

衣壳变异的重组腺相关病毒及其应用

本申请要求 2021 年 05 月 28 日提交的申请号为 202110594986.X 的中国专利申请的优先权。

5

技术领域

本公开涉及重组腺相关病毒(rAAV)技术领域,具体涉及一种衣壳变异的 rAAV 病毒粒子及其向靶细胞(例如,视网膜细胞)递送基因产物的应用。

10 背景技术

腺相关病毒(AAV)属于细小病毒科(Parvoviridae)的依赖病毒属(Dependovirus),是小的(25nm)、无包膜的单链 DNA 病毒,核酸被二十面体衣壳(cap)所包裹。AAV 含有两个开放阅读框 rep 和 cap,rep 用于编码基因组复制所必需的四个蛋白(Rep78、Rep68、Rep52、和 Rep40),cap 用于编码病毒衣壳装配所需的三个结构蛋白(VP1-3)。

15 AAV 作为载体,表现出对于遗传病、基因病的治疗的巨大潜力,可用于基因补充疗法(也称为基因增强疗法)。其通过补足缺失的基因功能来恢复由突变导致的基因功能丧失或失调,从而使目标细胞的生物学功能恢复到正常的生理状态。已有研究显示,基于 AAV 的载体在临床前疾病模型及在人临床试验中的表现,显示出若干疾病治疗的应用的前景,例如,可以实现将基因高效转移到视网膜细胞中并长期持续地进行表达(Boye et al. Mol Ther.2013 Mar;21(3): 509-19. Trapani et al. Prog Retin Eye Res. 2014 Nov;43: 108-28.)。目前用于临床的 AAV 载体经过基因工程改造,无辅助病毒时,呈潜伏状态,其安全性和转基因的长期表达已在啮齿类动物模型、非人灵长类动物和多项人类试验中被广泛测试(MacLaren et al. Lancet.2014 Mar 29;383(9923): 1129-37.; Maguire et al. N Engl J Med. 2008 May 22;358(21): 2240-8; Simonelli et al. Mol Ther. 2010 Mar;18(3): 643-50; Nathwani et al. N Engl J Med. 2014 Nov 20;371(21): 1994-2004)。

AAV 衣壳蛋白在组成和结构上有不同的天然存在,不同的衣壳具有不同的组织嗜性。已经鉴别出了多个同源的灵长类 AAV 血清型和非人灵长类 AAV 血清型,开发了不同的工程化改造的 AAV 变体(也称 AAV 血清型)。虽然 AAV 载体在不同生物体和组织中均表现出一定程度的扩散感染,但多数临床试验中均采用病灶区局部注射的方式,尤其是在眼部基因治疗的应用。目前主要的眼部临床实施方式为视网膜下腔注射(即,在 RPE 和感光之间注射液体后形成的空腔)和玻璃体腔注射。视网膜下腔注射使得 AAV 对 RPE 和感光细胞得到充分接触,局部感染效果较好,但相对注射操作风险较高,易造成视网膜脱落。玻璃体腔注射,AAV 制剂会首先均匀分布分布于玻璃体液中而后逐层对视网膜进行扩散感染。由于视网膜层致密的结构和复杂的细胞种群构成,天然血清型中对眼部浸染能力较强的 AAV8、AAV2,在玻璃体腔注射时效果较差,而部分工程改造的衣壳却在实施玻璃体腔注

射时显露出较强的感染能力(如 AAVDJ、AAV2.7M8 等)。

本领域仍然需要新型 AAV 变体。本公开提供了这样的重组腺相关病毒(rAAV)病毒粒子,其具有新结构的衣壳蛋白,对于视网膜组织具有高感染噬性,对异源基因具有高表达效率,从而为临床治疗提供了更有潜力的方案。

5

发明内容

本公开提供变体 AAV 衣壳蛋白及其可携带的基因产物,含有所述衣壳蛋白的 rAAV 病毒粒子,药物组合物,所述 rAAV 病毒粒子侵染细胞(例如,视网膜细胞)、治疗和预防疾病(例如,眼部疾病)的方法和制药用途。

10

变体腺相关病毒(AAV)衣壳蛋白

本公开提供一种腺相关病毒(AAV)衣壳蛋白,其相对于亲本 AAV 衣壳蛋白包含插入的多肽,所述插入的多肽包含选自如下 1)-6)任一项所示的多肽或其任意组合:

- 1) LAETTRP(SEQ ID NO: 11)或由 SEQ ID NO: 11 组成的多肽;
- 15 2) LGDTTRP(SEQ ID NO: 12)或由 SEQ ID NO: 12 组成的多肽;
- 3) LGETTRN(SEQ ID NO: 13)或由 SEQ ID NO: 13 组成的多肽;
- 4) KADTTKN(SEQ ID NO: 14)或由 SEQ ID NO: 14 组成的多肽;
- 5) KDDTTRN(SEQ ID NO: 15)或由 SEQ ID NO: 15 组成的多肽;
- 6) LADTTKN(SEQ ID NO: 16)或由 SEQ ID NO: 16 组成的多肽。

- 20 本公开提供一种腺相关病毒(AAV)衣壳蛋白,其相对于亲本 AAV 衣壳蛋白包含插入的多肽,所述多肽包含 $X_1X_2X_3TTX_4X_5$ (SEQ ID NO: 35)或由 SEQ ID NO: 35 组成,其中, X_1 选自 L 或 K, X_2 选自 G、D 或 A, X_3 选自 D 或 E, X_4 选自 R 或 K, X_5 选自 P 或 N。

- 一些实施方案中,前述 1)-6)或 SEQ ID NO: 35 所示的多肽在其氨基末端和/或羧基末端具有 1-4 个间隔氨基酸(Y_1 - Y_4)。一些具体实施方案中,所述间隔氨基酸包括但不限于 A、L、G、S 和 T。

- 30 本公开提供一种腺相关病毒(AAV)衣壳蛋白,其相对于亲本 AAV 衣壳蛋白包含插入的多肽,所述多肽包含 $Y_1Y_2X_1X_2X_3TTX_4X_5Y_3Y_4$ (SEQ ID NO: 36)或由 SEQ ID NO: 36 组成,其中, X_1 选自 L 或 K, X_2 选自 G、D 或 A, X_3 选自 D 或 E, X_4 选自 R 或 K, X_5 选自 P 或 N; Y_1 、 Y_2 、 Y_3 、 Y_4 可以独立的存在或不存在, Y_1 、 Y_2 、 Y_3 、 Y_4 可以独立的选自 A、L、G、S 和 T。例如, Y_1 为 L, Y_2 为 A, Y_3 为 A, Y_4 不存在。

- 35 本公开提供一种腺相关病毒(AAV)衣壳蛋白,其相对于亲本 AAV 衣壳蛋白包含插入的多肽,所述插入的多肽选自如下 1-1)至 1-6)任一项所示的多肽或其任意组合:

1-1) 包含 LALAETTRPA(SEQ ID NO: 17)或由 SEQ ID NO: 17 组成的多肽;

- 2-1) 包含 LALGDTTRPA(SEQ ID NO: 18)或由 SEQ ID NO: 18 组成的多肽;
3-1) 包含 LALGETTRNA(SEQ ID NO: 19)或由 SEQ ID NO: 19 组成的多肽;
4-1) 包含 LAKADTTKNA(SEQ ID NO: 20)或由 SEQ ID NO: 20 组成的多肽;
5-1) 包含 LAKDDTTRNA(SEQ ID NO: 21)或由 SEQ ID NO: 21 组成的多肽;
5 6-1) 包含 LALADTTKNA(SEQ ID NO: 22)或由 SEQ ID NO: 22 组成的多肽。

一些实施方案中,上述 AAV 衣壳蛋白为 AAV2 衣壳蛋白,或 AAV9 衣壳蛋白。

- 一些实施方案中,前述 1)至 6)、1-1)至 6-1)、SEQ ID NO: 35 或 36 的多肽位于亲本 AAV 衣壳蛋白的 GH 环或 IV 环(环状结构域 IV)中,例如在 AAV 衣壳蛋白的 GH 环或 IV 环的溶剂可及部分(参见 van Vliet 等(2006)*Mol.Ther.*14: 809; Padron
10 等(2005)*J.Virol.*79: 5047; 和 Shen 等(2007)*Mol.Ther.*15: 1955)。“亲本 AAV 衣壳蛋白”指无插入的多肽的相同 AAV 血清型的衣壳蛋白(包括野生型 AAV 血清型或其变体衣壳蛋白,例如本公开的 SEQ ID NO: 1 所示的 AAV2 衣壳蛋白或 SEQ ID NO: 37 所示的 AAV9 衣壳蛋白,所述 AAV2 衣壳蛋白是可以具有或不具有 V708I 突变的)。

- 15 一些实施方案中,所述 AAV 选自 AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10(包括 AAVrh10)。

- 一些实施方案中,前述插入的多肽位于亲本 AAV 衣壳蛋白的氨基酸残基 411 位至 650 位内、或 432 位至 640 位内、或 570 位至 671 位内、或 570 位至 614 位内、或 570 位至 610 位内、或 580 位至 600 位内、或 570 位至 575 内、或 575 位
20 至 580 内、或 580 位至 585 内、或 585 位至 590 内、或 590 位至 600 内、或 600 位至 614 内。例如,位于亲本 AAV2 衣壳蛋白的氨基酸残基 570 位至 611 位内、亲本 AAV1 衣壳蛋白的氨基酸残基 571 位至 612 位内、亲本 AAV5 衣壳蛋白的氨基酸残基 560 位至 601 位内、亲本 AAV6 衣壳蛋白的氨基酸 571 位至 612 位内、亲本 AAV7 衣壳蛋白的氨基酸残基 572 位至 613 位内、亲本 AAV8 衣壳蛋白的氨基酸残基 573 位至 614 位内、亲本 AAV9 衣壳蛋白的氨基酸残基 571 位至 612 位
25 内,或亲本 AAV10(包括 AAVrh10)衣壳蛋白的氨基酸残基 573 位至 614 位内。

- 一些实施方案中,前述插入的多肽位于亲本 AAV2 衣壳蛋白的氨基酸残基 587 位和 588 位之间,或亲本 AAV9 衣壳蛋白的氨基酸残基 588 位和 589 位之间,或其他亲本 AAV 血清型的衣壳蛋白的对应位置。一些实施方案中,插入的多肽位于
30 亲本 AAV2 衣壳蛋白的氨基酸残基 587 位和 588 位之间,或亲本 AAV9 衣壳蛋白的氨基酸残基 588 位和 589 位之间,或其他亲本 AAV 血清型的衣壳蛋白的对应位置。其他血清型例如选自 AAV1、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8 或 AAV10(包括 AAVrh10)。不同 AAV 血清型中与 AAV2 的衣壳蛋白 VP1 的氨基酸 570-611 相对应的序列是本领域公知的(参见例如 WO2012145601A 的图 6,以及,
35 GenBank 登记号 NP_049542 的 AAV1; AAD13756 的 AAV5; AAB95459 的 AAV6; YP_077178 的 AAV7; YP_077180 的 AAV8; AAS99264 的 AAV9 和 AAT46337 的

AAV10)。一些实施方案中，插入的多肽位于亲本 AAV1 衣壳蛋白的氨基酸残基 590 位和 591 位之间、亲本 AAV5 衣壳蛋白的氨基酸残基 575 位和 576 位之间、亲本 AAV6 衣壳蛋白的氨基酸残基 590 位和 591 位之间、亲本 AAV7 衣壳蛋白的氨基酸残基 589 位和 590 位之间、亲本 AAV8 衣壳蛋白的氨基酸残基 590 位和 591 位之间、亲本 AAV10(包括 AAVrh10)衣壳蛋白的氨基酸残基 588 位和 589 位之间。

本公开的氨基酸残基数是相对于 AAV 衣壳蛋白的 VP1 的编码氨基酸序列从 N 端的自然计数。例如“插入的多肽位于 AAV2 衣壳蛋白的氨基酸残基 587 位和 588 位之间”是指所述多肽位于 AAV2 衣壳蛋白的 VP1 的编码氨基酸序列的第 587 位至 588 位之间，其在 VP2 的编码氨基酸序列中对应的是第 450 位至 451 位之间，在 VP3 的编码氨基酸序列中对应的是第 385 位至 386 位之间。

一些实施方案中，前述插入的多肽位于亲本 AAV 衣壳蛋白的氨基酸残基 450 位和 460 位之间，例如亲本 AAV2 的氨基酸残基 453 位、亲本 AAV1 的氨基酸残基 454 位、亲本 AAV6 的氨基酸残基 454 位、亲本 AAV7 的氨基酸残基 456 位、亲本 AAV8 的氨基酸残基 456 位、亲本 AAV9 的氨基酸残基 454 位、亲本 AAV10(包括 AAVrh10)的氨基酸残基 456 位之前或之后。这里全文引入 WO2012145601A 的图 17 所展示的不同 AAV 血清型的衣壳蛋白的氨基酸残基及其对应关系。

一些实施方案中，本公开提供变体 AAV 衣壳蛋白，其含有前述 1)至 6)、1-1)至 6-1)、SEQ ID NO: 35 或 36 的多肽。

一些实施方案中，前述本公开的变体 AAV 衣壳蛋白还包含一个或多个氨基酸残基的点突变(包括取代、缺失和/或添加)。

一些实施方案中，所述氨基酸残基的点突变位于第 1、15、34、57、66、81、101、109、144、164、176、188、196、226、236、240、250、312、363、368、449、456、463、472、484、524、535、551、593、698、708、719、721 和 735 位的一个或其任意组合。

一些实施方案中，所述氨基酸残基的点突变(取代)选自 1L、15P、34A、57D、66K、81Q、101R、109T、144K 或 M、164K、176P、188I、196Y、226E、236V、240T、250S、312K、363L、368H、449D、456K、463Y、472N、484C、524T、535S、551S、593E、698V、708I、719M、721L 和 735Q 的一个或其任意组合，例如 312K、449D、472N、551S、698V、735Q、273F、444F、500F、730F、708I 的一个或多个，例如 708I。

一些实施方案中，所述氨基酸残基的点突变(取代)选自 M1L、L15P、P34A、N57D、N66K、R81Q、Q101R、S109T、R144K、R144M、Q164K、T176P、L188I、S196Y、G226E、G236V、I240T、P250S、N312K、P363L、D368H、N449D、T456K、S463Y、D472N、R484C、A524T、P535S、N551S、A593E、I698V、V708I、V719M、S721L、L735Q、Y273F、Y444F、Y500F、Y730F 的一个或其任意组合，例如 N312K、N449D、D472N、N551S、I698V、L735Q、Y273F、Y444F、Y500F、Y730F、V708I

的一个或多个。

一些具体实施方案中，所述氨基酸残基的点突变(取代)为 708I 和/或 449D，或 V708I 和/或 N449D。一些具体实施方案中，所述氨基酸残基的突变(取代)为 273F、444F、500F 和/或 730F，或 Y273F、Y444F、Y500F 和/或 Y730F。

- 5 上述点突变是相对于相应亲本 AAV 衣壳蛋白的相应位置的，例如相对于亲本 AAV2 衣壳蛋白的相应位置的。

本公开全文引入 WO2012145601A、WO2017197355A、WO2018022905A 中衣壳蛋白的突变。

- 10 一些实施方案中，本公开的 AAV 衣壳是嵌合衣壳。例如，衣壳包括第一 AAV 血清型 AAV 衣壳的一部分和第二血清型 AAV 衣壳的一部分，所述血清型包括但不限于 AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10(包括 AAVrh10)。例如，所述 AAV 衣壳可以是 AAV2G9，其包含来自 AAV2 和 AAV9 的序列，此处全文引入 US20160017005 中 AAV2G9 的序列。

一些实施方案中，本公开的变体 AAV 衣壳蛋白是经分离的，和/或经纯化的。

- 15 一些实施方案中，本公开提供变体 AAV2 衣壳蛋白，相对于相应亲本 AAV2 衣壳蛋白(例如 SEQ ID NO: 1 所示)，在衣壳蛋白 GH 环或 IV 环(环状结构域 IV)中包含多肽，所述多肽选自前述 1)至 6)、1-1)至 6-1)、SEQ ID NO: 35 或 36。一些具体实施方案中，多肽包含或为 SEQ ID NO: 12 或 18。一些具体实施方案中，多肽位于亲本 AAV2 衣壳蛋白的 VP1 的第 587 位和 588 位氨基酸残基之间。

- 20 一些实施方案中，本公开提供变体 AAV9 衣壳蛋白，相对于相应亲本 AAV9 衣壳蛋白(例如 SEQ ID NO:37 所示)，在衣壳蛋白 GH 环或 IV 环(环状结构域 IV)中包含多肽，所述多肽选自前述 1)至 6)、1-1)至 6-1)、SEQ ID NO: 35 或 36。一些具体实施方案中，多肽包含或为 SEQ ID NO: 12 或 18。一些具体实施方案中，多肽位于亲本 AAV9 衣壳蛋白的 VP1 的第 588 位和 589 位氨基酸残基之间。

- 25 一些实施方案中，本公开提供变体 AAV 衣壳蛋白，其氨基酸序列为 SEQ ID NO: 3-9、23-27 任一所示或与之具有至少 85%、至少 90%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%的序列同一性。

一些实施方案中，本公开提供的变体 AAV 衣壳蛋白具有如下的特性：

- 30 (a) 与包括对应的亲本 AAV 衣壳蛋白的 AAV 病毒粒子的眼组织(例如，视网膜细胞)的感染性相比，增加的眼组织(例如，视网膜细胞)的感染性，特别是增加的视网膜神经细胞层感染性；

(b) 与包括对应的亲本 AAV 衣壳蛋白的 AAV 病毒粒子的趋向性相比，改变的细胞趋向性；

- 35 (c) 与包括对应的亲本 AAV 衣壳蛋白的 AAV 病毒粒子相比，增加的结合和/或穿过内界膜(ILM)的能力；和/或

(d) 与包括对应的亲本 AAV 衣壳蛋白的 AAV 病毒粒子相比，在眼组织(例如，

视网膜组织、房水、玻璃体)中具有对其包裹或携带的基因产物具有增加的表达量。

重组腺相关病毒(rAAV)病毒粒子

本公开提供重组腺相关病毒(rAAV)病毒粒子, 其包含:

(a) 前述本公开任意的变体 AAV 衣壳蛋白; 可选地, 包含

5 (b) 异源多核苷酸。

一些实施方案中, 所述异源多核苷酸包含表达或编码基因产物的多核苷酸。

一些实施方案中, 所述基因产物是相对于 AAV 是异源的。一些实施方案中, 所述基因产物是相对于靶细胞是异源的或内源的。一些实施方案中, 所述基因产物为一个或多个(例如, 2、3、4 个)。

10 一些实施方案中, 所述异源多核苷酸中包含调控序列, 所述调控序列调控基因产物的表达或编码。

关于基因产物:

一些实施方案中, 所述基因产物是针对疾病、病症治疗性的或预防性的。

一些实施方案中, 所述基因产物选自干扰 RNA(RNAi)、适配体、多肽。

15 一些具体实施方案中, 所述基因产物是 RNAi, 例如减少、降低细胞中凋亡因子或血管生成因子的水平的 RNAi。例如, RNAi 可以是降低细胞中诱导或促凋亡的基因产物的水平的 shRNA 或 siRNA, 所述促凋亡基因产物包含例如 Bax、Bid、Bak 和 Bad 基因产物(参见 US7,846,730, 并全文引入)。又例如, RNAi 可以是针对血管生成产物的, 例如 VEGF (例如 Cand5, 参见 US2011/0143400、US2008/0188437, 并全文引入)、VEGFR1(例如 Sirna-027, 参见 Kaiser 等(2010)Am.J.Ophthalmol.150: 33; 和 Shen 等(2006)Gene Ther.13: 225, 并全文引入)或 VEGFR2(例如 Kou 等(2005)Biochem.44: 15064, 并全文引入)。

一些具体实施方案中, 所述基因产物是适配体, 例如是针对 VEGF 的特异性适配体(例如 5'-cgcaucagugaugcuuaucauccg-3', 参见 Ng 等(2006)Nat.Rev.Drug 25 Discovery5: 123, 和 Lee 等(2005)Proc.Natl.Acad.Sci.USA 102: 18902, 并全文引入)、针对 PDGF 的特异性适配体(例如 E10030, 参见 Ni 和 Hui(2009)Ophthalmologica 223: 401, 和 Akiyama 等(2006)J.CellPhysiol.207: 407, 并全文引入)。

一些具体实施方案中, 所述基因产物是多肽。

30 一些具体实施方案中, 所述基因产物是神经保护多肽、抗血管生成多肽或增强视网膜细胞功能的多肽。

一些具体实施方案中, 所述多肽可以增强视网膜细胞的功能, 例如, 增强视杆或视锥感光细胞、视网膜神经节细胞、Muller 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。

一些具体实施方案中, 所述多肽包含或选自: 神经保护多肽(例如, GDNF、CNTF、NT4、NGF 和 NTN); 抗血管生成多肽(例如, 可溶性血管内皮生长因子(VEGF) 35 受体、抗 VEGF 抗体或其抗原结合片段、内皮抑素(endostatin)、肿瘤抑素(tumstatin)、

血管抑素(angiotatin)、可溶性 Flt 多肽及其与 Fc 区的融合蛋白(参见 Lai 等(2005)Mol.Ther.12: 659, Pechan 等(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剂-3(TIMP-3);光响应性视蛋白(例如视紫红质);抗凋亡多肽(例如 Bcl-2、Bcl-Xl)等。

5 一些具体实施方案中,所述多肽包括但不限于:表皮生长因子、视紫红质、X 连锁凋亡抑制蛋白。

一些具体实施方案中,所述多肽包括但不限于:视网膜劈裂蛋白(retinoschisin)、视网膜色素变性 GTP 酶调节剂(RGPR)相互作用蛋白-1(GenBank Accession NO. Q96KN7、Q9EPQ2、Q9GLM3)、外周蛋白-2(Prph2)(GenBank Accession NO. NP_000313)、视网膜色素上皮特异性蛋白(RPE65)(GenBank Accession NO. AAC39660)。

一些具体实施方案中,所述多肽包括但不限于:缺损或缺失时,引发无脉络膜的多肽,如 CHM(脉络膜缺损(choroideremia)(Rab 护送蛋白 1))(Donnelly 等(1994)Hum.Mol.Genet.3: 1017);缺损或缺失时,引发利伯氏先天性黑内障的多肽和色素性视网膜炎的多肽,如碎屑同源物 1(CRB1)(GenBank Accession NO. CAM23328);缺损或缺失时,引发色盲的多肽,如杆状感光 cGMP-门控通道亚基 α (CNGA3)(GenBank Accession NO. NP_001289)、杆状感光 cGMP-门控通道 β 亚基(CNGB3)、鸟嘌呤核苷酸结合蛋白(G 蛋白)、 α 转导活性多肽 2(GNAT2)(ACHM4)、ACHM5、L-视蛋白、M-视蛋白和 S-视蛋白。

20 一些具体实施方案中,所述基因产物提供用于对基因功能进行位点特异性敲减的位点特异性内切核酸酶,例如,其中内切核酸酶敲除了与视网膜疾病相关的等位基因。例如,在显性等位基因对当属于野生型时是视网膜结构蛋白和/或提供正常的视网膜功能的基因的缺陷副本进行编码的情况下,位点特异性内切核酸酶可以靶向到缺陷等位基因并且敲除缺陷等位基因。所述位点特异性内切核酸酶例如为:锌指核酸酶(ZFNs),以及转录激活因子样效应物核酸酶(TALENs),其中这种位点特异性内切核酸酶是非天然存在的并且被修饰以靶向特定基因。

此外,本公开全文引入 WO2012145601A、WO2017197355A、WO2018022905A 中的前述基因产物的序列、来源。

30 一些具体实施方案中,本公开的基因产物为抗血管生成剂,包括抗血管生成多肽,例如,抗 VEGF 抗体或其抗原结合片段;又例如 VEGF 拮抗剂(例如 VEGF-A、B、C 拮抗剂)或 PDGF 拮抗剂。

一些具体实施方案中,VEGF 拮抗剂包括但不限于兰尼单抗(ranibizumab)、贝伐单抗(bevacizumab)、阿柏西普(aflibercept)、KH902 VEGF 受体-Fc 融合蛋白、2C3 抗体、ORA102、哌加他尼钠(pegaptanib)、贝伐西尼(bevasiranib)、SIRNA-027、紫花前胡素(decursin)、紫花前胡醇(decursinol)、苦鬼臼脂素(picropodophyllin)、没药甾酮(guggulsterone)、PLG101、类二十烷酸 LXA4、PTK787、帕唑帕尼(pazopanib)、

阿西替尼(axitinib)、CDDO-Me、CDDO-Imm、紫草素(shikonin)、 β -羟异戊酰紫草素(beta-hydroxyisovalerylshikonin)、或神经节苷脂 GM3、DC101 抗体、Mab25 抗体、Mab73 抗体、4A5 抗体、4E10 抗体、5F12 抗体、VA01 抗体、BL2 抗体、VEGF 相关蛋白、sFLT01、sFLT02、肽 B3、TG100801、索拉非尼(sorafenib)、舒尼替尼(sunitinib)、G6-31 抗体, 或其药学上可接受的盐。此处全文引入 WO2018160686A 中的 VEGF 拮抗剂信息。

兰尼单抗(Lucentis®)的序列信息参见 US7,060,269(其图 1), 贝伐单抗(Avastin®)的序列信息参见 US6,054,297 (其图 1), 阿柏西普的(Eylea®)的序列信息参见 Do 等人(Br J Ophthalmol.2009,93: 144-9), 此处通过全文引用的方式并入。

一些实施方案中, VEGF 拮抗剂包含或为天然存在的蛋白质 sFlt-1, 或其功能片段(例如, sFlt-1 结构域 2, 参见 US5,861,484 中的 sFlt-1 序列信息、US2013/0323302 中的 sFlt-1 结构域 2 序列信息, 全文引入)。

一些实施方案中, VEGF 拮抗剂为 VEGF 结合融合蛋白, 此处全文引入 US7,635,474 中结合 VEGF 的融合蛋白的序列信息。

一些具体实施方案中, 阿柏西普(afibercept)的氨基酸序列如 SEQ ID NO: 38 所示, 并提供编码 SEQ ID NO: 38 的多核苷酸序列, 例如, 经密码子优化的多核苷酸序列, 其如 SEQ ID NO: 39-41 任一所示。

关于调控序列:

一些实施方案中, 前述多核苷酸包含如下(a)-(h)的任一项的多核苷酸(调控序列)或任意组合:

(a) 5'反向末端重复(5'ITR)和/或 3'反向末端重复(3'ITR);

(b) 5'非翻译区(5'UTR)和/或 3'非翻译区(3'UTR);

(c) 启动子;

(d) 增强子;

(e) 内含子(intron);

(f) 转录后调控元件;

(g) 多聚腺苷酸化信号(polyA);

(h) Kozak 序列。

一些实施方案中, AAV ITR 不需要具有野生型核苷酸序列并且可以通过核苷酸的插入、删除或取代来改变, 或者 AAV ITR 可以从几种 AAV 血清型中的任何一种得到, 例如 AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9 和 AAV10。一些具体实施方案中, 5'ITR 和 3'ITR 为 AAV2 的 5'ITR 和 3'ITR。

一些实施方案中, 编码基因产物的核苷酸序列与组织特异性或细胞类型特异性调控元件可操作地连接, 例如感光特异性调控元件(例如, 感光特异性启动子)可操作地连接, 又例如赋予感光细胞内可操作地连接的基因选择性表达的调控元件。

一些具体实施方案中，(a)-(h)的任意组合能够满足使得基因产物(例如，抗血管生成剂(或抗血管生成多肽)，又例如，抗 VEGF 抗体或其抗原结合片段、阿柏西普)表达的功能，例如在受试者眼组织(例如房水、视网膜组织)表达的功能。

一些具体实施方案中，所述(a)-(h)的任一项的多核苷酸或任意组合与前述编码
5 基因产物的多核苷酸可操作地连接。

一些具体实施方案中，所述 5'ITR 和/或 3'ITR 源自 AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV9.47、AAV9(hu14)、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAV-DJ 或 AAV-DJ8；例如，源自 AAV2、AAV9。

10 一些具体实施方案中，所述 5'UTR 和/或 3'UTR 源自血红蛋白结合蛋白(hemopexin, HPX)、 β 珠蛋白(hemoglobin subunit beta, HBB)、HSPB1、CCL13、爪蟾球蛋白(Xenopus globin)。例如，5'UTR 的序列如 SEQ ID NO: 30 所示，3'UTR 的序列如 SEQ ID NO: 32 所示。

一些具体实施方案中，所述启动子可以是组成型启动子或诱导型启动子。

15 一些具体实施方案中，所述启动子选自巨细胞病毒(CMV)启动子、劳斯氏肉瘤病毒(RSV)启动子、UB6 启动子、鸡 β -肌动蛋白启动子、CAG 启动子、RPE65 启动子、CBh 启动子、EFS 启动子、EF1(例如 EF-1 α)启动子、PGK 启动子、SV40 启动子、Ubi 启动子、视蛋白启动子或其任意组合。所述视蛋白启动子包括但不限于视紫红质启动子、视紫红质激酶启动子(Young 等(2003)Ophthalmol. Vis. Sci. 44: 4076)、 β 磷酸二酯酶基因启动子(Nicoud 等(2007)J. Gene Med. 9: 1015)、色素性视网膜炎基因启动子(Nicoud 等(2007)同上)、感光间维甲酸结合蛋白(IRBP)基因启动子(Yokoyama 等(1992)Exp Eye Res. 55: 225)。例如，启动子为 CMV 启动子，其序列例如 SEQ ID NO: 29 所示。

20 一些具体实施方案中，所述增强子选自 Ubi、CMV、RSV、IRBP 基因增强子(Nicoud 等(2007) J. Gene Med. 9: 1015)或其任意组合。例如，增强子为 CMV 增强子，其序列例如 SEQ ID NO: 28 所示。

一些具体实施方案中，所述内含子选自 MVM、SV40、 β Globin、EF1(例如 EF-1 α)、hybrid 内含子，或其任意组合。

30 一些具体实施方案中，所述 polyA 选自 PA75 polyA、SV40 polyA、hGH polyA、BGH polyA、rbGlob polyA 或其任意组合。例如，polyA 为 SV40 polyA，其序列例如 SEQ ID NO: 34 所示。

一些具体实施方案中，所述转录后调控元件选自 WPRE、HPRE 或其组合。例如，转录后调控元件为 WPRE，其序列例如 SEQ ID NO: 33 所示。

35 一些实施方案中，前述多核苷酸包含如下(a)-(g)的任一项的多核苷酸(调控序列)或任意组合：

(a) 源自 AAV2 的 5'ITR 和/或 3'ITR；

(b) 源自 *Xenopus globin* 的 5'UTR 和/或 3'UTR;

(c) CMV 启动子;

(d) CMV 增强子;

(e) WPRE;

5 (f) Kozak 序列;

(g) SV40 polyA。

示例性的, 多核苷酸从 5'端和 3'端的顺序为(括号内为可选的调控序列):

5'ITR- 增强子 - 启动子 -5'UTR-(Kozak 序列)- 基因产物
-3'UTR-(WPRE)-polyA-3'ITR, 或

10 增强子-启动子-5'UTR-(Kozak 序列)-基因产物-3'UTR-(WPRE)-polyA;

启动子-(5'UTR)-(Kozak 序列)-基因产物-(3'UTR)-(WPRE)-polyA。

关于 rAAV 病毒粒子的效果:

一些实施方案中, rAAV 病毒粒子具有感染性; 另一些实施方案中, rAAV 病毒粒子不具有感染性。

15 一些实施方案中, 与包含相应亲本 AAV 衣壳蛋白的 AAV 病毒粒子对细胞的感染性相比, 前述本公开的 rAAV 病毒粒子表现出增强至少 2 倍、至少 3 倍、至少 4 倍、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 50 倍以上的细胞感染性, 所述细胞选自:

1) 视网膜细胞;

20 2) 感光细胞;

3) RPE 细胞;

4) 双极细胞;

5) 无长突细胞; 和/或

6) 水平细胞。

25 一些实施方案中, 与包含相应亲本 AAV 衣壳蛋白的 AAV 病毒粒子穿过内界膜(ILM)的能力相比, 前述本公开的 rAAV 病毒粒子表现出增强至少 2 倍、至少 3 倍、至少 4 倍、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 50 倍以上的穿过 ILM 的能力。

一些实施方案中, 前述本公开的 rAAV 病毒粒子选择性感染视网膜细胞, 例如, 30 前述本公开的 rAAV 病毒粒子以比非视网膜细胞(例如非眼组织细胞)强 2 倍、3 倍、4 倍、5 倍、10 倍、15 倍、20 倍、25 倍、50 倍或 50 倍以上的特异性感染视网膜细胞。

一些实施方案中, 与包含 AAV2.7m8 衣壳蛋白的 rAAV 病毒粒子相比, 前述本公开 rAAV 病毒粒子(例如衣壳蛋白包含 seq2 的 rAAV 病毒粒子), 对基因产物(例如阿柏西普)在全眼组织的表达量强 2 倍、3 倍、4 倍、5 倍、10 倍、15 倍、20 倍、35 25 倍、50 倍或 50 倍以上。

一些实施方案中，与包含 AAV2.7m8 衣壳蛋白的 rAAV 病毒粒子相比，前述本公开 rAAV 病毒粒子(例如衣壳蛋白包含 seq2 的 rAAV 病毒粒子)，对基因产物(例如阿柏西普)在房水的表达量强 2 倍、3 倍、4 倍、5 倍、10 倍、15 倍、20 倍、25 倍、50 倍或 50 倍以上。

5 一些实施方案中，与包含 AAV2.7m8 衣壳蛋白的 rAAV 病毒粒子相比，前述本公开 rAAV 病毒粒子(例如衣壳蛋白包含 seq2 的 rAAV 病毒粒子)，对基因产物(例如阿柏西普)在玻璃体的表达量强 2 倍、3 倍、4 倍、5 倍、10 倍、15 倍、20 倍、25 倍、50 倍或 50 倍以上。

一些实施方案中，与包含 AAV2.7m8 衣壳蛋白的 rAAV 病毒粒子相比，前述
10 本公开 rAAV 病毒粒子(例如衣壳蛋白包含 seq2 的 rAAV 病毒粒子)，对基因产物(例如阿柏西普)在视网膜组织的表达量强 2 倍、3 倍、4 倍、5 倍、10 倍、15 倍、20 倍、25 倍、50 倍或 50 倍以上。

一些实施方案中，上述 rAAV 病毒粒子的给药方法是经玻璃体内注射。

多核苷酸和(表达)载体

15 本公开提供编码阿柏西普(aflibercept)的多核苷酸，其编码的氨基酸序列如 SEQ ID NO: 38 所示。

一些实施方案中，本公开提供编码阿柏西普(aflibercept)的经密码子优化的多核苷酸，如 SEQ ID NO: 39-41 任一所示或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%序列同一性。

20 本公开提供编码前述本公开的任意基因产物的多核苷酸。

本公开提供编码前述本公开任意的 AAV 衣壳蛋白的多核苷酸。

一些实施方案中，上述多核苷酸可以为 RNA、DNA 或 cDNA。

一些实施方案中，上述多核苷酸是分离的多核苷酸。

本公开的多核苷酸也可呈载体形式，可存在于载体中和/或可为载体的一部分，
25 该载体例如质粒、粘端质粒、YAC 或病毒载体。载体可例如为表达载体，即可提供多核苷酸在体外和/或体内(即在适合宿主细胞、宿主有机体和/或表达系统中)表达的载体。该表达载体通常包含至少一种本公开的多核苷酸，其可操作地连接至一个或多个适合的表达调控元件(例如启动子、增强子、终止子等)。

本公开的多核苷酸可通过已知的方式(例如通过自动 DNA 合成和/或重组 DNA
30 技术)制备或获得，和/或可从适合的天然来源加以分离。

本公开提供载体，其包含：

(a) 分离的多核苷酸，其编码前述本公开任意的变体 AAV 衣壳蛋白(如 SEQ ID NO: 3-9 和 23-27 中任一所示或与之具有至少 90%或 95%的序列同一性)；

(b) 分离的多核苷酸，其编码前述本公开任意的变体 AAV 衣壳蛋白(如 SEQ ID
35 NO: 3-9 和 23-27 中任一所示或与之具有至少 90%或 95%的序列同一性)；和/或前述编码基因产物的异源多核苷酸(如编码 SEQ ID NO: 38 所示阿柏西普(aflibercept))

的多核苷酸、SEQ ID NO: 39-41 中任一所示或与之具有至少 90%或 95%序列同一性的多核苷酸)。

一些实施方案中, (b)中编码变体 AAV 衣壳蛋白的多核苷酸和编码基因产物的异源多核苷酸在不同的载体中。

5 宿主细胞

本公开提供宿主细胞, 其包含前述本公开任意的多核苷酸或(表达)载体。

一些实施方案中, 所述宿主细胞可以是分离的细胞, 例如来自体外细胞培养物的细胞。这种细胞用于产生前述本公开任意的 rAAV 衣壳蛋白、基因产物或 rAAV 病毒粒子, 又称生产细胞。

10 一些具体实施方案中, 生产细胞为细菌细胞、真菌细胞或哺乳动物细胞。示例性生产细胞包括但不限于 HeLa、CHO、293(包括 293T)、Vero、NIH 3T3、Huh-7、BHK、PC12、COS(包括 COS-7)、RAT1、HepG2 细胞等。示例性哺乳动物细胞包括但不限于 293(293T)、COS、HeLa、Vero、3T3、C3H10T1/2、CHO 细胞。生产细胞也可使用两栖类细胞、昆虫细胞、植物细胞及本领域中用于表达蛋白、病毒

15 粒子的任何其他细胞。示例性昆虫细胞包括但不限于草地贪夜蛾、果蝇(*Drosophila*)细胞系, 或蚊细胞系, 如白纹伊蚊(*Aedes albopictus*)衍生细胞系, 包括但不限于 Se301、SeIZD2109、SeUCR1、Sf9、Sf900+、Sf21、BTI-TN-5B1-4、MG-1、Tn368、HzAm1、Ha2302、Hz2E5、HighFive(Invitrogen, CA, USA)、AO38 和 BM-N 细胞。

rAAV 病毒粒子的生产、制备方法

20 生产 rAAV 病毒粒子的方法是本领域常规的。本公开全文引入 WO200028004、WO200123001、WO2004112727、WO 2005005610、WO2005072364、WO2013123503、WO2015191508 和 US20130195801 中的 rAAV 病毒粒子生产、制备方法。所述 rAAV 病毒粒子可以具有增强递送效率的特性, 能进行有效包装, 能以高频率和最小的毒性成功地感染靶细胞(例如哺乳动物或人细胞)。

25 本公开提供生产、制备 rAAV 病毒粒子的方法, 包括将本公开任意多核苷酸包装至 AAV 衣壳中。

一些实施方案中, 提供生产、制备 rAAV 病毒粒子的方法, 包括: 将前述本公开包含编码基因产物的多核苷酸或其(表达)载体, 编码前述本公开任意 AAV 衣壳蛋白的多核苷酸或其(表达)载体, 和辅助功能质粒(例如, pHelper), 引入生产细胞

30 (例如 293 细胞), 包装、纯化, 获得 rAAV 病毒粒子。

一些实施方案中, 提供生产、制备 rAAV 病毒粒子的生产方法, 包括:

1) 将本公开任意包含编码基因产物的多核苷酸或(表达)载体, 表达 Rep 和 Cap 基因的载体(例如, pR2C9), 以及辅助载体(以实现辅助功能, 例如 pHelper)同时共转染哺乳动物细胞(例如 293 细胞), 所述表达 Rep 和 Cap 基因的载体含有编码前述

35 本公开任意 AAV 衣壳蛋白的多核苷酸;

2) 收获、纯化包含编码基因产物的多核苷酸的 rAAV 病毒粒子。

一些实施方案中, 提供 rAAV 病毒粒子生产系统, 用于生产前述本公开任意的 rAAV 病毒粒子, 包含:

- 1) 编码 AAV 衣壳蛋白的多核苷酸;
- 2) 本公开任意的包含编码基因产物的异源多核苷酸或其(表达)载体(例如, pGOI 质粒); 和
- 3) 辅助元件, 其具有足够的 AAV rep 功能和辅助功能, 以将 2) 中包含编码基因产物的异源多核苷酸包装到 AAV 衣壳中。

一些具体实施方案中, 足够的 AAV rep 功能和辅助功能通过包装细胞或 pHelper、pR2C9、pGOI 三质粒提供, 所述包装细胞可以包含 pHelper、pR2C9、pGOI 这三种质粒。

- 一些实施方案中, Rep 基因编码调节功能的非结构蛋白, 例如 AAV 基因组的复制, 可以选自 Rep78、Rep68、Rep52、Rep40。Rep78 和 Rep68 通常从 p5 启动子转录, 而 Rep52 和 Rep40 通常从 p19 启动子转录。Cap 基因编码组装形成病毒壳体的结构蛋白 VP1、VP2 和/或 VP3。Cap 基因通常从 p40 启动子转录。

一些实施方案中, 提供上述 AAV 生产方法或生产系统生产的 rAAV 病毒粒子。

药物组合物

- 本公开提供药物组合物, 其含有:

- (a) 前述本公开任意的 AAV 衣壳蛋白或 rAAV 病毒粒子; 和
- (b) 一种或多种药学上可接受的载剂、稀释剂、赋形剂或缓冲液。

本公开提供药物组合物, 其含有:

- (a) 预防或治疗有效量的活性成分; 和
- (b) 一种或多种药学上可接受的载剂、稀释剂、赋形剂或缓冲液;

所述活性成分例如选自: 前述本公开任意的 rAAV 病毒粒子, 编码 SEQ ID NO: 39-41 或其至少 80%、85%、90%、95%、96%、97%、98%、99% 同一性序列的基因产物的多核苷酸。

- 一些实施方案中, 所述药物组合物单位剂量中可含有 0.01 至 99 重量%的多核苷酸(例如 VEGF 抑制剂, 又例如阿伯西普的编码多核苷酸)或 rAAV 病毒粒子。一些具体实施方案中, 药物组合物单位剂量中含基因产物(例如 VEGF 抑制剂, 又例如阿伯西普)拷贝量为 0.1 至 10×10^{13} 。一些具体实施方案中, 药物组合物中的 rAAV 病毒粒子浓度为每毫升 1×10^8 个或更多, 通常不多于每毫升 1×10^{15} 个。

- 一些实施方案中, 可以用前述本公开任意的 rAAV 病毒粒子转染细胞, 而后将所述细胞转移或移植入受试者。

一些实施方案中, 药物组合物中含有前述本公开任意的 rAAV 病毒粒子, 所述 rAAV 病毒粒子包裹有多核苷酸, 所述多核苷酸为 SEQ ID NO: 39-41 任一所示或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99% 同一性序列。

治疗、预防疾病的方法和制药用途

- 本公开提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿伯西普(afibercept)的多核苷酸用于治疗、缓解或改善疾病或症状的方法或用途。

一些实施方案中, 提供向有需要的受试者递送基因产物的方法, 包括向所述

受试者施用有效量的前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸。

一些实施方案中，提供向靶细胞递送基因产物的方法，其包括使所述靶细胞与前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸接触。一些具体实施方案中，靶细胞选自肝细胞、胰腺细胞、骨骼肌细胞、心肌细胞、成纤维细胞、视网膜细胞、滑膜关节细胞、肺细胞、T 细胞、神经元、神经胶质细胞、干细胞、内皮细胞或癌细胞。一些具体实施方案中，靶细胞是体外的；另一些具体实施方案中，靶细胞是体内的。同时提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸制备向靶细胞递送异源核酸的制药用途。

一些实施方案中，提供特异性感染视网膜细胞的方法，包括向眼内施用预防或治疗有效量的前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸，例如通过玻璃体内注射或视网膜下注射。同时提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸制备特异性感染视网膜细胞药物的制药用途。

一些具体实施方案中，提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸用于治疗选自由以下组成的组的视网膜细胞的疾病或病症：感光细胞、视网膜神经节细胞、Muller 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，视网膜细胞是感光细胞，例如视杆细胞或视锥细胞。

一些具体实施方案中，提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸用于治疗选自由以下组成的组的视网膜细胞的疾病或病症：急性黄斑神经视网膜病变；白塞氏病；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，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，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；眼外伤，其影响眼睛后部位点或位置；眼部肿瘤；视网膜病症，如视网膜中央静脉阻塞、糖尿病性视网膜病变(包含增生性糖尿病视网膜病变)、增生性玻璃体视网膜病变(PVR)、视网膜动脉闭塞性疾病、视网膜脱离和葡萄膜炎性视网膜疾病；交感性眼炎；伏格特-小柳-原田三氏(Vogt Koyanagi-Harada)(VKH)综合征；葡萄膜扩散；由眼部激光治疗引起的或受其影响的眼后病状；由光动力学治疗引起或受其影响的眼后病状；光凝固法、放射性视网膜病变；视网膜前膜病症；视网膜分支静脉阻塞；前部缺血性视神经病变；非视网膜病变糖尿病视网膜功能障碍；视网膜劈裂症；视网膜色素变性；青光眼；乌谢尔综合征、锥体-杆体营养不良；斯图加特氏疾病(眼底黄色斑点症)；遗传性黄斑变性；脉络膜视网膜变性；莱伯先天性黑朦；先天性静止性夜盲；无脉络膜；巴比二氏综合征；黄斑毛细血管扩张症；莱伯遗传性视神经病变；早

产儿视网膜病变；以及色觉病症，包含全色盲、红色盲、绿色盲和蓝色盲，遗传性视网膜色素变性。

一些实施方案中，提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸用于治疗眼部疾病的方法或用途，所述眼部疾病包括但不限于：年龄相关性黄斑变性(AMD)、湿性 AMD、干性 AMD、视网膜新生血管、脉络膜新生血管、糖尿病性视网膜病变、增生性糖尿病视网膜病变、视网膜静脉阻塞、视网膜中央静脉阻塞、视网膜分支静脉阻塞、糖尿病性黄斑水肿、糖尿病性视网膜缺血、缺血性视网膜病变或糖尿病性视网膜水肿。可选地，是通过玻璃体内注射或视网膜下注射。

一些实施方案中，提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸治疗视网膜相关疾病的方法或用途。

同时提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸制备治疗或预防上述疾病的药物的制药用途。

给药方法

本公开提供前述本公开任意的 AAV 衣壳蛋白、rAAV 病毒粒子、药物组合物、编码阿柏西普(aflibercept)的多核苷酸用于治疗前述眼部疾病(例如视网膜相关疾病)的给药方法，其是通过眼内注射施用。

一些实施方案中，眼内注射包括通过玻璃体内注射、通过视网膜下注射、通过脉络膜上注射或通过将导致 rAAV 病毒粒子递送至眼睛的任何其它方便的施用模式或途径。其它方便的施用模式或途径包含但不限于静脉内、动脉内、眼周、前房内、结膜下和眼球囊下注射和局部施用和鼻内施用。

附图说明

图 1 为 AAV2 衣壳蛋白序列及 VP1、VP2、VP3 的起始位点，环状结构域 I-V 的示意图。

图 2 为本公开变体 AAV 衣壳蛋白在 587 到 588 位之间插入肽段，具有 708 位点突变的示意图。

图 3 为本公开的 rAAV 载体病毒 AAV2 seq1、AAV2 seq2、AAV2 seq3、AAV2 seq4、AAV2 seq5、AAV2 seq6 与 AAV2、AAV2.7m8 感染小鼠视网膜的荧光信号检测结果图。

图 4A 和图 4B 为本公开的 rAAV 载体病毒 AAV2 seq1、AAV2 seq2、AAV2 seq3、AAV2 seq4、AAV9 seq2 与 AAV2、AAV2.7m8 感染小鼠视网膜荧光信号检测结果图，图 4A 为第一周眼底镜检测结果，图 4B 为第四周眼底镜检测结果。

图 5A 为含有表达 VEGF Trap(阿柏西普)的核酸分子的载体结构示意图；图 5B

为 AAV 包装质粒示意图。

图 6 为 rAAV 载体病毒 AAV2 seq2 和 AAV2.7m8 在小鼠眼球表达目的基因(阿柏西普)的检测结果。

图 7 为 rAAV 载体病毒 AAV2 seq2 和 AAV 2.7m8 在兔眼球中表达目的基因蛋白的检测结果，其是病毒注射 14 天后，对房水中目的基因(阿柏西普)蛋白含量的检测。

图 8 为 rAAV 载体病毒 AAV2 seq2 和 AAV 2.7m8 在兔眼球中表达目的基因蛋白的检测结果，其是病毒注射 28 天后，对房水中目的基因蛋白摩尔浓度的检测。

10 具体实施方式

为了更容易理解本公开，以下具体定义了一些技术和科学。除非在本文中另有明确定义，本文使用的所有其它技术和科学都具有本公开所属领域的一般技术人员通常理解的含义。

“AAV”是腺相关病毒的缩写，且可用于指代病毒本身或其衍生物。除非另有所指，否则该涵盖所有 AAV 亚型以及天然存在形式和重组形式。“AAV”包含但不限于 AAV 1 型(AAV1)、AAV 2 型(AAV2)、AAV 3 型(AAV3)、AAV 4 型(AAV4)、AAV 5 型(AAV5)、AAV 6 型(AAV6)、AAV 7 型(AAV7)、AAV 8 型(AAV8)、AAV 9 型(AAV9)、AAV10 型(AAV10)、AAVrh10 型(AAVrh10)，以及不同种属的禽类 AAV、牛类 AAV、犬类 AAV、马类 AAV、灵长类 AAV、非灵长类 AAV 和绵羊类 AAV。

AAV 是由在非包膜二十面体衣壳内的 4.7kb 单链 DNA 基因组构成的非致病性细小病毒，基因组含有侧接有用作病毒复制起点和包装信号的反向末端重复(ITR)的三个开放阅读框(ORF)。Rep ORF 编码四个非结构蛋白，这些蛋白在病毒复制、转录调控、位点特异性整合、及病毒粒子装配中发挥作用。Cap ORF 编码三个结构蛋白(VP1-3)，这些蛋白装配形成 60 聚体病毒衣壳。最终，作为 cap 基因内的替代阅读框存在的 ORF 产生装配活化蛋白(AAP)，这是一种将 AAV 衣壳蛋白定位在细胞核中并在衣壳装配过程中发挥作用的病毒蛋白。

各种血清型 AAV 的基因组序列以及天然末端重复(ITR)序列、Rep 蛋白序列和衣壳亚基序列是本领域已知的。这种序列可以在文献或如 GenBank 等公共数据库中 30 找到。参见例如 GenBank 登录号 NC_002077.1(AAV1)、AF063497.1(AAV1)、NC_001401.2(AAV2)、AF043303.1(AAV2)、J01901.1(AAV2)、U48704.1(AAV3)、NC_001729.1(AAV3)、NC_001829.1(AAV4)、U89790.1(AAV4)、NC_006152.1(AAV5)、AF085716.1(AAV5)、AF028704.1(AAV6)、NC_006260.1(AAV7)、AF513851.1(AAV7)、AF513852.1(AAV8)、NC_006261.1(AAV8)、以及 AY530579.1(AAV9)。以及，35 Srivistava 等人(1983)J.Virology45: 555; Chiorini 等人(1998)J.Virology71: 6823;

Chiorini 等人(1999)*J.Virology*73: 1309; Bantel-Schaal 等人(1999)*J.Virology*73: 939; Xiao 等人(1999)*J.Virology*73: 3994; Muramatsu 等人(1996)*Virology*221: 208; Shade 等人,(1986)*J.Virol.*58: 921; Gao 等人(2002)*Proc.Nat.Acad.Sci.USA*99: 11854; Moris 等人(2004)*Virology*33: 375-383; 以及, WO00/28061、WO99/61601、WO98/11244、
5 US6156303、WO2012145601A、WO2017197355A、WO2018022905A。以上内容全文引入本公开。

关于 AAV 衣壳的 GH 环或环 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。

10 “AAV 病毒颗粒”或“AAV 病毒粒子”是指由至少一种 AAV 衣壳蛋白和衣壳化的 AAV 多核苷酸构成的病毒颗粒。

“rAAV”是指重组腺相关病毒, 应用于多核苷酸的“重组”是指多核苷酸是克隆、限制或连接步骤的各种组合的产物、以及导致不同于在自然中找到的多核苷酸的构建体的其它过程。重组病毒是包括重组多核苷酸的病毒颗粒。

15 如果“AAV 病毒粒子”包括异源多核苷酸(即除野生型 AAV 基因组以外的多核苷酸, 例如待递送到靶细胞的基因产物(例如转基因、RNAi 等), 它通常被称为“重组 AAV(rAAV)病毒粒子”或“rAAV 病毒颗粒”或“rAAV 载体病毒”。通常, 异源多核苷酸的两侧是至少一个, 并且通常是两个 AAV 反向末端重复序列(ITR)。

20 “rAAV 载体”涵盖 rAAV 病毒粒子, 其包含 rAAV 多核苷酸; 并且还涵盖对 rAAV 进行编码的多核苷酸(例如, 对 rAAV 进行编码的单链多核苷酸(ss-rAAV); 对 rAAV 进行编码的双链多核苷酸(ds-rAAV), 例如对 rAAV 进行编码的质粒; 等等)。

“AAV 变体”、“AAV 突变体”或“衣壳变异的 rAAV”是指由以下构成的
25 病毒颗粒:(a)变体 AAV 衣壳蛋白, 其中变体 AAV 衣壳蛋白相对于对应的亲本 AAV 衣壳蛋白包括至少一个氨基酸差异(例如氨基酸取代、插入或缺失), 其中 AAV 衣壳蛋白与天然存在的 AAV 衣壳蛋白的氨基酸序列不同或不对应; 以及任选地, (b)包括编码异源基因产物的异源多核苷酸, 其中与包括对应的亲本 AAV 衣壳蛋白的 AAV 病毒粒子的结合相比, 变体 AAV 衣壳蛋白赋予了与类肝素或硫酸类肝素蛋白多糖的增加结合。
30

“包装”是指导致 AAV 颗粒进行组装和衣壳化的一系列细胞内事件。

“rep”和“cap”基因是指对腺相关病毒的复制和衣壳化蛋白进行编码的多核苷酸序列。AAV rep 和 cap 在本文中被称为 AAV “包装基因”。

“辅助病毒”是指允许 AAV(例如野生型 AAV)被哺乳动物细胞复制和包装的
35 病毒。AAV 的各种辅助病毒是本领域已知的, 包含腺病毒、疱疹病毒和如牛痘等痘病毒。腺病毒涵盖许多不同的亚群, 但是最常用的是 C 亚群中的腺病毒 5 型;

疱疹家族的病毒包含例如单纯疱疹病毒(HSV)和艾巴氏病毒(EBV)以及巨细胞病毒(CMV)和假狂犬病毒(PRV)。人类、非人类哺乳动物和禽类来源的许多腺病毒、疱疹病毒等是已知的并且可从如 ATCC 等机构获得。

“辅助病毒功能”或“辅助功能”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。本公开中,“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。

“感染性”病毒或病毒颗粒是包括适合地组装的病毒衣壳并且能够将多核苷酸组分递送到病毒种类具有趋向性的细胞中的病毒或病毒颗粒,不必然暗示病毒有任何复制能力。本领域公知对感染性病毒颗粒进行计数的测定。病毒感染性可以表达为感染性病毒颗粒与总病毒颗粒之比。确定感染性病毒颗粒与总病毒颗粒之比的方法是本领域已知的。参见例如 Grainger 等人(2005)《分子疗法(Mol. Ther.)》11: S337(描述了 TCID₅₀ 感染滴度测定); Zolotukhin 等人(1999)《基因疗法(Gene Ther.)》6: 973。

“趋向性”或“特异性”是指病毒(例如, AAV)优先靶向特定宿主物种的细胞或宿主物种内的特定细胞类型。例如,相对于仅可能感染肺和肌肉细胞的病毒,可以感染心脏、肺、肝和肌肉细胞的病毒具有更宽(即增加)的趋向性。趋向性还可以包含病毒对宿主的特定类型的细胞表面分子的依赖性。例如,一些病毒只可能感染具有表面糖胺聚糖的细胞,而其它病毒只可能感染具有唾液酸的细胞(这种依赖性可以使用缺乏特定类别分子的各种细胞系作为病毒感染的潜在宿主细胞进行测试)。在某些情况下,病毒的趋向性描述了病毒的相对偏好。例如,第一种病毒可能能够感染所有细胞类型,但在用表面糖胺聚糖感染这些细胞方面更成功。如果第二种病毒也更喜欢相同的特征(例如,第二种病毒在用表面糖胺聚糖感染这些细胞方面也更成功),则可以认为第二种病毒与第一种病毒具有相似(或一致)的趋向性,即使绝对转导效率不相似。例如,第二种病毒在感染每种测试的给定细胞类型时可能比第一种病毒更有效,但如果相对偏好相似(或一致),则第二种病毒仍可被认为具有与第一种病毒相似(或一致)的趋向性。在一些实施例中,相对于天然存在的病毒粒子,包括本公开变体 AAV 衣壳蛋白的病毒粒子的趋向性没有改变。在一些实施例中,相对于天然存在的病毒粒子,包括本公开变体 AAV 衣壳蛋白的病毒粒子的趋向性被扩展(即,变宽)。在一些实施例中,相对于天然存在的病毒粒子,包括本公开变体 AAV 衣壳蛋白的病毒粒子的趋向性降低。

“多核苷酸”是指任何长度的核苷酸的聚合形式或其类似物,包含脱氧核糖核苷酸或核糖核苷酸。多核苷酸可以包括经过修饰的核苷酸,如甲基化核苷酸和核苷酸类似物,并且可以被非核苷酸组分打断。对核苷酸结构的修饰可以在组装聚合物之前或之后进行。多核苷酸可互换地指双链分子和单链分子,除非另外指明,否则本公开的多核苷酸涵盖双链形式以及构成双链形式的两种互补的单链形

式。

“同源性”或“同一性”是指两个多核苷酸序列之间或两个多肽之间的序列相似性。当两个比较序列中的位置均被相同核苷酸或氨基酸单体亚基占据时，例如如果两个 DNA 分子的每一个位置都被相同核苷酸占据时，那么所述分子在该位置是同源的。两个序列之间的同源性百分率是两个序列共有的匹配或同源位置数除以比较的位置数 $\times 100\%$ 的函数。例如，在序列最佳比对时，如果两个序列中的 10 个位置有 6 个匹配或同源，那么两个序列为 60%同源。序列相似性可以许多不同的方式测定。为了测定序列同一性，可使用方法和计算机程序比对序列，计算机程序包括可由因特网地址 ncbi.nlm.nih.gov/BLAST/ 获得的 BLAST。另一种比对算法是 FASTA，可由 Madison, Wisconsin, USA, a wholly owned subsidiary of Oxford Molecular Group, Inc 的 Genetics Computing Group (GCG) 包装中获得。其他比对技术描述于 Enzymology, 第 266 卷: Computer Methods for Macromolecular Sequence Analysis (1996), Doolittle 编著, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USA 的方法中。特别关注的是容许序列中的缺口的对比程序。Smith-Waterman 是容许序列比对中的缺口的一种类型的算法。参见 Meth.Mol.Biol.70: 173-187(1997)。使用 Needleman 和 Wunsch 对比方法的 GAP 程序也可用于比对序列。参见 J.Mol.Biol.48: 443-453(1970)。

“基因”是指含有至少一个开放读码框的能够在转录并且有时翻译后对特定基因产物进行编码的多核苷酸。“基因”或“编码序列”是指对基因产物进行编码的体外或体内核苷酸序列。一些情况下，基因由编码序列组成或基本上由编码序列组成，所述编码序列即对基因产物进行编码的序列。另一些情况下，基因包括另外的非编码序列。例如，基因可以包含或不包含在编码区域之前和之后的区域(例如 5'UTR、3'UTR 以及各个编码区段(外显子)之间的插入序列(内含子))。

“基因产物”是由特定基因的表达产生的分子，例如多肽、适配体、干扰 RNA、mRNA 等。一些实施方案中，“基因产物”是多肽、肽、蛋白质或干扰 RNA，包含短干扰 RNA(siRNA)、miRNA 或小发夹 RNA(shRNA)。一些具体实施方案中，基因产物是治疗性基因产物，例如治疗性多肽。所述治疗性基因产物对其所在的细胞、组织或哺乳动物赋予了有益效果，所述有益效果包含改善病状或疾病的体征或症状、预防或抑制病状或疾病或者赋予期望的特性。当基因编码多肽时，“基因产物”和“目的基因表达的产物”可互用。

“RNA 干扰剂”或“RNAi 试剂”包括可用于改变基因(如上所定义)表达的任何试剂(或编码这种试剂的多核苷酸)。本领域普通技术人员已知的 RNAi 试剂的实例包括但不限于(i)siRNA 试剂(“小干扰”或“短干扰 RNA”(或 siRNA)); (ii)反义 RNA; (iii)CRISPR 试剂; (iv)锌指核酸酶试剂; 以及(v)转录激活因子样效应子核酸酶(TALEN)试剂。这里全文引入 WO2017197355A 对(i)-(v)的定义。

其中，siRNA 试剂是靶向所关注基因(“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形成的结构，从而形成双链 RNA 的区域(dsRNA)。siRNA “靶向”基因是因为 siRNA 的双链体部分的核苷酸序列与靶基因的核苷酸序列互补。在一些实施方案中，siRNA 的双链体长度小于 30 个核苷酸。在一些实施方案中，双链体可以是 29、28、27、26、25、
5 24、23、22、21、20、19、18、17、16、15、14、13、12、11 或 10 个核苷酸长度。在一些实施方案中，双链体的长度是 19-25 个核苷酸长度。siRNA 的 RNA 双链体部分可以是发夹结构的一部分。含有发夹的 siRNA 试剂也可被称为“shRNA(短发夹 RNA)试剂”。除双链体部分之外，发夹结构还可含有位于形成双链体的两条序列之间的环部分。环在长度方面可变化。在一些实施方案中，环的长度是 5、6、7、
10 8、9、10、11、12 或 13 个核苷酸。发夹结构还可含有 3'或 5'突出端部分。在一些实施方案中，突出端是 0、1、2、3、4 或 5 个核苷酸长度的 3'或 5'突出端。一般说来，靶基因的表达产物(例如，mRNA、多肽等等)水平通过 siRNA 试剂(例如，siRNA、shRNA 等等)而降低，所述 siRNA 试剂含有至少与靶基因转录产物的 19-25 个核苷酸长度的区段(例如，20-21 个核苷酸序列)互补的特定双链核苷酸序列，包括 5'非
15 翻译(UTR)区、ORF、或 3'UTR 区。在一些实施方案中，短干扰 RNA 的长度为约 19-25nt。参见，例如 PCT 申请 WO0/44895、WO99/32619、WO01/75164、WO01/92513、WO01/29058、WO01/89304、WO02/16620、及 WO02/29858；以及美国专利公布号 20040023390(关于 siRNA 技术的描述)。siRNA 和/或 shRNA 可由核酸序列编码，且该核酸序列还可包括启动子。该核酸序列还可包括聚腺苷酸化
20 信号。在一些实施方案中，聚腺苷酸化信号是合成的最小聚腺苷酸化信号。

其中，反义 RNA 是与基因表达产物互补的 RNA。举例来说，靶向特定 mRNA 的反义 RNA 是与 mRNA 互补的基于 RNA 的试剂(或可以是修饰的 RNA)，其中反义 RNA 与 mRNA 的杂交改变 mRNA 的表达(例如，经由改变 RNA 的稳定性、改变 RNA 的翻译，等等)。编码反义 RNA 的核酸也包括在“反义 RNA”中。

25 “VEGF”是指诱导血管生成或血管生成过程的血管内皮生长因子，包含通过例如替代性地剪接 VEGF-A/VPF 基因产生的各种亚型的 VEGF(也称为血管通透因子(VPF)和 VEGF-A)(参见美国专利申请公开第 20120100136 号的图 2(A)和(B))，包含 VEGF121、VEGF165 和 VEGF189。进一步地，“VEGF”包含 VEGF 相关血管生成因子，如 PlGF(胎盘生长因子)、VEGF-B、VEGF-C、VEGF-D 和 VEGF-E，
30 所述 VEGF 相关血管生成因子通过同源 VEGF 受体(即 VEGFR)起作用以诱导血管生成或血管生成过程。

“sFlt-1”或“sFlt-1 蛋白”在本文中是指与天然存在的人 sFLT-1 序列至少 90% 或更高同源，使得 sFlt-1 蛋白或多肽与 VEGF 和/或 VEGF 受体结合的多肽序列或其功能片段。

35 “多肽”、“肽”和“蛋白质”是指任何长度的氨基酸的聚合物，还包含被修饰的氨基酸聚合物，例如二硫键形成、糖基化、脂化、磷酸化或与标记组分缀

合。

“调控元件”或“调控序列”是参与分子的相互作用的核苷酸序列，所述相互作用有助于对多核苷酸进行功能调节，包含多核苷酸的复制、重复、转录、剪接、翻译或降解。调节会影响过程的频率、速度或特异性并且在本质上会具有增强或抑制性。已知的控制元件包含例如转录调节序列，如启动子和增强子。启动子是在某些条件下能够结合 RNA 聚合酶并且启动对通常位于启动子下游(沿 3'方向)的编码区进行转录的 DNA 区域。

“表达载体”是包括对目的基因产物进行编码的区域的载体，并且用于在预期靶细胞中实现基因产物的表达，所述载体包括对关注的基因产物进行编码的多核苷酸。表达载体还包括与编码区域操作性地连接的以促进基因产物在靶标中表达的控制元件。控制元件例如启动子、增强子、UTR、miRNA 靶向序列等和与其可操作地连接以供表达的一个或多个基因的组合有时被称为“表达盒”。许多表达盒在本领域中是已知且可获得的，或者可以由本领域中可获得的组分容易构建的。

“操作性地连接”或“可操作地连接”是指遗传元件的并列，其中元件处于允许其按预期方式操作的关系中。例如，如果启动子帮助启动编码多核苷酸序列的转录，则启动子与编码多核苷酸序列操作性地连接，在启动子与编码多核苷酸序列之间可能有插入核酸残基，只要维持此功能关系即可。

“施用”或“引入”是指将用于重组基因或蛋白表达的载体递送到细胞、受试者的细胞和/或器官或者受试者。这种施用或引入可以在体内、在体外或离体地发生。用于表达基因产物的载体可以通过以下引入到细胞中：转染，其通常意指通过物理手段(例如，磷酸钙转染、电穿孔、微注射或脂质转染)将异源 DNA 插入到细胞中；感染，其通常是指通过感染剂即病毒引入；或者转导，其通常意指用细胞稳定感染细胞或将来自一个微生物的基因材料通过病毒剂(例如细菌噬菌体)转移到另一个微生物。“转化”通常用于指包括异源 DNA 的细菌或表达致癌基因并且已经转换为连续生长模式的细胞，如肿瘤细胞。用于“转化”细胞的载体可以是质粒、病毒或其它媒介。根据用于将异源 DNA(即载体)施用、引入或插入到细胞中的方式，细胞通常被称为“转导”、“感染”、“转染”或“转化”。“转导”、“转染”和“转化”可以在本文中可互换地使用，不考虑异源 DNA 的引入方法。

“宿主细胞”是指已经用载体转导、感染、转染或转化的细胞，涵盖最初转导、感染、转染或转化的细胞和其子代。载体可以是质粒、病毒颗粒、噬菌体等。如温度、pH 等培养条件对于本领域技术人员而言将是显而易见的。

“治疗”通常用于意指获得期望的药理学和/或生理学效果。效果在完全或部分地预防疾病或其症状方面可以是预防性的，例如降低受试者体内发生疾病或其症状的可能性，和/或在部分或完全治愈疾病和/或由疾病引起的不良反应方面可以

是治疗性的。“治疗”覆盖对哺乳动物的疾病的任何治疗，并且包含：(a)防止疾病在可能倾向于患有疾病但尚未被诊断为患有疾病的受试者身上发生；(b)抑制或中止疾病发展；或者(c)缓解疾病(或其引起的症状)或使疾病消退。可以在疾病或损伤发作之前、期间或之后施用治疗剂。特别是对发展中的疾病的治疗，其中所述治疗稳定或减少了患者的不期望的临床症状。一些优选方案是，在受影响组织的功能完全丧失之前执行这种治疗。一些优选方案是，将在疾病的症状期期间并且在一些情况下在疾病的症状期之后施用本公开的治疗。

“有效量”包含足以改善或预防医学病症的症状或病症的量。有效量还意指足以允许或促进诊断的量。用于特定受试者的有效量可依据以下因素而变化：如待治疗的病症、受试者的总体健康情况、给药的方法途径和剂量以及副作用严重性。有效量可以是避免显著副作用或毒性作用的最大剂量或给药方案。

“视网膜细胞”在本公开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。

“个体”、“受试者”和“患者”在本文中可互换地使用，包括但不限于人类和非人类灵长类动物，如猿猴、人类和其他哺乳类动物(例如马、绵羊、山羊、狗、猫，以及啮齿动物(例如小鼠、大鼠等))，优选为人。

本公开中，“多肽”和“蛋白”可互换使用。

20

实施例

以下结合实施例进一步描述本公开，但这些实施例并非限制着本公开的范围。本公开实施例中未注明具体条件的实验方法，通常按照常规条件或按照原料或商品制造厂商所建议的条件。未注明具体来源的试剂，则该试剂可自任意分子生物学试剂的供应商以用于分子生物学应用的质量/纯度而获得。

25

实施例 1. AAV 衣壳改造

分子动力学模拟和蛋白质指纹技术在蛋白质工程、抗体亲和性改造中被广泛应用，通过对抗体和受体的理化性质的动力学模拟，在一定程度上可以对氨基酸改变造成的局部亲和力影响进行预测。

30

已证实硫酸乙酰肝素受体为介导 AAV 细胞感染的最主要受体，在视网膜和是视神经细胞的感染中也起到重要的作用，研究证实在 AAV2 天然血清型衣壳蛋白氨基酸序列依照其空间结构特征，有 5 个环状结构域(图 1 所示)与病毒的感染能力相关。其中，环状结构域 IV 的第 587 和 588 位点中间插入 7-12 个氨基酸的短肽不会影响病毒衣壳蛋白的组装，同时能够大幅度的改变 AAV 衣壳与细胞受体(硫酸乙酰肝素)的相互作用方式，从而影响 AAV 对宿主细胞的感染能力。

35

本公开利用自主研发的生物信息学算法，在 AAV2 天然血清型衣壳蛋白氨基酸序列中的第 587 和 588 位点之间插入若干随机氨基酸序列(7-12 个氨基酸长度)，并预测衣壳蛋白与硫酸乙酰肝素受体间的相互作用。选取算法中预测能够较大提升感染能力的序列进行后续验证。根据算法预测，选取如表 1 所示的 6 个氨基酸序列 seq1 至 seq6，每个序列的长度均为 10 个氨基酸。同时，将 AAV2 天然血清型衣壳蛋白氨基酸序列中第 708 位氨基酸-缬氨酸(V)替换为异亮氨酸(I)。587 和 588 位之间的插入方式和 708 位的氨基酸替换方式如图 2 所示，获得经工程改造的 AAV 衣壳蛋白序列。同时，制备获得 seq2 插入 AAV9 相应位置的衣壳。

表 1. AAV 衣壳的第 587 和 588 位之间插入的氨基酸序列及其对应编号

插入多肽的编号及其氨基酸序列		对应的 AAV 衣壳编号及其氨基酸序列号	
seq1	<u>L</u> <u>A</u> L <u>A</u> E <u>T</u> T <u>R</u> P <u>A</u> (SEQ ID NO: 17)	AAV2 seq1	SEQ ID NO: 3
seq2	<u>L</u> <u>A</u> L <u>G</u> D <u>T</u> T <u>R</u> P <u>A</u> (SEQ ID NO: 18)	AAV2 seq2	SEQ ID NO: 4
seq3	<u>L</u> <u>A</u> L <u>G</u> E <u>T</u> T <u>R</u> N <u>A</u> (SEQ ID NO: 19)	AAV2 seq3	SEQ ID NO: 5
seq4	<u>L</u> <u>A</u> K <u>A</u> D <u>T</u> T <u>K</u> N <u>A</u> (SEQ ID NO: 20)	AAV2 seq4	SEQ ID NO: 6
seq5	<u>L</u> <u>A</u> K <u>D</u> D <u>T</u> T <u>R</u> N <u>A</u> (SEQ ID NO: 21)	AAV2 seq5	SEQ ID NO: 7
seq6	<u>L</u> <u>A</u> L <u>A</u> D <u>T</u> T <u>K</u> N <u>A</u> (SEQ ID NO: 22)	AAV2 seq6	SEQ ID NO: 8
seq2	<u>L</u> <u>A</u> L <u>G</u> D <u>T</u> T <u>R</u> P <u>A</u> (SEQ ID NO: 18)	AAV9 seq2	SEQ ID NO: 9

(注：下划线为连接子。)

由于 AAV 衣壳蛋白表达后是由 VP1、VP2、VP3 三种单体蛋白按一定比例组装而成，相比于普通的重组蛋白具有更高级的结构，因此发生在 AAV 衣壳蛋白的多肽插入造成的功能影响不只限于其带来的理化性质改变，更有可能在 VP1、VP2、VP3 三种单体蛋白的组装过程中施加更为复杂的影响。

15

实施例 2. rAAV 载体病毒感染小鼠模型视网膜能力检测

为验证功能，我们将具有 SEQ ID NO: 3 至 SEQ ID NO: 8、AAV2(SEQ ID NO: 1)、AAV2.7m8(SEQ ID NO: 2)这 8 种序列的衣壳，分别包装质粒 pGOI，所述质粒含有并能表达 CMV 启动子驱动的 EGFP 目的基因。将上述不同 AAV 衣壳包装的具有 EGFP 表达能力的 AAV 载体对小鼠玻璃体进行等剂量注射，通过荧光眼底照相来检测 EGFP 的表达，以 EGFP 荧光信号的强度差异来评估不同 AAV 衣壳的对视网膜组织感染效率的差异。

rAAV 载体病毒的包装和纯化步骤包括：将 pHelper、pR2C9、pGOI 三质粒系统转染 293T 贴壁细胞，收获细胞和上清，通过碘克沙醇离心法进行纯化，再经超滤浓缩，将缓冲液置换为 DPBS。病毒包装纯化完成后，通过 qPCR 检测滴度(单位 vg)。将表达 EGFP 目的基因的 8 种病毒(衣壳分别为 AAV2、AAV2.7m8、AAV2 seq1、AAV2 seq2、AAV2 seq3、AAV2 seq4、AAV2 seq5、AAV2 seq6)，分别用

25

DPBS 缓冲液稀释至同一滴度 $1.03E+12$ 。对周龄为 10 周的 C57 小鼠(购于北京维通利华实验技术有限公司)进行单侧眼玻璃体注射,注射体积为 $1\mu L$,每种 AAV 载体注射 3 只($n=3$),注射后 1 周至 4 周,每周进行活体荧光眼底照相。

部分结果如图 3 所示,其中 AAV2 组、AAV2 seq5 组、AAV2 seq6 组,荧光信号强度明显弱于其他组。而 AAV2 seq1 组、AAV2 seq2 组、AAV2 seq3 组、AAV2 seq4 组荧光信号强度与 AAV2.7m8 较为相似,显示其感染视网膜组织的效率较高。

将表 1 中的“LALGDTTRPA”序列插入 AAV9 衣壳中的同一位置(第 588 和 589 位之间),获得 AAV9 seq2。将 AAV2、AAV2.7m8、AAV2 seq1、AAV2 seq2、AAV2 seq3、AAV2 seq4、AAV9 seq2 共 7 种衣壳,分别包装质粒载体,所述质粒载体含有并能表达 CMV 启动子驱动的 EGFP 目的基因。将上述不同 AAV 衣壳包装的具有 EGFP 表达能力的 AAV 载体对小鼠玻璃体进行等剂量注射,通过荧光眼底照相来检测 EGFP 的表达,以 EGFP 信号的强度差异来评估不同 AAV 衣壳的对视网膜组织感染效率的差异。

使用上述方法对 AAV 病毒进行载体包装和纯化。将表达 EGFP 目的基因的 7 种病毒(衣壳分别为 AAV2、AAV2.7m8、AAV2 seq1、AAV2 seq2、AAV2 seq3、AAV2 seq4、AAV2 seq5、AAV9 seq2),分别用 DPBS 缓冲液稀释至同一滴度 $1.03E+12$ 。对周龄为 10 周的 C57 小鼠进行单侧眼玻璃体注射,注射体积为 $1\mu L$,每种 AAV 载体注射 3 只($n=3$),注射后 1 周至 4 周,每周进行活体荧光眼底照相。

结果如图 4 A、图 4B 所示,其中 AAV2 组、AAV2 seq2 组、AAV2 seq3 组,荧光信号强度相较其他组明显较强,且与 AAV2.7m8 相似。

实施例 3. rAAV 载体病毒对小鼠模型视网膜感染的生物学功能检测

为进一步验证 seq2 衣壳对视网膜组织的感染能力和其所关联的生物学功能,将 AAV2 seq2 和 AAV2.7m8 分别包装含有 CMV 启动子驱动的阿柏西普(aflibercept)目的质粒载体(AAV2 seq2-aflibercept, AAV2.7m8-aflibercept),两种病毒分别对新西兰兔和 C57 小鼠进行玻璃体注射,并于注射两周后检测注射眼组织中的阿柏西普蛋白含量,以评估对应 AAV 病毒对视网膜的感染和其关联的生物学功能。

其中,rAAV 递送的基因表达盒的结构从其从 5'端至 3'端含有(如图 5A):CMV 增强子、启动子、5'UTR、Kozak 序列(gccacc)、VEGF 抑制剂阿柏西普、3'UTR、WPRE 以及 SV40polyA。其中,CMV 增强子的碱基序列如 SEQ ID NO: 28 所示,CMV 启动子碱基序列如 SEQ ID NO: 29 所示,5'UTR 如 SEQ ID NO: 30 所示,Kozak 序列如 SEQ ID NO: 31 所示、3'UTR 如 SEQ ID NO: 32 所示、WPRE 如 SEQ ID NO: 33 所示、SV40polyA 如 SEQ ID NO: 34 所示。

表达载体的构建为:构建表达盒,通过酶切、连接、转化和克隆筛选鉴定等常规分子生物学操作构建了表达阿柏西普的 AAV 包装质粒,其依次包括:CMV 增强子、启动子、5'UTR、Kozak 序列(gccacc)、VEGF trap(阿柏西普)、3'UTR、

WPRE 以及 SV40polyA，表达盒两侧是反向末端重复序列(ITR)，其结构图示意图见图 5B，EcoRV 和 BSMI 为酶切位点。

按照实施例 2 中的方法包装载体和纯化 rAAV 载体病毒。

将纯化后的 AAV seq2-aflibercept 和 AAV2.7m8-aflibercept 这两种病毒分别用 DPBS 缓冲液稀释至同一滴度 9E +12 vg/mL，对周龄为 8-10 周的 C57 小鼠进行双眼玻璃体注射，注射体积为 1μL，每种 rAAV 载体病毒注射 4 只(n=4)。注射后 2 周，处死小鼠，取眼组织分别研磨，每个注射组左眼和右眼分别用两种不同的方法进行组织研磨处理。左眼使用方法 1：PBS+蛋白酶抑制剂研磨，离心后取上清。右眼使用方法 2：RIP Buffer+蛋白酶抑制剂研磨，离心后取上清。使用酶联免疫分析法检测上清中的阿柏西普蛋白含量(每个注射组 4 个样本,每个样本检测三重复)。结果如图 6 所示。其中空白对照为 PBS 注射组，纵坐标 HRAMD/总蛋白为每 mg 组织总蛋白中含有多少 ng 阿柏西普蛋白。

表 2. 小鼠眼组织中目的基因蛋白(阿柏西普)的表达量

组织	空白对照	AAV2 seq2	AAV2.7m8
左眼	0.01±0.01	1.10±0.59	0.39±0.17
右眼	0	1.66±1.08	0.44±0.11

图 6、表 2 结果显示，小鼠中 AAV seq2-aflibercept 组的表达量高于 AAV2.7m8-aflibercept 组。

实施例 4. rAAV 载体病毒对兔模型视网膜感染的生物学功能检测

按照实施例 2 中的方法包装载体和纯化 rAAV 载体病毒。

将纯化后的 AAV seq2-aflibercept 和 AAV2.7m8-aflibercept 这两种病毒分别用 DPBS 缓冲液稀释至 4E +9 vg/mL，对周龄为 8-10 周的新西兰兔(购自北京维通利华实验技术有限公司)进行双眼玻璃体注射，注射体积为 50μL，每种 rAAV 载体病毒注射 2 只(n=2)。注射后 2 周，处死兔，取房水、玻璃体和全眼组织，分别研磨，每个注射组左眼和右眼分别用两种不同的方法进行组织研磨处理。左眼使用方法 1：PBS+蛋白酶抑制剂研磨，离心后取上清。右眼使用方法 2：RIP Buffer+蛋白酶抑制剂研磨，离心后取上清。使用酶联免疫分析法检测上清中的阿柏西普蛋白含量(每个注射组 2 个样本,每个样本检测三重复)。结果如图 7 所示。其中空白对照为 PBS 注射组，纵坐标 HRAMD/总蛋白为每 mg 组织总蛋白中含有多少 ng 阿柏西普蛋白。

图 7 结果显示，新西兰兔中 AAV seq2-aflibercept 组的表达量均高于 AAV2.7m8-aflibercept 组。由于兔眼体积较大，阿柏西普在全眼研磨处理后被稀释在组织液中，因此浓度远低于玻璃体和房水。由于 AAV 包装的载体只有在 AAV 病毒成功感染细胞后才会成功表达所携带的目的基因蛋白，因此图 7 结果表明，AAV2 seq2 衣壳对细胞的感染效率要高于 AAV2.7m8 衣壳。

此外，我们还比对了 AAV2 seq2 和 AAV2.7m8 衣壳递送目的基因蛋白(VEGF 抑制剂)的表达量进行了检测。实验方法为：新西兰兔 8 只，按照表 3 剂量分别进行双眼注射，每个剂量组注射 2 只兔，四只眼，共计 4 组。注射后第 28 天处死实验动物并抽取房水，使用 ELISA 方法检测房水样本中的目的基因蛋白的摩尔浓度。

5 结果如图 8 所示，等剂量注射下，AAV2 seq2 所表达的目的基因蛋白的摩尔浓度明显高于 AAV2.7m8。

表 3.

组别	病毒滴度	注射体积	vg/眼
AAV2 seq2 (高)	6E12 vg/mL	100μL	6E11 vg/眼
AAV2 seq2 (低)	2E12 vg/mL	100μL	2E11 vg/眼
AAV2.7m8 (高)	6E12 vg/mL	100μL	6E11 vg/眼
AAV2.7m8 (低)	2E12 vg/mL	100μL	2E11 vg/眼

以下为本公开的部分序列：

10 > AAV2 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
15 NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRQDWQLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
20 VSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGFFGLK
HPPQILIKNTVPANPSTTFSAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYNKS
VNVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 1

25 > AAV2.7m8 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
30 NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRQDWQLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
35 VSTNLQRGNLALGETTRPARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSP
PLMGFFGLKHPPQILIKNTVPANPSTTFSAKFASFITQYSTGQVSVEIEWELQKENSKRWNPEI
QYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 2

> AAV2 seq1 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNLDKGE
5 PVNEADAAALEHDKAYDRQLDSGDNPYLYKNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRDLWQRLINNNWGFRPKRLNFKLFN
10 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALAEETTPARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPS
PLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEI
15 QYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 3

> AAV2 seq2 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNLDKGE
20 PVNEADAAALEHDKAYDRQLDSGDNPYLYKNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRDLWQRLINNNWGFRPKRLNFKLFN
25 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALGDTTPARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPS
PLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEI
30 QYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL*

SEQ ID NO: 4

> AAV2 seq3 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNLDKGE
35 PVNEADAAALEHDKAYDRQLDSGDNPYLYKNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRDLWQRLINNNWGFRPKRLNFKLFN
40 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALGETTRNARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPS
45 PLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEI
QYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 5

> AAV2 seq4 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGE
5 PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRQWQRLINNNWGFRPKRLNFKLFN
10 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLAKADTTKNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHP
SPLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPE
15 IQYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 6

> AAV2 seq5 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGE
20 PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRQWQRLINNNWGFRPKRLNFKLFN
25 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLAKDDTTRNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHP
30 SPLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPE
IQYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 7

> AAV2 seq6 衣壳氨基酸序列

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGE
35 PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRQWQRLINNNWGFRPKRLNFKLFN
40 IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALADTTKNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHPS
45 PLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEI
QYTSNYNKSINVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 8

> AAV9 seq2 衣壳氨基酸序列

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGNGLDKG
 5 EPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKR
 LLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGE
 PPAAPSGVGSMTMASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRVITSTRTWALPTYN
 NHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDNRFHCHFSRWDQRLINNWGFRPKRLNFKL
 FNIQVKEVTDNNGVKTIANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTL
 10 NDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYYL
 SKTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWPGASSWAL
 NGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATES
 YGQVATNHQSAQLALGDTTRPAAQAQTGWVQNQGILPGMVWQDRDVYLGQPIWAKIPHTDGN
 FHPSPLMGFGFMKHPPPQILIKNTVPADPPTAFNKKLNSFITQYSTGQVSVEIEWELQKENSKR
 15 WNPEIQYTSNYYKSNNVEFAVNTGCVYSEPRPIGTRYLTRNL

SEQ ID NO: 9

> EGFP 氨基酸序列

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLLKFICTTGKLPVPWPTLVTTL
 TYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGI
 20 DFKEDGNILGHKLEYNNSHNHYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGP
 VLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITLGMDELYK

SEQ ID NO: 10

> seq1(无连接子)

LAETTRP
 25

SEQ ID NO: 11

> seq2(无连接子)

LGDTRP
 30

SEQ ID NO: 12

> seq3(无连接子)

LGETTRN
 35

SEQ ID NO: 13

> seq4(无连接子)

KADTTKN
 40

SEQ ID NO: 14

> seq5(无连接子)

KDDTRN
 45

SEQ ID NO: 15

> seq6(无连接子)

LADTTKN
 50

SEQ ID NO: 16

> seq1

LALAETTRPA

SEQ ID NO: 17

> seq2
LALGDTTRPA

SEQ ID NO: 18

5 > seq3
LALGETTRNA

SEQ ID NO: 19

> seq4
LAKADTTKNA

10 > seq5
LAKDDTTRNA

SEQ ID NO: 20

> seq6
LALADTTKNA

15

SEQ ID NO: 21

SEQ ID NO: 22

> AAV2 seq2 衣壳氨基酸序列(不包含 V708I)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
20 EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
25 NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVPNGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALGDTTRPARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHP
PLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS
KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

30

SEQ ID NO: 23

> AAV2 seq3 衣壳氨基酸序列(不包含 V708I)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQP
35 PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
40 RDSLVPNGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALGETTRNARQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHP
PLMGGFGLKHPPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS
KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO: 24

> AAV2 seq4 衣壳氨基酸序列(不包含 V708I)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
5 PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
10 RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLAKADTTKNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHP
SPLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSkrwnPE
IQYTSNYNKSvNVDFTVDTNGVYSEPRIGTRYLTRNL

SEQ ID NO: 25

15

> AAV2 seq5 衣壳氨基酸序列(不包含 V708I)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
20 EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
25 NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLAKDDTTRNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHP
SPLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSkrwnPE
IQYTSNYNKSvNVDFTVDTNGVYSEPRIGTRYLTRNL

SEQ ID NO: 26

30

> AAV2 seq6 衣壳氨基酸序列(不包含 V708I)

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNGLDKGE
PVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVL
EPLGLVEEPVKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQP
35 PAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYN
NHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFN
IQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNN
GSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRT
NTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNG
40 RDSLVNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGS
VSTNLQRGNLALADTTKNARQAATADVNTQGVLPGMVWQDRDVYLQGPIWAKIPHTDGHFHPS
PLMGGFGLKHPPQILIKNTVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENSkrwnPEI
QYTSNYNKSvNVDFTVDTNGVYSEPRIGTRYLTRNL

SEQ ID NO: 27

45

> CMV 增强子序列

CGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCCGCCATTGAC

GTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG
GAGTATTTACGGTAAACTGCCCCTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGA
CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG

5

SEQ ID NO : 28

>CMV 启动子序列

TGCTGATGCGGTTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC
CAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTTC
CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGACGGTGGGAG
10 GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGAGGCCATCCACGCTGT
TTTGACCTCCATAGTGGACACCGGGACCGATCCAGCCTCCGCGTCTCAGGGGAAGCTT

SEQ ID NO : 29

>5'UTR 序列

CTTGTTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG

15

SEQ ID NO : 30

>Kozak 序列

GCCACC

SEQ ID NO : 31

>3'UTR 序列

20 GTCTAGAAACCAGCCTCAAGAACACCCGAATGGAGTCTCTAAGCTACATAATACCAACTTACA
CTTTACAAAATGTTGTCCCCCAAATGTAGCCATTCGTATCTGCTCCTAATAAAAAGAAAGTTT
CTTCAC

SEQ ID NO : 32

>WPRE 序列

25 AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAAGTATGTTGCTCCTT
TTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCA
TTTTCTCCTCCTTGATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAG
GCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCACTGGTTGGGGCATTGCCA
CCACCTGTCAGCTCCTTTCCGGGACTTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACATCAT
30 CGCCGCTGCTTGGCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG
TGTTGTGCGGGGAAATCATCGTCCTTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCG
CGGGACGTCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCT
GCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCTTCGCCCCTCAGACGAGTCGGATCTCCCT
TTGGGCCGCTCCCCGC

35

SEQ ID NO : 33

>SV40polyA 序列:

CCTAGGTAAGATACATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTGAATAAATGCT
TTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTAAGCTGCAATAAACAAGTT
ATCGATAGATCT

40

SEQ ID NO : 34

>AAV9 衣壳氨基酸序列

MAADGYLPDWLEDNLSEGIREWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGPNGLDKG
EPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKR
LLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTESVPDPQPIGE
PPAAPSGVGLTMASSGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRTVTTSTRTWALPTYN
5 NHLYKQISNSTSGSSNDNAYFGYSTPWGYFDNRFHCHFSRDLNINNNGFRPKRLNFKL
FNIQVKEVTDNNGVKTIANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTL
NDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQLDRLMNPLIDQYLYL
SKTINGSGQNQQLKFSVAGPSNMAVQGRNYPGPSYRQQRVSTTVTQNNNEFAWPGASSWAL
NGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATES
10 YGQVATNHQSAQAQAQTGWVQNGILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPLMGGF
GMKHPPQILIKNTPVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSN
YYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

SEQ ID NO: 37

>阿柏西普蛋白序列

15 MPLLLLLPLWAGALASDTGRPFVEMYSEIPEIIHMTGRELVIPCRVTSPNITVTLKKFPLDTLIPD
GKRIIWDNRKGFHISNATYKEIGLLTCEATVNGHLYKTNLYTHRQNTIIDVVLSPSHGIELSVGEKL
VLNCTARTELVNVDGDFNWEYPSSKHQHKKLVRDLKTQSGSEMKKFLSTLTIDGVTRSDQGLYTC
AASSGLMTKKNSTFVRVHEKDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD
VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKA
20 LPAPIEKTIKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT
TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSVSMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 38

>阿柏西普核酸序列 1

ATGCCCCCTGCTCCTGCTCCTGCCCCCTGCTGTGGGCGCGGCCCTGGCTAGCGACACCGGCAG
25 ACCCTTCGTGGAGATGTACAGCGAGATCCCCGAGATCATCCACATGACCGAGGGCAGAGAGC
TGGTGATCCCCCTGCAGAGTGACAAGCCCCAACATCACCGTGACCCTGAAAAAATTCCTCTG
GACACCCTGATCCCCGACGGCAAGAGAATCATCTGGGACAGCAGAAAGGGCTTCATCATCAG
CAACGCCACCTACAAGGAGATCGGCCCTGCTGACCTGCGAGGCCACCGTGAACGGCCACCTGT
ACAAGACCAACTACCTGACCCACAGACAGACCAACACCATCATCGACGTGGTGCTGAGCCCT
30 AGCCACGGCATCGAGCTGAGCGTGGGCGAGAAGCTGGTGCTGAAGTGCACCGCTAGAACCG
AGCTGAACGTGGGCATCGACTTCAACTGGGAGTACCCTAGCAGCAAGCATCAGCACAAGAAG
CTGGTGAAACAGAGACCTGAAGACACAGAGCGGCAGCGAGATGAAGAAGTTCCTGAGCACCC
TGACCATCGACGGCGTGACAAGAAGCGACCAAGGCCTGTACACCTGCGCCGCTAGCAGCGG
CCTGATGACCAAGAAGAAGACAGCACCTTCGTGAGAGTGCATGAGAAGGACAAGACCCACACC
35 TGCCCCCCCCCTGCCCCGCTCCTGAAGTGTGGGCGGCCCTAGCGTGTTCTGTTCCCCCCCCAAG
CCCAAGGACACACTCATGATCAGCAGAACCCCCGAGGTGACCTGCGTGGTCTGGACGTGAG
CCACGAGGACCCCGAGGTGAAGTTCAACTGGTACGTGGACGGCGTGAGGTGCACAACGCC
AAGACCAAGCCTAGAGAGGAGCAGTACAACAGCACCTACAGAGTGGTGAGCGTGCTGACCG
TGCTGCACCAAGACTGGCTGAACGGCAAGGAGTACAAGTGCAAGGTGAGCAACAAGGCCCT
40 GCCCGCCCCCATCGAGAAGACCATCAGCAAGGCCAAGGGGCAGCCTAGAGAGCCCCAAGTG
TACACCCTGCCCCCTAGCAGAGACGAGCTGACCAAGAACCAAGTGAGCCTGACCTGCCTGGT
GAAGGGCTTCTACCTAGCGACATCGCCGTGGAGTGGGAGAGCAACGGGCAGCCCGAGAAC
AACTACAAGACCACCCCCCGTGCTGGACAGCGACGGCAGCTTCTTCTGTACAGCAAGCT
GACCGTGGACAAGAGCAGATGGCAGCAAGGCAACGTGTTAGCTGCAGCGTGATGCACGAG
45 GCCCTGCACAACCACTACACACAGAAGAGCCTGAGCCTGAGCCCCGCAAGTGA

SEQ ID NO: 39

>阿柏西普核酸序列 2

ATGCTCTGCTCCTGCTGCTCCCTCTGCTGTGGGCGGCGCCCTGGCCAGCGACACAGGCCG
CCCCTTCGTGGAAATGTACAGCGAGATCCCTGAGATCATCCACATGACAGAGGGCCGGGAAC
TGGTCATCCCTGTAGAGTGACATCTCCAAACATCACCGTGACCCTGAAAAAGTTCCCCCTGG
ACACACTGATCCCTGACGGCAAAAGAATCATCTGGGATTCCCGGAAGGGCTTCATCATCAGCA
5 ATGCCACCTACAAGGAAATCGGACTGCTGACCTGTGAAGCCACAGTGAACGGCCACCTGTAC
AAGACCAACTACCTGACCCATAGGCAGACCAATACCATCATCGATGTGGTGCTGTCTCCTTCT
CACGGCATCGAGCTCAGCGTCGGCGAGAACTGGTGCTGAACTGCACCGCCAGAACCGAGC
TGAACGTGGGCATCGACTTCAACTGGGAATACCCAGTTCTAAGCACCAGCACAAAGAACTG
GTTAATAGAGATCTGAAGACCCAAAGCGGCAGCGAGATGAAAAAGTTTCTGTCTACACTGAC
10 AATTGACGGTGTTACCAGAAAGCGATCAGGGCCTGTATACATGCGCCGCTTCTAGCGGCCTGAT
GACCAAGAAGAACAGCACCTTCTGTGCGGGTGCACGAGAAGGACAAGACCCACACCTGCCCC
CCCTGTCCAGCTCCTGAACTGCTGGGAGGCCCTAGCGTGTTCTGTTCCTCCAAAGCCTAAG
GACACCCCTATGATTAGCAGAACACCTGAGGTGACCTGCGTGCTGGTTCGACGTGTCCACGA
GGACCCCGAGGTGAAATTTAACTGGTACGTGGATGGCGTGGAAGTTCACAACGCCAAGACCA
15 AGCCCAGAGAGGAACAGTACAACAGCACATATAGAGTGGTGTCCGTGCTGACCGTTCTGCAC
CAGGACTGGCTGAATGGAAAAGAGTACAAGTGCAAGGTGTCCAACAAGGCCCTGCCTGCCC
CTATCGAGAAGACTATCAGCAAGGCTAAGGGCCAACCTCGGGAACCCAGGTGTACACCCTG
CCCCCAGCCGGGACGAGCTGACAAAGAACCAGGTGTCCCTGACATGCCTGGTGAAGGGATT
TTACCCTAGCGATATCGCCGTGGAATGGGAGAGCAACGGCCAGCCTGAGAACAATAAAAA
20 CGACCCCTCCTGTGCTGGACTCTGATGGCAGCTTCTTCTGTACAGCAAGCTGACCGTGGACA
AGAGCAGATGGCAGCAGGGCAACGTGTTCTCATGCAGCGTGATGCACGAGGCCCTGCATAAT
CACTACACCCAGAAAAGCCTGAGCCTGAGCCCTGGCAAGTGA

SEQ ID NO: 40

>阿柏西普核酸序列 3

25 ATGCCACTGCTCTTGCTTTTGGCCCTTGCTTTGGGCTGGTGCCCTTGCTCAGACACTGGAAGG
CCGTTCTAGAAATGTACTCCGAGATCCCTGAAATAATCCACATGACGGAGGGGCGAGAGTTG
GTGATACCATGTGAGTTACGAGCCGAATATTACAGTAACATTGAAGAAATTCCCACTCGAC
ACTCTCATTCCAGATGGTAAGAGAATTATATGGGATTACAGAAAAGGGCTTCATTATATCTAACG
CTACTTACAAAGAGATAGGCTTGCTCACATGCGAGGCAACTGTCAATGGCCACTTGTATAAAA
30 CCAATTATCTCACACATAGGCAGACTAATACGATAATAGACGTTGTCCTTTCTCCGAGCCATGG
CATTGAATTGAGTGTGGTGAGAAGCTCGTATTGAACTGCACTGCTAGGACCGAGCTTAACGT
AGGGATCGACTTTAACTGGGAGTACCCATCATCAAAGCATCAGCATAAAAAGCTCGTCAACCG
GGACCTTAAGACGCAGAGTGGCTCAGAGATGAAGAAGTTCTTTCTACACTGACCATCGATG
GTGTAACCAGAAGTGACCAGGGACTCTATACCTGTGCCGCATCCTCCGGCCTTATGACGAAGA
35 AAAACAGCACATTTGTTTCGGGTCCACGAGAAGGACAAGACTCATACTGTCCGCCGTGTCCA
GCCCCAGAATTGCTCGGGGGGCCCTCCGTGTTTCTGTTCCCTCCAAAGCCAAAGGACACGCTT
ATGATTTCAAGGACTCCCGAAGTCACCTGCGTGTGGTTGACGTTTCTCATGAGGACCCAGAG
GTGAAATTCAACTGGTACGTGGATGGAGTAGAGGTACATAACGCCAAAAACAAACCACGAGA
AGAGCAGTATAACTCAACGTATCGGGTAGTTTCTGTGCTTACGGTTCTCCACCAAGATTGGTT
40 GAATGGAAGAGATATAAGTGTAAGTCTCAAACAAAGCCCTGCCGGCACCCATCGAGAAGA
CCATTAGCAAAGCAAAAGGCCAACCTAGAGAACCGCAGGTGTACACGTTGCCACCGTCAAG
GGATGAACTTACGAAGAATCAAGTATCACTCACATGTTTGGTTAAAGGATTTATCCCAGCGA
CATCGCCGTAGAATGGGAGAGTAACGGCCAGCCTGAAAATAACTATAAGACGACACCCCTG
TTTTGGATTCTGACGGTTCATTTTTTTGTATAGTAACTCACAGTCGACAAGAGTCGCTGGCA
45 ACAAGGTAATGTATTTAGCTGTAGCGTGATGCATGAAGCTTTGCACAACCACTATACAGAA
GTCCTTGCTCTGTCTCCGGGAAAATGA

SEQ ID NO: 41

权利要求书:

1. 一种变体腺相关病毒(AAV)衣壳蛋白,其相对于亲本 AAV 衣壳蛋白包含插入的多肽,所述插入的多肽包含选自如下 1)-6)任一项所示的多肽或其任意组合:

- 5 1) LGDTTRP(SEQ ID NO: 12)或 LALGDTTRPA(SEQ ID NO: 18);
 2) LAETTRP(SEQ ID NO: 11)或 LALAETTRPA(SEQ ID NO: 17);
 3) LGETTRN(SEQ ID NO: 13)或 LALGETTRNA(SEQ ID NO: 19);
 4) KADTTKN(SEQ ID NO: 14)或 LAKADTTKNA(SEQ ID NO: 20);
 5) KDDTTRN(SEQ ID NO: 15)或 LAKDDTTRNA(SEQ ID NO: 21);
10 6) LADTTKN(SEQ ID NO: 16)或 LALADTTKNA(SEQ ID NO: 22)。

2. 如权利要求 1 所述的变体 AAV 衣壳蛋白,所述 1)-6)任一项的多肽或其任意组合位于所述 AAV 衣壳蛋白的 GH 环中;

 优选地,位于亲本 AAV2 衣壳蛋白的 VP1 编码氨基酸序列的第 570 位至 611
15 位之间的任一个氨基酸残基之后;

 更优选地,位于亲本 AAV2 衣壳蛋白的 VP1 编码氨基酸序列的第 587 位和 588 位之间,或亲本 AAV9 衣壳蛋白的 VP1 编码氨基酸序列的第 588 位和 589 位之间,或其他亲本 AAV 血清型衣壳蛋白的相应位置。

20 3. 如权利要求 1 或 2 所述的变体 AAV 衣壳蛋白,其进一步包含选自 1L、15P、34A、57D、66K、81Q、101R、109T、144K、144M、164K、176P、188I、196Y、226E、236V、240T、250S、312K、363L、368H、449D、456K、463Y、472N、484C、524T、535S、551S、593E、698V、708I、719M、721L 和 735Q 中的一个或多个氨基酸残基的点突变,所述突变的位置基于 SEQ ID NO: 1 所示的 AAV2
25 的衣壳蛋白,或其他 AAV 血清型衣壳蛋白的相应位置;

 优选地,所述点突变是 708I。

4. 如权利要求 1-3 中任一项所述的变体 AAV 衣壳蛋白,其氨基酸序列为 SEQ ID NO: 3-9 和 23-27 中任一所示或与之具有至少 90%或 95%的序列同一性。

30

5. 一种重组腺相关病毒(rAAV)病毒粒子,其包含:

- (a) 权利要求 1-4 任一项所述的变体 AAV 衣壳蛋白,和
(b) 异源多核苷酸。

35 6. 如权利要求 5 所述的 rAAV 病毒粒子,其中所述异源多核苷酸包含编码基因产物的多核苷酸,所述基因产物选自多肽、干扰 RNA 或适配体;

 所述多肽优选为抗体或其抗原结合片段、融合蛋白。

7. 如权利要求 6 所述的 rAAV 病毒粒子，其中所述异源多核苷酸进一步包含选自以下(a)至(g)的一个或多个：

- (a) 5'反向末端重复(5'ITR)和/或 3'反向末端重复(3'ITR)，
 - 5 (b) 5'非翻译区(5'UTR)和/或 3'非翻译区(3'UTR)，
 - (c) 启动子，
 - (d) 增强子，
 - (e) 转录后调控元件，
 - (f) 多聚腺苷酸化信号(polyA)，和
 - 10 (g) Kozak 序列；
- 优选地，所述异源多核苷酸进一步包含选自以下(a)至(g)的一个或多个：
- (a) 源自 AAV2 的 5'ITR 和/或 3'ITR，
 - (b) 源自 *Xenopus globin* 的 5'UTR 和/或 3'UTR，
 - (c) CMV 启动子，
 - 15 (d) CMV 增强子，
 - (e) WPRE，
 - (f) SV40 polyA，和
 - (g) Kozak 序列。

20 8. 如权利要求 7 所述的 rAAV 病毒粒子，其中所述异源多核苷酸从 5'端到 3'端包含如下可操作地连接并按如下(a)至(h)顺序排列的多核苷酸：

- (a) 增强子，
- (b) 启动子，
- (c) 5'UTR，
- 25 (d) Kozak 序列，
- (e) 编码基因产物的多核苷酸，
- (f) 3'UTR，
- (g) WPRE，
- (h) polyA。

30 9. 如权利要求 6-8 中任一项所述的 rAAV 病毒粒子，所述多肽为神经保护多肽、抗血管生成多肽、增强视网膜细胞功能的多肽；

优选地，所述抗血管生成多肽为 VEGF 拮抗剂；

更优选地，所述 VEGF 拮抗剂为阿柏西普(aflibercept)。

35 10. 如权利要求 5-9 中任一项所述的 rAAV 病毒粒子，所述异源多核苷酸包含

编码 SEQ ID NO: 38 所示的阿柏西普(aflibercept)的多核苷酸，

优选地，所述异源多核苷酸包含 SEQ ID NO: 39-41 中任一所示或与之具有至少 90% 或 95% 序列同一性的多核苷酸。

5 11. 分离的多核苷酸，其编码权利要求 1-4 中任一项所述的变体 AAV 衣壳蛋白。

12. 载体，其包含：

(a) 权利要求 11 所述的分离的多核苷酸，或

10 (b) 权利要求 11 所述的分离的多核苷酸和权利要求 5-10 中任一项所述的异源多核苷酸。

13. 宿主细胞，其包含权利要求 12 所述的载体。

15 14. 药物组合物，其含有

(a) 权利要求 5-10 中任一项所述 rAAV 病毒粒子；和

(b) 一种或多种药学上可接受的载剂、稀释剂、赋形剂或缓冲液。

20 15. 特异性感染视网膜细胞的方法，所述方法包括眼内注射有效量的根据权利要求 5-10 中任一项所述的 rAAV 病毒粒子、权利要求 5-10 中任一项所述的异源多核苷酸或权利要求 14 所述的药物组合物的步骤；优选通过玻璃体内注射或视网膜下注射。

25 16. 治疗眼部相关疾病的方法，包括向有需要的受试者施用预防或治疗有效量的权利要求 5-10 中任一项所述的 rAAV 病毒粒子、权利要求 5-10 中任一项所述的异源多核苷酸或权利要求 14 所述的药物组合物；优选通过玻璃体内注射或视网膜下注射。

30 17. 如权利要求 16 所述的方法，其中所述眼部相关疾病为视网膜细胞疾病；优选地，所述视网膜细胞疾病选自感光细胞、视网膜神经节细胞、Muller 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疾病。

35 18. 如权利要求 16 或 17 所述的方法，其中所述眼部相关疾病选自：视网膜色素变性、黄斑变性、湿性 AMD、干性 AMD、视网膜新生血管、脉络膜新生血管、糖尿病性视网膜病变、增生性糖尿病视网膜病变、视网膜静脉阻塞、视网膜中央静脉阻塞、视网膜分支静脉阻塞、糖尿病性黄斑水肿、糖尿病性视网膜缺血、缺血性视网膜病变、糖尿病性视网膜水肿、视网膜劈裂症、青光眼、利伯氏先天性

黑朦和/或色盲。

19. 在体外和/或体内向靶细胞递送异源多核苷酸的方法，其包括使所述靶细胞与根据权利要求 5-10 中任一项的 rAAV 病毒粒子、权利要求 5-10 中任一项所述的异源多核苷酸或权利要求 14 所述的药物组合物接触。

20. rAAV 病毒粒子生产系统，包含：

(a) 如权利要求 11 所述的分离的多核苷酸；

(b) 编码基因产物的异源多核苷酸；和

10 (c) 辅助元件，其具有足够的 AAV rep 功能和辅助功能，以将(b)的编码基因产物的异源多核苷酸包装到(a)的变体 AAV 衣壳中。

21. 生产或制备权利要求 5-10 中任一项所述 rAAV 病毒粒子的方法，包括：将权利要求 12 所述的载体、包含辅助元件的载体引入生产细胞中，包装、纯化所述 rAAV 病毒粒子。