(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117295529 A
(43)申请公布日 2023.12.26

(21)申请号 202280028998.9
(22)申请日 2022.02.18
(30)优先权数据
63/151589 2021.02.19 US
(85)PCT国际申请进入国家阶段日
2023.10.17
(86)PCT国际申请的申请数据
PCT/US2022/017058 2022.02.18
(87)PCT国际申请的公布数据
WO2022/178298 EN 2022.08.25
(71)申请人 分贝治疗公司
地址 美国马萨诸塞州
(72)发明人 A·巴勒莫 N·潘 A·塞纳帕蒂 J·惠顿 X·张
(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 黄登高 杨思捷
(51)Int.Cl.
A61K 48/00(2006.01)

权利要求书4页 说明书162页
序列表139页 附图6页

(54)发明名称

使用耳畸蛋白双载体系统治疗感觉神经性听力损失的方法

(57)摘要

本公开提供借助于OTOF基因疗法来治疗具有耳畸蛋白(OTOF)中的双等位基因突变的25岁或25岁以上的受试者的组合物和方法。本公开提供了多种组合物,所述组合物包括含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体。这些载体能够用于治疗具有双等位基因OTOF突变的受试者的听力损失或听神经病变。

CN 117295529 A

1.一种治疗具有双等位基因耳畸蛋白(OTOF)突变的25岁或25岁以上的人类受试者的方法,所述方法包括向所述受试者施用治疗有效量的双载体系统,所述双载体系统包含:

第一核酸载体,所述第一核酸载体包含可操作地连接至编码OTOF蛋白的N末端部分的第一编码多核苷酸的启动子;以及

第二核酸载体,所述第二核酸载体包含编码OTOF蛋白的C末端部分的第二编码多核苷酸和定位于所述第二编码多核苷酸3'的poly(A)序列;

其中所述第一核酸载体和所述第二核酸载体都不编码全长OTOF蛋白。

2.如权利要求1所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸并不重叠。

3.如权利要求1或2所述的方法,其中所述第一核酸载体包含定位于所述第一编码多核苷酸3'的剪接供体信号序列,并且所述第二核酸载体包含定位于所述第二编码多核苷酸5'的剪接受体信号序列。

4.如权利要求3所述的方法,其中所述第一核酸载体包含定位于所述剪接供体信号序列3'的第一重组基因区,并且所述第二核酸载体包含定位于所述剪接受体信号序列5'的第二重组基因区。

5.如权利要求4所述的方法,其中所述第一重组基因区和所述第二重组基因区是相同的。

6.如权利要求4或5所述的方法,其中所述第一重组基因区或所述第二重组基因区是AP基因片段或F1噬菌体AK基因。

7.如权利要求4-6中任一项所述的方法,其中所述第一核酸载体还包含定位于所述重组基因区3'的降解信号序列;并且其中所述第二核酸载体还包含定位于所述重组基因区与所述剪接受体信号序列之间的降解信号序列。

8.如权利要求1-7中任一项所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸在OTOF外显子边界处分开。

9.如权利要求1所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸部分地重叠。

10.如权利要求9所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸重叠至少1千碱基(kb)。

11.如权利要求9或10所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸之间的重叠区以OTOF外显子边界为中心。

12.如权利要求11所述的方法,其中所述第一编码多核苷酸编码所述OTOF蛋白的N末端部分,并且包含OTOF N末端至所述重叠区中心处的所述外显子边界3'的500bp;并且所述第二编码多核苷酸编码所述OTOF蛋白的C末端部分,并且包含所述重叠区中心处的所述外显子边界5'的500bp至所述OTOF C末端。

13.如权利要求8、11和12中任一项所述的方法,其中选择所述OTOF外显子边界,使得所述第一编码多核苷酸编码整个C2C结构域,并且所述第二编码多核苷酸编码整个C2D结构域。

14.如权利要求8和11-13中任一项所述的方法,其中所述OTOF外显子边界是外显子19/20边界、外显子20/21边界或外显子21/22边界。

15.如权利要求8、11和12中任一项所述的方法,其中选择所述OTOF外显子边界,使得所述第一编码多核苷酸编码整个C2D结构域,并且所述第二编码多核苷酸编码整个C2E结构域。

16.如权利要求8、11、12和15中任一项所述的方法,其中所述OTOF外显子边界是外显子26/27边界或外显子28/29边界。

17.如权利要求8、11和12中任一项所述的方法,其中所述OTOF外显子边界在编码C2D结构域的所述第一编码多核苷酸和所述第二编码多核苷酸的一部分内。

18.如权利要求8、11、12和17中任一项所述的方法,其中所述OTOF外显子边界是外显子24/25边界或外显子25/26边界。

19.如权利要求1-18中任一项所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸中的每一者编码约一半的所述OTOF蛋白序列。

20.如权利要求1-19中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体不包含OTOF非翻译区(UTR)。

21.如权利要求1-19中任一项所述的方法,其中所述第一核酸载体包含OTOF 5'UTR。

22.如权利要求1-19和21中任一项所述的方法,其中所述第二核酸载体包含OTOF 3'UTR。

23.如权利要求1-22中任一项所述的方法,其中编码所述OTOF蛋白的所述第一编码多核苷酸和所述第二编码多核苷酸不包含内含子。

24.如权利要求1-23中任一项所述的方法,其中所述OTOF蛋白是哺乳动物OTOF蛋白。

25.如权利要求24所述的方法,其中所述OTOF蛋白是人类OTOF蛋白。

26.如权利要求1-25中任一项所述的方法,其中所述OTOF蛋白与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列具有至少85%同一性。

27.如权利要求26所述的方法,其中所述OTOF蛋白具有SEQ ID NO:1的序列。

28.如权利要求26所述的方法,其中所述OTOF蛋白具有SEQ ID NO:5的序列。

29.如权利要求1-25中任一项所述的方法,其中所述OTOF蛋白包含SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列或其具有一个或多个保守氨基酸取代的变体。

30.如权利要求29所述的方法,其中所述OTOF蛋白变体中不超过10%的氨基酸是保守氨基酸取代。

31.如权利要求1-30中任一项所述的方法,其中所述第一核酸载体包含所述启动子3'和编码所述OTOF蛋白的所述N末端部分的所述第一编码多核苷酸5'的Kozak序列。

32.如权利要求1-31中任一项所述的方法,其中所述启动子是普遍存在的启动子。

33.如权利要求32所述的方法,其中所述普遍存在的启动子是CAG启动子、巨细胞病毒(CMV)启动子、鸡β-肌动蛋白启动子、截短CMV-鸡β-肌动蛋白启动子(smCBA)、CB7启动子、杂合CMV增强子/人类β-肌动蛋白启动子、人类β-肌动蛋白启动子、延伸因子-1α(EF1α)启动子或磷酸甘油酸激酶(PGK)启动子。

34.如权利要求1-31中任一项所述的方法,其中所述启动子是耳蜗毛细胞特异性启动子。

35.如权利要求34所述的方法,其中所述耳蜗毛细胞特异性启动子是肌凝蛋白15

(Myo15)启动子、肌凝蛋白7A(Myo7A)启动子、肌凝蛋白6(Myo6)启动子、POU 4类同源盒3(POU4F3)启动子、atonal BHLH转录因子1(ATOH1)启动子、LIM同源盒3(LHX3)启动子、α9乙酰胆碱受体(α9AChR)启动子或α10乙酰胆碱受体(α10AChR)启动子。

36.如权利要求1-31中任一项所述的方法,其中所述启动子是内毛细胞特异性启动子。

37.如权利要求36所述的方法,其中所述内毛细胞特异性启动子是成纤维细胞生长因子8(FGF8)启动子、囊泡谷氨酸转运蛋白3(VGLUT3)启动子、OTOF启动子或钙结合蛋白2(CABP2)启动子。

38.如权利要求1-37中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体在所述核酸序列的每个末端处包含反向末端重复(ITR)。

39.如权利要求38所述的方法,其中所述ITR是AAV2 ITR或与AAV2 ITR具有至少80%序列同一性。

40.如权利要求1-39中任一项所述的方法,其中所述第二核酸载体包含土拨鼠肝炎病毒转录后调控元件(WPRE)。

41.如权利要求1-40中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体是腺相关病毒(AAV)载体。

42.如权利要求1-41中任一项所述的方法,其中所述受试者是30岁或30岁以上。

43.如权利要求1-42中任一项所述的方法,其中所述受试者是35岁或35岁以上。

44.如权利要求1-43中任一项所述的方法,其中所述受试者是40岁或40岁以上。

45.如权利要求1-44中任一项所述的方法,其中所述受试者是45岁或45岁以上。

46.如权利要求1-45中任一项所述的方法,其中所述受试者不大于50岁。

47.如权利要求1-46中任一项所述的方法,其中所述受试者已被鉴定为具有双等位基因OTOF突变。

48.如权利要求1-47中任一项所述的方法,其中所述方法还包括在施用所述双载体系统之前将所述受试者鉴定为具有双等位基因OTOF突变。

49.如权利要求1-48中任一项所述的方法,其中所述受试者被鉴定为具有可检测的耳声发射。

50.如权利要求1-49中任一项所述的方法,其中所述受试者被鉴定为具有可检测的耳蜗微音效应。

51.如权利要求1-50中任一项所述的方法,其中所述受试者被鉴定为具有可检测的总和电位。

52.如权利要求1-51中任一项所述的方法,其中所述受试者患有或被鉴定为患有9型常染色体隐性耳聋(DFNB9)。

53.如权利要求1-52中任一项所述的方法,其中将所述双载体系统施用于内耳。

54.如权利要求1-53中任一项所述的方法,其中所述第一载体和所述第二载体是同时施用的。

55.如权利要求1-54中任一项所述的方法,其中所述第一载体和所述第二载体是以约1×10^{7}个载体基因组(VG)/耳至约2×10^{15}个VG/耳的浓度施用。

56.如权利要求1-55中任一项所述的方法,其中所述第一载体和所述第二载体是以合起来足以用所述第一载体和所述第二载体两者转导至少20%的所述受试者的内毛细胞的

量施用。

使用耳畸蛋白双载体系统治疗感觉神经性听力损失的方法

[0001] 序列表

[0002] 本申请含有序列表,所述序列表已以ASCII格式以电子方式提交并且特此通过引用整体并入本文。ASCII副本创建于2022年2月18日,命名为51471-008WO2_Sequence_Listing_2_17_22_ST25并且大小为366,491个字节。

技术领域

[0003] 本文描述了借助于OTOF基因疗法治疗25岁或25岁以上的人类受试者的感觉神经性听力损失和听神经病变、特别是与耳畸蛋白(OTOF)突变相关的疾病形式的组合物和方法。本公开提供了双载体系统,所述双载体系统包括含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体。这些载体可用于增加野生型OTOF的表达或向受试者(诸如患有感觉神经性听力损失的人类受试者)提供野生型OTOF。

背景技术

[0004] 感觉神经性听力损失是一种由内耳细胞或从内耳投射至脑的神经通路的缺陷引起的听力损失。尽管感觉神经性听力损失通常是获得性的并且可由噪声、感染、头部创伤、耳毒性药物或衰老引起,但也存在与常染色体隐性突变相关的先天性形式的感觉神经性听力损失。常染色体隐性感觉神经性听力损失的一种此类形式与耳畸蛋白(OTOF)基因的突变相关,所述耳畸蛋白基因与语前非综合征性听力损失有关。近年来,治疗听力损失的努力越来越多地集中在基因疗法作为一种可能的解决方案;然而,OTOF因太大而不允许使用标准基因治疗方法来治疗。需要新的治疗剂来治疗OTOF相关的感觉神经性听力损失。

发明内容

[0005] 本发明提供了用于治疗具有双等位基因耳畸蛋白(OTOF)突变的25岁或25岁以上的人类受试者的组合物和方法,已知所述突变引起听力损失和听神经病变。本文所述的组合物可用于通过基因疗法将野生型(WT)OTOF递送至受试者,并且因此可用于治疗受试者的听力损失和听神经病变。用于治疗双等位基因OTOF突变的基因疗法被认为在生命的第一年期间需要以恢复听力;然而,本发明人类已经确定基因疗法可恢复由于双等位基因OTOF突变而丧失的听力,即使治疗是在很多年之后开始。本文所述的组合物还可用于治疗具有双等位基因OTOF突变的受试者,所述受试者被鉴定为具有可检测的耳声发射、可检测的耳蜗微音效应和/或可检测的总和电位。

[0006] 在第一方面,本发明提供了一种通过向具有双等位基因耳畸蛋白(OTOF)突变的25岁或25岁以上的人类受试者施用治疗有效量的双载体系统来治疗所述受试者的方法,所述双载体系统包括:第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白的N末端部分的第一编码多核苷酸的启动子;以及第二核酸载体,所述第二核酸载体含有编码OTOF蛋白的C末端部分的第二编码多核苷酸和定位于所述第二编码多核苷酸3'的聚腺苷酸

化(poly(A))序列;其中所述第一核酸载体和所述第二核酸载体都不编码全长OTOF蛋白。

[0007] 在另一方面,本发明提供了一种通过向具有双等位基因耳畸蛋白(OTOF)突变并且被鉴定为具有可检测的耳声发射、可检测的耳蜗微音效应和/或可检测的总和电位的人类受试者施用治疗有效量的双载体系统来治疗所述受试者的方法,所述双载体系统包括:第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白的N末端部分的第一编码多核苷酸的启动子;以及第二核酸载体,所述第二核酸载体含有编码OTOF蛋白的C末端部分的第二编码多核苷酸和定位于所述第二编码多核苷酸3'的聚腺苷酸化(poly(A))序列;其中所述第一核酸载体和所述第二核酸载体都不编码全长OTOF蛋白。

[0008] 在前述方面中任一者的一些实施方案中,第一编码多核苷酸和第二编码多核苷酸并不重叠。

[0009] 在前述方面中任一者的一些实施方案中,第一核酸载体包括定位于第一编码多核苷酸3'的剪接供体信号序列,并且第二核酸载体包括定位于第二编码多核苷酸5'的剪接受体信号序列。在一些实施方案中,第一核酸载体包括定位于剪接供体信号序列3'的第一重组基因区,并且第二核酸载体包括定位于剪接受体信号序列5'的第二重组基因区。在一些实施方案中,第一重组基因区和第二重组基因区是相同的。在一些实施方案中,第一重组基因区和/或第二重组基因区是AP基因片段或F1噬菌体AK基因。在一些实施方案中,F1噬菌体AK基因包括或具有SEQ ID NO:19的序列。在一些实施方案中,AP基因片段包括或具有SEQ ID NO:62-67中任一者的序列。在一些实施方案中,AP基因片段包括或具有SEQ ID NO:65的序列。在一些实施方案中,剪接供体序列包括或具有SEQ ID NO:20或SEQ ID NO:68的序列。在一些实施方案中,剪接受体序列包括或具有SEQ ID NO:21或SEQ ID NO:69的序列。在一些实施方案中,第一核酸载体还包括定位于重组基因区3'的降解信号序列,并且第二核酸载体还包括定位于重组基因区与剪接受体信号序列之间的降解信号序列。在一些实施方案中,降解信号序列包括或具有SEQ ID NO:22的序列。

[0010] 在前述方面中任一者的一些实施方案中,第一编码多核苷酸和第二编码多核苷酸在OTOF外显子边界处分开。在一些实施方案中,OTOF外显子边界不在编码C2结构域的第一编码多核苷酸或第二编码多核苷酸的一部分内。

[0011] 在前述方面中任一者的一些实施方案中,第一编码多核苷酸与第二编码多核苷酸部分重叠。在一些实施方案中,第一编码多核苷酸与第二编码多核苷酸重叠至少1千碱基(kb)。在一些实施方案中,第一编码多核苷酸与第二编码多核苷酸之间的重叠区以OTOF外显子边界为中心。在一些实施方案中,第一编码多核苷酸编码OTOF蛋白的N末端部分,并且包括OTOF N末端至重叠区中心处的外显子边界3'的500bp;并且第二编码多核苷酸编码OTOF蛋白的C末端部分,并且包括重叠区中心处的外显子边界5'的500bp至OTOF C末端。在一些实施方案中,重叠区中心处的OTOF外显子边界不在编码C2结构域的第一编码多核苷酸或第二编码多核苷酸的一部分内。

[0012] 在前述方面中任一者的一些实施方案中,选择OTOF外显子边界,使得第一编码多核苷酸编码整个C2C结构域,并且第二编码多核苷酸编码整个C2D结构域。在一些实施方案中,OTOF外显子边界是外显子19/20边界、外显子20/21边界或外显子21/22边界。

[0013] 在前述方面中任一者的一些实施方案中,选择OTOF外显子边界,使得第一编码多核苷酸编码整个C2D结构域,并且第二编码多核苷酸编码整个C2E结构域。在一些实施方案

中,OTOF外显子边界是外显子26/27边界或外显子28/29边界。

[0014] 在前述方面中任一者的一些实施方案中,OTOF外显子边界在编码C2D结构域的第一编码多核苷酸和第二编码多核苷酸的一部分内。在一些实施方案中,OTOF外显子边界是外显子24/25边界或外显子25/26边界。

[0015] 在前述方面中任一者的一些实施方案中,第一编码多核苷酸和第二编码多核苷酸中的每一者编码约一半的OTOF蛋白序列。

[0016] 在前述方面中任一者的一些实施方案中,第一核酸载体和第二核酸载体不包括OTOF非翻译区(UTR)。

[0017] 在前述方面中任一者的一些实施方案中,第一核酸载体包括OTOF 5'UTR。

[0018] 在前述方面中任一者的一些实施方案中,第二核酸载体包括OTOF 3'UTR。

[0019] 在前述方面中任一者的一些实施方案中,编码OTOF蛋白的第一编码多核苷酸和第二编码多核苷酸不包括内含子。

[0020] 在前述方面中任一者的一些实施方案中,编码OTOF蛋白的第一编码多核苷酸和第二编码多核苷酸不含内含子。

[0021] 在前述方面中任一者的一些实施方案中,OTOF蛋白是哺乳动物OTOF蛋白。

[0022] 在前述方面中任一者的一些实施方案中,OTOF蛋白是鼠OTOF蛋白。在前述方面中任一者的一些实施方案中,鼠OTOF蛋白与SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9的序列具有至少85%序列同一性(例如,85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性)。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:6的序列或由其组成。

[0023] 在前述方面中任一者的一些实施方案中,OTOF蛋白是人类OTOF蛋白。在前述方面中任一者的一些实施方案中,人类OTOF蛋白与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列具有至少85%序列同一性(例如,85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性)。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:1的序列或由其组成。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:2的序列或由其组成。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:3的序列或由其组成。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:4的序列或由其组成。在前述方面中任一者的一些实施方案中,OTOF蛋白包含SEQ ID NO:5的序列或由其组成。在前述方面中任一者的一些实施方案中,人类OTOF蛋白包含SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列或其具有一个或多个(例如,1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)保守氨基酸取代的变体。在一些实施方案中,OTOF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。

[0024] 在前述方面中任一者的一些实施方案中,OTOF蛋白由SEQ ID NO:10-14中的任一者编码。在一些实施方案中,OTOF蛋白由SEQ ID NO:10编码。在一些实施方案中,OTOF蛋白由SEQ ID NO:14编码。

[0025] 在前述方面中任一者的一些实施方案中,OTOF蛋白由SEQ ID NO:15-18中的任一者编码。

[0026] 在前述方面中任一者的一些实施方案中,第一编码多核苷酸编码SEQ ID NO:1或SEQ ID NO:5的氨基酸1-802,并且第二编码多核苷酸编码SEQ ID NO:1或SEQ ID NO:5的氨基酸803-1997。在一些实施方案中,第一编码多核苷酸编码SEQ ID NO:1的氨基酸1-802,并且第二编码多核苷酸编码SEQ ID NO:1的氨基酸803-1997。在一些实施方案中,第一编码多核苷酸编码SEQ ID NO:5的氨基酸1-802,并且第二编码多核苷酸编码SEQ ID NO:5的氨基酸803-1997。

[0027] 在前述方面中任一者的一些实施方案中,OTOF蛋白的N末端部分由SEQ ID NO:73的序列或其具有一个或多个(例如,1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)保守氨基酸取代的变体组成。在一些实施方案中,OTOF蛋白变体的N末端部分中不超过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,OTOF蛋白的N末端部分由SEQ ID NO:73的序列组成。在一些实施方案中,OTOF蛋白的N末端部分由SEQ ID NO:71的序列编码。

[0028] 在前述方面中任一者的一些实施方案中,OTOF蛋白的C末端部分由SEQ ID NO:74的序列或其具有一个或多个(例如,1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)保守氨基酸取代的变体组成。在一些实施方案中,OTOF蛋白变体的C末端部分中不超过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,OTOF蛋白的C末端部分由SEQ ID NO:74的序列组成。在一些实施方案中,OTOF蛋白的C末端部分由SEQ ID NO:72的序列编码。

[0029] 在前述方面中任一者的一些实施方案中,第一核酸载体包括定位于启动子3'和编码OTOF蛋白的N末端部分的第一编码多核苷酸5'的Kozak序列。

[0030] 在前述方面中任一者的一些实施方案中,启动子是普遍存在的启动子。在一些实施方案中,普遍存在的启动子是CAG启动子、巨细胞病毒(CMV)启动子、鸡β-肌动蛋白启动子、截短CMV-鸡β-肌动蛋白启动子(smCBA)、CB7启动子、杂合CMV增强子/人类β-肌动蛋白启动子、人类β-肌动蛋白启动子、延伸因子-1α(EF1α)启动子或磷酸甘油酸激酶(PGK)启动子。在一些实施方案中,普遍存在的启动子是CAG启动子。在一些实施方案中,普遍存在的启动子是smCBA启动子。在一些实施方案中,smCBA启动子具有SEQ ID NO:70的序列。

[0031] 在前述方面中任一者的一些实施方案中,启动子是耳蜗毛细胞特异性启动子。在一些实施方案中,耳蜗毛细胞特异性启动子是肌凝蛋白15(Myo15)启动子、肌凝蛋白7A(Myo7A)启动子、肌凝蛋白6(Myo6)启动子、POU 4类同源盒3(POU4F3)启动子、atonal BHLH转录因子1(ATOH1)启动子、LIM同源盒3(LHX3)启动子、α9乙酰胆碱受体(α9AChR)启动子或α10乙酰胆碱受体(α10AChR)启动子。在一些实施方案中,耳蜗毛细胞特异性启动子是Myo15启动子。

[0032] 在前述方面中任一者的一些实施方案中,启动子是内毛细胞特异性启动子。在一些实施方案中,内毛细胞特异性启动子是成纤维细胞生长因子8(FGF8)启动子、囊泡谷氨酸转运蛋白3(VGLUT3)启动子、OTOF启动子或钙结合蛋白2(CABP2)启动子。在一些实施方案中,内毛细胞特异性启动子是CABP2启动子。

[0033] 在前述方面中任一者的一些实施方案中,启动子是短启动子(例如,1kb或更短,例如大约1kb、950bp、900bp、850bp、800bp、750bp、700bp、650bp、600bp、550bp 500bp、450bp、400bp、350bp、300bp或更短的启动子)。在一些实施方案中,短启动子是CAG启动子。在一些

实施方案中,短启动子是CMV启动子。在一些实施方案中,短启动子是smCBA启动子。在一些实施方案中,短启动子是1kb或更短的Myo15启动子(例如,具有与SEQ ID NO:38、39或49-60中的任一者具有至少85%序列同一性(例如,85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性)的序列的Myo15启动子)。

[0034] 在前述方面中任一者的一些实施方案中,启动子是长启动子(例如,长于1kb,例如1.1kb、1.25kb、1.5kb、1.75kb、2kb、2.5kb、3kb或更长的启动子)。在一些实施方案中,长启动子是长于1kb的Myo15启动子(例如,包含与SEQ ID NO:36的序列具有至少85%序列同一性(例如,85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性)的序列或由其组成的Myo15启动子)。

[0035] 在前述方面中任一者的一些实施方案中,第一核酸载体和第二核酸载体是表4中列出的一对核酸载体。

[0036] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:75的核苷酸2272至6041的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:75的核苷酸2049至6264的序列或由其组成。

[0037] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:77的核苷酸182至3949的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:77的核苷酸19至4115的序列或由其组成。

[0038] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:79的核苷酸2267至6014的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:79的核苷酸2049至6237的序列或由其组成。

[0039] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:80的核苷酸177至3924的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:80的核苷酸19至4090的序列或由其组成。

[0040] 在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:76的核苷酸2267至6476的序列。在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:76的核苷酸2049至6693的序列或由其组成。

[0041] 在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:78的核苷酸187至4396的序列。在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:78的核苷酸19至4589的序列或由其组成。

[0042] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:81的核苷酸235至4004的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:81的核苷酸12至4227的序列或由其组成。

[0043] 在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:83的核苷酸230至3977的序列。在前述方面中任一者的一些实施方案中,第一核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:83的核苷酸12至4200的序列或由其组成。

[0044] 在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:72的核苷酸229至4438的序列。在前述方面中任一者的一些实施方案中,第二核酸载体含有多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:82的核苷酸12至4655的序列或由其组成。

[0045] 在前述方面中任一者的一些实施方案中,第一核酸载体和第二核酸载体在核酸序列的每个末端处包含反向末端重复(ITR)。在一些实施方案中,第一载体包括启动子5'的第一反向末端重复(ITR)序列和重组基因区3'的第二ITR序列,并且第二载体包括重组基因区5'的第一ITR序列和poly(A)序列3'的第二ITR序列。在一些实施方案中,第一载体和第二载体中的ITR是AAV2 ITR。在一些实施方案中,第一载体和第二载体中的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)。

[0046] 在前述方面中任一者的一些实施方案中,poly(A)序列是牛生长激素(bGH)poly(A)信号序列。

[0047] 在前述方面中任一者的一些实施方案中,第二核酸载体包括土拨鼠肝炎病毒转录后调控元件(WPRE)。在一些实施方案中,WPRE包含SEQ ID NO:23或SEQ ID NO:61的序列或由其组成。

[0048] 在前述方面中任一者的一些实施方案中,核酸载体是重叠双载体。

[0049] 在前述方面中任一者的一些实施方案中,核酸载体是反式剪接双载体。

[0050] 在前述方面中任一者的一些实施方案中,核酸载体是双杂合载体。

[0051] 在前述方面中任一者的一些实施方案中,核酸载体是腺相关病毒(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和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AAV2quad(Y-F)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AAV8衣壳。

[0052] 在前述方面中任一者的一些实施方案中,第一核酸载体和第二核酸载体具有相同衣壳(例如,第一核酸载体和第二核酸载体都是具有AAV1衣壳或AAV9衣壳的AAV载体)。在前述方面中任一者的一些实施方案中,第一核酸载体和第二核酸载体具有不同的衣壳(例如,第一核酸载体是具有AAV1衣壳的AAV,并且第二核酸载体是具有AAV9衣壳的AAV)。

[0053] 在前述方面中任一者的一些实施方案中,受试者是30岁或30岁以上。

[0054] 在前述方面中任一者的一些实施方案中,受试者是35岁或35岁以上。

[0055] 在前述方面中任一者的一些实施方案中,受试者是40岁或40岁以上。

[0056] 在前述方面中任一者的一些实施方案中,受试者是45岁或45岁以上。

[0057] 在前述方面中任一者的一些实施方案中,受试者不大于50岁。

[0058] 在前述方面中任一者的一些实施方案中,受试者已被鉴定为具有双等位基因OTOF突变。

[0059] 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双载体系统之前将受试者鉴定为具有双等位基因OTOF突变的步骤。

[0060] 在前述方面中任一者的一些实施方案中,受试者被鉴定为具有可检测的耳声发射。

[0061] 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双载体系统之前将受试者鉴定为具有可检测的耳声发射的步骤。

[0062] 在前述方面中任一者的一些实施方案中,受试者被鉴定为具有可检测的耳蜗微音效应。

[0063] 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双载体系统之前将受试者鉴定为具有可检测的耳蜗微音效应的步骤。

[0064] 在前述方面中任一者的一些实施方案中,受试者被鉴定为具有可检测的总和电位。

[0065] 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双载体系统之前将受试者鉴定为具有可检测的总和电位的步骤。

[0066] 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双载体系统之前评估受试者的听力的步骤。

[0067] 在前述方面中任一者的一些实施方案中,受试者患有或被鉴定为患有9型常染色体隐性耳聋(DFNB9)。

[0068] 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双载体系统之前评估受试者的听力的步骤。

[0069] 在前述方面中任一者的一些实施方案中,双载体系统是经局部施用于中耳或内耳。在一些实施方案中,双载体系统是通过注射穿过圆窗膜、注射至半规管中、小管造口术、通过圆窗膜插入导管、经鼓室注射或鼓室内注射来施用。

[0070] 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双矢量系统之后评估受试者的听力的步骤。

[0071] 在前述方面中任一者的一些实施方案中,所述方法还包括增加耳蜗毛细胞中的OTOF表达。在前述方面中任一者的一些实施方案中,所述耳蜗毛细胞是内毛细胞。

[0072] 在前述方面中任一者的一些实施方案中,双载体系统增加细胞(例如耳蜗毛细胞)中的OTOF表达,改善听力(例如,如通过标准测试(诸如测听术、听性脑干反应(ABR)、耳蜗电描记术(ECOG)和耳声发射)所评价),预防或减轻听力损失,延迟听力损失的发生,减缓听力损失的进展,改进言语识别或改进毛细胞功能。

[0073] 在前述方面中任一者的一些实施方案中,双载体系统是以足以增加耳蜗毛细胞中的OTOF表达、预防或减轻听力损失、延迟听力损失的发生、减缓听力损失的进展、改善听力(例如如通过标准测试(诸如测听术、ABR、ECOG和耳声发射)所评价)、改进言语识别或改进毛细胞功能的量施用。

[0074] 在前述方面中任一者的一些实施方案中,第一载体和第二载体是同时施用的。

[0075] 在前述方面中任一者的一些实施方案中,第一载体和第二载体是相继施用的。

[0076] 在前述方面中任一者的一些实施方案中,第一载体和第二载体是以下列浓度施用:$1x10^{7}$个载体基因组(VG)/耳至约$2x10^{15}$VG/耳(例如,$1x10^{7}$VG/耳、$2x10^{7}$VG/耳、$3x10^{7}$VG/耳、$4x10^{7}$VG/耳、$5x10^{7}$VG/耳、$6x10^{7}$VG/耳、$7x10^{7}$VG/耳、$8x10^{7}$VG/耳、$9x10^{7}$VG/耳、$1x10^{8}$VG/耳、$2x10^{8}$VG/耳、$3x10^{8}$VG/耳、$4x10^{8}$VG/耳、$5x10^{8}$VG/耳、$6x10^{8}$VG/耳、$7x10^{8}$VG/耳、$8x10^{8}$VG/耳、$9x10^{8}$VG/耳、$1x10^{9}$VG/耳、$2x10^{9}$VG/耳、$3x10^{9}$VG/耳、$4x10^{9}$VG/耳、$5x10^{9}$VG/耳、$6x10^{9}$VG/耳、$7x10^{9}$VG/耳、$8x10^{9}$VG/耳、$9x10^{9}$VG/耳、$1x10^{10}$VG/耳、$2x10^{10}$VG/耳、$3x10^{10}$VG/耳、$4x10^{10}$VG/耳、$5x10^{10}$VG/耳、$6x10^{10}$VG/耳、$7x10^{10}$VG/耳、$8x10^{10}$VG/耳、$9x10^{10}$VG/耳、$1x10^{11}$VG/耳、$2x10^{11}$VG/耳、$3x10^{11}$VG/耳、$4x10^{11}$VG/耳、$5x10^{11}$VG/耳、$6x10^{11}$VG/耳、$7x10^{11}$VG/耳、$8x10^{11}$VG/耳、$9x10^{11}$VG/耳、$1x10^{12}$VG/耳、$2x10^{12}$VG/耳、$3x10^{12}$VG/耳、$4x10^{12}$VG/耳、$5x10^{12}$VG/耳、$6x10^{12}$VG/耳、$7x10^{12}$VG/耳、$8x10^{12}$VG/耳、$9x10^{12}$VG/耳、$1x10^{13}$VG/耳、$2x10^{13}$VG/耳、$3x10^{13}$VG/耳、$4x10^{13}$VG/耳、$5x10^{13}$VG/耳、$6x10^{13}$VG/耳、$7x10^{13}$VG/耳、$8x10^{13}$VG/耳、$9x10^{13}$VG/耳、$1x10^{14}$VG/耳、$2x10^{14}$VG/耳、$3x10^{14}$VG/耳、$4x10^{14}$VG/耳、$5x10^{14}$VG/耳、$6x10^{14}$VG/耳、$7x10^{14}$VG/耳、$8x10^{14}$VG/耳、$9x10^{14}$VG/耳、$1x10^{15}$VG/耳或$2x10^{15}$VG/耳)。

[0077] 在前述方面中任一者的一些实施方案中,第一载体和第二载体以合起来足以用第一载体和第二载体两者转导至少20%的受试者的内毛细胞的量施用(例如,用两种载体转导至少20%、30%、40%、50%、60%、70%、80%、90%、95%或更多的受试者的内毛细胞)。

[0078] 在前述方面中任一者的一些实施方案中,双载体以包含药学上可接受的赋形剂的组合物形式施用。

[0079] 在前述方面中任一者的一些实施方案中,Myo15启动子包含第一区或由其组成,所述第一区与SEQ ID NO:24或其包括SEQ ID NO:26和/或SEQ ID NO:27的序列的功能性部分或衍生物具有至少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),所述第一区可操作地连接至第二区,所述第二区与SEQ ID NO:25或其包括SEQ ID NO:31和/或SEQ ID NO:32的序列的功能性部分或衍生物具有至少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),任选地含有接头,所述接头包含第一区与第二区之间的1至100个核苷酸(例如,1-5个、1-10个、1-15个、1-20个、1-25个、1-30个、1-35个、1-40个、1-45个、1-50个、1-60个、1-70个、1-80个、1-90个、10-20个、10-30个、10-40个、10-50个、10-60个、10-70个、10-80个、10-90个、10-100个、20-30个、20-40个、20-50个、20-60个、20-70个、20-80个、20-90个或20-100个核苷酸)。在一些实施方案中,第一区包含SEQ ID NO:24的序列或由其组成。在一些实施方案中,第二区包含SEQ ID NO:25的序列或由其组成。

[0080] 在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:36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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:36的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:38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)。在

前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:38的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:39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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:39的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:53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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:53的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:54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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:54的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:59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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:59的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:60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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:60的序列或由其组成。

[0081] 在前述方面中任一者的一些实施方案中,Myo15启动子包含第一区或由其组成,所述第一区与SEQ ID NO:25或其包括SEQ ID NO:31和/或SEQ ID NO:32的序列的功能性部分或衍生物具有至少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),所述第一区可操作地连接至第二区,所述第二区与SEQ ID NO:24或其包括SEQ ID NO:26和/或SEQ ID NO:27的序列的功能性部分或衍生物具有至少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),任选地含有接头,所述接头包含第一区与第二区之间的1至100个核苷酸(例如,1-5个、1-10个、1-15个、1-20个、1-25个、1-30个、1-35个、1-40个、1-45个、1-50个、1-60个、1-70个、1-80个、1-90个、10-20个、10-30个、10-40个、10-50个、10-60个、10-70个、10-80个、10-90个、10-100个、20-30个、20-40个、20-50个、20-60个、20-70个、20-80个、20-90个或20-100个核苷酸)。在一些实施方案中,第一区包含SEQ ID NO:25的序列或由其组成。在一些实施方案中,第二区包含SEQ ID NO:24的序列或由其组成。

[0082] 在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:37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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:37的序列或由其组成。在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:58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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:58的序列或由其组成。

[0083] 在前述方面中任一者的一些实施方案中,Myo15启动子包含与SEQ ID NO:24或其包括SEQ ID NO:26和/或SEQ ID NO:27的序列的功能性部分或衍生物具有至少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)的区域或由其组成。在一些实施方案中,所述区域包含SEQ ID NO:24的序列或由其组成。

[0084] 在前述方面中任一者的一些实施方案中,Myo15启动子包含与SEQ ID NO:25或其包括SEQ ID NO:31和/或SEQ ID NO:32的序列的功能性部分或衍生物具有至少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)的区域或由其组成。在一些实施方案中,所述区域包含SEQ ID NO:25的序列或由其组成。

[0085] 在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:26的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:27的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:26的序列和SEQ ID NO:27的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:28的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:29的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:30的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:24的功能性部分含有SEQ ID NO:50的序列。

[0086] 在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:31的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:32的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:51的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:51的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:31的序列和SEQ ID NO:32的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:33的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:34的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:35的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:25的功能性部分含有SEQ ID NO:55的序列。

[0087] 在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:50-58中任一者的核酸序列具有至少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)。在一些实施方案中,Myo15启动子包含SEQ ID NO:50的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:51的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:52的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:53的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:54的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:55的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:56的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:57的序列或由其组成。在一些实施方案中,Myo15启动子包含SEQ ID NO:58的序列或由其组成。

[0088] 在前述方面中任一者的一些实施方案中,Myo15启动子包含第一区或由其组成,所述第一区与SEQ ID NO:40或其包括SEQ ID NO:42的序列的功能性部分或衍生物具有至少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),所述第一区接合(例如,可操作地连接)至第二区,所述第二区与SEQ ID NO:41或其包括SEQ ID NO:43和/或SEQ ID NO:44的序列的功能性部分或衍生物具有至少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),任选地含有接头,所述接头包含第一区与第二区之间的1至400个核苷酸(例如,1-5个、1-10个、1-15个、1-20个、1-25个、1-30个、1-35个、1-40个、1-45个、1-50个、1-60个、1-70个、1-80个、1-90个、1-100个、1-125个、1-150个、1-175个、1-200个、1-225个、1-250个、1-275个、1-300个、1-325个、1-350个、1-375个、1-400个、10-20个、10-30个、10-40个、10-50个、10-60个、10-70个、10-80个、10-90个、10-100个、20-30个、20-40个、20-50个、20-60个、20-70个、20-80个、20-90个、20-100个、30-100个、40-100个、50-100个、50-150个、50-200个、50-250个、50-300个、50-350个、50-400个、100-150个、100-200个、100-250个、100-300个、100-350个、100-400个、150-200个、150-250个、150-300个、150-350个、150-400个、200-250个、200-300个、200-350个、200-400个、250-300个、250-350个、250-400个、300-400个或350-400个核苷酸)。在一些实施方案中,第一区包含SEQ ID NO:40的序列或由其组成。在一些实施方案中,第二区包含SEQ ID NO:41的序列或由其组成。

[0089] 在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:48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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:48的序列或由其组成。

[0090] 在前述方面中任一者的一些实施方案中,Myo15启动子与SEQ ID NO:49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)。在前述方面中任一者的一些实施方案中,Myo15启动子包含SEQ ID NO:49的序列或由其组成。

[0091] 在前述方面中任一者的一些实施方案中,Myo15启动子包含第一区或由其组成,所述第一区与SEQ ID NO:41或其包括SEQ ID NO:43和/或SEQ ID NO:44的序列的功能性部分或衍生物具有至少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),所述第一区接合(例如,可操作地连接)至第二区,所述第二区与SEQ ID NO:40或其包括SEQ ID NO:42的序列的功能性部分或衍生物具有至少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),任选地含有接头,所述接头包含第一区与第二区之间的1至400个核苷酸(例如,1-5个、1-10个、1-15个、1-20个、1-25个、1-30个、1-35个、1-40个、1-45个、1-50个、1-60个、1-70个、1-80个、1-90个、1-100个、1-125个、1-150个、1-175个、1-200个、1-225个、1-250个、1-275个、1-300个、1-325个、1-350个、1-375个、1-400个、10-20个、10-30个、10-40个、10-50个、10-60个、10-70个、10-80个、10-90个、10-100个、20-30个、20-40个、20-50个、20-60个、20-70个、20-80个、20-90个、20-100个、30-100个、40-100个、50-100个、50-150个、50-200个、50-250个、50-300个、50-

350个、50-400个、100-150个、100-200个、100-250个、100-300个、100-350个、100-400个、150-200个、150-250个、150-300个、150-350个、150-400个、200-250个、200-300个、200-350个、200-400个、250-300个、250-350个、250-400个、300-400个或350-400个核苷酸)。在一些实施方案中,第一区包含SEQ ID NO:41的序列或由其组成。在一些实施方案中,第二区包含SEQ ID NO:40的序列或由其组成。

[0092] 在前述方面中任一者的一些实施方案中,Myo15启动子包含与SEQ ID NO:40或其包括SEQ ID NO:42的序列的功能性部分或衍生物具有至少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)的区域或由其组成。在一些实施方案中,所述区域包含SEQ ID NO:40的序列或由其组成。

[0093] 在前述方面中任一者的一些实施方案中,Myo15启动子包含与SEQ ID NO:41或其包括SEQ ID NO:43和/或SEQ ID NO:44的序列的功能性部分或衍生物具有至少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)的区域或由其组成。在一些实施方案中,所述区域包含SEQ ID NO:41的序列或由其组成。

[0094] 在前述方面中任一者的一些实施方案中,SEQ ID NO:40的功能性部分含有SEQ ID NO:42的序列。

[0095] 在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:43的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:44的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:43的序列和SEQ ID NO:44的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:45的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:46的序列。在前述方面中任一者的一些实施方案中,SEQ ID NO:41的功能性部分含有SEQ ID NO:47的序列。

[0096] 在前述方面中任一者的一些实施方案中,当可操作地连接至转基因并引入毛细胞中时,Myo15启动子诱导转基因表达。

[0097] 定义

[0098] 如本文所用,术语“约”是指在所述值以上或以下10%内的值。

[0099] 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(例如组合物,其含有第一核酸载体,所述第一核酸载体含有编码耳畸蛋白的N末端部分的多核苷酸;和第二核酸载体,所述第二核酸载体含有编码耳畸蛋白的C末端部分的多核苷酸)。下文描述示例性施用途径。

[0100] 如本文所用,术语“双等位基因OTOF突变”是指其中突变存在于OTOF基因的两个等位基因(拷贝)中的状况。具有双等位基因OTOF突变的受试者可具有携带相同突变的两个OTOF等位基因,或者可在每个等位基因上具有不同的突变。

[0101] 如本文所用,短语“施用至内耳”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。施用至内耳的示例性途径包括施用至外淋巴或内淋巴中,诸如施用至或施用通过卵圆窗、圆窗或半规管(例如,水平半规管),或通过经鼓室或鼓室内注射施用,例如施用至毛细胞。

[0102] 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。

[0103] 如本文所用,术语“耳蜗毛细胞”是指内耳中参与感知声音的特化细胞的组。存在两种类型的耳蜗毛细胞:内毛细胞和外毛细胞。耳蜗毛细胞损伤和破坏耳蜗毛细胞功能的遗传突变与听力损失和耳聋相关。

[0104] 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。20种天然存在氨基酸中每一者的这些性质汇总于下表1中。

[0105] 表1:天然存在氨基酸的代表性物理化学性质

[0106]

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

[0107] 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)。

[0108] 如本文所用,术语“降解信号序列”是指介导包含其的多肽的降解的序列(例如,可翻译成氨基酸序列的核苷酸序列)。降解信号序列可包括在本发明的核酸载体中以减少或防止尚未经历重组和/或剪接的耳畸蛋白部分的表达。用于本发明的示例性降解信号序列是GCCTG CAAGAACTGGTTCAGCAGCCTGAGCCACTTCGTGATCCACCTG(SEQ ID NO:22)。

[0109] 如本文所用,术语本文所述的组合物、载体构建体或病毒载体的“有效量”、“治疗有效量”和“足量”是指在施用于有需要的受试者(包括哺乳动物,例如人类)时足以实现有益或期望的结果(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,在治疗感觉神经性听力损失的情形中,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)或所治疗的宿主等而变化,但仍可以常规方式由本领域技术人类员确定。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。应注意,当施用活性成分的组合时,组合的有效量可包括或可不包括在个别施用时有效的每一成分的量。如本文所定义,本公开的组合物、载体构建体、病毒载体或细胞的治疗有效量可容易地由本领域技术人类员通过本领域中已知的常规方法确定。可调整给药方案以提供最佳治疗反应。

[0110] 如本文所用,术语“内源”描述了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,例如人类耳蜗毛细胞)的分子(例如多肽、核酸或辅因子)。

[0111] 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;和(4)多肽或蛋白质的翻译后修饰。

[0112] 如本文所用,术语“外源”描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,例如人类耳蜗毛细胞)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。

[0113] 如本文所用,术语“毛细胞特异性表达”是指与内耳的其他细胞类型(例如螺旋神经节神经元、神经胶质或其他内耳细胞类型)相比,主要在毛细胞(例如耳蜗毛细胞)内产生RNA转录物或多肽。转基因的毛细胞特异性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(例如毛细胞对非毛细胞)的转基因表达(例如RNA或蛋白质表达)来确认。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,毛细胞特异性启动子在毛细胞(例如耳蜗毛细胞)中诱导的可操作地连接至其的转基因的表达(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):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、螺旋神经节神经元、施旺细胞(Schwann cell)。

[0114] 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,相对于施用前标志物的量,如本文所述度量(例如OTOF表达或听性脑干反应)的标志物的量在受试者中可增加或

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。

[0115] 如本文所用，术语“内含子”是指基因的编码区内的区域，其核苷酸序列不会翻译成相应蛋白质的氨基酸序列。术语内含子还指RNA的从基因转录的相应区域。内含子转录成前mRNA，但在加工期间去除，并且不包括在成熟mRNA中。

[0116] 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。

[0117] 如本文所用，术语“可操作地连接”是指可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。术语“可操作地连接”包括两种或更多种组分(例如启动子和另一序列元件)的并置，使得两种组分正常发挥功能并允许至少一种组分可介导对其他组分中的至少一者发挥的功能的可能性。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。在其他实施方案中，如果接合转录调控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。两个转录调控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。

[0118] 如本文所用，术语“耳畸蛋白”和“OTOF”是指与非综合征性隐性耳聋DNFB9相关的基因。术语“耳畸蛋白”和“OTOF”还指野生型OTOF蛋白的变体和编码其的核酸，诸如与野生型OTOF蛋白(例如，SEQ ID NO:1-5中任一者)的氨基酸序列具有至少85％序列同一性(例如，85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或99.9％同一性或更大)的变体蛋白或与野生型OTOF基因的核酸序列具有至少85％序列同一性(例如，85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或99.9％同一性或更大)的多核苷酸，条件是所编码的OTOF类似物保留野生型OTOF的治疗功能。如本文所用，OTOF可指定位于内毛细胞的蛋白质或编码此蛋白质的基因，这取决于上下文，如本领域技术人类员将理解的。

[0119] 如本文所用，术语“耳畸蛋白同功型5”和“OTOF同功型5”是指与非综合征性隐性耳聋DFNB9相关的基因的同功型。基因的人类同功型与参照序列NM_001287489相关，并且转录物包括人类耳畸蛋白的外显子1-45和47，但缺少OTOF基因的外显子46。人类OTOF同功型5蛋白也称为耳畸蛋白同功型e。术语“耳畸蛋白同功型5”和“OTOF同功型5”也指野生型OTOF同功型5蛋白和编码其的多核苷酸的变体，例如与野生型OTOF同功型5蛋白的氨基酸序列(例如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99.9％或更大同一性)的变体蛋白或与野生型OTOF同功型5基因的多核苷酸序列具有至少85％序列同一性(例如85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或99.9％或更大同一性)的多核苷酸，条件是所编码的OTOF同功型5类似物保留野生型OTOF同功型5的治疗功能。OTOF同功型5蛋白变体相对于野生型OTOF同功型5(例如SEQ ID NO:1)可具有一个或多个

(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,条件是OTOF同功型5变体保留野生型OTOF同功型5的治疗功能并且在氨基酸序列的N末端部分中具有不超过10%的氨基酸取代并在氨基酸序列的C末端部分中具有不超过10%的氨基酸取代。如本文所用,OTOF同功型5可指定位于内毛细胞的蛋白质或编码此蛋白质的基因,这根据上下文而定,如本领域技术人类员将了解。OTOF同功型5可指人类OTOF同功型5或来自另一哺乳动物物种的同源物。鼠类耳畸蛋白相对于人类耳畸蛋白含有一个额外外显子(在鼠类耳畸蛋白中有48个外显子),并且对应于编码人类OTOF同功型5的外显子的鼠类耳畸蛋白外显子为1-5、7-46和48。本文所用的外显子编号惯例是基于外显子目前理解为存在于人类OTOF的共有转录物中。

[0120] 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)可在引入宿主细胞中后整合至宿主细胞的基因组中,并且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。

[0121] 如本文所用,术语"核酸"和"多核苷酸"在本文中可互换使用并且是指任何长度的核苷的聚合形式。通常,多核苷酸是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。然而,所述术语涵盖含有核苷或含有经化学或生物修饰的碱基、经修饰主链等的核苷类似物的分子,无论是否发现于天然存在核酸中,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。

[0122] 如本文所用,术语核酸的"互补性"或"互补"意指核酸的一条链中的核苷酸序列因其核碱基组的取向而与相对核酸链上的另一序列形成氢键。DNA中的互补碱基通常是A与T和C与G。在RNA中,其通常是C与G和U与A。互补可为完全的或实质的/充分的。两个核酸之间的完全互补意指两个核酸可形成双链体,其中双链体中的每个碱基通过沃森-克里克配对(Watson-Crick pairing)键结至互补碱基。"实质"或"充分"互补意指一条链中的序列与相对链中的序列不完全(completely和/或perfectly)互补,但在一组杂交条件(例如盐浓度和温度)下在两条链之间的碱基之间发生充分键结以形成稳定的杂合复合物。此类条件可通过使用预测杂交链的Tm(熔融温度)的序列和标准数学计算或通过使用常规方法凭经验确定Tm来预测。Tm包括在两个核酸链之间形成的杂交复合物群体为50%变性(即双链核酸分子群体变得半解离成单链)时的温度。在低于Tm的温度下,形成杂交复合物是有利的,而在高于Tm的温度下,杂交复合物中链的熔融或分离是有利的。可通过使用例如Tm=81.5+0.41(G+C%)估计在1M NaCl水溶液具有已知G+C含量的核酸的Tm,尽管其他已知Tm计算将核酸结构特征考虑在内。

[0123] 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。适用于本文所述的组合物和方法的示例性启动子包括普遍存在的启动子(例如,CAG启动子、巨细胞病毒(CMV)启动子和嵌合CMV-鸡β-肌动蛋白启动子(CBA)的截短

形式,其中杂合鸡β-肌动蛋白/兔β-珠蛋白内含子被大大缩短以产生称为smCBA的较小型式的启动子)、耳蜗毛细胞特异性启动子(例如,肌凝蛋白15(Myo15)启动子、肌凝蛋白7A(Myo7A)启动子、肌凝蛋白6(Myo6)启动子、POU 4类同源盒3(POU4F3)启动子)和内毛细胞特异性启动子(例如,成纤维细胞生长因子8(FGF8)启动子、囊泡谷氨酸转运蛋白3(VGLUT3)启动子和OTOF启动子)。

[0124] 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类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。本领域技术人类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:

[0125] 100×(分数X/Y)

[0126] 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%。

[0127] 如本文所用的术语"衍生物"是指与相应全长野生型核酸、肽或蛋白质相比包含一个或多个突变和/或化学修饰的核酸、肽或蛋白质或其变体或类似物。涉及核酸的化学修饰的非限制性实例包括例如对碱基部分、糖部分、磷酸部分、磷酸-糖主链或其组合的修饰。

[0128] 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。

[0129] 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。优选地,术语"药学上可接受"意指经联邦或州政府管理机构批准或列于美国药典(U.S.Pharmacopeia)或其他公认药典中用于哺乳动物并且更特别是用于人类中。

[0130] 如本文所用,术语"重组基因区"是指介导两个不同序列之间的重组的同源区域。

[0131] 如本文所用,术语"调控序列"包括启动子、增强子和控制编码OTOF的多核苷酸的转录或翻译的其他表达控制元件(例如多聚腺苷酸化信号)。此类调控序列描述于例如Goeddel,Gene Expression Technology:Methods in Enzymology 185(Academic Press,San Diego,CA,1990)中;所述文献通过引用并入本文。

[0132] 如本文所用,术语"样品"是指从受试者分离的样本(例如血液、血液组分(例如血清或血浆)、尿液、唾液、羊水、脑脊髓液、组织(例如胎盘或皮肤)、胰液、绒毛膜绒毛样品和细胞)。

[0133] 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、核转染(Nucleofection)、挤压穿孔、声致穿孔、光学转染、磁性转染、刺穿转染等。

[0134] 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)、兽医动物(例如猫、狗、牛、马、绵羊、猪等)和实验动物疾病模型(例如小鼠、大鼠)。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听力损失(例如与OTOF突变相关的听力损失)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。本领域技术人类员应理解,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。

[0135] 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且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。

[0136] 如本文所用,状态、病症或疾患的“治疗(treatment)”和“治疗(treating)”可包括:(1)预防、延迟或减少在受试者中发展的状态、病症或疾患的至少一种临床或亚临床症状的出现的发生和/或可能性,所述受试者可能患有或易患所述状态、病症或疾患,但尚未经历或显示所述状态、病症或疾患的临床或亚临床症状;或(2)抑制状态、病症或疾患,即停止、减少或延迟疾病或其复发或其至少一种临床或亚临床症状的发展;或(3)减轻疾病,即使状态、病症或疾患或其临床或亚临床症状中的至少一者消退。对于待治疗受试者的益处是统计学上显著的或至少为患者或医师可察觉的。

[0137] 如本文所用,术语“载体”包括核酸载体,例如DNA载体,例如质粒、RNA载体、病毒或其他适宜复制子(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公开于例如WO94/11026中;所述文献关于适于表达目标基因的载体通过引用并入本文。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OTOF的某些载体包括含有引导基因转录的调控序列(例如启动子和增强子区域)的载体。可用于表达OTOF的其他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mRN A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。适宜标志物的实例包括编码抗生素(例如氨苄青霉素(ampicillin)、氯霉素(chloramphenico l)、卡那霉素(kanamycin)或诺尔丝菌素(nourseothricin))抗性的基因。

[0138] 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

[0139] 图1是显示在32或52周龄时治疗的纯合OTOF-Q828X突变型小鼠中ABR阈值恢复的图。在异氟烷麻醉下通过圆窗向动物给予媒介物(n=5/年龄组)或双杂合AAV-Myo15-hOTOF

载体(n＝10/年龄组)。在治疗后4周,在用OTOF双载体治疗的10/10的32周龄动物和9/10的52周龄动物中观察到ABR恢复。

[0140] 图2A-2B是显示纯合(Otof-Q828X hom)和杂合(Otof-Q828X het)Otof-Q828X小鼠中随时间推移的内毛细胞和外毛细胞数量的一系列图。IHC(图2A)和OHC(图2B)的数量在5至42周龄动物的50只小鼠耳中计数。显示了对应于5.6kHz、8kHz、11.3kHz、16kHz、22.6kHz、32kHz和45.2kHz的耳蜗区域中的计数。对于Otof-Q828X hom小鼠中的所有测试频率,IHC计数随年龄增加而出现统计学显著性损失。对于更低频率,在het小鼠中观察到IHC计数的类似趋势(Kendall秩相关)。Otof-828X hom和het动物中的IHC计数在16周内保持稳定。从16周后开始,Otof-Q828X hom动物显示在22.6-45.2kHz下并且在24周后在更低频率(8-16kHz)下开始IHC计数的减少。对于16和32kHz,在24周后Otof-Q828X het小鼠的IHC计数开始减少。在32周后,对于大多数测试频率(<45.2kHz),保留了超过75%的IHC(图2A)。在所有频率内,Otof-Q828X hom和het小鼠中的外毛细胞数量在6个月的研究时间过程中保持恒定,但8kHz除外,在8kHz下het小鼠显示计数的年龄相关性降低(Kendall秩相关)。5.6kHz和45.2kHz中的OHC计数与较大变异性相关,显示了het与hom之间的计数差异,所述差异在所测试的年龄内散布。大多数OHC在32周后仍保留(图2B)。

[0141] 图3是显示在纯合OTOF-Q828X突变型小鼠中ABR阈值恢复的图。将在22.6kHz下测量的ABR阈值相对于在用OTOF双载体系统治疗的成年纯合OTOF-Q828X突变型小鼠的多项研究中表达耳畸蛋白的内毛细胞(IHC)的百分比进行作图。在治疗后≥4周进行测量。

具体实施方式

[0142] 本文描述了通过向至少25岁(例如25-50岁、25-45岁、25-40岁、25-35岁、25-30岁、30-50岁、30-45岁、30-40岁、30-35岁、35-50岁、35-45岁、35-40岁、40-50岁、40-45岁或45-50岁,例如25岁、26岁、27岁、28岁、29岁、30岁、31岁、32岁、33岁、34岁、35岁、36岁、37岁、38岁、39岁、40岁、41岁、42岁、43岁、44岁、45岁、46岁、47岁、48岁、49岁或50岁)的人类受试者施用含有启动子和编码耳畸蛋白(OTOF)蛋白(例如野生型(WT)OTOF蛋白)的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸和聚腺苷酸化(poly(A))序列的第二核酸载体来治疗所述受试者中由于双等位基因耳畸蛋白(OTOF)突变所致的感觉神经性听力损失或听神经病变的组合物和方法。当引入至哺乳动物细胞(诸如耳蜗毛细胞)中时,由两种核酸载体编码的多核苷酸可组合形成编码全长OTOF蛋白的多核苷酸。因此,本文所述的组合物和方法可用于诱导或增加具有OTOF缺陷(例如,OTOF中的纯合或复合杂合突变)的受试者的耳蜗毛细胞中WT OTOF的表达。本文所述的组合物和方法还可用于治疗具有双等位基因OTOF突变的受试者,所述受试者被鉴定为具有可检测的耳声发射、可检测的耳蜗微音效应和/或可检测的总和电位。

[0143] 耳畸蛋白

[0144] OTOF是含有至少六个参与钙、磷脂和蛋白质结合的C2结构域的230kDa膜蛋白。其由含有48个外显子的基因编码,并且全长蛋白质由1,997个氨基酸构成。OTOF位于内毛细胞中的带状突触处,其中认为其用作突触囊泡融合物中的钙传感器,从而触发含有神经传递质的囊泡与质膜的融合。其还已参与囊泡补充和格形蛋白介导的胞吞作用,并且已显示与肌凝蛋白VI、Rab8b、SNARE蛋白、钙通道Cav1.3、Ergic2和AP-2相互作用。OTOF介导胞吐作用

的机制和其与其结合伴侣相互作用的生理意义仍待确定。

[0145] 耳畸蛋白相关的听力损失

[0146] OTOF首次是由研究耳聋的非综合征形式常染色体隐性耳聋-9(DFNB9)的遗传学的研究鉴定出。此后,在世界各地的患者中发现OTOF突变引起感觉神经性听力损失,其中许多携带OTOF突变的患者患有听神经病变,其是内耳检测到声音、但无法适当地将声音从耳传递至脑的病症。这些患者具有异常听性脑干反应(ABR)和受损的言语识别和最初正常的耳声发射。携带OTOF中的纯合或复合杂合突变的患者通常在儿童早期患上听力损失,并且已发现听力损害的严重程度随着OTOF突变的位置和类型而变化。已经鉴定了OTOF中的至少220个突变,包括引起截短的突变和不引起截短的突变。

[0147] 本发明部分地基于以下发现:向成年(32周龄)和中年(52周龄)耳畸蛋白缺陷型小鼠施用含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体在挽救听力损失方面是有效的。这些数据表明,将耳畸蛋白递送至成年和中年耳畸蛋白缺陷型小鼠能够恢复听力,尽管随着耳畸蛋白缺陷型小鼠衰老而发生毛细胞损失,并且表明耳畸蛋白基因疗法也可用于治疗类似年龄的人类受试者(32周龄和52周龄小鼠大致对应于30-50岁的人类受试者)。人类也经历年龄依赖性毛细胞损失,这预期将限制基因治疗方法在老年人类中的功效。然而,本发明人类还发现,当约20%的内毛细胞表达耳畸蛋白时,耳畸蛋白缺陷型小鼠的听力恢复,表明即使转导相对小百分比的内毛细胞,听力也可得到挽救。总之,这些数据表明可使用编码OTOF的双载体系统来治疗具有双等位基因OTOF突变的成人类受试者(例如,年龄为25岁或更大的人类受试者,诸如25-50岁、25-45岁、25-40岁、25-35岁、25-30岁、30-50岁、30-45岁、30-40岁、30-35岁、35-50岁、35-45岁、35-40岁、40-50岁、40-45岁或45-50岁,例如25岁、26岁、27岁、28岁、29岁、30岁、31岁、32岁、33岁、34岁、35岁、36岁、37岁、38岁、39岁、40岁、41岁、42岁、43岁、44岁、45岁、46岁、47岁、48岁、49岁或50岁)。

[0148] 本文所述的组合物和方法可用于通过施用含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体来治疗由双等位基因OTOF突变引起的感觉神经性听力损失或听神经病变。全长OTOF编码序列因太大而无法包括在常用于基因疗法的载体类型(例如腺相关病毒(AAV)载体,认为其具有5kb的包装限值)中。本文所述的组合物和方法通过将OTOF编码序列分在可重组于细胞中以重构全长OTOF序列的两个不同核酸载体之间克服了此问题。这些组合物和方法可用于治疗具有一个或多个OTOF基因突变(例如减少OTOF表达、降低OTOF功能或与听力损失相关的OTOF突变)的受试者。当第一核酸载体和第二核酸载体是以组合物形式施用时,编码OTOF的N末端和C末端部分的多核苷酸可组合于细胞(例如人类细胞,例如耳蜗毛细胞)内以形成含有全长OTOF编码序列的单个核酸分子(例如经由同源重组和/或剪接)。

[0149] 用于本文所述的组合物和方法中的核酸载体包括编码野生型OTOF或其变体的核酸序列,例如在组合时编码与野生型人类或小鼠OTOF的氨基酸序列具有至少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)的蛋白质的核酸序列。用于本文所述核酸载体中的多核苷酸编码下表2中OTOF氨基酸序列的N末端部分和C末端部分(例如在组合时编码表2中所列出的全长OTOF氨基酸序列(例如SEQ ID NO:1-5中任一者)的两个部分)。

[0150] 根据本文所述的方法,可向受试者施用组合物,所述组合物含有第一核酸载体和第二核酸载体,所述第一核酸载体和第二核酸载体分别含有编码SEQ ID NO:1-5中任一者的氨基酸序列的多核苷酸序列或编码与SEQ ID NO:1-5中任一者的氨基酸序列具有至少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)的氨基酸序列的多核苷酸序列或编码相对于SEQ ID NO:1-5中的任一者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个或10个或更多个保守氨基酸取代)的氨基酸序列的多核苷酸序列的N末端和C末端部分,条件是所编码的OTOF类似物保留野生型OTOF的治疗功能(例如调节带状突触处的胞吐作用或拯救或改进与耳畸蛋白基因缺陷(例如OTOF突变)相关的听力损失动物模型中的ABR反应的能力)。OTOF蛋白的N末端部分中不超过10%的氨基酸和OTOF蛋白的C末端部分中不超过10%的氨基酸可经保守氨基酸取代代替。OTOF蛋白可由具有SEQ ID NO:10-14中任一者的序列的多核苷酸编码。OTOF蛋白也可由具有已发现在人类受试者中为非致病性的单核苷酸变体(SNV)的多核苷酸编码。OTOF蛋白可为人类OTOF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、狗、兔、天竺鼠或其他哺乳动物)的人类OTOF蛋白的同源物。在一些实施方案中,所编码的OTOF蛋白具有SEQ ID NO:1(OTOF同功型1)的序列。在一些实施方案中,所编码的OTOF蛋白具有SEQ ID NO:5(OTOF同功型5)的序列。

[0151] 表2:OTOF序列

[0152]

SEQ ID NO.	序列名称	序列
1	OTOF-201蛋白(NP_919224.1),人类耳畸蛋白同功型a,也称为人类OTOF同功型1,1997 aa	MALLIHLKTVSELRGRGDRIAKVTFRGQSFYSRVLE NCEDVADFDETFRWPVASSIDRNEMLEIQVFNYSKV FSNKLIGTFRMVLQKVVEESHVEVTDTLIDDNNAIIK TSLCVEVRYQATDGTVGSWDDGDFLGDESLQEEEK DSQETDGLLPGSRPSSRPPGEKSFRRAGRSVFSAMK LGKNRSHKEEPQRPDEPAVLEMEDLDHLAIRLGDG LDPDSVSLASVTALTTNVSNKRSKPDIKMEPSAGRP MDYQVSITVIEARQLVGLNMDPVVCVEVGDDKKY TSMKESTNCPYYNEYFVFDFHVSPDVMFDKIIKISVI HSKNLLRSGTLVGSFKMDVGTVYSQPEHQFHHKW AILSDPDDISSGLKGYVKCDVAVVGKGDNIKTPHKA NETDEDDIEGNLLLPEGVPPERQWARFYVKIYRAEG LPRMNTSLMANVKKAFIGENKDLVDPYVQVFFAGQ KGKTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQ IRDSDKVNDVAIGTHFIDLRKISNDGDKGFLPTLGPA WVNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARL LLGLAVEIVDTSNPELTSSTEVQVEQATPISESCAGK MEEFFLFGAFLEASMIDRRNGDKPITFEVTIGNYGNE VDGLSRPQRPRPRKEPGDEEEVDLIQNASDDEAGDA GDLASVSSTPPMRPQVTDRNYFHLPYLERKPCIYIKS WWPDQRRRLYNANIMDHIADKLEEGLNDIQEMIKT

[0153]

SEQ ID NO.	序列名称	序列
		EKSYPERRLRGVLEELSCGCCRFLSLADKDQGHSSR TRLDRERLKSCMRELENMGQQARMLRAQVKRHTV RDKLRLCQNFLQKLRFLADEPQHSIPDIFIWMMSNN KRVAYARVPSKDLLFSIVEEETGKDCAKVKTLFLKL PGKRGFGSAGWTVQAKVELYLWLGLSKQRKEFLC GLPCGFQEVKAAQGLGLHAFPPVSLVYTKKQAFQL RAHMYQARSLFAADSSGLSDPFARVFFINQSQCTEV LNETLCPTWDQMLVFDNLELYGEAHELRDDPPIIVI EIYDQDSMGKADFMGRTFAKPLVKMADEAYCPPRF PPQLEYYQIYRGNATAGDLLAAFELLQIGPAGKADL PPINGPVDVDRGPIMPVPMGIRPVLSKYRVEVLFWG LRDLKRVNLAQVDRPRVDIECAGKGVQSSLIHNYK KNPNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAF GRYTLVGSHAVSSLRRFIYRPPDRSAPSWNTTVRLL RRCRVLCNGGSSSHSTGEVVVTMEPEVPIKKLETMV KLDATSEAVVKVDVAEEEKEKKKKKKGTAEEPEEE EPDESMLDWWSKYFASIDTMKEQLRQQEPSGIDLEE KEEVDNTEGLKGSMKGKEKARAAKEEKKKKTQSS GSGQGSEAPEKKKPKIDELKVYPKELESEFDNFEDW LHTFNLLRGKTGDDEDGSTEEERIVGRFKGSLCVYK VPLPEDVSREAGYDSTYGMFQGIPSNDPINVLVRVY VVRATDLHPADINGKADPYIAIRLGKTDIRDKENYIS KQLNPVFGKSFDIEASFPMESMLTVAVYDWDLVGT DDLIGETKIDLENRFYSKHRATCGIAQTYSTHGYNI WRDPMKPSQILTRLCKDGKVDGPHFGPPGRVKVAN RVFTGPSEIEDENGQRKPTDEHVALLALRHWEDIPR AGCRLVPEHVETRPLLNPDKPGIEQGRLELWVDMFP MDMPAPGTPLDISPRKPKKYELRVIIWNTDEVVLED DDFFTGEKSSDIFVRGWLKGQQEDKQDTDVHYHSL TGEGNFNWRYLFPFDYLAAEEKIVISKKESMFSWDE TEYKIPARLTLQIWDADHFSADDFLGAIELDLNRFPR GAKTAKQCTMEMATGEVDVPLVSIFKQKRVKGWW PLLARNENDEFELTGKVEAELHLLTAEEAEKNPVGL ARNEPDPLEKPNRPDTSFIWFLNPLKSARYFLWHTY RWLLLKLLLLLLLLLLLALFLYSVPGYLVKKILGA
2	OTOF-202蛋白(NP_004793.2)，人类耳畴蛋白同功型b，1230 aa	MIKTEKSYPERRLRGVLEELSCGCCRFLSLADKDQG HSSRTRLDRERLKSCMRELENMGQQARMLRAQVK RHTVRDKLRLCQNFLQKLRFLADEPQHSIPDIFIWM MSNNKRVAYARVPSKDLLFSIVEEETGKDCAKVKT LFLKLPGKRGFGSAGWTVQAKVELYLWLGLSKQR KEFLCGLPCGFQEVKAAQGLGLHAFPPVSLVYTKK QAFQLRAHMYQARSLFAADSSGLSDPFARVFFINQS QCTEVLNETLCPTWDQMLVFDNLELYGEAHELRDD PPIIVIEIYDQDSMGKADFMGRTFAKPLVKMADEAY CPPRFPPQLEYYQIYRGNATAGDLLAAFELLQIGPA GKADLPPINGPVDVDRGPIMPVPMGIRPVLSKYRVE VLFWGLRDLKRVNLAQVDRPRVDIECAGKGVQSSL IHNYKKNPNFNTLVKWFEVDLPENELLHPPLNIRVV

SEQ ID NO.	序列名称	序列
		DCRAFGRYTLVGSHAVSSLRRFIYRPPDRSAPSWNT TGEVVVTMEPEVPIKKLETMVKLDATSEAVVKVDV AEEEKEKKKKKKGTAEEPEEEEPDESMLDWWSKYF ASIDTMKEQLRQQEPSGIDLEEKEEVDNTEGLKGSM KGKEKARAAKEEKKKKTQSSGSGQGSEAPEKKKPK IDELKVYPKELESEFDNFEDWLHTFNLLRGKTGDDE DGSTEEERIVGRFKGSLCVYKVPLPEDVSREAGYDS TYGMFQGIPSNDPINVLVRVYVVRATDLHPADINGK ADPYIAIRLGKTDIRDKENYISKQLNPVFGKSFDIEAS FPMESMLTVAVYDWDLVGTDDLIGETKIDLENRFY SKHRATCGIAQTYSTHGYNIWRDPMKPSQILTRLCK DGKVDGPHFGPPGRVKVANRVFTGPSEIEDENGQR KPTDEHVALLALRHWEDIPRAGCRLVPEHVETRPLL NPDKPGIEQGRLELWVDMFPMDMPAPGTPLDISPRK PKKYELRVIIWNTDEVVLEDDDFFTGEKSSDIFVRG WLKGQQEDKQDTDVHYHSLTGEGNFNWRYLFPFD YLAAEEKIVISKKESMFSWDETEYKIPARLTLQIWD ADHFSADDFLGAIELDLNRFPRGAKTAKQCTMEMA TGEVDVPLVSIFKQKRVKGWWPLLARNENDEFELT GKVEAELHLLTAEEAEKNPVGLARNEPDPLEKPNRP DTSFIWFLNPLKSARYFLWHTYRWLLLKLLLLLLLL LLLALFLYSVPGYLVKKILGA
3	OTOF-203蛋白(NP_919304.1)，人类耳畴蛋白同功型d，1230 aa	MIKTEKSYPERRLRGVLEELSCGCCRFLSLADKDQG HSSRTRLDRERLKSCMRELENMGQQARMLRAQVK RHTVRDKLRLCQNFLQKLRFLADEPQHSIPDIFIWM MSNNKRVAYARVPSKDLLFSIVEEETGKDCAKVKT LFLKLPGKRGFGSAGWTVQAKVELYLWLGLSKQR KEFLCGLPCGFQEVKAAQGLGLHAFPPVSLVYTKK QAFQLRAHMYQARSLFAADSSGLSDPFARVFFINQS QCTEVLNETLCPTWDQMLVFDNLELYGEAHELRDD PPIIVIEIYDQDSMGKADFMGRTFAKPLVKMADEAY CPPRFPPQLEYYQIYRGNATAGDLLAAFELLQIGPA GKADLPPINGPVDVDRGPIMPVPMGIRPVLSKYRVE VLFWGLRDLKRVNLAQVDRPRVDIECAGKGVQSSL IHNYKKNPNFNTLVKWFEVDLPENELLHPPLNIRVV DCRAFGRYTLVGSHAVSSLRRFIYRPPDRSAPSWNT TGEVVVTMEPEVPIKKLETMVKLDATSEAVVKVDV AEEEKEKKKKKKGTAEEPEEEEPDESMLDWWSKYF ASIDTMKEQLRQQEPSGIDLEEKEEVDNTEGLKGSM KGKEKARAAKEEKKKKTQSSGSGQGSEAPEKKKPK IDELKVYPKELESEFDNFEDWLHTFNLLRGKTGDDE DGSTEEERIVGRFKGSLCVYKVPLPEDVSREAGYDS TYGMFQGIPSNDPINVLVRVYVVRATDLHPADINGK ADPYIAIRLGKTDIRDKENYISKQLNPVFGKSFDIEAS FPMESMLTVAVYDWDLVGTDDLIGETKIDLENRFY SKHRATCGIAQTYSTHGYNIWRDPMKPSQILTRLCK DGKVDGPHFGPPGRVKVANRVFTGPSEIEDENGQR KPTDEHVALLALRHWEDIPRAGCRLVPEHVETRPLL

[0154]

[0155]

SEQ ID NO.	序列名称	序列
		NPDKPGIEQGRLELWVDMFPMDMPAPGTPLDISPRK PKKYELRVIIWNTDEVVLEDDDFFTGEKSSDIFVRG WLKGQQEDKQDTDVHYHSLTGEGNFNWRYLFPFD YLAAEEKIVISKKESMFSWDETEYKIPARLTLQIWD ADHFSADDFLGAIELDLNRFPRGAKTAKQCTMEMA TGEVDVPLVSIFKQKRVKGWWPLLARNENDEFELT GKVEAELHLLTAEEAEKNPVGLARNEPDPLEKPNRP DTAFVWFLNPLKSIKYLICTRYKWLIIKIVLALLGLL MLGLFLYSLPGYMVKKLLGA
4	OTOF-208 蛋白 (NP_919303 .1)，人类耳畸蛋白同功型 c，1307 aa	MMTDTQDGPSESSQIMRSLTPLINREEAFGEAGEAG LWPSITHTPDSQEEGLNDIQEMIKTEKSYPERRLRGV LEELSCGCCRFLSLADKDQGHSSRTRLDRERLKSCM RELENMGQQARMLRAQVKRHTVRDKLRLCQNFLQ KLRFLADEPQHSIPDIFIWMMSNNKRVAYARVPSKD LLFSIVEEETGKDCAKVKTLFLKLPGKRGFGSAGWT VQAKVELYLWLGLSKQRKEFLCGLPCGFQEVKAAQ GLGLHAFPPVSLVYTKKQAFQLRAHMYQARSLFAA DSSGLSDPFARVFFINQSQCTEVLNETLCPTWDQML VFDNLELYGEAHELRDDPPIIVIEIYDQDSMGKADF MGRTFAKPLVKMADEAYCPPRFPPQLEYYQIYRGN ATAGDLLAAFELLQIGPAGKADLPPINGPVDVDRGP IMPVPMGIRPVLSKYRVEVLFWGLRDLKRVNLAQV DRPRVDIECAGKGVQSSLIHNYKKNPNFNTLVKWF EVDLPENELLHPPLNIRVVDCRAFGRYTLVGSHAVS SLRRFIYRPPDRSAPSWNTTVRLLRRCRVLCNGGSSS HSTGEVVVTMEPEVPIKKLETMVKLDATSEAVVKV DVAEEEKEKKKKKKGTAEEPEEEEPDESMLDWWS KYFASIDTMKEQLRQQEPSGIDLEEKEEVDNTEGLK GSMKGKEKARAAKEEKKKKTQSSGSGQGSEAPEK KKPKIDELKVYPKELESEFDNFEDWLHTFNLLRGKT GDDEDGSTEEERIVGRFKGSLCVYKVPLPEDVSREA GYDSTYGMFQGIPSNDPINVLVRVYVVRATDLHPA DINGKADPYIAIRLGKTDIRDKENYISKQLNPVFGKS FDIEASFPMESMLTVAVYDWDLVGTDDLIGETKIDL ENRFYSKHRATCGIAQTYSTHGYNIWRDPMKPSQIL TRLCKDGKVDGPHFGPPGRVKVANRVFTGPSEIEDE NGQRKPTDEHVALLALRHWEDIPRAGCRLVPEHVE TRPLLNPDKPGIEQGRLELWVDMFPMDMPAPGTPL DISPRKPKKYELRVIIWNTDEVVLEDDDFFTGEKSSD IFVRGWLKGQQEDKQDTDVHYHSLTGEGNFNWRY LFPFDYLAAEEKIVISKKESMFSWDETEYKIPARLTL QIWDADHFSADDFLGAIELDLNRFPRGAKTAKQCT MEMATGEVDVPLVSIFKQKRVKGWWPLLARNEND EFELTGKVEAELHLLTAEEAEKNPVGLARNEPDPLE KPNRPDTSFIWFLNPLKSARYFLWHTYRWLLLKLLL LLLLLLLLALFLYSVPGYLVKKILGA
5	OTOF-205 蛋白	MALLIHLKTVSELRGRGDRIAKVTFRGQSFYSRVLE NCEDVADFDETFRWPVASSIDRNEMLEIQVFNYSKV

[0156]

SEQ ID NO.	序列名称	序列
	(NP_001274418.1)，人类耳畸蛋白同功型 e，1997 aa，也称为人类 OTOF 同功型 5	FSNKLIGTFRMVLQKVVEESHVEVTDTLIDDNNAIIK TSLCVEVRYQATDGTVGSWDDGDFLGDESLQEEEK DSQETDGLLPGSRPSSRPPGEKSFRRAGRSVFSAMK LGKNRSHKEEPQRPDEPAVLEMEDLDHLAIRLGDG LDPDSVSLASVTALTTNVSNKRSKPDIKMEPSAGRP MDYQVSITVIEARQLVGLNMDPVVCVEVGDDKKY TSMKESTNCPYYNEYFVFDFHVSPDVMFDKIIKISVI HSKNLLRSGTLVGSFKMDVGTVYSQPEHQFHHKW AILSDPDDISSGLKGYVKCDVAVVGKGDNIKTPHKA NETDEDDIEGNLLLPEGVPPERQWARFYVKIYRAEG LPRMNTSLMANVKKAFIGENKDLVDPYVQVFFAGQ KGKTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQ IRDSDKVNDVAIGTHFIDLRKISNDGDKGFLPTLGPA WVNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARL LLGLAVEIVDTSNPELTSSTEVQVEQATPISESCAGK MEEFFLFGAFLEASMIDRRNGDKPITFEVTIGNYGNE VDGLSRPQRPRPRKEPGDEEEVDLIQNASDDEAGDA GDLASVSSTPPMRPQVTDRNYFHLPYLERKPCIYIKS WWPDQRRRLYNANIMDHIADKLEEGLNDIQEMIKT EKSYPERRLRGVLEELSCGCCRFLSLADKDQGHSSR TRLDRERLKSCMRELENMGQQARMLRAQVKRHTV RDKLRLCQNFLQKLRFLADEPQHSIPDIFIWMMSNN KRVAYARVPSKDLLFSIVEEETGKDCAKVKTLFLKL PGKRGFGSAGWTVQAKVELYLWLGLSKQRKEFLC GLPCGFQEVKAAQGLGLHAFPPVSLVYTKKQAFQL RAHMYQARSLFAADSSGLSDPFARVFFINQSQCTEV LNETLCPTWDQMLVFDNLELYGEAHELRDDPPIIVI EIYDQDSMGKADFMGRTFAKPLVKMADEAYCPPRF PPQLEYYQIYRGNATAGDLLAAFELLQIGPAGKADL PPINGPVDVDRGPIMPVPMGIRPVLSKYRVEVLFWG LRDLKRVNLAQVDRPRVDIECAGKGVQSSLIHNYK KNPNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAF GRYTLVGSHAVSSLRRFIYRPPDRSAPSWNTTVRLL RRCRVLCNGGSSSHSTGEVVVTMEPEVPIKKLETMV KLDATSEAVVKVDVAEEEKEKKKKKKGTAEEPEEE EPDESMLDWWSKYFASIDTMKEQLRQQEPSGIDLEE KEEVDNTEGLKGSMKGKEKARAAKEEKKKKTQSS GSGQGSEAPEKKKPKIDELKVYPKELESEFDNFEDW LHTFNLLRGKTGDDEDGSTEEERIVGRFKGSLCVYK VPLPEDVSREAGYDSTYGMFQGIPSNDPINVLVRVY VVRATDLHPADINGKADPYIAIRLGKTDIRDKENYIS KQLNPVFGKSFDIEASFPMESMLTVAVYDWDLVGT DDLIGETKIDLENRFYSKHRATCGIAQTYSTHGYNI WRDPMKPSQILTRLCKDGKVDGPHFGPPGRVKVAN RVFTGPSEIEDENGQRKPTDEHVALLALRHWEDIPR AGCRLVPEHVETRPLLNPDKPGIEQGRLELWVDMFP MDMPAPGTPLDISPRKPKKYELRVIIWNTDEVVLED DDFFTGEKSSDIFVRGWLKGQQEDKQDTDVHYHSL

[0157]

SEQ ID NO.	序列名称	序列
		TGEGNFNWRYLFPFDYLAAEEKIVISKKESMFSWDE TEYKIPARLTLQIWDADHFSADDFLGAIELDLNRFPR GAKTAKQCTMEMATGEVDVPLVSIFKQKRVKGWW PLLARNENDEFELTGKVEAELHLLTAEEAEKNPVGL ARNEPDPLEKPNRPDTAFVWFLNPLKSIKYLICTRY KWLIIKIVLALLGLLMLGLFLYSLPGYMVKKLLGA
6	mOTOF-201_1 蛋白 (NP_114081.2)，小鼠耳畴蛋白同功型 2，1997 aa	MALIVHLKTVSELRGKGDRIAKVTFRGQSFYSRVLE NCEGVADFDETFRWPVASSIDRNEVLEIQIFNYSKVF SNKLIGTFCMVLQKVVEENRVEVTDTLMDDSNAIIK TSLSMEVRYQATDGTVGPWDDGDFLGDESLQEEKD SQETDGLLPGSRPSTRISGEKSFRRAGRSVFSAMKLG KTRSHKEEPQRQDEPAVLEMEDLDHLAIQLGDGLD PDSVSLASVTALTSNVSNKRSKPDIKMEPSAGRPMD YQVSITVIEARQLVGLNMDPVVCVEVGDDKKYTSM KESTNCPYYNEYFVFDFHVSPDVMFDKIIKISVIHSK NLLRSGTLVGSFKMDVGTVYSQPEHQFHHKWAILS DPDDISAGLKGYVKCDVAVVGKGDNIKTPHKANET DEDDIEGNLLLPEGVPPERQWARFYVKIYRAEGLPR MNTSLMANVKKAFIGENKDLVDPYVQVFFAGQKG KTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQIR DSDKVNDVAIGTHFIDLRKISNDGDKGFLPTLGPAW VNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARLM LGLAVEILDTSNPELTSSTEVQVEQATPVSESCTGRM EEFFLFGAFLEASMIDRKNGDKPITFEVTIGNYGNEV DGMSRPLRPRPRKEPGDEEEVDLIQNSSDDEGDEAG DLASVSSTPPMRPQITDRNYFHLPYLERKPCIYIKSW WPDQRRRLYNANIMDHIADKLEEGLNDVQEMIKTE KSYPERRLRGVLEELSCGCHRFLSLSDKDQGRSSRT RLDRERLKSCMRELESMGQQAKSLRAQVKRHTVR DKLRSCQNFLQKLRFLADEPQHSIPDVFIWMMSNN KRIAYARVPSKDLLFSIVEEELGKDCAKVKTLFLKLP GKRGFGSAGWTVQAKLELYLWLGLSKQRKDFLCG LPCGFEEVKAAQGLGLHSFPPISLVYTKKQAFQLRA HMYQARSLFAADSSGLSDPFARVFFINQSQCTEVLN ETLCPTWDQMLVFDNLELYGEAHELRDDPPIIVIEIY DQDSMGKADFMGRTFAKPLVKMADEAYCPPRFPP QLEYYQIYRGSATAGDLLAAFELLQIGPSGKADLPPI NGPVDMDRGPIMPVPVGIRPVLSKYRVEVLFWGLR DLKRVNLAQVDRPRVDIECAGKGVQSSLIHNYKKN PNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAFGR YTLVGSHAVSSLRRFIYRPPDRSAPNWNTTVRLLRG CHRLRNGGPSSRPTGEVVVSMEPEEPVKKLETMVK LDATSDAVVKVDVAEDEKERKKKKKKGPSEEPEEE EPDESMLDWWSKYFASIDTMKEQLRQHETSGTDLE EKEEMESAEGLKGPMKSKEKSRAAKEEKKKKNQSP GPGQGSEAPEKKKAKIDELKVYPKELESEFDSFEDW LHTFNLLRGKTGDDEDGSTEEERIVGRFKGSLCVYK VPLPEDVSREAGYDPTYGMFQGIPSNDPINVLVRIY

[0158]

SEQ ID NO.	序列名称	序列
		VVRATDLHPADINGKADPYIAIKLGKTDIRDKENYIS KQLNPVFGKSFDIEASFPMESMLTVAVYDWDLVGT DDLIGETKIDLENRFYSKHRATCGIAQTYSIHGYNIW RDPMKPSQILTRLCKEGKVDGPHFGPHGRVRVANR VFTGPSEIEDENGQRKPTDEHVALSALRHWEDIPRV GCRLVPEHVETRPLLNPDKPGIEQGRLELWVDMFP MDMPAPGTPLDISPRKPKKYELRVIVWNTDEVVLE DDDFFTGEKSSDIFVRGWLKGQQEDKQDTDVHYHS LTGEGNFNWRYLFPFDYLAAEEKIVMSKKESMFSW DETEYKIPARLTLQIWDADHFSADDFLGAIELDLNR FPRGAKTAKQCTMEMATGEVDVPLVSIFKQKRVKG WWPLLARNENDEFELTGKVEAELHLLTAEEAEKNP VGLARNEPDPLEKPNRPDTSFIWFLNPLKSARYFLW HTYRWLLLKFLLLFLLLLLFALFLYSLPGYLAKKILG A
7	mOTOF-201_2 蛋白 (NP_001273350.1)，小鼠耳畸蛋白同功型 3，1977 aa	MALIVHLKTVSELRGKGDRIAKVTFRGQSFYSRVLE NCEGVADFDETFRWPVASSIDRNEVLEIQIFNYSKVF SNKLIGTFCMVLQKVVEENRVEVTDTLMDDSNAIIK TSLSMEVRYQATDGTVGPWDDGDFLGDESLQEEKD SQETDGLLPGSRPSTRISGEKSFRRAGRSVFSAMKLG KTRSHKEEPQRQDEPAVLEMEDLDHLAIQLGDGLD PDSVSLASVTALTSNVSNKRSKPDIKMEPSAGRPMD YQVSITVIEARQLVGLNMDPVVCVEVGDDKKYTSM KESTNCPYYNEYFVFDFHVSPDVMFDKIIKISVIHSK NLLRSGTLVGSFKMDVGTVYSQPEHQFHHKWAILS DPDDISAGLKGYVKCDVAVVGKGDNIKTPHKANET DEDDIEGNLLLPEGVPPERQWARFYVKIYRAEGLPR MNTSLMANVKKAFIGENKDLVDPYVQVFFAGQKG KTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQIR DSDKVNDVAIGTHFIDLRKISNDGDKGFLPTLGPAW VNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARLM LGLAVEILDTSNPELTSSTEVQVEQATPVSESCTGRM EEFFLFGAFLEASMIDRKNGDKPITFEVTIGNYGNEV DGMSRPLRPRPRKEPGDEEEVDLIQNSSDDEGDEAG DLASVSSTPPMRPQITDRNYFHLPYLERKPCIYIKSW WPDQRRRLYNANIMDHIADKLEEGLNDVQEMIKTE KSYPERRLRGVLEELSCGCHRFLSLSDKDQGRSSRT RLDRERLKSCMRELESMGQQAKSLRAQVKRHTVR DKLRSCQNFLQKLRFLADEPQHSIPDVFIWMMSNN KRIAYARVPSKDLLFSIVEEELGKDCAKVKTLFLKLP GKRGFGSAGWTVQAKLELYLWLGLSKQRKDFLCG LPCGFEEVKAAQGLGLHSFPPISLVYTKKQAFQLRA HMYQARSLFAADSSGLSDPFARVFFINQSQCTEVLN ETLCPTWDQMLVFDNLELYGEAHELRDDPPIIVIEIY DQDSMGKADFMGRTFAKPLVKMADEAYCPPRFPP QLEYYQIYRGSATAGDLLAAFELLQIGPSGKADLPPI NGPVDMDRGPIMPVPVGIRPVLSKYRVEVLFWGLR DLKRVNLAQVDRPRVDIECAGKGVQSSLIHNYKKN

[0159]

SEQ ID NO.	序列名称	序列
		PNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAFGR YTLVGSHAVSSLRRFIYRPPDRSAPNWNTTGEVVVS MEPEEPVKKLETMVKLDATSDAVVKVDVAEDEKE RKKKKKKGPSEEPEEEEPDESMLDWWSKYFASIDT MKEQLRQHETSGTDLEEKEEMESAEGLKGPMKSKE KSRAAKEEKKKKNQSPGPGQGSEAPEKKKAKIDEL KVYPKELESEFDSFEDWLHTFNLLRGKTGDDEDGST EEERIVGRFKGSLCVYKVPLPEDVSREAGYDPTYGM FQGIPSNDPINVLVRIYVVRATDLHPADINGKADPYI AIKLGKTDIRDKENYISKQLNPVFGKSFDIEASFPME SMLTVAVYDWDLVGTDDLIGETKIDLENRFYSKHR ATCGIAQTYSIHGYNIWRDPMKPSQILTRLCKEGKV DGPHFGPHGRVRVANRVFTGPSEIEDENGQRKPTDE HVALSALRHWEDIPRVGCRLVPEHVETRPLLNPDKP GIEQGRLELWVDMFPMDMPAPGTPLDISPRKPKKY ELRVIVWNTDEVVLEDDDFFTGEKSSDIFVRGWLK GQQEDKQDTDVHYHSLTGEGNFNWRYLFPFDYLA AEEKIVMSKKESMFSWDETEYKIPARLTLQIWDAD HFSADDFLGAIELDLNRFPRGAKTAKQCTMEMATG EVDVPLVSIFKQKRVKGWWPLLARNENDEFELTGK VEAELHLLTAEEAEKNPVGLARNEPDPLEKPNRPDT SFIWFLNPLKSARYFLWHTYRWLLLKFLLLFLLLLL FALFLYSLPGYLAKKILGA
8	mOTOF-202_1 蛋白 (NP_001093865.1)，小鼠耳畸蛋白同功型 1，1992 aa	MALIVHLKTVSELRGKGDRIAKVTFRGQSFYSRVLE NCEGVADFDETFRWPVASSIDRNEVLEIQIFNYSKVF SNKLIGTFCMVLQKVVEENRVEVTDTLMDDSNAIIK TSLSMEVRYQATDGTVGPWDDGDFLGDESLQEEKD SQETDGLLPGSRPSTRISGEKSFRSKGREKTKGGRDG EHKAGRSVFSAMKLGKTRSHKEEPQRQDEPAVLEM EDLDHLAIQLGDGLDPDSVSLASVTALTSNVSNKRS KPDIKMEPSAGRPMDYQVSITVIEARQLVGLNMDPV VCVEVGDDKKYTSMKESTNCPYYNEYFVFDFHVSP DVMFDKIIKISVIHSKNLLRSGTLVGSFKMDVGTVY SQPEHQFHHKWAILSDPDDISAGLKGYVKCDVAVV GKGDNIKTPHKANETDEDDIEGNLLLPEGVPPERQW ARFYVKIYRAEGLPRMNTSLMANVKKAFIGENKDL VDPYVQVFFAGQKGKTSVQKSSYEPLWNEQVVFTD LFPPLCKRMKVQIRDSDKVNDVAIGTHFIDLRKISND GDKGFLPTLGPAWVNMYGSTRNYTLLDEHQDLNE GLGEGVSFRARLMLGLAVEILDTSNPELTSSTEVQV EQATPVSESCTGRMEEFFLFGAFLEASMIDRKNGDK PITFEVTIGNYGNEVDGMSRPLRPRPRKEPGDEEEV DLIQNSSDDEGDEAGDLASVSSTPPMRPQITDRNYF HLPYLERKPCIYIKSWWPDQRRRLYNANIMDHIADK LEEGLNDVQEMIKTEKSYPERRLRGVLEELSCGCHR FLSLSDKDQGRSSRTRLDRERLKSCMRELESMGQQ AKSLRAQVKRHTVRDKLRSCQNFLQKLRFLADEPQ HSIPDVFIWMMSNNKRIAYARVPSKDLLFSIVEEELG

[0160]

SEQ ID NO.	序列名称	序列
		KDCAKVKTLFLKLPGKRGFGSAGWTVQAKLELYL WLGLSKQRKDFLCGLPCGFEEVKAAQGLGLHSFPPI SLVYTKKQAFQLRAHMYQARSLFAADSSGLSDPFA RVFFINQSQCTEVLNETLCPTWDQMLVFDNLELYGE AHELRDDPPIIVIEIYDQDSMGKADFMGRTFAKPLV KMADEAYCPPRFPPQLEYYQIYRGSATAGDLLAAFE LLQIGPSGKADLPPINGPVDMDRGPIMPVPVGIRPVL SKYRVEVLFWGLRDLKRVNLAQVDRPRVDIECAGK GVQSSLIHNYKKNPNFNTLVKWFEVDLPENELLHPP LNIRVVDCRAFGRYTLVGSHAVSSLRRFIYRPPDRS APNWNTTGEVVVSMEPEEPVKKLETMVKLDATSD AVVKVDVAEDEKERKKKKKKGPSEEPEEEEPDESM LDWWSKYFASIDTMKEQLRQHETSGTDLEEKEEME SAEGLKGPMKSKEKSRAAKEEKKKKNQSPGPGQGS EAPEKKKAKIDELKVYPKELESEFDSFEDWLHTFNL LRGKTGDDEDGSTEEERIVGRFKGSLCVYKVPLPED VSREAGYDPTYGMFQGIPSNDPINVLVRIYVVRATD LHPADINGKADPYIAIKLGKTDIRDKENYISKQLNPV FGKSFDIEASFPMESMLTVAVYDWDLVGTDDLIGET KIDLENRFYSKHRATCGIAQTYSIHGYNIWRDPMKP SQILTRLCKEGKVDGPHFGPHGRVRVANRVFTGPSE IEDENGQRKPTDEHVALSALRHWEDIPRVGCRLVPE HVETRPLLNPDKPGIEQGRLELWVDMFPMDMPAPG TPLDISPRKPKKYELRVIVWNTDEVVLEDDDFFTGE KSSDIFVRGWLKGQQEDKQDTDVHYHSLTGEGNFN WRYLFPFDYLAAEEKIVMSKKESMFSWDETEYKIP ARLTLQIWDADHFSADDFLGAIELDLNRFPRGAKTA KQCTMEMATGEVDVPLVSIFKQKRVKGWWPLLAR NENDEFELTGKVEAELHLLTAEEAEKNPVGLARNEP DPLEKPNRPDTAFVWFLNPLKSIKYLICTRYKWLIIK IVLALLGLLMLALFLYSLPGYMVKKLLGA
9	mOTOF-202_2 蛋白 (NP_001300696.1)，小鼠耳畸蛋白同功型 4，1977 aa	MALIVHLKTVSELRGKGDRIAKVTFRGQSFYSRVLE NCEGVADFDETFRWPVASSIDRNEVLEIQIFNYSKVF SNKLIGTFCMVLQKVVEENRVEVTDTLMDDSNAIIK TSLSMEVRYQATDGTVGPWDDGDFLGDESLQEEKD SQETDGLLPGSRPSTRISGEKSFRRAGRSVFSAMKLG KTRSHKEEPQRQDEPAVLEMEDLDHLAIQLGDGLD PDSVSLASVTALTSNVSNKRSKPDIKMEPSAGRPMD YQVSITVIEARQLVGLNMDPVVCVEVGDDKKYTSM KESTNCPYYNEYFVFDFHVSPDVMFDKIIKISVIHSK NLLRSGTLVGSFKMDVGTVYSQPEHQFHHKWAILS DPDDISAGLKGYVKCDVAVVGKGDNIKTPHKANET DEDDIEGNLLLPEGVPPERQWARFYVKIYRAEGLPR MNTSLMANVKKAFIGENKDLVDPYVQVFFAGQKG KTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQIR DSDKVNDVAIGTHFIDLRKISNDGDKGFLPTLGPAW VNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARLM LGLAVEILDTSNPELTSSTEVQVEQATPVSESCTGRM

[0161]

SEQ ID NO.	序列名称	序列
		EEFFLFGAFLEASMIDRKNGDKPITFEVTIGNYGNEV DGMSRPLRPRPRKEPGDEEEVDLIQNSSDDEGDEAG DLASVSSTPPMRPQITDRNYFHLPYLERKPCIYIKSW WPDQRRRLYNANIMDHIADKLEEGLNDVQEMIKTE KSYPERRLRGVLEELSCGCHRFLSLSDKDQGRSSRT RLDRERLKSCMRELESMGQQAKSLRAQVKRHTVR DKLRSCQNFLQKLRFLADEPQHSIPDVFIWMMSNN KRIAYARVPSKDLLFSIVEEELGKDCAKVKTLFLKLP GKRGFGSAGWTVQAKLELYLWLGLSKQRKDFLCG LPCGFEEVKAAQGLGLHSFPPISLVYTKKQAFQLRA HMYQARSLFAADSSGLSDPFARVFFINQSQCTEVLN ETLCPTWDQMLVFDNLELYGEAHELRDDPPIIVIEIY DQDSMGKADFMGRTFAKPLVKMADEAYCPPRFPP QLEYYQIYRGSATAGDLLAAFELLQIGPSGKADLPPI NGPVDMDRGPIMPVPVGIRPVLSKYRVEVLFWGLR DLKRVNLAQVDRPRVDIECAGKGVQSSLIHNYKKN PNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAFGR YTLVGSHAVSSLRRFIYRPPDRSAPNWNTTGEVVVS MEPEEPVKKLETMVKLDATSDAVVKVDVAEDEKE RKKKKKKGPSEEPEEEEPDESMLDWWSKYFASIDT MKEQLRQHETSGTDLEEKEEMESAEGLKGPMKSKE KSRAAKEEKKKKNQSPGPGQGSEAPEKKKAKIDEL KVYPKELESEFDSFEDWLHTFNLLRGKTGDDEDGST EEERIVGRFKGSLCVYKVPLPEDVSREAGYDPTYGM FQGIPSNDPINVLVRIYVVRATDLHPADINGKADPYI AIKLGKTDIRDKENYISKQLNPVFGKSFDIEASFPME SMLTVAVYDWDLVGTDDLIGETKIDLENRFYSKHR ATCGIAQTYSIHGYNIWRDPMKPSQILTRLCKEGKV DGPHFGPHGRVRVANRVFTGPSEIEDENGQRKPTDE HVALSALRHWEDIPRVGCRLVPEHVETRPLLNPDKP GIEQGRLELWVDMFPMDMPAPGTPLDISPRKPKKY ELRVIVWNTDEVVLEDDDFFTGEKSSDIFVRGWLK GQQEDKQDTDVHYHSLTGEGNFNWRYLFPFDYLA AEEKIVMSKKESMFSWDETEYKIPARLTLQIWDAD HFSADDFLGAIELDLNRFPRGAKTAKQCTMEMATG EVDVPLVSIFKQKRVKGWWPLLARNENDEFELTGK VEAELHLLTAEEAEKNPVGLARNEPDPLEKPNRPDT AFVWFLNPLKSIKYLICTRYKWLIIKIVLALLGLLML ALFLYSLPGYMVKKLLGA
10	编码人类耳畸蛋白同功型 1 蛋白 (SEQ ID NO: 1)的 DNA 序列，5979 bp，对应于以	ATGGCCCTGATTGTTCACCTCAAGACTGTCTCAGA GCTCCGAGGCAAAGGTGACCGGATTGCCAAAGTC ACTTTCCGAGGGCAGTCTTTCTACTCCCGGGTCCT GGAGAACTGCGAGGGTGTGGCTGACTTTGATGAG ACGTTCCGGTGGCCAGTGGCCAGCAGCATCGACC GGAATGAAGTGTTGGAGATTCAGATTTTCAACTA CAGCAAAGTCTTCAGCAACAAGCTGATAGGGACC TTCTGCATGGTGCTGCAGAAAGTGGTGGAGGAGA ATCGGGTAGAGGTGACCGACACGCTGATGGATGA

[0162]

SEQ ID NO.	序列名称	序列
	NM_001100395 记载的编码序列	CAGCAATGCTATCATCAAGACCAGCCTGAGCATG GAGGTCCGGTATCAGGCCACAGATGGCACTGTGG GCCCCTGGGATGATGGAGACTTCCTGGGAGATGA ATCCCTCCAGGAGGAGAAGGACAGCCAGGAGAC AGATGGGCTGCTACCTGGTTCCCGACCCAGCACC CGGATATCTGGCGAGAAGAGCTTTCGCAGCAAAG GCAGAGAGAAGACCAAGGGAGGCAGAGATGGCG AGCACAAAGCGGGAAGGAGTGTGTTCTCGGCCAT GAAACTCGGCAAAACTCGGTCCCACAAAGAGGA GCCCCAAAGACAAGATGAGCCAGCAGTGCTGGA GATGGAGGACCTGGACCACCTAGCCATTCAGCTG GGGGATGGGCTGGATCCTGACTCCGTGTCTCTAG CCTCGGTCACCGCTCTCACCAGCAATGTCTCCAAC AAACGGTCTAAGCCAGATATTAAGATGGAGCCCA GTGCTGGAAGGCCCATGGATTACCAGGTCAGCAT CACAGTGATTGAGGCTCGGCAGCTGGTGGGCTTG AACATGGACCCTGTGGTGTGTGTGGAGGTGGGTG ATGACAAGAAATACACGTCAATGAAGGAGTCCAC AAACTGCCCTTACTACAACGAGTACTTTGTCTTCG ACTTCCATGTCTCTCCTGATGTCATGTTTGACAAG ATCATCAAGATCTCGGTTATCCATTCTAAGAACCT GCTTCGGAGCGGCACCCTGGTGGGTTCCTTCAAA ATGGATGTGGGGACTGTGTATTCCCAGCCTGAAC ACCAGTTCCATCACAAATGGGCCATCCTGTCAGA CCCCGATGACATCTCTGCTGGGTTGAAGGGTTAT GTAAAGTGTGATGTCGCTGTGGTGGGCAAGGGAG ACAACATCAAGACACCCCACAAGGCCAACGAGA CGGATGAGGACGACATTGAAGGGAACTTGCTGCT CCCCGAGGGCGTGCCCCCCGAACGGCAGTGGGCA CGGTTCTATGTGAAAATTTACCGAGCAGAGGGAC TGCCCCGGATGAACACAAGCCTCATGGCCAACGT GAAGAAGGCGTTCATCGGTGAGAACAAGGACCTC GTCGACCCCTATGTGCAAGTCTTCTTTGCTGGACA AAAGGGCAAAACATCAGTGCAGAAGAGCAGCTA TGAGCCGCTATGGAATGAGCAGGTCGTCTTCACA GACTTGTTCCCCCCACTCTGCAAACGCATGAAGG TGCAGATCCGGGACTCTGACAAGGTCAATGATGT GGCCATCGGCACCCACTTCATCGACCTGCGCAAG ATTTCCAACGATGGAGACAAAGGCTTCCTGCCTA CCCTCGGTCCAGCCTGGGTGAACATGTACGGCTC CACGCGCAACTACACACTGCTGGACGAGCACCAG GACTTGAATGAAGGCCTGGGGGAGGGTGTGTCCT TCCGGGCCCGCCTCATGTTGGGACTAGCTGTGGA GATCCTGGACACCTCCAACCCAGAGCTCACCAGC TCCACGGAGGTGCAGGTGGAGCAGGCCACGCCTG TCTCGGAGAGCTGCACAGGGAGAATGGAAGAATT TTTTCTATTTGGAGCCTTCTTGGAAGCCTCAATGA TTGACCGGAAAAATGGGGACAAGCCAATTACCTT

[0163]

SEQ ID NO.	序列名称	序列
		TGAGGTGACCATAGGAAACTACGGCAATGAAGTC GATGGTATGTCCCGGCCCCTGAGGCCTCGGCCCC GGAAAGAGCCTGGGGATGAAGAAGAGGTAGACC TGATTCAGAACTCCAGTGACGATGAAGGTGACGA AGCCGGGGACCTGGCCTCGGTGTCCTCCACCCCA CCTATGCGGCCCCAGATCACGGACAGGAACTATT TCCACCTGCCCTACCTGGAGCGCAAGCCCTGCAT CTATATCAAGAGCTGGTGGCCTGACCAGAGGCGG CGCCTCTACAATGCCAACATCATGGATCACATTG CTGACAAGCTGGAAGAAGGCCTGAATGATGTACA GGAGATGATCAAAACGGAGAAGTCCTACCCGGA GCGCCGCCTGCGGGGTGTGCTAGAGGAACTCAGC TGTGGCTGCCACCGCTTCCTCTCCCTCTCGGACAA GGACCAGGGCCGCTCGTCCCGCACCAGGCTGGAT CGAGAGCGTCTTAAGTCCTGTATGAGGGAGTTGG AGAGCATGGGACAGCAGGCCAAGAGCCTGAGGG CTCAGGTGAAGCGGCACACTGTTCGGGACAAGCT GAGGTCATGCCAGAACTTTCTGCAGAAGCTACGC TTCCTGGCGGATGAGCCCCAGCACAGCATTCCTG ATGTGTTCATTTGGATGATGAGCAACAACAAACG TATCGCCTATGCCCGCGTGCCTTCCAAAGACCTGC TCTTCTCCATCGTGGAGGAGGAACTGGGCAAGGA CTGCGCCAAAGTCAAGACCCTCTTCCTGAAGCTG CCAGGGAAGAGGGGCTTCGGCTCGGCAGGCTGGA CAGTACAGGCCAAGCTGGAGCTCTACCTGTGGCT GGGCCTCAGCAAGCAGCGAAAGGACTTCCTGTGT GGTCTGCCCTGTGGCTTCGAGGAGGTCAAGGCAG CCCAAGGCCTGGGCCTGCATTCCTTTCCGCCCATC AGCCTAGTCTACACCAAGAAGCAAGCCTTCCAGC TCCGAGCACACATGTATCAGGCCCGAAGCCTCTT TGCTGCTGACAGCAGTGGGCTCTCTGATCCCTTTG CCCGTGTCTTCTTCATCAACCAGAGCCAATGCACT GAGGTTCTAAACGAGACACTGTGTCCCACCTGGG ACCAGATGCTGGTATTTGACAACCTGGAGCTGTA CGGTGAAGCTCACGAGTTACGAGATGATCCCCCC ATCATTGTCATTGAAATCTACGACCAGGACAGCA TGGGCAAAGCCGACTTCATGGGCCGGACCTTCGC CAAGCCCCTGGTGAAGATGGCAGATGAAGCATAC TGCCCACCTCGCTTCCCGCCGCAGCTTGAGTACTA CCAGATCTACCGAGGCAGTGCCACTGCCGGAGAC CTACTGGCTGCCTTCGAGCTGCTGCAGATTGGGCC ATCAGGGAAGGCTGACCTGCCACCCATCAATGGC CCAGTGGACATGGACAGAGGGCCCATCATGCCTG TGCCCGTGGGAATCCGGCCAGTGCTCAGCAAGTA CCGAGTGGAGGTGCTGTTCTGGGGCCTGAGGGAC CTAAAGAGGGTGAACCTGGCCCAGGTGGACCGAC CACGGGTGGACATCGAGTGTGCAGGAAAGGGGG TACAATCCTCCCTGATTCACAATTATAAGAAGAA

[0164]

SEQ ID NO.	序列名称	序列
		CCCCAACTTCAACACGCTGGTCAAGTGGTTTGAA GTGGACCTCCCGGAGAATGAGCTCCTGCACCCAC CCTTGAACATCCGAGTGGTAGATTGCCGGGCCTTT GGACGATACACCCTGGTGGGTTCCCACGCAGTCA GCTCACTGAGGCGCTTCATCTACCGACCTCCAGA CCGCTCAGCCCCCAACTGGAACACCACAGGGGAG GTTGTAGTAAGCATGGAGCCTGAGGAGCCAGTTA AGAAGCTGGAGACCATGGTGAAACTGGATGCGAC TTCTGATGCTGTGGTCAAGGTGGATGTGGCTGAA GATGAGAAGGAAAGGAAGAAGAAGAAAAAGAA AGGCCCGTCAGAGGAGCCAGAGGAGGAAGAGCC CGATGAGAGCATGCTGGATTGGTGGTCCAAGTAC TTCGCCTCCATCGACACAATGAAGGAGCAACTTC GACAACATGAGACCTCTGGAACTGACTTGGAAGA GAAGGAAGAGATGGAAAGCGCTGAGGGCCTGAA GGGACCAATGAAGAGCAAGGAGAAGTCCAGAGC TGCAAAGGAGGAGAAAAAGAAGAAAAACCAGAG CCCTGGCCCTGGCCAGGGATCGGAGGCTCCTGAG AAGAAGAAAGCCAAGATCGATGAGCTTAAGGTGT ACCCCAAGGAGCTGGAATCGGAGTTTGACAGCTT TGAGGACTGGCTGCACACCTTCAACCTGTTGAGG GGCAAGACGGGAGATGATGAGGATGGCTCCACA GAGGAGGAGCGCATAGTAGGCCGATTCAAGGGCT CCCTCTGTGTGTACAAAGTGCCACTCCCAGAAGA TGTATCTCGAGAAGCTGGCTATGATCCCACCTATG GAATGTTCCAGGGCATCCCAAGCAATGACCCCAT CAATGTGCTGGTCCGAATCTATGTGGTCCGGGCC ACAGACCTGCACCCGGCCGACATCAATGGCAAAG CTGACCCCTATATTGCCATCAAGTTAGGCAAGAC CGACATCCGAGACAAGGAGAACTACATCTCCAAG CAGCTCAACCCTGTGTTTGGGAAGTCCTTTGACAT TGAGGCCTCCTTCCCCATGGAGTCCATGTTGACAG TGGCCGTGTACGACTGGGATCTGGTGGGCACTGA TGACCTCATCGGAGAAACCAAGATTGACCTGGAA AACCGCTTCTACAGCAAGCATCGCGCCACCTGCG GCATCGCACAGACCTATTCCATACATGGCTACAA TATCTGGAGGGACCCCATGAAGCCCAGCCAGATC CTGACACGCCTCTGTAAAGAGGGCAAAGTGGACG GCCCCCACTTTGGTCCCCATGGGAGAGTGAGGGT TGCCAACCGTGTCTTCACGGGGCCTTCAGAAATA GAGGATGAGAATGGTCAGAGGAAGCCCACAGAT GAGCACGTGGCACTGTCTGCTCTGAGACACTGGG AGGACATCCCCCGGGTGGGCTGCCGCCTTGTGCC GGAACACGTGGAGACCAGGCCGCTGCTCAACCCT GACAAGCCAGGCATTGAGCAGGGCCGCCTGGAGC TGTGGGTGGACATGTTCCCCATGGACATGCCAGC CCCTGGGACACCTCTGGATATATCCCCCAGGAAA CCCAAGAAGTACGAGCTGCGGGTCATCGTGTGGA

[0165]

SEQ ID NO.	序列名称	序列
		ACACAGACGAGGTGGTCCTGGAAGACGATGATTT CTTCACGGGAGAGAAGTCCAGTGACATTTTTGTG AGGGGGTGGCTGAAGGGCCAGCAGGAGGACAAA CAGGACACAGATGTCCACTATCACTCCCTCACGG GGGAGGGCAACTTCAACTGGAGATACCTCTTCCC CTTCGACTACCTAGCGGCCGAAGAGAAGATCGTT ATGTCCAAAAAGGAGTCTATGTTCTCCTGGGATG AGACGGAGTACAAGATCCCTGCGCGGCTCACCCT GCAGATCTGGGACGCTGACCACTTCTCGGCTGAC GACTTCCTGGGGGCTATCGAGCTGGACCTGAACC GGTTCCCGAGGGGCGCTAAGACAGCCAAGCAGTG CACCATGGAGATGGCCACCGGGGAGGTGGACGTA CCCCTGGTTTCCATCTTTAAACAGAAACGTGTCAA AGGCTGGTGGCCCCTCCTGGCCCGCAATGAGAAT GATGAGTTTGAGCTCACAGGCAAAGTGGAGGCGG AGCTACACCTACTCACGGCAGAGGAGGCAGAGA AGAACCCTGTGGGCCTGGCTCGCAATGAACCTGA TCCCCTAGAAAAACCCAACCGGCCTGACACGGCA TTCGTCTGGTTCCTGAACCCACTCAAATCTATCAA GTACCTCATCTGCACCCGGTACAAGTGGCTGATC ATCAAGATCGTGCTGGCGCTGCTGGGGCTGCTCA TGCTGGCCCTCTTCCTTTACAGCCTCCCAGGCTAC ATGGTCAAGAAGCTCCTAGGGGCCTGA
11	OTOF-202 转录物 (NM_004802.3)，人类耳畴蛋白转录物变体 2，4969 bp，编码 SEQ ID NO: 2 的蛋白质	CCGTGAGTTCTGCCCAGGCCCTGTGAGCTCACCA GAGCCACAGACTCACAGCCCAGAGGTGGCTTCTT CCTTCAGGAACTGAAGAACCCCCATGAACACCAA CATCTCCAGGTTCTGAGAACAGAACCTGGGAAAT TGATGACTTCCTCATGATGACCGATACTCAGGAT GGCCCTAGCGAGAGCTCCCAGATCATGAGGAAGA AGGCCTGAACGACATACAGGAGATGATCAAAAC GGAGAAGTCCTACCCTGAGCGTCGCCTGCGGGGC GTCCTGGAGGAGCTGAGCTGTGGCTGCTGCCGCT TCCTCTCCCTCGCTGACAAGGACCAGGGCCACTC ATCCCGCACCAGGCTTGACCGGGAGCGCCTCAAG TCCTGCATGAGGGAGCTGGAAAACATGGGGCAGC AGGCCAGGATGCTGCGGGCCCAGGTGAAGCGGC ACACGGTGCGGGACAAGCTGAGGCTGTGCCAGAA CTTCCTGCAGAAGCTGCGCTTCCTGGCGGACGAG CCCCAGCACAGCATTCCCGACATCTTCATCTGGAT GATGAGCAACAACAAGCGTGTCGCCTATGCCCGT GTGCCCTCCAAGGACCTGCTCTTCTCCATCGTGGA GGAGGAGACTGGCAAGGACTGCGCCAAGGTCAA GACGCTCTTCCTTAAGCTGCCAGGGAAGCGGGGC TTCGGCTCGGCAGGCTGGACAGTGCAGGCCAAGG TGGAGCTGTACCTGTGGCTGGGCCTCAGCAAACA GCGCAAGGAGTTCCTGTGCGGCCTGCCCTGTGGC TTCCAGGAGGTCAAGGCAGCCCAGGGCCTGGGCC TGCATGCCTTCCCACCCGTCAGCCTGGTCTACACC

[0166]

SEQ ID NO.	序列名称	序列
		AAGAAGCAGGCGTTCCAGCTCCGAGCGCACATGT ACCAGGCCCGCAGCCTCTTTGCCGCCGACAGCAG CGGACTCTCAGACCCCTTTGCCCGCGTCTTCTTCA TCAATCAGAGTCAGTGCACAGAGGTGCTGAATGA GACCCTGTGTCCCACCTGGGACCAGATGCTGGTG TTCGACAACCTGGAGCTCTATGGTGAAGCTCATG AGCTGAGGGACGATCCGCCCATCATTGTCATTGA AATCTATGACCAGGATTCCATGGGCAAAGCTGAC TTCATGGGCCGGACCTTCGCCAAACCCCTGGTGA AGATGGCAGACGAGGCGTACTGCCCACCCCGCTT CCCACCTCAGCTCGAGTACTACCAGATCTACCGT GGCAACGCCACAGCTGGAGACCTGCTGGCGGCCT TCGAGCTGCTGCAGATTGGACCAGCAGGGAAGGC TGACCTGCCCCCCATCAATGGCCCGGTGGACGTG GACCGAGGTCCCATCATGCCCGTGCCCATGGGCA TCCGGCCCGTGCTCAGCAAGTACCGAGTGGAGGT GCTGTTCTGGGGCCTACGGGACCTAAAGCGGGTG AACCTGGCCCAGGTGGACCGGCCACGGGTGGACA TCGAGTGTGCAGGGAAGGGGGTGCAGTCGTCCCT GATCCACAATTATAAGAAGAACCCCAACTTCAAC ACCCTCGTCAAGTGGTTTGAAGTGGACCTCCCAG AGAACGAGCTGCTGCACCCGCCCTTGAACATCCG TGTGGTGGACTGCCGGGCCTTCGGTCGCTACACA CTGGTGGGCTCCCATGCCGTCAGCTCCCTGCGAC GCTTCATCTACCGGCCCCCAGACCGCTCGGCCCCC AGCTGGAACACCACGGGGGAGGTTGTGGTGACTA TGGAGCCAGAGGTACCCATCAAGAAACTGGAGAC CATGGTGAAGCTGGACGCGACTTCTGAAGCTGTT GTCAAGGTGGATGTGGCTGAGGAGGAGAAGGAG AAGAAGAAGAAGAAGAAGGGCACTGCGGAGGAG CCAGAGGAGGAGGAGCCAGACGAGAGCATGCTG GACTGGTGGTCCAAGTACTTTGCCTCCATTGACAC CATGAAGGAGCAACTTCGACAACAAGAGCCCTCT GGAATTGACTTGGAGGAGAAGGAGGAAGTGGAC AATACCGAGGGCCTGAAGGGGTCAATGAAGGGC AAGGAGAAGGCAAGGGCTGCCAAAGAGGAGAAG AAGAAGAAAACTCAGAGCTCTGGCTCTGGCCAGG GGTCCGAGGCCCCCGAGAAGAAGAAACCCAAGA TTGATGAGCTTAAGGTATACCCCAAAGAGCTGGA GTCCGAGTTTGATAACTTTGAGGACTGGCTGCAC ACTTTCAACTTGCTTCGGGGCAAGACCGGGGATG ATGAGGATGGCTCCACCGAGGAGGAGCGCATTGT GGGACGCTTCAAGGGCTCCCTCTGCGTGTACAAA GTGCCACTCCCAGAGGACGTGTCCCGGGAAGCCG GCTACGACTCCACCTACGGCATGTTCCAGGGCAT CCCGAGCAATGACCCCATCAATGTGCTGGTCCGA GTCTATGTGGTCCGGGCCACGGACCTGCACCCTG CTGACATCAACGGCAAAGCTGACCCCTACATCGC

[0167]

SEQ ID NO.	序列名称	序列
		CATCCGGCTAGGCAAGACTGACATCCGCGACAAG GAGAACTACATCTCCAAGCAGCTCAACCCTGTCT TTGGGAAGTCCTTTGACATCGAGGCCTCCTTCCCC ATGGAATCCATGCTGACGGTGGCTGTGTATGACT GGGACCTGGTGGGCACTGATGACCTCATTGGGGA AACCAAGATCGACCTGGAGAACCGCTTCTACAGC AAGCACCGCGCCACCTGCGGCATCGCCCAGACCT ACTCCACACATGGCTACAATATCTGGCGGGACCC CATGAAGCCCAGCCAGATCCTGACCCGCCTCTGC AAAGACGGCAAAGTGGACGGCCCCCACTTTGGGC CCCCTGGGAGAGTGAAGGTGGCCAACCGCGTCTT CACTGGGCCCTCTGAGATTGAGGACGAGAACGGT CAGAGGAAGCCCACAGACGAGCATGTGGCGCTGT TGGCCCTGAGGCACTGGGAGGACATCCCCCGCGC AGGCTGCCGCCTGGTGCCAGAGCATGTGGAGACG AGGCCGCTGCTCAACCCCGACAAGCCGGGCATCG AGCAGGGCCGCCTGGAGCTGTGGGTGGACATGTT CCCCATGGACATGCCAGCCCCTGGGACGCCTCTG GACATCTCACCTCGGAAGCCCAAGAAGTACGAGC TGCGGGTCATCATCTGGAACACAGATGAGGTGGT CTTGGAGGACGACGACTTCTTCACAGGGGAGAAG TCCAGTGACATCTTCGTGAGGGGGTGGCTGAAGG GCCAGCAGGAGGACAAGCAGGACACAGACGTCC ACTACCACTCCCTCACTGGCGAGGGCAACTTCAA CTGGCGCTACCTGTTCCCCTTCGACTACCTGGCGG CGGAGGAGAAGATCGTCATCTCCAAGAAGGAGTC CATGTTCTCCTGGGACGAGACCGAGTACAAGATC CCCGCGCGGCTCACCCTGCAGATCTGGGATGCGG ACCACTTCTCCGCTGACGACTTCCTGGGGGCCATC GAGCTGGACCTGAACCGGTTCCCGCGGGGCGCAA AGACAGCCAAGCAGTGCACCATGGAGATGGCCAC CGGGGAGGTGGACGTGCCCCTCGTGTCCATCTTC AAGCAAAAGCGCGTCAAAGGCTGGTGGCCCCTCC TGGCCCGCAATGAGAACGATGAGTTTGAGCTCAC GGGCAAGGTGGAGGCTGAGCTGCATTTACTGACA GCAGAGGAGGCAGAGAAGAACCCAGTGGGCCTG GCCCGCAATGAACCTGACCCCCTAGAGAAACCCA ACCGGCCCGACACGAGCTTCATCTGGTTCCTGAA CCCTCTCAAGTCGGCTCGCTACTTCTTGTGGCACA CGTATCGCTGGCTGCTCCTCAAACTGTTGCTGCTC CTGCTGCTGCTCCTCCTCCTCGCCCTGTTCCTCTAC TCTGTGCCTGGCTACCTGGTCAAGAAAATCCTCG GGGCCTGAGCCCAGTGGCCTCCTGGCCGGCCCGA CACGGCCTTCGTCTGGTTCCTCAACCCTCTCAAGT CCATCAAGTACCTCATCTGCACCCGGTACAAGTG GCTCATCATCAAGATCGTGCTGGCGCTGTTGGGG CTGCTCATGTTGGGGCTCTTCCTCTACAGCCTCCC TGGCTACATGGTCAAAAAGCTCCTTGGGGCATGA

[0168]

SEQ ID NO.	序列名称	序列
		AGGCCGCCAGCTCCCGCCAGCCGCTCCCCAGCCC TGCCGCATTTCCTTTCAGTGGCTTGGACTCTTTCC CATCTCCCCTGGGGAGCCTGAGGAGCCCAGCGTC CACTCTTCATGCCTTGGGCCGAGCCTGCCTCCTGC TTGCGGGGGCCGCCTGTCCTCACTGCCCCAGGCT GCGGCTTGCCCAGTCCCGCCCCTCTGACCCCTGCC TGTGGGCTGGGGAGCCTTGGATGGGGTGGGGACC TGGAATGGGTCTCTCTTGCCCCACCTGGCTGAGGC GCCACCCTTCTTCAGGCCCAGGCTCCAGAGGAAG ACTCCTGAAACCCTCCCCAGGTCTTCCAAGTACA GGATTGAAGCTTTAGTGAAATTAACCAAGGACCA TGGGTCAGTGCCCAGGGCTTTAAAAAGAATGAAC GAGCAAAAGGTATCCCCGCCGTGACCCCTGCAGA TAGCACCGGTCTTTGATCCGCAGCAGGGGCCAGA CCCTGCCCACAAGTCCCAGCGCGGCTGCTTCTGCC ACTGCTGGGCTCCACTTGGCTCCTCTCACTTCCCA GGGGGTCGCCTGTCCTGCCTGTGGGTTTCCATGGC TTCCCAGAGCTCCCTCTGCCCCAGCCAGCGCCTCC AGGCCCAGCTGAGGAGCTGTGAGAAGCAGCAGA GGGGACTCCCCATCCCGGGCACACCCTGTCCTCC CACCCCTGCCCCCTTGCCCTTCCAGCCCTTTCAGC TGCAGCTGGGAGCTGGCCCGTCAAGTGCTGCCCC TGCCTGTGTCTGGGTTTCTGTTGGCTGTTTTTCTTT TCTTGAGTGGTGATTTTTCTCTAAATAAAAGAAGT CAAGCACTGAAAAAAAAAAAAAAAA
12	OTOF-203 转录物 (NM_194323.2)，人类耳畸蛋白转录物变体 4，4771 bp，编码 SEQ ID NO: 3 的蛋白质	CCGTGAGTTCTGCCCAGGCCCTGTGAGCTCACCA GAGCCACAGACTCACAGCCCAGAGGTGGCTTCTT CCTTCAGGAACTGAAGAACCCCCATGAACACCAA CATCTCCAGGTTCTGAGAACAGAACCTGGGAAAT TGATGACTTCCTCATGATGACCGATACTCAGGAT GGCCCTAGCGAGAGCTCCCAGATCATGAGGAAGA AGGCCTGAACGACATACAGGAGATGATCAAAAC GGAGAAGTCCTACCCTGAGCGTCGCCTGCGGGGC GTCCTGGAGGAGCTGAGCTGTGGCTGCTGCCGCT TCCTCTCCCTCGCTGACAAGGACCAGGGCCACTC ATCCCGCACCAGGCTTGACCGGGAGCGCCTCAAG TCCTGCATGAGGGAGCTGGAAAACATGGGGCAGC AGGCCAGGATGCTGCGGGCCCAGGTGAAGCGGC ACACGGTGCGGGACAAGCTGAGGCTGTGCCAGAA CTTCCTGCAGAAGCTGCGCTTCCTGGCGGACGAG CCCCAGCACAGCATTCCCGACATCTTCATCTGGAT GATGAGCAACAACAAGCGTGTCGCCTATGCCCGT GTGCCCTCCAAGGACCTGCTCTTCTCCATCGTGGA GGAGGAGACTGGCAAGGACTGCGCCAAGGTCAA GACGCTCTTCCTTAAGCTGCCAGGGAAGCGGGGC TTCGGCTCGGCAGGCTGGACAGTGCAGGCCAAGG TGGAGCTGTACCTGTGGCTGGGCCTCAGCAAACA GCGCAAGGAGTTCCTGTGCGGCCTGCCCTGTGGC

[0169]

SEQ ID NO.	序列名称	序列
		TTCCAGGAGGTCAAGGCAGCCCAGGGCCTGGGCC TGCATGCCTTCCCACCCGTCAGCCTGGTCTACACC AAGAAGCAGGCGTTCCAGCTCCGAGCGCACATGT ACCAGGCCCGCAGCCTCTTTGCCGCCGACAGCAG CGGACTCTCAGACCCCTTTGCCCGCGTCTTCTTCA TCAATCAGAGTCAGTGCACAGAGGTGCTGAATGA GACCCTGTGTCCCACCTGGGACCAGATGCTGGTG TTCGACAACCTGGAGCTCTATGGTGAAGCTCATG AGCTGAGGGACGATCCGCCCATCATTGTCATTGA AATCTATGACCAGGATTCCATGGGCAAAGCTGAC TTCATGGGCCGGACCTTCGCCAAACCCCTGGTGA AGATGGCAGACGAGGCGTACTGCCCACCCCGCTT CCCACCTCAGCTCGAGTACTACCAGATCTACCGT GGCAACGCCACAGCTGGAGACCTGCTGGCGGCCT TCGAGCTGCTGCAGATTGGACCAGCAGGGAAGGC TGACCTGCCCCCCATCAATGGCCCGGTGGACGTG GACCGAGGTCCCATCATGCCCGTGCCCATGGGCA TCCGGCCCGTGCTCAGCAAGTACCGAGTGGAGGT GCTGTTCTGGGGCCTACGGGACCTAAAGCGGGTG AACCTGGCCCAGGTGGACCGGCCACGGGTGGACA TCGAGTGTGCAGGGAAGGGGGTGCAGTCGTCCCT GATCCACAATTATAAGAAGAACCCCAACTTCAAC ACCCTCGTCAAGTGGTTTGAAGTGGACCTCCCAG AGAACGAGCTGCTGCACCCGCCCTTGAACATCCG TGTGGTGGACTGCCGGGCCTTCGGTCGCTACACA CTGGTGGGCTCCCATGCCGTCAGCTCCCTGCGAC GCTTCATCTACCGGCCCCCAGACCGCTCGGCCCCC AGCTGGAACACCACGGGGGAGGTTGTGGTGACTA TGGAGCCAGAGGTACCCATCAAGAAACTGGAGAC CATGGTGAAGCTGGACGCGACTTCTGAAGCTGTT GTCAAGGTGGATGTGGCTGAGGAGGAGAAGGAG AAGAAGAAGAAGAAGAAGGGCACTGCGGAGGAG CCAGAGGAGGAGGAGCCAGACGAGAGCATGCTG GACTGGTGGTCCAAGTACTTTGCCTCCATTGACAC CATGAAGGAGCAACTTCGACAACAAGAGCCCTCT GGAATTGACTTGGAGGAGAAGGAGGAAGTGGAC AATACCGAGGGCCTGAAGGGGTCAATGAAGGGC AAGGAGAAGGCAAGGGCTGCCAAAGAGGAGAAG AAGAAGAAAACTCAGAGCTCTGGCTCTGGCCAGG GGTCCGAGGCCCCCGAGAAGAAGAAACCCAAGA TTGATGAGCTTAAGGTATACCCCAAAGAGCTGGA GTCCGAGTTTGATAACTTTGAGGACTGGCTGCAC ACTTTCAACTTGCTTCGGGGCAAGACCGGGGATG ATGAGGATGGCTCCACCGAGGAGGAGCGCATTGT GGGACGCTTCAAGGGCTCCCTCTGCGTGTACAAA GTGCCACTCCCAGAGGACGTGTCCCGGGAAGCCG GCTACGACTCCACCTACGGCATGTTCCAGGGCAT CCCGAGCAATGACCCCATCAATGTGCTGGTCCGA

[0170]

SEQ ID NO.	序列名称	序列
		GTCTATGTGGTCCGGGCCACGGACCTGCACCCTG CTGACATCAACGGCAAAGCTGACCCCTACATCGC CATCCGGCTAGGCAAGACTGACATCCGCGACAAG GAGAACTACATCTCCAAGCAGCTCAACCCTGTCT TTGGGAAGTCCTTTGACATCGAGGCCTCCTTCCCC ATGGAATCCATGCTGACGGTGGCTGTGTATGACT GGGACCTGGTGGGCACTGATGACCTCATTGGGGA AACCAAGATCGACCTGGAGAACCGCTTCTACAGC AAGCACCGCGCCACCTGCGGCATCGCCCAGACCT ACTCCACACATGGCTACAATATCTGGCGGGACCC CATGAAGCCCAGCCAGATCCTGACCCGCCTCTGC AAAGACGGCAAAGTGGACGGCCCCCACTTTGGGC CCCCTGGGAGAGTGAAGGTGGCCAACCGCGTCTT CACTGGGCCCTCTGAGATTGAGGACGAGAACGGT CAGAGGAAGCCCACAGACGAGCATGTGGCGCTGT TGGCCCTGAGGCACTGGGAGGACATCCCCCGCGC AGGCTGCCGCCTGGTGCCAGAGCATGTGGAGACG AGGCCGCTGCTCAACCCCGACAAGCCGGGCATCG AGCAGGGCCGCCTGGAGCTGTGGGTGGACATGTT CCCCATGGACATGCCAGCCCCTGGGACGCCTCTG GACATCTCACCTCGGAAGCCCAAGAAGTACGAGC TGCGGGTCATCATCTGGAACACAGATGAGGTGGT CTTGGAGGACGACGACTTCTTCACAGGGGAGAAG TCCAGTGACATCTTCGTGAGGGGGTGGCTGAAGG GCCAGCAGGAGGACAAGCAGGACACAGACGTCC ACTACCACTCCCTCACTGGCGAGGGCAACTTCAA CTGGCGCTACCTGTTCCCCTTCGACTACCTGGCGG CGGAGGAGAAGATCGTCATCTCCAAGAAGGAGTC CATGTTCTCCTGGGACGAGACCGAGTACAAGATC CCCGCGCGGCTCACCCTGCAGATCTGGGATGCGG ACCACTTCTCCGCTGACGACTTCCTGGGGGCCATC GAGCTGGACCTGAACCGGTTCCCGCGGGGCGCAA AGACAGCCAAGCAGTGCACCATGGAGATGGCCAC CGGGGAGGTGGACGTGCCCCTCGTGTCCATCTTC AAGCAAAAGCGCGTCAAAGGCTGGTGGCCCCTCC TGGCCCGCAATGAGAACGATGAGTTTGAGCTCAC GGGCAAGGTGGAGGCTGAGCTGCATTTACTGACA GCAGAGGAGGCAGAGAAGAACCCAGTGGGCCTG GCCCGCAATGAACCTGACCCCCTAGAGAAACCCA ACCGGCCCGACACGGCCTTCGTCTGGTTCCTCAAC CCTCTCAAGTCCATCAAGTACCTCATCTGCACCCG GTACAAGTGGCTCATCATCAAGATCGTGCTGGCG CTGTTGGGGCTGCTCATGTTGGGGCTCTTCCTCTA CAGCCTCCCTGGCTACATGGTCAAAAAGCTCCTT GGGGCATGAAGGCCGCCAGCTCCCGCCAGCCGCT CCCCAGCCCTGCCGCATTTCCTTTCAGTGGCTTGG ACTCTTTCCCATCTCCCCTGGGGAGCCTGAGGAGC CCAGCGTCCACTCTTCATGCCTTGGGCCGAGCCTG

[0171]

SEQ ID NO.	序列名称	序列
		CCTCCTGCTTGCGGGGGCCGCCTGTCCTCACTGCC CCAGGCTGCGGCTTGCCCAGTCCCGCCCCTCTGAC CCCTGCCTGTGGGCTGGGGAGCCTTGGATGGGGT GGGGACCTGGAATGGGTCTCTCTTGCCCCACCTG GCTGAGGCGCCACCCTTCTTCAGGCCCAGGCTCC AGAGGAAGACTCCTGAAACCCTCCCCAGGTCTTC CAAGTACAGGATTGAAGCTTTAGTGAAATTAACC AAGGACCATGGGTCAGTGCCCAGGGCTTTAAAAA GAATGAACGAGCAAAAGGTATCCCCGCCGTGACC CCTGCAGATAGCACCGGTCTTTGATCCGCAGCAG GGGCCAGACCCTGCCCACAAGTCCCAGCGCGGCT GCTTCTGCCACTGCTGGGCTCCACTTGGCTCCTCT CACTTCCCAGGGGGTCGCCTGTCCTGCCTGTGGGT TTCCATGGCTTCCCAGAGCTCCCTCTGCCCCAGCC AGCGCCTCCAGGCCCAGCTGAGGAGCTGTGAGAA GCAGCAGAGGGGACTCCCCATCCCGGGCACACCC TGTCCTCCCACCCCTGCCCCCTTGCCCTTCCAGCC CTTTCAGCTGCAGCTGGGAGCTGGCCCGTCAAGT GCTGCCCCTGCCTGTGTCTGGGTTTCTGTTGGCTG TTTTTCTTTTCTTGAGTGGTGATTTTTCTCTAAATA AAAGAAGTCAAGCACTGAAAAAAAAAAAAAAAAA
13	OTOF-208 转录物 (NM_194322.2)，人类耳畴蛋白转录物变体 3，5123 bp，编码 SEQ ID NO: 4 的蛋白质	CCGTGAGTTCTGCCCAGGCCCTGTGAGCTCACCA GAGCCACAGACTCACAGCCCAGAGGTGGCTTCTT CCTTCAGGAACTGAAGAACCCCCATGAACACCAA CATCTCCAGGTTCTGAGAACAGAACCTGGGAAAT TGATGACTTCCTCATGATGACCGATACTCAGGAT GGCCCTAGCGAGAGCTCCCAGATCATGAGGTCCC TCACTCCCCTGATCAACAGGGAGGAGGCATTTGG GGAGGCTGGGGAGGCGGGGCTGTGGCCCAGCATC ACCCACACTCCTGATTCACAGGAAGAAGGCCTGA ACGACATACAGGAGATGATCAAAACGGAGAAGT CCTACCCTGAGCGTCGCCTGCGGGGCGTCCTGGA GGAGCTGAGCTGTGGCTGCTGCCGCTTCCTCTCCC TCGCTGACAAGGACCAGGGCCACTCATCCCGCAC CAGGCTTGACCGGGAGCGCCTCAAGTCCTGCATG AGGGAGCTGGAAAACATGGGGCAGCAGGCCAGG ATGCTGCGGGCCCAGGTGAAGCGGCACACGGTGC GGGACAAGCTGAGGCTGTGCCAGAACTTCCTGCA GAAGCTGCGCTTCCTGGCGGACGAGCCCCAGCAC AGCATTCCCGACATCTTCATCTGGATGATGAGCA ACAACAAGCGTGTCGCCTATGCCCGTGTGCCCTC CAAGGACCTGCTCTTCTCCATCGTGGAGGAGGAG ACTGGCAAGGACTGCGCCAAGGTCAAGACGCTCT TCCTTAAGCTGCCAGGGAAGCGGGGCTTCGGCTC GGCAGGCTGGACAGTGCAGGCCAAGGTGGAGCT GTACCTGTGGCTGGGCCTCAGCAAACAGCGCAAG GAGTTCCTGTGCGGCCTGCCCTGTGGCTTCCAGGA GGTCAAGGCAGCCCAGGGCCTGGGCCTGCATGCC

[0172]

SEQ ID NO.	序列名称	序列
		TTCCCACCCGTCAGCCTGGTCTACACCAAGAAGC AGGCGTTCCAGCTCCGAGCGCACATGTACCAGGC CCGCAGCCTCTTTGCCGCCGACAGCAGCGGACTC TCAGACCCCTTTGCCCGCGTCTTCTTCATCAATCA GAGTCAGTGCACAGAGGTGCTGAATGAGACCCTG TGTCCCACCTGGGACCAGATGCTGGTGTTCGACA ACCTGGAGCTCTATGGTGAAGCTCATGAGCTGAG GGACGATCCGCCCATCATTGTCATTGAAATCTATG ACCAGGATTCCATGGGCAAAGCTGACTTCATGGG CCGGACCTTCGCCAAACCCCTGGTGAAGATGGCA GACGAGGCGTACTGCCCACCCCGCTTCCCACCTC AGCTCGAGTACTACCAGATCTACCGTGGCAACGC CACAGCTGGAGACCTGCTGGCGGCCTTCGAGCTG CTGCAGATTGGACCAGCAGGGAAGGCTGACCTGC CCCCCATCAATGGCCCGGTGGACGTGGACCGAGG TCCCATCATGCCCGTGCCCATGGGCATCCGGCCC GTGCTCAGCAAGTACCGAGTGGAGGTGCTGTTCT GGGGCCTACGGGACCTAAAGCGGGTGAACCTGGC CCAGGTGGACCGGCCACGGGTGGACATCGAGTGT GCAGGGAAGGGGGTGCAGTCGTCCCTGATCCACA ATTATAAGAAGAACCCCAACTTCAACACCCTCGT CAAGTGGTTTGAAGTGGACCTCCCAGAGAACGAG CTGCTGCACCCGCCCTTGAACATCCGTGTGGTGG ACTGCCGGGCCTTCGGTCGCTACACACTGGTGGG CTCCCATGCCGTCAGCTCCCTGCGACGCTTCATCT ACCGGCCCCCAGACCGCTCGGCCCCCAGCTGGAA CACCACGGTCAGGCTTCTCCGGCGCTGCCGTGTG CTGTGCAATGGGGGCTCCTCCTCTCACTCCACAGG GGAGGTTGTGGTGACTATGGAGCCAGAGGTACCC ATCAAGAAACTGGAGACCATGGTGAAGCTGGACG CGACTTCTGAAGCTGTTGTCAAGGTGGATGTGGC TGAGGAGGAGAAGGAGAAGAAGAAGAAGAAGA AGGGCACTGCGGAGGAGCCAGAGGAGGAGGAGC CAGACGAGAGCATGCTGGACTGGTGGTCCAAGTA CTTTGCCTCCATTGACACCATGAAGGAGCAACTTC GACAACAAGAGCCCTCTGGAATTGACTTGGAGGA GAAGGAGGAAGTGGACAATACCGAGGGCCTGAA GGGGTCAATGAAGGGCAAGGAGAAGGCAAGGGC TGCCAAAGAGGAGAAGAAGAAGAAAACTCAGAG CTCTGGCTCTGGCCAGGGGTCCGAGGCCCCCGAG AAGAAGAAACCCAAGATTGATGAGCTTAAGGTAT ACCCCAAAGAGCTGGAGTCCGAGTTTGATAACTT TGAGGACTGGCTGCACACTTTCAACTTGCTTCGGG GCAAGACCGGGGATGATGAGGATGGCTCCACCGA GGAGGAGCGCATTGTGGGACGCTTCAAGGGCTCC CTCTGCGTGTACAAAGTGCCACTCCCAGAGGACG TGTCCCGGGAAGCCGGCTACGACTCCACCTACGG CATGTTCCAGGGCATCCCGAGCAATGACCCCATC

[0173]

SEQ ID NO.	序列名称	序列
		AATGTGCTGGTCCGAGTCTATGTGGTCCGGGCCA CGGACCTGCACCCTGCTGACATCAACGGCAAAGC TGACCCCTACATCGCCATCCGGCTAGGCAAGACT GACATCCGCGACAAGGAGAACTACATCTCCAAGC AGCTCAACCCTGTCTTTGGGAAGTCCTTTGACATC GAGGCCTCCTTCCCCATGGAATCCATGCTGACGG TGGCTGTGTATGACTGGGACCTGGTGGGCACTGA TGACCTCATTGGGGAAACCAAGATCGACCTGGAG AACCGCTTCTACAGCAAGCACCGCGCCACCTGCG GCATCGCCCAGACCTACTCCACACATGGCTACAA TATCTGGCGGGACCCCATGAAGCCCAGCCAGATC CTGACCCGCCTCTGCAAAGACGGCAAAGTGGACG GCCCCCACTTTGGGCCCCCTGGGAGAGTGAAGGT GGCCAACCGCGTCTTCACTGGGCCCTCTGAGATT GAGGACGAGAACGGTCAGAGGAAGCCCACAGAC GAGCATGTGGCGCTGTTGGCCCTGAGGCACTGGG AGGACATCCCCCGCGCAGGCTGCCGCCTGGTGCC AGAGCATGTGGAGACGAGGCCGCTGCTCAACCCC GACAAGCCGGGCATCGAGCAGGGCCGCCTGGAG CTGTGGGTGGACATGTTCCCCATGGACATGCCAG CCCCTGGGACGCCTCTGGACATCTCACCTCGGAA GCCCAAGAAGTACGAGCTGCGGGTCATCATCTGG AACACAGATGAGGTGGTCTTGGAGGACGACGACT TCTTCACAGGGGAGAAGTCCAGTGACATCTTCGT GAGGGGGTGGCTGAAGGGCCAGCAGGAGGACAA GCAGGACACAGACGTCCACTACCACTCCCTCACT GGCGAGGGCAACTTCAACTGGCGCTACCTGTTCC CCTTCGACTACCTGGCGGCGGAGGAGAAGATCGT CATCTCCAAGAAGGAGTCCATGTTCTCCTGGGAC GAGACCGAGTACAAGATCCCCGCGCGGCTCACCC TGCAGATCTGGGATGCGGACCACTTCTCCGCTGA CGACTTCCTGGGGGCCATCGAGCTGGACCTGAAC CGGTTCCCGCGGGGCGCAAAGACAGCCAAGCAGT GCACCATGGAGATGGCCACCGGGGAGGTGGACGT GCCCCTCGTGTCCATCTTCAAGCAAAAGCGCGTC AAAGGCTGGTGGCCCCTCCTGGCCCGCAATGAGA ACGATGAGTTTGAGCTCACGGGCAAGGTGGAGGC TGAGCTGCATTTACTGACAGCAGAGGAGGCAGAG AAGAACCCAGTGGGCCTGGCCCGCAATGAACCTG ACCCCCTAGAGAAACCCAACCGGCCCGACACGAG CTTCATCTGGTTCCTGAACCCTCTCAAGTCGGCTC GCTACTTCTTGTGGCACACGTATCGCTGGCTGCTC CTCAAACTGTTGCTGCTCCTGCTGCTGCTCCTCCT CCTCGCCCTGTTCCTCTACTCTGTGCCTGGCTACC TGGTCAAGAAAATCCTCGGGGCCTGAGCCCAGTG GCCTCCTGGCCGGCCCGACACGGCCTTCGTCTGGT TCCTCAACCCTCTCAAGTCCATCAAGTACCTCATC TGCACCCGGTACAAGTGGCTCATCATCAAGATCG

[0174]

SEQ ID NO.	序列名称	序列
		TGCTGGCGCTGTTGGGGCTGCTCATGTTGGGGCTC TTCCTCTACAGCCTCCCTGGCTACATGGTCAAAAA GCTCCTTGGGGCATGAAGGCCGCCAGCTCCCGCC AGCCGCTCCCCAGCCCTGCCGCATTTCCTTTCAGT GGCTTGGACTCTTTCCCATCTCCCCTGGGGAGCCT GAGGAGCCCAGCGTCCACTCTTCATGCCTTGGGC CGAGCCTGCCTCCTGCTTGCGGGGGCCGCCTGTCC TCACTGCCCCAGGCTGCGGCTTGCCCAGTCCCGCC CCTCTGACCCCTGCCTGTGGGCTGGGGAGCCTTG GATGGGGTGGGGACCTGGAATGGGTCTCTCTTGC CCCACCTGGCTGAGGCGCCACCCTTCTTCAGGCCC AGGCTCCAGAGGAAGACTCCTGAAACCCTCCCCA GGTCTTCCAAGTACAGGATTGAAGCTTTAGTGAA ATTAACCAAGGACCATGGGTCAGTGCCCAGGGCT TTAAAAAGAATGAACGAGCAAAAGGTATCCCCGC CGTGACCCCTGCAGATAGCACCGGTCTTTGATCC GCAGCAGGGGCCAGACCCTGCCCACAAGTCCCAG CGCGGCTGCTTCTGCCACTGCTGGGCTCCACTTGG CTCCTCTCACTTCCCAGGGGGTCGCCTGTCCTGCC TGTGGGTTTCCATGGCTTCCCAGAGCTCCCTCTGC CCCAGCCAGCGCCTCCAGGCCCAGCTGAGGAGCT GTGAGAAGCAGCAGAGGGGACTCCCCATCCCGGG CACACCCTGTCCTCCCACCCCTGCCCCCTTGCCCT TCCAGCCCTTTCAGCTGCAGCTGGGAGCTGGCCC GTCAAGTGCTGCCCCTGCCTGTGTCTGGGTTTCTG TTGGCTGTTTTTCTTTTCTTGAGTGGTGATTTTTCT CTAAATAAAAGAAGTCAAGCACTGAAAAAAAAA AAAAAAA
14	编码人类耳畸蛋白同功型 5 蛋白 (SEQ ID NO: 5)的 DNA 序列，5994 bp，对应于以 NM_001287489 记载的编码序列	ATGGCCTTGCTCATCCACCTCAAGACAGTCTCGG AGCTGCGGGGCAGGGGCGACCGGATCGCCAAAG TGACTTTCCGAGGGCAATCCTTCTACTCTCGGGTC CTGGAGAACTGTGAGGATGTGGCTGACTTTGATG AGACATTTCGGTGGCCGGTGGCCAGCAGCATCGA CAGAAATGAGATGCTGGAGATTCAGGTTTTCAAC TACAGCAAAGTCTTCAGCAACAAGCTCATCGGGA CCTTCCGCATGGTGCTGCAGAAGGTGGTAGAGGA GAGCCATGTGGAGGTGACTGACACGCTGATTGAT GACAACAATGCTATCATCAAGACCAGCCTGTGCG TGGAGGTCCGGTATCAGGCCACTGACGGCACAGT GGGCTCCTGGGACGATGGGGACTTCCTGGGAGAT GAGTCTCTTCAAGAGGAAGAGAAGGACAGCCAA GAGACGGATGGACTGCTCCCAGGCTCCCGGCCCA GCTCCCGGCCCCCAGGAGAGAAGAGCTTCCGGAG AGCCGGGAGGAGCGTGTTCTCCGCCATGAAGCTC GGCAAAAACCGGTCTCACAAGGAGGAGCCCCAA AGACCAGATGAACCGGCGGTGCTGGAGATGGAA GACCTTGACCATCTGGCCATTCGGCTAGGAGATG GACTGGATCCCGACTCGGTGTCTCTAGCCTCAGTC

[0175]

SEQ ID NO.	序列名称	序列
		ACAGCTCTCACCACTAATGTCTCCAACAAGCGAT CTAAGCCAGACATTAAGATGGAGCCAAGTGCTGG GCGGCCCATGGATTACCAGGTCAGCATCACGGTG ATCGAGGCCCGGCAGCTGGTGGGCTTGAACATGG ACCCTGTGGTGTGCGTGGAGGTGGGTGACGACAA GAAGTACACATCCATGAAGGAGTCCACTAACTGC CCCTATTACAACGAGTACTTCGTCTTCGACTTCCA TGTCTCTCCGGATGTCATGTTTGACAAGATCATCA AGATTTCGGTGATTCACTCCAAGAACCTGCTGCG CAGTGGCACCCTGGTGGGCTCCTTCAAAATGGAC GTGGGAACCGTGTACTCGCAGCCAGAGCACCAGT TCCATCACAAGTGGGCCATCCTGTCTGACCCCGAT GACATCTCCTCGGGGCTGAAGGGCTACGTGAAGT GTGACGTTGCCGTGGTGGGCAAAGGGGACAACAT CAAGACGCCCCACAAGGCCAATGAGACCGACGA AGATGACATTGAGGGGAACTTGCTGCTCCCCGAG GGGGTGCCCCCCGAACGCCAGTGGGCCCGGTTCT ATGTGAAAATTTACCGAGCAGAGGGGCTGCCCCG TATGAACACAAGCCTCATGGCCAATGTAAAGAAG GCTTTCATCGGTGAAAACAAGGACCTCGTGGACC CCTACGTGCAAGTCTTCTTTGCTGGCCAGAAGGG CAAGACTTCAGTGCAGAAGAGCAGCTATGAGCCC CTGTGGAATGAGCAGGTCGTCTTTACAGACCTCTT CCCCCCACTCTGCAAACGCATGAAGGTGCAGATC CGAGACTCGGACAAGGTCAACGACGTGGCCATCG GCACCCACTTCATTGACCTGCGCAAGATTTCTAAT GACGGAGACAAAGGCTTCCTGCCCACACTGGGCC CAGCCTGGGTGAACATGTACGGCTCCACACGTAA CTACACGCTGCTGGATGAGCATCAGGACCTGAAC GAGGGCCTGGGGGAGGGTGTGTCCTTCCGGGCCC GGCTCCTGCTGGGCCTGGCTGTGGAGATCGTAGA CACCTCCAACCCTGAGCTCACCAGCTCCACAGAG GTGCAGGTGGAGCAGGCCACGCCCATCTCGGAGA GCTGTGCAGGTAAAATGGAAGAATTCTTTCTCTTT GGAGCCTTCCTGGAGGCCTCAATGATCGACCGGA GAAACGGAGACAAGCCCATCACCTTTGAGGTCAC CATAGGCAACTATGGGAACGAAGTTGATGGCCTG TCCCGGCCCCAGCGGCCTCGGCCCCGGAAGGAGC CGGGGGATGAGGAAGAAGTAGACCTGATTCAGA ACGCAAGTGATGACGAGGCCGGTGATGCCGGGG ACCTGGCCTCAGTCTCCTCCACTCCACCAATGCGG CCCCAGGTCACCGACAGGAACTACTTCCATCTGC CCTACCTGGAGCGAAAGCCCTGCATCTACATCAA GAGCTGGTGGCCGGACCAGCGCCGCCGCCTCTAC AATGCCAACATCATGGACCACATTGCCGACAAGC TGGAAGAAGGCCTGAACGACATACAGGAGATGA TCAAAACGGAGAAGTCCTACCCTGAGCGTCGCCT GCGGGGCGTCCTGGAGGAGCTGAGCTGTGGCTGC

[0176]

SEQ ID NO.	序列名称	序列
		TGCCGCTTCCTCTCCCTCGCTGACAAGGACCAGG GCCACTCATCCCGCACCAGGCTTGACCGGGAGCG CCTCAAGTCCTGCATGAGGGAGCTGGAAAACATG GGGCAGCAGGCCAGGATGCTGCGGGCCCAGGTG AAGCGGCACACGGTGCGGGACAAGCTGAGGCTGT GCCAGAACTTCCTGCAGAAGCTGCGCTTCCTGGC GGACGAGCCCCAGCACAGCATTCCCGACATCTTC ATCTGGATGATGAGCAACAACAAGCGTGTCGCCT ATGCCCGTGTGCCCTCCAAGGACCTGCTCTTCTCC ATCGTGGAGGAGGAGACTGGCAAGGACTGCGCC AAGGTCAAGACGCTCTTCCTTAAGCTGCCAGGGA AGCGGGGCTTCGGCTCGGCAGGCTGGACAGTGCA GGCCAAGGTGGAGCTGTACCTGTGGCTGGGCCTC AGCAAACAGCGCAAGGAGTTCCTGTGCGGCCTGC CCTGTGGCTTCCAGGAGGTCAAGGCAGCCCAGGG CCTGGGCCTGCATGCCTTCCCACCCGTCAGCCTGG TCTACACCAAGAAGCAGGCGTTCCAGCTCCGAGC GCACATGTACCAGGCCCGCAGCCTCTTTGCCGCC GACAGCAGCGGACTCTCAGACCCCTTTGCCCGCG TCTTCTTCATCAATCAGAGTCAGTGCACAGAGGT GCTGAATGAGACCCTGTGTCCCACCTGGGACCAG ATGCTGGTGTTCGACAACCTGGAGCTCTATGGTG AAGCTCATGAGCTGAGGGACGATCCGCCCATCAT TGTCATTGAAATCTATGACCAGGATTCCATGGGC AAAGCTGACTTCATGGGCCGGACCTTCGCCAAAC CCCTGGTGAAGATGGCAGACGAGGCGTACTGCCC ACCCCGCTTCCCACCTCAGCTCGAGTACTACCAG ATCTACCGTGGCAACGCCACAGCTGGAGACCTGC TGGCGGCCTTCGAGCTGCTGCAGATTGGACCAGC AGGGAAGGCTGACCTGCCCCCCATCAATGGCCCG GTGGACGTGGACCGAGGTCCCATCATGCCCGTGC CCATGGGCATCCGGCCCGTGCTCAGCAAGTACCG AGTGGAGGTGCTGTTCTGGGGCCTACGGGACCTA AAGCGGGTGAACCTGGCCCAGGTGGACCGGCCAC GGGTGGACATCGAGTGTGCAGGGAAGGGGGTGC AGTCGTCCCTGATCCACAATTATAAGAAGAACCC CAACTTCAACACCCTCGTCAAGTGGTTTGAAGTG GACCTCCCAGAGAACGAGCTGCTGCACCCGCCCT TGAACATCCGTGTGGTGGACTGCCGGGCCTTCGG TCGCTACACACTGGTGGGCTCCCATGCCGTCAGCT CCCTGCGACGCTTCATCTACCGGCCCCCAGACCG CTCGGCCCCCAGCTGGAACACCACGGTCAGGCTT CTCCGGCGCTGCCGTGTGCTGTGCAATGGGGGCT CCTCCTCTCACTCCACAGGGGAGGTTGTGGTGACT ATGGAGCCAGAGGTACCCATCAAGAAACTGGAG ACCATGGTGAAGCTGGACGCGACTTCTGAAGCTG TTGTCAAGGTGGATGTGGCTGAGGAGGAGAAGGA GAAGAAGAAGAAGAAGAAGGGCACTGCGGAGGA

[0177]

SEQ ID NO.	序列名称	序列
		GCCAGAGGAGGAGGAGCCAGACGAGAGCATGCT GGACTGGTGGTCCAAGTACTTTGCCTCCATTGACA CCATGAAGGAGCAACTTCGACAACAAGAGCCCTC TGGAATTGACTTGGAGGAGAAGGAGGAAGTGGA CAATACCGAGGGCCTGAAGGGGTCAATGAAGGG CAAGGAGAAGGCAAGGGCTGCCAAAGAGGAGAA GAAGAAGAAAACTCAGAGCTCTGGCTCTGGCCAG GGGTCCGAGGCCCCCGAGAAGAAGAAACCCAAG ATTGATGAGCTTAAGGTATACCCCAAAGAGCTGG AGTCCGAGTTTGATAACTTTGAGGACTGGCTGCA CACTTTCAACTTGCTTCGGGGCAAGACCGGGGAT GATGAGGATGGCTCCACCGAGGAGGAGCGCATTG TGGGACGCTTCAAGGGCTCCCTCTGCGTGTACAA AGTGCCACTCCCAGAGGACGTGTCCCGGGAAGCC GGCTACGACTCCACCTACGGCATGTTCCAGGGCA TCCCGAGCAATGACCCCATCAATGTGCTGGTCCG AGTCTATGTGGTCCGGGCCACGGACCTGCACCCT GCTGACATCAACGGCAAAGCTGACCCCTACATCG CCATCCGGCTAGGCAAGACTGACATCCGCGACAA GGAGAACTACATCTCCAAGCAGCTCAACCCTGTC TTTGGGAAGTCCTTTGACATCGAGGCCTCCTTCCC CATGGAATCCATGCTGACGGTGGCTGTGTATGAC TGGGACCTGGTGGGCACTGATGACCTCATTGGGG AAACCAAGATCGACCTGGAGAACCGCTTCTACAG CAAGCACCGCGCCACCTGCGGCATCGCCCAGACC TACTCCACACATGGCTACAATATCTGGCGGGACC CCATGAAGCCCAGCCAGATCCTGACCCGCCTCTG CAAAGACGGCAAAGTGGACGGCCCCCACTTTGGG CCCCCTGGGAGAGTGAAGGTGGCCAACCGCGTCT TCACTGGGCCCTCTGAGATTGAGGACGAGAACGG TCAGAGGAAGCCCACAGACGAGCATGTGGCGCTG TTGGCCCTGAGGCACTGGGAGGACATCCCCCGCG CAGGCTGCCGCCTGGTGCCAGAGCATGTGGAGAC GAGGCCGCTGCTCAACCCCGACAAGCCGGGCATC GAGCAGGGCCGCCTGGAGCTGTGGGTGGACATGT TCCCCATGGACATGCCAGCCCCTGGGACGCCTCT GGACATCTCACCTCGGAAGCCCAAGAAGTACGAG CTGCGGGTCATCATCTGGAACACAGATGAGGTGG TCTTGGAGGACGACGACTTCTTCACAGGGGAGAA GTCCAGTGACATCTTCGTGAGGGGGTGGCTGAAG GGCCAGCAGGAGGACAAGCAGGACACAGACGTC CACTACCACTCCCTCACTGGCGAGGGCAACTTCA ACTGGCGCTACCTGTTCCCCTTCGACTACCTGGCG GCGGAGGAGAAGATCGTCATCTCCAAGAAGGAGT CCATGTTCTCCTGGGACGAGACCGAGTACAAGAT CCCCGCGCGGCTCACCCTGCAGATCTGGGATGCG GACCACTTCTCCGCTGACGACTTCCTGGGGGCCAT CGAGCTGGACCTGAACCGGTTCCCGCGGGGCGCA

[0178]

SEQ ID NO.	序列名称	序列
		AAGACAGCCAAGCAGTGCACCATGGAGATGGCC ACCGGGGAGGTGGACGTGCCCCTCGTGTCCATCT TCAAGCAAAAGCGCGTCAAAGGCTGGTGGCCCCT CCTGGCCCGCAATGAGAACGATGAGTTTGAGCTC ACGGGCAAGGTGGAGGCTGAGCTGCATTTACTGA CAGCAGAGGAGGCAGAGAAGAACCCAGTGGGCC TGGCCCGCAATGAACCTGACCCCCTAGAGAAACC CAACCGGCCCGACACGGCCTTCGTCTGGTTCCTCA ACCCTCTCAAGTCCATCAAGTACCTCATCTGCACC CGGTACAAGTGGCTCATCATCAAGATCGTGCTGG CGCTGTTGGGGCTGCTCATGTTGGGGCTCTTCCTC TACAGCCTCCCTGGCTACATGGTCAAAAAGCTCC TTGGGGCATGA
15	mOTOF-201_1 转录物 (NM_031875.2)，小鼠耳畴蛋白转录物变体 2，7125 bp，编码 SEQ ID NO:6 的蛋白质	TTGGTTGCCTTGGTCTCTGTGGGCAGCAGCAGGA GGAGGCGGCAGCAGCCAGAGAAGAGGGAGGCGT GTGAGCCACACTCCACCAGCGAGCTTCTTCCCGCT GCTCTGGAACTGCCCAGGCTCTCCCCACCAGCAT GGCCCTGATTGTTCACCTCAAGACTGTCTCAGAGC TCCGAGGCAAAGGTGACCGGATTGCCAAAGTCAC TTTCCGAGGGCAGTCTTTCTACTCCCGGGTCCTGG AGAACTGCGAGGGTGTGGCTGACTTTGATGAGAC GTTCCGGTGGCCAGTGGCCAGCAGCATCGACCGG AATGAAGTGTTGGAGATTCAGATTTTCAACTACA GCAAAGTCTTCAGCAACAAGCTGATAGGGACCTT CTGCATGGTGCTGCAGAAAGTGGTGGAGGAGAAT CGGGTAGAGGTGACCGACACGCTGATGGATGACA GCAATGCTATCATCAAGACCAGCCTGAGCATGGA GGTCCGGTATCAGGCCACAGATGGCACTGTGGGC CCCTGGGATGATGGAGACTTCCTGGGAGATGAAT CCCTCCAGGAGGAGAAGGACAGCCAGGAGACAG ATGGGCTGCTACCTGGTTCCCGACCCAGCACCCG GATATCTGGCGAGAAGAGCTTTCGCAGAGCGGGA AGGAGTGTGTTCTCGGCCATGAAACTCGGCAAAA CTCGGTCCCACAAAGAGGAGCCCCAAAGACAAG ATGAGCCAGCAGTGCTGGAGATGGAGGACCTGGA CCACCTAGCCATTCAGCTGGGGGATGGGCTGGAT CCTGACTCCGTGTCTCTAGCCTCGGTCACCGCTCT CACCAGCAATGTCTCCAACAAACGGTCTAAGCCA GATATTAAGATGGAGCCCAGTGCTGGAAGGCCCA TGGATTACCAGGTCAGCATCACAGTGATTGAGGC TCGGCAGCTGGTGGGCTTGAACATGGACCCTGTG GTGTGTGTGGAGGTGGGTGATGACAAGAAATACA CGTCAATGAAGGAGTCCACAAACTGCCCTTACTA CAACGAGTACTTTGTCTTCGACTTCCATGTCTCTC CTGATGTCATGTTTGACAAGATCATCAAGATCTCG GTTATCCATTCTAAGAACCTGCTTCGGAGCGGCA CCCTGGTGGGTTCCTTCAAAATGGATGTGGGGAC TGTGTATTCCCAGCCTGAACACCAGTTCCATCACA

[0179]

SEQ ID NO.	序列名称	序列
		AATGGGCCATCCTGTCAGACCCCGATGACATCTC TGCTGGGTTGAAGGGTTATGTAAAGTGTGATGTC GCTGTGGTGGGCAAGGGAGACAACATCAAGACA CCCCACAAGGCCAACGAGACGGATGAGGACGAC ATTGAAGGGAACTTGCTGCTCCCCGAGGGCGTGC CCCCCGAACGGCAGTGGGCACGGTTCTATGTGAA AATTTACCGAGCAGAGGGACTGCCCCGGATGAAC ACAAGCCTCATGGCCAACGTGAAGAAGGCGTTCA TCGGTGAGAACAAGGACCTCGTCGACCCCTATGT GCAAGTCTTCTTTGCTGGACAAAAGGGCAAAACA TCAGTGCAGAAGAGCAGCTATGAGCCGCTATGGA ATGAGCAGGTCGTCTTCACAGACTTGTTCCCCCCA CTCTGCAAACGCATGAAGGTGCAGATCCGGGACT CTGACAAGGTCAATGATGTGGCCATCGGCACCCA CTTCATCGACCTGCGCAAGATTTCCAACGATGGA GACAAAGGCTTCCTGCCTACCCTCGGTCCAGCCT GGGTGAACATGTACGGCTCCACGCGCAACTACAC ACTGCTGGACGAGCACCAGGACTTGAATGAAGGC CTGGGGGAGGGTGTGTCCTTCCGGGCCCGCCTCA TGTTGGGACTAGCTGTGGAGATCCTGGACACCTC CAACCCAGAGCTCACCAGCTCCACGGAGGTGCAG GTGGAGCAGGCCACGCCTGTCTCGGAGAGCTGCA CAGGGAGAATGGAAGAATTTTTTCTATTTGGAGC CTTCTTGGAAGCCTCAATGATTGACCGGAAAAAT GGGGACAAGCCAATTACCTTTGAGGTGACCATAG GAAACTACGGCAATGAAGTCGATGGTATGTCCCG GCCCCTGAGGCCTCGGCCCCGGAAAGAGCCTGGG GATGAAGAAGAGGTAGACCTGATTCAGAACTCCA GTGACGATGAAGGTGACGAAGCCGGGGACCTGG CCTCGGTGTCCTCCACCCCACCTATGCGGCCCCAG ATCACGGACAGGAACTATTTCCACCTGCCCTACCT GGAGCGCAAGCCCTGCATCTATATCAAGAGCTGG TGGCCTGACCAGAGGCGGCGCCTCTACAATGCCA ACATCATGGATCACATTGCTGACAAGCTGGAAGA AGGCCTGAATGATGTACAGGAGATGATCAAAACG GAGAAGTCCTACCCGGAGCGCCGCCTGCGGGGTG TGCTAGAGGAACTCAGCTGTGGCTGCCACCGCTT CCTCTCCCTCTCGGACAAGGACCAGGGCCGCTCG TCCCGCACCAGGCTGGATCGAGAGCGTCTTAAGT CCTGTATGAGGGAGTTGGAGAGCATGGGACAGCA GGCCAAGAGCCTGAGGGCTCAGGTGAAGCGGCA CACTGTTCGGGACAAGCTGAGGTCATGCCAGAAC TTTCTGCAGAAGCTACGCTTCCTGGCGGATGAGC CCCAGCACAGCATTCCTGATGTGTTCATTTGGATG ATGAGCAACAACAAACGTATCGCCTATGCCCGCG TGCCTTCCAAAGACCTGCTCTTCTCCATCGTGGAG GAGGAACTGGGCAAGGACTGCGCCAAAGTCAAG ACCCTCTTCCTGAAGCTGCCAGGGAAGAGGGGCT

[0180]

SEQ ID NO.	序列名称	序列
		TCGGCTCGGCAGGCTGGACAGTACAGGCCAAGCT GGAGCTCTACCTGTGGCTGGGCCTCAGCAAGCAG CGAAAGGACTTCCTGTGTGGTCTGCCCTGTGGCTT CGAGGAGGTCAAGGCAGCCCAAGGCCTGGGCCTG CATTCCTTTCCGCCCATCAGCCTAGTCTACACCAA GAAGCAAGCCTTCCAGCTCCGAGCACACATGTAT CAGGCCCGAAGCCTCTTTGCTGCTGACAGCAGTG GGCTCTCTGATCCCTTTGCCCGTGTCTTCTTCATC AACCAGAGCCAATGCACTGAGGTTCTAAACGAGA CACTGTGTCCCACCTGGGACCAGATGCTGGTATTT GACAACCTGGAGCTGTACGGTGAAGCTCACGAGT TACGAGATGATCCCCCCATCATTGTCATTGAAATC TACGACCAGGACAGCATGGGCAAAGCCGACTTCA TGGGCCGGACCTTCGCCAAGCCCCTGGTGAAGAT GGCAGATGAAGCATACTGCCCACCTCGCTTCCCG CCGCAGCTTGAGTACTACCAGATCTACCGAGGCA GTGCCACTGCCGGAGACCTACTGGCTGCCTTCGA GCTGCTGCAGATTGGGCCATCAGGGAAGGCTGAC CTGCCACCCATCAATGGCCCAGTGGACATGGACA GAGGGCCCATCATGCCTGTGCCCGTGGGAATCCG GCCAGTGCTCAGCAAGTACCGAGTGGAGGTGCTG TTCTGGGGCCTGAGGGACCTAAAGAGGGTGAACC TGGCCCAGGTGGACCGACCACGGGTGGACATCGA GTGTGCAGGAAAGGGGGTACAATCCTCCCTGATT CACAATTATAAGAAGAACCCCAACTTCAACACGC TGGTCAAGTGGTTTGAAGTGGACCTCCCGGAGAA TGAGCTCCTGCACCCACCCTTGAACATCCGAGTG GTAGATTGCCGGGCCTTTGGACGATACACCCTGG TGGGTTCCCACGCAGTCAGCTCACTGAGGCGCTT CATCTACCGACCTCCAGACCGCTCAGCCCCCAAC TGGAACACCACAGTCAGGCTGCTCCGGGGCTGCC ACAGGCTGCGCAATGGGGGCCCCTCTTCTCGCCC CACAGGGGAGGTTGTAGTAAGCATGGAGCCTGAG GAGCCAGTTAAGAAGCTGGAGACCATGGTGAAAC TGGATGCGACTTCTGATGCTGTGGTCAAGGTGGA TGTGGCTGAAGATGAGAAGGAAAGGAAGAAGAA GAAAAAGAAAGGCCCGTCAGAGGAGCCAGAGGA GGAAGAGCCCGATGAGAGCATGCTGGATTGGTGG TCCAAGTACTTCGCCTCCATCGACACAATGAAGG AGCAACTTCGACAACATGAGACCTCTGGAACTGA CTTGGAAGAGAAGGAAGAGATGGAAAGCGCTGA GGGCCTGAAGGGACCAATGAAGAGCAAGGAGAA GTCCAGAGCTGCAAAGGAGGAGAAAAAGAAGAA AAACCAGAGCCCTGGCCCTGGCCAGGGATCGGAG GCTCCTGAGAAGAAGAAAGCCAAGATCGATGAG CTTAAGGTGTACCCCAAGGAGCTGGAATCGGAGT TTGACAGCTTTGAGGACTGGCTGCACACCTTCAA CCTGTTGAGGGGCAAGACGGGAGATGATGAGGAT

[0181]

SEQ ID NO.	序列名称	序列
		GGCTCCACAGAGGAGGAGCGCATAGTAGGCCGAT TCAAGGGCTCCCTCTGTGTGTACAAAGTGCCACTC CCAGAAGATGTATCTCGAGAAGCTGGCTATGATC CCACCTATGGAATGTTCCAGGGCATCCCAAGCAA TGACCCCATCAATGTGCTGGTCCGAATCTATGTGG TCCGGGCCACAGACCTGCACCCGGCCGACATCAA TGGCAAAGCTGACCCCTATATTGCCATCAAGTTA GGCAAGACCGACATCCGAGACAAGGAGAACTAC ATCTCCAAGCAGCTCAACCCTGTGTTTGGGAAGT CCTTTGACATTGAGGCCTCCTTCCCCATGGAGTCC ATGTTGACAGTGGCCGTGTACGACTGGGATCTGG TGGGCACTGATGACCTCATCGGAGAAACCAAGAT TGACCTGGAAAACCGCTTCTACAGCAAGCATCGC GCCACCTGCGGCATCGCACAGACCTATTCCATAC ATGGCTACAATATCTGGAGGGACCCCATGAAGCC CAGCCAGATCCTGACACGCCTCTGTAAAGAGGGC AAAGTGGACGGCCCCCACTTTGGTCCCCATGGGA GAGTGAGGGTTGCCAACCGTGTCTTCACGGGGCC TTCAGAAATAGAGGATGAGAATGGTCAGAGGAA GCCCACAGATGAGCACGTGGCACTGTCTGCTCTG AGACACTGGGAGGACATCCCCCGGGTGGGCTGCC GCCTTGTGCCGGAACACGTGGAGACCAGGCCGCT GCTCAACCCTGACAAGCCAGGCATTGAGCAGGGC CGCCTGGAGCTGTGGGTGGACATGTTCCCCATGG ACATGCCAGCCCCTGGGACACCTCTGGATATATC CCCCAGGAAACCCAAGAAGTACGAGCTGCGGGTC ATCGTGTGGAACACAGACGAGGTGGTCCTGGAAG ACGATGATTTCTTCACGGGAGAGAAGTCCAGTGA CATTTTTGTGAGGGGGTGGCTGAAGGGCCAGCAG GAGGACAAACAGGACACAGATGTCCACTATCACT CCCTCACGGGGGAGGGCAACTTCAACTGGAGATA CCTCTTCCCCTTCGACTACCTAGCGGCCGAAGAG AAGATCGTTATGTCCAAAAAGGAGTCTATGTTCT CCTGGGATGAGACGGAGTACAAGATCCCTGCGCG GCTCACCCTGCAGATCTGGGACGCTGACCACTTCT CGGCTGACGACTTCCTGGGGGCTATCGAGCTGGA CCTGAACCGGTTCCCGAGGGGCGCTAAGACAGCC AAGCAGTGCACCATGGAGATGGCCACCGGGGAG GTGGACGTACCCCTGGTTTCCATCTTTAAACAGAA ACGTGTCAAAGGCTGGTGGCCCCTCCTGGCCCGC AATGAGAATGATGAGTTTGAGCTCACAGGCAAAG TGGAGGCGGAGCTACACCTACTCACGGCAGAGGA GGCAGAGAAGAACCCTGTGGGCCTGGCTCGCAAT GAACCTGATCCCCTAGAAAAACCCAATCGGCCGG ACACAAGCTTCATCTGGTTCTTGAACCCTCTCAAG TCTGCCCGCTACTTCCTGTGGCATACCTACCGCTG GCTACTCCTCAAATTCCTGCTGCTCTTCCTCCTGC TGCTGCTCTTCGCCCTGTTTCTCTACTCTCTGCCTG

[0182]

SEQ ID NO.	序列名称	序列
		GCTACCTGGCCAAGAAGATCCTTGGGGCCTGAGC CCTGCAGTCGCCTAGGCCTGCCGGCCTGACACGG CATTCGTCTGGTTCCTGAACCCACTCAAATCTATC AAGTACCTCATCTGCACCCGGTACAAGTGGCTGA TCATCAAGATCGTGCTGGCGCTGCTGGGGCTGCT CATGCTGGCCCTCTTCCTTTACAGCCTCCCAGGCT ACATGGTCAAGAAGCTCCTAGGGGCCTGAAGTGT GCCCCACCCCAGCCCGCTCCAGCATCCCTCCAGG GGCTGCTGCGTATTTTGCCTTCCCTCACCTGGACT CTCTCCCAACTCCCTGAGGAGCCCTCCCACGCCTG CCAGCCTTGAGCAAGACACCTGCTTGCTGGACTT CATCCCCACCCCACACCCAAACTGTTGCTTGCCTG ATCTTGTCCCAGGCCTGCCTGGGGTTTGGGGCAC AGTTGGCCTCCAAAACCAGATACCCTCTTGTCTAA AGTACCAGGTTCCTCTGCCCAACCCCAAGAGTGG TAGTGGCCCAACCCTCCCTGTGCTTTCCAAATCTT GTCTTAAGGCACCAGTGAAATTAACCAAGAAACG CGGAGCAATGCCCAAGGCTCTGATGAGTAGGAAC ACGTGGAAAGCACCAGGAATGCCAGCAGAGGCG AGGCGGCACACCTCTCTGCAGAGCATCCAGGCCG AGCGGCGGGCAGCGGCCAGCTGCTTCTGCGCATG CTCTCCTCTTGGCTCTGCTTCTTTCTCACAGTCACA GTCACTTCACAGCTTAGCCTTGGGCTTCCCATCAC TTCCAGGGGTGCCTCTGCCTTGGCCAGTGTGTGTC AGCTAGTACACAAGCTCCAAGTGTGAATCAGGTG TACTGGCCGTCCTGAAGACTGACTGCCCTGTCCTT CCTGCCGACAGCCACACCCGAGTGTACACTTAAA GCGGTGCCCTTCTGCCTCTGTGGGCCTGCTGGCTG CTGTTCCTTTCTTGAGTGTGATTTTTTTTTTCTCTC CCTCAATAAAATAAATCAAACTCTGAGAC
16	mOTOF-201_2 转录物 (NM_001286421.1)，小鼠耳畸蛋白转录物变体 3，7065 bp，编码 SEQ ID NO: 7 的蛋白质	TTGGTTGCCTTGGTCTCTGTGGGCAGCAGCAGGA GGAGGCGGCAGCAGCCAGAGAAGAGGGAGGCGT GTGAGCCACACTCCACCAGCGAGCTTCTTCCCGCT GCTCTGGAACTGCCCAGGCTCTCCCCACCAGCAT GGCCCTGATTGTTCACCTCAAGACTGTCTCAGAGC TCCGAGGCAAAGGTGACCGGATTGCCAAAGTCAC TTTCCGAGGGCAGTCTTTCTACTCCCGGGTCCTGG AGAACTGCGAGGGTGTGGCTGACTTTGATGAGAC GTTCCGGTGGCCAGTGGCCAGCAGCATCGACCGG AATGAAGTGTTGGAGATTCAGATTTTCAACTACA GCAAAGTCTTCAGCAACAAGCTGATAGGGACCTT CTGCATGGTGCTGCAGAAAGTGGTGGAGGAGAAT CGGGTAGAGGTGACCGACACGCTGATGGATGACA GCAATGCTATCATCAAGACCAGCCTGAGCATGGA GGTCCGGTATCAGGCCACAGATGGCACTGTGGGC CCCTGGGATGATGGAGACTTCCTGGGAGATGAAT CCCTCCAGGAGGAGAAGGACAGCCAGGAGACAG ATGGGCTGCTACCTGGTTCCCGACCCAGCACCCG

[0183]

SEQ ID NO.	序列名称	序列
		GATATCTGGCGAGAAGAGCTTTCGCAGAGCGGGA AGGAGTGTGTTCTCGGCCATGAAACTCGGCAAAA CTCGGTCCCACAAAGAGGAGCCCCAAAGACAAG ATGAGCCAGCAGTGCTGGAGATGGAGGACCTGGA CCACCTAGCCATTCAGCTGGGGGATGGGCTGGAT CCTGACTCCGTGTCTCTAGCCTCGGTCACCGCTCT CACCAGCAATGTCTCCAACAAACGGTCTAAGCCA GATATTAAGATGGAGCCCAGTGCTGGAAGGCCCA TGGATTACCAGGTCAGCATCACAGTGATTGAGGC TCGGCAGCTGGTGGGCTTGAACATGGACCCTGTG GTGTGTGTGGAGGTGGGTGATGACAAGAAATACA CGTCAATGAAGGAGTCCACAAACTGCCCTTACTA CAACGAGTACTTTGTCTTCGACTTCCATGTCTCTC CTGATGTCATGTTTGACAAGATCATCAAGATCTCG GTTATCCATTCTAAGAACCTGCTTCGGAGCGGCA CCCTGGTGGGTTCCTTCAAAATGGATGTGGGGAC TGTGTATTCCCAGCCTGAACACCAGTTCCATCACA AATGGGCCATCCTGTCAGACCCCGATGACATCTC TGCTGGGTTGAAGGGTTATGTAAAGTGTGATGTC GCTGTGGTGGGCAAGGGAGACAACATCAAGACA CCCCACAAGGCCAACGAGACGGATGAGGACGAC ATTGAAGGGAACTTGCTGCTCCCCGAGGGCGTGC CCCCCGAACGGCAGTGGGCACGGTTCTATGTGAA AATTTACCGAGCAGAGGGACTGCCCCGGATGAAC ACAAGCCTCATGGCCAACGTGAAGAAGGCGTTCA TCGGTGAGAACAAGGACCTCGTCGACCCCTATGT GCAAGTCTTCTTTGCTGGACAAAAGGGCAAAACA TCAGTGCAGAAGAGCAGCTATGAGCCGCTATGGA ATGAGCAGGTCGTCTTCACAGACTTGTTCCCCCCA CTCTGCAAACGCATGAAGGTGCAGATCCGGGACT CTGACAAGGTCAATGATGTGGCCATCGGCACCCA CTTCATCGACCTGCGCAAGATTTCCAACGATGGA GACAAAGGCTTCCTGCCTACCCTCGGTCCAGCCT GGGTGAACATGTACGGCTCCACGCGCAACTACAC ACTGCTGGACGAGCACCAGGACTTGAATGAAGGC CTGGGGGAGGGTGTGTCCTTCCGGGCCCGCCTCA TGTTGGGACTAGCTGTGGAGATCCTGGACACCTC CAACCCAGAGCTCACCAGCTCCACGGAGGTGCAG GTGGAGCAGGCCACGCCTGTCTCGGAGAGCTGCA CAGGGAGAATGGAAGAATTTTTTCTATTTGGAGC CTTCTTGGAAGCCTCAATGATTGACCGGAAAAAT GGGGACAAGCCAATTACCTTTGAGGTGACCATAG GAAACTACGGCAATGAAGTCGATGGTATGTCCCG GCCCCTGAGGCCTCGGCCCCGGAAAGAGCCTGGG GATGAAGAAGAGGTAGACCTGATTCAGAACTCCA GTGACGATGAAGGTGACGAAGCCGGGGACCTGG CCTCGGTGTCCTCCACCCCACCTATGCGGCCCCAG ATCACGGACAGGAACTATTTCCACCTGCCCTACCT

[0184]

SEQ ID NO.	序列名称	序列
		GGAGCGCAAGCCCTGCATCTATATCAAGAGCTGG TGGCCTGACCAGAGGCGGCGCCTCTACAATGCCA ACATCATGGATCACATTGCTGACAAGCTGGAAGA AGGCCTGAATGATGTACAGGAGATGATCAAAACG GAGAAGTCCTACCCGGAGCGCCGCCTGCGGGGTG TGCTAGAGGAACTCAGCTGTGGCTGCCACCGCTT CCTCTCCCTCTCGGACAAGGACCAGGGCCGCTCG TCCCGCACCAGGCTGGATCGAGAGCGTCTTAAGT CCTGTATGAGGGAGTTGGAGAGCATGGGACAGCA GGCCAAGAGCCTGAGGGCTCAGGTGAAGCGGCA CACTGTTCGGGACAAGCTGAGGTCATGCCAGAAC TTTCTGCAGAAGCTACGCTTCCTGGCGGATGAGC CCCAGCACAGCATTCCTGATGTGTTCATTTGGATG ATGAGCAACAACAAACGTATCGCCTATGCCCGCG TGCCTTCCAAAGACCTGCTCTTCTCCATCGTGGAG GAGGAACTGGGCAAGGACTGCGCCAAAGTCAAG ACCCTCTTCCTGAAGCTGCCAGGGAAGAGGGGCT TCGGCTCGGCAGGCTGGACAGTACAGGCCAAGCT GGAGCTCTACCTGTGGCTGGGCCTCAGCAAGCAG CGAAAGGACTTCCTGTGTGGTCTGCCCTGTGGCTT CGAGGAGGTCAAGGCAGCCCAAGGCCTGGGCCTG CATTCCTTTCCGCCCATCAGCCTAGTCTACACCAA GAAGCAAGCCTTCCAGCTCCGAGCACACATGTAT CAGGCCCGAAGCCTCTTTGCTGCTGACAGCAGTG GGCTCTCTGATCCCTTTGCCCGTGTCTTCTTCATC AACCAGAGCCAATGCACTGAGGTTCTAAACGAGA CACTGTGTCCCACCTGGGACCAGATGCTGGTATTT GACAACCTGGAGCTGTACGGTGAAGCTCACGAGT TACGAGATGATCCCCCCATCATTGTCATTGAAATC TACGACCAGGACAGCATGGGCAAAGCCGACTTCA TGGGCCGGACCTTCGCCAAGCCCCTGGTGAAGAT GGCAGATGAAGCATACTGCCCACCTCGCTTCCCG CCGCAGCTTGAGTACTACCAGATCTACCGAGGCA GTGCCACTGCCGGAGACCTACTGGCTGCCTTCGA GCTGCTGCAGATTGGGCCATCAGGGAAGGCTGAC CTGCCACCCATCAATGGCCCAGTGGACATGGACA GAGGGCCCATCATGCCTGTGCCCGTGGGAATCCG GCCAGTGCTCAGCAAGTACCGAGTGGAGGTGCTG TTCTGGGGCCTGAGGGACCTAAAGAGGGTGAACC TGGCCCAGGTGGACCGACCACGGGTGGACATCGA GTGTGCAGGAAAGGGGGTACAATCCTCCCTGATT CACAATTATAAGAAGAACCCCAACTTCAACACGC TGGTCAAGTGGTTTGAAGTGGACCTCCCGGAGAA TGAGCTCCTGCACCCACCCTTGAACATCCGAGTG GTAGATTGCCGGGCCTTTGGACGATACACCCTGG TGGGTTCCCACGCAGTCAGCTCACTGAGGCGCTT CATCTACCGACCTCCAGACCGCTCAGCCCCCAAC TGGAACACCACAGGGGAGGTTGTAGTAAGCATGG

[0185]

SEQ ID NO.	序列名称	序列
		AGCCTGAGGAGCCAGTTAAGAAGCTGGAGACCAT GGTGAAACTGGATGCGACTTCTGATGCTGTGGTC AAGGTGGATGTGGCTGAAGATGAGAAGGAAAGG AAGAAGAAGAAAAAGAAAGGCCCGTCAGAGGAG CCAGAGGAGGAAGAGCCCGATGAGAGCATGCTG GATTGGTGGTCCAAGTACTTCGCCTCCATCGACAC AATGAAGGAGCAACTTCGACAACATGAGACCTCT GGAACTGACTTGGAAGAGAAGGAAGAGATGGAA AGCGCTGAGGGCCTGAAGGGACCAATGAAGAGC AAGGAGAAGTCCAGAGCTGCAAAGGAGGAGAAA AAGAAGAAAAACCAGAGCCCTGGCCCTGGCCAG GGATCGGAGGCTCCTGAGAAGAAGAAAGCCAAG ATCGATGAGCTTAAGGTGTACCCCAAGGAGCTGG AATCGGAGTTTGACAGCTTTGAGGACTGGCTGCA CACCTTCAACCTGTTGAGGGGCAAGACGGGAGAT GATGAGGATGGCTCCACAGAGGAGGAGCGCATA GTAGGCCGATTCAAGGGCTCCCTCTGTGTGTACA AAGTGCCACTCCCAGAAGATGTATCTCGAGAAGC TGGCTATGATCCCACCTATGGAATGTTCCAGGGC ATCCCAAGCAATGACCCCATCAATGTGCTGGTCC GAATCTATGTGGTCCGGGCCACAGACCTGCACCC GGCCGACATCAATGGCAAAGCTGACCCCTATATT GCCATCAAGTTAGGCAAGACCGACATCCGAGACA AGGAGAACTACATCTCCAAGCAGCTCAACCCTGT GTTTGGGAAGTCCTTTGACATTGAGGCCTCCTTCC CCATGGAGTCCATGTTGACAGTGGCCGTGTACGA CTGGGATCTGGTGGGCACTGATGACCTCATCGGA GAAACCAAGATTGACCTGGAAAACCGCTTCTACA GCAAGCATCGCGCCACCTGCGGCATCGCACAGAC CTATTCCATACATGGCTACAATATCTGGAGGGAC CCCATGAAGCCCAGCCAGATCCTGACACGCCTCT GTAAAGAGGGCAAAGTGGACGGCCCCCACTTTGG TCCCCATGGGAGAGTGAGGGTTGCCAACCGTGTC TTCACGGGGCCTTCAGAAATAGAGGATGAGAATG GTCAGAGGAAGCCCACAGATGAGCACGTGGCACT GTCTGCTCTGAGACACTGGGAGGACATCCCCCGG GTGGGCTGCCGCCTTGTGCCGGAACACGTGGAGA CCAGGCCGCTGCTCAACCCTGACAAGCCAGGCAT TGAGCAGGGCCGCCTGGAGCTGTGGGTGGACATG TTCCCCATGGACATGCCAGCCCCTGGGACACCTCT GGATATATCCCCCAGGAAACCCAAGAAGTACGAG CTGCGGGTCATCGTGTGGAACACAGACGAGGTGG TCCTGGAAGACGATGATTTCTTCACGGGAGAGAA GTCCAGTGACATTTTTGTGAGGGGGTGGCTGAAG GGCCAGCAGGAGGACAAACAGGACACAGATGTC CACTATCACTCCCTCACGGGGGAGGGCAACTTCA ACTGGAGATACCTCTTCCCCTTCGACTACCTAGCG GCCGAAGAGAAGATCGTTATGTCCAAAAAGGAGT

[0186]

SEQ ID NO.	序列名称	序列
		CTATGTTCTCCTGGGATGAGACGGAGTACAAGAT CCCTGCGCGGCTCACCCTGCAGATCTGGGACGCT GACCACTTCTCGGCTGACGACTTCCTGGGGGCTAT CGAGCTGGACCTGAACCGGTTCCCGAGGGGCGCT AAGACAGCCAAGCAGTGCACCATGGAGATGGCC ACCGGGGAGGTGGACGTACCCCTGGTTTCCATCT TTAAACAGAAACGTGTCAAAGGCTGGTGGCCCCT CCTGGCCCGCAATGAGAATGATGAGTTTGAGCTC ACAGGCAAAGTGGAGGCGGAGCTACACCTACTCA CGGCAGAGGAGGCAGAGAAGAACCCTGTGGGCC TGGCTCGCAATGAACCTGATCCCCTAGAAAAACC CAATCGGCCGGACACAAGCTTCATCTGGTTCTTG AACCCTCTCAAGTCTGCCCGCTACTTCCTGTGGCA TACCTACCGCTGGCTACTCCTCAAATTCCTGCTGC TCTTCCTCCTGCTGCTGCTCTTCGCCCTGTTTCTCT ACTCTCTGCCTGGCTACCTGGCCAAGAAGATCCTT GGGGCCTGAGCCCTGCAGTCGCCTAGGCCTGCCG GCCTGACACGGCATTCGTCTGGTTCCTGAACCCAC TCAAATCTATCAAGTACCTCATCTGCACCCGGTAC AAGTGGCTGATCATCAAGATCGTGCTGGCGCTGC TGGGGCTGCTCATGCTGGCCCTCTTCCTTTACAGC CTCCCAGGCTACATGGTCAAGAAGCTCCTAGGGG CCTGAAGTGTGCCCCACCCCAGCCCGCTCCAGCA TCCCTCCAGGGGCTGCTGCGTATTTTGCCTTCCCT CACCTGGACTCTCTCCCAACTCCCTGAGGAGCCCT CCCACGCCTGCCAGCCTTGAGCAAGACACCTGCT TGCTGGACTTCATCCCCACCCCACACCCAAACTGT TGCTTGCCTGATCTTGTCCCAGGCCTGCCTGGGGT TTGGGGCACAGTTGGCCTCCAAAACCAGATACCC TCTTGTCTAAAGTACCAGGTTCCTCTGCCCAACCC CAAGAGTGGTAGTGGCCCAACCCTCCCTGTGCTTT CCAAATCTTGTCTTAAGGCACCAGTGAAATTAAC CAAGAAACGCGGAGCAATGCCCAAGGCTCTGATG AGTAGGAACACGTGGAAAGCACCAGGAATGCCA GCAGAGGCGAGGCGGCACACCTCTCTGCAGAGCA TCCAGGCCGAGCGGCGGGCAGCGGCCAGCTGCTT CTGCGCATGCTCTCCTCTTGGCTCTGCTTCTTTCTC ACAGTCACAGTCACTTCACAGCTTAGCCTTGGGCT TCCCATCACTTCCAGGGGTGCCTCTGCCTTGGCCA GTGTGTGTCAGCTAGTACACAAGCTCCAAGTGTG AATCAGGTGTACTGGCCGTCCTGAAGACTGACTG CCCTGTCCTTCCTGCCGACAGCCACACCCGAGTGT ACACTTAAAGCGGTGCCCTTCTGCCTCTGTGGGCC TGCTGGCTGCTGTTCCTTTCTTGAGTGTGATTTTTT TTTTCTCTCCCTCAATAAAATAAATCAAACTCTGA GAC
17	mOTOF-202_1 转录物	TTGGTTGCCTTGGTCTCTGTGGGCAGCAGCAGGA GGAGGCGGCAGCAGCCAGAGAAGAGGGAGGCGT

[0187]

SEQ ID NO.	序列名称	序列
	(NM_001100395.1)，小鼠耳畸蛋白转录物变体1，6907 bp，编码 SEQ ID NO: 8 的蛋白质	GTGAGCCACACTCCACCAGCGAGCTTCTTCCCGCT GCTCTGGAACTGCCCAGGCTCTCCCCACCAGCAT GGCCCTGATTGTTCACCTCAAGACTGTCTCAGAGC TCCGAGGCAAAGGTGACCGGATTGCCAAAGTCAC TTTCCGAGGGCAGTCTTTCTACTCCCGGGTCCTGG AGAACTGCGAGGGTGTGGCTGACTTTGATGAGAC GTTCCGGTGGCCAGTGGCCAGCAGCATCGACCGG AATGAAGTGTTGGAGATTCAGATTTTCAACTACA GCAAAGTCTTCAGCAACAAGCTGATAGGGACCTT CTGCATGGTGCTGCAGAAAGTGGTGGAGGAGAAT CGGGTAGAGGTGACCGACACGCTGATGGATGACA GCAATGCTATCATCAAGACCAGCCTGAGCATGGA GGTCCGGTATCAGGCCACAGATGGCACTGTGGGC CCCTGGGATGATGGAGACTTCCTGGGAGATGAAT CCCTCCAGGAGGAGAAGGACAGCCAGGAGACAG ATGGGCTGCTACCTGGTTCCCGACCCAGCACCCG GATATCTGGCGAGAAGAGCTTTCGCAGCAAAGGC AGAGAGAAGACCAAGGGAGGCAGAGATGGCGAG CACAAAGCGGGAAGGAGTGTGTTCTCGGCCATGA AACTCGGCAAAACTCGGTCCCACAAAGAGGAGCC CCAAAGACAAGATGAGCCAGCAGTGCTGGAGAT GGAGGACCTGGACCACCTAGCCATTCAGCTGGGG GATGGGCTGGATCCTGACTCCGTGTCTCTAGCCTC GGTCACCGCTCTCACCAGCAATGTCTCCAACAAA CGGTCTAAGCCAGATATTAAGATGGAGCCCAGTG CTGGAAGGCCCATGGATTACCAGGTCAGCATCAC AGTGATTGAGGCTCGGCAGCTGGTGGGCTTGAAC ATGGACCCTGTGGTGTGTGTGGAGGTGGGTGATG ACAAGAAATACACGTCAATGAAGGAGTCCACAA ACTGCCCTTACTACAACGAGTACTTTGTCTTCGAC TTCCATGTCTCTCCTGATGTCATGTTTGACAAGAT CATCAAGATCTCGGTTATCCATTCTAAGAACCTGC TTCGGAGCGGCACCCTGGTGGGTTCCTTCAAAAT GGATGTGGGGACTGTGTATTCCCAGCCTGAACAC CAGTTCCATCACAAATGGGCCATCCTGTCAGACC CCGATGACATCTCTGCTGGGTTGAAGGGTTATGT AAAGTGTGATGTCGCTGTGGTGGGCAAGGGAGAC AACATCAAGACACCCCACAAGGCCAACGAGACG GATGAGGACGACATTGAAGGGAACTTGCTGCTCC CCGAGGGCGTGCCCCCCGAACGGCAGTGGGCACG GTTCTATGTGAAAATTTACCGAGCAGAGGGACTG CCCCGGATGAACACAAGCCTCATGGCCAACGTGA AGAAGGCGTTCATCGGTGAGAACAAGGACCTCGT CGACCCCTATGTGCAAGTCTTCTTTGCTGGACAAA AGGGCAAAACATCAGTGCAGAAGAGCAGCTATG AGCCGCTATGGAATGAGCAGGTCGTCTTCACAGA CTTGTTCCCCCCACTCTGCAAACGCATGAAGGTGC AGATCCGGGACTCTGACAAGGTCAATGATGTGGC

[0188]

SEQ ID NO.	序列名称	序列
		CATCGGCACCCACTTCATCGACCTGCGCAAGATTT CCAACGATGGAGACAAAGGCTTCCTGCCTACCCT CGGTCCAGCCTGGGTGAACATGTACGGCTCCACG CGCAACTACACACTGCTGGACGAGCACCAGGACT TGAATGAAGGCCTGGGGGAGGGTGTGTCCTTCCG GGCCCGCCTCATGTTGGGACTAGCTGTGGAGATC CTGGACACCTCCAACCCAGAGCTCACCAGCTCCA CGGAGGTGCAGGTGGAGCAGGCCACGCCTGTCTC GGAGAGCTGCACAGGGAGAATGGAAGAATTTTTT CTATTTGGAGCCTTCTTGGAAGCCTCAATGATTGA CCGGAAAAATGGGGACAAGCCAATTACCTTTGAG GTGACCATAGGAAACTACGGCAATGAAGTCGATG GTATGTCCCGGCCCCTGAGGCCTCGGCCCCGGAA AGAGCCTGGGGATGAAGAAGAGGTAGACCTGATT CAGAACTCCAGTGACGATGAAGGTGACGAAGCCG GGGACCTGGCCTCGGTGTCCTCCACCCCACCTATG CGGCCCCAGATCACGGACAGGAACTATTTCCACC TGCCCTACCTGGAGCGCAAGCCCTGCATCTATATC AAGAGCTGGTGGCCTGACCAGAGGCGGCGCCTCT ACAATGCCAACATCATGGATCACATTGCTGACAA GCTGGAAGAAGGCCTGAATGATGTACAGGAGATG ATCAAAACGGAGAAGTCCTACCCGGAGCGCCGCC TGCGGGGTGTGCTAGAGGAACTCAGCTGTGGCTG CCACCGCTTCCTCTCCCTCTCGGACAAGGACCAG GGCCGCTCGTCCCGCACCAGGCTGGATCGAGAGC GTCTTAAGTCCTGTATGAGGGAGTTGGAGAGCAT GGGACAGCAGGCCAAGAGCCTGAGGGCTCAGGT GAAGCGGCACACTGTTCGGGACAAGCTGAGGTCA TGCCAGAACTTTCTGCAGAAGCTACGCTTCCTGGC GGATGAGCCCCAGCACAGCATTCCTGATGTGTTC ATTTGGATGATGAGCAACAACAAACGTATCGCCT ATGCCCGCGTGCCTTCCAAAGACCTGCTCTTCTCC ATCGTGGAGGAGGAACTGGGCAAGGACTGCGCC AAAGTCAAGACCCTCTTCCTGAAGCTGCCAGGGA AGAGGGGCTTCGGCTCGGCAGGCTGGACAGTACA GGCCAAGCTGGAGCTCTACCTGTGGCTGGGCCTC AGCAAGCAGCGAAAGGACTTCCTGTGTGGTCTGC CCTGTGGCTTCGAGGAGGTCAAGGCAGCCCAAGG CCTGGGCCTGCATTCCTTTCCGCCCATCAGCCTAG TCTACACCAAGAAGCAAGCCTTCCAGCTCCGAGC ACACATGTATCAGGCCCGAAGCCTCTTTGCTGCTG ACAGCAGTGGGCTCTCTGATCCCTTTGCCCGTGTC TTCTTCATCAACCAGAGCCAATGCACTGAGGTTCT AAACGAGACACTGTGTCCCACCTGGGACCAGATG CTGGTATTTGACAACCTGGAGCTGTACGGTGAAG CTCACGAGTTACGAGATGATCCCCCCATCATTGTC ATTGAAATCTACGACCAGGACAGCATGGGCAAAG CCGACTTCATGGGCCGGACCTTCGCCAAGCCCCT

[0189]

SEQ ID NO.	序列名称	序列
		GGTGAAGATGGCAGATGAAGCATACTGCCCACCT CGCTTCCCGCCGCAGCTTGAGTACTACCAGATCTA CCGAGGCAGTGCCACTGCCGGAGACCTACTGGCT GCCTTCGAGCTGCTGCAGATTGGGCCATCAGGGA AGGCTGACCTGCCACCCATCAATGGCCCAGTGGA CATGGACAGAGGGCCCATCATGCCTGTGCCCGTG GGAATCCGGCCAGTGCTCAGCAAGTACCGAGTGG AGGTGCTGTTCTGGGGCCTGAGGGACCTAAAGAG GGTGAACCTGGCCCAGGTGGACCGACCACGGGTG GACATCGAGTGTGCAGGAAAGGGGGTACAATCCT CCCTGATTCACAATTATAAGAAGAACCCCAACTT CAACACGCTGGTCAAGTGGTTTGAAGTGGACCTC CCGGAGAATGAGCTCCTGCACCCACCCTTGAACA TCCGAGTGGTAGATTGCCGGGCCTTTGGACGATA CACCCTGGTGGGTTCCCACGCAGTCAGCTCACTG AGGCGCTTCATCTACCGACCTCCAGACCGCTCAG CCCCCAACTGGAACACCACAGGGGAGGTTGTAGT AAGCATGGAGCCTGAGGAGCCAGTTAAGAAGCTG GAGACCATGGTGAAACTGGATGCGACTTCTGATG CTGTGGTCAAGGTGGATGTGGCTGAAGATGAGAA GGAAAGGAAGAAGAAGAAAAAGAAAGGCCCGTC AGAGGAGCCAGAGGAGGAAGAGCCCGATGAGAG CATGCTGGATTGGTGGTCCAAGTACTTCGCCTCCA TCGACACAATGAAGGAGCAACTTCGACAACATGA GACCTCTGGAACTGACTTGGAAGAGAAGGAAGA GATGGAAAGCGCTGAGGGCCTGAAGGGACCAAT GAAGAGCAAGGAGAAGTCCAGAGCTGCAAAGGA GGAGAAAAAGAAGAAAAACCAGAGCCCTGGCCC TGGCCAGGGATCGGAGGCTCCTGAGAAGAAGAA AGCCAAGATCGATGAGCTTAAGGTGTACCCCAAG GAGCTGGAATCGGAGTTTGACAGCTTTGAGGACT GGCTGCACACCTTCAACCTGTTGAGGGGCAAGAC GGGAGATGATGAGGATGGCTCCACAGAGGAGGA GCGCATAGTAGGCCGATTCAAGGGCTCCCTCTGT GTGTACAAAGTGCCACTCCCAGAAGATGTATCTC GAGAAGCTGGCTATGATCCCACCTATGGAATGTT CCAGGGCATCCCAAGCAATGACCCCATCAATGTG CTGGTCCGAATCTATGTGGTCCGGGCCACAGACC TGCACCCGGCCGACATCAATGGCAAAGCTGACCC CTATATTGCCATCAAGTTAGGCAAGACCGACATC CGAGACAAGGAGAACTACATCTCCAAGCAGCTCA ACCCTGTGTTTGGGAAGTCCTTTGACATTGAGGCC TCCTTCCCCATGGAGTCCATGTTGACAGTGGCCGT GTACGACTGGGATCTGGTGGGCACTGATGACCTC ATCGGAGAAACCAAGATTGACCTGGAAAACCGCT TCTACAGCAAGCATCGCGCCACCTGCGGCATCGC ACAGACCTATTCCATACATGGCTACAATATCTGG AGGGACCCCATGAAGCCCAGCCAGATCCTGACAC

[0190]

SEQ ID NO.	序列名称	序列
		GCCTCTGTAAAGAGGGCAAAGTGGACGGCCCCCA CTTTGGTCCCCATGGGAGAGTGAGGGTTGCCAAC CGTGTCTTCACGGGGCCTTCAGAAATAGAGGATG AGAATGGTCAGAGGAAGCCCACAGATGAGCACG TGGCACTGTCTGCTCTGAGACACTGGGAGGACAT CCCCCGGGTGGGCTGCCGCCTTGTGCCGGAACAC GTGGAGACCAGGCCGCTGCTCAACCCTGACAAGC CAGGCATTGAGCAGGGCCGCCTGGAGCTGTGGGT GGACATGTTCCCCATGGACATGCCAGCCCCTGGG ACACCTCTGGATATATCCCCCAGGAAACCCAAGA AGTACGAGCTGCGGGTCATCGTGTGGAACACAGA CGAGGTGGTCCTGGAAGACGATGATTTCTTCACG GGAGAGAAGTCCAGTGACATTTTTGTGAGGGGGT GGCTGAAGGGCCAGCAGGAGGACAAACAGGACA CAGATGTCCACTATCACTCCCTCACGGGGGAGGG CAACTTCAACTGGAGATACCTCTTCCCCTTCGACT ACCTAGCGGCCGAAGAGAAGATCGTTATGTCCAA AAAGGAGTCTATGTTCTCCTGGGATGAGACGGAG TACAAGATCCCTGCGCGGCTCACCCTGCAGATCT GGGACGCTGACCACTTCTCGGCTGACGACTTCCT GGGGGCTATCGAGCTGGACCTGAACCGGTTCCCG AGGGGCGCTAAGACAGCCAAGCAGTGCACCATG GAGATGGCCACCGGGGAGGTGGACGTACCCCTGG TTTCCATCTTTAAACAGAAACGTGTCAAAGGCTG GTGGCCCCTCCTGGCCCGCAATGAGAATGATGAG TTTGAGCTCACAGGCAAAGTGGAGGCGGAGCTAC ACCTACTCACGGCAGAGGAGGCAGAGAAGAACC CTGTGGGCCTGGCTCGCAATGAACCTGATCCCCT AGAAAAACCCAACCGGCCTGACACGGCATTCGTC TGGTTCCTGAACCCACTCAAATCTATCAAGTACCT CATCTGCACCCGGTACAAGTGGCTGATCATCAAG ATCGTGCTGGCGCTGCTGGGGCTGCTCATGCTGG CCCTCTTCCTTTACAGCCTCCCAGGCTACATGGTC AAGAAGCTCCTAGGGGCCTGAAGTGTGCCCCACC CCAGCCCGCTCCAGCATCCCTCCAGGGGCTGCTG CGTATTTTGCCTTCCCTCACCTGGACTCTCTCCCA ACTCCCTGAGGAGCCCTCCCACGCCTGCCAGCCTT GAGCAAGACACCTGCTTGCTGGACTTCATCCCCA CCCCACACCCAAACTGTTGCTTGCCTGATCTTGTC CCAGGCCTGCCTGGGGTTTGGGGCACAGTTGGCC TCCAAAACCAGATACCCTCTTGTCTAAAGTACCA GGTTCCTCTGCCCAACCCCAAGAGTGGTAGTGGC CCAACCCTCCCTGTGCTTTCCAAATCTTGTCTTAA GGCACCAGTGAAATTAACCAAGAAACGCGGAGC AATGCCCAAGGCTCTGATGAGTAGGAACACGTGG AAAGCACCAGGAATGCCAGCAGAGGCGAGGCGG CACACCTCTCTGCAGAGCATCCAGGCCGAGCGGC GGGCAGCGGCCAGCTGCTTCTGCGCATGCTCTCCT

[0191]

SEQ ID NO.	序列名称	序列
		CTTGGCTCTGCTTCTTTCTCACAGTCACAGTCACT TCACAGCTTAGCCTTGGGCTTCCCATCACTTCCAG GGGTGCCTCTGCCTTGGCCAGTGTGTGTCAGCTAG TACACAAGCT CCAAGTGTGAATCAGGTGTACTGGCCGTCCTGAA GACTGACTGCCCTGTCCTTCCTGCCGACAGCCACA CCCGAGTGTACACTTAAAGCGGTGCCCTTCTGCCT CTGTGGGCCTGCTGGCTGCTGTTCCTTTCTTGAGT GTGATTTTTTTTTTCTCTCCCTCAATAAAATAAAT CAAACTCTGAGAC
18	mOTOF-202_2 转录物 (NM_001313767.1)，小鼠耳畸蛋白转录物变体 4，6862 bp，编码 SEQ ID NO: 9 的蛋白质	TTGGTTGCCTTGGTCTCTGTGGGCAGCAGCAGGA GGAGGCGGCAGCAGCCAGAGAAGAGGGAGGCGT GTGAGCCACACTCCACCAGCGAGCTTCTTCCCGCT GCTCTGGAACTGCCCAGGCTCTCCCCACCAGCAT GGCCCTGATTGTTCACCTCAAGACTGTCTCAGAGC TCCGAGGCAAAGGTGACCGGATTGCCAAAGTCAC TTTCCGAGGGCAGTCTTTCTACTCCCGGGTCCTGG AGAACTGCGAGGGTGTGGCTGACTTTGATGAGAC GTTCCGGTGGCCAGTGGCCAGCAGCATCGACCGG AATGAAGTGTTGGAGATTCAGATTTTCAACTACA GCAAAGTCTTCAGCAACAAGCTGATAGGGACCTT CTGCATGGTGCTGCAGAAAGTGGTGGAGGAGAAT CGGGTAGAGGTGACCGACACGCTGATGGATGACA GCAATGCTATCATCAAGACCAGCCTGAGCATGGA GGTCCGGTATCAGGCCACAGATGGCACTGTGGGC CCCTGGGATGATGGAGACTTCCTGGGAGATGAAT CCCTCCAGGAGGAGAAGGACAGCCAGGAGACAG ATGGGCTGCTACCTGGTTCCCGACCCAGCACCCG GATATCTGGCGAGAAGAGCTTTCGCAGAGCGGGA AGGAGTGTGTTCTCGGCCATGAAACTCGGCAAAA CTCGGTCCCACAAAGAGGAGCCCCAAAGACAAG ATGAGCCAGCAGTGCTGGAGATGGAGGACCTGGA CCACCTAGCCATTCAGCTGGGGGATGGGCTGGAT CCTGACTCCGTGTCTCTAGCCTCGGTCACCGCTCT CACCAGCAATGTCTCCAACAAACGGTCTAAGCCA GATATTAAGATGGAGCCCAGTGCTGGAAGGCCCA TGGATTACCAGGTCAGCATCACAGTGATTGAGGC TCGGCAGCTGGTGGGCTTGAACATGGACCCTGTG GTGTGTGTGGAGGTGGGTGATGACAAGAAATACA CGTCAATGAAGGAGTCCACAAACTGCCCTTACTA CAACGAGTACTTTGTCTTCGACTTCCATGTCTCTC CTGATGTCATGTTTGACAAGATCATCAAGATCTCG GTTATCCATTCTAAGAACCTGCTTCGGAGCGGCA CCCTGGTGGGTTCCTTCAAAATGGATGTGGGGAC TGTGTATTCCCAGCCTGAACACCAGTTCCATCACA AATGGGCCATCCTGTCAGACCCCGATGACATCTC TGCTGGGTTGAAGGGTTATGTAAAGTGTGATGTC GCTGTGGTGGGCAAGGGAGACAACATCAAGACA

[0192]

SEQ ID NO.	序列名称	序列
		CCCCACAAGGCCAACGAGACGGATGAGGACGAC ATTGAAGGGAACTTGCTGCTCCCCGAGGGCGTGC CCCCCGAACGGCAGTGGGCACGGTTCTATGTGAA AATTTACCGAGCAGAGGGACTGCCCCGGATGAAC ACAAGCCTCATGGCCAACGTGAAGAAGGCGTTCA TCGGTGAGAACAAGGACCTCGTCGACCCCTATGT GCAAGTCTTCTTTGCTGGACAAAAGGGCAAAACA TCAGTGCAGAAGAGCAGCTATGAGCCGCTATGGA ATGAGCAGGTCGTCTTCACAGACTTGTTCCCCCCA CTCTGCAAACGCATGAAGGTGCAGATCCGGGACT CTGACAAGGTCAATGATGTGGCCATCGGCACCCA CTTCATCGACCTGCGCAAGATTTCCAACGATGGA GACAAAGGCTTCCTGCCTACCCTCGGTCCAGCCT GGGTGAACATGTACGGCTCCACGCGCAACTACAC ACTGCTGGACGAGCACCAGGACTTGAATGAAGGC CTGGGGGAGGGTGTGTCCTTCCGGGCCCGCCTCA TGTTGGGACTAGCTGTGGAGATCCTGGACACCTC CAACCCAGAGCTCACCAGCTCCACGGAGGTGCAG GTGGAGCAGGCCACGCCTGTCTCGGAGAGCTGCA CAGGGAGAATGGAAGAATTTTTTCTATTTGGAGC CTTCTTGGAAGCCTCAATGATTGACCGGAAAAAT GGGGACAAGCCAATTACCTTTGAGGTGACCATAG GAAACTACGGCAATGAAGTCGATGGTATGTCCCG GCCCCTGAGGCCTCGGCCCCGGAAAGAGCCTGGG GATGAAGAAGAGGTAGACCTGATTCAGAACTCCA GTGACGATGAAGGTGACGAAGCCGGGGACCTGG CCTCGGTGTCCTCCACCCCACCTATGCGGCCCCAG ATCACGGACAGGAACTATTTCCACCTGCCCTACCT GGAGCGCAAGCCCTGCATCTATATCAAGAGCTGG TGGCCTGACCAGAGGCGGCGCCTCTACAATGCCA ACATCATGGATCACATTGCTGACAAGCTGGAAGA AGGCCTGAATGATGTACAGGAGATGATCAAAACG GAGAAGTCCTACCCGGAGCGCCGCCTGCGGGGTG TGCTAGAGGAACTCAGCTGTGGCTGCCACCGCTT CCTCTCCCTCTCGGACAAGGACCAGGGCCGCTCG TCCCGCACCAGGCTGGATCGAGAGCGTCTTAAGT CCTGTATGAGGGAGTTGGAGAGCATGGGACAGCA GGCCAAGAGCCTGAGGGCTCAGGTGAAGCGGCA CACTGTTCGGGACAAGCTGAGGTCATGCCAGAAC TTTCTGCAGAAGCTACGCTTCCTGGCGGATGAGC CCCAGCACAGCATTCCTGATGTGTTCATTTGGATG ATGAGCAACAACAAACGTATCGCCTATGCCCGCG TGCCTTCCAAAGACCTGCTCTTCTCCATCGTGGAG GAGGAACTGGGCAAGGACTGCGCCAAAGTCAAG ACCCTCTTCCTGAAGCTGCCAGGGAAGAGGGGCT TCGGCTCGGCAGGCTGGACAGTACAGGCCAAGCT GGAGCTCTACCTGTGGCTGGGCCTCAGCAAGCAG CGAAAGGACTTCCTGTGTGGTCTGCCCTGTGGCTT

[0193]

SEQ ID NO.	序列名称	序列
		CGAGGAGGTCAAGGCAGCCCAAGGCCTGGGCCTG CATTCCTTTCCGCCCATCAGCCTAGTCTACACCAA GAAGCAAGCCTTCCAGCTCCGAGCACACATGTAT CAGGCCCGAAGCCTCTTTGCTGCTGACAGCAGTG GGCTCTCTGATCCCTTTGCCCGTGTCTTCTTCATC AACCAGAGCCAATGCACTGAGGTTCTAAACGAGA CACTGTGTCCCACCTGGGACCAGATGCTGGTATTT GACAACCTGGAGCTGTACGGTGAAGCTCACGAGT TACGAGATGATCCCCCCATCATTGTCATTGAAATC TACGACCAGGACAGCATGGGCAAAGCCGACTTCA TGGGCCGGACCTTCGCCAAGCCCCTGGTGAAGAT GGCAGATGAAGCATACTGCCCACCTCGCTTCCCG CCGCAGCTTGAGTACTACCAGATCTACCGAGGCA GTGCCACTGCCGGAGACCTACTGGCTGCCTTCGA GCTGCTGCAGATTGGGCCATCAGGGAAGGCTGAC CTGCCACCCATCAATGGCCCAGTGGACATGGACA GAGGGCCCATCATGCCTGTGCCCGTGGGAATCCG GCCAGTGCTCAGCAAGTACCGAGTGGAGGTGCTG TTCTGGGGCCTGAGGGACCTAAAGAGGGTGAACC TGGCCCAGGTGGACCGACCACGGGTGGACATCGA GTGTGCAGGAAAGGGGGTACAATCCTCCCTGATT CACAATTATAAGAAGAACCCCAACTTCAACACGC TGGTCAAGTGGTTTGAAGTGGACCTCCCGGAGAA TGAGCTCCTGCACCCACCCTTGAACATCCGAGTG GTAGATTGCCGGGCCTTTGGACGATACACCCTGG TGGGTTCCCACGCAGTCAGCTCACTGAGGCGCTT CATCTACCGACCTCCAGACCGCTCAGCCCCCAAC TGGAACACCACAGGGGAGGTTGTAGTAAGCATGG AGCCTGAGGAGCCAGTTAAGAAGCTGGAGACCAT GGTGAAACTGGATGCGACTTCTGATGCTGTGGTC AAGGTGGATGTGGCTGAAGATGAGAAGGAAAGG AAGAAGAAGAAAAAGAAAGGCCCGTCAGAGGAG CCAGAGGAGGAAGAGCCCGATGAGAGCATGCTG GATTGGTGGTCCAAGTACTTCGCCTCCATCGACAC AATGAAGGAGCAACTTCGACAACATGAGACCTCT GGAACTGACTTGGAAGAGAAGGAAGAGATGGAA AGCGCTGAGGGCCTGAAGGGACCAATGAAGAGC AAGGAGAAGTCCAGAGCTGCAAAGGAGGAGAAA AAGAAGAAAAACCAGAGCCCTGGCCCTGGCCAG GGATCGGAGGCTCCTGAGAAGAAGAAAGCCAAG ATCGATGAGCTTAAGGTGTACCCCAAGGAGCTGG AATCGGAGTTTGACAGCTTTGAGGACTGGCTGCA CACCTTCAACCTGTTGAGGGGCAAGACGGGAGAT GATGAGGATGGCTCCACAGAGGAGGAGCGCATA GTAGGCCGATTCAAGGGCTCCCTCTGTGTGTACA AAGTGCCACTCCCAGAAGATGTATCTCGAGAAGC TGGCTATGATCCCACCTATGGAATGTTCCAGGGC ATCCCAAGCAATGACCCCATCAATGTGCTGGTCC

[0194]

SEQ ID NO.	序列名称	序列
		GAATCTATGTGGTCCGGGCCACAGACCTGCACCC GGCCGACATCAATGGCAAAGCTGACCCCTATATT GCCATCAAGTTAGGCAAGACCGACATCCGAGACA AGGAGAACTACATCTCCAAGCAGCTCAACCCTGT GTTTGGGAAGTCCTTTGACATTGAGGCCTCCTTCC CCATGGAGTCCATGTTGACAGTGGCCGTGTACGA CTGGGATCTGGTGGGCACTGATGACCTCATCGGA GAAACCAAGATTGACCTGGAAAACCGCTTCTACA GCAAGCATCGCGCCACCTGCGGCATCGCACAGAC CTATTCCATACATGGCTACAATATCTGGAGGGAC CCCATGAAGCCCAGCCAGATCCTGACACGCCTCT GTAAAGAGGGCAAAGTGGACGGCCCCCACTTTGG TCCCCATGGGAGAGTGAGGGTTGCCAACCGTGTC TTCACGGGGCCTTCAGAAATAGAGGATGAGAATG GTCAGAGGAAGCCCACAGATGAGCACGTGGCACT GTCTGCTCTGAGACACTGGGAGGACATCCCCCGG GTGGGCTGCCGCCTTGTGCCGGAACACGTGGAGA CCAGGCCGCTGCTCAACCCTGACAAGCCAGGCAT TGAGCAGGGCCGCCTGGAGCTGTGGGTGGACATG TTCCCCATGGACATGCCAGCCCCTGGGACACCTCT GGATATATCCCCCAGGAAACCCAAGAAGTACGAG CTGCGGGTCATCGTGTGGAACACAGACGAGGTGG TCCTGGAAGACGATGATTTCTTCACGGGAGAGAA GTCCAGTGACATTTTTGTGAGGGGGTGGCTGAAG GGCCAGCAGGAGGACAAACAGGACACAGATGTC CACTATCACTCCCTCACGGGGGAGGGCAACTTCA ACTGGAGATACCTCTTCCCCTTCGACTACCTAGCG GCCGAAGAGAAGATCGTTATGTCCAAAAAGGAGT CTATGTTCTCCTGGGATGAGACGGAGTACAAGAT CCCTGCGCGGCTCACCCTGCAGATCTGGGACGCT GACCACTTCTCGGCTGACGACTTCCTGGGGGCTAT CGAGCTGGACCTGAACCGGTTCCCGAGGGGCGCT AAGACAGCCAAGCAGTGCACCATGGAGATGGCC ACCGGGGAGGTGGACGTACCCCTGGTTTCCATCT TTAAACAGAAACGTGTCAAAGGCTGGTGGCCCCT CCTGGCCCGCAATGAGAATGATGAGTTTGAGCTC ACAGGCAAAGTGGAGGCGGAGCTACACCTACTCA CGGCAGAGGAGGCAGAGAAGAACCCTGTGGGCC TGGCTCGCAATGAACCTGATCCCCTAGAAAAACC CAACCGGCCTGACACGGCATTCGTCTGGTTCCTG AACCCACTCAAATCTATCAAGTACCTCATCTGCAC CCGGTACAAGTGGCTGATCATCAAGATCGTGCTG GCGCTGCTGGGGCTGCTCATGCTGGCCCTCTTCCT TTACAGCCTCCCAGGCTACATGGTCAAGAAGCTC CTAGGGGCCTGAAGTGTGCCCCACCCCAGCCCGC TCCAGCATCCCTCCAGGGGCTGCTGCGTATTTTGC CTTCCCTCACCTGGACTCTCTCCCAACTCCCTGAG GAGCCCTCCCACGCCTGCCAGCCTTGAGCAAGAC

[0195]

SEQ ID NO.	序列名称	序列
		ACCTGCTTGCTGGACTTCATCCCCACCCCACACCC AAACTGTTGCTTGCCTGATCTTGTCCCAGGCCTGC CTGGGGTTTGGGGCACAGTTGGCCTCCAAAACCA GATACCCTCTTGTCTAAAGTACCAGGTTCCTCTGC CCAACCCCAAGAGTGGTAGTGGCCCAACCCTCCC TGTGCTTTCCAAATCTTGTCTTAAGGCACCAGTGA AATTAACCAAGAAACGCGGAGCAATGCCCAAGG CTCTGATGAGTAGGAACACGTGGAAAGCACCAGG AATGCCAGCAGAGGCGAGGCGGCACACCTCTCTG CAGAGCATCCAGGCCGAGCGGCGGGCAGCGGCC AGCTGCTTCTGCGCATGCTCTCCTCTTGGCTCTGC TTCTTTCTCACAGTCACAGTCACTTCACAGCTTAG CCTTGGGCTTCCCATCACTTCCAGGGGTGCCTCTG CCTTGGCCAGTGTGTGTCAGCTAGTACACAAGCT CCAAGTGTGAATCAGGTGTACTGGCCGTCCTGAA GACTGACTGCCCTGTCCTTCCTGCCGACAGCCACA CCCGAGTGTACACTTAAAGCGGTGCCCTTCTGCCT CTGTGGGCCTGCTGGCTGCTGTTCCTTTCTTGAGT GTGATTTTTTTTTTCTCTCCCTCAATAAAATAAAT CAAACTCTGAGAC

[0196] OTOF在哺乳动物细胞中的表达

[0197] OTOF突变与感觉神经性听力损失和听神经病变相关。本文所述的组合物和方法通过施用含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体来增加WT OTOF同功型5蛋白的表达。为将核酸载体用于治疗感觉神经性听力损失和听神经病变的治疗应用,可将其引导至细胞内部,并且特别是引导至特定细胞类型。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。

[0198] 编码OTOF的多核苷酸

[0199] 可用于在哺乳动物细胞中实现OTOF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OTOF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。转染或转化细胞的适宜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。此类方法更详细描述于例如Green等人类,Molecular Cloning:A Laboratory Manual,第4版(Cold Spring Harbor University Press,New York 2014);和Ausubel等人类,Current Protocols in Molecular Biology(John Wiley&Sons,New York 2015),这些文献各自的公开内容通过引用并入本文。

[0200] 也可通过使含有编码OTOF蛋白的基因的部分的载体靶向细胞膜磷脂将OTOF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类员众所周知的方法产生。

[0201] 哺乳动物RNA聚合酶对编码OTOF蛋白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。

[0202] 适用于本文所述的组合物和方法中的多核苷酸也包括编码哺乳动物启动子下游的OTOF蛋白的多核苷酸(例如编码哺乳动物启动子下游的OTOF蛋白的N末端部分的多核苷酸)。可用于在哺乳动物细胞中表达OTOF蛋白的启动子包括普遍存在的启动子、耳蜗毛细胞特异性启动子和内毛细胞特异性启动子。普遍存在的启动子包括CAG启动子、巨细胞病毒(CMV)启动子(例如,CMV立即早期增强子和启动子、CMVmini启动子、minCMV启动子、CMV-TATA+INR启动子或min CMV-T6启动子)、鸡β-肌动蛋白启动子、smCBA启动子、CB7启动子、杂合CMV增强子/人类β-肌动蛋白启动子、CASI启动子、二氢叶酸还原酶(DHFR)启动子、人类β-肌动蛋白启动子、β-珠蛋白启动子(例如,最小β-珠蛋白启动子)、HSV启动子(例如,最小HSV ICP0启动子或截短HSV ICP0启动子)、SV40启动子(例如,SV40最小启动子)、EF1α启动子和PGK启动子。耳蜗毛细胞特异性启动子包括肌凝蛋白15(Myo15)启动子、肌凝蛋白7A(Myo7A)启动子、肌凝蛋白6(Myo6)启动子、POU4F3启动子、Atonal BHLH转录因子1(ATOH1)启动子、LIM同源盒3(LHX3)启动子、α9乙酰胆碱受体(α9AChR)启动子和α10乙酰胆碱受体(α10AChR)启动子。内毛细胞特异性启动子包括FGF8启动子、VGLUT3启动子、OTOF启动子和钙结合蛋白2(CABP2)启动子(描述于国际专利申请公布号WO2021/091940中,其通过引用并入本文)。或者,衍生自病毒基因组的启动子也可用于在哺乳动物细胞中稳定表达这些剂。可用于促进这些剂的哺乳动物表达的功能性病毒启动子的实例包括腺病毒晚期启动子、牛痘病毒7.5K启动子、SV40启动子、HSV的tk启动子、小鼠乳腺肿瘤病毒(MMTV)启动子、HIV的LTR启动子、莫罗尼病毒(moloney virus)的启动子、爱泼斯坦巴尔病毒(Epstein barr virus,EBV)启动子和劳斯肉瘤病毒(Rous sarcoma virus,RSV)启动子。

[0203] 鼠类肌凝蛋白15启动子

[0204] 在一些实施方案中,用于本文所述的组合物和方法中的Myo15启动子包括来自能够在毛细胞中特异性表达转基因的鼠类Myo15基因座区域的核酸序列,或其变体,例如与能够在毛细胞中特异性表达转基因的鼠类Myo15基因座区域具有至少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。这些区域包括紧接在鼠类Myo15翻译起始位点之前的核酸序列和位于距鼠类Myo15翻译起始位点超过5kb的上游调控元件。用于本文所述的组合物和方法中的Myo15启动子可任选地包括可操作地连接能够在毛细胞中特异性表达转基因的鼠类Myo15基因座区域的接头,或鼠类Myo15基因座的区域可无需间插接头直接接合。

[0205] 在一些实施方案中,用于本文所述的组合物和方法中的Myo15启动子含有第一区(上游调控元件),所述第一区与含有鼠类Myo15基因的第一非编码外显子的区域(相对于鼠类Myo15翻译起始位点-6755至-7209的核酸,其序列示于SEQ ID NO:24中)或其功能性部分或衍生物具有至少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),所述第一区接合(例如可操作地连接)至第二区,所述第二区与紧接在鼠类Myo15翻译起始位点之前的核酸序列(相对于鼠类Myo15翻译起始位点-1至-1157的核酸,其序列示于SEQ ID NO:25中)或其功能性部分或衍生物具有至少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)。SEQ ID NO:24的功能性部分可具有相对于鼠类Myo15翻译起始位点-7166至-7091的核酸的序列(示于SEQ ID NO:26中)和/或相对于鼠类Myo15翻译起始位点-7077至-6983的核酸的序列(示于SEQ ID NO:27中)。第一区可含有融合至SEQ ID NO:27的核酸序列的SEQ ID NO:26的核酸序列而无间插核酸,如SEQ ID NO:28中所示,或第一区可含有融合至SEQ ID NO:26的核酸序列的SEQ ID NO:27的核酸序列而无间插核酸,如SEQ ID NO:29中所示。或者,第一区可含有通过内源间插核酸序列(例如,第一区可具有或包括相对于鼠类Myo15翻译起始位点-7166至-6983的核酸的序列,如SEQ ID NO:30和SEQ ID NO:50中所示)或核酸接头接合的SEQ ID NO:26和SEQ ID NO:27的序列。在第一区含有SEQ ID NO:26和SEQ ID NO:27的鼠类Myo15启动子中,两个序列可以任何顺序包括(例如,SEQ ID NO:26可接合至(例如先于)SEQ ID NO:27,或SEQ ID NO:27可接合至(例如先于)SEQ ID NO:26)。SEQ ID NO:25的功能性部分可具有相对于鼠类Myo15翻译起始位点-590至-509的核酸的序列(示于SEQ ID NO:31中)和/或相对于鼠类Myo15翻译起始位点-266至-161的核酸的序列(示于SEQ ID NO:32中)。在一些实施方案中,含有SEQ ID NO:31的序列具有SEQ ID NO:51的序列。在一些实施方案中,含有SEQ ID NO:32的序列具有SEQ ID NO:52的序列。第二区可含有融合至SEQ ID NO:32的核酸序列的SEQ ID NO:31的核酸序列而无间插核酸,如SEQ ID NO:33中所示,或者第二区可含有融合至SEQ ID NO:31的核酸序列的SEQ ID NO:32的核酸序列而无间插核酸,如SEQ ID NO:34中所示。第二区可含有融合至SEQ ID NO:52的核酸序列的SEQ ID NO:51的核酸序列而无间插核酸,如SEQ ID NO:55中所示,或者第二区可含有融合至SEQ ID NO:51的核酸序列的SEQ ID NO:52的核酸序列而无间插核酸。或者,第二区可含有通过内源间插核酸序列(例如,第二区可具有相对于鼠类Myo15翻译起始位点-590至-161的核酸的序列,如SEQ ID NO:35中所示)或核酸接头接合的SEQ ID NO:31和SEQ ID NO:32的序列。在第二区含有SEQ ID NO:31和SEQ ID NO:32的鼠类Myo15启动子中,两个序列可以任何顺序包括(例如,SEQ ID NO:31可接合至(例如先于)SEQ ID NO:32,或SEQ ID NO:32可接合至(例如先于)SEQ ID NO:31)。

[0206] 鼠类Myo15启动子的第一区和第二区可直接接合或可通过核酸接头接合。例如,鼠类Myo15启动子可含有融合至SEQ ID NO:25的序列或其功能性部分或衍生物(例如SEQ ID NO:31-35、51、52和55中的任一者或多者,例如SEQ ID NO 31和32)的SEQ ID NO:24的序列或其功能性部分或衍生物(例如SEQ ID NO:26-30和50中的任一者或多者,例如SEQ ID NO 26和27)而无间插核酸。例如,鼠类Myo15启动子的源自SEQ ID NO:24与SEQ ID NO:25的直接融合的核酸序列示于SEQ ID NO:36中。或者,接头可用于将SEQ ID NO:24的序列或其功能性部分或衍生物(例如SEQ ID NO:26-30和50中的任一者或多者,例如SEQ ID NO 26和27)接合至SEQ ID NO:25的序列或其功能性部分或衍生物(例如SEQ ID NO:31-35、51、52和55中的任一者或多者,例如SEQ ID NO 31和32)。含有SEQ ID NO:24和SEQ ID NO:25的功能性部分的示例性Myo15启动子提供于SEQ ID NO:38、39、53、54、59和60中。

[0207] 用于本文所述鼠类Myo15启动子中的核酸接头的长度可为约5kb或更短(例如约5kb、4.5kb、4kb、3.5kb、3kb、2.5kb、2kb、1.5kb、1kb、900bp、800bp、700bp、600bp、500bp、450bp、400bp、350bp、300bp、250bp、200bp、150bp、100bp、90bp、80bp、70bp、60bp、50bp、40bp、30bp、25bp、20bp、15bp、10bp、5bp、4bp、3bp、2bp或更短)。可用于本文所述鼠类Myo15启动子中的核酸接头不会破坏本发明的鼠类Myo15启动子诱导毛细胞中的转基因表达的能

力。

[0208] 在一些实施方案中,SEQ ID NO:24的序列或其功能性部分或衍生物(例如SEQ ID NO:26-30和50中的任一者或多者,例如SEQ ID NO 26和27)接合(例如可操作地连接)至SEQ ID NO:25的序列或其功能性部分或衍生物(例如SEQ ID NO:31-35、51、52和55中的任一者或多者,例如SEQ ID NO 31和32),并且在一些实施方案中,区域的顺序是相反的(例如SEQ ID NO:25的序列或其功能性部分或衍生物(例如SEQ ID NO:31-35、51、52和55中的任一者或多者,例如SEQ ID NO 31和32)接合(例如可操作地连接)至SEQ ID NO:24的序列或其功能性部分或衍生物(例如SEQ ID NO:26-30和50中的任一者或多者,例如SEQ ID NO 26和27))。例如,鼠类Myo15启动子的源自SEQ ID NO:25与SEQ ID NO:24的直接融合的核酸序列示于SEQ ID NO:37中。SEQ ID NO:25的功能性部分或衍生物先于SEQ ID NO:24的功能性部分或衍生物的鼠类Myo15启动子的实例提供于SEQ ID NO:58中。无论顺序如何,SEQ ID NO:24的序列或其功能性部分或衍生物和SEQ ID NO:25的序列或其功能性部分或衍生物可通过直接融合或核酸接头接合,如上文所述。

[0209] 在一些实施方案中,用于本文所述的组合物和方法中的鼠类Myo15启动子含有与含有鼠类Myo15基因的第一非编码外显子的区域(相对于鼠类Myo15翻译起始位点-6755至-7209的核酸,其序列示于SEQ ID NO:24中)或其功能性部分或衍生物具有至少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)的区域。SEQ ID NO:24的功能性部分可具有相对于鼠类Myo15翻译起始位点-7166至-7091的核酸的序列(示于SEQ ID NO:26中)和/或相对于鼠类Myo15翻译起始位点-7077至-6983的核酸的序列(示于SEQ ID NO:27中)。鼠类Myo15启动子可含有融合至SEQ ID NO:27的核酸序列的SEQ ID NO:26的核酸序列而无间插核酸,如SEQ ID NO:28中所示,或鼠类Myo15启动子可含有融合至SEQ ID NO:26的核酸序列的SEQ ID NO:27的核酸序列而无间插核酸,如SEQ ID NO:29中所示。或者,鼠类Myo15启动子可含有通过内源间插核酸序列(例如,第一区可具有或包括相对于鼠类Myo15翻译起始位点-7166至-6983的核酸的序列,如SEQ ID NO:30和SEQ ID NO:50中所示)或核酸接头接合的SEQ ID NO:26和SEQ ID NO:27的序列。在含有SEQ ID NO:26和SEQ ID NO:27的鼠类Myo15启动子中,两个序列可以任何顺序包括(例如,SEQ ID NO:26可接合至(例如先于)SEQ ID NO:27,或SEQ ID NO:27可接合至(例如先于)SEQ ID NO:26)。

[0210] 在一些实施方案中,用于本文所述的组合物和方法中的鼠类Myo15启动子含有与紧接在鼠类Myo15翻译起始位点上游的多核苷酸序列(相对于鼠类Myo15翻译起始位点的-1至-1157的核酸,其序列示于SEQ ID NO:25中)或其功能性部分或衍生物具有至少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)的区域。SEQ ID NO:25的功能性部分可具有相对于鼠类Myo15翻译起始位点-590至-509的核酸的序列(示于SEQ ID NO:31中)和/或相对于鼠类Myo15翻译起始位点-266至-161的核酸的序列(示于SEQ ID NO:32中)。在一些实施方案中,含有SEQ ID NO:31的序列具有SEQ ID NO:51的序列。在一些实施方案中,含有SEQ ID NO:32的序列具有SEQ ID NO:52的序列。鼠类Myo15启动子可含有融合至SEQ ID NO:32的核酸序列的SEQ ID NO:31的核酸序列而无间插核酸,如SEQ ID NO:33中所示,或者鼠类Myo15启动子可含有融合至SEQ ID NO:31的核酸序列的SEQ ID NO:32的核酸序列而无间插核酸,如

SEQ IDNO:34中所示。鼠类Myo15启动子可含有融合至SEQ ID NO:52的核酸序列的SEQ ID NO:51的核酸序列而无间插核酸,如SEQ ID NO:55中所示,或者鼠类Myo15启动子可含有融合至SEQ ID NO:51的核酸序列的SEQ ID NO:52的核酸序列而无间插核酸。或者,鼠类Myo15启动子可含有通过内源间插核酸序列(例如,第二区可具有相对于鼠类Myo15翻译起始位点-590至-161的核酸的序列,如SEQ ID NO:35中所示)或核酸接头接合的SEQ ID NO:31和SEQ ID NO:32的序列。在含有SEQ ID NO:31和SEQ ID NO:32的鼠类Myo15启动子中,两个序列可以任何顺序包括(例如,SEQ ID NO:31可接合至(例如先于)SEQ ID NO:32,或SEQ ID NO:32可接合至(例如先于)SEQ ID NO:31)。

[0211] 在一些实施方案中,用于本文所述的组合物和方法中的鼠类Myo15启动子含有与含有Myo15基因的第一非编码外显子的区域(相对于鼠类Myo15翻译起始位点的-6755至-7209的核酸,其序列示于SEQ ID NO:24中)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)的区域的功能性部分或衍生物,所述含有Myo15基因的第一非编码外显子的区域在两侧侧接有与紧接在鼠类Myo15翻译起始位点上游的核酸序列(相对于鼠类Myo15翻译起始位点的-1至-1157的核酸,其序列示于SEQ ID NO:25中)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)的区域的功能性部分或衍生物。例如,SEQ ID NO:25的功能性部分或衍生物(例如SEQ ID NO:31或51)可直接融合或通过核酸接头接合至SEQ ID NO:24的部分,例如SEQ ID NO:26-30和50中的任一者,其直接融合或通过核酸接头接合至SEQ ID NO:25的不同功能性部分,例如SEQ ID NO:32或52。在其他实施方案中,SEQ ID NO:25的功能性部分或衍生物,例如SEQ ID NO:32或52可直接融合或通过核酸接头接合至SEQ ID NO:24的部分,例如SEQ ID NO:26-30和50中的任一者,其直接融合或通过核酸接头接合至SEQ ID NO:25的不同功能性部分,例如SEQ ID NO:31或51。例如,与SEQ ID NO:51、50和52的核酸序列具有至少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)的多核苷酸可融合产生与SEQ ID NO:56的核酸序列具有至少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,与SEQ ID NO:52、50和51的核酸序列具有至少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)的多核苷酸可融合产生与SEQ ID NO:57的核酸序列具有至少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。

[0212] 人类肌凝蛋白15启动子

[0213] 在一些实施方案中,用于本文所述的组合物和方法的Myo15启动子包括来自能够在毛细胞中特异性表达转基因的人类Myo15基因座区域或其变体的核酸序列,例如与能够在毛细胞中特异性表达转基因的人类Myo15基因座区域具有至少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。用于本文所述的组合物和方法的Myo15启动子可任选地包括可操作地连接能够在毛细胞中特异性表达转基因的人类Myo15基因座区域的接头,或者

人类Myo15基因座的区域可无需间插接头直接接合。

[0214] 在一些实施方案中,用于本文所述的组合物和方法中的Myo15启动子含有第一区,所述第一区与SEQ ID NO:40中所示的序列或其功能性部分或衍生物具有至少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),所述第一区接合(例如可操作地连接)至第二区,所述第二区与SEQ ID NO:41中所示的序列或其功能性部分或衍生物具有至少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)。SEQ ID NO:40的功能性部分可具有SEQ ID NO:42中所示的序列。SEQ ID NO:41的功能性部分可具有SEQ ID NO:43中所示的序列和/或SEQ ID NO:44中所示的序列。第二区可含有融合至SEQ ID NO:44的核酸序列而无间插核酸的SEQ ID NO:43的核酸序列,如SEQ ID NO:45中所示,或者第二区可含有融合至SEQ ID NO:43的核酸序列而无间插核酸的SEQ ID NO:44的核酸序列,如SEQ ID NO:46中所示。或者,第二区可含有通过内源间插核酸序列(如SEQ ID NO:47中所示)或核酸接头接合的SEQ ID NO:43和SEQ ID NO:44的序列。在第二区含有SEQ ID NO:43和SEQ ID NO:44的人类Myo15启动子中,两个序列可以任何顺序包括(例如,SEQ ID NO:43可接合至(例如先于)SEQ ID NO:44,或SEQ ID NO:44可接合至(例如先于)SEQ ID NO:43)。

[0215] 人类Myo15启动子的第一区和第二区可直接接合或可通过核酸接头接合。例如,人类Myo15启动子可含有融合至SEQ ID NO:41的序列或其功能性部分或衍生物(例如SEQ ID NO:43-47中的任一者或多者,例如SEQ ID NO:43和/或44)而无间插核酸的SEQ ID NO:40的序列或其功能性部分或衍生物(例如SEQ ID NO:42)。或者,可使用接头将SEQ ID NO:40的序列或其功能性部分或衍生物(例如SEQ ID NO:42)接合至SEQ ID NO:41的序列或其功能性部分或衍生物(例如SEQ ID NO:43-47中的任一者或多者,例如SEQ ID NO:43和/或44)。含有SEQ ID NO:40和SEQ ID NO:41的功能性部分的示例性人类Myo15启动子提供于SEQ ID NO:48和49中。

[0216] 在一些实施方案中,SEQ ID NO:40的序列或其功能性部分或衍生物(例如SEQ ID NO:42)接合(例如可操作地连接)至SEQ ID NO:41的序列或其功能性部分或衍生物(例如SEQ ID NO:43-47中的任一者或多者,例如SEQ ID NO:43和44),并且在一些实施方案中,区域的顺序是相反的(例如,SEQ ID NO:41的序列或其功能性部分或衍生物(例如SEQ ID NO:43-47中的任一者或多者,例如SEQ ID NO:43和/或44)接合(例如可操作地连接)至SEQ ID NO:40的序列或其功能性部分或衍生物(例如SEQ ID NO:42)。无论顺序如何,SEQ ID NO:40的序列或其功能性部分或衍生物和SEQ ID NO:41的序列或其功能性部分或衍生物可通过直接融合或核酸接头接合,如上文所述。

[0217] 在一些实施方案中,用于本文所述的组合物和方法中的人类Myo15启动子含有与含有SEQ ID NO:40中所示的序列或其功能性部分或衍生物的区域具有至少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)的区域。SEQ ID NO:40的功能性部分可具有SEQ ID NO:42中所示的核酸的序列。

[0218] 在一些实施方案中,用于本文所述的组合物和方法中的人类Myo15启动子含有与SEQ ID NO:41中所示的序列或其功能性部分或衍生物具有至少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)的区域。SEQ ID NO:41的功能性部分可具有SEQ ID NO:43中所示的序列和/或SEQ ID NO:44中所示的序列。人类Myo15启动子可含有融合至SEQ ID NO:44的核酸序列而无间插核酸的SEQ ID NO:43的核酸序列,如SEQ ID NO:45中所示,或者人类Myo15启动子可含有融合至SEQ ID NO:43的核酸序列而无间插核酸的SEQ ID NO:44的核酸序列,如SEQ ID NO:46中所示。或者,人类Myo15启动子可含有通过内源间插核酸序列(例如如SEQ ID NO:47中所示)或核酸接头接合的SEQ ID NO:43和SEQ ID NO:44的序列。在含有SEQ ID NO:43和SEQ ID NO:44的人类Myo15启动子中,两个序列可以任何顺序包括(例如SEQ ID NO:43可接合至(例如先于)SEQ ID NO:44,或SEQ ID NO:44可接合至(例如先于)SEQ ID NO:43)。

[0219] 用于本文所述人类Myo15启动子中的核酸接头的长度可为约5kb或更短(例如约5kb、4.5kb、4kb、3.5kb、3kb、2.5kb、2kb、1.5kb、1kb、900bp、800bp、700bp、600bp、500bp、450bp、400bp、350bp、300bp、250bp、200bp、150bp、100bp、90bp、80bp、70bp、60bp、50bp、40bp、30bp、25bp、20bp、15bp、10bp、5bp、4bp、3bp、2bp或更短)。可用于本文所述人类Myo15启动子中的核酸接头不会破坏本发明的人类Myo15启动子诱导毛细胞中的转基因表达的能力。

[0220] 前述Myo15启动子序列汇总于下表3中。

[0221] 表3:用于本文所述的Myo15启动子中的示例性核苷酸序列

[0222]

SEQ ID NO.	核酸序列的描述	核酸序列
24	含有 Myo15 的非编码外显子 1 的区域(-6755 至 -7209)	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT AAGCTTCTGCCACTGGCTCCGGCATTGCAGAGAG AAGAGAAGGGGCGGCAGAGCTGAACCTTAGCCTT GCCTTCCTGGGTACCCTTCTGAGCCTCACTGTCTT CTGTGAGATGGGCAAAGTGCGGGTGTGACTCCTT GGCAACGGTGTTACACCAGGGCAGGTAAAGTTGT AGTTATTTGTGGGGTACACCAGGACTGTTAAAGG TGTAACTAT
25	紧接在 Myo15 的翻译起始位点之前的区域(-1 至-1157)	GGTCTCACCCAGCATTTTCACTTCTAATAAGTTCA AATGTGATACGGCACCTTTCTAAAAATTAGTTTTC AGGGAAATAGGGTTCAAAACTGGTAGTGGTAGG GTCCATTCTCACGACCCCCAGGCCTGCTAACCCTG ACCAAGCTACCTATTACTTACCCTCCTCTTTCTCC TCCTCCTCTTTCTCCTTCTCCTGCTTCCCCTCTTCC TTCTCCCTCCCTTCCTCTCCCTCCTCCCCCTCCTTG GCTGTGATCAGATCCAGAGCCTGAATGAGCCTCC TGACCCCACACCCCCACTAGCATGGGCCTGCAAG TGCCCAGAAGTCCCTCCTGCCTCCTAAACTGCCCA GCCGATCCATTAGCTCTTCCTTCTTCCCAGTGAAA GAAGCAGGCACAGCCTGTCCCTCCCGTTCTACAG AAAGGAAGCTACAGCACAGGGAGGGCCAAAGGC CTTCCTGGGACTAGACAGTTGATCAACAGCAGGA

[0223]

SEQ ID NO.	核酸序列的描述	核酸序列
		CTGGAGAGCTGGGCTCCATTTTTGTTCCTTGGTGC CCTGCCCCTCCCCATGACCTGCAGAGACATTCAG CCTGCCAGGCTTTATGAGGTGGGAGCTGGGCTCT CCCTGATGTATTATTCAGCTCCCTGGAGTTGGCCA GCTCCTGTTACACTGGCCACAGCCCTGGGCATCC GCTTCTCACTTCTAGTTTCCCCTCCAAGGTAATGT GGTGGGTCATGATCATTCTATCCTGGCTTCAGGG ACCTGACTCCACTTTGGGGCCATTCGAGGGGTCT AGGGTAGATGATGTCCCCCTGTGGGGATTAATGT CCTGCTCTGTAAAACTGAGCTAGCTGAGATCCAG GAGGGCTTGGCCAGAGACAGCAAGTTGTTGCCAT GGTGACTTTAAAGCCAGGTTGCTGCCCCAGCACA GGCCTCCCAGTCTACCCTCACTAGAAAACAACAC CCAGGCACTTTCCACCACCTCTCAAAGGTGAAAC CCAAGGCTGGTCTAGAGAATGAATTATGGATCCT CGCTGTCCGTGCCACCCAGCTAGTCCCAGCGGCT CAGACACTGAGGAGAGACTGTAGGTTCAGCTACA AGCAAAAAGACCTAGCTGGTCTCCAAGCAGTGTC TCCAAGTCCCTGAACCTGTGACACCTGCCCCAGG CATCATCAGGCACAGAGGGCCACC
26	SEQ ID NO: 24 的部分 (-7166 至 -7091)	CCCATGTCAGCTGCTTGTGCTTTCCAGAGACAAA ACAGGAATAATAGATGTCATTAAATATACATTGG GCCCCAGG
27	SEQ ID NO: 24 的部分 (-7077 至 -6983)	AGCCTGAGCCTCCTTTCCATCTCTGTGGAGGCAG ACATAGGACCCCCAACAAACAGCATGCAGGTTGG GAGCCAGCCACAGGACCCAGGTAAGGG
28	SEQ ID NO: 24 的部分 (SEQ ID NO: 26 融合至 SEQ ID NO: 27)	CCCATGTCAGCTGCTTGTGCTTTCCAGAGACAAA ACAGGAATAATAGATGTCATTAAATATACATTGG GCCCCAGGAGCCTGAGCCTCCTTTCCATCTCTGTG GAGGCAGACATAGGACCCCCAACAAACAGCATG CAGGTTGGGAGCCAGCCACAGGACCCAGGTAAG GG
29	SEQ ID NO: 24 的部分 (SEQ ID NO: 27 融合至 SEQ ID NO: 26)	AGCCTGAGCCTCCTTTCCATCTCTGTGGAGGCAG ACATAGGACCCCCAACAAACAGCATGCAGGTTGG GAGCCAGCCACAGGACCCAGGTAAGGGCCCATGT CAGCTGCTTGTGCTTTCCAGAGACAAAACAGGAA TAATAGATGTCATTAAATATACATTGGGCCCCAG G
30	SEQ ID NO: 24 的部分 (-7166 至 -6983)	CCCATGTCAGCTGCTTGTGCTTTCCAGAGACAAA ACAGGAATAATAGATGTCATTAAATATACATTGG GCCCCAGGCGGTCAATGTGGCAGCCTGAGCCTCC TTTCCATCTCTGTGGAGGCAGACATAGGACCCCC AACAAACAGCATGCAGGTTGGGAGCCAGCCACA GGACCCAGGTAAGGG

[0224]

SEQ ID NO.	核酸序列的描述	核酸序列
31	SEQ ID NO: 25 的部分 (-590 至-509)	TGAGGTGGGAGCTGGGCTCTCCCTGATGTATTATT CAGCTCCCTGGAGTTGGCCAGCTCCTGTTACACTG GCCACAGCCCTG
32	SEQ ID NO: 25 的部分 (-266 至-161)	CACAGGCCTCCCAGTCTACCCTCACTAGAAAACA ACACCCAGGCACTTTCCACCACCTCTCAAAGGTG AAACCCAAGGCTGGTCTAGAGAATGAATTATGGA TCCT
33	SEQ ID NO: 25 的部分 (SEQ ID NO: 31 融合至 SEQ ID NO: 32)	TGAGGTGGGAGCTGGGCTCTCCCTGATGTATTATT CAGCTCCCTGGAGTTGGCCAGCTCCTGTTACACTG GCCACAGCCCTGCACAGGCCTCCCAGTCTACCCT CACTAGAAAACAACACCCAGGCACTTTCCACCAC CTCTCAAAGGTGAAACCCAAGGCTGGTCTAGAGA ATGAATTATGGATCCT
34	SEQ ID NO: 25 的部分 (SEQ ID NO: 32 融合至 SEQ ID NO: 31)	CACAGGCCTCCCAGTCTACCCTCACTAGAAAACA ACACCCAGGCACTTTCCACCACCTCTCAAAGGTG AAACCCAAGGCTGGTCTAGAGAATGAATTATGGA TCCTTGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTG
35	SEQ ID NO: 25 的部分 (-590 至-161)	TGAGGTGGGAGCTGGGCTCTCCCTGATGTATTATT CAGCTCCCTGGAGTTGGCCAGCTCCTGTTACACTG GCCACAGCCCTGGGCATCCGCTTCTCACTTCTAGT TTCCCCTCCAAGGTAATGTGGTGGGTCATGATCAT TCTATCCTGGCTTCAGGGACCTGACTCCACTTTGG GGCCATTCGAGGGGTCTAGGGTAGATGATGTCCC CCTGTGGGGATTAATGTCCTGCTCTGTAAAACTG AGCTAGCTGAGATCCAGGAGGGCTTGGCCAGAGA CAGCAAGTTGTTGCCATGGTGACTTTAAAGCCAG GTTGCTGCCCCAGCACAGGCCTCCCAGTCTACCCT CACTAGAAAACAACACCCAGGCACTTTCCACCAC CTCTCAAAGGTGAAACCCAAGGCTGGTCTAGAGA ATGAATTATGGATCCT
36	SEQ ID NO: 24 融合至 SEQ ID NO: 25	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT AAGCTTCTGCCACTGGCTCCGGCATTGCAGAGAG AAGAGAAGGGGCGGCAGAGCTGAACCTTAGCCTT GCCTTCCTGGGTACCCTTCTGAGCCTCACTGTCTT CTGTGAGATGGGCAAAGTGCGGGTGTGACTCCTT GGCAACGGTGTTACACCAGGGCAGGTAAAGTTGT AGTTATTTGTGGGGTACACCAGGACTGTTAAAGG TGTAACTATGGTCTCACCCAGCATTTTCACTTCTA

[0225]

SEQ ID NO.	核酸序列的描述	核酸序列
		ATAAGTTCAAATGTGATACGGCACCTTTCTAAAA ATTAGTTTTCAGGGAAATAGGGTTCAAAACTGGT AGTGGTAGGGTCCATTCTCACGACCCCCAGGCCT GCTAACCCTGACCAAGCTACCTATTACTTACCCTC CTCTTTCTCCTCCTCCTCTTTCTCCTTCTCCTGCTT CCCCTCTTCCTTCTCCCTCCCTTCCTCTCCCTCCTC CCCCTCCTTGGCTGTGATCAGATCCAGAGCCTGA ATGAGCCTCCTGACCCCACACCCCCACTAGCATG GGCCTGCAAGTGCCCAGAAGTCCCTCCTGCCTCC TAAACTGCCCAGCCGATCCATTAGCTCTTCCTTCT TCCCAGTGAAAGAAGCAGGCACAGCCTGTCCCTC CCGTTCTACAGAAAGGAAGCTACAGCACAGGGA GGGCCAAAGGCCTTCCTGGGACTAGACAGTTGAT CAACAGCAGGACTGGAGAGCTGGGCTCCATTTTT GTTCCTTGGTGCCCTGCCCCTCCCCATGACCTGCA GAGACATTCAGCCTGCCAGGCTTTATGAGGTGGG AGCTGGGCTCTCCCTGATGTATTATTCAGCTCCCT GGAGTTGGCCAGCTCCTGTTACACTGGCCACAGC CCTGGGCATCCGCTTCTCACTTCTAGTTTCCCCTC CAAGGTAATGTGGTGGGTCATGATCATTCTATCCT GGCTTCAGGGACCTGACTCCACTTTGGGGCCATT CGAGGGGTCTAGGGTAGATGATGTCCCCCTGTGG GGATTAATGTCCTGCTCTGTAAAACTGAGCTAGC TGAGATCCAGGAGGGCTTGGCCAGAGACAGCAA GTTGTTGCCATGGTGACTTTAAAGCCAGGTTGCTG CCCCAGCACAGGCCTCCCAGTCTACCCTCACTAG AAAACAACACCCAGGCACTTTCCACCACCTCTCA AAGGTGAAACCCAAGGCTGGTCTAGAGAATGAAT TATGGATCCTCGCTGTCCGTGCCACCCAGCTAGTC CCAGCGGCTCAGACACTGAGGAGAGACTGTAGGT TCAGCTACAAGCAAAAAGACCTAGCTGGTCTCCA AGCAGTGTCTCCAAGTCCCTGAACCTGTGACACC TGCCCCAGGCATCATCAGGCACAGAGGGCCACC
37	SEQ ID NO: 25 融合至 SEQ ID NO: 24	GGTCTCACCCAGCATTTTCACTTCTAATAAGTTCA AATGTGATACGGCACCTTTCTAAAAATTAGTTTTC AGGGAAATAGGGTTCAAAACTGGTAGTGGTAGG GTCCATTCTCACGACCCCCAGGCCTGCTAACCCTG ACCAAGCTACCTATTACTTACCCTCCTCTTTCTCC TCCTCCTCTTTCTCCTTCTCCTGCTTCCCCTCTTCC TTCTCCCTCCCTTCCTCTCCCTCCTCCCCCTCCTTG GCTGTGATCAGATCCAGAGCCTGAATGAGCCTCC TGACCCCACACCCCCACTAGCATGGGCCTGCAAG TGCCCAGAAGTCCCTCCTGCCTCCTAAACTGCCCA GCCGATCCATTAGCTCTTCCTTCTTCCCAGTGAAA GAAGCAGGCACAGCCTGTCCCTCCCGTTCTACAG AAAGGAAGCTACAGCACAGGGAGGGCCAAAGGC CTTCCTGGGACTAGACAGTTGATCAACAGCAGGA

[0226]

SEQ ID NO.	核酸序列的描述	核酸序列
		CTGGAGAGCTGGGCTCCATTTTTGTTCCTTGGTGC CCTGCCCCTCCCCATGACCTGCAGAGACATTCAG CCTGCCAGGCTTTATGAGGTGGGAGCTGGGCTCT CCCTGATGTATTATTCAGCTCCCTGGAGTTGGCCA GCTCCTGTTACACTGGCCACAGCCCTGGGCATCC GCTTCTCACTTCTAGTTTCCCCTCCAAGGTAATGT GGTGGGTCATGATCATTCTATCCTGGCTTCAGGG ACCTGACTCCACTTTGGGGCCATTCGAGGGGTCT AGGGTAGATGATGTCCCCCTGTGGGGATTAATGT CCTGCTCTGTAAAACTGAGCTAGCTGAGATCCAG GAGGGCTTGGCCAGAGACAGCAAGTTGTTGCCAT GGTGACTTTAAAGCCAGGTTGCTGCCCCAGCACA GGCCTCCCAGTCTACCCTCACTAGAAAACAACAC CCAGGCACTTTCCACCACCTCTCAAAGGTGAAAC CCAAGGCTGGTCTAGAGAATGAATTATGGATCCT CGCTGTCCGTGCCACCCAGCTAGTCCCAGCGGCT CAGACACTGAGGAGAGACTGTAGGTTCAGCTACA AGCAAAAAGACCTAGCTGGTCTCCAAGCAGTGTC TCCAAGTCCCTGAACCTGTGACACCTGCCCCAGG CATCATCAGGCACAGAGGGCCACCCTGCAGCTCA GCCTACTACTTGCTTTCCAGGCTGTTCCTAGTTCC CATGTCAGCTGCTTGTGCTTTCCAGAGACAAAAC AGGAATAATAGATGTCATTAAATATACATTGGGC CCCAGGCGGTCAATGTGGCAGCCTGAGCCTCCTT TCCATCTCTGTGGAGGCAGACATAGGACCCCCAA CAAACAGCATGCAGGTTGGGAGCCAGCCACAGG ACCCAGGTAAGGGGCCCTGGGTCCTTAAGCTTCT GCCACTGGCTCCGGCATTGCAGAGAGAAGAGAA GGGGCGGCAGAGCTGAACCTTAGCCTTGCCTTCC TGGGTACCCTTCTGAGCCTCACTGTCTTCTGTGAG ATGGGCAAAGTGCGGGTGTGACTCCTTGGCAACG GTGTTACACCAGGGCAGGTAAAGTTGTAGTTATT TGTGGGGTACACCAGGACTGTTAAAGGTGTAACT AT
38	含有 SEQ ID NO: 26 和 SEQ ID NO: 27 的 SEQ ID NO: 24 的部分融合至含有 SEQ ID NO: 31 和 SEQ ID NO: 32 的 SEQ ID NO: 25 的部分	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT AAGCTTCTGCCACTGGCTCCGGCATTGCAGAGAG AAGAGAAGGGGCGGCAGACTGGAGAGCTGGGCT CCATTTTTGTTCCTTGGTGCCCTGCCCCTCCCCAT GACCTGCAGAGACATTCAGCCTGCCAGGCTTTAT GAGGTGGGAGCTGGGCTCTCCCTGATGTATTATT CAGCTCCCTGGAGTTGGCCAGCTCCTGTTACACTG

[0227]

SEQ ID NO.	核酸序列的描述	核酸序列
		GCCACAGCCCTGGGCATCCGCTTCTCACTTCTAGT TTCCCCTCCAAGGTAATGTGGTGGGTCATGATCAT TCTATCCTGGCTTCAGGGACCTGACTCCACTTTGG GGCCATTCGAGGGGTCTAGGGTAGATGATGTCCC CCTGTGGGGATTAATGTCCTGCTCTGTAAAACTG AGCTAGCTGAGATCCAGGAGGGCTTGGCCAGAGA CAGCAAGTTGTTGCCATGGTGACTTTAAAGCCAG GTTGCTGCCCCAGCACAGGCCTCCCAGTCTACCCT CACTAGAAAACAACACCCAGGCACTTTCCACCAC CTCTCAAAGGTGAAACCCAAGGCTGGTCTAGAGA ATGAATTATGGATCCTCGCTGTCCGTGCCACCCA GCTAGTCCCAGCGGCTCAGACACTGAGGAGAGAC TGTAGGTTCAGCTACAAGCAAAAAGACCTAGCTG GTCTCCAAGCAGTGTCTCCAAGTCCCTGAACCTGT GACACCTGCCCCAGGCATCATCAGGCACAGAGGG CCACC
39	含有 SEQ ID NO: 26 和 SEQ ID NO: 27 的 SEQ ID NO: 24 的部分融合至含有 SEQ ID NO: 31 和 SEQ ID NO: 32 的 SEQ ID NO: 25 的部分	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGCTGCCATGGT GACTTTAAAGCCAGGTTGCTGCCCCAGCACAGGC CTCCCAGTCTACCCTCACTAGAAAACAACACCCA GGCACTTTCCACCACCTCTCAAAGGTGAAACCCA AGGCTGGTCTAGAGAATGAATTATGGATCCTCGC TGTCCGTGCCACCCAGCTAGTCCCAGCGGCTCAG ACACTG
40	人类 Myo15 启动子的区域 1	GTATGCCTTTTGAGATGGATGCAGCAGGTTCTGT GAGGCTGCCAGGAGGGGTAGAGTTCCCGGGGGC CTCGGGCCCCGCTGGAGTGTGGAGCAGGCCCATG CTCAGCTCTCCAGGCTGTTCGTGGCTCCCCTGTCA GCTGCTCACTCCTTTCCAGAGACAAAACAGGAAT AATAGACATCATTAAATATACATAGGGCCCCAGG CGGTCGGCGTGGTGGGCTGGGCCTCCCTTCC
41	人类 Myo15 启动子的区域 2	TGCCCTGCCTTCTGAGCCGGCAGCCTGGCTCCCCA CCCCATGTATTATTCAGCTCCTGAGAGCCAGCCA GCTCCTGTTACACTGACCGCAGCCCAGCACCTGC TCTGCCCATTCCCCTCCTCCCTTGCCTAGGACCTA GAGGGTTCAAAGTTCTCCTCCAAGATGACTTGGT GGGCTTTGGCCATCCCACCCTAGGCCCCACTTCTG GCCCAGTGCAGGTGTGCTGGTGATTTAGGGCAGG TGGCATTCCATCTCTGTGGCTCAATGTCTTCCTCT

[0228]

SEQ ID NO.	核酸序列的描述	核酸序列
		GTGAAGCCGAAGTGACCCAAGGGCTCCCTTCATG GGGTTGAGCCAGCTGTGGCCCAGGGAGGGCCTAA CCAGGATGAGCACTGATGTTGCCATGACGACTCC GAGGCCAGAATGTCTCCCCCAGCACAGGCCTCAT AGGCAGGCTTCCCCATCCTGGTAAACAACACCCA CACACTTTCTACTACTGCTCTAGGGTGAAACCCA AGGCGCTCTAGAGGAGATGAATTATGGATCCGCC CTCCCGGAATCCTGGCTCGGCCCTCCCCACGCCA CCCAGGGCCAGTCGGGTCTGCTCACAGCCCGAGG AGGCCGCGTGTCCAGCCGCGGGCAAGAGACAGA GCAGGTCCCTGTGTCTCCAAGTCCCTGAGCCCGT GACACCGGCCCCAGGCCCTGTAGAGAGCAGGCA GCCACC
42	SEQ ID NO: 40 的部分	CCCCTGTCAGCTGCTCACTCCTTTCCAGAGACAAA ACAGGAATAATAGACATCATTAAATATACATAGG GCCCCAGG
43	SEQ ID NO: 41 的部分	TGAGCCGGCAGCCTGGCTCCCCACCCCATGTATT ATTCAGCTCCTGAGAGCCAGCCAGCTCCTGTTAC ACTGACCGCAGCCC
44	SEQ ID NO: 41 的部分	CACAGGCCTCATAGGCAGGCTTCCCCATCCTGGT AAACAACACCCACACACTTTCTACTACTGCTCTA GGGTGAAACCCAAGGCGCTCTAGAGGAGATGAA TTATGGATCC
45	SEQ ID NO: 41 的部分 (SEQ ID NO: 43 融合至 SEQ ID NO: 44)	TGAGCCGGCAGCCTGGCTCCCCACCCCATGTATT ATTCAGCTCCTGAGAGCCAGCCAGCTCCTGTTAC ACTGACCGCAGCCCCACAGGCCTCATAGGCAGGC TTCCCCATCCTGGTAAACAACACCCACACACTTTC TACTACTGCTCTAGGGTGAAACCCAAGGCGCTCT AGAGGAGATGAATTATGGATCC
46	SEQ ID NO: 41 的部分 (SEQ ID NO: 44 融合至 SEQ ID NO: 43)	CACAGGCCTCATAGGCAGGCTTCCCCATCCTGGT AAACAACACCCACACACTTTCTACTACTGCTCTA GGGTGAAACCCAAGGCGCTCTAGAGGAGATGAA TTATGGATCCTGAGCCGGCAGCCTGGCTCCCCAC CCCATGTATTATTCAGCTCCTGAGAGCCAGCCAG CTCCTGTTACACTGACCGCAGCCC
47	SEQ ID NO: 41 的部分(包括 SEQ ID NO: 43 和 SEQ ID NO: 44 的邻接序列)	TGAGCCGGCAGCCTGGCTCCCCACCCCATGTATT ATTCAGCTCCTGAGAGCCAGCCAGCTCCTGTTAC ACTGACCGCAGCCCAGCACCTGCTCTGCCCATTC CCCTCCTCCCTTGCCTAGGACCTAGAGGGTTCAA AGTTCTCCTCCAAGATGACTTGGTGGGCTTTGGCC ATCCCACCCTAGGCCCCACTTCTGGCCCAGTGCA GGTGTGCTGGTGATTTAGGGCAGGTGGCATTCCA TCTCTGTGGCTCAATGTCTTCCTCTGTGAAGCCGA AGTGACCCAAGGGCTCCCTTCATGGGGTTGAGCC AGCTGTGGCCCAGGGAGGGCCTAACCAGGATGA GCACTGATGTTGCCATGACGACTCCGAGGCCAGA

[0229]

SEQ ID NO.	核酸序列的描述	核酸序列
		ATGTCTCCCCCAGCACAGGCCTCATAGGCAGGCT TCCCCATCCTGGTAAACAACACCCACACACTTTCT ACTACTGCTCTAGGGTGAAACCCAAGGCGCTCTA GAGGAGATGAATTATGGATCC
48	含有 SEQ ID NO: 40 和 SEQ ID NO: 41 的多核苷酸	GTATGCCTTTTGAGATGGATGCAGCAGGTTCTGT GAGGCTGCCAGGAGGGGTAGAGTTCCCGGGGGC CTCGGGCCCCGCTGGAGTGTGGAGCAGGCCCATG CTCAGCTCTCCAGGCTGTTCGTGGCTCCCCTGTCA GCTGCTCACTCCTTTCCAGAGACAAAACAGGAAT AATAGACATCATTAAATATACATAGGGCCCCAGG CGGTCGGCGTGGTGGGCTGGGCCTCCCTTCCCCA TAACACTGAGCTGCTCTGCTGGGCCAACCGTGCT CCTGGGCCAGCCAGAGGACCCCCATGAGGCGGCA TGCAGGCGGGGAGCAGGCCACAGAACGCAGGTA AGGAGACCTTAGCCTAGAGTCCTTGGGGTCTGTC ACTGGCCACCCTCGCATCCCAGGCTGCAGGAAAC TGAGGCCCAGAGAGGACAAGGACTTTCCTGGACC CACACAGCCAGTCAGTGACAGAGCCTAGGGTCTG AGCCAGGCCTGACCCAACCTCCATTTCTGCCTCTC TACCCCTGCCCCCGCCCCAACACACACACACACA CAAGTGGAGTTCCACTGAAACGCCCCTCCTTGCC CTGCCTTCTGAGCCGGCAGCCTGGCTCCCCACCCC ATGTATTATTCAGCTCCTGAGAGCCAGCCAGCTC CTGTTACACTGACCGCAGCCCAGCACCTGCTCTG CCCATTCCCCTCCTCCCTTGCCTAGGACCTAGAGG GTTCAAAGTTCTCCTCCAAGATGACTTGGTGGGCT TTGGCCATCCCACCCTAGGCCCCACTTCTGGCCCA GTGCAGGTGTGCTGGTGATTTAGGGCAGGTGGCA TTCCATCTCTGTGGCTCAATGTCTTCCTCTGTGAA GCCGAAGTGACCCAAGGGCTCCCTTCATGGGGTT GAGCCAGCTGTGGCCCAGGGAGGGCCTAACCAG GATGAGCACTGATGTTGCCATGACGACTCCGAGG CCAGAATGTCTCCCCCAGCACAGGCCTCATAGGC AGGCTTCCCCATCCTGGTAAACAACACCCACACA CTTTCTACTACTGCTCTAGGGTGAAACCCAAGGC GCTCTAGAGGAGATGAATTATGGATCCGCCCTCC CGGAATCCTGGCTCGGCCCTCCCCACGCCACCCA GGGCCAGTCGGGTCTGCTCACAGCCCGAGGAGGC CGCGTGTCCAGCCGCGGGCAAGAGACAGAGCAG GTCCCTGTGTCTCCAAGTCCCTGAGCCCGTGACAC CGGCCCCAGGCCCTGTAGAGAGCAGGCAGCCACC
49	含有 SEQ ID NO: 42、SEQ ID NO: 43 和 SEQ ID NO: 44 的多核苷酸	GCAGGCCCATGCTCAGCTCTCCAGGCTGTTCGTG GCTCCCCTGTCAGCTGCTCACTCCTTTCCAGAGAC AAAACAGGAATAATAGACATCATTAAATATACAT AGGGCCCCAGGCGGTCGGCGTGGTGGGCTGGGCC TCCCTTCCCCATAACACTGAGCTGCTCTGCTGGGC CAACCGTGCTCCTGGGCCAGCCAGAGGACCCCCA

[0230]

SEQ ID NO.	核酸序列的描述	核酸序列
		TGAGGCGGCATGCAGGCGGGGAGCAGGCCACAG AACGCAGGTAAGGAGACCTTGCCTTCTGAGCCGG CAGCCTGGCTCCCCACCCCATGTATTATTCAGCTC CTGAGAGCCAGCCAGCTCCTGTTACACTGACCGC AGCCCAGCACCTGCTCTGCCCATTCCCCTCCTCCC TTGCCTAGGACCTAGAGGGTTCAAAGTTCTCCTCC AAGATGACTTGGTGGGCTTTGGCCATCGGGCCTA ACCAGGATGAGCACTGATGTTGCCATGACGACTC CGAGGCCAGAATGTCTCCCCCAGCACAGGCCTCA TAGGCAGGCTTCCCCATCCTGGTAAACAACACCC ACACACTTTCTACTACTGCTCTAGGGTGAAACCC AAGGCGCTCTAGAGGAGATGAATTATGGATCCGC CCTCCCGGAATCCTGGCTCGGCCCTCCCCACGC
50	含有 SEQ ID NO: 26 和 SEQ ID NO: 27 的 SEQ ID NO: 24 的部分	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT
51	含有 SEQ ID NO: 31 的 SEQ ID NO: 25 的部分	TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGC
52	含有 SEQ ID NO: 32 的 SEQ ID NO: 25 的部分	TGCCATGGTGACTTTAAAGCCAGGTTGCTGCCCC AGCACAGGCCTCCCAGTCTACCCTCACTAGAAAA CAACACCCAGGCACTTTCCACCACCTCTCAAAGG TGAAACCCAAGGCTGGTCTAGAGAATGAATTATG GATCCTCGCTGTCCGTGCCACCCAGCTAGTCCCA GCGGCTCAGACACTG
53	SEQ ID NO: 50 融合至 SEQ ID NO: 51	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGC
54	SEQ ID NO: 50 融合至 SEQ ID NO: 52	CTGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGT TCCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAG AGACAAAACAGGAATAATAGATGTCATTAAATAT ACATTGGGCCCCAGGCGGTCAATGTGGCAGCCTG AGCCTCCTTTCCATCTCTGTGGAGGCAGACATAG GACCCCCAACAAACAGCATGCAGGTTGGGAGCCA GCCACAGGACCCAGGTAAGGGGCCCTGGGTCCTT

[0231]

SEQ ID NO.	核酸序列的描述	核酸序列
		TGCCATGGTGACTTTAAAGCCAGGTTGCTGCCCC AGCACAGGCCTCCCAGTCTACCCTCACTAGAAAA CAACACCCAGGCACTTTCCACCACCTCTCAAAGG TGAAACCCAAGGCTGGTCTAGAGAATGAATTATG GATCCTCGCTGTCCGTGCCACCCAGCTAGTCCCA GCGGCTCAGACACTG
55	SEQ ID NO: 51 融合至 SEQ ID NO: 52	TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGCTGCCATGGT GACTTTAAAGCCAGGTTGCTGCCCCAGCACAGGC CTCCCAGTCTACCCTCACTAGAAAACAACACCCA GGCACTTTCCACCACCTCTCAAAGGTGAAACCCA AGGCTGGTCTAGAGAATGAATTATGGATCCTCGC TGTCCGTGCCACCCAGCTAGTCCCAGCGGCTCAG ACACTG
56	SEQ ID NO: 51 融合至 SEQ ID NO: 50，所述 SEQ ID NO: 50 融合至 SEQ ID NO: 52	TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGCCTGCAGCTC AGCCTACTACTTGCTTTCCAGGCTGTTCCTAGTTC CCATGTCAGCTGCTTGTGCTTTCCAGAGACAAAA CAGGAATAATAGATGTCATTAAATATACATTGGG CCCCAGGCGGTCAATGTGGCAGCCTGAGCCTCCT TTCCATCTCTGTGGAGGCAGACATAGGACCCCCA ACAAACAGCATGCAGGTTGGGAGCCAGCCACAG GACCCAGGTAAGGGGCCCTGGGTCCTTTGCCATG GTGACTTTAAAGCCAGGTTGCTGCCCCAGCACAG GCCTCCCAGTCTACCCTCACTAGAAAACAACACC CAGGCACTTTCCACCACCTCTCAAAGGTGAAACC CAAGGCTGGTCTAGAGAATGAATTATGGATCCTC GCTGTCCGTGCCACCCAGCTAGTCCCAGCGGCTC AGACACTG
57	SEQ ID NO: 52 融合至 SEQ ID NO: 50，所述 SEQ ID NO: 50 融合至 SEQ ID NO: 51	TGCCATGGTGACTTTAAAGCCAGGTTGCTGCCCC AGCACAGGCCTCCCAGTCTACCCTCACTAGAAAA CAACACCCAGGCACTTTCCACCACCTCTCAAAGG TGAAACCCAAGGCTGGTCTAGAGAATGAATTATG GATCCTCGCTGTCCGTGCCACCCAGCTAGTCCCA GCGGCTCAGACACTGCTGCAGCTCAGCCTACTAC TTGCTTTCCAGGCTGTTCCTAGTTCCCATGTCAGC TGCTTGTGCTTTCCAGAGACAAAACAGGAATAAT AGATGTCATTAAATATACATTGGGCCCCAGGCGG TCAATGTGGCAGCCTGAGCCTCCTTTCCATCTCTG TGGAGGCAGACATAGGACCCCCAACAAACAGCA TGCAGGTTGGGAGCCAGCCACAGGACCCAGGTAA GGGGCCCTGGGTCCTTTTTATGAGGTGGGAGCTG GGCTCTCCCTGATGTATTATTCAGCTCCCTGGAGT TGGCCAGCTCCTGTTACACTGGCCACAGCCCTGG GCATCCGC

[0232]

SEQ ID NO.	核酸序列的描述	核酸序列
58	SEQ ID NO: 51 融合至 SEQ ID NO: 52，所述 SEQ ID NO: 52 融合至 SEQ ID NO: 50	TTTATGAGGTGGGAGCTGGGCTCTCCCTGATGTAT TATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTAC ACTGGCCACAGCCCTGGGCATCCGCTGCCATGGT GACTTTAAAGCCAGGTTGCTGCCCCAGCACAGGC CTCCCAGTCTACCCTCACTAGAAAACAACACCCA GGCACTTTCCACCACCTCTCAAAGGTGAAACCCA AGGCTGGTCTAGAGAATGAATTATGGATCCTCGC TGTCCGTGCCACCCAGCTAGTCCCAGCGGCTCAG ACACTGCTGCAGCTCAGCCTACTACTTGCTTTCCA GGCTGTTCCTAGTTCCCATGTCAGCTGCTTGTGCT TTCCAGAGACAAAACAGGAATAATAGATGTCATT AAATATACATTGGGCCCCAGGCGGTCAATGTGGC AGCCTGAGCCTCCTTTCCATCTCTGTGGAGGCAG ACATAGGACCCCCAACAAACAGCATGCAGGTTGG GAGCCAGCCACAGGACCCAGGTAAGGGGCCCTG GGTCCTT
59	含有 SEQ ID NO: 26 和 SEQ ID NO: 27 的 SEQ ID NO: 24 的部分融合至含有 SEQ ID NO: 31 和 SEQ ID NO: 32 的 SEQ ID NO: 25 的部分	TGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGTT CCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAGA GACAAAACAGGAATAATAGATGTCATTAAATATA CATTGGGCCCCAGGCGGTCAATGTGGCAGCCTGA GCCTCCTTTCCATCTCTGTGGAGGCAGACATAGG ACCCCCAACAAACAGCATGCAGGTTGGGAGCCAG CCACAGGACCCAGGTAAGGGGCCCTGGGTCCTTA AGCTTCTGCCACTGGCTCCGGCATTGCAGAGAGA AGAGAAGGGGCGGCAGACTGGAGAGCTGGGCTC CATTTTTGTTCCTTGGTGCCCTGCCCCTCCCCATG ACCTGCAGAGACATTCAGCCTGCCAGGCTTTATG AGGTGGGAGCTGGGCTCTCCCTGATGTATTATTC AGCTCCCTGGAGTTGGCCAGCTCCTGTTACACTG GCCACAGCCCTGGGCATCCGCTTCTCACTTCTAGT TTCCCCTCCAAGGTAATGTGGTGGGTCATGATCAT TCTATCCTGGCTTCAGGGACCTGACTCCACTTTGG GGCCATTCGAGGGGTCTAGGGTAGATGATGTCCC CCTGTGGGGATTAATGTCCTGCTCTGTAAAACTG AGCTAGCTGAGATCCAGGAGGGCTTGGCCAGAGA CAGCAAGTTGTTGCCATGGTGACTTTAAAGCCAG GTTGCTGCCCCAGCACAGGCCTCCCAGTCTACCCT CACTAGAAAACAACACCCAGGCACTTTCCACCAC CTCTCAAAGGTGAAACCCAAGGCTGGTCTAGAGA ATGAATTATGGATCCTCGCTGTCCGTGCCACCCA GCTAGTCCCAGCGGCTCAGACACTGAGGAGAGAC TGTAGGTTCAGCTACAAGCAAAAAGACCTAGCTG GTCTCCAAGCAGTGTCTCCAAGTCCCTGAACCTGT GACACCTGCCCCAGGCATCATCAGGCACAGAGGG CCACC
60	含有 SEQ ID NO: 26 和	TGCAGCTCAGCCTACTACTTGCTTTCCAGGCTGTT CCTAGTTCCCATGTCAGCTGCTTGTGCTTTCCAGA

[0233]

SEQ ID NO.	核酸序列的描述	核酸序列
	SEQ ID NO: 27 的 SEQ ID NO: 24 的部分融合至含有 SEQ ID NO: 31 和 SEQ ID NO: 32 的 SEQ ID NO: 25 的部分	GACAAAACAGGAATAATAGATGTCATTAAATATA CATTGGGCCCCAGGCGGTCAATGTGGCAGCCTGA GCCTCCTTTCCATCTCTGTGGAGGCAGACATAGG ACCCCCAACAAACAGCATGCAGGTTGGGAGCCAG CCACAGGACCCAGGTAAGGGGCCCTGGGTCCTTT TTATGAGGTGGGAGCTGGGCTCTCCCTGATGTATT ATTCAGCTCCCTGGAGTTGGCCAGCTCCTGTTACA CTGGCCACAGCCCTGGGCATCCGCTGCCATGGTG ACTTTAAAGCCAGGTTGCTGCCCCAGCACAGGCC TCCCAGTCTACCCTCACTAGAAAACAACACCCAG GCACTTTCCACCACCTCTCAAAGGTGAAACCCAA GGCTGGTCTAGAGAATGAATTATGGATCCTCGCT GTCCGTGCCACCCAGCTAGTCCCAGCGGCTCAGA CACTG

[0234] 可与本文所述的组合物和方法结合使用的额外Myo15启动子包括与表3中所示的核酸序列以及表3中所示的核酸序列的功能性部分或衍生物具有至少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)的核酸分子。表3中所列出的Myo15启动子表征于国际申请公开第WO2019210181A1号和第WO2020163761A1号中,这些国际申请公开通过引用并入本文。

[0235] 在smCBA启动子包括在本文所述的双载体系统中(例如在双载体系统中的第一载体中)的实施方案中,smCBA启动子可具有美国专利第8,298,818号中所述的smCBA启动子的序列,所述专利通过引用并入本文。在一些实施方案中,smCBA启动子具有以下序列:

[0236] GGTACCTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCG TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTA TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTG GCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT ATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT CGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTATT TTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGGGCGAGG GGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTTTATGGC GAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGCGCTGCCTTCGCCCC GTGCCCCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTGAGCGGGC GGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTTGTTTCTTTTCTGTGGCTGCGTGA AAGCCTTGAGGGGCTCCGGGAGCTAGAGCCTCTGCTAACCATGTTCATGCCTTCTTCTTTTTCCTACAGCTCCTGG GCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCA(SEQ ID NO:70)。

[0237] 在一些实施方案中,smCBA启动子具有以下序列:

[0238] AATTCGGTACCCTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGT TCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAAT GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGGTAAACTGCC CACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC

CATGGTCGAGGTGAGCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTATTTA
TTTATTTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGGCGCGCGCCAGGCGGGGCGGGGCGGG
GCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGGCGCGCTCCGAAAGTTTCCTTT
TATGGCGAGGCGGCGGCGGCGGCGGCCCTATAAAAAGCGAAGCGCGCGGCGGGCGGGAGTCGCTGCGACGCTGCCT
TCGCCCCGTGCCCCGCTCCGCCGCCGCCTCGCGCCGCCCGCCCCGGCTCTGACTGACCGCGTTACTCCCACAGGTG
AGCGGGCGGGACGGCCCTTCTCCTCCGGGCTGTAATTAGCGCTTGGTTTAATGACGGCTTGTTTCTTTTCTGTGGC
TGCGTGAAAGCCTTGAGGGGCTCCGGGAGCTAGAGCCTCTGCTAACCATGTTCATGCCTTCTTCTTTTTCCTACAG
CTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAAAG(SEQ ID NO:84)。

[0239] 一旦将编码OTOF的多核苷酸并入哺乳动物细胞的核DNA中或稳定在游离型单体或多联体中,就可立即通过本领域中已知的方法来诱导此多核苷酸的转录。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脱氧羟四环素。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。

[0240] 可包括在用于本文所述的组合物和方法中的核酸载体中的其他DNA序列元件包括增强子序列。增强子代表另一类调控元件,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,用于本文所述的组合物和方法中的多核苷酸包括编码OTOF蛋白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。用于本文所述的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额外增强子序列公开于Yaniv等人类,Nature 297:17(1982)中。增强子可剪接至含有编码OTOF蛋白的多核苷酸的载体中,例如在此基因的位置5'或3'处。在优选取向中,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OTOF蛋白的多核苷酸的5'。

[0241] 本文所述的核酸载体可包括土拨鼠转录后调控元件(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。WPRE可位于编码OTOF蛋白的C末端部分的多核苷酸与poly(A)序列之间的第二核酸载体中。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:

[0242] GATCCAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCT
TTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCT
CCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCAC
TGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTC
CCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCA

CTGACAATTCCGTGGTGTTGTCGGGGAAATCATCGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGA(SEQ ID NO:23)。

[0243] 在其他实施方案中,WPRE具有以下序列:

[0244] AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTAGTTCTTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATCTAGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAA(SEQ ID NO:61)。

[0245] 在一些实施方案中,用于本文所述的组合物和方法中的核酸载体包括报告基因序列,其可用于验证例如特定细胞和组织中(例如耳蜗毛细胞中)的OTOF基因表达。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众所周知的其他蛋白质的DNA序列。当与驱动其表达的调控元件缔合时,报告基因序列提供可通过常规方法检测到的信号,这些常规方法包括酶促、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。当转基因为绿色荧光蛋白或荧光素酶时,携带信号的载体可以视觉方式通过光度计中的色彩或光产生来测量。

[0246] 重叠双载体

[0247] 一种用于在哺乳动物细胞中表达大蛋白的方法涉及使用重叠双载体。这种方法基于使用两种核酸载体,两种核酸载体中的每一者含有编码目标蛋白质的多核苷酸的一部分,并且具有与另一多核苷酸重叠的限定序列区域。同源重组可在重叠区域发生,并且导致编码全长目标蛋白质的单一核酸分子的形成。

[0248] 用于本文所述的方法和组合物的重叠双载体含有至少1千碱基(kb)的重叠序列(例如,1kb、1.5kb、2kb、2.5kb、3kb或更多的重叠序列)。核酸载体经设计,使得重叠区域以OTOF外显子边界为中心,在边界的任一侧具有等量的重叠。基于启动子的大小和编码OTOF C2结构域的多核苷酸部分的位置来选择边界。重叠区域以外显子边界为中心,所述外显子边界出现在编码C2C结构域的多核苷酸部分之外(例如,在编码C2C结构域的多核苷酸部分之后)。可选择编码C2D结构域的多核苷酸部分内的外显子边界作为重叠区域的中心,或者位于编码C2D结构域的多核苷酸部分之后和编码C2E结构域的多核苷酸部分之前的外显子边界可用作重叠区域的中心。用于本文所述的方法和组合物的核酸载体还经设计,使得大约一半的OTOF基因包含在每个载体内(例如,每个载体含有编码大约一半OTOF蛋白的多核苷酸)。

[0249] 一种示例性重叠双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至外显子1-28的CAG启动子和紧邻编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子28/29边界3'的500个碱基对(bp);和第二核酸载体,所述第二核酸载体含有紧邻编码OTOF蛋白(例如,人类OTOF,例

如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子28/29边界5'的500bp和其余外显子(例如,小鼠OTOF的外显子29-48,人类OTOF的外显子29-45和47或外显子29-46)以及poly(A)序列(例如,牛生长激素(bGH)poly(A)信号序列)。在这种重叠双载体系统中,重叠序列以外显子28/29边界为中心,所述外显子28/29边界在编码C2D结构域的多核苷酸部分之后。另一种示例性重叠双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至外显子1-24的CAG启动子和紧邻编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子24/25边界3'的500bp;和第二核酸载体,所述第二核酸载体含有紧邻编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子24/25边界5'的500bp和其余外显子(例如,小鼠OTOF的外显子25-48,人类OTOF的外显子25-45和47或外显子25-46)以及poly(A)序列(例如,bGH poly(A)信号序列)。在这种重叠双载体系统中,重叠序列以外显子24/25边界为中心,所述外显子24/25边界在编码C2D结构域的多核苷酸部分之内。上述两个外显子边界可与大小与CAG启动子相似的任何启动子(例如,CMV启动子或smCBA启动子)一起使用,诸如1kb或更短(例如大约1kb、950bp、900bp、850bp、800bp、750bp、700bp、650bp、600bp、550bp 500bp、450bp、400bp、350bp、300bp或更短)的启动子。例如,在前述双载体系统中的任一者中,CMV启动子或smCBA启动子可用于代替CAG启动子。具有1kb或更短的序列的Myo15启动子(例如,上文所述的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于代替CAG启动子。或者,可选择不同的外显子边界,所述外显子边界在编码C2D结构域的多核苷酸的部分之内或之后并且在编码C2E结构域的多核苷酸部分之前。含有此大小的启动子的核酸载体可任选地含有OTOF UTR。例如,在重叠区域以OTOF的外显子28/29边界为中心的前述重叠双载体系统中,第二核酸载体可含有全长OTOF 3'UTR(例如,编码人类OTOF的双载体系统中的1035bp人类OTOF 3'UTR,或编码小鼠OTOF的双载体系统中的1001bp小鼠OTOF 3'UTR)。在重叠区域以OTOF的外显子24/25边界为中心的前述重叠双载体系统中,第一核酸载体和第二核酸载体都不含OTOF UTR。

[0250] 在一些实施方案中,重叠双载体系统中的第一核酸载体含有长启动子(例如,长于1kb,例如1.1kb、1.25kb、1.5kb、1.75kb、2kb、2.5kb、3kb或更长的启动子)。在此类重叠双载体系统中,重叠区域可以位于编码C2C结构域的多核苷酸部分之后且编码C2D结构域的多核苷酸部分之前的外显子边界为中心。例如,用于本文所述的方法和组合物的重叠双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至外显子1-21的长于1kb的Myo15启动子(例如,SEQ ID NO:36)和紧邻编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子21/22边界3'的500bp;和第二核酸载体,所述第二核酸载体含有紧邻编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子21/22边界5'的500bp和其余外显子(例如,小鼠OTOF的外显子29-48,人类OTOF的外显子22-45和47或外显子22-46)以及poly(A)序列(例如,bGH poly(A)信号序列)。外显子20/21边界也可被选择作为重叠区域的中心。在此类重叠双载体系统中,第一核酸载体和第二核酸载体都不可以包括OTOF UTR。短启动子(例如,CMV启动子、CAG启动子、smCBA启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动

子)也可用于这些双载体系统(例如,其中重叠区域以外显子21/22或外显子20/21边界为中心的双载体系统)中。如果使用短启动子,则可在第一载体中包括另外的元件,诸如5'OTOF UTR(例如,载体含有编码OTOF蛋白的多核苷酸的外显子1-21和紧邻外显子21/22边界3'的500bp,或外显子1-20和紧邻外显子20/21边界3'的500bp)。

[0251] 反式剪接双载体

[0252] 用于在哺乳动物细胞中表达大蛋白的第二方法涉及使用反式剪接双载体。在这种方法中,使用含有不同核酸序列的两种核酸载体,并且编码目标蛋白质的N末端部分的多核苷酸和编码目标蛋白质的C末端部分的多核苷酸并不重叠。相反,第一核酸载体包括编码目标蛋白质的N末端部分的多核苷酸3'的剪接供体序列,并且第二核酸载体包括编码目标蛋白质的C末端部分的多核苷酸5'的剪接受体序列。当第一核酸和第二核酸存在于同一细胞中时,它们的ITR可多联体化,从而形成单一核酸结构,其中多联体化的ITR定位于剪接供体与剪接受体之间。然后在转录期间发生反式剪接,从而产生其中编码目标蛋白质的N末端部分和C末端部分的多核苷酸是邻接核酸分子,从而形成全长编码序列。

[0253] 用于本文所述的方法和组合物的反式剪接双载体经设计,使得大约一半的OTOF基因包含在每个载体内(例如,每个载体含有编码大约一半OTOF蛋白的多核苷酸)。基于启动子的大小和编码OTOF C2结构域的多核苷酸部分的位置来确定如何将多核苷酸序列分流于两种核酸载体之间。当在反式剪接双载体系统中使用短启动子(例如,1kb或更短,例如大约1kb、950bp、900bp、850bp、800bp、750bp、700bp、650bp、600bp、550bp 500bp、450bp、400bp、350bp、300bp或更短的启动子),诸如CAG启动子、CMV启动子、smCBA启动子或具有1kb或更短的序列的Myo15启动子(例如,上文所述的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)时,可在外显子边界处将OTOF多核苷酸序列分在两种核酸载体之间,所述外显子边界出现在编码C2D结构域的多核苷酸部分之后和编码C2E结构域的多核苷酸部分之前,例如外显子26/27边界。含有此大小的启动子的核酸载体可任选地含有OTOF UTR(例如,5'和3'OTOF UTR,例如全长UTR)。当在反式剪接双载体系统中使用长启动子(例如,长于1kb,例如,1.1kb、1.25kb、1.5kb、1.75kb、2kb、2.5kb、3kb或更长的启动子),诸如长于1kb的Myo15启动子(例如SEQ ID NO:36)时,可在外显子边界处将OTOF多核苷酸序列分在两种核酸载体之间,所述外显子边界出现在编码C2C结构域的多核苷酸部分之后和编码C2D结构域的多核苷酸部分之前(例如外显子19/20边界、外显子20/21边界或外显子21/22边界),或在编码C2D结构域的多核苷酸部分之内(例如外显子25/26边界)。短启动子(例如,CMV启动子、smCBA启动子、CAG启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的双载体系统中,在所述情况下可在第一载体(例如,含有编码C2C结构域的多核苷酸部分的载体)中包括额外的元件(例如,OTOF UTR序列)。

[0254] 使用短启动子的一种示例性反式剪接双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-26的CAG启动子和多核苷酸序列3'的剪接供体序列;和第二核酸载体,所述第二核酸载体含有编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子27-48,或人类OTOF的外显子27-45和47或外显子27-46)5'

的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列)。使用短启动子的一种示例性反式剪接双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-28的CAG启动子和多核苷酸序列3'的剪接供体序列;和第二核酸载体,所述第二核酸载体含有编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子29-48,或人类OTOF的外显子29-45和47或外显子29-46)5'的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列)。可在前述双载体系统中的任一者中使用CMV启动子、smCBA启动子或具有1kb或更短的序列的Myo15启动子(例如,上文所述的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)来代替CAG启动子。这些核酸载体还可分别在第一核酸载体和第二核酸载体中含有全长5'和3'OTOF UTR(例如,第一核酸载体可在编码人类OTOF的双载体系统中含有5'人类OTOF UTR(127bp),或在编码小鼠OTOF的双载体系统中含有5'小鼠UTR(134bp);并且第二核酸载体可在编码人类OTOF的双载体系统中含有3'人类OTOF UTR(1035bp),或在编码小鼠OTOF的双载体系统中含有3'小鼠OTOF UTR(1001bp)。

[0255] 使用长启动子的示例性反式剪接双载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-19的长于1kb的Myo15启动子和多核苷酸序列3'的剪接供体序列;和第二核酸载体,所述第二核酸载体含有编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子20-48,或人类OTOF的外显子20-45和47或外显子20-46)5'的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列)。或者,反式剪接双载体系统可包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-20的长于1kb的Myo15启动子和多核苷酸序列3'的剪接供体序列;和第二核酸载体,所述第二核酸载体含有编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子21-48,或人类OTOF的外显子21-45和47或外显子21-46)5'的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列)。前述Myo15启动子反式剪接双载体系统中任一者中的第一核酸载体和第二核酸载体都不含OTOF UTR。短启动子(例如,CMV启动子、smCBA启动子、CAG启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的前述双载体系统中。如果这些双载体系统含有短启动子,则其也可在第一载体中包括5'OTOF UTR或类似大小的另一元件。

[0256] 为适应OTOF UTR,OTOF编码序列可在不同位置中分开。例如,在反式剪接双载体系统中,其中第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-25的长于1kb的Myo15启动子和多核苷酸序列3'的剪接供体序列;并且第二核酸载体含有编码OTOF蛋白(例如,人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子26-48,或人类OTOF的外显子26-45和47或26-

46)5'的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列),第二核酸也可含有全长OTOF 3'UTR(例如,1035bp人类OTOF 3'UTR)。对于小鼠OTOF,如果第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-24的长于1kb的Myo15启动子(例如,SEQ ID NO:36)和多核苷酸序列3'的剪接供体序列;并且第二核酸载体含有编码OTOF蛋白(例如,小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子25-48 5'的剪接受体序列和poly(A)序列(例如,bGH poly(A)信号序列),则反式剪接双载体系统也可含有3'UTR。在这种双载体系统中,第二核酸还可含有全长OTOF 3'UTR(例如1001bp小鼠OTOF 3'UTR)。短启动子(例如,CMV启动子、smCBA启动子、CAG启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的前述双载体系统中。如果这些双载体系统含有短启动子,则其也可在第一载体中包括5'OTOF UTR。

[0257] 双杂合载体

[0258] 用于在哺乳动物细胞中表达大蛋白的第三方法涉及使用双杂合载体。这种方法结合了重叠双载体策略和反式剪接策略的元件,因为它的特征在于可进行同源重组的重叠区域以及剪接供体序列和剪接受体序列。在双杂合载体系统中,重叠区域是包含在第一核酸载体和第二核酸载体中的重组基因区,而不是编码目标蛋白质的多核苷酸序列的一部分-在这种方法中编码目标蛋白质的N末端部分的多核苷酸和编码目标蛋白质的C末端部分的多核苷酸并不重叠。重组基因区是第一核酸载体中的剪接供体序列的3'和第二核酸载体中的剪接受体序列的5'。第一和第二多核苷酸序列则可基于两种机制中的一者接合形成单一序列:1)在重叠区域重组,或2)ITR的多联体化。可通过剪接去除剩余重组基因区和/或多联体化ITR,从而形成编码目标全长蛋白的邻接多核苷酸序列。

[0259] 可用于本文所述的组合物和方法中的重组基因区包括具有以下序列的F1噬菌体AK基因:GGGATTTTGCCGATTTCGGCCTATTGG TTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA ACAAAAT(SEQ ID NO:19),和如美国专利第8,236,557号中所述的碱性磷酸酶(AP)基因片段,所述美国专利通过引用并入本文。在一些实施方案中,AP基因片段具有以下序列:

[0260] CCCCGGGTGCGCGGCGTCGGTGGTGCCGGCGGGGGGCGCCAGGTCGCAGGCGGTGTAGGGCTCCAGGCAGGCGGCGAAGGCCATGACGTGCGCTATGAAGGTCTGCTCCTGCACGCCGTGAACCAGGTGCGCCTGCGGGCCGCGCGCGAACACCGCCACGTCCTCGCCTGCGTGGGTCTCTTCGTCCAGGGGCACTGCTGACTGCTGCCGATACTCGGGGCTCCCGCTCTCGCTCTCGGTAACATCCGGCCGGGCGCCGTCCTTGAGCACATAGCCTGGACCGTTTCCGTATAGGAGGACCGTGTAGGCCTTCCTGTCCCGGGCCTTGCCAGCGGCCAGCCCGATGAAGGAGCTCCCTCGCAGGGGGTAGCCTCCGAAGGAGAAGACGTGGGAGTGGTCGGCAGTGACGAGGCTCAGCGTGTCCTCCTCGCTGGTGAGCTGGCCCGCCCTCTCAATGGCGTCGTCGAACATGATCGTCTCAGTCAGTGCCCGGTAAGCCCTGCTTTCATGATGACCATGGTCGATGCGACCACCCTCCACGAAGAGGAAGAAGCCGCGGGGGGTGTCTGCTCAGCAGGCGCAGGGCAGCCTCTGTCATCTCCATCAGGGAGGGGTCCAGTGTGGAGTCTCGGTGGATCTCGTATTTCATGTCTCCAGGCTCAAAGAGACCCATGAGATGGGTCACAGACGGGTCCAGGGAAGCCTGCATGAGCTCAGTGCGGTTCCACACGTACCGGGCACCCTGGCGTTCGCCGAGCCATTCCTGCACCAGATTCTTCCCGTCCAGCCTGGTCCCACCTTGGCTGTAGTCATCTGGGTACTCAGGGTCTGGGGTTCCCATGCGAAACATGTACTTTCGGCCTCCA(SEQ ID NO:62)。

[0261] 在一些实施方案中,AP基因片段具有以下序列:

[0262] CCCCGGGTGCGCGGCGTCGGTGGTGCCGGCGGGGGGCGCCAGGTCGCAGGCGGTGTAGGGCTCCAGGC

AGGCGGCGAAGGCCATGACGTGCGCTATGAAGGTCTGCTCCTGCACGCCGTGAACCAGGTGCGCCTGCGGGCCGCG CGCGAACACCGCCACGTCCTCGCCTGCGTGGGTCTCTTCGTCCAGGGGCACTGCTGACTGCTGCCGATACTCGGGG CTCCCGCTCTCGCTCTCGGTAACATCCGGCCGGGCGCCGTCCTTGAGCACATAGCCTGGACCGTTTCCGTATAGGA GGACCGTGTAGGCCTTCCTGTCCCGGGCCTTGCCAGCGGCCAGCCCGATGAAGGAGCTCCCTCGCAGGGGGTAGCC TCCGAAGGAGAAGACGTGGGAGTGGTCGGCAGTGACGAGGCTCAGCGTGTCCTCCTCG CTGGTGA(SEQ ID NO: 63)。

[0263] 在一些实施方案中,AP基因片段具有以下序列:

[0264] GCTGGCCCGCCCTCTCAATGGCGTCGTCGAACATGATCGTCTCAGTCAGTGCCCGGTAAGCCCTGCTT TCATGATGACCATGGTCGATGCGACCACCCTCCACGAAGAGGAAGAAGCCGCGGGGGTGTCTGCTCAGCAGGCGCA GGGCAGCCTCTGTCATCTCCATCAGGGAGGGGTCCAGTGTGGAGTCTCGGTGGATCTCGTATTTCATGTCTCCAGG CTCAAAGAGACCCATGAGATGGGTCACAGACGGGTCCAGGGAAGCCTGCATGAGCTCAGTGCGGTTCCACACGTAC CGGGCACCCTGGCGTTCGCCGAGCCATTCCTGCACCAGATTCTTCCCGTCCAGCCTGGTCCCACCTTGGCTGTAGT CATCTGGGTACTCAGGGTCTGGGGTTCCCATGCGAAACATGTACTTTCGGCCTCCA(SEQ ID NO:64)。

[0265] 在一些实施方案中,AP基因片段具有以下序列:

[0266] CCCCGGGTGCGCGGCGTCGGTGGTGCCGGCGGGGGGCGCCAGGTCGCAGGCGGTGTAGGGCTCCAGGC AGGCGGCGAAGGCCATGACGTGCGCTATGAAGGTCTGCTCCTGCACGCCGTGAACCAGGTGCGCCTGCGGGCCGCG CGCGAACACCGCCACGTCCTCGCCTGCGTGGGTCTCTTCGTCCAGGGGCACTGCTGACTGCTGCCGATACTCGGGG CTCCCGCTCTCGCTCTCGGTAACATCCGGCCGGGCGCCGTCCTTGAGCACATAGCCTGGACCGTTTC(SEQ ID NO:65)在一些实施方案中,AP基因片段具有以下序列:

[0267] CGTATAGGAGGACCGTGTAGGCCTTCCTGTCCCGGGCCTTGCCAGCGGCCAGCCCGATGAAGGAGCTC CCTCGCAGGGGGTAGCCTCCGAAGGAGAAGACGTGGGAGTGGTCGGCAGTGACGAGGCTCAGCGTGTCCTCCTCGC TGGTGAGCTGGCCCGCCCTCTCAATGGCGTCGTCGAACATGATCGTCTCAGTCAGTGCCCGGTAAGCCCTGCTTTC ATGATGACCATGGTCGATGCGACCACCCTCCACGAAGAGGAAGAAGCCGCGGGGGTGTCTGCTCAGCAGG(SEQ ID NO:66)。

[0268] 在一些实施方案中,AP基因片段具有以下序列:

[0269] CGCAGGGCAGCCTCTGTCATCTCCATCAGGGAGGGGTCCAGTGTGGAGTCTCGGTGGATCTCGTATTT CATGTCTCCAGGCTCAAAGAGACCCATGAGATGGGTCACAGACGGGTCCAGGGAAGCCTGCATGAGCTCAGTGCGG TTCCACACGTACCGGGCACCCTGGCGTTCGCCGAGCCATTCCTGCACCAGATTCTTCCCGTCCAGCCTGGTCCCAC CTTGGCTGTAGTCATCTGGGTACTCAGGGTCTGGGGTTCCCATGCGAAACATGTACTTTCGGCCTCCA(SEQ ID NO:67)。

[0270] 用于本文所述的方法和组合物中的示例性剪接供体序列(例如,在反式剪接和双杂合方法中)具有以下序列:AGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTT GCGTTTCT(SEQ ID NO:20)。用于本文所述的方法和组合物中的示例性剪接受体序列(例如,在反式剪接和双杂合方法中)具有以下序列:GATAGGCACCTATTGGTCTTACTGACATCCACTTTGCC TTTCTCTCCACAG(SEQ ID NO:21)。剪接供体序列GTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGAC CAATAGAAACTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGA(SEQ ID NO:68)和剪接受体序列 TAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCACAG(SEQ ID NO:69)也可用于本文所述的方法和组合物中。剪接供体和剪接受体序列的其他实例是本领域中已知的。

[0271] 用于本文所述方法和组合物中的双杂合载体经设计,使得大约一半的OTOF基因含

于每个载体内(例如,每个载体含有编码大约一半的OTOF蛋白的多核苷酸)。基于启动子的大小和编码OTOF C2结构域的多核苷酸部分的位置来确定如何将多核苷酸序列分流于两种核酸载体之间。当将短启动子用于双杂合载体系统中时(例如1kb或更短、例如大约1kb、950bp、900bp、850bp、800bp、750bp、700bp、650bp、600bp、550bp、500bp、450bp、400bp、350bp、300bp或更短的启动子),例如CAG、CMV、smCBA或具有1kb或更短的序列的Myo15启动子(例如上文所述的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子),在外显子边界处将OTOF多核苷酸序列分在两种核酸载体之间,所述外显子边界出现在编码C2D结构域的多核苷酸部分之后和编码C2E结构域的多核苷酸部分之前,例如外显子26/27边界。含有此大小的启动子的核酸载体可任选地含有OTOF UTR(例如全长5'和3'UTR)。当将长启动子用于双杂合载体系统中时(例如长于1kb,例如1.1kb、1.25kb、1.5kb、1.75kb、2kb、2.5kb、3kb或更长的启动子),例如长于1kb的Myo15启动子(例如SEQ ID NO:36),将在外显子边界处将OTOF多核苷酸序列分在两种核酸载体之间,所述外显子边界出现在编码C2C结构域的多核苷酸部分之后和编码C2D结构域的多核苷酸部分之前(例如外显子19/20边界、外显子20/21边界或外显子21/22边界),或在编码C2D结构域的多核苷酸部分之内(例如外显子25/26边界)。短启动子(例如CMV启动子、CAG启动子、smCBA启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的双载体系统中,在所述情况下可使额外元件(例如OTOF UTR序列)包括在第一载体(例如含有编码C2C结构域的多核苷酸部分的载体)中。

[0272] 使用短启动子的一种示例性双杂合载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-26的CAG启动子、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;和第二核酸载体,所述第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子27-48,或人类OTOF的外显子27-45和47或外显子27-46)的多核苷酸和poly(A)序列(例如bGH poly(A)信号序列)。第一核酸载体和第二核酸载体也可分别含有全长5'和3'OTOF UTR(例如,127bp人类OTOF 5'UTR可包括在第一核酸载体中,并且1035bp人类OTOF 3'UTR可包括在第二核酸载体中)。使用短启动子的另一种示例性双杂合载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-28的CAG启动子、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;和第二核酸载体,所述第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的29-48,或人类OTOF的外显子29-45和47或外显子29-46)的多核苷酸和poly(A)序列(例如bGH poly(A)信号序列)。第一核酸载体和第二核酸载体也可分别含有全长5'和3'OTOF UTR(例如,134bp小鼠OTOF 5'UTR可包括在第一核酸载体中,并且1001bp小鼠OTOF 3'UTR可包括在第二核酸载体中)。可在前述双载体系统中的任一者中使用CMV启动子、smCBA启动子或具有1kb或更

短的序列的Myo15启动子(例如,上文所述的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)来代替CAG启动子。

[0273] 使用长启动子的一种示例性双杂合载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-19的长于1kb的Myo15启动子(例如SEQ ID NO:36)、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;和第二核酸载体,所述第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的20-48外显子,或人类OTOF的外显子20-45和47或20-46)的多核苷酸和poly(A)序列(例如bGH poly(A)信号序列)。使用长启动子的另一示例性双杂合载体系统包括第一核酸载体,所述第一核酸载体含有可操作地连接至编码OTOF蛋白(例如小鼠OTOF,例如SEQ ID NO:6,或人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5)的多核苷酸的外显子1-20的长于1kb的Myo15启动子(例如SEQ ID NO:36)、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;和第二核酸载体,所述第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如小鼠OTOF,例如SEQ ID NO:6,或人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子20-48,或人类OTOF的外显子21-45和47或21-46)的多核苷酸和poly(A)序列(例如bGH poly(A)信号序列)。前述Myo15启动子双杂合载体系统中任一者中的第一核酸载体和第二核酸载体都不含OTOF UTR。短启动子(例如CMV启动子、smCBA启动子、CAG启动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的前述双载体系统中。如果这些双载体系统含有短启动子,则其也可在第一载体中包括另外的元件(例如5'OTOF UTR)。

[0274] 为适应OTOF UTR,OTOF编码序列可在不同位置中分开。例如,在双杂合载体系统中,其中第一核酸载体含有可操作地连接至编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-25的长于1kb的Myo15启动子(例如SEQ ID NO:36)、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;并且第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5,或小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的其余外显子(例如,小鼠OTOF的外显子26-48,或人类OTOF的外显子26-45和47或外显子26-46)的多核苷酸和poly(A)序列(例如bGH poly(A)信号序列),第二核酸也可含有全长OTOF 3'UTR(例如1035bp人类OTOF UTR)。对于小鼠OTOF,如果第一核酸载体含有可操作地连接至编码OTOF蛋白(例如,小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子1-24的长于1kb的Myo15启动子(例如,SEQ ID NO:36)、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的重组基因区;并且第二核酸载体含有重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF蛋白(例如小鼠OTOF,例如SEQ ID NO:6)的多核苷酸的外显子25-48的多核苷酸和poly(A)序列(例如,bGH poly(A)信号序列),则双杂合载体系统可含有3'UTR。在这种双杂合载体系统中,第二核酸还可含有全长OTOF 3'UTR(例如1001bp小鼠OTOF UTR)。短启动子(例如CMV启动子、smCBA启动子、CAG启

动子或具有1kb或更短的序列的Myo15启动子,例如具有SEQ ID NO:38、39或49-60中任一者的序列的Myo15启动子)也可用于经设计用于大启动子的前述双载体系统中。如果这些双载体系统含有短启动子,则其也可在第一载体中包括另外的元件(例如5'OTOF UTR)。

[0275] 本文所述的方法和组合物中使用的双杂合载体可任选地在第一核酸载体和第二核酸载体两者中包括降解信号序列。可包括降解信号序列以防止或减少来自未能重组和/或经历剪接的多核苷酸的OTOF蛋白部分的表达。降解信号序列定位于第一核酸载体中的重组基因区的3',并且定位于第二核酸载体中的重组基因区与剪接受体之间。可用于本文所述的组合物和方法中的降解信号序列具有以下序列:GCCTGCAAGAACTGGTTCAGCAGCCTGAGCCACTTCGTGATCCACCTG(SEQ ID NO:22)。

[0276] 示例性重叠、反式剪接和双杂合载体对描述于下表4中。

[0277] 表4:用于本文所述的组合物和方法中的示例性重叠、反式剪接和杂合双载体对

[0278]

载体对编号	载体类型	载体对
1	重叠	第一核酸载体含有:可操作地连接至外显子 1-24 的 CAG 启动子和编码人类 OTOF 蛋白的多核苷酸的外显子 24/25 边界 3'的 500 bp
		第二核酸载体含有:编码人类 OTOF 蛋白的多核苷酸的外显子 24/25 边界 5'的 500 bp 和其余外显子以及 bGH poly(A) 序列

[0279]

载体对编号	载体类型	载体对
2	重叠	第一核酸载体含有：可操作地连接至外显子 1-28 的 CAG 启动子和编码人类 OTOF 蛋白的多核苷酸的外显子 28/29 边界 3’的 500 bp
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的外显子 28/29 边界 5’的 500 bp 和其余外显子以及 bGH poly(A)序列
3	重叠	第一核酸载体含有：可操作地连接至外显子 1-21 的 Myo15 启动子和编码人类 OTOF 蛋白的多核苷酸的外显子 21/22 边界 3’的 500 bp
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的外显子 21/22 边界 5’的 500 bp 和其余外显子以及 bGH poly(A)序列
4	重叠	第一核酸载体含有：可操作地连接至外显子 1-20 的 Myo15 启动子和编码人类 OTOF 蛋白的多核苷酸的外显子 20/21 边界 3’的 500 bp
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的外显子 20/21 边界 5’的 500 bp 和其余外显子以及 bGH poly(A)序列
5	反式剪接	第一核酸载体含有：可操作地连接至编码人类 OTOF 蛋白的多核苷酸的外显子 1-26 的 CAG 启动子和多核苷酸 3’的剪接供体序列
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的其余外显子 5’的剪接受体序列和 bGH poly(A)序列
6	反式剪接	第一核酸载体含有：可操作地连接至编码人类 OTOF 蛋白的多核苷酸的外显子 1-19 的 Myo15 启动子和多核苷酸 3’的剪接供体序列
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的其余外显子 5’的剪接受体序列和 bGH poly(A)序列
7	反式剪接	第一核酸载体含有：可操作地连接至编码人类 OTOF 蛋白的多核苷酸的外显子 1-25 的 Myo15 启动子和多核苷酸 3’的剪接供体序列
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的其余外显子 5’的剪接受体序列和 bGH poly(A)序列
8	反式剪接	第一核酸载体含有：可操作地连接至编码人类 OTOF 蛋白的多核苷酸的外显子 1-20 的 Myo15 启动子和多核苷酸 3’的剪接供体序列
		第二核酸载体含有：编码人类 OTOF 蛋白的多核苷酸的其余外显子 5’的剪接受体序列和 bGH poly(A)序列
9	杂合	第一核酸载体含有：可操作地连接至编码人类 OTOF 蛋白的多核苷酸的外显子 1-26 的 CAG 启动子、多核苷酸 3’的剪接供体序列和剪接供体序列 3’的重组基因区
		第二核酸载体含有：重组基因区、重组基因区 3’的剪接受

[0280]

载体对编号	载体类型	载体对
		体序列、剪接受体序列3'的编码人类OTOF蛋白的多核苷酸的其余外显子和bGH poly(A)序列
10	杂合	第一核酸载体含有：可操作地连接至编码人类OTOF蛋白的多核苷酸的外显子1-19的Myo15启动子、多核苷酸3'的剪接供体序列和剪接供体序列3'的重组基因区
		第二核酸载体含有：重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的编码人类OTOF蛋白的多核苷酸的其余外显子和bGH poly(A)序列
11	杂合	第一核酸载体含有：可操作地连接至编码人类OTOF蛋白的多核苷酸的外显子1-25的Myo15启动子、多核苷酸3'的剪接供体序列和剪接供体序列3'的重组基因区
		第二核酸载体含有：重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的编码人类OTOF蛋白的多核苷酸的其余外显子和bGH poly(A)序列
12	杂合	第一核酸载体含有：可操作地连接至编码人类OTOF蛋白的多核苷酸的外显子1-20的Myo15启动子、多核苷酸3'的剪接供体序列和剪接供体序列3'的重组基因区
		第二核酸载体含有：重组基因区、重组基因区3'的剪接受体序列、剪接受体序列3'的编码人类OTOF蛋白的多核苷酸的其余外显子和bGH poly(A)序列

[0281] 在一些实施方案中，编码OTOF蛋白的多核苷酸序列为cDNA序列(例如不包括内含子的序列)。在一些实施方案中，双载体系统中的第一核酸载体和/或第二核酸载体可包括内含子序列。内含子序列可包括在OTOF编码序列中的一个或多个外显子之间，或者内含子序列可包括在编码序列的外显子与核酸载体的另一组分之间(例如，第一核酸载体中的OTOF编码序列的外显子与剪接供体序列之间或第二核酸载体中的OTOF编码序列的外显子与剪接受体序列之间)。

[0282] 在一些实施方案中，编码OTOF蛋白的多核苷酸在外显子20/21边界处分在双载体系统中的第一核酸载体(例如AAV载体)与第二核酸载体(例如AAV载体)之间。当编码OTOF蛋白的多核苷酸编码OTOF同功型5并且在外显子20/21边界处分在第一核酸载体(例如AAV载体)与第二核酸载体(例如AAV载体)之间时，编码OTOF的N末端部分的多核苷酸序列具有以下序列：

[0283] ATGGCCTTGCTCATCCACCTCAAGACAGTCTCGGAGCTGCGGGGCAGGGGCGACCGGATCGCCAAAGT
GACTTTCCGAGGGCAATCCTTCTACTCTCGGGTCCTGGAGAACTGTGAGGATGTGGCTGACTTTGATGAGACATTT
CGGTGGCCGGTGGCCAGCAGCATCGACAGAAATGAGATGCTGGAGATTCAGGTTTTCAACTACAGCAAAGTCTTCA
GCAACAAGCTCATCGGGACCTTCCGCATGGTGCTGCAGAAGGTGGTAGAGGAGAGCCATGTGGAGGTGACTGACAC
GCTGATTGATGACAACAATGCTATCATCAAGACCAGCCTGTGCGTGGAGGTCCGGTATCAGGCCACTGACGGCACA
GTGGGCTCCTGGGACGATGGGGACTTCCTGGGAGATGAGTCTCTTCAAGAGGAAGAGAAGGACAGCCAAGAGACGG
ATGGACTGCTCCCAGGCTCCCGGCCCAGCTCCCGGCCCCCAGGAGAGAAGAGCTTCCGGAGAGCCGGGAGGAGCGT
GTTCTCCGCCATGAAGCTCGGCAAAAACCGGTCTCACAAGGAGGAGCCCCAAAGACCAGATGAACCGGCGGTGCTG
GAGATGGAAGACCTTGACCATCTGGCCATTCGGCTAGGAGATGGACTGGATCCCGACTCGGTGTCTCTAGCCTCAG
TCACAGCTCTCACCACTAATGTCTCCAACAAGCGATCTAAGCCAGACATTAAGATGGAGCCAAGTGCTGGGCGGCC

CATGGATTACCAGGTCAGCATCACGGTGATCGAGGCCCGGCAGCTGGTGGGCTTGAACATGGACCCTGTGGTGTGC
GTGGAGGTGGGTGACGACAAGAAGTACACATCCATGAAGGAGTCCACTAACTGCCCCTATTACAACGAGTACTTCG
TCTTCGACTTCCATGTCTCTCCGGATGTCATGTTTGACAAGATCATCAAGATTTCGGTGATTCACTCCAAGAACCT
GCTGCGCAGTGGCACCCTGGTGGGCTCCTTCAAAATGGACGTGGGAACCGTGTACTCGCAGCCAGAGCACCAGTTC
CATCACAAGTGGGCCATCCTGTCTGACCCCGATGACATCTCCTCGGGGCTGAAGGGCTACGTGAAGTGTGACGTTG
CCGTGGTGGGCAAAGGGGACAACATCAAGACGCCCCACAAGGCCAATGAGACCGACGAAGATGACATTGAGGGGAA
CTTGCTGCTCCCCGAGGGGGTGCCCCCCGAACGCCAGTGGGCCCGGTTCTATGTGAAAATTTACCGAGCAGAGGGG
CTGCCCCGTATGAACACAAGCCTCATGGCCAATGTAAAGAAGGCTTTCATCGGTGAAAACAAGGACCTCGTGGACC
CCTACGTGCAAGTCTTCTTTGCTGGCCAGAAGGGCAAGACTTCAGTGCAGAAGAGCAGCTATGAGCCCCTGTGGAA
TGAGCAGGTCGTCTTTACAGACCTCTTCCCCCCACTCTGCAAACGCATGAAGGTGCAGATCCGAGACTCGGACAAG
GTCAACGACGTGGCCATCGGCACCCACTTCATTGACCTGCGCAAGATTTCTAATGACGGAGACAAAGGCTTCCTGC
CCACACTGGGCCCAGCCTGGGTGAACATGTACGGCTCCACACGTAACTACACGCTGCTGGATGAGCATCAGGACCT
GAACGAGGGCCTGGGGGAGGGTGTGTCCTTCCGGGCCCGGCTCCTGCTGGGCCTGGCTGTGGAGATCGTAGACACC
TCCAACCCTGAGCTCACCAGCTCCACAGAGGTGCAGGTGGAGCAGGCCACGCCCATCTCGGAGAGCTGTGCAGGTA
AAATGGAAGAATTCTTTCTCTTTGGAGCCTTCCTGGAGGCCTCAATGATCGACCGGAGAAACGGAGACAAGCCCAT
CACCTTTGAGGTCACCATAGGCAACTATGGGAACGAAGTTGATGGCCTGTCCCGGCCCCAGCGGCCTCGGCCCCGG
AAGGAGCCGGGGGATGAGGAAGAAGTAGACCTGATTCAGAACGCAAGTGATGACGAGGCCGGTGATGCCGGGGACC
TGGCCTCAGTCTCCTCCACTCCACCAATGCGGCCCCAGGTCACCGACAGGAACTACTTCCATCTGCCCTACCTGGA
GCGAAAGCCCTGCATCTACATCAAGAGCTGGTGGCCGGACCAGCGCCGCCGCCTCTACAATGCCAACATCATGGAC
CACATTGCCGACAAGCTGGAAGAAGGCCTGAACGACATACAGGAGATGATCAAAACGGAGAAGTCCTACCCTGAGC
GTCGCCTGCGGGGCGTCCTGGAGGAGCTGAGCTGTGGCTGCTGCCGCTTCCTCTCCCTCGCTGACAAGGACCAGGG
CCACTCATCCCGCACCAGGCTTGACCGGGAGCGCCTCAAGTCCTGCATGAGGGAGCTG(SEQ ID NO:71)。

[0284] 上述序列也对应于OTOF同功型1的外显子1-20。

[0285] 当编码OTOF蛋白的多核苷酸编码OTOF同功型5并且在外显子20/21边界处分在第一核酸载体(例如AAV载体)与第二核酸载体(例如AAV载体)之间时,编码OTOF的C末端部分的多核苷酸序列具有以下序列:

[0286] GAAAACATGGGGCAGCAGGCCAGGATGCTGCGGGCCCAGGTGAAGCGGCACACGGTGCGGGACAAGCT
GAGGCTGTGCCAGAACTTCCTGCAGAAGCTGCGCTTCCTGGCGGACGAGCCCCAGCACAGCATTCCCGACATCTTC
ATCTGGATGATGAGCAACAACAAGCGTGTCGCCTATGCCCGTGTGCCCTCCAAGGACCTGCTCTTCTCCATCGTGG
AGGAGGAGACTGGCAAGGACTGCGCCAAGGTCAAGACGCTCTTCCTTAAGCTGCCAGGGAAGCGGGGCTTCGGCTC
GGCAGGCTGGACAGTGCAGGCCAAGGTGGAGCTGTACCTGTGGCTGGGCCTCAGCAAACAGCGCAAGGAGTTCCTG
TGCGGCCTGCCCTGTGGCTTCCAGGAGGTCAAGGCAGCCCAGGGCCTGGGCCTGCATGCCTTCCCACCCGTCAGCC
TGGTCTACACCAAGAAGCAGGCGTTCCAGCTCCGAGCGCACATGTACCAGGCCCGCAGCCTCTTTGCCGCCGACAG
CAGCGGACTCTCAGACCCCTTTGCCCGCGTCTTCTTCATCAATCAGAGTCAGTGCACAGAGGTGCTGAATGAGACC
CTGTGTCCCACCTGGGACCAGATGCTGGTGTTCGACAACCTGGAGCTCTATGGTGAAGCTCATGAGCTGAGGGACG
ATCCGCCCATCATTGTCATTGAAATCTATGACCAGGATTCCATGGGCAAAGCTGACTTCATGGGCCGGACCTTCGC
CAAACCCCTGGTGAAGATGGCAGACGAGGCGTACTGCCCACCCCGCTTCCCACCTCAGCTCGAGTACTACCAGATC
TACCGTGGCAACGCCACAGCTGGAGACCTGCTGGCGGCCTTCGAGCTGCTGCAGATTGGACCAGCAGGGAAGGCTG
ACCTGCCCCCCATCAATGGCCCGGTGGACGTGGACCGAGGTCCCATCATGCCCGTGCCCATGGGCATCCGGCCCGT

GCTCAGCAAGTACCGAGTGGAGGTGCTGTTCTGGGGCCTACGGGACCTAAAGCGGGTGAACCTGGCCCAGGTGGAC
CGGCCACGGGTGGACATCGAGTGTGCAGGGAAGGGGGTGCAGTCGTCCCTGATCCACAATTATAAGAAGAACCCCA
ACTTCAACACCCTCGTCAAGTGGTTTGAAGTGGACCTCCCAGAGAACGAGCTGCTGCACCCGCCCTTGAACATCCG
TGTGGTGGACTGCCGGGCCTTCGGTCGCTACACACTGGTGGGCTCCCATGCCGTCAGCTCCCTGCGACGCTTCATC
TACCGGCCCCCAGACCGCTCGGCCCCCAGCTGGAACACCACGGTCAGGCTTCTCCGGCGCTGCCGTGTGCTGTGCA
ATGGGGGCTCCTCCTCTCACTCCACAGGGGAGGTTGTGGTGACTATGGAGCCAGAGGTACCCATCAAGAAACTGGA
GACCATGGTGAAGCTGGACGCGACTTCTGAAGCTGTTGTCAAGGTGGATGTGGCTGAGGAGGAGAAGGAGAAGAAG
AAGAAGAAGAAGGGCACTGCGGAGGAGCCAGAGGAGGAGGAGCCAGACGAGAGCATGCTGGACTGGTGGTCCAAGT
ACTTTGCCTCCATTGACACCATGAAGGAGCAACTTCGACAACAAGAGCCCTCTGGAATTGACTTGGAGGAGAAGGA
GGAAGTGGACAATACCGAGGGCCTGAAGGGGTCAATGAAGGGCAAGGAGAAGGCAAGGGCTGCCAAAGAGGAGAAG
AAGAAGAAAACTCAGAGCTCTGGCTCTGGCCAGGGGTCCGAGGCCCCCGAGAAGAAGAAACCCAAGATTGATGAGC
TTAAGGTATACCCCAAAGAGCTGGAGTCCGAGTTTGATAACTTTGAGGACTGGCTGCACACTTTCAACTTGCTTCG
GGGCAAGACCGGGGATGATGAGGATGGCTCCACCGAGGAGGAGCGCATTGTGGGACGCTTCAAGGGCTCCCTCTGC
GTGTACAAAGTGCCACTCCCAGAGGACGTGTCCCGGGAAGCCGGCTACGACTCCACCTACGGCATGTTCCAGGGCA
TCCCGAGCAATGACCCCATCAATGTGCTGGTCCGAGTCTATGTGGTCCGGGCCACGGACCTGCACCCTGCTGACAT
CAACGGCAAAGCTGACCCCTACATCGCCATCCGGCTAGGCAAGACTGACATCCGCGACAAGGAGAACTACATCTCC
AAGCAGCTCAACCCTGTCTTTGGGAAGTCCTTTGACATCGAGGCCTCCTTCCCCATGGAATCCATGCTGACGGTGG
CTGTGTATGACTGGGACCTGGTGGGCACTGATGACCTCATTGGGGAAACCAAGATCGACCTGGAGAACCGCTTCTA
CAGCAAGCACCGCGCCACCTGCGGCATCGCCCAGACCTACTCCACACATGGCTACAATATCTGGCGGGACCCCATG
AAGCCCAGCCAGATCCTGACCCGCCTCTGCAAAGACGGCAAAGTGGACGGCCCCCACTTTGGGCCCCCTGGGAGAG
TGAAGGTGGCCAACCGCGTCTTCACTGGGCCCTCTGAGATTGAGGACGAGAACGGTCAGAGGAAGCCCACAGACGA
GCATGTGGCGCTGTTGGCCCTGAGGCACTGGGAGGACATCCCCCGCGCAGGCTGCCGCCTGGTGCCAGAGCATGTG
GAGACGAGGCCGCTGCTCAACCCCGACAAGCCGGGCATCGAGCAGGGCCGCCTGGAGCTGTGGGTGGACATGTTCC
CCATGGACATGCCAGCCCCTGGGACGCCTCTGGACATCTCACCTCGGAAGCCCAAGAAGTACGAGCTGCGGGTCAT
CATCTGGAACACAGATGAGGTGGTCTTGGAGGACGACGACTTCTTCACAGGGGAGAAGTCCAGTGACATCTTCGTG
AGGGGGTGGCTGAAGGGCCAGCAGGAGGACAAGCAGGACACAGACGTCCACTACCACTCCCTCACTGGCGAGGGCA
ACTTCAACTGGCGCTACCTGTTCCCCTTCGACTACCTGGCGGCGGAGGAGAAGATCGTCATCTCCAAGAAGGAGTC
CATGTTCTCCTGGGACGAGACCGAGTACAAGATCCCCGCGCGGCTCACCCTGCAGATCTGGGATGCGGACCACTTC
TCCGCTGACGACTTCCTGGGGGCCATCGAGCTGGACCTGAACCGGTTCCCGCGGGGCGCAAAGACAGCCAAGCAGT
GCACCATGGAGATGGCCACCGGGGAGGTGGACGTGCCCCTCGTGTCCATCTTCAAGCAAAAGCGCGTCAAAGGCTG
GTGGCCCCTCCTGGCCCGCAATGAGAACGATGAGTTTGAGCTCACGGGCAAGGTGGAGGCTGAGCTGCATTTACTG
ACAGCAGAGGAGGCAGAGAAGAACCCAGTGGGCCTGGCCCGCAATGAACCTGACCCCCTAGAGAAACCCAACCGGC
CCGACACGGCCTTCGTCTGGTTCCTCAACCCTCTCAAGTCCATCAAGTACCTCATCTGCACCCGGTACAAGTGGCT
CATCATCAAGATCGTGCTGGCGCTGTTGGGGCTGCTCATGTTGGGGCTCTTCCTCTACAGCCTCCCTGGCTACATG
GTCAAAAAGCTCCTTGGGGCATGA(SEQ ID NO:72)。

[0287] 在多核苷酸编码OTOF同功型5并且在外显子20/21边界处分在第一核酸载体(例如AAV载体)与第二核酸载体(例如AAV载体)之间的实施方案中,OTOF多肽的N末端部分具有以下序列:

[0288] MALLIHLKTVSELRGRGDRIAKVTFRGQSFYSRVLENCEDVADFDETFRWPVASSIDRNEMLEIQVFN

YSKVFSNKLIGTFRMVLQKVVEESHVEVTDTLIDDNNAIIKTSLCVEVRYQATDGTVGSWDDGDFLGDESLQEEEK
DSQETDGLLPGSRPSSRPPGEKSFRRAGRSVFSAMKLGKNRSHKEEPQRPDEPAVLEMEDLDHLAIRLGDGLDPDS
VSLASVTALTTNVSNKRSKPDIKMEPSAGRPMDYQVSITVIEARQLVGLNMDPVVCVEVGDDKKYTSMKESTNCPY
YNEYFVFDFHVSPDVMFDKIIKISVIHSKNLLRSGTLVGSFKMDVGTVYSQPEHQFHHKWAILSDPDDISSGLKGY
VKCDVAVVGKGDNIKTPHKANETDEDDIEGNLLLPEGVPPERQWARFYVKIYRAEGLPRMNTSLMANVKKAFIGEN
KDLVDPYVQVFFAGQKGKTSVQKSSYEPLWNEQVVFTDLFPPLCKRMKVQIRDSDKVNDVAIGTHFIDLRKISNDG
DKGFLPTLGPAWVNMYGSTRNYTLLDEHQDLNEGLGEGVSFRARLLLGLAVEIVDTSNPELTSSTEVQVEQATPIS
ESCAGKMEEFFLFGAFLEASMIDRRNGDKPITFEVTIGNYGNEVDGLSRPQRPRPRKEPGDEEEVDLIQNASDDEA
GDAGDLASVSSTPPMRPQVTDRNYFHLPYLERKPCIYIKSWWPDQRRRLYNANIMDHIADKLEEGLNDIQEMIKTE
KSYPERRLRGVLEELSCGCCRFLSLADKDQGHSSRTRLDRERLKSCMREL(SEQ ID NO:73)。

[0289] 上述序列也对应于由外显子1-20编码的OTOF同功型1蛋白的N末端部分。

[0290] 在多核苷酸编码OTOF同功型5并且在外显子20/21边界处分在第一核酸载体(例如AAV载体)与第二核酸载体(例如AAV载体)之间的实施方案中,OTOF多肽的C末端部分具有以下序列:

[0291] ENMGQQARMLRAQVKRHTVRDKLRLCQNFLQKLRFLADEPQHSIPDIFIWMMSNNKRVAYARVPSKDL
LFSIVEEETGKDCAKVKTLFLKLPGKRGFGSAGWTVQAKVELYLWLGLSKQRKEFLCGLPCGFQEVKAAQGLGLHA
FPPVSLVYTKKQAFQLRAHMYQARSLFAADSSGLSDPFARVFFINQSQCTEVLNETLCPTWDQMLVFDNLELYGEA
HELRDDPPIIVIEIYDQDSMGKADFMGRTFAKPLVKMADEAYCPPRFPPQLEYYQIYRGNATAGDLLAAFELLQIG
PAGKADLPPINGPVDVDRGPIMPVPMGIRPVLSKYRVEVLFWGLRDLKRVNLAQVDRPRVDIECAGKGVQSSLIHN
YKKNPNFNTLVKWFEVDLPENELLHPPLNIRVVDCRAFGRYTLVGSHAVSSLRRFIYRPPDRSAPSWNTTVRLLRR
CRVLCNGGSSSHSTGEVVVTMEPEVPIKKLETMVKLDATSEAVVKVDVAEEEKEKKKKKKGTAEEPEEEEPDESML
DWWSKYFASIDTMKEQLRQQEPSGIDLEEKEEVDNTEGLKGSMKGKEKARAAKEEKKKKTQSSGSGQGSEAPEKKK
PKIDELKVYPKELESEFDNFEDWLHTFNLLRGKTGDDEDGSTEEERIVGRFKGSLCVYKVPLPEDVSREAGYDSTY
GMFQGIPSNDPINVLVRVYVVRATDLHPADINGKADPYIAIRLGKTDIRDKENYISKQLNPVFGKSFDIEASFPME
SMLTVAVYDWDLVGTDDLIGETKIDLENRFYSKHRATCGIAQTYSTHGYNIWRDPMKPSQILTRLCKDGKVDGPHF
GPPGRVKVANRVFTGPSEIEDENGQRKPTDEHVALLALRHWEDIPRAGCRLVPEHVETRPLLNPDKPGIEQGRLEL
WVDMFPMDMPAPGTPLDISPRKPKKYELRVIIWNTDEVVLEDDDFFTGEKSSDIFVRGWLKGQQEDKQDTDVHYHS
LTGEGNFNWRYLFPFDYLAAEEKIVISKKESMFSWDETEYKIPARLTLQIWDADHFSADDFLGAIELDLNRFPRGA
KTAKQCTMEMATGEVDVPLVSIFKQKRVKGWWPLLARNENDEFELTGKVEAELHLLTAEEAEKNPVGLARNEPDPL
EKPNRPDTAFVWFLNPLKSIKYLICTRYKWLIIKIVLALLGLLMLGLFLYSLPGYMVKKLLGA(SEQ ID NO:
74)。

[0292] 可用于产生用于本文所述的组合物和方法中的核酸载体的转移质粒提供于表5中。这些转移质粒经设计用于表达OTOF同功型5。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)。核酸载体(例如含有编码OTOF的N末端部分的多核苷酸的核酸载体(例如AAV载体)和含有编码OTOF的C末端部分的多核苷酸的核酸载体(例如AAV载体))可在施用之前组合(例如于单一制剂中)。以下转移质粒经设计以产生在双杂合载体系统中共配制或共施用(例如同时或依序施

用)的核酸载体(例如AAV载体):SEQ ID NO:75和SEQ ID NO:76;SEQ ID NO:77和SEQ ID NO:78;SEQ ID NO:79和SEQ ID NO:76;SEQ ID NO:80和SEQ ID NO:78;SEQ ID NO:81和SEQ ID NO:82;以及SEQ ID NO:83和SEQ ID NO:82。

[0293] 表5:用于产生双杂合载体系统的转移质粒

[0294]

SEQ ID NO.	描述	质粒序列
75	含有以下特征的5'转基因质粒: 位置53-2027处的λ(衍生自噬菌体λ的生物插入物和不活化的DNA以减少脱靶DNA壳体化) 位置2049-2178处的ITR 位置2272-3236处的Myo15启动子 位置3253-3262处的Kozak序列(起始蛋白质翻译的位点) 位置3259-5664处的人类OTOF同功型5的N末端部分 位置5665-5748处	TCGCGCGTTTCGGTGATGACGGTGAAAACC TCTGACACATGCAGCTCCCGGATAGAGGTC ATCCTTCCTGACCATTTCCATCATTCCAGTC GAACTCACACACAACACCAAATGCATTTA AGTCGCTTGAAATTGCTATAAGCAGAGCAT GTTGCGCCAGCATGATTAATACAGCATTTA ATACAGAGCCGTGTTTATTGAGTCGGTATT CAGAGTCTGACCAGAAATTATTAATCTGGT GAAGTTATTCCTCTGTCATTACGTCATGGT CGATTTCAATTTCTATTGATGCTTTCCAGTC GTAATCAATGATGTATTTTTTGATGTTTGA CCTCTGTTCATATCCTCACAGATAAAAAAT CGCCCTCACACTGGAGGGCAAAGAAGATT TCCAATAATCAGAACAAGTCGGCTCCTGTT TAGTTACGAGCGACATTGCTCCGTGTATTC ACTCGTTGGAATGAATACACAGTGCAGTGT TTATTCTGTTATTTATGCCAAAAATTAAGG CCACTATCAGGCAGCTTTGTTGTTCTGTTTA CCAAGTTCTCTGGCAATCATTGCCGTCGTT CGTATTGCCCATTTATCGACATATTTCCCAT CTTCCTATACAGGAAACATTTCTTCAGGCT

[0295]

SEQ ID NO.	描述	质粒序列
	的剪接供体(APSD)序列 位置 5755-6041 处的 AP 头部序列(重组基因区) 位置 6135-6264 处的 ITR 位置 6275-8287 处的λ 位置 8344-8932 处的 Ori (复制起点) 位置 9110-9919 处的 KanR (抗生素抗性基因) 位置 2049-6264 处的待转移至双载体系统中的载体中的转基因	TAACCATGCATTCCGATTGCAGCTTGCATC CATTGCATCGCTTGAATTGTCCACACCATT GATTTTTATCAATAGTCGTAGTTTAACGGA TAGTCCTGGTATTGTTCCATCACATCCTGA GGATGCCCTTCGAACTCTTCAAATTCTTCTT CCTAATATCACCTTAAATAGTGGATTGCGG TAGTAAAGATTGTGCCTGTCTTTTAACCAC ATCAGGCTCGGTGGTTCTCGTGTACCCCTA CAGCGAGAAATCGGATAAACTATTACAAC CCCTACAGTTTGTAGAGTATAGAAAATGAT CCACTCGTTATTCTCGGACGAGTGTTCAGT AATGAACCTCTGGAGAGAACCATCTATATG ATCGTTATCTGGGTTTGACTTCTGCTTTTAA GCCCAGATAACTTGCCTGAATATGTTAATG AGAGAATCGGTATTCCTCATGTGTGGCATG TTTTCGTCTTTGCTCTTGCATTTTCACTAGC AATTAATGTGCATCGATTATCAGCTATTGC CAGCGCCAGATATAAGCGATTTAAGCTAA GAAAACGCATTAAGGTGCAAAACGATAAA GTGCGATCAGTAATTCAAAACCTTACAGGA GAGCAATCTATGGTTTTGTGCTCAGCCCTT AATGAAGGCAGGTAGTATGTGGTTACATC AAAACAATTCCCATACATTAGTGAGTTGAT TGAGCTTGGTGTGTTGAACAAAACTTTTTC CCGATGGAATGGAAAGCATATATTATTCCC TATTGAGGATATTTACTGGACTGAATTAGT TGCCAGCTATGATCCATATAATATTGAGAT AAAGCCAAGGCCAATATCTAAGTAACTAG ATAAGAGGAATCGATTTTCCCTTAATTTTC TGGCGTCCACTGCATGTTATGCCGCGTTCG CCAGGCTTGCTGTACCATGTGCGCTGATTC TTGCGCTCAATACGTTGCAGGTTGCTTTCA ATCTGTTTGTGGTATTCAGCCAGCACTGTA AGGTCTATCGGATTTAGTGCGCTTTCTACT CGTGATTTCGGTTTGCGATTCAGCGAGAGA ATAGGGCGGTTAACTGGTTTTGCGCTTACC CCAACCAACAGGGGATTTGCTGCTTTCCAT TGAGCCTGTTACTCTGCGCGACGTTCGCGG CGGCGTGTTTGTGCATCCATCTGGATTCTC CTGTCAGTTAGCTTTGGTGGTGTGTGGCAG TTGTAGTCCTGAACGAAAACCCCCCGCGAT TGGCACGTTGGCAGCTAATCCGGAATCGCA CTTACGGCCAATGCTTCGTTTCGTATCACA CACCCCAAAGCCTTCTGCTTTGAATGCTGC CCTTCTTCAGGGCTTAATTTTTAAGAGCGT CACCTTCATGGTGGTCAGTGCGTCCTGCTG ATGTGCTCAGGCACGATTTAATTAAGGCCT TAATTAGGCTGCGCGCTCGCTCGCTCACTG

[0296]

SEQ ID NO.	描述	质粒序列
		AGGCCGCCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAA CTCCATCACTAGGGGTTCCTTGTAGTTAAT GATTAACCCGCCATGCTACTTATCTACGTA GCCATGCTCTAGGAAGATCGGAATTCGCCC TTAAGCTAGCGGCGCGCCCAATTCTGCAGC TCAGCCTACTACTTGCTTTCCAGGCTGTTCC TAGTTCCCATGTCAGCTGCTTGTGCTTTCCA GAGACAAAACAGGAATAATAGATGTCATT AAATATACATTGGGCCCCAGGCGGTCAAT GTGGCAGCCTGAGCCTCCTTTCCATCTCTG TGGAGGCAGACATAGGACCCCCAACAAAC AGCATGCAGGTTGGGAGCCAGCCACAGGA CCCAGGTAAGGGGCCCTGGGTCCTTAAGCT TCTGCCACTGGCTCCGGCATTGCAGAGAGA AGAGAAGGGGCGGCAGACTGGAGAGCTGG GCTCCATTTTTGTTCCTTGGTGCCCTGCCCC TCCCCATGACCTGCAGAGACATTCAGCCTG CCAGGCTTTATGAGGTGGGAGCTGGGCTCT CCCTGATGTATTATTCAGCTCCCTGGAGTT GGCCAGCTCCTGTTACACTGGCCACAGCCC TGGGCATCCGCTTCTCACTTCTAGTTTCCCC TCCAAGGTAATGTGGTGGGTCATGATCATT CTATCCTGGCTTCAGGGACCTGACTCCACT TTGGGGCCATTCGAGGGGTCTAGGGTAGAT GATGTCCCCCTGTGGGGATTAATGTCCTGC TCTGTAAAACTGAGCTAGCTGAGATCCAGG AGGGCTTGGCCAGAGACAGCAAGTTGTTG CCATGGTGACTTTAAAGCCAGGTTGCTGCC CCAGCACAGGCCTCCCAGTCTACCCTCACT AGAAAACAACACCCAGGCACTTTCCACCA CCTCTCAAAGGTGAAACCCAAGGCTGGTCT AGAGAATGAATTATGGATCCTCGCTGTCCG TGCCACCCAGCTAGTCCCAGCGGCTCAGAC ACTGAGGAGAGACTGTAGGTTCAGCTACA AGCAAAAAGACCTAGCTGGTCTCCAAGCA GTGTCTCCAAGTCCCTGAACCTGTGACACC TGCCCCAGGCATCATCAGGCACAGAGGGC CACCAAGAATTCTAGCGGCCGCCACCATG GCCTTGCTCATCCACCTCAAGACAGTCTCG GAGCTGCGGGGCAGGGGCGACCGGATCGC CAAAGTGACTTTCCGAGGGCAATCCTTCTA CTCTCGGGTCCTGGAGAACTGTGAGGATGT GGCTGACTTTGATGAGACATTTCGGTGGCC GGTGGCCAGCAGCATCGACAGAAATGAGA TGCTGGAGATTCAGGTTTTCAACTACAGCA AAGTCTTCAGCAACAAGCTCATCGGGACCT

[0297]

SEQ ID NO.	描述	质粒序列
		TCCGCATGGTGCTGCAGAAGGTGGTAGAG GAGAGCCATGTGGAGGTGACTGACACGCT GATTGATGACAACAATGCTATCATCAAGAC CAGCCTGTGCGTGGAGGTCCGGTATCAGGC CACTGACGGCACAGTGGGCTCCTGGGACG ATGGGGACTTCCTGGGAGATGAGTCTCTTC AAGAGGAAGAGAAGGACAGCCAAGAGAC GGATGGACTGCTCCCAGGCTCCCGGCCCAG CTCCCGGCCCCCAGGAGAGAAGAGCTTCC GGAGAGCCGGGAGGAGCGTGTTCTCCGCC ATGAAGCTCGGCAAAAACCGGTCTCACAA GGAGGAGCCCCAAAGACCAGATGAACCGG CGGTGCTGGAGATGGAAGACCTTGACCAT CTGGCCATTCGGCTAGGAGATGGACTGGAT CCCGACTCGGTGTCTCTAGCCTCAGTCACA GCTCTCACCACTAATGTCTCCAACAAGCGA TCTAAGCCAGACATTAAGATGGAGCCAAG TGCTGGGCGGCCCATGGATTACCAGGTCAG CATCACGGTGATCGAGGCCCGGCAGCTGG TGGGCTTGAACATGGACCCTGTGGTGTGCG TGGAGGTGGGTGACGACAAGAAGTACACA TCCATGAAGGAGTCCACTAACTGCCCCTAT TACAACGAGTACTTCGTCTTCGACTTCCAT GTCTCTCCGGATGTCATGTTTGACAAGATC ATCAAGATTTCGGTGATTCACTCCAAGAAC CTGCTGCGCAGTGGCACCCTGGTGGGCTCC TTCAAAATGGACGTGGGAACCGTGTACTCG CAGCCAGAGCACCAGTTCCATCACAAGTG GGCCATCCTGTCTGACCCCGATGACATCTC CTCGGGGCTGAAGGGCTACGTGAAGTGTG ACGTTGCCGTGGTGGGCAAAGGGGACAAC ATCAAGACGCCCCACAAGGCCAATGAGAC CGACGAAGATGACATTGAGGGGAACTTGC TGCTCCCCGAGGGGGTGCCCCCCGAACGCC AGTGGGCCCGGTTCTATGTGAAAATTTACC GAGCAGAGGGGCTGCCCCGTATGAACACA AGCCTCATGGCCAATGTAAAGAAGGCTTTC ATCGGTGAAAACAAGGACCTCGTGGACCC CTACGTGCAAGTCTTCTTTGCTGGCCAGAA GGGCAAGACTTCAGTGCAGAAGAGCAGCT ATGAGCCCCTGTGGAATGAGCAGGTCGTCT TTACAGACCTCTTCCCCCCACTCTGCAAAC GCATGAAGGTGCAGATCCGAGACTCGGAC AAGGTCAACGACGTGGCCATCGGCACCCA CTTCATTGACCTGCGCAAGATTTCTAATGA CGGAGACAAAGGCTTCCTGCCCACACTGG GCCCAGCCTGGGTGAACATGTACGGCTCCA CACGTAACTACACGCTGCTGGATGAGCATC

[0298]

SEQ ID NO.	描述	质粒序列
		AGGACCTGAACGAGGGCCTGGGGGAGGGT GTGTCCTTCCGGGCCCGGCTCCTGCTGGGC CTGGCTGTGGAGATCGTAGACACCTCCAAC CCTGAGCTCACCAGCTCCACAGAGGTGCA GGTGGAGCAGGCCACGCCCATCTCGGAGA GCTGTGCAGGTAAAATGGAAGAATTCTTTC TCTTTGGAGCCTTCCTGGAGGCCTCAATGA TCGACCGGAGAAACGGAGACAAGCCCATC ACCTTTGAGGTCACCATAGGCAACTATGGG AACGAAGTTGATGGCCTGTCCCGGCCCCAG CGGCCTCGGCCCCGGAAGGAGCCGGGGGA TGAGGAAGAAGTAGACCTGATTCAGAACG CAAGTGATGACGAGGCCGGTGATGCCGGG GACCTGGCCTCAGTCTCCTCCACTCCACCA ATGCGGCCCCAGGTCACCGACAGGAACTA CTTCCATCTGCCCTACCTGGAGCGAAAGCC CTGCATCTACATCAAGAGCTGGTGGCCGGA CCAGCGCCGCCGCCTCTACAATGCCAACAT CATGGACCACATTGCCGACAAGCTGGAAG AAGGCCTGAACGACATACAGGAGATGATC AAAACGGAGAAGTCCTACCCTGAGCGTCG CCTGCGGGGCGTCCTGGAGGAGCTGAGCT GTGGCTGCTGCCGCTTCCTCTCCCTCGCTG ACAAGGACCAGGGCCACTCATCCCGCACC AGGCTTGACCGGGAGCGCCTCAAGTCCTGC ATGAGGGAGCTGGTAAGTATCAAGGTTAC AAGACAGGTTTAAGGAGACCAATAGAAAC TGGGCTTGTCGAGACAGAGAAGACTCTTGC GTTTCTGAGCTAGCCCCCGGGTGCGCGGCG TCGGTGGTGCCGGCGGGGGGCGCCAGGTC GCAGGCGGTGTAGGGCTCCAGGCAGGCGG CGAAGGCCATGACGTGCGCTATGAAGGTC TGCTCCTGCACGCCGTGAACCAGGTGCGCC TGCGGGCCGCGCGCGAACACCGCCACGTC CTCGCCTGCGTGGGTCTCTTCGTCCAGGGG CACTGCTGACTGCTGCCGATACTCGGGGCT CCCGCTCTCGCTCTCGGTAACATCCGGCCG GGCGCCGTCCTTGAGCACATAGCCTGGACC GTTTCGTCGACCTCGAGTTAAGGGCGAATT CCCGATAAGGATCTTCCTAGAGCATGGCTA CGTAGATAAGTAGCATGGCGGGTTAATCAT TAACTACAAGGAACCCCTAGTGATGGAGTT GGCCACTCCCTCTCTGCGCGCTCGCTCGCT CACTGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCTCAG TGAGCGAGCGAGCGCGCAGCCTTAATTAA ATCCACATCTGTATGTTTTTTATATTAATTT ATTTTTTGCAGGGGGGCATTGTTTGGTAGG

[0299]

SEQ ID NO.	描述	质粒序列
		TGAGAGTTCTGAATTGCTATGTTTAGTGAG TTGTATCTATTTATTTTTCAATAAATACAAT TAGTTATGTGTTTTGGGGGCGATCGTGAGG CAAAGAAAACCCGGCGCTGAGGCCGGGTT ATTCTTGTTCTCTGGTCAAATTATATAGTTG GAAAACAAGGATGCATATATGAATGAACG ATGCAGAGGCAATGCCGATGGCGATAGTG GGTATCAGGTAGCCGCTTATGCTGGAAAG AAGCAATAACCCGCAGAAAAACAAAGCTC CAAGCTCAACAAAACTAAGGGCATAGACA ATAACTACCTATGTCATATACCCATACTCT CTAATCTTGGCCAGTCGGCGCGTTCTGCTT CCGATTAGAAACGTCAAGGCAGCAATCAG GATTGCAATCTTGGTTCCTGCATAGGATGA CAATGTCGCCCCAAGACCATCTCTATGAGC TGAAAAAGAAACACAAGGAATGTAGTGGC GGAAAAGGAGATAGCAAATGCTTACGATA ACGTAAGGAATTATTACTATGTAAACACCA GGCAAGATTCTGTTCCGTATAATTACTCCT GATAATTAATCCTTAACTTTGCCCACCTGC CTTTTAAAACATTCCAGTATATCACTTTTCA TTCTTGCGTAGCAATATGCCCTCTCTTCAG CTATCTCAGCATTGGTGACCTTGTTCAGAG GCGCTGAGAGATGGCCTTTTTCTGATAGAT AATGTTCTGTTAAAATATCTCCGGCCTCAT CTTTTGCCCGCAGGCTAATGTCTGAAAATT GAGGTGACGGGTTAAAAATAATATCCTTG GCAACCTTTTTTATATCCCTTTTAAATTTTG GCTTAATGACTATATCCAATGAGTCAAAAA GCTCCCCTTCAATATCTGTTGCCCCTAAGA CCTTTAATATATCGCCAAATACAGGTAGCT TGGCTTCTACCTTCACCGTTGTTCTGCCGAT GAAATGCTAATGCATAACATCGTCTTTGGT GGTTCCCCTCATCAGTGGCTCTATCTGAAC GCGCTCTCCACTGCTTAATGACATTCCTTTC CCGATTAAAAAATCTGTCAGATCGGATGTG GTCGGCCCGAAAACAGTTCTGGCAAAACC AATGGTGTCGCCTTCAACAAACAAAAAAG ATGGGAATCCCAATGATTCGTCATCTGCGA GGCTGTTCTTAATATCTTCAACTGTAGCTTT AGAGCGATTTATCTTCTGAACCAGACTCTT GTCATTTGTTTTGGTAAAGAGAAAAGTTTT TCCATCGATTTTATGAATATACAAATAATT GGAGCCAACCTTCAGGTGATGATTATCAGC CAGCAGAGAATTAAGGAAAACAGACAGGT TTATTGAGCACTTATCTTTCCCTTTATTTTT GCTGCGGTAAGTCGCATAAAAACCATTCTT CACAATTCAATCCATTTACTATGTTATGTTC

[0300]

SEQ ID NO.	描述	质粒序列
		TGAGGGGAGTGAAAATTCCCCTAATTCGAT GAAGATTCTTGCTAAATTGTTATCAGCTAT GCGCCGACCAGAACACCTTGCCGATCAGC CAAACGTCTAATCAGGCCACTGACTAGCG ATAACTTTCCCCACAACGGAACAACTCTCA TTGCATGGGATAATTGGGTACTGTGGGTTT AGTGGTTGTAAAAACACCTGACCGCTATCC CTGATCAGTTTCTTGAAGGTAAACTCATCA CCCCCAAGTCTGGCTATACAGAAATCACCT GGCTCAACAGCCTGCTCAGGGTCAACGAG AATTTACATTCCGTCAGGATAGCTTGGCTT GGAGCCTGTTGGTGCGGTCACGGAATTACC TTCAACCTCAAGCCAGAATGCAGAATCACT GGCTTTTTTGGTTGTGCTTACCCATCTCTCC GCATCACCTTTGGTAAAGGTTCTAAGCTAA GGTGAGAACATCCCTGCCTGAACATGAGA AAAAACAGGGTACTCATACTCACTTATTAG TGACGGCTATGAGCAAAAGGCCAGCAAAA GGCCAGGAACCGTAAAAAGGCCGCGTTGC TGGCGTTTTTCCATAGGCTCCGCCCCCCTG ACGAGCATCACAAAAATCGACGCTCAAGT CAGAGGTGGCGAAACCCGACAGGACTATA AAGATACCAGGCGTTTCCCCCTGGAAGCTC CCTCGTGCGCTCTCCTGTTCCGACCCTGCC GCTTACCGGATACCTGTCCGCCTTTCTCCCT TCGGGAAGCGTGGCGCTTTCTCATAGCTCA CGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAA CCCCCCGTTCAGCCCGACCGCTGCGCCTTA TCCGGTAACTATCGTCTTGAGTCCAACCCG GTAAGACACGACTTATCGCCACTGGCAGC AGCCACTGGTAACAGGATTAGCAGAGCGA GGTATGTAGGCGGTGCTACAGAGTTCTTGA AGTGGTGGCCTAACTACGGCTACACTAGA AGAACAGTATTTGGTATCTGCGCTCTGCTG AAGCCAGTTACCTTCGGAAAAAGAGTTGG TAGCTCTTGATCCGGCAAACAAACCACCGC TGGTAGCGGTGGTTTTTTTGTTTGCAAGCA GCAGATTACGCGCAGAAAAAAAGGATCTC AAGAAGATCCTTTGATCTTTTCTACGGGGT CTGACGCTCAGTGGAACGAAAACTCACGTT AAGGGATTTTGGTCATGAGATTATCAAAAA GGATCTTCACCTAGATCCTTTTAAATTAAA AATGAAGTTTTAAATCAAGCCCAATCTGAA TAATGTTACAACCAATTAACCAATTCTGAT TAGAAAAACTCATCGAGCATCAAATGAAA CTGCAATTTATTCATATCAGGATTATCAAT ACCATATTTTTGAAAAAGCCGTTTCTGTAA

[0301]

SEQ ID NO.	描述	质粒序列
		TGAAGGAGAAAACTCACCGAGGCAGTTCC ATAGGATGGCAAGATCCTGGTATCGGTCTG CGATTCCGACTCGTCCAACATCAATACAAC CTATTAATTTCCCCTCGTCAAAAATAAGGT TATCAAGTGAGAAATCACCATGAGTGACG ACTGAATCCGGTGAGAATGGCAAAAGTTT ATGCATTTCTTTCCAGACTTGTTCAACAGG CCAGCCATTACGCTCGTCATCAAAATCACT CGCATCAACCAAACCGTTATTCATTCGTGA TTGCGCCTGAGCAAGACGAAATACGCGAT CGCTGTTAAAAGGACAATTACAAACAGGA ATCGAATGCAACCGGCGCAGGAACACTGC CAGCGCATCAACAATATTTTCACCTGAATC AGGATATTCTTCTAATACCTGGAATGCTGT TTTTCCGGGGATCGCAGTGGTGAGTAACCA TGCATCATCAGGAGTACGGATAAAATGCTT GATGGTCGGAAGAGGCATAAATTCCGTCA GCCAGTTTAGTCTGACCATCTCATCTGTAA CATCATTGGCAACGCTACCTTTGCCATGTT TCAGAAACAACTCTGGCGCATCGGGCTTCC CATACAAGCGATAGATTGTCGCACCTGATT GCCCGACATTATCGCGAGCCCATTTATACC CATATAAATCAGCATCCATGTTGGAATTTA ATCGCGGCCTCGACGTTTCCCGTTGAATAT GGCTCATAACACCCCTTGTATTACTGTTTA TGTAAGCAGACAGTTTTATTGTTCATGATG ATATATTTTTATCTTGTGCAATGTAACATC AGAGATTTTGAGACACGGGCCAGAGCTGC A
76	含有以下特征的 3’转基因质粒: 位置 53-2027 处的λ 位置 2049-2178 处的 ITR 位置 2267-2553 处的 AP 头部序列 位置 2576-2624 处的剪接受体序列 人类 OTOF 同功型 5 的位置 2625-6212 处的 C 末端部分 位置 6255-6476 处的 bGH poly(A)序列 位置 6564-6693 处的 ITR	TCGCGCGTTTCGGTGATGACGGTGAAAACC TCTGACACATGCAGCTCCCGGATAGAGGTC ATCCTTCCTGACCATTTCCATCATTCCAGTC GAACTCACACACAACACCAAATGCATTTA AGTCGCTTGAAATTGCTATAAGCAGAGCAT GTTGCGCCAGCATGATTAATACAGCATTTA ATACAGAGCCGTGTTTATTGAGTCGGTATT CAGAGTCTGACCAGAAATTATTAATCTGGT GAAGTTATTCCTCTGTCATTACGTCATGGT CGATTTCAATTTCTATTGATGCTTTCCAGTC GTAATCAATGATGTATTTTTTGATGTTTGA CCTCTGTTCATATCCTCACAGATAAAAAAT CGCCCTCACACTGGAGGGCAAAGAAGATT TCCAATAATCAGAACAAGTCGGCTCCTGTT TAGTTACGAGCGACATTGCTCCGTGTATTC ACTCGTTGGAATGAATACACAGTGCAGTGT TTATTCTGTTATTTATGCCAAAAATTAAGG CCACTATCAGGCAGCTTTGTTGTTCTGTTTA CCAAGTTCTCTGGCAATCATTGCCGTCGTT

[0302]

SEQ ID NO.	描述	质粒序列
	位置 6704-8716 处的λ 位置 8773-9361 处的 Ori 位置 9539-10,348 处的 KanR 位置 2049-6693 处的待转移至双载体系统中的载体中的转基因	CGTATTGCCCATTTATCGACATATTTCCCAT CTTCCTATACAGGAAACATTTCTTCAGGCT TAACCATGCATTCCGATTGCAGCTTGCATC CATTGCATCGCTTGAATTGTCCACACCATT GATTTTTATCAATAGTCGTAGTTTAACGGA TAGTCCTGGTATTGTTCCATCACATCCTGA GGATGCCCTTCGAACTCTTCAAATTCTTCTT CCTAATATCACCTTAAATAGTGGATTGCGG TAGTAAAGATTGTGCCTGTCTTTTAACCAC ATCAGGCTCGGTGGTTCTCGTGTACCCCTA CAGCGAGAAATCGGATAAACTATTACAAC CCCTACAGTTTGTAGAGTATAGAAAATGAT CCACTCGTTATTCTCGGACGAGTGTTCAGT AATGAACCTCTGGAGAGAACCATCTATATG ATCGTTATCTGGGTTTGACTTCTGCTTTTAA GCCCAGATAACTTGCCTGAATATGTTAATG AGAGAATCGGTATTCCTCATGTGTGGCATG TTTTCGTCTTTGCTCTTGCATTTTCACTAGC AATTAATGTGCATCGATTATCAGCTATTGC CAGCGCCAGATATAAGCGATTTAAGCTAA GAAAACGCATTAAGGTGCAAAACGATAAA GTGCGATCAGTAATTCAAAACCTTACAGGA GAGCAATCTATGGTTTTGTGCTCAGCCCTT AATGAAGGCAGGTAGTATGTGGTTACATC AAAACAATTCCCATACATTAGTGAGTTGAT TGAGCTTGGTGTGTTGAACAAAACTTTTTC CCGATGGAATGGAAAGCATATATTATTCCC TATTGAGGATATTTACTGGACTGAATTAGT TGCCAGCTATGATCCATATAATATTGAGAT AAAGCCAAGGCCAATATCTAAGTAACTAG ATAAGAGGAATCGATTTTCCCTTAATTTTC TGGCGTCCACTGCATGTTATGCCGCGTTCG CCAGGCTTGCTGTACCATGTGCGCTGATTC TTGCGCTCAATACGTTGCAGGTTGCTTTCA ATCTGTTTGTGGTATTCAGCCAGCACTGTA AGGTCTATCGGATTTAGTGCGCTTTCTACT CGTGATTTCGGTTTGCGATTCAGCGAGAGA ATAGGGCGGTTAACTGGTTTTGCGCTTACC CCAACCAACAGGGGATTTGCTGCTTTCCAT TGAGCCTGTTACTCTGCGCGACGTTCGCGG CGGCGTGTTTGTGCATCCATCTGGATTCTC CTGTCAGTTAGCTTTGGTGGTGTGTGGCAG TTGTAGTCCTGAACGAAAACCCCCCGCGAT TGGCACGTTGGCAGCTAATCCGGAATCGCA CTTACGGCCAATGCTTCGTTTCGTATCACA CACCCCAAAGCCTTCTGCTTTGAATGCTGC CCTTCTTCAGGGCTTAATTTTTAAGAGCGT CACCTTCATGGTGGTCAGTGCGTCCTGCTG

[0303]

SEQ ID NO.	描述	质粒序列
		ATGTGCTCAGGCACGATTTAATTAAGGCCT TAATTAGGCTGCGCGCTCGCTCGCTCACTG AGGCCGCCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAA CTCCATCACTAGGGGTTCCTTGTAGTTAAT GATTAACCCGCCATGCTACTTATCTACGTA GCCATGCTCTAGGAAGATCGGAATTCGCCC TTAAGCTAGCGGCGCGCCCCCCGGGTGCGC GGCGTCGGTGGTGCCGGCGGGGGGCGCCA GGTCGCAGGCGGTGTAGGGCTCCAGGCAG GCGGCGAAGGCCATGACGTGCGCTATGAA GGTCTGCTCCTGCACGCCGTGAACCAGGTG CGCCTGCGGGCCGCGCGCGAACACCGCCA CGTCCTCGCCTGCGTGGGTCTCTTCGTCCA GGGGCACTGCTGACTGCTGCCGATACTCGG GGCTCCCGCTCTCGCTCTCGGTAACATCCG GCCGGGCGCCGTCCTTGAGCACATAGCCTG GACCGTTTCCTTAAGCGACGCATGCTCGCG ATAGGCACCTATTGGTCTTACTGACATCCA CTTTGCCTTTCTCTCCACAGGAAAACATGG GGCAGCAGGCCAGGATGCTGCGGGCCCAG GTGAAGCGGCACACGGTGCGGGACAAGCT GAGGCTGTGCCAGAACTTCCTGCAGAAGCT GCGCTTCCTGGCGGACGAGCCCCAGCACA GCATTCCCGACATCTTCATCTGGATGATGA GCAACAACAAGCGTGTCGCCTATGCCCGTG TGCCCTCCAAGGACCTGCTCTTCTCCATCG TGGAGGAGGAGACTGGCAAGGACTGCGCC AAGGTCAAGACGCTCTTCCTTAAGCTGCCA GGGAAGCGGGGCTTCGGCTCGGCAGGCTG GACAGTGCAGGCCAAGGTGGAGCTGTACC TGTGGCTGGGCCTCAGCAAACAGCGCAAG GAGTTCCTGTGCGGCCTGCCCTGTGGCTTC CAGGAGGTCAAGGCAGCCCAGGGCCTGGG CCTGCATGCCTTCCCACCCGTCAGCCTGGT CTACACCAAGAAGCAGGCGTTCCAGCTCC GAGCGCACATGTACCAGGCCCGCAGCCTCT TTGCCGCCGACAGCAGCGGACTCTCAGACC CCTTTGCCCGCGTCTTCTTCATCAATCAGA GTCAGTGCACAGAGGTGCTGAATGAGACC CTGTGTCCCACCTGGGACCAGATGCTGGTG TTCGACAACCTGGAGCTCTATGGTGAAGCT CATGAGCTGAGGGACGATCCGCCCATCATT GTCATTGAAATCTATGACCAGGATTCCATG GGCAAAGCTGACTTCATGGGCCGGACCTTC GCCAAACCCCTGGTGAAGATGGCAGACGA GGCGTACTGCCCACCCCGCTTCCCACCTCA

[0304]

SEQ ID NO.	描述	质粒序列
		GCTCGAGTACTACCAGATCTACCGTGGCAA CGCCACAGCTGGAGACCTGCTGGCGGCCTT CGAGCTGCTGCAGATTGGACCAGCAGGGA AGGCTGACCTGCCCCCCATCAATGGCCCGG TGGACGTGGACCGAGGTCCCATCATGCCCG TGCCCATGGGCATCCGGCCCGTGCTCAGCA AGTACCGAGTGGAGGTGCTGTTCTGGGGCC TACGGGACCTAAAGCGGGTGAACCTGGCC CAGGTGGACCGGCCACGGGTGGACATCGA GTGTGCAGGGAAGGGGGTGCAGTCGTCCC TGATCCACAATTATAAGAAGAACCCCAACT TCAACACCCTCGTCAAGTGGTTTGAAGTGG ACCTCCCAGAGAACGAGCTGCTGCACCCG CCCTTGAACATCCGTGTGGTGGACTGCCGG GCCTTCGGTCGCTACACACTGGTGGGCTCC CATGCCGTCAGCTCCCTGCGACGCTTCATC TACCGGCCCCCAGACCGCTCGGCCCCCAGC TGGAACACCACGGTCAGGCTTCTCCGGCGC TGCCGTGTGCTGTGCAATGGGGGCTCCTCC TCTCACTCCACAGGGGAGGTTGTGGTGACT ATGGAGCCAGAGGTACCCATCAAGAAACT GGAGACCATGGTGAAGCTGGACGCGACTT CTGAAGCTGTTGTCAAGGTGGATGTGGCTG AGGAGGAGAAGGAGAAGAAGAAGAAGAA GAAGGGCACTGCGGAGGAGCCAGAGGAGG AGGAGCCAGACGAGAGCATGCTGGACTGG TGGTCCAAGTACTTTGCCTCCATTGACACC ATGAAGGAGCAACTTCGACAACAAGAGCC CTCTGGAATTGACTTGGAGGAGAAGGAGG AAGTGGACAATACCGAGGGCCTGAAGGGG TCAATGAAGGGCAAGGAGAAGGCAAGGGC TGCCAAAGAGGAGAAGAAGAAGAAAACTC AGAGCTCTGGCTCTGGCCAGGGGTCCGAG GCCCCCGAGAAGAAGAAACCCAAGATTGA TGAGCTTAAGGTATACCCCAAAGAGCTGG AGTCCGAGTTTGATAACTTTGAGGACTGGC TGCACACTTTCAACTTGCTTCGGGGCAAGA CCGGGGATGATGAGGATGGCTCCACCGAG GAGGAGCGCATTGTGGGACGCTTCAAGGG CTCCCTCTGCGTGTACAAAGTGCCACTCCC AGAGGACGTGTCCCGGGAAGCCGGCTACG ACTCCACCTACGGCATGTTCCAGGGCATCC CGAGCAATGACCCCATCAATGTGCTGGTCC GAGTCTATGTGGTCCGGGCCACGGACCTGC ACCCTGCTGACATCAACGGCAAAGCTGAC CCCTACATCGCCATCCGGCTAGGCAAGACT GACATCCGCGACAAGGAGAACTACATCTC CAAGCAGCTCAACCCTGTCTTTGGGAAGTC

[0305]

SEQ ID NO.	描述	质粒序列
		CTTTGACATCGAGGCCTCCTTCCCCATGGA ATCCATGCTGACGGTGGCTGTGTATGACTG GGACCTGGTGGGCACTGATGACCTCATTGG GGAAACCAAGATCGACCTGGAGAACCGCT TCTACAGCAAGCACCGCGCCACCTGCGGC ATCGCCCAGACCTACTCCACACATGGCTAC AATATCTGGCGGGACCCCATGAAGCCCAG CCAGATCCTGACCCGCCTCTGCAAAGACGG CAAAGTGGACGGCCCCCACTTTGGGCCCCC TGGGAGAGTGAAGGTGGCCAACCGCGTCT TCACTGGGCCCTCTGAGATTGAGGACGAG AACGGTCAGAGGAAGCCCACAGACGAGCA TGTGGCGCTGTTGGCCCTGAGGCACTGGGA GGACATCCCCCGCGCAGGCTGCCGCCTGGT GCCAGAGCATGTGGAGACGAGGCCGCTGC TCAACCCCGACAAGCCGGGCATCGAGCAG GGCCGCCTGGAGCTGTGGGTGGACATGTTC CCCATGGACATGCCAGCCCCTGGGACGCCT CTGGACATCTCACCTCGGAAGCCCAAGAA GTACGAGCTGCGGGTCATCATCTGGAACAC AGATGAGGTGGTCTTGGAGGACGACGACT TCTTCACAGGGGAGAAGTCCAGTGACATCT TCGTGAGGGGGTGGCTGAAGGGCCAGCAG GAGGACAAGCAGGACACAGACGTCCACTA CCACTCCCTCACTGGCGAGGGCAACTTCAA CTGGCGCTACCTGTTCCCCTTCGACTACCT GGCGGCGGAGGAGAAGATCGTCATCTCCA AGAAGGAGTCCATGTTCTCCTGGGACGAG ACCGAGTACAAGATCCCCGCGCGGCTCAC CCTGCAGATCTGGGATGCGGACCACTTCTC CGCTGACGACTTCCTGGGGGCCATCGAGCT GGACCTGAACCGGTTCCCGCGGGGCGCAA AGACAGCCAAGCAGTGCACCATGGAGATG GCCACCGGGGAGGTGGACGTGCCCCTCGT GTCCATCTTCAAGCAAAAGCGCGTCAAAG GCTGGTGGCCCCTCCTGGCCCGCAATGAGA ACGATGAGTTTGAGCTCACGGGCAAGGTG GAGGCTGAGCTGCATTTACTGACAGCAGA GGAGGCAGAGAAGAACCCAGTGGGCCTGG CCCGCAATGAACCTGACCCCCTAGAGAAA CCCAACCGGCCCGACACGGCCTTCGTCTGG TTCCTCAACCCTCTCAAGTCCATCAAGTAC CTCATCTGCACCCGGTACAAGTGGCTCATC ATCAAGATCGTGCTGGCGCTGTTGGGGCTG CTCATGTTGGGGCTCTTCCTCTACAGCCTC CCTGGCTACATGGTCAAAAAGCTCCTTGGG GCATGAACGGCCGCTATGCTAGCTTGGTAC CAAGGGCGGATCCTGCATAGAGCTCGCTG

[0306]

SEQ ID NO.	描述	质粒序列
		ATCAGCCTCGACTGTGCCTTCTAGTTGCCA GCCATCTGTTGTTTGCCCCTCCCCCGTGCCT TCCTTGACCCTGGAAGGTGCCACTCCCACT GTCCTTTCCTAATAAAATGAGGAAATTGCA TCGCATTGTCTGAGTAGGTGTCATTCTATT CTGGGGGGTGGGGTGGGGCAGGACAGCAA GGGGGAGGATTGGGAAGACAATAGCAGGC ATCTCGAGTTAAGGGCGAATTCCCGATAAG GATCTTCCTAGAGCATGGCTACGTAGATAA GTAGCATGGCGGGTTAATCATTAACTACAA GGAACCCCTAGTGATGGAGTTGGCCACTCC CTCTCTGCGCGCTCGCTCGCTCACTGAGGC CGGGCGACCAAAGGTCGCCCGACGCCCGG GCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCAGCCTTAATTAAATCCACATCT GTATGTTTTTTATATTAATTTATTTTTTGCA GGGGGGCATTGTTTGGTAGGTGAGAGTTCT GAATTGCTATGTTTAGTGAGTTGTATCTAT TTATTTTTCAATAAATACAATTAGTTATGT GTTTTGGGGGCGATCGTGAGGCAAAGAAA ACCCGGCGCTGAGGCCGGGTTATTCTTGTT CTCTGGTCAAATTATATAGTTGGAAAACAA GGATGCATATATGAATGAACGATGCAGAG GCAATGCCGATGGCGATAGTGGGTATCAG GTAGCCGCTTATGCTGGAAAGAAGCAATA ACCCGCAGAAAAACAAAGCTCCAAGCTCA ACAAAACTAAGGGCATAGACAATAACTAC CTATGTCATATACCCATACTCTCTAATCTTG GCCAGTCGGCGCGTTCTGCTTCCGATTAGA AACGTCAAGGCAGCAATCAGGATTGCAAT CTTGGTTCCTGCATAGGATGACAATGTCGC CCCAAGACCATCTCTATGAGCTGAAAAAG AAACACAAGGAATGTAGTGGCGGAAAAGG AGATAGCAAATGCTTACGATAACGTAAGG AATTATTACTATGTAAACACCAGGCAAGAT TCTGTTCCGTATAATTACTCCTGATAATTA ATCCTTAACTTTGCCCACCTGCCTTTTAAA ACATTCCAGTATATCACTTTTCATTCTTGCG TAGCAATATGCCCTCTCTTCAGCTATCTCA GCATTGGTGACCTTGTTCAGAGGCGCTGAG AGATGGCCTTTTTCTGATAGATAATGTTCT GTTAAAATATCTCCGGCCTCATCTTTTGCC CGCAGGCTAATGTCTGAAAATTGAGGTGA CGGGTTAAAAATAATATCCTTGGCAACCTT TTTTATATCCCTTTTAAATTTTGGCTTAATG ACTATATCCAATGAGTCAAAAAGCTCCCCT TCAATATCTGTTGCCCCTAAGACCTTTAAT ATATCGCCAAATACAGGTAGCTTGGCTTCT

[0307]

SEQ ID NO.	描述	质粒序列
		ACCTTCACCGTTGTTCTGCCGATGAAATGC TAATGCATAACATCGTCTTTGGTGGTTCCC CTCATCAGTGGCTCTATCTGAACGCGCTCT CCACTGCTTAATGACATTCCTTTCCCGATT AAAAAATCTGTCAGATCGGATGTGGTCGG CCCGAAAACAGTTCTGGCAAAACCAATGG TGTCGCCTTCAACAAACAAAAAAGATGGG AATCCCAATGATTCGTCATCTGCGAGGCTG TTCTTAATATCTTCAACTGTAGCTTTAGAG CGATTTATCTTCTGAACCAGACTCTTGTCA TTTGTTTTGGTAAAGAGAAAAGTTTTTCCA TCGATTTTATGAATATACAAATAATTGGAG CCAACCTTCAGGTGATGATTATCAGCCAGC AGAGAATTAAGGAAAACAGACAGGTTTAT TGAGCACTTATCTTTCCCTTTATTTTTGCTG CGGTAAGTCGCATAAAAACCATTCTTCACA ATTCAATCCATTTACTATGTTATGTTCTGAG GGGAGTGAAAATTCCCCTAATTCGATGAA GATTCTTGCTAAATTGTTATCAGCTATGCG CCGACCAGAACACCTTGCCGATCAGCCAA ACGTCTAATCAGGCCACTGACTAGCGATAA CTTTCCCCACAACGGAACAACTCTCATTGC ATGGGATAATTGGGTACTGTGGGTTTAGTG GTTGTAAAAACACCTGACCGCTATCCCTGA TCAGTTTCTTGAAGGTAAACTCATCACCCC CAAGTCTGGCTATACAGAAATCACCTGGCT CAACAGCCTGCTCAGGGTCAACGAGAATTT ACATTCCGTCAGGATAGCTTGGCTTGGAGC CTGTTGGTGCGGTCACGGAATTACCTTCAA CCTCAAGCCAGAATGCAGAATCACTGGCTT TTTTGGTTGTGCTTACCCATCTCTCCGCATC ACCTTTGGTAAAGGTTCTAAGCTAAGGTGA GAACATCCCTGCCTGAACATGAGAAAAAA CAGGGTACTCATACTCACTTATTAGTGACG GCTATGAGCAAAAGGCCAGCAAAAGGCCA GGAACCGTAAAAAGGCCGCGTTGCTGGCG TTTTTCCATAGGCTCCGCCCCCCTGACGAG CATCACAAAAATCGACGCTCAAGTCAGAG GTGGCGAAACCCGACAGGACTATAAAGAT ACCAGGCGTTTCCCCCTGGAAGCTCCCTCG TGCGCTCTCCTGTTCCGACCCTGCCGCTTA CCGGATACCTGTCCGCCTTTCTCCCTTCGG GAAGCGTGGCGCTTTCTCATAGCTCACGCT GTAGGTATCTCAGTTCGGTGTAGGTCGTTC GCTCCAAGCTGGGCTGTGTGCACGAACCCC CCGTTCAGCCCGACCGCTGCGCCTTATCCG GTAACTATCGTCTTGAGTCCAACCCGGTAA GACACGACTTATCGCCACTGGCAGCAGCC

[0308]

SEQ ID NO.	描述	质粒序列
		ACTGGTAACAGGATTAGCAGAGCGAGGTA TGTAGGCGGTGCTACAGAGTTCTTGAAGTG GTGGCCTAACTACGGCTACACTAGAAGAA CAGTATTTGGTATCTGCGCTCTGCTGAAGC CAGTTACCTTCGGAAAAAGAGTTGGTAGCT CTTGATCCGGCAAACAAACCACCGCTGGTA GCGGTGGTTTTTTTGTTTGCAAGCAGCAGA TTACGCGCAGAAAAAAAGGATCTCAAGAA GATCCTTTGATCTTTTCTACGGGGTCTGAC GCTCAGTGGAACGAAAACTCACGTTAAGG GATTTTGGTCATGAGATTATCAAAAAGGAT CTTCACCTAGATCCTTTTAAATTAAAAATG AAGTTTTAAATCAAGCCCAATCTGAATAAT GTTACAACCAATTAACCAATTCTGATTAGA AAAACTCATCGAGCATCAAATGAAACTGC AATTTATTCATATCAGGATTATCAATACCA TATTTTTGAAAAAGCCGTTTCTGTAATGAA GGAGAAAACTCACCGAGGCAGTTCCATAG GATGGCAAGATCCTGGTATCGGTCTGCGAT TCCGACTCGTCCAACATCAATACAACCTAT TAATTTCCCCTCGTCAAAAATAAGGTTATC AAGTGAGAAATCACCATGAGTGACGACTG AATCCGGTGAGAATGGCAAAAGTTTATGC ATTTCTTTCCAGACTTGTTCAACAGGCCAG CCATTACGCTCGTCATCAAAATCACTCGCA TCAACCAAACCGTTATTCATTCGTGATTGC GCCTGAGCAAGACGAAATACGCGATCGCT GTTAAAAGGACAATTACAAACAGGAATCG AATGCAACCGGCGCAGGAACACTGCCAGC GCATCAACAATATTTTCACCTGAATCAGGA TATTCTTCTAATACCTGGAATGCTGTTTTTC CGGGGATCGCAGTGGTGAGTAACCATGCA TCATCAGGAGTACGGATAAAATGCTTGATG GTCGGAAGAGGCATAAATTCCGTCAGCCA GTTTAGTCTGACCATCTCATCTGTAACATC ATTGGCAACGCTACCTTTGCCATGTTTCAG AAACAACTCTGGCGCATCGGGCTTCCCATA CAAGCGATAGATTGTCGCACCTGATTGCCC GACATTATCGCGAGCCCATTTATACCCATA TAAATCAGCATCCATGTTGGAATTTAATCG CGGCCTCGACGTTTCCCGTTGAATATGGCT CATAACACCCCTTGTATTACTGTTTATGTA AGCAGACAGTTTTATTGTTCATGATGATAT ATTTTTATCTTGTGCAATGTAACATCAGAG ATTTTGAGACACGGGCCAGAGCTGCA
77	含有以下特征的 5’ 转基因质粒:	GGGGGGGGGGGGGGGGGGTTGGCCACTCC CTCTCTGCGCGCTCGCTCGCTCACTGAGGC CGGGCGACCAAAGGTCGCCCGACGCCCGG

[0309]

SEQ ID NO.	描述	质粒序列
	位置 19-161 处的 ITR 位置 182-1146 处的 Myo15 启动子 位置 1167-3572 处的人类 OTOF 同功型 5 的 N 末端部分 位置 3573-3656 处的剪接供体序列 位置 3663-3949 处的 AP 头部序列 位置 3973-4115 处的 ITR 位置 4641-5435 处的 KanR 位置 5821-6491 处的 pUC ori 位置 19-4115 处的待转移至双载体系统中的载体中的转基因	GCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCC ATCACTAGGGGTTCCTCAGATCTGAATTCG GTACCTGCAGCTCAGCCTACTACTTGCTTT CCAGGCTGTTCCTAGTTCCCATGTCAGCTG CTTGTGCTTTCCAGAGACAAAACAGGAATA ATAGATGTCATTAAATATACATTGGGCCCC AGGCGGTCAATGTGGCAGCCTGAGCCTCCT TTCCATCTCTGTGGAGGCAGACATAGGACC CCCAACAAACAGCATGCAGGTTGGGAGCC AGCCACAGGACCCAGGTAAGGGGCCCTGG GTCCTTAAGCTTCTGCCACTGGCTCCGGCA TTGCAGAGAGAAGAGAAGGGGCGGCAGAC TGGAGAGCTGGGCTCCATTTTTGTTCCTTG GTGCCCTGCCCCTCCCCATGACCTGCAGAG ACATTCAGCCTGCCAGGCTTTATGAGGTGG GAGCTGGGCTCTCCCTGATGTATTATTCAG CTCCCTGGAGTTGGCCAGCTCCTGTTACAC TGGCCACAGCCCTGGGCATCCGCTTCTCAC TTCTAGTTTCCCCTCCAAGGTAATGTGGTG GGTCATGATCATTCTATCCTGGCTTCAGGG ACCTGACTCCACTTTGGGGCCATTCGAGGG GTCTAGGGTAGATGATGTCCCCCTGTGGGG ATTAATGTCCTGCTCTGTAAAACTGAGCTA GCTGAGATCCAGGAGGGCTTGGCCAGAGA CAGCAAGTTGTTGCCATGGTGACTTTAAAG CCAGGTTGCTGCCCCAGCACAGGCCTCCCA GTCTACCCTCACTAGAAAACAACACCCAG GCACTTTCCACCACCTCTCAAAGGTGAAAC CCAAGGCTGGTCTAGAGAATGAATTATGG ATCCTCGCTGTCCGTGCCACCCAGCTAGTC CCAGCGGCTCAGACACTGAGGAGAGACTG TAGGTTCAGCTACAAGCAAAAAGACCTAG CTGGTCTCCAAGCAGTGTCTCCAAGTCCCT GAACCTGTGACACCTGCCCCAGGCATCATC AGGCACAGAGGGCCACCGAATTCTAGCGG CCGCCACCATGGCCTTGCTCATCCACCTCA AGACAGTCTCGGAGCTGCGGGGCAGGGGC GACCGGATCGCCAAAGTGACTTTCCGAGG GCAATCCTTCTACTCTCGGGTCCTGGAGAA CTGTGAGGATGTGGCTGACTTTGATGAGAC ATTTCGGTGGCCGGTGGCCAGCAGCATCGA CAGAAATGAGATGCTGGAGATTCAGGTTTT CAACTACAGCAAAGTCTTCAGCAACAAGC TCATCGGGACCTTCCGCATGGTGCTGCAGA AGGTGGTAGAGGAGAGCCATGTGGAGGTG ACTGACACGCTGATTGATGACAACAATGCT ATCATCAAGACCAGCCTGTGCGTGGAGGTC

[0310]

SEQ ID NO.	描述	质粒序列
		CGGTATCAGGCCACTGACGGCACAGTGGG CTCCTGGGACGATGGGGACTTCCTGGGAG ATGAGTCTCTTCAAGAGGAAGAGAAGGAC AGCCAAGAGACGGATGGACTGCTCCCAGG CTCCCGGCCCAGCTCCCGGCCCCCAGGAGA GAAGAGCTTCCGGAGAGCCGGGAGGAGCG TGTTCTCCGCCATGAAGCTCGGCAAAAACC GGTCTCACAAGGAGGAGCCCCAAAGACCA GATGAACCGGCGGTGCTGGAGATGGAAGA CCTTGACCATCTGGCCATTCGGCTAGGAGA TGGACTGGATCCCGACTCGGTGTCTCTAGC CTCAGTCACAGCTCTCACCACTAATGTCTC CAACAAGCGATCTAAGCCAGACATTAAGA TGGAGCCAAGTGCTGGGCGGCCCATGGAT TACCAGGTCAGCATCACGGTGATCGAGGC CCGGCAGCTGGTGGGCTTGAACATGGACC CTGTGGTGTGCGTGGAGGTGGGTGACGAC AAGAAGTACACATCCATGAAGGAGTCCAC TAACTGCCCCTATTACAACGAGTACTTCGT CTTCGACTTCCATGTCTCTCCGGATGTCAT GTTTGACAAGATCATCAAGATTTCGGTGAT TCACTCCAAGAACCTGCTGCGCAGTGGCAC CCTGGTGGGCTCCTTCAAAATGGACGTGGG AACCGTGTACTCGCAGCCAGAGCACCAGTT CCATCACAAGTGGGCCATCCTGTCTGACCC CGATGACATCTCCTCGGGGCTGAAGGGCTA CGTGAAGTGTGACGTTGCCGTGGTGGGCA AAGGGGACAACATCAAGACGCCCCACAAG GCCAATGAGACCGACGAAGATGACATTGA GGGGAACTTGCTGCTCCCCGAGGGGGTGC CCCCCGAACGCCAGTGGGCCCGGTTCTATG TGAAAATTTACCGAGCAGAGGGGCTGCCC CGTATGAACACAAGCCTCATGGCCAATGTA AAGAAGGCTTTCATCGGTGAAAACAAGGA CCTCGTGGACCCCTACGTGCAAGTCTTCTT TGCTGGCCAGAAGGGCAAGACTTCAGTGC AGAAGAGCAGCTATGAGCCCCTGTGGAAT GAGCAGGTCGTCTTTACAGACCTCTTCCCC CCACTCTGCAAACGCATGAAGGTGCAGAT CCGAGACTCGGACAAGGTCAACGACGTGG CCATCGGCACCCACTTCATTGACCTGCGCA AGATTTCTAATGACGGAGACAAAGGCTTCC TGCCCACACTGGGCCCAGCCTGGGTGAAC ATGTACGGCTCCACACGTAACTACACGCTG CTGGATGAGCATCAGGACCTGAACGAGGG CCTGGGGGAGGGTGTGTCCTTCCGGGCCCG GCTCCTGCTGGGCCTGGCTGTGGAGATCGT AGACACCTCCAACCCTGAGCTCACCAGCTC

[0311]

SEQ ID NO.	描述	质粒序列
		CACAGAGGTGCAGGTGGAGCAGGCCACGC CCATCTCGGAGAGCTGTGCAGGTAAAATG GAAGAATTCTTTCTCTTTGGAGCCTTCCTG GAGGCCTCAATGATCGACCGGAGAAACGG AGACAAGCCCATCACCTTTGAGGTCACCAT AGGCAACTATGGGAACGAAGTTGATGGCC TGTCCCGGCCCCAGCGGCCTCGGCCCCGGA AGGAGCCGGGGGATGAGGAAGAAGTAGAC CTGATTCAGAACGCAAGTGATGACGAGGC CGGTGATGCCGGGGACCTGGCCTCAGTCTC CTCCACTCCACCAATGCGGCCCCAGGTCAC CGACAGGAACTACTTCCATCTGCCCTACCT GGAGCGAAAGCCCTGCATCTACATCAAGA GCTGGTGGCCGGACCAGCGCCGCCGCCTCT ACAATGCCAACATCATGGACCACATTGCCG ACAAGCTGGAAGAAGGCCTGAACGACATA CAGGAGATGATCAAAACGGAGAAGTCCTA CCCTGAGCGTCGCCTGCGGGGCGTCCTGGA GGAGCTGAGCTGTGGCTGCTGCCGCTTCCT CTCCCTCGCTGACAAGGACCAGGGCCACTC ATCCCGCACCAGGCTTGACCGGGAGCGCCT CAAGTCCTGCATGAGGGAGCTGGTAAGTA TCAAGGTTACAAGACAGGTTTAAGGAGAC CAATAGAAACTGGGCTTGTCGAGACAGAG AAGACTCTTGCGTTTCTGAGCTAGCCCCCG GGTGCGCGGCGTCGGTGGTGCCGGCGGGG GGCGCCAGGTCGCAGGCGGTGTAGGGCTC CAGGCAGGCGGCGAAGGCCATGACGTGCG CTATGAAGGTCTGCTCCTGCACGCCGTGAA CCAGGTGCGCCTGCGGGCCGCGCGCGAAC ACCGCCACGTCCTCGCCTGCGTGGGTCTCT TCGTCCAGGGGCACTGCTGACTGCTGCCGA TACTCGGGGCTCCCGCTCTCGCTCTCGGTA ACATCCGGCCGGGCGCCGTCCTTGAGCACA TAGCCTGGACCGTTTCGTCGACTGGGGAGA GATCTGAGGAACCCCTAGTGATGGAGTTG GCCACTCCCTCTCTGCGCGCTCGCTCGCTC ACTGAGGCCGCCCGGGCAAAGCCCGGGCG TCGGGCGACCTTTGGTCGCCCGGCCTCAGT GAGCGAGCGAGCGCGCAGAGAGGGAGTGG CCAACCCCCCCCCCCCCCCCCTGCAGCCT GGCGTAATAGCGAAGAGGCCCGCACCGAT CGCCCTTCCCAACAGTTGCGTAGCCTGAAT GGCGAATGGCGCGACGCGCCCTGTAGCGG CGCATTAAGCGCGGCGGGTGTGGTGGTTAC GCGCAGCGTGACCGCTACACTTGCCAGCGC CCTAGCGCCCGCTCCTTTCGCTTTCTTCCCT TCCTTTCTCGCCACGTTCGCCGGCTTTCCCC

[0312]

SEQ ID NO.	描述	质粒序列
		GTCAAGCTCTAAATCGGGGGCTCCCTTTAG GGTTCCGATTTAGTGCTTTACGGCACCTCG ACCCCAAAAAACTTGATTAGGGTGATGGTT CACGTAGTGGGCCATCGCCCTGATAGACG GTTTTTCGCCCTTTGACGTTGGAGTCCACG TTCTTTAATAGTGGACTCTTGTTCCAAACT GGAACAACACTCAACCCTATCGCGGTCTAT TCTTTTGATTTATAAGGGATGTTGCCGATTT CGGCCTATTGGTTAAAAAATGAGCTGATTT AACAAAAATTTTAACAAAATTCAGAAGAA CTCGTCAAGAAGGCGATAGAAGGCGATGC GCTGCGAATCGGGAGCGGCGATACCGTAA AGCACGAGGAAGCGGTCAGCCCATTCGCC GCCAAGCTCTTCAGCAATATCACGGGTAGC CAACGCTATGTCCTGATAGCGGTCCGCCAC ACCCAGCCGGCCACAGTCGATGAATCCAG AAAAGCGGCCATTTTCCACCATGATATTCG GCAAGCAGGCATCGCCATGGGTCACGACG AGATCCTCGCCGTCGGGCATGCTCGCCTTG AGCCTGGCGAACAGTTCGGCTGGCGCGAG CCCCTGATGCTCTTCGTCCAGATCATCCTG ATCGACAAGACCGGCTTCCATCCGAGTACG TGCTCGCTCGATGCGATGTTTCGCTTGGTG GTCGAATGGGCAGGTAGCCGGATCAAGCG TATGCAGCCGCCGCATTGCATCAGCCATGA TGGATACTTTCTCGGCAGGAGCAAGGTGA GATGACAGGAGATCCTGCCCCGGCACTTCG CCCAATAGCAGCCAGTCCCTTCCCGCTTCA GTGACAACGTCGAGCACAGCTGCGCAAGG AACGCCCGTCGTGGCCAGCCACGATAGCC GCGCTGCCTCGTCTTGCAGTTCATTCAGGG CACCGGACAGGTCGGTCTTGACAAAAAGA ACCGGGCGCCCCTGCGCTGACAGCCGGAA CACGGCGGCATCAGAGCAGCCGATTGTCT GTTGTGCCCAGTCATAGCCGAATAGCCTCT CCACCCAAGCGGCCGGAGAACCTGCGTGC AATCCATCTTGTTCAATCATGCGAAACGAT CCTCATCCTGTCTCTTGATCAGATCTTGATC CCCTGCGCCATCAGATCCTTGGCGGCGAGA AAGCCATCCAGTTTACTTTGCAGGGCTTCC CAACCTTACCAGAGGGCGCCCCAGCTGGC AATTCCGGTTCGCTTGCTGTCCATAAAACC GCCCAGTCTAGCTATCGCCATGTAAGCCCA CTGCAAGCTACCTGCTTTCTCTTTGCGCTTG CGTTTTCCCTTGTCCAGATAGCCCAGTAGC TGACATTCATCCGGGGTCAGCACCGTTTCT GCGGACTGGCTTTCTACGTGAAAAGGATCT AGGTGAAGATCCTTTTTGATAATCTCATGA

[0313]

SEQ ID NO.	描述	质粒序列
		CCAAAATCCCTTAACGTGAGTTTTCGTTCC ACTGAGCGTCAGACCCCGTAGAAAAGATC AAAGGATCTTCTTGAGATCCTTTTTTTCTGC GCGTAATCTGCTGCTTGCAAACAAAAAAA CCACCGCTACCAGCGGTGGTTTGTTTGCCG GATCAAGAGCTACCAACTCTTTTTCCGAAG GTAACTGGCTTCAGCAGAGCGCAGATACC AAATACTGTCCTTCTAGTGTAGCCGTAGTT AGGCCACCACTTCAAGAACTCTGTAGCACC GCCTACATACCTCGCTCTGCTAATCCTGTT ACCAGTGGCTGCTGCCAGTGGCGATAAGTC GTGTCTTACCGGGTTGGACTCAAGACGATA GTTACCGGATAAGGCGCAGCGGTCGGGCT GAACGGGGGGTTCGTGCACACAGCCCAGC TTGGAGCGAACGACCTACACCGAACTGAG ATACCTACAGCGTGAGCTATGAGAAAGCG CCACGCTTCCCGAAGGGAGAAAGGCGGAC AGGTATCCGGTAAGCGGCAGGGTCGGAAC AGGAGAGCGCACGAGGGAGCTTCCAGGGG GAAACGCCTGGTATCTTTATAGTCCTGTCG GGTTTCGCCACCTCTGACTTGAGCGTCGAT TTTTGTGATGCTCGTCAGGGGGGCGGAGCC TATGGAAAAACGCCAGCAACGCGGCCTTTT TACGGTTCCTGGGCTTTTGCTGGCCTTTTGC TCACATGTTCTTTCCTGCGTTATCCCCTGAT TCTGTGGATAACCGTATTACCGCCTTTGAG TGAGCTGATACCGCTCGCCGCAGCCGAAC GACCGAGCGCAGCGAGTCAGTGAGCGAGG AAGCGGAAGAGCGCCCAATACGCAAACCG CCTCTCCCCGCGCGTTGGCCGATTCATTAA TGCAGGGCTGCA
78	含有以下特征的 3’转基因质粒: 位置 19-161 处的 ITR 位置 187-473 处的 AP 头部序列 位置 496-544 处的剪接受体序列 人类 OTOF 同功型 5 的位置 545-4132 处的 C 末端部分 位置 4175-4396 处的 bGH poly(A)序列 位置 4447-4589 处	GGGGGGGGGGGGGGGGGGTTGGCCACTCC CTCTCTGCGCGCTCGCTCGCTCACTGAGGC CGGGCGACCAAAGGTCGCCCGACGCCCGG GCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCC ATCACTAGGGGTTCCTCAGATCTGAATTCT AGCGGCCGCCCCCGGGTGCGCGGCGTCGG TGGTGCCGGCGGGGGGCGCCAGGTCGCAG GCGGTGTAGGGCTCCAGGCAGGCGGCGAA GGCCATGACGTGCGCTATGAAGGTCTGCTC CTGCACGCCGTGAACCAGGTGCGCCTGCG GGCCGCGCGCGAACACCGCCACGTCCTCG CCTGCGTGGGTCTCTTCGTCCAGGGGCACT GCTGACTGCTGCCGATACTCGGGGCTCCCG CTCTCGCTCTCGGTAACATCCGGCCGGGCG CCGTCCTTGAGCACATAGCCTGGACCGTTT CCTTAAGCGACGCATGCTCGCGATAGGCAC

[0314]

SEQ ID NO.	描述	质粒序列
	的 ITR 位置 5115-5909 处的 KanR 位置 6295-6965 处的 pUC ori 位置 19-4589 处的待转移至双载体系统中的载体中的转基因	CTATTGGTCTTACTGACATCCACTTTGCCTT TCTCTCCACAGGAAAACATGGGGCAGCAG GCCAGGATGCTGCGGGCCCAGGTGAAGCG GCACACGGTGCGGGACAAGCTGAGGCTGT GCCAGAACTTCCTGCAGAAGCTGCGCTTCC TGGCGGACGAGCCCCAGCACAGCATTCCC GACATCTTCATCTGGATGATGAGCAACAAC AAGCGTGTCGCCTATGCCCGTGTGCCCTCC AAGGACCTGCTCTTCTCCATCGTGGAGGAG GAGACTGGCAAGGACTGCGCCAAGGTCAA GACGCTCTTCCTTAAGCTGCCAGGGAAGCG GGGCTTCGGCTCGGCAGGCTGGACAGTGC AGGCCAAGGTGGAGCTGTACCTGTGGCTG GGCCTCAGCAAACAGCGCAAGGAGTTCCT GTGCGGCCTGCCCTGTGGCTTCCAGGAGGT CAAGGCAGCCCAGGGCCTGGGCCTGCATG CCTTCCCACCCGTCAGCCTGGTCTACACCA AGAAGCAGGCGTTCCAGCTCCGAGCGCAC ATGTACCAGGCCCGCAGCCTCTTTGCCGCC GACAGCAGCGGACTCTCAGACCCCTTTGCC CGCGTCTTCTTCATCAATCAGAGTCAGTGC ACAGAGGTGCTGAATGAGACCCTGTGTCCC ACCTGGGACCAGATGCTGGTGTTCGACAAC CTGGAGCTCTATGGTGAAGCTCATGAGCTG AGGGACGATCCGCCCATCATTGTCATTGAA ATCTATGACCAGGATTCCATGGGCAAAGCT GACTTCATGGGCCGGACCTTCGCCAAACCC CTGGTGAAGATGGCAGACGAGGCGTACTG CCCACCCCGCTTCCCACCTCAGCTCGAGTA CTACCAGATCTACCGTGGCAACGCCACAGC TGGAGACCTGCTGGCGGCCTTCGAGCTGCT GCAGATTGGACCAGCAGGGAAGGCTGACC TGCCCCCCATCAATGGCCCGGTGGACGTGG ACCGAGGTCCCATCATGCCCGTGCCCATGG GCATCCGGCCCGTGCTCAGCAAGTACCGA GTGGAGGTGCTGTTCTGGGGCCTACGGGAC CTAAAGCGGGTGAACCTGGCCCAGGTGGA CCGGCCACGGGTGGACATCGAGTGTGCAG GGAAGGGGGTGCAGTCGTCCCTGATCCAC AATTATAAGAAGAACCCCAACTTCAACAC CCTCGTCAAGTGGTTTGAAGTGGACCTCCC AGAGAACGAGCTGCTGCACCCGCCCTTGA ACATCCGTGTGGTGGACTGCCGGGCCTTCG GTCGCTACACACTGGTGGGCTCCCATGCCG TCAGCTCCCTGCGACGCTTCATCTACCGGC CCCCAGACCGCTCGGCCCCCAGCTGGAAC ACCACGGTCAGGCTTCTCCGGCGCTGCCGT GTGCTGTGCAATGGGGGCTCCTCCTCTCAC

[0315]

SEQ ID NO.	描述	质粒序列
		TCCACAGGGGAGGTTGTGGTGACTATGGA GCCAGAGGTACCCATCAAGAAACTGGAGA CCATGGTGAAGCTGGACGCGACTTCTGAA GCTGTTGTCAAGGTGGATGTGGCTGAGGA GGAGAAGGAGAAGAAGAAGAAGAAGAAG GGCACTGCGGAGGAGCCAGAGGAGGAGGA GCCAGACGAGAGCATGCTGGACTGGTGGT CCAAGTACTTTGCCTCCATTGACACCATGA AGGAGCAACTTCGACAACAAGAGCCCTCT GGAATTGACTTGGAGGAGAAGGAGGAAGT GGACAATACCGAGGGCCTGAAGGGGTCAA TGAAGGGCAAGGAGAAGGCAAGGGCTGCC AAAGAGGAGAAGAAGAAGAAAACTCAGA GCTCTGGCTCTGGCCAGGGGTCCGAGGCCC CCGAGAAGAAGAAACCCAAGATTGATGAG CTTAAGGTATACCCCAAAGAGCTGGAGTCC GAGTTTGATAACTTTGAGGACTGGCTGCAC ACTTTCAACTTGCTTCGGGGCAAGACCGGG GATGATGAGGATGGCTCCACCGAGGAGGA GCGCATTGTGGGACGCTTCAAGGGCTCCCT CTGCGTGTACAAAGTGCCACTCCCAGAGG ACGTGTCCCGGGAAGCCGGCTACGACTCC ACCTACGGCATGTTCCAGGGCATCCCGAGC AATGACCCCATCAATGTGCTGGTCCGAGTC TATGTGGTCCGGGCCACGGACCTGCACCCT GCTGACATCAACGGCAAAGCTGACCCCTA CATCGCCATCCGGCTAGGCAAGACTGACAT CCGCGACAAGGAGAACTACATCTCCAAGC AGCTCAACCCTGTCTTTGGGAAGTCCTTTG ACATCGAGGCCTCCTTCCCCATGGAATCCA TGCTGACGGTGGCTGTGTATGACTGGGACC TGGTGGGCACTGATGACCTCATTGGGGAA ACCAAGATCGACCTGGAGAACCGCTTCTAC AGCAAGCACCGCGCCACCTGCGGCATCGC CCAGACCTACTCCACACATGGCTACAATAT CTGGCGGGACCCCATGAAGCCCAGCCAGA TCCTGACCCGCCTCTGCAAAGACGGCAAA GTGGACGGCCCCCACTTTGGGCCCCCTGGG AGAGTGAAGGTGGCCAACCGCGTCTTCACT GGGCCCTCTGAGATTGAGGACGAGAACGG TCAGAGGAAGCCCACAGACGAGCATGTGG CGCTGTTGGCCCTGAGGCACTGGGAGGAC ATCCCCCGCGCAGGCTGCCGCCTGGTGCCA GAGCATGTGGAGACGAGGCCGCTGCTCAA CCCCGACAAGCCGGGCATCGAGCAGGGCC GCCTGGAGCTGTGGGTGGACATGTTCCCCA TGGACATGCCAGCCCCTGGGACGCCTCTGG ACATCTCACCTCGGAAGCCCAAGAAGTAC

[0316]

SEQ ID NO.	描述	质粒序列
		GAGCTGCGGGTCATCATCTGGAACACAGA TGAGGTGGTCTTGGAGGACGACGACTTCTT CACAGGGGAGAAGTCCAGTGACATCTTCG TGAGGGGGTGGCTGAAGGGCCAGCAGGAG GACAAGCAGGACACAGACGTCCACTACCA CTCCCTCACTGGCGAGGGCAACTTCAACTG GCGCTACCTGTTCCCCTTCGACTACCTGGC GGCGGAGGAGAAGATCGTCATCTCCAAGA AGGAGTCCATGTTCTCCTGGGACGAGACCG AGTACAAGATCCCCGCGCGGCTCACCCTGC AGATCTGGGATGCGGACCACTTCTCCGCTG ACGACTTCCTGGGGGCCATCGAGCTGGACC TGAACCGGTTCCCGCGGGGCGCAAAGACA GCCAAGCAGTGCACCATGGAGATGGCCAC CGGGGAGGTGGACGTGCCCCTCGTGTCCAT CTTCAAGCAAAAGCGCGTCAAAGGCTGGT GGCCCCTCCTGGCCCGCAATGAGAACGAT GAGTTTGAGCTCACGGGCAAGGTGGAGGC TGAGCTGCATTTACTGACAGCAGAGGAGG CAGAGAAGAACCCAGTGGGCCTGGCCCGC AATGAACCTGACCCCCTAGAGAAACCCAA CCGGCCCGACACGGCCTTCGTCTGGTTCCT CAACCCTCTCAAGTCCATCAAGTACCTCAT CTGCACCCGGTACAAGTGGCTCATCATCAA GATCGTGCTGGCGCTGTTGGGGCTGCTCAT GTTGGGGCTCTTCCTCTACAGCCTCCCTGG CTACATGGTCAAAAAGCTCCTTGGGGCATG AACGGCCGCTATGCTAGCTTGGTACCAAGG GCGGATCCTGCATAGAGCTCGCTGATCAGC CTCGACTGTGCCTTCTAGTTGCCAGCCATC TGTTGTTTGCCCCTCCCCCGTGCCTTCCTTG ACCCTGGAAGGTGCCACTCCCACTGTCCTT TCCTAATAAAATGAGGAAATTGCATCGCAT TGTCTGAGTAGGTGTCATTCTATTCTGGGG GGTGGGGTGGGGCAGGACAGCAAGGGGGA GGATTGGGAAGACAATAGCAGGCATGCTG GGGAGAGATCTGAGGACTAGTCCGTCGAC TGGGGAGAGATCTGAGGAACCCCTAGTGA TGGAGTTGGCCACTCCCTCTCTGCGCGCTC GCTCGCTCACTGAGGCCGCCCGGGCAAAG CCCGGGCGTCGGGCGACCTTTGGTCGCCCG GCCTCAGTGAGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACCCCCCCCCCCCCCCCC TGCAGCCTGGCGTAATAGCGAAGAGGCCC GCACCGATCGCCCTTCCCAACAGTTGCGTA GCCTGAATGGCGAATGGCGCGACGCGCCC TGTAGCGGCGCATTAAGCGCGGCGGGTGT GGTGGTTACGCGCAGCGTGACCGCTACACT

[0317]

SEQ ID NO.	描述	质粒序列
		TGCCAGCGCCCTAGCGCCCGCTCCTTTCGC TTTCTTCCCTTCCTTTCTCGCCACGTTCGCC GGCTTTCCCCGTCAAGCTCTAAATCGGGGG CTCCCTTTAGGGTTCCGATTTAGTGCTTTAC GGCACCTCGACCCCAAAAAACTTGATTAG GGTGATGGTTCACGTAGTGGGCCATCGCCC TGATAGACGGTTTTTCGCCCTTTGACGTTG GAGTCCACGTTCTTTAATAGTGGACTCTTG TTCCAAACTGGAACAACACTCAACCCTATC GCGGTCTATTCTTTTGATTTATAAGGGATG TTGCCGATTTCGGCCTATTGGTTAAAAAAT GAGCTGATTTAACAAAAATTTTAACAAAAT TCAGAAGAACTCGTCAAGAAGGCGATAGA AGGCGATGCGCTGCGAATCGGGAGCGGCG ATACCGTAAAGCACGAGGAAGCGGTCAGC CCATTCGCCGCCAAGCTCTTCAGCAATATC ACGGGTAGCCAACGCTATGTCCTGATAGCG GTCCGCCACACCCAGCCGGCCACAGTCGAT GAATCCAGAAAAGCGGCCATTTTCCACCAT GATATTCGGCAAGCAGGCATCGCCATGGG TCACGACGAGATCCTCGCCGTCGGGCATGC TCGCCTTGAGCCTGGCGAACAGTTCGGCTG GCGCGAGCCCCTGATGCTCTTCGTCCAGAT CATCCTGATCGACAAGACCGGCTTCCATCC GAGTACGTGCTCGCTCGATGCGATGTTTCG CTTGGTGGTCGAATGGGCAGGTAGCCGGA TCAAGCGTATGCAGCCGCCGCATTGCATCA GCCATGATGGATACTTTCTCGGCAGGAGCA AGGTGAGATGACAGGAGATCCTGCCCCGG CACTTCGCCCAATAGCAGCCAGTCCCTTCC CGCTTCAGTGACAACGTCGAGCACAGCTGC GCAAGGAACGCCCGTCGTGGCCAGCCACG ATAGCCGCGCTGCCTCGTCTTGCAGTTCAT TCAGGGCACCGGACAGGTCGGTCTTGACA AAAAGAACCGGGCGCCCCTGCGCTGACAG CCGGAACACGGCGGCATCAGAGCAGCCGA TTGTCTGTTGTGCCCAGTCATAGCCGAATA GCCTCTCCACCCAAGCGGCCGGAGAACCT GCGTGCAATCCATCTTGTTCAATCATGCGA AACGATCCTCATCCTGTCTCTTGATCAGAT CTTGATCCCCTGCGCCATCAGATCCTTGGC GGCGAGAAAGCCATCCAGTTTACTTTGCAG GGCTTCCCAACCTTACCAGAGGGCGCCCCA GCTGGCAATTCCGGTTCGCTTGCTGTCCAT AAAACCGCCCAGTCTAGCTATCGCCATGTA AGCCCACTGCAAGCTACCTGCTTTCTCTTT GCGCTTGCGTTTTCCCTTGTCCAGATAGCC CAGTAGCTGACATTCATCCGGGGTCAGCAC

[0318]

SEQ ID NO.	描述	质粒序列
		CGTTTCTGCGGACTGGCTTTCTACGTGAAA AGGATCTAGGTGAAGATCCTTTTTGATAAT CTCATGACCAAAATCCCTTAACGTGAGTTT TCGTTCCACTGAGCGTCAGACCCCGTAGAA AAGATCAAAGGATCTTCTTGAGATCCTTTT TTTCTGCGCGTAATCTGCTGCTTGCAAACA AAAAAACCACCGCTACCAGCGGTGGTTTGT TTGCCGGATCAAGAGCTACCAACTCTTTTT CCGAAGGTAACTGGCTTCAGCAGAGCGCA GATACCAAATACTGTCCTTCTAGTGTAGCC GTAGTTAGGCCACCACTTCAAGAACTCTGT AGCACCGCCTACATACCTCGCTCTGCTAAT CCTGTTACCAGTGGCTGCTGCCAGTGGCGA TAAGTCGTGTCTTACCGGGTTGGACTCAAG ACGATAGTTACCGGATAAGGCGCAGCGGT CGGGCTGAACGGGGGGTTCGTGCACACAG CCCAGCTTGGAGCGAACGACCTACACCGA ACTGAGATACCTACAGCGTGAGCTATGAG AAAGCGCCACGCTTCCCGAAGGGAGAAAG GCGGACAGGTATCCGGTAAGCGGCAGGGT CGGAACAGGAGAGCGCACGAGGGAGCTTC CAGGGGGAAACGCCTGGTATCTTTATAGTC CTGTCGGGTTTCGCCACCTCTGACTTGAGC GTCGATTTTTGTGATGCTCGTCAGGGGGGC GGAGCCTATGGAAAAACGCCAGCAACGCG GCCTTTTTACGGTTCCTGGGCTTTTGCTGGC CTTTTGCTCACATGTTCTTTCCTGCGTTATC CCCTGATTCTGTGGATAACCGTATTACCGC CTTTGAGTGAGCTGATACCGCTCGCCGCAG CCGAACGACCGAGCGCAGCGAGTCAGTGA GCGAGGAAGCGGAAGAGCGCCCAATACGC AAACCGCCTCTCCCCGCGCGTTGGCCGATT CATTAATGCAGGGCTGCA
79	含有以下特征的 5’转基因质粒: 位置 53-2027 处的λ 位置 2049-2178 处的 ITR 位置 2267-2636 处的 CMV，即增强子(smCBA 启动子的一部分) 位置 2633-2915 处的鸡 β-肌动蛋白启动子(smCBA 启动子的一部分)	TCGCGCGTTTCGGTGATGACGGTGAAAACC TCTGACACATGCAGCTCCCGGATAGAGGTC ATCCTTCCTGACCATTTCCATCATTCCAGTC GAACTCACACACAACACCAAATGCATTTA AGTCGCTTGAAATTGCTATAAGCAGAGCAT GTTGCGCCAGCATGATTAATACAGCATTTA ATACAGAGCCGTGTTTATTGAGTCGGTATT CAGAGTCTGACCAGAAATTATTAATCTGGT GAAGTTATTCCTCTGTCATTACGTCATGGT CGATTTCAATTTCTATTGATGCTTTCCAGTC GTAATCAATGATGTATTTTTTGATGTTTGA CCTCTGTTCATATCCTCACAGATAAAAAAT CGCCCTCACACTGGAGGGCAAAGAAGATT TCCAATAATCAGAACAAGTCGGCTCCTGTT TAGTTACGAGCGACATTGCTCCGTGTATTC

[0319]

SEQ ID NO.	描述	质粒序列
	位置 2916-3008 处的外显子 1 (smCBA 启动子的一部分) 位置 3008-3209 处的嵌合内含子 (smCBA 启动子的一部分) 位置 3226-3235 处的 Kozak 序列 位置 3232-5637 处的人类 OTOF 同功型 5 的 N 末端部分 位置 5638-5721 处的剪接供体序列 位置 5728-6014 处的 AP 头部序列 位置 6108-6237 处的 ITR 位置 6248-8260 处的 λ 位置 8317-8905 处的 Ori 位置 9083-9892 处的 KanR 位置 2049-6237 处的待转移至双载体系统中的载体中的转基因	ACTCGTTGGAATGAATACACAGTGCAGTGT TTATTCTGTTATTTATGCCAAAAATTAAGG CCACTATCAGGCAGCTTTGTTGTTCTGTTTA CCAAGTTCTCTGGCAATCATTGCCGTCGTT CGTATTGCCCATTTATCGACATATTTCCCAT CTTCCTATACAGGAAACATTTCTTCAGGCT TAACCATGCATTCCGATTGCAGCTTGCATC CATTGCATCGCTTGAATTGTCCACACCATT GATTTTTATCAATAGTCGTAGTTTAACGGA TAGTCCTGGTATTGTTCCATCACATCCTGA GGATGCCCTTCGAACTCTTCAAATTCTTCTT CCTAATATCACCTTAAATAGTGGATTGCGG TAGTAAAGATTGTGCCTGTCTTTTAACCAC ATCAGGCTCGGTGGTTCTCGTGTACCCCTA CAGCGAGAAATCGGATAAACTATTACAAC CCCTACAGTTTGTAGAGTATAGAAAATGAT CCACTCGTTATTCTCGGACGAGTGTTCAGT AATGAACCTCTGGAGAGAACCATCTATATG ATCGTTATCTGGGTTTGACTTCTGCTTTTAA GCCCAGATAACTTGCCTGAATATGTTAATG AGAGAATCGGTATTCCTCATGTGTGGCATG TTTTCGTCTTTGCTCTTGCATTTTCACTAGC AATTAATGTGCATCGATTATCAGCTATTGC CAGCGCCAGATATAAGCGATTTAAGCTAA GAAAACGCATTAAGGTGCAAAACGATAAA GTGCGATCAGTAATTCAAAACCTTACAGGA GAGCAATCTATGGTTTTGTGCTCAGCCCTT AATGAAGGCAGGTAGTATGTGGTTACATC AAAACAATTCCCATACATTAGTGAGTTGAT TGAGCTTGGTGTGTTGAACAAAACTTTTTC CCGATGGAATGGAAAGCATATATTATTCCC TATTGAGGATATTTACTGGACTGAATTAGT TGCCAGCTATGATCCATATAATATTGAGAT AAAGCCAAGGCCAATATCTAAGTAACTAG ATAAGAGGAATCGATTTTCCCTTAATTTTC TGGCGTCCACTGCATGTTATGCCGCGTTCG CCAGGCTTGCTGTACCATGTGCGCTGATTC TTGCGCTCAATACGTTGCAGGTTGCTTTCA ATCTGTTTGTGGTATTCAGCCAGCACTGTA AGGTCTATCGGATTTAGTGCGCTTTCTACT CGTGATTTCGGTTTGCGATTCAGCGAGAGA ATAGGGCGGTTAACTGGTTTTGCGCTTACC CCAACCAACAGGGGATTTGCTGCTTTCCAT TGAGCCTGTTACTCTGCGCGACGTTCGCGG CGGCGTGTTTGTGCATCCATCTGGATTCTC CTGTCAGTTAGCTTTGGTGGTGTGTGGCAG TTGTAGTCCTGAACGAAAACCCCCCGCGAT TGGCACGTTGGCAGCTAATCCGGAATCGCA

[0320]

SEQ ID NO.	描述	质粒序列
		CTTACGGCCAATGCTTCGTTTCGTATCACA CACCCCAAAGCCTTCTGCTTTGAATGCTGC CCTTCTTCAGGGCTTAATTTTTAAGAGCGT CACCTTCATGGTGGTCAGTGCGTCCTGCTG ATGTGCTCAGGCACGATTTAATTAAGGCCT TAATTAGGCTGCGCGCTCGCTCGCTCACTG AGGCCGCCCGGGCAAAGCCCGGGCGTCGG GCGACCTTTGGTCGCCCGGCCTCAGTGAGC GAGCGAGCGCGCAGAGAGGGAGTGGCCAA CTCCATCACTAGGGGTTCCTTGTAGTTAAT GATTAACCCGCCATGCTACTTATCTACGTA GCCATGCTCTAGGAAGATCGGAATTCGCCC TTAAGCTAGCGGCGCGCCGGTACCTAGTTA TTAATAGTAATCAATTACGGGGTCATTAGT TCATAGCCCATATATGGAGTTCCGCGTTAC ATAACTTACGGTAAATGGCCCGCCTGGCTG ACCGCCCAACGACCCCCGCCCATTGACGTC AATAATGACGTATGTTCCCATAGTAACGCC AATAGGGACTTTCCATTGACGTCAATGGGT GGAGTATTTACGGTAAACTGCCCACTTGGC AGTACATCAAGTGTATCATATGCCAAGTAC GCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGAC CTTATGGGACTTTCCTACTTGGCAGTACAT CTACGTATTAGTCATCGCTATTACCATGGT CGAGGTGAGCCCCACGTTCTGCTTCACTCT CCCCATCTCCCCCCCCTCCCCACCCCCAAT TTTGTATTTATTTATTTTTTAATTATTTTGTG CAGCGATGGGGGCGGGGGGGGGGGGGGGG GCGCGCGCCAGGCGGGGCGGGGCGGGGCG AGGGGCGGGGCGGGGCGAGGCGGAGAGG TGCGGCGGCAGCCAATCAGAGCGGCGCGC TCCGAAAGTTTCCTTTTATGGCGAGGCGGC GGCGGCGGCGGCCCTATAAAAAGCGAAGC GCGCGGCGGGCGGGAGTCGCTGCGCGCTG CCTTCGCCCCGTGCCCCGCTCCGCCGCCGC CTCGCGCCGCCCGCCCCGGCTCTGACTGAC CGCGTTACTCCCACAGGTGAGCGGGCGGG ACGGCCCTTCTCCTCCGGGCTGTAATTAGC GCTTGGTTTAATGACGGCTTGTTTCTTTTCT GTGGCTGCGTGAAAGCCTTGAGGGGCTCC GGGAGCTAGAGCCTCTGCTAACCATGTTCA TGCCTTCTTCTTTTTCCTACAGCTCCTGGGC AACGTGCTGGTTATTGTGCTGTCTCATCAT TTTGGCAAAGAATTCTAGCGGCCGCCACCA TGGCCTTGCTCATCCACCTCAAGACAGTCT CGGAGCTGCGGGGCAGGGGCGACCGGATC GCCAAAGTGACTTTCCGAGGGCAATCCTTC

[0321]

SEQ ID NO.	描述	质粒序列
		TACTCTCGGGTCCTGGAGAACTGTGAGGAT GTGGCTGACTTTGATGAGACATTTCGGTGG CCGGTGGCCAGCAGCATCGACAGAAATGA GATGCTGGAGATTCAGGTTTTCAACTACAG CAAAGTCTTCAGCAACAAGCTCATCGGGA CCTTCCGCATGGTGCTGCAGAAGGTGGTAG AGGAGAGCCATGTGGAGGTGACTGACACG CTGATTGATGACAACAATGCTATCATCAAG ACCAGCCTGTGCGTGGAGGTCCGGTATCAG GCCACTGACGGCACAGTGGGCTCCTGGGA CGATGGGGACTTCCTGGGAGATGAGTCTCT TCAAGAGGAAGAGAAGGACAGCCAAGAG ACGGATGGACTGCTCCCAGGCTCCCGGCCC AGCTCCCGGCCCCCAGGAGAGAAGAGCTT CCGGAGAGCCGGGAGGAGCGTGTTCTCCG CCATGAAGCTCGGCAAAAACCGGTCTCAC AAGGAGGAGCCCCAAAGACCAGATGAACC GGCGGTGCTGGAGATGGAAGACCTTGACC ATCTGGCCATTCGGCTAGGAGATGGACTGG ATCCCGACTCGGTGTCTCTAGCCTCAGTCA CAGCTCTCACCACTAATGTCTCCAACAAGC GATCTAAGCCAGACATTAAGATGGAGCCA AGTGCTGGGCGGCCCATGGATTACCAGGTC AGCATCACGGTGATCGAGGCCCGGCAGCT GGTGGGCTTGAACATGGACCCTGTGGTGTG CGTGGAGGTGGGTGACGACAAGAAGTACA CATCCATGAAGGAGTCCACTAACTGCCCCT ATTACAACGAGTACTTCGTCTTCGACTTCC ATGTCTCTCCGGATGTCATGTTTGACAAGA TCATCAAGATTTCGGTGATTCACTCCAAGA ACCTGCTGCGCAGTGGCACCCTGGTGGGCT CCTTCAAAATGGACGTGGGAACCGTGTACT CGCAGCCAGAGCACCAGTTCCATCACAAG TGGGCCATCCTGTCTGACCCCGATGACATC TCCTCGGGGCTGAAGGGCTACGTGAAGTGT GACGTTGCCGTGGTGGGCAAAGGGGACAA CATCAAGACGCCCCACAAGGCCAATGAGA CCGACGAAGATGACATTGAGGGGAACTTG CTGCTCCCCGAGGGGGTGCCCCCCGAACGC CAGTGGGCCCGGTTCTATGTGAAAATTTAC CGAGCAGAGGGGCTGCCCCGTATGAACAC AAGCCTCATGGCCAATGTAAAGAAGGCTTT CATCGGTGAAAACAAGGACCTCGTGGACC CCTACGTGCAAGTCTTCTTTGCTGGCCAGA AGGGCAAGACTTCAGTGCAGAAGAGCAGC TATGAGCCCCTGTGGAATGAGCAGGTCGTC TTTACAGACCTCTTCCCCCCACTCTGCAAA CGCATGAAGGTGCAGATCCGAGACTCGGA

[0322]

SEQ ID NO.	描述	质粒序列
		CAAGGTCAACGACGTGGCCATCGGCACCC ACTTCATTGACCTGCGCAAGATTTCTAATG ACGGAGACAAAGGCTTCCTGCCCACACTG GGCCCAGCCTGGGTGAACATGTACGGCTCC ACACGTAACTACACGCTGCTGGATGAGCAT CAGGACCTGAACGAGGGCCTGGGGGAGGG TGTGTCCTTCCGGGCCCGGCTCCTGCTGGG CCTGGCTGTGGAGATCGTAGACACCTCCAA CCCTGAGCTCACCAGCTCCACAGAGGTGCA GGTGGAGCAGGCCACGCCCATCTCGGAGA GCTGTGCAGGTAAAATGGAAGAATTCTTTC TCTTTGGAGCCTTCCTGGAGGCCTCAATGA TCGACCGGAGAAACGGAGACAAGCCCATC ACCTTTGAGGTCACCATAGGCAACTATGGG AACGAAGTTGATGGCCTGTCCCGGCCCCAG CGGCCTCGGCCCCGGAAGGAGCCGGGGGA TGAGGAAGAAGTAGACCTGATTCAGAACG CAAGTGATGACGAGGCCGGTGATGCCGGG GACCTGGCCTCAGTCTCCTCCACTCCACCA ATGCGGCCCCAGGTCACCGACAGGAACTA CTTCCATCTGCCCTACCTGGAGCGAAAGCC CTGCATCTACATCAAGAGCTGGTGGCCGGA CCAGCGCCGCCGCCTCTACAATGCCAACAT CATGGACCACATTGCCGACAAGCTGGAAG AAGGCCTGAACGACATACAGGAGATGATC AAAACGGAGAAGTCCTACCCTGAGCGTCG CCTGCGGGGCGTCCTGGAGGAGCTGAGCT GTGGCTGCTGCCGCTTCCTCTCCCTCGCTG ACAAGGACCAGGGCCACTCATCCCGCACC AGGCTTGACCGGGAGCGCCTCAAGTCCTGC ATGAGGGAGCTGGTAAGTATCAAGGTTAC AAGACAGGTTTAAGGAGACCAATAGAAAC TGGGCTTGTCGAGACAGAGAAGACTCTTGC GTTTCTGAGCTAGCCCCCGGGTGCGCGGCG TCGGTGGTGCCGGCGGGGGGCGCCAGGTC GCAGGCGGTGTAGGGCTCCAGGCAGGCGG CGAAGGCCATGACGTGCGCTATGAAGGTC TGCTCCTGCACGCCGTGAACCAGGTGCGCC TGCGGGCCGCGCGCGAACACCGCCACGTC CTCGCCTGCGTGGGTCTCTTCGTCCAGGGG CACTGCTGACTGCTGCCGATACTCGGGGCT CCCGCTCTCGCTCTCGGTAACATCCGGCCG GGCGCCGTCCTTGAGCACATAGCCTGGACC GTTTCGTCGACCTCGAGTTAAGGGCGAATT CCCGATAAGGATCTTCCTAGAGCATGGCTA CGTAGATAAGTAGCATGGCGGGTTAATCAT TAACTACAAGGAACCCCTAGTGATGGAGTT GGCCACTCCCTCTCTGCGCGCTCGCTCGCT

[0323]

SEQ ID NO.	描述	质粒序列
		CACTGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCTCAG TGAGCGAGCGAGCGCGCAGCCTTAATTAA ATCCACATCTGTATGTTTTTTATATTAATTT ATTTTTTGCAGGGGGGCATTGTTTGGTAGG TGAGAGTTCTGAATTGCTATGTTTAGTGAG TTGTATCTATTTATTTTTCAATAAATACAAT TAGTTATGTGTTTTGGGGGCGATCGTGAGG CAAAGAAAACCCGGCGCTGAGGCCGGGTT ATTCTTGTTCTCTGGTCAAATTATATAGTTG GAAAACAAGGATGCATATATGAATGAACG ATGCAGAGGCAATGCCGATGGCGATAGTG GGTATCAGGTAGCCGCTTATGCTGGAAAG AAGCAATAACCCGCAGAAAAACAAAGCTC CAAGCTCAACAAAACTAAGGGCATAGACA ATAACTACCTATGTCATATACCCATACTCT CTAATCTTGGCCAGTCGGCGCGTTCTGCTT CCGATTAGAAACGTCAAGGCAGCAATCAG GATTGCAATCTTGGTTCCTGCATAGGATGA CAATGTCGCCCCAAGACCATCTCTATGAGC TGAAAAAGAAACACAAGGAATGTAGTGGC GGAAAAGGAGATAGCAAATGCTTACGATA ACGTAAGGAATTATTACTATGTAAACACCA GGCAAGATTCTGTTCCGTATAATTACTCCT GATAATTAATCCTTAACTTTGCCCACCTGC CTTTTAAAACATTCCAGTATATCACTTTTCA TTCTTGCGTAGCAATATGCCCTCTCTTCAG CTATCTCAGCATTGGTGACCTTGTTCAGAG GCGCTGAGAGATGGCCTTTTTCTGATAGAT AATGTTCTGTTAAAATATCTCCGGCCTCAT CTTTTGCCCGCAGGCTAATGTCTGAAAATT GAGGTGACGGGTTAAAAATAATATCCTTG GCAACCTTTTTTATATCCCTTTTAAATTTTG GCTTAATGACTATATCCAATGAGTCAAAAA GCTCCCCTTCAATATCTGTTGCCCCTAAGA CCTTTAATATATCGCCAAATACAGGTAGCT TGGCTTCTACCTTCACCGTTGTTCTGCCGAT GAAATGCTAATGCATAACATCGTCTTTGGT GGTTCCCCTCATCAGTGGCTCTATCTGAAC GCGCTCTCCACTGCTTAATGACATTCCTTTC CCGATTAAAAAATCTGTCAGATCGGATGTG GTCGGCCCGAAAACAGTTCTGGCAAAACC AATGGTGTCGCCTTCAACAAACAAAAAAG ATGGGAATCCCAATGATTCGTCATCTGCGA GGCTGTTCTTAATATCTTCAACTGTAGCTTT AGAGCGATTTATCTTCTGAACCAGACTCTT GTCATTTGTTTTGGTAAAGAGAAAAGTTTT TCCATCGATTTTATGAATATACAAATAATT

[0324]

SEQ ID NO.	描述	质粒序列
		GGAGCCAACCTTCAGGTGATGATTATCAGC CAGCAGAGAATTAAGGAAAACAGACAGGT TTATTGAGCACTTATCTTTCCCTTTATTTTT GCTGCGGTAAGTCGCATAAAAACCATTCTT CACAATTCAATCCATTTACTATGTTATGTTC TGAGGGGAGTGAAAATTCCCCTAATTCGAT GAAGATTCTTGCTAAATTGTTATCAGCTAT GCGCCGACCAGAACACCTTGCCGATCAGC CAAACGTCTAATCAGGCCACTGACTAGCG ATAACTTTCCCCACAACGGAACAACTCTCA TTGCATGGGATAATTGGGTACTGTGGGTTT AGTGGTTGTAAAAACACCTGACCGCTATCC CTGATCAGTTTCTTGAAGGTAAACTCATCA CCCCCAAGTCTGGCTATACAGAAATCACCT GGCTCAACAGCCTGCTCAGGGTCAACGAG AATTTACATTCCGTCAGGATAGCTTGGCTT GGAGCCTGTTGGTGCGGTCACGGAATTACC TTCAACCTCAAGCCAGAATGCAGAATCACT GGCTTTTTTGGTTGTGCTTACCCATCTCTCC GCATCACCTTTGGTAAAGGTTCTAAGCTAA GGTGAGAACATCCCTGCCTGAACATGAGA AAAAACAGGGTACTCATACTCACTTATTAG TGACGGCTATGAGCAAAAGGCCAGCAAAA GGCCAGGAACCGTAAAAAGGCCGCGTTGC TGGCGTTTTTCCATAGGCTCCGCCCCCCTG ACGAGCATCACAAAAATCGACGCTCAAGT CAGAGGTGGCGAAACCCGACAGGACTATA AAGATACCAGGCGTTTCCCCCTGGAAGCTC CCTCGTGCGCTCTCCTGTTCCGACCCTGCC GCTTACCGGATACCTGTCCGCCTTTCTCCCT TCGGGAAGCGTGGCGCTTTCTCATAGCTCA CGCTGTAGGTATCTCAGTTCGGTGTAGGTC GTTCGCTCCAAGCTGGGCTGTGTGCACGAA CCCCCCGTTCAGCCCGACCGCTGCGCCTTA TCCGGTAACTATCGTCTTGAGTCCAACCCG GTAAGACACGACTTATCGCCACTGGCAGC AGCCACTGGTAACAGGATTAGCAGAGCGA GGTATGTAGGCGGTGCTACAGAGTTCTTGA AGTGGTGGCCTAACTACGGCTACACTAGA AGAACAGTATTTGGTATCTGCGCTCTGCTG AAGCCAGTTACCTTCGGAAAAAGAGTTGG TAGCTCTTGATCCGGCAAACAAACCACCGC TGGTAGCGGTGGTTTTTTTGTTTGCAAGCA GCAGATTACGCGCAGAAAAAAAGGATCTC AAGAAGATCCTTTGATCTTTTCTACGGGGT CTGACGCTCAGTGGAACGAAAACTCACGTT AAGGGATTTTGGTCATGAGATTATCAAAAA GGATCTTCACCTAGATCCTTTTAAATTAAA

[0325]

SEQ ID NO.	描述	质粒序列
		AATGAAGTTTTAAATCAAGCCCAATCTGAA TAATGTTACAACCAATTAACCAATTCTGAT TAGAAAAACTCATCGAGCATCAAATGAAA CTGCAATTTATTCATATCAGGATTATCAAT ACCATATTTTTGAAAAAGCCGTTTCTGTAA TGAAGGAGAAAACTCACCGAGGCAGTTCC ATAGGATGGCAAGATCCTGGTATCGGTCTG CGATTCCGACTCGTCCAACATCAATACAAC CTATTAATTTCCCCTCGTCAAAAATAAGGT TATCAAGTGAGAAATCACCATGAGTGACG ACTGAATCCGGTGAGAATGGCAAAAGTTT ATGCATTTCTTTCCAGACTTGTTCAACAGG CCAGCCATTACGCTCGTCATCAAAATCACT CGCATCAACCAAACCGTTATTCATTCGTGA TTGCGCCTGAGCAAGACGAAATACGCGAT CGCTGTTAAAAGGACAATTACAAACAGGA ATCGAATGCAACCGGCGCAGGAACACTGC CAGCGCATCAACAATATTTTCACCTGAATC AGGATATTCTTCTAATACCTGGAATGCTGT TTTTCCGGGGATCGCAGTGGTGAGTAACCA TGCATCATCAGGAGTACGGATAAAATGCTT GATGGTCGGAAGAGGCATAAATTCCGTCA GCCAGTTTAGTCTGACCATCTCATCTGTAA CATCATTGGCAACGCTACCTTTGCCATGTT TCAGAAACAACTCTGGCGCATCGGGCTTCC CATACAAGCGATAGATTGTCGCACCTGATT GCCCGACATTATCGCGAGCCCATTTATACC CATATAAATCAGCATCCATGTTGGAATTTA ATCGCGGCCTCGACGTTTCCCGTTGAATAT GGCTCATAACACCCCTTGTATTACTGTTTA TGTAAGCAGACAGTTTTATTGTTCATGATG ATATATTTTTATCTTGTGCAATGTAACATC AGAGATTTTGAGACACGGGCCAGAGCTGC A
80	含有以下特征的 5’转基因质粒: 位置 19-161 处的 ITR 位置 177-546 处的 CMV 增强子 (smCBA 启动子的一部分) 位置 548-825 处的鸡 β-肌动蛋白启动子(smCBA 启动子的一部分)	GGGGGGGGGGGGGGGGGGTTGGCCACTCC CTCTCTGCGCGCTCGCTCGCTCACTGAGGC CGGGCGACCAAAGGTCGCCCGACGCCCGG GCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCAGAGAGGGAGTGGCCAACTCC ATCACTAGGGGTTCCTCAGATCTGAATTCG GTACCTAGTTATTAATAGTAATCAATTACG GGGTCATTAGTTCATAGCCCATATATGGAG TTCCGCGTTACATAACTTACGGTAAATGGC CCGCCTGGCTGACCGCCCAACGACCCCCGC CCATTGACGTCAATAATGACGTATGTTCCC ATAGTAACGCCAATAGGGACTTTCCATTGA CGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCAT

[0326]

SEQ ID NO.	描述	质粒序列
	位置 826-918 处的外显子 1 (smCBA 启动子的一部分) 位置 918-1119 处的嵌合内含子 (smCBA 启动子的一部分) 位置 1142-3547 处的人类 OTOF 同功型 5 的 N 末端部分 位置 3548-3631 处的剪接供体序列 位置 3638-3924 处的 AP 头部序列 位置 3948-4090 处的 ITR 位置 4616-5410 处的 KanR 位置 5796-6466 处的 pUC ori 位置 19-4090 处的待转移至双载体系统中的载体中的转基因	ATGCCAAGTACGCCCCCTATTGACGTCAAT GACGGTAAATGGCCCGCCTGGCATTATGCC CAGTACATGACCTTATGGGACTTTCCTACT TGGCAGTACATCTACGTATTAGTCATCGCT ATTACCATGGTCGAGGTGAGCCCCACGTTC TGCTTCACTCTCCCCATCTCCCCCCCCTCCC CACCCCCAATTTTGTATTTATTTATTTTTTA ATTATTTTGTGCAGCGATGGGGGCGGGGG GGGGGGGGGGGCGCGCGCCAGGCGGGGCG GGGCGGGGCGAGGGGCGGGGCGGGGCGA GGCGGAGAGGTGCGGCGGCAGCCAATCAG AGCGGCGCGCTCCGAAAGTTTCCTTTTATG GCGAGGCGGCGGCGGCGGCGGCCCTATAA AAAGCGAAGCGCGCGGCGGGCGGGAGTCG CTGCGCGCTGCCTTCGCCCCGTGCCCCGCT CCGCCGCCGCCTCGCGCCGCCCGCCCCGGC TCTGACTGACCGCGTTACTCCCACAGGTGA GCGGGCGGGACGGCCCTTCTCCTCCGGGCT GTAATTAGCGCTTGGTTTAATGACGGCTTG TTTCTTTTCTGTGGCTGCGTGAAAGCCTTG AGGGGCTCCGGGAGCTAGAGCCTCTGCTA ACCATGTTCATGCCTTCTTCTTTTTCCTACA GCTCCTGGGCAACGTGCTGGTTATTGTGCT GTCTCATCATTTTGGCAAAGAATTCTAGCG GCCGCCACCATGGCCTTGCTCATCCACCTC AAGACAGTCTCGGAGCTGCGGGGCAGGGG CGACCGGATCGCCAAAGTGACTTTCCGAG GGCAATCCTTCTACTCTCGGGTCCTGGAGA ACTGTGAGGATGTGGCTGACTTTGATGAGA CATTTCGGTGGCCGGTGGCCAGCAGCATCG ACAGAAATGAGATGCTGGAGATTCAGGTT TTCAACTACAGCAAAGTCTTCAGCAACAAG CTCATCGGGACCTTCCGCATGGTGCTGCAG AAGGTGGTAGAGGAGAGCCATGTGGAGGT GACTGACACGCTGATTGATGACAACAATG CTATCATCAAGACCAGCCTGTGCGTGGAGG TCCGGTATCAGGCCACTGACGGCACAGTG GGCTCCTGGGACGATGGGGACTTCCTGGG AGATGAGTCTCTTCAAGAGGAAGAGAAGG ACAGCCAAGAGACGGATGGACTGCTCCCA GGCTCCCGGCCCAGCTCCCGGCCCCCAGGA GAGAAGAGCTTCCGGAGAGCCGGGAGGAG CGTGTTCTCCGCCATGAAGCTCGGCAAAAA CCGGTCTCACAAGGAGGAGCCCCAAAGAC CAGATGAACCGGCGGTGCTGGAGATGGAA GACCTTGACCATCTGGCCATTCGGCTAGGA GATGGACTGGATCCCGACTCGGTGTCTCTA GCCTCAGTCACAGCTCTCACCACTAATGTC

[0327]

SEQ ID NO.	描述	质粒序列
		TCCAACAAGCGATCTAAGCCAGACATTAA GATGGAGCCAAGTGCTGGGCGGCCCATGG ATTACCAGGTCAGCATCACGGTGATCGAG GCCCGGCAGCTGGTGGGCTTGAACATGGA CCCTGTGGTGTGCGTGGAGGTGGGTGACG ACAAGAAGTACACATCCATGAAGGAGTCC ACTAACTGCCCCTATTACAACGAGTACTTC GTCTTCGACTTCCATGTCTCTCCGGATGTC ATGTTTGACAAGATCATCAAGATTTCGGTG ATTCACTCCAAGAACCTGCTGCGCAGTGGC ACCCTGGTGGGCTCCTTCAAAATGGACGTG GGAACCGTGTACTCGCAGCCAGAGCACCA GTTCCATCACAAGTGGGCCATCCTGTCTGA CCCCGATGACATCTCCTCGGGGCTGAAGGG CTACGTGAAGTGTGACGTTGCCGTGGTGGG CAAAGGGGACAACATCAAGACGCCCCACA AGGCCAATGAGACCGACGAAGATGACATT GAGGGGAACTTGCTGCTCCCCGAGGGGGT GCCCCCCGAACGCCAGTGGGCCCGGTTCTA TGTGAAAATTTACCGAGCAGAGGGGCTGC CCCGTATGAACACAAGCCTCATGGCCAATG TAAAGAAGGCTTTCATCGGTGAAAACAAG GACCTCGTGGACCCCTACGTGCAAGTCTTC TTTGCTGGCCAGAAGGGCAAGACTTCAGTG CAGAAGAGCAGCTATGAGCCCCTGTGGAA TGAGCAGGTCGTCTTTACAGACCTCTTCCC CCCACTCTGCAAACGCATGAAGGTGCAGA TCCGAGACTCGGACAAGGTCAACGACGTG GCCATCGGCACCCACTTCATTGACCTGCGC AAGATTTCTAATGACGGAGACAAAGGCTT CCTGCCCACACTGGGCCCAGCCTGGGTGAA CATGTACGGCTCCACACGTAACTACACGCT GCTGGATGAGCATCAGGACCTGAACGAGG GCCTGGGGGAGGGTGTGTCCTTCCGGGCCC GGCTCCTGCTGGGCCTGGCTGTGGAGATCG TAGACACCTCCAACCCTGAGCTCACCAGCT CCACAGAGGTGCAGGTGGAGCAGGCCACG CCCATCTCGGAGAGCTGTGCAGGTAAAAT GGAAGAATTCTTTCTCTTTGGAGCCTTCCT GGAGGCCTCAATGATCGACCGGAGAAACG GAGACAAGCCCATCACCTTTGAGGTCACCA TAGGCAACTATGGGAACGAAGTTGATGGC CTGTCCCGGCCCCAGCGGCCTCGGCCCCGG AAGGAGCCGGGGGATGAGGAAGAAGTAG ACCTGATTCAGAACGCAAGTGATGACGAG GCCGGTGATGCCGGGGACCTGGCCTCAGTC TCCTCCACTCCACCAATGCGGCCCCAGGTC ACCGACAGGAACTACTTCCATCTGCCCTAC

[0328]

SEQ ID NO.	描述	质粒序列
		CTGGAGCGAAAGCCCTGCATCTACATCAA GAGCTGGTGGCCGGACCAGCGCCGCCGCC TCTACAATGCCAACATCATGGACCACATTG CCGACAAGCTGGAAGAAGGCCTGAACGAC ATACAGGAGATGATCAAAACGGAGAAGTC CTACCCTGAGCGTCGCCTGCGGGGCGTCCT GGAGGAGCTGAGCTGTGGCTGCTGCCGCTT CCTCTCCCTCGCTGACAAGGACCAGGGCCA CTCATCCCGCACCAGGCTTGACCGGGAGCG CCTCAAGTCCTGCATGAGGGAGCTGGTAA GTATCAAGGTTACAAGACAGGTTTAAGGA GACCAATAGAAACTGGGCTTGTCGAGACA GAGAAGACTCTTGCGTTTCTGAGCTAGCCC CCGGGTGCGCGGCGTCGGTGGTGCCGGCG GGGGGCGCCAGGTCGCAGGCGGTGTAGGG CTCCAGGCAGGCGGCGAAGGCCATGACGT GCGCTATGAAGGTCTGCTCCTGCACGCCGT GAACCAGGTGCGCCTGCGGGCCGCGCGCG AACACCGCCACGTCCTCGCCTGCGTGGGTC TCTTCGTCCAGGGGCACTGCTGACTGCTGC CGATACTCGGGGCTCCCGCTCTCGCTCTCG GTAACATCCGGCCGGGCGCCGTCCTTGAGC ACATAGCCTGGACCGTTTCGTCGACTGGGG AGAGATCTGAGGAACCCCTAGTGATGGAG TTGGCCACTCCCTCTCTGCGCGCTCGCTCG CTCACTGAGGCCGCCCGGGCAAAGCCCGG GCGTCGGGCGACCTTTGGTCGCCCGGCCTC AGTGAGCGAGCGAGCGCGCAGAGAGGGAG TGGCCAACCCCCCCCCCCCCCCCCCTGCAG CCTGGCGTAATAGCGAAGAGGCCCGCACC GATCGCCCTTCCCAACAGTTGCGTAGCCTG AATGGCGAATGGCGCGACGCGCCCTGTAG CGGCGCATTAAGCGCGGCGGGTGTGGTGG TTACGCGCAGCGTGACCGCTACACTTGCCA GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTT CCCTTCCTTTCTCGCCACGTTCGCCGGCTTT CCCCGTCAAGCTCTAAATCGGGGGCTCCCT TTAGGGTTCCGATTTAGTGCTTTACGGCAC CTCGACCCCAAAAAACTTGATTAGGGTGAT GGTTCACGTAGTGGGCCATCGCCCTGATAG ACGGTTTTTCGCCCTTTGACGTTGGAGTCC ACGTTCTTTAATAGTGGACTCTTGTTCCAA ACTGGAACAACACTCAACCCTATCGCGGTC TATTCTTTTGATTTATAAGGGATGTTGCCG ATTTCGGCCTATTGGTTAAAAAATGAGCTG ATTTAACAAAAATTTTAACAAAATTCAGAA GAACTCGTCAAGAAGGCGATAGAAGGCGA TGCGCTGCGAATCGGGAGCGGCGATACCG

[0329]

SEQ ID NO.	描述	质粒序列
		TAAAGCACGAGGAAGCGGTCAGCCCATTC GCCGCCAAGCTCTTCAGCAATATCACGGGT AGCCAACGCTATGTCCTGATAGCGGTCCGC CACACCCAGCCGGCCACAGTCGATGAATC CAGAAAAGCGGCCATTTTCCACCATGATAT TCGGCAAGCAGGCATCGCCATGGGTCACG ACGAGATCCTCGCCGTCGGGCATGCTCGCC TTGAGCCTGGCGAACAGTTCGGCTGGCGCG AGCCCCTGATGCTCTTCGTCCAGATCATCC TGATCGACAAGACCGGCTTCCATCCGAGTA CGTGCTCGCTCGATGCGATGTTTCGCTTGG TGGTCGAATGGGCAGGTAGCCGGATCAAG CGTATGCAGCCGCCGCATTGCATCAGCCAT GATGGATACTTTCTCGGCAGGAGCAAGGT GAGATGACAGGAGATCCTGCCCCGGCACT TCGCCCAATAGCAGCCAGTCCCTTCCCGCT TCAGTGACAACGTCGAGCACAGCTGCGCA AGGAACGCCCGTCGTGGCCAGCCACGATA GCCGCGCTGCCTCGTCTTGCAGTTCATTCA GGGCACCGGACAGGTCGGTCTTGACAAAA AGAACCGGGCGCCCCTGCGCTGACAGCCG GAACACGGCGGCATCAGAGCAGCCGATTG TCTGTTGTGCCCAGTCATAGCCGAATAGCC TCTCCACCCAAGCGGCCGGAGAACCTGCGT GCAATCCATCTTGTTCAATCATGCGAAACG ATCCTCATCCTGTCTCTTGATCAGATCTTGA TCCCCTGCGCCATCAGATCCTTGGCGGCGA GAAAGCCATCCAGTTTACTTTGCAGGGCTT CCCAACCTTACCAGAGGGCGCCCCAGCTG GCAATTCCGGTTCGCTTGCTGTCCATAAAA CCGCCCAGTCTAGCTATCGCCATGTAAGCC CACTGCAAGCTACCTGCTTTCTCTTTGCGCT TGCGTTTTCCCTTGTCCAGATAGCCCAGTA GCTGACATTCATCCGGGGTCAGCACCGTTT CTGCGGACTGGCTTTCTACGTGAAAAGGAT CTAGGTGAAGATCCTTTTTGATAATCTCAT GACCAAAATCCCTTAACGTGAGTTTTCGTT CCACTGAGCGTCAGACCCCGTAGAAAAGA TCAAAGGATCTTCTTGAGATCCTTTTTTTCT GCGCGTAATCTGCTGCTTGCAAACAAAAA AACCACCGCTACCAGCGGTGGTTTGTTTGC CGGATCAAGAGCTACCAACTCTTTTTCCGA AGGTAACTGGCTTCAGCAGAGCGCAGATA CCAAATACTGTCCTTCTAGTGTAGCCGTAG TTAGGCCACCACTTCAAGAACTCTGTAGCA CCGCCTACATACCTCGCTCTGCTAATCCTG TTACCAGTGGCTGCTGCCAGTGGCGATAAG TCGTGTCTTACCGGGTTGGACTCAAGACGA

[0330]

SEQ ID NO.	描述	质粒序列
		TAGTTACCGGATAAGGCGCAGCGGTCGGG CTGAACGGGGGGTTCGTGCACACAGCCCA GCTTGGAGCGAACGACCTACACCGAACTG AGATACCTACAGCGTGAGCTATGAGAAAG CGCCACGCTTCCCGAAGGGAGAAAGGCGG ACAGGTATCCGGTAAGCGGCAGGGTCGGA ACAGGAGAGCGCACGAGGGAGCTTCCAGG GGGAAACGCCTGGTATCTTTATAGTCCTGT CGGGTTTCGCCACCTCTGACTTGAGCGTCG ATTTTTGTGATGCTCGTCAGGGGGGCGGAG CCTATGGAAAAACGCCAGCAACGCGGCCT TTTTACGGTTCCTGGGCTTTTGCTGGCCTTT TGCTCACATGTTCTTTCCTGCGTTATCCCCT GATTCTGTGGATAACCGTATTACCGCCTTT GAGTGAGCTGATACCGCTCGCCGCAGCCG AACGACCGAGCGCAGCGAGTCAGTGAGCG AGGAAGCGGAAGAGCGCCCAATACGCAAA CCGCCTCTCCCCGCGCGTTGGCCGATTCAT TAATGCAGGGCTGCA
81	含有以下特征的 5’转基因质粒: 位置 12-141 处的 ITR 位置 235-1199 处的 Myo15 启动子 位置 1216-1225 处的 Kozak 序列 位置 1222-3627 处的人类 OTOF 同功型 5 的 N 末端部分 位置 3628-3711 处的剪接供体序列 位置 3718-4004 处的 AP 头部序列 位置 4098-4227 处的 ITR 位置 4246-4262 处的 M13 fwd 位置 4404-4859 处的 f1 ori 位置 4885-4989 处的 AmpR 启动子 位置 4990-5799 处的 KanR 位置 5430-5430 处	CCTTAATTAGGCTGCGCGCTCGCTCGCTCA CTGAGGCCGCCCGGGCAAAGCCCGGGCGT CGGGCGACCTTTGGTCGCCCGGCCTCAGTG AGCGAGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCTTGTAGTT AATGATTAACCCGCCATGCTACTTATCTAC GTAGCCATGCTCTAGGAAGATCGGAATTCG CCCTTAAGCTAGCGGCGCGCCCAATTCTGC AGCTCAGCCTACTACTTGCTTTCCAGGCTG TTCCTAGTTCCCATGTCAGCTGCTTGTGCTT TCCAGAGACAAAACAGGAATAATAGATGT CATTAAATATACATTGGGCCCCAGGCGGTC AATGTGGCAGCCTGAGCCTCCTTTCCATCT CTGTGGAGGCAGACATAGGACCCCCAACA AACAGCATGCAGGTTGGGAGCCAGCCACA GGACCCAGGTAAGGGGCCCTGGGTCCTTA AGCTTCTGCCACTGGCTCCGGCATTGCAGA GAGAAGAGAAGGGGCGGCAGACTGGAGA GCTGGGCTCCATTTTTGTTCCTTGGTGCCCT GCCCCTCCCCATGACCTGCAGAGACATTCA GCCTGCCAGGCTTTATGAGGTGGGAGCTGG GCTCTCCCTGATGTATTATTCAGCTCCCTG GAGTTGGCCAGCTCCTGTTACACTGGCCAC AGCCCTGGGCATCCGCTTCTCACTTCTAGT TTCCCCTCCAAGGTAATGTGGTGGGTCATG ATCATTCTATCCTGGCTTCAGGGACCTGAC TCCACTTTGGGGCCATTCGAGGGGTCTAGG GTAGATGATGTCCCCCTGTGGGGATTAATG TCCTGCTCTGTAAAACTGAGCTAGCTGAGA

[0331]

SEQ ID NO.	描述	质粒序列
	的 mutBsmBI 位置 5970-6558 处的 ori 位置 6846-6867 处的 CAP 结合位点 位置 6882-6912 处的 lac 启动子 位置 6920-6936 处的 lac 操纵子 位置 6944-6960 处的 M13 rev 位置 12-4227 处的待转移至双载体系统中的载体中的转基因	TCCAGGAGGGCTTGGCCAGAGACAGCAAG TTGTTGCCATGGTGACTTTAAAGCCAGGTT GCTGCCCCAGCACAGGCCTCCCAGTCTACC CTCACTAGAAAACAACACCCAGGCACTTTC CACCACCTCTCAAAGGTGAAACCCAAGGC TGGTCTAGAGAATGAATTATGGATCCTCGC TGTCCGTGCCACCCAGCTAGTCCCAGCGGC TCAGACACTGAGGAGAGACTGTAGGTTCA GCTACAAGCAAAAAGACCTAGCTGGTCTC CAAGCAGTGTCTCCAAGTCCCTGAACCTGT GACACCTGCCCCAGGCATCATCAGGCACA GAGGGCCACCAAGAATTCTAGCGGCCGCC ACCATGGCCTTGCTCATCCACCTCAAGACA GTCTCGGAGCTGCGGGGCAGGGGCGACCG GATCGCCAAAGTGACTTTCCGAGGGCAATC CTTCTACTCTCGGGTCCTGGAGAACTGTGA GGATGTGGCTGACTTTGATGAGACATTTCG GTGGCCGGTGGCCAGCAGCATCGACAGAA ATGAGATGCTGGAGATTCAGGTTTTCAACT ACAGCAAAGTCTTCAGCAACAAGCTCATC GGGACCTTCCGCATGGTGCTGCAGAAGGT GGTAGAGGAGAGCCATGTGGAGGTGACTG ACACGCTGATTGATGACAACAATGCTATCA TCAAGACCAGCCTGTGCGTGGAGGTCCGGT ATCAGGCCACTGACGGCACAGTGGGCTCCT GGGACGATGGGGACTTCCTGGGAGATGAG TCTCTTCAAGAGGAAGAGAAGGACAGCCA AGAGACGGATGGACTGCTCCCAGGCTCCC GGCCCAGCTCCCGGCCCCCAGGAGAGAAG AGCTTCCGGAGAGCCGGGAGGAGCGTGTT CTCCGCCATGAAGCTCGGCAAAAACCGGT CTCACAAGGAGGAGCCCCAAAGACCAGAT GAACCGGCGGTGCTGGAGATGGAAGACCT TGACCATCTGGCCATTCGGCTAGGAGATGG ACTGGATCCCGACTCGGTGTCTCTAGCCTC AGTCACAGCTCTCACCACTAATGTCTCCAA CAAGCGATCTAAGCCAGACATTAAGATGG AGCCAAGTGCTGGGCGGCCCATGGATTAC CAGGTCAGCATCACGGTGATCGAGGCCCG GCAGCTGGTGGGCTTGAACATGGACCCTGT GGTGTGCGTGGAGGTGGGTGACGACAAGA AGTACACATCCATGAAGGAGTCCACTAACT GCCCCTATTACAACGAGTACTTCGTCTTCG ACTTCCATGTCTCTCCGGATGTCATGTTTG ACAAGATCATCAAGATTTCGGTGATTCACT CCAAGAACCTGCTGCGCAGTGGCACCCTG GTGGGCTCCTTCAAAATGGACGTGGGAAC CGTGTACTCGCAGCCAGAGCACCAGTTCCA

[0332]

SEQ ID NO.	描述	质粒序列
		TCACAAGTGGGCCATCCTGTCTGACCCCGA TGACATCTCCTCGGGGCTGAAGGGCTACGT GAAGTGTGACGTTGCCGTGGTGGGCAAAG GGGACAACATCAAGACGCCCCACAAGGCC AATGAGACCGACGAAGATGACATTGAGGG GAACTTGCTGCTCCCCGAGGGGGTGCCCCC CGAACGCCAGTGGGCCCGGTTCTATGTGAA AATTTACCGAGCAGAGGGGCTGCCCCGTAT GAACACAAGCCTCATGGCCAATGTAAAGA AGGCTTTCATCGGTGAAAACAAGGACCTC GTGGACCCCTACGTGCAAGTCTTCTTTGCT GGCCAGAAGGGCAAGACTTCAGTGCAGAA GAGCAGCTATGAGCCCCTGTGGAATGAGC AGGTCGTCTTTACAGACCTCTTCCCCCCAC TCTGCAAACGCATGAAGGTGCAGATCCGA GACTCGGACAAGGTCAACGACGTGGCCAT CGGCACCCACTTCATTGACCTGCGCAAGAT TTCTAATGACGGAGACAAAGGCTTCCTGCC CACACTGGGCCCAGCCTGGGTGAACATGT ACGGCTCCACACGTAACTACACGCTGCTGG ATGAGCATCAGGACCTGAACGAGGGCCTG GGGGAGGGTGTGTCCTTCCGGGCCCGGCTC CTGCTGGGCCTGGCTGTGGAGATCGTAGAC ACCTCCAACCCTGAGCTCACCAGCTCCACA GAGGTGCAGGTGGAGCAGGCCACGCCCAT CTCGGAGAGCTGTGCAGGTAAAATGGAAG AATTCTTTCTCTTTGGAGCCTTCCTGGAGG CCTCAATGATCGACCGGAGAAACGGAGAC AAGCCCATCACCTTTGAGGTCACCATAGGC AACTATGGGAACGAAGTTGATGGCCTGTCC CGGCCCCAGCGGCCTCGGCCCCGGAAGGA GCCGGGGGATGAGGAAGAAGTAGACCTGA TTCAGAACGCAAGTGATGACGAGGCCGGT GATGCCGGGGACCTGGCCTCAGTCTCCTCC ACTCCACCAATGCGGCCCCAGGTCACCGAC AGGAACTACTTCCATCTGCCCTACCTGGAG CGAAAGCCCTGCATCTACATCAAGAGCTG GTGGCCGGACCAGCGCCGCCGCCTCTACA ATGCCAACATCATGGACCACATTGCCGACA AGCTGGAAGAAGGCCTGAACGACATACAG GAGATGATCAAAACGGAGAAGTCCTACCC TGAGCGTCGCCTGCGGGGCGTCCTGGAGG AGCTGAGCTGTGGCTGCTGCCGCTTCCTCT CCCTCGCTGACAAGGACCAGGGCCACTCAT CCCGCACCAGGCTTGACCGGGAGCGCCTC AAGTCCTGCATGAGGGAGCTGGTAAGTAT CAAGGTTACAAGACAGGTTTAAGGAGACC AATAGAAACTGGGCTTGTCGAGACAGAGA

[0333]

SEQ ID NO.	描述	质粒序列
		AGACTCTTGCGTTTCTGAGCTAGCCCCCGG GTGCGCGGCGTCGGTGGTGCCGGCGGGGG GCGCCAGGTCGCAGGCGGTGTAGGGCTCC AGGCAGGCGGCGAAGGCCATGACGTGCGC TATGAAGGTCTGCTCCTGCACGCCGTGAAC CAGGTGCGCCTGCGGGCCGCGCGCGAACA CCGCCACGTCCTCGCCTGCGTGGGTCTCTT CGTCCAGGGGCACTGCTGACTGCTGCCGAT ACTCGGGGCTCCCGCTCTCGCTCTCGGTAA CATCCGGCCGGGCGCCGTCCTTGAGCACAT AGCCTGGACCGTTTCGTCGACCTCGAGTTA AGGGCGAATTCCCGATAAGGATCTTCCTAG AGCATGGCTACGTAGATAAGTAGCATGGC GGGTTAATCATTAACTACAAGGAACCCCTA GTGATGGAGTTGGCCACTCCCTCTCTGCGC GCTCGCTCGCTCACTGAGGCCGGGCGACCA AAGGTCGCCCGACGCCCGGGCTTTGCCCGG GCGGCCTCAGTGAGCGAGCGAGCGCGCAG CCTTAATTAACCTAATTCACTGGCCGTCGT TTTACAACGTCGTGACTGGGAAAACCCTGG CGTTACCCAACTTAATCGCCTTGCAGCACA TCCCCCTTTCGCCAGCTGGCGTAATAGCGA AGAGGCCCGCACCGATCGCCCTTCCCAACA GTTGCGCAGCCTGAATGGCGAATGGGACG CGCCCTGTAGCGGCGCATTAAGCGCGGCG GGTGTGGTGGTTACGCGCAGCGTGACCGCT ACACTTGCCAGCGCCCTAGCGCCCGCTCCT TTCGCTTTCTTCCCTTCCTTTCTCGCCACGT TCGCCGGCTTTCCCCGTCAAGCTCTAAATC GGGGGCTCCCTTTAGGGTTCCGATTTAGTG CTTTACGGCACCTCGACCCCAAAAAACTTG ATTAGGGTGATGGTTCACGTAGTGGGCCAT CGCCCTGATAGACGGTTTTTCGCCCTTTGA CGTTGGAGTCCACGTTCTTTAATAGTGGAC TCTTGTTCCAAACTGGAACAACACTCAACC CTATCTCGGTCTATTCTTTTGATTTATAAGG GATTTTGCCGATTTCGGCCTATTGGTTAAA AAATGAGCTGATTTAACAAAAATTTAACGC GAATTTTAACAAAATATTAACGCTTACAAT TTAGGTGGCACTTTTCGGGGAAATGTGCGC GGAACCCCTATTTGTTTATTTTTCTAAATAC ATTCAAATATGTATCCGCTCATGAGACAAT AACCCTGATAAATGCTTCAATAATATTGAA AAAGGAAGAGTATGAGCCATATTCAACGG GAAACGTCGAGGCCGCGATTAAATTCCAA CATGGATGCTGATTTATATGGGTATAAATG GGCTCGCGATAATGTCGGGCAATCAGGTG CGACAATCTATCGCTTGTATGGGAAGCCCG

[0334]

SEQ ID NO.	描述	质粒序列
		ATGCGCCAGAGTTGTTTCTGAAACATGGCA AAGGTAGCGTTGCCAATGATGTTACAGATG AGATGGTCAGACTAAACTGGCTGACGGAA TTTATGCCTCTTCCGACCATCAAGCATTTTA TCCGTACTCCTGATGATGCATGGTTACTCA CCACTGCGATCCCCGGAAAAACAGCATTCC AGGTATTAGAAGAATATCCTGATTCAGGTG AAAATATTGTTGATGCGCTGGCAGTGTTCC TGCGCCGGTTGCATTCGATTCCTGTTTGTA ATTGTCCTTTTAACAGCGATCGCGTATTTC GTCTTGCTCAGGCGCAATCACGAATGAATA ACGGTTTGGTTGATGCGAGTGATTTTGATG ACGAGCGTAATGGCTGGCCTGTTGAACAA GTCTGGAAAGAAATGCATAAACTTTTGCCA TTCTCACCGGATTCAGTCGTCACTCATGGT GATTTCTCACTTGATAACCTTATTTTTGACG AGGGGAAATTAATAGGTTGTATTGATGTTG GACGAGTCGGAATCGCAGACCGATACCAG GATCTTGCCATCCTATGGAACTGCCTCGGT GAGTTTTCTCCTTCATTACAGAAACGGCTT TTTCAAAAATATGGTATTGATAATCCTGAT ATGAATAAATTGCAGTTTCATTTGATGCTC GATGAGTTTTTCTAACTGTCAGACCAAGTT TACTCATATATACTTTAGATTGATTTAAAA CTTCATTTTTAATTTAAAAGGATCTAGGTG AAGATCCTTTTTGATAATCTCATGACCAAA ATCCCTTAACGTGAGTTTTCGTTCCACTGA GCGTCAGACCCCGTAGAAAAGATCAAAGG ATCTTCTTGAGATCCTTTTTTTCTGCGCGTA ATCTGCTGCTTGCAAACAAAAAAACCACC GCTACCAGCGGTGGTTTGTTTGCCGGATCA AGAGCTACCAACTCTTTTTCCGAAGGTAAC TGGCTTCAGCAGAGCGCAGATACCAAATA CTGTTCTTCTAGTGTAGCCGTAGTTAGGCC ACCACTTCAAGAACTCTGTAGCACCGCCTA CATACCTCGCTCTGCTAATCCTGTTACCAG TGGCTGCTGCCAGTGGCGATAAGTCGTGTC TTACCGGGTTGGACTCAAGACGATAGTTAC CGGATAAGGCGCAGCGGTCGGGCTGAACG GGGGGTTCGTGCACACAGCCCAGCTTGGA GCGAACGACCTACACCGAACTGAGATACC TACAGCGTGAGCTATGAGAAAGCGCCACG CTTCCCGAAGGGAGAAAGGCGGACAGGTA TCCGGTAAGCGGCAGGGTCGGAACAGGAG AGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTC GCCACCTCTGACTTGAGCGTCGATTTTTGT GATGCTCGTCAGGGGGGCGGAGCCTATGG

[0335]

SEQ ID NO.	描述	质粒序列
		AAAAACGCCAGCAACGCGGCCTTTTTACG GTTCCTGGCCTTTTGCTGGCCTTTTGCTCAC ATGTTCTTTCCTGCGTTATCCCCTGATTCTG TGGATAACCGTATTACCGCCTTTGAGTGAG CTGATACCGCTCGCCGCAGCCGAACGACC GAGCGCAGCGAGTCAGTGAGCGAGGAAGC GGAAGAGCGCCCAATACGCAAACCGCCTC TCCCCGCGCGTTGGCCGATTCATTAATGCA GCTGGCACGACAGGTTTCCCGACTGGAAA GCGGGCAGTGAGCGCAACGCAATTAATGT GAGTTAGCTCACTCATTAGGCACCCCAGGC TTTACACTTTATGCTTCCGGCTCGTATGTTG TGTGGAATTGTGAGCGGATAACAATTTCAC ACAGGAAACAGCTATGACCATGATTACGC CAGATTTAATTAAGG
82	含有以下特征的 3’转基因质粒: 位置 12-141 处的 ITR 位置 229-515 处的 AP 头部序列 位置 538-586 处的剪接受体序列 人类 OTOF 同功型 5 的位置 587-4174 处的 C 末端部分 位置 4217-4438 处的 bGH poly(A)序列 位置 4526-4655 处的 ITR 位置 4674-4690 处的 M13 fwd 位置 4832-5287 处的 f1 ori 位置 5313-5417 处的 AmpR 启动子 位置 5418-6227 处的 KanR 位置 5858-5858 处的 mutBsmBI 位置 6398-6986 处的 ori 位置 7274-7295 处	CCTTAATTAGGCTGCGCGCTCGCTCGCTCA CTGAGGCCGCCCGGGCAAAGCCCGGGCGT CGGGCGACCTTTGGTCGCCCGGCCTCAGTG AGCGAGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCTTGTAGTT AATGATTAACCCGCCATGCTACTTATCTAC GTAGCCATGCTCTAGGAAGATCGGAATTCG CCCTTAAGCTAGCGGCGCGCCCCCGGGTGC GCGGCGTCGGTGGTGCCGGCGGGGGGCGC CAGGTCGCAGGCGGTGTAGGGCTCCAGGC AGGCGGCGAAGGCCATGACGTGCGCTATG AAGGTCTGCTCCTGCACGCCGTGAACCAGG TGCGCCTGCGGGCCGCGCGCGAACACCGC CACGTCCTCGCCTGCGTGGGTCTCTTCGTC CAGGGGCACTGCTGACTGCTGCCGATACTC GGGGCTCCCGCTCTCGCTCTCGGTAACATC CGGCCGGGCGCCGTCCTTGAGCACATAGCC TGGACCGTTTCCTTAAGCGACGCATGCTCG CGATAGGCACCTATTGGTCTTACTGACATC CACTTTGCCTTTCTCTCCACAGGAAAACAT GGGGCAGCAGGCCAGGATGCTGCGGGCCC AGGTGAAGCGGCACACGGTGCGGGACAAG CTGAGGCTGTGCCAGAACTTCCTGCAGAAG CTGCGCTTCCTGGCGGACGAGCCCCAGCAC AGCATTCCCGACATCTTCATCTGGATGATG AGCAACAACAAGCGTGTCGCCTATGCCCGT GTGCCCTCCAAGGACCTGCTCTTCTCCATC GTGGAGGAGGAGACTGGCAAGGACTGCGC CAAGGTCAAGACGCTCTTCCTTAAGCTGCC AGGGAAGCGGGGCTTCGGCTCGGCAGGCT GGACAGTGCAGGCCAAGGTGGAGCTGTAC CTGTGGCTGGGCCTCAGCAAACAGCGCAA GGAGTTCCTGTGCGGCCTGCCCTGTGGCTT

[0336]

SEQ ID NO.	描述	质粒序列
	的 CAP 结合位点 位置 7310-7340 处的 lac 启动子 位置 7348-7364 处的 lac 操纵子 位置 7372-7388 处的 M13 rev 位置 12-4655 处的待转移至双载体系统中的载体中的转基因	CCAGGAGGTCAAGGCAGCCCAGGGCCTGG GCCTGCATGCCTTCCCACCCGTCAGCCTGG TCTACACCAAGAAGCAGGCGTTCCAGCTCC GAGCGCACATGTACCAGGCCCGCAGCCTCT TTGCCGCCGACAGCAGCGGACTCTCAGACC CCTTTGCCCGCGTCTTCTTCATCAATCAGA GTCAGTGCACAGAGGTGCTGAATGAGACC CTGTGTCCCACCTGGGACCAGATGCTGGTG TTCGACAACCTGGAGCTCTATGGTGAAGCT CATGAGCTGAGGGACGATCCGCCCATCATT GTCATTGAAATCTATGACCAGGATTCCATG GGCAAAGCTGACTTCATGGGCCGGACCTTC GCCAAACCCCTGGTGAAGATGGCAGACGA GGCGTACTGCCCACCCCGCTTCCCACCTCA GCTCGAGTACTACCAGATCTACCGTGGCAA CGCCACAGCTGGAGACCTGCTGGCGGCCTT CGAGCTGCTGCAGATTGGACCAGCAGGGA AGGCTGACCTGCCCCCCATCAATGGCCCGG TGGACGTGGACCGAGGTCCCATCATGCCCG TGCCCATGGGCATCCGGCCCGTGCTCAGCA AGTACCGAGTGGAGGTGCTGTTCTGGGGCC TACGGGACCTAAAGCGGGTGAACCTGGCC CAGGTGGACCGGCCACGGGTGGACATCGA GTGTGCAGGGAAGGGGGTGCAGTCGTCCC TGATCCACAATTATAAGAAGAACCCCAACT TCAACACCCTCGTCAAGTGGTTTGAAGTGG ACCTCCCAGAGAACGAGCTGCTGCACCCG CCCTTGAACATCCGTGTGGTGGACTGCCGG GCCTTCGGTCGCTACACACTGGTGGGCTCC CATGCCGTCAGCTCCCTGCGACGCTTCATC TACCGGCCCCCAGACCGCTCGGCCCCCAGC TGGAACACCACGGTCAGGCTTCTCCGGCGC TGCCGTGTGCTGTGCAATGGGGGCTCCTCC TCTCACTCCACAGGGGAGGTTGTGGTGACT ATGGAGCCAGAGGTACCCATCAAGAAACT GGAGACCATGGTGAAGCTGGACGCGACTT CTGAAGCTGTTGTCAAGGTGGATGTGGCTG AGGAGGAGAAGGAGAAGAAGAAGAAGAA GAAGGGCACTGCGGAGGAGCCAGAGGAGG AGGAGCCAGACGAGAGCATGCTGGACTGG TGGTCCAAGTACTTTGCCTCCATTGACACC ATGAAGGAGCAACTTCGACAACAAGAGCC CTCTGGAATTGACTTGGAGGAGAAGGAGG AAGTGGACAATACCGAGGGCCTGAAGGGG TCAATGAAGGGCAAGGAGAAGGCAAGGGC TGCCAAAGAGGAGAAGAAGAAGAAAACTC AGAGCTCTGGCTCTGGCCAGGGGTCCGAG GCCCCCGAGAAGAAGAAACCCAAGATTGA

[0337]

SEQ ID NO.	描述	质粒序列
		TGAGCTTAAGGTATACCCCAAAGAGCTGG AGTCCGAGTTTGATAACTTTGAGGACTGGC TGCACACTTTCAACTTGCTTCGGGGCAAGA CCGGGGATGATGAGGATGGCTCCACCGAG GAGGAGCGCATTGTGGGACGCTTCAAGGG CTCCCTCTGCGTGTACAAAGTGCCACTCCC AGAGGACGTGTCCCGGGAAGCCGGCTACG ACTCCACCTACGGCATGTTCCAGGGCATCC CGAGCAATGACCCCATCAATGTGCTGGTCC GAGTCTATGTGGTCCGGGCCACGGACCTGC ACCCTGCTGACATCAACGGCAAAGCTGAC CCCTACATCGCCATCCGGCTAGGCAAGACT GACATCCGCGACAAGGAGAACTACATCTC CAAGCAGCTCAACCCTGTCTTTGGGAAGTC CTTTGACATCGAGGCCTCCTTCCCCATGGA ATCCATGCTGACGGTGGCTGTGTATGACTG GGACCTGGTGGGCACTGATGACCTCATTGG GGAAACCAAGATCGACCTGGAGAACCGCT TCTACAGCAAGCACCGCGCCACCTGCGGC ATCGCCCAGACCTACTCCACACATGGCTAC AATATCTGGCGGGACCCCATGAAGCCCAG CCAGATCCTGACCCGCCTCTGCAAAGACGG CAAAGTGGACGGCCCCCACTTTGGGCCCCC TGGGAGAGTGAAGGTGGCCAACCGCGTCT TCACTGGGCCCTCTGAGATTGAGGACGAG AACGGTCAGAGGAAGCCCACAGACGAGCA TGTGGCGCTGTTGGCCCTGAGGCACTGGGA GGACATCCCCCGCGCAGGCTGCCGCCTGGT GCCAGAGCATGTGGAGACGAGGCCGCTGC TCAACCCCGACAAGCCGGGCATCGAGCAG GGCCGCCTGGAGCTGTGGGTGGACATGTTC CCCATGGACATGCCAGCCCCTGGGACGCCT CTGGACATCTCACCTCGGAAGCCCAAGAA GTACGAGCTGCGGGTCATCATCTGGAACAC AGATGAGGTGGTCTTGGAGGACGACGACT TCTTCACAGGGGAGAAGTCCAGTGACATCT TCGTGAGGGGGTGGCTGAAGGGCCAGCAG GAGGACAAGCAGGACACAGACGTCCACTA CCACTCCCTCACTGGCGAGGGCAACTTCAA CTGGCGCTACCTGTTCCCCTTCGACTACCT GGCGGCGGAGGAGAAGATCGTCATCTCCA AGAAGGAGTCCATGTTCTCCTGGGACGAG ACCGAGTACAAGATCCCCGCGCGGCTCAC CCTGCAGATCTGGGATGCGGACCACTTCTC CGCTGACGACTTCCTGGGGGCCATCGAGCT GGACCTGAACCGGTTCCCGCGGGGCGCAA AGACAGCCAAGCAGTGCACCATGGAGATG GCCACCGGGGAGGTGGACGTGCCCCTCGT

[0338]

SEQ ID NO.	描述	质粒序列
		GTCCATCTTCAAGCAAAAGCGCGTCAAAG GCTGGTGGCCCCTCCTGGCCCGCAATGAGA ACGATGAGTTTGAGCTCACGGGCAAGGTG GAGGCTGAGCTGCATTTACTGACAGCAGA GGAGGCAGAGAAGAACCCAGTGGGCCTGG CCCGCAATGAACCTGACCCCCTAGAGAAA CCCAACCGGCCCGACACGGCCTTCGTCTGG TTCCTCAACCCTCTCAAGTCCATCAAGTAC CTCATCTGCACCCGGTACAAGTGGCTCATC ATCAAGATCGTGCTGGCGCTGTTGGGGCTG CTCATGTTGGGGCTCTTCCTCTACAGCCTC CCTGGCTACATGGTCAAAAAGCTCCTTGGG GCATGAACGGCCGCTATGCTAGCTTGGTAC CAAGGGCGGATCCTGCATAGAGCTCGCTG ATCAGCCTCGACTGTGCCTTCTAGTTGCCA GCCATCTGTTGTTTGCCCCTCCCCCGTGCCT TCCTTGACCCTGGAAGGTGCCACTCCCACT GTCCTTTCCTAATAAAATGAGGAAATTGCA TCGCATTGTCTGAGTAGGTGTCATTCTATT CTGGGGGGTGGGGTGGGGCAGGACAGCAA GGGGGAGGATTGGGAAGACAATAGCAGGC ATCTCGAGTTAAGGGCGAATTCCCGATAAG GATCTTCCTAGAGCATGGCTACGTAGATAA GTAGCATGGCGGGTTAATCATTAACTACAA GGAACCCCTAGTGATGGAGTTGGCCACTCC CTCTCTGCGCGCTCGCTCGCTCACTGAGGC CGGGCGACCAAAGGTCGCCCGACGCCCGG GCTTTGCCCGGGCGGCCTCAGTGAGCGAGC GAGCGCGCAGCCTTAATTAACCTAATTCAC TGGCCGTCGTTTTACAACGTCGTGACTGGG AAAACCCTGGCGTTACCCAACTTAATCGCC TTGCAGCACATCCCCCTTTCGCCAGCTGGC GTAATAGCGAAGAGGCCCGCACCGATCGC CCTTCCCAACAGTTGCGCAGCCTGAATGGC GAATGGGACGCGCCCTGTAGCGGCGCATT AAGCGCGGCGGGTGTGGTGGTTACGCGCA GCGTGACCGCTACACTTGCCAGCGCCCTAG CGCCCGCTCCTTTCGCTTTCTTCCCTTCCTT TCTCGCCACGTTCGCCGGCTTTCCCCGTCA AGCTCTAAATCGGGGGCTCCCTTTAGGGTT CCGATTTAGTGCTTTACGGCACCTCGACCC CAAAAAACTTGATTAGGGTGATGGTTCACG TAGTGGGCCATCGCCCTGATAGACGGTTTT TCGCCCTTTGACGTTGGAGTCCACGTTCTTT AATAGTGGACTCTTGTTCCAAACTGGAACA ACACTCAACCCTATCTCGGTCTATTCTTTTG ATTTATAAGGGATTTTGCCGATTTCGGCCT ATTGGTTAAAAAATGAGCTGATTTAACAAA

[0339]

SEQ ID NO.	描述	质粒序列
		AATTTAACGCGAATTTTAACAAAATATTAA CGCTTACAATTTAGGTGGCACTTTTCGGGG AAATGTGCGCGGAACCCCTATTTGTTTATT TTTCTAAATACATTCAAATATGTATCCGCT CATGAGACAATAACCCTGATAAATGCTTCA ATAATATTGAAAAAGGAAGAGTATGAGCC ATATTCAACGGGAAACGTCGAGGCCGCGA TTAAATTCCAACATGGATGCTGATTTATAT GGGTATAAATGGGCTCGCGATAATGTCGG GCAATCAGGTGCGACAATCTATCGCTTGTA TGGGAAGCCCGATGCGCCAGAGTTGTTTCT GAAACATGGCAAAGGTAGCGTTGCCAATG ATGTTACAGATGAGATGGTCAGACTAAACT GGCTGACGGAATTTATGCCTCTTCCGACCA TCAAGCATTTTATCCGTACTCCTGATGATG CATGGTTACTCACCACTGCGATCCCCGGAA AAACAGCATTCCAGGTATTAGAAGAATAT CCTGATTCAGGTGAAAATATTGTTGATGCG CTGGCAGTGTTCCTGCGCCGGTTGCATTCG ATTCCTGTTTGTAATTGTCCTTTTAACAGCG ATCGCGTATTTCGTCTTGCTCAGGCGCAAT CACGAATGAATAACGGTTTGGTTGATGCGA GTGATTTTGATGACGAGCGTAATGGCTGGC CTGTTGAACAAGTCTGGAAAGAAATGCAT AAACTTTTGCCATTCTCACCGGATTCAGTC GTCACTCATGGTGATTTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGAGTCGGAATCGCA GACCGATACCAGGATCTTGCCATCCTATGG AACTGCCTCGGTGAGTTTTCTCCTTCATTAC AGAAACGGCTTTTTCAAAAATATGGTATTG ATAATCCTGATATGAATAAATTGCAGTTTC ATTTGATGCTCGATGAGTTTTTCTAACTGTC AGACCAAGTTTACTCATATATACTTTAGAT TGATTTAAAACTTCATTTTTAATTTAAAAG GATCTAGGTGAAGATCCTTTTTGATAATCT CATGACCAAAATCCCTTAACGTGAGTTTTC GTTCCACTGAGCGTCAGACCCCGTAGAAA AGATCAAAGGATCTTCTTGAGATCCTTTTT TTCTGCGCGTAATCTGCTGCTTGCAAACAA AAAAACCACCGCTACCAGCGGTGGTTTGTT TGCCGGATCAAGAGCTACCAACTCTTTTTC CGAAGGTAACTGGCTTCAGCAGAGCGCAG ATACCAAATACTGTTCTTCTAGTGTAGCCG TAGTTAGGCCACCACTTCAAGAACTCTGTA GCACCGCCTACATACCTCGCTCTGCTAATC CTGTTACCAGTGGCTGCTGCCAGTGGCGAT AAGTCGTGTCTTACCGGGTTGGACTCAAGA

[0340]

SEQ ID NO.	描述	质粒序列
		CGATAGTTACCGGATAAGGCGCAGCGGTC GGGCTGAACGGGGGGTTCGTGCACACAGC CCAGCTTGGAGCGAACGACCTACACCGAA CTGAGATACCTACAGCGTGAGCTATGAGA AAGCGCCACGCTTCCCGAAGGGAGAAAGG CGGACAGGTATCCGGTAAGCGGCAGGGTC GGAACAGGAGAGCGCACGAGGGAGCTTCC AGGGGGAAACGCCTGGTATCTTTATAGTCC TGTCGGGTTTCGCCACCTCTGACTTGAGCG TCGATTTTTGTGATGCTCGTCAGGGGGGCG GAGCCTATGGAAAAACGCCAGCAACGCGG CCTTTTTACGGTTCCTGGCCTTTTGCTGGCC TTTTGCTCACATGTTCTTTCCTGCGTTATCC CCTGATTCTGTGGATAACCGTATTACCGCC TTTGAGTGAGCTGATACCGCTCGCCGCAGC CGAACGACCGAGCGCAGCGAGTCAGTGAG CGAGGAAGCGGAAGAGCGCCCAATACGCA AACCGCCTCTCCCCGCGCGTTGGCCGATTC ATTAATGCAGCTGGCACGACAGGTTTCCCG ACTGGAAAGCGGGCAGTGAGCGCAACGCA ATTAATGTGAGTTAGCTCACTCATTAGGCA CCCCAGGCTTTACACTTTATGCTTCCGGCT CGTATGTTGTGTGGAATTGTGAGCGGATAA CAATTTCACACAGGAAACAGCTATGACCAT GATTACGCCAGATTTAATTAAGG
83	含有以下特征的 5’转基因质粒: 位置 12-141 处的 ITR 位置 230-594 处的 CMV，即增强子 位置 296-599 处的 CMV 增强子 位置 596-878 处的鸡 β-肌动蛋白启动子 位置 879-971 处的外显子 1 位置 971-1172 处的嵌合内含子 位置 1189-1198 处的 Kozak 序列 位置 1195-3600 处的人类 OTOF 同功型 5 的 N 末端部分	CCTTAATTAGGCTGCGCGCTCGCTCGCTCA CTGAGGCCGCCCGGGCAAAGCCCGGGCGT CGGGCGACCTTTGGTCGCCCGGCCTCAGTG AGCGAGCGAGCGCGCAGAGAGGGAGTGGC CAACTCCATCACTAGGGGTTCCTTGTAGTT AATGATTAACCCGCCATGCTACTTATCTAC GTAGCCATGCTCTAGGAAGATCGGAATTCG CCCTTAAGCTAGCGGCGCGCCGGTACCTAG TTATTAATAGTAATCAATTACGGGGTCATT AGTTCATAGCCCATATATGGAGTTCCGCGT TACATAACTTACGGTAAATGGCCCGCCTGG CTGACCGCCCAACGACCCCCGCCCATTGAC GTCAATAATGACGTATGTTCCCATAGTAAC GCCAATAGGGACTTTCCATTGACGTCAATG GGTGGAGTATTTACGGTAAACTGCCCACTT GGCAGTACATCAAGTGTATCATATGCCAAG TACGCCCCCTATTGACGTCAATGACGGTAA ATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTATGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCAT GGTCGAGGTGAGCCCCACGTTCTGCTTCAC TCTCCCCATCTCCCCCCCCTCCCCACCCCCA ATTTTGTATTTATTTATTTTTTAATTATTTTG

[0341]

SEQ ID NO.	描述	质粒序列
	位置 3601-3684 处的剪接供体序列 位置 3691-3977 处的 AP 头部序列 位置 4071-4200 处的 ITR 位置 4219-4235 处的 M13 fwd 位置 4377-4832 处的 f1 ori 位置 4858-4962 处的 AmpR 启动子 位置 4963-5772 处的 KanR 位置 4963-5772 处的 mutBsmBI 位置 5943-6531 处的 ori 位置 6819-6840 处的 CAP 结合位点 位置 6855-6885 处的 lac 启动子 位置 6835-6909 处的 lac 操纵子 位置 6917-6933 处的 M13 rev 位置 12-4200 处的待转移至双载体系统中的载体中的转基因	TGCAGCGATGGGGGCGGGGGGGGGGGGGG GGCGCGCGCCAGGCGGGGCGGGGCGGGGC GAGGGGCGGGGCGGGGCGAGGCGGAGAG GTGCGGCGGCAGCCAATCAGAGCGGCGCG CTCCGAAAGTTTCCTTTTATGGCGAGGCGG CGGCGGCGGCGGCCCTATAAAAAGCGAAG CGCGCGGCGGGCGGGAGTCGCTGCGCGCT GCCTTCGCCCCGTGCCCCGCTCCGCCGCCG CCTCGCGCCGCCCGCCCCGGCTCTGACTGA CCGCGTTACTCCCACAGGTGAGCGGGCGG GACGGCCCTTCTCCTCCGGGCTGTAATTAG CGCTTGGTTTAATGACGGCTTGTTTCTTTTC TGTGGCTGCGTGAAAGCCTTGAGGGGCTCC GGGAGCTAGAGCCTCTGCTAACCATGTTCA TGCCTTCTTCTTTTTCCTACAGCTCCTGGGC AACGTGCTGGTTATTGTGCTGTCTCATCAT TTTGGCAAAGAATTCTAGCGGCCGCCACCA TGGCCTTGCTCATCCACCTCAAGACAGTCT CGGAGCTGCGGGGCAGGGGCGACCGGATC GCCAAAGTGACTTTCCGAGGGCAATCCTTC TACTCTCGGGTCCTGGAGAACTGTGAGGAT GTGGCTGACTTTGATGAGACATTTCGGTGG CCGGTGGCCAGCAGCATCGACAGAAATGA GATGCTGGAGATTCAGGTTTTCAACTACAG CAAAGTCTTCAGCAACAAGCTCATCGGGA CCTTCCGCATGGTGCTGCAGAAGGTGGTAG AGGAGAGCCATGTGGAGGTGACTGACACG CTGATTGATGACAACAATGCTATCATCAAG ACCAGCCTGTGCGTGGAGGTCCGGTATCAG GCCACTGACGGCACAGTGGGCTCCTGGGA CGATGGGGACTTCCTGGGAGATGAGTCTCT TCAAGAGGAAGAGAAGGACAGCCAAGAG ACGGATGGACTGCTCCCAGGCTCCCGGCCC AGCTCCCGGCCCCCAGGAGAGAAGAGCTT CCGGAGAGCCGGGAGGAGCGTGTTCTCCG CCATGAAGCTCGGCAAAAACCGGTCTCAC AAGGAGGAGCCCCAAAGACCAGATGAACC GGCGGTGCTGGAGATGGAAGACCTTGACC ATCTGGCCATTCGGCTAGGAGATGGACTGG ATCCCGACTCGGTGTCTCTAGCCTCAGTCA CAGCTCTCACCACTAATGTCTCCAACAAGC GATCTAAGCCAGACATTAAGATGGAGCCA AGTGCTGGGCGGCCCATGGATTACCAGGTC AGCATCACGGTGATCGAGGCCCGGCAGCT GGTGGGCTTGAACATGGACCCTGTGGTGTG CGTGGAGGTGGGTGACGACAAGAAGTACA CATCCATGAAGGAGTCCACTAACTGCCCCT ATTACAACGAGTACTTCGTCTTCGACTTCC

[0342]

SEQ ID NO.	描述	质粒序列
		ATGTCTCTCCGGATGTCATGTTTGACAAGA TCATCAAGATTTCGGTGATTCACTCCAAGA ACCTGCTGCGCAGTGGCACCCTGGTGGGCT CCTTCAAAATGGACGTGGGAACCGTGTACT CGCAGCCAGAGCACCAGTTCCATCACAAG TGGGCCATCCTGTCTGACCCCGATGACATC TCCTCGGGGCTGAAGGGCTACGTGAAGTGT GACGTTGCCGTGGTGGGCAAAGGGGACAA CATCAAGACGCCCCACAAGGCCAATGAGA CCGACGAAGATGACATTGAGGGGAACTTG CTGCTCCCCGAGGGGGTGCCCCCCGAACGC CAGTGGGCCCGGTTCTATGTGAAAATTTAC CGAGCAGAGGGGCTGCCCCGTATGAACAC AAGCCTCATGGCCAATGTAAAGAAGGCTTT CATCGGTGAAAACAAGGACCTCGTGGACC CCTACGTGCAAGTCTTCTTTGCTGGCCAGA AGGGCAAGACTTCAGTGCAGAAGAGCAGC TATGAGCCCCTGTGGAATGAGCAGGTCGTC TTTACAGACCTCTTCCCCCCACTCTGCAAA CGCATGAAGGTGCAGATCCGAGACTCGGA CAAGGTCAACGACGTGGCCATCGGCACCC ACTTCATTGACCTGCGCAAGATTTCTAATG ACGGAGACAAAGGCTTCCTGCCCACACTG GGCCCAGCCTGGGTGAACATGTACGGCTCC ACACGTAACTACACGCTGCTGGATGAGCAT CAGGACCTGAACGAGGGCCTGGGGGAGGG TGTGTCCTTCCGGGCCCGGCTCCTGCTGGG CCTGGCTGTGGAGATCGTAGACACCTCCAA CCCTGAGCTCACCAGCTCCACAGAGGTGCA GGTGGAGCAGGCCACGCCCATCTCGGAGA GCTGTGCAGGTAAAATGGAAGAATTCTTTC TCTTTGGAGCCTTCCTGGAGGCCTCAATGA TCGACCGGAGAAACGGAGACAAGCCCATC ACCTTTGAGGTCACCATAGGCAACTATGGG AACGAAGTTGATGGCCTGTCCCGGCCCCAG CGGCCTCGGCCCCGGAAGGAGCCGGGGGA TGAGGAAGAAGTAGACCTGATTCAGAACG CAAGTGATGACGAGGCCGGTGATGCCGGG GACCTGGCCTCAGTCTCCTCCACTCCACCA ATGCGGCCCCAGGTCACCGACAGGAACTA CTTCCATCTGCCCTACCTGGAGCGAAAGCC CTGCATCTACATCAAGAGCTGGTGGCCGGA CCAGCGCCGCCGCCTCTACAATGCCAACAT CATGGACCACATTGCCGACAAGCTGGAAG AAGGCCTGAACGACATACAGGAGATGATC AAAACGGAGAAGTCCTACCCTGAGCGTCG CCTGCGGGGCGTCCTGGAGGAGCTGAGCT GTGGCTGCTGCCGCTTCCTCTCCCTCGCTG

[0343]

SEQ ID NO.	描述	质粒序列
		ACAAGGACCAGGGCCACTCATCCCGCACC AGGCTTGACCGGGAGCGCCTCAAGTCCTGC ATGAGGGAGCTGGTAAGTATCAAGGTTAC AAGACAGGTTTAAGGAGACCAATAGAAAC TGGGCTTGTCGAGACAGAGAAGACTCTTGC GTTTCTGAGCTAGCCCCCGGGTGCGCGGCG TCGGTGGTGCCGGCGGGGGGCGCCAGGTC GCAGGCGGTGTAGGGCTCCAGGCAGGCGG CGAAGGCCATGACGTGCGCTATGAAGGTC TGCTCCTGCACGCCGTGAACCAGGTGCGCC TGCGGGCCGCGCGCGAACACCGCCACGTC CTCGCCTGCGTGGGTCTCTTCGTCCAGGGG CACTGCTGACTGCTGCCGATACTCGGGGCT CCCGCTCTCGCTCTCGGTAACATCCGGCCG GGCGCCGTCCTTGAGCACATAGCCTGGACC GTTTCGTCGACCTCGAGTTAAGGGCGAATT CCCGATAAGGATCTTCCTAGAGCATGGCTA CGTAGATAAGTAGCATGGCGGGTTAATCAT TAACTACAAGGAACCCCTAGTGATGGAGTT GGCCACTCCCTCTCTGCGCGCTCGCTCGCT CACTGAGGCCGGGCGACCAAAGGTCGCCC GACGCCCGGGCTTTGCCCGGGCGGCCTCAG TGAGCGAGCGAGCGCGCAGCCTTAATTAA CCTAATTCACTGGCCGTCGTTTTACAACGT CGTGACTGGGAAAACCCTGGCGTTACCCA ACTTAATCGCCTTGCAGCACATCCCCCTTT CGCCAGCTGGCGTAATAGCGAAGAGGCCC GCACCGATCGCCCTTCCCAACAGTTGCGCA GCCTGAATGGCGAATGGGACGCGCCCTGT AGCGGCGCATTAAGCGCGGCGGGTGTGGT GGTTACGCGCAGCGTGACCGCTACACTTGC CAGCGCCCTAGCGCCCGCTCCTTTCGCTTT CTTCCCTTCCTTTCTCGCCACGTTCGCCGGC TTTCCCCGTCAAGCTCTAAATCGGGGGCTC CCTTTAGGGTTCCGATTTAGTGCTTTACGG CACCTCGACCCCAAAAAACTTGATTAGGGT GATGGTTCACGTAGTGGGCCATCGCCCTGA TAGACGGTTTTTCGCCCTTTGACGTTGGAG TCCACGTTCTTTAATAGTGGACTCTTGTTCC AAACTGGAACAACACTCAACCCTATCTCGG TCTATTCTTTTGATTTATAAGGGATTTTGCC GATTTCGGCCTATTGGTTAAAAAATGAGCT GATTTAACAAAAATTTAACGCGAATTTTAA CAAAATATTAACGCTTACAATTTAGGTGGC ACTTTTCGGGGAAATGTGCGCGGAACCCCT ATTTGTTTATTTTTCTAAATACATTCAAATA TGTATCCGCTCATGAGACAATAACCCTGAT AAATGCTTCAATAATATTGAAAAAGGAAG

[0344]

SEQ ID NO.	描述	质粒序列
		AGTATGAGCCATATTCAACGGGAAACGTC GAGGCCGCGATTAAATTCCAACATGGATG CTGATTTATATGGGTATAAATGGGCTCGCG ATAATGTCGGGCAATCAGGTGCGACAATCT ATCGCTTGTATGGGAAGCCCGATGCGCCAG AGTTGTTTCTGAAACATGGCAAAGGTAGCG TTGCCAATGATGTTACAGATGAGATGGTCA GACTAAACTGGCTGACGGAATTTATGCCTC TTCCGACCATCAAGCATTTTATCCGTACTC CTGATGATGCATGGTTACTCACCACTGCGA TCCCCGGAAAAACAGCATTCCAGGTATTAG AAGAATATCCTGATTCAGGTGAAAATATTG TTGATGCGCTGGCAGTGTTCCTGCGCCGGT TGCATTCGATTCCTGTTTGTAATTGTCCTTT TAACAGCGATCGCGTATTTCGTCTTGCTCA GGCGCAATCACGAATGAATAACGGTTTGG TTGATGCGAGTGATTTTGATGACGAGCGTA ATGGCTGGCCTGTTGAACAAGTCTGGAAA GAAATGCATAAACTTTTGCCATTCTCACCG GATTCAGTCGTCACTCATGGTGATTTCTCA CTTGATAACCTTATTTTTGACGAGGGGAAA TTAATAGGTTGTATTGATGTTGGACGAGTC GGAATCGCAGACCGATACCAGGATCTTGC CATCCTATGGAACTGCCTCGGTGAGTTTTC TCCTTCATTACAGAAACGGCTTTTTCAAAA ATATGGTATTGATAATCCTGATATGAATAA ATTGCAGTTTCATTTGATGCTCGATGAGTT TTTCTAACTGTCAGACCAAGTTTACTCATA TATACTTTAGATTGATTTAAAACTTCATTTT TAATTTAAAAGGATCTAGGTGAAGATCCTT TTTGATAATCTCATGACCAAAATCCCTTAA CGTGAGTTTTCGTTCCACTGAGCGTCAGAC CCCGTAGAAAAGATCAAAGGATCTTCTTGA GATCCTTTTTTTCTGCGCGTAATCTGCTGCT TGCAAACAAAAAAACCACCGCTACCAGCG GTGGTTTGTTTGCCGGATCAAGAGCTACCA ACTCTTTTTCCGAAGGTAACTGGCTTCAGC AGAGCGCAGATACCAAATACTGTTCTTCTA GTGTAGCCGTAGTTAGGCCACCACTTCAAG AACTCTGTAGCACCGCCTACATACCTCGCT CTGCTAATCCTGTTACCAGTGGCTGCTGCC AGTGGCGATAAGTCGTGTCTTACCGGGTTG GACTCAAGACGATAGTTACCGGATAAGGC GCAGCGGTCGGGCTGAACGGGGGGTTCGT GCACACAGCCCAGCTTGGAGCGAACGACC TACACCGAACTGAGATACCTACAGCGTGA GCTATGAGAAAGCGCCACGCTTCCCGAAG GGAGAAAGGCGGACAGGTATCCGGTAAGC

[0345]

SEQ ID NO.	描述	质粒序列
		GGCAGGGTCGGAACAGGAGAGCGCACGAG GGAGCTTCCAGGGGGAAACGCCTGGTATC TTTATAGTCCTGTCGGGTTTCGCCACCTCTG ACTTGAGCGTCGATTTTTGTGATGCTCGTC AGGGGGGCGGAGCCTATGGAAAAACGCCA GCAACGCGGCCTTTTTACGGTTCCTGGCCT TTTGCTGGCCTTTTGCTCACATGTTCTTTCC TGCGTTATCCCCTGATTCTGTGGATAACCG TATTACCGCCTTTGAGTGAGCTGATACCGC TCGCCGCAGCCGAACGACCGAGCGCAGCG AGTCAGTGAGCGAGGAAGCGGAAGAGCGC CCAATACGCAAACCGCCTCTCCCCGCGCGT TGGCCGATTCATTAATGCAGCTGGCACGAC AGGTTTCCCGACTGGAAAGCGGGCAGTGA GCGCAACGCAATTAATGTGAGTTAGCTCAC TCATTAGGCACCCCAGGCTTTACACTTTAT GCTTCCGGCTCGTATGTTGTGTGGAATTGT GAGCGGATAACAATTTCACACAGGAAACA GCTATGACCATGATTACGCCAGATTTAATT AAGG

[0346] 用于表达OTOF的载体

[0347] 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。表达载体的实例公开于例如WO 1994/011026中并通过引用并入本文。用于本文所述的组合物和方法中的表达载体含有编码OTOF的一部分的多核苷酸序列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额外序列元件。可用于表达OTOF的某些载体包括含有引导基因转录的调控序列(如启动子和增强子区域)的质粒。可用于表达OTOF的其他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。适宜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。

[0348] 用于核酸递送的AAV载体

[0349] 在一些实施方案中,将本文所述组合物和方法的核酸并入重组AAV(rAAV)载体和/或病毒粒子中以促进其引入细胞中。可用于本文所述的组合物和方法中的rAAV载体是重组核酸构建体,其包括(1)待表达的异源序列(例如编码OTOF蛋白的N末端或C末端部分的多核苷酸)和(2)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。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载体的方法描述于例如Tal等人类,J.Biomed.Sci.7:279(2000)以及Monahan和Samulski,Gene Delivery 7:24(2000),这些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。

[0350] 本文所述的核酸和载体可并入rAAV病毒粒子中以促进核酸或载体引入细胞中。AAV的衣壳蛋白构成外部,病毒粒子的非核酸部分是由AAV cap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的。rAAV病毒粒子的构建已描述于例如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等人类,J.Virol.76:791(2002)和Bowles等人类,J.Virol.77:423(2003)中,这些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。

[0351] 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,这些AAV血清型包括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和PHP.S。为靶向耳蜗毛细胞,AAV1、AAV2、AAV6、AAV9、Anc80、Anc80L65、DJ/9、7m8和PHP.B可能是特别是有用的。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。本文所述组合物和方法中的第一核酸载体和第二核酸载体可具有相同血清型或不同血清型。不同血清型的AAV载体和AAV蛋白的构建和用途描述于例如Chao等人类,Mol.Ther.2:619(2000);Davidson等人类,Proc.Natl.Acad.Sci.USA 97:3428(2000);Xiao等人类,J.Virol.72:2224(1998);Halbert等人类,J.Virol.74:1524(2000);Halbert等人类,J.Virol.75:6615(2001);和Auricchio等人类,Hum.Molec.Genet.10:3075(2001)中,这些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。

[0352] 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、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描述于例如Duan等人类,J.Virol.75:7662(2001);Halbert等人类,J.Virol.74:1524(2000);Zolotukhin等人类,Methods,28:158(2002);和Auricchio等人类,Hum.Molec.Genet.10:3075(2001)中。

[0353] 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,适宜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)的构建和表征描述于Wu等人类,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等人类,Nat.Genet.,25:436(2000)以及Kolman和Stemmer,Nat.Biotechnol.19:423(2001)。

[0354] 在一些实施方案中,使用AAV载体来递送功能性OTOF蛋白需要使用双载体系统,其中双载体系统的第一成员编码OTOF蛋白的N末端部分并且第二成员编码OTOF蛋白的C末端部分,使得在将双载体系统施用于细胞后,含于两种载体内的多核苷酸序列可接合形成产生全长OTOF蛋白的单一序列。在一些实施方案中,蛋白质是OTOF同功型5蛋白。在一些实施方案中,蛋白质是OTOF同功型1蛋白。

[0355] 在一些实施方案中,双载体系统的第一成员还将以5'至3'顺序包括第一反向末端重复("ITR");启动子(例如Myo15启动子);Kozak序列;OTOF编码序列的N末端部分;剪接供

体序列;AP基因片段(例如AP头部序列);和第二ITR;并且双载体系统的第二成员将以5'至3'顺序包括第一ITR;AP基因片段(例如AP头部序列);剪接受体序列;OTOF编码序列的C末端部分;polyA序列;和第二ITR。在一些实施方案中,OTOF编码序列的N末端部分和OTOF编码序列的C末端部分并不重叠并且在细胞中接合(例如通过在重叠区域(AP基因片段)重组或通过ITR的多联体化)以产生全长OTOF氨基序列(例如,对于OTOF同功型1,SEQ ID NO:1中所示的序列,或者对于OTOF同功型5,SEQ ID NO:5中所示的序列)。在特定实施方案中,OTOF编码序列的N末端部分编码OTOF的氨基酸1-802(例如,SEQ ID NO:1或SEQ ID NO:5的氨基酸1-802,对应于SEQ ID NO:73),并且OTOF编码序列的C末端部分编码OTOF的氨基酸803-1997(例如,SEQ ID NO:1的氨基酸803-1997,或SEQ ID NO:5的氨基酸803-1997,对应于SEQ ID NO:74)。

[0356] 在一些实施方案中,双载体系统的第一成员包括SEQ ID NO:38的Myo15启动子(也由SEQ ID NO:81的核苷酸235-1199表示),其可操作地连接至编码OTOF同功型5蛋白的N末端802个氨基酸(SEQ ID NO:5的氨基酸1-802)的核苷酸,这些核苷酸由编码所述蛋白质的原生多核苷酸序列的外显子1-20编码。在某些实施方案中,编码OTOF同功型5蛋白的N末端氨基酸的核苷酸序列是SEQ ID NO:81的核苷酸1222-3627。在一些实施方案中,编码OTOF同功型5蛋白的N末端氨基酸的核苷酸序列是因遗传密码的冗余性编码SEQ ID NO:5的氨基酸1-802的任何核苷酸序列。编码OTOF同功型5蛋白的核苷酸序列可经部分或完全密码子优化以用于表达。在一些实施方案中,双载体系统的第一成员包括对应于SEQ ID NO:81的核苷酸1216-1225的Kozak序列。在一些实施方案中,双载体系统的第一成员包括对应于SEQ ID NO:81的核苷酸3628-3711的剪接供体序列。在一些实施方案中,双载体系统的第一成员包括对应于SEQ ID NO:81的核苷酸3718-4004的AP头部序列。在特定实施方案中,双载体系统的第一成员包括在5'和3'侧中的每一者上侧接有反向末端重复的SEQ ID NO:81的核苷酸235-4004。在一些实施方案中,侧接反向末端重复是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。在某些实施方案中,5'侧接反向末端重复具有对应于SEQ ID NO:81的核苷酸12-141的序列或与其具有至少80%序列同一性(至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列;并且3'侧接反向末端重复具有对应于SEQ ID NO:81的核苷酸4098-4227的序列或与其具有至少80%序列同一性(至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。本领域技术人类员应理解,对于用于产生病毒载体的转移质粒(通常通过用所述质粒以及携带病毒载体形成所需的AAV基因的其他质粒转染细胞)中任何给定对的反向末端重复序列(例如SEQ ID NO:75、77、79、80、81或83中的任一者),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。然而,在一些非常具体的实施方案中,双载体系统的第一成员包括SEQ ID NO:81的核苷酸12-4227。

[0357] 在一些实施方案中,双载体系统的第二成员包括编码OTOF同功型5蛋白的C末端1195个氨基酸(SEQ ID NO:5的氨基酸803-1997)的核苷酸,紧接着是终止密码子。在某些实施方案中,编码OTOF同功型5蛋白的C末端氨基酸的核苷酸序列是SEQ ID NO:82的核苷酸587-4174。在一些实施方案中,编码OTOF同功型5蛋白的C末端氨基酸的核苷酸序列是因遗

传密码的冗余性编码SEQ ID NO:5的氨基酸803-1997的任何核苷酸序列。编码OTOF同功型5蛋白的核苷酸序列可经部分或完全密码子优化以用于表达。在一些实施方案中,双载体系统的第二成员包括对应于SEQ ID NO:82的核苷酸538-586的剪接受体序列。在一些实施方案中,双载体系统的第二成员包括对应于SEQ ID NO:82的核苷酸229-515的AP头部序列。在一些实施方案中,双载体系统的第二成员包括对应于SEQ ID NO:82的核苷酸4217-4438的poly(A)序列。在特定实施方案中,双载体系统的第二成员包括在5'和3'侧中的每一者上侧接有反向末端重复的SEQ ID NO:82的核苷酸229-4438。在一些实施方案中,侧接反向末端重复是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。在某些实施方案中,5'侧接反向末端重复具有对应于SEQ ID NO:82的核苷酸12-141的序列或与其具有至少80%序列同一性(至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列;并且3'侧接反向末端重复具有对应于SEQ ID NO:82的核苷酸4526-4655的序列或与其具有至少80%序列同一性(至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性)的序列。本领域技术人类员应理解,对于用于产生病毒载体的转移质粒(通常通过用所述质粒以及携带病毒载体形成所需的AAV基因的其他质粒转染细胞)中任何给定对的反向末端重复序列(例如SEQ ID NO:76、78或82中的任一者),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。然而,在一些非常具体的实施方案中,双载体系统的第一成员包括SEQ ID NO:82的核苷酸12-4655。

[0358] 在一些实施方案中,双载体系统是AAV1双载体系统。

[0359] 在一些实施方案中,双载体系统是AAV9双载体系统。

[0360] 药物组合物

[0361] 本文所述的核酸载体(例如AAV载体)可并入媒介物中来施用于患者,例如患有双等位基因OTOF突变的人类患者,如本文所述。含有含编码OTOF蛋白的一部分的多核苷酸的载体(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's Pharmaceut ical Sciences第16版,Osol,A.编辑.(1980);通过引用并入本文)并且以期望形式(以冻干制剂或水溶液的形式)来制备。

[0362] 本文所述核酸载体(例如AAV载体)的混合物可于水与一种或多种赋形剂、载体或稀释剂的适宜混合物中制备。分散液也可于甘油、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(描述于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)、其适宜混合物和/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)、在分散液的情况下通过维持所需粒度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,优选应包括等渗剂,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。

[0363] 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类员已知的。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。为局部施用于内耳,组合物可经配制以含有合成外淋巴溶液。示例性合成外淋巴溶液包括20-200mM NaCl、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。负责施用的人类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。

[0364] 治疗方法

[0365] 本文所述的组合物可通过多种途径施用于患有双等位基因OTOF突变的受试者,例如局部施用于内耳(例如施用于外淋巴或内淋巴中,例如施用至或穿过卵圆窗、圆窗或水平半规管,例如施用于耳蜗毛细胞)、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)。在一些实施方案中,第一核酸载体和第二核酸载体是同时施用的(例如在一种组合物中)。在一些实施方案中,第一核酸载体和第二核酸载体是相继施用的(例如,第二核酸载体是紧跟第一核酸载体或在第一核酸载体后5分钟、10分钟、15分钟、20分钟、25分钟、30分钟、45分钟、1小时、2小时、8小时、12小时、1天、2天、7天、2周、1个月或更长时间施用)。第一核酸载体和第二核酸载体可具有相同血清型或不同血清型(例如AAV血清型)。

[0366] 可如本文所述治疗的受试者是由于双等位基因OTOF突变而患有感觉神经性听力损失或听神经病变或具有患上所述疾病的风险的25岁或25岁以上(例如25-50岁、25-45岁、25-40岁、25-35岁、25-30岁、30-50岁、30-45岁、30-40岁、30-35岁、35-50岁、35-45岁、35-40岁、40-50岁、40-45岁或45-50岁,例如25岁、26岁、27岁、28岁、29岁、30岁、31岁、32岁、33岁、34岁、35岁、36岁、37岁、38岁、39岁、40岁、41岁、42岁、43岁、44岁、45岁、46岁、47岁、48岁、49岁或50岁)的受试者。如果受试者具有双等位基因OTOF突变并且被鉴定为具有外毛细胞完整性(耳声发射和/或耳蜗微音效应的存在)和/或内毛细胞完整性(总和电位的存在)的可检测指标(例如,在治疗之前,被鉴定为具有可检测的耳声发射、耳蜗微音效应和/或总和电位),则受试者也可如本文所述进行治疗。因此,本文所述的方法可包括在治疗受试者之前评估外毛细胞完整性和内毛细胞完整性的步骤。本文所述的组合物和方法可用于治疗具有OTOF突变(例如降低OTOF功能或表达的突变或与感觉神经性听力损失或听神经病变相关的OTOF突变)的受试者、具有常染色体隐性感觉神经性听力损失或听神经病变家族史(例如OTOF相关听力损失家族史)的受试者或OTOF突变状态和/或OTOF活性水平未知的受试者。本文所述的方法可包括在用本文所述的组合物治疗之前或施用本文所述的组合物之前筛选

受试者的OTOF突变的步骤。受试者的OTOF突变可使用本领域技术人类员已知的标准方法(例如遗传性测试)来筛选。本文所述的方法还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听力的步骤。听力可使用标准测试(例如测听术、ABR、耳蜗电描记术(ECOG)和耳声发射)来评价。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听力损失或听神经病变的风险的患者,例如具有遗传性听力损失家族史的患者或尚未表现出听力损失或损害的携带OTOF突变的患者。

[0367] 治疗可包括以多个单位剂量施用含有本文所述的核酸载体(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类员的技能范围内。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耳蜗的损伤。在核酸载体为AAV载体(例如AAV1、AAV2、AAV 2quad(Y-F)、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AA V9、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,AAV载体可具有例如以下滴度:约1×10^{9}个载体基因组(VG)/mL至约1×10^{16}VG/mL(例如1×10^{9}VG/mL、2×10^{9}VG/mL、3×10^{9}VG/mL、4×10^{9}VG/mL、5×10^{9}VG/mL、6×10^{9}VG/mL、7×10^{9}VG/mL、8×10^{9}VG/mL、9×10^{9}VG/mL、1×10^{10}VG/mL、2×10^{10}VG/mL、3×10^{10}VG/mL、4×10^{10}VG/mL、5×10^{10}VG/mL、6×10^{10}VG/mL、7×10^{10}VG/mL、8×10^{10}VG/mL、9×10^{10}VG/mL、1×10^{11}VG/mL、2×10^{11}VG/mL、3×10^{11}VG/mL、4×10^{11}VG/mL、5×10^{11}VG/mL、6×10^{11}VG/mL、7×10^{11}VG/mL、8×10^{11}VG/mL、9×10^{11}VG/mL、1×10^{12}VG/mL、2×10^{12}VG/mL、3×10^{12}VG/mL、4×10^{12}VG/mL、5×10^{12}VG/mL、6×10^{12}VG/mL、7×10^{12}VG/mL、8×10^{12}VG/mL、9×10^{12}VG/mL、1×10^{13}VG/mL、2×10^{13}VG/mL、3×10^{13}VG/mL、4×10^{13}VG/mL、5×10^{13}VG/mL、6×10^{13}VG/mL、7×10^{13}VG/mL、8×10^{13}VG/mL、9×10^{13}VG/mL、1×10^{14}VG/mL、2×10^{14}VG/mL、3×10^{14}VG/mL、4×10^{14}VG/mL、5×10^{14}VG/mL、6×10^{14}VG/mL、7×10^{14}VG/mL、8×10^{14}VG/mL、9×10^{14}VG/mL、1×10^{15}VG/mL、2×10^{15}VG/mL、3×10^{15}VG/mL、4×10^{15}VG/mL、5×10^{15}VG/mL、6×10^{15}VG/mL、7×10^{15}VG/mL、8×10^{15}VG/mL、9×10^{15}VG/mL或1×10^{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1×10^{7}VG/耳至约2×10^{15}VG/耳(例如1×10^{7}VG/耳、2×10^{7}VG/耳、3×10^{7}VG/耳、4×10^{7}VG/耳、5×10^{7}VG/耳、6×10^{7}VG/耳、7×10^{7}VG/耳、8×10^{7}VG/耳、9×10^{7}VG/耳、1×10^{8}VG/耳、2×10^{8}VG/耳、3×10^{8}VG/耳、4×10^{8}VG/耳、5×10^{8}VG/耳、6×10^{8}VG/耳、7×10^{8}VG/耳、8×10^{8}VG/耳、9×10^{8}VG/耳、1×10^{9}VG/耳、2×10^{9}VG/耳、3×10^{9}VG/耳、4×10^{9}VG/耳、5×10^{9}VG/耳、6×10^{9}VG/耳、7×10^{9}VG/耳、8×10^{9}VG/耳、9×10^{9}VG/耳、1×10^{10}VG/耳、2×10^{10}VG/耳、3×10^{10}VG/耳、4×10^{10}VG/耳、5×10^{10}VG/耳、6×10^{10}VG/耳、7×10^{10}VG/耳、8×10^{10}VG/耳、9×10^{10}VG/耳、1×10^{11}VG/耳、2×10^{11}VG/耳、3×10^{11}VG/耳、4×10^{11}VG/耳、5×10^{11}VG/耳、6×10^{11}VG/耳、7×10^{11}VG/耳、8×10^{11}VG/耳、9×10^{11}VG/耳、1×10^{12}VG/耳、2×10^{12}VG/耳、3×10^{12}VG/耳、4×10^{12}VG/耳、5×10^{12}VG/耳、6×10^{12}VG/耳、7×10^{12}VG/耳、8×10^{12}VG/耳、9×10^{12}VG/耳、1×10^{13}VG/耳、2×10^{13}VG/耳、3×10^{13}VG/耳、4×10^{13}VG/耳、5×10^{13}VG/耳、6×10^{13}VG/耳、7×10^{13}VG/耳、8×10^{13}VG/耳、9×10^{13}VG/耳、1×10^{14}VG/耳、2×10^{14}VG/耳、3×10^{14}VG/耳、4×10^{14}VG/耳、5×10^{14}VG/耳、6×10^{14}VG/耳、7×10^{14}VG/耳、$8\times$

10^{14}VG/耳、9×10^{14}VG/耳、1×10^{15}VG/耳或2×10^{15}VG/耳)。在一些实施方案中,核酸载体(例如AAV载体)以足以用第一载体和第二载体两者转导至少20%的受试者的内毛细胞的量施用(例如,用双载体系统的两种载体转导至少20%、30%、40%、50%、60%、70%、80%、90%、95%或更多的受试者的内毛细胞)。

[0368] 本文所述的组合物是以足以改善听力、改进言语识别、增加WT OTOF表达(例如耳蜗毛细胞、例如内毛细胞中的表达)或增加OTOF功能的量来施用。听力可使用标准听力测试(例如测听术、ABR、耳蜗电描记术(ECOG)和耳声发射)来评估并且与治疗之前获得的听力测量值相比可提高5%或更大(例如5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100%或更大)。在一些实施方案中,组合物是以足以改进受试者理解言语的能力的量施用。本文所述的组合物也可以足以缓解或防止感觉神经性听力损失或听神经病变的发生或进展的量施用(例如在携带OTOF突变或具有常染色体隐性听力损失家族史但未表现出听力损害的受试者中,或在表现出轻度至中度听力损失的受试者中)。OTOF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OTOF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%或更大)。OTOF功能可直接(例如使用电生理方法或成像方法来评价胞吐作用)或间接基于听力测试来评估,并且与施用本文所述的组合物之前的OTOF功能相比可增加5%或更大(例如5%、10%、15%、20%、30%、40%、50%、60%、70%、80%、90%、100%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。

[0369] 药盒

[0370] 本文所述的组合物可以药盒提供以用于治疗患有双等位基因OTOF突变的25岁或25岁以上的受试者(例如以治疗这种受试者的感觉神经性听力损失或听神经病变),或用于治疗具有双等位基因OTOF突变的受试者,所述受试者被鉴定为具有可检测的耳声发射、可检测的耳蜗微音效应和/或可检测的总和电位(例如,以治疗这种受试者的感觉神经性听力损失或听神经病)。组合物可包括本文所述的核酸载体(例如AAV载体)(例如含有编码OTOF蛋白的N末端部分的多核苷酸的第一核酸载体和含有编码OTOF蛋白的C末端部分的多核苷酸的第二核酸载体),任选地包装于AAV病毒衣壳(例如AAV1、AAV9、AAV2、AAV8、Anc80、Anc80L65、DJ/9或7m8)中。所述药盒可还包括指导药盒的使用者(例如医师)实施本文所述方法的包装插页。所述药盒可任选地包括注射器或用于施用组合物的其他器件。

[0371] 实施例

[0372] 提出以下实施例以向本领域的普通技术人类员提供可如何使用、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类视为其发明的范围。

[0373] 实施例1-用OTOF双载体治疗的32周龄和52周龄OTOF缺陷型小鼠中的ABR恢复

[0374] 动物由于年龄的增长而逐渐丧失听力,这部分是由于丧失外毛细胞功能。耳畸蛋白缺陷型动物缺乏ABR(内毛细胞功能),但存在畸变产物耳声发射(DPOAE)(外毛细胞功能)。与其他衰老动物一样,耳畸蛋白无效动物随着年龄的增长而丧失外毛细胞功能和

DPOAE。

[0375] 以3.9x10^{10}vg/耳的剂量向多达52周龄的老年OTOF纯合突变型(OTOF-Q828X)动物给予双杂合AAV1-Myo15-hOTOF载体以测试功效治疗窗,考虑到后期可能的外毛细胞损失和DPOAE升高。第一载体含有可操作地连接至含有编码OTOF同功型5蛋白的多核苷酸的外显子1-20的多核苷酸(SEQ ID NO:71)的SEQ ID NO:38的Myo15启动子、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的AP重组基因区(SEQ ID NO:65);并且第二载体含有AP重组基因区(SEQ ID NO:65)、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF同功型5蛋白的多核苷酸的外显子21-45的多核苷酸(SEQ ID NO:72)和poly(A)序列。

[0376] 在32周龄(n=15)和52周龄(n=15)动物中记录基线DPOAE。在32周龄动物中,4/15具有升高的基线DPOAE,而7/15 52周龄动物具有升高的基线DPOAE。

[0377] 在异氟烷麻醉下通过圆窗向动物给予媒介物(n=5/年龄组)或双杂合AAV1-Myo15-hOTOF(n=10/年龄组)。根据方案,允许动物在手术后恢复。在产后4周和8周测试DPOAE和ABR。

[0378] 在10/10的32周龄病毒治疗动物和9/10的52周龄病毒治疗动物中观察到ABR恢复,包括在术后4周和8周时具有基线DPOAE升高的动物。治疗后4周的ABR恢复显示于图1中。在22.6kHz音频下观察到最佳恢复,并且与在较年轻动物中观察到的恢复相似。

[0379] 实施例2-OTOF-Q828X小鼠模型毛细胞损失的表征

[0380] 开发OTOF-Q828X小鼠模型以模拟由耳畸蛋白损失引起的人类先天性耳聋。人类耳畸蛋白Q829X突变(参考SNP rs80356593)是外显子22中的充分研究的停止增益突变,导致在1997个氨基酸的编码序列的828个氨基酸之后的耳畸蛋白的截短。使用CRISPR介导的敲入在FVB品系背景上产Otof-Q828X小鼠品系,其在小鼠OTOF(mOtof)中具有模拟这种人类等位基因的靶向突变。

[0381] 进行实验以评估纯合Otof-Q828X(Otof-Q828X hom)和杂合(Otof-Q828X het)小鼠中的毛细胞损失。在对应于Otof-Q828X het或hom小鼠的5.6kHz、8kHz、11.3kHz、16kHz、22.6kHz、32kHz和45.2kHz的耳蜗区域中,从5至42周龄对IHC(图2A)和OHC(图2B)的数量进行计数。通过毛细胞特异性标记物染色对离散IHC和OHC进行计数。本分析评估了50只耳朵。

[0382] 对于Otof-Q828X hom小鼠中的所有测试频率,IHC计数随年龄增加而出现统计学显著性损失。对于更低频率,在het小鼠中观察到IHC计数的类似趋势(Kendall秩相关)。Otof-828X hom和het动物中的IHC计数在16周内保持稳定(图2A)。从16周后开始,Otof-Q828Xhom动物显示在22.6-45.2kHz下并且在24周后在更低频率(8-16kHz)下开始IHC计数的减少。对于16和32kHz,在24周后Otof-Q828Xhet小鼠的IHC计数开始减少。在32周后,对于大多数测试频率(<45.2kHz),保留了超过75%的IHC(图2A)。

[0383] 在所有频率内,Otof-Q828X hom和het小鼠中的外毛细胞数量在6个月的研究时间过程中保持恒定,但8kHz除外,在8kHz下het小鼠显示计数的年龄相关性降低(Kendall秩相关)。5.6kHz和45.2kHz中的OHC计数与较大变异性相关,显示了het与hom之间的计数差异,所述差异在所测试的年龄内散布。大多数OHC在32周后仍保留(图2B)。

[0384] 实施例3-纯合OTOF-Q828X突变型小鼠的ABR阈值恢复与表达耳畸蛋白的细胞数量之间的关系

[0385] 使用具有AAV1或AAV2quadYF衣壳和smCBA或Myo15启动子的耳畸蛋白双杂合载体

系统,在几项研究中在n＝76只年龄>4周(介于4周龄与34周龄之间)并且接受介于1.0x10^{9}与6.4x10^{10}vg/耳之间剂量的纯合OTOF-Q828X突变型小鼠中检查ABR阈值恢复与表达耳畸蛋白的细胞数量之间的关系。4至34周后,在小鼠介于10周龄与44周龄之间时测量ABR阈值。在这些研究期间施用的双杂合载体系统包括AAV2quadYF-smCBA(SEQ ID NO:70)-mOTOF(施用于34和29周龄小鼠)、AAV2quadYF-Myo15(SEQ ID NO:38)-mOTOF(施用于29周龄小鼠)、AAV2quadYF-Myo15(SEQ ID NO:48)-mOTOF(施用于29周龄小鼠)、AAV1-smCBA(SEQ ID NO:70)-hOTOF(施用于4和8周龄小鼠)、AAV1-Myo15(SEQ ID NO:38)-hOTOF(施用于4和5周龄小鼠)和AAV1-Myo15(SEQ ID NO:38)-mOTOF(施用于9周龄小鼠)。

[0386] 在22kHz,当约20％的IHC表达耳畸蛋白时,ABR阈值进入正常范围(平均值±2SD)(图3)。随着表达耳畸蛋白的IHC的比例更高,阈值没有进一步改善。

[0387] 实施例4-向具有双等位基因OTOF突变的25岁以上的受试者施用OTOF双载体系统

[0388] 根据本文所述的方法,本领域技术的医师可治疗25岁以上(例如25-50岁、25-45岁、25-40岁、25-35岁、25-30岁、30-50岁、30-45岁、30-40岁、30-35岁、35-50岁、35-45岁、35-40岁、40-50岁、40-45岁或45-50岁,例如25岁、26岁、27岁、28岁、29岁、30岁、31岁、32岁、33岁、34岁、35岁、36岁、37岁、38岁、39岁、40岁、41岁、42岁、43岁、44岁、45岁、46岁、47岁、48岁、49岁或50岁)的具有双等位基因OTOF突变的患者,以预防、减轻或治疗听力损失或听神经病变。为此,本领域技术人类员可向人类患者施用组合物,所述组合物含有第一核酸载体(例如,AAV1或AAV9载体),所述第一核酸载体含有可操作地连接至编码OTOF蛋白的N末端部分(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5的N末端部分)的多核苷酸的启动子;和第二核酸载体(例如,AAV1或AAV9载体),所述第二核酸载体含有编码OTOF蛋白的C末端部分(例如人类OTOF,例如SEQ ID NO:1或SEQ ID NO:5的C末端部分)的多核苷酸和poly(A)序列。双载体可以是如本文所述的重叠双载体、反式剪接双载体或双杂合载体。例如,载体可以是双杂合载体,其中第一载体含有可操作地连接至编码OTOF同功型5蛋白的多核苷酸(例如人类OTOF同功型5,例如SEQ ID NO:5,例如具有SEQ ID NO:71的序列的多核苷酸)的外显子1-20的Myo15启动子(例如,SEQ ID NO:36、38、39、48或49)、多核苷酸序列3'的剪接供体序列和剪接供体序列3'的AP重组基因区(例如AP基因片段,诸如SEQ ID NO:62-67中的任一者,例如SEQ ID NO:65),并且其中第二载体含有AP重组基因区(AP基因片段,诸如SEQ ID NO:62-67中的任一者,例如SEQ ID NO:65)、重组基因区3'的剪接受体序列、剪接受体序列3'的含有编码OTOF同功型5蛋白的多核苷酸(例如人类OTOF同功型5,例如SEQ ID NO:5,例如具有SEQ ID NO:72的序列的多核苷酸)的外显子21-45和47的多核苷酸和bGH poly(A)序列。含有重叠双AAV载体的组合物可例如通过局部施用于内耳(例如,通过圆窗膜注射)而施用于患者来治疗或预防与双等位基因OTOF突变相关的感觉神经性听力损失或听神经病的发展。

[0389] 向患者施用组合物后,本领域技术的专业人类员可通过多种方法监测患者对疗法的反应的改进。例如,医师可在施用组合物后通过实施标准测试(例如测听术、ABR、耳蜗电描记术(ECOG))和耳声发射来监测患者的听力。与施用组合物之前的听力测试结果相比,患者在施用组合物后在一种或多种测试中表现出改善的听力的发现指示,患者对治疗具有有利反应。可按需要确定和施用后续剂量。

[0390] 本发明的示例性实施方案描述于以下所列举段落中。

[0391] E1.一种治疗具有双等位基因耳畸蛋白(OTOF)突变的25岁或25岁以上的人类受试者的方法,所述方法包括向所述受试者施用治疗有效量的双载体系统,所述双载体系统包含:

[0392] 第一核酸载体,所述第一核酸载体包含可操作地连接至编码OTOF蛋白的N末端部分的第一编码多核苷酸的启动子;以及

[0393] 第二核酸载体,所述第二核酸载体包含编码OTOF蛋白的C末端部分的第二编码多核苷酸和定位于所述第二编码多核苷酸3'的poly(A)序列;

[0394] 其中所述第一核酸载体和所述第二核酸载体都不编码全长OTOF蛋白。

[0395] E2.如E1所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸并不重叠。

[0396] E3.如E1或E2所述的方法,其中所述第一核酸载体包含定位于所述第一编码多核苷酸3'的剪接供体信号序列,并且所述第二核酸载体包含定位于所述第二编码多核苷酸5'的剪接受体信号序列。

[0397] E4.如E3所述的方法,其中所述第一核酸载体包含定位于所述剪接供体信号序列3'的第一重组基因区,并且所述第二核酸载体包含定位于所述剪接受体信号序列5'的第二重组基因区。

[0398] E5.如E4所述的方法,其中所述第一重组基因区和所述第二重组基因区是相同的。

[0399] E6.如E4或E5所述的方法,其中所述第一重组基因区或所述第二重组基因区是AP基因片段或F1噬菌体AK基因。

[0400] E7.如E6所述的方法,其中所述F1噬菌体AK基因包含SEQ ID NO:19的序列或由其组成。

[0401] E8.如E6所述的方法,其中所述AP基因片段包含SEQ ID NO:62-67中任一者的序列或由其组成。

[0402] E9.如E8所述的方法,其中所述AP基因片段包含SEQ ID NO:65的序列或由其组成。

[0403] E10.如E3-E9中任一项所述的方法,其中所述剪接供体序列包含SEQ ID NO:20或SEQ ID NO:68的序列或由其组成。

[0404] E11.如E3-E10中任一项所述的方法,其中所述剪接受体序列包含SEQ ID NO:21或SEQ ID NO:69的序列或由其组成。

[0405] E12.如E4-E11中任一项所述的方法,其中所述第一核酸载体还包含定位于所述重组基因区3'的降解信号序列;并且其中所述第二核酸载体还包含定位于所述重组基因区与所述剪接受体信号序列之间的降解信号序列。

[0406] E13.如E12所述的方法,其中所述降解信号序列包含SEQ ID NO:22的序列或由其组成。

[0407] E14.如E1-E13中任一项所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸在OTOF外显子边界处分开。

[0408] E15.如E14所述的方法,其中所述OTOF外显子边界不在编码C2结构域的第一编码多核苷酸或第二编码多核苷酸的一部分内。

[0409] E16.如E1所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸部分重叠。

[0410] E17.如E16所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸重叠至少1千碱基(kb)。

[0411] E18.如E16或E17所述的方法,其中所述第一编码多核苷酸与所述第二编码多核苷酸之间的重叠区以OTOF外显子边界为中心。

[0412] E19.如E18所述的方法,其中所述第一编码多核苷酸编码所述OTOF蛋白的N末端部分,并且包含OTOF N末端至所述重叠区中心处的所述外显子边界3'的500bp;并且所述第二编码多核苷酸编码所述OTOF蛋白的C末端部分,并且包含所述重叠区中心处的所述外显子边界5'的500bp至所述OTOF C末端。

[0413] E20.如E18或E19所述的方法,其中所述重叠区中心处的OTOF外显子边界不在编码C2结构域的第一编码多核苷酸或第二编码多核苷酸的一部分内。

[0414] E21.如E14、E15和E18-E20中任一项所述的方法,其中选择所述OTOF外显子边界,使得所述第一编码多核苷酸编码整个C2C结构域,并且所述第二编码多核苷酸编码整个C2D结构域。

[0415] E22.如E14、E15和E18-E21中任一项所述的方法,其中所述OTOF外显子边界是外显子19/20边界、外显子20/21边界或外显子21/22边界。

[0416] E23.如E14、E15和E18-E20中任一项所述的方法,其中选择所述OTOF外显子边界,使得所述第一编码多核苷酸编码整个C2D结构域,并且所述第二编码多核苷酸编码整个C2E结构域。

[0417] E24.如E14、E15、E18-E20和E23中任一项所述的方法,其中所述OTOF外显子边界是外显子26/27边界或外显子28/29边界。

[0418] E25.如E14、E18和E19中任一项所述的方法,其中所述OTOF外显子边界在编码C2D结构域的第一编码多核苷酸和第二编码多核苷酸的一部分内。

[0419] E26.如E14、E18、E19和E25中任一项所述的方法,其中所述OTOF外显子边界是外显子24/25边界或外显子25/26边界。

[0420] E27.如E1-E26中任一项所述的方法,其中所述第一编码多核苷酸和所述第二编码多核苷酸中的每一者编码约一半的所述OTOF蛋白序列。

[0421] E28.如E1-E27中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体不包含OTOF非翻译区(UTR)。

[0422] E29.如E1-E27中任一项所述的方法,其中所述第一核酸载体包含OTOF 5'UTR。

[0423] E30.如E1-E27和E29中任一项所述的方法,其中所述第二核酸载体包含OTOF 3'UTR。

[0424] E31.如E1-E30中任一项所述的方法,其中编码所述OTOF蛋白的所述第一编码多核苷酸和所述第二编码多核苷酸不包含内含子。

[0425] E32.如E1-E31中任一项所述的方法,其中所述OTOF蛋白是哺乳动物OTOF蛋白。

[0426] E33.如E32所述的方法,其中所述OTOF蛋白是人类OTOF蛋白。

[0427] E34.如E1-E33中任一项所述的方法,其中所述人类OTOF蛋白与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列具有至少85%序列同一性(例如,85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性)。

[0428] E35.如E34所述的方法,其中所述OTOF蛋白具有SEQ ID NO:1的序列。

[0429] E36.如E34所述的方法,其中所述OTOF蛋白具有SEQ ID NO:5的序列。

[0430] E37.如E1-E33中任一项所述的方法,其中所述OTOF蛋白包含SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4或SEQ ID NO:5的序列或其具有一个或多个保守氨基酸取代的变体。

[0431] E38.如E37所述的方法,其中所述OTOF蛋白变体中不超过10%(例如,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。

[0432] E39.如E1-E33中任一项所述的方法,其中所述OTOF蛋白由SEQ ID NO:10-14中的任一者编码。

[0433] E40.如E1-E33中任一项所述的方法,其中所述第一编码多核苷酸编码SEQ ID NO:1或SEQ ID NO:5的氨基酸1-802,并且第二编码多核苷酸编码SEQ ID NO:1或SEQ ID NO:5的氨基酸803-1997。

[0434] E41.如E1-E33中任一项所述的方法,其中所述OTOF蛋白的N末端部分由SEQ ID NO:73的序列或其具有一个或多个保守氨基酸取代的变体组成。

[0435] E42.如E41所述的方法,其中所述OTOF蛋白变体的N末端部分中不超过10%(例如,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。

[0436] E43.如E41所述的方法,其中所述OTOF蛋白的N末端部分由SEQ ID NO:73的序列组成。

[0437] E44.如E1-E33和E43中任一项所述的方法,所述OTOF蛋白的N末端部分由SEQ ID NO:71的序列编码。

[0438] E45.如E1-E33和E41-E44中任一项所述的方法,其中所述OTOF蛋白的C末端部分由SEQ ID NO:74的序列或其具有一个或多个保守氨基酸取代的变体组成。

[0439] E46.如E45所述的方法,其中所述OTOF蛋白变体的C末端部分中不超过10%(例如,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。

[0440] E47.如E45所述的方法,其中所述OTOF蛋白的C末端部分由SEQ ID NO:74的序列组成。

[0441] E48.如E1-E33、E41-E44和E47中任一项所述的方法,其中所述OTOF蛋白的C末端部分由SEQ ID NO:72的序列编码。

[0442] E49.如E1-E48中任一项所述的方法,其中所述第一核酸载体包含所述启动子3'和编码所述OTOF蛋白的所述N末端部分的所述第一编码多核苷酸5'的Kozak序列。

[0443] E50.如E1-E49中任一项所述的方法,其中所述启动子是普遍存在的启动子。

[0444] E51.如E50所述的方法,其中所述普遍存在的启动子是CAG启动子、巨细胞病毒(CMV)启动子、鸡β-肌动蛋白启动子、截短CMV-鸡β-肌动蛋白启动子(smCBA)、CB7启动子、杂合CMV增强子/人类β-肌动蛋白启动子、人类β-肌动蛋白启动子、延伸因子-1α(EF1α)启动子或磷酸甘油酸激酶(PGK)启动子。

[0445] E52.如E1-E49中任一项所述的方法,其中所述启动子是耳蜗毛细胞特异性启动子。

[0446] E53.如E52所述的方法,其中所述耳蜗毛细胞特异性启动子是肌凝蛋白15(Myo15)启动子、肌凝蛋白7A(Myo7A)启动子、肌凝蛋白6(Myo6)启动子、POU 4类同源盒3(POU4F3)启

动子、atonal BHLH转录因子1(ATOH1)启动子、LIM同源盒3(LHX3)启动子、α9乙酰胆碱受体(α9AChR)启动子或α10乙酰胆碱受体(α10AChR)启动子。

[0447] E54.如E1-E49中任一项所述的方法,其中所述启动子是内毛细胞特异性启动子。

[0448] E55.如E54所述的方法,其中所述内毛细胞特异性启动子是成纤维细胞生长因子8(FGF8)启动子、囊泡谷氨酸转运蛋白3(VGLUT3)启动子、OTOF启动子或钙结合蛋白2(CABP2)启动子。

[0449] E56.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:75的核苷酸2272至6041的序列。

[0450] E57.如E1或E56所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:75的核苷酸2049至6264的序列或由其组成。

[0451] E58.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:77的核苷酸182至3949的序列。

[0452] E59.如E1或E58所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:77的核苷酸19至4115的序列或由其组成。

[0453] E60.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:79的核苷酸2267至6014的序列。

[0454] E61.如E1或E60所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:79的核苷酸2049至6237的序列或由其组成。

[0455] E62.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:80的核苷酸177至3924的序列。

[0456] E63.如E1或E62所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:80的核苷酸19至4090的序列或由其组成。

[0457] E64.如E1、E56、E57、E60和E61中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:76的核苷酸2267至6476的序列。

[0458] E65.如E1、E56、E57、E60、E61和E64中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:76的核苷酸2049至6693的序列或由其组成。

[0459] E66.如E1、E58、E59、E62和E63中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:78的核苷酸187至4396的序列。

[0460] E67.如E1、E58、E59、E62、E63和E66中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:78的核苷酸19至4589的序列或由其组成。

[0461] E68.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:81的核苷酸235至4004的序列。

[0462] E69.如E1或E62所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:81的核苷酸12至4227的序列或由其组成。

[0463] E70.如E1所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:83的核苷酸230至3977的序列。

[0464] E71.如E1或E70所述的方法,其中所述第一核酸载体包含多核苷酸序列,所述多核

苷酸序列包含SEQ ID NO:83的核苷酸12至4200的序列或由其组成。

[0465] E72.如E1和E68-E71中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:82的核苷酸229至4438的序列。

[0466] E73.如E1和E68-E72中任一项所述的方法,其中所述第二核酸载体包含多核苷酸序列,所述多核苷酸序列包含SEQ ID NO:82的核苷酸12至4655的序列或由其组成。

[0467] E74.如E1-E73中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体在所述核酸序列的每个末端处包含反向末端重复(ITR)。

[0468] E75.如E74所述的方法,其中所述ITR是AAV2 ITR或与AAV2ITR具有至少80%序列同一性(例如,80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%序列同一性)。

[0469] E76.如E1-E75中任一项所述的方法,其中所述poly(A)序列是牛生长激素(bGH)poly(A)信号序列。

[0470] E77.如E1-E76中任一项所述的方法,其中所述第二核酸载体包含土拨鼠肝炎病毒转录后调控元件(WPRE)。

[0471] E78.如E77所述的方法,其中所述WPRE包含SEQ ID NO:23或SEQ ID NO:61的序列或由其组成。

[0472] E79.如E1-E78中任一项所述的方法,其中所述第一核酸载体和所述第二核酸载体是腺相关病毒(AAV)载体。

[0473] E80.如E79所述的方法,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和PHP.S衣壳。

[0474] E81.如E1-E80中任一项所述的方法,其中所述受试者是30岁或30岁以上。

[0475] E82.如E1-E80中任一项所述的方法,其中所述受试者是35岁或35岁以上。

[0476] E83.如E1-E80中任一项所述的方法,其中所述受试者是40岁或40岁以上。

[0477] E84.如E1-E80中任一项所述的方法,其中所述受试者是45岁或45岁以上。

[0478] E85.如E1-E84中任一项所述的方法,其中所述受试者不大于50岁。

[0479] E86.如E1-E85中任一项所述的方法,其中所述受试者已被鉴定为具有双等位基因OTOF突变。

[0480] E87.如E1-E85中任一项所述的方法,其中所述方法还包括在施用所述双载体系统之前将所述受试者鉴定为具有双等位基因OTOF突变。

[0481] E88.如E1-E87中任一项所述的方法,其中所述受试者被鉴定为具有可检测的耳声发射。

[0482] E89.如E1-E88中任一项所述的方法,其中所述受试者被鉴定为具有可检测的耳蜗微音效应。

[0483] E90.如E1-E89中任一项所述的方法,其中所述受试者被鉴定为具有可检测的总和电位。

[0484] E91.如E1-E90中任一项所述的方法,其中所述受试者患有或被鉴定为患有9型常染色体隐性耳聋(DFNB9)。

[0485] E92.如E1-E91中任一项所述的方法,其中所述方法还包括在施用所述双载体系统

之前评估所述受试者的听力。

[0486] E93.如E1-E92中任一项所述的方法,其中将所述双载体系统施用至内耳。

[0487] E94.如E93所述的方法,其中所述双载体系统是通过注射穿过圆窗膜、注射至半规管中、小管造口术、通过圆窗膜插入导管、经鼓室注射或鼓室内注射来施用。

[0488] E95.如E1-E94中任一项所述的方法,其中所述方法还包括在施用所述双载体系统之后评估所述受试者的听力。

[0489] E96.如E1-E95中任一项所述的方法,其中所述双载体系统是以足以预防或减轻听力损失、延迟听力损失的发生、减缓听力损失的进展、改善听力、改进言语识别或改进毛细胞功能的量施用。

[0490] E97.如E1-E96中任一项所述的方法,其中所述第一载体和所述第二载体是同时施用的。

[0491] E98.如E1-E96中任一项所述的方法,其中所述第一载体和所述第二载体是相继施用的。

[0492] E99.如E1-E98中任一项所述的方法,其中所述第一载体和所述第二载体以约$1x10^{7}$个载体基因组(VG)/耳至约$2x10^{15}$VG/耳的浓度施用。

[0493] E100.如E1-E99中任一项所述的方法,其中所述第一载体和所述第二载体以合起来足以用所述第一载体和所述第二载体两者转导至少20%的所述受试者的内毛细胞的量施用。

[0494] 其他实施方案

[0495] 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类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,对本领域技术人类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> 使用耳畸蛋白双载体系统治疗感觉神经性听力损失的方法

<130> 51471-008WO2

<150> US 63/151,589

<151> 2021-02-19

<160> 84

<170> PatentIn version 3.5

<210> 1

<211> 1997

<212> PRT

<213> 智人(Homo sapiens)

<400> 1

Met Ala Leu Leu Ile His Leu Lys Thr Val Ser Glu Leu Arg Gly Arg
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Asp Val Ala Asp Phe Asp Glu Thr Phe
35 40 45
Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Met Leu Glu Ile
50 55 60
Gln Val Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80
Phe Arg Met Val Leu Gln Lys Val Val Glu Glu Ser His Val Glu Val
85 90 95
Thr Asp Thr Leu Ile Asp Asp Asn Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110
Cys Val Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Ser Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Glu Lys
130 135 140
Asp Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Ser
145 150 155 160
Arg Pro Pro Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe
165 170 175
Ser Ala Met Lys Leu Gly Lys Asn Arg Ser His Lys Glu Glu Pro Gln
180 185 190
Arg Pro Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu

195 200 205

Ala Ile Arg Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala

210 215 220

Ser Val Thr Ala Leu Thr Thr Asn Val Ser Asn Lys Arg Ser Lys Pro

225 230 235 240

Asp Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val

245 250 255

Ser Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp

260 265 270

Pro Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met

275 280 285

Lys Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp

290 295 300

Phe His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser

305 310 315 320

Val Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser

325 330 335

Phe Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe

340 345 350

His His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ser Gly

355 360 365

Leu Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp

370 375 380

Asn Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile

385 390 395 400

Glu Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp

405 410 415

Ala Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met

420 425 430

Asn Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn

435 440 445

Lys Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys

450 455 460

Gly Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu

465 470 475 480

Gln Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys

485 490 495

Val Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr

500 505 510

His Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe
515 520 525
Leu Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg
530 535 540
Asn Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly
545 550 555 560
Glu Gly Val Ser Phe Arg Ala Arg Leu Leu Leu Gly Leu Ala Val Glu
565 570 575
Ile Val Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln
580 585 590
Val Glu Gln Ala Thr Pro Ile Ser Glu Ser Cys Ala Gly Lys Met Glu
595 600 605
Glu Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg
610 615 620
Arg Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr
625 630 635 640
Gly Asn Glu Val Asp Gly Leu Ser Arg Pro Gln Arg Pro Arg Pro Arg
645 650 655
Lys Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ala Ser
660 665 670
Asp Asp Glu Ala Gly Asp Ala Gly Asp Leu Ala Ser Val Ser Ser Thr
675 680 685
Pro Pro Met Arg Pro Gln Val Thr Asp Arg Asn Tyr Phe His Leu Pro
690 695 700
Tyr Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp
705 710 715 720
Gln Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp
725 730 735
Lys Leu Glu Glu Gly Leu Asn Asp Ile Gln Glu Met Ile Lys Thr Glu
740 745 750
Lys Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser
755 760 765
Cys Gly Cys Cys Arg Phe Leu Ser Leu Ala Asp Lys Asp Gln Gly His
770 775 780
Ser Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg
785 790 795 800
Glu Leu Glu Asn Met Gly Gln Gln Ala Arg Met Leu Arg Ala Gln Val
805 810 815
Lys Arg His Thr Val Arg Asp Lys Leu Arg Leu Cys Gln Asn Phe Leu

820 825 830

Gln Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp

835 840 845

Ile Phe Ile Trp Met Met Ser Asn Asn Lys Arg Val Ala Tyr Ala Arg

850 855 860

Val Pro Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Thr Gly

865 870 875 880

Lys Asp Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys

885 890 895

Arg Gly Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Val Glu Leu

900 905 910

Tyr Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys Glu Phe Leu Cys Gly

915 920 925

Leu Pro Cys Gly Phe Gln Glu Val Lys Ala Ala Gln Gly Leu Gly Leu

930 935 940

His Ala Phe Pro Pro Val Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe

945 950 955 960

Gln Leu Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp

965 970 975

Ser Ser Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln

980 985 990

Ser Gln Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp

995 1000 1005

Gln Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His

1010 1015 1020

Glu Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp

1025 1030 1035

Gln Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala

1040 1045 1050

Lys Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg

1055 1060 1065

Phe Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala

1070 1075 1080

Thr Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly

1085 1090 1095

Pro Ala Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp

1100 1105 1110

Val Asp Arg Gly Pro Ile Met Pro Val Pro Met Gly Ile Arg Pro

1115 1120 1125

Val Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg
1130 1135 1140
Asp Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val
1145 1150 1155
Asp Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His
1160 1165 1170
Asn Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe
1175 1180 1185
Glu Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn
1190 1195 1200
Ile Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val
1205 1210 1215
Gly Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro
1220 1225 1230
Pro Asp Arg Ser Ala Pro Ser Trp Asn Thr Thr Val Arg Leu Leu
1235 1240 1245
Arg Arg Cys Arg Val Leu Cys Asn Gly Gly Ser Ser Ser His Ser
1250 1255 1260
Thr Gly Glu Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys
1265 1270 1275
Lys Leu Glu Thr Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val
1280 1285 1290
Val Lys Val Asp Val Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys
1295 1300 1305
Lys Lys Gly Thr Ala Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu
1310 1315 1320
Ser Met Leu Asp Trp Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr
1325 1330 1335
Met Lys Glu Gln Leu Arg Gln Gln Glu Pro Ser Gly Ile Asp Leu
1340 1345 1350
Glu Glu Lys Glu Glu Val Asp Asn Thr Glu Gly Leu Lys Gly Ser
1355 1360 1365
Met Lys Gly Lys Glu Lys Ala Arg Ala Ala Lys Glu Glu Lys Lys
1370 1375 1380
Lys Lys Thr Gln Ser Ser Gly Ser Gly Gln Gly Ser Glu Ala Pro
1385 1390 1395
Glu Lys Lys Lys Pro Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys
1400 1405 1410
Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu Asp Trp Leu His Thr

1415 1420 1425

Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser

1430 1435 1440

Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly Ser Leu Cys

1445 1450 1455

Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu Ala Gly

1460 1465 1470

Tyr Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn Asp

1475 1480 1485

Pro Ile Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp

1490 1495 1500

Leu His Pro Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala

1505 1510 1515

Ile Arg Leu Gly Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile

1520 1525 1530

Ser Lys Gln Leu Asn Pro Val Phe Gly Lys Ser Phe Asp Ile Glu

1535 1540 1545

Ala Ser Phe Pro Met Glu Ser Met Leu Thr Val Ala Val Tyr Asp

1550 1555 1560

Trp Asp Leu Val Gly Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile

1565 1570 1575

Asp Leu Glu Asn Arg Phe Tyr Ser Lys His Arg Ala Thr Cys Gly

1580 1585 1590

Ile Ala Gln Thr Tyr Ser Thr His Gly Tyr Asn Ile Trp Arg Asp

1595 1600 1605

Pro Met Lys Pro Ser Gln Ile Leu Thr Arg Leu Cys Lys Asp Gly

1610 1615 1620

Lys Val Asp Gly Pro His Phe Gly Pro Pro Gly Arg Val Lys Val

1625 1630 1635

Ala Asn Arg Val Phe Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn

1640 1645 1650

Gly Gln Arg Lys Pro Thr Asp Glu His Val Ala Leu Leu Ala Leu

1655 1660 1665

Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys Arg Leu Val Pro

1670 1675 1680

Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly

1685 1690 1695

Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe Pro Met

1700 1705 1710

Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys
1715 1720 1725
Pro Lys Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu
1730 1735 1740
Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser
1745 1750 1755
Asp Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys
1760 1765 1770
Gln Asp Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn
1775 1780 1785
Phe Asn Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu
1790 1795 1800
Glu Lys Ile Val Ile Ser Lys Lys Glu Ser Met Phe Ser Trp Asp
1805 1810 1815
Glu Thr Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp
1820 1825 1830
Asp Ala Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu
1835 1840 1845
Leu Asp Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln
1850 1855 1860
Cys Thr Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val
1865 1870 1875
Ser Ile Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu
1880 1885 1890
Ala Arg Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu
1895 1900 1905
Ala Glu Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro
1910 1915 1920
Val Gly Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn
1925 1930 1935
Arg Pro Asp Thr Ser Phe Ile Trp Phe Leu Asn Pro Leu Lys Ser
1940 1945 1950
Ala Arg Tyr Phe Leu Trp His Thr Tyr Arg Trp Leu Leu Leu Lys
1955 1960 1965
Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Ala Leu Phe Leu
1970 1975 1980
Tyr Ser Val Pro Gly Tyr Leu Val Lys Lys Ile Leu Gly Ala
1985 1990 1995
<210> 2

<211> 1230
<212> PRT
<213> 智人(Homo sapiens)
<400> 2

Met Ile Lys Thr Glu Lys Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val
1 5 10 15
Leu Glu Glu Leu Ser Cys Gly Cys Cys Arg Phe Leu Ser Leu Ala Asp
20 25 30
Lys Asp Gln Gly His Ser Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu
35 40 45
Lys Ser Cys Met Arg Glu Leu Glu Asn Met Gly Gln Gln Ala Arg Met
50 55 60
Leu Arg Ala Gln Val Lys Arg His Thr Val Arg Asp Lys Leu Arg Leu
65 70 75 80
Cys Gln Asn Phe Leu Gln Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln
85 90 95
His Ser Ile Pro Asp Ile Phe Ile Trp Met Met Ser Asn Asn Lys Arg
100 105 110
Val Ala Tyr Ala Arg Val Pro Ser Lys Asp Leu Leu Phe Ser Ile Val
115 120 125
Glu Glu Glu Thr Gly Lys Asp Cys Ala Lys Val Lys Thr Leu Phe Leu
130 135 140
Lys Leu Pro Gly Lys Arg Gly Phe Gly Ser Ala Gly Trp Thr Val Gln
145 150 155 160
Ala Lys Val Glu Leu Tyr Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys
165 170 175
Glu Phe Leu Cys Gly Leu Pro Cys Gly Phe Gln Glu Val Lys Ala Ala
180 185 190
Gln Gly Leu Gly Leu His Ala Phe Pro Pro Val Ser Leu Val Tyr Thr
195 200 205
Lys Lys Gln Ala Phe Gln Leu Arg Ala His Met Tyr Gln Ala Arg Ser
210 215 220
Leu Phe Ala Ala Asp Ser Ser Gly Leu Ser Asp Pro Phe Ala Arg Val
225 230 235 240
Phe Phe Ile Asn Gln Ser Gln Cys Thr Glu Val Leu Asn Glu Thr Leu
245 250 255
Cys Pro Thr Trp Asp Gln Met Leu Val Phe Asp Asn Leu Glu Leu Tyr
260 265 270
Gly Glu Ala His Glu Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu

275 280 285

Ile Tyr Asp Gln Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr

290 295 300

Phe Ala Lys Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro

305 310 315 320

Arg Phe Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala

325 330 335

Thr Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro

340 345 350

Ala Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Val Asp

355 360 365

Arg Gly Pro Ile Met Pro Val Pro Met Gly Ile Arg Pro Val Leu Ser

370 375 380

Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp Leu Lys Arg

385 390 395 400

Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp Ile Glu Cys Ala

405 410 415

Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn Tyr Lys Lys Asn Pro

420 425 430

Asn Phe Asn Thr Leu Val Lys Trp Phe Glu Val Asp Leu Pro Glu Asn

435 440 445

Glu Leu Leu His Pro Pro Leu Asn Ile Arg Val Val Asp Cys Arg Ala

450 455 460

Phe Gly Arg Tyr Thr Leu Val Gly Ser His Ala Val Ser Ser Leu Arg

465 470 475 480

Arg Phe Ile Tyr Arg Pro Pro Asp Arg Ser Ala Pro Ser Trp Asn Thr

485 490 495

Thr Gly Glu Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys Lys

500 505 510

Leu Glu Thr Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val Val Lys

515 520 525

Val Asp Val Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys Lys Lys Gly

530 535 540

Thr Ala Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp

545 550 555 560

Trp Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu

565 570 575

Arg Gln Gln Glu Pro Ser Gly Ile Asp Leu Glu Glu Lys Glu Glu Val

580 585 590

Asp Asn Thr Glu Gly Leu Lys Gly Ser Met Lys Gly Lys Glu Lys Ala
595 600 605
Arg Ala Ala Lys Glu Glu Lys Lys Lys Lys Thr Gln Ser Ser Gly Ser
610 615 620
Gly Gln Gly Ser Glu Ala Pro Glu Lys Lys Lys Pro Lys Ile Asp Glu
625 630 635 640
Leu Lys Val Tyr Pro Lys Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu
645 650 655
Asp Trp Leu His Thr Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp
660 665 670
Glu Asp Gly Ser Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly
675 680 685
Ser Leu Cys Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu
690 695 700
Ala Gly Tyr Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn
705 710 715 720
Asp Pro Ile Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp
725 730 735
Leu His Pro Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile
740 745 750
Arg Leu Gly Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys
755 760 765
Gln Leu Asn Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe
770 775 780
Pro Met Glu Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val
785 790 795 800
Gly Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg
805 810 815
Phe Tyr Ser Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr Ser
820 825 830
Thr His Gly Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser Gln Ile
835 840 845
Leu Thr Arg Leu Cys Lys Asp Gly Lys Val Asp Gly Pro His Phe Gly
850 855 860
Pro Pro Gly Arg Val Lys Val Ala Asn Arg Val Phe Thr Gly Pro Ser
865 870 875 880
Glu Ile Glu Asp Glu Asn Gly Gln Arg Lys Pro Thr Asp Glu His Val
885 890 895
Ala Leu Leu Ala Leu Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys

900 905 910
Arg Leu Val Pro Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp
915 920 925
Lys Pro Gly Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe
930 935 940
Pro Met Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg
945 950 955 960
Lys Pro Lys Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu
965 970 975
Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp
980 985 990
Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp
995 1000 1005
Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn
1010 1015 1020
Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys
1025 1030 1035
Ile Val Ile Ser Lys Lys Glu Ser Met Phe Ser Trp Asp Glu Thr
1040 1045 1050
Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala
1055 1060 1065
Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp
1070 1075 1080
Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr
1085 1090 1095
Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val Ser Ile
1100 1105 1110
Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu Ala Arg
1115 1120 1125
Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu Ala Glu
1130 1135 1140
Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro Val Gly
1145 1150 1155
Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro
1160 1165 1170
Asp Thr Ser Phe Ile Trp Phe Leu Asn Pro Leu Lys Ser Ala Arg
1175 1180 1185
Tyr Phe Leu Trp His Thr Tyr Arg Trp Leu Leu Leu Lys Leu Leu
1190 1195 1200

Leu Leu Leu Leu Leu Leu Leu Leu Leu Ala Leu Phe Leu Tyr Ser
1205 1210 1215
Val Pro Gly Tyr Leu Val Lys Lys Ile Leu Gly Ala
1220 1225 1230

<210> 3
<211> 1230
<212> PRT
<213> 智人(Homo sapiens)
<400> 3

Met Ile Lys Thr Glu Lys Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val
1 5 10 15
Leu Glu Glu Leu Ser Cys Gly Cys Cys Arg Phe Leu Ser Leu Ala Asp
20 25 30
Lys Asp Gln Gly His Ser Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu
35 40 45
Lys Ser Cys Met Arg Glu Leu Glu Asn Met Gly Gln Gln Ala Arg Met
50 55 60
Leu Arg Ala Gln Val Lys Arg His Thr Val Arg Asp Lys Leu Arg Leu
65 70 75 80
Cys Gln Asn Phe Leu Gln Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln
85 90 95
His Ser Ile Pro Asp Ile Phe Ile Trp Met Met Ser Asn Asn Lys Arg
100 105 110
Val Ala Tyr Ala Arg Val Pro Ser Lys Asp Leu Leu Phe Ser Ile Val
115 120 125
Glu Glu Glu Thr Gly Lys Asp Cys Ala Lys Val Lys Thr Leu Phe Leu
130 135 140
Lys Leu Pro Gly Lys Arg Gly Phe Gly Ser Ala Gly Trp Thr Val Gln
145 150 155 160
Ala Lys Val Glu Leu Tyr Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys
165 170 175
Glu Phe Leu Cys Gly Leu Pro Cys Gly Phe Gln Glu Val Lys Ala Ala
180 185 190
Gln Gly Leu Gly Leu His Ala Phe Pro Pro Val Ser Leu Val Tyr Thr
195 200 205
Lys Lys Gln Ala Phe Gln Leu Arg Ala His Met Tyr Gln Ala Arg Ser
210 215 220
Leu Phe Ala Ala Asp Ser Ser Gly Leu Ser Asp Pro Phe Ala Arg Val
225 230 235 240

Phe Phe Ile Asn Gln Ser Gln Cys Thr Glu Val Leu Asn Glu Thr Leu
245 250 255
Cys Pro Thr Trp Asp Gln Met Leu Val Phe Asp Asn Leu Glu Leu Tyr
260 265 270
Gly Glu Ala His Glu Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu
275 280 285
Ile Tyr Asp Gln Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr
290 295 300
Phe Ala Lys Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro
305 310 315 320
Arg Phe Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala
325 330 335
Thr Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro
340 345 350
Ala Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Val Asp
355 360 365
Arg Gly Pro Ile Met Pro Val Pro Met Gly Ile Arg Pro Val Leu Ser
370 375 380
Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp Leu Lys Arg
385 390 395 400
Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp Ile Glu Cys Ala
405 410 415
Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn Tyr Lys Lys Asn Pro
420 425 430
Asn Phe Asn Thr Leu Val Lys Trp Phe Glu Val Asp Leu Pro Glu Asn
435 440 445
Glu Leu Leu His Pro Pro Leu Asn Ile Arg Val Val Asp Cys Arg Ala
450 455 460
Phe Gly Arg Tyr Thr Leu Val Gly Ser His Ala Val Ser Ser Leu Arg
465 470 475 480
Arg Phe Ile Tyr Arg Pro Pro Asp Arg Ser Ala Pro Ser Trp Asn Thr
485 490 495
Thr Gly Glu Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys Lys
500 505 510
Leu Glu Thr Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val Val Lys
515 520 525
Val Asp Val Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys Lys Lys Gly
530 535 540
Thr Ala Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp

545 550 555 560
Trp Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu
565 570 575
Arg Gln Gln Glu Pro Ser Gly Ile Asp Leu Glu Glu Lys Glu Glu Val
580 585 590
Asp Asn Thr Glu Gly Leu Lys Gly Ser Met Lys Gly Lys Glu Lys Ala
595 600 605
Arg Ala Ala Lys Glu Glu Lys Lys Lys Lys Thr Gln Ser Ser Gly Ser
610 615 620
Gly Gln Gly Ser Glu Ala Pro Glu Lys Lys Lys Pro Lys Ile Asp Glu
625 630 635 640
Leu Lys Val Tyr Pro Lys Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu
645 650 655
Asp Trp Leu His Thr Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp
660 665 670
Glu Asp Gly Ser Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly
675 680 685
Ser Leu Cys Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu
690 695 700
Ala Gly Tyr Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn
705 710 715 720
Asp Pro Ile Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp
725 730 735
Leu His Pro Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile
740 745 750
Arg Leu Gly Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys
755 760 765
Gln Leu Asn Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe
770 775 780
Pro Met Glu Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val
785 790 795 800
Gly Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg
805 810 815
Phe Tyr Ser Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr Ser
820 825 830
Thr His Gly Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser Gln Ile
835 840 845
Leu Thr Arg Leu Cys Lys Asp Gly Lys Val Asp Gly Pro His Phe Gly
850 855 860

Pro Pro Gly Arg Val Lys Val Ala Asn Arg Val Phe Thr Gly Pro Ser
865 870 875 880
Glu Ile Glu Asp Glu Asn Gly Gln Arg Lys Pro Thr Asp Glu His Val
885 890 895
Ala Leu Leu Ala Leu Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys
900 905 910
Arg Leu Val Pro Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp
915 920 925
Lys Pro Gly Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe
930 935 940
Pro Met Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg
945 950 955 960
Lys Pro Lys Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu
965 970 975
Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp
980 985 990
Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp
995 1000 1005
Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn
1010 1015 1020
Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys
1025 1030 1035
Ile Val Ile Ser Lys Lys Glu Ser Met Phe Ser Trp Asp Glu Thr
1040 1045 1050
Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala
1055 1060 1065
Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp
1070 1075 1080
Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr
1085 1090 1095
Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val Ser Ile
1100 1105 1110
Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu Ala Arg
1115 1120 1125
Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu Ala Glu
1130 1135 1140
Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro Val Gly
1145 1150 1155
Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro

1160 1165 1170
Asp Thr Ala Phe Val Trp Phe Leu Asn Pro Leu Lys Ser Ile Lys
1175 1180 1185
Tyr Leu Ile Cys Thr Arg Tyr Lys Trp Leu Ile Ile Lys Ile Val
1190 1195 1200
Leu Ala Leu Leu Gly Leu Leu Met Leu Gly Leu Phe Leu Tyr Ser
1205 1210 1215
Leu Pro Gly Tyr Met Val Lys Lys Leu Leu Gly Ala
1220 1225 1230

<210> 4
<211> 1307
<212> PRT
<213> 智人(Homo sapiens)
<400> 4

Met Met Thr Asp Thr Gln Asp Gly Pro Ser Glu Ser Ser Gln Ile Met
1 5 10 15
Arg Ser Leu Thr Pro Leu Ile Asn Arg Glu Glu Ala Phe Gly Glu Ala
20 25 30
Gly Glu Ala Gly Leu Trp Pro Ser Ile Thr His Thr Pro Asp Ser Gln
35 40 45
Glu Glu Gly Leu Asn Asp Ile Gln Glu Met Ile Lys Thr Glu Lys Ser
50 55 60
Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser Cys Gly
65 70 75 80
Cys Cys Arg Phe Leu Ser Leu Ala Asp Lys Asp Gln Gly His Ser Ser
85 90 95
Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg Glu Leu
100 105 110
Glu Asn Met Gly Gln Gln Ala Arg Met Leu Arg Ala Gln Val Lys Arg
115 120 125
His Thr Val Arg Asp Lys Leu Arg Leu Cys Gln Asn Phe Leu Gln Lys
130 135 140
Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Ile Phe
145 150 155 160
Ile Trp Met Met Ser Asn Asn Lys Arg Val Ala Tyr Ala Arg Val Pro
165 170 175
Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Thr Gly Lys Asp
180 185 190
Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg Gly

195 200 205

Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Val Glu Leu Tyr Leu
210 215 220

Trp Leu Gly Leu Ser Lys Gln Arg Lys Glu Phe Leu Cys Gly Leu Pro
225 230 235 240

Cys Gly Phe Gln Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His Ala
245 250 255

Phe Pro Pro Val Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln Leu
260 265 270

Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser Ser
275 280 285

Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser Gln
290 295 300

Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln Met
305 310 315 320

Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu Leu Arg
325 330 335

Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln Asp Ser Met
340 345 350

Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys Pro Leu Val Lys
355 360 365

Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe Pro Pro Gln Leu Glu
370 375 380

Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala Thr Ala Gly Asp Leu Leu Ala
385 390 395 400

Ala Phe Glu Leu Leu Gln Ile Gly Pro Ala Gly Lys Ala Asp Leu Pro
405 410 415

Pro Ile Asn Gly Pro Val Asp Val Asp Arg Gly Pro Ile Met Pro Val
420 425 430

Pro Met Gly Ile Arg Pro Val Leu Ser Lys Tyr Arg Val Glu Val Leu
435 440 445

Phe Trp Gly Leu Arg Asp Leu Lys Arg Val Asn Leu Ala Gln Val Asp
450 455 460

Arg Pro Arg Val Asp Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser
465 470 475 480

Leu Ile His Asn Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys
485 490 495

Trp Phe Glu Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu
500 505 510

Asn Ile Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val
515 520 525
Gly Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro
530 535 540
Asp Arg Ser Ala Pro Ser Trp Asn Thr Thr Val Arg Leu Leu Arg Arg
545 550 555 560
Cys Arg Val Leu Cys Asn Gly Gly Ser Ser Ser His Ser Thr Gly Glu
565 570 575
Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys Lys Leu Glu Thr
580 585 590
Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val Val Lys Val Asp Val
595 600 605
Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys Lys Lys Gly Thr Ala Glu
610 615 620
Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp Trp Trp Ser
625 630 635 640
Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu Arg Gln Gln
645 650 655
Glu Pro Ser Gly Ile Asp Leu Glu Glu Lys Glu Glu Val Asp Asn Thr
660 665 670
Glu Gly Leu Lys Gly Ser Met Lys Gly Lys Glu Lys Ala Arg Ala Ala
675 680 685
Lys Glu Glu Lys Lys Lys Lys Thr Gln Ser Ser Gly Ser Gly Gln Gly
690 695 700
Ser Glu Ala Pro Glu Lys Lys Lys Pro Lys Ile Asp Glu Leu Lys Val
705 710 715 720
Tyr Pro Lys Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu Asp Trp Leu
725 730 735
His Thr Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp Glu Asp Gly
740 745 750
Ser Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly Ser Leu Cys
755 760 765
Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu Ala Gly Tyr
770 775 780
Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn Asp Pro Ile
785 790 795 800
Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp Leu His Pro
805 810 815
Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile Arg Leu Gly

820 825 830

Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys Gln Leu Asn

835 840 845

Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe Pro Met Glu

850 855 860

Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val Gly Thr Asp

865 870 875 880

Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg Phe Tyr Ser

885 890 895

Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr Ser Thr His Gly

900 905 910

Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser Gln Ile Leu Thr Arg

915 920 925

Leu Cys Lys Asp Gly Lys Val Asp Gly Pro His Phe Gly Pro Pro Gly

930 935 940

Arg Val Lys Val Ala Asn Arg Val Phe Thr Gly Pro Ser Glu Ile Glu

945 950 955 960

Asp Glu Asn Gly Gln Arg Lys Pro Thr Asp Glu His Val Ala Leu Leu

965 970 975

Ala Leu Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys Arg Leu Val

980 985 990

Pro Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly

995 1000 1005

Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe Pro Met

1010 1015 1020

Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys

1025 1030 1035

Pro Lys Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu

1040 1045 1050

Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser

1055 1060 1065

Asp Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys

1070 1075 1080

Gln Asp Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn

1085 1090 1095

Phe Asn Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu

1100 1105 1110

Glu Lys Ile Val Ile Ser Lys Lys Glu Ser Met Phe Ser Trp Asp

1115 1120 1125

Glu Thr Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp
1130 1135 1140
Asp Ala Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu
1145 1150 1155
Leu Asp Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln
1160 1165 1170
Cys Thr Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val
1175 1180 1185
Ser Ile Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu
1190 1195 1200
Ala Arg Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu
1205 1210 1215
Ala Glu Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro
1220 1225 1230
Val Gly Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn
1235 1240 1245
Arg Pro Asp Thr Ser Phe Ile Trp Phe Leu Asn Pro Leu Lys Ser
1250 1255 1260
Ala Arg Tyr Phe Leu Trp His Thr Tyr Arg Trp Leu Leu Leu Lys
1265 1270 1275
Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Ala Leu Phe Leu
1280 1285 1290
Tyr Ser Val Pro Gly Tyr Leu Val Lys Lys Ile Leu Gly Ala
1295 1300 1305

<210> 5
<211> 1997
<212> PRT
<213> 智人(Homo sapiens)
<400> 5

Met Ala Leu Leu Ile His Leu Lys Thr Val Ser Glu Leu Arg Gly Arg
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Asp Val Ala Asp Phe Asp Glu Thr Phe
35 40 45
Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Met Leu Glu Ile
50 55 60
Gln Val Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80

Phe Arg Met Val Leu Gln Lys Val Val Glu Glu Ser His Val Glu Val
85 90 95
Thr Asp Thr Leu Ile Asp Asp Asn Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110
Cys Val Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Ser Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Glu Lys
130 135 140
Asp Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Ser
145 150 155 160
Arg Pro Pro Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe
165 170 175
Ser Ala Met Lys Leu Gly Lys Asn Arg Ser His Lys Glu Glu Pro Gln
180 185 190
Arg Pro Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu
195 200 205
Ala Ile Arg Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala
210 215 220
Ser Val Thr Ala Leu Thr Thr Asn Val Ser Asn Lys Arg Ser Lys Pro
225 230 235 240
Asp Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val
245 250 255
Ser Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp
260 265 270
Pro Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met
275 280 285
Lys Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp
290 295 300
Phe His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser
305 310 315 320
Val Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser
325 330 335
Phe Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe
340 345 350
His His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ser Gly
355 360 365
Leu Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp
370 375 380
Asn Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile

385 390 395 400
Glu Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp
405 410 415
Ala Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met
420 425 430
Asn Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn
435 440 445
Lys Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys
450 455 460
Gly Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu
465 470 475 480
Gln Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys
485 490 495
Val Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr
500 505 510
His Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe
515 520 525
Leu Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg
530 535 540
Asn Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly
545 550 555 560
Glu Gly Val Ser Phe Arg Ala Arg Leu Leu Leu Gly Leu Ala Val Glu
565 570 575
Ile Val Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln
580 585 590
Val Glu Gln Ala Thr Pro Ile Ser Glu Ser Cys Ala Gly Lys Met Glu
595 600 605
Glu Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg
610 615 620
Arg Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr
625 630 635 640
Gly Asn Glu Val Asp Gly Leu Ser Arg Pro Gln Arg Pro Arg Pro Arg
645 650 655
Lys Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ala Ser
660 665 670
Asp Asp Glu Ala Gly Asp Ala Gly Asp Leu Ala Ser Val Ser Ser Thr
675 680 685
Pro Pro Met Arg Pro Gln Val Thr Asp Arg Asn Tyr Phe His Leu Pro
690 695 700

Tyr Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp
705 710 715 720
Gln Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp
725 730 735
Lys Leu Glu Glu Gly Leu Asn Asp Ile Gln Glu Met Ile Lys Thr Glu
740 745 750
Lys Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser
755 760 765
Cys Gly Cys Cys Arg Phe Leu Ser Leu Ala Asp Lys Asp Gln Gly His
770 775 780
Ser Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg
785 790 795 800
Glu Leu Glu Asn Met Gly Gln Gln Ala Arg Met Leu Arg Ala Gln Val
805 810 815
Lys Arg His Thr Val Arg Asp Lys Leu Arg Leu Cys Gln Asn Phe Leu
820 825 830
Gln Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp
835 840 845
Ile Phe Ile Trp Met Met Ser Asn Asn Lys Arg Val Ala Tyr Ala Arg
850 855 860
Val Pro Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Thr Gly
865 870 875 880
Lys Asp Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys
885 890 895
Arg Gly Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Val Glu Leu
900 905 910
Tyr Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys Glu Phe Leu Cys Gly
915 920 925
Leu Pro Cys Gly Phe Gln Glu Val Lys Ala Ala Gln Gly Leu Gly Leu
930 935 940
His Ala Phe Pro Pro Val Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe
945 950 955 960
Gln Leu Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp
965 970 975
Ser Ser Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln
980 985 990
Ser Gln Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp
995 1000 1005
Gln Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His

1010 1015 1020

Glu Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp

1025 1030 1035

Gln Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala

1040 1045 1050

Lys Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg

1055 1060 1065

Phe Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala

1070 1075 1080

Thr Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly

1085 1090 1095

Pro Ala Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp

1100 1105 1110

Val Asp Arg Gly Pro Ile Met Pro Val Pro Met Gly Ile Arg Pro

1115 1120 1125

Val Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg

1130 1135 1140

Asp Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val

1145 1150 1155

Asp Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His

1160 1165 1170

Asn Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe

1175 1180 1185

Glu Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn

1190 1195 1200

Ile Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val

1205 1210 1215

Gly Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro

1220 1225 1230

Pro Asp Arg Ser Ala Pro Ser Trp Asn Thr Thr Val Arg Leu Leu

1235 1240 1245

Arg Arg Cys Arg Val Leu Cys Asn Gly Gly Ser Ser Ser His Ser

1250 1255 1260

Thr Gly Glu Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys

1265 1270 1275

Lys Leu Glu Thr Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val

1280 1285 1290

Val Lys Val Asp Val Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys

1295 1300 1305

Lys Lys Gly Thr Ala Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu
1310 1315 1320
Ser Met Leu Asp Trp Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr
1325 1330 1335
Met Lys Glu Gln Leu Arg Gln Gln Glu Pro Ser Gly Ile Asp Leu
1340 1345 1350
Glu Glu Lys Glu Glu Val Asp Asn Thr Glu Gly Leu Lys Gly Ser
1355 1360 1365
Met Lys Gly Lys Glu Lys Ala Arg Ala Ala Lys Glu Glu Lys Lys
1370 1375 1380
Lys Lys Thr Gln Ser Ser Gly Ser Gly Gln Gly Ser Glu Ala Pro
1385 1390 1395
Glu Lys Lys Lys Pro Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys
1400 1405 1410
Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu Asp Trp Leu His Thr
1415 1420 1425
Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser
1430 1435 1440
Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly Ser Leu Cys
1445 1450 1455
Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu Ala Gly
1460 1465 1470
Tyr Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn Asp
1475 1480 1485
Pro Ile Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp
1490 1495 1500
Leu His Pro Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala
1505 1510 1515
Ile Arg Leu Gly Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile
1520 1525 1530
Ser Lys Gln Leu Asn Pro Val Phe Gly Lys Ser Phe Asp Ile Glu
1535 1540 1545
Ala Ser Phe Pro Met Glu Ser Met Leu Thr Val Ala Val Tyr Asp
1550 1555 1560
Trp Asp Leu Val Gly Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile
1565 1570 1575
Asp Leu Glu Asn Arg Phe Tyr Ser Lys His Arg Ala Thr Cys Gly
1580 1585 1590
Ile Ala Gln Thr Tyr Ser Thr His Gly Tyr Asn Ile Trp Arg Asp

1595 1600 1605

Pro Met Lys Pro Ser Gln Ile Leu Thr Arg Leu Cys Lys Asp Gly

1610 1615 1620

Lys Val Asp Gly Pro His Phe Gly Pro Pro Gly Arg Val Lys Val

1625 1630 1635

Ala Asn Arg Val Phe Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn

1640 1645 1650

Gly Gln Arg Lys Pro Thr Asp Glu His Val Ala Leu Leu Ala Leu

1655 1660 1665

Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys Arg Leu Val Pro

1670 1675 1680

Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly

1685 1690 1695

Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe Pro Met

1700 1705 1710

Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys

1715 1720 1725

Pro Lys Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu

1730 1735 1740

Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser

1745 1750 1755

Asp Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys

1760 1765 1770

Gln Asp Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn

1775 1780 1785

Phe Asn Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu

1790 1795 1800

Glu Lys Ile Val Ile Ser Lys Lys Glu Ser Met Phe Ser Trp Asp

1805 1810 1815

Glu Thr Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp

1820 1825 1830

Asp Ala Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu

1835 1840 1845

Leu Asp Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln

1850 1855 1860

Cys Thr Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val

1865 1870 1875

Ser Ile Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu

1880 1885 1890

Ala Arg Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu
1895 1900 1905
Ala Glu Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro
1910 1915 1920
Val Gly Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn
1925 1930 1935
Arg Pro Asp Thr Ala Phe Val Trp Phe Leu Asn Pro Leu Lys Ser
1940 1945 1950
Ile Lys Tyr Leu Ile Cys Thr Arg Tyr Lys Trp Leu Ile Ile Lys
1955 1960 1965
Ile Val Leu Ala Leu Leu Gly Leu Leu Met Leu Gly Leu Phe Leu
1970 1975 1980
Tyr Ser Leu Pro Gly Tyr Met Val Lys Lys Leu Leu Gly Ala
1985 1990 1995

<210> 6
<211> 1997
<212> PRT
<213> 小家鼠(Mus musculus)
<400> 6

Met Ala Leu Ile Val His Leu Lys Thr Val Ser Glu Leu Arg Gly Lys
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Gly Val Ala Asp Phe Asp Glu Thr Phe
35 40 45
Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Val Leu Glu Ile
50 55 60
Gln Ile Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80
Phe Cys Met Val Leu Gln Lys Val Val Glu Glu Asn Arg Val Glu Val
85 90 95
Thr Asp Thr Leu Met Asp Asp Ser Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110
Ser Met Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Pro Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Lys Asp
130 135 140
Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Thr Arg
145 150 155 160

Ile Ser Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe Ser
165 170 175
Ala Met Lys Leu Gly Lys Thr Arg Ser His Lys Glu Glu Pro Gln Arg
180 185 190
Gln Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu Ala
195 200 205
Ile Gln Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala Ser
210 215 220
Val Thr Ala Leu Thr Ser Asn Val Ser Asn Lys Arg Ser Lys Pro Asp
225 230 235 240
Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val Ser
245 250 255
Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp Pro
260 265 270
Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met Lys
275 280 285
Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp Phe
290 295 300
His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser Val
305 310 315 320
Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser Phe
325 330 335
Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe His
340 345 350
His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ala Gly Leu
355 360 365
Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp Asn
370 375 380
Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile Glu
385 390 395 400
Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp Ala
405 410 415
Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met Asn
420 425 430
Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn Lys
435 440 445
Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys Gly
450 455 460
Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu Gln

465 470 475 480
Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys Val
485 490 495
Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr His
500 505 510
Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe Leu
515 520 525
Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg Asn
530 535 540
Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly Glu
545 550 555 560
Gly Val Ser Phe Arg Ala Arg Leu Met Leu Gly Leu Ala Val Glu Ile
565 570 575
Leu Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln Val
580 585 590
Glu Gln Ala Thr Pro Val Ser Glu Ser Cys Thr Gly Arg Met Glu Glu
595 600 605
Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg Lys
610 615 620
Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr Gly
625 630 635 640
Asn Glu Val Asp Gly Met Ser Arg Pro Leu Arg Pro Arg Pro Arg Lys
645 650 655
Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ser Ser Asp
660 665 670
Asp Glu Gly Asp Glu Ala Gly Asp Leu Ala Ser Val Ser Ser Thr Pro
675 680 685
Pro Met Arg Pro Gln Ile Thr Asp Arg Asn Tyr Phe His Leu Pro Tyr
690 695 700
Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp Gln
705 710 715 720
Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp Lys
725 730 735
Leu Glu Glu Gly Leu Asn Asp Val Gln Glu Met Ile Lys Thr Glu Lys
740 745 750
Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser Cys
755 760 765
Gly Cys His Arg Phe Leu Ser Leu Ser Asp Lys Asp Gln Gly Arg Ser
770 775 780

Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg Glu
785 790 795 800
Leu Glu Ser Met Gly Gln Gln Ala Lys Ser Leu Arg Ala Gln Val Lys
805 810 815
Arg His Thr Val Arg Asp Lys Leu Arg Ser Cys Gln Asn Phe Leu Gln
820 825 830
Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Val
835 840 845
Phe Ile Trp Met Met Ser Asn Asn Lys Arg Ile Ala Tyr Ala Arg Val
850 855 860
Pro Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Leu Gly Lys
865 870 875 880
Asp Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg
885 890 895
Gly Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Leu Glu Leu Tyr
900 905 910
Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys Asp Phe Leu Cys Gly Leu
915 920 925
Pro Cys Gly Phe Glu Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His
930 935 940
Ser Phe Pro Pro Ile Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln
945 950 955 960
Leu Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser
965 970 975
Ser Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser
980 985 990
Gln Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln
995 1000 1005
Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu
1010 1015 1020
Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln
1025 1030 1035
Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys
1040 1045 1050
Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe
1055 1060 1065
Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Ser Ala Thr
1070 1075 1080
Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro

1085 1090 1095

Ser Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Met

1100 1105 1110

Asp Arg Gly Pro Ile Met Pro Val Pro Val Gly Ile Arg Pro Val

1115 1120 1125

Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp

1130 1135 1140

Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp

1145 1150 1155

Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn

1160 1165 1170

Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe Glu

1175 1180 1185

Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn Ile

1190 1195 1200

Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val Gly

1205 1210 1215

Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro

1220 1225 1230

Asp Arg Ser Ala Pro Asn Trp Asn Thr Thr Val Arg Leu Leu Arg

1235 1240 1245

Gly Cys His Arg Leu Arg Asn Gly Gly Pro Ser Ser Arg Pro Thr

1250 1255 1260

Gly Glu Val Val Val Ser Met Glu Pro Glu Glu Pro Val Lys Lys

1265 1270 1275

Leu Glu Thr Met Val Lys Leu Asp Ala Thr Ser Asp Ala Val Val

1280 1285 1290

Lys Val Asp Val Ala Glu Asp Glu Lys Glu Arg Lys Lys Lys Lys

1295 1300 1305

Lys Lys Gly Pro Ser Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu

1310 1315 1320

Ser Met Leu Asp Trp Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr

1325 1330 1335

Met Lys Glu Gln Leu Arg Gln His Glu Thr Ser Gly Thr Asp Leu

1340 1345 1350

Glu Glu Lys Glu Glu Met Glu Ser Ala Glu Gly Leu Lys Gly Pro

1355 1360 1365

Met Lys Ser Lys Glu Lys Ser Arg Ala Ala Lys Glu Glu Lys Lys

1370 1375 1380

Lys Lys Asn Gln Ser Pro Gly Pro Gly Gln Gly Ser Glu Ala Pro
1385 1390 1395
Glu Lys Lys Lys Ala Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys
1400 1405 1410
Glu Leu Glu Ser Glu Phe Asp Ser Phe Glu Asp Trp Leu His Thr
1415 1420 1425
Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser
1430 1435 1440
Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly Ser Leu Cys
1445 1450 1455
Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu Ala Gly
1460 1465 1470
Tyr Asp Pro Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn Asp
1475 1480 1485
Pro Ile Asn Val Leu Val Arg Ile Tyr Val Val Arg Ala Thr Asp
1490 1495 1500
Leu His Pro Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala
1505 1510 1515
Ile Lys Leu Gly Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile
1520 1525 1530
Ser Lys Gln Leu Asn Pro Val Phe Gly Lys Ser Phe Asp Ile Glu
1535 1540 1545
Ala Ser Phe Pro Met Glu Ser Met Leu Thr Val Ala Val Tyr Asp
1550 1555 1560
Trp Asp Leu Val Gly Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile
1565 1570 1575
Asp Leu Glu Asn Arg Phe Tyr Ser Lys His Arg Ala Thr Cys Gly
1580 1585 1590
Ile Ala Gln Thr Tyr Ser Ile His Gly Tyr Asn Ile Trp Arg Asp
1595 1600 1605
Pro Met Lys Pro Ser Gln Ile Leu Thr Arg Leu Cys Lys Glu Gly
1610 1615 1620
Lys Val Asp Gly Pro His Phe Gly Pro His Gly Arg Val Arg Val
1625 1630 1635
Ala Asn Arg Val Phe Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn
1640 1645 1650
Gly Gln Arg Lys Pro Thr Asp Glu His Val Ala Leu Ser Ala Leu
1655 1660 1665
Arg His Trp Glu Asp Ile Pro Arg Val Gly Cys Arg Leu Val Pro

1670 1675 1680
Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly
1685 1690 1695
Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe Pro Met
1700 1705 1710
Asp Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys
1715 1720 1725
Pro Lys Lys Tyr Glu Leu Arg Val Ile Val Trp Asn Thr Asp Glu
1730 1735 1740
Val Val Leu Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser
1745 1750 1755
Asp Ile Phe Val Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys
1760 1765 1770
Gln Asp Thr Asp Val His Tyr His Ser Leu Thr Gly Glu Gly Asn
1775 1780 1785
Phe Asn Trp Arg Tyr Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu
1790 1795 1800
Glu Lys Ile Val Met Ser Lys Lys Glu Ser Met Phe Ser Trp Asp
1805 1810 1815
Glu Thr Glu Tyr Lys Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp
1820 1825 1830
Asp Ala Asp His Phe Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu
1835 1840 1845
Leu Asp Leu Asn Arg Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln
1850 1855 1860
Cys Thr Met Glu Met Ala Thr Gly Glu Val Asp Val Pro Leu Val
1865 1870 1875
Ser Ile Phe Lys Gln Lys Arg Val Lys Gly Trp Trp Pro Leu Leu
1880 1885 1890
Ala Arg Asn Glu Asn Asp Glu Phe Glu Leu Thr Gly Lys Val Glu
1895 1900 1905
Ala Glu Leu His Leu Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro
1910 1915 1920
Val Gly Leu Ala Arg Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn
1925 1930 1935
Arg Pro Asp Thr Ser Phe Ile Trp Phe Leu Asn Pro Leu Lys Ser
1940 1945 1950
Ala Arg Tyr Phe Leu Trp His Thr Tyr Arg Trp Leu Leu Leu Lys
1955 1960 1965

Phe Leu Leu Leu Phe Leu Leu Leu Leu Leu Phe Ala Leu Phe Leu
1970 1975 1980
Tyr Ser Leu Pro Gly Tyr Leu Ala Lys Lys Ile Leu Gly Ala
1985 1990 1995

<210> 7
<211> 1977
<212> PRT
<213> 小家鼠(Mus musculus)
<400> 7

Met Ala Leu Ile Val His Leu Lys Thr Val Ser Glu Leu Arg Gly Lys
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Gly Val Ala Asp Phe Asp Glu Thr Phe
35 40 45
Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Val Leu Glu Ile
50 55 60
Gln Ile Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80
Phe Cys Met Val Leu Gln Lys Val Val Glu Glu Asn Arg Val Glu Val
85 90 95
Thr Asp Thr Leu Met Asp Asp Ser Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110
Ser Met Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Pro Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Lys Asp
130 135 140
Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Thr Arg
145 150 155 160
Ile Ser Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe Ser
165 170 175
Ala Met Lys Leu Gly Lys Thr Arg Ser His Lys Glu Glu Pro Gln Arg
180 185 190
Gln Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu Ala
195 200 205
Ile Gln Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala Ser
210 215 220
Val Thr Ala Leu Thr Ser Asn Val Ser Asn Lys Arg Ser Lys Pro Asp
225 230 235 240

Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val Ser
245 250 255
Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp Pro
260 265 270
Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met Lys
275 280 285
Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp Phe
290 295 300
His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser Val
305 310 315 320
Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser Phe
325 330 335
Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe His
340 345 350
His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ala Gly Leu
355 360 365
Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp Asn
370 375 380
Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile Glu
385 390 395 400
Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp Ala
405 410 415
Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met Asn
420 425 430
Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn Lys
435 440 445
Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys Gly
450 455 460
Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu Gln
465 470 475 480
Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys Val
485 490 495
Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr His
500 505 510
Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe Leu
515 520 525
Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg Asn
530 535 540
Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly Glu

545 550 555 560
Gly Val Ser Phe Arg Ala Arg Leu Met Leu Gly Leu Ala Val Glu Ile
565 570 575
Leu Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln Val
580 585 590
Glu Gln Ala Thr Pro Val Ser Glu Ser Cys Thr Gly Arg Met Glu Glu
595 600 605
Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg Lys
610 615 620
Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr Gly
625 630 635 640
Asn Glu Val Asp Gly Met Ser Arg Pro Leu Arg Pro Arg Pro Arg Lys
645 650 655
Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ser Ser Asp
660 665 670
Asp Glu Gly Asp Glu Ala Gly Asp Leu Ala Ser Val Ser Ser Thr Pro
675 680 685
Pro Met Arg Pro Gln Ile Thr Asp Arg Asn Tyr Phe His Leu Pro Tyr
690 695 700
Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp Gln
705 710 715 720
Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp Lys
725 730 735
Leu Glu Glu Gly Leu Asn Asp Val Gln Glu Met Ile Lys Thr Glu Lys
740 745 750
Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser Cys
755 760 765
Gly Cys His Arg Phe Leu Ser Leu Ser Asp Lys Asp Gln Gly Arg Ser
770 775 780
Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg Glu
785 790 795 800
Leu Glu Ser Met Gly Gln Gln Ala Lys Ser Leu Arg Ala Gln Val Lys
805 810 815
Arg His Thr Val Arg Asp Lys Leu Arg Ser Cys Gln Asn Phe Leu Gln
820 825 830
Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Val
835 840 845
Phe Ile Trp Met Met Ser Asn Asn Lys Arg Ile Ala Tyr Ala Arg Val
850 855 860

Pro Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Leu Gly Lys
865 870 875 880
Asp Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg
885 890 895
Gly Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Leu Glu Leu Tyr
900 905 910
Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys Asp Phe Leu Cys Gly Leu
915 920 925
Pro Cys Gly Phe Glu Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His
930 935 940
Ser Phe Pro Pro Ile Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln
945 950 955 960
Leu Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser
965 970 975
Ser Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser
980 985 990
Gln Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln
995 1000 1005
Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu
1010 1015 1020
Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln
1025 1030 1035
Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys
1040 1045 1050
Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe
1055 1060 1065
Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Ser Ala Thr
1070 1075 1080
Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro
1085 1090 1095
Ser Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Met
1100 1105 1110
Asp Arg Gly Pro Ile Met Pro Val Pro Val Gly Ile Arg Pro Val
1115 1120 1125
Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp
1130 1135 1140
Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp
1145 1150 1155
Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn

1160 1165 1170

Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe Glu

1175 1180 1185

Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn Ile

1190 1195 1200

Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val Gly

1205 1210 1215

Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro

1220 1225 1230

Asp Arg Ser Ala Pro Asn Trp Asn Thr Thr Gly Glu Val Val Val

1235 1240 1245

Ser Met Glu Pro Glu Glu Pro Val Lys Lys Leu Glu Thr Met Val

1250 1255 1260

Lys Leu Asp Ala Thr Ser Asp Ala Val Val Lys Val Asp Val Ala

1265 1270 1275

Glu Asp Glu Lys Glu Arg Lys Lys Lys Lys Lys Lys Gly Pro Ser

1280 1285 1290

Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp Trp

1295 1300 1305

Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu

1310 1315 1320

Arg Gln His Glu Thr Ser Gly Thr Asp Leu Glu Glu Lys Glu Glu

1325 1330 1335

Met Glu Ser Ala Glu Gly Leu Lys Gly Pro Met Lys Ser Lys Glu

1340 1345 1350

Lys Ser Arg Ala Ala Lys Glu Glu Lys Lys Lys Lys Asn Gln Ser

1355 1360 1365

Pro Gly Pro Gly Gln Gly Ser Glu Ala Pro Glu Lys Lys Lys Ala

1370 1375 1380

Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys Glu Leu Glu Ser Glu

1385 1390 1395

Phe Asp Ser Phe Glu Asp Trp Leu His Thr Phe Asn Leu Leu Arg

1400 1405 1410

Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser Thr Glu Glu Glu Arg

1415 1420 1425

Ile Val Gly Arg Phe Lys Gly Ser Leu Cys Val Tyr Lys Val Pro

1430 1435 1440

Leu Pro Glu Asp Val Ser Arg Glu Ala Gly Tyr Asp Pro Thr Tyr

1445 1450 1455

Gly Met Phe Gln Gly Ile Pro Ser Asn Asp Pro Ile Asn Val Leu
1460 1465 1470
Val Arg Ile Tyr Val Val Arg Ala Thr Asp Leu His Pro Ala Asp
1475 1480 1485
Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile Lys Leu Gly Lys
1490 1495 1500
Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys Gln Leu Asn
1505 1510 1515
Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe Pro Met
1520 1525 1530
Glu Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val Gly
1535 1540 1545
Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg
1550 1555 1560
Phe Tyr Ser Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr
1565 1570 1575
Ser Ile His Gly Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser
1580 1585 1590
Gln Ile Leu Thr Arg Leu Cys Lys Glu Gly Lys Val Asp Gly Pro
1595 1600 1605
His Phe Gly Pro His Gly Arg Val Arg Val Ala Asn Arg Val Phe
1610 1615 1620
Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn Gly Gln Arg Lys Pro
1625 1630 1635
Thr Asp Glu His Val Ala Leu Ser Ala Leu Arg His Trp Glu Asp
1640 1645 1650
Ile Pro Arg Val Gly Cys Arg Leu Val Pro Glu His Val Glu Thr
1655 1660 1665
Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly Ile Glu Gln Gly Arg
1670 1675 1680
Leu Glu Leu Trp Val Asp Met Phe Pro Met Asp Met Pro Ala Pro
1685 1690 1695
Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys Pro Lys Lys Tyr Glu
1700 1705 1710
Leu Arg Val Ile Val Trp Asn Thr Asp Glu Val Val Leu Glu Asp
1715 1720 1725
Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp Ile Phe Val Arg
1730 1735 1740
Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp Thr Asp Val

1745 1750 1755
His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn Trp Arg Tyr
1760 1765 1770
Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys Ile Val Met
1775 1780 1785
Ser Lys Lys Glu Ser Met Phe Ser Trp Asp Glu Thr Glu Tyr Lys
1790 1795 1800
Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala Asp His Phe
1805 1810 1815
Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp Leu Asn Arg
1820 1825 1830
Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr Met Glu Met
1835 1840 1845
Ala Thr Gly Glu Val Asp Val Pro Leu Val Ser Ile Phe Lys Gln
1850 1855 1860
Lys Arg Val Lys Gly Trp Trp Pro Leu Leu Ala Arg Asn Glu Asn
1865 1870 1875
Asp Glu Phe Glu Leu Thr Gly Lys Val Glu Ala Glu Leu His Leu
1880 1885 1890
Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro Val Gly Leu Ala Arg
1895 1900 1905
Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro Asp Thr Ser
1910 1915 1920
Phe Ile Trp Phe Leu Asn Pro Leu Lys Ser Ala Arg Tyr Phe Leu
1925 1930 1935
Trp His Thr Tyr Arg Trp Leu Leu Leu Lys Phe Leu Leu Leu Phe
1940 1945 1950
Leu Leu Leu Leu Leu Phe Ala Leu Phe Leu Tyr Ser Leu Pro Gly
1955 1960 1965
Tyr Leu Ala Lys Lys Ile Leu Gly Ala
1970 1975

<210> 8
<211> 1992
<212> PRT
<213> 小家鼠(Mus musculus)
<400> 8

Met Ala Leu Ile Val His Leu Lys Thr Val Ser Glu Leu Arg Gly Lys
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser

20 25 30

Arg Val Leu Glu Asn Cys Glu Gly Val Ala Asp Phe Asp Glu Thr Phe
35 40 45

Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Val Leu Glu Ile
50 55 60

Gln Ile Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80

Phe Cys Met Val Leu Gln Lys Val Val Glu Glu Asn Arg Val Glu Val
85 90 95

Thr Asp Thr Leu Met Asp Asp Ser Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110

Ser Met Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Pro Trp
115 120 125

Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Lys Asp
130 135 140

Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Thr Arg
145 150 155 160

Ile Ser Gly Glu Lys Ser Phe Arg Ser Lys Gly Arg Glu Lys Thr Lys
165 170 175

Gly Gly Arg Asp Gly Glu His Lys Ala Gly Arg Ser Val Phe Ser Ala
180 185 190

Met Lys Leu Gly Lys Thr Arg Ser His Lys Glu Glu Pro Gln Arg Gln
195 200 205

Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu Ala Ile
210 215 220

Gln Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala Ser Val
225 230 235 240

Thr Ala Leu Thr Ser Asn Val Ser Asn Lys Arg Ser Lys Pro Asp Ile
245 250 255

Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val Ser Ile
260 265 270

Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp Pro Val
275 280 285

Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met Lys Glu
290 295 300

Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp Phe His
305 310 315 320

Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser Val Ile
325 330 335

His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser Phe Lys
340 345 350
Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe His His
355 360 365
Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ala Gly Leu Lys
370 375 380
Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp Asn Ile
385 390 395 400
Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile Glu Gly
405 410 415
Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp Ala Arg
420 425 430
Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met Asn Thr
435 440 445
Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn Lys Asp
450 455 460
Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys Gly Lys
465 470 475 480
Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu Gln Val
485 490 495
Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys Val Gln
500 505 510
Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr His Phe
515 520 525
Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe Leu Pro
530 535 540
Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg Asn Tyr
545 550 555 560
Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly Glu Gly
565 570 575
Val Ser Phe Arg Ala Arg Leu Met Leu Gly Leu Ala Val Glu Ile Leu
580 585 590
Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln Val Glu
595 600 605
Gln Ala Thr Pro Val Ser Glu Ser Cys Thr Gly Arg Met Glu Glu Phe
610 615 620
Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg Lys Asn
625 630 635 640
Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr Gly Asn

645 650 655

Glu Val Asp Gly Met Ser Arg Pro Leu Arg Pro Arg Pro Arg Lys Glu

660 665 670

Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ser Ser Asp Asp

675 680 685

Glu Gly Asp Glu Ala Gly Asp Leu Ala Ser Val Ser Ser Thr Pro Pro

690 695 700

Met Arg Pro Gln Ile Thr Asp Arg Asn Tyr Phe His Leu Pro Tyr Leu

705 710 715 720

Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp Gln Arg

725 730 735

Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp Lys Leu

740 745 750

Glu Glu Gly Leu Asn Asp Val Gln Glu Met Ile Lys Thr Glu Lys Ser

755 760 765

Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser Cys Gly

770 775 780

Cys His Arg Phe Leu Ser Leu Ser Asp Lys Asp Gln Gly Arg Ser Ser

785 790 795 800

Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg Glu Leu

805 810 815

Glu Ser Met Gly Gln Gln Ala Lys Ser Leu Arg Ala Gln Val Lys Arg

820 825 830

His Thr Val Arg Asp Lys Leu Arg Ser Cys Gln Asn Phe Leu Gln Lys

835 840 845

Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Val Phe

850 855 860

Ile Trp Met Met Ser Asn Asn Lys Arg Ile Ala Tyr Ala Arg Val Pro

865 870 875 880

Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Leu Gly Lys Asp

885 890 895

Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg Gly

900 905 910

Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Leu Glu Leu Tyr Leu

915 920 925

Trp Leu Gly Leu Ser Lys Gln Arg Lys Asp Phe Leu Cys Gly Leu Pro

930 935 940

Cys Gly Phe Glu Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His Ser

945 950 955 960

Phe Pro Pro Ile Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln Leu
965 970 975
Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser Ser
980 985 990
Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser Gln
995 1000 1005
Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln
1010 1015 1020
Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu
1025 1030 1035
Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln
1040 1045 1050
Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys
1055 1060 1065
Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe
1070 1075 1080
Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Ser Ala Thr
1085 1090 1095
Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro
1100 1105 1110
Ser Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Met
1115 1120 1125
Asp Arg Gly Pro Ile Met Pro Val Pro Val Gly Ile Arg Pro Val
1130 1135 1140
Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp
1145 1150 1155
Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp
1160 1165 1170
Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn
1175 1180 1185
Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe Glu
1190 1195 1200
Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn Ile
1205 1210 1215
Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val Gly
1220 1225 1230
Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro
1235 1240 1245
Asp Arg Ser Ala Pro Asn Trp Asn Thr Thr Gly Glu Val Val Val

1250 1255 1260

Ser Met Glu Pro Glu Glu Pro Val Lys Lys Leu Glu Thr Met Val

1265 1270 1275

Lys Leu Asp Ala Thr Ser Asp Ala Val Val Lys Val Asp Val Ala

1280 1285 1290

Glu Asp Glu Lys Glu Arg Lys Lys Lys Lys Lys Lys Gly Pro Ser

1295 1300 1305

Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp Trp

1310 1315 1320

Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu

1325 1330 1335

Arg Gln His Glu Thr Ser Gly Thr Asp Leu Glu Glu Lys Glu Glu

1340 1345 1350

Met Glu Ser Ala Glu Gly Leu Lys Gly Pro Met Lys Ser Lys Glu

1355 1360 1365

Lys Ser Arg Ala Ala Lys Glu Glu Lys Lys Lys Lys Asn Gln Ser

1370 1375 1380

Pro Gly Pro Gly Gln Gly Ser Glu Ala Pro Glu Lys Lys Lys Ala

1385 1390 1395

Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys Glu Leu Glu Ser Glu

1400 1405 1410

Phe Asp Ser Phe Glu Asp Trp Leu His Thr Phe Asn Leu Leu Arg

1415 1420 1425

Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser Thr Glu Glu Glu Arg

1430 1435 1440

Ile Val Gly Arg Phe Lys Gly Ser Leu Cys Val Tyr Lys Val Pro

1445 1450 1455

Leu Pro Glu Asp Val Ser Arg Glu Ala Gly Tyr Asp Pro Thr Tyr

1460 1465 1470

Gly Met Phe Gln Gly Ile Pro Ser Asn Asp Pro Ile Asn Val Leu

1475 1480 1485

Val Arg Ile Tyr Val Val Arg Ala Thr Asp Leu His Pro Ala Asp

1490 1495 1500

Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile Lys Leu Gly Lys

1505 1510 1515

Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys Gln Leu Asn

1520 1525 1530

Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe Pro Met

1535 1540 1545

Glu Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val Gly
1550 1555 1560
Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg
1565 1570 1575
Phe Tyr Ser Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr
1580 1585 1590
Ser Ile His Gly Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser
1595 1600 1605
Gln Ile Leu Thr Arg Leu Cys Lys Glu Gly Lys Val Asp Gly Pro
1610 1615 1620
His Phe Gly Pro His Gly Arg Val Arg Val Ala Asn Arg Val Phe
1625 1630 1635
Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn Gly Gln Arg Lys Pro
1640 1645 1650
Thr Asp Glu His Val Ala Leu Ser Ala Leu Arg His Trp Glu Asp
1655 1660 1665
Ile Pro Arg Val Gly Cys Arg Leu Val Pro Glu His Val Glu Thr
1670 1675 1680
Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly Ile Glu Gln Gly Arg
1685 1690 1695
Leu Glu Leu Trp Val Asp Met Phe Pro Met Asp Met Pro Ala Pro
1700 1705 1710
Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys Pro Lys Lys Tyr Glu
1715 1720 1725
Leu Arg Val Ile Val Trp Asn Thr Asp Glu Val Val Leu Glu Asp
1730 1735 1740
Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp Ile Phe Val Arg
1745 1750 1755
Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp Thr Asp Val
1760 1765 1770
His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn Trp Arg Tyr
1775 1780 1785
Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys Ile Val Met
1790 1795 1800
Ser Lys Lys Glu Ser Met Phe Ser Trp Asp Glu Thr Glu Tyr Lys
1805 1810 1815
Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala Asp His Phe
1820 1825 1830
Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp Leu Asn Arg

1835 1840 1845
Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr Met Glu Met
1850 1855 1860
Ala Thr Gly Glu Val Asp Val Pro Leu Val Ser Ile Phe Lys Gln
1865 1870 1875
Lys Arg Val Lys Gly Trp Trp Pro Leu Leu Ala Arg Asn Glu Asn
1880 1885 1890
Asp Glu Phe Glu Leu Thr Gly Lys Val Glu Ala Glu Leu His Leu
1895 1900 1905
Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro Val Gly Leu Ala Arg
1910 1915 1920
Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro Asp Thr Ala
1925 1930 1935
Phe Val Trp Phe Leu Asn Pro Leu Lys Ser Ile Lys Tyr Leu Ile
1940 1945 1950
Cys Thr Arg Tyr Lys Trp Leu Ile Ile Lys Ile Val Leu Ala Leu
1955 1960 1965
Leu Gly Leu Leu Met Leu Ala Leu Phe Leu Tyr Ser Leu Pro Gly
1970 1975 1980
Tyr Met Val Lys Lys Leu Leu Gly Ala
1985 1990

<210> 9
<211> 1977
<212> PRT
<213> 小家鼠(Mus musculus)
<400> 9

Met Ala Leu Ile Val His Leu Lys Thr Val Ser Glu Leu Arg Gly Lys
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Gly Val Ala Asp Phe Asp Glu Thr Phe
35 40 45
Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Val Leu Glu Ile
50 55 60
Gln Ile Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80
Phe Cys Met Val Leu Gln Lys Val Val Glu Glu Asn Arg Val Glu Val
85 90 95
Thr Asp Thr Leu Met Asp Asp Ser Asn Ala Ile Ile Lys Thr Ser Leu

100 105 110
Ser Met Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Pro Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Lys Asp
130 135 140
Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Thr Arg
145 150 155 160
Ile Ser Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe Ser
165 170 175
Ala Met Lys Leu Gly Lys Thr Arg Ser His Lys Glu Glu Pro Gln Arg
180 185 190
Gln Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu Ala
195 200 205
Ile Gln Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala Ser
210 215 220
Val Thr Ala Leu Thr Ser Asn Val Ser Asn Lys Arg Ser Lys Pro Asp
225 230 235 240
Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val Ser
245 250 255
Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp Pro
260 265 270
Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met Lys
275 280 285
Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp Phe
290 295 300
His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser Val
305 310 315 320
Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser Phe
325 330 335
Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe His
340 345 350
His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ala Gly Leu
355 360 365
Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp Asn
370 375 380
Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile Glu
385 390 395 400
Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp Ala
405 410 415

Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met Asn
420 425 430
Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn Lys
435 440 445
Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys Gly
450 455 460
Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu Gln
465 470 475 480
Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys Val
485 490 495
Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr His
500 505 510
Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe Leu
515 520 525
Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg Asn
530 535 540
Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly Glu
545 550 555 560
Gly Val Ser Phe Arg Ala Arg Leu Met Leu Gly Leu Ala Val Glu Ile
565 570 575
Leu Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln Val
580 585 590
Glu Gln Ala Thr Pro Val Ser Glu Ser Cys Thr Gly Arg Met Glu Glu
595 600 605
Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg Lys
610 615 620
Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr Gly
625 630 635 640
Asn Glu Val Asp Gly Met Ser Arg Pro Leu Arg Pro Arg Pro Arg Lys
645 650 655
Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ser Ser Asp
660 665 670
Asp Glu Gly Asp Glu Ala Gly Asp Leu Ala Ser Val Ser Ser Thr Pro
675 680 685
Pro Met Arg Pro Gln Ile Thr Asp Arg Asn Tyr Phe His Leu Pro Tyr
690 695 700
Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp Gln
705 710 715 720
Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp Lys

725 730 735
Leu Glu Glu Gly Leu Asn Asp Val Gln Glu Met Ile Lys Thr Glu Lys
740 745 750
Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser Cys
755 760 765
Gly Cys His Arg Phe Leu Ser Leu Ser Asp Lys Asp Gln Gly Arg Ser
770 775 780
Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg Glu
785 790 795 800
Leu Glu Ser Met Gly Gln Gln Ala Lys Ser Leu Arg Ala Gln Val Lys
805 810 815
Arg His Thr Val Arg Asp Lys Leu Arg Ser Cys Gln Asn Phe Leu Gln
820 825 830
Lys Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Val
835 840 845
Phe Ile Trp Met Met Ser Asn Asn Lys Arg Ile Ala Tyr Ala Arg Val
850 855 860
Pro Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Leu Gly Lys
865 870 875 880
Asp Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg
885 890 895
Gly Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Leu Glu Leu Tyr
900 905 910
Leu Trp Leu Gly Leu Ser Lys Gln Arg Lys Asp Phe Leu Cys Gly Leu
915 920 925
Pro Cys Gly Phe Glu Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His
930 935 940
Ser Phe Pro Pro Ile Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln
945 950 955 960
Leu Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser
965 970 975
Ser Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser
980 985 990
Gln Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln
995 1000 1005
Met Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu
1010 1015 1020
Leu Arg Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln
1025 1030 1035

Asp Ser Met Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys
1040 1045 1050
Pro Leu Val Lys Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe
1055 1060 1065
Pro Pro Gln Leu Glu Tyr Tyr Gln Ile Tyr Arg Gly Ser Ala Thr
1070 1075 1080
Ala Gly Asp Leu Leu Ala Ala Phe Glu Leu Leu Gln Ile Gly Pro
1085 1090 1095
Ser Gly Lys Ala Asp Leu Pro Pro Ile Asn Gly Pro Val Asp Met
1100 1105 1110
Asp Arg Gly Pro Ile Met Pro Val Pro Val Gly Ile Arg Pro Val
1115 1120 1125
Leu Ser Lys Tyr Arg Val Glu Val Leu Phe Trp Gly Leu Arg Asp
1130 1135 1140
Leu Lys Arg Val Asn Leu Ala Gln Val Asp Arg Pro Arg Val Asp
1145 1150 1155
Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser Leu Ile His Asn
1160 1165 1170
Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys Trp Phe Glu
1175 1180 1185
Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu Asn Ile
1190 1195 1200
Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val Gly
1205 1210 1215
Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro
1220 1225 1230
Asp Arg Ser Ala Pro Asn Trp Asn Thr Thr Gly Glu Val Val Val
1235 1240 1245
Ser Met Glu Pro Glu Glu Pro Val Lys Lys Leu Glu Thr Met Val
1250 1255 1260
Lys Leu Asp Ala Thr Ser Asp Ala Val Val Lys Val Asp Val Ala
1265 1270 1275
Glu Asp Glu Lys Glu Arg Lys Lys Lys Lys Lys Lys Gly Pro Ser
1280 1285 1290
Glu Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp Trp
1295 1300 1305
Trp Ser Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu
1310 1315 1320
Arg Gln His Glu Thr Ser Gly Thr Asp Leu Glu Glu Lys Glu Glu

1325 1330 1335

Met Glu Ser Ala Glu Gly Leu Lys Gly Pro Met Lys Ser Lys Glu

1340 1345 1350

Lys Ser Arg Ala Ala Lys Glu Glu Lys Lys Lys Lys Asn Gln Ser

1355 1360 1365

Pro Gly Pro Gly Gln Gly Ser Glu Ala Pro Glu Lys Lys Lys Ala

1370 1375 1380

Lys Ile Asp Glu Leu Lys Val Tyr Pro Lys Glu Leu Glu Ser Glu

1385 1390 1395

Phe Asp Ser Phe Glu Asp Trp Leu His Thr Phe Asn Leu Leu Arg

1400 1405 1410

Gly Lys Thr Gly Asp Asp Glu Asp Gly Ser Thr Glu Glu Glu Arg

1415 1420 1425

Ile Val Gly Arg Phe Lys Gly Ser Leu Cys Val Tyr Lys Val Pro

1430 1435 1440

Leu Pro Glu Asp Val Ser Arg Glu Ala Gly Tyr Asp Pro Thr Tyr

1445 1450 1455

Gly Met Phe Gln Gly Ile Pro Ser Asn Asp Pro Ile Asn Val Leu

1460 1465 1470

Val Arg Ile Tyr Val Val Arg Ala Thr Asp Leu His Pro Ala Asp

1475 1480 1485

Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile Lys Leu Gly Lys

1490 1495 1500

Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys Gln Leu Asn

1505 1510 1515

Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe Pro Met

1520 1525 1530

Glu Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val Gly

1535 1540 1545

Thr Asp Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg

1550 1555 1560

Phe Tyr Ser Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr

1565 1570 1575

Ser Ile His Gly Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser

1580 1585 1590

Gln Ile Leu Thr Arg Leu Cys Lys Glu Gly Lys Val Asp Gly Pro

1595 1600 1605

His Phe Gly Pro His Gly Arg Val Arg Val Ala Asn Arg Val Phe

1610 1615 1620

Thr Gly Pro Ser Glu Ile Glu Asp Glu Asn Gly Gln Arg Lys Pro
1625 1630 1635
Thr Asp Glu His Val Ala Leu Ser Ala Leu Arg His Trp Glu Asp
1640 1645 1650
Ile Pro Arg Val Gly Cys Arg Leu Val Pro Glu His Val Glu Thr
1655 1660 1665
Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly Ile Glu Gln Gly Arg
1670 1675 1680
Leu Glu Leu Trp Val Asp Met Phe Pro Met Asp Met Pro Ala Pro
1685 1690 1695
Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys Pro Lys Lys Tyr Glu
1700 1705 1710
Leu Arg Val Ile Val Trp Asn Thr Asp Glu Val Val Leu Glu Asp
1715 1720 1725
Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp Ile Phe Val Arg
1730 1735 1740
Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp Thr Asp Val
1745 1750 1755
His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn Trp Arg Tyr
1760 1765 1770
Leu Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys Ile Val Met
1775 1780 1785
Ser Lys Lys Glu Ser Met Phe Ser Trp Asp Glu Thr Glu Tyr Lys
1790 1795 1800
Ile Pro Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala Asp His Phe
1805 1810 1815
Ser Ala Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp Leu Asn Arg
1820 1825 1830
Phe Pro Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr Met Glu Met
1835 1840 1845
Ala Thr Gly Glu Val Asp Val Pro Leu Val Ser Ile Phe Lys Gln
1850 1855 1860
Lys Arg Val Lys Gly Trp Trp Pro Leu Leu Ala Arg Asn Glu Asn
1865 1870 1875
Asp Glu Phe Glu Leu Thr Gly Lys Val Glu Ala Glu Leu His Leu
1880 1885 1890
Leu Thr Ala Glu Glu Ala Glu Lys Asn Pro Val Gly Leu Ala Arg
1895 1900 1905
Asn Glu Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro Asp Thr Ala

1910 1915 1920

Phe Val Trp Phe Leu Asn Pro Leu Lys Ser Ile Lys Tyr Leu Ile

1925 1930 1935

Cys Thr Arg Tyr Lys Trp Leu Ile Ile Lys Ile Val Leu Ala Leu

1940 1945 1950

Leu Gly Leu Leu Met Leu Ala Leu Phe Leu Tyr Ser Leu Pro Gly

1955 1960 1965

Tyr Met Val Lys Lys Leu Leu Gly Ala

1970 1975

<210> 10

<211> 5979

<212> DNA

<213> 智人(Homo sapiens)

<400> 10

atggccctga ttgttcacct caagactgtc tcagagctcc gaggcaaagg tgaccggatt 60

gccaaagtca ctttccgagg gcagtctttc tactcccggg tcctggagaa ctgcgagggt 120

gtggctgact ttgatgagac gttccggtgg ccagtggcca gcagcatcga ccggaatgaa 180

gtgttggaga ttcagatttt caactacagc aaagtcttca gcaacaagct gatagggacc 240

ttctgcatgg tgctgcagaa agtggtggag gagaatcggg tagaggtgac cgacacgctg 300

atggatgaca gcaatgctat catcaagacc agcctgagca tggaggtccg gtatcaggcc 360

acagatggca ctgtgggccc ctgggatgat ggagacttcc tgggagatga atccctccag 420

gaggagaagg acagccagga gacagatggg ctgctacctg gttcccgacc cagcacccgg 480

atatctggcg agaagagctt tcgcagcaaa ggcagagaga agaccaaggg aggcagagat 540

ggcgagcaca aagcgggaag gagtgtgttc tcggccatga aactcggcaa aactcggtcc 600

cacaaagagg agccccaaag acaagatgag ccagcagtgc tggagatgga ggacctggac 660

cacctagcca ttcagctggg ggatgggctg gatcctgact ccgtgtctct agcctcggtc 720

accgctctca ccagcaatgt ctccaacaaa cggtctaagc cagatattaa gatggagccc 780

agtgctggaa ggcccatgga ttaccaggtc agcatcacag tgattgaggc tcggcagctg 840

gtgggcttga acatggaccc tgtggtgtgt gtggaggtgg gtgatgacaa gaaatacacg 900

tcaatgaagg agtccacaaa ctgcccttac tacaacgagt actttgtctt cgacttccat 960

gtctctcctg atgtcatgtt tgacaagatc atcaagatct cggttatcca ttctaagaac 1020

ctgcttcgga gcggcaccct ggtgggttcc ttcaaaatgg atgtggggac tgtgtattcc 1080

cagcctgaac accagttcca tcacaaatgg gccatcctgt cagaccccga tgacatctct 1140

gctgggttga agggttatgt aaagtgtgat gtcgctgtgg tgggcaaggg agacaacatc 1200

aagacacccc acaaggccaa cgagacggat gaggacgaca ttgaagggaa cttgctgctc 1260

cccgagggcg tgcccccga acggcagtgg gcacggttct atgtgaaaat ttaccgagca 1320

gagggactgc cccggatgaa cacaagcctc atggccaacg tgaagaaggc gttcatcggt 1380

gagaacaagg acctcgtcga cccctatgtg caagtcttct ttgctggaca aaagggcaaa 1440

acatcagtgc agaagagcag ctatgagccg ctatggaatg agcaggtcgt cttcacagac 1500

ttgttccccc cactctgcaa acgcatgaag gtgcagatcc gggactctga caaggtcaat 1560
gatgtggcca tcggcaccca cttcatcgac ctgcgcaaga tttccaacga tggagacaaa 1620
ggcttcctgc ctaccctcgg tccagcctgg gtgaacatgt acggctccac gcgcaactac 1680
acactgctgg acgagcacca ggacttgaat gaaggcctgg gggagggtgt gtccttccgg 1740
gcccgcctca tgttgggact agctgtggag atcctggaca cctccaaccc agagctcacc 1800
agctccacgg aggtgcaggt ggagcaggcc acgcctgtct cggagagctg cacagggaga 1860
atggaagaat tttttctatt tggagccttc ttggaagcct caatgattga ccggaaaaat 1920
ggggacaagc caattacctt tgaggtgacc ataggaaact acggcaatga agtcgatggt 1980
atgtcccggc ccctgaggcc tcggccccgg aaagagcctg gggatgaaga agaggtagac 2040
ctgattcaga actccagtga cgatgaaggt gacgaagccg gggacctggc ctcggtgtcc 2100
tccaccccac ctatgcggcc ccagatcacg gacaggaact atttccacct gccctacctg 2160
gagcgcaagc cctgcatcta tatcaagagc tggtggcctg accagaggcg gcgcctctac 2220
aatgccaaca tcatggatca cattgctgac aagctggaag aaggcctgaa tgatgtacag 2280
gagatgatca aaacggagaa gtcctacccg gagcgccgcc tgcggggtgt gctagaggaa 2340
ctcagctgtg gctgccaccg cttcctctcc ctctcggaca aggaccaggg ccgctcgtcc 2400
cgcaccaggc tggatcgaga gcgtcttaag tcctgtatga gggagttgga gagcatggga 2460
cagcaggcca agagcctgag ggctcaggtg aagcggcaca ctgttcggga caagctgagg 2520
tcatgccaga actttctgca gaagctacgc ttcctggcgg atgagcccca gcacagcatt 2580
cctgatgtgt tcatttggat gatgagcaac aacaaacgta tcgcctatgc ccgcgtgcct 2640
tccaaagacc tgctcttctc catcgtggag gaggaactgg gcaaggactg cgccaaagtc 2700
aagaccctct tcctgaagct gccagggaag aggggcttcg gctcggcagg ctggacagta 2760
caggccaagc tggagctcta cctgtggctg ggcctcagca agcagcgaaa ggacttcctg 2820
tgtggtctgc cctgtggctt cgaggaggtc aaggcagccc aaggcctggg cctgcattcc 2880
tttccgccca tcagcctagt ctacaccaag aagcaagcct tccagctccg agcacacatg 2940
tatcaggccc gaagcctctt tgctgctgac agcagtgggc tctctgatcc ctttgcccgt 3000
gtcttcttca tcaaccagag ccaatgcact gaggttctaa acgagacact gtgtcccacc 3060
tgggaccaga tgctggtatt tgacaacctg gagctgtacg gtgaagctca cgagttacga 3120
gatgatcccc ccatcattgt cattgaaatc tacgaccagg acagcatggg caaagccgac 3180
ttcatgggcc ggaccttcgc caagcccctg gtgaagatgg cagatgaagc atactgccca 3240
cctcgcttcc cgccgcagct tgagtactac cagatctacc gaggcagtgc cactgccgga 3300
gacctactgg ctgccttcga gctgctgcag attgggccat cagggaaggc tgacctgcca 3360
cccatcaatg gcccagtgga catggacaga gggcccatca tgcctgtgcc cgtgggaatc 3420
cggccagtgc tcagcaagta ccgagtggag gtgctgttct ggggcctgag ggacctaaag 3480
agggtgaacc tggcccaggt ggaccgacca cgggtggaca tcgagtgtgc aggaaagggg 3540
gtacaatcct ccctgattca caattataag aagaacccca acttcaacac gctggtcaag 3600
tggtttgaag tggacctccc ggagaatgag ctcctgcacc cacccttgaa catccgagtg 3660
gtagattgcc gggcctttgg acgatacacc ctggtgggtt cccacgcagt cagctcactg 3720
aggcgcttca tctaccgacc tccagaccgc tcagccccca actggaacac cacaggggag 3780
gttgtagtaa gcatggagcc tgaggagcca gttaagaagc tggagaccat ggtgaaactg 3840

gatgcgactt ctgatgctgt ggtcaaggtg gatgtggctg aagatgagaa ggaaaggaag 3900
aagaagaaaa agaaaggccc gtcagaggag ccagaggagg aagagcccga tgagagcatg 3960
ctggattggt ggtccaagta cttcgcctcc atcgacacaa tgaaggagca acttcgacaa 4020
catgagacct ctggaactga cttggaagag aaggaagaga tggaaagcgc tgagggcctg 4080
aagggaccaa tgaagagcaa ggagaagtcc agagctgcaa aggaggagaa aaagaagaaa 4140
aaccagagcc ctggccctgg ccagggatcg gaggctcctg agaagaagaa agccaagatc 4200
gatgagctta aggtgtaccc caaggagctg gaatcggagt ttgacagctt tgaggactgg 4260
ctgcacacct tcaacctgtt gaggggcaag acgggagatg atgaggatgg ctccacagag 4320
gaggagcgca tagtaggccg attcaagggc tccctctgtg tgtacaaagt gccactccca 4380
gaagatgtat ctcgagaagc tggctatgat cccacctatg gaatgttcca gggcatccca 4440
agcaatgacc ccatcaatgt gctggtccga atctatgtgg tccgggccac agacctgcac 4500
ccggccgaca tcaatggcaa agctgacccc tatattgcca tcaagttagg caagaccgac 4560
atccgagaca aggagaacta catctccaag cagctcaacc ctgtgtttgg gaagtccttt 4620
gacattgagg cctccttccc catggagtcc atgttgacag tggccgtgta cgactgggat 4680
ctggtgggca ctgatgacct catcggagaa accaagattg acctggaaaa ccgcttctac 4740
agcaagcatc gcgccacctg cggcatcgca cagacctatt ccatacatgg ctacaatatc 4800
tggagggacc ccatgaagcc cagccagatc ctgacacgcc tctgtaaaga gggcaaagtg 4860
gacggccccc actttggtcc ccatgggaga gtgagggttg ccaaccgtgt cttcacgggg 4920
ccttcagaaa tagaggatga gaatggtcag aggaagccca cagatgagca cgtggcactg 4980
tctgctctga gacactggga ggacatcccc cgggtgggct gccgccttgt gccggaacac 5040
gtggagacca ggccgctgct caaccctgac aagccaggca ttgagcaggg ccgcctggag 5100
ctgtgggtgg acatgttccc catggacatg ccagcccctg ggacacctct ggatatatcc 5160
cccaggaaac ccaagaagta cgagctgcgg gtcatcgtgt ggaacacaga cgaggtggtc 5220
ctggaagacg atgatttctt cacgggagag aagtccagtg acatttttgt gaggggtgg 5280
ctgaagggcc agcaggagga caaacaggac acagatgtcc actatcactc cctcacgggg 5340
gagggcaact tcaactggag atacctcttc cccttcgact acctagcggc cgaagagaag 5400
atcgttatgt ccaaaaagga gtctatgttc tcctgggatg agacggagta caagatccct 5460
gcgcggctca ccctgcagat ctgggacgct gaccacttct cggctgacga cttcctgggg 5520
gctatcgagc tggacctgaa ccggttcccg aggggcgcta agacagccaa gcagtgcacc 5580
atggagatgg ccaccgggga ggtggacgta cccctggttt ccatctttaa acagaaacgt 5640
gtcaaaggct ggtggcccct cctggcccgc aatgagaatg atgagtttga gctcacaggc 5700
aaagtggagg cggagctaca cctactcacg gcagaggagg cagagaagaa ccctgtgggc 5760
ctggctcgca atgaacctga tcccctagaa aaacccaacc ggcctgacac ggcattcgtc 5820
tggttcctga acccactcaa atctatcaag tacctcatct gcacccggta caagtggctg 5880
atcatcaaga tcgtgctggc gctgctgggg ctgctcatgc tggccctctt cctttacagc 5940
ctcccaggct acatggtcaa gaagctccta ggggcctga 5979

<210> 11
<211> 4969
<212> DNA

<213> 智人(Homo sapiens)

<400> 11

ccgtgagttc tgcccaggcc ctgtgagctc accagagcca cagactcaca gcccagaggt 60
ggcttcttcc ttcaggaact gaagaacccc catgaacacc aacatctcca ggttctgaga 120
acagaacctg ggaaattgat gacttcctca tgatgaccga tactcaggat ggccctagcg 180
agagctccca gatcatgagg aagaaggcct gaacgacata caggagatga tcaaaacgga 240
gaagtcctac cctgagcgtc gcctgcgggg cgtcctggag gagctgagct gtggctgctg 300
ccgcttcctc tccctcgctg acaaggacca gggccactca tcccgcacca ggcttgaccg 360
ggagcgcctc aagtcctgca tgagggagct ggaaaacatg gggcagcagg ccaggatgct 420
gcgggcccag gtgaagcggc acacggtgcg ggacaagctg aggctgtgcc agaacttcct 480
gcagaagctg cgcttcctgg cggacgagcc ccagcacagc attcccgaca tcttcatctg 540
gatgatgagc aacaacaagc gtgtcgccta tgcccgtgtg ccctccaagg acctgctctt 600
ctccatcgtg gaggaggaga ctggcaagga ctgcgccaag gtcaagacgc tcttccttaa 660
gctgccaggg aagcggggct tcggctcggc aggctggaca gtgcaggcca aggtggagct 720
gtacctgtgg ctgggcctca gcaaacagcg caaggagttc ctgtgcggcc tgccctgtgg 780
cttccaggag gtcaaggcag cccagggcct gggcctgcat gccttcccac ccgtcagcct 840
ggtctacacc aagaagcagg cgttccagct ccgagcgcac atgtaccagg cccgcagcct 900
ctttgccgcc gacagcagcg gactctcaga ccccctttgcc cgcgtcttct tcatcaatca 960
gagtcagtgc acagaggtgc tgaatgagac cctgtgtccc acctgggacc agatgctggt 1020
gttcgacaac ctggagctct atggtgaagc tcatgagctg agggacgatc cgcccatcat 1080
tgtcattgaa atctatgacc aggattccat gggcaaagct gacttcatgg gccggacctt 1140
cgccaaaccc ctggtgaaga tggcagacga ggcgtactgc ccaccccgct tcccacctca 1200
gctcgagtac taccagatct accgtggcaa cgccacagct ggagacctgc tggcggcctt 1260
cgagctgctg cagattggac cagcagggaa ggctgacctg ccccccatca atggcccggt 1320
ggacgtggac cgaggtccca tcatgcccgt gcccatgggc atccggcccg tgctcagcaa 1380
gtaccgagtg gaggtgctgt tctggggcct acgggaccta aagcgggtga acctggccca 1440
ggtggaccgg ccacgggtgg acatcgagtg tgcagggaag ggggtgcagt cgtccctgat 1500
ccacaattat aagaagaacc ccaacttcaa caccctcgtc aagtggtttg aagtggacct 1560
cccagagaac gagctgctgc acccgccctt gaacatccgt gtggtggact gccgggcctt 1620
cggtcgctac acactggtgg gctcccatgc cgtcagctcc ctgcgacgct tcatctaccg 1680
gcccccagac cgctcggccc ccagctggaa caccacgggg gaggttgtgg tgactatgga 1740
gccagaggta cccatcaaga aactggagac catggtgaag ctggacgcga cttctgaagc 1800
tgttgtcaag gtggatgtgg ctgaggagga gaaggagaag aagaagaaga agaagggcac 1860
tgcggaggag ccagaggagg aggagccaga cgagagcatg ctggactggt ggtccaagta 1920
ctttgcctcc attgacacca tgaaggagca acttcgacaa caagagccct ctggaattga 1980
cttggaggag aaggaggaag tggacaatac cgagggcctg aaggggtcaa tgaagggcaa 2040
ggagaaggca agggctgcca aagaggagaa gaagaagaaa actcagagct ctggctctgg 2100
ccaggggtcc gaggcccccg agaagaagaa acccaagatt gatgagctta aggtataccc 2160
caaagagctg gagtccgagt ttgataactt tgaggactgg ctgcacactt tcaacttgct 2220

tcggggcaag accggggatg atgaggatgg ctccaccgag gaggagcgca ttgtgggacg 2280
cttcaagggc tccctctgcg tgtacaaagt gccactccca gaggacgtgt cccgggaagc 2340
cggctacgac tccacctacg gcatgttcca gggcatcccg agcaatgacc ccatcaatgt 2400
gctggtccga gtctatgtgg tccgggccac ggacctgcac cctgctgaca tcaacggcaa 2460
agctgacccc tacatcgcca tccggctagg caagactgac atccgcgaca aggagaacta 2520
catctccaag cagctcaacc ctgtctttgg gaagtccttt gacatcgagg cctccttccc 2580
catggaatcc atgctgacgg tggctgtgta tgactgggac ctggtgggca ctgatgacct 2640
cattggggaa accaagatcg acctggagaa ccgcttctac agcaagcacc gcgccacctg 2700
cggcatcgcc cagacctact ccacacatgg ctacaatatc tggcgggacc ccatgaagcc 2760
cagccagatc ctgacccgcc tctgcaaaga cggcaaagtg gacggccccc actttgggcc 2820
ccctgggaga gtgaaggtgg ccaaccgcgt cttcactggg ccctctgaga ttgaggacga 2880
gaacggtcag aggaagccca cagacgagca tgtggcgctg ttggccctga ggcactggga 2940
ggacatcccc cgcgcaggct gccgcctggt gccagagcat gtggagacga ggccgctgct 3000
caaccccgac aagccgggca tcgagcaggg ccgcctggag ctgtgggtgg acatgttccc 3060
catggacatg ccagcccctg ggacgcctct ggacatctca cctcggaagc ccaagaagta 3120
cgagctgcgg gtcatcatct ggaacacaga tgaggtggtc ttggaggacg acgacttctt 3180
cacaggggag aagtccagtg acatcttcgt gagggggtgg ctgaagggcc agcaggagga 3240
caagcaggac acagacgtcc actaccactc cctcactggc gagggcaact tcaactggcg 3300
ctacctgttc cccttcgact acctggcggc ggaggagaag atcgtcatct ccaagaagga 3360
gtccatgttc tcctgggacg agaccgagta caagatcccc gcgcggctca ccctgcagat 3420
ctgggatgcg gaccacttct ccgctgacga cttcctgggg gccatcgagc tggacctgaa 3480
ccggttcccg cggggcgcaa agacagccaa gcagtgcacc atggagatgg ccaccgggga 3540
ggtggacgtg cccctcgtgt ccatcttcaa gcaaaagcgc gtcaaaggct ggtggcccct 3600
cctggcccgc aatgagaacg atgagtttga gctcacgggc aaggtggagg ctgagctgca 3660
tttactgaca gcagaggagg cagagaagaa cccagtgggc ctggcccgca atgaacctga 3720
ccccctagag aaacccaacc ggcccgacac gagcttcatc tggttcctga accctctcaa 3780
gtcggctcgc tacttcttgt ggcacacgta tcgctggctg ctcctcaaac tgttgctgct 3840
cctgctgctg ctcctcctcc tcgccctgtt cctctactct gtgcctggct acctggtcaa 3900
gaaaatcctc ggggcctgag cccagtggcc tcctggccgg cccgacacgg ccttcgtctg 3960
gttcctcaac cctctcaagt ccatcaagta cctcatctgc acccggtaca agtggctcat 4020
catcaagatc gtgctggcgc tgttggggct gctcatgttg gggctcttcc tctacagcct 4080
ccctggctac atggtcaaaa agctccttgg ggcatgaagg ccgccagctc ccgccagccg 4140
ctccccagcc ctgccgcatt tcctttcagt ggcttggact ctttcccatc tcccctgggg 4200
agcctgagga gcccagcgtc cactcttcat gccttgggcc gagcctgcct cctgcttgcg 4260
ggggccgcct gtcctcactg ccccaggctg cggcttgccc agtcccgccc ctctgacccc 4320
tgcctgtggg ctggggagcc ttggatgggg tggggacctg gaatgggtct ctcttgcccc 4380
acctggctga ggcgccaccc ttcttcaggc ccaggctcca gaggaagact cctgaaaccc 4440
tccccaggtc ttccaagtac aggattgaag ctttagtgaa attaaccaag gaccatgggt 4500
cagtgcccag ggctttaaaa agaatgaacg agcaaaaggt atccccgccg tgacccctgc 4560

agatagcacc ggtctttgat ccgcagcagg ggccagaccc tgcccacaag tcccagcgcg 4620
gctgcttctg ccactgctgg gctccacttg gctcctctca cttcccaggg ggtcgcctgt 4680
cctgcctgtg ggtttccatg gcttcccaga gctccctctg ccccagccag cgcctccagg 4740
cccagctgag gagctgtgag aagcagcaga ggggactccc catcccgggc acaccctgtc 4800
ctcccacccc tgccccttg cccttccagc cctttcagct gcagctggga gctggcccgt 4860
caagtgctgc ccctgcctgt gtctgggttt ctgttggctg tttttctttt cttgagtggt 4920
gatttttctc taaataaaag aagtcaagca ctgaaaaaaa aaaaaaaaa 4969

<210> 12
<211> 4771
<212> DNA
<213> 智人(Homo sapiens)
<400> 12

ccgtgagttc tgcccaggcc ctgtgagctc accagagcca cagactcaca gcccagaggt 60
ggcttcttcc ttcaggaact gaagaacccc catgaacacc aacatctcca ggttctgaga 120
acagaacctg ggaaattgat gacttcctca tgatgaccga tactcaggat ggccctagcg 180
agagctccca gatcatgagg aagaaggcct gaacgacata caggagatga tcaaaacgga 240
gaagtcctac cctgagcgtc gcctgcgggg cgtcctggag gagctgagct gtggctgctg 300
ccgcttcctc tccctcgctg acaaggacca gggccactca tcccgcacca ggcttgaccg 360
ggagcgcctc aagtcctgca tgagggagct ggaaaacatg gggcagcagg ccaggatgct 420
gcgggcccag gtgaagcggc acacggtgcg ggacaagctg aggctgtgcc agaacttcct 480
gcagaagctg cgcttcctgg cggacgagcc ccagcacagc attcccgaca tcttcatctg 540
gatgatgagc aacaacaagc gtgtcgccta tgcccgtgtg ccctccaagg acctgctctt 600
ctccatcgtg gaggaggaga ctggcaagga ctgcgccaag gtcaagacgc tcttccttaa 660
gctgccaggg aagcggggct tcggctcggc aggctggaca gtgcaggcca aggtggagct 720
gtacctgtgg ctgggcctca gcaaacagcg caaggagttc ctgtgcggcc tgccctgtgg 780
cttccaggag gtcaaggcag cccagggcct gggcctgcat gccttcccac ccgtcagcct 840
ggtctacacc aagaagcagg cgttccagct ccgagcgcac atgtaccagg cccgcagcct 900
ctttgccgcc gacagcagcg gactctcaga ccccctttgcc cgcgtcttct tcatcaatca 960
gagtcagtgc acagaggtgc tgaatgagac cctgtgtccc acctgggacc agatgctggt 1020
gttcgacaac ctggagctct atggtgaagc tcatgagctg agggacgatc cgcccatcat 1080
tgtcattgaa atctatgacc aggattccat gggcaaagct gacttcatgg gccggacctt 1140
cgccaaaccc ctggtgaaga tggcagacga ggcgtactgc ccaccccgct tcccacctca 1200
gctcgagtac taccagatct accgtggcaa cgccacagct ggagacctgc tggcggcctt 1260
cgagctgctg cagattggac cagcagggaa ggctgacctg ccccccatca atggcccggt 1320
ggacgtggac cgaggtccca tcatgcccgt gcccatgggc atccggcccg tgctcagcaa 1380
gtaccgagtg gaggtgctgt tctggggcct acgggaccta aagcgggtga acctggccca 1440
ggtggaccgg ccacgggtgg acatcgagtg tgcagggaag ggggtgcagt cgtccctgat 1500
ccacaattat aagaagaacc ccaacttcaa caccctcgtc aagtggtttg aagtggacct 1560
cccagagaac gagctgctgc acccgccctt gaacatccgt gtggtggact gccgggcctt 1620

cggtcgctac acactggtgg gctcccatgc cgtcagctcc ctgcgacgct tcatctaccg 1680
gcccccagac cgctcggccc ccagctggaa caccacgggg gaggttgtgg tgactatgga 1740
gccagaggta cccatcaaga aactggagac catggtgaag ctggacgcga cttctgaagc 1800
tgttgtcaag gtggatgtgg ctgaggagga gaaggagaag aagaagaaga agaagggcac 1860
tgcggaggag ccagaggagg aggagccaga cgagagcatg ctggactggt ggtccaagta 1920
ctttgcctcc attgacacca tgaaggagca acttcgacaa caagagccct ctggaattga 1980
cttggaggag aaggaggaag tggacaatac cgagggcctg aaggggtcaa tgaagggcaa 2040
ggagaaggca agggctgcca aagaggagaa gaagaagaaa actcagagct ctggctctgg 2100
ccaggggtcc gaggcccccg agaagaagaa acccaagatt gatgagctta aggtataccc 2160
caaagagctg gagtccgagt ttgataactt tgaggactgg ctgcacactt tcaacttgct 2220
tcggggcaag accggggatg atgaggatgg ctccaccgag gaggagcgca ttgtgggacg 2280
cttcaagggc tccctctgcg tgtacaaagt gccactccca gaggacgtgt cccgggaagc 2340
cggctacgac tccacctacg gcatgttcca gggcatcccg agcaatgacc ccatcaatgt 2400
gctggtccga gtctatgtgg tccgggccac ggacctgcac cctgctgaca tcaacggcaa 2460
agctgacccc tacatcgcca tccggctagg caagactgac atccgcgaca aggagaacta 2520
catctccaag cagctcaacc ctgtctttgg gaagtccttt gacatcgagg cctccttccc 2580
catggaatcc atgctgacgg tggctgtgta tgactgggac ctggtgggca ctgatgacct 2640
cattggggaa accaagatcg acctggagaa ccgcttctac agcaagcacc gcgccacctg 2700
cggcatcgcc cagacctact ccacacatgg ctacaatatc tggcgggacc ccatgaagcc 2760
cagccagatc ctgacccgcc tctgcaaaga cggcaaagtg gacggccccc actttgggcc 2820
ccctgggaga gtgaaggtgg ccaaccgcgt cttcactggg ccctctgaga ttgaggacga 2880
gaacggtcag aggaagccca cagacgagca tgtggcgctg ttggccctga ggcactggga 2940
ggacatcccc cgcgcaggct gccgcctggt gccagagcat gtggagacga ggccgctgct 3000
caaccccgac aagccgggca tcgagcaggg ccgcctggag ctgtgggtgg acatgttccc 3060
catggacatg ccagcccctg ggacgcctct ggacatctca cctcggaagc ccaagaagta 3120
cgagctgcgg gtcatcatct ggaacacaga tgaggtggtc ttggaggacg acgacttctt 3180
cacaggggag aagtccagtg acatcttcgt gaggggtgg ctgaagggcc agcaggagga 3240
caagcaggac acagacgtcc actaccactc cctcactggc gagggcaact tcaactggcg 3300
ctacctgttc cccttcgact acctggcggc ggaggagaag atcgtcatct ccaagaagga 3360
gtccatgttc tcctgggacg agaccgagta caagatcccc gcgcggctca ccctgcagat 3420
ctgggatgcg gaccacttct ccgctgacga cttcctgggg gccatcgagc tggacctgaa 3480
ccggttcccg cggggcgcaa agacagccaa gcagtgcacc atggagatgg ccaccgggga 3540
ggtggacgtg cccctcgtgt ccatcttcaa gcaaaagcgc gtcaaaggct ggtggcccct 3600
cctggcccgc aatgagaacg atgagtttga gctcacgggc aaggtggagg ctgagctgca 3660
tttactgaca gcagaggagg cagagaagaa cccagtgggc ctggcccgca atgaacctga 3720
ccccctagag aaacccaacc ggcccgacac ggccttcgtc tggttcctca accctctcaa 3780
gtccatcaag tacctcatct gcacccggta caagtggctc atcatcaaga tcgtgctggc 3840
gctgttgggg ctgctcatgt tggggctctt cctctacagc ctccctggct acatggtcaa 3900
aaagctcctt ggggcatgaa ggccgccagc tcccgccagc cgctccccag ccctgccgca 3960

tttcctttca gtggcttgga ctctttccca tctcccctgg ggagcctgag gagcccagcg 4020
tccactcttc atgccttggg ccgagcctgc ctcctgcttg cgggggccgc ctgtcctcac 4080
tgccccaggc tgcggcttgc ccagtcccgc ccctctgacc cctgcctgtg ggctggggag 4140
ccttggatgg ggtggggacc tggaatgggt ctctcttgcc ccacctggct gaggcgccac 4200
ccttcttcag gcccaggctc cagaggaaga ctcctgaaac cctccccagg tcttccaagt 4260
acaggattga agctttagtg aaattaacca aggaccatgg gtcagtgccc agggctttaa 4320
aaagaatgaa cgagcaaaag gtatccccgc cgtgacccct gcagatagca ccggtctttg 4380
atccgcagca ggggccagac cctgcccaca agtcccagcg cggctgcttc tgccactgct 4440
gggctccact tggctcctct cacttcccag ggggtcgcct gtcctgcctg tgggtttcca 4500
tggcttccca gagctccctc tgccccagcc agcgcctcca ggcccagctg aggagctgtg 4560
agaagcagca gaggggactc cccatcccgg gcacaccctg tcctcccacc cctgcccccc 4620
tgcccttcca gcccctttcag ctgcagctgg gagctggccc gtcaagtgct gcccctgcct 4680
gtgtctgggt ttctgttggc tgtttttctt ttcttgagtg gtgatttttc tctaaataaa 4740
agaagtcaag cactgaaaaa aaaaaaaaaa a 4771

<210> 13
<211> 5123
<212> DNA
<213> 智人(Homo sapiens)
<400> 13

ccgtgagttc tgcccaggcc ctgtgagctc accagagcca cagactcaca gcccagaggt 60
ggcttcttcc ttcaggaact gaagaacccc catgaacacc aacatctcca ggttctgaga 120
acagaacctg ggaaattgat gacttcctca tgatgaccga tactcaggat ggccctagcg 180
agagctccca gatcatgagg tccctcactc ccctgatcaa cagggaggag gcatttgggg 240
aggctgggga ggcggggctg tggcccagca tcacccacac tcctgattca caggaagaag 300
gcctgaacga catacaggag atgatcaaaa cggagaagtc ctaccctgag cgtcgcctgc 360
ggggcgtcct ggaggagctg agctgtggct gctgccgctt cctctccctc gctgacaagg 420
accagggcca ctcatcccgc accaggcttg accgggagcg cctcaagtcc tgcatgaggg 480
agctggaaaa catggggcag caggccagga tgctgcgggc ccaggtgaag cggcacacgg 540
tgcgggacaa gctgaggctg tgccagaact tcctgcagaa gctgcgcttc ctggcggacg 600
agccccagca cagcattccc gacatcttca tctggatgat gagcaacaac aagcgtgtcg 660
cctatgcccg tgtgccctcc aaggacctgc tcttctccat cgtggaggag gagactggca 720
aggactgcgc caaggtcaag acgctcttcc ttaagctgcc agggaagcgg ggcttcggct 780
cggcaggctg gacagtgcag gccaaggtgg agctgtacct gtggctgggc ctcagcaaac 840
agcgcaagga gttcctgtgc ggcctgccct gtggcttcca ggaggtcaag gcagcccagg 900
gcctgggcct gcatgccttc ccacccgtca gcctggtcta caccaagaag caggcgttcc 960
agctccgagc gcacatgtac caggcccgca gcctctttgc cgccgacagc agcggactct 1020
cagacccctt tgcccgcgtc ttcttcatca atcagagtca gtgcacagag gtgctgaatg 1080
agaccctgtg tcccacctgg gaccagatgc tggtgttcga caacctggag ctctatggtg 1140
aagctcatga gctgagggac gatccgccca tcattgtcat tgaaatctat gaccaggatt 1200

ccatgggcaa agctgacttc atgggccgga ccttcgccaa acccctggtg aagatggcag 1260
acgaggcgta ctgcccaccc cgcttcccac ctcagctcga gtactaccag atctaccgtg 1320
gcaacgccac agctggagac ctgctggcgg ccttcgagct gctgcagatt ggaccagcag 1380
ggaaggctga cctgcccccc atcaatggcc cggtggacgt ggaccgaggt cccatcatgc 1440
ccgtgcccat gggcatccgg cccgtgctca gcaagtaccg agtggaggtg ctgttctggg 1500
gcctacggga cctaaagcgg gtgaacctgg cccaggtgga ccggccacgg gtggacatcg 1560
agtgtgcagg gaaggggtg cagtcgtccc tgatccacaa ttataagaag aacccccaact 1620
tcaacaccct cgtcaagtgg tttgaagtgg acctcccaga gaacgagctg ctgcacccgc 1680
ccttgaacat ccgtgtggtg gactgccggg ccttcggtcg ctacacactg gtgggctccc 1740
atgccgtcag ctccctgcga cgcttcatct accggccccc agaccgctcg gcccccagct 1800
ggaacaccac ggtcaggctt ctccggcgct gccgtgtgct gtgcaatggg ggctcctcct 1860
ctcactccac aggggaggtt gtggtgacta tggagccaga ggtacccatc aagaaactgg 1920
agaccatggt gaagctggac gcgacttctg aagctgttgt caaggtggat gtggctgagg 1980
aggagaagga gaagaagaag aagaagaagg gcactgcgga ggagccagag gaggaggagc 2040
cagacgagag catgctggac tggtggtcca agtactttgc ctccattgac accatgaagg 2100
agcaacttcg acaacaagag ccctctggaa ttgacttgga ggagaaggag gaagtggaca 2160
ataccgaggg cctgaagggg tcaatgaagg gcaaggagaa ggcaagggct gccaaagagg 2220
agaagaagaa gaaaactcag agctctggct ctggccaggg gtccgaggcc cccgagaaga 2280
agaaacccaa gattgatgag cttaaggtat accccaaaga gctggagtcc gagtttgata 2340
actttgagga ctggctgcac actttcaact tgcttcgggg caagaccggg gatgatgagg 2400
atggctccac cgaggaggag cgcattgtgg gacgcttcaa gggctccctc tgcgtgtaca 2460
aagtgccact cccagaggac gtgtcccggg aagccggcta cgactccacc tacggcatgt 2520
tccagggcat cccgagcaat gaccccatca atgtgctggt ccgagtctat gtggtccggg 2580
ccacggacct gcaccctgct gacatcaacg gcaaagctga cccctacatc gccatccggc 2640
taggcaagac tgacatccgc gacaaggaga actacatctc caagcagctc aaccctgtct 2700
ttgggaagtc ctttgacatc gaggcctcct tccccatgga atccatgctg acggtggctg 2760
tgtatgactg ggacctggtg ggcactgatg acctcattgg ggaaaccaag atcgacctgg 2820
agaaccgctt ctacagcaag caccgcgcca cctgcggcat cgcccagacc tactccacac 2880
atggctacaa tatctggcgg gaccccatga agcccagcca gatcctgacc cgcctctgca 2940
aagacggcaa agtggacggc ccccactttg ggccccctgg gagagtgaag gtggccaacc 3000
gcgtcttcac tgggccctct gagattgagg acgagaacgg tcagaggaag cccacagacg 3060
agcatgtggc gctgttggcc ctgaggcact gggaggacat cccccgcgca ggctgccgcc 3120
tggtgccaga gcatgtggag acgaggccgc tgctcaaccc cgacaagccg ggcatcgagc 3180
agggccgcct ggagctgtgg gtggacatgt tccccatgga catgccagcc cctgggacgc 3240
ctctggacat ctcacctcgg aagcccaaga agtacgagct gcgggtcatc atctggaaca 3300
cagatgaggt ggtcttggag gacgacgact tcttcacagg ggagaagtcc agtgacatct 3360
tcgtgagggg gtggctgaag ggccagcagg aggacaagca ggacacagac gtccactacc 3420
actccctcac tggcgagggc aacttcaact ggcgctacct gttccccttc gactacctgg 3480
cggcggagga gaagatcgtc atctccaaga aggagtccat gttctcctgg gacgagaccg 3540

agtacaagat ccccgcgcgg ctcaccctgc agatctggga tgcggaccac ttctccgctg 3600
acgacttcct gggggccatc gagctggacc tgaaccggtt cccgcggggc gcaaagacag 3660
ccaagcagtg caccatggag atggccaccg gggaggtgga cgtgcccctc gtgtccatct 3720
tcaagcaaaa gcgcgtcaaa ggctggtggc ccctcctggc ccgcaatgag aacgatgagt 3780
ttgagctcac gggcaaggtg gaggctgagc tgcatttact gacagcagag gaggcagaga 3840
agaacccagt gggcctggcc cgcaatgaac ctgaccccct agagaaaccc aaccggcccg 3900
acacgagctt catctggttc ctgaaccctc tcaagtcggc tcgctacttc ttgtggcaca 3960
cgtatcgctg gctgctcctc aaactgttgc tgctcctgct gctgctcctc ctcctcgccc 4020
tgttcctcta ctctgtgcct ggctacctgg tcaagaaaat cctcggggcc tgagcccagt 4080
ggcctcctgg ccggcccgac acggccttcg tctggttcct caaccctctc aagtccatca 4140
agtacctcat ctgcacccgg tacaagtggc tcatcatcaa gatcgtgctg gcgctgttgg 4200
ggctgctcat gttggggctc ttcctctaca gcctccctgg ctacatggtc aaaaagctcc 4260
ttggggcatg aaggccgcca gctcccgcca gccgctcccc agccctgccg catttccttt 4320
cagtggcttg gactctttcc catctcccct ggggagcctg aggagcccag cgtccactct 4380
tcatgccttg ggccgagcct gcctcctgct tgcgggggcc gcctgtcctc actgccccag 4440
gctgcggctt gcccagtccc gcccctctga cccctgcctg tgggctgggg agccttggat 4500
ggggtgggga cctggaatgg gtctctcttg ccccacctgg ctgaggcgcc acccttcttc 4560
aggcccaggc tccagaggaa gactcctgaa accctcccca ggtcttccaa gtacaggatt 4620
gaagctttag tgaaattaac caaggaccat gggtcagtgc ccagggcttt aaaaagaatg 4680
aacgagcaaa aggtatcccc gccgtgaccc ctgcagatag caccggtctt tgatccgcag 4740
caggggccag accctgccca caagtcccag cgcggctgct tctgccactg ctgggctcca 4800
cttggctcct ctcacttccc agggggtcgc ctgtcctgcc tgtgggtttc catggcttcc 4860
cagagctccc tctgccccag ccagcgcctc caggcccagc tgaggagctg tgagaagcag 4920
cagaggggac tccccatccc gggcacaccc tgtcctccca cccctgcccc cttgcccttc 4980
cagccctttc agctgcagct gggagctggc ccgtcaagtg ctgcccctgc ctgtgtctgg 5040
gtttctgttg gctgtttttc ttttcttgag tggtgatttt tctctaaata aaagaagtca 5100
agcactgaaa aaaaaaaaaa aaa 5123

<210> 14
<211> 5994
<212> DNA
<213> 智人(Homo sapiens)
<400> 14

atggccttgc tcatccacct caagacagtc tcggagctgc ggggcagggg cgaccggatc 60
gccaaagtga ctttccgagg gcaatccttc tactctcggg tcctggagaa ctgtgaggat 120
gtggctgact ttgatgagac atttcggtgg ccggtggcca gcagcatcga cagaaatgag 180
atgctggaga ttcaggtttt caactacagc aaagtcttca gcaacaagct catcgggacc 240
ttccgcatgg tgctgcagaa ggtggtagag gagagccatg tggaggtgac tgacacgctg 300
attgatgaca acaatgctat catcaagacc agcctgtgcg tggaggtccg gtatcaggcc 360
actgacggca cagtgggctc ctgggacgat ggggacttcc tgggagatga gtctcttcaa 420

gaggaagaga aggacagcca agagacggat ggactgctcc caggctcccg gcccagctcc 480
cggcccccag gagagaagag cttccggaga gccgggagga gcgtgttctc cgccatgaag 540
ctcggcaaaa accggtctca caaggaggag ccccaaagac cagatgaacc ggcggtgctg 600
gagatggaag accttgacca tctggccatt cggctaggag atggactgga tcccgactcg 660
gtgtctctag cctcagtcac agctctcacc actaatgtct ccaacaagcg atctaagcca 720
gacattaaga tggagccaag tgctgggcgg cccatggatt accaggtcag catcacggtg 780
atcgaggccc ggcagctggt gggcttgaac atggaccctg tggtgtgcgt ggaggtgggt 840
gacgacaaga agtacacatc catgaaggag tccactaact gcccctatta caacgagtac 900
ttcgtcttcg acttccatgt ctctccggat gtcatgtttg acaagatcat caagatttcg 960
gtgattcact ccaagaacct gctgcgcagt ggcaccctgg tgggctcctt caaaatggac 1020
gtgggaaccg tgtactcgca gccagagcac cagttccatc acaagtgggc catcctgtct 1080
gaccccgatg acatctcctc ggggctgaag ggctacgtga agtgtgacgt tgccgtggtg 1140
ggcaaagggg acaacatcaa gacgccccac aaggccaatg agaccgacga agatgacatt 1200
gaggggaact tgctgctccc cgagggggtg ccccccgaac gccagtgggc ccggttctat 1260
gtgaaaattt accgagcaga ggggctgccc cgtatgaaca caagcctcat ggccaatgta 1320
aagaaggctt tcatcggtga aaacaaggac ctcgtggacc cctacgtgca agtcttcttt 1380
gctggccaga agggcaagac ttcagtgcag aagagcagct atgagcccct gtggaatgag 1440
caggtcgtct ttacagacct cttcccccca ctctgcaaac gcatgaaggt gcagatccga 1500
gactcggaca aggtcaacga cgtggccatc ggcacccact tcattgacct gcgcaagatt 1560
tctaatgacg gagacaaagg cttcctgccc acactgggcc cagcctgggt gaacatgtac 1620
ggctccacac gtaactacac gctgctggat gagcatcagg acctgaacga gggcctgggg 1680
gagggtgtgt ccttccgggc ccggctcctg ctgggcctgg ctgtggagat cgtagacacc 1740
tccaaccctg agctcaccag ctccacagag gtgcaggtgg agcaggccac gcccatctcg 1800
gagagctgtg caggtaaaat ggaagaattc tttctctttg gagccttcct ggaggcctca 1860
atgatcgacc ggagaaacgg agacaagccc atcacctttg aggtcaccat aggcaactat 1920
gggaacgaag ttgatggcct gtcccggccc cagcggcctc ggccccggaa ggagccgggg 1980
gatgaggaag aagtagacct gattcagaac gcaagtgatg acgaggccgg tgatgccggg 2040
gacctggcct cagtctcctc cactccacca atgcggcccc aggtcaccga caggaactac 2100
ttccatctgc cctacctgga gcgaaagccc tgcatctaca tcaagagctg gtggccggac 2160
cagcgccgcc gcctctacaa tgccaacatc atggaccaca ttgccgacaa gctggaagaa 2220
ggcctgaacg acatacagga gatgatcaaa acggagaagt cctaccctga gcgtcgcctg 2280
cggggcgtcc tggaggagct gagctgtggc tgctgccgct tcctctccct cgctgacaag 2340
gaccagggcc actcatcccg caccaggctt gaccgggagc gcctcaagtc ctgcatgagg 2400
gagctggaaa acatggggca gcaggccagg atgctgcggg cccaggtgaa gcggcacacg 2460
gtgcgggaca agctgaggct gtgccagaac ttcctgcaga agctgcgctt cctggcggac 2520
gagccccagc acagcattcc cgacatcttc atctggatga tgagcaacaa caagcgtgtc 2580
gcctatgccc gtgtgccctc caaggacctg ctcttctcca tcgtggagga ggagactggc 2640
aaggactgcg ccaaggtcaa gacgctcttc cttaagctgc cagggaagcg gggcttcggc 2700
tcggcaggct ggacagtgca ggccaaggtg gagctgtacc tgtggctggg cctcagcaaa 2760

cagcgcaagg agttcctgtg cggcctgccc tgtggcttcc aggaggtcaa ggcagcccag 2820
ggcctgggcc tgcatgcctt cccacccgtc agcctggtct acaccaagaa gcaggcgttc 2880
cagctccgag cgcacatgta ccaggcccgc agcctctttg ccgccgacag cagcggactc 2940
tcagacccct ttgcccgcgt cttcttcatc aatcagagtc agtgcacaga ggtgctgaat 3000
gagaccctgt gtcccacctg ggaccagatg ctggtgttcg acaacctgga gctctatggt 3060
gaagctcatg agctgaggga cgatccgccc atcattgtca ttgaaatcta tgaccaggat 3120
tccatgggca aagctgactt catgggccgg accttcgcca aacccctggt gaagatggca 3180
gacgaggcgt actgcccacc ccgcttccca cctcagctcg agtactacca gatctaccgt 3240
ggcaacgcca cagctggaga cctgctggcg gccttcgagc tgctgcagat tggaccagca 3300
gggaaggctg acctgccccc catcaatggc ccggtggacg tggaccgagg tcccatcatg 3360
cccgtgccca tgggcatccg gcccgtgctc agcaagtacc gagtggaggt gctgttctgg 3420
ggcctacggg acctaaagcg ggtgaacctg gcccaggtgg accggccacg ggtggacatc 3480
gagtgtgcag ggaaggggct gcagtcgtcc ctgatccaca attataagaa gaaccccaac 3540
ttcaacaccc tcgtcaagtg gtttgaagtg gacctcccag agaacgagct gctgcacccg 3600
cccttgaaca tccgtgtggt ggactgccgg gccttcggtc gctacacact ggtgggctcc 3660
catgccgtca gctccctgcg acgcttcatc taccggcccc cagaccgctc ggcccccagc 3720
tggaacacca cggtcaggct tctccggcgc tgccgtgtgc tgtgcaatgg gggctcctcc 3780
tctcactcca caggggaggt tgtggtgact atggagccag aggtacccat caagaaactg 3840
gagaccatgg tgaagctgga cgcgacttct gaagctgttg tcaaggtgga tgtggctgag 3900
gaggagaagg agaagaagaa gaagaagaag ggcactgcgg aggagccaga ggaggaggag 3960
ccagacgaga gcatgctgga ctggtggtcc aagtactttg cctccattga caccatgaag 4020
gagcaacttc gacaacaaga gccctctgga attgacttgg aggagaagga ggaagtggac 4080
aataccgagg gcctgaaggg gtcaatgaag ggcaaggaga aggcaagggc tgccaaagag 4140
gagaagaaga agaaaactca gagctctggc tctggccagg ggtccgaggc ccccgagaag 4200
aagaaaccca agattgatga gcttaaggta taccccaaag agctggagtc cgagtttgat 4260
aactttgagg actggctgca cactttcaac ttgcttcggg gcaagaccgg ggatgatgag 4320
gatggctcca ccgaggagga gcgcattgtg ggacgcttca agggctccct ctgcgtgtac 4380
aaagtgccac tcccagagga cgtgtcccgg gaagccggct acgactccac ctacggcatg 4440
ttccagggca tcccgagcaa tgaccccatc aatgtgctgg tccgagtcta tgtggtccgg 4500
gccacggacc tgcaccctgc tgacatcaac ggcaaagctg acccctacat cgccatccgg 4560
ctaggcaaga ctgacatccg cgacaaggag aactacatct ccaagcagct caaccctgtc 4620
tttgggaagt cctttgacat cgaggcctcc ttccccatgg aatccatgct gacggtggct 4680
gtgtatgact gggacctggt gggcactgat gacctcattg gggaaaccaa gatcgacctg 4740
gagaaccgct tctacagcaa gcaccgcgcc acctgcggca tcgcccagac ctactccaca 4800
catggctaca atatctggcg ggaccccatg aagcccagcc agatcctgac ccgcctctgc 4860
aaagacggca aagtggacgg cccccacttt gggccccctg ggagagtgaa ggtggccaac 4920
cgcgtcttca ctgggccctc tgagattgag gacgagaacg gtcagaggaa gcccacagac 4980
gagcatgtgg cgctgttggc cctgaggcac tgggaggaca tcccccgcgc aggctgccgc 5040
ctggtgccag agcatgtgga gacgaggccg ctgctcaacc ccgacaagcc gggcatcgag 5100

cagggccgcc tggagctgtg ggtggacatg ttccccatgg acatgccagc ccctgggacg 5160
cctctggaca tctcacctcg gaagcccaag aagtacgagc tgcgggtcat catctggaac 5220
acagatgagg tggtcttgga ggacgacgac ttcttcacag gggagaagtc cagtgacatc 5280
ttcgtgaggg ggtggctgaa gggccagcag gaggacaagc aggacacaga cgtccactac 5340
cactccctca ctggcgaggg caacttcaac tggcgctacc tgttcccctt cgactacctg 5400
gcggcggagg agaagatcgt catctccaag aaggagtcca tgttctcctg ggacgagacc 5460
gagtacaaga tccccgcgcg gctcaccctg cagatctggg atgcggacca cttctccgct 5520
gacgacttcc tgggggccat cgagctggac ctgaaccggt tcccgcgggg cgcaaagaca 5580
gccaagcagt gcaccatgga gatggccacc ggggaggtgg acgtgcccct cgtgtccatc 5640
ttcaagcaaa agcgcgtcaa aggctggtgg cccctcctgg cccgcaatga gaacgatgag 5700
tttgagctca cgggcaaggt ggaggctgag ctgcatttac tgacagcaga ggaggcagag 5760
aagaacccag tgggcctggc ccgcaatgaa cctgaccccc tagagaaacc caaccggccc 5820
gacacggcct tcgtctggtt cctcaaccct ctcaagtcca tcaagtacct catctgcacc 5880
cggtacaagt ggctcatcat caagatcgtg ctggcgctgt tggggctgct catgttgggg 5940
ctcttcctct acagcctccc tggctacatg gtcaaaaagc tccttggggc atga 5994

<210> 15
<211> 7125
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 15

ttggttgcct tggtctctgt gggcagcagc aggaggaggc ggcagcagcc agagaagagg 60
gaggcgtgtg agccacactc caccagcgag cttcttcccg ctgctctgga actgcccagg 120
ctctccccac cagcatggcc ctgattgttc acctcaagac tgtctcagag ctccgaggca 180
aaggtgaccg gattgccaaa gtcactttcc gagggcagtc tttctactcc cgggtcctgg 240
agaactgcga gggtgtggct gactttgatg agacgttccg gtggccagtg gccagcagca 300
tcgaccggaa tgaagtgttg gagattcaga ttttcaacta cagcaaagtc ttcagcaaca 360
agctgatagg gaccttctgc atggtgctgc agaaagtggt ggaggagaat cgggtagagg 420
tgaccgacac gctgatggat gacagcaatg ctatcatcaa gaccagcctg agcatggagg 480
tccggtatca ggccacagat ggcactgtgg gcccctggga tgatggagac ttcctgggag 540
atgaatccct ccaggaggag aaggacagcc aggagacaga tgggctgcta cctggttccc 600
gacccagcac ccggatatct ggcgagaaga gctttcgcag agcgggaagg agtgtgttct 660
cggccatgaa actcggcaaa actcggtccc acaaagagga gccccaaaga caagatgagc 720
cagcagtgct ggagatggag gacctggacc acctagccat tcagctgggg gatgggctgg 780
atcctgactc cgtgtctcta gcctcggtca ccgctctcac cagcaatgtc tccaacaaac 840
ggtctaagcc agatattaag atggagccca gtgctggaag gcccatggat taccaggtca 900
gcatcacagt gattgaggct cggcagctgg tgggcttgaa catggaccct gtggtgtgtg 960
tggaggtggg tgatgacaag aaatacacgt caatgaagga gtccacaaac tgcccttact 1020
acaacgagta ctttgtcttc gacttccatg tctctcctga tgtcatgttt gacaagatca 1080
tcaagatctc ggttatccat tctaagaacc tgcttcggag cggcaccctg gtgggttcct 1140

tcaaaatgga tgtggggact gtgtattccc agcctgaaca ccagttccat cacaaatggg 1200
ccatcctgtc agaccccgat gacatctctg ctgggttgaa gggttatgta aagtgtgatg 1260
tcgctgtggt gggcaaggga gacaacatca agacacccca caaggccaac gagacggatg 1320
aggacgacat tgaagggaac ttgctgctcc ccgagggcgt gccccccgaa cggcagtggg 1380
cacggttcta tgtgaaaatt taccgagcag agggactgcc ccggatgaac acaagcctca 1440
tggccaacgt gaagaaggcg ttcatcggtg agaacaagga cctcgtcgac ccctatgtgc 1500
aagtcttctt tgctggacaa aagggcaaaa catcagtgca gaagagcagc tatgagccgc 1560
tatggaatga gcaggtcgtc ttcacagact tgttcccccc actctgcaaa cgcatgaagg 1620
tgcagatccg ggactctgac aaggtcaatg atgtggccat cggcacccac ttcatcgacc 1680
tgcgcaagat ttccaacgat ggagacaaag gcttcctgcc taccctcggt ccagcctggg 1740
tgaacatgta cggctccacg cgcaactaca cactgctgga cgagcaccag gacttgaatg 1800
aaggcctggg ggagggtgtg tccttccggg cccgcctcat gttgggacta gctgtggaga 1860
tcctggacac ctccaaccca gagctcacca gctccacgga ggtgcaggtg gagcaggcca 1920
cgcctgtctc ggagagctgc acagggagaa tggaagaatt ttttctattt ggagccttct 1980
tggaagcctc aatgattgac cggaaaaatg gggacaagcc aattaccttt gaggtgacca 2040
taggaaacta cggcaatgaa gtcgatggta tgtcccggcc cctgaggcct cggccccgga 2100
aagagcctgg ggatgaagaa gaggtagacc tgattcagaa ctccagtgac gatgaaggtg 2160
acgaagccgg ggacctggcc tcggtgtcct ccaccccacc tatgcggccc cagatcacgg 2220
acaggaacta tttccacctg ccctacctgg agcgcaagcc ctgcatctat atcaagagct 2280
ggtggcctga ccagaggcgg cgcctctaca atgccaacat catggatcac attgctgaca 2340
agctggaaga aggcctgaat gatgtacagg agatgatcaa aacggagaag tcctacccgg 2400
agcgccgcct gcggggtgtg ctagaggaac tcagctgtgg ctgccaccgc ttcctctccc 2460
tctcggacaa ggaccagggc cgctcgtccc gcaccaggct ggatcgagag cgtcttaagt 2520
cctgtatgag ggagttggag agcatgggac agcaggccaa gagcctgagg gctcaggtga 2580
agcggcacac tgttcgggac aagctgaggt catgccagaa ctttctgcag aagctacgct 2640
tcctggcgga tgagccccag cacagcattc ctgatgtgtt catttggatg atgagcaaca 2700
acaaacgtat cgcctatgcc cgcgtgcctt ccaaagacct gctcttctcc atcgtggagg 2760
aggaactggg caaggactgc gccaaagtca agaccctctt cctgaagctg ccagggaaga 2820
ggggcttcgg ctcggcaggc tggacagtac aggccaagct ggagctctac ctgtggctgg 2880
gcctcagcaa gcagcgaaag gacttcctgt gtggtctgcc ctgtggcttc gaggaggtca 2940
aggcagccca aggcctgggc ctgcattcct ttccgcccat cagcctagtc tacaccaaga 3000
agcaagcctt ccagctccga gcacacatgt atcaggcccg aagcctcttt gctgctgaca 3060
gcagtgggct ctctgatccc tttgcccgtg tcttcttcat caaccagagc caatgcactg 3120
aggttctaaa cgagacactg tgtcccacct gggaccagat gctggtattt gacaacctgg 3180
agctgtacgg tgaagctcac gagttacgag atgatccccc catcattgtc attgaaatct 3240
acgaccagga cagcatgggc aaagccgact tcatgggccg gaccttcgcc aagcccctgg 3300
tgaagatggc agatgaagca tactgcccac ctcgcttccc gccgcagctt gagtactacc 3360
agatctaccg aggcagtgcc actgccggag acctactggc tgccttcgag ctgctgcaga 3420
ttgggccatc agggaaggct gacctgccac ccatcaatgg cccagtggac atggacagag 3480

ggcccatcat gcctgtgccc gtgggaatcc ggccagtgct cagcaagtac cgagtggagg 3540
tgctgttctg gggcctgagg gacctaaaga gggtgaacct ggcccaggtg gaccgaccac 3600
gggtggacat cgagtgtgca ggaaaggggg tacaatcctc cctgattcac aattataaga 3660
agaaccccaa cttcaacacg ctggtcaagt ggtttgaagt ggacctcccg gagaatgagc 3720
tcctgcaccc acccttgaac atccgagtgg tagattgccg ggcctttgga cgatacaccc 3780
tggtgggttc ccacgcagtc agctcactga ggcgcttcat ctaccgacct ccagaccgct 3840
cagcccccaa ctggaacacc acagtcaggc tgctccgggg ctgccacagg ctgcgcaatg 3900
ggggcccctc ttctcgcccc acaggggagg ttgtagtaag catggagcct gaggagccag 3960
ttaagaagct ggagaccatg gtgaaactgg atgcgacttc tgatgctgtg gtcaaggtgg 4020
atgtggctga agatgagaag gaaaggaaga agaagaaaaa gaaaggcccg tcagaggagc 4080
cagaggagga agagcccgat gagagcatgc tggattggtg gtccaagtac ttcgcctcca 4140
tcgacacaat gaaggagcaa cttcgacaac atgagacctc tggaactgac ttggaagaga 4200
aggaagagat ggaaagcgct gagggcctga agggaccaat gaagagcaag gagaagtcca 4260
gagctgcaaa ggaggagaaa aagaagaaaa accagagccc tggccctggc cagggatcgg 4320
aggctcctga gaagaagaaa gccaagatcg atgagcttaa ggtgtacccc aaggagctgg 4380
aatcggagtt tgacagcttt gaggactggc tgcacacctt caacctgttg aggggcaaga 4440
cgggagatga tgaggatggc tccacagagg aggagcgcat agtaggccga ttcaagggct 4500
ccctctgtgt gtacaaagtg ccactcccag aagatgtatc tcgagaagct ggctatgatc 4560
ccacctatgg aatgttccag ggcatcccaa gcaatgaccc catcaatgtg ctggtccgaa 4620
tctatgtggt ccgggccaca gacctgcacc cggccgacat caatggcaaa gctgacccct 4680
atattgccat caagttaggc aagaccgaca tccgagacaa ggagaactac atctccaagc 4740
agctcaaccc tgtgtttggg aagtcctttg acattgaggc ctccttcccc atggagtcca 4800
tgttgacagt ggccgtgtac gactgggatc tggtgggcac tgatgacctc atcggagaaa 4860
ccaagattga cctggaaaac cgcttctaca gcaagcatcg cgccacctgc ggcatcgcac 4920
agacctattc catacatggc tacaatatct ggagggaccc catgaagccc agccagatcc 4980
tgacacgcct ctgtaaagag ggcaaagtgg acggccccca ctttggtccc catgggagag 5040
tgagggttgc caaccgtgtc ttcacggggc cttcagaaat agaggatgag aatggtcaga 5100
ggaagcccac agatgagcac gtggcactgt ctgctctgag acactgggag gacatccccc 5160
gggtgggctg ccgccttgtg ccggaacacg tggagaccag gccgctgctc aaccctgaca 5220
agccaggcat tgagcagggc cgcctggagc tgtgggtgga catgttcccc atggacatgc 5280
cagcccctgg gacacctctg gatatatccc ccaggaaacc caagaagtac gagctgcggg 5340
tcatcgtgtg gaacacagac gaggtggtcc tggaagacga tgatttcttc acgggagaga 5400
agtccagtga cattttgtg agggggtggc tgaagggcca gcaggaggac aaacaggaca 5460
cagatgtcca ctatcactcc ctcacggggg agggcaactt caactggaga tacctcttcc 5520
ccttcgacta cctagcggcc gaagagaaga tcgttatgtc caaaaaggag tctatgttct 5580
cctgggatga gacggagtac aagatccctg cgcggctcac cctgcagatc tgggacgctg 5640
accacttctc ggctgacgac ttcctggggg ctatcgagct ggacctgaac cggttcccga 5700
ggggcgctaa gacagccaag cagtgcacca tggagatggc caccggggag gtggacgtac 5760
ccctggtttc catctttaaa cagaaacgtg tcaaaggctg gtggcccctc ctggcccgca 5820

atgagaatga tgagtttgag ctcacaggca aagtggaggc ggagctacac ctactcacgg 5880
cagaggaggc agagaagaac cctgtgggcc tggctcgcaa tgaacctgat cccctagaaa 5940
aacccaatcg gccggacaca agcttcatct ggttcttgaa ccctctcaag tctgcccgct 6000
acttcctgtg gcatacctac cgctggctac tcctcaaatt cctgctgctc ttcctcctgc 6060
tgctgctctt cgccctgttt ctctactctc tgcctggcta cctggccaag aagatccttg 6120
gggcctgagc cctgcagtcg cctaggcctg ccggcctgac acggcattcg tctggttcct 6180
gaacccactc aaatctatca agtacctcat ctgcacccgg tacaagtggc tgatcatcaa 6240
gatcgtgctg gcgctgctgg ggctgctcat gctggccctc ttcctttaca gcctcccagg 6300
ctacatggtc aagaagctcc taggggcctg aagtgtgccc caccccagcc cgctccagca 6360
tccctccagg ggctgctgcg tattttgcct tccctcacct ggactctctc ccaactccct 6420
gaggagccct cccacgcctg ccagccttga gcaagacacc tgcttgctgg acttcatccc 6480
caccccacac ccaaactgtt gcttgcctga tcttgtccca ggcctgcctg gggtttgggg 6540
cacagttggc ctccaaaacc agataccctc ttgtctaaag taccaggttc ctctgcccaa 6600
ccccaagagt ggtagtggcc caaccctccc tgtgctttcc aaatcttgtc ttaaggcacc 6660
agtgaaatta accaagaaac gcggagcaat gcccaaggct ctgatgagta ggaacacgtg 6720
gaaagcacca ggaatgccag cagaggcgag gcggcacacc tctctgcaga gcatccaggc 6780
cgagcggcgg gcagcggcca gctgcttctg cgcatgctct cctcttggct ctgcttcttt 6840
ctcacagtca cagtcacttc acagcttagc cttgggcttc ccatcacttc caggggtgcc 6900
tctgccttgg ccagtgtgtg tcagctagta cacaagctcc aagtgtgaat caggtgtact 6960
ggccgtcctg aagactgact gccctgtcct tcctgccgac agccacaccc gagtgtacac 7020
ttaaagcggt gcccttctgc ctctgtgggc ctgctggctg ctgttccttt cttgagtgtg 7080
attttttttt tctctccctc aataaaataa atcaaactct gagac 7125

<210> 16
<211> 7065
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 16

ttggttgcct tggtctctgt gggcagcagc aggaggaggc ggcagcagcc agagaagagg 60
gaggcgtgtg agccacactc caccagcgag cttcttcccg ctgctctgga actgcccagg 120
ctctccccac cagcatggcc ctgattgttc acctcaagac tgtctcagag ctccgaggca 180
aaggtgaccg gattgccaaa gtcactttcc gagggcagtc tttctactcc cgggtcctgg 240
agaactgcga gggtgtggct gactttgatg agacgttccg gtggccagtg gccagcagca 300
tcgaccggaa tgaagtgttg gagattcaga ttttcaacta cagcaaagtc ttcagcaaca 360
agctgatagg gaccttctgc atggtgctgc agaaagtggt ggaggagaat cgggtagagg 420
tgaccgacac gctgatggat gacagcaatg ctatcatcaa gaccagcctg agcatggagg 480
tccggtatca ggccacagat ggcactgtgg gcccctggga tgatggagac ttcctgggag 540
atgaatccct ccaggaggag aaggacagcc aggagacaga tgggctgcta cctggttccc 600
gacccagcac ccggatatct ggcgagaaga gctttcgcag agcgggaagg agtgtgttct 660
cggccatgaa actcggcaaa actcggtccc acaaagagga gccccaaaga caagatgagc 720

cagcagtgct ggagatggag gacctggacc acctagccat tcagctgggg gatgggctgg 780
atcctgactc cgtgtctcta gcctcggtca ccgctctcac cagcaatgtc tccaacaaac 840
ggtctaagcc agatattaag atggagccca gtgctggaag gcccatggat taccaggtca 900
gcatcacagt gattgaggct cggcagctgg tgggcttgaa catggaccct gtggtgtgtg 960
tggaggtggg tgatgacaag aaatacacgt caatgaagga gtccacaaac tgcccttact 1020
acaacgagta ctttgtcttc gacttccatg tctctcctga tgtcatgttt gacaagatca 1080
tcaagatctc ggttatccat tctaagaacc tgcttcggag cggcaccctg gtgggttcct 1140
tcaaaatgga tgtggggact gtgtattccc agcctgaaca ccagttccat cacaaatggg 1200
ccatcctgtc agaccccgat gacatctctg ctgggttgaa gggttatgta aagtgtgatg 1260
tcgctgtggt gggcaaggga gacaacatca agacacccca caaggccaac gagacggatg 1320
aggacgacat tgaagggaac ttgctgctcc ccgagggcgt gccccccgaa cggcagtggg 1380
cacggttcta tgtgaaaatt taccgagcag agggactgcc ccggatgaac acaagcctca 1440
tggccaacgt gaagaaggcg ttcatcggtg agaacaagga cctcgtcgac ccctatgtgc 1500
aagtcttctt tgctggacaa aagggcaaaa catcagtgca gaagagcagc tatgagccgc 1560
tatggaatga gcaggtcgtc ttcacagact tgttcccccc actctgcaaa cgcatgaagg 1620
tgcagatccg ggactctgac aaggtcaatg atgtggccat cggcacccac ttcatcgacc 1680
tgcgcaagat ttccaacgat ggagacaaag gcttcctgcc taccctcggt ccagcctggg 1740
tgaacatgta cggctccacg cgcaactaca cactgctgga cgagcaccag gacttgaatg 1800
aaggcctggg ggagggtgtg tccttccggg cccgcctcat gttgggacta gctgtggaga 1860
tcctggacac ctccaaccca gagctcacca gctccacgga ggtgcaggtg gagcaggcca 1920
cgcctgtctc ggagagctgc acagggagaa tggaagaatt ttttctattt ggagccttct 1980
tggaagcctc aatgattgac cggaaaaatg gggacaagcc aattaccttt gaggtgacca 2040
taggaaacta cggcaatgaa gtcgatggta tgtcccggcc cctgaggcct cggccccgga 2100
aagagcctgg ggatgaagaa gaggtagacc tgattcagaa ctccagtgac gatgaaggtg 2160
acgaagccgg ggacctggcc tcggtgtcct ccaccccacc tatgcggccc cagatcacgg 2220
acaggaacta tttccacctg ccctacctgg agcgcaagcc ctgcatctat atcaagagct 2280
ggtggcctga ccagaggcgg cgcctctaca atgccaacat catggatcac attgctgaca 2340
agctggaaga aggcctgaat gatgtacagg agatgatcaa aacggagaag tcctacccgg 2400
agcgccgcct gcggggtgtg ctagaggaac tcagctgtgg ctgccaccgc ttcctctccc 2460
tctcggacaa ggaccagggc cgctcgtccc gcaccaggct ggatcgagag cgtcttaagt 2520
cctgtatgag ggagttggag agcatgggac agcaggccaa gagcctgagg gctcaggtga 2580
agcggcacac tgttcgggac aagctgaggt catgccagaa ctttctgcag aagctacgct 2640
tcctggcgga tgagccccag cacagcattc ctgatgtgtt catttggatg atgagcaaca 2700
acaaacgtat cgcctatgcc cgcgtgcctt ccaaagacct gctcttctcc atcgtggagg 2760
aggaactggg caaggactgc gccaaagtca agaccctctt cctgaagctg ccagggaaga 2820
ggggcttcgg ctcggcaggc tggacagtac aggccaagct ggagctctac ctgtggctgg 2880
gcctcagcaa gcagcgaaag gacttcctgt gtggtctgcc ctgtggcttc gaggaggtca 2940
aggcagccca aggcctgggc ctgcattcct ttccgcccat cagcctagtc tacaccaaga 3000
agcaagcctt ccagctccga gcacacatgt atcaggcccg aagcctcttt gctgctgaca 3060

gcagtgggct ctctgatccc tttgcccgtg tcttcttcat caaccagagc caatgcactg 3120
aggttctaaa cgagacactg tgtcccacct gggaccagat gctggtattt gacaacctgg 3180
agctgtacgg tgaagctcac gagttacgag atgatccccc catcattgtc attgaaatct 3240
acgaccagga cagcatgggc aaagccgact tcatgggccg gaccttcgcc aagcccctgg 3300
tgaagatggc agatgaagca tactgcccac ctcgcttccc gccgcagctt gagtactacc 3360
agatctaccg aggcagtgcc actgccggag acctactggc tgccttcgag ctgctgcaga 3420
ttgggccatc agggaaggct gacctgccac ccatcaatgg cccagtggac atggacagag 3480
ggcccatcat gcctgtgccc gtgggaatcc ggccagtgct cagcaagtac cgagtggagg 3540
tgctgttctg gggcctgagg gacctaaaga gggtgaacct ggcccaggtg gaccgaccac 3600
gggtggacat cgagtgtgca ggaaaggggg tacaatcctc cctgattcac aattataaga 3660
agaaccccaa cttcaacacg ctggtcaagt ggtttgaagt ggacctcccg gagaatgagc 3720
tcctgcaccc acccttgaac atccgagtgg tagattgccg ggcctttgga cgatacaccc 3780
tggtgggttc ccacgcagtc agctcactga ggcgcttcat ctaccgacct ccagaccgct 3840
cagcccccaa ctggaacacc acaggggagg ttgtagtaag catggagcct gaggagccag 3900
ttaagaagct ggagaccatg gtgaaactgg atgcgacttc tgatgctgtg gtcaaggtgg 3960
atgtggctga agatgagaag gaaaggaaga agaagaaaaa gaaaggcccg tcagaggagc 4020
cagaggagga agagcccgat gagagcatgc tggattggtg gtccaagtac ttcgcctcca 4080
tcgacacaat gaaggagcaa cttcgacaac atgagacctc tggaactgac ttggaagaga 4140
aggaagagat ggaaagcgct gagggcctga agggaccaat gaagagcaag gagaagtcca 4200
gagctgcaaa ggaggagaaa aagaagaaaa accagagccc tggccctggc cagggatcgg 4260
aggctcctga gaagaagaaa gccaagatcg atgagcttaa ggtgtacccc aaggagctgg 4320
aatcggagtt tgacagcttt gaggactggc tgcacacctt caacctgttg aggggcaaga 4380
cgggagatga tgaggatggc tccacagagg aggagcgcat agtaggccga ttcaagggct 4440
ccctctgtgt gtacaaagtg ccactcccag aagatgtatc tcgagaagct ggctatgatc 4500
ccacctatgg aatgttccag ggcatcccaa gcaatgaccc catcaatgtg ctggtccgaa 4560
tctatgtggt ccgggccaca gacctgcacc cggccgacat caatggcaaa gctgacccct 4620
atattgccat caagttaggc aagaccgaca tccgagacaa ggagaactac atctccaagc 4680
agctcaaccc tgtgtttggg aagtcctttg acattgaggc ctccttcccc atggagtcca 4740
tgttgacagt ggccgtgtac gactgggatc tggtgggcac tgatgacctc atcggagaaa 4800
ccaagattga cctggaaaac cgcttctaca gcaagcatcg cgccacctgc ggcatcgcac 4860
agacctattc catacatggc tacaatatct ggagggaccc catgaagccc agccagatcc 4920
tgacacgcct ctgtaaagag ggcaaagtgg acggcccccа ctttggtccc catgggagag 4980
tgagggttgc caaccgtgtc ttcacggggc cttcagaaat agaggatgag aatggtcaga 5040
ggaagcccac agatgagcac gtggcactgt ctgctctgag acactgggag gacatccccc 5100
gggtgggctg ccgccttgtg ccggaacacg tggagaccag gccgctgctc aaccctgaca 5160
agccaggcat tgagcagggc cgcctggagc tgtgggtgga catgttcccc atggacatgc 5220
cagcccctgg gacacctctg gatatatccc ccaggaaacc caagaagtac gagctgcggg 5280
tcatcgtgtg gaacacagac gaggtggtcc tggaagacga tgatttcttc acgggagaga 5340
agtccagtga cattttgtg agggggtggc tgaagggcca gcaggaggac aaacaggaca 5400

cagatgtcca ctatcactcc ctcacggggg agggcaactt caactggaga tacctcttcc 5460
ccttcgacta cctagcggcc gaagagaaga tcgttatgtc caaaaaggag tctatgttct 5520
cctgggatga gacggagtac aagatccctg cgcggctcac cctgcagatc tgggacgctg 5580
accacttctc ggctgacgac ttcctggggg ctatcgagct ggacctgaac cggttcccga 5640
ggggcgctaa gacagccaag cagtgcacca tggagatggc caccggggag gtggacgtac 5700
ccctggtttc catctttaaa cagaaacgtg tcaaaggctg gtggcccctc ctggcccgca 5760
atgagaatga tgagtttgag ctcacaggca aagtggaggc ggagctacac ctactcacgg 5820
cagaggaggc agagaagaac cctgtgggcc tggctcgcaa tgaacctgat cccctagaaa 5880
aacccaatcg gccggacaca agcttcatct ggttcttgaa ccctctcaag tctgcccgct 5940
acttcctgtg gcatacctac cgctggctac tcctcaaatt cctgctgctc ttcctcctgc 6000
tgctgctctt cgccctgttt ctctactctc tgcctggcta cctggccaag aagatccttg 6060
gggcctgagc cctgcagtcg cctaggcctg ccggcctgac acggcattcg tctggttcct 6120
gaacccactc aaatctatca agtacctcat ctgcacccgg tacaagtggc tgatcatcaa 6180
gatcgtgctg gcgctgctgg ggctgctcat gctggccctc ttcctttaca gcctcccagg 6240
ctacatggtc aagaagctcc taggggcctg aagtgtgccc caccccagcc cgctccagca 6300
tccctccagg ggctgctgcg tattttgcct tccctcacct ggactctctc ccaactccct 6360
gaggagccct cccacgcctg ccagccttga gcaagacacc tgcttgctgg acttcatccc 6420
caccccacac ccaaactgtt gcttgcctga tcttgtccca ggcctgcctg gggtttgggg 6480
cacagttggc ctccaaaacc agataccctc ttgtctaaag taccaggttc ctctgcccaa 6540
ccccaagagt ggtagtggcc caaccctccc tgtgctttcc aaatcttgtc ttaaggcacc 6600
agtgaaatta accaagaaac gcggagcaat gcccaaggct ctgatgagta ggaacacgtg 6660
gaaagcacca ggaatgccag cagaggcgag gcggcacacc tctctgcaga gcatccaggc 6720
cgagcggcgg gcagcggcca gctgcttctg cgcatgctct cctcttggct ctgcttcttt 6780
ctcacagtca cagtcacttc acagcttagc cttgggcttc ccatcacttc caggggtgcc 6840
tctgccttgg ccagtgtgtg tcagctagta cacaagctcc aagtgtgaat caggtgtact 6900
ggccgtcctg aagactgact gccctgtcct tcctgccgac agccacaccc gagtgtacac 6960
ttaaagcggt gcccttctgc ctctgtgggc ctgctggctg ctgttccttt cttgagtgtg 7020
attttttttt tctctccctc aataaaataa atcaaactct gagac 7065

<210> 17
<211> 6907
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 17

ttggttgcct tggtctctgt gggcagcagc aggaggaggc ggcagcagcc agagaagagg 60
gaggcgtgtg agccacactc caccagcgag cttcttcccg ctgctctgga actgcccagg 120
ctctccccac cagcatggcc ctgattgttc acctcaagac tgtctcagag ctccgaggca 180
aaggtgaccg gattgccaaa gtcactttcc gagggcagtc tttctactcc cgggtcctgg 240
agaactgcga gggtgtggct gactttgatg agacgttccg gtggccagtg gccagcagca 300
tcgaccggaa tgaagtgttg gagattcaga ttttcaacta cagcaaagtc ttcagcaaca 360

agctgatagg gaccttctgc atggtgctgc agaaagtggt ggaggagaat cgggtagagg 420
tgaccgacac gctgatggat gacagcaatg ctatcatcaa gaccagcctg agcatggagg 480
tccggtatca ggccacagat ggcactgtgg gcccctggga tgatggagac ttcctgggag 540
atgaatccct ccaggaggag aaggacagcc aggagacaga tgggctgcta cctggttccc 600
gacccagcac ccggatatct ggcgagaaga gctttcgcag caaaggcaga gagaagacca 660
agggaggcag agatggcgag cacaaagcgg gaaggagtgt gttctcggcc atgaaactcg 720
gcaaaactcg gtcccacaaa gaggagcccc aaagacaaga tgagccagca gtgctggaga 780
tggaggacct ggaccaccta gccattcagc tgggggatgg gctggatcct gactccgtgt 840
ctctagcctc ggtcaccgct ctcaccagca atgtctccaa caaacggtct aagccagata 900
ttaagatgga gcccagtgct ggaaggccca tggattacca ggtcagcatc acagtgattg 960
aggctcggca gctggtgggc ttgaacatgg accctgtggt gtgtgtggag gtgggtgatg 1020
acaagaaata cacgtcaatg aaggagtcca caaactgccc ttactacaac gagtactttg 1080
tcttcgactt ccatgtctct cctgatgtca tgtttgacaa gatcatcaag atctcggtta 1140
tccattctaa gaacctgctt cggagcggca ccctggtggg ttccttcaaa atggatgtgg 1200
ggactgtgta ttcccagcct gaacaccagt tccatcacaa atgggccatc ctgtcagacc 1260
ccgatgacat ctctgctggg ttgaagggtt atgtaaagtg tgatgtcgct gtggtgggca 1320
agggagacaa catcaagaca ccccacaagg ccaacgagac ggatgaggac gacattgaag 1380
ggaacttgct gctccccgag ggcgtgcccc ccgaacggca gtgggcacgg ttctatgtga 1440
aaatttaccg agcagaggga ctgccccgga tgaacacaag cctcatggcc aacgtgaaga 1500
aggcgttcat cggtgagaac aaggacctcg tcgaccccta tgtgcaagtc ttctttgctg 1560
gacaaaaggg caaaacatca gtgcagaaga gcagctatga gccgctatgg aatgagcagg 1620
tcgtcttcac agacttgttc cccccactct gcaaacgcat gaaggtgcag atccgggact 1680
ctgacaaggt caatgatgtg gccatcggca cccacttcat cgacctgcgc aagatttcca 1740
acgatggaga caaaggcttc ctgcctaccc tcggtccagc ctgggtgaac atgtacggct 1800
ccacgcgcaa ctacacactg ctggacgagc accaggactt gaatgaaggc ctgggggagg 1860
gtgtgtcctt ccgggcccgc ctcatgttgg gactagctgt ggagatcctg gacacctcca 1920
acccagagct caccagctcc acggaggtgc aggtggagca ggccacgcct gtctcggaga 1980
gctgcacagg gagaatggaa gaatttttc tatttggagc cttcttggaa gcctcaatga 2040
ttgaccggaa aaatggggac aagccaatta cctttgaggt gaccatagga aactacggca 2100
atgaagtcga tggtatgtcc cggcccctga ggcctcggcc ccggaaagag cctggggatg 2160
aagaagaggt agacctgatt cagaactcca gtgacgatga aggtgacgaa gccggggacc 2220
tggcctcggt gtcctccacc ccacctatgc ggccccagat cacggacagg aactatttcc 2280
acctgcccta cctggagcgc aagccctgca tctatatcaa gagctggtgg cctgaccaga 2340
ggcggcgcct ctacaatgcc aacatcatgg atcacattgc tgacaagctg gaagaaggcc 2400
tgaatgatgt acaggagatg atcaaaacgg agaagtccta cccggagcgc cgcctgcggg 2460
gtgtgctaga ggaactcagc tgtggctgcc accgcttcct ctccctctcg gacaaggacc 2520
agggccgctc gtcccgcacc aggctggatc gagagcgtct taagtcctgt atgagggagt 2580
tggagagcat gggacagcag gccaagagcc tgagggctca ggtgaagcgg cacactgttc 2640
gggacaagct gaggtcatgc cagaactttc tgcagaagct acgcttcctg gcggatgagc 2700

cccagcacag cattcctgat gtgttcattt ggatgatgag caacaacaaa cgtatcgcct 2760
atgcccgcgt gccttccaaa gacctgctct tctccatcgt ggaggaggaa ctgggcaagg 2820
actgcgccaa agtcaagacc ctcttcctga agctgccagg gaagaggggc ttcggctcgg 2880
caggctggac agtacaggcc aagctggagc tctacctgtg gctgggcctc agcaagcagc 2940
gaaaggactt cctgtgtggt ctgccctgtg gcttcgagga ggtcaaggca gcccaaggcc 3000
tgggcctgca ttcctttccg cccatcagcc tagtctacac caagaagcaa gccttccagc 3060
tccgagcaca catgtatcag gcccgaagcc tctttgctgc tgacagcagt gggctctctg 3120
atcccttttgc ccgtgtcttc ttcatcaacc agagccaatg cactgaggtt ctaaacgaga 3180
cactgtgtcc cacctgggac cagatgctgg tatttgacaa cctggagctg tacggtgaag 3240
ctcacgagtt acgagatgat ccccccatca ttgtcattga aatctacgac caggacagca 3300
tgggcaaagc cgacttcatg ggccggacct tcgccaagcc cctggtgaag atggcagatg 3360
aagcatactg cccacctcgc ttcccgccgc agcttgagta ctaccagatc taccgaggca 3420
gtgccactgc cggagaccta ctggctgcct tcgagctgct gcagattggg ccatcaggga 3480
aggctgacct gccacccatc aatggcccag tggacatgga cagagggccc atcatgcctg 3540
tgcccgtggg aatccggcca gtgctcagca agtaccgagt ggaggtgctg ttctggggcc 3600
tgagggacct aaagagggtg aacctggccc aggtggaccg accacgggtg gacatcgagt 3660
gtgcaggaaa gggggtacaa tcctccctga ttcacaatta taagaagaac cccaacttca 3720
acacgctggt caagtggttt gaagtggacc tcccggagaa tgagctcctg cacccaccct 3780
tgaacatccg agtggtagat tgccgggcct ttggacgata caccctggtg ggttcccacg 3840
cagtcagctc actgaggcgc ttcatctacc gacctccaga ccgctcagcc cccaactgga 3900
acaccacagg ggaggttgta gtaagcatgg agcctgagga gccagttaag aagctggaga 3960
ccatggtgaa actggatgcg acttctgatg ctgtggtcaa ggtggatgtg gctgaagatg 4020
agaaggaaag gaagaagaag aaaaagaaag gcccgtcaga ggagccagag gaggaagagc 4080
ccgatgagag catgctggat tggtggtcca agtacttcgc ctccatcgac acaatgaagg 4140
agcaacttcg acaacatgag acctctggaa ctgacttgga agagaaggaa gagatggaaa 4200
gcgctgaggg cctgaaggga ccaatgaaga gcaaggagaa gtccagagct gcaaaggagg 4260
agaaaaagaa gaaaaaccag agccctggcc ctggccaggg atcggaggct cctgagaaga 4320
agaaagccaa gatcgatgag cttaaggtgt accccaagga gctggaatcg gagtttgaca 4380
gctttgagga ctggctgcac accttcaacc tgttgagggg caagacggga gatgatgagg 4440
atggctccac agaggaggag cgcatagtag gccgattcaa gggctccctc tgtgtgtaca 4500
aagtgccact cccagaagat gtatctcgag aagctggcta tgatcccacc tatggaatgt 4560
tccagggcat cccaagcaat gaccccatca atgtgctggt ccgaatctat gtggtccggg 4620
ccacagacct gcacccggcc gacatcaatg gcaaagctga cccctatatt gccatcaagt 4680
taggcaagac cgacatccga gacaaggaga actacatctc caagcagctc aaccctgtgt 4740
ttgggaagtc ctttgacatt gaggcctcct tccccatgga gtccatgttg acagtggccg 4800
tgtacgactg ggatctggtg ggcactgatg acctcatcgg agaaaccaag attgacctgg 4860
aaaaccgctt ctacagcaag catcgcgcca cctgcggcat cgcacagacc tattccatac 4920
atggctacaa tatctggagg gaccccatga agcccagcca gatcctgaca cgcctctgta 4980
aagagggcaa agtggacggc ccccactttg gtccccatgg gagagtgagg gttgccaacc 5040

gtgtcttcac ggggccttca gaaatagagg atgagaatgg tcagaggaag cccacagatg 5100
agcacgtggc actgtctgct ctgagacact gggaggacat cccccgggtg ggctgccgcc 5160
ttgtgccgga acacgtggag accaggccgc tgctcaaccc tgacaagcca ggcattgagc 5220
agggccgcct ggagctgtgg gtggacatgt tccccatgga catgccagcc cctgggacac 5280
ctctggatat atcccccagg aaacccaaga agtacgagct gcgggtcatc gtgtggaaca 5340
cagacgaggt ggtcctggaa gacgatgatt tcttcacggg agagaagtcc agtgacattt 5400
ttgtgagggg gtggctgaag ggccagcagg aggacaaaca ggacacagat gtccactatc 5460
actccctcac gggggagggc aacttcaact ggagatacct cttccccttc gactacctag 5520
cggccgaaga gaagatcgtt atgtccaaaa aggagtctat gttctcctgg gatgagacgg 5580
agtacaagat ccctgcgcgg ctcaccctgc agatctggga cgctgaccac ttctcggctg 5640
acgacttcct gggggctatc gagctggacc tgaaccggtt cccgaggggc gctaagacag 5700
ccaagcagtg caccatggag atggccaccg gggaggtgga cgtacccctg gtttccatct 5760
ttaaacagaa acgtgtcaaa ggctggtggc ccctcctggc ccgcaatgag aatgatgagt 5820
ttgagctcac aggcaaagtg gaggcggagc tacacctact cacggcagag gaggcagaga 5880
agaaccctgt gggcctggct cgcaatgaac ctgatcccct agaaaaaccc aaccggcctg 5940
acacggcatt cgtctggttc ctgaacccac tcaaatctat caagtacctc atctgcaccc 6000
ggtacaagtg gctgatcatc aagatcgtgc tggcgctgct ggggctgctc atgctggccc 6060
tcttccttta cagcctccca ggctacatgg tcaagaagct cctaggggcc tgaagtgtgc 6120
cccaccccag cccgctccag catccctcca ggggctgctg cgtattttgc cttccctcac 6180
ctggactctc tcccaactcc ctgaggagcc ctcccacgcc tgccagcctt gagcaagaca 6240
cctgcttgct ggacttcatc cccaccccac acccaaactg ttgcttgcct gatcttgtcc 6300
caggcctgcc tggggtttgg ggcacagttg gcctccaaaa ccagataccc tcttgtctaa 6360
agtaccaggt tcctctgccc aaccccaaga gtggtagtgg cccaaccctc cctgtgcttt 6420
ccaaatcttg tcttaaggca ccagtgaaat taaccaagaa acgcggagca atgcccaagg 6480
ctctgatgag taggaacacg tggaaagcac caggaatgcc agcagaggcg aggcggcaca 6540
cctctctgca gagcatccag gccgagcggc gggcagcggc cagctgcttc tgcgcatgct 6600
ctcctcttgg ctctgcttct ttctcacagt cacagtcact tcacagctta gccttgggct 6660
tcccatcact tccaggggtg cctctgcctt ggccagtgtg tgtcagctag tacacaagct 6720
ccaagtgtga atcaggtgta ctggccgtcc tgaagactga ctgccctgtc cttcctgccg 6780
acagccacac ccgagtgtac acttaaagcg gtgcccttct gcctctgtgg gcctgctggc 6840
tgctgttcct ttcttgagtg tgatttttt tttctctccc tcaataaaat aaatcaaact 6900
ctgagac 6907

<210> 18
<211> 6862
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 18

ttggttgcct tggtctctgt gggcagcagc aggaggaggc ggcagcagcc agagaagagg 60
gaggcgtgtg agccacactc caccagcgag cttcttcccg ctgctctgga actgcccagg 120

ctctccccac cagcatggcc ctgattgttc acctcaagac tgtctcagag ctccgaggca 180
aaggtgaccg gattgccaaa gtcactttcc gagggcagtc tttctactcc cgggtcctgg 240
agaactgcga gggtgtggct gactttgatg agacgttccg gtggccagtg gccagcagca 300
tcgaccggaa tgaagtgttg gagattcaga ttttcaacta cagcaaagtc ttcagcaaca 360
agctgatagg gaccttctgc atggtgctgc agaaagtggt ggaggagaat cgggtagagg 420
tgaccgacac gctgatggat gacagcaatg ctatcatcaa gaccagcctg agcatggagg 480
tccggtatca ggccacagat ggcactgtgg gcccctggga tgatggagac ttcctgggag 540
atgaatccct ccaggaggag aaggacagcc aggagacaga tgggctgcta cctggttccc 600
gacccagcac ccggatatct ggcgagaaga gctttcgcag agcgggaagg agtgtgttct 660
cggccatgaa actcggcaaa actcggtccc acaaagagga gccccaaaga caagatgagc 720
cagcagtgct ggagatggag gacctggacc acctagccat tcagctgggg gatgggctgg 780
atcctgactc cgtgtctcta gcctcggtca ccgctctcac cagcaatgtc tccaacaaac 840
ggtctaagcc agatattaag atggagccca gtgctggaag gcccatggat taccaggtca 900
gcatcacagt gattgaggct cggcagctgg tgggcttgaa catggaccct gtggtgtgtg 960
tggaggtggg tgatgacaag aaatacacgt caatgaagga gtccacaaac tgcccttact 1020
acaacgagta ctttgtcttc gacttccatg tctctcctga tgtcatgttt gacaagatca 1080
tcaagatctc ggttatccat tctaagaacc tgcttcggag cggcaccctg gtgggttcct 1140
tcaaaatgga tgtggggact gtgtattccc agcctgaaca ccagttccat cacaaatggg 1200
ccatcctgtc agaccccgat gacatctctg ctgggttgaa gggttatgta aagtgtgatg 1260
tcgctgtggt gggcaaggga gacaacatca agacacccca caaggccaac gagacggatg 1320
aggacgacat tgaagggaac ttgctgctcc ccgagggcgt gccccccgaa cggcagtggg 1380
cacggttcta tgtgaaaatt taccgagcag agggactgcc ccggatgaac acaagcctca 1440
tggccaacgt gaagaaggcg ttcatcggtg agaacaagga cctcgtcgac ccctatgtgc 1500
aagtcttctt tgctggacaa aagggcaaaa catcagtgca gaagagcagc tatgagccgc 1560
tatggaatga gcaggtcgtc ttcacagact tgttcccccc actctgcaaa cgcatgaagg 1620
tgcagatccg ggactctgac aaggtcaatg atgtggccat cggcacccac ttcatcgacc 1680
tgcgcaagat ttccaacgat ggagacaaag gcttcctgcc taccctcggt ccagcctggg 1740
tgaacatgta cggctccacg cgcaactaca cactgctgga cgagcaccag gacttgaatg 1800
aaggcctggg ggagggtgtg tccttccggg cccgcctcat gttgggacta gctgtggaga 1860
tcctggacac ctccaaccca gagctcacca gctccacgga ggtgcaggtg gagcaggcca 1920
cgcctgtctc ggagagctgc acagggagaa tggaagaatt ttttctattt ggagccttct 1980
tggaagcctc aatgattgac cggaaaaatg gggacaagcc aattaccttt gaggtgacca 2040
taggaaacta cggcaatgaa gtcgatggta tgtcccggcc cctgaggcct cggccccgga 2100
aagagcctgg ggatgaagaa gaggtagacc tgattcagaa ctccagtgac gatgaaggtg 2160
acgaagccgg ggacctggcc tcggtgtcct ccaccccacc tatgcggccc cagatcacgg 2220
acaggaacta tttccacctg ccctacctgg agcgcaagcc ctgcatctat atcaagagct 2280
ggtggcctga ccagaggcgg cgcctctaca atgccaacat catggatcac attgctgaca 2340
agctggaaga aggcctgaat gatgtacagg agatgatcaa aacggagaag tcctacccgg 2400
agcgccgcct gcggggtgtg ctagaggaac tcagctgtgg ctgccaccgc ttcctctccc 2460

tctcggacaa ggaccagggc cgctcgtccc gcaccaggct ggatcgagag cgtcttaagt 2520
cctgtatgag ggagttggag agcatgggac agcaggccaa gagcctgagg gctcaggtga 2580
agcggcacac tgttcgggac aagctgaggt catgccagaa ctttctgcag aagctacgct 2640
tcctggcgga tgagccccag cacagcattc ctgatgtgtt catttggatg atgagcaaca 2700
acaaacgtat cgcctatgcc cgcgtgcctt ccaaagacct gctcttctcc atcgtggagg 2760
aggaactggg caaggactgc gccaaagtca agaccctctt cctgaagctg ccagggaaga 2820
ggggcttcgg ctcggcaggc tggacagtac aggccaagct ggagctctac ctgtggctgg 2880
gcctcagcaa gcagcgaaag gacttcctgt gtggtctgcc ctgtggcttc gaggaggtca 2940
aggcagccca aggcctgggc ctgcattcct ttccgcccat cagcctagtc tacaccaaga 3000
agcaagcctt ccagctccga gcacacatgt atcaggcccg aagcctcttt gctgctgaca 3060
gcagtgggct ctctgatccc tttgcccgtg tcttcttcat caaccagagc caatgcactg 3120
aggttctaaa cgagacactg tgtcccacct gggaccagat gctggtattt gacaacctgg 3180
agctgtacgg tgaagctcac gagttacgag atgatccccc catcattgtc attgaaatct 3240
acgaccagga cagcatgggc aaagccgact tcatgggccg gaccttcgcc aagcccctgg 3300
tgaagatggc agatgaagca tactgcccac ctcgcttccc gccgcagctt gagtactacc 3360
agatctaccg aggcagtgcc actgccggag acctactggc tgccttcgag ctgctgcaga 3420
ttgggccatc agggaaggct gacctgccac ccatcaatgg cccagtggac atggacagag 3480
ggcccatcat gcctgtgccc gtgggaatcc ggccagtgct cagcaagtac cgagtggagg 3540
tgctgttctg gggcctgagg gacctaaaga gggtgaacct ggcccaggtg gaccgaccac 3600
gggtggacat cgagtgtgca ggaaaggggg tacaatcctc cctgattcac aattataaga 3660
agaaccccaa cttcaacacg ctggtcaagt ggtttgaagt ggacctcccg gagaatgagc 3720
tcctgcaccc acccttgaac atccgagtgg tagattgccg ggcctttgga cgatacaccc 3780
tggtgggttc ccacgcagtc agctcactga ggcgcttcat ctaccgacct ccagaccgct 3840
cagcccccaa ctggaacacc acaggggagg ttgtagtaag catggagcct gaggagccag 3900
ttaagaagct ggagaccatg gtgaaactgg atgcgacttc tgatgctgtg gtcaaggtgg 3960
atgtggctga agatgagaag gaaaggaaga agaagaaaaa gaaaggcccg tcagaggagc 4020
cagaggagga agagcccgat gagagcatgc tggattggtg gtccaagtac ttcgcctcca 4080
tcgacacaat gaaggagcaa cttcgacaac atgagacctc tggaactgac ttggaagaga 4140
aggaagagat ggaaagcgct gagggcctga agggaccaat gaagagcaag gagaagtcca 4200
gagctgcaaa ggaggagaaa aagaagaaaa accagagccc tggccctggc cagggatcgg 4260
aggctcctga gaagaagaaa gccaagatcg atgagcttaa ggtgtacccc aaggagctgg 4320
aatcggagtt tgacagcttt gaggactggc tgcacacctt caacctgttg aggggcaaga 4380
cgggagatga tgaggatggc tccacagagg aggagcgcat agtaggccga ttcaagggct 4440
ccctctgtgt gtacaaagtg ccactcccag aagatgtatc tcgagaagct ggctatgatc 4500
ccacctatgg aatgttccag ggcatcccaa gcaatgaccc catcaatgtg ctggtccgaa 4560
tctatgtggt ccgggccaca gacctgcacc cggccgacat caatggcaaa gctgacccct 4620
atattgccat caagttaggc aagaccgaca tccgagacaa ggagaactac atctccaagc 4680
agctcaaccc tgtgtttggg aagtcctttg acattgaggc ctccttcccc atggagtcca 4740
tgttgacagt ggccgtgtac gactgggatc tggtgggcac tgatgacctc atcggagaaa 4800

ccaagattga cctggaaaac cgcttctaca gcaagcatcg cgccacctgc ggcatcgcac 4860
agacctattc catacatggc tacaatatct ggagggaccc catgaagccc agccagatcc 4920
tgacacgcct ctgtaaagag ggcaaagtgg acggccccca ctttggtccc catgggagag 4980
tgagggttgc caaccgtgtc ttcacggggc cttcagaaat agaggatgag aatggtcaga 5040
ggaagcccac agatgagcac gtggcactgt ctgctctgag acactgggag gacatccccc 5100
gggtgggctg ccgccttgtg ccggaacacg tggagaccag gccgctgctc aaccctgaca 5160
agccaggcat tgagcagggc cgcctggagc tgtgggtgga catgttcccc atggacatgc 5220
cagcccctgg gacacctctg gatatatccc ccaggaaacc caagaagtac gagctgcggg 5280
tcatcgtgtg gaacacagac gaggtggtcc tggaagacga tgatttcttc acgggagaga 5340
agtccagtga cattttgtg agggggtggc tgaagggcca gcaggaggac aaacaggaca 5400
cagatgtcca ctatcactcc ctcacggggg agggcaactt caactggaga tacctcttcc 5460
ccttcgacta cctagcggcc gaagagaaga tcgttatgtc caaaaaggag tctatgttct 5520
cctgggatga gacggagtac aagatccctg cgcggctcac cctgcagatc tgggacgctg 5580
accacttctc ggctgacgac ttcctggggg ctatcgagct ggacctgaac cggttcccga 5640
ggggcgctaa gacagccaag cagtgcacca tggagatggc caccggggag gtggacgtac 5700
ccctggtttc catctttaaa cagaaacgtg tcaaaggctg gtggcccctc ctggcccgca 5760
atgagaatga tgagtttgag ctcacaggca aagtggaggc ggagctacac ctactcacgg 5820
cagaggaggc agagaagaac cctgtgggcc tggctcgcaa tgaacctgat cccctagaaa 5880
aacccaaccg gcctgacacg gcattcgtct ggttcctgaa cccactcaaa tctatcaagt 5940
acctcatctg cacccggtac aagtggctga tcatcaagat cgtgctggcg ctgctggggc 6000
tgctcatgct ggccctcttc ctttacagcc tcccaggcta catggtcaag aagctcctag 6060
gggcctgaag tgtgccccac cccagcccgc tccagcatcc ctccaggggc tgctgcgtat 6120
tttgccttcc ctcacctgga ctctctccca actccctgag gagccctccc acgcctgcca 6180
gccttgagca agacacctgc ttgctggact tcatccccac cccacaccca aactgttgct 6240
tgcctgatct tgtcccaggc ctgcctgggg tttggggcac agttggcctc caaaaccaga 6300
taccctcttg tctaaagtac caggttcctc tgcccaaccc caagagtggt agtggcccaa 6360
ccctccctgt gctttccaaa tcttgtctta aggcaccagt gaaattaacc aagaaacgcg 6420
gagcaatgcc caaggctctg atgagtagga acacgtggaa agcaccagga atgccagcag 6480
aggcgaggcg gcacacctct ctgcagagca tccaggccga gcggcgggca gcggccagct 6540
gcttctgcgc atgctctcct cttggctctg cttctttctc acagtcacag tcacttcaca 6600
gcttagcctt gggcttccca tcacttccag gggtgcctct gccttggcca gtgtgtgtca 6660
gctagtacac aagctccaag tgtgaatcag gtgtactggc cgtcctgaag actgactgcc 6720
ctgtccttcc tgccgacagc cacacccgag tgtacactta aagcggtgcc cttctgcctc 6780
tgtgggcctg ctggctgctg ttcctttctt gagtgtgatt ttttttttct ctccctcaat 6840
aaaataaatc aaactctgag ac 6862

<210> 19
<211> 77
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>
<223> 合成构建体
<400> 19
gggattttgc cgatttcggc ctattggtta aaaaatgagc tgatttaaca aaaatttaac 60
gcgaatttta acaaaat 77
<210> 20
<211> 82
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 20
gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg cttgtcgaga 60
cagagaagac tcttgcgttt ct 82
<210> 21
<211> 51
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 21
gataggcacc tattggtctt actgacatcc actttgcctt tctctccaca g 51
<210> 22
<211> 48
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 22
gcctgcaaga actggttcag cagcctgagc cacttcgtga tccacctg 48
<210> 23
<211> 548
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 23
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> 24
<211> 454
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 24

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240
aagcttctgc cactggctcc ggcattgcag agagaagaga aggggcggca gagctgaacc 300
ttagccttgc cttcctgggt acccttctga gcctcactgt cttctgtgag atgggcaaag 360
tgcgggtgtg actccttggc aacggtgtta caccagggca ggtaaagttg tagttatttg 420
tggggtacac caggactgtt aaaggtgtaa ctat 454

<210> 25
<211> 1157
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 25

ggtctcaccc agcattttca cttctaataa gttcaaatgt gatacggcac ctttctaaaa 60
attagtttc agggaaatag ggttcaaaac tggtagtggt agggtccatt ctcacgaccc 120
ccaggcctgc taaccctgac caagctacct attacttacc ctcctctttc tcctcctcct 180
ctttctcctt ctcctgcttc ccctcttcct tctccctccc ttcctctccc tcctcccccc 240
ccttggctgt gatcagatcc agagcctgaa tgagcctcct gaccccacac ccccactagc 300
atgggcctgc aagtgcccag aagtccctcc tgcctcctaa actgcccagc cgatccatta 360
gctcttcctt cttcccagtg aaagaagcag gcacagcctg tccctcccgt tctacagaaa 420
ggaagctaca gcacagggag ggccaaaggc cttcctggga ctagacagtt gatcaacagc 480
aggactggag agctgggctc catttttgtt ccttggtgcc ctgcccctcc ccatgacctg 540
cagagacatt cagcctgcca ggctttatga ggtgggagct gggctctccc tgatgtatta 600
ttcagctccc tggagttggc cagctcctgt tacactggcc acagccctgg gcatccgctt 660
ctcacttcta gtttcccctc caaggtaatg tggtgggtca tgatcattct atcctggctt 720
cagggacctg actccacttt ggggccattc gaggggtcta gggtagatga tgtcccctg 780

tggggattaa tgtcctgctc tgtaaaactg agctagctga gatccaggag ggcttggcca 840
gagacagcaa gttgttgcca tggtgacttt aaagccaggt tgctgcccca gcacaggcct 900
cccagtctac cctcactaga aaacaacacc caggcacttt ccaccacctc tcaaaggtga 960
aacccaaggc tggtctagag aatgaattat ggatcctcgc tgtccgtgcc acccagctag 1020
tcccagcggc tcagacactg aggagagact gtaggttcag ctacaagcaa aaagacctag 1080
ctggtctcca agcagtgtct ccaagtccct gaacctgtga cacctgcccc aggcatcatc 1140
aggcacagag ggccacc 1157

<210> 26
<211> 76
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 26
cccatgtcag ctgcttgtgc tttccagaga caaaacagga ataatagatg tcattaaata 60
tacattgggc cccagg 76

<210> 27
<211> 95
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 27
agcctgagcc tcctttccat ctctgtggag gcagacatag gacccccaac aaacagcatg 60
caggttggga gccagccaca ggacccaggt aaggg 95

<210> 28
<211> 171
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 28
cccatgtcag ctgcttgtgc tttccagaga caaaacagga ataatagatg tcattaaata 60
tacattgggc cccaggagcc tgagcctcct ttccatctct gtggaggcag acataggacc 120
cccaacaaac agcatgcagg ttgggagcca gccacaggac ccaggtaagg g 171

<210> 29
<211> 171
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 29
agcctgagcc tcctttccat ctctgtggag gcagacatag gacccccaac aaacagcatg 60

caggttggga gccagccaca ggacccaggt aagggcccat gtcagctgct tgtgctttcc 120
agagacaaaa caggaataat agatgtcatt aaatatacat tgggccccag g 171
<210> 30
<211> 184
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 30
cccatgtcag ctgcttgtgc tttccagaga caaaacagga ataatagatg tcattaaata 60
tacattgggc cccaggcggt caatgtggca gcctgagcct cctttccatc tctgtggagg 120
cagacatagg acccccaaca aacagcatgc aggttgggag ccagccacag gacccaggta 180
aggg 184
<210> 31
<211> 82
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 31
tgaggtggga gctgggctct ccctgatgta ttattcagct ccctggagtt ggccagctcc 60
tgttacactg gccacagccc tg 82
<210> 32
<211> 106
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 32
cacaggcctc ccagtctacc ctcactagaa aacaacaccc aggcactttc caccacctct 60
caaaggtgaa acccaaggct ggtctagaga atgaattatg gatcct 106
<210> 33
<211> 188
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 33
tgaggtggga gctgggctct ccctgatgta ttattcagct ccctggagtt ggccagctcc 60
tgttacactg gccacagccc tgcacaggcc tcccagtcta ccctcactag aaaacaacac 120
ccaggcactt tccaccacct ctcaaaggtg aaacccaagg ctggtctaga gaatgaatta 180
tggatcct 188
<210> 34
<211> 188
<212> DNA

<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 34
cacaggcctc ccagtctacc ctcactagaa aacaacaccc aggcactttc caccacctct 60
caaaggtgaa acccaaggct ggtctagaga atgaattatg gatccttgag gtgggagctg 120
ggctctccct gatgtattat tcagctccct ggagttggcc agctcctgtt acactggcca 180
cagccctg 188
<210> 35
<211> 430
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 35
tgaggtggga gctgggctct ccctgatgta ttattcagct ccctggagtt ggccagctcc 60
tgttacactg gccacagccc tgggcatccg cttctcactt ctagtttccc ctccaaggta 120
atgtggtggg tcatgatcat tctatcctgg cttcagggac ctgactccac tttggggcca 180
ttcgaggggt ctagggtaga tgatgtcccc ctgtggggat taatgtcctg ctctgtaaaa 240
ctgagctagc tgagatccag gagggcttgg ccagagacag caagttgttg ccatggtgac 300
tttaaagcca ggttgctgcc ccagcacagg cctcccagtc taccctcact agaaaacaac 360
acccaggcac tttccaccac ctctcaaagg tgaaacccaa ggctggtcta gagaatgaat 420
tatggatcct 430
<210> 36
<211> 1611
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 36
ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240
aagcttctgc cactggctcc ggcattgcag agagaagaga aggggcggca gagctgaacc 300
ttagccttgc cttcctgggt acccttctga gcctcactgt cttctgtgag atgggcaaag 360
tgcgggtgtg actccttggc aacggtgtta caccagggca ggtaaagttg tagttatttg 420
tggggtacac caggactgtt aaaggtgtaa ctatggtctc acccagcatt ttcacttcta 480
ataagttcaa atgtgatacg gcacctttct aaaaattagt tttcagggaa atagggttca 540
aaactggtag tggtagggtc cattctcacg acccccaggc ctgctaaccc tgaccaagct 600
acctattact taccctcctc tttctcctcc tcctctttct ccttctcctg cttcccctct 660

tccttctccc tcccttcctc tccctcctcc ccctccttgg ctgtgatcag atccagagcc 720
tgaatgagcc tcctgacccc acacccccac tagcatgggc ctgcaagtgc ccagaagtcc 780
ctcctgcctc ctaaactgcc cagccgatcc attagctctt ccttcttccc agtgaaagaa 840
gcaggcacag cctgtccctc ccgttctaca gaaaggaagc tacagcacag ggagggccaa 900
aggccttcct gggactagac agttgatcaa cagcaggact ggagagctgg gctccatttt 960
tgttccttgg tgccctgccc ctccccatga cctgcagaga cattcagcct gccaggcttt 1020
atgaggtggg agctgggctc tccctgatgt attattcagc tccctggagt tggccagctc 1080
ctgttacact ggccacagcc ctgggcatcc gcttctcact tctagtttcc cctccaaggt 1140
aatgtggtgg gtcatgatca ttctatcctg gcttcaggga cctgactcca ctttggggcc 1200
attcgagggg tctagggtag atgatgtccc cctgtgggga ttaatgtcct gctctgtaaa 1260
actgagctag ctgagatcca ggagggcttg gccagagaca gcaagttgtt gccatggtga 1320
ctttaaagcc aggttgctgc cccagcacag gcctcccagt ctaccctcac tagaaaacaa 1380
cacccaggca ctttccacca cctctcaaag gtgaaaccca aggctggtct agagaatgaa 1440
ttatggatcc tcgctgtccg tgccacccag ctagtcccag cggctcagac actgaggaga 1500
gactgtaggt tcagctacaa gcaaaaagac ctagctggtc tccaagcagt gtctccaagt 1560
ccctgaacct gtgacacctg ccccaggcat catcaggcac agagggccac c 1611

<210> 37
<211> 1611
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 37

ggtctcaccc agcattttca cttctaataa gttcaaatgt gatacggcac ctttctaaaa 60
attagttttc agggaaatag ggttcaaaac tggtagtggt agggtccatt ctcacgaccc 120
ccaggcctgc taaccctgac caagctacct attacttacc ctcctctttc tcctcctcct 180
ctttctcctt ctcctgcttc ccctcttcct tctccctccc ttcctctccc tcctcccccc 240
ccttggctgt gatcagatcc agagcctgaa tgagcctcct gaccccacac ccccactagc 300
atgggcctgc aagtgcccag aagtccctcc tgcctcctaa actgcccagc cgatccatta 360
gctcttcctt cttcccagtg aaagaagcag gcacagcctg tccctcccgt tctacagaaa 420
ggaagctaca gcacagggag ggccaaaggc cttcctggga ctagacagtt gatcaacagc 480
aggactggag agctgggctc catttttgtt ccttggtgcc ctgcccctcc ccatgacctg 540
cagagacatt cagcctgcca ggctttatga ggtgggagct gggctctccc tgatgtatta 600
ttcagctccc tggagttggc cagctcctgt tacactggcc acagccctgg gcatccgctt 660
ctcacttcta gtttcccctc caaggtaatg tggtgggtca tgatcattct atcctggctt 720
cagggacctg actccacttt ggggccattc gaggggtcta gggtagatga tgtcccctg 780
tggggattaa tgtcctgctc tgtaaaactg agctagctga gatccaggag ggcttggcca 840
gagacagcaa gttgttgcca tggtgacttt aaagccaggt tgctgcccca gcacaggcct 900
cccagtctac cctcactaga aaacaacacc caggcacttt ccaccacctc tcaaaggtga 960

aacccaaggc tggtctagag aatgaattat ggatcctcgc tgtccgtgcc acccagctag 1020
tcccagcggc tcagacactg aggagagact gtaggttcag ctacaagcaa aaagacctag 1080
ctggtctcca agcagtgtct ccaagtccct gaacctgtga cacctgcccc aggcatcatc 1140
aggcacagag ggccaccctg cagctcagcc tactacttgc tttccaggct gttcctagtt 1200
cccatgtcag ctgcttgtgc tttccagaga caaaacagga ataatagatg tcattaaata 1260
tacattgggc cccaggcggt caatgtggca gcctgagcct cctttccatc tctgtggagg 1320
cagacatagg acccccaaca aacagcatgc aggttgggag ccagccacag gacccaggta 1380
aggggccctg ggtccttaag cttctgccac tggctccggc attgcagaga gaagagaagg 1440
ggcggcagag ctgaacctta gccttgcctt cctgggtacc cttctgagcc tcactgtctt 1500
ctgtgagatg ggcaaagtgc gggtgtgact ccttggcaac ggtgttacac cagggcaggt 1560
aaagttgtag ttatttgtgg ggtacaccag gactgttaaa ggtgtaacta t 1611

<210> 38
<211> 965
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 38

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggacccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240
aagcttctgc cactggctcc ggcattgcag agagaagaga aggggcggca gactggagag 300
ctgggctcca tttttgttcc ttggtgccct gcccctcccc atgacctgca gagacattca 360
gcctgccagg ctttatgagg tgggagctgg gctctccctg atgtattatt cagctccctg 420
gagttggcca gctcctgtta cactggccac agccctgggc atccgcttct cacttctagt 480
ttcccctcca aggtaatgtg gtgggtcatg atcattctat cctggcttca gggacctgac 540
tccactttgg ggccattcga ggggtctagg gtagatgatg tccccctgtg gggattaatg 600
tcctgctctg taaaactgag ctagctgaga tccaggaggg cttggccaga gacagcaagt 660
tgttgccatg gtgactttaa agccaggttg ctgccccagc acaggcctcc cagtctaccc 720
tcactagaaa acaacaccca ggcactttcc accacctctc aaaggtgaaa cccaaggctg 780
gtctagagaa tgaattatgg atcctcgctg tccgtgccac ccagctagtc ccagcggctc 840
agacactgag gagagactgt aggttcagct acaagcaaaa agacctagct ggtctccaag 900
cagtgtctcc aagtccctga acctgtgaca cctgccccag gcatcatcag gcacagaggg 960
ccacc 965

<210> 39
<211> 520
<212> DNA
<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 39

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60

tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120

ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180

acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240

tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 300

ctcctgttac actggccaca gccctgggca tccgctgcca tggtgacttt aaagccaggt 360

tgctgcccca gcacaggcct cccagtctac cctcactaga aaacaacacc caggcacttt 420

ccaccacctc tcaaaggtga aacccaaggc tggtctagag aatgaattat ggatcctcgc 480

tgtccgtgcc acccagctag tcccagcggc tcagacactg 520

<210> 40

<211> 235

<212> DNA

<213> 智人(Homo sapiens)

<400> 40

gtatgccttt tgagatggat gcagcaggtt ctgtgaggct gccaggaggg gtagagttcc 60

cgggggcctc gggccccgct ggagtgtgga gcaggcccat gctcagctct ccaggctgtt 120

cgtggctccc ctgtcagctg ctcactcctt tccagagaca aaacaggaat aatagacatc 180

attaaatata catagggccc caggcggtcg gcgtggtggg ctgggcctcc cttcc 235

<210> 41

<211> 688

<212> DNA

<213> 智人(Homo sapiens)

<400> 41

tgccctgcct tctgagccgg cagcctggct ccccacccca tgtattattc agctcctgag 60

agccagccag ctcctgttac actgaccgca gcccagcacc tgctctgccc attcccctcc 120

tcccttgcct aggacctaga gggttcaaag ttctcctcca agatgacttg gtgggctttg 180

gccatcccac cctaggcccc acttctggcc cagtgcaggt gtgctggtga tttagggcag 240

gtggcattcc atctctgtgg ctcaatgtct tcctctgtga agccgaagtg acccaagggc 300

tcccttcatg gggttgagcc agctgtggcc cagggagggc ctaaccagga tgagcactga 360

tgttgccatg acgactccga ggccagaatg tctcccccag cacaggcctc ataggcaggc 420

ttccccatcc tggtaaacaa cacccacaca ctttctacta ctgctctagg gtgaaaccca 480

aggcgctcta gaggagatga attatggatc cgccctcccg gaatcctggc tcggccctcc 540

ccacgccacc cagggccagt cgggtctgct cacagcccga ggaggccgcg tgtccagccg 600

cgggcaagag acagagcagg tccctgtgtc tccaagtccc tgagcccgtg acaccggccc 660

caggccctgt agagagcagg cagccacc 688

<210> 42

<211> 77
<212> DNA
<213> 智人(Homo sapiens)
<400> 42
cccctgtcag ctgctcactc ctttccagag acaaaacagg aataatagac atcattaaat 60
atacataggg ccccagg 77
<210> 43
<211> 82
<212> DNA
<213> 智人(Homo sapiens)
<400> 43
tgagccggca gcctggctcc ccaccccatg tattattcag ctcctgagag ccagccagct 60
cctgttacac tgaccgcagc cc 82
<210> 44
<211> 111
<212> DNA
<213> 智人(Homo sapiens)
<400> 44
cacaggcctc ataggcaggc ttccccatcc tggtaaacaa cacccacaca ctttctacta 60
ctgctctagg gtgaaaccca aggcgctcta gaggagatga attatggatc c 111
<210> 45
<211> 193
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 45
tgagccggca gcctggctcc ccaccccatg tattattcag ctcctgagag ccagccagct 60
cctgttacac tgaccgcagc cccacaggcc tcataggcag gcttccccat cctggtaaac 120
aacacccaca cactttctac tactgctcta gggtgaaacc caaggcgctc tagaggagat 180
gaattatgga tcc 193
<210> 46
<211> 193
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 46
cacaggcctc ataggcaggc ttccccatcc tggtaaacaa cacccacaca ctttctacta 60

ctgctctagg gtgaaaccca aggcgctcta gaggagatga attatggatc ctgagccggc 120
agcctggctc cccaccccat gtattattca gctcctgaga gccagccagc tcctgttaca 180
ctgaccgcag ccc 193

<210> 47
<211> 499
<212> DNA
<213> 智人(Homo sapiens)
<400> 47

tgagccggca gcctggctcc ccaccccatg tattattcag ctcctgagag ccagccagct 60
cctgttacac tgaccgcagc ccagcacctg ctctgcccat tcccctcctc ccttgcctag 120
gacctagagg gttcaaagtt ctcctccaag atgacttggt gggctttggc catcccaccc 180
taggccccac ttctggccca gtgcaggtgt gctggtgatt tagggcaggt ggcattccat 240
ctctgtggct caatgtcttc ctctgtgaag ccgaagtgac ccaagggctc ccttcatggg 300
gttgagccag ctgtggccca gggagggcct aaccaggatg agcactgatg ttgccatgac 360
gactccgagg ccagaatgtc tcccccagca caggcctcat aggcaggctt ccccatcctg 420
gtaaacaaca cccacacact ttctactact gctctagggt gaaacccaag gcgctctaga 480
ggagatgaat tatggatcc 499

<210> 48
<211> 1262
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 48

gtatgccttt tgagatggat gcagcaggtt ctgtgaggct gccaggaggg gtagagttcc 60
cgggggcctc gggccccgct ggagtgtgga gcaggcccat gctcagctct ccaggctgtt 120
cgtggctccc ctgtcagctg ctcactcctt tccagagaca aaacaggaat aatagacatc 180
attaaatata catagggccc caggcggtcg gcgtggtggg ctgggcctcc cttccccata 240
acactgagct gctctgctgg gccaaccgtg ctcctgggcc agccagagga cccccatgag 300
gcggcatgca ggcggggagc aggccacaga acgcaggtaa ggagacctta gcctagagtc 360
cttggggtct gtcactggcc accctcgcat cccaggctgc aggaaactga ggcccagaga 420
ggacaaggac tttcctggac ccacacagcc agtcagtgac agagcctagg gtctgagcca 480
ggcctgaccc aacctccatt tctgcctctc taccccctgcc cccgccccaa cacacacaca 540
cacacaagtg gagttccact gaaacgcccc tccttgccct gccttctgag ccggcagcct 600
ggctccccac cccatgtatt attcagctcc tgagagccag ccagctcctg ttacactgac 660
cgcagcccag cacctgctct gcccattccc ctcctccctt gcctaggacc tagagggttc 720
aaagttctcc tccaagatga cttggtgggc tttggccatc ccaccctagg ccccacttct 780
ggcccagtgc aggtgtgctg gtgatttagg gcaggtggca ttccatctct gtggctcaat 840
gtcttcctct gtgaagccga agtgacccaa gggctccctt catggggttg agccagctgt 900

ggcccaggga gggcctaacc aggatgagca ctgatgttgc catgacgact ccgaggccag 960
aatgtctccc ccagcacagg cctcataggc aggcttcccc atcctggtaa acaacaccca 1020
cacactttct actactgctc tagggtgaaa cccaaggcgc tctagaggag atgaattatg 1080
gatccgccct cccggaatcc tggctcggcc ctccccacgc cacccagggc cagtcgggtc 1140
tgctcacagc ccgaggaggc cgcgtgtcca gccgcgggca agagacagag caggtccctg 1200
tgtctccaag tccctgagcc cgtgacaccg gccccaggcc ctgtagagag caggcagcca 1260
cc 1262

<210> 49
<211> 649
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 49

gcaggcccat gctcagctct ccaggctgtt cgtggctccc ctgtcagctg ctcactcctt 60
tccagagaca aaacaggaat aatagacatc attaaatata catagggccc caggcggtcg 120
gcgtggtggg ctgggcctcc cttccccata acactgagct gctctgctgg gccaaccgtg 180
ctcctgggcc agccagagga cccccatgag gcggcatgca ggcggggagc aggccacaga 240
acgcaggtaa ggagaccttg ccttctgagc cggcagcctg gctccccacc ccatgtatta 300
ttcagctcct gagagccagc cagctcctgt tacactgacc gcagcccagc acctgctctg 360
cccattcccc tcctcccttg cctaggacct agagggttca aagttctcct ccaagatgac 420
ttggtgggct ttggccatcg ggcctaacca ggatgagcac tgatgttgcc atgacgactc 480
cgaggccaga atgtctcccc cagcacaggc ctcataggca ggcttcccca tcctggtaaa 540
caacacccac acactttcta ctactgctct agggtgaaac ccaaggcgct ctagaggaga 600
tgaattatgg atccgccctc ccggaatcct ggctcggccc tccccacgc 649

<210> 50
<211> 240
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 50

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240

<210> 51
<211> 95
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 51

tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 60
ctcctgttac actggccaca gccctgggca tccgc 95

<210> 52
<211> 185
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 52

tgccatggtg actttaaagc caggttgctg ccccagcaca ggcctcccag tctaccctca 60
ctagaaaaca acacccaggc actttccacc acctctcaaa ggtgaaaccc aaggctggtc 120
tagagaatga attatggatc ctcgctgtcc gtgccaccca gctagtccca gcggctcaga 180
cactg 185

<210> 53
<211> 335
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 53

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240
tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 300
ctcctgttac actggccaca gccctgggca tccgc 335

<210> 54
<211> 425
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 54

ctgcagctca gcctactact tgctttccag gctgttccta gttcccatgt cagctgcttg 60
tgctttccag agacaaaaca ggaataatag atgtcattaa atatacattg ggccccaggc 120
ggtcaatgtg gcagcctgag cctcctttcc atctctgtgg aggcagacat aggaccccca 180
acaaacagca tgcaggttgg gagccagcca caggacccag gtaaggggcc ctgggtcctt 240
tgccatggtg actttaaagc caggttgctg ccccagcaca ggcctcccag tctaccctca 300
ctagaaaaca acacccaggc actttccacc acctctcaaa ggtgaaaccc aaggctggtc 360
tagagaatga attatggatc ctcgctgtcc gtgccaccca gctagtccca gcggctcaga 420
cactg 425

<210> 55
<211> 280
<212> DNA
<213> 小家鼠(Mus musculus)
<400> 55
tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 60
ctcctgttac actggccaca gccctgggca tccgctgcca tggtgacttt aaagccaggt 120
tgctgcccca gcacaggcct cccagtctac cctcactaga aaacaacacc caggcacttt 180
ccaccacctc tcaaaggtga aacccaaggc tggtctagag aatgaattat ggatcctcgc 240
tgtccgtgcc acccagctag tcccagcggc tcagacactg 280
<210> 56
<211> 520
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 56
tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 60
ctcctgttac actggccaca gccctgggca tccgcctgca gctcagccta ctacttgctt 120
tccaggctgt tcctagttcc catgtcagct gcttgtgctt tccagagaca aaacaggaat 180
aatagatgtc attaaatata cattgggccc caggcggtca atgtggcagc ctgagcctcc 240
tttccatctc tgtggaggca gacataggac ccccaacaaa cagcatgcag gttgggagcc 300
agccacagga cccaggtaag gggccctggg tcctttgcca tggtgacttt aaagccaggt 360
tgctgcccca gcacaggcct cccagtctac cctcactaga aaacaacacc caggcacttt 420
ccaccacctc tcaaaggtga aacccaaggc tggtctagag aatgaattat ggatcctcgc 480
tgtccgtgcc acccagctag tcccagcggc tcagacactg 520
<210> 57
<211> 520
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 57
tgccatggtg actttaaagc caggttgctg ccccagcaca ggcctcccag tctaccctca 60
ctagaaaaca acacccaggc actttccacc acctctcaaa ggtgaaaccc aaggctggtc 120
tagagaatga attatggatc ctcgctgtcc gtgccaccca gctagtccca gcggctcaga 180
cactgctgca gctcagccta ctacttgctt tccaggctgt tcctagttcc catgtcagct 240
gcttgtgctt tccagagaca aaacaggaat aatagatgtc attaaatata cattgggccc 300
caggcggtca atgtggcagc ctgagcctcc tttccatctc tgtggaggca gacataggac 360

ccccaacaaa cagcatgcag gttgggagcc agccacagga cccaggtaag gggccctggg 420
tccttttttat gaggtgggag ctgggctctc cctgatgtat tattcagctc cctggagttg 480
gccagctcct gttacactgg ccacagccct gggcatccgc 520

<210> 58
<211> 520
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 58

tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg agttggccag 60
ctcctgttac actggccaca gccctgggca tccgctgcca tggtgacttt aaagccaggt 120
tgctgcccca gcacaggcct cccagtctac cctcactaga aaacaacacc caggcacttt 180
ccaccacctc tcaaaggtga aacccaaggc tggtctagag aatgaattat ggatcctcgc 240
tgtccgtgcc acccagctag tcccagcggc tcagacactg ctgcagctca gcctactact 300
tgctttccag gctgttccta gttcccatgt cagctgcttg tgctttccag agacaaaaca 360
ggaataatag atgtcattaa atatacattg ggccccaggc ggtcaatgtg gcagcctgag 420
cctcctttcc atctctgtgg aggcagacat aggaccccca acaaacagca tgcaggttgg 480
gagccagcca caggacccag gtaaggggcc ctgggtcctt 520

<210> 59
<211> 964
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 59

tgcagctcag cctactactt gctttccagg ctgttcctag ttcccatgtc agctgcttgt 60
gctttccaga gacaaaacag gaataataga tgtcattaaa tatacattgg gccccaggcg 120
gtcaatgtgg cagcctgagc ctcctttcca tctctgtgga ggcagacata ggacccccaa 180
caaacagcat gcaggttggg agccagccac aggacccagg taaggggccc tgggtcctta 240
agcttctgcc actggctccg gcattgcaga gagaagagaa ggggcggcag actggagagc 300
tgggctccat ttttgttcct tggtgccctg cccctcccca tgacctgcag agacattcag 360
cctgccaggc tttatgaggt gggagctggg ctctccctga tgtattattc agctccctgg 420
agttggccag ctcctgttac actggccaca gccctgggca tccgcttctc acttctagtt 480
tcccctccaa ggtaatgtgg tgggtcatga tcattctatc ctggcttcag ggacctgact 540
ccactttggg gccattcgag gggtctaggg tagatgatgt ccccctgtgg ggattaatgt 600
cctgctctgt aaaactgagc tagctgagat ccaggagggc ttggccagag acagcaagtt 660
gttgccatgg tgactttaaa gccaggttgc tgccccagca caggcctccc agtctaccct 720
cactagaaaa caacacccag gcactttcca ccacctctca aaggtgaaac ccaaggctgg 780

tctagagaat gaattatgga tcctcgctgt ccgtgccacc cagctagtcc cagcggctca 840
gacactgagg agagactgta ggttcagcta caagcaaaaa gacctagctg gtctccaagc 900
agtgtctcca agtccctgaa cctgtgacac ctgccccagg catcatcagg cacagagggc 960
cacc 964

<210> 60
<211> 519
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 60

tgcagctcag cctactactt gctttccagg ctgttcctag ttcccatgtc agctgcttgt 60
gctttccaga gacaaaacag gaataataga tgtcattaaa tatacattgg gccccaggcg 120
gtcaatgtgg cagcctgagc ctcctttcca tctctgtgga ggcagacata ggacccccaa 180
caaacagcat gcaggttggg agccagccac aggacccagg taaggggccc tgggtccttt 240
ttatgaggtg ggagctgggc tctccctgat gtattattca gctccctgga gttggccagc 300
tcctgttaca ctggccacag ccctgggcat ccgctgccat ggtgacttta aagccaggtt 360
gctgccccag cacaggcctc ccagtctacc ctcactagaa aacaacaccc aggcactttc 420
caccacctct caaaggtgaa acccaaggct ggtctagaga atgaattatg gatcctcgct 480
gtccgtgcca cccagctagt cccagcggct cagacactg 519

<210> 61
<211> 425
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 61

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> 62
<211> 865
<212> DNA
<213> 智人(Homo sapiens)

<400> 62

ccccgggtgc gcggcgtcgg tggtgccggc gggggggcgcc aggtcgcagg cggtgtaggg 60
ctccaggcag gcggcgaagg ccatgacgtg cgctatgaag gtctgctcct gcacgccgtg 120
aaccaggtgc gcctgcgggc cgcgcgcgaa caccgccacg tcctcgcctg cgtgggtctc 180
ttcgtccagg ggcactgctg actgctgccg atactcgggg ctcccgctct cgctctcggt 240
aacatccggc cgggcgccgt ccttgagcac atagcctgga ccgtttccgt ataggaggac 300
cgtgtaggcc ttcctgtccc gggccttgcc agcggccagc ccgatgaagg agctccctcg 360
caggggggtag cctccgaagg agaagacgtg ggagtggtcg gcagtgacga ggctcagcgt 420
gtcctcctcg ctggtgagct ggcccgccct ctcaatggcg tcgtcgaaca tgatcgtctc 480
agtcagtgcc cggtaagccc tgctttcatg atgaccatgg tcgatgcgac caccctccac 540
gaagaggaag aagccgcggg ggtgtctgct cagcaggcgc agggcagcct ctgtcatctc 600
catcagggag gggtccagtg tggagtctcg gtggatctcg tatttcatgt ctccaggctc 660
aaagagaccc atgagatggg tcacagacgg gtccagggaa gcctgcatga gctcagtgcg 720
gttccacacg taccgggcac cctggcgttc gccgagccat tcctgcacca gattcttccc 780
gtccagcctg gtcccacctt ggctgtagtc atctgggtac tcagggtctg gggttcccat 840
gcgaaacatg tactttcggc ctcca 865

<210> 63
<211> 437
<212> DNA
<213> 智人(Homo sapiens)

<400> 63

ccccgggtgc gcggcgtcgg tggtgccggc gggggggcgcc aggtcgcagg cggtgtaggg 60
ctccaggcag gcggcgaagg ccatgacgtg cgctatgaag gtctgctcct gcacgccgtg 120
aaccaggtgc gcctgcgggc cgcgcgcgaa caccgccacg tcctcgcctg cgtgggtctc 180
ttcgtccagg ggcactgctg actgctgccg atactcgggg ctcccgctct cgctctcggt 240
aacatccggc cgggcgccgt ccttgagcac atagcctgga ccgtttccgt ataggaggac 300
cgtgtaggcc ttcctgtccc gggccttgcc agcggccagc ccgatgaagg agctccctcg 360
caggggggtag cctccgaagg agaagacgtg ggagtggtcg gcagtgacga ggctcagcgt 420
gtcctcctcg ctggtga 437

<210> 64
<211> 428
<212> DNA
<213> 智人(Homo sapiens)

<400> 64

gctggcccgc cctctcaatg gcgtcgtcga acatgatcgt ctcagtcagt gcccggtaag 60
ccctgctttc atgatgacca tggtcgatgc gaccaccctc cacgaagagg aagaagccgc 120
gggggtgtct gctcagcagg cgcagggcag cctctgtcat ctccatcagg gaggggtcca 180
gtgtggagtc tcggtggatc tcgtatttca tgtctccagg ctcaaagaga cccatgagat 240
gggtcacaga cgggtccagg gaagcctgca tgagctcagt gcggttccac acgtaccggg 300

caccctggcg ttcgccgagc cattcctgca ccagattctt cccgtccagc ctggtcccac 360
cttggctgta gtcatctggg tactcagggt ctggggttcc catgcgaaac atgtactttc 420
ggcctcca 428

<210> 65
<211> 287
<212> DNA
<213> 智人(Homo sapiens)
<400> 65

ccccgggtgc gcggcgtcgg tggtgccggc ggggggcgcc aggtcgcagg cggtgtaggg 60
ctccaggcag gcggcgaagg ccatgacgtg cgctatgaag gtctgctcct gcacgccgtg 120
aaccaggtgc gcctgcgggc cgcgcgcgaa caccgccacg tcctcgcctg cgtgggtctc 180
ttcgtccagg ggcactgctg actgctgccg atactcgggg ctcccgctct cgctctcggt 240
aacatccggc cgggcgccgt ccttgagcac atagcctgga ccgtttc 287

<210> 66
<211> 290
<212> DNA
<213> 智人(Homo sapiens)
<400> 66

cgtataggag gaccgtgtag gccttcctgt cccgggcctt gccagcggcc agcccgatga 60
aggagctccc tcgcaggggg tagcctccga aggagaagac gtgggagtgg tcggcagtga 120
cgaggctcag cgtgtcctcc tcgctggtga gctggcccgc cctctcaatg gcgtcgtcga 180
acatgatcgt ctcagtcagt gcccggtaag ccctgctttc atgatgacca tggtcgatgc 240
gaccaccctc cacgaagagg aagaagccgc gggggtgtct gctcagcagg 290

<210> 67
<211> 288
<212> DNA
<213> 智人(Homo sapiens)
<400> 67

cgcagggcag cctctgtcat ctccatcagg gaggggtcca gtgtggagtc tcggtggatc 60
tcgtatttca tgtctccagg ctcaaagaga cccatgagat gggtcacaga cgggtccagg 120
gaagcctgca tgagctcagt gcggttccac acgtaccggg caccctggcg ttcgccgagc 180
cattcctgca ccagattctt cccgtccagc ctggtcccac cttggctgta gtcatctggg 240
tactcagggt ctggggttcc catgcgaaac atgtactttc ggcctcca 288

<210> 68
<211> 84
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体

<400> 68
gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg cttgtcgaga 60
cagagaagac tcttgcgttt ctga 84
<210> 69
<211> 49
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 69
taggcaccta ttggtcttac tgacatccac tttgcctttc tctccacag 49
<210> 70
<211> 943
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 70
ggtacctagt tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga 60
gttccgcgtt acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg 120
cccattgacg tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg 180
acgtcaatgg gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca 240
tatgccaagt acgcccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc 300
ccagtacatg accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc 360
tattaccatg gtcgaggtga gccccacgtt ctgcttcact ctccccatct ccccccccctc 420
cccaccccca attttgtatt tatttatttt ttaattattt tgtgcagcga tgggggcggg 480
gggggggggg gggcgcgcgc caggcggggc ggggcggggc gaggggcggg gcggggcgag 540
gcggagaggt gcggcggcag ccaatcagag cggcgcgctc cgaaagtttc cttttatggc 600
gaggcggcgg cggcggcggc cctataaaaa gcgaagcgcg cggcgggcgg gagtcgctgc 660
gcgctgcctt cgccccgtgc cccgctccgc cgccgcctcg cgccgcccgc cccggctctg 720
actgaccgcg ttactcccac aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa 780
ttagcgcttg gtttaatgac ggcttgtttc ttttctgtgg ctgcgtgaaa gccttgaggg 840
gctccgggag ctagagcctc tgctaaccat gttcatgcct tcttcttttt cctacagctc 900
ctgggcaacg tgctggttat tgtgctgtct catcattttg gca 943
<210> 71
<211> 2406
<212> DNA
<213> 智人(Homo sapiens)
<400> 71

atggccttgc tcatccacct caagacagtc tcggagctgc ggggcagggg cgaccggatc 60
gccaaagtga ctttccgagg gcaatccttc tactctcggg tcctggagaa ctgtgaggat 120
gtggctgact ttgatgagac atttcggtgg ccggtggcca gcagcatcga cagaaatgag 180
atgctggaga ttcaggtttt caactacagc aaagtcttca gcaacaagct catcgggacc 240
ttccgcatgg tgctgcagaa ggtggtagag gagagccatg tggaggtgac tgacacgctg 300
attgatgaca acaatgctat catcaagacc agcctgtgcg tggaggtccg gtatcaggcc 360
actgacggca cagtgggctc ctgggacgat ggggacttcc tgggagatga gtctcttcaa 420
gaggaagaga aggacagcca agagacggat ggactgctcc caggctcccg gcccagctcc 480
cggcccccag gagagaagag cttccggaga gccgggagga gcgtgttctc cgccatgaag 540
ctcggcaaaa accggtctca caaggaggag ccccaaagac cagatgaacc ggcggtgctg 600
gagatggaag accttgacca tctggccatt cggctaggag atggactgga tcccgactcg 660
gtgtctctag cctcagtcac agctctcacc actaatgtct ccaacaagcg atctaagcca 720
gacattaaga tggagccaag tgctgggcgg cccatggatt accaggtcag catcacggtg 780
atcgaggccc ggcagctggt gggcttgaac atggaccctg tggtgtgcgt ggaggtgggt 840
gacgacaaga agtacacatc catgaaggag tccactaact gcccctatta caacgagtac 900
ttcgtcttcg acttccatgt ctctccggat gtcatgtttg acaagatcat caagatttcg 960
gtgattcact ccaagaacct gctgcgcagt ggcaccctgg tgggctcctt caaaatggac 1020
gtgggaaccg tgtactcgca gccagagcac cagttccatc acaagtgggc catcctgtct 1080
gaccccgatg acatctcctc ggggctgaag ggctacgtga agtgtgacgt tgccgtggtg 1140
ggcaaagggg acaacatcaa gacgccccac aaggccaatg agaccgacga agatgacatt 1200
gaggggaact tgctgctccc cgagggggtg ccccccgaac gccagtgggc ccggttctat 1260
gtgaaaattt accgagcaga ggggctgccc cgtatgaaca caagcctcat ggccaatgta 1320
aagaaggctt tcatcggtga aaacaaggac ctcgtggacc cctacgtgca agtcttcttt 1380
gctggccaga agggcaagac ttcagtgcag aagagcagct atgagcccct gtggaatgag 1440
caggtcgtct ttacagacct cttcccccca ctctgcaaac gcatgaaggt gcagatccga 1500
gactcggaca aggtcaacga cgtggccatc ggcacccact tcattgacct gcgcaagatt 1560
tctaatgacg gagacaaagg cttcctgccc acactgggcc cagcctgggt gaacatgtac 1620
ggctccacac gtaactacac gctgctggat gagcatcagg acctgaacga gggcctgggg 1680
gagggtgtgt ccttccgggc ccggctcctg ctgggcctgg ctgtggagat cgtagacacc 1740
tccaaccctg agctcaccag ctccacagag gtgcaggtgg agcaggccac gcccatctcg 1800
gagagctgtg caggtaaaat ggaagaattc tttctctttg gagccttcct ggaggcctca 1860
atgatcgacc ggagaaacgg agacaagccc atcacctttg aggtcaccat aggcaactat 1920
gggaacgaag ttgatggcct gtcccggccc cagcggcctc ggccccggaa ggagccgggg 1980
gatgaggaag aagtagacct gattcagaac gcaagtgatg acgaggccgg tgatgccggg 2040
gacctggcct cagtctcctc cactccacca atgcggcccc aggtcaccga caggaactac 2100
ttccatctgc cctacctgga gcgaaagccc tgcatctaca tcaagagctg gtggccggac 2160
cagcgccgcc gcctctacaa tgccaacatc atggaccaca ttgccgacaa gctggaagaa 2220
ggcctgaacg acatacagga gatgatcaaa acggagaagt cctaccctga gcgtcgcctg 2280
cggggcgtcc tggaggagct gagctgtggc tgctgccgct tcctctccct cgctgacaag 2340

gaccagggcc actcatcccg caccaggctt gaccgggagc gcctcaagtc ctgcatgagg 2400
gagctg 2406

<210> 72
<211> 3588
<212> DNA
<213> 智人(Homo sapiens)
<400> 72

gaaaacatgg ggcagcaggc caggatgctg cgggcccagg tgaagcggca cacggtgcgg 60
gacaagctga ggctgtgcca gaacttcctg cagaagctgc gcttcctggc ggacgagccc 120
cagcacagca ttcccgacat cttcatctgg atgatgagca acaacaagcg tgtcgcctat 180
gcccgtgtgc cctccaagga cctgctcttc tccatcgtgg aggaggagac tggcaaggac 240
tgcgccaagg tcaagacgct cttccttaag ctgccaggga agcggggctt cggctcggca 300
ggctggacag tgcaggccaa ggtggagctg tacctgtggc tgggcctcag caaacagcgc 360
aaggagttcc tgtgcggcct gccctgtggc ttccaggagg tcaaggcagc ccagggcctg 420
ggcctgcatg ccttcccacc cgtcagcctg gtctacacca agaagcaggc gttccagctc 480
cgagcgcaca tgtaccaggc ccgcagcctc tttgccgccg acagcagcgg actctcagac 540
ccctttgccc gcgtcttctt catcaatcag agtcagtgca cagaggtgct gaatgagacc 600
ctgtgtccca cctgggacca gatgctggtg ttcgacaacc tggagctcta tggtgaagct 660
catgagctga gggacgatcc gcccatcatt gtcattgaaa tctatgacca ggattccatg 720
ggcaaagctg acttcatggg ccggaccttc gccaaacccc tggtgaagat ggcagacgag 780
gcgtactgcc caccccgctt cccacctcag ctcgagtact accagatcta ccgtggcaac 840
gccacagctg gagacctgct ggcggccttc gagctgctgc agattggacc agcagggaag 900
gctgacctgc cccccatcaa tggcccggtg gacgtggacc gaggtcccat catgcccgtg 960
cccatgggca tccggcccgt gctcagcaag taccgagtgg aggtgctgtt ctggggccta 1020
cgggacctaa agcgggtgaa cctggcccag gtggaccggc cacgggtgga catcgagtgt 1080
gcagggaagg gggtgcagtc gtccctgatc cacaattata agaagaaccc caacttcaac 1140
accctcgtca agtggtttga agtggacctc ccagagaacg agctgctgca cccgcccttg 1200
aacatccgtg tggtggactg ccgggccttc ggtcgctaca cactggtggg ctcccatgcc 1260
gtcagctccc tgcgacgctt catctaccgg cccccagacc gctcggcccc cagctggaac 1320
accacggtca ggcttctccg gcgctgccgt gtgctgtgca atgggggctc ctcctctcac 1380
tccacagggg aggttgtggt gactatggag ccagaggtac ccatcaagaa actggagacc 1440
atggtgaagc tggacgcgac ttctgaagct gttgtcaagg tggatgtggc tgaggaggag 1500
aaggagaaga agaagaagaa gaagggcact gcggaggagc cagaggagga ggagccagac 1560
gagagcatgc tggactggtg gtccaagtac tttgcctcca ttgacaccat gaaggagcaa 1620
cttcgacaac aagagccctc tggaattgac ttggaggaga aggaggaagt ggacaatacc 1680
gagggcctga aggggtcaat gaagggcaag gagaaggcaa gggctgccaa agaggagaag 1740
aagaagaaaa ctcagagctc tggctctggc caggggtccg aggcccccga gaagaagaaa 1800
cccaagattg atgagcttaa ggtatacccc aaagagctgg agtccgagtt tgataacttt 1860
gaggactggc tgcacacttt caacttgctt cggggcaaga ccggggatga tgaggatggc 1920

tccaccgagg aggagcgcat tgtgggacgc ttcaagggct ccctctgcgt gtacaaagtg 1980
ccactcccag aggacgtgtc ccgggaagcc ggctacgact ccacctacgg catgttccag 2040
ggcatcccga gcaatgaccc catcaatgtg ctggtccgag tctatgtggt ccgggccacg 2100
gacctgcacc ctgctgacat caacggcaaa gctgacccct acatcgccat ccggctaggc 2160
aagactgaca tccgcgacaa ggagaactac atctccaagc agctcaaccc tgtctttggg 2220
aagtcctttg acatcgaggc ctccttcccc atggaatcca tgctgacggt ggctgtgtat 2280
gactgggacc tggtgggcac tgatgacctc attggggaaa ccaagatcga cctggagaac 2340
cgcttctaca gcaagcaccg cgccacctgc ggcatcgccc agacctactc cacacatggc 2400
tacaatatct ggcgggaccc catgaagccc agccagatcc tgacccgcct ctgcaaagac 2460
ggcaaagtgg acggccccca ctttgggccc cctgggagag tgaaggtggc caaccgcgtc 2520
ttcactgggc cctctgagat tgaggacgag aacggtcaga ggaagcccac agacgagcat 2580
gtggcgctgt tggccctgag gcactgggag gacatccccc gcgcaggctg ccgcctggtg 2640
ccagagcatg tggagacgag gccgctgctc aaccccgaca agccgggcat cgagcagggc 2700
cgcctggagc tgtgggtgga catgttcccc atggacatgc cagcccctgg gacgcctctg 2760
gacatctcac ctcggaagcc caagaagtac gagctgcggg tcatcatctg gaacacagat 2820
gaggtggtct tggaggacga cgacttcttc acaggggaga agtccagtga catcttcgtg 2880
agggggtggc tgaagggcca gcaggaggac aagcaggaca cagacgtcca ctaccactcc 2940
ctcactggcg agggcaactt caactggcgc tacctgttcc ccttcgacta cctggcggcg 3000
gaggagaaga tcgtcatctc caagaaggag tccatgttct cctgggacga gaccgagtac 3060
aagatccccg cgcggctcac cctgcagatc tgggatgcgg accacttctc cgctgacgac 3120
ttcctggggg ccatcgagct ggacctgaac cggttcccgc ggggcgcaaa gacagccaag 3180
cagtgcacca tggagatggc caccggggag gtggacgtgc ccctcgtgtc catcttcaag 3240
caaaagcgcg tcaaaggctg gtggcccctc ctggcccgca atgagaacga tgagtttgag 3300
ctcacgggca aggtggaggc tgagctgcat ttactgacag cagaggaggc agagaagaac 3360
ccagtgggcc tggcccgcaa tgaacctgac cccctagaga aacccaaccg gcccgacacg 3420
gccttcgtct ggttcctcaa ccctctcaag tccatcaagt acctcatctg cacccggtac 3480
aagtggctca tcatcaagat cgtgctggcg ctgttggggc tgctcatgtt ggggctcttc 3540
ctctacagcc tccctggcta catggtcaaa aagctccttg gggcatga 3588

<210> 73
<211> 802
<212> PRT
<213> 智人(Homo sapiens)
<400> 73

Met Ala Leu Leu Ile His Leu Lys Thr Val Ser Glu Leu Arg Gly Arg
1 5 10 15
Gly Asp Arg Ile Ala Lys Val Thr Phe Arg Gly Gln Ser Phe Tyr Ser
20 25 30
Arg Val Leu Glu Asn Cys Glu Asp Val Ala Asp Phe Asp Glu Thr Phe
35 40 45

Arg Trp Pro Val Ala Ser Ser Ile Asp Arg Asn Glu Met Leu Glu Ile
50 55 60
Gln Val Phe Asn Tyr Ser Lys Val Phe Ser Asn Lys Leu Ile Gly Thr
65 70 75 80
Phe Arg Met Val Leu Gln Lys Val Val Glu Glu Ser His Val Glu Val
85 90 95
Thr Asp Thr Leu Ile Asp Asp Asn Asn Ala Ile Ile Lys Thr Ser Leu
100 105 110
Cys Val Glu Val Arg Tyr Gln Ala Thr Asp Gly Thr Val Gly Ser Trp
115 120 125
Asp Asp Gly Asp Phe Leu Gly Asp Glu Ser Leu Gln Glu Glu Glu Lys
130 135 140
Asp Ser Gln Glu Thr Asp Gly Leu Leu Pro Gly Ser Arg Pro Ser Ser
145 150 155 160
Arg Pro Pro Gly Glu Lys Ser Phe Arg Arg Ala Gly Arg Ser Val Phe
165 170 175
Ser Ala Met Lys Leu Gly Lys Asn Arg Ser His Lys Glu Glu Pro Gln
180 185 190
Arg Pro Asp Glu Pro Ala Val Leu Glu Met Glu Asp Leu Asp His Leu
195 200 205
Ala Ile Arg Leu Gly Asp Gly Leu Asp Pro Asp Ser Val Ser Leu Ala
210 215 220
Ser Val Thr Ala Leu Thr Thr Asn Val Ser Asn Lys Arg Ser Lys Pro
225 230 235 240
Asp Ile Lys Met Glu Pro Ser Ala Gly Arg Pro Met Asp Tyr Gln Val
245 250 255
Ser Ile Thr Val Ile Glu Ala Arg Gln Leu Val Gly Leu Asn Met Asp
260 265 270
Pro Val Val Cys Val Glu Val Gly Asp Asp Lys Lys Tyr Thr Ser Met
275 280 285
Lys Glu Ser Thr Asn Cys Pro Tyr Tyr Asn Glu Tyr Phe Val Phe Asp
290 295 300
Phe His Val Ser Pro Asp Val Met Phe Asp Lys Ile Ile Lys Ile Ser
305 310 315 320
Val Ile His Ser Lys Asn Leu Leu Arg Ser Gly Thr Leu Val Gly Ser
325 330 335
Phe Lys Met Asp Val Gly Thr Val Tyr Ser Gln Pro Glu His Gln Phe
340 345 350
His His Lys Trp Ala Ile Leu Ser Asp Pro Asp Asp Ile Ser Ser Gly

355 360 365
Leu Lys Gly Tyr Val Lys Cys Asp Val Ala Val Val Gly Lys Gly Asp
370 375 380
Asn Ile Lys Thr Pro His Lys Ala Asn Glu Thr Asp Glu Asp Asp Ile
385 390 395 400
Glu Gly Asn Leu Leu Leu Pro Glu Gly Val Pro Pro Glu Arg Gln Trp
405 410 415
Ala Arg Phe Tyr Val Lys Ile Tyr Arg Ala Glu Gly Leu Pro Arg Met
420 425 430
Asn Thr Ser Leu Met Ala Asn Val Lys Lys Ala Phe Ile Gly Glu Asn
435 440 445
Lys Asp Leu Val Asp Pro Tyr Val Gln Val Phe Phe Ala Gly Gln Lys
450 455 460
Gly Lys Thr Ser Val Gln Lys Ser Ser Tyr Glu Pro Leu Trp Asn Glu
465 470 475 480
Gln Val Val Phe Thr Asp Leu Phe Pro Pro Leu Cys Lys Arg Met Lys
485 490 495
Val Gln Ile Arg Asp Ser Asp Lys Val Asn Asp Val Ala Ile Gly Thr
500 505 510
His Phe Ile Asp Leu Arg Lys Ile Ser Asn Asp Gly Asp Lys Gly Phe
515 520 525
Leu Pro Thr Leu Gly Pro Ala Trp Val Asn Met Tyr Gly Ser Thr Arg
530 535 540
Asn Tyr Thr Leu Leu Asp Glu His Gln Asp Leu Asn Glu Gly Leu Gly
545 550 555 560
Glu Gly Val Ser Phe Arg Ala Arg Leu Leu Leu Gly Leu Ala Val Glu
565 570 575
Ile Val Asp Thr Ser Asn Pro Glu Leu Thr Ser Ser Thr Glu Val Gln
580 585 590
Val Glu Gln Ala Thr Pro Ile Ser Glu Ser Cys Ala Gly Lys Met Glu
595 600 605
Glu Phe Phe Leu Phe Gly Ala Phe Leu Glu Ala Ser Met Ile Asp Arg
610 615 620
Arg Asn Gly Asp Lys Pro Ile Thr Phe Glu Val Thr Ile Gly Asn Tyr
625 630 635 640
Gly Asn Glu Val Asp Gly Leu Ser Arg Pro Gln Arg Pro Arg Pro Arg
645 650 655
Lys Glu Pro Gly Asp Glu Glu Glu Val Asp Leu Ile Gln Asn Ala Ser
660 665 670

Asp Asp Glu Ala Gly Asp Ala Gly Asp Leu Ala Ser Val Ser Ser Thr
675 680 685
Pro Pro Met Arg Pro Gln Val Thr Asp Arg Asn Tyr Phe His Leu Pro
690 695 700
Tyr Leu Glu Arg Lys Pro Cys Ile Tyr Ile Lys Ser Trp Trp Pro Asp
705 710 715 720
Gln Arg Arg Arg Leu Tyr Asn Ala Asn Ile Met Asp His Ile Ala Asp
725 730 735
Lys Leu Glu Glu Gly Leu Asn Asp Ile Gln Glu Met Ile Lys Thr Glu
740 745 750
Lys Ser Tyr Pro Glu Arg Arg Leu Arg Gly Val Leu Glu Glu Leu Ser
755 760 765
Cys Gly Cys Cys Arg Phe Leu Ser Leu Ala Asp Lys Asp Gln Gly His
770 775 780
Ser Ser Arg Thr Arg Leu Asp Arg Glu Arg Leu Lys Ser Cys Met Arg
785 790 795 800
Glu Leu

<210> 74
<211> 1195
<212> PRT
<213> 智人(Homo sapiens)
<400> 74

Glu Asn Met Gly Gln Gln Ala Arg Met Leu Arg Ala Gln Val Lys Arg
1 5 10 15
His Thr Val Arg Asp Lys Leu Arg Leu Cys Gln Asn Phe Leu Gln Lys
20 25 30
Leu Arg Phe Leu Ala Asp Glu Pro Gln His Ser Ile Pro Asp Ile Phe
35 40 45
Ile Trp Met Met Ser Asn Asn Lys Arg Val Ala Tyr Ala Arg Val Pro
50 55 60
Ser Lys Asp Leu Leu Phe Ser Ile Val Glu Glu Glu Thr Gly Lys Asp
65 70 75 80
Cys Ala Lys Val Lys Thr Leu Phe Leu Lys Leu Pro Gly Lys Arg Gly
85 90 95
Phe Gly Ser Ala Gly Trp Thr Val Gln Ala Lys Val Glu Leu Tyr Leu
100 105 110
Trp Leu Gly Leu Ser Lys Gln Arg Lys Glu Phe Leu Cys Gly Leu Pro
115 120 125
Cys Gly Phe Gln Glu Val Lys Ala Ala Gln Gly Leu Gly Leu His Ala

130 135 140

Phe Pro Pro Val Ser Leu Val Tyr Thr Lys Lys Gln Ala Phe Gln Leu
145 150 155 160
Arg Ala His Met Tyr Gln Ala Arg Ser Leu Phe Ala Ala Asp Ser Ser
165 170 175
Gly Leu Ser Asp Pro Phe Ala Arg Val Phe Phe Ile Asn Gln Ser Gln
180 185 190
Cys Thr Glu Val Leu Asn Glu Thr Leu Cys Pro Thr Trp Asp Gln Met
195 200 205
Leu Val Phe Asp Asn Leu Glu Leu Tyr Gly Glu Ala His Glu Leu Arg
210 215 220
Asp Asp Pro Pro Ile Ile Val Ile Glu Ile Tyr Asp Gln Asp Ser Met
225 230 235 240
Gly Lys Ala Asp Phe Met Gly Arg Thr Phe Ala Lys Pro Leu Val Lys
245 250 255
Met Ala Asp Glu Ala Tyr Cys Pro Pro Arg Phe Pro Pro Gln Leu Glu
260 265 270
Tyr Tyr Gln Ile Tyr Arg Gly Asn Ala Thr Ala Gly Asp Leu Leu Ala
275 280 285
Ala Phe Glu Leu Leu Gln Ile Gly Pro Ala Gly Lys Ala Asp Leu Pro
290 295 300
Pro Ile Asn Gly Pro Val Asp Val Asp Arg Gly Pro Ile Met Pro Val
305 310 315 320
Pro Met Gly Ile Arg Pro Val Leu Ser Lys Tyr Arg Val Glu Val Leu
325 330 335
Phe Trp Gly Leu Arg Asp Leu Lys Arg Val Asn Leu Ala Gln Val Asp
340 345 350
Arg Pro Arg Val Asp Ile Glu Cys Ala Gly Lys Gly Val Gln Ser Ser
355 360 365
Leu Ile His Asn Tyr Lys Lys Asn Pro Asn Phe Asn Thr Leu Val Lys
370 375 380
Trp Phe Glu Val Asp Leu Pro Glu Asn Glu Leu Leu His Pro Pro Leu
385 390 395 400
Asn Ile Arg Val Val Asp Cys Arg Ala Phe Gly Arg Tyr Thr Leu Val
405 410 415
Gly Ser His Ala Val Ser Ser Leu Arg Arg Phe Ile Tyr Arg Pro Pro
420 425 430
Asp Arg Ser Ala Pro Ser Trp Asn Thr Thr Val Arg Leu Leu Arg Arg
435 440 445

Cys Arg Val Leu Cys Asn Gly Gly Ser Ser Ser His Ser Thr Gly Glu
450 455 460
Val Val Val Thr Met Glu Pro Glu Val Pro Ile Lys Lys Leu Glu Thr
465 470 475 480
Met Val Lys Leu Asp Ala Thr Ser Glu Ala Val Val Lys Val Asp Val
485 490 495
Ala Glu Glu Glu Lys Glu Lys Lys Lys Lys Lys Lys Gly Thr Ala Glu
500 505 510
Glu Pro Glu Glu Glu Glu Pro Asp Glu Ser Met Leu Asp Trp Trp Ser
515 520 525
Lys Tyr Phe Ala Ser Ile Asp Thr Met Lys Glu Gln Leu Arg Gln Gln
530 535 540
Glu Pro Ser Gly Ile Asp Leu Glu Glu Lys Glu Glu Val Asp Asn Thr
545 550 555 560
Glu Gly Leu Lys Gly Ser Met Lys Gly Lys Glu Lys Ala Arg Ala Ala
565 570 575
Lys Glu Glu Lys Lys Lys Lys Thr Gln Ser Ser Gly Ser Gly Gln Gly
580 585 590
Ser Glu Ala Pro Glu Lys Lys Lys Pro Lys Ile Asp Glu Leu Lys Val
595 600 605
Tyr Pro Lys Glu Leu Glu Ser Glu Phe Asp Asn Phe Glu Asp Trp Leu
610 615 620
His Thr Phe Asn Leu Leu Arg Gly Lys Thr Gly Asp Asp Glu Asp Gly
625 630 635 640
Ser Thr Glu Glu Glu Arg Ile Val Gly Arg Phe Lys Gly Ser Leu Cys
645 650 655
Val Tyr Lys Val Pro Leu Pro Glu Asp Val Ser Arg Glu Ala Gly Tyr
660 665 670
Asp Ser Thr Tyr Gly Met Phe Gln Gly Ile Pro Ser Asn Asp Pro Ile
675 680 685
Asn Val Leu Val Arg Val Tyr Val Val Arg Ala Thr Asp Leu His Pro
690 695 700
Ala Asp Ile Asn Gly Lys Ala Asp Pro Tyr Ile Ala Ile Arg Leu Gly
705 710 715 720
Lys Thr Asp Ile Arg Asp Lys Glu Asn Tyr Ile Ser Lys Gln Leu Asn
725 730 735
Pro Val Phe Gly Lys Ser Phe Asp Ile Glu Ala Ser Phe Pro Met Glu
740 745 750
Ser Met Leu Thr Val Ala Val Tyr Asp Trp Asp Leu Val Gly Thr Asp

755 760 765

Asp Leu Ile Gly Glu Thr Lys Ile Asp Leu Glu Asn Arg Phe Tyr Ser
770 775 780

Lys His Arg Ala Thr Cys Gly Ile Ala Gln Thr Tyr Ser Thr His Gly
785 790 795 800

Tyr Asn Ile Trp Arg Asp Pro Met Lys Pro Ser Gln Ile Leu Thr Arg
805 810 815

Leu Cys Lys Asp Gly Lys Val Asp Gly Pro His Phe Gly Pro Pro Gly
820 825 830

Arg Val Lys Val Ala Asn Arg Val Phe Thr Gly Pro Ser Glu Ile Glu
835 840 845

Asp Glu Asn Gly Gln Arg Lys Pro Thr Asp Glu His Val Ala Leu Leu
850 855 860

Ala Leu Arg His Trp Glu Asp Ile Pro Arg Ala Gly Cys Arg Leu Val
865 870 875 880

Pro Glu His Val Glu Thr Arg Pro Leu Leu Asn Pro Asp Lys Pro Gly
885 890 895

Ile Glu Gln Gly Arg Leu Glu Leu Trp Val Asp Met Phe Pro Met Asp
900 905 910

Met Pro Ala Pro Gly Thr Pro Leu Asp Ile Ser Pro Arg Lys Pro Lys
915 920 925

Lys Tyr Glu Leu Arg Val Ile Ile Trp Asn Thr Asp Glu Val Val Leu
930 935 940

Glu Asp Asp Asp Phe Phe Thr Gly Glu Lys Ser Ser Asp Ile Phe Val
945 950 955 960

Arg Gly Trp Leu Lys Gly Gln Gln Glu Asp Lys Gln Asp Thr Asp Val
965 970 975

His Tyr His Ser Leu Thr Gly Glu Gly Asn Phe Asn Trp Arg Tyr Leu
980 985 990

Phe Pro Phe Asp Tyr Leu Ala Ala Glu Glu Lys Ile Val Ile Ser Lys
995 1000 1005

Lys Glu Ser Met Phe Ser Trp Asp Glu Thr Glu Tyr Lys Ile Pro
1010 1015 1020

Ala Arg Leu Thr Leu Gln Ile Trp Asp Ala Asp His Phe Ser Ala
1025 1030 1035

Asp Asp Phe Leu Gly Ala Ile Glu Leu Asp Leu Asn Arg Phe Pro
1040 1045 1050

Arg Gly Ala Lys Thr Ala Lys Gln Cys Thr Met Glu Met Ala Thr
1055 1060 1065

Gly Glu Val Asp Val Pro Leu Val Ser Ile Phe Lys Gln Lys Arg
1070 1075 1080
Val Lys Gly Trp Trp Pro Leu Leu Ala Arg Asn Glu Asn Asp Glu
1085 1090 1095
Phe Glu Leu Thr Gly Lys Val Glu Ala Glu Leu His Leu Leu Thr
1100 1105 1110
Ala Glu Glu Ala Glu Lys Asn Pro Val Gly Leu Ala Arg Asn Glu
1115 1120 1125
Pro Asp Pro Leu Glu Lys Pro Asn Arg Pro Asp Thr Ala Phe Val
1130 1135 1140
Trp Phe Leu Asn Pro Leu Lys Ser Ile Lys Tyr Leu Ile Cys Thr
1145 1150 1155
Arg Tyr Lys Trp Leu Ile Ile Lys Ile Val Leu Ala Leu Leu Gly
1160 1165 1170
Leu Leu Met Leu Gly Leu Phe Leu Tyr Ser Leu Pro Gly Tyr Met
1175 1180 1185
Val Lys Lys Leu Leu Gly Ala
1190 1195

<210> 75
<211> 10032
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 75

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gatagaggtc 60
atccttcctg accatttcca tcattccagt cgaactcaca cacaacacca aatgcattta 120
agtcgcttga aattgctata agcagagcat gttgcgccag catgattaat acagcattta 180
atacagagcc gtgtttattg agtcggtatt cagagtctga ccagaaatta ttaatctggt 240
gaagttattc ctctgtcatt acgtcatggt cgatttcaat ttctattgat gctttccagt 300
cgtaatcaat gatgtatttt ttgatgtttg acctctgttc atatcctcac agataaaaaa 360
tcgccctcac actggagggc aaagaagatt tccaataatc agaacaagtc ggctcctgtt 420
tagttacgag cgacattgct ccgtgtattc actcgttgga atgaatacac agtgcagtgt 480
ttattctgtt atttatgcca aaaattaagg ccactatcag gcagctttgt tgttctgttt 540
accaagttct ctggcaatca ttgccgtcgt tcgtattgcc catttatcga catatttccc 600
atcttcctat acaggaaaca tttcttcagg cttaaccatg cattccgatt gcagcttgca 660
tccattgcat cgcttgaatt gtccacacca ttgattttta tcaatagtcg tagtttaacg 720
gatagtcctg gtattgttcc atcacatcct gaggatgccc ttcgaactct tcaaattctt 780
cttcctaata tcaccttaaa tagtggattg cggtagtaaa gattgtgcct gtcttttaac 840

cacatcaggc tcggtggttc tcgtgtaccc ctacagcgag aaatcggata aactattaca 900
acccctacag tttgtagagt atagaaaatg atccactcgt tattctcgga cgagtgttca 960
gtaatgaacc tctggagaga accatctata tgatcgttat ctgggtttga cttctgcttt 1020
taagcccaga taacttgcct gaatatgtta atgagagaat cggtattcct catgtgtggc 1080
atgttttcgt ctttgctctt gcattttcac tagcaattaa tgtgcatcga ttatcagcta 1140
ttgccagcgc cagatataag cgatttaagc taagaaaacg cattaaggtg caaaacgata 1200
aagtgcgatc agtaattcaa aaccttacag gagagcaatc tatggttttg tgctcagccc 1260
ttaatgaagg caggtagtat gtggttacat caaaacaatt cccatacatt agtgagttga 1320
ttgagcttgg tgtgttgaac aaaacttttt cccgatggaa tggaaagcat atattattcc 1380
ctattgagga tatttactgg actgaattag ttgccagcta tgatccatat aatattgaga 1440
taaagccaag gccaatatct aagtaactag ataagaggaa tcgattttcc cttaattttc 1500
tggcgtccac tgcatgttat gccgcgttcg ccaggcttgc tgtaccatgt gcgctgattc 1560
ttgcgctcaa tacgttgcag gttgctttca atctgtttgt ggtattcagc cagcactgta 1620
aggtctatcg gatttagtgc gctttctact cgtgatttcg gtttgcgatt cagcgagaga 1680
atagggcggt taactggttt tgcgcttacc ccaaccaaca ggggatttgc tgctttccat 1740
tgagcctgtt actctgcgcg acgttcgcgg cggcgtgttt gtgcatccat ctggattctc 1800
ctgtcagtta gctttggtgg tgtgtggcag ttgtagtcct gaacgaaaac cccccgcgat 1860
tggcacgttg gcagctaatc cggaatcgca cttacggcca atgcttcgtt tcgtatcaca 1920
caccccaaag ccttctgctt tgaatgctgc ccttcttcag ggcttaattt ttaagagcgt 1980
caccttcatg gtggtcagtg cgtcctgctg atgtgctcag gcacgattta attaaggcct 2040
taattaggct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg 2100
cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg cagagaggga gtggccaact 2160
ccatcactag gggttccttg tagttaatga ttaacccgcc atgctactta tctacgtagc 2220
catgctctag gaagatcgga attcgcccct aagctagcgg cgcgcccaat tctgcagctc 2280
agcctactac ttgctttcca ggctgttcct agttcccatg tcagctgctt gtgctttcca 2340
gagacaaaac aggaataata gatgtcatta aatatacatt gggccccagg cggtcaatgt 2400
ggcagcctga gcctcctttc catctctgtg gaggcagaca taggaccccc aacaaacagc 2460
atgcaggttg ggagccagcc acaggaccca ggtaaggggc cctgggtcct taagcttctg 2520
ccactggctc cggcattgca gagagaagag aaggggcggc agactggaga gctgggctcc 2580
atttttgttc cttggtgccc tgcccctccc catgacctgc agagacattc agcctgccag 2640
gctttatgag gtgggagctg ggctctccct gatgtattat tcagctccct ggagttggcc 2700
agctcctgtt acactggcca cagccctggg catccgcttc tcacttctag tttcccctcc 2760
aaggtaatgt ggtgggtcat gatcattcta tcctggcttc agggacctga ctccactttg 2820
gggccattcg aggggtctag ggtagatgat gtccccctgt ggggattaat gtcctgctct 2880
gtaaaactga gctagctgag atccaggagg gcttggccag agacagcaag ttgttgccat 2940
ggtgacttta aagccaggtt gctgccccag cacaggcctc ccagtctacc ctcactagaa 3000
aacaacaccc aggcactttc caccacctct caaaggtgaa acccaaggct ggtctagaga 3060
atgaattatg gatcctcgct gtccgtgcca cccagctagt cccagcggct cagacactga 3120
ggagagactg taggttcagc tacaagcaaa aagacctagc tggtctccaa gcagtgtctc 3180

caagtccctg aacctgtgac acctgcccca ggcatcatca ggcacagagg gccaccaaga 3240
attctagcgg ccgccaccat ggccttgctc atccacctca agacagtctc ggagctgcgg 3300
ggcaggggcg accggatcgc caaagtgact ttccgagggc aatccttcta ctctcgggtc 3360
ctggagaact gtgaggatgt ggctgacttt gatgagacat ttcggtggcc ggtggccagc 3420
agcatcgaca gaaatgagat gctggagatt caggttttca actacagcaa agtcttcagc 3480
aacaagctca tcgggacctt ccgcatggtg ctgcagaagg tggtagagga gagccatgtg 3540
gaggtgactg acacgctgat tgatgacaac aatgctatca tcaagaccag cctgtgcgtg 3600
gaggtccggt atcaggccac tgacggcaca gtgggctcct gggacgatgg ggacttcctg 3660
ggagatgagt ctcttcaaga ggaagagaag gacagccaag agacggatgg actgctccca 3720
ggctcccggc ccagctcccg gcccccagga gagaagagct tccggagagc cgggaggagc 3780
gtgttctccg ccatgaagct cggcaaaaac cggtctcaca aggaggagcc ccaaagacca 3840
gatgaaccgg cggtgctgga gatggaagac cttgaccatc tggccattcg gctaggagat 3900
ggactggatc ccgactcggt gtctctagcc tcagtcacag ctctcaccac taatgtctcc 3960
aacaagcgat ctaagccaga cattaagatg gagccaagtg ctgggcggcc catggattac 4020
caggtcagca tcacggtgat cgaggcccgg cagctggtgg gcttgaacat ggaccctgtg 4080
gtgtgcgtgg aggtgggtga cgacaagaag tacacatcca tgaaggagtc cactaactgc 4140
ccctattaca acgagtactt cgtcttcgac ttccatgtct ctccggatgt catgtttgac 4200
aagatcatca agatttcggt gattcactcc aagaacctgc tgcgcagtgg caccctggtg 4260
ggctccttca aaatggacgt gggaaccgtg tactcgcagc cagagcacca gttccatcac 4320
aagtgggcca tcctgtctga ccccgatgac atctcctcgg ggctgaaggg ctacgtgaag 4380
tgtgacgttg ccgtggtggg caaaggggac aacatcaaga cgccccacaa ggccaatgag 4440
accgacgaag atgacattga ggggaacttg ctgctccccg agggggtgcc ccccgaacgc 4500
cagtgggccc ggttctatgt gaaaatttac cgagcagagg ggctgccccg tatgaacaca 4560
agcctcatgg ccaatgtaaa gaaggctttc atcggtgaaa acaaggacct cgtggacccc 4620
tacgtgcaag tcttctttgc tggccagaag ggcaagactt cagtgcagaa gagcagctat 4680
gagcccctgt ggaatgagca ggtcgtcttt acagacctct tccccccact ctgcaaacgc 4740
atgaaggtgc agatccgaga ctcggacaag gtcaacgacg tggccatcgg cacccacttc 4800
attgacctgc gcaagatttc taatgacgga gacaaaggct tcctgcccac actgggccca 4860
gcctgggtga acatgtacgg ctccacacgt aactacacgc tgctggatga gcatcaggac 4920
ctgaacgagg gcctgggggа gggtgtgtcc ttccgggccc ggctcctgct gggcctggct 4980
gtggagatcg tagacacctc caaccctgag ctcaccagct ccacagaggt gcaggtggag 5040
caggccacgc ccatctcgga gagctgtgca ggtaaaatgg aagaattctt tctctttgga 5100
gccttcctgg aggcctcaat gatcgaccgg agaaacggag acaagcccat cacctttgag 5160
gtcaccatag gcaactatgg gaacgaagtt gatggcctgt cccggcccca gcggcctcgg 5220
ccccggaagg agccgggggа tgaggaagaa gtagacctga ttcagaacgc aagtgatgac 5280
gaggccggtg atgccgggga cctggcctca gtctcctcca ctccaccaat gcggccccag 5340
gtcaccgaca ggaactactt ccatctgccc tacctggagc gaaagccctg catctacatc 5400
aagagctggt ggccggacca gcgccgccgc ctctacaatg ccaacatcat ggaccacatt 5460
gccgacaagc tggaagaagg cctgaacgac atacaggaga tgatcaaaac ggagaagtcc 5520

taccctgagc gtcgcctgcg gggcgtcctg gaggagctga gctgtggctg ctgccgcttc 5580
ctctccctcg ctgacaagga ccagggccac tcatcccgca ccaggcttga ccgggagcgc 5640
ctcaagtcct gcatgaggga gctggtaagt atcaaggtta caagacaggt ttaaggagac 5700
caatagaaac tgggcttgtc gagacagaga agactcttgc gtttctgagc tagcccccgg 5760
gtgcgcggcg tcggtggtgc cggcgggggg cgccaggtcg caggcggtgt agggctccag 5820
gcaggcggcg aaggccatga cgtgcgctat gaaggtctgc tcctgcacgc cgtgaaccag 5880
gtgcgcctgc gggccgcgcg cgaacaccgc cacgtcctcg cctgcgtggg tctcttcgtc 5940
caggggcact gctgactgct gccgatactc ggggctcccg ctctcgctct cggtaacatc 6000
cggccgggcg ccgtccttga gcacatagcc tggaccgttt cgtcgacctc gagttaaggg 6060
cgaattcccg ataaggatct tcctagagca tggctacgta gataagtagc atggcgggtt 6120
aatcattaac tacaaggaac ccctagtgat ggagttggcc actccctctc tgcgcgctcg 6180
ctcgctcact gaggccgggc gaccaaaggt cgcccgacgc ccgggctttg cccgggcggc 6240
ctcagtgagc gagcgagcgc gcagccttaa ttaaatccac atctgtatgt tttttatatt 6300
aatttatttt ttgcaggggg gcattgtttg gtaggtgaga gttctgaatt gctatgttta 6360
gtgagttgta tctatttatt tttcaataaa tacaattagt tatgtgtttt gggggcgatc 6420
gtgaggcaaa gaaaacccgg cgctgaggcc gggttattct tgttctctgg tcaaattata 6480
tagttggaaa acaaggatgc atatatgaat gaacgatgca gaggcaatgc cgatggcgat 6540
agtgggtatc aggtagccgc ttatgctgga aagaagcaat aacccgcaga aaaacaaagc 6600
tccaagctca acaaaactaa gggcatagac aataactacc tatgtcatat acccatactc 6660
tctaatcttg gccagtcggc gcgttctgct tccgattaga aacgtcaagg cagcaatcag 6720
gattgcaatc ttggttcctg cataggatga caatgtcgcc ccaagaccat ctctatgagc 6780
tgaaaaagaa acacaaggaa tgtagtggcg gaaaaggaga tagcaaatgc ttacgataac 6840
gtaaggaatt attactatgt aaacaccagg caagattctg ttccgtataa ttactcctga 6900
taattaatcc ttaactttgc ccacctgcct tttaaaacat tccagtatat cacttttcat 6960
tcttgcgtag caatatgccc tctcttcagc tatctcagca ttggtgacct tgttcagagg 7020
cgctgagaga tggccttttt ctgatagata atgttctgtt aaaatatctc cggcctcatc 7080
ttttgcccgc aggctaatgt ctgaaaattg aggtgacggg ttaaaaataa tatccttggc 7140
aacctttttt atatcccttt taaattttgg cttaatgact atatccaatg agtcaaaaag 7200
ctcccccttca atatctgttg cccctaagac ctttaatata tcgccaaata caggtagctt 7260
ggcttctacc ttcaccgttg ttctgccgat gaaatgctaa tgcataacat cgtctttggt 7320
ggttcccctc atcagtggct ctatctgaac gcgctctcca ctgcttaatg acattccttt 7380
cccgattaaa aaatctgtca gatcggatgt ggtcggcccg aaaacagttc tggcaaaacc 7440
aatggtgtcg ccttcaacaa acaaaaaaga tgggaatccc aatgattcgt catctgcgag 7500
gctgttctta atatcttcaa ctgtagcttt agagcgattt atcttctgaa ccagactctt 7560
gtcatttgtt ttggtaaaga gaaaagtttt tccatcgatt ttatgaatat acaaataatt 7620
ggagccaacc ttcaggtgat gattatcagc cagcagagaa ttaaggaaaa cagacaggtt 7680
tattgagcac ttatctttcc ctttattttt gctgcggtaa gtcgcataaa aaccattctt 7740
cacaattcaa tccatttact atgttatgtt ctgaggggag tgaaaattcc cctaattcga 7800
tgaagattct tgctaaattg ttatcagcta tgcgccgacc agaacacctt gccgatcagc 7860

caaacgtcta atcaggccac tgactagcga taactttccc cacaacggaa caactctcat 7920
tgcatgggat aattgggtac tgtgggttta gtggttgtaa aaacacctga ccgctatccc 7980
tgatcagttt cttgaaggta aactcatcac ccccaagtct ggctatacag aaatcacctg 8040
gctcaacagc ctgctcaggg tcaacgagaa tttacattcc gtcaggatag cttggcttgg 8100
agcctgttgg tgcggtcacg gaattacctt caacctcaag ccagaatgca gaatcactgg 8160
ctttttggt tgtgcttacc catctctccg catcaccttt ggtaaaggtt ctaagctaag 8220
gtgagaacat ccctgcctga acatgagaaa aaacagggta ctcatactca cttattagtg 8280
acggctatga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc 8340
gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag 8400
gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt 8460
gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg 8520
aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt aggtcgttcg 8580
ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 8640
taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac 8700
tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg 8760
gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc tgaagccagt 8820
taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 8880
tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 8940
tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt 9000
ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 9060
taaatcaagc ccaatctgaa taatgttaca accaattaac caattctgat tagaaaaact 9120
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 9180
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 9240
gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 9300
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 9360
agaatggcaa aagtttatgc atttctttcc agacttgttc aacaggccag ccattacgct 9420
cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc gcctgagcaa 9480
gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc 9540
gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat tcttctaata 9600
cctggaatgc tgtttttccg gggatcgcag tggtgagtaa ccatgcatca tcaggagtac 9660
ggataaaatg cttgatggtc ggaagaggca taaattccgt cagccagttt agtctgacca 9720
tctcatctgt aacatcattg gcaacgctac ctttgccatg tttcagaaac aactctggcg 9780
catcgggctt cccatacaag cgatagattg tcgcacctga ttgcccgaca ttatcgcgag 9840
cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc ctcgacgttt 9900
cccgttgaat atggctcata acacccccttg tattactgtt tatgtaagca gacagtttta 9960
ttgttcatga tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacgg 10020
gccagagctg ca 10032

<210> 76

<211> 10461

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 76

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gatagaggtc 60
atccttcctg accatttcca tcattccagt cgaactcaca cacaacacca aatgcattta 120
agtcgcttga aattgctata agcagagcat gttgcgccag catgattaat acagcattta 180
atacagagcc gtgtttattg agtcggtatt cagagtctga ccagaaatta ttaatctggt 240
gaagttattc ctctgtcatt acgtcatggt cgatttcaat ttctattgat gctttccagt 300
cgtaatcaat gatgtatttt ttgatgtttg acctctgttc atatcctcac agataaaaaa 360
tcgccctcac actggagggc aaagaagatt tccaataatc agaacaagtc ggctcctgtt 420
tagttacgag cgacattgct ccgtgtattc actcgttgga atgaatacac agtgcagtgt 480
ttattctgtt atttatgcca aaaattaagg ccactatcag gcagctttgt tgttctgttt 540
accaagttct ctggcaatca ttgccgtcgt tcgtattgcc catttatcga catatttccc 600
atcttcctat acaggaaaca tttcttcagg cttaaccatg cattccgatt gcagcttgca 660
tccattgcat cgcttgaatt gtccacacca ttgattttta tcaatagtcg tagtttaacg 720
gatagtcctg gtattgttcc atcacatcct gaggatgccc ttcgaactct tcaaattctt 780
cttcctaata tcaccttaaa tagtggattg cggtagtaaa gattgtgcct gtcttttaac 840
cacatcaggc tcggtggttc tcgtgtaccc ctacagcgag aaatcggata aactattaca 900
acccctacag tttgtagagt atagaaaatg atccactcgt tattctcgga cgagtgttca 960
gtaatgaacc tctggagaga accatctata tgatcgttat ctgggtttga cttctgcttt 1020
taagcccaga taacttgcct gaatatgtta atgagagaat cggtattcct catgtgtggc 1080
atgttttcgt ctttgctctt gcattttcac tagcaattaa tgtgcatcga ttatcagcta 1140
ttgccagcgc cagatataag cgatttaagc taagaaaacg cattaaggtg caaaacgata 1200
aagtgcgatc agtaattcaa aaccttacag gagagcaatc tatggttttg tgctcagccc 1260
ttaatgaagg caggtagtat gtggttacat caaaacaatt cccatacatt agtgagttga 1320
ttgagcttgg tgtgttgaac aaaacttttt cccgatggaa tggaaagcat atattattcc 1380
ctattgagga tatttactgg actgaattag ttgccagcta tgatccatat aatattgaga 1440
taaagccaag gccaatatct aagtaactag ataagaggaa tcgattttcc cttaattttc 1500
tggcgtccac tgcatgttat gccgcgttcg ccaggcttgc tgtaccatgt gcgctgattc 1560
ttgcgctcaa tacgttgcag gttgctttca atctgtttgt ggtattcagc cagcactgta 1620
aggtctatcg gatttagtgc gctttctact cgtgatttcg gtttgcgatt cagcgagaga 1680
atagggcggt taactggttt tgcgcttacc ccaaccaaca ggggatttgc tgctttccat 1740
tgagcctgtt actctgcgcg acgttcgcgg cggcgtgttt gtgcatccat ctggattctc 1800
ctgtcagtta gctttggtgg tgtgtggcag ttgtagtcct gaacgaaaac cccccgcgat 1860
tggcacgttg gcagctaatc cggaatcgca cttacggcca atgcttcgtt tcgtatcaca 1920
caccccaaag ccttctgctt tgaatgctgc ccttcttcag ggcttaattt ttaagagcgt 1980
caccttcatg gtggtcagtg cgtcctgctg atgtgctcag gcacgattta attaaggcct 2040

taattaggct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg 2100
cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg cagagaggga gtggccaact 2160
ccatcactag gggttccttg tagttaatga ttaacccgcc atgctactta tctacgtagc 2220
catgctctag gaagatcgga attcgcccct aagctagcgg cgcgcccccc gggtgcgcgg 2280
cgtcggtggt gccggcgggg ggcgccaggt cgcaggcggt gtagggctcc aggcaggcgg 2340
cgaaggccat gacgtgcgct atgaaggtct gctcctgcac gccgtgaacc aggtgcgcct 2400
gcgggccgcg cgcgaacacc gccacgtcct cgcctgcgtg ggtctcttcg tccaggggca 2460
ctgctgactg ctgccgatac tcggggctcc cgctctcgct ctcggtaaca tccggccggg 2520
cgccgtcctt gagcacatag cctggaccgt ttccttaagc gacgcatgct cgcgataggc 2580
acctattggt cttactgaca tccactttgc ctttctctcc acaggaaaac atggggcagc 2640
aggccaggat gctgcgggcc caggtgaagc ggcacacggt gcgggacaag ctgaggctgt 2700
gccagaactt cctgcagaag ctgcgcttcc tggcggacga gccccagcac agcattcccg 2760
acatcttcat ctggatgatg agcaacaaca agcgtgtcgc ctatgcccgt gtgccctcca 2820
aggacctgct cttctccatc gtggaggagg agactggcaa ggactgcgcc aaggtcaaga 2880
cgctcttcct taagctgcca gggaagcggg gcttcggctc ggcaggctgg acagtgcagg 2940
ccaaggtgga gctgtacctg tggctgggcc tcagcaaaca gcgcaaggag ttcctgtgcg 3000
gcctgccctg tggcttccag gaggtcaagg cagcccaggg cctgggcctg catgccttcc 3060
cacccgtcag cctggtctac accaagaagc aggcgttcca gctccgagcg cacatgtacc 3120
aggcccgcag cctctttgcc gccgacagca gcggactctc agaccccttt gcccgcgtct 3180
tcttcatcaa tcagagtcag tgcacagagg tgctgaatga gaccctgtgt cccacctggg 3240
accagatgct ggtgttcgac aacctggagc tctatggtga agctcatgag ctgagggacg 3300
atccgcccat cattgtcatt gaaatctatg accaggattc catgggcaaa gctgacttca 3360
tgggccggac cttcgccaaa cccctggtga agatggcaga cgaggcgtac tgcccacccc 3420
gcttcccacc tcagctcgag tactaccaga tctaccgtgg caacgccaca gctggagacc 3480
tgctggcggc cttcgagctg ctgcagattg gaccagcagg gaaggctgac ctgccccccа 3540
tcaatggccc ggtggacgtg gaccgaggtc ccatcatgcc cgtgcccatg ggcatccggc 3600
ccgtgctcag caagtaccga gtggaggtgc tgttctgggg cctacgggac ctaaagcggg 3660
tgaacctggc ccaggtggac cggccacggg tggacatcga gtgtgcaggg aaggggtgc 3720
agtcgtccct gatccacaat tataagaaga accccaactt caacaccctc gtcaagtggt 3780
ttgaagtgga cctcccagag aacgagctgc tgcacccgcc cttgaacatc cgtgtggtgg 3840
actgccgggc cttcggtcgc tacacactgg tgggctccca tgccgtcagc tccctgcgac 3900
gcttcatcta ccggccccca gaccgctcgg cccccagctg gaacaccacg gtcaggcttc 3960
tccggcgctg ccgtgtgctg tgcaatgggg gctcctcctc tcactccaca ggggaggttg 4020
tggtgactat ggagccagag gtacccatca agaaactgga gaccatggtg aagctggacg 4080
cgacttctga agctgttgtc aaggtggatg tggctgagga ggagaaggag aagaagaaga 4140
agaagaaggg cactgcggag gagccagagg aggaggagcc agacgagagc atgctggact 4200
ggtggtccaa gtactttgcc tccattgaca ccatgaagga gcaacttcga caacaagagc 4260
cctctggaat tgacttggag gagaaggagg aagtggacaa taccgagggc ctgaaggggt 4320
caatgaaggg caaggagaag gcaagggctg ccaaagagga gaagaagaag aaaactcaga 4380

gctctggctc tggccagggg tccgaggccc ccgagaagaa gaaacccaag attgatgagc 4440
ttaaggtata ccccaaagag ctggagtccg agtttgataa ctttgaggac tggctgcaca 4500
ctttcaactt gcttcggggc aagaccgggg atgatgagga tggctccacc gaggaggagc 4560
gcattgtggg acgcttcaag ggctccctct gcgtgtacaa agtgccactc ccagaggacg 4620
tgtcccggga agccggctac gactccacct acggcatgtt ccagggcatc ccgagcaatg 4680
accccatcaa tgtgctggtc cgagtctatg tggtccgggc cacggacctg caccctgctg 4740
acatcaacgg caaagctgac ccctacatcg ccatccggct aggcaagact gacatccgcg 4800
acaaggagaa ctacatctcc aagcagctca accctgtctt tgggaagtcc tttgacatcg 4860
aggcctcctt ccccatggaa tccatgctga cggtggctgt gtatgactgg gacctggtgg 4920
gcactgatga cctcattggg gaaaccaaga tcgacctgga gaaccgcttc tacagcaagc 4980
accgcgccac ctgcggcatc gcccagacct actccacaca tggctacaat atctggcggg 5040
accccatgaa gcccagccag atcctgaccc gcctctgcaa agacggcaaa gtggacggcc 5100
cccactttgg gccccctggg agagtgaagg tggccaaccg cgtcttcact gggccctctg 5160
agattgagga cgagaacggt cagaggaagc ccacagacga gcatgtggcg ctgttggccc 5220
tgaggcactg ggaggacatc ccccgcgcag gctgccgcct ggtgccagag catgtggaga 5280
cgaggccgct gctcaacccc gacaagccgg gcatcgagca gggccgcctg gagctgtggg 5340
tggacatgtt ccccatggac atgccagccc ctgggacgcc tctggacatc tcacctcgga 5400
agcccaagaa gtacgagctg cgggtcatca tctggaacac agatgaggtg gtcttggagg 5460
acgacgactt cttcacaggg gagaagtcca gtgacatctt cgtgaggggg tggctgaagg 5520
gccagcagga ggacaagcag gacacagacg tccactacca ctccctcact ggcgagggca 5580
acttcaactg gcgctacctg ttccccttcg actacctggc ggcggaggag aagatcgtca 5640
tctccaagaa ggagtccatg ttctcctggg acgagaccga gtacaagatc cccgcgcggc 5700
tcaccctgca gatctgggat gcggaccact tctccgctga cgacttcctg ggggccatcg 5760
agctggacct gaaccggttc ccgcgggggcg caaagacagc caagcagtgc accatggaga 5820
tggccaccgg ggaggtggac gtgcccctcg tgtccatctt caagcaaaag cgcgtcaaag 5880
gctggtggcc cctcctggcc cgcaatgaga acgatgagtt tgagctcacg ggcaaggtgg 5940
aggctgagct gcatttactg acagcagagg aggcagagaa gaacccagtg ggcctggccc 6000
gcaatgaacc tgacccccta gagaaaccca accggcccga cacggccttc gtctggttcc 6060
tcaaccctct caagtccatc aagtacctca tctgcacccg gtacaagtgg ctcatcatca 6120
agatcgtgct ggcgctgttg gggctgctca tgttggggct cttcctctac agcctccctg 6180
gctacatggt caaaaagctc cttggggcat gaacggccgc tatgctagct tggtaccaag 6240
ggcggatcct gcatagagct cgctgatcag cctcgactgt gccttctagt tgccagccat 6300
ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact cccactgtcc 6360
tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat tctattctgg 6420
ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc aggcatctcg 6480
agttaagggc gaattcccga taaggatctt cctagagcat ggctacgtag ataagtagca 6540
tggcgggtta atcattaact acaaggaacc cctagtgatg gagttggcca ctccctctct 6600
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggctttgc 6660
ccgggcggcc tcagtgagcg agcgagcgcg cagccttaat taaatccaca tctgtatgtt 6720

ttttatatta atttatttt tgcaggggg cattgtttgg taggtgagag ttctgaattg 6780
ctatgtttag tgagttgtat ctatttattt ttcaataaat acaattagtt atgtgttttg 6840
ggggcgatcg tgaggcaaag aaaacccggc gctgaggccg ggttattctt gttctctggt 6900
caaattatat agttggaaaa caaggatgca tatatgaatg aacgatgcag aggcaatgcc 6960
gatggcgata gtgggtatca ggtagccgct tatgctggaa agaagcaata acccgcagaa 7020
aaacaaagct ccaagctcaa caaaactaag ggcatagaca ataactacct atgtcatata 7080
cccatactct ctaatcttgg ccagtcggcg cgttctgctt ccgattagaa acgtcaaggc 7140
agcaatcagg attgcaatct tggttcctgc ataggatgac aatgtcgccc caagaccatc 7200
tctatgagct gaaaaagaaa cacaaggaat gtagtggcgg aaaaggagat agcaaatgct 7260
tacgataacg taaggaatta ttactatgta aacaccaggc aagattctgt tccgtataat 7320
tactcctgat aattaatcct taactttgcc cacctgcctt ttaaaacatt ccagtatatc 7380
acttttcatt cttgcgtagc aatatgccct ctcttcagct atctcagcat tggtgacctt 7440
gttcagaggc gctgagagat ggccttttтc tgatagataa tgttctgtta aaatatctcc 7500
ggcctcatct tttgcccgca ggctaatgtc tgaaaattga ggtgacgggt taaaaataat 7560
atccttggca acctttttta tatccctttt aaattttggc ttaatgacta tatccaatga 7620
gtcaaaaagc tcccccttcaa tatctgttgc ccctaagacc tttaatatat cgccaaatac 7680
aggtagcttg gcttctacct tcaccgttgt tctgccgatg aaatgctaat gcataacatc 7740
gtctttggtg gttcccctca tcagtggctc tatctgaacg cgctctccac tgcttaatga 7800
cattcctttc ccgattaaaa aatctgtcag atcggatgtg gtcggcccga aaacagttct 7860
ggcaaaacca atggtgtcgc cttcaacaaa caaaaaagat gggaatccca atgattcgtc 7920
atctgcgagg ctgttcttaa tatcttcaac tgtagcttta gagcgattta tcttctgaac 7980
cagactcttg tcatttgttt tggtaaagag aaaagttttt ccatcgattt tatgaatata 8040
caaataattg gagccaacct tcaggtgatg attatcagcc agcagagaat taaggaaaac 8100
agacaggttt attgagcact tatctttccc tttatttttg ctgcggtaag tcgcataaaa 8160
accattcttc acaattcaat ccatttacta tgttatgttc tgaggggagt gaaaattccc 8220
ctaattcgat gaagattctt gctaaattgt tatcagctat gcgccgacca gaacaccttg 8280
ccgatcagcc aaacgtctaa tcaggccact gactagcgat aactttcccc acaacggaac 8340
aactctcatt gcatgggata attgggtact gtgggtttag tggttgtaaa aacacctgac 8400
cgctatccct gatcagtttc ttgaaggtaa actcatcacc cccaagtctg gctatacaga 8460
aatcacctgg ctcaacagcc tgctcagggt caacgagaat ttacattccg tcaggatagc 8520
ttggcttgga gcctgttggt gcggtcacgg aattaccttc aacctcaagc cagaatgcag 8580
aatcactggc ttttttggtt gtgcttaccc atctctccgc atcacctttg gtaaaggttc 8640
taagctaagg tgagaacatc cctgcctgaa catgagaaaa aacagggtac tcatactcac 8700
ttattagtga cggctatgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc 8760
gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc 8820
aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag 8880
ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct 8940
cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta 9000
ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc 9060

cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc 9120
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt 9180
gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct 9240
gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc 9300
tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca 9360
agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta 9420
agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa 9480
atgaagtttt aaatcaagcc caatctgaat aatgttacaa ccaattaacc aattctgatt 9540
agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac 9600
catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata 9660
ggatggcaag atcctggtat cggtctgcga ttccgactcg tccaacatca atacaaccta 9720
ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg 9780
aatccggtga gaatggcaaa agtttatgca tttctttcca gacttgttca acaggccagc 9840
cattacgctc gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg 9900
cctgagcaag acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaat 9960
gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt 10020
cttctaatac ctggaatgct gtttttccgg ggatcgcagt ggtgagtaac catgcatcat 10080
caggagtacg gataaaatgc ttgatggtcg gaagaggcat aaattccgtc agccagttta 10140
gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca 10200
actctggcgc atcgggcttc ccatacaagc gatagattgt cgcacctgat tgcccgacat 10260
tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc 10320
tcgacgtttc ccgttgaata tggctcataa caccccttgt attactgttt atgtaagcag 10380
acagttttat tgttcatgat gatatatttt tatcttgtgc aatgtaacat cagagatttt 10440
gagacacggg ccagagctgc a 10461

<210> 77
<211> 6678
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 77

gggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 60
gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 120
gcgcgcagag agggagtggc caactccatc actaggggtt cctcagatct gaattcggta 180
cctgcagctc agcctactac ttgctttcca ggctgttcct agttcccatg tcagctgctt 240
gtgctttcca gagacaaaac aggaataata gatgtcatta aatatacatt gggccccagg 300
cggtcaatgt ggcagcctga gcctcctttc catctctgtg gaggcagaca taggaccccc 360
aacaaacagc atgcaggttg ggagccagcc acaggaccca ggtaaggggc cctgggtcct 420
taagcttctg ccactggctc cggcattgca gagagaagag aaggggcggc agactggaga 480

gctgggctcc atttttgttc cttggtgccc tgcccctccc catgacctgc agagacattc 540
agcctgccag gctttatgag gtgggagctg ggctctccct gatgtattat tcagctccct 600
ggagttggcc agctcctgtt acactggcca cagccctggg catccgcttc tcacttctag 660
tttcccctcc aaggtaatgt ggtgggtcat gatcattcta tcctggcttc agggacctga 720
ctccactttg gggccattcg aggggtctag ggtagatgat gtccccctgt ggggattaat 780
gtcctgctct gtaaaactga gctagctgag atccaggagg gcttggccag agacagcaag 840
ttgttgccat ggtgacttta aagccaggtt gctgccccag cacaggcctc ccagtctacc 900
ctcactagaa aacaacaccc aggcactttc caccacctct caaaggtgaa acccaaggct 960
ggtctagaga atgaattatg gatcctcgct gtccgtgcca cccagctagt cccagcggct 1020
cagacactga ggagagactg taggttcagc tacaagcaaa aagacctagc tggtctccaa 1080
gcagtgtctc caagtccctg aacctgtgac acctgcccca ggcatcatca ggcacagagg 1140
gccaccgaat tctagcggcc gccaccatgg ccttgctcat ccacctcaag acagtctcgg 1200
agctgcgggg caggggcgac cggatcgcca aagtgacttt ccgagggcaa tccttctact 1260
ctcgggtcct ggagaactgt gaggatgtgg ctgactttga tgagacattt cggtggccgg 1320
tggccagcag catcgacaga aatgagatgc tggagattca ggttttcaac tacagcaaag 1380
tcttcagcaa caagctcatc gggaccttcc gcatggtgct gcagaaggtg gtagaggaga 1440
gccatgtgga ggtgactgac acgctgattg atgacaacaa tgctatcatc aagaccagcc 1500
tgtgcgtgga ggtccggtat caggccactg acggcacagt gggctcctgg gacgatgggg 1560
acttcctggg agatgagtct cttcaagagg aagagaagga cagccaagag acggatggac 1620
tgctcccagg ctcccggccc agctcccggc ccccaggaga gaagagcttc cggagagccg 1680
ggaggagcgt gttctccgcc atgaagctcg gcaaaaaccg gtctcacaag gaggagcccc 1740
aaagaccaga tgaaccggcg gtgctggaga tggaagacct tgaccatctg gccattcggc 1800
taggagatgg actggatccc gactcggtgt ctctagcctc agtcacagct ctcaccacta 1860
atgtctccaa caagcgatct aagccagaca ttaagatgga gccaagtgct gggcggccca 1920
tggattacca ggtcagcatc acggtgatcg aggcccggca gctggtgggc ttgaacatgg 1980
accctgtggt gtgcgtggag gtgggtgacg acaagaagta cacatccatg aaggagtcca 2040
ctaactgccc ctattacaac gagtacttcg tcttcgactt ccatgtctct ccggatgtca 2100
tgtttgacaa gatcatcaag atttcggtga ttcactccaa gaacctgctg cgcagtggca 2160
ccctggtggg ctccttcaaa atggacgtgg gaaccgtgta ctcgcagcca gagcaccagt 2220
tccatcacaa gtgggccatc ctgtctgacc ccgatgacat ctcctcgggg ctgaagggct 2280
acgtgaagtg tgacgttgcc gtggtgggca aaggggacaa catcaagacg ccccacaagg 2340
ccaatgagac cgacgaagat gacattgagg ggaacttgct gctccccgag ggggtgcccc 2400
ccgaacgcca gtgggcccgg ttctatgtga aaatttaccg agcagagggg ctgccccgta 2460
tgaacacaag cctcatggcc aatgtaaaga aggctttcat cggtgaaaac aaggacctcg 2520
tggaccccta cgtgcaagtc ttctttgctg gccagaaggg caagacttca gtgcagaaga 2580
gcagctatga gcccctgtgg aatgagcagg tcgtctttac agacctcttc cccccactct 2640
gcaaacgcat gaaggtgcag atccgagact cggacaaggt caacgacgtg gccatcggca 2700
cccacttcat tgacctgcgc aagatttcta atgacggaga caaaggcttc ctgcccacac 2760
tgggcccagc ctgggtgaac atgtacggct ccacacgtaa ctacacgctg ctggatgagc 2820

atcaggacct gaacgagggc ctgggggagg gtgtgtcctt ccgggcccgg ctcctgctgg 2880
gcctggctgt ggagatcgta gacacctcca accctgagct caccagctcc acagaggtgc 2940
aggtggagca ggccacgccc atctcggaga gctgtgcagg taaaatggaa gaattctttc 3000
tctttggagc cttcctggag gcctcaatga tcgaccggag aaacggagac aagcccatca 3060
cctttgaggt caccataggc aactatggga acgaagttga tggcctgtcc cggccccagc 3120
ggcctcggcc ccggaaggag ccgggggatg aggaagaagt agacctgatt cagaacgcaa 3180
gtgatgacga ggccggtgat gccggggacc tggcctcagt ctcctccact ccaccaatgc 3240
ggccccaggt caccgacagg aactacttcc atctgcccta cctggagcga aagccctgca 3300
tctacatcaa gagctggtgg ccggaccagc gccgccgcct ctacaatgcc aacatcatgg 3360
accacattgc cgacaagctg gaagaaggcc tgaacgacat acaggagatg atcaaaacgg 3420
agaagtccta ccctgagcgt cgcctgcggg gcgtcctgga ggagctgagc tgtggctgct 3480
gccgcttcct ctccctcgct gacaaggacc agggccactc atcccgcacc aggcttgacc 3540
gggagcgcct caagtcctgc atgagggagc tggtaagtat caaggttaca agacaggttt 3600
aaggagacca atagaaactg ggcttgtcga gacagagaag actcttgcgt ttctgagcta 3660
gcccccgggt gcgcggcgtc ggtggtgccg gcgggggcgg ccaggtcgca ggcggtgtag 3720
ggctccaggc aggcggcgaa ggccatgacg tgcgctatga aggtctgctc ctgcacgccg 3780
tgaaccaggt gcgcctgcgg gccgcgcgcg aacaccgcca cgtcctcgcc tgcgtgggtc 3840
tcttcgtcca ggggcactgc tgactgctgc cgatactcgg ggctcccgct ctcgctctcg 3900
gtaacatccg gccgggcgcc gtccttgagc acatagcctg gaccgtttcg tcgactgggg 3960
agagatctga ggaaccccta gtgatggagt tggccactcc ctctctgcgc gctcgctcgc 4020
tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag 4080
tgagcgagcg agcgcgcaga gagggagtgg ccaacccccc ccccccccccc cctgcagcct 4140
ggcgtaatag cgaagaggcc cgcaccgatc gcccttccca acagttgcgt agcctgaatg 4200
gcgaatggcg cgacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc 4260
gcagcgtgac cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt 4320
cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag 4380
ggttccgatt tagtgcttta cggcacctcg accccaaaaa acttgattag ggtgatggtt 4440
cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt 4500
tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc gcggtctatt 4560
cttttgattt ataagggatg ttgccgattt cggcctattg gttaaaaaat gagctgattt 4620
aacaaaaatt ttaacaaaat tcagaagaac tcgtcaagaa ggcgatagaa ggcgatgcgc 4680
tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca ttcgccgcca 4740
agctcttcag caatatcacg ggtagccaac gctatgtcct gatagcggtc cgccacaccc 4800
agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat attcggcaag 4860
caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gcatgctcgc cttgagcctg 4920
gcgaacagtt cggctggcgc gagcccctga tgctcttcgt ccagatcatc ctgatcgaca 4980
agaccggctt ccatccgagt acgtgctcgc tcgatgcgat gtttcgcttg gtggtcgaat 5040
gggcaggtag ccggatcaag cgtatgcagc cgccgcattg catcagccat gatggatact 5100
ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc gcccaatagc 5160

agccagtccc ttcccgcttc agtgacaacg tcgagcacag ctgcgcaagg aacgcccgtc 5220
gtggccagcc acgatagccg cgctgcctcg tcttgcagtt cattcagggc accggacagg 5280
tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca gccggaacac ggcggcatca 5340
gagcagccga ttgtctgttg tgcccagtca tagccgaata gcctctccac ccaagcggcc 5400
ggagaacctg cgtgcaatcc atcttgttca atcatgcgaa acgatcctca tcctgtctct 5460
tgatcagatc ttgatcccct gcgccatcag atccttggcg gcgagaaagc catccagttt 5520
actttgcagg gcttcccaac cttaccagag ggcgccccag ctggcaattc cggttcgctt 5580
gctgtccata aaaccgccca gtctagctat cgccatgtaa gcccactgca agctacctgc 5640
tttctctttg cgcttgcgtt ttcccttgtc cagatagccc agtagctgac attcatccgg 5700
ggtcagcacc gtttctgcgg actggctttc tacgtgaaaa ggatctaggt gaagatcctt 5760
tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 5820
cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc 5880
ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca 5940
actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta 6000
gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct 6060
ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg 6120
gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc 6180
acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta 6240
tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg 6300
gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt 6360
cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggg 6420
cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggg cttttgctgg 6480
ccttttgctc acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc 6540
gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg 6600
agcgaggaag cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt 6660
cattaatgca gggctgca 6678

<210> 78
<211> 7152
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 78

ggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 60
gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 120
gcgcgcagag agggagtggc caactccatc actaggggtt cctcagatct gaattctagc 180
ggccgccccc gggtgcgcgg cgtcggtggt gccggcgggg ggcgccaggt cgcaggcggt 240
gtagggctcc aggcaggcgg cgaaggccat gacgtgcgct atgaaggtct gctcctgcac 300
gccgtgaacc aggtgcgcct gcgggccgcg cgcgaacacc gccacgtcct cgcctgcgtg 360

ggtctcttcg tccaggggca ctgctgactg ctgccgatac tcggggctcc cgctctcgct 420
ctcggtaaca tccggccggg cgccgtcctt gagcacatag cctggaccgt ttccttaagc 480
gacgcatgct cgcgataggc acctattggt cttactgaca tccactttgc ctttctctcc 540
acaggaaaac atggggcagc aggccaggat gctgcgggcc caggtgaagc ggcacacggt 600
gcgggacaag ctgaggctgt gccagaactt cctgcagaag ctgcgcttcc tggcggacga 660
gccccagcac agcattcccg acatcttcat ctggatgatg agcaacaaca agcgtgtcgc 720
ctatgcccgt gtgccctcca aggacctgct cttctccatc gtggaggagg agactggcaa 780
ggactgcgcc aaggtcaaga cgctcttcct taagctgcca gggaagcggg gcttcggctc 840
ggcaggctgg acagtgcagg ccaaggtgga gctgtacctg tggctgggcc tcagcaaaca 900
gcgcaaggag ttcctgtgcg gcctgccctg tggcttccag gaggtcaagg cagcccaggg 960
cctgggcctg catgccttcc cacccgtcag cctggtctac accaagaagc aggcgttcca 1020
gctccgagcg cacatgtacc aggcccgcag cctctttgcc gccgacagca gcggactctc 1080
agaccccttt gcccgcgtct tcttcatcaa tcagagtcag tgcacagagg tgctgaatga 1140
gaccctgtgt cccacctggg accagatgct ggtgttcgac aacctggagc tctatggtga 1200
agctcatgag ctgagggacg atccgcccat cattgtcatt gaaatctatg accaggattc 1260
catgggcaaa gctgacttca tgggccggac cttcgccaaa cccctggtga agatggcaga 1320
cgaggcgtac tgcccacccc gcttcccacc tcagctcgag tactaccaga tctaccgtgg 1380
caacgccaca gctggagacc tgctggcggc cttcgagctg ctgcagattg gaccagcagg 1440
gaaggctgac ctgcccccca tcaatggccc ggtggacgtg gaccgaggtc ccatcatgcc 1500
cgtgcccatg ggcatccggc ccgtgctcag caagtaccga gtggaggtgc tgttctgggg 1560
cctacgggac ctaaagcggg tgaacctggc ccaggtggac cggccacggg tggacatcga 1620
gtgtgcaggg aaggggtgc agtcgtccct gatccacaat tataagaaga accccaactt 1680
caacaccctc gtcaagtggt ttgaagtgga cctcccagag aacgagctgc tgcacccgcc 1740
cttgaacatc cgtgtggtgg actgccgggc cttcggtcgc tacacactgg tgggctccca 1800
tgccgtcagc tccctgcgac gcttcatcta ccggccccca gaccgctcgg cccccagctg 1860
gaacaccacg gtcaggcttc tccggcgctg ccgtgtgctg tgcaatgggg gctcctcctc 1920
tcactccaca ggggaggttg tggtgactat ggagccagag gtacccatca agaaactgga 1980
gaccatggtg aagctggacg cgacttctga agctgttgtc aaggtggatg tggctgagga 2040
ggagaaggag aagaagaaga agaagaaggg cactgcggag gagccagagg aggaggagcc 2100
agacgagagc atgctggact ggtggtccaa gtactttgcc tccattgaca ccatgaagga 2160
gcaacttcga caacaagagc cctctggaat tgacttggag gagaaggagg aagtggacaa 2220
taccgagggc ctgaaggggt caatgaaggg caaggagaag gcaagggctg ccaaagagga 2280
gaagaagaag aaaactcaga gctctggctc tggccagggg tccgaggccc ccgagaagaa 2340
gaaacccaag attgatgagc ttaaggtata ccccaaagag ctggagtccg agtttgataa 2400
ctttgaggac tggctgcaca ctttcaactt gcttcggggc aagaccgggg atgatgagga 2460
tggctccacc gaggaggagc gcattgtggg acgcttcaag ggctccctct gcgtgtacaa 2520
agtgccactc ccagaggacg tgtcccggga agccggctac gactccacct acggcatgtt 2580
ccagggcatc ccgagcaatg accccatcaa tgtgctggtc cgagtctatg tggtccgggc 2640
cacggacctg caccctgctg acatcaacgg caaagctgac ccctacatcg ccatccggct 2700

aggcaagact gacatccgcg acaaggagaa ctacatctcc aagcagctca accctgtctt 2760
tgggaagtcc tttgacatcg aggcctcctt ccccatggaa tccatgctga cggtggctgt 2820
gtatgactgg gacctggtgg gcactgatga cctcattggg gaaaccaaga tcgacctgga 2880
gaaccgcttc tacagcaagc accgcgccac ctgcggcatc gcccagacct actccacaca 2940
tggctacaat atctggcggg accccatgaa gcccagccag atcctgaccc gcctctgcaa 3000
agacggcaaa gtggacggcc cccactttgg gccccctggg agagtgaagg tggccaaccg 3060
cgtcttcact gggccctctg agattgagga cgagaacggt cagaggaagc ccacagacga 3120
gcatgtggcg ctgttggccc tgaggcactg ggaggacatc ccccgcgcag gctgccgcct 3180
ggtgccagag catgtggaga cgaggccgct gctcaacccc gacaagccgg gcatcgagca 3240
gggccgcctg gagctgtggg tggacatgtt ccccatggac atgccagccc ctgggacgcc 3300
tctggacatc tcacctcgga agcccaagaa gtacgagctg cgggtcatca tctggaacac 3360
agatgaggtg gtcttggagg acgacgactt cttcacaggg gagaagtcca gtgacatctt 3420
cgtgaggggg tggctgaagg gccagcagga ggacaagcag gacacagacg tccactacca 3480
ctccctcact ggcgagggca acttcaactg gcgctacctg ttccccttcg actacctggc 3540
ggcggaggag aagatcgtca tctccaagaa ggagtccatg ttctcctggg acgagaccga 3600
gtacaagatc cccgcgcggc tcaccctgca gatctgggat gcggaccact tctccgctga 3660
cgacttcctg ggggccatcg agctggacct gaaccggttc ccgcggggcg caaagacagc 3720
caagcagtgc accatggaga tggccaccgg ggaggtggac gtgcccctcg tgtccatctt 3780
caagcaaaag cgcgtcaaag gctggtggcc cctcctggcc cgcaatgaga acgatgagtt 3840
tgagctcacg ggcaaggtgg aggctgagct gcatttactg acagcagagg aggcagagaa 3900
gaacccagtg ggcctggccc gcaatgaacc tgaccccctа gagaaaccca accggcccga 3960
cacggccttc gtctggttcc tcaaccctct caagtccatc aagtacctca tctgcacccg 4020
gtacaagtgg ctcatcatca agatcgtgct ggcgctgttg gggctgctca tgttggggct 4080
cttcctctac agcctccctg gctacatggt caaaaagctc cttggggcat gaacggccgc 4140
tatgctagct tggtaccaag ggcggatcct gcatagagct cgctgatcag cctcgactgt 4200
gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga 4260
aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag 4320
taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga 4380
agacaatagc aggcatgctg gggagagatc tgaggactag tccgtcgact ggggagagat 4440
ctgaggaacc cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg 4500
aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg 4560
agcgagcgcg cagagaggga gtggccaacc ccccccccccc ccccccctgca gcctggcgta 4620
atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgtagcctg aatggcgaat 4680
ggcgcgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg 4740
tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc 4800
tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc 4860
gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta 4920
gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta 4980
atagtggact cttgttccaa actggaacaa cactcaaccc tatcgcggtc tattcttttg 5040

atttataagg gatgttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa 5100
aattttaaca aaattcagaa gaactcgtca agaaggcgat agaaggcgat gcgctgcgaa 5160
tcgggagcgg cgataccgta aagcacgagg aagcggtcag cccattcgcc gccaagctct 5220
tcagcaatat cacgggtagc caacgctatg tcctgatagc ggtccgccac acccagccgg 5280
ccacagtcga tgaatccaga aaagcggcca ttttccacca tgatattcgg caagcaggca 5340
tcgccatggg tcacgacgag atcctcgccg tcgggcatgc tcgccttgag cctggcgaac 5400
agttcggctg gcgcgagccc ctgatgctct tcgtccagat catcctgatc gacaagaccg 5460
gcttccatcc gagtacgtgc tcgctcgatg cgatgtttcg cttggtggtc gaatgggcag 5520
gtagccggat caagcgtatg cagccgccgc attgcatcag ccatgatgga tactttctcg 5580
gcaggagcaa ggtgagatga caggagatcc tgccccggca cttcgcccaa tagcagccag 5640
tcccttcccg cttcagtgac aacgtcgagc acagctgcgc aaggaacgcc cgtcgtggcc 5700
agccacgata gccgcgctgc ctcgtcttgc agttcattca gggcaccgga caggtcggtc 5760
ttgacaaaaa gaaccgggcg cccctgcgct gacagccgga acacggcggc atcagagcag 5820
ccgattgtct gttgtgccca gtcatagccg aatagcctct ccacccaagc ggccggagaa 5880
cctgcgtgca atccatcttg ttcaatcatg cgaaacgatc ctcatcctgt ctcttgatca 5940
gatcttgatc ccctgcgcca tcagatcctt ggcggcgaga aagccatcca gtttactttg 6000
cagggcttcc caaccttacc agagggcgcc ccagctggca attccggttc gcttgctgtc 6060
cataaaaccg cccagtctag ctatcgccat gtaagcccac tgcaagctac ctgctttctc 6120
tttgcgcttg cgttttccct tgtccagata gcccagtagc tgacattcat ccggggtcag 6180
caccgtttct gcggactggc tttctacgtg aaaaggatct aggtgaagat ccttttttgat 6240
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta 6300
gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa 6360
acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt 6420
tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag 6480
ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta 6540
atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca 6600
agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag 6660
cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa 6720
agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga 6780
acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc 6840
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc 6900
ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tgggcttttg ctggcctttt 6960
gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt 7020
gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag 7080
gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa 7140
tgcagggctg ca 7152

<210> 79
<211> 10005
<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 79

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gatagaggtc 60
atccttcctg accatttcca tcattccagt cgaactcaca cacaacacca aatgcattta 120
agtcgcttga aattgctata agcagagcat gttgcgccag catgattaat acagcattta 180
atacagagcc gtgtttattg agtcggtatt cagagtctga ccagaaatta ttaatctggt 240
gaagttattc ctctgtcatt acgtcatggt cgatttcaat ttctattgat gctttccagt 300
cgtaatcaat gatgtatttt ttgatgtttg acctctgttc atatcctcac agataaaaaa 360
tcgccctcac actggagggc aaagaagatt tccaataatc agaacaagtc ggctcctgtt 420
tagttacgag cgacattgct ccgtgtattc actcgttgga atgaatacac agtgcagtgt 480
ttattctgtt atttatgcca aaaattaagg ccactatcag gcagctttgt tgttctgttt 540
accaagttct ctggcaatca ttgccgtcgt tcgtattgcc catttatcga catatttccc 600
atcttcctat acaggaaaca tttcttcagg cttaaccatg cattccgatt gcagcttgca 660
tccattgcat cgcttgaatt gtccacacca ttgatttta tcaatagtcg tagtttaacg 720
gatagtcctg gtattgttcc atcacatcct gaggatgccc ttcgaactct tcaaattctt 780
cttcctaata tcaccttaaa tagtggattg cggtagtaaa gattgtgcct gtcttttaac 840
cacatcaggc tcggtggttc tcgtgtaccc ctacagcgag aaatcggata aactattaca 900
acccctacag tttgtagagt atagaaaatg atccactcgt tattctcgga cgagtgttca 960
gtaatgaacc tctggagaga accatctata tgatcgttat ctgggtttga cttctgcttt 1020
taagcccaga taacttgcct gaatatgtta atgagagaat cggtattcct catgtgtggc 1080
atgttttcgt ctttgctctt gcatttcac tagcaattaa tgtgcatcga ttatcagcta 1140
ttgccagcgc cagatataag cgatttaagc taagaaaacg cattaaggtg caaaacgata 1200
aagtgcgatc agtaattcaa aaccttacag gagagcaatc tatggttttg tgctcagccc 1260
ttaatgaagg caggtagtat gtggttacat caaaacaatt cccatacatt agtgagttga 1320
ttgagcttgg tgtgttgaac aaaacttttt cccgatggaa tggaaagcat atattattcc 1380
ctattgagga tatttactgg actgaattag ttgccagcta tgatccatat aatattgaga 1440
taaagccaag gccaatatct aagtaactag ataagaggaa tcgattttcc cttaattttc 1500
tggcgtccac tgcatgttat gccgcgttcg ccaggcttgc tgtaccatgt gcgctgattc 1560
ttgcgctcaa tacgttgcag gttgctttca atctgtttgt ggtattcagc cagcactgta 1620
aggtctatcg gatttagtgc gctttctact cgtgatttcg gtttgcgatt cagcgagaga 1680
atagggcggt taactggttt tgcgcttacc ccaaccaaca ggggatttgc tgctttccat 1740
tgagcctgtt actctgcgcg acgttcgcgg cggcgtgttt gtgcatccat ctggattctc 1800
ctgtcagtta gctttggtgg tgtgtggcag ttgtagtcct gaacgaaaac cccccgcgat 1860
tggcacgttg gcagctaatc cggaatcgca cttacggcca atgcttcgtt tcgtatcaca 1920
caccccaaag ccttctgctt tgaatgctgc ccttcttcag ggcttaattt ttaagagcgt 1980
caccttcatg gtggtcagtg cgtcctgctg atgtgctcag gcacgattta attaaggcct 2040
taattaggct gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg 2100

cgacctttgg tcgcccggcc tcagtgagcg agcgagcgcg cagagaggga gtggccaact 2160
ccatcactag gggttccttg tagttaatga ttaacccgcc atgctactta tctacgtagc 2220
catgctctag gaagatcgga attcgccctt aagctagcgg cgcgccggta cctagttatt 2280
aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc cgcgttacat 2340
aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca ttgacgtcaa 2400
taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt caatgggtgg 2460
agtatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg ccaagtacgc 2520
cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag tacatgacct 2580
tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt accatggtcg 2640
aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca cccccaattt 2700
tgtatttatt tatttttaa ttattttgtg cagcgatggg ggcgggggggg gggggggggc 2760
gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg agaggtgcgg 2820
cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg cggcggcggc 2880
ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgcgc tgccttcgcc 2940
ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg accgcgttac 3000
tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag cgcttggttt 3060
aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc cgggagctag 3120
agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg gcaacgtgct 3180
ggttattgtg ctgtctcatc attttggcaa agaattctag cggccgccac catggccttg 3240
ctcatccacc tcaagacagt ctcggagctg cggggcaggg gcgaccggat cgccaaagtg 3300
actttccgag ggcaatcctt ctactctcgg gtcctggaga actgtgagga tgtggctgac 3360
tttgatgaga catttcggtg gccggtggcc agcagcatcg acagaaatga gatgctggag 3420
attcaggttt tcaactacag caaagtcttc agcaacaagc tcatcgggac cttccgcatg 3480
gtgctgcaga aggtggtaga ggagagccat gtggaggtga ctgacacgct gattgatgac 3540
aacaatgcta tcatcaagac cagcctgtgc gtggaggtcc ggtatcaggc cactgacggc 3600
acagtgggct cctgggacga tggggacttc ctgggagatg agtctcttca agaggaagag 3660
aaggacagcc aagagacgga tggactgctc ccaggctccc ggcccagctc ccggccccca 3720
ggagagaaga gcttccggag agccgggagg agcgtgttct ccgccatgaa gctcggcaaa 3780
aaccggtctc acaaggagga gccccaaaga ccagatgaac cggcggtgct ggagatggaa 3840
gaccttgacc atctggccat tcggctagga gatggactgg atcccgactc ggtgtctcta 3900
gcctcagtca cagctctcac cactaatgtc tccaacaagc gatctaagcc agacattaag 3960
atggagccaa gtgctgggcg gcccatggat taccaggtca gcatcacggt gatcgaggcc 4020
cggcagctgg tgggcttgaa catggaccct gtggtgtgcg tggaggtggg tgacgacaag 4080
aagtacacat ccatgaagga gtccactaac tgcccctatt acaacgagta cttcgtcttc 4140
gacttccatg tctctccgga tgtcatgttt gacaagatca tcaagatttc ggtgattcac 4200
tccaagaacc tgctgcgcag tggcaccctg gtgggctcct tcaaaatgga cgtgggaacc 4260
gtgtactcgc agccagagca ccagttccat cacaagtggg ccatcctgtc tgaccccgat 4320
gacatctcct cggggctgaa gggctacgtg aagtgtgacg ttgccgtggt gggcaaaggg 4380
gacaacatca agacgcccca caaggccaat gagaccgacg aagatgacat tgaggggaac 4440

ttgctgctcc ccgaggggct gccccccgaa cgccagtggg cccggttcta tgtgaaaatt 4500
taccgagcag aggggctgcc ccgtatgaac acaagcctca tggccaatgt aaagaaggct 4560
ttcatcggtg aaaacaagga cctcgtggac ccctacgtgc aagtcttctt tgctggccag 4620
aagggcaaga cttcagtgca gaagagcagc tatgagcccc tgtggaatga gcaggtcgtc 4680
tttacagacc tcttcccccc actctgcaaa cgcatgaagg tgcagatccg agactcggac 4740
aaggtcaacg acgtggccat cggcacccac ttcattgacc tgcgcaagat ttctaatgac 4800
ggagacaaag gcttcctgcc cacactgggc ccagcctggg tgaacatgta cggctccaca 4860
cgtaactaca cgctgctgga tgagcatcag gacctgaacg agggcctggg ggagggtgtg 4920
tccttccggg cccggctcct gctgggcctg gctgtggaga tcgtagacac ctccaaccct 4980
gagctcacca gctccacaga ggtgcaggtg gagcaggcca cgcccatctc ggagagctgt 5040
gcaggtaaaa tggaagaatt ctttctcttt ggagccttcc tggaggcctc aatgatcgac 5100
cggagaaacg gagacaagcc catcaccttt gaggtcacca taggcaacta tgggaacgaa 5160
gttgatggcc tgtcccggcc ccagcggcct cggccccgga aggagccggg ggatgaggaa 5220
gaagtagacc tgattcagaa cgcaagtgat gacgaggccg gtgatgccgg ggacctggcc 5280
tcagtctcct ccactccacc aatgcggccc caggtcaccg acaggaacta cttccatctg 5340
ccctacctgg agcgaaagcc ctgcatctac atcaagagct ggtggccgga ccagcgccgc 5400
cgcctctaca atgccaacat catggaccac attgccgaca agctggaaga aggcctgaac 5460
gacatacagg agatgatcaa aacggagaag tcctaccctg agcgtcgcct gcggggcgtc 5520
ctggaggagc tgagctgtgg ctgctgccgc ttcctctccc tcgctgacaa ggaccagggc 5580
cactcatccc gcaccaggct tgaccgggag cgcctcaagt cctgcatgag ggagctggta 5640
agtatcaagg ttacaagaca ggtttaagga gaccaataga aactgggctt gtcgagacag 5700
agaagactct tgcgtttctg agctagcccc cgggtgcgcg gcgtcggtgg tgccggcggg 5760
gggcgccagg tcgcaggcgg tgtagggctc caggcaggcg gcgaaggcca tgacgtgcgc 5820
tatgaaggtc tgctcctgca cgccgtgaac caggtgcgcc tgcgggccgc gcgcgaacac 5880
cgccacgtcc tcgcctgcgt gggtctcttc gtccaggggc actgctgact gctgccgata 5940
ctcggggctc ccgctctcgc tctcggtaac atccggccgg gcgccgtcct tgagcacata 6000
gcctggaccg tttcgtcgac ctcgagttaa gggcgaattc ccgataagga tcttcctaga 6060
gcatggctac gtagataagt agcatggcgg gttaatcatt aactacaagg aacccctagt 6120
gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg ggcgaccaaa 6180
ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg agcgagcgag cgcgcagcct 6240
taattaaatc cacatctgta tgtttttttat attaatttat tttttgcagg ggggcattgt 6300
ttggtaggtg agagttctga attgctatgt ttagtgagtt gtatctattt atttttcaat 6360
aaatacaatt agttatgtgt tttgggggcg atcgtgaggc aaagaaaacc cggcgctgag 6420
gccgggttat tcttgttctc tggtcaaatt atatagttgg aaaacaagga tgcatatatg 6480
aatgaacgat gcagaggcaa tgccgatggc gatagtgggt atcaggtagc cgcttatgct 6540
ggaaagaagc aataacccgc agaaaaacaa agctccaagc tcaacaaaac taagggcata 6600
gacaataact acctatgtca tatacccata ctctctaatc ttggccagtc ggcgcgttct 6660
gcttccgatt agaaacgtca aggcagcaat caggattgca atcttggttc ctgcatagga 6720
tgacaatgtc gccccaagac catctctatg agctgaaaaa gaaacacaag gaatgtagtg 6780

gcggaaaagg agatagcaaa tgcttacgat aacgtaagga attattacta tgtaaacacc 6840
aggcaagatt ctgttccgta taattactcc tgataattaa tccttaactt tgcccacctg 6900
ccttttaaaa cattccagta tatcactttt cattcttgcg tagcaatatg ccctctcttc 6960
agctatctca gcattggtga ccttgttcag aggcgctgag agatggcctt tttctgatag 7020
ataatgttct gttaaaatat ctccggcctc atcttttgcc cgcaggctaa tgtctgaaaa 7080
ttgaggtgac gggttaaaaa taatatcctt ggcaaccttt tttatatccc ttttaaattt 7140
tggcttaatg actatatcca atgagtcaaa aagctcccct tcaatatctg ttgcccctaa 7200
gacctttaat atatcgccaa atacaggtag cttggcttct accttcaccg ttgttctgcc 7260
gatgaaatgc taatgcataa catcgtcttt ggtggttccc ctcatcagtg gctctatctg 7320
aacgcgctct ccactgctta atgacattcc tttcccgatt aaaaaatctg tcagatcgga 7380
tgtggtcggc ccgaaaacag ttctggcaaa accaatggtg tcgccttcaa caaacaaaaa 7440
agatgggaat cccaatgatt cgtcatctgc gaggctgttc ttaatatctt caactgtagc 7500
tttagagcga tttatcttct gaaccagact cttgtcattt gttttggtaa agagaaaagt 7560
ttttccatcg attttatgaa tatacaaata attggagcca accttcaggt gatgattatc 7620
agccagcaga gaattaagga aaacagacag gtttattgag cacttatctt tccctttatt 7680
tttgctgcgg taagtcgcat aaaaaccatt cttcacaatt caatccattt actatgttat 7740
gttctgaggg gagtgaaaat tcccctaatt cgatgaagat tcttgctaaa ttgttatcag 7800
ctatgcgccg accagaacac cttgccgatc agccaaacgt ctaatcaggc cactgactag 7860
cgataacttt ccccacaacg gaacaactct cattgcatgg gataattggg tactgtgggt 7920
ttagtggttg taaaaacacc tgaccgctat ccctgatcag tttcttgaag gtaaactcat 7980
caccccccaag tctggctata cagaaatcac ctggctcaac agcctgctca gggtcaacga 8040
gaatttacat tccgtcagga tagcttggct tggagcctgt tggtgcggtc acggaattac 8100
cttcaacctc aagccagaat gcagaatcac tggctttttt ggttgtgctt acccatctct 8160
ccgcatcacc tttggtaaag gttctaagct aaggtgagaa catccctgcc tgaacatgag 8220
aaaaaacagg gtactcatac tcacttatta gtgacggcta tgagcaaaag gccagcaaaa 8280
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga 8340
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 8400
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 8460
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 8520
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 8580
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 8640
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 8700
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 8760
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 8820
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 8880
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 8940
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 9000
cacctagatc cttttaaatt aaaaatgaag ttttaaatca agcccaatct gaataatgtt 9060
acaaccaatt aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt 9120

tattcatatc aggattatca ataccatatt tttgaaaaag ccgtttctgt aatgaaggag 9180
aaaactcacc gaggcagttc cataggatgg caagatcctg gtatcggtct gcgattccga 9240
ctcgtccaac atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagtg 9300
agaaatcacc atgagtgacg actgaatccg gtgagaatgg caaaagttta tgcatttctt 9360
tccagacttg ttcaacaggc agccattac gctcgtcatc aaaatcactc gcatcaacca 9420
aaccgttatt cattcgtgat tgcgcctgag caagacgaaa tacgcgatcg ctgttaaaag 9480
gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa 9540
tattttcacc tgaatcagga tattcttcta atacctggaa tgctgttttt ccggggatcg 9600
cagtggtgag taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag 9660
gcataaattc cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc 9720
tacctttgcc atgtttcaga aacaactctg gcgcatcggg cttcccatac aagcgataga 9780
ttgtcgcacc tgattgcccg acattatcgc gagcccattt atacccatat aaatcagcat 9840
ccatgttgga atttaatcgc ggcctcgacg tttcccgttg aatatggctc ataacacccc 9900
ttgtattact gtttatgtaa gcagacagtt ttattgttca tgatgatata tttttatctt 9960
gtgcaatgta acatcagaga ttttgagaca cgggccagag ctgca 10005

<210> 80
<211> 6653
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 80

gggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc 60
gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 120
gcgcgcagag agggagtggc caactccatc actaggggtt cctcagatct gaattcggta 180
cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 240
cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca 300
ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt 360
caatgggtgg agtatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg 420
ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag 480
tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 540
accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc cccctcccca 600
cccccaattt tgtatttatt tatttttaa ttattttgtg cagcgatggg ggcgggggg 660
ggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg ggcgaggcgg 720
agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt tatggcgagg 780
cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt cgctgcgcgc 840
tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg 900
accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg ctgtaattag 960
cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc 1020

cgggagctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1080
gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattctag cggccgccac 1140
catggccttg ctcatccacc tcaagacagt ctcggagctg cggggcaggg gcgaccggat 1200
cgccaaagtg actttccgag ggcaatcctt ctactctcgg gtcctggaga actgtgagga 1260
tgtggctgac tttgatgaga catttcggtg gccggtggcc agcagcatcg acagaaatga 1320
gatgctggag attcaggttt tcaactacag caaagtcttc agcaacaagc tcatcgggac 1380
cttccgcatg gtgctgcaga aggtggtaga ggagagccat gtggaggtga ctgacacgct 1440
gattgatgac aacaatgcta tcatcaagac cagcctgtgc gtggaggtcc ggtatcaggc 1500
cactgacggc acagtgggct cctgggacga tggggacttc ctgggagatg agtctcttca 1560
agaggaagag aaggacagcc aagagacgga tggactgctc ccaggctccc ggcccagctc 1620
ccggccccca ggagagaaga gcttccggag agccgggagg agcgtgttct ccgccatgaa 1680
gctcggcaaa aaccggtctc acaaggagga gccccaaaga ccagatgaac cggcggtgct 1740
ggagatggaa gaccttgacc atctggccat tcggctagga gatggactgg atcccgactc 1800
ggtgtctcta gcctcagtca cagctctcac cactaatgtc tccaacaagc gatctaagcc 1860
agacattaag atggagccaa gtgctgggcg gcccatggat taccaggtca gcatcacggt 1920
gatcgaggcc cggcagctgg tgggcttgaa catggaccct gtggtgtgcg tggaggtggg 1980
tgacgacaag aagtacacat ccatgaagga gtccactaac tgcccctatt acaacgagta 2040
cttcgtcttc gacttccatg tctctccgga tgtcatgttt gacaagatca tcaagatttc 2100
ggtgattcac tccaagaacc tgctgcgcag tggcaccctg gtgggctcct tcaaaatgga 2160
cgtgggaacc gtgtactcgc agccagagca ccagttccat cacaagtggg ccatcctgtc 2220
tgaccccgat gacatctcct cggggctgaa gggctacgtg aagtgtgacg ttgccgtggt 2280
gggcaaaggg gacaacatca agacgcccca caaggccaat gagaccgacg aagatgacat 2340
tgaggggaac ttgctgctcc ccgagggggt gccccccgaa cgccagtggg cccggttcta 2400
tgtgaaaatt taccgagcag aggggctgcc ccgtatgaac acaagcctca tggccaatgt 2460
aaagaaggct ttcatcggtg aaaacaagga cctcgtggac ccctacgtgc aagtcttctt 2520
tgctggccag aagggcaaga cttcagtgca gaagagcagc tatgagcccc tgtggaatga 2580
gcaggtcgtc tttacagacc tcttcccccc actctgcaaa cgcatgaagg tgcagatccg 2640
agactcggac aaggtcaacg acgtggccat cggcacccac ttcattgacc tgcgcaagat 2700
ttctaatgac ggagacaaag gcttcctgcc cacactgggc ccagcctggg tgaacatgta 2760
cggctccaca cgtaactaca cgctgctgga tgagcatcag gacctgaacg agggcctggg 2820
ggagggtgtg tccttccggg cccggctcct gctgggcctg gctgtggaga tcgtagacac 2880
ctccaaccct gagctcacca gctccacaga ggtgcaggtg gagcaggcca cgcccatctc 2940
ggagagctgt gcaggtaaaa tggaagaatt ctttctcttt ggagccttcc tggaggcctc 3000
aatgatcgac cggagaaacg gagacaagcc catcaccttt gaggtcacca taggcaacta 3060
tgggaacgaa gttgatggcc tgtcccggcc ccagcggcct cggccccgga aggagccggg 3120
ggatgaggaa gaagtagacc tgattcagaa cgcaagtgat gacgaggccg gtgatgccgg 3180
ggacctggcc tcagtctcct ccactccacc aatgcggccc caggtcaccg acaggaacta 3240
cttccatctg ccctacctgg agcgaaagcc ctgcatctac atcaagagct ggtggccgga 3300
ccagcgccgc cgcctctaca atgccaacat catggaccac attgccgaca agctggaaga 3360

aggcctgaac gacatacagg agatgatcaa aacggagaag tcctaccctg agcgtcgcct 3420
gcggggcgtc ctggaggagc tgagctgtgg ctgctgccgc ttcctctccc tcgctgacaa 3480
ggaccagggc cactcatccc gcaccaggct tgaccgggag cgcctcaagt cctgcatgag 3540
ggagctggta agtatcaagg ttacaagaca ggtttaagga gaccaataga aactgggctt 3600
gtcgagacag agaagactct tgcgtttctg agctagcccc cgggtgcgcg gcgtcggtgg 3660
tgccggcggg gggcgccagg tcgcaggcgg tgtagggctc caggcaggcg gcgaaggcca 3720
tgacgtgcgc tatgaaggtc tgctcctgca cgccgtgaac caggtgcgcc tgcgggccgc 3780
gcgcgaacac cgccacgtcc tcgcctgcgt gggtctcttc gtccaggggc actgctgact 3840
gctgccgata ctcggggctc ccgctctcgc tctcggtaac atccggccgg gcgccgtcct 3900
tgagcacata gcctggaccg tttcgtcgac tggggagaga tctgaggaac ccctagtgat 3960
ggagttggcc actccctctc tgcgcgctcg ctcgctcact gaggccgccc gggcaaagcc 4020
cgggcgtcgg gcgacctttg gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg 4080
agtggccaac ccccccccccc ccccccctgc agcctggcgt aatagcgaag aggcccgcac 4140
cgatcgccct tcccaacagt tgcgtagcct gaatggcgaa tggcgcgacg cgccctgtag 4200
cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag 4260
cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgccggctt 4320
tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg ctttacggca 4380
cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat cgccctgata 4440
gacggttttt cgcccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca 4500
aactggaaca acactcaacc ctatcgcggt ctattctttt gatttataag ggatgttgcc 4560
gatttcggcc tattggttaa aaaatgagct gatttaacaa aaattttaac aaaattcaga 4620
agaactcgtc aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt 4680
aaagcacgag gaagcggtca gcccattcgc cgccaagctc ttcagcaata tcacgggtag 4740
ccaacgctat gtcctgatag cggtccgcca cacccagccg gccacagtcg atgaatccag 4800
aaaagcggcc attttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga 4860
gatcctcgcc gtcgggcatg ctcgccttga gcctggcgaa cagttcggct ggcgcgagcc 4920
cctgatgctc ttcgtccaga tcatcctgat cgacaagacc ggcttccatc cgagtacgtg 4980
ctcgctcgat gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgtat 5040
gcagccgccg cattgcatca gccatgatgg atactttctc ggcaggagca aggtgagatg 5100
acaggagatc ctgccccggc acttcgccca atagcagcca gtcccttccc gcttcagtga 5160
caacgtcgag cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg 5220
cctcgtcttg cagttcattc agggcaccgg acaggtcggt cttgacaaaa agaaccgggc 5280
gcccctgcgc tgacagccgg aacacggcgg catcagagca gccgattgtc tgttgtgccc 5340
agtcatagcc gaatagcctc tccacccaag cggccggaga acctgcgtgc aatccatctt 5400
gttcaatcat gcgaaacgat cctcatcctg tctcttgatc agatcttgat cccctgcgcc 5460
atcagatcct tggcggcgag aaagccatcc agtttacttt gcagggcttc ccaaccttac 5520
cagagggcgc cccagctggc aattccggtt cgcttgctgt ccataaaacc gcccagtcta 5580
gctatcgcca tgtaagccca ctgcaagcta cctgctttct ctttgcgctt gcgttttccc 5640
ttgtccagat agcccagtag ctgacattca tccggggtca gcaccgtttc tgcggactgg 5700

ctttctacgt gaaaaggatc taggtgaaga tccttttga taatctcatg accaaaatcc 5760
cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt 5820
cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac 5880
cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct 5940
tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta ggccaccact 6000
tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg 6060
ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata 6120
aggcgcagcg gtcgggctga acgggggttt cgtgcacaca gcccagcttg gagcgaacga 6180
cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag 6240
ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg 6300
agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac 6360
ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca 6420
acgcggcctt tttacggttc ctgggctttt gctggccttt tgctcacatg ttctttcctg 6480
cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct gataccgctc 6540
gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa gagcgcccaa 6600
tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagggct gca 6653

<210> 81
<211> 6986
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 81

ccttaattag gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacgt 180
agccatgctc taggaagatc ggaattcgcc cttaagctag cggcgcgccc aattctgcag 240
ctcagcctac tacttgcttt ccaggctgtt cctagttccc atgtcagctg cttgtgcttt 300
ccagagacaa aacaggaata atagatgtca ttaaatatac attgggcccc aggcggtcaa 360
tgtggcagcc tgagcctcct ttccatctct gtggaggcag acataggacc cccaacaaac 420
agcatgcagg ttgggagcca gccacaggac ccaggtaagg ggccctgggt ccttaagctt 480
ctgccactgg ctccggcatt gcagagagaa gagaaggggc ggcagactgg agagctgggc 540
tccatttttg ttccttggtg ccctgcccct ccccatgacc tgcagagaca ttcagcctgc 600
caggctttat gaggtgggag ctgggctctc cctgatgtat tattcagctc cctggagttg 660
gccagctcct gttacactgg ccacagccct gggcatccgc ttctcacttc tagtttcccc 720
tccaaggtaa tgtggtgggt catgatcatt ctatcctggc ttcagggacc tgactccact 780
ttggggccat tcgaggggtc tagggtagat gatgtccccc tgtggggatt aatgtcctgc 840
tctgtaaaac tgagctagct gagatccagg agggcttggc cagagacagc aagttgttgc 900
catggtgact ttaaagccag gttgctgccc cagcacaggc ctcccagtct accctcacta 960

gaaaacaaca cccaggcact ttccaccacc tctcaaaggt gaaacccaag gctggtctag 1020
agaatgaatt atggatcctc gctgtccgtg ccacccagct agtcccagcg gctcagacac 1080
tgaggagaga ctgtaggttc agctacaagc aaaaagacct agctggtctc caagcagtgt 1140
ctccaagtcc ctgaacctgt gacacctgcc ccaggcatca tcaggcacag agggccacca 1200
agaattctag cggccgccac catggccttg ctcatccacc tcaagacagt ctcggagctg 1260
cggggcaggg gcgaccggat cgccaaagtg actttccgag ggcaatcctt ctactctcgg 1320
gtcctggaga actgtgagga tgtggctgac tttgatgaga catttcggtg gccggtggcc 1380
agcagcatcg acagaaatga gatgctggag attcaggttt tcaactacag caaagtcttc 1440
agcaacaagc tcatcgggac cttccgcatg gtgctgcaga aggtggtaga ggagagccat 1500
gtggaggtga ctgacacgct gattgatgac aacaatgcta tcatcaagac cagcctgtgc 1560
gtggaggtcc ggtatcaggc cactgacggc acagtgggct cctgggacga tggggacttc 1620
ctgggagatg agtctcttca agaggaagag aaggacagcc aagagacgga tggactgctc 1680
ccaggctccc ggcccagctc ccggccccca ggagagaaga gcttccggag agccgggagg 1740
agcgtgttct ccgccatgaa gctcggcaaa aaccggtctc acaaggagga gccccaaaga 1800
ccagatgaac cggcggtgct ggagatggaa gaccttgacc atctggccat tcggctagga 1860
gatggactgg atcccgactc ggtgtctcta gcctcagtca cagctctcac cactaatgtc 1920
tccaacaagc gatctaagcc agacattaag atggagccaa gtgctgggcg gcccatggat 1980
taccaggtca gcatcacggt gatcgaggcc cggcagctgg tgggcttgaa catggaccct 2040
gtggtgtgcg tggaggtggg tgacgacaag aagtacacat ccatgaagga gtccactaac 2100
tgcccctatt acaacgagta cttcgtcttc gacttccatg tctctccgga tgtcatgttt 2160
gacaagatca tcaagatttc ggtgattcac tccaagaacc tgctgcgcag tggcaccctg 2220
gtgggctcct tcaaaatgga cgtgggaacc gtgtactcgc agccagagca ccagttccat 2280
cacaagtggg ccatcctgtc tgaccccgat gacatctcct cggggctgaa gggctacgtg 2340
aagtgtgacg ttgccgtggt gggcaaaggg gacaacatca agacgcccca caaggccaat 2400
gagaccgacg aagatgacat tgaggggaac ttgctgctcc ccgaggggt gccccccgaa 2460
cgccagtggg cccggttcta tgtgaaaatt taccgagcag aggggctgcc ccgtatgaac 2520
acaagcctca tggccaatgt aaagaaggct ttcatcggtg aaaacaagga cctcgtggac 2580
ccctacgtgc aagtcttctt tgctggccag aagggcaaga cttcagtgca gaagagcagc 2640
tatgagcccc tgtggaatga gcaggtcgtc tttacagacc tcttcccccc actctgcaaa 2700
cgcatgaagg tgcagatccg agactcggac aaggtcaacg acgtggccat cggcacccac 2760
ttcattgacc tgcgcaagat ttctaatgac ggagacaaag gcttcctgcc cacactgggc 2820
ccagcctggg tgaacatgta cggctccaca cgtaactaca cgctgctgga tgagcatcag 2880
gacctgaacg agggcctggg ggagggtgtg tccttccggg cccggctcct gctgggcctg 2940
gctgtggaga tcgtagacac ctccaaccct gagctcacca gctccacaga ggtgcaggtg 3000
gagcaggcca cgcccatctc ggagagctgt gcaggtaaaa tggaagaatt ctttctcttt 3060
ggagccttcc tggaggcctc aatgatcgac cggagaaacg gagacaagcc catcaccttt 3120
gaggtcacca taggcaacta tgggaacgaa gttgatggcc tgtcccggcc ccagcggcct 3180
cggccccgga aggagccggg ggatgaggaa gaagtagacc tgattcagaa cgcaagtgat 3240
gacgaggccg gtgatgccgg ggacctggcc tcagtctcct ccactccacc aatgcggccc 3300

caggtcaccg acaggaacta cttccatctg ccctacctgg agcgaaagcc ctgcatctac 3360
atcaagagct ggtggccgga ccagcgccgc cgcctctaca atgccaacat catggaccac 3420
attgccgaca agctggaaga aggcctgaac gacatacagg agatgatcaa aacggagaag 3480
tcctaccctg agcgtcgcct gcggggcgtc ctggaggagc tgagctgtgg ctgctgccgc 3540
ttcctctccc tcgctgacaa ggaccagggc cactcatccc gcaccaggct tgaccgggag 3600
cgcctcaagt cctgcatgag ggagctggta agtatcaagg ttacaagaca ggtttaagga 3660
gaccaataga aactgggctt gtcgagacag agaagactct tgcgtttctg agctagcccc 3720
cgggtgcgcg gcgtcggtgg tgccggcggg gggcgccagg tcgcaggcgg tgtagggctc 3780
caggcaggcg gcgaaggcca tgacgtgcgc tatgaaggtc tgctcctgca cgccgtgaac 3840
caggtgcgcc tgcgggccgc gcgcgaacac cgccacgtcc tcgcctgcgt gggtctcttc 3900
gtccaggggc actgctgact gctgccgata ctcggggctc ccgctctcgc tctcggtaac 3960
atccggccgg gcgccgtcct tgagcacata gcctggaccg tttcgtcgac ctcgagttaa 4020
gggcgaattc ccgataagga tcttcctaga gcatggctac gtagataagt agcatggcgg 4080
gttaatcatt aactacaagg aacccctagt gatggagttg gccactccct ctctgcgcgc 4140
tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc 4200
ggcctcagtg agcgagcgag cgcgcagcct taattaacct aattcactgg ccgtcgtttt 4260
acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccttg cagcacatcc 4320
cccttcgcc agctggcgta atagcgaaga ggcccgcacc gatcgccctt cccaacagtt 4380
gcgcagcctg aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg cggcgggtgt 4440
ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg ctcctttcgc 4500
tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc taaatcgggg 4560
gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa aacttgatta 4620
gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt 4680
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat 4740
ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa 4800
tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc ttacaattta 4860
ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttatttttt ctaaatacat 4920
tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa 4980
aggaagagta tgagccatat tcaacgggaa acgtcgaggc cgcgattaaa ttccaacatg 5040
gatgctgatt tatatgggta taaatgggct cgcgataatg tcgggcaatc aggtgcgaca 5100
atctatcgct tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt 5160
agcgttgcca atgatgttac agatgagatg gtcagactaa actggctgac ggaatttatg 5220
cctcttccga ccatcaagca ttttatccgt actcctgatg atgcatggtt actcaccact 5280
gcgatccccg gaaaaacagc attccaggta ttagaagaat atcctgattc aggtgaaaat 5340
attgttgatg cgctggcagt gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt 5400
ccttttaaca gcgatcgcgt atttcgtctt gctcaggcgc aatcacgaat gaataacggt 5460
ttggttgatg cgagtgattt tgatgacgag cgtaatggct ggcctgttga acaagtctgg 5520
aaagaaatgc ataaactttt gccattctca ccggattcag tcgtcactca tggtgatttc 5580
tcacttgata accttatttt tgacgagggg aaattaatag gttgtattga tgttggacga 5640

gtcggaatcg cagaccgata ccaggatctt gccatcctat ggaactgcct cggtgagttt 5700
tctccttcat tacagaaacg gctttttcaa aaatatggta ttgataatcc tgatatgaat 5760
aaattgcagt ttcatttgat gctcgatgag tttttctaac tgtcagacca agtttactca 5820
tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 5880
ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 5940
gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 6000
tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 6060
ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 6120
ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatacctc 6180
gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 6240
ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 6300
tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 6360
ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 6420
agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 6480
agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 6540
gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 6600
tggccttttg ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 6660
accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 6720
gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg 6780
attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac 6840
gcaattaatg tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg 6900
gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac 6960
catgattacg ccagatttaa ttaagg 6986

<210> 82
<211> 7414
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 82

ccttaattag gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacgt 180
agccatgctc taggaagatc ggaattcgcc cttaagctag cggcgcgccc ccgggtgcgc 240
ggcgtcggtg gtgccggcgg ggggcgccag gtcgcaggcg gtgtagggct ccaggcaggc 300
ggcgaaggcc atgacgtgcg ctatgaaggt ctgctcctgc acgccgtgaa ccaggtgcgc 360
ctgcgggccg cgcgcgaaca ccgccacgtc ctcgcctgcg tgggtctctt cgtccagggg 420
cactgctgac tgctgccgat actcggggct cccgctctcg ctctcggtaa catccggccg 480
ggcgccgtcc ttgagcacat agcctggacc gtttccttaa gcgacgcatg ctcgcgatag 540

gcacctattg gtcttactga catccacttt gcctttctct ccacaggaaa acatggggca 600
gcaggccagg atgctgcggg cccaggtgaa gcggcacacg gtgcgggaca agctgaggct 660
gtgccagaac ttcctgcaga agctgcgctt cctggcggac gagccccagc acagcattcc 720
cgacatcttc atctggatga tgagcaacaa caagcgtgtc gcctatgccc gtgtgccctc 780
caaggacctg ctcttctcca tcgtggagga ggagactggc aaggactgcg ccaaggtcaa 840
gacgctcttc cttaagctgc cagggaagcg gggcttcggc tcggcaggct ggacagtgca 900
ggccaaggtg gagctgtacc tgtggctggg cctcagcaaa cagcgcaagg agttcctgtg 960
cggcctgccc tgtggcttcc aggaggtcaa ggcagcccag ggcctgggcc tgcatgcctt 1020
cccacccgtc agcctggtct acaccaagaa gcaggcgttc cagctccgag cgcacatgta 1080
ccaggcccgc agcctctttg ccgccgacag cagcggactc tcagacccct ttgcccgcgt 1140
cttcttcatc aatcagagtc agtgcacaga ggtgctgaat gagaccctgt gtcccacctg 1200
ggaccagatg ctggtgttcg acaacctgga gctctatggt gaagctcatg agctgaggga 1260
cgatccgccc atcattgtca ttgaaatcta tgaccaggat tccatgggca aagctgactt 1320
catgggccgg accttcgcca aacccctggt gaagatggca gacgaggcgt actgcccacc 1380
ccgcttccca cctcagctcg agtactacca gatctaccgt ggcaacgcca cagctggaga 1440
cctgctggcg gccttcgagc tgctgcagat tggaccagca gggaaggctg acctgccccc 1500
catcaatggc ccggtggacg tggaccgagg tcccatcatg cccgtgccca tgggcatccg 1560
gcccgtgctc agcaagtacc gagtggaggt gctgttctgg ggcctacggg acctaaagcg 1620
ggtgaacctg gcccaggtgg accggccacg ggtggacatc gagtgtgcag ggaaggggt 1680
gcagtcgtcc ctgatccaca attataagaa gaaccccaac ttcaacaccc tcgtcaagtg 1740
gtttgaagtg gacctcccag agaacgagct gctgcacccg cccttgaaca tccgtgtggt 1800
ggactgccgg gccttcggtc gctacacact ggtgggctcc catgccgtca gctccctgcg 1860
acgcttcatc taccggcccc cagaccgctc ggcccccagc tggaacacca cggtcaggct 1920
tctccggcgc tgccgtgtgc tgtgcaatgg gggctcctcc tctcactcca caggggaggt 1980
tgtggtgact atggagccag aggtacccat caagaaactg gagaccatgg tgaagctgga 2040
cgcgacttct gaagctgttg tcaaggtgga tgtggctgag gaggagaagg agaagaagaa 2100
gaagaagaag ggcactgcgg aggagccaga ggaggaggag ccagacgaga gcatgctgga 2160
ctggtggtcc aagtactttg cctccattga caccatgaag gagcaacttc gacaacaaga 2220
gccctctgga attgacttgg aggagaagga ggaagtggac aataccgagg gcctgaaggg 2280
gtcaatgaag ggcaaggaga aggcaagggc tgccaaagag gagaagaaga agaaaactca 2340
gagctctggc tctggccagg ggtccgaggc ccccgagaag aagaaaccca agattgatga 2400
gcttaaggta taccccaaag agctggagtc cgagtttgat aactttgagg actggctgca 2460
cactttcaac ttgcttcggg gcaagaccgg ggatgatgag gatggctcca ccgaggagga 2520
gcgcattgtg ggacgcttca agggctccct ctgcgtgtac aaagtgccac tcccagagga 2580
cgtgtcccgg gaagccggct acgactccac ctacggcatg ttccagggca tcccgagcaa 2640
tgaccccatc aatgtgctgg tccgagtcta tgtggtccgg gccacggacc tgcaccctgc 2700
tgacatcaac ggcaaagctg acccctacat cgccatccgg ctaggcaaga ctgacatccg 2760
cgacaaggag aactacatct ccaagcagct caaccctgtc tttgggaagt cctttgacat 2820
cgaggcctcc ttccccatgg aatccatgct gacggtggct gtgtatgact gggacctggt 2880

gggcactgat gacctcattg gggaaaccaa gatcgacctg gagaaccgct tctacagcaa 2940
gcaccgcgcc acctgcggca tcgcccagac ctactccaca catggctaca atatctggcg 3000
ggaccccatg aagcccagcc agatcctgac ccgcctctgc aaagacggca aagtggacgg 3060
cccccacttt gggccccctg ggagagtgaa ggtggccaac cgcgtcttca ctgggccctc 3120
tgagattgag gacgagaacg gtcagaggaa gcccacagac gagcatgtgg cgctgttggc 3180
cctgaggcac tgggaggaca tcccccgcgc aggctgccgc ctggtgccag agcatgtgga 3240
gacgaggccg ctgctcaacc ccgacaagcc gggcatcgag cagggccgcc tggagctgtg 3300
ggtggacatg ttccccatgg acatgccagc ccctgggacg cctctggaca tctcacctcg 3360
gaagcccaag aagtacgagc tgcgggtcat catctggaac acagatgagg tggtcttgga 3420
ggacgacgac ttcttcacag gggagaagtc cagtgacatc ttcgtgaggg ggtggctgaa 3480
gggccagcag gaggacaagc aggacacaga cgtccactac cactccctca ctggcgaggg 3540
caacttcaac tggcgctacc tgttcccctt cgactacctg gcggcggagg agaagatcgt 3600
catctccaag aaggagtcca tgttctcctg ggacgagacc gagtacaaga tccccgcgcg 3660
gctcaccctg cagatctggg atgcggacca cttctccgct gacgacttcc tgggggccat 3720
cgagctggac ctgaaccggt tcccgcgggg cgcaaagaca gccaagcagt gcaccatgga 3780
gatggccacc ggggaggtgg acgtgccccc cgtgtccatc ttcaagcaaa agcgcgtcaa 3840
aggctggtgg cccctcctgg cccgcaatga gaacgatgag tttgagctca cgggcaaggt 3900
ggaggctgag ctgcatttac tgacagcaga ggaggcagag aagaacccag tgggcctggc 3960
ccgcaatgaa cctgaccccc tagagaaacc caaccggccc gacacggcct tcgtctggtt 4020
cctcaaccct ctcaagtcca tcaagtacct catctgcacc cggtacaagt ggctcatcat 4080
caagatcgtg ctggcgctgt tggggctgct catgttgggg ctcttcctct acagcctccc 4140
tggctacatg gtcaaaaagc tccttggggc atgaacggcc gctatgctag cttggtacca 4200
agggcggatc ctgcatagag ctcgctgatc agcctcgact gtgccttcta gttgccagcc 4260
atctgttgtt tgcccctccc ccgtgccttc cttgaccctg gaaggtgcca ctcccactgt 4320
cctttcctaa taaaatgagg aaattgcatc gcattgtctg agtaggtgtc attctattct 4380
gggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata gcaggcatct 4440
cgagttaagg gcgaattccc gataaggatc ttcctagagc atggctacgt agataagtag 4500
catggcgggt taatcattaa ctacaaggaa cccctagtga tggagttggc cactccctct 4560
ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg cccgggcttt 4620
gcccgggcgg cctcagtgag cgagcgagcg cgcagcctta attaacctaa ttcactggcc 4680
gtcgttttac aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca 4740
gcacatcccc ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc 4800
caacagttgc gcagcctgaa tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg 4860
gcgggtgtgg tggttacgcg cagcgtgacc gctacacttg ccagcgccct agcgcccgct 4920
cctttcgctt tcttcccttc ctttctcgcc acgttcgccg gctttccccg tcaagctcta 4980
aatcgggggc tccctttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa 5040
cttgattagg gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgccct 5100
ttgacgttgg agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc 5160
aaccctatct cggtctattc ttttgattta taagggattt tgccgatttc ggcctattgg 5220

ttaaaaaatg agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgctt 5280
acaatttagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt ttatttttct 5340
aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg cttcaataat 5400
attgaaaaag gaagagtatg agccatattc aacgggaaac gtcgaggccg cgattaaatt 5460
ccaacatgga tgctgattta tatgggtata aatgggctcg cgataatgtc gggcaatcag 5520
gtgcgacaat ctatcgcttg tatgggaagc ccgatgcgcc agagttgttt ctgaaacatg 5580
gcaaaggtag cgttgccaat gatgttacag atgagatggt cagactaaac tggctgacgg 5640
aatttatgcc tcttccgacc atcaagcatt ttatccgtac tcctgatgat gcatggttac 5700
tcaccactgc gatccccgga aaaacagcat tccaggtatt agaagaatat cctgattcag 5760
gtgaaaatat tgttgatgcg ctggcagtgt tcctgcgccg gttgcattcg attcctgttt 5820
gtaattgtcc ttttaacagc gatcgcgtat ttcgtcttgc tcaggcgcaa tcacgaatga 5880
ataacggttt ggttgatgcg agtgattttg atgacgagcg taatggctgg cctgttgaac 5940
aagtctggaa agaaatgcat aaacttttgc cattctcacc ggattcagtc gtcactcatg 6000
gtgatttctc acttgataac cttatttttg acgaggggaa attaataggt tgtattgatg 6060
ttggacgagt cggaatcgca gaccgatacc aggatcttgc catcctatgg aactgcctcg 6120
gtgagtttc tccttcatta cagaaacggc tttttcaaaa atatggtatt gataatcctg 6180
atatgaataa attgcagttt catttgatgc tcgatgagtt tttctaactg tcagaccaag 6240
tttactcata tatactttag attgatttaa aacttcattt ttaatttaaa aggatctagg 6300
tgaagatcct ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact 6360
gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg 6420
taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc 6480
aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata 6540
ctgttcttct agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta 6600
catacctcgc tctgctaatc ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc 6660
ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg 6720
ggggttcgtg cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac 6780
agcgtgagct atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg 6840
taagcggcag ggtcggaaca ggagagcgca cgagggagct tccagggga aacgcctggt 6900
atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct 6960
cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggcctttta cggttcctgg 7020
ccttttgctg gccttttgct cacatgttct ttcctgcgtt atcccctgat tctgtggata 7080
accgtattac cgcctttgag tgagctgata ccgctcgccg cagccgaacg accgagcgca 7140
gcgagtcagt gagcgaggaa gcggaagagc gcccaatacg caaaccgcct ctccccgcgc 7200
gttggccgat tcattaatgc agctggcacg acaggtttcc cgactggaaa gcgggcagtg 7260
agcgcaacgc aattaatgtg agttagctca ctcattaggc accccaggct ttacacttta 7320
tgcttccggc tcgtatgttg tgtggaattg tgagcggata acaatttcac acaggaaaca 7380
gctatgacca tgattacgcc agatttaatt aagg 7414

〈210〉 83

〈211〉 6959

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 合成构建体

<400> 83

ccttaattag gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacgt 180
agccatgctc taggaagatc ggaattcgcc cttaagctag cggcgcgccg gtacctagtt 240
attaatagta atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta 300
cataacttac ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt 360
caataatgac gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg 420
tggagtattt acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagta 480
cgccccctat tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga 540
ccttatggga ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg 600
tcgaggtgag ccccacgttc tgcttcactc tccccatctc ccccccctcc ccacccccaa 660
ttttgtattt atttatttttt taattatttt gtgcagcgat gggggcgggg ggggggggggg 720
ggcgcgcgcc aggcggggcg gggcggggcg aggggcgggg cggggcgagg cggagaggtg 780
cggcggcagc caatcagagc ggcgcgctcc gaaagtttcc ttttatggcg aggcggcggc 840
ggcggcggcc ctataaaaag cgaagcgcgc ggcgggcggg agtcgctgcg cgctgccttc 900
gccccgtgcc ccgctccgcc gccgcctcgc gccgcccgcc ccggctctga ctgaccgcgt 960
tactcccaca ggtgagcggg cgggacggcc cttctcctcc gggctgtaat tagcgcttgg 1020
tttaatgacg gcttgtttct tttctgtggc tgcgtgaaag ccttgagggg ctccgggagc 1080
tagagcctct gctaaccatg ttcatgcctt cttcttttttc ctacagctcc tgggcaacgt 1140
gctggttatt gtgctgtctc atcattttgg caaagaattc tagcggccgc caccatggcc 1200
ttgctcatcc acctcaagac agtctcggag ctgcggggca ggggcgaccg gatcgccaaa 1260
gtgactttcc gagggcaatc cttctactct cgggtcctgg agaactgtga ggatgtggct 1320
gactttgatg agacatttcg gtggccggtg gccagcagca tcgacagaaa tgagatgctg 1380
gagattcagg ttttcaacta cagcaaagtc ttcagcaaca agctcatcgg gaccttccgc 1440
atggtgctgc agaaggtggt agaggagagc catgtggagg tgactgacac gctgattgat 1500
gacaacaatg ctatcatcaa gaccagcctg tgcgtggagg tccggtatca ggccactgac 1560
ggcacagtgg gctcctggga cgatggggac ttcctgggag atgagtctct tcaagaggaa 1620
gagaaggaca gccaagagac ggatggactg ctcccaggct cccggcccag ctcccggccc 1680
ccaggagaga agagcttccg gagagccggg aggagcgtgt tctccgccat gaagctcggc 1740
aaaaaccggt ctcacaagga ggagccccaa agaccagatg aaccggcggt gctggagatg 1800
gaagaccttg accatctggc cattcggcta ggagatggac tggatcccga ctcggtgtct 1860
ctagcctcag tcacagctct caccactaat gtctccaaca agcgatctaa gccagacatt 1920
aagatggagc caagtgctgg gcggcccatg gattaccagg tcagcatcac ggtgatcgag 1980
gcccggcagc tggtgggctt gaacatggac cctgtggtgt gcgtggaggt gggtgacgac 2040

aagaagtaca catccatgaa ggagtccact aactgcccct attacaacga gtacttcgtc 2100
ttcgacttcc atgtctctcc ggatgtcatg tttgacaaga tcatcaagat ttcggtgatt 2160
cactccaaga acctgctgcg cagtggcacc ctggtgggct ccttcaaaat ggacgtggga 2220
accgtgtact cgcagccaga gcaccagttc catcacaagt gggccatcct gtctgacccc 2280
gatgacatct cctcggggct gaagggctac gtgaagtgtg acgttgccgt ggtgggcaaa 2340
ggggacaaca tcaagacgcc ccacaaggcc aatgagaccg acgaagatga cattgagggg 2400
aacttgctgc tccccgaggg ggtgcccccc gaacgccagt gggcccggtt ctatgtgaaa 2460
atttaccgag cagaggggct gccccgtatg aacacaagcc tcatggccaa tgtaaagaag 2520
gctttcatcg gtgaaaacaa ggacctcgtg gacccctacg tgcaagtctt ctttgctggc 2580
cagaagggca agacttcagt gcagaagagc agctatgagc ccctgtggaa tgagcaggtc 2640
gtctttacag acctcttccc cccactctgc aaacgcatga aggtgcagat ccgagactcg 2700
gacaaggtca acgacgtggc catcggcacc cacttcattg acctgcgcaa gatttctaat 2760
gacggagaca aaggcttcct gcccacactg ggcccagcct gggtgaacat gtacggctcc 2820
acacgtaact acacgctgct ggatgagcat caggacctga acgagggcct gggggagggt 2880
gtgtccttcc gggcccggct cctgctgggc ctggctgtgg agatcgtaga cacctccaac 2940
cctgagctca ccagctccac agaggtgcag gtggagcagg ccacgcccat ctcggagagc 3000
tgtgcaggta aaatggaaga attctttctc tttggagcct tcctggaggc ctcaatgatc 3060
gaccggagaa acggagacaa gcccatcacc tttgaggtca ccataggcaa ctatgggaac 3120
gaagttgatg gcctgtcccg gccccagcgg cctcggcccc ggaaggagcc gggggatgag 3180
gaagaagtag acctgattca gaacgcaagt gatgacgagg ccggtgatgc cggggacctg 3240
gcctcagtct cctccactcc accaatgcgg ccccaggtca ccgacaggaa ctacttccat 3300
ctgccctacc tggagcgaaa gccctgcatc tacatcaaga gctggtggcc ggaccagcgc 3360
cgccgcctct acaatgccaa catcatggac cacattgccg acaagctgga agaaggcctg 3420
aacgacatac aggagatgat caaaacggag aagtcctacc ctgagcgtcg cctgcggggc 3480
gtcctggagg agctgagctg tggctgctgc cgcttcctct ccctcgctga caaggaccag 3540
ggccactcat cccgcaccag gcttgaccgg gagcgcctca agtcctgcat gagggagctg 3600
gtaagtatca aggttacaag acaggtttaa ggagaccaat agaaactggg cttgtcgaga 3660
cagagaagac tcttgcgttt ctgagctagc ccccgggtgc gcggcgtcgg tggtgccggc 3720
gggggcgcc aggtcgcagg cggtgtaggg ctccaggcag gcggcgaagg ccatgacgtg 3780
cgctatgaag gtctgctcct gcacgccgtg aaccaggtgc gcctgcgggc cgcgcgcgaa 3840
caccgccacg tcctcgcctg cgtgggtctc ttcgtccagg ggcactgctg actgctgccg 3900
atactcgggg ctcccgctct cgctctcggt aacatccggc cgggcgccgt ccttgagcac 3960
atagcctgga ccgtttcgtc gacctcgagt taagggcgaa ttcccgataa ggatcttcct 4020
agagcatggc tacgtagata agtagcatgg cgggttaatc attaactaca aggaacccct 4080
agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc 4140
aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag 4200
ccttaattaa cctaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg 4260
cgttacccaa cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga 4320
agaggcccgc accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatgggacgc 4380

gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 4440
acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt 4500
cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 4560
tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc 4620
gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 4680
cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 4740
gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc 4800
gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aaatgtgcgc 4860
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 4920
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagcca tattcaacgg 4980
gaaacgtcga ggccgcgatt aaattccaac atggatgctg atttatatgg gtataaatgg 5040
gctcgcgata atgtcgggca atcaggtgcg acaatctatc gcttgtatgg gaagcccgat 5100
gcgccagagt tgtttctgaa acatggcaaa ggtagcgttg ccaatgatgt tacagatgag 5160
atggtcagac taaactggct gacggaattt atgcctcttc cgaccatcaa gcattttatc 5220
cgtactcctg atgatgcatg gttactcacc actgcgatcc ccggaaaaac agcattccag 5280
gtattagaag aatatcctga ttcaggtgaa aatattgttg atgcgctggc agtgttcctg 5340
cgccggttgc attcgattcc tgtttgtaat tgtcctttta acagcgatcg cgtatttcgt 5400
cttgctcagg cgcaatcacg aatgaataac ggtttggttg atgcgagtga ttttgatgac 5460
gagcgtaatg gctggcctgt tgaacaagtc tggaaagaaa tgcataaact tttgccattc 5520
tcaccggatt cagtcgtcac tcatggtgat ttctcacttg ataaccttat ttttgacgag 5580
gggaaattaa taggttgtat tgatgttgga cgagtcggaa tcgcagaccg ataccaggat 5640
cttgccatcc tatggaactg cctcggtgag ttttctcctt cattacagaa acggcttttt 5700
caaaaatatg gtattgataa tcctgatatg aataaattgc agtttcattt gatgctcgat 5760
gagtttttct aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt 5820
cattttaat ttaaaaggat ctaggtgaag atccttttg ataatctcat gaccaaaatc 5880
ccttaacgtg agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct 5940
tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta 6000
ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc 6060
ttcagcagag cgcagatacc aaatactgtt cttctagtgt agccgtagtt aggccaccac 6120
ttcaagaact ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct 6180
gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat 6240
aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac agcccagctt ggagcgaacg 6300
acctacaccg aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa 6360
gggagaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg 6420
gagcttccag ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga 6480
cttgagcgtc gatttttgtg atgctcgtca gggggggcgga gcctatggaa aaacgccagc 6540
aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct 6600
gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct 6660
cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgccca 6720

atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg 6780
tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat 6840
taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc 6900
ggataacaat ttcacacagg aaacagctat gaccatgatt acgccagatt taattaagg 6959

<210> 84
<211> 953
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 合成构建体
<400> 84

aattcggtac cctagttatt aatagtaatc aattacgggg tcattagttc atagcccata 60
tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga 120
cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt 180
ccattgacgt caatgggtgg actatttacg gtaaactgcc cacttggcag tacatcaagt 240
gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca 300
ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt 360
catcgctatt accatggtcg aggtgagccc cacgttctgc ttcactctcc ccatctcccc 420
cccctcccca cccccaattt tgtatttatt tatttttaa ttattttgtg cagcgatggg 480
ggcgggggg ggggggggc gcgcgccagg cggggcgggg cggggcgagg ggcggggcgg 540
ggcgaggcgg agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa agtttccttt 600
tatggcgagg cggcggcggc ggcggcccta taaaaagcga agcgcgcggc gggcgggagt 660
cgctgcgacg ctgccttcgc cccgtgcccc gctccgccgc cgcctcgcgc cgcccgcccc 720
ggctctgact gaccgcgtta ctcccacagg tgagcgggcg ggacggccct tctcctccgg 780
gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc 840
ttgaggggct ccgggagcta gagcctctgc taaccatgtt catgccttct tctttttcct 900
acagctcctg ggcaacgtgc tggttattgt gctgtctcat cattttggca aag 953

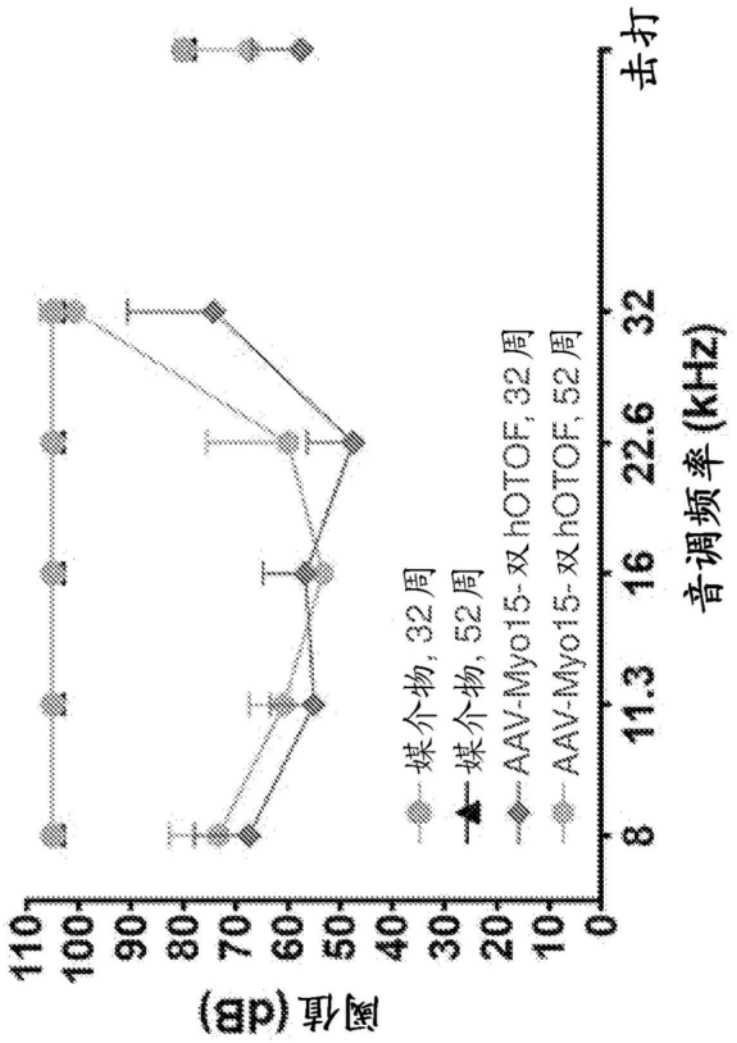

阈值 (dB)
110
100
90
80
70
60
50
40
30
20
10
0
8
11.3
16
22.6
32
音调频率 (kHz)
击打
媒介物, 32 周
媒介物, 52 周
AAV-Myo15-双 hOTOF, 32 周
AAV-Myo15-双 hOTOF, 52 周

图1

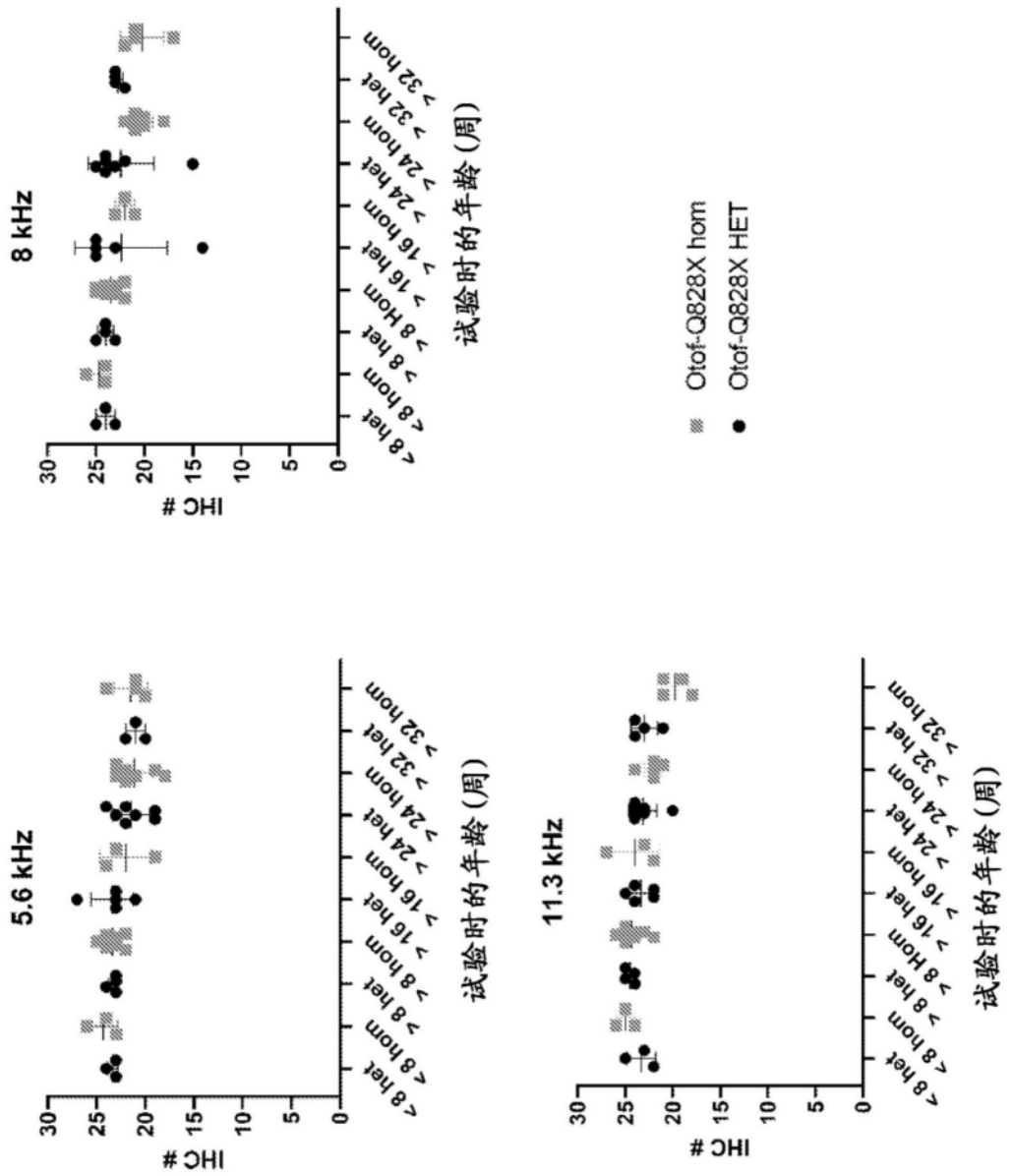


图2A

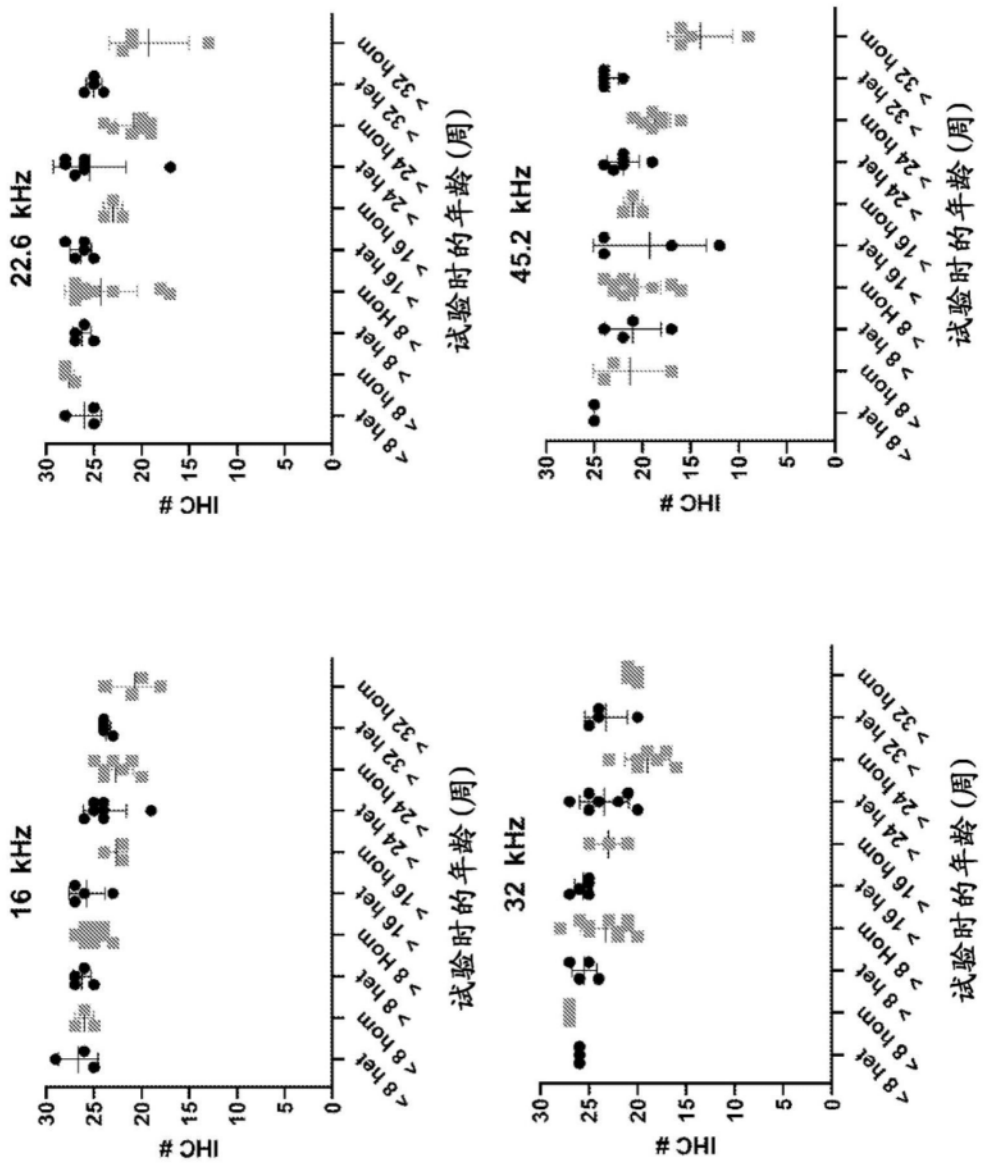

22.6 kHz
45.2 kHz
16 kHz
32 kHz
IHC #
试验时的年龄(周)
<8 het
<8 hom
>8 het
>8 Hom
>16 het
>16 hom
>24 het
>24 hom
>32 het
>32 hom

图2A

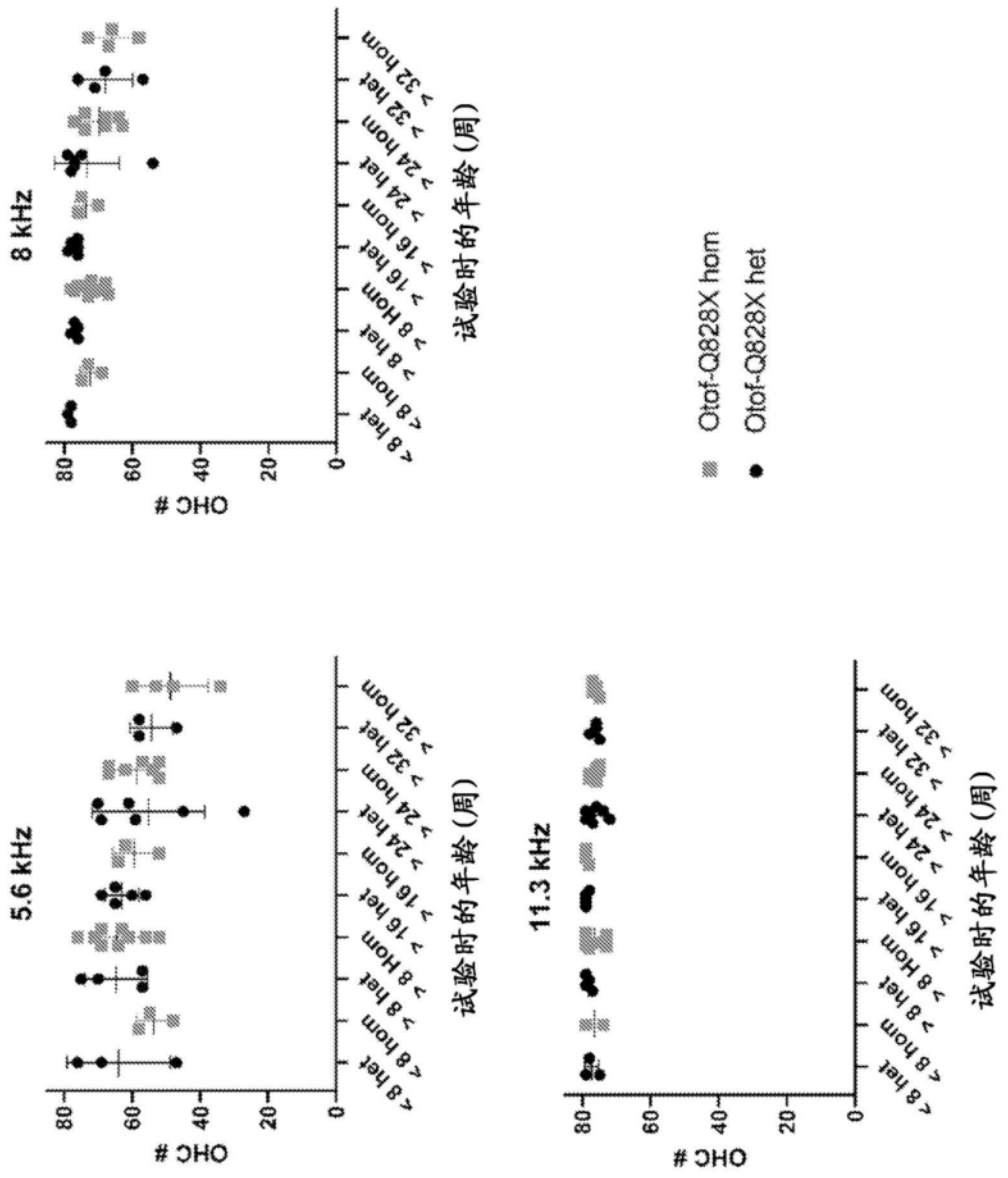

8 kHz
5.6 kHz
11.3 kHz
OHC #
0
20
40
60
80
<8 het
<8 hom
>8 het
>8 Hom
>16 het
>16 hom
>24 het
>24 hom
>32 het
>32 hom
试验时的年龄(周)
Otof-Q828X hom
Otof-Q828X het

图2B

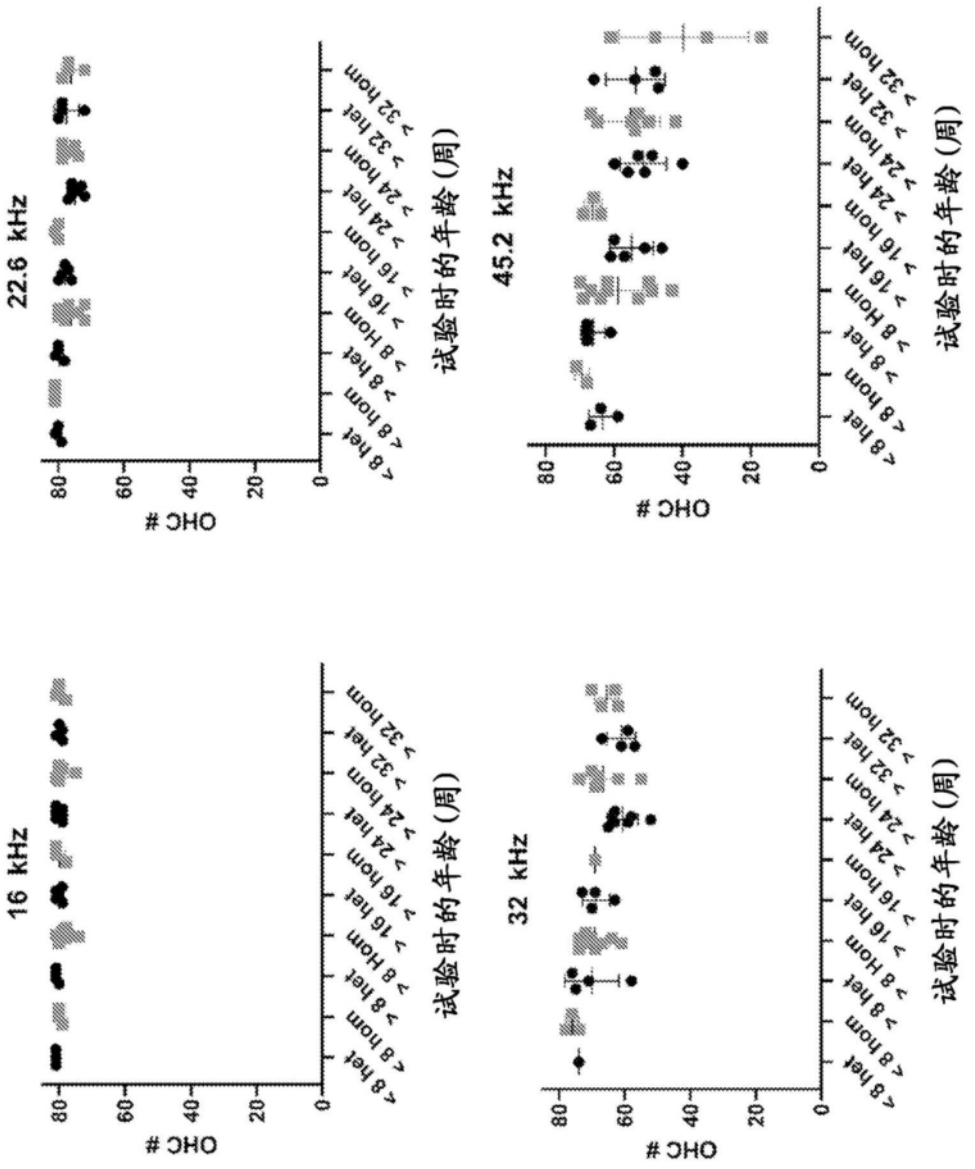

22.6 kHz
45.2 kHz
16 kHz
32 kHz
OHC
试验时的年龄(周)
<8 het
<8 hom
>8 het
>8 Hom
>16 het
>16 hom
>24 het
>24 hom
>32 het
>32 hom

图2B

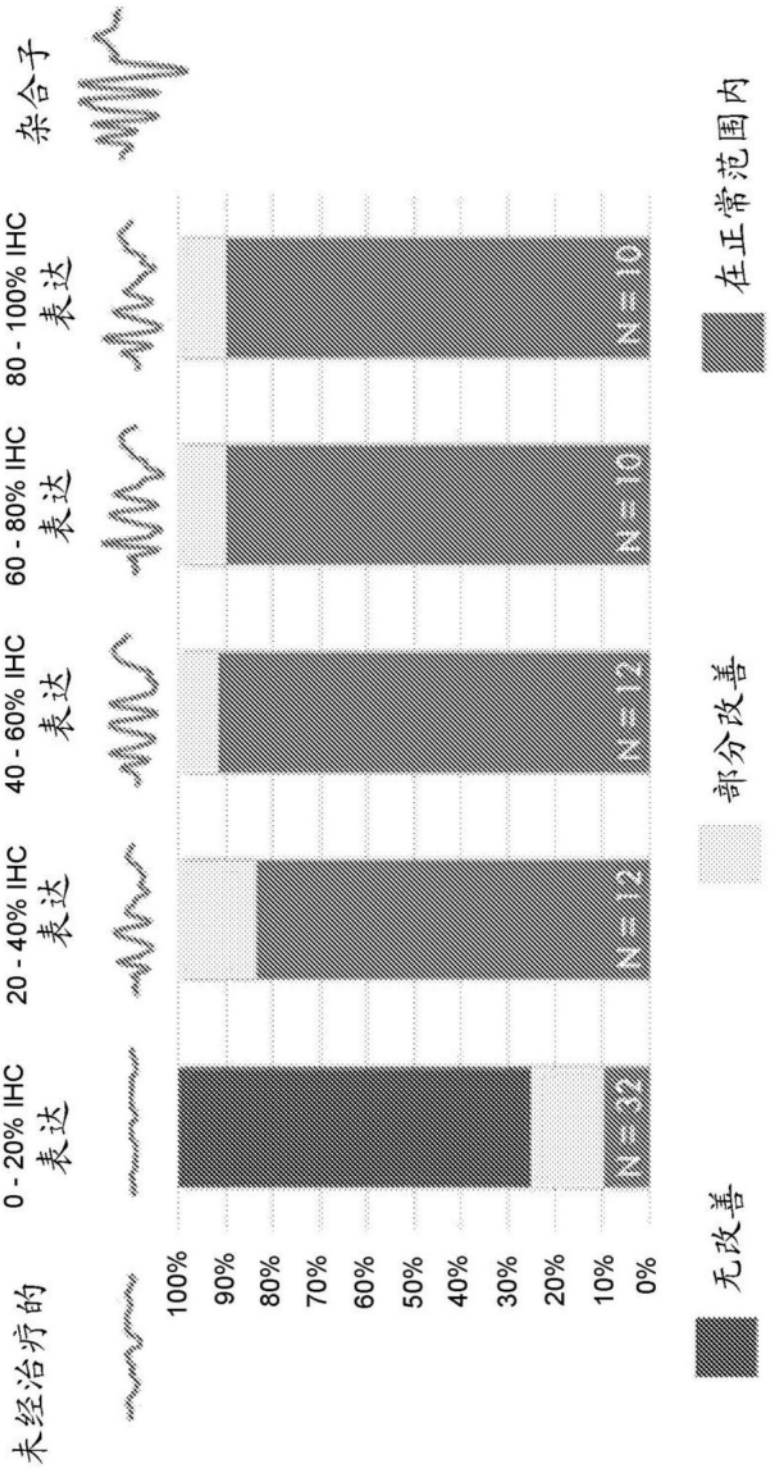

未经治疗的
0 - 20% IHC 表达
20 - 40% IHC 表达
40 - 60% IHC 表达
60 - 80% IHC 表达
80 - 100% IHC 表达
杂合子
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
N = 32
N = 12
N = 12
N = 10
N = 10
无改善
部分改善
在正常范围内

图3