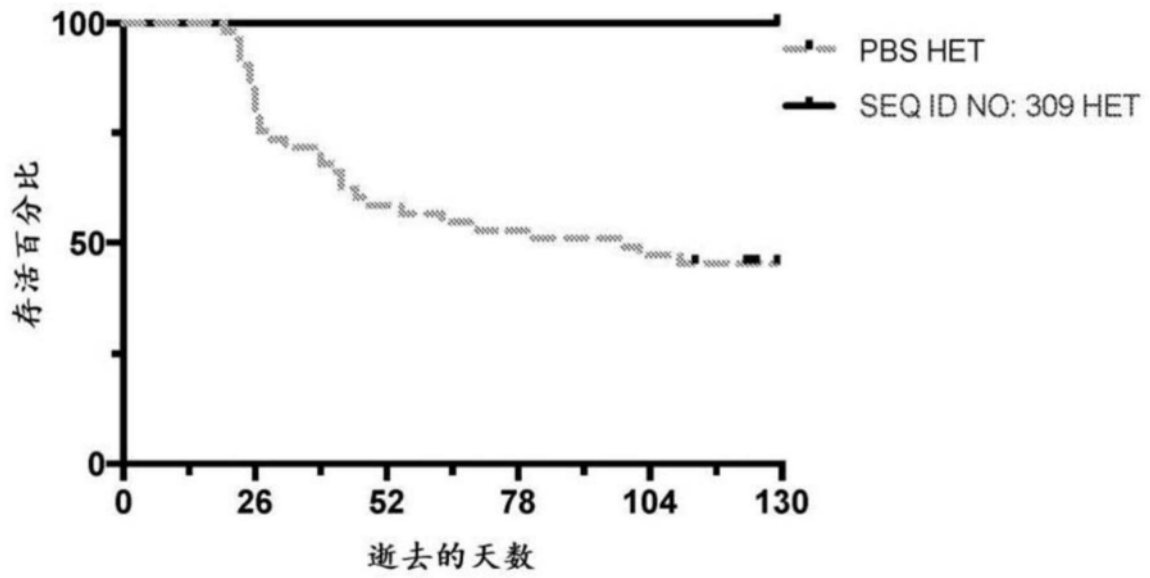


图21E

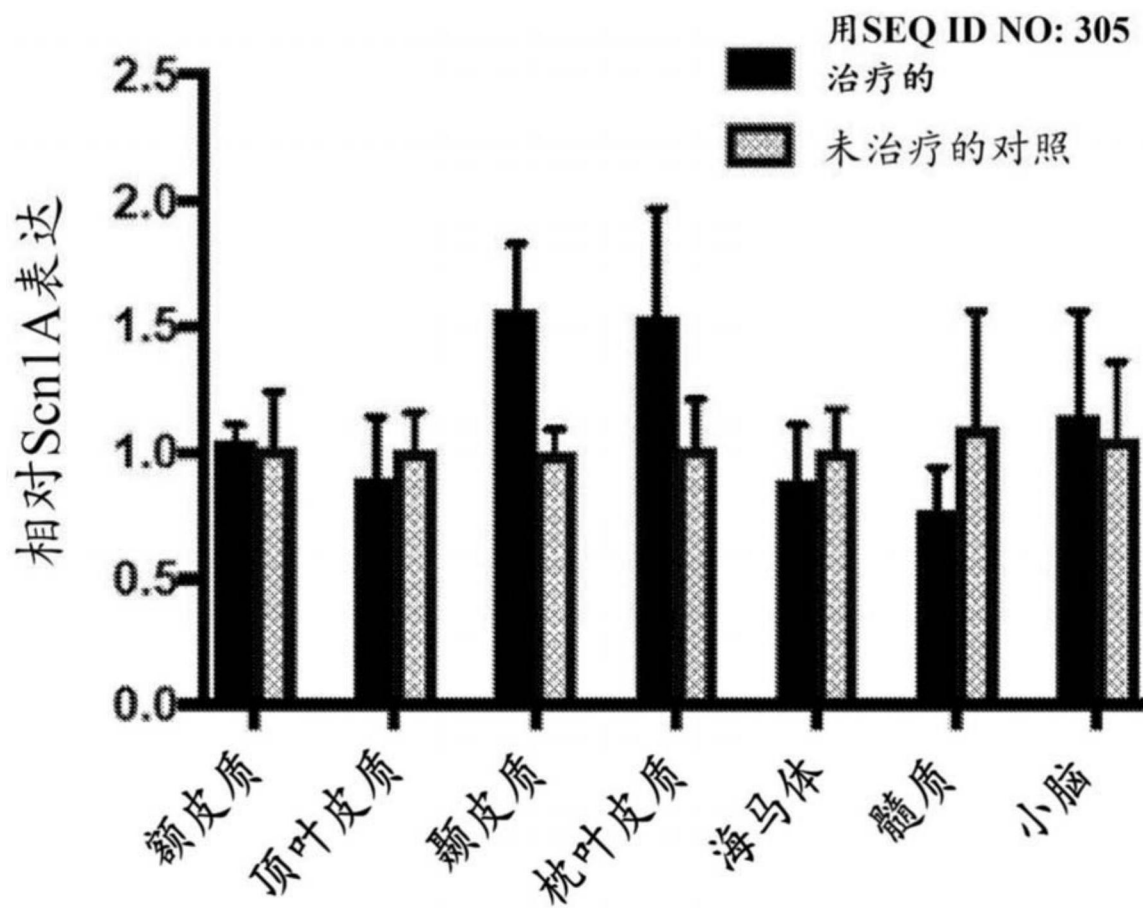


图22

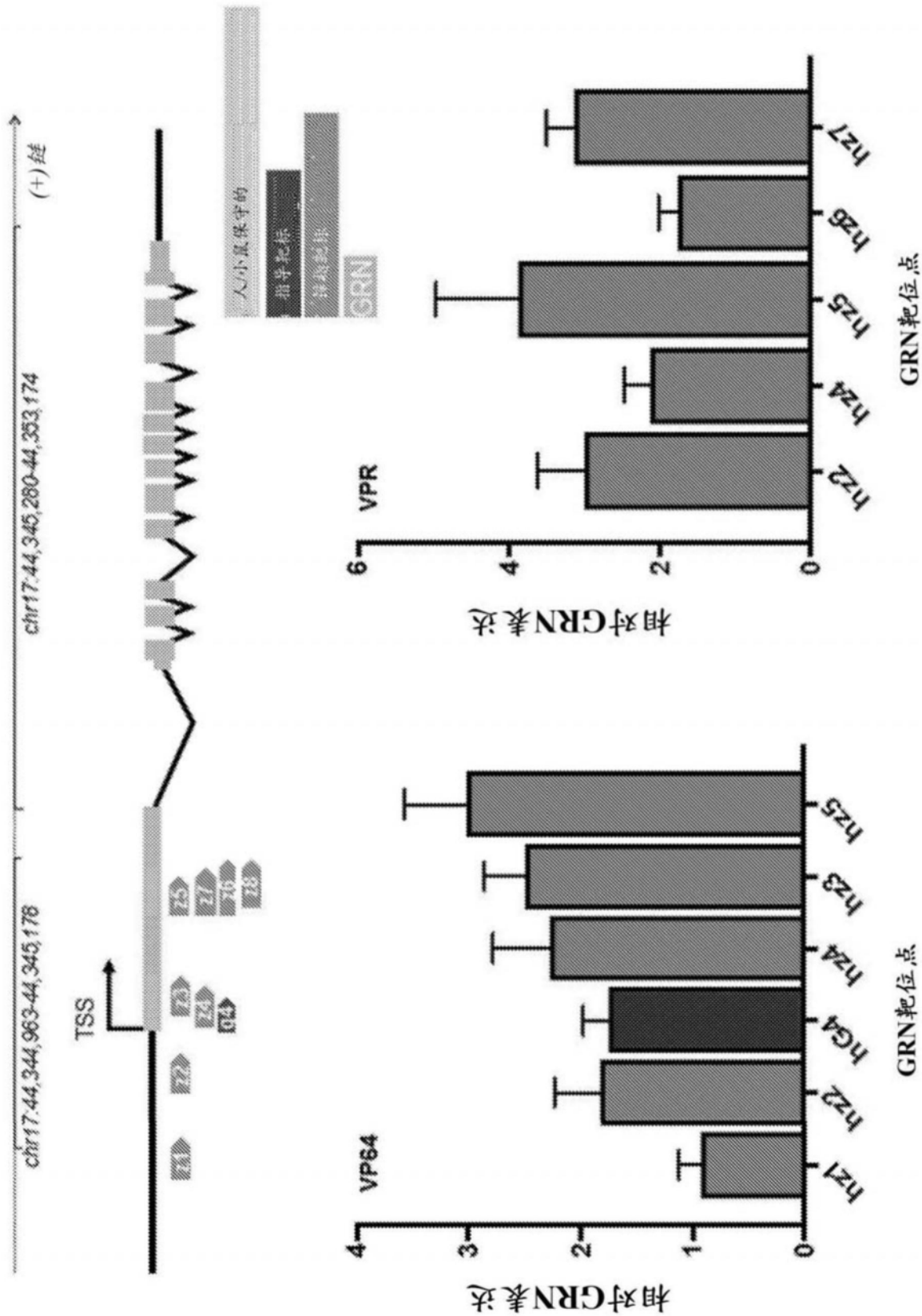


图23

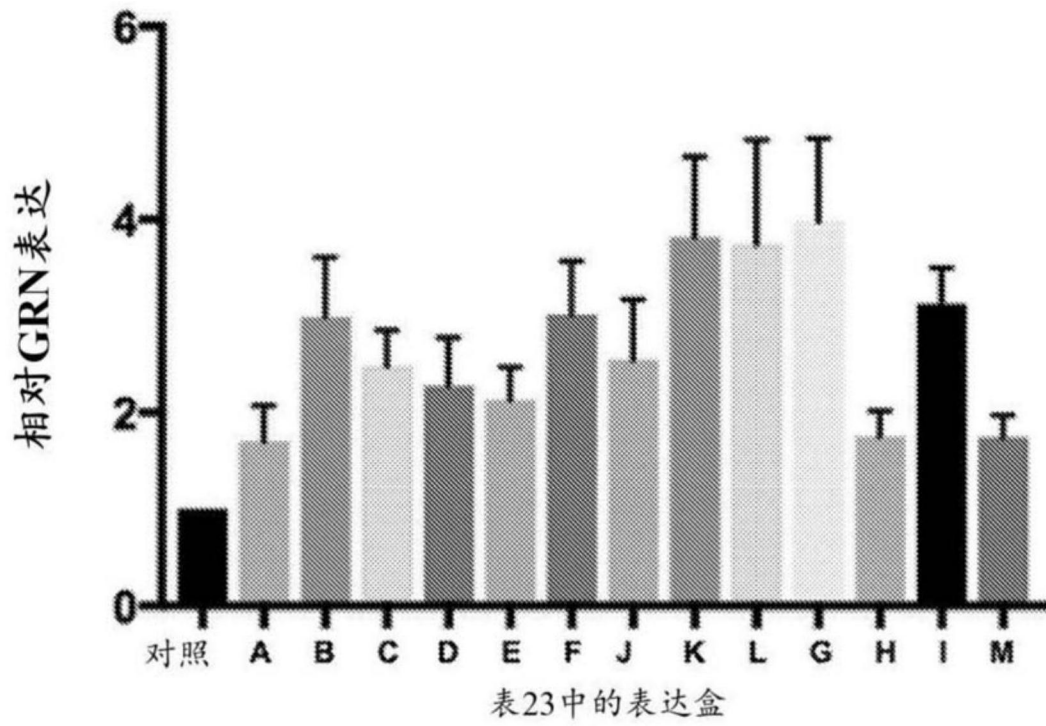
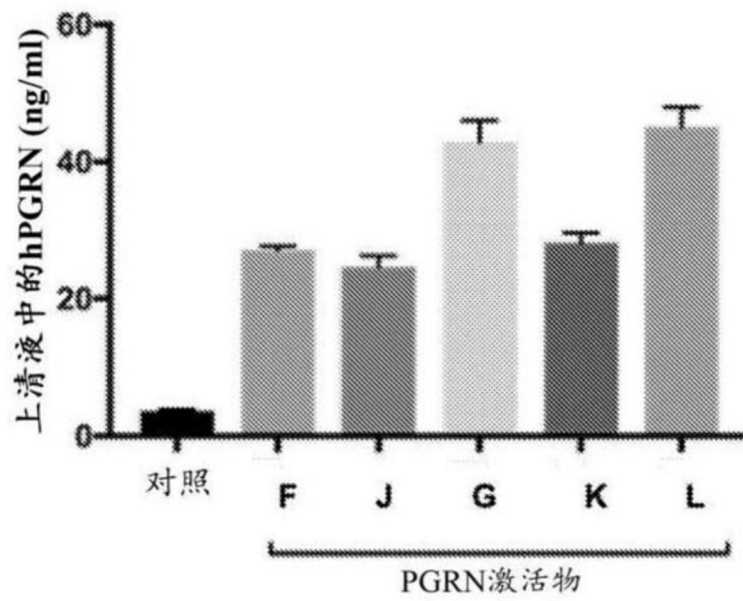


图24A



转染后2天收获培养基
以进行ELISA

图24B

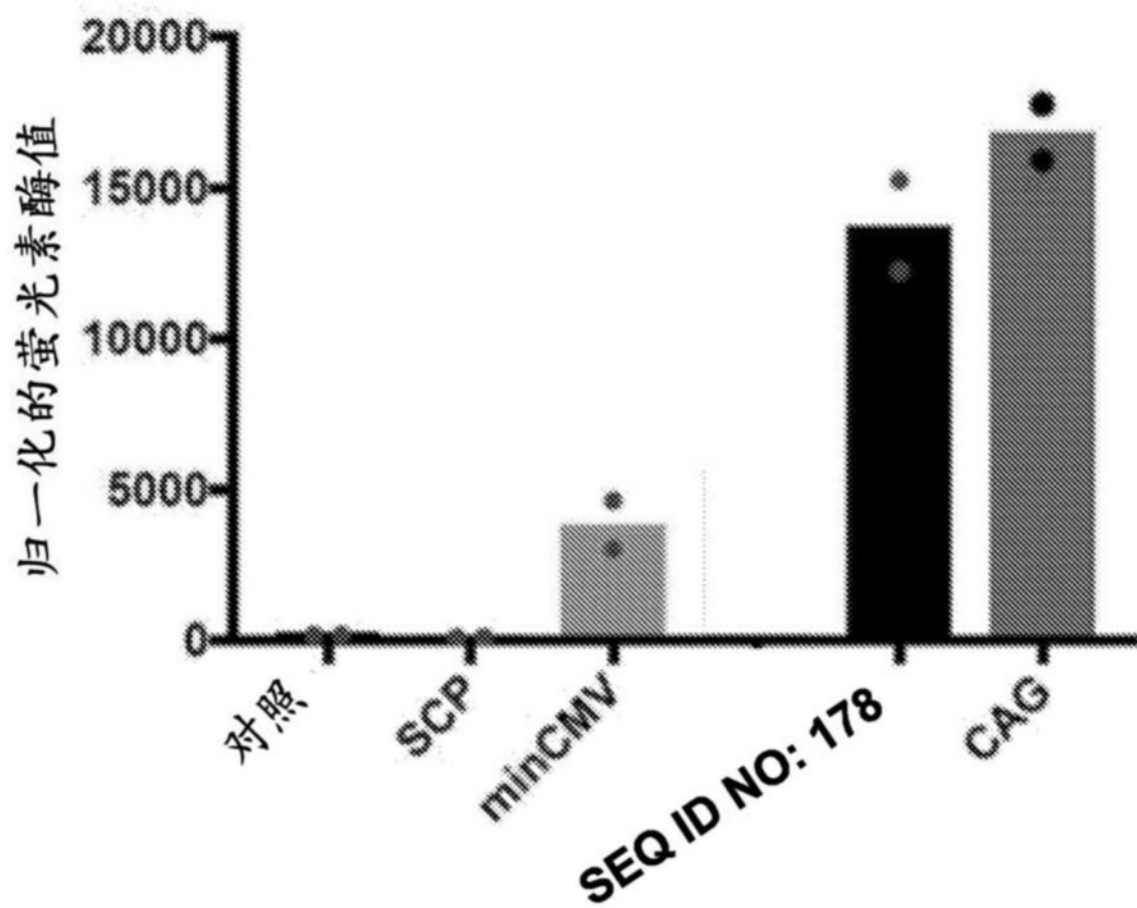


图25

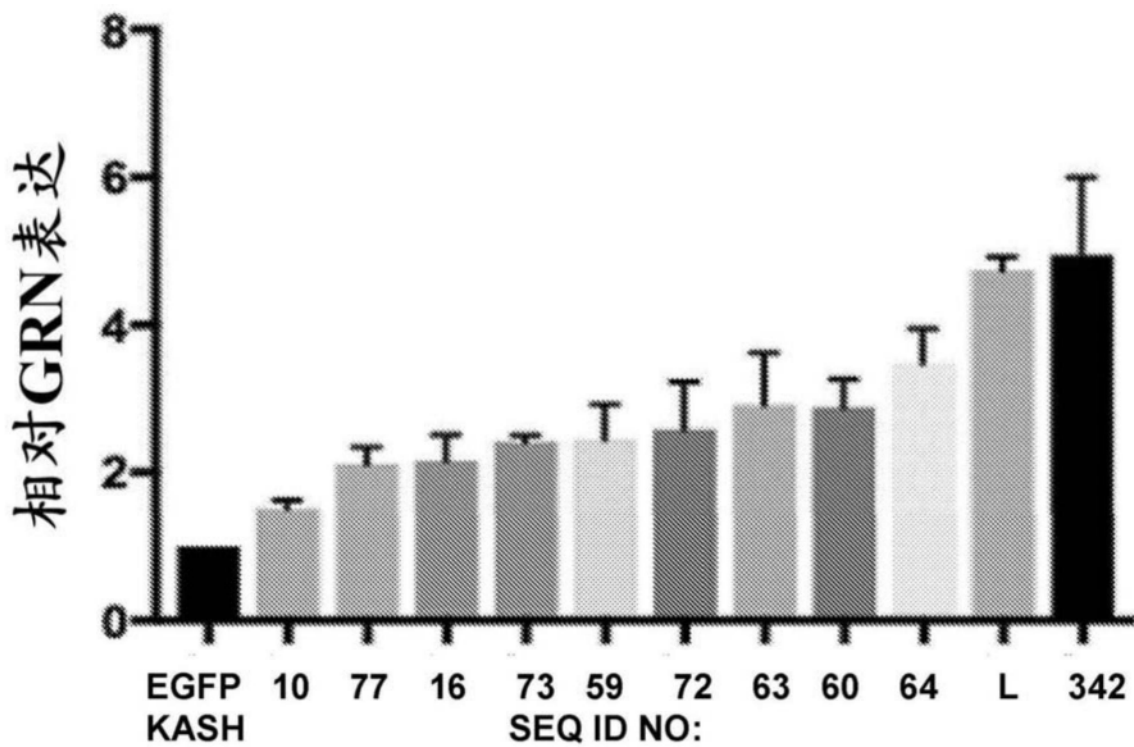


图26

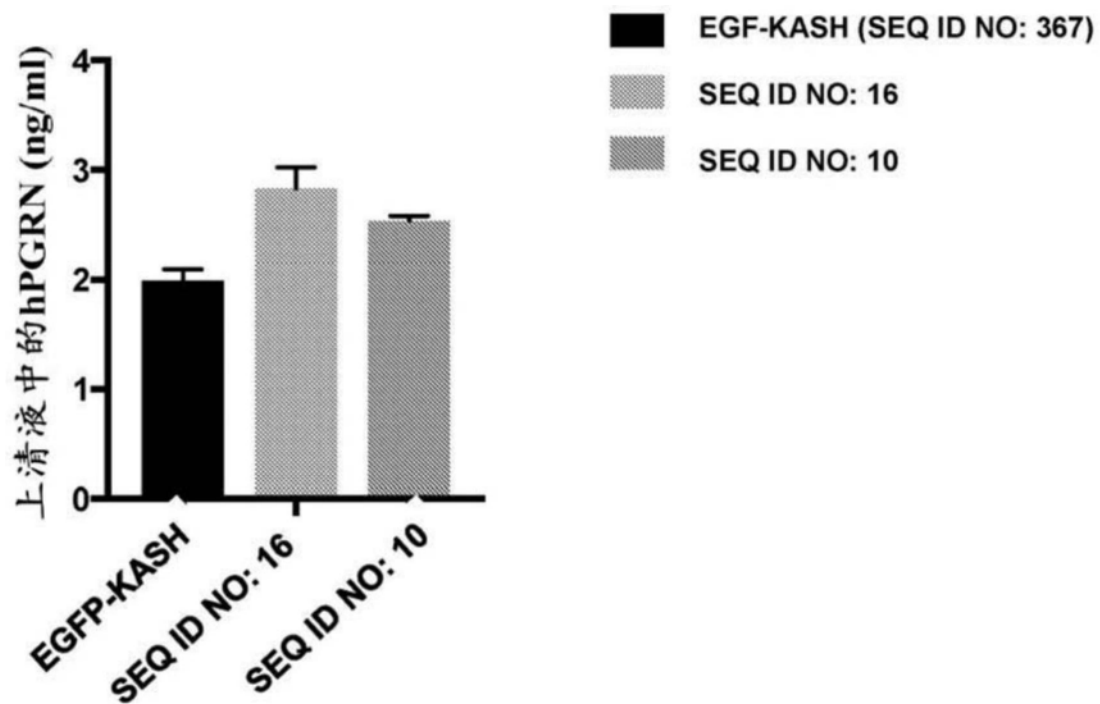


图27

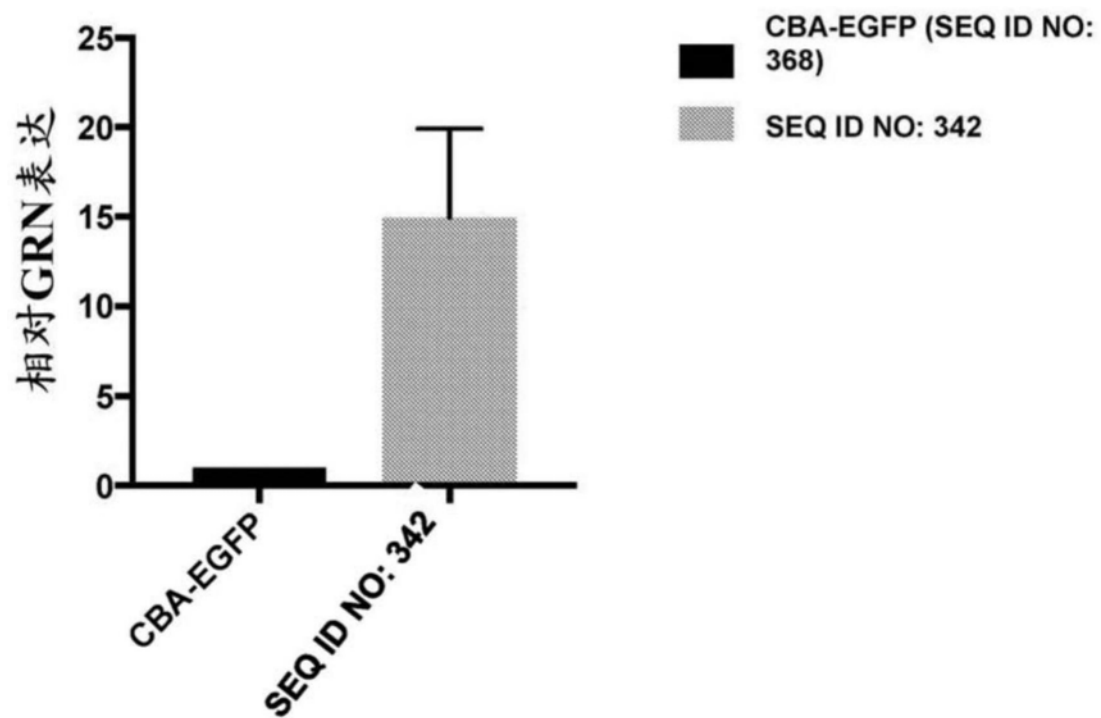


图28

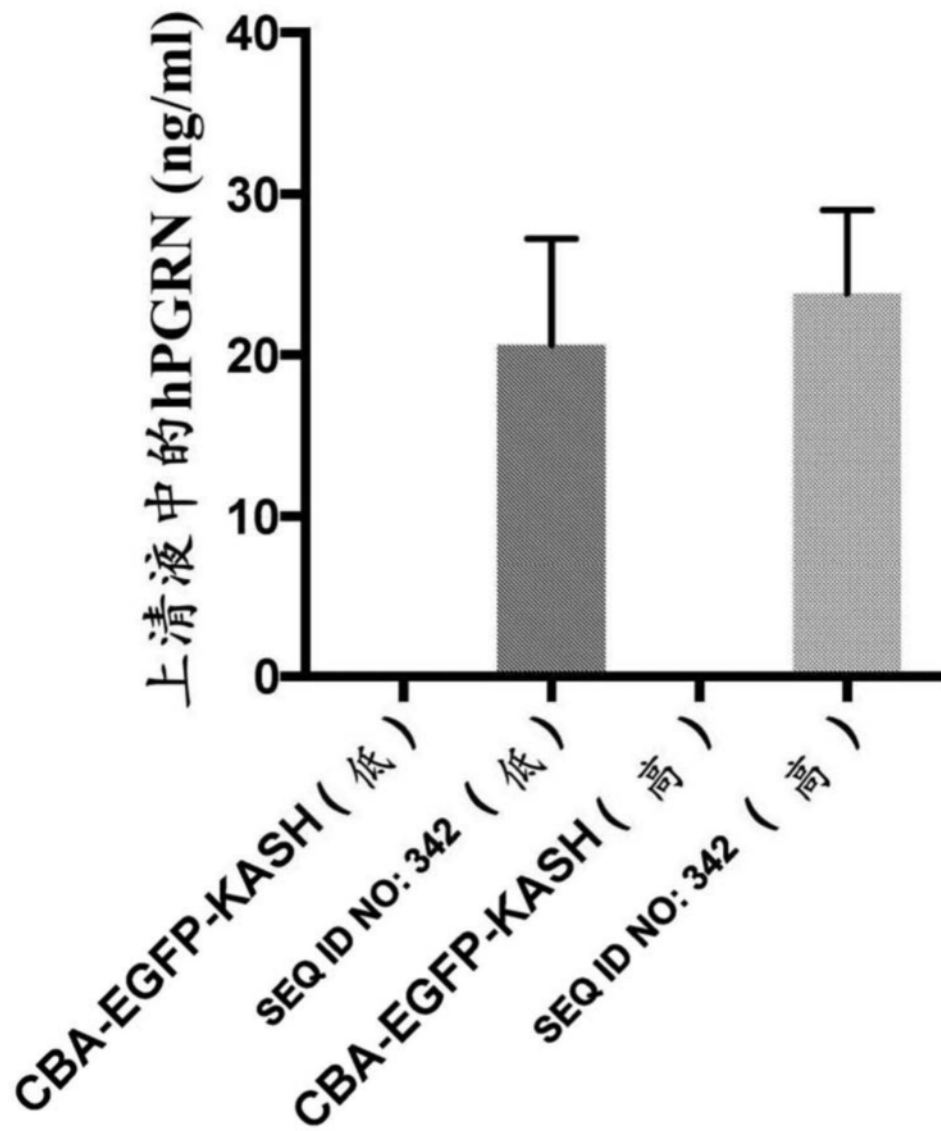


图29

尾静脉注射后1周

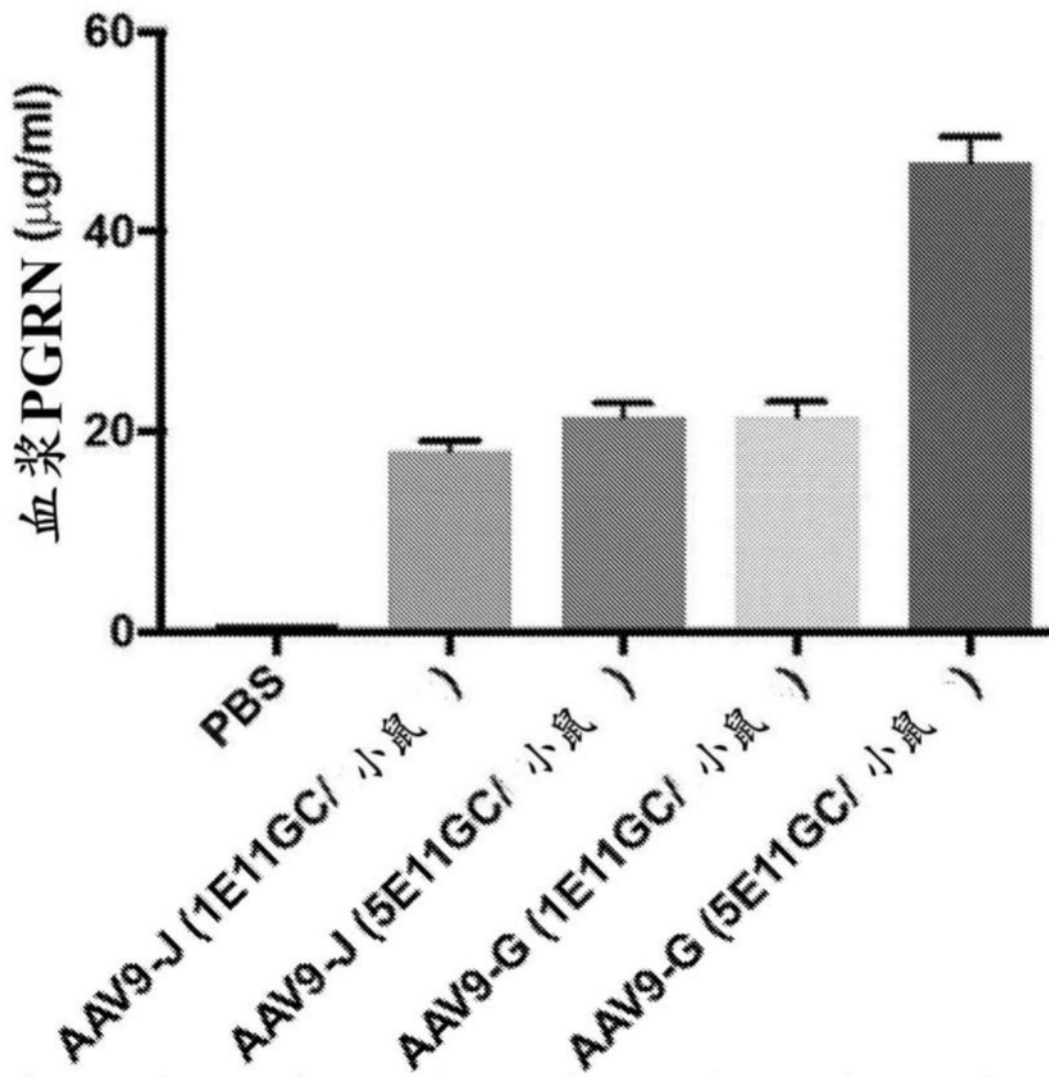


图30

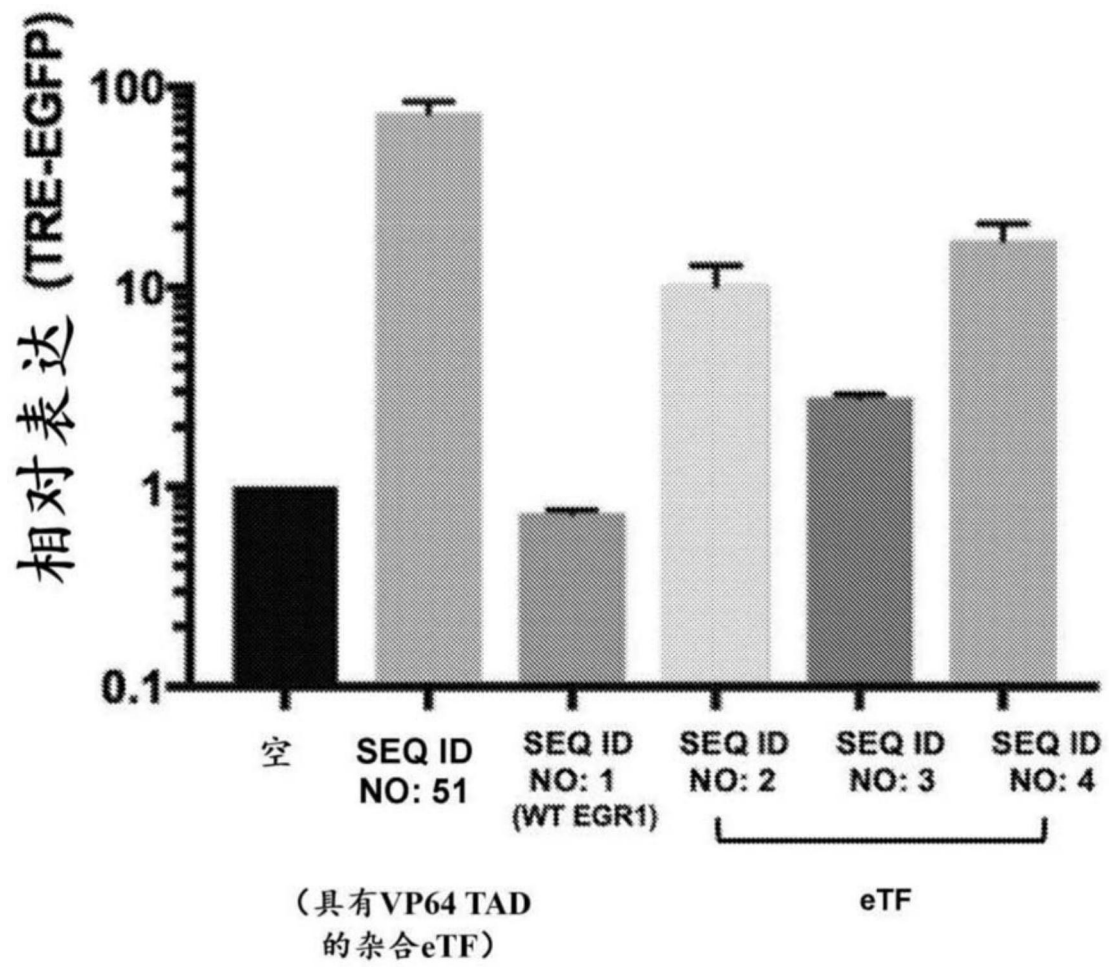


图31A

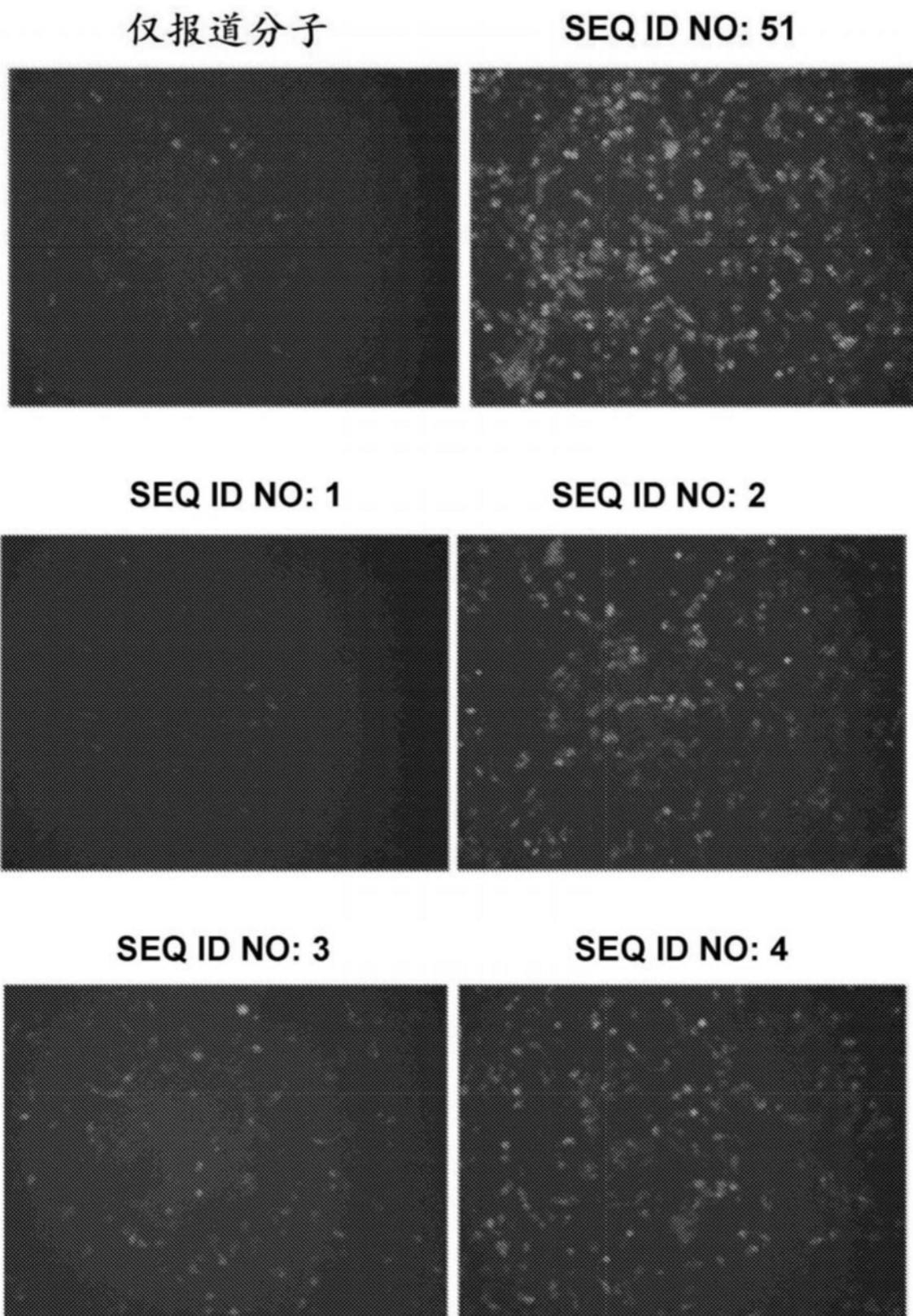


图31B

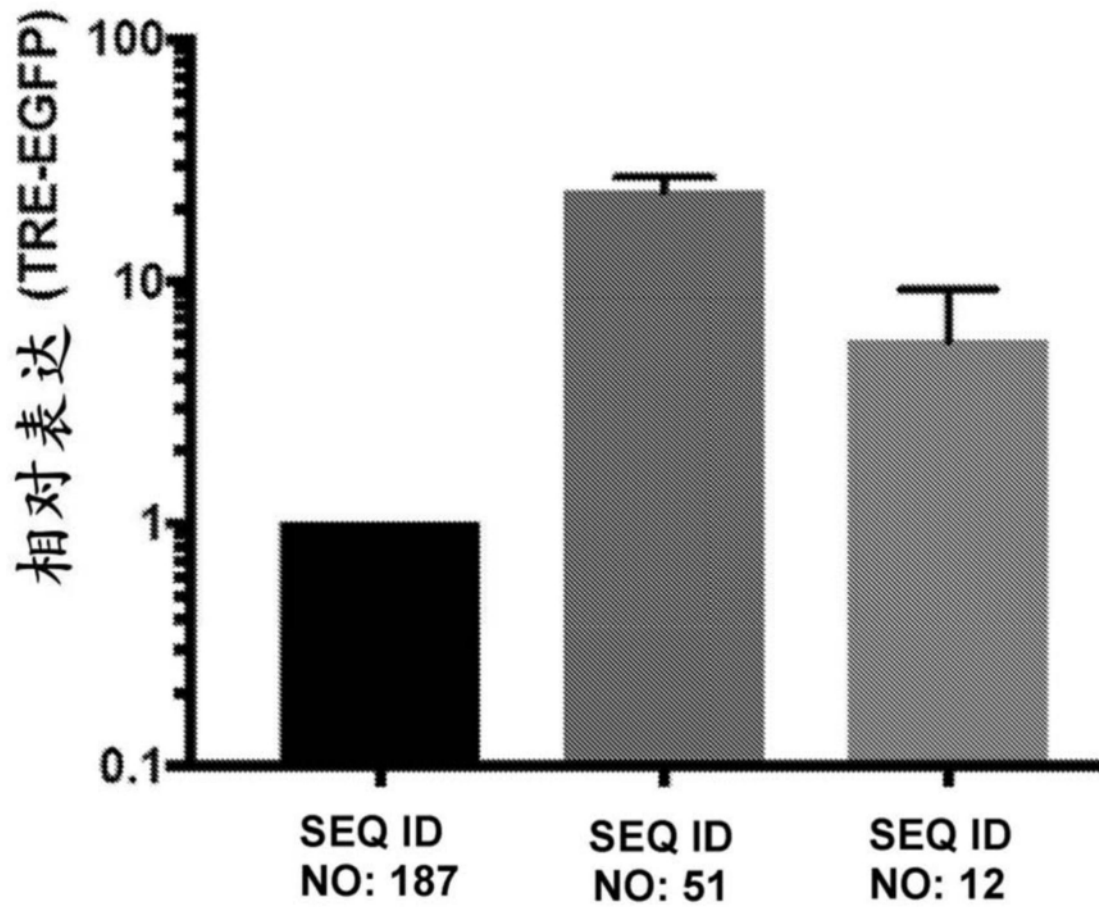


图32

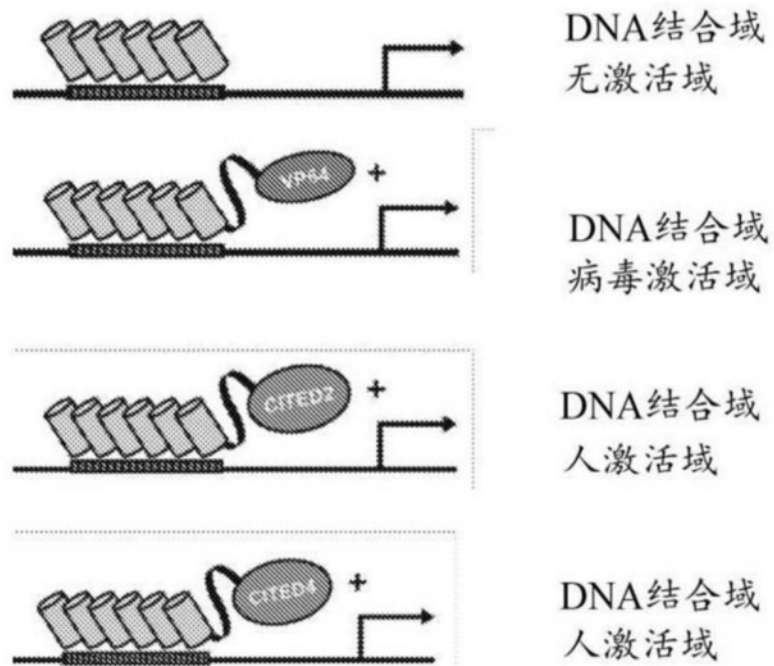


图33A

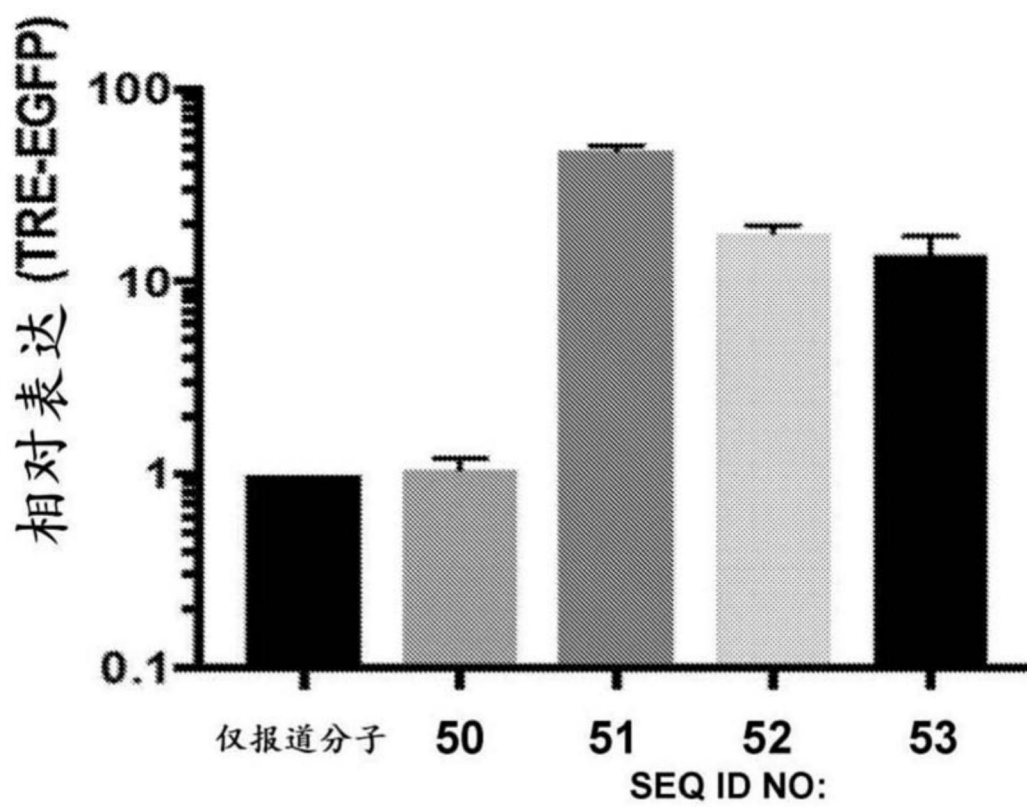


图33B

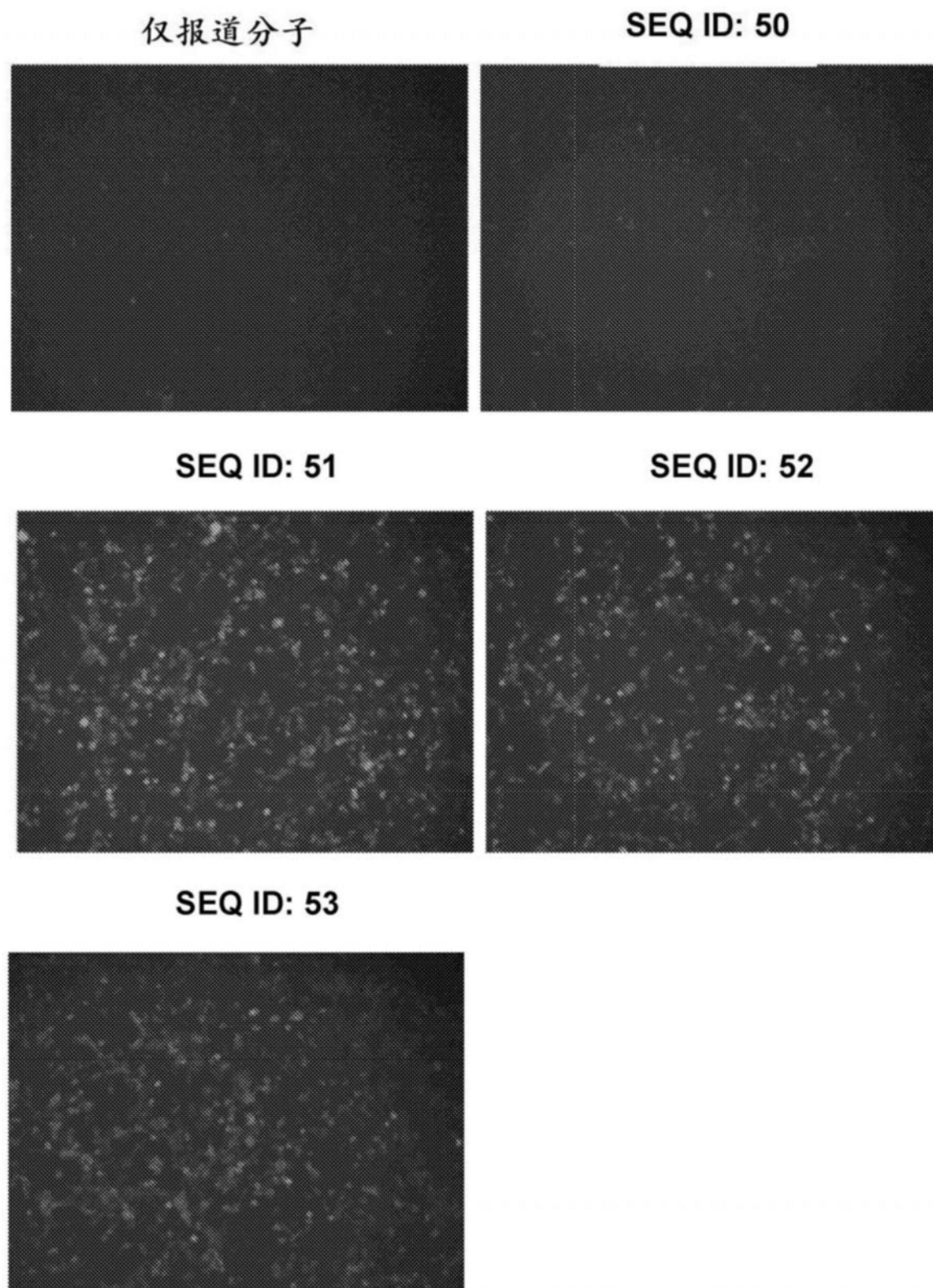
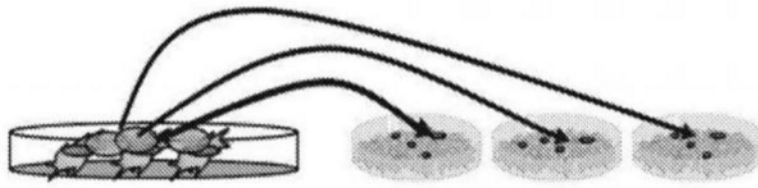
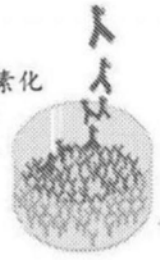


图33C

1. 将抗原刺激的细胞转移到预包被的板上



2. 添加生物素化的抗细胞因子抗体



3. 添加显色试剂



4. 斑点形成，即细胞因子分泌

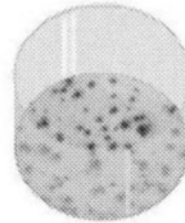


图35A

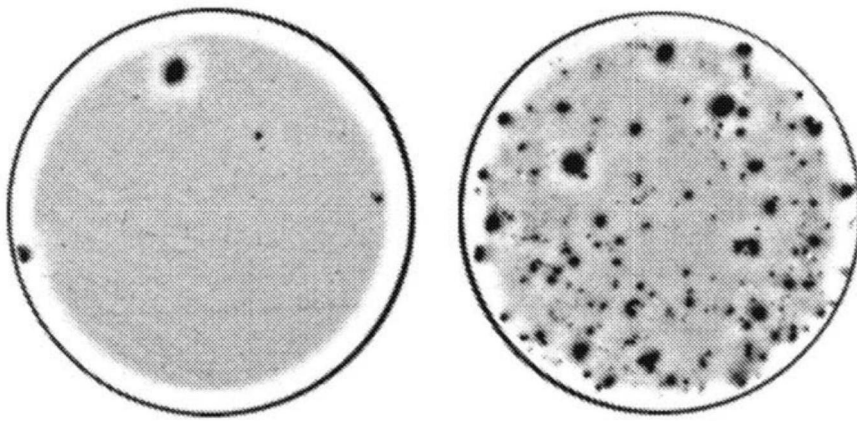


图35B

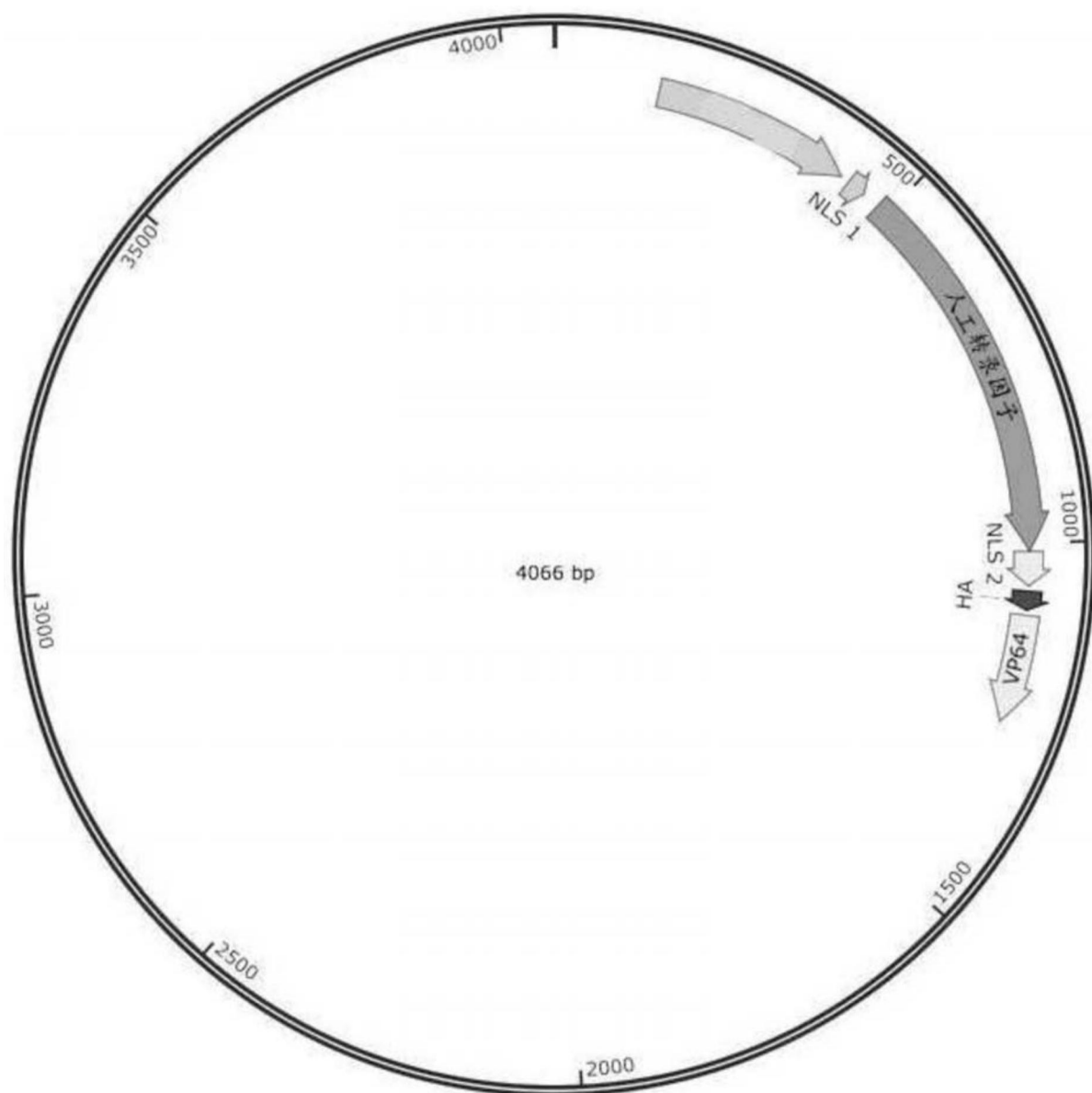


图36

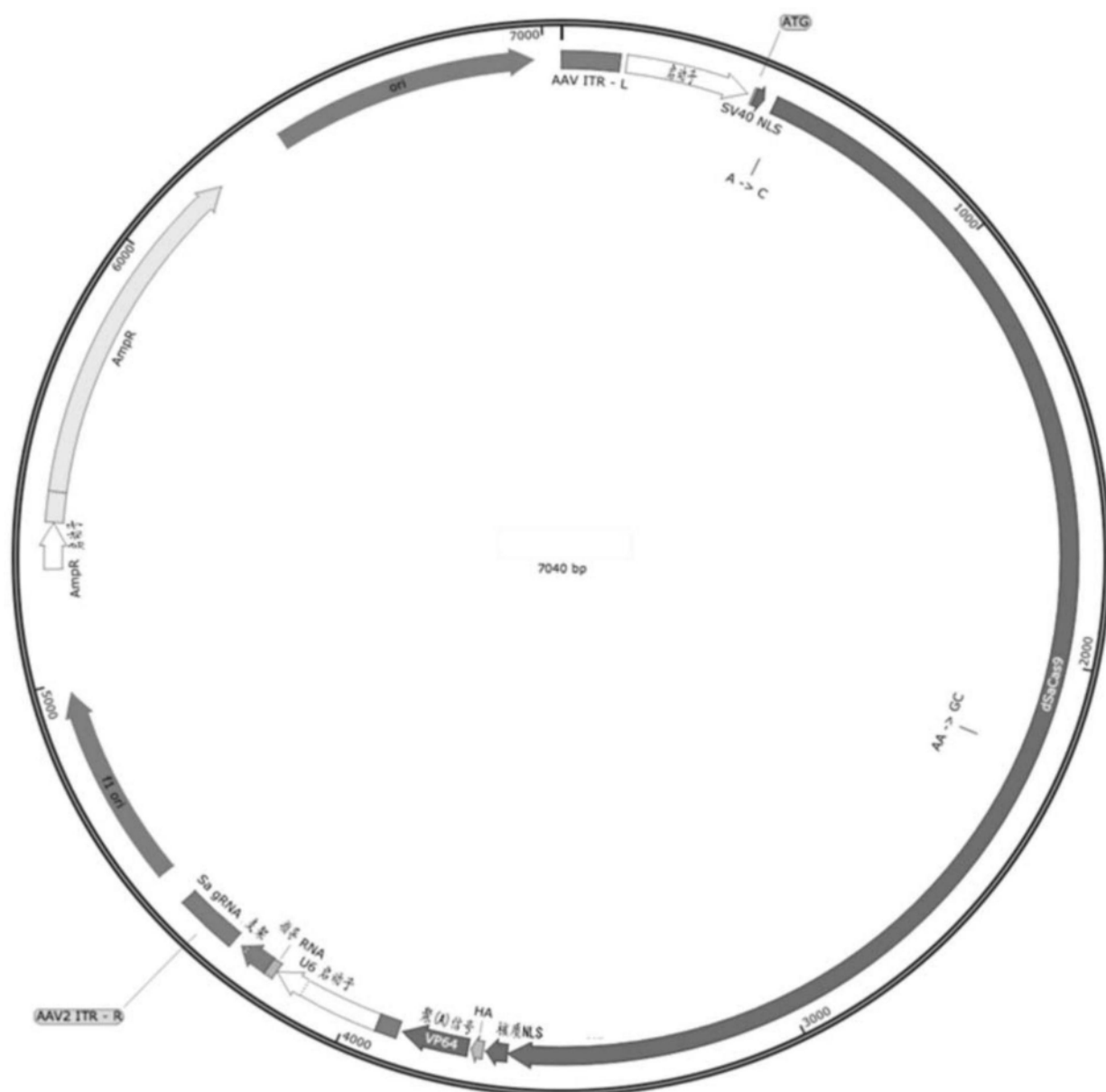


图37