



(12) 发明专利

(10) 授权公告号 CN 111118016 B

(45) 授权公告日 2022.04.15

(21) 申请号 201811280540.4

(22) 申请日 2018.10.30

(65) 同一申请的已公布的文献号
申请公布号 CN 111118016 A

(43) 申请公布日 2020.05.08

(73) 专利权人 上海朗昇生物科技有限公司
地址 201203 上海市浦东新区中国(上海)
自由贸易试验区张衡路1077号2幢2层
B210、B212单元

(72) 发明人 罗学廷 孙晓东

(74) 专利代理机构 上海一平知识产权代理有限公司 31266
代理人 华珊 徐迅

(51) Int. Cl.
C12N 15/12 (2006.01)
C12N 15/864 (2006.01)
C12N 5/10 (2006.01)
A61K 48/00 (2006.01)
A61K 38/17 (2006.01)
A61P 27/02 (2006.01)

(56) 对比文件
CN 101245350 A, 2008.08.20
CN 101610793 A, 2009.12.23
CN 103260649 A, 2013.08.21
Lee YM et al..Homo sapiens prominin 1
(PROM1), transcript variant 6, mRNA, NM_
001145850.1.《GenBank》.2018,

Jon Permanyer et al..Autosomal
Recessive Retinitis Pigmentosa with Early
Macular Affectation Caused by Premature
Truncation in PROM1.《Investigative
Ophthalmology & Visual Science》.2010,第51
卷(第5期),

M. Dominik Fischer et al..Codon-
Optimized RPGR Improves Stability and
Efficacy of AAV8 Gene Therapy in Two
Mouse Models of X-Linked Retinitis
Pigmentosa.《Molecular Therapy》.2017,第25
卷(第8期),

Lee YM et al..Homo sapiens prominin 1
(PROM1), transcript variant 6, mRNA, NM_
001145850.1.《GenBank》.2018,

Jon Permanyer et al..Autosomal
Recessive Retinitis Pigmentosa with Early
Macular Affectation Caused by Premature
Truncation in PROM1.《Investigative
Ophthalmology & Visual Science》.2010,第51
卷(第5期), (续)

审查员 琚羽

权利要求书2页 说明书19页
序列表8页 附图3页

(54) 发明名称
治疗视网膜色素变性疾病的基因治疗载体

(57) 摘要
本发明涉及治疗视网膜色素变性疾病的基
因治疗载体。具体地,对PROM1基因编码序列进行了
针对性特殊优化设计,从而获得了一种特别适
合在哺乳动物(如人)细胞(尤其是感光细胞)中
高效表达PROM1蛋白的核苷酸序列,并构建了一
种表达正常人源PROM1蛋白的重组AAV病毒,并在

PROM1敲除小鼠模型中证明其有效性。实验结果
表明,相对于未优化的编码序列,经过特殊优化
后的PROM1编码序列(SEQ ID NO.:1)的表达量显
著提高了3倍以上,非常适合在哺乳动物(尤其是
人)细胞内表达,能有效治疗视网膜色素变性疾
病等眼部疾病。

CN 111118016 B

[接上页]

(56) 对比文件

M. Dominik Fischer et al..Codon-Optimized RPGR Improves Stability and Efficacy of AAV8 Gene Therapy in Two Mouse Models of X-Linked Retinitis

Pigmentosa.《Molecular Therapy》.2017,第25卷(第8期),

常美会等.翻译速率对重组蛋白表达影响的研究进展.《中国农业科技导报》.2017,第19卷(第4期),

1. 一种多核苷酸,其特征在於,所述多核苷酸编码PROM1蛋白,且所述多核苷酸的序列选自下组:

(a) 所述多核苷酸的序列如SEQ ID NO:1所示;

(b) 与SEQ ID No:1所示核苷酸序列互补的核苷酸序列。

2. 一种表达盒,其特征在於,所述表达盒含有权利要求1所述的多核苷酸。

3. 如权利要求2所述的表达盒,其特征在於,所述表达盒从5' -3' 端具有式I结构:

Z1-Z2-Z3-Z4-Z5 (I)

式中,各“-”独立地为键或核苷酸连接序列;

Z1为无或增强子;

Z2为启动子;

Z3为无或内含子;

Z4为权利要求1所述的多核苷酸;和

Z5为无或polyA。

4. 如权利要求3所述的表达盒,其特征在於,所述增强子为CMV增强子,和/或所述启动子为鸡β肌动蛋白启动子。

5. 如权利要求3所述的表达盒,其特征在於,所述内含子为sv40内含子。

6. 一种载体,其特征在於,所述载体含有如权利要求1所述的多核苷酸或权利要求2所述的表达盒。

7. 一种腺相关病毒载体,其特征在於,所述腺相关病毒载体含有如权利要求1所述的多核苷酸或权利要求2所述的表达盒。

8. 如权利要求7所述的腺相关病毒载体,其特征在於,所述腺相关病毒载体的序列如SEQ ID NO:4所示。

9. 一种宿主细胞,其特征在於,所述宿主细胞含有权利要求6所述的载体或权利要求7所述的腺相关病毒载体,或其染色体中整合有外源的权利要求1所述的多核苷酸或权利要求2所述的表达盒。

10. 如权利要求9所述的宿主细胞,其特征在於,所述宿主细胞选自下组:HEK细胞、感光细胞、双极细胞、水平细胞、视神经细胞、或其组合。

11. 如权利要求10所述的宿主细胞,其特征在於,所述的感光细胞包括视锥细胞和/或视杆细胞。

12. 如权利要求6所述的载体或权利要求7所述的腺相关病毒载体的用途,其特征在於,用于制备一制剂或组合物,所述制剂或组合物用于治疗眼部疾病和/或恢复受试者视力或感光能力。

13. 如权利要求12所述的用途,其特征在於,所述制剂或组合物用于扩大或恢复视网膜的感光细胞功能、恢复受试者视力或感光能力、和/或治疗视网膜退化性疾病。

14. 如权利要求13所述的用途,其特征在於,所述视网膜退化性疾病选自下组:视网膜营养不良、视网膜或黄斑退化、视网膜色素变性、由感光细胞能力的丧失导致的其他疾病、或其组合。

15. 如权利要求14所述的用途,其特征在於,所述视网膜营养不良选自下组:视杆营养不良、视锥营养不良或黄斑营养不良。

16. 如权利要求12所述的用途,其特征在于,所述制剂或组合物用于治疗PROM1基因突变导致的视网膜色素变性疾病。

17. 一种药物制剂,其特征在于,所述的制剂含有 (a) 权利要求6所述的载体或权利要求7所述的腺相关病毒载体,以及 (b) 药学上可接受的载体或赋形剂。

18. 如权利要求17所述的药物制剂,其特征在于,所述药物制剂的剂型选自下组:冻干制剂、液体制剂、或其组合。

19. 如权利要求17所述的药物制剂,其特征在于,所述药物制剂中载体的含量为 1×10^9 - 1×10^{16} 个病毒/毫升。

治疗视网膜色素变性疾病的基因治疗载体

技术领域

[0001] 本发明涉及基因工程技术领域,尤其涉及治疗视网膜色素变性疾病的基因治疗载体。

背景技术

[0002] 视网膜色素变性(retinitis pigmentosa,RP)是一组以遗传性、进行性感光细胞丧失且最终导致视网膜变性萎缩为主要特征的疾病,伴有视网膜色素上皮(RPE)细胞营养不良退行性病变,文献报道在全球范围内RP的患病率约为1/4 000,超过150万患者患有进行性视觉障碍。病变通常从视网膜赤道部开始,并向黄斑和中央凹推进。其主要表现为夜盲、进行性周围视野下降等,因感光细胞死亡导致视力下降。眼底可见视网膜血管变细呈白线样狭窄,黄斑区萎缩样改变,始于赤道部的骨细胞样色素沉着并逐渐向后极部发展。视网膜电图可见a、b波振幅下降,潜伏期延长。患者还可伴有听力丧失、代谢异常、神经病变以及肝脏或肾脏病变等全身表现。

[0003] 随着科学技术的不断进步,人们对RP的疾病特征已有一定了解,感光细胞的死亡是RP的主要病理学特征,最终将导致患者失明。既往研究发现RPE细胞对感光细胞外节脱落膜盘的吞噬功能下降,未被吞噬的膜盘在视网膜外层堆积,破坏了原有的视网膜结构,导致感光细胞变性死亡,从而引发RP。其致病机制主要为基因突变所导致的相应正常蛋白功能异常或缺失。由于其显著的遗传异质性及复杂的致病机制,目前尚无明确的针对此病的治疗手段。

[0004] RP有多种遗传方式,常染色体显性遗传、常染色体隐性遗传及X染色体连锁遗传,目前为止已发现64个基因与RP的发生相关(<https://sph.uth.edu/retnet/sum-dis.htm>),包括PROM1。随着人类基因图谱成功构建、分子生物学技术迅猛发展,病毒载体成功应用,使得RP的基因治疗方面已经取得了一定进展。很多实验研究为RP的基因疗法应用于临床做了很多铺垫和前期工作。腺相关病毒(Adeno-associated virus,AAV)较其他病毒载体小,无致病性,可转染正在分裂和未分裂的细胞等特性,基于AAV载体的针对眼部特别是遗传性视网膜退行性病变的基因治疗方法受到了广泛的关注。

[0005] 因此,本领域急需开发能有效治疗遗传性视网膜退行性病变的基因治疗方法和治疗药物。

发明内容

[0006] 本发明的目的是提供一种能有效治疗遗传性视网膜退行性病变的基因治疗药物。

[0007] 本发明的另一目的是提供一种表达正常人源PROM1蛋白的重组AAV病毒。

[0008] 本发明的第一方面,提供了一种核苷酸序列,所述核苷酸序列编码PROM1蛋白,且所述核苷酸序列选自下组:

[0009] (a) 所述核苷酸序列如SEQ ID NO.:1所示;和

[0010] (b) 所述核苷酸序列与SEQ ID NO.:1所示的核苷酸序列有 $\geq 95\%$ 相同性,优选地

≥98%，更优选地≥99%；

[0011] (c) 与(a)或(b)所述的核苷酸序列互补的核苷酸序列。

[0012] 在另一优选例中，所述核苷酸序列包括DNA序列、cDNA序列、或mRNA序列。

[0013] 在另一优选例中，所述核苷酸序列包括单链序列和双链序列。

[0014] 在另一优选例中，所述核苷酸序列包括与SEQ ID NO.:1完全互补的核苷酸序列。

[0015] 本发明的第二方面，提供了一种表达盒，所述表达盒含有本发明第一方面所述的核苷酸序列。

[0016] 在另一优选例中，所述表达盒从5' -3' 端具有式I结构：

[0017] Z1-Z2-Z3-Z4-Z5 (I)

[0018] 式中，各“-”独立地为键或核苷酸连接序列；

[0019] Z1为无或增强子；

[0020] Z2为启动子；

[0021] Z3为无或内含子；

[0022] Z4为本发明第一方面所述的核苷酸序列；和

[0023] Z5为无或polyA。

[0024] 在另一优选例中，所述增强子为CMV增强子。

[0025] 在另一优选例中，所述启动子为鸡β肌动蛋白启动子(CBA启动子)。

[0026] 在另一优选例中，所述内含子为sv40内含子。

[0027] 在另一优选例中，所述polyA为sv40polyA。

[0028] 在另一优选例中，各个核苷酸连接序列的长度为0-30nt，较佳地1-15nt。

[0029] 本发明的第三方面，提供了一种载体，所述载体含有如本发明第一方面所述的核苷酸序列或本发明第二方面所述的表达盒。

[0030] 在另一优选例中，所述载体包含一个或多个启动子，所述启动子可操作地与所述核苷酸序列、增强子、内含子、转录终止信号、多腺苷酸化序列、复制起点、选择性标记、核酸限制性位点、和/或同源重组位点连接。

[0031] 在另一优选例中，所述的载体包括质粒、病毒载体。

[0032] 在另一优选例中，所述的载体包括DNA病毒、逆转录病毒载体。

[0033] 在另一优选例中，所述的载体选自下组：慢病毒载体、腺病毒载体、腺相关病毒载体(AAV)、或其组合。较佳地，所述载体为AAV载体。

[0034] 在另一优选例中，所述载体为含有或插入有本发明第一方面所述的核苷酸序列的AAV载体。

[0035] 在另一优选例中，所述载体用于表达人PROM1蛋白。

[0036] 本发明的第四方面，提供了一种腺相关病毒载体，所述腺相关病毒载体含有如本发明第一方面所述的核苷酸序列或本发明第二方面所述的表达盒。

[0037] 在另一优选例中，所述腺相关病毒的血清型选自：AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、rh10或其组合。

[0038] 在另一优选例中，所述腺相关病毒载体用于治疗眼部疾病和/或恢复受试者视力或感光能力。

[0039] 在另一优选例中，所述腺相关病毒载体的序列如SEQ ID NO.:4所示。

- [0040] 其中,第1-141位为左ITR序列;
- [0041] 第153-532位为CMV增强子;
- [0042] 第535-810位为鸡 β 肌动蛋白启动子(CBA启动子);
- [0043] 第813-1006位为SV40内含子;
- [0044] 第1019-3523位为优化的人PROM1编码序列;
- [0045] 第3530-3757位为SV40PolyA;
- [0046] 第3758-3898位为右ITR序列。
- [0047] 本发明的第五方面,提供了一种宿主细胞,所述宿主细胞含有本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体,或其染色体中整合有外源的本发明第一方面所述的核苷酸序列或本发明第二方面所述的表达盒。
- [0048] 在另一优选例中,所述宿主细胞为哺乳动物细胞,所述哺乳动物包括人和非人哺乳动物。
- [0049] 在另一优选例中,所述宿主细胞选自下组:HEK细胞、感光细胞(包括视锥细胞和/或视杆细胞)、其他视觉细胞(如双极细胞、水平细胞)、(视)神经细胞、或其组合。
- [0050] 在另一优选例中,所述宿主细胞选自下组:视杆细胞、视锥细胞、给光双极细胞、撤光双极细胞、水平细胞、神经节细胞、无长突细胞、或其组合。
- [0051] 在另一优选例中,所述宿主细胞为感光细胞(即光感受器细胞)。
- [0052] 本发明的第六方面,提供了如本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体的用途,用于制备一制剂或组合物,所述制剂或组合物用于治疗眼部疾病和/或恢复受试者视力或感光能力。
- [0053] 在另一优选例中,所述制剂或组合物用于扩大或恢复视网膜的感光细胞功能、恢复受试者视力(或感光能力)、和/或治疗视网膜退化性疾病。
- [0054] 在另一优选例中,所述视网膜退化性疾病选自下组:视网膜营养不良(如视杆营养不良、视杆视锥营养不良、视锥视杆营养不良、视锥营养不良或黄斑营养不良)、视网膜或黄斑退化、视网膜色素变性、由感光细胞能力的丧失导致的其他疾病、或其组合。
- [0055] 在另一优选例中,所述制剂或组合物用于治疗视网膜色素变性疾病,较佳地治疗PROM1基因突变导致的视网膜色素变性疾病。
- [0056] 本发明的第七方面,提供了一种药物制剂,所述的制剂含有(a)本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体,以及(b)药学上可接受的载体或赋形剂。
- [0057] 在另一优选例中,所述药物制剂的剂型选自下组:冻干制剂、液体制剂、或其组合。
- [0058] 在另一优选例中,所述药物制剂中载体的含量为 1×10^9 - 1×10^{16} 个病毒/毫升,较佳地 1×10^{12} - 1×10^{13} 个病毒/毫升。
- [0059] 在另一优选例中,所述药物制剂用于治疗眼部疾病和/或恢复受试者视力或感光能力。
- [0060] 在另一优选例中,所述眼部疾病由PROM1基因突变引起。
- [0061] 在另一优选例中,所述药物制剂用于扩大或恢复视网膜的感光细胞功能、恢复受试者视力(或感光能力)、和/或治疗视网膜退化性疾病。
- [0062] 在另一优选例中,所述药物制剂用于治疗视网膜色素变性疾病,较佳地治疗PROM1

基因突变导致的视网膜色素变性疾病。

[0063] 本发明的第八方面,提供了一种治疗方法,所述方法包括将本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体施用于需要的对象。

[0064] 在另一优选例中,将所述腺相关病毒载体引入到需要的对象的眼睛内。

[0065] 在另一优选例中,所述需要的对象包括人和非人哺乳动物。

[0066] 在另一优选例中,所述治疗方法为治疗眼部疾病的方法。

[0067] 在另一优选例中,所述眼部疾病由PROM1基因突变引起。

[0068] 在另一优选例中,所述眼部疾病为视网膜退化性疾病,较佳地视网膜色素变性疾病。

[0069] 在另一优选例中,所述视网膜退化性疾病选自下组:视网膜营养不良(如视杆营养不良、视杆视锥营养不良、视锥视杆营养不良、视锥营养不良或黄斑营养不良)、视网膜或黄斑退化、视网膜色素变性、由感光细胞能力的丧失导致的其他疾病、或其组合。

[0070] 在另一优选例中,所述治疗方法用于扩大或恢复视网膜的感光细胞功能、恢复受试者视力或感光能力。

[0071] 本发明的第九方面,提供了一种PROM1蛋白的制备方法,包括培养本发明第五方面所述的宿主细胞,从而得到PROM1蛋白。

[0072] 应理解,在本发明范围内中,本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合,从而构成新的或优选的技术方案。限于篇幅,在此不再一一累述。

附图说明

[0073] 图1显示了rAAV/CBA.opti-hPROM1载体的结构示意图。

[0074] 图2显示了序列优化的opti-hPROM1表达效率比原始hPROM1高。图A为Western Blot蛋白电泳结果,泳道1:对照HEK-293细胞;泳道2:rAAV/CBA.hPROM1转染HEK-293细胞;泳道3:rAAV/CBA.opti-hPROM1转染HEK-293细胞。B为rAAV/CBA.opti-hPROM1优化的与rAAV/CBA.hPROM1的相对表达量,图B为对图A中信号强度的量化,把泳道1信号设定为1,泳道2和泳道3以1为参考的比对值。

[0075] 图3显示了小鼠视网膜电图比对结果。其中图A从上往下分别为正常鼠、rAAV/CBA.opti-hPROM1注射的PROM1 KO鼠、rAAV/CBA.hPROM1注射的PROM1 KO鼠及未注射的PROM1 KO鼠暗适应下视网膜电图对比。图B为正常鼠、rAAV/CBA.opti-hPROM1注射的PROM1 KO鼠、rAAV/CBA.hPROM1注射的PROM1 KO鼠及未注射的PROM1 KO鼠暗适应下视网膜电图a波峰值对比,注射眼的a波峰值远高于未注射眼。其中rAAV/CBA.opti-hPROM1与正常对照相比没有显著性差异,而rAAV/CBA.hPROM1的治疗效果不及rAAV/CBA.opti-hPROM1,并存在显著性差异($p < 0.01$)。

[0076] 图4显示了小鼠视网膜荧光染色结果。其中,A:正常C57小鼠左眼在生后14天时注射治疗载体,注射后3周进行视网膜免疫组化检测;B:右眼未注射眼;标尺=2微米;C和D:hPROM1表达检测,标尺=100微米。

[0077] 图5显示了小鼠视网膜荧光染色结果。图A、B和C分别为正常小鼠、注射rAAV/CBA.opti-hPROM1的PROM1 KO小鼠和对照PROM1 KO小鼠(未注射)的视网膜切片DAPI染色。

注射rAAV/CBA.opti-hPROM1的PROM1 KO小鼠(图B)网膜厚度显著大于未注射的PROM1 KO小鼠(图C)。

具体实施方式

[0078] 本发明人经过广泛而深入的研究,对PROM1基因编码序列进行了针对性特殊优化设计,从而获得了一种特别适合在哺乳动物(如人)细胞(尤其是感光细胞)中高效表达PROM1蛋白的核苷酸序列,并构建了一种表达正常人源PROM1蛋白的重组AAV病毒,并在PROM1敲除小鼠模型中证明其有效性。实验结果表明,相对于未优化的编码序列,经过特殊优化后的PROM1编码序列(SEQ ID NO.:1)的表达量显著提高了3倍以上,非常适合在哺乳动物(尤其是人)细胞内表达,能有效治疗视网膜色素变性疾病等眼部疾病。在此基础上,发明人完成了本发明。

[0079] 术语

[0080] 为了可以更容易地理解本公开,首先定义某些术语。如本申请中所使用的,除非本文另有明确规定,否则以下术语中的每一个应具有下面给出的含义。在整个申请中阐述了其它定义。

[0081] 术语“约”可以是指在本领域普通技术人员确定的特定值或组成的可接受误差范围内的值或组成,其将部分地取决于如何测量或测定值或组成。例如,如本文所用,表述“约100”包括99和101和之间的全部值(例如,99.1、99.2、99.3、99.4等)。

[0082] 如本文所用,术语“含有”或“包括(包含)”可以是开放式、半封闭式和封闭式的。换言之,所述术语也包括“基本上由…构成”、或“由…构成”。

[0083] 序列同一性通过沿着预定的比较窗(其可以是参考核苷酸序列或蛋白的长度的50%、60%、70%、80%、90%、95%或100%)比较两个对齐的序列,并且确定出现相同的残基的位置的数目来确定。通常地,这表示为百分比。核苷酸序列的序列同一性的测量是本领域技术人员熟知的方法。

[0084] 如本文使用的,术语“受试者”、“需要的对象”指任何哺乳动物或非哺乳动物。哺乳动物包括但不限于人类、脊椎动物诸如啮齿类、非人类灵长类、牛、马、狗、猫、猪、绵羊、山羊。

[0085] 如本文所用,术语“光感受器”、“感光细胞”和“光感受细胞”可互换使用,包括视杆细胞和视锥细胞。

[0086] PROM1

[0087] 如本文所用,术语“PROM1蛋白”、“多肽”、“本发明蛋白”、“人PROM1蛋白”具有相同的意义,在本文中可互转使用。

[0088] PROM1基因编码的蛋白(PROM1蛋白)为一种跨膜的糖蛋白,一直以来被当做造血干细胞的生物标记,其在生物体的许多组织中均有表达,在视网膜上其编码蛋白位于感光细胞外节基底部及连接纤毛,其在感光细胞外节膜盘的形成中起了重要作用。既往研究中PROM1基因敲除小鼠可出现类似RP的疾病表现,而突变PROM1基因敲入小鼠则出现感光细胞外节膜盘发育异常。

[0089] 核酸编码序列

[0090] 本发明的要解决的技术问题是克服现有技术中PROM1表达量不高、治疗效果不佳

的技术缺陷。本发明的目的是提供一种PROM1优化基因序列。本发明优化后的PROM1编码序列如SEQ ID NO:1所示,其大小为2505bp。经研究发现,本发明优化的PROM1编码序列(SEQ ID NO.:1),PROM1蛋白表达效率更高,有更多的PROM1蛋白在患者视网膜感光细胞发挥生理作用。

[0091] 本发明所述的编码PROM1蛋白的核酸,其核苷酸序列如SEQ ID NO.:1所示。在另一优选例中,所述核苷酸序列与SEQ ID NO.:1所示的核苷酸序列有 $\geq 95\%$ 相同性,优选地 $\geq 98\%$,更优选地 $\geq 99\%$ 。在本发明中,所述编码PROM1蛋白的核酸又称作PROM1优化基因、PROM1优化核酸或opti-hPROM1。

[0092] 本发明的多核苷酸可以是DNA形式或RNA形式。在另一优选例中,所述核苷酸为DNA。DNA形式包括cDNA、基因组DNA或人工合成的DNA。DNA可以是单链的或是双链的。DNA可以是编码链或非编码链。本发明所述的核苷酸序列编码SEQ ID NO.:3所示的氨基酸序列。

[0093] PROM1蛋白的NCBI参考序列号:NP_001139322.1,氨基酸序列如SEQ ID NO.:3所示。

[0094]

```

1 malvlgslll lglcgnsfsg gqpsstdapk awnyelpatn yetqdshkag pigilfelvh
61 iflyvvqprd fpedtlrkfl qkayeskidy dkpetvilgl kivyyeagii lccvlgllfi
121 ilmplvgyff cmcrcnckcg gemhqrqken gpflrkcfai sllviciiis igifygfvan
181 hqvrtrikrs rkladsnfkd lrllnetpe qikyilaqyn ttkdkaftdl nsinsvlggg
241 ildrlrpni pvldeiksm taiketkeal enmstlksl hqqstqlsss ltsvktslrs
301 slndplclvh pssetensir lslsqlnsnp elrqlppvda eldnvnnvlr tdldglvqqg
361 yqslndipdr vqrqtttva gikrvlnsig sdidnvtqrl piqdilsafs vyvnntesyi
421 hrnlptleey dsywwlgglv icslltlivi fyyllgllcgv cgydrhatpt trgcvsntgg
481 vflmvgvpls flfcwilmii vltfvfgan veklicept skelfrvldt pyllnedwey
541 ylsyklfnks kmkltfeqvy sdckknrgty gtlhlqnsfn isehlnineh tgsisseles
601 lkvnlnifll gaagrknld faacgidrmn ydsylaqtgk spagvnllsf aydleakans
661 lppgnlrnsl krdaqtikti hqqrvpieq slstlyqsvk ilqrtgngll ervtrilasl
721 dfaqnfitnn tssviieetk kygrtiigyf ehylqwiefs isekvasckp vataldtavd
781 vflcsyiidp lnlfwfgigk atvflpali favklakyyr rmdsedvydd psqh (SEQ ID NO.: 3)

```

[0095] 核酸序列可以是DNA、RNA、cDNA或PNA。核酸序列可以是基因组的、重组的或合成的。核酸序列可以是分离的或纯化的。核酸序列可以是单链或双链的。优选地,核酸序列将编码如本文描述的PROM1蛋白。核酸序列可以通过克隆衍生,例如使用包括限制性酶切、连接、凝胶电泳的标准的分子克隆技术,例如在Sambrook等Molecular Cloning:A laboratory manual,Cold Spring Harbour Laboratory Press)中描述的。核酸序列可是分离的,例如使用PCR技术分离的。分离意指从任何杂质和从被自然地发现与其来源中的核酸序列缔合的其他核酸序列和/或蛋白中分离核酸序列。优选地,其还将不含细胞材料、培养基或来自纯化/生产过程的其他化学物质。核酸序列可以是合成的,例如通过直接的化学合成产生。核酸序列可以作为裸露的核酸被提供,或可与蛋白或脂质复合提供。

[0096] 本发明的核苷酸全长序列或其片段通常可以用PCR扩增法、重组法或人工合成的方法获得。对于PCR扩增法,可根据已公开的有关核苷酸序列,尤其是开放阅读框序列来设计引物,并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板,扩增而得有关序列。当序列较长时,常常需要进行两次或多次PCR扩增,然后再将各次扩增出的片段按正确次序拼接在一起。目前,已经可以完全通过化学合成来得到编码本发明多肽(或其片段,或其衍生物)的DNA序列。然后可将该DNA序列引入本领域中已知的各种现

有的DNA分子(或如载体)和细胞中。

[0097] 本发明也涉及包含本发明的多核苷酸的载体,以及用本发明的载体或多肽编码序列经基因工程产生的宿主细胞。上述多核苷酸、载体或宿主细胞可以是分离的。

[0098] 如本文所用,“分离的”是指物质从其原始环境中分离出来(如果是天然物质,原始环境即是天然环境)。如活体细胞内的天然状态下的多核苷酸和多肽是没有分离纯化的,但同样的多核苷酸或多肽如从天然状态中同存在的其他物质中分开,则为分离纯化的。

[0099] 在本发明较佳的实施方式中,所述核苷酸序列如SEQ ID NO.:1所示。

[0100] 一旦获得了有关的序列,就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体,再转入细胞,然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

[0101] 此外,还可用人工合成的方法来合成有关序列,尤其是片段长度较短时。通常,通过先合成多个小片段,然后再进行连接可获得序列很长的片段。

[0102] 应用PCR技术扩增DNA/RNA的方法被优选用于获得本发明的基因。用于PCR的引物可根据本文所公开的本发明的序列信息适当地选择,并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的DNA/RNA片段。

[0103] 本发明也涉及包含本发明的多核苷酸的载体,以及用本发明的载体或蛋白编码序列经基因工程产生的宿主细胞,以及经重组技术利用所述宿主细胞表达PROM1蛋白的方法。

[0104] 通过常规的重组DNA技术,可利用本发明的多核苷酸序列获得表达本发明PROM1蛋白的宿主细胞(如哺乳动物细胞)。一般来说包括步骤:将本发明第一方面所述的多核苷酸或本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体转导入宿主细胞内。

[0105] 本领域的技术人员熟知的方法能用于构建含本发明多肽的编码DNA序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组DNA技术、DNA合成技术、体内重组技术等。所述的DNA序列可有效连接到表达载体中的适当启动子上,以指导mRNA合成。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

[0106] 此外,表达载体优选地包含一个或多个选择性标记基因,以提供用于选择转化的宿主细胞的表型性状,如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP),或用于大肠杆菌的四环素或氨苄青霉素抗性。

[0107] 包含上述的适当DNA序列以及适当启动子或者控制序列的载体,可以用于转化适当的宿主细胞,以使其能够表达多肽。

[0108] 宿主细胞可以是原核细胞,或是低等真核细胞,或是高等真核细胞,如哺乳动物细胞(包括人和非人哺乳动物)。代表性例子有:CHO、NS0、COS7、或293细胞的动物细胞等。在本发明的一个优选实施方式中,选择293T细胞、感光细胞(包括锥状细胞和/或杆状细胞)、其他视觉细胞(如双节细胞)、神经细胞为宿主细胞。在另一优选例中,所述宿主细胞选自下组:视杆细胞、视锥细胞、给光双极细胞、撤光双极细胞、水平细胞、神经节细胞、无长突细胞、或其组合。

[0109] 用重组DNA转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时,能吸收DNA的感受态细胞可在指数生长期后收获,用CaCl₂法处理,所用的步骤在本领域众所周知。另一种方法是使用MgCl₂。如果需要,转化也可用电穿孔的方法进行。当宿主是真核生物,可选用如下的DNA转染方法:磷酸钙共沉淀法,常规机械方法如

显微注射、电穿孔、脂质体包装等。

[0110] 获得的转化子可以用常规方法培养,表达本发明的基因所编码的蛋白质。根据所用的宿主细胞,培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后,用合适的方法(如温度转换或化学诱导)诱导选择的启动子,将细胞再培养一段时间。

[0111] 在上面的方法中的多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要,可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但不限于:常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

[0112] 序列优化

[0113] 在本发明中,提供了优化的、特别适合在哺乳动物细胞中表达的PROM1的编码序列,所述编码序列如SEQ ID NO.:1所示。

[0114] 如本文所用,所述“优化的PROM1编码序列”、“优化PROM1编码基因”均指一种用于编码PROM1的核苷酸序列,所述的核苷酸序列编码SEQ ID NO.:3所示的氨基酸序列。

[0115] 在本发明中,所述PROM1的野生DNA编码序列(未优化的DNA编码序列)如SEQ ID NO.:2所示,所述未优化的野生DNA编码序列的表达量很低。

[0116] PROM1野生编码序列来自NCBI参考序列:NM_001145850.1的编码区CDS:213-2717,具体核酸序列如SEQ ID NO.:2所示。

[0117] ATGGCCCTCGTACTCGGCTCCCTGTTGCTGCTGGGGCTGTGCGGAACTCCTTTTCAGGAGGGCAGCC
TTCATCCACAGATGCTCCTAAGGCTTGAATTATGAATTGCCTGCAACAAATTATGAGACCCAAGACTCCCATAAA
GCTGGACCCATTGGCATTCTCTTTGAACTAGTGCATATCTTTCTCTATGTGGTACAGCCGCGTGATTTCCCAGAAG
ATACTTTGAGAAAATTCTTACAGAAGGCATATGAATCCAAAATTGATTATGACAAGCCAGAACTGTAATCTTAGG
TCTAAAGATTGTCTACTATGAAGCAGGGATTATTCTATGCTGTGTCCTGGGGCTGCTGTTTATTATTCTGATGCCT
CTGGTGGGGTATTTCTTTTGTATGTGTCGTTGCTGTAACAAATGTGGTGGAGAAATGCACCAGCGACAGAAGGAAA
ATGGGCCCTTCTGAGGAAATGCTTTGCAATCTCCCTGTTGGTGATTTGTATAATAAAGCATTGGCATCTTCTA
TGGTTTTGTGGCAAAATCACCAGGTAAGAACCCGGATCAAAAGGAGTCGGAAACTGGCAGATAGCAATTTCAAGGAC
TTGCGAACTCTCTGAATGAACTCCAGAGCAAATCAAATATATATTGGCCCAGTACAACACTACCAAGGACAAGG
CGTTCACAGATCTGAACAGTATCAATTCAGTGCTAGGAGGCGGAATTTCTTGACCGACTGAGACCCAACATCATCCC
TGTTCTTGATGAGATTAAGTCCATGGCAACAGCGATCAAGGAGACCAAAGAGGCGTTGGAGAACATGAACAGCACC
TTGAAGAGCTTGACCAACAAAGTACACAGCTTAGCAGCAGTCTGACCAGCGTGAAAAGTACCTGCGGTCATCTC
TCAATGACCCTCTGTGCTTGGTGCATCCATCAAGTGAACCTGCAACAGCATCAGATTGTCTCTAAGCCAGCTGAA
TAGCAACCCTGAACTGAGGCAGCTTCCACCCGTTGATGCAGAACTTGACAACGTTAATAACGTTCTTAGGACAGAT
TTGGATGGCCTGGTCCAACAGGGCTATCAATCCCTTAATGATATACCTGACAGAGTACAACGCCAAACCACGACTG
TCGTAGCAGGTATCAAAAGGGTCTTGAATTCATTGGTTTCAAGATATCGACAATGTAACCTCAGCGTCTTCTATTCA
GGATATACTCTCAGCATTCTCTGTTTATGTTAATAACACTGAAAGTTACATCCACAGAAATTTACCTACATTGGAA
GAGTATGATTACTACTGGTGGCTGGGTGGCCTGGTCATCTGCTCTCTGCTGACCCTCATCGTGATTTTTACTACC
TGGGCTTACTGTGTGGCGTGTGCGGCTATGACAGGCATGCCACCCCGACCACCCGAGGCTGTGTCTCCAACACCGG
AGGCGTCTTCTCATGGTTGGAGTTGGATTAAGTTTCTCTTTTCTGCTGGATATTGATGATCATTGTGGTTCTTACC

TTTGTCTTTGGTGCAAATGTGGAAAACTGATCTGTGAACCTTACACGAGCAAGGAATTATCCGGGTTTTGGATA
CACCTACTTACTAAATGAAGACTGGGAATACTATCTCTCTGGGAAGCTATTTAATAAAATCAAAAATGAAGCTCAC
TTTTGAACAAGTTTACAGTGACTGCAAAAAAATAGAGGCACTTACGGCACTCTTCACCTGCAGAACAGCTTCAAT
ATCAGTGAACATCTCAACATTAATGAGCATACTGGAAGCATAAGCAGTGAATTGGAAAAGTCTGAAGGTAAATCTTA
ATATCTTTCTGTTGGGTGCAGCAGGAAGAAAAACCTTCAGGATTTTGTCTGCTTGTGGAATAGACAGAATGAATTA
TGACAGCTACTTGGCTCAGACTGGTAAATCCCCCGCAGGAGTGAATCTTTTATCATTTGCATATGATCTAGAAGCA
AAAGCAAACAGTTTGCCCCAGGAAATTTGAGGAACTCCCTGAAAAGAGATGCACAACTATTAACAATTCCACC
AGCAACGAGTCCTTCTATAGAACAATCACTGAGCACTCTATACCAAAGCGTCAAGATACTTCAACGCACAGGGAA
TGGATTGTTGGAGAGAGTAACTAGGATTCTAGCTTCTCTGGATTTTGTCTCAGAACTTCATCACAAACAATACTTCC
TCTGTTATTATTGAGGAACTAAGAAGTATGGGAGAACAATAATAGGATATTTTGAACATTATCTGCAGTGGATCG
AGTTCTCTATCAGTGAGAAAGTGGCATCGTGCAAACCTGTGGCCACCGCTCTAGATACTGCTGTTGATGTCTTTCT
GTGTAGCTACATTATCGACCCCTGAATTTGTTTTGGTTTGGCATAGGAAAAGCTACTGTATTTTTACTTCCGGCT
CTAATTTTTGCGGTAAAACTGGCTAAGTACTATCGTCAATGGATTTCGGAGGACGTGTACGATGACCCATCACAAC
ATTGA (SEQ ID NO.:2)

[0118] 本发明优化了影响基因表达的序列片段,这些序列片段包括但不限于,密码子使用偏好性,消除不利于表达的二级结构(如发夹结构),改变GC含量,CpG二核苷酸含量,mRNA的二级结构,隐蔽剪接位点,早期多聚腺苷化位点,内部核糖体进入位点和结合位点,负CpG岛,RNA不稳定区,重复序列(直接重复、反向重复等)和可能影响克隆的限制性位点。通过分析和试验筛选,最终得到如SEQ ID NO.:1所示的特别优化的DNA编码序列。特别优化后的SEQ ID NO.:1所示的编码序列与SEQ ID NO.:2所示的野生编码序列相似度为75% (1879/2505)。

[0119] ATGGCCCTGGTGTGGGGAGCCTGCTGCTGCTGGGGCTGTGCGGAACTCCTTCTCCGGGGCCAGCC
CTCCAGCACCGACGCTCCTAAGGCCTGGAACACTACGAGCTGCCCCGCCACCACTACGAAACCCAAGACTCCCACAAA
GCCGGCCCCATCGGCATCCTGTTTGAACCTCGTGCATATTTTCTCTACGTGGTTCAACCCAGAGATTTTCCCGAGG
ACACCCTGAGAAAAGTTCTTGCAGAAGGCCTATGAGAGCAAGATTGACTACGACAAGCCCAGAACCGTGATCCTGGG
CCTGAAGATCGTGATTATGAGGCCGGAATTATCCTCTGTTGCGTGCTGGGCCTGCTGTTTATCATCCTGATGCCA
CTGGTGGGCTACTTCTTTTGCATGTGCAGATGTTGCAACAAGTGTGGCGGCGAGATGCACCAGCGTCAGAAGGAGA
ACGGGCCTTTCTCCGAAATGCTTTGCCATCTCCCTGCTGGTGATTTGTATCATTATCAGCATCGGGATCTTCTA
CGGATTCGTGGCTAACCATCAGGTCAGAACCCGCATCAAGCGCAGTAGAAAAGCTGGCCGACTCCAACCTCAAGGAC
CTGCGGACCCTGCTGAACGAGACCCCCGAGCAGATCAAGTACATTCTGGCCCAATACAACACCACCAAGGACAAAG
CCTTACAGACCTGAACTCCATCAACAGCGTGCTCGGCGGAGGCATACTGGACCGGCTGAGACCCAACATAATACC
CGTGCTGGACGAAATCAAAAAGCATGGCCACCGCCATAAAGGAGACCAAAGAAGCCCTCGAAAACATGAACTCCACC
CTGAAAAGCCTCCACCAACAAAGCACCCAGCTCAGCAGCTCCCTGACCAGCGTGAAAACAAGCCTGAGAAGCAGCC
TGAACGACCCCCTGTGCCTCGTCCACCCAGCAGCGAGACCTGCAACAGCATCAGACTCAGCCTCAGCCAACCTCAA
CAGCAACCCCGAACTCAGACAACCTCCCCCGTGGACGCCGAACTGGACAACGTCAACAACGTGCTCAGAACAGAC
CTGGACGGCCTCGTGCAGCAGGGCTACCAAAGCCTCAACGACATCCCCGACAGAGTGACAGACAAACCACCACCG
TGGTGGCCGGAATTAAGAGAGTCTGAATAGCATCGGCAGCGACATTGACAACGTGACACAAAGACTCCCCATCCA
AGACATACTGAGCGCCTTCAGTGTGTACGTCAACAACACCGAGAGTTACATACACAGAAACCTGCCACCCTGGAG
GAGTACGACAGCTACTGGTGGCTGGGCGGACTCGTCATCTGCAGCCTCCTGACCCTGATCGTATTTTCTATTACC

TGGGCCTGCTCTGCGGCGTCTGCGGCTACGACCGACACGCCACACCCACCACAAGGGGCTGCGTGTCTAATACCGG
CGGCGTGTTCCTCATGGTGGGCGTTCGACTGTCCTTCTGTTCTGTTGGATCCTGATGATTATTGTTGTGCTGACC
TTCGTTTTTCGGCGCAACGTGGAGAAGCTGATCTGCGAGCCCTACACCTCCAAAGAGCTGTTTCAGAGTGCTGGACA
CCCCCTATCTGCTGAACGAAGACTGGGAGTATTACCTGAGCGGCAAGCTGTTTAATAAGAGTAAGATGAAACTGAC
CTTCGAGCAGGTGTATAGCGACTGCAAGAAAAACCGCGAACCTACGGCACCTGCACCTGCAGAACAGCTTCAAC
ATTTTCAGAGCACCTCAACATCAACGAGCACACCGGCTCCATCAGCAGCGAACTGGAGAGCCTGAAGGTGAACCTGA
ACATCTTCTGCTGGGCGCCGAGGCAGAAAAACCTGCAGGACTTCGCCCGCTGCGGCATCGACAGAATGAACTA
CGACAGCTACCTGGCCCAGACCGGCAAGAGCCCCGCGGCGTGAACCTGCTGAGCTTCGCCTACGACCTGGAGGCC
AAGGCCAACAGCCTGCCCCCGGCAACCTGAGAAACAGCCTGAAGAGAGACGCCAGACCATCAAGACCATCCACC
AGCAGAGAGTGCTGCCCATCGAGCAGAGCCTGAGCACCTGTACCAGAGCGTGAAGATCCTGCAGAGAACCGGCAA
CGGCCTGCTGGAGAGAGTGACCAGAATCCTGGCCAGCCTGGACTTCGCCCAGAACTTCATACCAACAACACCAGC
AGCGTGATCATCGAGGAGACCAAGAAGTACGGCAGAACCATCATCGGCTACTTCGAGCACTACCTGCAGTGGATCG
AGTTCAGCATCAGCGAGAAGGTGGCCAGCTGCAAGCCCGTGGCCACCGCCCTGGACACCGCCGTGGACGTGTTCT
GTGCAGCTACATCATCGACCCCCTGAACCTGTTCTGGTTTCGGCATCGGCAAGGCCACCGTGTTCCTGCTGCCCGCC
CTGATCTTCGCCGTGAAGCTGGCCAAGTACTACAGAAGAATGGACAGCGAGGACGTGTACGACACCCCAGCCAGC
ACTGA (SEQ ID NO.:1)

[0120] 腺相关病毒

[0121] 因腺相关病毒 (Adeno-associated virus, AAV) 较其他病毒载体小, 无致病性, 可转染正在分裂和未分裂的细胞等特性, 基于AAV载体的针对眼部特别是遗传性视网膜退行性病变的基因治疗方法受到了广泛的关注。

[0122] 腺相关病毒 (adeno-associated virus, AAV), 也称腺伴随病毒, 属于微小病毒科依赖病毒属, 是目前发现的一类结构最简单的单链DNA缺陷型病毒, 需要辅助病毒 (通常为腺病毒) 参与复制。它编码两个末端的反向重复序列 (ITR) 中的cap和rep基因。ITRs对于病毒的复制和包装具有决定性作用。cap基因编码病毒衣壳蛋白, rep基因参与病毒的复制和整合。AAV能感染多种细胞。

[0123] 重组腺相关病毒载体 (rAAV) 源于非致病的野生型腺相关病毒, 由于其安全性好、宿主细胞范围广 (分裂和非分裂细胞)、免疫源性低, 在体内表达外源基因时间长等特点, 被视为最有前途的基因转移载体之一, 在世界范围内的基因治疗和疫苗研究中得到广泛应用。经过10余年的研究, 重组腺相关病毒的生物学特性已被深入了解, 尤其是其在各种细胞、组织和体内实验中的应用效果方面已经积累了许多资料。在医学研究中, rAAV被用于多种疾病的基因治疗的研究 (包括体内、体外实验); 同时作为一种有特点的基因转移载体, 还广泛用于基因功能研究、构建疾病模型、制备基因敲除鼠等方面。

[0124] 在本发明一个优选的实施例中, 载体为重组AAV载体。AAV是相对较小的DNA病毒, 其可以稳定和位点特异性方式整合到它们所感染的细胞的基因组中。它们能够感染一大系列的细胞而不对细胞生长、形态或分化产生任何影响, 并且它们似乎并不涉及人体病理学。AAV基因组已被克隆、测序及表征。AAV在每个末端包含约145个碱基的反向末端重复序列 (ITR) 区域, 其作为病毒的复制起点。该基因组的其余被分成两个带有衣壳化功能的重要区域: 包含涉及病毒复制和病毒基因表达的rep基因的基因组左边部分; 以及包含编码病毒衣壳蛋白的cap基因的基因组右边部分。

[0125] AAV载体可采用本领域的标准方法制备。任何血清型的腺相关病毒均是合适的。用于纯化载体的方法可见于例如美国专利No. 6566118、6989264和6995006，它们的公开内容整体以引用方式并入本文。杂合载体的制备在例如PCT申请No. PCT/US2005/027091中有所描述，该申请的公开内容整体以引用方式并入本文。用于体外和体内转运基因的衍生自AAV的载体的使用已有描述（参见例如国际专利申请公布No. W091/18088和W093/09239；美国专利No. 4,797,368、6,596,535和5,139,941，以及欧洲专利No. 0488528，它们均整体以引用方式并入本文）。这些专利公布描述了其中rep和/或cap基因缺失并被所关注的基因替换的各种来源于AAV的构建体，以及这些构建体在体外（进入培养的细胞中）或体内（直接进入生物体）转运所关注的基因的用途。复制缺陷重组AAV可通过将以下质粒共转染进被人类辅助病毒（例如腺病毒）感染的细胞系而制备：所含的所关注核酸序列的侧翼为两个AAV反向末端重复序列（ITR）区域的质粒，和携带AAV衣壳化基因（rep和cap基因）的质粒。然后通过标准技术纯化所产生的AAV重组体。

[0126] 在一些实施方案中，重组载体被衣壳化到病毒粒子（例如包括但不限于AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAV13、AAV14、AAV15和AAV16的AAV病毒粒子）中。因此，本公开包括含有本文所述的任何载体的重组病毒粒子（因其包含重组多核苷酸而为重组的）。产生这样的粒子的方法是本领域已知的，并在美国专利No. 6,596,535中有所描述。

[0127] 表达载体和宿主细胞

[0128] 本发明还提供了一种用于PROM1蛋白的表达载体，它含有本发明的优化PROM1编码序列。

[0129] 通过提供的序列信息，熟练的技术人员可以使用可用的克隆技术以产生适于转导进入细胞的核酸序列或载体。

[0130] 优选地，编码PROM1蛋白的核酸序列作为载体，优选地表达载体被提供。优选地，其可作为优选地适用于在视网膜靶细胞中转导和表达的基因治疗载体被提供。载体可以是病毒的或非病毒的（例如质粒）。病毒载体包括源自以下的那些病毒载体：腺病毒、包括突变的形式腺相关病毒（AAV）、逆转录病毒、慢病毒、疱疹病毒、牛痘病毒、MMLV、GaLV、猿猴免疫缺陷病毒（SIV）、HIV、痘病毒和SV40。优选地，病毒载体是复制缺陷的（replication defective），尽管设想其可以是复制缺乏的（replication deficient）、能够复制或条件性复制的。病毒载体通常可以保持染色体外状态而不整合进入靶视网膜细胞的基因组。用于向视网膜靶细胞引入编码PROM1蛋白的核酸序列的优选的病毒载体是AAV载体，例如自身互补的腺相关病毒（scAAV）。使用特定的AAV血清型（AAV血清型2到AAV血清型12）或这些血清型中的任何一个的修饰的版本（包括AAV 4YF和AAV 7m8载体）可以实现选择性靶向。

[0131] 病毒载体可被修饰以缺失任何非必需的序列。例如，AAV中，病毒可被修饰以缺失全部或部分的IX基因、E1a和/或E1b基因。对于野生型AAV，没有辅助病毒诸如腺病毒的存在，复制是非常低效率的。对于重组的腺相关病毒，优选地，复制基因和衣壳基因以反式被提供（在pRep/Cap质粒中），并且仅AAV基因组的2ITR被保留并且包装进入病毒体，同时需要的腺病毒基因被腺病毒或另一个质粒提供。也可对慢病毒载体做出类似的修饰。

[0132] 病毒载体具有进入细胞的能力。然而，非病毒载体诸如质粒可与剂复合以有利于病毒载体被靶细胞的摄取。此类剂包括聚阳离子剂。可选地，递送系统诸如基于脂质体的递

送系统可被使用。用于在本发明中使用的载体优选地适于在体内或体外使用,并且优选地适于在人类中使用。

[0133] 载体将优选地包含一个或多个调节序列以指导核酸序列在视网膜靶细胞中的表达。调节序列可以包括与核酸序列可操作地连接的启动子、内含子、增强子、转录终止信号、多腺苷酸化序列、复制起点、核酸限制性位点、和同源重组位点。载体还可包括选择性标记,例如来确定载体在生长系统(例如细菌细胞)中或在视网膜靶细胞中的表达。

[0134] “可操作地连接”意指,核酸序列在功能上与其可操作地连接的序列相关,以使得它们以使得它们影响彼此的表达或功能的方式连接。例如,与启动子可操作地连接的核酸序列将具有被启动子影响的表达模式。

[0135] 启动子介导与其连接的核酸序列的表达。启动子可以是组成型的或可以是诱导型的。启动子可以指导在内视网膜细胞中遍在的表达,或神经元特异的表达。在后一种情况中,启动子可以指导细胞类型特异的表达,例如对给视神经节细胞。合适的启动子将是本领域技术人员已知的。例如,合适的启动子可以选自由以下组成的组:L7、thy-1、恢复蛋白、钙结合蛋白、人类CMV、GAD-67、鸡β肌动蛋白、hSyn、Grm6、Grm6增强子SV40融合蛋白。使用细胞特异的启动子可以实现靶向,例如Grm6-SV40用于选择性靶向给视神经细胞。Grm6启动子是Grm6基因的200碱基对增强子序列和SV40真核启动子的融合体,Grm6基因编码给视神经细胞特异的代谢型谷氨酸受体mGluR6。Grm6基因的优选的来源是小鼠和人类。使用泛-神经元的启动子可以实现遍在的表达,其实例在本领域是已知的并且可得的。一个此类实例是CAG。CAG启动子是CMV早期增强子和鸡β肌动蛋白启动子的融合体。

[0136] 合适的启动子的一个例子为即时早期巨细胞病毒(CMV)启动子序列。该启动子序列为能够驱动可操作地连接至其上的任何多核苷酸序列高水平表达的强组成型启动子序列。合适的启动子的另一个例子为延伸生长因子-1α(EF-1α)。然而,也可使用其他组成型启动子序列,包括但不限于类人猿病毒40(SV40)早期启动子、小鼠乳癌病毒(MMTV)、人免疫缺陷病毒(HIV)长末端重复(LTR)启动子、MoMuLV启动子、鸟类白血病毒启动子、艾伯斯坦-巴尔(Epstein-Barr)病毒即时早期启动子、鲁斯氏肉瘤病毒启动子、以及人基因启动子,诸如但不限于肌动蛋白启动子、肌球蛋白启动子、血红素启动子和肌酸激酶启动子。进一步地,本发明不应被限于组成型启动子的应用。诱导型启动子也被考虑为本发明的一部分。诱导型启动子的使用提供了分子开关,其能够当这样的表达是期望的时,打开可操作地连接诱导型启动子的多核苷酸序列的表达,或当表达是不期望的时关闭表达。诱导型启动子的例子包括但不限于金属硫蛋白启动子、糖皮质激素启动子、孕酮启动子和四环素启动子。

[0137] 许多表达载体可应用PROM1蛋白在哺乳动物细胞(较佳地为人,更佳地为人视神经细胞或感光细胞)表达。本发明优选腺相关病毒作为表达载体。

[0138] 本发明还提供一种PROM1优化编码序列的重组应用腺相关病毒载体的构建方法,该方法能快速、简便地构建携带PROM1优化编码序列的重组腺相关病毒载体,并包装获得复杂缺陷腺相关病毒载体。

[0139] 在另一优选例中,本发明所述的携带PROM1优化编码序列的腺相关病毒载体的序列如SEQ ID NO.:4所示。其中,第1-141位为左ITR序列;第153-532位为CMV增强子(下划线部分);第535-810位为鸡β肌动蛋白启动子(斜体部分);第813-1006位为SV40内含子(双下划线部分);第1019-3523位为优化的人PROM1编码序列(粗体部分);第3530-3757位为

SV40PolyA;第3758-3898位为右ITR序列。

[0140] CCTGCAGGCAGCTGCGGCTCGCTCGCTACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGGACCTTTGGTCGCCCCGGC
 CTCAGTGAGCGAGCGAGCGCAGAGAGGGAGTGCCAACTCCATCACTAGGGGTTCTGCGGCCGCGTCGACATTGATTATTGACTA
 GTTATTAATAGTAATCAATTACGGGGTCATTAGTTTACATAGCCCATATATGGAGTTCGCGGTTACATAACTACGGTAAATGGCCCCG
 TGGTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGT
 CAATGGGTGGACTATTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATG
 ACGGTAAATGGCCCCGCTGGCATTATGCCAGTACATGACCTTATGGGACTTTCCACTTGGCAGTACATCTACGTATTAGTCATCGC
 TATTACCATGGGTCGAGGTGAGCCCCACGTTCTGCTTCACTTCCCCATCTCCCCCCCCCTCCCCACCCCAATTTTGATTATTAT
 TTTTAAATATTTTGTGACGATGGGGGGGGGGGGGGGGGGGGCGCGCCAGGCGGGCGGGGGCGAGGGGGGGGGGGGGGGGG
 GAGGCGGAGAGGTGCGGGCGGAGCAATCAGAGCGGGCGCTCCGAAAGTTTCTTTTATGGCGAGGCGGGCGGGCGGGCGGGCTAT
 AAAAAGCGAAGCGCGGGGGGGGctagaggatccggtactagaggaactgaaaaccagaaaagtaactggttaagttagtctttt
 tgtcttttatttcaggtcccggatccgggtggtggtgcaaatcaagaactgctcctcagtggtggtgcttacttcttagcctgta
 cggaggtgtacttctgctctaaaagctgcggaattgtaccgcggaattccaccATGGCCCTGGTGGTGGGGAGCCTGCTGCTGC
 TGGGGCTGTGCGGAAACTCCTTCTCGGGGGCCAGCCCTCCAGCAGCCGACGCTCCTAAGCCCTGGAACACGAGCTGCCCGCCACCA
 ACTAGAAAACCAAGACTCCCAAAAGCCGGCCCATCGGCATCCTGTTGCAACTCGTGCAATTTTCTCTACGTGGTTCAACCCA
 GAGATTTTCCCGAGGACACCCTGAGAAAGTTCTGCAGAAGCCATGAGAGCAAGATTGACTACGACAAGCCCGAAACCGTGATCC
 TGGGCTGAAAGATCGTGTATTATGAGGCCGGAATTATCCTCTGTTGGCTGCTGGGCTGCTGTTTATCATCTGATGCCACTGGTGG
 GCTACTTCTTTTGCATGTGCAGATGTTGCAACAAGTGTGGCGCGAGATGCACCAGCGTCAGAAGGAGAAGCCGCTTCTCCGGA
 AATGCTTTGCCATCTCCCTGCTGGTATTGTATCATTATCAGCATCGGGATCTTCTACGGATTCTGGCTAACATCAGTCAAGT
 CCCCATCAAGCGCAGTAGAAAGCTGGCCGACTCCAACCTCAAGGACTGCGGACCTGCTGAACGAGACCCCGAGCAGATCAAGT
 ACATTCTGGCCCAATACAACACCACCAAGGACAAAGCCTTCAAGACTGAACTCCATCAACAGCGTCTCGCGGAGGCATACTGG
 ACCGGCTGAGACCAACATAATACCGTGTGGACGAAATCAAAGCATGGCCACCACCATAAAGGAGACCAAGAAGCCCTCGAAA
 ACATGAACTCCACCCTGAAAAGCCTCCACCAACAAAGCAGCTCAGCAGCTCCCTGACCAGCGTGAAAAACAAGCCTGAGAAGCA
 GCCTGAACGACCCCTGTGCTTCCACCCAGCAGCGAGACTGCAACAGCATCAGACTCAGCCTCAGCCAACCAACAGCAACC
 CCGAAGTCAAGCAACTCCCCCGTGGACGCCAAGTGGACAACGTCAACAACGTGCTCAGAACAGACTGGACGGCCTCGTGACG
 AGGGTACCAAAGCCTCAACGACATCCCGACAGAGTGCAGAGACAAACCACCACCGTGGTGGCCGAATTAAGAGAGTCTGAATA
 GCATCGGACGGACATTGACAACGTGACACAAAGACTCCCCATCAAGACATACTGAGCGCCTTCAGTGTGATCGTCAACAACACCG
 AGAGTTACATACAGAAAACCTGCCACCCTGGAGGAGTACGACAGTACTGGTGGTGGGCGGACTCGTCATCTGCAGCCTCCTGA
 CCCTGATCGTGATTTTCTATTACCTGGGCTGCTCTGCGGCTGCTGCGGCTACGACCGACACGCCACCCACCACAAGGGGCTGCG
 TGCTAATACCGGGCGGTGTTCTCATGGTGGGCGTGGACTGCTTCTGTTTCTGTTGGATCCTGATGATTATTGTTGTGCTGA
 CCTTCGTTTTCGGCGCAAAGTGGAGAAGCTGATCTGCGAGCCCTACACCTCAAAGAGCTGTTCAAGTGTGGACACCCCTATC
 TGCTGAACGAAGACTGGAGTATTACCTGAGCGGCAAGCTGTTAATAAGAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT
 ACTGCAAGAAAAACCGGAACCTACGGCACCTGCACCTGCAGAACAGCTTCAACATTTCAAGACCTCAACATCAACGAGCACA
 CCGGCTCCATCAGCAGCAACTGGAGGCTGAAGGTGAACCTGAACATCTTCTGCTGGGCGCGCAGGAGAAAAACCTGCAGG
 ACTTCGCGCCTGCGGCATCGACAGAATGAATACGACAGTACTGGCCAGACCGGAAGAGCCCGCGCGGTGAACCTGCTGA
 GCTTCGCTACGACTGGAGGCAAGGCCAACAGCCTGCCCCCGGCAACCTGAGAAACAGCTGAAGAGAGACGCCAGACCATCA
 AGACCATCCACCAGCAGAGAGTGTGCCATCGAGCAGAGCTGAGCACCCTGTACCAGAGCGTGAAGATCTGCAGAGAACCGCA
 ACGGCTGCTGGAGAGAGTACCCAGAACTCGGCCAGCTGGACTTCGCCCAGAACTTATCACCAACAACACCAGCAGCGTATCA
 TCGAGGAGACCAAGAAGTACGGCAGAACCATCATCGGCTACTTCGAGCACTACCTGCAGTGGATCGAGTTACAGATCAGCGAGAAG
 TGGCCAGCTGCAAGCCCGTGGCCACCGCCCTGGACACCGCGTGGACGTGTTCTGTCAGCTACATCATGACCCCTGAACCTGT
 TCTGGTTCGGCATCGGCAAGGCCACCGTGTCTGCTGCCGCCCTGATCTTCGCGTGAAGCTGGCCAAGTACTACAGAAGAATGG
 ACAGCGAGGACGTGACGACGACCCAGCCAGCACTGActcagagcgcgatccagacatgataagatacatgatgagtttgacaaa
 accacaactagaatgcagtgaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgaaccattataagctgcaata
 aacaagttacaacaacaattgcattcattttatgtttcaggttcaggggaggtgtgggaggttttttagtgcactggggagagatc
 tGCGGCCGAGGAAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGGCACCAGGTCGC
 CCGACGCCGGGCTTTGCCGGGGCGCCTCAGTGAGCGAGCGAGCGCAGCTGCCTGCAGGGCGCCTGATGCGGTAATTTCTCCTT
 ACGCATCTGTGCGGTATTTACACCCGATACGTCAAAGCAACCATAGTACGCGCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGTT
 GGTTACGCGCAGCGTACCGCTACACTTGCAGCGCCTAGCGCCGCTCCTTTTCGCTTCTTCCCTTCTTCTCGCCACGTTTCGCC
 GGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGATTTAGTGTCTTACGGCACCTCGACCCCAAAAAACTTGATT
 TGGGTGATGGTTACAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTAAATAGTGGACT
 CTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTGATTTATAAGGGATTTTGGCGATTTCGGCCTATTGGTTA
 AAAAATGAGCTGATTTAACAATAATTAACCGCAATTTAACAATAATTAACGTTTACAATTTTATGGTGCCTCTCAGTACAATCT
 GCTCTGATGCCCATAGTTAAGCCAGCCCCGACCCGCCAACCCCGTGCAGCGCCCTGACGGGCTGTCTGCTCCCGCATCCGC
 TTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGTTTTACCGTTCATCACCAGAACCGCGAGACGAAAGGGCCTC

[0141]

[0142] GTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA
 CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAA
 AAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATCCCTTTTTTTCGGCATTTCCTTCTGTTTTTGTCTACCCAGAA
 ACGTGGTAAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGTTACATCGAACTGGATCTAACAGCGGTAAGATCCTTG
 AGAGTTTTCCGCCCGAAGAACGTTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGG
 GCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGC
 ATGACAGTAAGAGAATTATGCAGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTACTTCTGACAACGATCGGAGGACCGA
 AGGAGCTAACCGCTTTTTTGCACAACATGGGGATCATGTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA
 CGACGAGGTGACACCAGATGCTGTAGCAATGGCAACAACGTTGCGCAAATAAAGTGGGCAACTACTTACTCTAGCTTCCCGG
 CAACAATTAATAGACTGGATGGAGCGGATAAAGTTGCAGGACCCTTCTGCGCTCGGCCCTCCGGCTGGCTGTTTTATTGCTGATA
 AATCTGGAGCCGTTGAGCGTGGGTCTCGCGGTATCATGTCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACAC
 GACGGGAGTCAAGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGTAAGTGTGACAGAC
 CAAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAGGATCTAGTGAAGATCCTTTTTGATAATCTCA
 TGACAAAAATCCCTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTT
 TCTGCGCGTAATCTGCTGCTTGCACAAAAAACCCCGTACCAGCGGTGGTTTTGTTGCGGATCAAGAGTACCAACTCTTTTT
 CCGAAGGTAACCTGCTCAGCAGAGCGAGATACCAATACTGTCTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAATCTG
 TAGCACCCTACATACCTCGCTCTGTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTGTGCTTACCAGGTTGGACTC
 AAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGCTGAACGGGGGTTCTGTCACACAGCCAGCTTGAGCGAACGACCTACACC
 GAACTGAGATACCTACAGCGTGAAGTATGAGAAAGCGCCACGCTTCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGACGG
 TCGAACAGGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGTCTGTCGGGTTTCGCCACCTCTGACTGA
 GCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCTTTTTACGGTTCTGGCCTTTTGC
 TGGCCTTTTGTCTACATGT (SEQ ID NO. : 4)

[0143] 本发明还提供了一种宿主细胞,用于表达PROM1蛋白。优选地,所述宿主细胞为哺乳动物细胞(较佳地为人,更佳地为人视神经细胞或感光细胞),提高PROM1蛋白的表达量。

[0144] 制剂和组合物

[0145] 本发明提供一种制剂或组合物,所述制剂或组合物含有(a)本发明第三方面所述的载体或本发明第四方面所述的腺相关病毒载体,以及(b)药学上可接受的载体或赋形剂。

[0146] 在另一优选例中,所述药物制剂用于治疗眼部疾病,所述眼部疾病是由PROM1基因突变导致的眼部疾病。

[0147] 在另一优选例中,所述药物制剂用于治疗视网膜色素变性(Retinitis Pigmentosa, RP),较佳地由PROM1基因突变导致的视网膜色素变性。

[0148] 本发明所述药物组合物中的“活性成分”是指本发明所述的载体(vector),例如病毒载体(包括腺相关病毒载体)。本发明所述的“活性成分”、制剂和/或组合物可用于治疗眼部疾病。“安全有效量”指的是:活性成分的量足以明显改善病情或症状,而不至于产生严重的副作用。“药学上可接受的载体或赋形剂(excipient)”指的是:一种或多种相容性固体或液体填料或凝胶物质,它们适合于人使用,而且必须有足够的纯度和足够低的毒性。“相容性”在此指的是组合物中各组份能和本发明的活性成分以及它们之间相互掺和,而不明显降低活性成分的药效。

[0149] 组合物可以是液体或固体,例如粉末、凝胶或糊剂。优选地,组合物是液体,优选地可注射液体。合适的赋形剂将是本领域技术人员已知的。

[0150] 在本发明中,所述载体可通过视网膜下或玻璃体内施用向眼睛施用。在任一种施用模式中,优选地,载体作为可注射液体被提供。优选地,可注射液体作为胶囊或注射器被提供。

[0151] 药学上可以接受的载体部分例子有纤维素及其衍生物(如羧甲基纤维素钠、乙基纤维素钠、纤维素乙酸酯等)、明胶、滑石、固体润滑剂(如硬脂酸、硬脂酸镁)、硫酸钙、植物油(如豆油、芝麻油、花生油、橄榄油等)、多元醇(如丙二醇、甘油、甘露醇、山梨醇等)、乳化

剂(如吐温®)、润湿剂(如十二烷基硫酸钠)、着色剂、调味剂、稳定剂、抗氧化剂、防腐剂、无热原水等。

[0152] 组合物可包含生理上可接受的无菌含水或无水溶液、分散液、悬浮液或乳液,和用于重新溶解成无菌的可注射溶液或分散液的无菌粉末。适宜的含水和非水载体、稀释剂、溶剂或赋形剂包括水、乙醇、多元醇及其适宜的混合物。

[0153] 本发明提供的编码PROM1的核酸或融合核酸,可以体外或体内生产PROM1蛋白,含所述PROM1优化编码序列的制剂可应用于制备治疗眼部疾病的药物。

[0154] 经优化的编码人PROM1蛋白的核酸表达量更高,从而翻译出更多的PROM1蛋白,比现有技术表达更多的PROM1蛋白,能较好地治疗视网膜色素变性。

[0155] 治疗方法

[0156] 本发明提供了向细胞提供感光细胞功能的方法,所述方法包括将包含编码PROM1的优化序列的载体引入到眼睛内。所述方法可包括向眼睛的内视网膜细胞视网膜下或玻璃体内施用核酸载体。

[0157] 本发明提供了用于通过向细胞提供感光细胞功能在治疗视网膜退化的方法中使用的核酸载体,所述核酸载体包含编码PROM1的优化序列。本发明组合物可以单独给药,或者与其他治疗药物联合给药(如配制在同一药物组合物中)。

[0158] 本发明还提供了扩大视网膜中的感光细胞功能的方法,特别是在视杆和/或视锥细胞退化之后扩大视网膜中的感光细胞功能的方法,所述方法包括将核酸载体引入到眼睛的玻璃体腔内,所述核酸载体包含编码PROM1的优化序列。所述方法可包括向眼睛的内视网膜细胞,视网膜下或玻璃体内施用核酸载体。本发明提供了用于通过扩大视网膜中的感光细胞功能在治疗视网膜退化中使用的核酸载体,所述核酸载体包含编码PROM1的优化序列。

[0159] 本发明还提供了对受试者恢复视力的方法,所述方法包括将包含编码PROM1的优化序列的载体引入到眼睛内。方法可包括视网膜下或玻璃体内施用核酸载体到眼睛的内视网膜细胞。本发明提供了用于在对受试者恢复视力中使用的核酸载体,所述核酸载体包含编码PROM1的优化序列。

[0160] 本发明还提供了治疗受试者中的视网膜疾病的方法,所述方法包括将包含编码PROM1的优化序列的载体引入到眼睛内。方法可包括视网膜下或玻璃体内施用核酸载体到眼睛的内视网膜细胞。疾病可以是视网膜营养不良,包括视杆营养不良、视杆视锥营养不良、视锥视杆营养不良、视锥营养不良和黄斑营养不良;其他形式的视网膜或黄斑退化、缺血性状况、视网膜色素变性、葡萄膜炎和由感光细胞能力的丧失导致的任何其他疾病。

[0161] 如本文所用,向细胞提供感光细胞功能意指,之前不具有感光细胞能力或其感光细胞能力已经完全地或部分地退化的细胞,在其中表达编码PROM1的外来核酸序列后,变成感光的。此类细胞在本文中可被称作转化的细胞,因为其在其中包含非天然的核酸。优选地,转化的视网膜细胞展现天然的感光细胞的一些或全部的感光细胞能力。优选地,转化的细胞展现天然的视网膜感光细胞的至少相同或大体上相同的感光能力。优选地,转化的细胞展现比患病的或正在退化的天然的视网膜感光细胞高的感光能力。因此,转化的细胞相比于来自同一来源、保持在同一条件下、未经处理的退化的或患病的细胞,将优选地具有增加的感光细胞。转化的细胞通过其中的外源核酸的存在可与天然的细胞区分。

[0162] 如本文所用,扩大感光细胞功能意指通过增加感光细胞诸如视杆或视锥细胞中的

功能和/或通过向细胞提供感光细胞功能,增加视网膜的感光细胞功能。因此,视网膜相比于未用如本文描述的方法处理的视网膜,将具有增加的接收光信号并且传送此类信号的能力,增加可以是任何量。

[0163] 如本文所用,恢复受试者中的视力意指,受试者相比于治疗之前,例如使用如本文描述的视力测试,显示改进的视力。恢复包括任何程度的改进,包括视力的完全恢复到完美的或接近完美的视力。

[0164] 如本文所用,治疗疾病意指施用如本文描述的核酸或载体以改善或减轻疾病的一种或多种症状,所述疾病选自由以下组成的组:视网膜营养不良,包括视杆营养不良、视杆视锥营养不良、视锥视杆营养不良、视锥营养不良和黄斑营养不良;另一种形式的视网膜或黄斑退化、视网膜色素变性、缺血性状况、葡萄膜炎和由感光细胞能力的丧失导致的任何其他疾病。改善或减轻可导致外周或中央视力、和/或白天或夜间视力的改善。

[0165] 本发明的方法包括将编码PROM1蛋白的核酸序列引入到眼睛的玻璃体腔内。优选地,方法包括使细胞与包含编码PROM1蛋白的核酸序列的载体(较佳地为病毒,更佳地为腺相关病毒)接触。优选地,细胞是视网膜细胞,优选地视锥细胞、视杆细胞、给光双极细胞、撤光双极细胞、水平细胞、神经节细胞和/或无长突细胞。

[0166] 当核酸序列和一种或多种酶以多个(两个或多个)剂量被提供时,这些剂量可分隔合适的时间间隔,例如30秒到若干小时或1天或多天。

[0167] 每个剂量可包含有效量的核酸序列或病毒载体。核酸序列或病毒载体的有效剂量可以在每治疗方案 1×10^9 - 1×10^{16} 病毒的范围。

[0168] 本发明是基于靶向编码PROM1的优化的核酸序列到视网膜细胞,以补偿视网膜中感光细胞的退化。核酸序列被靶向至的细胞是视网膜的细胞,其是活的并且能够表达外来核酸序列。视网膜细胞是视网膜的细胞,其是神经或神经元细胞并且能够变兴奋并传送电信号。优选地,靶视网膜细胞将能够产生电信号并且起始信号级联,导致信号向视神经的传送。优选地,靶视网膜细胞是内视网膜的细胞。靶细胞可以是视杆或视锥细胞,和/或可以是非感光细胞(即呈其天然形式的对光不响应的视网膜细胞)。靶视网膜细胞可以包括一种或多种细胞类型,所述细胞类型选自由以下组成的组:视杆细胞、视锥细胞、给光双极细胞、撤光双极细胞、水平细胞、神经节细胞、米勒细胞和/或无长突细胞。

[0169] 因此,当靶视网膜细胞是靶向视网膜的给光双极细胞、撤光双极细胞、水平细胞、神经节细胞和/或无长突细胞时,编码PROM1的核酸的表达可以被称作异位表达。因此,本发明在其范围内包括在非感光细胞中异位表达编码PROM1的核酸序列的方法。此类异位表达通过其中的异源PROM1蛋白的表达,具有向细胞提供感光细胞功能的作用。这用于增加观察到退化的视网膜的感光能力。

[0170] 水平细胞是内视网膜细胞,参与信号加工和反馈到感光细胞;双极细胞是内视网膜细胞并且在视杆/视锥细胞和无长突和/或神经节细胞间通信;无长突细胞发现于内视网膜并且允许在感光细胞途径和神经节细胞间的通信;神经节细胞是最内部的视网膜细胞,其将信号从感光细胞传递到视神经。

[0171] 本文对细胞的提及包括细胞的后代。优选地,根据本发明的对细胞的修饰还发生在转化的宿主细胞以后的代中。后代细胞可以不与最初的靶向细胞一致,但优选地将也展现非天然的PROM1的表达。

[0172] 与现有技术相比,本发明主要具有以下优点:

[0173] 1. 本发明对PROM1基因序列进行了针对性的特殊优化,基因序列不同于现有技术。与PROM1蛋白的未优化的DNA编码序列SEQ ID NO.2相比,优化后的序列(SEQ ID NO.:1)的PROM1蛋白表达量和含有该优化序列的细胞的感光能力显著提高,具有更强的视网膜电图a波峰值。

[0174] 2. 本发明优化后的PROM1蛋白编码序列非常适合在哺乳动物体内表达。在本发明中,表达的PROM1蛋白既保持了天然PROM1蛋白的活性,且PROM1蛋白表达量高,细胞感光能力强,能够非常有效地治疗由PROM1突变导致的眼部疾病,并且安全性好。

[0175] 下面结合具体实施,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件,例如Sambrook等人,分子克隆:实验室手册(New York:Cold Spring Harbor Laboratory Press,1989)中所述的条件,或按照制造厂商所建议的条件。除非另外说明,否则百分比和份数按重量计算。

[0176] 实施例1序列优化

[0177] 在本实施例中,基于PROM1蛋白的氨基酸序列(SEQ ID No.:3)和天然的编码序列(SEQ ID No.:2),本发明人对编码序列进行了优化。具体地,本发明人优化了影响基因表达的序列片段,这些序列片段包括但不限于:密码子使用偏好性,消除不利于表达的二级结构(如发夹结构),改变GC含量,CpG二核苷酸含量,mRNA的二级结构,隐蔽剪接位点,早期多聚腺苷化位点,内部核糖体进入位点和结合位点,负CpG岛,RNA不稳定区,重复序列(直接重复、反向重复等)和可能影响克隆的限制性位点。通过分析和试验筛选,最终得到如SEQ ID NO.:1所示的特别优化的DNA编码序列。

[0178] 如图1所示,构建携带有优化的PROM1编码序列的重组腺相关病毒载体rAAV/CBA.opti-hPROM1,序列如SEQ ID NO.:4所示。

[0179] 此外,还构建了携带有未优化的PROM1编码序列(SEQ ID NO.:2)的重组腺相关病毒载体rAAV/CBA.hPROM1,所述rAAV/CBA.hPROM1中除PROM1蛋白编码序列不同(即将SEQ ID NO.:1替换为SEQ ID NO.:2),其它与rAAV/CBA.opti-hPROM1相同。

[0180] 实施例2

[0181] 1. 实验分组:

[0182] 实验组1为出生2周的PROM1 KO小鼠接受rAAV/CBA.opti-hPROM1注射,实验组2为出生2周的PROM1 KO小鼠接受rAAV/CBA.hPROM1注射,对照组为同龄PROM1KO小鼠和正常小鼠。饲养环境为清洁级,予以国家标准饲料、过滤无菌水饲养,温度、湿度恒定,光照接近自然光,强度181lux,12h/12h昼/夜循环交替光照。

[0183] 2. 视网膜下腔注射

[0184] 小鼠经充分散瞳后全身麻醉,在眼科专用手术显微镜直视下,在角巩膜缘内侧瞳孔范围内用301/2gauge的一次性尖针头穿刺角膜,避免伤及虹膜和晶状体,然后用带33gauge平针头的微量进样器沿穿刺口进入,针头绕过晶体后到达玻璃体,然后逐渐进针至神经视网膜层和视网膜色素上皮(RPE)层之间的潜在视网膜腔隙并缓慢推注,注射量为1ul。注射载体悬浊液中添加0.1%荧光素钠染料(安全浓度),方便观察注射是否成功及网脱范围。注射过程中,眼表滴2.5%羟丙基甲基纤维素以便于随时观察眼底情况。手术显微

镜下如清楚看到眼底局部视网膜呈圆形隆起且网膜隆起下方可见绿色,证明注射成功。一定时间后小泡消失,视网膜局部隆起变平。术中如不能看到视网膜隆起及其下方的绿色或看到视网膜有大出血等并发症,则另选小鼠重新注射。术毕涂1%阿托品眼膏和四环素可的松眼膏,以后每隔1天重复一次,共三次,以减少炎症反应、防止感染。注射后18个月处死动物取眼球进行病理检测。每只小鼠一侧眼注射1 μ l滴度为 1×10^{13} 的rAAV/CBA.opti-hPROM1或rAAV/CBA.hPROM1(治疗眼),另一侧眼则不注射(未治疗眼或未注射眼)。

[0185] 3. Western Blot蛋白表达检测

[0186] 用rAAV/CBA.opti-hPROM1和rAAV/CBA.hPROM1分别感染HEK-293细胞,未转染的HEK-293细胞作为对照。感染2天后,分别提取蛋白,进行Western Blot检测蛋白表达,并检测PROM1蛋白的相对表达量。

[0187] 用标准Western Blot实验方法对结果进行检测,蛋白样品用蛋白提取试剂盒(天根)处理,人源PROM1一抗(1:1000稀释)。

[0188] 结果如图2所示,opti-hPROM1表达量明显高于未优化的hPROM1表达量,表达量约为未优化的hPROM1序列的3.2倍。

[0189] 4. 视觉电生理的检测

[0190] 采用Reti.port系统(德国Roland公司),刺激器为Ganzfeld Q450SC UV全视野球形刺激器。记录电极为金箔环状角膜电极,直径3rain;参考电极和接地电极均为不锈钢针状电极,各电极阻抗均小于5Q,干扰波幅小于20pV。双眼同时进行视网膜电图(Full.field electroretinograms, F-ERGs)记录,记录时间保持一致性,减小因昼夜节律造成的差异,均选在每天的14:00~17:00之间。F.ERGs记录前将小鼠暗适应过夜,实验时将10%水合氯醛作为麻醉注射入小鼠腹腔进行麻醉,同时用复方托吡卡胺滴眼液散瞳。待小鼠全身肌肉松弛,用沾有生理盐水的棉签轻擦眼周,使眼球突出后,置于37℃恒温水浴台上。暗红光下安插电极,记录电极分别置于左右眼角膜,局部滴一滴1%羧甲基纤维素钠滴眼液保持角膜滋润,增加导电性:参考电极刺入前额正中皮下,接地电极置于尾部。再次暗适应5min,按顺序记录暗适应ERG(Scotopic electroretinogram, Scot-ERG)和明适应ERG(Photopic electroretinogram, Phot.ERG)。记录Scot.ERG时,光强为 $.2.0 \log \text{ cd} \cdot \text{s}/\text{m}^2$;记录Phot.ERG时,首先明适应10min,刺激光强度为 $1.08 \log \text{ cd} \cdot \text{s}/\text{m}^2$ 。

[0191] PROM1 KO小鼠在生后14天接受视网膜下腔注射rAAV/CBA.opti-hPROM1载体、rAAV/CBA.hPROM1载体,对照组为未注射的PROM1 KO鼠和正常小鼠。注射2周后行ERG检查。

[0192] 结果如图3所示,正常小鼠、rAAV/CBA.opti-hPROM1注射的PROM1 KO鼠、rAAV/CBA.hPROM1注射的PROM1 KO鼠及未注射的PROM1 KO鼠暗适应下视网膜电图a波峰值对比,注射眼的a波峰值远高于未注射眼,且未注射眼基本无任何视网膜杆细胞功能。其中rAAV/CBA.opti-hPROM1与正常对照相比没有显著性差异,而rAAV/CBA.hPROM1的治疗效果不及rAAV/CBA.opti-hPROM1,并存在显著性差异($p < 0.01$)。

[0193] 5. 免疫组化检测:

[0194] (1)眼球冰冻切片制备:颈椎脱臼法处死小鼠,取出眼球前用大头针在小鼠眼角巩缘上方顶点12:00处留烙印作标记。用弯镊快速取出眼球,立即放入0.01mol/L PBS中;将眼球即刻浸泡在新鲜配制4%多聚甲醛溶液中,并用5ml一次性注射器上的16.19号粗针头在角膜缘穿刺,4℃过夜;分别使用含梯度为10%、20%、30%蔗糖的PBS溶液脱水;剪去角

膜,分离出晶状体,冷冻包埋剂(Optimal cutting temperature compound, OCT)进行组织包埋2h,用液氮速冻;冰冻切片机制备12 μ m的眼球冰冻切片置于粘附载玻片上,储存于-80 $^{\circ}$ C冰箱中。

[0195] (2) 免疫荧光染色:将冰冻切片从-80 $^{\circ}$ C冰箱取出,晾30min后放入0.01mol/L PBS中漂洗,快洗1次,5min \times 2次,10min \times 2次;0.3% Triton X.100中室温浸泡30min;使用5%牛血清蛋白封闭,室温下浸泡2h;吸去血清,暗室中用移液枪加兔抗人PROM1一抗(1:500),每个标本30 μ l,4 $^{\circ}$ C孵育过夜;0.01mol/L PBS中快洗1次,5min \times 2次,10min \times 2次;暗室中滴加山羊抗兔IgG.Cy3二抗稀释液(1:2000)、FITC(1:100),每个标本30 μ l,湿盒内孵育1h;0.01mol/L PBS中快速漂洗5次;最后滴加DAPI稀释液(1:500),每个标本30 μ l,孵育2min,0.01mol/L PBS中快速漂洗5次,用移液枪加少许抗荧光淬灭液,盖玻片封片,盖玻片四角滴加指甲油固定。

[0196] 结果:(1) PROM1表达检测:PROM1 KO小鼠在生后14天一只眼接受视网膜下腔注射rAAV/CBA.opti-hPROM1载体,另一只眼不注射为对照,3周时进行视网膜免疫组化检测。注射眼外核层皆可见到PROM1蛋白表达,未注射眼无表达迹象(图4)。(2) DAPI染色检测:PROM1 KO小鼠在生后14天一只眼接受视网膜下腔注射rAAV/CBA.opti-hPROM1载体,未注射PROM1 KO小鼠作为对照,18个月时进行视网膜免疫组化检测。注射rAAV/CBA.opti-hPROM1载体小鼠视网膜厚度显著大于未注射小鼠,与正常对照相比没有显著性差异(图5)。

[0197] 在本发明提及的所有文献都在本申请中引用作为参考,就如同每一篇文献被单独引用作为参考那样。此外应理解,在阅读了本发明的上述讲授内容之后,本领域技术人员可以对本发明作各种改动或修改,这些等价形式同样落于本申请所附权利要求书所限定的范围。

[0001]	序列表	
[0002]	<110> 上海市第一人民医院	
[0003]	<120> 治疗视网膜色素变性疾病的基因治疗载体	
[0004]	<130> P2018-1831	
[0005]	<160> 4	
[0006]	<170> PatentIn version 3.5	
[0007]	<210> 1	
[0008]	<211> 2505	
[0009]	<212> DNA	
[0010]	<213> 人工序列(artificial sequence)	
[0011]	<400> 1	
[0012]	atggccctgg tgctggggag cctgctgctg ctggggctgt gcggaactc cttctccggg	60
[0013]	ggccagccct ccagaccga cgctcctaag gcctggaact acgagctgcc cgccaccaac	120
[0014]	tacgaaacc aagactcca caaagccgc cccatcggca tcctgttcga actcgtgcat	180
[0015]	atcttctct acgtggttca acccagagat tttcccgagg acaccctgag aaagttcctg	240
[0016]	cagaaggcct atgagagca gattgactac gacaagccc aaaccgtgat cctgggcctg	300
[0017]	aagatcgtgt attatgagc cggaattatc ctctgttgcg tgctgggcct gctgtttatc	360
[0018]	atctgatgc cactggtgg ctacttcttt tgcatgtgca gatgttgca caagtgtggc	420
[0019]	ggcgagatgc accagcgtca gaaggagaac gggcctttcc tccgaaatg ctttgccatc	480
[0020]	tcctgctgg tgattgtat cattatcagc atcgggatct tctacggatt cgtggctaac	540
[0021]	catcaggtca gaaccgcat caagcgcagt agaaagctgg ccgactcca cttcaaggac	600
[0022]	ctgcggacc tgctgaacga gaccccgag cagatcaagt acattctggc ccaatacaac	660
[0023]	accaccaagg acaagcctt cacagacctg aactccatca acagcgtgct cggcggaggc	720
[0024]	atactggacc ggctgagacc caacataata cccgtgctgg acgaaatcaa aagcatggcc	780
[0025]	accgccataa aggagacca agaagccctc gaaaacatga actccacct gaaaagcctc	840
[0026]	caccaacaaa gcaccagct cagcagctcc ctgaccagcg tgaaaacaag cctgagaagc	900
[0027]	agcctgaacg acccctgtg cctcgtccac cccagcagcg agacctgcaa cagcatcaga	960
[0028]	ctcagcctca gccaaactca cagcaacccc gaactcagac aactcccc cgtggacgcc	1020
[0029]	gaactggaca acgtcaacaa cgtgctcaga acagacctgg acggcctcgt gcagcagggc	1080
[0030]	taccaaaagg tcaacgacat ccccgacaga gtgcagagac aaaccaccac cgtggtggcc	1140
[0031]	ggaattaaga gactcctgaa tagcatcggc agcgacattg acaacgtgac acaagactc	1200
[0032]	cccatccaag acatactgag cgccttcagt gtgtacgtca acaacaccga gagttacata	1260
[0033]	cacagaaacc tgcccacct ggaggagtac gacagctact ggtggctggg cggactcgtc	1320
[0034]	atctgcagcc tctgacct gatcgtgatt ttctattacc tgggctgct ctgcggcgte	1380
[0035]	tgcggtacg accgacacgc cacaccacc acaagggct gcgtgtctaa taccggcggc	1440
[0036]	gtgttctca tgggtggcgt cggactgtcc ttctgttct gttggatcct gatgattatt	1500
[0037]	gttgtgctga cttcgtttt cggcgccaac gtggagaagc tgatctgca gccctacacc	1560
[0038]	tccaaagagc gtctcagagt gctggacacc ccctatctgc tgaacgaaga ctgggagtat	1620
[0039]	tactgagcg gcaagctgtt taataagagt aagatgaaac tgaccttca gcaggtgtat	1680
[0040]	agcgactgca agaaaaacc cggaacctac ggcacctgc acctgcagaa cagcttcaac	1740
[0041]	atctcagagc acctcaacat caacgagcac accggtcca tcagcagcga actggagagc	1800

[0042] ctgaaggatga acctgaacat cttcctgctg ggcgccgag gcagaaaaa cctgcaggac 1860
 [0043] ttgccgcct gggcatcga cagaatgaac tacgacagct acctggccca gaccggcaag 1920
 [0044] agccccgcg gcgtgaacct gctgagcttc gcctacgacc tggaggccaa ggccaacagc 1980
 [0045] ctgcccccg gcaacctgag aaacagcctg aagagagacg cccagacat caagaccatc 2040
 [0046] caccagcaga gagtgtgcc catcgagcag agcctgagca ccctgtacca gagcgtgaag 2100
 [0047] atcctgcaga gaaccggcaa cggcctgctg gagagagtga ccagaatcct ggcagcctg 2160
 [0048] gacttcgcc agaactcat caccaaac accagcagcg tgatcatcga ggagaccaag 2220
 [0049] aagtacggca gaaccatcat cggctacttc gagcactacc tgcagtggat cgagttcagc 2280
 [0050] atcagcgaga aggtggccag ctgcaagccc gtggccaccg ccctggacac cgccgtggac 2340
 [0051] gtgttcctgt gcagctacat catcgacccc ctgaacctgt tctggttcgg catcggaag 2400
 [0052] gccaccgtgt tctgtgcc cgccctgatc ttgccgtga agctggccaa gtactacaga 2460
 [0053] agaattggaca gcgaggacgt gtacgacgac cccagccagc actga 2505
 [0054] <210> 2
 [0055] <211> 2505
 [0056] <212> DNA
 [0057] <213> 智人 (Homo sapiens)
 [0058] <400> 2
 [0059] atggcctcg tactcgctc cctgttctg ctggggctgt gcgggaactc cttttcagga 60
 [0060] gggcagcctt catccacaga tgctcctaag gcttgaatt atgaattgc tgcaacaaat 120
 [0061] tatgagacc aagactcca taaagctgga ccattggca ttctctttga actagtgcatt 180
 [0062] atctttctct atgtgttaca gccgctgat ttcccagaag atactttgag aaaattctta 240
 [0063] cagaaggcat atgaatcaa aattgattat gacaagccag aaactgtaat cttaggtcta 300
 [0064] aagattgtct actatgaagc agggattatt ctatgctgtg tcttggggct gctgtttatt 360
 [0065] attctgatgc ctctgtggg gtatttcttt tgtatgtgc gttgctgtaa caaatgtggt 420
 [0066] ggagaaatgc accagcaga gaagaaaat gggccctcc tgaggaaatg ctttgcaatc 480
 [0067] tccctgttgg tgatttgtat aataataagc attggcatct tctatggtt tgtggcaaat 540
 [0068] caccaggtaa gaaccggat caaaaggagt cggaaactgg cagatagcaa tttcaaggac 600
 [0069] ttgcgaactc tcttgaatga aactccagag caaatcaaat atatattggc ccagtacaac 660
 [0070] actaccaagg acaaggcgtt cacagatctg aacagtatca attcagtgt aggaggcgga 720
 [0071] attcttgacc gactgagacc caacatcatc cctgttcttg atgagattaa gtccatggca 780
 [0072] acagcgatca aggagaccaa agaggcgtg gagaacatga acagcacct gaagagcttg 840
 [0073] caccaacaaa gtacacagct tagcagcagt ctgaccagcg tgaaaactag cctgcggtca 900
 [0074] tctctcaatg accctctgtg cttggtgcat ccatcaatg aaacctgcaa cagcatcaga 960
 [0075] ttgtctctaa gccagctgaa tagcaaccct gaactgagc agcttccacc cgtggatgca 1020
 [0076] gaactgaca acgttaataa cgttcttagg acagatttg atggcctggt ccaacagggc 1080
 [0077] tatcaatccc ttaatgatat acctgacaga gtacaacgc aaaccagcagc tgctcgtagca 1140
 [0078] ggtatcaaaa ggtcttgaa ttccattggt tcagatatcg acaatgtaac tcagcgtctt 1200
 [0079] cctattcagg atatactctc agcattctct gtttatgta ataactga aagttacatc 1260
 [0080] cacagaaatt tacctacatt ggaagagtat gattcact ggtggctggg tggcctggtc 1320
 [0081] atctgctctc tgctgacct catcgtgatt ttttactacc tgggcttact gtgtggcgtg 1380
 [0082] tgcgctatg acaggcatgc caccgagc acccgagct gtgtctcaa caccggaggc 1440
 [0083] gtcttctca tggttggagt tggattaagt ttctctttt gctggatatt gatgatcatt 1500

[0126]	Ser Leu Leu Val Ile Cys Ile Ile Ile Ser Ile Gly Ile Phe Tyr Gly
[0127]	165 170 175
[0128]	Phe Val Ala Asn His Gln Val Arg Thr Arg Ile Lys Arg Ser Arg Lys
[0129]	180 185 190
[0130]	Leu Ala Asp Ser Asn Phe Lys Asp Leu Arg Thr Leu Leu Asn Glu Thr
[0131]	195 200 205
[0132]	Pro Glu Gln Ile Lys Tyr Ile Leu Ala Gln Tyr Asn Thr Thr Lys Asp
[0133]	210 215 220
[0134]	Lys Ala Phe Thr Asp Leu Asn Ser Ile Asn Ser Val Leu Gly Gly Gly
[0135]	225 230 235 240
[0136]	Ile Leu Asp Arg Leu Arg Pro Asn Ile Ile Pro Val Leu Asp Glu Ile
[0137]	245 250 255
[0138]	Lys Ser Met Ala Thr Ala Ile Lys Glu Thr Lys Glu Ala Leu Glu Asn
[0139]	260 265 270
[0140]	Met Asn Ser Thr Leu Lys Ser Leu His Gln Gln Ser Thr Gln Leu Ser
[0141]	275 280 285
[0142]	Ser Ser Leu Thr Ser Val Lys Thr Ser Leu Arg Ser Ser Leu Asn Asp
[0143]	290 295 300
[0144]	Pro Leu Cys Leu Val His Pro Ser Ser Glu Thr Cys Asn Ser Ile Arg
[0145]	305 310 315 320
[0146]	Leu Ser Leu Ser Gln Leu Asn Ser Asn Pro Glu Leu Arg Gln Leu Pro
[0147]	325 330 335
[0148]	Pro Val Asp Ala Glu Leu Asp Asn Val Asn Asn Val Leu Arg Thr Asp
[0149]	340 345 350
[0150]	Leu Asp Gly Leu Val Gln Gln Gly Tyr Gln Ser Leu Asn Asp Ile Pro
[0151]	355 360 365
[0152]	Asp Arg Val Gln Arg Gln Thr Thr Thr Val Val Ala Gly Ile Lys Arg
[0153]	370 375 380
[0154]	Val Leu Asn Ser Ile Gly Ser Asp Ile Asp Asn Val Thr Gln Arg Leu
[0155]	385 390 395 400
[0156]	Pro Ile Gln Asp Ile Leu Ser Ala Phe Ser Val Tyr Val Asn Asn Thr
[0157]	405 410 415
[0158]	Glu Ser Tyr Ile His Arg Asn Leu Pro Thr Leu Glu Glu Tyr Asp Ser
[0159]	420 425 430
[0160]	Tyr Trp Trp Leu Gly Gly Leu Val Ile Cys Ser Leu Leu Thr Leu Ile
[0161]	435 440 445
[0162]	Val Ile Phe Tyr Tyr Leu Gly Leu Leu Cys Gly Val Cys Gly Tyr Asp
[0163]	450 455 460
[0164]	Arg His Ala Thr Pro Thr Thr Arg Gly Cys Val Ser Asn Thr Gly Gly
[0165]	465 470 475 480
[0166]	Val Phe Leu Met Val Gly Val Gly Leu Ser Phe Leu Phe Cys Trp Ile
[0167]	485 490 495

[0168]	Leu Met Ile Ile Val Val Leu Thr Phe Val Phe Gly Ala Asn Val Glu
[0169]	500 505 510
[0170]	Lys Leu Ile Cys Glu Pro Tyr Thr Ser Lys Glu Leu Phe Arg Val Leu
[0171]	515 520 525
[0172]	Asp Thr Pro Tyr Leu Leu Asn Glu Asp Trp Glu Tyr Tyr Leu Ser Gly
[0173]	530 535 540
[0174]	Lys Leu Phe Asn Lys Ser Lys Met Lys Leu Thr Phe Glu Gln Val Tyr
[0175]	545 550 555 560
[0176]	Ser Asp Cys Lys Lys Asn Arg Gly Thr Tyr Gly Thr Leu His Leu Gln
[0177]	565 570 575
[0178]	Asn Ser Phe Asn Ile Ser Glu His Leu Asn Ile Asn Glu His Thr Gly
[0179]	580 585 590
[0180]	Ser Ile Ser Ser Glu Leu Glu Ser Leu Lys Val Asn Leu Asn Ile Phe
[0181]	595 600 605
[0182]	Leu Leu Gly Ala Ala Gly Arg Lys Asn Leu Gln Asp Phe Ala Ala Cys
[0183]	610 615 620
[0184]	Gly Ile Asp Arg Met Asn Tyr Asp Ser Tyr Leu Ala Gln Thr Gly Lys
[0185]	625 630 635 640
[0186]	Ser Pro Ala Gly Val Asn Leu Leu Ser Phe Ala Tyr Asp Leu Glu Ala
[0187]	645 650 655
[0188]	Lys Ala Asn Ser Leu Pro Pro Gly Asn Leu Arg Asn Ser Leu Lys Arg
[0189]	660 665 670
[0190]	Asp Ala Gln Thr Ile Lys Thr Ile His Gln Gln Arg Val Leu Pro Ile
[0191]	675 680 685
[0192]	Glu Gln Ser Leu Ser Thr Leu Tyr Gln Ser Val Lys Ile Leu Gln Arg
[0193]	690 695 700
[0194]	Thr Gly Asn Gly Leu Leu Glu Arg Val Thr Arg Ile Leu Ala Ser Leu
[0195]	705 710 715 720
[0196]	Asp Phe Ala Gln Asn Phe Ile Thr Asn Asn Thr Ser Ser Val Ile Ile
[0197]	725 730 735
[0198]	Glu Glu Thr Lys Lys Tyr Gly Arg Thr Ile Ile Gly Tyr Phe Glu His
[0199]	740 745 750
[0200]	Tyr Leu Gln Trp Ile Glu Phe Ser Ile Ser Glu Lys Val Ala Ser Cys
[0201]	755 760 765
[0202]	Lys Pro Val Ala Thr Ala Leu Asp Thr Ala Val Asp Val Phe Leu Cys
[0203]	770 775 780
[0204]	Ser Tyr Ile Ile Asp Pro Leu Asn Leu Phe Trp Phe Gly Ile Gly Lys
[0205]	785 790 795 800
[0206]	Ala Thr Val Phe Leu Leu Pro Ala Leu Ile Phe Ala Val Lys Leu Ala
[0207]	805 810 815
[0208]	Lys Tyr Tyr Arg Arg Met Asp Ser Glu Asp Val Tyr Asp Asp Pro Ser
[0209]	820 825 830

[0210]	Gln His	
[0211]	<210>	4
[0212]	<211>	6495
[0213]	<212>	DNA
[0214]	<213>	人工序列(artificial sequence)
[0215]	<400>	4
[0216]	cctgcaggca gctgcgcgct cgctcgetca ctgaggccgc ccgggcaaag cccgggcgtc	60
[0217]	gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca	120
[0218]	actccatcac taggggttcc tgcggcccgc tcgacattga ttattgacta gttattaata	180
[0219]	gtaatcaatt acggggtcatt tagttcatag cccatataat gagttccgcg ttacataact	240
[0220]	tacggtaaat ggcccgcctg gctgaccgcc caacgacccc cgcccattga cgtcaataat	300
[0221]	gacgtatggt cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggacta	360
[0222]	tttacggtaa actgcccact tggcagtaca tcaagtgtat catatgcca gtagcggccc	420
[0223]	tattgacgtc aatgacggta aatggcccgc ctggcattat gccagtagca tgaccttatg	480
[0224]	ggactttcct acttggcagt acatctacgt attagtcatc gctattacca tgggtcgagg	540
[0225]	tgagccccac gttctgcttc actctcccca tctccccccc ctccccacc ccaattttgt	600
[0226]	atattttat tttttaatta ttttgtcag cgatgggggc gggggggggg ggggcgcgcg	660
[0227]	ccaggcgggg cggggcgggg cgagggcgcg ggcggggcga ggcggagagg tgcggcgcca	720
[0228]	gccaatcaga gcggcgcgct ccgaaagttt cttttatgg cgaggcggcg gcggcggcg	780
[0229]	ccctataaaa agcgaagcgc gcggcggcg tctagaggat ccggtactag aggaactgaa	840
[0230]	aaaccagaaa gttaactggt aagtttagtc tttttgtctt ttatttcagg tcccggatcc	900
[0231]	ggtggtggtg caaatcaaag aactgctcct cagtggatgt tgcctttact tctaggcctg	960
[0232]	tacggaagtg ttacttctgc tctaaaagct gcggaattgt acccgcgga attccacat	1020
[0233]	ggcctggtg ctggggagcc tgctgtctct ggggctgtgc ggaaactcct tctccgggg	1080
[0234]	ccagccctcc agcaccgacg ctctaaagc ctggaactac gagctgccc ccaccaacta	1140
[0235]	cgaaaccaa gactcccaca aagccggccc catcgcatc ctgttcgaac tcgtgcatat	1200
[0236]	tttctctac gtggttcaac ccagagattt tcccaggac accctgagaa agttcctgca	1260
[0237]	gaaggcctat gagagcaaga ttgactacga caagcccga accgtgatcc tgggcctgaa	1320
[0238]	gatcgtgat tatgagccg gaattatcct ctgttgctg ctgggectgc tgtttatcat	1380
[0239]	cctgatgcca ctggtgggct acttctttg catgtgcaga tgttgcaaca agtgtggcgg	1440
[0240]	cgagatgcac cagcgtcaga aggagaacgg gcctttcctc cggaaatgct ttgccatctc	1500
[0241]	cctgctggtg atttgtatca ttatcagcat cgggatcttc tacggattcg tggctaacca	1560
[0242]	tcaggtcaga accgcatca agcgcagtag aaagctggcc gactccaact tcaaggacct	1620
[0243]	gcggacctg ctgaacgaga cccccagca gatcaagtac attctggccc aatacaaac	1680
[0244]	caccaaggac aaagccttca cagacctgaa ctccatcaac agcgtgctcg gcggaggcat	1740
[0245]	actggaccgg ctgagacca acataatacc cgtgctggac gaaatcaaaa gcatggccac	1800
[0246]	cgccataaag gagaccaaag aagcctcga aaacatgaac tccaccctga aaagcctcca	1860
[0247]	ccaacaaagc acccagctca gcagctcct gaccagcgtg aaaacaagcc tgagaagcag	1920
[0248]	cctgaacgac cccctgtgcc tcgtccacc cagcagcgag acctgcaaca gcatcagact	1980
[0249]	cagcctcagc caactcaaca gcaaccccga actcagacaa ctccccccg tggacgccga	2040
[0250]	actggacaac gtcaacaacg tgctcagaac agacctggac ggcctcgtgc agcagggcta	2100
[0251]	ccaaagcctc aacgacatcc ccgacagagt gcagagacaa accaccaccg tgggtggccg	2160

[0252]	aattaagaga gtcctgaata gcatcggcag cgacattgac aacgtgacac aaagactccc	2220
[0253]	catccaagac atactgagcg ctttcagtgt gtacgtcaac aacaccgaga gttacataca	2280
[0254]	cagaaacctg cccaccctgg aggagtacga cagctactgg tggctgggcg gactcgtcat	2340
[0255]	ctgcagcctc ctgaccctga tcgtgatttt ctattacctg ggcttgcctc gcggcgtctg	2400
[0256]	cggctacgac cgacacgcca caccaccac aaggggctgc gtgtctaata ccggcggcgt	2460
[0257]	gttctcatg gtggcgctcg gactgtcctt cctgttctgt tggatcctga tgattattgt	2520
[0258]	tgtgetgacc ttctgttctg gcgccaactg ggagaactg atctgcgagc cctacacctc	2580
[0259]	caaagactg ttctagactg tggacacccc ctatctgctg aacgaagact gggagtatta	2640
[0260]	cctgagcggc aagctgttta ataagagtaa gatgaaactg accttcgagc aggtgtatag	2700
[0261]	cgactgcaag aaaaaccgcg gaacctacgg caccctgcac ctgcagaaca gcttcaacat	2760
[0262]	ttcagagcac ctcaacatca acgagcacac cggctccatc agcagcgaac tggagagcct	2820
[0263]	gaagtgaaac ctgaacatct tcttgcctgg cgccgcagc agaaaaaac tgcaggactt	2880
[0264]	cgccgctgc ggcctgcaca gaatgaacta cgacagctac ctggcccaga ccggcaagag	2940
[0265]	ccccgccggc gtgaacctgc tgagcttgc ctacgacctg gaggccaagg ccaacagcct	3000
[0266]	gcccccggc aacctgagaa acagcctgaa gagagacgcc cagaccatca agaccatcca	3060
[0267]	ccagcagaga gtgctgcca tcgagcagag cctgagcacc ctgtaccaga gcgtgaagat	3120
[0268]	cctgcagaga accggcaacg gcctgctgga gagagtgacc agaactcctg ccagcctgga	3180
[0269]	cttcgcccag aacttcatca ccaacaacac cagcagcgtg atcatcgagg agaccaagaa	3240
[0270]	gtacggcaga accatcatcg gctacttcca gactacctg cagtggatcg agttcagcat	3300
[0271]	cagcagaaag gtggccagct gcaagcccgt ggccaccgcc ctggacaccg ccgtggacgt	3360
[0272]	gttctgtgc agctacatca tcgacccct gaacctgttc tggttcggca tcggcaaggc	3420
[0273]	caccgtgttc ctgctgccc cctgatctt cgccgtgaag ctggccaagt actacagaag	3480
[0274]	aatggacagc gaggacgtgt acgacaccc cagccagcac tgactcgagc gcgatccag	3540
[0275]	acatgataag atacattgat gagtttgac aaaccacaac tagaatgcag tgaaaaaaat	3600
[0276]	gctttatttg tgaaattgt gatgctattg ctttattgt aaccattata agctgcaata	3660
[0277]	aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg gaggtgtggg	3720
[0278]	aggtttttta gtcgactggg gagagatctg cggccgcagg aaccctagt gatggagtg	3780
[0279]	gccactcct ctctgcgcgc tcgctcctc actgaggccg ggcgacaaa ggtcgcccga	3840
[0280]	cgccgggct ttgcccggc ggccctcagtg agcagcagc cgcgcagct cctgcagggg	3900
[0281]	cgctgatgc ggtattttct ccttacgat ctgtgcgta tttcacacc catacgtcaa	3960
[0282]	agcaaccata gtacgcgcc tgtagcggc cattaagcgc ggcggtgtg gtggttacgc	4020
[0283]	gcagcgtgac cgctacactt gccagcggc tagcggcgc tctttcgt tcttccct	4080
[0284]	cctttctgc cagttcgcc ggctttccc gtcaagctt aaatcggggg ctccctttag	4140
[0285]	ggttccgatt tagtcttta cggcacctg acccaaaaa acttgattg ggtgatggtt	4200
[0286]	cacgtagtgg gccatgcgcc tgatagcgg tttttcgcc tttgacgtg gactccacgt	4260
[0287]	tctttaatag tggactctg ttccaaactg gaacaacact caaccctatc tcgggctatt	4320
[0288]	ctttgattt ataaggatt ttgcccatt cgccctatt gttaaaaaat gagctgattt	4380
[0289]	aacaaaaatt taacgcgaat tttacaaaa tattaactt tacaatttta tgggtgactc	4440
[0290]	tcagtacaat ctgctctgat gccgatagt taagccagc ccgacaccg ccaacaccg	4500
[0291]	ctgacgcgcc ctgacgggct tgtctgtcc cggcatccg ttacagaaa gctgtgaccg	4560
[0292]	tctccgggag ctgcatgtgt cagagtttt caccgtcatc accgaaacgc gcgagacgaa	4620
[0293]	aggcctcgt gatacgccct tttttatagg ttaatgtcat gataataatg gtttcttaga	4680

[0294] cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 4740
[0295] tacattcaaa tatgtatccg ctcatgagac aataaccctg ataatgctt caataatatt 4800
[0296] gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc cttattccc ttttttgceg 4860
[0297] cattttgcct tcctgtttt gctcaccag aaacgctggt gaaagtaaaa gatgctgaag 4920
[0298] atcagttggg tgcacgagt ggttacatcg aactggatct caacagcggg aagatccttg 4980
[0299] agagtttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 5040
[0300] gcgcggtatt atcccgtatt gacgccggg aagagcaact cggtcgccg atacactatt 5100
[0301] ctcagaatga cttggttgag tactcaccag tcacagaaaa gcattcttac gatggcatga 5160
[0302] cagtaagaga attatgcagt gctgccataa ccatgagtga taactctgc gccaaacttac 5220
[0303] ttctgacaac gatcggagga ccgaaggagc taaccgctt tttgcacaac atgggggagc 5280
[0304] atgtaactcg cttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 5340
[0305] gtgacaccac gatgcctgta gcaatggcaa caacgttgc caaactatta actggcgaac 5400
[0306] tacttactct agcttcccg caacaattaa tagactggat ggaggcggat aaagttgcag 5460
[0307] gaccacttct gcgctcggc cttccggctg gctggtttat tgctgataaa tctggagccg 5520
[0308] gtgagcgtgg gtctcgcgg atcattgcag cactggggcc agatggtaag ccctcccgtg 5580
[0309] tcgtagttat ctacacgagc gggagtcagg caactatgga tgaacgaaat agacagatcg 5640
[0310] ctgagatagg tcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 5700
[0311] tactttagat tgatttaaaa cttcattttt aatttaaaag gatctagtg aagatccttt 5760
[0312] ttgataatct catgacccaa atccctaac gtgagtttc gttccactga gcgtcagacc 5820
[0313] ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcta atctgctgct 5880
[0314] tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgtt gccgatcaa gagctaccaa 5940
[0315] ctcttttcc gaaggtact ggcttcagca gagcgcagat accaaatact gtccttctag 6000
[0316] tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgtc 6060
[0317] tgctaatcct gttaccagt gctgctgcca gtggcgataa gtcgtgtctt accgggttg 6120
[0318] actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 6180
[0319] cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat 6240
[0320] gagaaagcgc cacgctccc gaaggagaa aggcggacag gtatccgta agcggcaggg 6300
[0321] tcggaacagg agagcgcac agggagctc cagggggaaa cgcctggtat ctttatagtc 6360
[0322] ctgtcgggtt tcgccacctc tgacttgagc gtcgatttt gtgatgctc tcaggggggc 6420
[0323] ggagcctatg gaaaaacgc agcaacgcgg cttttttac gttcctggc ttttgctggc 6480
[0324] cttttgctca catgt 6495

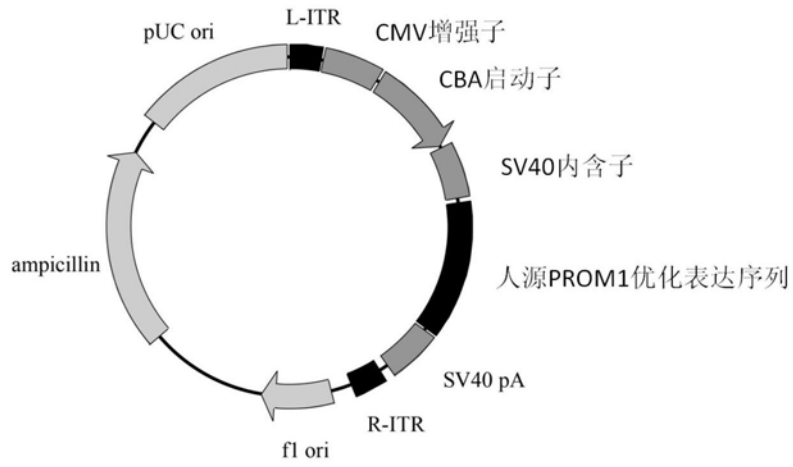


图1

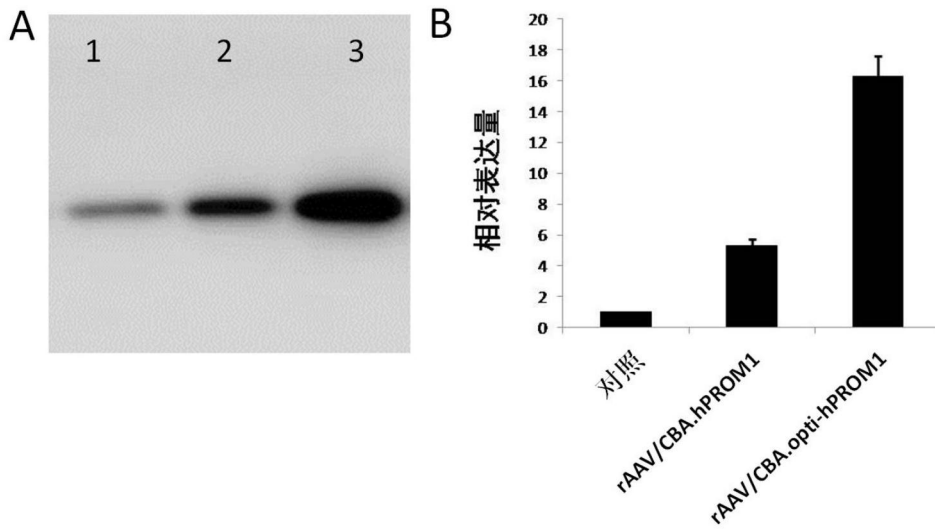


图2

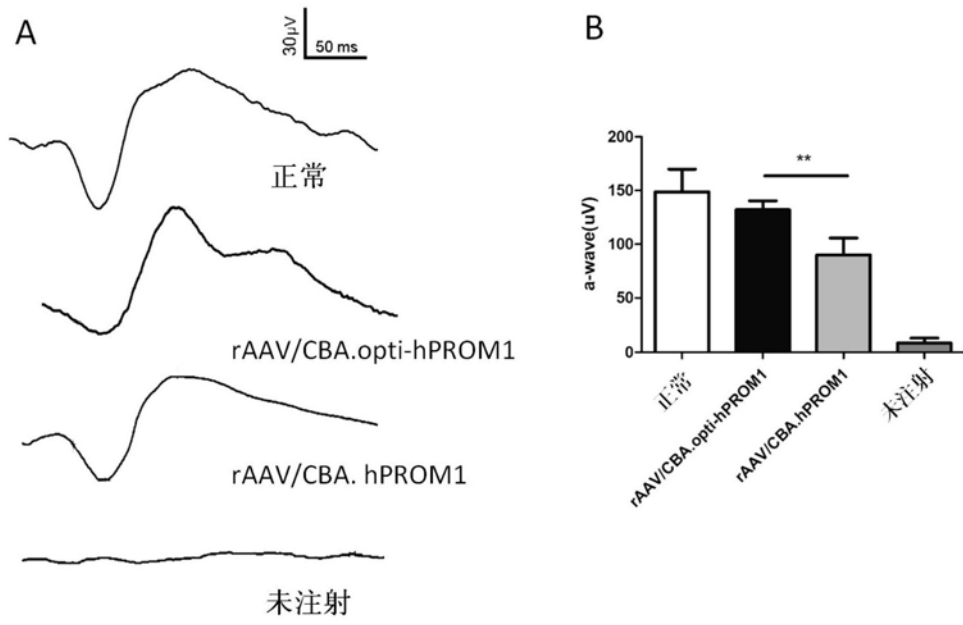


图3

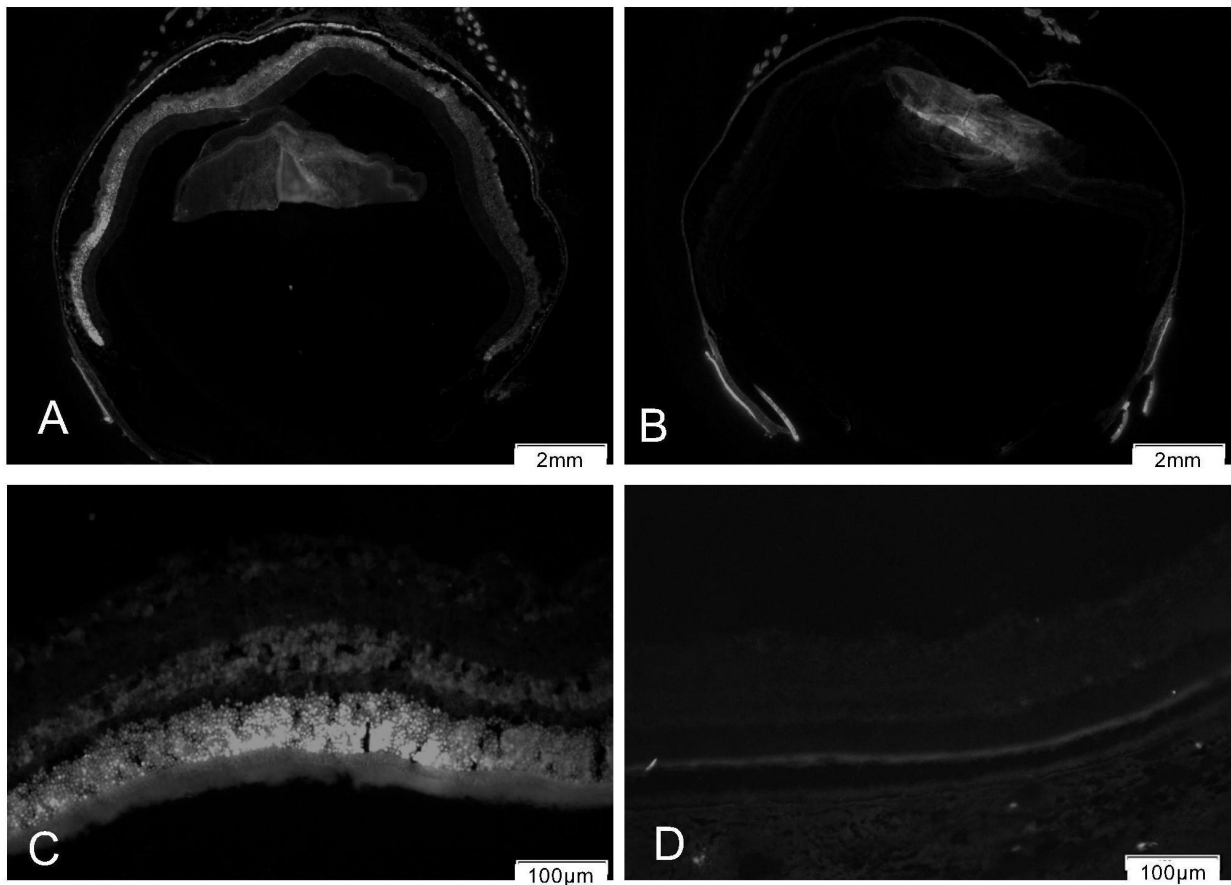


图4

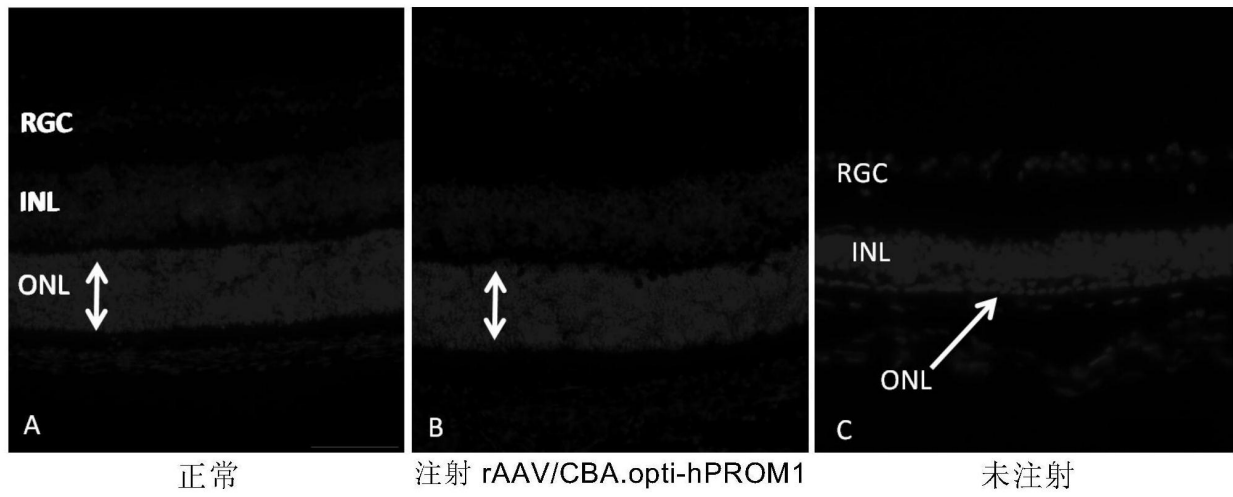


图5