(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117616127 A
(43)申请公布日 2024.02.27

(21)申请号 202280036423.1
(22)申请日 2022.05.04
(30)优先权数据
63/184015 2021.05.04 US
(85)PCT国际申请进入国家阶段日
2023.11.20
(86)PCT国际申请的申请数据
PCT/US2022/027679 2022.05.04
(87)PCT国际申请的公布数据
WO2022/235805 EN 2022.11.10
(71)申请人 分贝治疗公司
地址 美国马萨诸塞州
(72)发明人 J·伯恩斯 T·吉布森 G·普雷格尼格 K·索
(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 任晓华 杨思捷
(51)Int.Cl.
C12N 15/85(2006.01)

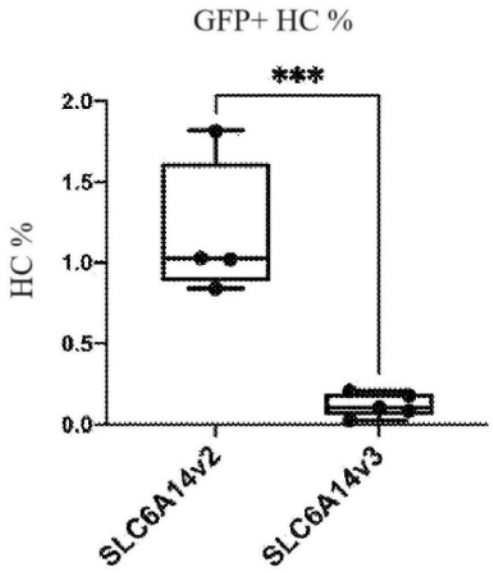
权利要求书2页 说明书52页
序列表16页 附图9页

(54)发明名称

前庭支持细胞启动子和其用途

(57)摘要

本公开提供了含有SLC6A14启动子的多核苷酸以及含有所述多核苷酸的载体,其可用于促进转基因在前庭支持细胞中的表达。本文所述的多核苷酸可以可操作地连接至转基因,例如编码目标蛋白的转基因,以促进所述转基因的前庭支持细胞表达。本文所述的多核苷酸可以可操作地连接至转基因并用于治疗患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者。



CN 117616127 A

1.一种核酸载体,所述核酸载体包含溶质载体家族6成员14(SLC6A14)启动子,其包含与SEQ ID NO:1具有至少85%序列同一性的多核苷酸序列。

2.如权利要求1所述的核酸载体,其中所述SLC6A14启动子具有SEQ ID NO:1的序列。

3.如权利要求1或2所述的核酸载体,其中所述SLC6A14启动子可操作地连接至转基因。

4.如权利要求3所述的核酸载体,其中所述转基因是异源转基因。

5.如权利要求3或4所述的核酸载体,其中所述转基因编码蛋白质、短发夹RNA(shRNA)、反义寡核苷酸(ASO)、核酸酶,或微小RNA。

6.如权利要求5所述的核酸载体,其中所述转基因编码蛋白质。

7.如权利要求6所述的核酸载体,其中所述蛋白质是无调BHLH转录因子1(Atoh1)、婆罗双树样转录因子2(Sall2)、钙调蛋白结合转录激活因子1(Camta1)、具有YRPW基序2的Hes相关家族BHLH转录因子(Hey2)、Gata结合蛋白2(Gata2)、具有YRPW基序1的Hes相关家族BHLH转录因子(Hey1)、神经酰胺合酶2(Lass2)、SRY盒10(Sox10)、GATA结合蛋白3(Gata3)、Cut样同源异形盒1(Cux1)、核受体亚家族2F组成员(Nr2f1)、Hes相关家族BHLH转录因子(Hes1)、RAR相关孤儿受体B(Rorb)、Jun原癌基因AP-1转录因子亚基(Jun)、锌指蛋白667(Zfp667)、LIM同源异形盒3(Lhx3)、无义螺旋-环-螺旋1(Nhlh1)、MAX二聚化蛋白4(Mxd4)、MIZ-1型锌指(Zmiz1)、髓磷脂转录因子1(Myt1)、信号转导子和转录激活因子3(Stat3)、BarH样同源异形盒1(Barhl1)、胸腺细胞选择相关的高迁移率族盒(Tox)、Prospero同源异形盒1(Prox1)、核因子I A(Nfia)、甲状腺激素受体β(Thrb)、MYCL原癌基因BHLH转录因子(Mycl1)、赖氨酸去甲基酶5A(Kdm5a)、CAMP反应元件结合蛋白3样4(Creb3l4)、ETS变体1(Etv1)、父系表达的3(Peg3)、BTB结构域和CNC同源物2(Bach2)、ISLLIM同源异形盒1(Isl1)、锌指和含BTB结构域的38(Zbtb38)、肢芽和心脏发育(Lbh)、Tubby二分转录因子(Tub)、泛素C(Hmg20)、RE1沉默转录因子(Rest)、锌指蛋白827(Zfp827)、AF4/FMR2家族成员3(Aff3)、PBX/有节的1同源异形盒2(Pknox2)、富含AT的相互作用结构域3B(Arid3b)、MLX相互作用蛋白(Mlxip)、锌指蛋白(Zfp532)、IKAROS家族锌指2(Ikzf2)、婆罗双树样转录因子1(Sall1)、SIX同源异形盒2(Six2)、婆罗双树样转录因子3(Sall3)、Lin-28同源物B(Lin28b)、调节因子X7(Rfx7)、脑源性神经营养因子(Bdnf)、生长因子非依赖性1转录阻遏子(Gfi1)、POU4类同源异形盒3(Pou4f3)、MYC原癌基因BHLH转录因子(Myc)、β-连环蛋白(Ctnnb1)、SRY盒2(Sox2)、SRY盒4(Sox4)、SRY盒11(Sox11)、TEA结构域转录因子2(Tead2)或Atoh1变体。

8.如权利要求7所述的核酸载体,其中所述蛋白质是Atoh1。

9.如权利要求3-8中任一项所述的核酸载体,其中所述核酸载体另外包含所述SLC6A14启动子的5'的第一反向末端重复;和所述转基因的3'并且以5'至3'顺序,任选的转录后调节元件、多聚腺苷酸化信号和第二反向末端重复。

10.如权利要求9所述的核酸载体,所述核酸载体包含SEQ ID NO:10的核苷酸219-3831、SEQ ID NO:10的核苷酸219-3831的5'的第一反向末端重复,其中所述5'反向末端重复与SEQ ID NO:10的核苷酸1-130具有至少80%序列同一性;和SEQ ID NO:10的核苷酸219-3831的3'的第二反向末端重复,其中所述3'反向末端重复与SEQ ID NO:10的核苷酸3919-4048具有至少80%序列同一性。

11.如权利要求9所述的核酸载体,所述核酸载体包含SEQ ID NO:11的核苷酸219-3822、SEQ ID NO:11的核苷酸219-3822的5'的第一反向末端重复,其中所述5'反向末端重

复与SEQ ID NO:11的核苷酸1-130具有至少80％序列同一性;和SEQ ID NO:11的核苷酸219-3822的3'的第二反向末端重复,其中所述3'反向末端重复与SEQ ID NO:11的核苷酸3910-4039具有至少80％序列同一性。

12.如权利要求1-11中任一项所述的核酸载体,其中所述核酸载体是质粒。

13.如权利要求1-11中任一项所述的核酸载体,其中所述核酸载体是腺相关病毒(AAV)载体。

14.如权利要求13所述的核酸载体,其中所述AAV载体具有AAV8衣壳。

15.一种药物组合物,所述药物组合物包含权利要求1-14中任一项所述的核酸载体和药学上可接受的载体、稀释剂或赋形剂。

16.一种在哺乳动物前庭支持细胞(VSC)中表达转基因的方法,所述方法包括使所述哺乳动物VSC与权利要求1-14中任一项所述的核酸载体或权利要求15所述的组合物接触。

17.如权利要求16所述的方法,其中所述哺乳动物VSC是人类VSC。

18.一种治疗患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者的方法,所述方法包括向所述受试者的内耳施用有效量的权利要求1-14中任一项所述的核酸载体或权利要求15所述的组合物。

19.一种诱导或增加有需要的受试者的前庭毛细胞再生的方法,所述方法包括向所述受试者的内耳施用有效量的权利要求1-14中任一项所述的核酸载体或权利要求15所述的组合物。

20.一种治疗患有双侧前庭病或具有患上双侧前庭病的风险的受试者的方法,所述方法包括向所述受试者的内耳施用有效量的权利要求1-14中任一项所述的核酸载体或权利要求15所述的组合物。

前庭支持细胞启动子和其用途

[0001] 序列表

[0002] 本申请含有序列表,其已以ASCII格式以电子方式提交并特此通过引用整体并入。ASCII拷贝创建于2022年4月28日,命名为51471-010WO2_Sequence_Listing_4_28_22_ST25并且大小为41,804个字节。

背景技术

[0003] 前庭功能障碍是重大公共健康问题,其对生活质量有深远影响。大约35%的40岁及以上的美国成年人表现出平衡障碍,并且这个比例随年龄增长而显著增加,导致日常活动中断、情绪和认知下降以及老年人跌倒的盛行率增加。前庭功能障碍通常是获得性的,并且具有多种病因,包括疾病或感染、头部创伤、耳毒性药物和衰老。前庭功能障碍病因中的常见因素是内耳前庭毛细胞损伤。因此,旨在恢复毛细胞功能的疗法将有益于患有前庭功能障碍的患者。已知前庭支持细胞在损伤后自发分化为毛细胞,并因此可用作适于恢复毛细胞功能的治疗靶标。

发明内容

[0004] 本发明提供了促进目标基因(例如促进或改进毛细胞或支持细胞功能、再生、成熟、增殖或存活的基因)在特定细胞类型中的表达的组合物和方法。本文所述的组合物和方法涉及可用于诱导转基因在内耳前庭支持细胞(VSC)中的表达的溶质载体家族6成员14(SLC6A14)启动子序列。本文所述的SLC6A14启动子序列可以可操作地连接至转基因,并且可施用于患者以治疗前庭功能障碍(例如眩晕、头晕、失衡、双侧前庭病、双侧前庭功能低下、振动幻视或平衡障碍)。本文所述的SLC6A14启动子序列通过驱动可操作地连接的转基因在前庭支持细胞中的高表达(在其他内耳细胞类型如毛细胞中的表达要低得多)而显示出高的细胞类型特异性。

[0005] 在第一方面,本发明提供了核酸载体,其包含与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸。在一些实施方案中,多核苷酸具有SEQ ID NO:1的至少90%序列。在一些实施方案中,多核苷酸具有SEQ ID NO:1的至少95%序列。在一些实施方案中,多核苷酸具有SEQ ID NO:1的序列。

[0006] 在一些实施方案中,多核苷酸可操作地连接至转基因。在一些实施方案中,转基因是异源转基因。在一些实施方案中,转基因编码蛋白质(例如治疗蛋白或报告蛋白)、短发夹RNA(shRNA)、反义寡核苷酸(ASO)、核酸酶(例如CRISPR相关蛋白9(Cas9)、转录激活因子样效应物核酸酶(TALEN)、锌指核酸酶(ZFN)或向导RNA(gRNA)),或微小RNA。在一些实施方案中,转基因编码蛋白质。

[0007] 在一些实施方案中,多核苷酸引导蛋白质(例如治疗蛋白或报告蛋白)、shRNA、ASO、核酸酶或微小RNA在哺乳动物VSC中的前庭支持细胞(VSC)特异性表达。在一些实施方案中,VSC是人类VSC。

[0008] 在一些实施方案中,由可操作地连接至多核苷酸的转基因编码的蛋白质是婆罗双树样转录因子2(Sall2)、钙调蛋白结合转录激活因子1(Camta1)、具有YRPW基序2的Hes相关家族BHLH转录因子(Hey2)、Gata结合蛋白2(Gata2)、具有YRPW基序1的Hes相关家族BHLH转录因子(Hey1)、神经酰胺合酶2(Lass2)、SRY盒10(Sox10)、GATA结合蛋白3(Gata3)、Cut样同源异形盒1(Cux1)、核受体亚家族2F组成员(Nr2f1)、Hes相关家族BHLH转录因子(Hes1)、RAR相关孤儿受体B(Rorb)、Jun原癌基因AP-1转录因子亚基(Jun)、锌指蛋白667(Zfp667)、LIM同源异形盒3(Lhx3)、无义螺旋-环-螺旋1(Nhlh1)、MAX二聚化蛋白4(Mxd4)、MIZ-1型锌指(Zmiz1)、髓磷脂转录因子1(Myt1)、信号转导子和转录激活因子3(Stat3)、BarH样同源异形盒1(Barhl1)、胸腺细胞选择相关的高迁移率族盒(Tox)、Prospero同源异形盒1(Prox1)、核因子IA(Nfia)、甲状腺激素受体β(Thrb)、MYCL原癌基因BHLH转录因子(Mycl1)、赖氨酸去甲基酶5A(Kdm5a)、CAMP反应元件结合蛋白3样4(Creb3l4)、ETS变体1(Etv1)、父系表达的3(Peg3)、BTB结构域和CNC同源物2(Bach2)、ISL LIM同源异形盒1(Isl1)、锌指和含BTB结构域的38(Zbtb38)、肢芽和心脏发育(Lbh)、Tubby二分转录因子(Tub)、泛素C(Hmg20)、RE1沉默转录因子(Rest)、锌指蛋白827(Zfp827)、AF4/FMR2家族成员3(Aff3)、PBX/有节的1同源异形盒2(Pknox2)、富含AT的相互作用结构域3B(Arid3b)、MLX相互作用蛋白(Mlxip)、锌指蛋白(Zfp532)、IKAROS家族锌指2(Ikzf2)、婆罗双树样转录因子1(Sall1)、SIX同源异形盒2(Six2)、婆罗双树样转录因子3(Sall3)、Lin-28同源物B(Lin28b)、调节因子X7(Rfx7)、脑源性神经营养因子(Bdnf)、生长因子非依赖性1转录阻遏子(Gfi1)、POU4类同源异形盒3(Pou4f3)、MYC原癌基因BHLH转录因子(Myc)、β-连环蛋白(Ctnnb1)、SRY盒2(Sox2)、SRY盒4(Sox4)、SRY盒11(Sox11)、TEA结构域转录因子2(Tead2)、无调BHLH转录因子1(Atoh1)或Atoh1变体。

[0009] 在一些实施方案中,由可操作地连接至多核苷酸的转基因编码的蛋白质是Atoh1或Atoh1变体。在一些实施方案中,Atoh1变体具有相对于SEQ ID NO:4的一个或多个选自由以下组成的组的氨基酸取代:S328A、S331A、S334A、S328A/S331A、S328A/S334A、S331A/S334A和S328A/S331A/S334。在一些实施方案中,蛋白质是Atoh1(例如人类Atoh1)。在一些实施方案中,Atoh1蛋白包含SEQ ID NO:4的序列或其具有一个或多个(例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多个)保守氨基酸取代的变体。在一些实施方案中,Atoh1蛋白包含SEQ ID NO:6的序列或其具有一个或多个(例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多个)保守氨基酸取代的变体。在一些实施方案中,Atoh1蛋白变体中不超过10%(10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更少)的氨基酸是保守氨基酸取代。在一些实施方案中,Atoh1蛋白具有SEQ ID NO:4的序列。在一些实施方案中,Atoh1蛋白是由SEQ ID NO:5的序列编码。在一些实施方案中,Atoh1蛋白具有SEQ ID NO:6的序列。在一些实施方案中,Atoh1蛋白是由SEQ ID NO:7的序列编码。

[0010] 在一些实施方案中,核酸载体还包括反向末端重复序列(ITR)。在核酸载体包括可操作地连接至转基因的本发明多核苷酸的实施方案中,核酸载体包括多核苷酸5'的第一ITR序列和转基因3'的第二ITR序列。在一些实施方案中,ITR是AAV2 ITR。在一些实施方案中,ITR与AAV2 ITR具有至少80%序列同一性(例如至少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)。

[0011] 在一些实施方案中,核酸载体还包括多聚腺苷酸化(poly(A))序列。在一些实施方案中,poly(A)序列是牛生长激素(bGH)poly(A)信号序列。在核酸载体包括可操作地连接至转基因的本发明多核苷酸的实施方案中,poly(A)序列定位于转基因的3'。在核酸载体包括第一和第二ITR序列以及可操作地连接至转基因的本发明多核苷酸的实施方案中,poly(A)序列定位于转基因的3'和第二ITR序列的5'。

[0012] 在一些实施方案中,核酸载体还包括土拨鼠(Woodchuck)转录后调节元件(WPRE)。在一些实施方案中,WPRE具有SEQ ID NO:8或SEQ ID NO:9的序列。在核酸载体包括可操作地连接至转基因的本发明多核苷酸的实施方案中,WPRE定位于转基因的3'。在核酸载体包括可操作地连接至转基因的本发明多核苷酸和poly(A)序列的实施方案中,WPRE定位于转基因的3'和poly(A)序列的5'。

[0013] 在一些实施方案中,核酸载体含有包含SEQ ID NO:10的核苷酸219-3831的序列的多核苷酸序列。

[0014] 在一些实施方案中,核酸载体含有包含SEQ ID NO:11的核苷酸219-3822的序列的多核苷酸序列。

[0015] 在一些实施方案中,本发明的核酸载体包括SLC6A14启动子(例如SEQ ID NO:1的多核苷酸),其可操作地连接至编码人类Atoh1(人类ATOH1蛋白=RefSeq登录号NP_005163(SEQ ID NO:4);mRNA序列=RefSeq登录号NM_005172)的多核苷酸序列。在一些更具体实施方案中,本发明的核酸载体包括SEQ ID NO:1的SLC6A14启动子,其可操作地连接至编码人类Atoh1的多核苷酸序列(例如编码SEQ ID NO:4的多核苷酸序列,例如SEQ ID NO:5的多核苷酸序列)。在一些甚至更具体实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码人类Atoh1的多核苷酸序列;多聚腺苷酸化序列;和第二反向末端重复。在其他更具体实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码人类Atoh1的多核苷酸序列;WPRE;多聚腺苷酸化序列;和第二反向末端重复。在甚至更具体实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:10的核苷酸219-3831。在甚至更具体实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:10的核苷酸219-3831,其中5'反向末端重复与SEQ ID NO:10的核苷酸1-130具有至少80%序列同一性(例如至少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);并且其中3'反向末端重复与SEQ ID NO:10的核苷酸3919-4048具有至少80%序列同一性(例如至少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)。

[0016] 在一些实施方案中,本发明的核酸载体包括SLC6A14启动子(例如SEQ ID NO:1的多核苷酸),其可操作地连接至编码鼠类Atoh1(鼠类ATOH1蛋白=UniProt P48985(SEQ ID NO:6);mRNA序列=RefSeq登录号NM_007500.5)的多核苷酸序列。在一些更具体实施方案中,本发明的核酸载体包括SEQ ID NO:1的SLC6A14启动子,其可操作地连接至编码鼠类Atoh1的多核苷酸序列(例如编码SEQ ID NO:6的多核苷酸序列,例如SEQ ID NO:7的多核苷酸序列)。在一些甚至更具体实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;

SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码鼠类Atoh1的多核苷酸序列;多聚腺苷酸化序列;和第二反向末端重复。在其他更具体实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码鼠类Atoh1的多核苷酸序列;WPRE;多聚腺苷酸化序列;和第二反向末端重复。在甚至更具体实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:11的核苷酸219-3822。在甚至更具体实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:11的核苷酸219-3822,其中5'反向末端重复与SEQ ID NO:11的核苷酸1-130具有至少80%序列同一性(例如至少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);并且其中3'反向末端重复与SEQ ID NO:11的核苷酸3910-4039具有至少80%序列同一性(例如至少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)。

[0017] 在一些实施方案中,核酸载体是病毒载体、质粒、粘粒或人工染色体。在一些实施方案中,核酸载体是选自包括腺相关病毒(AAV)、腺病毒和慢病毒的组的病毒载体。在一些实施方案中,病毒载体是AAV载体。在一些实施方案中,AAV载体具有AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、rh10、rh39、rh43、rh74、Anc80、Anc80L65、DJ/8、DJ/9、7m8、PHP.B、PHP.eB或PHP.S衣壳。在一些实施方案中,AAV载体具有AAV1衣壳。在一些实施方案中,AAV载体具有AAV9衣壳。在一些实施方案中,AAV载体具有AAV6衣壳。在一些实施方案中,AAV载体具有AAV8衣壳。在一些实施方案中,AAV载体具有Anc80衣壳。在一些实施方案中,AAV载体具有Anc80L65衣壳。在一些实施方案中,AAV载体具有DJ/9衣壳。在一些实施方案中,AAV载体具有7m8衣壳。在一些实施方案中,AAV载体具有AAV2衣壳。在一些实施方案中,AAV载体具有PHP.B衣壳。在一些实施方案中,AAV载体具有AAV2quad(Y-F)衣壳。在一些实施方案中,AAV载体具有PHP.S衣壳。在一些实施方案中,AAV载体具有PHP.eB衣壳。在一些实施方案中,AAV载体具有AAV3衣壳。在一些实施方案中,AAV载体具有AAV4衣壳。在一些实施方案中,AAV载体具有AAV5衣壳。在一些实施方案中,AAV载体具有AAV7衣壳。

[0018] 本领域的普通技术人员应理解,产生本发明的病毒载体通常需要使用本发明的质粒以及提供适合的病毒包装和活力所需的元件的额外质粒(例如对于AAV,提供适当的AAV rep基因、cap基因和其他基因(例如E2A和E4)的质粒)。产生细胞系中那些质粒的组合产生病毒载体。然而,本领域技术人员应理解,对于用于产生病毒载体的本发明转移质粒中的任何给定对的反向末端重复序列,病毒载体中的相应序列可因ITR在重组期间采取"翻转(flip)"或"倒转(flop)"取向而改变。因此,转移质粒中ITR的序列不必为在从其制备的病毒载体中发现的相同序列。

[0019] 在另一方面,本发明提供了含有本发明核酸载体的组合物。在一些实施方案中,组合物还包括药学上可接受的载体、稀释剂或赋形剂。

[0020] 在另一方面,本发明提供了与可操作地连接至转基因的SEQ IDNO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸。在一些实施方案中,多核苷酸具有SEQ ID NO:1的序列。

[0021] 在一些实施方案中,转基因是异源转基因。在前述方面的一些实施方案中,转基因编码蛋白质(例如治疗蛋白或报告蛋白)、shRNA、ASO、核酸酶(例如Cas9、TALEN、ZFN或gRNA),或微小RNA。在一些实施方案中,转基因编码蛋白质。

[0022] 在一些实施方案中,蛋白质是Sox9、Sall2、Camta1、Hey2、Gata2、Hey1、Lass2、Sox10、Gata3、Cux1、Nr2f1、Hes1、Rorb、Jun、Zfp667、Lhx3、Nhlh1、Mxd4、Zmiz1、Myt1、Stat3、Barhl1、Tox、Prox1、Nfia、Thrb、Mycl1、Kdm5a、Creb314、Etv1、Peg3、Bach2、Isl1、Zbtb38、Lbh、Tub、Hmg20、Rest、Zfp827、Aff3、Pknox2、Arid3b、Mlxip、Zfp532、Ikzf2、Sall1、Six2、Sall3、Lin28b、Rfx7、Bdnf、Gfi1、Pou4f3、Myc、Ctnnb1、Sox2、Sox4、Sox11、Tead2、Atoh1或Atoh1变体(例如具有一个或多个选自由以下组成的组的氨基酸取代的Atoh1变体:S328A、S331A、S334A、S328A/S331A、S328A/S334A、S331A/S334A和S328A/S331A/S334)。在一些实施方案中,蛋白质是Atoh1(例如人类Atoh1)。在一些实施方案中,Atoh1蛋白包含SEQ ID NO:4的序列或其具有一个或多个(例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多个)保守氨基酸取代的变体。在一些实施方案中,Atoh1蛋白变体中不超过10%(10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更少)的氨基酸是保守氨基酸取代。在一些实施方案中,Atoh1蛋白具有SEQ ID NO:4的序列。在一些实施方案中,Atoh1蛋白是由SEQ ID NO:5的序列编码。

[0023] 在另一方面,本发明提供了包括前述方面和实施方案中任一者的多核苷酸或核酸载体的细胞(例如哺乳动物细胞,例如人类细胞,例如VSC)。在一些实施方案中,细胞是哺乳动物VSC。在一些实施方案中,哺乳动物VSC是人类VSC。在一些实施方案中,多核苷酸与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)。

[0024] 在另一方面,本发明提供了通过使哺乳动物VSC与前述方面和实施方案中任一者的核酸载体或组合物接触而在哺乳动物VSC中表达转基因的方法。在一些实施方案中,转基因在VSC中特异性表达(例如,在比观察到表达的一种或多种其他内耳细胞(例如毛细胞)的百分比高至少2倍、5倍、10倍、50倍、100倍、200倍、300倍、400倍、500倍或更大百分比的VSC中表达)。在一些实施方案中,哺乳动物VSC是人类VSC。

[0025] 在另一方面,本发明提供了通过向患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来治疗受试者的方法。在一些实施方案中,前庭功能障碍是眩晕、头晕、失衡、双侧前庭病(也称为双侧前庭功能低下)、振动幻视或平衡障碍。在一些实施方案中,前庭功能障碍是年龄相关的前庭功能障碍、头部创伤相关的前庭功能障碍、疾病或感染相关的前庭功能障碍或耳毒性药物诱导的前庭功能障碍。在一些实施方案中,前庭功能障碍与遗传突变相关。在一些实施方案中,前庭功能障碍是特发性前庭功能障碍。

[0026] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来诱导或增加受试者的前庭毛细胞再生的方法。

[0027] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来诱导或增加受试者的VSC增殖的方法。

[0028] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来诱导或增加受试者的前庭毛细胞增殖的方法。

[0029] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来诱导或增加受试者的前庭毛细胞成熟的方法。在一些实施方案中,前庭毛细胞是再生前庭毛细胞。

[0030] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来增加受试者的VSC存活的方法。

[0031] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来增加受试者的前庭毛细胞存活的方法。

[0032] 在另一方面,本发明提供了通过向有需要的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来诱导或增加受试者的前庭毛细胞神经支配的方法。

[0033] 在前述方面中任一者的一些实施方案中,受试者患有前庭功能障碍(例如头晕、眩晕、失衡、双侧前庭病(双侧前庭功能低下)、振动幻视或平衡障碍)或具有患上前庭功能障碍的风险。

[0034] 在另一方面,本发明提供了通过向患有双侧前庭病或具有患上双侧前庭病的风险的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来治疗受试者的方法。在一些实施方案中,双侧前庭病是耳毒性药物诱导的双侧前庭病。

[0035] 在前述方面中任一者的一些实施方案中,耳毒性药物选自由以下组成的组:氨基糖苷、抗肿瘤药物、依他尼酸(ethacrynic acid)、呋塞米(furosemide)、水杨酸盐和奎宁。

[0036] 在另一方面,本发明提供了通过向患有振动幻视或具有患上振动幻视的风险的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来治疗受试者的方法。

[0037] 在另一方面,本发明提供了通过向患有平衡障碍(例如失衡)或具有患上平衡障碍(例如失衡)的风险的受试者的内耳施用有效量的前述方面和实施方案中任一者的核酸载体或组合物来治疗受试者的方法。

[0038] 在前述方面中任一者的一些实施方案中,所述方法还包括在施用核酸载体或组合物之前评估受试者的前庭功能。在一些实施方案中,所述方法还包括在施用核酸载体或组合物之后评估受试者的前庭功能。

[0039] 在前述方面中任一者的一些实施方案中,核酸载体或组合物是局部施用的。在一些实施方案中,核酸载体或组合物被施用于半规管。在一些实施方案中,核酸载体或组合物是经鼓室或鼓室内(例如经由经鼓室或鼓室内注射)施用的。在一些实施方案中,核酸载体或组合物被施用于外淋巴或内淋巴,例如经由卵圆窗、圆窗或半规管(例如外半规管),例如施用于前庭支持细胞。在一些实施方案中,本发明的核酸载体或组合物被施用于外淋巴中。在一些实施方案中,本发明的核酸载体或组合物被施用于内淋巴中。在一些实施方案中,本发明的核酸载体或组合物被施用于卵圆窗或经由卵圆窗施用。在一些实施方案中,本发明的核酸载体或组合物被施用于圆窗或经由圆窗施用。

[0040] 在前述方面中任一者的一些实施方案中,核酸载体或组合物是以足以预防或减轻前庭功能障碍、延迟前庭功能障碍的发生、减缓前庭功能障碍的进展、改进前庭功能、增加前庭毛细胞数、增加前庭毛细胞成熟、增加前庭毛细胞增殖、增加前庭毛细胞再生、增加前庭毛细胞神经支配、增加VSC增殖或增加VSC数的量施用。

[0041] 在前述方面中任一者的一些实施方案中,受试者是人类。

[0042] 在另一方面,本发明提供了含有本发明核酸载体或本发明组合物的试剂盒。

[0043] 定义

[0044] 如本文所用,“施用”是指通过任何有效途径提供或给予受试者治疗剂(例如含有可操作地连接至转基因的溶质载体家族6成员14(SLC6A14)启动子的核酸载体)。下文阐述示例性施用途径。

[0045] 如本文所用,短语“向内耳施用”是指通过允许内耳细胞转导的任何途径向受试者提供或给予本文所述的治疗剂。向内耳施用的示例性途径包括施用到外淋巴或内淋巴中,例如到达或通过卵圆窗、圆窗或半规管(例如,水平管),或通过经鼓膜或鼓室内注射,例如施用到前庭,例如施用到前庭支持细胞。

[0046] 如本文所用,术语“细胞类型”是指共享基于基因表达数据在统计学上可分离的表型的细胞群。例如,常见细胞类型的细胞可共享相似的结构和/或功能特征,例如相似的基因活化模式和抗原呈递谱。常见细胞类型的细胞可包括从常见组织(例如上皮组织、神经组织、结缔组织或肌肉组织)分离的那些细胞和/或从常见器官、组织系统、血管或生物体中的其他结构和/或区域分离的那些细胞。

[0047] 如本文所用,术语“保守突变”、“保守取代”和“保守氨基酸取代”是指用一个或多个氨基酸取代一个或多个表现出相似物理化学性质(例如极性、静电荷和空间体积)的不同氨基酸。二十种天然存在氨基酸中每一者的这些性质汇总于下表1中。

[0048] 表1.天然存在氨基酸的代表性物理化学性质

[0049]

氨基酸	3字母代码	1字母代码	侧链极性	生理pH (7.4)下的静电特征	空间体积†

[0050]

氨基酸	3字母代码	1字母代码	侧链极性	生理pH (7.4)下的静电特征	空间体积†
丙氨酸	Ala	A	非极性	中性	小
精氨酸	Arg	R	极性	阳离子	大
天冬酰胺	Asn	N	极性	中性	中等
天冬氨酸	Asp	D	极性	阴离子	中等
半胱氨酸	Cys	C	非极性	中性	中等
谷氨酸	Glu	E	极性	阴离子	中等
谷氨酰胺	Gln	Q	极性	中性	中等
甘氨酸	Gly	G	非极性	中性	小
组氨酸	His	H	极性	在pH 7.4下平衡的中性和阳离子形式两者	大
异亮氨酸	Ile	I	非极性	中性	大
亮氨酸	Leu	L	非极性	中性	大
赖氨酸	Lys	K	极性	阳离子	大
甲硫氨酸	Met	M	非极性	中性	大
苯丙氨酸	Phe	F	非极性	中性	大
脯氨酸	Pro	P	非极性	中性	中等
丝氨酸	Ser	S	极性	中性	小
苏氨酸	Thr	T	极性	中性	中等
色氨酸	Trp	W	非极性	中性	巨大
酪氨酸	Tyr	Y	极性	中性	大
缬氨酸	Val	V	非极性	中性	中等
†基于A^3的体积：50-100为小，100-150为中等，150-200为大，并且>200为巨大					

[0051] 根据此表应了解,保守氨基酸家族包括(i)G、A、V、L和I;(ii)D和E;(iii)C、S和T;(iv)H、K和R;(v)N和Q;以及(vi)F、Y和W。因此,保守突变或取代是用一种氨基酸取代同一氨基酸家族的成员者(例如用Ser取代Thr或用Lys取代Arg)。

[0052] 如本文所用,本文所述的术语组合物、载体构建体或病毒载体的“有效量”、“治疗有效量”和“足量”是指在施用于受试者(包括哺乳动物,例如人类)时足以实现有益或期望结果(包括临床结果)的量,并且因此“有效量”或其同义词取决于应用其的情形。例如,在治疗前庭功能障碍的情形中,其是与未施用组合物、载体构建体或病毒载体时获得的反应相比,足以实现治疗反应的组合物、载体构建体或病毒载体的量。将对应于此量的本文所述的给定组合物的量将根据多种因素而变化,例如给定剂、药物制剂、施用途径、疾病或病症的类型、受试者的身份(例如年龄、性别、体重)或所治疗的宿主等,但仍可以常规方式由本领

域技术人员确定。此外,如本文所用,本公开的组合物、载体构建体或病毒载体的“治疗有效量”是与对照相比在受试者中产生有益或期望结果的量。如本文所定义,本公开的组合物、载体构建体或病毒载体的治疗有效量可容易地由本领域的普通技术人员通过本领域中已知的常规方法来确定。可调整给药方案以提供最佳治疗反应。

[0053] 如本文所用,术语“内源”阐述天然存在于特定生物体(例如人类)或生物体内的特定位置(例如器官、组织或细胞,例如人类细胞,例如人类前庭支持细胞)中的分子(例如多肽、核酸或辅因子)。

[0054] 如本文所用,术语“表达”是指以下事件中的一者或多者:(1)从DNA序列产生RNA模板(例如通过转录);(2)RNA转录物的加工(例如通过剪接、编辑、5'帽形成和/或3'端加工);(3)将RNA翻译成多肽或蛋白质;以及(4)多肽或蛋白质的翻译后修饰。

[0055] 如本文所用,术语“外源”阐述并非天然存在于特定生物体(例如人类)或生物体内的特定位置(例如器官、组织或细胞,例如人类细胞,例如人类前庭支持细胞)中的分子(例如多肽、核酸或辅因子)。外源材料包括从生物体的外部来源或从其提取的培养物提供的那些材料。

[0056] 如本文所用,术语“外显子”是指基因编码区内的区域,其核苷酸序列决定相应蛋白质的氨基酸序列。术语外显子也指RNA的从基因转录的相应区域。外显子转录成前mRNA,并且可包括在成熟mRNA中,这取决于基因的替代剪接。在加工后包括在成熟mRNA中的外显子被翻译成蛋白质,其中外显子的序列决定蛋白质的氨基酸组成。

[0057] 如本文所用,术语“异源”是指非天然元件的组合。例如,异源转基因是指并非由与其可操作地连接的启动子天然表达的转基因。

[0058] 如本文所用,术语“增加”和“减少”是指调节分别产生度量的功能、表达或活性相对于参照更大或更小的量。例如,在本文所述的方法中施用组合物后,受试者中如本文所述度量(例如转基因表达)的标志物的量相对于施用前标志物的量可增加或减少至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或98%或更大。通常,度量是在施用后施用已具有所列举效应时(例如在治疗方案开始后至少一周、一个月、3个月或6个月)进行测量。

[0059] 如本文所用,“局部地”或“局部施用”意指旨在用于局部效应而非全身效应的身体特定位点处的施用。局部施用的实例是皮上、吸入、关节内、鞘内、阴道内、玻璃体内、子宫内、病灶内施用、淋巴结施用、肿瘤内施用、施用于中耳或内耳和施用于受试者的粘膜,其中施用旨在具有局部效应而非全身效应。

[0060] 如本文所用,术语“可操作地连接”是指接合至第二分子的第一分子,其中所述分子如此排列以使得第一分子影响第二分子的功能。两种分子可为或可不为单个邻接分子的一部分并且可为或可不为相邻的。例如,如果启动子调节目标可转录多核苷酸分子在细胞中的转录,则启动子可操作地连接至可转录多核苷酸分子。另外,如果接合转录调节元件的两部分使得一部分的转录活化功能不因另一部分的存在而受到不利影响,则所述两部分彼此可操作地连接。两个转录调节元件可借助于接头核酸(例如间插的非编码核酸)彼此可操作地连接或者可在不存在间插的核苷酸时彼此可操作地连接。

[0061] 如本文所用,术语“质粒”是指额外DNA区段可连接至其中的染色体外环状双链DNA分子。质粒是一种载体,其是能够运输与其连接的另一核酸的核酸分子。某些质粒能够在引

入其的宿主细胞中自主复制(例如具有细菌复制起点的细菌质粒和游离型哺乳动物质粒)。其他载体(例如非游离型哺乳动物载体)可在引入至宿主细胞中后整合至宿主细胞的基因组中,并由此与宿主基因组一起复制。某些质粒能够引导与其可操作地连接的基因的表达。

[0062] 如本文所用,术语“多核苷酸”是指核苷的聚合物。通常,多核苷酸由天然存在于通过磷酸二酯键接合的DNA或RNA中的核苷(例如腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷和脱氧胞苷)组成。所述术语涵盖包含含有经化学或生物修饰的碱基、经修饰主链等的核苷或核苷类似物的分子,无论是否发现于天然存在的核酸中,并且此类分子对于某些应用可为优选的。当此应用涉及多核苷酸时应理解,提供DNA、RNA和在每种情况下单链和双链形式(和每个单链分子的补体)。如本文所用的“多核苷酸序列”可指多核苷酸材料自身和/或生物化学表征特定核酸的序列信息(即用作碱基的缩写的字母串)。除非另有指示,否则本文所呈现的多核苷酸序列是以5’至3’方向呈现。

[0063] 如本文所用,术语“启动子”是指DNA上由RNA聚合酶结合的识别位点。聚合酶驱动转基因的转录。

[0064] 相对于参照多核苷酸或多肽序列的“序列同一性百分比(%)”定义为在比对序列和引入空位(如果需要)以实现最大序列同一性百分比后,候选序列中与参照多核苷酸或多肽序列中的核酸或氨基酸同一的核酸或氨基酸的百分比。出于确定核酸或氨基酸序列同一性百分比的目的,比对可以本领域技术人员的能力范围内的各种方式来实现,例如使用可公开获得的计算机软件,例如BLAST、BLAST-2或Megalign软件。本领域技术人员可确定适用于比对序列的参数,包括在所比较序列的全长范围内实现最大比对所需的任何算法。例如,序列同一性百分比值可使用序列比较计算机程序BLAST来产生。作为说明,给定核酸或氨基酸序列A相对于(to)、与(with)或针对(against)给定核酸或氨基酸序列B的序列同一性%(替代地可表达为相对于、与或针对给定核酸或氨基酸序列B具有特定序列同一性%的给定核酸或氨基酸序列A)计算如下:

[0065] 100×(分数X/Y)

[0066] 其中X是在A与B的程序比对中由序列比对程序(例如BLAST)评定为同一性匹配的核苷酸或氨基酸的数量,并且其中Y是B中核酸的总数。应了解,当核酸或氨基酸序列A的长度不等于核酸或氨基酸序列B的长度时,A相对于B的序列同一性%将不等于B相对于A的序列同一性%。

[0067] 如本文所用,术语“药物组合物”是指含有治疗剂、任选地与一种或多种药学上可接受的赋形剂、稀释剂和/或载体组合的混合物,有待施用于受试者(例如哺乳动物,例如人类)来预防、治疗或控制侵袭或可能侵袭受试者的特定疾病或疾患。

[0068] 如本文所用,术语“药学上可接受”是指适于与受试者(例如哺乳动物,例如人类)的组织接触而无过度毒性、刺激性、过敏反应和其他问题并发症并且与合理益处/风险比相称的那些化合物、材料、组合物和/或剂型。

[0069] 如本文所用,术语“转录调节元件”是指至少部分地控制目标基因的转录的核酸。转录调节元件可包括启动子、增强子和控制或帮助控制基因转录的其他核酸(例如多聚腺苷酸化信号)。转录调节元件的实例阐述于例如Lorence,Recombinant Gene Expression: Reviews and Protocols(HumanaPress,NewYork,NY,2012)中。

[0070] 如本文所用,术语“转染”是指常用于将外源DNA引入原核或真核宿主细胞中的众

多种技术中的任一者,例如电穿孔、脂转染、磷酸钙沉淀、DEAE聚葡萄糖转染、Nucleofection、挤压穿孔、声致穿孔、光学转染、磁性转染、刺穿转染等。

[0071] 如本文所用,术语“受试者”和“患者”是指动物(例如哺乳动物,例如人类)。有待根据本文所述的方法治疗的受试者可为已经诊断患有前庭功能障碍(例如头晕、眩晕或失衡)的受试者或具有患上这些疾患的风险的受试者。诊断可通过本领域中已知的任何方法或技术来实施。本领域技术人员应理解,有待根据本公开治疗的受试者可能已经受标准测试或者可能在无检查的情况下已经鉴定为因存在一种或多种与疾病或疾患相关的风险因子而具有风险的受试者。

[0072] 如本文所用,术语“转导(transduction)”和“转导(transduce)”是指将载体构建体或其一部分引入细胞中的方法。其中载体构建体含于病毒载体(例如AAV载体)中,转导是指用病毒感染细胞并随后将载体构建体或其一部分转移和整合至细胞基因组中。

[0073] 如本文所用,关于疾病或疾患的“治疗(treatment)”和“治疗(treating)”是指用于获得有益或期望结果(例如临床结果)的方法。有益或期望结果可包括(但不限于)缓和或改善一种或多种症状或疾患;降低疾病或疾患的程度;稳定(即不恶化)疾病、病症或疾患的状态;防止疾病或疾患的扩散;延迟或减缓疾病或疾患的进展;改善或减轻疾病或疾患;以及缓解(无论部分或完全),无论可检测或不可检测。“改善”或“减轻”疾病或疾患意指与无治疗时的程度或时程相比,减小疾病、病症或疾患的程度和/或不期望临床表现和/或减缓或延长进展的时程。“治疗”也可意指与未接受治疗时的预期存活相比提供存活。需要治疗的那些包括已患有疾患或病症的那些以及易患疾患或病症的那些或待预防疾患或病症的那些。

[0074] 如本文所用,术语“载体”包括核酸载体,例如DNA载体,例如质粒、粘粒或人工染色体、RNA载体、病毒或任何其他适宜复制子(例如病毒载体)。已开发出多种载体用于将编码外源蛋白质的多核苷酸递送至原核或真核细胞中。此类表达载体的实例阐述于例如Gellissen,Production of Recombinant Proteins:Novel Microbial and Eukaryotic Expression Systems(John Wiley&Sons,Marblehead,MA,2006)。适用于本文所述组合物和方法的表达载体含有多核苷酸序列以及例如用于表达蛋白质和/或将这些多核苷酸序列整合至哺乳动物细胞的基因组中的额外序列元件。可用于表达如本文所述转基因的某些载体包括含有引导基因转录的调节序列(例如启动子和增强子区域)的载体。可用于表达转基因的其他载体含有增强转基因的翻译速率或改进源自基因转录的mRNA的稳定性或核输出的多核苷酸序列。这些序列元件包括例如5’和3’非翻译区和引导表达载体上所携带基因的有效转录的多聚腺苷酸化信号位点。适用于本文所述组合物和方法的表达载体也可含有编码用于选择含有这种载体的细胞的标志物的多核苷酸。合适的标志物的实例包括编码抗生素(例如氨苄青霉素(ampicillin)、氯霉素(chloramphenicol)、卡那霉素(kanamycin)或诺尔丝菌素(nourseothricin))抗性的基因。

[0075] 如本文所用,术语“前庭支持细胞”和“VSC”是指内耳的前庭系统中参与前庭毛细胞发育、存活、功能、死亡和吞噬作用的特化上皮细胞的集合。VSC通过将前庭毛细胞锚定在感觉上皮中并释放对于毛细胞神经支配至关重要的神经营养因子来为所述前庭毛细胞提供结构支持。

[0076] 如本文所用,术语“前庭支持细胞特异性表达”和“VSC特异性表达”是指与内耳的

其他细胞类型(例如前庭毛细胞、耳蜗毛细胞、耳蜗支持细胞、神经胶质或其他内耳细胞类型)相比,主要在前庭支持细胞内产生RNA转录物或多肽。转基因的VSC表达可通过使用任何标准技术(例如定量RT PCR、免疫组织化学、蛋白质印迹分析(western blot analysis)或测量可操作地连接至启动子的报告基因(例如GFP)的荧光)比较内耳的各种细胞类型之间(例如VSC对非VSC细胞)的转基因表达(例如RNA或蛋白质表达)来确认。诱导VSC特异性表达的启动子("VSC特异性启动子")是这样的启动子,其(i)在VSC中诱导的与其可操作地连接的转基因的表达(例如RNA或蛋白质表达)大至少50%(例如大50%、75%、100%、125%、150%、175%、200%或更大),或(ii)在VSC中诱导的与其可操作地连接的转基因的表达大至少2倍(例如大5倍、10倍、50倍、100倍、200倍、300倍、400倍、500倍或更大),各自与以下内耳细胞类型中的至少3者(例如3者、4者、5者、6者、7者、8者、9者、10者或更多者)相比:前庭神经节细胞、前庭器官的非感觉上皮细胞、前庭器官的暗细胞、前庭器官的间充质细胞、螺旋神经节细胞、边缘细胞、内指细胞、内柱细胞、外柱细胞、第一排戴特氏细胞(first row Deiter cell)、第二排戴特氏细胞、第三排戴特氏细胞、亨森氏细胞(Hensen's cell)、克劳氏细胞(Claudius cell)、内沟细胞、外沟细胞、螺旋突细胞、根细胞、齿间细胞、血管纹基底细胞、血管纹中间细胞、血管纹缘细胞、内毛细胞、外毛细胞、前庭毛细胞和施旺细胞(Schwann cell)。

[0077] 如本文所用,术语"野生型"是指在给定生物体中具有特定基因的最高频率的基因型。

附图说明

[0078] 图1A至图1C是一系列单平面共焦荧光图像,其比较了用在SLC6A14v2(SEQ ID NO:3;顶行)或SLC6A14v3(SEQ ID NO:1;底行)启动子控制下编码核GFP的病毒载体转导的成年小鼠椭圆囊中的核GFP表达。支持细胞核用针对SRY-Box转录因子2(Sox2)蛋白产生的抗体进行免疫标记,并且毛细胞核用针对POU4类同源异形盒3(Pou4f3)蛋白产生的抗体进行免疫标记。图1A显示了支持细胞(SC)核层。图1B显示了毛细胞(HC)核层,并且图1C显示了间充质层。

[0079] 图2A至图2D是一系列图,其显示了在用在SLC6A14v2(SEQ ID NO:3)或SLC6A14v3(SEQ ID NO:1)启动子控制下编码核GFP的病毒载体转导的成年小鼠椭圆囊中的支持细胞中核GFP表达的定量(图2A)、支持细胞中核GFP表达的强度(图2B)、毛细胞中核GFP表达的定量(图2C),和除支持细胞之外的所有细胞的定量(图2D)。

[0080] 图3是质粒P530的图谱。

[0081] 图4是质粒P919的图谱。

[0082] 图5是质粒P990的图谱。

[0083] 图6是质粒P1071的图谱。

具体实施方式

[0084] 本文阐述了用于特异性诱导转基因在内耳的前庭支持细胞(VSC)中的表达的组合物和方法。本发明的特征在于多核苷酸,其含有溶质载体家族6成员14(SLC6A14)启动子的能够在VSC中特异性表达转基因的区域。本发明的特征也在于核酸载体,其含有可操作地连

接至编码多肽或RNA分子的多核苷酸的所述启动子。本文所述的组合物和方法可用于在VSC中表达编码蛋白质(例如治疗蛋白、报告蛋白或其他目标蛋白质)或RNA分子(例如,抑制性RNA分子)的多核苷酸,所述蛋白质向前庭毛细胞提供结构和营养支持并且可使其分化成毛细胞,并且因此可将本文所述的组合物施用于受试者(例如哺乳动物受试者,例如人类)来治疗由前庭毛细胞的功能障碍引起的病症,例如眩晕、头晕、失衡、双侧前庭病、振动幻视或平衡障碍。

[0085] 支持细胞

[0086] 前庭系统的支持细胞是驻留于内耳中的特化上皮细胞。VSC构成介导正常前庭功能所需的关键结构、发育和营养活性的一类解剖学和形态学均一的细胞。VSC位于内耳的椭圆囊、小囊和半规管中并用作前庭毛细胞的结构锚,所述前庭毛细胞是外周前庭系统的主要感觉细胞,参与有助于平衡和空间定向感的运动感觉。突触形成至前庭耳蜗神经的毛细胞上是由VSC分泌的神经营养因子来介导,由此有助于建立和维持适合的前庭功能。此外,VSC借助于其控制细胞外和细胞内钙信号传导并形成吞噬细胞多细胞结构(称为吞噬体)用作前庭毛细胞存活、死亡和吞噬细胞清除的重要介体,所述吞噬细胞多细胞结构通过去除死亡或濒死的毛细胞来维持感觉上皮的完整性。前庭毛细胞损伤和破坏前庭毛细胞功能的遗传突变与前庭功能障碍(例如平衡缺失和眩晕(例如头晕))相关。基因疗法最近已成为用于治疗前庭功能障碍的有吸引力的方法;然而,所述领域缺少使用于基因疗法中的核酸载体靶向前庭系统的支持细胞的方法。

[0087] 本发明部分地基于SLC6A14在内耳VSC中特异性表达的发现。SLC6A14是编码钠和氯依赖性神经递质转运蛋白的基因,所述钠和氯依赖性神经递质转运蛋白能够以先前尚未鉴定为在内耳中表达的钠和氯依赖性方式转运中性和带正电的氨基酸。本文所公开的SLC6A14启动子序列以VSC特异性方式诱导内耳中的基因表达。因此,本文所述的组合物和方法可用于在VSC中表达目标基因(例如在患有前庭功能障碍的受试者中,参与前庭毛细胞发育、前庭毛细胞命运特化、前庭毛细胞再生、前庭毛细胞和/或VSC增殖、前庭毛细胞神经支配或前庭毛细胞成熟的基因或已知被破坏、例如突变的基因)以治疗患有前庭功能障碍(例如眩晕、头晕、失衡、双侧前庭病(例如双侧前庭功能低下)、振动幻视或平衡障碍)或具有患上所述疾病的风险的受试者。细胞类型特异性基因表达可通过降低与脱靶表达相关的毒性而改善基因疗法的安全性和功效。

[0088] 本文所述的组合物和方法包括SEQ ID NO:1的SLC6A14启动子,其能够在VSC或其变体(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的核酸序列)中特异性表达转基因。

[0089] 前述核酸序列提供于下表2中。

[0090] 表2.SLC6A14启动子序列

[0091]

SEQ ID NO:	启动子序列的描述	启动子序列
1	SLC6A14v3	ATACACTTATGTATATGTGCGATGTCAGTG

[0092]

SEQ ID NO:	启动子序列的描述	启动子序列
	(鼠类启动子)	TGTGTGCATATAAAGTCCCAAACAAGCCT GTATGATATTGACCAACAAGGTCAAGGCA AAGTTTTGATACTTTCAGGTCACAACCTCT CCGCATCCCTCTCTACTTTGCTCTATCTGC CTGAACTCCTGAGGACATGTTTCTACTGC AAATGGAAAATCCTTGTCAGCCAGTGAGG AACAAAGGGACTATACATAGATGAAAACT TGGCTCTCTGCTGGTTCCTTTGTTTGTATG AATTTATACAATTTGGTAAAACTGCCACCA TGTCTTACATGGACAGATTGAGTGTAGATT CTTTGAATTTTTGATGAAGAGGCGCTGCA CTGGTGATCGGAATTGCAGTCTTTCCTCTG TAGGTAACCTGGCTTGTTTCCTTACAGTTT ACTTTCTAGGCCTCGCCTTTCTCACAGAG TGAAGTCCTTTGTTAAGGTTCGAATTTCCC ATAAACCTGCTCAATAATTTGTTTGTGTTT GGCTTCTTTGAAATACTACACAAAGCAAT CCTTGTAAAAGGCAAAACTATTCCGAAGG CTGAGAAAGGAGCTCCAGGACATAGATTC AAAGTCGCTCTTTTCAGGTAGAGACAGCT GGGTAATCTTATCTTAACTGGCTACATTTC AAGGTTCCCAATTCAGGGGCTTTCCCCTC TGGGAGCAGCATTCTCTCCGGGTGATGAA GAGCTTTCTAGTGAGGAGCAAAACTTTCA GAAAACCGGAGGGCCCAGAGCAGTCTGG TCTGTTCACAAAAATTATAGCAAACAAAA TAAGCCCGGCGGATTGGGTCTCTCCTACC TCCAGCACCAGGGGAGATCAGCACTTGG CCCCAGGACAGAGACCTGAGAAGTGAGG TTTGGAAGAAGCCAGGAATCCAGGAAAG GAGGCAAGATTGCTAAGGCACCGGCACA GCTCTGAGTCAAAAGTTGTCAGTCTTCTT TGGCTCTGGCTGCGGAGCTCAATTGCTCA CAGCCCTGCCCTTTCCTAGGGCTGGGGCA AGGAATTGCTACATTCAGGATTACCTGGG GGAAAAACCAGAGGCTTGCTTTGGTCCCT TCCGGTAATTGAAAGGACTGGCCGTCAGC GAGGGGGAGGAGAGAGCTTCCCTCCATA AATGGTCCCACCCCTGGGCAAGGTGGCTC ACTTTGGCAGGTAGCAACCGGGGAGTGT

[0093]

SEQ ID NO:	启动子序列的描述	启动子序列
		GCACCTGCCACCAGTCAAGCTCAGCCAG ACTGTGAGAAGAGGAGAGGCGAGGCACA CCAAGGGATCCAGTGAACCaacGACAGATT GAAGTGCCCGAACTTCTTCAAGTGCAGAC AGAAGGAGGTAGGGTTCTGGAAGTTTCTG GTGGTGTAGGGGAGTCAGGAAGGGAAAA CAAGGAGGGAGAGTGAGTCTTAGTTTTTT GCTTTCTGTAGCTGTTCCTTATTTTGCATAT TTCTTTCTCTTCAACTCTTTTCAAGTATGC CTGATACGTTGTTCTCACGAAGTTGACGT GAAAAACAACTTTCCTGCTGGTAGTTAGG AAACTTAGGAGCACCTCAACCTGTACCTT GAGAACACCCAGAGAATGCTGCTCTTTGT CGTTCTCTATACCGTGTTCATATGCTGCAG GGAAATGCAAAGAATGTACTGTCCTTATC TGACCCTGGGAGCATTCCATAGTCAAGCA GCAGCTATCAGGTTGGGAAAGAGCTCCTC TCCAAGGTGTAACAGAAAAGGAAAATGT TGATATTTTTCTTGTTTAGAAAGTGACAGC TTCATCCGAGAAC
3	SLC6A14v2(人类启动子)	AAGCTGGGATGTTTCCTCATAGTTTACTTT CTAGGCCTCATCTTTCTTACAGAGTGTGCT CCTTTGTTAAGGTTAGAATTTCCCATAAAC CTGCTCAATAATTTGTTTGTGTTTGGCTTC TTTGAAATACTACACAAAGCAATCCCTGT AAAAGGCAAAGCTGTCCTGAAGGCTGAG AAAGGAGCCTGAGACATAGGCTCCAAGTT GCTCTTTTCAGGCAGAGCCAGCTGGGTAA TCTTATCTCAGATGGCTGCTTTTCAAGGTG CCCAATTCAGGGGCTTTTCCTCTGGGAGC AGCATTTGCCCCAGGGAATCAAGTGCTTT CTAGTCAGGGGCAAAACTTTGGGAAATCT GAGGACCCAGGGTGGTATGGTCTGTTCAG GAGAATTTTGGGGAACAGAATGGCCCCCT TCTCCCTCCAGCACTTGTACAGATCAGCA CTTGGCCCCAGAACAGAGACCAGACTGA GAGGCGAGGTTAGGAGGAAACAGGGGAC CCAGGAAAGGCGGCTAGATTGCAAACGT ACCTACACAGCTCTGAGTCAAAGGCTGTC AGTCATCTCGGCTCAGACTGCTCTGCTCT

[0094]

SEQ ID NO:	启动子序列的描述	启动子序列
		CCAGCAGCCCAGCCCTTTCCCAGGGCTGG GGCAGGAGATTGCTACATGTAGGCTTATCT GGGGAAAAACCAGAGCCTCACTTTAGTCC CTTCCGGTAATTGACACTACTGGACACCC AGGAGGGGGAGGAGAGAGCTTCTCTTCAT AAATGTTCCCACCCCTGGGCAAGGTGGCT CACTCTGGCAGGTAGGAACAGGGGAGAG TGCACCTGCTACCAGTCAAGCTCAGCCAG ACTGCAAGAGGAGGCGAGGCG

[0095] 外源核酸在哺乳动物细胞中的表达

[0096] 本文所述的组合物和方法可用于通过施用含有可操作地连接至编码目标蛋白质的核酸序列的SLC6A14启动子(例如与SEQ ID NO:1具有至少85％序列同一性(例如85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更大的序列同一性)的多核苷酸)的核酸载体,诱导或增加由目标基因(例如参与前庭功能障碍的基因的野生型形式,或参与前庭毛细胞发育、前庭毛细胞命运特化、前庭毛细胞再生、前庭毛细胞和/或VSC增殖、前庭毛细胞神经支配或前庭毛细胞成熟的基因)编码的蛋白质在VSC中的表达。已建立将蛋白质递送至哺乳动物细胞和在哺乳动物细胞中稳定表达编码蛋白质的基因的众多种方法。可与本文所述的组合物结合表达(例如当编码蛋白质的转基因可操作地连接至SLC6A14启动子(例如与SEQ ID NO:1具有至少85％序列同一性(例如85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更大的序列同一性)的多核苷酸)的蛋白质是在健康VSC中表达的蛋白质(例如在前庭毛细胞发育、前庭毛细胞命运特化、前庭毛细胞再生、前庭毛细胞和/或VSC增殖、前庭毛细胞神经支配或前庭毛细胞成熟中起作用的蛋白质,或在患有前庭功能障碍的受试者中缺失的蛋白质)或其他目标蛋白质。可使用本文所述的组合物和方法在VSC中表达的蛋白质包括婆罗双树样转录因子2(Sall2)、钙调蛋白结合转录激活因子1(Camta1)、具有YRPW基序2的Hes相关家族BHLH转录因子(Hey2)、Gata结合蛋白2(Gata2)、具有YRPW基序1的Hes相关家族BHLH转录因子(Hey1)、神经酰胺合酶2(Lass2)、SRY盒10(Sox10)、GATA结合蛋白3(Gata3)、Cut样同源异形盒1(Cux1)、核受体亚家族2族F成员(Nr2f1)、Hes相关家族BHLH转录因子(Hes1)、RAR相关孤儿受体B(Rorb)、Jun原癌基因AP-1转录因子亚基(Jun)、锌指蛋白667(Zfp667)、LIM同源异形盒3(Lhx3)、无义螺旋-环-螺旋1(Nhlh1)、MAX二聚化蛋白4(Mxd4)、MIZ-1型锌指(Zmiz1)、髓磷脂转录因子1(Myt1)、信号转导子和转录激活因子3(Stat3)、BarH样同源异形盒1(Barhl1)、胸腺细胞选择相关的高迁移率族盒(Tox)、Prospero同源异形盒1(Prox1)、核因子I A(Nfia)、甲状腺激素受体β(Thrb)、MYCL原癌基因BHLH转录因子(Mycl1)、赖氨酸去甲基酶5A(Kdm5a)、CAMP反应元件结合蛋白3样4(Creb3l4)、ETS变体1(Etv1)、父系表达的3(Peg3)、BTB结构域和CNC同源物2(Bach2)、ISLLIM同源异形盒1(Isl1)、锌指和含BTB结构域的38(Zbtb38)、肢芽和心脏发育(Lbh)、Tubby二分转录因子(Tub)、泛素C(Hmg20)、RE1沉默转录因子(Rest)、锌指蛋白827(Zfp827)、AF4/FMR2家族成员3(Aff3)、PBX/有节的1同源异

形盒2(Pknox2)、富含AT的相互作用结构域3B(Arid3b)、MLX相互作用蛋白(Mlxip)、锌指蛋白(Zfp532)、IKAROS家族锌指2(Ikzf2)、婆罗双树样转录因子1(Sall1)、SIX同源异形盒2(Six2)、婆罗双树样转录因子3(Sall3)、Lin-28同源物B(Lin28b)、调节因子X7(Rfx7)、脑源性神经营养因子(Bdnf)、生长因子非依赖性1转录阻遏子(Gfi1)、POU4类同源异形盒3(Pou4f3)、MYC原癌基因BHLH转录因子(Myc)、β-连环蛋白(Ctnnb1)、SRY盒2(Sox2)、SRY盒4(Sox4)、SRY盒11(Sox11)、TEA结构域转录因子2(Tead2)、无调BHLH转录因子1(Atoh1)和在氨基酸328、331和/或334处含有取代(例如S328A、S331A、S334A、S328A/S331A、S328A/S334A、S331A/S334A和S328A/S331A/S334)的Atoh1变体。本文所述的多核苷酸(例如SLC6A14启动子)也可用于在VSC中表达抑制性RNA分子(例如,短发夹RNA(shRNA)、反义寡核苷酸(ASO))、核酸酶(例如CRISPR相关蛋白9(Cas9)、转录激活因子样效应物核酸酶(TALEN)、锌指核酸酶(ZFN)或向导RNA(gRNA))或微小RNA。

[0097] 在一些实施方案中,使用本文所述的组合物和方法在VSC中表达的蛋白质是Atoh1。SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)可操作地连接至编码野生型Atoh1或其变体的多核苷酸序列,例如编码与野生型哺乳动物(例如人类或小鼠)Atoh1的氨基酸序列(例如SEQ ID NO:4或SEQ ID NO:6)具有至少85%序列同一性(例如85%、90%、95%、96%、97%、98%、99%或更大的序列同一性)的蛋白质的多核苷酸序列。示例性Atoh1氨基酸和多核苷酸序列列于下表3中。

[0098] 在一些实施方案中,编码Atoh1蛋白的多核苷酸序列编码相对于SEQ ID NO:4含有一个或多个保守氨基酸取代(例如1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多个保守氨基酸取代)的氨基酸序列,条件是编码的Atoh1类似物保留野生型Atoh1的治疗功能(例如促进毛细胞发育的能力)。Atoh1蛋白中不超过10%的氨基酸可用保守氨基酸取代来代替。在一些实施方案中,编码Atoh1的多核苷酸序列是因遗传密码的冗余性而编码SEQ ID NO:4的任何多核苷酸序列。编码Atoh1的多核苷酸序列可经部分或完全密码子优化用于表达(例如,在人类VSC中)。Atoh1可由具有SEQ ID NO:5的序列的多核苷酸编码。Atoh1蛋白可为人类Atoh1蛋白或者可为来自另一哺乳动物物种(例如小鼠、大鼠、牛、马、山羊、绵羊、驴、猫、犬、兔、天竺鼠或其他哺乳动物)的人类Atoh1蛋白的同源物。

[0099] 表3.Atoh1序列

[0100]

SEQ ID NO:	启动子序列的描述	序列
4	人类Atoh1氨基酸序列，RefSeq登录号NP_005163.1	MSRLLHAEEWAEVKELGDHHRQPQPHHLPQ PPPPPQPPATLQAREHPVYPPELSLLDSTDPR AWLAPTLQGICTARAAQYLLHSPELGASEAA APRDEVDGRGELVRRSSGGASSSKSPGPVKV REQLCKLKGGVVVDELGCSRQRAPSSKQVN GVQKQRRLAANARERRRMHGLNHAFDQLR NVIPSFNNDKKLSKYETLQMAQIYINALSEL LQTPSGGEQPPPPPASCKSDHHHLRTAASYE GGAGNATAAGAQQASGGSQRPTPPGSCRTR FSAPASAGGYSVQLDALHFSTFEDSALTAMM AQKNLSPSLPGSILQPVQEENSKTSPRSHRSD GEFSPHSHYSDSDEAS

[0101]

SEQ ID NO:	启动子序列的描述	序列
5	人类 ATOH1 蛋白编码序列，也记载于 RefSeq 登录号 NM_005172.2 下	ATGTCCCGCCTGCTGCATGCAGAAGAGTGG GCTGAAGTGAAGGAGTTGGGAGACCACCA TCGCCAGCCCCAGCCGCATCATCTCCCGCA ACCGCCGCCGCCGCCGCAGCCACCTGCAA CTTTGCAGGCGAGAGAGCATCCCGTCTACC CGCCTGAGCTGTCCCTCCTGGACAGCACC GACCCACGCGCCTGGCTGGCTCCCACTTTG CAGGGCATCTGCACGGCACGCGCCGCCCA GTATTTGCTACATTCCCCGGAGCTGGGTGC CTCAGAGGCCGCTGCGCCCCGGGACGAGG TGGACGGCCGGGGGGAGCTGGTAAGGAGG AGCAGCGGCGGTGCCAGCAGCAGCAAGA GCCCCGGGCCGGTGAAAGTGCGGGAACAG CTGTGCAAGCTGAAAGGCGGGGTGGTGGT AGACGAGCTGGGCTGCAGCCGCCAACGGG CCCCTTCCAGCAAACAGGTGAATGGGGTG CAGAAGCAGAGACGGCTAGCAGCCAACGC CAGGGAGCGGCGCAGGATGCATGGGCTGA ACCACGCCTTCGACCAGCTGCGCAATGTTA TCCCGTCGTTCAACAACGACAAGAAGCTG TCCAAATATGAGACCCTGCAGATGGCCCAA ATCTACATCAACGCCTTGTCCGAGCTGCTA CAAACGCCCAGCGGAGGGGAACAGCCACC GCCGCCTCCAGCCTCCTGCAAAAGCGACC ACCACCACCTTCGCACCGCGGCCTCCTATG AAGGGGGCGCGGGCAACGCGACCGCAGCT GGGGCTCAGCAGGCTTCCGGAGGGAGCCA GCGGCCGACCCCGCCCGGGAGTTGCCGGA CTCGCTTCTCAGCCCCAGCTTCTGCGGGAG GGTACTCGGTGCAGCTGGACGCTCTGCACT TCTCGACTTTCGAGGACAGCGCCCTGACA GCGATGATGGCGCAAAAGAATTTGTCTCCT TCTCTCCCCGGGAGCATCTTGCAGCCAGTG CAGGAGGAAAACAGCAAAACTTCGCCTCG GTCCCACAGAAGCGACGGGGAATTTTCCC CCCATTCCCATTACAGTGACTCGGATGAGG CAAGT

[0102]

SEQ ID NO:	启动子序列的描述	序列
6	鼠类Atoh1氨基酸序列，UniProt P48985	MSRLLHAEEWAEVKELGDHHRHPQPHHVPP LTPQPPATLQARDLPVYPAELSLLDSTDPRAW LTPTLQGLCTARAAQYLLHSPELGASEAAAP RDEADSQGELVRRSGCGGLSKSPGPVKVRE QLCKLKGGVVVDELGCSRQRAPSSKQVNGV QKQRRLAANARERRRMHGLNHAFDQLRNV IPSFNNDKKLSKYETLQMAQIYINALSELLQT PNVGEQPPPPTASCKNDHHHLRTASSYEGGA GASAVAGAQPAPGGGPRPTPPGPCRTRFSGPA SSGGYSVQLDALHFPAFEDRALTAMMAQKD LSPSLPGGILQPVQEDNSKTSPRSHRSDGEFS PHSHYSDSDEAS
7	鼠类ATOH1蛋白编码序列，也记载于RefSeq登录号NM_007500.5下	ATGTCCCGCCTGCTGCATGCAGAAGAGTGG GCTGAGGTAAAAGAGTTGGGGGACCACCA TCGCCATCCCCAGCCGCACCACGTCCCGCC GCTGACGCCACAGCCACCTGCTACCCTGC AGGCGAGAGACCTTCCCGTCTACCCGGCA GAACTGTCCCTCCTGGATAGCACCGACCCA CGCGCCTGGCTGACTCCCACTTTGCAGGGC CTCTGCACGGCACGCGCCGCCCAGTATCTG CTGCATTCTCCCGAGCTGGGTGCCTCCGAG GCCGCGGCGCCCCGGGACGAGGCTGACAG CCAGGGTGAGCTGGTAAGGAGAAGCGGCT GTGGCGGCCTCAGCAAGAGCCCCGGGCCC GTCAAAGTACGGGAACAGCTGTGCAAGCT GAAGGGTGGGGTTGTAGTGGACGAGCTTG GCTGCAGCCGCCAGCGAGCCCCTTCCAGC AAACAGGTGAATGGGGTACAGAAGCAAAG GAGGCTGGCAGCAAACGCAAGGGAACGG CGCAGGATGCACGGGCTGAACCACGCCTT CGACCAGCTGCGCAACGTTATCCCGTCCTT CAACAACGACAAGAAGCTGTCCAAATATG AGACCCTACAGATGGCCCAGATCTACATCA ACGCTCTGTCGGAGTTGCTGCAGACTCCCA ATGTCGGAGAGCAACCGCCGCCGCCCACA GCTTCCTGCAAAAATGACCACCATCACCTT CGCACCGCCTCCTCCTATGAAGGAGGTGCG GGCGCCTCTGCGGTAGCTGGGGCTCAGCC AGCCCCGGGAGGGGGCCCGAGACCTACCC CGCCCGGGCCTTGCCGGACTCGCTTCTCAG GCCCAGCTTCCTCTGGGGGTTACTCGGTGC AGCTGGACGCTTTGCACTTCCCAGCCTTCG AGGACAGGGCCCTAACAGCGATGATGGCA CAGAAGGACCTGTCGCCTTCGCTGCCCGG GGGCATCCTGCAGCCTGTACAGGAGGACA ACAGCAAAACATCTCCCAGATCCCACAGA AGTGACGGAGAGTTTTCCCCCCACTCTCAT TACAGTGACTCTGATGAGGCCAGT

[0103] 编码目标蛋白质的多核苷酸

[0104] 可用于在哺乳动物细胞中实现目标蛋白质的治疗有效的细胞内浓度的一个平台

是经由稳定表达编码目标蛋白质的基因(例如,通过整合至哺乳动物细胞的核或线粒体基因组中,或通过在哺乳动物细胞核中形成游离型多联体)。所述基因是编码相应蛋白质的一级氨基酸序列的多核苷酸。为将外源基因引入哺乳动物细胞中,可将基因并入载体中。可通过多种方法将载体引入细胞中,包括转化、转染、转导、直接摄取、射弹轰击和通过将载体囊封于脂质体中。转染或转化细胞的合适方法的实例包括磷酸钙沉淀、电穿孔、显微注射、感染、脂转染和直接摄取。此类方法更详细阐述于例如Green等人,Molecular Cloning:A Laboratory Manual,第四版(Cold Spring Harbor University Press,NewYork 2014);和Ausubel等人,CurrentProtocols inMolecular Biology(JohnWiley&Sons,NewYork2015),所述文献各自的公开内容通过引用并入本文。

[0105] 也可通过使含有编码目标蛋白质的基因的载体靶向细胞膜磷脂将目标蛋白质引入哺乳动物细胞中。例如,可通过将载体分子连接至VSV-G蛋白使载体靶向细胞膜的细胞外表面上的磷脂,所述VSV-G蛋白是对所有细胞膜磷脂具有亲和力的病毒蛋白。这种构建体可使用本领域技术人员所熟知的方法产生。

[0106] 哺乳动物RNA聚合酶对编码目标蛋白质的多核苷酸的识别和结合对于基因表达至关重要。因此,可在多核苷酸内包括对招募RNA聚合酶并促进转录复合物在转录起始位点组装的转录因子表现出高亲和力的序列元件。此类序列元件包括例如哺乳动物启动子,其序列可由特异性转录起始因子和最终RNA聚合酶识别和结合。哺乳动物启动子的实例已阐述于Smith等人,Mol.Sys.Biol.,3:73,在线出版中,其公开内容通过引用并入本文。用于本文所述方法和组合物中的启动子是SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)。

[0107] 一旦将编码目标蛋白质的多核苷酸并入哺乳动物细胞的核DNA中,就可立即通过本领域中已知的方法来诱导这种多核苷酸的转录。例如,可通过使哺乳动物细胞暴露于外部化学试剂(例如调节转录因子和/或RNA聚合酶与哺乳动物启动子的结合并因此调节基因表达的剂)来诱导表达。化学试剂可用于促进RNA聚合酶和/或转录因子与哺乳动物启动子的结合,例如通过去除已结合启动子的抑制蛋白。或者,化学试剂可用于增强哺乳动物启动子对RNA聚合酶和/或转录因子的亲和力,使得位于启动子下游的基因的转录速率在化学试剂存在下增加。加强通过上述机制的多核苷酸转录的化学试剂的实例包括四环素和强力霉素。这些试剂可购得(Life Technologies,Carlsbad,CA)并且可根据所建立的方案施用于哺乳动物细胞来促进基因表达。

[0108] 可包括在用于本文所述组合物和方法中的多核苷酸中的其他DNA序列元件包括增强子序列。增强子代表另一类调节元件,其诱导含有目标基因的多核苷酸的构象变化,使得DNA采取有利于转录因子和RNA聚合酶在转录起始位点结合的三维取向。因此,用于本文所述组合物和方法中的多核苷酸包括编码目标蛋白质并且另外包括哺乳动物增强子序列的那些多核苷酸。现在从哺乳动物基因已知许多增强子序列,并且实例包括来自编码哺乳动物珠蛋白、弹性蛋白酶、白蛋白、α-胎儿蛋白和胰岛素的基因的增强子。用于本文所述组合物和方法中的增强子也包括衍生自能够感染真核细胞的病毒的遗传材料的那些增强子。实例包括在复制起点后侧(bp 100-270)的SV40增强子、巨细胞病毒早期启动子增强子、在复制起点后侧的多瘤增强子和腺病毒增强子。诱导真核基因转录活化的其他增强子序列包括

CMV增强子和RSV增强子。增强子可剪接至含有编码目标蛋白质的多核苷酸的载体中,例如在此基因的位置5'或3'处。在一个优选取向中,增强子定位于启动子的5'侧,所述启动子进而位于相对于编码目标蛋白质的多核苷酸的5'。

[0109] 本文所述含有SLC6A14启动子的核酸载体可包括WPRE。WPRE在mRNA水平上通过促进转录物的核输出和/或通过增加新生转录物的多聚腺苷酸化效率、由此增加细胞中mRNA的总量起作用。将WPRE添加至载体可在体外和体内实质上改进来自若干不同启动子的转基因表达的水平。在本文所述的组合物和方法的一些实施方案中,WPRE具有以下序列:

[0110] GATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGA(SEQ ID NO:8)。

[0111] 在其他实施方案中,WPRE具有以下序列:

[0112] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATCTAGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAA(SEQ ID NO:9)

[0113] 在一些实施方案中,本文所述含有SLC6A14启动子的核酸载体包括报告基因序列,其可用于验证可操作地连接至VSC中的SLC6A14启动子的基因的表达或者其可用于确定或证实启动子的前庭支持细胞特异性。可提供于转基因中的报告基因序列包括编码β-内酰胺酶、β-半乳糖苷酶(LacZ)、碱性磷酸酶、胸苷激酶、绿色荧光蛋白(GFP)、氯霉素乙酰基转移酶(CAT)、荧光素酶和本领域中所熟知的其他蛋白质的DNA序列。当与驱动其表达的调节元件(例如SLC6A14启动子)缔合时,报告基因序列提供可通过常规方法检测到的信号,所述常规方法包括酶促、放射照相、比色、荧光或其他光谱测定、荧光激活细胞分选测定和免疫测定,包括酶联免疫吸附测定(ELISA)、放射免疫测定(RIA)和免疫组织化学。例如,当标志物序列是LacZ基因时,通过测定β-半乳糖苷酶活性检测到携带信号的载体的存在。当转基因是绿色荧光蛋白或荧光素酶时,携带信号的载体可在视觉上通过光度计中的色彩或光产生来测量。

[0114] 可用于产生用于本文所述组合物和方法中的核酸载体(例如AAV载体)的转移质粒提供于表4中。转移质粒(例如含有有待通过核酸载体递送、例如有待通过AAV递送的DNA序列的质粒)可与辅助质粒(例如提供AAV制造所需的蛋白质的质粒)和rep/cap质粒(例如提供AAV衣壳蛋白和将转移质粒DNA序列插入衣壳外壳中的蛋白质的质粒)共递送至产生细胞中以产生用于施用的核酸载体(例如AAV载体)。可使用表4中所提供的转移质粒来产生含有可操作地连接至转基因(例如编码Atoh1的多核苷酸(鼠类(SEQ ID NO:11)或人类(SEQ ID

NO:10)Atoh1)或编码GFP的多核苷酸(SEQ ID NO:2))的SLC6A14启动子的核酸载体(例如AAV载体)。

[0115] 表4.转移质粒

[0116]

SEQ ID NO.	描述	质粒序列
2	质粒 P919 位置 1-130 处的 5' ITR 位置 219-1977 处的 SLC6A14v3 启动子 位置 1992-2369 处的 H2B 序列 位置 2388-3107 处的 eGFP 序列 位置 3116-3663 处的 WPRE 位置 3748-3955 处的 polyA 位置 4043-4172 处的 3' ITR	CTGCGCGCTCGCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCGGGCGACCTTT GGTCGCCCGGCCTCAGTGAGCGAGCGAGCG CGCAGAGAGGGAGTGGCCAACTCCATCACT AGGGGTTCCTTGTAGTTAATGATTAACCCG CCATGCTACTTATCTACGTAGCCATGCTCTA GGAAGATCGGAATTCGCCCTTAAGCTAGCG GCGCGCCATACACTTATGTATATGTGCGAT GTCAGTGTGTGTGCATATAAAGTCCCAAAC AAGCCTGTATGATATTGACCAACAAGGTCA AGGCAAAGTTTTGATACTTTCAGGTCACAA CCTCTCCGCATCCCTCTCTACTTTGCTCTAT CTGCCTGAACTCCTGAGGACATGTTTCTACT GCAAATGGAAAATCCTTGTCAGCCAGTGAG GAACAAAGGGACTATACATAGATGAAAACT TGGCTCTCTGCTGGTTCCTTTGTTTGTATGA ATTTATACAATTTGGTAAAACTGCCACCAT GTCTTACATGGACAGATTGAGTGTAGATTC TTTGAATTTTTGATGAAGAGGCGCTGCACT GGTGATCGGAATTGCAGTCTTTCCTCTGTAG GTAACCTGGCTTGTTTCCTTACAGTTTACTT TCTAGGCCTCGCCTTTCTCACAGAGTGAAGT CCTTTGTTAAGGTTCGAATTTCCCATAAACC TGCTCAATAATTTGTTTGTGTTTGGCTTCTTT GAAATACTACACAAAGCAATCCTTGTAAAA GGCAAAACTATTCCGAAGGCTGAGAAAGG AGCTCCAGGACATAGATTCAAAGTCGCTCT TTTCAGGTAGAGACAGCTGGGTAATCTTAT CTTAACTGGCTACATTTCAAGGTTCCCAATT CAGGGGCTTTCCCCTCTGGGAGCAGCATTC TCTCCGGGTGATGAAGAGCTTTCTAGTGAG GAGCAAAACTTTCAGAAAACCGGAGGGCCC AGAGCAGTCTGGTCTGTTCACAAAAATTAT AGCAAACAAAATAAGCCCGGCGGATTGGGT CTCTCCTACCTCCAGCACCAGGGGAGATCA GCACTTGGCCCCAGGACAGAGACCTGAGAA GTGAGGTTTGGAAGAAGCCAGGAATCCAGG AAAGGAGGCAAGATTGCTAAGGCACCGGC ACAGCTCTGAGTCAAAAGTTGTCAGTCTTCT TTGGCTCTGGCTGCGGAGCTCAATTGCTCAC AGCCCTGCCCTTTCCTAGGGCTGGGGCAAG GAATTGCTACATTCAGGATTACCTGGGGGA AAAACCAGAGGCTTGCTTTGGTCCCTTCCG GTAATTGAAAGGACTGGCCGTCAGCGAGGG GGAGGAGAGAGCTTCCCTCCATAAATGGTC CCACCCCTGGGCAAGGTGGCTCACTTTGGC AGGTAGCAACCGGGGAGTGTGCACCTGCCA CCAGTCAAGCTCAGCCAGACTGTGAGAAGA GGAGAGGCGAGGCACACCAAGGGATCCAG TGAACCAACGACAGATTGAAGTGCCCGAAC TTCTTCAAGTGCAGACAGAAGGAGGTAGGG TTCTGGAAGTTTCTGGTGGTGTAGGGGAGT

[0117]

SEQ ID NO.	描述	质粒序列
		CAGGAAGGGAAAACAAGGAGGGAGAGTGA GTCTTAGTTTTTTGCTTTCTGTAGCTGTTCCT TATTTTGCATATTTCTTTCTCTTCAACTCTTT TCAAGTATGCCTGATACGTTGTTCTCACGAA GTTGACGTGAAAAACAACTTTCCTGCTGGT AGTTAGGAAACTTAGGAGCACCTCAACCTG TACCTTGAGAACACCCAGAGAATGCTGCTC TTTGTCGTTCTCTATACCGTGTTCATATGCT GCAGGGAAATGCAAAGAATGTACTGTCCTT ATCTGACCCTGGGAGCATTCCATAGTCAAG CAGCAGCTATCAGGTTGGGAAAGAGCTCCT CTCCAAGGTGTAACAGAAAAGGAAAATGTT GATATTTTTCTTGTTTAGAAAGTGACAGCTT CATCCGAGAACGCGGCCGCGCCACCATGCC AGAGCCAGCGAAGTCTGCTCCCGCCCCGAA AAAGGGCTCCAAGAAGGCGGTGACTAAGG CGCAGAAGAAAGGCGGCAAGAAGCGCAAG CGCAGCCGCAAGGAGAGCTATTCCATCTAT GTGTACAAGGTTCTGAAGCAGGTCCACCCT GACACCGGCATTTCGTCCAAGGCCATGGGC ATCATGAATTCGTTTGTGAACGACATTTTCG AGCGCATCGCAGGTGAGGCTTCCCGCCTGG CGCATTACAACAAGCGCTCGACCATCACCT CCAGGGAGATCCAGACGGCCGTGCGCCTGC TGCTGCCTGGGGAGTTGGCCAAGCACGCCG TGTCCGAGGGTACTAAGGCCATCACCAAGT ACACCAGCGCTAAGGATCCACCGGTCGCCA CCATGGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAGCTGG ACGGCGACGTAAACGGCCACAAGTTCAGCG TGTCCGGCGAGGGCGAGGGCGATGCCACCT ACGGCAAGCTGACCCTGAAGTTCATCTGCA CCACCGGCAAGCTGCCCGTGCCCTGGCCCA CCCTCGTGACCACCCTGACCTACGGCGTGC AGTGCTTCAGCCGCTACCCCGACCACATGA AGCAGCACGACTTCTTCAAGTCCGCCATGC CCGAAGGCTACGTCCAGGAGCGCACCATCT TCTTCAAGGACGACGGCAACTACAAGACCC GCGCCGAGGTGAAGTTCGAGGGCGACACCC TGGTGAACCGCATCGAGCTGAAGGGCATCG ACTTCAAGGAGGACGGCAACATCCTGGGGC ACAAGCTGGAGTACAACTACAACAGCCACA ACGTCTATATCATGGCCGACAAGCAGAAGA ACGGCATCAAGGTGAACTTCAAGATCCGCC ACAACATCGAGGACGGCAGCGTGCAGCTCG CCGACCACTACCAGCAGAACACCCCCATCG GCGACGGCCCCGTGCTGCTGCCCGACAACC ACTACCTGAGCACCCAGTCCGCCCTGAGCA AAGACCCCAACGAGAAGCGCGATCACATG GTCCTGCTGGAGTTCGTGACCGCCGCCGGG ATCACTCTCGGCATGGACGAGCTGTACAAG TAATAAGCTTGGATCCAATCAACCTCTGGA

[0118]

SEQ ID NO.	描述	质粒序列
		TTACAAAATTTGTGAAAGATTGACTGGTAT TCTTAACTATGTTGCTCCTTTTACGCTATGT GGATACGCTGCTTTAATGCCTTTGTATCATG CTATTGCTTCCCGTATGGCTTTCATTTTCTCC TCCTTGTATAAATCCTGGTTGCTGTCTCTTT ATGAGGAGTTGTGGCCCGTTGTCAGGCAAC GTGGCGTGGTGTGCACTGTGTTTGCTGACG CAACCCCCACTGGTTGGGGCATTGCCACCA CCTGTCAGCTCCTTTCCGGGACTTTCGCTTT CCCCCTCCCTATTGCCACGGCGGAACTCATC GCCGCCTGCCTTGCCCGCTGCTGGACAGGG GCTCGGCTGTTGGGCACTGACAATTCCGTG GTGTTGTCGGGGAAATCATCGTCCTTTCCTT GGCTGCTCGCCTGTGTTGCCACCTGGATTCT GCGCGGGACGTCCTTCTGCTACGTCCCTTCG GCCCTCAATCCAGCGGACCTTCCTTCCCGCG GCCTGCTGCCGGCTCTGCGGCCTCTTCCGCG TCTTCGAACAATTGCATCGGACACATCTTG GCGTTTTACAACGTCGTGACTGGGAAAACC CTGGCGTTACCCAACTTAAGATCTGCCTCG ACTGTGCCTTCTAGTTGCCAGCCATCTGTTG TTTGCCCCTCCCCCGTGCCTTCCTTGACCCT GGAAGGTGCCACTCCCACTGTCCTTTCCTAA TAAAATGAGGAAATTGCATCGCATTGTCTG AGTAGGTGTCATTCTATTCTGGGGGGTGGG GTGGGGCAGGACAGCAAGGGGGAGGATTG GGAAGACAATAGCAGGCATGCTGGGGACTC GAGTTAAGGGCGAATTCCCGATAAGGATCT TCCTAGAGCATGGCTACGTAGATAAGTAGC ATGGCGGGTTAATCATTAACTACAAGGAAC CCCTAGTGATGGAGTTGGCCACTCCCTCTCT GCGCGCTCGCTCGCTCACTGAGGCCGGGCG ACCAAAGGTCGCCCGACGCCCGGGCTTTGC CCGGGCGGCCTCAGTGAGCGAGCGAGCGCG CAGCCTTAATTAACCTAATTCACTGGCCGTC GTTTTACAACGTCGTGACTGGGAAAACCCT GGCGTTACCCAACTTAATCGCCTTGCAGCA CATCCCCCTTTCGCCAGCTGGCGTAATAGC GAAGAGGCCCGCACCGATCGCCCTTCCCAA CAGTTGCGCAGCCTGAATGGCGAATGGGAC GCGCCCTGTAGCGGCGCATTAAGCGCGGCG GGTGTGGTGGTTACGCGCAGCGTGACCGCT ACACTTGCCAGCGCCCTAGCGCCCGCTCCTT TCGCTTTCTTCCCTTCCTTTCTCGCCACGTTC GCCGGCTTTCCCCGTCAAGCTCTAAATCGG GGGCTCCCTTTAGGGTTCCGATTTAGTGCTT TACGGCACCTCGACCCCAAAAAACTTGATT AGGGTGATGGTTCACGTAGTGGGCCATCGC CCTGATAGACGGTTTTTCGCCCTTTGACGTT GGAGTCCACGTTCTTTAATAGTGGACTCTTG TTCCAAACTGGAACAACACTCAACCCTATC TCGGTCTATTCTTTTGATTTATAAGGGATTT

[0119]

SEQ ID NO.	描述	质粒序列
		TGCCGATTTCGGCCTATTGGTTAAAAAATG AGCTGATTTAACAAAAATTTAACGCGAATT TTAACAAAATATTAACGCTTACAATTTAGG TGGCACTTTTCGGGGAAATGTGCGCGGAAC CCCTATTTGTTTATTTTTCTAAATACATTCA AATATGTATCCGCTCATGAGACAATAACCC TGATAAATGCTTCAATAATATTGAAAAAGG AAGAGTATGAGTATTCAACATTTCCGTGTC GCCCTTATTCCCTTTTTTGCGGCATTTTGCCT TCCTGTTTTTGCTCACCCAGAAACGCTGGTG AAAGTAAAAGATGCTGAAGATCAGTTGGGT GCACGAGTGGGTTACATCGAACTGGATCTC AACAGCGGTAAGATCCTTGAGAGTTTTCGC CCCGAAGAACGTTTTCCAATGATGAGCACT TTTAAAGTTCTGCTATGTGGCGCGGTATTAT CCCGTATTGACGCCGGGCAAGAGCAACTCG GTCGCCGCATACACTATTCTCAGAATGACTT GGTTGAGTACTCACCAGTCACAGAAAAGCA TCTTACGGATGGCATGACAGTAAGAGAATT ATGCAGTGCTGCCATAACCATGAGTGATAA CACTGCGGCCAACTTACTTCTGACAACGAT CGGAGGACCGAAGGAGCTAACCGCTTTTTT GCACAACATGGGGGATCATGTAACTCGCCT TGATCGTTGGGAACCGGAGCTGAATGAAGC CATACCAAACGACGAGCGTGACACCACGAT GCCTGTAGCAATGGCAACAACGTTGCGCAA ACTATTAACTGGCGAACTACTTACTCTAGCT TCCCGGCAACAATTAATAGACTGGATGGAG GCGGATAAAGTTGCAGGACCACTTCTGCGC TCGGCCCTTCCGGCTGGCTGGTTTATTGCTG ATAAATCTGGAGCCGGTGAGCGTGGGTCTC GCGGTATCATTGCAGCACTGGGGCCAGATG GTAAGCCCTCCCGTATCGTAGTTATCTACAC GACGGGGAGTCAGGCAACTATGGATGAAC GAAATAGACAGATCGCTGAGATAGGTGCCT CACTGATTAAGCATTGGTAACTGTCAGACC AAGTTTACTCATATATACTTTAGATTGATTT AAAACTTCATTTTTAATTTAAAAGGATCTAG GTGAAGATCCTTTTTGATAATCTCATGACCA AAATCCCTTAACGTGAGTTTTCGTTCCACTG AGCGTCAGACCCCGTAGAAAAGATCAAAG GATCTTCTTGAGATCCTTTTTTTCTGCGCGT AATCTGCTGCTTGCAAACAAAAAAACCACC GCTACCAGCGGTGGTTTGTTTGCCGGATCA AGAGCTACCAACTCTTTTTCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATAC TGTTCTTCTAGTGTAGCCGTAGTTAGGCCAC CACTTCAAGAACTCTGTAGCACCGCCTACA TACCTCGCTCTGCTAATCCTGTTACCAGTGG CTGCTGCCAGTGGCGATAAGTCGTGTCTTA CCGGGTTGGACTCAAGACGATAGTTACCGG ATAAGGCGCAGCGGTCGGGCTGAACGGGG

SEQ ID NO.	描述	质粒序列
		GGTTCGTGCACACAGCCCAGCTTGGAGCGA ACGACCTACACCGAACTGAGATACCTACAG CGTGAGCTATGAGAAAGCGCCACGCTTCCC GAAGGGAGAAAGGCGGACAGGTATCCGGT AAGCGGCAGGGTCGGAACAGGAGAGCGCA CGAGGGAGCTTCCAGGGGGAAACGCCTGGT ATCTTTATAGTCCTGTCGGGTTTCGCCACCT CTGACTTGAGCGTCGATTTTTGTGATGCTCG TCAGGGGGGCGGAGCCTATGGAAAAACGC CAGCAACGCGGCCTTTTTACGGTTCCTGGCC TTTTGCTGGCCTTTTGCTCACATGTTCTTTCC TGCGTTATCCCCTGATTCTGTGGATAACCGT ATTACCGCCTTTGAGTGAGCTGATACCGCTC GCCGCAGCCGAACGACCGAGCGCAGCGAG TCAGTGAGCGAGGAAGCGGAAGAGCGCCC AATACGCAAACCGCCTCTCCCCGCGCGTTG GCCGATTCATTAATGCAGCTGGCACGACAG GTTTCCCGACTGGAAAGCGGGCAGTGAGCG CAACGCAATTAATGTGAGTTAGCTCACTCA TTAGGCACCCCAGGCTTTACACTTTATGCTT CCGGCTCGTATGTTGTGTGGAATTGTGAGC GGATAACAATTTCACACAGGAAACAGCTAT GACCATGATTACGCCAGATTTAATTAAGGC CTTAATTAGG
[0120] 10	质粒 P990 位置 1-130 处的 5' ITR 位置 219-1977 处的 SLC6A14v3 启动子 位置 1992-3053 处的人类 Atoh1 序列 位置 3064-3611 处的 WPRE 位置 3624-3831 处的 polyA 位置 3919-4048 处的 3' ITR	CTGCGCGCTCGCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCGGGCGACCTTT GGTCGCCCGGCCTCAGTGAGCGAGCGAGCG CGCAGAGAGGGAGTGGCCAACTCCATCACT AGGGGTTCCTTGTAGTTAATGATTAACCCG CCATGCTACTTATCTACGTAGCCATGCTCTA GGAAGATCGGAATTCGCCCTTAAGCTAGCG GCGCGCCATACACTTATGTATATGTGCGAT GTCAGTGTGTGTGCATATAAAGTCCCAAAC AAGCCTGTATGATATTGACCAACAAGGTCA AGGCAAAGTTTTGATACTTTCAGGTCACAA CCTCTCCGCATCCCTCTCTACTTTGCTCTAT CTGCCTGAACTCCTGAGGACATGTTTCTACT GCAAATGGAAAATCCTTGTCAGCCAGTGAG GAACAAAGGGACTATACATAGATGAAAACT TGGCTCTCTGCTGGTTCCTTTGTTTGTATGA ATTTATACAATTTGGTAAAACTGCCACCAT GTCTTACATGGACAGATTGAGTGTAGATTC TTTGAATTTTTGATGAAGAGGCGCTGCACT GGTGATCGGAATTGCAGTCTTTCCTCTGTAG GTAACCTGGCTTGTTTCCTTACAGTTTACTT TCTAGGCCTCGCCTTTCTCACAGAGTGAAGT CCTTTGTTAAGGTTCGAATTTCCCATAAACC TGCTCAATAATTTGTTTGTGTTTGGCTTCTTT GAAATACTACACAAAGCAATCCTTGTAAAA GGCAAAACTATTCCGAAGGCTGAGAAAGG AGCTCCAGGACATAGATTCAAAGTCGCTCT TTTCAGGTAGAGACAGCTGGGTAATCTTAT

[0121]

SEQ ID NO.	描述	质粒序列
		CTTAACTGGCTACATTTCAAGGTTCCCAATT CAGGGGCTTTCCCCTCTGGGAGCAGCATTC TCTCCGGGTGATGAAGAGCTTTCTAGTGAG GAGCAAAACTTTCAGAAAACCGGAGGGCCC AGAGCAGTCTGGTCTGTTCACAAAAATTAT AGCAAACAAAATAAGCCCGGCGGATTGGGT CTCTCCTACCTCCAGCACCAGGGGAGATCA GCACTTGGCCCCAGGACAGAGACCTGAGAA GTGAGGTTTGGAAGAAGCCAGGAATCCAGG AAAGGAGGCAAGATTGCTAAGGCACCGGC ACAGCTCTGAGTCAAAAGTTGTCAGTCTTCT TTGGCTCTGGCTGCGGAGCTCAATTGCTCAC AGCCCTGCCCTTTCCTAGGGCTGGGGCAAG GAATTGCTACATTCAGGATTACCTGGGGGA AAAACCAGAGGCTTGCTTTGGTCCCTTCCG GTAATTGAAAGGACTGGCCGTCAGCGAGGG GGAGGAGAGAGCTTCCCTCCATAAATGGTC CCACCCCTGGGCAAGGTGGCTCACTTTGGC AGGTAGCAACCGGGGAGTGTGCACCTGCCA CCAGTCAAGCTCAGCCAGACTGTGAGAAGA GGAGAGGCGAGGCACACCAAGGGATCCAG TGAACCAACGACAGATTGAAGTGCCCGAAC TTCTTCAAGTGCAGACAGAAGGAGGTAGGG TTCTGGAAGTTTCTGGTGGTGTAGGGGAGT CAGGAAGGGAAAACAAGGAGGGAGAGTGA GTCTTAGTTTTTTGCTTTCTGTAGCTGTTCCT TATTTTGCATATTTCTTTCTCTTCAACTCTTT TCAAGTATGCCTGATACGTTGTTCTCACGAA GTTGACGTGAAAAACAACTTTCCTGCTGGT AGTTAGGAAACTTAGGAGCACCTCAACCTG TACCTTGAGAACACCCAGAGAATGCTGCTC TTTGTCGTTCTCTATACCGTGTTCATATGCT GCAGGGAAATGCAAAGAATGTACTGTCCTT ATCTGACCCTGGGAGCATTCCATAGTCAAG CAGCAGCTATCAGGTTGGGAAAGAGCTCCT CTCCAAGGTGTAACAGAAAAGGAAAATGTT GATATTTTTCTTGTTTAGAAAGTGACAGCTT CATCCGAGAACGCGGCCGCGCCACCATGTC CCGCCTGCTGCATGCAGAAGAGTGGGCTGA AGTGAAGGAGTTGGGAGACCACCATCGCCA GCCCCAGCCGCATCATCTCCCGCAACCGCC GCCGCCGCCGCAGCCACCTGCAACTTTGCA GGCGAGAGAGCATCCCGTCTACCCGCCTGA GCTGTCCCTCCTGGACAGCACCGACCCACG CGCCTGGCTGGCTCCCACTTTGCAGGGCAT CTGCACGGCACGCGCCGCCCAGTATTTGCT ACATTCCCCGGAGCTGGGTGCCTCAGAGGC CGCTGCGCCCCGGGACGAGGTGGACGGCCG GGGGGAGCTGGTAAGGAGGAGCAGCGGCG GTGCCAGCAGCAGCAAGAGCCCCGGGCCG GTGAAAGTGCGGGAACAGCTGTGCAAGCTG AAAGGCGGGGTGGTGGTAGACGAGCTGGG

[0122]

SEQ ID NO.	描述	质粒序列
		CTGCAGCCGCCAACGGGCCCCTTCCAGCAA ACAGGTGAATGGGGTGCAGAAGCAGAGAC GGCTAGCAGCCAACGCCAGGGAGCGGCGC AGGATGCATGGGCTGAACCACGCCTTCGAC CAGCTGCGCAATGTTATCCCGTCGTTCAAC AACGACAAGAAGCTGTCCAAATATGAGACC CTGCAGATGGCCCAAATCTACATCAACGCC TTGTCCGAGCTGCTACAAACGCCCAGCGGA GGGGAACAGCCACCGCCGCCTCCAGCCTCC TGCAAAAGCGACCACCACCACCTTCGCACC GCGGCCTCCTATGAAGGGGGCGCGGGCAAC GCGACCGCAGCTGGGGCTCAGCAGGCTTCC GGAGGGAGCCAGCGGCCGACCCCGCCCGG GAGTTGCCGGACTCGCTTCTCAGCCCCAGC TTCTGCGGGAGGGTACTCGGTGCAGCTGGA CGCTCTGCACTTCTCGACTTTCGAGGACAGC GCCCTGACAGCGATGATGGCGCAAAAGAAT TTGTCTCCTTCTCTCCCCGGGAGCATCTTGC AGCCAGTGCAGGAGGAAAACAGCAAAACT TCGCCTCGGTCCCACAGAAGCGACGGGGAA TTTTCCCCCCATTCCCATTACAGTGACTCGG ATGAGGCAAGTTAGAAGCTTGGATCCAATC AACCTCTGGATTACAAAATTTGTGAAAGAT TGACTGGTATTCTTAACTATGTTGCTCCTTT TACGCTATGTGGATACGCTGCTTTAATGCCT TTGTATCATGCTATTGCTTCCCGTATGGCTT TCATTTTCTCCTCCTTGTATAAATCCTGGTT GCTGTCTCTTTATGAGGAGTTGTGGCCCGTT GTCAGGCAACGTGGCGTGGTGTGCACTGTG TTTGCTGACGCAACCCCCACTGGTTGGGGC ATTGCCACCACCTGTCAGCTCCTTTCCGGGA CTTTCGCTTTCCCCCTCCCTATTGCCACGGC GGAACTCATCGCCGCCTGCCTTGCCCGCTG CTGGACAGGGGCTCGGCTGTTGGGCACTGA CAATTCCGTGGTGTTGTCGGGGAAATCATC GTCCTTTCCTTGGCTGCTCGCCTGTGTTGCC ACCTGGATTCTGCGCGGGACGTCCTTCTGCT ACGTCCCTTCGGCCCTCAATCCAGCGGACC TTCCTTCCCGCGGCCTGCTGCCGGCTCTGCG GCCTCTTCCGCGTCTTCGAGATCTGCCTCGA CTGTGCCTTCTAGTTGCCAGCCATCTGTTGT TTGCCCCTCCCCCGTGCCTTCCTTGACCCTG GAAGGTGCCACTCCCACTGTCCTTTCCTAAT AAAATGAGGAAATTGCATCGCATTGTCTGA GTAGGTGTCATTCTATTCTGGGGGGTGGGG TGGGGCAGGACAGCAAGGGGGAGGATTGG GAAGACAATAGCAGGCATGCTGGGGACTCG AGTTAAGGGCGAATTCCCGATAAGGATCTT CCTAGAGCATGGCTACGTAGATAAGTAGCA TGGCGGGTTAATCATTAACTACAAGGAACC CCTAGTGATGGAGTTGGCCACTCCCTCTCTG CGCGCTCGCTCGCTCACTGAGGCCGGGCGA

[0123]

SEQ ID NO.	描述	质粒序列
		CCAAAGGTCGCCCGACGCCCGGGCTTTGCC CGGGCGGCCTCAGTGAGCGAGCGAGCGCGC AGCCTTAATTAACCTAATTCACTGGCCGTCG TTTTACAACGTCGTGACTGGGAAAACCCTG GCGTTACCCAACTTAATCGCCTTGCAGCAC ATCCCCCTTTCGCCAGCTGGCGTAATAGCG AAGAGGCCCGCACCGATCGCCCTTCCCAAC AGTTGCGCAGCCTGAATGGCGAATGGGACG CGCCCTGTAGCGGCGCATTAAGCGCGGCGG GTGTGGTGGTTACGCGCAGCGTGACCGCTA CACTTGCCAGCGCCCTAGCGCCCGCTCCTTT CGCTTTCTTCCCTTCCTTTCTCGCCACGTTCG CCGGCTTTCCCCGTCAAGCTCTAAATCGGG GGCTCCCTTTAGGGTTCCGATTTAGTGCTTT ACGGCACCTCGACCCCAAAAAACTTGATTA GGGTGATGGTTCACGTAGTGGGCCATCGCC CTGATAGACGGTTTTTCGCCCTTTGACGTTG GAGTCCACGTTCTTTAATAGTGGACTCTTGT TCCAAACTGGAACAACACTCAACCCTATCT CGGTCTATTCTTTTGATTTATAAGGGATTTT GCCGATTTCGGCCTATTGGTTAAAAAATGA GCTGATTTAACAAAAATTTAACGCGAATTT TAACAAAATATTAACGCTTACAATTTAGGT GGCACTTTTCGGGGAAATGTGCGCGGAACC CCTATTTGTTTATTTTTCTAAATACATTCAA ATATGTATCCGCTCATGAGACAATAACCCT GATAAATGCTTCAATAATATTGAAAAAGGA AGAGTATGAGCCATATTCAACGGGAAACGT CGAGGCCGCGATTAAATTCCAACATGGATG CTGATTTATATGGGTATAAATGGGCTCGCG ATAATGTCGGGCAATCAGGTGCGACAATCT ATCGCTTGTATGGGAAGCCCGATGCGCCAG AGTTGTTTCTGAAACATGGCAAAGGTAGCG TTGCCAATGATGTTACAGATGAGATGGTCA GACTAAACTGGCTGACGGAATTTATGCCTC TTCCGACCATCAAGCATTTTATCCGTACTCC TGATGATGCATGGTTACTCACCACTGCGAT CCCCGGAAAAACAGCATTCCAGGTATTAGA AGAATATCCTGATTCAGGTGAAAATATTGT TGATGCGCTGGCAGTGTTCCTGCGCCGGTT GCATTCGATTCCTGTTTGTAATTGTCCTTTT AACAGCGATCGCGTATTTCGTCTTGCTCAG GCGCAATCACGAATGAATAACGGTTTGGTT GATGCGAGTGATTTTGATGACGAGCGTAAT GGCTGGCCTGTTGAACAAGTCTGGAAAGAA ATGCATAAACTTTTGCCATTCTCACCGGATT CAGTCGTCACTCATGGTGATTTCTCACTTGA TAACCTTATTTTTGACGAGGGGAAATTAAT AGGTTGTATTGATGTTGGACGAGTCGGAAT CGCAGACCGATACCAGGATCTTGCCATCCT ATGGAACTGCCTCGGTGAGTTTTCTCCTTCA TTACAGAAACGGCTTTTTCAAAAATATGGT

[0124]

SEQ ID NO.	描述	质粒序列
		ATTGATAATCCTGATATGAATAAATTGCAG TTTCATTTGATGCTCGATGAGTTTTTCTAAC TGTCAGACCAAGTTTACTCATATATACTTTA GATTGATTTAAAACTTCATTTTTAATTTAAA AGGATCTAGGTGAAGATCCTTTTTGATAAT CTCATGACCAAAATCCCTTAACGTGAGTTTT CGTTCCACTGAGCGTCAGACCCCGTAGAAA AGATCAAAGGATCTTCTTGAGATCCTTTTTT TCTGCGCGTAATCTGCTGCTTGCAAACAAA AAAACCACCGCTACCAGCGGTGGTTTGTTT GCCGGATCAAGAGCTACCAACTCTTTTTCC GAAGGTAACTGGCTTCAGCAGAGCGCAGAT ACCAAATACTGTTCTTCTAGTGTAGCCGTAG TTAGGCCACCACTTCAAGAACTCTGTAGCA CCGCCTACATACCTCGCTCTGCTAATCCTGT TACCAGTGGCTGCTGCCAGTGGCGATAAGT CGTGTCTTACCGGGTTGGACTCAAGACGAT AGTTACCGGATAAGGCGCAGCGGTCGGGCT GAACGGGGGGTTCGTGCACACAGCCCAGCT TGGAGCGAACGACCTACACCGAACTGAGAT ACCTACAGCGTGAGCTATGAGAAAGCGCCA CGCTTCCCGAAGGGAGAAAGGCGGACAGG TATCCGGTAAGCGGCAGGGTCGGAACAGGA GAGCGCACGAGGGAGCTTCCAGGGGGAAA CGCCTGGTATCTTTATAGTCCTGTCGGGTTT CGCCACCTCTGACTTGAGCGTCGATTTTTGT GATGCTCGTCAGGGGGGCGGAGCCTATGGA AAAACGCCAGCAACGCGGCCTTTTTACGGT TCCTGGCCTTTTGCTGGCCTTTTGCTCACAT GTTCTTTCCTGCGTTATCCCCTGATTCTGTG GATAACCGTATTACCGCCTTTGAGTGAGCT GATACCGCTCGCCGCAGCCGAACGACCGAG CGCAGCGAGTCAGTGAGCGAGGAAGCGGA AGAGCGCCCAATACGCAAACCGCCTCTCCC CGCGCGTTGGCCGATTCATTAATGCAGCTG GCACGACAGGTTTCCCGACTGGAAAGCGGG CAGTGAGCGCAACGCAATTAATGTGAGTTA GCTCACTCATTAGGCACCCCAGGCTTTACA CTTTATGCTTCCGGCTCGTATGTTGTGTGGA ATTGTGAGCGGATAACAATTTCACACAGGA AACAGCTATGACCATGATTACGCCAGATTT AATTAAGGCCTTAATTAGG
11	质粒 P1071 位置 1-130 处的 5' ITR 位置 219-1977 处的 SLC6A14v3 启动子 位置 1992-3044 处的鼠类 Atoh1 序列	CTGCGCGCTCGCTCGCTCACTGAGGCCGCC CGGGCAAAGCCCGGGCGTCGGGCGACCTTT GGTCGCCCGGCCTCAGTGAGCGAGCGAGCG CGCAGAGAGGGAGTGGCCAACTCCATCACT AGGGGTTCCTTGTAGTTAATGATTAACCCG CCATGCTACTTATCTACGTAGCCATGCTCTA GGAAGATCGGAATTCGCCCTTAAGCTAGCG GCGCGCCATACACTTATGTATATGTGCGAT GTCAGTGTGTGTGCATATAAAGTCCCAAAC AAGCCTGTATGATATTGACCAACAAGGTCA

[0125]

SEQ ID NO.	描述	质粒序列
	位置 3055-3602 处的 WPRE 位置 3615-3822 处的 polyA 位置 3910-4039 处的 3' ITR	AGGCAAAGTTTTGATACTTTCAGGTCACAA CCTCTCCGCATCCCTCTCTACTTTGCTCTAT CTGCCTGAACTCCTGAGGACATGTTTCTACT GCAAATGGAAAATCCTTGTCAGCCAGTGAG GAACAAAGGGACTATACATAGATGAAAACT TGGCTCTCTGCTGGTTCCTTTGTTTGTATGA ATTTATACAATTTGGTAAAACTGCCACCAT GTCTTACATGGACAGATTGAGTGTAGATTC TTTGAATTTTTGATGAAGAGGCGCTGCACT GGTGATCGGAATTGCAGTCTTTCCTCTGTAG GTAACCTGGCTTGTTTCCTTACAGTTTACTT TCTAGGCCTCGCCTTTCTCACAGAGTGAAGT CCTTTGTTAAGGTTCGAATTTCCCATAAACC TGCTCAATAATTTGTTTGTGTTTGGCTTCTTT GAAATACTACACAAAGCAATCCTTGTAAAA GGCAAAACTATTCCGAAGGCTGAGAAAGG AGCTCCAGGACATAGATTCAAAGTCGCTCT TTTCAGGTAGAGACAGCTGGGTAATCTTAT CTTAACTGGCTACATTTCAAGGTTCCCAATT CAGGGGCTTTCCCCTCTGGGAGCAGCATTC TCTCCGGGTGATGAAGAGCTTTCTAGTGAG GAGCAAAACTTTCAGAAAACCGGAGGGCCC AGAGCAGTCTGGTCTGTTCACAAAAATTAT AGCAAACAAAATAAGCCCGGCGGATTGGGT CTCTCCTACCTCCAGCACCAGGGGAGATCA GCACTTGGCCCCAGGACAGAGACCTGAGAA GTGAGGTTTGGAAGAAGCCAGGAATCCAGG AAAGGAGGCAAGATTGCTAAGGCACCGGC ACAGCTCTGAGTCAAAAGTTGTCAGTCTTCT TTGGCTCTGGCTGCGGAGCTCAATTGCTCAC AGCCCTGCCCTTTCCTAGGGCTGGGGCAAG GAATTGCTACATTCAGGATTACCTGGGGGA AAAACCAGAGGCTTGCTTTGGTCCCTTCCG GTAATTGAAAGGACTGGCCGTCAGCGAGGG GGAGGAGAGAGCTTCCCTCCATAAATGGTC CCACCCCTGGGCAAGGTGGCTCACTTTGGC AGGTAGCAACCGGGGAGTGTGCACCTGCCA CCAGTCAAGCTCAGCCAGACTGTGAGAAGA GGAGAGGCGAGGCACACCAAGGGATCCAG TGAACCAACGACAGATTGAAGTGCCCGAAC TTCTTCAAGTGCAGACAGAAGGAGGTAGGG TTCTGGAAGTTTCTGGTGGTGTAGGGGAGT CAGGAAGGGAAAACAAGGAGGGAGAGTGA GTCTTAGTTTTTTGCTTTCTGTAGCTGTTCCT TATTTTGCATATTTCTTTCTCTTCAACTCTTT TCAAGTATGCCTGATACGTTGTTCTCACGAA GTTGACGTGAAAAACAACTTTCCTGCTGGT AGTTAGGAAACTTAGGAGCACCTCAACCTG TACCTTGAGAACACCCAGAGAATGCTGCTC TTTGTCGTTCTCTATACCGTGTTCATATGCT GCAGGGAAATGCAAAGAATGTACTGTCCTT ATCTGACCCTGGGAGCATTCCATAGTCAAG

[0126]

SEQ ID NO.	描述	质粒序列
		CAGCAGCTATCAGGTTGGGAAAGAGCTCCT CTCCAAGGTGTAACAGAAAAGGAAAATGTT GATATTTTTCTTGTTTAGAAAGTGACAGCTT CATCCGAGAACGCGGCCGCGCCACCATGTC CCGCCTGCTGCATGCAGAAGAGTGGGCTGA GGTAAAAGAGTTGGGGGACCACCATCGCCA TCCCCAGCCGCACCACGTCCCGCCGCTGAC GCCACAGCCACCTGCTACCCTGCAGGCGAG AGACCTTCCCGTCTACCCGGCAGAACTGTC CCTCCTGGATAGCACCGACCCACGCGCCTG GCTGACTCCCACTTTGCAGGGCCTCTGCAC GGCACGCGCCGCCCAGTATCTGCTGCATTC TCCCGAGCTGGGTGCCTCCGAGGCCGCGGC GCCCCGGGACGAGGCTGACAGCCAGGGTG AGCTGGTAAGGAGAAGCGGCTGTGGCGGCC TCAGCAAGAGCCCCGGGCCCGTCAAAGTAC GGGAACAGCTGTGCAAGCTGAAGGGTGGG GTTGTAGTGGACGAGCTTGGCTGCAGCCGC CAGCGAGCCCCTTCCAGCAAACAGGTGAAT GGGGTACAGAAGCAAAGGAGGCTGGCAGC AAACGCAAGGGAACGGCGCAGGATGCACG GGCTGAACCACGCCTTCGACCAGCTGCGCA ACGTTATCCCGTCCTTCAACAACGACAAGA AGCTGTCCAAATATGAGACCCTACAGATGG CCCAGATCTACATCAACGCTCTGTCGGAGT TGCTGCAGACTCCCAATGTCGGAGAGCAAC CGCCGCCGCCCACAGCTTCCTGCAAAAATG ACCACCATCACCTTCGCACCGCCTCCTCCTA TGAAGGAGGTGCGGGCGCCTCTGCGGTAGC TGGGGCTCAGCCAGCCCCGGGAGGGGGCCC GAGACCTACCCCGCCCGGGCCTTGCCGGAC TCGCTTCTCAGGCCCAGCTTCCTCTGGGGGT TACTCGGTGCAGCTGGACGCTTTGCACTTCC CAGCCTTCGAGGACAGGGCCCTAACAGCGA TGATGGCACAGAAGGACCTGTCGCCTTCGC TGCCCGGGGGCATCCTGCAGCCTGTACAGG AGGACAACAGCAAAACATCTCCCAGATCCC ACAGAAGTGACGGAGAGTTTTCCCCCCACT CTCATTACAGTGACTCTGATGAGGCCAGTT AGAAGCTTGGATCCAATCAACCTCTGGATT ACAAAATTTGTGAAAGATTGACTGGTATTC TTAACTATGTTGCTCCTTTTACGCTATGTGG ATACGCTGCTTTAATGCCTTTGTATCATGCT ATTGCTTCCCGTATGGCTTTCATTTTCTCCTC CTTGTATAAATCCTGGTTGCTGTCTCTTTAT GAGGAGTTGTGGCCCGTTGTCAGGCAACGT GGCGTGGTGTGCACTGTGTTTGCTGACGCA ACCCCCACTGGTTGGGGCATTGCCACCACC TGTCAGCTCCTTTCCGGGACTTTCGCTTTCC CCCTCCCTATTGCCACGGCGGAACTCATCG CCGCCTGCCTTGCCCGCTGCTGGACAGGGG CTCGGCTGTTGGGCACTGACAATTCCGTGG

[0127]

SEQ ID NO.	描述	质粒序列
		TGTTGTCGGGGAAATCATCGTCCTTTCCTTG GCTGCTCGCCTGTGTTGCCACCTGGATTCTG CGCGGGACGTCCTTCTGCTACGTCCCTTCGG CCCTCAATCCAGCGGACCTTCCTTCCCGCGG CCTGCTGCCGGCTCTGCGGCCTCTTCCGCGT CTTCGAGATCTGCCTCGACTGTGCCTTCTAG TTGCCAGCCATCTGTTGTTTGCCCCTCCCCC GTGCCTTCCTTGACCCTGGAAGGTGCCACTC CCACTGTCCTTTCCTAATAAAATGAGGAAA TTGCATCGCATTGTCTGAGTAGGTGTCATTC TATTCTGGGGGGTGGGGTGGGGCAGGACAG CAAGGGGGAGGATTGGGAAGACAATAGCA GGCATGCTGGGGACTCGAGTTAAGGGCGAA TTCCCGATAAGGATCTTCCTAGAGCATGGC TACGTAGATAAGTAGCATGGCGGGTTAATC ATTAACTACAAGGAACCCCTAGTGATGGAG TTGGCCACTCCCTCTCTGCGCGCTCGCTCGC TCACTGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCTCAG TGAGCGAGCGAGCGCGCAGCCTTAATTAAC CTAATTCACTGGCCGTCGTTTTACAACGTCG TGACTGGGAAAACCCTGGCGTTACCCAACT TAATCGCCTTGCAGCACATCCCCCTTTCGCC AGCTGGCGTAATAGCGAAGAGGCCCGCACC GATCGCCCTTCCCAACAGTTGCGCAGCCTG AATGGCGAATGGGACGCGCCCTGTAGCGGC GCATTAAGCGCGGCGGGTGTGGTGGTTACG CGCAGCGTGACCGCTACACTTGCCAGCGCC CTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT CCTTTCTCGCCACGTTCGCCGGCTTTCCCCG TCAAGCTCTAAATCGGGGGCTCCCTTTAGG GTTCCGATTTAGTGCTTTACGGCACCTCGAC CCCAAAAAACTTGATTAGGGTGATGGTTCA CGTAGTGGGCCATCGCCCTGATAGACGGTT TTTCGCCCTTTGACGTTGGAGTCCACGTTCT TTAATAGTGGACTCTTGTTCCAAACTGGAA CAACACTCAACCCTATCTCGGTCTATTCTTT TGATTTATAAGGGATTTTGCCGATTTCGGCC TATTGGTTAAAAAATGAGCTGATTTAACAA AAATTTAACGCGAATTTTAACAAAATATTA ACGCTTACAATTTAGGTGGCACTTTTCGGG GAAATGTGCGCGGAACCCCTATTTGTTTATT TTTCTAAATACATTCAAATATGTATCCGCTC ATGAGACAATAACCCTGATAAATGCTTCAA TAATATTGAAAAAGGAAGAGTATGAGCCAT ATTCAACGGGAAACGTCGAGGCCGCGATTA AATTCCAACATGGATGCTGATTTATATGGG TATAAATGGGCTCGCGATAATGTCGGGCAA TCAGGTGCGACAATCTATCGCTTGTATGGG AAGCCCGATGCGCCAGAGTTGTTTCTGAAA CATGGCAAAGGTAGCGTTGCCAATGATGTT ACAGATGAGATGGTCAGACTAAACTGGCTG

[0128]

SEQ ID NO.	描述	质粒序列
		ACGGAATTTATGCCTCTTCCGACCATCAAG CATTTTATCCGTACTCCTGATGATGCATGGT TACTCACCACTGCGATCCCCGGAAAAACAG CATTCCAGGTATTAGAAGAATATCCTGATT CAGGTGAAAATATTGTTGATGCGCTGGCAG TGTTCCTGCGCCGGTTGCATTCGATTCCTGT TTGTAATTGTCCTTTTAACAGCGATCGCGTA TTTCGTCTTGCTCAGGCGCAATCACGAATG AATAACGGTTTGGTTGATGCGAGTGATTTT GATGACGAGCGTAATGGCTGGCCTGTTGAA CAAGTCTGGAAAGAAATGCATAAACTTTTG CCATTCTCACCGGATTCAGTCGTCACTCATG GTGATTTCTCACTTGATAACCTTATTTTTGA CGAGGGGAAATTAATAGGTTGTATTGATGT TGGACGAGTCGGAATCGCAGACCGATACCA GGATCTTGCCATCCTATGGAACTGCCTCGGT GAGTTTTCTCCTTCATTACAGAAACGGCTTT TTCAAAAATATGGTATTGATAATCCTGATAT GAATAAATTGCAGTTTCATTTGATGCTCGAT GAGTTTTTCTAACTGTCAGACCAAGTTTACT CATATATACTTTAGATTGATTTAAAACTTCA TTTTTAATTTAAAAGGATCTAGGTGAAGAT CCTTTTTGATAATCTCATGACCAAAATCCCT TAACGTGAGTTTTCGTTCCACTGAGCGTCAG ACCCCGTAGAAAAGATCAAAGGATCTTCTT GAGATCCTTTTTTTCTGCGCGTAATCTGCTG CTTGCAAACAAAAAAACCACCGCTACCAGC GGTGGTTTGTTTGCCGGATCAAGAGCTACC AACTCTTTTTCCGAAGGTAACTGGCTTCAGC AGAGCGCAGATACCAAATACTGTTCTTCTA GTGTAGCCGTAGTTAGGCCACCACTTCAAG AACTCTGTAGCACCGCCTACATACCTCGCTC TGCTAATCCTGTTACCAGTGGCTGCTGCCAG TGGCGATAAGTCGTGTCTTACCGGGTTGGA CTCAAGACGATAGTTACCGGATAAGGCGCA GCGGTCGGGCTGAACGGGGGGTTCGTGCAC ACAGCCCAGCTTGGAGCGAACGACCTACAC CGAACTGAGATACCTACAGCGTGAGCTATG AGAAAGCGCCACGCTTCCCGAAGGGAGAA AGGCGGACAGGTATCCGGTAAGCGGCAGG GTCGGAACAGGAGAGCGCACGAGGGAGCT TCCAGGGGGAAACGCCTGGTATCTTTATAG TCCTGTCGGGTTTCGCCACCTCTGACTTGAG CGTCGATTTTTGTGATGCTCGTCAGGGGGG CGGAGCCTATGGAAAAACGCCAGCAACGC GGCCTTTTTACGGTTCCTGGCCTTTTGCTGG CCTTTTGCTCACATGTTCTTTCCTGCGTTATC CCCTGATTCTGTGGATAACCGTATTACCGCC TTTGAGTGAGCTGATACCGCTCGCCGCAGC CGAACGACCGAGCGCAGCGAGTCAGTGAG CGAGGAAGCGGAAGAGCGCCCAATACGCA AACCGCCTCTCCCCGCGCGTTGGCCGATTC

[0129]

SEQ ID NO.	描述	质粒序列
		ATTAATGCAGCTGGCACGACAGGTTTCCCG ACTGGAAAGCGGGCAGTGAGCGCAACGCA ATTAATGTGAGTTAGCTCACTCATTAGGCA CCCCAGGCTTTACACTTTATGCTTCCGGCTC GTATGTTGTGTGGAATTGTGAGCGGATAAC AATTTCACACAGGAAACAGCTATGACCATG ATTACGCCAGATTTAATTAAGGCCTTAATTA GG

[0130] 将外源核酸递送至靶细胞的方法

[0131] 可用于将转基因(例如可操作地连接至本文所述的SLC6A14启动子的转基因)引入靶细胞(例如哺乳动物细胞)中的技术是本领域中众所周知的。例如,可使用电穿孔通过将静电势施加至目标细胞(例如人类靶细胞)使哺乳动物细胞透化。以此方式经受外部电场的哺乳动物细胞(例如人类细胞)随后易于摄取外源核酸。哺乳动物细胞的电穿孔详细阐述于例如Chu等人,Nucleic Acids Research 15:1311(1987)中,所述文献的公开内容通过引用并入本文。类似技术Nucleofection™利用施加的电场来刺激外源多核苷酸摄取至真核细胞的核中。Nucleofection™和可用于实施此技术的方案详细阐述于例如Distler等人,Experimental Dermatology 14:315(2005)以及US 2010/0317114中,所述文献各自的公开内容通过引用并入本文。

[0132] 可用于转染靶细胞的其他技术包括挤压-穿孔方法。此技术诱导细胞的快速机械变形以刺激经由响应于所施加应力形成的膜孔摄取外源DNA。此技术的优点在于,载体并非将核酸递送至细胞(例如人类靶细胞)中所必需。挤压-穿孔详细阐述于例如Sharei等人,Journal of Visualized Experiments 81:e50980(2013)中,所述文献的公开内容通过引用并入本文。

[0133] 脂转染代表可用于转染靶细胞的另一种技术。此方法涉及将核酸加载至脂质体中,所述脂质体通常呈现朝向脂质体外部的阳离子官能团,例如季胺或质子化胺。这由于细胞膜的阴离子性质而促进脂质体与细胞之间的静电相互作用,最终例如通过引导脂质体与细胞膜融合或通过复合物的胞吞作用摄取外源核酸。脂转染详细阐述于例如美国专利第7,442,386号中,所述专利的公开内容通过引用并入本文。利用与细胞膜的离子相互作用来引起外源核酸摄取的类似技术包括使细胞与阳离子聚合物-核酸复合物接触。与多核苷酸缔合以赋予有利于与细胞膜相互作用的正电荷的示例性阳离子分子包括活化的树枝状聚合物(阐述于例如Dennig,Topics in Current Chemistry 228:227(2003)中,所述文献的公开内容通过引用并入本文)、聚乙烯亚胺和二乙基氨基乙基(DEAE)-聚葡萄糖,其作为转染剂的用途详细阐述于例如Gulick等人,Current Protocols in Molecular Biology 40:I:9.2:9.2.1(1997)中,所述文献的公开内容通过引用并入本文。磁珠是可用于以温和且有效的方式转染靶细胞的另一种工具,这是因为这种方法利用所施加的磁场来引导核酸的摄取。这种技术详细阐述于例如US 2010/0227406中,所述专利的公开内容通过引用并入本文。

[0134] 另一种可用于诱导靶细胞对外源核酸的摄取的工具是激光转染,也称为光学转染,其是涉及使细胞暴露于特定波长的电磁辐射以温和地透化细胞并允许多核苷酸透过细

胞膜的技术。这种技术的生物活性类似于电穿孔,并且在一些情况下发现优于电穿孔。

[0135] 刺穿转染是可用于将遗传材料递送至靶细胞的另一种技术。其依赖于纳米材料(例如碳纳米纤维、碳纳米管和纳米线)的使用。垂直于基底的表面来合成针样纳米结构。将含有旨在细胞内递送的基因的DNA附接至纳米结构表面。然后将具有这些针的阵列的芯片按压在细胞或组织上。由纳米结构刺穿的细胞可表达所递送的基因。这种技术的实例阐述于Shalek等人,PNAS107:1870(2010)中,所述文献的公开内容通过引用并入本文。

[0136] 也可使用磁性转染将核酸递送至靶细胞。磁性转染原理是使核酸与阳离子磁性纳米粒子缔合。磁性纳米粒子由完全生物可降解的氧化铁制成,并且涂覆有根据应用变化的特定阳离子专有分子。其与基因载体(DNA、RNA、病毒载体等)的缔合是通过盐诱导的胶体聚集和静电相互作用来实现。然后通过影响磁铁所产生的外部磁场使磁性粒子集中于靶细胞上。这种技术详细阐述于Scherer等人,Gene Therapy9:102(2002),所述文献的公开内容通过引用并入本文。

[0137] 另一种可用于诱导靶细胞对外源核酸的摄取的工具是声致穿孔,其是涉及使用声音(通常超声波频率)来改变细胞质膜的透性以透化细胞并允许多核苷酸透过细胞膜的技术。这种技术详细阐述于例如Rhodes等人,Methods in Cell Biology 82:309(2007),所述文献的公开内容通过引用并入本文。

[0138] 微泡代表可用于根据本文所述的方法修饰靶细胞的基因组的另一种潜在媒介物。例如,已通过糖蛋白VSV-G与例如基因组修饰蛋白(例如核酸酶)的共同过表达诱导的微泡可用于将蛋白质有效地递送至细胞中,随后催化内源多核苷酸序列的位点特异性裂解以使细胞的基因组准备用于共价并入目标多核苷酸(例如基因或调节序列)。此类囊泡(也称为纳米囊泡(Gesicle))用于真核细胞的遗传修饰的用途详细阐述于例如Quinn等人,Genetic Modification of Target Cells by Direct Delivery ofActive Protein[摘要].Methylation changes in early embryonic genes in cancer[摘要],Proceedings of the 18th Annual Meeting ofthe American SocietyofGene and Cell Therapy;2015年5月13日,摘要编号122。

[0139] 将外源核酸递送至靶细胞的载体

[0140] 除实现高转录和翻译速率以外,外源基因在哺乳动物细胞中的稳定表达可通过将含有基因的多核苷酸整合至哺乳动物细胞的核基因组中来实现。已开发出将编码外源蛋白质的多核苷酸递送并整合至哺乳动物细胞的核DNA中的多种载体。表达载体的实例阐述于例如Gellissen,Production of Recombinant Proteins:Novel Microbial and Eukaryotic Expression Systems(JohnWiley&Sons,Marblehead,MA,2006)中。用于本文所述组合物和方法中的表达载体含有可操作地连接至编码目标蛋白质的多核苷酸序列的SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸),以及例如用于表达这些剂和/或将这些多核苷酸序列整合至哺乳动物细胞的基因组中的其他序列元件。可含有可操作地连接至编码目标蛋白质的转基因的SLC6A14启动子的载体包括质粒(例如可在细胞内自主复制的环状DNA分子)、粘粒(例如pWE或sCos载体)、人工染色体(例如人类人工染色体(HAC)、酵母人工染色体(YAC)、细菌人工染色体(BAC)或P1源性人工染色体(PAC))和病毒载体。可用于表达目标蛋白质的某些载体包

括含有引导基因转录的调节序列(例如增强子区域)的质粒。其他可用于表达目标蛋白质的载体含有增强这些基因的翻译速率或改进源自基因转录的mRNA的稳定性或核输出的多核苷酸序列。这些序列元件包括例如5'和3'非翻译区、内部核糖体进入位点(IRES)和引导表达载体上所携带基因的有效转录的多聚腺苷酸化信号位点。适用于本文所述组合物和方法的表达载体也可含有编码用于选择含有这种载体的细胞的标志物的多核苷酸。合适的标志物的实例包括编码抗生素(例如氨苄青霉素、氯霉素、卡那霉素或诺尔丝菌素)抗性的基因。

[0141] 用于核酸递送的病毒载体

[0142] 病毒基因组提供可用于将目标基因有效递送至靶细胞(例如哺乳动物细胞,例如人类细胞)的基因组中的载体的丰富来源。病毒基因组是特别可用于基因递送的载体,这是因为含于此类基因组内的多核苷酸通常通过一般或专门转导并入哺乳动物细胞的核基因组中。这些过程作为天然病毒复制周期的一部分出现,并且无需添加蛋白质或试剂来诱导基因整合。病毒载体的实例包括逆转录病毒(例如逆转录病毒科病毒载体)、腺病毒(例如Ad5、Ad26、Ad34、Ad35和Ad48)、小病毒(例如腺相关病毒)、冠状病毒、负链RNA病毒如正粘液病毒(例如流行性感冒病毒)、棒状病毒(例如狂犬病和水疱性口炎病毒)、副粘液病毒(例如麻疹和仙台病毒(Sendai))、正链RNA病毒如微小RNA病毒和α病毒以及双链DNA病毒,包括腺病毒、疱疹病毒(例如1型和2型单纯疱疹病毒、爱泼斯坦-巴尔病毒(Epstein-Barr virus)、巨细胞病毒)和痘病毒(例如牛痘、经修饰安卡拉牛痘(modified vaccinia Ankara,MVA)、鸡痘和金丝雀痘)。其他病毒包括例如诺沃克病毒(Norwalk virus)、披衣病毒、黄病毒、呼肠孤病毒(reovirus)、乳多泡病毒、嗜肝DNA病毒、人类乳头瘤病毒、人类泡沫病毒和肝炎病毒。逆转录病毒的实例包括:鸟白血病-肉瘤、鸟C型病毒、哺乳动物C型、B型病毒、D型病毒、致癌逆转录病毒、HTLV-BLV族、慢病毒、α逆转录病毒、γ逆转录病毒、泡沫病毒(Coffin,J.M.,Retroviridae:The viruses and their replication,Virology,第三版(Lippincott-Raven,Philadelphia,1996))。其他实例包括鼠类白血病病毒、鼠类肉瘤病毒、小鼠乳腺肿瘤病毒、牛白血病病毒、猫白血病病毒、猫肉瘤病毒、鸟白血病病毒、人类T细胞白血病病毒、狒狒内源病毒、长臂猿白血病病毒、梅森菲舍猴病毒(MasonPfizermonkeyvirus)、猿猴免疫缺失病毒、猿猴肉瘤病毒、劳斯肉瘤病毒(Rous sarcoma virus)和慢病毒。载体的其他实例阐述于例如美国专利第5,801,030号中,所述专利关于用于基因疗法中的病毒载体的公开内容通过引用并入本文。

[0143] 用于核酸递送的AAV载体

[0144] 在一些实施方案中,将本文所述组合物和方法的多核苷酸并入rAAV载体和/或病毒粒子中以促进其引入细胞(例如VSC)中。可用于本文所述组合物和方法中的rAAV载体是重组核酸构建体,其包括(1)本文所述的SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸),(2)待表达的异源序列,和(3)促进异源基因的稳定性和表达的病毒序列。病毒序列可包括DNA顺式复制和包装(例如功能性ITR)至病毒粒子中所需的那些AAV序列。在典型应用中,转基因编码可促进或增加前庭毛细胞发育、前庭毛细胞命运特化、前庭毛细胞再生、前庭毛细胞和/或VSC增殖、前庭毛细胞神经支配或前庭毛细胞成熟的蛋白质,或在患有多种形式的遗传性前庭功能障碍的受试者中突变的前庭毛细胞蛋白的野生型形式,其可用于改进携带与前庭功能障碍(例如头晕、眩晕、失

衡、双侧前庭病、双侧前庭功能低下、振动幻视或平衡障碍)相关的突变的受试者的前庭功能。此类rAAV载体也可含有标志物或报告基因。有用的rAAV载体具有一个或多个整体或部分缺失的AAV WT基因,但保留功能性侧接ITR序列。AAV ITR可具有适于特定应用的任何血清型。为用于本文所述方法和组合物中,ITR可为AAV2 ITR。使用rAAV载体的方法阐述于例如Tal等人,J.Biomed.Sci.7:279(2000)以及Monahan和Samulski,Gene Delivery7:24(2000)中,所述文献中的每一者关于用于基因递送的AAV载体的公开内容都通过引用并入本文。

[0145] 本文所述的多核苷酸和载体(例如可操作地连接至编码目标蛋白质的转基因的SLC6A14启动子)可并入rAAV病毒粒子中以促进多核苷酸或载体引入细胞(例如VSC)中。AAV的衣壳蛋白构成外部,病毒粒子的非核酸部分是由AAVcap基因编码。cap基因编码三种病毒外壳蛋白VP1、VP2和VP3,其为病毒粒子组装所必需。rAAV病毒粒子的构建已阐述于例如US 5,173,414;US 5,139,941;US 5,863,541;US 5,869,305;US 6,057,152;和US 6,376,237;以及Rabinowitz等人,J.Virol.76:791(2002)和Bowles等人,J.Virol.77:423(2003)中,所述文献中的每一者关于用于基因递送的AAV载体的公开内容都通过引用并入本文。

[0146] 可与本文所述的组合物和方法结合使用的rAAV病毒粒子包括衍生自多种AAV血清型的那些rAAV病毒粒子,所述AAV血清型包括AAV 1、2、3、4、5、6、7、8、9、10、11、rh10、rh39、rh43、rh74、Anc80、Anc80L65、DJ/8、DJ/9、7m8、PHP.B、PHP.eB和PHP.S。为靶向VSC,AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、Anc80、Anc80L65、7m8、PHP.B、PHP.eB或PHP.S血清型可能是特别有用的。经进化用于转导视网膜的血清型也可用于本文所述方法和组合物中。不同血清型的AAV载体和AAV蛋白的构建和用途阐述于例如Chao等人,Mol.Ther.2:619(2000);Davidson等人,Proc.Natl.Acad.Sci.USA97:3428(2000);Xiao等人,J.Virol.72:2224(1998);Halbert等人,J.Virol.74:1524(2000);Halbert等人,J.Virol.75:6615(2001);和Auricchio等人,Hum.Molec.Genet.10:3075(2001)中,所述文献中的每一者关于用于基因递送的AAV载体的公开内容都通过引用并入本文。

[0147] 假型rAAV载体也可与本文所述的组合物和方法结合使用。假型载体包括经衍生自除给定血清型(例如AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8等)外的血清型的衣壳基因假型化的给定血清型(例如AAV9)的AAV载体。涉及假型rAAV病毒粒子的构建和用途的技术是本领域中已知的并且阐述于例如Duan等人,J.Virol.75:7662(2001);Halbert等人,J.Virol.74:1524(2000);Zolotukhin等人,Methods,28:158(2002);和Auricchio等人,Hum.Molec.Genet.10:3075(2001)中。

[0148] 在病毒粒子衣壳内具有突变的AAV病毒粒子可用于比非突变衣壳病毒粒子更有效地感染特定细胞类型。例如,合适的AAV突变体可具有帮助AAV靶向特定细胞类型的配体插入突变。AAV衣壳突变体(包括插入突变体、丙氨酸筛选突变体和表位标签突变体)的构建和表征阐述于Wu等人,J.Virol.74:8635(2000)中。可用于本文所述方法中的其他rAAV病毒粒子包括通过病毒的分子育种以及通过外显子改组产生的那些衣壳杂合体。参见例如Soong等人,Nat.Genet.,25:436(2000)以及Kolman和Stemmer,Nat.Biotechnol.19:423(2001)。

[0149] 在一些实施方案中,核酸载体(例如AAV载体)包括本文所述的SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、

92％、93％、94％、95％、96％、97％、98％、99％或更大的序列同一性)的多核苷酸),其可操作地连接至编码人类Atoh1的多核苷酸序列(人类ATOH1蛋白＝RefSeq登录号NP_005163(SEQ ID NO:4);mRNA序列＝RefSeq登录号NM_005172)。在一些实施方案中,SLC6A14启动子是SEQ ID NO:1的SLC6A14启动子(也由SEQ ID NO:10的核苷酸219-1977表示)并且其可操作地连接至编码人类Atoh1的多核苷酸序列。在一些实施方案中,编码人类Atoh1的多核苷酸序列是SEQ ID NO:5。在一些实施方案中,编码人类Atoh1的多核苷酸序列是因遗传密码的冗余性而编码SEQ ID NO:4的任何多核苷酸序列。编码人类Atoh1的多核苷酸序列可经部分或完全密码子优化用于表达。在一些实施方案中,载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码人类Atoh1的多核苷酸序列;多聚腺苷酸化序列;和第二反向末端重复。在一些实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码人类Atoh1的多核苷酸序列;土拨鼠转录后调节元件(WPRE);多聚腺苷酸化序列;和第二反向末端重复。在一些实施方案中,WPRE具有SEQ ID NO:8或SEQ ID NO:9的序列。在一些实施方案中,WPRE具有SEQ ID NO:8的序列。在一些实施方案中,WPRE具有SEQ ID NO:10的核苷酸3064-3611的序列。在一些实施方案中,多聚腺苷酸化序列具有SEQ ID NO:10的核苷酸3624-3831的序列。在某些实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:10的核苷酸219-3831。在一些实施方案中,反向末端重复是AAV2反向末端重复。在一些实施方案中,反向末端重复是AAV2反向末端重复的可经携带AAV2 Rep基因的质粒壳体化的任何变体。在特定实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:10的核苷酸219-3831,其中5'反向末端重复与SEQ ID NO:10的核苷酸1-130具有至少80％序列同一性(例如至少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);并且其中3'反向末端重复与SEQ ID NO:10的核苷酸3919-4048具有至少80％序列同一性(例如至少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)。在一些实施方案中,核酸载体是病毒载体。在一些实施方案中,病毒载体是AAV载体。在一些实施方案中,AAV载体具有AAV8衣壳。

[0150] 在一些实施方案中,核酸载体(例如AAV载体)包括本文所述的SLC6A14启动子(例如与SEQ ID NO:1具有至少85％序列同一性(例如85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或更大的序列同一性)的多核苷酸),其可操作地连接至编码鼠类Atoh1的多核苷酸序列(鼠类ATOH1蛋白＝UniProt P48985(SEQ ID NO:6);mRNA序列＝RefSeq登录号NM_007500.5)。在一些实施方案中,SLC6A14启动子是SEQ ID NO:1的SLC6A14启动子(也由SEQ ID NO:11的核苷酸219-1977表示)并且其可操作地连接至编码鼠类Atoh1的多核苷酸序列。在一些实施方案中,编码鼠类Atoh1的多核苷酸序列是SEQ ID NO:7。在一些实施方案中,编码鼠类Atoh1的多核苷酸序列是因遗传密码的冗余性而编码SEQ ID NO:6的任何多核苷酸序列。编码鼠类Atoh1的多核苷酸序列可经部分或完全密码子优化用于表达。在一些实施方案中,载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动子的编码鼠类Atoh1的多核苷酸序列;多聚腺苷酸化序列;和第二反向末端重复。在一些实施方案中,核酸载体以5'至3'顺序包括第一反向末端重复;SEQ ID NO:1的SLC6A14启动子;可操作地连接至SLC6A14启动

子的编码鼠类Atoh1的多核苷酸序列;土拨鼠转录后调节元件(WPRE);多聚腺苷酸化序列;和第二反向末端重复。在一些实施方案中,WPRE具有SEQ ID NO:8或SEQ ID NO:9的序列。在一些实施方案中,WPRE具有SEQ ID NO:8的序列。在一些实施方案中,WPRE具有SEQ ID NO:11的核苷酸3055-3602的序列。在一些实施方案中,多聚腺苷酸化序列具有SEQ ID NO:11的核苷酸3615-3822的序列。在某些实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:11的核苷酸219-3822。在一些实施方案中,反向末端重复是AAV2反向末端重复。在一些实施方案中,反向末端重复是AAV2反向末端重复的可经携带AAV2 Rep基因的质粒壳体化的任何变体。在特定实施方案中,核酸载体包括侧接有反向末端重复的SEQ ID NO:11的核苷酸219-3822,其中5'反向末端重复与SEQ ID NO:11的核苷酸1-130具有至少80%序列同一性(例如至少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);并且其中3'反向末端重复与SEQ ID NO:11的核苷酸3910-4039具有至少80%序列同一性(例如至少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)。在一些实施方案中,核酸载体是病毒载体。在一些实施方案中,病毒载体是AAV载体。在一些实施方案中,AAV载体具有AAV8衣壳。

[0151] 本领域的普通技术人员应理解,产生本发明的病毒载体通常需要使用本发明的质粒以及提供适合的病毒包装和活力所需的元件的额外质粒(例如对于AAV,提供适当的AAV rep基因、cap基因和其他基因(例如E2A和E4)的质粒)。产生细胞系中那些质粒的组合产生病毒载体。然而,本领域技术人员应理解,对于用于产生病毒载体的本发明转移质粒中的任何给定对的反向末端重复序列,病毒载体中的相应序列可因ITR在重组期间采取"翻转"或"倒转"取向而改变。因此,转移质粒中ITR的序列不必为在从其制备的病毒载体中发现的相同序列。

[0152] 药物组合物

[0153] 本文所述的多核苷酸(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的SLC6A14启动子)可以可操作地连接至转基因(例如编码目标蛋白质、shRNA、ASO或核酸酶(例如Cas9、TALEN、ZFN或gRNA)的转基因,或者可经转录以产生微小RNA的转基因)并且并入媒介物中以施用于患者,例如患有前庭功能障碍的人类患者。含有含可操作地连接至转基因的本文所述多核苷酸的载体(例如病毒载体)的药物组合物可使用本领域中已知的方法来制备。例如,此类组合物可使用例如生理上可接受的载体、赋形剂或稳定剂(Remington:The Science and Practice of Pharmacology,第22版,Allen,L.编辑,(2013);通过引用并入本文)并且以期望形式(例如以冻干制剂或水溶液的形式)来制备。

[0154] 含有可操作地连接至转基因的本文所述SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)的核酸载体(例如病毒载体)的混合物可在与一种或多种赋形剂、载体或稀释剂适当混合的水中制备。分散液也可在甘油、液体聚乙二醇和其混合物中以及油中制备。在普通储存和使用条件下,这些制剂可含有防止微生物生长的防腐剂。适于可注射用途的药物形式包括无菌水溶液或分散液和用于临时制备无菌可注射溶液或分散液的无菌粉末(阐述于US 5,466,468中,其公开内容通过

引用并入本文)。在任何情况下,制剂可为无菌的并且可为达到易于注射程度的流体。制剂在制造和储存条件下可为稳定的并且可防止微生物(例如细菌和真菌)的污染作用。载体可为含有例如水、乙醇、多元醇(例如甘油、丙二醇和液体聚乙二醇等)、其合适混合物和/或植物油的溶剂或分散介质。可例如通过使用包衣(例如卵磷脂)、在分散液的情形下通过维持所需粒径和通过使用表面活性剂来维持适当流动性。可通过多种抗细菌和抗真菌剂(例如对羟基苯甲酸酯、氯丁醇、苯酚、山梨酸、硫柳汞等)来防止微生物的作用。在许多情况下,应优选包括等渗剂,例如糖或氯化钠。可注射组合物的延长吸收可通过在组合物中使用延迟吸收剂(例如单硬脂酸铝和明胶)来实现。

[0155] 例如,如果需要,含有本文所述药物组合物的溶液可经适当缓冲,并且首先用足量盐水或葡萄糖使液体稀释剂等渗。这些特定水溶液尤其适用于静脉内、肌内、皮下和腹膜内施用。就此而言,根据本公开,可采用的无菌水性介质将为本领域技术人员已知的。例如,可将一个剂量溶解在1ml NaCl等渗溶液中并添加至1000ml皮下灌注液中或在所提出的输注位点处注射。根据所治疗受试者的疾患,剂量必然将发生一些变化。为局部施用于中耳或内耳,组合物可经配制以含有合成外淋巴溶液。示例性的合成外淋巴溶液包括20-200mMNaCl、1-5mM KCl、0.1-10mM $CaCl_2$、1-10mM葡萄糖和2-50mM HEPE,其中pH介于约6与9之间并且摩尔渗透压浓度为约300mOsm/kg。负责施用的人在任何情形下将决定个别受试者的适当剂量。此外,对于人类施用,制剂应符合FDA生物制品标准办公室(FDA Office of Biologics standards)所要求的无菌性、热原性、一般安全性和纯度标准。

[0156] 治疗方法

[0157] 本文所述的组合物可通过多种途径施用于患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者,例如局部施用于中耳或内耳(例如施用于外淋巴或内淋巴中,例如经由卵圆窗、圆窗或半规管(例如外半规管),或通过经鼓室或鼓室内注射,例如施用于前庭支持细胞或毛细胞)、静脉内、肠胃外、真皮内、透皮、肌内、鼻内、皮下、经皮、气管内、腹膜内、动脉内、血管内、吸入、灌注、灌洗和口服施用。在任何给定情况下最适于施用的途径将取决于所施用的特定组合物、患者、药物配制方法、施用方法(例如施用时间和施用途径)、患者的年龄、体重、性别、所治疗疾病的严重程度、患者的饮食和患者的排泄速率。组合物可施用一次或一次以上(例如每年一次、每年两次、每年三次、每两月、每月或每两周)。

[0158] 可如本文所述治疗的受试者是患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者。本文所述的组合物和方法可用于治疗患有前庭毛细胞损伤(例如与疾病或感染、头部创伤、耳毒性药物(例如氨基糖苷)或衰老相关的损伤)或具有患上所述疾病的风险的受试者,患有前庭功能障碍(例如头晕、眩晕、失衡、双侧前庭病(也称为双侧前庭功能低下)、振动幻视或平衡障碍)或具有患上所述疾病的风险的受试者,携带与前庭功能障碍相关的遗传突变的受试者或具有遗传性前庭功能障碍家族史的受试者。在一些实施方案中,与毛细胞(例如前庭毛细胞)损伤或损失相关的疾病是其中自身免疫反应导致毛细胞损伤或死亡的自身免疫疾病或疾患。与前庭功能障碍相关的自身免疫疾病包括自身免疫内耳疾病(AIED)、结节性多动脉炎(PAN)、科根综合征(Cogan's syndrome)、复发性多发性软骨炎、全身性红斑狼疮(SLE)、韦格纳肉芽肿病(Wegener's granulomatosis)、舍格伦综合征(Sjögren's syndrome)和白塞病(Behçet's disease)。一些传染病(例如莱姆病(Lyme disease)和梅毒)也可引起前庭功能障碍(例如通过触发自身抗体产生)。病毒感染(例如风

疹、巨细胞病毒(CMV)、淋巴细胞性脉络丛脑膜炎病毒(LCMV)、1型和2型HSV、西尼罗病毒(West Nile virus,WNV)、人类免疫缺失病毒(HIV)、水痘带状疱疹病毒(VZV)、麻疹和腮腺炎)也可引起前庭功能障碍。在一些实施方案中,受试者患有与毛细胞(例如前庭毛细胞)损失相关或源自其的前庭功能障碍。在一些实施方案中,本文所述的组合物和方法可用于治疗患有振动幻视或具有患上振动幻视的风险的受试者。在一些实施方案中,本文所述的组合物和方法可用于治疗患有双侧前庭病或具有患上双侧前庭病的风险的受试者。在一些实施方案中,本文所述的组合物和方法可用于治疗患有平衡障碍(例如失衡)或具有患上平衡障碍(例如失衡)的风险的受试者。本文所述的方法可包括在用本文所述的组合物治疗或施用本文所述的组合物之前筛选受试者的已知与前庭功能障碍相关的基因的一个或多个突变的步骤。受试者的遗传突变可使用本领域技术人员已知的标准方法(例如遗传性测试)来筛选。本文所述的方法也可包括在用本文所述的组合物治疗之前或施用本文所述的组合物之前评价受试者的前庭功能的步骤。前庭功能可使用例如以下的标准测试来评价:眼动测试(例如眼震图(ENG)或视频眼震图(VNG))、前庭眼反射(VOR)测试(例如可在床边实施的头部冲击测试(哈-克测试(Halmagyi-Curthoys test))或使用视频头部冲击测试(VHIT)或热量反射测试)、体位图、旋转椅测试、ECOG、前庭引发的肌源性电势(VEMP)和专门临床平衡测试,例如Mancini和Horak,Eur JPhys Rehabil Med,46:239(2010)中所述的那些测试。这些测试也可用于在用本文所述的组合物治疗或施用本文所述的组合物后评价受试者的前庭功能。本文所述的组合物和方法也可作为预防性治疗施用于具有患上前庭功能障碍的风险的患者,例如具有前庭功能障碍家族史(例如遗传性前庭功能障碍)的患者、尚未表现出前庭功能障碍症状的携带与前庭功能障碍相关的遗传突变的患者或暴露于获得性前庭功能障碍的风险因子(例如疾病或感染、头部创伤、耳毒性药物或衰老)的患者。本文所述的组合物和方法也可用于治疗患有特发性前庭功能障碍的受试者。

[0159] 本文所述的组合物和方法可用于诱导或增加受试者中的毛细胞再生(例如前庭毛细胞再生)和/或诱导或增加前庭毛细胞和/或VSC的增殖。可从促进或诱导前庭毛细胞再生、前庭毛细胞神经支配和/或前庭毛细胞和/或VSC增殖的组合物获益的受试者包括因毛细胞损失(例如与创伤(例如头部创伤)、疾病或感染、耳毒性药物或衰老相关的前庭毛细胞损失)而患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者,和具有异常前庭毛细胞(例如与正常前庭毛细胞相比不能适当地发挥功能的前庭毛细胞)、受损前庭毛细胞(例如与创伤(例如头部创伤)、疾病或感染、耳毒性药物或衰老相关的前庭毛细胞损伤)或因遗传突变或先天异常导致的前庭毛细胞数减少的受试者。本文所述的组合物和方法也可用于促进或增加前庭毛细胞成熟,这可改进前庭功能。在一些实施方案中,本文所述的组合物和方法促进或增加再生前庭毛细胞的成熟(例如促进或增加响应于本文所述的组合物(例如含有可操作地连接至转基因的SLC6A14启动子的组合物)在VSC中的表达形成的前庭毛细胞的成熟)。本文所述的组合物和方法也可促进或增加VSC和/或前庭毛细胞存活和/或改进VSC功能。

[0160] 本文所述的组合物和方法也可用于预防或减轻已用耳毒性药物治疗或当前正在用耳毒性药物治疗或很快开始用耳毒性药物治疗的受试者中由耳毒性药物诱导的毛细胞损伤或死亡(例如前庭毛细胞损伤或死亡)引起的前庭功能障碍。耳毒性药物对内耳细胞有毒性,并且可引起前庭功能障碍(例如眩晕、头晕、失衡、双侧前庭病或振动幻视)。已发现具

有耳毒性的药物包括氨基糖苷抗生素(例如庆大霉素、新霉素(neomycin)、链霉素(streptomycin)、妥布霉素(tobramycin)、卡那霉素、万古霉素(vancomycin)和阿米卡星(amikacin))、紫霉素(viomycin)、抗肿瘤药物(例如含铂化学治疗剂,例如顺铂、卡铂和奥沙利铂(oxaliplatin))、袢利尿剂(例如依他尼酸和呋塞米)、水杨酸盐(例如阿司匹林(aspirin),特别是在高剂量下)和奎宁。在一些实施方案中,本文所述的方法和组合物可用于治疗双侧前庭病。在一些实施方案中,本文所述的方法和组合物可用于治疗由氨基糖苷耳毒性诱导的双侧前庭病或振动幻视(例如在患有氨基糖苷诱导的双侧前庭病或振动幻视的受试者中,可使用本文所述的方法和组合物来减少氨基糖苷诱导的前庭毛细胞损伤或死亡,或促进或增加毛细胞再生和/或毛细胞或VSC增殖)。

[0161] 如本文所述可操作地连接至SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)来治疗受试者的转基因可为编码在健康VSC中表达的蛋白质(例如在前庭毛细胞发育、前庭毛细胞命运特化、前庭毛细胞再生、前庭毛细胞和/或VSC增殖、前庭毛细胞成熟或前庭毛细胞神经支配中起作用的蛋白质或患有前庭功能障碍的受试者中缺失的蛋白质)、另一目标蛋白质(例如治疗蛋白或报告蛋白,例如荧光蛋白、lacZ或荧光素酶)、shRNA、ASO、核酸酶或微小RNA的转基因。转基因可基于受试者前庭功能障碍的病因(例如如果受试者的前庭功能障碍与特定遗传突变相关,则转基因可为在受试者中突变的基因的野生型形式,或者如果受试者患有与毛细胞损失相关的前庭功能障碍,则转基因可编码促进前庭毛细胞再生、前庭毛细胞神经支配或前庭毛细胞和/或VSC增殖的蛋白质)、受试者前庭功能障碍的严重程度、受试者的毛细胞的健康状况、受试者的年龄、受试者的前庭功能障碍家族史或其他因素来选择。如本文所述可由可操作地连接SLC6A14启动子的转基因表达来治疗受试者的蛋白质包括Sox9、Sall2、Camta1、Hey2、Gata2、Hey1、Lass2、Sox10、Gata3、Cux1、Nr2f1、Hes1、Rorb、Jun、Zfp667、Lhx3、Nhlh1、Mxd4、Zmiz1、Myt1、Stat3、Barhl1、Tox、Prox1、Nfia、Thrb、Mycl1、Kdm5a、Creb314、Etv1、Peg3、Bach2、Isl1、Zbtb38、Lbh、Tub、Hmg20、Rest、Zfp827、Aff3、Pknox2、Arid3b、Mlxip、Zfp532、Ikzf2、Sall1、Six2、Sall3、Lin28b、Rfx7、Bdnf、Gfi1、Pou4f3、Myc、Ctnnb1、Sox2、Sox4、Sox11、Tead2、Atoh1和在氨基酸328、331和/或334处含有取代的Atoh1变体(例如S328A、S331A、S334A、S328A/S331A、S328A/S334A、S331A/S334A和328A/S331A/S334)。

[0162] 治疗可包括以不同单位剂量施用含有含本文所述的SLC6A14启动子(例如SEQ ID NO:1)的核酸载体(例如AAV病毒载体)的组合物。每个单位剂量通常将含有预定量的治疗性组合物。待施用的量和特定施用途径和制剂在临床领域的技术人员的技能范围内。单位剂量无需作为单次注射来施用,但可包括在设定的时间段内连续输注。可使用注射泵进行给药来控制输注速率以最小化内耳(例如前庭迷路)的损伤。在核酸载体为AAV载体(例如AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、rh10、rh39、rh43、rh74、Anc80、Anc80L65、DJ/8、DJ/9、7m8、PHP.B、PHP.eB或PHP.S载体)的情况下,病毒载体可以例如下列剂量施用于患者:约1×10^{9}个载体基因组(VG)/mL至约1×10^{16}VG/mL(例如$1x10^{9}$VG/mL、$2x\ 10^{9}$VG/mL、$3x\ 10^{9}$VG/mL、$4x\ 10^{9}$VG/mL、$5x\ 10^{9}$VG/mL、$6x\ 10^{9}$VG/mL、$7x\ 10^{9}$VG/mL、$8x\ 10^{9}$VG/mL、$9x\ 10^{9}$VG/mL、$1x\ 10^{10}$VG/mL、$2x\ 10^{10}$VG/mL、$3x\ 10^{10}$VG/mL、$4x\ 10^{10}$VG/mL、$5x\ 10^{10}$VG/mL、$6x\ 10^{10}$VG/mL、$7x\ 10^{10}$VG/mL、$8x\ 10^{10}$VG/mL、$9x\ 10^{10}$VG/mL、1x

10^{11}VG/mL、2x 10^{11}VG/mL、3x 10^{11}VG/mL、4x 10^{11}VG/mL、5x 10^{11}VG/mL、6x 10^{11}VG/mL、7x 10^{11}VG/mL、8x 10^{11}VG/mL、9x 10^{11}VG/mL、1x 10^{12}VG/mL、2x 10^{12}VG/mL、3x 10^{12}VG/mL、4x 10^{12}VG/mL、5x 10^{12}VG/mL、6x 10^{12}VG/mL、7x 10^{12}VG/mL、8x 10^{12}VG/mL、9x 10^{12}VG/mL、1x 10^{13}VG/mL、2x 10^{13}VG/mL、3x 10^{13}VG/mL、4x 10^{13}VG/mL、5x 10^{13}VG/mL、6x 10^{13}VG/mL、7x 10^{13}VG/mL、8x 10^{13}VG/mL、9x 10^{13}VG/mL、1x 10^{14}VG/mL、2x 10^{14}VG/mL、3x 10^{14}VG/mL、4x 10^{14}VG/mL、5x 10^{14}VG/mL、6x 10^{14}VG/mL、7x 10^{14}VG/mL、8x 10^{14}VG/mL、9x 10^{14}VG/mL、1x 10^{15}VG/mL、2x 10^{15}VG/mL、3x 10^{15}VG/mL、4x 10^{15}VG/mL、5x 10^{15}VG/mL、6x 10^{15}VG/mL、7x 10^{15}VG/mL、8x 10^{15}VG/mL、9x 10^{15}VG/mL或1x 10^{16}VG/mL)于1μL至200μL(例如1μL、2μL、3μL、5μL、6μL、7μL、8μL、9μL、10μL、15μL、20μL、25μL、30μL、35μL、40μL、45μL、50μL、55μL、60μL、65μL、70μL、75μL、80μL、85μL、90μL、95μL、100μL、110μL、120μL、130μL、140μL、150μL、160μL、170μL、180μL、190μL或200μL)的体积中。AAV载体可以下列剂量施用于受试者:约1x 10^{7}VG/耳至约2x 10^{15}VG/耳(例如,1x 10^{7}VG/耳、2x 10^{7}VG/耳、3x 10^{7}VG/耳、4x 10^{7}VG/耳、5x 10^{7}VG/耳、6x 10^{7}VG/耳、7x 10^{7}VG/耳、8x 10^{7}VG/耳、9x 10^{7}VG/耳、1x 10^{8}VG/耳、2x10^{8}VG/耳、3x 10^{8}VG/耳、4x 10^{8}VG/耳、5x 10^{8}VG/耳、6x 10^{8}VG/耳、7x 10^{8}VG/耳、8x 10^{8}VG/耳、9x 10^{8}VG/耳、1x 10^{9}VG/耳、2x10^{9}VG/耳、3x 10^{9}VG/耳、4x 10^{9}VG/耳、5x 10^{9}VG/耳、6x 10^{9}VG/耳、7x 10^{9}VG/耳、8x 10^{9}VG/耳、9x 10^{9}VG/耳、1x 10^{10}VG/耳、2x10^{10}VG/耳、3x 10^{10}VG/耳、4x 10^{10}VG/耳、5x 10^{10}VG/耳、6x 10^{10}VG/耳、7x 10^{10}VG/耳、8x 10^{10}VG/耳、9x 10^{10}VG/耳、1x 10^{11}VG/耳、2x 10^{11}VG/耳、3x 10^{11}VG/耳、4x 10^{11}VG/耳、5x 10^{11}VG/耳、6x 10^{11}VG/耳、7x 10^{11}VG/耳、8x 10^{11}VG/耳、9x 10^{11}VG/耳、1x 10^{12}VG/耳、2x 10^{12}VG/耳、3x 10^{12}VG/耳、4x 10^{12}VG/耳、5x 10^{12}VG/耳、6x 10^{12}VG/耳、7x 10^{12}VG/耳、8x 10^{12}VG/耳、9x 10^{12}VG/耳、1x 10^{13}VG/耳、2x 10^{13}VG/耳、3x 10^{13}VG/耳、4x 10^{13}VG/耳、5x 10^{13}VG/耳、6x 10^{13}VG/耳、7x 10^{13}VG/耳、8x 10^{13}VG/耳、9x 10^{13}VG/耳、1x 10^{14}VG/耳、2x 10^{14}VG/耳、3x 10^{14}VG/耳、4x 10^{14}VG/耳、5x 10^{14}VG/耳、6x 10^{14}VG/耳、7x 10^{14}VG/耳、8x 10^{14}VG/耳、9x 10^{14}VG/耳、1x 10^{15}VG/耳或2x 10^{15}VG/耳)。

[0163] 本文所述的组合物是以足以改进前庭功能(例如改进平衡或减少头晕或眩晕)、治疗双侧前庭病、治疗振动幻视、治疗平衡障碍、增加由可操作地连接至SLC6A14启动子的转基因编码的蛋白质的表达、增加由可操作地连接至SLC6A14启动子的转基因编码的蛋白质的功能、促进或增加毛细胞发育、增加毛细胞数(例如促进或诱导毛细胞再生或增殖)、增加或诱导毛细胞成熟(例如再生毛细胞的成熟)、改进毛细胞功能、改进VSC功能、促进或增加VSC和/或前庭毛细胞存活和/或促进或增加VSC增殖的量施用。前庭功能可使用标准平衡和眩晕测试(例如眼动测试(例如ENG或VNG)、VOR测试(例如头部冲击测试(哈-克测试,例如VHIT)或热量反射测试)、体位图、旋转椅测试、ECOG、VEMP和专门临床平衡测试)来评估,并且与在治疗之前获得的测量相比可改进5%或更大(例如5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%或更大)。本文所述的组合物也可以足以减缓或防止前庭功能障碍的发生或进展的量来施用(例如在携带与前庭功能障碍相关的遗传突变、具有前庭功能障碍家族史(例如遗传性前庭功能障碍)或已暴露于与前庭功能障碍相关的风险因子(例如耳毒性药物、头部创伤或疾病或感染)但未表现出前庭功能障碍(例如眩晕、头晕或失衡)的受试者中,或在表现出轻度至中度前庭功能障碍的受试者中)。在施用于受试者的核酸载体中由可操作地连接至SLC6A14启动子的转基因编码的蛋白

质的表达可使用免疫组织化学、蛋白质印迹分析、定量实时PCR或本领域中已知用于检测蛋白质或mRNA的其他方法来评估,并且与施用本文所述的组合物之前的表达相比可增加5%或更大(例如5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%或更大)。毛细胞数、毛细胞功能、毛细胞成熟、毛细胞再生或由施用于受试者的核酸载体编码的蛋白质的功能可间接基于前庭功能的测试来评估,并且与施用本文所述的组合物之前的毛细胞数、毛细胞功能、毛细胞成熟、毛细胞再生或蛋白质功能相比或与未经治疗的受试者相比,可增加5%或更大(例如5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%或更大)。这些效应可在例如施用本文所述的组合物之后1周、2周、3周、4周、5周、6周、7周、8周、9周、10周、15周、20周、25周或更长时间内出现。可在施用组合物之后1个月、2个月、3个月、4个月、5个月、6个月或更长时间来评估患者,这取决于用于治疗的剂量和施用途径。根据评估的结果,患者可接受额外治疗。

[0164] 试剂盒

[0165] 本文所述的组合物可于用于治疗前庭功能障碍的试剂盒中提供。组合物可包括本文所述的SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)、含有此类多核苷酸的核酸载体和含有可操作地连接至编码目标蛋白质(例如可在VSC中表达以治疗前庭功能障碍的蛋白质)的转基因的本文所述多核苷酸的核酸载体。核酸载体可包装于AAV病毒衣壳(例如AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、Anc80、7m8、PHP.B、PHP.eB或PHP.S)中。所述试剂盒还可包括指导试剂盒的使用者(例如医师)实施本文所述方法的包装插页。所述试剂盒可任选地包括注射器或用于施用组合物的其他器件。

[0166] 实施例

[0167] 提出以下实施例以向本领域的普通技术人员提供可如何使用、制造和评估本文所述的组合物和方法的描述,并且旨在纯粹示例本发明,而不旨在限制本发明人视为其发明的范围。

[0168] 实施例1.体内鼠类前庭器官中SLC6A14v2和SLC6A14v3启动子活性的测定

[0169] 为了比较SLC6A14v2和SLC6A14v3启动子的体内活性,将驱动核导向的H2B-GFP融合蛋白的人类SLC6A14启动子(SEQ ID NO:3)(来自质粒P530;图3)和驱动核导向的H2B-GFP融合蛋白的鼠类SLC6A14启动子(SEQ ID NO:1)(来自质粒P919(SEQ ID NO:2);图4)分别包装到AAV8中,并通过以2.0×10^{10}vg/耳的剂量注射到雄性八周龄C57BL/6小鼠的后半规管中来递送1μL病毒(n=6只小鼠/病毒)。两周后,随后利用CO_2对动物进行安乐死并且依次灌注PBS和中性缓冲福尔马林(NBF)。提取颞骨,对椭圆囊进行显微解剖,并对毛细胞核标志物Pou4f3(1:200,sc-1980,Santa Cruz Biotechnology,Dallas,Texas,USA)和支持细胞核标志物Sox2(1:200,AF2018,R&DSystems,Inc.,Minneapolis,Minnesota,USA)进行荧光免疫标记。

[0170] 将器官整个安装在载玻片上并在Zeiss LSM 800共焦显微镜上成像(图1A至图1C)。对于每个椭圆囊,使用20x/0.8NA物镜收集z堆栈的共焦图像,该物镜具有足够的视野来捕获整个椭圆囊。每个z堆栈跨越毛细胞核层、支持细胞核层和间充质层,其中z厚度设置为Nyquist标准。图1A-1C中的图像显示了z堆栈内的单个z平面在指示深度处的缩放区域。

核GFP表达显示在感觉上皮内的支持细胞核层(图1A)、感觉上皮内的毛细胞核层(图1B),以及感觉上皮下方的间充质层(图1C)。在两个启动子的支持细胞核层中检测到了相当水平的核GFP表达。毛细胞核层和间充质中的核GFP表达显著减少;然而,与SLC6A14v3启动子(图1C,底行)相比,在具有SLC6A14v2启动子的间充质中可以看到更多的标记(图1C,顶行)。

[0171] 使用Imaris软件中的3D计数自动算法对表达GFP的细胞核进行定量测量,该软件确定整个椭圆囊内GFP阳性细胞核的数量、GFP荧光的强度以及Pou4f3和/或Sox2免疫标记的共阳性。在GraphPad Prism中进行统计分析。

[0172] SLC6A14v2和SLC6A14v3启动子之间具有可检测水平的核GFP的支持细胞的百分比相当(图2A;箱形图上的点代表单个椭圆囊)。然而,与SLC6A14v2相比,SLC6A14v3启动子在可检测水平高于背景的支持细胞中核GFP的平均强度显著更大(图2B;散点图上的圆圈代表所有样品中的单个细胞,黑线是群体平均值;$p<0.0001$,Student t检验)。此外,与SLC6A14v3相比,SLC6A14v2启动子的可检测到的GFP水平高于背景的毛细胞核的百分比显著更高(图2C;$p=0.001$,Studentt检验),如通过阳性Sox2免疫标记确定的非支持细胞的所有GFP+细胞核的百分比也是如此(图2D;$p=0.022$;Studentt检验)。

[0173] 实施例2.向患有前庭功能障碍的受试者施用含有含SLC6A14启动子的核酸载体的组合物

[0174] 根据本文所公开的方法,本领域技术的医师可治疗患有前庭功能障碍的患者(例如人类患者)以改进或恢复前庭功能。为此,本领域技术的医师可向人类患者施用含有AAV载体(例如AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、Anc80、7m8、PHP.B、PHP.eB或PHP.S)的组合物,所述AAV载体含有可操作地连接至编码治疗蛋白(例如无调BHLH转录因子1(Atoh1))的转基因的SLC6A14启动子(例如与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸)。在一个实施例中,载体具有AAV8衣壳并且含有SEQ ID NO:10的核苷酸219-3831。含有AAV载体的组合物可例如通过局部施用于内耳(例如注射至半规管中)来施用于患者,以治疗前庭功能障碍。

[0175] 在将组合物施用于患者后,本领域技术的专业人员可通过多种方法监测由转基因编码的治疗蛋白的表达和患者响应于疗法的改善。例如,医师可通过实施标准测试(例如眼震电图、视频眼震图、VOR测试(例如头部冲击测试(哈-克测试,例如VHIT)或热量反射测试)、旋转测试、前庭引发的肌源性电势或计算机化动态姿势仪)来监测患者的前庭功能。与施用组合物之前获得的测试结果相比,患者在施用组合物后在一个或多个测试中表现出改进的前庭功能指示患者对治疗的反应良好。可按需要确定和施用后续剂量。

[0176] 本发明的示例性实施方案阐述于以下所列举段落中。

[0177] E1.一种核酸载体,所述核酸载体包含与SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)的多核苷酸。

[0178] E2.如E1中任一项所述的核酸载体,其中所述多核苷酸可操作地连接至转基因。

[0179] E3.如E2所述的核酸载体,其中所述转基因是异源转基因。

[0180] E4.如E2或E3所述的核酸载体,其中所述转基因编码蛋白质、短发夹RNA(shRNA)、反义寡核苷酸(ASO)、核酸酶,或微小RNA。

[0181] E5.如E4所述的核酸载体,其中所述多核苷酸能够引导蛋白质、shRNA、ASO、核酸酶或微小RNA在哺乳动物VSC中的前庭支持细胞(VSC)特异性表达。

[0182] E6.如E5所述的核酸载体,其中所述VSC是人类VSC。

[0183] E7.如E4-E6中任一项所述的核酸载体,其中所述蛋白质是婆罗双树样转录因子2(Sall2)、钙调蛋白结合转录激活因子1(Camta1)、具有YRPW基序2的Hes相关家族BHLH转录因子(Hey2)、Gata结合蛋白2(Gata2)、具有YRPW基序1的Hes相关家族BHLH转录因子(Hey1)、神经酰胺合酶2(Lass2)、SRY盒10(Sox10)、GATA结合蛋白3(Gata3)、Cut样同源异形盒1(Cux1)、核受体亚家族2F组成员(Nr2f1)、Hes相关家族BHLH转录因子(Hes1)、RAR相关孤儿受体B(Rorb)、Jun原癌基因AP-1转录因子亚基(Jun)、锌指蛋白667(Zfp667)、LIM同源异形盒3(Lhx3)、无义螺旋-环-螺旋1(Nhlh1)、MAX二聚化蛋白4(Mxd4)、MIZ-1型锌指(Zmiz1)、髓磷脂转录因子1(Myt1)、信号转导子和转录激活因子3(Stat3)、BarH样同源异形盒1(Barhl1)、胸腺细胞选择相关的高迁移率族盒(Tox)、Prospero同源异形盒1(Prox1)、核因子IA(Nfia)、甲状腺激素受体β(Thrb)、MYCL原癌基因BHLH转录因子(Mycl1)、赖氨酸去甲基酶5A(Kdm5a)、CAMP反应元件结合蛋白3样4(Creb3I4)、ETS变体1(Etv1)、父系表达的3(Peg3)、BTB结构域和CNC同源物2(Bach2)、ISL LIM同源异形盒1(Isl1)、锌指和含BTB结构域的38(Zbtb38)、肢芽和心脏发育(Lbh)、Tubby二分转录因子(Tub)、泛素C(Hmg20)、RE1沉默转录因子(Rest)、锌指蛋白827(Zfp827)、AF4/FMR2家族成员3(Aff3)、PBX/有节的1同源异形盒2(Pknox2)、富含AT的相互作用结构域3B(Arid3b)、MLX相互作用蛋白(Mlxip)、锌指蛋白(Zfp532)、IKAROS家族锌指2(Ikzf2)、婆罗双树样转录因子1(Sall1)、SIX同源异形盒2(Six2)、婆罗双树样转录因子3(Sall3)、Lin-28同源物B(Lin28b)、调节因子X7(Rfx7)、脑源性神经营养因子(Bdnf)、生长因子非依赖性1转录阻遏子(Gfi1)、POU 4类同源异形盒3(Pou4f3)、MYC原癌基因BHLH转录因子(Myc)、β-连环蛋白(Ctnnb1)、SRY盒2(Sox2)、SRY盒4(Sox4)、SRY盒11(Sox11)、TEA结构域转录因子2(Tead2)、无调BHLH转录因子1(Atoh1)或Atoh1变体。

[0184] E8.如E7所述的核酸载体,其中所述蛋白质是Atoh1。

[0185] E9.如E7所述的核酸载体,其中所述Atoh1变体具有一个或多个选自由以下组成的组的氨基酸取代:S328A、S331A、S334A、S328A/S331A、S328A/S334A、S331A/S334A和S328A/S331A/S334。

[0186] E10.如E2-E9中任一项所述的核酸载体,其中所述核酸载体另外包含所述多核苷酸的5'的第一反向末端重复;和所述转基因的3'并且以5'至3'顺序,任选的转录后调节元件、多聚腺苷酸化信号和第二反向末端重复。

[0187] E11.如E10所述的核酸载体,所述核酸载体包含SEQ ID NO:10的核苷酸219-3831,SEQ ID NO:10的核苷酸219-3831的5'的第一反向末端重复,其中所述5'反向末端重复与SEQ ID NO:10的核苷酸1-130具有至少80％序列同一性;和SEQ ID NO:10的核苷酸219-3831的3'的第二反向末端重复,其中所述3'反向末端重复与SEQ ID NO:10的核苷酸3919-4048具有至少80％的序列同一性。

[0188] E12.如E10所述的核酸载体,所述核酸载体包含SEQ ID NO:11的核苷酸219-3822,SEQ ID NO:10的核苷酸219-3822的5'的第一反向末端重复,其中所述5'反向末端重复与SEQ ID NO:11的核苷酸1-130具有至少80％序列同一性;和SEQ ID NO:11的核苷酸219-

3822的3'的第二反向末端重复,其中所述3'反向末端重复与SEQ ID NO:11的核苷酸3919-4039具有至少80%的序列同一性。

[0189] E13.如E1-E12中任一项所述的核酸载体,其中所述核酸载体是病毒载体、质粒、粘粒或人工染色体。

[0190] E14.如E13所述的核酸载体,其中所述核酸载体是选自由以下组成的组的病毒载体:腺相关病毒(AAV)、腺病毒和慢病毒。

[0191] E15.如E14所述的核酸载体,其中所述病毒载体是AAV载体。

[0192] E16.如E15所述的核酸载体,其中所述AAV载体具有AAV1、AAV2、AAV2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、rh10、rh39、rh43、rh74、Anc80、Anc80L65、DJ/8、DJ/9、7m8、PHP.B、PHP.eB或PHP.S衣壳。

[0193] E17.一种组合物,所述组合物包含如E1-E16中任一项所述的核酸载体。

[0194] E18.如E17所述的组合物,所述组合物还包含药学上可接受的载体、稀释剂或赋形剂。

[0195] E19.一种多核苷酸,所述多核苷酸与可操作地连接至转基因的SEQ ID NO:1具有至少85%序列同一性(例如85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更大的序列同一性)。

[0196] E20.如E19所述的多核苷酸,其中所述转基因是异源转基因。

[0197] E21.如E20所述的多核苷酸,其中所述转基因编码蛋白质、shRNA、ASO、核酸酶,或微小RNA。

[0198] E22.如E21所述的多核苷酸,其中所述蛋白质是Sox9、Sall2、Camta1、Hey2、Gata2、Hey1、Lass2、Sox10、Gata3、Cux1、Nr2f1、Hes1、Rorb、Jun、Zfp667、Lhx3、Nhlh1、Mxd4、Zmiz1、Myt1、Stat3、Barhl1、Tox、Prox1、Nfia、Thrb、Mycl1、Kdm5a、Creb314、Etv1、Peg3、Bach2、Isl1、Zbtb38、Lbh、Tub、Hmg20、Rest、Zfp827、Aff3、Pknox2、Arid3b、Mlxip、Zfp532、Ikzf2、Sall1、Six2、Sall3、Lin28b、Rfx7、Bdnf、Gfi1、Pou4f3、Myc、Ctnnb1、Sox2、Sox4、Sox11、Tead2、Atoh1或Atoh1变体。

[0199] E23.如E22所述的多核苷酸,其中所述蛋白质是Atoh1。

[0200] E24.一种细胞,所述细胞包含如E19-E23中任一项所述的多核苷酸或如E1-E16中任一项所述的核酸载体。

[0201] E25.如E24所述的细胞,其中所述细胞是哺乳动物VSC。

[0202] E26.如E25所述的细胞,其中所述哺乳动物VSC是人类VSC。

[0203] E27.一种在哺乳动物VSC中表达转基因的方法,所述方法包括使所述哺乳动物VSC与如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物接触。

[0204] E28.如E27所述的方法,其中所述转基因在VSC中特异性表达。

[0205] E29.如E27或E28所述的方法,其中所述哺乳动物VSC是人类VSC。

[0206] E30.一种治疗患有前庭功能障碍或具有患上前庭功能障碍的风险的受试者的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0207] E31.如E30所述的方法,其中所述前庭功能障碍包括眩晕、头晕、失衡、双侧前庭病(也称为双侧前庭功能低下)、振动幻视或平衡障碍。

[0208] E32.如E30或E31所述的方法,其中所述前庭功能障碍是年龄相关的前庭功能障碍、头部创伤相关的前庭功能障碍、疾病或感染相关的前庭功能障碍或耳毒性药物诱导的前庭功能障碍。

[0209] E33.如E30-E32中任一项所述的方法,其中所述前庭功能障碍与遗传突变相关。

[0210] E34.如E30或E31所述的方法,其中所述前庭功能障碍是特发性前庭功能障碍。

[0211] E35.一种诱导或增加有需要的受试者的前庭毛细胞再生的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0212] E36.一种诱导或增加有需要的受试者的VSC增殖的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0213] E37.一种诱导或增加有需要的受试者的前庭毛细胞增殖的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0214] E38.一种诱导或增加有需要的受试者的前庭毛细胞成熟(例如再生毛细胞的成熟)的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0215] E39.一种诱导或增加有需要的受试者的前庭毛细胞神经支配的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0216] E40.一种增加有需要的受试者的VSC和/或前庭毛细胞存活的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0217] E41.如E35-E40中任一项所述的方法,其中所述受试者患有前庭功能障碍(例如眩晕、头晕、失衡、双侧前庭病(双侧前庭功能低下)、振动幻视或平衡障碍)或具有患上所述疾病的风险。

[0218] E42.一种治疗患有双侧前庭病(也称为双侧前庭功能低下)或具有患上双侧前庭病(也称为双侧前庭功能低下)的风险的受试者的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0219] E43.一种治疗患有振动幻视或具有患上振动幻视的风险的受试者的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0220] E44.如E42或E43所述的方法,其中所述双侧前庭病或所述振动幻视是耳毒性药物诱导的双侧前庭病或耳毒性药物诱导的振动幻视。

[0221] E45.如E32或E44所述的方法,其中所述耳毒性药物选自由以下组成的组:氨基糖苷、抗肿瘤药物、依他尼酸、呋塞米、水杨酸盐和奎宁。

[0222] E46.一种治疗患有平衡障碍或具有患上平衡障碍的风险的受试者的方法,所述方法包括向所述受试者施用有效量的如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0223] E47.如E30-E46中任一项所述的方法,其中所述方法还包括在施用所述核酸载体或组合物之前评估所述受试者的前庭功能。

[0224] E48.如E30-E47中任一项所述的方法,其中所述方法还包括在施用所述核酸载体

或组合物之后评估所述受试者的前庭功能。

[0225] E49.如E30-E48中任一项所述的方法,其中所述核酸载体或组合物是经局部施用的。

[0226] E50.如E49所述的方法,其中所述核酸载体或组合物被施用于半规管。

[0227] E51.如E49所述的方法,其中所述核酸载体或组合物是经鼓室或鼓室内施用的。

[0228] E52.如E49所述的方法,其中所述核酸载体或组合物被施用于外淋巴中。

[0229] E53.如E49所述的方法,其中所述核酸载体或组合物被施用于内淋巴中。

[0230] E54.如E49所述的方法,其中所述核酸载体或组合物被施用于卵圆窗或经由卵圆窗施用。

[0231] E55.如E49所述的方法,其中所述核酸载体或组合物被施用于圆窗或经由圆窗施用。

[0232] E56.如E30-E55中任一项所述的方法,其中所述核酸载体或组合物是以足以预防或减轻前庭功能障碍、延迟前庭功能障碍的发生、减缓前庭功能障碍的进展、改进前庭功能、增加前庭毛细胞数、增加前庭毛细胞成熟(例如再生毛细胞的成熟)、增加前庭毛细胞增殖、增加前庭毛细胞再生、增加前庭毛细胞神经支配、增加VSC增殖、增加VSC数、增加VSC存活、增加前庭毛细胞存活或改进VSC功能的量施用。

[0233] E57.如E30-E56中任一项所述的方法,其中所述受试者是人类。

[0234] E58.一种试剂盒,所述试剂盒包含如E1-E16中任一项所述的核酸载体或如E17或E18所述的组合物。

[0235] 其他实施方案

[0236] 在不脱离本发明的范围和精神的情况下,所描述的发明的各种修改和变化对于本领域技术人员来说将是显而易见的。尽管已结合具体实施方案描述了本发明,但应理解,所要求保护的本发明不应不适当地限于这些具体实施方案。实际上,对本领域技术人员来说显而易见的用于实施本发明的所述模式的各种修改旨在落入本发明的范围内。其他实施方案在权利要求中。

序列表

<110> 分贝治疗公司(Decibel Therapeutics, Inc.)

<120> 前庭支持细胞启动子和其用途

<130> 51471-010WO2

<150> US 63/184,015

<151> 2021-05-04

<160> 11

<170> PatentIn version 3.5

<210> 1

<211> 1759

<212> DNA

<213> 小家鼠(Mus musculus)

<400> 1

atacacttat gtatatgtgc gatgtcagtg tgtgtgcata taaagtccca aacaagcctg 60
tatgatattg accaacaagg tcaaggcaaa gttttgatac tttcaggtca caacctctcc 120
gcatccctct ctactttgct ctatctgcct gaactcctga ggacatgttt ctactgcaaa 180
tggaaaatcc ttgtcagcca gtgaggaaca aagggactat acatagatga aaacttggct 240
ctctgctggt tcctttgttt gtatgaattt atacaatttg gtaaaactgc caccatgtct 300
tacatggaca gattgagtgt agattctttg aatttttgat gaagaggcgc tgcactggtg 360
atcggaattg cagtctttcc tctgtaggta acctggcttg tttccttaca gtttactttc 420
taggcctcgc ctttctcaca gagtgaagtc ctttgttaag gttcgaattt cccataaacc 480
tgctcaataa tttgtttgtg tttggcttct ttgaaatact acacaaagca atccttgtaa 540
aaggcaaaac tattccgaag gctgagaaag gagctccagg acatagattc aaagtcgctc 600
ttttcaggta gagacagctg ggtaatctta tcttaactgg ctacatttca aggttcccaa 660
ttcaggggct ttcccctctg ggagcagcat tctctccggg tgatgaagag ctttctagtg 720
aggagcaaaa ctttcagaaa accggagggc ccagagcagt ctggtctgtt cacaaaaatt 780
atagcaaaca aaataagccc ggcggattgg gtctctccta cctccagcac caggggagat 840
cagcacttgg ccccaggaca gagacctgag aagtgaggtt tggaagaagc caggaatcca 900
ggaaaggagg caagattgct aaggcaccgg cacagctctg agtcaaaagt tgtcagtctt 960
ctttggctct ggctgcggag ctcaattgct cacagccctg ccctttccta gggctggggc 1020
aaggaattgc tacattcagg attacctggg ggaaaaacca gaggcttgct ttggtccctt 1080
ccggtaattg aaaggactgg ccgtcagcga gggggaggag agagcttccc tccataaatg 1140
gtcccacccc tgggcaaggt ggctcacttt ggcaggtagc aaccggggag tgtgcacctg 1200
ccaccagtca agctcagcca gactgtgaga agaggagagg cgaggcacac caagggatcc 1260
agtgaaccaa cgacagattg aagtgcccga acttcttcaa gtgcagacag aaggaggtag 1320
ggttctggaa gtttctggtg gtgtagggga gtcaggaagg gaaaacaagg agggagagtg 1380
agtcttagtt ttttgctttc tgtagctgtt ccttattttg catatttctt tctcttcaac 1440
tcttttcaag tatgcctgat acgttgttct cacgaagttg acgtgaaaaa caactttcct 1500

gctggtagtt aggaaactta ggagcacctc aacctgtacc ttgagaacac ccagagaatg 1560
ctgctctttg tcgttctcta taccgtgttc atatgctgca gggaaatgca aagaatgtac 1620
tgtccttatc tgaccctggg agcattccat agtcaagcag cagctatcag gttgggaaag 1680
agctcctctc caaggtgtaa cagaaaagga aaatgttgat atttttcttg tttagaaagt 1740
gacagcttca tccgagaac 1759

<210> 2
<211> 6993
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 2

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg gaattcgccc ttaagctagc ggcgcgccat acacttatgt atatgtgcga 240
tgtcagtgtg tgtgcatata aagtcccaaa caagcctgta tgatattgac caacaaggtc 300
aaggcaaagt tttgatactt tcaggtcaca acctctccgc atccctctct actttgctct 360
atctgcctga actcctgagg acatgtttct actgcaaatg gaaaatcctt gtcagccagt 420
gaggaacaaa gggactatac atagatgaaa acttggctct ctgctggttc ctttgtttgt 480
atgaatttat acaatttggt aaaactgcca ccatgtctta catggacaga ttgagtgtag 540
attctttgaa tttttgatga agaggcgctg cactggtgat cggaattgca gtctttcctc 600
tgtaggtaac ctggcttgtt tccttacagt ttactttcta ggcctcgcct ttctcacaga 660
gtgaagtcct ttgttaaggt tcgaatttcc cataaacctg ctcaataatt tgtttgtgtt 720
tggcttcttt gaaatactac acaaagcaat ccttgtaaaa ggcaaaacta ttccgaaggc 780
tgagaaagga gctccaggac atagattcaa agtcgctctt ttcaggtaga gacagctggg 840
taatcttatc ttaactggct acatttcaag gttcccaatt caggggcttt cccctctggg 900
agcagcattc tctccgggtg atgaagagct ttctagtgag gagcaaaact ttcagaaaac 960
cggagggccc agagcagtct ggtctgttca caaaaattat agcaaacaaa ataagcccgg 1020
cggattgggt ctctcctacc tccagcacca ggggagatca gcacttggcc ccaggacaga 1080
gacctgagaa gtgaggtttg gaagaagcca ggaatccagg aaaggaggca agattgctaa 1140
ggcaccggca cagctctgag tcaaaagttg tcagtcttct ttggctctgg ctgcggagct 1200
caattgctca cagccctgcc ctttcctagg gctggggcaa ggaattgcta cattcaggat 1260
tacctggggg aaaaaccaga ggcttgcttt ggtcccttcc ggtaattgaa aggactggcc 1320
gtcagcgagg gggaggagag agcttccctc cataaatggt cccacccctg ggcaaggtgg 1380
ctcactttgg caggtagcaa ccggggagtg tgcacctgcc accagtcaag ctcagccaga 1440
ctgtgagaag aggagaggcg aggcacacca agggatccag tgaaccaacg acagattgaa 1500
gtgcccgaac ttcttcaagt gcagacagaa ggaggtaggg ttctggaagt ttctggtggt 1560
gtaggggagt caggaaggga aaacaaggag ggagagtgag tcttagtttt ttgctttctg 1620

tagctgttcc ttattttgca tatttctttc tcttcaactc ttttcaagta tgcctgatac 1680
gttgttctca cgaagttgac gtgaaaaaca actttcctgc tggtagttag gaaacttagg 1740
agcacctcaa cctgtacctt gagaacaccc agagaatgct gctctttgtc gttctctata 1800
ccgtgttcat atgctgcagg gaaatgcaaa gaatgtactg tccttatctg accctgggag 1860
cattccatag tcaagcagca gctatcaggt tgggaaagag ctcctctcca aggtgtaaca 1920
gaaaaggaaa atgttgatat ttttcttgtt tagaaagtga cagcttcatc cgagaacgcg 1980
gccgcgccac catgccagag ccagcgaagt ctgctcccgc cccgaaaaag ggctccaaga 2040
aggcggtgac taaggcgcag aagaaaggcg gcaagaagcg caagcgcagc cgcaaggaga 2100
gctattccat ctatgtgtac aaggttctga agcaggtcca ccctgacacc ggcatttcgt 2160
ccaaggccat gggcatcatg aattcgtttg tgaacgacat tttcgagcgc atcgcaggtg 2220
aggcttcccg cctggcgcat tacaacaagc gctcgaccat cacctccagg gagatccaga 2280
cggccgtgcg cctgctgctg cctggggagt tggccaagca cgccgtgtcc gagggtacta 2340
aggccatcac caagtacacc agcgctaagg atccaccggt cgccaccatg gtgagcaagg 2400
gcgaggagct gttcaccggg gtggtgccca tcctggtcga gctggacggc gacgtaaacg 2460
gccacaagtt cagcgtgtcc ggcgagggcg agggcgatgc cacctacggc aagctgaccc 2520
tgaagttcat ctgcaccacc ggcaagctgc ccgtgccctg gcccaccctc gtgaccaccc 2580
tgacctacgg cgtgcagtgc ttcagccgct accccgacca catgaagcag cacgacttct 2640
tcaagtccgc catgcccgaa ggctacgtcc aggagcgcac catcttcttc aaggacgacg 2700
gcaactacaa gacccgcgcc gaggtgaagt tcgagggcga caccctggtg aaccgcatcg 2760
agctgaaggg catcgacttc aaggaggacg gcaacatcct ggggcacaag ctggagtaca 2820
actacaacag ccacaacgtc tatatcatgg ccgacaagca gaagaacggc atcaaggtga 2880
acttcaagat ccgccacaac atcgaggacg gcagcgtgca gctcgccgac cactaccagc 2940
agaacacccc catcggcgac ggccccgtgc tgctgcccga caaccactac ctgagcaccc 3000
agtccgccct gagcaaagac cccaacgaga agcgcgatca catggtcctg ctggagttcg 3060
tgaccgccgc cgggatcact ctcggcatgg acgagctgta caagtaataa gcttggatcc 3120
aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct 3180
ccttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat tgcttcccgt 3240
atggctttca ttttctcctc cttgtataaa tcctggttgc tgtctcttta tgaggagttg 3300
tggcccgttg tcaggcaacg tggcgtggtg tgcactgtgt ttgctgacgc aacccccact 3360
ggttggggca ttgccaccac ctgtcagctc ctttccggga ctttcgcttt ccccctccct 3420
attgccacgg cggaactcat cgccgcctgc cttgcccgct gctggacagg ggctcggctg 3480
ttgggcactg acaattccgt ggtgttgtcg gggaaatcat cgtcctttcc ttggctgctc 3540
gcctgtgttg ccacctggat tctgcgcggg acgtccttct gctacgtccc ttcggccctc 3600
aatccagcgg accttccttc ccgcggcctg ctgccggctc tgcggcctct tccgcgtctt 3660
cgaacaattg catcggacac atcttggcgt tttacaacgt cgtgactggg aaaaccctgg 3720
cgttacccaa cttaagatct gcctcgactg tgccttctag ttgccagcca tctgttgttt 3780
gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat 3840
aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg 3900
tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggactcga 3960

gttaagggcg aattcccgat aaggatcttc ctagagcatg gctacgtaga taagtagcat 4020
ggcgggttaa tcattaacta caaggaaccc ctagtgatgg agttggccac tccctctctg 4080
cgcgctcgct cgctcactga ggccgggcga ccaaaggtcg cccgacgccc gggctttgcc 4140
cgggcggcct cagtgagcga gcgagcgcgc agccttaatt aacctaattc actggccgtc 4200
gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca 4260
catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa 4320
cagttgcgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt aagcgcggcg 4380
ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct 4440
ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat 4500
cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt 4560
gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg 4620
acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac 4680
cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta 4740
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca 4800
atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta ttttctaaa 4860
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 4920
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg 4980
cattttgcct tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag 5040
atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg 5100
agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 5160
gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt 5220
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga 5280
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac 5340
ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc 5400
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 5460
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac 5520
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag 5580
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg 5640
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta 5700
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 5760
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 5820
tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt 5880
ttgataatct catgaccaaa atcccttaac gtgagtttc gttccactga gcgtcagacc 5940
ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct 6000
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa 6060
ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gttcttctag 6120
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 6180
tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg 6240
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 6300

cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat 6360
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg 6420
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc 6480
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 6540
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc 6600
cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg 6660
cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga 6720
gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc 6780
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 6840
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc 6900
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 6960
attacgccag atttaattaa ggccttaatt agg 6993

<210> 3
<211> 834
<212> DNA
<213> 智人(Homo sapiens)
<400> 3

aagctgggat gtttcctcat agtttacttt ctaggcctca tctttcttac agagtgtgct 60
cctttgttaa ggttagaatt tcccataaac ctgctcaata atttgtttgt gtttggcttc 120
tttgaaatac tacacaaagc aatccctgta aaaggcaaag ctgtcctgaa ggctgagaaa 180
ggagcctgag acataggctc caagttgctc ttttcaggca gagccagctg ggtaatctta 240
tctcagatgg ctgcttttca aggtgcccaa ttcaggggct tttcctctgg gagcagcatt 300
tgccccaggg aatcaagtgc tttctagtca ggggcaaaac tttgggaaat ctgaggaccc 360
agggtggtat ggtctgttca ggagaatttt ggggaacaga atggcccccc tctccctcca 420
gcacttgtac agatcagcac ttggccccag aacagagacc agactgagag gcgaggttag 480
gaggaaacag gggacccagg aaaggcggct agattgcaaa cgtacctaca cagctctgag 540
tcaaaggctg tcagtcatct cggctcagac tgctctgctc tccagcagcc cagccctttc 600
ccagggctgg ggcaggagat tgctacatgt aggcttatct ggggaaaaac cagagcctca 660
ctttagtccc ttccggtaat tgacactact ggacacccag gagggggagg agagagcttc 720
tcttcataaa tgttcccacc cctgggcaag gtggctcact ctggcaggta ggaacagggg 780
agagtgcacc tgctaccagt caagctcagc cagactgcaa gaggaggcga ggcg 834

<210> 4
<211> 354
<212> PRT
<213> 智人(Homo sapiens)
<400> 4

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1 5 10 15
Gly Asp His His Arg Gln Pro Gln Pro His His Leu Pro Gln Pro Pro

20 25 30

Pro Pro Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Glu His Pro Val
35 40 45

Tyr Pro Pro Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp
50 55 60

Leu Ala Pro Thr Leu Gln Gly Ile Cys Thr Ala Arg Ala Ala Gln Tyr
65 70 75 80

Leu Leu His Ser Pro Glu Leu Gly Ala Ser Glu Ala Ala Ala Pro Arg
85 90 95

Asp Glu Val Asp Gly Arg Gly Glu Leu Val Arg Arg Ser Ser Gly Gly
100 105 110

Ala Ser Ser Ser Lys Ser Pro Gly Pro Val Lys Val Arg Glu Gln Leu
115 120 125

Cys Lys Leu Lys Gly Gly Val Val Val Asp Glu Leu Gly Cys Ser Arg
130 135 140

Gln Arg Ala Pro Ser Ser Lys Gln Val Asn Gly Val Gln Lys Gln Arg
145 150 155 160

Arg Leu Ala Ala Asn Ala Arg Glu Arg Arg Arg Met His Gly Leu Asn
165 170 175

His Ala Phe Asp Gln Leu Arg Asn Val Ile Pro Ser Phe Asn Asn Asp
180 185 190

Lys Lys Leu Ser Lys Tyr Glu Thr Leu Gln Met Ala Gln Ile Tyr Ile
195 200 205

Asn Ala Leu Ser Glu Leu Leu Gln Thr Pro Ser Gly Gly Glu Gln Pro
210 215 220

Pro Pro Pro Pro Ala Ser Cys Lys Ser Asp His His His Leu Arg Thr
225 230 235 240

Ala Ala Ser Tyr Glu Gly Gly Ala Gly Asn Ala Thr Ala Ala Gly Ala
245 250 255

Gln Gln Ala Ser Gly Gly Ser Gln Arg Pro Thr Pro Pro Gly Ser Cys
260 265 270

Arg Thr Arg Phe Ser Ala Pro Ala Ser Ala Gly Gly Tyr Ser Val Gln
275 280 285

Leu Asp Ala Leu His Phe Ser Thr Phe Glu Asp Ser Ala Leu Thr Ala
290 295 300

Met Met Ala Gln Lys Asn Leu Ser Pro Ser Leu Pro Gly Ser Ile Leu
305 310 315 320

Gln Pro Val Gln Glu Glu Asn Ser Lys Thr Ser Pro Arg Ser His Arg
325 330 335

Ser Asp Gly Glu Phe Ser Pro His Ser His Tyr Ser Asp Ser Asp Glu
340 345 350

Ala Ser

<210> 5

<211> 1062

<212> DNA

<213> 智人(Homo sapiens)

<400> 5

atgtcccgcc tgctgcatgc agaagagtgg gctgaagtga aggagttggg agaccaccat 60
cgccagcccc agccgcatca tctcccgcaa ccgccgccgc cgccgcagcc acctgcaact 120
ttgcaggcga gagagcatcc cgtctacccg cctgagctgt ccctcctgga cagcaccgac 180
ccacgcgcct ggctggctcc cactttgcag ggcatctgca cggcacgcgc cgcccagtat 240
ttgctacatt ccccggagct gggtgcctca gaggccgctg cgccccggga cgaggtggac 300
ggccgggggg agctggtaag gaggagcagc ggcggtgcca gcagcagcaa gagccccggg 360
ccggtgaaag tgcgggaaca gctgtgcaag ctgaaaggcg gggtggtggt agacgagctg 420
ggctgcagcc gccaacgggc cccttccagc aaacaggtga atggggtgca gaagcagaga 480
cggctagcag ccaacgccag ggagcggcgc aggatgcatg ggctgaacca cgccttcgac 540
cagctgcgca atgttatccc gtcgttcaac aacgacaaga agctgtccaa atatgagacc 600
ctgcagatgg cccaaatcta catcaacgcc ttgtccgagc tgctacaaac gcccagcgga 660
ggggaacagc caccgccgcc tccagcctcc tgcaaaagcg accaccacca ccttcgcacc 720
gcggcctcct atgaaggggg cgcgggcaac gcgaccgcag ctggggctca gcaggcttcc 780
ggagggagcc agcggccgac cccgcccggg agttgccgga ctcgcttctc agccccagct 840
tctgcgggag ggtactcggt gcagctggac gctctgcact tctcgacttt cgaggacagc 900
gccctgacag cgatgatggc gcaaaagaat ttgtctcctt ctctccccgg gagcatcttg 960
cagccagtgc aggaggaaaa cagcaaaact tcgcctcggt cccacagaag cgacggggaa 1020
ttttcccccc attcccatta cagtgactcg gatgaggcaa gt 1062

<210> 6

<211> 351

<212> PRT

<213> 小家鼠(Mus musculus)

<400> 6

Met Ser Arg Leu Leu His Ala Glu Glu Trp Ala Glu Val Lys Glu Leu
1 5 10 15
Gly Asp His His Arg His Pro Gln Pro His His Val Pro Pro Leu Thr
20 25 30
Pro Gln Pro Pro Ala Thr Leu Gln Ala Arg Asp Leu Pro Val Tyr Pro
35 40 45
Ala Glu Leu Ser Leu Leu Asp Ser Thr Asp Pro Arg Ala Trp Leu Thr
50 55 60

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

<210> 7
<211> 1053
<212> DNA

<213> 小家鼠(Mus musculus)

<400> 7

atgtcccgcc tgctgcatgc agaagagtgg gctgaggtaa aagagttggg ggaccaccat 60
cgccatcccc agccgcacca cgtcccgccg ctgacgccac agccacctgc taccctgcag 120
gcgagagacc ttcccgtcta cccggcagaa ctgtccctcc tggatagcac cgacccacgc 180
gcctggctga ctcccacttt gcagggcctc tgcacggcac gcgccgccca gtatctgctg 240
cattctcccg agctgggtgc ctccgaggcc gcggcgcccc gggacgaggc tgacagccag 300
ggtgagctgg taaggagaag cggctgtggc ggcctcagca agagccccgg gcccgtcaaa 360
gtacgggaac agctgtgcaa gctgaagggt ggggttgtag tggacgagct tggctgcagc 420
cgccagcgag ccccttccag caaacaggtg aatggggtac agaagcaaag gaggctggca 480
gcaaacgcaa gggaacggcg caggatgcac gggctgaacc acgccttcga ccagctgcgc 540
aacgttatcc cgtccttcaa caacgacaag aagctgtcca aatatgagac cctacagatg 600
gcccagatct acatcaacgc tctgtcggag ttgctgcaga ctcccaatgt cggagagcaa 660
ccgccgccgc ccacagcttc ctgcaaaaat gaccaccatc accttcgcac cgcctcctcc 720
tatgaaggag gtgcgggcgc ctctgcggta gctggggctc agccagcccc gggagggggc 780
ccgagaccta ccccgcccgg gccttgccgg actcgcttct caggcccagc ttcctctggg 840
ggttactcgg tgcagctgga cgctttgcac ttcccagcct tcgaggacag ggccctaaca 900
gcgatgatgg cacagaagga cctgtcgcct tcgctgcccg ggggcatcct gcagcctgta 960
caggaggaca acagcaaaac atctcccaga tcccacagaa gtgacggaga gttttccccc 1020
cactctcatt acagtgactc tgatgaggcc agt 1053

<210> 8

<211> 548

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 8

gatccaatca acctctggat tacaaaattt gtgaaagatt gactggtatt cttaactatg 60
ttgctccttt tacgctatgt ggatacgctg ctttaatgcc tttgtatcat gctattgctt 120
cccgtatggc tttcattttc tcctccttgt ataaatcctg gttgctgtct ctttatgagg 180
agttgtggcc cgttgtcagg caacgtggcg tggtgtgcac tgtgtttgct gacgcaaccc 240
ccactggttg gggcattgcc accacctgtc agctcctttc cgggactttc gctttccccc 300
tccctattgc cacggcggaa ctcatcgccg cctgccttgc ccgctgctgg acaggggctc 360
ggctgttggg cactgacaat tccgtggtgt tgtcggggaa atcatcgtcc tttccttggc 420
tgctcgcctg tgttgccacc tggattctgc gcgggacgtc cttctgctac gtcccttcgg 480
ccctcaatcc agcggacctt ccttcccgcg gcctgctgcc ggctctgcgg cctcttccgc 540
gtcttcga 548

<210> 9

<211> 425

<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 9

aatcaacctc tggattacaa aatttgtgaa agattgactg gtattcttaa ctatgttgct 60
ccttttacgc tatgtggata cgctgcttta atgcctttgt atcatgctat tgcttcccgt 120
atggctttca ttttctcctc cttgtataaa tcctggttag ttcttgccac ggcggaactc 180
atcgccgcct gccttgcccg ctgctggaca ggggctcggc tgttgggcac tgacaattcc 240
gtggtgttta tttgtgaaat ttgtgatgct attgctttat ttgtaaccat ctagctttat 300
ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt 360
taacaacaac aattgcattc attttatgtt tcaggttcag ggggagatgt gggaggtttt 420
ttaaa 425

<210> 10
<211> 6818
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 10

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg gaattcgccc ttaagctagc ggcgcgccat acacttatgt atatgtgcga 240
tgtcagtgtg tgtgcatata aagtcccaaa caagcctgta tgatattgac caacaaggtc 300
aaggcaaagt tttgatactt tcaggtcaca acctctccgc atccctctct actttgctct 360
atctgcctga actcctgagg acatgtttct actgcaaatg gaaaatcctt gtcagccagt 420
gaggaacaaa gggactatac atagatgaaa acttggctct ctgctggttc ctttgtttgt 480
atgaatttat acaatttggt aaaactgcca ccatgtctta catggacaga ttgagtgtag 540
attctttgaa tttttgatga agaggcgctg cactggtgat cggaattgca gtctttcctc 600
tgtaggtaac ctggcttgtt tccttacagt ttactttcta ggcctcgcct ttctcacaga 660
gtgaagtcct ttgttaaggt tcgaatttcc cataaacctg ctcaataatt tgtttgtgtt 720
tggcttcttt gaaatactac acaaagcaat ccttgtaaaa ggcaaaacta ttccgaaggc 780
tgagaaagga gctccaggac atagattcaa agtcgctctt ttcaggtaga gacagctggg 840
taatcttatc ttaactggct acatttcaag gttcccaatt caggggcttt cccctctggg 900
agcagcattc tctccgggtg atgaagagct ttctagtgag gagcaaaact ttcagaaaac 960
cggagggccc agagcagtct ggtctgttca caaaaattat agcaaacaaa ataagcccgg 1020
cggattgggt ctctcctacc tccagcacca ggggagatca gcacttggcc ccaggacaga 1080
gacctgagaa gtgaggtttg gaagaagcca ggaatccagg aaaggaggca agattgctaa 1140

ggcaccggca cagctctgag tcaaaagttg tcagtcttct ttggctctgg ctgcggagct 1200
caattgctca cagccctgcc ctttcctagg gctggggcaa ggaattgcta cattcaggat 1260
tacctggggg aaaaaccaga ggcttgcttt ggtcccttcc ggtaattgaa aggactggcc 1320
gtcagcgagg gggaggagag agcttccctc cataaatggt cccacccctg ggcaaggtgg 1380
ctcactttgg caggtagcaa ccggggagtg tgcacctgcc accagtcaag ctcagccaga 1440
ctgtgagaag aggagaggcg aggcacacca agggatccag tgaaccaacg acagattgaa 1500
gtgcccgaac ttcttcaagt gcagacagaa ggaggtaggg ttctggaagt ttctggtggt 1560
gtaggggagt caggaaggga aaacaaggag ggagagtgag tcttagtttt ttgctttctg 1620
tagctgttcc ttattttgca tatttctttc tcttcaactc ttttcaagta tgcctgatac 1680
gttgttctca cgaagttgac gtgaaaaaca actttcctgc tggtagttag gaaacttagg 1740
agcacctcaa cctgtacctt gagaacaccc agagaatgct gctctttgtc gttctctata 1800
ccgtgttcat atgctgcagg gaaatgcaaa gaatgtactg tccttatctg accctgggag 1860
cattccatag tcaagcagca gctatcaggt tgggaaagag ctcctctcca aggtgtaaca 1920
gaaaaggaaa atgttgatat ttttcttgtt tagaaagtga cagcttcatc cgagaacgcg 1980
gccgcgccac catgtcccgc ctgctgcatg cagaagagtg ggctgaagtg aaggagttgg 2040
gagaccacca tcgccagccc cagccgcatc atctcccgca accgccgccg ccgccgcagc 2100
cacctgcaac tttgcaggcg agagagcatc ccgtctaccc gcctgagctg tccctcctgg 2160
acagcaccga cccacgcgcc tggctggctc ccactttgca gggcatctgc acggcacgcg 2220
ccgcccagta tttgctacat tccccggagc tgggtgcctc agaggccgct gcgccccggg 2280
acgaggtgga cggccggggg gagctggtaa ggaggagcag cggcggtgcc agcagcagca 2340
agagccccgg gccggtgaaa gtgcgggaac agctgtgcaa gctgaaaggc ggggtggtgg 2400
tagacgagct gggctgcagc cgccaacggg ccccttccag caaacaggtg aatggggtgc 2460
agaagcagag acggctagca gccaacgcca gggagcggcg caggatgcat gggctgaacc 2520
acgccttcga ccagctgcgc aatgttatcc cgtcgttcaa caacgacaag aagctgtcca 2580
aatatgagac cctgcagatg gcccaaatct acatcaacgc cttgtccgag ctgctacaaa 2640
cgcccagcgg aggggaacag ccaccgccgc ctccagcctc ctgcaaaagc gaccaccacc 2700
accttcgcac cgcggcctcc tatgaagggg gcgcgggcaa cgcgaccgca gctggggctc 2760
agcaggcttc cggagggagc cagcggccga ccccgcccgg gagttgccgg actcgcttct 2820
cagccccagc ttctgcggga gggtactcgg tgcagctgga cgctctgcac ttctcgactt 2880
tcgaggacag cgccctgaca gcgatgatgg cgcaaaagaa tttgtctcct tctctccccg 2940
ggagcatctt gcagccagtg caggaggaaa acagcaaaac ttcgcctcgg tcccacagaa 3000
gcgacgggga attttccccc cattcccatt acagtgactc ggatgaggca agttagaagc 3060
ttggatccaa tcaacctctg gattacaaaa tttgtgaaag attgactggt attcttaact 3120
atgttgctcc ttttacgcta tgtggatacg ctgctttaat gcctttgtat catgctattg 3180
cttcccgtat ggctttcatt ttctcctcct tgtataaatc ctggttgctg tctctttatg 3240
aggagttgtg gcccgttgtc aggcaacgtg gcgtggtgtg cactgtgttt gctgacgcaa 3300
cccccactgg ttggggcatt gccaccacct gtcagctcct ttccgggact ttcgctttcc 3360
ccctccctat tgccacggcg gaactcatcg ccgcctgcct tgcccgctgc tggacagggg 3420
ctcggctgtt gggcactgac aattccgtgg tgttgtcggg gaaatcatcg tcctttcctt 3480

ggctgctcgc ctgtgttgcc acctggattc tgcgcgggac gtccttctgc tacgtccctt 3540
cggccctcaa tccagcggac cttccttccc gcggcctgct gccggctctg cggcctcttc 3600
cgcgtcttcg agatctgcct cgactgtgcc ttctagttgc cagccatctg ttgtttgccc 3660
ctccccctgtg ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa 3720
tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg 3780
gcaggacagc aaggggggagg attgggaaga caatagcagg catgctgggg actcgagtta 3840
agggcgaatt cccgataagg atcttcctag agcatggcta cgtagataag tagcatggcg 3900
ggttaatcat taactacaag gaacccctag tgatggagtt ggccactccc tctctgcgcg 3960
ctcgctcgct cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg 4020
cggcctcagt gagcgagcga gcgcgcagcc ttaattaacc taattcactg gccgtcgttt 4080
tacaacgtcg tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc 4140
ccccttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt 4200
tgcgcagcct gaatggcgaa tgggacgcgc cctgtagcgg cgcattaagc gcggcgggtg 4260
tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg 4320
ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg 4380
ggctcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt 4440
agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt 4500
tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaacccta 4560
tctcggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa 4620
atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt 4680
aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 4740
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 4800
aaggaagagt atgagccata ttcaacggga aacgtcgagg ccgcgattaa attccaacat 4860
ggatgctgat ttatatgggt ataaatgggc tcgcgataat gtcgggcaat caggtgcgac 4920
aatctatcgc ttgtatggga agcccgatgc gccagagttg tttctgaaac atggcaaagg 4980
tagcgttgcc aatgatgtta cagatgagat ggtcagacta aactggctga cggaatttat 5040
gcctcttccg accatcaagc attttatccg tactcctgat gatgcatggt tactcaccac 5100
tgcgatcccc ggaaaaacag cattccaggt attagaagaa tatcctgatt caggtgaaaa 5160
tattgttgat gcgctggcag tgttcctgcg ccggttgcat tcgattcctg tttgtaattg 5220
tccttttaac agcgatcgcg tatttcgtct tgctcaggcg caatcacgaa tgaataacgg 5280
tttggttgat gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg aacaagtctg 5340
gaaagaaatg cataaacttt tgccattctc accggattca gtcgtcactc atggtgattt 5400
ctcacttgat aaccttattt ttgacgaggg gaaattaata ggttgtattg atgttggacg 5460
agtcggaatc gcagaccgat accaggatct tgccatccta tggaactgcc tcggtgagtt 5520
ttctccttca ttacagaaac ggctttttca aaaatatggt attgataatc ctgatatgaa 5580
taaattgcag tttcatttga tgctcgatga gtttttctaa ctgtcagacc aagtttactc 5640
atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat 5700
cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc 5760
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg 5820

ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct 5880
accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgttct 5940
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct 6000
cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg 6060
gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc 6120
gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 6180
gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg 6240
cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta 6300
tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 6360
ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg 6420
ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat 6480
taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc 6540
agtgagcgag gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc 6600
gattcattaa tgcagctggc acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa 6660
cgcaattaat gtgagttagc tcactcatta ggcaccccag gctttacact ttatgcttcc 6720
ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga 6780
ccatgattac gccagattta attaaggcct taattagg 6818

<210> 11
<211> 6809
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 11

ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctacgta gccatgctct 180
aggaagatcg gaattcgccc ttaagctagc ggcgcgccat acacttatgt atatgtgcga 240
tgtcagtgtg tgtgcatata aagtcccaaa caagcctgta tgatattgac caacaaggtc 300
aaggcaaagt tttgatactt tcaggtcaca acctctccgc atccctctct actttgctct 360
atctgcctga actcctgagg acatgtttct actgcaaatg gaaaatcctt gtcagccagt 420
gaggaacaaa gggactatac atagatgaaa acttggctct ctgctggttc ctttgtttgt 480
atgaatttat acaatttggt aaaactgcca ccatgtctta catggacaga ttgagtgtag 540
attctttgaa tttttgatga agaggcgctg cactggtgat cggaattgca gtctttcctc 600
tgtaggtaac ctggcttgtt tccttacagt ttactttcta ggcctcgcct ttctcacaga 660
gtgaagtcct ttgttaaggt tcgaatttcc cataaacctg ctcaataatt tgtttgtgtt 720
tggcttcttt gaaatactac acaaagcaat ccttgtaaaa ggcaaaacta ttccgaaggc 780
tgagaaagga gctccaggac atagattcaa agtcgctctt ttcaggtaga gacagctggg 840
taatcttatc ttaactggct acatttcaag gttcccaatt caggggcttt cccctctggg 900

agcagcattc tctccgggtg atgaagagct ttctagtgag gagcaaaact ttcagaaaac 960
cggagggccc agagcagtct ggtctgttca caaaaattat agcaaacaaa ataagcccgg 1020
cggattgggt ctctcctacc tccagcacca ggggagatca gcacttggcc ccaggacaga 1080
gacctgagaa gtgaggtttg gaagaagcca ggaatccagg aaaggaggca agattgctaa 1140
ggcaccggca cagctctgag tcaaaagttg tcagtcttct ttggctctgg ctgcggagct 1200
caattgctca cagccctgcc ctttcctagg gctggggcaa ggaattgcta cattcaggat 1260
tacctggggg aaaaaccaga ggcttgcttt ggtcccttcc ggtaattgaa aggactggcc 1320
gtcagcgagg gggaggagag agcttccctc cataaatggt cccacccctg ggcaaggtgg 1380
ctcactttgg caggtagcaa ccggggagtg tgcacctgcc accagtcaag ctcagccaga 1440
ctgtgagaag aggagaggcg aggcacacca agggatccag tgaaccaacg acagattgaa 1500
gtgcccgaac ttcttcaagt gcagacagaa ggaggtaggg ttctggaagt ttctggtggt 1560
gtaggggagt caggaaggga aaacaaggag ggagagtgag tcttagtttt ttgctttctg 1620
tagctgttcc ttattttgca tatttctttc tcttcaactc ttttcaagta tgcctgatac 1680
gttgttctca cgaagttgac gtgaaaaaca actttcctgc tggtagttag gaaacttagg 1740
agcacctcaa cctgtacctt gagaacaccc agagaatgct gctctttgtc gttctctata 1800
ccgtgttcat atgctgcagg gaaatgcaaa gaatgtactg tccttatctg accctgggag 1860
cattccatag tcaagcagca gctatcaggt tgggaaagag ctcctctcca aggtgtaaca 1920
gaaaaggaaa atgttgatat ttttcttgtt tagaaagtga cagcttcatc cgagaacgcg 1980
gccgcgccac catgtcccgc ctgctgcatg cagaagagtg ggctgaggta aaagagttgg 2040
gggaccacca tcgccatccc cagccgcacc acgtcccgcc gctgacgcca cagccacctg 2100
ctaccctgca ggcgagagac cttcccgtct acccggcaga actgtccctc ctggatagca 2160
ccgacccacg cgcctggctg actcccactt tgcagggcct ctgcacggca cgcgccgccc 2220
agtatctgct gcattctccc gagctgggtg cctccgaggc cgcggcgccc cgggacgagg 2280
ctgacagcca gggtgagctg gtaaggagaa gcggctgtgg cggcctcagc aagagccccg 2340
ggcccgtcaa agtacgggaa cagctgtgca agctgaaggg tggggttgta gtggacgagc 2400
ttggctgcag ccgccagcga gcccccttcca gcaaacaggt gaatggggta cagaagcaaa 2460
ggaggctggc agcaaacgca agggaacggc gcaggatgca cgggctgaac cacgccttcg 2520
accagctgcg caacgttatc ccgtccttca acaacgacaa gaagctgtcc aaatatgaga 2580
ccctacagat ggcccagatc tacatcaacg ctctgtcgga gttgctgcag actcccaatg 2640
tcggagagca accgccgccg cccacagctt cctgcaaaaa tgaccaccat caccttcgca 2700
ccgcctcctc ctatgaagga ggtgcgggcg cctctgcggt agctggggct cagccagccc 2760
cgggaggggg cccgagacct accccgcccg ggccttgccg gactcgcttc tcaggcccag 2820
cttcctctgg gggttactcg gtgcagctgg acgctttgca cttcccagcc ttcgaggaca 2880
gggccctaac agcgatgatg gcacagaagg acctgtcgcc ttcgctgccc gggggcatcc 2940
tgcagcctgt acaggaggac aacagcaaaa catctcccag atcccacaga agtgacggag 3000
agttttcccc ccactctcat tacagtgact ctgatgaggc cagttagaag cttggatcca 3060
atcaacctct ggattacaaa atttgtgaaa gattgactgg tattcttaac tatgttgctc 3120
cttttacgct atgtggatac gctgctttaa tgcctttgta tcatgctatt gcttcccgta 3180
tggctttcat tttctcctcc ttgtataaat cctggttgct gtctctttat gaggagttgt 3240

ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt tgctgacgca acccccactg 3300
gttggggcat tgccaccacc tgtcagctcc tttccgggac tttcgctttc cccctcccta 3360
ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg ctggacaggg gctcggctgt 3420
tgggcactga caattccgtg gtgttgtcgg ggaaatcatc gtcctttcct tggctgctcg 3480
cctgtgttgc cacctggatt ctgcgcggga cgtccttctg ctacgtccct tcggccctca 3540
atccagcgga ccttccttcc cgcggcctgc tgccggctct gcggcctctt ccgcgtcttc 3600
gagatctgcc tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctcccccgt 3660
gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat 3720
tgcatcgcat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag 3780
caagggggag gattgggaag acaatagcag gcatgctggg gactcgagtt aagggcgaat 3840
tcccgataag gatcttccta gagcatggct acgtagataa gtagcatggc gggttaatca 3900
ttaactacaa ggaaccccta gtgatggagt tggccactcc ctctctgcgc gctcgctcgc 3960
tcactgaggc cgggcgacca aaggtcgccc gacgcccggg ctttgcccgg gcggcctcag 4020
tgagcgagcg agcgcgcagc cttaattaac ctaattcact ggccgtcgtt ttacaacgtc 4080
gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat ccccctttcg 4140
ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc 4200
tgaatggcga atgggacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta 4260
cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc 4320
cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctccctt 4380
tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg 4440
gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca 4500
cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcggtct 4560
attcttttga tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga 4620
tttaacaaaa atttaacgcg aattttaaca aaatattaac gcttacaatt taggtggcac 4680
ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac attcaaatat 4740
gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa aaaggaagag 4800
tatgagccat attcaacggg aaacgtcgag gccgcgatta aattccaaca tggatgctga 4860
tttatatggg tataaatggg ctcgcgataa tgtcgggcaa tcaggtgcga caatctatcg 4920
cttgtatggg aagcccgatg cgccagagtt gtttctgaaa catggcaaag gtagcgttgc 4980
caatgatgtt acagatgaga tggtcagact aaactggctg acggaattta tgcctcttcc 5040
gaccatcaag cattttatcc gtactcctga tgatgcatgg ttactcacca ctgcgatccc 5100
cggaaaaaca gcattccagg tattagaaga atatcctgat tcaggtgaaa atattgttga 5160
tgcgctggca gtgttcctgc gccggttgca ttcgattcct gtttgtaatt gtccttttaa 5220
cagcgatcgc gtatttcgtc ttgctcaggc gcaatcacga atgaataacg gtttggttga 5280
tgcgagtgat tttgatgacg agcgtaatgg ctggcctgtt gaacaagtct ggaaagaaat 5340
gcataaactt ttgccattct caccggattc agtcgtcact catggtgatt tctcacttga 5400
taaccttatt tttgacgagg ggaaattaat aggttgtatt gatgttggac gagtcggaat 5460
cgcagaccga taccaggatc ttgccatcct atggaactgc ctcggtgagt tttctccttc 5520
attacagaaa cggctttttc aaaaatatgg tattgataat cctgatatga ataaattgca 5580

gtttcatttg atgctcgatg agtttttcta actgtcagac caagtttact catatatact 5640
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga 5700
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 5760
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 5820
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 5880
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgta 5940
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 6000
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 6060
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggttt cgtgcacaca 6120
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 6180
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 6240
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 6300
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 6360
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 6420
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 6480
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 6540
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 6600
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 6660
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 6720
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 6780
cgccagattt aattaaggcc ttaattagg 6809

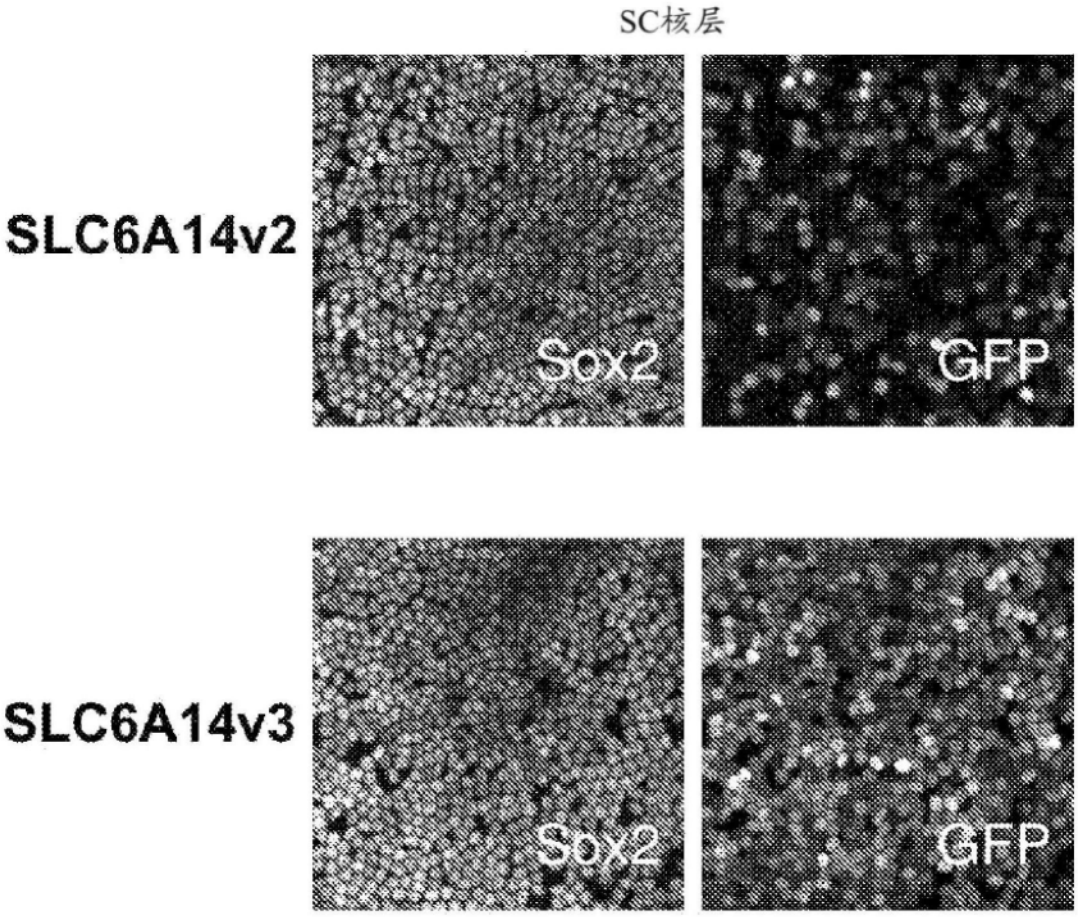


图1A

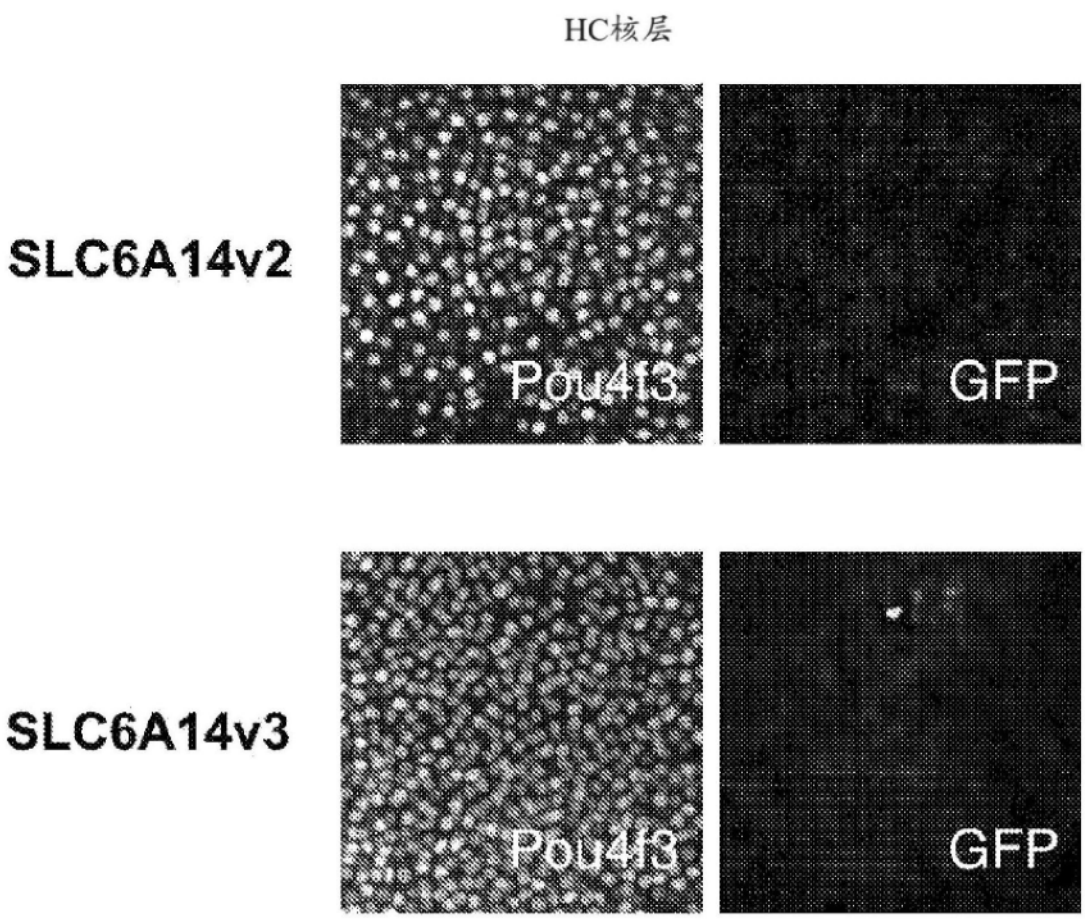


图1B

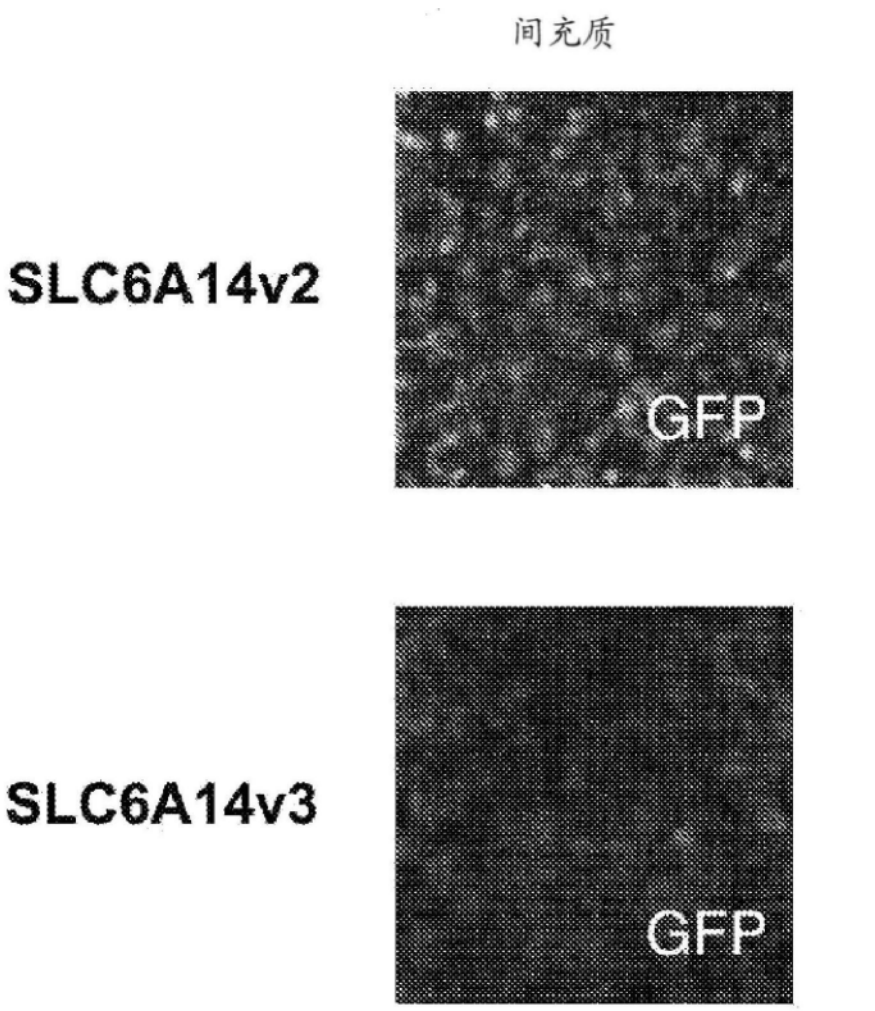


图1C

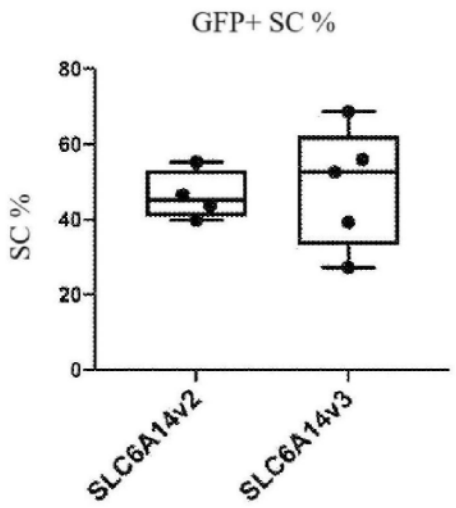

GFP+ SC %
SC %
80
60
40
20
0
SLC6A14v2
SLC6A14v3

图2A

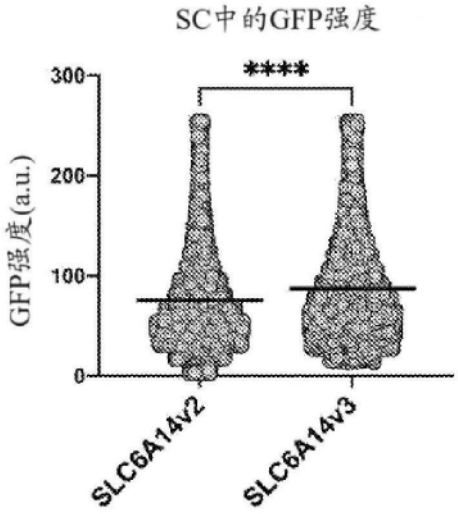

SC中的GFP强度
****
GFP强度(a.u.)
300
200
100
0
SLC6A14v2
SLC6A14v3

图2B

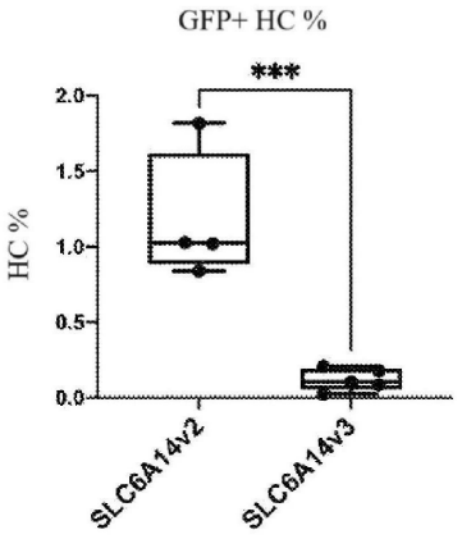

GFP+ HC %
***
HC %
2.0
1.5
1.0
0.5
0.0
SLC6A14v2
SLC6A14v3

图2C

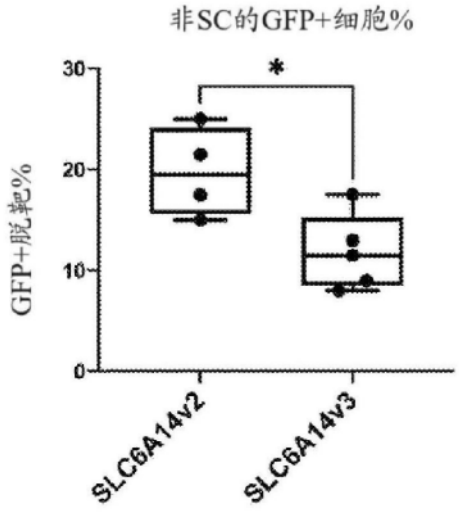

非SC的GFP+细胞%
*
GFP+脱靶%
30
20
10
0
SLC6A14v2
SLC6A14v3

图2D

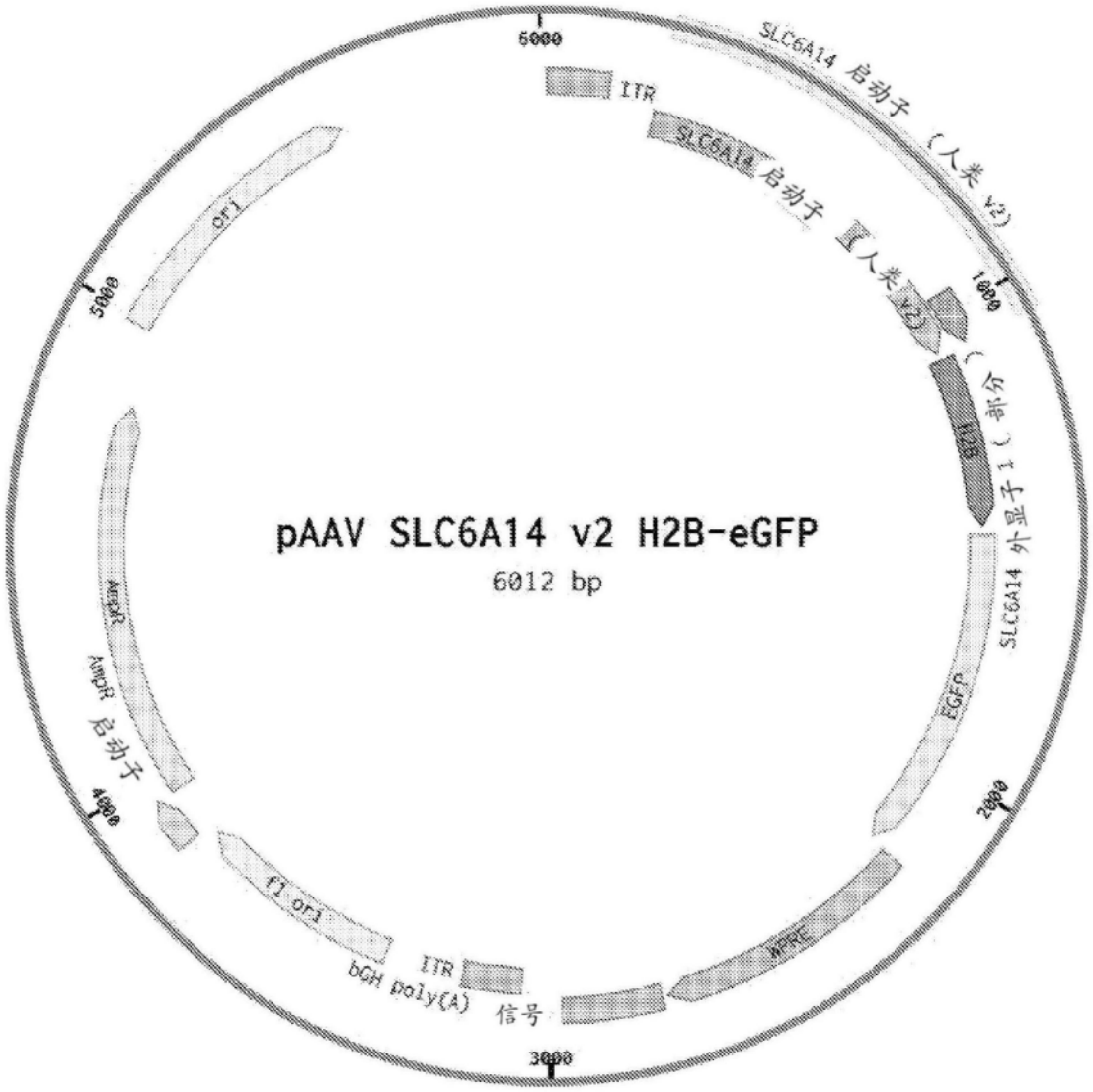

pAAV SLC6A14 v2 H2B-eGFP
6012 bp
6000
1000
2000
3000
4000
5000
ITR
SLC6A14 启动子（人类 v2）
SLC6A14 启动子（人类 v2）
H2B
SLC6A14 外显子 1（部分）
EGFP
WPRE
信号
ITR
bGH poly(A)
f1 ori
AMPR 启动子
AMPR
ori

图3

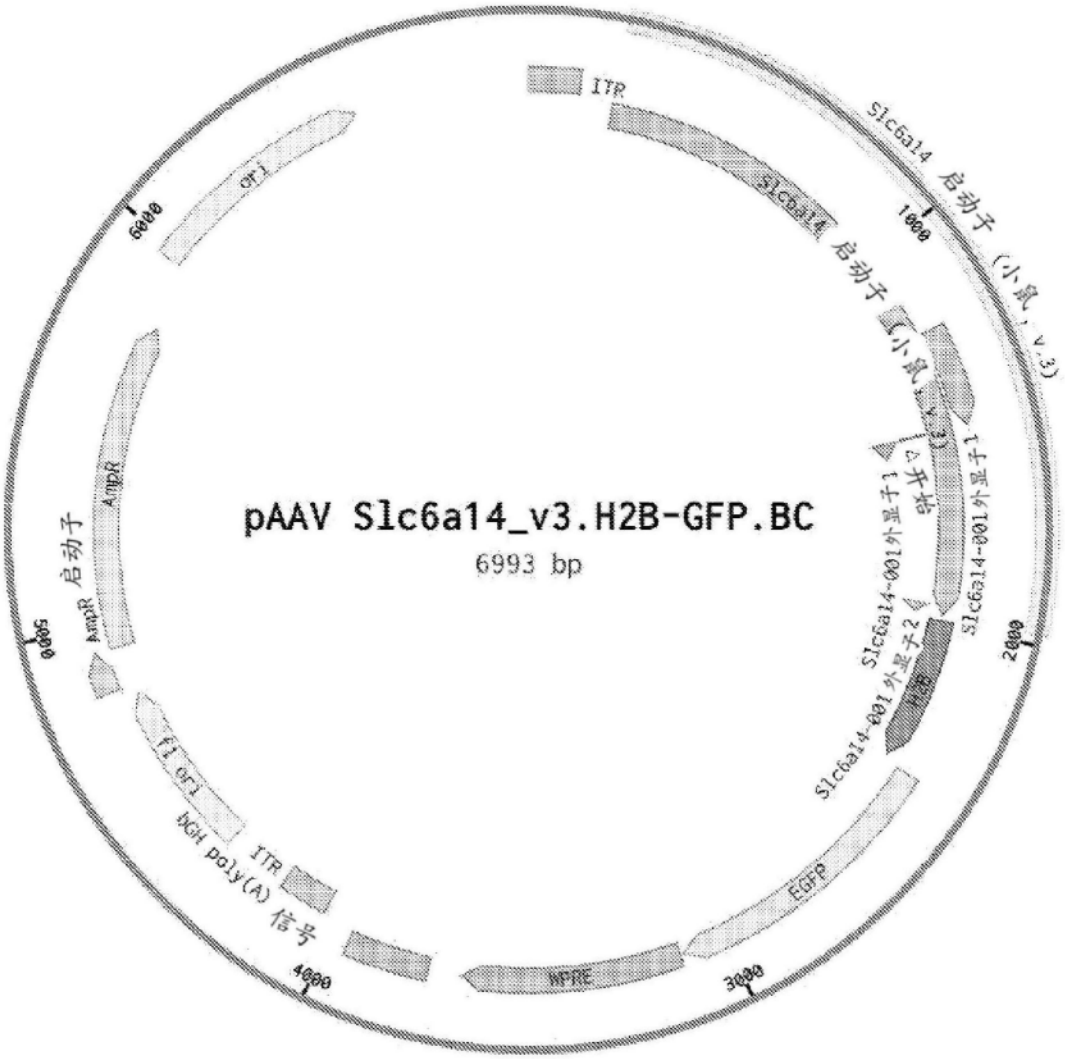

pAAV Slc6a14_v3.H2B-GFP.BC
6993 bp
ITR
Slc6a14 启动子 (小鼠, v3)
Slc6a14-001外显子1
Slc6a14-001外显子2
H2B
EGFP
WPRE
bGH poly(A) 信号
ITR
f1 ori
启动子
AmpR
ori
1000
2000
3000
4000
5000
6000

图4

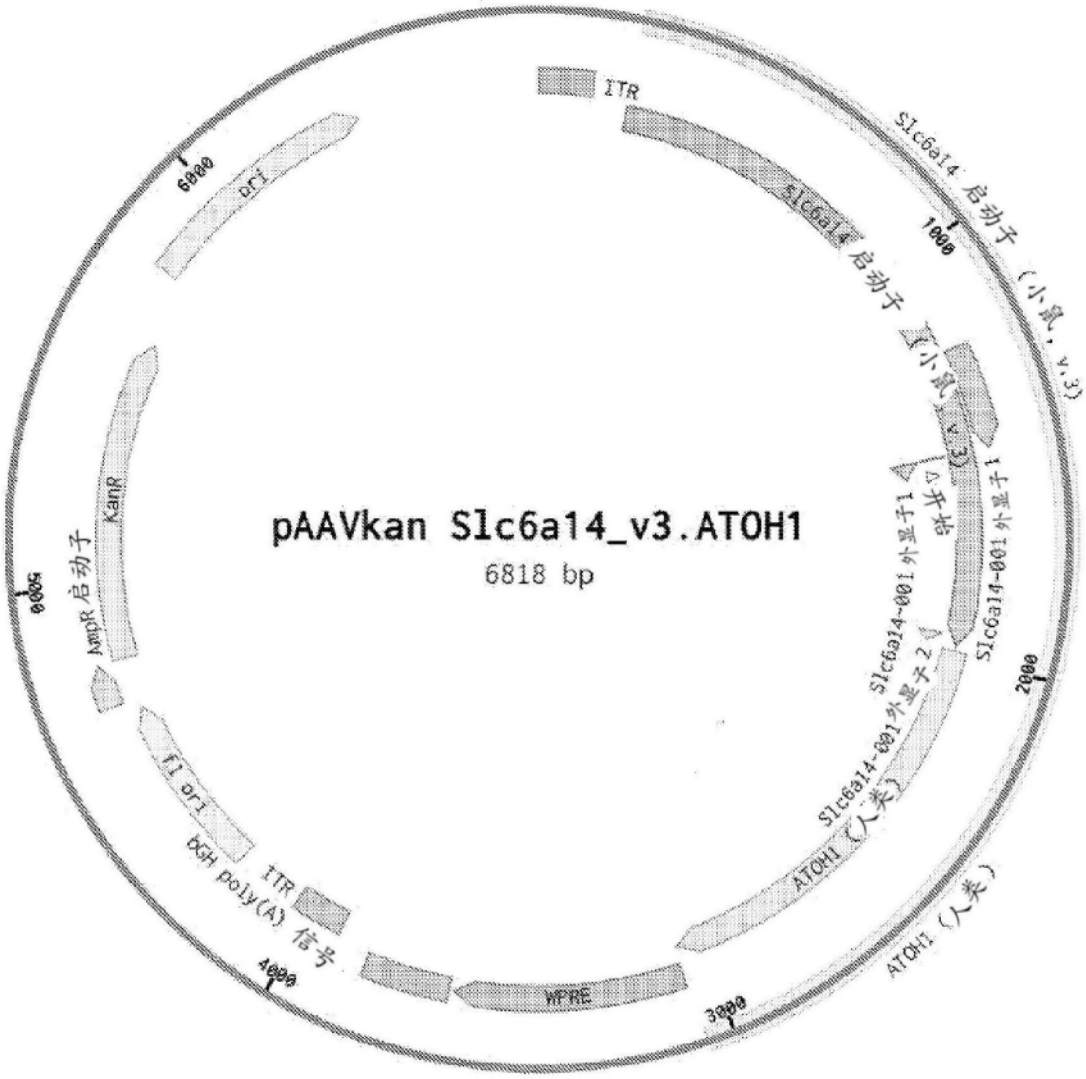

pAAVkan Slc6a14_v3.ATOH1
6818 bp
ITR
Slc6a14 启动子 (小鼠, v.3)
1000
2000
3000
4000
5000
6000
ori
Slc6a14 启动子 (小鼠, v.3)
Slc6a14-001外显子1
缺失区
Slc6a14-001外显子1
Slc6a14-001外显子2
Slc6a14-001外显子2
ATOH1 (人类)
ATOH1 (人类)
WPRE
ITR
bGH poly(A) 信号
f1 ori
KanR
AmpR 启动子

图5

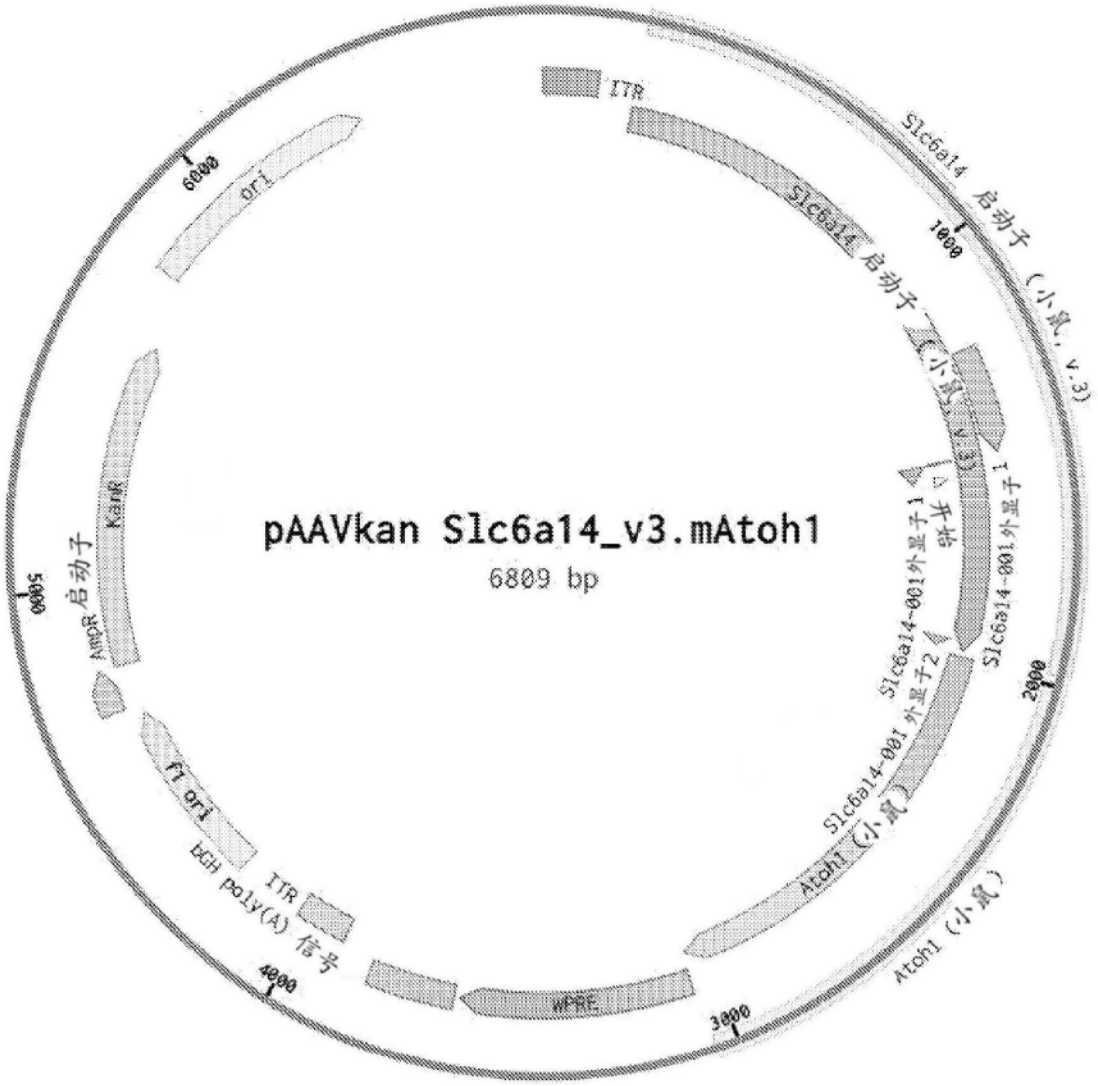

pAAVkan Slc6a14_v3.mAtoh1
6809 bp
ITR
Slc6a14 启动子 (小鼠, v.3)
Slc6a14 启动子 (小鼠, v.3)
Slc6a14-001外显子1
Slc6a14-001外显子1
Slc6a14-001外显子2
Slc6a14-001外显子2
Atoh1 (小鼠)
Atoh1 (小鼠)
WPRE
ITR
bGH poly(A) 信号
f1 ori
AmpR 启动子
KanR
ori
1000
2000
3000
4000
5000
6000

图6