



(10) 申请公布号 CN 116670152 A

(43) 申请公布日 2023. 08. 29

(21) 申请号 202180085871.6

(22) 申请日 2021.12.01

(30) 优先权数据

63/119863 2020.12.01 US

(85) PCT国际申请进入国家阶段日

2023.06.19

(86) PCT国际申请的申请数据

PCT/US2021/061312 2021.12.01

(87) PCT国际申请的公布数据

W02022/119871 EN 2022.06.09

(71) 申请人 宾夕法尼亚州大学信托人

地址 美国宾夕法尼亚州

(72) 发明人 J·M·威尔逊 J·J·西姆斯

Y·元

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

专利代理师 黄登高 梅黎

(51) Int.Cl.

C07K 14/005 (2006.01)

权利要求书2页 说明书46页
序列表130页 附图11页

(54) 发明名称

具有组织特异性靶向基序的新型组合物和含有其的组合

(57) 摘要

本文提供了组合物,所述组合物包含与重组载体的靶向蛋白连接或插入重组载体的靶向蛋白中的靶向肽,所述重组载体具有至少一个外源肽,所述外源肽包括N-x-(T/I/V/A)-(K/R) (SEQ ID NO:47)的氨基酸序列。提供了提供具有突变衣壳或包膜蛋白的此类缀合物、靶向肽或重组载体的组合物以及其用途。

1. 一种重组腺相关病毒颗粒(rAAV),其具有包括氨基酸序列的衣壳,所述氨基酸序列包括基序N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47),其中所述氨基酸序列至少是所述衣壳中AAV vp3蛋白和包装在所述衣壳中的载体基因组的一部分,所述载体基因组包括在指导基因产物表达的序列控制下编码所述基因产物的核酸序列,条件是所述衣壳不是包括NDVRAVS(SEQ ID NO:48)序列的突变AAV2衣壳。

2. 根据权利要求1所述的rAAV,其中将包括所述N-x-(T/I/V/A)-(K/R)基序的所述氨基酸序列插入AAV衣壳vp3区域,任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸。

3. 根据权利要求1或2所述的rAAV,其中插入所述衣壳的所述序列包括:

- (a) SSNTVKLTSGH(SEQ ID NO:40);
- (b) EFSSNTVKLTS(SEQ ID NO:38);
- (c) GGVLNIARGEYMRGG(SEQ ID NO:46);
- (d) GGIEINATRAGTNLGG(SEQ ID NO:43);
- (e) GGSSNTVKLTSGHGG(SEQ ID NO:39);
- (f) IEINATRAGTNL(SEQ ID NO:42);或
- (g) SANFIKPTSY(SEQ ID NO:41)。

4. 根据权利要求1至3中任一项所述的rAAV,其中所述基序的所述氨基酸序列是NTVK。

5. 根据权利要求1所述的rAAV,其中任选在所述基序N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)的羧基和/或氨基末端侧接两个至七个氨基酸并且插入AAV9衣壳蛋白的氨基酸588和589之间,基于以下氨基酸序列的编号:SEQ ID NO:44。

6. 一种组合物,其包括根据权利要求1至5中任一项所述的rAAV的原液以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

7. 一种内皮细胞靶向肽,其中所述内皮细胞靶向肽包括基序,所述基序包括N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)的氨基酸序列,任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸并且任选进一步与纳米颗粒、第二分子或病毒衣壳蛋白缀合。

8. 根据权利要求7所述的内皮细胞靶向肽,其中所述内皮细胞靶向肽包括:

- (a) SSNTVKLTSGH(SEQ ID NO:40);
- (b) EFSSNTVKLTS(SEQ ID NO:38);
- (c) GGVLNIARGEYMRGG(SEQ ID NO:46);
- (d) GGIEINATRAGTNLGG(SEQ ID NO:43);
- (e) GGSSNTVKLTSGHGG(SEQ ID NO:39);
- (f) IEINATRAGTNL(SEQ ID NO:42);或
- (g) SANFIKPTSY(SEQ ID NO:41)。

9. 根据权利要求7或8所述的内皮细胞靶向肽,其中所述基序的所述氨基酸序列是NTVK。

10. 一种组合物,其包括根据权利要求7至9中任一项所述的内皮细胞靶向肽以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

11. 一种融合多肽或蛋白质,其包括根据权利要求7至9中任一项所述的脑内皮细胞靶向肽和包括至少一种多肽或蛋白质的融合伴侣。

12. 一种组合物,其包括根据权利要求11所述的融合多肽或蛋白质以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

13. 根据权利要求1至5中任一项所述的rAAV的原液、根据权利要求7至9中任一项所述的内皮细胞靶向肽、或根据权利要求11所述的融合多肽或蛋白质、或根据权利要求6、10或12中任一项所述的组合物用于向有此需要的患者递送治疗的用途。

14. 一种用于对脑内皮细胞的靶向疗法的方法,所述方法包括向有此需要的患者施用根据权利要求1至5中任一项所述的rAAV的原液。

15. 一种用于通过向有此需要的受试者递送根据权利要求1至5中任一项所述的rAAV的原液来治疗Allan-Herndon-Dudley病的方法,其中所述编码的基因产物是MCT8蛋白。

16. 一种用于对肺的靶向疗法的方法,其包括向有此需要的患者施用根据权利要求1至5中任一项所述的rAAV的原液。

17. 一种用于通过向有此需要的受试者递送根据权利要求1至5中任一项所述的rAAV的原液来治疗肺疾病的方法,其中所述编码的基因产物是可溶性Ace2蛋白、抗SARS抗体、抗SARS-CoV2抗体、抗流感抗体或囊性纤维化跨膜蛋白。

18. 一种用于增加体外AAV生产细胞转导的方法,其包括将N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)基序插入AAV衣壳中。

19. 根据权利要求16所述的方法,其中所述生产细胞是293细胞。

具有组织特异性靶向基序的新型组合物和含有其的组合物

背景技术

[0001] 腺相关病毒(AAV)目前是精选的基因疗法载体。这是因为AAV可以从非整合基因组中递送稳定表达几十年的转基因,并且因为AAV是非常安全的且没有免疫原性。然而,由于递送和趋向性方面的挑战,AAV基因疗法目前局限于少数疾病。这对于中枢神经系统(CNS)病症尤其如此。可以通过将载体直接注射到脑脊液(CSF)中来直接递送AAV基因疗法载体,但这种方法通常转导1%或更少的脑细胞。此外,转导大多集中在与CSF直接接触的细胞上。“脑深部”中的细胞很少被转导。这限制了可通过基因疗法治疗的CNS病症的数量。

[0002] 与CSF网络相比,脑的血管系统到达CNS中的几乎每一个细胞。这是因为这些组织对葡萄糖、氧和其它营养物质的需求很大。然而,脑和脊髓中的细胞受到专门的血管单元(血脑屏障(BBB))的保护,不受循环系统影响。BBB限制了大分子如病毒载体和蛋白质的扩散,甚至限制了许多小分子药物通过围绕脑和脊髓血管的紧密连接的复杂细胞网络扩散。因此,工程化能够高效穿越BBB并转导脑深部细胞的AAV变体已经成为向CNS递送基因疗法的一个巨大挑战。

[0003] 加州理工学院(CalTech)开发的一种AAV衣壳在AAV9衣壳上的高变环8(HVR8)中插入了七个氨基酸的肽以生成称为AAV9-PHP.B的rAAV。据报道,rAAV介导与Ly6a(一些小鼠品系的脑血管上的GPI锚定受体)的相互作用。美国专利公开申请号2017/0166926A1。这种相互作用驱动AAV9-PHP.B跨越BBB运输,导致脑细胞的转导是AAV9的约50倍。然而,这一发现还没有转化至更大的动物或人类。

[0004] 仍然需要可以特异性靶向所选组织和细胞类型的载体。

发明内容

[0005] 在某些实施例中,提供了一种重组腺相关病毒颗粒(rAAV),其具有包括氨基酸序列的衣壳,所述氨基酸序列包括基序N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)。适当地,氨基酸序列至少是衣壳中AAV vp3蛋白和包装在衣壳中的载体基因组的一部分,所述载体基因组包括在指导基因产物表达的序列控制下编码所述基因产物的核酸序列,条件是所述衣壳不是包括NDVRAVS(SEQ ID NO:48)序列的突变AAV2衣壳。在某些实施例中,将包括N-x-(T/I/V/A)-(K/R)基序的氨基酸序列插入AAV衣壳vp3区域,任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸。在某些实施例中,插入衣壳的序列包括:(a) SSNTVKLTSGH(SEQ ID NO:40);(b) EFSSNTVKLTS(SEQ ID NO:38);(c) GGVLTNIARGEYMRGG(SEQ ID NO:46);(d) GGIEINATRAGTNLGG(SEQ ID NO:43);(e) GGSSNTVKLTSGHGG(SEQ ID NO:39);(f) IEINATRAGTNL(SEQ ID NO:42);或(g) SANFIKPTSY(SEQ ID NO:41)。在某些实施例中,基序的氨基酸序列是NTVK,任选在其羧基和/或氨基末端侧接两个至七个氨基酸并且插入AAV9衣壳蛋白的氨基酸588和589之间,基于以下氨基酸序列的编号:SEQ ID NO:44。

[0006] 在某些实施例中,rAAV在其衣壳中具有插入的NTVK序列,所述序列任选在其羧基和/或氨基末端侧接两个至七个氨基酸并且插入AAV9衣壳蛋白的氨基酸588和589之间,基于以下氨基酸序列的编号:SEQ ID NO:44。

[0007] 在某些实施例中,组合物包括具有插入基序和任选侧翼序列的rAAV以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

[0008] 在某些实施例中,提供了一种内皮细胞靶向肽,所述肽包括N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)的氨基酸序列,其任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸并且任选进一步与纳米颗粒、第二分子或病毒衣壳蛋白缀合。在某些实施例中,内皮细胞靶向肽包括:(a) SSNTVKLTSGH(SEQ ID NO:40);(b) EFSSNTVKLTS(SEQ ID NO:38);(c) GGVLNIARGEYMRGG(SEQ ID NO:46);(d) GGIEINATRAGTNLGG(SEQ ID NO:43);(e) GGSSNTVKLTSGHGG(SEQ ID NO:39);(f) IEINATRAGTNL(SEQ ID NO:42);或(g) SANFIKPTSY(SEQ ID NO:41)。在某些实施例中,基序的氨基酸序列是NTVK。在某些实施例中,提供了一种组合物,其包括内皮细胞靶向肽以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

[0009] 在某些实施例中,本文提供了一种融合多肽或蛋白质,其包括脑内皮细胞靶向肽和包括至少一种多肽或蛋白质的融合伴侣。在某些实施例中,组合物包括根据权利要求11所述的融合多肽或蛋白质以及生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

[0010] 本文提供了用于使用rAAV、内皮细胞靶向肽、融合多肽或蛋白质和/或本文所述的组合物向有此需要的患者递送治疗的组合物和方法。在某些实施例中,使治疗靶向脑内皮细胞。

[0011] 在某些实施例中,提供了一种用于通过向有此需要的受试者递送本文所述的rAAV来治疗Allan-Herndon-Dudley病的组合物和方法,其中编码的基因产物是MCT8蛋白。

[0012] 在某些实施例中,提供了一种用于肺的靶向疗法的方法,其包括向有此需要的患者施用本文所述的rAAV。

[0013] 在某些实施例中,提供了一种用于通过向有此需要的受试者递送rAAV来治疗肺疾病的方法,所述rAAV具有带有插入靶向肽的衣壳并编码治疗性基因产物,其中编码的基因产物是可溶性Ace2蛋白、抗SARS抗体、抗SARS-CoV2抗体、抗流感抗体或囊性纤维化跨膜蛋白。

[0014] 在某些实施例中,提供了一种用于增加AAV生产细胞体外转导的方法,其包括将N-x-(T/I/V/A)-(K/R)基序插入AAV衣壳。在某些实施例中,生产细胞是293细胞。

[0015] 根据说明书(包含但不限于本发明的详细描述),本发明的这些和其它实施例和优点将是显而易见的。

附图说明

[0016] 图1A至1B显示了筛选中小鼠脑中表现最好的肽命中以及参考肽的富集分数。图1A显示了C57BL/6J小鼠的富集分数。图1B显示了Ba1b/c小鼠的富集分数。

[0017] 图2A和2B显示了筛选中NHP组织中表现最好的命中的富集分数。图2A显示了NHP脑的富集分数。图2B显示了NHP脊髓组织的富集分数。

[0018] 图3A至3D显示了在包括GFP报告转基因的AAV衣壳中表现最好的肽命中的转导水平的二次验证。结果相对于AAV9转导绘制。图3A显示了在Ba1b/c小鼠中脑组织的所选肽靶向的二次验证筛选。图3B显示了在C57BL/6小鼠中脑组织的所选肽靶向的二次验证筛选。图

3C显示了在Ba1b/c小鼠中肝组织的所选肽靶向的二次验证筛选。图3D显示了在C57BL/6小鼠中肝组织的所选肽靶向的二次验证筛选。

[0019] 图4显示了AA9、AAV8、AAV7、AAV6、AAV5、AAV4、AAV3B、AAV2和AAV1的各种AAV衣壳蛋白的氨基酸序列的比对区域,其侧重于可能插入有靶向肽的区域HVRVIII(基于结构分析)。

[0020] 图5显示了“NxTK”基序是SAN插入物中脑生物分布的关键基序,并且显示了取代的平均影响(与原始序列相比的变化倍数)。

[0021] 图6显示了“NxTK”基序控制SAN肽插入物中的质粒到AAV转化,并且显示了取代的平均影响(与原始序列相比的变化倍数)。

[0022] 图7A至7D显示了“NxTK”基序赋予跨越细胞系的广泛转导优势。图7A显示了与293细胞中AAV9衣壳相比的相对转导水平。图7B显示了与NIH3T3细胞中AAV9衣壳相比的相对转导水平。图7C显示了与HUH7细胞中AAV9衣壳相比的相对转导水平。图7D显示了猕猴原代气道细胞在转导后第3天(3DPT)和转导后第7天(7DPT)的转导水平。图7E显示了用载剂处理的对照样品(即无载体)中猕猴原代气道上皮细胞的显微镜分析。图7F显示了用AAV9-GFP载体转导后的猕猴原代气道上皮细胞的显微镜分析。图7G显示了用包括EFS肽插入物的AAV9-GFP载体转导后的猕猴原代气道上皮细胞的显微镜分析。图7H显示了用包括SAN肽插入物的AAV9-GFP转导后的猕猴原代气道上皮细胞的显微镜分析。

[0023] 图8显示了用GFP载体在培养的人细胞(鼻、支气管和气管)中进行的初步转导测试,所述测试以mRNA拷贝数与以微克计的总mRNA的比率来绘制。

具体实施方式

[0024] 本文提供了一种靶向肽序列。本文还提供了与外源靶向肽基序N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)可操作地连接的融合蛋白、修饰蛋白、突变病毒衣壳和其它部分。在某些实施例中,这种外源基序赋予这些组合物对源(亲本)蛋白、病毒载体或其它部分的天然组织特异性进行调节。在某些实施例中,这种基序中的靶向肽提供了增强或改变的內皮细胞靶向。在某些实施例中,这种基序中的靶向肽提供了增强或改变的肺、支气管、气管和/或鼻上皮靶向。在某些实施例中,具有带有这种基序的修饰衣壳的病毒载体表现出在体外AAV生产细胞的转导增加。

[0025] 靶向肽可以与重组蛋白(例如,用于酶替代疗法)或多肽(例如,免疫球蛋白)连接以靶向所需组织(例如,CNS或肺),从而形成融合蛋白或缀合物。此外,靶向肽可以与脂质体和/或纳米颗粒(脂质纳米颗粒,LNP)连接,从而形成肽包被的脂质体和/或LNP以靶向所需组织。编码靶向肽的至少一个拷贝的序列和任意的连接序列可以与重组蛋白的编码序列框内融合,并与蛋白质或多肽共同表达以提供融合蛋白或缀合物。可替代地,可以用其它合成方法来形成与蛋白质、多肽或其它部分(例如,DNA、RNA或小分子)的缀合物。在某些实施例中,靶向肽的多个拷贝在融合蛋白/缀合物中。使靶向肽与重组蛋白缀合的合适方法包含用第一交联剂修饰重组人蛋白(例如,酶)上的氨基(N)末端和一个或多个残基,以产生第一交联剂修饰的重组人蛋白,用第二交联剂修饰在靶向肽前面的短延伸接头区域的氨基(N)末端以产生第二交联剂修饰的变体靶肽,然后使第一交联剂修饰的重组人蛋白与第二交联剂修饰的含有短延伸接头的变体靶向肽缀合。使靶向肽与重组蛋白缀合的其它合适方法包含使第一交联剂修饰的重组人蛋白与一个或多个第二交联剂修饰的变体靶向肽缀合,其中第

一交联剂修饰的重组蛋白包括特征在于具有化学修饰的N末端和一个或多个修饰的赖氨酸残基的重组蛋白,并且一个或多个第二交联剂修饰的变体靶向肽包括一个或多个变体靶向肽,所述变体靶向肽包括在靶向肽前面的短延伸接头的修饰的N末端氨基酸。可以选择用于使靶向肽与蛋白质、多肽、纳米颗粒或其它生物学上有用的化学部分缀合的其它合适方法。参见例如美国专利号US 9,545,450 B2(NHS-膦交联剂;NHS-叠氮化物交联剂);美国公开专利申请号US 2018/0185503 A1(醛-酰肼交联)。

[0026] 在某些实施例中,靶向肽可以插入到蛋白质或多肽(例如,病毒衣壳蛋白)内的合适位点。在某些这些实施例中并且在某些其它实施例中,靶向肽可以在其羧基(COO-)和/或氨基(N)末端侧接短延伸接头。此类接头的长度可以是1至20个氨基酸残基,或者长度可以是约2至20个氨基酸残基,或约1至15个氨基酸残基,或约2至12个氨基酸残基,或2至7个氨基酸残基。短延伸接头的长度也可以是约10个氨基酸。N末端的接头的存在和长度独立于羧基末端的接头来选择,并且羧基末端的接头的存在和长度独立于N末端的接头来选择。可以使用5个氨基酸的柔性GS延伸接头(甘氨酸-甘氨酸-甘氨酸-甘氨酸-丝氨酸)、包括2个柔性GS接头的10氨基酸延伸接头、包括3个柔性GS接头的15氨基酸延伸接头、包括4个柔性GS接头的20氨基酸延伸接头或其任何组合来提供合适的短延伸接头。

[0027] 在某些实施例中,提供了一种组合物,其可用于靶向内皮细胞。组合物是突变衣壳、融合蛋白或另一缀合物,其包括至少一个外源靶向肽,所述外源靶向肽包括:N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)的氨基酸序列,任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸并且任选进一步与纳米颗粒、第二分子或病毒衣壳蛋白缀合。靶向肽包括以下序列与任选连接序列:

[0028] (a) SSNTVKLTSGH(SEQ ID NO:40);

[0029] (b) EFSSNTVKLTS(SEQ ID NO:38);

[0030] (c) GGVLTNARGEYMRGG(SEQ ID NO:46);

[0031] (d) GGIEINATRAGTNLGG(SEQ ID NO:43);

[0032] (e) GGSSNTVKLTSGHGG(SEQ ID NO:39);

[0033] (f) IEINATRAGTNL(SEQ ID NO:42);或

[0034] (g) SANFIKPTSY(SEQ ID NO:41)。

[0035] 在某些实施例中,靶向肽基序由选自以下的核酸序列编码:

[0036] (a) agcagcaacaccgtgaagctgaccagcggacac(SEQ ID NO:54);

[0037] (b) gagttcagcagcaacaccgtgaagctgaccagc(SEQ ID NO:50);

[0038] (c) ggaggagtgcctgaccaacatcgctagaggagagtacatgagaggagga(SEQ ID NO:56);

[0039] (d) ggaggaatcgagatcaacgctaccagagctggaaccaacctgggagga(SEQ ID NO:52);

[0040] (e) ggaggaagcagcaacaccgtgaagctgaccagcggacacggagga(SEQ ID NO:55);

[0041] (f) atcgagatcaacgctaccagagctggaaccaacctg(SEQ ID NO:51);或

[0042] (g) agcgctaacttcatcaagcctaccagctac(SEQ ID NO:53)。

[0043] 在某些实施例中,靶向肽由SEQ ID NO:50的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽由SEQ ID NO:51的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽由SEQ ID NO:52的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽由SEQ ID NO:53的核酸序列或与其至少约70%相同的序

列编码。在某些实施例中，靶向肽由SEQ ID NO:54的核酸序列或与其至少约70%相同的序列编码。在某些实施例中，靶向肽由SEQ ID NO:55的核酸序列或与其至少约70%相同的序列编码。在某些实施例中，靶向肽由SEQ ID NO:56的核酸序列或与其至少约70%相同的序列编码。在一些实施例中，编码靶向肽基序的核酸序列任选在所述基序的核酸序列的5'和/或3'端侧接六个至二十一个核苷酸的延伸接头。

[0044] 在某些实施例中，靶向肽是NTVK。在某些实施例中，靶向肽是NTVR。在某些实施例中，在缀合物或修饰蛋白质(例如，细小病毒衣壳)中提供了这种基序内的靶向肽的超过一个拷贝。在某些实施例中，存在两个或更多个不同的靶向肽。

[0045] 在某些实施例中，提供了一种组合物，其可用于靶向鼻上皮和/或肺上皮细胞。组合物是突变衣壳、融合蛋白或另一缀合物，其包括至少一个外源靶向肽，所述外源靶向肽包括：N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)的氨基酸序列，任选在所述基序的氨基末端和/或羧基末端侧接两个氨基酸至七个氨基酸并且任选进一步与纳米颗粒、第二分子或病毒衣壳蛋白缀合。靶向肽包括：(a) SSNTVKLTSGH(SEQ ID NO:40)；(b) EFSSNTVKLTS(SEQ ID NO:38)；(c) GGVL TN IARGEYMRGG(SEQ ID NO:46)；(d) GGIEINATRAGTNLGG(SEQ ID NO:43)；(e) GGSSNTVKLTSGHGG(SEQ ID NO:39)；(f) IEINATRAGTNL(SEQ ID NO:42)；或(g) SANFIKPTSY(SEQ ID NO:41)。

[0046] 在某些实施例中，靶向肽是NTVK。在某些实施例中，靶向肽是NTVR，其任选侧接本文所述的间隔子氨基酸。在某些实施例中，在缀合物或修饰蛋白质(例如，细小病毒衣壳)中提供了这种基序内的靶向肽的超过一个拷贝。在某些实施例中，存在两个或更多个不同的靶向肽。

[0047] 合适的用于靶向的蛋白质的实例，包含酶、免疫球蛋白、治疗性蛋白、免疫原性多肽、纳米颗粒、DNA、RNA和其它部分(例如，小分子等)将在下面更详细描述。这些和其它生物和化学部分适合与本文提供的靶向肽一起使用。

[0048] 在某些实施例中，组合物是含有与核酸分子连接的靶向肽序列基序的核酸序列分子，其中核酸序列是DNA分子或RNA分子，例如，裸DNA、裸质粒DNA、信使RNA(mRNA)。在一些实施例中，核酸分子进一步与各种组合物和纳米颗粒偶联，包含例如胶束、脂质体、阳离子脂质-核酸组合物、聚多糖组合物和其它聚合物、脂质和/或胆固醇基-核酸缀合物，以及其它诸如本文所述的构建体。参见例如，W02014/089486、US 2018/0353616A1、US2013/0037977A1、W02015/074085A1、US9670152B2和US 8,853,377B2，X.Su等人，《分子药剂学(Mol.Pharmaceutics)》，2011,8(3)，第774-787页；网络发布：2011年3月21日；W02013/182683、W0 2010/053572和W0 2012/170930，它们均通过引用并入本文。在某些实施例中，靶向肽基序与纳米颗粒表面化学连接，其中纳米颗粒包封核酸分子。在一些实施例中，包括与表面连接的靶向肽的纳米颗粒被设计用于靶向组织特异性递送。在一些实施例中，两个或更多个不同的靶向肽与纳米颗粒表面连接。合适的化学连接或交联包含本领域技术人员已知的那些。

[0049] 衣壳

[0050] 在某些实施例中，提供了一种重组细小病毒，其具有修饰的细小病毒衣壳，所述衣壳具有至少来自N-x-(T/I/V/A)-(K/R)靶向基序的外源肽。此类重组细小病毒可以是杂合博卡病毒/AAV或重组AAV载体。在其它实施例中，可以生成其它在暴露的衣壳蛋白中具有一

个或多个来自N-x-(T/I/V/A)-(K/R)基序的外源靶向肽(所述外源靶向肽可以相同或不同,或其组合)的病毒载体,以调节和/或改变病毒载体与亲本载体相比的靶向特异性。

[0051] 靶向肽可以插入到高变环(HVR)VIII的任何合适位置。例如,基于AAV9衣壳的编号,在AAV9衣壳蛋白的氨基酸588和589(Q-A)之间给肽插入不同长度的接头,基于以下AAV9 VP1氨基酸序列的编号:SEQ ID NO:44。还参见2019年9月6日公开的W0 2019/168961,包含提供AAV9脱酰胺模式的表G,以及2018年9月7日提交的W0 2020/160582。AAVhu68(SEQ ID NO:45)中的氨基酸残基位置是相同的。然而,可以在HVRVIII内选择另一位点。可替代地,可以选择另一暴露的环HVR(例如,HVRIV)用于插入。可以在其它衣壳中选择可比的HVR区域。在某些实施例中,HVRVIII和HVRIV的位置使用美国专利号US 9,737,618 B2(第15栏第3-23行)和美国专利号US 10,308,958B2(第15栏第46行-第16栏第6行)中描述的算法和/或比对技术来确定,它们均通过引用以其整体并入本文。在某些实施例中,AAV1衣壳蛋白被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸582至585的HVRVIII区域或氨基酸456至459的HVRIV区域的合适位置,基于vp1编号(Gurda,BL.等人,不同的腺相关病毒血清型的衣壳抗体结合共同的区域(Capsid Antibodies to Different Adeno-Associated Virus Serotypes Bind Common Regions),2012,《病毒学杂志(Journal of Virology)》,2013年6月12日,87(16):9111-91114)。在某些实施例中,AAV8被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸586至591(例如,590-591(N-T))的HVRVIII区域或氨基酸456至460的HVRIV区域的合适位置,基于VP1编号(Gurda,BL.等人,绘制在腺相关病毒血清型8的衣壳上的中和表位(Mapping a Neutralizing epitope onto the Capsid of Adeno-Associated Virus Serotype 8),2012,《病毒学杂志》,2012年5月16日,86(15):7739-7751)。在某些实施例中,AAV7被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸589至590(N-T)的合适位置。在某些实施例中,AAV6被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸588至589(S-T)的合适位置。在某些实施例中,AAV5被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸577至578(T-T)的合适位置。在某些实施例中,AAV4被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸586至587(S-N)的合适位置。在某些实施例中,AAV3B被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸588至589(N-T)的合适位置。在某些实施例中,AAV2被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸587至588(N-R)的合适位置。在某些实施例中,AAV1被选作亲本衣壳,其中带有不同长度的接头的靶向肽被插入到氨基酸589至589(S-T)的合适位置。还参见图4。

[0052] 在某些实施例中,被修饰以含有N-x-(T/I/V/A)-(K/R)基序以及任选侧翼序列的亲本衣壳选自天然靶向CNS的细小病毒(例如,进化枝F AAV(例如,AAVhu68或AAV9)、进化枝E(例如,AAV8)或某些进化枝A AAV(例如,AAV1、AAVrh91))衣壳,或非细小病毒衣壳(例如,单纯疱疹病毒等),以增强CNS靶向细胞的表达和/或以其它方式调节CNS靶向细胞的类型。在其它实施例中,衣壳选自非天然靶向CNS的细小病毒(例如,进化枝F AAV,例如AAVhu68或AAV9,或某些进化枝A AAV,例如AAV1、AAVrh91)衣壳或非细小病毒衣壳(例如,单纯疱疹病毒(HSV)等)。参见例如2020年11月5日公开的W0 2020/223231(rh91,包含具有脱酰胺模式的表格),2020年8月14日提交的美国临时专利申请号63/065,616,和2020年11月4日提交的美国临时专利申请号63/109734。在某些实施例中,衣壳选自AAV进化枝F AAVhu95和

AAVhu96衣壳。参见例如2021年10月2日提交的美国临时申请号63/251,599。

[0053] 在某些实施例中,被修饰以含有N-x-(T/I/V/A)-(K/R)基序的亲本衣壳选自天然靶向鼻上皮细胞、鼻咽细胞和/或肺细胞的病毒(例如AAV),以便与亲本AAV(例如,进化枝AAV,例如AAV1、AAVrh32.33、AAV6.2、AAV6、AAVrh91)或AAV5或某些进化枝F AAV(例如,AAVhu68或AAV9)衣壳或非细小病毒衣壳(例如,腺病毒、HSV、RSV等)相比增强靶向性。参见例如2020年11月5日公开的WO 2020/223231(rh91,包含具有脱酰胺模式的表格),2020年8月14日提交的美国临时专利申请号63/065,616,和2020年11月4日提交的美国临时专利申请号63/109734。

[0054] 在某些实施例中,AAV衣壳不是包括NDVRAVS(SEQ ID NO:48)序列的突变AAV2衣壳。

[0055] 例如,可以选择来自进化枝F AAV如AAVhu68或AAV9的衣壳。已经描述了生成具有AAV9衣壳或AAVhu68衣壳和/或源自AAV9的嵌合衣壳的载体的方法。参见例如US 7,906,111,其通过引用并入本文。可以选择其它转导鼻细胞或另一合适靶标(例如,肌肉或肺)的AAV血清型作为AAV病毒载体的衣壳来源,包含例如AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV6.2、AAV7、AAV8、AAV9、rh10、AAVrh64R1、AAVrh64R2、rh8、AAVrh32.33(参见例如美国公开专利申请号2007-0036760-A1;美国公开专利申请号2009-0197338-A1;和EP 1310571)。还参见WO 2003/042397(AAV7和其它类人猿AAV)、美国专利7790449和美国专利7282199(AAV8)、WO 2005/033321(AAV9)和WO 2006/110689或尚未发现的或基于此的重组AAV可以用作AAV衣壳的来源。参见例如WO 2020/223232 A1(AAV rh90)、WO 2020/223231 A1和2021年8月13日提交的国际申请号PCT/US21/45945(AAV rh91)以及WO 2020/223236 A1(AAV rh92、AAV rh93、AAV rh9193),它们均通过引用以其整体并入本文。这些文档还描述了可以选择用于产生AAV的其它AAV,并且通过引用并入。在一些实施例中,用于病毒载体的AAV衣壳(cap)可以通过诱变(即通过插入、缺失或取代)上述AAV cap之一或其编码核酸来生成。在一些实施例中,所述AAV衣壳是嵌合的,包括来自上述AAV衣壳蛋白中的两种或三种或四种或更多种AAV衣壳蛋白的结构域。在一些实施例中,所述AAV衣壳是来自两种或三种不同的AAV或重组AAV的Vp1单体、Vp2单体和Vp3单体的嵌合体。在一些实施例中,rAAV组合物包括多于一种上述cap。

[0056] 如本文所用,与AAV的组有关的术语“进化枝”是指如基于对AAV vp1氨基酸序列进行的比对,通过(至少1000个复制品的)至少75%的自举值(bootstrap value)和不大于0.05的泊松校正距离测量结果(Poisson correction distance measurement)使用邻接算法(Neighbor-Joining algorithm)确定的一组在系统发育上彼此相关的AAV。在文献中已经描述了邻接算法。参见例如M.Nei和S.Kumar,《分子进化和系统发育学(Molecular Evolution and Phylogenetics)》(牛津大学出版社(Oxford University Press),纽约(2000))。提供了可以用于实施此算法的可用计算机程序。例如,MEGA v2.1程序实施了经修改的Nei-Gojobori方法。使用这些技术和计算机程序以及AAV vp1衣壳蛋白的序列,本领域技术人员可以容易地确定所选AAV是含在本文所鉴定的进化枝之一中,还是在这些进化枝之外的另一个进化枝中。参见例如,G Gao等人,《病毒学杂志(JVirology)》,2004年6月;78(12):6381-6388,其鉴定了进化枝A、B、C、D、E和F,并提供了新型AAV的核酸序列,GenBank登录号AY530553至AY530629。还参见WO 2005/033321。

[0057] 如本文所用,“AAV9衣壳”是由多种AAV9 vp蛋白构成的自组装AAV衣壳。AAV9vp蛋白通常被表达为替代性剪接变体,所述剪接变体由编码GenBank登录号:AAS99264的vp1氨基酸序列的核酸序列编码。这些剪接变体产生不同长度的蛋白质。在某些实施例中,“AAV9衣壳”包含具有与AAS99264 99%相同或与其99%相同的氨基酸序列AAV。还参见2019年9月6日公开的WO 2019/168961,包含提供AAV9的脱酰胺模式的表G。还参见US7906111和WO 2005/033321。如本文所用,“AAV9变体”包含在例如WO2016/049230、US 8,927,514、US 2015/0344911和US 8,734,809中描述的变体。

[0058] rAAVhu68由AAVhu68衣壳和载体基因组构成。AAVhu68衣壳是vp1异源群、vp2异源群和vp3蛋白异源群的组装体。如本文所用,当用于指vp衣壳蛋白时,术语“异源”或其任何语法变体是指由不相同的元件组成的群体,例如具有带有不同的修饰氨基酸序列的vp1、vp2或vp3单体(蛋白质)。还参见PCT/US2018/019992,WO 2018/160582,标题为“腺相关病毒(AAV)进化枝F载体及其用途(Adeno-Associated Virus (AAV) Clade F Vector and Uses Therefor)”,并且其通过引用以其整体并入本文。

[0059] 对于其它重组病毒载体,选择负责靶向特异性的病毒衣壳或包膜蛋白的适当暴露部分以用于靶向肽的插入。例如,在腺病毒中,可能需要修饰六邻体蛋白。在慢病毒中,包膜融合蛋白可以修饰为包括靶向基序的一个或多个拷贝。对于疫苗病毒,主要的糖蛋白可以修饰为包括靶向基序的一个或多个拷贝。适当地,为了安全起见,这些重组病毒载体是复制缺陷型的。

[0060] 表达盒和载体

[0061] 包装入AAV衣壳并递送至宿主细胞的载体基因组序列通常至少由转基因及其调控序列和AAV反向末端重复(ITR)构成。单链AAV和自互补(sc)AAV两者都涵盖在rAAV内。转基因是与载体序列异源的核酸编码序列,所述核酸编码序列编码多肽、蛋白质、功能性RNA分子(例如,miRNA、miRNA抑制剂)或其它所关注的基因产物。核酸编码序列以允许转基因在靶组织的细胞中转录、翻译和/或表达的方式与调控组分可操作地连接。

[0062] 载体的AAV序列通常包括顺式作用的5'和3'反向末端重复(ITR)序列(参见例如,B.J.Carter,“细小病毒手册(Handbook of Parvoviruses)”,P.Tijsser编辑,CRC Press,pp.155-168(1990))。ITR序列的长度为约145个碱基对(bp)。优选地,分子中使用了编码ITR的基本上整个序列,尽管允许对这些序列进行一定程度的微小修饰。修饰这些ITR序列的能力在本领域的技术范围内。(参见例如,文本,如Sambrook等人,“《分子克隆:实验室手册(Molecular Cloning:A Laboratory Manual)》”,第2版,纽约的冷泉港实验室(Cold Spring Harbor Laboratory,New York)(1989);以及K.Fisher等人,《病毒学杂志》,70:520-532(1996))。在本发明中采用的此类分子的实例是含有转基因的“顺式作用”质粒,其中所选转基因序列和相关的调控元件侧接5'和3'AAV ITR序列。在一个实施例中,ITR来自与供应衣壳的AAV不同的AAV。在一个实施例中,ITR序列来自AAV2。已经描述了被称为 Δ ITR的5'ITR的缩短版本,其中缺失了D序列和末端解析位点(trs)。在某些实施例中,载体基因组(例如,质粒的载体基因组)包含130个碱基对的缩短的AAV2 ITR,其中外部A元件被删除。在使用内部A元件作为模板并包装入衣壳以形成病毒颗粒的载体DNA扩增期间,缩短的ITR回复到145个碱基对的野生型长度。在其它实施例中,使用了全长AAV 5'和3'ITR。然而,可以选择来自其它AAV来源的ITR。在ITR的来源来自AAV2并且AAV衣壳来自另一个AAV来源的情况

下,所得载体可以被称为假型的。然而,这些元件的其它构型可以是合适的。

[0063] 除重组AAV载体的上文鉴定的主要元件外,AAV载体还包含必需的常规控制元件,所述常规控制元件以允许其在用质粒载体转染或用由本发明产生的病毒感染的细胞中转录、翻译和/或表达的方式与转基因可操作地连接。如本文所用,“可操作地连接的”序列包含与所关注的基因邻接的表达控制序列和以反式或在远处起作用以控制所关注的基因的表达控制序列。

[0064] 调控控制元件通常含有作为表达控制序列的一部分的启动子序列,例如定位在所选5' ITR序列与编码序列之间。可以在本文所描述的载体中使用组成型启动子、可调控启动子[参见例如WO 2011/126808和WO 2013/04943]、组织特异性启动子或响应于生理学诱因的启动子。

[0065] 适用于控制治疗产物的表达的组成型启动子的实例包含但不限于鸡 β -肌动蛋白(CB)启动子、CB7启动子、人巨细胞病毒(CMV)启动子、泛素C启动子(UbC)、猿猴病毒40(SV40)的早期和晚期启动子、U6启动子、金属硫蛋白启动子、EF1 α 启动子、泛素启动子、次黄嘌呤磷酸核糖基转移酶(HPRT)启动子、二氢叶酸还原酶(DHFR)启动子(Scharfmann等人,《美国国家科学研究院学报(Proc.Natl.Acad.Sci.USA 88:4626-4630(1991)、腺苷脱氨酶启动子、磷酸甘油激酶(PGK)启动子、丙酮酸激酶启动子、磷酸甘油突变酶启动子、 β -肌动蛋白启动子(Lai等人,《美国国家科学研究院学报(Proc.Natl.Acad.Sci.USA)》86:10006-10010(1989))、莫洛尼白血病病毒和其它逆转录病毒的长末端重复(LTR)、单纯疱疹病毒的胸苷激酶启动子和本领域技术人员已知的其它组成型启动子。适用于本发明的组织或细胞特异性启动子的实例包含但不限于内皮素-I(ET-I)和Flt-I(其是内皮细胞特异性的)、FoxJ1(其靶向纤毛细胞)。适合用于本发明的组织特异性启动子的其它实例包含但不限于肝脏特异性启动子。肝脏特异性启动子的实例可以包含例如甲状腺激素结合球蛋白(TBG),白蛋白,Miyatake等人,(1997)《病毒学杂志》,71:5124-32;乙肝病毒核心启动子,Sandig等人,(1996)《基因疗法(Gene Ther.)》,3:1002-9;或人类 α 1-抗胰蛋白酶、磷酸烯醇丙酮酸羧基激酶(PECK)或甲胎蛋白(AFP),Arbuthnot等人,(1996)《人类基因疗法(Hum.Gene Ther.)》,7:1503-14)。优选地,这种启动子是人源的。

[0066] 适用于控制治疗产物的表达的诱导型启动子包含响应于外源剂(例如,药理学剂)或生理学诱因的启动子。这些响应元件包含但不限于结合HIF-1 α 和 β 的低氧响应元件(HRE),金属离子响应元件,如Mayo等人(1982,《细胞(Cell)》29:99-108);Brinster等人(1982,《自然(Nature)》296:39-42)和Searle等人(1985)《分子细胞生物学(Mol.Cell.Biol.5:1480-1489)所述;或热休克响应元件,如Nouer等人所述(于:《热休克响应(Heat Shock Response)》,编辑Nouer,L.,CRC,Boca Raton,Fla.,pp167-220,1991)。

[0067] 在一个实施例中,基因产物的表达由可调控启动子控制,所述启动子严格控制编码基因产物(例如药理学剂)的序列的转录或被药理学剂或在替代实施例中被生理学诱因激活的转录因子。优选非渗漏且可被严格控制的启动子系统。

[0068] 可用于本发明的作为配体依赖性转录因子复合物的可调控启动子的实例包含但不限于被其各自配体(例如,糖皮质激素、雌激素、孕激素、类维生素A、蜕皮激素及其类似物和模拟物)激活的核受体超家族成员和被四环素激活的rTTA。在本发明的一个方面,基因开关是基于EcR的基因开关。此类系统的实例包含但不限于在美国专利号6,258,603、7,045,

315,美国公开专利申请号2006/0014711、2007/0161086和国际公开申请号WO 01/70816中描述的系统。嵌合蜕皮激素受体系统的实例描述于以下中:美国专利号7,091,038、美国公开专利申请号2002/0110861、2004/0033600、2004/0096942、2005/0266457和2006/0100416,以及国际公开申请号WO 01/70816、WO 02/066612、WO 02/066613、WO 02/066614、WO 02/066615、WO 02/29075和WO 2005/108617,其各自通过引用以其整体并入。非甾体蜕皮激素激动剂调控系统的实例是RheoSwitch®哺乳动物诱导型表达系统(马塞诸塞州伊普斯威奇新英格兰生物实验室(New England Biolabs,Ipswich,MA))。

[0069] 又其它启动子系统可以包含响应元件,包含但不限于四环素(tet)响应元件(例如,由Gossen&Bujard(1992,《美国国家科学院学报(Proc.Natl.Acad.Sci.USA)》89:5547-551描述);或激素响应元件,例如由Lee等人(1981,《自然》294:228-232);Hynes等人(1981,《美国国家科学院学报(Proc.Natl.Acad.Sci.USA)》78:2038-2042);Klock等人(1987,《自然》329:734-736);以及Israel&Kaufman(1989,《核酸研究(Nucl.Acids Res.)》17:2589-2604)描述,以及本领域已知的其它诱导型启动子。使用此类启动子可以控制可溶性hACE2构建体的表达,例如,通过Tet-开/关系统(Gossen等人,1995,《科学(Science)》268:1766-9;Gossen等人,1992,《美国国家科学院学报(Proc.Natl.Acad.Sci.USA.)》,89(12):5547-51);TetR-KRAB系统(Urrutia R.,2003,《基因组生物学(Genome Biol.)》,4(10):231;Deuschle U等人,1995,《分子细胞生物学》(4):1907-14);美服培酮(RU486)可调控系统(Geneswitch;Wang Y等人,1994,《美国国家科学院学报(Proc.Natl.Acad.Sci.USA.,91(17):8180-4;Schillinger等人,2005,《美国国家科学院学报(Proc.Natl.Acad.Sci.U S A.102(39):13789-94);以及人源化他莫昔芬-dep可调控系统(Roscilli等人,2002,《分子疗法(Mol.Ther.)》6(5):653-63)。

[0070] 在另一方面,基因开关基于FK506结合蛋白(FKBP)与FKBP雷帕霉素相关蛋白(FRAP)的异二聚化,并通过雷帕霉素或其非免疫抑制类似物来调控。此类系统的实例包含但不限于ARGENT™转录技术(ARIAD Pharmaceuticals,Cambridge,Mass.)以及描述于以下中的系统:美国专利号6,015,709、6,117,680、6,479,653、6,187,757和6,649,595,美国公开号2002/0173474、美国公开号200910100535、美国专利号5,834,266、美国专利号7,109,317、美国专利号7,485,441、美国专利号5,830,462、美国专利号5,869,337、美国专利号5,871,753、美国专利号6,011,018、美国专利号6,043,082、美国专利号6,046,047、美国专利号6,063,625、美国专利号6,140,120、美国专利号6,165,787、美国专利号6,972,193、美国专利号6,326,166、美国专利号7,008,780、美国专利号6,133,456、美国专利号6,150,527、美国专利号6,506,379、美国专利号6,258,823、美国专利号6,693,189、美国专利号6,127,521、美国专利号6,150,137、美国专利号6,464,974、美国专利号6,509,152、美国专利号6,015,709、美国专利号6,117,680、美国专利号6,479,653、美国专利号6,187,757、美国专利号6,649,595、美国专利号6,984,635、美国专利号7,067,526、美国专利号7,196,192、美国专利号6,476,200、美国专利号6,492,106、WO 94/18347、WO 96/20951、WO 96/06097、WO 97/31898、WO 96/41865、WO 98/02441、WO 95/33052、WO 99110508、WO 99110510、WO 99/36553、WO 99/41258、WO 01114387,ARGENT™调控转录逆转录病毒试剂盒,2.0版(9109102),和ARGENT™调控转录质粒试剂盒,2.0版(9109/02),其各自通过引用以其整体并入本文。Ariad系统被设计为由雷帕霉素及其类似物(称为“rapalogs”)诱导。合适的雷帕霉素的实

例在以上结合ARGENT™系统的描述列出的文件中提供。在一个实施例中，分子是雷帕霉素[例如，由Pfizer以Rapamune™销售]。在另一个实施例中，使用被称为AP21967[ARIAD]的rapalog。可用于本发明的这些二聚体分子的实例包含但不限于雷帕霉素、FK506、FK1012(FK506的同二聚体)、雷帕霉素类似物(“rapalogs”)，其容易地通过对天然产物进行化学修饰以添加“隆起物(bump)”来制备，所述隆起物减少或消除对内源FKBP和/或FRAP的亲合力。rapalog的实例包括但不限于，如AP26113(Ariad)、AP1510(Amara, J.F. 等人, 1997, 《美国国家科学研究所学报(Proc. Natl. Acad. Sci. USA)》, 94(20):10618-23)、AP22660、AP22594、AP21370、AP22594、AP23054、AP1855、AP1856、AP1701、AP1861、AP1692和AP1889，其具有设计的使与内源FKBP的相互作用最小化的‘隆起物’。还可以选择又其它rapalog，例如AP23573[Merck]。在某些实施例中，雷帕霉素或合适的类似物可被局部递送至鼻咽的AAV转染的细胞。这种局部递送可以通过鼻内注射经由团注、乳霜或凝胶局部递送到细胞。参见美国专利申请US 2019/0216841 A1，其通过引用并入本文。

[0071] 其它合适的增强子包含那些适合于所需靶组织适应症的增强子。在一个实施例中，表达盒包括一种或多种表达增强子。在一个实施例中，表达盒含有两种或更多种表达增强子。这些增强子可以相同或彼此不同。例如，增强子可以包含CMV立即早期增强子。这种增强子可以存在于彼此相邻定位的两个拷贝中。可替代地，增强子的双拷贝可以被一个或多个序列分离。在又一个实施例中，表达盒进一步含有内含子，例如，鸡β-肌动蛋白内含子。其它合适的内含子包含本领域已知的内含子，例如，WO 2011/126808中所描述的内含子。合适的多聚腺苷酸化(polyA)序列的实例包含例如兔结合球蛋白(也称为兔β球蛋白或rBG)、SV40、SV50、牛生长激素(bGH)、人生长激素和合成polyA。任选地，可以选择一个或多个序列来稳定mRNA。此类序列的实例是修饰的WPRE序列，其可以在polyA序列的上游和编码序列的下游被工程化(参见例如MA Zanta-Boussif等人, 《基因疗法》(2009) 16:605-619)。

[0072] AAV病毒载体可以包含多个转基因。在某些情况下，可以使用不同的转基因编码蛋白质的每个亚基(例如，免疫球蛋白结构域、免疫球蛋白重链、免疫球蛋白轻链)。在一个实施例中，在用含不同亚基中的每一种的病毒感染/转染后，细胞产生多亚基蛋白质。在另一个实施例中，蛋白质的不同亚基可以由相同的转基因编码。当编码每个亚基的DNA的大小较小，例如，编码亚基和IRES的DNA的总大小小于5千碱基时，IRES是需要的。作为IRES的替代方案，DNA可以由编码2A肽的序列分离，所述2A肽在翻译后事件中自切割。参见例如ML Donnelly等人(1997年1月)《基因病毒学杂志(J. Gen. Virol.)》, 78(Pt 1):13-21; S. Furler, S等人(2001年6月)《基因疗法》, 8(11):864-873; H. Klump等人(2001年5月)《基因疗法》, 8(10):811-817。这种2A肽显著小于IRES，使其非常适合在空间是限制因素的情况下使用。更通常地，当转基因大、由多亚基组成，或共同递送两个转基因时，共同施用携带所需的转基因或亚基的rAAV，以允许它们在体内串联以形成单一载体基因组。在此类实施例中，第一AAV可以携带表达单个转基因的表达盒，并且第二AAV可以携带表达不同转基因的表达盒，用于在宿主细胞中共表达。然而，所选择的转基因可以编码任何生物活性产物或其它产物，例如研究所需的产物。

[0073] 除了上文针对表达盒所鉴定的元件外，载体还包含常规控制元件，其以允许在用质粒载体转染或者用本发明产生的病毒感染的细胞中转录、翻译和/或表达编码产物(例如，可溶性hACE2构建体、抗流感抗体、抗COVID19抗体)的方式与编码序列可操作地连接。本

文提供了其它合适转基因的实例。如本文所用,“可操作地连接的”序列包含与所关注的基因邻接的表达控制序列和以反式或在远处起作用以控制所关注的基因的表达控制序列。

[0074] 表达控制序列包含适当的增强子;转录因子;转录终止子;启动子;有效的RNA处理信号,如剪接和多聚腺苷酸化(polyA)信号;使细胞质mRNA稳定的序列,例如土拨鼠肝炎病毒(WHP)转录后调控元件(WPRE);增强翻译效率的序列(即Kozak共有序列);增强蛋白质稳定性的序列;以及当需要时,增强编码产物分泌的序列。

[0075] 在一个实施例中,选择调控序列,使得总rAAV载体基因组的大小为约2.0至约5.5千碱基。在一个实施例中,需要rAAV载体基因组接近天然AAV基因组的大小。因此,在一个实施例中,选择调控序列使得总rAAV载体基因组的大小为约4.7kb。在另一个实施例中,总rAAV载体基因组的大小小于约5.2kb。可以基于包含启动子、增强子、内含子、poly A等的调控序列的大小来操纵载体基因组的大小。参见Wu等人,《分子疗法》,2010年1月,18(1):80-6,其通过引用并入本文。

[0076] 因此,在一个实施例中,内含子包含在载体中。合适的内含子包含鸡 β -肌动蛋白内含子,人 β 球蛋白IVS2(Kelly等人,《核酸研究》,43(9):4721-32(2015));Promega嵌合内含子(Almond, B. 和Schenborn, E. T.《pCI-neo载体和pcDNA4/HisMax载体的比较(A Comparison of pCI-neo Vector and pcDNA4/HisMax Vector)》);和hFIX内含子。适于本文的多种内含子是本领域已知的并且包含但不限于在**bp.g.utoledo.edu/~afedorov/lab/eid.html**上发现的内含子,其通过引用并入本文。还参见Shepelev V., Fedorov A. 外显子-内含子数据库的进展(Advances in the Exon-Intron Database).《生物信息简报(Briefings in Bioinformatics)》2006,7:178-185,其通过引用并入本文。

[0077] 在本文描述的研究中生成了若干不同的病毒基因组。然而,本领域技术人员将理解,可以用其它基因组构型(包含其它调控序列)来代替启动子、增强子,并且可以选择其它编码序列。

[0078] rAAV载体产生

[0079] 供产生AAV病毒载体(例如,重组(r)AAV)之用,表达盒可以携带在递送到包装宿主细胞的任何合适的载体(例如,质粒)上。可在本发明中使用的质粒可以被工程化,使得其适合于在原核细胞、昆虫细胞、哺乳动物细胞以及其它细胞中进行体外复制和包装。合适的转染技术和包装宿主细胞是已知的和/或可以由本领域的技术人员容易地设计。

[0080] 在某些实施例中,与不在AAV衣壳中纳入基序的至少一个拷贝的方法相比,将N-x-(T/I/V/A)-(K/R)基序的至少一个拷贝纳入AAV衣壳中提供了生产优势,并且其中生产细胞是293细胞。

[0081] 制备基于AAV的载体(例如,具有AAV9或另一AAV衣壳)的方法是已知的。参见例如美国公开专利申请号2007/0036760(2007年2月15日),其通过引用并入本文。本发明不限于使用AAV9或其它进化枝F AAV氨基酸序列,而是涵盖通过本领域已知的其它方法(包含例如通过化学合成、通过其它合成技术或通过其它方法)生成的含有末端 β -半乳糖结合的肽和/或蛋白质。本文提供的任何AAV衣壳的序列可以使用多种技术容易地生成。合适的生产技术是本领域技术人员众所周知的。参见例如Sambrook等人,《分子克隆实验指南(Molecular Cloning: A Laboratory Manual)》,纽约冷泉港的冷泉港实验室出版社(Cold Spring Harbor Press, Cold Spring Harbor, NY)。可替代地,也可以通过众所周知的固相肽合成方

法(Merrifield, (1962)《美国化学会志(J. Am. Chem. Soc.)》, 85:2149; Stewart和Young,《固相肽合成(Solid Phase Peptide Synthesis)》(旧金山弗里曼, 1969) pp.27-62)来合成肽。这些方法可以涉及例如培养宿主细胞, 所述宿主细胞含有编码AAV衣壳的核酸序列; 功能性rep基因; 至少由AAV反向末端重复(ITR)和转基因组成的小基因(minigene); 以及足以允许将小基因包装到AAV衣壳蛋白中的辅助功能。这些和其它合适的生产方法在本领域技术人员知识范围内, 而不是对本发明的限制。

[0082] 在宿主细胞中培养以将AAV小基因包装至AAV衣壳中所需的组分可以反式形式提供至宿主细胞。或者, 可通过稳定宿主细胞提供所需组分(例如小基因、rep序列、cap序列和/或辅助功能)中的任一种或多种, 所述宿主细胞已使用本领域的技术人员已知的方法工程化为含有所需组分中的一种或多种。最合适的是, 此类稳定的宿主细胞将含有在诱导型启动子的控制下的所需组分。然而, 所需组分可在组成型启动子的控制下。在适用于转基因的调控元件的讨论中, 本文提供了合适的诱导型和组成型启动子的实例。在另一替代方案中, 所选稳定宿主细胞可含有在组成型启动子控制下的所选成分和在一或多个诱导型启动子控制下的其它所选成分。例如, 可产生稳定宿主细胞, 所述宿主细胞来源于293细胞(其含有在组成型启动子控制下的E1辅助功能), 但含有在诱导型启动子控制下的rep和/或cap蛋白。其它稳定的宿主细胞可由本领域的技术人员产生。

[0083] 这些rAAV特别适用于基因递送以用于治疗目的和用于防止感染。进一步地, 本发明的组合物还可以用于体外产生所需的基因产物。对于体外产生, 可以在用含有对所需的产物进行编码的分子的rAAV转染宿主细胞并在允许表达的条件下培养细胞培养物之后从所期望的培养物中获得所期望的产物(例如, 蛋白质)。然后可以根据需要纯化和分离所表达的产物。用于转染、细胞培养、纯化和分离的合适的技术是本领域的技术人员已知的。用于产生和分离适合于用作载体的AAV的方法是本领域已知的。通常参见例如, Grieger & Samulski, 2005, “腺相关病毒作为基因疗法载体: 载体研发、产生及临床应用(Adeno-associated virus as a gene therapy vector: Vector development, production and clinical applications)”, 《生物化学工程/生物技术进展(Adv. Biochem. Engin/Biotechnol.)》99:119-145; Buning等人, 2008, “腺相关病毒载体技术的最新发展(Recent developments in adeno-associated virus vector technology)”, 《基因医学杂志(J. Gene Med.)》10:717-733; 以及下文引用的参考文献, 这些参考文献中的每个参考文献通过引用整体并入本文。为了将转基因包装到病毒颗粒中, ITR是在与含有表达盒的核酸分子相同的构建体中需要的顺式的唯一AAV组分。cap和rep基因可以反式供应。

[0084] 在一个实施例中, 本文所描述的表达盒被工程化到基因元件(例如, 穿梭质粒), 所述基因元件将其上携带的免疫球蛋白构建体序列转移到包装宿主细胞中以产生病毒载体。在一个实施例中, 所选基因元件可以通过任何合适的方法递送至AAV包装细胞, 所述方法包含转染、电穿孔、脂质体递送、膜融合技术、高速DNA包被的团粒、病毒感染和原生质体融合。也可以制备稳定的AAV包装细胞。可替代地, 表达盒可以用于产生除AAV之外的病毒载体, 或用于在体外产生抗体的混合物。用于制备此类构建体的方法对核酸操纵技术人员而言是已知的并且包含基因工程、重组工程以及合成技术。参见例如, 《分子克隆: 实验室手册》, Green和Sambrook编辑, 纽约冷泉港的冷泉港实验室出版社(Cold Spring Harbor Press, Cold Spring Harbor, NY) (2012)。

[0085] 术语“AAV中间体”或“AAV载体中间体”是指缺少包装在其中的所期望的基因组序列的组装的rAAV衣壳。这些也可以被称为“空”衣壳。此类衣壳可以不含有表达盒的可检测基因组序列,或者含有不足以实现基因产物的表达的仅部分包装的基因组序列。这些空衣壳是无功能性的以将所关注的基因转移至宿主细胞。

[0086] 本文描述的重组AAV可以使用已知的技术生成。参见例如WO 2003/042397;WO 2005/033321、WO 2006/110689;US 7588772 B2。此类方法涉及培养宿主细胞,其含有编码AAV衣壳的核酸序列;功能性rep基因;表达盒,其至少由AAV反向末端重复(ITR)和转基因构成;以及足以允许将表达盒包装到AAV衣壳蛋白中的辅助功能。因此已经描述了产生衣壳的方法、编码序列以及用于产生rAAV病毒载体的方法。参见例如,Gao等人,《美国国家科学院院刊(Proc.Natl.Acad.Sci.)》U.S.A),100(10),6081-6086(2003)和US 2013/0045186A1。

[0087] 在一个实施例中,在合适的细胞培养物(例如,HEK 293细胞)中制造细胞。用于制造本文所描述的基因疗法载体的方法包含本领域众所周知的方法,如产生用于产生基因疗法载体的质粒DNA、产生载体以及纯化载体。在一些实施例中,基因疗法载体是AAV载体,并且生成的质粒是编码AAV基因组和所关注的基因以用于包装入衣壳的AAV顺式质粒、含有AAV rep和cap基因的AAV反式质粒,以及腺病毒辅助质粒。载体产生过程可以包含方法步骤,如开始细胞培养、进行细胞传代、接种细胞、用质粒DNA转染细胞、将转染后培养基交换为无血清培养基以及采集含载体的细胞和培养基。所采集的含载体的细胞和培养基在本文中被称为粗细胞采集物。在又另一种系统中,通过用基于杆状病毒的载体进行感染来将基因疗法载体引入到昆虫细胞中。关于这些生产系统的综述,一般参见例如Zhang等人,2009,“用于大规模重组腺相关病毒生产的腺病毒-腺相关病毒杂合体(Adenovirus-adenovirus-associated virus hybrid for large-scale recombinant adeno-associated virus production),”《人基因疗法(Human Gene Therapy)》20:922-929,其通过引用以其整体并入本文。在以下美国专利中也描述了制备和使用这些及其它AAV产生系统的方法,所述美国专利中的每个美国专利的内容通过引用整体并入本文:5,139,941;5,741,683;6,057,152;6,204,059;6,268,213;6,491,907;6,660,514;6,951,753;7,094,604;7,172,893;7,201,898;7,229,823;和7,439,065。在某些实施例中,制备和使用AAV生产系统的方法包含使用2020年4月28日提交的美国专利申请63/016,894中描述的伪狂犬病病毒(rPRV)的方法,其通过引用并入本文。

[0088] 此后,粗细胞采集物可以是本主题的方法步骤,如浓缩载体采集物、渗滤载体采集物、微流化载体采集物、核酸酶消化载体采集物、过滤经微流化的中间体、通过色谱粗纯化、通过超速离心法粗纯化、通过切向流过滤进行缓冲液交换和/或调配和过滤以制备大量载体。

[0089] 在高盐浓度下进行两步亲和色谱法纯化,随后使用阴离子交换树脂色谱法来纯化载体药物产物并去除空衣壳。这些方法更详细描述在以下中:2016年12月9日提交的国际专利申请号PCT/US2016/065970,标题为“AAV9的可扩展纯化方法(Scalable Purification Method for AAV9)”,其通过引用并入本文。针对于2016年12月9日提交的国际专利申请第PCT/US2016/065976号中的AAV8、以及于2016年12月9日提交、还于2015年12月11日提交的题为“AAVrh10的可扩展纯化方法(Scalable Purification Method for AAVrh10)”的国际专利申请第PCT/US16/66013号中的rh10、以及于2016年12月9日提交的国际专利申请第

PCT/US2016/065974号“AAV1的可扩展纯化方法(Scalable Purification Method for AAV1)”、于2015年12月11日提交中的AAV1的纯化方法全部通过引用并入本文。

[0090] 为了计算空颗粒和完整颗粒含量,将所选样品的vp3带体积(例如,在本文的实例中,碘克沙醇梯度纯化的制剂,其中GC数=颗粒数)相对于负载的GC颗粒作图。所得线性等式($y=mx+c$)用于计算测试制品峰的带状体积中的颗粒的数量。然后将加载的每20 μ L颗粒数量(pt)乘以50,以得到颗粒(pt)/mL。将Pt/mL除以GC/mL得到颗粒与基因组拷贝的比率(pt/GC)。Pt/mL-GC/mL得到空pt/mL。空pt/mL除以pt/mL并且 $\times 100$ 得到空颗粒的百分比。

[0091] 通常,用于测定具有包装的基因组的空衣壳和AAV载体颗粒的方法是本领域已知的。参见例如Grimm等人,《基因疗法》(1999)6:1322-1330;和Sommer等人,《分子疗法(Molec. Ther.)》(2003)7:122-128。为了测试变性的衣壳,所述方法包含使经处理的AAV原液经受SDS-聚丙烯酰胺凝胶电泳(由能够分离三种衣壳蛋白的任何凝胶组成,例如在缓冲液中含有3-8%三乙酸盐的梯度凝胶),然后运行凝胶直到分离出样品材料,并且将凝胶印迹到尼龙或硝酸纤维素膜(优选地尼龙)上。然后,将抗AAV衣壳抗体用作与变性的衣壳蛋白结合的初级抗体,优选地抗AAV衣壳单克隆抗体,最优选地B1抗AAV2单克隆抗体(Wobus等人,《病毒学杂志》(2000)74:9281-9293)。然后使用次级抗体,所述次级抗体与初级抗体结合并且含有一种用于检测与初级抗体的结合的装置,更优选地是含有与其共价结合的检测分子的抗IgG抗体,最优选地是与辣根过氧化物酶共价连接的绵羊抗小鼠IgG抗体。一种用于检测结合的方法用于半定量地确定初级抗体与次级抗体之间的结合,优选地是能够检测放射性同位素发射、电磁辐射或比色变化的检测方法,最优选地是化学发光检测试剂盒。例如,对于SDS-PAGE,可以从柱级分中提取样品并在含有还原剂(例如,DTT)的SDS-PAGE上样缓冲液中加热,并且在预制的梯度聚丙烯酰胺凝胶(例如,Novex)上解析衣壳蛋白。可以根据制造商的说明使用SilverXpress(加利福尼亚州英杰公司(Invitrogen, CA))或其它合适的染色方法(即,SYPRO红宝石色或考马斯染色)进行银染色。在一个实施例中,可以通过定量实时PCR(Q-PCR)测量柱级分中的AAV载体基因组(vg)的浓度。将样品稀释并用DNA酶I(或另一种合适的核酸酶)消化以去除外源DNA。在核酸酶失活后,使用引物和对引物之间的DNA序列具有特异性的TaqManTM荧光探针进一步稀释和扩增样品。在Applied Biosystems Prism 7700序列检测系统上测量每种样品达到所定义的荧光水平所需的周期的数量(阈值周期,Ct)。含有与AAV载体中所含序列相同的序列的质粒DNA用于在Q-PCR反应中产生标准曲线。从样品获得的周期阈值(Ct)的值用于通过相对于质粒标准曲线的Ct值对其进行归一化来确定载体基因组效价。也可以使用基于数字PCR的端点测定。

[0092] 另外,本领域中还已知其它测量空颗粒与完整颗粒比率的实例。在分析超速离心机(AUC)中测量的沉降速度可以检测聚集体、其它次要组分以及基于其不同沉降系数提供不同颗粒物质的相对量的良好定量。这是基于长度和时间的基本单位的绝对方法,不需要标准分子作为参考。将载体样品加载到具有12mm光程长度的2通道炭-epon中心片(centerpiece)的细胞中。将提供的稀释缓冲液加载到每个细胞的参考通道中。然后将加载的细胞置于AN-60Ti分析转子中并装入配有吸光度和RI检测器的Beckman-Coulter ProteomeLab XL-I分析超速离心机中。在20 $^{\circ}$ C下完全温度平衡后,使转子达到12,000rpm的最终运行速度。大约每3分钟记录 A_{280} 扫描,持续约5.5小时(每个样品扫描总共110次)。使用c(s)方法分析原始数据,并在分析程序SEDFIT中实施。将得到的尺寸分布绘制成图并进行

峰积分。与每个峰相关联的百分比值表示在所有峰下的总面积的峰面积分数,并且基于在280nm处生成的原始数据;许多实验室使用这些值来计算空颗粒:完整颗粒比率。然而,因为空颗粒和完整颗粒在这一波长处具有不同的消光系数,所以可以相应地调整原始数据。消光系数调整之前和之后的空颗粒与完整单体峰值的比率用于确定空颗粒-完整颗粒比率。

[0093] 一方面,使用了经优化的q-PCR方法,所述方法利用了广谱丝氨酸蛋白酶,例如蛋白酶K(如可从凯杰公司(Qiagen)商购获得)。更具体地,经过优化的qPCR基因组效价测定与标准测定类似,不同之处在于在DNA酶I消化之后,将样品用蛋白酶K缓冲液稀释并用蛋白酶K处理,然后进行热失活。合适地,以等于样品大小的量用蛋白酶K缓冲液稀释样品。蛋白酶K缓冲液可以浓缩2倍或更多倍。通常,蛋白酶K处理为约0.2mg/mL,但是可以在0.1mg/mL至约1mg/mL之间变化。处理步骤通常在约55°C下进行持续约15分钟,但是可以在较低温度(例如,约37°C至约50°C)下进行持续较长的时间段(例如,约20分钟至约30分钟),或者在较高的温度(例如,至多约60°C)下进行持续较短的时间段(例如,约5分钟至10分钟)。类似地,热失活通常在约95°C下持续约15分钟,但是温度可以降低(例如,约70°C至约90°C)并且时间延长(例如,约20分钟至约30分钟)。然后将样品稀释(例如,1000倍),并如标准测定中所描述的进行TaqMan分析。也可以使用ViroCyt或流式细胞术进行定量。

[0094] 另外地或可替代地,可以使用液滴数字PCR(ddPCR)。例如,已经描述了用于通过ddPCR确定单链和自互补AAV载体基因组效价的方法。参见例如M.Lock等人,Hu基因治疗方法(Hu Gene Therapy Methods),《人类基因治疗方法(Hum.Gene Ther.Methods)》.2014年4月;25(2):115-25.doi:10.1089/hgtb.2013.131.电子版2014年2月14日。

[0095] 治疗性蛋白和递送系统

[0096] 含有本文提供的靶向基序(即N-x-(T/I/V/A)-(K/R)基序)的融合伴侣、缀合伴侣和重组载体可用于各种不同的治疗性蛋白、多肽、纳米颗粒和递送系统。可用于本文提供的组合物和靶向递送的蛋白质和化合物的实例包含以下。可以理解的是,病毒载体、纳米颗粒和其它递送系统含有编码用于在体内表达的所选蛋白质(或缀合物)的序列。

[0097] 在某些实施例中,蛋白质是MCT8蛋白(SLC16A2基因)和其它用于治疗Allan-Herndon-Dudley病及其症状的化合物。

[0098] 在某些实施例中,蛋白质选自与运输缺陷相关的疾病,如例如囊性纤维化(囊性纤维化跨膜调节因子)、 α -1-抗胰蛋白酶(遗传性肺气肿)、FE(遗传性血色病)、酪氨酸酶(眼皮肤白化病)、蛋白C(蛋白C缺乏症)、补体C抑制剂(I型遗传性血管性水肿)、 α -D-半乳糖苷酶(Fabry病)、 β 己糖胺酶(Tay-Sachs)、蔗糖酶-异麦芽糖酶(先天性蔗糖酶-异麦芽糖酶缺乏症)、UDP-葡萄糖醛酸转移酶(克果纳杰氏II型)、胰岛素受体(糖尿病)、生长激素受体(莱伦综合征)等。其它基因和蛋白质的实例是与例如以下有关的那些:脊髓性肌肉萎缩症(SMA, SMN1)、亨廷顿病、雷特综合征(例如,甲基-CpG-结合蛋白2(MeCP2);UniProtKB-P51608)、肌萎缩性脊髓侧索硬化症(ALS)、杜氏肌营养不良、弗里德赖希共济失调(例如,frataxin)、与脊髓小脑共济失调2型(SCA2)/ALS相关的ATXN2;与ALS相关的TDP-43、颗粒蛋白前体(PRGN)(与非阿尔茨海默氏型脑退化(包含额颞叶痴呆(FTD)、进行性非流利性失语(PNFA)和语义痴呆)相关)等。参见例如,www.orpha.net/consor/cgi-bin/Disease_Search_List.php;rarediseases.info.nih.gov/diseases。可以经由rAAV递送的进一步的示例性基因包含但不限于与糖原贮积病或1A型缺陷(GSD1)相关的葡萄糖6磷酸酶、与PEPCK缺陷相关的磷酸烯

醇丙酮酸羧激酶 (PEPCK) ;与发作和严重的神经发育障碍相关的细胞周期蛋白依赖性激酶样5 (CDKL5) ,也称为丝氨酸/苏氨酸激酶9 (STK9) ;与半乳糖血症相关的半乳糖-1磷酸尿苷转移酶;与苯丙酮尿症 (PKU) 相关的苯丙氨酸羟化酶 (PAH) ;与原发性高草酸尿症1型相关的基因产物,包含羟基酸氧化酶1 (GO/HAO1) 和AGXT,与枫糖浆尿病相关的支链 α -酮酸脱氢酶,包含BCKDH、BCKDH E2、BAKDH E1a和BAKDH E1b;与1型酪氨酸血症相关的延胡索酰乙酰乙酸酶;与甲基丙二酸血症相关的甲基丙二酰辅酶A变位酶;与中链乙酰辅酶A缺乏相关的中链酰基辅酶A脱氢酶;与鸟氨酸转氨甲酰酶缺乏症相关的鸟氨酸转氨甲酰酶 (OTC) ;与瓜氨酸血症相关的精氨酸琥珀酸合成酶 (ASS1) ;卵磷脂-胆固醇酰基转移酶 (LCAT) 缺乏症;甲基丙二酸血症 (MMA) ;与尼曼匹克病C1型相关的NPC1) ;丙酸血症 (PA) ;与转甲状腺素蛋白 (TTR) 相关的遗传性淀粉样变性相关的TTR;与家族性高胆固醇血症 (FH) 相关的低密度脂蛋白受体 (LDLR) 蛋白,LDLR变体,如WO 2015/164778中描述的那些;PCSK9;与痴呆症相关的ApoE和ApoC蛋白;与克果纳杰氏病相关的UDP-葡萄糖醛酸转移酶;与严重的联合免疫缺陷病相关的腺苷脱氢酶;与痛风和莱施-奈恩综合征相关的次黄嘌呤鸟嘌呤磷酸核糖转移酶;与生物素酶缺乏症相关的生物素酶;与法布里病相关的 α -半乳糖苷酶A (α -Gal A)) ;与GM1神经节苷脂沉积病相关的 β -半乳糖苷酶 (GLB1) ;与威尔逊病相关的ATP7B;与戈谢病2型和3型相关的 β -葡萄糖脑苷脂酶;与齐薇格综合征相关的过氧化物酶体膜蛋白70kDa;与异染性脑白质营养不良相关的芳基硫酸酯酶A (ARSA) 、与克拉伯病相关的半乳糖脑苷脂酶 (GALC) 、与庞贝病相关的 α -葡萄糖苷酶 (GAA) ;与尼曼匹克病A型相关的鞘磷脂酶 (SMPD1) 基因;与成人发病的II型瓜氨酸血症 (CTLN2) 相关的精氨酸琥珀酸合成酶;与尿素循环障碍相关的氨基甲酰磷酸合成酶1 (CPS1) ;与脊髓性肌萎缩症相关的运动神经元存活 (SMN) 蛋白;与法伯脂肪肉芽肿相关的神经酰胺酶;与GM2神经节苷脂沉积病和泰-萨二氏病和桑德霍夫病相关的b-氨基己糖苷酶;与天冬氨酰葡萄糖胺相关的天冬氨酰氨基葡萄糖苷酶;与岩藻糖苷贮积症相关的 α -岩藻糖苷酶;与 α -甘露糖苷贮积症相关的 α -甘露糖苷酶;与急性间歇性卟啉症 (AIP) 相关的胆色素原脱氢酶;用于治疗 α -1抗胰蛋白酶缺乏症 (肺气肿) 的 α -1抗胰蛋白酶;用于治疗地中海贫血或肾功能衰竭引起的贫血的红细胞生成素;用于治疗缺血性疾病的血管内皮生长因子、血管生成素-1和成纤维细胞生长因子;用于治疗闭塞血管的血栓调节蛋白和组织因子通路抑制剂,如例如动脉粥样硬化、血栓形成或栓塞中所见;用于治疗帕金森病的芳香族氨基酸脱羧酶 (AADC) 和酪氨酸羟化酶 (TH) 。

[0099] 可用于本文提供的组合物和靶向递送的蛋白质和化合物的实例包含以下呼吸道相关传染病的治疗性蛋白和其它化合物及疫苗蛋白衍生物,以及直接针对这些传染病的被动免疫球蛋白。合适的治疗性蛋白的实例包含例如 α -1-抗胰蛋白酶、囊性纤维化跨膜蛋白及其变体、表面活性剂-B、骨形态发生蛋白受体II型 (与肺动脉高压相关),以及各种癌症治疗。

[0100] 合适的疫苗或被动免疫的实例包含源自已经与严重急性呼吸综合征 (SARS-CoV1) 、普通感冒和非甲型、乙型或丙型肝炎相关的空气传播病原体的蛋白质,所述病原体包含人类呼吸道冠状病毒。SARS-CoV2是COVID-19的致病因子,并且对这种病毒具有特异性的抗体已经被描述。已被描述为可用于结合SARS-CoV2的人ACE2的刺突蛋白并具有中和活性的IgG抗体的实例包含例如LY-CoV555 (Eli Lilly) 、TY027 (Tychon) 、STI-1499和STI-2020 (COVI-GUARD; Sorrento) 、80R、ADI055689/56046 (Adimab) (Renn等人,《药理科学趋势

(Trends in Pharmacological Sciences)》,2020);BD-217、BD-218、BD-236(Cao等人,《细胞(Cell)》,182,73-84(2020))。已被描述为可用于结合SARS-COV2的人ACE2的受体结合结构域(RBD)并具有中和活性的IgG抗体的实例包含例如COV2-2196、COV2-2130、COV2-2165(Zost等人,《自然》,584,443-465(2020));BD-361、BD-368、BD-368-2(Cao等人,《细胞(Cell)》,182,73-84(2020));B38、H4(Y.Wu等人,《科学》10.1126/science.abc2241(2020);Jahanshahlu和Rezaei,《生物医学和药物治疗法(Biomedicine and Pharmacotherapy)》129(2020));S309、S315、S304(Pinto等人,《自然》,583,290-311(2020));CC6.29、CC6.30、CC6.33、CC12.1、CC12.3(Rogers等人,《科学》369,956-963(2020));JS016(Eli Lilly)、CA1、CB6-LALA、P2C-1F11/P2B-2F6/P2A-1A3、311mab-31B5311/32D4、COVA 2-15、414-1(Renn等人,《药理科学趋势(Trends in Pharmacological Sciences)》,2020)。已被描述为可用于结合SARS-COV1的人ACE2的刺突蛋白并具有中和活性的IgG抗体的实例包含例如m396和CR3104(Prabakaran等人,《生物化学杂志(Journal of Biological Chemistry)》,281,15829-15836(2006);ter Meulen等人,《美国科学公共图书馆(PLoS)》,3,7(2006))。已被描述为可用于结合SARS-COV1和SARS-CoV2的人ACE2的RBD或刺突蛋白并具有中和活性的IgG抗体的实例包含例如CR3022和47D11(Wang等人,《自然通讯(Nature Communications)》,11,nature.com/naturecommunications(2020))。

[0101] 其它靶病毒的实例包含来自正黏液病毒科(orthomyxoviridae family)的流感病毒,其包含:甲型流感、乙型流感和丙型流感。甲型病毒是毒性最强的人类病原体。已与大流行相关的甲型流感血清型包含:H1N1,其引起1918年的西班牙流感(Spanish Flu)和2009年的猪流感(Swine Flu);H2N2,其引起1957年的亚洲流感(Asian Flu);H3N2,其引起1968年的香港流感(Hong Kong Flu);H5N1,其引起2004年的禽流感(Bird Flu);H7N7;H1N2;H9N2;H7N2;H7N3;以及H10N7。已经描述了针对甲型流感的广泛中和抗体(broadly neutralizing antibody)。如本文所用,“广泛中和抗体”是指可以中和来自多个亚型的多个毒株的中和抗体。例如,CR6261[The Scripps Institute/Crucell]已被描述为单克隆抗体,其与广泛的流感病毒结合,包含1918年“西班牙流感(Spanish flu)”(SC1918/H1)以及2004年在越南从鸡传染给人的禽流感H5N1类别的病毒(Viet04/H5)。CR6261识别血凝素近膜端茎中高度保守的螺旋区,血凝素是流感病毒表面上的主要蛋白质。这一抗体描述于WO 2010/130636中,其通过引用并入本文。已描述了可用于针对H1N1和H5N1的另一种中和抗体F10[XOMA Ltd]。[Sui等人,《自然结构和分子生物学(Nature Structural and Molecular Biology)》(Sui等人,2009,16(3):265-73)]可以选择其它针对流感的抗体,例如Fab28和Fab49。参见例如WO 2010/140114和WO 2009/115972,其通过引用并入本文。还可以容易地选择又其它抗体,例如WO 2010/010466、美国公开专利公布US/2011/076265和WO 2008/156763中所述的那些抗体。

[0102] 其它靶病原体病毒包含砂粒病毒(包含funin、马丘波病毒和拉沙病毒(Lassa))、丝状病毒(包含马尔堡病毒(Marburg)和埃博拉病毒(Ebola))、汉坦病毒、小核糖核酸病毒(包含鼻病毒、艾柯病毒)、冠状病毒、副粘病毒、麻疹病毒、呼吸道合胞体病毒、披盖病毒、柯萨奇病毒、细小病毒B19、副流感病毒、腺病毒、呼肠孤病毒、来自痘病毒科(poxvirus family)的天花病毒(重型天花(Variola major)(Smallpox)) and 牛痘(Vaccinia)(牛痘(Cowpox))、以及水痘-带状疱疹病毒(伪狂犬病)。病毒性出血热是由沙粒病毒科

(arenavirus family) (拉沙热(Lassa fever)) (所述沙粒病毒科也与淋巴细胞性脉络丛脑膜炎(LCM)相关)、丝状病毒属(filovirus) (埃博拉病毒),以及汉坦病毒属(hantavirus) (普马拉病毒(puremala))的成员引起的。细小核糖核酸病毒(鼻病毒亚科)的成员与人类的感冒有关。冠状病毒科包含许多非人类病毒,如传染性支气管炎病毒(家禽)、猪传染性胃肠炎病毒(猪)、猪血凝脑脊髓炎病毒(猪)、猫传染性腹膜炎病毒(猫)、猫肠道冠状病毒(猫)、狗冠状病毒(狗)。副粘病毒科包含1型副流感病毒、3型副流感病毒、3型牛副流感病毒、风疹病毒(rubelavirus) (腮腺炎病毒(mumps virus))、2型副流感病毒、4型副流感病毒、新城疫病毒(鸡)、牛瘟、麻疹病毒(包含麻疹和犬瘟病病毒),以及肺病毒(包含呼吸道合胞体病毒(RSV))。细小病毒科包含猫细小病毒(猫肠炎)、猫泛白细胞减少病病毒(feline panleucopeniavirus)、犬细小病毒,以及猪细小病毒。腺病毒科包含引起呼吸性疾病的病毒(EX、AD7、ARD、O.B.)。

[0103] 还可以选择针对细菌病原体的中和抗体构建体用于本发明中。在一个实施例中,中和抗体构建体针对细菌本身。在另一个实施例中,中和抗体构建体针对由细菌产生的毒素。空气传播的细菌病原体的实例包含例如脑膜炎奈瑟菌(*Neisseria meningitidis*) (脑膜炎)、肺炎克雷伯菌(*Klebsiella pneumonia*) (肺炎)、铜绿假单胞菌(*Pseudomonas aeruginosa*) (肺炎)、伪鼻疽假单胞菌(*Pseudomonas pseudomallei*) (肺炎)、鼻疽假单胞菌(肺炎)、不动杆菌属(*Acinetobacter*) (肺炎)、卡他莫拉菌(*Moraxella catarrhalis*)、腔隙莫拉菌(*Moraxella lacunata*)、产碱杆菌属(*Alkaligenes*)、心杆菌属(*Cardiobacterium*)、流感嗜血杆菌(*Haemophilus influenzae*) (流感)、副流感嗜血杆菌(*Haemophilus parainfluenzae*)、百日咳博代氏杆菌(*Bordetella pertussis*) (百日咳)、土拉弗朗西斯菌(*Francisella tularensis*) (肺炎/发热)、肺炎军团菌(*Legionella pneumonia*) (军团病)、鹦鹉热衣原体(*Chlamydia psittaci*) (肺炎)、肺炎衣原体(*Chlamydia pneumoniae*) (肺炎)、结核分枝杆菌(*Mycobacterium tuberculosis*) (结核病(TB))、堪萨斯分枝杆菌(*Mycobacterium kansasii*) (TB)、鸟分枝杆菌(*Mycobacterium avium*) (肺炎)、星状诺卡氏菌(*Nocardia asteroides*) (肺炎)、炭疽杆菌(*Bacillus anthracis*) (炭疽)、金黄色葡萄球菌(*Staphylococcus aureus*) (肺炎)、酿脓链球菌(*Streptococcus pyogenes*) (猩红热)、肺炎链球菌(*Streptococcus pneumoniae*) (肺炎)、白喉杆菌(*Corynebacteria diphtheria*) (白喉)、肺炎支原体(*Mycoplasma pneumoniae*) (肺炎)。炭疽的致病因子是由炭疽杆菌产生的毒素。已经描述了针对保护剂(PA)的中和抗体,三种肽之一形成类毒素。另外两种多肽由致死因子(LF)和水肿因子(EF)组成。抗PA中和抗体已经描述为进行针对炭疽的被动免疫有效。参见例如美国专利号7,442,373;R.Sawada-Hirai等人,《基于免疫的疗法和疫苗的杂志(J Immune Based Ther Vaccines.)》2004;2:5。(2004年5月12日在线)。已经描述和/或可以生成又其它抗炭疽毒素中和抗体。类似地,可以使用针对其它细菌和/或细菌毒素的中和抗体来生成如本文所述的非IgG抗体。

[0104] 其它传染病可以由空气传播的真菌引起,所述真菌包含例如曲霉属物种(*Aspergillus species*)、伞枝梨头霉(*Absidia corymbifera*)、匍枝根霉(*Rhizopus stolonifer*)、密丛毛霉(*Mucor plumbeus*)、新型隐球菌(*Cryptococcus neoformans*)、荚膜组织胞浆菌(*Histoplasma capsulatum*)、皮炎芽生菌(*Blastomyces dermatitidis*)、粗球孢子菌(*Coccidioides immitis*)、青霉菌属物种(*Penicillium species*)、干草小多孢菌

(*Micropolyspora faeni*)、普通高温放线菌(*Thermoactinomyces vulgaris*)、交链格孢菌(*Alternaria alternate*)、分枝孢子菌属物种(*Cladosporium species*)、长蠕孢属(*Helminthosporium*),以及葡萄状穗霉属物种(*Stachybotrys species*)。

[0105] 除了影响人类的空气传播的传染病病况之外(上面描述了其中许多病况),可以使用根据本发明的被动免疫预防与直接接种鼻道有关的病况,例如可通过手指与鼻道的直接接触传播的病况。这些病况可以包含真菌感染(例如,运动员脚)、癣菌病,或者病毒、细菌、寄生虫、真菌以及可以通过直接接触传播的其它病原体。另外,多种病况影响家庭宠物、牛和其它家畜,以及其它动物。例如,在狗中,狗鼻窦曲霉菌对上呼吸道的感染引起显著疾病。在猫中,如果未处理,在鼻中起源的上呼吸道疾病或猫呼吸性疾病症候群会引起发病和死亡。牛易于被传染性牛鼻气管炎(通常称为IBR或红鼻)感染,所述鼻气管炎是一种急性牛传染性病毒疾病。另外,牛易于感染牛呼吸道合胞体病毒(BRSV),所述病毒引起轻度至重度呼吸性疾病并且可损害对其它疾病的抗性。又其它病原体和疾病对于本领域技术人员而言将是显而易见的。

[0106] 针对病原体的抗体并且特别是中和抗体(如本文具体确定的那些(例如,抗SARS-CoV2、抗SARS-CoV1、抗流感、抗埃博拉病毒、抗RSV))可用于生成类别转换或非IgG抗体。可以使用抗体噬菌体展示从最近接种季节性流感疫苗的供体、非免疫人类或自然感染的幸存者中筛选文库来鉴定具有广泛中和能力的单克隆抗体(mAb)。就流感而言,通过阻断病毒与宿主细胞的融合,已经鉴定了中和多于一种流感亚型的抗体。这种技术可用于其它感染以获得中和的单克隆抗体。参见例如US 5,811,524,该专利描述了抗呼吸道合胞体病毒(RSV)中和抗体的生成。本文所述的技术适用于其它病原体。此类抗体可以完整使用或其被修饰以生成人工或重组中和抗体构建体的序列(支架)。已经描述了此类方法[参见例如WO 2010/13036;WO 2009/115972;WO 2010/140114]。在一个实施例中,用免疫剂对小鼠、大鼠、仓鼠或其它宿主动物进行免疫以生成产生对免疫抗原具有结合特异性的抗体的淋巴细胞。在可替代的方法中,可以在体外对淋巴细胞进行免疫。人类抗体可以使用如噬菌体展示文库的技术来生产(Hoogenboom和Winter,《分子生物学杂志(J.Mol.Biol)》,1991,227:381, Marks等人,《分子生物学杂志(J.Mol.Biol.)》1991,222:581)。

[0107] 组合物和用途

[0108] 本文提供了含有至少一种rAAV原液(例如,rAAV9或rAAVhu68突变原液)和任选的载剂、赋形剂和/或防腐剂的组合物。rAAV原液是指多个rAAV载体,所述多个rAAV载体的量与例如在下文关于浓度和剂量单位的讨论中描述的量相同。

[0109] 在某些实施例中,组合物可含有至少第二种不同的rAAV原液。这第二种载体原液可以通过具有不同的AAV衣壳和/或不同的载体基因组而与第一种载体原液不同。在某些实施例中,本文所述的组合物可以含有表达本文所述的表达盒的不同载体,或另一种活性组分(例如,抗体构建体、另一种生物制剂和/或小分子药物)。

[0110] 如本文所用,“载剂”包含任何和所有溶剂、分散介质、媒剂、涂层、稀释剂、抗菌剂和抗真菌剂、等渗剂和吸收延迟剂、缓冲液、载体溶液、悬浮液、胶质物等。将此类培养基和药剂用于药物活性物质在本领域中是众所周知的。还可以将补充性活性成分并入到组合物中。短语“药学上可接受的”是指当向宿主施用时不会产生过敏或类似不良反应的分子实体和组合物。如脂质体、纳米胶囊、微颗粒、微球、脂质颗粒、囊泡等递送媒剂可以用于将本

发明的组合物引入到合适的宿主细胞中。具体地，rAAV载体递送的转基因可以被调配成用于递送或包封在脂质颗粒、脂质体、囊泡、纳米球或纳米颗粒等中。

[0111] 在一个实施例中，组合物包含适合于递送到受试者的最终调配物，所述组合物是例如缓冲到生理学上相容的pH和盐浓度的水性液体悬浮液。任选地，调配物中存在一种或多种表面活性剂。在另一个实施例中，可以将组合物作为稀释以向受试者施用的浓缩物运输。在其它实施例中，可以在施用时将组合物冻干并重构。

[0112] 可以从无毒的非离子型表面活性剂中选择合适的表面活性剂或表面活性剂的组合。在一个实施例中，选择以伯羟基为末端的双管能嵌段共聚物表面活性剂，例如如Pluronic®F68[BASF]，也称为Poloxamer 188，其具有中性pH，平均分子量为8400。可以选择其它表面活性剂和其它泊洛沙姆，即非离子型三嵌段共聚物，所述非离子型三嵌段共聚物由侧接聚氧乙烯(聚(环氧乙烷))的两个亲水链的聚氧丙烯(聚(环氧丙烷))的中心疏水链、SOLUTOL HS 15(聚乙二醇-15羟基硬脂酸酯)、LABRASOL(聚氧辛酸甘油酯)、聚氧10油醚、TWEEN(聚氧乙烯山梨糖脂脂肪酸酯)、乙醇和聚乙二醇构成。在一个实施例中，调配物含有泊洛沙姆。这些共聚物通常以字母“P”(对于泊洛沙姆)命名，后跟三个数字：前两位数字x 100给出了聚氧丙烯核的近似分子量，并且最后一位数字x 10给出了聚氧乙烯含量的百分比。在一个实施例中，选择了泊洛沙姆188。表面活性剂可以以悬浮液的至多约0.0005%至约0.001%的量存在。

[0113] 在一个实施例中，调配物缓冲液是总盐浓度为200mM，含0.001% (w/v) Pluronic F68的磷酸盐缓冲液(PBS) (最终调配物缓冲液，FFB)。

[0114] 以足够的量施用载体以转染细胞并提供足够水平的基因转移和表达，以便提供治疗益处，而不会产生过度的副作用或具有医学上可接受的生理作用，这可以由医学领域的技术人员确定。在某些实施例中，载体被配制用于经由鼻内递送装置递送，以靶向递送到鼻和/或鼻咽上皮细胞。在某些实施例中，载体被配制用于气溶胶递送装置，例如经由喷雾器或通过其它合适的装置。其它常规的和药学上可接受的施用途径包含但不限于直接输送到所需器官(例如，肺)、口服吸入、鞘内、气管内、动脉内、眼内、静脉内、肌内、皮下、皮内和其它肠胃外施用途径。在一个实施例中，使用鼻内粘膜雾化装置(LMA®MAD Nasal™-MAD110)鼻内施用载体。在另一个实施例中，使用振动网状喷雾器(Vibrating Mesh Nebulize) (Acrogen®Solo)或MADgic™喉粘膜雾化器(Laryngeal Mucosal Atomizer)以雾化形式在肺内施用载体。如果期望，可以组合施用途径。施用途径和其递送rAAV载体的使用也描述于以下公开的美国专利申请中：US 2018/0155412A1、US 2018/0243416A1、US 2014/0031418 A1和US 2019/0216841A1，其各自通过引用以其整体并入本文。

[0115] 病毒载体的剂量将主要取决于如所治疗的病况、患者的年龄、体重和健康状况等因素，并且因此在患者之间可能有所不同。例如，病毒载体的治疗有效人类剂量通常在以下范围内：含有约 10^9 至 4×10^{14} GC的AAV载体的剂量的约25至约1000微升到约5mL的水性悬浮液体。将调整剂量以平衡治疗益处与任何副作用，并且此类剂量可以根据采用重组载体的治疗应用而变化。可以监测转基因的表达水平以确定所得病毒载体的剂量频率，优选地含有微基因的AAV载体。任选地，与出于治疗目的描述的剂量方案类似的剂量方案可以用于使用本发明的组合物进行免疫。

[0116] 复制缺陷病毒组合物可以以剂量单位配制，所述剂量单位含有在以下范围内的量

的复制缺陷病毒:约 10^9 GC至约 10^{16} GC(用于治疗平均体重为70kg的受试者),包含所述范围内的所有整数或分数量,并且对于人类患者优选 10^{12} GC至 10^{14} GC。在一个实施例中,组合物被配制为每剂含有至少 10^9 、 2×10^9 、 3×10^9 、 4×10^9 、 5×10^9 、 6×10^9 、 7×10^9 、 8×10^9 或 9×10^9 GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{10} 、 2×10^{10} 、 3×10^{10} 、 4×10^{10} 、 5×10^{10} 、 6×10^{10} 、 7×10^{10} 、 8×10^{10} 或 9×10^{10} GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{11} 、 2×10^{11} 、 3×10^{11} 、 4×10^{11} 、 5×10^{11} 、 6×10^{11} 、 7×10^{11} 、 8×10^{11} 或 9×10^{11} GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{12} 、 2×10^{12} 、 3×10^{12} 、 4×10^{12} 、 5×10^{12} 、 6×10^{12} 、 7×10^{12} 、 8×10^{12} 或 9×10^{12} GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{13} 、 2×10^{13} 、 3×10^{13} 、 4×10^{13} 、 5×10^{13} 、 6×10^{13} 、 7×10^{13} 、 8×10^{13} 或 9×10^{13} GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{14} 、 2×10^{14} 、 3×10^{14} 、 4×10^{14} 、 5×10^{14} 、 6×10^{14} 、 7×10^{14} 、 8×10^{14} 或 9×10^{14} GC,包含所述范围内的所有整数或分数量。在一个实施例中,组合物被配制为每剂含有至少 10^{15} 、 2×10^{15} 、 3×10^{15} 、 4×10^{15} 、 5×10^{15} 、 6×10^{15} 、 7×10^{15} 、 8×10^{15} 或 9×10^{15} GC,包含所述范围内的所有整数或分数量。在一个实施例中,对于人类应用,剂量范围可以为每剂 10^{10} 至约 10^{12} GC,包含所述范围内的所有整数或分数量。在一个实施例中,对于人类应用,剂量范围可以为每剂 10^9 至约 7×10^{13} GC,包含所述范围内的所有整数或分数量。在一个实施例中,对于人类应用,剂量范围为 6.25×10^{12} GC至 5.00×10^{13} GC。在进一步的实施例中,剂量为约 6.25×10^{12} GC、约 1.25×10^{13} GC、约 2.50×10^{13} GC或约 5.00×10^{13} GC。在某些实施例中,剂量被平均分成两半,并施用于每个鼻孔。在某些实施例中,对于人类应用,剂量范围为 6.25×10^{12} GC至 5.00×10^{13} GC,以每个鼻孔0.2ml的两个等份施用,每个受试者递送的总体积为0.8ml。

[0117] 这些上述剂量可以以各种体积的载剂、赋形剂或缓冲液调配物施用,范围为约25至约1000微升或更大的体积,包含所述范围内的所有数字,这取决于待治疗区域的大小、所使用的病毒效价、施用途径以及所述方法的所期望的效果。在一个实施例中,载剂、赋形剂或缓冲液的体积为至少约25 μ L。在一个实施例中,体积为约50 μ L。在另一个实施例中,体积为约75 μ L。在另一个实施例中,体积为约100 μ L。在另一个实施例中,体积为约125 μ L。在另一个实施例中,体积为约150 μ L。在另一个实施例中,体积为约175 μ L。在又一个实施例中,体积为约200 μ L。在另一个实施例中,体积为约225 μ L。在又一个实施例中,体积为约250 μ L。在又一个实施例中,体积为约275 μ L。在又一个实施例中,体积为约300 μ L。在又一个实施例中,体积为约325 μ L。在另一个实施例中,体积为约350 μ L。在另一个实施例中,体积为约375 μ L。在另一个实施例中,体积为约400 μ L。在另一个实施例中,体积为约450 μ L。在另一个实施例中,体积为约500 μ L。在另一个实施例中,体积为约550 μ L。在另一个实施例中,体积为约600 μ L。在另一个实施例中,体积为约650 μ L。在另一个实施例中,体积为约700 μ L。在另一个实施例中,体积介于约700与约1000 μ L之间。

[0118] 在某些实施例中,重组载体可以通过向每个鼻孔喷射两次来进行鼻内给药。在一个实施例中,两次喷射通过交替喷向每个鼻孔来施用,例如,左鼻孔喷射,右鼻孔喷射,然后左鼻孔喷射,右鼻孔喷射。在某些实施例中,交替喷射之间可以存在延迟。例如,每个鼻孔可以接受多次喷射,间隔时间为约10至60秒或20至40秒或约30秒至数分钟或更长。此类喷射在每次喷射中可以递送例如约150 μ L至300 μ L或约250 μ L,以实现约200 μ L至约600 μ L、400 μ L

至700 μ L,或450 μ L至1000 μ L的总给药体积。

[0119] 在某些实施例中,重组AAV载体可以鼻内给药以实现给药后(例如,施用载体后一周至四周或约两周)在洗鼻液中测量的转基因表达产物的浓度为5-20ng/ml。获得受试者的洗鼻液的方法以及转基因表达产物的定量方法都是常规的。

[0120] 对于其它施用途径,例如静脉内或肌内,剂量水平会比鼻内递送高。例如,此类悬浮液的剂量体积可以为约1mL至约25mL,剂量至多约 2.5×10^{15} GC。

[0121] 在某些实施例中,鼻内递送装置提供了喷雾雾化器,其递送平均尺寸范围为约30微米至约100微米的颗粒雾尺寸。在某些实施例中,平均尺寸范围为约10微米至约50微米。文献中已经描述了合适的装置,并且一些装置是可商购获得的,例如LMA MAD NASAL™ (Teleflex Medical;Ireland);Teleflex VaxINator™ (Teleflex Medical;Ireland);来自Kurve Technologies的Controlled Particle Dispersion® (CPD)。还参见PG Djupesland,《药物递送和移动研究 (Drug Deliv and Transl.Res)》(2013)3:42-62。在某些实施例中,控制递送颗粒尺寸和体积以便优先靶向鼻上皮细胞并且最小化对肺的靶向。在其它实施例中,颗粒雾为约0.1微米至约20微米或更小以便递送到肺细胞。此类较小的颗粒尺寸可以最小化在鼻上皮的滞留。

[0122] 一种装置以约16微米至约22微米的平均直径对颗粒进行雾化。雾可以直接递送到通过3.5-mm柔性纤维支气管镜(纽约梅尔维尔奥林巴斯(Olympus,Melville,NY))的抽吸通道插入的气管支气管树。其它合适的递送装置可以包含喉-气管粘膜雾化器,其提供了经由声带跨越上呼吸道的施用。其穿过声带并沿喉罩向下或进入鼻腔。以约30微米至约100微米的平均直径雾化液滴。标准装置的尖端直径为约0.18in(4.6mm),长度为约4.5-8.5英寸,并且通过抽吸通道插入并越过窥镜的远侧尖端推进大约3mm。可以施用的剂量是10个等份(每个大约150 μ l)的喷射到右主干支气管的含盐水的对照或rAAV。

[0123] 在一个实施例中,提供了一种冷冻组合物,其含有本文所述的缓冲溶液中的呈冷冻形式的rAAV。任选地,这一组合物中存在一种或多种表面活性剂(例如,Pluronic F68)、稳定剂或防腐剂。适当地,在使用时,将组合物解冻,并用适当的稀释剂(例如,无菌盐水或缓冲盐水)滴定到所需剂量。

[0124] 在一个实施例中,提供了一种组合物,其包括一个或多个来自以下基序的外源上皮细胞靶向肽:N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)和任选的侧翼接头序列,以及一种或多种生理学上相容的载剂、赋形剂和/或水性悬浮基质。进一步提供了包括编码所述靶向肽的核酸序列的组合物。在某些实施例中,靶向肽是SEQ ID NO:40的靶向肽并且被SEQ ID NO:54的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:38的靶向肽并且被SEQ ID NO:50的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:46的靶向肽并且被SEQ ID NO:56的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:43的靶向肽并且被SEQ ID NO:52的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:39的靶向肽并且被SEQ ID NO:55的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:42的靶向肽并且被SEQ ID NO:51的核酸序列或与其至少约70%相同的序列编码。在某些实施例中,靶向肽是SEQ ID NO:41的靶向肽并且被SEQ ID NO:53的核酸序列或与其至少约70%相同的序列编码。

[0125] 在另一个实施例中,提供了一种融合多肽或蛋白质,其包含一个或多个来自以下基序的外源脑内皮细胞靶向肽: $N\text{-}x\text{-}(T/I/V/A)\text{-}(K/R)$ (SEQ ID NO:47) 和包括至少一种多肽或蛋白质的融合伴侣。进一步提供了编码所述融合多肽或蛋白质的核酸序列。

[0126] 在某些实施例中,提供了一种组合物,其包括融合多肽或蛋白质,或编码所述融合多肽或蛋白质的核酸序列,或含有所述融合多肽或蛋白质的纳米颗粒。组合物可进一步包括生理学上相容的载剂、赋形剂和/或水性悬浮基质中的一种或多种。

[0127] 在某些实施例中,编码融合多肽蛋白的核酸序列被包封在脂质纳米颗粒(LNP)中。如本文所用,短语“脂质纳米颗粒”或“纳米颗粒”是指包括一种或多种脂质(例如,阳离子脂质、非阳离子脂质和PEG修饰的脂质)的转移载体。优选地,脂质纳米颗粒被配制以将一个或多个核酸序列递送至一个或多个靶细胞(例如,肝和/或肌肉)。合适的脂质的实例包含例如磷脂酰化合物(例如,磷脂酰甘油、磷脂酰胆碱、磷脂酰丝氨酸、磷脂酰乙醇胺、鞘脂、脑苷脂和神经节苷脂)。还考虑使用聚合物作为转移媒剂,无论是单独使用还是与其它转移媒剂组合使用。合适的聚合物可以包含例如聚丙烯酸酯、聚氰基丙烯酸烷基酯、聚交酯、聚交酯-聚乙交酯共聚物、聚己内酯、葡聚糖、白蛋白、明胶、藻酸盐、胶原、壳聚糖、环糊精、树枝状大分子和聚乙烯亚胺。在一个实施例中,转移载体的选择基于其促进其中包封的核酸序列转染到靶细胞的能力。用于核酸序列的可用脂质纳米颗粒包括阳离子脂质以包封此类核酸序列和/或增强此类核酸序列到靶细胞的递送,所述靶细胞将充当蛋白质生产的储库。如本文所用,短语“阳离子脂质”是指在所选pH,如生理pH下携带净正电荷的多种脂质物种中的任何脂质物种。可以通过包含采用一种或多种阳离子脂质、非阳离子脂质和经PEG修饰的脂质的不同比率的多组分脂质混合物来制备所考虑的脂质纳米颗粒。文献中已经描述了若干阳离子脂质,其中许多是可商购获得的。参见例如,WO2014/089486、US 2018/0353616A1和US 8,853,377B2,所述文献通过引用并入本文。在某些实施例中,LNP调配物使用常规程序进行,包括胆固醇、可离子化脂质、辅助脂质、PEG-脂质和聚合物,在包封的核酸序列周围形成脂质双层(Kowalski等人,2019,《分子疗法(Mol. Ther.)》27(4):710-728)。在一些实施例中,LNP包括阳离子脂质(即 $N\text{-}[1\text{-}(2,3\text{-二油酰氧基})\text{丙基}]\text{-}N,N,N\text{-三甲基氯化铵}$ (DOTMA) 或 $1,2\text{-二油酰基-}3\text{-三甲基铵丙烷}$ (DOTAP)) 与辅助脂质DOPE。在一些实施例中,LNP包括可电离脂质Dlin-MC3-DMA可电离脂质,或基于二酮哌嗪的可电离脂质(cKK-E12)。在一些实施例中,聚合物包括聚乙烯亚胺(PEI) 或聚(β -氨基)酯(PBAE)。参见例如,WO2014/089486、US 2018/0353616A1、US2013/0037977A1、WO2015/074085A1、US9670152B2和US 8,853,377B2,所述文献通过引用并入本文。

[0128] 在某些实施例中,组合物,例如具有带有 $N\text{-}x\text{-}(T/I/V/A)\text{-}(K/R)$ (SEQ ID NO:47) 肽和任选接头序列的修饰衣壳的rAAV,融合多肽或蛋白质,或包括纳米颗粒或化学部分的缀合物,可用于向有此需要的患者递送治疗。在某些实施例中,所述方法用于对脑内皮细胞的靶向疗法。在某些实施例中,所述方法用于通过递送MCT8蛋白(例如UniProt ID No: P36021) 或体内表达MCT8的基因来治疗Allan-Herndon-Dudley病。在其它实施例中,所述方法用于对肺的靶向疗法。在某些实施例中,递送的产物是可溶性Ace2蛋白(例如,hAce2诱饵或hAce2诱饵融合物)、抗SARS抗体、抗SARS-CoV2抗体、抗流感抗体或囊性纤维化跨膜蛋白。还参见omim.org/entry/300523,其内容通过引用并入本文。还参见2021年1月29日提交的美国临时申请号63/143,614,2021年3月12日提交的美国临时申请号63/165,051,以及

2021年3月26日提交的美国专利申请号63/166,686,2021年6月25日提交的美国临时专利申请号63/215,159,2021年10月8日提交的美国临时申请号63/253,654,它们均通过引用并入本文。

[0129] 在某些实施例中,具有本文所述修饰衣壳的rAAV可在联合治疗方案中递送,所述方案进一步包括一种或多种其它活性组分。在某些实施例中,所述方案可能涉及共同施用免疫调节组分。此类免疫调节方案可以包含例如但不限于免疫抑制剂如糖皮质激素、类固醇、抗代谢药、T细胞抑制剂、大环内酯类药物(例如,雷帕霉素或雷帕霉素类似物(rapalog)),以及包含烷化剂、抗代谢药、细胞毒抗生素、抗体或对免疫亲和素有活性的药剂在内的细胞抑制剂。免疫抑制剂可以包含氮芥、亚硝基脲、铂类化合物、甲氨蝶呤、硫唑嘌呤、巯基嘌呤、氟尿嘧啶、更生霉素、葱环类药物、丝裂霉素C、博莱霉素、光神霉素、IL-2受体(CD25)或CD3导向的抗体、抗IL-2抗体、环孢素、他克莫司、西罗莫司、IFN- β 、IFN- γ 、阿片类药物或TNF- α (肿瘤坏死因子- α)结合剂。在某些实施例中,免疫抑制疗法可以在基因疗法施用之前开始。此类疗法可以涉及在同一天共同施用两种或更多种药物(例如,泼尼松、麦考酚吗乙酯(MMF)和/或西罗莫司(即雷帕霉素))。可以在基因疗法施用之后以相同的剂量或经过调整的剂量继续使用这些药物中的一种或多种。根据需要,此类疗法可以持续约1周、约15天、约30天、约45天、60天或更长时间。又其它共治疗可以包含例如抗IgG酶,其已被描述为可用于耗尽抗AAV抗体(并且因此可以允许对所选AAV衣壳的抗体的测试高于阈值水平的患者施用),和/或递送抗FcRN抗体,其例如描述于2020年6月17日提交的题为“治疗基因疗法患者的组合物和方法”的美国临时专利申请号63/040,381号中,和/或a)类固醇或类固醇组合和/或(b) IgG清除酶,(c) Fc-IgE结合抑制剂;(d) Fc-IgM结合抑制剂;(e) Fc-IgA结合抑制剂;和/或(f) γ 干扰素中的一种或多种。

[0130] 抗体“Fc区”是指可结晶的片段,其是抗体中与细胞表面受体(Fc受体)相互作用的区域。在一个实施例中,Fc区是人IgG1 Fc。在一个实施例中,Fc区是人IgG2 Fc。在一个实施例中,Fc区是人IgG4 Fc。在一个实施例中,Fc区是工程化的Fc片段参见例如Lobner, Elisabeth等人,“工程化的IgG1-Fc——一个片段就能把它们全部结合起来(Engineered IgG1-Fc-one fragment to bind them all.)”《免疫综述(Immunological reviews)》270.1(2016):113-131;Saxena,Abhishek和Donghui Wu.“治疗性Fc工程化的进展——调节IgG相关的效应功能和血清半衰期(Advances in therapeutic Fc engineering-modulation of IgG-Associated effector functions and serum half-life.)”《免疫学前沿(Frontiers in immunology)》7(2016);Irani,Vashti等人,“人IgG亚类的分子特性及其对设计针对传染病的治疗性单克隆抗体的影响(Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases.)”《分子免疫学(Molecular immunology)》67.2(2015):171-182;Rath,Timo等人,“Fc融合蛋白和FcRn:对更持久和更有效的治疗的结构见解(Fc-fusion proteins and FcRn:structural insights for longer-lasting and more effective therapeutics.)”《生物技术关键综述(Critical reviews in biotechnology)》35.2(2015):235-254;和Invivogen,《用于治疗用途的IgG-Fc工程化(IgG-Fc Engineering For Therapeutic Use)》,invivogen.com/docs/Insight200605.pdf,2006年4月;其各自通过引用并入本文。

[0131] 抗体“铰链区”是IgG和IgA免疫球蛋白类重链的柔性氨基酸部分,其通过二硫键连接这两条链。

[0132] “免疫球蛋白分子”是含有共价偶联在一起并且能够与抗原特异性结合的免疫球蛋白重链和免疫球蛋白轻链的免疫活性部分的蛋白质。免疫球蛋白分子是任何类型(例如,IgG、IgE、IgM、IgD、IgA和IgY)、类(例如,IgG1、IgG2、IgG3、IgG4、IgA1和IgA2)或亚类的免疫球蛋白分子。术语“抗体”和“免疫球蛋白”在本文可互换使用。

[0133] “免疫球蛋白重链”是含有免疫球蛋白抗原结合结构域的至少一部分和免疫球蛋白重链可变区的至少一部分或免疫球蛋白重链恒定区的至少一部分的多肽。因此,免疫球蛋白衍生的重链与免疫球蛋白基因超家族的成员具有显著的氨基酸序列同源性区域。例如,Fab片段中的重链是免疫球蛋白衍生的重链。

[0134] “免疫球蛋白轻链”是含有免疫球蛋白抗原结合结构域的至少一部分和免疫球蛋白轻链可变区的至少一部分或免疫球蛋白轻链恒定区的至少一部分的多肽。因此,免疫球蛋白衍生的轻链与免疫球蛋白基因超家族的成员具有显著的氨基酸序列同源性区域。

[0135] “中和抗体效价”(NAb效价)是产生多少中和抗体(例如,抗AAV NAb)来中和其靶表位(例如,AAV)的生理效应的量度。抗AAV NAb效价可以如在例如Calcedo,R等人,“针对腺相关病毒的中和抗体的世界范围的流行病学(Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses.)”《传染病杂志(Journal of Infectious Diseases)》,2009,199(3):p.381-390中所述测量,其通过引用并入本文。

[0136] 如本文所用,除非另有说明,否则vp蛋白的“亚群体”是指一组vp蛋白,所述一组vp蛋白具有至少一个限定的共同特性,并且由至少一个组成员到少于参考组的所有成员组成。例如,除非另有说明,否则vp1蛋白的“亚群体”是至少一种(1)vp1蛋白,并且少于组装的AAV衣壳中的所有vp1蛋白。除非另有说明,否则vp3蛋白的“亚群体”可以是少于组装的AAV衣壳中的所有vp3蛋白的一种(1)vp3蛋白。例如,vp1蛋白可以是vp蛋白的亚群体;vp2蛋白可以是vp蛋白的单独的亚群体,并且vp3是组装的AAV衣壳中的vp蛋白的又另外的亚群体。在另一个实例中,vp1、vp2和vp3蛋白可以含有具有不同修饰的亚群体,例如,至少一种、两种、三种或四种高度脱酰胺化的天冬酰胺,例如在天冬酰胺-甘氨酸对处。除非另有说明,否则高度脱酰胺是指与参考氨基酸位置的预测氨基酸序列相比,在参考氨基酸位置处至少45%脱酰胺、至少50%脱酰胺、至少60%脱酰胺、至少65%脱酰胺、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、97%、99%、至多约100%脱酰胺。此类百分比可以使用2D凝胶、质谱技术或其它合适的技术来确定。

[0137] 本文所使用的rAAV的“原种”是指rAAV群体。尽管它们的衣壳蛋白由于脱酰胺作用而具有异源性,但是预计原种中的rAAV共有相同的载体基因组。原种可以包括具有衣壳的rAAV,所述衣壳具有例如所选AAV衣壳蛋白和所选生产系统的异源脱酰胺模式特征。原种可以由单一生产系统生产或由生产系统的多次运行汇集。可以选择各种生产系统,包括但不限于本文所描述的那些。参见例如2019年9月6日公开的WO 2019/168961,包含提供AAV9的脱酰胺模式的表G和2018年9月7日提交的WO 2020/160582。还参见例如于2020年11月5日公开的WO 2020/223231(rh91,包含带有脱酰胺模式的表格),于2020年8月14日提交的美国临时专利申请号63/065,616,和于2020年11月4日提交的美国临时专利申请号63/109,734,以及于2021年8月13日提交的国际专利申请号PCT/US21/45945,其均通过引用以其整体并入

本文。

[0138] 缩写“sc”是指自身互补的。“自互补AAV”是指其中由重组AAV核酸序列所携带的编码区已经被设计成形成分子内双链DNA模板的构建体。感染后,未等待细胞介导的第二条链合成,而是两条互补的半scAAV将缔合以形成易于立即复制和转录的一条双链DNA(dsDNA)。参见例如,D M McCarty等人,“自互补重组腺相关病毒(scAAV)载体独立于DNA合成而促进高效转导(Self-complementary recombinant adeno-associated virus(scAAV) vectors promote efficient transduction independently of DNA synthesis)”,《基因疗法》,(2001年8月),第8卷,第16期,第1248-1254页。自互补AAV描述于例如美国专利号6,596,535;7,125,717;和7,456,683,其各自通过引用以其整体并入本文。

[0139] 如本文所用,术语“可操作地连接”是指与所关注的基因邻接的表达控制序列以及以反式或在远处起作用以控制所关注的基因的表达控制序列两者。

[0140] 术语“异源性”当结合蛋白质或核酸使用时指示蛋白质或核酸包括在自然界中未发现彼此间的相同关系的两个或更多个序列或子序列。例如,核酸通常是重组产生的,具有两个或更多个来自无关基因的序列,这些序列被排列以产生新的功能性核酸。例如,在一个实施例中,所述核酸具有来自一个基因的启动子,其被排列以指导来自不同基因的编码序列的表达。因此,关于编码序列,启动子是异源的。

[0141] “复制缺陷型病毒”或“病毒载体”是指其中含有所关注的基因的表达盒包装在病毒衣壳或包膜中的合成或人工病毒颗粒,其中同样包装在病毒衣壳或包膜内的任何病毒基因组序列均是复制缺陷型的;即,其不能产生子代病毒颗粒,但保留了感染靶细胞的能力。在一个实施例中,病毒载体的基因组不包含对复制所需的酶进行编码的基因(基因组可以被工程化成“无肠的(gutless)”-仅含有所关注的转基因,其侧接扩增和包装人工基因组所需的信号),但是这些基因可以在产生期间供应。因此,这被认为可以安全地用于基因疗法,因为除非存在复制所需的病毒酶,否则不会发生通过子代病毒颗粒进行的复制和感染。

[0142] “重组AAV”或“rAAV”是含有两个元件的DNA酶抗性病毒颗粒,所述两个元件即AAV衣壳和至少含有包装在AAV衣壳内的非AAV编码序列的载体基因组。在某些实施例中,衣壳含有约60个由vp1蛋白、vp2蛋白和vp3蛋白构成的蛋白质,它们自组装形成衣壳。除非另有说明,否则“重组AAV”或“rAAV”可与短语“rAAV载体”互换使用。rAAV是“复制缺陷型病毒”或“病毒载体”,因为其缺少任何功能性AAV rep基因或功能性AAV cap基因并且不能产生后代。在某些实施例中,仅AAV序列是AAV反向末端重复序列(ITR),通常定位在载体基因组的5'和3'最端处,以允许定位在ITR之间的基因和调控序列包装在AAV衣壳内。

[0143] 术语“核酸酶抗性”表示AAV衣壳组装在表达盒周围,所述表达盒被设计用于将转基因递送至宿主细胞并保护这些包装的基因组序列在核酸酶孵育步骤期间免于降解(消化),核酸酶孵育步骤被设计用于从生产过程中除去可能存在的污染核酸。

[0144] 如本文所用,“载体基因组”是指包装在形成病毒颗粒的细小病毒(例如,rAAV)衣壳内部的核酸序列。此类核酸序列含有AAV反向末端重复序列(ITR)。在本文的实例中,载体基因组至少含有从5'至3'的AAV 5' ITR、编码序列(即,转基因)和AAV 3' ITR。可以选择来自AAV2(与所述衣壳来源不同的AAV)或除非全长ITR之外的ITR。在某些实施例中,ITR来自与在产生或反式补充AAV期间提供rep功能的AAV来源相同的AAV。进一步地,可以使用其它ITR,例如,自互补(scAAV) ITR。单链AAV和自互补(sc)AAV两者都涵盖在rAAV内。转基因是与

载体序列异源的核酸编码序列,所述核酸编码序列编码多肽、蛋白质、功能性RNA分子(例如,miRNA、miRNA抑制剂)或其它所关注的基因产物。核酸编码序列以允许转基因在靶组织的细胞中转录、翻译和/或表达的方式与调控组分可操作地连接。本文更详细地讨论了载体基因组的合适组分。在一个例子中,“载体基因组”从5'到3'至少含有载体特异性序列,与调控控制序列(指导序列在靶细胞中的表达)可操作地连接的编码所关注的蛋白质的核酸序列,其中载体特异性序列可以是将载体基因组特异性包装入病毒载体衣壳或包膜蛋白中的末端重复序列。例如,AAV反向末端重复序列用于包装到AAV和某些其它细小病毒衣壳中。

[0145] 如本文所用,“可操作地连接的”序列包含与所关注的基因邻接的表达控制序列和以反式或在远处起作用以控制所关注的基因的表达控制序列。

[0146] 在某些实施例中,用于制造rAAV的非病毒遗传元件将被称为载体(例如,生产载体)。在某些实施例中,这些载体是质粒,但考虑使用其它合适的遗传元件。此类生产质粒可以编码在rAAV生产期间表达的序列,例如生产rAAV所需的没有被包装入rAAV的AAV衣壳或rep蛋白。可替代地,此类生产质粒可以携带包装入rAAV的载体基因组。

[0147] 如本文所用,“亲本衣壳”是指选自细小病毒或其它病毒(例如,AAV、腺病毒、HSV、RSV等)的非突变或非修饰的衣壳。在某些实施例中,亲本衣壳包含任何天然存在的AAV衣壳,其包括编码衣壳蛋白(即vp蛋白)的野生型基因组,其中衣壳蛋白指导AAV转导和/或组织特异性趋向性。在一些实施例中,亲本衣壳选自天然靶向CNS的AAV。在其它实施例中,亲本衣壳选自非天然靶向CNS的AAV。

[0148] 如本文所用,“变体衣壳”或“变体AAV”或“变体AAV衣壳”是指修饰的衣壳或突变的衣壳,其中衣壳蛋白包括组织特异性靶向肽的插入。

[0149] 如本文所用,“表达盒”是指包括生物学上有用的核酸序列(例如,编码蛋白质、酶或其它有用的基因产物的基因cDNA、mRNA等)和与其可操作连接的调控序列的核酸分子,所述调控序列引导或调节核酸序列和其基因产物的转录、翻译和/或表达。如本文所用,“可操作地连接的”序列包含与核酸序列邻接或不邻接的调控序列和以反式或顺式核酸序列作用的调控序列两者。此类调控序列通常包含例如启动子、增强子、内含子、Kozak序列、聚腺苷酸化序列和TATA信号中的一者或多者。表达盒可以含有基因序列上游(5'处)的调控序列,例如启动子、增强子、内含子等中的一者或多者以及增强子,或基因序列下游(3'处)的调控序列中的一者或多者,例如包括聚腺苷酸化位点的3'非翻译区(3'UTR),以及其它元件。在某些实施例中,调控序列与基因产物的核酸序列可操作地连接,其中调控序列通过插入的核酸序列,即5'非翻译区(5'UTR)与基因产物的核酸序列分离。在某些实施例中,表达盒包括一种或多种基因产物的核酸序列。在一些实施例中,表达盒可以是单顺反子表达盒或双顺反子表达盒。在其它实施例中,术语“转基因”是指来自插入到靶细胞中的外源的一个或多个DNA序列。

[0150] 在本发明的上下文中,术语“翻译”涉及核糖体的过程,其中mRNA链控制氨基酸序列的组装以产生蛋白质或肽。

[0151] 术语“表达”在本文中以其最广泛的含义使用,并且包括RNA或RNA和蛋白质的产生。表达可以是暂时的或可以是稳定的。

[0152] 当提及核酸或其片段时,术语“基本同源性”或“基本类似性”表示当与另一个核酸(或其互补链)的适当核苷酸插入或缺失进行最佳比对时,至少约95至99%的比对序列具有

核苷酸序列同一性。优选地,同源性在全长序列、或其开放阅读框或长度为至少15个核苷酸的另一个合适的片段上。本文描述了合适片段的实例。

[0153] 在核酸序列的上下文中,术语“序列同一性”、“序列同一性百分比”或“相同百分比”是指两个序列中的残基在比对以获得最大对应性时是相同的。期望序列同一性比较的长度可以超过基因组的全长、基因编码序列的全长或至少约500至5000个核苷酸的片段。然而,也可能期望较小片段之间的同一性,例如至少约九个核苷酸,通常至少约20至24个核苷酸、至少约28至32个核苷酸、至少约36个或更多个核苷酸。类似地,对于氨基酸序列,在蛋白质的全长或其片段上,可以容易地确定“序列同一性百分比”。合适地,片段长度为至少约8个氨基酸,并且可以至多约700个氨基酸。本文描述了合适片段的实例。

[0154] 当提及氨基酸或其片段时,术语“基本同源性”或“基本类似性”表示当与另一个氨基酸(或其互补链)的适当氨基酸插入或缺失进行最佳比对时,至少约95到99%的比对序列具有氨基酸序列同一性。优选地,同源性是超过全长序列或其蛋白质,例如,免疫球蛋白区域或结构域,AAV cap蛋白或其长度为至少8个氨基酸或更期望地至少15个氨基酸的片段。本文描述了合适片段的实例。

[0155] 术语“高度保守的”意指至少80%同一性、优选地至少90%同一性,并且更优选地超过97%同一性。通过使用本领域技术人员已知的算法和计算机程序,本领域技术人员可以容易地确定同一性。

[0156] 通常,当提及两种不同的腺相关病毒之间的“同一性”、“同源性”或“类似性”时,参考“比对”序列来确定“同一性”、“同源性”或“类似性”。“比对”序列或“比对”是指与参考序列相比,通常含有对丢失的或另外的碱基或氨基酸的校正的多个核酸序列或蛋白质(氨基酸)序列。使用多种公开或可商购获得的多序列比对程序中的任一种进行比对。此类程序的实例包含“Clustal Ω ”、“Clustal W”、“CAP序列组装”、“MAP”和“MEME”,这些程序可通过因特网上的Web服务器进行访问。此类程序的其它来源是本领域技术人员已知的。可替代地,也使用了载体NTI实用程序。本领域已知的许多算法可以用于测量核苷酸序列同一性,包含上述程序中包含的那些算法。作为另一个实例,可以使用GCG 6.1版本的程序FastaTM比较多核苷酸序列。FastaTM提供了查询序列与搜索序列之间最佳重叠区的比对和序列同一性百分比。例如,核酸序列之间的序列同一性百分比可以是使用如GCG 6.1版本中所提供的采用其默认参数(字号6和评分矩阵的NOPAM系数)的FastaTM所确定的,所述程序通过引用并入本文。多个序列比对程序也可用于氨基酸序列,这些程序例如“Clustal Ω ”、“Clustal X”、“MAP”、“PIMA”、“MSA”、“BLOCKMAKER”、“MEME”和“Match-Box”程序。通常,以默认设置使用这些程序中的任何程序,尽管本领域技术人员可以根据需要改变这些设置。可替代地,本领域技术人员可以利用另一种算法或计算机程序,所述算法或计算机程序提供至少与通过参考算法和程序所提供的一样水平的同一性或比对。参见例如,J.D.Thomson等人,《核酸研究(Nucl.Acids.Res.)》,“多个序列比对的全面比较(A comprehensive comparison of multiple sequence alignments)”,27(13):2682-2690(1999)。

[0157] 可以基于动物模型而不是人类患者来确定有效量。

[0158] 如上所述,除非另有说明,否则术语“约”在用于修饰数值时意指相对于给定参考 $\pm 10\%$ 的变化($\pm 10\%$,例如, $\pm 1\%$ 、 $\pm 2\%$ 、 $\pm 3\%$ 、 $\pm 4\%$ 、 $\pm 5\%$ 、 $\pm 6\%$ 、 $\pm 7\%$ 、 $\pm 8\%$ 、 $\pm 9\%$ 、 ± 10 或其间的值)。

[0159] 在某些情况下,术语“E+#”或术语“e+#”用于指代指数。例如,“5E10”或“5e10”是 5×10^{10} 。这些术语可以互换地使用。

[0160] 在整个说明书和权利要求书中使用的术语“包括(comprise)”和“含有(contain)”及其变体(包含“包括(comprises)”,“包括(comprising)”,“含有(contains)”和“含有(containing)”)等其它变体)包括其它组分、元件、整数、步骤等。术语“由……组成(consists of)”或“由……组成(consisting of)”不包括其它组件、元件、整数、步骤等。

[0161] 应当注意,术语“一个”或“一种”是指一个或多个/一种或多种,例如“增强子”应理解为代表一个或多个/一种或多种增强子。如此,术语“一个”(或“一种”)、“一个或多个/一种或多种”和“至少一个/种”在本文中可互换地使用。

[0162] 关于这些发明的描述,意图是本文所述的每种组合物在另一个实施例中可用于本发明的方法。另外,还预期在另一个实施例中,本文所描述的用于所述方法中的组合物中的每种组合物本身也是本发明的实施例。

[0163] 除非在本说明书中另有定义,否则本文所用的技术和科学术语具有与本发明所属领域的普通技术人员通常所理解的含义相同的含义,并且通过参考公开的文本为本领域技术人员提供了对本申请中使用的许多术语的一般指引。

[0164] 实例

[0165] 下面的实例只是说明性的,并不是对本文所述发明的限制。

[0166] 实例1.初次筛选

[0167] 已经表明,小肽插入AAV衣壳表面的柔性环中可以介导与新的细胞受体的相互作用。在加州理工学院发现的一个案例(AAV9-PHP.B)中,插入AAV9上的HVR8环的七个氨基酸肽介导了与Ly6a(一些小鼠品系的脑血管上的GPI锚定受体)的相互作用。这种相互作用驱动AAV9-PHP.B跨越血脑屏障(BBB)运输,导致脑细胞的转导是AAV9的约50倍。在这项工作中,我们寻找能够与BBB上的细胞膜靶标结合,从而具有驱动AAV9衣壳跨越BBB的潜力的肽插入物。

[0168] 我们试图通过首先调查现有的学术和专利文献中可能有与脑血管细胞相互作用的潜力的肽序列来解决AAV-BBB问题。我们发现这些肽的以下来源:

[0169] • 公开的噬菌体展示实验的结果,其中针对原代脑内皮细胞淘选噬菌体展示文库;

[0170] • 与已知的BBB驻留膜蛋白的天然配体肽;

[0171] • 靶向BBB驻留膜蛋白的抗体的CDR;

[0172] • 导致脑炎的黄病毒的病毒外壳蛋白;以及

[0173] • 具有针对GPI锚定点的细胞结合活性的细菌毒素。

[0174] 我们生成了AAV9插入突变体文库,其含有来自这些来源的数百个肽,所有的肽都单独插入在HVR8位点(在588和589位之间,基于SEQ ID NO:44的AAV9衣壳氨基酸序列的编号)。每个肽通常以多种形式存在于文库中,其不同之处在于:1)插入的肽的长度;2)肽两侧柔性GSG或GG接头序列的存在。肽也使用多个同义密码子编码,使得我们可以在筛选中独立观察复制活性。

[0175] 此外,我们生成了在具有已知或疑似配体肽的AAVhu68衣壳的HVR8中的插入变体文库,所述配体肽靶向血脑屏障(BBB)受体(在588和589位之间,基于SEQ ID NO:45的

AAVhu68衣壳氨基酸序列的编号)。这些是：

- [0176] • 与哺乳动物脑内皮结合的肽(已公开的噬菌体显示数据)；
- [0177] • 经典的RMT受体配体(例如,Tf)；
- [0178] • 针对RMT受体的mAb的CDR(例如,抗TfR)；以及
- [0179] • 导致脑炎的黄病毒的外壳蛋白。

[0180] 作为对照,也包含PHP.B肽(C57/BL6的阳性对照,Balb/c&NHP的阴性对照)。每个肽以多种方式编码(有和无接头,并且在若干同义DNA序列中)。

[0181] 我们以高剂量向两个小鼠品系和一个非人灵长类动物静脉内(IV)注射这一文库。在2-3周的生存期后,对动物进行尸检,并收集组织。我们从CNS和其它组织中提取了AAV载体的DNA基因组,并对其进行了下一代测序(NGS)。载体变体使它们自己的衣壳基因变体壳体化,从而允许我们通过所关注的组织中衣壳基因变体的相对丰度来追踪衣壳活性。我们通过计算相对于其在注射的文库混合物中的丰度归一化的每个变体在CNS中的丰度,来对文库中每个变体的BBB活性(“富集分数”)进行评分。

[0182] 在小鼠研究中,C57/BL6小鼠中最多的脑富集的HVR8插入物为:TLAVPFK(SEQ ID NO:49)(PHP.B),阳性对照PHP.B独立出现了3次,作为富集最多的命中。三个具有同义密码子的PHP.B肽是独立富集的。其它若干肽也在脑中富集。图1A和1B显示了筛选中最佳小鼠脑命中与参考肽的富集分数(图1A为C57BL/6J小鼠;图1B为Balb/c小鼠。图2A和2B显示了表现最好的NHP脑(图2A)和脊髓(图2B)组织的富集分数。

[0183] 表1. 初次筛选的命中肽列表(如在NHP和小鼠筛选中鉴定的)。

缩略语	肽氨基酸序列	SEQ ID NO
EFS	EFSSNTVKLTS	38
SSN-L	GGSSNTVKLTSGHGG	39
SSN	SSNTVKLTSGH	40
SAN	SANFIKPTSY	41
VLT-L	GGVLTNIARGEYMRGG	46
IEI	IEINATRAGTNL	42
IEI-L	GGIEINATRAGTNLGG	43

[0185] 实例2. 二次验证

[0186] 我们通过为若干命中衣壳生成GFP报告载体继续跟进小鼠的初次筛选。以高剂量向C57BL/6J小鼠IV注射载体。2周后,我们对小鼠进行尸检,并收集脑切片的GFP图像(数据未显示)。在GFP研究中测试的所有命中载体都离开了肝,这一点从肝GFP染色中可以明显看出(数据未显示)。

[0187] 更高放大率的成像揭示出衣壳SSN和SAN明显在脑中定位,但限于内皮。RCA-外源凝集素是这些切片中脑内皮细胞的共染色标记物(数据未显示)。这些结果表明筛选中鉴定的AAV衣壳与BBB受体结合,但不会跨越BBB。突变体系列对Ly6a受体显示出一系列的亲和力。脑转导水平降低,这与观察到的肽与Ly6a受体更紧密的结合亲和力有关。观察到的紧密结合减少了转导,这是因为基因组可能卡在内皮中。

[0188] 这种内皮定位可能对某些疾病有用。另外,这种活性可以被优化以将这些从脑定位的载体转化为跨越BBB的载体。

[0189] 我们在条形码载体研究中证实了这些跨越BBB和脑定位的活性。简而言之,每个衣壳被用来单独生产含有GFP报告基因的载体,其中包含独特的DNA条形码(barcode)。将带条形码的衣壳制剂以相同比例混合,并注射到C57BL/6J或Balb/c小鼠中(图3A-3D)。在生存结束后,对小鼠组织进行NGS测序,以计算从组织中提取的载体基因组中每个条形码的丰度。结果证实了初次筛选中鉴定的所有命中衣壳的载体基因组的脑定位。在Balb/c小鼠中,二次验证筛选显示出对于在初次筛选中发现的所有命中序列的脑靶向性(图3A)。在C576BL/6小鼠中,二次验证筛选显示出对于在初次筛选中发现的所有命中序列的脑靶向性(图3B)。在Balb/c和C57BL/6小鼠中,相对于AVA9,对于所有命中序列的肝脱靶都与对脑血管的亲合力一致(图3C和3D)。

[0190] 实例3. 内皮靶向序列

[0191] 对于NHP二次验证,进行条形码研究。研究在注射有27个条形码载体的混合物的两个NHP中进行,所述混合物包含生存21天后的 4.5×10^{13} GC/kg的AAV9。进行全脑组织匀浆分析。虽然一些载体显示出载体生物分布的适度改善,但没有载体显示出全脑转导的改善。载体基因组在脑中的累积(DNA)与载体衍生的转录物的表达(mRNA)的相关性差(表2)。我们观察到内皮靶向载体的mRNA与DNA的相关性差。

[0192] 表2.

	DNA		mRNA	
	NHP1	NHP2	NHP1	NHP2
SSN	7.2	4.2	0.1	0.1
EFS	7.0	3.8	0.1	0.1
VLT-L	4.4	1.6	0.1	0
IEI-L	1.7	4.2	0.1	0.1
SSN-L	2.9	2.3	0.1	0.1
IEI	2.8	1.9	0.1	0.1
SAN	2.8	1.9	0.1	0.1
AAV9	1.0	1.0	1.0	1.0

[0194] 下表3显示了条形码研究中两个NHP的平均相对定位分数,同时相对于AAV9(等于1)归一化。与以脑为重点的文库设计相一致,大多数载体在非脑组织中的定位并不显著。这方面的一个例外是载体PMK,相对于AAV9,载体PMK显著重新定位到两个NHP的脾。

[0195] 表3.

AAV	眼睛	肾	肝	肺	胰腺	脾	睾丸	隔膜	Quadrcp	心脏
AAV9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SSN	1.5	0.3	0.3	0.3	0.2	0.0	0.4	1.1	1.0	0.6
SSN-L	1.3	0.2	0.5	0.2	0.2	0.1	0.6	0.9	0.9	0.6
EFS	2.6	0.4	0.4	0.7	0.2	0.1	0.9	1.6	1.4	1.0
VLT-L	0.9	0.2	0.5	0.3	0.2	0.1	0.4	0.6	0.6	0.4
IEI	0.6	0.2	0.1	0.5	0.1	0.0	0.2	0.3	0.5	0.2
IEI-L	1.0	0.2	0.1	0.4	0.1	0.6	0.8	0.5	0.7	0.5
SAN	1.0	0.3	0.3	0.5	0.1	0.1	0.3	0.7	0.8	0.4

[0197] 下表4显示脑内皮命中中具有共同的体外转导概况(如在293转导中所测量的)和共

同的生产概况。重要地,如NGS所测量的,具有内皮靶向活性的载体在AAV与质粒文库中的相对丰度均显示显著增加。

[0198] 表4.

载体名称	条形码DNA	条形码mRNA	插入物序列	293 转导	文库中质粒到AAV的转化	载体产量
SSN	5.7	0.1	SSNTVKLTSGH	20	54	0.9
EFS	5.4	0.1	EFSSNTVKLTS	17	35	0.8
VLT-L	3.0	0.0	GGVLTNIARGEYMRGG	21	16	0.6
IEI-L	3.0	0.1	GGIEINATRAGTNLGG	10	20	0.9
SSN-L	2.6	0.1	GGSSNTVKLTSGHGG	24	69	0.8
IEI	2.3	0.1	IEINATRAGTNL	10	14	0.9
SAN	2.3	0.1	SANFIKPTSY	43	26	0.7
AAV9	1.0	1.0		1		1.0

[0200] SAN肽的突变文库证实了“N_xTK”基序在脑靶向中的作用。在这项研究中,对SAN插入物进行每一个可能的单一氨基酸改变,在小鼠中注射优化的文库变体,并测量每个变体的生物分布f分数和产量。“N_xTK”基序是SAN插入物中脑生物分布的关键基序(表5和图5)。“N_xTK”基序控制SAN肽插入物中的质粒到AAV的转化(表6和图6)。此外,“N_xTK”基序将这些内皮载体在“N_xTK”类中的以下三个特性联系起来:内皮细胞转导、改善的293细胞转导和在文库生产期间传播的能力。图7A至7D显示“N_xTK”基序赋予跨越细胞系的广泛转导优势。与AAV9衣壳相比,在293细胞(图7A)、NIH3T3细胞(图7B)和HUH7细胞(图7C)中的相对转导水平得到改善。图7D显示在转导后第3天(3DPT)转导的早期显著改善,和转导后第7天(7DPT)大约10倍的改善。带有EFS和SAN肽插入物的AAV-GFP载体在转导原代猕猴气道上皮细胞时显示出改善的转导(图7E-7H)。

[0201] 表5. 脑生物分布(从上到下的平均值)

	S	A	N	F	I	K	P	T	S	Y
A	0.4	-0.4	-4.8	1.0	0.3	-3.8	-0.3	-0.5	-0.2	0.1
C	-6.0	0.0	-4.0	-1.5	-4.6	-3.1	-4.0	-4.0	-3.1	-1.1
D	0.7	-1.0	-3.7	0.5	-2.6	-8.3	-1.2	-0.8	-0.3	0.7
E	0.0	-0.8	-2.1	-0.5	-4.2	-5.3	-1.8	-0.4	0.2	0.9
F	-1.3	-2.4	-1.6	-0.4	-3.9	-1.9	-1.4	-0.9	-2.2	-0.2
G	0.1	-0.9	-2.7	1.3	-3.1	-4.6	0.5	0.4	-0.2	-0.7
H	-1.8	-1.0	-4.3	0.2	-2.4	-3.0	-1.1	-1.7	-1.9	-1.3
I	-1.0	-0.9	-5.2	0.8	-0.4	-3.5	-0.9	-1.2	-0.8	0.2

[0203]

K	-4.3	-2.8	-4.9	-0.4	-3.5	-4.0	-1.2	-2.1	-3.2	-2.7
L	-0.6	-0.8	-3.8	0.2	-3.7	-3.3	-1.1	-1.1	-0.5	0.4
M	-0.7	0.2	-4.2	0.6	-3.0	-4.1	-0.8	-0.8	-0.5	0.4
N	-0.3	-0.7	-0.4	1.0	-1.1	-4.4	-0.2	-0.5	-0.7	0.5
P	-1.1	-0.1	-4.5	-2.5	-2.3	-3.6	-0.4	-0.8	-0.2	-0.2
Q	0.0	-0.6	-3.4	1.2	-4.6	-3.6	-0.1	-0.3	-0.1	0.8
R	-2.3	-3.6	-3.9	-0.3	-1.1	-1.1	-2.3	-2.5	-3.1	-2.6
S	-0.4	-1.2	-3.6	0.0	-1.0	-4.6	-0.2	-0.5	-0.4	0.3
T	0.2	-0.9	-3.4	1.2	0.2	-4.2	-0.3	-0.4	-0.3	0.6
V	-1.0	-1.1	-3.9	1.0	-0.1	-4.7	-1.1	-0.6	-0.5	0.6
W	-2.0	-3.0	0.0	-2.0	-2.9	-4.8	0.0	-2.1	0.9	-1.3
Y	-1.6	-4.8	-1.9	0.0	-4.4	-4.1	-1.1	-1.4	-0.8	-0.4

[0204]

表6. 质粒到AAV产量转化

[0205]

	S	A	N	F	I	K	P	T	S	Y
A	0.6	0.4	-1.1	1.1	1.3	-1.4	1.3	-0.1	0.4	1.2
C	-3.1	11.0	-6.2	-2.6	-6.0	-6.8	-4.3	-5.4	-4.4	-1.7
D	-0.1	0.6	-0.7	0.3	-0.7	-1.1	1.0	0.2	0.8	0.9
E	0.3	-1.2	-1.0	0.7	-0.6	-0.6	0.9	0.1	0.8	0.8
F	-2.8	-5.5	-5.1	0.4	-3.7	-7.2	-5.5	-3.7	-3.2	0.2
G	0.4	0.7	-0.9	1.2	-0.4	-0.5	1.2	0.9	0.7	1.1
H	0.2	-0.9	-1.6	1.3	-0.6	-1.7	1.0	0.0	0.1	0.9
I	-0.9	-1.4	-2.9	1.1	0.4	-3.3	-1.9	-0.7	-0.9	1.0
K	-0.5	-0.5	-1.3	0.3	-1.6	-3	-1.2	-2.5	-1.8	-1.1
L	-0.5	-1.1	-2.6	1.1	-1.1	-4.5	-1.7	-1.2	-1.2	0.6
M	-0.6	-1.5	-2.4	1.3	-0.6	-3.0	-0.4	-0.9	-0.8	1.0
N	1.0	0.5	0.4	1.5	-0.1	-1.0	1.5	0.6	0.7	1.2
P	0.7	-0.2	-1.1	-0.4	-0.4	-0.9	0.4	-0.4	0.4	0.7
Q	0.8	1.2	-0.6	1.6	-0.1	-1.2	1.5	0.0	0.4	1.0
R	-2.1	-2.9	-4.6	-1.0	-5.2	-1.3	-3.2	-4.2	-3.9	-1.3
S	0.4	0.5	-1.5	1.0	1.3	-0.9	1.3	0.3	0.4	1.2
T	0.5	0.4	-1.0	1.3	1.1	-1.0	1.3	0.4	0.5	1.1
V	0.3	-0.3	-2.3	1.1	0.8	-3.7	-0.1	-0.2	0.2	1.1
W	-4.4	-5.6	-11.0	-2.0	-5.5	-5.8	-11.6	-4.9	-7.3	-1.1
Y	-1.8	-2.9	-6.3	0.2	-3.8	-6.8	-4.3	-2.9	-3.1	0.4

[0206]

综上所述,如表7所列,所选氨基酸序列都含有功能基序N-x-(T/I/V/A)-(K/R)(SEQ ID NO:47)。除了表7中所示的所选序列外,在筛选期间还鉴定了其它序列。我们有数据支持,这些插入物序列的许多替代也支持或甚至改善内皮靶向活性。此外,我们已经从大型随机插入物文库中发现了大约数千个符合这一基序的序列——这些序列可能都享有改善的转导特性。

[0207]

表7. 所选内皮靶向序列

[0208]	载体名称	插入物序列	SEQ ID NO
	SSN	SSNTVKLTSGH	40
	EFS	EFSSNTVKLTS	38
	VLT-L	GGVLTNIARGEYMRGG	46
	IEI-L	GGIEINATRAGTNLGG	43
	SSN-L	GGSSNTVKLTSGHGG	39
	IEI	IEINATRAGTNL	42
	SAN	SANFIKPTSY	41

[0209] 我们在NHP中完成了初次筛选命中的条形码评估。脑定位是这一文库中最突出的特征,而除了可能的AAV9-PMK的脾脏靶向外,周边组织的靶向没有明显的增强。我们定义了所有具有脑内皮靶向活性的肽插入物共同的序列基序。我们证实了这个基序在脑内皮靶向中的活性,以及在赋予广泛的体外转导优势中的活性。单一的序列基序“NxTK”定义了来自4个不相关来源的共享以下三个特性的插入物:

[0210] (1)显著改善的293s和其它细胞系的体外转导;

[0211] (2)文库生产期间的寄生性扩张——“扩散”表型;以及

[0212] (3)在小鼠和NHP体内的内皮生物分布。

[0213] 在SAN和EFS载体插入物的系统突变筛选中映射“NxTK”基序。在系统突变筛选中,显示“NxTK”基序是脑内皮生物分布的关键以及文库生产中丰度的关键。质粒到AAV转化,是指文库生产期间的衣壳产量,由以下2个因素控制:寄生性扩散基序的存在(主要因素)和内在的衣壳产量(次要因素)。扩散基序之一被鉴定为“NxTK”,并可能与293个细胞表面受体相互作用并赋予转导优势。293中的转导优势可能导致载体基因组(Cap)在生产期间传播到邻近的细胞,因为文库生产是在限制Cap的情况下进行的,所以大多数细胞最初除了Cap基因外什么都有。次要因素只有在数字过滤掉具有寄生性扩散基序的载体后才会显现出来。

[0214] 实例4.用于气道递送的AAV衣壳开发的工程化策略

[0215] 目前的AAV载体对鼻道和上呼吸道的细胞的转导差,限制了预防AAV策略对防御上呼吸道传染病如流感或COVID-19的功效。这一项目旨在工程化AAV衣壳以改善对这些组织的转导,特别追求定向进化和受体靶向策略。追求AAV气道特异性衣壳选择的工程化策略包括以下步骤:包含但不限于生成带有插入物的构建体多样化文库(10^3 至 10^9 个初始多样性),在原代灵长类或人类气道细胞中筛选和选择,鉴定所选构建体的基因组特征(DNA或RNA),进行额外的筛选以集中于命中,在NHP中验证改善的衣壳的命中。用于气道递送的衣壳多样性的来源包括追求两种方法:无偏和有偏方法。在无偏方法中,将随机肽插入到AAV衣壳表面以生成大于 10^7 个变体多样性的大型文库。在有偏方法中,将具有已知或疑似气道细胞结合活性的肽插入到AAV衣壳以生成具有大约 10^3 个变体多样性的小型文库。此类已知或疑似气道细胞结合肽的来源为:已公开的噬菌体展示结果、来自病毒受体结合结构域的肽、气道受体的已知配体以及以前体外文库筛选中生成的肽。为了在体外筛选生成的衣壳文库,使用利用气液界面(ALI)培养物中的原代气道细胞的测定。使用的细胞来源于人类和猕猴,并且包括鼻、气管和支气管细胞来源的细胞。用GFP载体在猕猴原代气道上皮细胞培养物中进行的初步转导测试显示出EFS和SAN插入肽到AAV衣壳的转导的显著早期改善(图7D和7E-7H)。图7E显示了用载剂处理的对照样品(即无载体)中猕猴原代气道上皮细胞的显微镜分

析。图7F显示了用AAV9-GFP载体转导后的猕猴原代气道上皮细胞的显微镜分析。图7G显示了用包括EFS肽插入物的AAV9-GFP载体转导后的猕猴原代气道上皮细胞的显微镜分析。图7H显示了用包括SAN肽插入物的AAV9-GFP转导后的猕猴原代气道上皮细胞的显微镜分析。用GFP载体在培养的人类细胞中进行的初步转导测试显示出总体较低的转导,其中在第7天,培养的人类细胞的mRNA拷贝数与以微克计的总mRNA的比率为 1×10^4 (图8),并且培养的猕猴原代气道上皮细胞的所述比率为 1×10^6 (图7D)。与AAV9转导相比,SAN基序在第7天的转导中显示出优势(图8)。EFS基序在支气管和气管培养的人类细胞中显示出较差的转导(图8)。

[0216] 此外,我们在其它项目中通过体外选择方案开发了许多插入物序列,所述序列赋予了显著改善的细胞结合和转导活性。在ALI培养物上的条形码池中测试生成的AAV-插入物载体。结果显示,所有体外选择的AAV9-插入物载体都优于AAV9载体(结果未显示)。具体地,Spr3L(NxTK)的AAV9-插入物载体在比较中是最好的,后者的转导是AAV9衣壳的大约50倍。

[0217] (序列表自由文本)

[0218] 针对含有在

[0219] 数字标识符<223>下的自由文本的序列提供以下信息。

SEQ ID NO:	<223>下的自由文本
1	<220> <223> AAV2/9 n.588.EFS 核酸序列表达盒

[0220]

[0221]

	<p><220> <221> misc_feature <222> (1)..(36) <223>截短启动子</p> <p><220> <221> promoter <222> (1)..(7) <223> p5 启动子</p> <p><220> <221> CDS <222> (37)..(1899) <223> AAV2-Rep</p> <p><220> <221> CDS <222> (1919)..(4162) <223> AAV9 Cap</p> <p><220> <221> misc_feature <222> (3683)..(3715) <223> EFS</p> <p><220> <221> misc_feature <222> (4253)..(4383) <223> p5 启动子</p> <p><220> <221> misc_feature <222> (4511)..(4725) <223> LacZ 启动子</p>
2	<p><220> <223>合成构建体</p>
3	<p><220> <223>合成构建体</p>
4	<p><220> <223>AAV2/9 n.588.IEI 核酸序列表达盒</p> <p><220> <221> misc_feature <222> (1)..(36) <223>截短启动子</p> <p><220> <221> promoter <222> (1)..(7)</p>

[0222]

	<p><223>p5 启动子</p> <p><220> <221>CDS <222> (37)..(1899) <223>AAV2-Rep</p> <p><220> <221>CDS <222> (1919)..(4165) <223>AAV9 Cap</p> <p><220> <221> misc_feature <222> (3683)..(3718) <223>IEI</p> <p><220> <221> misc_feature <222> (4256)..(4386) <223>p5 启动子</p> <p><220> <221> misc_feature <222> (4514)..(4728) <223>LacZ 启动子</p>
5	<p><220> <223>合成构建体</p>
6	<p><220> <223>合成构建体</p>
7	<p><220> <223>AAV2/9 n.588.IEI-L 核酸序列表达盒</p> <p><220> <221> misc_feature <222> (1)..(36) <223>截短启动子</p> <p><220> <221> promoter <222> (1)..(7) <223>p5 启动子</p> <p><220> <221>CDS <222> (37)..(1899) <223>AAV2 Rep</p> <p><220></p>

[0223]

	<221>CDS <222> (1919)..(4177) <223>AAV9 Cap <220> <221> misc_feature <222> (3683)..(3739) <223>IEI-L <220> <221> misc_feature <222> (4268)..(4398) <223>p5 启动子 <220> <221> misc_feature <222> (4526)..(4740) <223> LacZ 启动子
8	<220> <223>合成构建体
9	<220> <223>合成构建体
10	<210>10 <211>4722 <212>DNA <213>人工序列 <220> <223>AAV2/9 n.588.SAN 核酸序列表达盒 <220> <221> misc_feature <222> (1)..(36) <223>截短启动子 <220> <221> promoter <222> (1)..(7) <223>p5 启动子 <220> <221>CDS <222> (37)..(1899) <223>AAV2 Rep <220> <221>CDS <222> (1919)..(4159) <223>AAV 9 Cap

[0224]

	<p><220> <221> misc_feature <222> (3683)..(3712) <223> SAN</p> <p><220> <221> misc_feature <222> (4250)..(4380) <223> p5 启动子</p> <p><220> <221> misc_feature <222> (4508)..(4722) <223> LacZ 启动子</p>
11	<p><220> <223> 合成构建体</p>
12	<p><220> <223> 合成构建体</p>
13	<p><220> <223> AAV2/9 n.588.SSN 核酸序列表达盒</p> <p><220> <221> misc_feature <222> (1)..(36) <223> 截短启动子</p> <p><220> <221> promoter <222> (1)..(7) <223> p5 启动子</p> <p><220> <221> CDS <222> (37)..(1899) <223> AAV2 Rep</p> <p><220> <221> CDS <222> (1919)..(4162) <223> AAV9 Cap</p> <p><220> <221> misc_feature <222> (3683)..(3715) <223> SSN</p> <p><220> <221> misc_feature</p>

[0225]

	<222> (4253)..(4383) <223> p5 启动子 <220> <221> misc_feature <222> (4511)..(4725) <223> LacZ 启动子
14	<220> <223> 合成构建体
15	<220> <223> 合成构建体
16	<220> <223> AAV2/9 n.588.SSN-L 核酸序列表达盒 <220> <221> misc_feature <222> (1)..(36) <223> 截短启动子 <220> <221> promoter <222> (1)..(7) <223> p5 启动子 <220> <221> CDS <222> (37)..(1899) <223> AAV2 Rep <220> <221> CDS <222> (1919)..(4174) <223> AAV9 Cap <220> <221> misc_feature <222> (3683)..(3727) <223> SSN-L <220> <221> misc_feature <222> (4253)..(4737) <223> LacZ 启动子 <220> <221> misc_feature <222> (4265)..(4395) <223> p5 启动子

[0226]

17	<220> <223>合成构建体
18	<220> <223>合成构建体
19	<220> <223>AAV2/9 n.588.VLT-L 核酸序列表达盒 <220> <221> misc_feature <222> (1)..(36) <223>截短启动子 <220> <221> promoter <222> (1)..(7) <223> p5 启动子 <220> <221> CDS <222> (37)..(1899) <223> AAV2 Rep <220> <221> CDS <222> (1919)..(4177) <223> AAV9 Cap <220> <221> misc_feature <222> (3683)..(3730) <223> VLT-L <220> <221> misc_feature <222> (4268)..(4398) <223> p5 启动子 <220> <221> misc_feature <222> (4526)..(4740) <223> LacZ 启动子
20	<220> <223> 合成构建体
21	<220> <223> 合成构建体
22	<220> <223> AAV2 Rep 核酸序列
23	<220>

[0227]

	<223> AAV2 Rep 氨基酸序列
24	<220> <223> AAV9 Cap n.588.EFS 核酸序列 <220> <221> misc_feature <222> (1765)..(1797) <223> EFS
25	<220> <223> AAV9 Cap n.588.EFS 氨基酸序列 <220> <221> misc_feature <222> (499)..(599) <223> EFS
25	<220> <223> AAV9 Cap n588.IEI 核酸序列 <220> <221> misc_feature <222> (1765)..(1800) <223> IEI
27	<220> <223> AAV9 Cap n588.IEI 氨基酸序列 <220> <221> misc_feature <222> (499)..(600) <223> IEI
28	<220> <223> AAV9 Cap n.588.IEI-L 核酸序列 <220> <221> misc_feature <222> (1765)..(1812) <223> IEI-L
29	<220> <223> AAV9 Cap n.588.IEI-L 氨基酸序列 <220> <221> misc_feature <222> (499)..(604) <223> IEI-L

[0228]

30	<p><220> <223> AAV9 Cap n.588.SAN 核酸序列</p> <p><220> <221> misc_feature <222> (1765)..(1794) <223> SAN</p>
31	<p><220> <223> AAV9 Cap n.588.SAN 氨基酸序列</p> <p><220> <221> misc_feature <222> (499)..(598) <223> SAN</p>
32	<p><220> <223> AAV9 Cap n.588.SSN 核酸序列</p> <p><220> <221> misc_feature <222> (1765)..(1797) <223> SSN</p>
33	<p><220> <223> AAV9 Cap n.588.SSN 氨基酸序列</p> <p><220> <221> misc_feature <222> (499)..(599) <223> SSN</p>
34	<p><220> <223> AAV9 Cap n.588.SSN-L 核酸序列</p> <p><220> <221> misc_feature <222> (1765)..(1809) <223> SSN-L</p>
35	<p><220> <223> AAV9 Cap n.588.SSN-L 氨基酸序列</p> <p><220> <221> misc_feature <222> (499)..(603) <223> SSN-L</p>
36	<220>

[0229]

	<p><223> AAV9 Cap n588.VLT-L 核酸序列</p> <p><220> <221> misc_feature <222> (1765)..(1812) <223> VLT-L</p>
37	<p><220> <223> AAV9 Cap n588.VLT-L 氨基酸序列</p> <p><220> <221> misc_feature <222> (499)..(604) <223> VLT-L</p>
38	<p><220> <223> EFS 肽序列</p>
39	<p><220> <223> SSN-L 肽序列</p>
40	<p><220> <223> SSN 肽序列</p>
41	<p><220> <223> SAN 肽序列</p>
42	<p><220> <223> IEI 肽序列</p>
43	<p><220> <223> IEI-L 肽序列</p>
44	<p><220> <223> AAV9 衣壳</p>
45	<p><220> <223> AAVhu68 衣壳</p>
46	<p><220> <223> VLT-L 肽序列</p>
47	<p><220> <223> N-X-(T/I/V/A)-(K/R)基序</p> <p><220> <221> misc_feature <222> (2)..(2) <223>任意氨基酸</p> <p><220> <221> misc_feature <222> (3)..(3) <223> Xaa 选自苏氨酸 (T)、异亮氨酸 (I)、缬氨酸 (V) 或丙氨酸 (A)</p>

	<220> <221> misc_feature <222> (4)..(4) <223>Xaa 选自赖氨酸 (K) 或精氨酸 (R)
48	<220> <223>AAV2 变体肽 NDVRAVS
49	<220> <223>PHP.B 肽插入物
50	<220> <223>核酸序列 EFS
[0230] 51	<220> <223>核酸序列 IEI
52	<220> <223>核酸序列 IEI-L
53	<220> <223>核酸序列 SAN
54	<220> <223>核酸序列 SSN
55	<220> <223>核酸序列 SSN-L
56	<220> <223>核酸序列 VLT-L

[0231] 本说明书中引用的所有文献均通过引用并入本文。2020年12月1日提交的美国临时申请号63/119,863通过引用以其整体并入本文。随之提交的名为“20-9409PCT_ST25”的序列表以及其中的序列和文本均通过引用并入。虽然已经参考特定实施例描述了本发明，但是应理解，在不脱离本发明的精神的情况下可以进行修改。这些修改旨在落入所附权利要求的范围内。

序列表

<110> 宾夕法尼亚大学董事会 (The Trustees of the University of Pennsylvania)

<120> 具有内皮组织特异性靶向基序的新型AAV衣壳和含有其的组合物

<130> UPN-20-9409.PCT

<150> US 63/119,863

<151> 2020-12-01

<160> 56

<170> PatentIn版本3.5

<210> 1

<211> 4725

<212> DNA

<213> 人工序列

<220>

<223> AAV2/9 n.588.EFS核酸序列表达盒

<220>

<221> misc_feature

<222> (1) .. (36)

<223> 截短启动子

<220>

<221> promoter

<222> (1) .. (7)

<223> p5启动子

<220>

<221> CDS

<222> (37) .. (1899)

<223> AAV2-Rep

<220>

<221> CDS

<222> (1919) .. (4162)

<223> AAV9 Cap

<220>

<221> misc_feature

<222> (3683) .. (3715)

<223> EFS

<220>

<221> misc_feature

<222> (4253) .. (4383)

<223> p5启动子	
<220>	
<221> misc_feature	
<222> (4511)..(4725)	
<223> LacZ启动子	
<400> 1	
ccattttgaa gcgggagggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag	54
	Met Pro Gly Phe Tyr Glu
	1 5
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att	102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile	
	10 15 20
tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg	150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro	
	25 30 35
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc	198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr	
	40 45 50
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg	246
Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val	
	55 60 65 70
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag	294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu	
	75 80 85
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc	342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser	
	90 95 100
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag	390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln	
	105 110 115
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc	438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val	
	120 125 130
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag	486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu	
	135 140 145 150
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag	534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln	
	155 160 165

tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc	582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu	
170 175 180	
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tgc cag	630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln	
185 190 195	
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg	678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200 205 210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215 220 225 230	
ctc gtg gac aag ggg att acc tgc gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
235 240 245	
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tgc cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
250 255 260	
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys	
265 270 275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc	918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser	
280 285 290	
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa	966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln	
295 300 305 310	
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag	1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys	
315 320 325	
agg aac acc atc tgg ctg ttt ggg cct gca act acc ggg aag acc aac	1062
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn	
330 335 340	
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac	1110
Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn	
345 350 355	
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg	1158
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val	
360 365 370	

gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585 590 595	
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600 605 610	
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat	1930
Asp Asp Cys Ile Phe Glu Gln Met Ala Ala Asp	
615 620 625	
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc	1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg	
630 635 640	
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa	2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln	
645 650 655	
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac	2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr	
660 665 670	
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca	2122
Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala	
675 680 685	
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag	2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys	
690 695 700 705	
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc	2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe	
710 715 720	
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga	2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg	
725 730 735	
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt	2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val	
740 745 750	
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag	2362
Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln	
755 760 765	
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca	2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala	
770 775 780 785	

cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag	2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu	
790 795 800	
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca	2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser	
805 810 815	
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca	2554
Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala	
820 825 830	
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg	2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp	
835 840 845	
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc	2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr	
850 855 860 865	
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc	2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile	
870 875 880	
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc	2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly	
885 890 895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac	2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
900 905 910	
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc	2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe	
915 920 925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag	2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
930 935 940 945	
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc	2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser	
950 955 960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc	2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965 970 975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980 985 990	

atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995 1000 1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010 1015 1020	
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu	
1025 1030 1035	
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac	3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp	
1040 1045 1050	
cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
1055 1060 1065	
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc	3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe	
1070 1075 1080	
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac	3352
Ser Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr	
1085 1090 1095	
ata cct gga ccc agc tac cga caa caa cgt gtc tca acc act gtg	3397
Ile Pro Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val	
1100 1105 1110	
act caa aac aac aac agc gaa ttt gct tgg cct gga gct tct tct	3442
Thr Gln Asn Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser	
1115 1120 1125	
tgg gct ctc aat gga cgt aat agc ttg atg aat cct gga cct gct	3487
Trp Ala Leu Asn Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala	
1130 1135 1140	
atg gcc agc cac aaa gaa gga gag gac cgt ttc ttt cct ttg tct	3532
Met Ala Ser His Lys Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser	
1145 1150 1155	
gga tct tta att ttt ggc aaa caa gga act gga aga gac aac gtg	3577
Gly Ser Leu Ile Phe Gly Lys Gln Gly Thr Gly Arg Asp Asn Val	
1160 1165 1170	
gat gcg gac aaa gtc atg ata acc aac gaa gaa gaa att aaa act	3622
Asp Ala Asp Lys Val Met Ile Thr Asn Glu Glu Glu Ile Lys Thr	
1175 1180 1185	

act aac ccg gta gca acg gag tcc tat gga caa gtg gcc aca aac	3667
Thr Asn Pro Val Ala Thr Glu Ser Tyr Gly Gln Val Ala Thr Asn	
1190 1195 1200	
cac cag agt gcc caa gag ttc agc agc aac acc gtg aag ctg acc	3712
His Gln Ser Ala Gln Glu Phe Ser Ser Asn Thr Val Lys Leu Thr	
1205 1210 1215	
agc gca cag gcg cag acc ggc tgg gtt caa aac caa gga ata ctt	3757
Ser Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile Leu	
1220 1225 1230	
ccg ggt atg gtt tgg cag gac aga gat gtg tac ctg caa gga ccc	3802
Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro	
1235 1240 1245	
att tgg gcc aaa att cct cac acg gac ggc aac ttt cac cct tct	3847
Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser	
1250 1255 1260	
ccg ctg atg gga ggg ttt gga atg aag cac ccg cct cct cag atc	3892
Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile	
1265 1270 1275	
ctc atc aaa aac aca cct gta cct gcg gat cct cca acg gcc ttc	3937
Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe	
1280 1285 1290	
aac aag gac aag ctg aac tct ttc atc acc cag tat tct act ggc	3982
Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly	
1295 1300 1305	
caa gtc agc gtg gag atc gag tgg gag ctg cag aag gaa aac agc	4027
Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser	
1310 1315 1320	
aag cgc tgg aac ccg gag atc cag tac act tcc aac tat tac aag	4072
Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys	
1325 1330 1335	
tct aat aat gtt gaa ttt gct gtt aat act gaa ggt gta tat agt	4117
Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser	
1340 1345 1350	
gaa ccc cgc ccc att ggc acc aga tac ctg act cgt aat ctg taa	4162
Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu	
1355 1360 1365	
ttgcttgta atcaataaac cgtttaattc gtttcagttg aactttggtc tctgcgaagg	4222
gcgaattcgt ttaaacctgc aggactagag gtctgtatt agaggtcacg tgagtgtttt	4282
gcgacatttt gcgacacat gtggtcacgc tgggtattta agcccgagtg agcacgcagg	4342

gtctccattt tgaagcggga ggtttgaacg cgcagccgcc aagccgaatt ctgcagatat 4402
ccatcacact ggcggccgct cgactagagc ggccgccacc gcggtggagc tccagctttt 4462
gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag ctgtttcctg 4522
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg 4582
aagcctgggg tgcctaata gtagactaac tcacattaat tgcgttgccg tcaactgccc 4642
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 4702
gaggcggttt gcgtattggg cgc 4725

<210> 2

<211> 621

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 2

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

195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515 520 525

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530 535 540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545 550 555 560

Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565 570 575

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580 585 590

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595 600 605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610 615 620

<210> 3
<211> 747
<212> PRT
<213> 人工序列
<220>
<223> 合成构建体
<400> 3

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

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

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
 260 265 270
 Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
 290 295 300
 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
 305 310 315 320
 Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser

450	455	460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro		
465	470	475
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn		
	485	490
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn		
	500	505
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys		
	515	520
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly		
	530	535
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile		
545	550	555
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser		
	565	570
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Glu Phe Ser Ser		
	580	585
Asn Thr Val Lys Leu Thr Ser Ala Gln Ala Gln Thr Gly Trp Val Gln		
	595	600
Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr		
	610	615
Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe		
625	630	635
His Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro		
	645	650
Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala		
	660	665
Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly		
	675	680
Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys		
	690	695
Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn		
705	710	715
Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg		
	725	730
Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu		
	740	745

<210> 4

<211> 4728

- <212> DNA
 <213> 人工序列
 <220>
 <223> AAV2/9 n.588.IEI核酸序列表达盒
 <220>
 <221> misc_feature
 <222> (1) .. (36)
 <223> 截短启动子
 <220>
 <221> promoter
 <222> (1) .. (7)
 <223> p5启动子
 <220>
 <221> CDS
 <222> (37) .. (1899)
 <223> AAV2-Rep
 <220>
 <221> CDS
 <222> (1919) .. (4165)
 <223> AAV9 Cap
 <220>
 <221> misc_feature
 <222> (3683) .. (3718)
 <223> IEI
 <220>
 <221> misc_feature
 <222> (4256) .. (4386)
 <223> p5启动子
 <220>
 <221> misc_feature
 <222> (4514) .. (4728)
 <223> LacZ启动子
 <400> 4

```

ccattttgaa gcgggagggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag      54
                                     Met Pro Gly Phe Tyr Glu
                                     1           5
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att      102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile
                10                15                20
  
```

tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg	150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro	
25 30 35	
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc	198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr	
40 45 50	
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg	246
Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val	
55 60 65 70	
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag	294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu	
75 80 85	
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc	342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser	
90 95 100	
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag	390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln	
105 110 115	
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc	438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val	
120 125 130	
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag	486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu	
135 140 145 150	
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag	534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln	
155 160 165	
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc	582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu	
170 175 180	
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tcg cag	630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln	
185 190 195	
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg	678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200 205 210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215 220 225 230	

ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
235 240 245	
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
250 255 260	
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys	
265 270 275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc	918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser	
280 285 290	
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa	966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln	
295 300 305 310	
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag	1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys	
315 320 325	
agg aac acc atc tgg ctg ttt ggg cct gca act acc ggg aag acc aac	1062
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn	
330 335 340	
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac	1110
Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn	
345 350 355	
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg	1158
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val	
360 365 370	
atc tgg tgg gag gag ggg aag atg acc gcc aag gtc gtg gag tcg gcc	1206
Ile Trp Trp Glu Glu Gly Lys Met Thr Ala Lys Val Val Glu Ser Ala	
375 380 385 390	
aaa gcc att ctc gga gga agc aag gtg gcg gtg gac cag aaa tgc aag	1254
Lys Ala Ile Leu Gly Gly Ser Lys Val Arg Val Asp Gln Lys Cys Lys	
395 400 405	
tcc tcg gcc cag ata gac ccg act ccc gtg atc gtc acc tcc aac acc	1302
Ser Ser Ala Gln Ile Asp Pro Thr Pro Val Ile Val Thr Ser Asn Thr	
410 415 420	
aac atg tgc gcc gtg att gac ggg aac tca acg acc ttc gaa cac cag	1350
Asn Met Cys Ala Val Ile Asp Gly Asn Ser Thr Thr Phe Glu His Gln	
425 430 435	

cag ccg ttg caa gac cgg atg ttc aaa ttt gaa ctc acc cgc cgt ctg	1398
Gln Pro Leu Gln Asp Arg Met Phe Lys Phe Glu Leu Thr Arg Arg Leu	
440 445 450	
gat cat gac ttt ggg aag gtc acc aag cag gaa gtc aaa gac ttt ttc	1446
Asp His Asp Phe Gly Lys Val Thr Lys Gln Glu Val Lys Asp Phe Phe	
455 460 465 470	
cgg tgg gca aag gat cac gtg gtt gag gtg gag cat gaa ttc tac gtc	1494
Arg Trp Ala Lys Asp His Val Val Glu Val Glu His Glu Phe Tyr Val	
475 480 485	
aaa aag ggt gga gcc aag aaa aga ccc gcc ccc agt gac gca gat ata	1542
Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile	
490 495 500	
agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca	1590
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser	
505 510 515	
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt	1638
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys	
520 525 530	
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc	1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys	
535 540 545 550	
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa	1734
Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys	
555 560 565	
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc	1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val	
570 575 580	
gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585 590 595	
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600 605 610	
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat	1930
Asp Asp Cys Ile Phe Glu Gln Met Ala Ala Asp	
615 620 625	
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc	1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg	
630 635 640	

gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln 645 650 655	2026
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr 660 665 670	2074
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala 675 680 685	2122
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys 690 695 700 705	2170
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe 710 715 720	2218
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg 725 730 735	2266
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val 740 745 750	2314
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln 755 760 765	2362
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala 770 775 780 785	2410
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu 790 795 800	2458
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser 805 810 815	2506
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala 820 825 830	2554
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp 835 840 845	2602

cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc	2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr	
850 855 860 865	
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc	2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile	
870 875 880	
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc	2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly	
885 890 895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac	2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
900 905 910	
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc	2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe	
915 920 925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag	2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
930 935 940 945	
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc	2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser	
950 955 960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc	2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965 970 975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980 985 990	
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995 1000 1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010 1015 1020	
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu	
1025 1030 1035	
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac	3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp	
1040 1045 1050	

cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
1055 1060 1065	
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc	3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe	
1070 1075 1080	
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac	3352
Ser Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr	
1085 1090 1095	
ata cct gga ccc agc tac cga caa caa cgt gtc tca acc act gtg	3397
Ile Pro Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val	
1100 1105 1110	
act caa aac aac aac agc gaa ttt gct tgg cct gga gct tct tct	3442
Thr Gln Asn Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser	
1115 1120 1125	
tgg gct ctc aat gga cgt aat agc ttg atg aat cct gga cct gct	3487
Trp Ala Leu Asn Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala	
1130 1135 1140	
atg gcc agc cac aaa gaa gga gag gac cgt ttc ttt cct ttg tct	3532
Met Ala Ser His Lys Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser	
1145 1150 1155	
gga tct tta att ttt ggc aaa caa gga act gga aga gac aac gtg	3577
Gly Ser Leu Ile Phe Gly Lys Gln Gly Thr Gly Arg Asp Asn Val	
1160 1165 1170	
gat gcg gac aaa gtc atg ata acc aac gaa gaa gaa att aaa act	3622
Asp Ala Asp Lys Val Met Ile Thr Asn Glu Glu Glu Ile Lys Thr	
1175 1180 1185	
act aac ccg gta gca acg gag tcc tat gga caa gtg gcc aca aac	3667
Thr Asn Pro Val Ala Thr Glu Ser Tyr Gly Gln Val Ala Thr Asn	
1190 1195 1200	
cac cag agt gcc caa atc gag atc aac gct acc aga gct gga acc	3712
His Gln Ser Ala Gln Ile Glu Ile Asn Ala Thr Arg Ala Gly Thr	
1205 1210 1215	
aac ctg gca cag gcg cag acc ggc tgg gtt caa aac caa gga ata	3757
Asn Leu Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile	
1220 1225 1230	
ctt ccg ggt atg gtt tgg cag gac aga gat gtg tac ctg caa gga	3802
Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly	
1235 1240 1245	

ccc att tgg gcc aaa att cct cac acg gac ggc aac ttt cac cct	3847
Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro	
1250 1255 1260	
tct ccg ctg atg gga ggg ttt gga atg aag cac ccg cct cct cag	3892
Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln	
1265 1270 1275	
atc ctc atc aaa aac aca cct gta cct gcg gat cct cca acg gcc	3937
Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala	
1280 1285 1290	
ttc aac aag gac aag ctg aac tct ttc atc acc cag tat tct act	3982
Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr	
1295 1300 1305	
ggc caa gtc agc gtg gag atc gag tgg gag ctg cag aag gaa aac	4027
Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn	
1310 1315 1320	
agc aag cgc tgg aac ccg gag atc cag tac act tcc aac tat tac	4072
Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr	
1325 1330 1335	
aag tct aat aat gtt gaa ttt gct gtt aat act gaa ggt gta tat	4117
Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr	
1340 1345 1350	
agt gaa ccc cgc ccc att ggc acc aga tac ctg act cgt aat ctg	4162
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu	
1355 1360 1365	
taa ttgcttgta atcaataaac cgtttaattc gtttcagttg aactttggtc	4215
tctgcgaagg gcgaattcgt ttaaactgc aggactagag gtcctgtatt agaggtcacg	4275
tgagtgtttt gcgacathtt gcgacacat gtggtcacgc tgggtattta agcccagtg	4335
agcacgcagg gtctccattt tgaagcggga gtttgaacg cgcagccgcc aagccgaatt	4395
ctgcagatat ccatcacact ggcggccgct ctagtagagc ggccgccacc gcggtggagc	4455
tccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag	4515
ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc	4575
ataaagtgta aagcctgggg tgccaatga gtgagctaac tcacattaat tgcgttgccg	4635
tcaactgccc ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcgccaa	4695
cgcgcgggga gaggcggttt gcgtattggg cgc	4728
<210> 5	
<211> 621	
<212> PRT	
<213> 人工序列	
<220>	

<223> 合成构建体

<400> 5

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
 1 5 10 15
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
 20 25 30
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
 35 40 45
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
 50 55 60
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65 70 75 80
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
 85 90 95
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
 100 105 110
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
 115 120 125
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
 130 135 140
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
 145 150 155 160
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
 165 170 175
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
 195 200 205
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 225 230 235 240
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
 275 280 285
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu

290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
	325	330
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
	340	345
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
	355	360
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
	405	410
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
	420	425
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
	435	440
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
	485	490
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
	500	505
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
	515	520
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
	565	570
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
	580	585
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
 610 615 620
 <210> 6
 <211> 748
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 6
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240

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

545	550	555	560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser			
	565	570	575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ile Glu Ile Asn			
	580	585	590
Ala Thr Arg Ala Gly Thr Asn Leu Ala Gln Ala Gln Thr Gly Trp Val			
	595	600	605
Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val			
	610	615	620
Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn			
625	630	635	640
Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro			
	645	650	655
Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr			
	660	665	670
Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr			
	675	680	685
Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser			
	690	695	700
Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser			
705	710	715	720
Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro			
	725	730	735
Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	

<210> 7

<211> 4740

<212> DNA

<213> 人工序列

<220>

<223> AAV2/9 n.588.IEI-L核酸序列表达盒

<220>

<221> misc_feature

<222> (1) .. (36)

<223> 截短启动子

<220>

<221> promoter

<222> (1) .. (7)

<223> p5启动子

agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc	342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser	
90 95 100	
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag	390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln	
105 110 115	
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc	438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val	
120 125 130	
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag	486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu	
135 140 145 150	
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag	534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln	
155 160 165	
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc	582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu	
170 175 180	
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tcg cag	630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln	
185 190 195	
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg	678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200 205 210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215 220 225 230	
ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
235 240 245	
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
250 255 260	
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys	
265 270 275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc	918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser	
280 285 290	

agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca	1590
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser	
505 510 515	
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt	1638
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys	
520 525 530	
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc	1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys	
535 540 545 550	
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa	1734
Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys	
555 560 565	
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc	1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val	
570 575 580	
gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585 590 595	
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600 605 610	
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat	1930
Asp Asp Cys Ile Phe Glu Gln Met Ala Ala Asp	
615 620 625	
ggg tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc	1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg	
630 635 640	
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa	2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln	
645 650 655	
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac	2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr	
660 665 670	
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca	2122
Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala	
675 680 685	
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag	2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys	
690 695 700 705	

gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc	2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe	
710 715 720	
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga	2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg	
725 730 735	
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt	2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val	
740 745 750	
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag	2362
Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln	
755 760 765	
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca	2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala	
770 775 780 785	
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag	2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu	
790 795 800	
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca	2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser	
805 810 815	
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca	2554
Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala	
820 825 830	
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg	2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp	
835 840 845	
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc	2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr	
850 855 860 865	
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc	2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile	
870 875 880	
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc	2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly	
885 890 895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac	2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
900 905 910	

ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc	2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe	
915 920 925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag	2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
930 935 940 945	
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc	2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser	
950 955 960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc	2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965 970 975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980 985 990	
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995 1000 1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010 1015 1020	
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu	
1025 1030 1035	
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac	3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp	
1040 1045 1050	
cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
1055 1060 1065	
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc	3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe	
1070 1075 1080	
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac	3352
Ser Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr	
1085 1090 1095	
ata cct gga ccc agc tac cga caa caa cgt gtc tca acc act gtg	3397
Ile Pro Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val	
1100 1105 1110	

act	caa aac aac aac agc	gaa ttt gct tgg cct	gga gct tct tct	3442
Thr	Gln Asn Asn Asn Ser	Glu Phe Ala Trp Pro	Gly Ala Ser Ser	
1115		1120	1125	
tgg	gct ctc aat gga cgt	aat agc ttg atg aat	cct gga cct gct	3487
Trp	Ala Leu Asn Gly Arg	Asn Ser Leu Met Asn	Pro Gly Pro Ala	
1130		1135	1140	
atg	gcc agc cac aaa gaa	gga gag gac cgt ttc	ttt cct ttg tct	3532
Met	Ala Ser His Lys Glu	Gly Glu Asp Arg Phe	Phe Pro Leu Ser	
1145		1150	1155	
gga	tct tta att ttt ggc	aaa caa gga act gga	aga gac aac gtg	3577
Gly	Ser Leu Ile Phe Gly	Lys Gln Gly Thr Gly	Arg Asp Asn Val	
1160		1165	1170	
gat	gcg gac aaa gtc atg	ata acc aac gaa gaa	gaa att aaa act	3622
Asp	Ala Asp Lys Val Met	Ile Thr Asn Glu Glu	Glu Ile Lys Thr	
1175		1180	1185	
act	aac ccg gta gca acg	gag tcc tat gga caa	gtg gcc aca aac	3667
Thr	Asn Pro Val Ala Thr	Glu Ser Tyr Gly Gln	Val Ala Thr Asn	
1190		1195	1200	
cac	cag agt gcc caa gga	gga atc gag atc aac	gct acc aga gct	3712
His	Gln Ser Ala Gln Gly	Gly Ile Glu Ile Asn	Ala Thr Arg Ala	
1205		1210	1215	
gga	acc aac ctg gga gga	gca cag gcg cag acc	ggc tgg gtt caa	3757
Gly	Thr Asn Leu Gly Gly	Ala Gln Ala Gln Thr	Gly Trp Val Gln	
1220		1225	1230	
aac	caa gga ata ctt ccg	ggt atg gtt tgg cag	gac aga gat gtg	3802
Asn	Gln Gly Ile Leu Pro	Gly Met Val Trp Gln	Asp Arg Asp Val	
1235		1240	1245	
tac	ctg caa gga ccc att	tgg gcc aaa att cct	cac acg gac ggc	3847
Tyr	Leu Gln Gly Pro Ile	Trp Ala Lys Ile Pro	His Thr Asp Gly	
1250		1255	1260	
aac	ttt cac cct tct ccg	ctg atg gga ggg ttt	gga atg aag cac	3892
Asn	Phe His Pro Ser Pro	Leu Met Gly Gly Phe	Gly Met Lys His	
1265		1270	1275	
ccg	cct cct cag atc ctc	atc aaa aac aca cct	gta cct gcg gat	3937
Pro	Pro Pro Gln Ile Leu	Ile Lys Asn Thr Pro	Val Pro Ala Asp	
1280		1285	1290	
cct	cca acg gcc ttc aac	aag gac aag ctg aac	tct ttc atc acc	3982
Pro	Pro Thr Ala Phe Asn	Lys Asp Lys Leu Asn	Ser Phe Ile Thr	
1295		1300	1305	

cag tat tct act ggc caa gtc agc gtg gag atc gag tgg gag ctg	4027
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu	
1310 1315 1320	
cag aag gaa aac agc aag cgc tgg aac ccg gag atc cag tac act	4072
Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr	
1325 1330 1335	
tcc aac tat tac aag tct aat aat gtt gaa ttt gct gtt aat act	4117
Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr	
1340 1345 1350	
gaa ggt gta tat agt gaa ccc cgc ccc att ggc acc aga tac ctg	4162
Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu	
1355 1360 1365	
act cgt aat ctg taa ttgcttgta atcaataaac cgtttaattc gtttcagttg	4217
Thr Arg Asn Leu	
1370	
aactttggtc tctgcaagg gcaattcgt ttaaactgc aggactagag gtctgtatt	4277
agaggtcacg tgagtgtttt gcgacatttt gcgacacat gtggtcacgc tgggtattta	4337
agccccgagtg agcacgcagg gtctccattt tgaagcggga ggtttgaacg cgcagccgcc	4397
aagccgaatt ctgcagatat ccatcacact ggcgccgct cgactagagc ggccgccacc	4457
gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc	4517
atgggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg	4577
agccggaagc ataaagtgt aagcctgggg tgccaatga gtgagctaac tcacattaat	4637
tgcgttgccg tcaactgccc cttccagtc gggaacctg tegtgccagc tgcattaatg	4697
aatcgcccaa cgcgcgggga gaggcggttt gcgtattggg cgc	4740
<210> 8	
<211> 621	
<212> PRT	
<213> 人工序列	
<220>	
<223> 合成构建体	
<400> 8	
Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp	
1 5 10 15	
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu	
20 25 30	
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile	
35 40 45	
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu	
50 55 60	

Thr	Glu	Trp	Arg	Arg	Val	Ser	Lys	Ala	Pro	Glu	Ala	Leu	Phe	Phe	Val
65					70					75					80
Gln	Phe	Glu	Lys	Gly	Glu	Ser	Tyr	Phe	His	Met	His	Val	Leu	Val	Glu
				85					90						95
Thr	Thr	Gly	Val	Lys	Ser	Met	Val	Leu	Gly	Arg	Phe	Leu	Ser	Gln	Ile
			100					105						110	
Arg	Glu	Lys	Leu	Ile	Gln	Arg	Ile	Tyr	Arg	Gly	Ile	Glu	Pro	Thr	Leu
			115					120						125	
Pro	Asn	Trp	Phe	Ala	Val	Thr	Lys	Thr	Arg	Asn	Gly	Ala	Gly	Gly	Gly
								130						135	
Asn	Lys	Val	Val	Asp	Glu	Cys	Tyr	Ile	Pro	Asn	Tyr	Leu	Leu	Pro	Lys
145					150					155					160
Thr	Gln	Pro	Glu	Leu	Gln	Trp	Ala	Trp	Thr	Asn	Met	Glu	Gln	Tyr	Leu
					165					170					175
Ser	Ala	Cys	Leu	Asn	Leu	Thr	Glu	Arg	Lys	Arg	Leu	Val	Ala	Gln	His
					180					185					190
Leu	Thr	His	Val	Ser	Gln	Thr	Gln	Glu	Gln	Asn	Lys	Glu	Asn	Gln	Asn
					195					200					205
Pro	Asn	Ser	Asp	Ala	Pro	Val	Ile	Arg	Ser	Lys	Thr	Ser	Ala	Arg	Tyr
								210						215	
Met	Glu	Leu	Val	Gly	Trp	Leu	Val	Asp	Lys	Gly	Ile	Thr	Ser	Glu	Lys
225						230				235					240
Gln	Trp	Ile	Gln	Glu	Asp	Gln	Ala	Ser	Tyr	Ile	Ser	Phe	Asn	Ala	Ala
						245				250					255
Ser	Asn	Ser	Arg	Ser	Gln	Ile	Lys	Ala	Ala	Leu	Asp	Asn	Ala	Gly	Lys
					260					265					270
Ile	Met	Ser	Leu	Thr	Lys	Thr	Ala	Pro	Asp	Tyr	Leu	Val	Gly	Gln	Gln
					275					280					285
Pro	Val	Glu	Asp	Ile	Ser	Ser	Asn	Arg	Ile	Tyr	Lys	Ile	Leu	Glu	Leu
								290						295	
Asn	Gly	Tyr	Asp	Pro	Gln	Tyr	Ala	Ala	Ser	Val	Phe	Leu	Gly	Trp	Ala
305						310				315					320
Thr	Lys	Lys	Phe	Gly	Lys	Arg	Asn	Thr	Ile	Trp	Leu	Phe	Gly	Pro	Ala
						325				330					335
Thr	Thr	Gly	Lys	Thr	Asn	Ile	Ala	Glu	Ala	Ile	Ala	His	Thr	Val	Pro
					340					345					350
Phe	Tyr	Gly	Cys	Val	Asn	Trp	Thr	Asn	Glu	Asn	Phe	Pro	Phe	Asn	Asp
					355					360					365
Cys	Val	Asp	Lys	Met	Val	Ile	Trp	Trp	Glu	Glu	Gly	Lys	Met	Thr	Ala

370	375	380																		
Lys	Val	Val	Glu	Ser	Ala	Lys	Ala	Ile	Leu	Gly	Gly	Ser	Lys	Val	Arg					
385		390								395					400					
Val	Asp	Gln	Lys	Cys	Lys	Ser	Ser	Ala	Gln	Ile	Asp	Pro	Thr	Pro	Val					
		405								410					415					
Ile	Val	Thr	Ser	Asn	Thr	Asn	Met	Cys	Ala	Val	Ile	Asp	Gly	Asn	Ser					
		420								425					430					
Thr	Thr	Phe	Glu	His	Gln	Gln	Pro	Leu	Gln	Asp	Arg	Met	Phe	Lys	Phe					
		435								440					445					
Glu	Leu	Thr	Arg	Arg	Leu	Asp	His	Asp	Phe	Gly	Lys	Val	Thr	Lys	Gln					
		450								455					460					
Glu	Val	Lys	Asp	Phe	Phe	Arg	Trp	Ala	Lys	Asp	His	Val	Val	Glu	Val					
						470				475					480					
Glu	His	Glu	Phe	Tyr	Val	Lys	Lys	Gly	Gly	Ala	Lys	Lys	Arg	Pro	Ala					
						485				490					495					
Pro	Ser	Asp	Ala	Asp	Ile	Ser	Glu	Pro	Lys	Arg	Val	Arg	Glu	Ser	Val					
						500				505					510					
Ala	Gln	Pro	Ser	Thr	Ser	Asp	Ala	Glu	Ala	Ser	Ile	Asn	Tyr	Ala	Asp					
						515				520					525					
Arg	Tyr	Gln	Asn	Lys	Cys	Ser	Arg	His	Val	Gly	Met	Asn	Leu	Met	Leu					
						530				535					540					
Phe	Pro	Cys	Arg	Gln	Cys	Glu	Arg	Met	Asn	Gln	Asn	Ser	Asn	Ile	Cys					
						545				550					555					
Phe	Thr	His	Gly	Gln	Lys	Asp	Cys	Leu	Glu	Cys	Phe	Pro	Val	Ser	Glu					
						565				570					575					
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr					
						580				585					590					
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp					
						595				600					605					
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln								
						610				615					620					
<210>	9																			
<211>	752																			
<212>	PRT																			
<213>	人工序列																			
<220>																				
<223>	合成构建体																			
<400>	9																			
Met	Ala	Ala	Asp	Gly	Tyr	Leu	Pro	Asp	Trp	Leu	Glu	Asp	Asn	Leu	Ser					

1	5	10	15
Glu Gly Ile Arg	Glu Trp Trp Ala	Leu Lys Pro Gly	Ala Pro Gln Pro
20	25	30	
Lys Ala Asn Gln	Gln His Gln Asp	Asn Ala Arg Gly	Leu Val Leu Pro
35	40	45	
Gly Tyr Lys Tyr	Leu Gly Pro Gly	Asn Gly Leu Asp	Lys Gly Glu Pro
50	55	60	
Val Asn Ala Ala	Asp Ala Ala Ala	Leu Glu His Asp	Lys Ala Tyr Asp
65	70	75	80
Gln Gln Leu Lys	Ala Gly Asp Asn	Pro Tyr Leu Lys	Tyr Asn His Ala
85	90	95	
Asp Ala Glu Phe	Gln Glu Arg Leu	Lys Glu Asp Thr	Ser Phe Gly Gly
100	105	110	
Asn Leu Gly Arg	Ala Val Phe Gln	Ala Lys Lys Arg	Leu Leu Glu Pro
115	120	125	
Leu Gly Leu Val	Glu Glu Ala Ala	Lys Thr Ala Pro	Gly Lys Lys Arg
130	135	140	
Pro Val Glu Gln	Ser Pro Gln Glu	Pro Asp Ser Ser	Ala Gly Ile Gly
145	150	155	160
Lys Ser Gly Ala	Gln Pro Ala Lys	Lys Arg Leu Asn	Phe Gly Gln Thr
165	170	175	
Gly Asp Thr Glu	Ser Val Pro Asp	Pro Gln Pro Ile	Gly Glu Pro Pro
180	185	190	
Ala Ala Pro Ser	Gly Val Gly Ser	Leu Thr Met Ala	Ser Gly Gly Gly
195	200	205	
Ala Pro Val Ala	Asp Asn Asn Glu	Gly Ala Asp Gly	Val Gly Ser Ser
210	215	220	
Ser Gly Asn Trp	His Cys Asp Ser	Gln Trp Leu Gly	Asp Arg Val Ile
225	230	235	240
Thr Thr Ser Thr	Arg Thr Trp Ala	Leu Pro Thr Tyr	Asn Asn His Leu
245	250	255	
Tyr Lys Gln Ile	Ser Asn Ser Thr	Ser Gly Gly Ser	Ser Asn Asp Asn
260	265	270	
Ala Tyr Phe Gly	Tyr Ser Thr Pro	Trp Gly Tyr Phe	Asp Phe Asn Arg
275	280	285	
Phe His Cys His	Phe Ser Pro Arg	Asp Trp Gln Arg	Leu Ile Asn Asn
290	295	300	
Asn Trp Gly Phe	Arg Pro Lys Arg	Leu Asn Phe Lys	Leu Phe Asn Ile
305	310	315	320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Gly Gly Ile Glu
 580 585 590
 Ile Asn Ala Thr Arg Ala Gly Thr Asn Leu Gly Gly Ala Gln Ala Gln
 595 600 605
 Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
 610 615 620
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His

625	630	635	640
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met			
	645	650	655
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala			
	660	665	670
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr			
	675	680	685
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln			
	690	695	700
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn			
705	710	715	720
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val			
	725	730	735
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	750

<210> 10

<211> 4722

<212> DNA

<213> 人工序列

<220>

<223> AAV2/9 n.588.SAN核酸序列表达盒

<220>

<221> misc_feature

<222> (1) .. (36)

<223> 截短启动子

<220>

<221> promoter

<222> (1) .. (7)

<223> p5启动子

<220>

<221> CDS

<222> (37) .. (1899)

<223> AAV2 Rep

<220>

<221> CDS

<222> (1919) .. (4159)

<223> AAV 9 Cap

<220>

<221> misc_feature

<222>	(3683) .. (3712)	
<223>	SAN	
<220>		
<221>	misc_feature	
<222>	(4250) .. (4380)	
<223>	p5启动子	
<220>		
<221>	misc_feature	
<222>	(4508) .. (4722)	
<223>	LacZ启动子	
<400>	10	
ccattttgaa gcgggaggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag		54
	Met Pro Gly Phe Tyr Glu	
	1 5	
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att		102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile		
	10 15 20	
tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg		150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro		
	25 30 35	
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc		198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr		
	40 45 50	
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg		246
Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val		
	55 60 65 70	
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag		294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu		
	75 80 85	
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc		342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser		
	90 95 100	
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag		390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln		
	105 110 115	
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc		438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val		
	120 125 130	
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag		486

Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu	
135	140 145 150
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag	534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln	
155 160 165	
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc	582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu	
170 175 180	
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tgc cag	630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln	
185 190 195	
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg	678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200 205 210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215 220 225 230	
ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
235 240 245	
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
250 255 260	
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys	
265 270 275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc	918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser	
280 285 290	
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa	966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln	
295 300 305 310	
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag	1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys	
315 320 325	
agg aac acc atc tgg ctg ttt ggg cct gea act acc ggg aag acc aac	1062
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn	
330 335 340	
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac	1110

Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn	
345	350 355
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg	1158
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val	
360	365 370
atc tgg tgg gag gag ggg aag atg acc gcc aag gtc gtg gag tcg gcc	1206
Ile Trp Trp Glu Glu Gly Lys Met Thr Ala Lys Val Val Glu Ser Ala	
375	380 385 390
aaa gcc att ctc gga gga agc aag gtg cgc gtg gac cag aaa tgc aag	1254
Lys Ala Ile Leu Gly Gly Ser Lys Val Arg Val Asp Gln Lys Cys Lys	
395	400 405
tcc tcg gcc cag ata gac ccg act ccc gtg atc gtc acc tcc aac acc	1302
Ser Ser Ala Gln Ile Asp Pro Thr Pro Val Ile Val Thr Ser Asn Thr	
410	415 420
aac atg tgc gcc gtg att gac ggg aac tca acg acc ttc gaa cac cag	1350
Asn Met Cys Ala Val Ile Asp Gly Asn Ser Thr Thr Phe Glu His Gln	
425	430 435
cag ccg ttg caa gac cgg atg ttc aaa ttt gaa ctc acc cgc cgt ctg	1398
Gln Pro Leu Gln Asp Arg Met Phe Lys Phe Glu Leu Thr Arg Arg Leu	
440	445 450
gat cat gac ttt ggg aag gtc acc aag cag gaa gtc aaa gac ttt ttc	1446
Asp His Asp Phe Gly Lys Val Thr Lys Gln Glu Val Lys Asp Phe Phe	
455	460 465 470
cgg tgg gca aag gat cac gtg gtt gag gtg gag cat gaa ttc tac gtc	1494
Arg Trp Ala Lys Asp His Val Val Glu Val Glu His Glu Phe Tyr Val	
475	480 485
aaa aag ggt gga gcc aag aaa aga ccc gcc ccc agt gac gca gat ata	1542
Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile	
490	495 500
agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca	1590
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser	
505	510 515
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt	1638
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys	
520	525 530
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc	1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys	
535	540 545 550
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa	1734

Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys	
555	560
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc	1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val	
570	575
gac aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585	590
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600	605
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat	1930
Asp Asp Cys Ile Phe Glu Gln	Met Ala Ala Asp
615	620
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc	1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg	
630	635
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa	2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln	
645	650
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac	2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr	
660	665
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca	2122
Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala	
675	680
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag	2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys	
690	695
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc	2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe	
710	715
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga	2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg	
725	730
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt	2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val	
740	745
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag	2362

Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln 755	760	765	
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca			2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala 770	775	780	785
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag			2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu 790	795	800	
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca			2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser 805	810	815	
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca			2554
Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala 820	825	830	
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg			2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp 835	840	845	
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc			2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr 850	855	860	865
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc			2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile 870	875	880	
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc			2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly 885	890	895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac			2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His 900	905	910	
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc			2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe 915	920	925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag			2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu 930	935	940	945
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc			2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser 950	955	960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc			2986

Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965	970
975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980	985
990	
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995	1000
1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010	1015
1020	
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu	
1025	1030
1035	
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac	3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp	
1040	1045
1050	
cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
1055	1060
1065	
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc	3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe	
1070	1075
1080	
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac	3352
Ser Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr	
1085	1090
1095	
ata cct gga ccc agc tac cga caa caa cgt gtc tca acc act gtg	3397
Ile Pro Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val	
1100	1105
1110	
act caa aac aac aac agc gaa ttt gct tgg cct gga gct tct tct	3442
Thr Gln Asn Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser	
1115	1120
1125	
tgg gct ctc aat gga cgt aat agc ttg atg aat cct gga cct gct	3487
Trp Ala Leu Asn Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala	
1130	1135
1140	
atg gcc agc cac aaa gaa gga gag gac cgt ttc ttt cct ttg tct	3532
Met Ala Ser His Lys Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser	
1145	1150
1155	
gga tct tta att ttt ggc aaa caa gga act gga aga gac aac gtg	3577

Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu	
1355	1360 1365
ttgcttggtta atcaataaac cgtttaattc gtttcagttg aactttggtc tctgcgaagg	4219
gcgaattcgt ttaaacctgc aggactagag gtctgtatt agaggtcacg tgagtgtttt	4279
gcgacatttt gcgacaccat gtggtcacgc tgggtattta agcccagtg agcacgcagg	4339
gtctccattt tgaagcggga gtttgaacg cgcagccgcc aagccgaatt ctgcagatat	4399
ccatcacact ggcgccgct cgactagagc ggccgccacc gcggtggagc tccagctttt	4459
gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag ctgtttcctg	4519
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgt	4579
aagcctgggg tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcaactgccc	4639
ctttccagtc gggaaacctg tegtgccagc tgcattaatg aatcggccaa cgcgcgggga	4699
gaggcggttt gcgtattggg cgc	4722
<210>	11
<211>	621
<212>	PRT
<213>	人工序列
<220>	
<223>	合成构建体
<400>	11
Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp	
1	5 10 15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu	
	20 25 30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile	
	35 40 45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu	
	50 55 60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val	
65	70 75 80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu	
	85 90 95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile	
	100 105 110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu	
	115 120 125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly	
	130 135 140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys	
145	150 155 160

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

	100		105		110										
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Leu	Leu	Glu	Pro
	115		120		125										
Leu	Gly	Leu	Val	Glu	Glu	Ala	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg
	130		135		140										
Pro	Val	Glu	Gln	Ser	Pro	Gln	Glu	Pro	Asp	Ser	Ser	Ala	Gly	Ile	Gly
145			150		155										
Lys	Ser	Gly	Ala	Gln	Pro	Ala	Lys	Lys	Arg	Leu	Asn	Phe	Gly	Gln	Thr
			165		170										
Gly	Asp	Thr	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	Pro
	180		185		190										
Ala	Ala	Pro	Ser	Gly	Val	Gly	Ser	Leu	Thr	Met	Ala	Ser	Gly	Gly	Gly
	195		200		205										
Ala	Pro	Val	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	Ser
	210		215		220										
Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Gln	Trp	Leu	Gly	Asp	Arg	Val	Ile
225			230		235										
Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu
			245		250										
Tyr	Lys	Gln	Ile	Ser	Asn	Ser	Thr	Ser	Gly	Gly	Ser	Ser	Asn	Asp	Asn
	260		265		270										
Ala	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg
	275		280		285										
Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn
	290		295		300										
Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile
305			310		315										
Gln	Val	Lys	Glu	Val	Thr	Asp	Asn	Asn	Gly	Val	Lys	Thr	Ile	Ala	Asn
			325		330										
Asn	Leu	Thr	Ser	Thr	Val	Gln	Val	Phe	Thr	Asp	Ser	Asp	Tyr	Gln	Leu
	340		345		350										
Pro	Tyr	Val	Leu	Gly	Ser	Ala	His	Glu	Gly	Cys	Leu	Pro	Pro	Phe	Pro
	355		360		365										
Ala	Asp	Val	Phe	Met	Ile	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asp
	370		375		380										
Gly	Ser	Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe
385			390		395										
Pro	Ser	Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Gln	Phe	Ser	Tyr	Glu
			405		410										

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ser Ala Asn Phe
 580 585 590
 Ile Lys Pro Thr Ser Tyr Ala Gln Ala Gln Thr Gly Trp Val Gln Asn
 595 600 605
 Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu
 610 615 620
 Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His
 625 630 635 640
 Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln
 645 650 655
 Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe
 660 665 670
 Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln
 675 680 685
 Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg
 690 695 700
 Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn Asn
 705 710 715 720
 Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg Pro

	725	730	735
Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	
<210>	13		
<211>	4725		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV2/9 n.588.SSN核酸序列表达盒		
<220>			
<221>	misc_feature		
<222>	(1) .. (36)		
<223>	截短启动子		
<220>			
<221>	promoter		
<222>	(1) .. (7)		
<223>	p5启动子		
<220>			
<221>	CDS		
<222>	(37) .. (1899)		
<223>	AAV2 Rep		
<220>			
<221>	CDS		
<222>	(1919) .. (4162)		
<223>	AAV9 Cap		
<220>			
<221>	misc_feature		
<222>	(3683) .. (3715)		
<223>	SSN		
<220>			
<221>	misc_feature		
<222>	(4253) .. (4383)		
<223>	p5启动子		
<220>			
<221>	misc_feature		
<222>	(4511) .. (4725)		
<223>	LacZ启动子		
<400>	13		

ccattttgaa gcgggagggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag

54

	Met	Pro	Gly	Phe	Tyr	Glu	
	1				5		
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att							102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile							
	10		15		20		
tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg							150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro							
	25		30		35		
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc							198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr							
	40		45		50		
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg							246
Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val							
	55		60		65		70
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag							294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu							
	75		80		85		
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc							342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser							
	90		95		100		
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag							390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln							
	105		110		115		
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc							438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val							
	120		125		130		
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag							486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu							
	135		140		145		150
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag							534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln							
	155		160		165		
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc							582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu							
	170		175		180		
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tcg cag							630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln							
	185		190		195		
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg							678

Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200	205 210
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215	220 225 230
ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
	235 240 245
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
	250 255 260
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys	
	265 270 275
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc	918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser	
	280 285 290
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa	966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln	
295	300 305 310
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag	1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys	
	315 320 325
agg aac acc atc tgg ctg ttt ggg cct gca act acc ggg aag acc aac	1062
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn	
	330 335 340
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac	1110
Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn	
	345 350 355
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg	1158
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val	
	360 365 370
atc tgg tgg gag gag ggg aag atg acc gcc aag gtc gtg gag tcg gcc	1206
Ile Trp Trp Glu Glu Gly Lys Met Thr Ala Lys Val Val Glu Ser Ala	
375	380 385 390
aaa gcc att ctc gga gga agc aag gtg cgc gtg gac cag aaa tgc aag	1254
Lys Ala Ile Leu Gly Gly Ser Lys Val Arg Val Asp Gln Lys Cys Lys	
	395 400 405
tcc tcg gcc cag ata gac ccg act ccc gtg atc gtc acc tcc aac acc	1302

Ser Ser Ala Gln Ile Asp Pro Thr Pro Val Ile Val Thr Ser Asn Thr	
410	415
aac atg tgc gcc gtg att gac ggg aac tca acg acc ttc gaa cac cag	1350
Asn Met Cys Ala Val Ile Asp Gly Asn Ser Thr Thr Phe Glu His Gln	
425	430
cag ccg ttg caa gac cgg atg ttc aaa ttt gaa ctc acc cgc cgt ctg	1398
Gln Pro Leu Gln Asp Arg Met Phe Lys Phe Glu Leu Thr Arg Arg Leu	
440	445
gat cat gac ttt ggg aag gtc acc aag cag gaa gtc aaa gac ttt ttc	1446
Asp His Asp Phe Gly Lys Val Thr Lys Gln Glu Val Lys Asp Phe Phe	
455	460
cgg tgg gca aag gat cac gtg gtt gag gtg gag cat gaa ttc tac gtc	1494
Arg Trp Ala Lys Asp His Val Val Glu Val Glu His Glu Phe Tyr Val	
475	480
aaa aag ggt gga gcc aag aaa aga ccc gcc ccc agt gac gca gat ata	1542
Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile	
490	495
agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca	1590
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser	
505	510
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt	1638
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys	
520	525
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc	1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys	
535	540
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa	1734
Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys	
555	560
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc	1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val	
570	575
gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585	590
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600	605
gat gac tgc atc ttt gaa caa taaatgattt aatcaggt atg gct gcc gat	1930

Asp Asp Cys Ile Phe Glu Gln	Met Ala Ala Asp	
615	620	625
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc		1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg		
630	635	640
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa		2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln		
645	650	655
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac		2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr		
660	665	670
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca		2122
Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala		
675	680	685
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag		2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys		
690	695	700
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc		2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe		
710	715	720
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga		2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg		
725	730	735
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt		2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val		
740	745	750
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag		2362
Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln		
755	760	765
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca		2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala		
770	775	780
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag		2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu		
790	795	800
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca		2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser		
805	810	815
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca		2554

Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala	
820	825 830
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg	2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp	
835	840 845
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc	2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr	
850	855 860 865
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc	2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile	
870	875 880
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc	2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly	
885	890 895
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac	2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
900	905 910
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc	2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe	
915	920 925
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag	2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
930	935 940 945
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc	2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser	
950	955 960
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc	2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965	970 975
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980	985 990
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995	1000 1005
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010	1015 1020
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172

His Ala Gln Ala Gln Thr Gly Trp Val Gln Asn Gln Gly Ile Leu 1220 1225 1230	
ccg ggt atg gtt tgg cag gac aga gat gtg tac ctg caa gga ccc Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro 1235 1240 1245	3802
att tgg gcc aaa att cct cac acg gac ggc aac ttt cac cct tct Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His Pro Ser 1250 1255 1260	3847
ccg ctg atg gga ggg ttt gga atg aag cac ccg cct cct cag atc Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln Ile 1265 1270 1275	3892
ctc atc aaa aac aca cct gta cct gcg gat cct cca acg gcc ttc Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe 1280 1285 1290	3937
aac aag gac aag ctg aac tct ttc atc acc cag tat tct act ggc Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly 1295 1300 1305	3982
caa gtc agc gtg gag atc gag tgg gag ctg cag aag gaa aac agc Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser 1310 1315 1320	4027
aag cgc tgg aac ccg gag atc cag tac act tcc aac tat tac aag Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys 1325 1330 1335	4072
tct aat aat gtt gaa ttt gct gtt aat act gaa ggt gta tat agt Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser 1340 1345 1350	4117
gaa ccc cgc ccc att ggc acc aga tac ctg act cgt aat ctg taa Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu 1355 1360 1365	4162
ttgcttggtta atcaataaac cgtttaattc gtttcagttg aactttggtc tctgcaagg	4222
gcgaattcgt ttaaacctgc aggactagag gtctgtatt agaggtcacg tgagtgtttt	4282
gcgacatttt gcgacacat gtggtcacgc tgggtattta agcccgagtg agcacgcagg	4342
gtctccattht tgaagcggga ggtttgaacg cgcagccgcc aagccgaatt ctgcagatat	4402
ccatcacact ggcggccgct cgaactagagc ggccgccacc gcggtggagc tccagctttt	4462
gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag ctgtttcctg	4522
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta	4582
aagcctgggg tgccaatga gtgagetaac tcacattaat tgcgttgccg tccactgcccg	4642
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga	4702
gaggcggtht gcgtattggg cgc	4725

<210> 14
 <211> 621
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成构建体
 <400> 14
 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
 1 5 10 15
 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
 20 25 30
 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
 35 40 45
 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
 50 55 60
 Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
 65 70 75 80
 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
 85 90 95
 Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
 100 105 110
 Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
 115 120 125
 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
 130 135 140
 Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
 145 150 155 160
 Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
 165 170 175
 Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
 195 200 205
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 225 230 235 240
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
 275 280 285
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 290 295 300
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
 305 310 315 320
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 325 330 335
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
 340 345 350
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 355 360 365
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 370 375 380
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
 385 390 395 400
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
 405 410 415
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
 420 425 430
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 435 440 445
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
 450 455 460
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
 465 470 475 480
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
 485 490 495
 Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
 500 505 510
 Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
 515 520 525
 Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
 530 535 540
 Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
 545 550 555 560
 Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu

	565		570		575														
Ser	Gln	Pro	Val	Ser	Val	Val	Lys	Lys	Ala	Tyr	Gln	Lys	Leu	Cys	Tyr				
			580						585					590					
Ile	His	His	Ile	Met	Gly	Lys	Val	Pro	Asp	Ala	Cys	Thr	Ala	Cys	Asp				
			595						600					605					
Leu	Val	Asn	Val	Asp	Leu	Asp	Asp	Cys	Ile	Phe	Glu	Gln							
			610						615					620					
<210>	15																		
<211>	747																		
<212>	PRT																		
<213>	人工序列																		
<220>																			
<223>	合成构建体																		
<400>	15																		
Met	Ala	Ala	Asp	Gly	Tyr	Leu	Pro	Asp	Trp	Leu	Glu	Asp	Asn	Leu	Ser				
1				5					10					15					
Glu	Gly	Ile	Arg	Glu	Trp	Trp	Ala	Leu	Lys	Pro	Gly	Ala	Pro	Gln	Pro				
			20						25					30					
Lys	Ala	Asn	Gln	Gln	His	Gln	Asp	Asn	Ala	Arg	Gly	Leu	Val	Leu	Pro				
			35						40					45					
Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Gly	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro				
			50						55					60					
Val	Asn	Ala	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp				
65									70					75					80
Gln	Gln	Leu	Lys	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Lys	Tyr	Asn	His	Ala				
									85					90					95
Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Lys	Glu	Asp	Thr	Ser	Phe	Gly	Gly				
			100						105					110					
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Leu	Leu	Glu	Pro				
			115						120					125					
Leu	Gly	Leu	Val	Glu	Glu	Ala	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg				
			130						135					140					
Pro	Val	Glu	Gln	Ser	Pro	Gln	Glu	Pro	Asp	Ser	Ser	Ala	Gly	Ile	Gly				
145									150					155					160
Lys	Ser	Gly	Ala	Gln	Pro	Ala	Lys	Lys	Arg	Leu	Asn	Phe	Gly	Gln	Thr				
									165					170					175
Gly	Asp	Thr	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	Pro				
			180						185					190					
Ala	Ala	Pro	Ser	Gly	Val	Gly	Ser	Leu	Thr	Met	Ala	Ser	Gly	Gly	Gly				

195	200	205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser		
210	215	220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile		
225	230	235
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu		
	245	250
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn		
	260	265
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg		
	275	280
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn		
	290	295
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile		
305	310	315
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn		
	325	330
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu		
	340	345
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro		
	355	360
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp		
	370	375
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe		
385	390	395
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu		
	405	410
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu		
	420	425
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser		
	435	440
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser		
	450	455
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro		
465	470	475
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn		
	485	490
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn		
	500	505
		510

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

<210> 16

<211> 4737

<212> DNA

<213> 人工序列

<220>

<223> AAV2/9 n.588.SSN-L核酸序列表达盒

<220>

<221> misc_feature

<222> (1) .. (36)

<223> 截短启动子	
<220>	
<221> promoter	
<222> (1) .. (7)	
<223> p5启动子	
<220>	
<221> CDS	
<222> (37) .. (1899)	
<223> AAV2 Rep	
<220>	
<221> CDS	
<222> (1919) .. (4174)	
<223> AAV9 Cap	
<220>	
<221> misc_feature	
<222> (3683) .. (3727)	
<223> SSN-L	
<220>	
<221> misc_feature	
<222> (4253) .. (4737)	
<223> LacZ启动子	
<220>	
<221> misc_feature	
<222> (4265) .. (4395)	
<223> p5启动子	
<400> 16	
ccattttgaa gcgggagggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag	54
	Met Pro Gly Phe Tyr Glu
	1 5
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att	102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile	
	10 15 20
tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg	150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro	
	25 30 35
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc	198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr	
	40 45 50
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg	246

Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val	
55	60
65	70
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag	294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu	
75	80
85	
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc	342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser	
90	95
100	
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag	390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln	
105	110
115	
aga att tac cgc ggg atc gag ccg act ttg cca aac tgg ttc gcg gtc	438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val	
120	125
130	
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag	486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu	
135	140
145	150
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag	534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln	
155	160
165	
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc	582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu	
170	175
180	
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tcg cag	630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln	
185	190
195	
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg	678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro	
200	205
210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg	726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp	
215	220
225	230
ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac	774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp	
235	240
245	
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa	822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln	
250	255
260	
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa	870

Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys 265	270	275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc			918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser 280	285	290	
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa			966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln 295	300	305	310
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag			1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys 315	320	325	
agg aac acc atc tgg ctg ttt ggg cct gea act acc ggg aag acc aac			1062
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn 330	335	340	
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac			1110
Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn 345	350	355	
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg			1158
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val 360	365	370	
atc tgg tgg gag gag ggg aag atg acc gcc aag gtc gtg gag tcg gcc			1206
Ile Trp Trp Glu Glu Gly Lys Met Thr Ala Lys Val Val Glu Ser Ala 375	380	385	390
aaa gcc att ctc gga gga agc aag gtg cgc gtg gac cag aaa tgc aag			1254
Lys Ala Ile Leu Gly Gly Ser Lys Val Arg Val Asp Gln Lys Cys Lys 395	400	405	
tcc tcg gcc cag ata gac ccg act ccc gtg atc gtc acc tcc aac acc			1302
Ser Ser Ala Gln Ile Asp Pro Thr Pro Val Ile Val Thr Ser Asn Thr 410	415	420	
aac atg tgc gcc gtg att gac ggg aac tca acg acc ttc gaa cac cag			1350
Asn Met Cys Ala Val Ile Asp Gly Asn Ser Thr Thr Phe Glu His Gln 425	430	435	
cag ccg ttg caa gac cgg atg ttc aaa ttt gaa ctc acc cgc cgt ctg			1398
Gln Pro Leu Gln Asp Arg Met Phe Lys Phe Glu Leu Thr Arg Arg Leu 440	445	450	
gat cat gac ttt ggg aag gtc acc aag cag gaa gtc aaa gac ttt ttc			1446
Asp His Asp Phe Gly Lys Val Thr Lys Gln Glu Val Lys Asp Phe Phe 455	460	465	470
cgg tgg gca aag gat cac gtg gtt gag gtg gag cat gaa ttc tac gtc			1494

Arg Trp Ala Lys Asp His Val Val Glu Val Glu His Glu Phe Tyr Val	
475	480 485
aaa aag ggt gga gcc aag aaa aga ccc gcc ccc agt gac gca gat ata	1542
Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile	
490	495 500
agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca	1590
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser	
505	510 515
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt	1638
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys	
520	525 530
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc	1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys	
535	540 545 550
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa	1734
Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys	
555	560 565
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc	1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val	
570	575 580
gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga	1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly	
585	590 595
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg	1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu	
600	605 610
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat	1930
Asp Asp Cys Ile Phe Glu Gln	Met Ala Ala Asp
615	620 625
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc	1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg	
630	635 640
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa	2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln	
645	650 655
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac	2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr	
660	665 670
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca	2122

Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala	
675	680 685
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag	2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys	
690	695 700 705
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc	2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe	
	710 715 720
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga	2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg	
	725 730 735
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt	2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val	
	740 745 750
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag	2362
Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln	
	755 760 765
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca	2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala	
770	775 780 785
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag	2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu	
	790 795 800
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca	2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser	
	805 810 815
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca	2554
Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala	
	820 825 830
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg	2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp	
	835 840 845
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc	2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr	
850	855 860 865
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc	2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile	
	870 875 880
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc	2746

Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly	
885	890
895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac	2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His	
900	905
910	
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc	2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe	
915	920
925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag	2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu	
930	935
940	945
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc	2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser	
950	955
960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc	2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu	
965	970
975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc	3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe	
980	985
990	
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc	3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala	
995	1000
1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa	3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln	
1010	1015
1020	
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag	3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu	
1025	1030
1035	
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac	3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp	
1040	1045
1050	
cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca	3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser	
1055	1060
1065	
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc	3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe	
1070	1075
1080	
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac	3352

	20	25	30
Lys Glu Trp	Glu Leu Pro Pro	Asp Ser Asp Met	Asp Leu Asn Leu Ile
	35	40	45
Glu Gln Ala	Pro Leu Thr Val	Ala Glu Lys Leu	Gln Arg Asp Phe Leu
	50	55	60
Thr Glu Trp	Arg Arg Val Ser	Lys Ala Pro Glu	Ala Leu Phe Phe Val
65	70	75	80
Gln Phe Glu	Lys Gly Glu Ser	Tyr Phe His Met	His Val Leu Val Glu
	85	90	95
Thr Thr Gly	Val Lys Ser Met	Val Leu Gly Arg	Phe Leu Ser Gln Ile
	100	105	110
Arg Glu Lys	Leu Ile Gln Arg	Ile Tyr Arg Gly	Ile Glu Pro Thr Leu
	115	120	125
Pro Asn Trp	Phe Ala Val Thr	Lys Thr Arg Asn	Gly Ala Gly Gly Gly
	130	135	140
Asn Lys Val	Val Asp Glu Cys	Tyr Ile Pro Asn	Tyr Leu Leu Pro Lys
145	150	155	160
Thr Gln Pro	Glu Leu Gln Trp	Ala Trp Thr Asn	Met Glu Gln Tyr Leu
	165	170	175
Ser Ala Cys	Leu Asn Leu Thr	Glu Arg Lys Arg	Leu Val Ala Gln His
	180	185	190
Leu Thr His	Val Ser Gln Thr	Gln Glu Gln Asn	Lys Glu Asn Gln Asn
	195	200	205
Pro Asn Ser	Asp Ala Pro Val	Ile Arg Ser Lys	Thr Ser Ala Arg Tyr
	210	215	220
Met Glu Leu	Val Gly Trp Leu	Val Asp Lys Gly	Ile Thr Ser Glu Lys
225	230	235	240
Gln Trp Ile	Gln Glu Asp Gln	Ala Ser Tyr Ile	Ser Phe Asn Ala Ala
	245	250	255
Ser Asn Ser	Arg Ser Gln Ile	Lys Ala Ala Leu	Asp Asn Ala Gly Lys
	260	265	270
Ile Met Ser	Leu Thr Lys Thr	Ala Pro Asp Tyr	Leu Val Gly Gln Gln
	275	280	285
Pro Val Glu	Asp Ile Ser Ser	Asn Arg Ile Tyr	Lys Ile Leu Glu Leu
	290	295	300
Asn Gly Tyr	Asp Pro Gln Tyr	Ala Ala Ser Val	Phe Leu Gly Trp Ala
305	310	315	320
Thr Lys Lys	Phe Gly Lys Arg	Asn Thr Ile Trp	Leu Phe Gly Pro Ala
	325	330	335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415
Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser		
420	425	430
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe		
435	440	445
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln		
450	455	460
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val		
465	470	475
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala		
485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val		
500	505	510
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp		
515	520	525
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu		
530	535	540
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys		
545	550	555
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu		
565	570	575
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr		
580	585	590
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp		
595	600	605
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln		
610	615	620
<210> 18		
<211> 751		
<212> PRT		

275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn		
290	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile		
305	310	315
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn		
325	330	335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu		
340	345	350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro		
355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp		
370	375	380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe		
385	390	395
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu		
405	410	415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu		
420	425	430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser		
435	440	445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser		
450	455	460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro		
465	470	475
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn		
485	490	495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn		
500	505	510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys		
515	520	525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly		
530	535	540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile		
545	550	555
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser		
565	570	575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Gly Gly Ser Ser		
580	585	590

Asn Thr Val Lys Leu Thr Ser Gly His Gly Gly Ala Gln Ala Gln Thr
 595 600 605
 Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp
 610 615 620
 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
 625 630 635 640
 Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys
 645 650 655
 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp
 660 665 670
 Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln
 675 680 685
 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
 690 695 700
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
 705 710 715 720
 Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr
 725 730 735
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 740 745 750

<210> 19

<211> 4740

<212> DNA

<213> 人工序列

<220>

<223> AAV2/9 n.588.VLT-L核酸序列表达盒

<220>

<221> misc_feature

<222> (1) .. (36)

<223> 截短启动子

<220>

<221> promoter

<222> (1) .. (7)

<223> p5启动子

<220>

<221> CDS

<222> (37) .. (1899)

<223> AAV2 Rep

<220>

<221> CDS	
<222> (1919) .. (4177)	
<223> AAV9 Cap	
<220>	
<221> misc_feature	
<222> (3683) .. (3730)	
<223> VLT-L	
<220>	
<221> misc_feature	
<222> (4268) .. (4398)	
<223> p5启动子	
<220>	
<221> misc_feature	
<222> (4526) .. (4740)	
<223> LacZ启动子	
<400> 19	
ccattttgaa gcgggaggtt tgaacgcgca gccgcc atg ccg ggg ttt tac gag	54
	Met Pro Gly Phe Tyr Glu
	1 5
att gtg att aag gtc ccc agc gac ctt gac gag cat ctg ccc ggc att	102
Ile Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile	
	10 15 20
tct gac agc ttt gtg aac tgg gtg gcc gag aag gaa tgg gag ttg ccg	150
Ser Asp Ser Phe Val Asn Trp Val Ala Glu Lys Glu Trp Glu Leu Pro	
	25 30 35
cca gat tct gac atg gat ctg aat ctg att gag cag gca ccc ctg acc	198
Pro Asp Ser Asp Met Asp Leu Asn Leu Ile Glu Gln Ala Pro Leu Thr	
	40 45 50
gtg gcc gag aag ctg cag cgc gac ttt ctg acg gaa tgg cgc cgt gtg	246
Val Ala Glu Lys Leu Gln Arg Asp Phe Leu Thr Glu Trp Arg Arg Val	
	55 60 65 70
agt aag gcc ccg gag gct ctt ttc ttt gtg caa ttt gag aag gga gag	294
Ser Lys Ala Pro Glu Ala Leu Phe Phe Val Gln Phe Glu Lys Gly Glu	
	75 80 85
agc tac ttc cac atg cac gtg ctc gtg gaa acc acc ggg gtg aaa tcc	342
Ser Tyr Phe His Met His Val Leu Val Glu Thr Thr Gly Val Lys Ser	
	90 95 100
atg gtt ttg gga cgt ttc ctg agt cag att cgc gaa aaa ctg att cag	390
Met Val Leu Gly Arg Phe Leu Ser Gln Ile Arg Glu Lys Leu Ile Gln	

105	110	115	
aga att tac cgc ggg atc gag cgc act ttg cca aac tgg ttc gcg gtc			438
Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu Pro Asn Trp Phe Ala Val			
120	125	130	
aca aag acc aga aat ggc gcc gga ggc ggg aac aag gtg gtg gat gag			486
Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly Asn Lys Val Val Asp Glu			
135	140	145	150
tgc tac atc ccc aat tac ttg ctc ccc aaa acc cag cct gag ctc cag			534
Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys Thr Gln Pro Glu Leu Gln			
	155	160	165
tgg gcg tgg act aat atg gaa cag tat tta agc gcc tgt ttg aat ctc			582
Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu Ser Ala Cys Leu Asn Leu			
	170	175	180
acg gag cgt aaa cgg ttg gtg gcg cag cat ctg acg cac gtg tcg cag			630
Thr Glu Arg Lys Arg Leu Val Ala Gln His Leu Thr His Val Ser Gln			
	185	190	195
acg cag gag cag aac aaa gag aat cag aat ccc aat tct gat gcg ccg			678
Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn Pro Asn Ser Asp Ala Pro			
200	205	210	
gtg atc aga tca aaa act tca gcc agg tac atg gag ctg gtc ggg tgg			726
Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr Met Glu Leu Val Gly Trp			
215	220	225	230
ctc gtg gac aag ggg att acc tcg gag aag cag tgg atc cag gag gac			774
Leu Val Asp Lys Gly Ile Thr Ser Glu Lys Gln Trp Ile Gln Glu Asp			
	235	240	245
cag gcc tca tac atc tcc ttc aat gcg gcc tcc aac tcg cgg tcc caa			822
Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala Ser Asn Ser Arg Ser Gln			
	250	255	260
atc aag gct gcc ttg gac aat gcg gga aag att atg agc ctg act aaa			870
Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys Ile Met Ser Leu Thr Lys			
265	270	275	
acc gcc ccc gac tac ctg gtg ggc cag cag ccc gtg gag gac att tcc			918
Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln Pro Val Glu Asp Ile Ser			
280	285	290	
agc aat cgg att tat aaa att ttg gaa cta aac ggg tac gat ccc caa			966
Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu Asn Gly Tyr Asp Pro Gln			
295	300	305	310
tat gcg gct tcc gtc ttt ctg gga tgg gcc acg aaa aag ttc ggc aag			1014
Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala Thr Lys Lys Phe Gly Lys			

315	320	325	1062
agg aac acc atc tgg ctg ttt ggg cct gca act acc ggg aag acc aac			
Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala Thr Thr Gly Lys Thr Asn			
330	335	340	1110
atc gcg gag gcc ata gcc cac act gtg ccc ttc tac ggg tgc gta aac			
Ile Ala Glu Ala Ile Ala His Thr Val Pro Phe Tyr Gly Cys Val Asn			
345	350	355	1158
tgg acc aat gag aac ttt ccc ttc aac gac tgt gtc gac aag atg gtg			
Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp Cys Val Asp Lys Met Val			
360	365	370	1206
atc tgg tgg gag gag ggg aag atg acc gcc aag gtc gtg gag tcg gcc			
Ile Trp Trp Glu Glu Gly Lys Met Thr Ala Lys Val Val Glu Ser Ala			
375	380	385	1254
aaa gcc att ctc gga gga agc aag gtg cgc gtg gac cag aaa tgc aag			
Lys Ala Ile Leu Gly Gly Ser Lys Val Arg Val Asp Gln Lys Cys Lys			
395	400	405	1302
tcc tcg gcc cag ata gac ccg act ccc gtg atc gtc acc tcc aac acc			
Ser Ser Ala Gln Ile Asp Pro Thr Pro Val Ile Val Thr Ser Asn Thr			
410	415	420	1350
aac atg tgc gcc gtg att gac ggg aac tca acg acc ttc gaa cac cag			
Asn Met Cys Ala Val Ile Asp Gly Asn Ser Thr Thr Phe Glu His Gln			
425	430	435	1398
cag ccg ttg caa gac cgg atg ttc aaa ttt gaa ctc acc cgc cgt ctg			
Gln Pro Leu Gln Asp Arg Met Phe Lys Phe Glu Leu Thr Arg Arg Leu			
440	445	450	1446
gat cat gac ttt ggg aag gtc acc aag cag gaa gtc aaa gac ttt ttc			
Asp His Asp Phe Gly Lys Val Thr Lys Gln Glu Val Lys Asp Phe Phe			
455	460	465	1494
cgg tgg gca aag gat cac gtg gtt gag gtg gag cat gaa ttc tac gtc			
Arg Trp Ala Lys Asp His Val Val Glu Val Glu His Glu Phe Tyr Val			
475	480	485	1542
aaa aag ggt gga gcc aag aaa aga ccc gcc ccc agt gac gca gat ata			
Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala Pro Ser Asp Ala Asp Ile			
490	495	500	1590
agt gag ccc aaa cgg gtg cgc gag tca gtt gcg cag cca tcg acg tca			
Ser Glu Pro Lys Arg Val Arg Glu Ser Val Ala Gln Pro Ser Thr Ser			
505	510	515	1638
gac gcg gaa gct tcg atc aac tac gca gac agg tac caa aac aaa tgt			
Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp Arg Tyr Gln Asn Lys Cys			

520	525	530	
tct cgt cac gtg ggc atg aat ctg atg ctg ttt ccc tgc aga caa tgc			1686
Ser Arg His Val Gly Met Asn Leu Met Leu Phe Pro Cys Arg Gln Cys			
535	540	545	550
gag aga atg aat cag aat tca aat atc tgc ttc act cac gga cag aaa			1734
Glu Arg Met Asn Gln Asn Ser Asn Ile Cys Phe Thr His Gly Gln Lys			
	555	560	565
gac tgt tta gag tgc ttt ccc gtg tca gaa tct caa ccc gtt tct gtc			1782
Asp Cys Leu Glu Cys Phe Pro Val Ser Glu Ser Gln Pro Val Ser Val			
	570	575	580
gtc aaa aag gcg tat cag aaa ctg tgc tac att cat cat atc atg gga			1830
Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr Ile His His Ile Met Gly			
	585	590	595
aag gtg cca gac gct tgc act gcc tgc gat ctg gtc aat gtg gat ttg			1878
Lys Val Pro Asp Ala Cys Thr Ala Cys Asp Leu Val Asn Val Asp Leu			
600	605	610	
gat gac tgc atc ttt gaa caa taaatgattt aaatcaggt atg gct gcc gat			1930
Asp Asp Cys Ile Phe Glu Gln			Met Ala Ala Asp
615	620	625	
ggt tat ctt cca gat tgg ctc gag gac aac ctt agt gaa gga att cgc			1978
Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser Glu Gly Ile Arg			
	630	635	640
gag tgg tgg gct ttg aaa cct gga gcc cct caa ccc aag gca aat caa			2026
Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro Lys Ala Asn Gln			
	645	650	655
caa cat caa gac aac gct cga ggt ctt gtg ctt ccg ggt tac aaa tac			2074
Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro Gly Tyr Lys Tyr			
	660	665	670
ctt gga ccc ggc aac gga ctc gac aag ggg gag ccg gtc aac gca gca			2122
Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Ala Ala			
	675	680	685
gac gcg gcg gcc ctc gag cac gac aag gcc tac gac cag cag ctc aag			2170
Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp Gln Gln Leu Lys			
690	695	700	705
gcc gga gac aac ccg tac ctc aag tac aac cac gcc gac gcc gag ttc			2218
Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala Asp Ala Glu Phe			
	710	715	720
cag gag cgg ctc aaa gaa gat acg tct ttt ggg ggc aac ctc ggg cga			2266
Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly Asn Leu Gly Arg			

725	730	735	
gca gtc ttc cag gcc aaa aag agg ctt ctt gaa cct ctt ggt ctg gtt			2314
Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro Leu Gly Leu Val			
740	745	750	
gag gaa gcg gct aag acg gct cct gga aag aag agg cct gta gag cag			2362
Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg Pro Val Glu Gln			
755	760	765	
tct cct cag gaa ccg gac tcc tcc gcg ggt att ggc aaa tcg ggt gca			2410
Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly Lys Ser Gly Ala			
770	775	780	785
cag ccc gct aaa aag aga ctc aat ttc ggt cag act ggc gac aca gag			2458
Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr Gly Asp Thr Glu			
790	795	800	
tca gtc cca gac cct caa cca atc gga gaa cct ccc gca gcc ccc tca			2506
Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro Ala Ala Pro Ser			
805	810	815	
ggt gtg gga tct ctt aca atg gct tca ggt ggt ggc gca cca gtg gca			2554
Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly Ala Pro Val Ala			
820	825	830	
gac aat aac gaa ggt gcc gat gga gtg ggt agt tcc tcg gga aat tgg			2602
Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser Ser Gly Asn Trp			
835	840	845	
cat tgc gat tcc caa tgg ctg ggg gac aga gtc atc acc acc agc acc			2650
His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile Thr Thr Ser Thr			
850	855	860	865
cga acc tgg gcc ctg ccc acc tac aac aat cac ctc tac aag caa atc			2698
Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile			
870	875	880	
tcc aac agc aca tct gga gga tct tca aat gac aac gcc tac ttc ggc			2746
Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn Ala Tyr Phe Gly			
885	890	895	
tac agc acc ccc tgg ggg tat ttt gac ttc aac aga ttc cac tgc cac			2794
Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His Cys His			
900	905	910	
ttc tca cca cgt gac tgg cag cga ctc atc aac aac aac tgg gga ttc			2842
Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp Gly Phe			
915	920	925	
cgg cct aag cga ctc aac ttc aag ctc ttc aac att cag gtc aaa gag			2890
Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln Val Lys Glu			

930	935	940	945	
gtt acg gac aac aat gga gtc aag acc atc gcc aat aac ctt acc agc				2938
Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn Asn Leu Thr Ser				
	950	955	960	
acg gtc cag gtc ttc acg gac tca gac tat cag ctc ccg tac gtg ctc				2986
Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu Pro Tyr Val Leu				
	965	970	975	
ggg tcg gct cac gag ggc tgc ctc ccg ccg ttc cca gcg gac gtt ttc				3034
Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro Ala Asp Val Phe				
	980	985	990	
atg att cct cag tac ggg tat ctg acg ctt aat gat gga agc cag gcc				3082
Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp Gly Ser Gln Ala				
	995	1000	1005	
gtg ggt cgt tcg tcc ttt tac tgc ctg gaa tat ttc ccg tcg caa				3127
Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln				
1010	1015	1020		
atg cta aga acg ggt aac aac ttc cag ttc agc tac gag ttt gag				3172
Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu Phe Glu				
1025	1030	1035		
aac gta cct ttc cat agc agc tac gct cac agc caa agc ctg gac				3217
Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp				
1040	1045	1050		
cga cta atg aat cca ctc atc gac caa tac ttg tac tat ctc tca				3262
Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser				
1055	1060	1065		
aag act att aac ggt tct gga cag aat caa caa acg cta aaa ttc				3307
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe				
1070	1075	1080		
agt gtg gcc gga ccc agc aac atg gct gtc cag gga aga aac tac				3352
Ser Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr				
1085	1090	1095		
ata cct gga ccc agc tac cga caa caa cgt gtc tca acc act gtg				3397
Ile Pro Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val				
1100	1105	1110		
act caa aac aac aac agc gaa ttt gct tgg cct gga gct tct tct				3442
Thr Gln Asn Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser				
1115	1120	1125		
tgg gct ctc aat gga cgt aat agc ttg atg aat cct gga cct gct				3487
Trp Ala Leu Asn Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala				

1130	1135	1140	
atg gcc agc cac aaa gaa gga gag gac cgt ttc ttt cct ttg tct			3532
Met Ala Ser His Lys Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser			
1145	1150	1155	
gga tct tta att ttt ggc aaa caa gga act gga aga gac aac gtg			3577
Gly Ser Leu Ile Phe Gly Lys Gln Gly Thr Gly Arg Asp Asn Val			
1160	1165	1170	
gat gcg gac aaa gtc atg ata acc aac gaa gaa gaa att aaa act			3622
Asp Ala Asp Lys Val Met Ile Thr Asn Glu Glu Glu Ile Lys Thr			
1175	1180	1185	
act aac ccg gta gca acg gag tcc tat gga caa gtg gcc aca aac			3667
Thr Asn Pro Val Ala Thr Glu Ser Tyr Gly Gln Val Ala Thr Asn			
1190	1195	1200	
cac cag agt gcc caa gga gga gtg ctg acc aac atc gct aga gga			3712
His Gln Ser Ala Gln Gly Gly Val Leu Thr Asn Ile Ala Arg Gly			
1205	1210	1215	
gag tac atg aga gga gga gca cag gcg cag acc ggc tgg gtt caa			3757
Glu Tyr Met Arg Gly Gly Ala Gln Ala Gln Thr Gly Trp Val Gln			
1220	1225	1230	
aac caa gga ata ctt ccg ggt atg gtt tgg cag gac aga gat gtg			3802
Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val			
1235	1240	1245	
tac ctg caa gga ccc att tgg gcc aaa att cct cac acg gac ggc			3847
Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly			
1250	1255	1260	
aac ttt cac cct tct ccg ctg atg gga ggg ttt gga atg aag cac			3892
Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His			
1265	1270	1275	
ccg cct cct cag atc ctc atc aaa aac aca cct gta cct gcg gat			3937
Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp			
1280	1285	1290	
cct cca acg gcc ttc aac aag gac aag ctg aac tct ttc atc acc			3982
Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr			
1295	1300	1305	
cag tat tct act ggc caa gtc agc gtg gag atc gag tgg gag ctg			4027
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu			
1310	1315	1320	
cag aag gaa aac agc aag cgc tgg aac ccg gag atc cag tac act			4072
Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr			

1325	1330	1335	
tcc aac tat tac aag tct aat aat gtt gaa ttt gct gtt aat act			4117
Ser Asn Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr			
1340	1345	1350	
gaa ggt gta tat agt gaa ccc cgc ccc att ggc acc aga tac ctg			4162
Glu Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu			
1355	1360	1365	
act cgt aat ctg taa ttgcttgta atcaataaac cgtttaattc gtttcagttg			4217
Thr Arg Asn Leu			
1370			
aactttggtc tctgcaagg gcgaattcgt ttaaactgc aggactagag gtctctgtatt			4277
agaggtcacg tgagtgtttt gcgacatttt gcgacacat gtggtcacgc tgggtattta			4337
agccccgagtg agcacgcagg gtctccattt tgaagcggga ggtttgaacg cgcagccgcc			4397
aagccgaatt ctgcagatat ccatcacact ggcgccgct cgactagagc ggccgccacc			4457
gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct tggcgtaatc			4517
atggatcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg			4577
agccggaagc ataaagtga aagcctgggg tgcctaataa gtgagctaac tcacattaat			4637
tgcgttgccg tccactgccc ctttcagtc gggaaactg tccgtgccagc tgcattaatg			4697
aatcgcccaa cgcgcgggga gaggcggttt gcgtattggg cgc			4740
<210>	20		
<211>	621		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成构建体		
<400>	20		
Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp			
1	5	10	15
Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu			
	20	25	30
Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile			
	35	40	45
Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu			
	50	55	60
Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val			
65	70	75	80
Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu			
	85	90	95
Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile			

100	105	110
Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu		
115	120	125
Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly		
130	135	140
Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys		
145	150	155
Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu		
165	170	175
Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His		
180	185	190
Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn		
195	200	205
Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr		
210	215	220
Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys		
225	230	235
Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala		
245	250	255
Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys		
260	265	270
Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln		
275	280	285
Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu		
290	295	300
Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala		
305	310	315
Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala		
325	330	335
Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro		
340	345	350
Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp		
355	360	365
Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala		
370	375	380
Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg		
385	390	395
Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val		
405	410	415

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser																			
Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe																			
Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln																			
Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val																			
Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala																			
Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val																			
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp																			
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu																			
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys																			
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu																			
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr																			
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp																			
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln																			
<210> 21																			
<211> 752																			
<212> PRT																			
<213> 人工序列																			
<220>																			
<223> 合成构建体																			
<400> 21																			
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser																			
1																			
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro																			
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro																			

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255
 Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
 260 265 270
 Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
 290 295 300
 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
 305 310 315 320
 Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro

355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp		
370	375	380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe		
385	390	395
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu		
405	410	415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu		
420	425	430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser		
435	440	445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser		
450	455	460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro		
465	470	475
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn		
485	490	495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn		
500	505	510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys		
515	520	525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly		
530	535	540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile		
545	550	555
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser		
565	570	575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Gly Gly Val Leu		
580	585	590
Thr Asn Ile Ala Arg Gly Glu Tyr Met Arg Gly Gly Ala Gln Ala Gln		
595	600	605
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln		
610	615	620
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His		
625	630	635
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met		
645	650	655
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala		
660	665	670

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr			
675	680	685	
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln			
690	695	700	
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn			
705	710	715	720
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val			
	725	730	735
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	750

<210> 22

<211> 1863

<212> DNA

<213> 人工序列

<220>

<223> AAV2 Rep核酸序列

<400> 22

atgccggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc	60
ggcattttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat	120
tctgacatgg atctgaatct gattgagcag gcaccctga ccgtggccga gaagctgcag	180
cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggctct tttctttgtg	240
caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaac caccggggtg	300
aaatccatgg ttttgggacg tttcctgagt cagattecgg aaaaactgat tcagagaatt	360
taccgcggga tcgagccgac tttgccaac tggttcgcgg tcacaaagac cagaaatggc	420
gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctccccaaa	480
accagcctg agctccagtg ggcgtggact aatatggaac agtatttaag cgcctgtttg	540
aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag	600
gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggatgatcag atcaaaaact	660
tcagccaggt acatggagct ggtcgggtgg ctctgggaca aggggattac ctcgagaag	720
cagtggatcc aggaggacca ggctcatac atctcctta atcggcctc caactcgcgg	780
tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc	840
cccgactacc tggtagggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa	900
atthttggaac taaacgggta cgatcccaa tatgcggctt ccgtctttct gggatgggcc	960
acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggctgcaac taccgggaag	1020
accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc	1080
aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg	1140
aagatgaccg ccaaggtcgt ggagtgcgcc aaagccattc tcggaggaag caaggtgcgc	1200
gtggaccaga aatgcaagtc ctcgcccag atagaccga ctcccgtgat cgtcacctcc	1260
aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca ccagcagccg	1320

ttgcaagacc ggatgttcaa atttgaactc acccgccgctc tggatcatga ctttggaag 1380
 gtcaccaagc aggaagtcaa agactttttc cggtagggcaa aggatcacgt ggttgaggtg 1440
 gagcatgaat tctacgtcaa aaagggtgga gccaaagaaa gaccgcccc cagtgcgca 1500
 gatataagtg agcccaaacg ggtgcgagag tcagttgagc agccatcgac gtcagacgag 1560
 gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg 1620
 aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc 1680
 ttcactcacg gacagaaaaga ctgttttagag tgctttcccg tgcagaatc tcaaccggtt 1740
 tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg 1800
 ccagacgctt gcaactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa 1860
 caa 1863

<210> 23

<211> 621

<212> PRT

<213> 人工序列

<220>

<223> AAV2 Rep氨基酸序列

<400> 23

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

Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190
 Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
 195 200 205
 Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220
 Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 225 230 235 240
 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255
 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270
 Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
 275 280 285
 Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 290 295 300
 Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
 305 310 315 320
 Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 325 330 335
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
 340 345 350
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 355 360 365
 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 370 375 380
 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
 385 390 395 400
 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
 405 410 415
 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser
 420 425 430
 Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 435 440 445
 Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
 450 455 460
 Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
 465 470 475 480
 Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala

	485	490	495
Pro Ser Asp Ala Asp Ile Ser Glu	Pro Lys Arg Val Arg Glu Ser Val		
500	505	510	
Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp			
515	520	525	
Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu			
530	535	540	
Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys			
545	550	555	560
Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu			
565	570	575	
Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr			
580	585	590	
Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp			
595	600	605	
Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln			
610	615	620	

- <210> 24
- <211> 2244
- <212> DNA
- <213> 人工序列
- <220>
- <223> AAV9 Cap n.588.EFS核酸序列
- <220>
- <221> misc_feature
- <222> (1765)..(1797)
- <223> EFS
- <400> 24

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc	60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac	120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gaccggcaa cggactcgac	180
aaggggggagc cggccaacgc agcagacgcg gcggcctcg agcacgaaa ggcctacgac	240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc	300
caggagcggc tcaaagaaga tacgtttttt gggggcaacc tcgggcgagc agtcttccag	360
gccaaaaaga ggcttcttga acctcttggc ctggttgagg aagcggctaa gacggctcct	420
ggaaagaaga ggctgtaga gcagtctcct caggaaccgg actcctccgc gggattggc	480
aaatcgggtg cacagcccgc taaaagaga ctcaatttcg gtcagactgg cgacacagag	540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct	600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga	660

gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac	720
accaccagca cccgaacctg ggccctgcc acctacaaca atcacctcta caagcaaate	780
tccaacagca catctggagg atcttcaa atgacaacgcct acttcggcta cagcaccccc	840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga	900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt	960
cagggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa cttaccagc	1020
acgggtccagg tcttcacgga ctccagactat cagctcccgt acgtgctcgg gtcggctcac	1080
gagggtctgcc tcccgcggtt cccagcggac gtttcatga ttctcagta cgggtatctg	1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc	1200
ccgtcgcaaa tgctaagaac gggttaacaac ttccagttca gctacgagtt tgagaacgta	1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcacc	1320
gaccaatact tgtaactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg	1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct	1440
ggaccagct accgacaaca acgtgtctca accactgtga ctcaaaaca caacagcga	1500
tttgcctggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatctt	1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct	1620
ttaatttttg gcaaacaaag aactggaaga gacaacgtgg atgaggaca agtcatgata	1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg	1740
gccacaaacc accagagtgc ccaagagttc agcagcaaca ccgtgaagct gaccagcga	1800
caggcgcaga ccggctgggt tcaaaaccaa ggaatacttc cgggtatggt ttggcaggac	1860
agagatgtgt acctgcaagg acccatttgg gccaaaatte ctccacagga cggcaacttt	1920
cacccttctc cgctgatggg agggtttga atgaagcacc cgctcctca gatcctcacc	1980
aaaaacacac ctgtacctgc ggatcctcca acggccttca acaaggaca gctgaactct	2040
ttcatcacc agtattctac tggccaagtc agcgtggaga tcgagtggga gctgcagaag	2100
gaaaacagca agcgtgga cccggagatc cagtacactt ccaactatta caagtcta	2160
aatgttgaat ttgctgttaa tactgaaggt gtatatagtg aacccgccc cattggcacc	2220
agatacctga ctcgtaatct gtaa	2244

<210> 25

<211> 747

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n.588.EFS氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (599)

<223> EFS

<400> 25

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

1	5	10	15
Glu Gly Ile Arg	Glu Trp Trp Ala	Leu Lys Pro Gly	Ala Pro Gln Pro
	20	25	30
Lys Ala Asn Gln	Gln His Gln Asp	Asn Ala Arg Gly	Leu Val Leu Pro
	35	40	45
Gly Tyr Lys Tyr	Leu Gly Pro Gly	Asn Gly Leu Asp	Lys Gly Glu Pro
	50	55	60
Val Asn Ala Ala	Asp Ala Ala Ala	Leu Glu His Asp	Lys Ala Tyr Asp
65	70	75	80
Gln Gln Leu Lys	Ala Gly Asp Asn	Pro Tyr Leu Lys	Tyr Asn His Ala
	85	90	95
Asp Ala Glu Phe	Gln Glu Arg Leu	Lys Glu Asp Thr	Ser Phe Gly Gly
	100	105	110
Asn Leu Gly Arg	Ala Val Phe Gln	Ala Lys Lys Arg	Leu Leu Glu Pro
	115	120	125
Leu Gly Leu Val	Glu Glu Ala Ala	Lys Thr Ala Pro	Gly Lys Lys Arg
	130	135	140
Pro Val Glu Gln	Ser Pro Gln Glu	Pro Asp Ser Ser	Ala Gly Ile Gly
145	150	155	160
Lys Ser Gly Ala	Gln Pro Ala Lys	Lys Arg Leu Asn	Phe Gly Gln Thr
	165	170	175
Gly Asp Thr Glu	Ser Val Pro Asp	Pro Gln Pro Ile	Gly Glu Pro Pro
	180	185	190
Ala Ala Pro Ser	Gly Val Gly Ser	Leu Thr Met Ala	Ser Gly Gly Gly
	195	200	205
Ala Pro Val Ala	Asp Asn Asn Glu	Gly Ala Asp Gly	Val Gly Ser Ser
	210	215	220
Ser Gly Asn Trp	His Cys Asp Ser	Gln Trp Leu Gly	Asp Arg Val Ile
225	230	235	240
Thr Thr Ser Thr	Arg Thr Trp Ala	Leu Pro Thr Tyr	Asn Asn His Leu
	245	250	255
Tyr Lys Gln Ile	Ser Asn Ser Thr	Ser Gly Gly Ser	Ser Asn Asp Asn
	260	265	270
Ala Tyr Phe Gly	Tyr Ser Thr Pro	Trp Gly Tyr Phe	Asp Phe Asn Arg
	275	280	285
Phe His Cys His	Phe Ser Pro Arg	Asp Trp Gln Arg	Leu Ile Asn Asn
	290	295	300
Asn Trp Gly Phe	Arg Pro Lys Arg	Leu Asn Phe Lys	Leu Phe Asn Ile
305	310	315	320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Glu Phe Ser Ser
 580 585 590
 Asn Thr Val Lys Leu Thr Ser Ala Gln Ala Gln Thr Gly Trp Val Gln
 595 600 605
 Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr
 610 615 620
 Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe

625	630	635	640
His Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro			
	645	650	655
Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala			
	660	665	670
Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly			
	675	680	685
Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys			
	690	695	700
Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn			
705	710	715	720
Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg			
	725	730	735
Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	

- <210> 26
- <211> 2247
- <212> DNA
- <213> 人工序列
- <220>
- <223> AAV9 Cap n588.IEI核酸序列
- <220>
- <221> misc_feature
- <222> (1765)..(1800)
- <223> IEI
- <400> 26

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc	60
gagtgggtggg ctttgaacc tggagcccct caaccaagg caaatcaaca acatcaagac	120
aacgctcgag gtcttgtgt tccgggttac aaataccttg gaccggcaa cggactcgac	180
aaggggggagc cggtaaacgc agcagacgcg gcggcctcg agcacgaca ggcctacgac	240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc	300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag	360
gccaaaaaga ggcttcttga acctcttggg ctggttgagg aagcggctaa gacggctcct	420
ggaaaagaaga ggctgtaga gcagtctcct caggaaccgg actcctcgc gggattggc	480
aaatcgggtg cacagcccgc taaaagaga ctcaatttcg gtcagactgg cgacacagag	540
tcagtcccag accctcaacc aatcggagaa cctcccgcag ccccctcagg tgtgggatct	600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga	660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac	720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc	780

tccaacagca catctggagg atcttcaa at gacaacgcct acttcggcta cagcaccccc 840
 tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga 900
 ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960
 caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa cttaccagc 1020
 acgggtccagg tcttcacgga ctcaactat cagctcccgt acgtgctcgg gtcggctcac 1080
 gagggctgcc tcccgcggtt cccagcggac gtttcatga ttctcagta cgggtatctg 1140
 acgcttaatg atggaagcca ggccgtgggt cgttcgtcct ttactgcct ggaatatttc 1200
 ccgtcgcaaa tgctaagaac ggtaacaac ttccagttca gctacgagtt tgagaacgta 1260
 cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
 gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
 ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
 ggaccagct accgacaaca acgtgtctca accactgtga ctcaaaaca caacagcga 1500
 tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
 ggacctgcta tgccagcca caaagaagga gaggaccgtt tctttcctt gtctggatct 1620
 ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgaggaca agtcatgata 1680
 accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtctta tggacaagtg 1740
 gccacaaacc accagagtgc ccaaactgag atcaacgcta ccagagctgg aaccaacctg 1800
 gcacaggcgc agaccggctg ggttcaaac caaggaatac ttccgggtat ggtttggcag 1860
 gacagagatg tgtacctgca aggaccatt tggccaaaa ttctcacac ggacggcaac 1920
 tttcaccctt ctccgctgat gggagggtt ggaatgaagc accgcctcc tcagatcctc 1980
 atcaaaaaca cacctgtacc tgcggatcct ccaacgcct tcaacaagga caagctgaac 2040
 tctttcatca cccagtattc tactggccaa gtcagcgtgg agatcgagtg ggagctgcag 2100
 aaggaaaaca gcaagcgtg gaaccggag atccagtaca cttccaacta ttacaagtct 2160
 aataatgttg aatttgctgt taatactgaa ggtgtatata gtgaaccccg cccattggc 2220
 accagatacc tgactcgtaa tctgtaa 2247

<210> 27

<211> 748

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n588.IEI氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (600)

<223> IEI

<400> 27

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro

20					25					30									
Lys	Ala	Asn	Gln	Gln	His	Gln	Asp	Asn	Ala	Arg	Gly	Leu	Val	Leu	Pro				
35					40					45									
Gly	Tyr	Lys	Tyr	Leu	Gly	Pro	Gly	Asn	Gly	Leu	Asp	Lys	Gly	Glu	Pro				
50					55					60									
Val	Asn	Ala	Ala	Asp	Ala	Ala	Ala	Leu	Glu	His	Asp	Lys	Ala	Tyr	Asp				
65					70					75					80				
Gln	Gln	Leu	Lys	Ala	Gly	Asp	Asn	Pro	Tyr	Leu	Lys	Tyr	Asn	His	Ala				
85					90					95									
Asp	Ala	Glu	Phe	Gln	Glu	Arg	Leu	Lys	Glu	Asp	Thr	Ser	Phe	Gly	Gly				
100					105					110									
Asn	Leu	Gly	Arg	Ala	Val	Phe	Gln	Ala	Lys	Lys	Arg	Leu	Leu	Glu	Pro				
115					120					125									
Leu	Gly	Leu	Val	Glu	Glu	Ala	Ala	Lys	Thr	Ala	Pro	Gly	Lys	Lys	Arg				
130					135					140									
Pro	Val	Glu	Gln	Ser	Pro	Gln	Glu	Pro	Asp	Ser	Ser	Ala	Gly	Ile	Gly				
145					150					155					160				
Lys	Ser	Gly	Ala	Gln	Pro	Ala	Lys	Lys	Arg	Leu	Asn	Phe	Gly	Gln	Thr				
165					170					175									
Gly	Asp	Thr	Glu	Ser	Val	Pro	Asp	Pro	Gln	Pro	Ile	Gly	Glu	Pro	Pro				
180					185					190									
Ala	Ala	Pro	Ser	Gly	Val	Gly	Ser	Leu	Thr	Met	Ala	Ser	Gly	Gly	Gly				
195					200					205									
Ala	Pro	Val	Ala	Asp	Asn	Asn	Glu	Gly	Ala	Asp	Gly	Val	Gly	Ser	Ser				
210					215					220									
Ser	Gly	Asn	Trp	His	Cys	Asp	Ser	Gln	Trp	Leu	Gly	Asp	Arg	Val	Ile				
225					230					235					240				
Thr	Thr	Ser	Thr	Arg	Thr	Trp	Ala	Leu	Pro	Thr	Tyr	Asn	Asn	His	Leu				
245					250					255									
Tyr	Lys	Gln	Ile	Ser	Asn	Ser	Thr	Ser	Gly	Gly	Ser	Ser	Asn	Asp	Asn				
260					265					270									
Ala	Tyr	Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg				
275					280					285									
Phe	His	Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn				
290					295					300									
Asn	Trp	Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile				
305					310					315					320				
Gln	Val	Lys	Glu	Val	Thr	Asp	Asn	Asn	Gly	Val	Lys	Thr	Ile	Ala	Asn				
325					330					335									

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

645	650	655
Pro Gln Ile Leu Ile Lys Asn Thr	Pro Val Pro Ala Asp Pro Pro Thr	
660	665	670
Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr		
675	680	685
Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser		
690	695	700
Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser		
705	710	715
Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro		
725	730	735
Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu		
740	745	

- <210> 28
- <211> 2259
- <212> DNA
- <213> 人工序列
- <220>
- <223> AAV9 Cap n.588.IEI-L核酸序列
- <220>
- <221> misc_feature
- <222> (1765)..(1812)
- <223> IEI-L
- <400> 28

```

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc      60
gagtgggtggg ctttgaaac tggagcccct caaccaagg caaatcaaca acatcaagac      120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gaccggcaa cggactcgac      180
aagggggagc cggtaacgc agcagacgcg gcggccctcg agcacgaaa ggcctacgac      240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc      300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag      360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct      420
ggaaagaaga ggctgtaga gcagtctct caggaaccgg actctctcgc gggattggc      480
aaatcgggtg cacagcccgc taaaagaga ctcaatttcg gtcagactgg cgacacagag      540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct      600
cttacaatgg cttcaggtgg tggcgacca gtggcagaca ataacgaagg tgccgatgga      660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac      720
accaccagca cccgaacctg ggcctgccc acctacaaca atcacctcta caagcaaadc      780
tccaacagca catctggagg atcttcaa atgacaagcct acttcggcta cagcaccccc      840
tgggggtatt ttgacttcaa cagattccac tgccattct caccacgtga ctggcagcga      900
    
```

ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt 960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ctttaccagc 1020
acgggtccagg tcttcacgga ctcaactat cagctcccgt acgtgctcgg gtcggctcac 1080
gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg 1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200
ccgtcgcaaa tgctaagaac ggtaacaac ttccagttca gctacgagtt tgagaacgta 1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
ggaccagct accgacaaca acgtgtctca accactgtga ctcaaaaca caacagcgaa 1500
tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggaca agtcatgata 1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtecta tggacaagtg 1740
gccacaaacc accagagtgc ccaaggagga atcgagatca acgctaccag agctggaacc 1800
aacctgggag gagcacaggc gcagaccggc tgggttcaaa accaaggaat acttccgggt 1860
atggtttggc aggacagaga tgtgtacctg caaggacca tttgggcca aattcctcac 1920
acggacggca actttcacc ttctccgctg atgggagggt ttggaatgaa gcaccgcct 1980
cctcagatcc tcatcaaaaa cacacctgta cctgcggatc ctccaacggc cttcaacaag 2040
gacaagctga actctttcat caccagtat tctactggcc aagtcagcgt ggagatcgag 2100
tgggagctgc agaaggaaaa cagcaagcgc tggaaccgg agatccagta cacttccaac 2160
tattacaagt ctaataatgt tgaatttgct gtaatactg aaggtgtata tagtgaacce 2220
cgccccattg gcaccagata cctgactcgt aatctgtaa 2259

<210> 29

<211> 752

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n.588.IEI-L氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (604)

<223> IEI-L

<400> 29

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15
Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro

35	40	45
Gly Tyr Lys Tyr Leu Gly Pro	Gly Asn Gly Leu Asp Lys Gly Glu Pro	
50	55	60
Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp		
65	70	75
80		
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala		
85	90	95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly		
100	105	110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro		
115	120	125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg		
130	135	140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly		
145	150	155
160		
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr		
165	170	175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro		
180	185	190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly		
195	200	205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser		
210	215	220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile		
225	230	235
240		
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu		
245	250	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn		
260	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg		
275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn		
290	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile		
305	310	315
320		
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn		
325	330	335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu		
340	345	350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Gly Gly Ile Glu
 580 585 590
 Ile Asn Ala Thr Arg Ala Gly Thr Asn Leu Gly Gly Ala Gln Ala Gln
 595 600 605
 Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
 610 615 620
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 625 630 635 640
 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
 645 650 655
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala

660	665	670
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr		
675	680	685
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln		
690	695	700
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn		
705	710	715
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val		
	725	730
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu		
740	745	750

- <210> 30
- <211> 2241
- <212> DNA
- <213> 人工序列
- <220>
- <223> AAV9 Cap n.588.SAN核酸序列
- <220>
- <221> misc_feature
- <222> (1765)..(1794)
- <223> SAN
- <400> 30

```

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc      60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac      120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac      180
aaggggggagc cgggtcaacgc agcagacgcg gcggcctcg agcacgacaa ggcctacgac      240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc      300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtctttccag      360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct      420
ggaaaagaaga ggctgtaga gcagtctct caggaaccgg actcctccgc ggttattggc      480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag      540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct      600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga      660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac      720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaate      780
tccaacagca catctggagg atcttcaaat gacaacgect actteggcta cagcaccccc      840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga      900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt      960
cagggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc     1020
    
```

acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac 1080
 gagggctgcc tcccgccgtt cccagcggac gttttcatga ttctcagta cgggtatctg 1140
 acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc 1200
 ccgtcgcaaa tgctaagaac ggtaacaac ttccagttca gctacgagtt tgagaacgta 1260
 cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
 gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
 ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
 ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaaca caacagcgaa 1500
 tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
 ggacctgcta tgccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
 ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgctggaca agtcatgata 1680
 accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtecta tggacaagtg 1740
 gccacaaacc accagagtgc ccaaagecgt aacttcatca agcctaccag ctacgcacag 1800
 gcgcagaccg gctgggttca aaaccaagga atacttccgg gtatggtttg gcaggacaga 1860
 gatgtgtacc tgcaaggacc catttgggcc aaaattctc acacggacgg caactttcac 1920
 ccttctccgc tgatgggagg gtttggatg aagcaccgc ctctcagat cctcatcaaa 1980
 aacacacctg tacctgcgga tctccaacg gccttcaaca aggacaagct gaactctttc 2040
 atcaccagat attctactgg ccaagtcagc gtggagatcg agtgggagct gcagaaggaa 2100
 aacagcaagc gctggaacc ggagatccag tacacttcca actattaca gtctaataat 2160
 gttgaatttg ctgttaatac tgaagtgta tatagtgaac cccgccccat tggcaccaga 2220
 tacctgactc gtaatctgta a 2241

<210> 31

<211> 746

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n.588.SAN氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (598)

<223> SAN

<400> 31

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

50	55	60
Val Asn Ala Ala Asp	Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp	
65	70	75
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala		80
	85	90
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly		95
	100	105
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro		110
	115	120
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg		125
	130	135
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly		140
145	150	155
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr		160
	165	170
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro		175
	180	185
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly		190
	195	200
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser		205
210	215	220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile		230
225	235	240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu		250
	245	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn		260
	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg		275
	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn		290
	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile		305
	310	315
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn		320
	325	330
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu		335
	340	345
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro		350
	355	360
		365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ser Ala Asn Phe
 580 585 590
 Ile Lys Pro Thr Ser Tyr Ala Gln Ala Gln Thr Gly Trp Val Gln Asn
 595 600 605
 Gln Gly Ile Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu
 610 615 620
 Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp Gly Asn Phe His
 625 630 635 640
 Pro Ser Pro Leu Met Gly Gly Phe Gly Met Lys His Pro Pro Pro Gln
 645 650 655
 Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asp Pro Pro Thr Ala Phe
 660 665 670
 Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln

	675		680		685										
Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln	Lys	Glu	Asn	Ser	Lys	Arg
	690		695		700										
Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn	Tyr	Tyr	Lys	Ser	Asn	Asn
705			710		715										720
Val	Glu	Phe	Ala	Val	Asn	Thr	Glu	Gly	Val	Tyr	Ser	Glu	Pro	Arg	Pro
			725		730										735
Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn	Leu						
			740		745										

- <210> 32
- <211> 2244
- <212> DNA
- <213> 人工序列
- <220>
- <223> AAV9 Cap n.588.SSN核酸序列
- <220>
- <221> misc_feature
- <222> (1765) .. (1797)
- <223> SSN
- <400> 32

```

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc      60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac      120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gaccgcaaa cggactcgac      180
aagggggagc cgggtcaacgc agcagacgcg gcggccctcg agcacgaaa ggcctacgac      240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc      300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggagagc agtcttccag      360
gccaaaaaga ggcttcttga acctcttggg ctggttgagg aagcggctaa gacggctcct      420
ggaaagaaga ggctgtaga gcagtctcct caggaaccgg actcctccgc gggattggc      480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag      540
tcagtcccag accctcaacc aatcggagaa cctcccagc cccctcagg tgtgggatct      600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga      660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac      720
accaccagca cccgaacctg ggccctgcc acctacaaca atcactcta caagcaaadc      780
tccaacagca catctggagg atcttcaaat gacaacgct acttcggcta cagcaccccc      840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga      900
ctcatcaaca acaactgggg attccggcct aagcagctca acttcaagct cttaacatt      960
cagggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa cttaccagc     1020
acgggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac     1080
gagggtgcc tcccgcggtt cccagcggac gttttcatga ttctcagta cgggtatctg     1140
    
```

acgcttaatg atggaagcca ggccgtgggt cgttcgtcct ttactgcct ggaatatttc 1200
 ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta 1260
 cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
 gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
 ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
 ggacccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa 1500
 tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
 ggacctgcta tgccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
 ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcgacaa agtcatgata 1680
 accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtctta tggacaagtg 1740
 gccacaaacc accagagtgc ccaaagcagc aacaccgtga agctgaccag cggacacgca 1800
 caggcgcaga ccggttgggt tcaaaaccaa ggaatacttc cgggtatggt ttggcaggac 1860
 agagatgtgt acctgcaagg acccatttgg gccaaaattc ctacacgga cggcaacttt 1920
 cacccttctc cgctgatggg agggtttggg atgaagcacc cgctctctca gatcctcatc 1980
 aaaaacacac ctgtacctgc ggatcctcca acggccttca acaaggacaa gctgaactct 2040
 ttcatcacc agtattctac tggccaagtc agcgtggaga tcgagtggga gctgcagaag 2100
 gaaaacagca agcgtggaa cccggagatc cagtacactt ccaactatta caagtcta 2160
 aatgttgaat ttgctgtaa tactgaaggt gtatatagtg aaccccgccc cattggcacc 2220
 agatacctga ctcgtaatct gtaa 2244

<210> 33

<211> 747

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n.588.SSN氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (599)

<223> SSN

<400> 33

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

65	70	75	80
Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala			
	85	90	95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly			
	100	105	110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro			
	115	120	125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg			
	130	135	140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly			
145	150	155	160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr			
	165	170	175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro			
	180	185	190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly			
	195	200	205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser			
210	215	220	
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile			
225	230	235	240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu			
	245	250	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn			
	260	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg			
	275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn			
290	295	300	
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile			
305	310	315	320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn			
	325	330	335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu			
	340	345	350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro			
	355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp			
370	375	380	

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

690	695	700	
Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr Tyr Lys Ser Asn			
705	710	715	720
Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr Ser Glu Pro Arg			
	725	730	735
Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	
<210>	34		
<211>	2256		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	AAV9 Cap n.588.SSN-L核酸序列		
<220>			
<221>	misc_feature		
<222>	(1765) .. (1809)		
<223>	SSN-L		
<400>	34		
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc			60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac			120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gaccceggcaa cggactcgac			180
aaggggggagc cgggtcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac			240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc			300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag			360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct			420
ggaaagaaga ggcttcttga gcagtctcct caggaaccgg actcctccgc gggatttggc			480
aaatcgggtg cacagcccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag			540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct			600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga			660
gtgggtagtt cctcgggaaa ttggcattgc gattcccatt ggctggggga cagagtcatt			720
accaccagca cccgaacctg ggccctgcc acctacaaca atcacctcta caagcaaattc			780
tccaacagca catctggagg atcttcaaat gacaacgect acttcggcta cagcaccccc			840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga			900
ctcatcaaca acaactgggg attccggcct aagcagctca acttcaagct cttcaacatt			960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa cttaccagc			1020
acgggtccagg tcttcacgga ctcagaactat cagctcccgt acgtgctcgg gtcggctcac			1080
gagggtgcc tcccgcggtt cccagcggac gttttcatga ttctcagta cgggtatctg			1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgcct ggaatatttc			1200
ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta			1260

cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc 1320
 gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg 1380
 ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
 ggacccagct accgacaaca acgtgtctca accactgtga ctaaaacaa caacagcgaa 1500
 tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
 ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
 ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata 1680
 accaacgaag aagaaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740
 gccacaaacc accagagtgc ccaaggagga agcagcaaca ccgtgaagct gaccagcgga 1800
 cacggaggag cacagggcga gaccggctgg gttcaaaacc aaggaatact tccgggtatg 1860
 gtttggcagg acagagatgt gtacctgcaa ggaccattt gggccaaaat tcttcacacg 1920
 gacggcaact ttcacccttc tccgetgatg ggagggtttg gaatgaagca cccgcctcct 1980
 cagatcctca tcaaaaacac acctgtacct gcggatctc caacggcctt caacaaggac 2040
 aagetgaact ctttcatcac ccagtattct actggccaag tcagcgtgga gatcgagtgg 2100
 gagctgcaga aggaaaacag caagcgetgg aaccgggaga tccagtacac ttccaactat 2160
 tacaagtcta ataatgttga atttgctgtt aatactgaag gtgtatatag tgaacccgcg 2220
 cccattggca ccagatacct gactcgtaat ctgtaa 2256

<210> 35

<211> 751

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n.588.SSN-L氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (603)

<223> SSN-L

<400> 35

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

	85	90	95
Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly			
	100	105	110
Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro			
	115	120	125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg			
	130	135	140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly			
145	150	155	160
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr			
	165	170	175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro			
	180	185	190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly			
	195	200	205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser			
210	215	220	
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile			
225	230	235	240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu			
	245	250	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn			
	260	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg			
	275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn			
	290	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile			
305	310	315	320
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn			
	325	330	335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu			
	340	345	350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro			
	355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp			
	370	375	380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe			
385	390	395	400

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

705	710	715	720	
Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val Tyr				
	725	730	735	
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu				
	740	745	750	
<210> 36				
<211> 2259				
<212> DNA				
<213> 人工序列				
<220>				
<223> AAV9 Cap n588.VLT-L核酸序列				
<220>				
<221> misc_feature				
<222> (1765) .. (1812)				
<223> VLT-L				
<400> 36				
atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc				60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac				120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gaccgcaaa cggactcgac				180
aagggggagc cggtaaacgc agcagacgcg gcggccctcg agcacgaaa ggcctacgac				240
cagcagctca aggccggaga caaccgtac ctcaagtaca accacgccga cgccgagttc				300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggagagc agtcttccag				360
gccaaaaaga ggcttcttga acctcttggg ctggttgagg aagcggctaa gacggctcct				420
ggaaagaaga ggcttcttga gcagtctcct caggaaccgg actcctccgc gggatttggc				480
aaatcgggtg cacagcccgc taaaagaga ctcaatttcg gtcagactgg cgacacagag				540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct				600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga				660
gtgggtagtt cctcgggaaa ttggcattgc gattccaat ggctggggga cagagtcac				720
accaccagca cccgaacctg ggcctgccc acctacaaca atcacctcta caagcaaatc				780
tccaacagca catctggagg atcttcaaat gacaacgctt acttcggcta cagcaccccc				840
tgggggtatt ttgacttcaa cagattccac tgccattct caccagtgta ctggcagcga				900
ctcatcaaca acaactgggg attccggcct aagcagctca acttcaagct cttcaacatt				960
caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc				1020
acggtccagg tcttcacgga ctcagaactat cagctcccgt acgtgctcgg gtcggctcac				1080
gagggtgcc tcccgcggtt cccagcggac gttttcatga ttctcagta cgggtatctg				1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct tttactgctt ggaatatttc				1200
ccgtcgaaa tgctaagaac gggttaacaac ttccagttca gctacgagtt tgagaacgta				1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc				1320
gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg				1380

ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct 1440
 ggaccagct accgacaaca acgtgtctca accactgtga ctcaaaciaa caacagcgaa 1500
 tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct 1560
 ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct 1620
 ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgctggaciaa agtcatgata 1680
 accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg 1740
 gccacaaacc accagagtgc ccaaggagga gtgctgacca acatcgctag aggagagtac 1800
 atgagaggag gagcacaggc gcagaccggc tgggttcaaa accaaggaat acttccgggt 1860
 atggtttggc aggacagaga tgtgtacctg caaggacca tttgggcaa aattcctcac 1920
 acggacggca actttcacc tctccgctg atgggagggt ttggaatgaa gcaccgcct 1980
 cctcagatcc tcatcaaaaa cacacctgta cctgctgacc ctccaacggc cttcaacaag 2040
 gacaagctga actctttcat caccagtat tctactggcc aagtcagcgt ggagatcgag 2100
 tgggagctgc agaaggaaaa cagcaagcgc tggaaaccgg agatccagta cacttccaac 2160
 tattacaagt ctaataatgt tgaatttgc gttaatactg aaggtgtata tagtgaacce 2220
 cgccccattg gcaccagata cctgactcgt aatctgtaa 2259

<210> 37

<211> 752

<212> PRT

<213> 人工序列

<220>

<223> AAV9 Cap n588.VLT-L氨基酸序列

<220>

<221> MISC_FEATURE

<222> (499) .. (604)

<223> VLT-L

<400> 37

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
 565 570 575
 Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Gly Gly Val Leu
 580 585 590
 Thr Asn Ile Ala Arg Gly Glu Tyr Met Arg Gly Gly Ala Gln Ala Gln
 595 600 605
 Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
 610 615 620
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 625 630 635 640
 Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
 645 650 655
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
 660 665 670
 Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
 675 680 685
 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
 690 695 700
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
 705 710 715 720
 Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val

	725	730	735
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu			
	740	745	750
<210>	38		
<211>	11		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	EFS肽序列		
<400>	38		
Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser			
1	5	10	
<210>	39		
<211>	15		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	SSN-L肽序列		
<400>	39		
Gly Gly Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Gly Gly			
1	5	10	15
<210>	40		
<211>	11		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	SSN肽序列		
<400>	40		
Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His			
1	5	10	
<210>	41		
<211>	10		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	SAN肽序列		
<400>	41		
Ser Ala Asn Phe Ile Lys Pro Thr Ser Tyr			
1	5	10	

<210> 42
 <211> 12
 <212> PRT
 <213> 人工序列
 <220>
 <223> IEI肽序列
 <400> 42
 Ile Glu Ile Asn Ala Thr Arg Ala Gly Thr Asn Leu
 1 5 10
 <210> 43
 <211> 16
 <212> PRT
 <213> 人工序列
 <220>
 <223> IEI-L肽序列
 <400> 43
 Gly Gly Ile Glu Ile Asn Ala Thr Arg Ala Gly Thr Asn Leu Gly Gly
 1 5 10 15
 <210> 44
 <211> 736
 <212> PRT
 <213> 人工序列
 <220>
 <223> AAV9衣壳
 <400> 44
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro	115	120	125
Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg	130	135	140
Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly	145	150	155
Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr	165	170	175
Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro	180	185	190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly	195	200	205
Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser	210	215	220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile	225	230	235
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu	245	250	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn	260	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg	275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn	290	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile	305	310	315
Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn	325	330	335
Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu	340	345	350
Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro	355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp	370	375	380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe	385	390	395
Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu	405	410	415
Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu			

	420		425		430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser					
	435		440		445
Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser					
	450		455		460
Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro					
465		470		475	480
Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn					
	485		490		495
Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn					
	500		505		510
Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys					
	515		520		525
Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly					
	530		535		540
Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile					
545		550		555	560
Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser					
	565		570		575
Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln					
	580		585		590
Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln					
	595		600		605
Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His					
	610		615		620
Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met					
625		630		635	640
Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala					
	645		650		655
Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr					
	660		665		670
Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln					
	675		680		685
Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn					
	690		695		700
Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val					
705		710		715	720
Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu					
	725		730		735

<210> 45
 <211> 736
 <212> PRT
 <213> 人工序列
 <220>
 <223> AAVhu68衣壳
 <400> 45
 Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15
 Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30
 Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
 35 40 45
 Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60
 Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80
 Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95
 Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Val Gly Ile Gly
 145 150 155 160
 Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
 165 170 175
 Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190
 Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205
 Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
 210 215 220
 Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
 225 230 235 240
 Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
 245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
 260 265 270
 Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
 290 295 300
 Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
 305 310 315 320
 Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
 355 360 365
 Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
 370 375 380
 Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
 405 410 415
 Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
 435 440 445
 Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
 450 455 460
 Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
 465 470 475 480
 Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
 485 490 495
 Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
 500 505 510
 Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
 515 520 525
 Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
 530 535 540
 Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
 545 550 555 560
 Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser

<223> 任何氨基酸
 <220>
 <221> MISC_FEATURE
 <222> (3) .. (3)
 <223> Xaa选自苏氨酸(T)、异亮氨酸(I)、缬氨酸(V)或丙氨酸(A)
 <220>
 <221> MISC_FEATURE
 <222> (4) .. (4)
 <223> Xaa选自赖氨酸(K)或精氨酸(R)
 <400> 47
 Asn Xaa Xaa Xaa
 1
 <210> 48
 <211> 7
 <212> PRT
 <213> 人工序列
 <220>
 <223> AAV2变体肽NDVRAVS
 <400> 48
 Asn Asp Val Arg Ala Val Ser
 1 5
 <210> 49
 <211> 7
 <212> PRT
 <213> 人工序列
 <220>
 <223> PHP.B肽插入物
 <400> 49
 Thr Leu Ala Val Pro Phe Lys
 1 5
 <210> 50
 <211> 33
 <212> DNA
 <213> 人工序列
 <220>
 <223> 核酸序列EFS
 <400> 50
 gagttcagca gcaacaccgt gaagctgacc agc
 <210> 51

<211>	36	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	核酸序列IEI	
<400>	51	
	atcgagatca acgctaccag agctggaacc aacctg	36
<210>	52	
<211>	48	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	核酸序列IEI-L	
<400>	52	
	ggaggaatcg agatcaacgc taccagagct ggaaccaacc tgggagga	48
<210>	53	
<211>	30	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	核酸序列SAN	
<400>	53	
	agcgctaact tcatcaagcc taccagctac	30
<210>	54	
<211>	33	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	核酸序列SSN	
<400>	54	
	agcagcaaca ccgtgaagct gaccagcgga cac	33
<210>	55	
<211>	45	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	核酸序列SSN-L	
<400>	55	
	ggaggaagca gcaacaccgt gaagctgacc agcggacacg gagga	45

<210> 56

<211> 48

<212> DNA

<213> 人工序列

<220>

<223> 核酸序列VLT-L

<400> 56

ggaggagtgc tgaccaacat cgctagagga gagtacatga gaggagga

48

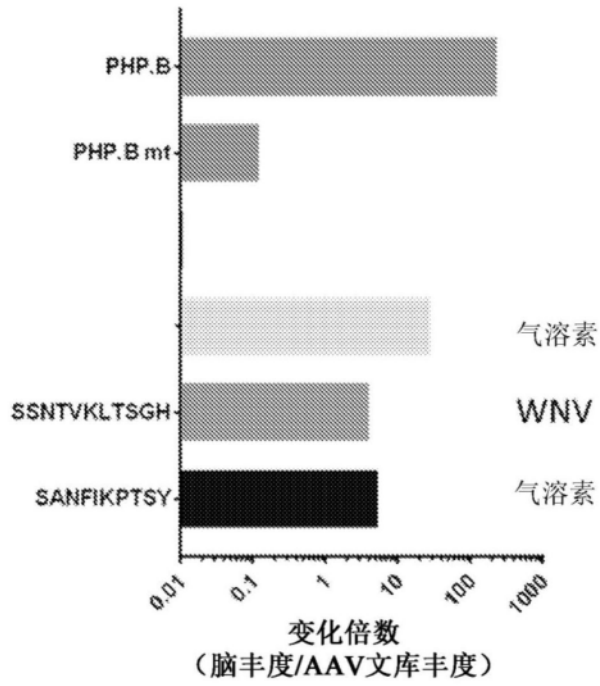


图1A

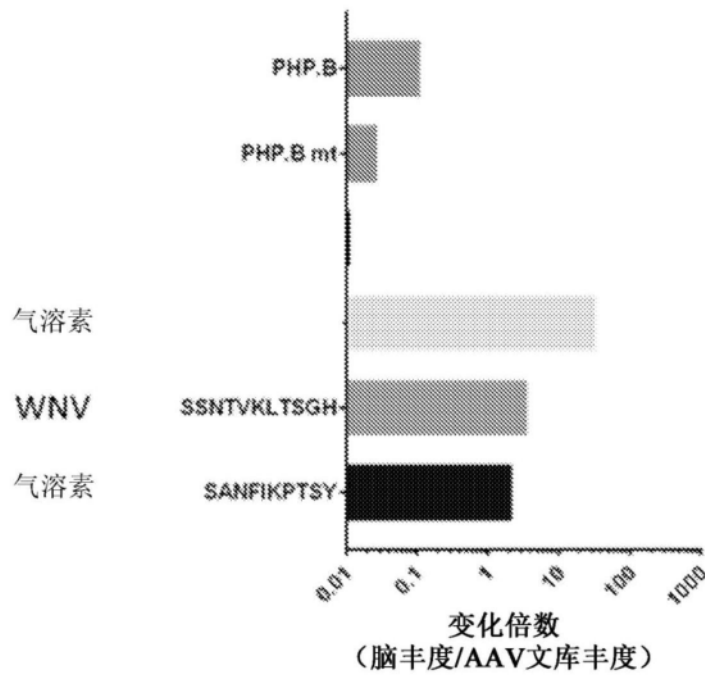


图1B

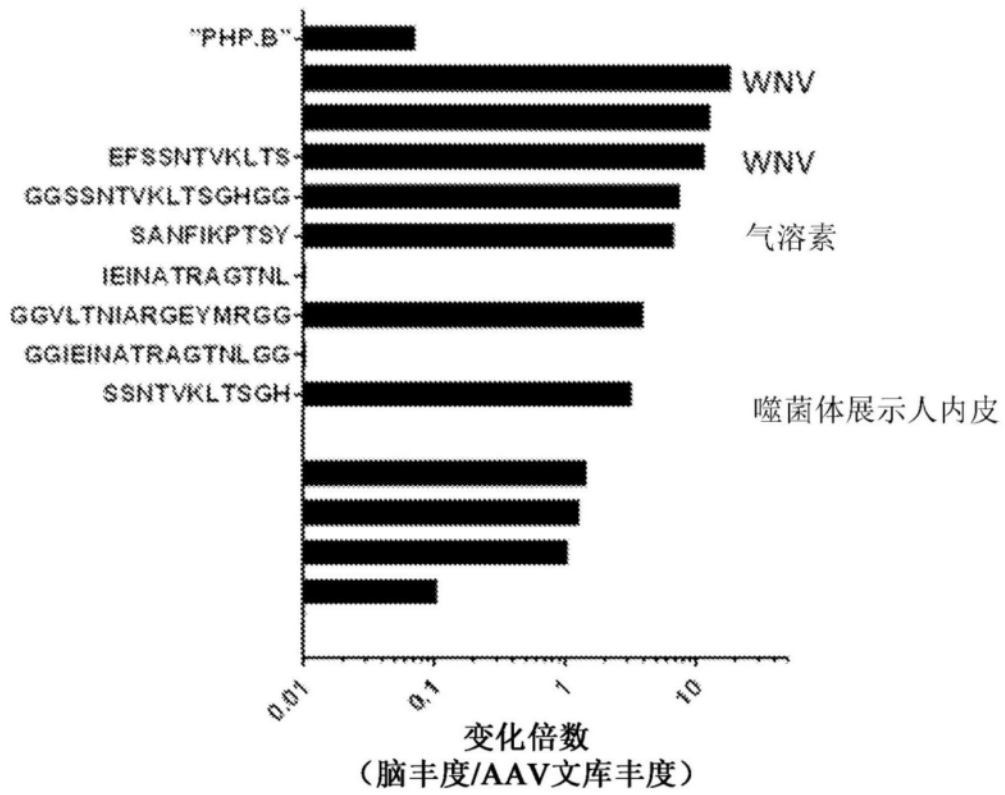


图2A

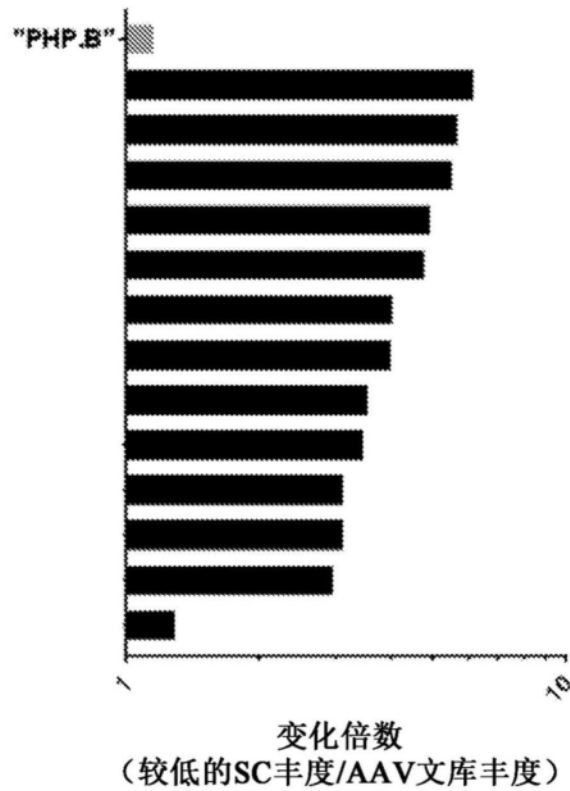


图2B

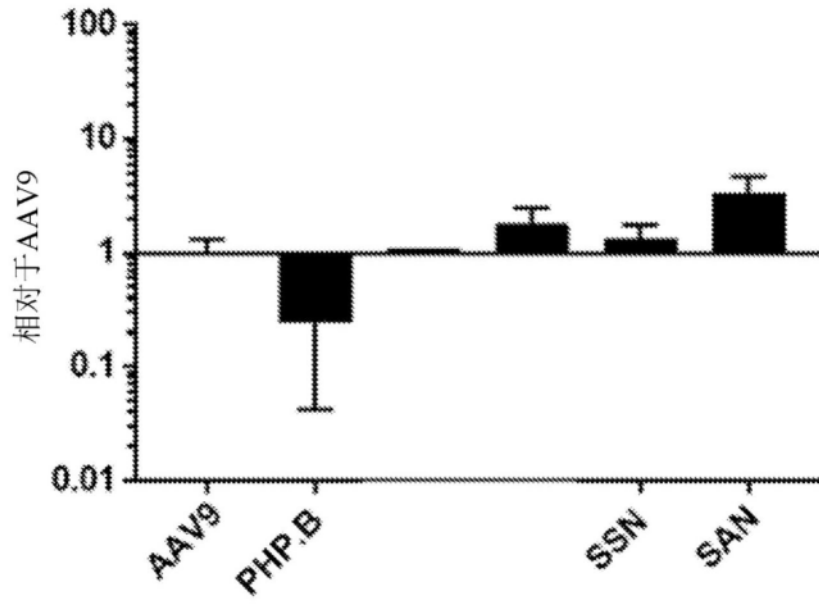


图3A

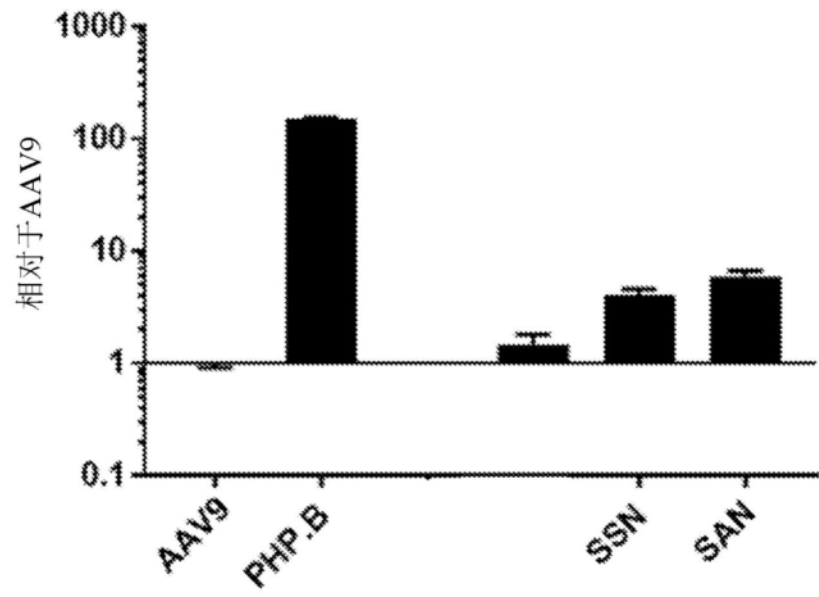


图3B

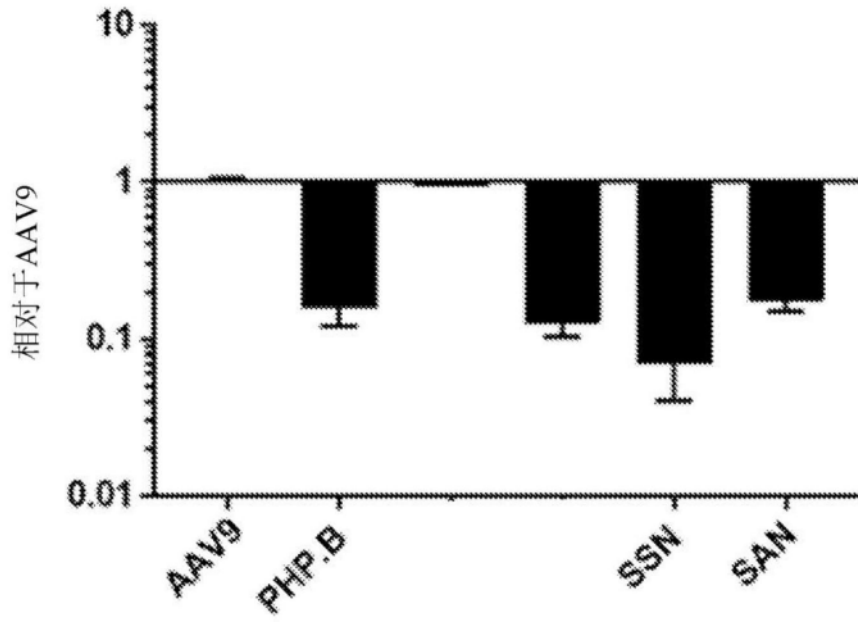


图3C

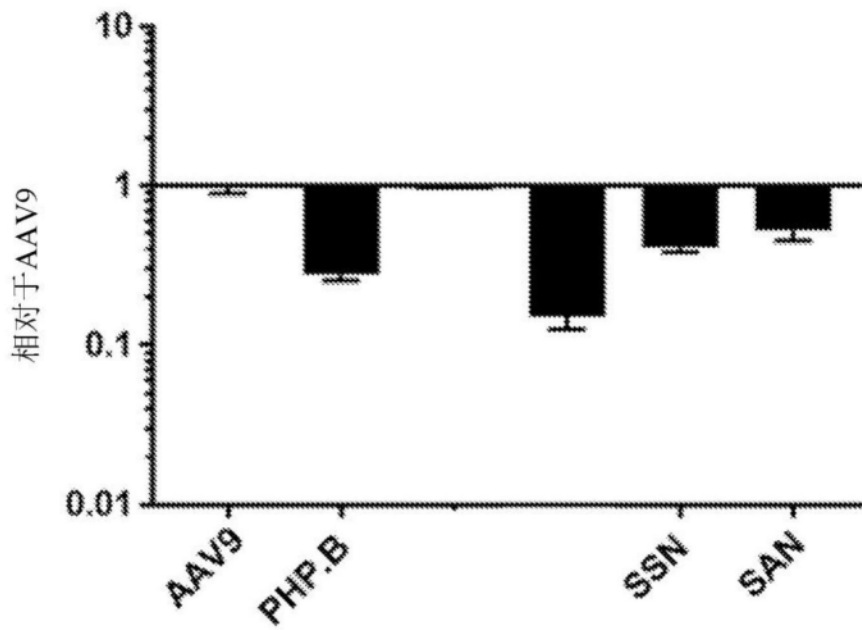


图3D

566 571 576 581 586 591 596 601 606 611
 IKTTNPVATESYGQVATNHQSAQAQTGWVQNGQILPGMWQDRDYYLQ
 566 571 576 581 586 591 596 601 606 611
 EEEIKTTNPVATEEYGIVADNLQQNTAPQIGTVNSQGALPGMWQNRDV
 66 571 576 581 586 591 596 601 606 611 61
 IRPTNPVATEEYGIVSSNLQAANTAAQTQVNNQGALPGMWQNRDYYLQ
 551 556 561 566 571 576 581 586 591 596
 ASNTALDNVMITDEEEIKATNPVATERFGTVAVNLQSSSTDPATGDVHVM
 556 561 566 571 576 581 586 591 596 601
 QPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIVPGSVMERDYYLQ
 6 561 566 571 576 581 586 591 596 601 60
 FTSEELAATNATDTDMWGNLPGGQSNLPTVDRLTALGAVPGMWQNR
 566 571 576 581 586 591 596 601 606 611
 EEIRTTNPVATEEQYGTVANLQSSNTAPTTRTVNDQGALPGMWQDRDYY
 566 571 576 581 586 591 596 601 606 611
 RTTNPVATEEQYGSVSTNLQRGNRQAATADVNTQGVLPGMWQDRDYYLQ
 566 571 576 581 586 591 596 601 606 611
 IKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMWQDRDYYLQ

AAV9

AAV8

AAV7

AAV6

AAV5

AAV4

AAV3B

AAV2

AAV1

图4

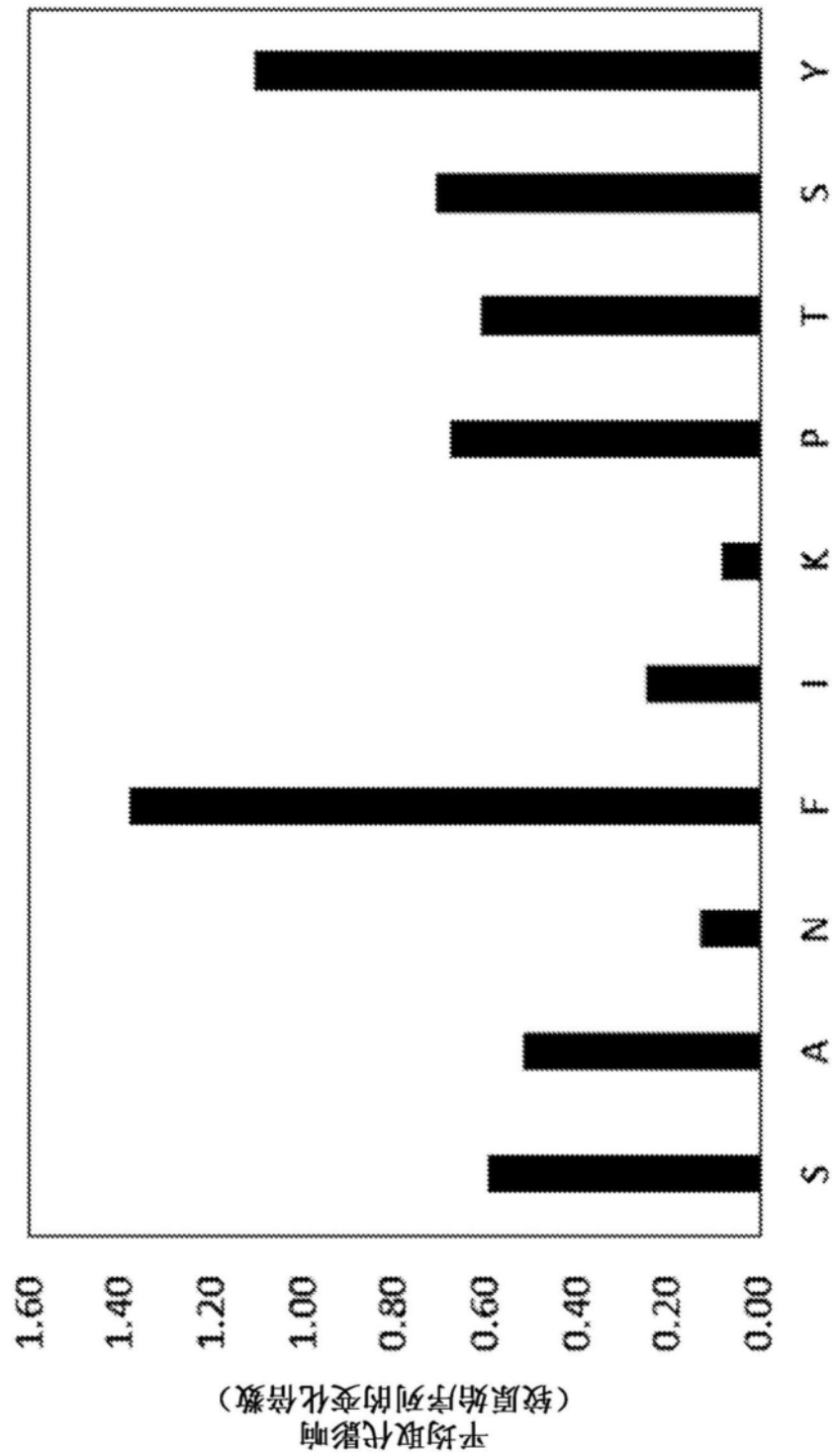


图5

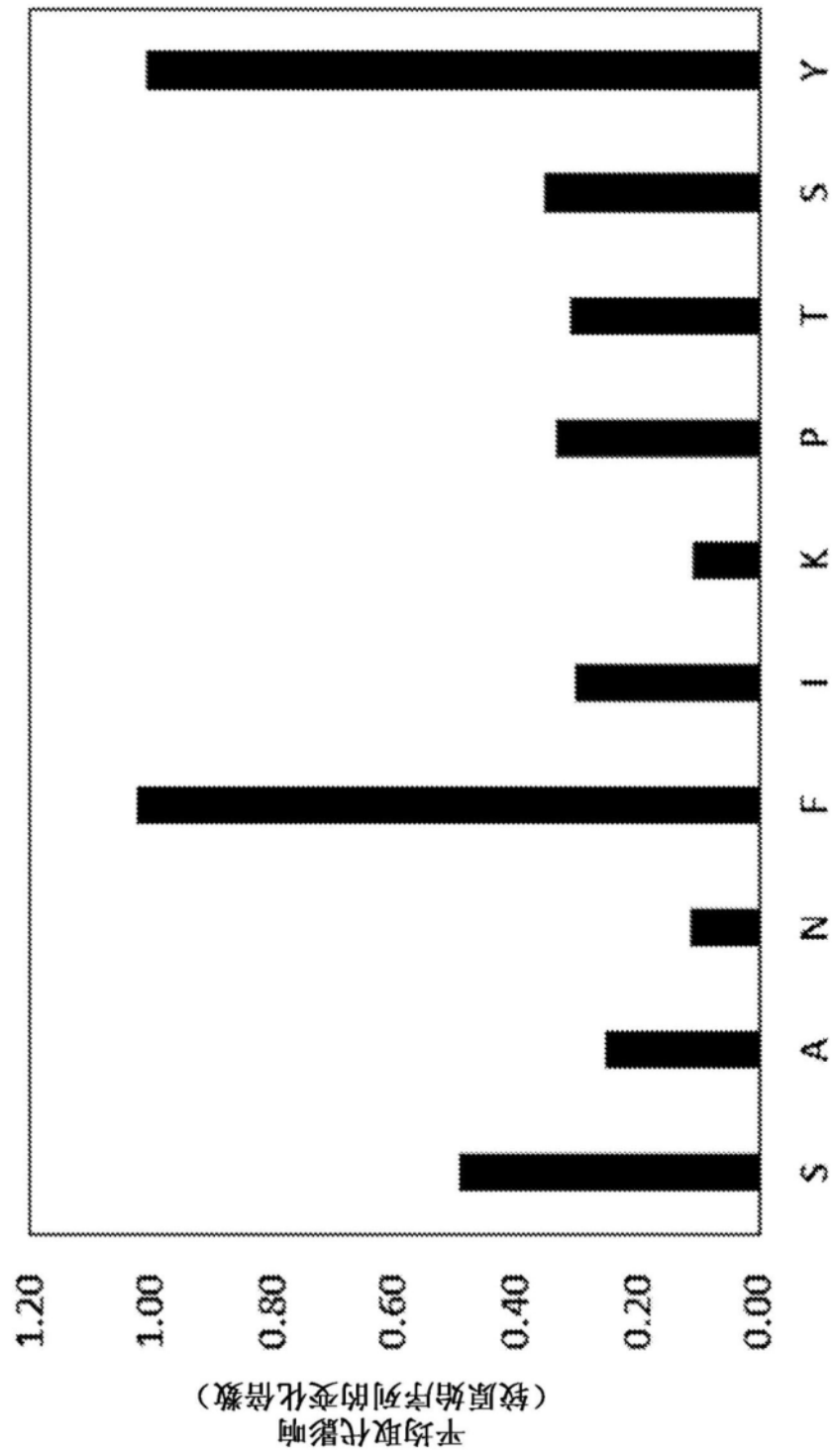


图6

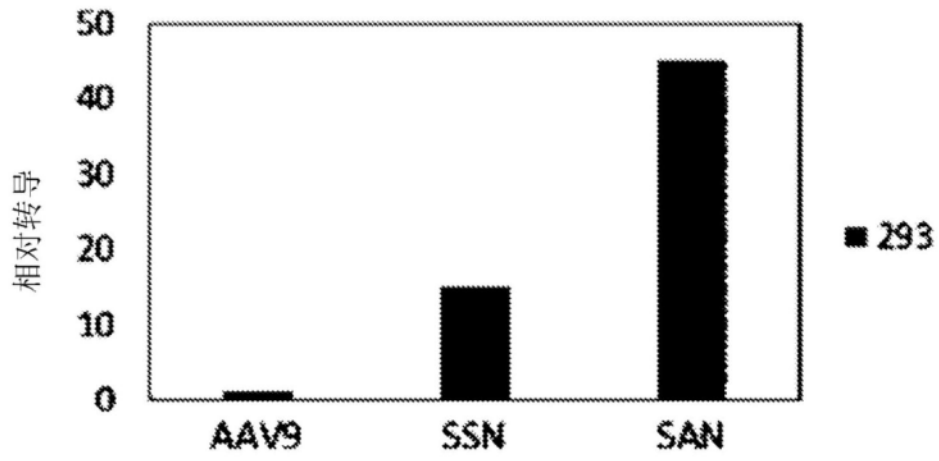


图7A

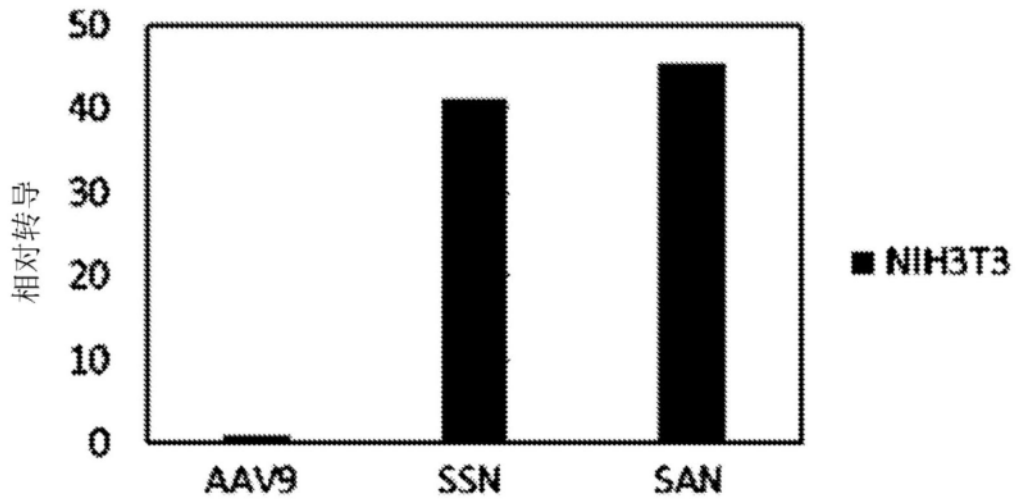


图7B

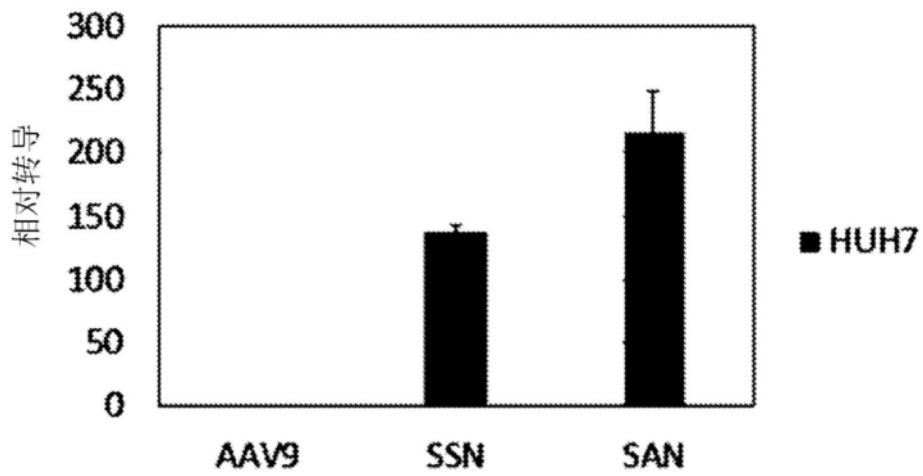


图7C

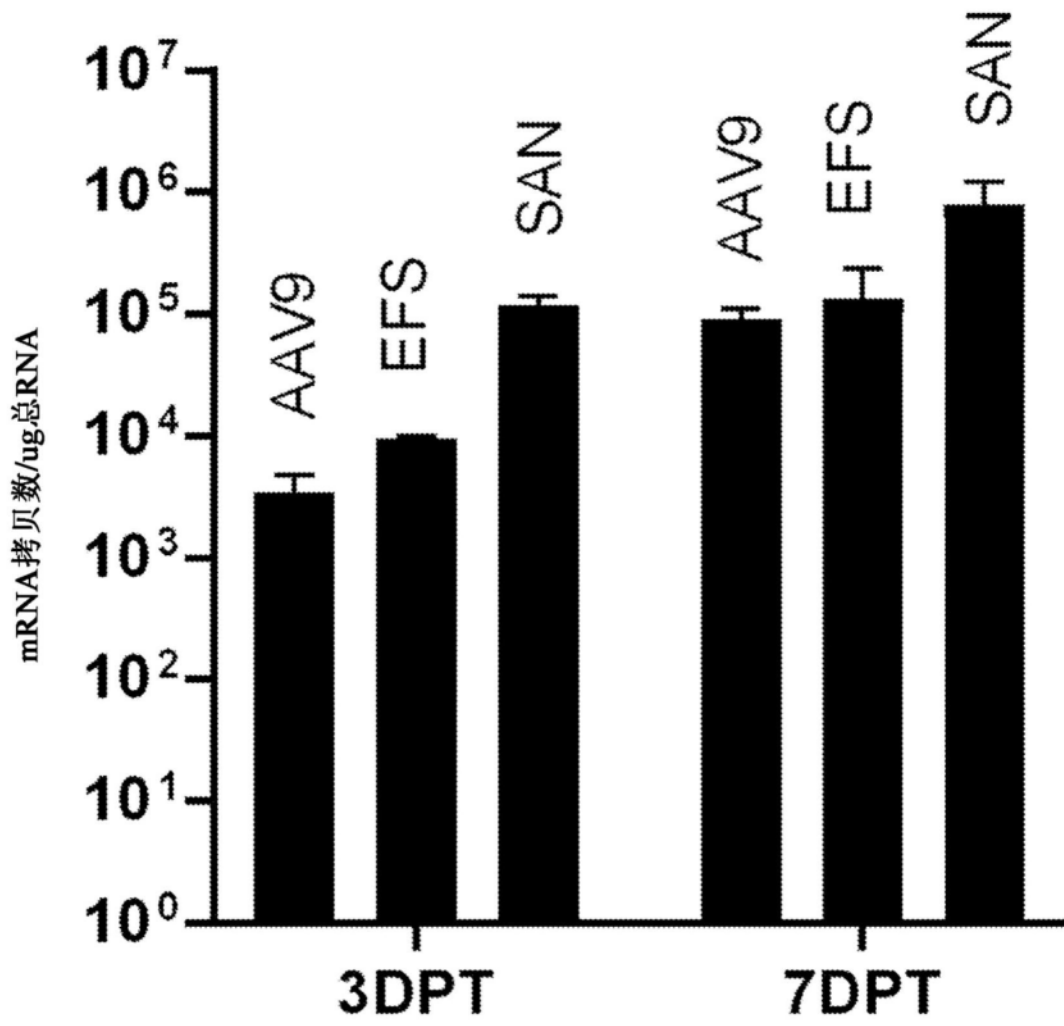


图7D

无载体

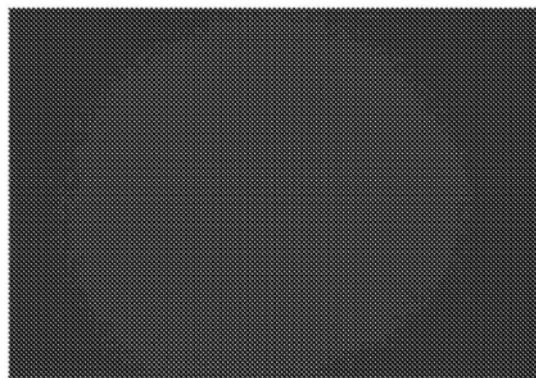


图7E

AAV9

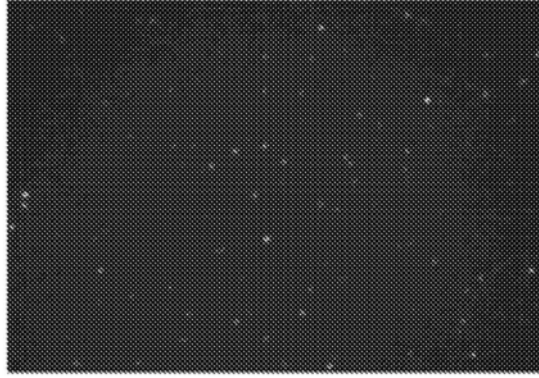
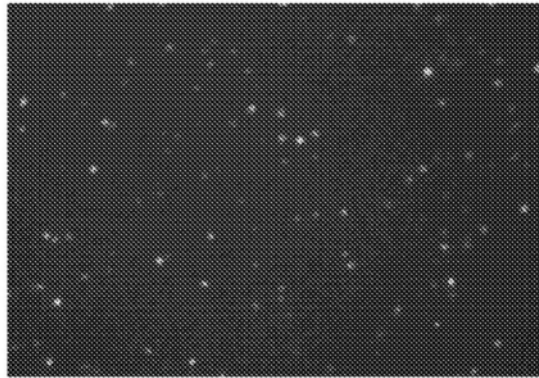
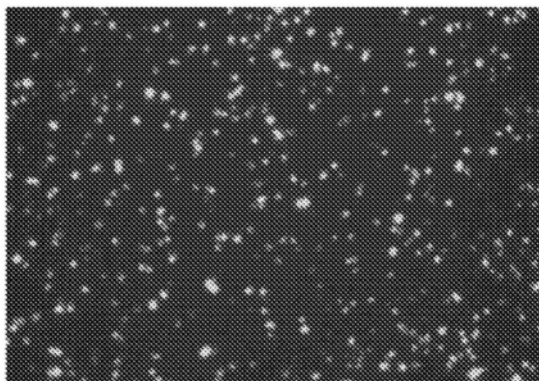


图7F



EFS

图7G



SAN

图7H

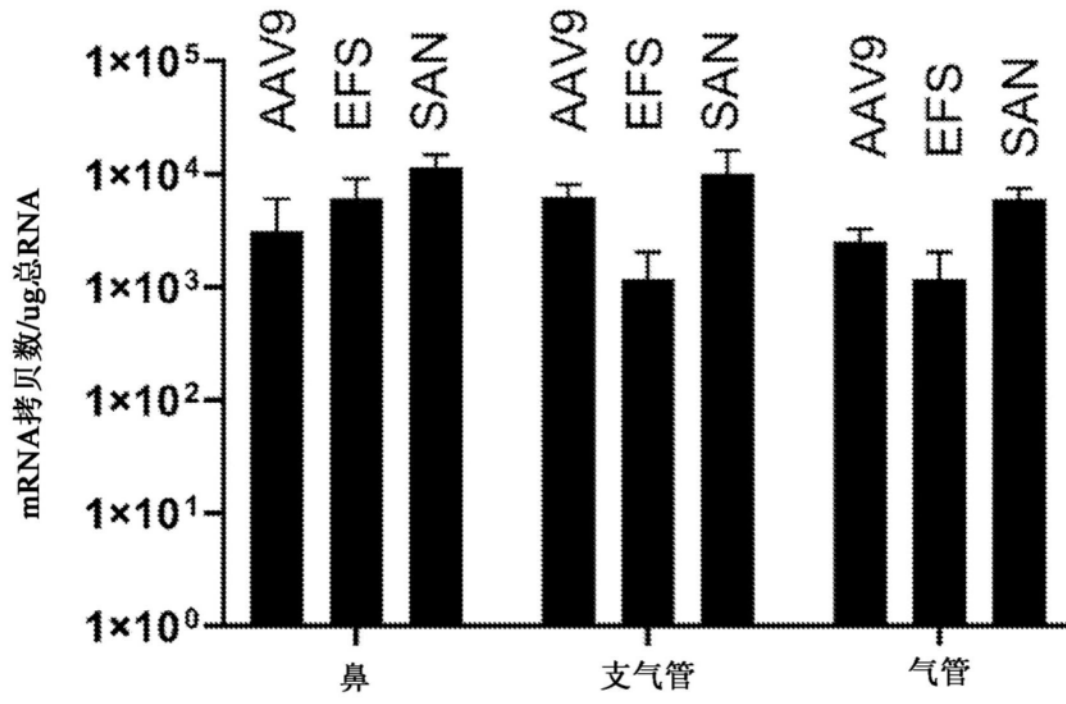


图8