



(12) 发明专利申请

(10) 申请公布号 CN 116286986 A

(43) 申请公布日 2023. 06. 23

(21) 申请号 202211640157.1

(22) 申请日 2017.07.27

(30) 优先权数据
62/368,929 2016.07.29 US

(62) 分案原申请数据
201780050806.3 2017.07.27

(71) 申请人 加利福尼亚大学董事会
地址 美国加利福尼亚
申请人 宾夕法尼亚州大学理事会

(72) 发明人 D·V·莎菲尔 J·G·弗兰纳里
W·A·贝尔特兰 L·C·拜内
M·菲泽尔

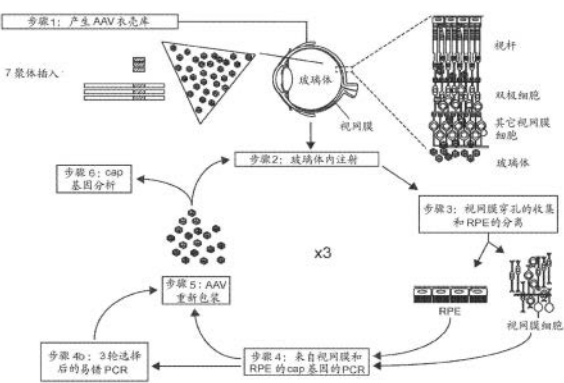
(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038
专利代理师 陈晓娜

(51) Int.Cl.
C12N 15/864 (2006.01)
A61K 48/00 (2006.01)
A61P 27/06 (2006.01)
A61P 27/02 (2006.01)
C12N 7/01 (2006.01)
C12N 15/35 (2006.01)
C07K 14/015 (2006.01)

权利要求书4页 说明书37页
序列表(电子公布) 附图33页

(54) 发明名称
具有变异衣壳的腺相关病毒病毒体和其使用
方法

(57) 摘要
本公开提供了具有改变的衣壳蛋白的腺相
关病毒(AAV)病毒体,其中与野生型AAV相比,所
述AAV病毒体展现对视网膜细胞更大的感染性。
本公开还提供了将基因产物递送至个体中的视
网膜细胞的方法,以及治疗眼病的方法。



1. 一种重组腺相关病毒(rAAV)病毒体,其包含:

a) 变异AAV衣壳蛋白,其中相对于对应的亲本AAV衣壳蛋白,所述变异AAV衣壳蛋白在衣壳蛋白GH环中包含异源肽的插入,其中与包含所述对应的亲本AAV衣壳蛋白的对照AAV病毒体对视网膜细胞的感染性相比,所述变异衣壳蛋白赋予增加的对视网膜细胞的感染性,并且

其中所述异源肽是式I的肽:

$$X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10},$$

其中:

X_1 为Leu、Ile、Pro或Gln;

X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val;

X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr;

X_4 如果存在则为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg;

X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr;

X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile;

X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys;

X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn;

X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala;且

X_{10} 为Ala、Phe、Asp、Thr、Val或Met,并且

其中所述异源肽不包含SEQ ID NO:51、60和53所示的氨基酸序列;以及

b) 异源核酸,所述异源核酸包含编码异源基因产物的核苷酸序列。

2. 根据权利要求1所述的rAAV病毒体,其中与包含所述对应的亲本AAV衣壳蛋白的对照AAV病毒体对视网膜细胞的感染性相比,所述rAAV病毒体展现增加至少5倍的对视网膜细胞的感染性。

3. 根据权利要求1所述的rAAV病毒体,其中与包含所述对应的亲本AAV衣壳蛋白的对照AAV病毒体对视网膜细胞的感染性相比,所述rAAV病毒体展现增加至少10倍的对视网膜细胞的感染性。

4. 根据权利要求1所述的rAAV病毒体,其中与包含所述对应的亲本AAV衣壳蛋白的AAV病毒体定位至内核层、外核层、光感受器层、神经节细胞层或视网膜色素上皮细胞的程度相比,所述rAAV病毒体展现增加至少5倍的至内核层、外核层、光感受器层、神经节细胞层和视网膜色素上皮细胞中的一个或多个的定位。

5. 根据权利要求1所述的rAAV病毒体,其中插入位点介于与AAV2的VP1的氨基酸570和611对应的氨基酸之间,或另一AAV血清型的衣壳蛋白中的对应位置。

6. 根据权利要求4所述的rAAV病毒体,其中插入位点位于与AAV2的VP1的氨基酸587和588对应的氨基酸之间,或另一AAV血清型的衣壳蛋白中的对应位置。

7. 根据权利要求1所述的rAAV病毒体,其中所述异源肽包含氨基酸序列LAANQPSKPA (SEQ ID NO:63)。

8. 根据权利要求1所述的rAAV病毒体,其中所述异源肽包含氨基酸序列PAPQDTTKKA (SEQ ID NO:57)。

9. 根据权利要求1所述的rAAV病毒体,其中所述异源肽包含氨基酸序列LAKGTELKPA

(SEQ ID NO:54)。

10. 根据权利要求1所述的rAAV病毒体,其中所述异源肽包含氨基酸序列LAVDGAQRSA (SEQ ID NO:56)。

11. 根据权利要求1所述的rAAV病毒体,其中所述异源肽包含氨基酸序列LATTSQNKPA (SEQ ID NO:50)。

12. 根据权利要求1至11中任一项所述的rAAV病毒体,其中所述基因产物是短干扰RNA或适体。

13. 根据权利要求1至11中任一项所述的rAAV病毒体,其中所述基因产物是多肽。

14. 根据权利要求13所述的rAAV病毒体,其中所述多肽是神经保护多肽、抗血管生成多肽或增强视网膜细胞的功能的多肽。

15. 根据权利要求13所述的rAAV病毒体,其中所述多肽是RNA引导的核酸内切酶。

16. 根据权利要求15所述的rAAV病毒体,其中所述RNA引导的核酸内切酶是Cas9多肽。

17. 根据权利要求16所述的rAAV病毒体,其中所述基因产物是RNA引导的核酸内切酶和向导RNA。

18. 一种药物组合物,其包含:

a) 权利要求1至17中任一项所述的重组腺相关病毒病毒体;和

b) 药学上可接受的赋形剂。

19. 权利要求1至17中任一项所述的重组腺相关病毒 (rAAV) 病毒体在制备用于将基因产物递送至个体中的视网膜细胞的药物中的用途。

20. 根据权利要求19所述的用途,其中所述基因产物是多肽。

21. 根据权利要求19所述的用途,其中所述基因产物是短干扰RNA或适体。

22. 根据权利要求20所述的用途,其中所述多肽是神经保护多肽、抗血管生成多肽、抗细胞凋亡因子或增强视网膜细胞的功能的多肽。

23. 根据权利要求20所述的用途,其中所述多肽是神经胶质源性神经营养因子、成纤维细胞生长因子2、神经秩蛋白、睫状神经营养因子、神经生长因子、脑源性神经营养因子、表皮生长因子、视紫红质、X连锁凋亡抑制因子、视网膜劈裂蛋白、RPE65、视网膜色素变性GTP酶相互作用蛋白-1、外周蛋白、外周蛋白-2、RdCVF、视网膜色素变性GTP酶调控因子 (RPGR) 或音猬因子。

24. 根据权利要求20所述的用途,其中所述多肽是RNA引导的核酸内切酶。

25. 权利要求1至17中任一项所述的重组腺相关病毒 (rAAV) 病毒体在制备用于治疗个体的眼病的药物中的用途。

26. 根据权利要求25所述的用途,其中所述药物通过眼内注射施用至所述个体。

27. 根据权利要求25所述的用途,其中所述药物通过玻璃体内注射施用。

28. 根据权利要求25所述的用途,其中所述眼病为青光眼、视网膜色素变性、黄斑变性、视网膜劈裂症、利伯氏先天性黑朦、糖尿病性视网膜病、全色盲或色盲。

29. 一种分离的核酸,其包含编码变异腺相关病毒 (AAV) 衣壳蛋白的核苷酸序列,其中相对于对应的亲本AAV衣壳蛋白,所述变异AAV衣壳蛋白在衣壳蛋白GH环中包含异源肽的插入,所述异源肽具有9个氨基酸至20个氨基酸的长度,并且其中所述变异衣壳蛋白在存在于AAV病毒体中时,提供所述AAV病毒体增加的对视网膜细胞的感染性,并且其中所述异源肽

的插入是在天然AAV衣壳的GH环中，

并且其中所述异源肽是式I的肽：

$X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$ ，

其中：

X_1 为Leu、Ile、Pro或Gln；

X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val；

X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr；

X_4 如果存在则为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg；

X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr；

X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile；

X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys；

X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn；

X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala；且

X_{10} 为Ala、Phe、Asp、Thr、Val或Met，并且

其中所述异源肽不包含SEQ ID NO:51、60和53所示的氨基酸序列。

30. 根据权利要求29所述的分离的核酸，其中插入位点介于AAV2的氨基酸587与588之间、AAV1的氨基酸590与591之间、AAV5的氨基酸575与576之间、AAV6的氨基酸590与591之间、AAV7的氨基酸589与590之间、AAV8的氨基酸590与591之间、AAV9的氨基酸588与589之间或AAV10的氨基酸588与589之间。

31. 根据权利要求29所述的分离的核酸，其中所述异源肽包含选自以下的氨基酸序列：

LAKSDQSKPA (SEQ ID NO:61)，LAANQPSKPA (SEQ ID NO:63)，PAPQDTTKKA (SEQ ID NO:57)，LAKGTELKPA (SEQ ID NO:54)，LAVDGAQRSA (SEQ ID NO:56) 和LATTSQNKPA (SEQ ID NO:50)。

32. 一种分离的经遗传修饰的宿主细胞，其包含权利要求29至31中任一项所述的核酸。

33. 一种变异腺相关病毒 (AAV) 衣壳蛋白，其中所述变异AAV衣壳蛋白包含异源肽的插入，所述异源肽具有9个氨基酸至20个氨基酸的长度，并且其中所述异源肽的插入是在天然AAV衣壳的GH环中，其中所述异源肽是式I的肽：

$X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$ ，

其中：

X_1 为Leu、Ile、Pro或Gln；

X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val；

X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr；

X_4 如果存在则为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg；

X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr；

X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile；

X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys；

X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn；

X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala；且

X_{10} 为Ala、Phe、Asp、Thr、Val或Met，并且

其中所述异源肽不包含SEQ ID NO:51、60和53所示的氨基酸序列。

34. 根据权利要求33所述的变异rAAV衣壳蛋白,其中所述异源肽包含选自以下的氨基酸序列:

LAKSDQSKPA (SEQ ID NO:61), LAANQPSKPA (SEQ ID NO:63), PAPQDTTKKA (SEQ ID NO:57), LAKGTELKPA (SEQ ID NO:54), LAVDGAQRSA (SEQ ID NO:56) 和LATTSQNKPA (SEQ ID NO:50)。

具有变异衣壳的腺相关病毒病毒体和其使用方法

[0001] 本申请是分案申请,其相应母案申请是申请号为201780050806.3的中国发明专利申请,申请日为2017年7月27日,发明名称为“具有变异衣壳的腺相关病毒病毒体和其使用方法”。

[0002] 交叉引用

[0003] 本申请要求2016年7月29日提交的美国临时专利申请No.62/368,929的权益,该申请以引用的方式整体并入本文中。

[0004] 关于联邦政府资助的研究的声明

[0005] 本发明是在政府支持下根据美国国家卫生研究院颁发的合同/许可No.EY022975、EY018241和EY06855进行的。政府在本发明中具有一定的权利。

背景技术

[0006] 光感受器是视网膜中接收和加工视觉信息的第一神经元,其将可见的电磁辐射通过光转换转变成超极化反应。绝大多数遗传性视网膜疾病引起这些细胞的丧失,直接引起,例如在影响视紫红蛋白折叠的显性突变中;或间接引起,例如在影响视网膜色素上皮细胞(retinal pigment epithelium,RPE)中的视网膜再循环通路的隐性突变中。

[0007] 腺相关病毒(Adeno-associated virus,AAV)属于成员需要与例如腺病毒等辅助病毒协同感染来促进复制的细小病毒科(Parvoviridae)和依赖性病毒(Dependovirus)属,且AAV在缺少辅助剂时建立潜伏性感染。病毒体由25nm二十面体衣壳环绕4.9kb单链DNA基因组与两个开放阅读框rep和cap构成。非结构性rep基因编码四种对病毒复制来说不可或缺的调节蛋白,而cap编码三种装配成60聚体衣壳的结构蛋白(VP1-3)。此病毒衣壳介导AAV载体克服病毒转导的许多生物障碍的能力,所述生物障碍包括细胞表面受体结合、胞吞作用、胞内运输和核中拆包。

发明内容

[0008] 本公开提供了具有改变的衣壳蛋白的重组腺相关病毒(AAV)病毒体,其中与野生型AAV相比,所述重组AAV(rAAV)病毒体展现更大的对视网膜细胞的感染性,且其中rAAV病毒体包含异源核酸。本公开还提供了将基因产物递送至个体中的视网膜细胞的方法,和治疗眼病的方法。本公开提供了一种rAAV病毒体,其中与包含对应的亲本AAV衣壳蛋白的AAV病毒体定位至内核层、外核层、光感受器层、神经节细胞层或视网膜色素上皮细胞的程度相比,所述rAAV病毒体展现增加至少5倍的至内核层、外核层、光感受器层、神经节细胞层和视网膜色素上皮细胞中的一个或多个的定位。

附图说明

[0009] 图1是用于研发与亲本AAV相比,展现增加的对眼细胞(例如视网膜细胞)的感染性的AAV变体的定向进化法的示意图。

[0010] 图2是含有绿色荧光蛋白(GFP)-条型码构建体的AAV变体的深度测序的示意图。

[0011] 图3描绘18成员AAV变体库对神经节细胞层、内核层、光感受器层和视网膜色素上皮细胞(RPE)层中细胞的感染。

[0012] 图4提供AAV2衣壳蛋白VP1的氨基酸序列。氨基酸587和588(NP)呈粗体且加下划线。

[0013] 图5提供与多种AAV血清型的AAV衣壳蛋白VP1的氨基酸570-610对应的氨基酸序列。

[0014] 图6A-6C提供AAV衣壳蛋白环IV(GH环)区的氨基酸序列的比对。插入位点以粗体和下划线显示。AAV1:SEQ ID NO:35;AAV6:SEQ ID NO:36;AAV3:SEQ ID NO:37;AAV2:SEQ ID NO:38;AAV8:SEQ ID NO:39;AAV8.1:SEQ ID NO:40;AAV8 rh8:SEQ ID NO:41;AAV10:SEQ ID NO:42;AAV7:SEQ ID NO:43;AAV9:SEQ ID NO:44;AAV 9.1:SEQ ID NO:45;AAV5:SEQ ID NO:46。

[0015] 图7A-7V提供例示性异源基因产物的氨基酸序列。

[0016] 图8A-8C提供例示性向导-RNA指导的核酸内切酶的氨基酸序列。

[0017] 定义

[0018] 术语“视网膜细胞”在本文中可以指包含视网膜的细胞类型中的任一种,例如视网膜神经节细胞;无长突细胞;水平细胞;双极细胞;包括视杆和视锥的感光细胞;米勒胶质细胞(Müller glial cell);星形细胞(例如视网膜星形细胞);和视网膜色素上皮细胞。

[0019] “AAV”是腺相关病毒的缩写,且可以用于指病毒本身或其衍生物。除非另外需要,否则该术语涵盖所有亚型和天然存在的形式与重组形式两者。缩写“rAAV”是指重组腺相关病毒,又称为重组AAV载体(或“rAAV载体”)。术语“AAV”包括AAV 1型(AAV-1)、AAV 2型(AAV-2)、AAV 3型(AAV-3)、AAV 4型(AAV-4)、AAV 5型(AAV-5)、AAV 6型(AAV-6)、AAV 7型(AAV-7)、AAV 8型(AAV-8)、AAV 9型(AAV-9)、鸟类AAV、牛类AAV、犬科AAV、马科AAV、灵长类AAV、非灵长类AAV和绵羊科AAV。“灵长类AAV”是指从灵长类动物分离的AAV,“非灵长类AAV”是指从非灵长类哺乳动物分离的AAV,“牛类AAV”是指从牛类哺乳动物(例如母牛)分离的AAV等等。

[0020] 如本文所用的“rAAV载体”是指包含非AAV来源的多核苷酸序列(即对AAV来说异源的多核苷酸),典型地用于细胞的遗传转化的所关注的序列的AAV载体。一般说来,异源多核苷酸侧接至少一个且一般侧接两个AAV反向末端重复序列(ITR)。术语rAAV载体涵盖rAAV载体粒子与rAAV载体质粒两者。

[0021] “AAV病毒”或“AAV病毒粒子”或“rAAV载体粒子”是指由至少一种AAV衣壳蛋白(典型地野生型AAV的全部衣壳蛋白)和包壳的多核苷酸rAAV载体构成的病毒粒子。如果粒子包含异源多核苷酸(即除野生型AAV基因组以外的多核苷酸,例如有待递送至哺乳动物细胞的转基因),那么其典型地称为“rAAV载体粒子”或简称为“rAAV载体”。因此,rAAV粒子的产生当然包括rAAV载体的产生,因为此类载体含于rAAV粒子内。

[0022] “包装”是指引起AAV粒子的装配和包壳的一系列细胞内事件。

[0023] AAV“rep”和“cap”基因是指编码腺相关病毒的复制和包壳蛋白的多核苷酸序列。AAV rep和cap在本文中称为AAV“包装基因”。

[0024] AAV“辅助病毒”是指允许哺乳动物细胞复制和包装AAV(例如野生型AAV)的病毒。多种此类AAV辅助病毒是本领域中已知的,包括腺病毒、疱疹病毒和痘病毒,例如牛痘。腺病毒涵盖许多不同的亚类,不过亚类C的腺病毒5型最常使用。已知人类、非人类哺乳动物和鸟

类来源的许多腺病毒,且可以从例如ATCC等保藏中心获得。疱疹科病毒包括例如单纯疱疹病毒(herpes simplex viruses,HSV)和埃-巴二氏病毒(Epstein-Barr virus,EBV)以及巨细胞病毒(cytomegalovirus,CMV)和假狂犬病毒(pseudorabies virus,PRV);其也可以从例如ATCC等保藏中心获得。

[0025] “辅助病毒功能”是指在辅助病毒基因组中编码的允许AAV复制和包装(结合对本文所述的复制和包装的其它需求)的功能。如本文所述,“辅助病毒功能”可以呈许多方式提供,包括通过提供辅助病毒或提供例如编码必需功能的多核苷酸序列至在转运中的生产细胞。

[0026] “感染”病毒或病毒粒子是包含能够递送至病毒物种向性的细胞中的多核苷酸组分的病毒或病毒粒子。该术语不一定暗示病毒的任何复制能力。如本文所用,“感染性”病毒或病毒粒子是可以接近靶细胞,可以感染靶细胞,且可以在靶细胞中表达异源核酸的病毒或病毒粒子。因此,“感染性”是指病毒粒子接近靶细胞、感染靶细胞和在靶细胞中表达异源核酸的能力。感染性可以指体外感染性或体内感染性。用于计数感染性病毒粒子的测定在本公开中和本领域中别处进行描述。病毒感染性可以表示为感染性病毒粒子与总病毒粒子的比率。总病毒粒子可以表示为病毒基因组(vg)拷贝的数目。病毒粒子在细胞中表达异源核酸的能力可以称为“转导”。病毒粒子在细胞中表达异源核酸的能力可以使用许多技术测定,所述技术包括评定标志物基因,例如绿色荧光蛋白(GFP)测定(例如其中病毒包含编码GFP的核苷酸序列),其中GFP在感染病毒粒子的细胞中产生且检测和/或测量;或例如通过酶联免疫吸附测定(ELISA)测量所产生的蛋白质。病毒感染性可以表示为感染性病毒粒子与总病毒粒子的比率。本领域中已知测定感染性病毒粒子与总病毒粒子的比率的方法。参见例如Grainger等人(2005)Mol. Ther. 11: S337(描述TCID50感染滴度测定);和Zolotukhin等人(1999)Gene Ther. 6: 973。

[0027] “有复制能力”的病毒(例如有复制能力的AAV)是指具有感染性的表型野生型病毒,且在受感染细胞中也能够复制(即在辅助病毒或辅助病毒功能存在下)。在AAV情况下,复制能力一般需要功能性AAV包装基因的存在。一般说来,由于缺乏一种或多种AAV包装基因,如本文所述的rAAV载体在哺乳动物细胞(特别是人类细胞中)是不能复制的。典型地,此类rAAV载体缺乏任何AAV基因序列,从而使通过AAV包装基因与引入的rAAV载体之间的重组产生有复制能力的AAV的可能性降至最低。在许多实施方案中,如本文所述的rAAV载体制剂为含有即使有也极少数的有复制能力的AAV(rcAAV,又称为RCA)(例如每 10^2 rAAV粒子少于约1rcAAV、每 10^4 rAAV粒子少于约1rcAAV、每 10^8 rAAV粒子少于约1rcAAV、每 10^{12} rAAV粒子少于约1rcAAV或无rcAAV)的制剂。

[0028] 术语“多核苷酸”是指任何长度的核苷酸的聚合物形式,包括脱氧核糖核苷酸或核糖核苷酸或其类似物。多核苷酸可以包含经修饰的核苷酸,例如甲基化核苷酸和核苷酸类似物,且可以间杂有非核苷酸组分。如果存在,那么可以在聚合物装配之前或之后给予核苷酸结构修饰。如本文所用的术语多核苷酸可互换地指双链和单链分子。除非另作说明或需要,作为多核苷酸的本文所述的本发明的任何实施方案涵盖双链形式和已知或预测组成双链形式的两种互补单链形式中的每一种。

[0029] 多核苷酸或多肽具有一定百分比的与另一多核苷酸或多肽的“序列一致性”,意指在比对时,当比较两种序列时该百分数的碱基或氨基酸相同。序列相似性可以用许多不同

的方式测定。为了测定序列一致性,可以使用包括可以在万维网的ncbi.nlm.nih.gov/BLAST/BLAST获得的BLAST在内的方法和计算机程序,比对序列。另一比对算法为FASTA,可以在来自Madison,Wisconsin,USA的Genetics Computing Group(GCG)程序包(Oxford Molecular Group公司的全资附属公司)获得。其它比对技术描述于Methods in Enzymology,第266卷:Computer Methods for Macromolecular Sequence Analysis(1996),编辑Doolittle,Academic Press公司,Harcourt Brace&Co.分公司, San Diego, California, USA。尤其关注允许序列中缺口的比对程序。Smith-Waterman为在序列比对中允许缺口的一类算法。参见Meth.Mol.Biol.70:173-187(1997)。此外,使用Needleman和Wunsch比对方法的GAP程序可以用于比对序列。参见J.Mol.Biol.48:443-453(1970)。

[0030] 关注BestFit程序,其使用Smith Waterman的局部同源性算法(Advances in Applied Mathematics 2:482-489(1981)来测定序列一致性。缺口产生罚分一般在1至5范围内,通常2至4且在许多实施方案中将为3。缺口延伸罚分一般在约0.01至0.20范围内,且在许多情况下将为0.10。程序具有由所输入的待比较的序列决定的缺省参数。优选地,使用由程序决定的缺省参数确定序列一致性。此程序也可以从来自Madison,Wisconsin,USA的Genetics Computing Group(GCG)程序包获得。

[0031] 另一所关注的程序为FastDB算法。FastDB描述于Current Methods in Sequence Comparison and Analysis,Macromolecule Sequencing and Synthesis,Selected Methods and Applications,第127-149页,1988,Alan R.Liss公司。序列一致性百分比通过FastDB基于以下参数计算:

[0032] 错配罚分:1.00;

[0033] 缺口罚分:1.00;

[0034] 缺口尺寸罚分:0.33;以及

[0035] 接合罚分:30.0。

[0036] “基因”是指含有至少一个能够在转录和翻译后编码特定蛋白质的开放阅读框的多核苷酸。

[0037] 如本文所用的术语“向导RNA”是指包含以下的RNA:i)结合于向导RNA指导的核酸内切酶(例如2类CRISPR/Cas核酸内切酶,例如II型、V型或VI型CRISPR/Cas核酸内切酶)且活化RNA指导的核酸内切酶的“活化”核苷酸序列;和ii)包含与靶核酸杂交的核苷酸序列的“靶”核苷酸序列。“活化”核苷酸序列和“靶”核苷酸序列可以在分开的RNA分子(例如“双向向导RNA”)上;或可以在相同的RNA分子(“单向向导RNA”)上。

[0038] “小干扰”或“短干扰RNA”或siRNA为靶向所关注的基因(“靶基因”)的核苷酸的RNA双链体。“RNA双链体”是指由RNA分子的两个区域之间的互补配对所形成的结构。siRNA由于siRNA的双链体部分的核苷酸序列与所靶向基因的核苷酸序列互补而“靶向”基因。在一些实施方案中,siRNA的双链体的长度少于30个核苷酸。在一些实施方案中,双链体长度可以为29个、28个、27个、26个、25个、24个、23个、22个、21个、20个、19个、18个、17个、16个、15个、14个、13个、12个、11个或10个核苷酸。在一些实施方案中,双链体的长度为19-25个核苷酸长。siRNA的RNA双链体部分可以为发夹结构的一部分。除双链体部分外,发夹结构可以含有环部分,该环部分位于形成双链体的两个序列之间。环的长度可以变化。在一些实施方案中,环长度为5个、6个、7个、8个、9个、10个、11个、12个或13个核苷酸。发夹结构还可以含有

3'或5'悬垂部分。在一些实施方案中,悬垂物为0个、1个、2个、3个、4个或5个核苷酸长的3'或5'悬垂物。

[0039] 如本文所用,术语“微型RNA”是指任何类型的干扰RNA,包括但不限于内源性微型RNA和人工微型RNA(例如合成miRNA)。内源性微型RNA为在基因组中天然编码的能够调节mRNA的生产性利用的小型RNA。人工微型RNA可以为除内源性微型RNA以外的能够调节mRNA的活性的任何类型的RNA序列。微型RNA序列可以为由这些序列中的任一种或多种构成的RNA分子。微型RNA(或“miRNA”)序列已经描述于例如以下的出版物中:Lim等人,2003, *Genes&Development*, 17, 991-1008; Lim等人,2003, *Science*, 299, 1540; Lee和Ambrose, 2001, *Science*, 294, 862; Lau等人,2001, *Science* 294, 858-861; Lagos-Quintana等人,2002, *Current Biology*, 12, 735-739; Lagos-Quintana等人,2001, *Science*, 294, 853-857; 和Lagos-Quintana等人,2003, *RNA*, 9, 175-179。微型RNA的实例包括作为较大RNA的片段的任何RNA,或为miRNA、siRNA、stRNA、sncRNA、tncRNA、snoRNA、smRNA、shRNA、snRNA或其它小型非编码RNA。参见例如美国专利申请20050272923、20050266552、20050142581和20050075492。“微型RNA前体”(或“前驱miRNA”)是指微型RNA序列并入其中的具有茎环结构的核酸。“成熟微型RNA”(或“成熟miRNA”)包括已经从微型RNA前体(“前驱miRNA”)裂解或已经合成(例如在实验室中通过无细胞合成而合成)的微型RNA,且具有约19个核苷酸至约27个核苷酸的长度,例如成熟微型RNA可以具有19nt、20nt、21nt、22nt、23nt、24nt、25nt、26nt或27nt的长度。成熟微型RNA可以结合于靶mRNA且抑制靶mRNA的翻译。

[0040] 如应用于多核苷酸的“重组”意指多核苷酸为克隆、限制或连接步骤和产生与在自然界中发现的多核苷酸不同的构建体的其它程序的多种组合的产物。重组病毒为包含重组多核苷酸的病毒粒子。所述术语分别包括原始多核苷酸构建体的复制物和原始病毒构建体的后代。

[0041] “控制元件”或“控制序列”为参与促进多核苷酸的功能调控的分子的相互作用,包括多核苷酸的复制、重复、转录、剪接、翻译或降解的核苷酸序列。调控可以影响过程的频率、速度或特异性,且可以为增强性或抑制性的。本领域中已知的控制元件包括例如转录调控序列,例如启动子和增强子。启动子为在一定条件下能够结合核糖核酸聚合酶且开始通常位于启动子下游(在3'方向上)的编码区的转录的DNA区域。

[0042] “可操作地连接(Operatively linked)”或“可操作地连接(operably linked)”是指遗传元件的并置,其中所述元件处于允许其以预期方式操作的关系中。举例来说,如果启动子帮助开始编码序列的转录,那么启动子可操作地连接于编码区。在启动子与编码区之间可能有插入残基,只要维持此功能关系即可。

[0043] “表达载体”为包含编码所关注的多肽的区域的载体,且用于实现蛋白质在预期靶细胞中的表达。表达载体还包含可操作地连接于编码区以促进蛋白质在靶中的表达的控制元件。控制元件与其可操作地连接以求表达的基因的组合有时称为“表达盒”,本领域中已知和可获得许多表达盒,或容易由本领域中可获得的组分构建。

[0044] “异源”意指衍生自一种实体,该实体在基因型上截然不同于正被比较实体的其余部分。举例来说,通过遗传工程技术引入源自于不同物种的质粒或载体中的多核苷酸为异源多核苷酸。从天然编码序列去除且可操作地连接于在自然界中未发现连接的编码序列的启动子为异源启动子。因此,举例来说,包括编码异源基因产物的异源核酸的rAAV为包括天

然存在的野生型AAV中通常不包括的核酸的rAAV,且编码的异源基因产物为天然存在的野生型AAV通常不编码的基因产物。作为另一实例,包含插入衣壳蛋白的GH环中的异源肽的变异AAV衣壳蛋白为包括天然存在的野生型AAV中通常不包括的肽插入的变异AAV衣壳蛋白。

[0045] 术语“遗传改变”和“遗传修饰”(和其语法变体)在本文中可互换使用,是指其中不通过有丝分裂或减数分裂,将遗传元件(例如多核苷酸)引入细胞的过程。所述元件可以与细胞异源,或其可能为细胞中已经存在的元件的额外拷贝或改进型式。遗传改变可以例如通过用重组质粒或其它多核苷酸通过本领域中已知的任何方法,例如电穿孔、磷酸钙沉淀或接触多核苷酸-脂质粒复合物转染细胞来实现。遗传改变还可以例如通过用DNA或RNA病毒或病毒载体转导或感染来实现。一般地,遗传元件被引入细胞中的染色体或极微染色体中;但改变细胞和其后代的表型和/或基因型的任何改变都包括在此术语中。

[0046] 如果在细胞的体外长期培养期间序列可以用来执行其功能,那么称细胞经遗传序列“稳定地”改变、转导、基因修饰或转化。一般地,此类细胞为“遗传上”改变(遗传上修饰),因为引入了所改变的细胞的后代也能够遗传的遗传改变。

[0047] 术语“多肽”、“肽”和“蛋白质”在本文中可互换使用,是指任何长度的氨基酸的聚合物。所述术语还涵盖已经修饰的氨基酸聚合物;例如二硫键形成、糖基化、脂质化、磷酸化或与标记组分偶联。在基因产物递送至哺乳动物受试者的背景下论述时例如抗血管生成多肽、神经保护多肽等多肽和其组合物是指相应的完整多肽或保留完整蛋白质的所需生物化学功能的其任何片段或基因工程衍生物。类似地,提及编码抗血管生成多肽的核酸、编码神经保护多肽的核酸和用于基因产物递送至哺乳动物受试者中的其它此类核酸(可以称为递送至受体细胞的“转基因”)包括编码完整多肽或具有所需生物化学功能的任何片段或基因工程衍生物的多核苷酸。

[0048] “分离”的质粒、核酸、载体、病毒、病毒体、宿主细胞或其它物质是指缺乏物质或类似物质天然存在或最初制备时也可能存在的其它组分中的至少一些的物质制剂。因此,举例来说,分离的物质可以通过使用纯化技术将其从来源混合物富集来制备。富集可以在绝对基础上测量,例如每体积溶液的重量,或其可以相对于来源混合物中存在的可能干扰的第二物质测量。本发明的实施方案的日益富集,日益更多地分离。在一些实施方案中,分离的质粒、核酸、载体、病毒、宿主细胞或其它物质经纯化,例如约80%至约90%纯、至少约90%纯、至少约95%纯、至少约98%纯或至少约99%纯或更纯。

[0049] 如本文所用,术语“治疗(treatment)”、“治疗(treating)”等等是指获得所需药理学和/或生理学作用。就完全或部分预防疾病或其症状而言,作用可以为预防性的,和/或就部分或完全治愈疾病和/或可归因于所述疾病的不利影响而言,作用可以为治疗性的。如本文所用,“治疗”涵盖哺乳动物、尤其人类中疾病的任何治疗,且包括:(a)预防疾病在易感染疾病或处于得病的风险中但尚未被诊断为患病的受试者中发生;(b)抑制疾病,即抑制其发展;和(c)减轻疾病,即引起疾病消退。

[0050] 术语“个体”、“宿主”、“受试者”和“患者”在本文中可互换使用,且是指哺乳动物,包括(但不限于)人类和非人类灵长类动物,包括猿和人类;哺乳动物运动动物(例如马、骆驼等);哺乳动物农畜(例如绵羊、山羊、母牛等);哺乳动物宠物(犬、猫等);和啮齿类动物(例如小鼠、大鼠等)。在一些情况下,个体为人类。

[0051] 在进一步描述本发明前,应了解本发明不局限于所述的具体实施方案,因而当然

可以变化。还应了解本文中使用的术语仅仅是出于描述具体实施方案的目的,且不意图限制,因为本发明的范围仅仅受随附权利要求书限制。

[0052] 在提供值的范围的情况下,应了解除非上下文另外清楚地规定,否则本发明内涵盖该范围的上下限之间的每个中间值,至下限的单位的十分之一,和所说明范围内的任何其它所说明或中间值。这些较小范围的上下限可以独立地包括在较小范围中,且也涵盖于本发明内,服从于所说明范围中的任何特别排除的界限。在所说明的范围包括界限中的一个或两个的情况下,排除那些包括的界限中的任一个或两个的范围也包括在本发明中。

[0053] 除非另外定义,否则本文中使用的所有技术和科学术语都具有与本发明所属领域的一般技术人员通常所了解的含义相同的含义。虽然与本文所述的方法和材料类似或同等的任何方法和材料也可以用于实践或测试本发明,但是现在描述优选的方法和材料。本文中提到的所有出版物都以引用的方式并入本文中以公开和描述引用出版物所相关的方法和/或材料。

[0054] 必须指出如本文和随附权利要求书中所用,除非上下文另外清楚规定,否则单数形式“一种(a/an)”、“或”和“所述”包括多个指示物。因此,举例来说,提及“一种AAV衣壳”包括多个此类衣壳,且提及“所述AAV病毒体”包括提及一种或多种AAV病毒体和本领域的技术人员已知的其同等物等等。进一步指出,权利要求书可以被撰写成排除任何任选的要素。因而,此陈述意图作为与权利要求要素的叙述一起使用排他性术语如“单独地”、“仅”等,或采用“负面”限制的先前基础。

[0055] 应了解,为清楚起见,在分开实施方案的上下文中描述的本发明的某些特征也可能组合提供于单个实施方案。相反地,为简便起见,在单个实施方案的上下文中描述的本发明的多种特征也可以分开或呈任何合适的子组合提供。本发明特别涵盖关于本发明的实施方案的所有组合且在本文中公开,如同每个组合个别且明确地公开一般。另外,本发明也特别涵盖多种实施方案和其要素的所有子组合,且在本文中公开,如同每个此类子组合个别且明确地在本文中公开一般。

[0056] 本文中论述的出版物仅仅针对其在本申请的提交日期前的公开内容提供。本文中不应有内容解释为承认本发明无权借助在先发明而先于此类出版物。此外,所提供的公布日期可以与实际公布日期不同,实际公布日期可以独立地证实。

具体实施方式

[0057] 本公开提供了具有改变的衣壳蛋白的重组腺相关病毒(AAV)病毒体,其中与野生型AAV相比,所述重组AAV(rAAV)病毒体展现更大的对视网膜细胞的感染性;且其中rAAV病毒体包含异源核酸。与对应的野生型AAV对视网膜细胞的感染性相比,rAAV病毒体展现更大的对视网膜细胞的感染性。视网膜细胞可以为光感受器(例如视杆;视锥)、视网膜神经节细胞(RGC)、米勒细胞(米勒胶质细胞)、星形细胞(例如视网膜星形细胞)、双极细胞、无长突细胞、水平细胞或视网膜色素上皮细胞(RPE)。本公开还提供了将基因产物递送至个体中的视网膜细胞的方法,和治疗眼病的方法。本公开提供了一种具有改变的衣壳蛋白的rAAV病毒体,其中与包含对应的亲本AAV衣壳蛋白的AAV病毒体定位至内核层、外核层、光感受器层、神经节细胞层或视网膜色素上皮细胞的程度相比,所述rAAV病毒体展现增加至少5倍的至内核层、外核层、光感受器层、神经节细胞层和视网膜色素上皮细胞中的一个或多个的定

位;且其中所述rAAV病毒体包含异源核酸。

[0058] 变异AAV衣壳多肽

[0059] 本公开提供了一种变异AAV衣壳蛋白。本公开的变异AAV衣壳蛋白在亲本AAV衣壳蛋白的表面可及(例如溶剂可及)部分中的插入位点中包含5个氨基酸至20个氨基酸长的异源肽的插入,使得与包含对应的亲本AAV衣壳蛋白的AAV病毒体对视网膜细胞的感染性相比,所述变异衣壳蛋白在存在于AAV病毒体中时赋予增加的对视网膜细胞的感染性。在其它情况下,视网膜细胞为米勒细胞。其它视网膜细胞包括无长突细胞、双极细胞和水平细胞。“约5个氨基酸至约20个氨基酸的插入”在本文中也称为“肽插入”(例如异源肽插入)。“对应的亲本AAV衣壳蛋白”是指没有肽插入的相同AAV血清型的AAV衣壳蛋白。在许多情况下,变异AAV衣壳包含5个氨基酸至20个氨基酸(例如5至7个、7至10个、10至12个、12至15个或15至20个氨基酸)长的单一异源肽插入。

[0060] 插入位点在AAV衣壳蛋白的GH环或环IV中,例如AAV衣壳蛋白的GH环或环IV的溶剂可及部分中。关于AAV衣壳的GH环/环IV,参见例如van Vliet等人(2006)Mol. Ther. 14:809; Padron等人(2005)J. Virol. 79:5047;和Shen等人(2007)Mol. Ther. 15:1955。举例来说,插入位点可以在AAV衣壳蛋白的氨基酸411-650内,如图6A-6C中所描绘。举例来说,插入位点可以在AAV2的氨基酸570-611内,AAV1的氨基酸571-612内,AAV5的氨基酸560-601内,AAV6的氨基酸571至612内,AAV7的氨基酸572至613内,AAV8的氨基酸573至614内,AAV9的氨基酸571至612内,或AAV10的氨基酸573至614内,如图5中所描绘。在一些情况下,插入位点介于AAV2衣壳蛋白的氨基酸588与589之间,或不同血清型的AAV中的对应插入位点。在一些情况下,插入位点介于AAV2衣壳蛋白的氨基酸587与588之间,或不同血清型的AAV中的对应插入位点。

[0061] 在一些情况下,相对于对应的亲本AAV衣壳蛋白,约5个氨基酸至约20个氨基酸(例如5至7个、7至10个、10至12个、12至15个或15至20个氨基酸)长的异源肽插入衣壳蛋白的GH环或环IV中的插入位点中。举例来说,插入位点可以介于AAV2的氨基酸587与588之间,或另一AAV血清型的衣壳子单元的对位位置。应注意到,插入位点587/588是基于AAV2衣壳蛋白。5个氨基酸至约20个氨基酸(例如5至7个、7至10个、10至12个、12至15个或15至20个氨基酸)长的异源肽可以插入除AAV2以外的AAV血清型(例如AAV8、AAV9等)中的对应位点中。基于多种AAV血清型的衣壳蛋白的氨基酸序列的比较,本领域的技术人员将知道“与AAV2的氨基酸587-588对应”的插入位点在任何既定的AAV血清型的衣壳蛋白中的位置。多种AAV血清型中与AAV2的衣壳蛋白VP1的氨基酸570-611(参见图4)对应的序列展示于图5中。参见例如GenBank登记号NP_049542, AAV1; GenBank登记号AAD13756, AAV5; GenBank登记号AAB95459, AAV6; GenBank登记号YP_077178, AAV7; GenBank登记号YP_077180, AAV8; GenBank登记号AAS99264, AAV9; 和GenBank登记号AAT46337, AAV10。

[0062] 举例来说,插入位点可以介于AAV2的氨基酸587与588之间,AAV1的氨基酸590与591之间,AAV5的氨基酸575与576之间,AAV6的氨基酸590与591之间,AAV7的氨基酸589与590之间,AAV8的氨基酸590与591之间,AAV9的氨基酸588与589之间,或AAV10的氨基酸588与589之间。图5中插入位点加下划线;氨基酸编号是基于图5中描绘的编号。

[0063] 在一些实施方案中,主题衣壳蛋白包括氨基酸序列与图6A-6C中阐述的氨基酸序列具有至少约85%、至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列

一致性且插入5至20个氨基酸(例如5至7个、7至10个、10至12个、12至15个或15至20个氨基酸)长的异源肽的GH环。

[0064] 插入肽

[0065] 如上所述,将约5个氨基酸至约20个氨基酸长的异源肽插入AAV衣壳的GH环中。在一些情况下,插入肽具有5个氨基酸至20个氨基酸的长度。在一些情况下,插入肽具有7个氨基酸至15个氨基酸的长度。在一些情况下,插入肽具有9个氨基酸至15个氨基酸的长度。在一些情况下,插入肽具有9个氨基酸至12个氨基酸的长度。插入肽具有5个氨基酸、6个氨基酸、7个氨基酸、8个氨基酸、9个氨基酸、10个氨基酸、11个氨基酸、12个氨基酸、13个氨基酸、14个氨基酸、15个氨基酸、16个氨基酸、17个氨基酸、18个氨基酸、19个氨基酸或20个氨基酸的长度。在一些情况下,插入肽具有7个氨基酸的长度。在一些情况下,插入肽具有8个氨基酸的长度。在一些情况下,插入肽具有9个氨基酸的长度。在一些情况下,插入肽具有10个氨基酸的长度。在一些情况下,插入肽具有11个氨基酸的长度。在一些情况下,插入肽具有12个氨基酸的长度。在一些情况下,插入肽具有13个氨基酸的长度。在一些情况下,插入肽具有14个氨基酸的长度。在一些情况下,插入肽具有15个氨基酸的长度。

[0066] 在一些情况下,肽插入物为式I的肽:

[0067] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0068] X_1 为Leu、Ile、Pro或Gln;

[0069] X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val;

[0070] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr;

[0071] X_4 (如果存在)为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg;

[0072] X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr;

[0073] X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile;

[0074] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys;

[0075] X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn;

[0076] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala;且

[0077] X_{10} 为Ala、Phe、Asp、Thr、Val或Met。

[0078] 式I的肽插入物包括(但不限于)(1) LAKDATKNA (SEQ ID NO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ ID NO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ ID NO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ ID NO:56); (11) PAPQDTTKKA (SEQ ID NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ ID NO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ ID NO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); (18) LAVSDSTKAA (SEQ ID NO:64); (19) LAAQGTAKKPA (SEQ ID NO:65); (20) LAPDQTTRNA (SEQ ID NO:66); (21) LAASDSTKAA (SEQ ID NO:67); (22) LAPQDTTKNA (SEQ ID NO:68); (23) LAKADETRPA (SEQ ID NO:69); (24) LAHQDTAKNA (SEQ ID NO:70); (25) LAHQDTKKNA (SEQ ID NO:71); (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ ID NO:74); (29) LAHQDTTNA (SEQ ID NO:75); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); (32) LAISDQSKPA (SEQ ID NO:78); (33) LADATKTA

(SEQ ID NO:79); (34) LAKDTTKNA (SEQ ID NO:80); (35) LAKSDQSRPA (SEQ IDNO:81); (36) LAPQDTTKNA (SEQ ID NO:82); (37) LATSDSTKAA (SEQ ID NO:83); (38) LAVDGSQRSA (SEQ ID NO:84); (39) LPISDQTKHA (SEQ ID NO:85); (40) LPKDATKTIA (SEQ IDNO:86); (41) LPPQDTTKNA (SEQ ID NO:87); (42) PAPQDTTKNA (SEQ ID NO:88); (43) QAHQDTTKNA (SEQ ID NO:89); (44) LAHETSPRPA (SEQ ID NO:90); (45) LAKSTSTAPA (SEQ IDNO:91); (46) LADQDTTKNA (SEQ ID NO:92); (47) LAESDQSKPA (SEQ ID NO:93); (48) LAHKDTTKNA (SEQ ID NO:94); (49) LAHKTQQKM (SEQ ID NO:95); (50) LAHQDTTENA (SEQ IDNO:96); (51) LAHQDTTINA (SEQ ID NO:97); (52) LAHQDTTKKT (SEQ ID NO:98); (53) LAHQDTTKND (SEQ ID NO:99); (54) LAHQDTTKNT (SEQ ID NO:100); (55) LAHQDTTKNV (SEQ IDNO:101); (56) LAHQDTTKTM (SEQ ID NO:102); (57) LAHQNTTKNA (SEQ ID NO:103); (58) LAHRDTTKNA (SEQ IDNO:104); (59) LAISDQTNHA (SEQ ID NO:105); (60) LAKQKSASTA (SEQ ID NO:106); (61) LAKSDQCKPA (SEQ ID NO:107); (62) LAKSDQSKPD (SEQ ID NO:108); (63) LAKSDQSNPA (SEQ IDNO:109); (64) LAKSYQSKPA (SEQ ID NO:110); (65) LANQDTTKNA (SEQ ID NO:111); (66) LAPQNTTKNA (SEQ ID NO:112); (67) LAPSSIQKPA (SEQ ID NO:113); (68) LAQQDTTKNA (SEQ IDNO:114); (69) LAYQDTTKNA (SEQ ID NO:115); (70) LDHQDTTKNA (SEQ ID NO:116); (71) LDHQDTTKSA (SEQ IDNO:117); (72) LGHQDTTKNA (SEQ ID NO:118); (73) LPHQDTTKND (SEQ ID NO:119); (74) LPHQDTTKNT (SEQ ID NO:120); (75) LPHQDTTNNNA (SEQ ID NO:121); (76) LTHQDTTKNA (SEQ IDNO:122); (77) LTKDATKTIA (SEQ ID NO:123); (78) LTPQDTTKNA (SEQ ID NO:124); 和 (79) LVHQDTTKNA (SEQ ID NO:125)。

[0079] 在一些情况下,肽插入物为式II的肽:

[0080] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0081] X_1 为Leu、Ile或Pro;

[0082] X_2 为Ala、Pro或Ser;

[0083] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg或Ala;

[0084] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile或Asn;

[0085] X_5 为Asp、Ser、Gln、Val、Thr、Gly或Ala;

[0086] X_6 为Thr、Ala、Gln、Ser、Glu或Pro;

[0087] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys或Ala;

[0088] X_8 为Lys、Ser、Arg或Thr;

[0089] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且

[0090] X_{10} 为Ala。

[0091] 式II的肽插入物包括(但不限于) (1) LAKDATKNA (SEQ IDNO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ ID NO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ IDNO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ ID NO:56); (11) PAPQDTTKKA (SEQ ID NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ IDNO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ ID NO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); (18) LAVSDSTKAA (SEQ IDNO:64); (19) LAAQGTAKKPA (SEQ ID NO:65); (20) LAPDQTTRNA (SEQ ID

NO:66); (21) LAASDSTKAA (SEQ ID NO:67); (22) LAPQDTTKNA (SEQ ID NO:68); (23) LAKADETRPA (SEQ IDNO:69); (24) LAHQDTAKNA (SEQ ID NO:70); (25) LAHQDTKKNA (SEQ ID NO:71); (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ IDNO:74); (29) LAHQDTTNA (SEQ ID NO:75); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); (32) LAISDQSKPA (SEQ ID NO:78); (33) LADATKTA (SEQ ID NO:79); (34) LAKDTTKNA (SEQ ID NO:80); (35) LAKSDQSRPA (SEQ IDNO:81); (36) LAPQDTKKNA (SEQ ID NO:82); (37) LATSDSTKAA (SEQ ID NO:83); (38) LAVDGSQRSA (SEQ ID NO:84); (39) LPISDQTKHA (SEQ ID NO:85); (40) LPKDATKTIA (SEQ IDNO:86); (41) LPPQDTTKNA (SEQ ID NO:87); 和 (42) PAPQDTTKNA (SEQ ID NO:88)。

[0092] 式II的肽包括(但不限于): (1) LAKDATKNA (SEQ ID NO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ IDNO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ ID NO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ IDNO:56); (11) PAPQDTTKKA (SEQ ID NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ ID NO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ IDNO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); 和 (18) LAVSDSTKAA (SEQ ID NO:64)。在一些情况下, 肽插入物为 (1) LAKDATKNA (SEQ ID NO:47)。在一些情况下, 肽插入物为 (2) PAHQDTTKNA (SEQ ID NO:48)。在一些情况下, 肽插入物为 (3) LAHQDTTKNA (SEQ ID NO:49)。在一些情况下, 肽插入物为 (4) LATTSQNKPA (SEQ ID NO:50)。在一些情况下, 肽插入物为 (5) LAISDQTKHA (SEQ ID NO:51)。在一些情况下, 肽插入物为 (6) IARGVAPSSA (SEQ ID NO:52)。在一些情况下, 肽插入物为 (7) LAPDSTTRSA (SEQ ID NO:53)。在一些情况下, 肽插入物为 (8) LAKGTELKPA (SEQ ID NO:54)。在一些情况下, 肽插入物为 (9) LAIIDATKNA (SEQ ID NO:55)。在一些情况下, 肽插入物为 (10) LAVDGAQRSA (SEQ ID NO:56)。在一些情况下, 肽插入物为 (11) PAPQDTTKKA (SEQ ID NO:57)。在一些情况下, 肽插入物为 (12) LPHQDTTKNA (SEQ ID NO:58)。在一些情况下, 肽插入物为 (13) LAKDATKTIA (SEQ ID NO:59)。在一些情况下, 肽插入物为 (14) LAKQQSASTA (SEQ ID NO:60)。在一些情况下, 肽插入物为 (15) LAKSDQSKPA (SEQ ID NO:61)。在一些情况下, 肽插入物为 (16) LSHQDTTKNA (SEQ ID NO:62)。在一些情况下, 肽插入物为 (17) LAANQPSKPA (SEQ ID NO:63)。在一些情况下, 肽插入物为 (18) LAVSDSTKAA (SEQ ID NO:64)。

[0093] 在一些情况下, 肽插入物为式III的肽:

[0094] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0095] X_1 为Leu、Ile或Pro;

[0096] X_2 为Ala、Pro或Ser;

[0097] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg或Ala;

[0098] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile或Asn;

[0099] X_5 为Asp、Ser、Gln、Val、Thr、Gly或Ala;

[0100] X_6 为Thr、Ala、Gln、Ser、Glu或Pro;

[0101] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys或Ala;

[0102] X_8 为Lys、Ser、Arg或Thr;

[0103] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala;且

[0104] X_{10} 为Ala、Thr、Asp Val或Met。

[0105] 式III的肽插入物包括(但不限于)(1)LAKDATKNA(SEQ IDNO:47);(2)PAHQDTTKNA(SEQ ID NO:48);(3)LAHQDTTKNA(SEQ ID NO:49);(6)IARGVAPSSA(SEQ ID NO:52);(7)LAPDSTTRSA(SEQ ID NO:53);(8)LAKGTELKPA(SEQ ID NO:54);(9)LAIIDATKNA(SEQ ID NO:55);(10)LAVDGAQRSA(SEQ IDNO:56);(11)PAPQDTTKKA(SEQ ID NO:57);(12)LPHQDTTKNA(SEQ ID NO:58);(13)LAKDATKTIA(SEQ ID NO:59);(14)LAKQQSASTA(SEQ ID NO:60);(16)LSHQDTTKNA(SEQ IDNO:62);(17)LAANQPSKPA(SEQ ID NO:63);(18)LAVSDSTKAA(SEQ ID NO:64);(19)LAAQGTAKPA(SEQ ID NO:65);(20)LAPDQTTRNA(SEQ ID NO:66);(24)LAHQDTAKNA(SEQ IDNO:70);(25)LAHQDTKKNA(SEQ ID NO:71);(26)LAHQDTTKHA(SEQ ID NO:72);(27)LAHQDTTKKA(SEQ ID NO:73);(28)LAHQDTTRNA(SEQ ID NO:74);(29)LAHQDTTTNA(SEQ IDNO:75);(30)LAHQGTTKNA(SEQ ID NO:76);(21)LAASDSTKAA(SEQ ID NO:67);(22)LAPQDTTKNA(SEQ ID NO:68);(31)LAHQVTTKNA(SEQ ID NO:77);(33)LAKDATKTA(SEQ IDNO:79);(34)LAKDTTKNA(SEQ ID NO:80);(36)LAPQDTKKNA(SEQ ID NO:82);(37)LATSDSTKAA(SEQ ID NO:83);(38)LAVDGSQRSA(SEQ ID NO:84);(41)LPPQDTTKNA(SEQ IDNO:87);(42)PAPQDTTKNA(SEQ ID NO:88);(52)LAHQDTTKKT(SEQ ID NO:98);(53)LAHQDTTKND(SEQ ID NO:99);(54)LAHQDTTKNT(SEQ ID NO:100);(55)LAHQDTTKNV(SEQ IDNO:101);(56)LAHQDTTKTM(SEQ ID NO:102);(73)LPHQDTTKND(SEQ ID NO:119);和(74)LPHQDTTKNT(SEQ ID NO:120)。

[0106] 在一些情况下,肽插入物为式IV的肽:

[0107] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$,其中:

[0108] X_1 为Leu;

[0109] X_2 为Ala;

[0110] X_3 为Lys、His、Thr、Ile、Pro或Val;

[0111] X_4 (如果存在)为Gln、Asp、Ser或Gly;

[0112] X_5 为Asp、Ser或Gln;

[0113] X_6 为Thr、Ala、Gln或Ser;

[0114] X_7 为Thr或Ser;

[0115] X_8 为Lys、Ser或Arg;

[0116] X_9 为Asn、Pro或Ser;且

[0117] X_{10} 为Ala。

[0118] 式IV的肽插入物包括(但不限于)(1)LAKDATKNA(SEQ IDNO:47);(3)LAHQDTTKNA(SEQ ID NO:49);(7)LAPDSTTRSA(SEQ ID NO:53);(15)LAKSDQSKPA(SEQ ID NO:61);(20)LAPDQTTRNA(SEQ ID NO:66);(22)LAPQDTTKNA(SEQ ID NO:68);(28)LAHQDTTRNA(SEQ ID NO:74);(32)LAISDQSKPA(SEQ IDNO:78);(34)LAKDTTKNA(SEQ ID NO:80);和(35)LAKSDQSRPA(SEQ ID NO:81)。

[0119] 在一些情况下,肽插入物为式V的肽:

[0120] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$,其中:

[0121] X_1 为Leu;

- [0122] X_2 为Ala;
- [0123] X_3 为Lys或His;
- [0124] X_4 (如果存在) 为Gln、Asp、Ser或Gly;
- [0125] X_5 为Asp、Ser或Gln;
- [0126] X_6 为Thr、Ala、Gln或Ser;
- [0127] X_7 为Thr或Ser;
- [0128] X_8 为Lys、Ser或Arg;
- [0129] X_9 为Asn、Pro或Ser;且
- [0130] X_{10} 为Ala。
- [0131] 式V的肽插入物包括(但不限于)(1) LAKDATKNA (SEQ ID NO:47); (15) LAKSDQSKPA (SEQ ID NO:51); (34) LAKDTTKNA (SEQ ID NO:80); 和(35) LAKSDQSRPA (SEQ ID NO:81)。
- [0132] 在一些情况下,肽插入物为式VI的肽:
- [0133] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:
- [0134] X_1 为Leu;
- [0135] X_2 为Ala;
- [0136] X_3 为Asn、Lys、Thr、Gln、Ser、Ile或Leu;
- [0137] X_4 为Ser、Ala、Thr、Glu、Gln、Gly、Lys或Pro;
- [0138] X_5 为Asp、Pro、Glu、Thr、Asn或Arg;
- [0139] X_6 为Ile、His、Thr、Gln、Asn、Tyr、Asp或Glu;
- [0140] X_7 为Gln、Thr、Asn、Ala或Lys;
- [0141] X_8 为Lys、Thr、Arg或Asp;
- [0142] X_9 为Pro、Asn、Thr、Arg、Lys或Ser;且
- [0143] X_{10} 为Ala。
- [0144] 式VI的肽包括(但不限于): (80) LAKANQNTPA (SEQ ID NO:126); (81) LATTPITKPA (SEQ ID NO:127); (82) LATTPIAKPA (SEQ ID NO:128); (83) LAIEDHTKSA (SEQ ID NO:129); (84) LAQSEHQRP (SEQ ID NO:130); (85) LAKSPNKDNA (SEQ ID NO:131); (86) LANQDYTKTA (SEQ ID NO:132); (87) LANSTDQTRA (SEQ ID NO:133); (88) LALGETTRPA (SEQ ID NO:134); (89) LANSTEQTRA (SEQ ID NO:135); (90) LAQADTTKNA (SEQ ID NO:136); (91) LASKDITKTA (SEQ ID NO:137); 和(92) LASPRHNKKC (SEQ ID NO:138)。
- [0145] 在一些情况下,肽插入物为式VII的肽: LAHQDTTKX₁X₂X₃ (SEQ ID NO:148), 其中 X_1 为Lys、Thr、Asn或His; X_2 为Ala、Thr、Val、Ile、Met或Asp; 且 X_3 如果存在, 那么为Ala。式VII的肽包括(但不限于): (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (52) LAHQDTTKKT (SEQ ID NO:98); (53) LAHQDTTKND (SEQ ID NO:99); (54) LAHQDTTKNT (SEQ ID NO:100); (55) LAHQDTTKNV (SEQ ID NO:101); (56) LAHQDTTKTM (SEQ ID NO:102); 和(93) LAHQDTTKTIA (SEQ ID NO:139)。
- [0146] 在一些情况下,肽插入物为式VIII的肽: LAX₁QX₂TX₃X₄X₅X₆ (SEQ ID NO:149), 其中 X_1 为Ala、Pro、Asp或His; X_2 为Gly或Asp; X_3 为Ala、Thr或Lys; X_4 为Asn、Glu、Lys、Arg或Thr; X_5 为Leu、Asn、Lys或Thr; 且 X_6 如果存在, 那么为Ala、Thr、Asp、Val或Met。式VIII的肽包括(但不限于)(94) LAAQGTANL (SEQ ID NO:140); (22) LAPQDTTKNA (SEQ ID NO:68); (46)

LADQDTTKNA (SEQ ID NO:92); (24) LAHQDTAKNA (SEQ ID NO:70); (25) LAHQDTKKNA (SEQ ID NO:71); (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ ID NO:74); (29) LAHQDTTNA (SEQ ID NO:75); (50) LAHQDTTENA (SEQ ID NO:96); (51) LAHQDTTINA (SEQ ID NO:97); (52) LAHQDTTKKT (SEQ ID NO:98); (53) LAHQDTTKND (SEQ ID NO:99); (54) LAHQDTTKNT (SEQ ID NO:100); (55) LAHQDTTKNV (SEQ ID NO:101); 和 (56) LAHQDTTKTM (SEQ ID NO:102)。

[0147] 在一些情况下,肽插入物为式IX的肽: $X_1AX_2X_3DX_4TKX_5A$ (SEQ ID NO:150), 其中 X_1 为Val或Leu; X_2 为Ile、Val、His或Asp; X_3 为Glu、Ser、Lys或Gln; X_4 为His、Ser或Thr;且 X_5 为Ser、Ala、Asn、His或Lys。式IX的肽包括(但不限于)(95) VAIEDHTKSA (SEQ ID NO:141); (18) LAVSDSTKAA (SEQ ID NO:64); (46) LADQDTTKNA (SEQ ID NO:92); (48) LAHKDTTKNA (SEQ ID NO:94); (26) LAHQDTTKHA (SEQ ID NO:72); 和 (27) LAHQDTTKKA (SEQ ID NO:73)。

[0148] 在一些情况下,肽插入物为式X的肽: $X_1X_2X_3AX_4QX_5TX_6KNA$ (SEQ ID NO:151), 其中 X_1 如果存在,那么为Leu; X_2 如果存在,那么为Ala; X_3 为Lys、Leu或Pro; X_4 为Asn、His、Pro或Tyr; X_5 为Asn、Gly、Val或Asp;且 X_6 为Pro或Thr。式X的肽包括(但不限于)(96) LAKANQNTPKNA (SEQ ID NO:142); (57) LAHQNTTKNA (SEQ ID NO:103); (66) LAPQNTTKNA (SEQ ID NO:112); (69) LAYQDTTKNA (SEQ ID NO:115); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); 和 (42) PAPQDTTKNA (SEQ ID NO:88)。

[0149] 在一些情况下,肽插入物为LAHQDTTKKX (SEQ ID NO:143), 其中X为任何氨基酸。在一些情况下,肽插入物为LAHQDTTKKX (SEQ ID NO:143), 其中X为Ala、Thr、Asp、Val或Met。在一些情况下,肽插入物为(27) LAHQDTTKKA (SEQ ID NO:73)。在一些情况下,肽插入物为(52) LAHQDTTKKT (SEQ ID NO:98)。在一些情况下,肽插入物为LAHQDTTKKD (SEQ ID NO:144)。在一些情况下,肽插入物为LAHQDTTKKV (SEQ ID NO:145)。在一些情况下,肽插入物为LAHQDTTKKM (SEQ ID NO:146)。

[0150] 在一些情况下,肽插入物不为(88) LALGETTRPA (SEQ ID NO:134)。在一些情况下,肽插入物不为LGETTRP (SEQ ID NO:147)。

[0151] 合适的肽插入物包括(但不限于)(1) LAKDATKNA (SEQ ID NO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ ID NO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ ID NO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ ID NO:56); (11) PAPQDTTKKA (SEQ ID NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ ID NO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ ID NO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); (18) LAVSDSTKAA (SEQ ID NO:64); (19) LAAQGTAKKPA (SEQ ID NO:65); (20) LAPDQTTTRNA (SEQ ID NO:66); (21) LAASDSTKAA (SEQ ID NO:67); (22) LAPQDTTKNA (SEQ ID NO:68); (23) LAKADETRPA (SEQ ID NO:69); (24) LAHQDTAKNA (SEQ ID NO:70); (25) LAHQDTKKNA (SEQ ID NO:71); (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ ID NO:74); (29) LAHQDTTNA (SEQ ID NO:75); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); (32) LAISDQSKPA (SEQ ID NO:78); (33) LADATKTA (SEQ ID NO:79); (34) LAKDTTKNA (SEQ ID NO:80); (35) LAKSDQSRPA (SEQ ID NO:81); (36)

LAPQDTKKNA (SEQ ID NO:82); (37) LATSDSTKAA (SEQ ID NO:83); (38) LAVDGSQRSA (SEQ ID NO:84); (39) LPISDQTKHA (SEQ ID NO:85); (40) LPKDATKTIA (SEQ ID NO:86); (41) LPPQDTTKNA (SEQ ID NO:87); (42) PAPQDTTKNA (SEQ ID NO:88); (43) QAHQDTTKNA (SEQ ID NO:89); (44) LAHETSPRPA (SEQ ID NO:90); (45) LAKSTSTAPA (SEQ ID NO:91); (46) LADQDTTKNA (SEQ ID NO:92); (47) LAESDQSKPA (SEQ ID NO:93); (48) LAHKDTTKNA (SEQ ID NO:94); (49) LAHKTQQKM (SEQ ID NO:95); (50) LAHQDTTENA (SEQ ID NO:96); (51) LAHQDTTINA (SEQ ID NO:97); (52) LAHQDTTKKT (SEQ ID NO:98); (53) LAHQDTTKND (SEQ ID NO:99); (54) LAHQDTTKNT (SEQ ID NO:100); (55) LAHQDTTKNV (SEQ ID NO:101); (56) LAHQDTTKTM (SEQ ID NO:102); (57) LAHQNTTKNA (SEQ ID NO:103); (58) LAHRDTTKNA (SEQ ID NO:104); (59) LAISDQTNHA (SEQ ID NO:105); (60) LAKQKSASTA (SEQ ID NO:106); (61) LAKSDQCKPA (SEQ ID NO:107); (62) LAKSDQSKPD (SEQ ID NO:108); (63) LAKSDQSNPA (SEQ ID NO:109); (64) LAKSYQSKPA (SEQ ID NO:110); (65) LANQDTTKNA (SEQ ID NO:111); (66) LAPQNTTKNA (SEQ ID NO:112); (67) LAPSSIQKPA (SEQ ID NO:113); (68) LAQQDTTKNA (SEQ ID NO:114); (69) LAYQDTTKNA (SEQ ID NO:115); (70) LDHQDTTKNA (SEQ ID NO:116); (71) LDHQDTTKSA (SEQ ID NO:117); (72) LGHQDTTKNA (SEQ ID NO:118); (73) LPHQDTTKND (SEQ ID NO:119); (74) LPHQDTTKNT (SEQ ID NO:120); (75) LPHQDTTNA (SEQ ID NO:121); (76) LTHQDTTKNA (SEQ ID NO:122); (77) LTKDATKTIA (SEQ ID NO:123); (78) LTPQDTTKNA (SEQ ID NO:124); (79) LVHQDTTKNA (SEQ ID NO:125); (80) LAKANQNTPA (SEQ ID NO:126); (81) LATTPITKPA (SEQ ID NO:127); (82) LATTPIAKPA (SEQ ID NO:128); (83) LAIEDHTKSA (SEQ ID NO:129); (84) LAQSEHQ RPA (SEQ ID NO:130); (85) LAKSPNKDNA (SEQ ID NO:131); (86) LANQDYTKTA (SEQ ID NO:132); (87) LANSTDQTRA (SEQ ID NO:133); (88) LALGETTRPA (SEQ ID NO:134); (89) LANSTEQTRA (SEQ ID NO:135); (90) LAQADTTKNA (SEQ ID NO:136); (91) LASKDITKTA (SEQ ID NO:137); (92) LASPRHNKKC (SEQ ID NO:138); (93) LAHQDTTKTIA (SEQ ID NO:139); (94) LAAQGTANL (SEQ ID NO:140); (95) VAIEDHTKSA (SEQ ID NO:141); 和 (96) LAKANQNTPKNA (SEQ ID NO:142)。

[0182] 在一些情况下,肽插入物为(11) PAPQDTTKKA (SEQ ID NO:57)。在一些情况下,肽插入物为(7) LAPDSTTRSA (SEQ ID NO:53)。

[0183] 在一些实施方案中,除相对于对应的亲本AAV衣壳蛋白, GH环或环IV中约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入外，主题rAAV病毒体衣壳不包括任何其它的氨基酸取代、插入或缺失。在其它实施方案中，除相对于对应的亲本AAV衣壳蛋白，GH环或环IV中约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入外，与亲本AAV衣壳蛋白相比，主题rAAV病毒体衣壳包括1至约25个氨基酸插入、缺失或取代。举例来说，在一些实施方案中，除相对于对应的亲本AAV衣壳蛋白，GH环或环IV中的约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入外，与亲本AAV衣壳蛋白相比，主题rAAV病毒体衣壳包括1至约5个、约5至约10个、约10至约15个、约15至约20个或约20至约25个氨基酸插入、缺失或取代。

[0184] 在一些情况下，主题rAAV病毒体衣壳不包括以下氨基酸取代中的一个、两个、三个或四个：Y273F、Y444F、Y500F和Y730F。

[0185] 在一些情况下，除如上所述的插入肽外，主题变异衣壳多肽包含以下氨基酸取代中的一个、两个、三个或四个：Y273F、Y444F、Y500F和Y730F。

[0186] 在一些情况下，主题rAAV病毒体衣壳为嵌合衣壳，例如衣壳包含第一AAV血清型的AAV衣壳的一部分和第二血清型的AAV衣壳的一部分；且相对于对应的亲本AAV衣壳蛋白，在GH环或环IV中包含约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)。

[0187] 在一些实施方案中，主题rAAV病毒体包含如下衣壳蛋白，其氨基酸序列具有与图4中提供的氨基酸序列至少约85%、至少约90%、至少约95%、至少约98%或至少约99%氨基酸序列一致性；且相对于对应的亲本AAV衣壳蛋白，在GH环或环IV中包含约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入。在一些实施方案中，主题rAAV病毒体包含如下衣壳蛋白，其氨基酸序列具有与图4中提供的氨基酸序列至少约85%、至少约90%、至少约95%、至少约98%或至少约99%氨基酸序列一致性；且相对于图4中描绘的氨基酸序列，在氨基酸587与588之间，或在相对于对应的亲本AAV衣壳蛋白的对应位点，包含约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入。

[0188] 在一些实施方案中，主题rAAV病毒体包含包括如下GH环的衣壳蛋白，所述GH环包含与图5中阐述的氨基酸序列具有至少约85%、至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列，且在粗体与加下划线的氨基酸之间包含约5个氨基酸至约20个氨基酸(例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸)的插入。

[0189] 在一些实施方案中，主题rAAV病毒体包含如下衣壳蛋白，其氨基酸序列具有与图6A-6C中提供的任一氨基酸序列至少约85%、至少约90%、至少约95%、至少约98%或至少

约99%氨基酸序列一致性；且在AAV2的氨基酸587与588之间，或在相对于另一AAV基因型的对应位点，包含约5个氨基酸至约20个氨基酸（例如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个或20个氨基酸；例如9个氨基酸、10个氨基酸、11个氨基酸或12个氨基酸）的插入。在一些情况下，对应插入位点为如图6B中粗体文字和下划线所指示的位点。

[0190] 与包含对应的亲本AAV衣壳蛋白的AAV病毒体对视网膜细胞的感染性相比，主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对视网膜细胞的感染性。

[0191] 在一些情况下，与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对视网膜细胞的感染性相比，主题rAAV病毒体在经由玻璃体内注射进行施用时展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对视网膜细胞的感染性。

[0192] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体对感光细胞的感染性相比，主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对感光（视杆或视锥）细胞的感染性。

[0193] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对感光细胞的感染性相比，主题rAAV病毒体在经由玻璃体内注射进行施用，展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对感光（视杆或视锥）细胞的感染性。

[0194] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体对RGC的感染性相比，主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对RGC的感染性。

[0195] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对RGC的感染性相比，主题rAAV病毒体在经由玻璃体内注射进行施用，展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对RGC的感染性。

[0196] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体对RPE细胞的感染性相比，主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对RPE细胞的感染性。

[0197] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对RPE细胞的感染性相比，主题rAAV病毒体在经由玻璃体内注射进行施用，展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对RPE细胞的感染性。

[0198] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体对米勒细胞的感染性相比，主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对米勒细胞的感染性。

[0199] 在一些实施方案中，与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对米勒细胞的感染性相比，主题rAAV病毒体在经由玻璃体内注射进行施用，展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对

米勒细胞的感染性。

[0200] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体对双极细胞的感染性相比,主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对双极细胞的感染性。

[0201] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对双极细胞的感染性相比,主题rAAV病毒体在经由玻璃体内注射进行施用展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对双极细胞的感染性。

[0202] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体对无长突细胞的感染性相比,主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对无长突细胞的感染性。

[0203] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时无长突细胞的感染性相比,主题rAAV病毒体在经由玻璃体内注射进行施用展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对无长突细胞的感染性。

[0204] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体对水平细胞的感染性相比,主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对水平细胞的感染性。

[0205] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对水平细胞的感染性相比,主题rAAV病毒体在经由玻璃体内注射进行施用展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对水平细胞的感染性。

[0206] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体对视网膜星形细胞的感染性相比,主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对视网膜星形细胞的感染性。

[0207] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对视网膜星形细胞的感染性相比,主题rAAV病毒体在经由玻璃体内注射进行施用展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对视网膜星形细胞的感染性。

[0208] 在一些情况下,与包含对应的亲本AAV衣壳蛋白的AAV病毒体跨越内界膜(ILM)的能力相比,主题rAAV病毒体展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的跨越ILM的能力。

[0209] 在一些情况下,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用跨越ILM的能力相比,主题rAAV病毒体在经由玻璃体内注射进行施用展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的跨越ILM的能力。

[0210] 主题rAAV病毒体可以跨越ILM,且还可以横越细胞层,包括米勒细胞、无长突细胞等等,到达感光细胞和或RPE细胞。举例来说,主题rAAV病毒体在经由玻璃体内注射进行施用可以跨越ILM,且还可以横越细胞层,包括米勒细胞、无长突细胞等等,到达感光细胞和

或RPE细胞。

[0211] 在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体经过ILM定位的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的经过ILM的定位。举例来说,在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体定位至视网膜色素上皮细胞(RPE)层的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的至RPE的定位。作为另一实例,在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体定位至光感受器(PR)层的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的至PR层的定位。作为另一实例,在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体定位至内核层的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的至内核层的定位。作为另一实例,在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体定位至外核层的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的至外核层的定位。作为另一实例,在一些情况下,与经玻璃体内注射的包含对应的亲本AAV衣壳蛋白的对照AAV病毒体定位至神经节细胞层的程度相比,主题rAAV病毒体在经玻璃体内注射时,展现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的至神经节细胞层的定位。

[0212] 在一些实施方案中,主题rAAV病毒体选择性地感染视网膜细胞,例如主题rAAV病毒体以比非视网膜细胞(例如在眼睛外的细胞)大10倍、15倍、20倍、25倍、50倍或超过50倍的特异性感染视网膜细胞。举例来说,在一些实施方案中,主题rAAV病毒体选择性地感染视网膜细胞,例如主题rAAV病毒体以比非视网膜细胞(例如在眼睛外的细胞)大10倍、15倍、20倍、25倍、50倍或超过50倍的特异性感染感光细胞。

[0213] 在一些实施方案中,主题rAAV病毒体选择性地感染感光细胞,例如主题rAAV病毒体以比眼睛中存在的非感光细胞(例如视网膜神经节细胞、米勒细胞等)大10倍、15倍、20倍、25倍、50倍或超过50倍的特异性感染感光细胞。

[0214] 在一些实施方案中,与包含对应的亲本AAV衣壳蛋白的AAV病毒体在经由玻璃体内注射进行施用时对感光细胞的感染性相比,主题rAAV病毒体在经由玻璃体内注射进行施用时代现增加至少5倍、至少10倍、至少15倍、至少20倍、至少25倍、至少50倍或超过50倍的对感光细胞的感染性。

[0215] 基因产物

[0216] 主题rAAV病毒体包含异源核酸,所述异源核酸包含编码基因产物(异源基因产物)的核苷酸序列。在一些情况下,基因产物为多肽。在一些情况下,基因产物为RNA。在基因产物为RNA的情况下,在一些情况下,RNA基因产物编码多肽。在一些情况下,本公开的rAAV病毒体包含单个异源核酸,该单个异源核酸包含编码单个异源基因产物的核苷酸序列。在一些情况下,本公开的rAAV病毒体包含单个异源核酸,该单个异源核酸包含编码两个异源基因产物的核苷酸序列。在一些情况下,本公开的rAAV病毒体包含两个异源核酸,每个异源核

酸包含编码异源基因产物的核苷酸序列。

[0217] 在一些实施方案中,基因产物为干扰RNA。在一些实施方案中,基因产物为适体。在一些实施方案中,基因产物为多肽。在一些实施方案中,基因产物为提供基因功能的位点特异性敲低的位点特异性核酸酶。在一些实施方案中,基因产物为提供靶核酸的修饰的RNA引导的核酸内切酶。

[0218] 干扰RNA

[0219] 在基因产物为干扰RNA (RNAi) 的情况下,合适的RNAi包括降低细胞中细胞凋亡或血管生成因子的水平的RNAi。举例来说, RNAi可以为降低细胞中诱发或促进细胞凋亡的基因产物的水平的shRNA或siRNA。基因产物诱发或促进细胞凋亡的基因在本文中称为“促细胞凋亡基因”,且那些基因的产物(mRNA;蛋白质)称为“促细胞凋亡基因产物”。促细胞凋亡基因产物包括例如Bax、Bid、Bak和Bad基因产物。参见例如美国专利No.7,846,730。

[0220] 干扰RNA还可以针对血管生成产物,例如血管内皮生长因子(VEGF) (例如Cand5;参见例如美国专利公布No.2011/0143400;美国专利公布No.2008/0188437;和Reich等人(2003) Mol.Vis.9:210); VEGF受体-1 (VEGFR1) (例如Sirna-027;参见例如Kaiser等人(2010) Am.J.Ophthalmol.150:33;和Shen等人(2006) Gene Ther.13:225);或VEGF受体-2 (VEGFR2) (Kou等人(2005) Biochem.44:15064)。又参见美国专利No.6,649,596、6,399,586、5,661,135、5,639,872和5,639,736;和美国专利No.7,947,659和7,919,473。

[0221] 适体

[0222] 在基因产物为适体的情况下,所关注的例示性适体包括针对VEGF的适体。参见例如Ng等人(2006) Nat.Rev.Drug Discovery 5:123;和Lee等人(2005) Proc.Natl.Acad.Sci.USA 102:18902。举例来说,VEGF适体可以包含核苷酸序列5'-cgcaaucagugaaugcuuauacau ccg-3' (SEQ ID NO://)。血小板源性生长因子(PDGF)特异性适体,例如E10030也是适用的;参见例如Ni和Hui(2009) Ophthalmologica223:401;和Akiyama等人(2006) J.Cell Physiol.207:407)。

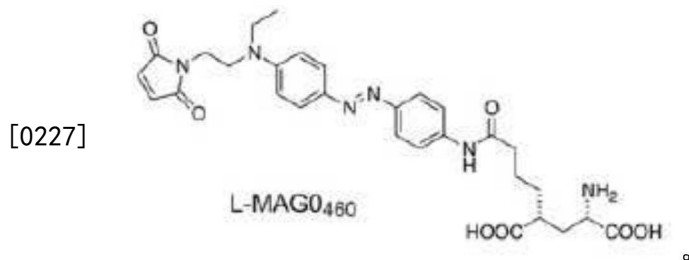
[0223] 多肽

[0224] 在基因产物为多肽的情况下,多肽一般为增强视网膜细胞的功能,例如视杆或视锥感光细胞、视网膜神经节细胞、米勒细胞、双极细胞、无长突细胞、水平细胞或视网膜色素上皮细胞的功能的多肽。例示性多肽包括神经保护多肽(例如胶质细胞源性神经营养因子(GDNF)、睫状神经营养因子(CNTF)、神经营养素-4(NT4)、神经生长因子(NGF)和神经秩蛋白(neurturin,NTN);抗血管生成多肽(例如可溶性VEGF受体;VEGF结合抗体;VEGF结合抗体片段(例如单链抗VEGF抗体);内皮抑素(endostatin);肿瘤抑素(tumstatin);血管抑素(angiostatin);可溶性Flt多肽(Lai等人(2005) Mol.Ther.12:659);包含可溶性Flt多肽的Fc融合蛋白(参见例如Pechan等人(2009) Gene Ther.16:10);色素上皮衍生因子(PEDF);可溶性Tie-2受体等);组织金属蛋白酶抑制因子-3(TIMP-3);光反应性视蛋白,例如视紫红质;抗细胞凋亡多肽(例如Bcl-2、Bcl-XL;XIAP)等等。合适的多肽包括(但不限于)神经胶质源性神经营养因子(GDNF);成纤维细胞生长因子;成纤维细胞生长因子2;神经秩蛋白(NTN);睫状神经营养因子(CNTF);神经生长因子(NGF);神经营养素-4(NT4);脑源性神经营养因子(BDNF;例如氨基酸序列具有与图7B中描绘的氨基酸序列(SEQ ID NO:11)的约200个氨基酸至247个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或

100%氨基酸序列一致性的多肽);表皮生长因子;视紫红质;X连锁凋亡抑制因子;和音猬因子。

[0225] 合适的光反应性视蛋白包括例如如以下中所述的光反应性视蛋白:美国专利公布No.2007/0261127(例如光敏感通道-2;ChR2;Chop2);美国专利公布No.2001/0086421;美国专利公布No.2010/0015095;美国专利公布No.2016/0002302;美国专利公布No.2013/0347137;美国专利公布No.2013/0019325;和Diester等人(2011)Nat.Neurosci.14:387。参见Thyagarajan等人(2010)J.Neurosci.30(26):8745-8758;Lagali等人(2008)Nat.Neurosci.11(6):667-675;Doroudchi等人(2011)Mol.Ther.19(7):1220-1229;Henriksen等人(2014)J.Ophthalmic Vis.Res.9:374;Tomita等人(2014)Mol.Ther.22:1434。

[0226] 合适的多肽包括光门控离子通道多肽。参见例如Gaub等人(2014)Proc.Natl.Acad.Sci.USA 111:E5574。举例来说,合适的多肽为光门控离子型谷氨酸受体(LiGluR)。在可光异构化化合物存在下LiGluR在视网膜神经节细胞和ON-双极细胞中的表达使细胞对光起反应。LiGluR包含L439C取代;参见Caporale等人(2011)Mol.Ther.19:1212-1219;Volgraf等人(2006)Nat.Chem Biol.2:47-52;和Gorostiza等人(2007)Proc.Natl.Acad.Sci.USA.104:10865-10870。可光异构化化合物包括例如在460nm下具有峰值效率的顺丁烯二酰亚胺-偶氮苯-谷氨酸酯0(MAGO₄₆₀)。MAGO₄₆₀具有以下结构:



[0228] 合适的多肽还包括视网膜劈裂蛋白(例如氨基酸序列具有与图7A中描绘的氨基酸序列(SEQ ID NO:10)的约200个氨基酸至224个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽)。合适的多肽包括例如视网膜色素变性GTP酶调控因子(RPGR)相互作用蛋白-1(参见例如GenBank登记号Q96KN7、Q9EPQ2和Q9GLM3)(例如氨基酸序列具有与图7F中描绘的氨基酸序列(SEQ ID NO:15)的约1150个氨基酸至1200个氨基酸或约1200个氨基酸至1286个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽);外周蛋白-2(Prph2)(参见例如GenBank登记号NP_000313(例如氨基酸序列具有与图7D中描绘的氨基酸序列(SEQ ID NO:13)的约300个氨基酸至346个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽);和Travis等人(1991)Genomics 10:733);外周蛋白(例如氨基酸序列具有与图7E中描绘的氨基酸序列(SEQ ID NO:14)的约400个氨基酸至约470个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽);视网膜色素上皮细胞特异性蛋白(RPE65)(例如氨基酸序列具有与图7C中描绘的氨基酸序列(SEQ ID NO:12)的约200个氨基酸至247个氨基酸的相邻延伸至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽)(参见例如GenBank AAC39660;和Morimura等人(1998)Proc.Natl.Acad.Sci.USA 95:3088);视杆源性视锥细胞活力因子(RdCVF)(例如氨基酸序

列具有与图7H、7I和7J中的任一个中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽；无脉络膜（例如氨基酸序列具有与图7G中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽）；视网膜色素变性GTP酶调控因子（RPGR）（例如氨基酸序列具有与图7S-7V中的一个中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的多肽）等等。举例来说，在一些情况下，合适的RPGR多肽包含具有与图7S中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例，在一些情况下，合适的RPGR多肽包含具有与图7T中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。举例来说，在一些情况下，合适的RPGR多肽包含具有与图7U中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。举例来说，在一些情况下，合适的RPGR多肽包含具有与图7V中描绘的氨基酸序列至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。

[0229] 合适的多肽还包括：CHM（无脉络膜（Rab护送蛋白1（REP1））），一种在有缺陷或缺失时引起无脉络膜的多肽（参见例如Donnelly等人（1994）*Hum. Mol. Genet.* 3:1017；和van Bokhoven等人（1994）*Hum. Mol. Genet.* 3:1041）；和Crumbs同源物1（CRB1），一种在有缺陷或缺失时引起利伯氏先天性黑朦和色素性视网膜炎的多肽（参见例如den Hollander等人（1999）*Nat. Genet.* 23:217；和GenBank登记号CAM23328）。举例来说，合适的REP1多肽可以包含具有与图7G中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。

[0230] 合适的多肽包括视杆cGMP特异性3', 5'-环状磷酸二酯酶子单元 α （PDE6 α ）、视杆cGMP特异性3', 5'-环状磷酸二酯酶子单元 β 同种型1（PDE6 β 同种型1）、视杆cGMP特异性3', 5'-环状磷酸二酯酶子单元 β 同种型2（PDE6 β 同种型2）、视杆cGMP特异性3', 5'-环状磷酸二酯酶子单元 β 同种型3（PDE6 β 同种型3）。举例来说，合适的PDE6 α 多肽可以包含具有与图7K中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例，合适的PDE6 β 同种型1多肽可以包含具有与图7L中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例，合适的PDE6 β 同种型2多肽可以包含具有与图7M中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例，合适的PDE6 β 同种型3多肽可以包含具有与图7N中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。

[0231] 合适的多肽还包括当有缺陷或缺失时引起全色盲的多肽，其中此类多肽包括例如视锥光感受器cGMP门控通道子单元 α （CNGA3）（参见例如GenBank登记号NP_001289；和Booi等人（2011）*Ophthalmology* 118:160-167）；视锥光感受器cGMP门控阳离子通道 β -子单元（CNGB3）（参见例如Kohl等人（2005）*Eur J Hum Genet.* 13(3):302）；鸟嘌呤核苷酸结合蛋白（G蛋白）、 α 转导活性多肽2（GNAT2）（ACHM4）；和ACHM5；和当有缺陷或缺乏时引起色盲的多种形式的多肽（例如L-视蛋白、M-视蛋白和S-视蛋白）。参见Mancuso等人（2009）*Nature* 461

(7265):784-787。

[0232] 举例来说,合适的CNGA3(又名ACHM2)同种型1多肽可以包含具有与图70中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例,合适的CNGA3(又名ACHM2)同种型2多肽可以包含具有与图7P中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。

[0233] 作为另一实例,合适的CNGB3(又名ACHM3)多肽可以包含具有与图7Q中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。作为另一实例,GNAT2(又名ACHM4)可以包含具有与图7R中描绘的氨基酸序列组至少约90%、至少约95%、至少约98%、至少约99%或100%氨基酸序列一致性的氨基酸序列。

[0234] 位点特异性核酸内切酶

[0235] 在一些情况下,所关注的基因产物为提供基因功能的位点特异性敲低的位点特异性核酸内切酶,例如其中核酸内切酶敲除与视网膜疾病有关的等位基因。举例来说,在显性等位基因编码当野生型时为视网膜结构蛋白和/或提供正常的视网膜功能的基因的缺陷拷贝的情况下,位点特异性核酸内切酶可以靶向缺陷等位基因且敲出缺陷等位基因。在一些情况下,位点特异性核酸内切酶为RNA引导的核酸内切酶。

[0236] 除敲除缺陷等位基因外,位点特异性核酸酶还可以用于刺激与编码由缺陷等位基因编码的蛋白质的功能性拷贝的供体DNA的同源重组。因此,例如,主题rAAV病毒体可以用于递送敲除缺陷等位基因的位点特异性核酸内切酶,且可以用于递送缺陷等位基因的功能性拷贝,引起缺陷等位基因的修复,从而产生功能性视网膜蛋白(例如功能性视网膜劈裂蛋白、功能性RPE65、功能性外周蛋白等等)。参见例如Li等人(2011)Nature 475:217。在一些实施方案中,主题rAAV病毒体包含编码位点特异性核酸内切酶的异源核苷酸序列;和编码缺陷等位基因的功能性拷贝的异源核苷酸序列,其中功能性拷贝编码功能性视网膜蛋白。功能性视网膜蛋白包括例如视网膜劈裂蛋白、RPE65、视网膜色素变性GTP酶调控因子(RGPR)相互作用蛋白-1、外周蛋白、外周蛋白-2、RdCVF等等。

[0237] 适用的位点特异性核酸内切酶包括例如锌指核酸酶(ZFN);大范围核酸酶;和转录激活因子样效应物核酸酶(TALEN),其中此类位点特异性核酸内切酶并非天然存在的且经修饰以靶向特定基因。此类位点特异性核酸酶可以进行工程化以切割基因组内的特定位置,且接着非同源性末端接合可以在插入或缺失若干核苷酸时修复断裂。接着此类位点特异性核酸内切酶(又称为“INDEL”)将蛋白质抛出框,且有效地敲除基因。参见例如美国专利公布No.2011/0301073。合适的位点特异性核酸内切酶包括工程化的大范围核苷酸、再次工程化的归巢核酸内切酶。合适的核酸内切酶包括I-Tev1核酸酶。合适的大范围核酸酶包括I-Sce1(参见例如Bellaiche等人(1999)Genetics 152:1037);和I-Cre1(参见例如Heath等人(1997)Nature Structural Biology 4:468)。

[0238] RNA引导的核酸内切酶

[0239] 在一些情况下,基因产物为RNA引导的核酸内切酶。在一些情况下,基因产物为包含编码RNA引导的核酸内切酶的核苷酸序列的RNA。在一些情况下,基因产物为向导RNA,例如单向导RNA。在一些情况下,基因产物为:1)向导RNA;和2)RNA引导的核酸内切酶。向导RNA

可以包含:a)结合于RNA引导的核酸内切酶的蛋白结合区;和b)结合于靶核酸的区域。RNA引导的核酸内切酶在本文中又称为“基因组编辑核酸酶”。

[0240] 合适的基因组编辑核酸酶的实例为CRISPR/Cas核酸内切酶(例如2类CRISPR/Cas核酸内切酶,例如II型、V型或VI型CRISPR/Cas核酸内切酶)。因此,基因组靶向组合物可以包括CRISPR/Cas核酸内切酶(例如2类CRISPR/Cas核酸内切酶,例如II型、V型或VI型CRISPR/Cas核酸内切酶)。在一些情况下,基因组靶向组合物包括2类CRISPR/Cas核酸内切酶。在一些情况下,基因组靶向组合物包括2类II型CRISPR/Cas核酸内切酶(例如Cas9蛋白)。在一些情况下,基因组靶向组合物包括2类V型CRISPR/Cas核酸内切酶(例如Cpf1蛋白、C2c1蛋白或C2c3蛋白)。在一些情况下,基因组靶向组合物包括2类VI型CRISPR/Cas核酸内切酶(例如C2c2蛋白)。

[0241] 在一些情况下,基因组编辑核酸酶为与异源多肽(又称为“融合搭配物”)融合的融合蛋白。在一些情况下,基因组编辑核酸酶与提供亚细胞定位的氨基酸序列(融合搭配物)融合,即融合搭配物为亚细胞定位序列(例如一种或多种用于靶向核的核定位信号(NLS)、两种或更多种NLS、三种或更多种NLS等)。

[0242] 在一些情况下,基因组编辑核酸内切酶为II型CRISPR/Cas核酸内切酶。在一些情况下,基因组编辑核酸内切酶为Cas9多肽。Cas9蛋白由于与Cas9向导RNA的蛋白结合区段缔合而被引导至靶核酸序列(例如染色体序列或染色体外序列,例如游离体序列、微环序列、线粒体序列、叶绿体序列等)内的靶位点(例如稳定在靶位点)。在一些情况下,Cas9多肽包含具有与图8A中描绘的酿脓链球菌(*Streptococcus pyogenes*)Cas9至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少98%、至少99%或超过99%氨基酸序列一致性的氨基酸序列。在一些情况下,用于本公开的组合物或方法中的Cas9多肽为金黄色葡萄球菌(*Staphylococcus aureus*)Cas9(saCas9)多肽。在一些情况下,saCas9多肽包含具有与图8B中描绘的saCas9氨基酸序列至少85%、至少90%、至少95%、至少98%、至少99%或100%氨基酸序列一致性的氨基酸序列。

[0243] 在一些情况下,合适的Cas9多肽为高保真度(HF)Cas9多肽。Kleinstiver等人(2016)Nature 529:490。举例来说,图8A中描绘的氨基酸序列的氨基酸N497、R661、Q695和Q926经例如丙氨酸取代。举例来说,HF Cas9多肽可以包含具有与图8A中描绘的氨基酸序列至少90%、至少95%、至少98%、至少99%或100%氨基酸序列一致性的氨基酸序列,其中氨基酸N497、R661、Q695和Q926经例如丙氨酸取代。

[0244] 在一些情况下,合适的Cas9多肽展现改变的PAM特异性。参见例如Kleinstiver等人(2015)Nature 523:481。

[0245] 在一些情况下,基因组编辑核酸内切酶为V型CRISPR/Cas核酸内切酶。在一些情况下,V型CRISPR/Cas核酸内切酶为Cpf1蛋白。在一些情况下,Cpf1蛋白包含具有与图8C中描绘的Cpf1氨基酸序列至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少90%或100%氨基酸序列一致性的氨基酸序列。

[0246] 结合于2类CRISPR/Cas核酸内切酶(例如Cas9蛋白;V型或VI型CRISPR/Cas蛋白;Cpf1蛋白等)且将复合物靶向靶核酸内的特定位置的核酸在本文中称为“向导RNA”或“CRISPR/Cas向导核酸”或“CRISPR/Cas向导RNA”。向导RNA通过包括靶向区段提供对复合物

(RNP复合物)的特异性,所述靶向区段包括向导序列(在本文中又称为靶向序列),向导序列为与靶核酸的序列互补的核苷酸序列。

[0247] 在一些情况下,向导RNA包括两个分开的核酸分子:“活化子”和“靶向子”,且在本文中称为“双向导RNA”、“双分子向导RNA”、“两分子向导RNA”或“dgRNA”。在一些情况下,向导RNA为一个分子(例如对于一些2类CRISPR/Cas蛋白,对应的向导RNA为单一分子;且在一些情况下,活化子和靶向子彼此例如经由插入核苷酸而共价连接),且向导RNA称为“单向导RNA”、“单分子向导RNA”、“一分子向导RNA”或简称为“sgRNA”。

[0248] 在基因产物为RNA引导的核酸内切酶或为RNA引导的核酸内切酶与向导RNA的情况下,基因产物可以修饰靶核酸。在一些情况下,例如在靶核酸包含缺陷等位基因中的有害突变(例如视网膜细胞靶核酸中的有害突变)的情况下,RNA引导的核酸内切酶/向导RNA复合物与包含校正有害突变的核苷酸序列的供体核酸(例如包含编码由缺陷等位基因编码的蛋白质的功能性拷贝的核苷酸序列的供体核酸)一起可以用于例如经由同源性定向修复(HDR)校正有害突变。

[0249] 在一些情况下,基因产物为RNA引导的核酸内切酶和2个分开的sgRNA,其中2个分开的sgRNA经由非同源性末端接合(NHEJ)提供靶核酸的缺失。

[0250] 本公开提供了一种修饰个体中的视网膜细胞中的靶核酸的方法,其中靶核酸包含有害突变,所述方法包括向所述个体施用(例如通过眼内、玻璃体内等施用)本公开的rAAV病毒体,其中所述rAAV病毒体包含含有以下的异源核酸:i)编码RNA引导的核酸内切酶(例如Cas9核酸内切酶)的核苷酸序列;ii)编码包含与靶核酸互补的核苷酸序列的sgRNA的核苷酸序列;和iii)编码包含校正有害突变的核苷酸序列的供体DNA模板的核苷酸序列。rAAV病毒体的施用通过HDR校正靶核酸中的有害突变。

[0251] 本公开提供了一种修饰个体中的视网膜细胞中的靶核酸的方法,其中靶核酸包含有害突变,所述方法包括向所述个体施用(例如通过眼内、玻璃体内等施用)本公开的rAAV病毒体,其中所述rAAV病毒体包含含有以下的异源核酸:i)编码RNA引导的核酸内切酶(例如Cas9核酸内切酶)的核苷酸序列;ii)编码包含与靶核酸中的第一序列互补的核苷酸序列的第一sgRNA的核苷酸序列;和iii)编码包含与靶核酸中的第二序列互补的核苷酸序列的第二sgRNA的核苷酸序列。rAAV病毒体的施用通过NHEJ切除靶核酸中的有害突变。

[0252] 调控序列

[0253] 在一些情况下,编码所关注的基因产物的核苷酸序列可操作地连接于转录控制元件。举例来说,在一些情况下,编码所关注的基因产物的核苷酸序列可操作地连接于组成型启动子。在其它情况下,编码所关注的基因产物的核苷酸序列可操作地连接于诱导型启动子。在一些情况下,编码所关注的基因产物的核苷酸序列可操作地连接于组织特异性或细胞类型特异性调控元件。举例来说,在一些情况下,编码所关注的基因产物的核苷酸序列可操作地连接于视网膜细胞特异性启动子。举例来说,在一些情况下,编码所关注的基因产物的核苷酸序列可操作地连接于光感受器特异性调控元件(例如光感受器特异性启动子),例如赋予可操作地连接的基因在感光细胞中选择性表达的调控元件。合适的光感受器特异性调控元件包括例如视紫红质启动子;视紫红质激酶启动子(Young等人(2003) *Ophthalmol.Vis.Sci.*44:4076); β 磷酸二酯酶基因启动子(Nicoud等人(2007) *J. Gene Med.*9:1015);视网膜色素变性基因启动子(Nicoud等人(2007) 上述);光感受器间类维生素

A结合蛋白 (IRBP) 基因增强子 (Nicoud等人 (2007) 上述); IRBP基因启动子 (Yokoyama等人 (1992) Exp Eye Res.55:225)。

[0254] 药物组合物

[0255] 本公开提供了一种药物组合物,其包含:a)如上所述的主体rAAV病毒体;和b)药学上可接受的载剂、稀释剂、赋形剂或缓冲剂。在一些实施方案中,药学上可接受的载剂、稀释剂、赋形剂或缓冲剂适合用于人类。

[0256] 此类赋形剂、载剂、稀释剂和缓冲剂包括可以施用而无过度毒性的任何制剂。药学上可接受的赋形剂包括(但不限于)例如水、盐水、丙三醇和乙醇等液体。其中可以包括药学上可接受的盐,例如无机酸盐,例如盐酸盐、氢溴酸盐、磷酸盐、硫酸盐等等;和有机酸盐,例如乙酸盐、丙酸盐、丙二酸盐、苯甲酸盐等等。另外,例如湿润剂或乳化剂、pH缓冲物质等等辅助物质可以存在于此类媒剂中。本领域中已知多种药学上可接受的赋形剂且本文中无需详细地论述。药学上可接受的赋形剂已经在多个公布中充分描述,所述公布包括例如 A.Gennaro (2000) “Remington: The Science and Practice of Pharmacy,” 第20版, Lippincott, Williams, & Wilkins; Pharmaceutical Dosage Forms and Drug Delivery Systems (1999) H.C. Ansel等人编辑, 第7版, Lippincott, Williams, & Wilkins; 和 Handbook of Pharmaceutical Excipients (2000) A.H. Kibbe等人编辑, 第3版 Amer. Pharmaceutical Assoc.

[0257] 将基因产物递送至视网膜细胞的方法和治疗方法

[0258] 本公开提供了一种将基因产物递送至个体中的视网膜细胞的方法,所述方法包括向所述个体施用如上所述的主体rAAV病毒体。如上所述,基因产物可以是多肽或干扰RNA (例如shRNA、siRNA等等)、适体或位点特异性核酸内切酶 (例如RNA引导的核酸内切酶)。基因产物递送至视网膜细胞可以治疗视网膜疾病。视网膜细胞可以是光感受器、视网膜神经节细胞、米勒细胞、双极细胞、无长突细胞、水平细胞或视网膜色素上皮细胞。在一些情况下,视网膜细胞为感光细胞,例如视杆或视锥细胞。

[0259] 本公开提供了一种修饰视网膜细胞中的靶核酸的方法,所述方法包括使视网膜细胞接触:1)本公开的rAAV病毒体,其中所述rAAV病毒体包含异源核酸,所述异源核酸包含编码结合向导RNA的RNA引导的核酸内切酶的核苷酸序列;和2)向导RNA。本公开提供了一种修饰视网膜细胞中的靶核酸的方法,所述方法包括使视网膜细胞接触本公开的rAAV病毒体,其中所述rAAV病毒体包含异源核酸,所述异源核酸包含编码以下的核苷酸序列:i)结合向导RNA的RNA引导的核酸内切酶;和ii)向导RNA。在一些情况下,所述方法包括使视网膜细胞与供体DNA模板接触。在一些情况下,RNA引导的核酸内切酶为Cas9多肽。在一些情况下,向导RNA为单向导RNA。

[0260] 本公开提供了一种治疗眼病 (例如视网膜疾病) 的方法,所述方法包括向有需要的个体施用有效量的如上所述的主体rAAV病毒体。主体rAAV病毒体可以经由眼球内注射、通过玻璃体内注射或通过任何其它合宜的施用模式或途径施用。其它合宜的施用模式或途径包括例如静脉内、鼻内等等。

[0261] “治疗有效量”将在可以通过实验和/或临床试验测定的相对广泛的范围内。举例来说,对于体内注射,即直接注射至眼中,治疗有效剂量将大约为约 10^6 至约 10^{15} rAAV病毒体,例如约 10^8 至 10^{12} rAAV病毒体。对于体外转导,待递送至细胞的rAAV病毒体的有效量将大

约为约 10^8 至约 10^{13} rAAV病毒体。其它有效剂量容易由本领域的一般技术人员通过常规试验,确定剂量反应曲线来确定。

[0262] 在一些实施方案中,超过一次施用(例如两次、三次、四次或更多次施用)可以用以实现所需的基因表达水平。在一些情况下,超过一次施用以多个时间间隔施用,例如每日、每周、半月、每月、每3个月、每6个月、每年施用等等。在一些情况下,多次施用在1个月至2个月、2个月至4个月、4个月至8个月、8个月至12个月、1年至2年、2年至5年或超过5年的时期内施用。

[0263] 可以使用主题方法治疗的眼病包括(但不限于)急性黄斑神经视网膜病变;白塞氏病(Behcet's disease);脉络膜新生血管;糖尿病性葡萄膜炎;组织胞浆菌病;黄斑变性,例如急性黄斑变性、非渗出性年龄相关性黄斑变性和渗出性年龄相关性黄斑变性;水肿,例如黄斑水肿、黄斑囊样水肿和糖尿病性黄斑水肿;多灶性脉络膜炎;影响后眼位点或位置的眼外伤;眼肿瘤;视网膜病症,例如视网膜中央静脉阻塞、糖尿病性视网膜病(包括增生性糖尿病性视网膜病)、增生性玻璃体视网膜病变(PVR)、视网膜动脉阻塞性疾病、视网膜剥离、葡萄膜炎视网膜病;交感性眼炎;伏格特-小柳-原田(Vogt Koyanagi-Harada, VKH)综合征;葡萄膜扩散;由眼睛激光治疗引起或受其影响的后眼病状;由光动力学疗法引起或受其影响的后眼病状;光凝、辐射性视网膜病变;视网膜外膜病症;视网膜分支静脉阻塞;前部缺血性视神经病变;非视网膜病变的糖尿病性视网膜功能障碍;视网膜劈裂;视网膜色素变性;青光眼;尤塞氏综合征(Usher syndrome)、视锥-视杆营养不良;斯特格氏病(Stargardt disease)(眼底黄色斑点症);遗传性黄斑变性;脉络膜视网膜变性;利伯氏先天性黑朦;先天性静止性夜盲;无脉络膜;巴德-毕德氏综合征(Bardet-Biedl syndrome);黄斑毛细血管扩张;利伯氏遗传性视神经病变;早产儿视网膜病;色觉障碍,包括全色盲、红色盲、绿色盲和蓝色盲;和比蒂氏结晶样视网膜营养不良(Bietti's crystalline dystrophy)。

[0264] 核酸和宿主细胞

[0265] 本公开提供了一种分离的核酸,其包含编码如上所述的主题变异腺相关病毒(AAV)衣壳蛋白的核苷酸序列,其中相对于对应的亲本AAV衣壳蛋白,所述变异AAV衣壳蛋白在GH环或环IV中包含约5个氨基酸至约20个氨基酸的插入,且其中与包含对应的亲本AAV衣壳蛋白的AAV病毒体对视网膜细胞的感染性相比,所述变异衣壳蛋白在存在于AAV病毒体中对视网膜细胞的感染性增加。主题分离的核酸可以为AAV载体,例如重组AAV载体。

[0266] 插入肽

[0267] 由主题核酸编码的变异AAV衣壳蛋白具有约5个氨基酸至约20个氨基酸长的插入肽,所述插入肽插入AAV衣壳的GH环中。插入肽具有5个氨基酸、6个氨基酸、7个氨基酸、8个氨基酸、9个氨基酸、10个氨基酸、11个氨基酸、12个氨基酸、13个氨基酸、14个氨基酸、15个氨基酸、16个氨基酸、17个氨基酸、18个氨基酸、19个氨基酸或20个氨基酸的长度。合适的插入肽如上所述。合适的插入肽包括如上所述的式I-X中的任一个的肽。

[0268] 主题重组AAV载体可以用于产生如上所述的主题重组AAV病毒体。因此,本公开提供一种重组AAV载体,其在引入合适的细胞中时,可以产生主题重组AAV病毒体。

[0269] 本发明还提供了宿主细胞,例如分离(经遗传修饰)的宿主细胞,其包含主题核酸。主题宿主细胞可以为分离的细胞,例如体外培养的细胞。主题宿主细胞可如下所述用于产生主题rAAV病毒体。在主题宿主细胞用于产生主题rAAV病毒体的情况下,其称为“包装细

胞”。在一些实施方案中,主题宿主细胞稳定地经主题核酸遗传修饰。在其它实施方案中,主题宿主细胞短暂地经主题核酸遗传修饰。

[0270] 使用建立的技术,包括(但不限于)电穿孔、磷酸钙沉淀、脂质粒介导的转染等等将主题核酸稳定或短暂地引入宿主细胞中。为稳定转化,主题核酸一般还包括可选择的标志物,例如若干众所周知的可选择的标志物中的任一种,例如新霉素抗性等等。

[0271] 主题宿主细胞通过将主题核酸引入多种细胞中的任一种,例如哺乳动物细胞,包括例如鼠科细胞和灵长类动物细胞(例如人类细胞)中而产生。合适的哺乳动物细胞包括(但不限于)初级细胞和细胞系,其中合适的细胞系包括(但不限于)293细胞、COS细胞、HeLa细胞、Vero细胞、3T3小鼠成纤维细胞、C3H10T1/2成纤维细胞、CHO细胞等等。合适的宿主细胞的非限制性实例包括例如HeLa细胞(例如美国菌种保藏中心(American Type Culture Collection, ATCC) No. CCL-2)、CHO细胞(例如ATCC No. CRL9618、CCL61、CRL9096)、293细胞(例如ATCC No. CRL-1573)、Vero细胞、NIH 3T3细胞(例如ATCC No. CRL-1658)、Huh-7细胞、BHK细胞(例如ATCC No. CCL10)、PC12细胞(ATCC No. CRL1721)、COS细胞、COS-7细胞(ATCC No. CRL1651)、RAT1细胞、小鼠L细胞(ATCC No. CCL1.3)、人类胚肾(HEK)细胞(ATCC No. CRL1573)、HLHepG2细胞等等。主题宿主细胞还可以使用杆状病毒感染昆虫细胞如Sf9细胞,产生AAV来制成(参见例如美国专利No. 7,271,002;美国专利申请12/297,958)。

[0272] 在一些实施方案中,除包含编码如上所述的变异AAV衣壳蛋白的核苷酸序列的核酸外,主题经遗传修饰的宿主细胞还包括包含编码一种或多种AAV rep蛋白的核苷酸序列的核酸。在其它实施方案中,主题宿主细胞还包含rAAV载体。rAAV病毒体可以使用主题宿主细胞产生。产生rAAV病毒体的方法描述于例如美国专利公布No. 2005/0053922和美国专利公布No. 2009/0202490中。

[0273] 本公开的非限制性方面的实例

[0274] 包括实施方案在内的上述本发明的主题的方面在单独或与一个或多个其它方面或实施方案组合下可以为有益的。不限制以上描述,下文提供本公开编号1-34的某些非限制性方面。如本领域的技术人员在阅读本公开后将显而易见,可以使用每个个别编号的方面或与前面或后面个别编号的方面中的任一个组合。此意图为方面的所有此类组合提供支持且不局限于以下明确提供的方面的组合:

[0275] 方面1. 一种重组腺相关病毒(rAAV)病毒体,其包含:a) 变异AAV衣壳蛋白,其中相对于对应的亲本AAV衣壳蛋白,所述变异AAV衣壳蛋白在衣壳蛋白GH环中包含具有5个氨基酸至约20个氨基酸长度的异源肽的插入,且其中与包含对应的亲本AAV衣壳蛋白的对照AAV病毒体对视网膜细胞的感染性相比,所述变异衣壳蛋白对视网膜细胞的感染性增加;和b) 包含编码异源基因产物的核苷酸序列的异源核酸。

[0276] 方面2. 方面1的rAAV病毒体,其中与包含对应的亲本AAV衣壳蛋白的对照AAV病毒体对视网膜细胞的感染性相比,所述rAAV病毒体展现增加5倍的对视网膜细胞的感染性。

[0277] 方面3. 方面1的rAAV病毒体,其中与包含对应的亲本AAV衣壳蛋白的AAV病毒体对视网膜细胞的感染性相比,所述rAAV病毒体展现增加至少10倍的对视网膜细胞的感染性。

[0278] 方面4. 方面1至3中任一项的rAAV病毒体,其中与包含对应的亲本AAV衣壳蛋白的AAV病毒体定位至内核层、外核层、光感受器层、神经节细胞层或视网膜色素上皮细胞的程度相比,所述rAAV病毒体展现增加至少5倍的至内核层、外核层、光感受器层、神经节细胞层

和视网膜色素上皮细胞中的一个或多个的定位。

[0279] 方面5. 方面1至4中任一项的rAAV病毒体, 其中插入位点介于与AAV2的VP1的氨基酸570和611对应的氨基酸之间, 或另一AAV血清型的衣壳蛋白中的对应位置。

[0280] 方面6. 方面1至5中任一项的rAAV病毒体, 其中插入位点位于与AAV2的VP1的氨基酸587和588对应的氨基酸之间, 或另一AAV血清型的衣壳蛋白中的对应位置。

[0281] 方面7. 方面1至6中任一项的rAAV病毒体, 其中基因产物为干扰RNA或适体。

[0282] 方面8. 方面1至6中任一项的rAAV病毒体, 其中所述基因产物为多肽。

[0283] 方面9. 方面8的rAAV病毒体, 其中所述多肽为神经保护多肽、抗血管生成多肽或增强视网膜细胞的功能的多肽。

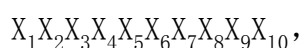
[0284] 方面10. 方面8的rAAV病毒体, 其中所述多肽为RNA引导的核酸内切酶。

[0285] 方面11. 方面10的rAAV病毒体, 其中所述RNA引导的核酸内切酶为Cas9多肽。

[0286] 方面12. 方面10的rAAV病毒体, 其中所述基因产物为RNA引导的核酸内切酶和向导RNA。

[0287] 方面13. 方面1至12中任一项的rAAV病毒体, 其中插入GH环中的所述异源肽具有式I-X中的任一个。

[0288] 方面14. 方面1至12中任一项的rAAV病毒体, 其中所述异源肽为式I的肽:



[0289] 其中:

[0290] X_1 为Leu、Ile、Pro或Gln;

[0291] X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val;

[0292] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr;

[0293] X_4 如果存在, 那么为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg;

[0294] X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr;

[0295] X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile;

[0296] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys;

[0297] X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn;

[0298] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且

[0299] X_{10} 为Ala、Phe、Asp、Thr、Val或Met。

[0300] 方面15. 方面14的rAAV病毒体, 其中所述异源肽包含以下氨基酸序列中的一种:

(1) LAKDATKNA (SEQ ID NO: 47); (2) PAHQDTTKNA (SEQ ID NO: 48); (3) LAHQDTTKNA (SEQ ID NO: 49); (4) LATTSQNKPA (SEQ ID NO: 50); (5) LAISDQTKHA (SEQ ID NO: 51); (6) IARGVAPSSA (SEQ ID NO: 52); (7) LAPDSTTRSA (SEQ ID NO: 53); (8) LAKGTELKPA (SEQ ID NO: 54); (9) LAIIDATKNA (SEQ ID NO: 55); (10) LAVDGAQRSA (SEQ ID NO: 56); (11) PAPQDTTKKA (SEQ ID NO: 57); (12) LPHQDTTKNA (SEQ ID NO: 58); (13) LAKDATKTIA (SEQ ID NO: 59); (14) LAKQQSASTA (SEQ ID NO: 60); (15) LAKSDQSKPA (SEQ ID NO: 61); (16) LSHQDTTKNA (SEQ ID NO: 62); (17) LAANQPSKPA (SEQ ID NO: 63); (18) LAVSDSTKAA (SEQ ID NO: 64); (19) LAAQGTAKKPA (SEQ ID NO: 65); (20) LAPDQTTRNA (SEQ ID NO: 66); (21) LAASDSTKAA (SEQ ID NO: 67); (22) LAPQDTTKNA (SEQ ID NO: 68); (23) LAKADETRPA (SEQ ID NO: 69); (24) LAHQDTAKNA (SEQ ID NO: 70); (25) LAHQDTKKNA (SEQ ID NO: 71); (26) LAHQDTTKHA (SEQ ID

NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ ID NO:74); (29) LAHQDTTNA (SEQ ID NO:75); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); (32) LAISDQSKPA (SEQ ID NO:78); (33) LADATKTA (SEQ ID NO:79); (34) LAKDTTKNA (SEQ ID NO:80); (35) LAKSDQSRPA (SEQ ID NO:81); (36) LAPQDTKKNA (SEQ ID NO:82); (37) LATSDSTKAA (SEQ ID NO:83); (38) LAVDGSQRSA (SEQ ID NO:84); (39) LPISDQTKHA (SEQ ID NO:85); (40) LPKDATKTIA (SEQ ID NO:86); (41) LPPQDTTKNA (SEQ ID NO:87); (42) PAPQDTTKNA (SEQ ID NO:88); (43) QAHQDTTKNA (SEQ ID NO:89); (44) LAHETSPRPA (SEQ ID NO:90); (45) LAKSTSTAPA (SEQ ID NO:91); (46) LADQDTTKNA (SEQ ID NO:92); (47) LAESDQSKPA (SEQ ID NO:93); (48) LAHKDTTKNA (SEQ ID NO:94); (49) LAHKTQQKM (SEQ ID NO:95); (50) LAHQDTTENA (SEQ ID NO:96); (51) LAHQDTTINA (SEQ ID NO:97); (52) LAHQDTTKKT (SEQ ID NO:98); (53) LAHQDTTKND (SEQ ID NO:99); (54) LAHQDTTKNT (SEQ ID NO:100); (55) LAHQDTTKNV (SEQ ID NO:101); (56) LAHQDTTKTM (SEQ ID NO:102); (57) LAHQNTTKNA (SEQ ID NO:103); (58) LAHRDTTKNA (SEQ ID NO:104); (59) LAISDQTNHA (SEQ ID NO:105); (60) LAKQKSASTA (SEQ ID NO:106); (61) LAKSDQCKPA (SEQ ID NO:107); (62) LAKSDQSKPD (SEQ ID NO:108); (63) LAKSDQSNPA (SEQ ID NO:109); (64) LAKSYQSKPA (SEQ ID NO:110); (65) LANQDTTKNA (SEQ ID NO:111); (66) LAPQNTTKNA (SEQ ID NO:112); (67) LAPSSIQKPA (SEQ ID NO:113); (68) LAQQDTTKNA (SEQ ID NO:114); (69) LAYQDTTKNA (SEQ ID NO:115); (70) LDHQDTTKNA (SEQ ID NO:116); (71) LDHQDTTKSA (SEQ ID NO:117); (72) LGHQDTTKNA (SEQ ID NO:118); (73) LPHQDTTKND (SEQ ID NO:119); (74) LPHQDTTKNT (SEQ ID NO:120); (75) LPHQDTTNNA (SEQ ID NO:121); (76) LTHQDTTKNA (SEQ ID NO:122); (77) LTKDATKTIA (SEQ ID NO:123); (78) LTPQDTTKNA (SEQ ID NO:124); 和 (79) LVHQDTTKNA (SEQ ID NO:125)。

[0301] 方面16. 方面1至12中任一项的rAAV病毒体, 其中所述异源肽为式II的肽:

[0302] $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0303] X_1 为Leu、Ile或Pro;

[0304] X_2 为Ala、Pro或Ser;

[0305] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg或Ala;

[0306] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile或Asn;

[0307] X_5 为Asp、Ser、Gln、Val、Thr、Gly或Ala;

[0308] X_6 为Thr、Ala、Gln、Ser、Glu或Pro;

[0309] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys或Ala;

[0310] X_8 为Lys、Ser、Arg或Thr;

[0311] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且

[0312] X_{10} 为Ala。

[0313] 方面17. 方面16的rAAV病毒体, 其中所述肽包含以下氨基酸序列中的一种: (1) LAKDATKNA (SEQ ID NO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ ID NO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ ID NO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ ID NO:56); (11) PAPQDTTKKA (SEQ ID

NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ ID NO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ ID NO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); (18) LAVSDSTKAA (SEQ ID NO:64); (19) LAAQGTAKKPA (SEQ ID NO:65); (20) LAPDQTTRNA (SEQ ID NO:66); (21) LAASDSTKAA (SEQ ID NO:67); (22) LAPQDTTKNA (SEQ ID NO:68); (23) LAKADETRPA (SEQ ID NO:69); (24) LAHQDTAKNA (SEQ ID NO:70); (25) LAHQDTKKNA (SEQ ID NO:71); (26) LAHQDTTKHA (SEQ ID NO:72); (27) LAHQDTTKKA (SEQ ID NO:73); (28) LAHQDTTRNA (SEQ ID NO:74); (29) LAHQDTTNA (SEQ ID NO:75); (30) LAHQGTTKNA (SEQ ID NO:76); (31) LAHQVTTKNA (SEQ ID NO:77); (32) LAISDQSKPA (SEQ ID NO:78); (33) LADATKTA (SEQ ID NO:79); (34) LAKDTTKNA (SEQ ID NO:80); (35) LAKSDQSRPA (SEQ ID NO:81); (36) LAPQDTKKNA (SEQ ID NO:82); (37) LATSDSTKAA (SEQ ID NO:83); (38) LAVDGSQRSA (SEQ ID NO:84); (39) LPISDQTKHA (SEQ ID NO:85); (40) LPKDATKTIA (SEQ ID NO:86); (41) LPPQDTTKNA (SEQ ID NO:87); 和 (42) PAPQDTTKNA (SEQ ID NO:88)。

[0314] 方面18. 方面16的rAAV病毒体, 其中所述肽包含以下氨基酸序列中的一种: (1) LAKDATKNA (SEQ ID NO:47); (2) PAHQDTTKNA (SEQ ID NO:48); (3) LAHQDTTKNA (SEQ ID NO:49); (4) LATTSQNKPA (SEQ ID NO:50); (5) LAISDQTKHA (SEQ ID NO:51); (6) IARGVAPSSA (SEQ ID NO:52); (7) LAPDSTTRSA (SEQ ID NO:53); (8) LAKGTELKPA (SEQ ID NO:54); (9) LAIIDATKNA (SEQ ID NO:55); (10) LAVDGAQRSA (SEQ ID NO:56); (11) PAPQDTTKKA (SEQ ID NO:57); (12) LPHQDTTKNA (SEQ ID NO:58); (13) LAKDATKTIA (SEQ ID NO:59); (14) LAKQQSASTA (SEQ ID NO:60); (15) LAKSDQSKPA (SEQ ID NO:61); (16) LSHQDTTKNA (SEQ ID NO:62); (17) LAANQPSKPA (SEQ ID NO:63); 和 (18) LAVSDSTKAA (SEQ ID NO:64)。

[0315] 方面19. 一种药物组合物, 其包含: a) 方面1至18中任一项的重组腺相关病毒病毒体; 和 b) 药学上可接受的赋形剂。

[0316] 方面20. 一种将基因产物递送至个体中的视网膜细胞的方法, 所述方法包括向所述个体施用方面1至18中任一项的重组腺相关病毒 (rAAV) 病毒体。

[0317] 方面21. 方面20的方法, 其中所述基因产物为多肽。

[0318] 方面22. 方面20的方法, 其中所述基因产物为短干扰RNA或适体。

[0319] 方面23. 方面21的方法, 其中所述多肽为神经保护因子、抗血管生成多肽、抗细胞凋亡因子或增强视网膜细胞的功能的多肽。

[0320] 方面24. 方面21的方法, 其中所述多肽为神经胶质源性神经营养因子、成纤维细胞生长因子2、神经粘蛋白、睫状神经营养因子、神经生长因子、脑源性神经营养因子、表皮生长因子、视紫红质、X连锁凋亡抑制因子、视网膜劈裂蛋白、RPE65、视网膜色素变性GTP酶相互作用蛋白-1、外周蛋白、外周蛋白-2、视紫红质、RdCVF、视网膜色素变性GTP酶调控因子 (RPGR) 或音猬因子。

[0321] 方面25. 方面21的方法, 其中所述多肽为RNA引导的核酸内切酶。

[0322] 方面26. 一种治疗眼病的方法, 所述方法包括向有需要的个体施用有效量的方面1至18中任一项所述的重组腺相关病毒 (rAAV) 病毒体。

[0323] 方面27. 方面26的方法, 其中所述施用是通过眼球内注射。

[0324] 方面28. 方面26的方法, 其中所述施用是通过玻璃体内注射。

[0325] 方面29. 方面26的方法, 其中所述眼病为青光眼、视网膜色素变性、黄斑变性、视网膜劈裂症、利伯氏先天性黑朦(Leber's Congenital Amaurosis)、糖尿病性视网膜病、全色盲或色盲。

[0326] 方面30. 一种分离的核酸, 其包含编码变异腺相关病毒(AAV)衣壳蛋白的核苷酸序列, 其中所述变异AAV衣壳蛋白相对于对应的亲本AAV衣壳蛋白, 在衣壳蛋白GH环中包含约5个氨基酸至约20个氨基酸的插入, 且其中所述变异衣壳蛋白在存在于AAV病毒体中时, 增加所述AAV病毒体对视网膜细胞的感染性, 且其中所述氨基酸插入在天然AAV衣壳的GH环中, 其中所述插入为具有式I-X中的任一个的肽。

[0327] 方面31. 方面30的分离的核酸, 其中所述插入位点介于AAV2的氨基酸587与588之间、AAV1的氨基酸590与591之间、AAV5的氨基酸575与576之间、AAV6的氨基酸590与591之间、AAV7的氨基酸589与590之间、AAV8的氨基酸590与591之间、AAV9的氨基酸588与589之间或AAV10的氨基酸588与589之间。

[0328] 方面32. 一种分离的经遗传修饰的宿主细胞, 其包含方面30或方面31的核酸。

[0329] 方面33. 一种变异腺相关病毒(AAV)衣壳蛋白, 其中所述变异AAV衣壳蛋白包含约5个氨基酸至约20个氨基酸的插入, 其中所述氨基酸插入在天然AAV衣壳的GH环中, 其中所述插入为具有式I-X中的任一个的肽。

[0330] 方面34. 在方面1至33中的任一项中, 插入GH环中的所述异源肽可以具有式I-X中的一个, 其中:

[0331] 式I为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0332] X_1 为Leu、Ile、Pro或Gln;

[0333] X_2 为Ala、Pro、Ser、Asp、Gly、Thr或Val;

[0334] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg、Ala、Asp、Glu、Asn、Gln或Tyr;

[0335] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile、Asn、Glu、Lys或Arg;

[0336] X_5 为Asp、Ser、Gln、Val、Thr、Gly、Ala、Asn、Lys或Tyr;

[0337] X_6 为Thr、Ala、Gln、Ser、Glu、Pro或Ile;

[0338] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys、Ala或Cys; X_8 为Lys、Ser、Arg、Thr、Ala、Glu、Ile或Asn;

[0339] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且 X_{10} 为Ala、Phe、Asp、Thr、Val或Met;

[0340] 式II为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

[0341] X_1 为Leu、Ile或Pro;

[0342] X_2 为Ala、Pro或Ser;

[0343] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg或Ala;

[0344] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile或Asn; X_5 为Asp、Ser、Gln、Val、Thr、Gly或Ala;

[0345] X_6 为Thr、Ala、Gln、Ser、Glu或Pro;

[0346] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys或Ala;

[0347] X_8 为Lys、Ser、Arg或Thr;

[0348] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且 X_{10} 为Ala;

[0349] 式III为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:

- [0350] X_1 为Leu、Ile或Pro;
- [0351] X_2 为Ala、Pro或Ser;
- [0352] X_3 为Lys、His、Thr、Ile、Pro、Val、Arg或Ala;
- [0353] X_4 (如果存在) 为Gln、Asp、Ser、Gly、Thr、Ile或Asn; X_5 为Asp、Ser、Gln、Val、Thr、Gly或Ala;
- [0354] X_6 为Thr、Ala、Gln、Ser、Glu或Pro;
- [0355] X_7 为Thr、Ser、Asn、Pro、Leu、Gln、Lys或Ala; X_8 为Lys、Ser、Arg或Thr;
- [0356] X_9 为Asn、Pro、Ser、Lys、His、Ile、Thr或Ala; 且 X_{10} 为Ala、Thr、Asp Val或Met;
- [0357] 式IV为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:
- [0358] X_1 为Leu;
- [0359] X_2 为Ala;
- [0360] X_3 为Lys、His、Thr、Ile、Pro或Val;
- [0361] X_4 (如果存在) 为Gln、Asp、Ser或Gly;
- [0362] X_5 为Asp、Ser或Gln;
- [0363] X_6 为Thr、Ala、Gln或Ser;
- [0364] X_7 为Thr或Ser;
- [0365] X_8 为Lys、Ser或Arg;
- [0366] X_9 为Asn、Pro或Ser; 且
- [0367] X_{10} 为Ala;
- [0368] 式V为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:
- [0369] X_1 为Leu;
- [0370] X_2 为Ala;
- [0371] X_3 为Lys或His;
- [0372] X_4 (如果存在) 为Gln、Asp、Ser或Gly;
- [0373] X_5 为Asp、Ser或Gln;
- [0374] X_6 为Thr、Ala、Gln或Ser;
- [0375] X_7 为Thr或Ser;
- [0376] X_8 为Lys、Ser或Arg;
- [0377] X_9 为Asn、Pro或Ser; 且
- [0378] X_{10} 为Ala;
- [0379] 式VI为 $X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}$, 其中:
- [0380] X_1 为Leu;
- [0381] X_2 为Ala;
- [0382] X_3 为Asn、Lys、Thr、Gln、Ser、Ile或Leu;
- [0383] X_4 为Ser、Ala、Thr、Glu、Gln、Gly、Lys或Pro; X_5 为Asp、Pro、Glu、Thr、Asn或Arg;
- [0384] X_6 为Ile、His、Thr、Gln、Asn、Tyr、Asp或Glu; X_7 为Gln、Thr、Asn、Ala或Lys;
- [0385] X_8 为Lys、Thr、Arg或Asp;
- [0386] X_9 为Pro、Asn、Thr、Arg、Lys或Ser; 且
- [0387] X_{10} 为Ala;

[0388] 式VII为LAHQDTTKX₁X₂X₃(SEQ ID NO:148),其中X₁为Lys、Thr、Asn或His;X₂为Ala、Thr、Val、Ile、Met或Asp;且X₃如果存在,那么为Ala;

[0389] 式VIII为LAX₁QX₂TX₃X₄X₅X₆(SEQ ID NO:149),其中X₁为Ala、Pro、Asp或His;X₂为Gly或Asp;X₃为Ala、Thr或Lys;X₄为Asn、Glu、Lys、Arg或Thr;X₅为Leu、Asn、Lys或Thr;且X₆如果存在,那么为Ala、Thr、Asp、Val或Met;

[0390] 式IX为X₁AX₂X₃DX₄TKX₅A(SEQ ID NO:150),其中X₁为Val或Leu;X₂为Ile、Val、His或Asp;X₃为Glu、Ser、Lys或Gln;X₄为His、Ser或Thr;且X₅为Ser、Ala、Asn、His或Lys;且

[0391] 式X为X₁X₂X₃AX₄QX₅TX₆KNA(SEQ ID NO:151),其中X₁如果存在,那么为Leu;X₂如果存在,那么为Ala;X₃为Lys、Leu或Pro;X₄为Asn、His、Pro或Tyr;X₅为Asn、Gly、Val或Asp;且X₆为Pro或Thr。

[0392] 实施例

[0393] 提出以下实施例以便为一般技术者提供关于如何进行和使用本发明的完整公开和描述,且既不意图限制本发明人视为其发明的范围,也不意图表示以下实验是所执行的全部或唯一实验。虽然努力确保关于所用数字(例如量、温度等)的准确性,但应考虑一些实验误差和偏差。除非另外指示,否则份为重量份,分子量为重量平均分子量,温度为摄氏温度,且压力为大气压或接近大气压。可能使用标准缩写,例如bp,碱基对;kb,千碱基;pL,皮升;s或sec,秒;min,分钟;h或hr,小时;aa,氨基酸;kb,千碱基;bp,碱基对;nt,核苷酸;i.m.,肌肉内;i.p.,腹膜内;s.c.,皮下等等。

[0394] 实施例1:具有AAV衣壳变体的AAV病毒体的产生和表征

[0395] 迭代体内筛选方法用于产生具有能克服显著和复杂障碍的衣壳变体的AAV,预防大型动物眼睛中的全视网膜AAV感染。犬为视网膜变性疾病重要的临床前模型,其眼睛尺寸和结构类似于人类,且视网膜疾病的许多形式在多种犬品种中是天然存在的。筛选法用于鉴别96种能够在犬科视网膜中进行全视网膜感染的AAV变体。深度测序用于定量来自筛选的AAV变体池的这些变体中的18种在犬科视网膜中的性能。基于玻璃体内注射后视网膜细胞中的病毒DNA和mRNA的水平定量感染性。这些变体可以用于大型动物和人眼中的多种基因递送策略。

[0396] 产生在AAV衣壳上的表面暴露位置含有随机21核苷酸插入(被5' 6-核苷酸连接子和3' 3-核苷酸连接子围绕)的肽呈现库。包装病毒,使得每个病毒基因组被包壳在该基因组编码的衣壳蛋白壳内。因此,通过选择而鉴别的功能改善可以连接于病毒衣壳内所含有的基因组序列。由此库,使用迭代体内筛选选择过程鉴别能够从玻璃体感染犬科视网膜的变体(图1)。每轮向犬科眼睛注射约250μL 10E+13-10E+14病毒基因组/mL(vg/mL)滴度病毒。注射三周后,摘出眼睛,且从视网膜的中央和边缘区域取出视网膜穿孔。将RPE细胞从视网膜组织分离,并将组织冷冻。接着从视网膜细胞收集DNA,且将cap基因从分离的样品进行聚合酶链反应(PCR)扩增。cap基因用于随后的AAV包装。

[0397] 图1.定向进化法的例示,用于研发犬科视网膜AAV变体。产生肽呈现库,包装成AAV载体,并经由玻璃体内注射来注射至犬科眼睛中。迭代数轮的选择用于从载体池阳性选择AAV变体。三轮选择后为一轮易错PCR,接着为额外的数轮选择。

[0398] 在5轮选择后,Illumina深度测序用于鉴别AAV变体库中相对表示在数轮中增加的变体。病毒库中表示的增加表明阳性选择和从玻璃体感染犬科视网膜的能力。从约10E+7变

体的库中,在体内筛选中选择的最高96种变体提供于表1中。

[0399] 表1

	肽 No.	SEQ ID NO:
LAKDATKNA	1	47
PAHQDTTKNA	2	48
LAHQDTTKNA	3	49
LATTSQNKPA	4	50
LAISDQTKHA	5	51
IARGVAPSSA	6	52
LAPDSTTRSA	7	53
LAKGTELKPA	8	54
LAIIDATKNA	9	55
LAVDGAQRSA	10	56
PAPQDTTKKA	11	57
LPHQDTTKNA	12	58
LAKDATKTIA	13	59
LAKQQSASTA	14	60
LAKSDQSKPA	15	61
LSHQDTTKNA	16	62
LAANQPSKPA	17	63
LAVSDSTKAA	18	64
LAAQGTAKPA	19	65
LAPDQTTRNA	20	66
LAASDSTKAA	21	67
LAPQDTTKNA	22	68
[0400] LAKADETRPA	23	69
LAHQDTAKNA	24	70
LAHQDTKKNA	25	71
LAHQDTTKHA	26	72
LAHQDTTKKA	27	73
LAHQDTTRNA	28	74
LAHQDTTNA	29	75
LAHQGTTKNA	30	76
LAHQVTTKNA	31	77
LAISDQSKPA	32	78
LAKDATKTA	33	79
LAKDTTKNA	34	80
LAKSDQSRPA	35	81
LAPQDTKKNA	36	82
LATSDSTKAA	37	83
LAVDGSQRSA	38	84
LPISDQTKHA	39	85
LPKDATKTIA	40	86
LPPQDTTKNA	41	87
PAPQDTTKNA	42	88
QAHQDTTKNA	43	89
LAHETSPRPA	44	90
LAKSTSTAPA	45	91
LADQDTTKNA	46	92

	肽 No.	SEQ ID NO:
LAESDQSKPA	47	93
LAHKDTTKNA	48	94
LAHKTQQKM	49	95
LAHQD TTENA	50	96
LAHQD TTINA	51	97
LAHQD TTKKT	52	98
LAHQD TTKND	53	99
LAHQD TTKNT	54	100
LAHQD TTKNV	55	101
LAHQD TTKTM	56	102
LAHQNTTKNA	57	103
LAHRD TTKNA	58	104
LAISDQTNHA	59	105
LAKQKSASTA	60	106
LAKSDQCKPA	61	107
LAKSDQSKPD	62	108
LAKSDQSNPA	63	109
LAKSYQSKPA	64	110
LANQD TTKNA	65	111
LAPQNTTKNA	66	112
LAPSSIQKPA	67	113
LAQQD TTKNA	68	114
LAYQD TTKNA	69	115
[0401] LDHQD TTKNA	70	116
LDHQD TTKSA	71	117
LGHQD TTKNA	72	118
LPHQD TTKND	73	119
LPHQD TTKNT	74	120
LPHQD TTNNA	75	121
LTHQD TTKNA	76	122
LTKDATKTIA	77	123
LTPQD TTKNA	78	124
LVHQD TTKNA	79	125
LAKANQNTPA	80	126
LATTPITKPA	81	127
LATTPIAKPA	82	128
LAIEDHTKSA	83	129
LAQSEHQRPA	84	130
LAKSPNKDNA	85	131
LANQDYTKTA	86	132
LANSTDQTRA	87	133
LALGETTRPA	88	134
LANSTEQTRA	89	135
LAQAD TTKNA	90	136
LASKDITKTA	91	137
LASPRHNKKC	92	138
LAHQD TTKTIA	93	139
LAAQGTANL	94	140

		肽 No.	SEQ ID NO:
[0402]	VAIEDHTKSA	95	141
	LAKANQNTPKNA	96	142

[0403] 使用高通量测序进一步定量表1中描绘的96种变体中的最高18种变体感染犬科视网膜的能力。表2描绘选择用于进一步定量的最高18种变体。

[0404] 表2

	LAKDATKNA (SEQ ID NO:47)	LAPDSTTRSA (SEQ ID NO:53)	LAKDATKTIA (SEQ ID NO:59)
	PAHQDTTKNA (SEQ ID NO:48)	LAKGTELKPA (SEQ ID NO:54)	LAKQQSASTA (SEQ ID NO:60)
	LAHQDTTKNA (SEQ ID NO:49)	LAIIDATKNA (SEQ ID NO:55)	LAKSDQSKPA (SEQ ID NO:61)
[0405]	LATTSQNKPA (SEQ ID NO:50)	LAVDGAQRSA (SEQ ID NO:56)	LSHQDTTKNA (SEQ ID NO:62)
	LAISDQTKHA (SEQ ID NO:51)	PAPQDTTKKA (SEQ ID NO:57)	LAANQPSKPA (SEQ ID NO:63)
	IARGVAPSSA (SEQ ID NO:52)	LPHQDTTKNA (SEQ ID NO:58)	LAVSDSTKAA (SEQ ID NO:64)

[0406] 用驱动GFP表达的普遍CAG启动子,包装十八种变体。GFP cDNA与独特的25碱基对条形码标识符融合。18种变体中的每一种用独特的GFP条型码包装。将包装的变体以同等比率混合,且与对照的基于AAV2的载体(代表天然存在的亲本血清型的阴性对照)一起注射至视网膜中。注射后,从光感受器和RPE细胞收集DNA和mRNA。对DNA和mRNA水平进行定量以测定犬科来源的载体递送DNA至视网膜和引起转基因表达的能力(图2)。

[0407] 图2. 含有GFP-条型码构建体的变体的深度测序。犬科来源的变体对犬科视网膜的感染通过标记GFP cDNA和mRNA的深度测序来定量。

[0408] 使用冷冻的视网膜切片的共焦显微术将18成员库的表达成像。GFP表达表明18成员库靶向内视网膜中的视网膜细胞和外视网膜中的光感受器(图3)。

[0409] 图3. 18成员犬科来源的AAV变体库感染神经节细胞层、内核层、光感受器层和RPE层中的细胞。

[0410] 在测试的最高18种变体中,2种变体引起最高水平的DNA和mRNA恢复。引起最高水平的DNA恢复的变体具有插入序列~588-PAPQDTTKKA (SEQ ID NO:57)。引起最高水平的mRNA表达的变体具有插入序列~588-LAPDSTTRSA (SEQ ID NO:53)。

[0411] 虽然本发明已经参考其特定的实施方案描述,但应了解本领域的技术人员可以进行多种改变且可以在不脱离本发明真实的精神和范围下用同等物替代。另外,可以进行许多改变以使具体的情况、材料、物质组合物、方法、方法步骤适应本发明的目标、精神和范围。所有这类的改变意图在这里附上的权利要求书的范围内。

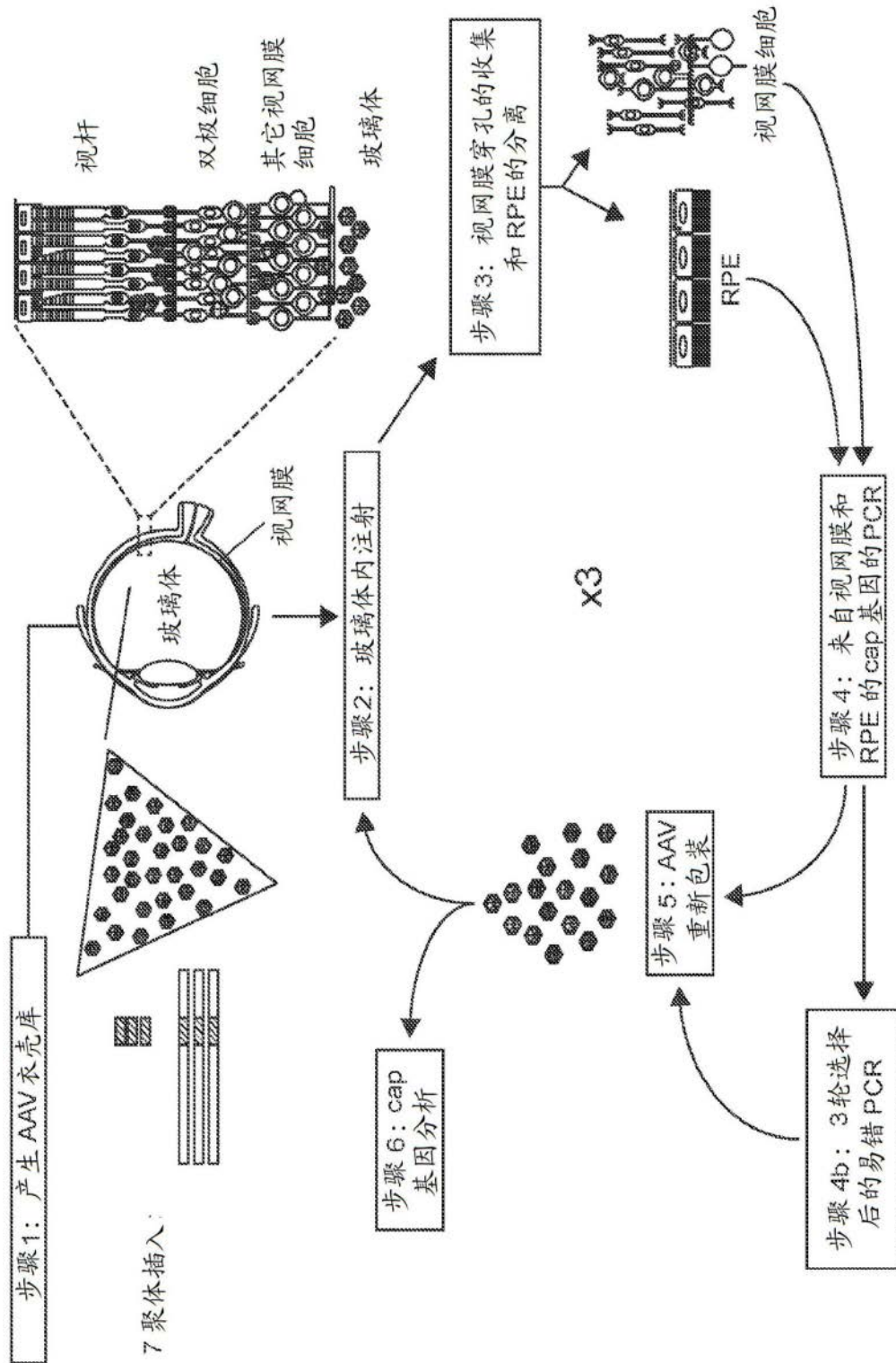


图1

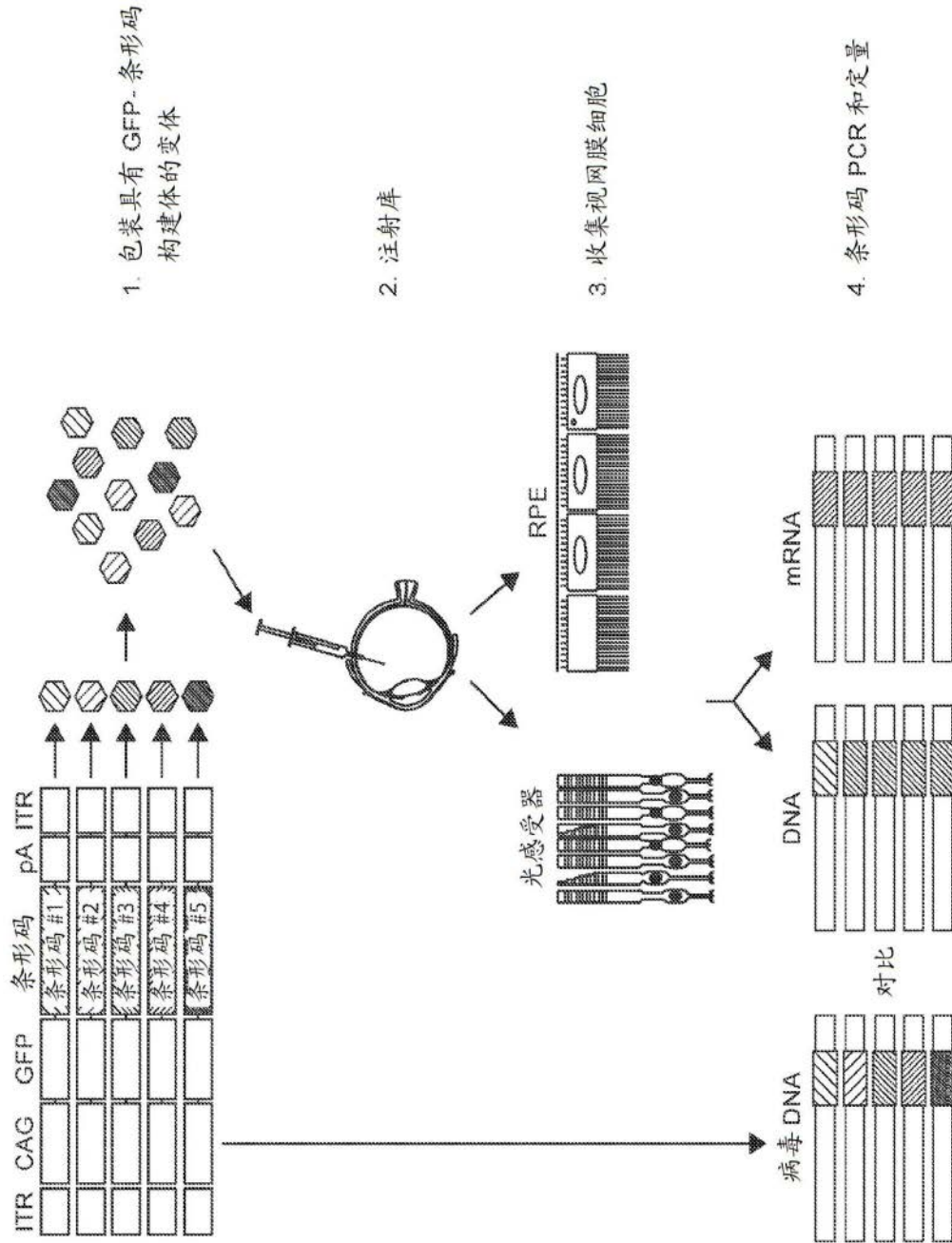


图2

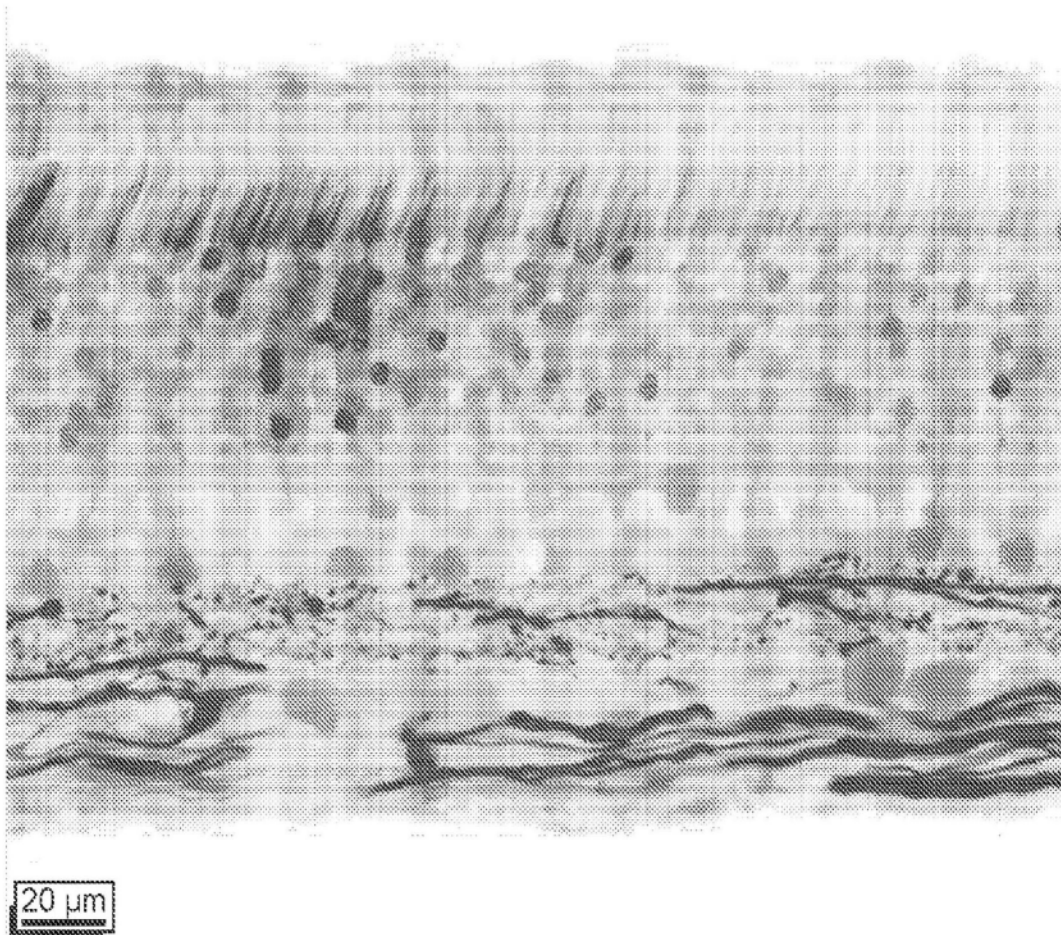


图3

AAV2	VP1	1	MAADGYLPDWLEDTLSEGIQWVKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD
AAV2	VP1	61	KGEPVNEADAAALEHDKAYDRQLDSCDNPYLYKYNHADAEEQERLKEDTSFGCNLGRAVFQ
AAV2	VP1	121	AKKRVLEPLGLVEEPVKTPAGKKRPVEHSPVEPSSSGTGKAGQQFARKRLNFGQTGDAD
AAV2	VP1	181	SVFDPQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNNWHCDSTWMGDRV
AAV2	VP1	241	TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDENRFHCHESPRDWQRLI
AAV2	VP1	301	NNNWGERPKRLNFKLFNIOQKEVTQNDGTTTIANNTSTVQVFTDSEYQLPYVLGSAHQG
AAV2	VP1	361	CLPFPFADVFMPQYGYLTINNGSQAVGRSSFYCLEYFPSQMLRTGNNFTSFYTFEDVFF
AAV2	VP1	421	HSSYAHSQSLDRIMNPLIDQYLYLSRNTPSGTTTQSRMQFSQAGASDIRDQSRNWLPG
AAV2	VP1	481	PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFFPQSGVL
AAV2	VP1	541	IFGKQGSEKTNVDIEKVMITDEEEIRTNPVATEQYGSVSTNLQRG NR QAATADVNTQGV
AAV2	VP1	601	LPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGFGGLKHPPQILLIKNTVPANPSTT
AAV2	VP1	661	FSAAKFASFITQYSTQGVSVIEIWEIQLKENSKRWNPEIQYTSYNKSVNVDFTVDTNGVI
AAV2	VP1	721	SEPRPIGTRYLTR (SEQ ID NO:1)

图4

AAV-2	570	PVATEQYGSVSTNLQRC NR QAATADVNTQGVLPGMVWQDRDV	611	(SEQ ID NO:2)
AAV-1	571	PVATERFGTVAVNFEQSSST DF ATG DVHAMGALPGMVWQDRDV	612	(SEQ ID NO:3)
AAV-5	560	RVAYNVGGQMATNNQ SS TTAPATGTYNLQEI VPGSVWMERDV	601	(SEQ ID NO:4)
AAV-6	571	PVATERFGTVAVNLEQSSST DF ATG DVHVMGALPGMVWQDRDV	612	(SEQ ID NO:5)
AAV-7	572	PVATEEYGIVSSNLQAA NT AAQTQVVNNQGALPGMVWQNRDV	613	(SEQ ID NO:6)
AAV-8	573	PVATEEYGIVADNLQQQ NT APQIGTVNSQGALPGMVWQNRDV	614	(SEQ ID NO:7)
AAV-9	571	PVATESYGQVATNHQSA QA QAQTGWVQNQGILPGMVWQDRDV	612	(SEQ ID NO:8)
AAV-10	573	PVATEQYGVVADNLQ QA NTGPIVGNVNSQGALPGMVWQNRDV	614	(SEQ ID NO:9)

图5

AAV1	KDDEDKFFPMSGVMI FGK--ESAGASNTALD-NVMITDEEEIKATNPVATERFGTVAVNF	584
AAV6	KDDKDKFFPMSGVMI FGK--ESAGASNTALD-NVMITDEEEIKATNPVATERFGTVAVNL	584
AAV3	KDDEEKFFPMSGVMI FGK--EGTTASNAELD-NVMITDEEEIRTTNPVATEQYGTVANL	584
AAV2	KDDEEKFFPMSGVMI FGK--QGSEKTNVDIE-KVMITDEEEIRTTNPVATEQYGSVSTNL	583
AAV8	KDDEERFFPMSGVMI FGK--QNAARDNADYS-DVMLTSEEEIKTTNPVATEEYGI VADNL	586
AAV8.1	KDDEERFFPMSGVMI FGK--QNAARDNADYS-DVMLTSEEEIKTTNPVATEEYGI VADNL	586
AAV8 rh8	KDDDDRRFFPSSGVLI FGK--QGAGNDGVDYS-QVLI TDEEEIKATNPVATEEYGA VAINN	584
AAV10	KDDEERFFPSSGVLMFGK--QGAGRDNVDYS-SVMLTSEEEIKTTNPVATEQYGVVADNL	586
AAV7	KDDEDRRFFPSSGVLI FGK--TGAT-NKTTLE-NVLMTNEEEIRPTNPVATEEYGI VSSNL	585
AAV9	KEGEDRRFFPLSGSLI FGK--QGTGRDNVDAD-KVMITNEEEIKTTNPVATESYGVVATNH	584
AAV9.1	KEGEDRRFFPLSGSLI FGK--QGTGRDNVDAD-KVMITNEEEIKTTNPVATESYGVVATNH	584
AAV5	NLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATTNN	573
	: : : : : : : : : : : : : : : : *	
AAV1	QSSSTDPATGVDVHAMGALPGMVWQDRDVY LQGPIWAKI PHTDGHFHPSPMLMGFGGLKNPP	644
AAV6	QSSSTDPATGVDVHAMGALPGMVWQDRDVY LQGPIWAKI PHTDGHFHPSPMLMGFGGLKNPP	644
AAV3	QSSNTAPTGTGNHQGALPGMVWQDRDVY LQGPIWAKI PHTDGHFHPSPMLMGFGGLKNPP	644
AAV2	QRCNRQAATAADVNTQGVLPGMVWQDRDVY LQGPIWAKI PHTDGHFHPSPMLMGFGGLKNPP	643
AAV8	QQQNTAPQIGTVNSQGALPGMVWQNRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	646
AAV8.1	QGRQAQIGTVNSQGALPGMVWQNRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	646
AAV8 rh8	QAANTQAQTGLVHNQGVIPGMVWQNRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	644
AAV10	QAANTGPIVGNVNSQGALPGMVWQNRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	646
AAV7	QAANTAAQTQVVNNQGALPGMVWQNRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	645
AAV9	QSAQAQAQTGWVQNQGILPGMVWQDRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	644
AAV9.1	QSGQAQAQTGWVQNQGILPGMVWQDRDVY LQGPIWAKI PHTDGNFHPSPMLMGFGGLKNPP	644
AAV5	QSSTTAPATGTYNLQEI VPGSVWMERDVY LQGPIWAKI PETGAHFHPSPAMGGFGGLKNPP	633
	* : *	

图6B

AAV1	PQILIK--	650	(SEQ	ID	NO: 35)
AAV6	PQILIK-	650	(SEQ	ID	NO: 36)
AAV3	PQIMIK-	650	(SEQ	ID	NO: 37)
AAV2	PQILIKN	650	(SEQ	ID	NO: 38)
AAV8	PQILIKN	653	(SEQ	ID	NO: 39)
AAV8.1	PQILIKN	653	(SEQ	ID	NO: 40)
AAV8 rh8	PQILIKN	651	(SEQ	ID	NO: 41)
AAV10	PQILIKN	653	(SEQ	ID	NO: 42)
AAV7	PQILIKN	652	(SEQ	ID	NO: 43)
AAV9	PQILIK--	650	(SEQ	ID	NO: 44)
AAV9.1	PQILIK-	650	(SEQ	ID	NO: 45)
AAV5	PMMLIKN	640	(SEQ	ID	NO: 46)
	* : *				

图6C

视网膜劈裂蛋白-1
智人

```

1  msrkiegfll  lllfgyeatl  glstedege  dpwyqkackc  dcgggnalw  sagatsldci
61  pecpyhkpig  fesgevtpdq  itcsnpeqyv  gwyswtank  arlnsqgfgc  awlskfgdss
121 qwlqidlkei  kvisgiltgg  rcdidewmtk  ysvqyrtder  lnwiyykdqt  gnnrvfygns
181 drtstvtqnll  rppiisrfir  liplgwhvri  airmellecv  skca (SEQ ID NO:10)

```

图7A

BDNF
智人

```

1  mtilfltmvi syfgcmkaap mkeanirggg glaypgvrth gtlesvngpk agsrgltsla
61  dtfehvieel ldedhkvrpn eenkdadly tszvmlesqv plep1lfl1 eeyknyldaa
121 nmsmmvlrhs dparrgelsv cdsisewvta adktavdms ggtvtvlekv pvskgqlkqy
181 fyetkcnpmg ytkegcrgid krhwnsqcrt tqsyvraltm dskkrigwrf iridtscvct
241 ltikrgr (SEQ ID NO:11)

```

图7B

RPE65
智人

```

1  msiaqvehpag gykklfetve elssplltahv tgriplwltg slircpglff evgsepfyhl
61 fdgqallhkf dfkeghvtyh rrfirtdayv ramtekriivi tefgtcafpd pcknifsrff
121 syfrgvevtd nalvnvypvg edyyactetn fitkinpetl etikqvdlcn yvsvngatah
181 phlendgtvy nigncfgnf siaynivkip plqadkedpi skseiivvqfp csdrfkpsyv
241 hsfgltpnyi vfvetpvkin lfkflsswsl wqanymdcfe snetmgvwlh iadkkrrkyl
301 nnkyrtspfn lfhhintyed ngflivdlcc wkgfefvyny lylanlrenw eevkknarka
361 pqpevriryvl plnidkadtg knlvtlpntt atailcsdet iwlepevlfs gprqafefpq
421 inyqkycgkp ytyayglgln hfvpdrlickl nvtketwvw qepdsypsep ifvshpdale
481 eddgvvlsvv vspgagqkpa yllilnakdl sevараеvei nipvtfhglf kks
      (SEQ ID NO:12)

```

图7C

外周蛋白-2
智人

```

1  mallkvkfdq kkrvklaggl wlmwfsvla giilfslglf lkielrkrsd vmnnseshfv
61  pnsliqngvl scvfnsilagk icydaldpak yarwkpwlkp ylaicvlfni ilflvalccf
121 llrgslentl ggglkngmk yrdtdtpgrc fmkktdmlq iefkccgng frdwfeiqwi
181 snryldfssk evkdriksnv dgrylvdgvp fscnpspr pciqyqitnn sahysydhqt
241 eelnlwvrgc raallsyss lmsngvvtl liwlfevtit iglrylqtsl dgvsnpese
301 sesqgwller svpetwkaf1 esvkkkqkgn qveaegadag qapeag (SEQ ID NO:13)

```

图7D

外周蛋白
智人

```

1  mshhpsglra gfsstsyrrt fgpppslspg afsyssssrf sssrllgsas psssvrlgsf
61  rspragagat lrlpserldf smaeanqef latrsnekqe lqelndrfan fiekvrfleq
121 qnaalrgels qargqepara dqlcqqlre lrrelellgr erdrvqverd glaedlaalk
181 qrleeetrkr edaehnlvlf rkdvddatls rlelerkies lmdeiefkk lheeelrdlq
241 vsvesqqvqq veveatvkpe ltaalrdira qyesiaaknl qeaeewyksk yadlsdaanr
301 nhealrqakq emnesrrqiq sltcevdglr gtneallrql releeqfale aggyqagaar
361 leeelrqlke emarhlreyq ellnvkmald ieiatyrrkll egeesrisvp vhsfaslnik
421 ttvpeveppq dshsrktvli ktietrngev vtesqkeqrs eldkssahsy (SEQ ID NO:14)

```

图7E

RPCR-相互作用蛋白-1 智人

```

1  mshlvdpstsg dlpvrddidai plvlpaskgk nmktqpplsr mnreeledsf frlrhdhnlv
61  kelswkqgde ikrlrttllr ltaagrldrv aeeaaaplset arrgqkagwr qrlsmhqrpq
121 mhrllqghfhc vgpasprraq prvqvghrql htagapvppek pkrqprdris ytappsfkeh
181 atnenrgeva skpselvsqs nsisfssvi smakpigliem pnsahimasn tmqveepkps
241 pekmpkden feqrssleca qkaaelrasi kekvelirlk kllhernasl vmtkaqltev
301 qeayetllqk nggillsaah allkqvnelr aelkeeskka vslksqledv silqmtlkef
361 qervedleke rkllndnydk llesmldssd sssqphwsne liaeqllqqv sqlqddqldae
421 ledkrkvllle lsrekaqned iklevtnllq khkqevellq naatisqppd rqsepathpa
481 vlqentqieq sepknqeek lsqvlnelqv shaettelle ktrdmlllqr kinvcygeel
541 eamntkadnd nrhkekler ltrlldlknk rikqllegilr shdlptseqi kdvaaygtrpl
601 slcletlpah gdedkvdisl lhqgenlfel hihqafitsa alaqagdtqp ttftctysfyd
661 fethctplsv gpqplydfts qyvmetdsif lhylqeasar ldihqamase hstlaagwic
721 fdrvletvek vhglatliga ggeefgvley wmlrfpikp slqacnkrkk agvylstdvl
781 ggrkaqeeef rseswepqne lwieitkccg lrsrwlgtqp spyavyrfft fsdhdtaaiip
841 asnnpyfrdq arfpvlvtsd ldhylrreal sihvddedl epgsylgrar vpllplakne
901 sikgdfnltd paekpnsgisq vqldwkfpvi ppsflkpea qtkgkdtkds skisseeka
961 sfpsqddqmas pevpieaggy rskrkpphgg erkekehqv erkekehqv sysrrkhgkr igvqgknme
1021 ylslnilngn tpeqvnyteu kfsetnsfig dgfnqheeee emtlshsalk qkeplhpvnd
1081 kesseqgsev seaqtdtdsd viivpmsqky pkadseknci eivslafype aevmsdenik
1141 qvyveykfyd lplsetetpv slrkpragee ihfhfskvid ldpqeqggr rflfdmlngq
1201 dpdqqghlkt vvsdpideek keceevgyay lqlwqilesq rdlileqeldi vspedlatpi
1261 grlkvslqaa avlhailykem tedlfs (SEQ ID NO:15)

```

图7F

Rab 护送蛋白-1

```

1 madtlpsefd vivigtglpe siiaaacrs qrrvlhvdsr syyggnwasf sfsqllswlk
61 eyqensdivs dspvwgdqil eneeaialsr kdktiqhvev fcyasqdlhe dveeagalqk
121 nhalvtsans teaadsaflp tedeslstms cemlteqtps sdpenalevn gaevtgeken
181 hcdkktcvps tsaedmsenv piaedtteqp kknritysqi ikegrrrfnid lvskllysrg
241 llidlliksn vsryaefkni trilafragegr veqvpcsrar vfnskqitmv ekrmlmkflt
301 fcmeyekypd eykgyeeitf yeylktqklt pnlqyivmhs iamtsetass tidglkatkn
361 flhclgrygn tpflfplygq gelpqcfcrm cavfggiycl rhsvqqlvvd kesrkckail
421 dqfggriise hflvedsyfp enmcsrvqyr qisravlitd rsvlktddsq qisiltvpae
481 epgtfavrvl elcsstmtcm kgtylvhltc tssktaredl esvvqqlfvp ytemeieneq
541 vekprillwal yfnmrdsdi srscyndlps nvvyvcsqgpc glgndnavkq aetlfqeicp
601 nedfcppppn pediildgds lqpeasessa ipeansetfk estnlgnee sse

```

(SEQ ID NO:16)

图7G

RdCVF 的 212 氨基酸同种型

```
1 maslfsgril irnnsdqdel dteaevsrrl enrllvllffg agacpqcqaf vpilkdffvr  
61 ltdefyvlra aqlalvyvsq dsteeqqdlf lkdnppkkwlf lpfeddllrrd lgrqfsverl  
121 pavvvlkpdg dvltrdgade lqrlgtacfa hwqaaaevld rnfqlpedle dgeprsltec  
181 lrrhkyrvek aarggrdp99 gggeeggagg lf (SEQ ID NO:17)
```

图7H

RdCVF 的 156 氨基酸同种型 (同种型 1)

```
1 mvdilgerhl vtckgatvea eaalqkvva lyfaaaarcap srdftpllcd fytalvaeear  
61 rpapfevvfv sadgssqeml dfmrelhgaw lalpfdpyr helrkrynvt aipklvivkq  
121 ngevltkngr kqirerglac fqdweaaadi fgnfsv (SEQ ID NO:18)
```

图7I

RdCVF 的 135 氨基酸同种型 (同种型 2)

```
1 mvdilgerhl vtckgatvea eaalqnkvva lyfaaarcap srdftplldc fytalvaeear  
61 rpapfevfvv sadgssqeml dfmrelhgaw lalpfhdpyr qrsllallprrl ecsgvillahc  
121 nlcllgssds lalas (SEQ ID NO:19)
```

图7J

视杆 cGMP 特异性 3',5'-环状磷酸二酯酶子单元 α (PDE6 α)
GenBank NP_00431

```

1  mgevtaeeve kfldsnigfa kqyynlhyra klisdllgak eaavdfsnyh spssmeesei
61 ifdl1lrdfqe nlqtekcifn vmkklcfl1q adrm1fmyr trngiaelat rlfvnhkdav
121 ledclvmpdq eivfpldngi vghvahskki anvpnteede hfcdfvdlit eyktnilias
181 pinngkdvva iimavnkvdg shftkrdeei llkylnfanl inkvyhlsyl hncetrrgqi
241 llwsgskvfe eltdierqfh kalytvrafl ncdrysvgll dmtkqkeffd vwpvlmgev
301 pysgprtpdg reinfykvid yilhgkedik vipnppdhw alvsq1payv aqnglicnim
361 napaedffaf qkepldesgw miknvlsmpl vnkkeeivgv atfynrkdgk pfdemdetlm
421 esltqflgws vlnpdtyesm nklenrkdif qdivkyhvk lpltelelvk trevygkep
481 eceeeelaai lqaelpdack yeinkfhfsd wrhgf1nvqgt mfsllvtgkl kryftdleal amvtaafchd
541 alvr1fmysls kgyrkityhn wrhgf1nvqgt mfsllvtgkl kryftdleal amvtaafchd
601 idhrgtnnly qmksqnplak lhgssilerh h1efgktllr deslnifgnl nrrqhehaih
661 nm1ia1iatd lalyfkkrtn fgkivdqskt yeseqewtqy m1leqtrkei vmamntacd
721 lsaitkpwev qsqvallvaa efweqgdler tvlqqnpi1m mdrnkadelp klqv1gfidfv
781 ctfvykefsr fheeitpml1d gitnnrkewk aladey1dak kvqeek1kqk qsa1saaagn
841 qp1gn1psp1g1 attsksc1ci1q (SEQ ID NO:20)

```

图7K

视杆 cGMP 特异性 3',5'- 环状磷酸二酯酶子单元 β 同种型 1 (PDE6 β 同种型 1)
GenBank NP_000274

```

1  mslseeqars fldqnpdfer qyfgkklspe nvaaacedgc ppdcslrdl cqveestall
61  elvqdmqesi nmervvfkvl rrlctllqad rcslfmyrqr ngvaelatrl fsvqpdsvle
121 dclvpdpdsei vfpldigvvg hvaqtckmyn vedvaecephf ssfadeltdy ktknmlatpi
181 mngkdvvavi mavnklnpgf tdsededvfl kylvfatlyl kiyhlsylhn cetrngqvll
241 wsankvfeel ivfykvidyi lhgkeekvi ptpsadhwal asglpsyvae sgfcnimna
301 sgprtpdgre sademfkfge galddsgwli knvlsmplvn kkeeivgvat fynrkdgkpf deqdevlmes
361 sademfkfge galddsgwli knvlsmplvn kkeeivgvat fynrkdgkpf deqdevlmes
421 ltqflgwsvm ntdtydkmnk lenrkdiag mvlvkvkdr deiglilptr arlgkepadc
481 dedelgeilk eelpgpttfd iye fhfsdle cteldlvkcg iqmyyvelgvv rkfqipqevl
541 vrflfsiskg yrityhnwr hgfnvaqtmf tllmtgklks yytdleafam vtaglchdid
601 hrgtnnlyqm ksqnplaklh gssillerhhl efgkflisee tlniyqnlnr rqbehvhlm
661 diaiatdla lyfkkramfq kivedsknyq dkkswvcyls lcttrkeivm ammtacdls
721 aitkpwewqs kvallvaaef weggdlertv ldqqpipmnd rnkaaelpkl qvgfidfvct
781 fvykefsrfrh eellpnmfdr1 qnnrkewkal adeyeakvka leekeeeerv aakkvgteic
841 nggpapksst ccil (SEQ ID NO:21)

```

图7L

视杆 cGMP 特异性 3',5'- 环状磷酸二酯酶子单元 β 同种型 2 (PDE6 β 同种型 2)
GenBank NP_001138763

```

1 mslseeqars fldqnpdfar qyfgkklspe nvaaacedgc ppdcslrdl cqveestall
61 elvqdmqesi nmervvfkvl rrlctllqad rcslfmyrqx ngvaelatrl fsvqpdsvle
121 dclvpdpdsei vfpldigvvg hvaqtckmvm vedvaeceph sfadeltdy ktknmlatpi
181 mngkdvvavi mavnklingpf ftseodedvfl kylvfatlyl kiyhlsylhn cetrngqvll
241 wsankvfeel tdierqfhka fytvraylnc erysvglldm tkekeffdvw svlmgessqpy
301 sgprtpdgre ivfykvidyi lhgkeekvi ptpsadhwal asglpsyvae sgficnimna
361 sademfkfge galddsgwli knvlsmplvn kkeeivgvat fynrkdgkpf deqdevlmes
421 ltqflqwsvm ntdtydkmnk lenrkdiaqd mvlvhyvkcdx deiqlilptr arlgkepadc
481 dedelgeilk eelpgpttfd iyefhfsdle cteldlvkcg iqmyyelgvv rkfgqipgevl
541 vrflfsiskg yrrityhnwr hqfnvaqtmf tllmtgkiks yytdleafam vtaglchdid
601 hrgtnnlyqm ksqnplaklh gssilerhhl efgkflisee tlniyqnlnr rqhehvihlm
661 diaiaatdia lyfkkrampf kivdesknyq dkkswveyis lettrkeivm ammtacdls
721 aitkpwevqs kvallvaaef weggdlertv ldqqpipmmd rnkaaelpk1 qvgfidfvct
781 fvykefsrfh eeilpmfdrl qnnrkewkal adeyeakvka leekeeeerv aakkgteicn
841 ggpapksstc cil (SEQ ID NO:22)

```

图7M

视杆 cGMP 特异性 3',5'- 环状磷酸二酯酶子单元 β 同种型 3 (PDE6 β 同种型 3)
GenBank NP_001138764

```

1  mtkekeffdv wsvlmgesqp ysgprtpdgr eivfykvidy ilhgkeekv iptpsadhwa
61  lasglpsyva esgficninn asademfkfq egaldsgwl iknvlmpiv nkkeeivgva
121  tfynrkdgkp fdeqdevlme sitqflgsv mntdtydknn klenrkdiaq dmvlyhvkcd
181  rdeiqililpt rarigkepad cdedelgeil keelpgpttf diyefhfsdl ecteldlvkc
241  gigmyyelgv vrkfqipqev lvrlfsisk gyrrityhnw rhgfnvaqtm ftllmtgklk
301  syytdleafa mvtaglchdi dhrgtnnlyq mksqnpalak hgssilerhh lefgkfilse
361  etlniyqnln rrghehvihl mdiaiatdl alyfkkrmf qkivdeskny gdkkswveyi
421  slettrkeiv mammtacd1 saitkpwvqv skvallvaae fweqgdlerf vldqqp1pmm
481  drnkaaelpk lqvgfidfvc tfvykefsrf heeilpmfdr lqnrkewka ladeyeakvk
541  aleekeeeer vaakkvgtei cnggpapkss tccil (SEQ ID NO:23)

```

图7N

环核苷酸门控阳离子通道 α -3 同种型 1 (CNGA3 同种型 1)

GenBank NP_001289

```

1 makintqysh psrthlkvkt sdrdinraen glsrahsse etssvlqpgi ametrqlads
61 qggsftgggi arlsrlifll rrwaarhvh qdggpdsfpd rfrgaekkev ssqesnaqan
121 vgsqepadrg rsawplakcn tntsnnteee kktkkkdaiv vdpssnlyyr witaialpvi
181 ynwyllicra cfdeiqseyl mlwlvldysa dvlyvldvlv rartgfleqg lmvsdtnrlw
241 qhyktttqfk ldvlslvptd laylkvgtgy pevrfnrllk fsrlfeffdr tetrtnypnm
301 frignlvlyi lllihwnaci yfaiskfigf gtdswvypni sipehgrlsr kyislywst
361 ltlttigetp ppvkdeeylf vvvdflvgvi ifatlvgnvg smisnmnesr aefgakidsi
421 kqymqfrkvt kdletrvlrw fdylwanckt vdekevlksl pdklkaeiai nvhldtlkky
481 rifgdceagi lvelvlklrp tvfspgdyic kkgdligkemy linegklavv addgvtqfvv
541 lsdgsyfgei silnikgsk gnrrtanirs igysdlfcis kddlmealte ypeakkalee
601 kgrqilmkdn lideelarag adpkdleekv eqlgssldtl qtrfarllae ynatqmkmkq
661 rlsqlesqvk gggdkpladg evpgdatkte dkqg (SEQ ID NO:24)

```

图70

环核苷酸门控阳离子通道 α -3 同种型 2 (CNGA3 同种型 2)
GenBank NP_001073347

```

1  makintqysh  psrthlkvkt  sdrdlhraen  glsrahssse  etssvlqpqi  ametrglads
61  gggsftgggi  arlsrlifll  rrwaarhvh  gdqgpdspdp  rfrgaelkev  ssqesnaqan
121 vgsqepadrg  rrkktkkkda  ivvdpssnly  yrwltaialp  vfynwylllc  racfdelqse
181 ylmwlvlvdy  sadvlyvldv  lvrartgfle  qglmvsdtnr  lwqhyktttq  fkldvlsivp
241 tdlaylkgvt  nypevrfrnrl  lkfsrlfeff  drtetrtnyp  nmfrignlvt  yiliiihwna
301 ciyfaiskfi  gfgtdswvyp  nisipehgrl  srkyiyslyw  stltlttge  tpppvkdeey
361 lfvvvdflvg  vlifativgn  vgsmlsmnna  sraefqakid  sikqymqfrk  vtkdlettrvi
421 rwdylwank  ktvdekevlk  slpdklkaei  ainvhldtlk  kvrifqdcea  glivelvlkl
481 rptvfspgdy  ickkgdigke  myiinegkla  vvaddgvtqf  vvlsgdsyfg  eisilnikgs
541 ksgnrirtani  rsigysdlfc  lskddlmeal  teypeakkal  eekgrqilmk  dnlideelar
601 agadpkdlee  kveqlgssld  tlqtrfarll  aeynatqmk  kqrlsqlesq  vkggggdkpla
661 dgevpgdatk  tedkqq (SEQ ID NO:25)

```

图7P

环核苷酸门控阳离子通道 β -3 (CNGB3)
GenBank NP_061971

```

1  mfksltkvnk vkpigenenn eqssrrneeg shpsnqsqqt taceenkgee kslktkstpv
61  tseeptniq dklskknssg dltnpdpqn aaeptgtvpe qkempgkeg pnsqgnkppa
121  apvineyada qlhnlvkrmr qrtalykkkl vegdlsspea spgtakptav ppvkesddkp
181  tehyyrllwf kvkkmpltey lkrikpnsi dsytdrlyll wlllvltayn wnccfiplr1
241  vfpyqtadni hyladiiic diilydmlf iqprlqfvrq gdiivdsnel rkhyrtstkf
301  qldvasiipf dicylffgfn pmfranrmlk ytsffefnhh lesimdkayi yrvirttgyl
361  lfilhinacv yywasnyegi gttrwvydge gneylrcyyw avrtlitigq lpepqtifei
421  vfqllnffsg vfvfssligq mrdvigaata nqnyfracmd dtiaymnys ipklivqkrvr
481  tweyetwdsq rmldesdllk tlptttvqlal aldvnfslis kvdlfkgedt gmiydmlrl1
541  ksvlylpgdf vckkgeigke mylikhgevq vlggpdgtkv lvtlkagsvf geisllaagg
601  gnrrtanvva hqfanlltld kktlqeilvh ypdserilmk karvllkqka ktaeatpprk
661  dlallfpkke etpklfktll ggtgkaslar ilklkregaa qkkensegge eegkenedkq
721  kenedkqken edkgkenedk dkgrepeekp ldrpectasp iaveeeephsv rrtvlprgts
781  rqsliismap saeggeevlt ievkekakq (SEQ ID NO:26)

```

图7Q

鸟嘌呤核苷酸结合蛋白 G(t) 子单元 α -2 (GNAT2)

GenBank NP_005263

```

1  mgsqasaedk elakrskel kklqedadke aktvkl1lllg agesgkstiv kqmkiilhgdg
61  yspceclefk aiiygnvlgq ilaiiramtt lgidyaepsc addgrqinnl adsieegtmp
121 pelvevirrl wkdgqvqacf eraaeyqlnd sasyylngle ritdpeylps eqdvlsrvk
181 ttgiietkfs vkdlnfrnfd vgqqrserkk wihefegvto iifcaalsay dmvlveddev
241 nrmheslhlf nsicnhkffa atsivlflnk kdlfeekikk vhlscfpey dgnsyddag
301 nyiksqqf1d1 nmrkdvkeiy shmtcatdtq nvkfvdavt diikenlkd cglf

```

(SEQ ID NO:27)

图7R

RPGR-815 氨基酸
GenBank NP_000319

```

1 mrepeelmpd sgavftfgks kfaennpgkf wfkndvpvhl scgdehsavv tgnnklymfg
61 snnwqqlglg sksaiskptc vkalkpekvv laacgrnhtl vsteggnyva tggnnegqlg
121 lgdteerntf hvisfftseh kikqlsagsn tsaaltedgr lfmwgdnsseg qiglknsnv
181 cvpqgvtiqk pvswiscgyy hsafvttdge lyvfgepeng klglpnqlig nhrtpqlvse
241 ipekviqvac ggehtvvlte navytfglgq fgqlglgtfl fetsepkvie nirdqtisyl
301 scgenhtali tdiglmtyfg dgrhgklglg lenftnhfip tlcsnflrfi vklvacggch
361 mvvfaaphrq vakeiefdei ndtclsvatf lpyssltsgn vlqrtlarm rrrererspd
421 sfsmrrtlpp iegtlglvac flpnsvfpro sernlqesvl seqdlmqpee pdylldemtk
481 eaeldnsstv eslgettdil nmthimslns neksklspv qkqkkqqtig eltqdtalte
541 nddsdeyeem semkegkack qhvsqgifmt qpattieafs deeveipeek egaedskgng
601 ieegeveane envkvhggrk ekteilsddl tdkaedhefs kteelkleedv deenaenve
661 skkktvgdde svptgyhskt egaertndds saetiekkk anleeraice ynenpkgym
721 ddadsslei lensettpsk dmkktkkifl fkrvpsinqk ivknnneplp eiksigidqii
781 lksdnkdadq nhmsqnhqni pptnterrsk sctil (SEQ ID NO:28)

```

图7S

RPGR -- 646 氨基酸
GenBank CAB54002

```

1  mrepeelmpd sgavftfgks kfaennpgkf wfkndvvpvhl scgdehsavv tgnnklymfg
61  snnwggqlglg sksaiskptc vkalkpekvk laacgrnhtl vsteggnyva tggnnegqlg
121  lgdteerntf hvisftseh kikqlsagsn tsaaltedgr lfmwgdnsseg qiglknavsnv
181  cvpqgvtigk pvswiscgyy hsafvttdge lyvfgepeng klglpnqllg nhrtplvlse
241  ipekviqvac ggehtvvlte navytfglgq fgqlglgtfl fetsepkvie nirdqtisyl
301  scgenhtali tdiglmtytg dgrhghklglg lenftnhfip tlcsnflrfi vklvacggch
361  mvvfaaphrg vakeiefdei ndtclsvatf lpyssltsgn vlqrtlarm rrrererspd
421  sfsmrtrtlpp iegtlglvac flpnsvfprc sernlqesvl seqdlmqpee pdylldemtk
481  eaeidnsstv eslgettdil nmthimslns nekslklspv qkqkkqqtig eltqdtalte
541  nddsdeyeem semkegkack qhvsqgifmt qpattieafs deeveipeek egaedskgng
601  ieeqeveane envkvhggrr ekteillsddl tdkaeysash sqivsv (SEQ ID NO:29)

```

图7T

RPGR-1152 氨基酸

```

1 mrepeelmpd sqavftfgks kfaennpqkf wfkndvpvhl scgdehsavv tgnnklymfq
61 snnwqqlglg sksaiskptc vkalkpekvk laacgrnhtl vsteggnyva tggnnegqlg
121 lgdteerntf hvisfftseh kikqlsagsn tsaaaltedgr lfmwgdnsseg qiglknavsnv
181 cvpqqvttgk pvswiscgyy hsafvttdge lyvfgepeng klglpnqllg nhrtpqvlvse
241 ipekviqvac ggehtvvlte navytfglgq fgqlglgtfl fetsepkvie nirdqtisyl
301 scgenhtali tdiglmtyfg dgrhgklglg lenfthfip tlcsnflrfi vklvacggch
361 mvvfaaphrg vakeiefdei ndtclsvatf lpyssltsgn vlqrtlarm rrrererspd
421 sfsmrtrtlpp iegtlglvac flpnsvfprc sernlgesvl seqdlmqpee pdylldemtk
481 eaeidnsstv eslgettdil nmthimslns neksklspv qkqkqqtig eltqdtalte
541 nddsdeyeem semkegkack qhvsqgifmt qpattieafs deeveipeek egaedskgng
601 ieeqeveane envkvhggrk ekteillsddl tdkaevsegk aksvgeaadg pegrgdgtce
661 egssgaehwq deerekged krgemerpg egekelaeke ewkrdgeeq eqkerekghq
721 kernqemeeg geeehggee eegdreeeee kegegkeeg geevegerek eegerkkeer
781 agkeekgeee gdqgegeeee tegrgeekke ggeveggeve egkgeresee eegegeeseg
841 egeeegege eegegkggee egeegegeee geegegegee eeggegeeee gegegeeseg
901 egeeegegeg egeeegegk geeegegeg egeeegegee gedgegeee eegewegeee
961 egeegeeeee egeegegee geeegegeg geeegegeg eegegeeseg gegegeeseg
1021 vgevegeeg egegeeeeee egeerekeg egeenrrnre eeeeeegkyq etgeeenery
1081 dgeeykkvsk ikgsvkygkh ktyqkksvtn tqngnkeqrs kmpvqskrll knqpsgskkf
1141 wnnvlphyle lk (SEQ ID NO:30)

```

图7U

RPGR-1020 氨基酸

```

1 mrepeelmpd sgavftfgks kfaennpgkf wfknvdpvhl sogdehsavv tgnnklymfg
61 snnwqqiglg sksaiskptc vkalkpekvk laacgrnhtl vsteggnyva tqgnnegqlg
121 lgdteerntf hvisfftseh kikqlsagsn tsaaltedgr lfmwgdnsseg qiglknsnv
181 cvpqqvtigk pswiscgyy hsafvtttddge lyvfgepeng klglpnqlig nhrtppqlvse
241 ipekviqvac ggehtvvlte navytfglgq fgqlglgtfl fetsepkvie nirdqtisyl
301 scgenhtali tdiglmtyfg dgrhghklglg lenftnhfip tlcsnflrfl vklvacqgch
361 mvvfaaphrg vakeiefdel ndtclsvatf lpyssltsgn vlqrtlarm rrrererspd
421 sfsmrzrtlpp iegtlglisac flpnsvfprc sernlqesvl seqdlmqpee pdyllidentk
481 eaeidnsstv eslgettdil nmthimsins nekslklspv qkqkkqqtig eltqddtalte
541 nddsdeyeem semkegkack qhvsqgifmt qpattieafs deevgndtgg vgpqadtgde
601 glqkevyrhe nnngvddqlda keiekesdgg hsqkesaeae idseketkla eiagmkdlre
661 rekstkkmsp ffgnlpdrgm nteseenkdf vkkresckqd vifdseresv ekpdsymega
721 sesqqgiadg fqqpeaiefs sgekeddeve tdqniryrgrk lieqgneket kpiisksmak
781 ydfkcdrlse ipeekegae skngieeeg veaneenvkv hggrkektei lsddltdkae
841 dhefskteel kledvdeen aenveskkkt vqddesvptg yhsktegaez tnddssaeti
901 ekkekanlee raiccynenp kgymlddads ssleilense ttpsksdmkkt kkiFlfkrvp
961 sinqkivknn neplpeiksi gdqiilkndn kdadqnhmsq nhqniptnt errskscitl
(SEQ ID NO:31)

```

图7V

酿脓链球菌 Cas9

```

1 mdkkysigld igtntsvgwav itdeykvpsk kfkvlqntdr hskknliga llfdsetae
61 atrlkrtrr rytrkrnric ylqeifsnem akvddsfhr leesflveed kkherhpifg
121 nivdevayhe kyptiyhlrk klvdstdkad lrliylalah mikfrghfli egdlndnsd
181 vdklfiqlvq tynqlfeenp inasgvdaka llsarlksr rlenliaqlp geknglfgn
241 liaislglt p nksnfdlae dakqlskdt yddldnlla qigdqyadlf laaknlsdai
301 llsdilrvnt eitkaplsas mikrydehmq ditiikalvr qqlpekykei ffdqskngya
361 gyidggasqe efykfikpil ekmdgteeli vknredlir kqrtfdngsi phqihlgelh
421 ailrrqedfy pflkdnreki ekiltfripy yvgplargns rfawmtrkse etitpwnfee
481 vvdkgasaqs fiermtndfk nlpnekvlpk hsllyeyftv yneltkvkyv tegmrkpaf1
541 sgeqkkaivd llfktnrkvt vkqlkedyfk kiecfdsvei sgvedrfnas lgtyhdl1ki
601 ikdkdflone enedilediv ltitlfedre mieerlktya hlfdckvmkq lkrriygtwg
661 rlsrklingi rdkqsgktii dflksdgfan rnfmqlihdd sltfkediqk aqvsgqgds1
721 behianlags paikkgilqt vkvdelvkv mgrhkpeniv iemarenqtt qkgqknsrer
781 mkrieegike lgsqilkehp ventqlqnek lylyylqngr dmyvvdqeldi nrlsdydvdh
841 ivpqsflkdd sidnkvlt rs dknrgksdnv pseevkvkmk nywrqllnak litqrkfdnl
901 tkaergglse ldkagfikrq lvetrqitkh vaqildsrnn tkydendkli revkvitlks
961 klvsdfrkdf qfykvrein yhhahdayln avvgtalikk ypklesefvy gdykvdydvk
1021 miakseqeig katakyffys nimnffktei tlangeirkr plietngetg eivwdkgrdf
1081 atvrkvlsm p qvniivkktev qtggfskesi lpkrnskli arkkdwdpkk yggfdsptva
1141 ysvlvvakve kgkskklksv kelligitime rssfeknpid fleakykev kkdliiklpk
1201 yslfelengr kmrlasagel qkgnelalps kymnflylas hyeklkgspe dneqqlfve
1261 qhkhyldeii eqisefskrv iladanldkv lsaynkhrdk pireqaenii hlftltnlga
1321 paafkyfdtt idrkrytstk evldatling sitglyetri dlsq1ggd (SEQ ID NO:32)

```

图8A

金黄色葡萄球菌 Cas9

```

1 mkrnyilgld igitsvgysi idyetrdivd agvrlfkean vennegrsk rgarrlkrir
61 rhriqrvkkl lfdynlltdh selsginpye arvkglsqkl seeefsaall hlakrrgvhn
121 vneveedtgn elstkeqisr nskaleekyv aelqlerlkk dgevrgsinr fktstdyvkea
181 kqllkvvqkay hqldqsfidt yidlletirt yyegpgegsp fgwkdkewy enlmghctyf
241 peelrsvkya ynadlynaln dlennlvitrd enekleyyek fqienvfkk kkkptlkqia
301 keillvneedi kgyrvtstgk peftnlkvvh dikditarke iienaelldq iakiltiyqs
361 sediqaeltl inseltqeei eqisnlkgyt gthnlslkai nlildelwht ndnqiaifnr
421 lklvpkkvdl sqkceipttl vddfilsppv krsfiqsikv inaiikkygl pndiilelar
481 eknskdaqkm inemqkrnrq tnerieeirr ttgkenakyl iekiklhdmq egcolyslea
541 ipledllnnp fnyevdhiip rsvsfdnsfn nkvlvkqeen skkgnrtpfq ylsssdskis
601 yetfkkhiln lakggrisk tkkeylleer dinrfsvqkd finrnlvdtr yatrglmnll
661 rsyfrvnnld vkksinggf tsflrrkwkf kkerknkykh haedalilian adfifkewkk
721 ldkakkvmen qmfeekqaes mpeieteqey keifitphqi khikdfkdyk yshrvdkkpn
781 relindtlys trkddkgntl ivnnlnglyd kdndklkkl i nkspeklmy hhdpgtyqkl
841 klimeqygde knplykyvee tgnyltkysk kdngpvikki kyygnkl nah lditddypns
901 rnkvvklslk pyrfdvyl dn gvykfvtvkn ldvikkenny evnskoyeea kklkkisnqa
961 efiasfyand likingelyr vigvnnldln rievnmidit yreylennnd krppriikti
1021 asktqsikky stdilgnlye vkskhhpqi i kkg (SEQ ID NO:33)

```

图8B

土拉热弗朗西丝菌 Cpf1

```

1  mslyqefvnk yslsktlrfe lipqgktlen ikarglildd ekrakdykka kqlldkyhqf
61  fieceilssvc isedllqnys dvfklkkksd ddnllqkdfks akdtikkkqis eyikdsekfk
121 nlfngnlida kkgqesdlil wlqskdngi elfkansdit didealeiik sfkgwttyfk
181 gfhenrknvy ssndiptsli yriivddnlpk flenkakyes lkdkapeain yeqikkdlae
241 eltfdidykt sevngrvfsl devfeianfn nylngsgitk fntliiggkfv ngentkrkgi
301 neyinlysqq indktlkkyk msvlfkqils dtesksfvid kledsdsvvt tmqsfyeqia
361 afktveeksi ketlslllfd lkaqklldsk iyfkndkslt dlsqqvfddy svigtavley
421 itqqiapknl dnpskkeqel iakktekaky isletiklal eefnkhrdid kqcrfeeila
481 nfaaipmifd eiaqnkdnla qisikyqngg kkdllqasae ddvkaikdli dqtlnllhkl
541 kifhisqsed kanlldkdeh fylvfeecyf elanivplyn kirnyitqkp ysdekfklnf
601 enstlangwd knkepdtai lfikddkyyi gvmnknkni fddkaikenk gegykkivyk
661 llpgankmlp kvffsaksik fynpsedilr irnhsthtkn gspqkgyekf efniedcrkf
721 idfykqsisk hpewkdfgr fsdtqrynsi defyrevenq gykltfenis esyidsvvng
781 gklylfqiyn kdfsayskgr pnhtlywka lfdernlqdv vyklngaeal fyrcqslpkk
841 ithpakeaia nknkdnpkke svfeydlikd krftedkfff hcpittinfks sgankfndei
901 nlllikekand vhlslidrge rhlayytlvd gkgnlikqdt fnliqndrmk tnyhdklaai
961 ekdrdsarkd wkinnikem kegylsqvvh eiaklvieyn aivvfedlnf gfkrgfrkve
1021 kqvyqklekm lieklnylvf kdnefdktgg vlrayqltap fetfkkmgkq tgliyyvpag
1081 ftskicpvtg fvnglypkye svksqeffs kfdkicynld kgyfefsfdy knfgdkaakg
1141 kwtiassfgr linfrnsdkn hnwdtrevyp tkelekllkd ysieyghgec ikaaiggesd
1201 kkffakltsv lntilqmns ktgteldyli spvadvnqnf fdsrqapknn pqdadangay
1261 higlklml1 griknqgegk klnlviknee yfefvqnrnn (SEQ ID NO:34)

```

图8C